Reliability Analysis of FinFET-Based SRAM PUFs for 16nm, 14nm, and 7nm Technology Nodes

Masoumian, Shayesteh; Selimis, Georgios; Wang, Rui; Schrijen, Geert-Jan; Hamdioui, Said; Taouil, Mottaqiallah

DOI
10.23919/DATE54114.2022.9774735

Publication date
2022

Document Version
Final published version

Published in
Proceedings of the 2022 Design, Automation & Test in Europe Conference & Exhibition (DATE)

Reliability Analysis of FinFET-Based SRAM PUFs for 16nm, 14nm, and 7nm Technology Nodes

Shayesteh Masoumian[1,2] Georgios Selimis[1] Rui Wang[1]
Geert-Jan Schrijen[1]
[1]Intrinsic ID B.V.
High Tech Campus 83, Eindhoven, The Netherlands
shayesteh.masoumian@intrinsic-id.com

Said Hamdioui[2] Mottaqiallah Taouil[2]
[2]Delft University of Technology
Faculty of EE, Mathematics and CS
Mekelweg 4, 2628 CD Delft,
The Netherlands

***Abstract*—SRAM Physical Unclonable Functions (PUFs) are among other things today commercially used for secure primitives such as key generation and authentication. The quality of the PUFs and hence the security primitives, depends on intrinsic variations which are technology dependent. Therefore, to sustain the commercial usage of PUFs for cutting-edge technologies, it is important to properly model and evaluate their reliability. In this work, we evaluate the SRAM PUF reliability using within class Hamming distance (WCHD) for 16nm, 14nm, and 7nm using simulations and silicon validation for both low-power and high-performance designs. The results show that our simulation models and expectations match with the silicon measurements. From the experiments, we conclude the following: (1) SRAM PUF is reliable in advanced FinFET technology nodes, i.e., the noise is low in 16nm, 14nm, and 7nm, (2) temperature variations have a marginal impact on the reliability, and (3) both low-power and high-performance SRAMs can be used as a PUF without excessive need of error correcting codes (ECCs).**

***Index Terms*—FinFET, measurements, reliability, simulation model, SRAM PUF**

I. Introduction

SRAM PUFs derive their entropy from random physical variations in transistors and wires as a side effect of the manufacturing process [1]. Therefore, they can be used to generate unique unforgeable root keys and device identities. SRAM PUFs are very popular and are deployed in many commercial products such as Microsemi [2] and NXP [3]. To ensure the reliability and reproducibility of the device keys and identities during the lifetime, typically error correcting code (ECC) is used to create a safe operational margin. The reliability of an SRAM PUF determines how much error correction is required and hence, how large the hardware overhead will be. SRAM PUF reliability depends on the technology, operational conditions, and environmental conditions. As the structure of FinFETs differ from planar transistors, the effect of process variation may differ at the electrical level. Hence, it is important to model and validate the reliability for FinFET based SRAM PUFs.

Research on SRAM PUF reliability has mainly focused on planar CMOS [4–6]. The results of these studies show that SRAM PUFs are reliable and that the noise can be overcome with ECCs. For example, in [4] the authors simulated and measured the maximum noise for SRAM PUFs for 40nm, 65nm, and 130nm planar CMOS technology nodes. In [6], the authors evaluated the SRAM PUF reliability and stability for 28nm planar and 16nm FinFET SRAM PUFs using predominately silicon measurements. The authors showed that the PUFs in 16nm FinFET and 28nm planar had a marginal difference in terms of reliability. 16nm FinFET based SRAM PUF reliability for different ramp-up times using only simulation models are evaluated in [5]. They showed that the within class Hamming distance (WCHD) is less than 10% in 16nm FinFET based SRAM PUFs. From the above it becomes evident that a limited number of research avenues have been explored.

In this work, we analyze the FinFET based SRAM PUF reliability using both simulations and silicon measures from 16nm to 7nm devices. We analyze the impact of temperature and study both low-power and high-performance devices. Our main contributions are: (1) a simulation model to assess the reliability of SRAM PUF for FinFET technology 16nm, 14nm, 7nm and their associated silicon measurement validation, (2) reliability analysis for high-performance and low-power SRAM PUF designs, and (3) reliability analysis for different temperatures, i.e., from 0°C to 85°C.

The remainder of paper is organized as follows. Section II provides a background on SRAM PUF reliability. Section III presents our simulation framework and simulation results. Section IV validates the simulation model using silicon results. Finally, Section V and Section VI discuss the results and conclude this paper, respectively.

II. Background

In this section we explain how SRAM PUFs work and how they are affected by process and environmental variation.

SRAM PUF Cell and Reliability: An SRAM PUF is based on the well-known 6 transistor (6T) cell [7]. The cell consists of two cross-coupled inverters and two access transistors. The inverters are asymmetrical as process variation causes non-uniform transistor threshold voltages. The asymmetry is exploited by PUFs; a cell has a preferred initial state (i.e., '0' or '1') during voltage ramp-up. Repeated responses of the same SRAM PUF are slightly different due to circuit noise and environmental changes (e.g., temperature), which have an impact on the reliability. The reliability can be evaluated using WCHD metric. WCHD represents the hamming distance between the enrollment (i.e., a reference measurement at 25°C in a trusted environment) and a reconstruction (i.e., a later measurement possibly with different environmental

978-3-9819263-6-1/DATE22/©2022 EDAA

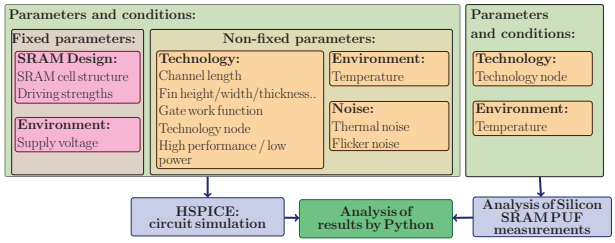

Figure 1: Simulation and Validation Process

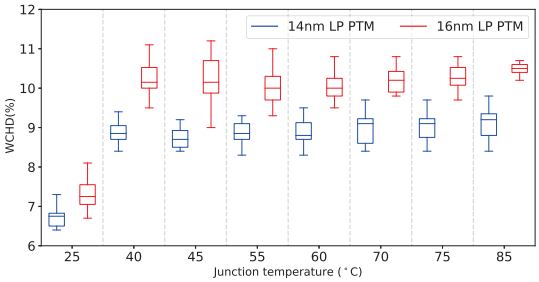

Figure 2: WCHD Simulation results

conditions). Note that the lower the WCDH value, the less amount of error correction is needed. Next, we discuss the important parameters that affect the reliability of SRAM PUF, namely process variation, environment variation (e.g., temperature) and circuit noise.

FinFET Process Variation: Variations in physical dimensions, dopant concentration, and gate work function are main sources of process variation in FinFET transistors [8]. Process variation affects the transistor's electrical parameters (e.g., threshold voltage of transistors (V_{th})) which typically has a higher impact in smaller technology nodes. In this work, we consider variations in channel length, channel width, thickness of oxide, and work function as the main sources of process variation in our simulation model.

Environmental Variations: Temperature and supply voltage are environmental sources of variation. In this paper, we limit ourselves to temperature variations as the impact of supply voltage on SRAM PUF reliability is marginal [9]. Temperature variation has an impact on the electrical parameters of transistors and hence, the SRAM PUF response is affected by it. The start-up value of SRAM PUF depends on the transistor currents through both inverters during voltage ramp-up, which further depends on V_{th} and mobility [4]. Note that V_{th} decreases and the mobility decreases with increasing temperatures [10, 11].

Circuit Noise: The SRAM start-up value may be affected by any noise source. The most relevant noise types in FinFETs are thermal noise and flicker noise. Thermal noise originates from the resistance that a current flow faces inside the channel and is mainly affected by the temperature. We ignore the power supply noise as its impact is negligible [9].

III. Simulations

In this section we describe the simulation model, simulation setup and the simulation results. The complete overview of the simulation and validation process is depicted in Fig. 1.

A. Simulation Model and Setup

Our simulation model consists of 6T SRAM cells with minimum sizes. In order to have reliable read and write operations, the ratio between the drive strengths of the pull up (PU), pull down (PD), and access transistors (AX) must satisfy certain conditions [12]. The ratio's that satisfy these conditions with minimum sizes are PU:AX:PD=1:1:2. The cell is simulated using the following FinFET libraries: 16nm, 14nm PTM for low-power [13], and 7nm ASAP7 [14] for SRAM. In addition, we consider 3 corners to simulate global variation for the ASAP7 library; they are: TT (typical nFET, typical pFET), FF (fast nFET, fast pFET), SS (slow nFET, slow pFET).

To model local process variation, a Gaussian distribution is used for the four main sources of variation (i.e., channel length, fin width, thickness of oxide, and work function).

The nominal library values have been used as mean value for them and for the standard deviations a sigma of 4%, 4%, 1.33%, and 1% respectively have been used [15]. To properly analyze the impact of these sources of variation, 1000 Monte Carlo samples have been used for each experiment and hence 1000 cells have been effectively simulated. The experiments have been performed using the HSPICE simulator [16].

The experiments are performed for a temperature range from -40°C to 85°C for the ASAP7 and from 25°C to 85°C for the PTM library. In all cases, the voltage ramp-up time has been fixed to 10μs [5].

The noise is created from the internal HSPICE flicker and thermal noise models [17]. As the start-up values of SRAM PUFs can only be retrieved in the time domain, transient simulations have been performed. For each cell, twenty different random noise seeds have been used.

B. Simulation Results

In this subsection, we provide the simulation results. More specifically, we explain the impact of temperature, and high-performance and low-power models on the WCHD metric.

Impact of temperature on WCHD: Figs. 2 and 3.a show the WCHD results for 16nm, 14nm, and 7nm technology nodes for the given temperature range. From these figures, we observe the following:

1) The impact of temperature on WCHD in 16nm, 14nm, 7nm is marginal and WCHD is below 14% for all temperatures. This can be explained as follows. On the one hand, at high temperatures the affect of process variation is typically larger (e.g., larger V_{th} mismatch) which reduces the WCHD value as cells become more stable. However, on the other hand, a higher temperature increases the circuit noise which reduces the cell stability and hence increases the WCHD. Apparently, both affects cancel each other more or less out resulting in a marginal impact of temperature on WCHD.
2) The impact of technology scaling on WCHD is also marginal. The WCHD slightly decreases when the technology is scaled. Note that for smaller nodes the impact

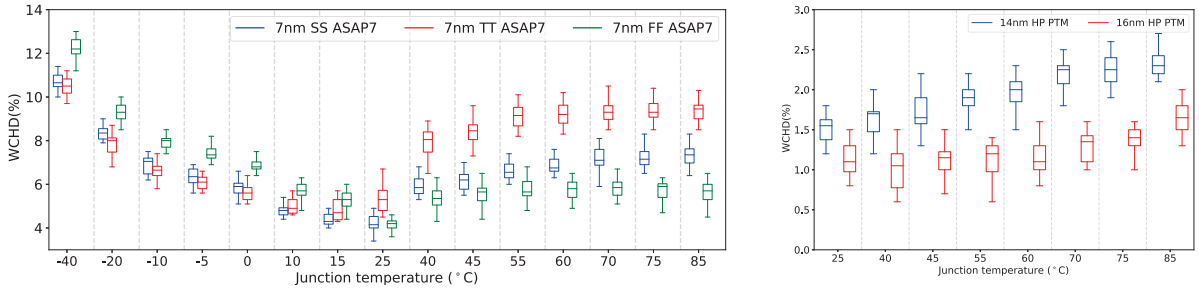

Figure 3: Simulation results - WCHD at different temperatures (a) 7nm ASAP7 (b)16nm and 14nm HP PTM technology node

of process variation is larger and hence, more cells get a larger asymmetry; this results in more predictable start-up values [18]. The slight variations in WCHD can be attributed to process variation and different noise levels in the different technologies.

Impact of high-performance and low-power PUF designs on WCHD: Fig. 3.b shows the simulation results for 14nm and 16nm HP SRAM PUFs. The noise model parameters for the LP designs have been used as well for the HP designs. With respect to Figure 2, we conclude the following:

1) WCHD is much lower in high-performance SRAM PUF designs. This is around 1 to 3% for HP as compared to 6 to 12% for LP. This can be explained by looking at the differences in electrical parameters such as V_{th}. As the V_{th} is lower in HP transistors, they are more impacted by process variation. This leads to more skewed cells and hence a better reliability.
2) The impact of temperature on WCHD in both HP and LP SRAM PUFs is similar and marginal.

IV. Silicon Validation

To validate the simulation results, the SRAM start-up values of three advanced technology nodes have been measured. They have been measured from a 16nm LP NVIDIA chip [19], 14nm LPC NXP chip [20] and 7nm Xilinx chip [21]. The NVIDIA and Xilinx chips have been manufactured by TSMC and the NXP chips by Samsung. Fig. 4 shows the WCHD for the three chips for different temperatures. The figure contains two x-axis; the x-axis on the bottom shows the ambient temperature (e.g., temperature of the oven) and the top x-axis the junction temperature (i.e., the temperature of the transistors). Unfortunately, the junction temperature was not available for the Xilinx chip.

Due to temperature safety requirements, negative temperatures for the Xilinx 7nm chip have not been measured. From Fig. 4 we conclude the following:

1) SRAM PUFs are reliable in 16nm, 14nm, and 7nm technology. The WCHD is less than 13% in all the silicon based measurements.
2) The temperature impact on WCHD is marginal in 16nm, 14nm, and 7nm and errors can be corrected using ECCs.
3) The impact of temperature on WCHD shows a slightly different trend for the 16nm NVIDIA chip as compared to the 14nm NXP chip. This can be most likely attributed to different manufactures.

Simulation vs Measurement: Both the simulation results and measurements show that the impact of temperature on the SRAM PUF reliability for advanced FinFET technologies is marginal and that SRAM PUFs are reliable for these technology nodes. Similar WCHD values have been reported for the simulations and measurements.

V. Discussion

In this work we analyzed the reliability of FinFET SRAM PUFs in terms of WCHD. From the results, we conclude the following:

Impact of temperature on reliability of SRAM PUF in FinFET technology: The impact of temperature on WCHD in 16nm, 14nm, and 7nm is less than 14% in the measurements. This can be easily overcome by ECCs [22]. The temperature influences the behavior of an SRAM PUF, i.e., a higher temperature leads to a larger impact of process variation which in turn leads to more skewed cells. On the other hand, increasing temperature results in more thermal noise and hence a higher WCHD and thus a lower reliability. Depending on the technology, one affect might be more dominant than the other. However, from our experiments we conclude that these affects have the tendency to cancel each other out.

High-performance vs low-power SRAM PUFs: Both high-performance and low-power SRAM are reliable to be used as PUF. The WCHD analysis for both of them shows that the error in different conditions is less than the acceptable ECC capacity. In [18], the authors showed that the impact of temperature on static behavior of HP SRAM PUFs is higher than LP SRAM PUFs. In this work we have shown this subsequently leads to less noise in high-performance SRAM PUFs as compared to LP SRAM PUFs.

FinFET vs planar technology: Cortez et al. in [4] showed that the maximum WCHD for 40nm planar technology is around 28%. In this work we have shown from SRAM PUF measurements that the maximum WCHD for 16nm, 14nm, and 7nm SRAM PUFs are 14%, 10%, and 11%, respectively. This shows that advanced FinFETs based SRAM PUF cells have a higher reliability than the older planar based ones.

Failure rate estimation: In practical PUF applications, a 128-bit secret key must be derived. We assume that we have 1000

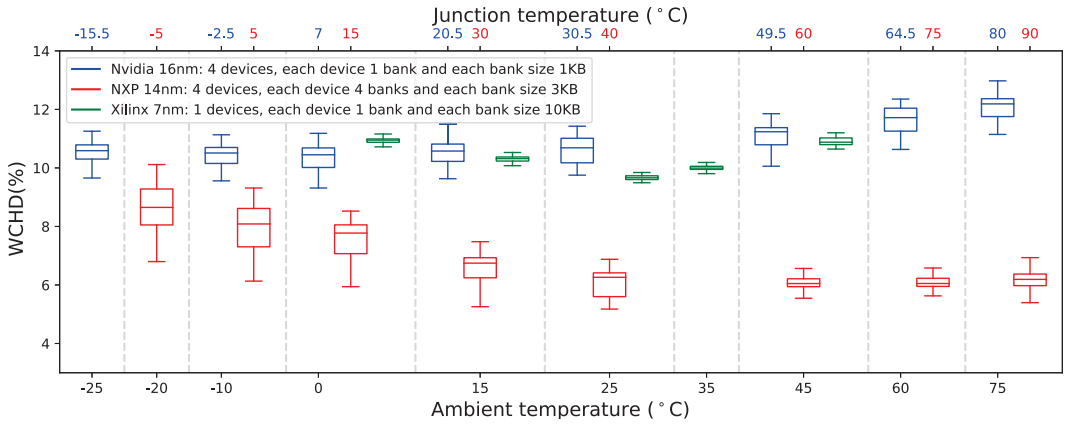

Figure 4: WCHD Measurement results (10, 10, and 100 experiments per temperature for Nvidia, NXP, and Xilinx respectively)

noisy SRAM PUF cells available to realize this. According to the simulated WCHD results presented in Fig. 3.a, the worst-case bit error rate (BER) of 7nm SRAM PUF data in the TT case is assumed to be 12%. To overcome this BER a helper data algorithm (HDA) based on ECCs and code-offset techniques is required to fully recover from the errors. To achieve this, we propose a concatenated code with soft information decoding, i.e., using Reed-Muller (RM) [32, 6] code and Reed-Solomon (RS) [31, 22] code [23] with 6-bit symbols ($k_{RS} = 128/6 = 22$). Therefore, the proposed HDAs require $31 \times 32 = 992$ SRAM cells to be able to generate a reliable 128-bit secret key. The concatenated code works as follows. In the first stage, the RM decoder uses a brute-force approach to find the most likely 32-bit RM code word candidate and decodes this into a 6-bit RS symbol together with the corresponding soft information output. During the second stage, the RS decoder erases 5 positions where the soft information of these positions shows the RS symbol is more likely to be erred compared to other symbols. Besides the erasures, up to 2 errors can be corrected with the RS decoder. The proposed decoders are simulated over different BERs to evaluate the reliability. When the BER is 12%, the key failure rate is approximately 10^{-9}, showing a very low failure rate.

VI. Conclusion

In this paper, we simulated and measured the SRAM PUF reliability in terms of WCHD for advanced FinFET technologies. The impact of process variation, circuit noise, and temperature on reliability of SRAM PUF has been investigated. To the best of our knowledge, this paper presented for the first time in the literature silicon results for 14nm and 7nm SRAM PUF designs. The results from both simulation and measurements show that the SRAM PUF is reliable in advanced FinFET nodes. Moreover, the temperature barely impacts the reliability both in simulations and measurements. Although there is a difference between low-power and high-performance SRAM PUF cells, the reliability in both cases can be easily guaranteed when ECC is used. For high-performance SRAM PUFs less ECC is needed as the reliability is higher. Overall, this research shows that FinFET based SRAMs PUFs can be reliably deployed.

Acknowledgment

The authors would like to thank Karthik Keni Yerriswamy from Intrinsic ID for performing the experiments.

References

[1] R. Maes, *Physically Unclonable Functions: Construction, Properties and Applications*, 2013, no. August.
[2] P. Overview, "Microsemi - Product Overview PolarFire FPGA," 2019.
[3] "NXP - LPC55S6x 32-bit ARM Cortex-M33+ MCU Data Sheet," 2019.
[4] M. Cortez *et al.*, "Intelligent Voltage Ramp-Up Time Adaptation for Temperature Noise Reduction on Memory-Based PUF Systems," *TCAD*, 2015.
[5] H. Ramanna, "A Study on Controlling Power Supply Ramp-Up Time in SRAM PUFs," 2019.
[6] B. Narasimham *et al.*, "SRAM PUF quality and reliability comparison for 28 nm planar vs. 16 nm FinFET CMOS processes," *IRPS*, 2017.
[7] D. Holcomb *et al.*, "Power-Up SRAM State as an Identifying Fingerprint and Source of True Random Numbers," *IEEE Trans. Comput.*, 2009.
[8] C. Shin, *Variation-Aware Advanced CMOS Devices and SRAM*. Springer Netherlands, 2016, vol. 56.
[9] M. S. Golanbari *et al.*, "Reliable memory PUF design for low-power applications," in *ISQED*, mar 2018.
[10] STMicroelectronics, "How to achieve the threshold voltage thermal coefficient of the MOSFET acting on design parameters Intr."
[11] I. M. Filanovsky *et al.*, "Mutual compensation of mobility and threshold voltage temperature effects with applications in CMOS circuits," *(IEEE Trans Circ Syst Fund Theor Appl*, vol. 48, pp. 876–884, 2001.
[12] L. Chang *et al.*, "Stable SRAM cell design for the 32 nm node and beyond," *Digest of Technical Papers - Symp. on VLSI Tech.*, 2005.
[13] "Predictive technology model." [Online]. Available: http://ptm.asu.edu/
[14] "ASAP: Arizona State Predictive PDK." [Online]. Available: http://asap.asu.edu/asap/
[15] H. W. Cheng *et al.*, "Random work function variation induced threshold voltage fluctuation in 16-nm bulk FinFET devices with high-κ-metal-gate material," *IWCE*, 2010.
[16] "HSPICE: The Gold Standard for Accurate Circuit Simulation." Synopsys. [Online]. Available: https://www.synopsys.com/content/dam/synopsys/verification/datasheets/hspice-ds.pdf
[17] *HSPICE® User Guide: RF Analysis*. Synopsys, 2010, no. December.
[18] S. Masoumian *et al.*, "Modeling static noise margin for finfet based sram pufs," in *ETS*, 2020.
[19] NVIDIA. "harness ai at the edge with the jetson tx2 developer kit.
[20] NXP. "evaluation kit for the i.mx 8m mini applications processor.
[21] Xilinx. "versal ai core series vck190 evaluation kit.
[22] M. D. Yu *et al.*, "Secure and robust error correction for physical unclonable functions," *IEEE Design and Test of Computers*, 2010.
[23] *BCH and Reed-Solomon Codes: Designer Cyclic Codes*. John Wiley & Sons, Ltd, 2005.