# Differential Fault Analysis on A.E.S. 

P. Dusart, G. Letourneux, O. Vivolo

Rapport de recherche ${ }^{\circ}$ 2003-01

# Differential Fault Analysis on A.E.S. 

P. Dusart, G. Letourneux ${ }^{\dagger}$ O. Vivolo ${ }^{\ddagger}$<br>01/10/2002


#### Abstract

We explain how a differential fault analysis (DFA) works on AES 128, 192 or 256 bits.


## Contents

1 Introduction ..... 1
2 The description of the AES ..... 2
2.1 Representation chosen for $G F\left(2^{8}\right)$ ..... 2
2.2 Notations used in this article ..... 3
2.3 AddRoundKey for $i^{\text {th }}$ round ..... 3
2.4 SubByte for $i^{\text {th }}$ round ..... 3
2.5 MixColumn for $i^{\text {th }}$ round ..... 3
2.6 ShiftRows for $i^{\text {th }}$ round ..... 4
3 The description of the attack on computation of AES ..... 4
3.1 Principle of the attack ..... 4
3.2 Example ..... 4
3.3 How the injected error acts on the final state ..... 5
3.3.1 Fault modification ..... 5
3.3.2 Effect on MixColumn ..... 5
3.3.3 Effect on AddRoundKey ..... 5
3.3.4 Effect on last SubBytes ..... 5
3.3.5 Effect after last ShiftRows ..... 5
3.3.6 Effect after last AddRoundKey ..... 5
3.4 Example ..... 6
3.5 Analysis on information brought by fault ..... 6
3.6 Example ..... 8
4 Generalisation ..... 9
4.1 Without fault location ..... 9
4.2 Hardware Device ..... 9
A Back to initial key with the last subkey ..... 9

## 1 Introduction

In September 1996, Boneh, Demillo, and Lipton [4] from Bellcore announced a new type of cryptanalytic attack which exploits computational errors to find cryptographic keys. Their attack is applicable to public key cryptosystems such as RSA, excluding secret key algorithms. In [3], E. Biham \& A.Shamir extend this attack to various secret key cryptosystems such as DES, and call it Differential Fault Analysis (DFA). They applied the differential cryptanalysis to Data Encryption Standard (DES) in case of hardware fault model.

We further assume that the attacker is in physical possession of the tamperproof-device, so that he can repeat the experiment with the same cleartext and key but without applying the external physical effects. As a result,

[^0]he obtains two ciphertexts derived from the same (unknown) cleartext and key, where one of the ciphertexts is correct and the other is the result of a computation corrupted by a single error during the computation.

The DES used a 56 -bits key which seems to be too short for future. Hence a mondial competition between secret key cryptosystems has been realized. The requirements of this standard is to replace DES standard: a symmetric cryptosystem with 128 to 256 key sizes, which can be easily implemented in hardware. On Oct. 2, 2000, NIST choose Rijndael to be the Advanced Encryption Standard (AES). AES uses a 128, 192 or 256 bits key with a 128 bits input message. It works on bytes with an algebraic structure which is the finite field $\mathrm{GF}\left(2^{8}\right)$. Rijndael has been choosed by NIST for its resistance to linear and differential cryptanalysis [5].

The major critique of DFA was the practical feasibility of the theory. But some authors [2] have designed practical experimentations of this kind of attack with the possibility to inject the fault in a temporal windows which can be clearly related with program running process. By exposing a sealed tamperproof device such as a smartcard to certain physical effects (e.g., ionizing or microwave radiation), one can induce with reasonable probability a fault at a short random bit location in one of the registers at some intermediate stage in the cryptographic computation. In practice, the perturbation can change more than one bit. We assume that it can change up to one byte anywhere between the last two MixColumn operations of AES.

For DFA on DES, the attacker knows the differential input and output of the touched SBox. For AES, contrary to DES, we don't have the value of the differential fault $\varepsilon$ which could be obtained by considering the left part of the final DES state at round 16. For AES, if we consider a single fault before the SubBytes transformation, we can't go back to the key (There are 127 possibilities of the injected fault and 256 possibilities of a single byte of the round key, so the AES is protected against classical differential analysis.).

When the injected fault is becoming several induced faults (at least two) occuring in different bytes of the state, we can intersect each set of possible induced faults (the cardinal of intersection is lower than 63) and so we find a set of possible values (at most 128) for several bytes of the last subkey.

Further we find the last subkey with enough pairs of correct cipher/fault cipher. Once known this subkey is, we can find easily the key. For the sake of simplicity, we first assume that the first byte of the state before the MixColumn transformation of the nine round is replaced by a unknown value. The induced fault is going to be propagated by the MixColumn and spread over four bytes of the state. There is a linear relation between the four induced faults. For each byte is possible to find a set of possible value of induced fault, and then a set of possible values for the roundkey 10 .

In this paper, we show that AES is sensitive to Fault Analysis. We have implemented this attack on a personal computer. Our analysis program found the full AES-128 key by analysing less than 50 ciphertexts.

## 2 The description of the AES

In this article, we use a description slightly different from the original AES submission FIPS PUB 197 [1]. We describe AES using matrix on $G F\left(2^{8}\right)$ but we try to keep the notations of [1].

The AES is a block cipher with block length to 128 bits, and support key lengths $N_{k}$ of 128,192 or 256 bits. The AES is a key-iterated block cipher : it consists of the repeated application of a round transformation on the state. The number of rounds is denoted $N_{r}$ and depends on the key length ( $N r=10$ for $128 \mathrm{bits}, N r=12$ for 192 bits and $N_{r}=14$ for 256 bits).

The AES transforms a state, noted $S \in M_{4}\left(G F\left(2^{8}\right)\right.$ ), (i.e. $S$ is a matrix 4 x 4 with its coefficients in $\left.G F\left(2^{8}\right)\right)$ to another state in $M_{4}\left(G F\left(2^{8}\right)\right)$. The key $K$ is expanded into $N_{r}+1$ subkeys noted $K_{i} \in M_{4}\left(G F\left(2^{8}\right)\right)$ $\left(i=0,1, \ldots, N_{r}\right)$.

A round of an encryption with AES is composed of four main operations :

1. AddRoundKey
2. MixColumn
3. SubBytes
4. ShiftRows

### 2.1 Representation chosen for $G F\left(2^{8}\right)$

The representation chosen in [1] of $G F\left(2^{8}\right)$ is $G F(2)[X] /<m>$, where $<m>$ is the ideal generated by the irreducible polynomial $m \in G F(2)[X], m=x^{8}+x^{4}+x^{3}+x+1$.

### 2.2 Notations used in this article

We use four notations for representing an element in $G F\left(2^{8}\right)$, which are equivalent to one another:

1. $x^{7}+x^{6}+x^{4}+x^{2}$, the polynomial notation
2. $\{11010100\}_{b}$, the binary notation
3. 'D4', the hexadecimal notation
4. 212 , the decimal notation

### 2.3 AddRoundKey for $i^{\text {th }}$ round

The AddRoundKey transformation consists of an addition of matrix in $M_{4}\left(G F\left(2^{8}\right)\right)$ between the state and the subkey of the $i^{\text {th }}$ round. We denote by $S_{i, A}$ the state after the $i^{t h}$ AddRoundKey.

$$
\begin{aligned}
M_{4}\left(G F\left(2^{8}\right)\right) & \longrightarrow M_{4}\left(G F\left(2^{8}\right)\right) \\
S & \longmapsto S_{i, A}=S+K_{i}
\end{aligned}
$$

### 2.4 SubByte for $i^{\text {th }}$ round

The SubByte transformation consists in applying on each element of the matrix $S$ an elementary transformation $s$. We denote by $S_{i, S u}$ the state after the $i^{t h}$ SubByte.

$$
S=\left(\right) \quad \longmapsto \quad S_{i, S u}=\left(\begin{array}{cccc}
s(S[1]) & s(S[5]) & s(S[9]) & s(S[13]) \\
s(S[2]) & s(S[6]) & s(S[10]) & s(S[14]) \\
s(S[3]) & s(S[7]) & s(S[11]) & s(S[15]) \\
s(S[4]) & s(S[8]) & s(S[12]) & s(S[16])
\end{array}\right),
$$

where $s$ is the non linear application defined by

$$
\begin{aligned}
G F\left(2^{8}\right) & \longrightarrow G F\left(2^{8}\right) \\
x & \longmapsto s(x)=\left\{\begin{aligned}
a * x^{-1}+b, & \text { if } x \neq 0, \\
b, & \text { if } x=0 .
\end{aligned}\right.
\end{aligned}
$$

$a$ is a linear invertible application over $G F(2), a \in M_{8}(G F(2)), *$ is the multiplication of matrices over $G F(2)$ and $x^{-1}=\left\{b_{0} b_{1} \ldots b_{7}\right\}_{b}$ is seen as a $G F(2)$-vector equal to tranpose of the vector $\left(b_{0}, \cdots, b_{7}\right)$. The value of $b=$ $' 63 ' \in G F\left(2^{8}\right)$ and

$$
a=\left(\begin{array}{llllllll}
1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\
1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\
1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\
1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \\
1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\
0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\
0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 \\
0 & 0 & 0 & 1 & 1 & 1 & 1 & 1
\end{array}\right)
$$

### 2.5 MixColumn for $i^{\text {th }}$ round

The MixColumn transformation consists of a multiplication of matrices in $M_{4}\left(G F\left(2^{8}\right)\right)$, between the state and a fixed matrix $A_{0}$ of $M_{4}\left(G F\left(2^{8}\right)\right)$. We denote by $S_{i, M}$ the state after the $i^{\text {th }}$ MixColumn.

$$
\begin{aligned}
& M_{4}\left(G F\left(2^{8}\right)\right) \longrightarrow M_{4}\left(G F\left(2^{8}\right)\right) \\
& S \longmapsto \\
& S_{i, M}=A_{0} . S
\end{aligned}
$$

where $A_{0}$ is defined by

$$
A_{0}=\left(\begin{array}{cccc}
02 & 03 & 01 & 01 \\
01 & 02 & 03 & 01 \\
01 & 01 & 02 & 03 \\
03 & 01 & 01 & 02
\end{array}\right)
$$

### 2.6 ShiftRows for $i^{t h}$ round

The ShiftRows transformation is a byte transposition that cyclically shifts the rows of the state over different offsets. We denote by $S_{i, S h}$ the state after the $i^{t h}$ ShiftRows.

$$
S=\left(\right) \quad \longmapsto \quad S_{i, S h}=\left(\begin{array}{cccc}
S[1] & S[5] & S[9] & S[13] \\
S[6] & S[10] & S[14] & S[2] \\
S[11] & S[15] & S[3] & S[7] \\
S[16] & S[4] & S[8] & S[12]
\end{array}\right) .
$$

## 3 The description of the attack on computation of AES

First, we are going describe an attack on AES in a simple case and after that we will see how we can generalize this attack. The goal of the attack is to recover the key $K_{N r}$. Once we discover the subkey $K_{N r}$, it is easy to get the key $K$, see appendix A.

### 3.1 Principle of the attack

We suppose that we can change a single byte of the state after the ShiftRow of the $N_{r}-1$ round and we know the index of the faulty element of state (this last supposition can be omitted, it is more easier to explain the mechanism). The new value of the element of the state is supposed unknown. The fault $\varepsilon$ is spread over four bytes on the output state. For each modified elements on the output state, we find a set of possible fault $\varepsilon$. Moreover we can intersect the possible values $\varepsilon$ for these four sets, we obtain a small set thus reducing the number of required ciphertext for the full analysis. Finally for each fault, we deduce some possible values of four elements of the last roundkey. Repeating ciphertexts, we find four bytes of roundkey 10.

This attack still works out, even with more general assumptions on the fault locations, such as faults without knowing the fault locations before the $9^{t h}$ MixColumn transformation. We also expect that faults in round 8 (before the $8^{t h}$ MixColumn transformation) might be useful for the analysis, thus growing the number of required ciphertext for the full analysis. With our example, we need ten ciphertexts to get four bytes of roundkey 10, when we don't make hypothesis about the fault locations.

### 3.2 Example

We use the same example as Appendix B of [1]. The following diagram shows the values in the final States array as the Cipher progresses for a block length and a Cipher Key length of 16 bytes each (i.e., $\mathrm{Nb}=4$ and $\mathrm{Nk}=4)$.

Input= '32 43 F 6 A 8885 A 30 8D 313198 A 2 E 0370734 '
Cipher Key= '2B 7E 151628 AE D2 A6 AB F7 158809 CF 4F 3C'
Output= '39 2584 1D 02 DC 09 FB DC 11859719 6A 0B 32'
The fault propagation appears in grey tint and in hexadecimal notation:
After ShiftRows 9

| 87 | F 2 | 4 D | 97 |
| :---: | :---: | :---: | :---: |
| 6 E | 4 C | 90 | EC |
| 46 | E 7 | 4 A | C 3 |
| A 6 | 8 C | D 8 | 95 |

Fault injected 1E

| 99 | F2 | 4 D | 97 |
| :---: | :---: | :---: | :---: |
| 6 E | 4 C | 90 | EC |
| 46 | E 7 | 4 A | C 3 |
| A 6 | 8 C | D 8 | 95 |


| After Mixcolumn |  |  |  |
| :---: | :---: | :---: | :---: |
| 7 B | 40 | A3 | 4 C |
| 29 | D 4 | 70 | 9 F |
| 8 A | E 4 | 3 A | 42 |
| CF | A5 | A6 | BC |


$\oplus$| AC | 19 | 28 | 57 |
| :---: | :---: | :---: | :---: |
| 77 | FA | D1 | 5 C |
| 66 | DC | 29 | 00 |
| F3 | 21 | 41 | 6 E |


| After Subsytes 10 |  |  |  |
| :---: | :---: | :---: | :---: |
| 0 E | CB | 3 D | AF |
| 58 | 31 | 32 | 2 E |
| CE | 07 | 7 D | 2 C |
| EB | 5 F | 94 | B 5 |


| After ShiftRows 10 |  |  |  |
| :---: | :---: | :---: | :---: |
| 0 E | CB | 3 D | AF |
| 31 | 32 | 2 E | 58 |
| 7 D | 2 C | CE | 07 |
| B5 | EB | 5 F | 94 |


| value of $K_{10}$ |  |  |  |
| :---: | :---: | :---: | :---: |
| D0 | C 9 | E1 | B6 |
| 14 | EE | 3 F | 63 |
| F9 | 25 | 0 C | 0 C |
| A8 | 89 | C8 | A6 |

Output with Faults

| DE | 02 | DC | 19 |
| :---: | :---: | :---: | :---: |
| 25 | DC | 11 | 3 B |
| 84 | 09 | C 2 | 0 B |
| 1D | 62 | 97 | 32 |

The injected error in the state, give four errors in the final state.

### 3.3 How the injected error acts on the final state

We denote by $F$ the faulty state. Now we describe each step from the $N_{r}-1^{t h}$ MixColumn to the end, and assume that we replace the first element of the state by an unknown value. Let $\varepsilon \in G F\left(2^{8}\right)-\{0\}$ defined by

$$
F_{N_{r}-1, S h}[1]=S_{N_{r}-1, S h}[1]+\varepsilon .
$$

### 3.3.1 Fault modification

Obviously

$$
F_{N_{r}-1, S h}=S_{N_{r}-1, S h}+\left(\begin{array}{cccc}
\varepsilon & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{array}\right)
$$

### 3.3.2 Effect on MixColumn

$$
F_{N_{r}-1, M}=S_{N_{r}-1, M}+A_{0} .\left(\begin{array}{cccc}
\varepsilon & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{array}\right)=S_{N_{r}-1, M}+\left(\begin{array}{cccc}
2 . \varepsilon & 0 & 0 & 0 \\
\varepsilon & 0 & 0 & 0 \\
\varepsilon & 0 & 0 & 0 \\
3 . \varepsilon & 0 & 0 & 0
\end{array}\right) .
$$

### 3.3.3 Effect on AddRoundKey

$$
F_{N_{r}-1, A}=S_{N_{r}-1, A}+\left(\begin{array}{cccc}
2 . \varepsilon & 0 & 0 & 0 \\
\varepsilon & 0 & 0 & 0 \\
\varepsilon & 0 & 0 & 0 \\
3 . \varepsilon & 0 & 0 & 0
\end{array}\right)
$$

### 3.3.4 Effect on last SubBytes

We can define $\varepsilon_{0}^{\prime}, \varepsilon_{1}^{\prime}, \varepsilon_{2}^{\prime}, \varepsilon_{3}^{\prime}$ (the differential faults) by the following equation

$$
F_{N_{r}, S u}=S_{N_{r}, S u}+\left(\begin{array}{cccc}
\varepsilon_{0}^{\prime} & 0 & 0 & 0 \\
\varepsilon_{1}^{\prime} & 0 & 0 & 0 \\
\varepsilon_{2}^{\prime} & 0 & 0 & 0 \\
\varepsilon_{3}^{\prime} & 0 & 0 & 0
\end{array}\right)
$$

### 3.3.5 Effect after last ShiftRows

$$
F_{N_{r}, S h}=S_{N_{r}, S h}+\left(\begin{array}{cccc}
\varepsilon_{0}^{\prime} & 0 & 0 & 0 \\
0 & 0 & 0 & \varepsilon_{1}^{\prime} \\
0 & 0 & \varepsilon_{2}^{\prime} & 0 \\
0 & \varepsilon_{3}^{\prime} & 0 & 0
\end{array}\right)
$$

### 3.3.6 Effect after last AddRoundKey

$$
F_{N_{r}, A}=S_{N_{r}, A}+\left(\begin{array}{cccc}
\varepsilon_{0}^{\prime} & 0 & 0 & 0 \\
0 & 0 & 0 & \varepsilon_{1}^{\prime} \\
0 & 0 & \varepsilon_{2}^{\prime} & 0 \\
0 & \varepsilon_{3}^{\prime} & 0 & 0
\end{array}\right)
$$

$F_{N_{r}, A}$ is the faulty output for a cipher. Comparing the states $F_{N_{r}, A}$ with $S_{N_{r}, A}$, it is easy to get the values of $\varepsilon_{0}^{\prime}, \varepsilon_{1}^{\prime}, \varepsilon_{2}^{\prime}$ and $\varepsilon_{3}^{\prime}$.

### 3.4 Example

Always, in hexadecimal notation, we find

|  | atput | h faults |  | $\oplus$ | Output without fault |  |  |  | $=$ | Error |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| DE | 02 | DC | 19 |  | 39 | 02 | DC | 19 |  | E7 | 00 | 00 | 00 |
| 25 | DC | 11 | 3B |  | 25 | DC | 11 | 6A |  | 00 | 00 | 00 | 51 |
| 84 | 09 | C2 | 0B |  | 84 | 09 | 85 | 0B |  | 00 | 00 | 47 | 00 |
| 1D | 62 | 97 | 32 |  | 1D | FB | 97 | 32 |  | 00 | 99 | 00 | 00 |

The differential faults are $\varepsilon_{0}^{\prime}={ }^{'} \mathrm{E} 7^{\prime}, \varepsilon_{1}^{\prime}={ }^{\prime} 51^{\prime}, \varepsilon_{2}^{\prime}={ }^{\prime} 47$ ' and $\varepsilon_{3}^{\prime}={ }^{\prime} 99^{\prime}$.

### 3.5 Analysis on information brought by fault

The only operation that could bring information about the key $K_{N_{r}}$ is the last SubBytes transformation. Consequently we have four equations where $x_{0}, x_{1}, x_{2}, x_{3}, \varepsilon$ are unknown variables. We want to solve the following equations (in $x_{i}$ and $\varepsilon$ ) :

$$
\left\{\begin{aligned}
s\left(x_{0}+2 . \varepsilon\right) & =s\left(x_{0}\right)+\varepsilon_{0}^{\prime} \\
s\left(x_{1}+\varepsilon\right) & =s\left(x_{1}\right)+\varepsilon_{1}^{\prime} \\
s\left(x_{2}+\varepsilon\right) & =s\left(x_{2}\right)+\varepsilon_{2}^{\prime} \\
s\left(x_{3}+3 . \varepsilon\right) & =s\left(x_{3}\right)+\varepsilon_{3}^{\prime}
\end{aligned}\right.
$$

All these equations belong to a generalized equation

$$
\begin{equation*}
s(x+c . \varepsilon)+s(x)=\varepsilon^{\prime}, \tag{1}
\end{equation*}
$$

where $c=$ ' 01 ', ' 02 ' or ' 03 ' and let us analyse it.
Definition 1 We define the set of solutions of (1) in $\varepsilon$ by

$$
S_{c, \varepsilon^{\prime}}=\left\{\varepsilon \in G F\left(2^{8}\right): \exists x \in G F\left(2^{8}\right), s(x+c . \varepsilon)+s(x)=\varepsilon^{\prime}\right\} .
$$

Definition 2 Consider the linear application over $G F(2)$ :

$$
\begin{aligned}
l: \quad G F\left(2^{8}\right) & \longrightarrow G F\left(2^{8}\right) \\
x & \longmapsto x^{2}+x
\end{aligned}
$$

Denote by $E_{1}=\operatorname{Im}(l)$ be the $G F(2)$-vector space image of $l$ and $\operatorname{dim}_{G F(2)}\left(E_{1}\right)=7$. If $\theta \in E_{1}$, then there are two solutions $x_{1}, x_{2} \in G F\left(2^{8}\right)$ of equation $x^{2}+x=\theta$, and the solutions verify $x_{2}=x_{1}+1$.

Definition 3 Let $\lambda \in G F\left(2^{8}\right), \lambda \neq 0$ and define $\phi_{\lambda}$ an isomorphism of $G F(2)$-vector spaces

$$
\begin{aligned}
\phi_{\lambda}: \quad G F\left(2^{8}\right) & \longrightarrow G F\left(2^{8}\right) \\
x & \longmapsto \lambda \cdot x
\end{aligned}
$$

and let $E_{\lambda}=\operatorname{Im}\left(\left.\phi_{\lambda}\right|_{E_{1}}\right)$ be the $G F(2)$-vector space image of $\phi_{\lambda}$ restricted to $E_{1}$. Moreover $\operatorname{dim}_{G F(2)}\left(E_{\lambda}\right)=7$.
Proposition 1 There is a bijective application $\phi$ between $E_{1}^{*}\left(=E_{1}-\{0\}\right)$ and $S_{c, \varepsilon^{\prime}}$.

$$
\begin{aligned}
\phi: \quad E_{1}^{*} & \longrightarrow S_{c, \varepsilon^{\prime}} \\
t & \longmapsto\left(c\left(a^{-1} * \varepsilon^{\prime}\right) \cdot t\right)^{-1} .
\end{aligned}
$$

$S_{c, \varepsilon^{\prime}}$ have 127 elements.
Proof: Let $\varepsilon \in S_{c, \varepsilon^{\prime}}$, then $\exists x \in G F\left(2^{8}\right)$ such that (1) holds.
Assume $x \neq 0$ and $x \neq c$. $\varepsilon$, we get

$$
x^{2}+c . \varepsilon . x=\left(a^{-1} * \varepsilon^{\prime}\right)^{-1} . c . \varepsilon .
$$

We denote by $t=x .(c . \varepsilon)^{-1} \in G F\left(2^{8}\right)-\{0\}$, then we have

$$
\begin{equation*}
t^{2}+t=\left(a^{-1} * \varepsilon^{\prime}\right)^{-1} \cdot(c \cdot \varepsilon)^{-1} \tag{2}
\end{equation*}
$$

Therefore $\left(a^{-1} * \varepsilon^{\prime}\right)^{-1}(c . \varepsilon)^{-1} \in E_{1}^{*}$. Reciprocally for $\theta \in E_{1}^{*}$ we can define $\left(a^{-1} * \varepsilon^{\prime}\right)^{-1} \cdot(c . \theta)^{-1} \in S_{c, \varepsilon^{\prime}}$.
Assume $x=0$ or $x=c . \varepsilon$, (1) becomes $a *(c . \varepsilon)^{-1}=\varepsilon^{\prime}$. We obtain $\varepsilon=\left(\left(a^{-1} * \varepsilon^{\prime}\right) . c\right)^{-1}$, this case is included in
the previous case because $1 \in E_{1}^{*}$. We see for the case $\theta=1$, the equation (1) has four solutions in $x$. In brief, there exists a bijection map between $E_{1}^{*}$ and $S_{c, \varepsilon^{\prime}}$ :

$$
\begin{array}{rlccc}
E_{1}^{*} & \xrightarrow[\phi_{\lambda}]{ } & E_{\lambda}-\{0\} & \longrightarrow & S_{c, \varepsilon^{\prime}} \\
t & \longmapsto & \lambda . t & \longmapsto & (\lambda . t)^{-1}
\end{array}
$$

where $\lambda=c\left(a^{-1} * \varepsilon^{\prime}\right)$.

Proposition 2 The following statements hold for $\lambda_{1}, \lambda_{2} \in G F\left(2^{8}\right)-\{0\}$ :

$$
\operatorname{dim}_{G F(2)}\left(E_{\lambda_{1}} \cap E_{\lambda_{2}}\right)= \begin{cases}7 & \text { If } \lambda_{1}=\lambda_{2} \\ 6 & \text { Otherwise }\end{cases}
$$

Proof: This proof comes from the following lemma:
Lemma 1 For $\lambda_{1}, \lambda_{2} \in G F\left(2^{8}\right)-\{0\}$, we get

$$
E_{\lambda_{1}}=E_{\lambda_{2}} \Longleftrightarrow \lambda_{1}=\lambda_{2}
$$

Proof: This lemma is equivalent to this assertion : for $\lambda \in G F\left(2^{8}\right)-\{0\}$,

$$
E_{\lambda}=E_{1} \Longleftrightarrow \lambda=1
$$

Let us prove this statement and assume that $\lambda E_{1}=E_{1}$. Remark that $E_{1}=\left\{e=\left\{e_{7} e_{6} \cdots e_{0}\right\}_{b} \in G F\left(2^{8}\right)-\right.$ $\left.\{0\}: e_{7}=e_{5}\right\}$. Hence $\left\{1, x, x^{2}, x^{3}, x^{4}, x^{6}, x^{5}+x^{7}\right\}$ is a basis of $E_{1}$. Multiply the basis's vectors $v_{i}$ with $\lambda=\left\{\lambda_{7} \cdots \lambda_{0}\right\}_{b}$. As $\lambda v_{i} \in E_{1}$, we have $\left(\lambda v_{i}\right)_{7}=\left(\lambda v_{i}\right)_{5}$. We obtain 7 relations $\left(\lambda_{7}=\lambda_{5}, \lambda_{6}=\lambda_{4}, \lambda_{5}=\lambda_{3}+\lambda_{7}\right.$, $\left.\lambda_{4}=\lambda_{6}+\lambda_{2}+\lambda_{7}, \lambda_{7}+\lambda_{3}=\lambda_{5}+\lambda_{1}+\lambda_{6}, \lambda_{5}+\lambda_{1}=\lambda_{3}+\lambda_{4}, \lambda_{6}+\lambda_{5}=\lambda_{7}+\lambda_{3}\right)$. We solve this system to obtain $\lambda_{7}=\lambda_{6}=\lambda_{5}=\lambda_{4}=\lambda_{3}=\lambda_{2}=\lambda_{1}=0$. The solution $\lambda=0$ doesn't match. We have $\lambda=1$.

Proposition 3 For $\lambda_{1}, \lambda_{2}, \lambda_{3} \in G F\left(2^{8}\right)-\{0\}$, we get:

$$
\operatorname{dim}_{G F(2)}\left(E_{\lambda_{1}} \cap E_{\lambda_{2}} \cap E_{\lambda_{3}}\right)= \begin{cases}7 & \text { If } \lambda_{1}=\lambda_{2}=\lambda_{3} \\ 6 & \text { If rank } \\ 5 F(2)\left\{\lambda_{1}^{-1}, \lambda_{2}^{-1}, \lambda_{3}^{-1}\right\}=2 \\ 5 & \text { Otherwise }\end{cases}
$$

Proof: It comes from proposition 2 and this following lemma
Lemma 2 For $\lambda_{1}, \lambda_{2}, \lambda_{3} \in G F\left(2^{8}\right)-\{0\}$, we get

$$
E_{\lambda_{1}} \cap E_{\lambda_{3}}=E_{\lambda_{2}} \cap E_{\lambda_{3}} \Longleftrightarrow \lambda_{3}^{-1}=\lambda_{1}^{-1}+\lambda_{2}^{-1} \text { or } \lambda_{1}=\lambda_{2} .
$$

Proof:

1. $\Leftarrow$

Let $x \in E_{\lambda_{1}} \cap E_{\lambda_{3}}$, then $\exists y, t \in E_{1}$ such that $x=\lambda_{1} \cdot y=\lambda_{3}$.t.

$$
\begin{gathered}
y=\lambda_{1}^{-1} \cdot \lambda_{3} \cdot t=\lambda_{2}^{-1} \cdot \lambda_{3} \cdot t+t \\
y-t=\lambda_{2}^{-1} \cdot \lambda_{3} \cdot t \in E_{1}
\end{gathered}
$$

and

$$
x=\lambda_{3} \cdot t=\lambda_{2} \cdot(y-t) \in E_{\lambda_{2}}
$$

2. $\Rightarrow$

Assume that $\lambda_{1} \neq \lambda_{2}$, and let us show $\forall t \in E_{1}, \lambda_{3} .\left(\lambda_{1}^{-1}+\lambda_{2}^{-1}\right) . t \in E_{1}$.
Let $x=\lambda_{3} . t \in E_{\lambda_{3}}$ :

- If $x \in E_{\lambda_{1}}$ then $x \in E_{\lambda_{2}}$ therefore $\exists s_{1}, s_{2} \in E_{1}$ such that $x=\lambda_{1} \cdot s_{1}=\lambda_{2} . s_{2}$ and we get $\lambda_{3} .\left(\lambda_{1}^{-1}+\right.$ $\left.\lambda_{2}^{-1}\right) \cdot t=s_{1}+s_{2} \in E_{1}$.
- If $x \notin E_{\lambda_{1}}$ then $x \notin E_{\lambda_{2}}$ therefore we get $\lambda_{1}^{-1} \cdot x \notin E_{1}$ and $\lambda_{2}^{-1} \cdot x \notin E_{1}$. We have $\lambda_{3} .\left(\lambda_{1}^{-1}+\lambda_{2}^{-1}\right) . t=$ $\lambda_{1}^{-1} \cdot x+\lambda_{2}^{-1} \cdot x \in E_{1}$ (because $\forall u \notin E_{1}$ and $\forall v \notin E_{1}$ then $u+v \in E_{1}$ ).

We showed that $E_{\lambda_{3} \cdot\left(\lambda_{1}^{-1}+\lambda_{2}^{-1}\right)}=E_{1}$ and with the lemma 1 we get $\lambda_{3}^{-1}=\lambda_{1}^{-1}+\lambda_{2}^{-1}$.

Proposition 4 Finally for $\lambda_{1}, \lambda_{2}, \lambda_{3}, \lambda_{4} \in G F\left(2^{8}\right)-\{0\}$, we get:

Definition 4 We considered four equations in a different way, but the committed fault is common to these four equations, that is why we introduce the set of possible committed faults $S$ :

$$
S=S_{2, \varepsilon_{0}^{\prime}} \bigcap S_{1, \varepsilon_{1}^{\prime}} \bigcap S_{1, \varepsilon_{2}^{\prime}} \bigcap S_{3, \varepsilon_{3}^{\prime} .} .
$$

Moreover the cardinal of $S$ is smaller than the cardinal of $S_{c, \varepsilon}$. It allows to reduce the space of the faults, and so to use fewer faultly calculations to go back up to the key.

Corollary 1 If two of the four following values $2^{-1} . \varepsilon_{0}^{\prime}, \varepsilon_{1}^{\prime}, \varepsilon_{2}^{\prime}, 3^{-1} . \varepsilon_{3}^{\prime}$ are not equal, we have

$$
\operatorname{Card}\left(S_{2, \varepsilon_{0}^{\prime}} \bigcap S_{1, \varepsilon_{1}^{\prime}} \bigcap S_{1, \varepsilon_{2}^{\prime}} \bigcap S_{3, \varepsilon_{3}^{\prime}}\right) \leq 63
$$

Proposition 5 For a differential fault $\varepsilon^{\prime}$, let $\varepsilon \in S \cap S_{c, \varepsilon^{\prime}}$ be a fault value and define $\theta=\left(\left(a^{-1} * \varepsilon^{\prime}\right) . c . \varepsilon\right)^{-1} \in E_{1}^{*}$ and $\alpha, \beta$ the two solutions (in $G F\left(2^{8}\right)$ ) of the equation $t^{2}+t=\theta$. The possible values of key $K_{N_{r}}[i]$ (for some $i$, it is the index of element in the state) are

- If $\theta \neq 1$, then there are two possible values of $K_{N_{r}}[i]$

$$
K_{N_{r}}[i]=s(c . \varepsilon . \alpha)+F_{N_{r}, A}[i] \text { or } K_{N_{r}}[i]=s(c . \varepsilon . \beta)+F_{N_{r}, A}[i]
$$

- If $\theta=1$, then there are four possible values of $K_{N_{r}}[i]$

$$
\begin{gathered}
K_{N_{r}}[i]=s(c . \varepsilon . \alpha)+F_{N_{r}, A}[i] \text { or } K_{N_{r}}[i]=s(c . \varepsilon . \beta)+F_{N_{r}, A}[i] \\
\text { or } K_{N_{r}}[i]=b+F_{N_{r}, A}[i] \text { or } K_{N_{r}}[i]=s(c . \varepsilon)+F_{N_{r}, A}[i]
\end{gathered}
$$

Proof:

- If $\theta \neq 1$, we know that $\theta \in E_{1}$, then there are two solutions $\alpha, \beta$ of $t^{2}+t=\theta$. We deduce two solutions from (1) noted $\left\{x_{1}, x_{2}\right\}$, by $x_{1}=c . \varepsilon . \alpha$ and $x_{2}=c . \varepsilon . \beta$.
- If $\theta=1$, we know that $1 \in E_{1}$, then there are two solutions $\alpha, \beta$ of $t^{2}+t=1$. We deduce two solutions from (1) noted $\left\{x_{1}, x_{2}\right\}$, by $x_{1}=c . \varepsilon . \alpha$ and $x_{2}=c . \varepsilon . \beta$. Moreover there are also two trivial solutions of (1) : $x_{3}=0$ and $x_{4}=c . \varepsilon$.

Once we get a solution $x$ of (1), it is easy to get a possible value of $K_{N_{r}}[i]$.
By applying this proposition to the four faulty elements of the state, we can deduce four sets of possible values for $K_{N_{r}}[0], K_{N_{r}}[7], K_{N_{r}}[10]$ and $K_{N_{r}}[13]$. Then by repeating the insertion of faults in a calculation, and by intersecting these four sets we get rather quickly a single value for $K_{N_{r}}[0], K_{N_{r}}[7], K_{N_{r}}[10]$ and $K_{N_{r}}[13]$.

### 3.6 Example

Remember our example:

$$
\begin{aligned}
s\left(x_{0} \oplus 2 . \varepsilon\right) & =s\left(x_{0}\right) \oplus{ }^{\prime} \mathrm{E} 7 \\
s\left(x_{1} \oplus \varepsilon\right) & =s\left(x_{1}\right) \oplus{ }^{\prime} 51 \\
s\left(x_{2} \oplus \varepsilon\right) & =s\left(x_{2}\right) \oplus{ }^{\prime} 47 \\
s\left(x_{3} \oplus 3 . \varepsilon\right) & =s\left(x_{3}\right) \oplus{ }^{\prime} 99^{\prime}
\end{aligned}
$$

Let $E_{1}=\left\{{ }^{\prime} 01^{\prime} . . .^{\prime} 1 \mathrm{~F}^{\prime},,^{\prime} 40^{\prime} . .{ }^{\prime} 5 \mathrm{~F}^{\prime},{ }^{\prime} \mathrm{A}^{\prime}{ }^{\prime} . . \mathrm{BF}^{\prime},{ }^{\prime} \mathrm{E} 0^{\prime} . .{ }^{\prime} \mathrm{FF} '\right\}$ and $S_{c, \varepsilon^{\prime}}=\left\{\left(c .\left(a^{-1} * \varepsilon^{\prime}\right) . e\right)^{-1}, \quad e \in E 1\right\}$.

We compute

$$
\begin{aligned}
& S_{2,{ }^{\prime}{ }^{\prime} 7^{\prime}} \bigcap S_{1,{ }^{\prime} 1_{1}, \bigcap} S_{1,{ }^{\prime} 477^{\prime}} \bigcap S_{3,{ }^{\prime} 99}, \\
& =\text { \{'01', '04', '13', '1E', '21', '27', '33', '3B', '48', '4D', '50', '53', '55', '5D', '64', '65', } \\
& \text { '7E', '7F', '80', '83', '8D', '8F', '93', 'A7', 'A8', 'A9', 'AB', 'B3', 'B8', 'C9', 'F6'\} }
\end{aligned}
$$

We get (the real value of $K_{10}[0]$ is 'D0')
$K_{10}[0] \in\left\{{ }^{\prime} 03 ', ~ ' 06 ', ~ ' 09 ', ~ ' 0 C ', ~ ' 10 ', ~ ' 15 ', ~ ' 1 A ', ~ ' 1 F ', ~ ' 21 ', ~ ' 24 ', ~ ' 2 B ', ~ ' 2 E ', ~ ' 32 ', ~ ' 37 ', ~ ' 38 ', ~ ' 3 D ', ~ ' 43 ', ~ ' 46 ', ~ ' 49 ', ~\right.$ '4C', '50', '55', '5F', '61', '64', '6B', '6E', '72', '77', '78', '7D', '83', '86', '89', '8C', '90', '95', '9A', '9F', 'A1', 'A4', 'AB', 'AE', 'B2', 'B7', 'B8', 'C3', 'C6', 'C9', 'CC', 'D0', 'D5', 'DA', 'DF', 'E1', 'E4', 'EB', 'EE', 'F2', 'F7', 'F8', 'FD'\}

With the five faults \{'1E', 'E1', 'B3', '16', '9E'\}, we obtain a correct and single value of $K_{10}[0], K_{10}[7], K_{10}[10]$, $K_{10}[13]$.

## 4 Generalisation

### 4.1 Without fault location

In this section, we assume that the fault is on a byte, between the last two MixColumn. It's the same case than previously except that the fault can be confined on the byte 1 to 16 . The fault is propagated by the MixColumn and spread on 4 bytes of the state. On the first line of the differential state matrix, we have a induced fault. We can determine from which column the injected fault belongs by considering the column of induced fault. Next we analyse the four possibilities of line position for the injected fault with the method presented in previous section.

### 4.2 Hardware Device

Suppose that you can physically modify an hardware AES device. First, compute ciphers for more than ten random plaintexts with AES device. Next, modify by example the design by cutting lines and connecting them to the earth (or Vcc) temporaly between two bytes during the round located two rounds before the end. It amounts to having a byte of round $N_{k}-2$, always replaced by ' 00 ' (or 'FF'). Compute an other time the same messages with the tampered device. With random plaintexts, the faulty byte is like an random fault. This fault is passed on four faults at round $N_{k}-1$ and sixteen faults at round $N_{k}$. It is this differential matrix we can analyse error by error to find the last round key.

## A Back to initial key with the last subkey

See [1] for additional informations about $w$ and RotWord, Rcon and SubWord functions.
Let us denote by $K_{n}[j]$ the $j^{t h}$ byte of the $n^{t h}$ roundkey and $w[i]$ as in [1]. We have

$$
K_{n}=\left(w\left[N_{k} n\right], w\left[N_{k} n+1\right], \cdots, w\left[N_{k} n+N_{k}-1\right]\right) .
$$

We have the following relations (for $N_{k}=4,6$ ):
for $N_{k} \leqslant i<N b *(N r+1), i \neq 0 \bmod N_{k}$,

$$
\begin{aligned}
w[i] & =w\left[i-N_{k}\right] \oplus w[i-1] \\
\text { i.e. } w\left[i-N_{k}\right] & =w[i] \oplus w[i-1]
\end{aligned}
$$

and for $i=0 \bmod N_{k}$,

$$
\begin{aligned}
w[i] & =w\left[i-N_{k}\right] \oplus \operatorname{SubWord}(\operatorname{RotWord}(w[i-1])) \oplus \operatorname{Rcon}\left[i / N_{k}\right] \\
\text { i.e. } w\left[i-N_{k}\right] & =w[i] \oplus \operatorname{SubWord}(\operatorname{RotWord}(w[i-1])) \oplus \operatorname{Rcon}\left[i / N_{k}\right]
\end{aligned}
$$

Hence, we have
for $0 \leqslant i<N b *(N r+1)-N k, i \neq 0 \bmod N_{k}$,

$$
\begin{equation*}
w[i]=w\left[i+N_{k}\right] \oplus w\left[i+N_{k}-1\right] \tag{3}
\end{equation*}
$$

and for $i=0 \bmod N_{k}$,

$$
\begin{equation*}
w[i]=w\left[i+N_{k}\right] \oplus \operatorname{SubWord}\left(\operatorname{Rot} \operatorname{Word}\left(w\left[i+N_{k}-1\right]\right)\right) \oplus \operatorname{Rcon}\left[\left(i+N_{k}\right) / N_{k}\right] \tag{4}
\end{equation*}
$$

With AES-256, you must add an Subword operation when $i \equiv 4 \bmod N_{k}$. So we can deduce previous key from the ending subkey and step by step obtain $K_{0}$ with is the cipherkey.

```
RecoverKey(byte Finalkey[4*Nk], word w[Nb*(Nr+1)], Nk)
begin
    word temp
    i = Nb * (Nr+1)-1
    j = Nk - 1
    while (j >= 0)
        w[i] = word(Finalkey[4*j], Finalkey[4*j+1],
                    Finalkey[4*j+2], Finalkey[4*j+3])
        i = i-1
        j = j-1
    end while
{here, "i" must be equal to Nb * (Nr+1) - Nk - 1}
    while (i >= 0)
        temp = w[i+Nk-1]
        if (i mod Nk = 0)
            temp = SubWord(RotWord(temp)) xor Rcon[i/Nk+1]
        else if (Nk > 6 and i mod Nk = 4)
            temp = SubWord(temp)
        end if
        w[i] = w[i+Nk] xor temp
        i = i - 1
    end while
end
```

Figure 1: Pseudo Code for Key Recovery.

Remark 1 On AES-128, it is sufficient to know $K_{10}$ to find the cipher key, but on $A E S$-256, you must know $K_{13}$ and $K_{14}$.

## References

[1] FIPS PUB 197 : Avanced Encryption Standard, http://csrc.nist.gov/publications/fips/fips197/fips-197.pdf
[2] Ross J. Anderson, Markus G. Kuhn: Tamper Resistance - a Cautionary Note, The Second USENIX Workshop on Electronic Commerce Proceedings, Oakland, California, November 18-21, 1996, pp 1-11, ISBN 1-880446-83-9.
[3] E. Biham \& A.Shamir, Differential Fault Analysis of Secret Key Cryptosystems, CS 0910, Proceedings of Crypto'97.
[4] Boneh, DeMillo, and Lipton, On the Importance of Checking Cryptographic Protocols for Faults, Lecture Notes in Computer Science, Advances in Cryptology, proceedings of EUROCRYPT'97, pp. 37-51, 1997.
[5] Joan Daemen, Annex to AES Proposal Rijndael, http://www.esat.kuleuven.ac.be/~rijmen/rijndael/PropCorr.PDF, 1998.
[6] G. Letourneux, Rapport de stage EDSI : Etude et implémentation de l'AES, Attaques DPA et DFA, August 30, 2002.


[^0]:    ${ }^{*}$ LACO (UMR CNRS n ${ }^{\circ} 6090$ ), Faculté des Sciences, 123, avenue Albert THOMAS, 87060 Limoges, France
    ${ }^{\dagger}$ E.D.S.I. Atalis 1, 1, rue de Paris, 35510 Cesson-Sévigné, France
    $\ddagger$ development@edsi-smartcards.com.fr, E.D.S.I. Atalis 1, 1, rue de Paris, 35510 Cesson-Sévigné, France

