SE LABS ®

EDR
PROTECTION
RANSOMWARE

Advanced Security Test Report

Palo Alto Networks
Cortex XDR

ONLINE REPORT

AUGUST 2026

SE LABS tested **Palo Alto Networks Cortex XDR** against a range of ransomware attacks designed to extort victims. These attacks were realistic, using the same tactics and techniques as those used against victims in recent months.

Target systems, protected by **Cortex XDR**, were attacked by testers acting in the same way as we observe ransomware groups to behave.

Attacks were initiated from the start of the attack chain, using phishing email links and attachments, as just two examples. Each attack was run from the very start to its obvious conclusion, which means attempting to steal, encrypt and destroy sensitive data on the target systems.

Contents

Introduction	04
Executive Summary	05
Advanced Security Test Award	05
1. How We Tested	06
Threat Responses	07
Attack Details	08
2. Total Accuracy Ratings	09
3. Protection Ratings (Ransomware Direct Attacks)	10
Protection Scores	11
Protection Details	11
4. Response Details (Ransomware Deep Attacks)	12
Detection Accuracy Rating	13
Legitimate Software Rating	13
5. Conclusion	14
Appendices	15
Appendix A: Threat Intelligence (Ransomware Deep Attacks)	15
Appendix B: Response Details	17
Appendix C: Legitimate Interaction Ratings	18
Appendix D: Terms Used	20
Appendix E: FAQs	20
Appendix F: Ransomware Deep Attack Details	21
Appendix G: Product Version	23

Document version 1.0 Written 25th August 2026

SE LABS

Management
Chief Executive Officer Simon Edwards
Chief Operations Officer Marc Briggs
Chief Human Resources Officer Magdalena Jurenko
Technical Director Thomas Bean

Testing Team
Nikki Albesa
Daniel Botez
Solandra Brewster
Billy Coyne
Jarred Earlington
Gia Gorbold
Jake Hearn
Anila Johny
Cameron Love
Jeremiah Morgan
Julian Owusu-Abrokwa
Joseph Pike
Enejda Torba

Marketing
Sara Claridge
Isobel Edwards

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 https://selabs.uk/blog/
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.

© 2026 SE Labs Ltd

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.

Ransomware vs. Endpoint Security

Results from the largest public ransomware test

Ransomware is the most visible, most easily understood cyber threat affecting businesses today. Paralysed computer systems mean stalled business and loss of earnings. On top of that, a ransom demand provides a clear, countable value to a threat. A demand for "one million dollars!" is easier to quantify than the possible leak of intellectual property to a competitor.

One reason why ransomware is so 'popular' is that the attackers don't have to produce their own. They outsource the production of ransomware to others, who provide Ransomware as a Service (RAAS). Attackers then usually trick targets into running it, or at least into providing a route for the attackers to run it for them. Artificial intelligence systems make the creation of such social engineering attacks easier, cheaper and more effective than ever before.

Given the global interest and terror around ransomware, we have created a comprehensive test that shows how effective security products are when faced with the whole range of threats posed by ransomware itself and the criminal groups operating in the shadows.

In this report we have taken two main approaches to assessing how well products can detect and protect against ransomware.

Ransomware Deep Attacks

For the first part of this test, we analysed the common tactics of ransomware gangs and created two custom gangs that use a wider variety of methods. In all cases we run the attack from the very start, including attempting to access targets with stolen credentials or other means. We then move through the system and sometimes the network, before deploying the ransomware as the final payload.

In the first two attacks for each group, we gain access and deploy ransomware onto the target immediately. In the third, fourth and fifth attacks we move through the network and deploy ransomware on a target deeper into the network.

The ransomware payloads used in this part of the report were known files from all of the families listed in **Attack Details** on page 8**.**

This test shows a product's ability to track the movement of the attacker through the entire attack chain. We disable the product's protection features and rely on its detection mode for this part of the test. The results demonstrate how incident response teams can use the product to gain visibility on ransomware attacks.

Ransomware Direct Attacks

The second part of the test takes a wide distribution of known malware and adds variations designed to evade detection. We've listed the ransomware families used in **Attack Details** on page 8. We sent each of these ransomware payloads directly to target systems using realistic techniques, such as through email social engineering attacks. This is a full but short attack chain. In this part of the test, we ensure any protection features are enabled in the product.

If products can detect and protect against the known version of each of these files, all well and good. But if they also detect and block each ransomware's two variations then we can conclude that the protection available is more proactive than simply reacting to yesterday's unlucky victims.

Executive Summary

We tested **Palo Alto Networks Cortex XDR** against a wide range of ransomware attacks. These included recent, prevalent ransomware payloads alongside new, never-before-seen variations.

We examined the product's abilities to:

- **Detect and protect against known ransomware**
- **Detect and protect against new ransomware variants**
- **Track full network breaches**
- **Detect deployment of ransomware on internal targets.**

Legitimate files were used alongside the threats to measure any false positive detections or other sub-optimum interactions.

Palo Alto Networks Cortex XDR excelled in both detection and protection. It was able to track attacks on targets other than the endpoint and provided thorough insight into the full network breaches that concluded with ransomware deployments. It also provided complete protection against all of the ransomware attacks.

By also allowing all legitimate applications, **Palo Alto Networks Cortex XDR** achieved a Total Accuracy Rating of 100% and an AAA rating for advanced security.

Executive Summary

Palo Alto Networks Cortex XDR		
	Accuracy Score	**Rating (%)**
Detection Accuracy	440/440	100%
Protection Accuracy	2,544/2,544	100%
Legitimate Accuracy	680/680	100%
Total Accuracy	3,664/3,664	100%

- The Detection rating shows how effective the product was at detecting the ransomware attacks.

Advanced Security Test Award

The following product wins the SE Labs award:

Palo Alto Networks
Cortex XDR

1. How We Tested

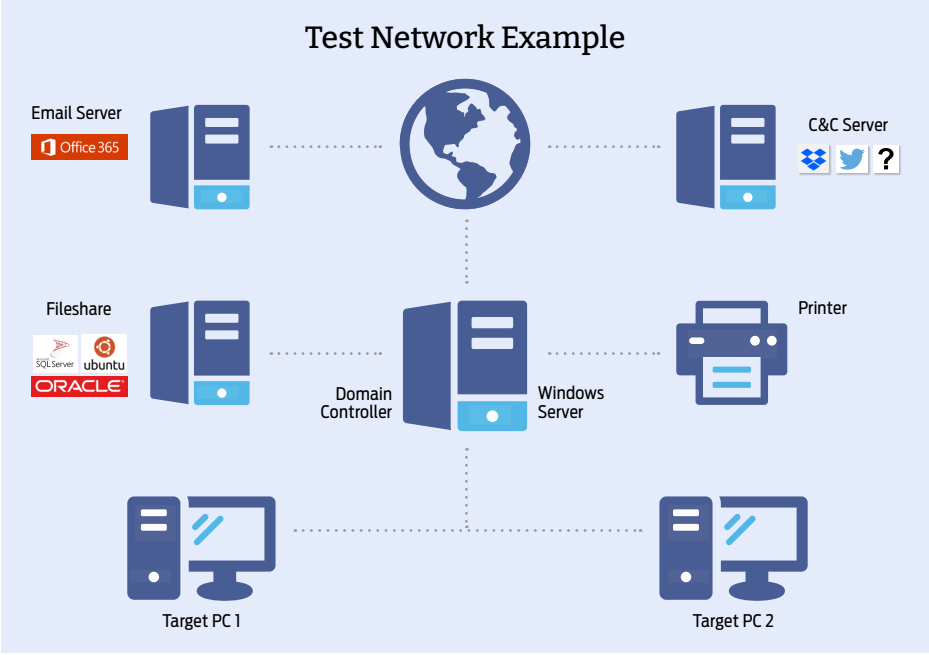
Testers can't assume that products will work a certain way, so running a realistic advanced security test means setting up real networks and hacking them in the same way that real adversaries behave.

In the diagram on the right you will see an example network that contains workstations, some basic infrastructure such as file servers and a domain controller, as well as cloud-based email and a malicious command and control (C&C) server, which may be a conventional computer or a service such as Dropbox, Twitter, Slack or something more imaginative.

As you will see in the **Threat Responses** section on page 7, attackers often jump from one compromised system to another in so-called 'lateral movement'. To allow products to detect this type of behaviour the network needs to be built realistically, with systems available, vulnerable and worth compromising.

It is possible to compromise devices such as enterprise printers and other so-called 'IoT' (internet of things) machines, which is why we've included a representative printer in the diagram.

The techniques that we choose for each test case are largely dictated by the real-world behaviour of online criminals. We observe their tactics and replicate what they do in this test. To see more details about how the specific attackers behaved, and how we copied them, see **Attack Details** on page 8 and, for a really detailed drill down on the details, **Appendix A: Threat Intelligence** on pages 15-16 and **Appendix F: Attack Details** on pages 21.



● **This example of a test network shows one possible topology and ways in which enterprises and criminals deploy resources**

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.

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.

Attack Details

When testing services against targeted attacks it is important to ensure that the attacks used are relevant. Anyone can run an attack randomly against someone else. It is the security vendor's challenge to identify common attack types and to protect against them. As testers, we need to generate threats that in some way relate to the real world.

All of the attacks used in this test are valid ways to compromise an organisation. Without any security in place, all would succeed in attacking the target. Outcomes would include systems infected with ransomware, remote access to networks and data theft.

But we didn't just sit down and brainstorm how we would attack different companies. Instead we used current threat intelligence to look at what the bad guys have been doing over the last few years and copied them quite closely. This way we can test the services' abilities to handle similar threats to those faced by global governments, financial institutions and national infrastructure.

The graphic on this page shows a summary of the attack groups that inspired the targeted attacks used in this test. If a service was able to detect these then there's a good chance they are on track to detect similar attacks in the real world. If they fail, then you might take their bold marketing claims about defeating hackers with a pinch of salt.

Attacker/ APT Group	Method	Target	Details
Akira	Compromised Credentials		First appearing in early 2023, Akira steals data and demands huge ransoms from large companies.
Babuk	Phishing		Multi-platform ransomware using a public RaaS model.
BlackBasta	Spear Phishing		Evolution of Hermes/ Ryuk/ Conti ransomware, practising double extortion.
Chaos	Phishing, Infected External Drives		Wiper malware with a low barrier to entry that began by masquerading as encryption ransomware before evolving into advanced detection-evading encryptors.
DragonForce	Public-Facing Remote Desktop Server		Hacktivist-style operation that expanded into criminal motives.
Fog	Compromised VPN Accounts		Multi-pronged extortion operation focussing (but not limited to) US entities.
LockBit	RDP, Brute Force		An RaaS threat used across a variety of industries and continues to be prolific.
Medusa	Compromised Legitimate Services		Originated as a closed variant, before progressing to an affiliate model employing a double encryption model and using Living-Off-The-Land techniques.
Phobos	RDP, Brute Force		Phobos operates as Ransomware as a Service (RaaS) and goes beyond simple data encryption, leading to disruption of critical systems.
Sodinokibi	Phishing, Vulnerabiltiy Exploitation		Russian-based group that expanded extremely quickly after its inception. RaaS approach leading to many variants.
Trigona	Spear Phishing, Exploiting MSSQL Servers		Multi-extortion group with Windows and Linux variants. First observed in June 2022, now inactive due to takedown by the Ukrainian Cyber Alliance.

KEY							
	Education		Financial Industries		Government		Infrastructure
	Large Organisations		Legal Industries		Manufacturing		Public Figures
	Public Organisations		Retail		Small Businesses		

2. 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 chart.

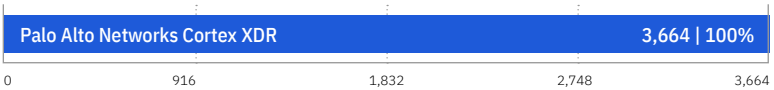
The chart below takes into account not only the product's ability to detect and protect against threats, but also its handling of non-malicious objects such as web addresses (URLs) and applications.

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. 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.

Scoring a product's response to a potential breach requires a granular method, which we outline in **4. Response Details (Ransomware Deep Attacks)** on page 12.

Total Accuracy Ratings



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

SE LABS PRESENTS

THE-C2

MARCH 2027

Connecting business with cyber security

The-C2 is an exclusive, invite-only threat intelligence event that connects multinational business executives with the cutting edge of the cyber security industry. The event enables frank and open discussion of the developing digital threat landscape among global security leaders.

The-C2 is hosted by SE Labs, the world's leading security testing lab. Its unique position in the industry provides a route to understanding both the developing threat landscape and the evolving security measures for defending against attackers.

THE-C2.COM

3. Protection Ratings (Ransomware Direct Attacks)

The following results relate to the direct ransomware attacks, in which ransomware payloads are sent directly to targets in realistic ways, such as via phishing emails.

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.

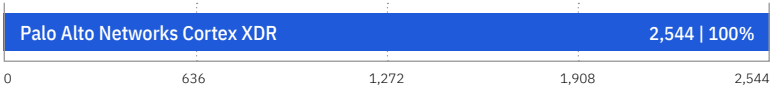
These ratings are based on our opinion of how important these different outcomes are. You may have a different view on how seriously you treat a 'Compromise' or 'Neutralisation without complete remediation'. If you want to create your own rating system, you can use the raw data from **Protection Details (Ransomware Direct Attacks)** on page 11 to roll your own set of personalised ratings.

Protection Score (Ransomware Direct Attacks)

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 Score



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

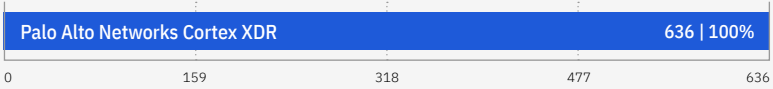
Protection Details (Ransomware Direct Attacks)

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.

Protection Details

Product	Detected	Blocked	Neutralised	Compromised	Protected
Palo Alto NetworksCortex XDR	636	636	0	0	636



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

4. Response Details (Ransomware Deep Attacks)

In this test security products are exposed to attacks, which comprise multiple stages. The perfect product will detect all relevant elements of an attack.
The term 'relevant' is important, because sometimes detecting one part of an attack means it's not necessary to detect another.

For example, in the table below certain stages of the attack chain have been grouped together. As mentioned in **2. Total Accuracy Ratings**, these groups are as follows:

Delivery/Execution (+10)
If the product detects either the delivery or execution of the initial attack stage then a detection for this stage is recorded.

Action (+10)
When the attack performs one or more actions, while remotely controlling the target, the product should detect at least one of those actions.

Privilege Escalation/Action (+10)
As the attack progresses there will likely be an attempt to escalate system privileges and to perform more powerful and insidious actions. If the product can detect either the escalation process itself, or any resulting actions, then a detection is recorded.

Lateral Movement/Action (+10)
The attacker may attempt to use the target as a launching system to other vulnerable systems.

If this attempt is discovered, or any subsequent action, a detection is reported.

The Detection Rating is calculated by adding points for each group in a threat chain that is detected. When at least one detection occurs in a single group, a 'group detection' is recorded and 10 points are awarded. Each test round contains one threat chain, which itself contains four groups (as shown below), meaning that complete visibility of each attack adds 40 points to the total value.

A product that detects the delivery of a threat, but nothing subsequently to that, wins only 10 points, while a product that detects delivery and action, but not privilege escalation or lateral behaviours, is rated at 20 for that test round.

Understanding Detection Groups

		First Group		Second Group	Third Group		Fourth Group	
Incident No:	Detection	Delivery	Execution	Action	Escalation	PE Action	Lateral Movement	Lateral Action
1	✓	✓	✓	—	✓	✓	✓	✓
2	✓	—	✓	✓	✓	✓	✓	✓
3	✓	—	✓	✓	✓	✓	✓	✓
4	✓	✓	✓	—	✓	✓	✓	✓

Attacker/ Apt Group	Number of Incidents	Attacks Detected	Delivery/ Execution	Action	Privilege Escalation/Action	Lateral Movement Action
Dragonfly & Dragonfly 2	4	4	4	2	4	4

Elements of the attack chain are put into groups. For example, the Delivery and Execution stages of an attack are in the same group. Similarly, we group the Post Escalation stage with the Post Escalation Action (PE Action) stage. When we count detections we look to see at least one detection (tick) in each group. One or two detections in a group is a success.

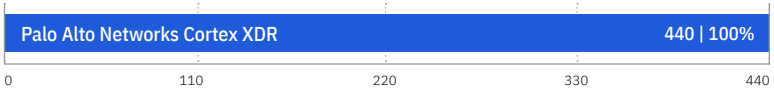
In this example we have four test cases, which we call 'incidents'. In Incident No. 1 there was a detection recorded for the delivery of the threat and when it was executed. These two results count as one detection. In Incident No. 2 the threat delivery was not detected, but its execution was. This also counts as one detection.

When no detection is registered in any part of a group the result will be a 'miss'. In Incident No. 1 there was no detection when the attacker performed the 'Action' stage of the attack. This is a miss for the product. In fact, this product only detected two of the four Action stages, which is why the Response Details table shows '2' in the Action column.

Detection Accuracy Rating

To understand how we calculate these ratings, see **Appendix B: Response Details** on page 17.

Detection Accuracy



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

Legitimate Software Rating

These ratings indicate how accurately the product classifies legitimate applications and URLs, while also taking into account the interactions that the 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.

Legitimate Software Rating

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

5. Conclusion

This report looks at how effectively a security product can protect against a wide range of ransomware attacks. It also investigates the product's capabilities in tracking the behaviour of attackers that use ransomware as a final payload.

Ransomware Deep Attacks

Attackers used to rely on random and widespread ransomware deployment to extort payment from as many hapless victims as they could. Today's ransomware attacks are much more targeted and persistent aimed at large organisations that can pay in the millions of dollars.

This test mimics how attacks breach large organisations by running full, advanced attacks against the target systems and by installing malware at the end of each attack. We wanted to assess how well **Palo Alto Networks Cortex XDR** could track the hacking attacks through the network, as well as registering the ransomware attacks at the end.

The methods of attacking the target systems were a combination of tactics used by a number of different ransomware groups. You can see a summary of these in **Appendix A: Threat Intelligence**, on pages 15 and 16, and a full rundown of each in **Appendix F: Ransomware Deep Attack Details**.

We jumped from target systems to others on the internal network (moving laterally) and ran ransomware on these deeper targets. We did this for the first two test cases for both the first and second group tests, where we installed ransomware directly on the main target systems. This is why the Lateral Movement and the Lateral Action results for these cases are not applicable (N/A).

Palo Alto Networks Cortex XDR's 100% Detection Accuracy Rating is a good indicator of its ability to deliver timely alerts when breaches occur. Timeliness is critical given how quickly targeted attacks, especially those with ransomware payloads, can damage endpoints and infect other resources in a network. It gives security professionals the best chance to investigate and handle a persistent attacker in real time.

However, the product's 100% Protection Accuracy Rating also points to its value as a defender - blocking the delivery of almost all of the threats before any damage is done. In the one instance when the breach shifted from delivery to execution, **Palo Alto Networks Cortex XDR** took over the task of containing and mitigating the damage caused by the BTQLock ransomware variant.

Ransomware Direct Attacks

In the second part of the test, we used a large group of ransomware attack files. The files formed a combination of malicious software both known and unknown by security researchers. Our goal was to see how well a product could identify ransomware that has already been analysed by security experts, as well as new, never-before-seen variations that represent potential future attacks.

We modified files from prevalent ransomware families using techniques designed to make the malware different (although the malware would perform the same malicious activities). We had 636 ransomware files – about a third of which were originals and the others, variations. We exposed target systems to these ransomware files using very direct methods of attack, such as sending the malware (or links to the malware) via phishing emails.

Palo Alto Networks Cortex XDR detected and blocked every single ransomware file, including all of the new variants. It blocked ransomware files upon delivery 99% of the time. Nine ransomware files managed to reach the execution stage but the product still managed to restrict their malicious actions. A fast-moving series of pop-ups showed the product attempting to contain file encryption and log corruption when the ransomware was successfully deployed. In the end, the threats were completely remediated and product ejected the ransomware files from the systems.

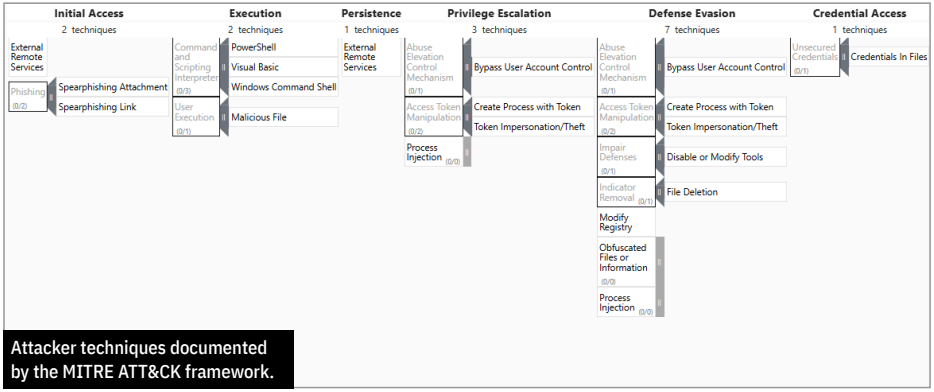
Palo Alto Networks Cortex XDR performed exceptionally well at protecting against known and new variants of ransomware, as well as tracking network attacks the concluded with ransomware payloads. It is fully deserving of its AAA award.

Appendices

Appendix A: Threat Intelligence (Ransomware Deep Attacks)

SE Labs
Ransomware Deep Attack Group 1

After the system was completely compromised, testers deployed ransomware from a sub-set of groups listed in **Attack Details**, page 8.

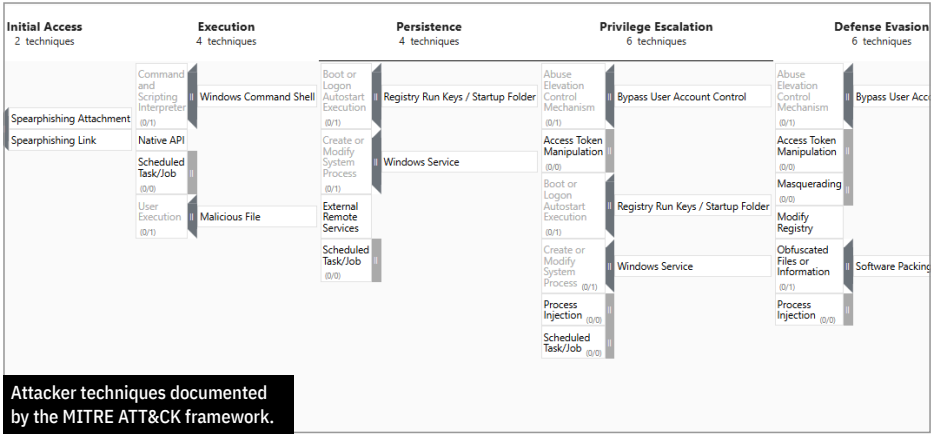


Attacker techniques documented by the MITRE ATT&CK framework.

Example SE Labs Ransomware Deep Attack Group 1						
Delivery	**Execution**	**Action**	**Privilege Escalation**	**Post-Escalation**	**Lateral Movement**	**Lateral Action**
Spear Phishing Attachment	Powershell	Process Injection	Access Token Manipulation - Create Process with Token	Disable or Modify Tools	N/A	N/A
	Obfuscated Files or Information	System Information Discovery		File Deletion		
	Malicious File	System Service Discovery		Exfiltration Over C2 Channel		
	Windows Command Shell			Data Destruction		
	Asymmetic Cryptography			Data Encrypted for Impact		
				Inhibit System Recovery		
				Service Stop		

SE Labs
Ransomware Deep Attack Group 2

After the system was completely compromised, testers deployed ransomware from a sub-set of groups listed in **Attack Details**, page 8.



Attacker techniques documented by the MITRE ATT&CK framework.

Example SE Labs Ransomware Deep Attack Group 2						
Delivery	**Execution**	**Action**	**Privilege Escalation**	**Post-Escalation**	**Lateral Movement**	**Lateral Action**
Spear Phishing Link	Malicious File	Process Discovery	Access Token Manipulation	Credentials In Files	N/A	N/A
		System Information Discovery		Exfiltration Over C2 Channel		
		Permission Groups Discovery		Data Destruction		
		System Network Configuration Discovery		Data Encrypted for Impact		
		File and Directory Discovery		Inhibit System Recovery		
		System Owner/User Discovery		Service Stop		

Appendix B: Response Details

Group 1

Incident No:	Detection	Delivery	Execution	Action	Escalation	PE Action	Lateral Movement	Lateral Action
1	✓	✓	✓	✓	✓	✓	N/A	N/A
2	✓	✓	✓	✓	✓	✓	N/A	N/A
3	✓	✓	✓	✓	✓	✓	✓	✓
4	✓	✓	✓	✓	✓	✓	✓	✓
5	✓	✓	✓	✓	✓	✓	✓	✓

Group 2

Incident No:	Detection	Delivery	Execution	Action	Escalation	PE Action	Lateral Movement	Lateral Action
1	✓	✓	✓	✓	✓	✓	N/A	N/A
2	✓	✓	✓	✓	✓	✓	N/A	N/A
3	✓	✓	✓	✓	✓	✓	✓	✓
4	✓	✓	✓	✓	✓	✓	✓	✓
5	✓	✓	✓	✓	✓	✓	✓	✓

Attacker/APT Group	Number of Incidents	Attacks Detected	Delivery/ Execution	Action	Privilege Escalation/ Action	Lateral Movement/ Action
Group 1	5	5	5	5	5	3
Group 2	5	5	5	5	5	3
TOTAL	**10**	**10**	**10**	**10**	**10**	**6**

- This data shows how the product handled different group stages of each APT. The Detection column shows the basic level of detection.

Detection Accuracy Rating Details

Attacker/ APT Group	Number of Incidents	Attacks Detected	Group Detections	Detection Rating
Group 1	5	5	18	220
Group 2	5	5	18	220
TOTAL	**10**	**10**	**36**	**440**

- Different levels of detection, and failure to detect, are used to calculate the Detection Rating.

Group Detections

We record detections in groups, as described above in Understanding Detection Groups. To get an overview of how a product handled the entire set of threats we then combine these detections into 'Group Detections'.

In a test with four incidents and four detection groups (Delivery/Execution; Action; Escalation/PE Action; and Lateral Movement/Lateral Action) the maximum score would be 16. This is because for each of the four threats a product that detects everything would score 4.

Our overall Detection Rating is based on the number of Detection Groups achieved.

Appendix C: 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 Software Rating** on page 13.

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)	Click to allow/block (no recommendation)	None (allowed)
Palo Alto Networks Cortex XDR	100	0	0

● **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
High Impact	40
Medium Impact	40
Low Impact	20

Appendix D: 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 E: 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 We are a customer considering buying or changing our endpoint protection and/ or endpoint detection and response (EDR) product. Can you help?

A Yes, we frequently run private testing for organisations that are considering changing their security products.
Please contact us at **info@selabs.uk** for more information.

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

- The test was conducted between 3rd and 15th July 2026.
- 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 F: SE Labs Ransomware Deep Attack Details

Group 1

Incident No.	Delivery	Execution	Action	Privilege Escalation	Post-Escalation	Lateral Movement	Lateral Action
1	Spear Phishing Attachment	Powershell	Process Injection	Access Token Manipulation - Create Process with Token	Disable or Modify Tools	N/A	N/A
		Obfuscated Files or Information	System Information Discovery		File Deletion		
		Malicious File	System Service Discovery		Exfiltration Over C2 Channel		
		Windows Command Shell			Data Destruction		
		Asymmetic Cryptography			Data Encrypted for Impact		
					Inhibit System Recovery		
					Service Stop		
2	Spear Phishing Link	Malicious File	File and Directory Discovery	Access Token Manipulation - Token Impersonation/Theft Process Injection	Credentials In Files	N/A	N/A
			System Information Discovery		Data Destruction		
			Process Discovery		Data Encrypted for Impact		
			System Owner/User Discovery		Inhibit System Recovery		
			Internet Connection Discovery		Service Stop		
			Query Registry		Exfiltration over C2 Channel		
			Domain Account				
			Domain Groups				
			Network Share Discovery				
3	Spear Phishing Link	Powershell	Query Registry	Access Token Manipulation - Create Process with Token	Modify Registry	External Remote Services	Exfiltration over C2 Channel
		Malicious File	System Information Discovery		Service Stop		Data Destruction
		Windows Command Shell	System Location Discovery - System Language Discovery				Data Encrypted for Impact
		Asymmetric Cryptography	File Deletion				Inhibit System Recovery
							Service Stop
4	Spear Phishing Attachment	Windows Command Shell	System Information Discovery	Access Token Manipulation - Token Impersonation/Theft Process Injection	Disable or Modify Tools	Lateral Tool Transfer	Exfiltration over C2 Channel
		Malicious File	Permission Groups Discovery - Domain Groups		Inhibit System Recovery	Remote Desktop Protocol	Data Destruction
		Visual Basic	Process Injection				Data Encrypted for Impact
			File Deletion				Inhibit System Recovery
							Service Stop

Incident No.	Delivery	Execution	Action	Privilege Escalation	Post-Escalation	Lateral Movement	Lateral Action
5	Spear Phishing Attachment	Windows Command Shell	System Information Discovery	Bypass User Account Control; Valid Accounts	Ingress Tool Transfer	External Remove Services	Exfiltration over C2 Channel
		Malicious File	Query Registry		Modify Registry		Data Destruction
							Data Encrypted for Impact
							Inhibit System Recovery
							Service Stop

Group 2

Incident No.	Delivery	Execution	Action	Privilege Escalation	Post-Escalation	Lateral Movement	Lateral Action
1	Spear Phishing Link	Malicious File	Process Discovery	Access Token Manipulation	Credentials In Files	N/A	N/A
			System Information Discovery		Exfiltration Over C2 Channel		
			Permission Groups Discovery		Data Destruction		
			System Network Configuration Discovery		Data Encrypted for Impact		
			File and Directory Discovery		Inhibit System Recovery		
			System Owner/User Discovery		Service Stop		
2	Spear Phishing Link	Malicious File	Process Discovery	Bypass User Account Control	Data From Local System	N/A	N/A
		Windows Command Shell	System Information Discovery	Valid Accounts	Exfiltration Over C2 Channel		
		Masquerading	Account Discovery - Local Account		Credentials from Web Browsers		
		Software Packing	System Network Configuration Discovery		Data Destruction		
		Native API			Data Encrypted for Impact		
		Symmetric Cryptography			Inhibit System Recovery		
					Service Stop		
3	Spear Phishing Link	Malicious File	Process Discovery	Bypass User Account Control	Data From Local System	External Remote Services	Exfiltration Over Alternative Protocol
		Windows Command Shell	System Information Discovery	Valid Accounts	Exfiltration Over C2 Channel		Automated Collection
		Software Packing	Network Share Discovery		Modify Registry		Data Destruction
		Obfuscated Files or Information	System Service Discovery				Data Encrypted for Impact
							Inhibit System Recovery
							Service Stop

Incident No.	Delivery	Execution	Action	Privilege Escalation	Post-Escalation	Lateral Movement	Lateral Action
4	Spear Phishing Attachment	Malicious File	Process Discovery	Bypass User Account Control	Credentials in Files	External Remote Services	Exfiltration Over Alternative Protocol
		Windows Command Shell	System Information Discovery	Valid Accounts	System Owner/User Discovery		Data Destruction
		Software Packing	Credentials from Web Browsers		Modify Registry		Data Encrypted for Impact
		Masquerading			Windows Service		Inhibit System Recovery
							Service Stop
5	Spear Phishing Link	Malicious File	Process Discovery	Bypass User Account Control	Scheduled Task	Lateral Tool Transfer	Exfiltration Over C2 Channel
		Windows Command Shell	System Information Discovery	Valid Accounts	Registry Run Keys / Startup Folder		Automated Collection
		Obfuscated Files or Information	Credentials from Web Browsers				Data Destruction
			System Owner/User Discovery				Data Encrypted for Impact
							Inhibit System Recovery
							Service Stop

Appendix G: Product Version

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)
Palo Alto Networks	Cortex XDR	9.2.0.120	9.2.0.120

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.