



MRG Effitas 360 Degree Assessment & Certification

Q1 2016

Contents

Introduction.....	3
Executive Summary	3
Certification.....	4
The Purpose of this Report.....	5
Tests Employed	6
Security Applications Tested.....	8
Test Results	9
QI 2016 In the Wild 360 / Full Spectrum Test Results.....	9
QI 2016 Time to Detect & Remediate Test.....	10
Appendix I	14
Methodology Used in the 360 Assessment & Certification Programme QI 2016.....	14

Introduction

MRG Effitas has a core focus on efficacy assessments in the anti-financial fraud space; however, we publish more traditional “Real World” detection tests as well. An example of such a test is our “Time to Detect Assessment Q4 2013” (Project 37).

This assessment measured security products’ ability to protect an endpoint from a live infection, but also, where a system was compromised, it measured the time taken to detect the infection and remediate the system. The time-to-detect-and-remediate component relied on each security product being manually forced to conduct a scan every thirty minutes over a twenty-four hour period.

For 2014, it was decided that a new approach was needed as the methodology applied in previous tests did not reflect how a security product would be used on an endpoint in the real world. In practice, many security applications will only detect an infection during a reboot/startup or if a scheduled scan has been set by default.

For this assessment, time-to-detect will employ a methodology based on the infected endpoint being rebooted once during a 24 hour period.

The methodology employed in this test maps more closely to real world use and although it may not be a 100% accurate model of how an “average” system is used, it gives a more realistic assessment of a security product’s ability to detect and remediate an infected endpoint.

This Programme is called a “360 Assessment” since it deals with the full spectrum of malware instead of just financial malware. In the 360 Assessments, trojans, backdoors, ransomware, PUAs, financial malware and “other” malware are used.

Executive Summary

This Certification Programme is designed to serve as a reflection of product efficacy based on what we have previously termed “metrics that matter”.

In many of our previous tests, particularly those that have focused on financial malware, we started with the assumption that the endpoint has already been compromised. Being the world’s largest supplier of early-life malicious binaries and malicious URLs, and from our own simulator development, we know that all endpoints can be infected, regardless of the security solution employed.

For us, a product’s ability to block initial infection (although critical in most cases) is not the only metric that matters. One also needs to measure the time taken for the security product to detect malware on a system and remediate it.

When conducting these tests, we tried to simulate normal user behaviour. We are aware that a “Real World” test cannot be conducted by a team of professionals inside a lab because we understand how certain types of malware work, how malware attacks and how such attacks could be prevented. Simulating normal user behaviour means that we paid special attention to all alerts given by security applications. A pass was given only when alerts were straightforward and clearly suggested that malicious action should be blocked.

We tested a group of internet security suites and complementary security applications. With internet security suites and complementary applications, it is very important to note that the best choice for an average user is to keep things very simple and for the product not to present many popup alerts or questions.

Out of sixteen products we tested, five managed to meet the specification to attain our Q1 2016 360 certification award, these being **Kaspersky Internet Security, SurfRight HitmanPro, Symantec Norton Security, Webroot SecureAnywhere Internet Security Plus** and **Watchdog Anti-Malware**.

All other security applications failed the test in that they were unable to detect the malware and/or remediate the system even after the end of a twenty-four hour period.

Certification

In order to attain a quarterly MRG Effitas 360 Degree certification award, a security application must either protect the system from initial infection (a level 1 pass) or detect any missed malware and fully remediate the system before or on the first user initiated reboot (a level 2 pass). Applications that meet this specification will be given certification for that quarter.

Under the MRG Effitas 360 Degree Assessment & Certification , the following products were certified for Q1 2016:

Certified: Kaspersky Internet Security, Norton Internet Security, Webroot SecureAnywhere Internet Security Plus, SurfRight HitmanPro and Watchdog Anti-Malware



The Purpose of this Report

Since its inception in 2009, MRG Effitas has strived to differentiate itself from traditional testing houses by having its primary focus on providing “*efficacy assessments*” and not just performing “*tests*”.

Traditionally, testing of security software has centred about measuring product ability to detect malware. Testing has evolved rapidly over the last two to three years, as most labs, under the guidance of AMTSO (of which MRG Effitas is a member) strived to conduct “Real World” testing.

Although there is no absolute definition of this kind of testing, loosely speaking, it involves the introduction of malware to an endpoint through a realistic vector, such as a browser or USB memory stick. Real World testing mostly involves “dynamic testing” (i.e. the malware is executed and then the ability of the security product to block the malware is measured).

Several testing labs also conduct “System Rescue” tests. These assess a security product’s ability to remediate a pre-infected endpoint.

Whilst both types of tests are useful and yield valid and meaningful data, MRG Effitas wanted to merge these tests and also take one step further by measuring the time security products take to detect infections and remediate the endpoint.

To make testing more realistic to real world scenarios, no manual scanning was conducted; instead, the system was rebooted once a day (exactly 12 hours after the system was compromised) , thereby giving security applications the opportunity to detect infections on restart.

As we have stated in our previous test reports, all malware has one primary objective, and that is to make money for the cybercriminals.

Measuring initial detection rates and also the time taken to detect active malware is important, particularly in today’s threat landscape with the mix of malware that is prevalent.

As we have repeated in our previous financial malware test reports, the longer a cybercriminal can have their malware on a system, the greater the opportunity for them to be able to capture private user information including banking passwords and social media credentials, etc.

There has been an increase in the prevalence of ransomware, such as “CryptoLocker”, which, once active on the system, holds the user at ransom to decrypt system data or unlock the system in some other way (interestingly, the most common way CryptoLocker is installed on an endpoint is via Zeus infections).

For these types of malware, it is initial detection that is of the greatest importance, since the vast majority of security solutions will be unable to rescue an encrypted or locked system. (In other internal tests, we have found that Webroot SecureAnywhere was in fact able to undo the encryption performed by some ransomware.)

In providing these quarterly certifications, the MRG Effitas 360 Assessment & Certification Programme is the *de facto* standard by which security vendors, financial institutions and other corporations can attain the most rigorous and accurate determination of a product’s efficacy against the full spectrum of malware that is prevalent during the period.

Tests Employed

In this assessment (Q1 2016), we ran the following tests:

In the Wild 360 / Full Spectrum Test

Approximately 50% of the malicious URLs used in this test were compromised legitimate websites which served malware. We believe that such URLs pose the greatest danger to users as this is the place where they expect to get infected the least. 10% of the URLs pose as fake porn websites serving visitors with various types of malware. Remaining 40% of the URLs come from our regular honeypots or, in case of ransomware and financial malware in particular, we used URLs from newly discovered distribution sites.

Malware delivered by URLs used in this test can be considered as Zero Day in the true meaning of that word, this posed a great challenge to all participants as new variants samples such as Locky (Ransomware) TeslaCrypt (Ransomware), Dridex (Banking Trojan) and many others caused most damage.

It is our opinion that Ransomware currently poses the greatest threat to users, for this reason we choose use more URLs serving this threat than before.

Because of the wide spectrum of malware used in this project and the freshness of the samples, we used a smaller set than usual.

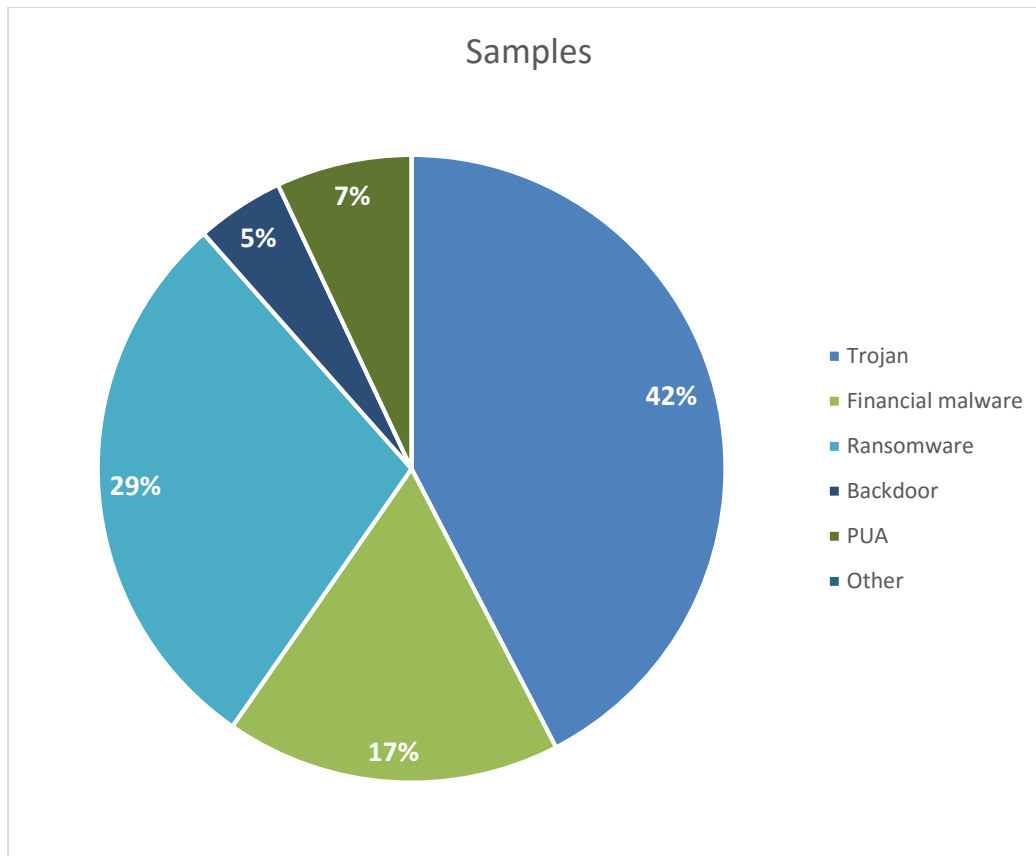
As a part of this project we used a set of PUA samples. PUA's (Potentially Unwanted Applications) are a gray area subject but one thing is clear, nobody wants them running on their system. For this reason, results for the PUA test are shown separately and are not a part of the certification process.

Applications that didn't protect the system from file encrypting ransomware cannot be certified because they could not remediate the threat; decrypt files.

Our testing environment supports the use of VM aware malware, this is the reason why we were able to use more sophisticated threats which wouldn't run on Virtual Machines.

10% of the threats used in this test were introduced to the system via USB flash memory sticks. These samples came originally from live URLs, but inside archives.

Testing was conducted as per the methodology detailed in Appendix I. In total, 243 live ITW samples were used. The stimulus load was comprised of the following: 103 trojans, 11 backdoors, 42 financial malware samples, 70 ransomware samples, and 17 PUAs.



Time to Detect and Remediate Testing was conducted as per the methodology detailed in Appendix I.

Security Applications Tested

(Last program build used in the project)

avast! Internet Security 2015 11.1.22.53

AVG Internet Security 16.6.1.7538

Avira Internet Security 14.0.16.208

Bitdefender Internet Security 2016 20.0.26.1418

ESET Smart Security 9 9.0.375.0

Kaspersky Internet Security 16.0.0.614 (d)

Malwarebytes Anti-Malware 2.2.1

McAfee Internet Security 14.0.7086

Microsoft Windows Defender

Panda Internet Security 16.1.2

SurfRight HitmanPro 3.7.14ⁱ

Symantec Norton Security 22.6.0.142

ThreatTrack VIPRE Internet Security 2016 9.0.1.4

Trend Micro Maximum Security 10.0.1186

Webroot SecureAnywhere with IdentityShield 9.0.7.46

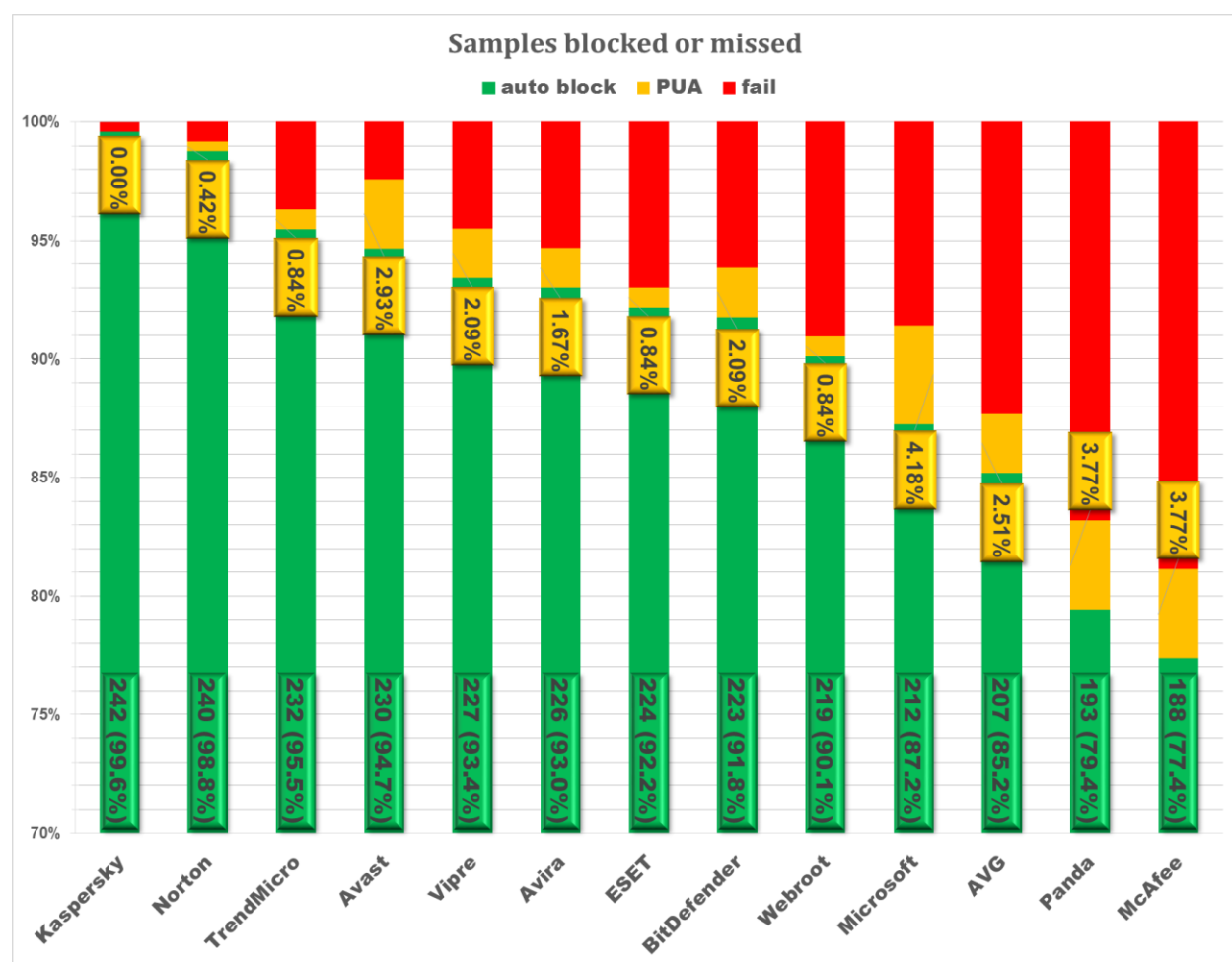
Watchdog Anti-Malware 2.20.2.613

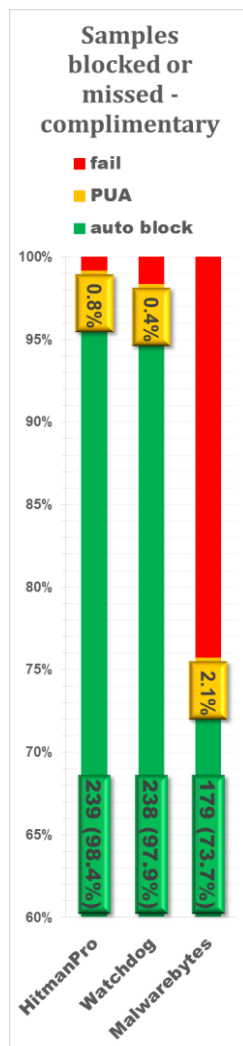
Test Results

The tables below show the results of testing under the MRG Effitas 360 Q1 Assessment Programme.

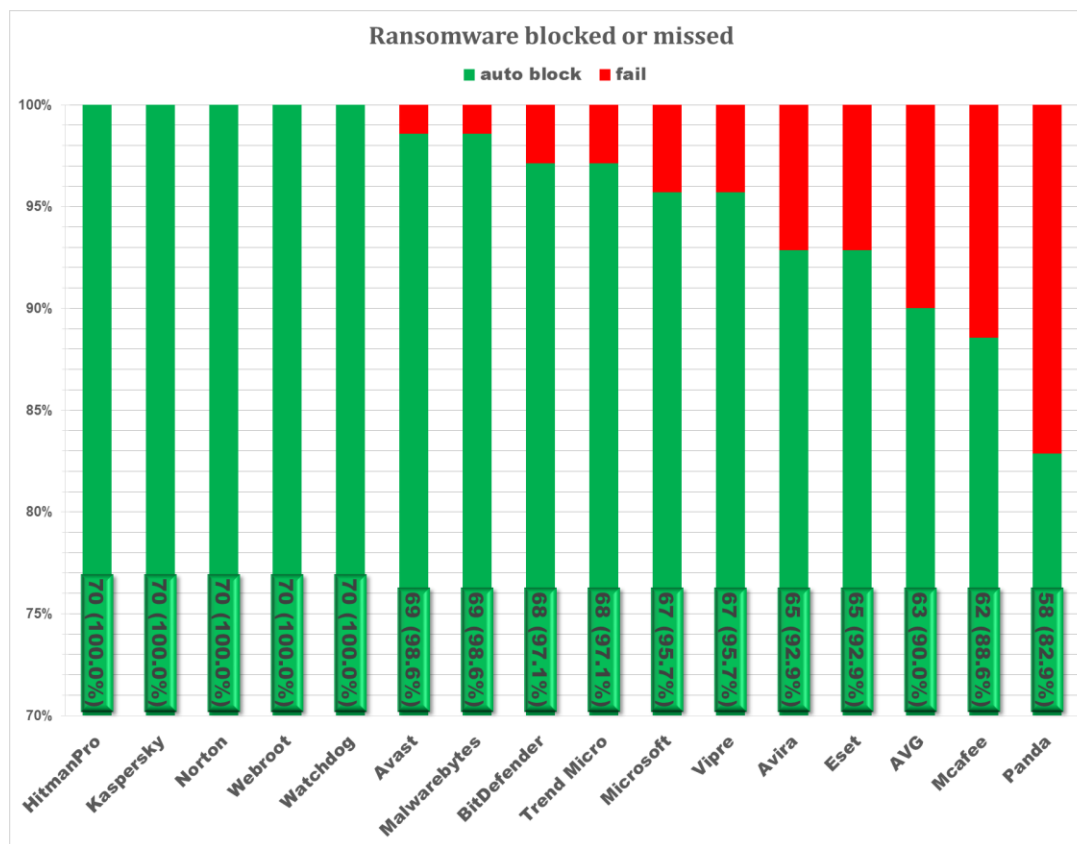
Q1 2016 In the Wild 360 / Full Spectrum Test Results

The table below shows the initial detection rates of the security products.

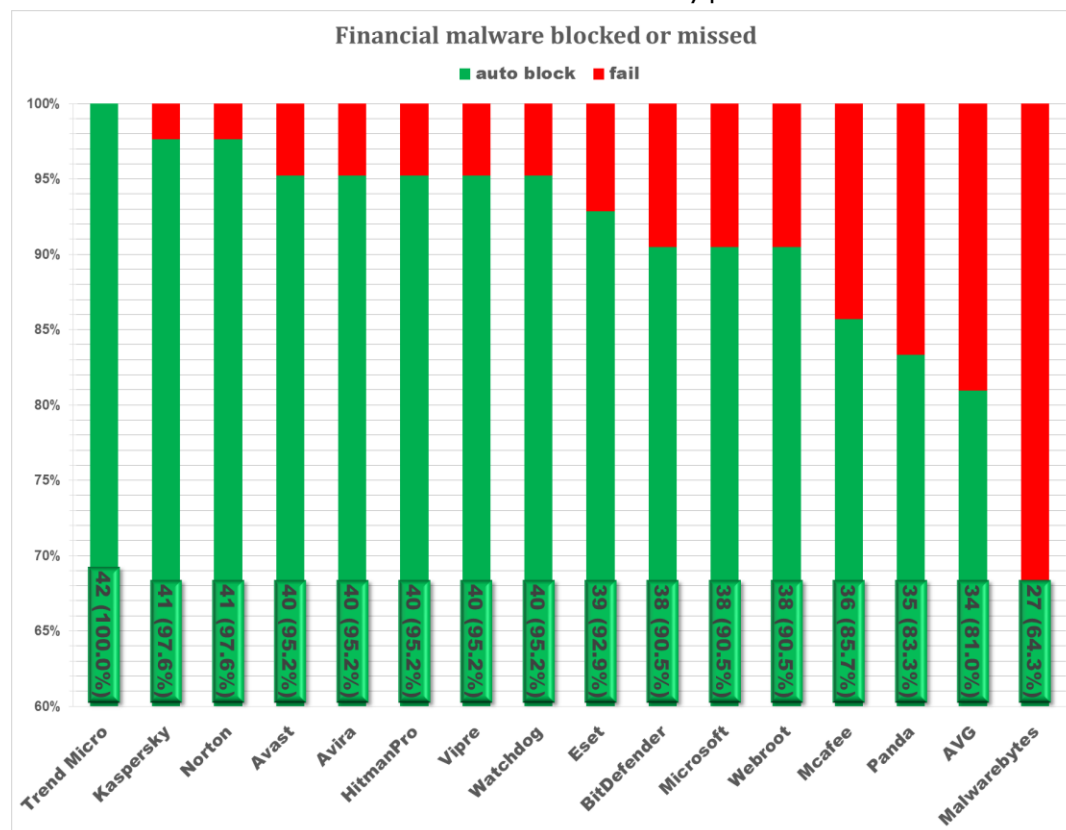




The table below shows the initial detection rates of the security products for ransomware.



The table below shows the initial detection rates of the security products for financial malware.



Q1 2016 Time to Detect & Remediate Test

Understanding Grade of Pass:

- **Level 1** = All threats detected on first exposure / system uncompromised.
- **Level 2** = All threats detected and neutralised / system remediated before or on the first user reboot

Kaspersky Internet Security

SurfRight HitmanPro

Symantec Norton Security

Webroot SecureAnywhere Internet Security Plus

Watchdog Anti-Malware

- **Failed** = Security product failed to detect all infections and remediate the system during the test procedure.

avast! Internet Security

AVG Internet Security

Avira Internet Security

Bitdefender Internet Security

ESET SmartSecurity

Malwarebytes Anti-Malware

McAfee Internet Security

Microsoft Security Essentials

Panda Internet Security

ThreatTrack VIPRE Internet Security

Trend Micro Premium Security

Appendix 1

Methodology Used in the 360 Assessment & Certification Programme Q1 2016

Methodology used in the assessment:

1. Windows 10 64 bit operating system was installed on a virtual machineⁱⁱ, all updates were applied and third party applications installed and updated according to our “Average Endpoint Specification”ⁱⁱⁱ
2. An image of the operating system was created.
3. A clone of the imaged systems was made for each of the security applications used in the test.
4. An individual security application was installed using default settings^{iv} on each of the systems created in 3 and then, where applicable, updated.
5. A clone of the system as at the end of 4 was created.
6. Each live URL test was conducted by:
 - a. Downloading a single malicious binary from its native URL using Microsoft Edge to the desktop, closing Microsoft Edge and then executing the binary.
 - b. The security application blocked the URL where the malicious binary was located.
 - c. The security application detected and blocked the malicious binary whilst it was being downloaded to the desktop.
 - d. The security application detected the malicious binary when it was executed according to the following criteria:

It identified the binary as being malicious and either automatically blocked it or postponed its execution and warned the user that the file was malicious and awaited user input.
7. The system under test was deemed to have been infected if:

The security application failed to detect or block the binary at any stage in 6 and allowed it to be executed.
8. Testing on infected systems continued for 24 hours. The system was rebooted once, exactly 12 hours after the system was compromised.
9. Remediation performance of an application was determined by manual inspection of the system in contrast to its pre-infected state and not by the logs and reports of the security application itself.^v
10. Testing was conducted with all systems having internet access.
11. Each individual test for each security application was conducted from a unique IP address.
12. All security applications were fully-functional unregistered versions or versions registered anonymously, with no connection to MRG Effitas.
13. All testing was conducted during Q1 2016.
14. As no user initiated scans were involved in this test, applications relied on various technologies to detect, block and remediate threats. Some of these technologies were: background scanning, startup scanning, scheduled scanning, system monitors, etc. A scheduled scan was used only if enabled by default.

ⁱ SurfRight HitmanPro was tested on-demand only and with startup scanning which is enabled by default

ⁱⁱ VM hardware spec is 4GB RAM & 2 core processor.

ⁱⁱⁱ AES includes Adobe Flash, Reader, Java, Microsoft Office 2010, Edge & VLC Player. All Microsoft components were fully updated; all third-party components were out of date by three months.

^{iv} During installation of the security application, if an option to detect PUAs was given, it was selected.

^v This is because in some instances, an application will claim to have removed an infection, but actually failed to do so and was still active on the system.