Method Taking into Account Process Dispersion to Detect Hardware Trojan Horse by Side-Channel

X. Ngo, Z. Najm, S. Guilley, S. Bhasin, J.-L.Danger

PROOFS’14, Busan, South Korea

Presentation Outline

Introduction to HTH and its detection

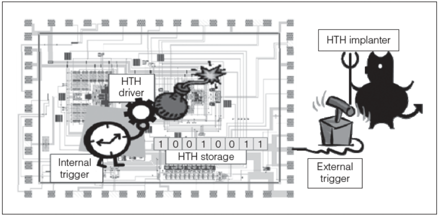
Proposed HTH Detection model

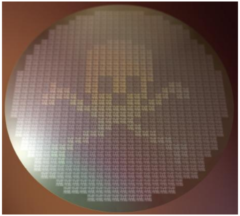
Setup and experimental results

Hardware Trojan Introduction

Hardware Trojan Horse (HTH) Definition

- Malicious modifications in Integrated Circuits (ICs).
- To extract a secret, alter the behaviour, ...
- HTH was born because of outsourcing design and fabrication process.

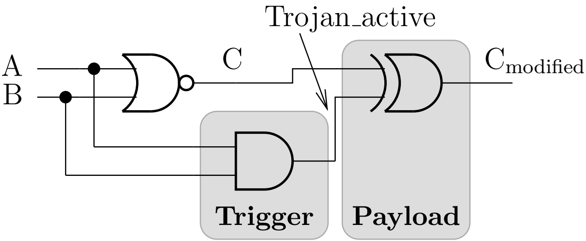




Hardware Trojan Structure

Any HTH is composed of two main components

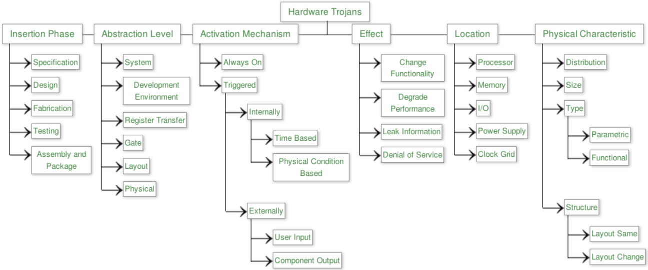
- **Trigger**: is the part of HTH used to activate the malicious activity.
- **Payload**: is the part of HTH used to realize / execute the malicious activity.



Hardware Trojan Taxonomy

- Classify all type of HTH [a]
- Help to develop suitable detection techniques for each HTH type

[a] Tehranipoor et al. [KRRT10]

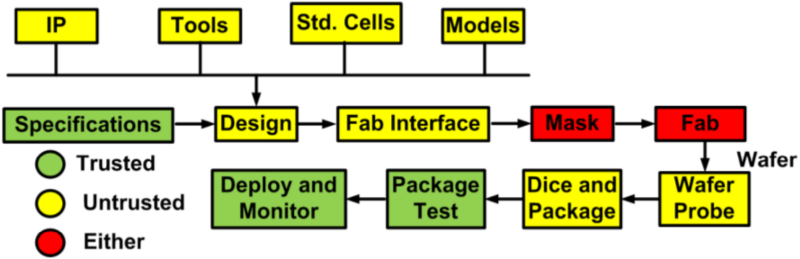


- Hardware Trojans
  - Insertion Phase
    - Specification
    - Design
    - Fabrication
    - Testing
    - Assembly and Package
  - Abstraction Level
    - System
    - Development Environment
    - Register Transfer
    - Gate
    - Layout
    - Physical
  - Activation Mechanism
    - Always On
    - Triggered
      - Internally
        - Time Based
        - Physical Condition Based
      - Externally
        - User Input
        - Component Output
  - Effect
    - Change Functionality
    - Degrade Performance
    - Leak Information
    - Denial of Service
  - Location
    - Processor
    - Memory
    - I/O
    - Power Supply
    - Clock Grid
  - Physical Characteristic
    - Distribution
    - Size
    - Type
      - Parametric
      - Functional
    - Structure
      - Layout Same
      - Layout Change

Trust in the design

HTH insertion in the fabrication flow of an ASIC. [a]

[a] Chakraborthy et al. [CNB09]



Hardware Trojan Detection

Classification of HTH Detection techniques

- **Destructive reverse engineering**: try to reconstruct netlist and layout of ICs.
- **Invasive methods**: try to (prophylactically) modify the design of IC to prevent the HTH or to assist another detection technique.
- **Non-Invasive methods**: are done by comparing the performance characteristics of an IC, possibly with a known good copy also known as the "golden circuit".

Invasive Methods

Examples

- To extend the state space
  - in two operating modes: Normal and Transparent mode.[a]
  - To consider either Q or QN of D flip-flops.[b]
- To insert dummy flip-flops into IC logic.[c]
- To add logic that will make the detection easier by using side-channel analysis.[d]

[a] Chakraborty et al. [CB09]
[b] Banga et al. [BH11]
[c] Salmani et al. [STP09]
[d] Lin et al. [LKG+09]

Non-Invasive Methods

Non-Invasive methods can be done either at **runtime** or during the **test phase**.

Non-invasive methods at runtime

- Use of OS features (Software approach).[a]
- Real-time security monitors: (**DEFENSE.**[b])

[a] Bloom et al. [BNS09]
[b] Abramovivi et al. [AB09]

Non-Invasive Methods

Non-invasive methods at test phase

Logic Testing:

- Compare the functionality of the design of the circuit with the implemented circuit.
- To test rare occurrences rather than correctness.[a]

Side Channel analysis Examples:

- To use power supply transient signal analysis.[b]
- To magnify the side-channel "sustained vector technique".[c]

[a] Chakraborthy et al [CWP$^+$09]
[b] Rad et al [RPT08]
[c] Banga et al [BH09]

Rationale

Side-Channel Detection Method Advantages

- Non-invasive method.
- Can detect almost HTH types, even untriggered.

Motivation

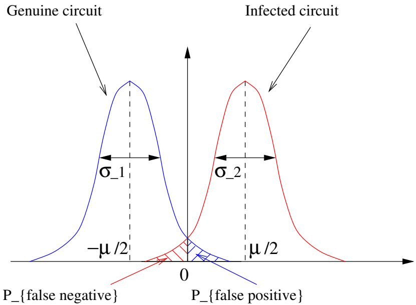
- Many Side-channel methods are based on power measurement or simulation results.
- Previous work did not take into account process variation and HTH placement.

Proposed detection Model

To take advantage of extra “load” due to HTH intrusion

- The HTH impact is an increase of current
- This effect comes from greater mean gate load,
- Which is mainly due to due to the complexity of the Trigger block
- Use of EM observation (spatial accuracy)
- $T^\circ C$ and V_{dd} should remain constant

Proposed detection Metrics



The metrics is a false negative and false positive probability, whose equation is:

$$P_{\text{false negative}} = P_{\text{false positive}} = \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp -\frac{(x-\frac{\mu}{2})^2}{2\sigma^2}\, \mathrm{d}x$$

Model flaws

The model is impacted by side effects

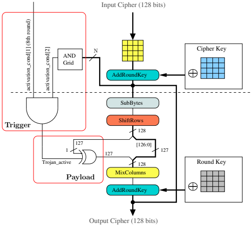
- $T^\circ C$ and V_{dd}
- Process variation
- HTH size and placement

⇒ we proposed to study theses potential flaws on the model, except the $T^\circ C$ and V_{dd} which are kept constant.

Setup description

HTH structure

- **Trigger part**: 8th computation round and N least significant bits (LSB) of 128 bits at the output of AddRoundKey are at “1”.
- **Payload part**:an XOR gate that will inject a fault in the inner eighth round when HT is activated.

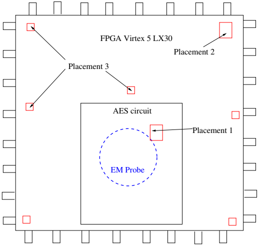


HTH with Different Sizes

- **Trojan 1**: HTH with the parameter **N = 32**, around **0.5 %** of the original circuit.
- **Trojan 2**: HTH with the parameter **N = 64**, around **1 %** of the original circuit.
- **Trojan 3**: HTH with the parameter **N = 128**, around **1.7 %** of the original circuit.

HTH with different Placement

- **Placement 1**: Trojan 3 placed **within** the boundary of AES crypto-processor.
- **Placement 2**: Trojan 3 placed **outside** the boundary of AES crypto-processor in a far-off corner of the FPGA.
- **Placement 3**: Trojan 3 placed outside the boundary of AES crypto-processor and **dispersed** over the FPGA.



Experimental Setup

Test platform setup

- 10 FPGA Virtex5LX30 for process variation evaluation.
- FF324 Virtex 5 board used to change the device under test.
- Frequency: 24 Mhz.
- EM measurement using Langer RFU-5-2 probe.
- Traces averaged 1000 times using Agilent 54853A.

HTH insertion

HTHs are inserted after the original circuit was placed and routed to minimize its impact on original circuit.

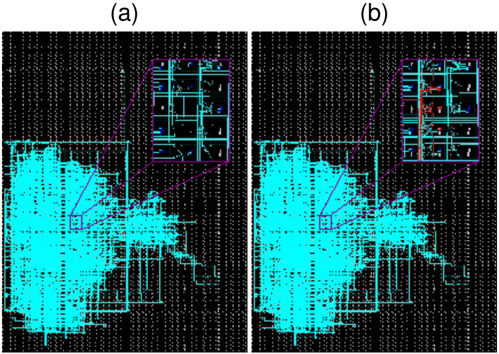
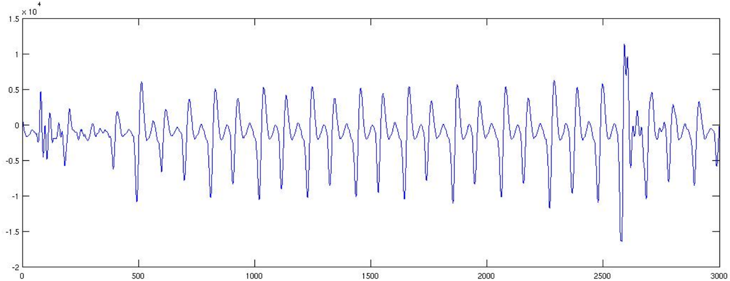


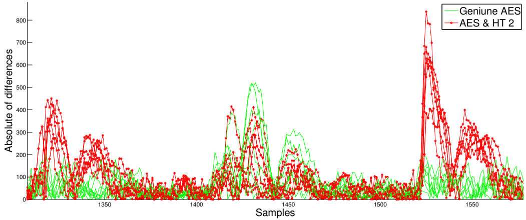
Figure : P/R for (a) AES 128 bit without HTH and (b) with HTH 1.7%

EM Leakage Trace



Impact of Process Variation on EM Measurement

- Calculate the golden mean trace over 10 FPGAs.
- In green: the difference between the golden circuit traces with the mean trace.
- In red: the difference between the HTH test circuit traces with the mean trace.



HTH Detection Using Sum of Absolute Differences

- Calculate the EM absolute differences.
- Calculate the sum of these differences.

	HTH 1 (0.5%)	HTH 2 (1%)	HTH 3 (1.7%)
1st Approach	43%	34%	9%

Table : False negative detection probability.

HTH Detection Using Threshold Technique

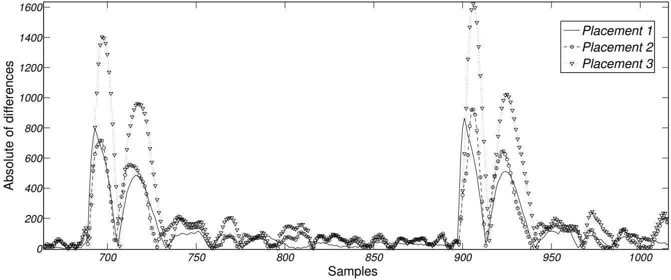
- Keep only the interesting points of EM differences.
- Re-calculate the sum of absolute differences of the interesting points.

	HT 1 (0.5%)	HT 2 (1%)	HT 3 (1.7%)
2nd approach	24%	0.017%	0.011%

Table : False negative detection probability with the Threshold technique.

Impact of HTH Placement

- The probe position affects directly to the result.
- The most distant HTH is more detectable (more buffers and lines) but has limited impact



Conclusion

Conclusion

- Proof of concept study for HTHs detection by EM measurement.
- Model based on the mean of EM activity
- HTH of different sizes: HTH greater than 1% can be detected with a false negative rate of 0.017%.
- Detection taking into account the process variation
- HTH placement has a little impact on HTH detection.

Références

[AB09] Miron Abramovici and Paul Bradley.
Integrated circuit security: new threats and solutions.
In Frederick T. Sheldon, Greg Peterson, Axel W. Krings, Robert K. Abercrombie, and Ali Mili, editors, *CSIIRW*, page 55. ACM, 2009.

[BH09] Mainak Banga and Michael S. Hsiao.
A Novel Sustained Vector Technique for the Detection of Hardware Trojans.
In *Proceedings of the 2009 22nd International Conference on VLSI Design*, VLSID '09, pages 327–332, Washington, DC, USA, 2009. IEEE Computer Society.

[BH11] M. Banga and M. S. Hsiao.
ODETTE : A Non-Scan Design-for-Test Methodology for Trojan Detection in ICs.
In *International Workshop on Hardware-Oriented Security and Trust (HOST), IEEE*, pages 18–23, 2011.

[BNS09] Gedare Bloom, Bhagirath Narahari, and Rahul Simha.
OS Support for Detecting Trojan Circuit Attacks.
In Mohammad Tehranipoor and Jim Plusquellic, editors, *HOST*, pages 100–103. IEEE Computer Society, 2009.

[CB09] R. S. Chakraborty and S. Bhunia.
Security against hardware trojan through a novel application of design obfuscation.
In *International Conference on Computer-Aided Design Digest of Technical Papers (ICCAD), IEEE*, pages 113–116, 2009.

[CNB09] Rajat Subhra Chakraborty, Seetharam Narasimhan, and Swarup Bhunia.
Hardware trojan: Threats and emerging solutions.
In *IEEE International High Level Design Validation and Test Workshop, HLDVT 2009, San Francisco, CA, USA, 4-6 November 2009*, pages 166–171. IEEE, 2009.

[CWP+09] R. S. Chakraborty, F. G. Wolff, S. Paul, C. A. Papachristou, and S. Bhunia.
MERO: A Statistical Approach for Hardware Trojan Detection.
In *Workshop on Cryptographic Hardware and Embedded Systems (CHES), LNCS*, volume 5747, pages 396–410, 2009.

[KRRT10] Ramesh Karri, Jeyavijayan Rajendran, Kurt Rosenfeld, and Mohammad Tehranipoor.
Trustworthy Hardware: Identifying and Classifying Hardware Trojans.
IEEE Computer, 43(10):39–46, 2010.

[LKG+09] Lang Lin, Markus Kasper, Tim Güneysu, Christof Paar, and Wayne Burleson.
Trojan Side-Channels: Lightweight Hardware Trojans through Side-Channel Engineering.
In *CHES*, volume 5747 of *Lecture Notes in Computer Science*, pages 382–395. Springer, September 6–9 2009.
Lausanne, Switzerland.

[RPT08] R. Rad, J. Plusquellic, and M. Tehranipoor.
Sensitivity analysis to hardware trojans using power supply transient signals.
In *International Workshop on Hardware-Oriented Security and Trust (HOST), IEEE*, pages 3–7, 2008.

[STP09] Hassan Salmani, Mohammad Tehranipoor, and Jim 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*, pages 66–73, 2009.