mcafee.com

Did You Check Your Quarantine?! | McAfee Blogs

By Prabhat Singh and Sameer Paranjape on Oct 28, 2019

11-13 minutes

A cost-effective way to detect targeted attacks in your enterprise

While it is easy to get caught up in the many waves of new and exciting protection strategies, we have recently discovered an interesting approach to detect a targeted attack and the related actor(s). Quite surprisingly, a big part of the solution already exists in most enterprises: the tried, tested endpoint security platform. In this case, we used our own McAfee Endpoint Security (ENS). Along with ENS, we used GetQuarantine, a freeware tool from McAfee, and a third-party threat analytics service.

The Problem

We will begin with a working definition of a targeted attack:

A targeted attack is a threat in which a threat actor actively pursues and compromises a specific target. To achieve the

goal, the adversary may adapt and improve their attack(s) to counter the victim's defenses and persist at it for a long period of time.

What does this say? First, the adversary's objective is to compromise a specific target, not just an arbitrary target. Second, the adversary is skilled enough to know how defenses work and is resourceful enough to actively adapt and improve their attack to beat defenses. Third, the adversary is determined enough to pursue the objective for perhaps an indefinite period of time.

Taken together, the above characteristics challenge most defense technologies. Why so? Because these characteristics run counter to the assumptions on which these technologies are based.

At the heart of it, most defense technologies are signature-based, where the signatures are created either by a human analyst, by a machine, or by using instances of known malicious behavior. The cost of constructing signatures is high and is amortized by using the same signature to defend against the same attack elsewhere.

Twenty years ago, when there were just a few thousand examples of malicious software around, it was relatively easy to find the origin, perpetrators, and the reason for the creation and release of a malicious file or application. Security researchers would manually analyze each sample, carefully identify similarities with previously known samples through

sheer memory and label each sample with a unique name. This method worked well because the attacks then were opportunistic and aimed at spreading as wide as possible. This meant that anti-virus companies could discover an attack in one place, extract relevant detection signatures, and send the signature updates to its install base.

Now, security threat intelligence companies receive hundreds of thousands of new malware samples every day. There is simply not enough time or resources to analyze each malware to answer who, what, when, and why? The best any anti-virus software can do is to classify a file into just two bins: good or bad. It is impossible for researchers to manually look at every sample and process them to the same detail as before. To make matters worse, attacks today are targeted. Attackers create one-off variants aimed at a specific enterprise. This makes it virtually impossible to connect attacks across enterprises to understand the attacker.

And therein lies an important problem. Just as the numbers and sophistication of attacks have increased exponentially, the objective of tracking who is behind an attack, and identifying linkages between different malware samples and their authors, and the intent behind an attack, have been lost.

Why should it matter? In the absence of information about who is attempting to breach an organization, defenders are left operating in the dark. They make security decisions based on breaches that happen elsewhere using threat intelligence that

is most often irrelevant, and when relevant, is most likely outdated.

The Solution

Analysis of targeted campaigns shows that programs that are part of an attack usually show a couple of similar characteristics. First, the malware or attack mechanism is focused on one enterprise or, at most, one sector and second, the malware program demonstrates evolutionary characteristics where the actor repeatedly unleashes different variants of it. Our proposed solution focuses on these characteristics and tries to uncover targeted campaigns by finding binary semantics between malware found in customer environments and known targeted campaigns.

Our solution strategy is:

Endpoint-security detects a malware sample. It is compared with a sample from a known targeted attack. If the similarity is high, it is a strong indication that the ENS detected sample is a part of that targeted attack and the threat actor is the same.

The strategy is implemented in three building blocks: sample collector, sample storage and targeted attack analysis using third-party threat analytics application.

Sample Collector (GetQuarantine)

Sample collection is performed using McAfee proprietary licensed freeware, GetQuarantine. GetQuarantine is a McAfee

e-Policy Orchestrator (ePO) deployable tool that can run on all endpoints protected by McAfee ENS. GetQuarantine runs as an ePO scheduled product deployment task. ENS cleans or deletes items that are detected as threats and saves copies in a non-executable format to the Quarantine folder. The GetQuarantine tool on a scheduled run, checks the quarantine folder and uploads items to the McAfee backend if items are not already uploaded during previous tool runs.

Sample Storage (McAfee Workflow & AWS)

The McAfee workflow backend receives sample submissions from GetQuarantine and stores them in an S3 bucket. Samples are segregated per enterprise and made available for further analysis. Third-party analytics applications like MAGIC can be run on samples to extract targeted attack insights. Analytics services are available to McAfee customers participating in a third-party analytics program. For customers that do not participate in a third-party analytics program, sample processing ends at the McAfee backend layer and the sample eventually gets deleted without further analysis.

Targeted Attack Analysis

For our pilot implementation we used Cythereal MAGIC services. The McAfee backend submits samples to MAGIC for binary similarity analysis. Customers can check analysis reports using Cythereal website. Cythereal is McAfee's Security Innovation Alliance (SIA) partner.

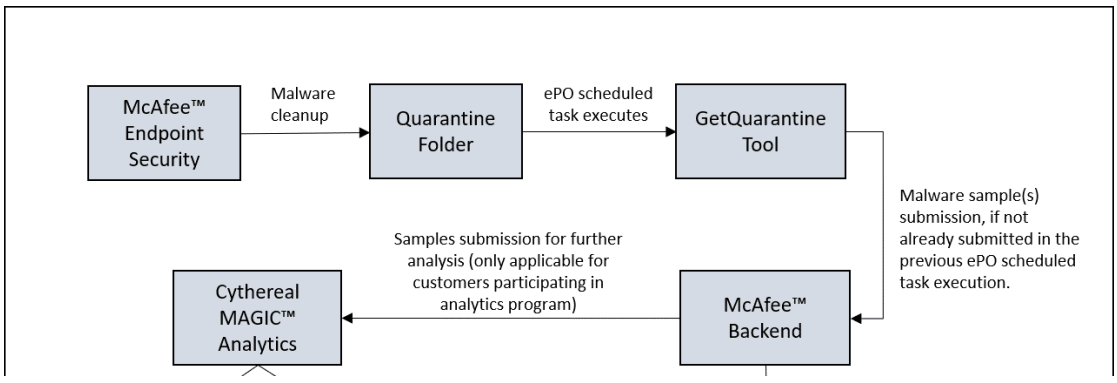
Cythereal MAGIC™ (malware genomic correlation) is a web-

service touted as “BinDiff on Steroids”. The system carries out semantic similarity analysis of programs using advanced program analysis techniques at the assembly instruction-level code. The semantics of the program gives more meaningful insights than structural or behavioral characteristics. MAGIC can find similarity between samples submitted using GetQuarantine and also find variants of those samples from MAGIC’s database. It has the facility to provide alerts for campaign detections and can generate YARA rules that can be used for searching other services, like VirusTotal.

We first tried human-driven in-house analysis using open source tools like SSDEEP, SDHASH, TLSH, etc. to prove the concept of identifying targeted attacks using the binary similarity of samples found in quarantine. Though we were successful in proving this concept with these open source tools, they were not very effective, especially with polymorphic variants, so we explored third-party options and identified Cythereal MAGIC™.

Architecture

Figure 1 shows the overall architecture of our system:



[Figure 1: McAfee ENS detects a suspicious sample by studying its behavior or other means and then moves the sample file to the quarantine folder. The scheduled execution of the GetQuarantine Tool configured in ePO as a scheduled task submits the sample to the McAfee backend. The third-party analytics app, periodically receives samples from McAfee backend for further analysis.]

Case Study

For a case study, we used samples from a McAfee discovered campaign, Oceansalt. We tested the solution's ability to group samples using semantic similarity and also tested the solution's ability to identify new variants of Oceansalt samples.

Illustration of the Solution's Ability to Group Malware From Quarantine

McAfee Endpoint Security (ENS) detected two samples of Oceansalt (as listed in Table 1). GetQuarantine submitted these samples to the McAfee backend. Targeted attack analysis of these files showed a semantic similarity of 95.1%. The comparison of their control-flow graphs in Figure 2 justifies the high semantic similarity score.

MD5 Hash	Compilation date (yyyy-mm-dd)

531DEE019792A089A4589C2CCE3DAC95 [VT]	2018-6-13
76C8DA4147B08E902809D1E80D96FBB4 [VT]	2018-7-17

[Table 1: Oceansalt samples reported by McAfee™ security operation center in June-July 2018.]

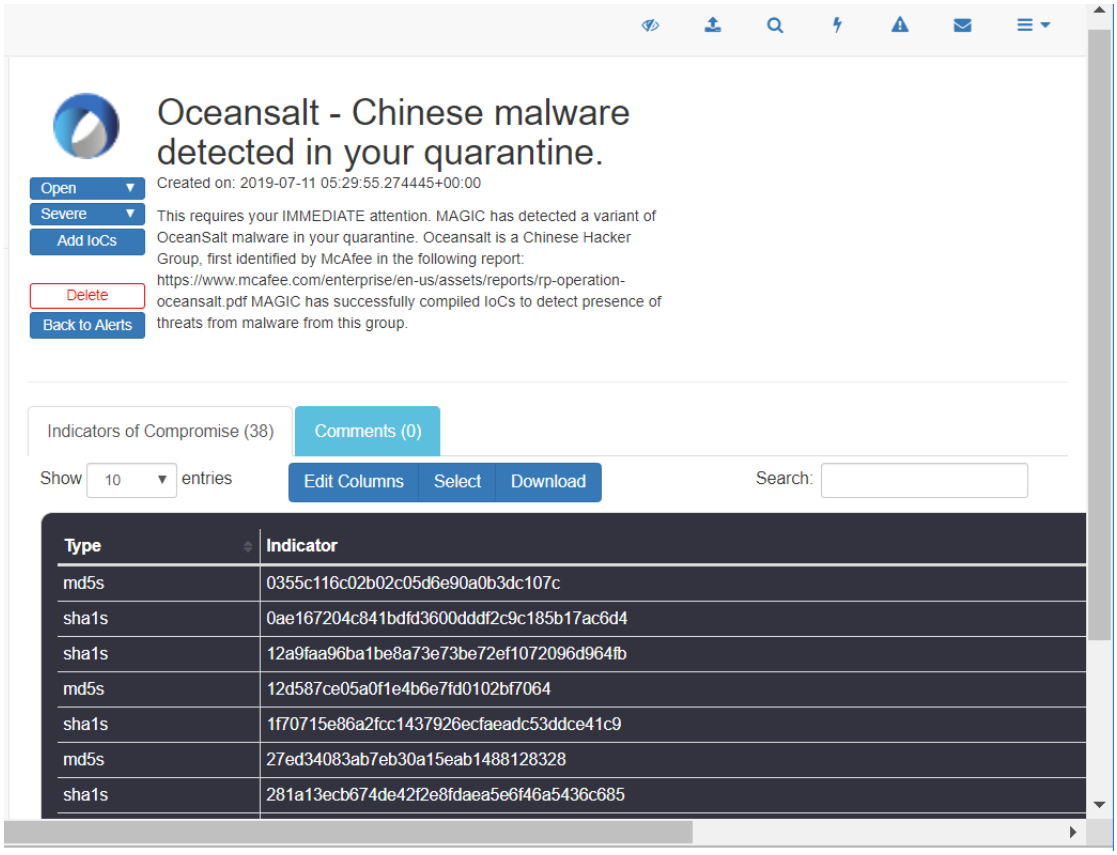
[Figure 2: Control-flow graph of Oceansalt samples from Table 1]

Illustration of the Solution's Ability to Link New Variants From the Wild to a Known Targeted Attack

Finally, we come to the use case that motivated this study. Malware belonging to a targeted attack is identified by its file-hashes. However, attackers use polymorphism and other obfuscations to create new variants. Though McAfee ENS may block such variants, it may not link it to the original attack. Targeted attack analytics can help fill this void.

To test the solution's ability to locate such targeted attacks, we uploaded an Oceansalt sample (MD5: 531DEE019792A089A4589C2CCE3DAC95 [VT]) to MAGIC and identified it as an APT. We then uploaded a large number (thousands) of malware samples via GetQuarantine. As we

had thought, targeted attack analytics sent an alert that it had detected variants of Oceansalt (Figure 3).



[Figure 3: Alert about detecting an Oceansalt variant in the quarantine]

MAGIC's alert was triggered because it found two Oceansalt variants from the wild which were not previously reported by the McAfee SOC or any other global threat intelligence.

MD5 hash	Compilation date (yyyy-mm-dd)	First seen on VirusTotal (yyyy-mm-dd)
7EFEB42FF8E04FAD22EFDE189C87CB23 [VT]	2018-6-14	2019-1-2
12D587CE05A0F1E4B6E7FD0102BF7064 [VT]	2018-7-12	2019-1-2

[Table 2: Two new variants of Oceansalt samples found using semantic similarity]

Try Your Quarantine

We tested the GetQuarantine-based solution in our lab and found encouraging results. If you would like to try out this solution use the following steps, along with McAfee Endpoint Security (ENS):

- Download the beta version of GetQuarantine, a proprietary licensed freeware.
- Deploy it using the ePO ecosystem.
- On successful sample submission to the McAfee backend, receive an acknowledgment email.

To obtain analysis results from the third-party analytics app, follow these steps:

- Visit Cythereal MAGIC™.
- The MAGIC dashboard contains plots with details about various ongoing campaigns.
- Upon selecting a campaign in the plot, a table with all the associated malware is displayed where the customer can download samples and YARA rules.
- Whenever MAGIC detects a targeted attack, it sends an alert email to the registered email address of the customer, along with additional threat intelligence, such as information on the threat group, third-party research on the group, and associated IoCs. Customers can also see the list of alerts on the MAGIC website.

Summary

As you can see from this exercise, traditional AV still has lot to offer and can play an important role in overall security strategy againt targeted attacks. We can amplify signals coming out of AV detections using tools like GetQuarantine and by running analytics on quarantine artifacts to uncover targeted attacks. We can take an incremental approach in solving targeted attack challenges.