BREACH DETECTION SYSTEMS COMPARATIVE REPORT

Security

OCTOBER 19, 2017

Author – Thomas Skybakmoen

Tested Products

Check Point Software Technologies 15600 Next Generation Threat Prevention & SandBlast™ (NGTX) Appliance R77.30

Cisco FirePower 8120 v.6 & Cisco AMP v.5.1.9.10430

FireEye Network Security NX 10450 v7.9.2 & EX 8400 v7.9.0

FireEye Network Security 6500NXES-VA v7.9.2Fortinet FortiSandbox-2000E v.FSA 2.4.1 & FortiClient (APT Agent) v.5.6.0.1075

Lastline Enterprise v7.25

Trend Micro Deep Discovery Inspector Model 4000 v3.8 SP5 & OfficeScan (OSCE) v.12.0.1807

Environment

Breach Detection Systems Test Methodology v4.0

Overview

Implementation of breach detection systems (BDS) can be a complex process, with multiple factors affecting the overall *Security Effectiveness* of the system. The following factors should be considered over the course of the useful life of the BDS:

- Detection rate
- Anti-evasion capabilities (resistance to common evasion techniques)
- Device stability and reliability
- Time to detect

In order to determine the relative *Security Effectiveness* of BDS products on the market and to facilitate accurate product comparisons, NSS Labs has developed a unique metric:

***Security Effectiveness** = Breach Detection Rate[1] x Evasions x Stability and Reliability*

Figure 1 – Security Effectiveness Formula

By focusing on *Security Effectiveness* as a whole instead of on detection rate alone, NSS is able to factor in the ease with which defenses can be bypassed, as well as the reliability of the system under test.

Product	Detection Rate	Anti-Evasion Rating	Stability and Reliability	Security Effectiveness
Check Point	99.7%	97.0%	100%	96.7%
Cisco	99.0%	97.0%	100%	96.0%
FireEye NX & EX	98.4%	83.0%	100%	81.7%
FireEye NXES-VA	96.6%	83.0%	100%	80.2%
Fortinet	99.0%	99.0%	100%	98.0%
Lastline	100.0%	100.0%	100%	100.0%
Trend Micro	100.0%	100.0%	100%	100.0%

Figure 2 – Security Effectiveness

As part of the initial BDS test setup, solutions are configured and tuned as deemed necessary by the vendor. Every effort is made to deploy policies that ensure the optimal combination of *Security Effectiveness* and performance, as would be the aim of a typical customer deploying the solution in a live environment. This provides readers with the most useful information on key BDS *Security Effectiveness* and performance capabilities based on their expected usage.

Evasion techniques are a means of disguising and modifying attacks in order to avoid detection and blocking by security products. Resistance to evasion is a critical component in a BDS. If a single evasion is missed, an attacker can utilize an entire class of exploits and/or malware to circumvent the BDS, rendering it virtually useless.

[1] The breach detection rate is calculated as the percentage of all malware and exploits that were detected under test (drive-by exploits, social exploits, HTTP malware, SMTP malware, and offline infections).

Average Time to Detect

BDS products attempts to discover malware that is bypassing traditional security controls by examining various indicators of compromise (IoCs) to determine whether files are malicious. For this reason, it is important for enterprises to consider the average time a product takes to detect attacks when they are evaluating the product's overall security performance.

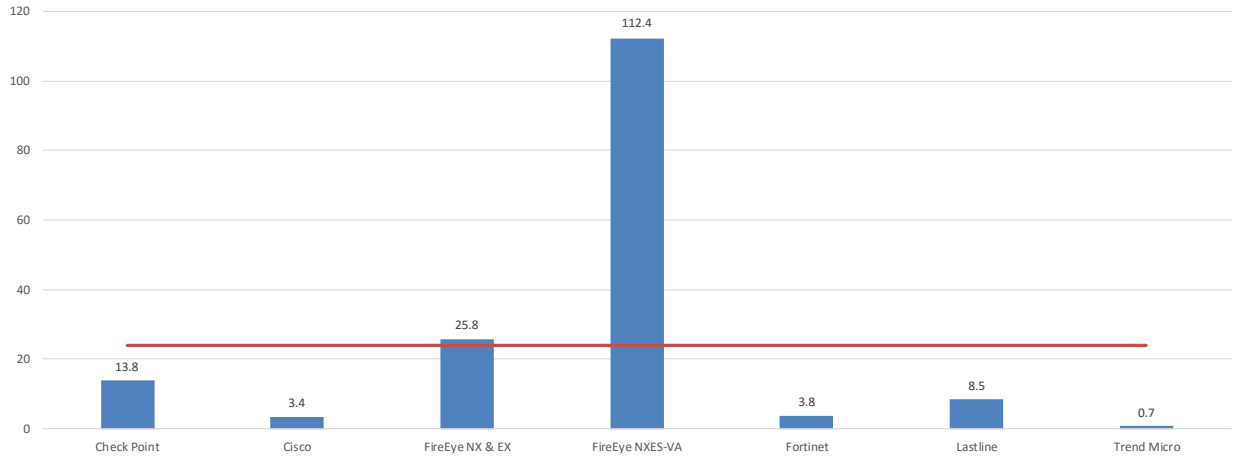

Figure 3 – Average Time to Detect in Minutes

Table of Contents

Table of Figures

Analysis

The threat landscape is evolving constantly; attackers are refining their strategies and increasing both the volume and intelligence of their attacks. Additionally, enterprises now must defend against targeted persistent attacks (TPAs). In the past, servers were the main target; however, attacks against desktop client applications are now mainstream and present a clear danger to organizations.

Tuning

Security products are often complex. Vendors have responded to customer needs by simplifying the user interface and security policy selection to meet the usability needs of a broadening user base. Indeed, many organizations accept and deploy the default settings, understanding these to be the best recommendations from the vendor. NSS research has found that BDS often require little or no tuning, and in fact, several are provided with little or no tuning options. However, where possible, all BDS products are tuned prior to testing to eliminate false positives and provide the most appropriate coverage for the systems to be protected. Typically, tuning is carried out by a vendor's system engineers, but where this is not possible, NSS engineers will perform any necessary tuning. NSS engineers may also adjust the configuration of a system under test, where specific characteristics of the system or its configuration interfere with the normal operation of any of the tests, or where the results obtained from those tests would, in the opinion of NSS engineers, misrepresent the true capabilities of the system. Every effort is made to ensure the optimal combination of *Security Effectiveness* and performance, as would be the aim of a typical customer deploying the system in a live network environment.

Detection Rate

This test utilizes threats and attack methods that exist in the wild and that are currently being used by cybercriminals and other threat actors. For live testing, NSS employs a unique live test harness, the CAWS Continuous Security Validation Platform, to measure how well security products protect against "drive-by" exploits that target client applications.

The CAWS Continuous Security Validation Platform captures thousands of suspicious URLs per day from threat data generated from NSS and its customers, as well as data from open-source and commercial threat feeds. This list of URLs is optimized and assigned to victim machines, each of which has a unique combination of operating system (including service pack/patch level), browser, and client application. For details on live testing, please refer to the latest Security Stack (Network) Test Methodology, which can be found at www.nsslabs.com.

The ability of the product to detect and report successful infections in a timely manner is critical to maintaining the security and functionality of the monitored network. Infection and transmission of malware should be reported quickly and accurately, giving administrators the opportunity to contain the infection and minimize impact on the network.

As response time is critical in halting the damage caused by malware infections, the system under test should be able to detect known samples, or analyze unknown samples, and report on them within 24 hours of initial infection and command and control (C&C) callback. Any system that does not alert on an attack, infection, or C&C callback within the detection window will not receive credit for the detection.

Time to Detect

Given that the purpose of these products is to detect breaches, both detection rate and time to detect are critical. NSS testing revealed that most products were able to detect breaches within 60 minutes, but some products took several hours to detect the same breaches. For further details on breach detection over time, please see the individual Test Reports.

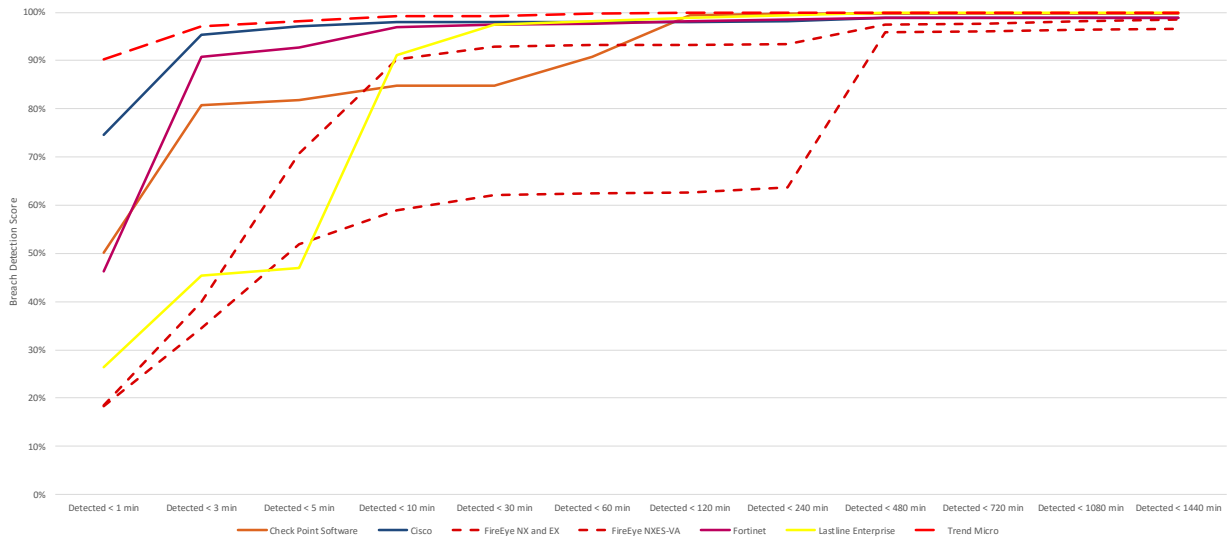

Figure 4 – Breach Detection over Time (I)

Product	Detected < 1 min	Detected < 3 min	Detected < 5 min	Detected < 10 min	Detected < 30 min	Detected < 60 min
Check Point	50.1%	80.7%	81.7%	84.9%	84.9%	90.9%
Cisco	74.7%	95.3%	97.1%	97.9%	97.9%	97.9%
FireEye NX & EX	18.5%	39.9%	70.8%	90.3%	93.0%	93.2%
FireEye NXES-VA	18.3%	34.5%	52.0%	59.0%	62.1%	62.4%
Fortinet	46.2%	90.9%	92.7%	96.9%	97.4%	97.7%
Lastline	26.4%	45.4%	47.0%	91.1%	97.4%	98.2%
Trend Micro	90.3%	97.1%	98.2%	99.2%	99.2%	99.7%

Figure 5 – Breach Detection over Time (II)

Product	Detected < 120 min	Detected < 240 min	Detected < 480 min	Detected < 720 min	Detected < 1080 min	Detected < 1440 min
Check Point	99.5%	99.7%	99.7%	99.7%	99.7%	99.7%
Cisco	97.9%	98.2%	99.0%	99.0%	99.0%	99.0%
FireEye NX & EX	93.2%	93.5%	97.4%	97.7%	98.2%	98.4%
FireEye NXES-VA	62.7%	63.7%	95.8%	96.1%	96.3%	96.6%
Fortinet	98.2%	98.4%	99.0%	99.0%	99.0%	99.0%
Lastline	99.0%	99.5%	100.0%	100.0%	100.0%	100.0%
Trend Micro	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Figure 6 – Breach Detection over Time (III)

Weighted Detection Score

Once a product's time to detect attacks is known, NSS can calculate its detection score. Furthermore, this detection score can be better represented by assigning it a weighting, depending on whether the product detects attacks more quickly or more slowly over a 24 hour period. By applying this weighting to the product's detection score, we can more accurately represent its "zero-hour" detection capability.

Figure 7 and Figure 8 depict the weighted protection score for each system under test. For each product, the first column depicts its original detection score and the second column depicts its detection score with weighting applied.

Detection over Time (Minutes)	Check Point Software		Cisco		FireEye NX and EX		FireEye NXES-VA	
	Detection Score	Weighted Detection Score	Detection Score	Weighted Detection Score	Detection Score	Weighted Detection Score	Detection Score	Weighted Detection Score
< 1	50.1%	15%	74.7%	22%	18.5%	5%	18.3%	5%
< 3	80.7%	37%	95.3%	48%	39.9%	17%	34.5%	15%
< 5	81.7%	52%	97.1%	66%	70.8%	30%	52.0%	25%
< 10	84.9%	63%	97.9%	78%	90.3%	41%	59.0%	32%
< 30	84.9%	68%	97.9%	84%	93.0%	47%	62.1%	36%
< 60	90.9%	71%	97.9%	87%	93.2%	49%	62.4%	38%
< 120	99.5%	72%	97.9%	89%	93.2%	51%	62.7%	39%
< 240	99.7%	73%	98.2%	90%	93.5%	52%	63.7%	39%
< 480	99.7%	73%	99.0%	90%	97.4%	52%	95.8%	39%
< 720	99.7%	73%	99.0%	90%	97.7%	52%	96.1%	39%
< 1080	99.7%	74%	99.0%	90%	98.2%	52%	96.3%	39%
< 1440	99.7%	74%	99.0%	90%	98.4%	52%	96.6%	40%

Figure 7– Weighted Detection Score (I)

Detection over Time (Minutes)	Fortinet		Lastline		Trend Micro	
	Detection Score	Weighted Detection Score	Detection Score	Weighted Detection Score	Detection Score	Weighted Detection Score
< 1	46.2%	14%	26.4%	8%	90.3%	26%
< 3	90.9%	39%	45.4%	20%	97.1%	53%
< 5	92.7%	56%	47.0%	29%	98.2%	72%
< 10	96.9%	68%	91.1%	40%	99.2%	84%
< 30	97.4%	74%	97.4%	46%	99.2%	90%
< 60	97.7%	77%	98.2%	49%	99.7%	93%
< 120	98.2%	78%	99.0%	51%	100.0%	95%
< 240	98.4%	79%	99.5%	52%	100.0%	95%
< 480	99.0%	79%	100.0%	52%	100.0%	96%
< 720	99.0%	80%	100.0%	52%	100.0%	96%
< 1080	99.0%	80%	100.0%	52%	100.0%	96%
< 1440	99.0%	80%	100.0%	52%	100.0%	96%

Figure 8 – Weighted Detection Score (II)

Malware Delivered by Drive-by Exploits

Drive-by exploits are defined as malicious software that is designed to take advantage of existing deficiencies in hardware or software systems, such as vulnerabilities or bugs. Figure 9 displays each product's detection score for malware delivered by drive-by exploits.

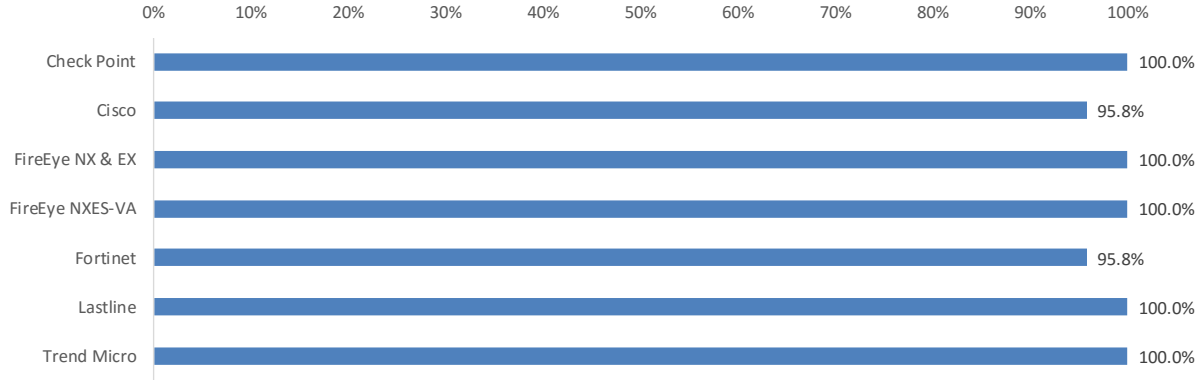

Figure 9 – Malware Delivered by Drive-by Exploits

Malware Delivered by Social Exploits

Social exploits are defined as malicious software that is designed to take advantage of existing deficiencies in hardware or software systems, such as vulnerabilities or bugs, for example, a user is deceived into clicking a malicious link to download and execute malware. Figure 10 displays each product's detection score for malware delivered by social exploits.

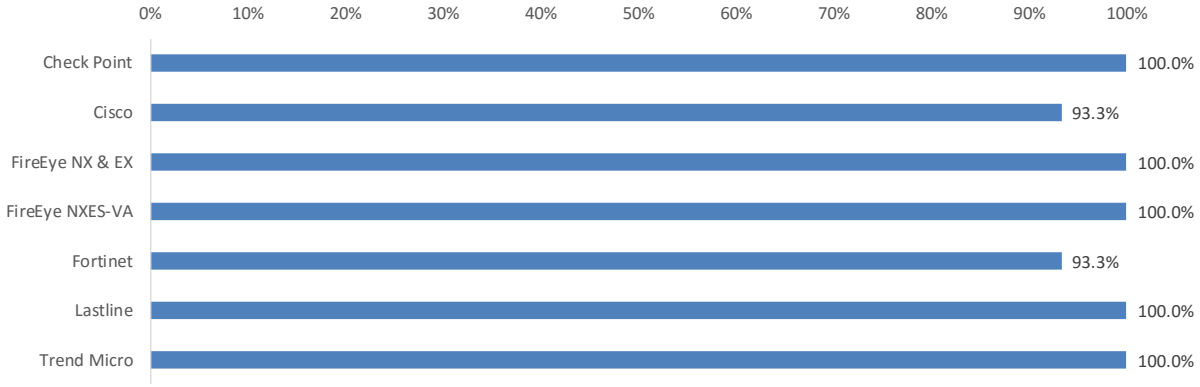

Figure 10 – Malware Delivered by Social Exploits

Malware Delivered over HTTP

Figure 11 displays each product's detection score for malware using the HTTP protocol as its transport mechanism, i.e., malware delivered through a web browser.

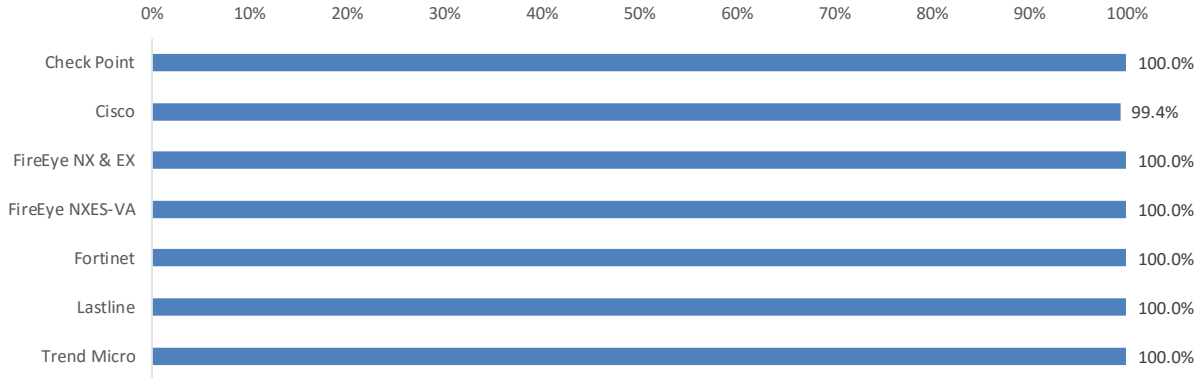

Figure 11 – Malware Delivered over HTTP

Malware Delivered over Email

Figure 12 displays each product's detection score for malware that uses email (SMTP) as a transport mechanism, such as a malicious email attachment.

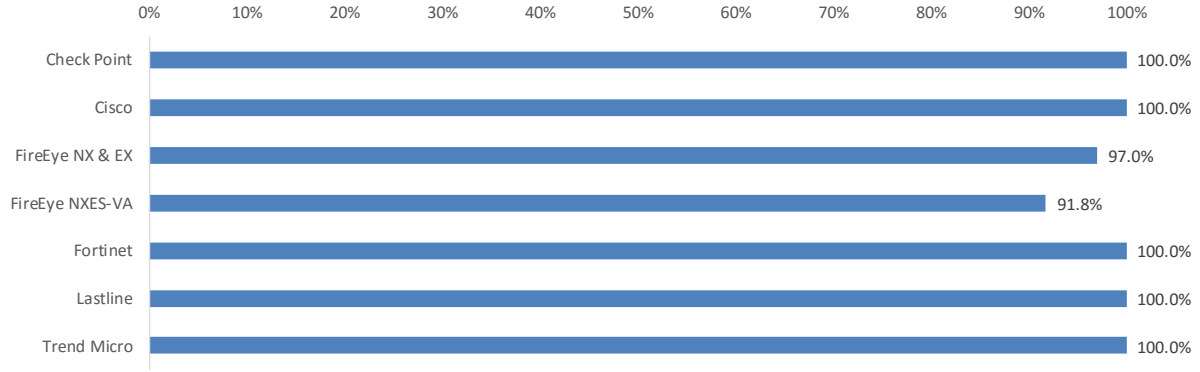

Figure 12 – Malware Delivered over Email

Off-line Infections

Remote users with mobile devices can become infected while outside the protection of the corporate network. When infected devices are subsequently reattached to the corporate network, the infection can spread. Figure 13 displays each product's detection score for off-line infections.

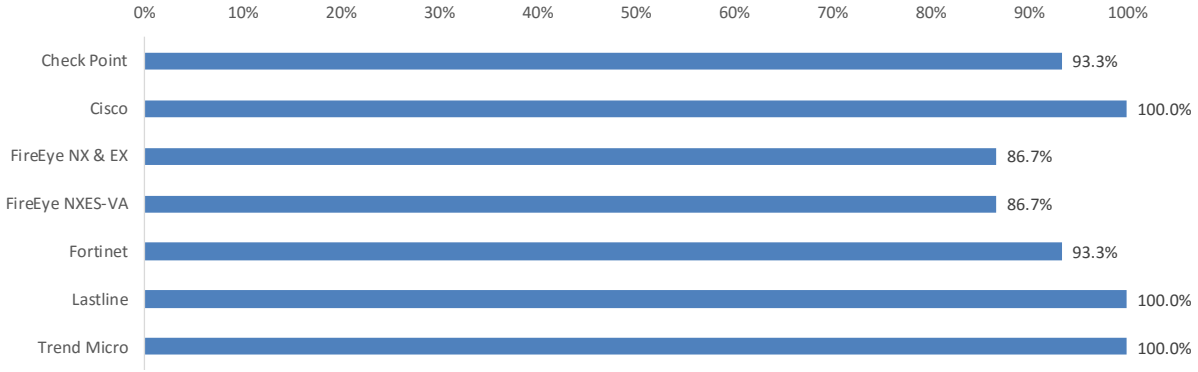

Figure 13 – Malware Delivered with Off-line Infections

Resistance to Evasion Techniques

Evasion techniques are a means of disguising and modifying attacks at the point of delivery to avoid detection by security products. Failure of a security device to correctly identify a specific type of evasion potentially allows an attacker to use an entire class of exploits and/or malware for which the device is assumed to have protection.

Providing protection results without fully factoring in evasion can be misleading. The more classes of evasion that are missed, such as packers, compressors, virtual machine, and sandbox, the less effective the device. For example, it is better to miss all techniques in one evasion category, such as packers, than one technique in each category, which would result in a broader attack surface.

A product's effectiveness is significantly handicapped if it fails to detect exploits and malware that employ obfuscation or evasion techniques, and the NSS product guidance is adjusted to reflect this.

Figure 14 and Figure 15 provide evasion resistance scores for each of the tested products. The impact of missing an evasion is weighted as high, medium, or low. A product's evasion score impacts its *Security Effectiveness* score and therefore its placement on the SVM. For details on which packing and compressing techniques were used and or missed, please see the individual Test Reports.

Product	Packers & Compressors	Virtual Machine	Sandbox	HTML Obfuscation	HTML5 Obfuscation
Check Point	100.0%	100.0%	100.0%	100.0%	100.0%
Cisco	100.0%	100.0%	94.1%	100.0%	100.0%
FireEye NX & EX	98.0%	100.0%	100.0%	100.0%	100.0%
FireEye NXES-VA	98.0%	100.0%	100.0%	100.0%	100.0%
Fortinet	100.0%	100.0%	98.0%	100.0%	100.0%
Lastline	100.0%	100.0%	100.0%	100.0%	100.0%
Trend Micro	100.0%	100.0%	100.0%	100.0%	100.0%

Figure 14 – Evasion Resistance (I)

Product	HTML5 HeapSpray	Web Socket Connection	HTTP Evasion	Layered Evasions
Check Point	100.0%	50.0%	100.0%	100.0%
Cisco	100.0%	100.0%	100.0%	100.0%
FireEye NX & EX	100.0%	0.0%	80.0%	100.0%
FireEye NXES-VA	100.0%	0.0%	80.0%	100.0%
Fortinet	100.0%	100.0%	100.0%	100.0%
Lastline	100.0%	100.0%	100.0%	100.0%
Trend Micro	100.0%	100.0%	100.0%	100.0%

Figure 15 – Evasion Resistance (II)

Stability and Reliability

Long-term stability is important, since a failure can result in serious breaches remaining undetected and thus not being remediated. These tests verify the stability of the system under test along with its ability to maintain *Security Effectiveness* while under normal load and while detecting malicious traffic. Products that cannot sustain logging of legitimate traffic (or that crash) while under hostile attack will not pass.

The system is required to remain operational and stable throughout these tests and to detect 100% of previously detected traffic, raising an alert for each. If any malicious traffic passes undetected, caused by either the volume of traffic or by the system failing for any reason, this will result in a fail.

Product	Detection Under Extended Attack	Power Failure and Persistence of Data
Check Point	PASS	PASS
Cisco	PASS	PASS
FireEye NX & EX	PASS	PASS
FireEye NXES-VA	PASS	PASS
Fortinet	PASS	PASS
Lastline	PASS	PASS
Trend Micro	PASS	PASS

Figure 16 – Stability and Reliability

Security Effectiveness

The *Security Effectiveness* of a device is determined by factoring the results of evasions testing and stability and reliability testing into its detection rate.

Product	Detection Rate	Anti-Evasion Rating	Stability and Reliability	Security Effectiveness
Check Point	99.7%	97.0%	100%	96.7%
Cisco	99.0%	97.0%	100%	96.0%
FireEye NX & EX	98.4%	83.0%	100%	81.7%
FireEye NXES-VA	96.6%	83.0%	100%	80.2%
Fortinet	99.0%	99.0%	100%	98.0%
Lastline	100.0%	100.0%	100%	100.0%
Trend Micro	100.0%	100.0%	100%	100.0%

Figure 17 – Security Effectiveness

Test Methodology

Breach Detection Systems Test Methodology v4.0

A copy of the test methodology is available on the NSS Labs website at www.nsslabs.com.

Contact Information

NSS Labs, Inc.
3711 South MoPac Expressway
Building 1, Suite 400
Austin, TX 78746-8022
USA
info@nsslabs.com
www.nsslabs.com