

360° Assessment & Certification

Q4 2022



MRG Effitas is a world-leading, independent IT security efficacy testing & assurance company. We are trusted by antimalware vendors across the world.

Contents

Contents	2
Introduction.....	3
The Purpose of this Report	4
Executive Summary	5
Tests Employed	6
Security Applications Tested	9
Malware sample types used to conduct the test	9
Test Results	10
360° Assessment Certification	13
360° Exploit Degree	14
360° Exploit Certification.....	15
360° Online Banking Degree.....	16
360° Online Banking Certification.....	18
360° Ransomware Degree.....	19
360° Ransomware Certification.....	22
Understanding the Grade of Pass.....	25
Appendix 1	28
Appendix 2	43

WEBSITE:
www.mrg-effitas.com

TEL:
+44 (0)20 3290 0447

EMAIL:
contact@mrg-effitas.com

TWITTER:
@mrgeffitas

Introduction

MRG Effitas is a world leader in independent IT research. Our core focus of antivirus efficacy assessment includes traditional “Real World” malware detection capabilities and financial fraud prevention.

The methodology employed in this test maps closely to Real World practice representing the valid threats endangering anyone using the Windows operating system. This evaluation aims to help users choose the most suitable enterprise security application for their needs.

This programme is called “360° Assessment & Certification” as it tests the capabilities of the participating security applications with a full spectrum of attack vectors. In-The-Wild Assessment, trojans, backdoors, spyware, financial malware, ransomware, and “other” malicious applications are all used. Alongside the traditional In-The-Wild (ITW) file-based attacks, our evaluation also contains scenarios where fileless cases and exploitation techniques, live botnets and financial malware simulators are applied.

In addition to malicious attacks, we also evaluate the practical accuracy of AV products, exposing them to potentially unwanted applications (PUA or Greyware) and clean files (FP) as well.

Our assessment also measures the footprint each security software has on a computer’s performance.



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 the testing of security software has been aimed at measuring a product’s ability to detect malware. Testing has evolved rapidly over the last couple of years with most labs, under the direction of AMTSO (of which MRG Effitas is a member) striving to conduct “Real World” testing, based on standardised guidelines. More information about the compliance status of this test can be found on the AMTSO website.

<https://www.amtso.org/tests/mrg-effitas-q4-2022-360-degree-assessment-and-certification/>

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 entry point, such as downloading the sample using a browser or getting it from a 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).

Whilst these types of tests are useful yielding valid and meaningful data, MRG Effitas wanted to merge standalone tests and go the extra mile by measuring the time security products take to detect infections and remediate the endpoint.

To make testing more akin to real world scenarios, no manual scanning was conducted. Instead, the system was retested 24 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, most malware has one primary objective, and that is to make money for the cybercriminals, thus making malware creation a lucrative business with its own unique economic models and traits.¹

Measuring initial detection rates and the time taken to detect active malware is important, particularly in today’s threat landscape with the mix of malware that is prevalent. The longer a cybercriminal can run their malware on a system, the greater the opportunity is for them to capture private user information, including banking logins and social media credentials, etc., or to encrypt user data.

For these types of malware, initial detection is of the utmost importance since the vast majority of security solutions will be unable to remediate the problem of an encrypted system.

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.

¹ For instance, in many ransomware campaigns, the criminals operate a 24x7 full blown customer help desk to help victims with buying Bitcoin, installing the TOR Browser etc., with a better „user experience” than traditional help desk services.

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”.

Based on decades of experience in IT security, our previous tests, and being one of the world’s largest suppliers of early-life malicious files and URLs, we know that all endpoints can and will be infected, regardless of the security solutions employed. The question is not ‘if’, but ‘when’ a malicious file hits the system.

A security product’s ability to block initial infection (although critical in most cases) is not the only metric that matters. Measuring the time taken to detect malicious files or actions is another metric that can also be crucial in evaluation. An additional key factor is the point in time when the fact of the infection and any associated malicious behaviour were detected.

When conducting these tests, we try 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 organised malware attacks are conducted, and how such attacks could be prevented. Simulating normal user behaviour means that we pay special attention to all alerts given by security applications. A pass is given only when alerts are straightforward, and clearly suggest that the malicious action should be blocked.

With this in mind, it is very important to note that the best choice for an average user is to keep things as simple as possible and not to overwhelm them with cryptic pop-ups, alerts or questions.

During our Q4 2022 360° Assessment, the following applications managed to attain our certifications.

360° Assessment Certification

- Avast Business Antivirus
- Avira Antivirus Pro
- Bitdefender Endpoint Security
- ESET Endpoint Security
- Malwarebytes Endpoint Protection
- Microsoft Windows Defender
- Symantec Endpoint Protection
- Trend Micro Security

360° Exploit Certification

- Avast Business Antivirus
- Bitdefender Endpoint Security
- ESET Endpoint Security
- Malwarebytes Endpoint Protection

360° Online Banking Certification

- ESET Endpoint Security
- Malwarebytes Endpoint Protection
- Symantec Endpoint Protection

360° Ransomware Certification

- Avast Business Antivirus
- Avira Antivirus Pro
- Bitdefender Endpoint Security
- ESET Endpoint Security
- Malwarebytes Endpoint Protection
- Microsoft Windows Defender
- Symantec Endpoint Protection
- Trend Micro Security

Tests Employed

In this assessment (Q4 2022), we ran the following tests.

In the Wild / Full Spectrum Test

Most of the malicious URLs used in this test were compromised legitimate websites, serving malware. We believe that such URLs pose the greatest danger to users, as this is the place where they least expect to get infected, and any URL based protection fails on them. Some URLs originate from our honeypots, or in case of ransomware and financial malware, 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 the phrase. This posed a significant challenge to the participant products.

~10% of the threats used in this test were introduced to the system via internal webmail sites. We have witnessed many SMBs being infected through internal webmail's and a lack of spam filtering. Downloading malware attachments from internal webmail sites bypass the URL blocking features of the products, and this happens in-the-wild.

During the In the Wild / Full Spectrum test, 360 live ITW samples were used. The stimulus load comprised the following: 106 trojans, 23 backdoors, 7 financial malware samples, 55 ransomware, 50 spyware, 78 malicious documents, 39 malicious script files and 2 others.

PUA / Adware Test

The PUA samples used in this test are deceptive, or potentially unwanted applications (PUA), that are not malicious, but are generally considered unsuitable for most home or business networks. They usually contain adware, install toolbars, or have other vague objectives. They may also contribute to consuming computing resources or network bandwidth. PUAs can be deceptive, harmful, hoax, show aggressive popups and mislead or scare the user. They may provide some unconventional ways of uninstalling the application, maybe retain some of their components on the device without the user's consent. We mainly use a filtered version of AppEsteem's feed, as they have developed a 'deceptor' benchmark as part of a cross-industry effort of many of the world's leading security companies which represents a minimum bar that all apps and services must meet to avoid being titled 'deceptive'.

AppEsteem, as a member of the AMTSO group, is dedicated to protecting consumers from harassment and objectionable material, enabling security companies to restrict access to these apps. MRG Effitas, as a member of the AMTSO group, is dedicated to the same cause.

In the PUA/Adware part we tested the products against 14 PUAs.

Exploit / Fileless Test

The main purpose of this test is to see how security products protect against a specific exploitation technique. To measure this, we developed test cases that simulate the corresponding exploit and post-exploitation techniques only.

Drive-by download exploits are the biggest threats for an enterprise environment, since no user interaction is needed to start the chain of infection on a victim machine. Outdated browsers and Office applications are widespread in enterprise environments, due to compatibility issues or the lack of proper updating and patch management.

We test the products' abilities to avoid any exposure to adversaries and to interrupt malicious payload delivery before they begin performing malicious actions. We focus explicitly on each product's ability to mitigate each attack technique. The results are not intended to evaluate the complete efficacy of the products, but rather the products' anti-exploit and anti-post-exploit features in isolation.

During this test we used 8 different exploitation techniques. A detailed description can be found in the 'Appendix'.

Real Botnet Test

A python based BYOB (Build Your Own Botnet) inspired tool was used as Botnet test. Its behaviour is parallel to any in-the-wild botnet. The main built-in feature is designed to steal credentials with two major components, a CnC Server, and a downloaded Portable Executable file.

Banking Simulator Test – Magecart v4

The Financial malware simulator used in Q4 2022 was the Magecart credit card-skimming attack. Magecart is the name of the collective who are targeting some eCommerce sites and ticketing companies such as Ticketmaster, British Airways, Newegg, Infowars etc. They use small JavaScript codes on the online store's checkout pages. These scripts pull personal and/or credit card data and send it to the attackers' servers.

Ransomware Simulator Test

To assess how the protection product manages ransomware, we created ransomware samples in-house, ensuring the security product could only rely on its behaviour scanning modules, without the help of possibly known signatures or community verdicts. During Q4 2022 we tested 4 ransomware simulator samples.

False Positive Ransomware Test

The False Positive test samples used are legitimate utilities with completely benign use cases. We use them to mimic malicious ransomware behaviour as closely as possible, to see how security applications react to them. In this quarter we used 3 FP-ransomware test cases.

False Positive Test

Perfect blocking of malicious content is only part of the story from a practical point of view for any decent AV product. In many cases all malware blocking is a result of a very aggressive filter which can block non-malicious, legitimate applications as well prohibiting everyday work by blocking legitimate, perhaps newly developed in-house software.

To test this feature, we pitched the security applications against completely clean, recently created applications.

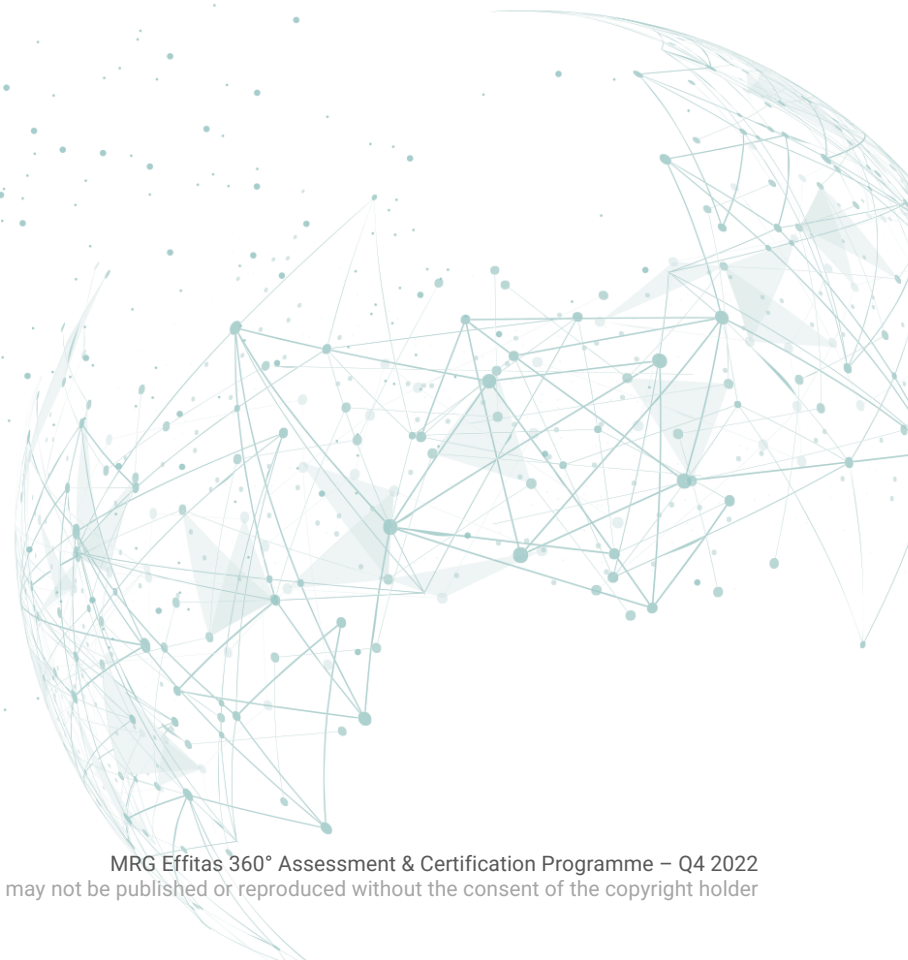
False positive assessment consisted of 500 clean and legitimate application samples. The selection has been focused on applications frequently found in enterprise environments (drivers, media editors, developer tools, etc.)

Performance Test

A security product's usefulness does not depend on protection abilities alone, but also on its resource footprint and its effect of the overall operating system performance.

To assess the products' influence on the operating system, we tested several performance factors on a physical machine and combined the results, based on a scoring approach. Detailed information can be found in the 'Appendix'.

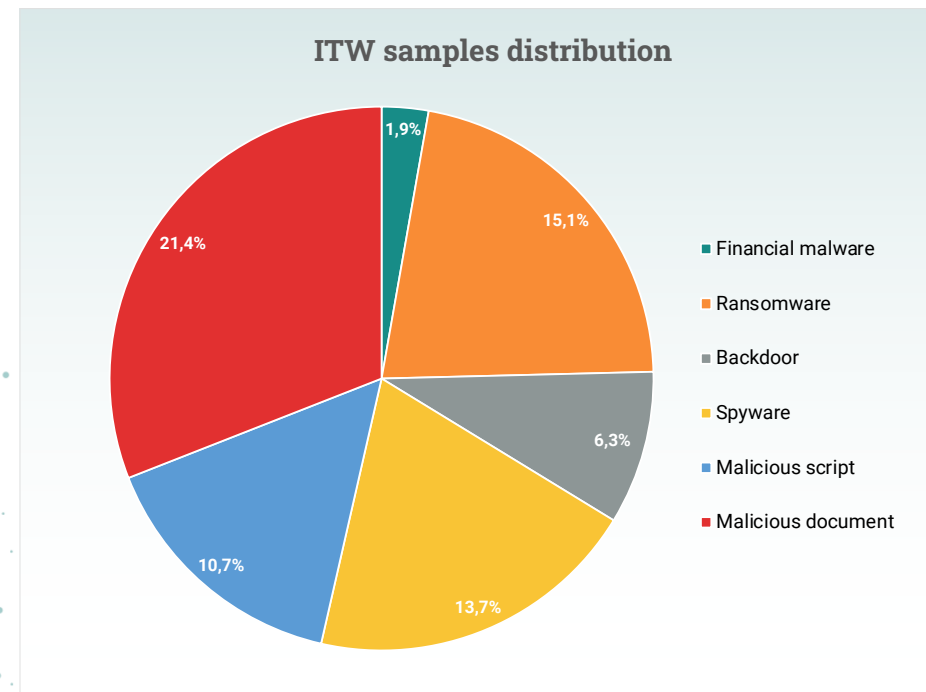
In every test case (except for the performance test), our testing environment supports the execution of VM-aware malware, this is why we are able to use more sophisticated threats which normally would not run on Virtual Machines.



Security Applications Tested

- Avast Business Antivirus 22.12.2733
- Avira Antivirus Pro 1.1.76.2
- Bitdefender Endpoint Security 7.8.1.244
- ESET Endpoint Security 9.0.2060.0
- Malwarebytes Endpoint Protection 1.2.0.1157
- Microsoft Windows Defender 4.18.2211.5
- Symantec Endpoint Protection 14.3.8268.5000
- Trend Micro Security 6.7.2151/14.2.2097

Malware sample types used to conduct the test

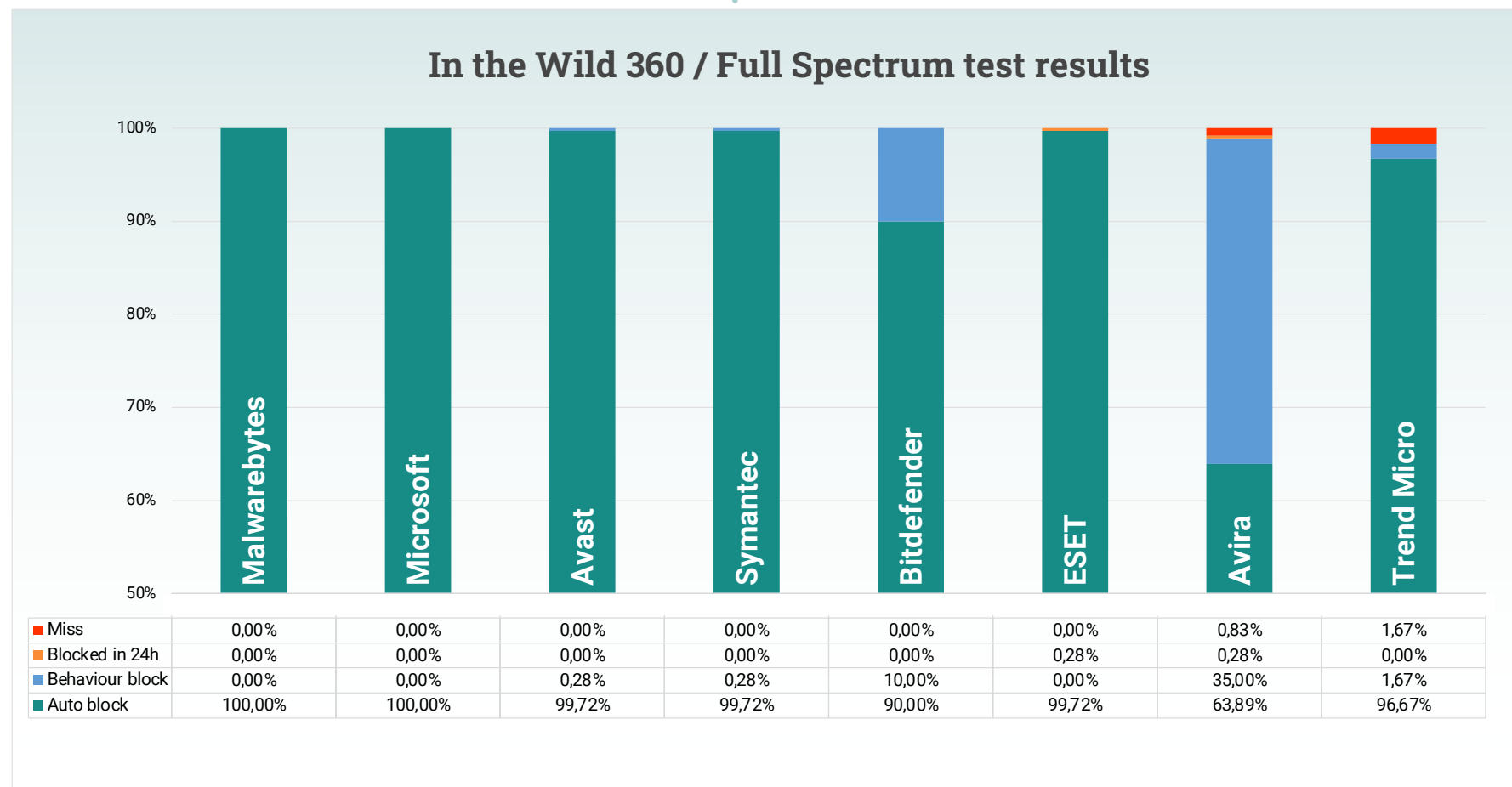


Test Results

The tables below show the results of testing under the MRG Effitas 360° Assessment Programme Q4 2022.

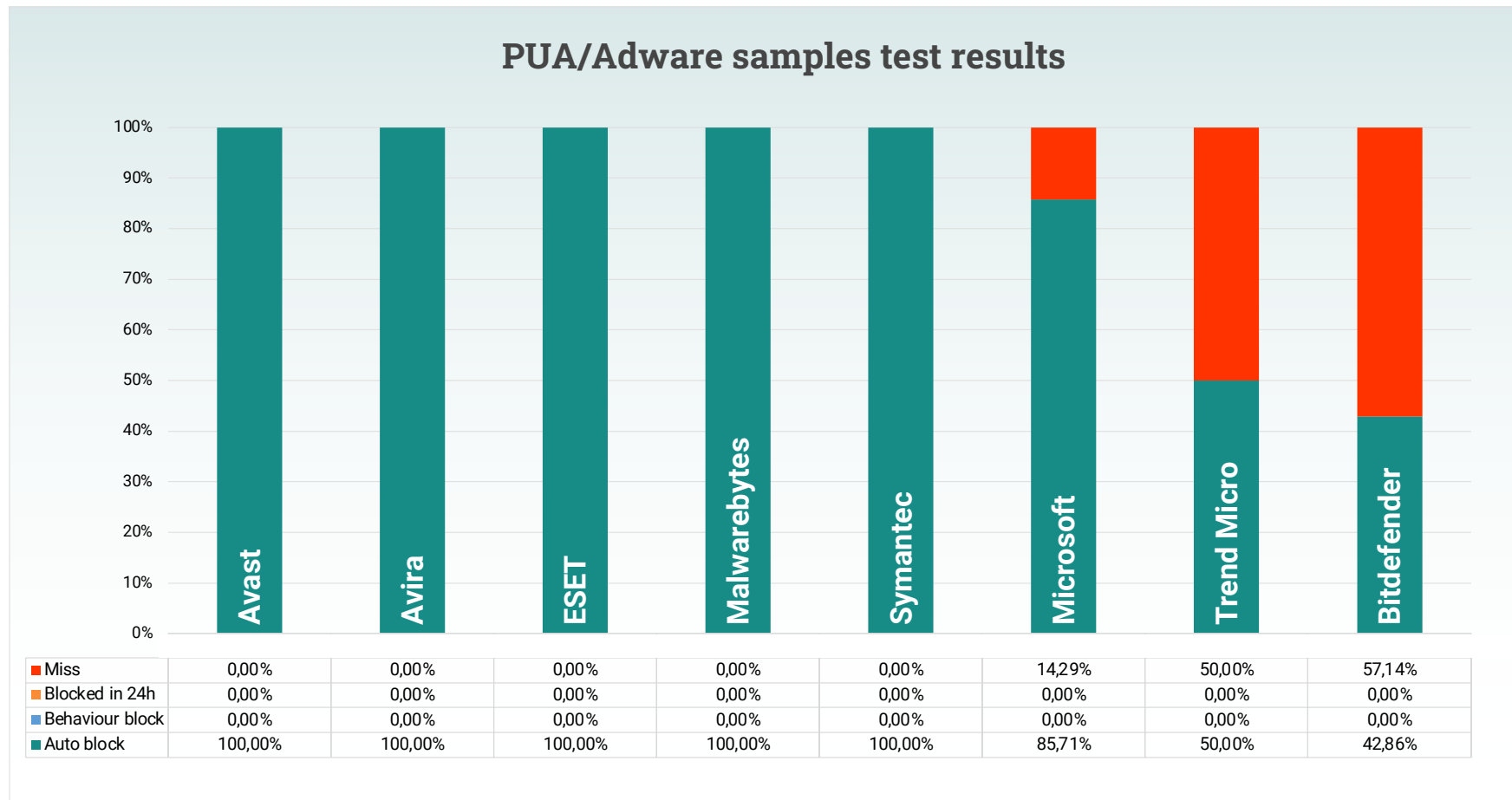
In the Wild / Full Spectrum test results

The table below shows the detection rates of the security products for 360 ITW samples. This table is sorted by smallest number of missed samples.



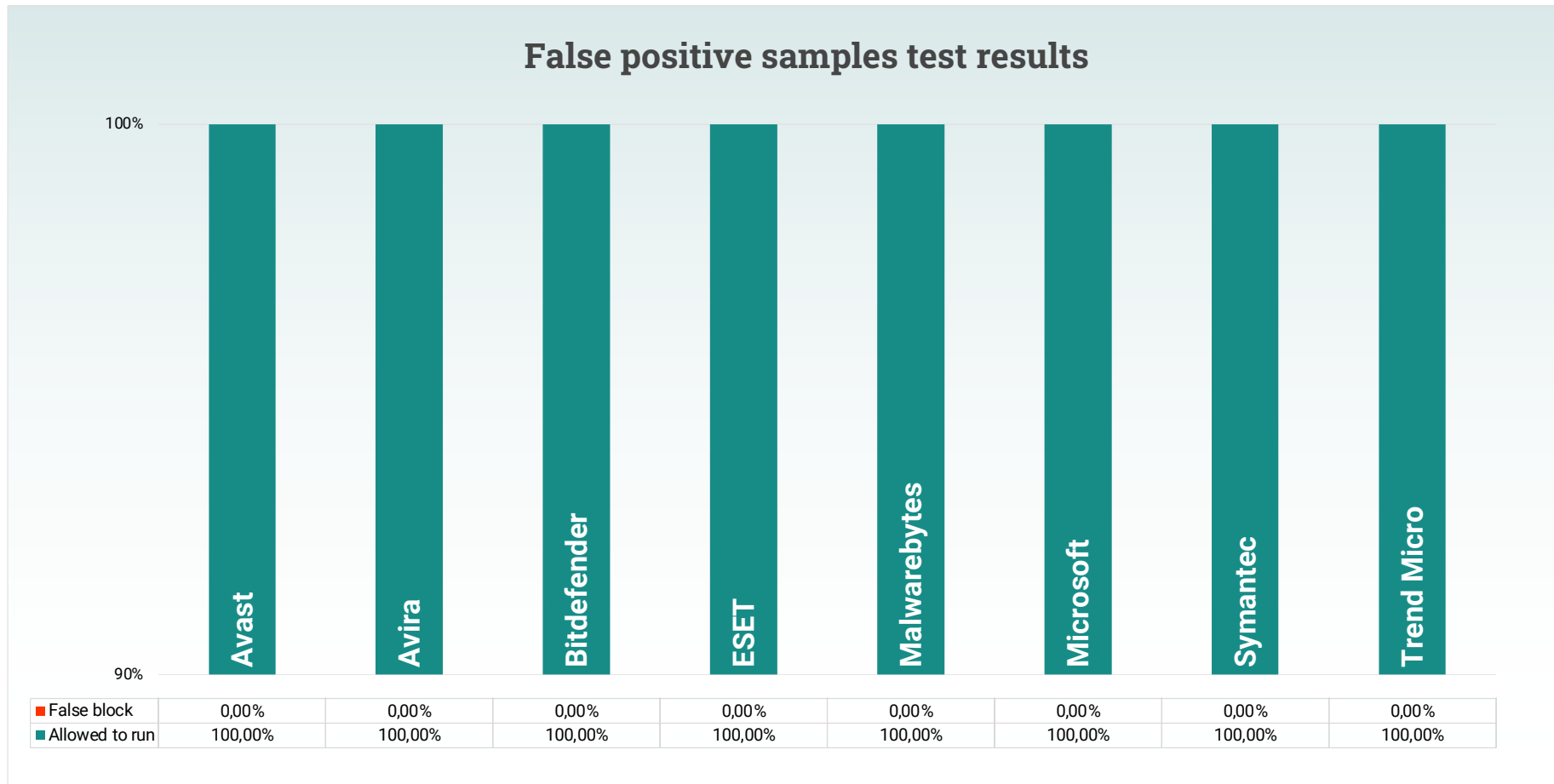
PUA/adware samples test results

The table below shows the detection rates of the security products for 14 PUA/Adware samples. This table is sorted by smallest number of missed samples.



False positive samples test results

The table below shows the initial detection rates of the security products for 500 false positive (clean) samples. This table is sorted by smallest number of false positive sample blocks.



360° Assessment Certification

To attain a quarterly MRG Effitas 360° Level 1 certification, a security application must completely protect the system from initial infection. This could be either by automatically blocking every ITW sample, or by blocking them based on their behaviour prior to any malicious actions. The product must also pass the Real Botnet test. (PUA, FP, Exploit/Fileless, Financial Malware Simulator, and Performance tests are not part of this certification.)

Level 2 certification is given if the application blocks or detects any initially missed malware in at least 98% of all cases on the 24-hour retest, while the initially missed test cases are less than 10%. If a ransomware/wiper successfully runs and the files are not available anymore, Level 2 certification is lost.

Under the MRG Effitas 360° Assessment & Certification, the following products were certified for Q4 2022.

Certified (Level 1)

- Bitdefender Endpoint Security
- Malwarebytes Endpoint Protection
- Symantec Endpoint Protection

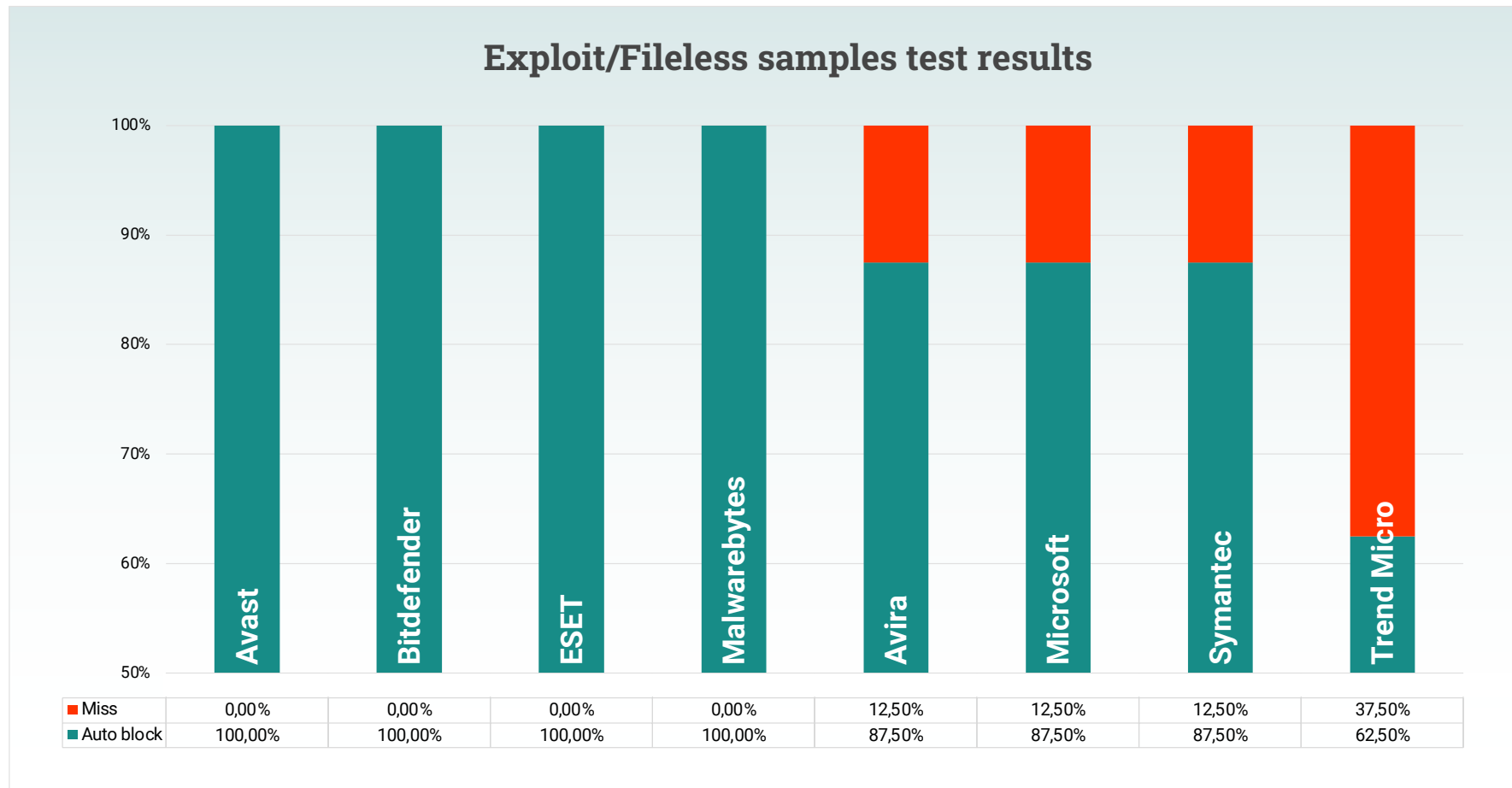
Certified (Level 2)

- Avast Business Antivirus
- Avira Antivirus Pro
- ESET Endpoint Security
- Microsoft Windows Defender
- Trend Micro Security



360° Exploit Degree

The table below shows the initial detection rates of the security products for 8 Exploit/Fileless test. This table is sorted by smallest number of missed attack vectors.



360° Exploit Certification

To attain a quarterly MRG Effitas 360° Exploit certification award, a security application must entirely protect the system from initial infection (autoblock, signature block, or behaviour block).

Under the MRG Effitas 360° Exploit Certification, the following products were certified for Q4 2022.

Certified

- Avast Business Antivirus
- Bitdefender Endpoint Security
- ESET Endpoint Security
- Malwarebytes Endpoint Protection

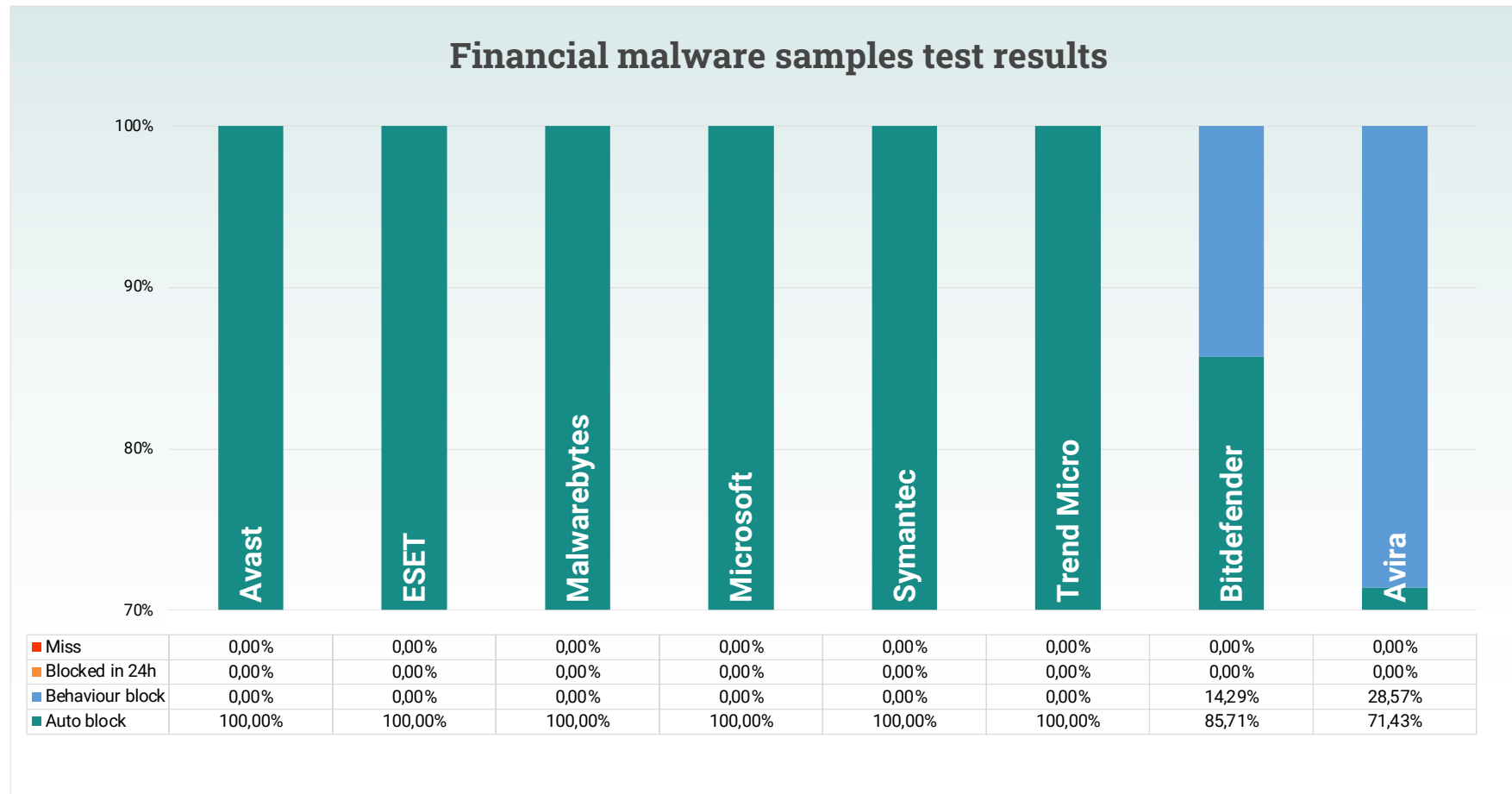


Q4 2022

360° Online Banking Degree

Financial malware samples test results

The table below shows the detection rates of the security products for 7 financial malware samples. This table is sorted by smallest number of missed samples.



Real Botnet test results

The table below shows the results of live Real Botnet test.

Real Botnet Test	
Product	Result
Avast Business Antivirus	✘
Avira Antivirus Pro	✘
Bitdefender Endpoint Security	✔
ESET Endpoint Security	✔
Malwarebytes Endpoint Protection	✔
Microsoft Windows Defender	✘
Symantec Endpoint Protection	✔
Trend Micro Security	✘
✔ The application prevented the malware from capturing login data	
✘ The application failed to prevent the malware from capturing login data	

Banking Simulator test results

The table shows the results of Banking Simulator test.

Banking Simulator Test	
Product	Result
Avast Business Antivirus	✘
Avira Antivirus Pro	✘
Bitdefender Endpoint Security	✘
ESET Endpoint Security	✔
Malwarebytes Endpoint Protection	✔
Microsoft Windows Defender	✘
Symantec Endpoint Protection	✔
Trend Micro Security	✘
✔ The application blocked the simulator	
✘ The application failed to block the simulator	

360° Online Banking Certification

To attain a quarterly MRG Effitas 360° Online Banking certification award, a security application must entirely protect the system from initial In-the-wild financial malware infection (autoblock or behaviour block) and the product must pass the Real Botnet and Banking simulator tests during the quarter.

Under the MRG Effitas 360° Online Banking Certification, the following products were certified for Q4 2022.

Certified

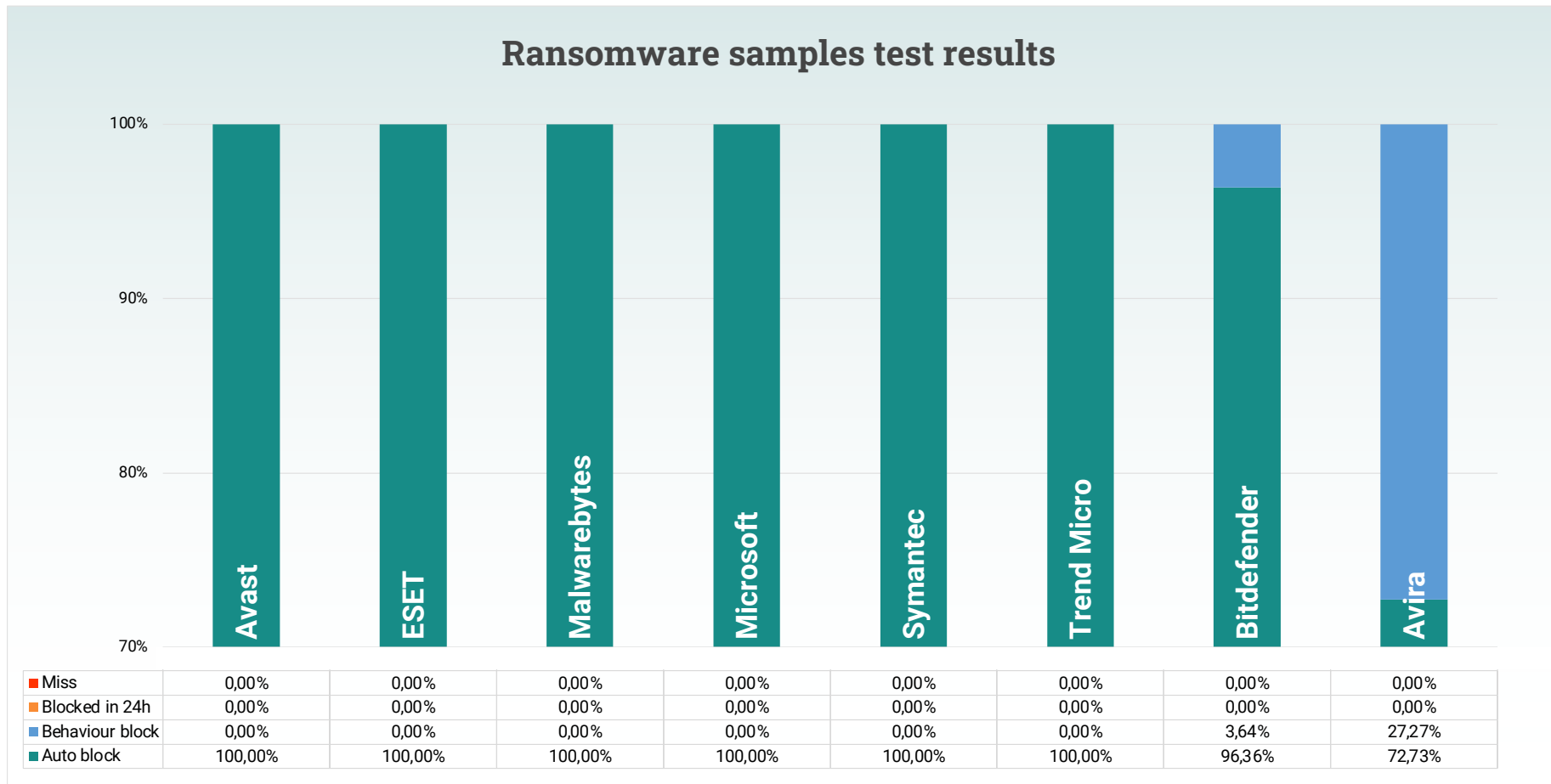
- ESET Endpoint Security
- Malwarebytes Endpoint Protection
- Symantec Endpoint Protection



360° Ransomware Degree

Ransomware samples test results

The table below shows the detection rates of the security products for 55 ransomware samples. This table is sorted by smallest number of missed samples.



Ransomware Simulator test results

The table below shows the detection rates of the security products for 4 Ransomware Simulator samples.

Ransomware Simulator test				
Product	Onyx (new file)	Onyx (rename)	Python (new file)	Python (rename)
Avast Business Antivirus	✓	✓	✓	✓
Avira Antivirus Pro	✓	✓	✓	✓
Bitdefender Endpoint Security	✓	✓	✓	✓
ESET Endpoint Security	✓	✓	✓	✓
Malwarebytes Endpoint Protection	✓	✓	✓	✓
Microsoft Windows Defender	✓	✓	✓	✓
Symantec Endpoint Protection	✓	✓	✓	✓
Trend Micro Security	✓	✓	✓	✓
<p>✓ The application blocked the simulator</p> <p>✗ The application failed to block the simulator</p>				

False Positive Ransomware test results

The table below shows the detection rates of the security products for 3 False Positive Ransomware samples.

False Positive Ransomware Test			
Product	encrypt0r	Kryptor	iSunshare SafeFile
Avast Business Antivirus	✓	✓	✓
Avira Antivirus Pro	✓	✓	✓
Bitdefender Endpoint Security	✓	✓	✓
ESET Endpoint Security	✓	✓	✓
Malwarebytes Endpoint Protection	✓	✓	✓
Microsoft Windows Defender	✓	✓	✓
Symantec Endpoint Protection	✓	✓	✓
Trend Micro Security	✓	✓	✓
<p>✓ The security application allowed the clean process to run</p> <p>✗ The security application falsely blocked the clean process</p>			

360° Ransomware Certification

In order to attain a quarterly MRG Effitas 360° Ransomware certification award, a security application must entirely protect the system from initial In-the-wild ransomware malware infection (autoblock or behaviour block) and pass the ransomware simulator and the false positive ransomware test during the quarter.

Under the MRG Effitas 360° Ransomware Certification, the following products were certified for Q4 2022.

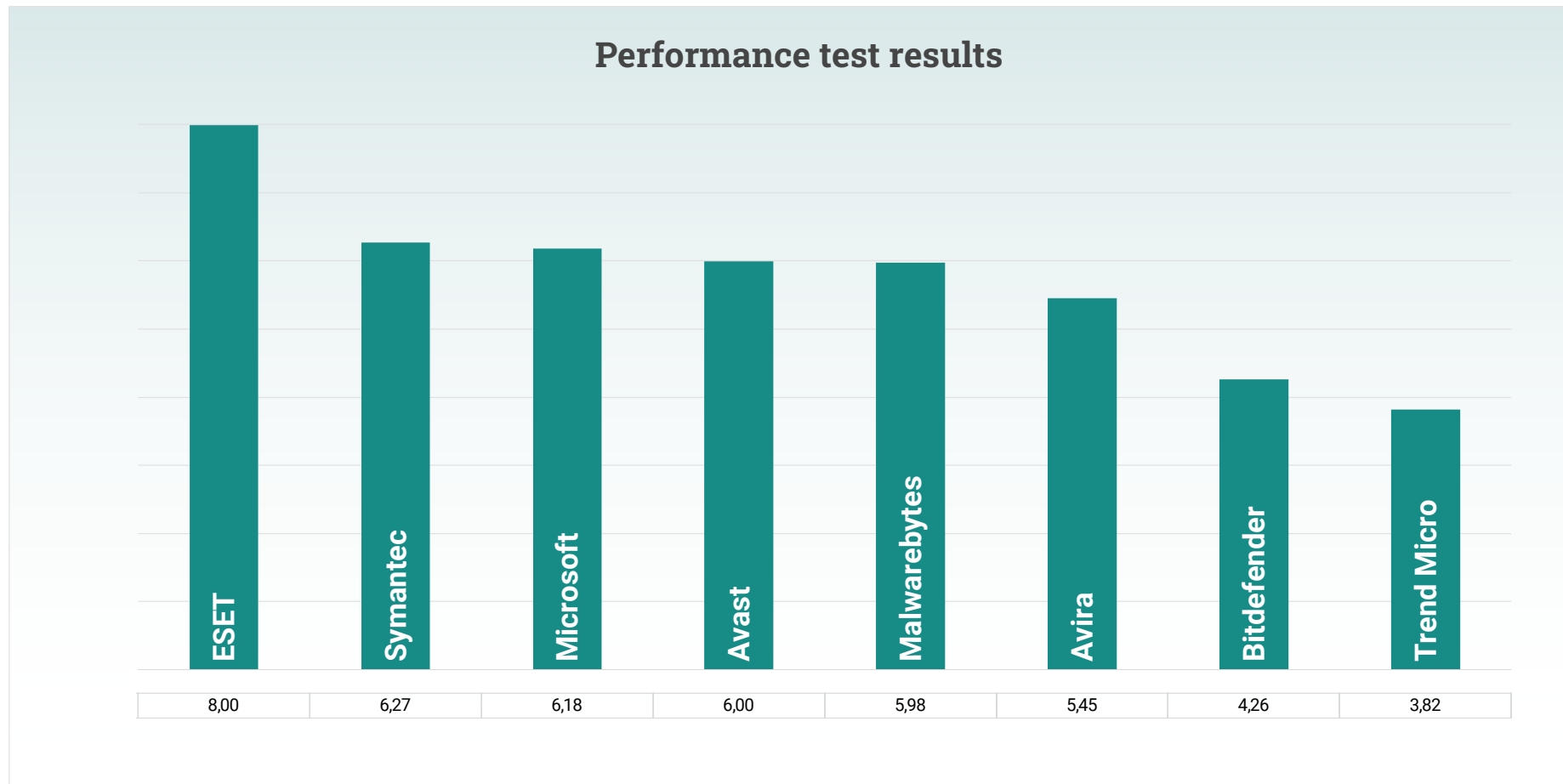
Certified

- Avast Business Antivirus
- Avira Antivirus Pro
- Bitdefender Endpoint Security
- ESET Endpoint Security
- Malwarebytes Endpoint Protection
- Microsoft Windows Defender
- Symantec Endpoint Protection
- Trend Micro Security



Performance and Footprint test results

This table is sorted from highest to lowest score where the highest score denotes the lowest impact on the system.



Scoring details can be found in the 'Appendix'.

Detailed results of the Performance test

The table below shows the detailed results of the performance test of the security products. This table is sorted alphabetically.

	Windows 10 Base	Avast	Avira	Bitdefender	ESET	Malwarebytes	Microsoft	Symantec	Trend Micro
Bootup time (s)	30,7	39,8	58,9	44,0	36,1	50,8	38,1	37,9	47,6
Security software size on disk (Mb)	n/a	1621,6	1057,6	1542,6	870,8	658,9	979,0	2320,2	958,3
Browser Operations (s)									
Website Open	2,7	3,1	4,0	7,7	3,7	2,6	2,0	2,7	5,7
File Download	10,3	11,5	11,7	12,1	10,9	12,1	10,1	13,6	16,0
File Operations (s)									
File Copy	2,4	12,7	2,0	2,1	2,4	2,1	2,8	2,0	2,4
File Compression	38,2	38,1	38,1	39,8	37,4	38,0	39,1	37,9	50,6
Archive Extraction	6,7	6,2	10,3	7,3	5,8	6,8	6,6	6,8	29,4
Office File Opening (s)									
Excel	6,3	3,4	3,8	4,6	3,3	4,8	3,5	3,4	5,3
Word	1,1	1,6	1,8	2,1	1,5	2,3	2,3	1,8	3,0

Detailed results of the Footprint test

	Windows 10 Base	Avast	Avira	Bitdefender	ESET	Malwarebytes	Microsoft	Symantec	Trend Micro
Security software update									
Time (s)	n/a	30,3	57,7	103,0	18,7	n/a	31,0	17,0	115,7
CPU usage (%)	n/a	31,4	36,6	30,8	42,9	n/a	34,6	38,4	19,5
Memory usage (Mb)	n/a	456,0	840,0	905,0	215,0	n/a	169,0	421,0	1287,0
Physical disk usage (%)	n/a	12,4	19,2	44,9	12 082.80	n/a	30,1	14,6	5,8
Network interface usage (B/s)	n/a	38 522.67	283 645.67	133 293.00	43 020.50	n/a	122 035.33	139 905.97	25 931.00
Security software scanning - C:\									
Time (s)	n/a	318,0	180,0	331,7	150,7	202,0	1 526.67	458,7	115,7
CPU usage (%)	n/a	19,3	81,2	37,5	29,9	33,3	85,0	22,9	19,5
Memory usage (Mb)	n/a	963,0	1177,0	1428,0	592,0	970,0	985,0	731,0	1287,0
Physical disk usage (%)	n/a	62,6	59,7	22,9	37,2	18,6	28,7	14,5	5,8
Network interface usage (B/s)	n/a	297 114.00	7 848.33	28 961.00	846,7	723,0	36 385.67	532,5	25 931.00

Understanding the Grade of Pass

360° Assessment - Level 1 certified

All threats detected on first exposure or via behaviour protection.

- Bitdefender Endpoint Security
- Malwarebytes Endpoint Protection
- Symantec Endpoint Protection

360° Assessment - Level 2 certified

At least 98% of the threats detected and neutralised / system remediated before or on the first rescan while the initially missed test cases are less than 10% and no ransomware was missed on first exposure.

- Avast Business Antivirus
- Avira Antivirus Pro
- ESET Endpoint Security
- Microsoft Windows Defender
- Trend Micro Security

360° Assessment - Not certified

Security product failed to detect at least 98% of the infections and remediate the system during the test procedure, or at least one ransomware was missed.

360° Exploit Degree - Certified

The application entirely protected the system from initial infection.

- Avast Business Antivirus
- Bitdefender Endpoint Security
- ESET Endpoint Security
- Malwarebytes Endpoint Protection

360° Exploit Degree - Not Certified

The application failed to protect the system from initial infection.

- Avira Antivirus Pro
- Microsoft Windows Defender
- Symantec Endpoint Protection
- Trend Micro Security

360° Online Banking Degree - Certified

The application entirely protected the system from initial In-the-wild financial malware infection and passed the Botnet and Financial malware simulator test.

- ESET Endpoint Security
- Malwarebytes Endpoint Protection
- Symantec Endpoint Protection

360° Online Banking Degree - Not Certified

The application failed to protect the system from initial In-the-wild financial malware infection, or it has not passed the Botnet or Financial malware simulator test.

- Avast Business Antivirus
- Avira Antivirus Pro
- Bitdefender Endpoint Security
- Microsoft Windows Defender
- Trend Micro Security

360° Ransomware Degree - Certified

The application must entirely protect the system from initial infection and pass the ransomware simulator and the false positive ransomware test.

- Avast Business Antivirus
- Avira Antivirus Pro
- Bitdefender Endpoint Security
- ESET Endpoint Security
- Malwarebytes Endpoint Protection
- Microsoft Windows Defender
- Symantec Endpoint Protection
- Trend Micro Security

360° Ransomware Degree - Not Certified

The application failed to protect the system from initial infection or pass the ransomware simulator, or the false positive ransomware test.

Appendix 1

Methodology used in the “In the Wild / Full Spectrum” test

1. Windows 10 Enterprise 64-bit operating system is installed on a hardened virtual machine, all updates are applied, and third-party applications installed and updated.
2. An image of the operating system is created.
3. A clone of the imaged systems is made for each of the security applications used in the test.
4. An individual security application is installed using default settings on each of the systems created in (3) and then, where applicable, updated. If the vendor provided a non-default setting, this setting is checked whether it is realistic. If yes, the changes are documented, applied, and added to the appendix section of the report.
5. A clone of the system as at the end of (4) is created.
6. Downloading a single binary executable (or document, script, etc.) from its native URL using Chrome to the Downloads folder and then executing the binary in the clean, unprotected system. If the sample works, the sample is saved in a replay proxy to provide the same binary throughout the test.

Live URL test is conducted by the following procedure.

- 6.1. The sample is selected for the test and tested in the systems where a security product is installed.
 - 6.2. The test case is retested 24 hours after the initial test if the security application failed to block the malicious binary.
- **The test case is marked as “Blocked”** if either the security application blocks the URL where the malicious binary was located, or the security application blocks the malicious binary whilst it was being downloaded to the machine.
 - **The test case is marked as “Behaviour Blocked”** if the security application blocks the malicious binary when it is executed and either automatically blocks it or postpones its execution and warns the user that the file is malicious and awaiting user input.
 - **The test case is marked as “Detected”** if the security application detects the threat and sends an alert to the central console or notifies the user, but the sample is allowed to run.
 - **The test case is marked as “Blocked in 24h”** if the security application fails to block or behaviour block the malicious sample but blocks it during the retest.

- **The test case is marked as “Missed”** if the security application fails to block or behaviour block the malicious sample during both tests.
7. Tests are conducted with all systems having Internet access.
 8. As no user-initiated scans are involved in this test, applications rely on various technologies to detect, block, and remediate threats. Some of these technologies are URL blacklisting, reputation, signature, machine learning, heuristics, behaviour, etc.

Methodology used in the “In-The-Wild PUA/Adware” test

1. Windows 10 Enterprise 64-bit operating system is installed on a hardened virtual machine, all updates are applied, and third-party applications installed and updated.
 2. An image of the operating system is created.
 3. A clone of the imaged systems is made for each of the security applications used in the test.
 4. An individual security application is installed using default settings on each of the systems created in (3) and then, where applicable, updated. If the vendor provided a non-default setting, this setting is checked whether it is realistic. If yes, the changes are documented, applied, and added to the appendix section of the report.
 5. A clone of the system as at the end of (4) is created.
 6. Downloading a single binary executable (or document, script, etc.) from its native URL using Chrome to the Downloads folder and then executing the binary in the clean, unprotected system. If the sample works, the sample is saved in a replay proxy to provide the same binary throughout the test.
 7. The sample is selected for the test and tested in the systems where a security product is installed.
 8. The test case is retested 24 hours after the initial test if the security application failed to block the malicious binary.
- **The test case is marked as “Blocked”** if either the security application blocks the URL where the malicious binary was located, or the security application blocks the malicious binary whilst it was being downloaded to the machine.
 - **The test case is marked as “Behaviour Blocked”** if the security application blocks the malicious binary when it is executed and either automatically blocks it or postpones its execution and warns the user that the file is malicious and awaiting user input.
 - **The test case is marked as “Detected”** if the security application detects the threat and sends an alert to the central console or notifies the user, but the sample is allowed to run.

- **The test case is marked as “Blocked in 24h”** if the security application fails to block or behaviour block the malicious sample but blocks it during the retest.
 - **The test case is marked as “Missed”** if the security application fails to block or behaviour block the malicious sample during both tests.
9. Tests are conducted with all systems having Internet access.

As no user-initiated scans are involved in this test, applications rely on various technologies to detect, block and remediate threats. Some of these technologies are URL blacklisting, reputation, signature, machine learning, heuristics, behaviour etc.

Methodology used in the False Positive test

1. Windows 10 Enterprise 64-bit operating system is installed on a hardened virtual machine, all updates are applied, and third-party applications installed and updated.
2. An image of the operating system is created.
3. A clone of the imaged systems is made for each of the security applications used in the test.
4. An individual security application is installed using default settings on each of the systems created in (3) and then, where applicable, updated. If the vendor provided a non-default setting, this setting is checked whether it is realistic. If yes, the changes are documented, applied, and added to the appendix section of the report.
5. A clone of the system as at the end of (4) is created.
6. Introducing the binary executables (or documents, scripts, etc.) to the clean, unprotected system via disk image or network share. If the sample works, the sample is saved to a different disk image or network share.

False Positive test is conducted by the following procedure.

- 6.1. Scanning the binary executables (or documents, scripts, etc.) on the disk image or on the network share.
 - 6.2. Executing the test samples.
 - 6.3. The sample is retested 24 hours after the initial test if the security application failed to permit the harmless file.
- **The test case is marked as “False block”** if the security application falsely identifies and blocks the binary at any stage during the test and retest.
 - **The test case is marked as “Detected”** if the security application falsely identifies and the binary at any stage during the test and retest but allows it to run.

- **The test case is marked as “Allowed to run in 24h”** if the security application falsely identifies and blocks the binary at any stage during the test but allows it to run upon the retest.
 - **The test case is marked as “Allowed to run”** if the security application correctly identifies the binary as harmless and allows it to run.
7. Tests are conducted with all systems having Internet access.

Methodology used in the Exploit/Fileless test

1. Windows 10 Enterprise 64-bit operating system is installed on a hardened virtual machine, all updates are applied, and third-party applications installed and updated.
2. An image of the operating system is created.
3. A clone of the imaged systems is made for each of the security applications used in the test.
4. An individual security application is installed using default settings on each of the systems created in (3) and then, where applicable, updated. If the vendor provided a non-default setting, this setting is checked whether it is realistic. If yes, the changes are documented, applied, and added in the report in an appendix.
5. A clone of the system as at the end of (4) is created.

Exploit / Fileless test is conducted by the following procedure.

6. Our payloads use an exploit for the one of an installed vulnerable application. To simulate a realistic attack scenario, a payload is constructed to include at least one of the common CnC frameworks.
7. The opening stage of the exploit is introduced to the system and we monitor if the vulnerable application starts the initial stage payload, the exploit is being executed and if a session is established to our CnC server.
8. After navigating to the exploit site, the system is supervised to see if there are any new processes, loaded DLLs or CnC traffic emerges. If the exploitation is successful, the following actions are executed:
 - 8.1. Upload a file to the victim.
 - 8.2. Download a file from the victim.
 - 8.3. Create a process remotely.
 - 8.4. Read the contents of a file on the victim.

9. When user interaction is needed from the endpoint protection (e.g. site visit not recommended, etc.) the default action is chosen. When user interaction is needed from the operating system, we chose the run/allow options.
10. Throughout the test, the Process Monitor from the Sysinternals Suite and Wireshark are running (both installed to non-default directories and modified not to be detected by default anti-debugging tools).
 - **The test case is marked as “Signature Block”** if the security application blocks the URL (infected URL, exploit kit URL, redirection URL, malware URL) by the URL database (local or cloud).
 - **The test case is marked as “Blocked”** if the security application blocks the page containing a malicious HTML code, JavaScript (redirects, iframes, obfuscated JavaScript, etc.) or Flash files. Or if the security application blocks the downloaded payload by analysing the malware before it can be started. (reputation-based block or heuristic based block).
 - **The test case is marked as “Behaviour Blocked”** if the security application blocks the downloaded payload after it has been started.
 - **The test case is marked as “Detected”** if the security application detects the threat and sends an alert to the central console or notifies the user but the attack is allowed to run.
 - **The test case is marked as “Missed”** if the security application fails to detect, block or behaviour block the attack and the malicious action is carried out.
11. Tests are conducted with all systems having Internet access.
12. As no user-initiated scans is involved in this test, applications rely on various technologies to detect, block and remediate threats. Some of these technologies are URL blacklisting, reputation, signature, machine learning, heuristics, behaviour etc.

Detailed description of the Exploit / Fileless cases.

Test case 001

Koadic / BAT

Koadic is a framework using VBScript stagers for increased stealth and limited footprint. In this test case, a Koadic connectback payload is instantiated using a wmic command.

In case the exploitation was successful, as a proof of that working session has been established, the following actions has been carried out through the connection.

- A directory list is queried
- A file is uploaded to the victim
- A file is downloaded
- A shell command is executed

Test case is flagged as MISSED if exploitation was successful and test machine had been successfully controlled via the new session.
References: <https://github.com/zerosum0x0/koadic>

Test case 002

Koadic / BAT v2

Koadic is a framework using VBScript stagers for increased stealth and limited footprint. In this test case, a Koadic connectback payload is instantiated using a wmic command.

In case the exploitation was successful, as a proof of that working session has been established, the following actions has been carried out through the connection.

- A directory list is queried
- A file is uploaded to the victim
- A file is downloaded
- A shell command is executed

Test case is flagged as MISSED if exploitation was successful and test machine had been successfully controlled via the new session.

References: <https://github.com/zerosum0x0/koadic>

Test case 003

Koadic / regsvr32

Koadic is a framework using VBScript stagers for increased stealth and limited footprint. In this test case, a Koadic connectback payload is instantiated using a regsvr32 remote object load call.

In case the exploitation was successful, as a proof of that working session has been established, the following actions were carried out through the connection.

- A directory list is queried
- A file is uploaded to the victim
- A file is downloaded
- A shell command is executed

The test case is flagged as MISSED if exploitation was successful and test machine had been successfully controlled via the new session.

References: <https://github.com/zerosum0x0/koadic>

Test case 004

Octopus / BAT

In this test case, we use the Octopus framework which is a framework using VBScript stagers for increased stealth and limited footprint. In this test case, a PowerShell connectback payload is instantiated.

In case the exploitation was successful, as a proof of a working session, the following steps were taken.

- A directory list is queried
- A file has been downloaded
- A file has been uploaded
- A shell command is executed

The test case is flagged as MISSED if exploitation was successful and the test machine had been successfully controlled via the new session.

References: <https://github.com/mhaskar/Octopus>

Test case 005

Koadic / BitsAdmin

Koadic is a framework using VBScript stagers for increased stealth and limited footprint. In this test case, a Powershell connectback payload is instantiated through the BITSadmin Windows management interface.

In case the exploitation was successful, as a proof of that working session has been established, the following actions has been carried out through the connection.

- A directory list is queried
- A file has been downloaded
- A file has been uploaded
- A shell command is executed

Test case is flagged as MISSED if exploitation was successful and test machine had been successfully controlled via the new session.

References: <https://0x1.gitlab.io/exploitation-tools/Koadic/>

Test case 004

Octopus / BAT v2

In this test case, we use the Octopus framework which is a framework using VBScript stagers for increased stealth and limited footprint. In this test case, a PowerShell connectback payload is instantiated.

In case the exploitation was successful, as a proof of a working session, the following steps were taken.

- A directory list is queried
- A file is uploaded to the victim
- A file is downloaded
- A shell command is executed

The test case is flagged as MISSED if exploitation was successful and the test machine had been successfully controlled via the new session.

References: <https://github.com/mhaskar/Octopus>

Test case 007

Powercat / BAT

Powercat is a pure-PowerShell implementation of netcat. In this test case, a .bat file is used to instantiate a powershell script to download the ps1 from the remote server and to use it to create a connectback shell to the attacker's box. The following activities are carried out.

- A directory list is queried
- A file is uploaded to the victim
- A file is downloaded
- A shell command is executed

Test case is flagged as MISSED if exploitation was successful and test machine had been successfully controlled via the new session.

References: <https://github.com/besimorhino/powercat>

Test case 008

Octopus / BAT v3

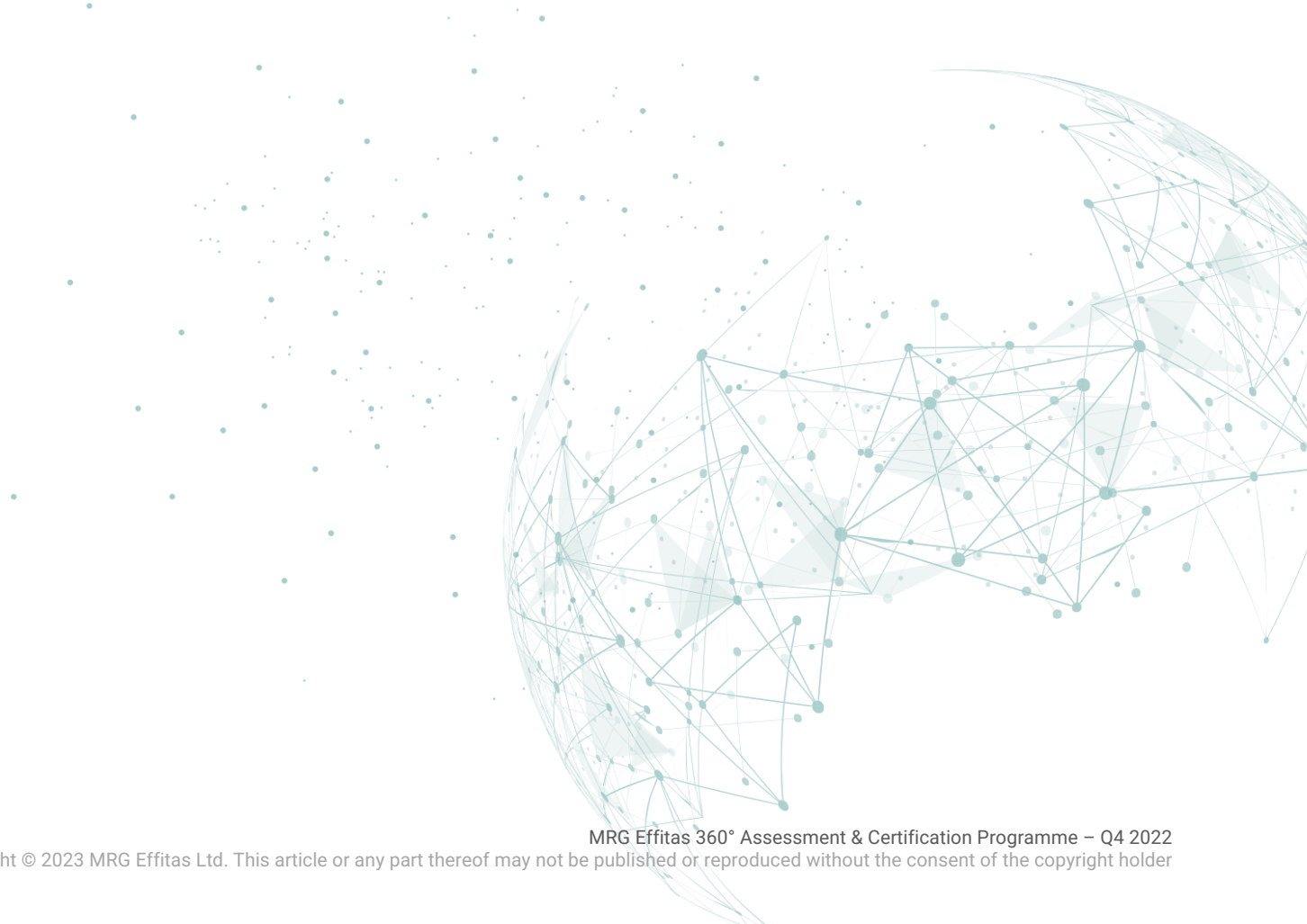
In this test case, we use the Octopus framework which is a framework using VBScript stagers for increased stealth and limited footprint. In this test case, a PowerShell connectback payload is instantiated.

In case the exploitation was successful, as a proof of a working session, the following steps were taken.

- A directory list is queried
- A file is uploaded to the victim
- A file is downloaded
- A shell command is executed

The test case is flagged as MISSED if exploitation was successful and the test machine had been successfully controlled via the new session.

References: <https://github.com/mhaskar/Octopus>



Methodology used in the Real Botnet Test

1. Windows 10 Enterprise 64-bit operating system is installed on a hardened virtual machine, all updates are applied, and third-party applications installed and updated.
2. An image of the operating system is created.
3. A Real botnet dropper is run on the clean, unprotected system, thus simulating a pre-infected state.
4. A clone of the imaged system is made for each of the security applications to be used in the test.
5. An individual security application is installed using default settings on each of the systems created in (4) and then, where applicable, updated. If the vendor provided a non-default setting, this setting is checked as to whether it is realistic. If yes, the changes are documented, applied, and added in the report in an appendix.
6. A clone of the system as at the end of (5) is created.

Real botnet test is conducted by the following procedure.

- 6.1. Starting a new instance of Chrome (or the Safe Browser) and navigating to a financial website. Where the security application offers a secured or dedicated banking browser, this is used.
 - 6.2. Text is entered into the Account login page of the financial website using the keyboard or using a virtual keyboard if the application under test provides such functionality, and then the “log in” button is pressed.
- **The test case is marked as passed – a green checkmark** if the security application detects the financial malware when the security application is installed, and a mandatory scan is made. Or the security application detects the real financial malware when it is executed according to the following criteria:
 - It identifies the real financial malware as being malicious and either automatically blocks it or postpones its execution, warns the user that the file is malicious and awaits user input.
 - It identifies the real financial malware as suspicious or unknown and gives the option to run in a sandbox or safe restricted mode, which prevents the real financial malware from capturing and sending the login data to the MRG CnC, whilst giving no alerts or giving informational alerts only. Or the security application intercepts the action of the real financial malware and displays warnings and user action input requests that are clearly different from those displayed in response to legitimate applications.
 - a. **The test case is marked as missed – a red cross** if the security application fails to detect the real financial malware according to the following criteria:

- The security application fails to prevent the real financial malware from capturing and sending the login data to the MRG CnC and gives no alert or provides informational alerts only.
 - The security application intercepts the action of the real financial malware but displays warnings and user action input requests that are indistinguishable in meaning from those displayed in response to legitimate applications.
 - The security application identifies the malware and gives the option to run in a sandbox or safe restricted mode which fails to prevent the real financial malware from capturing and sending the login data to the MRG CnC and gives no alert or provides informational alerts only.
7. Testing is conducted with all systems having Internet access.

Because we did not use zero-day malware in this test, but 1-2 years old or even older malware versions, when a security application provided both traditional AV engines and safe browser solutions, the security application was tested in two modes. In the first mode, all protections were turned on and the safe browser was used. In the second mode, all protections were turned on and the safe browser was not used. Thus, the second test simulated that if the user forgot to use the safe browser, but the AV engine is still on.

Methodology Used in the Banking Simulator Test

1. Windows 10 Enterprise 64-bit operating system is installed on a hardened virtual machine, all updates are applied, and third-party applications installed and updated.
2. An image of the operating system is created.
3. A clone of the imaged systems is made for each of the security applications used in the test.
4. An individual security application is installed using default settings on each of the systems created in (3) and then, where applicable, updated. If the vendor provided a non-default setting, this setting is checked as to whether it is realistic. If yes, the changes are documented, applied, and added to the appendix section of the report.
5. A clone of the system as at the end of (4) is created.

Financial malware simulator test is conducted by the following procedure.

6. Where the security application offers a secured or dedicated banking browser, this is used. If the security application is designed to protect the browser, only that component is tested.
 - 6.1. The simulator specific process is started.

- **The test case is marked as passed – a green checkmark** if the security application identifies the simulator as being malicious. It must then either automatically block it, postpone its execution, then warn the user that the file is malicious and awaits user input, or it identifies the simulator as suspicious or unknown and gives the option to run in a sandbox or safe restricted mode which does not allow the hooking/redirection. With successful hooking, the personal data must not be captured from the browser.
 - **The test case is marked as missed – a red cross** if the security application fails to identify the simulator based on the following criteria:
 - The security application allows the hooking/redirection of the event, and the personal data is captured from the browser, or, it fails to prevent the simulator from injecting itself into the browser process and gives no alert or provides informational alerts only.
 - The security application identifies the simulator as malware or unknown and gives the option to run in a sandbox or safe restricted mode which fails to prevent the simulator from injecting itself into the browser process and gives no alert or provides informational alerts only. Alternatively, the security application allows the hooking/redirection of the event, and the personal data is captured from the browser.
7. Testing is conducted with all systems having Internet access.

Methodology used in the Ransomware Simulator test

1. Windows 10 Enterprise 64-bit operating system is installed on a hardened virtual machine, all updates are applied, and third-party applications installed and updated.
 2. An image of the operating system is created.
 3. A clone of the imaged systems is made for each of the security applications used in the test.
 4. An individual security application is installed using default settings on each of the systems created in (3) and then, where applicable, updated. If the vendor provided a non-default setting, this setting is checked whether it is realistic. If yes, the changes are documented, applied, and added to the appendix section of the report.
 5. A clone of the system as at the end of (4) is created.
 6. Downloading a single binary executable (or document, script, etc.) from its native URL using Chrome to the Downloads folder and then executing the binary in the clean, unprotected system. If the sample works, the sample is saved in a replay proxy to provide the same binary throughout the test.
- **The test case is marked as “Blocked”** when either the security application blocks the URL where the malicious binary was located, or the security application blocks the malicious binary whilst it was being downloaded to the machine.

- **The test case is marked as “Behaviour Blocked”** if the security application blocks the malicious binary when it is executed and either automatically blocks it or postpones its execution and warns the user that the file is malicious and awaiting user input.
 - **The test case is marked as “Detected”** if the security application detects the threat and sends an alert to the central console or notifies the user, but the sample is allowed to run.
 - **The test case is marked as “Blocked in 24h”** if the security application fails to block or behaviour block the malicious sample but blocks it during the retest.
 - **The test case is marked as “Missed”** if the security application fails to block or behaviour block the malicious sample during both tests.
9. Tests are conducted with all systems having Internet access.
10. As no user-initiated scans are involved in this test, applications rely on various technologies to detect, block, and remediate threats. Some of these technologies are URL blacklisting, reputation, signature, machine learning, heuristics, behaviour etc.

Detailed description of the Ransomware Simulator cases

Test case 1 - Onyx (new file)

OnyxLocker is a proof-of-concept ransomware written in the C# language using the .NET framework.

The test malware encrypts the user documents by creating a new encrypted file and deleting the original. Newly created filenames will have additional extensions like .onyx, .locked, .crypted.

Test case 2 - Onyx (rename)

Just like the previous test case, OnyxLocker is a proof-of-concept ransomware written in the C# language using the .NET framework.

The test malware first encrypts the victim's file in place then appends the extensions like .onyx, .locked, .crypted.

Test case 3 – Python (new file)

We created a proof-of-concept test Python script to simulate ransomware activities. The product should have to block the application before it finishes its activities. The test malware encrypts the user documents by creating a new encrypted file and deleting the original.

Test case 4 - Python (rename)

We created a proof-of-concept test Python script to simulate ransomware activities. The product should have to block the application before it finishes its activities. The test malware first encrypts the victim's file in place then appends new extension.

Methodology used in the False Positive Ransomware test

1. Windows 10 Enterprise 64-bit operating system is installed on a hardened virtual machine, all updates are applied, and third-party applications installed and updated.
2. An image of the operating system is created.
3. A clone of the imaged systems is made for each of the security applications used in the test.
4. An individual security application is installed using default settings on each of the systems created in (3) and then, where applicable, updated. If the vendor provided a non-default setting, this setting is checked whether it is realistic. If yes, the changes are documented, applied, and added to the appendix section of the report.
5. A clone of the system as at the end of (4) is created.
6. Manually executing the benign application and initiating mass file operations.
 - **The test case is marked as “False block”** if the security application falsely identifies and blocks the binary at any stage during the test and retest.
 - **The test case is marked as “Detected”** if the security application falsely identifies and the binary at any stage during the test and retest but allows it to run.
 - **The test case is marked as “Allowed to run in 24h”** if the security application falsely identifies and blocks the binary at any stage during the test but allows it to run upon the retest.
 - **The test case is marked as “Allowed to run”** if the security application correctly identifies the binary as harmless and allows it to run.
8. Tests are conducted with all systems having Internet access.

Detailed description of the False Positive Ransomware cases

Test case 1 - encrypt0r (v3.11.11)

encrypt0r provides a simple drag and drop interface to encrypt/decrypt your files using a password.

<https://github.com/kunalnagar/encrypt0r>

Test case 2 - Kryptor (v4.0.1)

Kryptor is a simple, modern, and secure file encryption and signing tool for Windows, Linux, and macOS.

<https://github.com/samuel-lucas6/Kryptor>

Test case 3 - iSunshare SafeFile Genius (v3.1.1.2)

Powerful File encryption tool to secure your important and sensitive information in internal/external hard drives on Windows.

<https://www.isunshare.com/safefile-genius/index.html>

Methodology used in Performance and Footprint test

1. Windows 10 Enterprise 64-bit operating system is installed on a physical machine, all updates are applied, and third-party applications installed and updated.
2. A backup image of the operating system is created.
3. The security application is installed, with the same configuration it is used in the other tests.
4. The following performance metrics are measured.
 - Operating system boot time
 - Size of the files installed and created by the security application. The size is measured at least one week after the installation, after virus definition updates, scans, and time passed with normal computer usage.
 - Copy time of files
 - Archive operation time
 - Opening time for (clean) files in Office applications
 - Downloading files through browser
 - Website loading time in browser. The browser should fully load a popular, complex website, from a local network URL or replay proxy to eliminate network latency.

Every performance result is a calculated average of at least three measurements.

The performance chart was calculated as follows:

- The security product reaching the best result in the category was rewarded with 8 points, the second received 7 points and so on. Once every performance category was measured, the points were summed up, and the final calculation was made by dividing the total points by the number of tests the product's result could have been measured against.

Additionally, for informational purposes, the following footprint data is collected.

- AV product update time
- AV product update CPU usage (%)
- AV product update memory usage (Mb)
- AV product update physical disk usage (%)
- AV product update network interface usage (%)
- System disk scan time
- System disk scan CPU usage (%)
- System disk scan memory usage (Mb)
- System disk scan physical disk usage (%)
- System disk scan network interface usage (%)

Physical machine specification

- Browser: Google Chrome v109.0.5414.120
- OS: Windows 10 x64 20H2
- CPU: Intel Core i5
- Memory: 8GB
- Storage: 100GB SSD

Hardened virtual machine specification

- Browser: Google Chrome v109.0.5414.120
- OS: Windows 10 x64 20H2
- CPU: 4 core processor
- Memory: 8GB
- Storage: 100GB SSD



Appendix 2

Non-default endpoint protection configurations

Endpoint protection software was running on custom configuration if suggested by the vendor.

- **Avast Business Antivirus**
Detailed logging was enabled via configuration file and Self-defense module was turned off.
- **Avira Antivirus Pro**
Log level was set to 'Complete' instead of 'Default' in 'System Scanner' and in 'Real-Time Protection'.
- **Bitdefender Endpoint Security**
Sandbox detection set to monitor only.
- **ESET Endpoint Security**
Detection of 'Potentially unwanted applications' and 'Potentially unsafe applications' were turned on among with 'SSL/TLS protocol filtering'.
- **Microsoft Windows Defender**
Microsoft Defender ATP endpoint detection and response capabilities including ASR rules were turned on.

Default endpoint protection configurations

- **Malwarebytes Endpoint Protection**
- **Symantec Endpoint Protection**
- **Trend Micro Security**

Version History

Nr.	Modify date	Comment
1.0	24.02.2023	Report published

