AES-Based Security Coprocessor IC in 0.18-μm CMOS With Resistance to Differential Power Analysis Side-Channel Attacks

David D. Hwang, *Member, IEEE*, Kris Tiri, *Member, IEEE*, Alireza Hodjat, *Student Member, IEEE*, Bo-Cheng Lai, *Student Member, IEEE*, Shenglin Yang, *Student Member, IEEE*, Patrick Schaumont, *Member, IEEE*, and Ingrid Verbauwhede, *Senior Member, IEEE*

***Abstract*—Security ICs are vulnerable to side-channel attacks (SCAs) that find the secret key by monitoring the power consumption or other information that is leaked by the switching behavior of digital CMOS gates. This paper describes a side-channel attack resistant coprocessor IC fabricated in 0.18-μm CMOS consisting of an Advanced Encryption Standard (AES) based cryptographic engine, a fingerprint-matching engine, template storage, and an interface unit. Two functionally identical coprocessors have been fabricated on the same die. The first coprocessor was implemented using standard cells and regular routing techniques. The second coprocessor was implemented using a logic style called wave dynamic differential logic (WDDL) and a layout technique called differential routing to combat the differential power analysis (DPA) side-channel attack. Measurement-based experimental results show that a DPA attack on the insecure coprocessor requires only 8000 encryptions to disclose the entire 128-bit secret key. The same attack on the secure coprocessor does not disclose the entire secret key even after 1 500 000 encryptions.**

***Index Terms*—Advanced Encryption Standard (AES), biometrics, cryptography, differential power analysis, security, side-channel attacks.**

I. Introduction

IN RECENT YEARS, the integrated circuit (IC) has emerged as a weak link in embedded security applications. Due to its physical nature and characteristics, the IC broadcasts information that can be directly linked to the secret key being used in an encryption operation. Several attacks have been reported that use broadcasted information such as power consumption, time delay, and electromagnetic radiation to find the secret key. These side-channel attacks (SCAs) are noninvasive, require minimal equipment, and are a real threat for any device in which the security IC is easily observable, such as smart cards and other embedded devices [1]–[3].

As an example of the potency of SCAs, Lenstra and Verheul wrote that a 109-bit symmetric key should be able to guarantee the confidentiality of data encrypted with such a key until the year 2050 [4]. With the differential power analysis (DPA) side-channel attack, however, we were able to find the key of a standard cell IC AES implementation with a larger 128-bit key in less than three minutes with standard laboratory equipment. Clearly, SCAs pose serious concerns for the embedded IC security community.

Of all side-channel attacks, differential power analysis is a SCA of particular concern as it is very effective in finding a secret key. The attack is based on the fact that logic operations in standard static CMOS have power characteristics that depend on the input data. Dynamic power is only drawn from the power supply by a CMOS logic gate when a 0 to 1 output transition occurs. (During 0 to 0 and 1 to 1 transitions, no power is drawn. During a 1 to 0 transition, the stored capacitance is discharged to ground. There is also leakage power and short circuit power but currently for SCA analysis it is the data dependent dynamic power that matters.) Therefore, by measuring the power supply of an IC as it encrypts, and then performing statistical analysis of the measured power traces, the secret key can be determined. DPA has been proven effective in extracting the key of both microprocessor-based and ASIC-based encryption systems.

This paper discusses an embedded security coprocessor IC which implements two circuit-level techniques used to thwart differential power analysis. The first technique is called Wave Dynamic Differential Logic, and is used to create logic gates which dissipate a constant amount of power per cycle. The second technique is called differential routing and is used to ensure the interconnect capacitances of the true and false output nodes of the WDDL gates are equal. The coprocessor itself is used for embedded biometric authentication, and consists of an AES-based cryptographic engine, a fingerprint matching engine (which we call the oracle), a fingerprint storage element, and an interface module.

The remainder of the paper is outlined as follows. Section II describes the coprocessor IC system architecture and its com-

Manuscript received September 5, 2005; revised December 19, 2005. This work was supported in part by the National Science Foundation under CCR-0098361, UC-Micro under Grants 02-079, 03-088, and 04-095, Panasonic Foundation, Sun Microsystems, Atmel Corporation, and the Fannie and John Hertz Foundation (DH). The work was performed when the authors were with the Department of Electrical Engineering, University of California, Los Angeles, CA 90095 USA.

D. Hwang is with KeyEye Communications, Irvine, CA 92618 USA (e-mail: dhwang@ee.ucla.edu).

K. Tiri is with Intel, Hillsboro, OR 97124 (e-mail: tiri@ee.ucla.edu).

A. Hodjat is with Broadcom, Irvine, CA 92618 USA (e-mail: ahodjat@ee.ucla.edu).

B. Lai and S. Yang are with the Department of Electrical Engineering, University of California, Los Angeles, CA 90095 USA (e-mail: bclai@ee.ucla.edu; shengliny@ee.ucla.edu).

P. Schaumont is with the Department of Electrical and Computer Engineering, Virginia Tech, Blacksburg, VA 24061 USA (e-mail: schaum@vt.edu).

I. Verbauwhede is with the Department of Electrical Engineering, University of California, Los Angeles, CA 90095 USA. She is also with the Katholieke Universiteit Leuven, ESAT-COSIC, 3000 Leuven, Belgium (e-mail: ingrid@ee.ucla.edu).

Digital Object Identifier 10.1109/JSSC.2006.870913

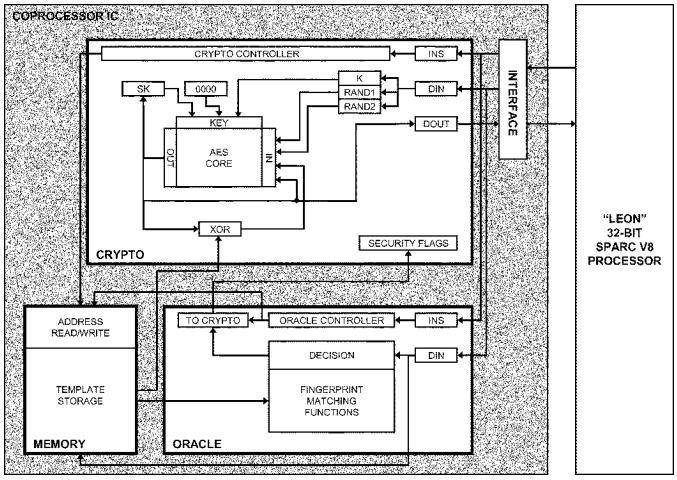


Fig. 1. ThumbPod system block diagram (fabricated IC is shaded).

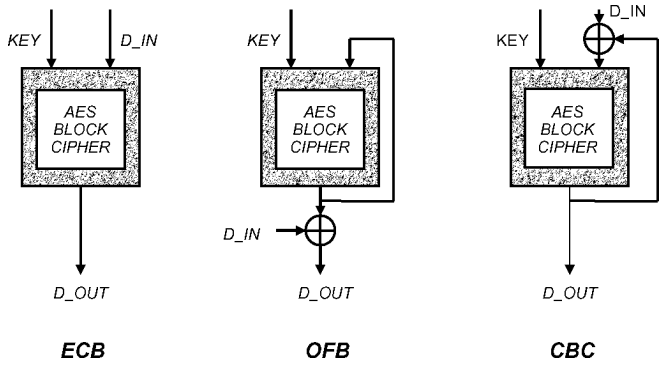


Fig. 2. Block cipher encryption modes of operation.

ponents. Section III describes the differential power analysis attack in further details and explains the circuit technique countermeasures of WDDL and differential routing. Subsequently, area, timing and power results are presented together with the power attack resistance. Finally, related state-of-the-art and a conclusion are presented.

II. Coprocessor IC Architecture

The described coprocessor IC is part of the ThumbPod embedded system, which is a portable biometric and cryptographic authentication device composed of a 32-bit SPARC processor coupled with a memory-mapped coprocessor IC, as shown in Fig. 1.

A. Overall System Architecture

The ThumbPod embedded system is used as a personal authenticator based on fingerprint biometrics and symmetric-key cryptography in the context of a client-server authentication system. For maximum security and privacy, all biometric components of the authentication system have been partitioned to the embedded device (versus on the server). The embedded device thus must have the capabilities to store the biometric data template, extract fingerprint minutiae from a candidate fingerprint, perform a matching operation of the candidate versus the template, implement various symmetric-key protocols (encryption and message authentication code generation), and communicate wirelessly via secure communication protocols to the server.

Since all sensitive data is localized on the embedded device, the device must be protected from both software and hardware attacks. However, as will be seen in further sections, providing this protection requires overhead in terms of power, area, and computational cost. Hence, a design technique called security partitioning has been applied to the device. Using security partitioning, the system is partitioned into two parts: a secure module (which stores secure data and processes secure information) and an insecure module (which stores insecure data and processes nonsensitive information). The partitioning is performed in order to isolate the sensitive data and functions of the device onto the secure module. The entire system does not need to be protected by the circuit techniques described in this paper. Only the secure module must be protected for the system to remain secure, thus minimizing such overhead.

In our system, the embedded system has been partitioned into an insecure SPARC processor (not fabricated) and secure coprocessor IC (the fabricated IC) [5]. The insecure SPARC processor implements the feature extraction algorithm and the wireless communication protocols. The secure coprocessor IC implements the remaining biometric and cryptographic components. In particular, as shown in Fig. 1, the coprocessor IC consists of four components: an Advanced Encryption Standard based cryptographic engine, fingerprint matching engine (oracle), template memory, and an interface unit. These components are described further in the following sections.

B. AES-Based Cryptographic Engine

The cryptographic engine consists of an AES core together with a controller, registers, and an interface to read/hash the memory. The datapath is based on one round of the AES-128 algorithm with on-the-fly key scheduling. The AES core is optimized for speed, with a goal of minimizing the delay for one round.

Different feedback and nonfeedback modes of operation are required for the secure encryption of data. In our application the crypto engine performs AES encryption in ECB (Electronic CodeBook), OFB (Output FeedBack), and CBC-MAC (Cipher Block Chaining Message Authentication Code) modes without any loss in throughput compared to a plain encryption. Fig. 2 shows how these modes are implemented for a typical block cipher. Due to the feedback in these modes of operations the block cipher core cannot be pipelined. Different registers that contain intermediate values of data and key are used as well as logic that implements the ECB, OFB, and CBC modes of operations.

Fig. 3 shows one round of the AES algorithm core together with the registers and feedback paths to implement the different modes of operation [6]. The architecture of one round contains two different datapaths, the encryption datapath and the key scheduling datapath. In the AES algorithm the data block is 128 bits long and the key size can be 128, 192, or 256 bits. The fabricated coprocessor implements the AES-128 algorithm, in which both the key length and the input size is 128 bits. The

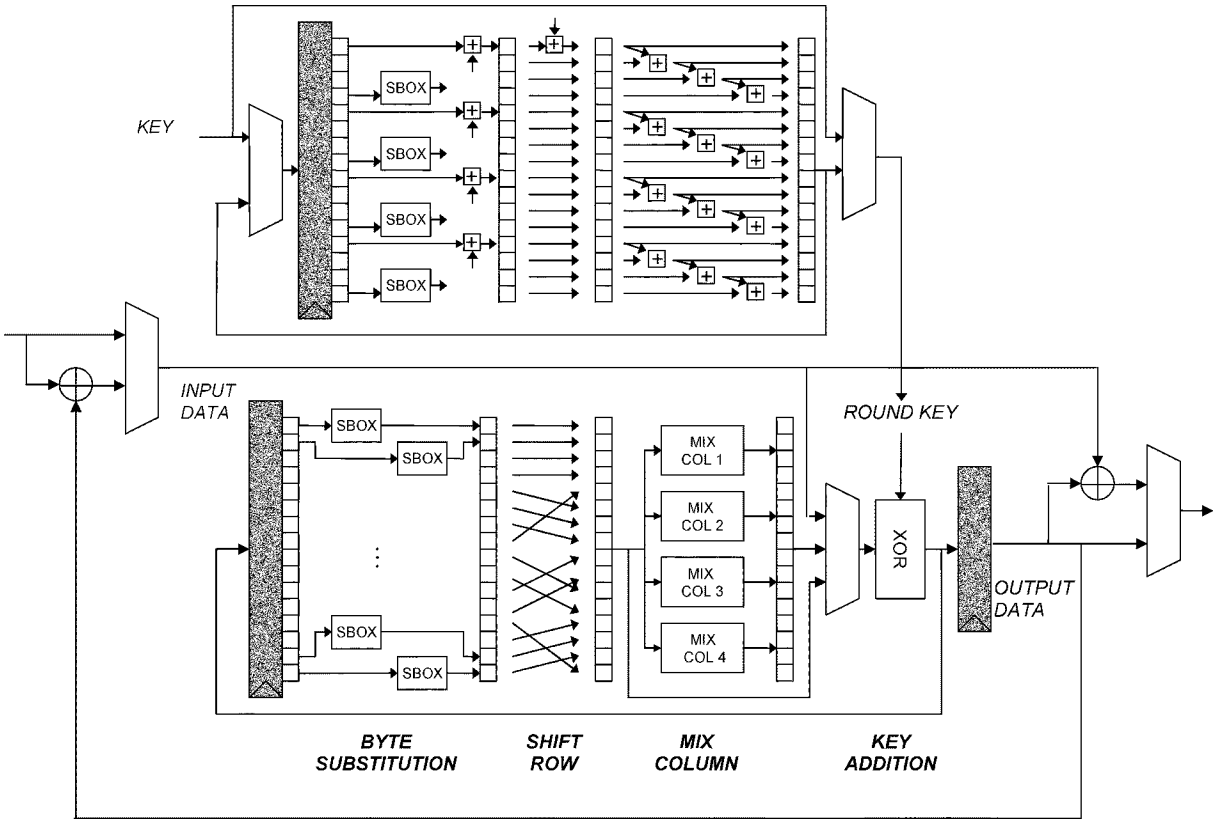


Fig. 3. AES round micro-architecture.

total number of processing rounds for AES-128 is 10 rounds plus a pre-processing round.

There are four sequential steps in each round of the encryption datapath. These are substitution, shift row, mix column, and key addition. Each 128-bit data value is operated on as an individual 8-bit byte (for a total of 16 bytes), a detail whose importance will be addressed in the DPA attack section of this paper. The description of each step of the algorithm is as follows.

Byte Substitution: This step is a nonlinear operation that substitutes each byte of the round data independently according to a substitution table (SBOX). The look-up table implementation of the byte substitution phase is used in the fabricated IC.

Shift Row: This step is a circular shifting of bytes in each row of the round data. The number of shifted bytes is different for each row and is accomplished by a gate-free permutation of physical wires.

Mix Column: In this step the bytes of each column are mixed together by multiplying the round data with a fixed polynomial modulo $x^4 + 1$. The mix column step is implemented using a chain of XORs which results in the minimum delay implementation for this unit.

AddKey: In this step the round data is XOR'd with the round key, which is generated from the key scheduling datapath.

All the above four steps are required for every round except the last round, which does not include the mix column phase. Similar steps are followed in the key scheduling flow. Each datapath round is completed in one coprocessor clock cycle, thus the total number of clock cycles required to complete an AES-128 encryption (including pre-processing) is 11.

C. Fingerprint Template Memory

The second module of the coprocessor IC is a memory element used to store a secure fingerprint template of up to 30 fingerprint minutiae, as shown in Fig. 4. Each minutia is composed of its own angle value (5 bit) relative to the horizontal axis, called *sita*, as well as a 19-bit field for each of its six closest minutia neighbors. The 19-bit field is composed of the distance to neighbor *dis* (8 bit), the angle to the neighbor *phi* (6 bit), and the angle of the neighbor relative to the horizontal axis *sita_nei* (5 bit). Thus, each minutia requires a storage size of $5+6\times 19 = 119$ bits. The maximum template size is thus 3570 bits for 30 minutiae. Since each minutiae possesses a common *sita* value and six different neighbor fields, the template memory was decomposed into a SELF memory bank of 30 words $\times$ 5 bits and a NEIGHBOR memory bank of 180 words $\times$ 19 bits, as shown in Fig. 4. The memory banks were implemented as register banks, and a memory controller was designed to allow secure access to the memory by the oracle engine and the cryptographic engine.

D. Fingerprint Matching Oracle

The fingerprint matching engine of the coprocessor IC is called the oracle and is able to perform a neighbor-based matching algorithm. The algorithm requires two parties: an untrusted feature extraction agent implemented on the SPARC processor and a secure matching agent (the oracle) implemented on the coprocessor IC. The operation of the oracle is as follows: after the insecure portion of the device obtains the candidate fingerprint of a user, the SPARC performs a neighbor-based

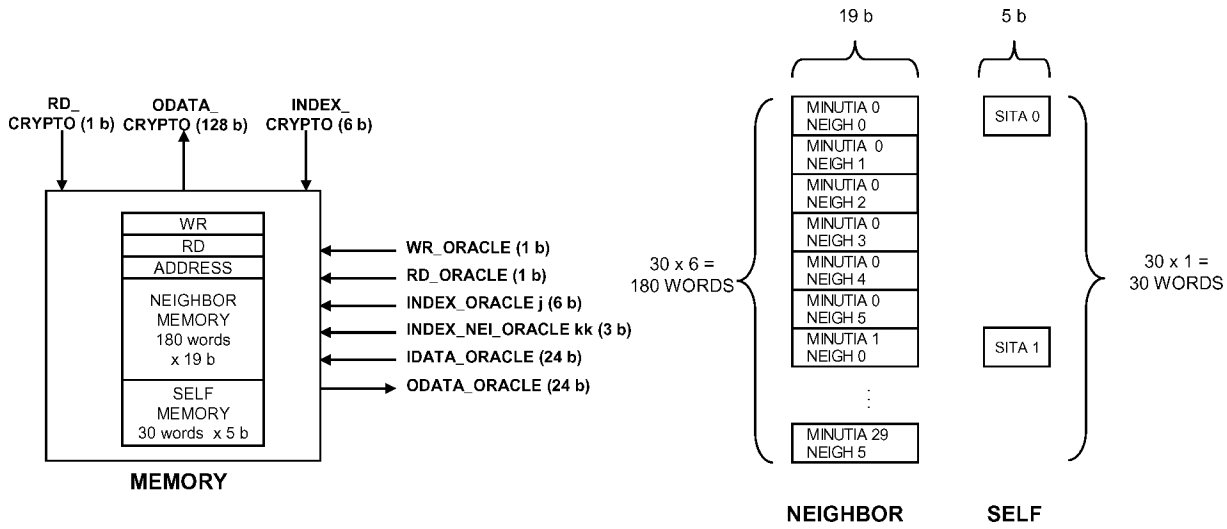


Fig. 4. Template memory decomposition.

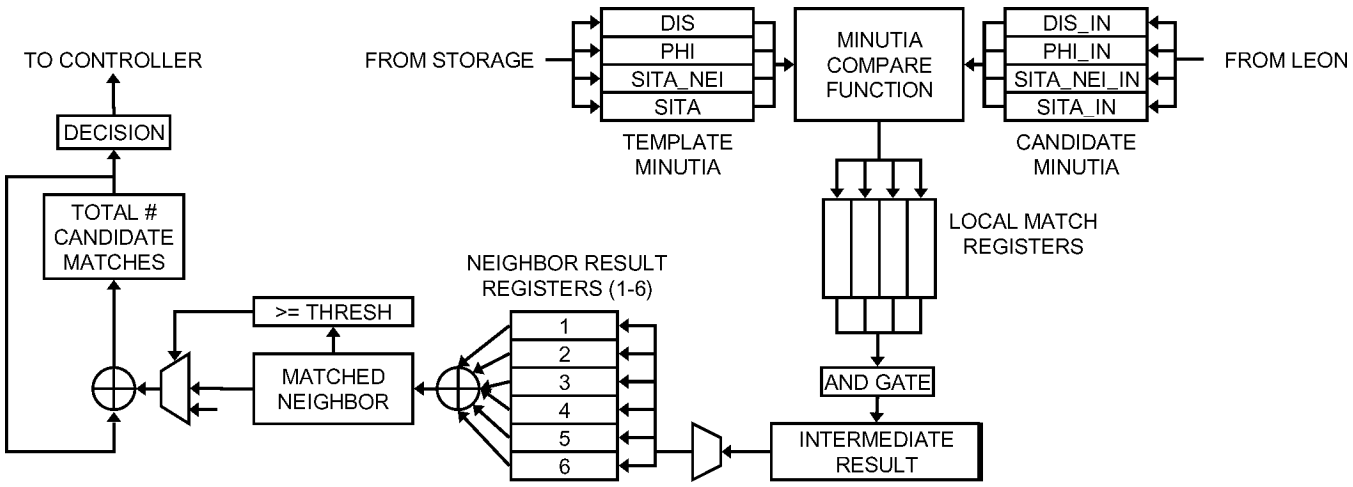


Fig. 5. Fingerprint matching oracle architecture.

feature extraction algorithm to extract the candidate minutiae set. This minutiae set is then mapped into a format suitable for the oracle.

At this point, the oracle operates to compare the candidate minutiae set with a pre-stored template minutiae set and to form a binary accept/reject decision. The oracle performs the comparison via a secure data exchange protocol between the insecure SPARC and the secure oracle engine. Data is sent from the SPARC to the oracle via a series of queries, each query consisting of a 24-bit data value and a 9-bit index value. The 24-bit data value is a single candidate minutia's 5-bit angle value *sita_in*, 8-bit distance to neighbor value *dis_in*, 6-bit angle to neighbor value *phi_in*, and 5-bit angle of neighbor value *sita_nei_in*. The 9-bit index value consists of two indexing terms, 6 bits to indicate which template minutia (j) and 3 bits to indicate which of the template minutia's neighbors (kk) should be compared for this particular query, as shown in Fig. 5. At each query the oracle implements a correlation function between the 24-bit candidate minutia and the requested 24-bit template minutia section, stores intermediate decision values, and waits for the next query from the SPARC processor. Upon finishing the query process, the oracle uses the intermediate decision values to produce a final accept/reject decision. This final decision is later passed to the cryptographic engine as the security flag to control access to the memory and enable the cryptographic engine.

Note that to prevent adaptive query attacks, the oracle does not provide intermediate feedback to the SPARC during the query phase, hence its name of oracle. The implemented matching oracle algorithm was tested to have a false accept rate (FAR) of 0.01% and a false reject rate (FRR) of 1.5%.

E. IC Interface Unit

The interface unit allows access to the IC by means of a 20-bit instruction/data input bus and a 17-bit output bus. The unit uses pipelined registers with logic gates to ensure stable data processing with one- or two-sided handshaking protocols. The coprocessor can operate with a 50-MHz SPARC processor within a range of clock frequencies from 1 to 288 MHz.

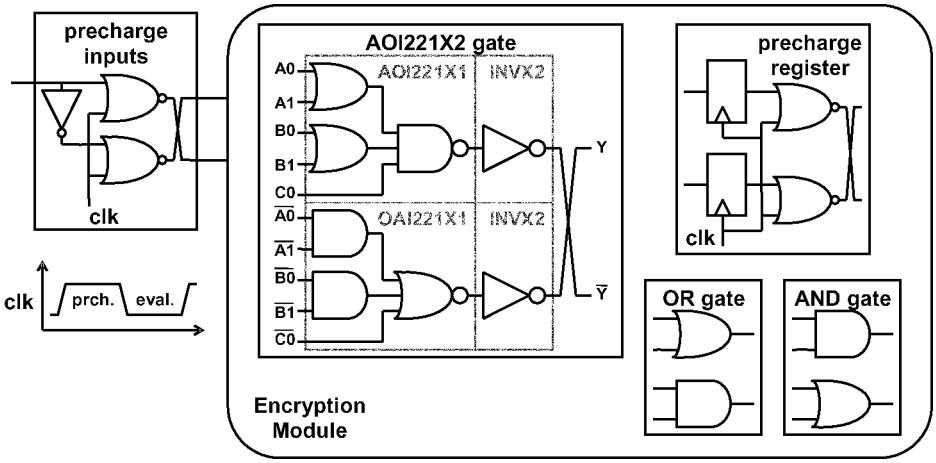


Fig. 6. Wave dynamic differential logic (WDDL).

III. Differential Power Analysis Countermeasures

This section of the paper describes the circuit-level countermeasures used to combat differential power analysis. As described earlier, in standard complementary CMOS logic, the building blocks of most modern ICs, the transition which causes the main dynamic power dissipation from the power supply is a 0 to 1 output transition. Kocher [3] has shown that the asymmetry in power demand causes information leakage. Therefore, the secret key of an encryption circuit can be successfully deduced by analyzing the statistical properties of power traces. At first, DPA was fought with ad hoc countermeasures. For instance, the addition of random power consuming operations obscured the data dependent power variations. Subsequently, countermeasures have been conceived at different abstraction levels. For instance at the algorithmic level, masking prevents intermediate variables from depending on an easily accessible subset of the secret key. Algorithmic countermeasures however, need to be reformulated for each algorithm and oftentimes the proposed solutions add overhead to the system and may be insecure afterwards [7]. Instead of masking at the algorithm level, the approach in this paper is to implement circuit techniques that avoid creating any side-channel information.

The goal of our approach is to make the power consumption of the individual logic gates constant and independent of their input signals (i.e., 0 to 0, 0 to 1, 1 to 0, and 1 to 1 transitions all draw the same power from the supply). The major advantages of this type of approach are that it is correct by construction, is independent of the cryptographic algorithm or arithmetic implemented, and is a distributed countermeasure which cannot be tampered with or corrupted. However, just as masking and other approaches incur overhead, the approach described in this paper also creates overhead in terms of power increase, area increase, and performance decrease.

Two conditions must be satisfied to have constant power dissipating logic: 1) a logic gate must have exactly one charging event per clock cycle, and 2) the logic gate must charge a constant capacitance in that event. The fabricated IC uses a technique called wave dynamic differential logic (WDDL) to fulfill the first condition, and a differential routing technique to fulfill the second condition.

A. Wave Dynamic Differential Logic: Constant Power Dissipating Logic

As described earlier, standard cell static CMOS logic used in a normal manner produces asymmetries in the power signature that can be exploited with DPA. Hence, other circuit topologies must be examined for DPA resistance. Dynamic logic has the property that during a precharge phase the output node is charged to VDD and during an evaluation phase the node is conditionally discharged to ground. Hence, a dynamic logic gate still possesses asymmetries in its power signature due to the input-dependent conditional discharge.

However, consider dynamic differential logic, also known as dual-rail with precharge logic. The dynamic differential logic gate takes in complementary inputs x and $\bar{x}$ and produces complementary outputs y and $\bar{y}$. In this topology, during the precharge phase exactly one node is precharged to VDD and during the evaluation phase exactly one output node is discharged to ground. Hence, dynamic differential logic possesses the desired property of one charging event per cycle.

The fabricated IC uses WDDL [8] to mimic dynamic differential behavior using static CMOS standard cells. A WDDL gate consists of a parallel combination of two positive complementary gates. A positive gate is defined as a gate that produces a zero output for an all-zero input. A complementary (or dual) gate computes the false output of the original logic gate using the false inputs of the original gate. Thus, a compound AND gate would consist of $y = a \cdot b$ and $\bar{y} = \bar{a} + \bar{b}$. Fig. 6 shows the WDDL AND and OR gates.

WDDL requires a precharge phase and an evaluation phase. In the precharge phase, both true and false inputs are set to 0. This causes the output of all gates (true and false) to be 0. This 0 precharge value travels as the input to the next gate, creating a precharge "wave." In the evaluation phase, each input signal is differential and the WDDL gate calculates a dynamic differential output.

Special registers and input converters, shown in Fig. 6, launch the precharge value. They produce an all-zero output in the precharge phase (clock high) but let the differential signal through during the evaluation phase (clock low). Since any gate in which the AND and OR operator are combined is positive and since all

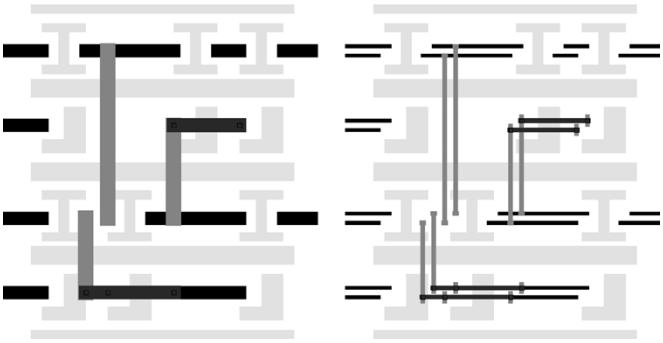


Fig. 7. Differential routing decomposition of "fat" wires into differential wires.

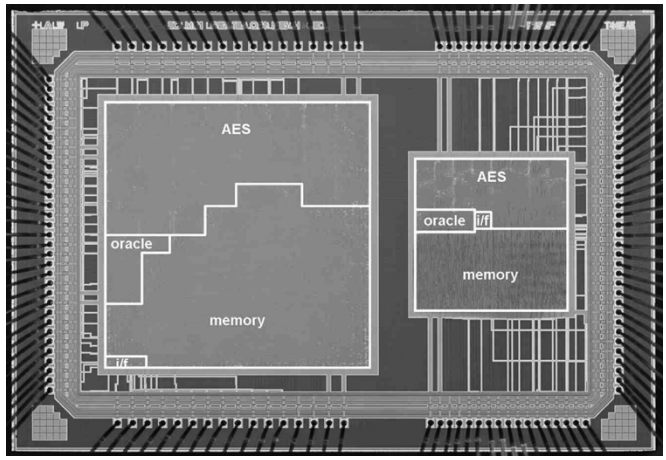


Fig. 8. IC micrograph. (left) Secure coprocessor using WDDL and differential routing. (right) Insecure coprocessor using standard cells and regular routing.

signals exist in differential form, any complex gate can be built. Thus, Fig. 3 also shows the composition of the WDDL AOI221 gate with drive strength 2.

WDDL has many advantages, including the fact that it can be readily implemented from an existing and fully supported standard cell library and results in a dynamic differential logic with only a small load capacitance on the precharge control signal and with the low power consumption and the high noise margins of static CMOS. It has the disadvantages of requiring more area and dissipating more power than a full-custom dynamic differential library designed for DPA resistance.

B. Differential Routing: Matching Interconnect Capacitances of Dual Rail Logic

Besides a 100% switching factor, to preserve power symmetry it is essential that a fixed amount of capacitance is charged during the transition. Thus, the total load at the true output of the differential gate should match the total load at the false output. The load capacitance has three main components: 1) the intrinsic output capacitance of the gate; 2) the interconnect capacitance; and 3) the intrinsic input capacitance of the load. For high security applications, the contribution of all components must be constant. Matching intrinsic input and output capacitances is a one-time task that can take place during the construction of the library. However, as the channel-length of transistors shrink, the share of the interconnect capacitance in the total load capacitance increases and the interconnect capacitances become a dominant capacitance for an increased percentage of routes [9]. Hence, the issue of matching the interconnect capacitances of the signal wires is crucial for the countermeasure to succeed.

Matched interconnect capacitances can be obtained by routing the true and false output signals with parallel routes that are at all times in adjacent tracks of the routing grid, on the same layers, and of the same length. Then independent of the placement, the two routes have the same first order parasitic effects.

Differential pair routing has been available through gridless routers. But the goal of gridless routers is to route a few critical signals, such as the clock or general reset signal. High-capacity gridded routers on the other hand have no or only limited capability to route differential pairs. We have recently presented a way to work around tool limitations [10] which we call differential routing. In the technique, each differential output pair is abstracted as a single "fat" wire, which has among other characteristics the width of two parallel wires plus spacing. The differential design is routed with the fat wire and during script-based post-processing the fat wire is decomposed into the differential wire (true and false nets). Fig. 7 demonstrates the place and route approach. At the left, the result of the fat routing is shown. At the right, the result after decomposition is shown. For the secure part of the prototype IC, the capacitances at the true and the corresponding false signal nets, directly reported from Silicon Ensemble using Simcap, have exactly the same values. The second order parasitics are not reported by this tool.

In summary, the secure digital design flow is completely supported by mainstream EDA tools and uses a commercially available static CMOS standard cell library. The differences with a regular synchronous CMOS standard cell design flow are minor. The secure digital design flow starts from a normal design in a hardware description language (HDL) and only a few key modifications are incorporated in the backend of the design flow. A cell substitution phase and an interconnect decomposition phase parse intermediate design files. The former procedure modifies the gate level description, the latter duplicates and translates the interconnect wires. The additional steps only required six minutes of CPU time for the prototype IC.

IV. Prototype IC Results

The prototype IC consists of two functionally identical coprocessors, fabricated on the same die using a TSMC 6M 0.18-μm process. An insecure coprocessor, which serves as a benchmark, was implemented using standard cells and regular routing techniques. A secure coprocessor was implemented using WDDL and differential routing. Both coprocessors have been implemented starting from the same synthesized gate level netlist. A die micrograph of the IC is shown in Fig. 8.

This section of the manuscript describes the test results of the fabricated IC. The results are divided into two sections: DPA resistance results and performance results.

A. DPA Resistance Test Results

A differential power analysis attack on the AES coprocessor core can be performed using a correlation attack on the transient

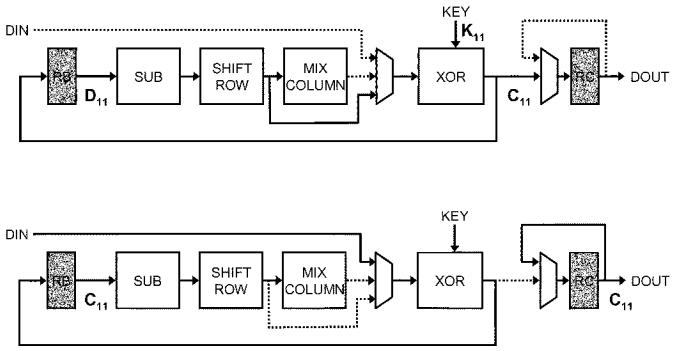


Fig. 9. AES core: round 11 (top); and round 11 + 1 (bottom).

signature of the core IC power supply. To attack our system, consider Fig. 9. During one encryption, the AES core encrypts a 128-bit plaintext P, using a 128-bit key K, to produce a 128-bit ciphertext C_{11} after 11 rounds. Note that the original 128-bit K is broken up into different 128-bit round keys (K_1 through K_{11}) corresponding to the 11 rounds of the AES algorithm, as shown in Fig. 3. Due to the reversible nature of the round key computation algorithm, once K_{11} is deduced, it is easy to reverse the algorithm and find the original key K. Thus, our attack attempts to obtain the 128-bit round key K_{11}. In addition, due to the byte-processing structure of AES, the attack can take place byte by byte. Using the same measured data, each of the sixteen bytes of K_{11} ($K_{11[0]}$ through $K_{11[15]}$) can be hacked separately.

1) DPA Attack Methodology: To perform an attack on the standard cell coprocessor, an estimation of the power consumption in round 11 + 1 was compared to a measurement of the power consumption in round 11 + 1, as shown in Fig. 9. To obtain an estimate of the power consumption we choose to attack register RB as it transitions from round 11 to round 11 + 1. After round 11 + 1, the value stored in RB is simply the known final ciphertext C_{11}. The value stored in RB in round 11 (D_{11}) can be found by tracing back the signal obtained after XOR-ing the final ciphertext (C_{11}) and the round key (K_{11}) through both the shift row operation and the substitution box (recall that the mix column is not performed in the last round). The Hamming weight of the difference between D_{11} and C_{11} is an indication of the power transitions which took place on RB, as bits switched from 0 to 1 and 1 to 0. Each key byte of K_{11} is 8-bit and thus can take on a value between 0x00 and 0xFF, for a total of 256 possibilities. Thus, for each key byte there are 256 possible power signatures for a particular plaintext-ciphertext-key combination, one of which is the correct signature. These groups of power signatures are called P_{guess}.

On the experimental side, the maximum value of the current was measured during round 11 + 1. This measured value, called P_{measured}, corresponds with the number of transitions on RB. From round 11 to 11+1, for each of the 16 bytes of K_{11}, a correlation is performed between P_{measured} and the 256 variants of P_{guess}. The value from 0x00 to 0xFF with the highest correlation is selected as the correct key-byte guess for K_{11}. Mathematically, this implies a search for

$$\max f_{\text{cost}}(K_{11}) = \text{correlation}(P_{\text{guess}}, P_{\text{measured}})$$

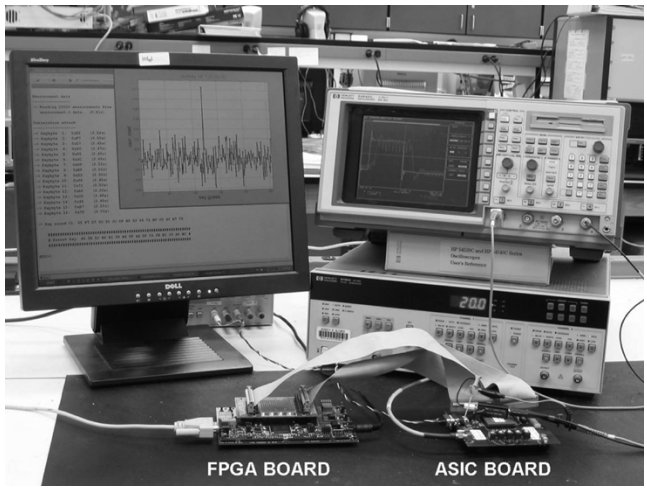


Fig. 10. DPA measurement and attack setup.

where

$$P_{\text{guess}} = \text{HamDist}(D_{11}, C_{11})$$
$$D_{11} = \text{sub}^{-1}(\text{shiftrow}^{-1}(K_{11} \oplus C_{11})$$
$$P_{\text{measured}} = \max(I_{\text{supply},11+1}).$$

Of course, the correlation may be inaccurate due to noise for only one measurement (i.e., one set of P and C_{11}). Hence, thousands of different (P, C_{11}) pairs were measured using the same key K (and hence the same K_{11}) in order to filter out the noise and provide a correct correlation.

For the WDDL coprocessor, we only need to look at a single round, as all signals are at 0 at the start of each evaluation phase. The number of changing bits of RB in round eleven, when we do the current measurements, is therefore the Hamming weight of RB.

2) DPA Attack Experimental Setup: The measurement and analysis setup is shown in Fig. 10. The core supply current is measured between a custom-designed printed circuit board's decoupling capacitances and the IC. A CT1 current probe from Tektronix with a 25-kHz to 1-GHz bandwidth measures the supply current variations. For every mA, it provides a 5-mV output to the HP54542C oscilloscope. The oscilloscope filters the waveform transients at 500 MHz and digitizes them at a 2-GHz sampling frequency. To facilitate the synchronization of the measurements, we also have access to the encryption start signal. A clock of 50 MHz is provided to the coprocessor under attack, for which only the AES core processes data. The attack works on one byte at a time. During that time the other 15 bytes operate and contribute to the noise. The other circuits and modules on the regular coprocessor are quiet, while for the attack on the WDDL coprocessor, they always have the same switching events.

3) DPA Attack Measurements and Results: As an illustration of the power-varying nature of standard cell CMOS, Fig. 11 shows the encryption start signal and the core supply current in the actual attack. The supply current of the standard cell coprocessor easily reveals the encryption operation: one can count exactly eleven peaks. The secure coprocessor has a continuous

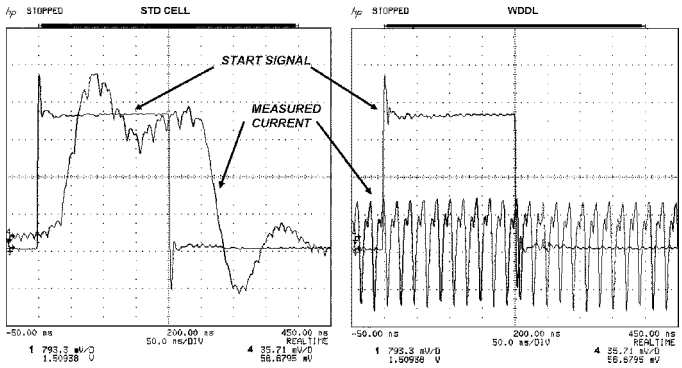


Fig. 11. Transient measurement of encryption start signal and core supply current for single encryption: (left) standard cells and regular routing, and (right) WDDL and differential routing.

current whether or not data is being processed, either cryptographic or other. If an attacker does not have access to the encryption start signal, it is almost impossible to know when the IC is encrypting.

The resistance against DPA can be quantified as the number of measurements to disclosure (MTD). We define the MTD as the crossover point between the correlation coefficient of the correct key and the maximum correlation coefficient of all the wrong key guesses. For both coprocessors, attacks on two key bytes are shown in Fig. 12; the results for the other fourteen key bytes are similar. The MTD is shown in the "Correlation versus Number of Measurements" graphs as the point where the black line (correct key) crosses the gray envelope (wrong keys). The maximum number of measurements we took is 15 000 and 1 500 000 for the standard cell and the WDDL coprocessor, respectively.

For the standard cell implementation, the correct key bytes are found easily, as shown by the large signal-to-noise ratio on the "Correlation versus Key Guess" graphs in Fig. 12(a) and (b). On average, about 2000 measurements are required to disclose a secret key byte for the insecure coprocessor. In one case, a mere 320 samples is sufficient to mount a successful attack. From the graphs, there is no doubt about which byte is the correct key byte. It should be noted that the MTD is a metric that can be used when the correct key is known *a priori* to the attack; in a blind attack on the system this metric may not be applicable as the largest peak after a certain span may in fact not be the correct one.

The WDDL coprocessor, on the other hand, substantially reduces this signal-to-noise ratio of correlation, shown by the small correlation peaks in the "Correlation versus Key Guess" graphs in Fig. 12(c) and Fig. 12(d). Our measurements show that out of sixteen key bytes, WDDL effectively protects five key bytes. In other words, after 1.5 million measurements, five key bytes could not be broken. One example of such a protected key byte is in Fig. 12(d); as seen in the "Correlation versus Number of Measurements" graph, the black line (correct key) never escapes the envelope of the gray lines (wrong keys). The eleven key bytes that were found required on average 255 000 measurements, an increase of more than two orders of magnitude when compared to the average for the standard cell coprocessor. One of these found key bytes for the WDDL

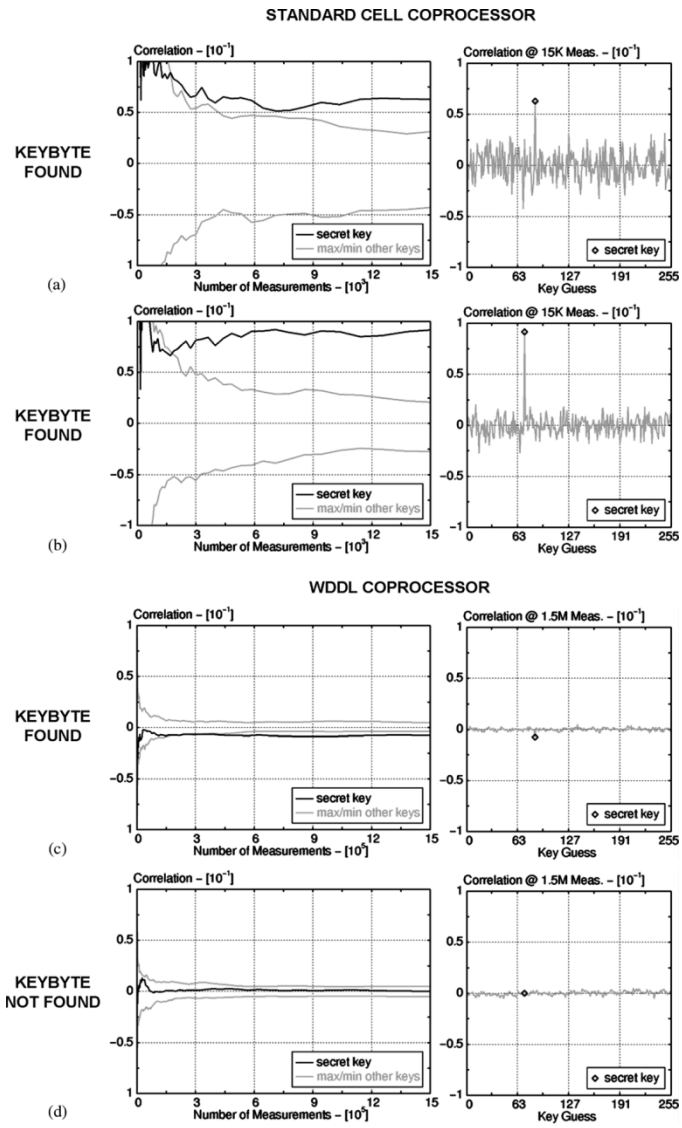


Fig. 12. Cracking the secret key. (a) Standard cells and regular routing using 15K measurements—keybyte found. (b) Standard cells and regular rouing using 15K measurements—keybyte found. (c) WDDL and differential routing using 1.5 M measurements—keybyte found. (d) WDDL and differential routing using 1.5 M measurements—keybyte not found.

coprocessor is shown in Fig. 12(c). A brute force attack can be mounted to reveal the remaining 5 keybytes. Without the synchronization signal, however, it is almost impossible to mount the attack as one does not know when to measure. In an actual attack, one does not have access to the synchronization signal.

The analysis also revealed that for a dual rail design, the correlation coefficient of the correct key guess can be negative, as shown in Fig. 12(c). This means that less power is consumed as more bits change. This implies that the 0 to 1 switching of the false net uses more power than the 0 to 1 switching of the true net. In WDDL, this means the parasitic capacitances affected by the false signals are larger than the ones affected by the true signals. On the other hand, for the five bytes that have not been found, the capacitances have an almost perfect matching between the differential nets. Hence, it is crucial to guarantee matched capacitances consistently for all the logic.

Further techniques to improve capacitance matching include making every other metal layer a ground plane, which would completely control the capacitance to other layers. Shielding the differential routes on either side with a power line would eliminate the cross-talk to adjacent wires in the same metal layer. Alternatively, increasing the distance between different differential pairs would reduce the effect, or an iterative design flow could be used to identify and correct mismatches.

At the circuit topology level, some power asymmetries are inherent in the WDDL structure. This is due to the fact that, while the AND and OR gates are logically complementary gates, in physical implementation the WDDL AND and OR gates differ in structure at the transistor level. To address this asymmetry, a full-custom library of gates can be built using transistor-level complementation.

Furthermore, it should be noted that WDDL avoids creating side-channel information specifically from dynamic power. However, leakage power may be data-dependent and represents a side-channel that is not directly addressed by WDDL. At the measured temperature and process technology our results show that leakage is not a viable side-channel. However, as temperature increases and/or as leakage power becomes more prominent in technologies 90 nm and below, leakage current may form a viable side-channel attack point. Whether WDDL is able to combat leakage side-channel information is uncertain.

B. IC Performance Results

Functionally, the full system architecture (including all blocks of the coprocessor and the interface with the 50-MHz SPARC) operated at 288 MHz for a full cryptographic and biometric protocol for the standard cell coprocessor. For the WDDL processor, the maximum clock frequency was 69 MHz. Using BIST, the standard cell AES was able to operate in all modes of operation (ECB, CBC-MAC, OFB) at a maximum of 330 MHz, which is equivalent to 3.84 Gb/s. As far as we know, this is the fastest nonpipelined AES encryption rate implemented in actual silicon. (Other work [11], however, based on simulations using 0.13-μm CMOS has shown a faster AES core.) The WDDL AES was able to operate at a maximum of 85.5 MHz, which is equivalent to 0.99 Gb/s.

In terms of power consumption, for the standard cell coprocessor at 50 MHz, the AES and full system architecture consumed 54 and 36 mW, respectively. The full system architecture consumed less power than the AES in feedback mode due to the fact that the AES core is not in full operation during the entire verification protocol. For WDDL at 50 MHz, the power consumption results are 200 and 486 mW, respectively.

Table I summarizes the results for the fabricated coprocessor IC in tabular format. WDDL and differential routing is a technique proven to thwart power attacks by improving DPA resistance orders of magnitude over a standard cell IC implementation. The trade-off of using such techniques is an increase in area by three times, an increase in power by four times, and a reduction of maximum clock frequency by four times. Recall that by performing security partitioning, the careful division of the architecture into two parts (a secure and an insecure part), this overhead is minimized for complex embedded systems.

TABLE I
IC RESULTS SUMMARY

Parameter	Standard Cell	WDDL
Gate Count (eq. gates) [K]	199	596
Area [mm^2]		
AES	0.79	2.45
Oracle	0.11	0.26
Memory	1.05	3.21
Entire System	1.98	5.95
Maximum Frequency (@1.8V) [MHz]		
AES	330.0	85.5[a]
Entire System	288.2	69.0[a]
Maximum Throughput (@1.8V) [Gb/s]		
AES	3.84	0.99
Power Consumption (@1.8V,50 MHz) [W]		
AES	0.054	0.200[b]
Entire System	0.036	0.486
Measurements to Disclosure[c]		
Min	320	21,185
Mean	2,133	255,391
Max	8,168	1,276,186
Keybytes not found (@1.5M meas.)	n/a	5

[a]Duty factor of clock > 50% to guarantee precharge of all gates
[b]Estimation based on area ratio AES vs. Entire System
[c]Based on correctly guessed key bytes

V. RELATED WORK

As far as we know, our work is the first published DPA-resistant circuit-plus-routing technique implemented and tested in actual silicon. Other published countermeasures have either never been implemented in silicon, never been measured and attacked, or did not offer any significant DPA resistance.

A dual rail asynchronous chip has been presented previously [12]. The implementation did not provide a significant increase in DPA resistance. This failure has been attributed to unbalanced signal paths caused by routing differences. Note that if asynchronous logic is used to increase the DPA resistance, dual rail encoded asynchronous logic must be used. Because of the dual rail logic, there is also a factor of three area increase compared with a single ended synchronous benchmark. As described earlier, masking is another technique proposed to protect IC circuits against DPA attacks. In [13], Mangard *et al.* have shown that masked implementations can be broken when glitches are present in the circuit, which is the case for regular CMOS implementations. In [14], this is confirmed and measured on an actual IC implementation.

VI. CONCLUSION

This paper has described a security coprocessor IC that does not leak cryptographic side-channel information through the power supply, which is a major and easy-to-access side-channel leakage source. The coprocessor IC contains processing engines for symmetric-key cryptography and biometrics for use in embedded security applications. Built in a 0.18-μm CMOS technology, we believe that this is the first IC that is practically immune to DPA attacks. Its immunity has been experimentally verified and compared to a functionally-identical coprocessor built with a regular standard cell approach. We have presented the measurement setup and analysis technique. Experimental results showed that 1 500 000 acquisitions are not sufficient to fully disclose the 128-bit secret key.

REFERENCES

[1] M. Renaudin, F. Bouesse, P. Proust, J. Tual, L. Sourgen, and F. Germain, "High security smart-cards," in *Proc. Design Automation and Test in Europe Conf. (DATE)*, Feb. 2004, pp. 228–233.

[2] P. Kocher, R. Lee, G. McGraw, A. Raghunathan, and S. Ravi, "Security as a new dimension in embedded system design," in *Proc. Design Automation Conf. (DAC)*, Jun. 2004, pp. 753–760.

[3] P. Kocher, J. Jaffe, and B. Jun, "Differential power analysis," in *Proc. Advances in Cryptology (CRYPTO)*, Aug. 1999, LNCS 1666, pp. 388–397.

[4] A. Lenstra and E. Verheul, "Selecting cryptographic key sizes," in *Int. Workshop on Practice and Theory in PublicKey Cryptography (PKC)*, 2000, LNCS 1751, pp. 446–465.

[5] K. Tiri, D. Hwang, A. Hodjat, B. Lai, S. Yang, P. Schaumont, and I. Verbauwhede, "AES-based cryptographic and biometric security coprocessor IC in 0.18-μm CMOS resistant to side-channel power analysis attacks," in *Symp. VLSI Circuits Dig.*, Jun. 2005, pp. 216–219.

[6] National Institute of Standards and Technology, Advanced Encryption Standard. [Online]. Available: http://csrc.nist.gov/publications/fips/fips197/fips-197.pdf 2001

[7] E. Oswald, S. Mangard, and N. Pramstaller, "Secure and efficient masking of AES-a mission impossible?," *IACR Cryptology ePrint*, 2004.

[8] K. Tiri and I. Verbauwhede, "A logic level design methodology for a secure DPA resistant ASIC or FPGA implementation," in *Proc. Design Automation and Test in Europe Conf. (DATE)*, Feb. 2004, pp. 246–251.

[9] International Technology Roadmap for Semiconductors (ITRS), Interconnect. [Online]. Available: http://public.itrs.net/Files/2003ITRS/Interconnect2003.pdf 2003

[10] K. Tiri and I. Verbauwhede, "Place and route for secure standard cell design," in *Proc. IFIP Smart Card Research and Advanced Application Conf. (CARDIS)*, Aug. 2004, pp. 143–158.

[11] S. Morioka and A. Satoh, "A 10-Gbps full-AES design with a twisted BDD S-box architecture," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 12, no. 7, pp. 686–691, Jul. 2004.

[12] J. Fournier, S. Moore, H. Li, R. Mullins, and G. Taylor, "Security evaluation of asynchronous circuits," in *Proc. Workshop on Cryptographic Hardware and Embedded Systems (CHES)*, Sep. 2003, LNCS 2779, pp. 137–151.

[13] S. Mangard, T. Popp, and B. Gammel, "Side-channel leakage of masked CMOS gates," in *Proc. RSA Conf. 2005 Cryptographers' Track (CT-RSA)*, Feb. 2005, LNCS 3376, pp. 351–365.

[14] S. Mangard, N. Pramstaller, and E. Oswald, "Successfully attacking masked AES hardware implementations," in *Proc. Workshop on Cryptographic Hardware and Embedded Systems (CHES)*, Aug. 2005, LNCS 3659, pp. 157–171.

David D. Hwang (S'02–M'06) received the M.S. degree from the University of California, Los Angeles (UCLA) in December 2001, researching architectures and ASIC implementations of VLSI digital signal processing systems. He received the Ph.D. degree in electrical engineering from UCLA in March 2005. His research focused on VLSI implementations and architectures for cryptographic and secure systems.

He is currently with KeyEye Communications investigating DSP architectures for multi-gigabit Ethernet transceivers.

Dr. Hwang was a UC Regents Scholar, an NSF Graduate Fellow from 1999 to 2000, and a Hertz Foundation Graduate Fellow from 2000 to 2005.

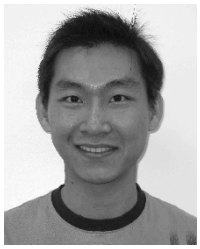
Kris Tiri (S'99–M'06) was born in Bree, Belgium, in 1976. He received the M.S. degree in electrical engineering from theKatholieke Universiteit Leuven, Belgium, in 1999, and the Ph.D. degree in electrical engineering from the University of California, Los Angeles (UCLA), in 2005. His doctoral research focused on design for side-channel attack resistant security ICs.

He is currently with the Trusted Platform Laboratory of Intel Corporation, Hillsboro, OR. From 1999 to 2005, he was a Research Assistant with the Electrical Engineering Department of UCLA. During the spring of 1999, he was with COMELEC of Ecole Nationale Supérieure des Télécommunications, Paris, France. During 2001–2002, he was with IMEC, Heverlee, Belgium, studying substrate noise modeling and reduction.

Dr. Tiri was awarded a Francqui Foundation Fellowship by the Belgian American Educational Foundation in 1999, and he received the 2005 EDAA Outstanding Dissertation Award.

Alireza Hodjat (S'99) received the B.S degree in electrical engineering from the University of Tehran, Iran, in 1999, and the M.S. degree in electrical engineering from the University of California, Los Angeles (UCLA), in 2002. He received the Ph.D. degree in 2005 in the field of embedded computing systems in the Emsec Group of the Electrical Engineering Department at UCLA.

He is currently with Broadcom Corporation, Irvine, CA. His research interests include hardware/software co-design and application specific instruction set coprocessor architectures and VLSI implementations for secure embedded systems.



Bo-Cheng Lai (S'01) received the B.S. degree in electronics engineering from National Chiao-Tung University, Hsinchu, Taiwan, R.O.C., in 1999, and the M.S. degree in electrical engineering from the University of California, Los Angeles (UCLA), in 2003, where he is currently pursuing the Ph.D. degree.

His research concerns the design of interconnect architectures for system on chip, multi-processor architectures, low power designs, networking embedded systems and wireless sensor networks.

Mr. Lai was the recipient of the Henry Samueli Fellowship for fall 2002. He also received a scholarship from the John Deere Foundation in 2003.

Shenglin Yang (S'03) received the B.S. and M.S. degrees in electronics from Beijing University, Beijing, China, in 1998 and 2001, respectively. She is currently pursuing the Ph.D. degree in electrical engineering at the University of California, Los Angeles.

Her research interests include biometrics, pattern recognition, embedded systems, and security.

Patrick Schaumont (M'97) received the M.S. degree in computer science from Rijksuniversiteit Ghent, Belgium, and the Ph.D. degree in electrical engineering from the University of California, Los Angeles (UCLA), in 1990 and 2004, respectively.

He is an Assistant Professor in the Electrical and Computer Engineering Department of Virginia Tech. Before joining UCLA in 2001, he was a researcher at IMEC, Belgium, from 1992. His research focuses on design methods and architectures for embedded systems, and he works in close cooperation with designers to demonstrate new methodologies on practical applications.

Ingrid Verbauwhede (M'92–SM'02) received the electrical engineering degree in 1984 and the Ph.D. degree in applied sciences from the Katholieke Universiteit Leuven (K.U.Leuven), Belgium, in 1991.

She was a Lecturer and Visiting Research Engineer at the University of California (UC), Berkeley, from 1992 to 1994. From 1994 to 1998, she was a Principal Engineer first with TCSI and then with Atmel in Berkeley, CA. She joined the University of California, Los Angeles (UCLA), in 1998 as an Associate Professor and joined the K.U.Leuven in 2003. At UCLA, she heads the embedded security group (EMSEC). At K.U.Leuven, she is co-director of the ESAT-COSIC research group. Her interests include circuits, processor architectures and design methodologies for real-time, embedded systems in application domains such as security, cryptography, digital signal processing, and wireless applications.

Prof. Verbauwhede was the general chair of the IEEE International Symposium on Low Power Electronic Devices (ISLPED) in 2003. She is or has been a member of several program committees, including DAC, ISSCC, DATE, CHES, ICASSP, SIPS, ASAP. She is the design community chair on the 42nd and 43th DAC executive community. For the IC presented in this paper, her students won Third Place in the operational category of the 2005 ISSCC/DAC student design contest.