Behind the horror are important and real hardware security questions.

Olivier Thomas - Founder and CEO @ Texplained

25th, Feb. 2016

As CEO of Texplained (a french company specialized in hardware reverse engineering and hardware attacks) but mainly as a hardware reverse-engineer, I wanted to comment on some of the recent articles I read about extracting the secrets from the iphone belonging to the murderer. As it turned out, writing a small comment is not an easy task, therefore here is some of my thoughts.

Apple's position:

Apple declared in its 16th, feb. 2016 letter that data confidentiality must be ensured and I can't agree more. But on the other side and because of the barbary that occurred, it also makes sense to gather any piece of information that can help in increasing people physical security. As a french citizen and after terrorist attacks striking France several times last year, I do understand the need for extracting secure devices. The question that remains is how to perform that critical operation without creating some backdoored devices that by nature would not guarantee a decent data confidentiality.

Apple refuses to create a dedicated iOS release with lower security and once again, I have to admit this is making sense. Apple is not generating revenue on its customer data which legitimates not to backdoor its devices.

Extracting complex chips as the Apple A6 processor is an exciting challenge for whoever works in the hardware security field.

That's where I read those articles proposing some methods to extract the apple chip. I have to admit that this casting light on hardware attacks is something I like a lot as Texplained is constantly advising its customers to pay attention to the hardware attacks threats (from piracy, to backdoor, IP violation…). Therefore, the following will be quite technical.

Time and effort are the two main questions that need clarification :

The proposed solutions are pretty wide, from power analysis to laser glitching and even micro-probing. This looks like a hardware security benchmark. When it comes to silicon security assessment, three types of attacks have to be evaluated : non-invasive, semi-invasive and invasive attacks.

Power analysis:

Power analysis belongs to the non-invasive type of attacks because the chip package will not be opened. It has the advantage that required equipments are affordable.

The all theory behind power analysis is that when bits are flipping inside a processor, power is consumed. By looking at traces of the consumed power, information about what the chip is doing can be recovered. Knowing what algorithm is being used can lead to the extraction of the cryptographic keys by using some DPA (Differential Power Analysis) techniques.

The apple chip is a very complex piece of silicon called a SOC (System On Chip). Therefore, the A6 processor is not a single processor but a combination of 2 ARM cores, 3 graphic cores plus peripherals. Its internal frequency is pretty high (GHz range) which means that transistors are switching around a billion time per second. This frequency is generated internally by PLLs (structures that create the clock signal which is used to synchronized the internal processes). The power consumption of the chip is related to the entire chip. Getting traces that are relevant is far from being trivial in this situation and can take a lot of time and effort. There is no guarantee of success there but as the technique is the cheapest of the proposed ones, it might be tried with a significant automation.

Micro-probing:

Another proposed solution is based on micro-probing the chip to extract some data. This process is known as an invasive attack and is a very efficient way to extract data. On the other side, it requires strong expertise and heavy equipments to be performed.

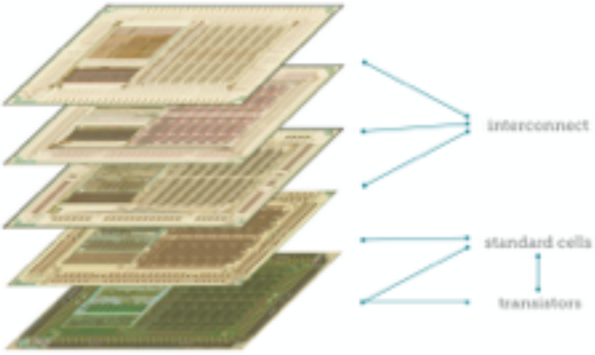
Invasive attacks on a SOC… it got me thinking. Let me share with you why by analyzing the different steps.

To start an invasive attack, the first thing to do is to get samples of the target chip because some of the needed steps involve the destruction of part of the samples. I expect that the number of sample for that study will be pretty high (several tens).

The second step is to remove the chip from its package by using some chemicals. So far so good.

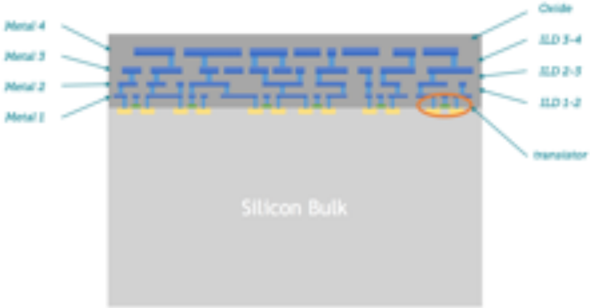
Then, reverse-engineering the chip is needed to understand what is going on and to be able to identify the few signals you want to read and / or manipulate to extract the data.

To be able to reverse-engineer a chip, one has to take pictures of every of its layers. An Integrated Circuit (IC) is a combination of transistors (several hundreds of million in our case) that are connected together to create bigger and bigger functions (until you get a processor). Those transistors are build on a silicon substrate and their interconnection is made by stacking layers of metal tracks (at least 10 layers in our case). Those metal layers are linked together by vias which are small plugs that creates junctions between tracks of two adjacent metal layers.



Optical pictures of several layers of an IC

To reach every of the layer, laboratory work has to be performed. The techniques are well known to the silicon designer and manufacturer as they have to investigate every bug before releasing a stable version of their chips. This is called Failure Analysis. But there is a major difference between deprocessing a chip for finding a defect and reverse engineering it. When a defect is localized, the deprocessing focus on reaching that particular point of interest. The area of interest for Reverse-Engineering is the entire chip. Therefore, the deprocessing must be perfectly planar. Here is the first challenge. Apple A6 is a 32 nm chip which means that the smallest transistor gate in the design is 32 nm wide. This also means that the height between two interconnect layers is pretty small. In addition, the chip surface is huge (almost 100 square mm). Being able to remove material evenly to expose only one layer is a task that only trained skilled professional can perform with some limitations (...).



Chip Cross-Section

After exposing one layer, pictures need to be taken. I love to see optical pictures in articles as they are colorful and nice looking but with this kind of technology, optical microscopes are banned and the laboratory conducting the test will use a Scanning Electron Microscope. This machine allows for higher magnification which is mandatory to be able to study the small geometrical structures embedded inside the chip. Using a high magnification results in a small field of view meaning that several thousands (if not tens of thousands) of pictures will be taken for each layer. That makes at least 10 scans for imaging the interconnect layers. But there is another trick here. Every layer is connected to the adjacent layers with some little plugs that also needs to be imaged. This makes at least 10 scans of several thousands pictures for the interconnect layers and at least 10 extra scans in between those layers.

SEM picture example

At Texplained, we perform Reverse-Engineering of chips using this process and we validated our technology on 40 nm node. We are currently looking for partners to investigate the feasibility of reverse-engineering the smallest geometries. There can be a problem when imaging the latest nodes and we honestly do not know if that problem would need to be solved for the A6. Chip manufacturers are using some materials that makes, as a side effect, the imagery a difficult operation. SEM pictures are usually distorted. When stitching pictures together, this can become an issue as one need to make sure that the different structures are linked properly. On small nodes (22nm for example), the picture distortion is much greater than on older chips and might be an issue.

We do think at Texplained that even if that problem is significant, it can be solved. So we will assume that everything went well so far : perfect deprocessing and high quality imaging.

Of course, stitching the huge collection of pictures together must be performed so as to be able to trace signals, understand part of the circuitry and strategize an attack. This is of course not trivial and using standard software to deal with the images is not an option. You probably do not have enough memory on your computer to hold so much pictures. Texplained expertise relies on the use of a custom, home-made software "ARES" (Automated Reverse Engineering Software) which has precisely be designed to display and process huge number of pictures. The software then converts the pictures into comprehensible files which represents the netlist of the circuit (list of transistors including their connection).

ARES

AUTOMATED REVERSE ENGINEERING SOFTWARE

I strongly encourage you to get more info on that on our website, there is a white paper and also the video of our last presentation at Black Hat - Las Vegas 2015 which show the all process on a secure smart-card.

Without a dedicated software, the study that can be performed will be limited to finding obvious structures such as memories, PLLs, and the position of the cores. Manual stitching and signal tracing is an option I really do not recommend unless you have unlimited time and ressources. This could be not enough to strategize an attack.. but we will assume that someone finds where to micro-probe to get keys and useful data.

Micro-probing an apple A6 processor is a big task. As suggested in the different articles, a Focus Ion Beam (FIB) will be used to add test pads to the chip. Those are the places where the tiny sharp needles will be placed to extract / modify the data. It may seem easy but FIB editing an Integrated Circuit is a dangerous task as the chip can die in the process with an Electro Static Discharge or by simply removing material at the wrong place or creating electric shorts.. As we are speaking here about a single chip to extract, the risk is huge.

The chip itself is inside a BGA (Ball Grid Array) package with more than a thousand pins. Inside the package, the chip is probably face down (flip chip). This means that the silicon bulk is easily accessible when the surface is not. It is possible to perform backside FIB editing but this process is far more complicated than frontside FIB editing.

One can argue that it is still possible to remove the chip from its board, make a hole to reach its surface with some acid and then perform a frontside attack. But this is an option that will increase the risk of killing the chip even more. Removing the chip is an easy operation. Making a hole to reach the correct spot is more difficult. The package has to be drilled carefully. There is a chance that it contains a small board (which will have to be studied) to route the signals from the chip to the pins. This PCB can be cut partially under the chip as this area is generally filled only by a ground plane. The connection between the chip and this PCB are probably made of copper and do not like nitric acid so much as they can get dissolved. This operation becomes tricky.

But when it comes to micro-probing, one can not solder the chip back as he will loose access to its probe points. Making another hole in the phone board to reach the re-soldered chip does not seem to be an option as there are parts on both side of the board. One will have to create kind of an adapter to be able to connect the chip back while keeping frontside access. This is done pretty often but we are talking here about a high end chip that operates at high frequencies. The adapter must then be carefully designed so as the overall phone signal integrity remains correct.

Last but not least, micro-probing a chip that runs in the Giga Hertz range is also not trivial as the tungsten probes (needles) can disturb the internal chip signal leading to malfunctions.

Invasive Attacks on A6 seems great and I am not saying here that this is not an option. I am only saying that it requires a very strong expertise:

- Deprocessing of sub 40 nm chip of 100 square mm.
- Deprocess more than 20 layers with perfect planarity
- SEM image all layers
- Reverse-Engineering a dual ARM core
- Strategizing an attack
- Backside FIB editting
- Or frontside FIB editting
  - Drilling inside the package
  - Accessing the chip surface with chemical without damaging bounding wires
  - Designing of an adapter to use the chip
- Micro-probing fast chip
- Exploiting the data.

As a hardware security evaluator, I am used to describe a risk vs cost situation. The risk here is to destroy the chip that needs to be extracted. This is a pretty high risk. The cost on the other side is also pretty high. We are speaking about several weeks of laboratory work with a daily rate of easily 1K dollars. But the reverse-engineering process could take even longer and finding people up to this task can be challenging.

Laser fault injection:

Looking at a flip chip, the other proposed method seems easier to perform : laser fault injection. This is the good occasion to speak a bit about semi-invasive attacks. Semi-invasive because only the package will be opened. The chip itself will remain untouched. The sample preparation is therefore lighter than for an invasive attack. The chip will not be removed from its board and the access to its backside will be easily made.

Laser fault injection is a technique made for injecting wrong values inside a device in order to create a situation that advantages the hacker.

When an infrared laser is shot at a transistor, electrons will be created and can make the transistor switch, thus modifying the regular values. The macro effect can be among others instruction skip or register modification. It is possible to perform such attacks from the frontside or the backside of the chip. If you do it from the frontside, the metal tracks will act as a shield and diffract the light beam making the attack difficult to setup. But we are probably dealing here with a flip chip which means direct access to the backside. Transistors are far from the bottom of the silicon bulk but silicon is transparent to infrared.

The difficulty of laser fault injection is to find the precise location and timing for the attack. Performing a reverse-engineering such as the one described for the invasive strategy can help finding the sweet spot and represents a systematic approach that increase the success rate. But that's not the only method. Someone can use several phones to test the attack blindly. Trial and error, pass-fail tests can be conducted until the correct location and timing are eventually found. There is still a risk of killing the device while performing that search but the final attack can be quite reliable. The risk of killing the only interesting phone is far lower in these conditions. This technique is not properly reducing the time of the study. In both cases (Reverse-engineering or blind testing), weeks of investigation have to be planned. The cost will still be high.

Other methods:

The proposed methods are time consuming. Part of this is due to the complexity and the size of the chip that needs to be extracted. The high incentive for reading that chip has to justify the time and necessary ressources for the extraction. If that time is too long, this will teach a lesson to the bad guys around the world. They will make sure to have the latest product. Apple is releasing new devices every year which means that already several chips of that complexity would potentially need to be analyzed. And that is a key point here. For someone who had studied every Ax processor, a new model will be analyzed faster because of the gathered knowledge. This is a strategy we use at Texplained to study complex chips. Of course, I just said that this one shot project is a time consuming and ressource intensive project and I am now saying that it should be extended to previous and latest generation chips. But this is the cost for understanding this kind of system and being able to bypass its protection.

I would finally add that when I first heard of that all story and I was thinking of how to look at the A6, I had several different strategies in mind (non invasive, semi-invasive, invasive) but the truth is that it is hard to elaborate a correct strategy without looking at the chip itself for some time.

We do consider at Texplained that with sufficient time and effort, Reverse-Engineering could give sufficient information to choose the most efficient and less risky attack vector.

If reading out an Apple Ax processor seems reasonable, it does not solve the security policy that we might want to put in place for guaranteeing data confidentiality and protection while using all necessary means for investigating criminal network to increase people physical security. Apple could for example as the chip designer skip the all Reverse-Engineering part and focus on an attack strategy. When one has been designed properly, a service laboratory could be rented to perform the actual attack. This would limit the time necessary to recover the data while avoiding the creation of any backdoored devices. You could argue here that if Apple can find a "weakness", hackers could find it too and you would be partially right. But if Apple does not look for this, hackers may find it anyway. This leads us to the "how to secure a chip?" question... But this is another topic..