Theory and Applications of Cellular Automata for Detection and Mitigation of Hardware Trojan Attacks Targeting Cache Performance in a Many-core System

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of

Doctor of Philosophy

in

Computer Science & Engineering

By

Suvadip Hazra

(NITD/PhD/CS/2018/01035)

Under the Supervision
of

Dr. Mamata Dalui

Assistant Professor

Department of Computer Science & Engineering, NIT Durgapur



DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
NATIONAL INSTITUTE OF TECHNOLOGY DURGAPUR
WEST BENGAL 713209
INDIA

September, 2022

Abstract

Integration of untrusted component(s) in the Integrated Circuit (IC) design cycle or malicious modification(s) of the ICs in the design house or foundry has emerged as a potential security threat which is popularly known as the Hardware Trojan. Hardware Trojans are implanted within a system to cause system malfunction, destruction or leakage of confidential information. In this dissertation, we have analyzed Hardware Trojan from both - the attacker's and defender's point of view with the final aim to thwart the attack and build a secure system. From an attacker's perspective, we aim to design stealthy Hardware Trojans that can be hosted in a real-time many-core system and cause severe degradation in system performance due to the Trojan payload. Furthermore, from a defender's perspective, we target to design effective defence mechanisms to safeguard the system.

In a real-time many-core system, three potential areas have been identified for insertion of Hardware Trojans - the coherence mechanisms, memory transactions and replacement policies. The many-core systems integrating multiple processor cores in a single chip exploit the shared memory programming model, thus compels to rivet on the cache coherence mechanism for many-core systems having multi-level caches. Insertion of Hardware Trojans in the cache coherence controller causing disruption in the coherence mechanisms, may seriously hamper the operational reliability of the system. On the other hand, inducing a fake memory transaction by a Trojan can essentially lead to system performance degradation or even a system crash. The cache replacement policies play an important role in the cache performance. Thus, cache replacement policy, if affected by the Trojan payload, may have serious consequences on the system performance.

As the Trojans remain inactive for most of the time span of the affected system, they can easily evade conventional post-silicon tests/verification processes. Hence, unconventional solution, such as Cellular Automata (CA) based solution for detection and mitigation of Trojans has been explored. CA as a powerful modelling tool has long been received attention from researchers of diverse fields - VLSI circuit testing, pattern recognition, study of fractals and chaos, image processing, etc. The continual quest for more scalability, robustness and versatility of CA based models to study physical systems around us have forced researchers to explore wide variety of CA based models. Modelling physical systems, inherently irreversible in nature, involves a special class of irreversible CA. This dissertation explores properties of irreversible CA to model solutions for Hardware Trojan attacks in the context of multi-core processors' cache system. Theory of irreversible CA has been developed that exhibits its power in the potential application domain.