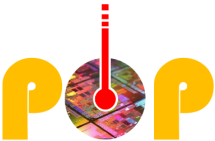
Advanced techniques for fault injection attacks on integrated circuits

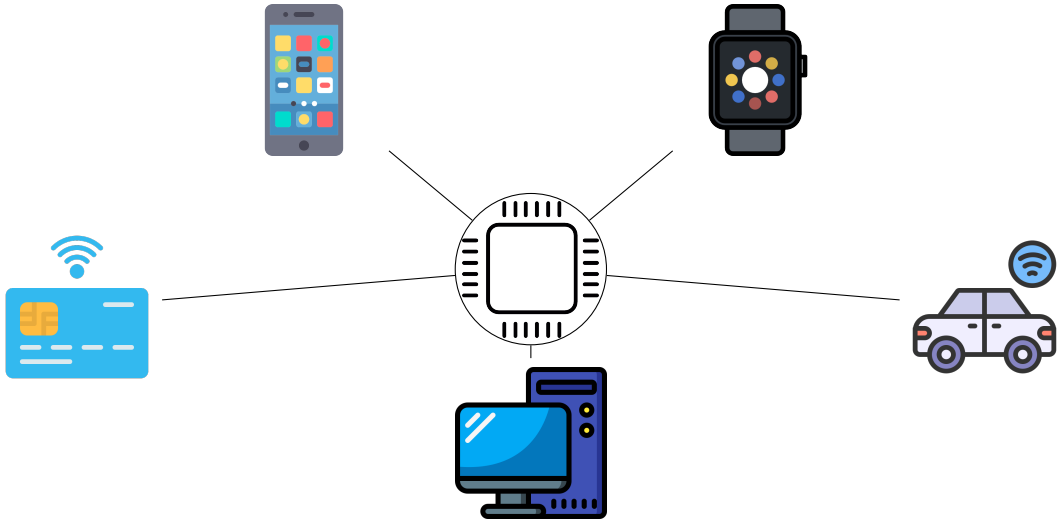
Paul GRANDAMME

April 25th 2025

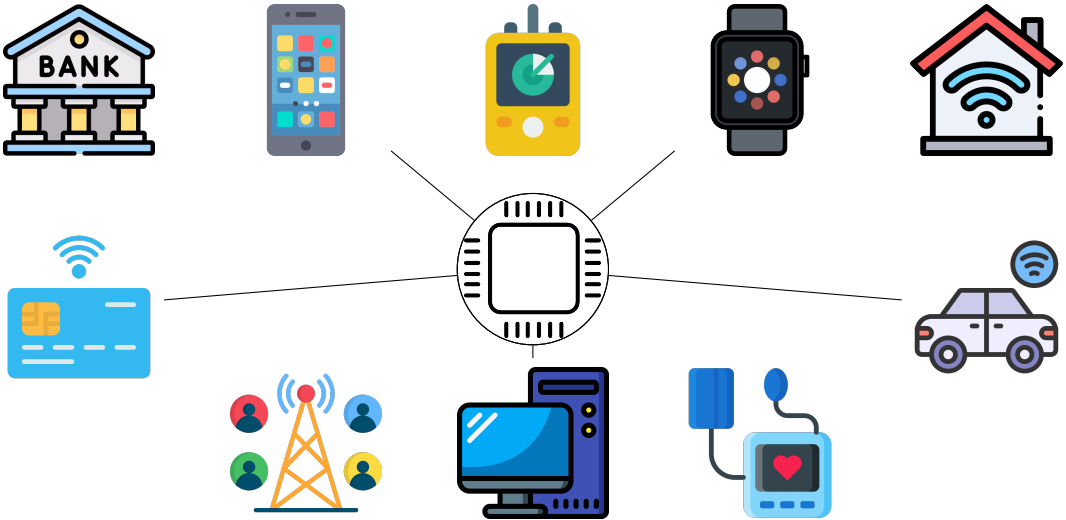


Context

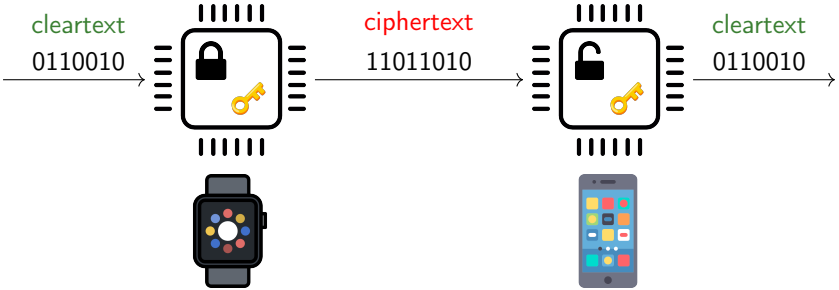
Context



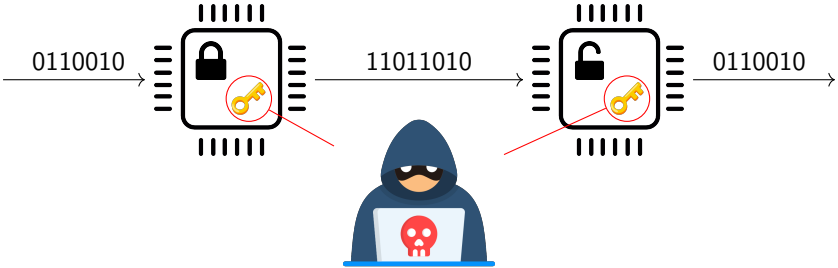
Context



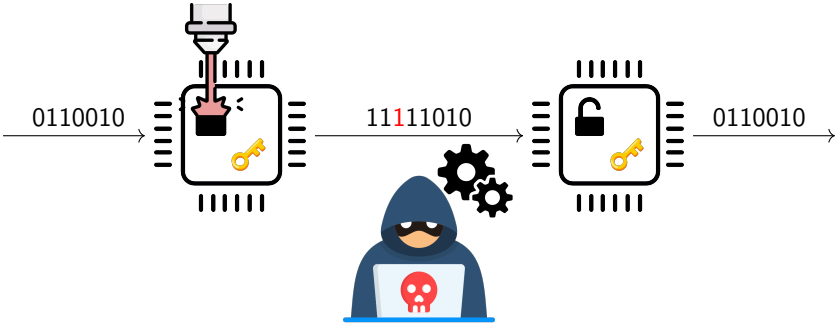
Secure communications



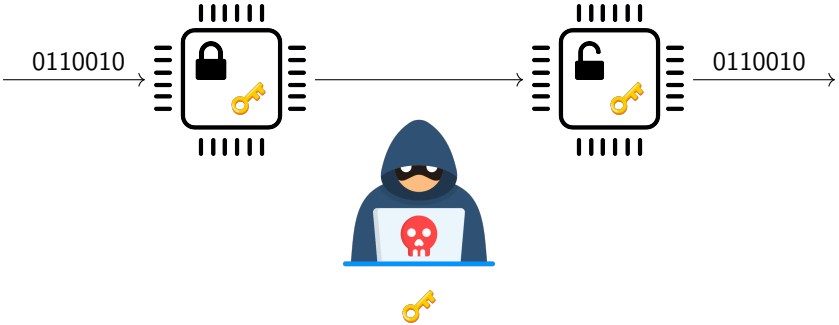
Secure communications



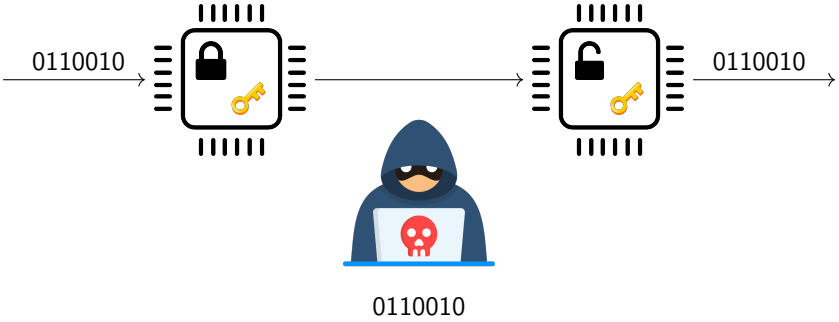
Secure communications



Secure communications



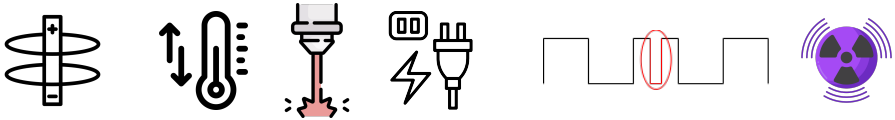
Secure communications



Fault attack

Definition

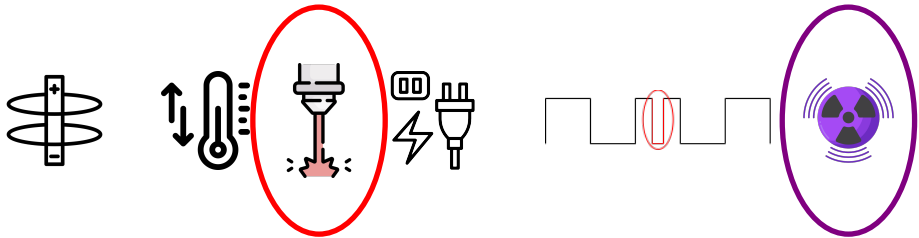
Disturbing the integrated circuit to **modify its behavior** in order to obtain information or disable internal protection mechanisms.



Fault attack

Definition

Disturbing the integrated circuit to **modify its behavior** in order to obtain information or disable internal protection mechanisms.

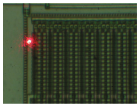


Chronology

2002 2009

Photoflash photo[1]

UV lamp[2] Laser[3]



[1] Sergei P. Skorobogatov and Ross J. Anderson. "Optical Fault Induction Attacks". In: *CHES 2002*.

[2] Jörn-Marc Schmidt, Michael Hutter, and Thomas Plos. "Optical Fault Attacks on AES: A Threat in Violet". In: *FDTC 2009*.

[3] Sergei P. Skorobogatov. "Local Heating Attacks on Flash Memory Devices". In: *IEEE HOST*. 2009.

[4] Brice Colombier et al. "Laser-induced Single-bit Faults in Flash Memory: Instructions Corruption on a 32-bit Microcontroller". In: *IEEE HOST 2019*.

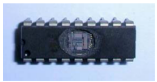
[5] Alexandre Menu et al. "Single-bit Laser Fault Model in NOR Flash Memories: Analysis and Exploitation". In: *FDTC 2020*.

[6] Brice Colombier et al. "Multi-Spot Laser Fault Injection Setup: New Possibilities for Fault Injection Attacks". In: *CARDIS*. 2022.

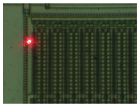
Chronology

2002 2009 2019 2020

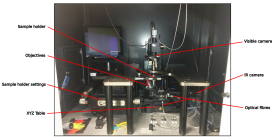
Photoflash photo[1]



UV lamp[2] Laser[3]



Full comprehension of the fault model[4,5]



[1] Sergei P. Skorobogatov and Ross J. Anderson. "Optical Fault Induction Attacks". In: *CHES 2002*.

[2] Jörn-Marc Schmidt, Michael Hutter, and Thomas Plos. "Optical Fault Attacks on AES: A Threat in Violet". In: *FDTC 2009*.

[3] Sergei P. Skorobogatov. "Local Heating Attacks on Flash Memory Devices". In: *IEEE HOST*. 2009.

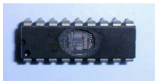
[4] Brice Colombier et al. "Laser-induced Single-bit Faults in Flash Memory: Instructions Corruption on a 32-bit Microcontroller". In: *IEEE HOST 2019*.

[5] Alexandre Menu et al. "Single-bit Laser Fault Model in NOR Flash Memories: Analysis and Exploitation". In: *FDTC 2020*.

[6] Brice Colombier et al. "Multi-Spot Laser Fault Injection Setup: New Possibilities for Fault Injection Attacks". In: *CARDIS*. 2022.

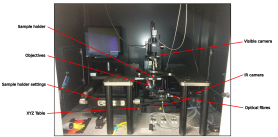
Chronology (CARDIS 2022)

2002 — Photoflash photo[1]

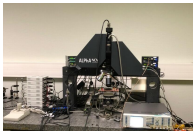


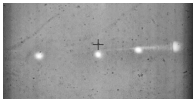
2009 — UV lamp[2] Laser[3]

2019 2020 — Full comprehension of the fault model[4,5]



2022 — Multispot laser bench[6]





ALPhANOV
Optics & Lasers Technology Center

[1] Sergei P. Skorobogatov and Ross J. Anderson. "Optical Fault Induction Attacks". In: *CHES 2002*.

[2] Jörn-Marc Schmidt, Michael Hutter, and Thomas Plos. "Optical Fault Attacks on AES: A Threat in Violet". In: *FDTC 2009*.

[3] Sergei P. Skorobogatov. "Local Heating Attacks on Flash Memory Devices". In: *IEEE HOST*. 2009.

[4] Brice Colombier et al. "Laser-induced Single-bit Faults in Flash Memory: Instructions Corruption on a 32-bit Microcontroller". In: *IEEE HOST 2019*.

[5] Alexandre Menu et al. "Single-bit Laser Fault Model in NOR Flash Memories: Analysis and Exploitation". In: *FDTC 2020*.

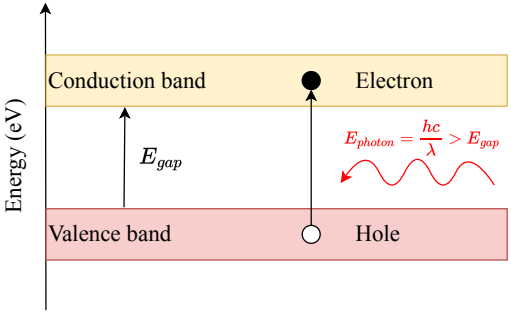
[6] Brice Colombier et al. "Multi-Spot Laser Fault Injection Setup: New Possibilities for Fault Injection Attacks". In: *CARDIS*. 2022.

Photoelectric effect

$$E_{photon} = \frac{hc}{\lambda}$$

Photoelectric effect

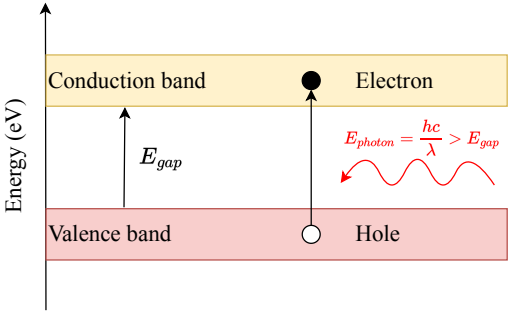
$$E_{photon} = \frac{hc}{\lambda}$$



Photoelectric effect

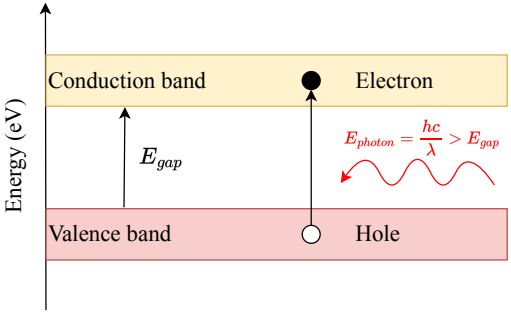
$$E_{photon} = \frac{hc}{\lambda}$$

$$E_{photon} > E_{gap} \Rightarrow \lambda < \frac{hc}{E_{gap}} \approx 1100\,\text{nm}$$



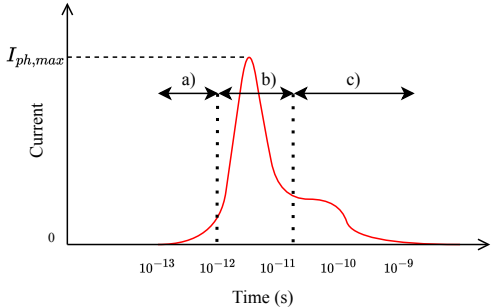
Photoelectric effect

$$E_{photon} = \frac{hc}{\lambda}$$



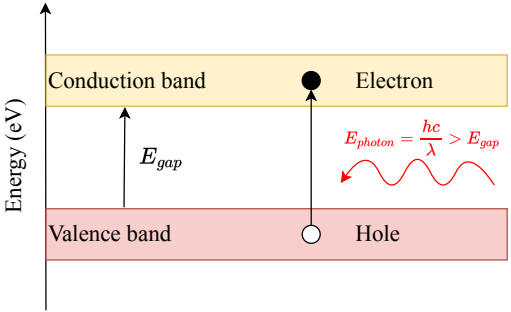
$$E_{photon} > E_{gap} \Rightarrow \lambda < \frac{hc}{E_{gap}} \approx 1100\,\text{nm}$$

If an electrical field is present :



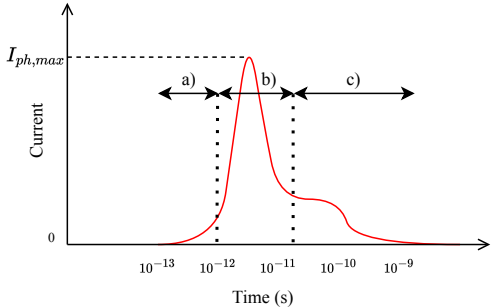
Photoelectric effect

$$E_{photon} = \frac{hc}{\lambda}$$



$$E_{photon} > E_{gap} \Rightarrow \lambda < \frac{hc}{E_{gap}} \approx 1100\,\text{nm}$$

If an electrical field is present :

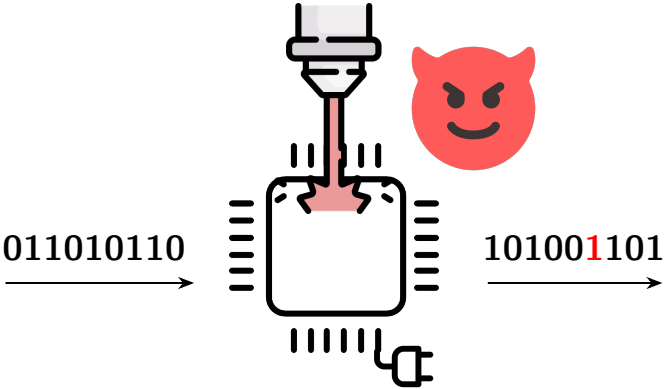


Thermal effect also present at 1064 nm

Context

State of the art

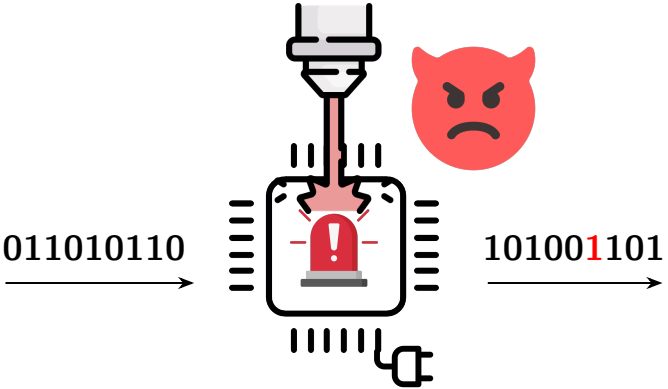
Almost all attacks are carried out on powered devices



Context

State of the art

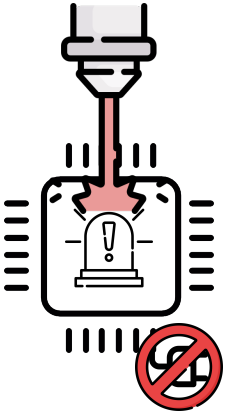
Almost all attacks are carried out on powered devices



Context

State of the art

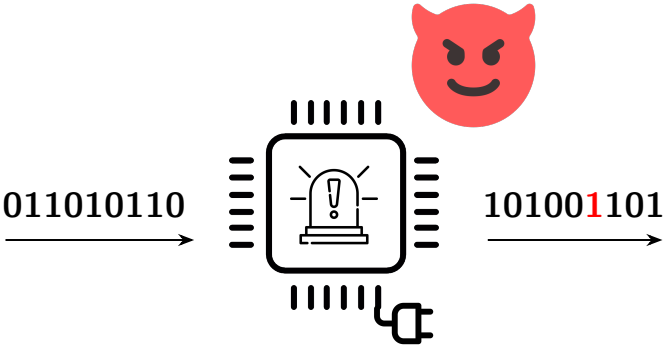
Almost all attacks are carried out on powered devices



Context

State of the art

Almost all attacks are carried out on powered devices



Context

Problem

Can laser faults be injected into unpowered devices?

POP project

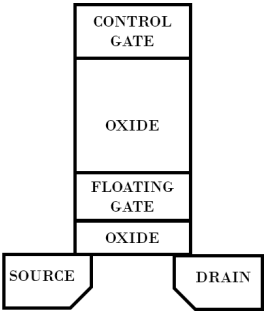
Attack unpowered devices

- No detection is possible ⇒ No reaction possible
- No synchronisation required

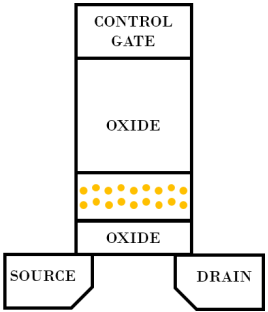
Idea

- Corrupt stored data : non-volatile memories (Flash)

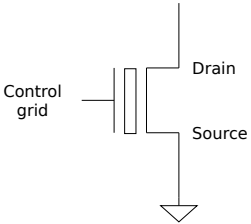
Floating gate transistors



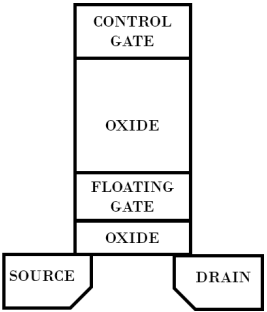
Discharged cell.



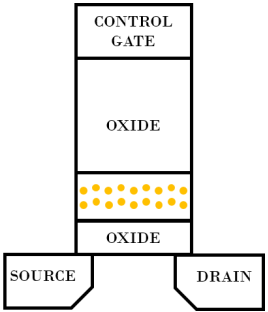
Charged cell.



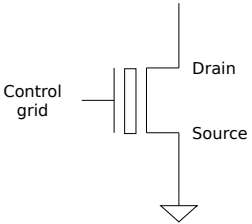
Floating gate transistors



Discharged cell.



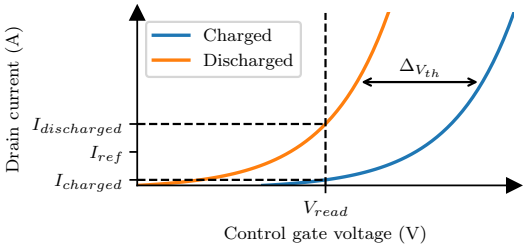
Charged cell.



Convention :

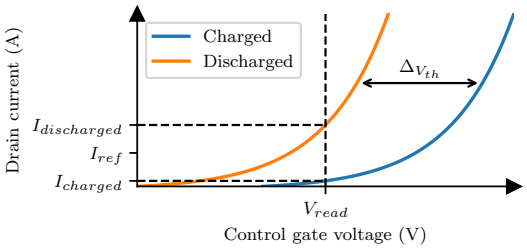
- Charged = ‘0’
- Discharged = ‘1’

Read mechanism

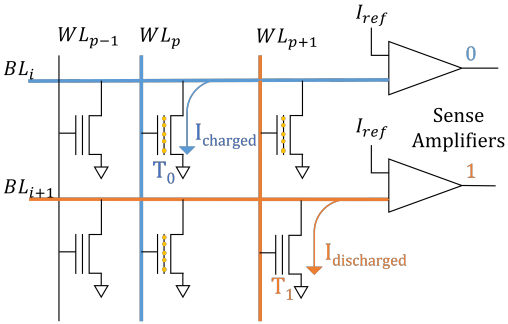


I-V characteristics

Read mechanism

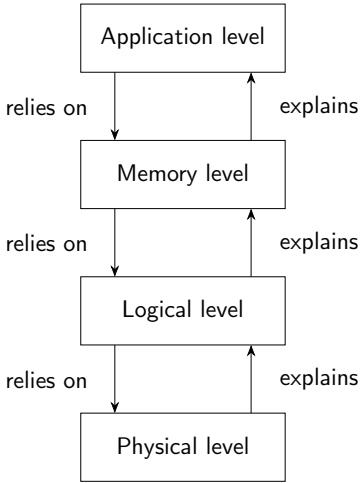


I-V characteristics

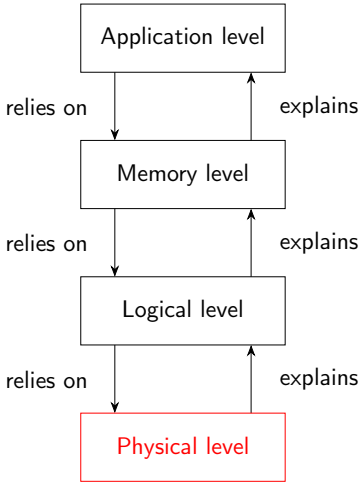


Read mechanism in NOR Flash memories

Abstraction levels



Abstraction levels



Physical level

- Energy of the laser beam

⇒ Temperature rise

⇒ Floating gate transistors discharge[a]

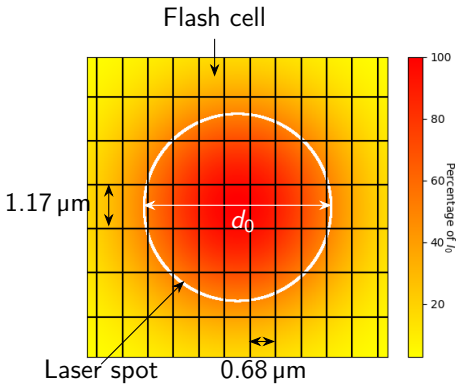
Laser beam intensity :

$$I(r) = I_0 \cdot e^{-\frac{2r^2}{\omega_0^2}} \text{ avec } \omega_0 = \frac{2\lambda}{\pi \times NA}$$

FWHM criterion ⇒

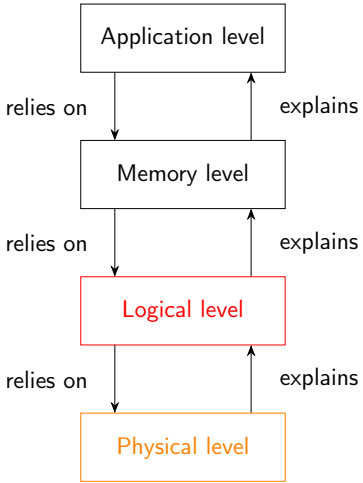
$$d_0 = \omega_0 \sqrt{\frac{\ln 2}{2}} \approx 5\,\mu\text{m}$$

[a] Sergei P. Skorobogatov. "Local Heating Attacks on Flash Memory Devices". In: *IEEE HOST*. 2009.



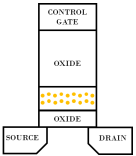
Heatmap induced by laser exposure (numerical simulation with $\lambda = 1064\,\text{nm}$ et $NA = 0.16$)

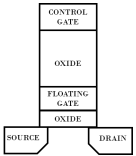
Abstraction levels



Logical level

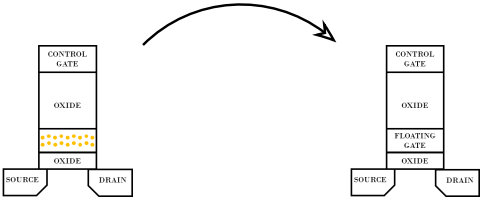
From the physical level :





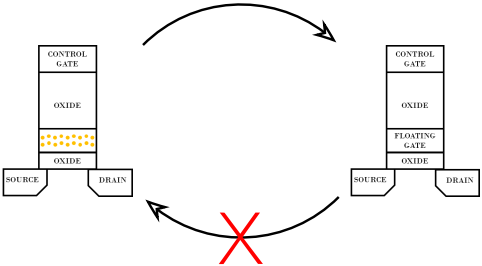
Logical level

From the physical level :



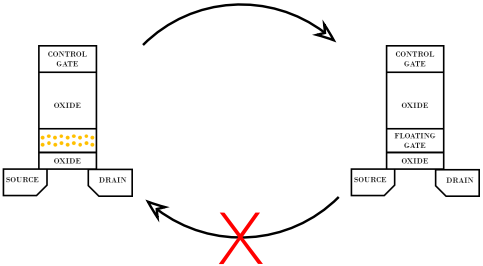
Logical level

From the physical level :



Logical level

From the physical level :

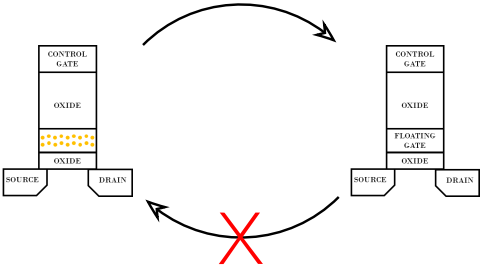


To the logical level :

‘0’ ‘1’

Logical level

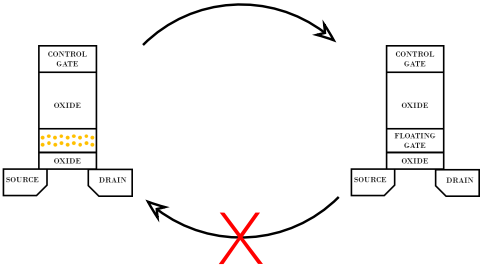
From the physical level :



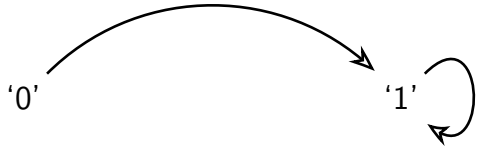
To the logical level :

Logical level

From the physical level :

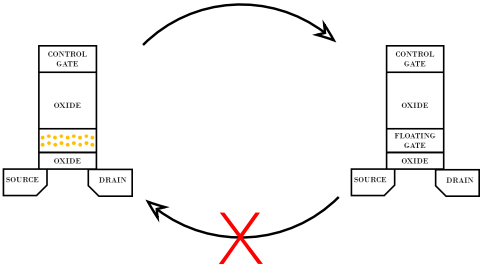


To the logical level :

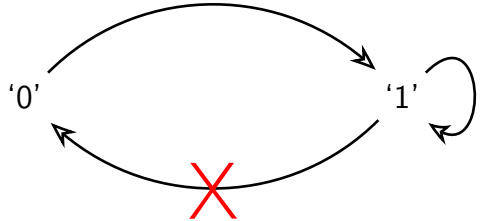


Logical level

From the physical level :

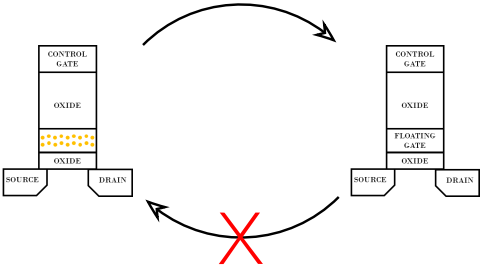


To the logical level :

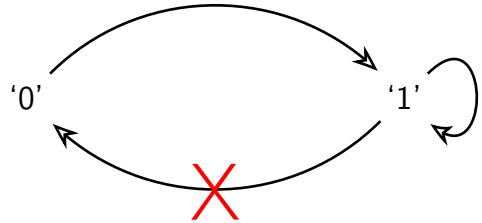


Logical level

From the physical level :



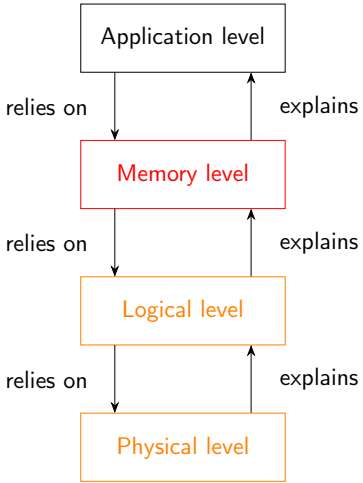
To the logical level :



⇒ Data-dependant and unidirectional fault model

- ⇒ Bitsets for our target
- ⇒ Permanent faults (Flash memory can be reprogrammed)

Abstraction levels



Memory level

Flash memory

- Used in particular to store code, constants, etc.

Memory level

Flash memory

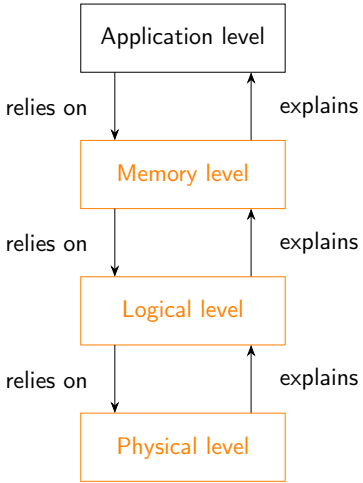
- Used in particular to store code, constants, etc.

Source code

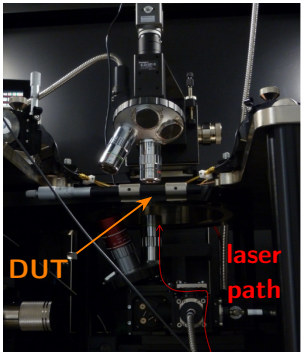
Stored in Flash

- Fault injection ⇒ Code corruption

Abstraction levels



Laser injection bench



Characteristics

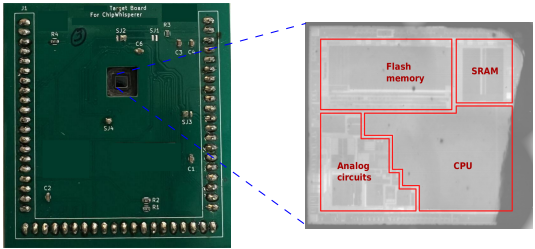
- Laser source 1064 nm (near-IR)
- Laser pulse of 0.9 s
- Spot of 5 µm with x20 magnification
- Backside view possible thanks to infrared camera
- Laser bench commonly used for laser injection on powered circuits

MINES Saint-Étienne
Une école de l'IMT

Experimental material

Target

- 32-bit microcontroller (CMOS 80 nm) dedicated to IoT applications
- ARM Cortex-M3 core
- 128 kB of Flash memory (128 pages of 1 kB)
- Open on the backside for access to the substrate



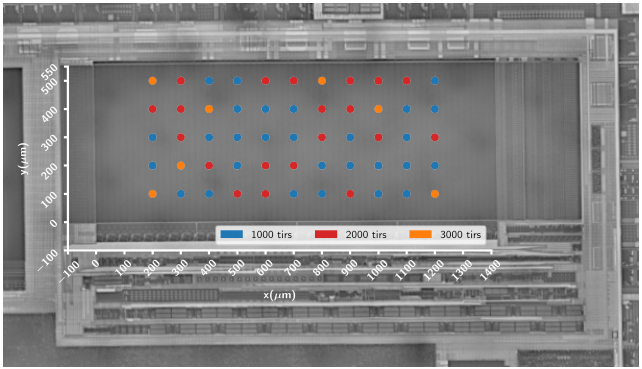
Injected faults mapping

Require: $X_{min}, X_{max}, X_{step}, Y_{min}, Y_{max}, Y_{step}$

1: **for** $x \in \texttt{range}(X_{min}, X_{max}, X_{step})$ **do**
2: **for** $y \in \texttt{range}(Y_{min}, Y_{max}, Y_{step})$ **do**
3: Memory initialization
4: **do**
5: Move the laser to (x,y)
6: Switch off the target
7: **for** $i \in [0, \ldots, 999]$ **do**
8: Laser shot
9: Switch on the target
10: Flash memory read
11: **while** #faults == 0
12: mapping[x][y] = #faults
13: **return** mapping[x][y]

Laser sensitivity map

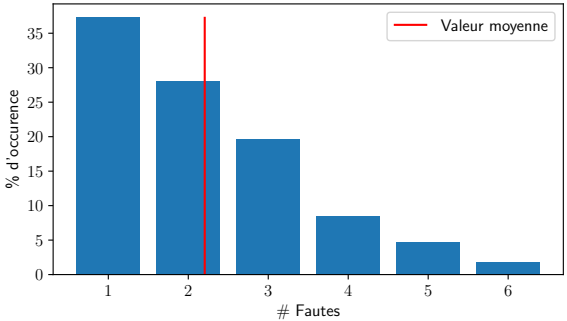
- Memory initialized to 0x00000000 (programmed) before laser exposure

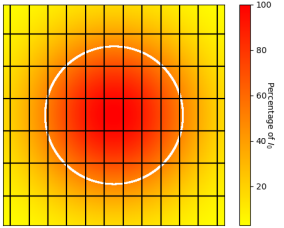


Mapping of injected faults. $P_{laser} = 1\,\mathrm{W}$, $f_{laser} = 1\,\mathrm{Hz}$, $T_{pulse} = 0.9\,\mathrm{s}$

$\Rightarrow$ Values and addresses of known faults (bitsets only)

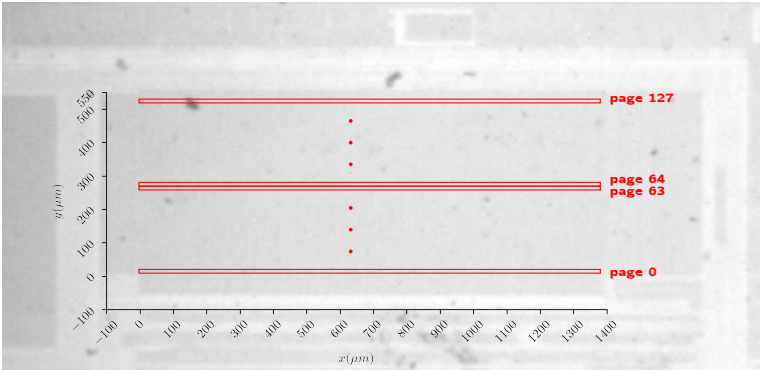
Experimental distribution





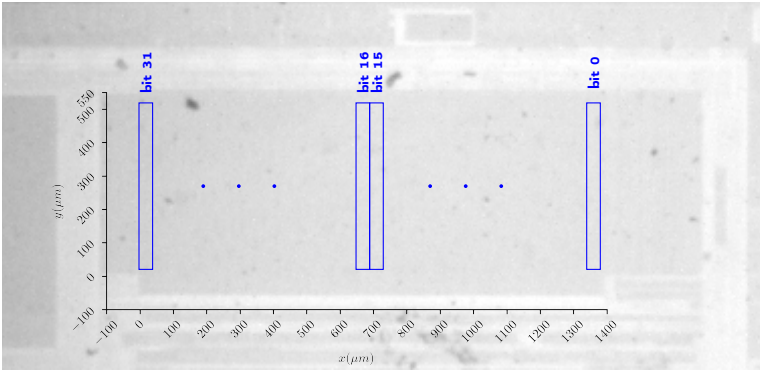
- 2.2 faulty bits on average on a memory initialized at 0x00000000
- Monobit faults in 33% of cases
- No correlation between #shots and #faults

Reverse engineering 1/2



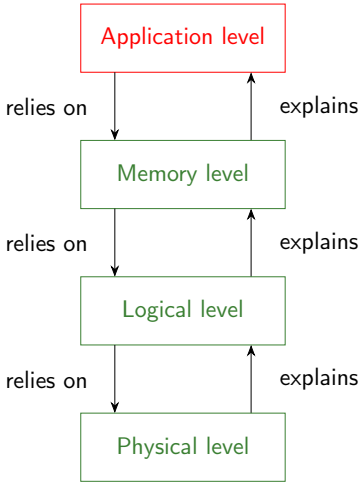
Reverse engineering at the page level

Reverse engineering 2/2



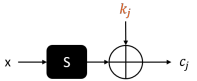
Reverse engineering at the bit level

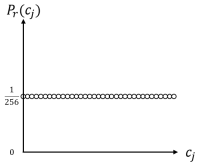
Abstraction levels



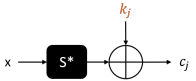
Persistant Fault Analysis (PFA) : Theory[7] (CHES 2018)

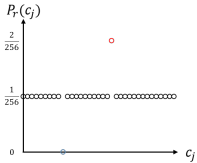
- Permanent fault injection in the substitution box (S-Box)
  - AES : 256 known values stored in Flash memory
- Statistical study of the bytes in ciphertext
- Exploiting the bias to find the encryption key





Without fault

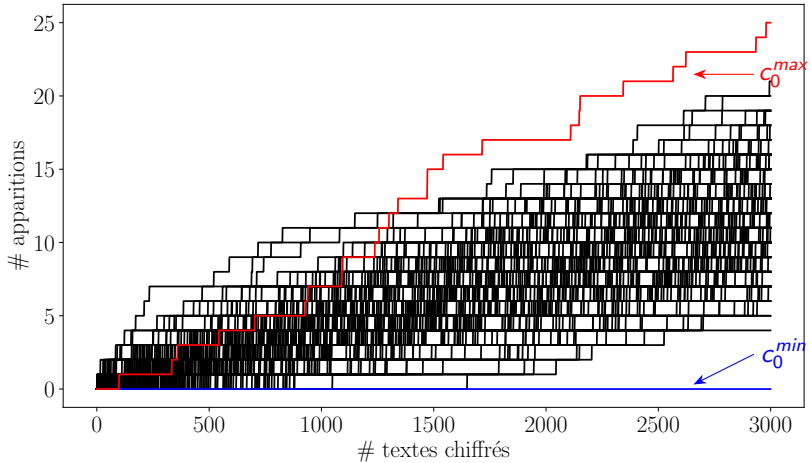




With one faulty byte

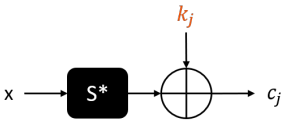
[7] Fan Zhang et al. "Persistent Fault Analysis on Block Ciphers". In: *IACR TCHES*. (2018).

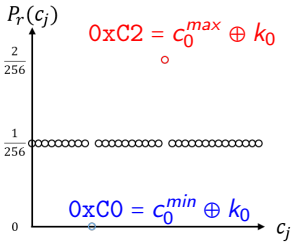
Appearance of byte values



Experimental results (IACR TCHES 2024)

1. Reverse engineering of the *firmware*
   ⇒ S-box stored between addresses 0x080012F4 and 0x080013F3
2. From the previous mapping
   ⇒ the position (x,y) = (44.3,300) is in the S-box range
3. Laser fault injection on unpowered devices
   ⇒ fault 0x00000002 at addresse 0x08001310
   32^{th} value of the S-box (0xC2 instead of 0xC0)

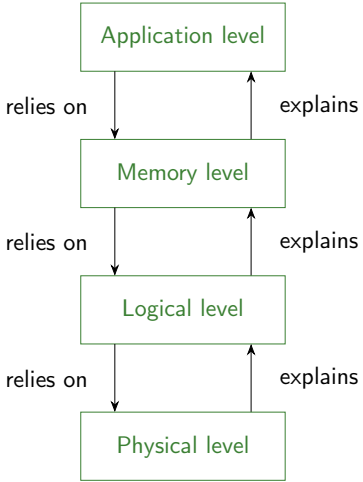




Successful PFA

First experimental use of PFA on an unpowered device

Abstraction levels



Conclusion

Synthesis

- Fault injection in unpowered components
- Possibility of injecting localized faults into Flash memories
- Single-bit, permanent, bitset faults
- Description of a complete fault model
- Application: Persistent Fault Analysis (PFA)

Consequences

- Securing circuits dedicated to IoT applications (EDAC codes)
- Increased interest in the PFA
  - Countermeasures specific to the PFA[8]

[8] Pierre-Antoine Tissot, Lilian Bossuet, and Vincent Grosso. "BALoo: First and Efficient Countermeasure Dedicated to Persistent Fault Attacks". In: *IEEE IOLTS*. 2023.

State of the art

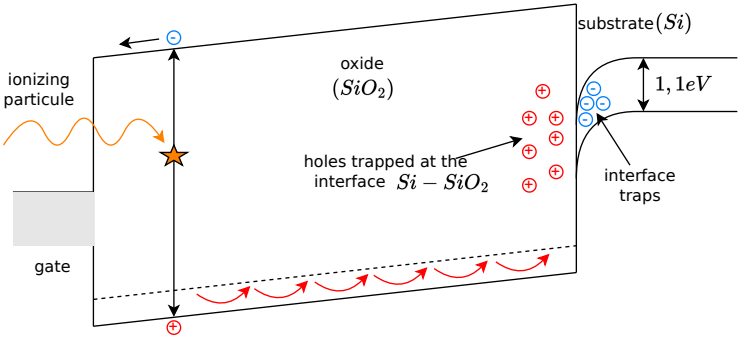
- Floating gate transistors erasure on powered devices thanks to X-ray exposure[9,10,11]

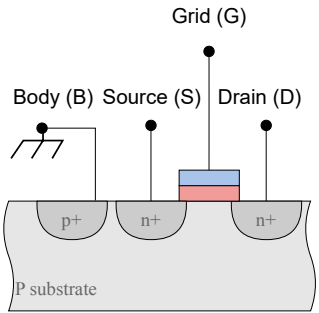
Objective

- Evaluate the possibility to inject faults on unpowered devices

[9] Stéphanie Anceau et al. "Nanofocused X-Ray Beam to Reprogram Secure Circuits". In: *IACR TCHES 2017*.
[10] Laurent Maingault et al. "Laboratory X-rays Operando Single Bit Attacks on Flash Memory Cells". In: *CARDIS 2021*.
[11] S. Bouat et al. "X ray nanoprobe for fault attacks and circuit edits on 28-nm integrated circuits". In: *IEEE DFT 2023*.

Total Ionizing Dose on MOS transistors[12]

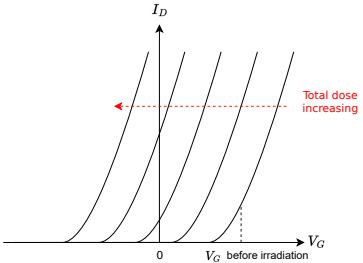




Energy band representation of a cross-sectionnal view of a MOS transistor

[12] H. J. Barnaby. "Total-Ionizing-Dose Effects in Modern CMOS Technologies". In: *IEEE Transactions on Nuclear Science* (2006).

TID effect on MOS characteristics[13]



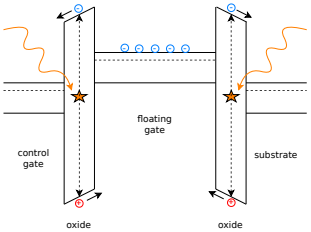
NMOS case

Consequences

- NMOS transistors become more easily conducting, even permanently conducting
- PMOS transistors become less easily conducting, even permanently blocking

[13] Ashok K. Sharma. *Semiconductor Memories: Technology, Testing and Reliability*. 2002.

TID effect on floating gate transistors[14]

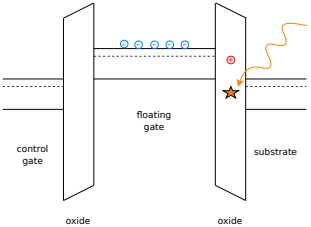


First effect

- e^+/h^- pairs creation by radiation
- Pairs separation by the electric field
- Electrons evacuation through the control gate
- Holes injection in the floating gate
- Recombinaison with the stored charges
- Decrease of the stored charge

[14] S. Gerardin et al. "Radiation Effects in Flash Memories". In: *IEEE Transactions on Nuclear Science* (2013).

TID effect on floating gate transistors[14]

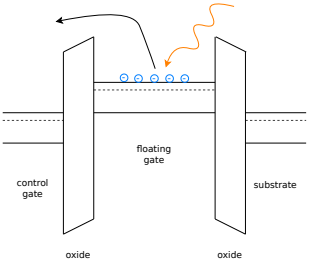


Second effect

- Charge trapping in oxides
- Not very significant given the thickness of the oxides

[14] S. Gerardin et al. "Radiation Effects in Flash Memories". In: *IEEE Transactions on Nuclear Science* (2013).

TID effect on floating gate transistors[14]



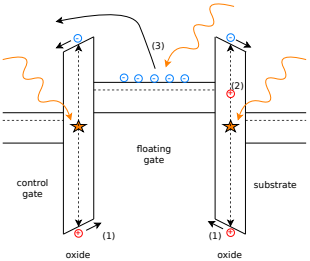
Third effect

- The stored charges obtain enough energy to overcome the potential barrier
- Decrease of the stored charge

$\Rightarrow$ Photoemission

[14] S. Gerardin et al. "Radiation Effects in Flash Memories". In: *IEEE Transactions on Nuclear Science* (2013).

TID effect on floating gate transistors[14]

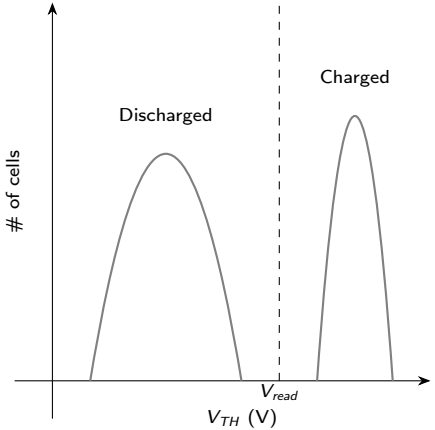


Three different mechanims:

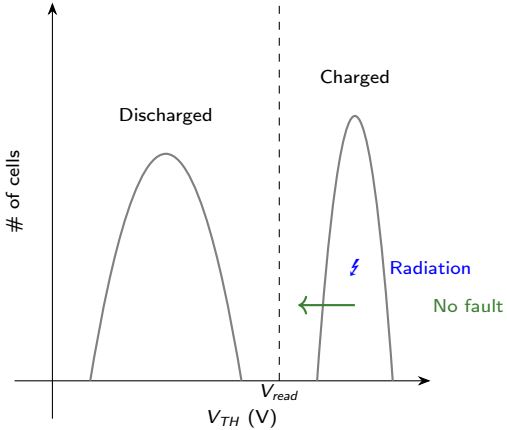
- Electron/hole pair generation in the oxides
- Charge trapping in the oxides
- Photoemission

[14] S. Gerardin et al. "Radiation Effects in Flash Memories". In: *IEEE Transactions on Nuclear Science* (2013).

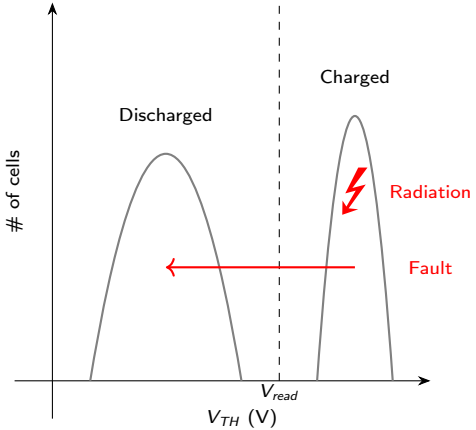
Impact on threshold voltage distribution



Impact on threshold voltage distribution



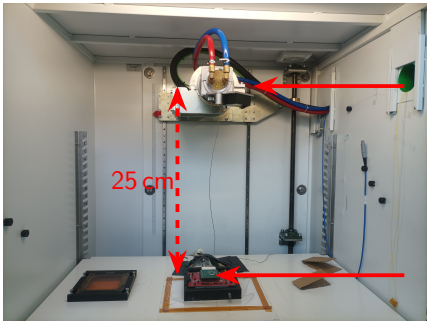
Impact on threshold voltage distribution



Experimental setup

X-ray tube	COMET MXR-165
Maximum voltage	160 kV
Maixmum current	45 mA
Anode material	Tungsten (W)
Anode angle	30°
Beam coverage	50°
Beam filtering	4 mm Beryllium (Be)

LABORATOIRE HUBERT CURIEN
UMR • CNRS • 5516 • SAINT-ETIENNE



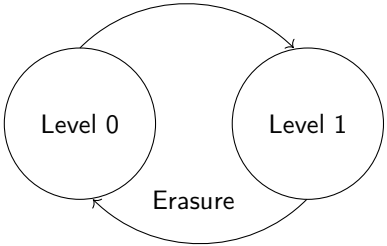
The entire target is irradiated !

Thanks to the MOPERE team Hubert Curien lab!

Target

2 levels of reading protection

- Level 0 : No restriction
- Level 1 : Flash read impossible when the *debugger* is connected



RDP	nRDP	Status	Level
0xFF	0xFF	Protected	1
0xA5	0x5A	Unprotected	0
0xXY	$\neq \overline{\texttt{0xXY}}$	Protected	1
0xXY	$\overline{\texttt{0xXY}}$	Not specified in the documentation	?

Table 1: Flash memory protection status according to the values of RDP and nRDP

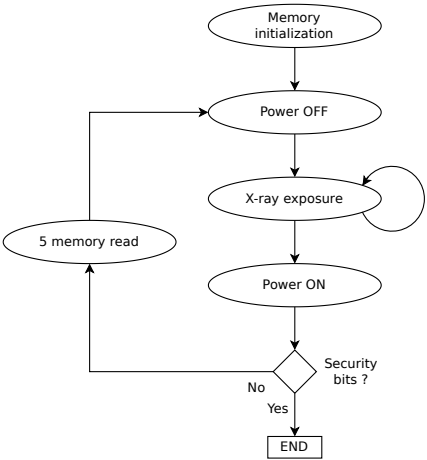
Experimental protocol

Source parameters

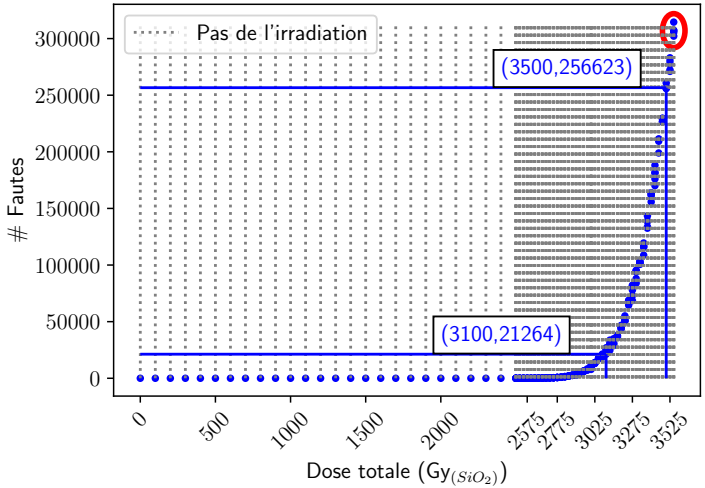
- 100 kV and 45 mA

⇒ Photons with 40 keV energy

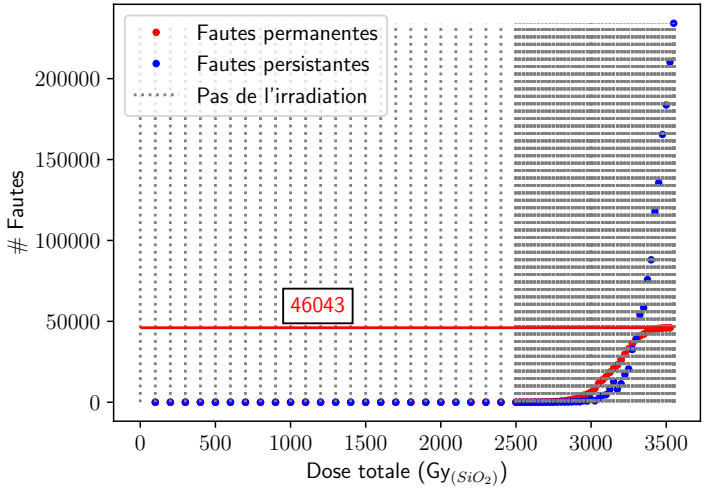
- Dose rate : 1 $Gy_{(SiO_2)}/s$



Results (Flash)

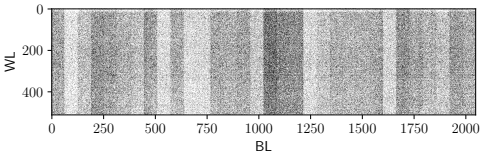


Results (Flash)



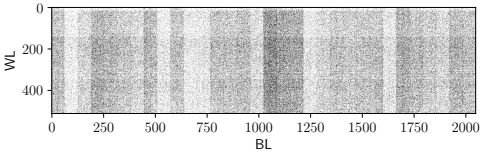
Recovery phenomenon

After irradiation



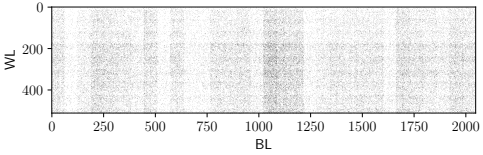
≤ 300 000 faults

After temporal recovery
7 days at room temperature



≤ 225 000 faults
(-25%)

After thermal recovery
2h at 150 °C



≤ 70 000 faults
(-69%)

Synthesis (IEEE PAINE 2023)

Two types of faults

- Floating gate discharge by photoemission
- MOS transistor threshold voltages drift by charge trapping
  - Temporal and thermal recoveries possible

Limits

- Injection of non-localized faults that cannot be exploited

Synthesis

Two types of faults

- Floating gate discharge by photoemission
- MOS transistor threshold voltages drift by charge trapping
  - Temporal and thermal recoveries possible

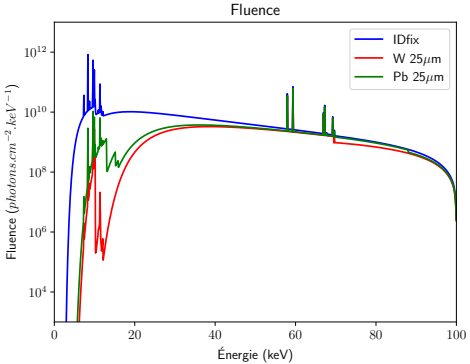
Limits

- Injection of non-localized faults that cannot be exploited

Focusing mask conception

- Two possible materials:
  - Tungsten (W)
  - Lead (Pb)

Shield simulation



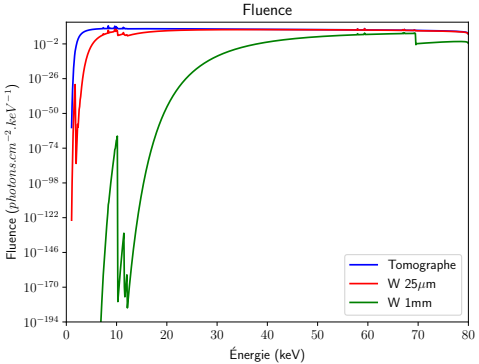
Shield	Overall attenuation
Tungsten (25 µm)	≲ 75%
Lead (25 µm)	≲ 70%

Conclusion

Tungsten mask is more efficient

Thanks to the team GPM2 of the SIMaP lab and to the CEA-Leti for the shields!

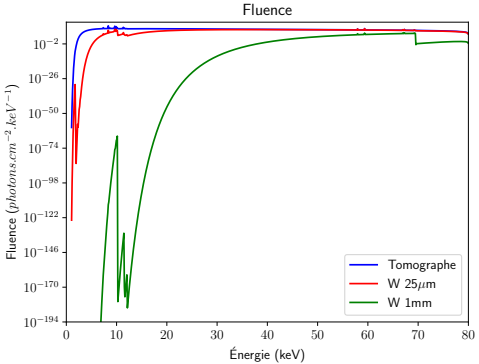
Shield simulation



Thickness	Overall attenuation
25 µm	≲ 80 %
1 mm	≲ 99.99 %

Thanks to the team GPM2 of the SIMaP lab and to the CEA-Leti for the shields!

Shield simulation



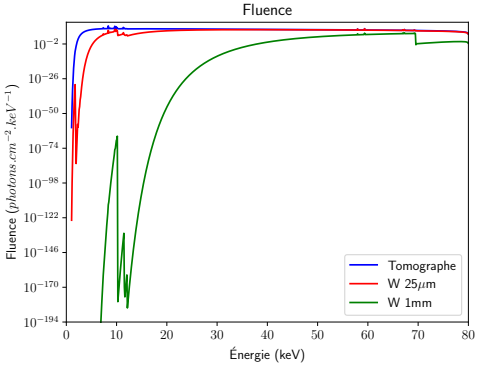
Thickness	Overall attenuation
25 µm	≃ 80 %
1 mm	≃ 99.99 %

Shield efficiency

- Experimental conditions not conducive to use the mask of 25 µm

Thanks to the team GPM2 of the SIMaP lab and to the CEA-Leti for the shields!

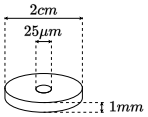
Shield simulation



Thickness	Overall attenuation
25 µm	≃ 80 %
1 mm	≃ 99.99 %

Shield efficiency

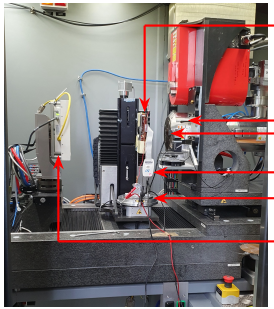
- Experimental conditions not conducive to use the mask of 25 µm

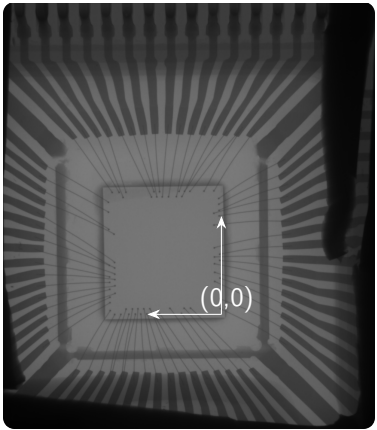


Shield dimensions

Thanks to the team GPM2 of the SIMaP lab and to the CEA-Leti for the shields!

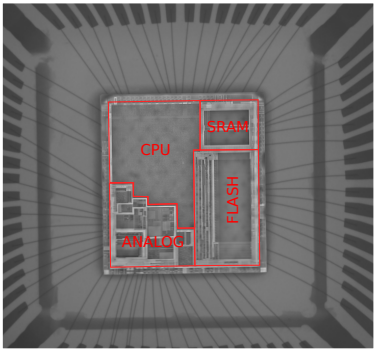
Tomograph description



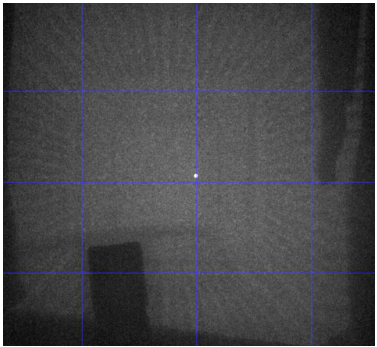


Thanks to the GPM2 team of the SIMaP lab !

Pictures obtained



Infrared image overlaid on the tomograph image

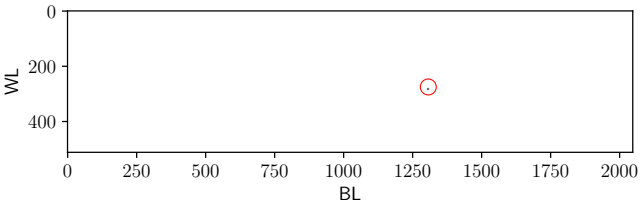


Shield with a thickness of 1 mm

Irradiation campaign : Powered device

3 differents positions

- Position ①: (0.6 mm,1.1 mm)
- Position ②: (0.6 mm,1.2 mm)
- Position ③: (0.7 mm,1.2 mm)



State of the Flash memory after 80 min exposure at the first position

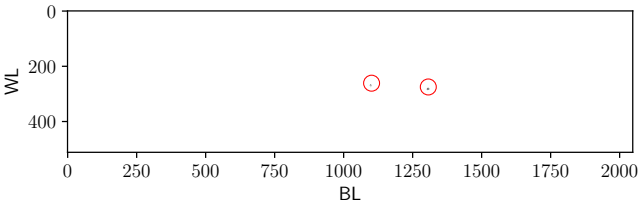
Results

Approx. 40 localized faults

Irradiation campaign : Powered device

3 differents positions

- Position ①: (0.6 mm,1.1 mm)
- Position ②: (0.6 mm,1.2 mm)
- Position ③: (0.7 mm,1.2 mm)



State of the Flash memory after 60 min exposure at the second position

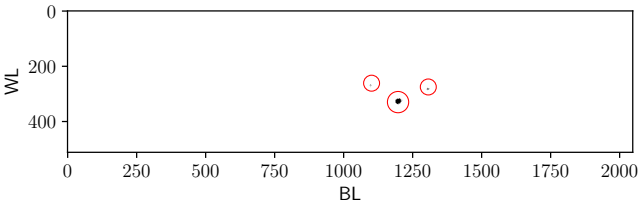
Results

Approx. 10 localized faults

Irradiation campaign : Powered device

3 differents positions

- Position ①: (0.6 mm,1.1 mm)
- Position ②: (0.6 mm,1.2 mm)
- Position ③: (0.7 mm,1.2 mm)



State of the Flash memory after 180 min exposure at the third position

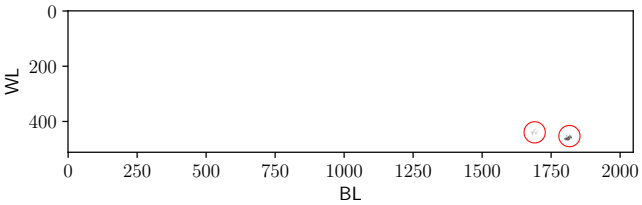
Results

Approx. 300 localized faults

Irradiation campaign : Unpowered device

3 different positions

- Position ①: (0.6 mm,1.2 mm) ⇒ 1 h of irradiation
- Position ②: (0.6 mm,1.1 mm) ⇒ 2 h15 min of irradiation

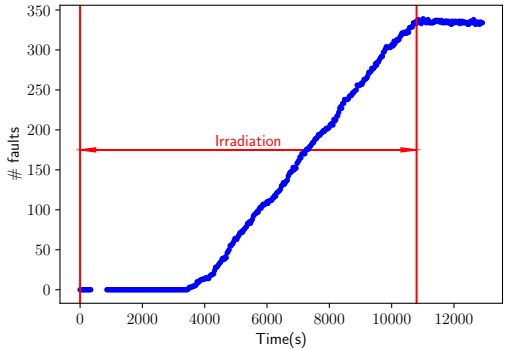


State of the Flash memory after X-ray exposure at the two positions

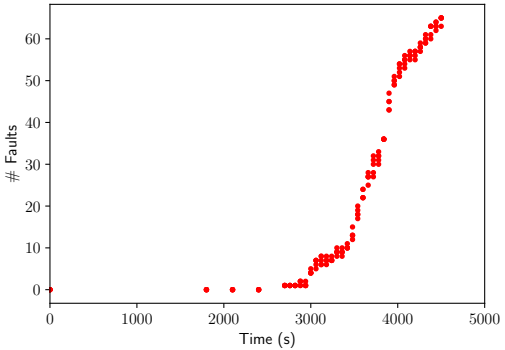
Results

- Approx. 70 faults localized at the position ①
- Approx. 300 faults localized at the position ②

Temporal evolution

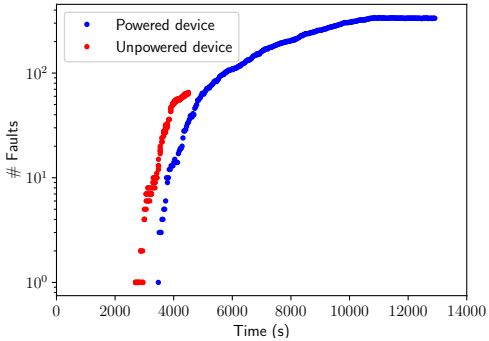


Powered device



Unpowered device

Temporal evolution



Powered and unpowered device in log scale.

Conclusion

Two kind of faults

- Floating gate discharge by photoemission
- MOS transistor threshold voltages drift by charge trapping
  - Temporal and thermal recoveries (more efficient)

Fault localization (under submission)

- Possibility of focusing the faults by usint a tungsten shield 1 mm

Limits

- Shield conception limits the localization
- Few possible security attack scenarios

Conclusion

Unpowered devices

- Real threat of these attacks
- Hardware sensors are not working
- No synchronisation required

Laser

- New fault model from physical to application level
- Validation of a PFA attack scenario

X-Rays

- Injection of permanent and non-permanent faults
- Possibility of focusing fault injection using a shield

Future works

Short term

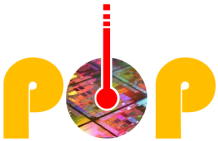
- Secure devices (including contermeasures)
  - Hardware sensors
  - EDAC codes
- Reconfigurable circuits (FPGA, SoC-FPGA)
  - Particularly in the context of PUFs and TRNGs (ring oscillators)

Long term

- Countermeasures design
  - Systematic read physical sensors at start-up (sensitive FGmos)
  - Analysis specific to attacks on non-powered circuits

Thank you for your attention !

This thesis is part of the *Power-Off laser attacks on security Primitives* (POP) project funded by the French Agence Nationale de la Recherche (ANR).



Advanced techniques for fault injection attacks on integrated circuits

Paul GRANDAMME

April 25th 2025