Start your ENGINEs: dynamically loadable contemporary crypto

Nicola Tuveri and Billy Bob Brumley
Tampere University of Technology, Finland
{nicola.tuveri,billy.brumley}@tut.fi

Abstract. Software ever-increasingly relies on building blocks implemented by security libraries, which provide access to evolving standards, protocols, and cryptographic primitives. These libraries are often subject to complex development models and long decision-making processes, which limit the ability of contributors to participate in the development process, hinder the deployment of scientific results and pose challenges for OS maintainers.

In this paper, focusing on OpenSSL as a de-facto standard, we analyze these limits, their impact on the security of modern systems, and their significance for researchers. We propose the OpenSSL ENGINE API as a tool in a framework to overcome these limits, describing how it fits in the OpenSSL architecture, its features, and a technical review of its internals.

We evaluate our methodology by instantiating libsuola, a new ENGINE providing support for emerging cryptographic standards such as X25519 and Ed25519 for currently deployed versions of OpenSSL, performing benchmarks to demonstrate the viability and benefits.

The results confirm that the ENGINE API offers (1) an ideal architecture to address wide-ranging security concerns; (2) a valuable tool to enhance future research by easing testing and facilitating the dissemination of novel results in real-world systems; and (3) a means to bridge the gaps between research results and currently deployed systems.

Keywords: applied cryptography, public key cryptography, elliptic curve cryptography, software engineering, software implementation, OpenSSL

1 Introduction

Following current common best practices for the development of secure systems, most applications rely on cryptographic primitives for which widely accepted standards have been published after extensive studies to ensure that breaking them is believed computationally infeasible. As the available computation power increases and new attacks and algorithms to improve the performance of theoretic attacks are developed, these standards are periodically revised to raise the security level of the recommended primitives.

Considering the complexity of protocols and algorithms and the number of potential pitfalls concerning details at every level of abstraction, the software implementation of such evolving standards is a daunting and complex task and application developers are strongly advised not to implement their own crypto but to rely on existing well-established cryptographic libraries, among which OpenSSL\(^1\) is the most widely adopted. Still, in addition to the complexity of their programming interfaces, bugs and defects in these libraries are often the cause of cryptographic failures in the security layer of modern information systems.

\(^1\)https://www.openssl.org
Previous research focused on the development of new cryptographic software libraries as a solution to these problems and, among the published results, the NaCl [BLS12] project is especially relevant for this paper, aiming at providing practical and efficient strong confidentiality and integrity with a special emphasis on the simplicity of the provided programming interface. Although the project gathered some momentum, especially due to the libsodium\(^2\) fork, unfortunately, after five years, we see that this approach fails in meeting the needs of mainstream projects, resulting in a comparatively low adoption rate. Another interesting development sprouting off this research is HACL*\(^3\) [ZBPB17], a verified portable C cryptographic library that implements the NaCl API, producing code that is as fast as state-of-the-art C implementations, while providing mathematical guarantees on the absence of whole classes of potential bugs, memory safety, timing side-channel safety and functional correctness with respect to the published primitive specification.

Alongside the development of NaCl, research efforts were also spent on the related SUPERCOP benchmarking suite as part of the eBACS project [BL] for the benchmarking of cryptographic systems: many researchers from industry and academia submitted new cryptographic primitives or alternative implementations for benchmarking. Included submissions are benchmarked across different architectures and toolchains and undergo several automated tests. But still, most of these programming and scientific efforts fail to reach widespread adoption through mainstream libraries.

While our intent is not to belittle the efforts of the OpenSSL project, to motivate our contributions, here we briefly highlight some limits of the project from the point of view of researchers that might explain this gap:

- lack of unified quality reference documentation on the overall software architecture of the library, on API design choices, and frequently on the details of public and internal functions and data structures;

- complex ad-hoc build system;

- strong constraints on the choice of programming languages (C and custom “augmented” ASM syntax) and coding style;

- being a complex and huge project ongoing very active development, it implicates higher maintenance costs for external developers to keep their submissions up-to-date during the contribution process, which in turn can become quite long due to the double review constraint (see e.g. [KS09] which, while contributed in early 2009, did not get mainlined until late 2011, finally reaching a release version in early 2012);

- contributing a new feature usually requires splitting the contributed code in smaller units to facilitate and speed up the review process and collect feedback and consensus on development choices for the subsequent units. In turn, generally, this might slow down the overall development process and increase the maintenance costs for contributing developers;

- inclusion of new implementations or features usually undergoes a long decision-making process rarely compatible with research timelines;

- trust, quality assurance, and IPR issues force the core development team to be very conservative in the decision-making process, thus even in cases where time is a minor issue, the final result can still be a rejection. In these cases, it is up to the researchers to maintain a custom set of patches and documentation to make their contribution available to users and other researchers for further research (see e.g. [BCD+16] where the proposed cipher suite is no longer compatible with the newer OpenSSL API).

\(^{2}\)https://libsodium.org/
\(^{3}\)https://github.com/mitls/hacl-star
We expand on some of these limits in Section 2. Consequences of these limits can be seen, for example, in the implementation of emerging cryptography standards based on Curve25519 [Ber06, BDL+12]. X25519, the Diffie-Hellman cryptosystem, was originally released in 2005 and the properties of this curve and of its design promised simpler and faster implementations with properties enhancing its resistance to side-channel attacks. Later, in 2011, Ed25519 a digital signature system based on the twisted Edwards equivalent of Curve25519, was formally introduced, delivering fast short digital signature generation (and fast verification) with interesting properties concerning the resistance to side-channel attacks. These cryptosystems have since gathered momentum, gaining official support in OpenSSH in 2014, in BoringSSL in 2015, in the Google Chrome Internet browser TLS/QUIC protocol support in April 2016, and finally becoming part of IETF RFC 7748 [LHT16] (X25519) in January 2016 and RFC 8032 [JL17] (Ed25519) in January 2017.

OpenSSL officially added support for X25519 in August 2016, with release 1.1.0, but to this date it still lacks support for Ed25519. It is planned for the next 1.1.1 release and is included in the current pre-release version.

While implementation details are examined in Section 3.5, here we highlight that the original implementation chosen by the development team favored portability over optimization, and as a result using these cryptosystems in applications built on top of the current stable release of OpenSSL, does not yield the expected performance in comparison with other elliptic curve implementations (e.g. NIST P-256).

This does not seem to be a consequence of a lack of trusted optimized implementations for the most widespread architectures supported by OpenSSL, as the mentioned SUPERCOP project includes several implementations tested on different architectures and alternative libraries like libsodium and BoringSSL ship optimized implementations for popular architectures.

Our contribution. To address these limits, in this work:

- we present an analysis of the OpenSSL ENGINE API and its benefits for bridging gaps between cryptographic research and practical real-world implementations of cryptosystems;
- we develop an ENGINE to demonstrate how to use the ENGINE API as a framework to transparently integrate alternative implementations or new functionality in OpenSSL, making them available to existing applications;
- we evaluate experimental results, by benchmarking our libsuola ENGINE, demonstrating the viability and the benefits of the proposed solution across different versions of OpenSSL.

Even though performance gains are a nice side-effect, the main values of using the proposed framework come from

- the integration of missing functionality in end-of-life, yet still widely deployed, versions of OpenSSL;
- transparent access to the integrated functionality to existing applications, requiring exclusively changes to system configuration;
- freedom of choosing alternative implementations for new or existing functionality (e.g. choosing a formally verified implementation like HACL\textsuperscript{\textdaggerstar});
- ease of testing and benchmarking in real-world scenarios and dissemination to a larger user audience of novel implementations and primitives for researchers.

\textsuperscript{4}https://github.com/mitls/hacl-star
Outline. Section 2 presents further claims motivating our contribution. Section 3 contains an analysis of the OpenSSL architecture and the ENGINE API. Section 4 presents libsuola, an ENGINE we developed to demonstrate how to use the ENGINE API as a framework to offer alternative or missing implementations in OpenSSL. In Section 5 we present and evaluate experimental results comparing the default implementation of X25519 and Ed-25519 (or the lack thereof) against the implementations provided through our custom ENGINE and analyze the limits of our proposed methodology and how it addresses the concerns exposed in Section 2. Section 6 presents a brief analysis of related work, focusing on the mechanisms deployed by other security libraries, frameworks and operating systems to allow the transparent adoption of alternative cryptographic implementations. We conclude in Section 7.

2 Motivation

As mentioned in Section 1, this work is partially motivated by the rigidity of currently deployed, ubiquitous cryptography software libraries. This rigidity impacts real-world security at least via two avenues, which this section summarizes: it amplifies software assurance issues due to the small circle of contributors, and furthermore restricts the practical ability of OS vendors to provide predictable and steady support throughout the product life cycle.

2.1 Software assurance

The OpenSSL codebase is maintained by a small set of core developers on a voluntary basis. This practically restricts the ability of contributors and end users to, e.g., control what alternative cryptographic primitive implementations are featured in the library. Upkeep of end user custom builds is costly and does not scale well, hence the logical choice is to stick to the version and feature set provided by the OS vendor, overwhelmingly driven by the choices of OpenSSL core developers themselves. This is a poor model for security-critical software: it leads to e.g. limited accountability in the decision-making process, and lack of assurance that the included code is functionally correct. In what follows, we give some examples of extremely security-critical code paths in OpenSSL that lack software assurance (see e.g. [MS13] for a broader survey). We strongly reiterate here that our work should not be construed as diminishing the contributions of the OpenSSL project, but rather we use these observations to motivate our research and later demonstrate how our framework can alleviate this security burden, shifting it back towards developers and cryptographers.

Software defects. Biham et al. introduced the concept of bug attacks in 2008 [BCS08], highlighting the importance of cryptography implementation correctness to protect private keys. In 2011, Brunley et al. presented the first (and only, as far as we are aware) real-world bug attack [BBPV12], remotely recovering P-256 private keys from a TLS server by exploiting a carry propagation software defect in OpenSSL 0.9.8g finite field arithmetic (CVE-2011-4354). The defect (and vulnerability) resurfaced later in the Nettle cryptography library [Dub17, Sec. 3.1]. CVE-2014-3570 discloses a defect in multi-precision integer squaring, potentially affecting RSA, DSA, and ECDH cryptosystems in OpenSSL 0.9.8+. CVE-2016-7055 discloses a carry propagation bug leading to incorrect Montgomery multiplication results for 256-bit inputs, affecting ECDH cryptosystems for Brainpool P-512 elliptic curve in OpenSSL 1.0.2+. CVE-2017-3736 discloses a carry propagation bug leading to incorrect Montgomery squaring results, affecting RSA, DSA, and DH cryptosystems in OpenSSL 1.0.2+; CVE-2017-3732 and CVE-2015-3193 are similar but distinct defects. CVE-2017-3738 discloses an overflow bug in Montgomery arithmetic for 1024-bit moduli leading to incorrect multiplication results, affecting RSA, DSA, and DH cryptosystems in OpenSSL 1.0.2+. To summarize, these security vulnerability disclosures
unfortunately document a track record of functional correctness failures in OpenSSL, putting the security of private keys at risk. Our proposed methodology allows any developer to provide alternative implementations of cryptosystems decoupled from the OpenSSL codebase with higher functional correctness assurance—e.g., formally verified libraries [ZBPB17, ABB+17]. We demonstrate this by providing an option to select HACL* [ZBPB17] as the backend provider of the X25519 and Ed25519 implementations.

Side-channel security. In 2004, Percival discovered [Per05] the first cache-timing attack on public key cryptography in OpenSSL, recovering RSA keys by exploiting key-dependent table lookups in sliding window exponentiation (CVE-2005-0109). OpenSSL responded by (1) implementing (against expert advice to avoid cache line-level lookups [Ber05, Sec. 15]) the scatter-gather technique [Gue12] during exponentiation to reduce the amount of leakage; (2) adding a runtime flag that controls whether or not to follow this secure code path. Scatter-gather is not a full mitigation but a hedge [BS13], and in 2016 Yarom et al. implemented an attack utilizing cache-bank conflicts [YGH16]. OpenSSL responded (CVE-2016-0702) by tweaking scatter-gather parameters to further reduce the amount of leakage, hence the root cause of the vulnerability is still present in OpenSSL as of this writing. Also in 2016, Pereida García et al. discovered a defect in the way OpenSSL handles the runtime flag added in 2005 and used it to recover DSA private keys from TLS and SSH servers [PGBY16], i.e. the fix to the 2005 vulnerability was in fact not being activated; a bug present for over a decade (CVE-2016-2178).

In 2009, Brumley and Hakala discovered the first cache-timing attack on ECC in OpenSSL, recovering private keys from scalar multiplication [BH09]. OpenSSL responded by implementing constant-time paths on certain architectures for elliptic curves P-224, P-256, and P-521 based initially on Käsper’s work [Käs11] and later a P-256 contribution from Intel [GK15]. The OpenSSL team declined contributed patches to blanketly mitigate the vulnerability [Bru15]. Leaving the vulnerability present for all other curves, in 2014 Benger et al. improved the 2009 result and targeted the 256-bit BitCoin curve [BvdPSY14]. In 2015, van de Pol et al. further reduced the number of required signatures to recover private keys [vdPSY15], and Allan et al. even further in 2016 [ABF+16]—all from the initial code path shown vulnerable in 2009 but never fully fixed.

To summarize, unfortunately OpenSSL has a dubious track record regarding side-channel security: from stopgap countermeasures that do not stand the test of time, to implemented mitigations that are never activated due to defects, to no response at all when code paths are shown to be vulnerable. Our proposed methodology allows any developer to trump these OpenSSL design decisions and provide side-channel secure implementations largely independent from the OpenSSL codebase.

2.2 Release strategies
Reconciling software package EOL with OS distribution EOL is a challenging task w.r.t. both feature set and security. For example, consider RedHat Enterprise Linux (RHEL) version 7 with a 10 June 2014 release date. At that time, the stable version of OpenSSL was 1.0.1 and RHEL7 shipped with that package. Once a distribution chooses a software package version, this is essentially written in stone—the stock version of OpenSSL on RHEL7 started at 1.0.1 and will stay at 1.0.1 until RHEL7 EOL through 2024. Security issues will be patched with backported fixes and applied to this distribution’s OpenSSL flavor by the OS vendor. The reason the version is fixed is that other applications will link against e.g. shared libraries and, since there is no guarantee newer versions will maintain backwards compatibility, the simplest solution is not to change the version number.

Considering bugs, it is an extremely challenging task for the OS vendor to determine which fixes to backport. On the security side, they essentially rely on the risk analysis that takes place during the responsible disclosure process. This process is perilous—e.g. the
Start your ENGINEs: dynamically loadable contemporary crypto

Table 1: Selective OpenSSL features across versions. X25519 only appears since 1.1.0, not currently shipping with the majority of OS vendors due in part to the drastic OpenSSL API change between 1.0.2 and 1.1.0. As of this writing, Ed25519 does not feature in any OpenSSL release version.

<table>
<thead>
<tr>
<th>Feature</th>
<th>1.0.1</th>
<th>1.0.2</th>
<th>1.1.0</th>
<th>master</th>
</tr>
</thead>
<tbody>
<tr>
<td>nistz256</td>
<td>—</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>X25519</td>
<td>—</td>
<td>—</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Ed25519</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>✓</td>
</tr>
</tbody>
</table>

Figure 1: OpenSSL release strategy vs. OS vendor release strategy. For the OS distributions in question, it demonstrates that in every case the OS version EOL exceeds the OpenSSL version EOL. The data motivates this work’s proposed framework to provided alternative, extended, and/or new functionality largely decoupled from the OpenSSL core codebase, yet fully compatible due to dynamic loading.

previously discussed P-256 defect exploited by [BBPV12] was not flagged by either the OpenSSL security team or the OS vendors as a security issue, and hence went unpatched over four years in the wild before backported fixes rolled out to address CVE-2011-4354.

This task is slightly easier when the software in question is still maintained and has not reached EOL—the burden falls upon the software package project in question. But what happens when the software package reaches EOL before the OS does? This implies that official fixes are not available from the software package project, and the vendors must rely on contributed and third-party fixes and security analysis.

Before OpenSSL 1.0.2, OpenSSL had no official release strategy and roadmap. Understandably, this made the OS vendor’s job difficult, as there was no guarantee on how long official security fixes would be available. With the new release strategy at the end of 2014, OS vendors can make more informed choices regarding OpenSSL version inclusion, yet there will still be a coverage gap. For example, consulting Table 1 and Figure 1, stock OpenSSL on RHEL7 will not feature X25519 or Ed25519, and since OpenSSL 1.0.1 reached EOL at the end of 2016, RedHat is on the hook for backporting third-party security fixes until the end of 2024 as the OpenSSL team will no longer provide official security fixes.

To summarize, as Figure 1 shows the expected lifetime of OpenSSL versions is in fact much longer than the official EOL stated by OpenSSL. Compounded with the feature sets in Table 1, OS vendors for currently deployed distributions are essentially stuck with little to no OpenSSL support for emerging cryptography standards such as X25519 and Ed25519.

3 OpenSSL and the ENGINE API

OpenSSL is an open source project consisting of a general-purpose cryptographic library, an SSL/TLS library and toolkit and a collection of command line tools to generate and
handle keys, certificates, PKIs and other cryptographic objects and execute cryptographic operations.

The project officially started in December 1998, with release 0.9.1c forking the SSLeay project by Eric Andrew Young and Tim Hudson. Since then the project has seen a total of twelve major releases, of which two are currently actively supported (1.0.2 is the current Long Term Support version, initially released in January 2015 and supported until the end of December 2019; 1.1.0 is the current stable version, released in August 2016 and officially supported until the end of April 2018), while the next 1.1.1 version is being actively developed.

Being ubiquitous on the server side (e.g. powering the HTTPS support for the Apache and nginx web servers which combined cover almost two-thirds of Internet active sites according to NETCRAFT’s July 2017 Web Server Survey\(^5\)) and in many client tools, the OpenSSL project has become a de facto standard for Internet security.

The project is written mainly in the C programming language and assembly for optimized implementations and supports a plethora of platforms, including a wide range of hardware architectures running most Unix and Unix-like operating systems, OpenVMS and Microsoft Windows.

### 3.1 Architecture of the OpenSSL project

The diagram in Figure 2 depicts the current architecture of the OpenSSL project. It is based on the 1.1.1 development version, but can be applied with minor changes also to the 1.0.2 and 1.1.0 release branches.

The three main blocks in the diagram are the OpenSSL binaries, consisting of the command-line tools included in the project, which are linked against the other two main blocks of the OpenSSL project diagram, representing the two software libraries implementing the core project functionality.

---

OpenSSL **libssl** implements the SSL/TLS library, dealing with the implementation details of the standardized network protocols and extensions, linking against OpenSSL **libcrypto** for the implementation of the underlying cryptographic functionality required by the network protocols.

OpenSSL **libcrypto** contains the actual cryptographic implementations and also acts as portability layer across the different platforms supported by the OpenSSL project, and is further organized into several independent modules.

Focusing on the OpenSSL **libcrypto** block, the diagram is organized so that, in general, the vertical position is correlated with the level of abstraction of each depicted module (going from the top application-oriented modules to the bottom modules providing lower-level implementations or low-level functionality mostly dealing with operating systems and system libraries abstractions), while the horizontal position, with exceptions, is coupled with the “topic” of each module or group of modules (the left part of the **libcrypto** block is more closely related to cryptographic implementations, while the right part comprises increasingly general purpose modules).

From a point of view of an application wishing to use OpenSSL to perform cryptographic operations including encryption and decryption, digital signature schemes, key derivation algorithms, and message digests, the high-level **EVP API** is the recommended interface: the EVP block provides an abstraction level to handle these operations using abstract methods and data types to decouple application programmers from the details of the actual low-level implementations of each cryptosystem. In particular, among the sub-modules included in the **EVP API**, **EVP_PKEY** abstracts asymmetric cryptosystems, **EVP_MD** abstracts message digest cryptosystems, and **EVP_CIPHER** abstracts symmetric encryption/decryption cryptosystems.

These abstractions are made possible by the use of abstract context and method structures which describe every EVP cryptosystem in terms of pointers to metadata structures describing the parameters of the specific cryptosystem, and functions to manipulate such metadata structures or derived data structures describing the internal status of an instance of a specific cryptosystem, implementing the actual functionality provided by a cryptosystem.

Therefore, most **EVP API** functions ultimately act as wrappers around the library internal **OBJECTS** table, which can be queried by a numeric identifier (**NID**) to retrieve the actual method structures associated with a particular indexed cryptosystem.

This indirection mechanism can also be used as a way to provide multiple alternative implementations for a given cryptosystem, by manipulating the querying algorithm to select among the method structures registered for the same **NID**: this approach is what ultimately powers the **ENGINE API** described in the following paragraphs.

### 3.2 What is an OpenSSL ENGINE?

OpenSSL 0.9.6 introduced a new component to support alternative cryptography implementations, most commonly for interfacing with external crypto devices (e.g. accelerator cards), in the form of **ENGINE** objects.⁷

These objects act as “containers for implementations of cryptographic algorithms” and can be statically linked in the OpenSSL library at compile-time or dynamically loaded at run-time in and out of the running application with low-overhead from external binary objects implementing the **ENGINE API**⁸ through a special built-in **ENGINE** called “dynamic”.

Such dynamic **ENGINEs** are particularly interesting due to the flexibility they provide:

---

⁶[https://www.openssl.org/docs/man1.1.0/crypto/evp.html](https://www.openssl.org/docs/man1.1.0/crypto/evp.html)
⁸[https://www.openssl.org/docs/man1.1.0/crypto/ENGINE_init.html](https://www.openssl.org/docs/man1.1.0/crypto/ENGINE_init.html)
they allow to replace compiled-in implementations affected by known problems with newer ones, maintaining compatibility with existing applications;
• they allow hardware vendors to release self-contained shared-libraries to add support for arbitrary hardware to work with applications based on OpenSSL, keeping their software outside of the main OpenSSL codebase;
• they allow to reduce the memory impact of OpenSSL, by avoiding to statically link support for unneeded hardware at compile-time in favor of system configuration or automatically probing for supported devices at run-time and dynamically loading only the required cryptographic modules.

Dynamic ENGINEs are also interesting for software implementations:
• they allow to replace compiled-in software implementations in case of bugs, vulnerabilities or sub-optimal performances with newer alternative implementations;
• they provide an alternative in case of issues with the OpenSSL core development team decision-making process, decoupling the decision to adopt the OpenSSL library from the choice of individual cryptosystem implementations, without incurring the costs of maintaining a fork of the OpenSSL project or patch sets, while also providing transparent binary compatibility with existing applications;
• they allow to backport newer cryptosystems in previous versions of the OpenSSL library and existing applications based on it;
• they allow to easily add new cryptosystems or new implementations for already compiled-in cryptosystems to the OpenSSL library, providing a convenient way to test and benchmark new software implementations in a real-world context;
• they offer a greater degree of freedom from the OpenSSL project toolchain, allowing developers to use different programming languages and build systems, further lowering the development and maintenance costs for developing plugin alternative implementations or new functionality;
• they offer flexibility to solve licensing issues: currently the OpenSSL project is released under a “dual license” scheme, under the OpenSSL License, (a derivate of the Apache License 1.0) and the SSLeay License (similar to a 4-clause BSD License), and is in the process of transitioning to the Apache License 2.0. Contributors are thus forced to release their work under these licenses, which may be an issue especially when reusing code from projects released under a proprietary license or an incompatible copyleft license. Being objects dynamically loaded at runtime, engines can benefit from usually more flexible licensing requirements, providing a bridge towards software released under different licenses.

What functionality do ENGINEs provide? The cryptographic functionality that can be provided by an ENGINE implementation includes the following abstractions:
• RSA_METHOD, DSA_METHOD, DH_METHOD, ECDH_METHOD, ECDSA_METHOD: providing alternative RSA/DSA/etc. implementations;
• RAND_METHOD: providing alternative (pseudo-)random number generation implementations;
• EVP_CIPHER: providing alternative (symmetric) cipher algorithms;
• EVP_MD: providing alternative message digest algorithms;
• EVP_PKEY: providing alternative public-key algorithms.
3.3 Structural and functional references to dynamic ENGINEs

Due to the modular nature of the ENGINE API, references to ENGINE objects require special handling, organized over two levels of reference-counting matching the ways such objects are used in the OpenSSL library.

Any ENGINE pointer is inherently a structural reference, providing a guarantee that the ENGINE structure cannot be deallocated until the reference is released. Structural references are sufficient to query or manipulate the data of an ENGINE object and are generally used to instantiate a new ENGINE, to iterate across the OpenSSL internal list of loaded ENGINEs, or to read information about an ENGINE. All structural references obtained through the ENGINE API should be released by a call to the ENGINE_free() function, and the ENGINE object itself will be automatically cleaned up and deallocated upon release of the last structural reference.

However, a structural reference does not guarantee that the ENGINE object has been initialized and is ready to execute any of the implemented cryptosystems: most ENGINEs are typically used to support specialized hardware, so they might be loaded but unable to be initialized unless the required hardware is actually present in the system.

As such, to use the actual functionality of an ENGINE, a functional reference is required: a functional reference is a specialized form of a structural reference (derived from it by calling the ENGINE_init() function) which guarantees that the ENGINE has been correctly initialized and is ready to perform the implemented cryptographic operations, until the functional reference is explicitly released by calling the ENGINE_finish() function (which in turn would also automatically release the implicit structural reference).

3.4 Anatomy of a dynamic ENGINE

At the highest level of abstraction, a dynamic ENGINE can be split into two functional blocks.

One block contains all the alternative implementations for the cryptosystems provided by the ENGINE: this part mainly consists of a collection of structs for each cryptosystem, each linking to the actual functions implementing its operations. For example, an EVP_MD message digest struct would reference the actual init(), update(), and final() functions implementing the OpenSSL message digest streaming API, in addition to some utility functions allowing the OpenSSL library to cleanly handle, clone and destroy instances of the provided message digest implementation.

Every such struct would be individually registered against the OpenSSL library during the bind() process, and structs of the same kind (e.g. all the EVP_MD structs, all the EVP_PKEY_meth structs, etc.) are glued together by functions registered in the ENGINE object that allow the OpenSSL library to query the engine for lists of provided algorithms or a specific algorithm indexed by nid.

The other block contains the bind() method and the initialization and deinitialization functions:

- the bind() method is called by the OpenSSL built-in dynamic ENGINE upon load and is used to set the internal state of the ENGINE object and allocate needed resources, and to set its id and name, pointers to the init(), finish(), and destroy() functions;

- the init() function is called to derive a fully initialized functional reference to the ENGINE from a structural reference;

- the finish() function is called when releasing an ENGINE functional reference, to free up any resource allocated to it;
• the `destroy()` function is called upon unloading the ENGINE, when the last structural reference to it is released, to cleanly free any resource allocated upon loading it into memory.

### 3.5 Curve25519, X25519 and Ed25519 in OpenSSL

OpenSSL 1.1.0 introduced support for X25519: as a consequence of the way Curve25519 and X25519 are defined, instead of adding the curve inside the EC module containing every other elliptic-curve cryptosystem implementation based on prime or binary fields, the OpenSSL developers decided to add a separate ECX sub-module defining dedicated EVP_PKEY_meth structures directly linking to a self-contained portable C implementation of Curve25519 and the underlying finite field arithmetic.

The current OpenSSL development version (1.1.1) expanded the ECX sub-module to add a new EVP_PKEY_meth supporting the Ed25519 digital signature scheme, reusing in large part the same portable C low-level implementation, and recently added an optimized implementation for X25519 for 64-bit architectures.

The actual low-level portable C implementation is closely based on the ref10 version of Ed25519 in SUPERCOP 20141124\(^9\) which, although portable, suffers huge speed penalties compared to implementations optimized for specific architectures or even portable 64-bit C code.

As a result, comparing the benchmarks of X25519 (in OpenSSL 1.1.0) and Ed25519 (in OpenSSL 1.1.1) cryptosystems with operations of analogous EC cryptosystem (e.g. over the NIST P-256 elliptic curve) does not yield the performance benefits expected from [Ber06, BDL+12], as shown later in Section 5.

### 4 The libsuola ENGINE

libsuola\(^10\) takes advantage of the ENGINE API described in Section 3 to provide alternative software implementations for X25519 and Ed25519, working as a bridge between OpenSSL and an external library.

`libsuola` itself is a shallow loadable module, i.e. it does not contain cryptographic implementations: the core part of `libsuola` is the code facing the OpenSSL APIs, that takes care of initializing data structures and registering abstract methods; these abstractions are routed to a ‘provider’ module in `libsuola`, which finally links against the selected external library to provide the actual cryptographic functionality. To demonstrate the potential of the proposed methodology, we provide three different providers, among which the user can select at build time, each linking against a different implementation, and further described in Section 4.4.

Figure 3 shows the architecture of `libsuola` (compiled selecting the `libsodium` provider) and its interactions with `libsodium` and OpenSSL. Selecting a different provider results in linking against another library or embedding, through static linking, an implementation inside the provider object, but the changes affect only the provider object, which is the only element providing cryptographic functionality.

For the installed applications that will benefit from the added functionality or alternative implementations, `libsuola` is completely transparent: while applications can include code to explicitly force their instance of OpenSSL to load the `libsuola` ENGINE, most applications will be completely unaltered and simply loading OpenSSL when they are executed; in this case the OpenSSL library is instructed to load at run-time the `libsuola` ENGINE through its system configuration files.

\(^9\)http://bench.cr.yp.to/supercop.html
\(^10\)https://github.com/romen/libsuola
Independently of the way it is instructed to load the libsuola ENGINE, internally the OpenSSL library creates an instance of the built-in dynamic ENGINE, which in turn uses the OS dynamic loader to load libsuola.

As soon as it is loaded into memory, the dynamic ENGINE calls the libsuola suola_bind() function which in turn:

- sets the id and name fields in the ENGINE structure;
- sets pointers to the destroy(), init() and finish() functions in the ENGINE structure;
- registers NIDs for X25519, Ed25519 and the identity message digest: NID_X25519 is defined in OpenSSL since version 1.1.0, while NID_ED25519 is only defined in the current development version. This step ensures that even in previous versions of OpenSSL the internal OBJECTS table includes definitions for both cryptosystems;
- creates the structures to handle each implemented cryptosystem: the EVP_MD md_identity describing the identity message digest and two (EVP_PKEY_meth, EVP_PKEY_ASN1_meth) pairs of structures for X25519 and Ed25519. Each of these structures is also registered in the internal OpenSSL OBJECTS table under the corresponding NID;
- sets pointers to the digests(), pkey_meths() and pkey_asn1_meths() callbacks in the ENGINE structure: these callbacks allow the OpenSSL library to manipulate the ENGINE handle querying for lists of implemented methods indexed by NID;
- calls the provider suola_implementation_init() function to initialize the internals of the provider implementation.

---

**Figure 3:** libsuola architecture and its interactions with libsodium and OpenSSL.
The rest of libsuola consists of the collection of functions used to implement the functionality described by each of the structures registered in the suola_bind() function, which map OpenSSL structures to the provider functionality. These are briefly described in the following paragraphs.

4.1 libsuola X25519

X25519 is implemented through an EVP_PKEY_meth and a corresponding EVP_PKEY_ASN1_meth. The first one mainly includes the definition of a keygen() and a derive() method, while the second includes methods for encoding and decoding X25519 private and public values.

The keygen() method delegates its core functionality to the suola_scalarmult_curve25519_base() function, which is then routed to the actual implementation through the selected provider.

The derive() method, called once the peer key has been set, performs a generic point scalar multiplication over Curve25519 through the provider suola_scalarmult_curve25519() function.

4.2 libsuola Ed25519

Similarly, Ed25519 is also implemented through an EVP_PKEY_meth and a corresponding EVP_PKEY_ASN1_meth. The first one mainly includes the definition of a keygen(), a sign() and a derive() method; the second includes methods for encoding and decoding Ed25519 private and public values, as well as a control method that the OpenSSL library uses to query for the default message digest algorithm associated with the digital signature cryptosystem, which in libsuola is set to return the NID for md_identity.

The keygen() method for Ed25519 is mostly a wrapper around the provider suola_sign_ed25519_seed_keypair() function.

The sign() and verify() methods are wrappers around suola_sign_ed25519_detached() and suola_sign_ed25519_verify_detached() functions in the provider. These functions implement a PureEdDSA signature scheme, i.e. where the message to be signed is directly used as an input to the signature generation algorithm, without a pre-hashing step as opposed to traditional digital signature schemes based on RSA, DSA or ECDSA or to the PreHashEdDSA signature scheme.

OpenSSL 1.1.1 introduced new API functions to support one-shot signature generation and verification cryptosystems like Ed25519, but using this approach in libsuola would make it incompatible with previous versions of OpenSSL. Instead, we chose to work around the limitations of the traditional API by adding a custom md_identity message digest method to be used in the default pre-hash step performed by OpenSSL before the actual signature generation or verification.

4.3 libsuola md_identity

The EVP_MD md_identity is a custom message digest algorithm behaving as an identity function to be used as the pre-hash algorithm in the libsuola implementation of PureEd25519: the output of this message digest is a copy of the input message.

The EVP_MD API in OpenSSL can be reduced to a streaming approach based on calling an init() function before starting the hash computation, repeated calls to an update() function until the whole input message is consumed, and a final call to a final() function to finalize the hash and retrieve its output. For md_identity these functions are implemented as described below:

init() Securely allocates an empty small buffer as the internal status of the message digest algorithm, tracks its length and sets to 0 the counter of used buffer bytes;
**update()** Depending on the size of the input block and the availability of unused space, the internal buffer is securely reallocated to sufficiently increase its size. A copy of the new input block is then concatenated to the existing data in the buffer while tracking the length of the whole buffer and of the data contained in it;

**final()** OpenSSL EVP_MD API limits the maximum size of a message digest algorithm output thus, considering that no actual finalization is required and that the md_identity algorithm is used only by code inside libsuola, we decided to work around this limit by gaining access to the internal md_identity buffer directly from the Ed25519 EVP_PKEY_meth implementation, without ever using the final() function. As a result, the implementation of this function simply returns an error.

### 4.4 Providers

The actual cryptographic functionality provided by libsuola is entirely contained in the provider module. It determines which library libsuola is linked against, and routes the function calls from the OpenSSL structures described above to the relevant functions in the selected implementation.

To demonstrate the potential of the proposed methodology, we provide three different alternative providers:

- the first provider, to which Figure 3 refers, links against libsodium;
- a second provider links against another fork of NaCl: the HACL* library, providing a formally verified implementation;
- as a proof of concept, we develop a third provider which does not depend on an external library and internally embeds the provided functionality through static linking.

**libsuola-sodium.** For our first provider, libsodium was selected as a modern cryptographic library, designed emphasizing high security and ease-of-use, and addressing from its inception as a fork of NaCl many of the shortcomings we described in OpenSSL.

On the platforms we tested, libsodium also happens to generally provide faster implementations than the equivalent OpenSSL ones. Thus while the aim of our project is to provide support for emerging cryptography standards such as X25519 and Ed25519 for currently deployed versions of OpenSSL, we also notice that the provided alternative implementations perform better than the built-in implementations included in the OpenSSL stable 1.1.0 version (supporting X25519 only). Even in the current development 1.1.1 version of OpenSSL (supporting both X25519 and Ed25519), the libsodium Ed25519 implementation provided through libsuola appears to have faster performance than the OpenSSL default implementation.

When this provider is selected, as depicted in Figure 3, libsuola is dynamically linked against libsodium, hence when OpenSSL loads the ENGINE the dynamic loader would automatically load both libsuola and libsodium.

The suola_implementation_init() function in this case calls sodium_init() during the binding process to initialize libsodium internals.

**libsuola-hacl.** HACL*11, presented in [ZBPB17], is a very interesting target for our framework, as it presents a practical implementation of a cryptographic library, which is formally verified for memory safety and functional correctness with respect to its published standard specification, and aims to be “as fast as state-of-the-art C implementations, while implementing standard countermeasures to timing side-channel attacks”.

11https://github.com/mitls/hacl-star
Through our proposed methodology it then becomes possible to integrate the guarantees of projects like this, that mathematically prove the absence of the defects we listed in Section 2.1, without needing to alter existing application to use a different cryptographic library.

HACL* uses the same C API as libsodium for the NaCl constructions, so excluding some obvious prefix renaming, the mappings described in Figure 3 with respect to libsodium are analogous to the ones provided by this provider for the HACL* library.

libsuola-donna. As a proof of concept, we implemented an alternative version of libsuola that statically links at compile-time against an implementation of Curve25519 and X25519 based on the work of Adam Langley and Andrew Moon.\footnote{https://github.com/floodyberry/ed25519-donna}

Through this provider, the contents of the EVP_PKEY_meth structures for Ed25519 and X25519 are routed as follows:

- the keygen() functions ultimately link to, respectively, ed25519_publickey() and curved25519_scalarmult_basepoint() from floodyberry/ed25519-donna/ed255-19.c;
- the Ed25519 sign() and verify() functions are implemented via, respectively, ed25519_sign() and ed25519_sign_open() from floodyberry/ed25519-donna/~ed25519.c;
- the X25519 derive() function is not supported by the code in the floodyberry/~ed25519-donna repository, so we opted to use the portable 64-bit C donna implementation included in SUPERCOP\footnote{http://bench.cr.yp.to/supercop.html} in crypto_scalarmult/curve25519/donna_c64/smult.c. This implementation is not compatible with our 32-bit ARM testing environment, so on this platform at build time we replace it with another implementation from SUPERCOP, in crypto_scalarmult/curve25519/neon2/scalarmult.s, contributed by Bernstein and Schwabe [BS12], optimized for the NEON Advanced SIMD extension.

5 Experimental results

Our last contribution consists in an experimental evaluation of the libsuola ENGINE, carried out by benchmarking the provided operations across different versions of OpenSSL on different target architectures, by analyzing how the proposed methodology effectively addresses the concerns listed in Section 2 and its limits.

5.1 Benchmarks

For the benchmarks, we targeted two platforms to address both the desktop/server scenario and a mobile/embedded system scenario: a 4-core/4-threads 3.2GHz Intel Core i5-6550 Skylake CPU (Table 2) and a quad-core 1.2GHz Broadcom BCM2837 CPU on a Raspberry Pi 3 on a 32-bit and a 64-bit environment (Table 3, Table 4).

We collected the measurements using a benchmarking application derived from the OpenSSL speed app, modified to use the EVP API for every operation and to measure CPU cycles for a fixed number of repeated runs of the specified operations (in contrast with measuring the number of repeated runs over a fixed amount of time). We linked the benchmarking application, and libsuola in its three different versions (i.e. one for each provider selected at compile time), against OpenSSL versions 1.0.2o (old stable), 1.1.0h (current stable) and 1.1.1-pre3 (latest beta snapshot for the development version).

\footnote{https://github.com/floodyberry/ed25519-donna}\footnote{http://bench.cr.yp.to/supercop.html}
The tables report the execution cost in CPU cycles for the following cryptosystem operations: key generation, ECDH shared secret derivation, digital signature generation, and digital signature verification. As a baseline reference, the tables also present the execution cost of these operations for cryptosystems based on the fast nistz256 implementation for the popular NIST P-256 elliptic curve.

The benchmarks demonstrate that our methodology achieves the goal of adding missing functionality to OpenSSL transparently for existing applications, and to replace the default implementations with alternative ones.

Excluding X25519 primitives in OpenSSL 1.1.1-pre3, we also discovered that using libsuola-sodium, on top of adding missing functionality, generally also improves the performance of the listed primitives, as libsodium selects implementations that are more optimized for the test platforms.

We also note that, as claimed in [ZBPB17], HACL* does generally achieve relatively good performance, comparable with the default implementations in OpenSSL, and, notably, in the case of X25519 derive() in OpenSSL 1.1.0h, the HACL* implementation is even twice as fast as the default implementation.

5.2 Analysis and security evaluation

In Section 1 and Section 2 we listed a series of concerns regarding software assurance, release strategies and limits of OpenSSL as a platform for researchers to motivate our work. In this section we analyze how our proposed methodology addresses those concerns and its limits.

Software assurance. We claim that the ENGINE approach is useful in preventing various vulnerabilities that plague OpenSSL, including e.g. traditional software defects, arithmetic defects in e.g. hand-coded assembly, and various side-channel attack vectors. To support this claim, we use the following metric. We first enumerate all the CVEs issued by OpenSSL and then fetch the related security metadata for each CVE. To semi-automate this, we utilize the vFeed tool: “The Correlated Vulnerability and Threat Intelligence Database Wrapper”. We then count the number of security advisories issued across selected vendors. We partially validated the results by examining the Debian and Ubuntu patches applied to package builds.

As a case study, we did the same analysis for libsodium, for which we found no CVEs issued—consistent with no patches found in said package builds. We carried out our analysis separately for both OpenSSL 1.0.2 and 1.1.0, the only two versions which are currently not EOL. Summarizing the below results, we conclude that our approach of dedicated backend crypto providers for an ENGINE do not suffer from the same classes of security issues as the analogous functionality in libcrypto, supporting our claim.

Security statistics: OpenSSL 1.0.2. OpenSSL issued 58 CVEs; 38 for libcrypto and 20 for libssl. Restricting to the libcrypto CVEs: Debian issued 8 Debian LTS Advisories; Debian issued 12 Debian Security Advisories; Gentoo issued 8 Gentoo Linux Security Advisories; RedHat issued 22 Redhat Security Advisories; SUSE issued 71 Security Updates; Ubuntu issued 12 Ubuntu Security Notices.

Security statistics: OpenSSL 1.1.0. OpenSSL issued 17 CVEs; 9 for libcrypto and 8 for libssl. Restricting to the libcrypto CVEs: Debian issued 3 Debian LTS Advisories; Debian issued 6 Debian Security Advisories; Gentoo issued 2 Gentoo Linux Security Advisories; RedHat issued 1 Redhat Security Advisory; SUSE issued 16 Security Updates; Ubuntu issued 4 Ubuntu Security Notices.

It can also be argued that libsuola itself is adding even more surface for bugs and defects, that might result in additional risks. While this is definitely true for any additional

14 https://github.com/toolswatch/vFeed
Table 2: Benchmark results on a 4-cores/4-threads Intel Core i5-6500 CPU (Skylake) running at 3.2GHz, with Enhanced Intel SpeedStep Technology and Intel Turbo Boost Technology disabled. The workstation is equipped with 16GB 2133MHz DRAM, and runs Ubuntu 16.04 64-bit with Linux x86_64 kernel 4.13.0-36-generic. For each operation we collected 12800 samples and computed the median value.

<table>
<thead>
<tr>
<th>Operation</th>
<th>CPU cycles per operation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OpenSSL-1.0.2o</td>
</tr>
<tr>
<td></td>
<td>default</td>
</tr>
<tr>
<td>nistz256 keygen</td>
<td>39678</td>
</tr>
<tr>
<td>X25519 keygen</td>
<td>127412 (4.4x)</td>
</tr>
<tr>
<td>Ed25519 keygen</td>
<td>81320 (4.0x)</td>
</tr>
<tr>
<td>nistz256 derive</td>
<td>172484</td>
</tr>
<tr>
<td>X25519 derive</td>
<td>134856 (4.0x)</td>
</tr>
<tr>
<td>nistz256 sign</td>
<td>121073</td>
</tr>
<tr>
<td>Ed25519 sign</td>
<td>83052 (5.0x)</td>
</tr>
<tr>
<td>nistz256 verify</td>
<td>254550</td>
</tr>
<tr>
<td>Ed25519 verify</td>
<td>209156 (4.0x)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Operation</th>
<th>CPU cycles per operation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OpenSSL-1.1.0h</td>
</tr>
<tr>
<td></td>
<td>default</td>
</tr>
<tr>
<td>nistz256 keygen</td>
<td>45904</td>
</tr>
<tr>
<td>X25519 keygen</td>
<td>127444 (4.4x)</td>
</tr>
<tr>
<td>Ed25519 keygen</td>
<td>81342 (4.0x)</td>
</tr>
<tr>
<td>nistz256 derive</td>
<td>172617</td>
</tr>
<tr>
<td>X25519 derive</td>
<td>134586 (2.6x)</td>
</tr>
<tr>
<td>nistz256 sign</td>
<td>127888</td>
</tr>
<tr>
<td>Ed25519 sign</td>
<td>83080 (4.0x)</td>
</tr>
<tr>
<td>nistz256 verify</td>
<td>256706</td>
</tr>
<tr>
<td>Ed25519 verify</td>
<td>207800 (4.0x)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Operation</th>
<th>CPU cycles per operation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OpenSSL-1.1.1-pre3</td>
</tr>
<tr>
<td></td>
<td>default</td>
</tr>
<tr>
<td>nistz256 keygen</td>
<td>52172</td>
</tr>
<tr>
<td>X25519 keygen</td>
<td>127455 (1.0x)</td>
</tr>
<tr>
<td>Ed25519 keygen</td>
<td>81256 (1.0x)</td>
</tr>
<tr>
<td>nistz256 derive</td>
<td>172413</td>
</tr>
<tr>
<td>X25519 derive</td>
<td>134636 (0.8x)</td>
</tr>
<tr>
<td>nistz256 sign</td>
<td>91084</td>
</tr>
<tr>
<td>Ed25519 sign</td>
<td>83236 (1.4x)</td>
</tr>
<tr>
<td>nistz256 verify</td>
<td>229935</td>
</tr>
<tr>
<td>Ed25519 verify</td>
<td>208592 (1.0x)</td>
</tr>
</tbody>
</table>
Table 3: Benchmark results on a Raspberry Pi 3, equipped with a quad-core 1.2GHz Broadcom BCM2837 64bit CPU and 1GB RAM, running Raspbian GNU/Linux 8 and 32-bit armv7l Linux kernel version 4.9.73-v7+. For each operation we collected 12800 samples and computed the median value.

<table>
<thead>
<tr>
<th>Operation</th>
<th>CPU cycles per operation</th>
<th>OpenSSL-1.0.2o</th>
<th>libsuola-sodium</th>
<th>libsuola-donna</th>
</tr>
</thead>
<tbody>
<tr>
<td>nistz256 keygen</td>
<td></td>
<td>3056253</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>X25519 keygen</td>
<td></td>
<td>—</td>
<td>646129</td>
<td>450518 (1.4x)</td>
</tr>
<tr>
<td>Ed25519 keygen</td>
<td></td>
<td>—</td>
<td>662416</td>
<td>456970 (1.4x)</td>
</tr>
<tr>
<td>nistz256 derive</td>
<td></td>
<td>3159502</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>X25519 derive</td>
<td></td>
<td>—</td>
<td>1740942</td>
<td>484376 (3.6x)</td>
</tr>
<tr>
<td>nistz256 sign</td>
<td></td>
<td>3488822</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Ed25519 sign</td>
<td></td>
<td>—</td>
<td>706955</td>
<td>446142 (1.6x)</td>
</tr>
<tr>
<td>nistz256 verify</td>
<td></td>
<td>3964863</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Ed25519 verify</td>
<td></td>
<td>—</td>
<td>1998464</td>
<td>1334290 (1.5x)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Operation</th>
<th>CPU cycles per operation</th>
<th>OpenSSL-1.1.0h</th>
<th>libsuola-sodium</th>
<th>libsuola-donna</th>
</tr>
</thead>
<tbody>
<tr>
<td>nistz256 keygen</td>
<td></td>
<td>222028</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>X25519 keygen</td>
<td></td>
<td>637946</td>
<td>646422 (1.0x)</td>
<td>452888 (1.4x)</td>
</tr>
<tr>
<td>Ed25519 keygen</td>
<td></td>
<td>—</td>
<td>662766</td>
<td>458888 (1.4x)</td>
</tr>
<tr>
<td>nistz256 derive</td>
<td></td>
<td>1148862</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>X25519 derive</td>
<td></td>
<td>1810028</td>
<td>1741014 (1.0x)</td>
<td>484337 (3.7x)</td>
</tr>
<tr>
<td>nistz256 sign</td>
<td></td>
<td>588156</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Ed25519 sign</td>
<td></td>
<td>—</td>
<td>707272</td>
<td>445408 (1.6x)</td>
</tr>
<tr>
<td>nistz256 verify</td>
<td></td>
<td>1501720</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Ed25519 verify</td>
<td></td>
<td>—</td>
<td>1971220</td>
<td>1338216 (1.5x)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Operation</th>
<th>CPU cycles per operation</th>
<th>OpenSSL-1.1.1-pre3</th>
<th>libsuola-sodium</th>
<th>libsuola-donna</th>
</tr>
</thead>
<tbody>
<tr>
<td>nistz256 keygen</td>
<td></td>
<td>224246</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>X25519 keygen</td>
<td></td>
<td>641314</td>
<td>646264 (1.0x)</td>
<td>451060 (1.4x)</td>
</tr>
<tr>
<td>Ed25519 keygen</td>
<td></td>
<td>645840</td>
<td>662971 (1.0x)</td>
<td>457103 (1.4x)</td>
</tr>
<tr>
<td>nistz256 derive</td>
<td></td>
<td>1150334</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>X25519 derive</td>
<td></td>
<td>1810896</td>
<td>1743616 (1.0x)</td>
<td>484084 (3.7x)</td>
</tr>
<tr>
<td>nistz256 sign</td>
<td></td>
<td>580100</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Ed25519 sign</td>
<td></td>
<td>652678</td>
<td>707880 (0.9x)</td>
<td>446438 (1.5x)</td>
</tr>
<tr>
<td>nistz256 verify</td>
<td></td>
<td>1505464</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Ed25519 verify</td>
<td></td>
<td>2046863</td>
<td>1991973 (1.0x)</td>
<td>1329866 (1.5x)</td>
</tr>
</tbody>
</table>
Table 4: Benchmark results on a Raspberry Pi 3, equipped with a quad-core 1.2GHz Broadcom BCM2837 64bit CPU and 1GB RAM, running Gentoo GNU/Linux and 64-bit aarch64 Linux kernel version 4.11.12-pi64+. For each operation we collected 12800 samples and computed the median value.

<table>
<thead>
<tr>
<th>Operation</th>
<th>OpenSSL-1.0.2o</th>
<th>default</th>
<th>libsuola-sodium</th>
<th>libsuola-donna</th>
<th>libsuola-hacl</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>CPU cycles per operation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nistz256 keygen</td>
<td>3836890</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>X25519 keygen</td>
<td>228510</td>
<td>183823 (1.2x)</td>
<td>526758 (0.4x)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Ed25519 keygen</td>
<td>232244</td>
<td>187966 (1.2x)</td>
<td>1008468 (0.2x)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>nistz256 derive</td>
<td>3946922</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>X25519 derive</td>
<td>585070</td>
<td>486426 (1.2x)</td>
<td>515296 (1.1x)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>nistz256 sign</td>
<td>4355668</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Ed25519 sign</td>
<td>236454</td>
<td>177357 (1.3x)</td>
<td>1987575 (0.1x)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>nistz256 verify</td>
<td>4832773</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Ed25519 verify</td>
<td>678238</td>
<td>553648 (1.2x)</td>
<td>2059746 (0.3x)</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Operation</th>
<th>OpenSSL-1.1.0h</th>
<th>default</th>
<th>libsuola-sodium</th>
<th>libsuola-donna</th>
<th>libsuola-hacl</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>CPU cycles per operation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nistz256 keygen</td>
<td>147018</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>X25519 keygen</td>
<td>257314</td>
<td>228961 (1.1x)</td>
<td>184167 (1.4x)</td>
<td>526014 (0.5x)</td>
<td>—</td>
</tr>
<tr>
<td>Ed25519 keygen</td>
<td>232954</td>
<td>188592 (1.2x)</td>
<td>1009943 (0.2x)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>nistz256 derive</td>
<td>547095</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>X25519 derive</td>
<td>542066</td>
<td>584903 (0.9x)</td>
<td>486298 (1.1x)</td>
<td>516660 (1.0x)</td>
<td>—</td>
</tr>
<tr>
<td>nistz256 sign</td>
<td>346759</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Ed25519 sign</td>
<td>236566</td>
<td>176528 (1.3x)</td>
<td>1988958 (0.1x)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>nistz256 verify</td>
<td>798062</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Ed25519 verify</td>
<td>682893</td>
<td>554762 (1.2x)</td>
<td>2058150 (0.3x)</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Operation</th>
<th>OpenSSL-1.1.1-pre3</th>
<th>default</th>
<th>libsuola-sodium</th>
<th>libsuola-donna</th>
<th>libsuola-hacl</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CPU cycles per operation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nistz256 keygen</td>
<td>151400</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>X25519 keygen</td>
<td>265149</td>
<td>228932 (1.2x)</td>
<td>189238 (1.4x)</td>
<td>531400 (0.5x)</td>
<td>—</td>
</tr>
<tr>
<td>Ed25519 keygen</td>
<td>268433</td>
<td>232831 (1.2x)</td>
<td>193442 (1.4x)</td>
<td>1013592 (0.3x)</td>
<td>—</td>
</tr>
<tr>
<td>nistz256 derive</td>
<td>546529</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>X25519 derive</td>
<td>545707</td>
<td>585558 (0.9x)</td>
<td>486117 (1.1x)</td>
<td>515283 (1.1x)</td>
<td>—</td>
</tr>
<tr>
<td>nistz256 sign</td>
<td>255810</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Ed25519 sign</td>
<td>251354</td>
<td>236868 (1.1x)</td>
<td>176992 (1.4x)</td>
<td>1988540 (0.1x)</td>
<td>—</td>
</tr>
<tr>
<td>nistz256 verify</td>
<td>728448</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Ed25519 verify</td>
<td>643690</td>
<td>684204 (0.9x)</td>
<td>558118 (1.2x)</td>
<td>2058942 (0.3x)</td>
<td>—</td>
</tr>
</tbody>
</table>
line of code in a software system, we designed \texttt{libsuola} as a “shallow” ENGINE to minimize the risk of introducing such defects: \texttt{libsuola} itself is not providing any cryptographic functionality, and we designed it to maximize readability and maintainability making it modular.

Also, compared with the process of patching OpenSSL directly to add missing or alternative primitives, modifying our proposed ENGINE template is generally an easier process: the actual cryptographic functionality can be implemented using any language and build system as long as it produces C bindings that can be mapped in a dedicated “provider” module, while the ENGINE functionality remains separate and mostly reusable, which further minimizes the risks of introducing software defects.

**Release strategies.** One of our original goals was to facilitate the integration of missing functionality in end-of-life, yet still widely deployed, versions of OpenSSL, motivated by different release strategies applied by software providers in real-world systems as described in Section 2.2.

We demonstrated how the proposed approach allows to easily add or backport functionality in OpenSSL 1.0.2, and how this is done transparently for existing applications.

Alternatives to our approach usually require to manually recompile and install more recent versions of OpenSSL and then repeat the same process for each component of the software stack of the target system that depends on OpenSSL. Sometime this is not even possible if existing applications do not support the newer release of OpenSSL, in which case the system administrator would need to design a patch (or retrieve a trusted one) for the target software. This approach is impractical in most medium- or large-scale deployments, as it is costly to maintain and definitely adds even more surface for the rise of critical software defects and security vulnerabilities.

The same can be told about planning to replace the cryptographic provider for existing applications based on OpenSSL. Excluding some exceptions where software is designed with a “cryptographic module abstraction layer” to easily swap the default cryptographic module with alternative supported ones, most applications do not usually allow this level of flexibility, and extensive patches would be required to modify, for example, an application using OpenSSL to serve TLS connections to use a different library for the TLS stack supporting \texttt{libsodium} as the cryptographic primitive provider.

**Accessibility for researchers.** Another goal of our research listed in Section 1 aimed at delivering a framework to enable researchers to do testing and benchmarks in real-world scenarios, and to easily disseminate novel implementations and primitives to a larger user audience.

Our proposed methodology achieve this by lowering the development and maintenance costs of adding functionality to the widely used OpenSSL library.

Freedom of choice in programming languages and building tools for the actual cryptographic implementation makes it easier to plug functionality into OpenSSL. It also lowers the long term maintenance costs, especially considering the usually long and possibly unfruitful process of proposing the addition of novel or alternative functionality into the main OpenSSL codebase. Moreover, the higher degree of freedom about licensing allows to further reduce the entry barrier for adding functionality to applications based on OpenSSL.

Finally, it provides an interesting framework to collect real data about researchers’ novel contributions, which are readily comparable with existing functionality by reusing existing benchmarking facilities. Moreover, being an optional component and easy to plug in at run-time, it makes it easier for researchers to reach a wider user audience for more extensive testing or for releasing their projects.

### 5.3 Limits

Everything has its limits, and our project does not escape this truth.
The scope is limited to OpenSSL. First and foremost, our proposed methodology applies to OpenSSL only. While it can be argued that this is sufficient considering the portion of the whole Internet that currently depends on OpenSSL as part of its security stack, we recognize that this is a strong limit of our proposed methodology. In Section 6 we discuss potential research directions to overcome this limit in the future.

Overhead of linking to other libraries. Another limit of our approach is the memory consumption: when using libsuola to integrate missing functionality in OpenSSL, in addition to the memory required by OpenSSL itself, libsuola and the library it depends on (libsodium or HACL*) are also loaded into memory.

One can argue that this cost should be justified, e.g. in comparison with loading into memory only the library actually providing the cryptographic implementation (in our case either libsodium or HACL*). It is true that if the application was rewritten to only use the primitives in libsodium—i.e. directly linking against libsodium—the memory cost would be more than halved compared to our approach. However: (1) this would require redesigning every single application that currently depends on OpenSSL, defying our goal of providing the additional functionality transparently to existing applications; (2) part of the design of libsodium (and NaCl) is based on providing only an opinionated set of primitives, therefore applications patched to rely exclusively on libsodium for their cryptographic needs would lose generality and compatibility with other applications, which is often not an option; (3) libsodium only provides cryptographic primitives, so redesigning the applications as has been proposed would also incur the additional cost (and security risks) of selecting and adapting to yet another library to implement the protocol functionality (e.g. TLS) on top of libsodium.

For these reasons we do not believe that such comparison is fair, because what forks of NaCl gain in elegance, conciseness, ease-of-use, performance and security comes at the cost of not being interoperable with projects that do not implement the same set of primitives, and with both the benefits and the drawbacks of being only a cryptographic library and not a full stack library.

libsuola memory overhead. Regarding the overhead of libsuola itself, with respect to memory consumption, its footprint is negligible compared with the memory requirements of OpenSSL across all the tested architectures and OpenSSL versions.

To give some figures, in the case of our x86-64 target platform and OpenSSL 1.1.1-pre3, the resident set size in memory of libsuola.so alone totals 60 KB, compared with over 2 MB of memory occupied by OpenSSL alone (excluding the additional 1.2 MB occupied by libc, libpthread, and libdl which are required by OpenSSL).

libsuola computational overhead. With respect to computation, excluding from our analysis md_identity which is not of particular relevance, libsuola itself adds some computational overhead once when it is dynamically loaded by OpenSSL during initialization, and a small computational cost every time the functions wrapping the actual cryptographic primitives in the provider module are called.

OpenSSL uses exactly the same level of indirection whether the function implementing a primitive resides in OpenSSL itself or in an external ENGINE. The additional cost comes from the fact that the implementing function is resolved with the memory address of the function in the provider object inside libsuola, which then calls the actual function in libsodium, HACL* or the internal object with the donna implementation.

With the libsuola-donna case the wrapper is actually required as the donna implementation does not implement the NaCl C API, but in the case of libsuola-sodium and libsuola-hacl, where the wrapper only calls the relevant function in the target library, this overhead could be erased by annotating the source code of the providers with compiler attributes to instead resolve the final function pointer at load time, so that OpenSSL would call directly the relevant function in the target library.
This further optimization would come at the cost of higher complexity, inferior code readability and limited portability, which we deemed to be not in line with the stated goals of our project. We also believe that researchers who strongly need to squeeze out every single superfluous assembly instruction, while testing their code against OpenSSL through our proposed methodology, will have no problem implementing the described expedient.

6 Related work

We focused our analysis on the OpenSSL project and on ENGINEs as the native mechanism to handle alternative cryptographic implementations for the supported primitives. In this section, we present an overview of how other security standards, libraries, and frameworks address extensibility.

Describing programming interfaces and application protocols to provide interoperability and transparency between applications and cryptographic providers in the form of hardware and software cryptographic modules is a decades-old problem. Standardization bodies, operating systems, security software and hardware vendors, manufacturers and various organizations over the years approach the issue aiming for akin goals but with different approaches (e.g. regarding the cryptographic awareness required to adopt the proposed solution), trust models and features.

Standards. One of the more relevant and widespread cryptographic standards is PKCS #11 [TC16]: it specifies the “Cryptoki” API for devices that hold cryptographic information and perform cryptographic functions, presenting applications an abstract “cryptographic token” view, thus providing a separation layer between applications and specific algorithms and implementations. The PKCS #11 API defines abstract object types to represent symmetric and asymmetric crypto keys, digital signature keys, X.509 certificates, hash functions, MACs, and other common cryptographic objects, and all the functions needed to generate, modify, use and delete such objects. Just like OpenSSL ENGINEs, PKCS #11 was originally designed for hardware devices, but the level of abstraction provided allows it to be used also for software cryptographic modules. By design, PKCS #11 defines use cases with many applications and many tokens, allowing interaction and the possibility to choose among alternative implementations.

Other related standards that aim to separate applications from cryptographic implementation details include GSS-API [Lin97, Lin00] and IDUP-GSS-API [Ada98], CDSA [Gro00], SSAPI [Tea96], and the Simple Cryptographic Program Interface [Smy99].

Operating System assisted crypto. Operating Systems are by design natural candidates to provide security services and abstraction of hardware capabilities to users and applications of a system. Modern OS designs provide a hardware abstraction layer (HAL), separation of security contexts and levels, access control, authentication of loadable modules and binaries, have privileged access to cryptographic units and coprocessors, and usually already require a set of cryptographic primitives for their internal functions.

The OpenBSD/FreeBSD Cryptographic Framework (OCF) provides convenient access to the kernel cryptographic functionalitites (including symmetric and public-key crypto, hashes, MACs, access to crypto hardware accelerators and RNGs) to userspace through a /dev/crypto pseudo-device [KWdR03, KwR06] and a standard ioctl interface, simplifying the development of applications and reducing the required cryptographic awareness of application developers.

Recent versions of the Linux kernel include the AF_ALG interface (providing access to the kernel symmetric crypto functions through the socket interface), while independent developers maintain a set of patches15 to provide a /dev/crypto ioctl interface compatible with the OCF one. See [DS13] for a performance analysis of this framework in

15http://cryptodev-linux.org
differently usage scenarios and [MTP12] for a proposal to enhance this framework to provide
decoupling of cryptographic keys from applications for increased forward secrecy properties.
A different project provides a complete port of OCF to the Linux kernel[16], including the
\texttt{ioctl} interface but using the ported implementations instead of the native Linux crypto
capabilities.

Microsoft Windows OSs expose a Cryptographic Application Programming Inter-
face (CAPI, a.k.a. CryptoAPI), providing an abstraction layer and a set of dynamically
linked libraries decoupling applications from the functionality provided by Cryptographic
Service Providers (CSP). The supported functionality includes RNGs, symmetric and
public-key crypto, hashes and MACs, authentication, PKI and key storage. The “Cryp-
tographic API: Next Generation” (CNG)[17] is the latest long-term replacement of the
original API, providing a compatibility layer and featuring better support for configuration,
cryptographic agility, and a wider selection of supported primitives, covering the whole
NSA Suite B [U.S09] (including ECC).

\textbf{Other security libraries and frameworks}. Mozilla NSS[18] is another widespread set
of libraries supporting cross-platform development of security-enabled client and server
applications. It has native PKCS \#11 support for hardware and software security modules,
and indeed the internal cryptosystem implementations are encapsulated in a software
PKCS \#11 token, allowing selection among alternative implementations of crypto primitives
during configuration.

GnuTLS[19] is a high-level library providing TLS support, and relies on other libraries for
crypto primitives. It features a multilayered architecture, based on a Cryptography Provider
Layer, which provides abstraction from the actual providers for individual primitives and
allows transparent access to implementations provided by the underlying cryptographic
software library (i.e. \texttt{libgcrypt} or \texttt{nettle}), to OS assisted crypto (e.g. \texttt{/dev/crypto}
or CAPI), non-privileged CPU crypto instructions (e.g. Intel AES-NI), and TPM or hardware
and software cryptographic modules through a rich native support for PKCS \#11. The
architecture also provides functions to override symmetric crypto implementations and
“abstract private keys” as a way to handle abstractions over keys stored in hardware
modules (PKCS \#11 or TPM) or over operations implemented directly using an external
API.

ARM mbed TLS[20] (formerly known as PolarSSL) has optional support for Hardware
Security Modules (HSM) with PKCS \#11, and supports the replacement of specific
functions or full cryptosystem modules with alternative implementations, but limited to
compile time,[21]

\texttt{cryptlib} [Gut17, Gut04] has built-in native support for selected hardware crypto
accelerators and for PKCS \#11 devices, ensuring that the API for any crypto device is identical to the API of cryptlib native crypto implementations, allowing easy and
transparent migration of applications from the native software implementations to the use
of crypto devices.

The Java SDK includes the Java Secure Socket Extension (JSSE),[22] and the Java
Cryptography Architecture (JCA).[23] These modular frameworks provide an abstraction
layer respectively for secure network protocols like TLS and for cryptographic primitive
implementations. Through the factory method design pattern and the Service Provider
Interface (SPI) programming interface, JCA supports the registration of Cryptographic

[19]https://www.gnutls.org/
[22]https://docs.oracle.com/javase/8/docs/technotes/guides/security/jsse/JSSERefGuide.html
Service Provider (CSP) instances. Applications using the framework can either select a specific CSP for a primitive or let the system configuration select the most suitable implementation at runtime, attaining complete decoupling between applications and crypto implementations.

Other libraries have a completely opposite design philosophy: NaCl and libsodium, for example, are intentionally designed as simplified and opinionated libraries to avoid “cryptographic disasters” due to insufficient cryptographic awareness of application developers. As such, cryptographic agility and configurability are considered “antifeatures” due to the inherent complexity costs and the risk of enabling adopters to select insecure combinations of primitives.

From this brief survey, PKCS #11, due to its widespread support, appears to be a suitable candidate to provide alternative crypto primitive implementations portable across different libraries. Unfortunately, OpenSSL does not natively support PKCS #11, only through a third party ENGINE implementation. Nonetheless, this application of PKCS #11 seems to be an interesting direction for further research on this topic.

7 Conclusion

In this work, we analyzed issues affecting real-world security across currently deployed, ubiquitous cryptography software libraries. The consequences derived from these issues include: (1) limited assurance that software implementations are functionally correct; (2) low or non-existent accountability in the software engineering decision-making process; (3) costly and nonscalable upkeep of custom builds augmented with new features; (4) and additional security risks due to conflicting release strategies between cryptographic libraries and OS vendors.

In addition, we also examined some factors that hinder researchers and practitioners from achieving timely and widespread dissemination of their scientific results in real-world applications, thus impeding the potential impact of current cryptography research.

As a possible solution, we presented the adoption of OpenSSL ENGINEs as a framework to overcome these limits and as a convenient tool for researchers to disseminate, test and benchmark their contributions in real-world applications. We demonstrated the usage, viability, limits and benefits of this framework through libsuola, a project that aims at providing support for emerging cryptography standards such as X25519 and Ed25519 for currently deployed versions of OpenSSL, while being completely transparent for existing applications.

Due to the nature of this research, we expect it to be the foundation for future work aiming at bridging the gaps between research results and real-world applications. For instance, we intend to explore automated ENGINE construction, potentially leading to tooling that rigs together popular research-driven APIs (e.g. SUPERCOP) with practice-driven APIs. We hope to apply said tooling in NIST’s ongoing post-quantum cryptosystem standardization competition [CJL+16], comparing candidates w.r.t. real-world performance in TLS cipher suites. Moreover, we also expect it to serve as a means to enhance future research efforts, by providing a framework to ease the testing and benchmarking of novel scientific results in real-world systems and settings.

Acknowledgments

Supported in part by Academy of Finland grant 303814.

This article is based in part upon work from COST Action IC1403 CRYPTACUS, supported by COST (European Cooperation in Science and Technology).
References


26 Start your Engines: dynamically loadable contemporary crypto


