Cache Timing Leakages in Zero-Knowledge Protocols

Shibam Mukherjee[1,3*], Christian Rechberger[1,4] and Markus Schofnegger[2]

[1] Graz University of Technology (Austria)
[2] Horizen Labs (United States)
[3] Know Center (Austria)
[4] TACEO (Austria)

Abstract. The area of modern zero-knowledge proof systems has seen a significant rise in popularity over the last couple of years, with new techniques and optimized constructions emerging on a regular basis.
As the field matures, the aspect of implementation attacks becomes more relevant, however side-channel attacks on zero-knowledge proof systems have seen surprisingly little treatment so far. In this paper we give an overview of potential attack vectors and show that some of the underlying finite field libraries, and implementations of heavily used components like hash functions, are vulnerable w.r.t. cache attacks on CPUs.
On the positive side, we demonstrate that the computational overhead to protect against these attacks is relatively small.

Keywords: No keywords given.

1 Introduction

Recent years have witnessed a significant development in the area of zero-knowledge proofs (ZKPs) [GMR85] and its applications in blockchains [Polb, Labc, Eth, Fou, Sta], decentralized apps [Min, Dus, Sem], anonymous credentials [Sui], crypto-assets [Zca, Mon, Pan], verifiable computation [Lur, ZKs], and decentralized storage [Sto, Swa, Sia, Labi], among others. A ZKP protocol allows a prover $\mathcal{P}$ to convince a verifier $\mathcal{V}$ of the validity of some statement without revealing any additional information about the statement to $\mathcal{V}$. More specifically, most ZKP applications rely on the technique of zero-knowledge succinct non-interactive argument of knowledge (zkSNARK) [BCCT12, AHIV17, GLS+23, XZS22] for sub-linear proof size and verification time. Since these properties are crucial for practicality, many recent developments focus on making these proofs even more efficient. This has also led to groups like Fabric [Fab], Ingonyama [Inga], and Supranational [Sup] working on dedicated hardware accelerators for ZK proof generation and verification.

At USENIX'20, Tramer, Boneh and Paterson [TBP20] demonstrated attacks on two privacy-focused cryptocurrencies, namely, Zcash [Zca] and Monero [Mon]. They exploited the timing-based side-channel leakages occurring when executing the protocol, allowing an attacker $\mathcal{A}$ to de-anonymize the secret payee (prover $\mathcal{P}$), violating the fundamental purpose of using private transactions. In their attack, the timing leakages stemmed from the high-level protocol properties such as changed communication patterns, or additional operations like commitments being computed only in certain cases. More recently at USENIX'24, [CET+24] took a broader look into the potential security vulnerabilities in

*shibam.mukherjee@iaik.tugraz.at

SNARKs protocols, in particular, they discuss various threat models and system-level vulnerabilities.

In this work, we discovered a zoo of ZK applications, potentially vulnerable to timing attacks [Koc96, BB03], in particular *cache-based timing side-channel attacks* [Ber05, TOS10]. We looked into leakages occurring in the fundamental components of ZK protocols like the underlying field arithmetic libraries used in the various constructions like Merkle trees or operations like in NTTs, polynomial evaluation, among others. More specifically, we looked into the Merkle tree construction using hash functions with potential leaky implementations.

Being the first work in the literature to explore such timing leakages, we only rely on the well-known Flush+Reload (F+R) [YF14] and Prime+Probe (P+P) [AES15, LYG+15, IGI+16] attacks and conjecture that other advanced timing attacks such as the ones with pre-fetching [CWS+24] or attaining sub-cache level resolution [SZB+24, YGH16] will also work, potentially with even fewer assumptions and higher precision. We show that when constructing Merkle Trees for ZK membership proofs [BdM93] with pairing-based [KZG10, CHM+20, Gro16] or FRI-based [BBHR18b] approaches, a non-constant time hash implementation[1] may leak the preimage containing sensitive information about $\mathcal{P}$'s private witness or potentially forfeiting the zero-knowledge property of the ZK protocols.

On scrutinizing the standard field arithmetic libraries like Rust `ff`, `ff-ce`, `ark-ff` crates, Python `galois` library, Typescript `circomlibjs`, we found several potential timing leakages in various field arithmetic operations like modulo addition, multiplication, reduction, exponentiation, among others. The same leaky field arithmetic libraries were used in several public implementation of ZK-friendly hash functions like *Rescue* [BBHR18b], Monolith [GKL+23], and even the de-facto industry-standard Poseidon [GKR+21], carrying forward the same timing vulnerabilities in the hash functions. It should be emphasised that the above hash functions themselves exhibit a constant-time algorithm, and the leakages occur only due to the use of underlying leaky field arithmetic libraries for the implementation. Additionally, we also looked into the family of lookup-based ZK-friendly hash functions, like Reinforced Concrete (RC) [GKL+22] and to some extent Tip5 [SLST23], which are inherently leaky due to the use of lookup-tables (LUTs) for efficient ZK proofs and verification. In such hash functions the timing leakage may originate both from the hashing algorithm itself (due to timing differences of memory-access) and also the underlying field arithmetic library.

Main Contributions. We summarize our main contributions in the following.

- For the first time in the literature, we discuss the potential implications of side-channel attacks on ZK applications, in particular, cache timing attacks on pairing-based and FRI-based ZK membership proofs.[2] In Section 3, we mount a black-box timing attack on the ZK protocols and show how $\mathcal{A}$ can forfeit the zero-knowledge property by leaking $\mathcal{P}$'s secret witness.

- In Section 4, we open the black box and show the existence of actual timing attacks in the fundamental components of the ZKP systems, like in the underlying field arithmetic libraries. We analyse the impact of cache based timing attacks, in particular F+R and P+P attacks on the public ZK protocols using leaky construction of Poseidon, and state the preimage recovery complexity of the same. We also show an attack on implementations of ZK-friendly hash functions that have faster plain performance due to use of lookup tables (Reinforced Concrete is the first

[1] Generally branched implementations are often faster than their constant time counterparts. This may serve as a motivation in the industry to use it for performance benefits.

[2] We also found some public discussions, like https://github.com/facebook/winterfell/issues/9, taking up the same concerns.

such proposal) and discuss the security implications. As part of our larger survey, a non-exhaustive list of such potentially vulnerable ZK applications found in the wild is also provided.

- Section 5 recommends the potential fixes for these vulnerabilities[3] and motivate the need for expedited research in the direction of micro-architecture based attacks [LSG+18, CWS+24, MWES19, SZB+24] or power side channel attacks [KJJ99, MOP07, RD20, LKO+21, AARR02, MDS99, CRR02] on ZK applications. Protecting against the above cache-based timing attacks, we also provide a constant-time implementation effort of some of the standard field arithmetic libraries (forked) used in the ZK protocols discussed in this work. On the bright side, in many cases we found that the constant-time implementations do not necessarily comes with significant performance penalty.

2 Preliminaries

We discuss the fundamentals of cache timing attacks like CPU caches, the concept of shared memory, and the two well-known cache timing attacks, F+R and P+P. Then we go through the fundamentals of zero-knowledge proofs and give a brief introduction on ZK-friendly hash functions and the role they play in pairing-based and FRI-based membership approaches.

2.1 Caches and Shared Memory

A standard CPU has a limited number of data registers, and thus cannot fit all the data it needs during a process execution. Any data that is required in the future is fetched from a low-latency memory into the CPU register before processing it. In most modern CPUs, caches provide this low-latency memory, however, they have a limited storage capacity and thus any large data must be fetched from the main memory using temporal and spatial data pre-fetching prediction algorithms. When executing a process, if the CPU finds the required data in the cache, we call it a *cache hit*. Otherwise, if the data needs to be fetched from the main memory, we call it a *cache miss*. The latter comes with a larger latency as the data needs to be fetched from the significantly slower main memory into the cache before the CPU can process it. For example, in a standard Intel CPU, the L1 cache has a latency of one to two clock cycles, whereas the L3 cache has a latency of around 30 to 40 clock cycles. In comparison, data fetching from the main memory may take up to 300 clock cycles.

Caches are usually partitioned into m smaller cache sets, which are further divided into n cache lines of b bytes each. We commonly refer to such caches as *n-way set-associative* caches, where the total cache size is $m \cdot n \cdot b$ bytes.[4] The L1 caches are shared between the sibling cores contained in a physical core, whereas the L3 cache is shared among all the physical cores and their respective sibling cores. Depending on the type of the cache-based timing attack, $\mathcal{A}$ and $\mathcal{P}$ may or may not have to share the same physical core in the threat model.

The concept of shared memory is heavily used by the operating system due to its performance benefits. For instance, several processes relying on the same library can access the shared physical memory where the library is mapped, averting the need for having several copies of the same library in the main memory for each process. Standardized

[3] Several micro-architectural mitigations have been proposed in the past, however, they come with their own trade-offs. In this work, we do not focus on these mitigations as it is a topic on its own interest, but only discuss the best constant-time implementation practices.

[4] For example, our test machine runs on an Intel i7-7600U CPU which has a 4 MB L3 cache with $m = 4096$, $n = 16$ and $b = 64$ bytes.

cryptographic implementations like OpenSSL [opeb], NSS [NSS], Libgcrypt [Lib], or even ZKP libraries can be examples of such shared libraries.

2.2 Cache Timing Attacks

Flush+Reload Attack. In 2016, Yarom and Falkner [YF14] presented the first Flush+Reload (F+R) attack to recover the RSA key. A F+R attack targets the shared instance, like a cryptographic library, mapped on the common cache accessible to both $\mathcal{A}$ and $\mathcal{P}$. $\mathcal{A}$ calls the `clflush` instruction to flush (remove) the cache lines containing the library data, for instance, the S-box lookup table (LUT) of AES or Reinforced Concrete. $\mathcal{A}$ later reloads the flushed data with `maccess` and measures the reload latency. If the data is not present in the cache, reload requires more time as the data needs to be fetched from the main memory. If a particular section of the flushed data has a smaller reload latency, $\mathcal{A}$ learns that $\mathcal{P}$ accessed that section, as the data was already loaded, potentially leaking the secret input to the cryptographic function, for example, the input to the S-box. Similar to the F+R attack, Gruss, Maurice, Wagner, and Mangard [GMWM16] proposed the Flush+Flush (F+F) attack, a window-less attack where reloading the data is not required, but instead $\mathcal{A}$ always flushes the memory, which has a lower latency than reloading the memory. This is particularly useful for achieving higher time resolution, especially when it is likely that $\mathcal{P}$ will access the data more than once between the flush and reload calls made by $\mathcal{A}$.

Prime+Probe Attack. In 2006, Osvik et al. [OST06] introduced the Prime+Probe (P+P) attack, followed by later improvements [AES15, LYG+15]. Compared to F+R, this attack does not require access to shared cryptographic library between $\mathcal{A}$ and $\mathcal{P}$, which naturally reduces the security assumptions one needs to make. In this attack, $\mathcal{A}$ primes (loads) all the cache sets with their eviction set (data) and then waits for $\mathcal{P}$ to load their data on the cache, replacing $\mathcal{A}$'s eviction set. $\mathcal{A}$ then sequentially probes the cache sets and looks for addresses where the eviction set was replaced. If the set was replaced by $\mathcal{P}$, $\mathcal{A}$'s access time to the set will be longer as the data needs to be fetched from the main memory into the cache. P+P attacks are generally more noisy and are limited only to the cache set resolution, giving low information leakage compared to F+R attacks which have a cache line resolution.

2.3 Zero-Knowledge Proof System

A zero-knowledge proof [GMR85] system allows a prover $\mathcal{P}$ to convince a verifier $\mathcal{V}$ the validity of a statement, using a public input (public data and the circuit) and potentially private witness data such that the public input reveals no information about the statement. For instance, $\mathcal{P}$ may want to convince $\mathcal{V}$ that they know a preimage x such that $y = H(x)$ for a publicly known value y and a cryptographic hash function H. A similar approach is also used for modern post-quantum signature schemes [CDG+17, BdSGK+21, DKR+22, BBM+24], especially the ones based on MPCitH and VOLEitH ZKP paradigms. Here $\mathcal{P}$ proves the knowledge of their secret key sk for a particular public key pk (plaintext-ciphertext pair), and the randomness is generated with a seed and the message to be signed. Two crucial properties of any ZKP protocol are *completeness* and *soundness*. Informally, the former ensures that an honest $\mathcal{V}$ will be convinced by an honest $\mathcal{P}$ and the latter ensures that a dishonest $\mathcal{P}$ cannot convince an honest $\mathcal{V}$ except with some small probability (usually designed to be negligible in the security parameter). Additionally, any ZKP protocol must also provide the *zero-knowledge* property.[5] Modern general-purpose ZKP systems can be

[5]We note that in some practical settings the ZK property may not be needed, but instead only the computational integrity property of modern general-purpose proof systems is used.

essentially split into two categories, namely pairing-based [Gro16, Coi, MBKM19, CHM+20, ABST23, Lee21, Set20] and FRI-based [BBHR18b, BCR+19, BFH+20] approaches. With the former, hash functions are often used to prove membership in zero knowledge, e.g. to prove on-chain coin ownership [Zca, Mon] or show group membership [Sem, Sui]. In contrast, FRI-based approaches often use these primitives internally in recursive proving approaches, which is a crucial building block of various ZK-rollup applications [Her, ZKs, Tus, Azt].

2.4 ZK-Friendly Hash Functions

Modern ZK protocols often rely on hash functions which exhibit a property referred to as *circuit friendliness*, *ZK friendliness*, or *arithmetization friendliness*, important for strong performance when proving particular statements in ZK. This essentially means that the algorithmic description of the hash function can be translated to a relatively small system of low-degree constraints, which makes these constructions efficient in proof systems for computational integrity. In the following, we give a brief overview of two such constructions, one of them being among the most popular choices and the other being a more modern approach using lookup arguments in the proof.

Poseidon. Poseidon [GKR+21] is a cryptographic hash function designed for efficient implementation as an arithmetic circuit, making it particularly suitable for zero-knowledge proof systems. It operates over a prime field p, where $p \approx 2^n$ and $n \geq 31$, mapping the input messages of arbitrary length to fixed-size digests. The underlying sponge permutation, Poseidon$^\pi$: $\mathbb{F}_p^* \to \mathbb{F}_p^o$, where o is the fixed output length measured in p elements, is based on a substitution-permutation network (SPN) first used for HadesMiMC [GLR+20]. It consists of several rounds with two different types of layers, namely *full* (or *external*) and *partial* (or *internal*) rounds. The former includes full nonlinear layers where S-boxes are applied to every word in the state. The latter consists of a partial nonlinear layer where the S-box is only applied to the first word of the state. The idea behind this approach is to provide simple arguments against statistical attacks using consecutive full rounds, while achieving the same degree growth using the more efficient partial rounds. The affine layer $M(\cdot)$ consists of a $t \times t$ MDS matrix multiplication and a round constant addition ARC$(\cdot)$. The nonlinear S-box function is defined as $S(x) = x^\alpha$, where $\alpha \geq 3$ is a small positive integer for which $\gcd(p-1, \alpha) = 1$. An overview of the Poseidon$^\pi$ permutation is shown in Fig. 1.

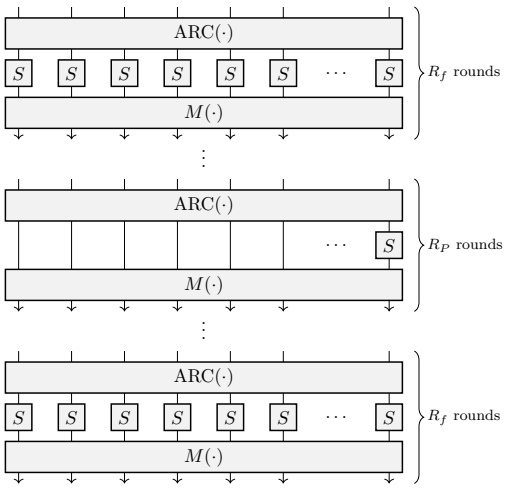


Figure 1: The Poseidon$^\pi$ permutation.

In computational integrity proof systems, the advantages of Poseidon stem from its low number of nonlinear operations, primarily due to the minimal use of S-boxes. This property makes it easier to express the hash function as a circuit and prove knowledge of its preimages within the proving framework. Indeed, nowadays Poseidon is used in various ZK-related applications, including RISC Zero [RIS], Plonky2 [Polb], Plonky3 [Polb], and for many on-chain use cases such as Filecoin [Labi], Dusk Network [Dus], Sovrin [Sov], Loopring [Loo].[6] We refer the reader to Poseidon [GKR+21] for more details.

Reinforced Concrete (RC). Reinforced Concrete [GKL+22] is another ZK-friendly sponge-based hash function which maps $\mathbb{F}_p^3 \to \mathbb{F}_p^3$ for some prime p, operating over two elliptic curves, namely BN254 and BLS12-381 and one special prime field (-ST) crafted for performance. Reinforced Concrete uses lookup-based S-boxes to take advantage of proof systems which use the lookup argument for efficient set membership proofs. The underlying permutation consists of a modified 7-round SPN with Concrete, Bricks and Bars layers, where Concrete $\circ$ Bricks is one round, and a single Bars layer is applied in the middle of the permutation. The Bricks function is a nonlinear permutation. The Concrete function is an MDS matrix multiplication with the state in $\mathbb{F}_p$. The Bar layer consists of composition/decomposition and S-box functions, where the Bar layer is implemented as a lookup table of the functions it contains. RC can be considered as an attractive drop-in replacement for Poseidon due to its increased efficiency for lookup-based protocols like Arya [BCG+18], Plookup [GW20, PFM+22], Halo2 [Zca], and logUp [Hab22]. A graphical representation of the RC function is provided in Fig. 2. We refer the reader to RC [GKL+22] for more details.

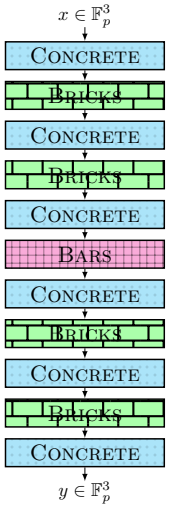


Figure 2: The Reinforced Concrete permutation.

3 Hash Functions in Zero-Knowledge Proofs

We distinguish between two types of general-purpose ZK proof constructions, namely pairing-based and FRI-based ones. Circuit-friendly hash functions are regularly used in both of them, and we give a summary of two main settings below.

[6] By "on-chain use cases" we refer to operations and data executed and stored on the main blockchain directly.

For the purpose of this section, we utilize the side-channel technique as a blackbox component in order to describe the setting and the general approach of the attack. In the next section, we go into the details of the side-channel part of the attack specifically.

3.1 Hash Functions in Pairing-Based Proofs

ZK-friendly hash functions defined over pairing-friendly elliptic curves are a powerful tool in many real-world scenarios. In particular, many proof systems [CHM+20, ABST23, Gro16] (especially those directly employed on chains like Ethereum) are based on this type of elliptic curves in order to support efficient verification of proofs. In this context, a ZK-friendly hash function can be translated to a circuit with minimal overhead.

To give an example, let us consider a simple membership proof where $\mathcal{P}$ wants to convince $\mathcal{V}$ that they hold some secret input (say, a coin). The input is part of a publicly verifiable set fixed by a Merkle root (say, a block on the blockchain containing information about the ownership of the coin). In this case, a hash function can be used to prove knowledge of the secret input without trivially revealing it to $\mathcal{V}$. In order to prove the opening path in the tree efficiently, a ZK-friendly hash function is used to construct the Merkle tree, resulting in easily verifiable constraints when used together with a pairing-based proof. Here, $\mathcal{V}$ verifies if the path from the leaf to the root of the Merkle tree was computed correctly by checking the correct execution of all the d hash function calls on the path in ZK, where d is the depth of the Merkle tree. In the context of cache timing attacks, any information leakage from the hash function also directly leaks $\mathcal{P}$'s secret witness (leaf) values. Using classical hash functions like SHA-3 makes this approach significantly more expensive in general, as these often result in a more complex and less efficient constraint system when computing the membership proof (i.e., the path from a leaf to the public root).

3.2 Hash Functions in FRI

Similar to pairing-based approaches, ZK-friendly hash functions are also used in FRI-based approaches [BBHR18a] for constructing Merkle trees during the commitment phase. However, unlike the pairing-based approach, FRI protocols do not directly take the secret witness as the input to the hash function, but instead, at a very high level, a polynomial f is constructed by interpolating the private witnesses, which is then evaluated at some special points (discussed later), and the outputs become the leaves of the Merkle tree fed to the hash functions for computing the Merkle root. Here $\mathcal{P}$ proves to $\mathcal{V}$ that they know a set of points (witnesses) which pass through f of a particular degree d (among other parts in the entire proof). In most applications, these witness points are execution traces of a program running on $\mathcal{P}$'s machine, where $\mathcal{P}$ proves the correct execution of the program to $\mathcal{V}$. The circuit friendliness of a given hash function is particularly important for efficient *recursive proofs*, which essentially include proving verifications of many other proofs.

Currently, FRI-based approaches are used in many popular projects widely adopted in the industry [Scr, RIS, Sta, Wit, Pola], mainly due to their advantages like simple setups and high prover performance. Indeed, many recent developments in the area of zero-knowledge protocols have focused on making FRI-based proving more efficient [HLP24, ACFY24].

Adding Zero Knowledge to FRI. While FRI is often used without the zero knowledge property (e.g., exploiting the succinctness property for arguments of computational integrity), in many cases ZK is needed and activated. For that, various other strategies have to be applied. When focusing on the trace in particular, one part consists of adding random values to an existing witness column in the trace, in order to change the length of the trace to a more suitable number (for example, a power of 2) and to mask the potentially private

witness values. We refer to [HK24] for more details on this approach and a summary of different techniques.

Another step consists of masking the Merkle tree leaves with random paddings, in order to prevent leaking information in the opening part of the proving protocol. A more detailed description of this approach is given in [BCS16].

Since our goal is to find private witnesses, we consider a scenario where zero knowledge is enabled.

Side-Channel Attack on FRI. Let us consider a trace domain $\mathcal{D}$, which we assume to be a multiplicative subgroup of order N generated by ω. The first step of $\mathcal{P}$ is to take the witnesses $\{w_i\}_{i=0}^{N-1}$ in a trace column and to find a polynomial f such that $f(\omega^i) = w_i$. This can be done by an inverse number-theoretic transform (NTT). Note that for simplicity, in this description we omit the random non-witness values added for zero knowledge. We will show that later this makes no difference in the context of our side-channel attack.

After this step, we define the evaluation domain to be a slightly larger domain $\mathcal{D}'$ of $\beta \cdot N$ elements, where β is called the *blowup factor*, generated by $\omega^{1/\beta}$ and shifted by s, i.e., $\mathcal{D}' = \{s \cdot \omega^{i/\beta}\}_{i=0}^{\beta \cdot N-1}$ and $|\mathcal{D}'| = \beta \cdot |\mathcal{D}|$. $\mathcal{P}$ then commits to the evaluations of f over $\mathcal{D}'$ (along with random padding, again added for zero knowledge) using a Merkle tree, where the Merkle root is the commitment. We refer to [BBHR18a, BBHR18c] for more details.

In this scenario, the task of $\mathcal{A}$ is to find the potentially private witness values. This can be done if the evaluations $\{f(s \cdot \omega^{i/\beta})\}_{i=0}^{\beta \cdot N-1}$ are leaked. Indeed, note that the polynomial f originates from the witnesses and, with overwhelming probability, is of degree $N-1$. Hence, it suffices to choose any $N-1$ points from $\{(s \cdot \omega^{i/\beta}, f(s \cdot \omega^{i/\beta}))\}_{i=0}^{\beta \cdot N-1}$ and to find a polynomial $f^\star$ which goes through these points. With high probability, $f = f^\star$, and thus evaluating $f^\star$ on $\mathcal{D} = \{\omega^i\}_{i=0}^{N-1}$ is sufficient to find the witness values $\{w_i\}_{i=0}^{N-1}$.

Summarizing, the attack consists of the following steps.

1. Use side-channel leakage to find the evaluations of the trace polynomial f over $\mathcal{D}' = \{s \cdot \omega^{i/\beta}\}_{i=0}^{\beta \cdot N-1}$.
2. Use $|\mathcal{D}| = N$ of these evaluations to find a highly likely candidate $f^\star$ for f (this can be done by interpolation).
3. Evaluate $f^\star$ on $\mathcal{D}$ to find $\{f^\star(\omega^i) = w_i^\star\}_{i=0}^{N-1}$. Again, with high probability, $\{w_i^\star\}_{i=0}^{N-1} = \{w_i\}_{i=0}^{N-1}$, which corresponds to the original witness values.

The last two steps can be handled relatively easily from $\mathcal{A}$'s point of view. However, the first step requires more assumptions, in particular, using high-resolution cache timing attacks, like sub cache line attacks [SZB+24, YGH16], mounted on hash implementations with byte-level leakages, like when using LUTs for the S-boxes in RC hash function. This allows $\mathcal{A}$ to directly learn $\{f(s \cdot \omega^{i/\beta})\}_{i=0}^{\beta \cdot N-1}$ and reconstruct f with trivial complexity as discussed above.

When using low-resolution attacks like F+R and P+P, or if the hash implementation leaks only a few bits of information, $\mathcal{A}$ obtains several candidates for every $\{f(s \cdot \omega^{i/\beta})\}_{i=0}^{\beta \cdot N-1}$ preimage. $\mathcal{A}$ can then iterate through all these candidates in order to reconstruct f. In particular, this is possible if the evaluations of f over $\mathcal{D}'$ exhibit a strong structure, which however is mainly mitigated due to the diffusion properties of the NTT application. Indeed, the evaluation of f over $\mathcal{D}'$ results in a highly unstructured table even for small variations in $\{w_i\}_{i=0}^{N-1}$. An obvious exception occurs if all values in $\{w_i\}_{i=0}^{N-1}$ are the same (i.e., $f(x) = w_0$), which however is an unrealistic scenario both due to the practical use cases for proof systems and due to additional steps, such as padding applied to trace columns.

Zero Knowledge and Side-Channel Resistance. In the context of our cache timing attack, some of the additional techniques to achieve zero knowledge are no obstacle, as recovering

the preimage includes both the evaluation points (evaluations of f at publicly known inputs) and the random paddings, which allows $\mathcal{A}$ to fully reconstruct f via interpolation. Once f is reconstructed, $\mathcal{A}$ can trivially retrieve the secret witness by evaluating f at the publicly known input points. The added random values to mask the interpolating polynomial are no obstacle here.

We also emphasize that the initial step of the interpolation is applied to potentially private witness data. It remains a future direction to investigate how this can be exploited in cache attacks with the high-resolution leakages and other side-channel approaches.

4 Cache Timing Attacks

We define the threat model under which the F+R and P+P attacks are mounted and discuss the complexity of recovering the preimage of the Poseidon and RC public implementations [Gra]. Finally, we provide a non-exhaustive list of all the open source ZK libraries that we found were using pairing-based and FRI-based ZK approaches with non-constant time Poseidon hash implementation. We should emphasise, we only point out the potential timing leakages in the implementation, especially critical for secure pairing-based approaches. For the FRI-based approaches, we took more interest in breaking the zero-knowledge property by recovering some bits of the preimage in the Merkle leaves.

Threat Model. We assume that $\mathcal{P}$ is generating the proof for their witness on a shared machine $\mathbf{M}$ (like a cloud), allowing $\mathcal{P}$ to access powerful hardware for generating faster proofs while minimizing the cost by sharing resources. An attacker $\mathcal{A}$ (e.g., pretending to be another legit prover), using $\mathbf{M}$ runs their malicious side-channel programs and learns some (if not all) bits about $\mathcal{P}$'s secret witness through cache-based timing attacks when $\mathcal{P}$ generates the ZK proof for $\mathcal{V}$. If $\mathbf{M}$ is not a cloud, but rather $\mathcal{P}$'s private machine, we assume that $\mathcal{A}$ is running on $\mathbf{M}$ through more sophisticated attacks like browser-based side-channel exploits [GCG+24, SAO+21]. Due to the zero-knowledge property, we know that $\mathcal{V}$ cannot extract any new information about the secret witness by reading the proof sent by $\mathcal{P}$. Likewise, if $\mathcal{A}$ gets access to the same proof sent to $\mathcal{V}$, $\mathcal{A}$ cannot learn anything new about the witness either, and thus any side-channel attack performed on $\mathcal{V}$'s end is meaningless.

Both $\mathcal{A}$ and $\mathcal{P}$ are required to run on the same CPU, however, they may run on different physical cores as F+R and P+P attacks target the LLC shared by all cores. For the F+R attack we assume that $\mathcal{A}$ and $\mathcal{P}$ share the cryptographic library containing the ZK functionalities like hash functions and arithmetic field operations as a shared object. Also, the `clflush` instruction must be available to $\mathcal{A}$, essential for F+R attacks, otherwise $\mathcal{A}$ may have to perform other similar attacks like Evict+Time [OST06] or Evict+Reload [GMM16]. For some implementations, as a common practice in the community, when demonstrating proof-of-concept cache-based timing attacks, we use additional `nop` instructions to align the branches on different cache lines. Naturally, such an instruction is equivalent to calling functions from different branches which may get aligned to a new cache line. Another example is compiler optimization which may put the secret-dependent branched code unfavourably on distinct cache lines, or even cache sets, leading to practical timing attack like in the case of the Kyber constant-time reference implementation.[7] When employing more modern side-channel attacks, like the ones with speculative execution [KHF+20, LSG+18], pre-fetching [CWS+24], sub-cache line attacks [YGH16, SZB+24], among others [GRBG18, vSGBR18], one may not require including such `nop` instructions. Before independently verifying the leakages in the hash

[7] https://pqshield.com/pqshield-plugs-timing-leaks-in-kyber-ml-kem-to-improve-pqc-implementation-maturity/

Table 1: A summary of the threat model involving all the 3 parties, $\mathcal{A}$, $\mathcal{P}$ and $\mathcal{V}$ in a ZKP protocol.

Parties	Machine	Cache SCA		Assumptions
		F+R	P+P	
Prover	Cloud	Yes	No	Shared ZKP library
		N/A	N/A	Additional instructions
		N/A	N/A	Attack resolution
Attacker	Cloud	Yes	No	Shared ZKP library
		`clflush,nop`	`nop`	Additional instructions
		Cache line level (64 bytes)	Cache set level (1024 bytes)	Attack resolution
Verifier	N/A	N/A	N/A	N/A

functions with F+R and P+P attacks, we used the DATA framework [WZS+18][8] to search for any secret-dependent timing leakages in the hash implementations. A summary of the threat model can be found in Table 1.

Reinforced Concrete Cache Timing Attack. The RC hash function uses lookup tables (LUTs) in the Bars function. Any LUT implementation is vulnerable to cache timing attacks, especially when the LUT lies on multiple cache lines/sets. For the F+R attack, we assume that $\mathcal{A}$ and $\mathcal{P}$ (victim) share the RC hash function as a shared library including the Bars layer LUT. In RC-BLS12-381, the $\mathbb{F}_p$ elements in the Bars layer are *decomposed* into 27 elements in $\mathbb{F}_{659}$ and then the S-box lookup is applied. Even though $\mathbb{F}_{659}$ can be represented with 10 bits, for efficiency purposes, the elements are mapped as 2-byte elements in the memory. This means the entire LUT is 1318 bytes in size, requiring 21 cache lines or 2 cache sets on the LLC of our test machine.

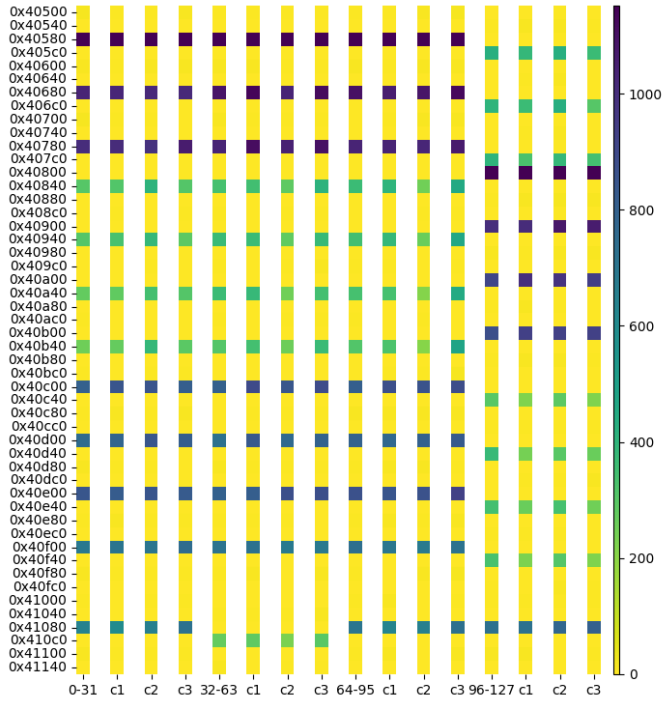
In the F+R attack, with every cache hit $\mathcal{A}$ learns ≈ 5 bits of equivalent (total input in bits$-$possible candidates in bits) S-box input information. In other words, out of 659 elements ($\approx 2^{9.32}$), $\mathcal{A}$ learns the possible $32 = 2^5$ or $19 \approx 2^{4.2}$ (the last cache line) input candidates for the S-box. In P+P attack, however, the recovery resolution is worse than F+R as the resolution is constrained to the cache set level only. In particular, $\mathcal{A}$ learns ≈ 2 bits of equivalent S-box input information, comparable to learning the possible $512 = 2^9$ or $147 \approx 2^7$ (last cache set) input candidates for the S-box. We observe somewhat similar results for RC-BN-254 and RC-ST. However, in the BN-254 implementation there exists a cache line with a single field element, allowing $\mathcal{A}$ to learn all 9.32 bits of equivalent input information in the best case F+R attack scenario occurring with a non-negligible probability of $(642)^{-1}$. We refer to Table 2 for an overview of the preimage recovery complexity. Refer to Fig. 3 for illustration of the F+F[9] and P+P attack on the RC hash function.

When considering stronger attacks, like the new sub-cache line resolution timing attack [SZB+24], here $\mathcal{A}$ may learn all the S-box input bits, fully recovering the preimage without additional search operations. Moreover, for the F+R and P+P attacks, the preimage recovery complexity also decreases when the preimage has some structure to it. This is particularly the case in many ZK applications where the Merkle leaves contain, for example, personally identifiable information like name or dates of birth. Padding some randomness (salt) to the leaves may help against side-channel attacks where $\mathcal{A}$ merely observes the hashes of the tree. However, because the padding is only applied to specific positions of the leaves, this mitigation does not protect against cache timing attacks.

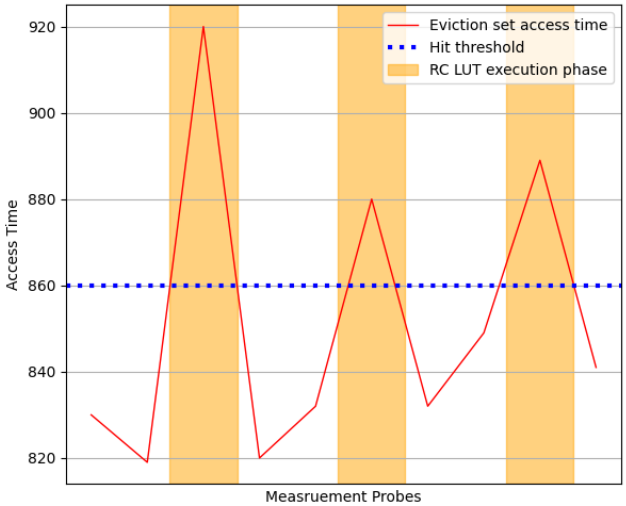
In the following subsections, we discuss some potentially vulnerable Merkle tree constructions used in real-world ZK applications.

[8] https://github.com/Fraunhofer-AISEC/DATA/tree/master

[9] We use F+F instead of F+R for attacking RC as the former gives less noisy measurements. The recovery complexity remains the same in both the cases.



(a) F+F cache-hit and cache-miss ratio for Reinforced Concrete BLS-12-381 hash function. Darker blocks represent higher cache-hits in party containing S-box LUT. The input of each S-box starts from `0x0000` and sequentially increases by a multiple of `0x0020` probing the first 4 cache lines, containing 32 inputs each, accessed by $\mathcal{P}$ when generating the Merkle tree. We repeat the attack (c1..c3) demonstrating consistency in the timing leakage when calling S-box inputs contained in the same cache line.



(b) Probe measurements of Reinforced Concrete BLS-12-381 hash function with P+P attack. In the orange phases, the LUT S-box inputs 0-511 are called. We probe the eviction set contained in the cache-set where the 0-511 S-box inputs lie, giving slower probe time compared to not calling the 0-511 S-box inputs.

Figure 3: Access time for F+F and P+P attacks on the Reinforced Concrete hash function. Here $\mathcal{A}$ probes the S-box LUT memory addresses and measures the time to distinguish if the same address was accessed by $\mathcal{P}$ or not.

Poseidon Cache Timing Attack. As a demonstration, we only attack the Poseidon Goldilocks and Mersenne non-constant time implementations [Gra] and summarize their leakages.[10] In particular, we attack the ARC$(\cdot)$ function performing modulo reduction after the field addition between the *state* and the round constants rc. Depending on the branch, $\mathcal{A}$ can determine if the sum between a state and a round constant is larger than the modulo prime p, where extreme values of round constants (very small or large values in the field) give the most information about the secret state. Additionally, we can also attack functions like $M(\cdot)$ and $S(\cdot)$ containing non-constant time modulo reductions after multiplication operations. However, here we focus on the ARC$(\cdot)$ function, as this is the first layer that directly operates on the preimage. We assume that the `if-else` branches lie on two different cache lines or cache sets for the F+R or P+P attacks respectively.

In the F+R attack, similar to RC, we assume that $\mathcal{A}$ and $\mathcal{P}$ (victim) are sharing the Poseidon library as a shared object, however, $\mathcal{A}$ instead of flushing the LUT addresses, they flush the addresses of the `if-else` branches to detect which branch is accessed by $\mathcal{P}$.

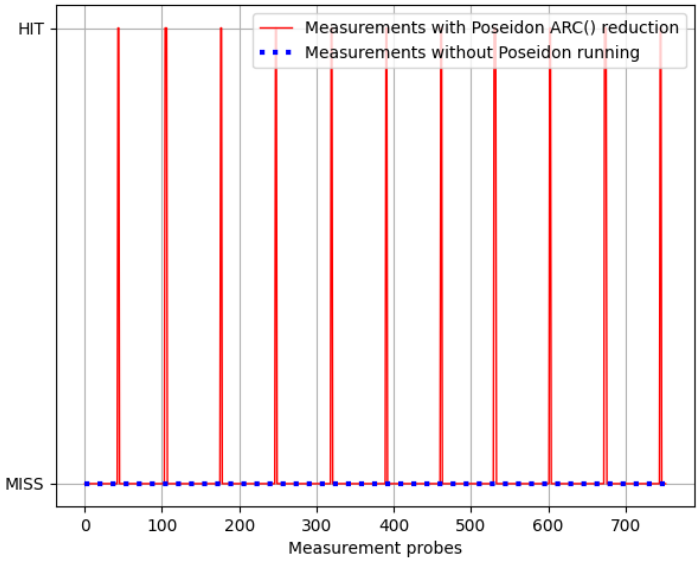
As an input to the Poseidon$^{\pi}$ permutation, Poseidon-(M, $t = 16$) takes an element in $\mathbb{F}_p^t$, where $p = 2^{31} - 1$ and $t = 16$, where the first layer of Poseidon$^{\pi}$ is the round constant addition. With the smallest round constant (`0x002f87c1`) in Poseidon-(M, $t = 16$), during a cache hit on the reduction branch, $\mathcal{A}$ learns the most significant ≈ 9 bits (all ones) of $\mathbb{F}_p$ input, where p is `0x7FFFFFFF`.[11] With the largest round constant (`0x7fc1a254`), $\mathcal{A}$ also learns the most significant ≈ 9 bits (all zeros) of the $\mathbb{F}_p$ input. If the round constant rc is centered in $\mathbb{F}_p$ (i.e., $rc \approx 2^{30}$), it leaks only one bit equivalent information. We observed similar leakages when inspecting Poseidon-(M, $t = 24$) and refer to Table 2 for more details and an overview of timing leakages using different versions of the Poseidon hash function.

The generated round constants will be large with an overwhelming probability. However, for large fields like in BLS-12-381 or BN254, usually custom small round constants are preferred as they provide better plain performance (fewer instructions are necessary for the round constant additions). Refer to Fig. 4 for the F+R and P+P attack on the Poseidon-GL hash function.

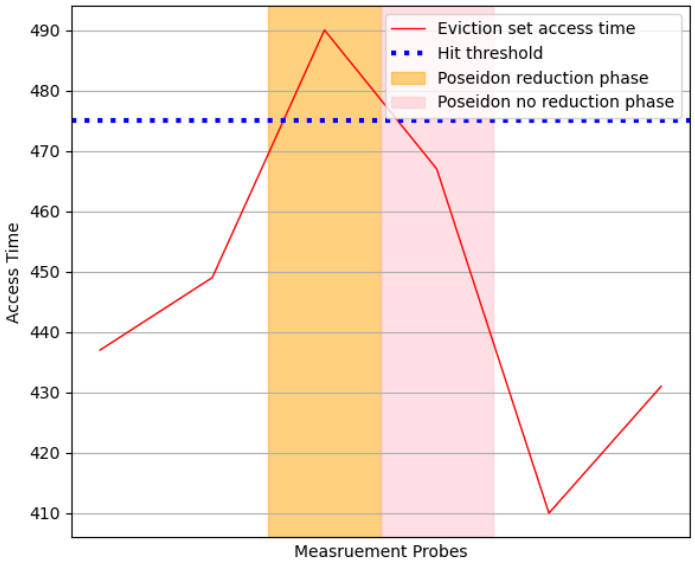
Merkle Tree Leaf Distinguisher. Here we discuss the potential attacks that can be mounted on real-world applications, in particular, targeting the Merkle tree construction using non-constant time hash implementations like for Poseidon. One specific aspect we particularly focus on is the entropy of the private values in the Merkle leaves. Low-entropy values, even when padded with randomness at certain points (a common security practice [BCG+14, Sia]) leak (partial) information about the preimage and hence the secret witness. For example, the Zerocash protocol [BCG+14], a foundational work of the Zcash cryptocurrency [Zca], hashes ($\mathcal{H}(\cdot)$) the coin values $v \in [0, 2^{64} - 1]$ with randomness $k \in [0, 2^{256} - 1]$ as $cm := \mathcal{H}(k \mid\mid 0^{192} \mid\mid v)$, where cm is the commitment value of the coin in the Merkle tree. The authors discuss using SHA-256 as the hash function, however in practice (Zcash), when proving Merkle tree commitments in ZK, Poseidon is the more efficient choice and any non-constant time implementation when constructing the Merkle tree, in this example, will leak the coin value v in accordance with Table 2. Similarly, Sia [Sia], a decentralized storage protocol, also pads its low-entropy preimage data used in the Merkle tree, such as a *time lock*, the number of public keys, and signatures required, by adding a random nonce to the leaves along with the secret input data. Another example is Semaphore [Sem], a ZK signaling framework, where the public id id_{pub} is padded with a random sequence of bytes $id_{\text{nullifier}}$ when committing to the user identity for generating the Merkle tree membership proofs. These random paddings were adopted to prevent

[10] The 64-bit Goldilocks prime $2^{64} - 2^{32} + 1$ and the 31-bit Mersenne prime $2^{31} - 1$ define two popular prime fields which are often used for efficient zero-knowledge protocol implementations based on FRI.

[11] Here $\mathcal{A}$ has the same leakage resolution for both F+R and P+P attacks.



(a) Cache-hit and cache-miss ratio for Poseidon-GL hash function with F+R attack. We set the input to the hash function such that modulo reduction (cache hit) is called after every 100 hash calls. We probe the cache address containing the modulo reduction operation branched from the `if-else` check.



(b) Probe measurements of Poseidon-GL hash function with P+P attack. In the orange phase, the round constant addition is followed by the modulo reduction operation. In the pink phase no such reduction is performed. We probe the eviction set contained in the cache-set where the reduction operation instruction also lies, giving slower probe time compared to the phase not executing the modulo reduction.

Figure 4: Access time graph for F+R and P+P attacks on Poseidon-GL hash function. Here $\mathcal{A}$ probes the `if-else` branch memory addresses and measures the time to distinguish if the same addresses were accessed by $\mathcal{P}$.

Table 2: Preimage recovery complexity after F+R and P+P attacks on the Reinforced Concrete and Poseidon hash functions. Here we provide both the average and *best* case input recovery complexity, where the *best* case occurs with a fairly large probability. The preimage leakage is given in terms equivalents bits of information learnt from the full state of the hash function.

Hash function	Timing vulnerability	Preimage leaked (in bits)			
		F+R	% leaked	P+P	% leaked
Poseidon-(M, $t = 16$)	Reduction in ARC(·)	1×16	3.22	1×16	3.22
		$\approx 9 \times 16$	29.03	$\approx 9 \times 16$	29.03
Poseidon-(M, $t = 24$)	Reduction in ARC(·)	1×24	3.22	1×24	3.22
		$\approx 9 \times 24$	29.03	$\approx 9 \times 24$	29.03
Poseidon-(GL, $t = 8$)	Reduction in ARC(·)	1×8	1.56	1×8	1.56
		$\approx 9 \times 8$	14.06	$\approx 9 \times 8$	14.06
Poseidon-(GL, $t = 12$)	Reduction in ARC(·)	1×12	1.56	1×12	1.56
		$\approx 9 \times 12$	14.06	$\approx 9 \times 12$	14.06
RC-(BLS12-381, $n = 27$)	Bars S-box LUT	4.36×27	45.98	0.36×27	3.79
		5.12×27	54	2.17×27	22.88
RC-(BN-254, $n = 27$)	Bars S-box LUT	4.32×27	45.92	0.32×27	3.4
		9.32×27	100	2.31×27	24.55
RC-(ST, $n = 25$)	Bars S-box LUT	4.98×25	49.8	0.98×25	9.8
		5.59×25	55.9	1.02×25	10.2

multiple signal broadcasts using the same public id id_{pub}.

In certain conditions, it is even sufficient if $\mathcal{A}$ is able to roughly distinguish the leaves instead of fully recovering the input. For example, let us assume the above Poseidon-(M, $t \in \{16, 24\}$) case where a Merkle tree contains four leaves, the first two leaves contain short phone numbers from Falkland Islands (5 digits, at most 17 bits) and the next two leaves contain phone numbers from Norway (8 digits, at most 27 bits).[12] Then with the smallest and largest Poseidon-(M, $t \in \{16, 24\}$) round constants, $\mathcal{A}$ can distinguish between the two groups of the leaves when $\mathcal{P}$ generates the Merkle tree.

Leakages in Zero-Knowledge Protocols. Table 3 shows a non-exhaustive list of ZK frameworks, spread across various ZK applications using various non-constant time Poseidon implementations. Most of the vulnerabilities root from the underlying field arithmetic library, except for ZK-kit, where they implement their own field library. In case of Panther and Sui, the underlying jdk native BigInt implementation is the source of leakage, known since 2020 [BRB20]. The poseidon-lite library acknowledges the timing leakage in its disclaimer, however, in both the Sui and the Panther ZK framework such declarations were found missing. Currently, we looked into only a handful of the ZK frameworks which use such leaky field libraries, however, we suspect that there are several other frameworks using them as well.

Even though most of the underlying field libraries state that they are actually vulnerable to timing attacks and thus should not be used for production, however, at the time of writing this paper, they were being used in many ZK libraries. Moreover, to the best of our knowledge, none of the ZK library security audits consider timing attacks in their threat models, which we believe in the future should be taken into consideration as a standard practice as is done in other areas of cryptographic implementations.

We also looked into other ZK protocols for potential timing leakages, in particular Plonky2, Plonky3 and Halo2. We found branching in the addition and the subtraction operations in both Plonky2 and Plonky3 using Poseidon-GL. However, the authors claimed, and we also independently verified from their public implementation, that this

[12] https://worldpopulationreview.com/country-rankings/phone-number-length-by-country

branching happens with a very low probability. In particular, for this to occur, there must exist a pair of a round constant and an input both $\geq (2^{64}-1)-(2^{30}-1)$, and hence it is more efficient to branch. In Halo2, we found the Poseidon S-box layer implementation to contain a non-constant time exponentiation `pow_vartime()`, but the exponent is already public and thus does not affect the security. However, special care must be taken in the future to not use this function for any secret dependent operation and thus we advice switching to constant time implementation as standard. We also looked into Aztec [Azt], and to the best of our knowledge, at the time of writing this paper, we did not find any non-constant time Poseidon implementation.

Currently, these particular findings do not result in a practical attack against the ZK proof systems. However, we still recommend having constant time implementations of the individual components to mitigate any risk associated with the leakages, especially given the fact that the performance differences is in many cases not significant [GKL+23] (Table 4). With this work, we are providing forked versions[13] of the `ff`, `ff-ce` and `ark-ff` crates which will include constant-time implementations of the functions discuss in Table 3. As a future collaboration with (but not limited to) the maintainers of the above libraries, we are currently looking into the possibility of integrating our constant-time proposal into the standard ZK libraries.

5 Further Discussion

Other Hash Functions. As the timing leakages mostly stem from the underlying arithmetic library, we also looked into other ZK-fiendly hash function candidates like Tip5, *Rescue*, and Monolith. Due to use of lookups operating over $\mathbb{F}_{2^8+1}$ in one of its type of S-boxes, Tip5 may also suffer from the same problems as RC. When performing the F+R attack, $\mathcal{A}$ learns 3 bits of equivalent information in the average case, and all 8 bits in the best case, occurring with a probability of $(2^8+1)^{-1}$. We also found some timing leakages in the public implementations of Monolith and *Rescue*. Depending on the *Rescue* implementation, for example, leakage in [Labd] may occur due to the use of crates like `ff-ce`, similar to Poseidon, or the leakage is only limited to the testing section of the framework [Wit] (`inv()` and `normalize()`). The non-constant time version of the Monolith implementation [Gra] relies on the same Goldilocks/Mersenne field implementations as Poseidon, thus it might be affected as well. However, Monolith also has a constant time version which is not affected.

Other Timing Attacks. This work focused on the well-known F+R and P+P timing attacks on ZKP protocols and their applications. We conjecture that more sophisticated attacks like [MWES19] with attacks on constant-time implementations, and cache timing attacks with sub-cache line resolution [YGH16, SZB+24] will reduce the preimage recovery complexity significantly, especially for lookup-based hash functions like Reinforced Concrete and potentially for Tip5. In the future, it might be interesting to further investigate the possibility of improving the preimage recovery complexity for the ZK-friendly implementations and their use cases in the wild. Power side-channel attacks might be another interesting direction.

Constant-Time Implementation: Best Practices. We observe that non-constant time implementations may leak sensitive input data in a ZKP protocol. Even though attacks like P+P and F+R were discussed several years ago, we still find exploits using these attack vectors on fairly recent protocols like in ZK and their applications. We believe that strengthening these protocols against side-channel attacks is paramount, especially

[13] https://extgit.iaik.tugraz.at/krypto/ffconstzksca

Table 3: We summarize the ZK frameworks (a non-exhaustive list) that may contain potential witness leakages due to F+R and P+P timing attacks on the various non-constant time Poseidon hash function implementations. We also state the source of the potential leakage affecting both the $ARC(\cdot)$ and $M(\cdot)$ layers.

ZK library (Application) - Source of leakage	Non-constant time implementation	Comments
Matter Labs[Labc] (Blockchain ZK) - rescue-poseidon [Labd] → ff-ce crate	Uses `ff-ce` crate with branched `reduce` in `add_assign`	`Readme.md` (`ff-ce`) - Does not provide constant-time guarantees
Lurk Labs [Lur] (Turing-complete ZK programming language) - neptune-poseidon [Laba] → ff crate [ff]	Uses `ff` crate with branched `reduce` in `add_assign`	`Readme.md` (`neptune-poseidon`) - Has been audited `Readme.md` (`ff`) - Does not provide constant time guarantees
Ingonyama poseidon-hash [Inga] (Hardware acceleration for ZKP) - poseidon-hash [Ingb] → galois [mho]	Uses `Galois` library with branched `add_modular`	`Readme.md` (`galois`) - The library could be vulnerable to timing attacks. Not intended for production.
Sui [Sui] (ZK login and authentication) - Sui [Labe] → poseidon-lite [Vim] → jdk (BigInt) [Opea]	Uses `poseidon-lite` library with branched addition using `jdk/../BigInteger.java`	`Readme.md` (`poseidon-lite`) - The code has not been audited and the native js `BigInt` is vulnerable to timing attacks.
ZK-kit [Zka] (General ZK library for developing various applications) - ZK-kit [zkb]	Field arithmetic `f1-field.ts` contains branched addition.	–
Panther [Pan] (Virtual asset private trading) - Panther [Proc] → circomlibjs [Idea] → ffjavascript [Ideb] → jdk (BigInt) [Opea]	Uses `circomlibjs` Poseidon implementation using `ffjavascript` for BN-128 and BLS12-381 with branched addition performed with `jdk/../BigInteger.java`.	–
Light Protocol [Labb] (ZK layer on Solana [Sol]) - Light Protocol [Prob] → light-poseidon [Proa] → ark-ff crate [af]	Uses `light-poseidon` using `ark-ff` crate for arithmetic operations. `add_assign` in `montgomery_backend.rs` has branching	`Readme.md` (`light-poseidon`) - Has been audited `Readme.md` (`ark-ff`) - Academic proof of concept, not ready for production use
Mina Protocol [Min] (Decentralized Apps) - O1-labs/proof-systems [Labf] → ark-ff crate [af]	Uses ark-ff crate, `add_assign` in `montgomery_backend.rs` has branching	`Readme.md` (`proof-systems`) - Security audit missing. `Readme.md` (`ark-ff`) - Academic proof of concept, not ready for production use.
Plonky2/3 [Labg, Labh] (Blockchain ZK) - Plonky2/3	`add` contains branching occurring with a small probability (similar to `reduce128`). In theory leaks the input to the hash function, and we believe **SHOULD NOT** lead to a practical attack.	During the time of writing this work Plonky3 finished its security audit and became production ready. We did not find constant time implementation as a part of the threat model.

when considering the rising popularity of applications in the Web3 domain and the potential financial impact of any vulnerabilities. The first step in this direction would be to follow the Intel guidelines on mitigation against timing attacks [Inta, Intb]. This includes implementing any secret-dependent code in constant time. In the context to this work, we recommend switching to a constant-time Poseidon implementation as an alternative to Reinforced Concrete. While this can be challenging and may not result in the most performant implementation, it is important from a security perspective and should be treated as a minimum requirement for any protocol handling sensitive data. Furthermore, one may also adopt free and easy-to-use tools like the DATA framework [WZS+18] for conducting initial checks for any secret-dependent cache-based timing leakages.

Table 4: Runtime comparison between the constant time and non-constant time implementation of Poseidon-M and Poseidon-GL [Gra]. Benchmarked on Intel Core i5-12450H running WSL2.0 with Ubuntu 22.04.3 LTS.

Hash function	Non-constant time	Constant time	% Slower
Poseidon-(M, $t = 16$)	4.577 µs	4.996 µs	9%
Poseidon-(M, $t = 24$)	10.183 µs	10.929 µs	7%
Poseidon-(GL, $t = 8$)	1.997 µs	2.318 µs	16%
Poseidon-(GL, $t = 12$)	3.708 µs	4.057 µs	9%

It is well-known that constant-time implementations often do not translate into constant-time assembly code due to compiler optimizations which are often unpredictable. The leakages may also originate from other sources like described in [CWS+24] and often for better constant time guarantees, protocols, or at least the secret dependent components, can be directly written in `asm`. Moreover, we also believe that the security audits of the ZK libraries should also take into consideration the basic cache timing attacks in their security model. Perhaps the best practice might be to include the ZK library along with its underlying arithmetic libraries when auditing for security vulnerabilities. Additionally, it might also be helpful to state the supported compiler options and the CPU architectures. As a last resort, one can also switch to dedicated side-channel secure hardware for generating zero-knowledge proofs.

Disclosures. This work looks into the potential cache timing side-channel vulnerabilities in the current ZKP systems. Even though we do not show any real-world attack on the existing ZK applications, however, following our footsteps, in the future one may find real-world exploits, potentially causing financial loss and privacy violation, among other damages. We have already informed some of the industry groups about our findings, however, it is infeasible for us to inform every potentially affected group as the scope of our attack spreads across a wide range of ZK applications. The informed industry groups have already acknowledged our findings, and we hope mitigations are out soon.

Acknowledgements

Special thanks to Andreas Kogler[14] for all the side-channel discussions we had. Also thanks to Erik Kraft and Lukas Giner[15] for the discussions on specific aspects of the attacks. We

[14] https://andreaskogler.com/
[15] https://ginerlukas.com/

would also very much like to thank Roman Walch[16] and the members of TACEO and other groups for their useful feedback, especially on the current state of the art ZK scenario.

References

[AARR02] Dakshi Agrawal, Bruce Archambeault, Josyula R. Rao, and Pankaj Rohatgi. The EM side-channel(s). In Burton S. Kaliski Jr., Çetin Kaya Koç, and Christof Paar, editors, *Cryptographic Hardware and Embedded Systems - CHES 2002, 4th International Workshop, Redwood Shores, CA, USA, August 13-15, 2002, Revised Papers*, volume 2523 of *Lecture Notes in Computer Science*, pages 29–45. Springer, 2002.

[ABST23] Miguel Ambrona, Marc Beunardeau, Anne-Laure Schmitt, and Raphael R. Toledo. a*PlonK*: Aggregated *PlonK* from multi-polynomial commitment schemes. In Junji Shikata and Hiroki Kuzuno, editors, *Advances in Information and Computer Security - 18th International Workshop on Security, IWSEC 2023, Yokohama, Japan, August 29-31, 2023, Proceedings*, volume 14128 of *Lecture Notes in Computer Science*, pages 195–213. Springer, 2023.

[ACFY24] Gal Arnon, Alessandro Chiesa, Giacomo Fenzi, and Eylon Yogev. STIR: reed-solomon proximity testing with fewer queries. *IACR Cryptol. ePrint Arch.*, page 390, 2024.

[AES15] Gorka Irazoqui Apecechea, Thomas Eisenbarth, and Berk Sunar. S$a: A shared cache attack that works across cores and defies VM sandboxing - and its application to AES. In *2015 IEEE Symposium on Security and Privacy, SP 2015, San Jose, CA, USA, May 17-21, 2015*, pages 591–604. IEEE Computer Society, 2015.

[af] ark ff. ark-ff. https://docs.rs/ark-ff/0.4.2/ark_ff/.

[AHIV17] Scott Ames, Carmit Hazay, Yuval Ishai, and Muthuramakrishnan Venkitasubramaniam. Ligero: Lightweight sublinear arguments without a trusted setup. In Bhavani Thuraisingham, David Evans, Tal Malkin, and Dongyan Xu, editors, *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, CCS 2017, Dallas, TX, USA, October 30 - November 03, 2017*, pages 2087–2104. ACM, 2017.

[Azt] Aztec. Aztec. https://aztec.network/.

[BB03] David Brumley and Dan Boneh. Remote timing attacks are practical. In *Proceedings of the 12th USENIX Security Symposium, Washington, D.C., USA, August 4-8, 2003*. USENIX Association, 2003.

[BBHR18a] Eli Ben-Sasson, Iddo Bentov, Yinon Horesh, and Michael Riabzev. Fast Reed-Solomon Interactive Oracle Proofs of Proximity. In *ICALP*, volume 107 of *LIPIcs*, pages 14:1–14:17. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2018.

[BBHR18b] Eli Ben-Sasson, Iddo Bentov, Yinon Horesh, and Michael Riabzev. Scalable, transparent, and post-quantum secure computational integrity. *IACR Cryptol. ePrint Arch.*, page 46, 2018.

[16] https://rwalch.at/

[BBHR18c] Eli Ben-Sasson, Iddo Bentov, Yinon Horesh, and Michael Riabzev. Scalable, transparent, and post-quantum secure computational integrity. *IACR Cryptol. ePrint Arch.*, page 46, 2018.

[BBM+24] Carsten Baum, Ward Beullens, Shibam Mukherjee, Emmanuela Orsini, Sebastian Ramacher, Christian Rechberger, Lawrence Roy, and Peter Scholl. One tree to rule them all: Optimizing GGM trees and owfs for post-quantum signatures. *IACR Cryptol. ePrint Arch.*, page 490, 2024.

[BCCT12] Nir Bitansky, Ran Canetti, Alessandro Chiesa, and Eran Tromer. From extractable collision resistance to succinct non-interactive arguments of knowledge, and back again. In Shafi Goldwasser, editor, *Innovations in Theoretical Computer Science 2012, Cambridge, MA, USA, January 8-10, 2012*, pages 326–349. ACM, 2012.

[BCG+14] Eli Ben-Sasson, Alessandro Chiesa, Christina Garman, Matthew Green, Ian Miers, Eran Tromer, and Madars Virza. Zerocash: Decentralized anonymous payments from bitcoin. In *2014 IEEE Symposium on Security and Privacy, SP 2014, Berkeley, CA, USA, May 18-21, 2014*, pages 459–474. IEEE Computer Society, 2014.

[BCG+18] Jonathan Bootle, Andrea Cerulli, Jens Groth, Sune K. Jakobsen, and Mary Maller. Arya: Nearly linear-time zero-knowledge proofs for correct program execution. In Thomas Peyrin and Steven D. Galbraith, editors, *Advances in Cryptology - ASIACRYPT 2018 - 24th International Conference on the Theory and Application of Cryptology and Information Security, Brisbane, QLD, Australia, December 2-6, 2018, Proceedings, Part I*, volume 11272 of *Lecture Notes in Computer Science*, pages 595–626. Springer, 2018.

[BCR+19] Eli Ben-Sasson, Alessandro Chiesa, Michael Riabzev, Nicholas Spooner, Madars Virza, and Nicholas P. Ward. Aurora: Transparent succinct arguments for R1CS. In Yuval Ishai and Vincent Rijmen, editors, *Advances in Cryptology - EUROCRYPT 2019 - 38th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Darmstadt, Germany, May 19-23, 2019, Proceedings, Part I*, volume 11476 of *Lecture Notes in Computer Science*, pages 103–128. Springer, 2019.

[BCS16] Eli Ben-Sasson, Alessandro Chiesa, and Nicholas Spooner. Interactive Oracle Proofs. In *TCC (B2)*, volume 9986 of *Lecture Notes in Computer Science*, pages 31–60, 2016.

[BdM93] Josh Cohen Benaloh and Michael de Mare. One-way accumulators: A decentralized alternative to digital sinatures (extended abstract). In Tor Helleseth, editor, *Advances in Cryptology - EUROCRYPT '93, Workshop on the Theory and Application of of Cryptographic Techniques, Lofthus, Norway, May 23-27, 1993, Proceedings*, volume 765 of *Lecture Notes in Computer Science*, pages 274–285. Springer, 1993.

[BdSGK+21] Carsten Baum, Cyprien Delpech de Saint Guilhem, Daniel Kales, Emmanuela Orsini, Peter Scholl, and Greg Zaverucha. Banquet: Short and fast signatures from AES. In Juan A. Garay, editor, *Public-Key Cryptography - PKC 2021 - 24th IACR International Conference on Practice and Theory of Public Key Cryptography, Virtual Event, May 10-13, 2021, Proceedings, Part I*, volume 12710 of *Lecture Notes in Computer Science*, pages 266–297. Springer, 2021.

[Ber05] Daniel J Bernstein. Cache-timing attacks on aes. 2005.

[BFH+20] Rishabh Bhadauria, Zhiyong Fang, Carmit Hazay, Muthuramakrishnan Venkitasubramaniam, Tiancheng Xie, and Yupeng Zhang. Ligero++: A new optimized sublinear IOP. In Jay Ligatti, Xinming Ou, Jonathan Katz, and Giovanni Vigna, editors, *CCS '20: 2020 ACM SIGSAC Conference on Computer and Communications Security, Virtual Event, USA, November 9-13, 2020*, pages 2025–2038. ACM, 2020.

[BRB20] Tegan Brennan, Nicolás Rosner, and Tevfik Bultan. JIT leaks: Inducing timing side channels through just-in-time compilation. In *2020 IEEE Symposium on Security and Privacy, SP 2020, San Francisco, CA, USA, May 18-21, 2020*, pages 1207–1222. IEEE, 2020.

[CDG+17] Melissa Chase, David Derler, Steven Goldfeder, Claudio Orlandi, Sebastian Ramacher, Christian Rechberger, Daniel Slamanig, and Greg Zaverucha. Post-quantum zero-knowledge and signatures from symmetric-key primitives. In Bhavani Thuraisingham, David Evans, Tal Malkin, and Dongyan Xu, editors, *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, CCS 2017, Dallas, TX, USA, October 30 - November 03, 2017*, pages 1825–1842. ACM, 2017.

[CET+24] Stefanos Chaliasos, Jens Ernstberger, David Theodore, David Wong, Mohammad Jahanara, and Benjamin Livshits. Sok: What don't we know? understanding security vulnerabilities in snarks. In Davide Balzarotti and Wenyuan Xu, editors, *33rd USENIX Security Symposium, USENIX Security 2024, Philadelphia, PA, USA, August 14-16, 2024*. USENIX Association, 2024.

[CHM+20] Alessandro Chiesa, Yuncong Hu, Mary Maller, Pratyush Mishra, Psi Vesely, and Nicholas P. Ward. Marlin: Preprocessing zksnarks with universal and updatable SRS. In Anne Canteaut and Yuval Ishai, editors, *Advances in Cryptology - EUROCRYPT 2020 - 39th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Zagreb, Croatia, May 10-14, 2020, Proceedings, Part I*, volume 12105 of *Lecture Notes in Computer Science*, pages 738–768. Springer, 2020.

[Coi] Electric Coin. Halo2. https://zcash.github.io/halo2/index.html.

[CRR02] Suresh Chari, Josyula R. Rao, and Pankaj Rohatgi. Template attacks. In Burton S. Kaliski Jr., Çetin Kaya Koç, and Christof Paar, editors, *Cryptographic Hardware and Embedded Systems - CHES 2002, 4th International Workshop, Redwood Shores, CA, USA, August 13-15, 2002, Revised Papers*, volume 2523 of *Lecture Notes in Computer Science*, pages 13–28. Springer, 2002.

[CWS+24] Boru Chen, Yingchen Wang, Pradyumna Shome, Christopher W Fletcher, David Kohlbrenner, Riccardo Paccagnella, and Daniel Genkin. Gofetch: Breaking constant-time cryptographic implementations using data memory-dependent prefetchers. In *Proc. USENIX Secur. Symp*, pages 1–21, 2024.

[DKR+22] Christoph Dobraunig, Daniel Kales, Christian Rechberger, Markus Schofnegger, and Greg Zaverucha. Shorter signatures based on tailor-made minimalist symmetric-key crypto. In Heng Yin, Angelos Stavrou, Cas Cremers, and Elaine Shi, editors, *Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security, CCS 2022, Los Angeles, CA, USA, November 7-11, 2022*, pages 843–857. ACM, 2022.

[Dus] Dusk. Dusk network. https://dusk.network/.

[Eth] Ethereum. Ethereum: A secure decentralised generalised transaction ledger. ethereum project yellow paper. https://ethereum.github.io/yellowpaper/paper.pdf.

[Fab] Fabric. Fabric cryptography. https://www.fabriccryptography.com/.

[ff] ff. ff crate. https://docs.rs/ff/0.13.0/ff/.

[Fou] Nil Foundation. Nil;'s zkevm1. https://cms.nil.foundation/uploads/zk_EVM_1_7d6f8caa16.pdf.

[GCG+24] Lukas Giner, Roland Czerny, Christoph Gruber, Fabian Rauscher, Andreas Kogler, Daniel De Almeida Braga, and Daniel Gruss. Generic and automated drive-by GPU cache attacks from the browser. In Jianying Zhou, Tony Q. S. Quek, Debin Gao, and Alvaro Cardenas, editors, *Proceedings of the 19th ACM Asia Conference on Computer and Communications Security, ASIA CCS 2024, Singapore, July 1-5, 2024*. ACM, 2024.

[GKL+22] Lorenzo Grassi, Dmitry Khovratovich, Reinhard Lüftenegger, Christian Rechberger, Markus Schofnegger, and Roman Walch. Reinforced concrete: A fast hash function for verifiable computation. In Heng Yin, Angelos Stavrou, Cas Cremers, and Elaine Shi, editors, *Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security, CCS 2022, Los Angeles, CA, USA, November 7-11, 2022*, pages 1323–1335. ACM, 2022.

[GKL+23] Lorenzo Grassi, Dmitry Khovratovich, Reinhard Lüftenegger, Christian Rechberger, Markus Schofnegger, and Roman Walch. Hash functions monolith for ZK applications: May the speed of SHA-3 be with you. *IACR Cryptol. ePrint Arch.*, page 1025, 2023.

[GKR+21] Lorenzo Grassi, Dmitry Khovratovich, Christian Rechberger, Arnab Roy, and Markus Schofnegger. Poseidon: A new hash function for zero-knowledge proof systems. In Michael D. Bailey and Rachel Greenstadt, editors, *30th USENIX Security Symposium, USENIX Security 2021, August 11-13, 2021*, pages 519–535. USENIX Association, 2021.

[GLR+20] Lorenzo Grassi, Reinhard Lüftenegger, Christian Rechberger, Dragos Rotaru, and Markus Schofnegger. On a generalization of substitution-permutation networks: The HADES design strategy. In Anne Canteaut and Yuval Ishai, editors, *Advances in Cryptology - EUROCRYPT 2020 - 39th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Zagreb, Croatia, May 10-14, 2020, Proceedings, Part II*, volume 12106 of *Lecture Notes in Computer Science*, pages 674–704. Springer, 2020.

[GLS+23] Alexander Golovnev, Jonathan Lee, Srinath T. V. Setty, Justin Thaler, and Riad S. Wahby. Brakedown: Linear-time and field-agnostic snarks for R1CS. In Helena Handschuh and Anna Lysyanskaya, editors, *Advances in Cryptology - CRYPTO 2023 - 43rd Annual International Cryptology Conference, CRYPTO 2023, Santa Barbara, CA, USA, August 20-24, 2023, Proceedings, Part II*, volume 14082 of *Lecture Notes in Computer Science*, pages 193–226. Springer, 2023.

[GMM16] Daniel Gruss, Clémentine Maurice, and Stefan Mangard. Rowhammer.js: A remote software-induced fault attack in javascript. In Juan Caballero, Urko Zurutuza, and Ricardo J. Rodríguez, editors, *Detection of Intrusions*

and Malware, and Vulnerability Assessment - 13th International Conference, DIMVA 2016, San Sebastián, Spain, July 7-8, 2016, Proceedings, volume 9721 of *Lecture Notes in Computer Science*, pages 300–321. Springer, 2016.

[GMR85] Shafi Goldwasser, Silvio Micali, and Charles Rackoff. The knowledge complexity of interactive proof-systems (extended abstract). In Robert Sedgewick, editor, *Proceedings of the 17th Annual ACM Symposium on Theory of Computing, May 6-8, 1985, Providence, Rhode Island, USA*, pages 291–304. ACM, 1985.

[GMWM16] Daniel Gruss, Clémentine Maurice, Klaus Wagner, and Stefan Mangard. Flush+flush: A fast and stealthy cache attack. In Juan Caballero, Urko Zurutuza, and Ricardo J. Rodríguez, editors, *Detection of Intrusions and Malware, and Vulnerability Assessment - 13th International Conference, DIMVA 2016, San Sebastián, Spain, July 7-8, 2016, Proceedings*, volume 9721 of *Lecture Notes in Computer Science*, pages 279–299. Springer, 2016.

[Gra] IAIK TU Graz. Zk friendly hash zoo. https://extgit.iaik.tugraz.at/krypto/zkfriendlyhashzoo.

[GRBG18] Ben Gras, Kaveh Razavi, Herbert Bos, and Cristiano Giuffrida. Translation leak-aside buffer: Defeating cache side-channel protections with TLB attacks. In William Enck and Adrienne Porter Felt, editors, *27th USENIX Security Symposium, USENIX Security 2018, Baltimore, MD, USA, August 15-17, 2018*, pages 955–972. USENIX Association, 2018.

[Gro16] Jens Groth. On the size of pairing-based non-interactive arguments. In Marc Fischlin and Jean-Sébastien Coron, editors, *Advances in Cryptology - EUROCRYPT 2016 - 35th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Vienna, Austria, May 8-12, 2016, Proceedings, Part II*, volume 9666 of *Lecture Notes in Computer Science*, pages 305–326. Springer, 2016.

[GW20] Ariel Gabizon and Zachary J. Williamson. plookup: A simplified polynomial protocol for lookup tables. *IACR Cryptol. ePrint Arch.*, page 315, 2020.

[Hab22] Ulrich Haböck. Multivariate Lookups Based on Logarithmic Derivatives. *IACR Cryptol. ePrint Arch.*, page 1530, 2022.

[Her] Hermez. Polygon hermez. https://hermez.io/.

[HK24] Ulrich Haböck and Al Kindi. A note on adding zero-knowledge to STARKs. *IACR Cryptol. ePrint Arch.*, page 1037, 2024.

[HLP24] Ulrich Haböck, David Levit, and Shahar Papini. Circle starks. *IACR Cryptol. ePrint Arch.*, page 278, 2024.

[Idea] Iden3. iden3/circomlibjs git. https://github.com/iden3/circomlibjs.

[Ideb] Iden3. iden3/ffjavascript git. https://github.com/iden3/ffjavascript.

[IGI+16] Mehmet Sinan Inci, Berk Gülmezoglu, Gorka Irazoqui, Thomas Eisenbarth, and Berk Sunar. Cache attacks enable bulk key recovery on the cloud. In Benedikt Gierlichs and Axel Y. Poschmann, editors, *Cryptographic Hardware and Embedded Systems - CHES 2016 - 18th International Conference, Santa Barbara, CA, USA, August 17-19, 2016, Proceedings*, volume 9813 of *Lecture Notes in Computer Science*, pages 368–388. Springer, 2016.

[Inga] Ingonyama. Ingonyama. https://www.ingonyama.com/.

[Ingb] Ingonyama. ingonyama-zk/poseidon-hash git. https://github.com/ingonyama-zk/poseidon-hash.

[Inta] Intel. Guidelines for mitigating timing side channels against cryptographic implementations. https://www.intel.com/content/www/us/en/developer/articles/technical/software-security-guidance/secure-coding/mitigate-timing-side-channel-crypto-implementation.html.

[Intb] Intel. Security best practices for side channel resistance. https://www.intel.com/content/www/us/en/developer/articles/technical/software-security-guidance/secure-coding/security-best-practices-side-channel-resistance.html.

[KHF+20] Paul Kocher, Jann Horn, Anders Fogh, Daniel Genkin, Daniel Gruss, Werner Haas, Mike Hamburg, Moritz Lipp, Stefan Mangard, Thomas Prescher, Michael Schwarz, and Yuval Yarom. Spectre attacks: exploiting speculative execution. *Commun. ACM*, 63(7):93–101, 2020.

[KJJ99] Paul C. Kocher, Joshua Jaffe, and Benjamin Jun. Differential power analysis. In Michael J. Wiener, editor, *Advances in Cryptology - CRYPTO '99, 19th Annual International Cryptology Conference, Santa Barbara, California, USA, August 15-19, 1999, Proceedings*, volume 1666 of *Lecture Notes in Computer Science*, pages 388–397. Springer, 1999.

[Koc96] Paul C. Kocher. Timing attacks on implementations of diffie-hellman, rsa, dss, and other systems. In Neal Koblitz, editor, *Advances in Cryptology - CRYPTO '96, 16th Annual International Cryptology Conference, Santa Barbara, California, USA, August 18-22, 1996, Proceedings*, volume 1109 of *Lecture Notes in Computer Science*, pages 104–113. Springer, 1996.

[KZG10] Aniket Kate, Gregory M. Zaverucha, and Ian Goldberg. Constant-size commitments to polynomials and their applications. In Masayuki Abe, editor, *Advances in Cryptology - ASIACRYPT 2010 - 16th International Conference on the Theory and Application of Cryptology and Information Security, Singapore, December 5-9, 2010. Proceedings*, volume 6477 of *Lecture Notes in Computer Science*, pages 177–194. Springer, 2010.

[Laba] Lurk Lab. lurk-lab/neptune git. https://github.com/lurk-lab/neptune.

[Labb] Light Protocol Labs. Light protocol. https://docs.lightprotocol.com/.

[Labc] Matter Labs. Matter labs. https://matter-labs.io/.

[Labd] Matter Labs. matter-labs/rescue-poseidon git. https://github.com/matter-labs/rescue-poseidon.

[Labe] Mysten Labs. mystenlabs/sui git. https://github.com/MystenLabs/sui.

[Labf] O1 Labs. o1-labs/proof-systems git. https://github.com/o1-labs/proof-systems.

[Labg] Polygon Labs. 0xpolygonzero/plonky2 git. https://github.com/0xPolygonZero/plonky2.

[Labh] Polygon Labs. Plonky3/plonky3 git. https://github.com/Plonky3/Plonky3.

[Labi] Protocl Labs. Filecoin. https://filecoin.io/filecoin.pdf.

[Lee21] Jonathan Lee. Dory: Efficient, transparent arguments for generalised inner products and polynomial commitments. In Kobbi Nissim and Brent Waters, editors, *Theory of Cryptography - 19th International Conference, TCC 2021, Raleigh, NC, USA, November 8-11, 2021, Proceedings, Part II*, volume 13043 of *Lecture Notes in Computer Science*, pages 1–34. Springer, 2021.

[Lib] Libgcrypt. Libgcrypt. https://gnupg.org/software/libgcrypt/index.html.

[LKO+21] Moritz Lipp, Andreas Kogler, David Oswald, Michael Schwarz, Catherine Easdon, Claudio Canella, and Daniel Gruss. PLATYPUS: Software-based Power Side-Channel Attacks on x86. In *2021 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2021.

[Loo] Loopring. Loopring. https://loopring.org/#/.

[LSG+18] Moritz Lipp, Michael Schwarz, Daniel Gruss, Thomas Prescher, Werner Haas, Anders Fogh, Jann Horn, Stefan Mangard, Paul Kocher, Daniel Genkin, Yuval Yarom, and Mike Hamburg. Meltdown: Reading kernel memory from user space. In William Enck and Adrienne Porter Felt, editors, *27th USENIX Security Symposium, USENIX Security 2018, Baltimore, MD, USA, August 15-17, 2018*, pages 973–990. USENIX Association, 2018.

[Lur] Lurk. Lurk. https://lurk-lang.org/.

[LYG+15] Fangfei Liu, Yuval Yarom, Qian Ge, Gernot Heiser, and Ruby B. Lee. Last-level cache side-channel attacks are practical. In *2015 IEEE Symposium on Security and Privacy, SP 2015, San Jose, CA, USA, May 17-21, 2015*, pages 605–622. IEEE Computer Society, 2015.

[MBKM19] Mary Maller, Sean Bowe, Markulf Kohlweiss, and Sarah Meiklejohn. Sonic: Zero-knowledge snarks from linear-size universal and updatable structured reference strings. In Lorenzo Cavallaro, Johannes Kinder, XiaoFeng Wang, and Jonathan Katz, editors, *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security, CCS 2019, London, UK, November 11-15, 2019*, pages 2111–2128. ACM, 2019.

[MDS99] Thomas S. Messerges, Ezzy A. Dabbish, and Robert H. Sloan. Power analysis attacks of modular exponentiation in smartcards. In Çetin Kaya Koç and Christof Paar, editors, *Cryptographic Hardware and Embedded Systems, First International Workshop, CHES'99, Worcester, MA, USA, August 12-13, 1999, Proceedings*, volume 1717 of *Lecture Notes in Computer Science*, pages 144–157. Springer, 1999.

[mho] mhostetter. galois git. https://github.com/mhostetter/galois.

[Min] Mina. Mina protocol. https://docs.minaprotocol.com/assets/technicalWhitepaper.pdf.

[Mon] Monero. The monero project. https://www.getmonero.org/.

[MOP07] Stefan Mangard, Elisabeth Oswald, and Thomas Popp. *Power analysis attacks - revealing the secrets of smart cards*. Springer, 2007.

[MWES19] Ahmad Moghimi, Jan Wichelmann, Thomas Eisenbarth, and Berk Sunar. Memjam: A false dependency attack against constant-time crypto implementations. *Int. J. Parallel Program.*, 47(4):538–570, 2019.

[NSS] NSS. Nss. https://github.com/nss-dev/nss.

[Opea] OpenJDK. openjdk/jdk git. https://github.com/openjdk/jdk/tree/master.

[opeb] openSSL. openssl. https://www.openssl.org/.

[OST06] Dag Arne Osvik, Adi Shamir, and Eran Tromer. Cache attacks and countermeasures: The case of AES. In David Pointcheval, editor, *Topics in Cryptology - CT-RSA 2006, The Cryptographers' Track at the RSA Conference 2006, San Jose, CA, USA, February 13-17, 2006, Proceedings*, volume 3860 of *Lecture Notes in Computer Science*, pages 1–20. Springer, 2006.

[Pan] Panther. Panther protocol. https://www.pantherprotocol.io/.

[PFM+22] Luke Pearson, Joshua Brian Fitzgerald, Héctor Masip, Marta Bellés-Muñoz, and Jose Luis Muñoz-Tapia. Plonkup: Reconciling plonk with plookup. *IACR Cryptol. ePrint Arch.*, page 86, 2022.

[Pola] Polygon. Polygon miden technologies. https://polygon.technology/polygon-miden.

[Polb] Polygon. Polygon technologies. https://polygon.technology/.

[Proa] Light Protocol. Lightprotocol/light-poseidon git. https://github.com/Lightprotocol/light-poseidon.

[Prob] Light Protocol. Lightprotocol/light-protocol git. https://github.com/Lightprotocol/light-protocol.

[Proc] Panther Protocol. pantherprotocol/panther-core git. https://github.com/pantherprotocol/panther-core.

[RD20] Mark Randolph and William Diehl. Power side-channel attack analysis: A review of 20 years of study for the layman. *Cryptogr.*, 4(2):15, 2020.

[RIS] RISCZero. Risczero. https://www.risczero.com/.

[SAO+21] Anatoly Shusterman, Ayush Agarwal, Sioli O'Connell, Daniel Genkin, Yossi Oren, and Yuval Yarom. Prime+probe 1, javascript 0: Overcoming browser-based side-channel defenses. In Michael D. Bailey and Rachel Greenstadt, editors, *30th USENIX Security Symposium, USENIX Security 2021, August 11-13, 2021*, pages 2863–2880. USENIX Association, 2021.

[Scr] Scroll. Scroll. https://scroll.io/.

[Sem] Semaphore. Semaphore. https://semaphore.pse.dev/whitepaper-v1.pdf.

[Set20] Srinath T. V. Setty. Spartan: Efficient and general-purpose zksnarks without trusted setup. In Daniele Micciancio and Thomas Ristenpart, editors, *Advances in Cryptology - CRYPTO 2020 - 40th Annual International Cryptology Conference, CRYPTO 2020, Santa Barbara, CA, USA, August 17-21, 2020, Proceedings, Part III*, volume 12172 of *Lecture Notes in Computer Science*, pages 704–737. Springer, 2020.

[Sia] Sia. Sia. https://sia.tech/sia.pdf.

[SLST23] Alan Szepieniec, Alexander Lemmens, Jan Ferdinand Sauer, and Bobbin Threadbare. The tip5 hash function for recursive starks. *IACR Cryptol. ePrint Arch.*, page 107, 2023.

[Sol] Solana. Solana. https://solana.com/.

[Sov] Sovrin. Sovrin. https://sovrin.org/.

[Sta] Starkware. Starkware. https://starkware.co/.

[Sto] Storj. Storj. https://www.storj.io/storjv3.pdf.

[Sui] Sui. Sui. https://sui.io/zklogin.

[Sup] Supranational. Supra national. https://www.supranational.net/.

[Swa] Swarm. Swarm. https://www.ethswarm.org/swarm-whitepaper.pdf.

[SZB+24] Florian Sieck, Zhiyuan Zhang, Sebastian Berndt, Chitchanok Chuengsatiansup, Thomas Eisenbarth, and Yuval Yarom. Teejam: Sub-cache-line leakages strike back. *IACR Trans. Cryptogr. Hardw. Embed. Syst.*, 2024(1):457–500, 2024.

[TBP20] Florian Tramèr, Dan Boneh, and Kenny Paterson. Remote side-channel attacks on anonymous transactions. In Srdjan Capkun and Franziska Roesner, editors, *29th USENIX Security Symposium, USENIX Security 2020, August 12-14, 2020*, pages 2739–2756. USENIX Association, 2020.

[TOS10] Eran Tromer, Dag Arne Osvik, and Adi Shamir. Efficient cache attacks on aes, and countermeasures. *J. Cryptol.*, 23(1):37–71, 2010.

[Tus] Tusima. Tusima. https://www.tusima.network/.

[Vim] Vimwitch. poseidon-lite git. https://github.com/vimwitch/poseidon-lite.

[vSGBR18] Stephan van Schaik, Cristiano Giuffrida, Herbert Bos, and Kaveh Razavi. Malicious management unit: Why stopping cache attacks in software is harder than you think. In William Enck and Adrienne Porter Felt, editors, *27th USENIX Security Symposium, USENIX Security 2018, Baltimore, MD, USA, August 15-17, 2018*, pages 937–954. USENIX Association, 2018.

[Wit] Witnerfell. Winterfell git. https://github.com/facebook/winterfell/tree/main.

[WZS+18] Samuel Weiser, Andreas Zankl, Raphael Spreitzer, Katja Miller, Stefan Mangard, and Georg Sigl. DATA - differential address trace analysis: Finding address-based side-channels in binaries. In William Enck and Adrienne Porter Felt, editors, *27th USENIX Security Symposium, USENIX Security 2018, Baltimore, MD, USA, August 15-17, 2018*, pages 603–620. USENIX Association, 2018.

[XZS22] Tiancheng Xie, Yupeng Zhang, and Dawn Song. Orion: Zero knowledge proof with linear prover time. In Yevgeniy Dodis and Thomas Shrimpton, editors, *Advances in Cryptology - CRYPTO 2022 - 42nd Annual International Cryptology Conference, CRYPTO 2022, Santa Barbara, CA, USA, August 15-18, 2022, Proceedings, Part IV*, volume 13510 of *Lecture Notes in Computer Science*, pages 299–328. Springer, 2022.

[YF14] Yuval Yarom and Katrina Falkner. FLUSH+RELOAD: A high resolution, low noise, L3 cache side-channel attack. In Kevin Fu and Jaeyeon Jung, editors, *Proceedings of the 23rd USENIX Security Symposium, San Diego, CA, USA, August 20-22, 2014*, pages 719–732. USENIX Association, 2014.

[YGH16] Yuval Yarom, Daniel Genkin, and Nadia Heninger. Cachebleed: A timing attack on openssl constant time RSA. In Benedikt Gierlichs and Axel Y. Poschmann, editors, *Cryptographic Hardware and Embedded Systems - CHES 2016 - 18th International Conference, Santa Barbara, CA, USA, August 17-19, 2016, Proceedings*, volume 9813 of *Lecture Notes in Computer Science*, pages 346–367. Springer, 2016.

[Zca] Zcash. Zcash protocol specification. https://zips.z.cash/protocol/protocol.pdf.

[Zka] ZK-kit. Zk-kit. https://zkkit.pse.dev/.

[zkb] zk kit. zk-kit git. https://github.com/privacy-scaling-explorations/zk-kit.

[ZKs] ZKsync. Zksync. https://zksync.io/.