Security and Vulnerability Implications of 3D ICs

Yang Xie, Chongxi Bao, Caleb Serafy, Tiantao Lu, Ankur Srivastava, *Senior Member, IEEE*, and Mark Tehranipoor, *Senior Member, IEEE*

Abstract—Physical limit of transistor miniaturization has driven chip design into the third dimension. 3D integration technology emerges as a viable option to improve chip performance and increase device density in a direction orthogonal to costly device scaling. As 3D integration is becoming a promising technology for next-generation chip design, recent years have seen a huge proliferation of research literature exploiting it from a security perspective. This paper presents a survey on the current state of 3D integration technology from a security perspective and summarizes its security opportunities and challenges. We report current research work on 3D integration based security in three major applications: supply chain attack prevention, side-channel attack mitigation, and trustworthy computing system design. The security advantages and opportunities of 3D integration in these security applications are highlighted. Besides, the paper discusses new vulnerabilities risen by 3D integration that require researchers' attention. Based on the survey result, we summarize the distinct characteristics of 3D ICs and investigate their impacts on security-aware 3D IC designs.

Index Terms—3D integration, hardware security, IP piracy, hardware trojan, side-channel attack, trustworthy computing system

1 Introduction

Technology scaling has allowed transistors to become smaller, faster and more power-efficient. However, device miniaturization has increased interconnect power and delay to such an extent that it is presenting a significant bottleneck to further performance and energy-efficiency gains. Moreover, it is expected in the near future that technology scaling will cease to reduce the cost per transistor [1]. These trends motivate the need for innovative approaches to reduce interconnect power/delay and at the same time increase transistor density, which has driven chip design into the third dimension. 3D integration is a promising technology to overcome many obstacles that have caused Moore's Law to slow down in recent years.

3D Integration technology expands circuit design into the third dimension by vertically stacking multiple functional device layers and interconnecting them using Through-Silicon-Vias (TSVs), as illustrated in Fig. 1. The vertical stacking structure is an attractive option for increasing transistor density. It also reduces interconnect wirelength hence scaling down power and delay. The reduction in interconnect wirelength can be leveraged by implementing a more highly connected architecture without increasing power or delay. Moreover, 3D integration allows separate layers to be fabricated using disparate materials and technologies. Heterogeneous integration optimizes existing System-on-Chip (SoC) designs by integrating components of different novel technologies into a single chip.

As 3D IC is becoming a promising technology for next-generation chip design, researchers have started to investigate it from a hardware security perspective. In general, hardware security research can be classified into three major categories: 1) preventing piracy, overbuilding, and malicious modification of a design during outsourced fabrication; 2) mitigating side-channel based secret information leakage; and 3) ensuring trustworthy operation of software through hardware-based security mechanisms such as trustworthy computer architectures and novel security primitives. While 3D integration is initially designed to improve chip performance, it has presented various potential in countering aforementioned security threats with its built-in security advantages:

- *Split fabrication:* With 3D integration, a designer can choose a portion of layers at his discretion and fabricate them in a trusted foundry while outsourcing the rest to untrusted foundries for low-cost fabrication. This split fabrication process prevents potential piracy, overbuilding or malicious modifications of a design during outsourced fabrication.
- *Stacking structure:* The stacking structure and high-density nature of 3D integration offer a natural defense for side-channel attacks as it adds significantly more complexity for an attacker to extract a meaningful signal from the complicated background noise. Moreover, reverse-engineering becomes challenging since hardware designs can be protected inside the firmly stacked substrates.
- *Heterogeneous integration for security:* With 3D heterogeneous integration, novel non-CMOS security primitives can be integrated with CMOS processor to achieve a comprehensive system with optimal security and performance.

3D integration not only boosts chip performance but also unlocks new opportunities to thwart security threats which

- *Y. Xie, C. Bao, C. Serafy, T. Lu, and A. Srivastava are with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742. E-mail: {yangxie, borisbcx, cserafy1, ttlu, ankurs}@umd.edu.*
- *M. Tehranipoor is with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611. E-mail: tehranipoor@ece.ufl.edu.*

Manuscript received 27 Sept. 2015; revised 18 Jan. 2016; accepted 22 Mar. 2016. Date of publication 5 Apr. 2016; date of current version 21 July 2016. Recommended for acceptance by S. Hu, Y. Jin, M.M. Tehranipoor, and K. Heffner.
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TMSCS.2016.2550460

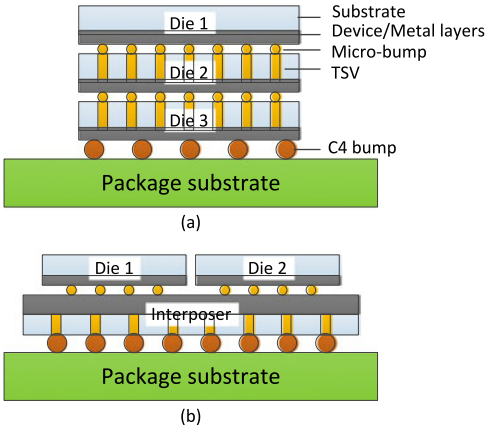


Fig. 1. Two common configurations of 3D ICs: (a) Stacked 3D IC and (b) Interposer-based 3D IC (2.5D IC).

occur in different phases of an IC's life cycle. But at the same time, it brings new reliability and security challenges for both design and fabrication of its own. The objective of this paper is to survey the current state of 3D integration based security applications and highlight potential security opportunities and challenges of this technology.

The rest of this paper is organized as follows. An overview of 3D integration technology is firstly provided, including its fabrication process (Section 2), architectures (Section 3), and its design flow and challenges (Section 4). Then, Section 5 highlights the security advantages of 3D IC and its opportunities in countering various security threats. Section 6 discusses potential security vulnerabilities in 3D ICs. Section 7 summarizes distinct characteristics of 3D ICs and investigate the security considerations in a security-aware 3D IC design. Finally, Section 8 concludes the paper.

2 3D IC Fabrication

3D integration is a technology that vertically integrates multiple layers of devices and metals to create a single high-performance chip, referred to as 3D IC. Currently, 3D IC can be fabricated in two ways: 1) die stacking based 3D fabrication and 2) monolithic 3D fabrication. *Die stacking based 3D fabrication* [2] utilizes conventional 2D IC fabrication process to fabricate each layer separately on different substrates and then interconnects them using Through-Silicon-Vias (TSV). On the contrary, *monolithic 3D fabrication* [3] grows multiple device layers vertically on the same substrate in a serial order, so it doesn't require alignment, thinning and bonding. Compared to the monolithic approach, die stacking approach can exploit existing fabrication processes and hence has lower fabrication cost and less time to market. Therefore, die stacking based 3D fabrication has received more attention from both academia and industry. Two key processes of the die stacking based 3D IC fabrication flow are TSV manufacturing and die stacking.

2.1 TSV Manufacturing

TSV is the most crucial element in die stacking based 3D integration. It can be used for signal communication, thermal conducting and power delivery. TSV is typically made of copper or tungsten. Its manufacturing process is similar to the process for a contact hole, except that TSV requires a much deeper hole that goes all the way through the substrate. In general, via-first and via-last are two mainstream TSV fabrication techniques [4]. The via-first approach etches into the silicon to form the TSV hole before processing the devices while the via-last approach does the opposite. Although two approaches differ in process details, they share a similar TSV formation process, which consists of via etching, dielectric isolation, via filling (metallization) and chemical-mechanical polishing (CMP) [5]. The performance and reliability of TSVs are related to their physical parameters, such as height, diameter and pitch size. A TSV example is of 5 μm diameter and 50 μm height at a 10 μm minimum pitch [6].

2.2 Die Stacking

Die stacking facilitates the integration of multiple homogeneous or heterogeneous dies that are designed and manufactured separately. Here we introduce the stacking process for two common configurations of 3D ICs: stacked 3D IC and interposer-based 3D IC, as shown in Fig. 1.

Stacked 3D IC integrates multiple dies by stacking and bonding them vertically, as shown in Fig. 1a. After substrate thinning and TSV manufacturing, multiple dies are aligned precisely and bonded vertically using Cu/CuSn micro-bumps. Various die stacking methods have been developed and in general, the main difference among these methods is related to the choice of stacking orientation [7]. The "face-to-face" approach bonds the via stubs in metal layers (face) of two dies directly, hence it doesn't require TSVs to connect two dies. The only TSVs that penetrate the substrate are those that are connected to external I/O pins. However, this approach only supports two-layer 3D integration. The alternative, the "face-to-back" approach bonds the metal layer (face) of one die to the TSVs in the substrate (back) of another die. Arbitrary layers of dies can be stacked vertically using the "face-to-back" approach, as long as the system meets its thermal and power constraints.

Interposer-based 3D IC, also known as 2.5D IC, is another configuration of 3D IC. It places multiple dies side-by-side and stacks them on a silicon interposer through fine-pitch micro-bumps. The structure of a 2.5D IC is shown in Fig. 1b. The interposer contains both horizontal chip-scale interconnect wires between dies as well as vertical interconnect TSVs to external I/O pins. The absence of TSVs in the dies of 2.5D IC makes it easier to design and fabricate than TSV-penetrated stacked 3D IC.

3 3D CPU Architecture

3D integration offers many new opportunities for high-performance CPU architectures. In the following sections, we discuss some examples of 3D CPU architectures that take advantage of the performance improvements in 3D ICs.

3.1 3D Memory on Chip

Heterogeneous integration involves the stacking of circuits fabricated in disparate manufacturing processes. Such an integration approach can provide massive bandwidth improvements between CPU and components that are traditionally fabricated off-chip. A well-known example of this is memory stacking on chip. Non-CMOS technologies such as

DRAM, phase-change RAM (PRAM) and magnetic RAM (MRAM) [8] can be stacked directly on top of a logic core and integrated using high-density TSVs, overcoming the well-known memory wall problem [9]. Not only do TSVs provide much lower latency than off-chip interconnects, but also provide high density interconnects, hence facilitating wider buses and parallel memory access using multiple memory controllers [9], [10]. Studies have shown that the performance improvements due to main memory stacking can be up to 2x [9], [11].

3.2 3D Network on Chip

Network on Chip (NoC) is a communication management system that utilizes novel network topologies and methodologies to interconnect a variety of functional intellectual property (IP) blocks and/or processing units such as processors, memory, and FPGA blocks. It's designed to overcome the bandwidth limitation of traditional hierarchal buses or the lack of scalability in a crossbar interconnect approach. Combining with 3D integration technology, 3D NoC has been shown to provide significant improvements in both latency, throughput and energy efficiency [12]. Due to the mismatch in dimensions between the vertical direction (hundreds of microns) and the planar directions (tens of millimeters), innovative 3D NoC topologies that are more highly connected in the vertical direction have been proposed [10], [12].

3.3 Fine-Grain 3D Integration

3D integration can be done at different granularities [13]. Coarse-grained 3D integration can be implemented at *core level*, such as the 3D memory on chip and 3D NoC models as mentioned in Sections 3.1 and 3.2. This approach could offer significant improvements to performance and power by removing the memory bandwidth wall, but does not take full advantage of the benefits of 3D ICs. A finer grained *functional block level* integration allows functional blocks to be distributed across multiple layers, but maintain each functional block as a 2D circuit. This can reduce intra-core wirelength and allow reduced clock period or power [13]. To take this idea even further, 3D integration at the *logic-gate level* (block folding) offers even more savings in power and delay. Block folding involves the implementation of an individual functional block across multiple layers, reducing intra-block delays and power. A recent study [14] of full-chip 3D design of a SPARC chip multiprocessors (CMP) showed that a 3D design using 2D functional blocks can reduce power by 14 percent compared to a baseline 2D design, however when block folding is applied this reduction in power becomes 20 percent. Even finer grained integration at the *transistor level* (e.g., separate layer for NMOS and PMOS) has been considered [13], but the ability to manufacture TSVs at the size and pitch required for such a scheme is yet to be realized. Moreover, the reliability and yield implications of such an approach are expected to be prohibitive [15].

4 3D IC Design Flow and Challenges

A typical 3D IC design flow is illustrated in Fig. 2. Given a design specification, *3D architectural design space exploration*

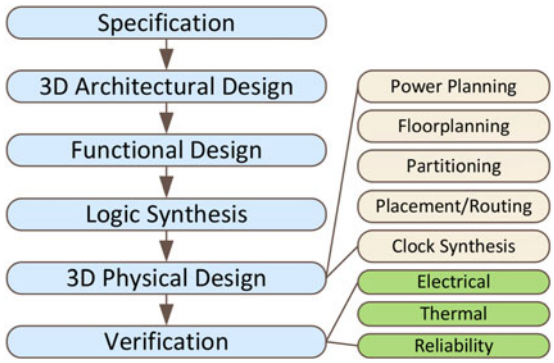


Fig. 2. Typical design flow for 3D ICs.

(DSE) is first performed to evaluate a multitude of design choices and choose the one that is optimal. Early estimation of the system's cost, performance, power, thermal and reliability characteristics is made to assist the architectural level decision making. After that, *functional design* (register-transfer-level abstraction) is performed to represent the logic flow between registers. The functionalities are converted into logic-gate level netlists in the subsequent *logic synthesis* step. This is followed by *3D physical design*, which is mainly composed of power/ground planning, floorplanning, partitioning, placement/routing, and clock network synthesis. It converts the gate-level netlists into physical layouts. The next step is *design verification*, where layout shapes and electrical parameters are extracted from the physical layout to verify the design rules and functionality. It also ensures that electrical performance (timing, power, IR-drop etc.), thermal behavior, and reliability can meet the design specification.

The increased device density due to multi-layer stacking and the introduction of TSVs pose new concerns and challenges in 3D IC design. Here we summarize five major design challenges in 3D ICs:

Thermal management: The thermal challenge in 3D ICs has two main sources: increased power flux due to stacking and interlayer thermal resistance due to interlayer oxide. Traditional air-cooling is a poor choice for 3D stacking as the cooling capacity can not scale with the number of layers. Various methods of thermally aware design have been proposed to overcome the thermal challenge [16], [17], [18], [19]. Another promising solution is embedded micro-fluidic (MF) cooling, which provides localized cooling to each layer of the stack and can unlock performance increases over 2x [10]. Runtime schemes also fill in the gaps to ensure thermal feasibility in 3D CPUs [20], [21], [22], [23], [24].

Power delivery: Increased power flux also puts increased stress on the power delivery network (PDN). Vertical interconnects change the electrical properties of the power network. Delivering reliable power to the layers far from the off-chip power pins is significantly difficult. Past works have characterized the 3D PDN with respect to number of power TSVs, number of layers, power/performance trade-offs and optimal placement/size of decoupling capacitance and on-chip voltage regulators [25], [26], [27], [28].

Reliability: TSV introduces new failure modes such as TSV electro-migration and stress-migration [29], thermal cycling [30], and stress-induced material fracture [31]. These new failure modes and other common device related failure

modes [32] such as hot carrier injection (HCI) and negative bias temperature instability (NBTI) are especially worsened in an elevated temperature, which is often the case in 3D ICs. The reliability issues in 3D ICs can be managed during design time and run time. During design-time, redundant TSVs insertion [33] and TSV placement [34] are two commonly used techniques. For run-time approach, task scheduling [35] and task migration [36] can effectively mitigate 3D IC's reliability degradation.

Signal integrity: Signal integrity issues can come about due to cross coupling between TSVs and the silicon substrate in 3D ICs. Studies have shown that unlike planar wires, increasing TSV pitch is not an efficient method for dealing with the signal integrity issue in 3D ICs [37]. Promising solutions to this problem are TSV shielding [37], [38], [39], [40], [41] and differential TSV pairs [42].

Clock synthesis: Clock tree synthesis is the design of a clock network that connects a clock source to all sequential logics such that the clock wirelength, clock skew/slew and clock power are minimized. Clock TSVs complicate the design of 3D ICs by introducing TSV capacitance, TSV-induced thermal-mechanical stress and TSV failure. State-of-the-art 3D clock tree synthesis algorithms include partition-based approach [43] and greedy method [44]. Besides, the impact of the TSV-induced stress on timing corners was considered in [45] and clock gating techniques were applied to 3D IC to minimize clock power [46].

These research efforts targeting different 3D IC design challenges ensure the feasibility and reliability of 3D integration technology.

5 Security Opportunities in 3D ICs

As 3D integration is becoming a promising technology for next-generation chip design, researchers have started to investigate it from a security perspective. 3D integration has been shown to possess various built-in security advantages in countering different security threats in ICs. These threats take place in different stages of an IC's life cycle, from design, fabrication to post-deployment. In this section, we focus on three security applications of 3D IC technology: supply chain attack prevention, side-channel attack mitigation and trustworthy computing system design. For each application, we describe the security threats, investigate the advantages of 3D IC in thwarting the attacks, and summarize various previously proposed 3D IC based countermeasure techniques.

5.1 Supply Chain Attack Prevention

Nowadays, IC designs are increasingly outsourced to an offshore fabrication foundry due to the increasing complexity of modern IC designs and the huge capital expenditure for developing an advanced semiconductor foundry [47]. In order to access advanced semiconductor technology at a lower cost, most IC design companies that once possessed their own foundries are now adopting a fab-less model: they concentrate their resources and efforts on IC designs while outsourcing the fabrication. Although such model is cost-effective, it poses new security threats on the outsourced designs since the offshore foundry might not be trustworthy. Without close monitoring and direct control, the outsourced designs are vulnerable to various attacks. The foundry can reverse-engineer a GDSII layout file to obtain its gate-level netlist (hardware IP), or it can overproduce the IC and sell illegal copies into the market, which are referred to as IP piracy and IC overbuilding [48], [49]. The economic loss due to piracy and overproduction was estimated up to US $4 billion per year in 2008 and was expected to grow significantly since then [50]. In addition, a malicious foundry can intentionally modify the layout design in order to produce a backdoor [51] or undermine the reliability [52] of the chip, referred to as hardware Trojan insertion. These attacks (also known as *supply chain attacks*) pose not only an economic risk to commercial IC design companies, but also security threats for sensitive electronic systems.

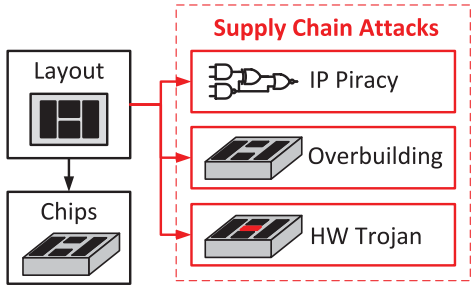


Fig. 3. Supply chain attacks.

5.1.1 Attack Model

The supply chain attack model assumes that the attacker is an untrusted foundry that has access to the layout files of a circuit design provided by a fab-less design company. These outsourced layout files are assumed to be correctly and securely designed using trusted Electronic Design Automation (EDA) tools for synthesis, partitioning, placement and routing. The attacker has the ability to reverse-engineer the layouts to obtain their gate-level netlists. Moreover, he is also capable of making malicious modifications to the layout files. As shown in Fig. 3, the supply chain attack consists of three types of attacks based on different attacker's goals:

1) *IP piracy:* An attacker can reverse-engineer the provided layout files to obtain their gate-level netlist (hardware IPs) using state-of-the-art reverse-engineering technique [53]. He then can claim the ownership of the pirated IPs and use them in his own designs or sell them to other IC design companies in order to gain profit.
2) *IC overbuilding:* The attacker can overbuild the ICs with the already available layout masks and then sell the illegal copies into the market. To reduce cost, these overbuilt ICs might not be subject to a complete testing process. As a result, some unauthorized and low-quality ICs (also known as *counterfeit ICs* [54]) may end up being packaged and sold to the market, which renders both economic and reputation loss to the design company.
3) *Hardware Trojan (HT) insertion:* In general, there are two types of HTs: *functional HTs* and *parametric HTs*. Functional HTs intend to modify or disable existing functionalities of a circuit. The attacker can modify one or more specific gates and/or wires in

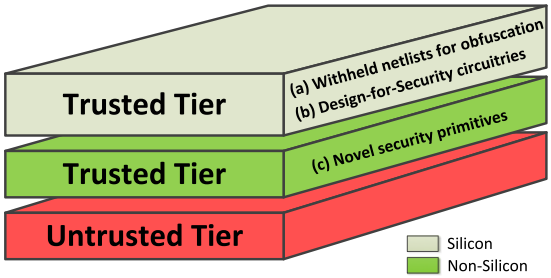


Fig. 4. The security applications of the trusted tier in a split-fabricated 3D IC: (a) Withholding partial netlist for netlist obfuscation and functionality obfuscation (Section 5.1). (b) Preventing design-for-security circuitries (HW Trojan detection circuitry and introspection circuitry) from malicious modification (Sections 5.1 and 5.3). (c) Heterogeneous integration of novel security primitives (Sections 5.2 and 5.4.2).

order to achieve certain purposeful attacks such as privilege level escalation and secret key recovery, referred to as *targeted HT* insertion. Parametric HTs are implemented by modifying the fabrication process parameters that can undermine the reliability or performance of the ICs. The detection of HTs through conventional IC testing is difficult since the HTs are normally designed to be triggered under extremely rare conditions (in the case of functional HTs) [55] or triggered after a relatively long time such as a few year (in the case of parametric HTs) [52].

5.1.2 New Opportunities with 3D Integration

In 3D integration, multiple functional layers can be fabricated independently on separate substrates and then integrated together into a single chip. This fabrication process offers inherent support for split fabrication, where different layers can be fabricated in different foundries. A designer can choose a portion of the circuit at his discretion and manufacture it in a trusted foundry for security while manufacturing the rest in an untrusted foundry for state-of-the-art fabrication technology. Because some information will be hidden from the untrusted foundry, 3D split fabrication can prevent the supply chain attacks such as piracy, overbuilding, and hardware Trojans.

Two split fabrication strategies of 3D IC have been presented in [56], [57]. In one embodiment, some active layers are fabricated in a trusted foundry, referred to as *trusted tier* while others are fabricated in an untrusted foundry, referred to as *untrusted tier*, as shown in Fig. 4. With that, the IC designs in trusted tier are not directly accessible to the untrusted foundry and hence they are protected from the supply chain attacks. In another embodiment, all active layers are outsourced to offshore foundries and then securely bonded in a trusted foundry. By doing so, the vertical connections between layers are kept secret. Although the offshore foundry can reverse-engineer the layout of each layer, the resultant incomplete netlist (lacking the inter-layer connections) is incomprehensible if a design is intelligently partitioned into different layers in an obfuscated manner. The second split-fabrication embodiment also applies to the interposer-based 3D IC (2.5D IC) technology [58], [59], [60], a cost-effective transition technology to 3D IC as discussed in Section 2.2. The silicon interposer of a 2.5D IC can be fabricated in the trusted foundry as the trusted tier while the dies are outsourced to offshore foundries as the untrusted tier. The final integration is implemented in the trusted foundry to maintain the secrecy of interconnections in the interposer.

In previous research literature, the "hidden" portion of design in the trusted tier is decided based on two considerations:

Firstly, the trusted tier can withhold a portion of original wires and/or gates that can maximally obfuscate the exposed portion of netlist in the untrusted tier. Without the complete knowledge of the original netlist, it's difficult for an attacker to reverse-engineer the ICs, or identify a specific location for HT insertion. To defend IP piracy and overbuilding, Xie et al. [58] proposed a security-aware physical design flow for 2.5D IC technology, which consists of a secure partitioning phases and a secure placement phase. The secure partitioning generates a bi-partitioning such that the corresponding cut-wires (which will be "hidden" in the interposer) have high controllability and observability in order to better obfuscate the functionality. The secure placement generates obfuscated chip layouts that can withstand the proximity attack [61], an attack that infers the hidden connection based on the physical proximity of two gates/pins. Compare to a conventional 2.5D design flow, their approach can achieve 3.87x functionality obfuscation (in terms of Hamming distance) and can mitigate the proximity attack. The obfuscation due to 2.5D split fabrication can also prevent the insertion of targeted HTs, i.e., malicious modification to a specific target gate. Imerson et al. [59] proposed a heuristic wire-lifting algorithm for 2.5D split fabrication to obfuscate the untrusted tier. By lifting and "hiding" sufficient number of wires into the interposer, the proposed technique ensures that every gate in the original netlist can be mapped to at least k indistinguishable gates in the incomplete netlist of the untrusted tier, thereby increasing the attacker's difficulty in identifying the target gate to attack.

Secondly, the trusted tier can obfuscate or conceal various security-critical circuitries such as HT detection sensors in order to protect them from being tampered or removed by the attacker. Recent years have seen a huge proliferation of hardware Trojan detection research based on functionality verification [62], side-channel signatures [63], [64], [65], built-in-self-authentication (BISA) [66] and so on. Most of these techniques require additional circuits to assist in Trojan activation and/or detection, including dummy flip-flops, sensors and authentication circuits, which are referred to as *design-for-security (DfS)* circuitries. However, these DfS circuitries may also be tampered or bypassed, which undermines the system's security. With 3D split fabrication, the DfS circuitries can be placed in the trusted tier of a 3D IC and fabricated in a trusted foundry. Narasimhan et al. [64] proposed to use on-chip transient current sensors to detect hardware Trojan and use 3D split fabrication to hide all the current monitors in the trusted tier. It prevents possible malicious modifications that can nullify the sensors, or upsize/downsize the sensors to undermine their measurement accuracy. Bilzor [67] proposed a similar technique that utilizes the trusted tier to conceal an execution monitor, which can detect specification violations of a target processor due to hardware Trojan. To reduce fabrication costs and brings down the semiconductor technology requirement for the trusted foundry, 2.5D technology can be utilized to conceal critical

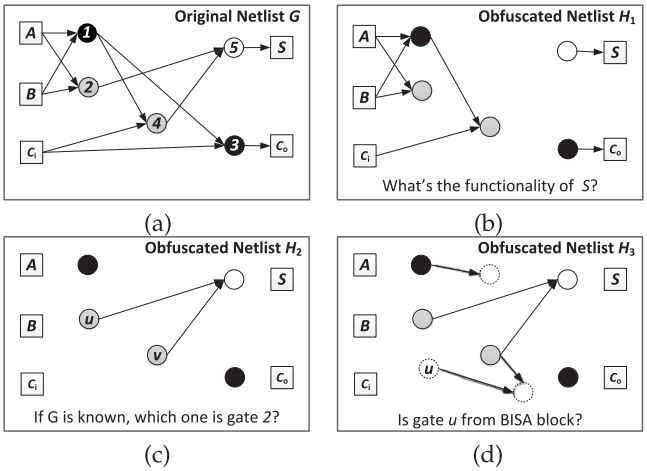


Fig. 5. 2.5D Split fabrication based obfuscations: (a) An original netlist of a full adder. Different colors indicate different gate types. (b) Functionality obfuscation to prevent IP piracy [58]. (c) Netlist obfuscation to prevent targeted HT insertion [59]. (d) Security-critical netlist (e.g., build-in-selt-authentication (BISA) circuits) obfuscation to prevent design-for-security circuitries from being identified or removed [60]. Dash circles are additional BISA gates.

wires of the DfS circuitries (instead of the complete circuitry). In [60], Xiao et al.proposed an obfuscated BISA (OBISA) technique that conceals critical wires of BISA circuitry [66] into the trusted tier. BISA block is a hardware authentication circuitry which guarantees that when it's modified by HTs, faults can be generated and observed. However, the BISA blocks consist of special circuit structures that are distinguishable from the original functional blocks and hence are subjected to removal by attackers. To make these structures less distinguishable, the author proposed to hide some critical wires of the BISA blocks to the trusted tier using split fabrication and construct an obfuscated BISA block.

5.1.3 Summary and Discussion

2.5D/3D IC technology provides an effective approach to hide partial circuitry into a trusted tier that's fabricated in a trusted foundry, which increases the burden of supply chain attacks (as illustrated in Fig. 5). A secure split-fabrication enhanced 2.5D/3D IC design flow consists of four phases: logic synthesis, netlist partitioning (wire and/or gate lifting), placement and routing. The core of the design flow is partitioning, which determines the secret information hidden from the attacker. The other three phases also have a significant impact on the security level of the split fabrication, i.e., they affect the difficulty of inferring the hidden information. Overall, it requires a comprehensive analysis and optimization to obtain a secure 2.5D/3D IC design flow to prevent the supply chain attacks.

Notice that the split fabrication strategy can also be applied to conventional 2D IC technology [61], [68], [69], [70]. 2D IC based split fabrication splits a 2D IC into a Front-End-Of-Line layer (FEOL) that contains active devices and lower metal layers, and a Back-End-Of-Line (BEOL) layer that contains higher metal layers. The FEOL layer is outsourced to an untrusted foundry for advanced fabrication technology while the fabrication of BEOL layer and final integration are securely implemented in a trusted foundry. Thus, interconnect wires in BEOL layer of a split 2D IC are

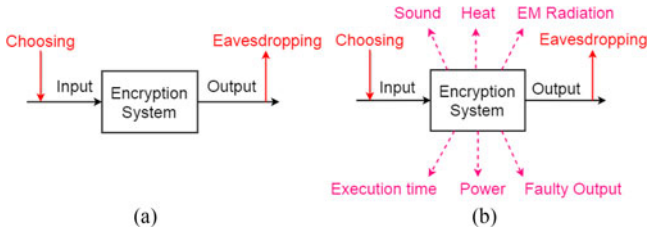


Fig. 6. (a) Black-box model and (b) side-channel attack model.

kept secret from the untrusted foundry. 2D split fabrication requires a new design and fabrication process [69]. The determination of which metal layer to split is a trade-off between security and fabrication cost. Although a low-layer split 2D IC (e.g., split after M1) offers better security since it hides almost all metal wires in the BEOL layer [68], it has dense connections of small pitch size between the trusted BEOL layer and the untrusted FEOL layer, which requires more precise alignment and integration techniques and hence leads to higher fabrication cost [60].

5.2 Side-Channel Attack Mitigation

While some attacks focus on attacking/modifying the hardware itself, many attacks try to leak valuable information from the system. One example of such valuable information is the secret key in an encryption algorithm. Nowadays encryption is used ubiquitously for the purpose of secure data transfer, identity authentication, etc. It is essential to many financial or military systems. Due to the nature of these systems, leaking the key is disastrous. Modern encryption systems are provably secure against brute-force attacks since it's computationally impossible to try all possible values of the key until the correct one is found. However, it is far from safe yet since side-channel attacks, where information from the physical implementations of the encryption methods is exploited to deduce the key, are still possible. Characteristics such as running time [71], [72], power consumption [73], [74], and electromagnetic (EM) emission [75], [76], [77] have all been utilized to recover the secret key from various encryption methods. In this section, we will go over different side-channel attacks and then investigate what opportunities 3D integration can offer to mitigate these attacks.

5.2.1 Attack Model

Similar to traditional cryptanalysis, we assume that the adversary has a complete description of the encryption protocol and the only thing he is trying to deduce is the secret key. As mentioned previously, almost all encryption algorithms provide security against brute-force attacks, where the adversary only has black-box access (shown in Fig. 6a) to the encryption system (i.e., the adversary can only observe the output of the system with the input chosen at his discretion). However, in side-channel attacks, the attacker has access to some side-channels in addition to only inputs and outputs. Side-channels are defined as the unintended output channels from the physical implementation of the encryption algorithm [78]. As shown in Fig. 6b, these side-channels include execution time, power consumption, electromagnetic radiation, sound, visible light, heat, faulty output, etc. Among these side-channels, execution time, power consumption and faulty output are the most exploited and thus

most dangerous ones. The rest of this section focuses on these three types of side-channel attacks.

Timing attack is the first studied and most widely applied kind of side-channel attacks. Kocher [79] pioneered the work in the timing attacks. Timing attacks are based on the fact that encryption methods are performed in non-constant time due to performance optimizations. The run time of one encryption usually depends on the key to some extent. For example, consider the square and multiply algorithm (shown in Algorithm 1) which is widely used in various encryption methods. Depending on the value of each bit in the secret key s, $z = z * y \bmod n$ may or may not be executed. This leads to variation in execution time. If the attacker somehow can measure the time it takes to perform each iteration of the loop, then the secret key s is leaked.

Another source of the variation in execution time comes from the cache behavior. The point is that a cache-miss takes much more time than a cache-hit. In cache-based timing side-channel attacks, the attacker measures some timing information and relates that to cache-access patterns. Since the cache-access patterns depend on the secret key, the key is eventually leaked from the measured timing information. More details can be found in [72], [80].

Power consumption attacks are based on the following observation: the amount of power consumed by a device is influenced by the data being processed [81]. For example, in Algorithm 1, if $s_i = 1$, then $z = z * y \bmod n$ will be executed, resulting in a larger power consumption than $s_i = 0$. Two types of power-based side-channel attacks have been proposed: simple power analysis (SPA) and differential power analysis (DPA). In SPA attacks, the attacker directly observes a system's power consumption and features such as AES rounds may be identified. In DPA attacks, a set of encryptions under the same secret key is performed and the power trace is collected. Then, a statistical method is applied to correlate the measurements to the secret key. Given enough traces, extreme small correlations can be identified and the key can eventually be deduced [73].

Faulty output is another side-channel that adversaries have widely exploited in the *fault injection attacks*. Fault injection attacks can be divided into two phases: fault injection and fault analysis [82]. In fault injection phase, a fault (e.g., flip of one bit in one register) is injected into the system using EM radiation, light illumination, focused ion beam, etc. Note that besides intentionally injected faults, unintentional flaws of a hardware system may also occur, which consist of soft errors, dynamic timing errors and hard errors [83]. Soft errors are caused by high-energy particles such as alpha particles and high-energy neutrons from cosmic radiation. Dynamic timing errors are delay violations that are caused by the fabrication process variation of devices or working condition variations such as temperature and supply voltage noise. Hard errors are due to the hardware aging rendered permanently faulty. Although these faults are not introduced by malicious attackers, they inevitably affect a system's reliability and may be exploited by the adversary. Therefore, we also include these types of errors in our fault injection attack model. In the fault analysis phase, the faulty output is analyzed and the secret key is deduced. Consider Algorithm 1 where the adversary injected a stuck-at-one fault into the register holding s at the ith iteration. If the encryption result is correct, the adversary knows that $s_i = 1$, otherwise $s_i = 0$. More details can be found in [82], [84], [85].

Algorithm 1. Square and Multiply: An Algorithm for Side-Channel Attack Illustration

Input: m, n, s
Output: $S = m^s \bmod n$
  $z = 1, y = m$;
  $t \leftarrow$ number of bits in the binary representation of s
  for $i = 1$ to t **do**
    if $s_{i-1} = 1$ **then**
      $z = z * y \bmod n$
    end if
    $y = y * y \bmod n$
  end for
  $S = z$

5.2.2 New Opportunities with 3D Integration

With heterogeneous integration, stacking structure, high bandwidth and shortened wire length, 3D integration not only boosts system performance, but also unlocks new opportunities to mitigate the side-channel attacks.

The stacking structure of 3D integration offers a natural defense against fault injection attacks. As mentioned in Section 5.2.1, the first phase of fault injection attacks is to inject faults into the device using light illumination, focused ion beam, etc. By placing the vulnerable tier (e.g., memory) in the lower stack, the stack above it will act as a shield, which may protect the vulnerable tier from being influenced by fault injection techniques. For instance, [86] utilized 3D technology to protect security-sensitive hardware from high-energy partials by concealing them in the inner layers of a 3D chip. Furthermore, the stacking structure enables system designers to stack a monitor tier on top of the original system. The monitor tier may watch the system for any intended faults. For example, in [83], the authors proposed to stack a redundant processor as a checker on top of a processor using 3D technology to detect random faults in the system under protection. The stacking structure is also a potential mitigation of the power, EM and heat-based side-channel attacks [87]. Since more elements are integrated on chip, the side-channel will become noisy. It will be much harder for the attacker to capture the targeted signal. In other words, the signal-to-noise ratio for the attacker will be much lower. This requires the attacker to obtain significantly more traces, making the attack much harder, even impossible.

Heterogeneous integration is another benefit enabled by 3D integration. As mentioned in Section 3.1, integrating different types of memories into the system is beneficial for performance. In addition to that, it may also increase system security. Those types of non-CMOS memory that are resilient to differential power attacks or fault injection attacks may now be integrated into the system, which adds resiliency to fault injection attacks. In [88], the authors propose to integrate resistive RAM (RRAM) into the system to mitigate DPA attacks. The authors propose an architecture with crypto-coprocessors on a dedicated CMOS layer and the associated memory on RRAM layer. The RRAM layer can dynamically reconfigure into a memory or sensing elements

and these two cannot be easily distinguished from the power consumption. This obscures the power signals and thus mitigates DPA attacks. RRAM's high susceptibility to voltage variations provides an added advantage for the DPA attacks since it will add more noise to the power side-channel leakage of the system.

Shortened wire-length offered by 3D integration can also mitigate some cache-based timing side-channel attacks. As mentioned in Section 5.2.1, the basic idea of these attacks is that cache-misses and cache-hits take dramatically different amount of time. Utilizing the memory-on-chip architecture as introduced in Section 3.1, the latency for accessing lower level of cache will be significantly decreased. Thus, the discrepancy between a cache-miss and a cache-hit will be much smaller, which lowers the timing side-channel leakage in these attacks. Shortened wire-length and high bandwidth also make several countermeasures possible because the performance overhead is smaller compared to a 2D setting. In [89], a 3D IC based cache random-eviction technique was proposed to mitigate cache-based timing side-channel attacks. The authors concluded that with 3D integration, the performance overhead for implementing a secure cache eviction policy is reduced from 25 percent (for 2D CPU and off-chip memory architecture) to about 0 percent (for 3D memory-on-chip architecture).

5.2.3 Summary and Discussion

Side-channel attacks have been demonstrated to be very powerful against many security-critical systems and various countermeasures have been proposed. Although some effort has been made to utilize the characteristics of 3D integration to mitigate the side-channel attacks, more effort is needed to fully exploit the security advantages it can potentially offer. We highlight some of these security advantages as follows. The stacking structure and high-density nature of 3D integration offer a natural defense for side-channel attacks as it adds significantly more complexity for the attacker to extract a meaningful signal from the background noise. The heterogeneous integration will enable the designer to investigate a broader range of security primitives and integrate them into the system to defend against side-channel attacks. Moreover, the performance overhead of previously proposed countermeasures may be significantly lowered due to shortened wire-length and increased bandwidth of 3D integration.

5.3 Trustworthy Computing Systems

One of many security challenges in computing systems is to ensure that software and their data stored on the hardware platform are securely allocated, managed and executed. Nowadays, secret data such as authentication keys and privacy information are collected, processed and exchanged among different devices, which are attractive targets to malicious software based attacks such as unauthorized memory access and malicious modification of execution environment. Conventional computing systems normally rely on their operating system (OS) or hypervisor to monitor and allocate time-share resources (e.g., CPU, memory and buses) to different software from multiple users with different privilege levels. However, modern OS kernels are often complex and have tens of millions of lines of code [90], which makes it practically impossible to be perfectly designed. Security vulnerabilities in OS kernels have been continuously revealed. For example, 133 Linux OS vulnerabilities have been discovered in 2014 [91]. Malicious software can exploit these vulnerabilities and pose serious security threats such as incorrect code execution, memory corruption and privilege escalation. To mitigate these threats, various software-based and hardware-based protection schemes have been proposed. Software-based approaches [92], [93] utilize a small-sized, privileged micro-hypervisor to monitor and provide isolation for sensitive software. On the other hand, hardware-based approaches such as secure architectures have been proposed to provide a secure execution environment to protect a program in the presence of external malicious software attacks [94], [95], [96], [97]. Besides, some security functions of OS kernels such as cryptographic primitives can be implemented in hardware [98]. These hardware approaches, although provide better security guarantees, inevitably increase the design and fabrication cost and might undermine the system's performance. Therefore, a low overhead hardware support to alleviate the vulnerabilities of OS kernel is essential for a trustworthy computing system.

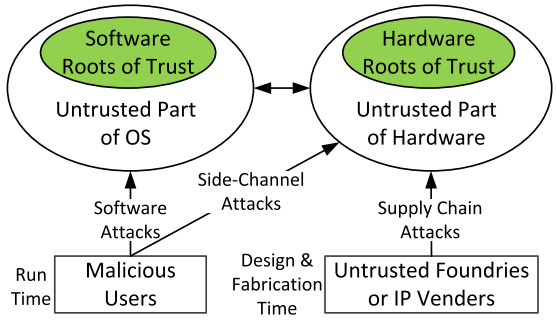


Fig. 7. Attack model for a computing system in design time, fabrication time and run time including supply chain attacks (Section 5.1), side-channel attacks (Section 5.2), and software attacks (Section 5.3).

5.3.1 Attack Model

The security vulnerabilities of computing systems addressed in this section mainly focus on malicious software that aims at exploiting vulnerabilities in the OS or hypervisor and compromising its isolated environment so as to audit, steal or corrupt critical information in another security-sensitive software. Specifically, this type of attacks can be unauthorized memory accesses, privilege escalation or any malicious modifications to the execution environment. A comprehensive attack model for a computing system in design time, fabrication time and run time is shown in Fig. 7.

5.3.2 New Opportunities with 3D Integration

3D IC technology offers various security and performance advantages in building a trustworthy computing system.

Firstly, the split fabrication ability of 3D IC technology can guarantee the integrity of a hardware root of trust, which is the security foundation of a trustworthy computing system. The hardware root of trust consists of hardware components that are inherently trusted. Every computer system is built with multiple levels of abstraction. Generally, higher layers must trust lower layers, and the hardware root of trust is composed of fundamental hardware components where the chain of trust is built on. For example, a system's root keys

shall be kept secret in hardware and the most fundamental security functions shall also be implemented as ROM code or custom hardware to ensure their integrity [94]. However, with the emergence of hardware Trojans introduced by untrusted third-party IP vendors and especially untrusted offshore foundries, the hardware root of trust might be compromised. 3D IC technology can help regaining this hardware root of trust. As described in Section 5.1, different layers of a 3D IC can be fabricated in different foundries and finally integrated by a trusted foundry. In order to prevent malicious modifications, the hardware root of trust can be designed by a trusted design team, and allocated to a trusted tier of 3D IC that's fabricated in a trusted foundry. Although the hardware root of trust can be achieved using an off-chip security module such as the Trusted Platform Module (TPM) [98], the exposed inter-chip connections make it vulnerable to probing or tampering attacks as reported in [99], [100]. On the contrary, 3D ICs are less susceptible to tampering or probing since the internal signals across layers are well concealed inside the chip. Delayering or prying the 3D IC to access the internal wires without destroying the chip is virtually impossible.

Secondly, the trusted tier can serve as a customized introspection layer that can ensure secure operations of software running in the untrusted tier. The applications of 3D IC in building trustworthy computing systems have been explored in various research [57], [87], [101], [102], [103]. In these research works, a 3D IC is constructed with two active layers: an untrusted *computation plane* and a trusted *control plane*, which are vertically stacked and connected using a face-to-face 3D integration technology. The computation plane contains a commercial high-performance processor and is outsourced to an untrusted offshore foundry for fabrication. On the other hand, the control plane which functions as an introspection layer is fabricated in a trusted foundry. The control plane can be designed to monitor and detect security-related policy violations [101], isolate and protect shared hardware resources such as memory and shared-buses [102], block unintended control and data flows such as malicious auditing between cores [102], and perform cryptographic functions [87].

Lastly, compared with 2D technology, 3D IC can significantly reduce the performance cost for hardware-based security approaches. Conventional 2D IC based off-chip and on-chip implementations of security modules will both result in a large performance overhead. 2D off-chip implementation such as TPM requires the utilization of off-chip interconnect buses, which have a high delay and power consumption. 2D on-chip implementation that inserts security modules inside a chip has larger performance overhead than 3D implementation since the former requires longer interconnect wires from the security module to other modules across the whole chip. Moreover, on-chip is not an economical solution for a large segment of customers who don't necessarily need the high-assurance hardware-based security modules. With 3D IC, the performance overhead can be alleviated. The vertical interconnection across layers substantially reduces the interconnect wirelength, and hence can reduce the overall delay and power consumptions. Besides, the control plane can be made optional depending on the security specification of the product [57]. It offers more flexibility than 2D on-chip implementation and provides an economical solution for system integrators to integrate critical security functions to a small portion of products for special interest groups.

5.3.3 Summary and Discussion

The split fabrication ability and the performance advantages of 3D IC ensure the security and performance of the hardware root of trust, which makes it a favorable and efficient hardware candidate for building a trustworthy computing system. Previous research has proposed to use 3D integration to stack a control plane on a commercial CPU (a computation plane) to monitor and control its behavior. In order to guarantee that the control plane can effectively function without being compromised, it's important to ensure two basic conditions: a) the security and functionality of the control plane are not dependent on the computation plane, and b) the computation plane should not bypass the monitoring and control of the control plane. Huffmire et al. [103] analyzed the security constraints for the design of control plane and computation plane from different perspectives. For example, it stated that the control plane shall function independently without the need of any functions from the computation plane. Also, global networks such as power and clock shall be ensured to supply reliable power and synchronization signals to the control plane to prevent malicious tampering. Moreover, the computation plane shall not be intentionally modified in order to mitigate the monitoring and control from the control plane. To maintain the effectiveness of the control plane, it's important to detect and remove hardware Trojans in the computation plane inserted by untrusted foundries or third-party IP vendors. Potential measures for hardware Trojan detection have been discussed in Section 5.1.

5.4 Other Security Opportunities

5.4.1 Reverse-Engineering Prevention

IC Reverse-engineering (RE) is the process of analyzing an ICs internal structures and connections in order to determine how it is designed and how it operates. It's usually performed by an end-user who can obtain the IC from the market. A typical RE flow has five major steps including *decapsulation*, *delayering*, *imaging*, *annotation* and *schematic creation*, as described in [53].

It's believed that the stacking structure of 3D IC is physically more difficult to be reverse-engineered [56], [57] since it conceals and protects valuable hardware IPs inside firmly stacked substrates. Delayering the 3D IC becomes challenging since the thick and tough substrates are difficult to be etched evenly [56]. Microscopy and imaging will be complicated by the existence of extra layers, bonding materials, and many layers of vertical and horizontal interconnections. Signal probing for functional analysis will be difficult because all active devices and internal wires can be encapsulated soundly inside substrates. The only exposure to the adversary is the primary I/O pins. This significantly limits the adversary's ability to probe inside the chip. Overall, the 3D stacking structure of multiple substrates increases the burden for each stage of state-of-the-art IC reverse-engineering technique [53], hence preventing the attacker from obtaining the original gate-level netlist.

5.4.2 Heterogeneous Integration of Security Primitives

3D heterogeneous integration can form highly integrated systems by vertically stacking and connecting multiple layers of various materials, technologies, and functional components [104]. The heterogeneous integration enables the integration of novel nano-technology based security primitives with a CMOS processor. In recent years, a huge proliferation of *physical unclonable function (PUF)* research is based on novel nano-technologies. These include PUF designs based on emerging non-volatile memories such as resistive random access memory (RRAM) [105], phase change memory (PCM) [106], and magnetic random access memory (MRAM) [107]. Compared to conventional silicon CMOS-based PUFs such as arbiter PUF, ring-oscillator PUF and SRAM PUF, the novel nano-PUFs offer security advantages such as larger number of challenge-response pairs and better stability and robustness [108]. With 3D heterogeneous integration, mainstream silicon CMOS-based ICs can efficiently access these advanced security hardware primitives. In addition, 3D heterogeneous integration of analog circuits and digital circuits offers potential opportunities for developing high-performance defense and security systems. In [109], the author proposed a 3D heterogeneous sensor system that combines sensors, preprocessing analog circuitry, digital logic and signal processing modules to build a comprehensive defense and security system. The performance advantages of 3D technology provide improvements in sensor data collection, detection, classification and autonomous decision-making of such systems.

6 Security Vulnerabilities in 3D ICs

While providing the great promise in terms of performance and security, 3D integration technology might also bring about adverse impacts. In this section, we discuss three potential security vulnerabilities in 3D ICs.

6.1 Testing for Counterfeit and Trojan-Inserted 3D ICs

IC testing is significant for detecting counterfeit components and hardware Trojans introduced in a global IC supply chain. For counterfeit components, physical testing and electrical testing techniques [54] are utilized to rule out used, unauthorized and low-quality products. For hardware Trojan detection, techniques based on functionality verification [62], layout verification [110], and side-channel signatures such as path-delay [111] requires different types of testing to identify malicious inclusions. The challenge of 3D IC testing stems from three aspects: test flows, test contents, and test accesses, as summarized in [6]:

1) *Test flows:* The test flows for 3D ICs are more complex since more potential intermediate tests can be performed in a complicated 3D fabrication process. Pre-bond test is performed before die stacking, which primarily focuses on the circuitries in the dies. Mid-bond test and post-bond test are performed after partial/complete die-stacking, which focus on testing defects in TSV interconnects. Once the stacked dies are packaged, a final test is performed to ensure the overall functionality and reliability of the IC. With split fabrication strategy, as discussed in Section 5.1.2, the design of a trusted and effective test flow becomes even more challenging.
2) *Test contents:* TSVs are the new contents in 3D test flows. TSV-related defects might occur during TSV fabrication and die stacking. Moreover, weak defects and timing faults in TSV-based interconnects and delivery networks (power, ground, and clocks) are difficult to be tested.
3) *Test accesses:* For external test access, the probing of micro-bump of a die is challenging since current probes (with minimum pitch size 50 μm) is unable to directly access the fine-pitch micro-bumps (20 μm). For internal test access, new design-for-test (DfT) architectures for 3D ICs are needed.

Emerging solutions to these challenges have been proposed. A test cost analysis has been performed to develop an economic and effective 3D test flow [112]. Redundant TSV has been proposed to reduce the yield loss due to TSV defects during fabrication [113]. Additional probe pads are integrated into each die to enable external test access and novel DfT architecture [114] for internal test access has been demonstrated.

6.2 Side-Channel Based Hardware Trojan Detection

The stacking of multiple active layers has a huge impact on hardware Trojan detection techniques that are based on side-channel sensors such as current and temperature sensors. Due to the stacking structure and high device density nature of 3D ICs, the signal-to-noise ratio for a side-channel is normally small. When the side-channel signal of Trojan is immersed in noise, the differentiation between Trojan-free and Trojan-active ICs becomes difficult. For example, the temperature of a target block under detection might be affected by other devices in adjacent active layers as well as nearby TSV interconnects. Therefore, the accuracy of temperature signature based hardware Trojan detection technique [65] might be impaired in the 3D context. The same problem also manifests for power signature based detection techniques [64]. IR-drop issue is especially problematic in 3D IC because of its high-current density (due to the increased number of integrated devices but a small number of pins and TSVs). Such IR-drop behaves as noise in the power side-channel of Trojan detection. Moreover, additional fabrication process variations due to the complicated fabrication process of 3D IC may also affect the side-channel signatures for different chips to different extents, which further complicates the analysis of side-channel signals.

6.3 Authenticating 3D IC Layers

3D ICs can be designed and fabricated by stacking conventional 2D dies and interconnecting them with TSVs. These 2D layers may contain functional IPs that are provided by third-party IP vendors and may be fabricated by different foundries. The complicated global supply chain introduces new chances for attackers to insert inauthentic (counterfeit and maliciously modified) designs to compromise the performance and security of the whole chip. Thus, a pre-bond testing for each 2D layer before stacking is important. Once all the layers are bonded, it's difficult to detect an inauthentic layer in the middle since the stacking structure of 3D ICs

TABLE 1
Summary of Security Advantages and Challenges for 3D ICs

2D IC Characteristics	3D IC Characteristics	3D IC Security Advantages	3D IC Security Challenges
2D structure	3D stacking structure	Support split fabrication [56], [57], [58], [59], [60], [64], [67]; Mitigate side channel attacks [83], [86], [87]; Prevent reverse engineering [56], [57]	Testing and authentication of counterfeit and Trojan inserted circuits
Lower device density	Higher device density	Mitigate side channel attacks [87]; Prevent reverse engineering [56], [57]	
Higher interconnect delay/power	Lower interconnect delay/power	Lower performance overhead for aggressive security policies [89] and mechanisms [57], [87], [101], [102], [103]	
No on-chip heterogeneous integration	Heterogeneous integration	Support integration of novel security primitives such as non-silicon PUFs and non-volatile RAMs [88], [105], [106], [107]	

complicates physical testing and electrical testing. However, ICs are nowadays fabricated and tested at offshore foundries in order to obtain high yield at a low cost. If the 2D layers are outsourced to an offshore foundry for fabrication and testing then their testing results might not be trustworthy. Therefore, effective techniques such as secure split-test (SST) [115] shall be developed to ensure a trustworthy prebond testing of these 2D layers. The design of such testing mechanism is more difficult if an IC design is partitioning into multiple layers in gate-level. More design-for-security mechanisms can also be developed to detect and/or isolate the inauthentic layers during runtime.

7 Security-Aware 3D IC Design

In Section 4, a performance-driven 3D IC design flow (Fig. 2) has been described, but security remains to be embedded in that process. In this section, we conclude four distinct characteristics of 3D ICs and investigate their impact on the 3D IC design flow. The distinct characteristics, security advantages and challenges of 3D IC design are summarized in Table 1.

1) *Stacking structure:* The most distinct feature that distinguishes 3D ICs from 2D ICs is the stacking structure. The vertical stacking structure not only offers a natural shield to mitigate side-channel attacks and reverse engineering, but also enables the split fabrication technique to defend the supply chain attacks. *3D partitioning* is the core design phase of a security-ware 3D IC design because it determines the layer assignment for functional blocks in the stacking structure. A system-level 3D partitioning can protect the security-critical functional blocks by placing them in a trusted inner tier. A gate-level partitioning for 2.5D/3D split fabrication can lift wires and/or gates into the trusted tier that can maximally obfuscate the netlist or functionality.
2) *High device density:* The high device density natural of 3D ICs add significant complexity to side-channel attacks, but it also makes side-channel based HT detection more difficult. *3D Placement and routing (P&R)* is the design phase that can alleviate this problem. Side-channel attack mitigation based P&R algorithms can be developed to reduce the signal-to-noise ratio (SNR) by placing more noisy circuitries around cryptographic functional blocks such as crypto-engines. On the other hand, HT detection based P&R algorithms can be implemented to increase the SNR ratio by placing the sensors in a less noisy environment.
3) *Reduced interconnect delay/power:* By vertically stacking multiple 2D chips, 3D integration improves chip performance and enables new architectures, thereby offering significant potential for the *design of secure CPU architectures*. High performance-overhead security policies and mechanisms that are once considered impractical in 2D ICs are now becoming viable with 3D integration [89]. A major challenge for secure 3D CPU design is posed by incorporation of security policy into novel architectures enabled by 3D ICs. An effective co-design approach shall be developed to incorporating security advantages in 3D CPUs while leveraging the challenges in power management, thermal dissipation and testing. This requires an accurate characterization and modeling on the impact of various distinct features of 3D ICs on security.
4) *Heterogeneous integration:* Heterogeneous integration enables mainstream silicon CMOS-based ICs to efficiently access advanced security primitives such as non-silicon PUFs and non-volatile RAMs. At the same time, 3D ICs present new security advantages and challenges for the *deisgn of security primitives*. For example, 3D ICs offer an additional source of entropy for PUF design such as process variation in TSVs and inter-layer variation [116], [117], which do not exist in 2D ICs.

8 Conclusion

While 3D integration is initially developed to overcome the obstacles in device miniaturization, it has presented various security advantages in different security techniques and applications. There is an emerging trend of research exploiting this technology from a security perspective. This paper presents a survey on the current state of 3D integration technology and highlights its opportunities and challenges in

various security applications. It summarizes current research works on 3D integration based security in three major applications including supply chain attack prevention, side-channel attack mitigation and trustworthy computing system design. It also addresses the potential vulnerabilities risen by 3D integration that call for future research efforts. Finally, the paper summarizes the distinct features of 3D ICs, highlight their security advantages/challenges and discuss their impact on a security-aware 3D IC design. With the effort made in 3D IC security characterization and modeling, future chip designers can take security into consideration at an early phase of the design while optimizing the chip for performance and power.

Acknowledgments

This material is based upon work supported by the US National Science Foundation under Grant No. 1223233 and Air Force Office of Scientific Research under Grant FA9550-14-1-0351.

References

[1] Z. Or-Bach. (2012). Is the cost reduction associated with IC scaling over? [Online]. Available: http://www.eetimes.com/author.asp?doc_id=1286363

[2] Y. Xie, J. Cong, and S. S. Sapatnekar, *Three-Dimensional Integrated Circuit Design*. New York, NY, USA: Springer, 2010.

[3] S. Bobba, A. Chakraborty, O. Thomas, P. Batude, V. F. Pavlidis, and G. De Micheli, "Performance analysis of 3D monolithic integrated circuits," in *Proc. IEEE Int. 3D Syst. Integration Conf.*, 2010, pp. 1–4.

[4] M. Puech, J.-M. Thevenoud, J. Gruffat, N. Launay, N. Arnal, and P. Godinat, "Fabrication of 3D packaging TSV using DRIE," in *Proc. Design, Test, Integration Packaging MEMS/MOEMS. Symp.*, 2008, pp. 109–114.

[5] V. S. Rao, H. S. Wee, L. Vincent, L. H. Yu, L. Ebin, R. Nagarajan, C. T. Chong, X. Zhang, and P. Damaruganath, "TSV interposer fabrication for 3D IC packaging," in *Proc. 11th Electron. Packaging Technol. Conf.*, 2009, pp. 431–437.

[6] E. J. Marinissen, "Challenges and emerging solutions in testing TSV-based 2 1/2D-and 3D-stacked ICs," in *Proc. Conf. Design, Automation Test Eur.*, 2012, pp. 1277–1282.

[7] W. R. Davis, J. Wilson, S. Mick, J. Xu, H. Hua, C. Mineo, A. M. Sule, M. Steer, and P. D. Franzon, "Demystifying 3D ICs: The pros and cons of going vertical," *IEEE Design Test Comput.*, vol. 22, no. 6, pp. 498–510, 2005.

[8] X. Wu, J. Li, L. Zhang, E. Speight, R. Rajamony, and Y. Xie, "Hybrid cache architecture with disparate memory technologies," in *Proc. 36th Annu. Int. Symp. Comput. Archit.*, 2009, pp. 34–45.

[9] G. H. Loh, "3D-Stacked memory architectures for multi-core processors," in *Proc. 35th Annu. Int. Symp. Comput. Archit.*, 2008, pp. 453–464.

[10] C. Serafy, A. Srivastava, and D. Yeung, "Unlocking the true potential of 3D CPUs with micro-fluidic cooling," in *Proc. IEEE/ACM Int. Symp. Low Power Electron. Des.*, 2014, pp. 323–326.

[11] C. Serafy, B. Shi, A. Srivastava, and D. Yeung, "High performance 3D stacked DRAM processor architectures with micro-fluidic cooling," in *Proc. IEEE Int. 3D Syst. Integration Conf.*, 2013, pp. 1–8.

[12] B. Feero and P. Pande, "Performance evaluation for three-dimensional networks-on-chip," in *Proc. IEEE Comput. Soc. Annu. Symp. VLSI*, 2007, pp. 305–310.

[13] G. H. Loh, Y. Xie, and B. Black, "Processor design in 3D die-stacking technologies," *IEEE Micro*, vol. 27, no. 3, pp. 31–48, 2007.

[14] M. Jung, T. Song, Y. Wan, Y. Peng, and S. K. Lim, "On enhancing power benefits in 3D ICs: Block folding and bonding styles perspective," in *Proc. 51st ACM/EDAC/IEEE Des. Automation Conf.*, 2014, pp. 1–6.

[15] T. Lu and A. Srivastava, "Electromigration-aware clock tree synthesis for TSV-based 3D-ICs," in *Proc. 25th Edition Great Lakes Symp. VLSI*, 2015, pp. 27–32.

[16] J. Cong and Y. Zhang, "Thermal via planning for 3-D ICs," in *Proc. IEEE/ACM Int. Conf. Comput.-Aid. Des.*, 2005, pp. 745–752.

[17] B. Goplen and S. Sapatnekar, "Thermal via placement in 3D ICs," in *Proc. Int. Symp. Phys. Des.*, 2005, pp. 167–174.

[18] C. Serafy, A. Srivastava, A. Bar-Cohen, and D. Yeung, "Design space exploration of 3D CPUs and micro-fluidic heatsinks with thermo-electrical-physical co-optimization," in *Proc. ASME Int. Tech. Conf. Exhib. Packaging Integr. Electron. Photonic Microsyst.*, 2015.

[19] G. Luo, Y. Shi, and J. Cong, "An analytical placement framework for 3D ICs and its extension on thermal awareness," *IEEE Trans. CAD Integrated Circuits Syst.*, vol. 32, no. 4, pp. 510–523, Apr. 2013.

[20] J. Meng, K. Kawakami, and A. Coskun, "Optimizing energy efficiency of 3-D multicore systems with stacked dram under power and thermal constraints," in *Proc. 49th Annu. Des. Automation Conf.*, 2012, pp. 648–655.

[21] H. J. Choi, Y. J. Park, H.-H. Lee, and C. H. Kim, "Adaptive dynamic frequency scaling for thermal-aware 3D multi-core processors," in *Proc. 12th Int. Conf. Comput. Sci. Its Appl. - Volume Part IV*, 2012, pp. 602–612.

[22] M. Sabry, A. Coskun, D. Atienza, T. Rosing, and T. Brunschwiler, "Energy-efficient multiobjective thermal control for liquid-cooled 3D stacked architectures," *IEEE Trans. Comput.-Aid. Des. Integrated Circuits Syst.*, vol. 30, no. 12, pp. 1883–1896, 2011.

[23] X. Zhou, Y. Xu, Y. Du, Y. Zhang, and J. Yang, "Thermal management for 3D processors via task scheduling," in *Proc. 37th Int. Conf. Parallel Process.*, 2008, pp. 115–122.

[24] S. Liu, J. Zhang, Q. Wu, and Q. Qiu, "Thermal-aware job allocation and scheduling for three dimensional chip multiprocessor," in *Proc. 11th Int. Symp.Quality Electron. Des.*, 2010, pp. 390–398.

[25] I. Savidis, S. Kose, and E. Friedman, "Power noise in TSV-based 3D integrated circuits," *IEEE J. Solid-State Circuits*, vol. 48, no. 2, pp. 587–597, Feb. 2013.

[26] R. Zhang, K. Wang, B. Meyer, M. Stan, and K. Skadron, "Architecture implications of pads as a scarce resource," in *Proc. ACM/IEEE 41st Int. Symp.Comput. Archit.*, 2014, pp. 373–384.

[27] P. Zhou, K. Sridharan, and S. Sapatnekar, "Optimizing decoupling capacitors in 3D circuits for power grid integrity," *IEEE Des. Test Comput.*, vol. 26, no. 5, pp. 15–25, Sep./Oct. 2009.

[28] P. Zhou, "Interconnect design techniques for multicore and 3D integrated circuits," Ph.D. dissertation, Citeseer, 2012.

[29] J. Pak, M. Pathak, S. K. Lim, and D. Z. Pan, "Modeling of electromigration in through-silicon-via based 3D IC," in *Proc. IEEE 61st Electron. Components Technol. Conf.*, 2011, pp. 1420–1427.

[30] T. Frank, S. Moreau, C. Chappaz, P. Leduc, L. Arnaud, A. Thuaire, E. Chery, F. Lorut, L. Anghel, and G. Poupon, "Reliability of TSV interconnects: Electromigration, thermal cycling, and impact on above metal level dielectric," *Microelectron. Rel.*, vol. 53, no. 1, pp. 17–29, 2013.

[31] M. Jung, X. Liu, S. K. Sitaraman, D. Z. Pan, and S. K. Lim, "Full-chip through-silicon-via interfacial crack analysis and optimization for 3D IC," in *Proc. Int. Conf. Comput.-Aid. Des.*, 2011, pp. 563–570.

[32] Y. Cao, J. Velamala, K. Sutaria, M. S.-W. Chen, J. Ahlbin, I. Sanchez Esqueda, M. Bajura, and M. Fritze, "Cross-layer modeling and simulation of circuit reliability," *IEEE Trans. Comput.-Aid. Des. Integrated Circuits Syst.*, vol. 33, no. 1, pp. 8–23, 2014.

[33] L. Jiang, Q. Xu, and B. Eklow, "On effective TSV repair for 3D-stacked ICs," in *Proc. Conf. Des. Automation Test Eur.*, 2012, pp. 793–798.

[34] Q. Zou, T. Zhang, E. Kursun, and Y. Xie, "Thermomechanical stress-aware management for 3D IC designs," in *Proc. Des. Automation Test Eur. Conf. Exhib.*, 2013, pp. 1255–1258.

[35] T. Chantem, Y. Xiang, X. Hu, and R. P. Dick, "Enhancing multicore reliability through wear compensation in online assignment and scheduling," in *Proc. Des. Automation Test Eur. Conf. Exhib.*, 2013, pp. 1373–1378.

[36] H. Tajik, H. Homayoun, and N. Dutt, "VAWOM: Temperature and process variation aware wearout management in 3D multicore architecture," in *Proc. 50th ACM/EDAC/IEEE Des. Automation Conf.*, 2013, pp. 1–8.

[37] C. Liu, T. Song, J. Cho, J. Kim, J. Kim, and S. K. Lim, "Full-chip TSV-to-TSV coupling analysis and optimization in 3D IC," in *Proc. 48th Des. Autom. Conf.*, 2011, pp. 783–788.

[38] N. Khan, S. Alam, and S. Hassoun, "Through-silicon via (TSV)-induced noise characterization and noise mitigation using coaxial TSVs," in *Proc. IEEE Int. Conf. 3D Syst. Integration*, 2009, pp. 1–7.

[39] J. Cho, E. Song, K. Yoon, J. S. Pak, J. Kim, W. Lee, T. Song, K. Kim, J. Lee, H. Lee, K. Park, S. Yang, M. Suh, K. Byun, and J. Kim, "Modeling and analysis of through-silicon via (TSV) noise coupling and suppression using a guard ring," *IEEE Trans. Components, Packag. Manuf. Technol.*, vol. 1, no. 2, pp. 220–233, Feb. 2011.

[40] C. Serafy, B. Shi, and A. Srivastava, "A geometric approach to chip-scale TSV shield placement for the reduction of TSV coupling in 3D-ICs," in *Proc. 23rd ACM Int. Conf. Great Lakes Symp. VLSI.*, 2014, pp. 275–280.

[41] C. Serafy and A. Srivastava, "TSV replacement and shield insertion for TSV-TSV coupling reduction in 3D global placement," *IEEE Trans. Comput.-Aid. Des. Integr. Circuits Syst.*, vol. 34, no. 4, pp. 554–562, Apr. 2015.

[42] Y. Peng, T. Song, D. Petranovic, and S. K. Lim, "Silicon effect-aware full-chip extraction and mitigation of TSV-to-TSV coupling," *IEEE Trans. Comput.-Aid. Des. Integr. Circuits Syst.*, vol. 33, no. 12, pp. 1900–1913, Dec. 2014.

[43] X. Zhao, J. Minz, and S. K. Lim, "Low-power and reliable clock network design for through-silicon via (TSV) based 3D ICs," *IEEE Trans. Compon. Packag. Manuf. Technol.*, vol. 1, no. 2, pp. 247–259, Feb. 2011.

[44] T.-Y. Kim and T. Kim, "Clock tree synthesis for TSV-based 3D IC designs," *ACM Trans. Des. Autom. Electron. Syst.*, vol. 16, no. 4, 2011, Art. no. 48.

[45] J. Yang, J. Pak, X. Zhao, S. K. Lim, and D. Pan, "Robust clock tree synthesis with timing yield optimization for 3D-ICs," in *Proc. 16th Asia South Pacific Des. Automation Conf.*, 2011, pp. 621–626.

[46] T. Lu and A. Srivastava, "Gated low-power clock tree synthesis for 3D-ICs," in *Proc. Int. Symp. Low Power Electron. Des.*, 2014, pp. 319–322.

[47] Gartner Inc. (2012). Market trends: Rising costs of production limit availability of leading-edge fabs [Online]. Available: https://www.gartner.com/doc/2163515

[48] M. Rostami, F. Koushanfar, and R. Karri, "A primer on hardware security: Models, methods, and metrics," *Proc. IEEE*, vol. 102, no. 8, pp. 1283–1295, Aug. 2014.

[49] SEMI. (2012). IP challenges for the semiconductor equipment and materials industry. [Online]. Available: http://www.semi.org/Issues/IntellectualProperty

[50] SEMI. (2008). Innovation at risk - Intellectual property challenges and opportunities [Online]. Available: http://www.semi.org/en/Press/P043775

[51] S. Skorobogatov and C. Woods, "Breakthrough silicon scanning discovers backdoor in military chip," in *Proc. 14th Int. Workshop Cryptographic Hardw. Embedded Syst.*, 2012, pp. 23–40.

[52] K. Vaidyanathan, B. P. Das, and L. Pileggi, "Detecting reliability attacks during split fabrication using test-only BEOL stack," in *Proc. 51st ACM/EDAC/IEEE Des. Autom. Conf.*, 2014, pp. 1–6.

[53] R. Torrance and D. James, "The state-of-the-art in IC reverse engineering," in *Proc. 11th Int. Workshop Cryptographic Hardware Embedded Syst.*, 2009, pp. 363–381.

[54] U. Guin, K. Huang, D. DiMase, J. M. Carulli, M. Tehranipoor, and Y. Makris, "Counterfeit integrated circuits: A rising threat in the global semiconductor supply chain," *Proc. IEEE*, vol. 102, no. 8, pp. 1207–1228, Aug. 2014.

[55] M. Tehranipoor and F. Koushanfar, "A survey of hardware trojan taxonomy and detection," *IEEE Des. Test Comput.*, vol. 27, no. 1, pp. 10–25, Jan./Feb. 2010.

[56] Tezzaron. (2008). 3D-ICs and integrated circuit security [Online]. Available: http://www.tezzaron.com/about/papers/3D-ICs_and_Integrated_Circuit_Security.pdf

[57] J. Valamehr, T. Sherwood, R. Kastner, D. Marangoni-Simonsen, T. Huffmire, C. Irvine, and T. Levin, "A 3D split manufacturing approach to trustworthy system development," *IEEE Trans. Comput.-Aid. Des. Integr. Circuits Syst.*, vol. 32, no. 4, pp. 611–615, 2013.

[58] Y. Xie, C. Bao, and A. Srivastava, "Security-aware design flow for 2.5D IC technology," in *Proc. 5th Int. Workshop Trustworthy Embedded Devices*, 2015, pp. 31–38.

[59] F. Imeson, A. Emtenan, S. Garg, and M. V. Tripunitara, "Securing computer hardware using 3D integrated circuit (IC) technology and split manufacturing for obfuscation." in *Proc. 22nd USENIX Conf. Security*, 2013, pp. 495–510.

[60] K. Xiao, D. Forte, and M. M. Tehranipoor, "Efficient and secure split manufacturing via obfuscated built-in self-authentication," in *Proc. IEEE Int. Symp. Hardware Oriented Security Trust*, 2015, pp. 14–19.

[61] J. Rajendran, O. Sinanoglu, and R. Karri, "Is split manufacturing secure?" in *Proc. Des. Automation Test Eur. Conf. Exhib.*, 2013, pp. 1259–1264.

[62] H. Salmani, M. Tehranipoor, and J. Plusquellic, "New design strategy for improving hardware trojan detection and reducing trojan activation time," in *Proc. IEEE Int. Workshop Hardware-Oriented Security Trust*, 2009, pp. 66–73.

[63] X. Zhang and M. Tehranipoor, "RON: An on-chip ring oscillator network for hardware trojan detection," in *Proc. Des. Autom. Test Eur. Conf. Exhib.*, 2011, pp. 1–6.

[64] S. Narasimhan, W. Yueh, X. Wang, S. Mukhopadhyay, and S. Bhunia, "Improving IC security against trojan attacks through integration of security monitors," *IEEE Design Test Comput.*, vol. 29, no. 5, pp. 37–46, Oct. 2012.

[65] C. Bao, D. Forte, and A. Srivastava, "Temperature tracking: Toward robust run-time detection of hardware trojans," *IEEE Trans. Comput.-Aid. Des. Integr. Circuits Syst.*, vol. 34, no. 10, pp. 1577–1585, 2015.

[66] K. Xiao and M. Tehranipoor, "BISA: Built-in self-authentication for preventing hardware trojan insertion," in *Proc. IEEE Int. Symp. Hardware-Oriented Security Trust*, 2013, pp. 45–50.

[67] M. Bilzor, "3D execution monitor (3D-EM): Using 3D circuits to detect hardware malicious inclusions in general purpose processors," in *Proc. 6th Int. Conf. Inform. Warfare Security*, 2011.

[68] K. Vaidyanathan, B. Das, E. Sumbul, R. Liu, and L. Pileggi, "Building trusted ICs using split fabrication," in *Proc. IEEE Int. Symp. Hardware-Oriented Security Trust*, 2014, pp. 1–6.

[69] K. Vaidyanathan, R. Liu, E. Sumbul, Q. Zhu, F. Franchetti, and L. Pileggi, "Efficient and secure intellectual property (IP) design with split fabrication," in *Proc. IEEE Int. Symp. Hardware-Oriented Security Trust*, 2014, pp. 13–18.

[70] M. Jagasivamani, P. Gadfort, M. Sika, M. Bajura, and M. Fritze, "Split-fabrication obfuscation: Metrics and techniques," in *Proc. IEEE Int. Symp. Hardware-Oriented Security Trust*, 2014, pp. 7–12.

[71] D. Brumley and D. Boneh, "Remote timing attacks are practical," *Comput. Netw. The Int. J. Comput. Telecommun. Netw. Web Security*, vol. 48, no. 5, pp. 701–716, 2005.

[72] D. J. Bernstein, "Cache-timing attacks on AES," Technical report, 2005. [Online]. Available: http://cr.yp.to/papers.html#cachetiming.

[73] P. Kocher, J. Jaffe, and B. Jun, "Differential power analysis," in *Proc. 19th Annu. Int. Cryptology Conf. Adv. Cryptology*, 1999, pp. 388–397.

[74] E. Brier, C. Clavier, and F. Olivier, "Correlation power analysis with a leakage model," in *Proc. 6th Int. Workshop Cryptographic Hardware Embedded Syst.*, 2004, pp. 16–29.

[75] J.-J. Quisquater and D. Samyde, "Electromagnetic analysis (EMA): Measures and counter-measures for smart cards," in *Proc. Int. Conf. Res. Smart Cards Programm. Security*, 2001, pp. 200–210.

[76] D. Agrawal, B. Archambeault, J. R. Rao, and P. Rohatgi, "The EM sidechannel (s)," in *Proc. 4th Int. Workshop Cryptographic Hardware Embedded Syst.*, 2003, pp. 29–45.

[77] K. Gandolfi, C. Mourtel, and F. Olivier, "Electromagnetic analysis: Concrete results," in *Proc. 3rd Int. Workshop Cryptographic Hardware Embedded Syst.*, 2001, pp. 251–261.

[78] Y. Zhou and D. Feng, "Side-channel attacks: Ten years after its publication and the impacts on cryptographic module security testing." *IACR Cryptology ePrint Archive*, vol. 2005, p. 388, 2005.

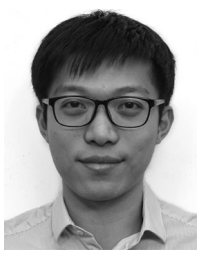
[79] P. C. Kocher, "Timing attacks on implementations of diffie-hellman, RSA, DSS, and other systems," in *Proc. 16th Annu. Int. Cryptology Conf. Adv. Cryptology*, 1996, pp. 104–113.

[80] E. Tromer, D. A. Osvik, and A. Shamir, "Efficient cache attacks on AES, and countermeasures," *J. Cryptol.*, vol. 23, no. 2, pp. 37–71, 2010.

[81] P. Kocher, J. Jaffe, B. Jun, and P. Rohatgi, "Introduction to differential power analysis," *J. Cryptographic Eng.*, vol. 1, no. 1, pp. 5–27, 2011.

[82] A. Barenghi, L. Breveglieri, I. Koren, and D. Naccache, "Fault injection attacks on cryptographic devices: Theory, practice, and countermeasures," *Proc. IEEE*, vol. 100, no. 1, pp. 3056–3076, Nov. 2012.

[83] N. Madan and R. Balasubramonian, "Leveraging 3D technology for improved reliability," in *Proc. 40th Annu. IEEE/ACM Int. Symp. Microarchit.*, 2007, pp. 223–235.
[84] D. Boneh, R. A. DeMillo, and R. J. Lipton, "On the importance of eliminating errors in cryptographic computations," *J. Cryptol.*, vol. 14, no. 2, pp. 101–119, 2001.
[85] E. Biham and A. Shamir, "Differential fault analysis of secret key cryptosystems," in *Proc. 17th Annu. Int. Cryptol. Conf. Adv. Cryptol.*, 1997, pp. 513–525.
[86] W. Zhang and T. Li, "Microarchitecture soft error vulnerability characterization and mitigation under 3D integration technology," in *Proc. 41st IEEE/ACM Int. Symp. Microarchit.*, 2008, pp. 435–446.
[87] J. Valamehr, T. Huffmire, C. Irvine, R. Kastner, Ç. K. Koç, T. Levin, and T. Sherwood, "A qualitative security analysis of a new class of 3D integrated crypto co-processors," in *Cryptography and Security: From Theory to Applications*, Berlin Germany: Springer, 2012, pp. 364–382.
[88] G. Khedkar and D. Kudithipudi, "RRAM motifs for mitigating differential power analysis attacks (DPA)," in *Proc. IEEE Comput. Soc. Annu. Symp. VLSI*, 2012, pp. 88–93.
[89] C. Bao and A. Srivastava, "3D integration: New opportunities in defense against cache-timing side-channel attacks," in *Proc. 33rd IEEE Int. Conf. Comput. Des.*, 2015, pp. 273–280.
[90] C. King and C. Beal, "Csi kernel: Finding a needle in a multiterabyte haystack," *IEEE Softw.*, vol. 29, no. 6, pp. 9–12, Nov./Dec. 2012.
[91] CVE Details: The ultimate security vulnerability datasource (2014) [Online]. Available: http://www.cvedetails.com
[92] U. Steinberg and B. Kauer, "NOVA: A microhypervisor-based secure virtualization architecture," in *Proc. 5th Eur. Conf. Comput. Syst.*, 2010, pp. 209–222.
[93] J. M. McCune, Y. Li, N. Qu, Z. Zhou, A. Datta, V. Gligor, and A. Perrig, "TrustVisor: Efficient TCB reduction and attestation," in *Proc. IEEE Symp. Security Privacy*, 2010, pp. 143–158.
[94] G. E. Suh, D. Clarke, B. Gassend, M. Van Dijk, and S. Devadas, "AEGIS: Architecture for tamper-evident and tamper-resistant processing," in *Proc. 17th Annu. Int. Conf. Supercomput.*, 2003, pp. 160–171.
[95] J. Winter, "Trusted computing building blocks for embedded linux-based ARM trustzone platforms," in *Proce. 3rd ACM Workshop Scalable Trusted Comput.*, 2008, pp. 21–30.
[96] E. Keller, J. Szefer, J. Rexford, and R. B. Lee, "NoHype: Virtualized cloud infrastructure without the virtualization," in *Proc. 37th Annu. Int. Symp. ACM SIGARCH Comput. Archit. News*, 201, pp. 350–361.
[97] E. Owusu, J. Guajardo, J. McCune, J. Newsome, A. Perrig, and A. Vasudevan, "Oasis: On achieving a sanctuary for integrity and secrecy on untrusted platforms," in *Proc. ACM SIGSAC Conf. Comput. Commun. Security*, 2013, pp. 13–24.
[98] (2008). Trusted Computing Group. Trusted platform module (TPM) summary [Online]. Available: http://www.trustedcomputinggroup.org/
[99] E. R. Sparks and E. R. Sparks, "A security assessment of trusted platform modules computer science technical report TR2007-597," Dept. Comput. Sci., Dartmouth College, Hanover, NH, USA, Tech. Rep., TR2007-597, 2007.
[100] J. Winter and K. Dietrich, "A hijackers guide to the LPC bus," in *Proc. 8th Eur. Conf. Public Key Infrastructures, Services Appl.*, 2012, pp. 176–193.
[101] S. Mysore, B. Agrawal, N. Srivastava, S.-C. Lin, K. Banerjee, and T. Sherwood, "Introspective 3D chips," in *Proc. ACM 12th Int. Conf. Archit. Support Programm. Languages Operating Syst.*, 2006, pp. 264–273.
[102] J. Valamehr, M. Tiwari, T. Sherwood, R. Kastner, T. Huffmire, C. Irvine, and T. Levin, "Hardware assistance for trustworthy systems through 3-D integration," in *Proc. 26th Annu. Comput. Security Appl. Conf.*, 2010, pp. 199–210.
[103] T. Huffmire, T. Levin, M. Bilzor, C. E. Irvine, J. Valamehr, M. Tiwari, T. Sherwood, and R. Kastner, "Hardware trust implications of 3D integration," in *Proc. 5th Workshop Embedded Syst. Security*, 2010, Art. no. 1.
[104] J.-Q. Lu, "3D hyperintegration and packaging technologies for micro-nano systems," *Proc. IEEE*, vol. 97, no. 1, pp. 18–30, 2009.
[105] J. Rajendran, G. S. Rose, R. Karri, and M. Potkonjak, "Nano-PPUF: A memristor-based security primitive," *VLSI (ISVLSI), 2012 IEEE Comput. Soc. Annu. Symp.*, pp. 84–87, 2012.
[106] L. Zhang, Z. H. Kong, C.-H. Chang, A. Cabrini, and G. Torelli, "Exploiting process variations and programming sensitivity of phase change memory for reconfigurable physical unclonable functions," *IEEE Trans. Inform. Forensics Security*, vol. 9, no. 6, pp. 921–932, 2014.
[107] J. Das, K. Scott, S. Rajaram, D. Burgett, and S. Bhanja, "MRAM PUF: A novel geometry based magnetic PUF with integrated CMOS," *IEEE Trans. Nanotechnol.*, vol. 14, no. 3, pp. 436–443, May. 2015.
[108] P.-Y. Chen, R. Fang, R. Liu, C. Chakrabarti, Y. Cao, and S. Yu, "Exploiting resistive cross-point array for compact design of physical unclonable function," in *Proc. IEEE Int. Symp. Hardware Oriented Security Trust*, 2015, pp. 26–31.
[109] S. Bhansali, G. H. Chapman, E. G. Friedman, Y. Ismail, P. Mukund, D. Tebbe, and V. K. Jain, "3D heterogeneous sensor system on a chip for defense and security applications," in *Proc. Defense and Security*, 2004.
[110] C. Bao, D. Forte, and A. Srivastava, "On application of one-class SVM to reverse engineering-based hardware trojan detection," in *Proc. 15th Int. Symp. Quality Electronic Des.*, 2014, pp. 47–54.
[111] J. Li and J. Lach, "At-speed delay characterization for IC authentication and trojan horse detection," in *Proc. IEEE Int. Workshop Hardware-Oriented Security Trust*, 2008, pp. 8–14.
[112] M. Taouil, S. Hamdioui, K. Beenakker, and E. J. Marinissen, "Test cost analysis for 3D die-to-wafer stacking," in *Proc. 19th IEEE Asian Test Symp.*, 2010, pp. 435–441.
[113] A.-C. Hsieh and T. Hwang, "TSV redundancy: Architecture and design issues in 3D IC," *IEEE Trans. Very Large Scale Integration Syst.*, 2012, pp. 711–722.
[114] E. J. Marinissen, B. De Wachter, S. O'Loughlin, S. Deutsch, C. Papameletis, and T. Burgherr, "Vesuvius-3D: A 3D-DfT demonstrator," in *Proc. IEEE Int. Test Conf.*, 2014, pp. 1–10.
[115] G. K. Contreras, M. T. Rahman, and M. Tehranipoor, "Secure split-test for preventing IC piracy by untrusted foundry and assembly," in *Proc. Int. Symp. Defect Fault Tolerance VLSI Nanotechnol. Syst.*, 2013, pp. 196–203.
[116] M. Wang, A. Yates, and I. L. Markov, "SuperPUF: Integrating heterogeneous physically unclonable functions," in *Proc. IEEE/ACM Int. Conf. Comput.-Aid. Des.*, 2014, pp. 454–461.
[117] C. Wang, J. Zhou, X. L. Katti Guruprasad, R. Weerasekera, and T. T. Kim, "TSV-based PUF circuit for 3DIC sensor nodes in IoT applications," in *Proc. IEEE Int. Conf. Electron Devices Solid-State Circuits*, 2015, pp. 313–316.



Yang Xie received the BS degree in electrical engineering from Zhejiang Univeristy, Hangzhou, China, in 2013. He is currently working toward the PhD degree in the Electrical and Computer Engineering Department, University of Maryland, College Park. His current research interests include hardware security, trustworthy hardware design, IP piracy prevention, and 3D IC security.

Chongxi Bao received the BS degree in mechanical engineering from Tsinghua University, Beijing, China, in 2012. He is currently working toward the PhD degree in the Electronic and Communication Engineering Department, University of Maryland, College Park. His current research interests include trusted hardware design, side-channel attack mitigation, and security issues in 3D integrated circuits.

Tiantao Lu received the BS degree from Peking University, Beijing, China, in 2011. He is currently working toward the PhD degree in computer engineering at the University of Maryland, College Park. His current research interests include designing high-performance, low-power, highly reliable, and secure 3D-IC.

Caleb Serafy received the BS degree in computer engineering and the MS degree in electrical engineering from Binghamton University in 2010 and 2011, respectively. He is currently working toward the PhD degree in the Electrical and Computer Engineering Department, University of Maryland (UMD), College Park. Since 2011, he has been a research assistant at UMD studying under Prof. Ankur Srivastava. His current research interests include thermal-electrical-physical co-design of 3D ICs. He received the Distinguished Graduate Fellowship, Summer Research Fellowship and Distinguished Dissertation Fellowship from the University of Maryland.

Ankur Srivastava received the BTech degree in electrical engineering from the Indian Institute of Technology Delhi in 1998, the MS degree in electronics and communication engineering from the Northwestern University in 2000, and the PhD degree in computer science from University of California, Los Angeles, in 2002. He is currently a full professor in the Electronics and Communication Engineering Department with joint appointment with the Institute for Systems Research, University of Maryland, College Park. He is the associate editor of the *IEEE Transactions on VLSI* and *INTEGRATION: VLSI Journal*. He is a senior member of the IEEE.

Mark Tehranipoor received the PhD degree from the University of Texas, Dallas, in 2004. He is currently the Intel Charles E. Young Preeminence Endowed professor in Cybersecurity, University of Florida (UF). His current research projects include: hardware security and trust, supply chain security, VLSI design, test, and reliability. He has published more than 300 journal articles and refereed conference papers and has given more than 150 invited talks and keynote addresses. He has published six books and 11 book chapters. He received several Best Paper Awards as well as the 2008 IEEE Computer Society (CS) Meritorious Service Award, the 2012 IEEE CS Outstanding Contribution, the 2009 US NSF CAREER Award, and the 2014 MURI Award. He serves on the program committee of more than a dozen leading conferences and workshops. He served as program chair of the 2007 IEEE Defect-Based Testing (DBT) workshop and the 2008 IEEE Defect and Data Driven Testing (D3T) workshop, co-program Chair of the 2008 International Symposium on Defect and Fault Tolerance in VLSI Systems (DFTS), general chair for D3T-2009 and DFTS-2009, and vice-general chair for NATW-2011. He co-founded the IEEE International Symposium on Hardware-Oriented Security and Trust (HOST) and served as HOST-2008 and HOST-2009 General Chair. He is currently serving as an associate editor for *JETTA*, *JOLPE*, *IEEE TVLSI*, and *ACM TODAES*. Prior to joining UF, he served as the founding director for CHASE and CSI centers at the University of Connecticut. He is currently serving as co-director for the Florida Institute for Cybersecurity Research (FICS). He is a Golden Core Member of the IEEE, and member of the ACM and ACM SIGDA. He is a senior member of the IEEE.