Targeting Hardware Trojans in Mixed-Signal Circuits for Security

Invited Paper

Sabyasachi Deyati, Barry J. Muldrey, Abhijit Chatterjee

School of Electrical & Computer Engineering, Georgia Institute of Technology, Atlanta Georgia 30332

e-mail: sdeyati3@gatech.edu, bmuldrey3@gatech.edu, abhijit.chatterjee@ece.gatech.edu

Abstract—The proliferation of third-party silicon manufacturing has increased the vulnerability of integrated circuits to malicious insertion of hardware for the purpose of leaking secret information or even rendering the circuits useless while deployed in the field. A key goal is to detect the presence of such circuits *before* they are activated for subversive reasons. One way to achieve this is to detect the presence of parasitic loads on internal nodes of a victim circuit. However, such detection becomes difficult in the presence of normal process variations of the silicon manufacturing process itself. In this work, we show how high-resolution detection of parasitic loads on internal circuit nodes can be achieved using a combination of test stimulus design and design-for-Trojan detection techniques. We illustrate our ideas on digital as well as analog/mixed-signal circuits and point to directions for future research.

Keywords—*Hardware security; hardware intrusion detection; hardware Trojan detection; side channel analysis*

I. INTRODUCTION

In order to reduce the cost of integrated circuit (IC) manufacturing and to meet aggressive time to market goals, IC design and manufacturing has been split into several vertical domains. Each specialized vertical domain (e.g. design verification) is handled by a separate team, across different organizations and geographical areas. Use of third party IPs in System on Chip (SoC) designs has become ubiquitous. The diversity in the IC manufacturing chain across various organizations and use of third party IPs has opened up the possibility of insertion or modification of hardware in such ICs for malicious reasons. The major motives behind malicious tampering of ICs are as follows: i) leaking secret information from the IC, breaching security of the system where the IC is deployed and ii) malicious disruption of the functions performed by the IC while deployed in the field, either temporarily or with permanent effects.

There are numerous motives for incorporating malicious hardware Trojans (HTs) into ICs. These include: i) incorporating extra circuitry to alter the logic value of an internal circuit node in response to a predefined input trigger condition, ii) changing the dopant level of certain gates in the design to defeat optical inspection mechanisms while compromising system level security [1] and iii) accelerating wear out mechanisms in victim ICs (reliability Trojan) [2].

Though the majority of research has focused on digital HTs, insidious insertion of HTs in analog circuits is also possible, by: i) designing circuits such as oscillators to have more than one stable operating point (depending on initial starting conditions) and triggering an unacceptable operating mode using a stealthy hardware Trojan to cause system malfunction [3], ii) enabling a

circuit to leak secret information through side channel methods, exploiting design margins of analog/RF circuits [4].

In general, Trojans are designed in such a way that their activation is a rare event. It is almost impossible to trigger a Trojan and detect it during functional manufacture test. However, it is possible to detect passive Trojans based on *how they load internal circuits nodes* of a victim circuit, either in a capacitive or resistive manner. In this paper, we explore the various dimensions as well as limitations of such an approach and show how the envelope of passive Trojan detection can be pushed using intelligent test stimulus generation algorithms combined with supporting on-chip hardware infrastructure for Trojan detection.

II. PRIOR WORK

Hardware Trojan detection for digital circuits:

HT detection techniques for digital circuits are based on mainly: i) Trojan activation time reduction [5-7], ii) power measurement [8-11] and iii) path delay measurement [12-14]. As hardware Trojans are mostly dormant, any artificially directed logic activation boosting will not be very effective to trigger it as the nature, location and manifestation of hardware Trojans are generally unknown. Side channel signature analysis (such as power, supply current, path delays) requires the use of signatures from trusted ICs for comparison purposes to train learning algorithms to differentiate between “good” and “bad” signatures. In the presence of large process variations, such signature differentiation becomes extremely difficult as the worst case delay or leakage current values of a digital circuit can easily mask the additional delay or leakage current incurred by a hardware Trojan. In this paper we discuss a high resolution pulse propagation based Trojan detection technique for digital systems that is resilient to the effects of large process variations on the detectability of Trojans inserted into the design.

Hardware Trojan detection for analog/RF circuits:

It is possible to design analog circuits with multiple stable operating points that can be reached from different circuit initial conditions (such as initial voltage values of capacitors). Driving the circuit to an undesirable operating condition using an externally triggered hardware Trojan can result in overall circuit malfunction. In [3], the authors have shown an example of a Wien bridge oscillator where altering the initial charge stored in feedback capacitors determines the final equilibrium state of the oscillator. A Trojan can control the initial charge of a capacitor and thereby cause circuit malfunction. In this

context, identification and removal of Trojan states in a positive feedback analog circuit using Homotopy methods has been investigated in [15, 16].

In [4], the authors have shown a concurrent Trojan detection technique that continuously monitors the side channel fingerprint of a wireless transmitter using an on-chip neural network. The neural network classifier is trained during testing using side channel fingerprints of trusted ICs. Systematic variations of the transmitted signal within acceptable signal guardbands (very hard to detect during manufacture test) are detected and flagged to reveal secret leakage of information to third party wireless receivers.

III. TROJAN DETECTION IN DIGITAL SYSTEMS

As described in [17, 18], a Trojan payload is the part of the circuit affected by the Trojan (part of the circuit where logic value is changed due to the Trojan) and the act of causing it to have incorrect logic values is initiated by a Trojan trigger. Trojans can be combinational or sequential. An example of each is given in Fig. 1. Whether combinational or sequential, the inputs to the Trojan trigger circuit are obtained from nodes of the circuit corresponding to low signal transition probability. When an original circuit node is tapped as a Trojan trigger, it experiences extra load capacitance. This load capacitance can be detected even when the Trojan is not activated. However, with existing techniques, only large values of capacitive loads can be detected allowing many Trojans to remain undetected. Trojan detection can be made difficult by an attacker by using minimum sized inverters to tap internal circuit nodes for Trojan activation (See Fig. 2). The gate capacitance of a minimum sized inverter adds 0.2-1 fF of capacitance to the tapped signal node (45nm PTM [19]). Such small load capacitance values cannot be detected in the presence of silicon manufacturing process variations. To combat this problem, we propose a novel pulse-propagation technique that allows 880 attofarads of load capacitance to be detected in the presence of (10%) random process variations.

Fig. 1. Trojan models (a) combinational Trojan, (b) sequential Trojan

Fig. 2. (a) Tapping original circuit node for Trojan inputs (b) Corresponding equivalent circuit when Trojan is not activated

Fig. 3. Inverter chain

To propagate a pulse through a long chain of identical gates, at every gate, the transition swing should be rail to rail i.e. a positive pulse should reach VDD and a negative pulse should reach ground potential in order to perpetuate pulse propagation. If at any intermediate gate, rail to rail swing is not achieved at the gate output, the pulse will attenuate as it progresses through the chain (Fig. 3) and eventually “die”. If the input pulse rise/fall time (t_r/t_f) is smaller than the (10 to 90%) rise/fall (τ_n/τ_p) time of a gate then the minimum pulse width required to propagate the pulse through the chain of gates is $\tau_n + \tau_p$ [20]. SPICE simulation reveals that this is a greatly relaxed constraint for 45nm technology nodes and below. If $t_r(t_f) \ll \tau_p(\tau_n)$ then the constraint is even more relaxed. Fig. 4 shows two examples for: (a) a pulse width above the required minimum pulse width for the logic circuit concerned with the pulse propagating indefinitely through a long logic chain of inverters and (b) a pulse width that is less than the required minimum width, resulting in the pulse being “killed” as it progress through the chain of inverters. The presence (or absence) of a pulse propagating through a logic chain can be detected by sensing the supply current of the output logic gate [8] or by using a pulse detector at the output of the chain [21].

Fig. 4. Pulse propagation (a) Input pulse width less than the required minimum width, (b) input pulse width greater than the minimum required pulse width

If the pulse is propagating through a logic gate then it will draw current (I_{pulse}) from the power plane and if it is not propagating then it will draw only leakage current ($I_{leakage}$) from the power supply. In 45nm PTM [19], the ratio of $I_{pulse}/I_{leakage}$ is above 200. For a “dying” pulse, the peak pulse voltage versus supply current at corresponding inverters of the chain of inverters are shown in Fig. 4. Both peak current and rms current show almost linear relationship with peak voltage. Fig. 5 corroborates the idea of detecting a pulse by

sensing supply current of a logic gate. Integration of on-chip pulse generators and pulse detectors into existing scan chain infrastructure to enable Trojan detection has been demonstrated in [21] (voltage sensing based pulse detection) and [8] (current sensing based pulse detection).

Table 1 shows a comparison of the minimum capacitive load that can be detected using the proposed pulse propagation techniques vs. existing techniques that rely on path delay measurement. In all the experiments, process variations are introduced by changing the V_t values of the transistors in simulation. It is assumed that the threshold voltage values of each transistor in the design can change by $\pm 20\%$ with half of the variation (10%) originating from systematic V_t variations and the other half (10%) stemming from random variability effects (the results are easily generalized to larger V_t variability statistics). In order to compare the proposed Trojan detection scheme with path delay measurement based methods, we compute the *average* value of the Trojan capacitance C_{Trojan} that can be detected using the proposed method with the same for path delay based tests. To do this, we set the V_t values of all transistors to their nominal (average) values and find the value of the capacitance C_{Trojan} at a target gate along the respective pulse-sensitized path that causes the injected pulse to be “killed”.

Fig. 5. Peak pulse voltage vs current drawn from power supply

Table 1: Comparison of Proposed Pulse Propagation Test and Delay Test Detection Accuracy for NAND Chain

Logic Depth	Pulse Test	Delay Test	Improvement of Pulse Test over Delay Test
	Min Cap Detected	Min Cap Detected	
5	880aF	1fF	1.13X
10	880aF	2fF	2.27X
16	880aF	3fF	3.40X
22	880aF	4.5fF	5.11X
38	880aF	9fF	10.22X
70	880aF	15f	17.04X

Table 2: Minimum Detectable Capacitance Comparison between Proposed Pulse Propagation and Delay Method (**PP: Pulse Propagation Method; **DM: Delay Method)

Trojan Cite	Minimum Detectable Trojan Capacitance (fF)					
	C1		C2		C3	
Detection Mechanism	PP	DM	PP	DM	PP	DM
4 bit Adder	1.3	5.5	1.9	5.5	2.5	5.5
8 bit Adder	1.0	13	1.3	13	1.7	13
16 bit Adder	0.95	26	1.2	27	1.5	27

The experiment is then repeated by finding the value of the capacitance C^*_{Trojan} that causes the delay of the sensitized path to be violated for two cases: (a) where the delay guardband is

the cumulative worst-case delay across all the gates in the sensitized path and (b) where the delay guardband is the expected path delay for the specified systematic and random V_t variability statistics. The ratio $(C^*_{\text{Trojan}}/C_{\text{Trojan}})$ is used to describe the resolution of the proposed Trojan detection scheme in relation to traditional path delay measurement based Trojan detection methods. Results shown in Table 1 validate the claim that Trojan detection capability of the proposed scheme is *independent of the number of gates* in the pulse-sensitized path and also the improvement over delay test grows with the number of gates in the sensitized path. Similar experimental results for a variable-length ripple carry adder (for nodes C1, C2 and C3 of Fig. 6) are shown in Table 2.

IV. TROJAN DETECTION IN ANALOG SYSTEMS

In analog circuits, a key problem is that load capacitances due to injected hardware Trojans can be hard to detect given guardbands in manufacturing tests imposed by nominal manufacturing process variations. In addition, analog victim circuitry, tends to be structurally large in comparison to digital logic in order to avoid process variability effects and meet matching requirements. Therefore, internal nodes tend to have larger inherent cap-to-ground values than the digital systems we have looked at. The problem of Trojan detection is therefore posed as a test stimulus generation problem: design tests that are sensitive to capacitive short-circuit defects of *minimum magnitude* between internal circuit nodes and ground. This is accomplished by controlling the primary inputs, power supply voltage and loading of the circuit under test and observing only its primary outputs or other observable test nodes. Stimulus optimization experiments were performed on the LNA of Fig. 7 to minimize the values of detectable capacitances on all the nodes shown.

Fig. 7. Minimizing values of detectable capacitive loads on victim nodes of 2.4 GHz LNA

Optimization was performed using a genetic algorithm with a cost function that reflected the minimum detectable capacitive load that could be detected for a given candidate test stimulus.

While optimization ran concurrently for all victim nodes, it was assumed that only one node is impacted by the presence of a Trojan at a time. The optimized test stimulus is shown in Fig. 8 and allowed detection of capacitive loads on all the nodes of Fig. 7, except one, down to the tens of femtofarads. Note that in general, it is possible to modify the circuit structure to enable more efficient detection of Trojans using test generation.

Fig. 8. Optimized transient test stimulus (baseband)

V. CONCLUSIONS AND FUTURE WORK

The proposed techniques can be combined with design for Trojan detection algorithms that modify the circuit topology to allow efficient Trojan detection. Future work will focus on combinations of voltage and current measurement algorithms along with the use of outlier detection (learning) algorithms for effective Trojan detection in mixed-signal circuits and systems.

REFERENCES

- [1] G. T. Becker, F. Regazzoni, C. Paar, and W. P. Bursleson, "Stealthy Dopant-Level Hardware Trojans," in *Cryptographic Hardware and Embedded Systems - CHES 2013: 15th International Workshop, Santa Barbara, CA, USA, August 20-23, 2013. Proceedings*, G. Bertoni and J.-S. Coron, Eds., ed Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 197-214.
- [2] Y. Shiyankovskii, F. Wolff, A. Rajendran, C. Papachristou, D. Weyer, and W. Clay, "Process reliability based trojans through NBTI and HCI effects," in *Adaptive Hardware and Systems (AHS), 2010 NASA/ESA Conference on*, 2010, pp. 215-222.
- [3] Q. Wang, R. L. Geiger, and D. Chen, "Hardware Trojans embedded in the dynamic operation of analog and mixed-signal circuits," in *2015 National Aerospace and Electronics Conference (NAECON)*, 2015, pp. 155-158.
- [4] Y. Liu, G. Volanis, K. Huang, and Y. Makris, "Concurrent hardware Trojan detection in wireless cryptographic ICs," in *Test Conference (ITC), 2015 IEEE International*, 2015, pp. 1-8.
- [5] H. Salmani, M. Tehranipoor, and J. Plusquellic, "A Novel Technique for Improving Hardware Trojan Detection and Reducing Trojan Activation Time," *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, vol. 20, pp. 112-125, 2012.
- [6] M. Banga and M. S. Hsiao, "A Novel Sustained Vector Technique for the Detection of Hardware Trojans," in *VLSI Design, 2009 22nd International Conference on*, 2009, pp. 327-332.
- [7] H. Salmani, M. Tehranipoor, and J. Plusquellic, "New design strategy for improving hardware Trojan detection and reducing Trojan activation time," in *Hardware-Oriented Security and Trust, 2009. HOST '09. IEEE International Workshop on*, 2009, pp. 66-73.
- [8] B. J. M. Sabyasachi Deyati, Abhijit Chatterjee, "Trojan Detection in Digital Systems Using Current Sensing of Pulse Propagation in Logic Gates," in *ISQED, Santa Clara USA*, 2016.
- [9] J. Aarestad, D. Acharyya, R. Rad, and J. Plusquellic, "Detecting Trojans Through Leakage Current Analysis Using Multiple Supply Pad $\langle \text{formula} \text{ formulatype=inline} \rangle \langle \text{img src=images/tex/19434.gif alt=I}_{\text{DDQ}} \rangle \langle \text{formula} \rangle$ s," *Information Forensics and Security, IEEE Transactions on*, vol. 5, pp. 893-904, 2010.
- [10] C. Yuan, C. Chip-Hong, and C. Shoushun, "A Cluster-Based Distributed Active Current Sensing Circuit for Hardware Trojan Detection," *Information Forensics and Security, IEEE Transactions on*, vol. 9, pp. 2220-2231, 2014.
- [11] J. Aarestad, D. Acharyya, R. Rad, and J. Plusquellic, "Aarestad, J.; Acharyya, D.; Rad, R.; Plusquellic, J., "Detecting Trojans Through Leakage Current Analysis Using Multiple Supply Pad $\langle \text{formula} \text{ formulatype=inline} \rangle \langle \text{img src=images/tex/19434.gif alt=I}_{\text{DDQ}} \rangle \langle \text{formula} \rangle$ s," *Information Forensics and Security, IEEE Transactions on*, vol. 5, pp. 893-904, 2010.
- [12] B. Cha and S. K. Gupta, "Trojan detection via delay measurements: A new approach to select paths and vectors to maximize effectiveness and minimize cost," in *Design, Automation & Test in Europe Conference & Exhibition (DATE), 2013, 2013*, pp. 1265-1270.
- [13] J. Yier and Y. Makris, "Hardware Trojan detection using path delay fingerprint," in *Hardware-Oriented Security and Trust, 2008. HOST 2008. IEEE International Workshop on*, 2008, pp. 51-57.
- [14] P. Kumar and R. Srinivasan, "Detection of hardware Trojan in SEA using path delay," in *Electrical, Electronics and Computer Science (SCEECS), 2014 IEEE Students' Conference on*, 2014, pp. 1-6.
- [15] Y. T. Wang, D. Chen, and R. L. Geiger, "Practical methods for verifying removal of Trojan stable operating points," in *2013 IEEE International Symposium on Circuits and Systems (ISCAS2013)*, 2013, pp. 2658-2661.
- [16] Y. T. Wang, D. J. Chen, and R. L. Geiger, "Effectiveness of circuit-level continuation methods for Trojan State Elimination verification," in *2013 IEEE 56th International Midwest Symposium on Circuits and Systems (MWSCAS)*, 2013, pp. 1043-1046.
- [17] R. S. Chakraborty, S. Narasimhan, and S. Bhunia, "Hardware Trojan: Threats and emerging solutions," in *High Level Design Validation and Test Workshop, 2009. HLDVT 2009. IEEE International*, 2009, pp. 166-171.
- [18] S. Narasimhan, D. Dongdong, R. S. Chakraborty, S. Paul, F. Wolff, C. Papachristou, et al., "Multiple-parameter side-channel analysis: A non-invasive hardware Trojan detection approach," in *Hardware-Oriented Security and Trust (HOST), 2010 IEEE International Symposium on*, 2010, pp. 13-18.
- [19] T. Hong, Y. Li, S. B. Park, D. Mui, D. Lin, Z. A. Kaleq, et al., "QED: Quick Error Detection tests for effective post-silicon validation," in *Test Conference (ITC), 2010 IEEE International*, 2010, pp. 1-10.
- [20] L. W. Massengill and P. W. Tuinenga, "Single-Event Transient Pulse Propagation in Digital CMOS," *Nuclear Science, IEEE Transactions on*, vol. 55, pp. 2861-2871, 2008.
- [21] S. Deyati, B. J. Muldrey, A. Singh, and A. Chatterjee, "High Resolution Pulse Propagation Driven Trojan Detection in Digital Logic: Optimization Algorithms and Infrastructure," in *Test Symposium (ATS), 2014 IEEE 23rd Asian*, 2014, pp. 200-205.