Unveiling the Vulnerability of Oxide-Breakdown-Based PUF

P. Saraza-Canflanca, F. Fodor, J. Diaz-Fortuny, B. Gierlichs, R. Degraeve, B. Kaczer, I. Verbauwhede, E. Bury

***Abstract*—This work reports a potential vulnerability of an oxide-breakdown-based Physical Unclonable Function (PUF). This generates a unique chip key based on the stochastic competition between the formation of oxide breakdown in pairs of identical transistors. Depending on which transistor breaks within each pair, a bit value of '0' or '1' is assigned to it. Combining the bits corresponding to several transistor pairs, the key is generated. This type of PUF had been considered secure until now. However, we show that, using Voltage Contrast Scanning Electron Microscopy (VC-SEM), it is possible to determine which transistor has oxide breakdown within each pair, and thus extract the PUF key with an accuracy larger than 99.9%. For this, the diffusion regions of the inspected transistors must have contacts. Furthermore, at least two contacts per cell (e.g., one for each of the two identical transistors) are needed due to the differential nature of the analysis.**

***Index Terms*—CMOS, Failure Analysis, Oxide Breakdown, Physical Unclonable Function (PUF), Security, Voltage Contrast Scanning Electron Microscopy (VC-SEM)**

I. Introduction

PHYSICAL unclonable functions (PUFs) have become a widespread security primitive in modern integrated circuits (ICs) [1]. They leverage the unavoidable and random process variations that occur during the circuit fabrication to generate secret keys that are unique to each IC. These circuits can then be used for device authentication [2] or for encryption/decryption schemes [3]. The most fundamental requirements for PUFs are randomness, uniqueness and stability. While the two first should be ideally satisfied as long as the bits of the key are truly randomly generated, most PUFs struggle to achieve perfect stability, specially when noise, temperature or voltage variations and circuit aging are considered [4].

To tackle this, most PUFs incorporate some error correction scheme involving logic and helper data stored in non-volatile memory [5]. This results in an undesirable overhead, specially considering that PUFs often target resource-constrained applications such as Internet of Things (IoT) devices [6].

To mitigate the need for such error correction, the most straightforward path is to increase the intrinsic stability of the entropy source. Two PUF concepts that achieve an extremely high stability in the presence of noise, temperature variations and even radiation are the Soft Oxide Breakdown PUF (SBD-PUF) [7], [8] and the Quantum Tunneling PUF (QT-PUF) [9], [10]. In fact, the NeoPUF, an implementation of the QT-PUF, has been presented commercially [11]. The key generated by both PUF designs is determined by which transistors within a predetermined set have oxide breakdown and which not.

It has been claimed that transistors with and without breakdown, specially when the hardness of the breakdown is limited, appear the same when inspected with physical inspection techniques such as Scanning Electron Microscopy (SEM) [7], [9], [12], and thus that PUFs based on this mechanism should be secure against that type of attack. However, we demonstrate hereby that some PUFs based on oxide breakdown can be vulnerable to such inspection techniques, and that their keys could be retrieved by a malicious attacker through them.

II. Oxide-Breakdown-Based PUFs

Both the SBD-PUF and the QT-PUF (Fig. 1) are based on the stochastic competition between the formation of oxide breakdown in pairs of identical transistors. For this, stress (i.e., overvoltage) is applied to the gate of these two identical transistors during a forming phase. The difference between the transistor oxides caused by process variability, together with the intrinsically stochastic nature of the breakdown process, causes one of the two transistors to suffer oxide breakdown before the other. A second breakdown is avoided by some compliance mechanism. In the case of the SBD-PUF, this is accomplished through the third transistor at the top in Fig. 1a. This transistor induces a reduction of the stress voltage at the gates of the stressed transistors after the first breakdown, preventing the unbroken transistor from also undergoing oxide breakdown [7] In the QT-PUF, the source/drain terminal shared by both identical devices causes a decrease in the stress

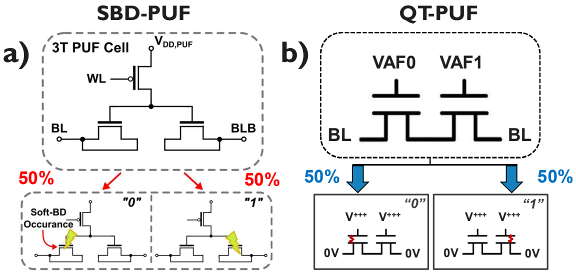

Fig. 1. a) Working principle of SBD-PUF, adapted from [7]. b) Working principle of QT-PUF, adapted from [9]. In both designs, a high voltage is applied to the gate of two identical devices until one of them forms a conductive breakdown/tunneling path. Which of the two devices forms the path determines the bit assigned to each cell: '0' or '1'.

This work was supported in part by the Cyber Security Research Flanders under Grant VR20192203 and in part by the European Commission through Horizon Europe grant agreement 101070008 ORSHIN.

P. Saraza-Canflanca (saraza68@imec.be), F. Fodor, J. Diaz-Fortuny, R. Degraeve, B. Kaczer and E. Bury are with imec, 3001, Leuven, Belgium.

B. Gierlichs and I. Verbauwhede are with the Computer Security and Industrial Cryptography (COSIC) Research Group, KU Leuven, 3001, Leuven, Belgium.

voltage after the first breakdown, thus preventing a second one [9]. The compliance mechanism has a second purpose: to limit the hardness of the breakdown, as softer breakdown spots are less visible than harder ones [7], [13]. Then, depending on which of the two transistors in the cell has oxide breakdown, a ‘0’ or a ‘1’ value is assigned to it. Both the SBD-PUF and the QT-PUF comprise an array of their basic cell so that a key can be constructed by combining the bit values of a number of cells. Notice that, due to the permanent nature of oxide breakdown, the forming phase occurs only once.

After the forming phase, the key can be read in the readout phase, where the currents flowing through both transistor gates are compared and transformed into a bit value of ‘0’ or ‘1’, depending on which of the two is larger (i.e., on which device has breakdown). Since the resistance values of the broken and unbroken gates are orders of magnitude apart, these PUFs are extremely stable [7], [9]. Notice that the readout phase can occur an unlimited number of times, as the difference between the transistor gate resistances will persist in time because of the permanent nature of oxide breakdown.

Apart from outstanding stability, oxide-breakdown-based PUFs have been reported to have close-to-ideal metrics in terms of randomness, uniformity, and uniqueness [7], [9].

III. Oxide-Breakdown PUF Vulnerability Study

One sample of the SBD-PUF fabricated in a commercial 28nm technology has been used in our study. This is an updated version of the design in [7]. The basic PUF cell (Fig. 1a) consists of one compliance PFET and two identical NFETs, one of which will break during the forming phase, determining the bit value associated to the cell.

A. Reference Electrical Measurement of the SBD-PUF

The PUF has been first formed in the laboratory. Then, the bit values of all cells have been read through the standard procedure, i.e., comparing the electrical currents flowing through the bit-lines *BL* and *BLB* (Fig. 1a) with a sense amplifier. These bit values are kept as a reference to evaluate the success of the subsequent VC-SEM inspection.

B. Passive VC-SEM inspection of an SBD-PUF

Passive VC-SEM is based on brightness differences in SEM images, and it is a widely accepted technique in the semiconductor failure analysis community [14]. It relies on scanning the sample with a primary electron beam that causes secondary electrons to leave the sample. This leads, usually, to the formation of positive charges, as for typical primary beam energies (0.7 kV in our test) the secondary electrons leaving the sample outnumber the incoming primary electrons. Depending on the local electrical configuration of the sample, the accumulation of positive charges varies, and thus the local electric potential. This impacts the secondary electrons leaving the sample, and thus the local brightness in the image (see Fig. 2). Although VC-SEM is a qualitative technique, it can be useful in failure analysis, as structures usually appearing as bright may be labelled as faulty if dark, and vice-versa [14].

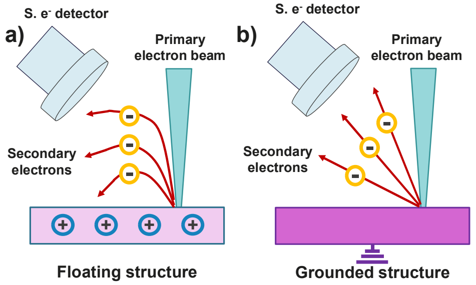

Fig. 2. Operation principle of VC-SEM and two extreme scenarios: a) In a floating structure, the initial secondary electrons cause a positive charge build-up, impeding further secondary electrons to leave, thus appearing dark in the image. b) In a grounded structure, there is no charge build-up and secondary electrons continue to leave the structure, which appears bright in the image. Thus, the local electrical configuration of the sample leads to different brightness levels.

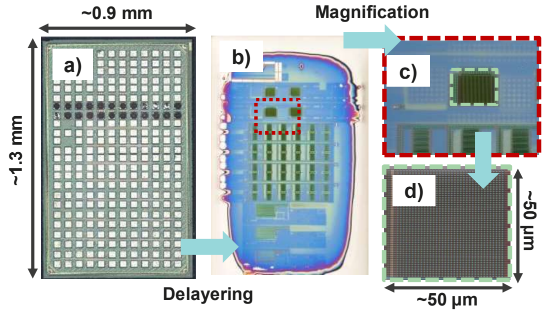

Fig. 3. Optical microscopy imaging of the sample a) before delayering, b) after delayering all metals of BEOL, c) magnified image of the PUF array and surrounding circuitry d) magnified image of the PUF array.

Before the VC-SEM inspection (performed by MA-tek Inc.), all the metals of the back end of line (BEOL) were delayered (Fig. 3). An array of 1,504 cells and an area of ~2,500 µm^2 was scanned. Fig. 4 shows the SEM image of one of the PUF cells, together with its layout to facilitate the identification of the different elements of the cell. The bright points in Fig. 4b correspond to the contacts between the different elements (transistor gate and diffusion areas, and biasing rings) and the first metal of the BEOL, which has been delayered.

It has been found that the brightness of the diffusion areas of the NFETs can be used to evaluate which of the two devices has breakdown. The diffusion contacts of NFETs without breakdown (even those of NFETs in the periphery circuitry such as sense amplifiers, not shown here) appear bright. Thus, NFETs with dark contacts can be labelled as faulty (i.e., with breakdown). Given the differential nature of the PUF cell, the NFET with darker diffusion contacts within each cell can be labelled as broken (and the other one as not-broken).

Fig. 5 shows the VC-SEM image of a portion of the array of 8×9 cells, together with a magnified view of two of those cells. For these two cells, it is straightforward to determine which of the NFET contacts are darker and, thus, which NFET has breakdown and what is the bit value associated to the cell.

To avoid any subjectivity in the analysis, we have developed an automated image analysis algorithm that determines which of the two identical transistors within each cell has

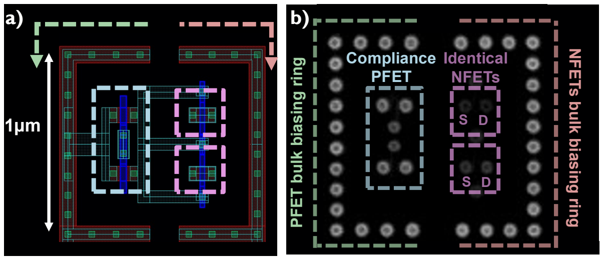

Fig. 4. a) Layout view and b) VC-SEM image of an SBD-PUF cell with its main elements highlighted. The PFET has a two-finger structure.

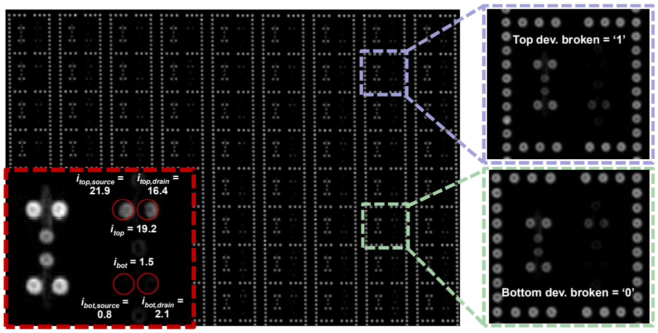

Fig. 5. VC-SEM image of a portion of the SBD-PUF array of 8x9 cells, together with a magnified view of two of those cells. The contacts of the diffusion areas of the broken NFETs appear darker. In the bottom-left inset, a magnified view of one cell with the NFET diffusion contact pixel intensities extracted by the image analysis algorithm. The red circles indicate the contact regions determined by the algorithm.

breakdown. This algorithm leverages that the PUF is composed of a grid of identical unit cells (Fig. 5), and that the distances in the SEM images can be accurately calculated by comparing these to the layout (Fig. 4). Thus, the algorithm can locate each cell in the array and, within each cell, the diffusion contacts of the NFETs. Then, the average pixel intensity of the contact regions of the two NFETs ($i_{top,drain}$, $\mathrm{i}_{\mathrm{top,source}}$, $i_{bot,drain}$ and $i_{bot,source}$) is extracted, which can take values from 0 to 255 (inset of Fig. 5). Then, the average intensity of the top (i_{top}) and the bottom (i_{bot}) transistors can be calculated as $i_{top} = (i_{top,drain} + \mathrm{i}_{\mathrm{top,source}})/2$ and $i_{bot} = (i_{bot,drain} + \mathrm{i}_{\mathrm{bot,source}})/2$.

Fig. 6a displays the extracted pixel intensity distributions of the contacts of transistors that had been labelled as broken/not-broken at the reference electrical measurement. Although broken devices display on average a lower pixel intensity than not-broken ones, there is some overlap between both distributions, which indicates that the pixel intensity value of a transistor is not enough to unequivocally determine whether it has breakdown or not.

To improve this result, a differential approach is taken, so that if $i_{top} > i_{bot}$, the bottom device is determined to have breakdown and the corresponding bit value '1' is extracted, and vice-versa (Fig. 6b). Using this method, the value of 1,503 out of 1,504 cells was correctly extracted, which leads to an accuracy larger than 99.9%. Furthermore, it is possible to remove dubious cells in which the pixel intensities of top and bottom devices are similar. This has been implemented by setting a "certainty threshold" so that i_{top} needs to be at least 20% larger or smaller than i_{bot}. Then, only 7 out of 1,504 cells are discarded as dubious (yellow region in Fig. 6b), including the one that had been evaluated erroneously, and the remaining 1,497 cells are correctly evaluated, demonstrating that the bit values of an oxide-breakdown-based PUF can be accurately extracted using VC-SEM and that these PUFs can be vulnerable to this technique. Furthermore, if an attacker attempts to extract the key, the certainty threshold allows them to distinguish between cells with a reliable and a dubious extracted value. Taking the 1,504-cell array as an example, in a worst case scenario in which the value of all dubious cells is extracted wrongly, the attacker only needs to consider $2^7 = 128$ possible keys (i.e., all possible combinations of the 7 dubious bits) instead of $2^{1,504}$ possibilities (i.e., all possible combinations of all the bits), facilitating the malicious retrieval of the key.

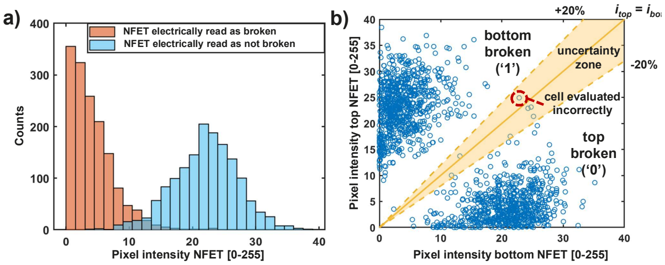

Fig. 6. a) Distributions of pixel intensity of contacts of transistors labelled as "broken" and "unbroken" through the reference electrical measurement. Broken appear in average darker than unbroken ones. However, there is some overlap between both distributions. b) Top vs bottom device pixel intensity for the 1,504 inspected cells. Comparing both values allows the extraction of the cell values. Setting a "certainty threshold" allows to discard dubious cells.

Finally, we have investigated whether the position of the breakdown along the channel of the broken NFET can also be determined through the ratio between the pixel intensities of its drain and source contacts (Fig. 4 and inset of Fig. 5). This is analogous to the method based on the ratio between drain/source currents [15]. Thus, the position is calculated as $x = i_{source}/(i_{drain} + i_{source})$, where $i_{source/drain}$ corresponds to the pixel intensity of the source/drain contact of the broken transistor. Then, $x = 0$ would correspond to breakdown exactly below the source terminal, and vice-versa, and x between 0 and 1 to an intermediate case. Fig. 7 shows the resulting distribution for x, and in the inset the distribution that was found through the current ratio method for NFETs of L_g = 40nm in [16]. The resemblance between both distributions suggests that VC-SEM inspection may allow the extraction of the breakdown position along the channel, which would render vulnerable PUFs implementations based on this [16], [17].

IV. Conclusion

The results reported in this work show that a type of oxide-breakdown-based PUFs (SBD-PUF) can be vulnerable to physical inspection attacks, although until now it had been claimed otherwise [7]. The secret key in these PUFs is determined by which transistors within a predetermined set have breakdown and which have not. We have shown that it is possible to extract this information through VC-SEM with an accuracy larger than 99.9%, which could lead to a source of vulnerability if performed by a malicious attacker. Such extraction can be achieved if the inspected transistors have contacts in their diffusion regions, and if at least two contacts are present (e.g., one for each of the two identical transistors with-

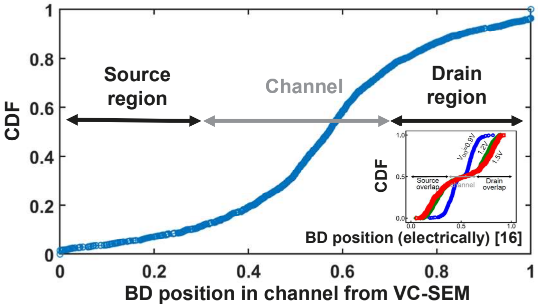

Fig. 7. Position of BD along the channel extracted a) by comparing the pixel intensity of the left (source) and right (drain) contacts of broken transistors and b) through the current ratio method (adapted from [16]).

in a cell) due to the differential nature of the analysis.

References

[1] Y. Gao, SF Al-Sarawi, D. Abbot, “Physical unclonable functions”, *Nature Electronics*, vol. 3, no. 2, pp. 81-91, 2020. DOI: 10.1038/s41928-020-0372-5.

[2] S. Sutar, A. Raha, V. Raghunathan, “Memory-based combination PUFs for device authentication in embedded systems”, *IEEE Transactions on Multi-Scale Computing Systems*, vol. 4, no. 4, pp. 793-810, 2018. DOI: 10.1109/TMSCS.2018.2885758.

[3] C. Herder, M.D. Yu, F. Koushanfar, S. Devadas, “Physical unclonable functions: A tutorial”, *Proceedings of the IEEE*, vol. 102, no. 8, pp. 1126-1141, 2014. DOI: 10.1109/JPROC.2014.2320516.

[4] M. Bhargava, C. Cakir, K. Mai, “Reliability enhancement of bi-stable PUFs in 65nm bulk CMOS”, in *IEEE International Symposium on Hardware-Oriented Security and Trust*, pp. 25-30, 2012. DOI: 10.1109/HST.2012.6224314.

[5] M. S. Mispan, S. Duan, B. Halak, M. Zwolinski, “A reliable PUF in a dual function SRAM”, *Integration*, vol. 68, pp. 12-21, 2019. DOI: 10.1016/j.vlsi.2019.06.001.

[6] A. Shamsoshoara, A. Korenda, F. Afghah, S. Zeadally, “A survey on physical unclonable function (PUF)-based security solutions for Internet of Things”, *Computer Networks*, vol. 183, 2020. DOI: 10.1016/j.comnet.2020.107593.

[7] K.H. Chuang, E. Bury, R. Degraeve, B. Kaczer, D. Linten, I. Verbauwhede, “A Physically unclonable function using soft oxide breakdown featuring 0% native BER and 51.8 fJ/bit in 40-nm CMOS”, *IEEE Journal of Solid-State Circuits*, vol. 54, no. 10, pp. 2765-2776, 2019. DOI: 10.1109/JSSC.2019.2920714.

[8] P. F. Wang, E. X. Zhang, K. H. Chuang, W. Liao, H. Gong, P. Wang, C. N. Arutt, K. Ni, M. W. McCurdy, I. Verbauwhede, E. Bury, D. Linten, D. M. Fleetwood, R. D. Schrimpf, R. A. Reed, “X-Ray and Proton Radiation Effects on 40 nm CMOS Physically Unclonable Function Devices”, *IEEE Transactions on Nuclear Sciences*, vol. 65, no. 8, pp. 1519-1524, 2018. DOI: 10.1109/TNS.2017.2789160.

[9] K. K. H. Chuang, H. M. Chen, M. Y. Wu, E. C. S. Yang, C. C. H. Hsu, “Quantum Tunneling PUF: A Chip Fingerprint for Hardware Security”, in *International Symposium on VLSI Technology, Systems and Applications (VLSI-TSA)*, pp. 1-2, 2021. DOI: 10.1109/VLSI-TSA51926.2021.9440114.

[10] M. Y. Wu, T. H. Yang, L. C. Chen, C. C. Lin, H. C. Hu, F. Y. Su, C. M. Wang, J. P. H. Huang, H. M. Chen, C. C. H. Lu, E. C. S. Yang, R. S. J. Shen, “A PUF Scheme Using Competing Oxide Rupture with Bit Error Rate Approaching Zero”, in *IEEE International Solid-State Circuits Conference (ISSCC)*, pp. 130-132, 2018. DOI: 10.1109/ISSCC.2018.8310218.

[11] https://www.pufsecurity.com/pufsecurity-leverages-neopuf-to-lead-hardware-security-technology/, consulted on the 22/11/2023.

[12] https://www.pufsecurity.com/technology/otp/, consulted on the 22/11/2023.

[13] K. L. Pey, C. H. Tung, M. K. Radhakrishnan, L. J. Tang, W. H. Lin, “Dielectric breakdown induced epitaxy in ultrathin gate oxide – a reliability concern”, in *Digest. International Electron Devices Meeting*, pp. 163-166, 2002. DOI: 10.1109/IEDM.2002.1175804.

[14] R. Rosenkranz, “Failure localization with active and passive voltage contrast in Fig and SEM”, *Journal of Materials Science: Materials in Electronics*, vol. 22, pp. 1523-1535, 2011. DOI: 10.1007/s10854-011-0459-x.

[15] M. A. Alam, D. Varghese, B. Kaczer, “Theory of breakdown position determination by voltage- and current-ratio methods”, *IEEE Transactions on Electron Devices*, vol. 55, no. 11, pp. 3150-3158, 2008. DOI: 10.1109/TED.2008.2004483.

[16] K. H. Chuang, E. Bury, R. Degraeve, B. Kaczer, T. Kallstenius, T. Groeseneken, D. Linten and I. Verbauwhede, “A Multi-bit/cell PUF using analog breakdown positions in CMOS”, in *IEEE International Reliability Physics Symposium (IRPS)*, pp. P-CR, 2018. DOI: 10.1109/IRPS.2018.8353655.

[17] H. M. Chen, M. Y. Wu, P. H. Huang, “Physically unclonable function unit with one single anti-fuse transistor”, *US Patent* 10,177,924, 2019.