Security and reliability – friend or foe?

I. Verbauwhede¹ and K.-H. Chuang^{1,2}

¹COSIC, KU Leuven, Leuven, Belgium, email: [ingrid.verbauwhede, kai.hsinchuang}@esat.kuleuven.be](mailto:{ingrid.verbauwhede, kai.hsinchuang}@esat.kuleuven.be)

²imec, Leuven, Belgium

Abstract— Security and reliability are more closely related than one might expect. The goal of this paper is to study interesting research topics at the boundaries between the two domains. Reliability concerns can benefit security to e.g. improve the performance of PUFs and TRNGs. Reliability issues can also trigger new security breaches, e.g. enabling Rowhammer or fault injection attacks. Co-optimizing security and reliability, these two communities need to work closer together.

I. INTRODUCTION

Security and reliability are two important aspects for electronic devices asking the same question: how to protect the device from failure? The security researcher focuses on the failures caused by adversary activities. The reliability researcher focuses on the failures occurring under normal operation conditions. A security failure is a breach of a system's security policy. E.g. a leak of the memory might disclose the secret key. A reliability failure is one where the normal operating conditions of a device no longer hold. Failure analysis in security is against an attacker model, i.e. an assumption we make about the capabilities of the adversary. Reliability failure analysis applies statistics and aims at modeling normal usage.

Given this similarity, it brings up interesting links between these two research fields. On the one hand, this paper discusses how reliability topics can be repurposed to benefit security in section II. On the other hand, other reliability issues may lead to severe security breaches if used outside the normal operating conditions, as illustrated in section III.

II. HOW RELIABILITY BENEFITS SECURITY

High entropy true randomness is an essential root-of-trust for almost every security application. Several effects that have a negative impact on circuit reliability, can be utilized to provide circuits that generate randomness, in particular to create physically unclonable functions (PUFs) and true random number generators (TRNGs).

A. Physically Unclonable Functions

PUFs harvest randomness from the inherent variability in ICs and produce unique fingerprints for each chip. In general, both the time-zero variability, i.e. the process variation, and the time-dependent variability, i.e. by aging, can contribute to generate unique chip fingerprints. The most well-known example of a PUF is the start-up value of the SRAM cell [1]. The raw PUF array is used as the source of entropy. But to generate a cryptographically perfect key, two important post-processing steps need to be applied, as is illustrated in Fig. 1. Note that we focus on PUFs for key generation, including strong crypto

based post-processing. Light-weight PUF constructions used for authentication are mostly broken by machine learning attacks [2]. PUF data can be affected by noise and become unstable. This instability needs to be corrected by a post-processing circuit and the helper data [1]. To cut down on the hardware overhead, a PUF with good native stability is essential. The following methods can help achieving this goal.

1) Burn-in enhancement techniques

Biased temperature instability (BTI) [3] and hot-carrier injection (HCI) [4] degradations have positive effects on the data stability of a PUF. The general idea is to make the overall mismatch in the PUF cells more significant by aging.

2) Breakdown PUF

A second technique uses the randomness from the gate oxide breakdown. In this case, the time-dependent variability completely dominates the time-zero variability. Soft oxide breakdown is used in [5] and shown in Fig. 2. It achieves zero instability under nominal operating conditions (Fig. 3). Similarly in [6] a hard breakdown is used: it achieves zero instability in a wider operating range, but it may leave more traces for visual inspection since the devices are more damaged.

3) Switching instability in memory

The memory cells of various emerging memory technologies suffer from switching instability problems, an issue for reliability. This variability again can be exploited as a PUF, e.g. RRAM PUFs [7] as shown in Fig. 4. Another benefit of this approach is the possibility of implementing a reconfigurable PUF. While it brings new opportunities, it is not ideal, since the simulations in [7] and the measurements in [8] both show non-ideal inter-configuration Hamming distance.

B. True Random Number Generators

A TRNG harvests the indeterministic fluctuations caused by noise, and transforms it into random numbers, as shown in Fig. 5. Besides an entropy source, important components for a TRNG are the post-processing to generate full-entropy bits and the on-line tests to detect on the fly weaknesses in the generated random bit stream [9, 10]. Jitter and metastability are the most used sources of true randomness. The characteristics of noise can be affected by aging. Yet the aging effects can also improve TRNG performances.

1) Metastability based TRNG

The metastability based TRNGs use circuits with metastability as the entropy source, as illustrated in Fig. 6 [11]. Using the metastability of a cross-coupled inverter pair, it has the advantage of reaching a high throughput.

2) Jitter based TRNG

Jitter is an important source of randomness, e.g. from Ring Oscillators and PLLs. In this case, it is important to only consider the Gaussian (white) noise and eliminate all global and deterministic components of the jitter [12].

3) RTN-based TRNG

Random telegraph noise (RTN) can also be exploited as an entropy source for TRNGs [13], as shown in Fig. 7. The main disadvantage is the lower throughput. Yet its main advantage is that the entropy source can be better modeled, by having good understandings in the underlying reliability physics.

4) Burn-in for TRNGs

Burn-in enhancement can be also applied to TRNGs to reduce the bias caused by offsets. By applying a stress condition, which cancels out the mismatches, the bias of a TRNG can be reduced, as the results shown in [11, 14].

III. RELIABILITY RELATED SECURITY BREACHES

Design for reliability assumes a certain usage pattern of a device. An attacker however can choose to go outside the normal operation conditions and trigger soft errors. This is illustrated with the *rowhammer* and the laser injection attacks. Both result in faulty operation, violating security boundaries.

A. Rowhammer attack

The rowhammer attack [15] targets commercial DRAMs. By repeatedly activating specific rows of a DRAM array, the charge stored in the physically adjacent cells can leak at a faster rate and will not retain enough charge before being refreshed, causing retention errors, as shown in Fig. 8. Such errors can be exploited to attack systems that use DRAM as the program memory as illustrated in Fig. 8 [16]. An attacker can inject errors into a non-accessible memory segment using rowhammers and eventually break the system. Moreover, if he has more freedom, e.g. to increase the operating temperature, the system is even more vulnerable.

It has been suggested that a similar concept can also be applied to attack NAND Flash, in which the data retention is also affected by the read disturbance issue. For the Flash memory using multi-level cells (MLCs) [17], the situation is even worse, since the error margin is reduced in this case.

RRAM, a NVM featuring also high memory density, is also susceptible to such attacks. RRAM arrays, where the cells are controlled by access transistors, provide good isolation against read disturbance. But when the density of RRAMs is increased by replacing the three-terminal access transistors with two-terminal selectors, as shown in Fig. 9, then less isolation between RRAM cells is expected and the RRAMs will be more vulnerable against rowhammer type attacks.

B. Laser fault injection attacks

In this attack, focused laser pulses are injected from the backside of a chip to the silicon substrate, as shown in Fig. 10. Depending on the injected positions, the circuit responds differently to the laser pulses, triggering typically soft-errors. E.g. a logic cell may temporarily give a wrong output value, and

a memory cell may encounter data flipping. These errors can also lead to security breaches, as shown in [18].

IV. RESEARCH CHALLENGES AND CONCLUSION

Reliability and security can be friend or foe. This was illustrated with a few examples in this paper. We see however many interesting topics for close collaboration between the two fields. Some examples are given below.

A. Silicon odometer with tamper resistance

Electronics recycling is important for the *circular economy*. Recycling cannot be done without knowing the remaining lifetime. Hence a good secure silicon odometer is important. Comprehensive know-how on reliability is therefore required. For security reason, these odometers also need to be tamper resistant, e.g., one can apply *thermal curing* effect to tamper the value of an odometer. We also need reliability know-how to design countermeasures.

B. PUFs and TRNGs resistant to physical attacks

PUF and TRNGs are basic security roots of trust and resistance to physical attacks is essential. E.g. soft and hard break down PUFs have zero instability. Yet it is unknown of the breakdowns could be visualized under a microscope.

C. Memory reliability testing including security aspects

Soft errors attacks can be avoided with improved reliability. The challenge here is to push security considerations when designing reliability tests, so that a product that passes the reliability tests can withstand the failure-based attacks.

D. Conclusion

Reliability phenomena can not only be used to improve security, but also can be abused to harm. As a result, security and reliability cannot be treated as two individual properties while designing. For the upcoming security challenges, the reliability has to be taken into account and these two properties need to be jointly optimized.

ACKNOWLEDGEMENTS

This work was supported in part by the Research Council KU Leuven: C16/15/058 and through the Horizon 2020 research program Cathedral ERC Advanced Grant 695305.

REFERENCES

- [1] R. Maes, "PUFs ...," in *Towards Hardware-Intrinsic Security*, Springer.
- [2] J. Delvaux et al., "A survey on ...," *ACM Computing Surveys* 48(2), 2015
- [3] R. Maes and V. van der Leest, "Countering the ...," *IEEE HOST*, 2014.
- [4] M. Bhargava and K. Mai, "A high reliability PUF ...," *LACR CHES*, 2013.
- [5] K.-H. Chuang et al., "A PUF using soft oxide break...," *IEEE JSSC*, 2019.
- [6] M.-Y. Wu et al., "A PUF scheme using ...," *IEEE ISSCC*, 2018.
- [7] K.-H. Chuang et al., "A cautionary note when ...," *LACR TCHES*, 2018.
- [8] Y. Pang et al., "A reconfigurable RRAM PUF ...," *IEEE ISSCC*, 2019.
- [9] M. S. Turan et al., "Recommendations...," *NIST Special Pub. 800-90B*.
- [10] W. Killmann et al., "A proposal for: functionality...," *Tech. report*, BSI,
- [11] S. K. Mathew et al., "μRNG: A 300-950 mV ...," *IEEE JSSC*, 2016.
- [12] B. Yang et al., "On-chip jitter measurement...," *IEEE AsianHOST*, 2017.
- [13] J. Brown et al., "A low-power and high-...," *IEEE Symp. VLSI-T*, 2018.
- [14] M. Bhargava et al., "Robust true random number ...," *IEEE HOST*, 2015
- [15] Y. Kim et al., "Flipping bits in memory ...," *ACM/IEEE ISCA*, 2014.
- [16] S. Govindavajhala and A. Appel, "Using memory ...," *IEEE S&P*, 2003.
- [17] Y. Cai et al., "Read disturb errors in MLC...," *IEEE/IFIP DSN*, 2015.
- [18] D. S. Kumar et al., "An in-depth and black-box...," *CARDIS*, 2018

Fig. 1. An example of a PUF-based key generation scheme. The raw entropy from process variation is readout as digital bits. The required 128-bit key will be extracted from these raw bits by the post-processing steps including error-correction and entropy extraction [1].

Fig. 3. The measured BER and unstable bit percentage of soft-BD PUF, under different operating voltages and temperatures [5]. In total 1024 PUF cells are measured for 20K readout cycles per cell.

Fig. 4. The concept of using resistance differences between programming cycles as the source of reconfigurable RRAM PUFs. The right hand figure shows that each RRAM cell is unique and will result in non-identical resistance distributions, which causes device dependent biases and limits the reconfigurability [7].

	Entropy source	Error correction	Reconfigurability	Additional process	Visibility
SRAM	Process variations	Yes	No	<u>No</u>	<u>Less</u>
Oxide BD	Oxide defects	<u>No</u>	No	<u>No</u>	More
RRAM	Switching instability	<u>No</u>	<u>Possible</u>	Yes	More

Table 1. Comparison between different PUF implementations. Both the PUFs based on oxide BD and RRAM can achieve ideal data stability, but both has advantages and disadvantages. The underlined texts show the properties which are considered as advantages.

Fig. 2. The schematic of the soft-BD PUF array including the periphery circuits and the 3T PUF cell [5].

Fig. 5. Generic block diagram of a true random number generator. The raw entropy from the stochastic processes is digitized and processed to generate qualified random numbers to be further used in various applications.

Fig. 6. A TRNG based on metastability, as proposed in [11]. The cross-coupled inverters are gated by an NMOS footer, which is switched on when clock signal is high. When clock is low, the inverters are not working and the output will be set to the initial state, i.e., $V_L = V_R = V_{DD}$. Once the clock rises, the two internal nodes start discharging, and depending on the transient fluctuations, the circuit will be moving to one of the two stable states. Since this process is indeterministic, the circuit can continuously generate random bits.

Fig. 7. Besides being a well-studied reliability concern, RTN can also be exploited to generate random bits. The advantage of using RTN as entropy source is that the underlying physics can be better modeled comparing to the thermal noise or clock jitter. This physical entropy model helps to prove randomness for this type of TRNGs.

	Process variation	Regular aging	Manual aging	Noise
PUFs	Good	Bad	Good	Bad
TRNGs	Bad	Bad	Good	Good
Crypto engines	Bad	Bad	-	Bad

Table 2. Summary of how the reliability concerns can affect different hardware security primitives.

Fig. 8. A Simplified illustration of the origin of the rowhammer attack in DRAMs. As the word-line spacing is scaled with technology, the disturbance from other rows become more severe in newer technology nodes. Once there is enough leakage accumulated through repetitive DRAM accessing, the targeted DRAM cell will lose the stored data.

Fig. 10. The read disturbance might also introduce security threats to RRAMs, especially if the RRAM structure is moving from the 1T1R configuration to the 1S1R configuration for higher density. Using selectors can indeed help to realize high density cross-bar RRAM arrays, but it has less isolation comparing to transistors. In this case, the read/write disturbance can have higher possibility to induce *retention failures*, which makes the subsequent attacks more feasible.

Fig. 9. Possible security threats lead by memory errors. Typically, the user software cannot access the memory allocated for the operating system (OS), but there are chances that the memory allocated for this application is adjacent to the memory for OS. Under this circumstance, the rowhammer attack can be utilized to induce errors for the OS and an adversary may find a way to hijack the OS.

Fig. 11. Illustration of how laser beams can be used to induce faults in an IC. By shooting lasers into the silicon substrate, different type of soft-errors can be induced in different circuit blocks. In general, when a reverse-biased depletion region is shot by laser, the photocurrent will be induced and might change the logic values. Laser attack is very powerful since it can affect most of the circuits, but it is also more expensive and more difficult to setup.

	Source	Target	Physical access	Example countermeasure
Rowhammer	Internal	Memory blocks	No	Increase refreshing rate
Fault injection	External	All circuits	Yes	Error correction

Table 3. Summary of the attacks enabled by the soft errors. The rowhammer attack can be performed by abusing the normal circuit operations, which works mainly on the DRAM and possibly other memory blocks. One of the proposed countermeasure for rowhammer attack is to increase the refreshing rate of DRAM cells to prevent too much charge leaking. The fault injection attack relies on the external excitations, e.g., laser pulses to induce soft errors. This type of attack is more powerful, which can works on most of the circuits, but it requires more delicate setup and usually needs physical interaction with the chips. One of the possible countermeasure is to add error correction functions, which typically requires redundancy.