Optical Fault Induction Attacks

Sergei Skorobogatov, Ross Anderson
University of Cambridge, Computer Laboratory
(sps32,rja14)@cl.cam.ac.uk

Abstract

We describe a new class of attacks on secure microcontrollers and smartcards. Illumination of a target transistor causes it to conduct, thereby inducing a transient fault. Such attacks are practical; they do not even require expensive laser equipment. We have carried them out using a flashgun bought second-hand from a camera store for $30. As an illustration of the power of this attack, we developed techniques to set or reset any individual bit of SRAM in a microcontroller. Unless suitable countermeasures are taken, optical probing may also be used to induce errors in cryptographic computations or protocols, and to disrupt the processor's control flow. It thus provides a powerful extension of existing glitching and fault analysis techniques. This vulnerability may pose a big problem for the industry, similar to those resulting from probing attacks in the mid-1990s and power analysis attacks in the late 1990s.

We have therefore developed a technology to block these attacks. We use self-timed dual-rail circuit design techniques whereby a logical 1 or 0 is not encoded by a high or low voltage on a single line, but by (HL) or (LH) on a pair of lines. The combination (HH) signals an alarm, which will typically reset the processor. Circuits can be designed so that single-transistor failures do not lead to security failure. This technology may also make power analysis attacks very much harder too.

1. Introduction

Secure microcontrollers and smartcards are designed to protect both the confidentiality and the integrity of sensitive information. It is not sufficient to prevent an attacker from finding out the value of a stored cryptographic key; she must also be unable to set part of the key to a known value, or to induce errors in the computation that enable sensitive information to be deduced. These errors may be data errors, such as an incorrect digital signature that leaks the value of the signing key [1], or errors in the code, such as a missed conditional jump that reduces the number of rounds in a block cipher [2]. Until now, the most widely known technique for inducing such errors was glitching – the introduction of voltage transients into the power of clock line of the target chip. Many chips are now designed to resist glitch attacks.

A review of the tamper-resistance of smartcard and secure microcontroller chips may be found in [3]. Attacks tend to be either invasive, using chip testing equipment such as probing stations and focused ion beam workstations to extract data from the chip directly, or else non-invasive processes involving the exploitation of unintentional electromagnetic emissions, protocol design flaws, and other vulnerabilities that manifest themselves externally. Either type of attack may be passive or active. The standard passive invasive attack involves using microprobes to monitor a smartcard's bus while a program is executing; in an active attack, signals may be also injected, the classic example being the use of a grounded microprobe needle on the clock line to the instruction latch to disable jump instructions. A passive non-invasive attack is analyzing the electromagnetic field in the neighborhood of the device under test [4], while glitching is the classic example of an active attack.

Until now, invasive attacks involved a relatively high capital investment for lab equipment plus a moderate investment of effort for each individual chip attacked. Non-invasive attacks such as power analysis require only a moderate capital investment, plus a moderate investment of effort in designing an attack on a particular type of device; thereafter the cost per device attacked is low. Non-invasive attacks are thus particularly attractive where they exist.

Unfortunately for the attacker, many chipmakers have now implemented defenses against the most obvious non-invasive attacks. These defenses include randomized clocking to make power analysis harder, and circuits that react to glitches by resetting the processor. Meanwhile invasive attacks are becoming constantly more demanding and expensive, as feature sizes shrink and device complexity increases, We therefore set out to find new, more powerful, ways of attacking chips.

We describe out new class of attacks as 'semi-invasive'. By this, we mean that, like invasive attacks, they require depackaging the chip to get access to the chip surface. But the passivation layer of the chip remains intact - semi-invasive methods do not require making an electrical contact to the metal surface or doing mechanical damage to the silicon.

Semi-invasive attacks are not entirely new. The electromagnetic analysis of [4] is best performed on a naked chip, and the old EPROM-hacking trick of exposing the write protect bit of a microcontroller to UV light usually entails depackaging it. Semi-invasive attacks could in theory be performed using such tools as UV light, X-rays, lasers, electromagnetic fields and local heating. They could be used individually or in conjunction with each other. However, this field has hardly been explored.

We will now show that extremely powerful attacks can be carried out quickly using very cheap and simple equipment.

2. Background

Once the semiconductor transistor had been invented, it was found to be more sensitive to ionizing radiation – whether caused by nuclear explosions, radioactive isotopes, X-rays or cosmic rays - than the thermionic valves used previously. In the middle sixties, during experiments with pulsed lasers, it was found that coherent light causes some similar phenomena. Lasers started to be used to simulate the effects of ionizing radiation on semiconductors [5].

Since then the technology has been improved dramatically. Expensive inert-gas-based lasers and solid-state lasers have been replaced with low-cost semiconductor lasers. As a result, the technology has moved from the laboratory all the way down to consumer electronics.

Laser radiation can ionize an IC's semiconductor regions if its photon energy exceeds the semiconductor band gap. The laser radiation with 1.06μm wavelength and 1.17eV photon energy used in [6] has a penetration depth of about 700μm and provides good spatial ionization uniformity for silicon devices. However its focusing is restricted by dispersion to several micrometers, and this is not precise enough for modern semiconductor devices. However, when moving from infrared to visible light, photon absorption dramatically increases [7], and it has become possible to use red and green lasers as the transistors in modern chips became thinner. Smaller devices also mean that less energy is required to achieve the same level of ionization.

In the case of CMOS devices, there is a danger of latching up the circuit, causing an open circuit that results in permanent damage. So the use of radiation with CMOS structures must be done with appropriate precautions.

Although there are many publications about using pulsed lasers to simulate ionizing radiation, we could find no published information about using them to control or change the behavior of integrated circuits. So we decided to apply an intense light source to a semiconductor chip, and particularly to CMOS logic, to see whether it would be possible to change the state of a memory cell and how easy, or difficult, it might be.

Our first experiments targeted SRAM. The structure of a standard six-transistor SRAM cell is shown on Figure 1 [8]. Two pairs of p- and n-channel transistors create a flip-flop, while two other n-channel transistors are used to read its state and write new values into it. The layout of the cell is shown on Figure 2 [9]. The transistors M_1 and M_2 create the CMOS inverter; together with the other similar pair, they create the flip-flop which is controlled by the transistors M_3 and M_6.

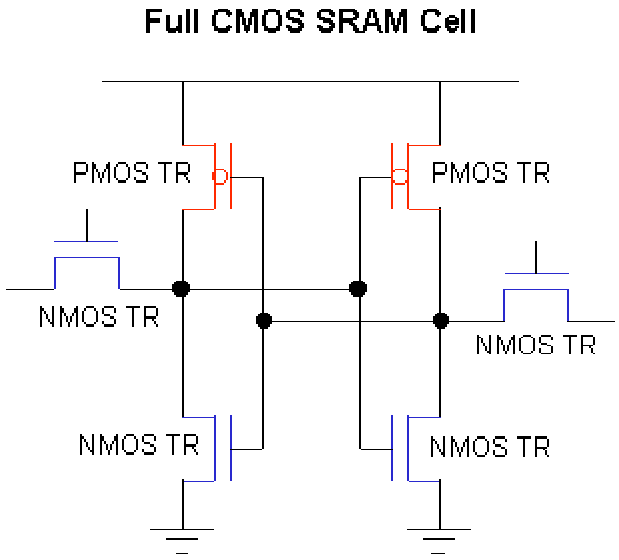

Figure 1. Structure of six-transistor SRAM cell

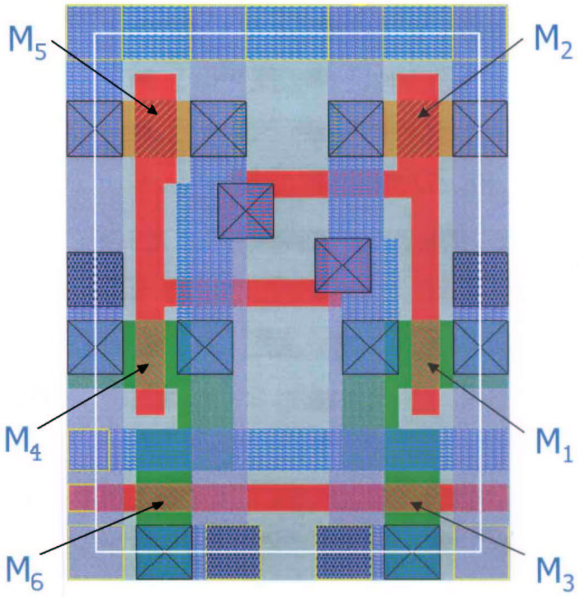

Figure 2. Layout of six-transistor SRAM cell

If the transistor M_1 could be opened for a very short time by an external stimulus, then it could cause the flip-flop to change state. By exposing the transistor M_4, the state of the cell would be changed to the opposite. The main difficulties we might anticipate are focusing the ionizing radiation down to several μm^2 and choosing the proper intensity.

3. Experimental Method

For our experiments we chose a common microcontroller, the PIC16F84, which has 68 bytes of SRAM memory on chip (Figure 3). A standard depackaging procedure was applied to the chip and the result of this operation is represented on Figure 4. The SRAM memory array is located in the middle of the bottom of the chip die. This area is shown with 80x magnification on Figure 5.

Figure 3. Microcontroller PIC16F84

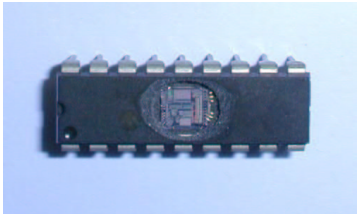

Figure 4. Depackaged PIC16F84 chip

Because we had a very limited equipment budget, and the laser we had appeared unsuitable, we decided to use a cheap photoflash lamp (a Vivitar 550FD, bought secondhand from a camera shop for 20 pounds). Although the luminosity of a flashlamp is much less than that of a pulsed laser, with appropriate magnification the necessary level of ionization might be achieved. We used duct tape to fix the photoflash lamp on the video port of a Wentworth Labs MP-901 manual probing station (Figure 6). Magnification was set to the maximum– 1500x.

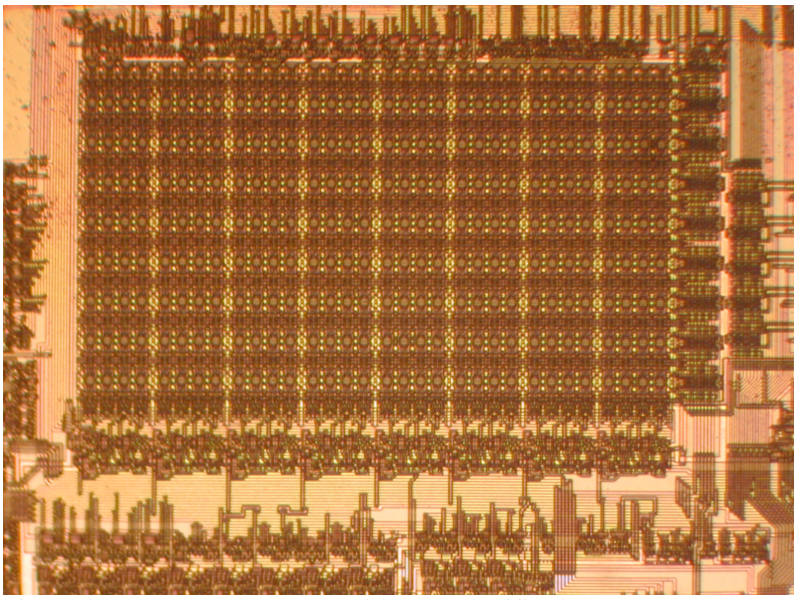

Figure 5. SRAM memory array with 80x magnification

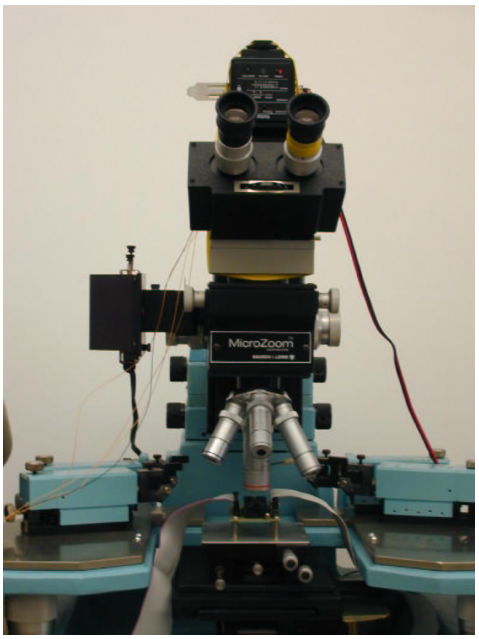

Figure 6. Wentworth Labs MP-901 manual prober with Vivitar 550FD photoflash lamp mounted on top

The microcontroller was programmed to upload and download its memory. By filling the whole memory with constant values, exposing it to the flash light, and downloading the result, we could observe which cells changed their state.

By shielding the light from the lamp with an aperture made from aluminum foil, we succeeded in changing the state of only one cell. The output power of the lamp was set to the maximum. The final state of the cell depended on the area exposed to the flash. This confirmed our intuition that it would be possible to change the contents of SRAM using a low cost semi-invasive attack.

4. Results

We found we could change any individual bit of an SRAM array. The array, under maximum magnification, is shown in Figure 7. Focusing the light spot from the lamp on the area shown by the red circle caused the cell to change its state from 1 to 0, with no change if the state was already 0. By focusing the spot on the area shown by blue circle, the cell changed its state from 0 to 1 or remained in state 1.

It can be seen from Figure 5 that the SRAM array is divided into eight equal blocks. By exposing cells in different blocks, we found that each block corresponds to one bit plane of information. The result of this operation is shown in Figure 8.

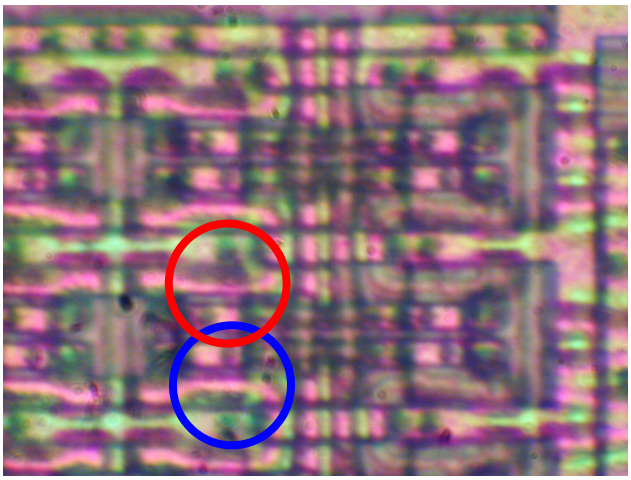

Figure 7. SRAM memory array with 1500x magnification

BIT 7	BIT 6	BIT 5	BIT 4	BIT 3	BIT 2	BIT 1	BIT 0

Figure 8. Allocation of data bits in SRAM memory array

We built a map of the addresses corresponding to the physical location of each memory cell by exposing each cell in turn to the photoflash light. The result is presented in Figure 9, with the left edge corresponding to the bottom side of the block. It can be seen that the addresses are not sequential, but divided into three groups.

0x30	0x34	0x38	0x3C	0x40	0x44	0x48	0x4C	0x10	0x14	0x18	0x1C	0x20	0x24	0x28	0x2C	0x0C
0x31	0x35	0x39	0x3D	0x41	0x45	0x49	0x4D	0x11	0x15	0x19	0x1D	0x21	0x25	0x29	0x2D	0x0D
0x32	0x36	0x3A	0x3E	0x42	0x46	0x4A	0x4E	0x12	0x16	0x1A	0x1E	0x22	0x26	0x2A	0x2E	0x0E
0x33	0x37	0x3B	0x3F	0x43	0x47	0x4B	0x4F	0x13	0x17	0x1B	0x1F	0x23	0x27	0x2B	0x2F	0x0F

Figure 9. Allocation of addresses in each bit block of SRAM memory array

This shows how simple semi-invasive attack methods can be used for reverse engineering a memory address map. The only limitation is that the photoflash lamp does not produce even and monochromatic light, so it is very difficult to control the area where the spot of the light will be applied. This problem can be solved by replacing the lamp with a suitable laser.

5. Further Work

The work reported above shows that optical probing attacks are possible using low-cost equipment. Using similar equipment, we plan to implement many of the fault induction attacks proposed in papers such as [1],[2]. In particular, we understand that our technique is effective at implementing the attack of Boneh et al on RSA signatures against at least one smartcard currently on the market. (The Digital Millennium Copyright Act might make the reporting of any further details imprudent, especially by a Russian author at a conference on US soil.)

Further scientific work in our plan includes a fuller investigation of the potential for attacks by an opponent with a moderately resourced laboratory, by which we mean a modern probing station with both a multiple wavelength laser and a motorized stage under program control. We hope to have such apparatus operational by the time of the conference, and intend to use it for testing a number of new attack ideas on different types of chip.

6. Countermeasures

The optical probing attack described above is a new and devastating technique for attacking smartcards and other security processors. We anticipate that, like the power analysis attacks reported by Kocher in [10], it could have a significant commercial effect on the industry, in that it will force a thorough reappraisal of security claims and the introduction of new defensive technology.

Following Kocher, we decided to delay the announcement of our attack until proper defenses were available. Existing high-end chip-defense techniques, such as top-layer metal shielding and bus encryption, may make an attack using these techniques more complicated, but are not enough. A sophisticated attacker can defeat metal shielding by going through the rear of a chip using an infrared laser, while bus encryption can be defeated by attacking registers directly.

The defensive technology that we have developed uses self-timed dual-rail logic. Conventional digital logic uses a clock to synchronize activities; but the cost of clocking rises as devices become more complex, and this has led to a surge of interest in design techniques for self-timed, or asynchronous, circuits – circuits that do not use clocks. Such circuits need some mechanism whereby functional components in a circuit can signal that they are ready to receive data, or are done. One way of doing this is to introduce redundancy into the data path.

In dual-rail logic, a 0 or 1 is signaled not by a low or high voltage on a single wire, but by a combination of signals on a pair of wires. For example, 0 may be `LH' and 1 may be `HL'. When used in self-timed circuits, `LL' signals quiescence. The principal drawback of this simple arrangement is fragility: bugs tend to cause the emergence of the unwanted `HH' state, which usually propagates rapidly throughout the circuit and locks it.

Our innovation was to turn this fragility to advantage, by making `HH' into an error signal. This signal can be raised deliberately by tamper sensors, causing the device to lock [11]. Of more interest here is the fact that matters can be so arranged that single device failures cause are unlikely to cause the output of sensitive information [12]. We believe that such robustness will be a requirement for many high-security devices in future.

The engineering details are non-trivial. For example, an obvious concern is that almost any undetected malfunction could be exploited by the attack of Boneh et al on RSA signatures. Colleagues have therefore developed a modular multiplication unit using our technology. Similarly, although bus encryption can remove the need to protect on-chip memory arrays, there remains the risk of attacks on the program counter and other registers. Other colleagues have therefore developed registers, and a memory management unit, that use our technology. These circuits have also been designed to make the power consumption independent of the input data. They will be described in separate papers that will be submitted to appropriate hardware design conferences.

7. Conclusion

Standard CMOS circuitry is extremely vulnerable to attack using optical probing. By exposing a transistor to a laser beam, or even the focused light from a flashlamp, it can be made to conduct. This gives rise to many effects that can be used by an attacker. We have described here in detail how the illumination of a certain area of an SRAM cell can be used to set it to either 0 or 1. This may be used, for example, to load a short program that outputs sensitive data. However, this is only the beginning.

It appears that, given only moderately expensive equipment, an attacker may be able to induce a fault in a CMOS integrated circuit, in any targeted transistor, and at precisely the clock cycle of her choice. This

is quite devastating. Hardware countermeasures will be necessary.

Acknowledgement

The first author thanks the European Union for financial support.

References

[1] D. Boneh, R. A. DeMillo, R. J. Lipton, ‘‘On the Importance of Checking Cryptographic Protocols for Faults, Advances in Cryptology – Eurocrypt 97, Springer LNCS vol 1233 pp 37-51

[2] R. J. Anderson, Markus G. Kuhn, “Low Cost Attacks on Tamper Resistant Devices”, in M.Lomas et al. (ed.), Security Protocols, 5th International Workshop, Paris, France, April 7-9, 1997

[3] R. J. Anderson, Security Engineering – A Guide to Building Dependable Distributed Systems’, Wiley 2001

[4] J. J. Quisquater, D. Samyde, “ElectroMagnetic Analysis (EMA): Measures and Countermeasures for Smart Cards”, International Conference on Research in Smart Cards, E-smart 2001, Cannes, France, pp.200-210, Sept. 2001

[5] D.H. Habing, “Use of Laser to Simulate Radiation-induced Transients In Semiconductors and Circuits”, IEEE Trans. Nuc. Sci., Vol. NS-12, No.6, pp.91-100, Dec. 1965

[6] A.H. Johnston, “Charge Generation and Collection in p-n Junctions Excited with Pulsed Infrared Lasers”, IEEE Trans. Nuc. Sci., Vol. NS-40, No.6, pp.1694-1702, 1993

[7] “Handbook of Optical Constants of Solids", edited by Edward D. Palik, Orlando: Academic Press, 1985, pp.547-569

[8] J. M. Rabaey, “Digital Integrated Circuits: A Design Perspective”, Prentice-Hall, 1995

[9] K. Yun, “Memory”, UC San Diego, Adapted from EE271 notes, Stanford University

[10] P. Kocher, “Differential Power Analysis”, Advances in Cryptology – Crypto 99, Springer LNCS v 1666 pp 388-397

[11] S. W Moore, R. J. Anderson, M. G. Kuhn, “Improving Smartcard Security using Self-Timed Circuit Technology”, Fourth AciD-WG Workshop, Grenoble, ISBN 2-913329-44-6, 2000

[12] S. W Moore, R. J. Anderson, P. Cunningham, R. Mullins, G. Taylor, “Improving Smartcard Security using Self-Timed Circuits”, Asynch 2002, proceedings published by IEEE Computer Society Press