Post-Layout Estimation of Side-Channel Power Supply Signatures

Sushmita Kadiyala Rao, Deepak Krishnankutty, Ryan Robucci, Nilanjan Banerjee and Chintan Patel
Department of Computer Science and Electrical Engineering
University of Maryland, Baltimore County
{sush1, deepakk1, robucci, nilanb, cpatel2}@umbc.edu

***Abstract*—Two major security challenges for integrated circuits (IC) that involve encryption cores are side-channel based attacks and malicious hardware insertions (trojans). Side-channel attacks predominantly use power supply measurements to exploit the correlation of power consumption with the underlying logic operations on an IC. Practical attacks have been demonstrated using power supply traces and either plaintext or cipher-text collected during encryption operations. Also, several techniques that detect trojans rely on detecting anomalies in the power supply in combination with other circuit parameters. Counter-measures against these side-channel attacks as well as detection schemes for hardware trojans are required and rely on accurate pre-fabrication power consumption predictions. However, available state-of-the-art techniques would require prohibitive full-chip SPICE simulations. In this work, we present an optimized technique to accurately estimate the power supply signatures that require significantly less computational resources, thus enabling integration of Design-for-Security (DfS) based paradigms. To demonstrate the effectiveness of our technique, we present data for a DES crypto-system that proves that our framework can identify vulnerabilities to Differential Power Analysis (DPA) attacks. Our framework can be generically applied to other crypto-systems and can handle larger IC designs without loss of accuracy.**

***Index Terms*—Hardware Security, Trojan Detection, Side-channel attacks, Power Supply analysis**

I. Introduction

The high cost of custom IC fabrication has given rise to a large number of fabless IC firms that outsource their design for manufacturing at third party foundries. This entails design transfer to untrusted, off-site foundries thus making fabricated ICs vulnerable to security compromise, malicious hardware modifications, and proprietary information leakage.

These vulnerabilities manifest themselves as hardware trojans or side channel leaks post-fabrication. In military systems, financial infrastructure, transportation and automotive devices, as well as household appliances these vulnerabilities can have deleterious effects. There is a need, therefore, for efficiently assessing these security vulnerabilities during design phase and aiding post-fabrication device testing to verify IC authenticity [1]. Conventional techniques for detecting trojans or side channel leakage that rely on processing power supply information require generation of *golden signatures* [1], [2], [3], [4]. These signatures quantify the power consumption of security conscious designs and untampered ICs. The golden signature is predominantly determined by performing exhaustive tests on a select set of ICs and measuring the power consumption. During device testing, the measured power consumption from the Chip-Under-Test (CUT) is compared against the golden signature to identify either malicious hardware insertions or side channel leaks. Such techniques suffer from two drawbacks: 1) they assume that this select set of ICs do not have malicious circuits and 2) exhaustive post-fabrication testing can be prohibitively expensive, even for a small subset of ICs.

Researchers have proposed on-chip power monitoring systems e.g. using ring oscillators [3] that measure the dynamic power in combination with off-chip measurement equipment to derive the golden signatures. These methods, though computationally cheaper, fundamentally suffer from the previously identified drawbacks. Additionally, these techniques cannot localize circuitry on the chip that is being attacked.

In this paper we present an accelerated pre-fabrication simulation framework useful for determination of golden signatures and evaluation of a design's side-channel leakage. We present our technique for augmenting IC design flow with Design for Security (DfS). We note that our technique supports augmenting an IC model with additional linear-components such as for package modeling. The design and evaluation of our technique presents the following research contributions.

- **An Efficient CAD Framework for performing simulations towards deriving the golden signatures that addresses security vulnerabilities:** We describe a technique for deriving the golden signatures used to detect IC-level security vulnerabilities such as side-channel attacks and hardware trojans. The framework can also be leveraged for applications in delay and transition based testing as well as power supply noise (PSN) estimation [11]. The technique optimizes power estimation using a partitioning technique that characterizes the power grid independent of the chip logic. It precomputes the gate-level current transients using path simulation which are then convolved with the Current-to-Current impulse responses to estimate the power consumption at each power pad in the IC.
- **Localize IC circuit elements under attack:** Our framework can estimate power consumption at each power pad on an IC accurately, and we leverage differences in the power signatures at multiple supply pads to localize the circuit element under attack. We demonstrate this feature in §IV by localizing the S-box that processes the bit under a Differential Power Attack (DPA).

978-1-4673-7421-7/15/$31.00

II. Background

Side-Channel Attacks allow an attacker to extract secret keys from a target device by monitoring the power supply, electromagnetic radiation or timing information.

Simple Power Analysis (SPA) involves direct interpretation of the power supply traces from the operation of interest. A Differential Power Analysis (DPA) attack was introduced by Kocher et al. [5] to identify secret keys from the CUT's power traces. They have since been used widely in the cryptographic community [8]. The single-bit and multi-bit DPA attack variations differ in the requirement for a larger number of guessed keys due to additional bits. These extra bits generate more intermediate values, unlike the two possibilities in the single-bit case, implying more than two groups to sort power traces. The groups are simplified and combined to create a single DPA trace for the guessed key. Correlation Power Analysis, first introduced in [6] and implemented in [7] utilizes a statistical approach to compute the correlation between power traces observed from the Device Under Test (DUT) and a power model based on Hamming Distance or Hamming Weight.

DPA attacks against the Data Encryption Standard (DES) algorithm [9] are primarily targeted at the first or final round of encryption. For the purpose of analysis, this paper focuses on the initial round of encryption.

III. Power Supply Prediction Technique

A convolution-based framework is presented that accurately predicts power consumption at each supply pad in the IC.

A. System Partitioning

Our estimation techniques are based on a modular framework proposed in [10]. They show that a digital chip can be partitioned into two independent subsystems, namely the linear Power Grid Circuit (PGC) and the non-linear Core Logic Circuit (CLC) as shown in Fig. 1.

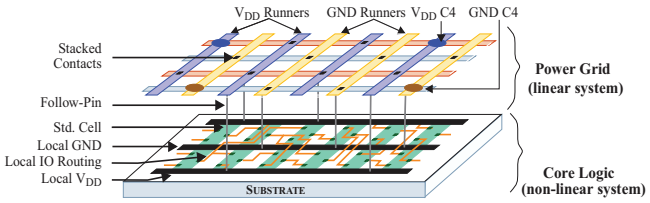

Fig. 1: Linear and non-linear components of a chip [10]

The power grid network is a multi-input, multi-output linear system. Since any linear time invariant (LTI) system can be characterized by its impulse response (IR), it is possible to characterize the response of the power grid to any arbitrary input signal (in this case, switching logic in the crypto-system).

B. Power Estimation Framework

Grid characterizations are carried out to compute IR responses between input and output locations of the power grid, Fig. 2. The current response to an input is computed using Current-to-Current impulse responses (C2C) that are computed by applying a step input and differentiating the response at an output. C2C IR provides the relationship between the current source applied at the input and the corresponding currents

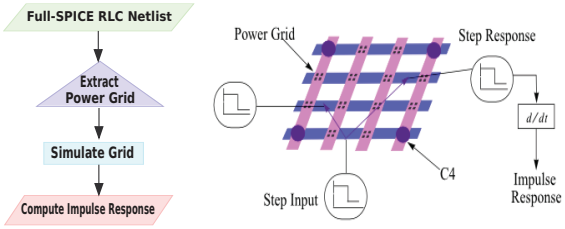

Fig. 2: Power Grid Characterization Flow

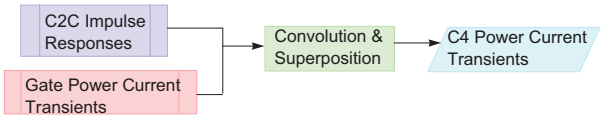

Fig. 3: Computation of Power Pad Transient Current

measured at any output location on the power grid. Switching gates are identified using a Verilog logic simulation and are extracted from a detailed RLC netlist. These isolated paths are simulated and provide current inputs to the power grid. Fig. 3 illustrates estimation of power transients by convolving grid IR responses with individual gate transients.

The layout of DES core is implemented in 180-nm technology using a commercial place-and-route tool. The chip has four V_{DD} and four GND power pads. In order to cater to much larger circuits, a deeper partitioning strategy can be used [11].

IV. Evaluation

In this section, we analyze the effectiveness of our power consumption estimation technique. The ground-truth is generated from full-chip SPICE simulations. Additionally, we evaluate the accuracy of our framework for application to DfS using a security metric described below.

A. Computational Performance

Full-chip SPICE simulations of the partitioned system were performed using the Cadence Spectre Accelerated Parallel Simulator^{TM}(APS) [12]. Table I compares the time required to simulate a full-chip SPICE netlist of the DES core, an individual path and the power grid.

TABLE I: Simulation Time for 60-ns

Component Simulated	Time
Full Chip SPICE	7hr 50m
Path SPICE using characterization	4hr 11m

The grid simulation for each gate location takes about 27 seconds. However, since this characterization is performed *only once* for a design, such run-time associated with the grid will be amortized over all the chip simulations. Thus, the computational advantage of our framework versus the full-chip SPICE will be substantial for a large number simulations which would be required for large designs. Furthermore, the segmentation of the design yields larger advantages for larger circuit netlists.

B. Correlation-Based Analysis of Side-Channel Attacks

The purpose of our simulation model is to predict power supply variations in response to both data and design variations. The ability to predict design dependent perturbations allows for an iterative design cycle towards improving security.

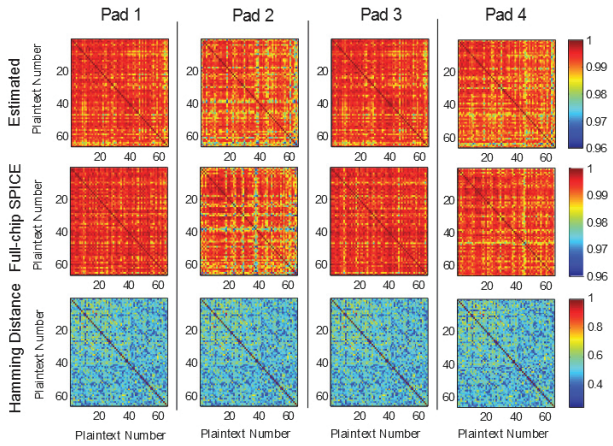

Fig. 4: Correlation analysis for 66 plain-texts: Row 1) our Framework, Row 2) Full-chip SPICE Row 3) Hamming Distance

We evaluate our framework using a security metric based on correlations and one that compares absolute similarity.

A system perfectly invulnerable to side-channel exploitation would either offer a data-invariant supply response, that is to say that no matter what the digital data is, the side-channel outputs (e.g. power transient, delay) are the same, or at least that the side-channel responses are uncorrelated to the data (perhaps randomized). Such an ideal system is likely impossible to design—though getting closer at a reasonable cost requires a tool for estimating leakage. Therefore, we compare the similarity/difference between a pair of waveforms using the correlation coefficient: $\rho(x,y) = \frac{\sum_i (x[i]-\mu_x)(y[i]-\mu_y)}{\sqrt{\sum_n (x[i]-\mu_x)^2 \sum_i (y[i]-\mu_y)^2}}$, where μ_x and μ_y are the mean of $x[i]$ and $y[i]$.

For each power pad, we compute the correlation between pairs of current transients using 66 different plain-texts, row 1 of Fig. 4. As a comparison, the same is done using full-chip SPICE. The values from the full-chip SPICE present the inherent challenge to the attacker. The correlation values are all high, meaning there is a low variation between waveforms as compared to the magnitudes of the transients. This means it is difficult to infer an input from a measurement of the power supply, particularly in the presence of noise in a physical system. Even so, previous work has shown that once enough transients are captured, hidden values may be uncovered, such as private keys [5]. Therefore, evaluating this weak, yet critical, dependency is important. In lieu of presenting analysis of vulnerability for every type of attack, we present, as a figure of merit, correlation between power draws using different data. Lastly, for comparison, the Hamming Distance of data present at the time of analysis is shown (such a predictor is independent of which power pad we are examining since it does not encode any information about position).

Fig. 4 presents the estimated data-dependent correlations in juxtaposition to SPICE and Hamming distance results. In order to quantify the similarities, we also present the correlation between the three sets of matrices considering the four pads as unique data rather than using total power draw. As quantified in Table II, our estimated results are similar to those from SPICE. In particular, we draw attention to the similarities between row 1 and 3 in the Table, which predicts the effectiveness of one particular Hamming Distance-based DPA attack. Therefore, *our system can effectively predict data-dependent perturbations as well as predict vulnerabilities to power supply-based side-channel attacks.* We next present data that evaluates the accuracy of our framework.

TABLE II: SPICE and Estimation transient correlation

	P1	P2	P3	P4
Hamming Distance vs SPICE	.1289	.1161	.1067	.1351
Estimated vs SPICE	.6321	.6883	.6234	.5556
Hamming Distance vs Estimated	.1433	.1316	.1491	.1687

C. Power Supply Estimation

To demonstrate the temporal similarities between full-chip SPICE and our power supply estimation framework, we present data for 66 plain-text input for a fixed secret key. For a particular combination of encryption key and input plain-text, the current waveforms estimated (represented by dotted lines) at the four power pads are compared to full-chip SPICE waveforms (represented by solid lines). Shown in Fig. 5 are waveforms (magnified) for plain-text 1 during the initial rounds of encryption. As seen in the figure, the estimated waveforms match full-chip SPICE waveforms closely.

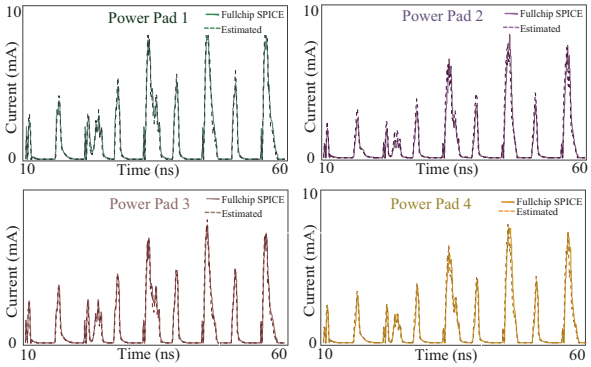

Fig. 5: Current Transients of plain-text 6 for all Power Pads

The metric used to evaluate the effectiveness of our prediction techniques is Mean Square Error (MSE). The MSE (%) compares the difference between the full-chip and estimated waveforms as defined in (1).

$$\frac{\sum_{n=0}^{N-1}(I_{FullChip}[n] - I_{Estimated}[n])^2}{\sum_{n=0}^{N-1}(I_{FullChip}[n])^2} \times 100\ \% \quad (1)$$

In (1), $I_{FullChip}$ is the full-chip SPICE transient current, $I_{Estimated}$ is the estimated transient current and N is the number of data points. We spent about 11 $\frac{1}{2}$ days of simulation time with our tool and achieved simulation of 66 paths. Maximum MSE % observed across all four power pads over 66 paths is 11 %. Fig. 6 shows the MSE % for all 66 plain-text inputs per power pad and the average is indicated by a horizontal line corresponding to each power pad.

For a demonstration of applicability of this technique to be utilized in vulnerability assessments of encryption systems, the DES current transients are subjected to DPA attacks described in the background section. The current peaks after the first round of DES for a correct key guess versus the incorrect guessed key at power pad 1 can be seen in Fig. 7.

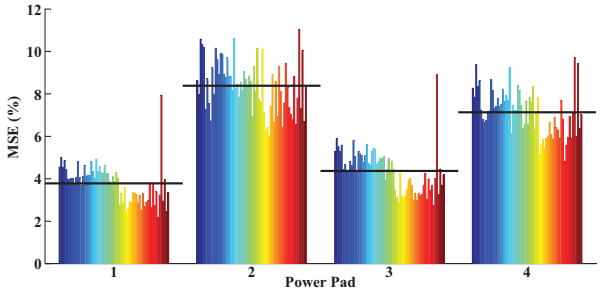

Fig. 6: Mean Square Error (%) for 66 paths for all Power Pads

Also noticeable are secondary peaks indicating the state of the intermediate values before moving over to the next round. Similar observations can be made for power pads 2, 3 and 4. Fig. 8 shows results for DPA attack based on 1000 guesses of a bit-sequence for the left input register to a given DES encryption clock cycle. actually be found to match one of the sequence of bits from the 32 left-registers.

Proven analysis has shown that if a bit sequence hypothesis matches a sequence found in operation, a correlation with power supply transients can be found. Using DPA analysis, 32 of the hypothesized sequences can actually be found to match one of the sequence of bits from the 32 left registers.

For Fig. 8, we inserted a set of correct hypothesized sequences, indexed 500^{th}-531^{st}, among a set of random hypothesis. With those that were inserted, we correctly predict higher correlations to the power supply as did SPICE. Peaks appearing prior and after the aforementioned interval indicate random partial sequence matching.

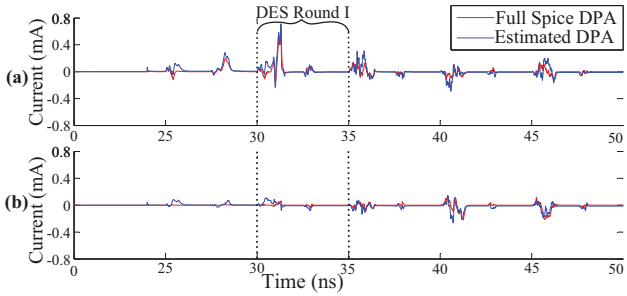

Fig. 7: DPA on a correct key guess (a) Vs incorrect key guess (b) for Power Pad 1

Registers in the layout with a closer proximity to a specific power pad tend to have higher current peaks in comparison to other registers placed at a distance from the power pad. As an example, guessed bit number 506 (which corresponds to the 7th s-box's bit after the first round of DES) at power pad 4 has a higher peak in relation with the other matched bits. This gives hints on localizing the position of core elements under attack in a Chip-Under-Test.

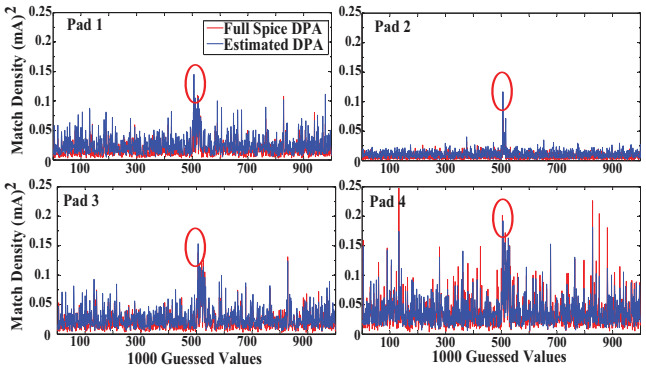

Fig. 8: DPA signal distribution over 1000 guessed values for all power pads

V. Conclusion

In this work we show results using our framework to accurately estimate power supply signatures of a DES system without running full-chip SPICE simulations. The estimated waveforms compared well against full-chip SPICE results. The predicted power supply signatures can be used to predict side-channel leakage, help in designing and evaluating counter-measures against attacks and provide golden signatures for trojan detection schemes. Furthermore, the current transients observed on each power pad also gives some hints on the proximity of core-elements being observed or attacked. The technique can be generically applied to other crypto-systems as well as industrial scale ICs by leveraging our partitioning scheme.

References

[1] Tehranipoor, M; Koushanfar, F; , *A Survey of Hardware Trojan Taxonomy and Detection*, Design & Test of Computers, IEEE , vol. 27, no. 1, pp. 10?25, Jan.-Feb. 2010.

[2] Agrawal, D.; Baktir, S.; Karakoyunlu, D.; Rohatgi, P.; Sunar, B.; , *Trojan Detection using IC Finger-printing*, Security and Privacy, 2007, SP 07, pp. 296-310, 20-23 May 2007.

[3] Xuehui Zhang; Tehranipoor, M.; *,RON: An on-chip ring oscillator network for hardware Trojan detection*, Design, Automation & Test in Europe Conference & Exhibition (DATE), 2011, pp.1-6, 14-18 March 2011.

[4] Davoodi, A.; Min Li; Tehranipoor, M. *A Sensor Assisted Self Authentication Framework for Hardware Trojan Detection, Design & Test*, IEEE, Vol. 30, Issue: 5, pp. 74-82, Oct. 2013.

[5] P. Kocher, J. Jaffe, B. Jun. *Differential power analysis*, CRYPTO '99, LNCS 1666, pp. 388-397,1999.

[6] J.-S. Coron, D. Naccache and P. Kocher,. *Statistics and secret leakage, ACM Transactions on Embedded Computing Systems*, Vol. 3, No. 3, pp. 492-508, August 2004,

[7] E. Brier, C. Clavier, F. Olivier, *Correlation Power Analysis with a Leakage Model*, proceedings of CHES 2004, LNCS, 3156 , pp. 16-29, 2004

[8] E. Peeters, Advanced DPA Theory and Practice - Towards the Security Limits of Secure Embedded Circuits, Chapter 2, pp. 11-13, 2013.

[9] National Bureau of Standards, "Data Encryption Standard", Federal Information Processing Standards Publication 46, January 1977.

[10] Abhishek Singh, Jim Plusquellic, Dhananjay Phatak and Chintan Patel, *Defect Simulation Methodology for iDDT Testing*, J.Electron Test., Vol. 22, pp. 255-272, June 2006.

[11] Rao, S.K.; Robucci, R.; Patel, C.; , *Framework for Estimation of Dynamic Power-Supply Noise and Path Delay*, Defect and Fault Tolerance in Symposium, pp. 272-277, Oct.2-4, 2013.

[12] http://www.cadence.com/products/cic/accelerated_parallel/pages/default.aspx, (Accessed October 2014)