<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>Bypassing Parity Protected Cryptography using Laser Fault Injection in Cyber-Physical System</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td>He, Wei; Breier, Jakub; Bhasin, Shivam; Chattopadhyay, Anupam</td>
</tr>
<tr>
<td><strong>Date</strong></td>
<td>2016</td>
</tr>
<tr>
<td><strong>URL</strong></td>
<td><a href="http://hdl.handle.net/10220/40620">http://hdl.handle.net/10220/40620</a></td>
</tr>
<tr>
<td><strong>Rights</strong></td>
<td>© 2016 Association for Computing Machinery (ACM). This is the author created version of a work that has been peer reviewed and accepted for publication by Proceedings of the 2nd ACM International Workshop on Cyber-Physical System Security (CPSS 2016), Association for Computing Machinery (ACM). It incorporates referee's comments but changes resulting from the publishing process, such as copyediting, structural formatting, may not be reflected in this document. The published version is available at: [<a href="http://dx.doi.org/10.1145/2899015.2899019">http://dx.doi.org/10.1145/2899015.2899019</a>].</td>
</tr>
</tbody>
</table>
Bypassing Parity Protected Cryptography using Laser Fault Injection in Cyber-Physical System

Wei He, Jakub Breier, Shivam Bhasin
PACE, Temasek Laboratories
Nanyang Technological University, Singapore
{he.wei,jbreier,sbhasin}@ntu.edu.sg

Anupam Chattopadhyay
School of Computer Engineering
Nanyang Technological University, Singapore
anupam@ntu.edu.sg

ABSTRACT
Lightweight cryptography has been widely utilized in resource constrained embedded devices of Cyber-Physical System (CPS) terminals. The hostile and unattended environment in many scenarios make those endpoints easy to be attacked by hardware based techniques. As a resource-efficient countermeasure against Fault Attacks, parity Concurrent Error Detection (CED) is preferably integrated with security-critical algorithm in CPS terminals. The parity bit changes if an odd number of faults occur during the cipher execution. In this paper, we analyze the effectiveness of fault detection of a parity CED protected cipher (PRESENT) using laser fault injection. The experimental results show that the laser perturbation to encryption can easily flip an even number of data bits, where the faults cannot be detected by parity. Due to the similarity of different parity structures, our attack can bypass almost all parity protections in block ciphers. Some suggestions are given to enhance the security of parity implementations.

Keywords
Laser Fault Injection, Concurrent Error Detection (CED), Parity, Register Bit-Flip, FPGA, Cyber-Physical System

1. INTRODUCTION
As a highly hybrid functional systems, cyber-physical system (CPS) has recently gained immense importance in both academia and industry owing to its integration in heterogeneous mobile networks and vast number of functional embedded endpoints (sensors, smartphones, RFIDs, etc). CPS dynamically integrates IT and physical processes, where the embedded devices and networks monitor/control the physical processes, and meanwhile receiving the feedback from the physical processes, to form closed functional loop for fulfilling assigned tasks. The context of applications range from conventional industrial automation, Mass Rapid Transit (MRT) security/safety management, war zone monitor, to the emerging Internet of Things (IoTs), such as fuel management in aircraft, barrier detection system in autonomous vehicle, and smart grid, etc [9, 16]. All those are combinations of sensors to monitor real-time environmental parameters, to be preprocessed by embedded processors in nodes, relayed by coordinator, and analyzed in central node or server.

Security-critical data are gathered, transmitted, processed and stored in these distributed hardware devices, which makes CPS attractive targets for deliberate attacks/compromises from both cyber or hardware layers. CPS security research has primarily focussed on the cyber layer, whereas hardware layer’s vulnerabilities were neglected [22, 23, 24, 29]. However CPS is particularly vulnerable against attack directly on hardware level, such as side-channel attacks (SCA): power or EM based SCA [3, 17], hardware trojan horse (HTH) [27], and fault based attacks (FA) [7]), as sketched in Fig. 1. Substantial motivations arise for securing CPS hardware layers, as summarized in the following.

- In CPS, embedded intelligent devices in terminals are deployed remotely, portably and typically powered by batteries, so they are usually constructed using low-power/low-security devices considering the overall costs. Especially, the unattended embedded devices are convenient for adversary to perform close manipulations.
- CPS is basically an evolving concept that merges new disciplines as its expanded functional parts. Security...
strategies are hence required to be consistently upgraded in time. As a matter of fact, thorough security coverage to all the nodes, especially to the newly integrated and tiny sensor terminals, become challenging.

As one of the most prominent attack styles at hardware level, Fault Attack (FA) exploits the intentionally triggered faulty behaviors from the target devices, in order to restore the data secrets, or reverse-engineer the internal circuit structures. Faults can be injected using a variety of solutions, which can be broadly categorized as global and local. In global fault perturbation, disturbance is injected into certain global system which possibly causes unfriendly behavior or logic errors concurrently at multiple logic locations, such as the glitch on global clock [10] and power supply [25], or under-powering [15] and overclocking [2] the cipher execution. In contrast, local fault perturbation aims at precisely injecting faults into specific logic points, in accordance to the requirements by adversary. Laser or EM injectors with 2D/3D motorized stages are needed in this context.

Many techniques have been proposed against the fault attacks [5, 12, 18]. Restricted by the limited power supply and insufficient computation source in the remote nodes, expensive countermeasures cannot be applied. As one of the low-cost solutions, parity prediction realizes the fault detection using redundant parity bit, precomputed concurrently with the crypto algorithm and compared with the true round output. In case the faults are detected by the parity bit/bits, the cipher execution can be immediately halted.

A successful parity detection against the laser fault attacks heavily depends on how the adversary can possibly trigger odd number of faults in the target device. If the chance of triggering a single fault or odd number of faults is high, the parity scheme provides acceptable protections, and vice versa. However, previous work mostly assess the achieved security from a purely theoretical standpoint, i.e., the practical implementation situations in real devices are ignored. In the following work, we explore a real laser attack on FPGA and investigate the probabilities of triggering different number of register bit flips using the pulse laser, to claim the vulnerabilities of the previously proposed parity schemes.

The rest of the paper is organized as follows: Section 2 gives the principles of CED countermeasure and laser based fault attacks in embedded system. The laser fault perturbation technique, allowing us to perform sophisticated evaluation of register bit flips in FPGA is detailed in Section 3. In Section 4, we characterize the fault properties of a parity enhanced cipher attacked by pulse laser in Virtex-5 FPGA, with some further discussions. Section 5 draws the conclusions and perspectives of future work.

2. TECHNICAL BACKGROUND

2.1 PRESENT Lightweight Cipher

Differential fault attacks render many block ciphers vulnerable. For all such attacks, a parity-based countermeasure is suggested by Wu et al. [28]. In the following, we describe a representative lightweight block cipher and show that parity-based countermeasure fails against our attack. PRESENT [8] is a widely adopted lightweight block cipher that is suitable for resource constrained hardware environment, like many cyber-physical systems where sensors and physical devices are deployed remotely and powered by batteries. The PRESENT cipher from algorithmic level is shown in Fig. 2, which is constructed following the substitution-permutation network (SPN), The block size of PRESENT is 64 bits and the key size can be either 80 or 128 bits. The non-linear Sbox is 4 bits. The entire encryption/decryption consists of 31 rounds, and the last round key is used as a post-whitening key. This cipher is a part of the international standard ISO/IEC 29192-2:2012 [1].

Two fault attack models towards PRESENT-80 have been described by Bagheri et al. [4]. In the first model, a single bit fault is required to be injected into the intermediate states at the beginning of the last round Sbox layer, and comparatively a fault needs to be injected into one nibble rather than a bit in the second model. And both of them require the injection into Sbox layer.

This work explores the possibility of injecting one, two, three and four bit-flips into slice registers of 65 nm FPGA, targeting a parity protection scheme. It helps in evaluating the effectiveness of parity schemes against an attacker with strong and precise laser equipment.

![PRESENT cipher algorithmic diagram](image)

**Figure 2:** PRESENT cipher algorithmic diagram.

2.2 Parity Concurrent Error Detection

Parity is one of the simplest solutions when it comes to error detection. It is heavily used in communication systems. When implemented with block ciphers, it can detect ciphering abnormality during the en-/de-cryption with efficient implementation, which makes it a good candidate for small embedded systems with low computational power. Different parity schemes have been proposed in literatures aiming at covering possible faults as much as possible, while staying low cost [6, 20, 21, 28]. Generally, fault detection capability is achieved at expense of either time or space redundancies. In the time redundancy, errors are detected by repeating the ciphering. In this case, no extra logic is required, and the throughput is reduced by 50%. Comparatively, extra parity computations or dual-rail duplications are required to compare the parity bit/bits with the the output parity/parities in space redundancy. As a matter of fact, the first solution can only detect transient errors and the latter one is capable of detecting both transient and permanent errors. Moreover,
the proposed parity scheme normally requires less cost than the dual-rail duplication.

The parity techniques can be implemented by two general directions [11]:

- Parity-1: Only 1 parity bit is required for all the bits of datapath in each ciphering round of the cryptographic algorithm, e.g., one parity bit checks the errors for all the 128 data bits for AES-128 [28].
- Parity-n: n parity bits are employed, and each parity bit is responsible for the error checking of a single data word in the cryptographic algorithm, e.g., Parity-16 implements 1 parity bit per byte of AES-128 [20, 21].

Nevertheless, both of the above schemes can only detect odd number of faults occurred in the 128 data bits in Parity-1, or in the data word for Parity-n. In other words, if even number of faults appear, the schemes will be compromised. Another scheme proposed by Karpovsky et al. [13] provides wider fault coverage, relying on a prediction circuit comprised of linear predictor, linear compressor and cubic function. Despite the uniform detection of both odd and even number of faults, the circuit overhead is too high to be applied in resource constrained scenarios. Considering the implementation efficiency and the fact that a majority of fault models aim at single bit-flips, we hereby focus on the Parity-1 scheme in our work, and the conclusions directly apply to Parity-n as well.

2.3 Parity-1 Detection Scheme

A low cost parity scheme for detecting fault injection attacks on implementation of AES was proposed by Wu et al. [28]. When considering AES-128, this scheme has 8% area overhead and 5% time overhead. However, as claimed by Malkin et al. [18], even though it can detect random technical failures during the execution, attacker capabilities are usually much better when it comes to triggering deliberate faults.

There are several other works proposing the usage of parity for fault detection ([14, 19]), showing that despite not being able to detect more complex fault models, its simplicity still attracts attention. Fig. 3 depicts the parity detection scheme proposed in [28], targeting AES. Actually it is universal to all the Secret Key Cryptography (SKC) constructions by substitution-permutation networks. The round inputs are noted as $X$. The round key $K$ is added with $X$ to produce $Y$. The nonlinear substitution box (Sbox) substitutes $Y$ by $Z$. The linear diffusion layer permutes the bits of $Z$ to give $U$, that is actually the $X$ for the next round. This parity prediction mainly consists in three computations. For clarity, parity of a bit vector \( \cdot \) is denoted as $P(\cdot)$.

- In key addition, the round input parity $P(X)$ is XORed with round key parity $P(K)$ to get $P(Y)$.
- In Sbox, $P(Y)$ is non-linearly changed. Since Sbox in SPN cipher is fixed and public, the Sbox output parity $P(Z)$ can be precomputed. To check the sanity of the processed data, the input and output parity can be combined $P(Y) \oplus P(Z)$ and precomputed as the extension of a standard Sbox.
- The linear diffusion layer simply perpetuate the bits, so parity is not changed in this step.

As shown in parity computation path in Fig. 3.

$$P(Y) \oplus (P(Y) \oplus P(Z)) = P(Z) = P(out)$$ (1)

If no odd number of errors occur,

$$P(out) = P(U)$$ (2)

Otherwise,

$$P(Y) \oplus (P(Y^*) \oplus P(Z)) = P(out) \neq P(U)$$ (3)

where $P(Y^*)$ represents the error infected bit vector $Y$. Similarly to key addition part,

$$P(Y) = P(X) \oplus P(K)$$ (4)

Since $P(X)$ is the $P(out)$ of the previous round, we get,

$$P(Y) = P(out) \oplus P(K)$$ (5)

By checking if $P(Y)$ is equal with $P(out) \oplus P(K)$, we can detect the ciphering faults caused by odd number of bit errors, in key addition in current round, or errors in Sbox or linear diffusion layer in previous round.

Figure 3: Parity based concurrent error detection in SPN block cipher.

3. HIGH-PRECISION LASER PERTURBATION

3.1 Precision Necessity for Countering Parity Check

In many cyber-physical systems, the hostile and unattended working environment of the physical endpoints (sensors with embedded processors) provides good opportunity for the adversaries to perform stealthy and deliberate attacks. When it comes to practical fault attacks, one of the most precise techniques is the laser fault injection. This technique was used for failure analysis before and was introduced for disturbing cryptographic circuits by Skorobogatov et al. [26]. The laser impact to circuit can either be bit-set/bit-reset, or delaying the signal propagation in routings to cause set-up time violation [25]. Due to the requirements for most fault models towards cryptography, bit flip should be precisely injected into the desired single bit or multiple
bits at the specific computation point. The power supply and clock disturbances normally cause unpredictable multiple faults over the entire affected logic network. Hence, laser-induced fault attacks provide the option with most precise controllability.

To perform successful laser fault injection, at least the following parameters need to be determined. (1) the intensity and duration of laser pulse determine the energy perturbed on the target logic cell; (2) the time delay from the trigger to the laser shoot; (3) the location where the interesting logics are deployed on the chip; (4) the penetration depth of the used laser and the effective laser spot size. In our work, a Xilinx Virtex-5 FPGA is selected as the implementation device for the parity protected cryptography. The generic structure of Xilinx FPGA mainly consists of an array of fundamental logic cells, naming configurable logic block (CLB). In each CLB, there are two slices. Each slice contains 4 look-up-tables, multiplexers and flip-flops. In the implementation, 4 flip-flops in each slice can be configured as 4-bit registers, to store a nibble of round computation. Since the 4 flip-flops are pre-fabricated into a single slice, the actual distance are extremely small, typically within several hundreds nms in 65 nm devices. Restricted to the detection limitation, only odd number of faults, occurring in different bits, can incur parity change. To avoid the change in detection parity, adversaries need to upset either 2 or 4 bits simultaneously, for an odd-bit parity detection scheme, if a single laser injection is mounted to a specific slice in FPGA.

3.2 Experiment Preparation

The laser setup for our work is shown in Fig. 4. The embedded device is emulated by a Xilinx Virtex-5 FPGA in 65 nm technology with flip-chip package. Since the substrate of the die inside the package is placed up, the chip can be preprocessed by mechanical solution for reducing the substrate thickness from 300 μm to 100 μm, so as to effectively impact the active layer of the FPGA logic array. Note that a thinner residual substrate leads to easier laser penetration, however it takes risk of destroying active logic resource or routing channels on chip. The laser utilized is a 10 W diode pulse laser with 5x5 times magnification lens. The wavelength is 1064 nm and the spot size of the laser beam is around $60 \times 14 \, \mu\text{m}^2$. It is emphasized that only the very center part ($\approx 1/10$) of the beam spot is effective to impact the logic mainly because of the laser refraction and energy absorption through the substrate. However the LUT or slice dimensions are unknown, that can only be empirically tested. A PRESENT-80 coprocessor enhanced with parity based error detection scheme (see Fig. 5) is implemented inside the FPGA. We only compare the cipher output immediately after the key addition layer and the corresponding computed parity $P(Y)$. The comparison detects the odd number of errors occur in $\text{Sbox}$, $\text{pLayer}$ and $\text{addRoundKey}$ in each round. The area overhead of the parity protected $\text{Sbox}$ (prime target of the study) is shown in Table 1.

The test platform is set up over a laser station, with a 2D motorized stage for high-precision chip scanning. The system is driven by a compatible software on the PC for performing the laser perturbation, operating the chip scan and recording the cipher output. By scanning the PRESENT implemented region in FPGA and observing the faulty outputs, we have successfully positioned the location of the slice

| Table 1: Area overhead of parity protected Sbox. |
|---------|---------|---------|
| Sbox    | LUT     | Flip-Flops |
| Unprotected | 64     | 64      |
| Parity-1 Protected | 78     | 65      |
| Overhead            | 21.75% | 1.5%    |

Figure 4: High-precision laser injection platform.

where the 4 flip-flops of the least significant nibble of the 64 bit data block of PRESENT-80 reside.

4. CHARACTERIZATION OF LASER BIT-FLIPS TO PARITY-1 CED SCHEME

4.1 Laser Fault Injection in FPGA Slice

A fine-grained laser scan is conducted to the navigated slice region where registers of bit 0, 1, 2, 3 are situated (the least significant nibble of the 64 bit data block). The purpose is to check the possibility to disturb specific register bits in a nibble after the $\text{AddRoundKey}$ operation, as seen in Fig. 5. Without loss of generality, we have targeted the last round of the cipher, with the results stated below. Register area sensitive to faults is $\approx 16 \times 7 \, \mu\text{m}^2$ on the thinned substate. This is actually the exact location where the FPGA slice is placed, as visually figured out in Fig. 6 from the FPGA editor view. The laser spot ($60 \times 14 \, \mu\text{m}^2$) is larger than the aforementioned sensitive region, but our experiment certifies that the effective laser core (sufficiently powerful to trigger...
data bit events) inside the optical beam is much smaller than this area, that is roughly 1/10 of the physical laser spot. We used the full laser power available for such lens type (10 W), with variable glitch length between 250 – 300ns.

The result from attacking the last round is shown in Fig. 7. Totally, 4947 faults have been recorded out of 12,000 scanned points. As can be seen in the plot, majority of faults are 4-bit flips (i.e., the fault mask is 1111, as represented by blue spots in Fig. 7), resulting to unchanged parity. The % of each faulty mask from all the faults are given in Table 2. Remark that the we are targeting the least significant nibble, i.e., \{bit0, bit1, bit2, bit3\}, and the scan also reached the neighboring slice where 2nd least nibble \{bit4, bit5, bit6, bit7\} reside, so there are also some faults from those bit, see the most right spots in Fig. 7.

In total, 99.293% (1111:99.212% + 0101:0.081%) of all the faults are even bit flips in the target nibble, therefore the parity would stay unchanged with high probability, but the ciphertext would be faulty. Attacker would be able to mount a fault attack with either 2-bit or 4-bit fault model, using a selected subset of the faulty outputs. Given a precise laser navigation to the interesting slice, attack would even exclusively trigger the useful faulty outputs for the specific fault models. A single bit flip is triggered with an extremely low probability of 0.1%.

### 4.2 Discussion

Considering countermeasures, even parity is an obvious solution. However, it doesn’t solve the problem, it only shifts the problem to different parameters like choice of fault injection technique or injection strength etc. In our implementation, a single slice is targeted for the evaluation, where 4 registers are used as the round registers in the PRESENT-80 datapath. This does not incur any loss of generality since commercial placement and routing tools deploy the bits of the same bit-vector close to each other, which is true for both ASIC and FPGA. This is owing to the requirement of area and timing optimizations in late design phase. In our case, we noticed the 4 bits of a bit vector to be always located in the same slice. As a matter of fact, the multiple bit flipping in registers or similar logics by a single laser injection is practical for the commercial chips. To elevate the security level of parity protected cipher, an implementation suggestion is given: Despite our experiment show that even number of bits can be easily flipped using a single laser injection, fault models based on even number of faults as well exist, such as the nibble fault model described in [4]. So it is highly recommended to carefully investigate the multiple-bit fault models against a specific cipher, and swap the non corresponding bits of nibbles into different slices in FPGA or deploy them further with each other during the placement phase of the ASIC design.

### 5. CONCLUSIONS

Cyber-physical systems have recently prevailed in almost all security-critical infrastructures, where the extensive security research is of paramount importance. Despite the wide security attention on cyber level, CPS is still vulnerable due to the deliberate attacks on the underlying hardware devices of the embedded terminals, such as side-channel analysis, hardware trojan horse, timing attack, as well as fault injection attacks. As one of the countermeasures against fault attacks, parity based fault detection scheme stands out due to low performance overhead. In this paper, we perform the practical security evaluation of the parity detection against laser fault injection attacks on a parity CED enhanced lightweight PRESENT-80 cipher in FPGA. Relying on the sophisticated chip preparation and profiling, we successfully flipped multiple data register bits on the security-critical algorithmic points in slice. The results show that the parity scheme can be practically defeated by injecting even number of faults into ciphers. We demonstrated fault injection into even number of bits of a slice with a probability as
high as 99.293%. In addition, implementation suggestion is provided for mitigating the security pitfalls, which helps to elevate the security of the resource constrained CPS terminals against deliberate laser fault attacks.

In the future work, we plan to explore the hardware vulnerabilities of the standard IEEE 802.15.4 wireless protocol, and perform successful laser fault attacks to a commercial cyber-physical system using the mentioned communication module.

6. REFERENCES


