Self-Referencing: A Scalable Side-Channel Approach for Hardware Trojan Detection

Dongdong Du, Seetharam Narasimhan*,

Rajat Subhra Chakraborty and Swarup Bhunia

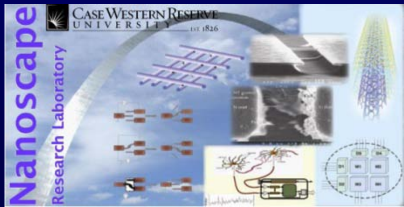
Department of Electrical Engineering and Computer Science

Case Western Reserve University

Cleveland, Ohio, USA

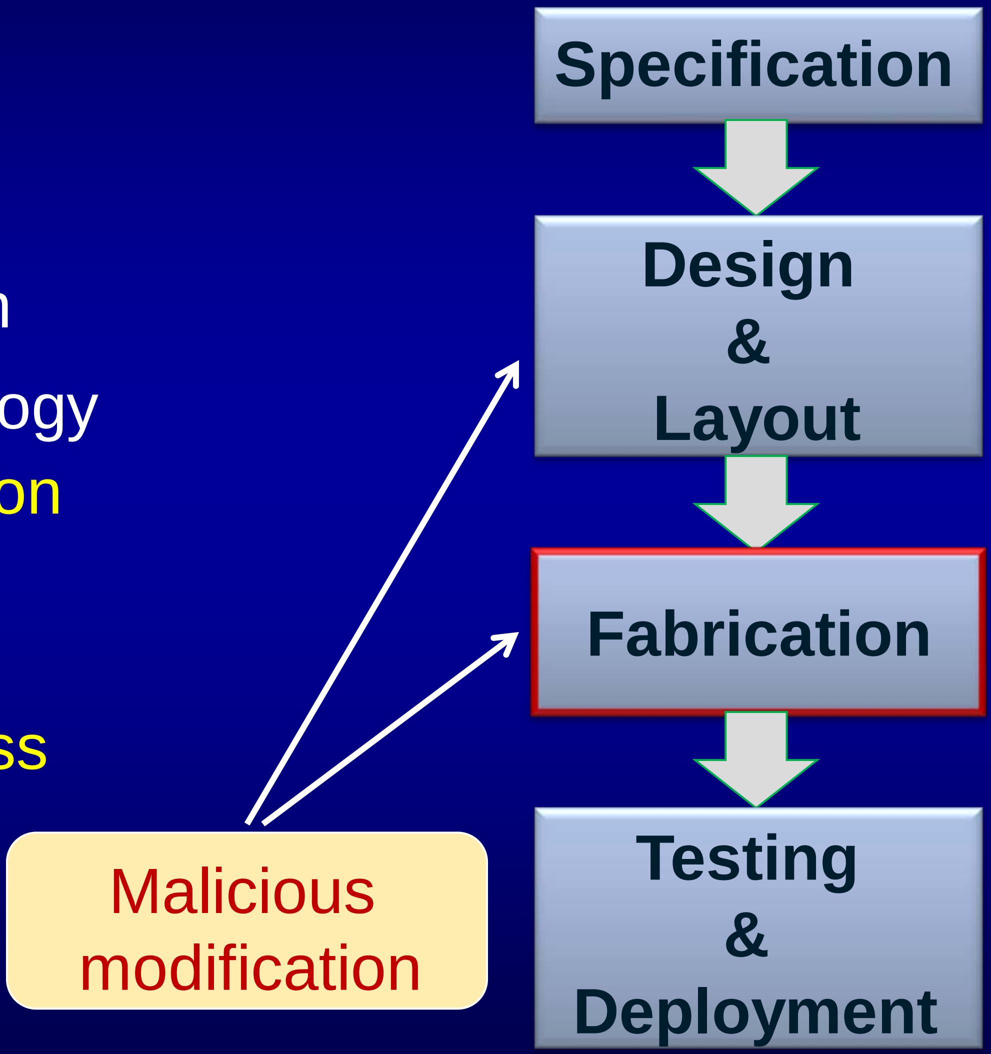
CASE WESTERN RESERVE UNIVERSITY
CASE SCHOOL OF ENGINEERING

19th Aug, 2010



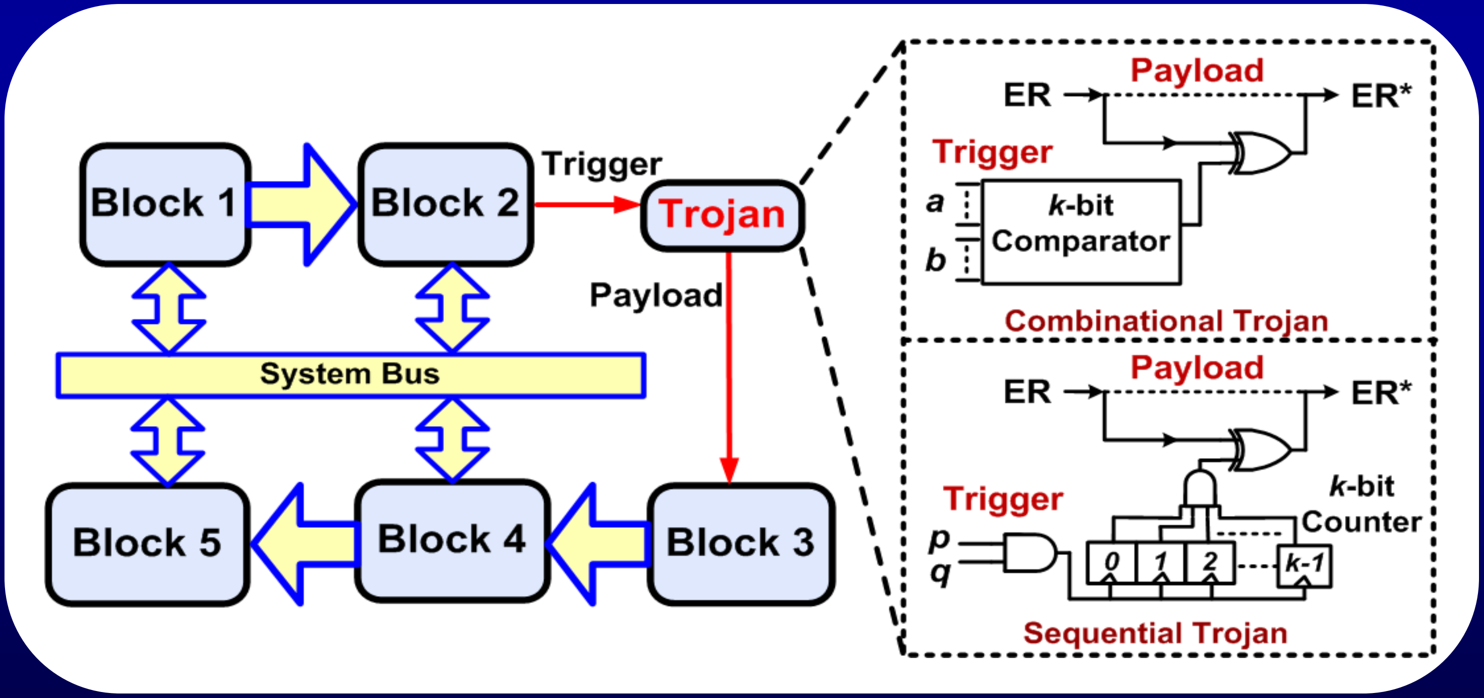
Outline

- Introduction
  - Hardware Trojans
  - Detection Methods
- Background and Motivation
- Self-Referencing Methodology
  - Functional Decomposition
  - Test Vector Generation
  - Side-Channel Analysis
  - Decision-making Process
- Results
- Conclusion



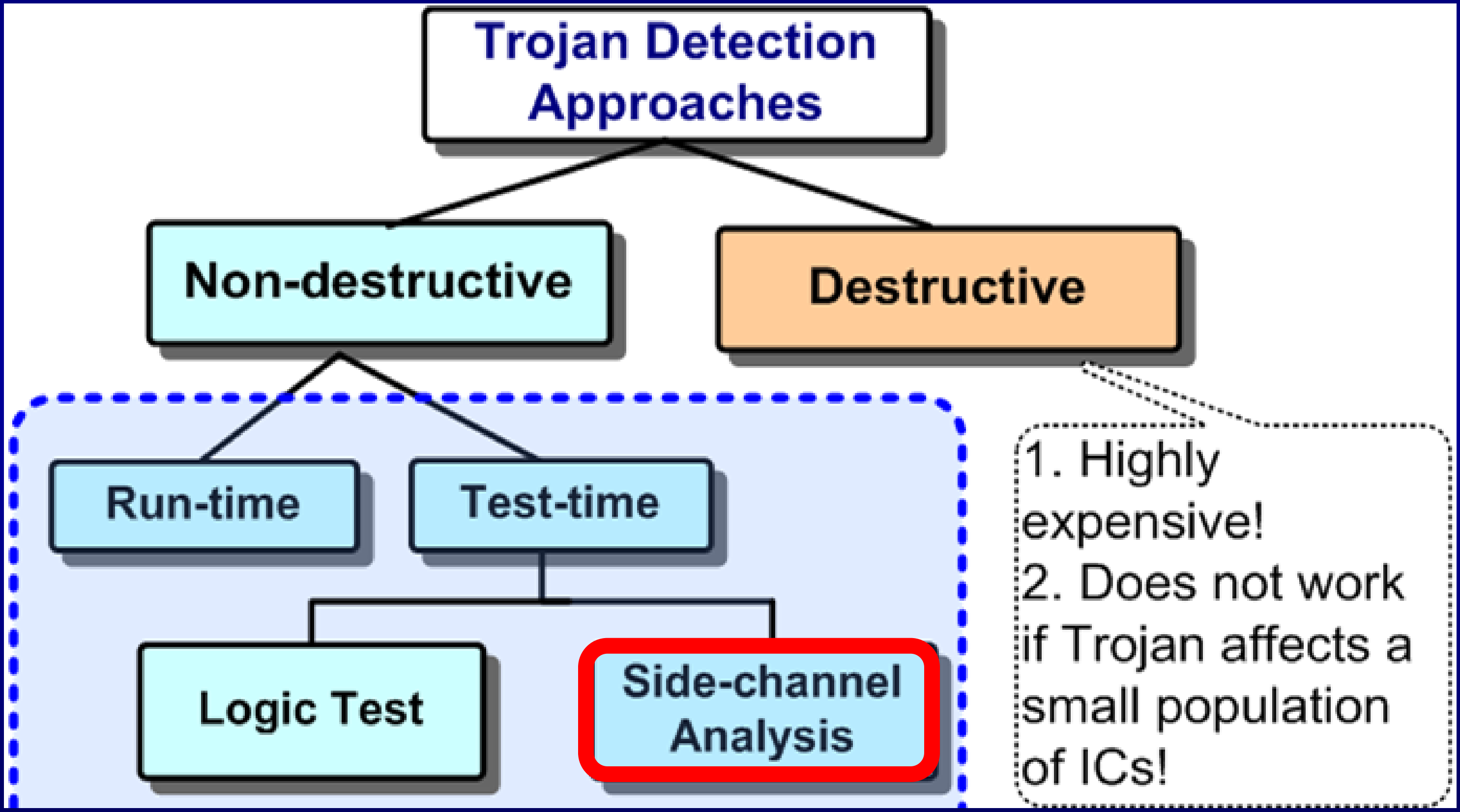
Introduction

- **Hardware Trojan**
  - Global outsourcing of fabrication of ICs raises potential for malicious modification which can cause malfunction in field or cause leakage of secret information (C. Paar *et al*, CHES'09).



Introduction

- **Trojan Detection Approaches**



Better for small combinational Trojans
(R.S. Chakraborty *et al*, CHES '09)

Better for large sequential Trojans

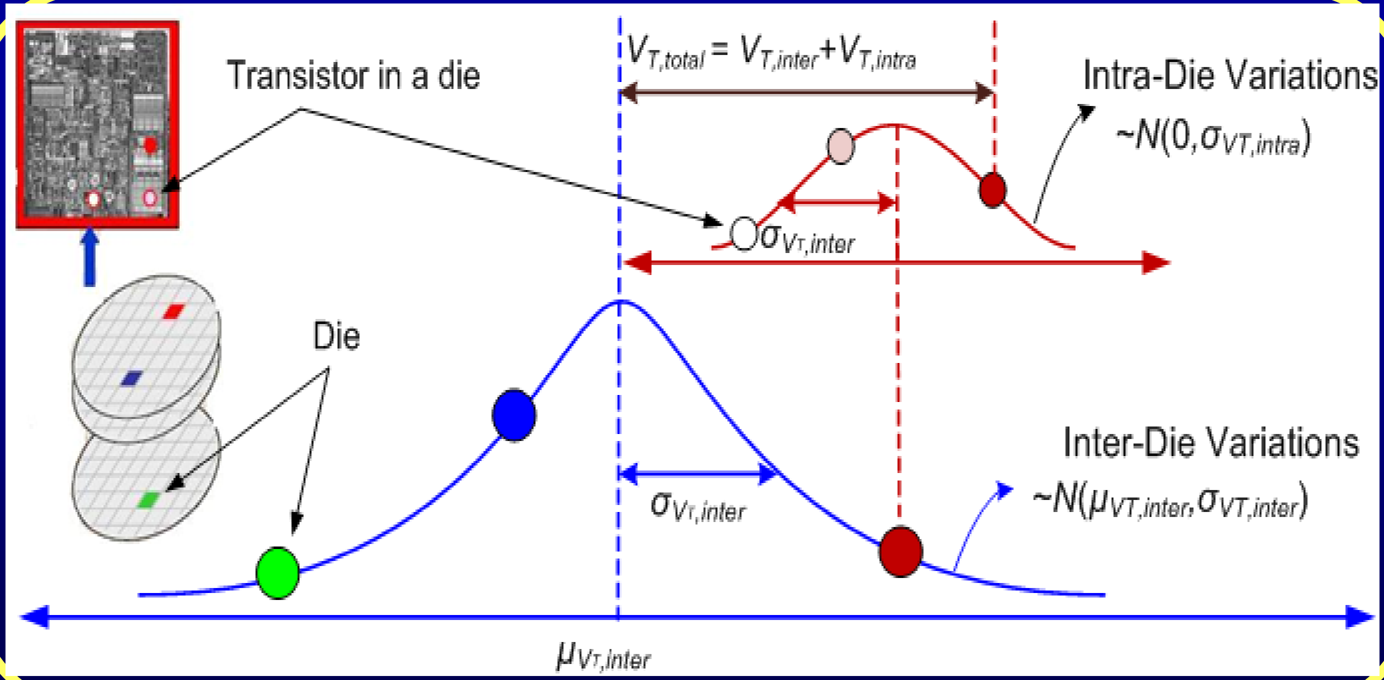
Background

- **Side-channel Analysis**
  - Measure effect of Trojan on some physical side-channel parameter, such as dynamic current, delay etc.
  - It does not require triggering the Trojan to observe its impact at primary output nodes.
  - Previous work:
    - IC Fingerprinting – D. Agarwal *et al*, Security and Privacy Symp. '07
    - Region-based approach – M. Banga *et al*, HOST '08
    - Multiple-parameter approach - S. Narasimhan *et al*, HOST '10
    - Multiple-power port approach - R. Rad *et al*, TVLSI '10
  - Power consumption in scaled technologies can vary by up to 20X due to process variations.

Background

➢ Effect of Process Variations

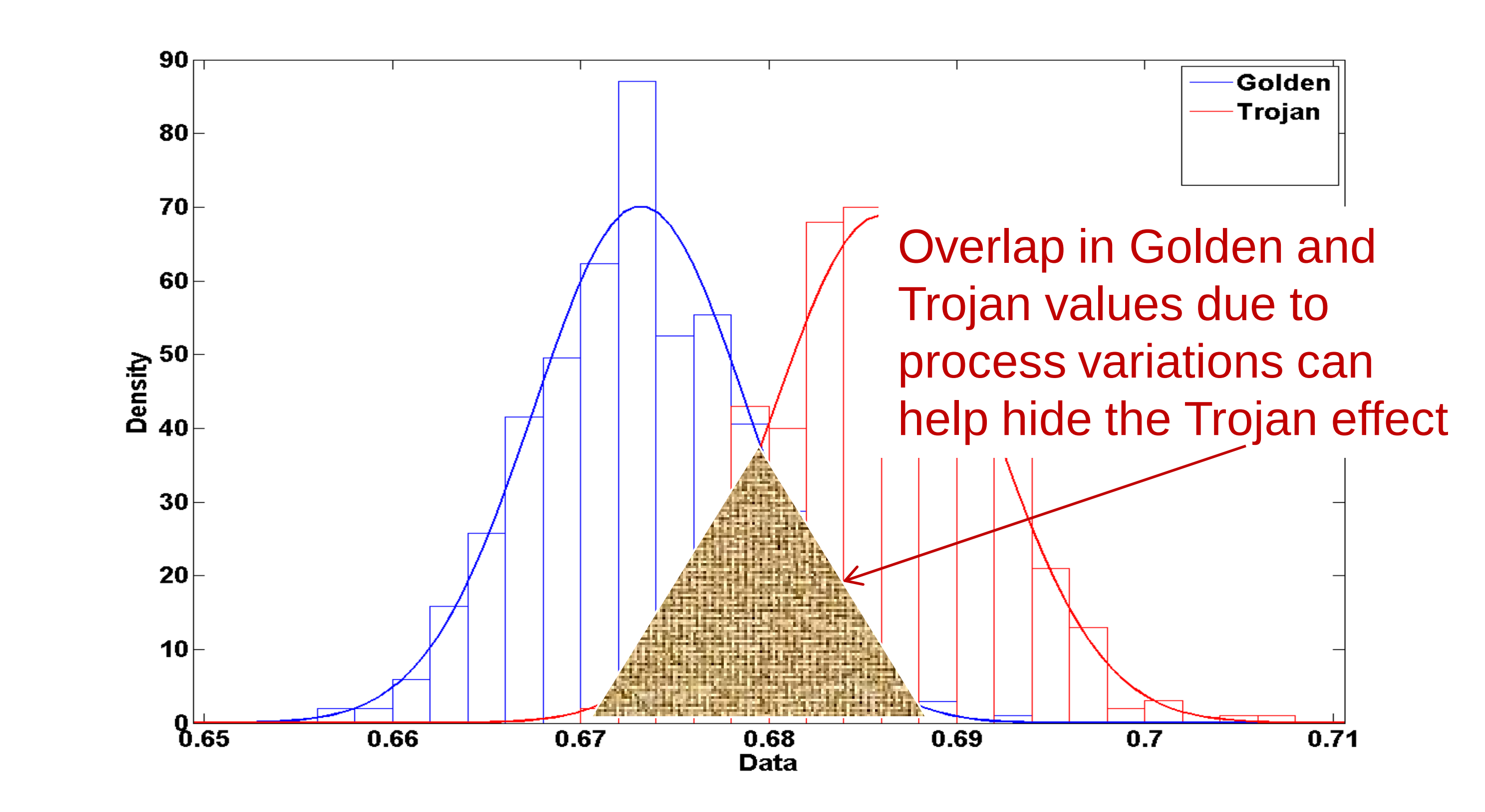
- Due to process variations, it is extremely challenging to detect the Trojan by only I_{DDT} individually.



Background

➢ Effect of Process Variations

- Due to process variations, it is extremely challenging to detect the Trojan by only I_{DDT} individually.



Background

- **Improving Detection Sensitivity**

$$Sensitivity = \frac{I_{tampered} - I_{original}}{\Delta I_{original(\mathbf{proc\_var})}} \times 100\%$$

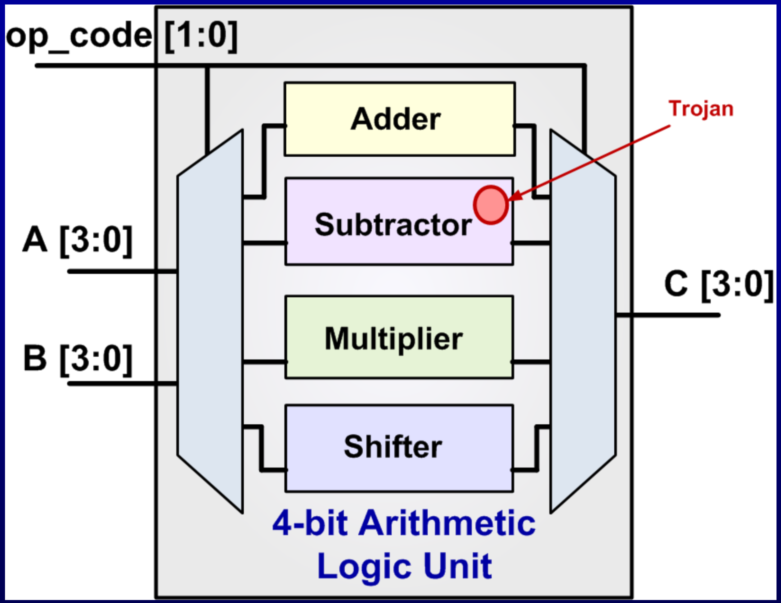
Trojan Size ↓ ➡ Sensitivity ↓

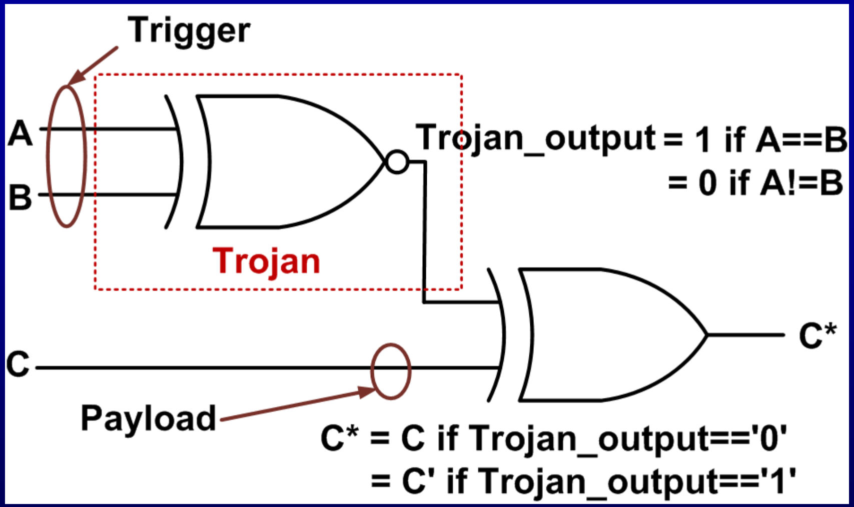
Circuit Size ↑ ➡ Sensitivity ↓

How to extend side-channel approach for detecting small Trojans in large circuits under process noise?

Motivational Example

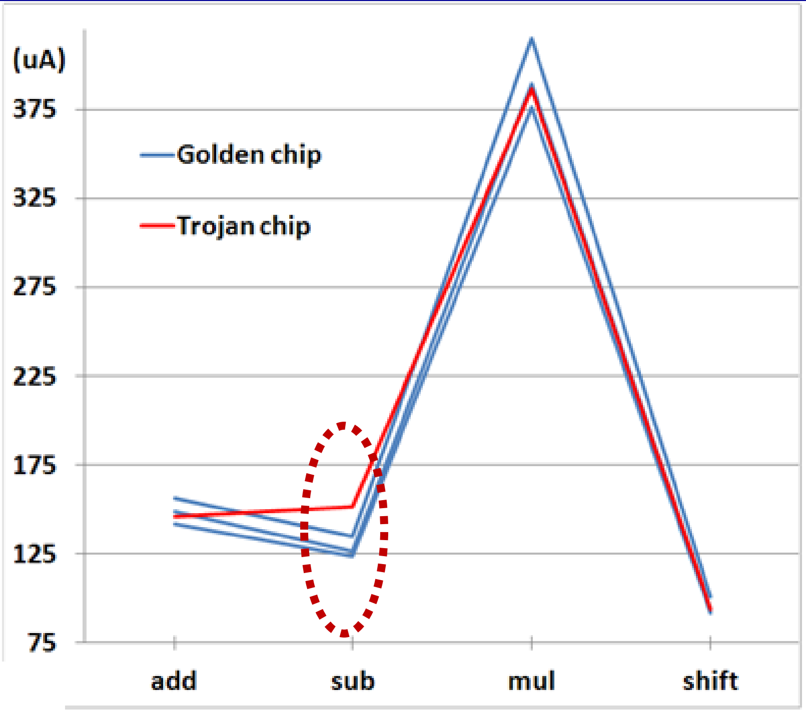
- Test circuit : 32-bit ALU.
- Trojan circuit : 1-bit comparator.
- The effect of process variations (both inter-die and intra-die) were simulated in HSPICE for the PTM 70nm technology by modulating the transistor V_{th}.

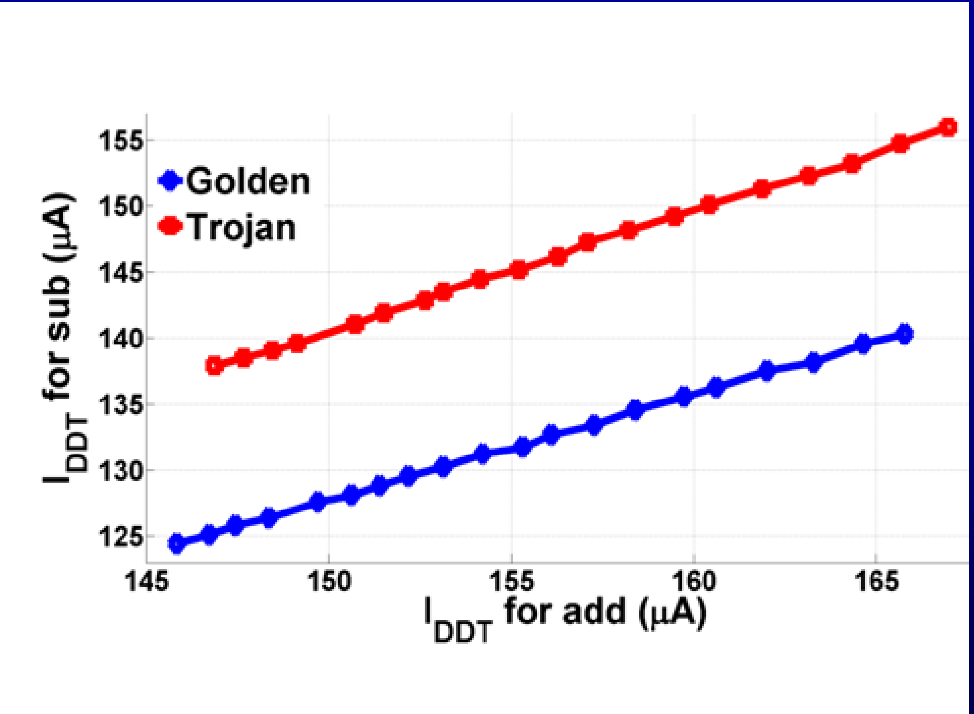




Motivational Example

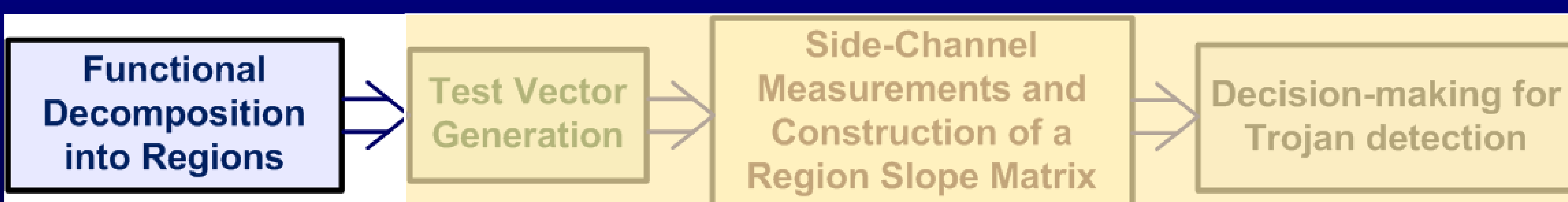
- Compare side-channel parameter I_{DDT} among different regions to isolate Trojan effect and location.
- The "slope" between the 4 regions shows that the Trojan is inserted in "sub" region. "I_{DDT} *for add*" acts as the reference.





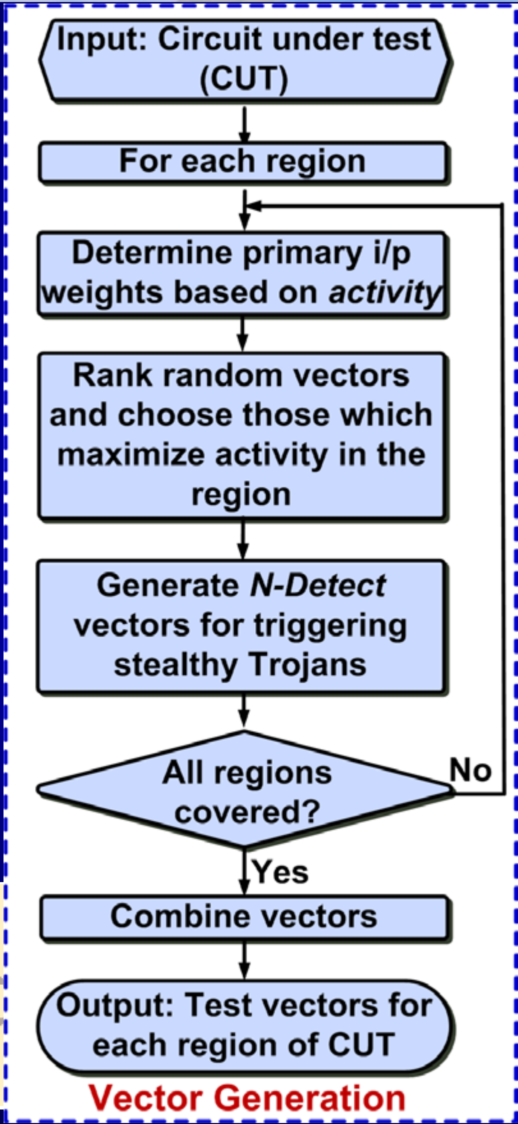
Methodology

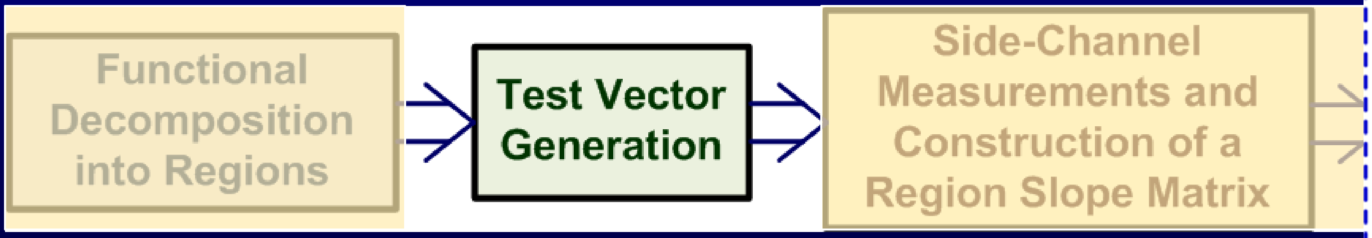
- **Functional Decomposition**
  - The circuit is broken into several small blocks which can be separately activated and compared against each other.
- **Main properties:**
  - Region size – Not too large and not too small
    – "*Goldilocks-sized*"
  - Functionally independent blocks
  - Hierarchical for larger SoCs



Methodology

- **Test Vector Generation**
  - The different regions need to be activated one-by-one.
- **Statistical Approach:**
  - In each region, the test vectors should cause some activity in all possible Trojan circuits.
  - The background current should be minimized.
  - For pipelined circuits, each stage is activated separately.





Methodology

➢ Self-Referencing

- The transient current I_i for each region is measured separately.
- The "*slope*" S_{ij} or relative difference in region currents is used to create a Region Slope Matrix.

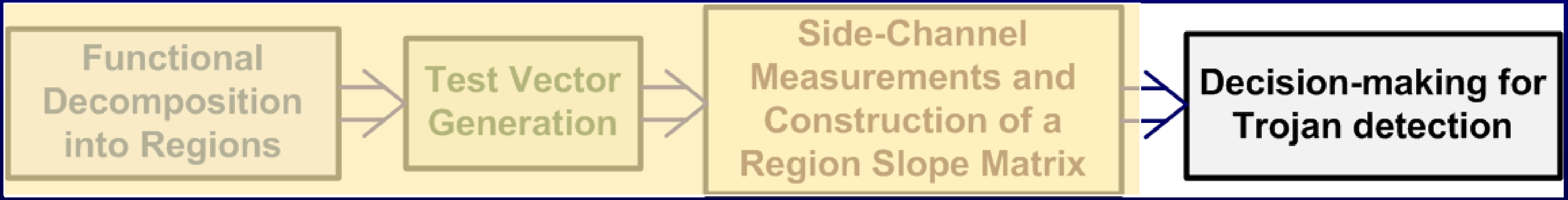
$$S_{ij} = \frac{I_i - I_j}{I_i}, \forall i, j \in [1, n]$$

- The region slope values are compared for golden ICs and threshold values are computed based on mean and σ values.
- The diagonal elements of the matrix are zeros.

Methodology

➢ Decision-making Process

- The Euclidean difference (L^2 norm) between the Region Slope Matrices of each IC with the golden nominal IC is computed.
- The Euclidean difference for a golden IC at a distant process corner is used as the Threshold value.
- Instead of a simple go/no-go decision, we come up with a confidence level regarding presence or absence of Trojan.
- The suspect ICs can be subject to hierarchical analysis.



Results

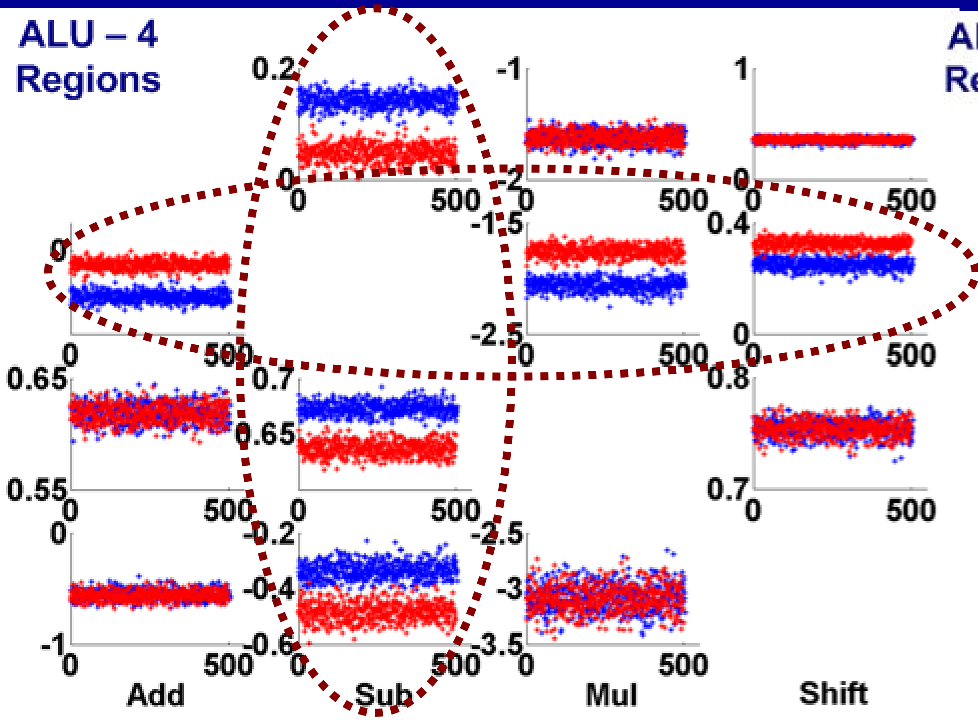
- The self-referencing approach was validated with simulation and experimental results.
- **Simulation Framework**
  - 32-bit Arithmetic Logic Unit (ALU) with 4 distinct regions for operations – add, sub, mul and shift.
  - 16-bit Finite Impulse Response (FIR) filter with 5 structural partitions.
  - A 32-bit DLX processor with 5 pipeline stages and the 32-bit ALU as its main execution stage.
  - The Trojan circuit consists of a small comparator to act as the trigger and an XOR gate for the payload.
  - To test sequential Trojans, we considered 16 flip-flops as a counter which are selectively activated by the trigger.

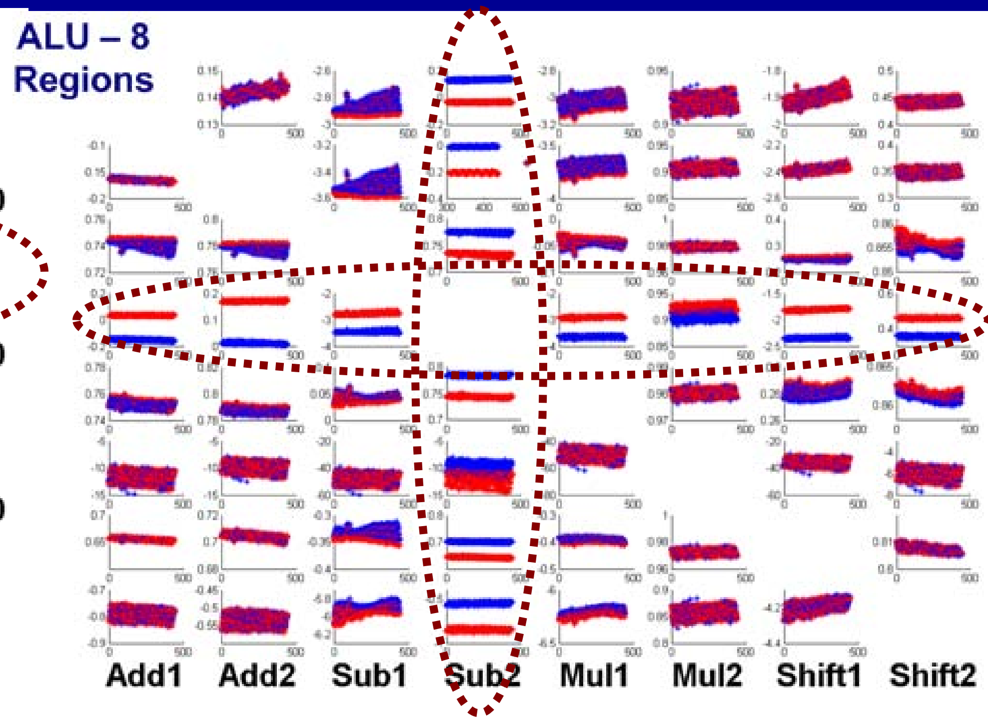
Results

➢ Simulation Results

- Region Slope Matrix for golden (blue) and Trojan (red)

32-bit ALU, Trojan in sub



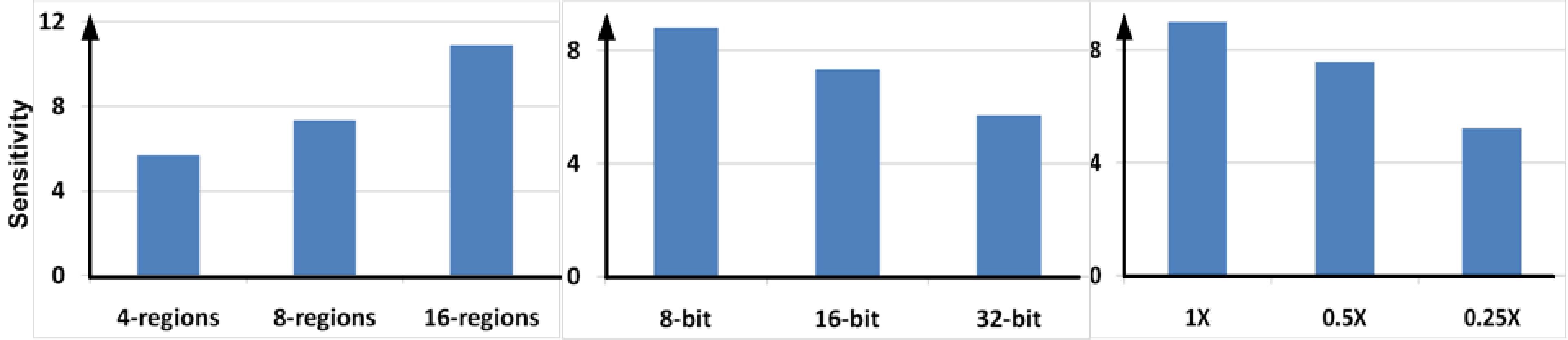


Number of regions can be increased to increase sensitivity.

Results

➢ Trojan Detection Sensitivity

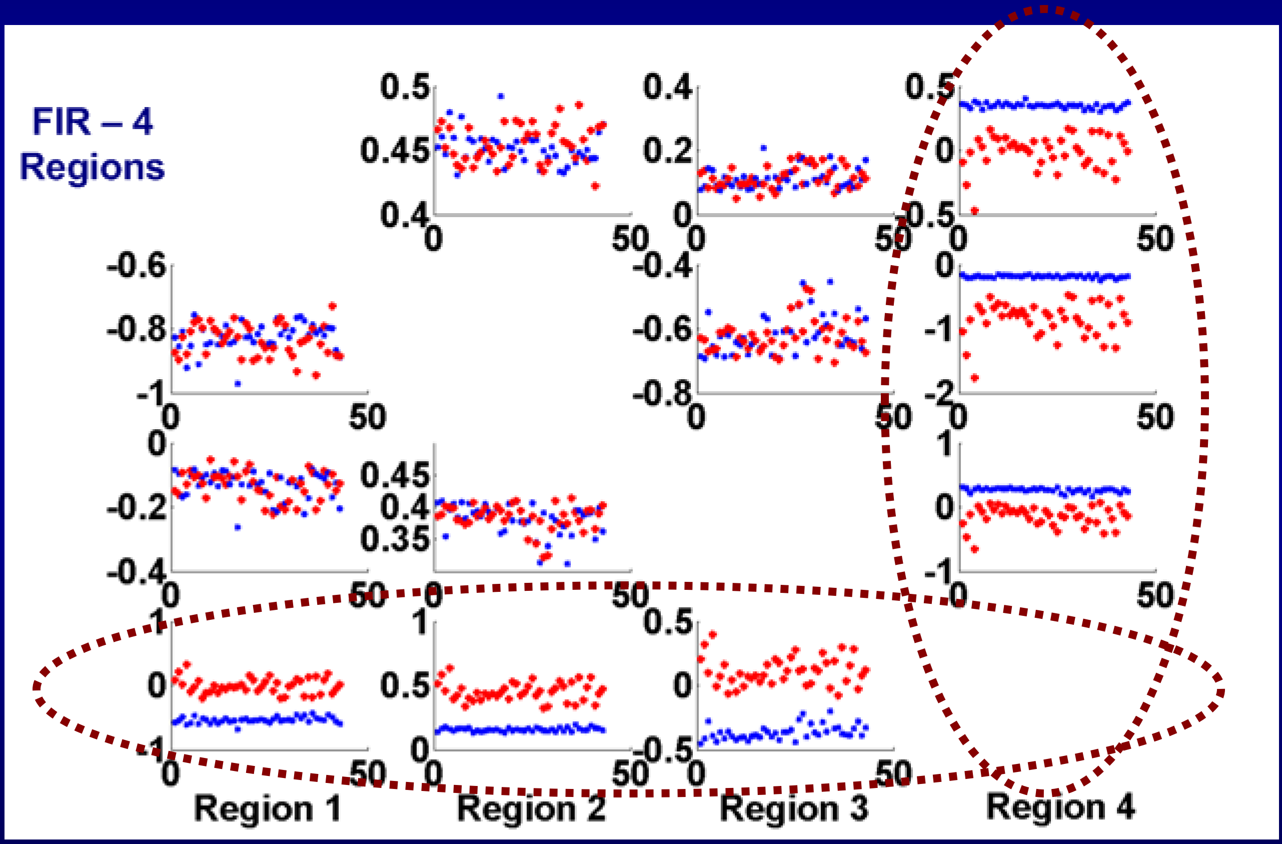
- Increases with increase in number of regions
- Decreases with increase in size of circuit
- Decreases with decrease in size of Trojan



16-bit ALU, Trojan in sub

Results

16-bit FIR filter, Trojan in 4th region



Results

- Monte Carlo simulations to observe effectiveness of self-referencing under both inter-die (σ =10%) and intra-die (σ= 6%) variations.
- The percentage of true negatives (correct detection of golden chip) and true positives (correct detection of Trojan) were noted.

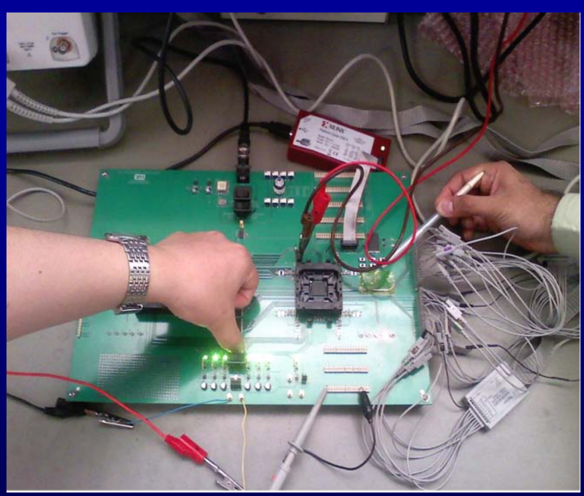
Circuit Name	True Negative	False Positive	True Positive	False Negative
32-bit ALU	99.10%	0.90%	5.90%	94.10%
FIR filter	97.72%	2.28%	6.60%	93.40%

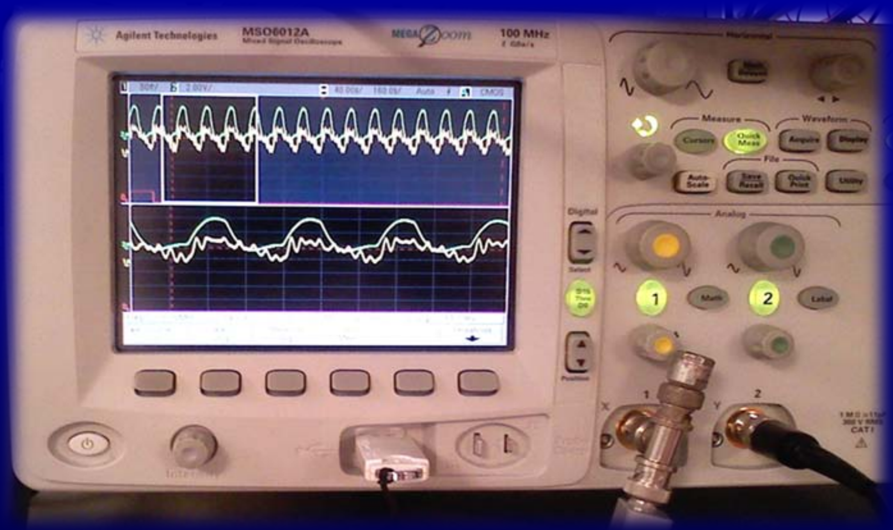
- The values are better for ALU, since the circuit is smaller, the regions can be separately activated.

Results

➢ Experimental Results

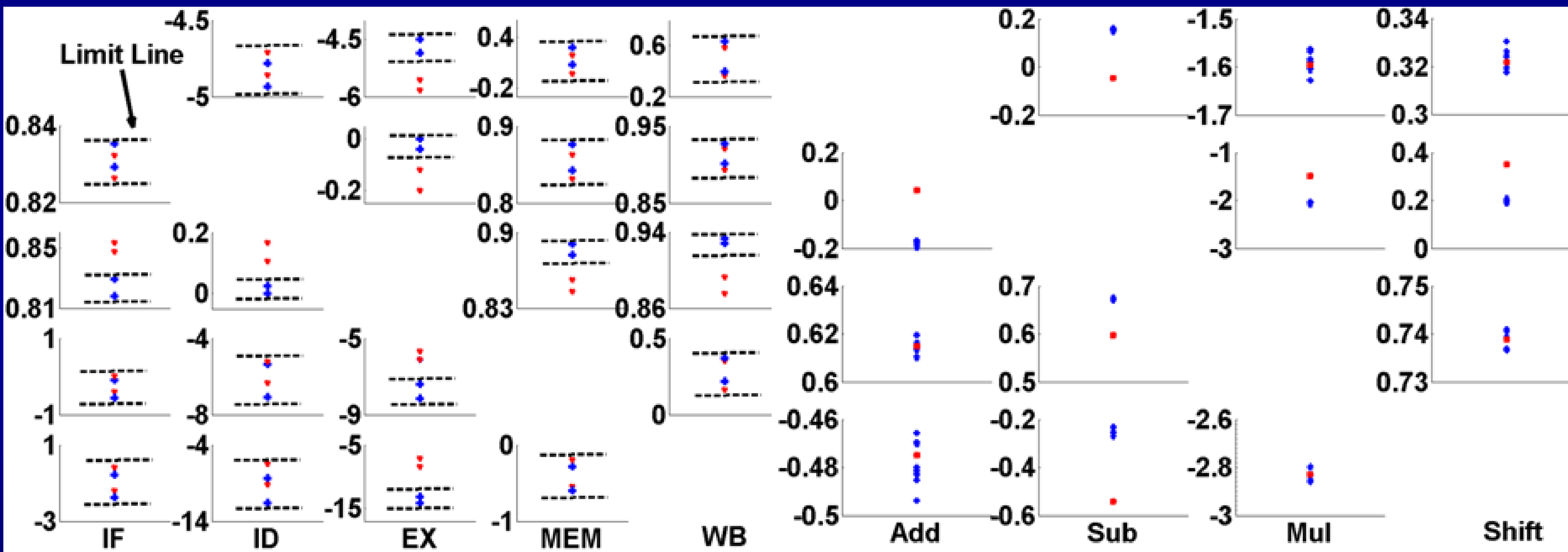
- Selected FPGA device was Xilinx Virtex-II XC2V500 fabricated in 120nm CMOS technology.
- We designed a custom test board with socketed FPGAs for measuring current from eight individual supply pins as well as the total current.





Results

➢ Experimental Results



32-bit DLX processor, Trojan in EX stage

32-bit ALU in EX stage, Trojan in ‘Sub’ region

Conclusion

- A novel side-channel analysis approach called *self-referencing* for hardware Trojan detection.
- The approach is scalable with respect to increasing die-to-die and within-die process variations in nanoscale technologies.
- We have also presented appropriate test vector generation method to improve the detection sensitivity.
- The approach is validated using both simulation as well as hardware measurements using 120nm FPGA chips.
- Combined with logic testing, it can detect ultra small Trojans for reliable detection of Trojans of all sizes.

Thank You

Questions ??