CYLANCE

Prevention vs. Detect & Respond

Introduction

The Risks of Letting Malware Execute

On the Internet, there are only victims and potential victims. Everyone has exposure, from individuals to large enterprises. Every minute sees more connected devices added to the attack surface. In the race to broaden and deepen defenses, security teams are faced with the additional challenge of increasing complexity. More products, more events and more monitoring are making it ever harder to find relevant and true indicators of compromise, pushing the security situation closer to bedlam. This spiraling complexity diminishes a security team's awareness and responsiveness, ultimately driving up the true cost of operational security.

Defenders didn't inflict this on themselves. The proliferation of vendors and products, driven by overwhelming need and a growing security economy, has created a wide diversity of approaches to solving various security problems. In seeking to find a novel, often information-centric angle, however, many of these solutions have inadvertently created new challenges and failed to improve security.

Some type of malware is involved in part of almost every security incident (considering that insiders are involved in 10.6% of incidents[1]). This means having an effective endpoint security strategy is one of the biggest challenges facing security teams. In this paper we examine current attack detection strategies that are rapidly gaining popularity. We will discuss the pitfalls of pure post-execution strategies, and review the underlying core of the malware detection problem that vexes most enterprises.

We contend true prevention lowers security costs and complexities, and is the best cure for malware.

An Evolution in Strategy

Attack detection strategies can be divided into three broad categories, based on when they are designed to intervene: Before the malware executes (pre-execution), as it executes (execution) or after it has already run (post-execution).

Pre-Execution	Execution	Post-Execution
▪ Hash-based - (semi) automated	▪ Memory-based (signature)	▪ IDS
▪ Human generics and heuristics (manual signature)	▪ IDS	▪ Memory analysis
▪ File reputation-based - cloud/local (late 2000s)	▪ Exploit detection	▪ Behavioral
▪ Policy (whitelist only) (late 2000s)		▪ SIEM (analytics)
▪ Machine learning-based models (2012)		▪ "Detect and respond" (ETDR)

Early malware detection technologies used generic and heuristic signatures to detect a malware file as soon as it was written to disk. Antivirus (AV) software vendors used to be able to get away with manually writing signatures because malware families did not change that often.

For a brief time, the cost of managing AV security solutions in the enterprise was largely fixed. Once installed, an AV solution would be on autopilot and detect and remediate known malware. AV solutions, with their heuristics and emulator-based signatures, had the upper hand in the battle. But not anymore.

The mid-2000's brought a wave of polymorphic, rapidly changing malware and rootkits, with many tricks for bypassing traditional AV's runtime protections. To keep detection rates up, security vendors created automated malware processing solutions with hashing technologies at their helm. The fatal shortcoming of hashing is that every malware specimen, no matter how slight its differences, looks completely new and different when hashed. The next wave of polymorphic and hash-busting malware preyed upon this limitation in automated hashing technologies. The cost of missed detections rose rapidly for enterprises, so vendors began to offer repair and remediation services. To counter the detection challenges, the industry added newer memory analysis, reputation and behavioral technologies. These security layers kept rising over time.

These newer malware detection strategies were created to address AV solutions' failure to prevent attacks. Unfortunately, from a collective security industry standpoint, this took the focus away from investing in preventive solutions and other technologies that kept security management costs lower.

This period saw the introduction of pure policy-based solutions where only known files could execute (whitelisting). That approach was severely limited to specific use cases that had very restrictive change control, such as point of sale systems.

Today, many "state-of-the-art" security solutions are responding to the increase in cyberattacks via post-execution malware analysis that includes continuous endpoint monitoring and rapid reactive response to attacks. Although it may seem like the need of the hour, we need to understand the implications of this new thrust and examine how we can best improve security.

Preparing for the Worst?

Malware execution on an endpoint has inherent risks. A compromise reaches the post-execution stage only after the failure of every preventive solution. The hope of this type of malware analysis is that its actions will give away the malicious behavior or that the affected enterprise will be able to use this new monitoring layer to recover from an attack. Post-execution monitoring logs and analyzes application behavior, and in many cases also analyzes and stores most of the network traffic. This is all done to detect and eventually recover from the inevitable worst-case compromise scenario.

> *The key word here is "worst." We believe the security industry can do better than allow the worst-case scenario to unfold before taking action.*

Coming back to the dilemma of malware execution and analysis on an endpoint, we must answer the question: What should these solutions monitor? What to monitor is an important issue. If a solution is logging most of the network, operating system and application behavior data in anticipation of the worst, it is collecting massive amounts of data.

Post-execution solutions are noisy. With limited autonomy, they simply can't risk missing something important, so they seek to collect and analyze an avalanche of data. This includes disk writes (and parent proc), execution events, some subset of reg, RPC communications, user activities (including URLs visited, cookies written), DNS requests and network trace data like pcap or NetFlow for every operation.

The amount of data quickly adds up. Let's say these systems collect 1 megabyte per hour per host (or 1,000 1-kilobyte records after compression). If you multiply that by 24 hours for 1,000 hosts, you have 24 million events, or 24 gigabytes of data a day. After 90 days you amass 2.1 billion records, or 2.1 terabytes of data. Imagine how much data an enterprise with 50,000 to 100,000 hosts might collect.

Even with this heavy-handed approach to data siphoning, defenders may never be able to find the needle in the haystack. This burdensome approach is extremely complex and wastes power, memory, disk space and network resources. Let's look at some of the hidden costs for an enterprise if a security solution purely uses the "detect and respond" approach after malware has executed.

Management Cost

Gathering and maintaining the volume of information needed to operate a "detect and respond" solution is a commitment that grows with time, as does the cost of extracting value from the information. Enterprises should be aware of the following hidden costs and concerns:

Security event analysis
More security events lead to more analysis and increased costs.

Endpoint system performance
Continuous endpoint monitoring leads to performance bottlenecks, while unnecessary data collection further strains the endpoint.

Cloud lookups/network bandwidth
Although the security vendor is paying for cloud storage, it is the enterprise that is paying for network data usage.

On-premises analysis
Hosting and managing a big data solution on premise to deal with the volume of data adds complexity, as well as hard and soft costs.

Privacy concerns
Solutions that collect and store most of the system events for detection and response may end up collecting more information than is necessary or desired. Access controls, locality, retention periods and encryption policies on the collected data may vary by vendor.

Some security solutions rely on open-source data to get information about suspicious files. These sources often do not turn up useful data since pre-execution malware detection has become a lost art and security vendors are not investing sufficiently in improving file detection capabilities.

For example, the Dyre family of malware was detected on two consecutive days in June. A post-execution system querying any vendor on June 4 would not have recognized the sample as malware, based on the industry's collective knowledge. There are many examples of malware that is not initially identified as malicious and fails to get reported as bad for weeks, months or even years. Such delays highlight the need for systems to fill the gap in malware identification without relying on reactive file-detection scanners.

What You See is What You Get

Permitting malware execution creates major technical challenges by expanding the playing field for malware instead of limiting its options. Here are some examples of weaknesses in newer technologies that are based on "detect and respond."

Good behavior / bad behavior
During malware post-execution analysis, endpoint security solutions need to monitor the suspect in its natural environment to detect, log and block events in order to recover from attacks. However, even under monitoring, it is very hard to predict when a malware specimen such as Rombertik may reveal its ugly side[2,3,4]. It may be days before it executes its malicious code or it could be dependent on some user action (like scrolling a document to the second page) to trigger the malware. Lots of new research has tried to solve some of these problems, only to be circumvented yet again[5,6]. An alternative would be to watch all applications all the time in the anticipation of a security event, which leads us back to the security management cost discussion.

How late is too late?
Can a monitoring technology detect the first "bad event"? Is installing a driver a malicious event by itself? Most often the answer is no, but the moment a malicious kernel driver runs, it's probably too late to save the system. These are just some of the disadvantages of post-execution monitoring. More often than not, a series of behaviors constitutes a malicious behavior. However, it may be too late to block the malware if that determination is not made in time and,

more importantly, every time. This yet again presents the old "signature detection," cat-and-mouse game, where defenders try to detect as early as possible and attackers try to evade by mixing good and bad events and sometimes gray events.

Examples of irreversible damage a piece of malware can cause if not blocked in time, every time, include:

Parasitic infection
By letting a parasitic infector run and infect critical files, the cost to recover and restore the system increases drastically. The file may become permanently damaged, requiring the system to be rebuilt or restored from backup.

Data destruction
We have recently seen two major attacks that could have not been prevented by defenders solely relying on a post-execution "detect and respond" approach. The Saudi Aramco and Sony Pictures attacks both used a signed commercial kernel driver to wipe and destroy the targeted machines[7]. Another simple example is running ransomware like CryptoWall/CryptoLocker, which encrypts every file on the system and then demands extortion money. Post-execution solutions detect these attacks too late because machines cannot be recovered later. This was demonstrated in the CylancePROTECT® vs. Ransomware video[8].

Security solution detection and hostile evasion
Recent research on the Rombertik malware shows a prime example of the havoc that can be caused when malware executes. Rombertik tries to bypass security checks, then attempts to destroy environments and machines that seek to analyze it.[9]

Data exfiltration
Over the years we've seen many types of point of sale malware such as Framework POS, which used a DNS mechanism to exfiltrate credit card data[10], and the notorious BlackPOS malware, which hit Target in 2013. Post-execution detection would not have reduced the risk with these types of malware. Once the data has left the system, the damage is almost impossible to undo. Once the malicious code is executing, it has many ways to exfiltrate data, and the security solution once again has to implement a myriad of defenses over time.

Attacks on security solutions
Allowing malware to execute has given some malicious programs an opportunity to directly attack security solutions and endpoint agents. For example, last year the Vawtrak malware attempted to disable security software using software restriction policies.[11]

Back to the Future

What have security vendors and practitioners learned from almost a decade of increased security costs and having to build newer solutions? If there was one thing we could have done differently, what would that be?

It's clear relying on solutions that only seek to detect malware AFTER it has executed is not viable. When the security industry started pulling away from pre-execution containment, technologies became reactionary and too dependent on manual sample analysis and signature creation. Security vendors hoped post-execution analysis and solutions would give them the necessary respite from the malware problem, only to find it made the system more complex, expensive, and more prone to attacks and bypasses.

Cylance® has overcome these challenges by building the first and only artificial intelligence and machine learning based pre-execution detection environment. The main challenge in the pre-execution environment is to analyze the program and determine if a file is good or bad based purely on the information in the file itself, and then do that at a sustainable, massive scale. The ability to do this across a huge number of samples is important because modern malware creation is automated. Today it requires very little effort for attackers to mutate a piece of malware. Manual generic signatures (emulation- or heuristic-based) were good for protection when malware creation was manual, but not anymore.

To be able to go back to the basics and stop malware before it ever gets a chance to execute, Cylance uses machine learning to generate models that can predict if a program is malicious. This approach for file detection has proven

extremely effective at stopping malware. Cylance has proven it is possible to identify malware with astounding accuracy, without ever having seen it before. Going back to our Dyre example, Cylance succeeded in detecting the sample using a machine-learning model that was released in August 2014 – 10 months before the variant was released. Cylance was able to detect Dyre, pre-execution, long before it was identified post-execution by traditional AV solutions.

Pre-execution malware detection is not a silver bullet that no malware can ever bypass. No single solution can be infallible. As security practitioners know, security is about minimizing risks, not implementing absolutes.

A pre-execution strategy is the first step in building an effective security portfolio. Identifying malicious applications before they get a chance to execute helps limit security management costs and system performance overhead. It can also reduce the challenges posed to post-execution analysis environments, greatly reducing both the number of samples that need post-execution monitoring and the odds that a malicious sample will ever make it past that final layer of defense. That can help reduce the number of security layers needed to successfully thwart hackers.

Conclusion

The security industry has come a long way in defending against malicious attacks. However, in the rush to create quick, easy, out-of-the-box solutions, the industry has gotten stuck in a vicious signature circle: Vendors develop new solutions, malware authors find techniques for defeating them, then the security industry designs another solution, leading to even more bypasses and attacks.

Many security solutions are now trying to tackle the malware detection problem with post-execution, "detect and respond" approaches. Malware has found multiple ways to attack and bypass these solutions.

Not only are they ineffective, but "detect and respond" solutions are generating so much data there is a growing, artificial need for big data security analytics that the industry may never be able to manage.

New security layers need to know what and when to monitor. If a solution is always blindly logging everything, it is merely preparing for an eventual recovery from a breach, not actually trying to stop it. This is a very risky stance because if defenders abandon all hope of ever preventing attacks, then they will always be too late and end up focusing only on containing enterprise losses, which is simply not an acceptable solution.

References

[1] http://www.verizonenterprise.com/DBIR/2015/

[2] http://blogs.cisco.com/security/talos/rombertik

[3] http://joe4security.blogspot.com/2012/10/defeating-sleeping-malware.html

[4] http://www.networkworld.com/article/2163341/byod/-sleeper--malware-like-nap-trojan-nothing-new.html

[5] http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.361.9423

[6] https://www.lastline.com/papers/acm_ccs11_hasten.pdf

[7] http://arstechnica.com/security/2014/12/sony-pictures-malware-tied-to-seoul-shamoon-cyber-attacks/

[8] https://www.youtube.com/watch?v=RkbB8pV09E8

[9] http://blogs.cisco.com/security/talos/rombertik

[10] https://blog.gdatasoftware.com/blog/article/new-frameworkpos-variant-exfiltrates-data-via-dns-requests

[11] http://blog.trendmicro.com/trendlabs-security-intelligence/windows-security-feature-abused-blocks-security-software/

About Cylance:

Cylance is the first company to apply artificial intelligence, algorithmic science and machine learning to cybersecurity and improve the way companies, governments and end-users proactively solve the world's most difficult security problems. Using a breakthrough predictive analysis process, Cylance quickly and accurately identifies what is safe and what is a threat, not just what is in a blacklist or whitelist. By coupling sophisticated machine learning and artificial intelligence with a unique understanding of a hacker's mentality, Cylance provides the technology and services to be truly predictive and preventive against advanced threats. For more information, visit cylance.com

+1 (877) 973-3336
sales@cylance.com
www.cylance.com
18201 Von Karman, Ste. 700, Irvine, CA 92612

CYLANCE