Software Benchmarking of the 2nd round CAESAR Candidates

Ralph Ankele1 and Robin Ankele2

1 Information Security Group, Royal Holloway, University of London
McCrea Building, Egham TW20 0EX, UK
ralph.ankele.2015@rhul.ac.uk
2 Department of Computer Science, University of Oxford
Wolfson Building, Oxford OX1 3QD, UK
robin.ankele@cs.ox.ac.uk

Abstract. The software performance of cryptographic schemes is an important factor in the decision to include such a scheme in real-world protocols like TLS, SSH or IPsec. In this paper, we develop a benchmarking framework to perform software performance measurements on authenticated encryption schemes. In particular, we apply our framework to independently benchmark the 29 remaining 2nd round candidates of the CAESAR competition. Many of these candidates have multiple parameter choices, or deploy software optimised versions raising our total number of benchmarked implementations to 232. We illustrate our results in various diagrams and hope that our contribution helps developers to find an appropriate cipher in their selection choice.

Keywords: Software benchmarks · CAESAR · 2nd round candidates · real-worlds usecases · tls · ssh · optimisations · aes-ni · sse · avx · neon

1 Introduction

An authenticated encryption (AE) scheme provides confidentiality, integrity and authenticity simultaneously. These goals are typically achieved by combining separate cryptographic primitives, an encryption scheme to ensure confidentiality and a message authentication code for integrity and authenticity. However, the generic composition, the combination of a confidentiality mode with an authentication mode, may lead to implementation errors and is often not efficient (e.g. the input stream has to be processed twice, for the encryption scheme and the message authentication code). Recently, some practical attacks [CHVV03], [Vau02] have confirmed the vulnerability of these AE schemes. In 2013 the CAESAR [Ber14] competition was announced to overcome these attacks and to achieve fast and efficient AE schemes and should be suitable for widespread adoption. In March 2014, 57 first round candidates were submitted to the CAESAR competition. Abed et al. [AFL14] provide a classification and a general overview of the first round candidates. In July 2015, the second round candidates were announced. The deadline for the announcement of the third round candidates is expected in end July 2016 that requires the CAESAR committee to make a selection decision which candidates will be included in the third round.

Beside of the security guarantees of the authenticated encryption mode, the software and hardware performance of the candidate becomes more and more important. One of the requirements imposed on all CAESAR submissions was that they should demonstrate the potential to be superior to AES-GCM in at least one significant aspect. One of those aspects that is particularly significant for the usage in widely deployed protocols
such as TLS, SSH and IPsec is software performance. Therefore, we have developed a benchmarking framework for automated and accurate software performance measurements and applied it to benchmark the remaining 29 second round candidates. Many of these candidates have multiple parameter choices, extending the number to 125 implementations. Overall, we have considered several software performance optimised implementations, raising the total number of benchmarked implementations to 232.

During our benchmarks, we perform and illustrate our measurements on a continuous function (with message and associated data sizes from 0 to 2048 bytes, in 128 byte steps) and on some fixed length sizes, relating to real-world usecases — 1 byte (one keystroke, SSH), 16 bytes (small payload), 557 bytes (average IP packet size), 1.5 kB (ethernet MTU size, TLS), 16 kB (max. TCP packet size) and 1 MB (e.g. file upload).

Related Work. In order to evaluate the software performance of the CAESAR candidates, Bernstein provides the SUPERCOP∗ [Ber16] framework, which is a general purpose framework for the evaluation of the software performance of cryptographic software. Saarinen presented the BRUTUS† [Saa16] framework, in order to test the CAESAR candidates for cryptographic weaknesses. Additionally to the software performance of the candidates, the hardware performance is also an important factor in the selection of the final portfolio. Homsirikamol et al. [HDF+15] present a hardware API to perform hardware benchmarks of the CAESAR candidates.

Contribution and Outline. In this paper, we present a complementary benchmarking framework to perform automatic and accurate measurements on authenticated encryption functions. We apply this framework to perform independent software benchmarks on the 2nd round CAESAR candidates and illustrate the results in various diagrams.

This paper is organised as follows. In Section 2, we introduce the 2nd round candidates of the CAESAR competition and classify them regarding their underlying cryptographic primitive. Section 3, presents our benchmarking framework and the measurement settings. In Section 4, we give an overview and discuss the software optimisation implemented in the 2nd round candidates. Finally, in Section 5 we present the results of the benchmarking and Section 6 concludes the paper.

2 Classification of the 2nd round CAESAR Candidates

In the CAESAR [Ber14] call for submissions the choice of underlying primitive and type of cipher was open for the designers. After 57 submissions for the first round, 29 submissions have been selected for the second round. Within these remaining candidates there are 15 based on block ciphers, 8 based on sponge constructions, 4 based on stream ciphers, 2 based permutations and 1 based on a compression function. Table 1 gives an overview of the 2nd round candidates of the CAESAR competition sorted by type.

Some of the candidates are online. An online cipher is defined such that the encryption of message block $M_i$ is only depending on already processed message blocks $M_1, \ldots, M_{i-1}$. Furthermore, an inverse-free scheme indicates if the inverse of an underlying primitive is needed (i.e. for the purpose of decryption of the underlying block cipher). An authenticated encryption scheme is nonce-misuse resistant, if the cipher is still secure under the repetition of nonces. Moreover, we can distinguish between nonce misuse resistance for online and offline ciphers. On the one hand, offline ciphers security can be claimed if the repetition of the ciphers leak only the ability to see a repeated message. On the other hand, for online ciphers security can be claimed, if nonces are repeated only the common prefix of the messages can be observed.

∗https://bench.cr.yp.to/supercop.html
†https://github.com/mjosaarinen/brutus
Table 1: Overview of the 2nd round CAESAR candidates. Possible types of schemes are: BC: block cipher based schemes, SC: stream cipher based schemes, Sponge: scheme based on Sponge construction (duplex or otherwise), P: permutation based scheme and CF: scheme based on compression function. Underlying primitives are: AES: when the full AES-128 is used, AES[r]: when a round reduced version of AES is used, other dedicated primitive or constructions: e.g. SHA2, SPN, LFSR. Entries for nonce misuse resistance are: None: when no security is claimed by the designers if nonces are repeated, Max: when for offline ciphers the repetition of nonces only leak the ability to see a repeated message, LCP: for online ciphers, all an adversary can observe is the longest common prefix of messages for repeated nonces.

<table>
<thead>
<tr>
<th>Name</th>
<th>Type</th>
<th>Primitive</th>
<th>Parallel Enc. / Dec.</th>
<th>Online Enc. / Dec.</th>
<th>Inverse-free</th>
<th>Security proof</th>
<th>Nonce misuse resistant</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>AEGIS</td>
<td>BC</td>
<td>AES[1]</td>
<td>✓ / -</td>
<td>✓ / ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>[WP14]</td>
</tr>
<tr>
<td>AES-COPA</td>
<td>BC</td>
<td>AES</td>
<td>partly / partly</td>
<td>✓ / ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>AES-JAMBU</td>
<td>BC</td>
<td>AES, Simon</td>
<td>✓ / -</td>
<td>✓ / ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>[W15]</td>
</tr>
<tr>
<td>AES-OTR</td>
<td>BC</td>
<td>AES</td>
<td>✓ / ✓</td>
<td>✓ / ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>[Min16]</td>
</tr>
<tr>
<td>AEZ</td>
<td>BC</td>
<td>AES[4]</td>
<td>✓ / ✓</td>
<td>✓ / ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>CLOC</td>
<td>BC</td>
<td>AES, TWINE</td>
<td>✓ / ✓</td>
<td>✓ / ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Decays</td>
<td>BC</td>
<td>AES</td>
<td>✓ / ✓</td>
<td>✓ / ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>ElGam</td>
<td>BC</td>
<td>AES</td>
<td>partly / partly</td>
<td>✓ / ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Joltik</td>
<td>BC</td>
<td>Joltik-BC</td>
<td>✓ / ✓</td>
<td>✓ / ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>OCB</td>
<td>BC</td>
<td>AES</td>
<td>✓ / ✓</td>
<td>✓ / ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>POET</td>
<td>BC</td>
<td>AES</td>
<td>partly / partly</td>
<td>✓ / ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>SCREAM</td>
<td>BC</td>
<td>SPN</td>
<td>✓ / ✓</td>
<td>✓ / ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>SHELL</td>
<td>BC</td>
<td>AES</td>
<td>partly / partly</td>
<td>✓ / ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>SILC</td>
<td>BC</td>
<td>AES, Present, LED</td>
<td>✓ / ✓</td>
<td>✓ / ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Taisen</td>
<td>BC</td>
<td>AES[1]</td>
<td>✓ / ✓</td>
<td>✓ / ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>OMD</td>
<td>CF</td>
<td>SHA2</td>
<td>✓ / ✓</td>
<td>✓ / ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Minalpher</td>
<td>P</td>
<td>SPN</td>
<td>✓ / ✓</td>
<td>✓ / ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>PAEQ</td>
<td>P</td>
<td>AES</td>
<td>✓ / ✓</td>
<td>✓ / ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>ACORN</td>
<td>SC</td>
<td>LFSR</td>
<td>✓ / ✓</td>
<td>✓ / ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>BSI-SIV</td>
<td>SC</td>
<td>ChaCha/Poly1305</td>
<td>✓ / ✓</td>
<td>✓ / ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>MORUS</td>
<td>SC</td>
<td>LRX</td>
<td>✓ / ✓</td>
<td>✓ / ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>TrivA-ck</td>
<td>SC</td>
<td>Trivium</td>
<td>partly / partly</td>
<td>✓ / ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Ascon</td>
<td>Sponge</td>
<td>SPN</td>
<td>✓ / ✓</td>
<td>✓ / ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>[DEMS15]</td>
</tr>
<tr>
<td>ICOLHE</td>
<td>Sponge</td>
<td>Keccak-like</td>
<td>✓ / ✓</td>
<td>✓ / ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Keccak</td>
<td>Sponge</td>
<td>Keccak-f</td>
<td>✓ / ✓</td>
<td>✓ / ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>[MGP15]</td>
</tr>
<tr>
<td>NORX</td>
<td>Sponge</td>
<td>LRX</td>
<td>✓ / ✓</td>
<td>✓ / ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>[BDP15]</td>
</tr>
<tr>
<td>π-cipher</td>
<td>Sponge</td>
<td>ARX</td>
<td>✓ / ✓</td>
<td>✓ / ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>PRIMATEs</td>
<td>Sponge</td>
<td>SPN</td>
<td>✓ / ✓</td>
<td>✓ / ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>[ABB14]</td>
</tr>
<tr>
<td>STRIBOB</td>
<td>Sponge</td>
<td>Streebog</td>
<td>✓ / ✓</td>
<td>✓ / ✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>[OSB15]</td>
</tr>
</tbody>
</table>

†† Data obtained from: https://aezoo.compute.dtu.dk/doku.php

3 Benchmarking Framework

Similar to the SUPERCOP [Ber16] and BRUTUS [Saa16] frameworks, we have developed a benchmarking framework to facilitate automatic software performance evaluations for authenticated encryption functions. Using our toolkit, one can perform independent and accurate software performance measurements (i.e. timing measurements), verify the implementation of the cipher under test, generate and illustrate statistical information about the software performance, and compare different parameters and optimisations of a cipher among its different optimisations or to other ciphers under test. Nevertheless, our benchmarking suite is currently optimised to perform measurements of authenticated encryption functions it remains open to be extended to perform benchmarking of other functionality and include other types of software tests.

Generally, our benchmarking framework provides the following benefits compared to the SUPERCOP and BRUTUS framework.

- Very simple with only focus on comparison of AE schemes
- Due to its simplicity open for extension of other software performance tests
4 Software Benchmarking of the 2nd round CAESAR Candidates

- Optimised TSC instruction (i.e. \texttt{rdtscp}) for accurate timing measurements
- Reduction of noise in the measurement using single user mode, averaging and median
- Benchmarking, verification, illustration and comparison of results

Using our framework, we built and benchmarked the 2nd round candidates (see Table 1) of the ongoing CAESAR competition. By July 2016, this includes the implementations of 29 candidates. Many of these candidates have multiple parameter choices, extending the number to 125 implementations. Overall, we have considered several software performance optimised implementations, raising the total number of benchmarked implementations to $232^\dagger$.

3.1 Measurement Setup

For our measurements we had two different measurement setups. The first measurements were taken on a 64-bit Sandy Bridge 2.3GHz dual-core Intel Core i5-2415M processor. The underlying hardware was a Mac Book Pro Early 2011 running OS X 10.11 (El Capitan) in single user mode. The second measurements were taken on a 64-bit Skylake 2.4GHz dual-core Intel Core i5-6300U processor. The underlying hardware was a DELL Latitude E7470 running Ubuntu 16.04 in single user mode. The implementations under test were built using the clang compiler version 6.1.0 (clang-602.0.53) and the gcc compiler version 5.4.0 (5.4.0-6ubuntu1-16.04.2). Additionally, we enabled the following compiler flags for compile time optimisation of the underlying source code: \texttt{-Ofast -fno-stack-protector -march=native}. In order to get rid of the noise (e.g. context switches) during the measurement and to get a clean and stable result in an uninterrupted environment, we choose to boot in the single user mode, which just starts the core operating system and provides a \texttt{sh} command interface, without starting other core elements of the OS X operating system (e.g. network services, multi-user support, graphical user interface) that may interfere with our benchmarking. Furthermore, we used the same method as described in the paper from Krovetz and Rogaway [KR11]. In their approach, they determine the performance of the underlying function as the median of 91 averaged timings of 200 measurements each.

3.2 High Resolution Methods for CPU Timing Information

In general, there are several ways to determine the software performance of a function under test. Each operating system offers several methods of accurate timing sources to be used to determine the runtime (e.g. latency) or throughput of a function. Using multi-core processors, one must be aware of handling hyper-threading and over-clocking of the CPU during their benchmarking, to avoid distorted or falsified measurements.

For the performance of high precision timing measurements on Windows platforms, one could use the HPET (i.e. High Precision Event Timer) or the QueryPerformanceCounter. On Unix based platforms one can use the posix conform \texttt{time()} and \texttt{clock()} functions, as used in the BRUTUS framework. However, a much more accurate timing source on x86 processors is the Timer Stamp Counter (TSC) as used in the SUPERCOP framework. The TSC is a 64 bit register and counts the number of cycles since the last reset of the machine. Using the \texttt{RDTSC} assembler instruction, one can read out the current TSC value and use it for accurate timing measurements.

In newer processor hardware, Intel processors practice out-of-order execution, where instructions are not always performed in the order as they appear in the source code. This could lead to a skewed measurement result. To avoid such errors, the \texttt{CPUID} instruction, a serialisation instruction, can be inserted, which

\footnote{$^\dagger$Number of implementations supported by our benchmarking hardware.}
Table 2: Real-world use case settings for our benchmarking.

<table>
<thead>
<tr>
<th>Message Size</th>
<th>Associated Data Size</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 byte</td>
<td>5 byte</td>
<td>one keystroke (e.g. SSH)</td>
</tr>
<tr>
<td>16 bytes</td>
<td>5 byte</td>
<td>small payload</td>
</tr>
<tr>
<td>557 byte</td>
<td>5 byte</td>
<td>average IP packet size**</td>
</tr>
<tr>
<td>1.5 kB</td>
<td>5 byte</td>
<td>ethernet MTU, TLS</td>
</tr>
<tr>
<td>16 kB</td>
<td>5 byte</td>
<td>max TCP packet size</td>
</tr>
<tr>
<td>1 MB</td>
<td>5 byte</td>
<td>file upload</td>
</tr>
</tbody>
</table>


forces every preceding instruction to complete before the execution of the \texttt{RDTSC} instruction.

In our framework, we make use of the \texttt{RDTSCP} \cite{Pao10} instruction, which is an optimised versions of the combination \texttt{CPUID} + \texttt{RDTSC}.

### 3.3 Authenticated Encryption Benchmarking Settings and Real-World Usecases

In our benchmarking, we implemented measurements on a continuous function (i.e. where we vary the message size and the associated data size) and over some fixed message and associated data sizes (i.e. representing some real-world usecases).

Overall, we consider message size from 0 to 1 MB and associated data size from 0 to 2 KB. The sizes of the key and nonce are fixed-size values, as determined by the CAESAR specification. In the case of the continuous function we consider message and associated data size from 0 to 2048 bytes in 128 byte steps. On the other hand, for the timing measurements with fixed message and associated data size, we consider the values in Table 2.

The associated data is fixed to 5 bytes§ for all measurements, which represents the typical size of the header of a TLS packet. Additionally, for better comparability, we represent the performance results of the authenticated encryption candidates in cycle per byte (cbp), which is a function of the throughput of the ciphers, instead of releasing the timings (i.e. latency) of the candidates.

### 4 Software Optimisations

Several ciphers are supported by software-optimised implementations. These optimisations are achieved by the usage of architecture-specific instructions and intrinsics. An intrinsic function is handled in a special manner by the compiler where the intrinsic function maps to some specific assembler instructions. Table 3 gives an overview of the mainly used instruction set extensions in the 2nd round candidates. Furthermore, in the following we give a short overview of the most common used instruction set extensions to achieve software optimised algorithms.

#### 4.1 AES-NI.

As many CAESAR candidates are AES based, they can obtain an increase in software performance by using the AES New Instructions \cite{Gue12}. These set of instructions are available beginning with the 2010 Intel processors based on the 32nm microarchitecture Westmere. The instruction set consists of six instructions that offer full hardware support

§ [http://netsekure.org/2010/03/tls-overhead](http://netsekure.org/2010/03/tls-overhead)
Table 3: Most common optimisations for the CAESAR candidates

<table>
<thead>
<tr>
<th>Instruction Set</th>
<th>Vendor</th>
<th>Microarchitecture</th>
<th>CAESAR Candidate</th>
</tr>
</thead>
<tbody>
<tr>
<td>AES-NI</td>
<td>Intel</td>
<td>Westmere</td>
<td>OCB, OTR, AEGIS, CLOC, SILC, POET, JAMBU, AEZ, Deoxys, PAEQ, Tiaoxin, TriviA-ck</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>OCB, CLOC, Keyak, Morus, OMD,</td>
</tr>
<tr>
<td>SSE</td>
<td>Intel</td>
<td>Pentium III</td>
<td>OCB, CLOC, Keyak, Morus, OMD,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Scream, Stribob, Tiaoxin, TriviA-ck</td>
</tr>
<tr>
<td>AVX</td>
<td>Intel</td>
<td>Sandy Bridge</td>
<td>NORX, OMD</td>
</tr>
<tr>
<td>AVX2</td>
<td>Intel</td>
<td>Haswell</td>
<td>Keyak, Minalpher, Morus, NORX</td>
</tr>
<tr>
<td>NEON</td>
<td>ARM</td>
<td>ARMv7-A</td>
<td>OCB, NORX, Scream, Stribob</td>
</tr>
</tbody>
</table>

for AES enabling fast and secure encryption and decryption. The AES New Instructions increase the performance by more than an order of magnitude for parallel modes (e.g. CTR mode), and provide roughly 2-3 fold gains for non-parallelisable modes (e.g. CBC mode). On a Westmere microarchitecture the AES-NI instructions can run with approximately 1.3 cycles per byte in parallel mode of operation.

4.2 SSE.

The Streaming SIMD Extensions [Sie09] contain around 70 instructions for high performance optimised implementations. The SSE instructions are available since Pentium III and the latest update SSE4.2 is available with the Nehalem microarchitecture. The extensions provide up to 16 128-bit registers (xmm0-15) and provide vector-mode operations, enabling parallel execution of one operation on multiple data. The speedup that can be achieved with the SSE instructions depends on whether the data can be parallelised, as operations are used in a vector-mode where the same operation is executed on multiple data.

4.3 AVX.

The Advanced Vector Extensions [Lom11] were introduced with the Sandy Bridge microarchitecture and extend the SSE instructions with a new set of instructions and a new coding scheme. The AVX1 instructions extend the previous xmm0-15 128-bit registers to 16 256-bit registers (ymm0-15). Furthermore, memory alignments have been relaxed and three-operand non-destructive operations have been added. Previously, with two-operand commands the source operand was always overwritten by the instruction (e.g. A = A + B). With the new three-operand instructions (e.g. A = B + C) an additional mov instruction to secure the source-value is not necessary. Additionally, the AVX2 [Bux11] instructions are an advancement of the AVX instructions supporting integer datatypes and adding vector shift operations. The new instructions were introduced with the Haswell microarchitecture.

4.4 NEON.

The implementation of the Advanced SIMD extensions used in ARM processors is called NEON [neo] and is available with the Cortex-A microarchitecture. ARM processors are used in many mobile devices such as tablets and mobile phones. The NEON instructions are 128-bit SIMD extensions providing a flexible and powerful acceleration as they perform packaged SIMD processing. Registers are therefore considered as vectors of elements with the same data type and perform the same operation on all lanes.
5 Results

We used our framework to benchmark all 2nd round CAESAR candidates. Figure 1 and Figure 2 show a comparison of the best implementations for each CEASAR candidate on SandyBridge and Skylake processors. To demonstrate the impact of a CAESAR candidate in a real-world setting, we also made benchmarks in the following settings. Figure 3 compares the best implementation per CAESAR candidate for the TLS setting on the Skylake architecture, while Figure 4 shows the same setting on the SandyBridge architecture. We chose the TLS setting, because we believe it is the most common usecase for authenticated encryption. Therefore, we set the associated data length to 5 bytes\(^5\) which represents the typical length of a TLS header. The message length is set to 1500 bytes, which corresponds to the MTU size of an ethernet frame. Moreover, we also made benchmarks for the SSH setting. In this setting, we fix the associated data length to 5 bytes and the message length to 1 byte (i.e. one key stroke). Figure 5 compares the best implementation per CAESAR candidate for the SSH setting on the Skylake architecture, and Figure 6 shows the same setting on the SandyBridge architecture. Furthermore, we created benchmarks for all candidates with increasing message/associated data lengths given in the appendix. Additionally, 3D plots are provided for the best implementation of each cipher, in order to illustrate the evolution in performance when the message length and the associated data length increase. We give detailed illustrations of these benchmarks in the appendices.

![Figure 1: Comparison of the best implementation (on the SandyBridge processor) for each CAESAR candidate. The associated data is fixed to 0 bytes.](http://netsekure.org/2010/03/tls-overhead)
Software Benchmarking of the 2nd round CAESAR Candidates

Figure 2: Comparison of the best implementation (on the SkyLake processor) for each CAESAR candidate. The associated data is fixed to 0 bytes.

Figure 3: Comparison of 2nd round CAESAR candidates in the TLS setting (on the SandyBridge processor).
Figure 4: Comparison of 2\textsuperscript{nd} round CAESAR candidates in the TLS setting (on the SkyLake processor).

Figure 5: Comparison of 2\textsuperscript{nd} round CAESAR candidates in the SSH setting (on the SandyBridge processor).
Software Benchmarking of the 2nd round CAESAR Candidates

Figure 6: Comparison of 2nd round CAESAR candidates in the SSH setting (on the SkyLake processor).

The benchmarking framework and our scripts to visualise the results are available online for verification purposes and to encourage other designers and researchers to use our framework for benchmarking their cryptographic schemes.

6 Conclusion

In this paper, we present a new benchmarking framework to perform software performance measurements of authenticated encryption schemes. We apply our framework to benchmark the 2nd round CAESAR candidates. Software performance is an important factor in real-world applications. Our benchmarks show that candidates supported by a software-optimised implementation achieve a huge speedup with regards to others, without an optimised implementation. The aim of the CAESAR competition is to provide a portfolio of authenticated encryption algorithms, including implementations optimised for software (and hardware), as in the eStream portfolio [est08]. Currently, nearly two thirds of the candidates are supported by at least one software optimisation. Nevertheless, with our benchmarks we want to encourage designers to increase the number of optimised implementations.

Acknowledgements

We would like to thank the designers of the CAESAR candidates, which supported us with their software implementations. The first author is supported by the Marie Sklodowska-Curie ITN ECRYPT-NET grant (Project Reference 643161).
References


[neo] Neon.


A Comparison of all Implementations per Candidate
Figure 13: deoxyseq128128v13_ref (blue), deoxyseq256128v13_ref (orange), deoxyseq128128v13_aesni (yellow), deoxyseq128128v13_ref (purple), deoxyseq256128v13_aesni (light blue), deoxyseq256128v13_ref (cyan)

Figure 14: elmd1000v2_ref (blue), elmd1001v2_ref (green), elmd101270v2_ref(yellow), elmd101271v2_ref (purple), elmd600v2_ref (green), elmd601v2_ref (cyan dotted), elmd61279v2_ref (magneta), elmd61271v2_ref (purple dotted)

Figure 15: hs1sivhiv1_ref (blue), hs1sivlov1_ref (orange), hs1sivv1_ref (yellow)

Figure 16: icepole128av2_ref (blue), icepole128v2_ref (orange), icepole256av2_ref (yellow)

Figure 17: aesjambuv2_aesni (blue), aesjambuv2_ref (orange), simonjambu128v2_ref (yellow), simonjambu64v2_ref (purple), simonjambu96v2_ref (green)

Figure 18: joltikeq12864v13_ref (light green), joltikeq6464v13_ref (orange), joltikeq80112v13_ref (yellow), joltikeq9696v13_ref (purple), joltikeq12864v13_ref (green), joltikeq6464v13_ref (light blue), joltikeq80112v13_ref (magenta), joltikeq9696v13_ref (blue)
Figure 19: ketjerv1_compact (blue), ketjerv1_reference (orange), ketjerv1_compact (yellow), ketjerv1_reference (purple)

Figure 20: lakekeyakv2_reference32bit (magenta), lakekeyakv2_compact (red), lakekeyakv2_generic32 (yellow), lakekeyakv2_generic32lc (purple), lakekeyakv2_generic64 (green), lakekeyakv2_generic64lc (light blue), lakekeyakv2_SandyBridge (dark blue)

Figure 21: riverkeyakv2_compact (orange), riverkeyakv2_generic32 (yellow), riverkeyakv2.generic32lc (purple), riverkeyakv2_generic64 (green), riverkeyakv2_generic64lc (light blue), riverkeyakv2_reference (brown), riverkeyakv2_reference32bits (dark blue), riverkeyakv2_SandyBridge (blue)

Figure 22: seakeyakv2_reference32bits (magenta), seakeyakv2_compact (orange), seakeyakv2_generic32 (yellow), seakeyakv2_generic32lc (purple), seakeyakv2_generic64 (green), seakeyakv2_generic64lc (light blue), seakeyakv2_SandyBridge (blue)

Figure 23: lunarkeyakv2_reference32bits (magenta), lunarkeyakv2_compact (orange), lunarkeyakv2_generic32 (yellow), lunarkeyakv2_generic32lc (purple), lunarkeyakv2_generic64 (green), lunarkeyakv2_generic64lc (light blue), lunarkeyakv2_SandyBridge (blue)

Figure 24: oceankeyakv2_reference32bits (magenta), oceankeyakv2_compact (orange), oceankeyakv2_generic32 (yellow), oceankeyakv2_generic32lc (purple), oceankeyakv2_generic64 (green), oceankeyakv2_generic64lc (light blue), oceankeyakv2_SandyBridge (blue)
Software Benchmarking of the 2nd round CAESAR Candidates

Figure 31: paeq128_aesni (pink), paeq128_ref (purple), paeq128t_aesni (black), paeq128t_ref (purple), paeq128tna_aesni (magenta), paeq160_ref (blue), paeq64_aesni (brown), paeq64_ref (yellow), paeq80_aesni (light green), paeq80_ref (green dotted)

Figure 32: omdsha256k128n128tau128v2_avx1 (red), omdsha256k128n128tau128v2_sse4 (yellow), omdsha256k128n128tau128v2_sse4 (purple), omdsha256k128n128tau64v2_ref (green dotted), omdsha256k128n128tau64v2_sse4 (red), omdsha256k128n128tau96v2_ref (blue dotted), omdsha256k128n128tau96v2_sse4 (red)

Figure 33: omdsha256k192n104tau128v2_avx1 (blue), omdsha256k192n104tau128v2_sse4 (yellow dotted)

Figure 34: omdsha256k256n104tau160v2_avx1 (red dotted), omdsha256k256n104tau160v2_sse4 (red dotted), omdsha256k256n248tau256v2_avx1 (yellow), omdsha256k256n248tau256v2_sse4 (red), omdsha256k256n256tau256v2_ref (green), omdsha256k256n256tau256v2_sse4 (green dotted)

Figure 35: omdsha512k128n128tau128v2_avx1 (light blue dotted), omdsha512k128n128tau128v2_sse4 (purple dotted), omdsha512k256n256tau256v2_avx1 (yellow), omdsha512k256n256tau256v2_sse4 (purple dotted), omdsha512k512n256tau256v2_avx1 (yellow), omdsha512k512n256tau256v2_sse4 (light blue dotted)

Figure 36: pi64cipher096v2_ref (red), pi64cipher128v2_ref (light green dotted), pi32cipher128v2_ref (pink dotted), pi32cipher256v2_ref (light green dotted), pi64cipher128v2_ref (yellow), pi64cipher256v2_ref (cyan dotted), pi64cipher128v2_goptv (black dotted), pi64cipher128v2_goptv (blue), pi32cipher128v2_goptv (magenta), pi64cipher256v2_goptv (purple), pi64cipher256v2_goptv (purple)
Figure 37: poetv2aes128_ni (blue), poetv2aes128_ref (orange), poetv2aes4_ni (yellow), poetv2aes4_ref (purple)

Figure 38: primatesv1ape120_ref (blue), primatesv1ape80_ref (orange), primatesv1gibbon120_ref (yellow), primatesv1gibbon80_ref (purple), primatesv1hanuman120_ref (green), primatesv1hanuman80_ref (cyan dotted)

Figure 39: scream10v3_ref (blue), scream10v3_sse (orange), scream12v3_ref (yellow), scream12v3_sse (purple)

Figure 40: shellsaes128v2d4n64_ref (red), shellsaes128v2d4n80_ref (blue), shellsaes128v2d5n64_ref (green), shellsaes128v2d5n80_ref (purple), shellsaes128v2d6n64_ref (yellow), shellsaes128v2d6n80_ref (cyan), shellsaes128v2d7n64_ref (light green), shellsaes128v2d7n80_ref (purple), shellsaes128v2d8n64_ref (orange), shellsaes128v2d8n80_ref (black)

Figure 41: present80n64silcv2_ref (green), aes128n12t8silcv2_aesni (blue), aes128n12t8silcv2_ref (red), aes128n8t8silcv2_aesni (yellow), aes128n8t8silcv2_ref (purple)

Figure 42: stribob192r2_8bit (blue), stribob192r2_bitslice (orange), stribob192r2_ref (yellow), stribob192r2_smaller (purple), stribob192r2_sse3 (green)
Figure 43: tiaoxinv2_nim (blue), tiaoxinv2_ref (orange)

Figure 44: trivia0v2_ref (blue), trivia128v2_ref (yellow), trivia128v2_sse4 (orange)

Figure 45: aesgcm128_ref (orange), aesgcm128_openssl (blue), aesgcm256_openssl (yellow)
B  3D Plots for Increasing Message and Associated Data Lengths

Figure 46: acorn128v2_opt

Figure 47: acdaaes256ochtaglen128v1_opt

Figure 48: aegin128l_aesnic

Figure 49: aes128n8t8clcv2_aesni

Figure 50: aes128n8t8silcv2_aesni

Figure 51: aescopav2_ref
Figure 52: aesjambuv2_aesni

Figure 53: aev4_aesni

Figure 54: ascon128av11_opt64

Figure 55: deoxysneq256128v13_aesni

Figure 56: elm600v2_ref

Figure 57: halsivlov1_ref
Figure 64: paeq64_aesni
Figure 65: pi64cipher128v2_ref
Figure 66: scream10v3_sse
Figure 67: shtellaes128v2d4n80_ref
Figure 68: stribo192r2_ssse3
Figure 69: tiaoxinv2_nim

Software Benchmarking of the 2nd round CAESAR Candidates
Figure 70: trivia0v2_ref

Figure 71: aes128otrsv2_ref

Figure 72: aes128poetv2aes4ls0lt0_ni
C Bar Charts for Real-World Message Sizes for each CAESAR Candidate

Figure 73: acorn128v2_opt

Figure 74: aeadaes128ocbtaglen128v1_opt

Figure 75: aegis128l_aesnic

Figure 76: aes128n12t8clocv2_aesni

Figure 77: aes128n8t8silcv2_aesni

Figure 78: aes128otrpv3_ref
28 Software Benchmarking of the 2nd round CAESAR Candidates

Figure 87: icepole128av2_ref

Figure 88: ketjesrv1_reference

Figure 89: seakeyakv2_SandyBridge

Figure 90: morus1280256v1_avx2

Figure 91: norx6441_ymm

Figure 92: omdsha512k512n256tau256v2_avx1

Figure 93: paeq80_aesni

Figure 94: p64cipher128v2_goptv
Figure 95: scream10v3_sse

Figure 96: shellaes128v2d8n80_ref

Figure 97: stribob192r2_ssse3

Figure 98: tiaoxinv2_nim

Figure 99: trivia0v2_sse4