Hardware security of silicon chips: progress, pitfalls and challenges for physical attacks

Dr Sergei Skorobogatov

http://www.cl.cam.ac.uk/~sps32 email: sps32@cam.ac.uk



Talk Outline

- Introduction
- Why do we need hardware security?
- Evolution of the hardware security
- Attack techniques or what to worry about
- Challenges: from old days to modern chips
- Defence techniques
- Pitfalls: something can always go wrong
- Future: glorying or glooming
- Conclusions

Introduction

- Semiconductor chips are everywhere
 - electronic locks and keys, smartcards for banking and service applications, phone cards, crypto-processors
- Protection of systems and devices against physical attacks at a hardware level
 - tamper detection
 - environmental sensors
 - preventing unauthorised access (e.g. password protection)
 - security fuses for data and intellectual property (IP) protection
 - data encryption
- Hardware security implementation
 - at a PCB level
 - on a silicon die
- Problem: the security comes as an extra feature

Why do we need hardware security?

Theft of service

 attacks on service providers (satellite TV, electronic meters, access cards, software protection dongles)

Access to information

- information recovery and extraction
- gaining trade secrets (IP piracy)
- ID theft

Cloning and overbuilding

- copying for making profit without investment in development
- low-cost mass production by subcontractors

Denial of service

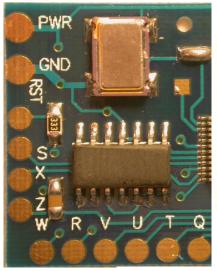
- dishonest competition
- electronic warfare

Who needs secure chips?

- car industry
 - anti-theft protection, spare parts identification
- accessory control
 - mobile phone batteries, printer toner cartridges, memory modules
- service and access control
 - RFID tags, access cards, payment tokens, software dongles
- home entertainment and consumer electronics
 - consumables, accessories, game consoles
- intellectual property protection
 - software copy protection
 - protection of algorithms
 - protection against cloning and reverse engineering

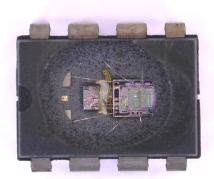
Hardware security evolution









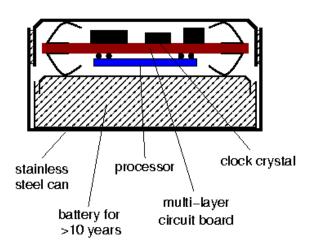


Hardware security evolution





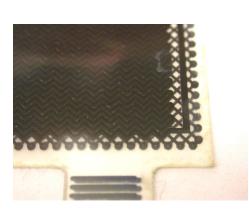




Hardware security evolution









Art of hardware security engineering

- What could be easier...
 - first understand the reason to attack your system
 - then find how your system is likely to be attacked, time and cost
 - after that develop adequate protection
 - finally perform security evaluation
 - ...and find your system has been hacked in a few months time
- Challenges in hardware security
 - choosing secure components
 - evolving attack technologies

Choosing secure components

- What has changed in the past?
 - too many devices on the market
 - vast majority of devices are claimed to be secure
 - security started to be used for marketing purposes
 - virtually impossible to test everything
- What are the problems?
 - certification does not provide guarantee against attacks
 - manufacturers do not carry any obligations or legal responsibility
 - no such thing as security benchmark
 - no ways of comparing devices from different manufacturers
 - no chip manufacturer will tell you the truth about security

Attack categories

Side-channel attacks

 techniques that allows the attacker to monitor the analog characteristics of supply and interface connections and any electromagnetic radiation

Software attacks

 use the normal communication interface and exploit security vulnerabilities found in the protocols, cryptographic algorithms, or their implementation

Fault generation

 use abnormal environmental conditions to generate malfunctions in the system that provide additional access

Microprobing

 can be used to access the chip surface directly, so we can observe, manipulate, and interfere with the device

Reverse engineering

 used to understand the inner structure of the device and learn or emulate its functionality; requires the use of the same technology available to semiconductor manufacturers and gives similar capabilities to the attacker

Attack methods

- Non-invasive attacks (low-cost)
 - observe or manipulate with the device without physical harm to it
 - require only moderately sophisticated equipment and knowledge to implement
- Invasive attacks (expensive)
 - almost unlimited capabilities to extract information from chips and understand their functionality
 - normally require expensive equipment, knowledgeable attackers and time
- Semi-invasive attacks (affordable)
 - semiconductor chip is depackaged but the internal structure of it remains intact
 - fill the gap between non-invasive and invasive types, being both inexpensive and easily repeatable

Non-invasive attacks

- Non-penetrative to the attacked device
 - normally do not leave tamper evidence of the attack
- Tools
 - digital multimeter
 - IC soldering/desoldering station
 - universal programmer and IC tester
 - oscilloscope, logic analyser, signal generator
 - programmable power supplies
 - PC with data acquisition board, FPGA board, prototyping boards
- Types of non-invasive attacks: passive and active
 - side-channel attacks: timing, power and emission analysis
 - data remanence
 - fault injection: glitching
 - brute forcing
- Comparing old days (late 90s) with today challenges

Non-invasive attacks: side-channel

- Timing attacks aimed at different computation time
 - incorrect password verification: termination on incorrect byte,
 different computation length for incorrect bytes
 - incorrect implementation of encryption algorithms: performance optimisation, cache memory usage, non-fixed time operations
- Today: timing attacks became harder to apply
 - common mistakes were fixed by manufacturers
 - internal clock sources and use of PLL made analysis difficult
 - countermeasures are in place: randomised clock, dummy cycles
 - careful selection of hardware eliminates many problems

Non-invasive attacks: side-channel

- Power analysis: measuring power consumption in time
 - very simple set of equipment a PC with an oscilloscope and a small resistor in power supply line; very effective against many cryptographic algorithms and password verification schemes
 - some knowledge in electrical engineering and digital signal processing is required
 - two basic methods: simple (SPA) and differential (DPA)
- Electro-magnetic analysis (EMA): measuring emission
 - similar to power analysis, but instead of resistor, a small magnetic coil is used allowing precise positioning over the chip
- Today: SPA/DPA and EMA became more challenging
 - higher operating frequency and noise: faster equipment is required
 - power supply is reduced from 5V to 1V: lower signal, more noise
 - 8-bit data vs 32-bit data: harder to distinguish single-bit change
 - more complex circuits: higher noise from other parts, hence, more signal averaging and digital signal processing are required
 - effective countermeasures for many cryptographic algorithms

Non-invasive attacks: data remanence

Data remanence in SRAM

- residual representation of data after erasure first discovered in magnetic media then appeared to be the case for other memories
- low temperature data remanence: cooling the device to −20°C increases the retention time from 1s to 100s, at −50°C to 1 hour
- dangerous to tamper resistant devices which store keys and secret data in a battery backed-up SRAM
- long period of time data storage causes the data to be "burned-in" and likely to appear after power up; dangerous to secure devices which store keys at the same memory location for years
- Today: data remanence in SRAM still exists
 - modern devices consume less power and have lower leakage
 - some countermeasures are in place to prevent burning-in
 - special memory chips with memory-clear input

Non-invasive attacks: data remanence

Data remanence in Flash and EEPROM

- levels of remanence threat: file system (undelete cmd), file backups (software features), smart memory (hardware buffers), memory cell
- floating-gate transistors store analog value charge of 10³–10⁵ e⁻
- widely used in microcontrollers and smartcards
- information can be recovered after memory bulk erase cycles, from PIC16F84A Flash memory even after 10 erase cycles
- Today: data remanence in Flash and EEPROM still exists
 - ineffective memory clean operations poses some threat
 - memory caching and buffering causes problems
 - power supply sensitivity in some chips
 - data recovery is more challenging due to higher density of cells
 - threat is ignored by many chip manufacturers

Non-invasive attacks: fault injection

- Glitch attacks
 - clock glitches
 - power supply glitches
 - corrupting data
- Security fuse verification in the Mask ROM bootloader of the Motorola MC68HC05B6 microcontroller
 - double frequency clock glitch causes incorrect instruction fetch
 - low-voltage power glitch results in corrupted EEPROM data read

```
#01h
           LDA
           AND
                      $0100
                                             ;the contents of the EEPROM byte is checked
           BEQ
                                             ;endless loop if bit 0 is zero
loop:
                      loop
           BRCLR
                      4, $0003, cont
                                             test mode of operation
           JMP
                      $0000
                                             ; direct jump to the preset address
cont:
```

Non-invasive attacks: fault injection

- Today: glitch attacks became harder to exploit
 - effective countermeasures are in place: clock and power supply monitors
 - internal clock sources, clock conditioning and PLL circuits
 - internal charge pumps and voltage regulators
 - asynchronous design
 - checksums (CRC, SHA-1)
 - encryption

Non-invasive attacks: brute forcing

Brute force attacks

- searching for keys and passwords, exploiting inefficient selection of keys and passwords
- recovering design from CPLDs, FPGAs and ASICs
- eavesdropping on communication to find hidden functions
- applying random signals and commands to find hidden functionality
- Today: brute force attacks became less feasible
 - longer keys make searching infeasible
 - moving from 8-bit base to 32-bit base means longer search
 - CPLDs, FPGAs and ASICs became too complex to analyse
 - too large search field for finding hidden functionality

Invasive attacks

Penetrative attacks

leave tamper evidence of the attack or even destroy the device

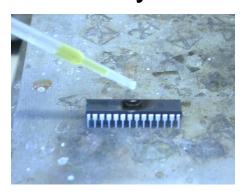
Tools

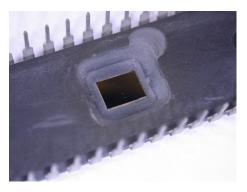
- IC soldering/desoldering station
- simple chemical lab
- high-resolution optical microscope
- wire bonding machine, laser cutting system, microprobing station
- oscilloscope, logic analyser, signal generator
- scanning electron microscope and focused ion beam workstation
- Types of invasive attacks: passive and active
 - decapsulation, optical imaging, reverse engineering
 - microprobing and internal fault injection
 - chip modification
- Comparing old days (late 90s) with today challenges

Invasive attacks: sample preparation

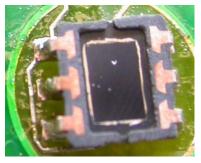
Decapsulation

- manual with fuming nitric acid (HNO₃) and acetone at 60°C
- automatic using mixture of HNO₃ and H₂SO₄
- full or partial
- from front side and from rear side
- Today: more challenging due to small and BGA packages

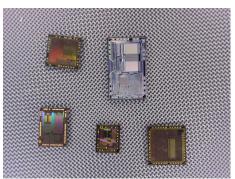


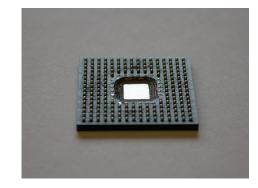






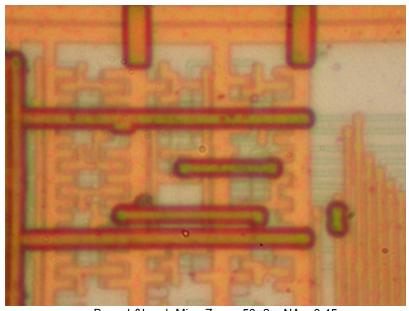




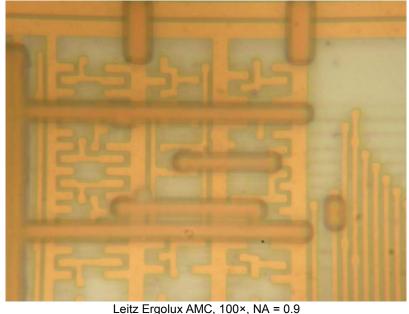


Invasive attacks: imaging

- Optical imaging
 - resolution is limited by optics and wavelength of a light: $R = 0.61 \lambda / NA = 0.61 \lambda / n \sin(\mu)$ best is $0.18\mu m$ technology
 - reduce wavelength of the light using UV sources
 - increasing the angular aperture, e.g. dry objectives have NA = 0.95
 - increase refraction index of the media using immersion oil (n = 1.5)
- Today: optical imaging is replaced by electron microscopy



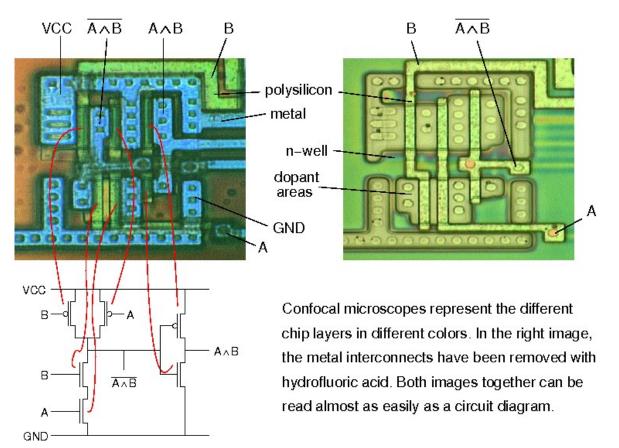
Bausch&Lomb MicroZoom, 50×2×, NA = 0.45



23

Invasive attacks: reverse engineering

- Reverse engineering understanding the structure of a semiconductor device and its functions
 - optical, using a confocal microscope (for > 0.5 µm chips)
 - deprocessing is necessary for chips with smaller technology



Invasive attacks: reverse engineering

Deprocessing

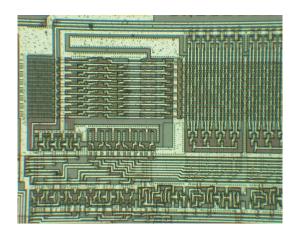
- removing passivation layer to expose the top metal layer for microprobing attacks
- decomposition of a chip for reverse engineering
- Mask ROM extraction

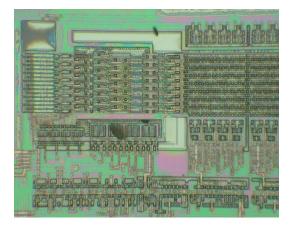
Methods

- wet chemical etching (KOH solutions, HCl, H₂O₂)
 - isotropic uniformity in all directions
 - uneven etching and undercuts metal wires lift off the surface
- plasma etching or dry etching (CF₄, C₂F₆, SF₆ or CCl₄ gases)
 - perpendicular to the surface
 - speed varies for different materials
- chemical-mechanical polishing (abrasives like Al₂0₃ or diamond)
 - good planarity and depth control, suitable for modern technologies
 - · difficult to maintain planarity of the surface, special tools are required

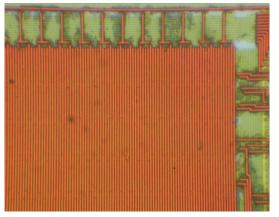
Invasive attacks: reverse engineering

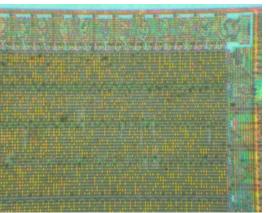
- Removing top metal layer using wet chemical etching
 - good uniformity over the surface, but works reliably only for chips fabricated with 0.8 μm or larger process (without polished layers)
- Today: plasma etching and chemical-mechanical polishing





Motorola MC68HC705C9A microcontroller 1.0 µm



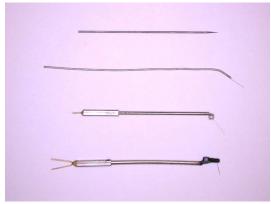


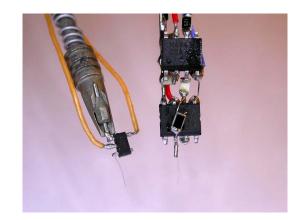
NEC μ PD78F9116 microcontroller 0.35 μ m

Invasive attacks: microprobing

- Microprobing with fine electrodes
 - eavesdropping on signals inside a chip
 - injection of test signals and observing the reaction
 - can be used for extraction of secret keys and memory contents
 - limited use for 0.35µm and smaller chips



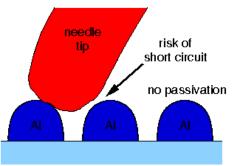


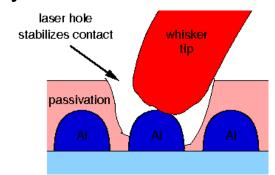


Invasive attacks: microprobing

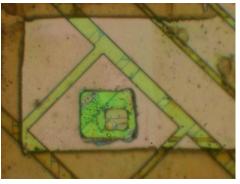
- Laser cutting systems
 - removing polymer layer from a chip surface
 - local removing of a passivation layer for microprobing attacks
 - cutting metal wires inside a chip
 - maximum can access the second metal layer

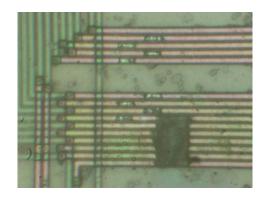






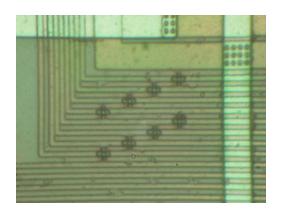
Picture courtesy of Dr Markus Kuhn

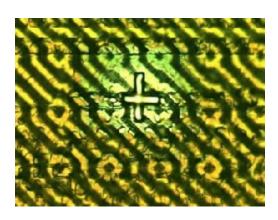




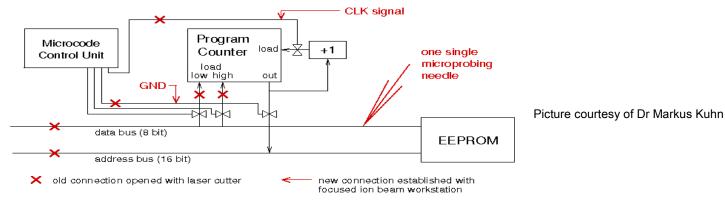
Invasive attacks: chip modification

- Today: Focused Ion Beam workstation
 - chip-level surgery with 10 nm precision
 - create probing points inside smartcard chips, read the memory
 - modern FIBs allow backside access, but require special chip preparation techniques to reduce the thickness of silicon





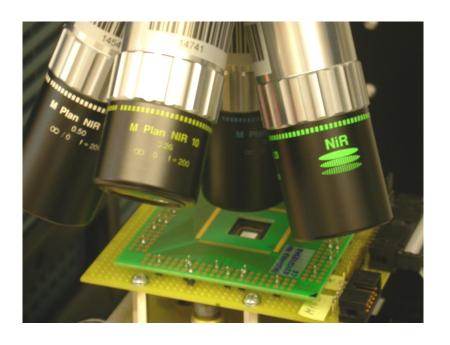
Picture: Oliver Kömmerling



Semi-invasive attacks

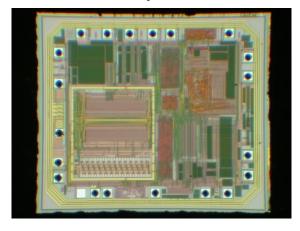
- Filling the gap between non-invasive and invasive attacks
 - less damaging to target device (decapsulation without penetration)
 - less expensive and easier to setup and repeat than invasive attacks
- Tools
 - IC soldering/desoldering station
 - simple chemical lab
 - high-resolution optical microscope
 - UV light sources, lasers
 - oscilloscope, logic analyser, signal generator
 - PC with data acquisition board, FPGA board, prototyping boards
 - special microscopes (laser scanning, infrared etc.)
- Types of semi-invasive attacks: passive and active
 - imaging: optical and laser techniques
 - fault injection: UV attack, photon injection, local heating
 - side-channel attacks: optical emission analysis, induced leakage
- Comparing old days (late 90s) with today challenges

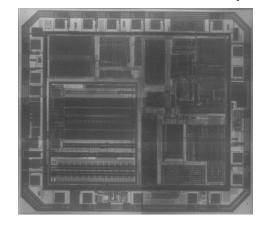
- Backside infrared imaging
 - microscopes with IR optics give better quality of image
 - IR-enhanced CCD cameras or special cameras must be used
 - resolution is limited to ~0.6µm by the wavelength of used light
 - view is not obstructed by multiple metal layers



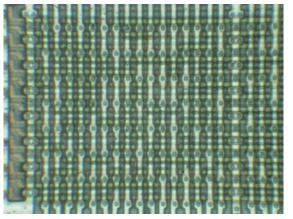


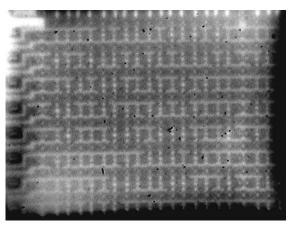
- Backside infrared imaging
 - Mask ROM extraction without chemical etching
- Today: the main option for 0.35µm and smaller chips
 - multiple metal wires do not block the optical path





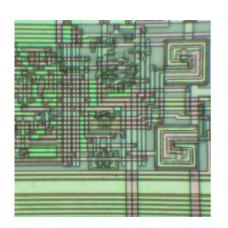
Texas Instruments MSP430F112 microcontroller $0.35 \ \mu m$

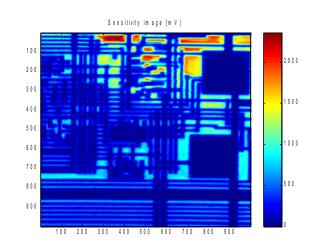


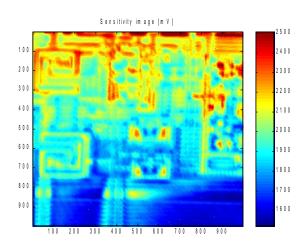


Motorola MC68HC705P6A microcontroller 1.2 µm

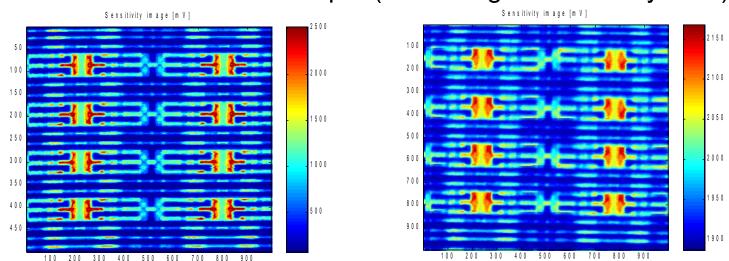
- Advanced imaging techniques active photon probing
 - Optical Beam Induced Current (OBIC)
 - photons with energy exceeding semiconductor band gap ionize IC's regions, which results in a photocurrent flow producing the image
 - used for localisation of active areas
 - also works from the rear side of a chip (using infrared lasers)







- Advanced imaging techniques active photon probing
 - light-induced current variation
 - alternative to light-induced voltage alteration (LIVA) technique
 - photon-induced photocurrent is dependable on the state of a transistor
 - reading logic state of CMOS transistors inside a powered-up chip
 - works from the rear side of a chip (using infrared lasers)
- Today: backside approach for 0.35µm and smaller chips
 - multiple metal wires do not block the optical path
 - resolution is limited to ~0.6µm (still enough for memory cells)

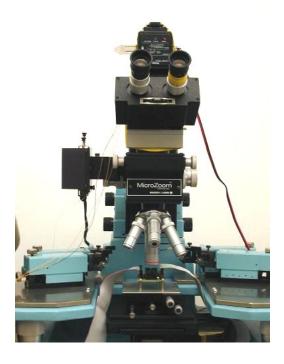


Semi-invasive attacks: fault injection

Optical fault injection attacks

- optical fault injection was observed in my experiments with microprobing attacks in early 2001, introduced as a new method in 2002
- lead to new powerful attack techniques and forced chip manufacturers to rethink their design and bring better protection
- original setup involved optical microscope with a photoflash and Microchip PIC16F84 microcontroller programmed to monitor its SRAM

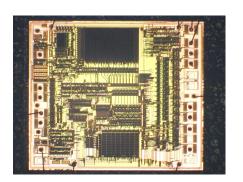


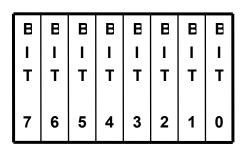


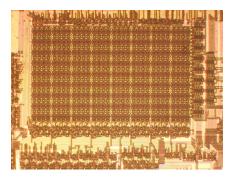
Semi-invasive attacks: fault injection

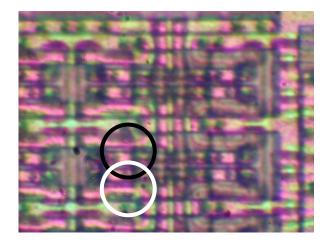
- Optical fault injection attacks
 - the chip was decapsulated and placed under a microscope
 - light from the photoflash was shaped with aluminium foil aperture
 - physical location of each memory address by modifying memory contents
 - the setup was later improved with various lasers and a better microscope
- Today: backside approach for 0.35µm and smaller chips
 - successfully tested on chips down to 130nm







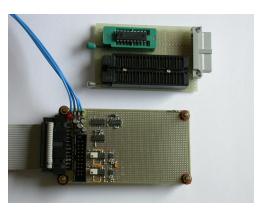


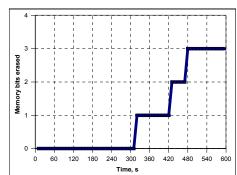


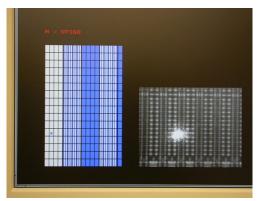
Semi-invasive attacks: fault injection

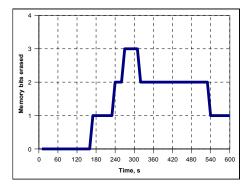
- Localised heating using cw lasers
 - test board with PIC16F628 and PC software for analysis
 - permanent change of a single memory cell on a 0.9µm chip
- Today: influence is limited for modern chips (<0.5µm)
 - adjacent cells are affected as well



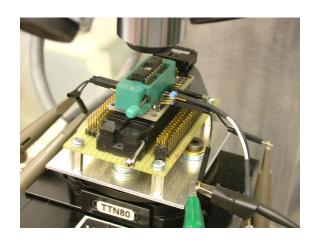


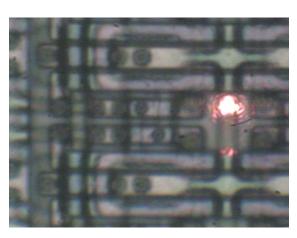


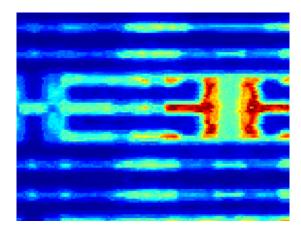




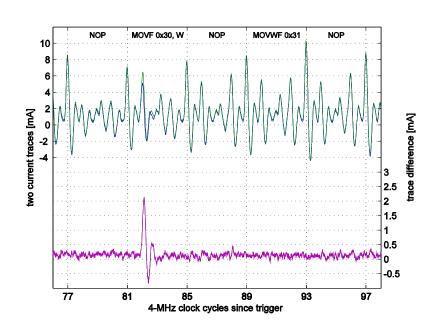
- Optically enhanced position-locked power analysis
 - Microchip PIC16F84 microcontroller with test program at 4 MHz
 - classic power analysis setup (10 Ω resistor in GND, digital storage oscilloscope) plus laser microscope scanning setup
 - test pattern
 - run the code inside the microcontroller and store the power trace
 - point the laser at a particular transistor and store the power trace
 - compare two traces

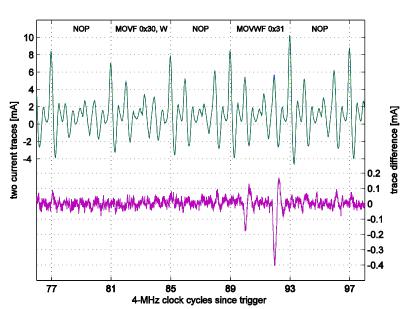






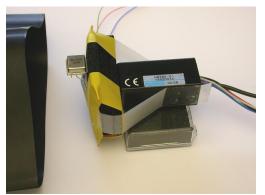
- Optically enhanced position-locked power analysis
 - results for memory read operations: non-destructive analysis of active memory locations ('0' and '1')
 - results for memory write operations: non-destructive analysis of active memory locations ('0→0', '0→1', '1→0' and '1→1')
- Today: backside approach for 0.35µm and smaller chips
 - single-cell access is limited to 0.5µm laser spot

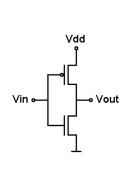


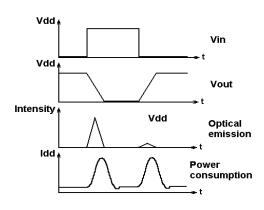


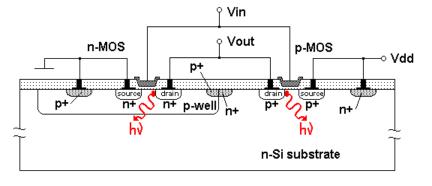
- Optical emission analysis
 - transistors emit photons when they switch
 - 10⁻² to 10⁻⁴ photons per switch with peak in NIR region (900–1200 nm)
 - optical emission can be detected with photomultipliers and CCD cameras
 - comes from area close to the drain and primarily from the NMOS transistor





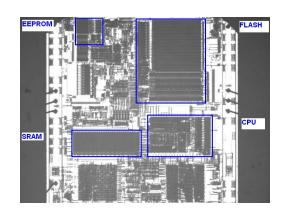


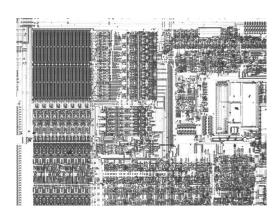




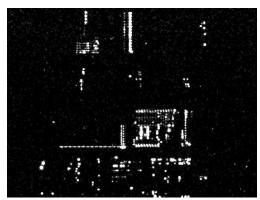
- Optical emission analysis
 - Microchip PIC16F628 microcontroller with test code at 20 Mhz;
 PMT vs SPA and CCD camera images in just 10 minutes
- Today: backside approach for 0.35µm and smaller chips
 - successfully tested on chips down to 130nm (higher Vcc and >1 hour)

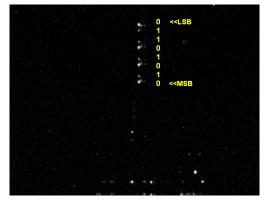












Semi-invasive attacks comparison

INVASIVE	SEMI-INVASIVE
Microprobing	Laser scanning
	Optical probing and emission analysis
Chip modification (laser cutter or FIB)	Fault injection
Reverse engineering	Special microscopy
Rear-side approach with a FIB	Infrared techniques

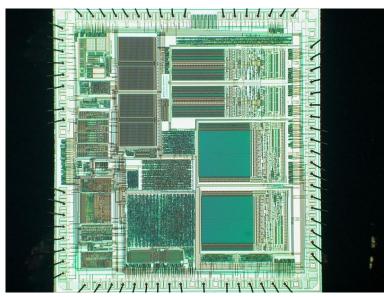
NON-INVASIVE	SEMI-INVASIVE
Power and clock glitching	Fault injection
Power analysis	Special microscopy
	Optical probing and emission analysis

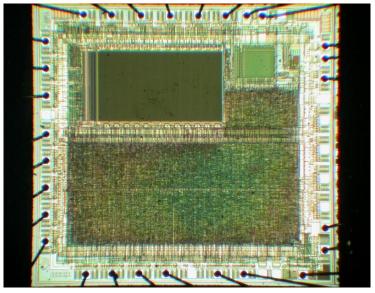
- Some semi-invasive attacks still effective on 130nm chips
- Recent publications showed that they still represent security threat to modern chips

Defence technologies: tamper protection

Old devices

- security fuse is placed separately from the memory array (easy to locate and defeat)
- security fuse is embedded into the program memory (hard to locate and defeat), similar approach is used in many smartcards in the form of password protection and encryption keys
- moving away from building blocks which are easily identifiable and have easily traceable data paths



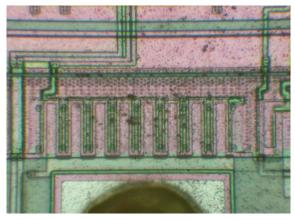


Motorola MC68HC908AZ60A microcontroller

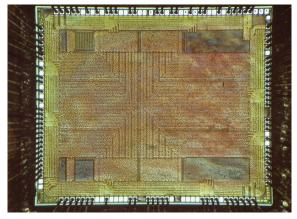
Scenix SX28 microcontroller

Defence technologies: tamper protection

- Help came from chip fabrication technology
 - planarisation as a part of modern chip fabrication process (0.5 µm or smaller feature size)
 - glue logic design makes reverse engineering much harder
 - multiple metal layers block any direct access
 - small size of transistors makes attacks less feasible
 - chips operate at higher frequency and consume less power
 - smaller and BGA packages scare off many attackers







0.9µm microcontroller

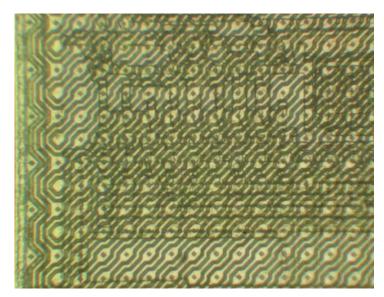
0.5µm microcontroller

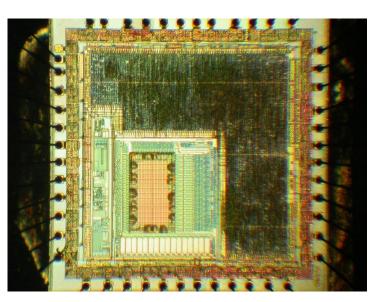
0.13µm FPGA

Defence technologies: tamper protection

Additional protections

- top metal layers with sensors
- voltage, frequency and temperature sensors
- memory access protection, crypto-coprocessors
- internal clocks, power supply pumps
- asynchronous logic design, symmetric design, dual-rail logic
- ASICs, secure FPGAs and custom-designed ICs
- software countermeasures





STMicroelectronics ST16 smartcard

Fujitsu secure microcontroller

Defence technologies: what goes wrong?

- Security advertising without proof
 - no means of comparing security, lack of independent analysis
 - no guarantee and no responsibility from chip manufacturers
 - wide use of magic words: protection, encryption, authentication, unique, highly secure, strong defence, cannot be, unbreakable, impossible, uncompromising, buried under x metal layers
- Constant economics pressure on cost reduction
 - less investment, hence, cheaper solutions and outsourcing
 - security via obscurity approach
- Quicker turnaround
 - less testing, hence, more bugs
- What about back-doors?
 - access to the on-chip data for factory testing purposes
 - how reliably was this feature disabled?
 - how difficult is to attack the access port?

Defence technologies: how it fails

- Microchip PIC microcontroller: security fuse bug
 - security fuse can be reset without erasing the code/data memory
 - solution: fixed in newer devices
- Hitachi smartcard: information leakage on a products CD
 - full datasheet on a smartcard was placed by mistake on the CD
- Actel secure FPGA: programming software bug
 - devices were always programmed with a 00..00 passkey
 - · solution: software update
- Xilinx secure CPLD: programming software bug
 - security fuse incorrectly programmed resulting in no protection
 - · solution: software update
- Dallas SHA-1 secure memory: factory initialisation bug
 - some security features were not activated resulting in no protection
 - · solution: recall of the batch
- Other examples
 - insiders, datasheets of similar products, development tools, patents
 - solution: test real devices and control the output

Defence technologies: why goes wrong?

- Ignorance of mistakes by chip manufacturers
- Unconditional trust from customers
- Reluctance to collaborate with people from academia
- Security perception and awareness levels
 - Level 1: attack is announced
 - lesson: nothing is absolutely secure
 - reaction: ignorance and disbelieve
 - Level 2: attack is confirmed and proved

 - lesson: something to worry about reaction: show no interest and develop some quick fix
 - Level 3: attack method is known (how to attack)
 - lesson: cost and time can be estimated
 - reaction: attempt to prevent disclosure and apply some measures
 - Level 4: technique for developing the method is known (know why)

 - lesson: security can be improved reaction: attempt to prevent disclosure and rethink security
 - Level 5: process of finding the technique is known
 - lesson: security can be redesigned and core of the problem fixed
 reaction: attempt to prevent disclosure and rethink strategy

Future work

- Improvements to semi-invasive attacks
 - some 180nm and 130nm chips tested
 - preparation for testing 90nm chips is under way
 - 65nm chips are in plans
- New challenges
 - is everything solved in side-channel attacks area?
 - what if a new attack can improve the existing methods?
 - normally you expect 10 times improvement every 3–5 years
 - by 10 times: this can be a publication
 - by 100 times: this can be a good publication
 - by 1000 times: this can be an outstanding publication
 - by 1000000 times: maybe better not to publish
 - What a million times improvement would mean for a real device?
 - 1 day for an attack which normally takes 2000 years to succeed
 - 1 second for an attack which normally takes 10 days to succeed
- More publications to come in 2010 and 2011

Conclusions

- There is no such a thing as absolute protection
 - given enough time and resources any protection can be broken
- Technical progress helps a lot, but has certain limits
 - do not overestimate capabilities of the silicon circuits
 - do not underestimate capabilities of the attackers
- Defence should be adequate to anticipated attacks
 - security hardware engineers must be familiar with attack technologies to develop adequate protection
 - choosing the correct protection saves money in development and manufacturing
- Attack technologies are constantly improving, so should the defence technologies
- Many vulnerabilities were found in various secure chips and more are to be found, that poses more challenges to hardware security engineers

References

- Slides
 - http://www.cl.cam.ac.uk/~sps32/lorentz_2010.pdf
- Literature:
 - http://www.cl.cam.ac.uk/techreports/UCAM-CL-TR-630.pdf
 - http://www.cl.cam.ac.uk/~sps32/#Publications