INTEGRATION the VLSI journal 59 (2017) 233–242

Contents lists available at ScienceDirect

INTEGRATION, the VLSI journal

journal homepage: www.elsevier.com/locate/vlsi

Effective usage of redundancy to aid neutralization of hardware Trojans in Integrated Circuits

Nagendra Babu Gunti, Karthikeyan Lingasubramanian*

University of Alabama at Birmingham, BEC 256, 1720 2nd Ave S, Birmingham, AL 35294, USA

ARTICLE INFO

Keywords:
Hardware Trojans
Neutralization
TMR

ABSTRACT

Hardware Trojans are malicious alterations in Integrated Circuits (ICs) that leak confidential information or disable the entire IC. The detection of these Trojans is performed through logic or side channel based testing. Under sub-nm technologies the detection of Hardware Trojans will face more problems due to process variations. Hence, there is a need to devise countermeasures which do not depend completely on detection. In order to achieve such a countermeasure, we propose to neutralize the effect of Hardware Trojans through redundancy. In this work, we present a Triple Modular Redundancy (TMR) based methodology to neutralize Hardware Trojans. In order to address the inevitable overhead on area, TMR will be implemented only on select paths of the circuit. Using a probabilistic model of a given digital circuit, we have measured the effect of Trojan on different paths of the circuit and found that equally probable output paths are vulnerable to Trojan placement. Therefore for security we propose that TMR should be implemented on the paths that lead to equally probable primary outputs. We have also shown that the detection of Trojans placed on predictable paths can be achieved through logic based testing methods. In order for the adversary to beat the proposed redundancy model, the size of the Trojan has to be larger. We have shown that such implementation can be detected using side channel based testing.

1. Introduction

The daily life is aided by the automation, monitoring or computational power provided by Electronic systems. These systems which are based on digital circuits are expected to be highly dependable and trustworthy. The extensive usage of them in almost all critical sectors including finance, military, and industry has only increased the expectations of their security. To be a part of a versatile infrastructure these systems involve integration of multiple components as System On Chip (SoC) or an Embedded system. Such applications contain several hardware elements which are manufactured in global foundries. The emergence of the usage of globalized business model for production of electronic devices has resulted in hardware security and trust issues. Without trusted foundries, the systems they support cannot necessarily be expected to perform as specified and may even be susceptible to attack by a malicious adversary. A Hardware Trojan is a covert malicious, modification of an electronic circuit or design, which results in undesired behavior of an electronic device [1]. These alterations can provide a back door entry to a SoC or an Embedded system. An adversary can also utilize Trojans to leak sensitive information from a system or can even deny providing the service during the execution of critical applications.

A hardware Trojan can be classified into three main categories according to their physical, activation and action characteristics [2–4]. The physical characteristics category describes the various hardware manifestations of Trojans according to their shape and size; the activation characteristics describe the conditions which activate the Trojans, and action characteristics refer to the behavior of the Trojans. There are several techniques to detect Trojans but it is difficult to device a single Trojan detection technique that is applicable to all the varieties of Hardware Trojans. The detection of Trojans at the post-manufacturing test and validation phase, of the supply chain, is based on the physical and activation characteristics. Trojan detection techniques can be classified in two categories: 1) logic testing which focuses on activation characteristics and 2) side-channel which focuses on both physical and activation characteristics. The logic testing based detection depends on rare conditions to activate Trojans occurring at internal nodes of the circuit under test [5,6]. Whereas, side-channel based techniques involve observing the effect of Hardware Trojan on one or more physical parameter(s) such as transient current, leakage current or delay [7–9]. Typically, the adversary would design a Trojan to evade detection by ensuring that 1) the rare activation of Trojan goes undetected by logic testing and 2) the physical characteristics, like size,

* Corresponding author.
E-mail address: klinga@uab.edu (K. Lingasubramanian).

http://dx.doi.org/10.1016/j.vlsi.2017.06.002
Received 23 January 2017; Received in revised form 2 May 2017; Accepted 3 June 2017
Available online 17 June 2017

are small enough to evade side channel based testing. Moreover, the inherent process variations due to device scaling will also make the detection process unsuccessful due to the lack of a proper golden model. To handle such security vulnerability, neutralization of the threats without depending on its detection will be necessary.

In order to achieve such a countermeasure, we propose to use redundancy based strategy employing Triple Modular Redundancy (TMR) which can neutralize the effect of Trojans without the need of detection. Though, this methodology has significant area and power overhead, it will be a useful method for critical applications like Industrial Control Systems which are not constrained by them. Also, such application specific strategies for security are essential since it is implausible to devise a single countermeasure for every application [2,3]. We have also diagnosed the outputs of the digital circuit to find the ones which are vulnerable to Trojan placement and Hence TMR can be implemented on such paths to reduce the area and power overheads. We achieve this by using a probabilistic model to realize security aware TMR scheme to narrow the options for Trojan insertion and to aid the detection of Trojans at the post-manufacturing phase.

In digital system, TMR is a fault tolerant redundancy scheme which has traditionally been used to mask the unpredictable malfunction of a system due to aspects like process variations. To utilize TMR, digital systems have three copies of same subsystems which perform identical functions. The outputs of these copies are fed to a majority voter which transmits the final output by neutralizing the effect of Trojan as shown in Fig. 1(a). In a TMR induced system, the Trojan has to be placed in at least two copies in order to be effective as shown in Fig. 1(b). Such placement makes the Trojan size relatively big, therefore making it relatively easy to detect through side channel based techniques. Therefore, this technique forces the adversary to place the Trojan on non-TMR paths. Trojans placed in these non-TMR paths should be easier to detect through traditional logic testing methods. The objective of this work is to understand the behavior of the primary outputs in reference to Trojan detection sensitivity. This can be achieved through a probabilistic representation of a given digital circuit. Such model can enable us to perform efficient diagnostic studies to understand the characteristics of the outputs based on given inputs. Using the proposed probabilistic model we show that the Trojans placed in the path leading to predictable outputs have better detection sensitivity compared to those placed in the path leading to unpredictable outputs. Therefore, in our security scheme TMR is performed on the paths which lead to random unpredictable outputs. Based on the above discussion, the major contributions of the proposed work are as follows:

- Identification of circuit paths which are vulnerable to hardware Trojan insertion.
- Design of a security aware TMR scheme which selectively introduces redundancy in the system to handle both security and reliability.
- Effective neutralization of Trojans apart from better detection capabilities at the post-manufacturing phase.
- Increasing the limitations for effective Trojan insertions by the adversary.
- A presentation of probabilistic model to efficiently analyze Trojan based security concerns.

The proposed method is tested on multiple benchmark circuits of different sizes and complexities. Our probabilistic model is efficiently able to diagnose the primary outputs in all these circuits and identify the unpredictable random outputs and the predictable biased outputs. Our experiments unanimously show that the Trojan Detection Sensitivity (TDS) is higher for biased outputs and lower for the random outputs. Simulations using the proposed TMR placement scheme have clearly shown better detection capabilities of hardware Trojans.

The rest of the paper is organized in the following manner. Section 2 gives the related work which will discuss different methods used for Trojan detection. In Section 3, we present the proposed model of leveraging TMR for Trojan detection. In Section 4, we discuss the probabilistic network model which is used for analyzing the circuit outputs. In Section 5, we discuss the results of experiments and Section 6 draws some important conclusions.

2. Related work

Trojan detection techniques can be classified in two categories: 1) logic testing based and 2) side-channel based. the logic testing based detection depends on rare conditions to activate trojans occurring at internal nodes of the circuit under test [5,6]. in this approach, the outputs of circuit under test are compared with the outputs of golden circuit. jha and jha [10] analyzed the circuit functionality using its probabilistic signature. in this method random set of input patterns are applied to the circuit under test and the corresponding probability of logic 1 of primary output is compared with the design of the circuit. side-channel based techniques involve observing the effect of hardware trojan on one or more physical parameter(s) such as transient current, leakage current or delay [7–9,11–13,10]. these parameter(s) from circuit under test are compared with the pre-characterized value(s) of the parameter(s) obtained from golden circuit. both the methods of detecting hardware trojans have positive and negative aspects. the logic testing based approach has very large hardware trojan design space and an extremely large number of input-output combinations which are required for testing. this makes test generation computationally infeasible due to time constraints. and, side channel based approaches are affected by large process-induced parameter variations [14]. In [15], it is proposed to use design obfuscation to increase the complexity of reverse engineering for the attacker which helps in avoiding the insertion of trojans. this can be done adding additional states which increases the complexity of the attacker to insert an ingenious trojan in the design. however, this technique involves adding additional circuitry and thus reducing the performance of the whole design. hence, we propose a redundancy based scheme to neutralize the effect of trojans. tmr has traditionally been used for protecting digital logics from process variation induced errors or (transient errors) in space born applications [16,17]. also, to reduce the area and power overheads, we have diagnosed the output paths of the digital circuits and found that the equally probable output paths are vulnerable to trojans compared to the predictable outputs. hence, to optimize the placement of tmr resources to neutralize the effect of hardware trojans, tmr has to be implemented on equally probable output paths. thus, the proposed method makes it difficult to insert effective hardware trojans and forces the adversary to alter the placement or size of the trojans which leads to better detection using logic or side-channel testing.

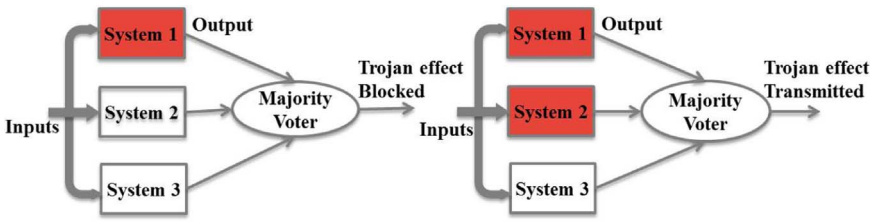

Fig. 1. (a) TMR with Trojan effect blocked (b) TMR with Trojan effect transmitted.