Ransomware continues to grow very rapidly—with the number of new samples rising 58% in Q2.

About McAfee Labs

McAfee Labs is one of the world’s leading sources for threat research, threat intelligence, and cybersecurity thought leadership. With data from millions of sensors across key threats vectors—file, web, message, and network—McAfee Labs delivers real-time threat intelligence, critical analysis, and expert thinking to improve protection and reduce risks.

McAfee is now part of Intel Security.


Follow McAfee Labs

Introduction

This month marks the five-year anniversary of Intel’s announcement that the company would acquire McAfee. Much has changed in the security space since then, so we decided to look back on these years and compare what we thought would happen with what actually happened.

We interviewed a dozen key people who have been with Intel or McAfee since the acquisition to get their views on the major developments of the past five years around the cyber threat landscape, including how the types of threat actors have changed, how attackers' behaviors and targets have changed, how the economics of cybercrime have changed, and how the industry has responded. We also wanted to know what they didn’t anticipate or what truly surprised them. We hope you enjoy the retrospective.

This quarter, we also discuss two very interesting Key Topics.

In McAfee Labs Threats Reports, we spend a lot of time examining ways in which attackers enter a trusted network or system, but we spend little time looking at how they exfiltrate the information they want to steal once they have successfully breached the network or system. In this Key Topic, we leverage the considerable experience of our McAfee Foundstone forensic consulting team to detail the specific tactics and techniques used by attackers to surreptitiously remove targeted data.
Malware attacks on graphics processing units (GPUs) have existed for several years without gaining much attention. Recently, proof-of-concept code was posted on GitHub purportedly demonstrating how GPUs can be used by attackers to evade detection by running malware and storing data on those devices. In this Key Topic, we dissect the claims and clarify what can and cannot be done through this form of attack.

Other items of note:

- **Black Hat USA 2015** took place at the beginning of August. Intel presented two sessions, one of which illustrates how research performed together by Intel and Intel Security leads to better hardware protection. The session “Attacking Hypervisors Using Firmware and Hardware” explores the attack surface of modern hypervisors from the perspective of vulnerabilities in system firmware, such as BIOS and in hardware emulation. The presentation will be available here soon after the conclusion of Black Hat.

- As we reported last quarter, the underlying cloud infrastructure of **McAfee Global Threat Intelligence** was replaced to handle more queries, more threat data, and more reputation types. It was also re-architected to be faster, safer, more secure, more resilient, and easier to manage. Foundational to that is its new RESTful architecture. In Q2, that architecture was fully implemented in McAfee GTI across the globe.

- In 2014, we formed a data sciences team to better understand and leverage the data inside McAfee GTI. The team has developed McAfee GTI cloud instrumentation coupled with a dashboard that allows us to see and analyze real-world attack patterns which will lead to better customer protection. These numbers provide insight into the attack volumes that our customers experience. In Q2, our customers saw the following attack volumes:
  - Every hour more than 6.7 million attempts were made to entice our customers into connecting to risky URLs (via emails, browser searches, etc.)
  - Every hour more than 19.2 million infected files were exposed to our customers’ networks.
  - Every hour an additional 7 million PUPs attempted installation or launch.
  - Every hour 2.3 million attempts were made by our customers to connect to risky IP addresses or those addresses attempted to connect to customers’ networks.

- We continue to receive valuable feedback from our readers through our **Threats Report user surveys**. If you would like to share your views about this **Threats Report**, please click here to complete a quick, five-minute survey.

  —Vincent Weafer, Senior Vice President, McAfee Labs
Contents

McAfee Labs Threats Report
August 2015

This report was researched and written by:
Brad Antoniewicz
Christiaan Beek
Dave Bull
Torry Campbell
Cedric Cochin
Carric Dooley
Douglas Frosst
Robert Gresham
Paula Greve
Steve Grobman
Dave Marcus
François Paget
Eric Peterson
Matthew Rosenquist
Raj Samani
Craig Schmugar
Mike Sentonas
Rick Simon
Bruce Snell
Dan Sommer
James Walter
Vincent Weafer

Executive Summary 5

Key Topics 6

Intel + McAfee: a five-year retrospective 7

Data exfiltration: an important step in the cyber thief’s journey 16

GPU malware: separating fact from fiction 26

Threats Statistics 30
Executive Summary

Intel + McAfee: a five-year retrospective

August marks the fifth anniversary of Intel’s announcement that it would buy McAfee. Since that time, much has changed in the cybersecurity world. For this retrospective, we brought together a dozen thought leaders from Intel and McAfee who have been here since before the acquisition to explain how the cybersecurity marketplace and our work together has evolved.

We discuss the evolution in our thinking about security in silicon, our views at the time about the “perfect storm” approaching in the cybersecurity world and how that storm has played out, the challenges we saw in emerging hard-to-detect attacks, and our 2010 expectations for new device types vs. the reality of the marketplace. We also discuss some of the things that surprised us, most notably the transformation of cybercrime into a full-fledged industry.

Data exfiltration: an important step in the cyber thief’s journey

The last 10 years have produced a monumental increase in the number of major data breaches and in the volume of records stolen, from TJ Maxx’s 2007 breach of 94 million records to this year’s theft of 80 million Anthem patient records. This Key Topic focuses on an important step in the data theft process: data exfiltration. It is the way in which a cyber thief copies or moves data from the owner’s network to one the thief controls. We examine attacker types, their motivations, and their likely targets; the methods and mechanisms they use to steal data; and policies businesses should embrace to better detect exfiltration.

GPU malware: separating fact from fiction

Malware attacks on graphics processing units (GPUs) have been around for years. In fact, a form of GPU malware has been active in the wild for at least four years—in the form of Bitcoin-mining Trojans that leverage GPU performance to increase the payout from each victim’s infected system.

Recently, a group published three proof-of-concept projects that together claim to use GPUs as an instrument of evasion by running code, and storing data, on GPUs—where no one is looking. In this Key Topic, we break down the projects’ claims into their components to establish what might be possible through the use of these software modules.
Key Topics

Intel + McAfee: a five-year retrospective

Data exfiltration: an important step in the cyber thief’s journey

GPU malware: separating fact from fiction
Intel + McAfee: a five-year retrospective

—McAfee Labs

On August 19, 2010, Intel announced that it would buy McAfee. At that time, McAfee and Intel were already working together on some projects and we realized we could enhance and accelerate our efforts if we made the arrangement permanent. Since then it has been fascinating to learn about each other’s capabilities. We worked through assumptions, resolved unrealistic expectations, and developed the confidence to build aggressive plans for the future.

It is now five years later: How goes the partnership? A dozen thought leaders from Intel and McAfee who have been here since before the acquisition collaborated to produce this look back at how the cybersecurity marketplace and our work together has evolved. In this retrospective, we examine what we expected to see across the threat landscape, what actually happened, and which developments have surprised us.

What Intel saw in McAfee

Intel thrives on the continued growth of the whole technology market. Throughout Intel’s history, we have taken steps to address anything that could slow the market or act as a barrier to continued growth. Processor speed, memory capacity, power consumption, peripheral connections, and chip size are barriers that we have overcome. Five years ago, we saw security as a looming impediment. If people started to lose confidence in their devices, connections, or services because of the loss of privacy, security, or even safety, it would slow the rest of the market. Unlike some of the hardware-specific issues that were straightforward for Intel to address organically, we concluded that we could not do this alone, and that we needed McAfee’s security expertise to help remove that growth barrier.

What McAfee saw in Intel

Five years ago, as attacks were improving their ability to evade defenses, the types of devices that needed protection were expanding quickly, and low-level threats such as rootkits were looming ever larger, McAfee realized that we needed to expand our security reach and coverage. Signature-based antimalware and perimeter defenses alone would not be able to guarantee a secure environment for much longer. We expected malware to become so sophisticated that it could break through the perimeter defense. We wanted to build security deeper down into the hardware and out into new platforms, to be able to stop attacks within the trusted network and repair the damage that they caused. To help us get there, we needed a much better understanding of hardware capabilities and behavior. We were already partnering with Intel on some processor-level security work, and realized that we would really benefit from their knowledge and capabilities.
Upon further review

When Intel announced the acquisition, we laid out our reasons to the tech, analyst, and investment communities. One of the top reasons was to bring software closer to the silicon to help strengthen security and more effectively counter increasingly sophisticated threats. These threats, combined with a significant increase in the number and type of devices, were creating the potential for a perfect storm of security breaches and vulnerabilities. We believed that these new threats would be more difficult to detect, requiring new approaches to cyber defense. We also expected the computing landscape to change dramatically, as billions of non-PC devices were connected to networks. Taken together, these elements would be a catalyst for further economic and technical development of the cyber threat landscape. So, what did we learn?

Security on silicon

Early on, an important focus of the acquisition was shifting security technology to silicon. This move was challenging given the rapid pace at which the cybercriminal community can copy and enhance the most sophisticated threats—often within days of their public discovery. Security hardware takes much longer to develop, market, and roll out than security software, and the security industry relies on the agility and adaptability of software to combat new and unanticipated threats. Customers need the ability to rapidly update their defenses to protect against attacks that were not envisioned yesterday, let alone in five-years, the typical hardware design cycle time.

Instead of working to embed antimalware in chips, we saw that it would be more logical to boost encryption performance with hardware assist, improve anti-tampering and kernel monitoring with low-level functions, and design security primitives into the next generation of chips that could then be leveraged by security and operating system software.

Low-level attacks in firmware and BIOS allow threats to remain persistent, making them attractive to cyberespionage and other long-view actors. As this type of malware moved deeper under the operating system in a quest to remain undetected and survive cleaning and rebooting, we released the open-source CHIPSEC framework, for analyzing hardware and firmware components and assessing low-level security risks, and Intel Kernel Guard, for ensuring runtime integrity; Intel Kernel Guard, for ensuring runtime integrity; and BIOS Guard, for authentication and protection. The combination of Intel's and McAfee's knowledge, expertise, and market footprint has provided us a unique vantage point to observe, adjust to, and anticipate changes on the threat landscape. Our objective remains to deliver security software for the new paradigms of mobile, Internet of Things, and cloud as security-enhanced chips penetrate the market.
Perfect storm approaching

We all thought that more users, more data, bigger networks, and many more types of devices and other targets like the cloud, combined with more attacks, clever new malware, and increasingly sophisticated actors were creating a perfect security storm. Most of these predictions came true. If anything, the adoption of cloud computing, Internet of Things devices, and mobile devices moved faster than we expected. Our 2010 prediction of 31 billion Internet-connected devices by 2020 now seems an underestimate.

Cyberattackers have certainly taken advantage of this massive increase in potential targets and expanding attack surface. At first, these threats were a concern mostly for governments, financial institutions, and security vendors, but they are now a major concern for enterprises and consumers, as they can significantly impact the value of businesses and can cause major headaches in our personal lives. Today, we face nation-state cyberwarfare that includes some highly visible, although actively denied, state-sponsored attacks as well as long-term espionage. Again, although we expected and predicted most of this development, the rapid evolution of malware, increase in attack volume, and large scale of nation-state attacks has been surprising.

Changing Attacker Profiles

Recreational
- Fame and notoriety
- Limited technical capabilities
- Limited technical resources
- Known exploits

Criminal
- Vandalism
- Limited technical capabilities
- Vandalism
- Targeted attacks

Hacktivist
- Statement
- Relentless, emotionally committed
- Vast networks
- Targeted attacks

Organized Crime
- Economic gain
- Significant technical resources and capabilities
- Established syndicates
- Adware, crimeware, IP theft

State Sponsored
- Cyberwar, state secrets, industrial espionage
- Highly sophisticated
- Nearly unlimited resources
- Advanced persistent threats

INCREASING RESOURCES AND SOPHISTICATION
The expansion of attacker types, their resources, and their sophistication.

Detecting the undetectable

In partial response to the perfect storm, we believed that we should quickly augment signature-based antimalware by adding technology to detect the undetectable, as malware evolved and adapted to avoid traditional security defenses. After all, unlike most attacks, the defenses are generally available for anyone to test and evaluate. Any attacker can put a security product in a lab and test it every way possible, looking for weakness to exploit or ways to evade.
In spite of this worry, the majority of security breaches during the past few years have been readily detectable. They were sophisticated in their planning, targeting, stalking, and execution; some were even highly technical or evasive. However, we have seen a change during the past two years, with a significant increase in the number of technically sophisticated attacks. Many of these have been designed purely to evade advanced defenses. They are infiltrating in pieces, hiding in seemingly inert code, and waiting for an unprotected moment to emerge. These threats also avoid the signature-based traps of their ancestors, employing encryption and dynamic code modification to change with each new deployment and hide incriminating data.

Retrospective: Five Years of Threats

<table>
<thead>
<tr>
<th>Trends and Notable Malware</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
<th>2014</th>
<th>2015</th>
</tr>
</thead>
<tbody>
<tr>
<td>MBR Infectors Increase</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drive-By Downloads</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Permanent Threats in the Browser</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exploit Kit Explosion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fileless Threats/Malware-Free Intrusion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Server Side Polymorphism/Hashbusters</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single-Use Malware</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Memory-Scraping Malware (Including POS Threats)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Macros and PowerShell Boost-Script Malware</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fake AV</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bitcoin/Digital Currencies Threats</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ransomware</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PoS Malware</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mac Threats on the Rise</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Malware Platform Diversification and Multiplatform Attacks</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mobile Threats (Malware, PHA, and PUA)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IoT Threats</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Key Vulnerabilities

<table>
<thead>
<tr>
<th>2011</th>
<th>2012</th>
<th>2013</th>
<th>2014</th>
<th>2015</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCrypt—CVE-2006-4339, CVE-2014-1568</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poodle—CVE-2013-3566, CVE-2014-8730</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FREAK—CVE-2015-0204, CVE-2015-1637</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logjam—CVE-2015-4000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Leading attacks against the core internet.
More and more we see long-running attacks that continue for many months, or those with a long-term view, willing to wait and watch before doing anything malicious. Most of these likely commit ongoing espionage instead of immediately selling exfiltrated data. Although in 2010 we expected to see long-running attacks, some of the tactics and techniques used in those attacks were unimagined five years ago. We documented one such attack in “The Equation Group: exploiting hard disk and solid state drive firmware,” a Key Topic in the McAfee Labs Threats Report, May 2015. Another long-running attack was documented in the report Operation Troy: Cyberespionage in South Korea.

Device type evolution

As we noted, one aspect of the perfect storm was the massive increase in the types and volume of devices, supported by a huge expansion in virtualization and public clouds.

Consumers have very quickly adopted cool technology. We did it with mobile phones, then smart phones, tablets, and now wearables. Rapid device adoption is connecting our homes and organizations to the Internet of Things—in healthcare, energy, logistics, retail, cities, transportation, automotive, and manufacturing. People have become so dependent on devices in their environments that they are willing to sacrifice security and privacy. We believed—and continue to believe—that whenever enough devices of a certain type create a lucrative market, attacks on those devices will begin. During the past five years, the results have been as we expected in some areas and surprising in others.

Mobile: Although mobile devices have seen very rapid growth in malware attacks, most of these are still at the exploratory stage or relatively minor in impact. The value of data recoverable from a smart phone is relatively low, and smart phones are not a prominent attack vector for the enterprise. The automatic backup capability of many smart phones and tablets make them straightforward to clean and recover if they are infected or ransomed, at least until criminals manage to attack the cloud-based backups. Application markets for smart phones and tablets are also much more restrictive, acting like whitelisting services to limit downloads of malicious apps. These restrictions are not 100% effective, but they do constrain the growth of mobile attacks. Although the volume of mobile devices has increased even faster than we expected, serious broad-based attacks on those devices has grown much more slowly than we thought.

Internet of things: IoT devices are just beginning to be exploited. The variety of devices, operating systems, and versions provides a near-term resistance to attack because few have a large enough installed base to attract cyber thieves. However, the sheer volume of devices has grown faster than we foresaw, and into industries that we did not expect, creating a massive attack surface—so it is only a matter of time until IoT device threats are widespread. Of course, attackers are not after the devices themselves, but the data or gateway capability that they enable. Attackers want the easiest way in and these devices often provide underdefended access to target-rich networks. We are seeing just the beginnings of attacks and breaches against them.

PCs and data center systems: Even with the incredible growth of non-PC devices, PCs and data center systems remain the most lucrative target for cybercriminals, as we expected. They have the best data, the most visible vulnerabilities, and the weakest patching regime.
Virtualization and the cloud: Device growth has been aided by a massive increase in virtualization and cloud computing. We expected fast growth in virtualization, especially in the data center, but were surprised by the rapid deployment and adoption of cloud computing and storage. Moving to virtualization made tremendous financial sense, and we optimized hardware for that purpose. Moving to the cloud also made operational and financial sense, but we thought that enterprises would be slower to adopt. Cloud adoption has changed the nature of some attacks, as devices are attacked not for the small amount of data that they store, but as a path to where the important data resides.

If an attacker gains access to a victim's cloud credentials, the attacker can eavesdrop on activities and transactions, manipulate data, return falsified information, and redirect clients to illegitimate sites. A victim's account or service instances can become a new base for the attacker, who can then leverage the power of the victim's reputation to launch subsequent attacks. We saw cloud vulnerabilities under attack even before the acquisition and they are still with us, as expected.

**Cyber threat evolution and economics**

We all expected significant growth in the volume and technical capabilities of cyberattacks. The conditions were just too tempting to ignore. Threats have evolved like a classic arms race, with criminals developing new attacks, the security industry responding with new defenses, and so on. The global Internet and the “dark web” were instrumental in fueling this race, making it easy for criminals to share techniques and learn from each other. As soon as an attack appeared in the wild, even if it was from the most technically sophisticated crime group, others could watch, decode, reuse, and even improve on it. Soon after a vulnerability was discovered, it was often sold to the bad guys for exploitation. Technology vendors began introducing bug bounties, and the buying of vulnerabilities by either vendors or criminals has become a much bigger business than we expected.

<table>
<thead>
<tr>
<th>Vulnerability Type</th>
<th>Price for Zero-Day Exploit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adobe Reader</td>
<td>$5,000–$30,000</td>
</tr>
<tr>
<td>Mac OS X</td>
<td>$20,000–$50,000</td>
</tr>
<tr>
<td>Android</td>
<td>$30,000–$60,000</td>
</tr>
<tr>
<td>Flash or Java Browser Plug-ins</td>
<td>$40,000–$100,000</td>
</tr>
<tr>
<td>Word</td>
<td>$50,000–$100,000</td>
</tr>
<tr>
<td>Windows</td>
<td>$60,000–$120,000</td>
</tr>
<tr>
<td>Firefox or Safari</td>
<td>$60,000–$150,000</td>
</tr>
<tr>
<td>Chrome or Internet Explorer</td>
<td>$80,000–$200,000</td>
</tr>
<tr>
<td>IOS</td>
<td>$100,000–$250,000</td>
</tr>
</tbody>
</table>

Quoted prices for zero-day exploits in 2013.
What we did not quite expect was the transformation of cybercrime into a full-fledged industry with suppliers, markets, service providers ("cybercrime as a service"), financing, trading systems, and a proliferation of business models. Of course, crime follows the path of least resistance to the money, and has to pay sufficiently well or people will stop doing it. Unfortunately, cybercrime has been paying very well. One security vendor detailed a 1,425% return on investment from a hypothetical, yet realistic, malware campaign. And in an Intel Security-commissioned study, the annual cost of cybercrime to the global economy was estimated to be around US$400 billion.

Although the Internet has been fundamental to cybercrime, attacks have been fueled by access to technologies that allow criminals to remain anonymous. Specifically, anonymizing networks—most notably TOR—and virtual currencies have become key to cybercriminals’ ability to remain hidden from law enforcement. Some of us noticed the early development of virtual currencies and immediately saw the potential for illegal transactions of many types. Bitcoin and anonymous brokerages have also reinvigorated the ransomware market, making it commercially viable and spurring unexpectedly high growth.

Five years ago, many high-profile thefts involved credit card details that were sold in bulk as quickly as possible to those who aimed to make fraudulent purchases. Credit card issuers have worked hard to quickly block the use of stolen cards, so the value of stolen cards now drops rapidly. Consequently, some attackers have started stealing other high-value data, such as personal health records, that does not lose value as fast. Learning from the business community, cybercriminals are also turning to data warehousing, combining and correlating multiple sets of stolen data into something much more valuable. Many recent high-profile data thefts, such as personal tax records or background checks, have not been immediately turned into cash, possibly indicating an increase in criminal maturity. This is something we did not anticipate.

Another indicator of cybercrime business maturity has been the drop in technical skills required to participate in the industry. Off-the-shelf toolkits for malware, affiliate programs for ransomware, fill-in-the-blank attack-creation programs, and other familiar business offerings have been showing up in the dark web to support faster, simpler and broader distribution of attacks. It now takes very little skill to be a cybercriminal. (For a look at packaged malware for sale, read “After the death of Blacole: the Angler exploit kit,” in the McAfee Labs Threats Report, February 2015.)
Generally, nation-states have different motives for their attacks, but they often leverage much of the same criminal infrastructure. Nation-states are not typically driven by direct monetization and can play a longer game, with vastly different resources. We think of espionage as something done quietly by a small number of people, and that has generally been true with cyberespionage. However, the scale of state-sponsored cyberespionage has exceeded our expectations and in just the last two years has become much more visible, even to the general public.

More surprises

Some surprises do not fit neatly into the sections we have discussed. Possibly the biggest is the continued lack of attention—by businesses and consumers alike—to updates, patches, password security, security alerts, default configurations, and other easy but critical ways to secure cyber and physical assets. This is not news to the security industry; we have banged this drum for decades, and yet these remain the most likely vectors for successful attacks.

Speaking of physical assets, we continue to wonder at the absence of a successful, catastrophic attack on critical infrastructure. Such attacks do not make sense for cybercriminals because there are no easy payoffs, but they almost certainly make sense for terrorists and perhaps for some nation-states. Although we have observed cyber reconnaissance on critical infrastructure, we suppose political or strategic considerations have kept this from happening—so far.

Speaking of infrastructure, the unexpected recent discovery and exploitation of core Internet vulnerabilities—in code that is decades old—has demonstrated how some foundational technologies are underfunded and understaffed. Acknowledging this risk has led to software sponsorship and increased collaboration among major organizations that depend on the Internet for everything they do.
We have been very pleased by the increasingly positive collaboration between the security industry, academia, law enforcement, and governments to take down cybercriminal operations.

Finally, we have been very pleased by the increasingly positive collaboration between the security industry, academia, law enforcement, and governments to take down cybercriminal operations. We see criminals able and willing to share their code and tips; we need to do the same on defense. “United we stand, divided we fall” may be a cliché, but it applies here.

**Conclusion**

We got some things right and some things wrong five years ago. Many of the predicted elements of the perfect storm materialized while others were unforeseen. Three forces have continued to challenge our cybersecurity landscape: the expanding attack surface, the industrialization of hacking, and the complexity and fragmentation of the IT security market. Cybercrime matured much more quickly than we expected from a hobby to an industry, trying different business models and operating under a mix of criminal, political, and military agendas.

Cybersecurity awareness is now at an all-time high, partly driven by the media, by new regulations requiring disclosure of breaches, and by increased knowledge and maturity. However, today the stakes are significantly higher, the landscape has transformed to the benefit of the attackers, and their skills and resources have increased as never before. Security battles continue to be a tremendous challenge, but the war is not over. A rise in awareness, along with the influx of more security professionals, technology innovations, and the recognition by governments of their role to protect citizens in cyberspace have all been beneficial. The merger of Intel and McAfee is part of the evolution that aims to provide security to protect people and technology for the future.
Data exfiltration: an important step in the cyber thief’s journey

—Brad Antoniewicz

During the last 10 years we’ve seen an unprecedented global adoption of technology. Internet usage has exploded from use by 15% to more than 40% of the world’s population and, with it, companies of all sizes have built Internet-connected networks to communicate with their customers and serve the data that fuels their businesses. This collection and digitalization of information combined with the vastness and reach of modern networks presents an enticing opportunity for thieves: stealing data.

The last 10 years has also shown a monumental increase in the number of major data breaches. In 2007, TJ Maxx experienced one of the first very large-scale breaches, in which credit and debit card information of up to 94 million customers was stolen. Just two years later, the payment processing giant Heartland Payment Systems was compromised, with data from an estimated 130 million customers exfiltrated. The years to follow would uncover even larger breaches with a wider a net of information targeted.

Beyond credit and debit card numbers, thieves have stolen nearly every other piece of information about individuals: names, dates of birth, addresses, phone numbers, social security numbers, health care information, account credentials, and even sexual preferences.

We now know that cyber thieves are not just profit-seeking groups or individuals. The motivations that fuel them have created distinct classifications of actors, each with a unique intent for stolen data. As shown in a recent foreign government job posting for a “US Intelligence Officer,” stolen personal data has a different purpose when the victims are employees of a government agency and the thieves are acting on behalf of nation-states.

The success of the Internet and evolution of cyber theft has also brought new life to cyberespionage, making digital intellectual property (IP) theft a realistic threat. Trade secrets have been stolen from organizations of all types—from Google, Microsoft, and Sony, to Boeing, Lockheed Martin, and DuPont—demonstrating that thieves find value in every place that it exists.

This Key Topic focuses on an important step in the data theft process: data exfiltration. In this step, cyber thieves copy or move data from the owner’s network to one they control. Data exfiltration is performed by actors with an intent to steal data—not to accidental data loss through misplaced or stolen equipment (in which the thief is more interested profiting from the hardware).

Threat actors

Threat sources, threat actors, and threat agents are terms to describe a group or individual who intends to gain unauthorized access to computer networks and systems. Across various publications in both the public and private sector that attempt to classify such threats, three major actors are consistent: nation-states, organized crime, and hacktivists.
**Motivation**

Motivation is one of the key characteristics that differentiate threat actors. Although not every actor needs to steal data during every campaign to accomplish an objective, many campaigns require it. When theft is required, a threat actor usually seeks the most appealing data types. However, data types pursued by an actor can change, so it is not uncommon to see an actor such as organized crime shift interests to IP theft, for example, to increase profits.

<table>
<thead>
<tr>
<th></th>
<th>Nation-State</th>
<th>Organized Crime</th>
<th>Hacktivists</th>
</tr>
</thead>
<tbody>
<tr>
<td>General motives</td>
<td>- Espionage</td>
<td>- Financial</td>
<td>- Reputational</td>
</tr>
<tr>
<td></td>
<td>- Influence</td>
<td></td>
<td>- Social</td>
</tr>
<tr>
<td>Example data types</td>
<td>- Source code</td>
<td>- Bank account information</td>
<td>- Emails</td>
</tr>
<tr>
<td></td>
<td>- Emails</td>
<td>- Credit card data</td>
<td>- Employee information</td>
</tr>
<tr>
<td></td>
<td>- Internal documents</td>
<td>- PII (including social security numbers, health data, etc.)</td>
<td>- Any sensitive internal data</td>
</tr>
<tr>
<td></td>
<td>- Military activity</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Government employee personally identifiable information (PII)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Volume of data pursued</td>
<td>Small-Large</td>
<td>Large</td>
<td>Small-Large</td>
</tr>
<tr>
<td>Sophistication of exfiltration techniques</td>
<td>High</td>
<td>Medium-Low</td>
<td>Medium-Low</td>
</tr>
<tr>
<td>Location on the network</td>
<td>Unknown/often scattered</td>
<td>Known</td>
<td>Both known and unknown/often scattered</td>
</tr>
</tbody>
</table>

Nation-state actors generally seek to gain a strategic advantage, which often translates into targeting intellectual property. Given the broad range of information that can benefit such an actor, the volume of data leaving the organization can be difficult to estimate—a simple plan or diagram of a new product can be relatively small, while the source code of a major application can be very large. Organizationally, this information is difficult to contain and is often scattered throughout multiple networks, requiring actors to spend significant time searching unless they have insider knowledge.

The financial goals of organized crime actors make their motives somewhat simpler to understand. These actors tend to focus on large troves of credit card, banking, or personally identifiable information. Most of these data types follow a standard, structured format that makes them easy to search for. The data is also commonly subject to regulations, which means it is in established locations on the network.
Hacktivists are perhaps the most difficult to stop because any internal data has the potential to impact an organization’s reputation. For this reason, all data types from credit cards to emails can be targeted by these actors, and data sizes can range from small to large.

**Physical access**

An actor’s ability to gain physical access to a system, even by proxy, yields a massive advantage. USB storage devices provide an easy way to exfiltrate large quantities of data while bypassing network security controls. A cyber thief can set up an attack by giving removable media to an unsuspecting employee, who then inadvertently launches the attack when he or she inserts the device.

**Environmental knowledge**

Social engineering, insider assistance, and open-source intelligence gathering can be used to gain environmental knowledge of networks and systems. These techniques reduce the time attackers spend on data discovery, obtaining system access, and exfiltration.

**Data exfiltration**

Copying data from a compromised network can be a complex task. It requires a strong understanding of an organization’s security configuration, the holes in its network segmentation, the placement and settings of security controls, and the privileges that yield access to systems along the way.

To better understand and categorize these complex techniques, we’ve divided the components into five major areas:

- **Data targets**: Systems containing data targeted by the attacker, such as file shares, repositories, point of sale (POS) systems, etc.
- **Staging infrastructure**: Company-owned systems used by the attacker to collect and transmit data from the company to dump servers.
- **Dump servers**: Attacker-accessible systems used to temporarily store data before it is fully under the attacker’s control.
- **Data transports**: Network protocols or data storage devices used to transport data from one location to another.
- **Data manipulation**: Techniques that alter or mask data, such as encryption, obfuscation, compression, and chunking.

**Components of Data Exfiltration**

The main building blocks for data exfiltration.
Data targets

Once an attacker compromises a system on the network, the door is open to explore other systems and uncover those that house attractive data. A complex network holds many types of data, making this a lengthy process for any actor without insider knowledge.

The main data targets include:

<table>
<thead>
<tr>
<th>Data Target</th>
<th>Types of Data</th>
<th>Actor Interest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Database systems</td>
<td>Varies: Protected health information, PII, credit cards, banking, and user accounts</td>
<td>Organized crime, hacktivists</td>
</tr>
<tr>
<td>Source code repositories</td>
<td>Source code, credentials, keys</td>
<td>Nation-states, hacktivists</td>
</tr>
<tr>
<td>Specialty systems</td>
<td>Varies</td>
<td>All, depending on endpoint type</td>
</tr>
<tr>
<td>File shares and similar systems</td>
<td>Source code, designs, communications, etc.</td>
<td>Nation-states, hacktivists</td>
</tr>
<tr>
<td>Email and communications</td>
<td>Designs, communications</td>
<td>Nation-states, hacktivists</td>
</tr>
</tbody>
</table>

Database systems

These systems store large amounts of structured data, which makes them an immediate target, especially for organized criminals. The systems perform many business functions:

- Authentication: Systems that contain information such as usernames and passwords associated with authenticating users.
- Patient tracking: Systems responsible for tracking the intake, management, and discharge of patients in the health care industry.
- Payment processing: Systems that accept, issue, and process financial transactions from customers or vendors.
- Customer processing/loyalty: Systems that contain customer data for tracking, marketing, or similar purposes.
- Human resource management/finance: Systems responsible for employee management and payment.
- Nonproduction/shadow IT: Testing systems and shadow IT systems that contain production and other company data can be just as valuable to an attacker and more vulnerable to attack.
Source code repositories

Internal source code repositories are sometimes left unprotected even though they contain highly valuable data such as application source code, API keys, database and authentication server credentials, and encryption keys.

Specialty systems

Attacks against large retailers and manufacturers show that actors target data that lies on specialty systems, or endpoint systems that have a well-defined purpose in a certain industry. These include:

- Point of sale systems: Perhaps the weakest link in payment processing is within the POS system, as credit card data is often unencrypted in memory after it is read from the card reader.
- Developer workstations: A wealth of intellectual property and environmental information can exist on developer workstations, making these systems high-value targets.
- Control systems: Set points and program logic provide valuable intelligence and can be modified with devastating impact in industrial attacks.

File repositories

To a thief, the sheer volume of data in large file repositories has advantages and disadvantages. On one hand, they may contain troves of information; while on the other hand, sifting through them manually may be a huge task because they contain unstructured information. Systems that fall under this category include:

- Network file shares: These systems contain group and user folders along with the documents, diagrams, and other company data contained in those folders.
- Content management systems: Microsoft SharePoint and others contain similar content as file shares but are usually more complicated for a cyber thief to unravel.
- Third-party cloud: Cloud-based file-sharing services such as Google Drive, Dropbox, and Box.com can also expose data but are often targeted by external attackers with insider knowledge rather than by attackers on an internal network.

Email and communications

User workstations, email servers, and instant-messaging systems such as Skype for Business are often a target because their caches store sensitive company data, operational information, and private communications.
The deeper and more segmented the target is on the network, the more complex it is for an actor to exfiltrate data. When needed, thieves build purpose-driven staging infrastructure, using hosts that act as intermediaries between network segments to an attacker-controlled dump server.

Staging infrastructure can be as complex or as simple as needed. In advanced exfiltration scenarios, we’ve seen systems with the following roles:

- **Endpoints**: Single or multiple data targets on the same or routable segment to the aggregator.
- **Aggregator**: Serves as a collection point for the data from the target endpoints and uploads the data to the exfiltrator. The aggregator may or may not have Internet access. In sophisticated campaigns, multiple aggregators may transfer data to several exfiltrators to obfuscate the outbound data path.
- **Exfiltrator**: Takes data from an aggregator and facilitates the transfer of it to the attacker’s dump server. This could be a simple transfer, or the exfiltrator may host the data for the attacker to retrieve.

This diagram represents a typical data exfiltration architecture, but there are others. For example, one publically dissected campaign established a geographically disperse ad-hoc network of exfiltrators and aggregators that operated on a rotational schedule when transferring data between systems and to the dump servers. In another instance, a company-owned Internet-accessible content distribution server was used as the exfiltrator by embedding data within the video stream of outward-facing content.
Dump servers

A dump server is the first point at which stolen data resides outside of the company’s control. However, it is not necessarily in the attacker’s control, either. It is simply in a place easily accessed by the attacker. Dump servers can be:

- Compromised systems: Systems that have been compromised by the attacker during a separate campaign. These systems can be everything from personal WordPress blogs to servers belonging to companies with weak security controls.
- Hosted systems in specific countries: Countries with strong privacy laws are attractive to attackers because they may be able to host systems within their borders and remain undisturbed while being afforded a certain level of protection.
- Temporarily hosted systems: Short-lived systems hosted in the cloud through providers such as AWS, Digital Ocean, or Azure.
- Cloud file-sharing services: General-access online file-sharing sites such as Dropbox, Box.com, or Paste Bin.
- Cloud-hosted services: Various other Internet-based services such as Twitter and Facebook that allow users to post data.

Compromised hosts, hosted systems in specific countries, and temporarily hosted systems work well as dump servers because they provide the greatest control over the data, allowing the attackers to fully customize the transports used from the exfiltrator. Cloud file-sharing and hosting services make it difficult for defenders to simply block the destination hosts due to their widespread locations. However, these services usually have easily accessible methods to report foul play and can quickly disable a malicious account.

The downside to using dedicated hosts as dump servers is that once they’re discovered, these systems can be blocked or shut down. One method to work around this is by using domain generation algorithms. These algorithms are built into malware running within the target company and generate a list of predictable domain names that can be queried to identify active control or dump servers.
### Data transports

Data transports are the protocols and methods used to copy data from one location or system to another, such as between an exfiltrator and the dump server or the endpoint and the aggregator. The following table summarizes many of the common transports used today and their networks:

<table>
<thead>
<tr>
<th>Transport</th>
<th>Description</th>
<th>Internal</th>
<th>External</th>
</tr>
</thead>
<tbody>
<tr>
<td>HTTP/HTTPS</td>
<td>HTTP's prevalence in network communications makes it an ideal protocol for hiding exfiltrated data with other traffic. It has been used as a general exfiltration transport by embedding commands in HTTP headers and within GET/POST/PUT methods.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FTP</td>
<td>FTP is commonly available on corporate servers and is easy to interact with using native system commands, making it a no-fuss transport.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>USB</td>
<td>USB storage devices are frequently used for exfiltration when traversing air-gapped networks. We have seen malware that looks for a USB storage device with a specific marker, then copies to-be-exfiltrated data to a hidden sector on the device. When the device is placed into another infected system with network access, the exfiltration begins. USB storage devices can also be used by insiders to easily copy large amounts of data and physically remove it from the organization.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNS</td>
<td>Specific DNS records such as TXT or even A and CNAME records can, to some extent, store data within them. With the control of a domain and a name server, an attacker can transmit small amounts of data by making specific lookups on the exfiltrating system.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOR</td>
<td>The use of the TOR network is becoming more popular. It allows attackers to post exfiltrated data to servers that are difficult to trace. However, TOR traffic on corporate networks is rarely legitimate and thus can be easily detected and stopped.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMTP/Email</td>
<td>Both company- and non-company-owned SMTP servers can be used to send data out of the organization as attachments or in the body of email messages.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMB</td>
<td>SMB is an extremely common protocol in Windows environments and may already be enabled on systems.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RDP</td>
<td>RDP supports various activities such as copy/paste and file sharing, and in some cases systems allowing RDP may be exposed to the Internet.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Custom transports</td>
<td>Custom transports are sometimes used in control server communications and sophisticated malware. A robust transport requires a great amount of effort and its uniqueness makes the protocol easy to identify on the network—tilting the scale toward an established transport.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Transports that offer encrypted alternatives (such as HTTPS) can increase detection difficulty for organizations. We’ve seen an increased but generally limited use of encryption at the transport level. The lack of encryption implies easy detection, but it also poses additional risk to the data because, in some cases, it is transmitted over the Internet.

Many transports require either valid credentials or some form of open/anonymous access to be enabled on the server. Thus if attackers want to automate exfiltration, they’ll need to leave a username/password on the compromised host or risk having the system remotely accessed by an unauthorized individual.

**Data manipulation**

Manipulating data before transfer can aid in avoiding detection, decrease transfer time, and increase analysis time. Although data is most frequently manipulated when it is transferred over the Internet, it is still very common to see manipulation in use on the internal network. Once the original data is manipulated, it is sent via the transport to its destination. Data manipulation techniques commonly in use:

<table>
<thead>
<tr>
<th>Technique</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compression</td>
<td>Compression using the standard ZIP format not only provides a level of obfuscation but also speeds file transfers.</td>
</tr>
<tr>
<td>Chunking</td>
<td>Splitting data into small pieces before sending helps the transfer blend in with regular network activity.</td>
</tr>
<tr>
<td>Encoding/Obfuscation</td>
<td>The most common type of data manipulation is a basic encoding or obfuscation algorithm. Using simple techniques such as performing an XOR operation with a static key, Base64 encoding, or simply converting each character to hex, the data can be manipulated just enough to avoid detection.</td>
</tr>
<tr>
<td>Encryption</td>
<td>It is surprising that encryption is not always used during exfiltration. Perhaps it is due to slower performance or just a lack of requirement. When used, it is common to see RC4 or AES encryption.</td>
</tr>
</tbody>
</table>

**Conclusion**

Digital information has become a prime target for thieves. Data being stolen ranges from large employee databases to volatile memory on POS systems. As soon as defenders build a new layer of security into their networks, attackers find ways to turn trusted systems against the organization by making them accomplices.
The first step in seizing control is to understand the actors, their motivations, and their techniques. Although data exfiltration may be a small component in an overall campaign, it is also one of the most important for the attacker to execute and for the defender to block. Establishing strong policies and procedures in addition to building a defense around critical assets and data targets allows organizations to prioritize their efforts so that the most important systems are treated with the greatest care.

<table>
<thead>
<tr>
<th>Key Topics</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Identify data sources</td>
<td>Conduct a risk assessment that involves interviewing key stakeholders to determine what sensitive data exists on your network and where it is housed.</td>
</tr>
<tr>
<td></td>
<td>- Asset inventory control</td>
</tr>
<tr>
<td></td>
<td>- Systems and network architecture</td>
</tr>
<tr>
<td></td>
<td>Consider using data-discovery software to identify sensitive information and its location.</td>
</tr>
<tr>
<td>Determine data flows</td>
<td>Identify the flow of sensitive data across and out of your network.</td>
</tr>
<tr>
<td></td>
<td>- Systems and network architecture</td>
</tr>
<tr>
<td></td>
<td>Consider using real-time data-flow monitoring software to understand data movements.</td>
</tr>
<tr>
<td>Identify regulatory and privacy requirements</td>
<td>Understand the regulatory requirements that apply to your organization and the security controls required.</td>
</tr>
<tr>
<td>Classify data</td>
<td>Establish a policy to classify data by sensitivity, type, and criticality.</td>
</tr>
<tr>
<td></td>
<td>- Data protection policy</td>
</tr>
<tr>
<td></td>
<td>- Data classification policy</td>
</tr>
<tr>
<td>Assign data owners</td>
<td>Develop a program that details data owners and their responsibilities.</td>
</tr>
<tr>
<td></td>
<td>- Data owners</td>
</tr>
<tr>
<td></td>
<td>- Data asset inventory and maintenance</td>
</tr>
<tr>
<td>Ensure data is protected</td>
<td>Establish a policy to define security requirements for data at rest and in motion.</td>
</tr>
<tr>
<td></td>
<td>- Data encryption policy</td>
</tr>
<tr>
<td></td>
<td>Implement data loss prevention software to prevent unauthorized data exfiltration.</td>
</tr>
<tr>
<td>Review access to data</td>
<td>Define a process in which access to data is formally tracked and authorized.</td>
</tr>
<tr>
<td></td>
<td>- Data authorization</td>
</tr>
<tr>
<td></td>
<td>- Change management</td>
</tr>
<tr>
<td>Regularly review the program</td>
<td>Define a data risk management process to annually review policies and procedures.</td>
</tr>
<tr>
<td></td>
<td>- Risk management</td>
</tr>
<tr>
<td></td>
<td>Consider using risk-management software to assess risks and manage compliance.</td>
</tr>
</tbody>
</table>
GPU malware: separating fact from fiction
—Craig Schmugar

Almost all of today’s malware is designed to run from main system memory and execute on the CPU. This has been the case for decades, and as such the vast majority of host-based defense and forensic tools are built around this assumption. Any deviation from this norm deserves a raised eyebrow—and many information security professionals had wrinkled foreheads earlier this year.

Malware attacks on graphics processing units (GPUs) have been around for a number of years, with attention flaring up now and then. In fact, such malware has been active in the wild for at least four years—in the form of Bitcoin-mining Trojans that leverage the awesome GPU throughput to increase the payout from each victim’s infected system.

A renewed interest in this subject arose recently, after the beginnings of proof-of-concept code appeared on GitHub, the world’s largest code host.

“Team JellyFish” published three such projects at the time. This new code claims to push the boundaries beyond merely leveraging the efficiencies of the GPU for raw processing power by using the architecture as an instrument of evasion by running code, and storing data, where no one is looking. The following text appears in their respective GitHub project pages:

**Demon**, a GPU keylogger described as containing the following features:
- CPU kernel module bootstrap to locate keyboard buffer via DMA in usb struct.
- Keyboard buffer stored in userland file.
- Kernel module deletes itself.
- **OpenCL** stores keyboard buffer inside GPU and deletes file.

**JellyFish**, described as a Linux-based user-mode GPU rootkit, with the GPU providing a number of advantages:
- Absence of GPU analysis tools online.
- Ability to access CPU host memory via Direct Memory Access.
- GPU performance benefits over CPU for mathematical calculations.
- Persistence across warm reboots.

**WIN_JELLY**, described as a Windows GPU remote access tool with persistent executable code storage in GPU that later can be mapped to userspace after rebooting.

GitHub project notes for proofs of concept regarding GPU attacks.
Subsequently, numerous articles were published reiterating the claims made by the authors. Out of context, it’s easy to twist these points together into a picture of an undetectable superbug, running autonomously and hidden from current defenses, but the truth is not as it first appears.

The claims can be distilled into four main points:

- CPU host memory access from the GPU.
- Subsequent deletion of CPU host files.
- Persistence across warm reboots.
- Absence of GPU analysis tools.

To address these claims, McAfee Labs enlisted members of Intel’s Visual and Parallel Computing Group (VPG) for their expertise to assist in verifying these assertions. The responses in the next section cover the area of expertise that Intel provides, namely integrated graphics and OpenCL, but most details apply to discrete graphics cards and nVidia’s CUDA platform.

**CPU host memory access from the GPU**

By design, programs accessing the GPU require a parent process running on the CPU. That parent process can operate in a similar fashion to other threats, reading and writing memory in ways that are often monitored or restricted by security products, but one benefit of using the GPU for this task is to conceal nefarious activity and circumvent such protections.

However, in order to deliver payloads often associated with malware using nontraditional methods and leveraging the GPU, physical memory must be mapped to the GPU. Furthermore, unprivileged code access is limited to memory pages mapped to a process’ virtual address space, making ring 0 access a requirement to map critical OS memory onto the GPU for read/write access, which adds to the malware’s footprint on the host. This dependency is subject to existing kernel protections. Which leads to the next point.

**Subsequent deletion of CPU host files**

Once a program is running on the GPU, the files required to install the application can be deleted. This includes the kernel driver responsible for memory mapping, as well as the parent usermode process. However, at this point the code running on the GPU will be orphaned and in the case of Microsoft Windows will initiate a Timeout Detection and Recovery process that resets the graphics card. On Windows, the default timeout is two seconds, although this is configurable. Microsoft states that TDR settings manipulation should be limited only to testing and debugging purposes. Thus any modification of these values can be considered a suspicious behavior: one that security products may choose to alert on, or even block. Furthermore, long-running GPU workloads will result in visual evidence of a problem because the GUI will become unresponsive. This is true for other operating systems as well.

In order to overcome these obstacles, usermode code (albeit minimal) must remain running, which provides something for endpoint protection to identify. Exceptions to this visual indication of a compromise scenario are multi-GPU and/or headless systems in which a lack of operating system access to the GPU may go unnoticed. Nonetheless, altered TDR values in Windows will remain as evidence of a potential problem.
**Persistence across warm reboots**

At face value, perhaps the most dubious claim of all is the notion that GPU-resident malware can remain persistent across an OS reboot. Recall the WIN_JELLY claim: “Persistent executable code storage in GPU that later can be mapped back to userspace after reboot.” Without careful parsing of the phrase, this sounds pretty bad from a security perspective, but upon further inspection this claim is not what it first appears. “Persistent” does not describe executing code, but rather data storage. Remember from the previous point: By default, a host process is required to keep the GPU program running. The idea of persistence claimed here is that a host application is running at system startup, retrieving data from GPU memory, and mapping it back to userspace, which is not nearly as daunting because malicious usermode code must also persist outside of the GPU.

**Absence of GPU analysis tools**

Although a number of tools provide GPU performance monitoring and debugging, there are very few for forensics and malware analysis. Historically such tools have been driven by necessity or by those looking to streamline a more mundane process. In this case, there is a need for threat analysis tools to more easily understand what a threat is doing on the GPU. But endpoint security products need not depend on such tools for threat classification and identification because using this attack vector provides other indicators of attack.

**Putting it all together**

GPU threats are a real concern. But this type of attack has not reached perfect storm status. On one hand reverse engineering and forensic analysis of such threats is much more complex and challenging than their pure CPU counterparts, and this may result in an infection going unnoticed for a longer period. By moving part of malicious code off of the CPU and host memory, the detection surface is reduced—making it more difficult for host-based defenses to detect attacks. On the other hand, the detection surface has not been completely eliminated. At a minimum, trace elements of malicious activity remain, allowing endpoint security products to detect and remediate the threat.

There are some parallels between GPU malware and Windows kernel rootkits about 10 years ago. Privileged code execution was a kernel rootkit requirement; once running, malicious code could conceal its presence; and rootkit analysis tools were more limited. It is now harder than ever for an attacker to gain and execute privileged code (which remains a requirement for doomsday GPU malware). Security products added rootkit-specific defenses and Microsoft released a number of kernel protections, including Patch Guard, driver signing enforcement, Early Launch Anti-Malware (ELAM), Secure Boot, and other defensive features. Many of these safeguards play a role in the defense against Windows kernel-assisted GPU attacks.

A recent advancement affecting the viability of GPU malware on certain systems is Microsoft’s Device Guard feature in Windows 10, which leverages the input/output memory management unit (in Intel’s case, the Virtualization Technology for Directed I/O, or Intel VT-d) in hardware to allow administrators to lock devices to allow only Microsoft signed and trusted applications to run. Although this feature is available only under certain circumstances, it can provide extra security for those securing critical information.
This response is not intended to summarily dismiss all claims made around GPU attacks, but rather to provide some context on the current threat and defenses. Undoubtedly there will be advancements made in this area by both attackers and defenders. If nothing else, the recent attention to this area has resulted in the security community’s reevaluating its current posture and exploring ways to improve upon it.

**Safe practices to protect against this type of attack**

McAfee Labs recommends several ways to protect systems against GPU attacks:

- Enable automatic operating system updates, or download OS updates regularly, to keep operating systems patched against known vulnerabilities.
- Install patches from other software manufacturers as soon as they are distributed.
- Place security software on all endpoints and keep AV signatures up to date.
- Consider application whitelisting to stop the execution of unauthorized applications.
- Avoid running applications in administrator mode whenever possible.
Threats Statistics

Mobile Threats
Malware
Web Threats

Share feedback
Mobile Threats

The total number of mobile malware samples grew 17% in Q2.
Mobile malware infection rates declined about 1% per region this quarter, with the exception of North America, which dropped almost 4%, and Africa, which was unchanged.
The McAfee Labs malware zoo grew 12% in the most recent quarter. It now contains more than 433 million samples.
Ransomware continues to grow very rapidly—with the number of new ransomware samples rising 58% in Q2. As first discussed in the McAfee Labs Threats Report: May 2015, we attribute the increase to fast-growing new families such as CTB-Locker, CryptoWall, and others. The total number of ransomware samples grew 127% in the past year.
Web Threats

New Suspect URLs

New Phishing URLs
New spam URLs and their domains leaped by 380% in Q2. Most of this increase is due to hundreds of thousands of autogenerated or sequential domains dedicated to spam campaigns we discovered after we improved our collection of Realtime Blackhole Lists.
The trend of decreasing botnet-generated spam volume continued through Q2, as the Kelihos botnet remained inactive. Slenfbot again claims the top rank, followed closely by Gamut, with Cutwail rounding out the top three. Slenfbot spam during the quarter was primarily on a theme of “male enhancement,” with a top subject line of “Tips to nights of happiness.”
About Intel Security

McAfee is now part of Intel Security. With its Security Connected strategy, innovative approach to hardware-enhanced security, and unique Global Threat Intelligence, Intel Security is intensely focused on developing proactive, proven security solutions and services that protect systems, networks, and mobile devices for business and personal use around the world. Intel Security combines the experience and expertise of McAfee with the innovation and proven performance of Intel to make security an essential ingredient in every architecture and on every computing platform. Intel Security’s mission is to give everyone the confidence to live and work safely and securely in the digital world.

www.intelsecurity.com

Feedback. To help guide our future work, we’re interested in your feedback. If you would like to share your views, please click here to complete a quick, five-minute Threats Report survey.

Follow McAfee Labs

Follow us on Twitter

Follow us on Facebook


The information in this document is provided only for educational purposes and for the convenience of McAfee customers. The information contained herein is subject to change without notice, and is provided “as is,” without guarantee or warranty as to the accuracy or applicability of the information to any specific situation or circumstance.

Intel and the Intel and McAfee logos are trademarks of Intel Corporation or McAfee, Inc. in the US and/or other countries. Other marks and brands may be claimed as the property of others. Copyright © 2015 McAfee, Inc. 62058rpt_qtr-q2_0815_PAIR