Triggerflow: Regression Testing by Advanced Execution Path Inspection

Iaroslav Gridin, Cesar Pereida García, Nicola Tuveri, and Billy Bob Brumley

Tampere University, Tampere, Finland

{iaroslav.gridin,cesar.pereidagarcia,nicola.tuveri,billy.brumley}@tuni.fi

Abstract. Cryptographic libraries often feature multiple implementations of primitives to meet both the security needs of handling private information and the performance requirements of modern services when the handled information is public. OpenSSL, the de-facto standard free and open source cryptographic library, includes mechanisms to differentiate the confidential data and its control flow, including runtime flags, designed for hardening against timing side-channels, but repeatedly accidentally mishandled in the past. To analyze and prevent these accidents, we introduce Triggerflow, a tool for tracking execution paths that, assisted by source annotations, dynamically analyzes the binary through the debugger. We validate this approach with case studies demonstrating how adopting our method in the development pipeline would have promptly detected such accidents. We further show-case the value of the tooling by presenting two novel discoveries facilitated by Triggerflow: one leak and one defect.

Keywords: software testing; regression testing; continuous integration; dynamic program analysis; applied cryptography; side-channel analysis; OpenSSL

1 Introduction

Attacks based on Side-Channel Analysis (SCA) are ubiquitous in microarchitectures and recent research [17, 15] suggest that they are much harder to mitigate than originally believed due to flawed system microarchitectures. Constant-time programming techniques are arguably the most effective and cheapest countermeasure against SCA. Functions implemented following this approach, execute and compute results time-independent from the secret inputs, thus avoiding information leakage.

Implementing constant-time code requires a highly specialized and ever growing skill set such as SCA techniques, operating systems, compilers, signal processing, and even hardware architecture; thus it is a difficult and error-prone task. Unfortunately, code is not always easily testable for SCA flaws due to code complexity and the difficulty of creating the tests themselves. Moreover, cryptography libraries tend to offer several versions of a single algorithm to be used in particular cases depending on the users' needs, thus amplifying the confusion and the possibility of using SCA vulnerable functions. To that end, we present Triggerflow, a tool that allows to selectively track code paths during program execution. The approach used by Triggerflow is elegant in its simplicity: it reports code paths taken by a given program according to the annotations defined by the user. This enables designing simple regression tests to track control flow skew. Moreover, the tool is extendable and can be integrated in the development pipeline, i.e. Continuous Integration (CI), to automatically test code paths in new builds. Triggerflow can be used both as a stand-alone tool to continuously test for known flaws, and as a support tool for other SCA tools when the source code is available. It easily allows examining code execution paths to pinpoint code flaws and regressions.

We motivate our work and demonstrate Triggerflow's effectiveness by adapting it to work with OpenSSL due to its rich history of known SCA attacks, its wide usage in the Internet, and its rapid and constant development stage. We start by back-testing OpenSSL's previously known and exploited code flaws, where our tool is able to easily find and corroborate the vulnerabilities. Additionally, using Triggerflow we identify new bugs and SCA vulnerabilities affecting the most recent OpenSSL 1.1.1a version.

In summary, Section 2 discusses previous problems and pitfalls in OpenSSL that lead to side-channel attacks. Section 3 describes the Triggerflow tool and Section 4 its application in a CI setting. We analyze in Section 5 the new bugs and vulnerabilities affecting OpenSSL, and in Section 6 we back-test known OpenSSL SCA vulnerabilities to validate the tool's effectiveness. Section 7 looks at related work. In Section 8 we discuss the limitations of our tool, and finally we conclude in Section 9.

2 Background

2.1 The OpenSSL BN_FLG_CONSTTIME Flag

In 2005, OpenSSL started considering SCA in their threat model, introducing code changes in OpenSSL version 0.9.7. The (then new) RSA cache-timing attack by Percival [20] allowed an attacker to recover secret exponent bits during the sliding-window exponentiation algorithm on systems supporting simultaneous multi-threading (SMT). As a countermeasure to this attack, the OpenSSL team adopted two important changes: Commit 3 introduced the constant-time exponentiation flag and BN_mod_exp_mont_consttime, a fixed-window modular exponentiation function; and Commit 4 implemented exponent padding. By combining these countermeasures, OpenSSL aimed for SCA resistant code path execution when performing secret key operations during DSA, RSA, and Diffie-Hellman (DH) key exchange, with the goal of performing exponentiation reasonably independent of the exponent weight or length.

The concept is to set the BN_FLG_EXP_CONSTTIME flag on BIGNUM variables containing secret information: e.g. private keys, secret prime values, nonces, and integer scalars. Once set, the flag drives access to the constant-time security critical modular exponentiation function supporting the flag. Due to performance

 $\mathbf{2}$

reasons, OpenSSL kept both functions: the constant-time version and the non constant-time version of the modular exponentiation operation. The library defaults to the non constant-time function since it assumes most operations are not secure critical, thus they can be done faster, but upon entry to the non constant-time function the input BN variables are checked for the flag and if the program detects the flag is set, it takes an early exit to the constant-time function, otherwise it continues the insecure code path.

As research and attacks on SCA improved, Aciiçmez et al. [1] demonstrated new SCA vulnerabilities in OpenSSL. More precisely, the authors showed that the default BN division function, and the Binary Extended Euclidean algorithm (BEEA) function—used in OpenSSL to perform modular inversion operations are highly dependent on their input values, therefore they leak enough information to perform a cache-timing attack. This discovery forced the introduction of Commit 14, implementing the BN_div_no_branch and BN_mod_inverse_no-_branch functions, offering a constant-time implementation for the respective operations. Moreover, BN_FLG_EXP_CONSTTIME was renamed to BN_FLG_CONST-TIME to reflect the fact that it offered protection not only to the modular exponentiation function, but to other functions as well.

2.2 Flag Exploitation

During the last three years, the BN_FLG_CONSTTIME flag has received a fair amount of attention due to its flawed effectiveness as an SCA countermeasure in OpenSSL. Pereida García et al. [22] showed the issues of having an insecureby-default approach in OpenSSL by exploiting a flaw during DSA signature generation due to a flag propagation issue. Performing a FLUSH+RELOAD [33] attack, the authors fully recover DSA private keys.

Following the previous work, Pereida García and Brumley [21] identified yet another flaw in OpenSSL, this time involving the BN_mod_inverse function. Failure to set the flag allowed the authors to successfully perform a cache-timing attack using FLUSH+RELOAD to recover secret keys during ECDSA P-256 signature generation in SSH and TLS protocols.

Building on top of the previous works, two research teams [29, 3] discovered independently several SCA flaws in OpenSSL. On the one hand, Aldaya et al. [3] developed and used a simple but effective methodology to find vulnerable code paths in OpenSSL. The authors tracked SCA vulnerable functions in OpenSSL using GDB by placing breakpoints on them. They executed the RSA key generation command, hitting the breakpoints and thus reveling flaws in OpenSSL's RSA key generation implementation. On the other hand, [29] analyzed the RSA key generation implementation and also discovered calls to the SCA vulnerable GCD function. In both cases, the authors noticed a combination of non constanttime functions in use, failure to set flags, and flags not propagated to BIGNUM variables caused OpenSSL to leak key bits. Moreover, both works demonstrate that it is possible to retrieve enough key bits to fully recover an RSA key after a single SCA trace using different cache techniques and threat models, i.e. page-level or FLUSH+RELOAD.

The previous works highlight a clear and serious issue surrounding the constanttime flag. The developers need to identify all the possible security critical cases in OpenSSL where the flag must be set in order to prevent SCA attacks, which has proven to be a laborious and clearly error-prone task. Even if done thoroughly and correctly, the developers must still ensure code changes do no introduce regressions surrounding the flag.

3 Tracking Execution Path with Triggerflow

OpenSSL's regression-testing framework has significantly improved over time, notably following the HeartBleed vulnerability. Nevertheless, the framework has its limitations, with real-world constraints largely imposed by portability requirements weighed against engineering effort. With respect to the BN_FLG_CONST-TIME flag, the testing framework does not provide a mechanism to track function calls or examine the call stack. This largely contributes to the root cause of the previously discussed vulnerabilities surrounding the BN_FLG_CONSTTIME flag: the testing framework cannot accommodate a reasonable regression test in these instances.

With this motivation, our work began by designing Triggerflow¹: a tool for tracking execution paths. After marking up the source code with special comments, its purpose is to detect when code hits paths of interest. We wrote Triggerflow in Ruby² and it uses GDB³ for inspecting code execution. Triggerflow is free and open source and distributed under MIT license.

We chose GDB since it provides all the required functionality: an established interface for choosing trace points and inspecting the program execution, as well as a machine-readable interface⁴. Additionally, GDB supports a wide variety of platforms, architectures, and languages.

Architecture. The high level concept of Triggerflow is as follows.

- 1. The inputs to Triggerflow are: a directory with annotated source code, instructions to build it, commands to run and debug, and optionally patches to apply before building.
- 2. Triggerflow scans the source code for special keywords, which are typically placed in comments near related lines of code, and builds a database of annotations.
- 3. Triggerflow commences the build, then runs the given commands ("triggers") under GDB, instructed to set breakpoints at all points of interest.
- 4. Upon hitting a breakpoint, Triggerflow inspects the context, makes decisions based on annotations, and possibly logs the code path that led to it.

In addition to verbose raw logging, Triggerflow provides output in Graphviz DOT format, allowing easy conversion to PDF, image, and other formats.

¹ https://gitlab.com/nisec/triggerflow

² https://www.ruby-lang.org/en/

³ https://www.gnu.org/software/gdb/

⁴ https://sourceware.org/gdb/onlinedocs/gdb/GDB 002fMI.html

Annotations. Using marked up source code allows leveraging existing tools for merging code changes to (semi)automatically update annotations to reflect codebase changes. Triggerflow can work with code that already includes annotations, or apply patches with annotations before compilation. Our tool currently supports four different annotations, described below and illustrated in Figure 1.

- 1. TRIGGERFLOW_POI is a point of interest and it is always tracked. The Triggerflow tool reports back every time the executing code steps into it.
- 2. TRIGGERFLOW_POI_IF is a conditional point of interest, thus it is conditionally tracked. The Triggerflow tool reports back every time the code annotated is stepped into and the given expression evaluates to true.
- 3. TRIGGERFLOW_IGNORE is an ignore annotation that allows to safely ignore specific code lines resulting in code execution paths that are not interesting, i.e. false positives.
- 4. TRIGGERFLOW_IGNORE_GROUP is a group ignore annotation that allows to safely ignore a specific code execution path if and only if every line marked with the same group ID is stepped into.



Fig. 1. Annotations currently supported by Triggerflow.

3.1 Annotating OpenSSL

Using the known vulnerable code paths previously discussed in Section 2.2, we created a set of annotations for OpenSSL with the intention to track potential leakage during secure critical operations in different public key cryptosystems such as DSA, ECDSA, RSA, as well as high-level CMS routines.

Following a direct approach, as Figure 2 illustrates we placed TRIGGERFLOW-_POI annotations to track the code path execution of the most prominent information-leaking functions previously exploited. We placed an annotation in the BN_mod_exp_mont function immediately after the early exit to its constant-time counterpart. In the BN_mod_inverse function, we placed a similar annotation after the early exit. We added an annotation at the top of the non constanttime BN_gcd function since it is known for being previously used during security critical operations but this function does not have an exit to a constant-time implementation, i.e., it is oblivious to the BN_FLG_CONSTTIME.

On the ECC code we annotated the ec_wNAF_mul function. This function implements wNAF scalar multiplication, a known SCA vulnerable function exploited several times in the past [10, 7, 23, 4, 2]. Similar to the previous cases, upon entry to this function, an early exit is available to a more SCA secure Montgomery ladder scalar multiplication $ec_scalar_mul_ladder$, thus we added the annotation immediately after the early exit.

The strategy to annotate BN_div varies depending on the OpenSSL branch. For branches up to and including 1.1.0, the function checks the flag on BN operands and assigns no_branch = 1 if it detects the flag. Hence we annotate with a no_branch != 1 conditional breakpoint. The master and 1.1.1 branches recently applied SCA hardening to its callee bn_div_fixed_top to make it oblivious to the flag. The corner case is when the number of words in BN operands are not equal, and inside the resulting data-dependent control flow we add an unconditional point of interest annotation.



Fig. 2. Top left: a TRIGGERFLOW_POI annotation in the wNAF scalar multiplication function after the early exit. Middle left: a TRIGGERFLOW_POI annotation during BN_div execution. Bottom left: a TRIGGERFLOW_POI annotation in OpenSSL's insecure BN_gcd function. Top right: a TRIGGERFLOW_POI annotation in OpenSSL's BN_mod_inverse function after the early exit should never be reached during secure critical operations. Bottom right: a TRIGGERFLOW_POI annotation in BN_mod_exp_mont after the early exit. Ideally, the previous annotations should never be reached, since we assume OpenSSL follows a constant-time code path during the execution of these secure critical operations. Yet one of the most security-critical parts of the process is marking false positive annotations. To give an idea of the scope of such marking, with the above described point of interest annotations applied to the OpenSSL 1.1.0 branch, and no ignore annotations, Triggerflow identifies 84 potentially errant code paths, provided with only a basic set of 25 triggers.

4 Continuous Integration

As previously discussed, our main motivation for Triggerflow is the need to test for regressions in OpenSSL surrounding the BN_FLG_CONSTTIME flag. From the software quality perspective, and given the previously exploited vulnerabilities discussed later in Section 6, there is a clear need for an automated approach that accounts for the time dimension and a rapidly changing codebase. Seemingly small and insignificant changes can suddenly shift codepaths, and when PRs are proposed and merged we want to be automatically informed. Using code marked up for Triggerflow allows establishing CI, automatically testing code for introducing unsafe codepaths. We propose (and deploy) the following approach to establish an automatic CI pipeline using Triggerflow and GitLab's infrastructure, illustrated in Figure 3.



Fig. 3. CI flow illustrated.

- 8 Iaroslav Gridin, Cesar Pereida García, Nicola Tuveri, and Billy Bob Brumley
 - Create a special Git repository containing Triggerflow configuration, trigger list, annotations in form of Quilt⁵ patch queue, and a submodule containing code to test (in our case, OpenSSL). This repository is hosted on a GitLab instance and includes the description of the testing process in Gitlab format, .gitlab.yml.
 - Two "runners" are established on separate machines, connected to the GitLab instance. A runner is automated testing software which creates a container and runs testing routines according to rules in .gitlab.yml. We maintain two runners with different architectures, x86_64 and aarch64. The runners are based in our infrastructure. When new code is pushed into the GitLab repository and .gitlab.yml is present, runners run the tests and report status back to GitLab, where results are then reviewed.
 - A separate software, "repatcher"⁶, is continuously monitoring main OpenSSL code repository for updates and adapting annotations to changed code. If changes can be applied automatically, repatcher pushes updated code to GitLab where it is tested. Otherwise, a human is notified to resolve conflicts and update the patches manually. After that, repatcher's work automatically continues. Repatcher is based in our infrastructure.

This process is independent of any support from the original developers. Of course, a better approach is to have developers themselves integrate and maintain Triggerflow annotations upstream, or potentially enforce them at compile time.

Unfortunately, successful deployment of such a CI pipeline depends on code being buildable on every upstream commit, which is sometimes not the case with OpenSSL. Still, with minimal manual inspection it makes a great automatic testing setup: Figure 4 illustrates our CI testing OpenSSL's master branch using Triggerflow. The results of our CI system instance are public⁷, monitoring master, 1.1.1 and 1.1.0 branches of OpenSSL.



Fig. 4. GitLab CI running: Triggerflow testing OpenSSL code.

⁵ https://savannah.nongnu.org/projects/quilt

⁶ https://gitlab.com/nisec/repatcher

⁷ https://gitlab.com/nisec/openssl-triggerflow-ci

5 New Bugs and Vulnerabilities

With the tooling in place, our first task was to examine functionality issues that could arise with applying the annotation patches to a shifting codebase. The EC module recently underwent a quite heavy overhaul regarding SCA security [27]. We used that as a case study, and in this section we present two discoveries facilitated by Triggerflow: one leak and one software defect.

5.1 A New Leak

We started from Commit 1 and the Triggerflow unit test in question is ECDSA signing in ecdsa_ossl.c. The test passed at that commit, hence the tooling proceeded with subsequent commits. They all passed unit testing, until reaching Commit 2. The purpose of said commit was to fix a regression in the padding of secret scalar inputs in the timing-resistant elliptic curve scalar multiplication, using the group cardinality rather than the generator order, supporting cryptosystems where the distinction is relevant (e.g., ECDH and cofactor variants). Figure 5 illustrates the failed unit test.



Fig. 5. Insecure flow: projective to affine point conversion (abridged).

The fix. In this case, what the tooling is telling us is that the code is traversing the insecure modular inversion path when converting from projective to affine coordinates. Examining this function, it has always been oblivious to the constant-time flag, yet academic results suggest that said conversion should be protected [19, 18]. Put another way, Commit 2 is not the culprit—the function is insecure by design. Instead of simply enabling the flag, we chose⁸ to add a field_inv function pointer inside the EC_METHOD structure, alongside existing pointers for other finite field operations such as field_mul and field_sqr. This allowed us to unify the finite field inversion across the EC module, instead of each function meticulously enabling the constant-time flag when calling BN_mod__ inverse. Once unified, we can ensure default SCA hardening through a single interface. We provided three different implementations for this pointer for three different EC_METHOD instances:

1. EC_GFp_mont_method is the default for prime curves and pre-computes a Montgomery arithmetic structure for finite field arithmetic. This is convenient for inversion via FLT, which is modular exponentiation with a fixed exponent and variable base—benefiting generously from the Montgomery

⁸ https://github.com/openssl/openssl/pull/8254

arithmetic. Hence our field_inv implementation is a straightforward version of FLT in this case.

- 2. EC_GFp_simple_method is a fallback method that contains much of the boilerplate code pointed to by several other EC_METHOD implementations. For example, those that implement their own custom arithmetic, such as NIST curves that use Mersenne-like primes. Here, no Montgomery structure is guaranteed to exist. Hence our field_inv implementation is blinding, computing $a^{-1} = b/(ab)$ with b chosen uniformly at random and the ab term inverted via BN_mod_inverse.
- EC_GF2m_simple_method is the only method for binary curves present in the OpenSSL codebase. Here field_inv is a simple wrapper around BN_GF2m-_mod_inv, which is already SCA-hardened with blinding.

With these SCA-hardened field_inv function pointers in place, we then transitioned all finite field inversions in the EC module from BN_mod_inverse and BN_GF2m_mod_inv to our new pointer, including that of the projective to affine conversion. After these changes, Triggerflow unit tests were successful.

5.2 A New Defect

The previous unit test failure is curious in the sense that Commit 2 was essentially unrelated to projective to affine conversion. As stated above, that conversion has always been oblivious to the constant-time flag. We were left with the question of how such a change could trigger an insecure behavior in an unrelated function.

Using the debugger to compare the internal state when executing EC_POINT-_get_affine_coordinates_GFp in Commit 2 and its parent, we discovered that, until the latter, a temporary variable storing one of the inputs to BN_mod-_inverse was flagged as constant-time even if the flag was not explicitly set with the dedicated function. The temporary variable in question was obtained through a BN_CTX object, a buffer shared among various functions that simulates a hardware stack to store BIGNUM variables, minimizing costly memory allocations—we defer to [11] for more details on the internals of the BN_CTX object.

In this case, the BN_CTX object is created in the top level function implementing signature generation for the ECDSA cryptosystem, and is shared among most of its callees and descendants; the analysis led to discover that the BN_CTX buffer retained the state of BN_FLG_CONSTTIME for each stored BIGNUM variable, allowing functions to alter the value of BN_FLG_CONSTTIME, and thus occasionally the execution flow, of subsequently called functions sharing the same BN_CTX.

The fix. This long-standing defect raises several concerns:

- as in the case that lead to its discovery, retrieving a BIGNUM variable from the BN_CTX with BN_FLG_CONSTTIME unexpectedly set, might lead to unintentional execution of a timing-resistant code-path. This could be perceived as a benign effect, but hides unexpected risks as it generates false negatives during security analysis. Moreover, changes as trivial as getting one more temporary variable from the shared BN_CTX—or even just changing the order by which temporary variables are retrieved—can influence the execution flow of seemingly unrelated functions, eluding manual analysis and defying developers' expectations;

- a BIGNUM variable with BN_FLG_CONSTTIME unexpectedly set could reach function implementations that execute in variable time and should never be called with confidential inputs marked with BN_FLG_CONSTTIME. Such functions diligently check for API abuse and raise exceptions at run time: this defect can then result in unexpected application crashes or potentially expose to bug attacks;
- automated testing is made fragile, in part for the false negatives already mentioned, but additionally because the test suite becomes not representative of external application usage of the library, as different usage patterns of a shared BN_CTX in unrelated functions lead to different execution paths. Finally, the generated failure reports could be misleading as changes in unrelated functions might end up triggering errors in other modules.

The fix itself was relatively straightforward, and consisted in unconditionally clearing BN_FLG_CONSTTIME every time a BIGNUM variable is retrieved from a BN_CTX⁹.

What is remarkable is how Triggerflow assisted in the discovery of a defect that had been unnoticed for over a decade, automating the interaction with the debugger to pinpoint which revisions triggered the anomalous behavior.

6 Validation

In order to validate our work, we present next a study of the known flaws briefly discussed in Section 2.2 that led to several SCA attacks and security advisories. We present these flaws as case studies, briefly discussing the root cause, security implications, and the results of running our tooling against an annotated OpenSSL. We separate the cases by cryptosystem and at the same we (mostly) follow the chronological discovery of these flaws.

As part of the validation, we used the same OpenSSL versions as in the original attacks. To that end, we forked OpenSSL branches on the respective versions and then, we applied the set of annotations previously discussed in Section 3.1. This approach allowed us to quickly back test and validate the effectiveness of our tooling to detect potential leakage in OpenSSL.

The list of cases presented here is not exhaustive but serves three purposes:

- 1. it gives insight on the type of flaws that our Triggerflow is able to find;
- 2. it shows it is not a trivial task to do; and
- 3. it demonstrates the fragility of the BN_FLG_CONSTTIME countermeasure introduced 14 years ago and the need of a secure-by-default approach in cryptography libraries such as OpenSSL.

⁹ https://github.com/openssl/openssl/pull/8253

6.1 DSA

The DSA signature generation implementation in OpenSSL has arguably the longest and most troubled history of SCA issues. In 2016, a decade after BN-_FLG_CONSTTIME and the constant-time exponentiation function countermeasures were introduced, Pereida García et al. [22] discovered that the constant-time path was not taken due to a flag propagation issue. The authors noticed that BN_copy effectively copies the content from a BIGNUM variable to another but it fails to copy the existing flags, thus flags are not propagated and the constant-time flag must be set again. This issue left the DSA signature generation vulnerable to cache-timing attacks for more than a decade. To test this issue, we pointed Triggerflow at our annotated OpenSSL_1_0_2k branch, resulting in Figure 6 and therefore correctly reporting the flaw.



Fig. 6. Triggerflow detecting CVE-2016-2178, i.e. the flawed CVE-2005-0109 fix (abridged).

The authors provided a fix for this issue in Commit 5, but at the same time they introduced a new flaw in the modular inversion operation during DSA signature generation. This new vulnerability was enabled due to a missing constanttime flag in one of the input values to the BN_mod_inverse function. At that time, the flaw was confined to the development branch, subsequently promptly fixed in Commit 6, thus it did not affect users. Figure 7 shows the result of pointing Triggerflow to OpenSSL in Commit 5, detecting the flawed fix.



Fig. 7. Triggerflow detecting the flawed CVE-2016-2178 fix (abridged).

Later in 2018, Weiser et al. [30] found additional SCA vulnerabilities in DSA. The authors exploited a timing variation due to the BIGNUM structure to recover DSA private keys, an unrelated issue to the BN_FLG_CONSTTIME flag. However, the fix provided for this issue in Commit 8 was incomplete, and moreover it introduced a new SCA flaw, once again due to not setting a flag properly. Triggerflow detected this flaw (see Figure 8) in the OpenSSL_1_1_1 branch, later fixed in Commit 9 but again only present briefly in development branches.

In the same work, the authors discovered that every time the library loads a DSA private key, it calculates the corresponding public key following a non



Fig. 8. Triggerflow detecting the flawed CVE-2018-0734 fix (abridged).

constant-time code path due to a missing flag, and therefore is also vulnerable to SCA attacks. In fact, Triggerflow previously detected this vulnerability while back-testing Commit 5, suggesting that this issue was long present in the codebase and could have been detected earlier. This issue was recently fixed in Commit 7.

6.2 ECDSA

OpenSSL's ECDSA implementation has also been affected by SCA leakage. In their work, Pereida García and Brumley [21] discovered that the BN_FLG-_CONSTTIME flag was not set at all during ECDSA P-256 signature generation. More specifically, the modular inversion operation was performed using the non constant-time path in the BN_mod_inverse function, thus leaving the scalar k vulnerable to SCA attacks.



Fig. 9. Triggerflow detecting CVE-2016-7056 (abridged).

Similar to the previous case and in order to back-test this issue, we pointed Triggerflow to the annotated $OpenSSL_1_0_1u$ branch and then we generated ECDSA signatures, triggering the breakpoints. The tool reported back an insecure usage of the modular inversion function as shown in Figure 9. The flag was not set in the nonce k prior to the modular inversion operation. Surprisingly, this issue is still present in the OpenSSL 1.0.1 branch although the authors provided a patch for it, mainlined by the vast majority of vendors. It is worth mentioning that that OpenSSL 1.0.1 branch reached EOL around the same time as the work and we assume that is the reason the OpenSSL team did not integrate it.

6.3 RSA

In 2018, two independent works [29, 3] discovered several SCA flaws during RSA key generation in OpenSSL. OpenSSL's RSA key generation is a fairly complex implementation due to the use of several different algorithms during the process. It requires the generation of random integers; testing the values for primality; computing the greatest common divisor and the least common multiple, using

secret values as input. For all of the previous reasons, it is not trivial to implement a constant-time RSA key generation algorithm. Both research works identified missing flags, flags set in the wrong variable, and a direct call to the non constanttime function BN_gcd as the culprits enabling the attacks.



Fig. 10. Triggerflow detecting CVE-2018-0737 (abridged).

During back testing we used an annotated OpenSSL_1_0_2k branch, and we pointed the Triggerflow tool at it. It successfully reported all the vulnerabilities discovered by the authors. The authors submitted a total of four commits to OpenSSL codebase to fully mitigate this issue—see Commit 10, Commit 11, Commit 12, and Commit 13 for more details.

7 Related Work

The Triggerflow framework differs from other existing tools in being a tool to assist the development process rather than a system for automated detection and quantification of security vulnerabilities, and aims at being more general purpose and not restricted to the field of cryptographic applications. As such, it should be viewed as complementary rather than alternative to the approaches listed below.

Programming languages. Various works propose and analyze the option of using specialized programming languages to achieve constant-time code generation and verification [9, 12], while others analyze the challenges [6] or opportunities [26] of translating human-readable code into machine instructions through compilers when dealing with cryptographic software and the need for SCA resistant implementations. They differ from this work in the goal: our evaluation is not based on a lack of timing-resistant implementations, but rather in assisting the development process and making sure that insecure paths are not executed, by mistake, with confidential inputs.

Black box testing. These practices are based on statistical analysis to estimate the SCA leakage. *dudect* [24] applies this methodology measuring the timing of the system under test for different inputs.

Static program analysis. These techniques refers to the analysis of the source code [5, 32, 25] (building on the capabilities of the LLVM project to perform the analysis) or annotated machine code [8] of a program to quantify leakages. An alternative to this approach is represented by *CacheAudit* [14, 13] based on symbolic execution, which is usually applied to smaller software or individual algorithms as it requires more resources. All of these methods share with Triggerflow the requirement of access to the source code of the tested software.

Dynamic program analysis. These techniques detect, measure, and accurately locate microarchitecture leakage during the execution of the code in the system. ctgrind [16], based on Valgrind memcheck, monitors control flow and memory accesses for dependencies on secret data. [31, 30] use Dynamic Binary Instrumentation, i.e. adding instrumentation at run-time to collect metadata and measurements directly to the binary code without altering the execution flow of the program, independently providing extensible frameworks with high accuracy and supporting leakage models for the most relevant microarchitecture attacks. CacheD [28] is an hybrid approach that combines DBI, symbolic execution and constraint solving to identify cache-based timing channels. Triggerflow is similar to Dynamic Program Analysis techniques with respect to performing the evaluation when the software is actively running on the target system. Although limited by requiring access to the source code, Triggerflow can leverage this property and avoid any instrumentation: the tested binary is exactly the one generated by the build process of the target, with the only requirement of including the debug symbols to aid GDB in mapping function names and the memory addresses of the routines included in the target software.

8 Limitations

Triggerflow requires access to the sources of the target software, and to annotate it with markup comments as described in Section 3. The latter can be a significant drawback as long as the annotations cannot be versioned together with the target software as the current version of Triggerflow has limited support for versioning the annotations externally; this is one of the future developments planned for the tool.

Triggerflow depends on the availability of GDB and ruby on the target platform, and is limited to the executables that can be debugged through GDB. This is arguably a minor concern, with the only remarkable exception that debugging through GDB inside a virtualized container usually requires overriding the default set of system call restrictions that is meant to isolate the supervisor from the container, raising security concerns when running Triggerflow for CI and partially limiting the selection of available CI platforms.

The tools developed during this work can also be applied to other software projects, not just OpenSSL. Triggerflow can work with any language GDB supports and would be useful for studying execution paths through any complex project that meets the minimal requirements.

9 Conclusion

Triggerflow complements the results offered by any of the analysis techniques described in Section 7: in large software projects like OpenSSL, pinpointing the location of a detected leak might not be sufficient. Similarly to other cryptographic libraries, OpenSSL often includes several implementations of the same primitive, many of which are designed for performance and safe to use only when all the inputs are public. When a leak is detected in one of these functions, developers are challenged with the task of discovering why and how secret data reached the insecure code path, rather than altering the location where the leakage is reported. As demonstrated in Sections 5 and 6, Triggerflow can be successfully and efficiently used to aid the developer in these situations and, through CI, prevent regressions in the handling of secret data.

Considering the high number of valid combinations of supported platforms and build time options for OpenSSL, and that the available implementations and control flow depend on these specific combinations, Triggerflow is a good solution to aid the developers by exhaustively automating the BN_FLG_CONST-TIME tests and prevent future regressions similar to the ones described in this work.

In the context of using Triggerflow with OpenSSL to monitor BN_FLG_CONST-TIME, it should be mentioned that, security-wise, a secure-by-default approach would be desirable: i.e., all BIGNUM are considered *constant-time* unless the programmer explicitly marks them as public, so that, where alternatives exist, the default implementation of each algorithm is the timing-resistant one, and insecure but more efficient ones need to be enabled explicitly and after careful examination. On the other hand, such change has the potential for being disruptive for existing applications, and is therefore likely to be rejected or implemented over a long period of time to meet the project release strategy.

Future work. On top of continued development of the tool as discussed, we plan to expand on this work in the future to widen the coverage of the OpenSSL library and of the project *apps* and their options, by setting more triggers and point of interest, across multiple architectures and build-time options. In parallel, to further demonstrate the capabilities of the tool, we plan to apply a similar methodology to other security libraries and cryptographic software, aiming at uncovering, fixing, and testing related timing leaks.

Acknowledgments

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 804476).

References

1. Aciçmez, O., Gueron, S., Seifert, J.: New branch prediction vulnerabilities in OpenSSL and necessary software countermeasures. In: Galbraith, S.D. (ed.) Cryptography and Coding, 11th IMA International Conference, Cirencester, UK, December 18-20, 2007, Proceedings. Lecture Notes in Computer Science, vol. 4887, pp. 185–203. Springer (2007), https://doi.org/10.1007/978-3-540-77272-9 12

- Aldaya, A.C., Brumley, B.B., ul Hassan, S., Pereida García, C., Tuveri, N.: Port contention for fun and profit. In: 2019 IEEE Symposium on Security and Privacy, SP 2019, Proceedings, 20-22 May 2019, San Francisco, California, USA. pp. 1037– 1054. IEEE (2019), https://doi.org/10.1109/SP.2019.00002
- Aldaya, A.C., Pereida García, C., Alvarez Tapia, L.M., Brumley, B.B.: Cachetiming attacks on RSA key generation. IACR Cryptology ePrint Archive 2018(367) (2018), https://eprint.iacr.org/2018/367
- Allan, T., Brumley, B.B., Falkner, K.E., van de Pol, J., Yarom, Y.: Amplifying side channels through performance degradation. In: Schwab, S., Robertson, W.K., Balzarotti, D. (eds.) Proceedings of the 32nd Annual Conference on Computer Security Applications, ACSAC 2016, Los Angeles, CA, USA, December 5-9, 2016. pp. 422–435. ACM (2016), http://doi.acm.org/10.1145/2991079.2991084
- Almeida, J.B., Barbosa, M., Barthe, G., Dupressoir, F., Emmi, M.: Verifying constant-time implementations. In: Holz, T., Savage, S. (eds.) 25th USENIX Security Symposium, USENIX Security 16, Austin, TX, USA, August 10-12, 2016. pp. 53–70. USENIX Association (2016), https://www.usenix.org/conference/ usenixsecurity16/technical-sessions/presentation/almeida
- Balakrishnan, G., Reps, T.W.: WYSINWYX: what you see is not what you execute. ACM Trans. Program. Lang. Syst. 32(6), 23:1–23:84 (2010), https://doi.org/10. 1145/1749608.1749612
- Benger, N., van de Pol, J., Smart, N.P., Yarom, Y.: "Ooh Aah... Just a Little Bit": A small amount of side channel can go a long way. In: Batina, L., Robshaw, M. (eds.) Cryptographic Hardware and Embedded Systems - CHES 2014 - 16th International Workshop, Busan, South Korea, September 23-26, 2014. Proceedings. Lecture Notes in Computer Science, vol. 8731, pp. 75–92. Springer (2014), https: //doi.org/10.1007/978-3-662-44709-3_5
- Blazy, S., Pichardie, D., Trieu, A.: Verifying constant-time implementations by abstract interpretation. In: Foley, S.N., Gollmann, D., Snekkenes, E. (eds.) Computer Security - ESORICS 2017 - 22nd European Symposium on Research in Computer Security, Oslo, Norway, September 11-15, 2017, Proceedings, Part I. Lecture Notes in Computer Science, vol. 10492, pp. 260–277. Springer (2017), https://doi.org/10.1007/978-3-319-66402-6_16
- Bond, B., Hawblitzel, C., Kapritsos, M., Leino, K.R.M., Lorch, J.R., Parno, B., Rane, A., Setty, S.T.V., Thompson, L.: Vale: Verifying high-performance cryptographic assembly code. In: Kirda, E., Ristenpart, T. (eds.) 26th USENIX Security Symposium, USENIX Security 2017, Vancouver, BC, Canada, August 16-18, 2017. pp. 917–934. USENIX Association (2017), https://www.usenix.org/conference/ usenixsecurity17/technical-sessions/presentation/bond
- Brumley, B.B., Hakala, R.M.: Cache-timing template attacks. In: Matsui, M. (ed.) Advances in Cryptology - ASIACRYPT 2009, 15th International Conference on the Theory and Application of Cryptology and Information Security, Tokyo, Japan, December 6-10, 2009. Proceedings. Lecture Notes in Computer Science, vol. 5912, pp. 667–684. Springer (2009), https://doi.org/10.1007/978-3-642-10366-7_39
- Brumley, B.B., Tuveri, N.: Cache-timing attacks and shared contexts. In: Constructive Side-Channel Analysis and Secure Design - 2nd International Workshop, COSADE 2011, Darmstadt, Germany, February 24-25, 2011. Proceedings. pp. 233– 242 (2011), https://tutcris.tut.fi/portal/files/15671512/cosade2011.pdf

- 18 Iaroslav Gridin, Cesar Pereida García, Nicola Tuveri, and Billy Bob Brumley
- Cauligi, S., Soeller, G., Brown, F., Johannesmeyer, B., Huang, Y., Jhala, R., Stefan, D.: Fact: A flexible, constant-time programming language. In: IEEE Cybersecurity Development, SecDev 2017, Cambridge, MA, USA, September 24-26, 2017. pp. 69– 76. IEEE Computer Society (2017), https://doi.org/10.1109/SecDev.2017.24
- Doychev, G., Köpf, B.: Rigorous analysis of software countermeasures against cache attacks. In: Cohen, A., Vechev, M.T. (eds.) Proceedings of the 38th ACM SIGPLAN Conference on Programming Language Design and Implementation, PLDI 2017, Barcelona, Spain, June 18-23, 2017. pp. 406–421. ACM (2017), https://doi.org/10.1145/3062341.3062388
- Doychev, G., Köpf, B., Mauborgne, L., Reineke, J.: Cacheaudit: A tool for the static analysis of cache side channels. ACM Trans. Inf. Syst. Secur. 18(1), 4:1–4:32 (2015), https://doi.org/10.1145/2756550
- Kocher, P., Genkin, D., Gruss, D., Haas, W., Hamburg, M., Lipp, M., Mangard, S., Prescher, T., Schwarz, M., Yarom, Y.: Spectre attacks: Exploiting speculative execution. CoRR abs/1801.01203 (2018), http://arxiv.org/abs/1801.01203
- 16. Langley, A.: ctgrind—checking that functions are constant time with Valgrind. https://github.com/agl/ctgrind (2010)
- Lipp, M., Schwarz, M., Gruss, D., Prescher, T., Haas, W., Fogh, A., Horn, J., Mangard, S., Kocher, P., Genkin, D., Yarom, Y., Hamburg, M.: Meltdown: Reading kernel memory from user space. In: Enck, W., Felt, A.P. (eds.) 27th USENIX Security Symposium, USENIX Security 2018, Baltimore, MD, USA, August 15-17, 2018. pp. 973–990. USENIX Association (2018), https://www.usenix.org/conference/ usenixsecurity18/presentation/lipp
- Maimut, D., Murdica, C., Naccache, D., Tibouchi, M.: Fault attacks on projectiveto-affine coordinates conversion. In: Prouff, E. (ed.) Constructive Side-Channel Analysis and Secure Design - 4th International Workshop, COSADE 2013, Paris, France, March 6-8, 2013, Revised Selected Papers. Lecture Notes in Computer Science, vol. 7864, pp. 46–61. Springer (2013), https://doi.org/10.1007/ 978-3-642-40026-1 4
- Naccache, D., Smart, N.P., Stern, J.: Projective coordinates leak. In: Cachin, C., Camenisch, J. (eds.) Advances in Cryptology - EUROCRYPT 2004, International Conference on the Theory and Applications of Cryptographic Techniques, Interlaken, Switzerland, May 2-6, 2004, Proceedings. Lecture Notes in Computer Science, vol. 3027, pp. 257–267. Springer (2004), https://doi.org/10. 1007/978-3-540-24676-3 16
- Percival, C.: Cache missing for fun and profit. In: BSDCan 2005, Ottawa, Canada, May 13-14, 2005, Proceedings (2005), http://www.daemonology.net/ papers/cachemissing.pdf
- Pereida García, C., Brumley, B.B.: Constant-time callees with variable-time callers. In: Kirda, E., Ristenpart, T. (eds.) 26th USENIX Security Symposium, USENIX Security 2017, Vancouver, BC, Canada, August 16-18, 2017. pp. 83–98. USENIX Association (2017), https://www.usenix.org/conference/usenixsecurity17/technical-sessions/presentation/garcia
- Pereida García, C., Brumley, B.B., Yarom, Y.: "Make sure DSA signing exponentiations really are constant-time". In: Weippl, E.R., Katzenbeisser, S., Kruegel, C., Myers, A.C., Halevi, S. (eds.) Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, Vienna, Austria, October 24-28, 2016. pp. 1639–1650. ACM (2016), http://doi.acm.org/10.1145/2976749.2978420
- 23. van de Pol, J., Smart, N.P., Yarom, Y.: Just a little bit more. In: Nyberg, K. (ed.) Topics in Cryptology CT-RSA 2015, The Cryptographer's Track at the

RSA Conference 2015, San Francisco, CA, USA, April 20-24, 2015. Proceedings. Lecture Notes in Computer Science, vol. 9048, pp. 3–21. Springer (2015), https://doi.org/10.1007/978-3-319-16715-2_1

- Reparaz, O., Balasch, J., Verbauwhede, I.: Dude, is my code constant time? In: Atienza, D., Natale, G.D. (eds.) Design, Automation & Test in Europe Conference & Exhibition, DATE 2017, Lausanne, Switzerland, March 27-31, 2017. pp. 1697– 1702. IEEE (2017), https://doi.org/10.23919/DATE.2017.7927267
- Rodrigues, B., Pereira, F.M.Q., Aranha, D.F.: Sparse representation of implicit flows with applications to side-channel detection. In: Zaks, A., Hermenegildo, M.V. (eds.) Proceedings of the 25th International Conference on Compiler Construction, CC 2016, Barcelona, Spain, March 12-18, 2016. pp. 110–120. ACM (2016), http: //doi.acm.org/10.1145/2892208.2892230
- Simon, L., Chisnall, D., Anderson, R.J.: What you get is what you C: controlling side effects in mainstream C compilers. In: 2018 IEEE European Symposium on Security and Privacy, EuroS&P 2018, London, United Kingdom, April 24-26, 2018. pp. 1–15. IEEE (2018), https://doi.org/10.1109/EuroSP.2018.00009
- Tuveri, N., ul Hassan, S., Pereida García, C., Brumley, B.B.: Side-channel analysis of SM2: A late-stage featurization case study. In: Proceedings of the 34th Annual Computer Security Applications Conference, ACSAC 2018, San Juan, PR, USA, December 03-07, 2018. pp. 147–160. ACM (2018), https://doi.org/10.1145/3274694.3274694.3274725
- Wang, S., Wang, P., Liu, X., Zhang, D., Wu, D.: Cached: Identifying cachebased timing channels in production software. In: Kirda, E., Ristenpart, T. (eds.) 26th USENIX Security Symposium, USENIX Security 2017, Vancouver, BC, Canada, August 16-18, 2017. pp. 235–252. USENIX Association (2017), https://www.usenix.org/conference/usenixsecurity17/technical-sessions/ presentation/wang-shuai
- Weiser, S., Spreitzer, R., Bodner, L.: Single trace attack against RSA key generation in Intel SGX SSL. In: Kim, J., Ahn, G., Kim, S., Kim, Y., López, J., Kim, T. (eds.) Proceedings of the 2018 on Asia Conference on Computer and Communications Security, AsiaCCS 2018, Incheon, Republic of Korea, June 04-08, 2018. pp. 575–586. ACM (2018), http://doi.acm.org/10.1145/3196494.3196524
- 30. Weiser, S., Zankl, A., Spreitzer, R., Miller, K., Mangard, S., Sigl, G.: DATA differential address trace analysis: Finding address-based side-channels in binaries. In: Enck, W., Felt, A.P. (eds.) 27th USENIX Security Symposium, USENIX Security 2018, Baltimore, MD, USA, August 15-17, 2018. pp. 603–620. USENIX Association (2018), https://www.usenix.org/conference/usenixsecurity18/presentation/weiser
- Wichelmann, J., Moghimi, A., Eisenbarth, T., Sunar, B.: MicroWalk: A framework for finding side channels in binaries. In: Proceedings of the 34th Annual Computer Security Applications Conference, ACSAC 2018, San Juan, PR, USA, December 03-07, 2018. pp. 161–173. ACM (2018), https://doi.org/10.1145/3274694.3274741
- 32. Wu, M., Guo, S., Schaumont, P., Wang, C.: Eliminating timing side-channel leaks using program repair. In: Tip, F., Bodden, E. (eds.) Proceedings of the 27th ACM SIGSOFT International Symposium on Software Testing and Analysis, IS-STA 2018, Amsterdam, The Netherlands, July 16-21, 2018. pp. 15–26. ACM (2018), https://doi.org/10.1145/3213846.3213851
- Yarom, Y., Falkner, K.: FLUSH+RELOAD: A high resolution, low noise, L3 cache side-channel attack. In: Proceedings of the 23rd USENIX Security Symposium, San Diego, CA, USA, August 20-22, 2014. pp. 719–732. USENIX Association

 $(2014), \quad https://www.usenix.org/conference/usenixsecurity14/technical-sessions/presentation/yarom$

A OpenSSL Commits

1.	fe2d3975880e6a89702f18ec58881307bf862542
2.	a766aab93a282774e63ba918d0bb1c6680a5f292
3.	46a643763de6d8e39ecf6f76fa79b4d04885aa59
4.	0ebfcc8f92736c900bae4066040b67f6e5db8edb
5.	621eaf49a289bfac26d4cbcdb7396e796784c534
6.	b7d0f2834e139a20560d64c73e2565e93715ce2b
7.	6364475a990449ef33fc270ac00472f7210220f2
8.	a9cfb8c2aa7254a4aa6a1716909e3f8cb78049b6
9.	00496b6423605391864fbbd1693f23631a1c5239
10.	e913d11f444e0b46ec1ebbf3340813693f4d869d
11.	8db7946ee879ce483f4c81141926e1357aa6b941
12.	54f007af94b8924a46786b34665223c127c19081
13.	6939eab03a6e23d2bd2c3f5e34fe1d48e542e787
14.	bd31fb21454609b125ade1ad569ebcc2a2b9b73c