The State-of-the-Art in IC Reverse Engineering

Randy Torrance and Dick James

Chipworks Inc.
3685 Richmond Road, Ottawa, Ontario, Canada K2H 5B7
`rtorrance@chipworks.com, djames@chipworks.com`

Abstract. This paper gives an overview of the place of reverse engineering (RE) in the semiconductor industry, and the techniques used to obtain information from semiconductor products.

The continuous drive of Moores law to increase the integration level of silicon chips has presented major challenges to the reverse engineer, obsolescing simple teardowns, and demanding the adoption of new and more sophisticated technology to analyse chips. Hardware encryption embedded in chips adds a whole other level of difficulty to IC analysis.

This paper covers product teardowns, and discusses the techniques used for system-level analysis, both hardware and software; circuit extraction, taking the chip down to the transistor level, and working back up through the interconnects to create schematics; and process analysis, looking at how a chip is made, and what it is made of. Examples are also given of each type of RE. The paper concludes with a case study of the analysis of an IC with embedded encryption hardware.

1 Introduction

One of the most basic business requirements is the need to know what the competition is doing. If a company wants to get into a new area of business, the simplest thing to do is buy an existing product and take it apart to see what is inside it. Having done that, we know the parts list involved, and the technological challenges to be faced in manufacturing the new version.

Reverse engineering (RE) can cover objects from as large as aircraft down to the smallest microchip, and the motivations have varied from the paranoia of the Cold War, through commercial piracy, to competitive intelligence, product verification, and courts of patent law. If we look back over the last few decades, reverse engineers around the world have had a significant influence on the dissemination of technology in the electronics industry.

RE is now a recognised part of the competitive intelligence field, and is commonly used to benchmark products and support patent licensing activities. A side area is the need to RE archaic parts that have gone out of service, and need replacing in long-lived equipment such as military systems, nuclear reactors, airliners, and ships.

A fact of life these days is that simple teardowns of products are just not good enough any more. Advances in semiconductor technology, namely the massive

integration of billions of individual devices and masses of functions into single components, have forced RE to evolve into a specialised niche of the engineering profession.

2   RE in the Semiconductor Industry

The question most often asked about reverse engineering is "is it legal?" The short answer is – yes! In the case of semiconductors, RE is protected in the US by the Semiconductor Chip Protection Act, which allows it "for the purpose of teaching, analyzing, or evaluating the concepts or techniques embodied in the mask work or circuitry..." There is similar legislation in Japan, the European Union, and other jurisdictions.

In the semiconductor business, RE customers fall into two groups: those who are interested in technical information, and those that are interested in patent-related information. The technical information customers are usually within manufacturing companies, performing product development, or strategic marketing or benchmarking studies. The patent clients are usually patent lawyers or intellectual property (IP) groups within companies. There are also companies that are purely licensing companies, and deal only in IP.

Types of RE

Reverse engineering of semiconductor-based products can broadly take several forms:

- Product teardowns – identify the product, package, internal boards, and components
- System level analysis – analyse operations, functions, timing, signal paths, and interconnections
- Process analysis – examine the structure and materials to see how it is manufactured, and what it is made of
- Circuit extraction – delayer to transistor level, then extract interconnections and components to create schematics and netlists.

3   Product Teardowns

Product teardowns are the simplest type of RE in the electronics arena; the unit is simply disassembled, the boards and sub-assemblies are photographed, and the components are inventoried. Reverse engineers are usually only interested in what components are in the device at this level, but there are also companies that use the data to provide a bill of materials and tentative costing for the manufacture.

Figure 1 shows an Apple 8 GB iPod nano personal media player, partly torn down to expose the internal board and the ICs used [1]. Optical and x-ray analyses (Fig. 2) showed that the 64 Gb flash memories were actually 2 x 32 Gb stacked packages, each containing four 8 Gb dice (total 64 Gb). In this case, we continued with detailed process analyses of the 8 Gb flash chips, since they were leading edge devices from Samsung and Toshiba.
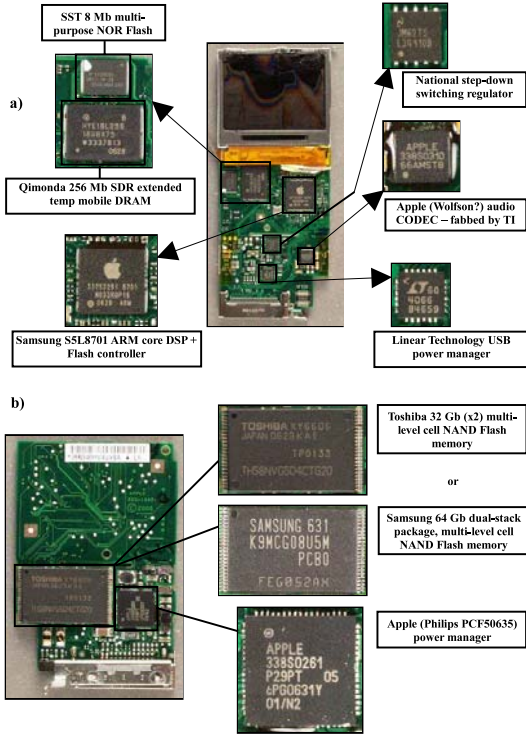
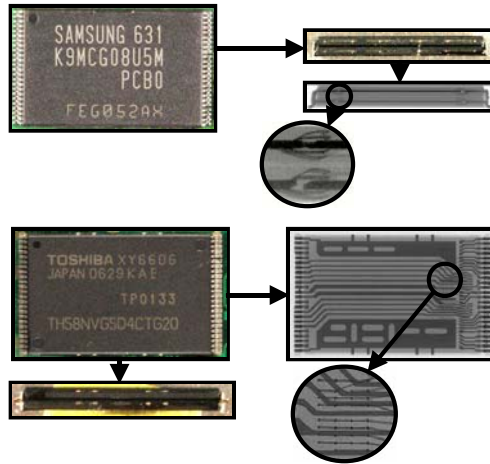**Fig. 1.** Partial Teardown of Apple 8 GB iPod Nano: a) Top b) Bottom

Fig. 2. Optical and X-Ray Images of 64 Gb Flash Devices

4   System Level Analysis

Just as there is a huge variation in electronic systems, there is also a variety of methods for system analysis. Electronic systems can consist of hardware, software, firmware, communications, transducers, etc. System analysis is useful for all of these.

4.1   Hardware

Hardware analysis takes one of two forms: reverse engineering or functional analysis.

Reverse engineering is a hierarchical analysis method. Take the example of a cell phone. The first phase of reverse engineering is to tear down the phone,
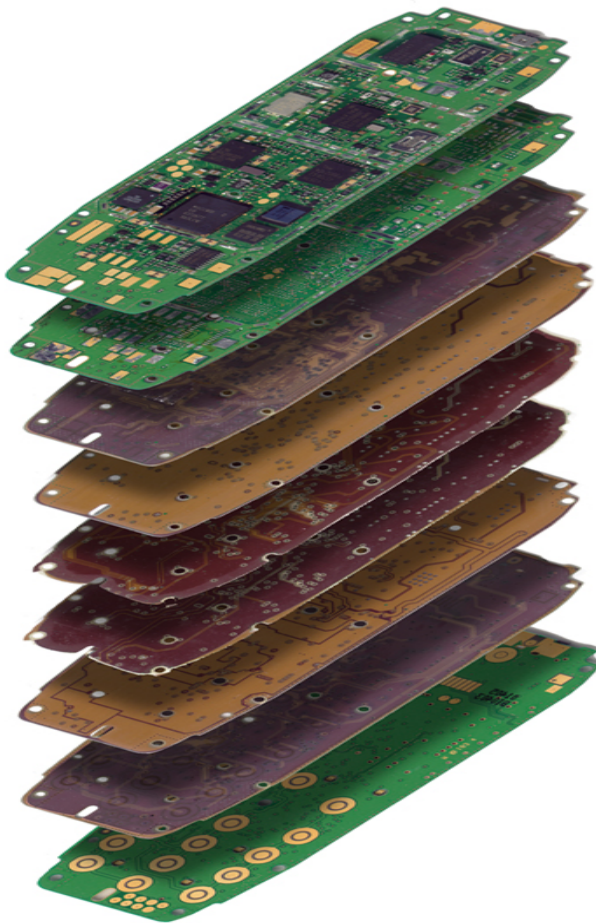

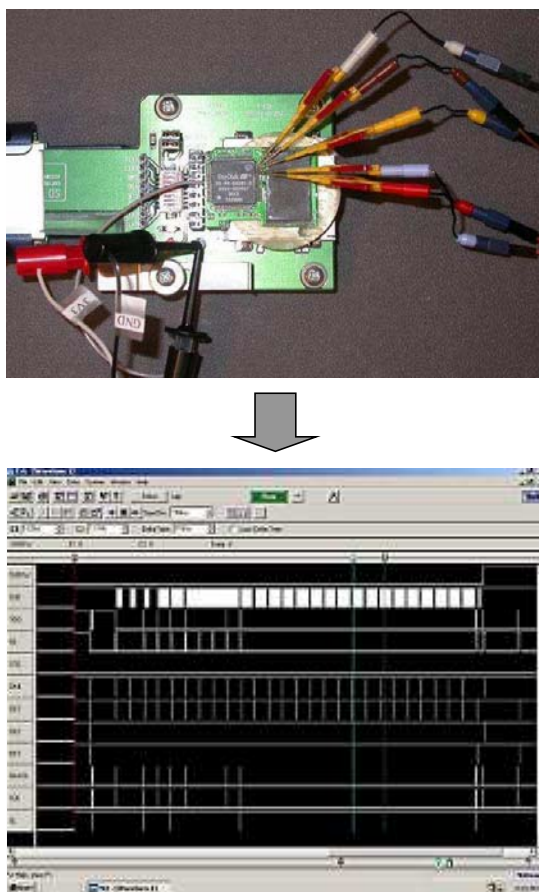**Fig. 3.** Delayered Nine Layer PCB from Cell Phone

Fig. 4. Probing SanDisk SD Memory Card for Functional Analysis

making notes of all connections between subsystems. Next, the main board is reverse engineered. Photos are taken of the board for future work. All components on the board are catalogued and then selectively removed. If the board is multi-layered, it can be delayered and imaged (Figure 3). The connections between all components are then identified and entered into the board schematic. Alternatively, electrical probing can sometimes be used to find the connections. Either way, a complete schematic of the board can be re-created.

Functional analysis entails system monitoring during functional operation. A system can be instrumented with probes wherever needed (sometimes with great difficulty, but it can usually be done, as shown in Figure 4). Microprobing is used to monitor on-chip signals. Test cases are developed, and stimulus created for operating the system in its functional modes. Signal generators, logic analyzers, and oscilloscopes are used to drive the system and collect the results. The signals and full system are then analyzed. Using the cell phone example once again, the phone can be partially disassembled, but still electrically connected to allow for

operation. Probes can be used to monitor key buses, pins of chips, and connectors. The phone can then be operated, and the signals analyzed, to understand the operation.

4.2 Software

As with hardware, software can be analyzed using the same two techniques; reverse engineering and functional analysis.

Software reverse engineering is the process of taking machine code and converting it back into human-readable form. The first task is often extraction of embedded code from an on-chip memory. Many techniques are available, such as EEPROM programmers, bus monitoring during code upload, and schematic extraction. Sometimes the code is protected with software or hardware locks. These can often be disabled via a collection of techniques. A chip's test port can be a good method of accessing its contents. IC microsurgery can be used to modify or bypass hardware locks. Usually these techniques require circuit analysis first, in order to identify the locks and find modifications that will disable them.

Encrypted code requires encryption analysis, followed by decryption. This requires both the keys and an understanding of the encryption algorithm. The keys can often be read from the embedded memory, along with the code, using the techniques described above. The encryption algorithm can sometimes be discovered via documentation or functional analysis. If these methods fail, then circuit extraction can often be used to reverse engineer the algorithm.

Once the code is extracted, disassemblers can be used as long as the processor and instruction set are known. Tools are then available to help take assembly code and structure it into a more C-like format. This structured code can then be analyzed by software experts. Code can be analyzed in either "static" ("dead") mode or "dynamic" ("live") mode. Live analysis is undertaken when it is possible to obtain the full control of the processor: starting and stopping code, inspecting registers, memory, tracing code execution. Live analysis is always preferable to dead code analysis which consists of analyzing just the instructions without the ability to inspect the code while running. Using software simulators enables another mode of software RE which is in between these two.

Software functional analysis is similar to hardware functional analysis. Test cases are designed, stimulus is created, the code can be instrumented, and the software executed. The outputs of this software can take many forms, from creating charts or driving a GUI, to controlling a robot or playing a song. These outputs can be analyzed to better understand the software or system.

5   Process Analysis

Process analysis of chips is straightforward in theory, since micro-analytical tools have been around for some time. Every wafer fab has a range of equipment for process control and failure analysis, and Chipworks uses the lab-scale equivalent.
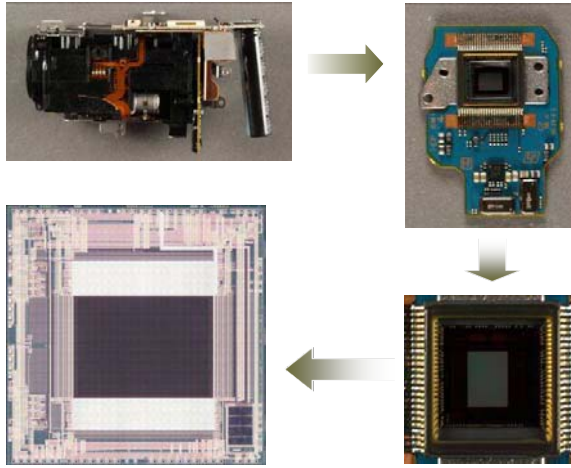
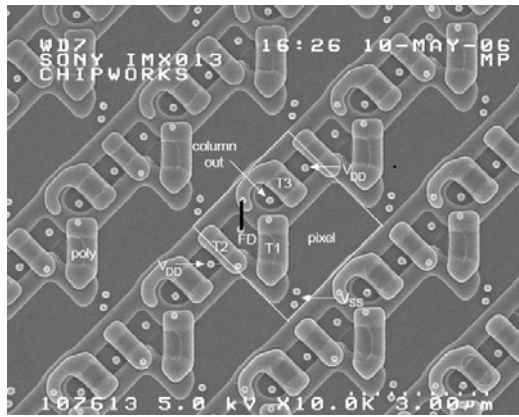**Fig. 5.** Disassembly of CMOS Image Sensor from Camera Module

Fig. 6. Plan-View SEM of Pixels at the Polysilicon Level

Using a Sony DCR-DVD505 Handycam as an example, we were interested in the CMOS image sensor in the camera.

We removed the camera module from the unit and took it apart, recording the details as we went, and ended up with the CMOS imager die (Figure 5), which turns out to be a Sony Clearvid IMX013 chip.

Then we get into the actual chip analysis. This part was a fairly leading-edge sensor, with a small pixel size of 2.85 μm x 2.85 μm, so the emphasis was on a detailed examination of the pixel. Figures 6 to 9 show some of the features seen in the pixel area.

When performing process analysis, plan-view imaging gives limited process information, so the primary source of data is cross-sectional analysis,
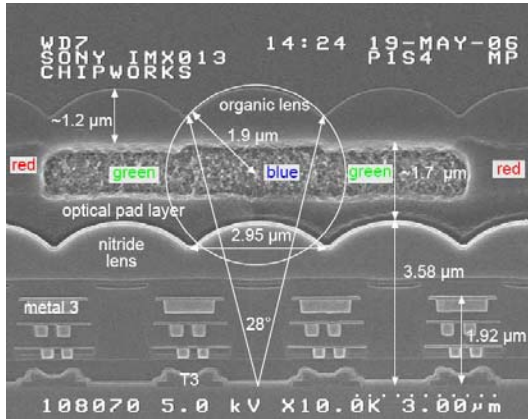
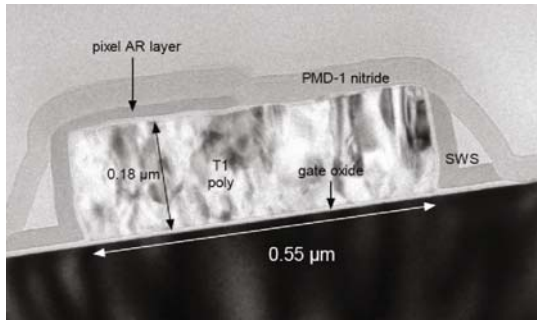**Fig. 7.** Cross-Sectional SEM of Pixels

Fig. 8. TEM Cross Section of Pixel Transfer Transistor

usually using SEM, TEM, and scanning capacitance microscopy (SCM). For details of the chemical composition, the most commonly used technique is energy dispersive x-ray analysis, although occasionally we use other methods such as secondary ion mass spectrometry or Auger analysis.

A few words of explanation here with respect to Figures 8 and 9. A TEM looks *through* the sample to give high resolution images of the device structure, and SCM is a way of seeing the positive and negative doping that makes up the actual working transistors, resistors, etc., in the silicon chip.

Looking at Figure 6, we see a plan-view image of part of the pixel array, showing the transfer transistor (T1), and the T2 reset transistor and T3 source follower transistors, comprising the 3 transistor pixel circuit. The short black line in the centre of the image represents a metal 1 strap joining the floating diffusion (FD), between T1 and T2, to the gate of T3.

Figure 7 shows a cross section of the pixel structure, illustrating the organic and nitride lenses, the colour filters, three layers of copper metallization in the array, and the T3 transistors on the substrate. There is also a fourth aluminium metal
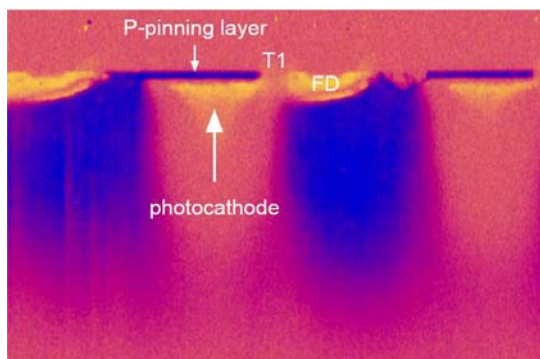
Fig. 9. SCM Cross Section of Pixels

layer, not shown in this section, used for bond pads and as a light shield (the white bars in the die photograph in Figure 4). The 28° angle of acceptance is also shown.

Figure 8 is a TEM image of the transfer transistor gate, and it is clear that the nitride layer used for the sidewall spacer has only been partially etched off the top of the gate; the residual nitride on the photocathode (left) side has been used as an antireflective (AR) layer in the photocathode area.

The doping structure of the pixels is illustrated in the SCM image in Figure 9. Chemical staining has been used for decades to highlight the doped areas in silicon, but even after many years of experiment, it is still more of an art than a science. The development of the SCM allows us to distinguish features such as the P-pinning layer above the photocathode, and the floating diffusion, more clearly. The deeper blue areas are the P-type isolation regions in the N-substrate.

There are two parallel trends in semiconductor processing. There is the well publicized Moores law shrinkage of dimensions, moving to the 45 nm node and below, with the introduction of high-k/metal gate transistors, and there is a drive to more process integration as RF/mixed signal and embedded memory processes are merged into CMOS logic processes.

As can be imagined, examining features deep into the nanometer scale (gate oxides are now 1.2 nm - 1.5 nm thick) stretches analytical capabilities to the limits. They can be imaged with high-resolution electron microscopy, but obtaining details of the chemical composition of the structure is now in the realm of counting atoms [5,6].

Similarly to the other forms of RE, our final documents can take several forms, from reports specifically focused on a feature described in a patent claim, to comprehensive reports detailing the full structural and process analysis of a high-end chip. It all depends on what the customer wants!

6   Circuit Extraction

Circuit extraction of semiconductor chips becomes increasingly more difficult with each new generation. In the "good old days" of 10 to 20 years ago, a circuit

Fig. 10. As RE Used to be Done!

analyst's life was much simpler. A typical IC of those days may have had one layer of metal, and used 1 μm - 2 μm technology. After package removal, all features could usually be seen from the top level metal planar view.

The die could then be put under optical imaging equipment in order to take multiple high-magnification images. The photographs were developed and taped together in an array to recreate an image of the chip. Engineers then used the "crawl-aroundon- the-floor" technique (Figure 10), where they annotated the wires and transistors. This was followed by drawing out the schematic first on paper, then in a schematic editor.

Life has changed since those days. The complexity of devices has followed Moores law, and we are now extracting circuits from 45 nm chips. Moreover, these devices now have up to 12 layers of metal, and use an esoteric combination of materials to create both the conductors and dielectrics [2,3]. They may have hundreds of millions of logic gates, plus huge analog, RF, memory, and other macrocell areas. MEMs, inductors, and other devices are also being integrated onchip.

The circuit extraction flow proceeds as follows:

- Package removal (known in the industry as device "depot")
- Delayering
- Imaging
- Annotation
- Schematic read-back and organization
- Analysis

6.1   Device Depot

Depot may well be the only step of the process that still follows the traditional methods. Typically, packages are etched off in a corrosive acid solution (Figure 11). A variety of acids at various temperatures are used depending on the composition and size of the particular package. These solutions dissolve away the packaging material, but do not damage the die.
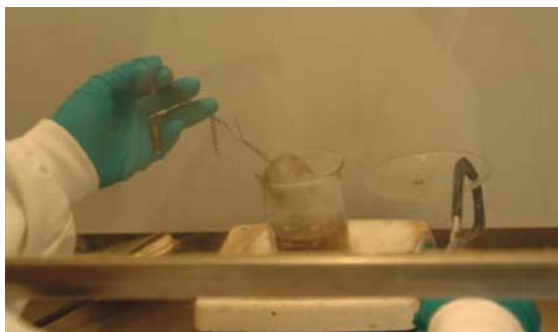
Fig. 11. Into the Acid Bath, My Pretty!

Hermetic and ceramic packages require different techniques that usually involve mechanical or thermal treatment to remove lids, or dice from substrates, or even polish away a ceramic substrate.

6.2   Device Delayering

Modern semiconductor devices range from 1.0 μm single metal bipolar chips, through 0.35 μm BiCMOS diffused MOS (BCDMOS) chips, to 45 nm 12 metal microprocessors, and everything in between. Both aluminum and copper can be used for metal on the same chip. Depending on the process generation, the polysilicon gates and source/drains can use different silicides. A variety of low-k dielectrics are now interspersed with fluorosilicate glass (FSG), phosphosilicate glass (PSG), and SiO_2. Layer thicknesses vary greatly. For instance, on a 7 metal 65 nm Texas Instruments (TI) [4] baseband processor chip we recently analyzed (Figure 12), we found:

- Interconnect layers included Cu, Al, TiN, and TaN
- Metal thicknesses ranged from 0.15 to 1.4 μm
- Dielectrics included silicon nitride, oxynitride, oxide, SiOC, SiONC, and PSG
- Dielectric thicknesses varied from ~ 0.3 μm to 2.6 μm (with individual layers of particular materials as thin as 47 nm), and gate oxide was 2.2 nm thick.

A delayering lab needs to create a single sample of the device at each metal layer, and at the polysilicon transistor gate level. As such, it needs to accurately strip off each layer, one at a time, while keeping the surface planar. This requires detailed recipes for removal of each layer. These recipes include a combination of methods such as plasma (dry) etching, wet etching, and polishing. As the complexity and variation of chips increases, so too does the number of recipes. A modern chipdelayering lab would now have over a hundred such recipes, specific to different processes and materials.

For unknown or unusual chips, it is advisable to start with a cross section (Figure 12). The cross section can be analyzed using scanning electron microscopes (SEM), transmission electron microscopes (TEM), and other techniques
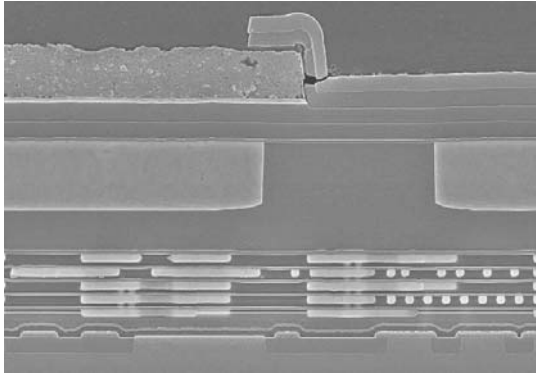
Fig. 12. SEM Cross Section of 65 nm TI Baseband Chip for Nokia

to determine the composition and thickness of all the layers. A delayering technician uses this information to choose the best delayering recipe for a chip. The recipe also varies depending on the type of imaging to be performed. Optical imaging looks best if the transparent dielectric is left on over the layer to be imaged. SEM, due to its operating methodology of electron reflection from a non-planar surface, requires the dielectric to be removed.

6.3   Imaging

Advanced RE labs currently use two types of imaging, optical and SEM. Up to and including the 0.25 μm generation of semiconductor chips, optical imaging was sufficient. However, for 0.18 μm technologies and smaller, optical imaging cannot resolve the smallest features, and SEM must be used (Figure 13).

The size of ICs, and the large magnifications required for the advanced feature sizes, now means that manually shooting images is no longer practical. Imaging systems now must have automated steppers integrated with the microscope. Our twodimensional steppers allow us to set up a shoot in the evening, and come back in the morning to find the entire layer imaged.

Next we use specially developed software to stitch the thousands of images per layer together, with minimal spatial error. Then more software is required to synchronize the multiple layers so that there is no misalignment between layers. Contacts and vias must be lined up with the layers above and below in order for extraction to proceed.

6.4   Annotation

Once all images are stitched and aligned, the actual work of reading back the circuit begins. Full circuit extraction requires taking note of all transistors, capacitors, diodes, and other components, all interconnect layers, and all contacts and vias. This can be done manually or using automation.
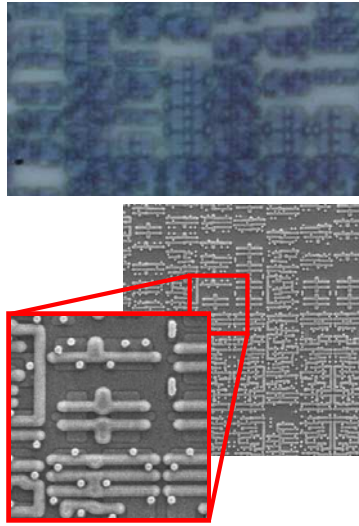
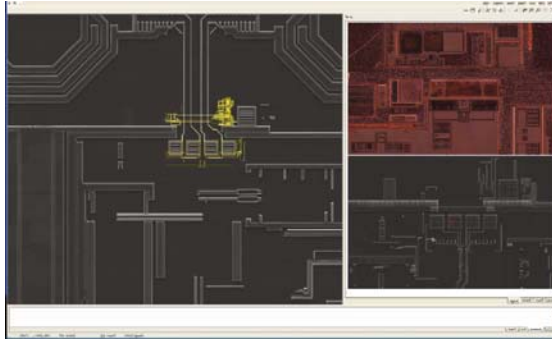**Fig. 13.** Optical (top) and SEM images of 130-nm chip

Fig. 14. Power Net of an RF-Switch Inductor and VCO Subsystem

There are multiple tools available to help with this process, including Chipworks' ICWorks Extractor. This tool is used to view all the imaged layers of a chip individually and aligned to each other. In one mode it allows several layers of a chip to be visible in multiple windows simultaneously (Figure 14). Each window shows the same two-dimensional area in each layer. A lock-step cursor allows the engineer to see exactly what lies above or below the feature he is looking at in one layer.

An extraction engineer can then use the tool to annotate and number all wires and devices in his area of interest (Figure 15). 2D and 3D image recognition and processing software can be used (Figure 16), or the engineer may do it manually. Image recognition software can also be used to recognize standard cells in digital logic. This can greatly aid the extraction of large blocks of digital cells.
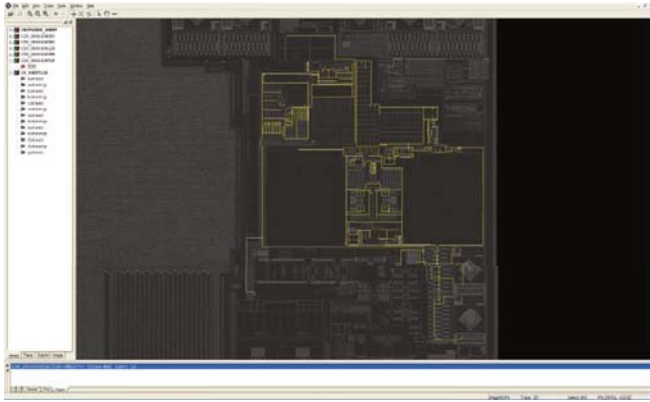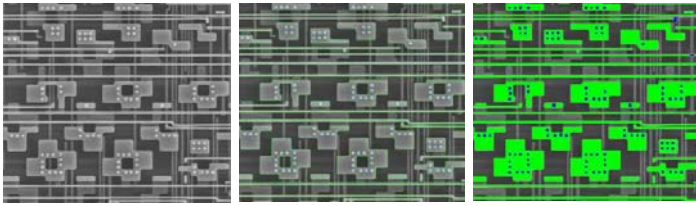
Fig. 15. Annotated VDD net on an RF Transceiver

Fig. 16. Automated feature extraction from SEM images

6.5 Verification and Schematic Creation

The annotation process can be error prone. Often the images are not perfect, manual techniques are used, bits of dust fall on the chip during imaging, or the image recognition software introduces an error. Hence, verification is performed at this stage. Design rule checks can find many issues, such as below minimum sized features or spaces, hanging wires, vias without wires, etc.

At this stage the ICWorks tool can automatically extract a netlist from the annotations, and from this netlist create a flat schematic (see Fig. 17). The schematic, netlist, and annotations are all associated with each other, such that one cannot be changed without changing all three.

The netlist and schematic can now be checked for other simple rule violations. Floating gates, shorted outputs, nets with no inputs or outputs, and shorted supplies can be checked.

6.6 Schematic Analysis and Organization

This is one of the steps requiring the most thought, since the schematic organization on a page, or in hierarchy, goes a long way to making a design coherent. Devices placed poorly on a schematic, or a strange hierarchy, can make the design
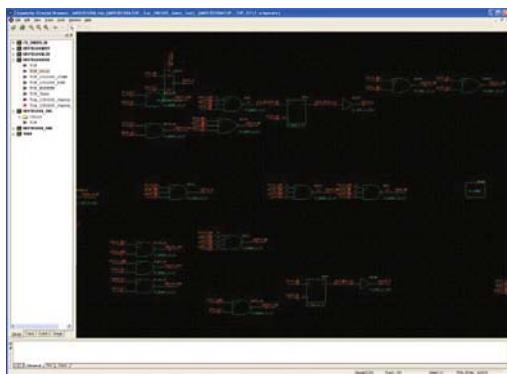
Fig. 17. Flat Schematic Auto-Exported from Annotated Images

very difficult to understand. Hence, this step usually requires very experienced analysts.

The analysis phase can be very iterative, and use many sources of information. Often public information is available for devices. This can take the form of marketing information, datasheets, technical papers, or patents. These can often help with the schematic organization, for instance if block diagrams are available. They can also help in the understanding of architectures and sometimes circuit designs.

Analysis can also be done using typical chip design techniques. A circuit can be hand analyzed using transistor and logic theory. Layout structures are often recognizable, for instance differential pairs, bipolar devices for bandgap references, etc. In fact, The ICWorks tool can find these structures automatically. Hierarchy can also sometimes be seen in the layout. If not, it can be created using a bottom-up schematic organization approach. Functional and timing analysis can be further validated using simulation. Multiple methods are usually used for verification.

The final product of circuit reverse engineering can take many forms. A complete set of hierarchical schematics can be delivered. This set of schematics can be used to also create a hierarchical netlist. Simulated waveforms, block diagrams, timing diagrams, analysis discussion, and circuit equations can be used to round out the report.

Since RE companies analyze so many ICs, they can also create comparative and trend reports. For instance, Chipworks has analyzed many CMOS image sensors over the years. As the technology and circuit designs evolve, they are monitored. The evolution can be shown from both a process point of view and a circuit point of view.

7   A Case Study

Used together, the above techniques can be very powerful. To illustrate that point, lets review a project we just finished; analyzing a digital ASIC with
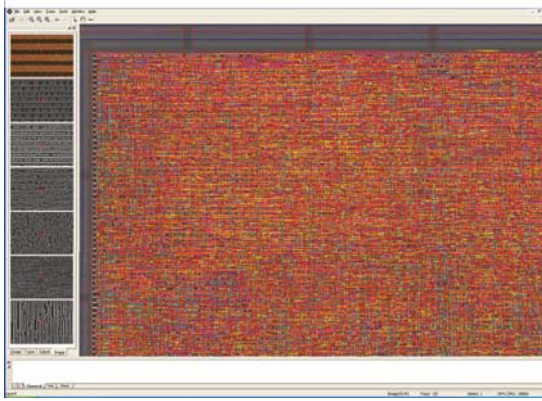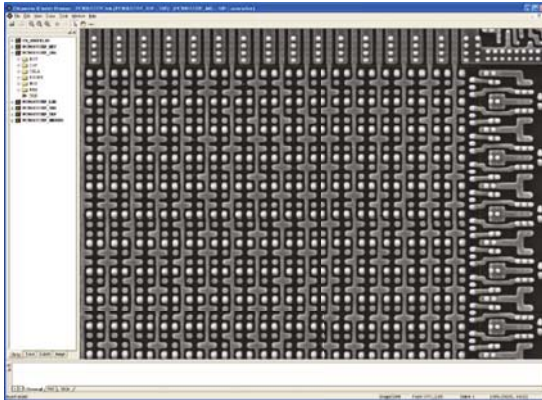
Fig. 18. Annotated digital logic

Fig. 19. Metal 1 image of mask programmed ROM

embedded analog and memory blocks, and including embedded encryption hardware. The goal of the project was to fully understand the ASIC, build a model of the ASIC, and get simulations up and running.

The first step was to run system level functional tests while the chip was still in its system. Logic probes were connected, the system was powered up, and vectors were collected which could be used later for simulations.

Next, the chip was depotted, delayered, imaged, stitched, and aligned. We found the chip contained 12,000 gates of digital logic and an embedded EEPROM. The entire chip was annotated, and the ICWorks tool created a netlist and flat schematic from this annotation. A portion of the digital logic annotation is shown in Figure 18. Annotation and schematic rule checks were used to verify a quality schematic starting point. In fact, for this project we annotated the entire chip twice, then compared the results to minimize annotation errors.
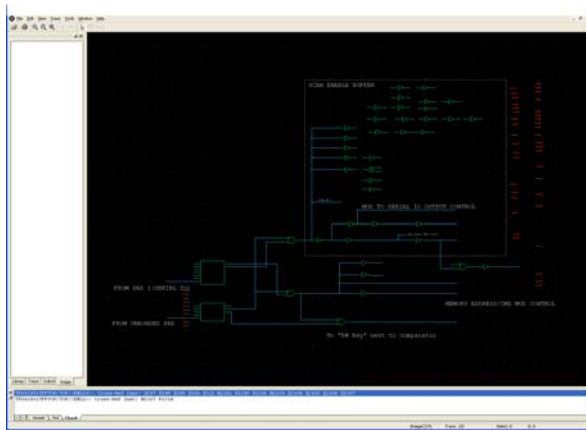
Fig. 20. Logic Scan Control Circuit

The schematics were then partially organized. The memory schematic was completely organized, and the main registers of the digital block were grouped. A few of the major busses were labeled and the I/Os were connected to the major blocks.

In order to run a full chip simulation on the netlist, we would need to extract all the contents of the chip, including both the hardware and memory contents. Different memory types have different challenges in reading them. Embedded SRAMs are the simplest. These memories are volatile, no data is stored in them during power down, so they do not need to be extracted. ROMs can be extracted using traditional RE techniques of physically reading back the mask programming. Figure 19 shows a metal 1 mask programmed ROM. Unfortunately EEPROMs are more difficult than either of these.

We knew up front that this chip included on-chip encryption, and that the keys were stored in the EEPROM. Hence, we anticipated a challenge in being able to access this memory. As expected, the memory was well protected, and much of this memory could not be directly read off-chip. Additionally, the interface to this chip was encrypted, so we had no idea how to generate a memory read command anyhow. The solution to this was to use the test hardware embedded in the chip.

This particular chip had both scan path test circuitry for the digital logic, and memory BIST for the EEPROM. Once we had organized the test and memory circuits, we set to work analyzing them. The scan test control circuit is shown in Figure 20. We found a method where we could almost read out the memory locations using a combination of the digital and memory test circuitry. A single application of microsurgery looked as though it would unlock the bits.

We took a single chip, used jet-etching to remove a portion of the package, then used focused ion beam (FIB) techniques to modify a connection on the chip (Figure 21). Next we used our analysis to create scan path vectors, with the appropriate control signals, and successfully read out the encryption keys and other memory contents via the test port.
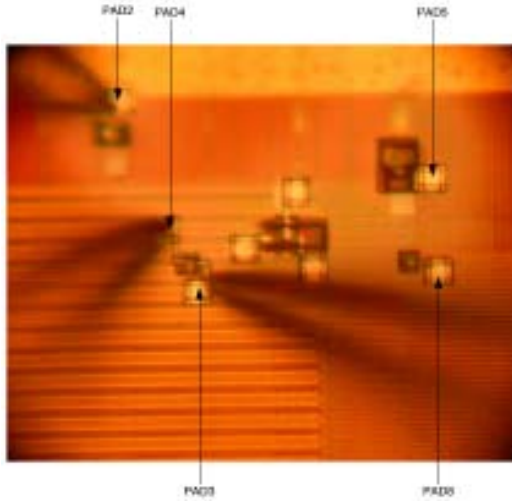
Fig. 21. Microsurgery Altered Chip Being Pico-probed

At this point, we created a memory model to use with our netlist. The vectors collected from the actual system were run on the netlist, and we verified that our chip model gave the same outputs as the actual chip tested. Hence, we confirmed our netlist and memory contents were correct.

The encryption algorithm also needs to be understood to be able to complete the analysis of this chip. This was accomplished via schematic organization and simula- tion. As we organized the chip, we found some interesting structures, such as a 56 bit register. Therefore, we ran our simulations, and monitored the busses in the area of this register. Sure enough, keys were read from our memory model, loaded into this embedded block, and a standard DES algorithm was observed.

Now we understood the encryption, had the keys, and had full chip simulations running. Since we had a full netlist, we were able to run full chip simulations and monitor any internal nodes required. This allowed us to complete the analysis of this chip and understand all the commands it could execute.

8   Summary

In this paper we have reviewed the different types of reverse engineering pertinent to the semiconductor industry. For reverse engineers, life will not get any easier in the electronics business. In semiconductors, the next challenge will be the 32 nm node devices already being ramped up in development fabs. The consumer electronics business keeps bouncing from new toy to yet another new toy, and it is necessary to be aware of all the new products that keep appearing.

As is shown in this paper, the RE business has to keep evolving to keep up with the changes in electronics and design, and it has become a discipline in

itself, created by the needs of the global market for competitive intelligence and IP support.

Acknowledgements

References

1. James, D.: A Case Study: Looking Inside Apple's iPOD Nano– a Teardown to the Atomic Scale,
   `http://electronics.wesrch.com/Paper/`
   `paper_details.php?id=EL1SE1KWRX174&paper_type=pdf&type=author`
2. Nii, H., et al.: A 45 nm High Performance Bulk Logic Platform Technology (CMOS6) using Ultra High NA (1.07) Immersion Lithography with Hybrid Dual-Damascene Structure and Porous Low-k BEOL. In: IEDM 2006 Technical Digest, pp. 685–688 (2006)
3. Narasimha, S., et al.: High Performance 45 nm SOI Technology with Enhanced Strain, Porous Low-k BEOL, and Immersion Lithography. In: IEDM 2006 Technical Digest, pp. 689–692 (2006)
4. Chatterjee, A., et al.: A 65 nm CMOS Technology for Mobile and Digital Signal Processing Applications. In: IEDM 2004 Technical Digest, pp. 665–668 (2004)
5. 2005 ITRS, Metrology section
6. Vartanian, V., et al.: Metrology Challenges for 45 nm Strained-Si Devices. In: 2005 International Conference on Characterization and Metrology for ULSI Technology (2005)