Hardware Trojan Horses in Cryptographic IP Cores

Shivam Bhasin, Jean-Luc Danger, Sylvain Guilley, Xuan Thuy Ngo and Laurent Sauvage

TELECOM-ParisTech and Secure-IC S.A.S.

August 21, 2013 09:55 – 10:20
Presentation Outline

1. Introduction
2. Hardware Trojan Detection State-of-the-art
3. Layout-GDS II Comparison Technique
4. Results
5. Conclusion
Hardware Trojan Introduction

Hardware Trojan (HT) Definition

- Malicious modifications in Integrated Circuits (ICs).
- Realize malicious functions (Leakage of sensible information, alteration of IC behaviours, etc.).
- HT was born because of outsourcing design and fabrication process.
Any HT 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.

![Diagram of Hardware Trojan Structure](image)
Hardware Trojan Taxonomy

- Classify all type of HT.
- Help to develop suitable detection techniques for each HT type.
Hardware Trojan Detection

Classification of HT 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 with a known good copy also known as the “golden circuit”.

Invasive Methods

- Chakraborty et al. propose a design with two operating modes (Normal and Transparent mode).
- Salmani et al. propose a procedure to insert dummy flip-flops into IC logic.
- Banga et al. propose using QN of D flip-flops.
- Other researchers also suggest logic additions that will make it easier to detect a HTH utilising side-channel analysis.
Non-Invasive Methods (1)

Non-Invasive methods can be done either at runtime or in the testing phase.

Non-invasive methods on runtime
- Bloom et al. detail a HTH detection approach that uses both hardware and software to detect HTs.
- Abramovici et al. propose real-time security monitors (DEFENSE).
Non-Invasive Methods (2)

Non-invasive methods on testing phase

Logic Testing:
- Compare the functionality of the design of the circuit with the implemented circuit.
- Chakraborty et al. suggest to test rare occurrences on an IC rather than testing for correctness.

Side Channel analysis use one or more side-channel parameters to obtain Fingerprint of ICs. We can cite:
- Rad et al. propose using power supply transient signal analysis.
- Banga and Hsiao propose the "sustained vector technique" that is able to magnify the side-channel.
Presentation Outline

1. Introduction
2. Hardware Trojan Detection State-of-the-art
3. Layout-GDS II Comparison Technique
4. Results
5. Conclusion
Introduction

Scenario

- Scenario: Attacker is founder.
- Study Hardware Trojan insertion in GDSII level.
- Impact of the insertion on the IC layout.
- The possibility to detect Hardware Trojan visually?
Case Study-AES 128 bits
Experiment setup

- Vary core utilization rate of AES (50% → 99%).
- Vary Hardware Trojan size (1 AND gate → 128 AND gates).
- Software used: Cadence / Encounter.

**Figure:** Hardware Trojan with $N - 1$ AND gates
Presentation Outline

1. Introduction
2. Hardware Trojan Detection State-of-the-art
3. Layout-GDS II Comparison Technique
4. Results
5. Conclusion
AES layouts for 6th metal layer

Figure: 6th metal layer AES layouts (1200 $\mu$m $\times$ 1200 $\mu$m) with 50% core utilization rate for (a) Original AES, (b) AES with 1 AND gate HTH, (c) AES with 128 AND gate HTH
Cross correlation between original AES layout and affected AES layout

<table>
<thead>
<tr>
<th>Core utilization rate</th>
<th>Hardware Trojan size (Nb of AND gates)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>50%</td>
<td>0.9991</td>
</tr>
<tr>
<td>60%</td>
<td>0.9987</td>
</tr>
<tr>
<td>70%</td>
<td>0.9989</td>
</tr>
<tr>
<td>80%</td>
<td>0.9999</td>
</tr>
<tr>
<td>90%</td>
<td>0.9988</td>
</tr>
<tr>
<td>95%</td>
<td>0.9997</td>
</tr>
<tr>
<td>99%</td>
<td>0.9917</td>
</tr>
</tbody>
</table>

- In black: ECO route works
- In red: total rerouting
- NC: routing impossible (not placement)
AES layout for all metal layer

Figure: AES layouts (1200 µm × 1200 µm) with 50% core utilization rate for (a) Original AES, (b) AES with 1 AND gate HTH, (c) AES with 128 AND gate HTH
Cross correlation between original AES layout and affected AES layout

<table>
<thead>
<tr>
<th>Core utilization rate</th>
<th>Hardware Trojan size (Nb of AND gates)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>50%</td>
<td>0.9996</td>
</tr>
<tr>
<td>60%</td>
<td>0.9993</td>
</tr>
<tr>
<td>70%</td>
<td>0.9991</td>
</tr>
<tr>
<td>80%</td>
<td>0.9997</td>
</tr>
<tr>
<td>90%</td>
<td>0.9990</td>
</tr>
<tr>
<td>95%</td>
<td>0.9995</td>
</tr>
<tr>
<td>99%</td>
<td>0.994</td>
</tr>
</tbody>
</table>
Pixelwise difference of AES layouts

Figure: Pixelwise difference of AES layouts with 50% core utilization rate for Original layout and Infected Layout with (a) 1 AND gate, (b) 128 AND gate.
Grid-correlation between layouts

Grid-correlation definition

- Improve Cross correlation coefficients.
- Split images on different pieces
- Compare these pieces of one with these corresponding pieces of others.
- Reverse Cross correlation coefficients are computed to visual improvement.
Grid-correlation examples

Grid-correlation example for CUR of 50% (a, b) and 95% (c, d)

(a) Trojan with 1 AND gate  (b) Trojan with 128 AND gates

(c) Trojan with 1 AND gate  (d) Trojan with 128 AND gates
Experiment on SECMAT circuit

Figure: Cross correlation based comparison between trojaned (left hand side) / genuine (right hand side) GDSII and an actual picture, a microscope image of an AES area where the inserted HTH shows up (center).

\[ |NCC| = 1.56\% > |NCC| = 0.67\% \]
Presentation Outline

1. Introduction
2. Hardware Trojan Detection State-of-the-art
3. Layout-GDS II Comparison Technique
4. Results
5. Conclusion
Conclusion

Hardware Trojan threat
- Become a serious threat in military, financial fields.
- For now, there is NO technique which can detect all type of Hardware Trojan.

GDSII comparison technique conclusion
- Can detect Hardware Trojan at layout level.
- No need golden model for detection.
- With a CUR superior than 80%, designer can prevent Hardware Trojan insertion.
Future Works

- Improve this technique with minutæ analyses.
- Insertion of Hardware Trojan in processors.

We wish to collaborate with https://www.trust-hub.org/
Acknowledgements: HOMERE (2011-2013)

- Collaborative project sponsored by the French government:
  - Partners from industry are: Cassidian, Gemalto, Secure-IC,
  - Partners from academia are: ENSMSE, Telecom-ParisTech, CEA LETI, Université de Montpellier,
Hardware Trojan Horses in Cryptographic IP Cores

Shivam Bhasin, Jean-Luc Danger, Sylvain Guilley, Xuan Thuy Ngo and Laurent Sauvage

TELECOM-ParisTech and Secure-IC S.A.S.

August 21, 2013 09:55 – 10:20