
University of Illinois at Chicago

Final Project Proposal

# AES attack through Hardware Trojan

ECE 594 - Hardware Security & Trust

Prof. Xiaolin Xu

Authors: Giorgio Bonomo Marco Montagna Francesca Pistilli

## Introduction

The current fabless trend adopted by IC design companies, forced to rely on worldwide third party to fabricate their design in order to reduce the increasing costs of fabrication of chips, has introduced several concerns about the possibility of malicious insertion or manipulation during the manufacturing process, for the purpose of reducing the reliability of a chip or steal information from it. This type of attacks takes the name of **Hardware Trojan** (HT), and represents one of the major threats in terms of hardware integrity, especially in the case of critical applications such as military, energetic and sensitive information protection.

Even though this topic has emerged only in recent years, several examples of HT insertion and detection techniques are available in literature, providing an interesting starting point for the development of ICs less susceptible to those typology of manipulations.
Following this mindset, the aim of the project here proposed is to give a proof of concept of two of the most promising HT insertion designs analyzed during the course, in order to simulate their behavior, test their realization on an FPGA and suggest possible improvements and detection techniques.

In particular, the target of those attacks will be the **AES** encryption system: considered a secure and trusted system, the study of how to reduce its key space or steal information from it is a challenging and stimulating topic.
The starting points of our design are based on the **A2: Analog Malicious Hardware** [4] and **Why You Should Care About Don't Cares: Exploiting Internal Don't Care Conditions for Hardware Trojans** [2] papers: the choice of these works as foundation of our work is related to the intrinsic stealthiness of the proposed solutions, together with a relatively new and interesting insertion methodology.

A more detailed explanation of the proposed project is postponed to the following section, with a more accurate description of the intended developments and declaration of intent.

## Proposal

In the following section, the planned design, simulation and verification steps are proposed.
In order to obtain a practical realization of the proposed HT, a Terasic DE-0 Board with Cyclone III FPGA has been made available.

### A2: Analog Trojan

Starting from the results reported in the cited paper, the aim of the proposed project is to replicate the insertion of an HT through an analog device (a capacitor) in an AES encryption system developed on FPGA. The capacitor is able to mimic the behavior of a counter, and is used as the trigger of the HT itself. When the switching frequency of the signal the capacitor is connected to overcomes a certain threshold, the *trigger* output becomes high and enables the *payload*. The explained functioning is clearly reported in *figure 1*.

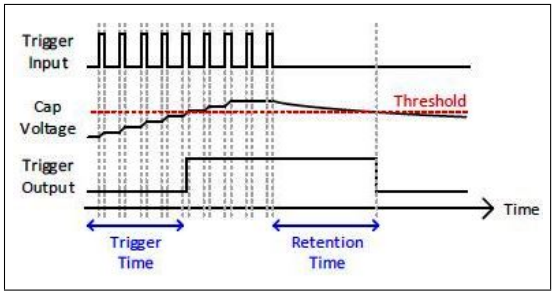


*Fig. 1: Functioning of the triggering system of the HT*

In our case, the payload is used to enable the reset/set of the Flip-Flops inserted to store the cipher key's bits: doing so, some bits of the encryption key will be fixed, reducing the key space and making possible to succeed with a brute force attack.

Before proceeding with the realization of the system, further study of which signal should be used as trigger input and about the size of the capacitor are going to be performed, in order to obtain a HT difficult to be detected during testing and verification phases, with a reasonable number of triggering cycles to be activated. Furthermore, the possibility to use a trigger signal given from an external source will be studied, in order to realize a system able to be triggered whenever the malicious attacker wants to.

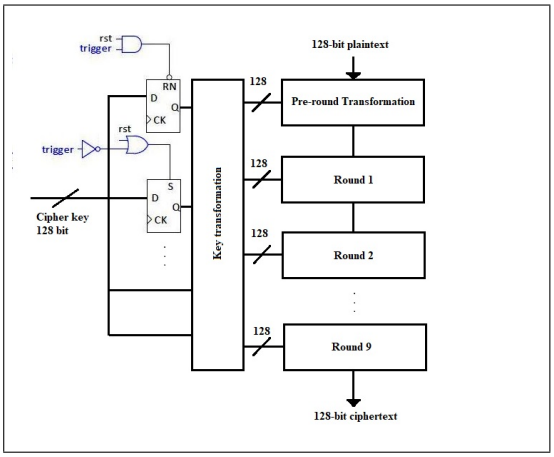


*Fig. 2: Basic idea of A2 HT insertion in the AES system*

The novelty introduced by our approach is that the Flip-Flops' output are not always fixed: in respect to the HT insertion in AES system performed in [1], the detection of such a HT insertion results difficult during Testing, unless the trigger input is stimulated. In all the other cases, the AES will be able to work without any visible variations caused by the insertion of the HT.

For the realization of the design, VHDL-AMS (an extension of VHDL that predicts the presence of Analog and Mixed signals) will be exploited; as far as it concerns simulation, one of the several possible tools available online will be used.

Finally, due to the lack of analog devices on a FPGA, the physical implementation of such a system presents some difficulties: during the project development, several possible solutions will be investigated (such as insertion of Flip-Flops to mimic the capacitor load or of external capacitors on the FPGA pins) in order to try to implement the designed system.

**Don't cares**

Starting from the interesting results obtained in the cited paper, the key idea of the second part of the project is based on a malicious exploitation of the lack of coverage of *satisfiability don't cares* (indicated in [2] as SDCs, representing never reached input conditions for a node): after the insertion of a state in the FSM used to leak information, normally hideous because never reached during normal behavior, the clock frequency of the realized system will be modified in order to violate the time requirements of the circuit and cause faults. Doing so, there is a probability that the malicious state will be reached and information leaked.

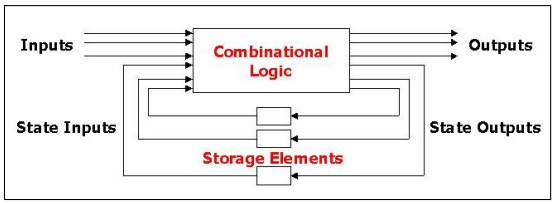


*Fig. 3: Basic view of an Finite State Machine (FSM)*

In particular, in order to prove the concept and make evident when the malicious state is reached, some bits of the encryption key will be leaked: it is however intended that, in order to make the presence of the Trojan less evident, information could be leaked in more stealthy ways, such as through code division multiple access (CDMA), as suggested in [3], or EM field emission.

In order to implement the system on the FPGA, both the possibilities to use an external stimulus as clock signal or derive it from the internal 50 MHz oscillator through successive clock divisions will be explored; this phase represents the less predictable of the whole project, due to the impossibility to strictly predict which clock frequency value will force the FSM to enter the malicious state.
Finally, the whole system will be implemented through VHDL description, and the malicious coverage of SDC will be forced through *others* statement.

It is well intended that the realized Trojan does not represent the best solution in terms of obfuscation and difficulty to be detected: however, the pursued objective is to underline the importance of a correct coverage of the don't care conditions during design and of testing the IC also for functioning conditions far from the nominal ones. Further modifications could exhibit the access to the malicious state only from particular configurations of the FSM (the least frequent for example), in order to make more challenging the detection of the inserted state.

## REFERENCES


[1] Georg Becker et al. “Stealthy dopant-level hardware Trojans: Extended version”. In: *Journal of Cryptographic Engineering* 4 (Apr. 2014), pp. 19–31. DOI: 10.1007/s13389-013-0068-0.

[2] Wei Hu et al. “Why you should care about don’t cares: Exploiting internal don’t care conditions for hardware Trojans”. In: *2017 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)* (2017), pp. 707–713.

[3] Lang Lin et al. “Trojan side-channels: lightweight hardware trojans through side-channel engineering”. In: *Cryptographic Hardware and Embedded Systems-CHES 2009*. Springer, 2009, pp. 382–395.

[4] Kaiyuan Yang et al. “A2: Analog Malicious Hardware”. In: *2016 IEEE Symposium on Security and Privacy (SP)* (2016), pp. 18–37.