Toward Fair and Comprehensive Benchmarking of CAESAR Candidates in Hardware: Standard API, High-Speed Implementations in VHDL/Verilog, and Benchmarking Using FPGAs

Ekawat Homsirikamol, William Diehl, Ahmed Ferozpuri,
Farnoud Farahmand, Michael X. Lyons, Panasayya Yalla, and Kris Gaj
George Mason University

Abstract

Hardware performance of candidates in cryptographic contests has always been a very important evaluation factor, especially at the final stages of the competitions, when all remaining algorithms have been found to have adequate security strength. In CAESAR, for the first time, an attempt has been made to conduct hardware benchmarking of candidates at the very early stages of the contest, when the number of competing algorithms was still very large, namely there were still 29 authenticated cipher families remaining, with multiple variants for many of them.

This early hardware evaluation has become possible because of the three novel approaches. First, a standard hardware *Application Programming Interface (API)* for authenticated ciphers has been approved by the CAESAR Committee [1–2]. Second, a comprehensive *Development Package* including VHDL and Python code, supporting the development of implementations compliant with the CAESAR API, has been developed and thoroughly documented [3–4]. Third, the design teams have been asked to submit their *own Verilog/VHDL code* before the end of Round 2 [5–8].

In this talk, we will first summarize an effort by the GMU Benchmarking Team on fair and comprehensive evaluation of Round 2 CAESAR candidates, with the focus on high-speed implementations in VHDL/Verilog and benchmarking using Field Programmable Gate Arrays (FPGAs). We will then discuss lessons learned and proposed modifications to the CAESAR Hardware API, the corresponding Development Package, and supporting documentation. Finally, we will propose the exact plan and timeline for hardware benchmarking of Round 3 CAESAR candidates.

In the first part of the talk, we will review the API Compliant Hardware Description Language (HDL) Code development, applicable to both Round 2 and Round 3 CAESAR candidates. We will then give an overview of all submitted Round 2 VHDL/Verilog designs and divide them into those fully compliant, partially compliant, and non-compliant with the CAESAR API. In summary, 43 hardware design packages have been submitted to the CAESAR's mailing list, covering the majority of primary variants of 28 out of 29 Round 2 candidate families (all except Tiaoxin) [6]. Additionally, a hardware design for AES-GCM, used as a reference point for evaluation, has been developed. Some of the submission packages supported multiple variants (understood as versions of the same algorithm producing different outputs for the same input) and/or architectures (understood as designs of the same variant, producing the same output, and differing only in terms of performance and/or resource utilization). As a result, a total of about

75 variant-architecture pairs have been implemented and benchmarked [7]. Six submission packages have been found to be non-compliant with the CAESAR API, one partially compliant, and two compliant but highly inefficient. Out of the remaining 35 submission packages, 20 were developed by the GMU Team, and four each by the groups from Nanyang Technological University (Singapore) and Laboratoire Hubert Curien (Saint Etienne, France). The remaining groups submitted either a single design package or a package covering two related algorithms (such as Deoxys/Joltik and CLOC/SILC) [8].

The majority of the implementations followed the basic iterative architecture and were optimized for the maximum throughput to area ratio and maximum throughput. Only two of the submitted implementations were truly lightweight, and optimized primarily for area. Eight representative FPGA devices, used in several major FPGA development boards, were selected as platforms for hardware benchmarking. For seven of these devices, ATHENa [9, 10] was used for the target frequency and option optimization. For a device representing Xilinx Virtex 7 family, 25 Vivado optimization strategies were used instead [11]. No embedded memories and no embedded DSP units were allowed inside of the AEAD core [1, 4, 12, 13], in order to allow for the straightforward calculation of the Throughput to Area ratio and to assure the good correlation of the obtained rankings with any future ASIC implementation rankings. To the best of our knowledge, the entire submitted VHDL and Verilog code was generic and could be easily used as a basis for future ASIC benchmarking.

Only the best variant, architecture, and implementation of each algorithm, satisfying the selected security level, has been used for comparison. Each result was divided by the corresponding result for AES-GCM, leading to the relative values of throughput, area, and throughput to area ratio, representing an improvement factor over the current NIST standard [14].

Our presentation covers results for several representative FPGA devices. It includes two-dimensional graphs showing the position of each algorithm in the Throughput-Area space. Additionally, simpler bar graphs are used to clearly illustrate relative rankings of all candidates according to each individual performance metric. A demonstration of the ATHENa database of results, and its major features will be performed as well [12, 13].

An on-line presentation, summarizing Benchmarking of Round 2 CAESAR Candidates in Hardware, is already available on the GMU ATHENa web page [15].

In this DIAC talk, we will go beyond the Round 2 benchmarking effort, and discuss lessons learned and our proposed plan for the hardware evaluation of Round 3 candidates. In particular, we will cover several proposed extensions to the CAESAR Hardware API, the corresponding Development Package, and the Implementer's Guide. We will also outline priorities, such as the development of lightweight implementations (at least for lightweight candidates), investigating throughput vs. area trade-offs through multiple hardware architectures per candidate, conducting power and energy measurements, performing ASIC benchmarking, and verifying the suitability of the CAESAR API for real-world applications through the development of a realistic experimental setup.

References

[1] E. Homsirikamol, W. Diehl, A. Ferozpuri, F. Farahmand, P. Yalla, J.-P. Kaps, and K. Gaj, "CAESAR Hardware API," Cryptology ePrint Archive: Report 2016/626, available at http://eprint.iacr.org/2016/626.

[2] E. Homsirikamol, W. Diehl, A. Ferozpuri, F. Farahmand, P. Yalla, J.-P. Kaps, and K. Gaj, "Addendum to the CAESAR Hardware API v1.0," available at https://cryptography.gmu.edu/athena/index.php?id=CAESAR

[3] "Development Package for the CAESAR Hardware API," available at https://cryptography.gmu.edu/athena/index.php?id=CAESAR

[4] E. Homsirikamol, W. Diehl, A. Ferozpuri, F. Farahmand, and K. Gaj, "Implementer's Guide to the CAESAR Hardware API," available at https://cryptography.gmu.edu/athena/index.php?id=CAESAR

[5] CAESAR: Competition for Authenticated Encryption: Security, Applicability, and Robustness, Timeline, available at https://competitions.cr.yp.to/caesar.html

[6] VHDL/Verilog Code of CAESAR Candidates: Summary I, available at https://cryptography.gmu.edu/athena/CAESAR_HW_Summary_1.html

[7] VHDL/Verilog Code of CAESAR Candidates: Summary II, available at https://cryptography.gmu.edu/athena/CAESAR_HW_Summary_2.html

[8] "GMU Source Code of Round 2 CAESAR Candidates, AES-GCM, AES, AES-HLS, and Keccak Permutation F available at https://cryptography.gmu.edu/athena/index.php?id=CAESAR

[9] K. Gaj, J.P. Kaps, V. Amirineni, M. Rogawski, E. Homsirikamol, B.Y. Brewster, "ATHENa – Automated Tool for Hardware EvaluatioN: Toward Fair and Comprehensive Benchmarking of Cryptographic Hardware using FPGAs," 20th International Conference on Field Programmable Logic and Applications, Milano, Italy, Aug. 31st - Sep. 2nd, 2010.

[10] ATHENa, available at https://cryptography.gmu.edu/athena/index.php?id=ATHENa

[11] Xilinx. Vivado Design Suite, available at http://www.xilinx.com/products/design-tools/vivado.html

[12] ATHENa Database of Results: Rankings View, available at https://cryptography.gmu.edu/athenadb/fpga_auth_cipher/rankings_view

[13] ATHENa Database of Results: Table View, available at https://cryptography.gmu.edu/athenadb/fpga_auth_cipher/table_view

[14] M. Dworkin, "Recommendation for Block Cipher Modes of Operation: Galois/Counter Mode (GCM) and GMAC," NIST Special Publication 800-38D, Nov. 2007, available at http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-38d.pdf

[15] E. Homsirikamol, W. Diehl, A. Ferozpuri, F. Farahmand, and K. Gaj, "Benchmarking of Round 2 CAESAR Candidates in Hardware: Methodology, Designs & Results," available at https://cryptography.gmu.edu/athena/index.php?id=CAESAR