TECHNICAL SCIENCES
Abbrev.: Techn. Sc., No 15(1), Y 2012

SECURITY CHARACTERISATION OF A HARDENED AES CRYPTOSYSTEM USING A LASER

Cyril Roscian[1], *Florian Praden*[1,2], *Jean-Max Dutertre*[1], *Jacques Fournier*[2], *Assia Tria*[1,2]

[1] ENSM.SE – École Nationale Supérieure des Mines de St-Étienne
[2] CEA-LETI Minatec, Gardanne, France

K e y w o r d s: Advanced Encryption Standard (AES), fault attacks, laser, security characterization, DFA.

A b s t r a k t

The AES is a standard encryption algorithm used in numerous cryptographic systems like smart cards, TPMs as well as in protocols like WPA2 or OpenSSL. Measuring the robustness of AES implementations against physical attacks is of utmost import-ance in order to guarantee the security of the system into which the AES is used. In this article, we describe how a hardware AES, embedding countermeasures against physical attacks, has been characterized using a laser. With the latter, we tried to implement a class of physical attacks called fault attacks which, when successful, allows an attacker to retrieve the secret key used by the AES module. Our experiments have allowed us to validate the efficiency of some of the countermeasures implemented in this AES implementation and have given us hints on how to further improve them.

OKREŚLANIE POZIOMU BEZPIECZEŃSTWA SPRZĘTOWEGO MODUŁU AES Z WYKORZYSTANIEM LASERA

Cyril Roscian[1], *Florian Praden*[1,2], *Jean-Max Dutertre*[1], *Jacques Fournier*[2], *Assia Tria*[1,2]

[1] ENSM.SE – École Nationale Supérieure des Mines de St-Étienne
[2] CEA-LETI Minatec, Gardanne, France

S ł o w a k l u c z o w e: Advanced Encryption Standard (AES), wstrzykiwanie błędów, laser, badanie bezpieczeństwa, DFA.

A b s t r a k t

AES to standardowy algorytm szyfrowania stosowany w wielu systemach kryptograficznych, np. kartach elektronicznych, TPM, oraz takich protokołach, jak WPA2 czy OpenSSL. Pomiar odporności implementacji algorytmu AES na ataki fizyczne jest najważniejszy do zapewnienia bezpieczeństwa systemowi opartemu na AES. W artykule opisano, jak sprzętowa implementacja AES z wbudowanymi

zabezpieczeniami przeciwko fizycznym atakom była badana z wykorzystaniem lasera. Następnie podjęto próby zaimplementowania ataków fizycznych polegających na wstrzykiwaniu błędów; ataki te – zakończone sukcesem – pozwalają atakującemu na przechwycenie tajnego klucza wykorzystywanego w module AES. Przeprowadzone eksperymenty pozwoliły na sprawdzenie efektywności zabezpieczeń zaimplementowanych w module sprzętowym AES oraz wskazały możliwości dalszego podniesienia poziomu bezpieczeństwa.

Introduction

Security, through the authenticity, confidentiality and integrity of communication systems, has become an essential component of all electronic systems. The vulnerability to attacks of the devices implementing the cryptographic algorithms (such as smart cards) has become a critical issue. Some malicious means or "physical attacks" could be used to retrieve sensitive information such as encryption keys and therefore lower the security of the whole protected transmission chain of information. There are three types of physical attacks: *invasive* attacks, which cover all the techniques based on the modification (ANDERSON 1998) and probing (HANDSCHUH 1999, SCHMIDT 2009, GAMMEL 2010) of integrated circuits (IC) by an invasive method (KÖMMERLING 1999, KOEUNE 2005); *observation or passive* attacks which exploit the fact that some physical characteristics such as power consumption (KOCHER 1999, LU 2010), electromagnetic radiation or the duration of computation depend on the chip's internal calculations (KOEUNE 2005); *perturbation or fault* attacks, which are based on changing the environmental conditions of the chip to infer information about the internal state of the IC. The latter ones are the most complex to implement as various and complex parameters must be taken into account such as the timing (i.e. the synchronization between the injection time and the calculation), the localization and the power level.

In this paper we mainly discuss about security characterisations based on fault attacks. One of the goals of such attacks could be to reveal the secret keys of a cryptographic device based on techniques like Differential Fault Analysis (DFA) (BIHAM 1997, BONEH 1997). Several methods can be used to induce faults into cryptographic ICs: use of a laser (SKOROBOGATOV 2005, BAR-EL 2006), voltage pulses (BLÖMER 2003), clock glitches (AGOYAN 2010) or electromagnetic (EM) disturbances (SCHMIDT 2007). The use of lasers is one of the most effective techniques. Lasers allow a good reproducibility, an accurate control on the timing of the injection (the instant of firing and the duration of the pulse) and a precise focalization (the ability to restrain its effect to a limited number of gates). Consequently, lasers generate precise local effects into the IC thus leaving the rest of the chip "undisturbed". Several theoretical attacks against cryptographic algorithms are based on such models of fault injection methods (PIRET 2003, GIRAUD 2005, 2010, MORADI 2010). Although protections

against these kinds of attacks exist (YEN 2006), advanced methods combine semi-invasive attacks and power or EM analysis (MORADI 2011).

In this article, we describe the security assessment of an AES hardware chip done to validate the efficiency of the embedded countermeasures that could be incorporated into the AES module of the SECRICOM'S secure docking module (SDM). The SDM contains authentication keys and access rights associated with each user of the SECRICOM secure communication network. To access such keys, a secure channel, based on AES encryption, has to be established between the SDM and its host device (Trusted Docking Module). Hence, guaranteeing the resistance of the AES used against physical attacks helps in hardening the security chain of SECRICOM's communication infrastructure.

The outline of this paper is as follows. First, we recall the physical phenomena triggered when a laser is used to inject a fault into an IC. Then we describe the AES implementation used as target of our laser-based characterization. Third, we provide a description of the laser-based tests made on the targeted AES. With attacks like those described in (PIRET 2003, GIRAUD 2005) in mind, we tested the countermeasure in a "black box" approach (i.e. without using the knowledge we had of the implemented countermeasures). We also tested the AES in a "white box" approach where we tried to take advantage of the knowledge we had of the implemented countermeasures in order to circumvent them. Finally we provide a discussion about the results that we obtained and the conclusions that could be drawn from them.

Physical phenomena induced by lasers into an IC

When analysing the effect of a laser beam onto an IC, two phenomena have to be considered: the photoelectric effect appearing and the fact that some parts of the IC are more sensitive to the laser than others. When a laser beam strikes the Silicon and that the photon;s energy is higher than the Silicon's band gap, electron-hole pairs are created. In general, these pairs recombine and there is no effect on the IC. However under specific conditions, some undesired effects can appear. We shall focus on one of those effects called the Single Event Effect (SEE).

Single Event Effect

A Single Event Effect can appear when the electron-hole pairs created by the laser beam are drifted in opposite directions by the electrical field in the PN

junction instead of immediately recombining. The consequence of that is the creation of a transient current as illustrated in Figure 1.

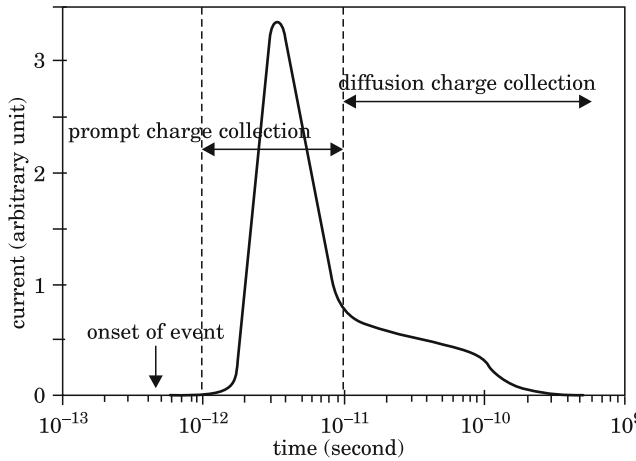

Fig. 1. Transient current resulting from charge collection after laser shoot

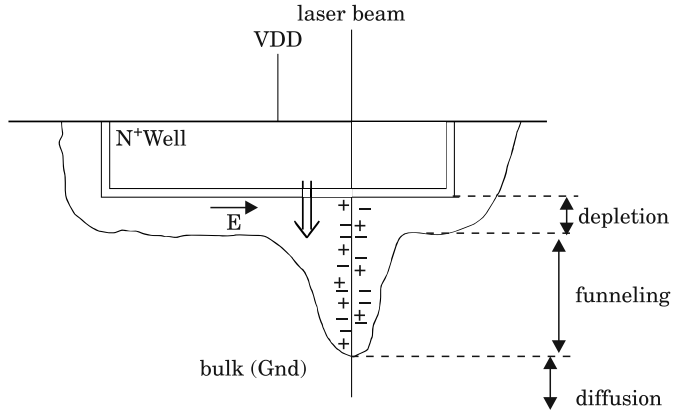

Fig. 2. Laser beam effect into the MOS structure

Along the laser beam shown in Figure 2, after the creation of the electron-hole pairs, two phenomena lead to the creation of the transient current. The first phenomenon stretches the depletion region (hence the extension of the electric field) along the laser beam and within a few picoseconds, charges are collected giving a current peak. In a second time, the rest of the charges are collected in a longer diffusing scheme. Figure 2 shows the transient current associated with the two phenomena of collection as given in (WANG 2008).

Sensitive zones

In CMOS technology, some parts are more sensitive than others to SEEs. To create an SEE, and then a fault, a strong electric field is needed. The reverse biased PN-junctions of the chip provide this required electric field. The position of these junctions can change, depending on the value of the data manipulated.

A good example to illustrate this data dependency is the CMOS inverter. The first case is a high state on the inverter's input: the NMOS transistor is in the "ON" state, its drain is grounded, the source and the bulk too and there is no reverse biased PN-junction. The PMOS transistor is in "OFF" state, then its source and the N well are in high potential, but its drain is grounded giving rise to a reverse biased junction. Hence, the drain of the PMOS transistor becomes sensitive to a laser shoot. In the same way, with a low state on the inverter's input, the drain of the NMOS transistor becomes sensitive to a laser shoot.

Figure 3(a) shows an inverter with a high state on its input and the "sensitive zone" is coloured in red. The second inverter (b) on the figure represents the other case of localisation of the "sensitive zone".

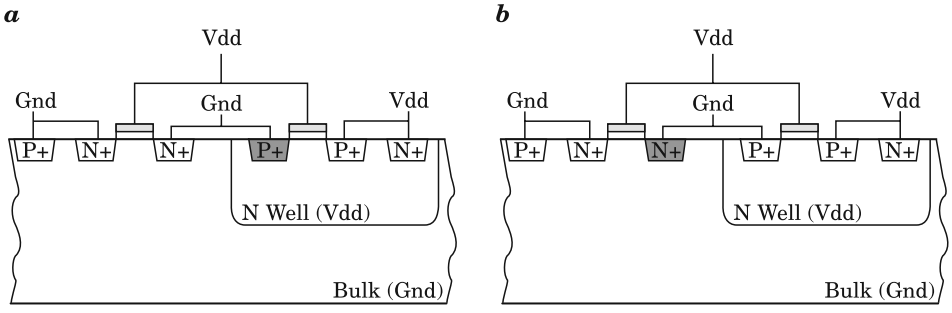

Fig. 3. CMOS inverter with a high state: *a* – or low state, *b* – on its input

From SEE to faults

We explained in the previous sections how a laser beam can create single events into CMOS structures and which parts of it are more sensitive to a laser. Even if an SEE is created by a laser, it is possible that the SEE has no effect on the chip's computations. An SEE can be transformed into a fault in two different ways. The first one is to generate an SEE directly into a register. In this case, the register's state is changed and this change is stored and propagated. The second way consists in creating, into the chip's logic, an SEE

which propagates through the logic up to the next register. Depending on the timing, if the SEE reaches the register's input on a clock's rising edge, an "faulty" value will be latched. Thus, a fault is injected into the chip's computations. Figure 4 illustrates the propagation of an SEE into the logic and the difficulty of transforming it into a fault. In the first case, the SEE generated in the logic is not captured by the D flip flop of the register cell and has no significant effect on the data processed. In the second case, with the adequate timing, the SEE is captured by the D flip flop. Thus, the value of the register is changed: the SEE has been turned into a fault.

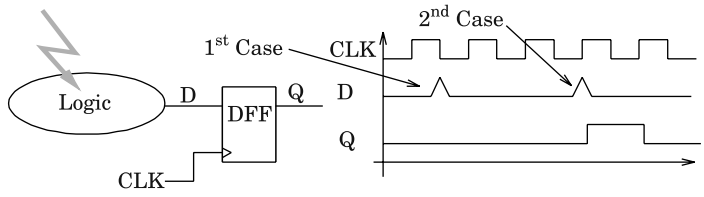

Fig. 4. Propagation of SEE through Logic

The hardened AES test chip

In this section, we describe the chip used as a device under test (DuT) for our laser-based characterizations. The DuT is a hardware module implementing the Advanced Encryption Standard (AES) algorithm used, for example, for encrypting the secure channel between a SECRICOM host device and the SECRICOM's SDM.

The AES algorithm

The AES algorithm is a symmetric key cryptography standard established by the NIST (NIST 2001). This algorithm is a substitution and permutation network based on four transformations (SubBytes, ShiftRows, MixColumns, AddRoundKey) used iteratively in rounds (Figure 5). In this paper, we focus on the 128-bit key version. This version processes data blocks of 128 bits, considered as matrices of 4x4 bytes called States, in ten rounds. The round keys (K1 to K10) used during every round are calculated by a key expansion routine (not detailed in this paper). We refer as M1 to M10 the States at the end of each round.

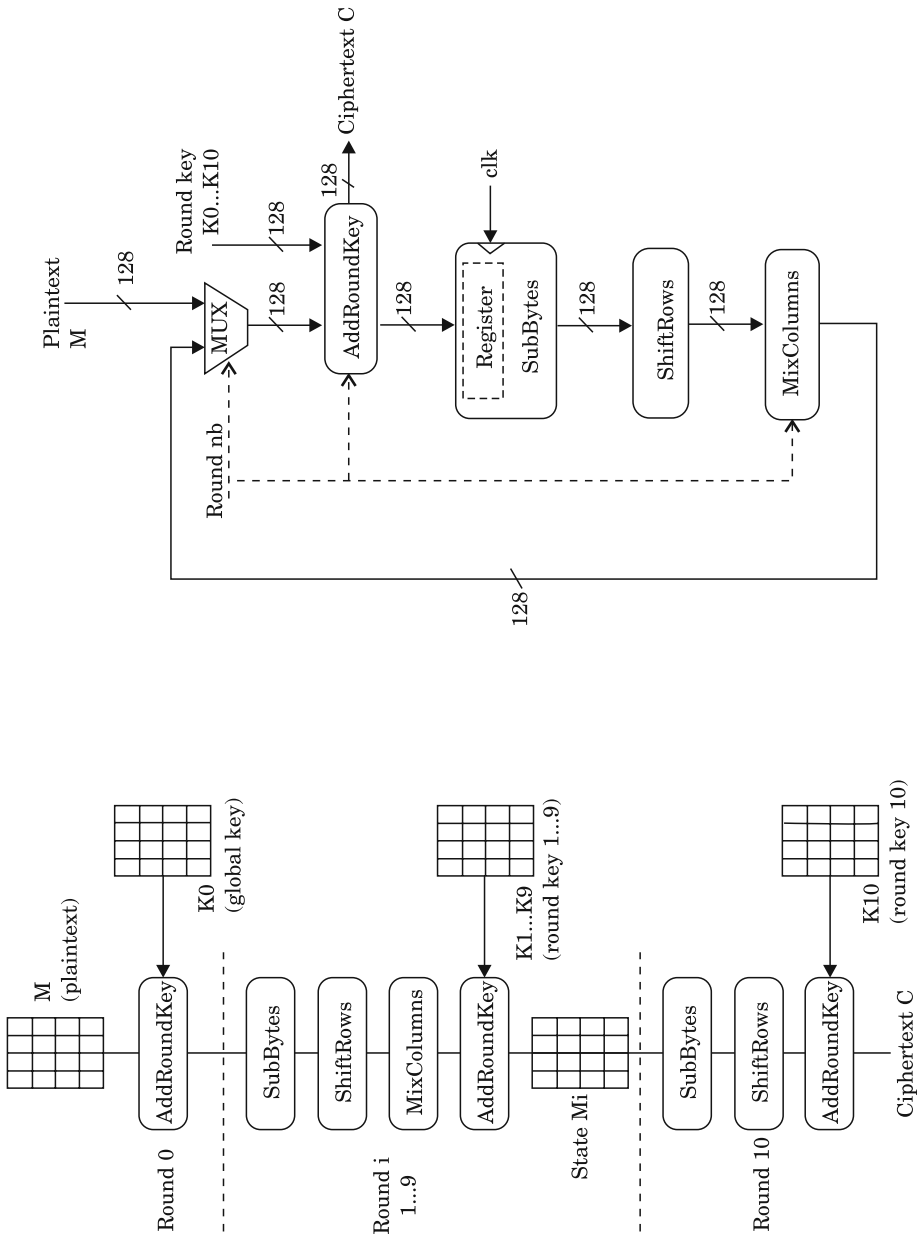

Fig. 5. The AES algorithm and its hardware implementation

The secure ASIC AES

In our study, we use the secure AES test chip described in (AGOYAN 2011) implemented in HCMOS9 gp 130 nm STM technology. The size of the die is 1336,m x 1411.8 μm and its working frequency is 25 MHz. A picture of the chip is shown in Figure 6.

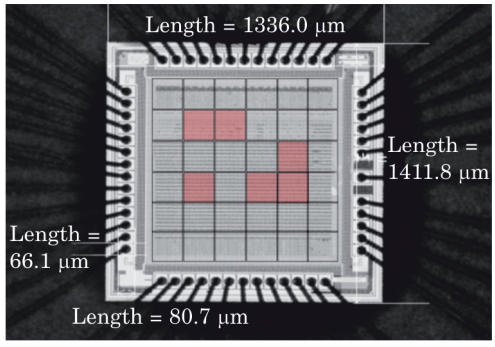

Fig. 6. Sensitive areas of the ASIC

The countermeasure against faults attacks implemented in this chip consists first in detecting errors and then in reacting in case of detection. The error detection is done by using spatial duplication: the AES is executed twice in parallel and at each round, the results of the two instances (the original path and the duplicated one) are compared. If an error is detected, the reaction consists in blurring the erroneous cipher text with a scrambled value of the detected error. The error detection mechanism is described in the following section.

Figure 7 illustrates the architecture of the implemented AES chip. We can see the two AES rounds executed in parallel with the error detection system.

The error detection mechanism

As mentioned in the previous section, when an error is detected (*XNOR* operation between the States from the two paths), the error detection mechanism scrambles the error value and then blurs the cipher text with the scrambled error. An error matrix is used whereby the error is spread across the rows and the columns as shown in the Figure 8. After that, the error matrix is XORed with the SubByte's results of the two data paths.

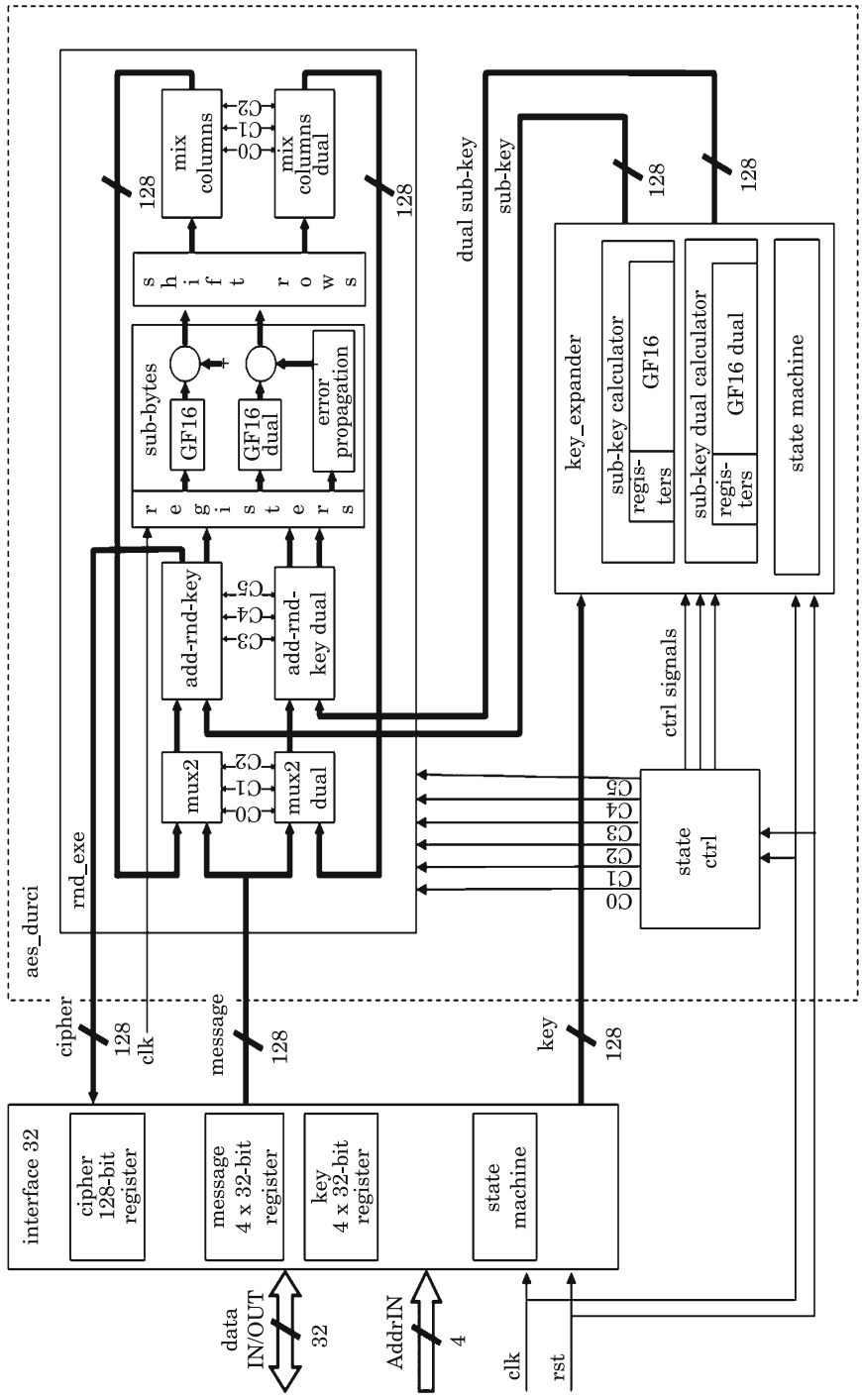

Fig. 7. Overview of the ASIC AES architecture

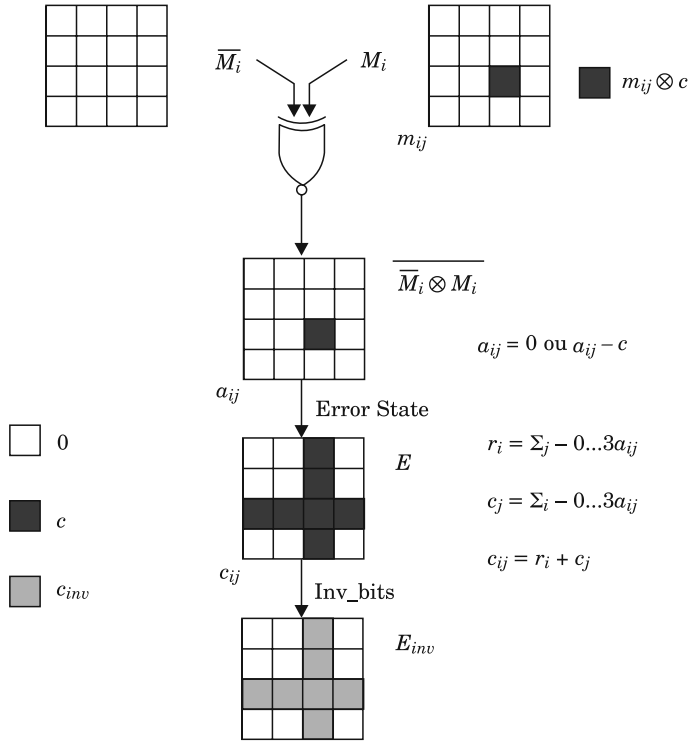

Fig. 8. The Error Detection and Spreading Mechanism

The Cross-ShiftRows operation

In addition to the above error detection and spreading mechanism, the ShiftRows operation is crossed between the two data paths. Half of the bits of each byte come from the other path and vice versa. This "crossing" operation is an additional security barrier to further scramble the error. If a fault is injected onto one of the paths, the fault is detected and propagated onto the two paths in parallel. With the Cross-ShiftRows, half of the information is lost due to the transfer to the other path. Figure 9 illustrates the injection of a fault on the last round of the AES and its propagation throughout the two data paths.

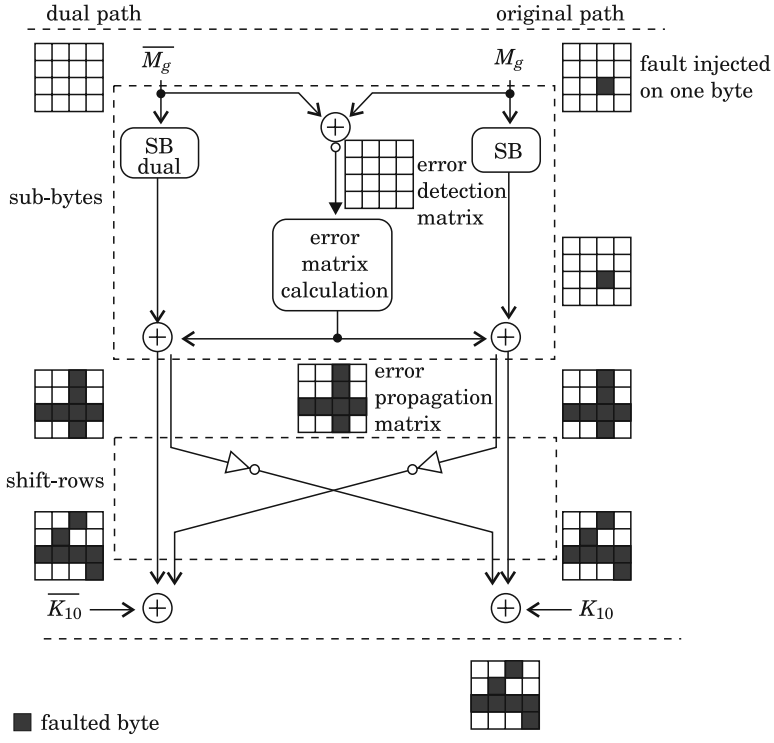

Fig. 9. Propagation of a fault on the last AES round (with the countermeasures)

Fault injection using laser

The Laser test bench

To perform the different tests on the AES chip, we used the laser facility of the MicroPackS™ platform (MICROPACKS 2012). The laser used is a YAG (Yttrium Aluminium Garnet) laser with three different sources: green, infrared and ultra-violet. We used the green source (with a wavelength of approximately 532 nm) with a 20× Mitutoyo lens. We obtained a spot size between 1 μm and 150 μm. With the largest spot size, we have approximately 15pJ of energy per laser shoot.

The AES is interfaced with a control PC. When an encryption is launched, a trigger signal is sent to an FPGA synchronization board, which sends a shoot signal to the laser after a delay defined by the control PC. This delay allows triggering the laser at different times during the encryption calculation.

We put ourselves into two configurations when doing those tests. In the first configuration we adopted a “black box” approach where we ignored any

of the implementation information we had on the DuT and tried to perform fault injections on the data path of the AES, in the "classical one", with the objective of collecting the erroneous cipher texts and doing differential cryptanalysis like in Giraud's or Piret's methods. In the second configuration, we used our knowledge of the implemented countermeasures, in a so-called "white box" approach, to try to circumvent the security mechanism by trying to fault the detection mechanism itself.

Fault injection on the data path

The easiest way to generate errors and trigger the detection mechanism is to inject faults into the register of the SubByte module. Despite the propagation of the error into the two data paths and the loss of half of the information due to the Cross-ShiftRows, we can always try to use the faulty cipher texts in Giraud's DFA (GIRAUD 2005). The *sine-qua-none* condition for this attack is to generate mono-bit faults (i.e. errors on only one bit of the State matrix). The error matrix can be found with a simple XOR operation between the correct cipher text and the faulty one. Figure 9 illustrates this kind of fault injection.

Fault injection on the detection mechanism

Another way of generating errors is to use the error detection mechanism itself. The DFA described in (PIRET 2003) needs a fault injection before the last MixColumns operation (Round 9). With our countermeasures and a fault injection on the data path, the faulty cipher texts cannot be exploited for this attack.

When looking closer at the Cross-ShiftRows operation, it appears that if the same fault is injected on the two data paths, the effect of the Cross-ShiftRows is "neutralized". Due to the dispersion of the lay-out of the two paths across the chip's surface, it's very hard to inject the same fault into the two paths with a laser which has a local effect. The solution is to inject the fault directly into the error matrix. By doing so, we could propagate the same error into the two paths and the Cross-ShiftRows will be "neutralized". In the last round, as the errors are the same on the two data paths, no detection appears and we have a faulty cipher text that could be used for DFA. Figure 10 depicts the propagation into the two paths of an error injected directly into the error matrix at the 9^{th} round.

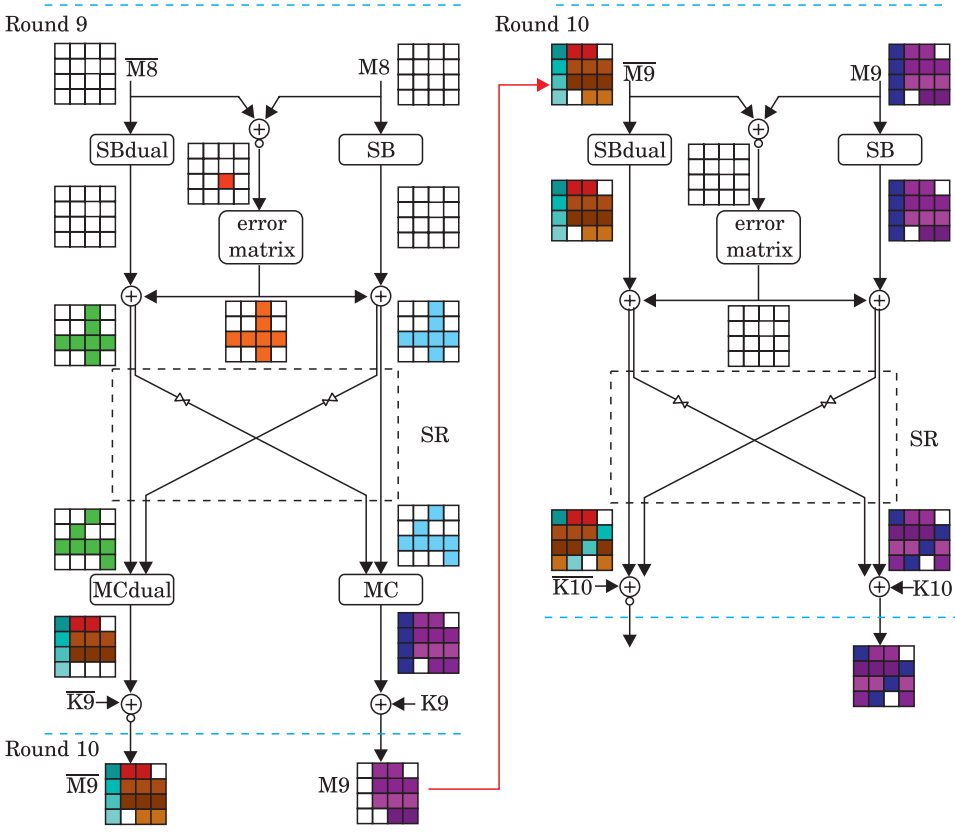

Fig. 10. Injection of an error into the error matrix and its propagation

Laser tests on the AES chip

Localization of the SubByte's registers

Since the registers of the SubByte are dispersed across the ASIC, the first step of the characterization work was to physically localize those registers. To do so, we made a cartography using the laser. The surface of the chip was partitioned into 36 zones of size 150 µm x 150 µm each as shown in Figure 6. For each zone, 50 different encryptions were performed among which we looked for specific faulty cipher texts: as shown in Figure 9, if a fault is injected on one byte at the beginning of the last round, in the end, we obtain the same error in six bytes of the resulting cipher matrix, due to the propagation mechanism, and one different error at the position of the injected fault. In Figure 6, the coloured regions highlight those where such

specific faulty cipher texts were obtained and which appeared to be more sensitive to this type of fault on the last round of encryption.

Results

Once the SubByte's registers have been localized, we started injecting faults on the data path. In the "black box" approach, we tried to perform DFA as described in (GIRAUD 2005) using only the faulty and correct cipher texts but in vain: the detection and error spreading mechanism proved to be efficient against such attacks.

However, when we used the knowledge of the implementation of the countermeasure (i.e. in a "white box" approach), especially the structure of the Cross-ShiftRows, we managed to recover a few bytes of the secret key of the AES. The complete knowledge of the Cross-ShiftRows is necessary because half of the information is lost in this operation and we need, for the attack, to keep all the information.

We also tried to inject faults directly into the error matrix in order to try another type of DFA (PIRET 2003) where in theory two faulty cipher texts are needed to recover 4 bytes of the secret key. To find all the 16 bytes of the key, we need to inject an error in one of the bytes of each column of the error matrix. Despite our efforts, we couldn't inject any fault into the error matrix with our laser test bench. One of the reasons for this is that the error matrix is not implemented using registers but with logic gates. Thus it is very hard to synchronize the laser shoot with the ASIC's encryption precisely enough to target separately each column of the error matrix.

Discussion and Conclusion

In this paper, we have described how countermeasures implemented in a hardware implementation of the AES have been tested using a laser as fault injection means. We have seen that, in a "black box" approach, classical DFA techniques are inefficient against such countermeasures. However our characterization work has also shown that in a "white box" scenario, some bytes of the secret keys could be recovered. This has led us to the conclusion that the error propagation should have been truly random and independent from the generated errors (requiring the implementation of a True Random Number Generator in the chip). We also investigated another attack path by trying to inject a fault in the error matrix itself but this has been unsuccessful illustrating the limits of current equipment with respect to current technolo-

gies. Such characterization works have provided valuable design rules for implementing secure encryption AES modules like those used in the SDM of the SECRICOM project.

Acknowledgements

This work was funded by the SECRICOM project (EC FP7-SEC-2007 grant 218123). The research work of Cyril Roscian was partly funded by the "Conseil Regional Provence-Alpes Côtes d'Azur".

Translated by AUTHORS

Accepted for print 30.06.2012

References

AGOYAN M., DUTERTRE J-M., NACCACHE D, ROBISSON B., TRIA A. 2010. *When Clocks Fail: On Critical Paths and Clock Faults*. SPRINGER VERLAG ed. Smart Card Research and Advanced Application.

AGOYAN M., BOUSQUET S., DUTERTRE J-Max., FOURNIER J., RIGAUD J-B., ROBISSON B., TRIA A. 2011. *Design and characterisation of an AES chip embedding countermeasures*. International Journal of Intelligent Engineering Informatics, 1, 3–4: 328–347.

AMIEL F., CLAVIER C., TUNSTALL M. *Collision fault analysis of DPA resistant algorithms*. In the proceedings of Fault Diagnosis ąand Tolerance in Cryptography 2006 – FDTC 2006.

ANDERSON R.J., KUHN M.G. 1998. *Low Cost Attacks on Tamper Resistant Devices*. In the Proceedings of the 5th International Workshop on Security Protocols.

BAR-EL H., CHOUKRI, H., NACCACHE D, TUNSTALL M., WHELAN C. 2004. *The Sorcerer's Apprentice Guide to Fault Attacks*. E-Print: 100.

BIHAM E., SHAMIR A. 1997. *Differential Fault Analysis of Secret Key Cryptosystems*. In the proceedings of the 17th Annual International Cryptology Conference on Advances in Cryptology.

BLÖMER J., SEIFERT J. 2003. *Fault Based Cryptanalysis of the Advanced Encryption Standard (AES)*. In the proceedings of Financial Cryptography.

BONEH D., DEMILLO R.A., LIPTON R.J. 1997. *On the Importance of Checking Cryptographic Protocols for Faults*. Advances in Cryptology – EUROCRYPT '97, International Conference on the Theory and Application of Cryptographic Techniques, Konstanz, Germany, May 11–15.

GAMMEL B.M., MANGARD S. 2010. *On the Duality of Probing and Fault Attacks*. J. Electron. Test., 26(4): 483-493 ISSN 0923-8174. DOI 10.1007/s10836-010-5160-0.

GIRAUD C. 2005. *DFA on AES*. In the proceedings of the 4th international conference on Advanced Encryption Standard. Bonn, Germany.

GIRAUD C., THILLARD A. 2010. *Piret and Quisquater's DFA on AES Revisited*. E-print: 440.

HANDSCHUH H., PAILLIER P., STERN J. 1999. *Probing Attacks on Tamper-Resistant Devices*. In the Proceedings of the First International Workshop on Cryptographic Hardware and Embedded Systems.

KIM C.H., QUISQUATER J-J. 2008. *New Differential Fault Analysis on AES Key Schedule: Two Faults Are Enough*. In the proceedings of the 8th IFIP WG 8.8/11.2 international conference on Smart Card ąResearch and Advanced Applications. London, UK.

KOCHER P.C., JAFFE J., JUN B. 1999. *Differential Power Analysis*. In the proceedings of CRYPTO.

KOEUNE F., STANDAERT F. *A Tutorial on Physical Security and Side-Channel Attacks*. In Foundations of Security Analysis and Design III: FOSAD 2004/2005, Nov 2006, 3655, 78–108

KÖMMERLING O., KUHN M.G. 1999. *Design principles for tamper-resistant smartcard processors*. In the Proceedings of the USENIX Workshop on Smartcard Technology on USENIX Workshop on Smartcard Technology. Chicago, Illinois.

Lu J., Pan J., Den Hartog J. 2010. *Principles on the security of AES against first and second-order differential power analysis.* In the Proceedings of the 8th international conference on Applied cryptography and network security. Beijing, China.

Micropacks. http://www.arcsis.org, last accessed 19th of April 2012.

Moradi A., Mischke O., Paar C., Li Y., Ohta K., Sakiyama K. 2011. *On the power of fault sensitivity analysis and collision side-channel attacks in a ącombined setting.* In the proceedings of the 13th international conference on Cryptographic hardware ąand embedded systems. Nara, Japan.

Moradi A., Shalmani M.T.M., Salmasizadeh M. 2006. *A generalized method of differential fault attack against AES cryptosystem.* In the Proceedings of the 8th international conference on Cryptographic Hardware and Embedded Systems. Yokohama, Japan.

Mukhopadhyay D. 2009. *An Improved Fault Based Attack of the Advanced Encryption Standard.* In the Proceedings of the 2nd International Conference on Cryptology in Africa: Progress in Cryptology. Gammarth, Tunisia.

NIST, National Institute of Standards and Technology. 2001. *Announcing the advanced encryption standard (AES)*, Federal Inf. Processing Standards Pub., Vol. 197.

Dusart P., Letourneux G., Vivolo O. 2003. *Differential Fault Analysis on A.E.S,* E-print: 010.

Piret G., Quisquater J-J. 2003. *A Differential Fault Attack Technique Against SPN Structures, with Application to the AES and KHAZAD.* In the proceedings of the 5th international conference on Cryptographic hardware and embedded systems, LNCS 2779.

Schmidt J., Hutter M. 2007. *Optical and EM Fault-Attacks on CRT-based RSA: Concrete Results.* Ed. Austrochip 2007, 15th Austrian Workhop on Microelectronics, 11 October 2007, Graz, Austria, Proceedings.

Schmidt J., Kim C.H. 2009. *Information Security Applications*. Chung K., Sohn K., Yung M. eds., Berlin, Heidelberg: Springer-Verlag, pp. 256-265 ISBN 978-3-642-00305-9. DOI 10.1007/978-3-642-00306-6–19.

Skorobogatov S.P. 2005. *Semi-Invasive Attacks – A New Approach to Hardware Security Analysis*. PhD thesis, University of Cambridge, Computer Laboratory.

Takahashi J., Fukunaga T. 2007. *Differential Fault Analysis on the AES Key Schedule*. E-print: 480.

Trichina E. 2003. *Combinational Logic Design for AES SubByte Transformation on Masked Data*. E-print: 236.

Tunstall M., Mukhopadhyay D., Ali S. 2011. *Differential fault analysis of the advanced encryption standard using a single fault.* In the Proceedings of the 5th IFIP WG 11.2 international conference on Information security theory and practice: security and privacy of mobile devices in wireless communication. Heraklion, Crete, Greece.

Wang F., Agrawal V.D. 2008. *Single Event Upset: An Embedded Tutorial.* Proc. of 21st International Conference on VLSI Design.

Yen C., Wu B. 2006. *Simple Error Detection Methods for Hardware Implementation of Advanced Encryption Standard.* IEEE Trans.Comput., jun, 55(6): 720–731 ISSN 0018-9340. DOI 10.1109/TC.2006.90.