TEST METHODOLOGY

Advanced Endpoint Protection (AEP)

JUNE 9, 2017

v2.0

Table of Contents

1 Introduction

1.1 The Need for Advanced Endpoint Protection

Cybercriminals are becoming ever more 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 current products and techniques are generally unable to stop even the least capable of the advanced threats that threaten the enterprise, let alone the truly determined advanced persistent threat.

In addition to the need for innovative protection technologies as a part of a thorough defense-in-depth strategy, the superior product must offer comprehensive monitoring and threat visibility. Why blindly block an infection or exploitation attempt when identifying metadata and incursion telemetry could mean the difference between the threat continuing to attack until it succeeds and permanently stopping it?

Today's enterprise requires both superior protection and robust, pervasive threat monitoring capabilities in its security posture. This is the purview of advanced endpoint protection.

1.2 Advanced Endpoint Protection Definition

Advanced endpoint protection (AEP) refers to endpoint security technology that combines the protective capabilities of anti-threat products with the detection, investigation, and prevention capabilities of endpoint monitoring products.

An AEP product is one which provides automatic threat prevention and threat event reporting capabilities to every endpoint system it protects. The following features are fundamental to an AEP product:

- Prevention
- Detection
- Forensics

In order to detect different types of malicious or threatening behaviors, an AEP product may utilize multiple approaches and technologies. Examples of such approaches include process monitoring, detecting communication with potentially malicious hosts and lateral movement to other machines, as well as auditing the file system and registry.

The level of detail provided within threat event reporting must provide threat forensic teams with the information they need to investigate suspicious activity. Note that presuming a certain level of responder capability for the purpose of a test conducted according to this methodology is beyond the scope of this document.

Protection and threat event forensic detail form the primary feature platform of AEP. While additional functionality may be available that enhances the overall security of the endpoint system, neither protection nor forensic monitoring may be removed without declassifying the AEP system.

1.3 About This Test Methodology

This document establishes a methodology for evaluating a security product that furnishes protection for some number of endpoint systems on an enterprise network and that meets the criteria for classification as an AEP product.

The output of a test conducted per the methodology outlined here should reflect the effectiveness of the product in its goal of securing its stock and/or assets.

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

- Security effectiveness
- Total cost of ownership (TCO)
- Secure communication
- Threat event reporting

1.4 Inclusion Criteria

NSS Labs welcomes the participation of any vendor whose product meets the minimum platform definition as written above. No product may consist of any appliance or sub-solution, physical or virtualized, beyond the endpoint agent/management station architecture.

1.5 Deployment

An AEP product should be deployed as an agent on the endpoint that reports to a central management apparatus, which resides on either a physical appliance, a virtualized appliance, or in the cloud.

The product should be deployed in a manner that reflects the “enterprise standard,” or “default” protection policy. This policy is reflective of common best practices seen throughout the enterprise and does not take into account the numerous adjustments and recommendations that an enterprise might employ for its own sake.

NSS must approve and validate the configuration of the product prior to it being tested. Records and configuration backups (if available) will be taken to ensure consistency across the testing process.

1.6 Test Metrics

A variety of metrics are collected during each testing phase. This data is used to understand the product’s prevention and reporting capabilities versus its cost of ownership.

The relationship between a product’s security effectiveness, which is derived from its measured preventative and reporting capabilities, and its cost of ownership, are used to compare the products tested against one another.

1.6.1 Security Effectiveness

The ability of the product under test to successfully secure its endpoints will be represented by the *Overall Prevention Rate* (OPR). The OPR is the ratio of successful preventative actions taken by the product against threats to the total number of threats employed during the test.

A product's OPR can be aggregated to represent the total performance against all threats at any time during the test, as well as segregated into individual ratings for each phase of the kill chain:

- Pre-infection OPR
- On-infection OPR
- Post-infection OPR

This provides visibility into the performance of a product under test before, during, and after the moment of exploitation or infection.

1.6.2 Threat Event Reporting

A product's *Overall Reporting Rate* (ORR) represents its ability to convey threat event and forensics data to the product's central management station.

The ORR is the ratio of threat event reports relayed to the central management apparatus to the total number of prevented threats targeting systems protected by the product under test.

Similar to the product's OPR, the ORR can be aggregated or segregated by threat life cycle phase for the purposes of underscoring a product's performance against threats at a certain point in time. The ORR can then be computed either as the aggregate ORR, the pre-infection ORR, the on-infection ORR, or the post-infection ORR.

1.6.3 False-Positive Rating

When a product incorrectly identifies benign data or programs on an endpoint as a threat to the system and network, this reduces the product's ability to defend against actual threats.[1]

The rate at which a product incorrectly identifies and prevents such benign endpoint traffic is defined by the ratio of successful false positive triggers to attempted false positive triggers, and is called the *Overall False-Positive Rate* (OFR).

A product's OFR is deducted from the product's OPR. Reporting on false-positive events reflects the correct function of a product's reporting apparatus, and therefore does not result in a penalty of any kind.

1.6.4 Total Cost of Ownership

Total cost of ownership, or TCO, refers to any data that can be used to quantify a dollar figure, which represents the expected costs of utilizing the product for endpoint defense. 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

[1] See section 2.1.6 for more information on false positives.

2 Test Components

2.1 Security Effectiveness

In the security effectiveness portion of the test, a product is assessed for its ability to provide continuous protection to secured endpoint systems. Here, diverse simulated threat scenarios are played out against endpoint systems designated as victims.

Products are tested against threats from the following categories:

- Malware
- Exploits
- Blended threats
- False positives

A test executed according to this methodology measures the effects of each of these threat categories against the endpoint systems. A combination of publicly available tools and proprietary software is used, which monitors the state of the endpoint at various stages of the attack, as well as the behavior of the threat as it plays out.

A product must successfully prevent a threat and generate a threat event report for each attack scenario it is tested against.

2.1.1 Socially Engineered Malware

One of the most common threats to the enterprise is system infection by malicious programs. It is the job of the AEP product to detect and prevent the malicious program's activity.

NSS evaluates the ability of the product under test to prevent malware from harming the victim endpoint system by exposing a set of endpoints to a continuous stream of live malware sourced from the wild.

There are numerous possible attack scenarios and numerous delivery vectors by which malware can infect the endpoint. Examples of malicious delivery vectors include binary attachments sent via email, as executable downloads from websites, in peer-to-peer transfers, or as self-extracting programs.

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 attempts to exploit certain vulnerabilities on a victim endpoint system. Various vulnerable applications and system features are used, and multiple payloads are employed in conjunction with them.

If a victim is successfully exploited, additional malicious activity may be attempted. 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, 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 and prevention capabilities of endpoint protection products. This methodology stipulates that all features of security effectiveness testing employ some form of the following methods of evasion, which include, but are not limited to, executable binary packing, file compression, in-memory execution, malware environmental analysis and awareness, code morphism, and web socket abuse.

The product under test must identify malicious behaviors and suspicious changes to the state of the endpoint machine and must work to prevent the evaded threat from damaging the endpoint.

2.1.5 Unknown Threats

This test evaluates the product's ability to detect and prevent against compromise by previously unknown threats. These threats are defined as those which have not been encountered by the product prior to the point of testing.

This test can be conducted in several ways, including, but not limited to, trailing product update testing, offline endpoint testing, and infected peripheral testing.

2.1.6 False Positives

An AEP product must be able to correctly identify and permit non-malicious activity on an endpoint system if it is to successfully protect against attack and subsequent compromise.

False positives are any content or activity on an endpoint system that the product perceives as malicious, which interrupt the user's experience .

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.), and scripts.

The rate at which a product incorrectly triggers alerts against benign files or programs directly affects the product's OPR. See section 1.6.3 for more information.

2.2 Total Cost of Ownership

Deploying an AEP product in an enterprise network can be a complex task, 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.

2.3 Threat Event Reporting

This test measures the completeness and timeliness of the product under test's threat event reporting apparatus. The product must be capable of cataloging threat events continuously, barring either technical failure (such as application fault or 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 given threat to the endpoint system must be included with the basic telemetric threat event data conveyed to the central management station. This constitutes the product's threat context awareness.

2.3.1 Threat Context Awareness

This test measures the ability of the product under test to assess the effects of a threat's activities on an endpoint system. The test also measures the product's ability to gather accurate forensic data about the threat's behavior on the system. The product must be able to correctly separate irrelevant data and system noise from the threat's own activities.

2.4 Secure Communication

If a product's feature set indicates that it can communicate sensitive data between the endpoint system agent and its central manager, NSS evaluates this claim. Network traffic between the endpoint and the manager is captured, recorded, and then analyzed to verify whether the data transmitted is truly encrypted.

3 Product Guidance

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

- Security effectiveness
- Secure communication
- Threat event reporting
- 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 a product has exhibited exemplary security effectiveness throughout the testing cycle. However, the product's costs of ownership are greater than average, making adoption of the product into an enterprise 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 one of 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.

Contact information

NSS Labs, Inc.
206 Wild Basin Rd.
Building A, Suite 200
Austin, TX 78746 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.