Keywords: anti-malware; compliance; assessment; testing; test plan; template; endpoint; security; SE Labs

SE Labs and AMTSO
Preparation Date: March 19th ,2020
Review Date: March 19th, 2020

Version 1.1

Q1 2020 Endpoint Protection Test Plan: Enterprise, Small Business, and Consumer

Authored by:
SE Labs (Simon Edwards, Stefan Dumitrascu)

Abstract:
This Test Plan Template provides the structure for constructing a Test Plan that is compliant with the AMTSO Testing Protocol Standards. This document is an informative reference to the AMTSO Testing Protocol Standard for the Testing of Anti-Malware Solutions (the "Standard"), and specifically to the requirements within such Standard for Test Plan construction and presentation. Wherever conflicts might exist between this Template and the Live Standards Version 1.3, the Testing Protocol Standards will provide the prevailing rule.

www.amtso.org

Table of Contents

SE Labs Test Plan for Q2 2020 Endpoint Protection

1. Introduction

SE Labs tests a variety of endpoint security products from a range of well-known vendors in an effort to judge which were the most effective. Each enterprise, small business, or consumer class product are exposed to the same threats, which are a mixture of targeted attacks using well-established techniques, public email, and web-based threats that are known or found to be live on the internet at the time of the test. The Test Reports indicate how effectively the products were at detecting and/or protecting against those threats in real time. The most recent quarterly reports for each class of application are listed below.

https://selabs.uk/download/consumers/epp/2019/oct-dec-2019-home.pdf

https://selabs.uk/download/small_business/epp/2019/oct-dec-2019-smb.pdf

https://selabs.uk/download/enterprise/epp/2019/oct-dec-2019-enterprise.pdf

2. Scope

The SE Labs Endpoint Test examines applications from the enterprise, small business, and consumer sector with the following companies and software products composing the latest tests.

ENTERPRISE PRODUCTS	
Vendor	**Product**
Bitdefender	Bitdefender GravityZone Endpoint Security
Crowdstrike	Crowdstrike Falcon
FireEye	Endpoint Security
Kaspersky Lab	Kaspersky Endpoint Security
McAfee	McAfee Endpoint Security
Microsoft	Microsoft Windows Defender Enterprise
Sophos	Intercept X Advanced
Symantec	Symantec Endpoint Security Enterprise Edition

SMALL BUSINESS PRODUCTS	
Vendor	**Product**
Bitdefender	Bitdefender Gravity Zone Endpoint Security
Kaspersky Lab	Kaspersky Small Office Security
Microsoft	Microsoft Windows Defender Enterprise
McAfee	McAfee Endpoint Security
Sophos	Intercept X Advanced
Trend Micro	Trend Micro Worry-Free Security Services
Webroot	SecureAnywhere Endpoint

CONSUMER PRODUCTS	
Vendor	**Product**
AVAST	Avast Free Antivirus
AVG	AVG Antivirus Free Edition
Avira	Avira Free Security Suite
Comodo	Comodo Internet Security Premium
F-Secure	F-Secure Safe
G-Data	Internet Security
Kaspersky Lab	Kaspersky Internet Security
McAfee	McAfee Total Protection
Microsoft	Microsoft Windows Defender Consumer
Symantec	Norton Security
Sophos	Home Premium
Trend Micro	Trend Micro Internet Security
Vipre	Vipre Endpoint Security
Zone Alarm	Zone Alarm Free Antivirus

3. Methodology and Strategy

Test Framework - The test framework collects threats, verifies that they will work against unprotected targets and exposes protected targets to the verified threats to determine the effectiveness of the protection mechanisms.

Threat Management Systems (TMS) - The Threat Management System is a database of attacks including live malicious URLs; malware attached to email messages; and a range of other attacks generated in the lab using a variety of tools and techniques. Threats are fed to the Threat Verification Network (TVN).

Threat Verification Network (TVN) - When threats arrive at the Threat Verification Network, they are sent to Vulnerable Target Systems in a realistic way. For example, a target would load the URL for an exploit-based web threat into a web browser and visit the page; while its email client would download, process and open email messages with malicious attachments, downloading and handling the attachment as if a naïve user was in control. Replay systems are used to ensure consistency when using threats that are likely to exhibit random behaviors and to make it simpler for other labs to replicate the attacks.

Target Systems (TS) - Target Systems (TS) are identical to the Vulnerable Target Systems used on the Threat Verification Network, except that they also have endpoint protection software installed.

Threat Selection - All of the following threats are considered valid for inclusion in the test, although the distribution of the different types will vary according to the test's specific purpose:

- Public exploit-based web threats (exploitation attacks)

- Public direct-download web threats (social engineering attacks)
- Public email attachment threats (exploitation and social-engineering attacks)
- Private exploit-based web threats (exploitation attacks)
- Private direct-download web threats (social engineering attacks)
- Private email attachment threats (exploitation and social-engineering attacks)

Public threats are sourced directly from attacking systems on the internet at the time of the test and can be considered 'live' attacks that were attacking members of the public at the time of the test run. Multiple versions of the same prevalent threats may be used in a single test run, but different domain names will always be used in each case. Private threats are generated in the lab according to threat intelligence gathered from a variety of sources and can be considered as similar to more targeted attacks that are in common use at the time of the test run.

All threats are identified, collected and analyzed independently of security vendors directly or indirectly involved in the test. The full threat sample selection will be confirmed by the Threat Verification Network as being malicious. False positive samples will be popular and non-malicious website URLs as well as applications downloaded directly from their source websites where possible.

4. Participation

SE Labs doesn't charge for testing so there are no contracts to be considered in this section. Details of post-testing consultancy can be discussed with SE Labs directly.

Voluntary participation rates have typically tracked the following profile.

- Business Tests : ~75% Voluntary, licenses provided for testing, post-test consulting services utilized.

- Consumer Tests : ~50% Voluntary, licenses provided for testing, post-test consulting services utilized with the others either neutral or unaware of the test itself.

The general process for participation in this, or any, test with SE Labs follows.

- SE Labs or Vendor approach one another in person, by email or on the phone.
- Both parties will then:
    - Discuss the desired testing methodology and configuration
    - SE Labs reserves the right that they may or may not test the product
    - Dispute processes precede
    - Report publication

Please contact us at info@SELabs.uk. We will be happy to arrange a phone call to discuss our methodology and the suitability of your product for inclusion.

Opt-Out Policy : If any Vendor can supply sufficient reason as to why SE Labs should not include their product in an upcoming or on-going test, SE Labs will honor that request provided that there is agreement on the facts. As an example, an acceptable Opt-Out request would include the documented and pending release of a substantially new version of Vendor software which would render the testing the previous version meaningless.

Conflict of Interest Disclosure : No known conflicts of interest exist at this time. Post-test consultancy services are available to all participants for a fee.

Funding : Products we consider of key interest to our readers are included in our tests at no cost to the vendor. Post-test consultancy services, rights to use our logos, and re-promote our results are made available to all participants, subject to consultancy or licensing fees.

5. Environment

Physical Configuration : The Target Systems are identical Windows PCs specified as below. Each system has unrestricted internet access and is isolated from other Target Systems using Virtual Local Area Networks (VLANs). Each system runs Windows 10 (64-bit), updated with security patches available up to January 2020 Update (version 1909). The general Target System specification includes Windows PCs with an Intel Core i3-9100 3.6GHz processor (4-Core, 6MB cache up to 4.2GHz), 4GB RAM, and a 2.5inch SATA 120GB SSD.

Popular but vulnerable third-party applications installed include Adobe Flash Player, Adobe Reader, Apple QuickTime and Oracle Java (32-bit). If a security product requires an updated file from Microsoft the tester will install the necessary file. A web session replay system will be used when exposing systems to web-based threats. This provides an accurate simulation of a live internet connection and allows each product to experience exactly the same threat. All products have real-time and unrestricted access to the internet.

Sample Relevance (SE Labs criteria defining legitimate sample selection) - Non-malicious website URLs and application files are used to check for false positive detection. The number of these URLs and files will match the number of malware samples used. Candidates for legitimate sample testing include newly released applications, ranging from free software to the latest commercial releases. Potentially unwanted programs, which are not clearly malicious but that exhibit dubious privacy policies and behaviors, will be excluded from the test.

Curation Process: Malicious URLs and legitimate applications and URLs were independently located and verified by SE Labs. Targeted attacks were selected and verified by SE Labs. They were created and managed by Metasploit Framework Edition using default settings. The choice of exploits was advised by public information about ongoing attacks. One notable source was the 2016 Data Breach Investigations Report from Verizon.

Details regarding Test Threat Selection and Management follow.

- **Sample numbers and sources** - The Target Systems will be exposed to a selection of threats. These are weighted heavily (~75 per cent) towards public web-based threats. A smaller set of the samples will include public threats attached to emails and private, targeted attacks delivered by web exploitation or as email attachments. There may also be some threats found via alternative routes, such as internet messaging (IM) or peer to peer (P2P) networks.

- **Sample verification** - Threats will be verified using Vulnerable Target Systems. Threat verification occurs throughout the test period, with live public threats being used shortly after they are verified as being effective against the Vulnerable Target Systems on the Threat Verification Network. In cases where a threat is initially verified to be effective, but which is found not to be effective during testing (e.g. its C&C server becomes unavailable), the threat sample will be excluded from the test results of each product.

- **Attack stage** - Threats will be introduced to the system in as realistic a method as possible. This means that threats found as email attachments will be sent to target systems in the same way – as attachments to email messages. Web-based threats are downloaded directly from their original sources. These downloads occur through a proxy system that includes a session replay service to ensure consistency. Public threats that run on the Target System are allowed 10 minutes to exhibit autonomous malicious behavior. This may include initiating connections to systems on the internet or making changes to the system to establish persistence.

Distribution of Test Data: Malicious and legitimate data will be provided to partner organizations once the full test is complete. SE Labs does not share data on one partner with other partners. We do not currently partner with organizations that do not engage in our testing. Any security vendor that has their product tested may request hashes of their missed samples.

6. Schedule

Start Date Range: Test configuration is scheduled to begin on 20th March, 2020 and the Test commencement is forecast for 30th March, 2020. Participant configuration is anticipated to take place during the intervening two-week period.

Test Duration and Calculated End Date: The final Test Report is anticipated during the week of 22nd June 2020.

Milestones : Interim schedule milestones are listed below.

Sample Schedule Summary for Test Project

Index	Test Activity	Start Date Range	Dependencies
1	*Test Commencement*	*30th March, 2020*	
2	*Confirm Vendor Configuration Feedback*	*20th March –27th March, 2020*	
3	*Milestone 1 – Preliminary Results*	*TBD*	*(1), (2)*
4	*Milestone 2 – Test Report First Edition – End of*	*03rd June 2020*	*(3)*

	Testing Period		
5	*Feedback and Dispute Resolution Time – Retests as Needed*	*5th June – 19th June 2020*	*(4)*
6	*Milestone 3 – Issue Final Report – End Date for Test*	*22nd June – 26th June, 2020*	*(5)*

Communications: All Participants will be notified when the schedule wanders by four week or more.

Risks and Risk Management: No additional risks are known at this time.

7. Control Procedures

Connectivity Validation: Automatic submission of data to vendors is disabled where possible unless this reduces the immediate effectiveness of the product. A means for confirming whether a Product's cloud connectivity or other features are functioning can be provided by the Vendor.

Logging: Products run with the default settings. Additional logging may be enabled if requested by the vendor of the product in question. Vendors of business software are invited to make configuration recommendations.

Updates: All products are updated fully using the latest definitions, patches and any other available updates. These updates are made immediately prior to each exposure to a threat or legitimate application. Products may be upgraded to the latest version, if the version changes during the test period.

8. Dependencies

Participant and Test Subject Vendors Required Actions: Vendors may contact SE Labs for inclusion, exclusion or to respond to an invitation, either accepting or declining.

9. Scoring Process

The following occurrences during the attack stage will be recorded and all contribute to the product effectiveness measure.

- The point of detection (e.g. before/after execution).
- Detection categorization, where possible (e.g. URL reputation, signature or heuristics).
- Details of the threat, as reported by the product (e.g. threat name; attack type).
- Unsuccessful detection of threats.
- Legitimate files allowed to run without problems.
- Legitimate files acted on in non-optimal ways (e.g. accusations of malicious behaviour; blocking of installation) and at what stage (e.g. before/after execution).

- User alerts/interaction prompts such as:
    - Pop-up information messages (even if not interactive).
    - Requests for action (take default option or follow testing policy of 'naïve user' if no default provided).
    - Default suggestions.
    - Time-out details (e.g. record if an alert/request for action disappears/takes a default action after n seconds of no user response).
- When an initial attack or attacker succeeds in downloading further malicious files, such downloads will be recorded along with the product's behavior (if any). This additional data will be presented alongside the main results, clearly labeled as representing a SE Labs Endpoint Anti-Malware Testing Methodology second attack. For statistical purposes, detection rates of these files will not be automatically added to the overall totals for each product (although doing so after the event will be possible).
- Any anomalies (e.g. strange or inconsistent behavior by the product).

Measuring Product Effectiveness: Each Target System is monitored to detect a product's ability to detect, block or neutralize threats that are allowed to execute. Third-party software records each Target System's state before, during and after the threat exposure stage. These results show the extent of an attacker's interaction with the target and the level of remediation provided by the product being tested. The same level of monitoring occurs when introducing legitimate URLs and files when assessing false positive detection rates.

Awards: SE Labs provides badges such as AAA, AA, and others based on the Test Scoring results. Partners can use the SE Labs awards logos for marketing purposes.

10. Dispute Process

The dispute process runs for two weeks from the end of the test. Please see Section 6 covering the Test Schedule for additional details and timing. The general Dispute Process works as follows.

1. Results are provided with hash values associated with any sub-optimal results.
2. Vendor responds within two weeks, arguing why some results are wrong.
3. Tester replies, accepting or denying the dispute.

Although discussions will follow, ultimately the data speaks for itself. The most closely argued disputes are over PUAs in false positive testing, where a file might be legitimate, do no harm, but does appear to contain a library with the potential for unwanted behavior. In such cases a vendor might detect it and argue that it's not an FP. In such cases we always remove the application from the test set.

Evidence Sharing Policy

Tested vendors will receive a results spreadsheet containing the hash values of the samples included the test and the rating achieved for each test case. Upon receiving a dispute on a test case SE Labs will provide screenshots of the analysis the tester performed during the exposure to a sample. Information from industry-accepted tools are used as evidence gathering such as the Sysinsternals suite, Wireshark, memory dumps (if available).The full output from the logs is not provided, only a subset of the information enough to validate the result will be shown to the vendor.

11. Attestations

I understand and agree that I am submitting this Test Plan, and the following Attestations, on behalf of the entity listed below, and I represent and warrant that I have authority to bind such entity to these Attestations. All references to “I” or “me” or similar language refer to such entity. I represent and warrant that the following Attestations are true, to the best of my knowledge and belief, and each of the following commitments will be upheld to the best of my ability.

1. I will provide public notification on the AMTSO website covering my obligation for notification of a Public Test, regardless of whether a potential Participant is in actual receipt of such notification prior to the Commencement Date of a Test. (Section 1, Section 4, Section 6)

2. All products included in this Test will be analyzed fairly and equally. (Section 2, Section 3, Section 5)

3. I will disclose any anticipated or known imbalance or inequity in the Test design to all Participants in the Test. (Section 2, Section 3)

4. Although I may charge for participation in a Test, I will not charge any additional fees for a vendor to be a test subject under the Standards. (Section 4)

5. I will disclose any material conflicts of interest or other information that could materially impact the reliability of the Test. (Section 4)

6. I will disclose how the Test was funded. (Section 4)

I hereby affirm, to the best of my knowledge and belief that this Test Plan complies with the AMTSO Testing Standards, as of the date hereof.

Signature: /s/ Simon Edwards

Name: Simon Edwards

Test Lab: SE Labs

AMTSO Test ID: AMTSO-LS1-TP023