Fault Attacks on Electronic Circuits

Jean-Luc DANGER, Sylvain GUILLEY
<jean-luc.danger@telecom-paristech.fr>
Outline

Introduction

Fault Injection Attacks examples

DFA on AES

Countermeasures
Outline

Introduction
  Physical Faults
    Faults Models and Tolerance

Fault Injection Attacks examples
  Fault Attack on the control part
  Differential Fault Attack on cryptographic functions

DFA on AES
  Theoretical Fault Attacks
  Practical Attacks

Countermeasures
  Redundancy
  Resilience
  Detection
How Attackers Actually Inject Faults

- **Glitch** Attacks on the Power or the Clock (synchronous circuits)
- High Energy **Particles**. However, they can be replaced by:
- Focused **Laser** (spot $\Phi \sim 1 \mu m$), front-side or back-side
- Using **bugged** HW/SW Components (Intel ® Pentium flawed floating point division, back to 1994)
- Eddy currents $\approx$ **EMI** (ElectroMagnetic Injection)
- etc.

See also: Fault Diagnosis and Tolerance in Cryptography (FDTC)
Realignement

[WISTP ’11, Guilley et al.] [GKLD11]
Laser station
Fault Injection Attacks  Jean-Luc Danger
Example @ TELECOM-ParisTech ASIC

Fault Injection Attacks Jean-Luc Danger
EMI disturbance system
Example of setup for EMI
Example of setup for EMI
Flipping Bits in Memory Without Accessing Them

Other faults on the RAM [Ver06]
Other faults on the RAM [Ver06]

Debian GNU/Linux comes with ABSolutely NO WARRANTY, to the extent permitted by applicable law.

bacchus:~# cd /home/briancea/
bacchus:/home/briancea# ls
mem
bacchus:/home/briancea# cd mem/
bacchus:/home/briancea/mem# ls

applitest checkmen.c flipbit.c
checkmen flipbit hs_err_pid1990.log
bacchus:/home/briancea/mem# make run
java -Xms112m -Xmx112m attaque.Demo

Attente d’un rayon cosmique ...

Bitflip detecte sur l’instance de B tableauB(69158), attribut 136
La nouvelle reference semble pointer sur une instance de B
Bitflip detecte sur l’instance de B tableauB(69158), attribut 137
La nouvelle reference semble pointer sur une instance de B
Recherche de l’instance de classe B sur laquelle on a la main ...
Instance trouvée (56870), attaque réalisable.

Menu
0 - Quitter
1 - Duniper une zone memoire
2 - Ecrire un mot de 4 octets en memoire

Hot air gun OFF  Hot air gun OFF

Institut Mines-Télécom
Fault Injection Attacks Jean-Luc Danger
Outline

Introduction
  Physical Faults
  Faults Models and Tolerance

Fault Injection Attacks examples
  Fault Attack on the control part
  Differential Fault Attack on cryptographic functions

DFA on AES
  Theoretical Fault Attacks
  Practical Attacks

Countermeasures
  Redundancy
  Resilience
  Detection
Fault impact

Single Event Effects (SEE)

**SET:** Single Event Transient Fault.

**SEU:** Single Event Upsets or Soft error. Transient Fault impacting a memory point or register by current peak (soft error) It is the useful fault for an attack.

**SEL:** Single Event Latchups. Short-circuit between \( V_{ss} \) and \( V_{dd} \), causing a permanent fault (**hard error**)
### SEU abstraction models

<table>
<thead>
<tr>
<th>Physical</th>
<th>Electrical</th>
<th>Logical</th>
<th>Behavioral</th>
</tr>
</thead>
<tbody>
<tr>
<td>Creation of $e^-/h^+$ pairs.</td>
<td>Current or voltage pulse.</td>
<td>Bit flips, signal inversion.</td>
<td>Erroneous transitions.</td>
</tr>
</tbody>
</table>
Observable impacts on an inverter

\[ V_{dd} \]  
\[ Gnd \]  
\[ '0' \]  
\[ '1' \]

\[ V_{dd} \]  
\[ Gnd \]  
\[ '1' \]  
\[ '0' \]  
\[ Cload \]
SRAM cell (two inverters)
Laser beam impact (if ‘1’, $\rightarrow 0$ / if ‘0’, $\rightarrow 1$)
Laser cartography
Fault model classification

- **Global Impact**
  - transient: SET
  - transient: SEU ⇒ useful fault to attack
  - permanent: as SEL, destructive

- **Number of impacted bits**
  - single bit
  - a few bits (word)

- **Type**
  - stuck-at, ’0’ or ’1’
  - Bit-flip
  - Random

- **Accuracy**
  - location
  - timing
  - energy
Fault injection intrusivity

- Non-invasive
  - No preparation
  - low-cost equipment
  - Examples: global fault, clock glitch, temperature...

- Semi-invasive
  - Unpackaging
  - Preparation: etching, cleaning
  - Affordable equipment
  - Examples: laser shots, near field EM

- Invasive
  - Unpackaging
  - Electrical contact with the heart of the chip
  - Expensive
  - Examples: Focus Ion Beam
Outline

Introduction

Fault Injection Attacks examples

DFA on AES

Countermeasures
Outline

Introduction
  Physical Faults
  Faults Models and Tolerance

Fault Injection Attacks examples
  Fault Attack on the control part
  Differential Fault Attack on cryptographic functions

DFA on AES
  Theoretical Fault Attacks
  Practical Attacks

Countermeasures
  Redundancy
  Resilience
  Detection
Reference: R. Anderson et al ”Tamper Resistance – a Cautionary Note” in USENIX 1996[AK96]:

```
1 b = answer_address
2 a = answer_length
3 if (a == 0) goto 8
4 transmit(*b)
5 b = b + 1
6 a = a - 1
7 goto 3
8 ...
```

- Fault Target: Conditionnal jump in line 3 not checked or decrementation in line 6 not executed.
- Attack goal: complete memory dump.

The same fault type can be applied at the PIN verify step of a smart card authentication.
Many more fault models in software

- Instruction skip
- Instruction change
- data change
- ...
- stall (most of the time but not exploitable)
Outline

Introduction
  Physical Faults
  Faults Models and Tolerance

Fault Injection Attacks examples
  Fault Attack on the control part
  Differential Fault Attack on cryptographic functions

DFA on AES
  Theoretical Fault Attacks
  Practical Attacks

Countermeasures
  Redundancy
  Resilience
  Detection
The Differential Fault Attack

- **Principle:**
  - one computation is without fault
  - at least one is faulty
  - the attack algorithms exploit the difference between the faulty and non-faulty ciphertexts

- **DFA targets** almost whatever algorithm (*known or even unknown*)
  - It works on complex bit operations, such as the ones involved in secret key cryptography

The So-Called “Bellcore” Attack

Bellcore = Bell Communications Research

- Three employees of Bellcore (and Pr. @ Stanford) find an attack that breaks RSA by injecting a single fault.


Bellcore attack against RSA signatures with CRT

- Signature $S$ of message $x$: $S \doteq x^d \mod N$, with $N = p \cdot q$, $d =$ private key.

- Using Chinese Remainder Theorem (CRT), the signature can be simplified as:
  - $S_1 = x^d \mod (p-1) \mod p$ and
  - $S_2 = x^d \mod (q-1) \mod q$, both operations working on half bitwidth.

- The signature is obtained back using the two constants:
  \[
  \begin{align*}
  a &= 0 \mod q \\
  a &= 1 \mod p \\
  b &= 1 \mod q \\
  b &= 0 \mod p
  \end{align*}
  \]

- $S = a \cdot S_1 + b \cdot S_2 \mod N$.

- Now, if $S_1$ happens to be faulty: $S_1 \rightarrow \hat{S}_1$ for whatever reason,

- $\gcd(S - \hat{S}, N) = \gcd(a \cdot (S_1 - \hat{S}_1), N) = q$. 
DFA on DES

DFA Assumptions

- Unrolled implementation.
- Single bit-flips on any right register $R_i$, for $i \in [1, 16]$.
- Ciphertext-only attack.

DES Properties

- All the DES constitutive boxes, but the S, are linear:
  $f(x \oplus a) = f(x) \oplus f(a)$, for $f \in \{\text{Id, P, E, FP}\}$.
- $L_{16} = R_{15}$. 

Unrolled implementation.
DFA on DES: One Fault Occurs in $R_{15}$

What Has Happened?

Bit $b \in [1, 32]$ of $R_{15}$ is flipped.

Attack Scenario

- Find $b$ by looking in $L_{16}$.
- Deduce which $S_i$, $i \in [1, 8]$, (there can be two of them) has output a wrong value.
- Solve the equation couple:
  \[
  \begin{align*}
  R_{16} &= L_{15} \oplus S_i(K_{16} \oplus R_{15}), \\
  \tilde{R}_{16} &= L_{15} \oplus S_i(K_{16} \oplus \tilde{R}_{15}).
  \end{align*}
  \]
- It has $\approx$ four 6-bit solutions.
DFA one DES: One Fault Occurs in $R_{14}$

What Has Happened?
Bit $b \in [1, 32]$ of $R_{14}$ is flipped.

Attack Scenario

- Previous attack on $R_{15}$ allowed a straightforward subkey $K_{16}$ retrieval at the input of $(b - 1)/8^{th}$ S-box.
- Current attack requires a differential analysis of the last two rounds of DES.
- Details to come...
Solving the “$R_{14}$ bit flip” plot (1/2)

Notations

- Tilded symbols, e.g. $\tilde{L}_{16}$, denote faulted quantities.
- $R_{14} \oplus \tilde{R}_{14} = 00 \cdots 010 \cdots 00 \oplus 1_b$, the “1” lying at position $b$.

Which bit $b$ was flipped?

- Notice that $\Delta = L_{16} \oplus \tilde{L}_{16}$ is also the difference at the output of $S$, at round 15.
- For each S-box $i$, $S_i(x) \oplus S_i(x \oplus 1_b) = \Delta$, $x$ being the unknown value $(R_{14} \oplus K_{15})[8 \cdot i, 8 \cdot (i + 1)]$, has few solutions $b$.
- Validate potential $b$ by verifying that $\Delta$ passed through $S$, at round 16, can generate the difference $R_{16} \oplus (\tilde{R}_{16} \oplus 1_b)$.
Solving the “$R_{14}$ bit flip” plot (2/2)

Retrieving Information on $K_{16}$ subkey

- Now that flipped bit $b$ in $R_{14}$ is known, the differences before and after S-boxes in round 16 are known.
- This property allows to eliminate many subkeys $K_{16}$ 6-bit parts at the input of activated S-boxes (6 out of 8.)

Attack Extension

- Basically the same differential attack can be used if the error occurs in round 14 (but not higher...).
- Not surprisingly, Eli Biham and Adi Shamir, inceptors of the DFA, are also the fathers of the differential cryptanalysis.

\[\rightarrow \text{“Differential Cryptanalysis of the Full 16-Round DES”, CS 708, December 1991, Proceedings of Crypto’92, LNCS 740.}\]
The DFA Efficiency

A Powerful Attack!

- According to authors, between 50 and 200 faults on whatever round are required to fully expose the last round subkey.
- Once $K_{16}$ is known, the key $K$ can be retrieved by an exhaustive search attacks on the $56 - 48 = 8$ remaining bits.

Generalization

- If $K_{16}$ is known, the DFA can be applied to the 15-round DES variant...
- The rounds are peeled off (and detected faults corrected).
- Thus, Triple-DES and DES with independent subkeys (768 bit) can be attacked.
The DFA Efficiency

Number of faulted ciphertexts ($C'$) to disclose the key (best attacker)

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Key space</th>
<th># $C'$</th>
<th>Fault model</th>
</tr>
</thead>
<tbody>
<tr>
<td>RSA (CRT) [BDL97]</td>
<td>$2^{1024}$</td>
<td>1</td>
<td>Any @ $S_p$ (or $S_q$)</td>
</tr>
<tr>
<td>RSA (L2R) [BDH$^+$97]</td>
<td></td>
<td>3083</td>
<td>Bit error @ each S&amp;M</td>
</tr>
<tr>
<td>DES [BS97]</td>
<td>$2^{56}$</td>
<td>11</td>
<td>Any @ 12$^{th}$ round</td>
</tr>
<tr>
<td>AES-256 [TMA11]</td>
<td>$2^{256}$</td>
<td>2</td>
<td>Byte error @ 8$^{th}$ round</td>
</tr>
<tr>
<td>ECDSA/P-192 [BBB$^+$11]</td>
<td>$2^{192}$</td>
<td>36</td>
<td>Any in key $d$ @ MULT</td>
</tr>
</tbody>
</table>
Outline

Introduction

Fault Injection Attacks examples

DFA on AES

Countermeasures
Outline

Introduction
- Physical Faults
- Faults Models and Tolerance

Fault Injection Attacks examples
- Fault Attack on the control part
- Differential Fault Attack on cryptographic functions

DFA on AES
- Theoretical Fault Attacks
- Practical Attacks

Countermeasures
- Redundancy
- Resilience
- Detection
AES Implementation

Plaintext (128 bits)

roundkey(0)

for i=1 to 9

SubBytes

ShiftRows

MixColumns

roundkey(i)

SubBytes

ShiftRows

roundkey(10)

Ciphertext (128 bits)
State-of-the-Art of Fault Attacks against AES

- Giraud in 2003: (50 faults) \[\text{[Gir04]}\]
- Dusart, Letourneux & Vivolo in 2002: \((5 \times 4\) faults) \[\text{[DLV03]}\]
- Piret & Quisquater in 2004: 2 faults \[\text{[PQ03]}\]
- Tunstall, Mukhopadhyay & Ali: 1 fault \[\text{[TMA11]}\]
Bit-fault $e_j$ attack on the last round

Regular encryption ($C$):
- $C_{\text{ShiftRow}}(i) = \text{SubBytes}(M_i^9) \oplus K_{\text{ShiftRow}}^{10}$ for $i \in [1, 16]$

Faulted encryption ($D$):
- $D_{\text{ShiftRow}}(i) = \text{SubBytes}(M_i^9) \oplus K_{\text{ShiftRow}}^{10}$ for $i \in [1, 16] \setminus \{j\}$ and
- $C_{\text{ShiftRow}}(j) = \text{SubBytes}(M_j^9 \oplus e_j) \oplus K_{\text{ShiftRow}}^{10}$.

Attack:
- $C_{\text{ShiftRow}}(j) \oplus D_{\text{ShiftRow}}(j) = \text{SubBytes}(M_j^9) \oplus \text{SubBytes}(M_j^9 \oplus e_j)$. 
Detail of Ch. Giraud’s attack

Goal: finding the value of \( M_j^9 \).

- \( C_{\text{ShiftRows}}(j) \oplus D_{\text{ShiftRows}}(j) = \Delta = \text{SubBytes}(M_j^9) \oplus \text{SubBytes}(M_j^9 \oplus e_j) \) has between 2 and 14 solutions in \((e_j, M_j^9)\) (set of \(8 \times 2^8\) unknown), and 8 in average.
- However, the exact value of \( e_j \) is of no importance.

Attack Strength

- Thus, with 2 faults, there is 50 % chance to get one \( M_j^9 \).
- With 3 faults, there is 97 % chance to get one \( M_j^9 \).
- Once \( M_j^9 \) is known, we have \( K_j^{10} = C_j \oplus \text{SubBytes}(M_j^9) \).
One faulty Byte at round 9 generates 4 faulty Bytes at the output. 255x4 candidates for $K_{10}$. 2 faults $\Rightarrow$ 98%. 8 faults $\Rightarrow$ 100%.
A fault at round 8 yields 4 faults at round 9! This is optimal...
Outline

Introduction
  Physical Faults
  Faults Models and Tolerance

Fault Injection Attacks examples
  Fault Attack on the control part
  Differential Fault Attack on cryptographic functions

DFA on AES
  Theoretical Fault Attacks
  Practical Attacks

Countermeasures
  Redundancy
  Resilience
  Detection
Faults injection: Local Over-Clocking

- jitter → ←
- jitter → ←
- jitter → ←
- step
Faults injection: Setup-Time Violation Attack Sketch

\[ V \downarrow \Rightarrow T_{\text{propagation}} \uparrow \]

Setup met

Setup violated
Occurrence (nominal voltage is 3.3 V)
Sbox statistics

% of faults

Sbox

Spatial localization

% of faults

S0 S1 S2 S3 S4 S5 S6 S7 S8 S9 S10 S11 S12 S13 S14 S15

51/67

Institut Mines-Télécom

Fault Injection Attacks Jean-Luc Danger
Outline

Introduction
  Physical Faults
  Faults Models and Tolerance

Fault Injection Attacks examples
  Fault Attack on the control part
  Differential Fault Attack on cryptographic functions

DFA on AES
  Theoretical Fault Attacks
  Practical Attacks

Countermeasures
  Redundancy
  Resilience
  Detection
Authenticated encryption is one protocol-level solution (see CAESAR competition).
Redundance in information [KKT04]
A protection against instruction-skip attacks

Instruction skip attacks and protection [HMER13, MHER14]

- One instruction can be skipped
- Replacement sequences for idempotent instructions:

<table>
<thead>
<tr>
<th>Instruction</th>
<th>Description</th>
<th>Replacement</th>
</tr>
</thead>
<tbody>
<tr>
<td>mov r1,r8</td>
<td>Copies r8 into r1</td>
<td>mov r1,r8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>mov r1,r8</td>
</tr>
<tr>
<td>ldr r1,[r8,r2]</td>
<td>Loads the value at the address</td>
<td>ldr r1,[r8,r2]</td>
</tr>
<tr>
<td></td>
<td>r8+r2 into r1</td>
<td>ldr r1,[r8,r2]</td>
</tr>
<tr>
<td>str r3,[r2,#10]</td>
<td>Stores r3 at the address</td>
<td>str r3,[r2,#10]</td>
</tr>
<tr>
<td></td>
<td>r2+10 into r1</td>
<td>str r3,[r2,#10]</td>
</tr>
<tr>
<td>add r3,r1,r2</td>
<td>Puts r1+r2 into r3</td>
<td>add r3,r1,r2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>add r3,r1,r2</td>
</tr>
</tbody>
</table>

Fault injection [RG14, Appendix 2.A]

- Pulse amplitude: ... 0 – 500 mV.
- Pulse duration: .............. 2 ns.
- Repeatability: ........... 500 MHz.
Outline

Introduction
  Physical Faults
  Faults Models and Tolerance

Fault Injection Attacks examples
  Fault Attack on the control part
  Differential Fault Attack on cryptographic functions

DFA on AES
  Theoretical Fault Attacks
  Practical Attacks

Countermeasures
  Redundancy
  Resilience
  Detection
Resilience

- Let the system output erroneous errors, as long as they convey no information about the internal sensitive values [GSDS10]
- at protocol level
- infective computation
- at logic level: DPL
Example: DPL

2 Networks: T and F

2 phases
Outline

Introduction
  Physical Faults
  Faults Models and Tolerance

Fault Injection Attacks examples
  Fault Attack on the control part
  Differential Fault Attack on cryptographic functions

DFA on AES
  Theoretical Fault Attacks
  Practical Attacks

Countermeasures
  Redundancy
  Resilience
  Detection
Digital Sensor [SBGD11]

$t_{\text{chain}} (N \text{ buffers}) > t_{\text{crit}}$

$I \rightarrow 0 \rightarrow 0 \rightarrow 0 \rightarrow 0 \rightarrow 1 \rightarrow O$

Monitoring DFF

error

$I = O$
Design Example: Frequency Sensitivity

- Nominal Sensor Trip Frequency
- Nominal Design Margin (~ 5 nSec)
- Design Frequency
- Less Sensitive – More Missed Events
- More Sensitive - More False Alarms
- Design Choice: 55 Lcells (~20 nSec)
2. Main technical characteristics

- Digital, hence:
  - Simple API
  - Stable
  - Small
  - Discreet, more difficult to recognize
  - Melted within the rest of the SoC, more difficult to by-pass
EM pulses should not be too long!
References on Fault Attacks I


References on Fault Attacks II


References on Fault Attacks III


References on Fault Attacks IV


References on Fault Attacks V