TEST METHODOLOGY

Endpoint Detection and Response (EDR)

July 12, 2018

v1.1

Table of Contents

1 Introduction

1.1 The Need for Endpoint Detection and Response

Cybercriminals have become adept at technical and social engineering, and contemporary threat actors are capable of carrying out sophisticated attacks that consistently breach modern network defenses. Oftentimes, these breaches lead to end user systems being infected and subsequently used as a stage for further compromise.

Strong anti-threat protection technologies on the endpoint are the best chance enterprises have at defeating these many incursions, but there are times when current products and techniques cannot stop even the least capable of the advanced threats, let alone the truly determined advanced persistent threat.

This creates considerable challenges for the security analyst who must also be a forensics expert in order to determine what occurred on the endpoint; whether any malicious and functional threat remnants exist; and what remediation is required. Endpoint detection and response (EDR) products address these challenges by capturing all of the relevant threat detail on an endpoint in order to assist with remediation and to determine whether any compromises occurred as a result of the threat.

EDR are often deployed alongside endpoint protection products and provide comprehensive visibility into the behaviors of threats that are not blocked so incident response investigations can focus less on how other security controls were bypassed and more on what the threat did and whether any data was compromised or lost. They provide actionable guidance for forensic security analysts, which allows the analysts to more efficiently deal with threats. At its core, an EDR product is both a cybercrime scene investigation toolkit and a means to improve the efficiency of incident response, which is one of the most challenging and resource-intensive activities in security today.

1.2 Definition

Endpoint detection and response (EDR) refers to endpoint security technology that supports the threat visibility needs of forensic security analysts. EDR products in this test are not allowed to perform automated blocking techniques.

An EDR product provides robust threat event reporting about every endpoint system it monitors. The following features are fundamental to an EDR product:

- Continuous monitoring of endpoint
- Detection or alerting on anomalous activity
- Forensic details that support incident response

In addition, many EDR products can remediate threats on the endpoint. However, this is not a requirement of the EDR solutions in this test since many large organizations do not rely on automated remediation tools, but instead rely on high-caliber forensic data to perform remediation themselves (e.g., scripts, backup tools).

In order to detect inherently malicious or potentially malicious behavior, an EDR product utilizes multiple approaches and technologies. Examples include process monitoring, detecting communication with potentially malicious hosts, detecting lateral movement to other machines, and auditing the file system and registry.

The threat event reporting must provide forensic teams with the information they need to investigate suspicious activity. Note that assessing the capabilities of these teams is beyond the scope of this methodology; it is assumed that they will be able to interpret the forensic details provided by these tools.

1.3 About This Test Methodology

This document establishes a methodology for evaluating security products that provide detection and response for endpoint systems on enterprise networks and that meet the criteria for classification as EDR products.

The output of a test conducted according to this methodology should reflect the effectiveness of the product in supporting investigation and forensic work.

In order to thoroughly evaluate products under test, several factors, attributes, and performance metrics are gathered and appraised. The scope of this methodology includes:

- Detailed threat visibility
- Total cost of ownership (TCO)
- Threat event reporting

1.4 Inclusion Criteria

NSS Labs welcomes the participation of any vendor whose product meets the minimum definition as written above.

1.5 Deployment

An EDR product is often deployed as an agent on an endpoint that reports to a central management apparatus, which can reside on a physical appliance, a virtualized appliance, or in the cloud.

The product will be deployed during testing in a manner that reflects the “enterprise standard,” or “default” policy.

NSS must approve and validate the configuration of products prior to their being tested. Records and configuration backups (if available) will be taken to ensure consistency during testing.

2 Test Components

2.1 Security Effectiveness

A variety of metrics are collected during each testing phase. This data is used to understand product reporting capabilities versus cost of ownership. Products are assessed for their ability to provide continuous visibility into monitored endpoint systems. Diverse threat scenarios are played out against endpoint systems designated as victims.

One of the most common threats to the enterprise is the infection of systems by malicious programs. The EDR product must detect activity by malicious programs. Products are assessed for their ability to detect and provide detailed information on threats from the following categories:

- Malware
- Exploits
- Blended threats
- False positives
- Evasions
- Unknown threats
- Any combination of the above in addition to secondary threat actions or behaviors (initial infection with secondary environment scan)

Where possible, these threats will be employed as “real world” in the test environment (i.e., using actual URLs and IP addresses). Furthermore, the threats will be introduced into host systems in a “real-world” manner (i.e., NSS will not disable the security product to get the malware on the host). The product’s ability to detect and provide forensic artefacts about threats from each of these categories will be measured. A combination of publicly available tools and proprietary software is used. The endpoint is monitored at various stages of the attack, as is the behavior of the threat.

Products are expected to alert on and provide supporting data on threats or suspicious activity. The product under test must identify malicious behaviors and suspicious changes to the state of the endpoint machine and must accurately capture all significant changes to the system as well the impact of these changes. The data should include sufficient forensic details to ensure that analysts can ascertain the nature of the threat and whether further actions or response is warranted. The value of an EDR product lies both in its capability to accurately detect threats and in its capability to support response efforts.

2.1.1 Socially Engineered Malware

Socially engineered malware can infect an endpoint using numerous attack vectors or delivery methods. Examples include:

- Binary attachments sent via email
- Executable downloads from websites

2.1.2 Exploits

An exploit is an attack against a computer that takes advantage of a vulnerability in some part of the system, such as a logical flaw in a program installed on the machine. Exploits do not require user intervention or knowledge.

This test exploits certain vulnerabilities on a victim endpoint system. Various vulnerable applications and system features are used, and multiple payloads are employed in conjunction with them.

When a victim is successfully exploited, the EDR product should capture sufficient forensic detail to inform security teams about the steps taken prior to, during, and after infection. Examples of such post-exploitation activity include data exfiltration, lateral movement, and pivoting.

2.1.3 Blended Threats

Blended threats combine multiple delivery vectors in attempts to compromise an endpoint system. For example, instead of relying on one malicious email or a single exploit attempt, the blended threat will leverage exploiting multiple vulnerabilities, such as spear phishing, infected peripherals, and sophisticated antivirus evasion techniques to infect the endpoint system.

2.1.4 Evading Protection

There are many ways to defeat the detection capabilities of endpoint products. This methodology stipulates that all features of security effectiveness testing employ some form of evasion method, including but not limited to:

- Executable binary packing
- File compression
- In-memory execution
- Malware environmental analysis and awareness
- Code morphism
- Web socket abuse

2.1.5 Unknown Threats

This test evaluates a product's ability to detect malicious behavior by previously unknown threats. Unknown threats are threats that have not been encountered by the product prior to testing.

2.1.6 False Positives

An EDR product must be able to correctly identify and permit non-malicious activity on an endpoint. It is expected that an EDR product will collect anomalous benign behavior about false positive samples. However, false positive samples that the EDR product perceives as malicious or that potentially result in unnecessary incident response or remediation efforts will be considered operational costs.

If a product incorrectly identifies benign data or programs on an endpoint as a threat to the system and the network, this reduces the analysts' ability to sift through the data and identify actual threats. The rate at which a product incorrectly identifies benign activity is defined by the ratio of successful false positive triggers to attempted false positive triggers. False positive alerts directly affect reporting and operational costs as they must be investigated.

Accurate reporting on false-positive events can reduce remediation efforts; however, inaccurate reporting on these events requires investigation, which leads to higher ownership costs.

Attempts to generate false-positive reports involve the use of benign software or data on the endpoint system by a simulated user.

Many file types commonly seen throughout the enterprise are examined in this portion of the test and include but will not be limited to:

- Portable executable (PE32) files
- Document files (.docx, .pdf, etc.)
- Scripts

2.2 Reporting

After an EDR product detects malicious or anomalous behavior, it should support follow-on investigation, incident response, and remediation efforts. EDR products allow analysts to better understand the context surrounding the intrusion.

The testing performed in the Reporting section of this methodology will determine whether data from either the management console/user interface or from centralized logging allows incident responders to build a clear picture of various post-exploitation activities. The EDR product should convey timely, detailed threat event and forensics data to its central management system. The product must be capable of cataloging threat events continuously, barring technical failure (such as network upset), or in very rare circumstances, administrative override (e.g., allowing for a lapse in logging functionality in the face of performance concerns).

Forensic data about a threat to an endpoint system must be included with the basic threat event data conveyed to the central management system. This supports the detection conviction and enables responders to act more quickly or establish incident severity more readily. Forensic test components will be observable, industry-accepted techniques such as those described in the MITRE Adversarial Tactics, Techniques, and Common Knowledge (ATT&CK™) Matrix. For malware samples collected "in the wild," this reporting category will measure observable, automated post-exploitation activity. For malware samples built in-house, NSS will manually run up to a dozen post-exploitation techniques per sample.

Reporting details can be aggregated or segregated according to where a threat is in its life cycle in order to reveal how a product observes a given threat (or group of threats) at specific points in time. EDR products do not require a one-to-one correlation between an event (such as lateral movement for example) and a user interface or log item. However, products should capture forensic artefacts to help analysts understand what happened during the intrusion.

The behaviors and activities of the threats submitted to a target environment have already been documented by NSS, and these will be assessed against the data captured by the EDR product.

While key attributes or activities are often used (and expected to be recorded) for a detection event or alert, all of the details captured leading up to and after a detection event are rarely, if ever, shared. However, during incident response, it is important for responders to be able to pivot from an alert to the event details and sort through them, and the EDR product should facilitate this. This portion of the test will document how well a product captures aspects of a threat according to common data attributes and how quickly the product allows this data to be accessed. The attributes assessed will depend on the threat type or the techniques used (see section 2.1), but may include source IP, application file system path, PowerShell commands, potential command-and-control or outbound connections attempted, and convicted blacklist IP match.

EDR products are often used to search through forensic and threat details to identify whether newly discovered or shared indicators (e.g., new hashes) have existed in the past on one of the EDR-managed endpoints.

2.3 Remediation (Optional)

Several EDR products possess the ability to take action once certain alert criteria are met. Depending on the maturity of their forensics or security analysis teams, some enterprises may leverage this capability. While validation of all remediation approaches is not in scope for this round of testing, NSS will document whether remediation capabilities exist for the tested product. These capabilities could include host isolation, process termination, and file and/or process isolation.

2.4 Total Cost of Ownership

Total cost of ownership, or TCO, refers to any data that can be used to quantify a dollar figure that represents the expected costs of utilizing the product for endpoint visibility. The scope of these metrics includes, but may not be limited to:

- Product purchase costs
- Product maintenance and update costs
- Installation costs
- Threat-associated costs
- Time to access key data points that support a threat conviction or malicious event
- Usability
- False positives

EDR products expedite the response process by helping responders rapidly determine what occurred after the exploitation, whether or not data loss has occurred, and what action needs to be taken post incident. The speed at which incidents are resolved depends on how quickly an EDR product provides accurate, high-quality data and how efficiently it handles data “distractions.”

However, it can be complex to deploy an EDR product in an enterprise network, with several factors affecting the overall cost of deployment, maintenance, and upkeep. All of these factors should be considered over the course of the useful life of the product.

3 Product Guidance

NSS issues summary product guidance based on the evaluation criteria that is important to information security professionals. These criteria include:

- Detection
- Incident detail and supporting data
- Total cost of ownership

Each product will be given a guidance rating.

3.1 Recommended

A *Recommended* rating from NSS indicates that a product has performed well and deserves strong consideration. Only the top technical products earn a *Recommended* rating from NSS – regardless of market share, company size, or brand recognition.

3.2 Security Recommended

A *Security Recommended* rating from NSS indicates that while a product has demonstrated excellent security effectiveness throughout the test, its cost of ownership is higher than the test average. This makes enterprise adoption of the product a costly proposition.

3.3 Neutral

A *Neutral* rating from NSS indicates that a product has performed reasonably well and should continue to be used if it is the incumbent within an organization. Products that earn a *Neutral* rating from NSS deserve consideration during the purchasing process.

3.4 Caution

A *Caution* rating from NSS indicates that a product has performed poorly. Organizations using these products should review their security posture and other threat mitigation factors, including possible alternative configurations and replacement. Products that earn a *Caution* rating from NSS should not be shortlisted or renewed.

Appendix A: Change Log

Version 1.1 – July 12, 2018

- 1.1: Clarifications concerning threat remnants
- 1.2: Explicitly specified that EDR products under test cannot perform automated blocking
- 1.5: Changed specific requirement for how an EDR product is deployed
- 2.1: Clarified how threats would be employed in the test environment
- 2.1.6: Clarified how EDR false positives will be identified
- 2.2: Elaborated on how EDR reporting will be measured using common manual and automated post-exploitation activity

Version 1.0 – June 6, 2018

Contact information

NSS Labs, Inc.
3711 South MoPac Expressway
Building 1, Suite 400
Austin, TX 78746-8022
USA
info@nsslabs.com
www.nsslabs.com

This and other related documents available at: **www.nsslabs.com**. To receive a licensed copy or report misuse, please contact NSS Labs.

Please read the disclaimer in this box because it contains important information that binds you. If you do not agree to these conditions, you should not read the rest of this report but should instead return the report immediately to us. "You" or "your" means the person who accesses this report and any entity on whose behalf he/she has obtained this report.

1. The information in this report is subject to change by us without notice, and we disclaim any obligation to update it.

2. The information in this report is believed by us to be accurate and reliable at the time of publication, but is not guaranteed. All use of and reliance on this report are at your sole risk. We are not liable or responsible for any damages, losses, or expenses of any nature whatsoever arising from any error or omission in this report.

3. NO WARRANTIES, EXPRESS OR IMPLIED ARE GIVEN BY US. ALL IMPLIED WARRANTIES, INCLUDING IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, AND NON-INFRINGEMENT, ARE HEREBY DISCLAIMED AND EXCLUDED BY US. IN NO EVENT SHALL WE BE LIABLE FOR ANY DIRECT, CONSEQUENTIAL, INCIDENTAL, PUNITIVE, EXEMPLARY, OR INDIRECT DAMAGES, OR FOR ANY LOSS OF PROFIT, REVENUE, DATA, COMPUTER PROGRAMS, OR OTHER ASSETS, EVEN IF ADVISED OF THE POSSIBILITY THEREOF.

4. This report does not constitute an endorsement, recommendation, or guarantee of any of the products (hardware or software) tested or the hardware and/or software used in testing the products. The testing does not guarantee that there are no errors or defects in the products or that the products will meet your expectations, requirements, needs, or specifications, or that they will operate without interruption.

5. This report does not imply any endorsement, sponsorship, affiliation, or verification by or with any organizations mentioned in this report.

6. All trademarks, service marks, and trade names used in this report are the trademarks, service marks, and trade names of their respective owners.