A Calibratable Detector for Invasive Attacks

Michael Weiner, Wolfgang Wieser, Emili Lupon, Georg Sigl, and Salvador Manich

***Abstract*— Microprobing is commonly used by adversaries to extract firmware or cryptographic keys from microcontrollers. We introduce the calibratable lightweight invasive attack detector (CaLIAD) to detect microprobing attacks. The CaLIAD measures timing imbalances between lines that are caused by the capacitive load of a probe. Compared to protection mechanisms from industry, it does not require an additional protection layer such as meshes do; in contrast to bus encryption, it does not introduce delay cycles. Compared to state-of-the-art low area probing detectors, it can be calibrated and, thus, allows compensating manufacturing variations as well as small layout imbalances. This capability allows us to significantly reduce the detection margin compared to the prior art while maintaining the low rate of false positives. We can finally show that capacitive loads of 23 fF or less can be detected, depending on how the CaLIAD is used. This includes all state-of-the-art commercial microprobes we are aware of.**

***Index Terms*— Data buses, digital integrated circuits, invasive attacks, microprobing, security, smart cards.**

I. Introduction

EMBEDDED security systems are often used in hostile environments. More than two decades ago, pay TV and telephone cards were attacked to get free phone calls and access to premium television programs. At present, the use cases of such systems are much more widely spread: contactless cards are used for various types of payment systems, secure microcontrollers control access to paid extra features of expensive products such as cars or test equipment, electronic keys serve as access control tokens, and smart cards are designed to store sensitive health information.

As the issuer of such devices cannot prevent adversaries from getting physical access to them, protection against physical attacks is a major design criterion.

Physical attacks can be classified into noninvasive, semi-invasive, and invasive attacks according to Skorobogatov [1]. Noninvasive attacks do not require decapsulation of the attacked integrated circuit. The attack methods include power and clock glitching and power analysis. In spite of the relatively low effort required by an attacker, they are quite powerful against unprotected circuits and systems. However, countermeasures against such attacks are already widely deployed in the field.

Semi-invasive attacks depend upon opening up the chip package but leave the passivation layer intact as they do not need electrical contact. Recent research results such as photonic emission analysis [2] have made them powerful also against well-protected circuits.

Invasive attacks need electrical contact; examples include focused ion beam (FIB) editing and attaching microprobes to lines containing secrets. Despite the fact that invasive attacks are quite powerful, they are also economic due to the existence of a second-hand market of laboratory equipment and the opportunity to pay FIB laboratories per hour. Invasive attacks have not only been analyzed in academia; instead, hackers have successfully performed complete memory dumps of hardened devices [3]. No successful invasive attacks are known against the next generation of hardened controllers that employ redundant cores and bus encryption to prevent fault injection and memory extraction through probing. While this approach has been able to prevent publicly known attacks in the past, it can be assumed to double the area and power consumption compared to unprotected implementations.

Wang *et al.* [4] expect microprobing assisted by FIB editing to gain importance due to their power compared to noninvasive and semi-invasive attacks, as well as due to the decreasing cost and better accessibility of laboratory equipment. A quite comprehensive survey of countermeasures in [4] and [5], where the advantages and drawbacks of various protection concepts, such as active shields or masking, are thoroughly discussed.

In this paper, we present a new invasive attack detector named calibratable lightweight invasive attack detector (CaLIAD) that combines the advantages of previous detectors and improves the detection sensitivity and reliability with respect to process, voltage, and temperature (PVT) variations.

This paper is structured as follows. Section II presents related previous work, Section III discusses the threat model, Section IV describes the concept behind the CaLIAD, Section V evaluates the results, Section VI shows a proof-of-concept implementation on a field-programmable gate array (FPGA), Section VII analyzes future work, and finally, Section VIII concludes this paper.

II. Related Work

Three different concepts exist as countermeasures against microprobing. Redundancy such as bus encryption or masking

Manuscript received June 29, 2018; revised October 9, 2018 and November 20, 2018; accepted December 31, 2018. Date of publication January 25, 2019; date of current version April 24, 2019. This work was supported by Spanish Government under Project TEC2013-J41209-P. *(Corresponding author: Michael Weiner.)*

M. Weiner is with the Chair of Security in Information Technology, Department of Electrical and Computer Engineering, Technical University of Munich, 80333 Munich, Germany, and also with the BMW Group, 80788 Munich, Germany (e-mail: m.weiner@tum.de).

W. Wieser is with Wieserlabs UG, 80687 Munich, Germany, and also with Optores GmbH, 80339 Munich, Germany (e-mail: priv@wieserlabs.com).

E. Lupon and S. Manich are with the Department of Electronic Engineering, ETSEIB, Universitat Politècnica de Catalunya, 08028 Barcelona, Spain (e-mail: emili.lupon@upc.edu; salvador.manich@upc.edu).

G. Sigl is with the Chair of Security in Information Technology, Department of Electrical and Computer Engineering, Technical University of Munich, 80333 Munich, Germany, and also with the Fraunhofer Institute for Applied and Integrated Security, 85748 Garching, Germany (e-mail: sigl@tum.de).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVLSI.2019.2892408

TABLE I
QUALITATIVE COMPARISON OF THE CaLIAD WITH OTHER INVASIVE ATTACK COUNTERMEASURES

alternative countermeasure	CaLIAD advantages
meshes	detection of backside attacks very low circuit overhead
bus encryption	no latency
PAD [7]	no analog design flow needed no capacitor
LAPD [5]	no transistor fine-tuning required better detection margin/reliability

can be introduced to hide information, physical access to lines can be obstructed, and it is possible to detect intrinsic probing effects.

Redundancy can be implemented by splitting data into different independent shares as, for example, proposed in [6]. However, this approach has a complexity of $O(t^2)$ when protecting against t probes and does not include protection against fault injection by probes. Another option is using redundant cores with mutual integrity checking and bus encryption; while a major semiconductor manufacturer implemented this in their flagship security controllers, its overhead may make this type of countermeasure unattractive for low-cost and low-power applications.

With respect to obstruction, one can distinguish between passive meshes that simply cover lines of interest and active meshes that are tamper protected. However, they both require an additional metal layer and thus increase production costs and do not protect against backside probing.

A rather unexplored field in the domain of microprobing protections is called “Analog Shields and Sensors” in [4]. The general idea is detecting inherent effects of a probe rather than making target lines inaccessible or making the data unusable. As an example, Wang *et al.* attested the probe attempt detector (PAD) [7] to be unique in backside attempt detection while keeping the area overhead comparably low. The PAD still uses analog circuit elements, though, which are larger than digital components and which might not be readily available in some technologies. The low area probing detector (LAPD) presented by Weiner *et al.* [5] fills this gap by presenting a purely digital circuit measuring race conditions between buffers and capacitively loaded lines. However, the LAPD needs the protected lines to be carefully balanced, and its transistor dimensions to be fine-tuned to work correctly. Also, the inability to compensate the effects of process variation by calibration leads to a lower sensitivity than the PAD. A qualitative summary of existing countermeasures is given in Table I.

What we consider missing is a detection circuit that can compensate process variations and line length imbalances by calibration, and that does not require fine-tuning transistor dimensions to function correctly, but that is still based on only digital components. Such a detector could be integrated into the design process of a digital IC much more smoothly than the LAPD while exhibiting robustness against PVT variations as well as enough sensitivity to detect state-of-the-art microprobes.

III. THREAT MODEL

We assume that an attacker with physical access to a security device aims at extracting its firmware or secrets such as cryptographic keys. The attack vectors offer a broad variety such as protocol level attacks, power analysis, glitching, localized electromagnetic analysis, laser fault injection, and microprobing.

We consider microprobing to be especially worth protecting against, as the second-hand laboratory equipment is available at low cost [8], and it was shown that microprobing can be used for full firmware dumps of high-security microcontrollers [3]. Furthermore, the attack is generic in that once an attacker can dump the memory contents from a specific microcontroller, he/she can access the secrets of all applications using that controller without repeatedly having to deal with application-specific attack details such as the timing of a fault attack.

An attacker uses a probe station together with micropositioners and microprobes to conduct a microprobing attack. A probe station consists of a microscope with high magnification and a flat surface for the micropositioners surrounding the stage with the device under attack.

The micropositioners are attached with a magnetic or vacuum fixation, and they are used to mount the actual microprobe; micrometer screws allow adjusting its position and placing it onto the line of interest. Passive microprobes consist of a tungsten tip that makes electrical contact to the tested line as well as an electrical connection to the measurement device (e.g., oscilloscope or logic analyzer). In addition, active microprobes contain an amplifier circuitry that significantly reduces distortion of the probed lines. This is especially helpful to probe internal signals such as bus lines as their drivers are not designed to be strong enough to drive the wiring between tungsten tip and measurement device. However, their influence on the circuit is still visible; microprobe manufacturers usually model the parasitics of the tungsten tip and amplifier circuitry as a capacitive load and a leakage current.

The microprobe model we found to have the smallest parasitics is the model “Picoprobe 18C/19C” by GGB, Inc. [9]. It exhibits an input capacitance of 20 fF if the specified transition time constraints are fulfilled, and a leakage current of 10 fA. While simulations have shown the effect of the leakage current to be negligible, the additional capacitive load on a bus line can be detected by precise timing measurements.

The tungsten wire connecting to the probe has a thickness in the range of 50 μm; its tip is sharpened to approximately 3 μm [9] by means of chemical etching. The small dimensions make the positioning of the probe fragile; in some occasions, for example, when probing deeply buried lines from the front side or when the exposed surface of the target wire is small, one can use an FIB to make an L-shaped metal to allow a stable mechanical and electrical contact. Fig. 1 shows this for a sample case of two lines.

As provably secure implementations often carry a huge resource overhead or are based on unstable assumptions, we rather focus on practical limitations of microprobing attacks. We claim that the limited space for micropositioners on a probe station and the increasing fragility of the

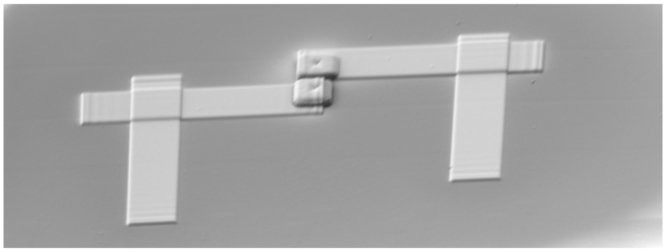


Fig. 1. Supporting L-shape deposits to probe two lines.

measurement setup with an increasing number of lines probed in parallel make simultaneous probing of multiple lines difficult in practice. For this reason, Tarnovsky [3] attacked a security microcontroller only using two probes at a time: one probe was used to inject faults, and the other probe was placed on one of the eight data bus lines, such that the data capturing had to be repeated eight times. The captured bus traces were merged during postprocessing [3].

Furthermore, we assume that the placement and positioning of a probe on the top of a target line are a fragile process that would disturb the circuit operation when done at runtime. If we consider that the probe detection circuit we present is triggered immediately prior to security critical operations (e.g., loading of cryptographic keys), an attacker would need the short time frame between detection run and target operation to place the probe (i.e., tens or hundreds of microseconds), which we supposed to be impractical.

IV. Calibratable Lightweight Invasive Attack Detector

A microprobe touching a line behaves as a capacitive load; see the Appendix for a comprehensive analysis of the parasitic effect. This makes transitions less steep and consequently delays the transitions in the connected gates. When one line out of a set of symmetric lines such as a bus is probed, a timing difference between the probed and the unprobed lines can be observed. This timing difference can be measured using a time-to-digital converter (TDC).

Similar to previous approaches [5], [7], [10], this can defend against probing at most $B-1$ lines when protecting a bus that consists of B lines. As described in Section III, we claim that to be sufficient as it is practically difficult to probe many lines simultaneously [5]. In some cases, e.g., when protecting serial buses, it is desirable to protect all B lines. In this case, an additional reference line can be added; however, it shall be dimensioned with a slightly different driver size, such that probing the reference line will change the timing behavior differently than probing the productive line.

A. Subgate Delay Measurement Circuits

Henzler [11] gives an overview of different types of TDCs; we are especially interested in circuits that allow subgate delay measurements because they allow a new concept of calibratable probing detectors that we are about to present here.

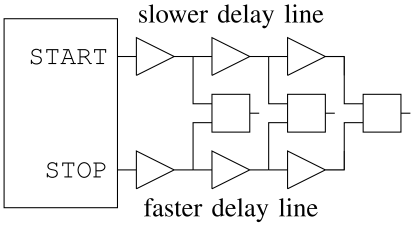


Fig. 2. Vernier delay line block diagram.

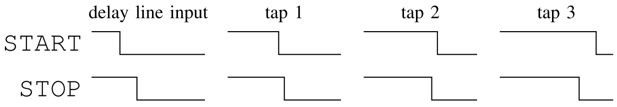


Fig. 3. Vernier delay line timing.

There are three types of subgate delay measurement concepts presented in [11]: local passive interpolation TDCs use voltage dividers between inputs and outputs of driver stages to reach below the resolution of an inverter delay. Its structure is comparatively resilient to manufacturing variations. However, it requires differential signals and stages that cause a significant area and power overhead.

Another idea to measure timings smaller than inverter delays is pulse shrinking delay lines [11], [12]. They use chains of unbalanced inverters with differing rise and fall times. A pulse applied to the input of a pulse shrinking delay line will be reduced in duration along the line before it eventually disappears completely. Its disadvantage is the need for gates with an unbalanced dimensioning of pMOS and nMOS transistors that might not be available in all technology libraries.

The Vernier delay line (VDL) has a comparatively small area footprint but shows inferior resilience against manufacturing variations when it comes to reliability of precise time measurements. This does not pose a problem when the goal is detecting deviations from a calibrated timing behavior.

We chose to base a probe detector on VDL: it has a comparably low area and power overhead and its bad performance with respect to manufacturing variations is not an issue in the use case of detecting microprobes. This effect can be compensated by storing calibration values as described later.

B. VDL Principle of Operation

A block diagram of the VDL is shown in Fig. 2. The device under test generates two output signals `START` and `STOP` that are passed through two differently dimensioned delay chains with different propagation delays. Both signals contain one transition; at the beginning of the line, the `START` transition occurs before the `STOP` transition. However, it propagates slower through the delay line than the `STOP` signal and eventually gets passed. Arbiters connected to each tap of the delay lines determine at which of the two taps the transition comes first. An example timing is shown in Fig. 3. In this figure, `STOP` passes `START` between tap 1 and tap 2.

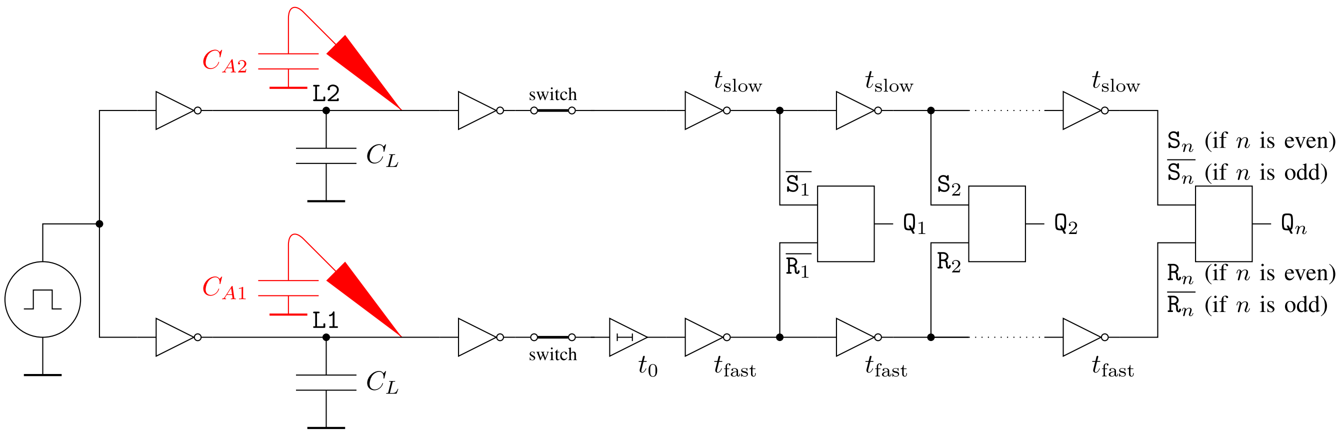


Fig. 4. Schematic of CaLIAD.

C. Delay Model

To model the delay that an attack capacitance C_A causes, we consider the case of two lines for reasons of simplicity. It was shown in [10] how to apply the detection concept to buses with more symmetric lines.

We assume to have two lines L1 and L2 that both have an intrinsic capacitance C_L, as depicted in Fig. 4. In addition, attack capacitances C_{A1} and C_{A2} are attached to L1 and L2, respectively,

$$C_{\text{L1}} = C_L + C_{A1} \tag{1}$$
$$C_{\text{L2}} = C_L + C_{A2}. \tag{2}$$

Assuming full-swing signals, we can estimate the propagation delay of a buffer using the alpha power model [13], [14] as follows:

$$t_{\text{P}} = \tilde{k}\frac{C\ V_{\text{DD}}}{(V_{\text{DD}} - V_t)^{\alpha}} \tag{3}$$

where α is the velocity saturation coefficient of the charge carriers, V_t is the threshold voltage of the transistors, and $\tilde{k}$ is the transresistance accumulating the other transistor parameters. The capacitive load at the output is denoted C and the supply voltage is called V_{DD}.

Using (3), we can approximate the overall delay of a bus line by summarizing the transistor parameters and supply voltage as Ω

$$t_{\text{L}i} = \Omega \cdot C_{\text{L}i}. \tag{4}$$

D. Probe Detection Concept

We adapt the use case of a VDL by evaluating one edge of a test signal; at line L2, this signal is passed through a delay line with slow incremental delay t_{slow} per stage, and at line L1, it is passed through a line with an initial delay t_0 and a fast incremental delay t_{fast} per stage, as shown in Fig. 4.

The two incremental delays t_{fast} and t_{slow} of the chain can be described as a common delay t_{C} and a delay difference Δt between the slow and the fast stage

$$t_{\text{fast}} = t_{\text{C}} \tag{5}$$
$$t_{\text{slow}} = t_{\text{C}} + \Delta t. \tag{6}$$

The transition at the faster line with initial delay passes the transition at the slower line between positions k and $k+1$ if the following equations hold:

$$t_{\text{L1}} + t_0 + k \cdot t_{\text{fast}} > t_{\text{L2}} + k \cdot t_{\text{slow}} \tag{7}$$
$$t_{\text{L1}} + t_0 + (k+1) \cdot t_{\text{fast}} < t_{\text{L2}} + (k+1) \cdot t_{\text{slow}}. \tag{8}$$

To insert C_{A1}, C_{A2}, and Δt into the inequalities, we solve (7) and (8) for the delay difference $t_{\text{L1}} - t_{\text{L2}}$ and use (4) to replace it by a capacitance difference. We use (1) and (2) to solve for $C_{A1} - C_{A2}$, then solve (5) and (6) for Δt and eventually insert the result into the transformed inequalities

$$k \cdot \Delta t - t_0 < \Omega \cdot (C_{A1} - C_{A2}) < (k+1) \cdot \Delta t - t_0. \tag{9}$$

It will be shown later that k can be stored as a calibration value. Therefore, we call the concept CaLIAD: on the one hand, calibration is possible, as it is with PAD [7], but which is not possible for LAPD [5]. On the other hand, it avoids large analog circuit components such as the PAD tank capacitor and only needs digital circuit elements that are comparably lightweight.

E. Circuit Realization

The schematic of the CaLIAD is shown in Fig. 4. The schematic is illustrated for a case study bus of two lines, which is the minimum number of lines required for its correct operation.

From left to right, it contains a test signal source, bus drivers, the bus lines that are under attack, output buffers, the initial delay element t_0, and eventually the VDL. One VDL chain element contains one RS latch that acts as an arbiter and two inverters: one with a slower propagation delay (t_{slow}) and a second which is slightly faster (t_{fast}).

We recommend to introduce the ability to switch off the evaluation circuitry, i.e., t_0 and the first stage of t_{slow}, when not in use. This is shown in the circuit diagram as “switch.” Being able to switch off the circuitry saves energy on the one hand, and on the other hand, it prevents attackers from probing the end of the delay to read signals from the protected lines.

Note that we tap the VDL after each inverter rather than after each buffer which is the usual approach found in the

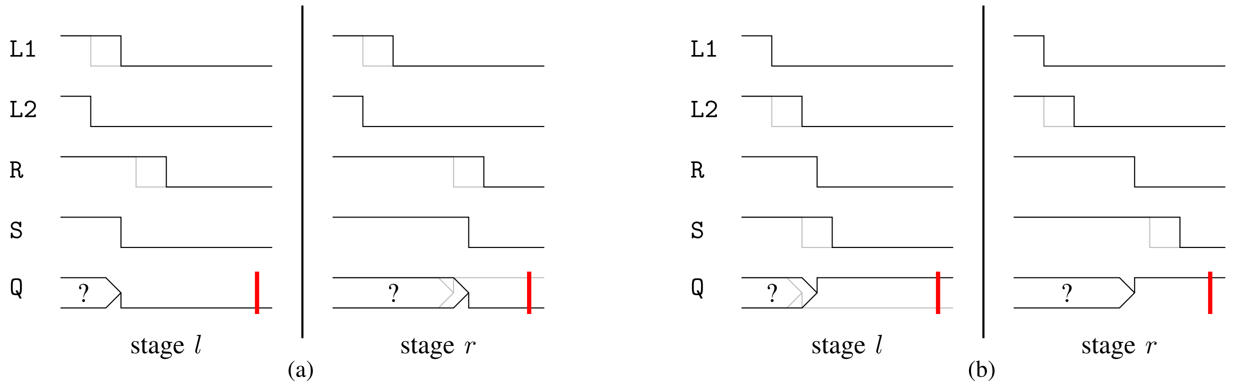


Fig. 5. CaLIAD timing. Gray signals: unprobed case. Red bars: latch output sampling time. (a) L1 probed. (b) L2 probed.

literature [11]. This optimization allows us to reduce the number of inverters in the VDL part by a factor of 2 but requires proper handling by:

1) either evaluating the rising edge of a test signal at even chain element positions k, and evaluating the falling edge of a test signal at odd chain element positions k, such that each arbiter can have high active inputs;
2) or alternating between different arbiter types: NAND RS latches have low active inputs and can be used at odd chain element positions, whereas NOR RS latches have high active inputs and can be used at even chain element positions.

We consider both optimization alternatives to be equally effective and concentrate on the latter option to keep simulation time low, as only one transition needs to be simulated instead of two.

Considering (9), the outputs $\mathtt{Q}_i$ are 0 if $i \leq k$ and 1 if $i \geq k+1$. Without attack, the transition from 0 to 1 will occur at positions k_0 and $k_0 + 1$. To check for an attack, we select one element left of that position that we call $l \leq k_0$ and a right element $r \geq k_0 + 1$. The following equation holds for the outputs of the selected elements:

$$(\mathtt{Q}_l, \mathtt{Q}_r) = \begin{cases} (0, 1), & \text{no probe attached} \\ (0, 0), & \text{probe at } \mathtt{L1} \\ (1, 1), & \text{probe at } \mathtt{L2}. \end{cases} \tag{10}$$

The timing of the CaLIAD is shown in Fig. 5 for an illustrative chain of two taps (say left and right taps). Fig. 5(a) highlights the case of probing L1, while Fig. 5(b) focuses on the effect of probing L2. The gray curves represent the behavior without a probe in both figures.

If more than two VDL chain elements are used, we get a thermometer code as a response as shown

$$(\mathtt{Q}_1, \mathtt{Q}_2, \ldots, \mathtt{Q}_k, \mathtt{Q}_{k+1}, \ldots, \mathtt{Q}_{n-1}, \mathtt{Q}_n) = \mathtt{00 \cdots 01 \cdots 11}. \tag{11}$$

The position $k + 1$ where the first 1 occurs will be referred to as position of first 1 (PF1).

F. Calibration

In the ideal case, i.e., assuming constant temperature and voltage as well as the absence of manufacturing variations, two chain elements are sufficient, such that $l = 1$ and $r = 2$ would hold.

In a real case, more than two chain elements are required: as explained later, we have seen that manufacturing variations bias the circuit in the same order of magnitude as the small scale probes that we want to detect. When no probe is attached, i.e., $C_A = 0$ fF, the transition occurs between a position k_0 and $k_0 + 1$, where k_0 will vary from device to device. This position can be stored as a calibration value to compensate manufacturing variations. It is required to perform $B - 1$ comparisons and thus store $B - 1$ calibration values to protect a B line bus.

Depending on the sensitivity of the design, that position may shift a small number of steps when varying voltage and temperature or under the presence of noise. To compensate that, an additional safety margin m is defined to avoid false positives: if $(\mathtt{Q}_{k_0}, \mathtt{Q}_{k_0+1}) = (0, 1)$ holds for a certain calibrated position k_0, we can skip m steps to the left and to the right of that position and configure the detector such that the actual evaluation takes place at the two latch outputs l and r as follows:

$$l = k_0 - m \tag{12}$$
$$r = k_0 + m + 1. \tag{13}$$

Voltage and temperature variations cannot be compensated using this method, but they only have second-order effects due to the differential nature of the circuit.

Another effect that cannot be completely compensated is a variation of the delay difference Δt between the slow and the fast inverters in the VDL chain, as described in [15].

The variation of Δt also implies that calibration accuracy is the best at the capacitive load C_A that is used for calibration and decreases with increasing distance to the calibration capacitance. The proposal above calibrates at $C_A = 0$ fF. The C_A values of which we expect most accuracy, though, are those that determine the transition between "no alarm" and "alarm."

Therefore, we propose a second calibration method using a reference capacitance C_{ref} at each protected line. We call this method two-point calibration and expect it to further increase the detection performance compared to the previously proposed single-point calibration method. It might not be

TABLE II
TWO-POINT CALIBRATION EXAMPLE

C_{L1}	C_{L2}	$\mathtt{Q}_1, \cdots, \mathtt{Q}_n$
$C_L + C_{\mathrm{ref}}$	C_L	`000`**`0`**`0000000000`**`0`**`11111`
C_L	C_L	`000`**`0`**`0000001111`**`1`**`11111`
C_L	$C_L + C_{\mathrm{ref}}$	`000`**`1`**`1111111111`**`1`**`11111`

applicable in all use cases, though, because the required reference capacitances increase the area usage; furthermore, they might help an attacker to spot the probing targets of interest.

C_{ref} is the smallest capacitance that is supposed to raise an alarm when attached to any line. To determine its value, we first define a minimum alarm capacitance $C_{0.99}$ for which we want to raise an alarm with an estimated alarm probability of $p_A \geq 0.99$. Second, we use Monte Carlo simulations to determine $C_{\mathrm{ref}} \leq C_{0.99}$ such that the probability constraint is satisfied for all corners. This is necessary as we assume that during manufacturing, only one set of voltage and temperature corners can be selected for efficiency reasons, whereas an attacker can freely choose the combination of corners to make a successful attack more likely.

To perform calibration, C_{ref} is alternately connected to `L1` and `L2`. The left calibration position k_L is determined as the PF1 when C_{ref} is connected to `L2`. Accordingly, the right calibration position k_R is determined as the position of last 0, i.e., the immediate left neighbor of the PF1, when C_{ref} is connected to `L1`. An example is shown in Table II where k_L and k_R are highlighted in bold.

The evaluation can then take place at k_L and k_R

$$l = k_L \tag{14}$$

$$r = k_R. \tag{15}$$

Note that (10) still applies.

G. Design Considerations

The performance of CaLIAD implementations can be evaluated based on different criteria. This section gives an overview of different such criteria and concludes with general design considerations.

The minimum detectable capacitance difference $\Delta C_{A,\mathrm{min}}$ determines the most "negative" capacitance difference, i.e., excess of attack capacitance attached to `L2` compared to `L1` that can be distinguished from range excess. When its value is applied, only the first VDL chain element is zero: $(\mathtt{Q}_1, \mathtt{Q}_2, \ldots, \mathtt{Q}_n) = \mathtt{01}\ldots\mathtt{1}$. From (9), one can conclude that the following inequality must hold in this case:

$$\Delta t - t_0 < \Omega \cdot \Delta C_{A,\mathrm{min}} < 2 \cdot \Delta t - t_0. \tag{16}$$

From this inequality, it can be seen that the value of $\Delta C_{A,\mathrm{min}}$ is determined by the initial delay t_0 as well as delay difference Δt between the slow and the fast inverter in the VDL chain. In the first place, this metric may not seem directly helpful—the CaLIAD only needs to give a pass/fail result and does not require providing quantitative information about the attached probe. However, it provides a lower bound of ability to calibrate the CaLIAD: upon calibration, a too high value of $\Delta C_{A,\mathrm{min}}$ can lead to a left stage $l < 1$ that refers to a nonexisting VDL chain element.

Similarly, the maximum detectable capacitance difference $\Delta C_{A,\mathrm{max}}$ determines the maximum excess of attack capacitance attached to `L1` compared to `L2` that can be distinguished from a range excess. When its value is applied, only the last VDL chain element is one: $(\mathtt{Q}_1, \ldots, \mathtt{Q}_{n-1}, \mathtt{Q}_n) = \mathtt{0}\ldots\mathtt{01}$. One can conclude from (17) that the value of $\Delta C_{A,\mathrm{max}}$ is determined by t_0, Δt as well as the total number of chain elements n

$$(n-1) \cdot \Delta t - t_0 < \Omega \cdot \Delta C_{A,\mathrm{max}} < n \cdot \Delta t - t_0 \tag{17}$$

where $\Delta C_{A,\mathrm{max}}$ determines the upper bound of ability to calibrate, as devices with a right calibration position $r > n$ refer to a nonexisting VDL stage and, therefore, cannot be calibrated.

Ideally, a symmetric detection range is desired, i.e., $\Delta C_{A,\mathrm{max}} = -\Delta C_{A,\mathrm{min}}$. This implies that $C_{A1} - C_{A2} = 0$ fF corresponds to a centered transition position $(\mathtt{Q}_1, \ldots, \mathtt{Q}_{(n/2-1)}, \mathtt{Q}_{(n/2)}, \ldots, \mathtt{Q}_n) = \mathtt{0}\ldots\mathtt{01}\ldots\mathtt{1}$.

Another important metric is the detector sensitivity, i.e., the extent of PF1 shift depending on the shift of the attack capacitance. It is determined by Δt. Decreasing Δt increases the sensitivity, which allows to compensate manufacturing variations more accurately but also implies longer chains, as shown in (16) and (17). Furthermore, circuit limitations impose a lower bound on Δt: it must be considered that the latches need a minimum time difference $\Delta t_{\mathtt{RS},\mathrm{min}}$ between the edges of $\mathtt{R}_i$ and $\mathtt{S}_i$ to avoid metastability and get a reliable output $\mathtt{Q}_i$. This has an implication on Δt between two VDL stages: while metastability cannot be avoided for *one* latch output, a lower bound $2 \cdot \Delta t_{\mathtt{RS},\mathrm{min}} < \Delta t$ shall be fulfilled to avoid that two or more latches become metastable. Otherwise, PF1 cannot be uniquely determined. This effect is called bubbles. An example is shown as follows:

$$(\mathtt{Q}_{k-3}, \mathtt{Q}_{k-2}, \mathtt{Q}_{k-1}, \mathtt{Q}_k, \mathtt{Q}_{k+1}, \mathtt{Q}_{k+2}, \mathtt{Q}_{k+3}) = \mathtt{0\ 010\ 011}. \tag{18}$$

Calibration would still be possible under the presence of bubbles: in the case of single-point calibration, one can use the middle between the position of the last consecutive 0 and the first consecutive 1 to estimate PF1. Obviously, m would have to be chosen such that l and r are outside the bubble region. The occurrence of bubbles indicates a limit in sensitivity. In that case, other measures would need to be taken to extend the sensitivity, e.g., an optimization of the used latches.

Other important metrics are area usage and power consumption of the CaLIAD. Aiming for good performances with respect to sensitivity and range of detectable capacitances will deteriorate the area and power performances.

H. System Integration

The concept as it was presented is able to protect two lines. This can be extended to B lines by either multiplexing the two lines onto the two detection lines, or by duplicating the detection circuitry, which is still feasible due to the small size of the CaLIAD.

The main use case we see is parallel buses in processor cores, which have a width of 8, 16, 32, or 64 bits, as these were the attack targets in real-life microprobing attacks. However, this circuit can be useful in other application areas as well, e.g., serial buses, network-on-chip platforms, or active shields on security controllers.

When protecting a parallel bus, one can distinguish between *offline* and *online* detection. In offline detection, the bus is disconnected from productive operation during a check. This can be accomplished by a bus master generating test signals and a bus slave that performs the evaluation. A more detailed description of offline detection is given in [5].

However, every situation in which two or more bus lines encounter simultaneous falling transitions can be used to perform a test. The concept of online detection makes use of this: instead of generating a dedicated test signal, the CaLIAD can run a test whenever it detects a suitable pattern on the bus. That can decrease latency and ensure continuous operation. While the data on the bus are not generally optimized for detection coverage, a dedicated sequence of data values can be placed on a bus before security-critical operations to ensure that all lines are tested.

V. Simulations and Results

CaLIAD was implemented on a 65-nm application-specified integrated circuit (ASIC) technology from STMicroelectronics. All circuit elements are based on the low-power standard threshold voltage transistors psvtlp and nsvtlp. Cadence spectre 11.1.0 was used to simulate the circuit on a workstation with four AMD Opteron 6274 CPUs and 256 GB of RAM. We automated the simulation sweeps using SALVADOR [16].

We assumed an ambient temperature of 27 °C and an intrinsic line capacitance of $C_L = 100$ fF, which corresponds to an approximate line length of 1.3 mm on the top layer, assuming a parallel GND line with the minimum distance.

A. Effects of Manufacturing Variations

We created the CaLIAD with 30 chain elements. The fast inverters in the chain elements were implemented with an aspect ratio of $(W/L) = 20$, the slow inverters had an aspect ratio of $(W/L) = 8$. All other circuit elements were given a default aspect ratio of $(W/L) = 10$. The initial delay element was constructed as a chain of four inverters to have a sufficient margin at the left end of the chain.

We expect to see a shift in PF1 depending on the attached attack capacitance difference ΔC_A. Our simulations are limited to the simple case of probing either L1 or L2, therefore, assuming $C_A = \Delta C_A$. In order to merge the results, we define that negative values of C_A mean $|C_A|$ being attached to L2 while C_L remains constant for both lines

$$\begin{cases} C_A > 0 \Rightarrow C_{\mathrm{L1}} = C_L + C_A & \wedge C_{\mathrm{L2}} = C_L \\ C_A < 0 \Rightarrow C_{\mathrm{L1}} = C_L & \wedge C_{\mathrm{L2}} = C_L + |C_A|. \end{cases} \tag{19}$$

We performed $N = 2000$ Monte Carlo simulations for each sweep point $C_A \in [-40\text{fF}; 40\text{fF}]$ with increments of 1 fF. The simulations included process (per-chip) and mismatch (per-transistor) variations.

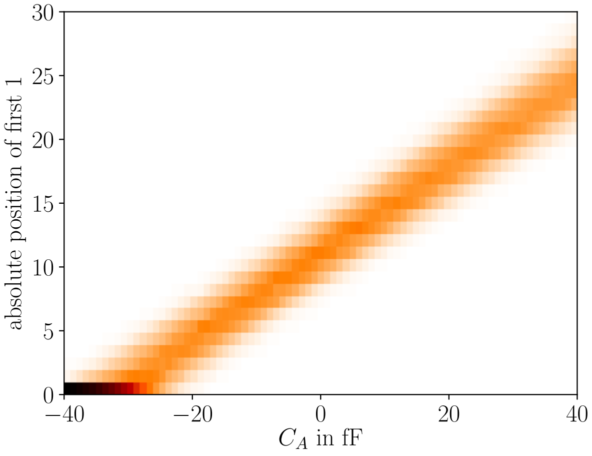


Fig. 6. Distribution of transition positions without calibration. Positive values of C_A represent a probe attached to L1 and negative values represent a probe attached to L2.

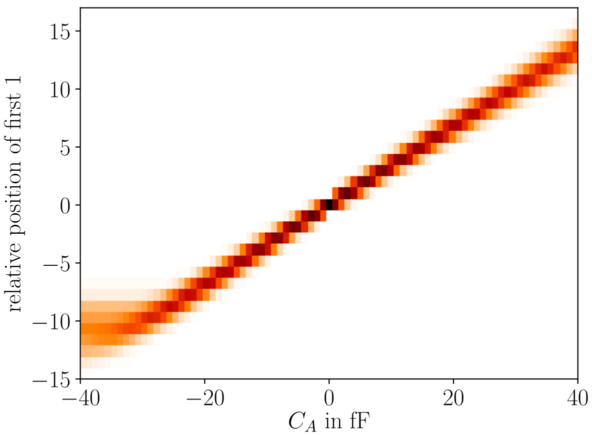


Fig. 7. Distribution of transition positions with single-point calibration at $C_A = 0$ fF.

Fig. 6 shows the distribution of the absolute PF1 in the chain as a function of the attack capacitance C_A. The larger C_A is, the further PF1 moves to the right. As described in (19), positive values of C_A represent a probe attached to L1 and negative values refer to a probe attached to L2. The outlier at the bottom left of Fig. 6 is caused by large absolute C_A values that already reach the beginning of the line, i.e., $Q_1 = 1$. In this case, PF1 cannot move further to the left such that all further increases of the absolute value of C_A accumulate in this point.

With our implementation, we did not notice any bubbles in the response of the VDL.

We can compensate the random offset of each device instance by subtracting PF1 at 0 fF for each device instance. Fig. 7 shows that the variation is significantly reduced by this calibration. However, one can also see that the variation increases with increasing distance to the calibration point. As explained before, this is due to the variation of Δt that cannot be compensated with the single-point calibration. However, Fig. 7 suggests that using calibration at the actual points of

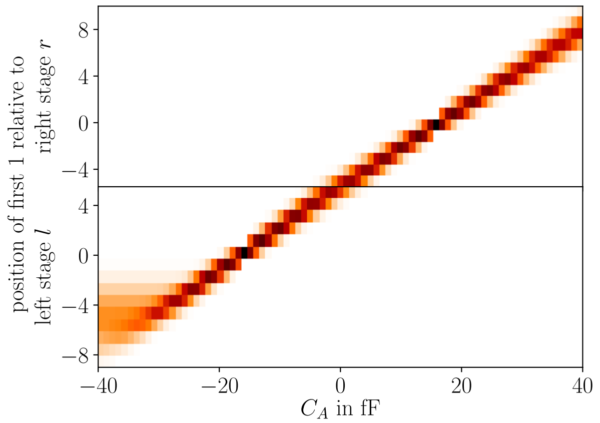


Fig. 8. Distribution of transition positions with two-point calibration ($C_{\text{ref}} = 16$ fF).

TABLE III

NOMINAL CORNER DETECTION PERFORMANCE OF DIFFERENT CALIAD CALIBRATION METHODS

calibration method	$C_{0.01}$	$C_{0.99}$	ΔC	$n_{\min}$
single-point ($m = 4$)	9 fF	18 fF	9 fF	23
single-point ($m = 5$)	11 fF	21 fF	10 fF	24
two-point	11 fF	16 fF	5 fF	25
LAPD [5]	16.4 fF	34.4 fF	18 fF	–
optimized LAPD [17]	6.4 fF	20.4 fF	14 fF	–

interest—values of C_A at which the transition between "no alarm" and "alarm" is desired to occur—can further reduce the noticed variation.

This effect is shown in Fig. 8. The bottom half shows PF1 relative to the left calibrated stage with index l and the top half shows PF1 relative to the right calibrated stage with index r. One can see a reduction of noise toward the extremes of C_A; more importantly, the PF1 becomes more accurate in the surrounding of the desired alarm threshold of 20 fF.

B. Detection Performance

Table III shows the detection boundaries of different CaLIAD calibration methods at the nominal corner, i.e., at a temperature of 27 °C and a supply voltage of 1.2 V. For the single-point calibration, the safety margins $m = 4$ and $m = 5$ are shown due to their proximity of the detection threshold to 20 fF, which is the smallest capacitive load of a commercial microprobe that we could find [9]. $C_{0.01}$ denotes the smallest capacitance with an estimated alarm probability of $p_A \leq 0.01$ at either of the lines `L1` or `L2`; accordingly, $C_{0.99}$ refers to the minimum capacitance with an estimated alarm probability of $p_A \geq 0.99$.

We used the Wilson method [18] to estimate the probability bounds using $N = 2000$ Monte Carlo iterations and assuming a confidence level of $\alpha = 0.01$.

Stricter bounds would imply a significant increase in simulation time. For cases in which the estimated probability bound is not sufficient, one can use simulation methods such as statistical blockade [19], [20] that can significantly improve the efficiency when analyzing rare events. In addition, one can use majority voting by repeated measurements and/or circuit duplication to get significantly stricter probability bounds [5].

TABLE IV

WORST CASE CORNER DETECTION PERFORMANCE OF DIFFERENT CALIAD CALIBRATION METHODS

calibration method	$C_{0.01}^{\text{WCC}}$	$C_{0.99}^{\text{WCC}}$	ΔC
single-point ($m = 4$) *worst case at corner*	6 fF 0 °C, 1.08 V	20 fF 0 °C, 1.08 V	14 fF
single-point ($m = 5$) *worst case at corner*	9 fF 0 °C, 1.08 V	23 fF 0 °C, 1.08 V	14 fF
two-point *worst case at corner*	9 fF 0 °C, 1.08 V	20 fF 0 °C, 1.08 V	11 fF
LAPD [5]	12.0 fF	39.4 fF	27.4 fF

$\Delta C = C_{0.99} - C_{0.01}$ is the region at which the CaLIAD output is not reliable. Smaller values of ΔC mean better performance. One can see in Table III that the two-point calibration method outperforms the single-point calibration method approximately by a factor of two with respect to ΔC.

$n_{\min}$ is the minimum required chain length. It is determined such that the right VDL evaluation stage r is still within the chain length for all Monte Carlo instances, i.e., $r \leq n_{\min}$ always holds.

The CaLIAD represents a significant improvement of detection performance compared to previous probing detectors. For example, the best previously published detector [17] exhibits an uncertainty region of $\Delta C = 14$ fF, whereas the CaLIAD has a worst case uncertainty region of 10 fF; when using two-point calibration, we were able to achieve 5 fF.

Another advantage of the CaLIAD is its flexibility with respect to calibration: when using single-point calibration, the detection thresholds can be shifted by adjusting the safety margin m; in the case of two-point calibration, the reference capacitance C_{ref} can be chosen according to the requirements.

We chose C_{ref} such that 20 fF is detected at all corners: the best microprobes with respect to input capacitance that we found were GGB Model 28/29 [21] with an input capacitance of 40 fF, as well as GGB Model 18C/19C [9] with an input capacitance between 20 and 60 fF depending on the transition time of the measured signals. Therefore, we consider the absolute worst case attack capacitance to be 20 fF.

C. Corners

We performed an analysis of voltage and temperature corners to determine the worst case corner values of $C_{0.01}^{\text{WCC}}$ and $C_{0.99}^{\text{WCC}}$ with respect to varying environmental conditions. The used corners were $\vartheta \in \{0\ °\text{C}, 27\ °\text{C}, 85\ °\text{C}\}$ and $V_{\text{DD}} \in \{1.08\text{V}, 1.2\text{V}, 1.32\text{V}\}$. This gives results comparable to the LAPD corners [5]. To obtain the threshold capacitances, we first performed a calibration at the nominal corner $\vartheta = 27$ °C and $V_{\text{DD}} = 1.2$V. Then, we looked at each corner and selected the worse of the two capacitance values of either probing `L1` or `L2`. Eventually, we took the worst case values of all nine corners. These are given in Table IV.

We can observe that the uncertainty region increases by at most 6 fF. Note that the two-point calibration value of $C_{0.99}^{\text{WCC}}$

TABLE V

AREA, TIMING, AND ENERGY COMPARISON OF CRYPTOGRAPHICALLY SECURE SHIELDS, PAD, LAPD, AND CALIAD

	area a [GE]	cycles	energy [fJ]
CSS [22]	8081	–	$7.01 \times 10^{12}\,\mathrm{s}^{-1}$
PAD [7]	549	50-100	n/a
LAPD [5]	48	2	981
optimized LAPD [17]	114	2	1311
CaLIAD	352	1	1154

TABLE VI

AREA OF THE CALIAD ELEMENTS IN GATE EQUIVALENTS

a_{drivers}	4
a_{t_0}	8
a_{fast}	4
a_{slow}	1.6
a_{latch}	8

is our desired target value that was used to determine the reference capacitance C_{ref}.

D. Resource Usage

Table V shows the comparison of the area usage, timing, and energy consumption of the CaLIAD with other probing detection circuits.

The cryptographically secure shields (CSSs) [22] aim at covering large parts of a chip with an active shield using Advanced Encryption Standard as a random pattern generator. Similar to the CaLIAD, PAD [7] and LAPD [5], [10] detect timing imbalances of symmetric lines, but they use different measurement concepts.

The area a in gate equivalents (GEs) can be expressed as

$$a_{\mathrm{total}} = a_{\mathrm{drivers}} + a_{t_0} + n \cdot (a_{\mathrm{fast}} + a_{\mathrm{slow}} + a_{\mathrm{latch}}) \qquad (20)$$

where a_{drivers} denotes the area of the driving stages before and after the bus, a_{t_0} represents the area of the initial delay element t_0, and a_{fast}, a_{slow}, and a_{latch} represent the area of one CaLIAD stage element.

Each element in (20) is computed as follows:

$$a = \frac{1}{A_{\mathrm{ref}}} \left(L^2 \cdot \frac{W}{L} \cdot t \cdot (1+s) \right) \qquad (21)$$

where A_{ref} is the absolute area of the smallest NAND gate HS65_LS_NAND2X2, L is the channel length of the transistors, which is a constant of $L = 0.06\ \mu$m in our case, (W/L) is the nMOS transistor aspect ratio, which is used to tune the delay of the chain elements and has a default value of $(W/L) = 10$, t is the count of nMOS transistors of the stage under consideration, and s is a technology-dependent pMOS transistor aspect ratio scale factor, which is $s = 2.2$ in our case.

Table VI shows the area of the stages based on (21).

We assumed a chain length of $n = 25$, which is the minimum length for which calibration was possible at all Monte Carlo simulations for all described calibration methods and safety margins. Inserting the values from Table VI into (20) results in a total area of 352 GEs. The data of CSS and PAD as well as the LAPD area are taken from [5]. We reimplemented the LAPD and the optimized LAPD to get an improved energy estimation. For example, we noticed that while the LAPD needs two test cycles, simulations had ended after the first edge of the second cycle. We include the last edge to measure the energy of two full cycles. The dimensions of the optimized LAPD [17] were obtained by inserting the raw transistor dimensions into (21). While no energy consumption is available for the PAD, we can expect it to be significantly higher as when compared to the CaLIAD as the PAD requires 50 to 100 cycles in which a large capacitor is charged. It is difficult to make a timing and energy consumption comparison with the CSS as the CSS is supposed to run continuously.

The CaLIAD is the second best with respect to the area after the LAPD implementations. Its power consumption is comparable to the LAPD. It has a superior detection performance and a faster response time compared to PAD and LAPD (the CSS are not directly comparable in these terms). Furthermore, the CaLIAD offers more parameters and, therefore, more degrees of freedom than the LAPD. We expect the CaLIAD to have a better benefit when it comes to systematic optimization.

VI. FPGA PROTOTYPE

We used an FPGA as a proof-of-concept implementation platform to demonstrate its functionality on real digital hardware. Furthermore, we performed a temperature sweep in a climate chamber to analyze its reliability over temperature changes.

A. Design

We used an Altera/Intel Cyclone III FPGA, model EP3C16F484C6. It is implemented in a 60-nm technology [23]. The majority of the 28 × 40 internal blocks of the FPGA are logic array blocks (LABs); other blocks implement memory and multipliers or are reserved. One LAB consists of 16 logic elements (LEs) and a local interconnect bus. Each LE contains a four-input lookup table (LUT) and a flip-flop.

The synthesis environment Quartus II allows the placement of components down to a granularity of LEs. Routing cannot be enforced, but a sparse regular architecture is a good practice to get balanced delays. Furthermore, postrouting simulations can be employed to estimate the timing behavior before testing hardware implementations.

As the CaLIAD is an asynchronous design taking advantage of timing properties, the following components must be designed with care.

1) It shall be possible to emulate an attack on either L1, L2, both lines, or to disable the attack. As it is not possible to emulate a real capacitive load on an FPGA, we decided to use an abstract model that introduces an additional small delay to the lines for which an attack is emulated. During early simulations, it was observed that different LUT inputs exhibit differences in internal propagation delay in the range between 61 and 122 ps without considering the output load. This is approximately in the same order of magnitude as the delay introduced by a probe in the ASIC technology we used. This makes

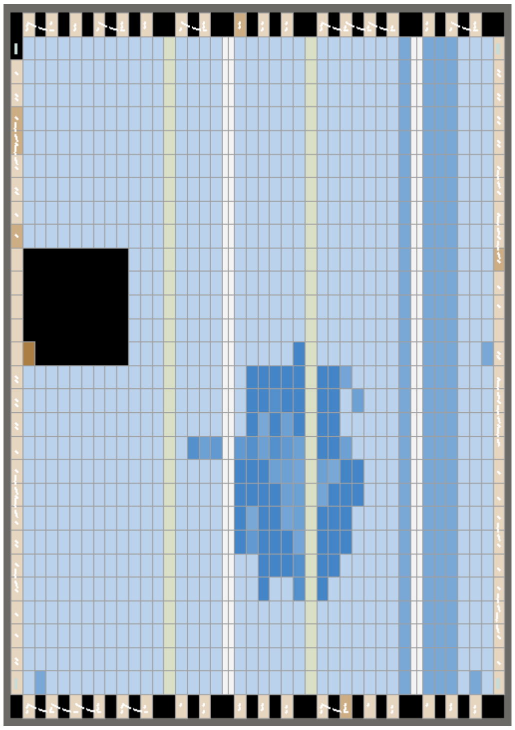


Fig. 9. Layout of CaLIAD FPGA implementation.

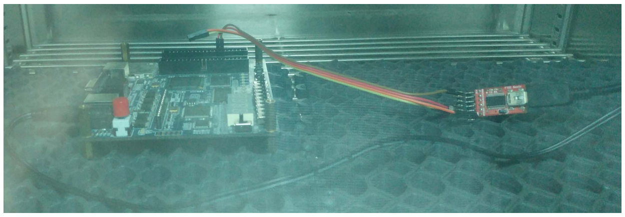


Fig. 10. FPGA in climate chamber.

TABLE VII

COUNTER VALUES CAPTURED AT 20 °C

k	emulated probe at line(s)			
	–	L1	L2	L1,L2
	count($\mathtt{Q}_k = 1$)			
1,2,...,11	0	0	10 000 000	0
12	10 000 000	0	10 000 000	10 000 000
13	11 593	0	10 000 000	4609
14	9 999 780	0	10 000 000	9 999 310
15	1	0	10 000 000	0
16	10 000 000	0	10 000 000	10 000 000
17	9 730 190	0	10 000 000	9 379 770
18	10 000 000	0	10 000 000	10 000 000
19	9 999 710	0	10 000 000	9 998 106
20,21,...,28	10 000 000	0	10 000 000	10 000 000

it possible to configure an LUT as a switchable delay element.

2) The initial delay t_0 can be implemented as a combination of a coarse and a fine delay. For the coarse delay, one can make use of the inherent routing delays between cells and by forcing signals to pass through additional LUTs acting as buffers. Fine delays can be implemented by additional dummy gates increasing the load and, hence, the interconnect delay.
3) The latches are implemented using an LUT with feedback from the output back to an input. Note that this approach results in an inherent imbalance of feedback paths compared to the ASIC implementation. Furthermore, simulations have shown that the interconnect and/or input load of the signals R and S (and, respectively, $\overline{\mathrm{R}}$ and $\overline{\mathrm{S}}$) is different.
4) There must be a delay difference Δt between the slow and fast inverters of the VDL chain. This can be accomplished by using the property of different LUT inputs having different propagation delays, by different interconnect delays of different LAB columns, and by introducing different loads at the inverter outputs. Simulations have shown that the different loads of the latch inputs can be used to accomplish different delays; this option turns out to exhibit the smallest $\Delta t \approx 5$ ps out of all three available options.

Four columns were used to implement the VDL chain; each row implements a chain element. As there exist 28 rows in the FPGA we use, we chose the chain that consists of 28 elements in total. The four columns are highlighted in darker blue on the right side of Fig. 9, which shows the layout of the CaLIAD implementation. From right to left, the columns implement the fast inverter chain, the latches, the slow inverter chain, and eventually a row of buffers that decouple the latches from the load of the postprocessing gates. The white column between the slow inverter chain and the buffers represents hardware multipliers and is not used in our design. The dark blue area in the middle of the design implements control and read out logic that was not subject to placement constraints.

B. Experimental Setup

A state machine was implemented that emulates the four cases, i.e., no attack, an emulated attack on L1, an emulated attack on L2, and an emulated attack on both lines. 10 000 000 iterations are performed for each case, and 28 counters capture the frequency (i.e., number of events) of a “1” at each latch output. The results are sent back to a PC through a universal asynchronous receiver/transmitter interface for further evaluation.

To verify the temperature stability of the design, the FPGA was placed into a climate chamber, as shown in Fig. 10. With this setup, we performed a temperature sweep from 0 °C to 70 °C.

C. Results

An initial CaLIAD implementation was tested outside the climate chamber. However, we detected irregularities that are described as follows.

An extra LUT tuning the initial delay t_0 had to be inserted at the slow inverter chain in order to see a transition between zeroes and ones in the range of the VDL chain. We expect this to be caused by a deviation of interconnection delays from the post place and route simulations.

Table VII shows the frequency of ones for 10 000 000 iterations after the aforementioned offset correction. One can see that emulating an attack on either L1 or L2 exceeds the range of the chain: all elements are either zero or one.

The other two cases, i.e., no attack or attacking both lines simultaneously, show an almost identical behavior. Latch outputs $Q_{13}, Q_{14}, Q_{15}, Q_{17}$, and Q_{19} are not constant—supposedly because of metastability and noise—while all other elements have a constant output of 0 or 1. Note that the observed numbers at the nonconstant positions may vary significantly when the experiment is repeated. However, the positions that remained constant or nonconstant did not change in all repetitions we made.

For the nonconstant positions, we expect to see a monotonic increase in counter values with increasing counter indices. However, the observed sequence of counter values is not monotonic. This implies there have to be results in the shape of $(Q_1, \ldots, Q_{28}) = 0000000000001110111111111111$, i.e., bubbles. We consider this to be the result of three different effects.

First, post place and route simulations have shown a timing irregularity between rows Y14 and Y15. This can explain the effect that $(Q_{14}, Q_{15}) = 10$.

Second, one can observe that the counter values in the middle of the chain alternate between increasing and decreasing from one element to another. Note that the CaLIAD implementation alternates between low-active NAND latches and high-active NOR latches; furthermore, post place and route simulations have shown different LUT propagation delays for rising and falling edges at the input. We assume a connection between these two effects.

Third, one can see that there exist situations where $(Q_{12}, Q_{14}) = 10$. Both latches are of the same type and the simulator does not point out any timing imbalances after routing. It is possible that metastability is the cause of this effect, but further investigations are required for clarification.

When performing the temperature sweep, we could notice that the positions with constant counter values of 0 or 10 000 000 do not change over the full temperature sweep between 0°C and 70 °C. The only temperature effect we could observe was an increase in variation of counter value with increasing temperature at certain nonconstant counter positions. Therefore, we consider the design to be temperature stable.

VII. Future Work

We used an FPGA to demonstrate the practical feasibility of the CaLIAD detection concept in digital hardware. This platform offers the advantage of low implementation time and cost. However, an FPGA is not the ideal platform to evaluate the CaLIAD: details of the interconnect architecture influence the timing behavior, but they are invisible to the designer. Timing information, which is critical for asynchronous circuits such as the CaLIAD, needs to be obtained implicitly through simulations. Toolchains are optimized for synchronous designs, though, and it is uncertain whether the timing estimations are accurate enough for asynchronous designs. Monte Carlo simulations that would allow collecting statistical information about timing are not possible. Finally, the inherent parasitics and the transistor parameters are unknown for FPGAs—this implies that no reliable quantitative statements about attack capacitances can be made even assuming that it is possible to characterize its timing behavior. A more accurate assessment would require implementing the CaLIAD in an ASIC technology.

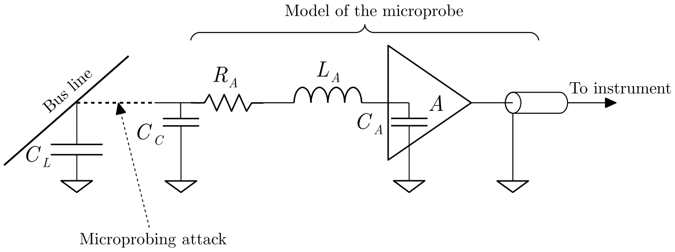


Fig. 11. Electrical model of a microprobe attack.

Another important aspect of future work is evaluating in detail how probing detectors like the LAPD and the CaLIAD can be integrated into different bus architectures.

The CaLIAD is also interesting from an optimization perspective: it combines discrete parameters such as the VDL chain length with continuous parameters such as transistor widths. Therefore, we assume that systematic optimization could lead to improvements with respect to detection performance and/or resource usage, as this was the case for the LAPD [17].

VIII. Conclusion

We have presented a microprobing detector that is able to detect probing attacks by timing measurements through a VDL. We have analyzed its performance with respect to manufacturing variations and environmental corners and were able to show that the CaLIAD is able to outperform comparable circuits such as the LAPD with respect to detection performance.

While its area usage is between the PAD and the LAPD, it only consists of digital components such as the LAPD; however, it can also be calibrated like the PAD, thus exhibiting an outstanding detection performance, such as an uncertainty region as low as $\Delta C = 5$ fF under nominal conditions and 11 fF at the worst case corner. Therefore, we recommend to use the CaLIAD for future probe detection designs as it is the best tradeoff between PAD and LAPD.

Appendix

In this paper, it has been assumed that attaching a probe to a line can be modeled as a lumped capacitor C_A connected to it. It is not evident how it is true, and in the following paragraphs, this simplified model is justified considering the order of magnitudes of different components existing in the attack.

An extended electrical model of the microprobe ac behavior connected to a bus line is presented in Fig. 11. After the attack, the parasitic components added to the bus line (C_L)

are constituted by: C_C (contact capacitance), R_A (contact resistance of the microprobe), L_A (tip inductance), and C_A (amplifier capacitance). If the microprobe is of an active type, it will have an amplifier (A) close to the tip which will decouple impedances for the rest of the instrumentation.

Contact capacitance C_C is produced during the preparation of the attack. As was exposed in Section III, L-shape platinum landing pads are deposited on the surface of the chip whose objectives are twofold: stabilize the contact of the microprobe and easy the contact to the bus line through the drilled vias. From our experience, we know that values lower than 10 fF may be expected.

The contact resistance R_A of the microprobe is highly dependent on the landing stability. The adversary aims to have a contact as stable as possible which maximizes the signal-to-noise ratio of the microprobe. Under this condition, values lower than 1 Ω are expected in tungsten tips [24]. The inductance of the tip L_A is caused by the piece of wire transmitting the signal to the amplifier. For short wires, about 2 mm in air, typical values of 100 nH can be obtained [25]. The amplifier capacitance C_A is provided in datasheets of microprobes and is highly dependent on amplifier technologies. Minimum values of 20 fF have been found for the technology [9].

After the attack, the bus line is loaded with two admittances that induce current in the bus drivers and, therefore, generate the corresponding delays in the transmission of the signals

$$\begin{aligned} Y_{\mathrm{BL}} &= Y_L + \Delta Y \\ \Delta Y &= Y_C + Y_A \end{aligned} \tag{22}$$

in which Y_{BL} is the total admittance of the bus line and ΔY is the added contribution due to the attack. A circuit analysis shows that the obtained admittances are

$$\begin{aligned} Y_C &= j\omega C_C \\ Y_A &= \frac{j\omega C_A}{\chi} \\ \Delta Y &= j\omega \left(C_C + \frac{C_A}{\chi} \right) \\ \chi &= (1 - \omega^2 L_A C_A) + j\omega C_A R_A. \end{aligned} \tag{23}$$

The term χ is a complex dimensionless number that modifies the reactive admittance of C_A as a function of L_A, R_A, and the signal frequency ω. For low frequencies, its real part tends to 1 and the imaginary part tends to 0 so that the only significant contribution of the microprobe becomes C_A, and R_A and L_A can be neglected since $\chi \rightsquigarrow 1$. In effect, the microprobe is a second-order filter with a given resonant frequency $\omega_c = \sqrt{L_A C_A}^{-1}$. If the excitation frequency is far below the resonant one, $\omega \ll \omega_c$ (as in our case), the gain tends to unity and the behavior becomes full reactive, depending only on the capacitor C_A.

For a probe bandwidth of 350 MHz [9], and considering the maximum values assumed before for the components, the estimation of χ is

$$\chi = 0.9903 + j0.00004398 \simeq 1$$

and, therefore, it can be assumed that

$$\Delta Y \simeq j\omega (C_C + C_A) > j\omega C_A$$

in which the right term of the inequality is the simplified admittance considered in this paper. Therefore, the assumption that the load contribution of the attack is due only to C_A is a conservative strategy that assures the reliable detection of the microprobing attack. However, a real attack will produce a larger load on the bus line and thus a delay larger than expected.

References

[1] S. P. Skorobogatov, "Semi-invasive attacks—A new approach to hardware security analysis," Dept. Comput. Lab., Univ. Cambridge, Cambridge, U.K., Tech. Rep. UCAM-CL-TR-630, 2005. [Online]. Available: https://www.cl.cam.ac.uk/techreports/UCAM-CL-TR-630.html

[2] J. Krämer, D. Nedospasov, A. Schlösser, and J.-P. Seifert, "Differential photonic emission analysis," in *Constructive Side-Channel Anal. Secure Design* (Lecture Notes in Computer Science), vol. 7864, E. Prouff, Ed. Berlin, Germany: Springer, 2013, pp. 1–16, doi: 10.1007/978-3-642-40026-1_1.

[3] C. Tarnovsky, *Deconstructing a 'Secure' Processor*. Washington, DC, USA: Blackhat, 2012.

[4] H. Wang, D. Forte, M. M. Tehranipoor, and Q. Shi, "Probing attacks on integrated circuits: Challenges and research opportunities," *IEEE Design Test*, vol. 34, no. 5, pp. 63–71, Oct. 2017.

[5] M. Weiner, S. Manich, R. Rodríguez-Montañés, and G. Sigl, "The low area probing detector as a countermeasure against invasive attacks," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 26, no. 2, pp. 392–403, Feb. 2017.

[6] Y. Ishai, A. Sahai, and D. Wagner, *Private Circuits: Securing Hardware against Probing Attacks*. Berlin, Germany: Springer, 2003, pp. 463–481, doi: 10.1007/978-3-540-45146-4_27.

[7] S. Manich, M. S. Wamser, and G. Sigl, "Detection of probing attempts in secure ICs," in *Proc. IEEE Int. Symp. Hardw.-Oriented Secur. Trust*, Jun. 2012, pp. 134–139.

[8] P. Maier and K. Nohl, "Low-cost chip microprobing," in *Proc. 29th Chaos Commun. Congr. (29C3)*, 2012. Accessed: Jan. 16, 2014. [Online]. Available: https://media.ccc.de/v/29c3-5124-en-low_cost_chip_microprobing_h264

[9] GGB Industries, *Picoprobe Model 18C & picoprobe Model 19C, Datasheet*. Accessed: Jun. 26, 2018. [Online]. Available: http://www.ggb.com/PdfIndex_files/mod18c.pdf

[10] M. Weiner, S. Manich, and G. Sigl, "A low area probing detector for power efficient security ICs," in *Proc. Radio Freq. Identification, Secur. Privacy Issues*, Oxford, U.K., vol. 8651, 2014, pp. 185–197.

[11] S. Henzler, *Time-to-Digital Converters* (Advanced Microelectronics), vol. 29, K. Itoh, T. Lee, S. Takayasu, W. M. Sansen, and D. Schmitt-Landsiedel, Eds. Berlin, Germany: Springer, 2010.

[12] T. E. Rahkonen and J. T. Kostamovaara, "The use of stabilized CMOS delay lines for the digitization of short time intervals," *IEEE J. Solid-State Circuits*, vol. 28, no. 8, pp. 887–894, Aug. 1993.

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

[14] K. A. Bowman, B. L. Austin, J. C. Eble, X. Tang, and J. D. Meindl, "A Physical Alpha-power Law MOSFET Model," in *Proc. Int. Symp. Low Power Electron. Design*, New York, NY, USA: ACM, 1999, pp. 218–222. [Online]. Available: http://doi.acm.org/10.1145/313817.313930

[15] T. Hashimoto, H. Yamazaki, A. Muramatsu, T. Sato, and A. Inoue, "Time-to-digital converter with vernier delay mismatch compensation for high resolution on-die clock jitter measurement," in *Proc. IEEE Symp. VLSI Circuits*, Jun. 2008, pp. 166–167.

[16] M. Weiner, S. Manich, and D. A. Delgado, "The SALVADOR simulation framework," in *Proc. TRUDEVICE*, 2016. [Online]. Available: https://upcommons.upc.edu/handle/2117/99186

[17] A. Herrmann, M. Weiner, M. Pehl, and H. Graeb, "Bringing analog design tools to security: Modeling and optimization of a low area probing detector," in *Proc. Int. Conf. Synth., Modeling, Anal. Simulation Methods Appl. Circuit Design (SMACD)*, 2018, pp. 1–4. [Online]. Available: https://ieeexplore.ieee.org/document/8434898

[18] E. B. Wilson, "Probable inference, the law of succession, and statistical inference," *J. Amer. Stat. Assoc.*, vol. 22, no. 158, pp. 209–212, 1927. [Online]. Available: http://amstat.tandfonline.com/doi/abs/10.1080/01621459.1927.10502953

[19] A. Singhee and R. A. Rutenbar, “Statistical blockade: A novel method for very fast monte carlo simulation of rare circuit events, and its application,” in *Proc. Des., Automat. Test Eur. Conf. Exhib.*, Apr. 2007, pp. 1–6.

[20] A. Singhee and R. A. Rutenbar, *Novel Algorithms for Fast Statistical Analysis of Scaled Circuits*. Amsterdam, The Netherlands: Springer, 2009. [Online]. Available: https://link.springer.com/book/10.1007/978-90-481-3100-6

[21] GGB Industries. *Picoprobe Models 28 & 29 Datasheet*. Accessed: Jun. 26, 2018. [Online]. Available: http://www.ggb.com/PdfIndex_files/mod28.pdf

[22] J.-M. Cioranesco *et al.*, “Cryptographically secure shields,” in *Proc. IEEE Int. Symp. Hardw.-Oriented Secur. Trust (HOST)*, May 2014, pp. 25–31.

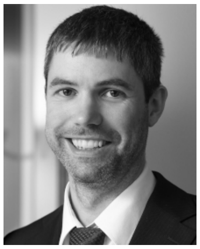
[23] Altera/Intel. (2015). *Altera Product Catalog-Devices: 60 nm Device Portfolio*. [Online]. Available: https://www.altera.com/content/dam/altera-www/global/en_US/pdfs/literature/pt/cyclone-iii-product-table.pdf

[24] Y. Zhang, Y. Zhang, and R. B. Marcus, “Thermally actuated microprobes for a new wafer probe card,” *J. Microelectromech. Syst.*, vol. 8, no. 1, pp. 43–49, Mar. 1999.

[25] A. Chandrasekhar, S. Brebels, E. Beyne, W. De Raedt, B. Nauwelaers, and T. Van Bever, “RF evaluation of low-cost leadless packages and development of distributed electrical models,” in *Proc. 53rd Electron. Compon. Technol. Conf.*, May 2003, pp. 1550–1558.

Michael Weiner received the B.Eng. degree in electrical engineering from Baden-Württemberg Cooperative State University, Stuttgart, Germany, in 2009, and the M.Sc. degree in electrical engineering from the Technical University of Munich, Munich, Germany, in 2012, where he is currently working toward the Ph.D. degree.

He is currently an Engineer at the BMW Group, Munich, where he is involved in the security architecture of electronic control units. His current research interests include embedded systems security, particularly detectors of invasive attacks and analyzing real-life products.



Wolfgang Wieser received the Dipl.-Phys. degree in physics from the Ludwig Maximilian University of Munich, Munich, Germany. During his Diploma thesis, he worked with ultracold atoms at the Chair of Laser Spectroscopy, Munich, under the supervision of Prof. Hänsch. In his Dr.rer.nat., he pioneered ultrahigh-speed optical coherence tomography, an optical 3-D imaging modality.

He owns an amateur radio license and is passionate about analog and digital electronics. He has started two companies, Wieserlabs UG, Munich, and Optores GmbH, Munich.

Emili Lupon received the M.S. and Ph.D. degrees in industrial engineering (electrical branch) from the Universitat Politècnica de Catalunya (UPC, Barcelona Tech), Barcelona, Spain, in 1975 and 1983, respectively.

Since 1986, he has been an Associate Professor at the Department of Electronic Engineering, UPC, where he teaches digital electronics at ETSEIB. He is currently a member of the research group Quality in Electronics, Munich, Germany. His current research interests include the design, test, and verification of digital and mixed-signal circuits, digital design based on field-programmable gate arrays, and error correction codes.

Georg Sigl received the Ph.D. degree in electrical engineering from the Technical University of Munich, Munich, Germany, in 1992, with a focus on layout synthesis.

He introduced new design-for-testability concepts in telecommunication application-specified integrated circuits at Siemens, Munich. In 1996, he joined the Automotive Microcontroller Department, Infineon, Munich, where he developed a universal library for peripherals to be used in 16- and 32-bit microcontrollers. Since 2000, he has been responsible for the development of new secure microcontroller platforms in the Chip Card and Security Division. Under his responsibility, two award-winning platforms have been designed. In 2010, he founded a new institute at the Technical University of Munich for Security in Electrical Engineering and Information Technology. He is currently the Director of the Fraunhofer Research Institute for Applied and Integrated Security, Munich, where he is involved in embedded security research.

Salvador Manich received the M.S. and Ph.D. degrees in industrial engineering from the Universitat Politècnica de Catalunya (UPC), Barcelona, Spain, in 1992 and 1998, respectively.

Since 2001, he has been an Associate Professor at the Department of Electronic Engineering, UPC. He develops his research activity at the Quality in Electronics Group, Munich, Germany. He has been an Invited Researcher at Instituto Superior Técnico, Lisbon, Portugal, and the Technical University of Munich, Munich. He is a member of the Center for Research in Nanoengineering, Munich. His current research interests include low-power design, test of digital systems, and security in hardware structures.