Defeating Microprobing Attacks using a Resource Efficient Detection Circuit

Michael Weiner*, Salvador Manich† and Georg Sigl*
*Lehrstuhl für Sicherheit in der Informationstechnik
Technische Universität München
{m.weiner,sigl}@tum.de
†Universitat Politècnica de Catalunya (UPC)
salvador.manich@upc.edu

***Abstract*—Microprobing is an attack technique against integrated circuits implementing security functions, such as OTP tokens or smartcards. It allows intercepting secrets from on-chip wires as well as injecting faults for other attacks. While the necessity to etch open chip packages and to remove the passivation layer makes microprobing appear expensive, it was shown that a successful attack can be run with equipment worth a few thousand euros. On the protector's side, however, appropriate countermeasures such as active shields, redundancy of core components, or analog detection circuits containing large capacitors, are still expensive.**

We present a resource efficient microbing detection circuit that we call Low Area Probing Detector (LAPD). It measures minimal timing differences between on-chip wires caused by the capacitive load of microprobes. Simulations show that it can detect up-to-date probes with capacitances as low as 10 fF. As a novelty, the LAPD is merely based on digital components and does not require analog circuitry, which reduces the required area and process steps compared to previous approaches.

***Index Terms*—Digital Integrated Circuits; Security; Smart Cards; Data Buses; Microprobing; Invasive Attacks**

I. Introduction

Microprobing is one of the most dangerous attacks that can be carried out on a secure chip. Its purpose is to violate the tamper-resistance characteristics of a chip. Despite its higher cost compared to other types of attacks, it has the advantage to achieve direct reading of internal data or writing on control signals. Furthermore, attackers can use microprobing to manipulate the behavior of an attacked chip by forcing on-chip wires to arbitrary voltages. [1]

When the reading of data is the objective, buses are a more interesting target than memory cells because on buses, all relevant data passes through a few single lines. Worse than that, buses are difficult to hide in lower metal layers due to their extension [2].

Since years ago, buses and other chip structures have been protected by different means like active/passive shields and other check mechanisms. Shields are top layers of metal that usually cover the whole surface of the chip [3]. Planarization and a dense mesh of conductive routes complicate the access to lower layers of it. In active shields, routes are periodically tested to detect breaks in them [4].

The performance improvement and accessibility of specialized laboratory equipment, from micropositioners over high-end optical and electron microscopes up to Focused Ion Beams (FIBs) has put in danger protection measures like shields. FIBs can drill holes with the necessary depth between meanders of metal routes which access underlying lines of interest without damaging the shield. Later on, they can deposit conductive material to route initially inaccessible signals to the surface, which allows the right alignment and contact of microprobes [5].

The aggregation of impediments against these attacks has forced attackers to search for other easier alternatives. Access through the backside of the chip has been recently investigated. Using the photonic emission of transistors, a mapping of transistors is elaborated and regions of interest located [6]. Then, the backside is thinned down close to transistors, approximately $50\,\mu\mathrm{m}$, and thereafter FIB machine edition completes the access to source and drain of the target ones. As formerly described, ad hoc metal contacts are added to ease the microprobe contacts [7]. In such a way, buses can be accessed too by locating the driving buffers, which usually produce larger photo-emission.

In this paper, a methodology to protect the lines of buses from the inside of itself is presented. It is based on the fact that the timing behavior of bus lines is mutually similar under normal conditions, while attaching a microprobe to some lines makes this significantly different.

In [8], a Probe Attempt Detector (PAD) was presented which is able to react when the tip of a microprobe is doing contact with a bus line. The advantage of the circuit is its differential behavior, autonomous operation and high sensitivity. Essentially, the PAD compares the delay differences between the bus lines and a digital word is produced according to the maximum delay difference existing between neighboring lines. Its limitations are the use of non-standard cell elements and its area requirement. Because the core of the PAD is analog, it needs to be designed at a transistor level using structures that do not exist in standard cell libraries. Additionally, in order to increase the sensitivity, an integration module is included that uses a large tank capacitor demanding a significant area amount.

In this paper, a Low Area Probing Detector (LAPD) is presented which is implemented only using digital standard cells and achieves a sensitivity degree in the order of magnitude of

the present commercial microprobes [9]. Owing to the digital scheme, the area requirement is much lower than the PAD as no analog components such as capacitors are required. In addition, the LAPD allows detecting probes in a single shot, while the PAD is slower and consumes more energy as it requires counting clock ticks of a ring oscillator.

The rest of the paper is organized as follows: In Section II the statement of the problem is formally presented. In Section III the LAPD is explained in detail and in Section IV simulation results are shown. Finally concluding remarks are presented in Section V.

II. Problem Statement

Consider an on-chip bus transmitting sensitive data that consists of n lines, to some of which an attacker can contact microprobes. Also assume that the test bench is static, i.e. the probes are not moved while the attacked chip is powered up.

The cost of an attacker faced with the need to contact k lines instead of one may appear linear in the first place – he needs to have access to k microprobes, micropositioners, amplifiers and other equipment. However, we assume it is more than linear, as he will also face additional and presumably time consuming practical challenges: the more probes are required, the more likely it is that different needles obstruct each other's way. Furthermore, it is difficult to position many micropositioners around the chip as they are orders of magnitude larger than the chip itself.

In order to minimize the harm of the information leakage by the bus, two possible strategies could be selected: the encryption of the data transmitted through the bus or the *detection of the microprobe presence* by electrical means. The first strategy needs very high resources in terms of chip area, power consumption and computation time, as well as price for certification. This is mainly used for the protection of high-value targets such as Pay TV smartcards [10], [11]. It is, however, not feasible for mass-market low performance processors – for example, SIM cards or RFID based public transport tickets – while the second strategy can be implemented at a vastly reduced cost in terms of area and power consumption. We focus on the latter case because we target low-cost secure chips.

The *detection of the microprobe presence* could be performed online, while data is transmitted through the bus, or *offline*, at time instants when the bus is idle. Like before, the online mode will typically require more power because it will be in continuous operation while the *offline* mode will consume power only during its activation. For this reason, the selected mode is the *offline* mode.

In conclusion, the LAPD presented in this paper is of type *offline detector of microprobe presence* by electrical means. Since the test bench is assumed static, typically a detection run must be performed after reset and/or in bus idle cycles before critical data is transfered over the bus.

III. The Low Area Probing Detector

Attaching a microprobe to a bus line increases its capacitive load. Different capacitive loads of equally sized lines lead to different delays of these lines. We present the Low Area Probing Detector (LAPD) that detects microprobing by observing the timing differences between two or more adjacent bus lines. This increases the complexity of a microprobing attack: If n lines are protected by the LAPD, $n-1$ microprobe connections can be detected such that the adversary would need to attach the same capacitive load to all n protected lines. We assume to protect buses consisting of lines with similar dimensions and delays.

In order to achieve the maximum level of security, the LAPD shall protect all lines that either transfer sensitive information or can be used for forcing or fault injection. This work is focused on the protection of bus lines on a security microcontroller: they transfer sensitive information between different components on the chip and they are easy targets as they are presumably situated on the top metal layers due to the distance they need to cross. Furthermore, their structure is well suitable for our symmetry assumptions that are used for this work. Alternatively, the LAPD can be used to enhance the security of active shields such that they do not only evaluate the existence of proper connections, but also validate their timing behavior [2].

A. Principle of Operation

The LAPD protects a set of bus lines in a system, as shown in Figure 1 for the example case of two lines. The lines to be monitored by the LAPD each have the parasitic capacitance C_L, while an attacker probing a line introduces the additional capacitance C_A, which increases the total capacitance of the probed line to $C_L + C_A$.

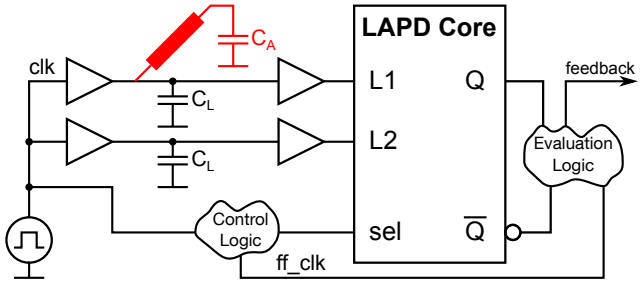


Figure 1. Overview of a System using the LAPD

During the attack, the line capacitances are

$$C_1 = C_L + C_A \tag{1}$$

$$C_2 = C_L \tag{2}$$

where C_1 is the capacitance of the victim line `L1` and C_2 is the capacitance of the reference line `L2`. Assuming the alpha-power model for the transistors [12], [13], the delay of the line buffers can be approximated by

$$d_i = \tilde{k}\frac{C_i\ V_{DD}}{(V_{DD} - V_t)^{\alpha}} \tag{3}$$

where α is the velocity saturation coefficient of the carriers, V_t is the threshold voltage of the transistors, $\tilde{k}$ is the transresistance including the remaining transistor parameters, V_{DD}

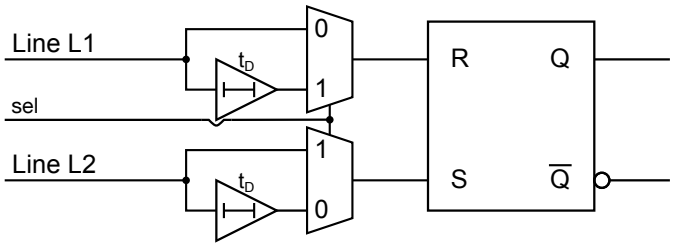
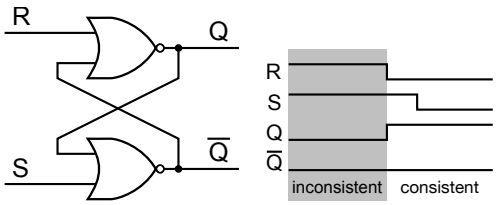


Figure 2. Conceptual Schematic of the LAPD Core



(a) Circuit (b) Timing

Figure 3. NOR RS Latch

the supply voltage and C_i the load of the driving buffer [14]. All technological parameters are balanced between nmos and pmos transistors. Equation (3), as explained in [14], assumes that signals approach voltage limits during swinging, which is the case when signals propagate through chains of gates.

After the attack the delay difference between lines `L1` and `L2` is

$$d_1 - d_2 = \tilde{k}\frac{(C_1 - C_2)V_{DD}}{(V_{DD} - V_t)^\alpha} = \Omega\ C_A \tag{4}$$

with

$$\Omega = \tilde{k}\frac{V_{DD}}{(V_{DD} - V_t)^\alpha} \tag{5}$$

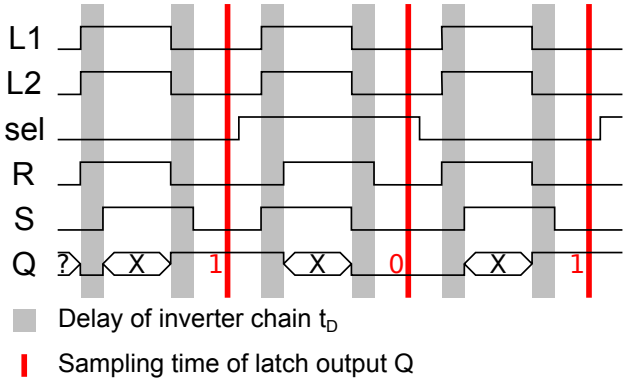
As shown in (4), the delay difference is, in a first approximation, proportional to the amount of capacitance of the microprobe. This relationship is valid for small values of C_A which is the characteristic property of advanced microprobes. For probes with larger C_A, Equation (4) tends to a saturation but in any case the increase of delay function is monotonic and therefore we expect the circuit to behave reliably.

The LAPD detects this delay difference by evaluating race conditions between the two inputs of an RS latch, as shown in Figure 2. A clock signal drives lines `L1` and `L2`, while a control logic alternates inserting intentional delays t_D in the end of these lines and before the `R` and the `S` input, such that the latch output shall alternate between 0 and 1 every cycle. It is preferrable that the clock is not externally accessible to avoid attacks such as glitching. The delay t_D is chosen such that its value is above the intrinsic timing jitter between the `R` and `S` inputs and below the minimum timing delay that is expected to be introduced by the microprobe.

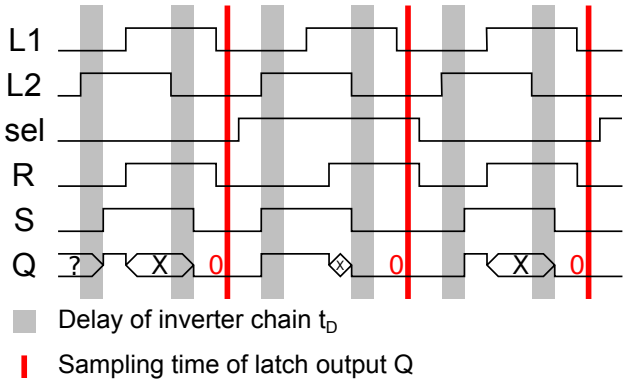
B. The LAPD Architecture

The LAPD is based on the timing behavior of a standard Reset-Set (RS) latch, as depicted in Figure 3(a) for the NOR implementation. The *Basic Concept* section explains the most basic case protecting two lines, *Control and Evaluation Logic* describes the components required for operation, and *Protection of n Lines* explains how to protect more than two lines.

1) Basic Concept: An RS latch, as composed of two NOR or NAND gates, is a memory cell that can be set by activating the `S` input and reset by activating the `R` input. As shown in Figure 3(b) for NOR RS latches, `Q` does not match $not(\overline{\texttt{Q}})$ during the time that `R` and `S` are active simulaneously – in other words, the output is inconsistent. However, as soon as the first of the two inputs returns to the inactive state, the other, still active input "wins the race" and the output becomes valid again.



(a) Without Probing



(b) With Probing of L1

Figure 4. LAPD Timing

The LAPD makes use of this behavior by providing both `R` and `S` with a square wave, e.g., a clock signal, where one of the `R` and `S` lines is alternately delayed. For our assumed case of balanced lines, the latch output `Q` will alternate between 0 and 1 every clock cycle if no probe is attached. The switchable delay driver is dimensioned to be smaller than the delay introduced by the target microprobe: If an adversary attaches such a probe, it will constantly delay one of the lines beyond the other line, such that `Q` will stop alternating and give a constant output of 0 or 1, depending on the line that is probed.

The timing of the LAPD is shown in Figure 4. The inconsistent output state of the latch is denoted "X", and an

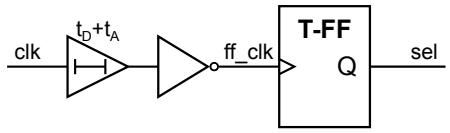


Figure 5. LAPD Control Logic

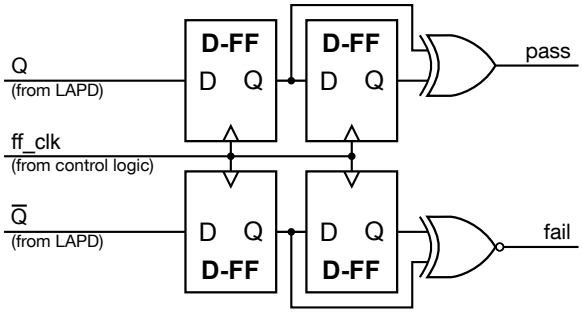


Figure 6. Redundant LAPD Evaluation Logic

unknown output state is denoted "?". Figure 4(a) shows its regular operation without any probe attached to `L1` or `L2`. Inputs `R` and `S` are alternately delayed such that `Q` alternates between 0 and 1 at the sampling time every clock cycle. In Figure 4(b), `L1` is probed, which induces an additional delay to `R`. In this case, `R` is *always slower* than `S`, such that `Q` stops alternating and keeps a constant value of 0.

2) Control and Evaluation Logic: The control logic provides the multiplexer input `sel` to the LAPD. `sel` controls whether latch input `R` or `S` shall be delayed.

Figure 5 depicts a schematic of a sample control logic implementation. It is designed such that `sel` is generated by a toggle flip-flop clocked by a delayed, inverted clock signal. The rising edge of the T flip-flop clock `ff_clk` shall occur after the falling edge of the delayed LAPD latch input. An additional delay t_A ensures this condition.

On the output side of the latch, the evaluation logic shall provide feedback about the absence or presence of a probe. Conceptually, this is a PASS/FAIL signal where PASS means that `Q` toggles every cycle and FAIL indicates that `Q` remains at a constant value over two subsequent cycles. Implementing a single PASS/FAIL output line is dangerous, though: if an attacker would force such a line to a constant PASS, for example by the means of a second microbe, the LAPD would become obsolete.

The circuit as provided in Figure 6 has two redundant outputs `pass` and `fail` to avoid this single point of failure. It is fed by the signals `Q` and $\overline{\texttt{Q}}$ and uses the clock `ff_clk` coming from the control logic. As a positive side effect of the symmetry of the evaluation logic, both outputs of the LAPD latch are equally loaded, which avoids introducing a bias to the circuit.

3) Protection of Multiple Lines: So far, only the protection of two lines has been discussed. In order to protect a bus, it is necessary to extend the scheme to the protection of n symmetric bus lines.

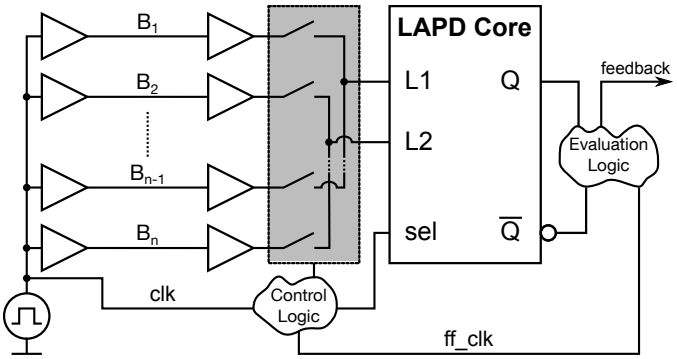


Figure 7. Bus protected by the LAPD

Using "switches" such as pass transistors, transmission gates or a combination of AND and OR gates, n lines can be protected by connecting $n/2$ lines to the `L1` input of the LAPD through such a gate, while connecting the other half to `L2`. Then, several delay comparisons are performed such that for each comparison, *one* of the bus lines is passed through to `L1` and *one other* bus line is passed through to `L2`. A schematic is depicted in Figure 7. With this approach, the LAPD protecting an n bit bus can detect up to $n-1$ attached probes.

A full probe detection coverage is obtained by verifying that the delays of all bus lines are equal. Due to the transitivity of equality, it is sufficient to perform a pairwise comparison of adjacent lines.

In practice, the length of bus lines is not exactly balanced and therefore, the comparison of two adjacent lines is assumed to be slightly biased. In the case this bias has a magnitude that affects the measurement accuracy, it can be compensated by fine-tuning the individual line delays t_D: Instead of having one constant delay t_D for all bus lines, an individual t_{Di} can be used for each bus line.

C. System Integration Example

Given that the LAPD can take over control of the bus for a limited time, it can be attached to the bus of a microprocessor system just like any peripheral component. The CPU core can trigger a probe detection run, for example, by a read operation to the LAPD which would give the LAPD full access to the data bus until the LAPD signals the end of the read operation.

A probe detection run can be triggered during startup or prior to transferring critical information such as keys over the bus.

A top-level view of the LAPD integration into a low-power smartcard chip is shown in Figure 8.

IV. Simulation Results

We simulated the function of the LAPD on a STMicroelectronics 65nm technology using standard cells in the Cadence environment with *spectre*.

We aimed at obtaining the dependency between the delay t_D and the minimum capacity $C_{A,min}$ that can be detected. From that, the delay t_D shall be determined.

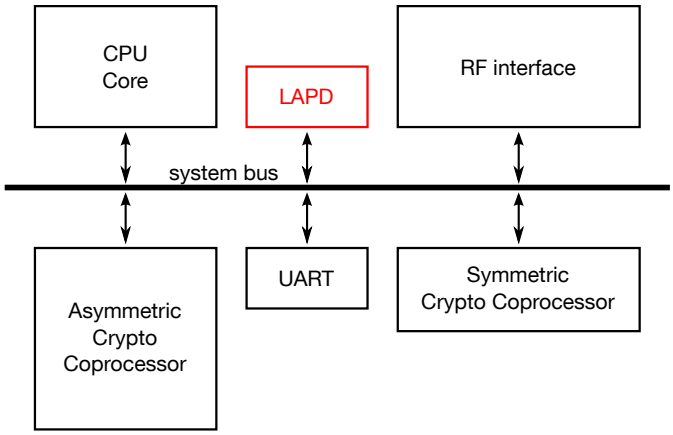


Figure 8. LAPD System Integration

t_D in ps	$C_{A,min}$ in fF	t_D in ps	$C_{A,min}$ in fF
10	5	160	25
20	5	170	25
30	5	180	30
40	5	190	30
50	10	200	30
60	10	210	35
70	10	220	35
80	15	230	35
90	15	240	35
100	15	250	40
110	20	260	40
120	20	270	40
130	20	280	45
140	20	290	45
150	25	300	45

Table I
MINIMUM DETECTED C_A DEPENDING ON t_D

The dependency between t_D and $C_{A,min}$ is determined by simulating a system as shown in Figure 1. For reasons of simplicity, the control logic is replaced by manually driving the `sel` input, while the evaluation logic is implemented in software that uses the *spectre* analog waveforms of `Q` as input data. Due to the symmetry of the RS latch, it is sufficient to simulate probing the line that is connected to the `L2` input of the LAPD. We assume probe capacitance values of $C_A \in \{0\,\text{fF}, 5\,\text{fF}, 10\,\text{fF}, \cdots, 60\,\text{fF}\}$. The LAPD itself is implemented according to Figure 2, but allows keeping the delay t_D variable. As an obvervation window, we chose values between 10 ps and 300 ps. Considering the line capacitance C_L, we assumed a value of 100 fF.

Table I shows $C_{A,min}$ in dependency of the delay t_D. A graphical representation of the detection coverage of probe attachments is shown in Figure 9. The x axis points out the configured delay t_D of the delay gate, while the y axis denotes the capacitance C_A of the attached probe. Blue circles point out undetected capacitive loads, which means that the detector output still behaves as normal, while white circles denote the successful detection of a probe attachment – on a technical level, this means that the LAPD outputs `Q` and $\overline{\texttt{Q}}$ stop alternating and keep at a constant value. From this figure, 10 fF can be spotted as the minimal value of C_A to be detected. The microprobe with the smallest input capacitance we found on the market [9] has an input capacitance value of at least 20 fF and therefore could be detected by the LAPD.

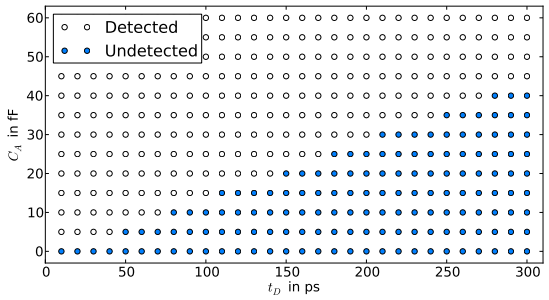


Figure 9. Nominal LAPD Detection Coverage

V. CONCLUSION

In this paper, we present the Low Area Probing Detector (LAPD), a new approach to detect microprobing on symmetric lines such as buses. It is the first detector measuring capacitances to detect tampering without relying on analog circuitry. This avoids large analog components, which makes the area required for the LAPD circuitry lower than for any other delay-based probe detection scheme.

According to our simulations, the LAPD detects state-of-the-art active microprobes with parasitic capatitances of 20 fF or less.

The scheme can be used to enhance the security of low-cost security controllers, as found on cheap mass market products such as SIM cards, but it is also possible to apply its concepts to improve the security of – already well-protected – high end security controllers, as they are found in Pay TV smart cards, for example.

As the LAPD increases the complexity for a successful bus attack, adversaries continue to look for other attack vectors. For an effective and comprehensive protection of security chips, other components need to be protected as well – this includes, for example, memory controllers, address decoders, control logic and arithmetic-logic units (ALUs), but also the signalling mechanisms of attack detectors themselves. Therefore, analyzing other microprobing attack targets and providing appropiate protection mechanisms appears as an important field for further work.

ACKNOWLEDGEMENTS

This work was partly funded by the Spanish research program TEC2010-18384 as well as by the German Federal Ministry of Education and Research (BMBF) in the project SIBASE through grant number 01S13020A.

REFERENCES

[1] O. Kömmerling and M. G. Kuhn, "Design Principles for Tamper-resistant Smartcard Processors," in *Proceedings of the USENIX Workshop on Smartcard Technology on USENIX Workshop on Smartcard Technology*, ser. WOST'99. Berkeley, CA, USA: USENIX Association, 1999, pp. 2–2. [Online]. Available: http://dl.acm.org/citation.cfm?id=1267115.1267117

[2] P. Maier and K. Nohl, "Low-Cost Chip Microprobing," 29th Chaos Communication Congress (29C3), 12 2012, accessed on 2014-01-16. [Online]. Available: http://events.ccc.de/congress/2012/Fahrplan/attachments/2247_29C3-Dexter_Nohl-Low_Cost_Chip_Microprobing.pdf

[3] R. Anderson, M. Bond, J. Clulow, and S. Skorobogatov, "Cryptographic Processors-A Survey," *Proceedings of the IEEE*, vol. 94, no. 2, pp. 357–369, 2006.

[4] M. Ling, L. Wu, X. Li, X. Zhang, J. Hou, and Y. Wang, "Design of Monitor and Protect Circuits against FIB Attack on Chip Security," in *Computational Intelligence and Security (CIS), 2012 Eighth International Conference on*, 2012, pp. 530–533.

[5] C. Tarnovsky, "Deconstructing a 'Secure' Processor," Blackhat DC, 2012.

[6] J. Krämer, D. Nedospasov, A. Schlösser, and J.-P. Seifert, "Differential Photonic Emission Analysis," in *Constructive Side-Channel Analysis and Secure Design*, ser. Lecture Notes in Computer Science, E. Prouff, Ed. Springer Berlin Heidelberg, 2013, vol. 7864, pp. 1–16. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-40026-1_1

[7] C. Helfmeier, D. Nedospasov, C. Tarnovsky, J. Krissler, C. Boit, and J.-P. Seifert, "Breaking and Entering Through the Silicon," in *Proceedings of the 2013 ACM SIGSAC Conference on Computer and Communications Security*, ser. CCS '13. New York, NY, USA: ACM, 2013, pp. 733–744. [Online]. Available: http://doi.acm.org/10.1145/2508859.2516717

[8] S. Manich, M. S. Wamser, and G. Sigl, "Detection of Probing Attempts in Secure ICs," in *Hardware-Oriented Security and Trust (HOST)*, 2012, pp. 134–139.

[9] "Picoprobe Model 18C & Picoprobe Model 19C," Datasheet, accessed on 2014-01-16. [Online]. Available: http://www.ggb.com/PdfIndex_files/mod18c.pdf

[10] H.-U. Buchmüller, "Security Target M7820 A11 and M11," August 2012, accessed on 2014-01-16. [Online]. Available: http://www.commoncriteriaportal.org/files/epfiles/0829b_pdf.pdf

[11] "Common Criteria for Information Technology Security Evaluation, Part 3: Security assurance components," 2012, https://www.niap-ccevs.org/Documents_and_Guidance/cc_docs.cfm, accessed on 25.08.2013.

[12] T. Sakurai and A. R. Newton, "Alpha-power law MOSFET model and its applications to CMOS inverter delay and other formulas," *IEEE Journal of Solid-State Circuits*, vol. 25, no. 2, pp. 584–594, 1990.

[13] K. A. Bowman, B. L. Austin, J. C. Eble, X. Tang, and J. D. Meindl, "A Physical Alpha-power Law MOSFET Model," in *Proceedings of the 1999 International Symposium on Low Power Electronics and Design*, ser. ISLPED '99. New York, NY, USA: ACM, 1999, pp. 218–222. [Online]. Available: http://doi.acm.org/10.1145/313817.313930

[14] A. Balankutty, T. C. Chih, C. Y. Chen, and P. Kinget, "Mismatch Characterization of Ring Oscillators," in *Custom Integrated Circuits Conference, 2007. CICC '07. IEEE*, 2007, pp. 515–518.