An Open-Source Platform for Evaluation of Hardware Implementations of Lightweight Authenticated Ciphers

Abubakr Abdulgadir*, William Diehl†, Jens-Peter Kaps*
*Cryptographic Engineering Research Group
George Mason University
Fairfax, Virginia 22030, USA
{aabdulga, jkaps}@gmu.edu
†Signatures Analysis Lab
Virginia Tech
Blacksburg, Virginia 24061, USA
wdiehl@vt.edu

***Abstract*—Lightweight implementations of cryptographic algorithms must be evaluated in terms of security, cost, and performance before their deployment in practical applications. The availability of open-source platforms for such evaluation saves researchers' time and increases reproducibility of results. In this work, we improve upon the previous version of the Flexible Opensource workBench fOr Side-channel analysis (FOBOS) to introduce "FOBOS 2," and utilize it to perform such evaluation tasks for hardware implementations of authenticated ciphers, with special focus on candidates submitted to the NIST Lightweight Cryptography standardization process. We perform power measurements on Artix7 FPGA, and countermeasure evaluation of lightweight hardware implementations of selected NIST Lightweight Cryptography Round-2 candidates and the current NIST standard AES-GCM on the Spartan6 and Artix7 FPGAs. Our results show that Ascon consumes the least power at 50 MHz, and has the lowest change in dynamic power per increase in frequency, while GIFT-COFB consumes the least energy-per-bit. We also show that side-channel countermeasures applied to implementations of Ascon and AES-GCM are effective using leakage detection tests.**

I. Introduction

Lightweight cryptography (LWC) is deployed in resource constrained devices like smart-cards and RFID tags. Power consumption and energy per bit (E/bit) determine power supply specification and battery life and hence, are crucial for such applications.

Also, adversaries can easily gain physical access to such systems and measure side-channels such as power consumption and Electro-Magnetic emanations (EM). This makes side-channel analysis [1] (SCA) especially concerning for lightweight applications. [1]

 Abubakr Abdulgadir, William Diehl, and and Jens-Peter Kaps. An Open-Source Platform for Evaluation of Hardware Implementations of Lightweight Authenticated Ciphers *2019 International Conference on ReConFigurable Computing and FPGAs (ReConFig), 2019*, pp. 1-5. https://doi.org/10.1109/ReConFig48160.2019.8994788

978-1-7281-1957-1/19/$31.00 ©2019 IEEE

Systematically evaluating SCA resistance is necessary for countermeasure designers. One of the most widely used methodologies is Test Vector Leakage Assessment (TVLA) [2] which applies statistical tests to measure the significance of leakage. Such methodologies and tools will be valuable for efforts like the NIST LWC project that aims to standardize algorithms for resource-constrained devices.

The availability of open-source hardware and software to perform security evaluation saves researchers' time and enables result reproducibility. Several solutions to perform SCA are already available for academia and industry. The DPA Workstation from Rambus [3] and Inspector from Riscure [4] are examples of commercial systems, however, they are too costly for many academic and low-end users. SAKURA boards [5] are also widely used in academia and support FPGAs and smart cards, however, they do not include integrated acquisition and analysis tools. NewAE Chipwhisperer is a platform that has many Design Under Test (DUT) options and allows DUT and sampling clocks to be synchronized for precise measurements [6].

The Flexible Opensource workBench fOr Side-channel analysis (FOBOS) is a comprehensive SCA platform that uses commercially available low-cost FPGA boards (e.g. Digilent Nexys-A7) whenever possible. Since FOBOS is directly compatible with CAESAR (Competition for Authenticated Encryption, Security, Applicability and Robustness) Hardware API [7] and expected to be directly compatible with the upcoming Lightweight Cryptography API, no time is needed to adapt cipher implementations to a new interface. Given the number of candidates in the NIST LWC project, time savings will be a significant factor in evaluating these ciphers. While Chipwhisperer is compatible with state-of-the-art target boards, work is needed to adapt NIST LWC ciphers interface to its interface. Therefore, using FOBOS will save time in the evaluation of NIST LWC candidates.

In this work, we improve the architecture of the FOBOS framework and upgrade FOBOS for compatibility with state-of-the-art Xilinx 7-series FPGAs resulting in the new FOBOS 2. We use FOBOS 2 to measure power and compute E/bit for Round-2 NIST LWC candidates Spoc, Spook, GIFT-COFB and Ascon, and compare them to the current standard, AES-GCM, as a benchmark. We also evaluate SCA countermeasures on protected implementations of Ascon and AES-GCM in the Spartan6 and Artix7. As a result, we claim the following contributions:

- An upgraded test platform capable of power measurement and SCA resistance evaluation that supports state-of-the-art, low-cost, commercially available FPGA boards.
- The first power measurements and energy computations of NIST LWC hardware implementations by 3rd party testers on actual advanced hardware.
- The first verification of SCA countermeasures of NIST LWC candidates in the Artix7 FPGA.

II. Background

A. Test Vector Leakage Assessment

The Test Vector Leakage Assessment (TVLA) methodology [2] has been used in many publications to test if there is significant information leakage from an implementation. This test is used in lieu of time-consuming Differential Power Analysis (DPA) attacks to evaluate leakage. Security against SCA implies that power traces collected when processing fixed data and traces collected when processing random data should be statistically indistinguishable. We call the two trace sets Q_f and Q_r respectively. A t value is calculated as follows:

$$t = (\mu_f - \mu_r)/\sqrt{s_f^2/n_f + s_r^2/n_r}$$

Where μ_f and μ_r are the means, s_f and s_r are the standard deviations and n_f and n_r are the number of samples in the sets. The null hypothesis is, that the means of the two trace sets Q_f and Q_r are equal. At values of $\mid t \mid > 4.5$ we can reject the null hypothesis at a confidence level of 99.999% and reason that the implementation is likely leaking information. However, this doesn't prove that the leakage is exploitable, and doesn't recover any secret information [2].

B. Frequency-based Leakage Detection

While moments-based leakage detection, e.g., computations on means and variances, can be used, frequency-based leakage detection can also be employed. An example of frequency-based leakage detection is the χ^2-test [8], which is based on frequency of occurrence. Frequencies of occurrences between "classes" are evaluated to χ values, and summed to get χ (normalized expected frequency of occurrence) and ν (degrees of freedom). Classes could include test vectors with fixed data or random data, for example. A probability p is computed to determine whether "classes are distinguishable." The χ^2-test is interpreted as "passing" for every instance in time where $p > 10^{-5}$, and "failing" when $p < 10^{-5}$.

III. Methodology

FOBOS is a free and open-source tool which provides a single "acquisition to analysis" platform to measure resistance to power analysis side-channel attacks. The system was described in [9] and demonstrated at [10]. FOBOS consists of a data acquisition module used to acquire power traces from the Device Under Test (DUT) and an analysis module used to process the traces, run attacks and assess SCA leakage.

The ongoing NIST LWC standardization process has 32 Round-2 candidates, and a presumably large number of future later-round candidates. NIST LWC candidates are evaluated partially based on performance (including power) and cost (including energy) [11]. To compare this large number of algorithms, in terms of power, E/bit and SCA resistance, one needs an efficient platform with flexible interfaces that is compatible with the hardware API in use. Academic efforts benefit from low-cost systems that can be assembled using commercially available components, which at the same time promotes result reproducibility. The previous version of FOBOS was limited in speed because of its PC $\leftrightarrow$ control board communication protocol, lack of support for fast oscilloscopes and use of now discontinued FPGA boards.

A. FOBOS 2

To address these issues, we have developed an upgraded system with similar architecture, but which runs much more efficiently and uses modern hardware. Our upgraded system is available for download at [12]. Specifically we have performed the following upgrades:

- The trace collection speed is improved. When collecting AES traces, we achieved 5x speedup when using the same oscilloscope as the previous version of FOBOS and 25x when using Picoscope.
- Fast USB3-based oscilloscope (Picoscope) is supported.
- Support for NewAE CW-305 Artix7-based DUT.
- New control-board based on Digilent Basys3 has been developed. Using hardware-software codesign based on Microblaze, future upgrades to control firmware can be made primarily in software.
- New analysis scripts have been added (e.g. χ^2-test script).

Below, we describe the upgraded system in detail.

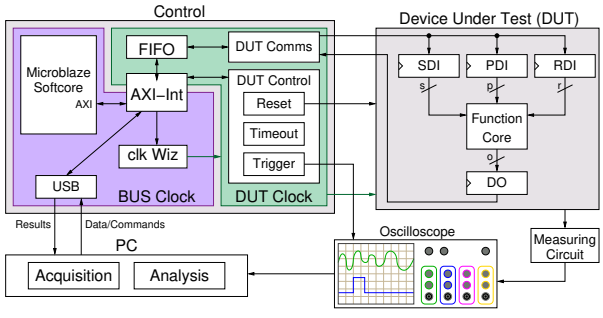


Fig. 1. FOBOS 2 Architecture

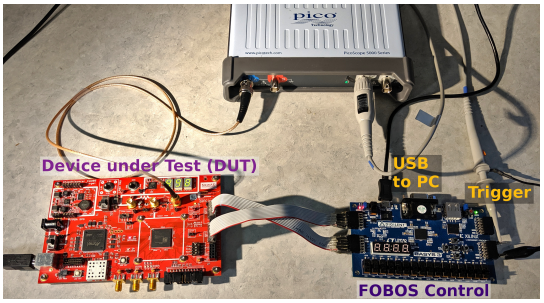


Fig. 2. Typical FOBOS 2 Setup

1) Data Acquisition module: The data acquisition module is used to capture traces. To reduce trace noise and for modularity, we use separate boards for the controller and the DUT. Fig. 1 shows the major components of the FOBOS capture system. Fig. 2 shows an example setup of the system. In this figure, control board appears on the right, the oscilloscope used is Picoscope 5000 (top), and the NewAE CW-305 (a low-noise Artix7-based SCA board) was used as DUT (left). The data acquisition module consists of the following components:

- **Control PC** The user interacts with the control PC which runs scripts to generate test vectors, communicate with the control board and retrieve traces from the oscilloscope. All scripts are written in Python which provides portability and good scientific computing libraries (e.g. NumPy). Traces are collected from the oscilloscope and stored in the control PC for analysis. The current version of the capture module supports two oscilloscopes models, Picoscope 5000 via USB and Agilent DSO6054A via Ethernet.
- **Control board** The control board is responsible for communication with the control PC and the DUT, and triggers the oscilloscope to capture power traces. FOBOS 2 supports Digilent Basys3 and Nexys-A7 control boards. Below, we describe the features of the control board.

  a) Communication: The control board handles communication with the control PC. It is connected to the PC using USB-UART. To process a test vector, the PC sends the vector to the control board. A simple protocol is used for PC-control board communication. The protocol provides headers to read/write configuration parameter (e.g. trigger mode) and instructions, i.e., encrypt using the DUT. The control board also handles communication with the DUT.

  b) Triggering: The control board is also responsible for generating the trigger signal which tells the oscilloscope when to start capturing the power waveform. The timing of the trigger signal relative to the beginning of data processing in the DUT is user-configurable.

  c) DUT Reset: For ciphers that take long time to execute, the controller can run the DUT for a configurable number of clock cycles and then reset it without waiting for it to complete. This helps reducing acquisition time.

  d) DUT Clock Generation: The control board is capable of supplying a clock signal to the DUT in the range from 400 KHz to 100 MHz.
- **DUT Board** The DUT board is where the function core (a.k.a victim algorithm) is instantiated. We provide a simple yet versatile wrapper to split data provided by the control board to separate streams. This wrapper is directly compatible with CAESAR Hardware API interface and is expected to be directly compatible with a future Hardware API for Lightweight Cryptography (LWC API). The wrapper receives data from the control board and distributes it into three FIFOs 1) the Public Data Input (PDI) FIFO (i.e. plaintext) 2) the Secret Data Input (SDI) FIFO (i.e. key) 3) the Random Data Input (RDI) FIFO to store random data used by protected implementations. The output is accumulated in the Data Out (DO) FIFO to be forwarded to the control board. To date we have validated Digilent Nexys3 boards (Spartan6 FPGA) and NewAE CW-305 SCA DUT (Artix7 FPGA) as DUT in FOBOS 2.

2) The Analysis module: The analysis module is used to process the traces acquired by the capture module and run SCA attacks or leakage assessment using TVLA and χ^2-test. The analysis software can run Correlation Power Analysis (CPA) [13]. CPA attacks are not performed in this work, but have been previously performed using FOBOS in [9].

B. Power Measurement

We measure the power consumption of the Vcc_{INT} rail by measuring the amplified voltage drop across a $1\,\Omega$ resistor while the DUT processes test vectors. Specifically, we used the XBP [14] which provides the $1\,\Omega$ resistor and TI-INA225 current sense amplifier. We connect Vcc_{INT} through the resistor in the XBP board to the DUT FPGA. When using the NewAE's CW-305 DUT, we cut the wire bridge between TP2 and TP3 and connected the Vcc wire from XBP to the FPGA through jumper JP7. Then the oscilloscope is used to measure the amplified voltage drop across the XBP's $1\,\Omega$ resistor. A Python script is used to calculate the power using the data collected from the oscilloscope.

C. Leakage Assessment Flow

FOBOS includes scripts that can perform fixed-vs-random TVLA. To perform this test, the user generates test vectors with fixed vectors randomly interleaved with random vectors. A meta file that records the type of each trace (i.e. fixed vs. random) is also generated. The test vectors are then fed to the capture module which processes them and produces power traces measured by the oscilloscope. The power traces are split into the random traces Q_r and fixed traces Q_f and passed to a script that calculates the t-values.

The current χ^2-test flow is based on two frequency classes "fixed" and "random;" as such, test vector generation and trace acquisition are identical to the TVLA. The two-class test differs only in the final analysis script, which calculates p-values for every sample, instead of t-values.

IV. Results

A. Power Measurements and Benchmarking

We performed power measurements on FPGA implementations of four NIST LWC Round-2 candidates plus AES-GCM with the following breakdown:

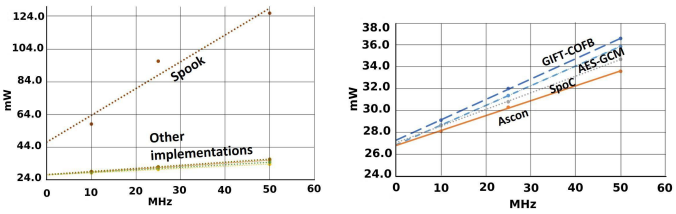
- 3 NIST LWC Round-2 implementations using basic-iterative architecture (SpoC, Spook, GIFT-COFB).
- 1 NIST LWC Round-2 implementation using a multi-cycle lightweight approach (Ascon-small).
- 1 existing standard AES-GCM, using a pipelined lightweight approach.

The implementation details for SpoC, Spook, and GIFT-COFB are documented in [15]. The implementation details for Ascon-small and AES-GCM are discussed in [16].

We used the upgraded FOBOS platform with a NewAE CW305 Artix7 target board. Picoscope 5000 oscilloscope with XBP was used to measure power. All five implementations use the same CAESAR LW Developer's package [17] and are benchmarked with Minerva hardware optimization tool [18] in Artix7 FPGA. Power is computed using the above methodology on 100 traces of four test vectors each (150 - 450 byte vectors) measured at 10, 25, and 50 MHz.

The measurements are found in Table I and are shown in Fig. 3. In Table I, abbreviations are Opt Freq (optimum frequency), LUT (look-up tables), P (power), E/bit (energy-per-bit). Opt Freq and area of SpoC, Spook, and GIFT-COFB are excerpted from [15]. P_{static} is estimated with linear interpolation. ΔP is calculated as $(|P_{max} - P_{min}|/P_{min}) * 100$. E/bit is calculated as $E/bit(nJ/bit) = P(mW)/TP(Mbps)$. Gradient $dP/dFreq$ is $dPwr(mW)/dFreq(MHz)$. Below, we discuss some observations:

- Ascon has the lowest power at 50 MHz, followed by SpoC, however, Ascon, AES-GCM, SpoC, and GIFT-COFB are relatively close. Spook is the outlier in Fig. 3a with power consumption much higher than other ciphers.
- Ascon has the smallest gradient, i.e., slope of increasing power with increasing frequency. However, Ascon has an area and minimum period larger than that of SpoC or AES-GCM, but has a lower dynamic power gradient.
- SpoC and Ascon have the smallest delta in percentage between maximum and mean power, which is a desirable design and security characteristic. Spook has up to a 28.4% delta between max and mean power at 50 MHz.
- GIFT-COFB has the lowest E/bit, 0.30 nJ/bit, versus the next lowest, Ascon, at 0.86 nJ/bit at 50 MHz. GIFT-COFB uses only slightly more power than Ascon at 36.6 vs 33.6 mW at 50 MHz. Since GIFT-COFB was implemented using basic-iterative architecture and Ascon using a multi-cycle approach, GIFT-COFB can probably be further optimized for power vs. E/bit.
- Static power of all ciphers (except Spook) is 27.0 mW, $\pm$ 1%. The static power of Spook is much higher, likely due to its larger area.



(a) All ciphers (b) Spook excluded

Fig. 3. Measured Power Consumption vs. Frequency.

B. Countermeasure Assessment

We performed leakage detection-based assessment on select lightweight implementations of the authenticated ciphers. We limit countermeasure assessment to AES-GCM and Ascon, since both unprotected and SCA-protected implementations of these ciphers are documented in [16], but no protected implementations of SpoC, Spook, or GIFT-COFB are documented in [15]. AES-GCM and Ascon were implemented in RTL-level hardware using VHDL and protected using threshold implementation against first-order DPA. Interested readers are referred to [16].

We instantiated the implementations above in the NewAE CW-305 Artix7 DUT and supplied a 1 MHz clock. This DUT features an Artix7 xc7a100tftg256-3 FPGA. We used a Basys3 control board and Picoscope 5000 oscilloscope to collect traces. The oscilloscope sampling frequency was 125 MS/s. Power measurements were taken from the CW-305 on-board low-noise amplifier (LNA) which amplifies the voltage drop across the on-board 0.1 Ω shunt resistor inserted between the core Voltage and the FPGA. We then collected 2000 traces using fixed-vs-random test vectors. Similar tests were done on the Spartan6 FPGA but this time, Tektronix CT-1 current probe was used for power measurement. In all cases, trace collection took less than 3 minutes for each cipher implementation.

TVLA results for Artix7 are shown in Fig 4. Spartan6 results were similar. The two horizontal lines at $|t| = 4.5$ mark the threshold. The χ^2-test has also been performed for the unprotected and protected Ascon implementations on Spartan6 FPGA using the same traces used for t-test. The results are shown in Fig. 5. In this figure, values of $p < 10^{-5}$ are considered a failure as discussed previously.

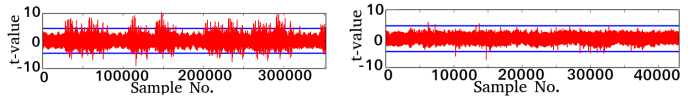
TVLA results show significant first order leakage in the unprotected versions as expected. On the other hand, the protected versions show t-values within the threshold which implies no significant leakage is detected. The χ^2-test on the unprotected Ascon detected leakage while the protected Ascon implementation shows no significant leakage which confirms the result obtained using TVLA.

V. Conclusions

We presented an upgraded FOBOS platform called FOBOS 2 suitable for performing power measurements and SCA resistance evaluation for hardware implementations of lightweight authenticated ciphers on modern Xilinx 7-series

TABLE I
CHARACTERISTICS OF AUTHENTICATED CIPHERS AND THEIR IMPLEMENTATIONS INVESTIGATED IN THIS WORK.

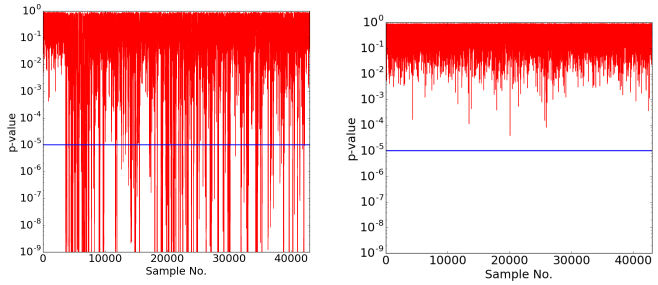
	Opt Freq MHz	Area LUTs	Cycles/ Block	Bits/ Block	P_{static} mW	Freq MHz	P_{mean} mW	P_{max} mW	ΔP %	TP Mbps	E/bit nJ/bit	Gradient dP/dFreq
AES-GCM	240	1532	205	128	26.9	10	28.6	29.7	3.7	6.2	4.59	0.1808
						25	31.4	33.2	5.9	15.6	2.01	
						50	35.9	38.3	6.8	31.2	1.15	
Ascon	232	1808	82	64	26.8	10	28.1	29.0	3.0	7.8	3.60	0.1369
						25	30.3	31.5	4.1	19.5	1.55	
						50	33.6	35.0	4.3	39.0	0.86	
SpoC	265	1344	111	64	27.0	10	28.6	29.4	2.8	5.8	4.96	0.1529
						25	30.8	31.7	2.9	14.4	2.14	
						50	34.7	36.1	4.1	28.8	1.20	
Spook	141	7082	145	256	47.0	10	58.8	71.0	20.8	17.7	3.33	1.642
						25	96.5	116.6	20.9	44.1	2.19	
						50	125.9	161.6	28.4	88.3	1.43	
GIFT-COFB	172	2695	53	128	27.3	10	29.1	30.1	3.5	24.2	1.20	0.1871
						25	32.0	33.5	4.6	60.4	0.53	
						50	36.6	38.5	5.2	120.8	0.30	



(a) AES-GCM Unprotected (b) Ascon Unprotected

(c) AES-GCM Protected (d) Ascon Protected

Fig. 4. TVLA results for AES-GCM and Ascon on Artix7 FPGA.



(a) Ascon Unprotected (b) Ascon Protected

Fig. 5. χ^2-test on Ascon. The blue line indicates the threshold.

FPGA and corresponding target boards. We used the platform above to measure power and compute energy-per-bit for (E/bit) for selected cipher candidates in the NIST LWC standardization process, including Spoc, Spook, GIFT-COFB and Ascon, and included a comparison to a current standard, AES-GCM. Through measurements on the Artix7 FPGA, we found that Ascon has the lowest power consumption at 50 MHz, and lowest incrementally increasing dynamic power with increasing frequency, and that GIFT-COFB has the lowest E/bit. We also reason that GIFT-COFB power can be further reduced through innovative architecture, without large sacrifices in energy efficiency. We additionally validated SCA protection countermeasures on Ascon and AES-GCM on FPGAs.

REFERENCES

[1] P. Kocher, J. Jaffe, and B. Jun, "Diferential Power Analysis," in *CRYPTO '99 - 19th International Conference on Cryptology*, Santa Barbara, CA, Aug. 1999.

[2] G. Goodwill, B. Jun, J. Jaffe, and P. Rohatgi, "A testing methodology for sidechannel resistance validation," in *NIST Non-Invasive Attack Testing Workshop*, 2011, p. 15.

[3] Rambus, "DPA Workstation Analysis Platform - Rambus," https://www.rambus.com/security/dpa-countermeasures/dpa-workstation-platform/, 2019.

[4] Riscure, "Side Channel Analysis Security Tools," https://www.riscure.com/security-tools/inspector-sca/, 2019.

[5] H. Guntur, J. Ishii, and A. Satoh, "Side-channel AttacK User Reference Architecture board SAKURA-G," in *2014 IEEE 3rd Global Conference on Consumer Electronics (GCCE)*. Tokyo, Japan: IEEE, Oct. 2014, pp. 271–274.

[6] C. O'Flynn, "A Framework for Embedded Hardware Security Analysis," Ph.D. dissertation, Dalhousie University, Halifax, Nova Scotia, Jun. 2017.

[7] E. Homsirikamol, P. Yalla, F. Farahmand, W. Diehl, A. Ferozpuri, J.-P. Kaps, and K. Gaj, "Implementer's Guide to Hardware Implementations Compliant with the CAESAR Hardware API," GMU, Fairfax, VA, GMU Report, 2016.

[8] A. Moradi, B. Richter, T. Schneider, and F.-X. Standaert, "Leakage Detection with the X2-Test," *IACR Transactions on Cryptographic Hardware and Embedded Systems*, pp. 209–237, 2018.

[9] R. Velegalati, J.-P. Kaps, and E. Department, "Towards a Flexible, Open-source BOard for Side-channel analysis (FOBOS)," in *Cryptographic Architectures Embedded in Reconfigurable Devices, CRYPTARCHI 2013*, Jun. 2013.

[10] A. Abdulgadir, W. Diehl, R. Velegalati, and J.-P. Kaps, "Flexible, Opensource workBench fOr Side-channel analysis," in *IEEE Hardware Oriented Security and Trust*, 2018.

[11] NIST, "Lightweight Cryptography — CSRC," https://csrc.nist.gov/projects/lightweight-cryptography/, 2019.

[12] CERG - GMU, "FOBOS Home Page," https://cryptography.gmu.edu/fobos/, 2019.

[13] E. Brier, C. Clavier, and F. Olivier, "Correlation Power Analysis with a Leakage Model," in *Cryptographic Hardware and Embedded Systems - CHES 2004*. Springer Berlin Heidelberg, 2004, pp. 16–29.

[14] J. Pham, "Benchmarking of Cryptographic Implementations on Embedded Platforms," Master's Thesis, GMU, 2015.

[15] B. Rezvani and W. Diehl, "Hardware Implementations of NIST Lightweight Cryptographic Candidates: A First Look," Tech. Rep., Jul. 2019.

[16] W. Diehl, F. Farahmand, A. Abdulgadir, J.-P. Kaps, and K. Gaj, "Face-off between the CAESAR Lightweight Finalists: ACORN vs. Ascon," in *2018 International Conference on Field Programmable Technology, FPT 2018*, Naha, Okinawa, Japan, Dec. 2018.

[17] P. Yalla and J.-P. Kaps, "Evaluation of the CAESAR hardware API for lightweight implementations," in *2017 International Conference on ReConFigurable Computing and FPGAs, ReConFig 2017*, Cancun, Mexico, Dec. 2017.

[18] F. Farahmand, A. Ferozpuri, W. Diehl, and K. Gaj, "Minerva: Automated hardware optimization tool," in *2017 International Conference on ReConFigurable Computing and FPGAs, ReConFig 2017*. Cancun: IEEE, Dec. 2017, pp. 1–8.