Issues in TECHNOLOGY Innovation

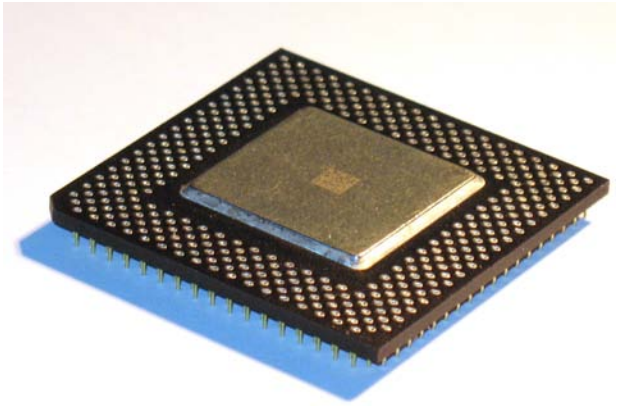
Number 9 May 2011

Ensuring Hardware Cybersecurity

John D. Villasenor

EXECUTIVE SUMMARY

More than ever, the electronic devices that are critical to everyday life, to the larger infrastructure, and to national defense are dependent on increasingly sophisticated semiconductor integrated circuits, also referred to as "chips". For example, laptop computers and tablets, smartphones, the financial system, the Internet, aircraft flight controls, automobile antilock braking, the power grid, and an almost endless list of other devices and systems can be trusted to run properly only if the chips they contain are free of hidden malicious circuits inserted during the design or manufacturing process.



© CPU Intel Celeron 400 socket 370

The combination of continued chip technology advances and an unprecedented level of globalization in the semiconductor industry has spurred enormous changes in the way chips are designed, manufactured, and used. These changes bring many benefits to the consumer including lower prices and faster time to market for products and services, but they have also created a widening set of opportunities for would-be attackers to insert malicious circuits during the chip design process that could be used to launch a hardware attack.

Despite the potentially devastating impact that a large-scale hardware attack could have on commerce, defense, and government function, the need to proactively address hardware security remains widely underappreciated. This document explains the causes and nature of the hardware security threat and outlines a multipronged approach to address it involving 1) a change in design practices within the semiconductor industry, 2) the establishment of a national-level capability to coordinate a quick response to an attack, 3) improved testing procedures to detect

Issues in Technology Innovation

The Center for Technology Innovation at Brookings has launched its inaugural paper series to seek and analyze public policy developments in technology innovation.

The Center for Technology Innovation

Founded in 2010, the Center for Technology Innovation at Brookings is at the forefront of shaping public debate on technology innovation and developing data-driven scholarship to enhance understanding of technology's legal, economic, social, and governance ramifications.

corrupted chips before they are placed into products, and 4) the inclusion of built-in defenses into chips to identify and thwart attacks as they occur.

Introduction

John D. Villasenor is a professor at the Electrical Engineering Department at UCLA and Director of the UCLA EE Image Communications Laboratory.

Modern electronic semiconductor integrated circuits are marvels of miniaturization, capable of packing into a circuit approximately the size of a nickel an amount of computational power far exceeding that delivered by an entire roomful of computers in the 1960s. By enabling the devices and networks that allow information to be almost instantly acquired, processed, searched, and shared, integrated circuits, or "chips", have revolutionized the technology landscape. In doing so, they have played a key behind-the-scenes role in reshaping social interactions and the political landscape as well.

The semiconductor industry has become an enormous global enterprise that generated nearly $300 billion in sales in 2010. There are approximately 1500 companies in the world today engaged in chip design. The full chip ecosystem, which includes chip designers and manufacturers, companies that use the resulting chips in products, and individuals, corporations, and governments that in turn purchase these products, relies on the assumption that the chips at the core of these products have integrity.

Until the very recent past, this has been a reasonable assumption. However, continued changes in the dynamics of the global semiconductor market have made it not only possible but inevitable that chips that have been intentionally and maliciously altered to contain hidden "Trojan" circuitry will be inserted into the supply chain. These Trojan circuits can then be triggered months or years later to launch attacks. It is imperative to put into place systems that can minimize the likelihood, number, and consequences of these attacks.

What is a Hardware Attack?

With respect to chips and the devices in which they reside, tasks such as performing an Internet search, editing a document on a computer, making a mobile phone call, and conducting a financial transaction rely on a complex interplay between software and hardware. "Software" refers to the set of instructions that describe how a task is performed, while "hardware" refers to the circuits that actually do the work to perform the task.

While software can be replaced, updated, altered, and downloaded from the Internet, chip hardware generally can't be changed after the chip leaves the factory. Thus, while malicious software can be created and disseminated by anyone with a computer and access to the Internet, malicious hardware can only be inserted by someone who can access and alter the design for a chip before it is manufactured and placed in a product.

Once malicious hardware has been built into a chip, a hardware attack can be initiated and act in a wide variety of ways. An attack can be internally triggered, based, for example on the arrival of a particular calendar day. Alternatively, an external trigger could be hidden within data sent by an attacker. More complex hybrid triggers could also be used. For example, a malicious circuit hidden within a GPS chip could be configured to attack only when the chip is located in a specific geographical area after a certain date.

There are multiple forms of potential attacks. In an overt attack, the malicious hardware could cause the device containing the corrupted chip to either cease functioning altogether or to continue to operate but in an obviously impaired manner. The existence of a problem would be clear, though its cause would not. In a personal electronics device such as a mobile phone such an attack could be nothing more than an inconvenience. If conducted on a large scale on thousands of chips within a critical portion of the national infrastructure, this form of attack could be devastating.

Once malicious hardware has been built into a chip, a hardware attack can be initiated and act in a wide variety of ways.

In contrast with an overt attack, in a covert attack the appearance of normal operation is maintained while malicious actions are quietly being performed in the background. For example, a corrupted chip within a communications system could be caused to send copies of confidential data to a third-party destination.

A third form of attack leaves the device operating, but introduces corruptions into the data. As an example, consider an attack hidden in a GPS chip with a location-based trigger that left the GPS functioning accurately until it was located in a certain geographical region. Upon receiving this geographical trigger, it could act by shifting GPS locations by a few hundred feet. An attack of this form would be extremely difficult to detect in advance and could have significant consequences across a wide range of application scenarios. There can also be attacks that exploit a hybrid of malicious hardware and software. Malicious hardware hidden within a chip could be triggered months or years later to open a back door allowing the installation of malicious software which in turn could launch the attack.

The Growth of Chip Complexity – and Vulnerability

It is the very size and complexity of modern chips that creates the vulnerabilities that make the insertion of malicious hardware possible. Gordon Moore, a co-founder of Intel, famously predicted in a 1965 paper that the amount of functionality that can be built into a single chip would double approximately every two years. Remarkably, "Moore's Law" has stood over a time span now approaching half a century. This level of sustained growth is made possible in large part due to continued advances in the ability to miniaturize the size of the structures within a chip. In the 1970s the smallest structure sizes that could be created were several thousand nanometers in width (one nanometer is one one-billionth of a meter). Today, the most advanced semiconductor manufacturing methods are able to create structures as small as a few tens of nanometers – hundreds of times smaller than what was possible in the 1970s.

The existence of chip vulnerabilities today can be traced in part to the history of chip design, and in this respect has analogies with the growth of the Internet. As is well known, the Internet was originally designed by a small community of researchers as an open environment accessed by a small number of trusted users. This assumption of trust became problematic with enormous growth in the size of the Internet, and with it, an increasing number of users willing to exploit weaknesses in the Internet for malicious purposes.

Chip design too has its heritage in small, self-contained teams with a shared interest in the success of the chip and in which there was no reason to question trust. Many of the practices used in chip design and many of the protocols for processing data and for communicating data among parts of a chip were established in that environment. In today's world, however, the design process for a single chip can involve contributions hundreds of people, many of whom may be employed by third party companies that simply provide functional blocks and who have little or no stake or interest in the success of the chip.

The Globalization of Chip Design

The process of creating a chip includes the two sequential phases of design and manufacturing. Design involves determining the functionality of the chip and mapping that functionality into a description in terms of electronic circuits. Manufacturing refers to physically producing the chip containing these circuits. Insertion of malicious hardware during manufacturing is very difficult because of the likelihood that the insertion process itself will lead to impairments that would be detected during post-manufacturing testing. For an attacker, the low hanging fruit lies in the design process, where there is the potential to create malicious circuits and bury them within the much larger set of healthy circuits in a non-disruptive manner.

The number of people who are in a position to access and therefore potentially compromise chip designs is vastly smaller than the number of people who could create malicious software. But, in absolute terms, thanks in large part to outsourcing, it is still very large, with many hundreds of thousands of people around the world directly employed in the chip design industry.

Chip design today relies heavily on outsourcing. Although a complex chip is a single, physically small device, it contains many different functional areas, called "blocks," that perform different tasks. A chip used in a smartphone, for example, may have a set of functional blocks devoted to receiving a wireless signal, processing that signal to extract the data it contains, decoding that data to produce audio and video signals, and sending those signals to a speaker and display screen. Much as an architectural firm charged with designing an enormous shopping complex might subcontract out portions of the design, a company overseeing the design of a complex chip typically designs some portions in house but obtains designs for other portions from third parties. While outsourced chip manufacturing has been common for several decades, the use of outsourcing in chip design has accelerated dramatically in

the last half decade. As with all outsourcing, the primary driver for this is economic. Labor costs are much lower in places like India and China, and companies located in high-cost labor markets cannot remain competitive if they rely exclusively on in-house chip designers. The combination of growth in both complexity and outsourcing means that the number of people with access to the design for a single chip during its development can easily number in the hundreds.

Given this landscape, there are multiple potential vectors for the insertion of malicious hardware. One possibility is that a company performing outsourced design services could intentionally provide a corrupted design with the full knowledge and participation of the company management. Alternatively, the design services company could act in good faith, but could store the designs on weakly secured networks, enabling the designs to be accessed and altered by an outside party. It is also possible for one or more individuals within the design services company to corrupt a design without the knowledge of their colleagues or managers.

The testing procedures are very good at identifying accidental design flaws, but are poorly suited to ferreting out intentionally hidden malicious circuitry.

Finally, malicious hardware insertion can also occur in non-outsourced portions of the design. An "inside job" perpetrated by one or more rogue employees employed at the home company overseeing the entire chip design could be particularly hard to detect because of the higher level of access and knowledge regarding the overall chip that these employees often possess. As an alternative, network vulnerabilities could be exploited to access and corrupt non-outsourced designs.

The Challenge of Testing

In an ideal world, corrupted designs would be detected, regardless of their source. However, the sheer complexity of modern chips greatly impedes such detection. While extensive – but not exhaustive – testing is performed during the design and manufacturing process, the goal of this testing is to confirm that a chip is behaving as expected. The testing procedures are very good at identifying accidental design flaws, but are poorly suited to ferreting out intentionally hidden malicious circuitry.

Consider the following example: Suppose that a company outsources the design for a block of the chip that is supposed to add the number six to any input. During testing, if 20 is provided to this block, the block outputs 26. When 127 is provided, the block outputs 133. One hundred thousand more inputs are provided, and in every case, the result comes back correct. This block will be deemed to have passed functional testing. But the block could have a hidden circuit triggered by an input with value 126,321,204. When that input – and that input alone – arrives, an attack is launched. Because testing can't possibly be exhaustive, this input will never be encountered until it is provided months later by an attacker.

Towards a Comprehensive Solution

If a significant hardware attack were to be launched today, we would be ill-equipped

to respond. In stark contrast with the large amount of attention and resources being directed to ensuring software security, efforts to address the potential impact of the contamination of the commercial chip supply by malicious hardware are in their infancy at best. The following steps could go a long way toward reducing the likelihood and impact of hardware attacks:

- A change in design practices within the semiconductor industry

  Current design practices in the semiconductor industry should be modified to specifically recognize and address the potential for malicious hardware insertion. First, companies engaged in chip design should adopt a need-to-know partitioning of information, much as occurs within the Department of Defense or in companies engaged in defense contract work. A designer working on a portion of a chip devoted to receiving wireless data does not need access to the internal details of a portion of the chip that processes video for display on the screen. But, in many companies, the barriers that separate design access are either nonexistent or insufficiently high. Second, companies engaged in chip design should recognize the existence of a real and significant threat from malicious hardware – a threat that is not generally even on the radar screen in any meaningful way at most such companies. This could lead to a more careful scrutiny of third-party suppliers as well as to measures that would reduce the odds that designs could be compromised under their own roofs.

- Establishment of a national-level capability to coordinate a quick response to an attack

  Currently, we do not have any national level capability to respond to an attack. This would greatly impede our ability to respond in an agile manner. For example, if an attack significant enough to require a national-level response were to occur today, it is not clear which governmental entity would oversee the response, and valuable time would be lost in creating the appropriate organizational framework and in establishing the appropriate lines of communication and information flow.

  To avoid this scenario, the entity that would be charged with overseeing the response to a hardware attack should be identified or created preemptively. Procedures should be put in place for reporting an attack and for engagement with the appropriate companies and governmental organizations.

  In the event of an attack, it would also be critically important to be able to rapidly identify the other chips containing designs received from a known supplier of corrupted hardware. Currently, this identification could take many weeks or longer – precious time that could confer significant advantages to an

attacker. However, a government-managed database of suppliers could allow this identification to be nearly instantaneous. It is not feasible to simply ask companies to furnish the government a list of suppliers every time they unveil a new chip, as this information is typically considered highly proprietary. However, with proper design using multiple layers of asymmetric encryption, companies could provide supplier information to the government that would enable rapid tracing while also avoiding the unnecessary disclosure of proprietary information.

- Improved testing procedures to detect corrupted chips before they are placed into products

  Today's commercial chip testing procedures are designed to identify accidental design flaws, not to discover intentionally hidden attacks. By developing new testing procedures that are specifically designed to look for attacks, the odds that corrupted hardware could escape pre-deployment testing would be significantly reduced. Fortunately, the research arm of the Department of Defense, the Defense Advanced Research Projects Agency (DARPA), has been funding a program for the past several years directed to testing untrusted circuits to look for malicious hardware. The solutions developed under this program have the potential to play a significant role in reducing the odds that compromised hardware will be deployed.

- Inclusion of built-in defenses into chips to identify and thwart attacks as they occur.

  While pre-deployment testing is extremely important, it cannot be relied on to find all instances of malicious hardware. Indeed, as stated in a 2005 report by the Defense Science Board Task Force on High Performance Microchip Supply, "[E]lectrical testing and reverse engineering cannot be relied upon to detect undesired alterations in military integrated circuits" (Defense Science Board Task Force on High Performance Microchip Supply, 2005). This challenge was also noted by U.S. Deputy Secretary of Defense William J. Lynn III in a September/October 2010 article in Foreign Affairs, who noted with respect to hardware that "[t]ampering is almost impossible to detect." (Lynn, September 2010)

  Given the inevitability that some number of compromised chips will slip past the testing process, it is important to build defenses into chips that can identify and respond to attacks within milliseconds (Villasenor, "The Hacker in Your Hardware", Scientific American, 2010). This could be accomplished by adding a modest amount of circuitry specifically charged with the task of monitoring the behavior of the chip and identifying behavior that may be indicative of an attack. When an attack is discovered, the offending portion of the chip could be identified and quarantined, and a notification sent to other

Given the inevitability that some number of compromised chips will slip past the testing process, it is important to build defenses into chips that can identify and respond to attacks within milliseconds.

devices containing similar circuits. Clearly, this requires that the circuits doing the policing are themselves free of corruption – which is more likely to occur if the "policing" portions of the chip are designed in-house using a very small group of highly trusted engineers.

In sum, while it is impossible to completely eliminate the potential for hardware attacks, the above measures, in combination, could significantly reduce the odds of a successful attack.

The Center for Technology Innovation
The Brookings Institution
1775 Massachusetts Ave., NW
Washington, DC 20036
Tel: 202.797.6090
Fax: 202.797.6144
http://www.brookings.edu/techinnovation

Editor
Christine Jacobs

Production & Layout
John S Seo

Tell us what you think of this *Issues in Technology Innovation*.

E-mail your comments to techinnovation@brookings.edu

References:

Report of the Defense Science Board Task Force on High Performance Microchip Supply, published by the Office of the Under Secretary of Defense For Acquisition, Technology, and Logistics, February 2005.
http://www.acq.osd.mil/dsb/reports/ADA435563.pdf

"Defending a New Domain: The Pentagon's Cyberstrategy," William J. Lynn III, Foreign Affairs, September/October 2010

Villasenor, John, "The Hacker in Your Hardware," Scientific American, August 2010, vol. 303, no. 2, pp. 82-87

Further reading:

Old Trick Threatens the Newest Weapons, by John Markoff, The NewYork Times, October 26, 2009, http://www.nytimes.com/2009/10/27/science/27trojan.html

"Securing the Information Highway," Gen. Wesley Clark and Peter Levin, Foreign Affairs, November/December 2009.