ISSN (Print) : 0974-6846
Indian Journal of Science and Technology, *Vol 8(24), DOI: 10.17485/ijst/2015/v8i24/80873, September 2015* ISSN (Online) : 0974-5645

Power based Self-Referencing Scheme for Hardware Trojan Detection and Diagnosis

P. K. Maneesh* and M. Nirmala Devi

Amrita School of Engineering, Amrita Vishwa Vidyapeetham University, Coimbatore - 641112, Tamil Nadu, India

Abstract

Due to the expeditious growth and recent trends in Integrated Circuit (IC) industry, intrusion in terms of Hardware Trojans (HT) has become a major threat for IC security and reliability. Modern VLSI trends make the design vulnerable for possible HT insertion in various design and manufacturing phases. Growing design complexity in terms of number of gates, high testing cost and increased process variation makes HT detection and diagnosis more challenging. Logic testing has become ineffective against current threats due to their rare activation and stealthy nature. Side channel analysis has emerged as a promising technique but most of which relies on availability of HT free golden chips. In this paper, a golden chip free self-referencing scheme for sequential Trojan detection by comparing the IC's power signature at different time windows is proposed. The technique is evaluated on a 4-bit ALU and a set of ISCAS'85 benchmarks.

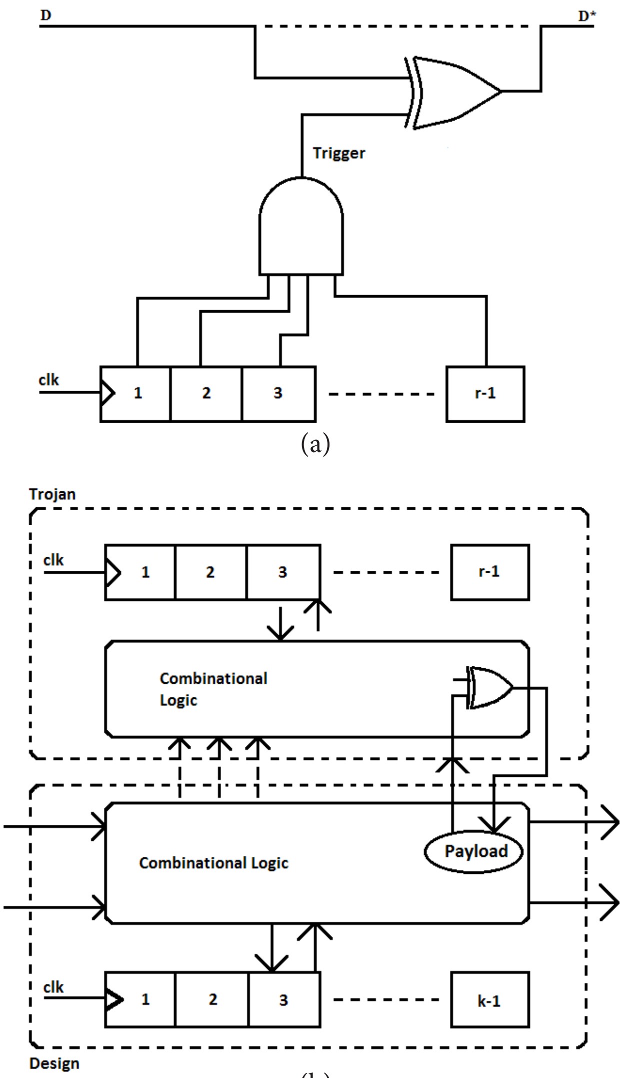
Keywords: Hardware Security, Hardware Trojan, Power Signature, Side Channel Analysis, Self Referencing

1. Introduction

Malicious alterations on integrated circuits that change the expected performance or functionality of the design are referred to as Hardware Trojans (HT). It poses serious concern about hardware reliability and trust especially in the field of military and security critical applications. The globalized modern trends in semiconductor industry makes the design vulnerable against Trojan insertion, most importantly from offshore fabrication facilities[1,2]. An HT may be designed either to alter the functionality or performance of a chip or to leak confidential information covertly to an adversary. Often HTs are efficiently designed to skip conventional post manufacturing structural and functional tests[1,3]. HT can either be of sequential or combinational type. Combinational Trojan may be triggered by the occurrence of specific logic states at rare internal points, while sequential type relies on a sequence of circuit events after extended period of operation. A simple and extended sequential Trojan models are shown in Figure 1(a) and (b) respectively. The former uses an r-bit counter which inverts the payload line when counter overflows, and the later uses a Finite State Machine (FSM) that is rarely triggered. The required internal state for Trojan activation is referred as trigger condition and the affected circuitry as payload.

Typically post manufacturing HT handling includes two steps. HT detection, which is identifying the presence of HT; and HT diagnosis, which is identifying the HT in terms of its physical parameters such as location, intrusion type, associated pins etc. Diagnosis is far more challenging especially when it lacks a golden reference for comparison. Sequential Trojan detection is more difficult in comparison with their combinational counter parts due to the difficulty of attaining rare event sequences that causes the Trojan activation. Several HT detection techniques have been proposed in the existing literatures[4]. Gate level characterization is the process of characterizing individual gates in a design, which has been used as a means for comparing golden chip and the one under authentication[5,6]. Parameter variation at Nano-metric feature sizes might mask the effects of Trojan on extracted GLC[7]. Side Channel analysis relies on the effect of Trojan circuitry on the side channel parameters of the IC such

**Author for correspondence*



(a)

(b)

Figure 1. (a) A simple sequential Trojan with r-bit counter (b) An extended sequential Trojan with independent FSM[11].

as power[8], delay[9], and current[10]. Existing Self-referencing schemes[11,12] fail when golden chip availability cannot be assured. New methods had been introduced that do not rely on the availability of golden reference. In[4] authors use gate level characterization upon segmentation to characterize overlapping circuit elements and checks for inconsistency for HT detection and diagnosis. Literature[13] extends this idea by using thermal conditioning to overcome the effect of linear dependency.

A complete solution to ease the process of both HT detection and diagnosis through a self-referencing scheme is presented. In this work, the method analyses the power signature of the design at different time windows and observe any inconsistency for HT detection. The procedure is extended by driving the circuit through smaller functional modules to diagnose the Trojan. The idea is that when a Trojan free design is exercised through same set of state transitions over different time windows their power signature should be consistent. Where as in a Trojan inserted circuitry, the signature varies over different time windows with same design state transitions due to independent Trojan FSM. Effectiveness of the scheme is verified on a 4-bit ALU and a set of ISCAS'85 benchmarks.

2. Trojan Detection Methodology

HT detection using self-referencing characterizes the circuitry in terms of its power signature at more than one time windows and checks for the inconsistency. The idea is based on the fact that most recent Trojans are of sequential type and they impose its effect on the IC at a later period of time, as a Time Bomb. So the Trojan might use the main clock or one derived from the main tree, but will not use system main reset. So the Trojan FSM runs completely independent of the main FSM. In this method, we exercised the design at two different time windows making sure the design traverse through same state transitions in both time. It is based on the fact that, if the IC is infected, the Trojan state transitions in those time windows will be different resulting in different power signatures. A Trojan-free IC only shows power variations caused by Process Variation and Noise. If the inconsistency in measured power values goes beyond a threshold, IC is identified as Untrusted.

2.1 Inconsisitency Metric

A new metric is proposed to analyze the effectiveness of the approach, called as Inconsistency Metric (IM), which is the ratio of maximum power variation exhibited by the infected circuitry with respect to that of a Trojan-free sample. The inconsistency matric can be written as follows:

$$IM = \frac{P_{\max,\,HT} - P_{\min,\,HT}}{P_{\max,\,HT-free} - P_{\min,\,HT-free}}$$

Where P_{max} and P_{min} are maximum and minimum measured power respectively for cases with or without HT as indicated by their subscript. So IM=4 means power variation with HT is four times that of the HT-free case.

Due to random process variation Trojan presence can only be confirmed if the IM value is above a specific threshold value, below which the effect of process variation might give a false-positive result.

3. Trojan Diagnosis Methodology

The proposed self-referencing based HT diagnosis scheme exploits the modern modular design architecture. By selecting proper input vectors, individual modules are activated at different time windows. Technique is based on the fact that when the infected section of the IC is activated, HT effect will be more reflected onto the measured parameter. Figure 2 shows a simple 4-bit ALU with addition, subtraction, multiplication and division blocks. A sequential Trojan is inserted in the addition module so that when the circuit is exercised to perform addition at different time windows, the power variation is more dominant when compared with similar measurements corresponding to other arithmetic modules which are Trojan free. This technique enables us to localize the Trojan so that suitable preventive measures can be taken, if possible.

4. Results and Analysis

Test circuits are setup, a simple 4-bit ALU and a set of ISCAS'85 benchmarks, and the detection and diagnosis approaches are validated. Sequential Trojans with FSM are inserted into all test designs and are analyzed. The circuit is forced through same input vectors at different time windows for HT detection.

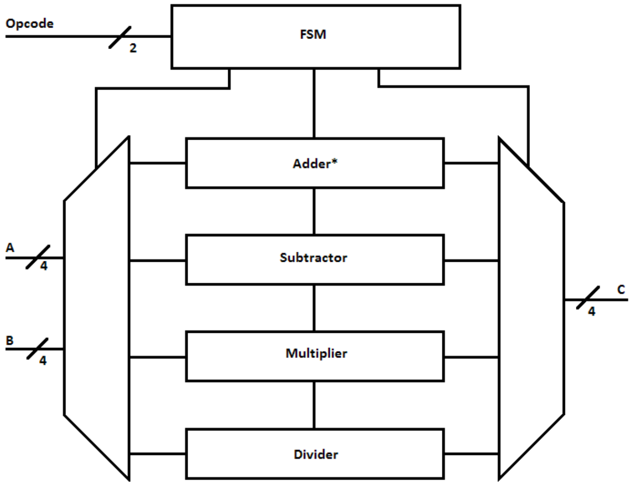


Figure 2. A 4-bit ALU with addition, subtraction, multiplication and division modules. A Trojan is embedded in the addition block.

4.1 4-Bit ALU

Trojan FSMs of different flip-flop sizes are designed and are inserted into a 4-bit ALU test circuit. Figure 3 shows the power readings of the same at four different time windows for the circuit with and without Trojan. The measurements corresponding to Trojan-free case stay consistent in all time slots, whereas inconsistency in the other graph confirms the presence of Trojan.

The design was exercised to activate individual modules separately to perform HT diagnosis. Same input combinations are applied to ensure similar switching of design elements in all time slots, with different Trojan FSM states. Measurements are taken by enabling individual modules and are repeated for four time windows. Result is shown in Figure 4. The inconsistency between readings corresponding to each module at all time slots is analyzed.

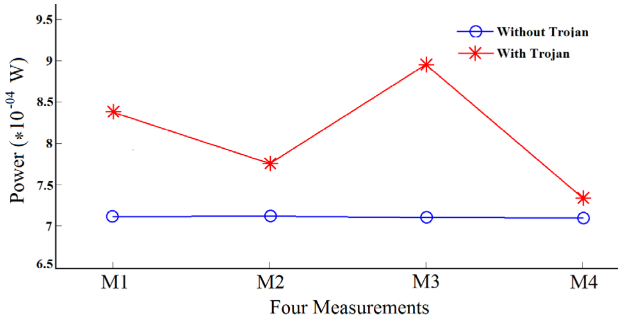


Figure 3. Power readings at four different time windows for a 4-bit ALU with and without sequential Trojan.

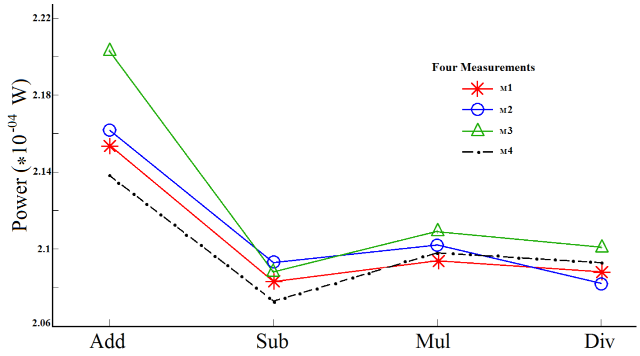


Figure 4. Four different power readings for a 4-bit ALU with a sequential Trojan. Four modules are individually activated.

Measurements of modules which are not infected preserve more consistency, whereas the infected addition block exhibits more variation enabling us to localize the Trojan. Readings corresponding to other modules show consistency because in their cases the Trojan is not directly activated.

4.2 ISCAS’85 Benchmarks

A set of ISCAS’85 benchmark circuits are setup embedding sequential Trojans, and are analyzed. Power measurements of c432 benchmark circuit is shown in Figure 5.

The circuit exhibits a power variation of 0.391x10-4 W with Trojan and 0.015x10-4 W without HT that gives an Inconsistency Metric value of 26 form (1). Complete results for all circuits under analysis are summarized in Table 1. When Process variation is simulated, as the design becomes larger the Trojan goes less visible. Trojan with more than three flip-flops are clearly detected in all the test circuits.

4.3 Inconsistency Analysis

Detection and diagnosis schemes are repeated with different Trojan sizes, by changing the number of flip-flops in the Trojan FSM. The Inconsistency Metric was generated so as to analyze the sensitivity of our approach. Inconsistency metric analysis for 4-bit ALU circuit with different Trojan sizes (in terms of number of flip-flops) is shown in Figure 6. With the simulated level of process variation Trojans with more than three flip-flops were found to be above the noise threshold.

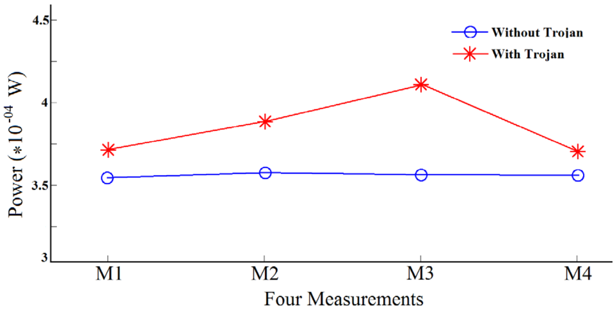


Figure 5. Sample power readings for c432 benchmark circuit.

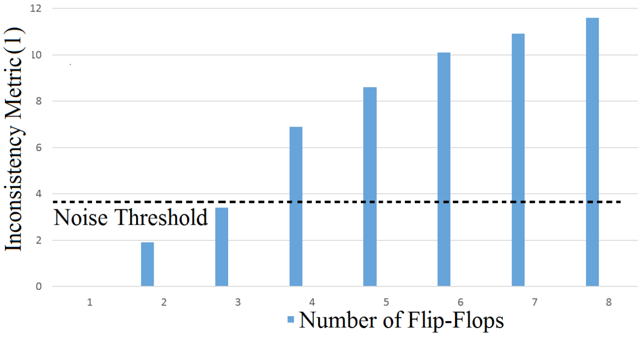


Figure 6. Inconsistency metric for varying Trojan size in terms of number of flip-flops (4-bit ALU).

Table 1. Inconsistency analysis for ISCAS’85 benchmarks

Benchmark	# Gates	Inconsistency (HT-free)	Inconsistency (HT)	Inconsistency Metric
c17	6	0.007	0.582	42.6
c432	160	0.015	0.391	26.0
c499	202	0.024	0.410	17.0
c880	383	0.031	0.421	13.5
C1908	880	0.051	0.568	11.1
C2670	1269	0.139	1.250	8.9
C3540	1669	0.173	1.231	7.1
C5315	2307	0.246	1.260	5.1
C6288	2416	0.273	1.370	5.0
C7552	3513	0.311	0.998	3.2

5. Conclusion

An effective and complete HT detection and diagnosis schemes are presented. Method is capable of detecting HTs with more than three flip-flops. Smaller Trojans with three or lesser number of flip-flops can be detected in post-manufacturing tests since the Trojan has limited number of states to switch through. The simulation results combined with the inconsistency analysis shows the effectiveness of proposed scheme for variable Trojan sizes without demanding a golden chip. Future work in this area will be to develop a complete golden chip free HT handling scheme by considering multiple parameters and devising more techniques to reduce the effect of process variation in analysis.

6. References

1. Tehranipoor M, Koushanfar F. A survey of hardware Trojan taxonomy and detection. IEEE Design and Test of Computers. 2010; 27:10–25.
2. Chakraborty RS, Narasimhan S, Bhunia S. Hardware Trojan: Threats and emerging solutions. IEEE International workshop High Level Validation and Test; 2009 Nov 4-6; San Francisco, CA; p. 166–71.
3. Wei S, Potkonjak M. The Undetectable and unprovable hardware Trojan horse. Procedeeings of 50th Annual DAC; 2013 May 29-Jun 7; Austin, TX, USA; p. 1–2.
4. Julien F, Frick F. Introduction to hardware Trojan detection methods. Proceedings of Design, Automation & Test in Europe Conference & Exhibition; EDA Consortium; 2015 Mar. p. 770–5.
5. Wei S, Meguerdichian S, Potkonjak M. Gate-level characterization: Foundations and hardware security applications. Proceedings of 47th ACM/IEEE DAC; 2010 Jun 13-18; Anaheim, CA, USA; p. 222–7.
6. Potkonjak M, Nahapetian A, Nelson M, Massey T. Hardware Trojan horse detection using gate–level characterization. Proceedings of DAC; 2009 Jul 26-31; San Francisco, CA; p. 688–93.
7. Borkar S, Karmik T, Narendra S, Tschanz J, Keshavarzi A, De V. Parameter variations and impact on circuits and micro architectures. Proceedings of DAC; 2003 Jun. p. 338–42.
8. Wei S, Potkonjak M. Self-consistency and consistency-based detection and diagnosis of malicious circuitry. IEEE Trans on VLSI Systems. 2014 Sep; 22:1845–53.
9. Jin Y, Makris Y. Hardware Trojan detection using path delay fingerprint. IEEE International Workshop Hardware-Oriented Security and Trust (HOST); 2008 Jun 9; Anaheim, CA; p. 51–7.
10. Agarwal D, Baktir S, Karakoyunlu D, Rohatgi P, Sunar B. Trojan detection using IC fingerprinting. IEEE Symposium on Security and Privacy; 2007 May 20-23; Berkeley, CA; p. 296–310.
11. Du D, Narasimhan S, Chakraborty RS, Bhunia S. Self-referencing: A scalable side-channel approach for hardware Trojan detection. CHES. 2010; p. 173–87.
12. Narasimhan S, Wang X, Du D, Chakraborty RS, Bhunia S. TeSR: A Robust temporal self-referencing approach for hardware Trojan detection. IEEE Symposium on Hardware-Oriented Security and Trust (HOST); 2011 Jun 5-6; San Diego CA; p. 71–4.
13. Wei S, Meguerdichian S, Potkonjak M. Malicious circuitry detection using thermal conditioning. IEEE Trans on Inf Forensics Security. 2011 Sep; 6(3):1136–45.