Hardware Trojan Horses: The New Face of Cyber Terrorism

Krishnendu Guha[1], Debasri Saha[2], Amlan Chakrabarti[3]
kgchem_rs@caluniv.ac.in[1], sahadebasri@gmail.com[2], acakcs@caluniv.ac.in[3]
A. K. Choudhury School of Information Technology
University of Calcutta

Introduction:

The present era is witnessing a merge of the physical and the cyber world. Previously, malicious operations of the cyber domain were confined to hacking and phishing operations. These were mainly associated with information stealing for monetary benefits but did not threaten innocent lives. Thus, there existed only cybercrimes but not cyber terrorism. However, with the involvement of technology in the day-to-day activities of common individuals, the nature of cyber crimes has changed. Technology malfunction can not only cause distress to common individuals but also can create mass terror and even lead to consequences, which may be catastrophic in nature. Thus, threats to physical entities is not only confined to the physical realm, but is also affected by malicious operations of the virtual domain. These led to the coining of the term “Cyber-Terrorism” in the late 1990s by Barry C. Colin [1].

As per FBI, cyber terrorism is defined as “*premeditated, politically motivated attack against information, computer systems, computer programs, and data which results in violence against non-combatant targets by subnational groups or clandestine agents*” [2]. Common cyber terror activities involve spreading of disputed propaganda via social media or by hacking common websites [3,4]. Such activities even create mass panic like the “*millennium bug*” incident. Though many are hoax, yet these are successful in creating fear and terror among common individuals. The scenario becomes quite serious when innocent lives are affected as side effects to such activities. Instances like the Saudi petrochemical sabotage attempt in 2018 and damage to Iran’s nuclear program by the malicious cyber worm, *Stuxnet* are quite alarming [5].

The breach of security was mainly confined to software and hardware was considered trusted. Researchers found direct execution of tasks in hardware was a convenient and safe solution for a secured system. Hence, the embedded era emerged. In the embedded regime, design of dedicated system on chips (SoCs) gained prominence. Design of an SoC involves several phases as evident from Figure 1. These design phases are not only complex, but also time consuming. Even many design sites lack the expertise to carry out all the phases of chip design. However, consumer demand increased with time. To meet stringent marketing deadlines and reduce cost, the semiconductor design industry adopted the globalization strategy for SoC designing [6].

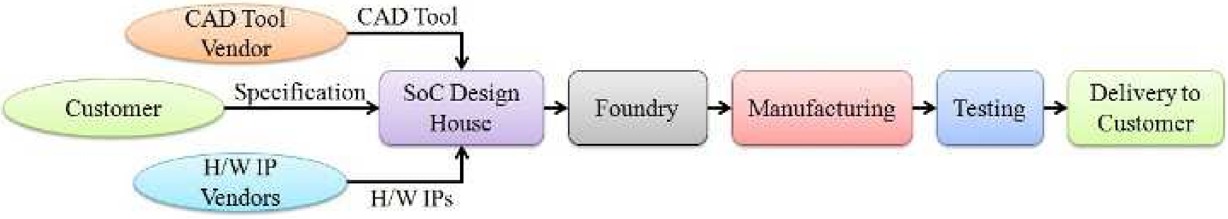

Figure 1: Stages in SoC Design

In the globalization technique, design modules or intellectual properties (IPs) are procured from various third party IP (3PIP) vendors, which are integrated to form the entire SoC. Even the various phases of chip design are outsourced to different parts of the world. Though such a technique reduced SoC design cost and facilitated meeting of stringent marketing deadlines, but the element of hardware trust had been evicted [6]. It is difficult to trust the 3PIP vendors who supply the IPs. Scenarios are common where malicious codes are introduced during the hardware description language (HDL) phase of IP design [7]. Adversaries in the outsourced foundries can even introduce malware during chip fabrication [8]. Such malicious elements are commonly known as Hardware Trojan Horses (HTH).

Hardware Trojan Horses:

HTHs possess the ability of remaining dormant during testing and the initial phases of operation, but get suddenly activated at runtime to jeopardize real time mission critical operations [6]. A trigger and a payload module are the essential elements of an HTH architecture as depicted in Figure 2. The trigger module can be either a combinational or a sequential circuit, which is basically an activation function. The trigger can either be internal like a rare combination of node values or external like activation with the aid of radio signals, which are received via an antenna. Only when the trigger criterion is satisfied, the malicious functionality encapsulated in the payload is activated. Effects of the payload may vary and generally depend on the nature and extent of the harm intended to be caused by the adversary.

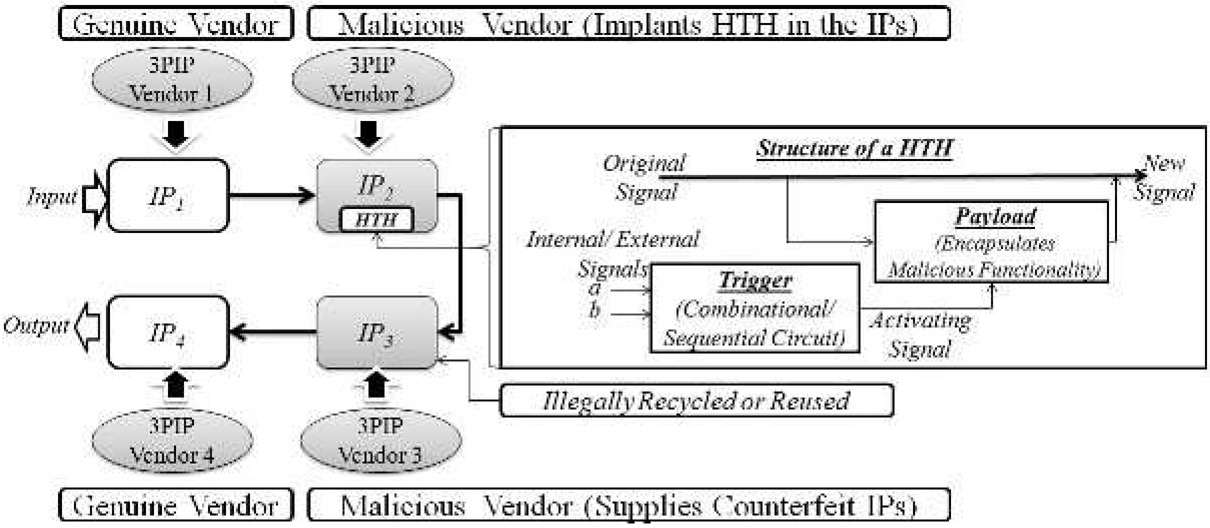
Genuine Vendor
Malicious Vendor (Implants HTH in the IPs)
3PIP Vendor 1
3PIP Vendor 2
Input
IP_1
IP_2
HTH
Structure of a HTH
Original Signal
New Signal
Internal/ External Signals
a
b
Payload
(Encapsulates Malicious Functionality)
Trigger
(Combinational/ Sequential Circuit)
Activating Signal
Output
IP_4
IP_3
Illegally Recycled or Reused
3PIP Vendor 4
3PIP Vendor 3
Genuine Vendor
Malicious Vendor (Supplies Counterfeit IPs)

Figure 2: Vulnerability in SoC Design and Structure of an HTH [16]

The US Government of Defense had even recognized HTHs as a significant threat to mission critical applications in 2005 [15]. HTHs may cause both active and passive attacks and possess the ability to jeopardize the basic security primitives of a system. Erroneous result generation may affect the integrity [9], while leakage of secret information may affect the confidentiality of the system [10,16]. Degrading system performance via sudden induced delays at runtime will affect system availability [11]. Even real time tasks of mixed critical systems may be affected by HTHs [12].

Difference from Traditional Faults:

Though effects of HTHs are quite similar to traditional faults occurring at runtime, yet their nature differs. Faults occurring in a system at runtime are unintentional and not preplanned, but attacks of a HTH are pre-planned and intentional. Faults may occur during testing and in such a scenario, the faulty IPs are replaced, but HTHs are designed in such a way that they will never exhibit their malicious characteristic during testing and will only get triggered at runtime. Moreover, it is possible to detect faults during post-mortem analysis, but HTHs may exhibit their malicious behavior and then regain their dormant nature during post-mortem analysis. Hence, detection of HTHs is quite difficult than faults.

Point of Concern:

With the entry into the embedded era, dependence on embedded devices has increased. Such embedded devices constantly monitor the operations of their host users. Many of these embedded devices like the embedded healthcare appliances are even attached to human bodies and directly affect human functionality. Implanted HTHs in such embedded devices may leak secret information related to their host users. This in turn affects the privacy of the users, without their knowledge. Moreover, if such embedded devices are associated with critical infrastructure, then their malfunction during critical stages of operation will lead to fatal consequences and loss of innocent lives.

Rate of import of embedded devices, which ranges from simple electronic devices to critical healthcare appliances and defense weaponry, is high for nations across the globe. Conventional tests which are performed, are unlikely to detect such malicious hardware implantations as these remain dormant unless the trigger is activated, whose control remains with the manufacturing countries. Thus, concern remains in the outbreak of emergency which may affect innocent lives, via the embedded devices utilized.

Defense against HTHs:

Existing strategies to counteract the threats of HTHs are mainly classified into three categories, i.e. (i) Test Time Detection Techniques, (ii) Protection based on Authentication and (iii) Runtime Mitigation Strategies. We briefly discuss each of these in this section.

(i) Test Time Detection Techniques: Test time detection techniques are basically of two types, namely logic testing and side channel analysis, as shown in Figure 3 [6]. As, HTHs remain dormant during test phases, hence, conventional testing strategies are unable to detect HTHs. In logic testing, special test vectors are generated to trigger the malicious effects of the HTHs before their activation. Though this is useful for simple IPs, but the methodology loses its effectiveness when the

IP is complex. This is because, with increase in complexity, it is difficult to generate test vectors for all the nodes of the IPs. Moreover, modern IPs are delivered in an unreadable format so that the users are unable to replicate them and without knowing the structure of the IPs, it’s difficult to generate the test vectors.

In side channel analysis, refuge is sought to side channel parameters like delay, power, leakage current, etc to detect the presence of HTHs in the procured IPs. In this technique, the values of the side channel parameters of the experimental IP are compared with a reference or golden IP. However, when the HTH size is small and the IP size is large, the merits of side channel analysis are limited. This is due to the fact that the difference in side channel parameters will be negligible and will be hard to detect.

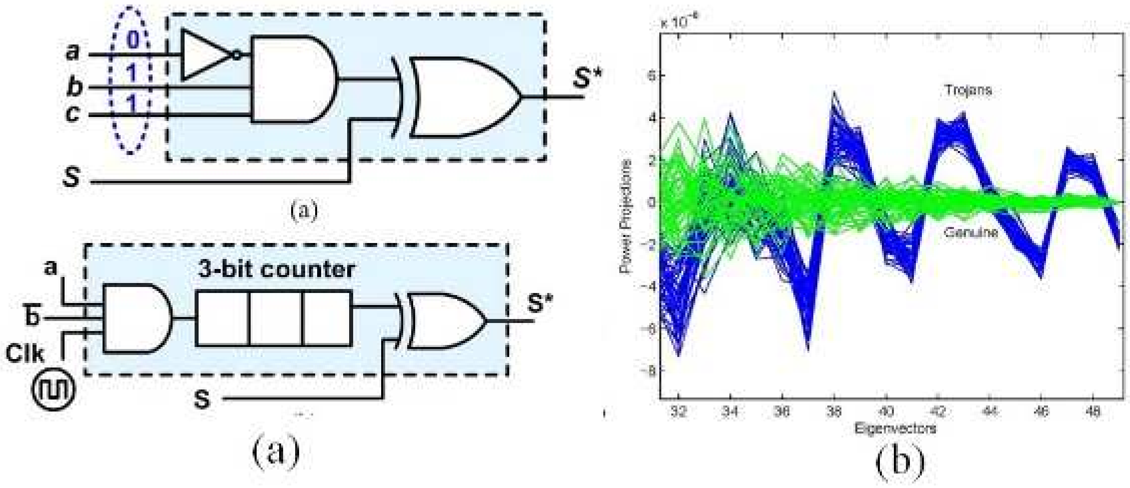

Figure 3: Diagrammatic Representation of Test Time Detection Strategies (a) Logic Testing (b) Side Channel Analysis [6]

(ii) Protection based on Authentication: Identification is necessary to confirm genuineness of the procured IPs. Moreover, tracking the vendors who supplied the malicious components in case of a system malfunction and taking appropriate action can ensure trust. To facilitate this, proof of authentication needs to be appended with the supplied IPs. Such proof of authentication must be unique and non- replicable. For this, physical unclonable functions (PUFs) are used, which utilizes the non-replicable properties of the semiconductor devices to generate an unique identity [13]. Other than PUFs, watermarking is also a convenient option [14]. In watermarking, the producer implants a watermark in the IP, which is duly verified by the user. Diagrammatic representations of these are shown in Figure 4.

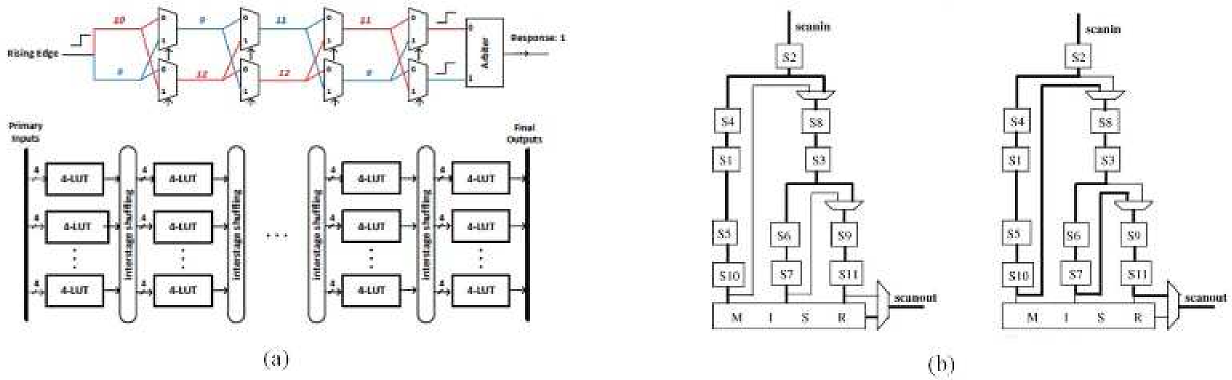

Figure 4: Diagrammatic Representation of Authentication Mechanisms (a) a 4 BIT Delay PUF implemented in FPGA architecture[13] (b) A Watermarking Technique which facilitates authentication by reordering Scan Cells S={S1, S2,…,S11}[14]

(iii) Runtime Mitigation Strategies: Such techniques are termed as the last line of defense by eminent researchers of this arena [6]. This can either be a redundant approach or a self-aware approach. In the former, multiple IPs are procured from different vendors for a particular functional operation. The same task is redundantly carried out in all the IPs and the correct result is generated after majority polling of all the results [9]. As the IPs are procured from different sources, hence, the same HTH cannot be implanted in all. Even if they cause the same malfunction, their activation times must be different. And as majority polling is carried out, the result generated is correct. More is the number of IPs procured for a particular task, more is its effectiveness. However, associated cost increases with increase in redundancy.

In the self-aware approach, no redundant operations are performed. Instead, a self-aware module is associated with the IPs, which works based on the Observe-Decide-Act (ODA) paradigm [10,11]. Operations of the IPs are constantly monitored in the Observe phase. Whenever an anomaly of operations is monitored, the Decide phase is triggered. In the Decide phase, it is deciphered whether the change in state of operations is associated with the objective of causing malfunction or not. Operations in the Act phase is carried out based on the inference of the Decide phase. If a malfunction is deciphered, then operations of the IPs are temporarily stopped or bypassed to prevent system damage, else the new state is learned and operations continue. Figure 5 demonstrates a redundancy approach and a self aware approach to ensure security.

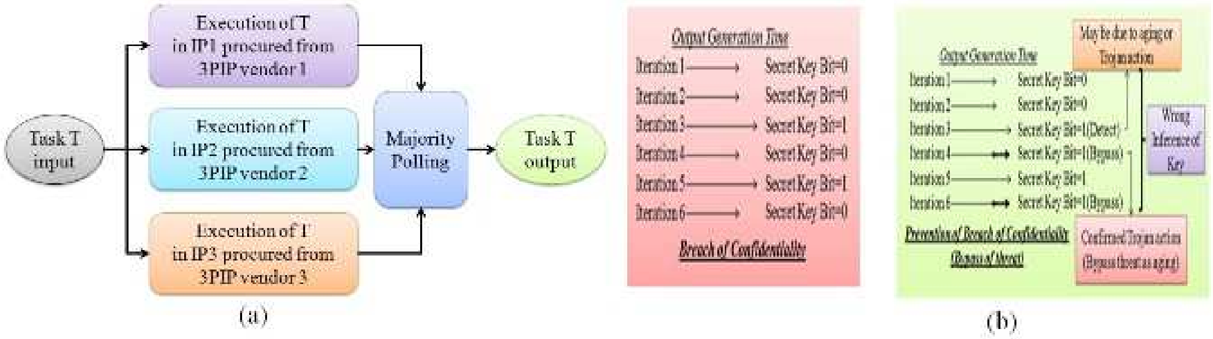

Figure 5: Diagrammatic Representation of Runtime Mitigation Strategies (a) Redundancy Approach (b) Self Aware Approach

Conclusion:

The nature of threat changes with time. With the entry into the embedded era, hardware threats gained prominence. Malware such as HTHs remain dormant during testing and gets activated at runtime. HTHs possess the capability to jeopardize basic security primitives of a system and its effects can be life threatening and fatal. With such consequences, HTHs can be considered as a new face of cyber terrorism in the recent embedded era. Though test time detection techniques are available, yet they are not full proof. Authentication techniques induce trust to a certain extent. But to ensure full proof trust, runtime mitigation strategies must be deployed, which are termed as the last line of defense.

References

1. William L. Tafoya, “Cyber Terror”, *FBI Law Enforcement Bulletin (FBI.gov)*, November 2011
2. Centre of Excellence Defence Against Terrorism, ed. (2008). *I. NATO science for peace and security series. Sub-series E: Human and societal dynamics*, ISSN 1874-6276
3. Warf, Barney, “Relational Geographies of Cyberterrorism and Cyberwar”, *Space & Polity*, vol. 20, no. 2, pp. 143–157, August 2016. *doi*:*10.1080/13562576.2015.1112113*
4. T. J. Holt, J. D. Freilich, S. M. Chermak, “Exploring the Subculture of Ideologically Motivated Cyber-Attackers”, *Journal of Contemporary Criminal Justice*”, vol. 33, no. 3, pp. 212–233, 2017. *doi*:*10.1177/1043986217699100*
5. A. Rudawski, “The Future of Cyber Threats: When Attacks Cause Physical Harm”, *New York Law Journal, 2018*
6. S. Bhunia, M. S. Hsiao, M. Banga, S. Narasimhan, “Hardware Trojan Attacks: Threat Analysis and Countermeasures,” *Proceedings of the IEEE*, vol. 102, no. 8, pp. 1229-1247, 2014
7. C. Liu, J. Rajendran, R. Karri, “Shielding Heterogeneous MPSoCs From Untrustworthy 3PIPs Through Security- Driven Task Scheduling”, *IEEE Tranctions. on Emerging Topics*, vol. 2, no. 4, pp. 461-472, 2014
8. K. Xiao and M. Tehranipoor, “BISA: Built-in self-authentication for preventing hardware Trojan insertion”, *2013 IEEE International Symposium on Hardware-Oriented Security and Trust (HOST)*, Austin, TX, 2013, pp. 45-50
9. D. McIntyre, F. Wolf, S. Bhunia, “Dynamic Evaluation of Hardware Trust”, *IEEE Hardware Oriented Security and Trust*, pp. 108-111, 2009
10. K. Guha, D. Saha, A. Chakrabarti, “RTNA: Securing SOC architectures from confidentiality attacks at runtime using ART1 neural networks”, *2015 19th International Symposium on VLSI Design and Test (VDAT)*, Ahmedabad, India, 2015, pp. 1-6
11. K. Guha, D. Saha, A. Chakrabarti, “Self Aware SoC Security to Counteract Delay Inducing Hardware Trojans at Runtime”, *2017 30th International Conference on VLSI Design and 2017 16th International Conference on Embedded Systems (VLSID)*, Hyderabad, 2017, pp. 417-422
12. K. Guha, A. Majumder, D. Saha, A. Chakrabarti, “Reliability Driven Mixed Critical Tasks Processing on FPGAs Against Hardware Trojan Attacks”, *2018 21st Euromicro Conference on Digital System Design (DSD)*, Prague, 2018, pp. 537-544
13. T. Xu, M. Potkonjak, “Robust and exible FPGA-based digital PUF,” in *2014 Proc. Field Programmable Logic and Applications*, pp.1-6
14. D. Saha, S. Sur-Kolay, “Embedding of signatures in reconfigurable scan architecture for authentication of intellectual properties in SoC”, *IET Computers and Digital Techniques*, vol. 10, no. 3, pp. 110-118, 2016.
15. Defense Science Board, “Task Force on High Performance Microchip Supply”, 2005, Available Online: http://www.acq.osd.mil/dsb/reports/ADA435563.pdf

16. K. Guha, D. Saha, A. Chakrabarti, “Real-Time SoC Security against Passive Threats Using Crypsis Behavior of Geckos”, *ACM Journal of Emerging Technologies in Computing Systems*, vol. 13, no. 3, Article 41, 2017

About the authors

Krishnendu Guha is presently a Senior Research Fellow in A. K. Choudhury School of Information Technology (AKCSIT), University of Calcutta. Prior to this, he has completed his MTech from University of Calcutta in 2014, where he received the University Gold Medal. He was awarded the prestigious INSPIRE Fellowship by the Department of Science and Technology, Government of India for carrying out his PhD work. He received the 2nd Runners Up PhD Forum Award in the premier international conference VLSID 2018. He is a student member of IEEE and ACM. His present research arena encompasses embedded security, with a flavour of artificial intelligence and nature inspired strategies.

Debasri Saha is presently an Assistant Professor in AKCSIT, University of Calcutta. Prior to this, she was associated with IIT Patna as an Assistant Professor after completing her PhD from ISI, Kolkata. She completed her M. Tech. From the Department of Computer Science and Engineering, University of Calcutta in 2006, where she received the University Gold Medal. She is a member of IEEE. Her research interests include VLSI design and its security issues, optimization and heuristic techniques.

Amlan Chakrabarti is presently Professor and Director of AKCSIT, University of Calcutta. He is also the Dean of Faculty for Engineering and Technology in University of Calcutta. Prior to this, he completed his post doctoral research in Princeton University after completing his PhD from University of Calcutta in association with ISI, Kolkata. He is the recipient of DST BOYSCAST fellowship award in Engineering Science in 2011, Indian National Science Academy (INSA) Visiting Faculty Fellowship in 2014, JSPS Invitation Research Award in 2016, Erasmus Mundus Leaders Award from EU in 2017 and Hamied Visiting Fellowship from Cambridge University in 2018. He has been associated with reputed international and national institutes of repute as a Visiting Professor to name a few, University of University of Cambridge, City University of London, University of Oradea Romania, SUNNY Buffalo USA, GSI Helmholtz Research Laboratory Germany, University of Bremen Germany, CERN Geneva, Kyushu Institute of Technology Japan etc. He has received multiple project grants in the areas of Embedded System Design,VLSI Design, Quantum Computing, Cybersecurity, and Computer Vision from various national and international agencies. He is a Senior Member of IEEE and ACM, Secretary of IEEE CEDA India Chapter and Vice President of Society for Data Science. His present research interests include VLSI Design, Quantum Computing and Embedded System Design.

IEEE Computer Society predicts top ten tech trends for 2019

Deep learning accelerators

Assisted transportation

The Internet of Bodies (IoB)

Social credit algorithms

Advanced (smart) materials and devices

Active security protection

Virtual reality (VR) and augmented reality (AR)

Chatbots

Automated voice spam (robocall) prevention

Technology for humanity (specifically machine learning)

Source: https://hub.packtpub.com/ieee-computer-society-predicts-top-ten-tech-trends-for-2019-assisted-transportation-chatbots-and-deep-learning-accelerators-among-others/