Can the SHIELD protect our integrated circuits?

Farinaz Koushanfar
Rice University
Houston, TX USA
Email: farinaz@rice.edu

Ramesh Karri
New York University
Brooklyn, NY USA
Email: rkarri@nyu.edu

***Abstract*—Mass production of Integrated Circuits (ICs) from a single blueprint (mask) renders inherent identification of the individual parts a challenge. Indelible marking of the ICs can enable fingerprinting, identification, authentication, metering, and tracing of components along the unascertained semiconductor supply chain. To enable these important objectives, DARPA is soliciting innovative proposals for a SHIELD that enables advanced supply chain hardware authentication capability. The envisioned SHIELD is intended to be a minuscule electronic chip that is physically and inseparably attached to the host electronic component. The desiderata for the SHIELD include providing an ineradicable hardware root-of-trust for cryptographic key storage and encryption, a compact structure encapsulating the keys, a physically-fragile but electrically-robust SHIELD dielet that self-destructs upon adversarial acts, an RF communication and remote charging interface, and sensors for recording the potential attack attempts. We discuss the SHIELD threat model and its potential for addressing a number of standing challenges in this area. We emphasize the dire need for open evaluation and thorough security analysis of SHIELD.**

I. Introduction

During the past two decades, there has been a continuous trend away from in-house integrated circuit (IC) design and fabrication towards out-sourcing various aspects of design, fabrication, testing and packaging of ICs. The emergence of such a globalized, horizontal semiconductor business model created hitherto unknown security and trust concerns in the ICs and the information systems (rooted in these ICs) on which modern society relies on for mission-critical functionality [1]. IC and system security and trust concerns include threats related to the malicious insertion of trojan circuits designed to act as silicon time bombs to disable an IC [2], [3], to intellectual property (IP) and IC piracy [4], [5], [6], [7], to untrustworthy 3rd party IPs [8], to exfiltrating sensitive material from an IC [9], [10], and to malicious system disruption and diversion.

Mass production of Integrated Circuits (ICs) from a single blueprint (mask) renders inherent identification and tracing of the individual parts a challenge [4], [5], [6], [7]. Traditional identification methods such as package/board serial numbers and storing IDs in the access-disabled non-volatile memory (e.g., by burning fuses) are subject to removal and remarking. Nondestructive tests for IC identification have been done [11] but those methods are typically slow and impractical for wide-scale adoption. Physical unclonable functions (PUFs) provide a promising way for chip unclonable identification if careful safeguarding against the attacks is provided [12], [13], [14], but they cannot provide stand-alone solutions as a root-of-trust without additional components.

Remarking is a particularly serious problem, giving rise to a surge in counterfeit chips; it is enabled by the advances and accessibility of inexpensive packaging technology that is used to repackage the old chips as new ones. The internals of ICs are opaque and hard to identify and test so detection of recycled ICs is a challenge. Even identification of the fabricating house is a hard problem [15]. What exacerbates the problems is the long chain of globally distributed (untrusted) suppliers that are involved in the IC production, testing, integration, and marketing. It is almost infeasible to monitor or control the chip supply chain entities.

The inability to identify, monitor, and trace authentic components in the IC supply chain has at least four important ramifications. First, the original IC part providers incur an irrecoverable loss due to the sale of (often) cheaper counterfeit components. Second, low performance of counterfeit products (that are mostly of lower quality and/or cheaper older generations of a chip family) affects the overall efficacy of the integrated systems that unintentionally use them; this also could harm the reputation of authentic providers. Third, unreliability of fake devices could make the integrated systems that unknowingly utilize those parts unreliable; this potentially affects the performance of weapons, airplanes, cars or other crucial applications that use the fake components [16]. Finally, untrusted fake components may have intentional malware or backdoors.

Indelible marking of the ICs can enable fingerprinting, identification, authentication, metering, and tracing of components along the unascertained semiconductor supply chain. To enable these important objectives, a recent call by DARPA is soliciting innovative research proposals for a SHIELD that enables advanced supply chain hardware authentication capability [17]. The SHIELD program has the potential to provide an exciting opportunity for ingraining trust in the unreliable chip design, manufacturing, testing, system integration, and distribution processes.

In a nutshell, the SHIELD program aims to create and develop a secure root-of-trust in hardware. This root-of-trust shall be co-packaged with the electronic components to provide them with permanent unique identification that is technologically infeasible (or prohibitively expensive) to alter or copy. The same root-of-trust can be utilized to authenticate the provenance of the containing electronic components in an assembly (through the long IC supply chain.) Furthermore, the final consumer of the assembly could use the SHIELD to re-authenticate the composition of the final installation of the IC throughout its lifetime.

To ensure practicality and resilience to attacks, several performance, cost, and attack vectors have to be consid-

978-1-4799-4132-2/14/$31.00 ©2014 IEEE

ered in the SHIELD design process. As a result, multiple methodological, testing, and automation challenges arise in the design of the envisioned SHIELD device. Although high level performance/cost metrics and adversarial models were described in the SHIELD BAA [17], the exact relevance of the SHIELD in thwarting the contemporary threats is not clear.

This paper aims to fit the SHIELD approach and its capabilities within the new classification of hardware threats, countermeasures, and metrics recently provided in [1]. Our classification of SHIELD provides a guide for the SHIELD designers to clearly state the assumptions, challenges, and solutions; it will allow them to fairly evaluate their solutions, find the susceptibility to attacks, and provide stronger countermeasures to the potential vulnerabilities. We conclude the paper by suggesting open analysis methods for ensuring a secure and protected SHIELD design.

The remainder of the paper is organized in the following way. We start by outlining the SHIELD desiderata in Section II. Section III discusses the relevance of the SHIELD to the classification in [1]. The open evaluation of design and thorough security analysis of the SHIELD are suggested in Section IV. Section sec:conc concludes the paper.

II. The SHIELD desiderata

The SHIELD program plans to provide a secure hardware root-of-trust that will be co-packaged with an electronic component. The SHIELD shall provide a unique, permanent identification that is infeasible, or prohibitively expensive to copy or modify. This root of trust shall be utilized to authenticate the component on the long IC supply chain, as well as by the final consumer of the product for the purpose of component re-identification throughout the chip's lifetime. The SHIELD will be encaged in a dielet.

The most important desired features of the SHIELD were described as follows [17]:

1) It must provide a hardware root-of-trust for cryptographic key storing that is prohibitively expensive or practically impossible (e.g., time consuming) to reverse-engineer.
2) The cryptographic key shall never leave the SHIELD dielet but this key must be utilized in a complete, compact and on-board key encryption engine that is capable to encrypting an external challenge using the on-board cryptographic key.
3) The dielet must be physically fragile so that it can self-destructs upon any attempts to physically remove, modify or open it while it must also be electronically robust to last the lifetime of the IC.
4) The design should include passive and unpowered sensors which record attempted compromises to the authenticator dielet. The sensors shall also potentially detect other operations on the package assembly, e.g., soldering or de-soldering.
5) The SHIELD design must include inductive or RF communication and powering that allows contactless operation.
6) The dielet must have built-in resiliency against power-based component exploits or attacks.

The SHIELD methodology is explicitly designed to target the counterfeit IC problem and it's variants [16]. To ensure a full applicability/efficiency in the full range of anti-counterfeiting scenarios, additional design properties were suggested as follows:

7) Since re-writable storage devices present a vulnerability in the system as it is subject to copying/modifications, any re-writable storage included in the design must be carefully assessed for its security.
8) It is requested that the SHIELD dielet is isolated from the original design to avoid potential reliability/coupling issues.
9) For operational security reasons, the inductive or RF probe must be next to the dielet to directly communicate. The key shall never leave the dielet, and the key information on the server must never be released.
10) For practicality and wide-adoptability reasons, the SHIELD dielet and its relevant design components need to be extremely inexpensive to acquire, implement, and execute.
11) Lastly, the authentication complexity and protocol overhead must be pushed to the server side as much as possible so that the SHIELD dielet's physical size, power consumption, and cost is minimized.

In addition to the above requirements, the tester (verifier) of the SHIELD dielet(s) must be a commodity inexpensive appliance such as smart phone with an inductive or RF probe connected to it. A cryptographically secure challenge response protocol between the verifier and the prover (i.e., the SHIELD dielet) is required.

III. Classification of SHIELD security model, threats, and metrics

Our recent relevant work provides a systemization of knowledge for a number of important contemporary problems in the hardware security field [1]. The paper classifies hardware-based threats, countermeasures, and metrics to evaluate the effectiveness of the developed defenses. This provides a guide for researchers and practitioners to enable fair evaluation of the suggested methodologies and defenses.

In brief, our contemporary work analyzes the following threats [1]: *(1) Hardware trojans*, which are addition of malicious circuits or modification of existing circuits thereby making them vulnerable to attacks; *(2) IP piracy and IC overbuilding*, where an IP user or foundry illegally pirates the IP without the knowledge of the designer, or a malicious foundry builds more than the required number of IC and sells the excess ICs; *(3) Reverse engineering*, where an attacker can reverse the IC/IP design to his/her desired abstraction level; *(4) Side-channel analysis*, where an attacker can extract the secret information by exploiting a physical modality (power consumption, timing or electro-magnetic emission) of the hardware that executes the target application; *(5) Counterfeiting*, where an attacker illegally forges or imitates the original component/design.

For each of the threats, various models/assumptions have been taken by the different research projects/groups. The paper classifies the different threat model scenarios that are

used within the context of each attack, and based on the assumptions/models suggests a classification of the relevant countermeasures to each scenario. For each of the countermeasures, the relevant method or metric used for quantification of the effectiveness of the proposed methodology is discussed.

Figure 1 demonstrates a high-level abstraction of the new systemazation of knowledge provided in [1]. The leftmost column demonstrates the attack (abstracting the various scenarios pertinent to each attack class), while the middle and right columns show the applicable countermeasures and metrics respectively. Note that the description of the attack scenarios is also application-dependent, and shall be analyzed for each target application class[1].

A SHIELD provides a new hardware root-of-trust that could potentially address a number of the hardware-based threats discussed in [1]. In Figure 1, we use the solid bold-black (block) edges on the squares to show the attack and countermeasures that can be directly addressed/affected by the SHIELD approach. The attacks and countermeasures that could indirectly benefit from the SHIELD approach are demonstrated using the dashed bold-black block edges. The metrics would become relevant in the context of using an aspect of the SHIELD as a suggested countermeasure. Further research and investigations shall be directed towards a fair evaluation of each SHIELD design, and in particular evaluation of its attack resiliency for various application(s) and adversarial scenarios.

IV. Towards an open SHIELD evaluation

When AES and other crypto primitives were designed, security assessment was public. This made adopters comfortable with the developed primitives. This also yielded attacks, optimizations, improvements, and implementations that were public and well whetted. In the end, these crypto primitives are being widely used. With this context, an open security assessment of SHIELD should include (i) technologies that make up SHIELD, (ii) the SHIELD hardware dielet and (iii) the complete system that includes the SHIELD dielet and will yield appropriate security metrics and trust assessment procedures.

An open evaluation of the SHIELD technologies using the Embedded Systems Security Challenge (ESC) (https://esc.isis.poly.edu/) framework will be beneficial in assessing the strengths and weaknesses of this technology. ESC is a one of a kind capture the chip red team-blue team platform. The ESC blue teams submit an embedded system or a defense for evaluation. The ESC red teams identify vulnerabilities in the target system or defense and demonstrate how to exploit them. Following are two examples:

- In ESC 2012, participants adapted VLSI testing techniques to detect hardware Trojans in a design to emulate the following supply-chain scenario: Alice, a designer, sends her design to an off-shore foundry owned by Mallory. Mallory (or a rogue in her foundry) inserts Trojans in some of the chips. Alice tests each chip and classifies it as either infected or Trojan-free. Alice may use one or more of the following techniques: (1) Functional tests to activate Trojans and observe their effect at the output and (2) Side-channel analysis of delay, power, or other chip parameters and differentiate the impact of Trojans on these parameters even in the presence of process variations.
- In ESC 2013, the FANCI tool[18], developed to scan register transfer level code for hardware trojans, was the focus of the red team blue team assessment. This challenge emulated the following scenario: Consider an undercover agent working in an intellectual property (IP) design house that makes IPs. The IPs are delivered to a system-on-chip (SoC) design house, which produces chips that will be deployed in secret missions. The security team in the SoC design company believes that it has techniques to detect any malicious elements in the IPs supplied by your IP design house. The challenge is to design a malicious IP that circumvents the security team's hardening techniques, while meeting the functional specification of the SoC designer.

VLSI testing methods cannot assess trustworthiness of SHIELD. In VLSI testing, the fault models are known a priori and fault detection targets the faults. Trust assessment has orthogonal goals: attack models, attacker intention, and attacker capabilities are unknown a priori. Also, the attacker can adapt to the defenses. Trust assurance provided by VLSI testing is inadequate at best! Consequently, submitting SHIELD to an open, red-team-blue-team trust evaluation is the best way to uncover potential weaknesses and indirectly help strengthen the technology before it is widely deployed. ESC has validated over 10 detection techniques and yielded over 500 trust and hardware security related benchmarks all of which are open (www.trust-hub.org).

V. Conclusion

Outsourcing of the integrated circuits mass fabrication along with the long supply chain of production introduce multiple vulnerabilities in the electronic components and critical applications that rely on them. To enable supply chain security assurance, DARPA has recently issued a solicitation for the design of a SHIELD device, enclosed in a dielet. The SHIELD is required to provide a secure root-of-trust that is physically and inseparably attached to the host electronic component. This paper summarized the desiderata for the SHIELD dielet as envisioned by DARPA. The SHIELD solution was discussed in the context of a recent comprehensive systemization of knowledge and classification of the contemporary hardware security threats. The attacks and countermeasures relevant to the SHIELD approach were identified and the pertinent evaluation metrics were presented. We suggest the need for open evaluations and thorough security analysis of the SHIELD to ensure it is protected against several potential adversarial acts and to ease its wide adoption.

Acknowledgment

This work was supported in parts by an Office of Naval Research grant (ONR R17460) and a NSF grants to Rice University (CNS-1059416) and NYU-Poly (CNS-1059328).

[1] A full description of the classifications in 1 is outside the scope of this paper. We refer the interested readers to the source paper for more information.

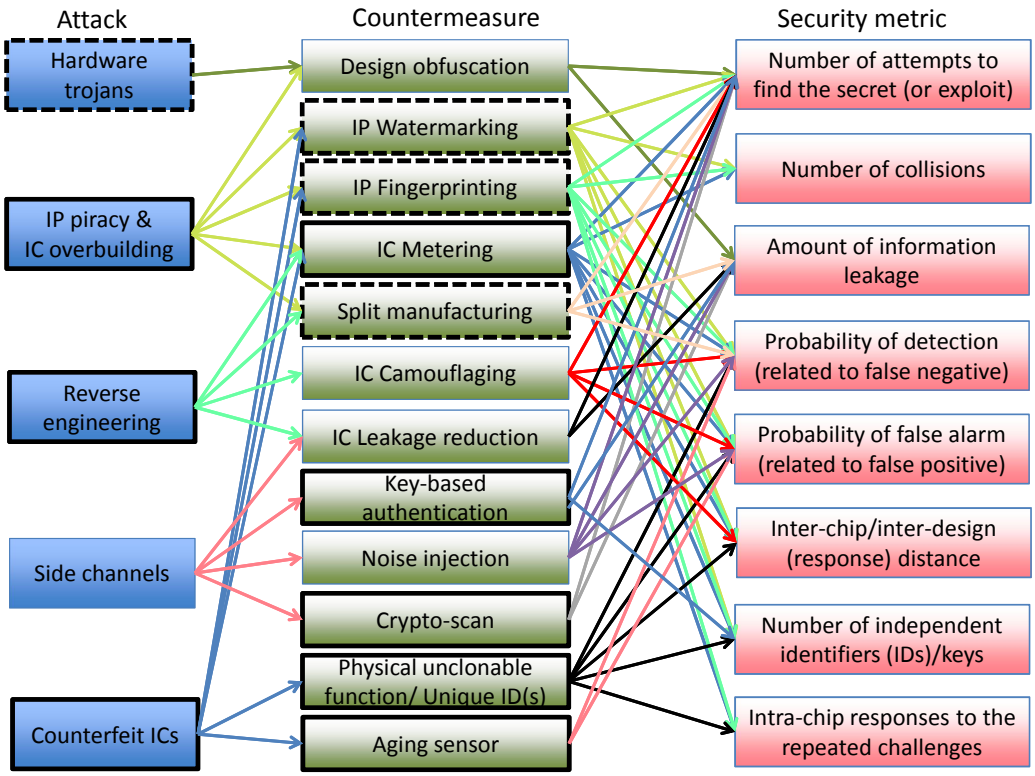

Fig. 1. A high-level view of the five classes of hardware attacks in [1] (left column), suggested countermeasures (middle column) and evaluation metrics (right column.) The blocks with bold black edges are relevant to the SHIELD, either directly (solid bold black) or indirectly (dashed bold black) lines.

REFERENCES

[1] M. Rostami, F. Koushanfar, and R. Karri, "A primer on hardware security: Threat models, metrics, and remedies," *Proceedings of the IEEE*, pp. 1–13, 2014, to appear.

[2] M. Tehranipoor and F. Koushanfar, "A survey of hardware trojan taxonomy and detection," *IEEE Design & Test of Computers*, vol. 27, no. 1, pp. 10–25, 2010.

[3] R. Karri, J. Rajendran, K. Rosenfeld, and M. Tehranipoor, "Trustworthy hardware: Identifying and classifying hardware trojans," *IEEE Computer*, vol. 43, no. 10, pp. 39–46, 2010.

[4] F. Koushanfar, G. Qu, and M. Potkonjak, "Intellectual property metering," *Information Hiding Workshop*, pp. 81–95, 2001 2001.

[5] Y. Alkabani and F. Koushanfar, "Active hardware metering for intellectual property protection and security," in *USENIX Security Symposium*, 2007, pp. 291–306.

[6] J. Roy, F. Koushanfar, and I. Markov, "EPIC: Ending Piracy of Integrated Circuits," *IEEE Computer*, vol. 43, no. 10, pp. 30–38, 2010.

[7] F. Koushanfar, "Provably secure active ic metering techniques for piracy avoidance and digital rights management," *IEEE Trans. on Information Forensics and Security*, vol. 7, no. 1, pp. 51–63, 2012.

[8] C. Sturton, M. Hicks, D. Wagner, and S. T. King, "Defeating uci: Building stealthy and malicious hardware," *IEEE Symposium on Security and Privacy*, pp. 64–77, 2011.

[9] B. Yang, K. Wu, and R. Karri, "Scan based side channel attack on dedicated hardware implementations of data encryption standard," in *Int. Test Conference*, 2004, pp. 339–344.

[10] ——, "Secure scan: A design-for-test architecture for crypto chips," *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, vol. 25, no. 10, pp. 2287–2293, 2006.

[11] Y. Alkabani, F. Koushanfar, N. Kiyavash, and M. Potkonjak, "Trusted Integrated Circuits: A Nondestructive Hidden Characteristics Extraction Approach," *Information Hiding*, pp. 102 – 117, 2008.

[12] B. Gassend, D. Clarke, M. Van Dijk, and S. Devadas, "Silicon physical random functions," in *Proc. of ACM Conf. on Computer and communications security*. ACM, 2002, pp. 148–160.

[13] U. Ruhrmair, S. Devadas, and F. Koushanfar, "Security based on physical unclonability and disorder," *Book Chapter in Introduction to Hardware Security and Trust*, 2011.

[14] F. Armknecht, R. Maes, A. Sadeghi, O.-X. Standaert, and C. Wachsmann, "A formalization of the security features of physical functions," in *IEEE Symp. on Security and Privacy*. IEEE, 2011, pp. 397–412.

[15] J. B. Wendt, F. Koushanfar, and M. Potkonjak, "Techniques for foundry identification," in *Proc IEEE/ACM Design Automation Conf.*, 2014.

[16] F. Koushanfar, S. Fazzari, C. McCants, W. Bryson, M. Sale, P. Song, and M. Potkonjak, "Can EDA Combat the Rise of Electronic Counterfeiting?" *IEEE/ACM Design Automation Conf*, pp. 133–138, 2012.

[17] Defense Advanced Research Projects Agency (DARPA), Microsystems Technology Office/MTO Broad Agency Announcement, "Supply chain hardware integrity for electronics defense (SHIELD)," 2014.

[18] A. Waksman, M. Suozzo, and S. Sethumadhavan, "Fanci: Identification of stealthy malicious logic using boolean functional analysis," in *Proc 2013 ACM SIGSAC Conf on Computer & Comm Security*.