Journal of Scientific & Industrial Research
Vol. 78, July 2019, pp. 415-418

Proposing a Novel Method for Hardware Trojan Detection

A Bazzazi[1]* and M T Manzuri[2]

Department of Computer, Gorgan Branch, Islamic Azad University, Gorgan, Iran
Computer Engineering Department, University of Sharif, Tehran, Iran

Received 19 June 2018; revised 28 January 2019; accepted 29March 2019

Hardware Trojan refers to an unintended or undesirable alteration in many circuits which turned out to be a major challenge for the design and fabrication of integrated circuits for the semiconductor industries. Such a challenge has been addressed in numerous critical applications. Due to their diversity and process variation, Trojans vary in terms of detection and prevention methods. This paper proposes a novel technique for the Trojan detection involving the application of self-examining circuits. This method adequately resists against the PV and scalability and do not need to any golden ICs. Segmenting the circuit enables the technique to both detect and locate the Trojans with a quick possible time. The implementation of this method indicates a significant improvement in the detection rate.

Keywords: Hardware Security, Hardware Trojans, Prevention and Detection Methods, Self-examining Circuit

Introduction

Nowadays, there are very few companies carrying out, on their own, the entire stages of chip design. The presence of Hardware Trojan is an issue in this regard turning into a whole challenge. With regard to their diversity, Trojans have so far been classified into a variety of categories[1-4]. The Trojans are impossible to detect through traditional methods, which is why obsolete methods fail to detect new Trojans. The common detection methods are various [1, 5]. This paper focuses on a new detection method. This paper has been organized as follows; Section 2 examines the Motivation; Section 3 introduces the proposed method; section 4 describes the result, Section 5 offers the conclusions.

Motivation

There are five major difficulties concerning the employment of conventional methods including[4]:

- In most methods such as side-channel and logical testing, a reference chip or golden IC is required. It is not always possible to employ such chip due to inaccessibility to users.
- In modern detection methods, scalability has become an essential problem, which is further highlighted by the fact that sub-micron technology is increasingly developing. For instance, in the analysis of power at 45-nanometer technology, the circuit current is numerically so insignificant that it might be mistaken with noise.
- Process variation during chip fabrication maintains a fixed parameter between several chips with different values compared to their nominal values. For example, in 180nm technology, the leakage current and frequency both vary from one chip to another up to 20 and 30 times, respectively[4].
- The scalability has intensified the chip complexity. Hence, it will be time-consuming and expensive to employ reverse-engineering methods.
- In a rare condition, Trojan activates in the circuit, making it impossible to detect merely through inspecting the nodes or common testing measures, because they are inactive during the test.

Whether a die on a wafer or the entire chip, the HT detection methods are expected to be effective immediately after fabrication and run-time. Although there is a wide variety of detection methods proposed so far, only a few have actually served to tackle the Trojan occurrence and there has been little debate concerning Trojan addressing and masking. According to the five level cases above, it seems necessary that the method will be introduced which has not been these disadvantages. In the proposed method of this paper, in addition to these problems, there is a Trojan location possibility on chips.

The Proposed method

The main idea involves the use of test vectors. Every circuit has n inputs and 2^n test vectors. As the

*Author for Correspondence
Email: bazzazi@gorganiau.ac.ir

circuit arrives at the test stage, the test vector is applied so as to examine the circuit output. The circuit will be deemed healthy if the output is consistent with the desired value. The underlying idea in this method involves additional points inside a circuit to be tested other than the output points. Trojans generally penetrate into nodes where there is a low level of controllability and observability. Hence, the circuit middle points with the slightly lower level of controllability and observability should be added to the points being tested. Having applied the test vector, a matrix will be created entailing the response from tested middle points. Every line in this matrix belongs to a specific node. Moreover, the columns in each line refer to the node response to the test vector. For instance, a circuit with four inputs will contain a total of sixteen test vectors, i.e. there are sixteen matrix columns. The final matrix will be 4*16, should the implementation of this method involve five middle points. In fact, the number of middle points signifies the number of matrix lines. Having selected the middle points and having applied the test vectors, the desired matrix forms. The numerical analysis indicate that certain lines can be found between the matrix lines either completely identical or bearing a high percentage of resemblance.

The two node A and B are completely identical or they are completely complementary.

However, there are often alternative modes. A third mode where A and B are identical only partially. Wherever A and B are not identical, there are C and D completely identical. There is always a set of points found to be in a relationship. If such relationship cannot be found between two points, then more points should be taken into account. For each of the triple modes mentioned above, a circuit can be designed to demonstrate the relationship. For first mode, the *XOR* function can be employed. The *XOR* value for two points is always one. For second mode, the *XNOR* function can be employed. For third case, the *XOR* value for A, B, C and D is always one. In order to obtain precise results, the points should be selected in a way the requirements do not overlap. If the selection of four points is not enough for meeting the requirement, then more points can be taken into account.

When a node is affected by Trojans, its value shifts from one to zero and vice versa. Therefore, the function output created above will change. If one of the points is infected by Trojan, the *XOR* output value will be zero determining the presence of HT while the output value in the normal model is expected to be one.

In the proposed method, detection is done through the output of circuits designed based on node interconnections. These circuits can be termed “self-examiners”. This detection method can be more effective when the main circuit is divided into several sections, in each of which the desired points are selected and the corresponding matrix is formed. Afterward, the required relationships are extracted and the self-examiner circuit is designed separately for each section. Higher detection accuracy will demand more segments. The self-examiner circuit response will be valid only when the circuit has passed through the transition mode into stability. By employing this method, not only the health status of a segment can be assessed, but also any attacked segment and its location in the main circuit can be identified, which is technically called *localization*. In cases where a chip is repeatedly used in a circuit design, the method can pinpoint the segment(s) prone to HT threat, thus making it possible to design safer circuits for vulnerable segments in subsequent improvements. This method is associated with difficulties such as *power and area overhead* and *sensitive to attacks* explained in the following section

Sensitive to attacks

These circuits take up no specific location in the main circuit, i.e. the positions depend on how the circuit has been segmented. Hence, the main circuit has several distinct self-examiners which can sporadically spread over the main circuit. The identification of self-examiner circuits is, therefore, an uneasy task. By placing the output of the self-examiner circuit in a node with the high level of controllability and observability, the likelihood of invasion can still be curtailed.

Area and power overhead

They operate only when the main circuit has reached a stable status. Therefore, it will activate at a certain time to examine the circuit and report the result. Moreover, it will keep running the examination until new input arrives, i.e. the examiner circuit operates for a shorter period as compared to the total chip running time, which allows the power to be overlooked. The self-examiner circuit activates at certain times when a specific number of clock pulses have passed. For that purpose, a counter can be employed to obtain the number of clock pulses. Moreover, a flip-flop can be used to synchronize the

clock pulse with the self-examiner circuit. Since the sum of inserted circuits operate at specific times, they bear no power overhead, while the self-examiner circuit has a slight area overhead. The counter area is considerable in small-scale circuits. However, since the enlarged circuit dimensions leave the counter size the same as before, the effect of area overhead can be over looked. FPGA can be used to the implementation of the proposed method[5-8].

To start with, the Trojans are not parametric. Second, there is no limitation on how the Trojans are distributed. In summary, this section states advantages of the proposed method:

1) Golden IC isn't required. 2) It is a scalable method. 3) It is independent of process variation. 4) It is not time-consuming and expensive method. 5) It is active at all time, even in a rare condition.

Simulation Results

The proposed idea is implemented about s27. At first, the input test vector is applied on the circuit and the wave response is extracted from the shape of given nodes. In Figure 1, the s27 waveforms are shown. Afterwards, the responses are inserted in the matrix as described earlier. Because the circuit is small and there is no segmentation and gradation. Having evaluated the equations between the matrix arrays, four nodes including G3, G13, G17 and G9 are considered for establishing the examiner circuit. The equation concerning the four points as follows:

$F_{self\text{-}examiner} = G9 \oplus G13 \oplus (G3+G17)$ (1).

In a Trojan-free circuit, the equation is always 1. If the logical values of a node differs from those nodes, the function will be zero indicating there is a Trojan involved. At the next step, the proposed circuit is added to s27. Figure 2 displays the waveforms of altered circuit. At this stage, the HT needs to be applied to the circuit so as to confirm the validity of the proposed circuit. The HT is applied to the circuit through three modes:

Case 1: Trojan can alter the circuit function through modifying one or more gates embedded in the circuit. For instance, the OR gate in the main circuit was changed to NOR gate. The altered circuit waveforms of checkers have been illustrated in Figure 3a. The variations in the output waveform as compared to the waveform of a Trojan-free circuit indicate the involvement of a Hardware Trojan.

Case 2: In this mode, the invader inserts one or more paths in the circuit. The altered circuit wave form have been illustrated in Figures 3b.

Case 3: The invader inserts one or more extra gates in the circuit. For instance, a XOR gate is added to the circuit. The checkers waveform of the new circuit has been illustrated in Figures 3c.

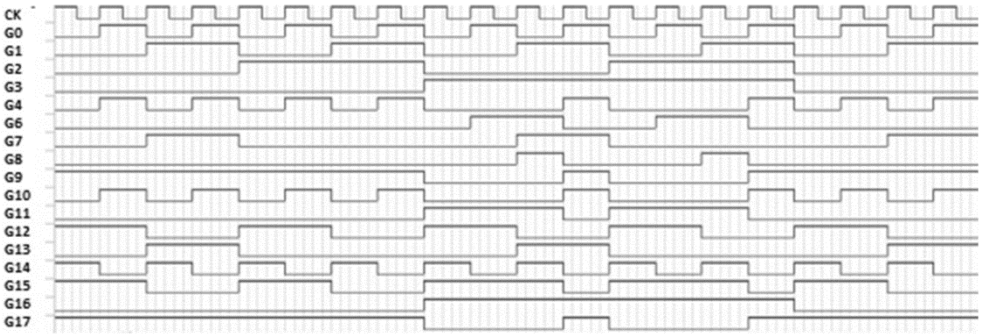

Fig. 1 — Waveforms of s27 nodes

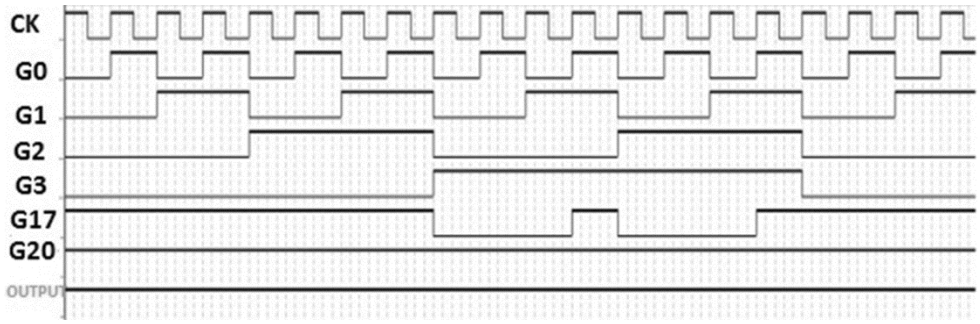

Fig. 2 — Waveforms of Self-Checker Circuit

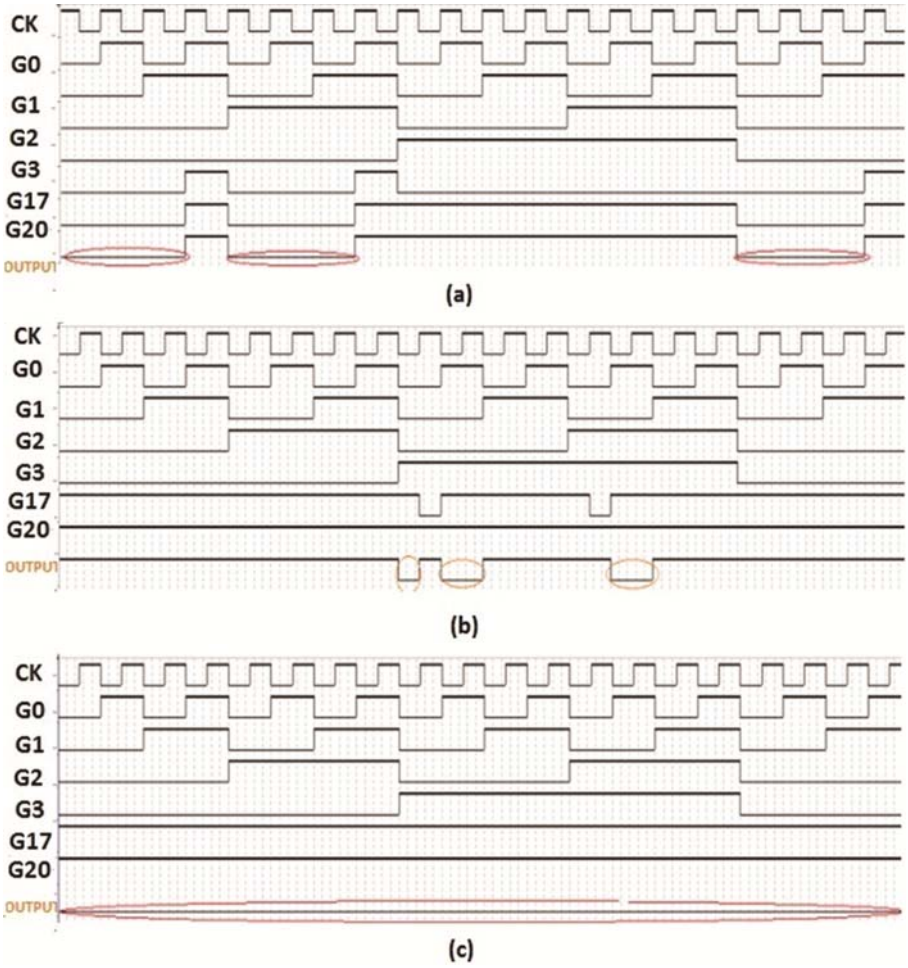

Fig. 3 — Hardware Trojan Detection by functional simulation waveforms (a) case 1, (b) case 2, (c) case 3

As can be seen in each mode, the proposed method is adequately capable of Trojan detection.

Conclusion

This paper presents a detection method based on test vector in the circuit. This method monitors certain nodes in the circuit. The main circuit is divided into several sections, in each of which an examiner circuit is designed considering the logical function between the corresponding nodes in that particular section. Any unintended alteration in a section can be traced by these circuits. This is regarded as a run-time and on-chip method, which has several key advantages leading to its popularity. The s27 quite successfully passed this method and it can be stated that such a method has proved successful in HT detection and localization.

References

1 Li H, Qiang L & Jiliang Zh, A survey of hardware Trojan threat and defense, *Integration*, **55**(2016) 426-437

2 Rishabh M, Agarwal K & Suman P, Hardware Trojans in IOT Devices: A Survey, *Int j adv res comput Sci*, **2** (2018) 68

3 Chaudhari F & Patel S, Survey: Trojan horse Detection Techniques in Network, *Int J Appl Math Comput Sci*, **9**(2017) 117-119

4 Bazzazi A, Manzuri MT & Hemmatyar AMA, Trojan counteraction in hardware: a survey and new taxonomy, *Indian J Sci Technol*, **9**(2016) 1-9

5 Sivanantham S & Tresa T, Built-in Self-Test Methodology for System-on-a-Chip Testing, *J Sci Ind Res*, **76**(2017) 149-153

6 Kuppuswamy C & Raghavendiran T, FPGA Implementation of Carrier Disposition PWM for Closed Loop Seven Level Diode Clamped Multilevel Inverter in Speed Control of Induction Motor, *J Sci Ind Res*, 77(2018) 504-509

7 Osornio-Rios RA, FPGA Lead-lag Compensator Design for Industrial Control Systems, *J Sci Ind Res*, **76**(2017) 733-736

8 Halim ISA, Kobayashi F, Watanabe M, Mashiko K & Yee OC, Small Area Implementation for Optically Reconfigurable Gate Array VLSI: FFT Case, *J Sci Ind Res*, **76**(2017) 697-700