Silicon stowaways

Trojans aren’t restricted to the software domain; in some instances, hardware Trojans could even open backdoors in custom silicon. By Chris Edwards.

 Outsourcing has reshaped the way electronics products are made – and helped to cut manufacturing costs massively. But, as production margins have fallen, so too has trust in the organisations that make up the supply chain. Companies which rely on outsourced manufacturing are having to come up with ways of ensuring that the products shipped to them have not had secure keys leaked or stuffed with viruses and compromised software. Even custom silicon is not safe.

Almost a decade ago, researchers from Case Western Reserve University described to delegates at the IEEE High-Level Design Validation and Test Workshop the ways in which they could see hardware malware – or Trojans – being introduced to an IC-design project. The widespread use of foundry services, third-party intellectual property and standard-cell libraries – as well as designers bribed to make circuit-level changes – all provide ways in which Trojans could be sneaked into circuitry.

Once it receives a trigger signal, the Trojan could open a backdoor to the group that wanted it introduced. In some use-cases, the Trojan may be introduced simply to compromise the product; no matter how it is used.

However, the nature of IC design makes hardware Trojans difficult to deploy as they require skills and levels of access that are probably out of reach of most cybercriminals. But other, lower hanging, fruit remains available to them. State actors have the skills, access and motive that may make the surreptitious deployment of some kinds of silicon highly attractive to them. In practice, such organisations may not bother trying to introduce backdoors without the knowledge of the manufacturer or find other ways to make gain access to secrets.

In 2015, the BBC identified declassified documents that confirmed the government convinced Crypto AG in the mid-1950s to compromise the security of its C-52 electromechanical encryption machines. Rather than making physical changes to the hardware itself, the company told the US National Security Agency (NSA) and the UK’s GCHQ which models target governments had bought – a practice that would allow the agencies to target decryption resources more effectively.

In 2013, the NSA came under suspicion of encouraging the use of algorithms supplied by specialist RSA that had been subtly weakened to make decryption easier.

While some ICs behave as if they have Trojans installed, in reality the backdoors were placed intentionally into the silicon by authorised designers. Usually, they are debug aids that were meant to stay secret, but often did not. Five years ago, Sergei Skorobogatov of the University of Cambridge and Christopher Woods of Quo Vadis Labs used side-channel emissions from the devices to uncover the key that would open the backdoor in the JTAG circuitry of an FPGA and to provide access to the encryption keys stored inside.

Although side-channel emissions provide one way to determine whether an IC has been compromised with a backdoor, designers have other options available through the deployment of EDA techniques with Trojan detection in mind. As with anything in cybersecurity, a cat-and-mouse game has produced ever more subtle ways of introducing Trojans and more powerful ways of detecting them.

The hardware Trojan is the subject of regular hacking competitions between research teams. For example, the Cyber Security Awareness Week (CSAW) organised by New York...
University has run several challenges around Trojans. In these challenges, red teams try to circumvent the detection mechanisms used by blue teams.

2013’s CSAW challenge focused on methods to beat FANCI, a largely effective detector developed at Columbia University and NYU. FANCI works on the basis that a Trojan would only have a loose connection to the design such that its logic would seem to be practically unreachable. Code-coverage analysis of the RTL can identify such unconnected lumps of circuitry and flag them up as possible Trojans.

A couple of years later, the DeTrust technique created by Jie Zhang and colleagues at the Chinese University of Hong Kong showed one method for fooling FANCI: spreading the suspicious logic across many otherwise independent gates.

Months later, Syed Haider and coworkers from the University of Connecticut developed the hardware Trojan catcher (HaTCH), designed to track down stealthier functions inserted at the logic level. Rather than isolate the Trojans before manufacture so they can be removed, HaTCH focuses on remediation. It adds tagging circuitry to legitimate cores that work to prevent any on chip Trojans from activating or succeeding in displaying malicious behaviour, such as opening a backdoor.

A more wide-ranging technique that could serve as a defence against Trojans is to insist that all IP be supplied with formal proofs that describe the operations it would be allowed to perform. Any changes would be flagged by formal-verification tools during design and prototyping. Such approaches could still be vulnerable to attacks that tweak designs below the abstraction of RTL.

As with the cases where the NSA is understood to have sought the help of manufacturers, weakened encryption is one of the most likely ways in which a practical hardware Trojan might work and evade detection by all but side-channel analysis. And it is possible using a tiny change at manufacture. At the 2013 International Workshop on Cryptographic Hardware and Embedded Systems, Georg Becker and colleagues from the University of Massachusetts at Amherst showed a proof of concept that simply switched dopants used for one of the transistors in an inverter.

In 2014, Takeshi Sugawara of Mitsubishi Electric and a team from the company and Ritsumeikan University showed such a tiny change in manufacturing could be detected after the fact. A combination of focused ion beams and scanning electron microscopy can reveal the dopants diffused into the substrate. It is an expensive proposition, involving delayering of the design and extensive analysis against a layout that contains a map of the expected dopants. However, for the kinds of high-value cryptographic IC that might be the targets of well-financed attackers, it is arguably one more in a list of checks that are readily justified.

As with other areas of embedded cybersecurity, the most feasible approach to dealing with the risk of hardware Trojans is one of focusing effort. Architectures such as Trustzone pull functions that need high levels of protection into a small portion of the overall SoC. In principle, this subset is much easier to verify than a design that calls for the either chip to be labelled as a whole.

If the secure core is guaranteed to not leak information or provide trapdoors, the value to an attacker of putting a Trojan in the more weakly protected part of the SoC diminishes greatly. For the user of SoCs and the buyer of IP to go into them, the question then becomes one of the level of expected risk and the degree of trust they can put in staff, suppliers and contractors.

---

**Figure 1:** By switching dopants, a transistor could be created that doesn’t switch.