SE LABS

ESS
PROTECTION

Advanced Security Test Report

Email Sandbox Protection

JANUARY 2025

SE LABS ® was commissioned to compare a number of email sandboxing products capable of running independently of cloud-based services. The products, provided by [redacted] and [redacted] were tested for their abilities to detect and stop email threats in the form of targeted malware attacks.

Each product was exposed to the same threats, which were a mixture of targeted attacks using well-established techniques and a number of popular evasion methods.

Contents

Document version 1.0 Written 8th January 2025

SE LABS

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

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

Marketing
Sara Claridge
Janice Sheridan

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.

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.

Inside the Fight Against email Threats

From phishing scams to malware, email attacks exploit our instincts and devices

Email provides a route right into the heart of our computers, phones and other devices. As such, it is frequently abused to perform a variety of attacks against potential victims of cybercrime. The sophistication of attacks varies, but many rely on our almost unbreakable instinct to open, read and interact with messages sent to work and personal email accounts. Businesses rely on email security services to filter out large numbers of such attacks.

The range of attack types in the real world is wide, but in general we consider there to be two main categories: targeted attacks, in which the attacker attempts to target a specific individual; and public attacks, which spread wide and far in an attempt to compromise as many people as possible.

Many of the same techniques are used in public and targeted attacks. The least technically sophisticated include requests for a money transfer or banking login credentials.

More credible attempts include professionally formatted emails and links to fake websites designed to trick users into entering their valuable details.

Attackers with more resources may use malware to achieve their goals, either in the form of attached files or by linking to websites that exploit visiting computers.

SE Labs monitors email threats in real time, analysing large numbers of messages and extracting samples that represent large groups of those threats. Human testers then manually verify that any malware included works properly before re-sending these threats to our own accounts through the tested services. We also generate targeted attacks using the same tools and techniques used by advanced attackers.

Executive Summary

All systems were kept as up to date as would be reasonably expected from an active and educated enterprise team.

We made best efforts to ensure that the latest firmware and updates were applied to each system to give the best possible outcome.

Sandbox products are designed to run potentially malicious code in a protected environment, providing an opportunity for an enterprise to detect and block attacks before they enter the network. However, in this relatively small test, no product was perfect at stopping all threats, while one incorrectly blocked a legitimate email message.

- **None of the products provided close to full protection**

Overall, the sandbox detection rates were poor, with around one third of the threats managing to evade detection. Fortinet's product stood out as superior in comparison, due to its higher and wider detection rates.

- **It was possible to bypass the sandbox protection without using special evasion techniques**

Different evasion techniques were used in an attempt to measure how much effort an attacker would need to expend in order to bypass the sandboxes' detection methods. However, in many cases threats could bypass without using evasion techniques.

- **False positives were not a significant issue for the sandboxes**

[redacted] product misclassified one legitimate email, but the strict weighting on our ratings system reduced its Legitimate Accuracy Rating to 75%.

Executive Summary

Product	Protection Accuracy Rating	Legitimate Accuracy Rating	Total Accuracy Rating
[redacted]	42%	100%	50%
[redacted]	9%	75%	18%
[redacted]	-11%	100%	4%

- **The ratings above are weighted to take into account the different levels of detection and protection provided by the products. Negative ratings are allocated when a product misses threats or misclassifies legitimate email.**

How We Tested

The common commodity threats were gathered from the wild and replayed through the email security services. Where possible, data about the original attackers' IP addresses were provided to allow services that have reliable IP address reputation systems to use their threat intelligence during testing.

Legitimate messages were constructed in-house.

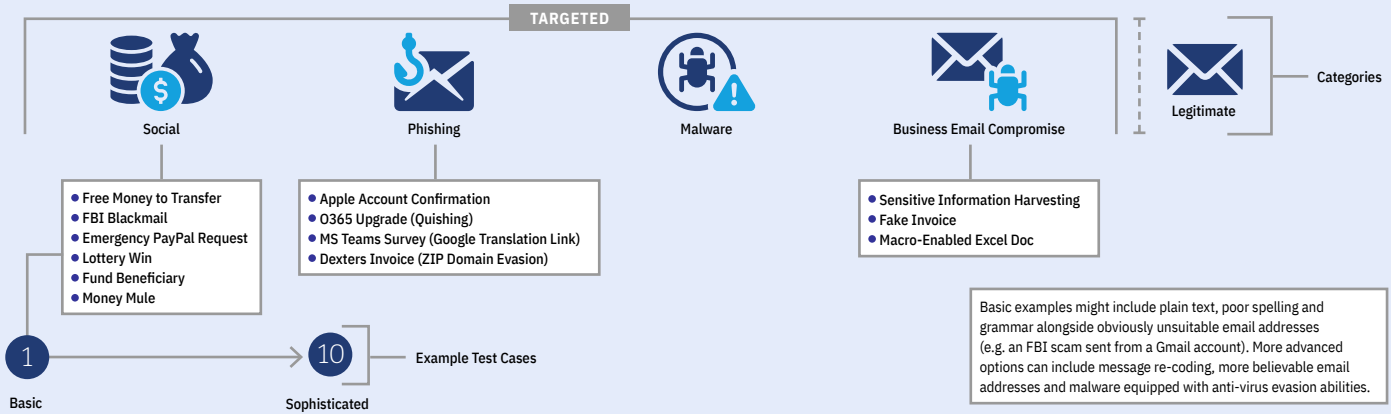
Targeted attacks comprise four distinct categories: Social Engineering; Phishing; Malware; and BEC. For each of these categories we created a number of main Test Case Structure variations.

In the example below you can see that the social engineering messages are formed into six groups (scenarios), including free money transfer, lottery win and law enforcement blackmail scams.

For each scenario we create variants that range in sophistication from extremely basic to very advanced. The goal is to test the effectiveness of each email security service and configuration when facing a range of different types of attacker, or at least a range of different attack approaches.

Email messages travel over the internet to their recipients. Before they reach the Inbox,

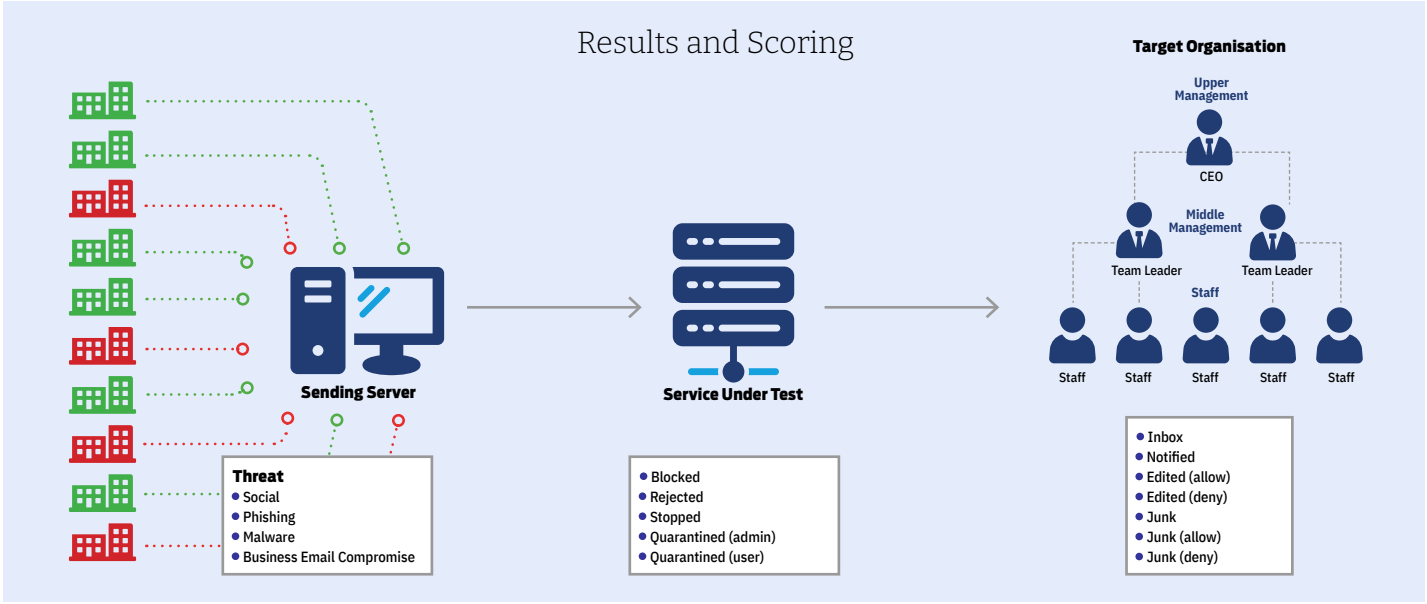
Test Case Structure



they negotiate their way through various security services before reaching the target's own infrastructure. There are opportunities for detection and protection at different stages in this journey.

Bad messages might be prevented from entering the 'service under test', being blocked or otherwise rejected. Once within the service, the message might be detected and prevented from progressing further, or it might be placed into a 'Quarantine' from which either a user or administrator may release it.

Messages may end up in the Inbox or Quarantine, with or without changes such as removed or rewritten URLs, attachments and other elements.



1. Total Accuracy Ratings

Judging the effectiveness of an email hosted protection service is a subtle art and many factors need to be considered when assessing how well it performs. To make things easier we've combined all of the different results into one easy-to-understand table.

The graphic below takes into account not only each service's ability to detect and protect against threats, but also its handling of non-malicious messages and components of those messages, such as attachments and links to websites.

Not all protection measures, or detections for that matter, are equal. A service might completely delete an incoming malicious email and never allow the intended recipient to see (and subsequently interact with) it. Services may condemn suspicious messages to a 'quarantine' area if it lacks the utter conviction that the message is unwanted. This keeps threats away from recipients unless the recipient judges that the message is really safe. At the weaker end of the scale, the service might simply add a warning to the email's Subject line.

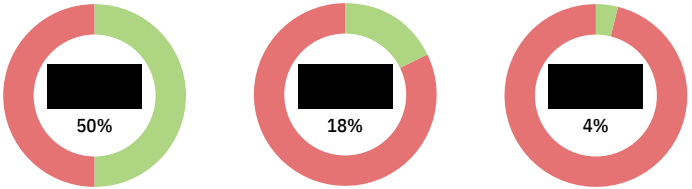
We take these different possible outcomes into account when attributing points that form final ratings.

For example, a service that completely blocks a malicious message from falling into the hands of its intended recipient is rated more highly than one that prefixes the Subject line with "Malware:" or "Phishing attempt:" or sends the message to a 'Junk' folder.

Categorising how a service handles legitimate messages is similar, but in reverse. Making a small change to the Subject line is much less serious a failing than deleting the message and failing to notify the recipient.

Total Accuracy Ratings

Products Tested	Total Accuracy Rating (%)
[redacted]	50%
[redacted]	18%
[redacted]	4%



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

2. Targeted Attacks

These results illustrate how each product handled the types of attacks that criminals use when attempting to compromise computers belonging to specific individuals. Tactics typically include sending email attachments containing customised malware that appears to be a legitimate document or other innocent file.

The results below use the following terms:

Stopped/Rejected The product prevented the threat from reaching the user's account without alerting the user.

Notified The product has detected the threat and alerted the user. It will not allow the threat into the network.

Warned The product has detected the threat and alerted the user. It will allow the threat into the network.

Inbox The service has failed to detect the threat.

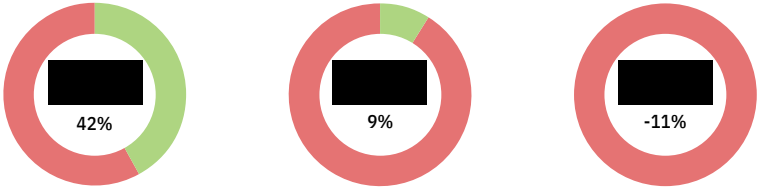
Each product is awarded four points for stopping a threat (Stopped or Notified), two for warning about it (but not blocking it) and -5 points for allowing the threat through.

Targeted Attacks

Products Tested	Stopped	Notified	Warned	Inbox
[redacted]	49	0	0	17
[redacted]	0	0	43	23
[redacted]	3	41	0	22

Protection Accuracy Ratings

Products Tested	Protection Accuracy Rating	Protection Accuracy Rating (%)
[redacted]	111	42%
[redacted]	66	9%
[redacted]	-29	-11%



- The table above shows how accurately the services handled legitimate email.

3. Legitimate Messages

A set of legitimate files, including documents of varying formats and executable programs were submitted to the sandboxes for analysis. It is important to test for false positives because too many indicate a product that is too aggressive and will block useful email as well as threats.

It would be easy to create a product that blocked all threats if it was also allowed to block all legitimate files. Finding the balance between allowing good and blocking bad is the key to almost every type of security system.

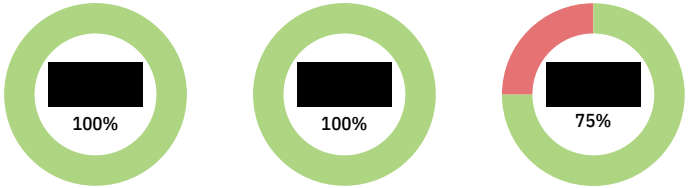
Each product is award two points for every legitimate file that it does not detect as malware. For each misclassification it loses eight points.

Legitimate File Ratings

Products Tested	Inbox	Stopped
[redacted]	19	1
[redacted]	20	0
[redacted]	20	0

Legitimate Accuracy Ratings

Products Tested	Legitimate Accuracy Rating	Legitimate Accuracy Rating (%)
[redacted]	40	100%
[redacted]	40	100%
[redacted]	30	75%



- Legitimate Accuracy Ratings give a weighted value to services based on how accurately they handle legitimate messages.

4. Evasion Effects

Attackers have a large range of techniques available to hide the malicious nature of their malware. This can involve using encryption, compression and other methods of 'scrambling' the code of a file to obfuscate its true nature.

In this test we used exploits embedded in files that appear otherwise innocent. The files were generated with Metasploit using default settings, and these files were changed using re-encoding techniques in an attempt to create stealthier malware. Additional re-encoding was used to create a potentially more advanced evasion method.

The three sets of default, encoded and re-encoded files were compressed into Zip files to create a total of six threat sets. These tables show how each product handled the different sets of threats.

(There are two more compressed threats than uncompressed. The additional two attacks require multiple files that must be extracted from their Zip files to work. Such multi-file attacks are not viable without using compression.)

Evasion Effect on Detection

[redacted]

Threats	Stopped	Notified	Warned	Inbox
1. Malware (default)	1	6	0	3
2. Malware (compressed)	0	9	0	3
3. Malware (basic encoding)	1	6	0	3
4. Malware (basic encoding/compressed)	0	8	0	4
5. Malware (advanced encoding)	1	5	0	4
6. Malware (advanced encoding/compressed)	0	7	0	5
Total	**3**	**41**	**0**	**22**

[redacted]

Threats	Stopped	Notified	Warned	Inbox
1. Malware (default)	7	0	0	3
2. Malware (compressed)	8	0	0	4
3. Malware (basic encoding)	9	0	0	1
4. Malware (basic encoding/compressed)	9	0	0	3
5. Malware (advanced encoding)	6	0	0	4
6. Malware (advanced encoding/compressed)	10	0	0	2
Total	**49**	**0**	**0**	**17**

[redacted]

Threats	Stopped	Notified	Warned	Inbox
1. Malware (default)	0	0	6	4
2. Malware (compressed)	0	0	8	4
3. Malware (basic encoding)	0	0	6	4
4. Malware (basic encoding/compressed)	0	0	8	4
5. Malware (advanced encoding)	0	0	7	3
6. Malware (advanced encoding/compressed)	0	0	8	4
Total	**0**	**0**	**43**	**23**

5. Conclusion

Sandbox products can be used in a network in a number of ways. This test was designed to evaluate the products for deployment in an email environment, whereby attachments will be submitted to the sandbox and checked for malicious behaviour before being either deleted or released to the recipient.

The test included threats that are commonly used to target organisations of all sizes. Each threat was changed in a number of ways to see how easy it would be to bypass the detection mechanisms of each sandbox. The methods and tools used are available to every internet user with an interest in exploiting systems. They are zero cost to obtain and uncomplicated to use. However, the results demonstrate that more sophisticated options are currently not necessary for many attacks to succeed. Indeed, in many cases no evasion techniques were necessary to avoid detection. Despite the limitations of this test, which used a small number of attacks generated by freely-available tools, none of the products provided close to full protection. Overall the sandbox detection rates were poor, with around one third of the threats managing to evade detection.

[redacted] product stood out as superior in comparison, due to its higher and wider detection rates. [redacted] appears to be particularly poor, however it is possible (as with all products) that its behaviour could be tuned in a real deployment.

For example, [redacted] product warned about 43 files. If the technology is capable of detecting that these files are suspicious enough to warrant a warning, it's likely that the product behaviour could be changed to block such files. That said, [redacted] still failed to detect one third of the files.

False positives were not a significant issue for the sandboxes. [redacted] product misclassified one legitimate email, but the strict weighting on our ratings system reduced its Legitimate Accuracy Rating to 75 per cent. Enterprises may consider this to be a lesser or greater issue and can apply their own weights to the results published here.

Appendices

Appendix A: 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 B: 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 product was configured according to its vendor's recommendations.
- The test was conducted throughout January2017.
- Targeted attacks were selected and verified by SE Labs.
- SE Labs conducted this endpoint test using physical systems.

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.