SE LABS ®

EPS
PROTECTION

Security Evaluation Test Report
Small Business Endpoint Security

ONLINE REPORT

MARCH 2025

SE LABS tested a variety of anti-malware (aka ‘anti-virus’; aka ‘endpoint security’) products from a range of well-known vendors in an effort to judge which were the most effective.

Each product was exposed to the same threats, which were a mixture of targeted attacks using well-established techniques and public email and web-based threats that were found to be live on the internet at the time of the test.

The results indicate how effectively the products were at detecting and/or protecting against those threats in real time.

Contents

Document version 1.0 Written 22nd April 2025

SE LABS

Management
Chief Executive Officer Simon Edwards
Chief Operations Officer Marc Briggs
Chief Human Resources Officer Magdalena Jurenko

Testing Team
Nikki Albesa
Thomas Bean
Solandra Brewster
Jarred Earlington
Gia Gorbold
Anila Johny
Erica Marotta
Jeremiah Morgan
Julian Owusu-Abrokwa
Joseph Pike
Georgios Sakatzidis
Enejda Torba
Dimitrios Tsarouchas
Stephen Withey

Marketing
Sara Claridge

Publication
Rahat Hussain
Colin Mackleworth

IT Support
Danny King-Smith
Chris Short

Website selabs.uk
Email info@SELabs.uk
LinkedIn www.linkedin.com/company/se-labs/
Blog blog.selabs.uk
Post SE Labs Ltd, 55A High Street, Wimbledon, SW19 5BA, UK

SE Labs is ISO/IEC 27001 : 2022 certified and BS EN ISO 9001 : 2015 certified for The Provision of IT Security Product Testing.

SE Labs is a member of the Microsoft Virus Initiative (MVI); the Anti-Malware Testing Standards Organization (AMTSO); the Association of anti Virus Asia Researchers (AVAR); and NetSecOPEN.

amtso
CONFIRMED COMPLIANT

CEO
Simon Edwards

If you spot a detail in this report that you don't understand, or would like to discuss, please contact us. SE Labs uses current threat intelligence to make our tests as realistic as possible. To learn more about how we test, how we define 'threat intelligence' and how we use it to improve our tests please visit our website and follow us on LinkedIn.

Can Your Endpoint Protection Stop a Real Hacker?

To find out, we test like hackers

In the enterprise security space, bold claims are everywhere. Most vendors say their endpoint protection stops ransomware, blocks phishing, and detects advanced threats. But when the stakes are high, how many tools can actually deliver?

At SE Labs, we don't rely on vendor claims. We Test Like Hackers.

That means replicating real-world attacks using threat intelligence and offensive tools. We create phishing emails, customise exploits, build backdoors and more. We don't cut corners. We mimic genuine adversaries to see how well products perform under realistic, high-pressure conditions.

Why do we go to all this trouble? Because businesses need answers grounded in reality, not synthetic benchmarks or scripted demos. We copy the bad guys to discover the truth.

In this comparative report, we put leading endpoint products through rigorous testing. Each product faced the same attack scenarios, allowing us to observe how early they detected threats, whether they blocked them effectively, and how well they protected the system overall.

If your organisation depends on endpoint security to protect sensitive data, this report will show you which solutions are worth your trust, and which ones may leave you exposed.

Executive Summary

Product Names

It is good practice to stay up to date with the latest version of your chosen endpoint security product. We made best efforts to ensure that each product tested was the very latest version running with the most recent updates to give the best possible outcome.

For specific build numbers, see **Appendix E: Product Versions** on page 16.

Executive Summary

Products Tested	Protection Accuracy Rating (%)	Legitimate Accuracy Rating (%)	Total Accuracy Rating (%)
ESET Endpoint Security	100%	100%	100%
Kaspersky Small Office Security	100%	100%	100%
Microsoft Defender Antivirus (enterprise)	100%	100%	100%
Sophos Intercept X	98%	100%	99%
Webroot SecureAnywhere Endpoint Protection	97%	98%	98%

- **Products highlighted in green were the most accurate, scoring 85 per cent or more for Total Accuracy. Those in yellow scored less than 85 but 75 or more. Products shown in red scored less than 75 per cent.**

For exact percentages, see **1.3 Total Accuracy Ratings** on page 9.

- **The endpoints were generally effective at handling general threats from cyber criminals...**
  All products were very capable of handling public email- and web-based threats such as those used by criminals to attack Windows PCs, tricking users into running malicious files or running scripts that download and run malicious files.

- **... and likewise provided complete protection against targeted attacks.**
  All of the endpoints proved highly effective against the targeted attacks used in this test. This is an encouraging trend since it only takes one targeted attack to breach an organisation.

- **False positives were not an issue for the products.**
  Almost all of the products were perfectly good at correctly classifying legitimate applications and websites.

- **Which products were the most effective?**
  Products from **ESET**, **Kaspersky** and **Microsoft** produced extremely good results due to a combination of their ability to block malicious URLs, handle exploits and correctly classify applications and websites. All products performed well enough to achieve AAA awards.

Security Evaluation
EPS Protection Small Business Awards

The following products win SE Labs awards:

ESET Endpoint Security

Kaspersky Small Office Security

Microsoft Defender Antivirus (enterprise)

Sophos Intercept X

Webroot SecureAnywhere Endpoint Protection

Threat Responses

Full Attack Chain: Testing Every Layer of Detection and Protection

Attackers start from a certain point and don't stop until they have either achieved their goal or have reached the end of their resources (which could be a deadline or the limit of their abilities). This means that, in a test, the tester needs to begin the attack from a realistic first position, such as sending a phishing email or setting up an infected website, and moving through many of the likely steps leading to actually stealing data or causing some other form of damage to the network.

If the test starts too far into the attack chain, such as executing malware on an endpoint, then many products will be denied opportunities to use the full extent of their protection and detection abilities. If the test concludes before any 'useful' damage or theft has been achieved, then similarly the product may be denied a chance to demonstrate its abilities in behavioural detection and so on.

Attack Stages

The illustration (below) shows typical stages of an attack. In a test, each of these should be attempted to determine the security solution's effectiveness. This test's results record detection and protection for each of these stages.

We measure how a product responds to the first stages of the attack with a detection and/or protection rating. Sometimes products allow threats to run yet still detect them. Other times they might allow the threat to run briefly before neutralising it. Ideally, they detect and block the threat before it has a chance to run. Products may delete threats or automatically contain them in a 'quarantine' or other safe holding mechanism for later analysis.

Should the initial attack phase succeed, we then measure post-exploitation stages, which are represented by steps two through to seven below. We broadly categorise these stages as: Access (step 2); Action (step 3); Escalation (step 4); and Post-Escalation (steps 5-6).

In figure 1. you can see a typical attack running from start to end, through various 'hacking' activities. This can be classified as a fully successful breach.

In figure 2. a product or service has interfered with the attack, allowing it to succeed only as far as stage 3, after which it was detected and neutralised. The attacker was unable to progress through stages 4 onwards.

It is possible for an attack to run in a different order with, for example, the attacker attempting to connect to other systems without needing to escalate privileges. However, it is common for password theft (see step 5) to occur before using stolen credentials to move further through the network.

How Hackers Progress

Figure 1. A typical attack starts with an initial contact and progresses through various stages, including reconnaissance, stealing data and causing damage.

Figure 2. This attack was initially successful but only able to progress as far as the reconnaissance phase.

1. Protection and Legitimate Handling Accuracy

1.1 Protection Details

These results break down how each product handled threats into some detail. You can see how many detected a threat and the levels of protection provided.

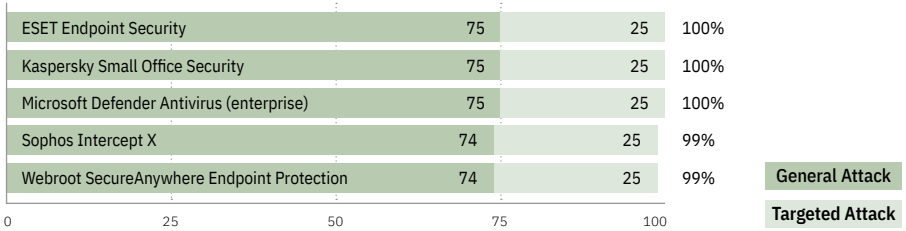
Products sometimes detect more threats than they protect against. This can happen when they recognise an element of the threat but aren't equipped to stop it. Products can also provide protection even if they don't detect certain threats. Some threats abort on detecting specific endpoint protection software.

Product	Detected	Blocked	Neutralised	Compromised	Protected
ESET Endpoint Security	100	100	0	0	100
Kaspersky Small Office Security	100	100	0	0	100
Microsoft Defender Antivirus (enterprise)	100	100	0	0	100
Sophos Intercept X	99	99	0	1	99
Webroot SecureAnywhere Endpoint Protection	99	97	2	1	99

- **This data shows in detail how each product handled the threats used.**

1.2 Attack Types

The graph shows how each product protected against the different types of attacks used in the test.

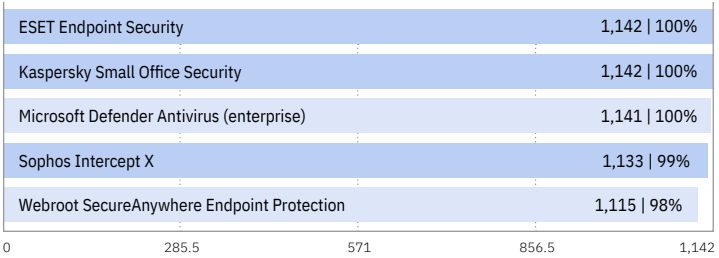


1.3 Total Accuracy Ratings

Judging the effectiveness of an endpoint security product is a subtle art, and many factors are at play when assessing how well it performs. To make things easier we've combined all the different results from this report into one easy-to-understand graph.

The graph takes into account not only each product's ability to detect and protect against threats, but also its handling of non-malicious objects such as web addresses (URLs) and applications.

Not all protections, or detections for that matter, are equal. A product might completely block a URL, which stops the threat before it can even start its intended series of malicious events. Alternatively, the product might allow a web-based exploit to execute but prevent it from downloading any further code to the target.



- **Total Accuracy Ratings combine protection and false positives.**

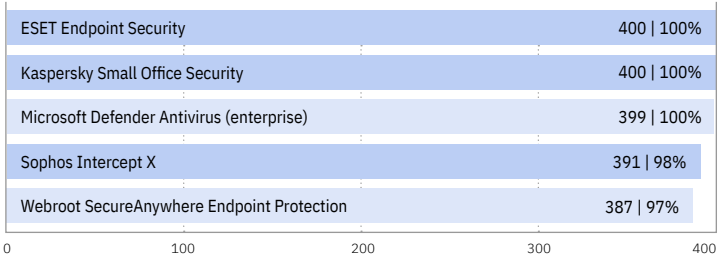
- **Categorising how a product handles legitimate objects is complex, and you can find out how we do it in Legitimate Accuracy Ratings on page 10.**

In another case malware might run on the target for a short while before its behaviour is detected and its code is deleted or moved to a safe 'quarantine' area for future analysis. We take these outcomes into account when attributing points that form final ratings.

For example, a product that completely blocks a threat is rated more highly than one that allows a threat to run for a while before eventually evicting it. Products that allow all malware infections, or that block popular legitimate applications, are penalised heavily.

1.4 Protection Accuracy

To understand how we calculate these ratings, see **Appendix A: Protection Ratings** on page 12.



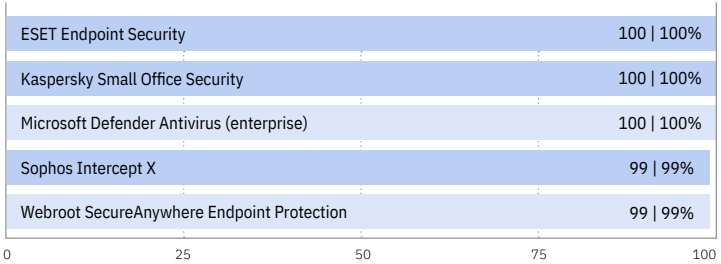
- **Protection Ratings are weighted to show that how products handle threats can be subtler than just 'win' or 'lose'.**

Average 99%

1.5 Protection Scores

This graph shows the overall level of protection, making no distinction between neutralised and blocked incidents.

For each product we add Blocked and Neutralised cases together to make one simple tally.



● **Protection Scores are a simple count of how many times a product protected the system.**

1.6 Legitimate Accuracy Ratings

These ratings indicate how accurately the products classify legitimate applications and URLs, while also taking into account the interactions that each product has with the user. Ideally a product will either not classify a legitimate object or will classify it as safe. In neither case should it bother the user.

We also take into account the prevalence (popularity) of the applications and websites used in this part of the test, applying stricter penalties for when products misclassify very popular software and sites.

To understand how we calculate these ratings, see **Accuracy Ratings** on page 14.

● **Legitimate Accuracy Ratings can indicate how well a vendor has tuned its detection engine.**

2. Conclusion

Attacks in this test included threats that affect the wider public and more closely targeted individuals and organisations. You could say that we tested the products with 'public' malware and full-on hacking attacks.

We introduced the threats in a realistic way such that threats seen in the wild on websites were downloaded from those same websites, while threats caught sending email were delivered to our target systems as emails.

All the products tested are well-known and should do well in this test. While we do 'create' threats by using publicly available free hacking tools, we do not write unique malware so there is no technical reason why any vendor being tested should do poorly.

The results were generally strong in the way the products handled both public threats and targeted attacks as three of the five products stopped all of the attacks this quarter. Products from **ESET**, **Kaspersky** and **Microsoft**, achieved 100% Protection Accuracy Ratings by being quick to block any detected intrusion.

By and large, products from **Sophos** and **Webroot** did the same. However, **Sophos Intercept X** failed to stop one malicious file from running its full course. This executable was embedded in a zip file downloaded from the web.

Webroot SecureAnywhere Endpoint Protection failed to stop one public threat from running and making multiple connections after reboot. A few points were also docked against the product for allowing two public threats to run briefly and leave artifacts in the system before neutralising them.

Public threats are live on the Internet the day that the products are tested. Consistently stopping public malware indicates both familiarity with common threats and frequent updates to keep databases current.

In previous quarters, almost all of the products would pass, with flying colours, the tests whereby public malware was used. More products would then falter when faced with the more persistent targeted attacks. It is too soon to call this a trend, but we did observe more outright wins against targeted attacks compared to public threats in this test.

This quarter, almost all of the products passed all the tests for false positives. Only **Webroot SecureAnywhere Endpoint Protection** misclassified two legitimate applications.

All the products in this test win AAA awards by virtue of scoring Total Accuracy Ratings of either 100% or in the very high 90s. The strongest, from **ESET**, **Kaspersky** and **Microsoft** scored 100% Total Accuracy Ratings. Those from **Sophos** and **Webroot** were only one or two points shy of a perfect Total Accuracy score.

Appendices

Appendix A: Protection Ratings

The results below indicate how effectively the products dealt with threats. Points are earned for detecting the threat and for either blocking or neutralising it.

Detected (+1) If the product detects the threat with any degree of useful information, we award it one point.

Blocked (+2) Threats that are disallowed from even starting their malicious activities are blocked. Blocking products score two points.

Complete Remediation (+1) If, in addition to neutralising a threat, the product removes all significant traces of the attack, it gains an additional one point.

Neutralised (+1) Products that kill all running malicious processes 'neutralise' the threat and win one point.

Persistent Neutralisation (-2) This result occurs when a product continually blocks a persistent threat from achieving its aim, while not removing it from the system.

Compromised (-5) If the threat compromises the system, the product loses five points. This loss may be reduced to four points if it manages to detect the threat (see Detected, above), as this at least alerts the user, who may now take steps to secure the system.

Rating Calculations

We calculate the protection ratings using the following formula:

Protection Rating =
(1x number of Detected) +
(2x number of Blocked) +
(1x number of Neutralised) +
(1x number of Complete Remediation) +
(-5x number of Compromised)

The 'Complete Remediation' number relates to cases of neutralisation in which all significant traces of the attack were removed from the target.

These ratings are based on our opinion of how important these different outcomes are. You may have a different view on how seriously you treat a 'Compromise' or 'Neutralisation without complete remediation'. If you want to create your own rating system, you can use the raw data from **1.1 Protection Details** on page 8 to roll your own set of personalised ratings.

Targeted Attack Scoring

The following scores apply only to targeted attacks and are cumulative, ranging from -1 to -5.

Access (-1) If any command that yields information about the target system is successful this score is applied. Examples of successful commands include listing current running processes, exploring the file system and so on. If the first command is attempted and the session is terminated by the product without the command being successful the score of Neutralised (see above) will be applied.

Action (-1) If the attacker is able to exfiltrate a document from the target's Desktop of the currently logged in user then an 'action' has been successfully taken.

Escalation (-2) The attacker attempts to escalate privileges to NT Authority/System. If successful, an additional two points are deducted.

Post-Escalation Action (-1) After escalation the attacker attempts actions that rely on escalated privileges. These include attempting to steal credentials, modifying the file system and recording keystrokes. If any of these actions are successful then a further penalty of one point deduction is applied.

Appendix B: Legitimate Interaction Ratings

It's crucial that anti-malware endpoint products not only stop – or at least detect – threats, but that they allow legitimate applications to install and run without misclassifying them as malware. Such an error is known as a 'false positive' (FP).

In reality, genuine FPs are quite rare in testing. In our experience it is unusual for a legitimate application to be classified as 'malware'. More often it will be classified as 'unknown', 'suspicious' or 'unwanted' (or terms that mean much the same thing).

We use a subtle system of rating an endpoint's approach to legitimate objects, which takes into account how it classifies the application and how it presents that information to the user. Sometimes the endpoint software will pass the buck and demand that the user decide if the application is safe or not. In such cases the product may make a recommendation to allow or block. In other cases, the product will make no recommendation, which is possibly even less helpful.

If a product allows an application to install and run with no user interaction, or with simply a brief notification that the application is likely to be safe, it has achieved an optimum result. Anything else is a Non-Optimal Classification/Action (NOCA). We think that measuring NOCAs is more useful than counting the rarer FPs.

Prevalence Ratings

There is a significant difference between an endpoint product blocking a popular application such as the latest version of Microsoft Word and condemning a rare Iranian dating toolbar for Internet Explorer 6. One is very popular all over the world and its detection as malware (or something less serious but still suspicious) is a big deal. Conversely, the outdated toolbar won't have had a comparably large user base even when it was new. Detecting this application as malware may be wrong, but it is less impactful in the overall scheme of things.

With this in mind, we collected applications of varying popularity and sorted them into five separate categories, as follows:

1. **Very High Impact**
2. **High Impact**
3. **Medium Impact**
4. **Low Impact**
5. **Very Low Impact**

	None (allowed)	Click to Allow (default allow)	Click to Allow/ Block (no recommendation)	Click to Block (default block)	None (blocked)	
Safe	2	1.5	1			A
Unknown	2	1	0.5	0	-0.5	B
Not Classified	2	0.5	0	-0.5	-1	C
Suspicious	0.5	0	-0.5	-1	-1.5	D
Unwanted	0	-0.5	1	-1.5	-2	E
Malicious				2	-2	F
	1	2	3	4	5	

Legitimate Software Prevalence Rating Modifiers

Very High Impact	5
High Impact	4
Medium Impact	3
Low Impact	2
Very Low Impact	1

Incorrectly handling any legitimate application will invoke penalties, but classifying Microsoft Word as malware and blocking it without any way for the user to override this will bring far greater penalties than doing the same for an ancient niche toolbar. In order to calculate these relative penalties, we assigned each impact category with a rating modifier, as shown in the table above.

Applications were downloaded and installed during the test, but third-party download sites were avoided and original developers' URLs were used where possible. Download sites will sometimes bundle additional components into applications' install files, which may correctly cause anti-malware products to flag adware. We remove adware from the test set because it is often unclear how desirable this type of code is.

The prevalence for each application and URL is estimated using metrics such as third-party download sites and the data from Tranco.com's global traffic ranking system.

Accuracy Ratings

We calculate legitimate software accuracy ratings by multiplying together the interaction and prevalence ratings for each download and installation:

Accuracy rating = Interaction rating x Prevalence rating

If a product allowed one legitimate, Medium impact application to install with zero interaction with the user, then its Accuracy rating would be calculated like this:

Accuracy rating = 2 x 3 = 6

This same calculation is made for each legitimate application/site in the test and the results are summed and used to populate the graph and table shown under **Legitimate Accuracy Ratings** on page 10.

Distribution of Impact Categories

Endpoint products that were most accurate in handling legitimate objects achieved the highest ratings. If all objects were of the highest prevalence, the maximum possible rating would be 1,000 (100 incidents x (2 interaction rating x 5 prevalence rating)).

In this test there was a range of applications with different levels of prevalence. The table below shows the frequency:

Legitimate Interaction Ratings

Product	None (allowed)	None (blocked)
ESET Endpoint Security	100	0
Kaspersky Small Office Security	100	0
Microsoft Defender Antivirus (enterprise)	100	0
Sophos Intercept X	100	0
Webroot SecureAnywhere Endpoint Protection	98	2

- **Products that do not bother users and classify most applications correctly earn more points than those that ask questions and condemn legitimate applications.**

Legitimate Software Category Frequency

Prevalence Rating	Frequency
Very High Impact	32
High Impact	32
Medium Impact	18
Low Impact	11
Very Low Impact	7

Appendix C: Terms Used

Compromised The attack succeeded, resulting in malware running unhindered on the target. In the case of a targeted attack, the attacker was able to take remote control of the system and carry out a variety of tasks without hindrance.

Blocked The attack was prevented from making any changes to the target.

False positive When a security product misclassifies a legitimate application or website as being malicious, it generates a 'false positive'.

Neutralised The exploit or malware payload ran on the target but was subsequently removed.

Complete Remediation If a security product removes all significant traces of an attack, it has achieved complete remediation.

Target The test system that is protected by a security product.

Threat A program or sequence of interactions with the target that is designed to take some level of unauthorised control of that target.

Update Security vendors provide information to their products in an effort to keep abreast of the latest threats. These updates may be downloaded in bulk as one or more files, or requested individually and live over the internet.

Appendix D: FAQs

Q What is a partner organisation? Can I become one to gain access to the threat data used in your tests?

A Partner organisations benefit from our consultancy services after a test has been run. Partners may gain access to low-level data that can be useful in product improvement initiatives and have permission to use award logos, where appropriate, for marketing purposes. We do not share data on one partner with other partners. We do not partner with organisations that do not engage in our testing.

Q I am a security vendor and you tested my product without permission. May I access the threat data to verify that your results are accurate?

A We are willing to share a certain level of test data with non-partner participants for free. The intention is to provide sufficient data to demonstrate that the results are accurate. For more in-depth data suitable for product improvement purposes we recommend becoming a partner.

A **full methodology** for this test is available from our website.

- The test was conducted between 17th January and 21st March 2025.
- All products were configured according to each vendor's recommendations, when such recommendations were provided.
- Targeted attacks were selected and verified by SE Labs.
- Malicious emails, URLs, attachments and legitimate messages were independently located and verified by SE Labs.
- Malicious and legitimate data was provided to partner organisations once the test was complete.

Appendix E: Product Versions

The table below shows the service's name as it was being marketed at the time of the test.

Vendor	Product	Build Version (start)	Build Version (end)
ESET	Endpoint Security	12.0.2038.0	12.0.2049.0
Kaspersky	Small Office Security	21.19.7.527(b)	21.19.7.527(b)
Microsoft	Defender Antivirus (enterprise)	Antimalware Client Version: 4.18.24090.11 Engine Version: 1.1.24090.11 Antivirus Version: 1.421.1284.0 Anti-spyware Version: 1.421.1284.0	Antimalware Client Version: 4.18.25010.11 Engine Version: 1.1.25020.1007 Antivirus Version: 1.425.152.0 Anti-spyware Version: 1.425.152.0
Sophos	Intercept X	Core Agent: 2024.3.1.3.0 Sophos Intercept: 2024.1.2.1.0 Device Encryption: 2024.3.0.71.0	Core Agent: 2024.3.2.3.0 Sophos Intercept: 2024.1.2.1.0 Device Encryption: 2024.3.0.71.0
Webroot	SecureAnywhere Endpoint Protection	9.0.38.42	9.0.38.42

SE LABS

SE Labs Report Disclaimer

1. The information contained in this report is subject to change and revision by SE Labs without notice.
2. SE Labs is under no obligation to update this report at any time.
3. SE Labs believes that the information contained within this report is accurate and reliable at the time of its publication, which can be found at the bottom of the contents page, but SE Labs does not guarantee this in any way.
4. All use of and any reliance on this report, or any information contained within this report, is solely at your own risk. SE Labs shall not be liable or responsible for any loss of profit (whether incurred directly or indirectly), any loss of goodwill or business reputation, any loss of data suffered, pure economic loss, cost of procurement of substitute goods or services, or other intangible loss, or any indirect, incidental, special or consequential loss, costs, damages, charges or expenses or exemplary damages arising his report in any way whatsoever.
5. The contents of this report does not constitute a recommendation, guarantee, endorsement or otherwise of any of the products listed, mentioned or tested.
6. The testing and subsequent results do not guarantee that there are no errors in the products, or that you will achieve the same or similar results. SE Labs does not guarantee in any way that the products will meet your expectations, requirements, specifications or needs.
7. Any trade marks, trade names, logos or images used in this report are the trade marks, trade names, logos or images of their respective owners.
8. The contents of this report are provided on an “AS IS” basis and accordingly SE Labs does not make any express or implied warranty or representation concerning its accuracy or completeness.