TEST METHODOLOGY

Software-Defined Wide Area Network (SD-WAN)

April 6, 2018

V1.2

Table of Contents

1 Introduction

1.1 The Need for the Software-Defined Wide Area Network (SD-WAN)

Since the advent of the Internet and, more importantly, since the emergence of the digital economy, the wide area network (WAN) has been a relatively static and predictable tool for enterprises. Without multi-site, assured performance connectivity, many applications would not function optimally, if at all. When architected well, WAN connectivity is expensive but generally predictable. Often, point-to-point circuits are built based on calculated performance loads, and they are accompanied by commercial service-level agreements (SLAs) that provide some guarantee of connectivity and performance. However, drawbacks to these circuits include time to provision (it can take weeks to set up a new circuit) and ongoing service fees. Historically, enterprises have had few options when it comes to commercial-grade service requirements. Alternatively, the consumer broadband offerings provide high-speed access links (e.g., cable and fiber to the home), which are affordable and widely available; however, they do not provide the service assurances that accompany commercial links.

The marriage of software-defined networking (SDN) benefits to WAN technology yields the software-defined wide area network (SD-WAN), which allows consumer-grade links (or links without assured performance) to be leveraged for business-class services. Through the use of common VPN capabilities and the separation of data and control planes within SDN, software-managed connections can be established and managed between multiple sites over any number of link types (e.g., fixed circuit, DSL, cable, mobile, MPLS, and so on) without the operational challenges of having to manage different links. SD-WANs easily establish links, manage traffic over links according to application or service requirements (e.g., VoIP vs. Facebook), and enforce policy control capabilities (e.g., limit web-based traffic to 50% of a given link), and they are simple to manage. SD-WAN options are part router, part WAN optimization, and part firewall. In addition, some SD-WAN offerings provide robust security, which makes for a compelling alternative to the multiple appliance approach that is often required at remote locations.

1.2 About This Test Methodology

NSS Labs test reports are designed to address the challenges faced by enterprise security and IT professionals in selecting and managing security products. The scope of this methodology includes:

- Performance
- Stability and reliability
- Total cost of ownership (TCO)
- Security effectiveness

As SD-WAN technology is expected to provide and manage the remote links connecting a site with either headquarters locations or the public Internet, its stability and reliability is imperative. Therefore, regardless of any security capabilities, the main requirement of any SD-WAN is that it must be as stable, as reliable, as fast, and as flexible as the edge device that it is replacing.

The following capabilities are considered essential in an SD-WAN:

- Traditional routing and policy control features, including:
    - Basic application identification and policy controls
    - Stateful networking controls
    - Virtual private network (VPN)

- Highly resilient remote office connectivity
- Prioritization of critical applications
- Remote configuration capabilities
- A consistent or improved performance experience for users

Tuning: For solutions submitted with security and inspection abilities, tuning of the security configuration will follow procedures from the NSS Labs Next Generation Firewall Test Methodology.

1.3 Inclusion Criteria

In order to encourage the greatest participation, and to allay any potential concerns of bias, NSS invites all vendors claiming SD-WAN capabilities to submit their products at no cost. Vendors with major market share, as well as challengers with new technology, will be included.

Where possible, the SD-WAN should be supplied as three separate appliances, with the appropriate number of physical interfaces capable of achieving the required level of connectivity for one HQ and two branch office locations. Each office location will have two links; one of which will be MPLS, and the other will be a standard broadband connection. Once installed in the test environment, solutions will be configured appropriately for the use cases as documented below. Figure 1 depicts the topology of an enterprise test environment.

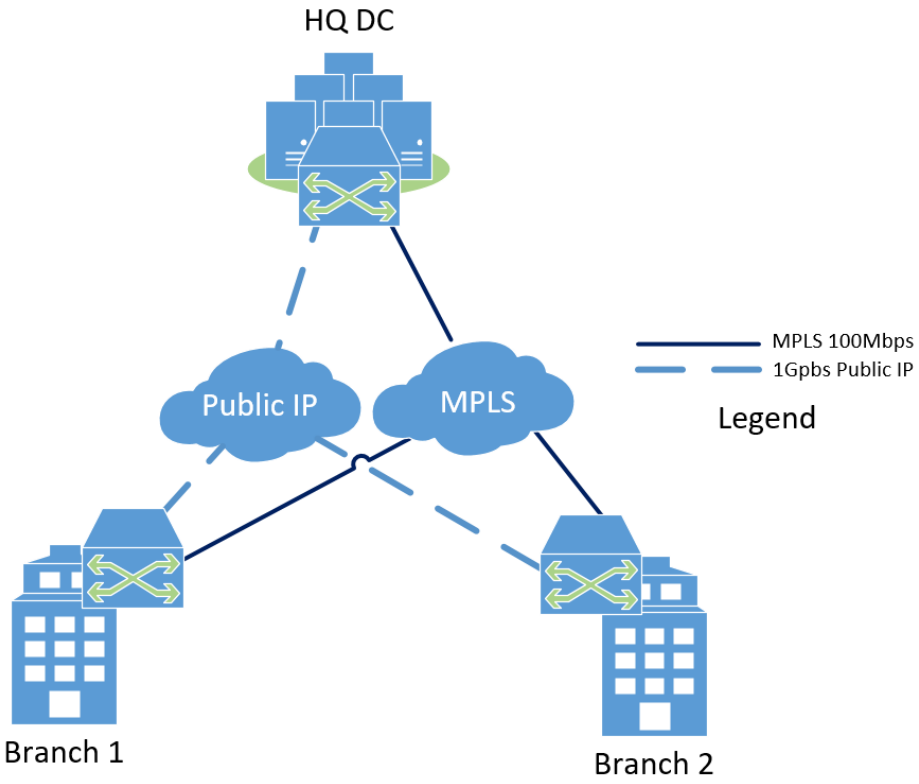

Figure 1 – Topology of Enterprise Test Environment

The WAN environment is provisioned with behavioral characteristics similar to those typically encountered over normal WAN link states. Next, the test harness baseline is recorded to ensure consistent behavior. Finally, the vendor solution is deployed and each test case is measured against the baseline.

2 Product Guidance

NSS issues summary product guidance based on evaluation criteria that is important to enterprise networking and security professionals. The evaluation criteria are as follows:

- **Network Policy Effectiveness** – The purpose of an SD-WAN is to securely connect multiple site networks through policy and routing over assured links and over commercial broadband links and to identify and manage applications and performance between multiple sites.
- **Security effectiveness** –SD-WAN solutions that have security and inspection capabilities are expected to be able to inspect and block malicious content, including evasion technique classes used to bypass inspection appliances today.
- **Stability and reliability** – Long-term stability is particularly important for an inline device, where failure can produce network outages.
- **Performance** – Correctly sizing an SD-WAN is essential.
- **Value** – Customers often seek low TCO and high effectiveness and performance rankings.

Products are listed in rank order according to their guidance rating.

2.1 Recommended

A *Recommended* rating from NSS indicates that a product has performed well and deserves strong consideration. Only the top technical products earn a *Recommended* rating from NSS, regardless of market share, company size, or brand recognition.

2.2 Neutral

A *Neutral* rating from NSS indicates that a product has performed reasonably well and should continue to be used if it is the incumbent within an organization. Products that earn a *Neutral* rating from NSS deserve consideration during the purchasing process.

2.3 Caution

A *Caution* rating from NSS indicates that a product has performed poorly. Organizations using one of these products should review their security posture and other threat mitigation factors, including possible alternative configurations and replacement. Products that earn a *Caution* rating from NSS should not be short-listed or renewed.

3 Network Policy Effectiveness

This section verifies that the solution performs and is capable of enforcing network configuration policies effectively. NSS analysis of SD-WAN technology is conducted by incrementally building upon a baseline configuration (simple routing with no policy restrictions) to create a complex, real-world, multiple-zone configuration that supports many addressing modes, policies, applications, and inspection engines. Any exception events that occur must also be accompanied by detailed log information in order to enable both network and security forensics. Additionally, administrative visibility is critical. To facilitate analysis and troubleshooting, NSS requires the logging of any function that results in dropped traffic. Every test scenario employs routing. In order to pass traffic and provide measurable results, a device must be capable of handling common enterprise routing protocol configurations (BGP, OSPF.)

The SD-WAN must support persistent policy capabilities that ensure the prioritization of application traffic is managed securely. The SD-WAN must be able to manage policy across multiple interfaces/zones. At a minimum, the firewall must provide a "trusted" internal interface, an "untrusted" external/Internet interface, and (optionally) one or more DMZ interfaces. In addition, a dedicated management interface (virtual or otherwise) is preferred.

Each of the following test cases intends to measure and record a vendor solution's ability to perform according to the parameters of the test case. In each case, the scenario is built to mimic real-world deployments and traffic experiences in addition to common configurations in order to reveal how well the configuration policy and overall solution perform.

Each of the following test cases possesses specific industry-accepted metrics that record the performance capabilities of solutions. In test cases where more than one measure is taken, individual test scenarios will possess unique metrics.

3.1.1 Baseline Policy

This is a routed configuration with an "allow all" policy. The solution is expected to pass traffic between all sites without incident.

3.1.2 Simple Policies

These are simple outbound and inbound policies that allow basic browsing and email access so that internal clients can access untrusted, external networks without giving external clients the ability to access internal network(s). A number of combinations will be run to assess a solution's ability to behave according to configuration.

Solution under test should provide the ability to control which applications and protocols are passed based on test case criteria.

A combination of policies from the following lists will be tested to simulate enterprise application prioritization:

Latency-sensitive applications and protocols (directed across service-assured link):

- VOIP (SIP)
- H.323
- RTSP
- RTP
- RTCP

Latency-tolerant applications and protocols (directed across public IP link)

- HTTP
- IMAP
- POP3
- SMB
- LDAP
- RDP
- SMTP
- FTP
- SNMPV2
- NetBIOS
- NTP
- SYSLOG
- SSH
- DNS
- SCP
- SFTP
- Telnet
- RADIUS
- TACACS+
- TFTP
- Social media applications

3.1.3 Remote Initial Configuration

SD-WAN helps organizations achieve operational savings is by enabling remote configuration of new locations rather than requiring engineers to be on site for setup. Many vendors offer "zero-touch provisioning (ZTP)," where no onsite engineering expertise is required other than the ability to correctly power up and connect a device to the right internal and external links. Once online, the device will call "home" whether that is the HQ or a cloud configuration service, where it will gather and download the operational configuration information. This can also be achieved via a central management system (CMS) and will be noted as such in the solution's test report. Both the ability to configure remotely and time to configure the site will be recorded. These metrics will be components of the operational cost model that is used to calculate the solution's overall TCO.

The solution is expected to be remotely configurable (solely via network link or "remote" keyboard) in order to receive a pass for the remote configuration test case.

4 Performance Effectiveness and Consistency

This section measures the performance of the solution using various traffic conditions that provide metrics for real-world performance. Individual implementations will vary based on usage; however, these quantitative metrics provide a gauge as to whether a particular solution is appropriate for a given environment and present a normative data set that is equal and comparable across all solutions.

All tests will be performed across the VPN links established according to the use case topology. Additionally, the harness baseline validation that is conducted prior to the introduction of the vendor solution will be documented in the test report.

The impairment test cases selected are the most stressful scenarios in which a WAN technology would ever be placed. The results of these tests and the application measures captured during each test case are indicative of a solution's ability to withstand punishing performance scenarios. In addition to these impairment scenarios, NSS will be recording standard traffic performance, and this will be included in the final test report and scorecard.

4.1 WAN Impairment

A critical function of any SD-WAN is the identification and correct routing of traffic based on policy prioritization (autonomous or configured), which is influenced by network performance characteristics (e.g., variability, latency, jitter, etc.). Link impairment tests subject connected links to testing that represents real-world conditions encountered by enterprises today. Latency, jitter, and variability are all commonly encountered on public link technologies. The solution is expected to route prioritized traffic according to the quality settings for an application based on link performance. If impairment is encountered over a link where priority traffic is being routed, the SD-WAN must either throttle lower-priority traffic to assure the quality of the prioritized application, or it must reroute traffic across an alternate link if it is available, depending on network cost and expected service levels.

The various impairments described in this section are applied in order to assess the adaptability of the SD-WAN solution..(i.e., path selection, QoS, failover, app steering, congestion avoidance). These tests go beyond packet loss, jitter, and latency to investigate misordered packets, link saturation, packet duplication, congestion, and bursty traffic .The goal is to verify how the solution adapts to varied impairments.

In each test case, background traffic will be introduced to populate links with sufficient activity as to represent typical enterprise network communications. Additionally, traffic-specific flows will be introduced in order to capture accurate measurements, including RTP MOS for voice over IP, relative MOS for video, and one-way delay for RTP. These measurements provide guidance as to how sensitive applications behave across a tested SD-WAN solutions when the solution is subjected to various impairments.

Different permutations of the following link impairments will be used to verify all SD-WAN performance features.

4.1.1 Packet Loss

This refers to the amount of data packets that do not reach their destination. The test will simulate various loss levels with a uniform distribution.

Any sustained packet loss should be identified by a solution and the links managed accordingly based on application or policy.

4.1.2 Packet Delay Variation (PDV)

This refers to the variation in delay of unidirectional, consecutive packets that flow between two hosts over an IP path. This impairment is most often referred to as jitter. The test will simulate a Gaussian delay with minimum and maximum values.

This test measures how a solution handles PDV impact on voice or video, both of which are delay-intolerant beyond the buffer capacity of the application.

4.1.3 Packet Reorder

This refers to the delivery of data packets out of the order in which they were originally sent. This test will simulate a periodic distribution of selected packets.

Solutions should be able to identify out-of-sequence packets and manage these according to the configured policy. This condition impacts voice and video applications significantly if the delay time exceeds the application buffer.

4.1.4 Packet Duplication

This test refers to a packet that is duplicated somewhere on the network and is received twice at the receiving host. This test will simulate a duplication of selected packets in a uniform distribution

Solutions should take the next-in-sequence packet and drop the duplicates in order to preserve the whole frame sequence.

4.1.5 Accumulate and Burst

This test refers to the accumulation of packets in a memory queue. Packets are burst once a configured condition is met. This test will simulate accumulation of packets until the buffer queue has (N) packets or until packets have been accumulated for a specified time (T) with a minimum interburst gap.

Burst capability testing is particularly stressful as it stresses network buffering capacity. Sustained burst behaviors reveal that there is link congestion or other issues and SD-WAN solutions should alter paths based on known good alternate paths.

4.1.6 “First-Mile” Network Behavior

The Policer limits the data rate of a network stream to ensure that it does not exceed the specified limits. This impairment emulates link saturation. Traffic policing is in accordance with the Metro Ethernet Forum (MEF) bandwidth profiles for Ethernet services. The solution is expected to utilize bandwidth appropriately.

To replicate congestion in the “first mile,” impairments will be applied to the HQ to site 1 links. These are the links from the HQ DC SD-WAN to the emulated aggregation point (MPLS and Public IP networks).c

4.1.7 “Last-Mile” Network Behavior

The Policer limits the data rate of a network stream to ensure that it does not exceed the specified limits. This impairment emulates link saturation. Traffic policing is in accordance with the Metro Ethernet Forum (MEF) bandwidth profiles for Ethernet services It is expected that the solution will utilize bandwidth appropriately

To replicate congestion in the “last mile,” impairments will be applied to the HQ to site 1 links. These are the links from the emulated aggregation point (MPLS and Public IP networks) to the branch 1 SD WAN site.

4.1.8 WAN Link Failover

In this test, an established link between sites is interrupted and the SD-WAN is observed to determine whether it is handling stateful session in a manner that is transparent to users. At the point of failure, the routed link should be altered without loss or interruption to the applications using the link. The only exception to this would be where the failover links are experiencing an exhaustion event and prioritized applications are consuming all available bandwidth based on policy configuration, which could impact the non-critical applications.

A WAN solution must be able to operate resiliently in spite of link outages. The session or application should be able to continue without interruption and there should be no noticeable user impact.

4.1.9 Dynamic Path Selection with SLA Measurements

The goal of this test is to determine how long it takes for traffic to move to an available link when preconfigured impairments are applied. To limit any visible user impact, the solution should support path decisions on a per-flow basis according to available links and according to the conditions that exist on those links.

The time to select a new path will be measured, as will any impact to applications.

4.1.10 Path Conditioning

SD-WAN solutions employ various techniques to condition WAN links in order to ensure reliability of data transmission. Some solutions employ packet duplication, forward error correction, bonding, or load balancing.

The vendor solution should identify the best path and guarantee priority policies (application, protocol, or other configured guidance) over known good links with other traffic transmitted as best effort.

4.1.11 Quality of Service (QoS)

Quality of service is important for business-critical applications such as voice and video. These applications must be prioritized if a link has bad performance indicators. This test will measure QoS using voice traffic and video stream. The test will include MOS scores for video and call measurements for VoIP.

The vendor solution should manage traffic according to configured QoS classification settings.

4.1.12 Link Saturation and Congestion

Global QoS awareness can prevent congestion during the last mile of data delivery; thus, the goal of this test is to ensure reliable use of bandwidth by the controller in the SD-WAN solution.

4.2 Application-Aware Traffic Steering

This test will verify how the solution directs various application traffic flows for applications besides video and VoIP. Behavior will be observed and recorded to establish whether voice/video and data are sent over the same link once impairments are applied and to establish which application takes precedence.

4.2.1 Application Control Policies

These are complex outbound and inbound policies that consist of many rules, objects, and applications and that verify whether the solution is capable of accurately determining the correct application (regardless of port/protocol used), and then taking the appropriate action.

- VoIP

- Business video (Cisco Spark, Microsoft Skype Professional, etc.)
- Popular social networking websites (web applications)
- Other basic legacy applications (e.g., FTP, Telnet)

For each application, NSS will test the solution's ability to perform the following functions:

4.2.1.1 Steer

The solution should be able to accurately identify the application and direct it over the correct link according to configured policy.

4.2.1.2 Block Specific Action (Depends on Application)

For example, in the case of instant messaging, the solution should allow text communications while blocking file transfers.

4.2.1.3 Drop Low-Priority Application During Congestion Event

The solution should recognize when link exhaustion occurs and ensure that high-priority applications take precedence over low-priority applications.

4.3 Site-to-Site VPN (IPSec, SSL, or Other)

An SD-WAN manages links between sites as VPN tunnels, thus providing secure connections over public links. While IPsec VPN is the dominant technology for securing site-to-site connections, testing will allow commonly accepted VPN configurations.

4.4 Raw Packet Processing Performance (UDP Throughput)

This test uses UDP packets of varying sizes generated by traffic generation appliances. A constant stream of the appropriate packet size—with variable source and destination IP addresses transmitting from a fixed source port to a fixed destination port—is transmitted bi-directionally through each port pair of the solution. Each packet contains dummy data and is targeted at a valid port on a valid IP address on the target subnet. The percentage load and frames per second (fps) figures across each inline port pair are verified by network monitoring tools before each test begins. Multiple tests are run and averages are taken where necessary.

This traffic does not attempt to simulate any form of real-world network condition. No TCP sessions are created during this test, and there is very little for the flow or policy engine to do.

The goal of this test is to determine the raw packet processing capability of each inline port pair of the solution, as well as its effectiveness at forwarding packets quickly in order to provide the highest level of network performance and with the lowest latency.

Every performance test will be conducted over the established tunnels between HQ DC and Branch Site 1 locations.

4.4.1 64-Byte Packets

Maximum 1,488,000 frames per second per Gigabit of traffic. This test determines the ability of a solution to process packets from the wire under the most challenging packet processing conditions.

4.4.2 128-Byte Packets

Maximum 844,000 frames per second per Gigabit of traffic

4.4.3 256-Byte Packets

Maximum 452,000 frames per second per Gigabit of traffic

4.4.4 512-Byte Packets

Maximum 234,000 frames per second per Gigabit of traffic. This test provides a reasonable indication of the ability of a solution to process packets from the wire on an "average" network.

4.4.5 1024-Byte Packets

Maximum 119,000 frames per second per Gigabit of traffic

4.4.6 1514-Byte Packets

Maximum 81,000 frames per second per Gigabit of traffic. This test has been included to demonstrate how easy it is to achieve good results using large packets. Readers should use caution when taking into consideration those test results that only quote performance figures using similar packet sizes.

4.5 Maximum Capacity

The use of traffic generation appliances allows NSS engineers to create "real-world" traffic at link-appropriate speeds as a background load for the tests.

The goal of these tests is to stress the policy or inspection engine and determine how it handles high volumes of TCP connections per second, application layer transactions per second, and concurrent open connections. All packets contain valid payload and address data, and these tests provide an excellent representation of a live network at various connection/transaction rates.

Note that in all tests, the following critical "breaking points"—where the final measurements are taken—are used:

- **Excessive concurrent TCP connections** – Latency within the solution is causing an unacceptable increase in open connections.
- **Excessive concurrent HTTP connections** – Latency within the solution is causing excessive delays and increased response time.
- **Unsuccessful HTTP transactions –** Normally, there should be zero unsuccessful transactions. Once these appear, it is an indication that excessive latency within the solution is causing connections to time out.

4.5.1 Theoretical Maximum Concurrent TCP Connections

This test is designed to determine the maximum concurrent TCP connections of the solution with no data passing across the connections. This type of traffic would not typically be found on a normal network, but it provides the means to determine the maximum possible concurrent connections figure.

An increasing number of Layer 4 TCP sessions are opened through the solution. Each session is opened normally and then held open for the duration of the test as additional sessions are added up to the maximum possible. Load is increased until no more connections can be established, and this number is recorded.

4.5.2 Maximum TCP Connections per Second

This test is designed to determine the maximum TCP connection rate of the solution with one byte of data passing across the connections. This type of traffic would not typically be found on a normal network, but it provides the means to determine the maximum possible TCP connection rate.

An increasing number of new sessions are established through the solution, ramped slowly to determine the exact point of failure. Each session is opened normally, one byte of data is passed to the host, and then the session is closed immediately. Load is increased until one or more of the breaking points defined earlier is reached.

4.5.3 Maximum HTTP Connections per Second

This test is designed to determine the maximum TCP connection rate of the solution with a one-byte HTTP response size. The response size defines the number of bytes contained in the body, excluding any bytes associated with the HTTP header. A one-byte response size is designed to provide a theoretical maximum HTTP connections per second rate.

Client and server are using HTTP 1.0 without keep-alive; the client will open a TCP connection, send one HTTP request, and close the connection. This ensures that all TCP connections are closed immediately upon the request being satisfied; and thus any concurrent TCP connections will be caused purely as a result of latency the solution introduces on the network. Load is increased until one or more of the breaking points defined earlier is reached.

4.5.4 Maximum HTTP Transactions per Second

This test is designed to determine the maximum HTTP transaction rate of the solution with a one-byte HTTP response size. The object size defines the number of bytes contained in the body, excluding any bytes associated with the HTTP header. A one-byte response size is designed to provide a theoretical maximum connections per second rate.

Client and server are using HTTP 1.1 with persistence, and the client will open a TCP connection, send 10 HTTP requests, and close the connection. This ensures that TCP connections remain open until all 10 HTTP transactions are complete, thus eliminating the maximum connection per second rate as a bottleneck (one TCP connection = 10 HTTP transactions). Load is increased until one or more of the breaking points defined earlier is reached.

4.6 HTTP Capacity

The aim of these tests is to stress the HTTP detection engine and determine how the solution copes with network loads of varying average packet size and varying connections per second. By creating genuine session-based traffic with varying session lengths, the solution is forced to track valid TCP sessions, thus ensuring a higher workload than for simple packet-based background traffic. This provides a test environment that is as close to real-world conditions as it is possible to achieve in a lab environment, while ensuring absolute accuracy and repeatability.

Each transaction consists of a single HTTP GET request ,and there are no transaction delays (i.e., the web server responds immediately to all requests). All packets contain valid payload (a mix of binary and ASCII objects) and address data, and this test provides an excellent representation of a live network (albeit one biased towards HTTP traffic) at various network loads.

4.6.1 44 KB HTTP Response Size – 2,500 Connections per Second

Maximum 2,500 new connections per second per Gigabit of traffic with a 44 KB HTTP response size—maximum 140,000 packets per second per Gigabit of traffic. With relatively low connection rates and large packet sizes, all hosts should be capable of performing well throughout this test.

4.6.2 21 KB HTTP Response Size – 5,000 Connections per Second

Maximum 5,000 new connections per second per Gigabit of traffic with a 21 KB HTTP response size—maximum 185,000 packets per second per Gigabit of traffic. With average connection rates and average packet sizes, this is a good approximation of a real-world production network, and all hosts should be capable of performing well throughout this test.

4.6.3 1.7 KB HTTP Response Size – 40,000 Connections per Second

Maximum 40,000 new connections per second per Gigabit of traffic with a 1.7 KB HTTP response size—maximum 445,000 packets per second per Gigabit of traffic. With small packet sizes and extremely high connection rates, this is an extreme test for any host.

4.7 Application Average Response Time: HTTP

Test traffic is passed across the infrastructure switches and through all inline port pairs of the solution simultaneously (the latency of the basic infrastructure is known and is constant throughout the tests). The results are recorded at each response size (44 KB, 21 KB, 10 KB, 4.5 KB, and 1.7 KB HTTP responses) at a load level of 90% of the maximum throughput with zero packet loss as previously determined in raw throughput testing.

4.8 HTTP Capacity with HTTP Persistent Connections

The aim of these tests is to determine how the solution copes with network loads of varying average packet size, varying connections per second while inspecting all traffic. By creating genuine session-based traffic with varying session lengths, the solution is forced to track valid TCP sessions, thus ensuring a higher workload than for simple packet-based background traffic. This provides a test environment that is as close to real-world conditions as it is possible to achieve in a lab environment, while ensuring absolute accuracy and repeatability.

This test will use HTTP persistent connections, with each TCP connection containing 10 HTTP GETs and associated responses. All packets contain valid payload (a mix of binary and ASCII objects) and address data, and this test provides an excellent representation of a live network at various network loads. The stated response size is the total of all HTTP responses within a single TCP session.

4.8.1 250 Connections per Second

This test will simulate HTTP persistent connections, each containing a total of 10 HTTP GET/responses of various sizes. The total HTTP response size for each persistent connection will be equal to four megabits, transmitted over a maximum of 25 connections per second for each 100 megabits of traffic.

4.8.2 500 Connections per Second

This test will simulate HTTP persistent connections, each containing a total of HTTP 10 GET/responses of various sizes. The total HTTP response size for each persistent connection will be equal to two megabits, transmitted over a maximum of 50 connections per second for each 100 megabits of traffic.

4.8.3 1000 Connections per Second

This test will simulate HTTP persistent connections, each containing a total of 10 HTTP GETs/responses of various sizes. The total HTTP response size for each persistent connection will be equal to one megabit, transmitted over a maximum of 100 connections per second, for each 100 megabits of traffic.

4.9 "Real-World" Single Application Flows

Where previous tests provide a pure HTTP environment with varying connection rates and average packet sizes, the goal of this test is to simulate real-world single application traffic. Each protocol will be run to capacity and failure, at which point the maximum supported throughput per protocol will be recorded.

4.9.1 Single Application SIP Flow

4.9.2 Single Application FIX Flow

4.9.3 Single Application SMTP Flow

4.9.4 Single Application FTP Flow

4.9.5 Single Application SMB Flow

4.9.6 Single Application RDP Flow

4.9.7 Single Application YouTube Flow

4.9.8 Single Application WebEx Flow

4.9.9 Single Application MSSQL Flow

5 Security Effectiveness (Optional)

5.1 Intrusion Prevention

These tests will accurately reflect the security capabilities of solutions that possess native deep packet inspection security capabilities.

5.2 False Positive Testing

The ability of the solution to identify and allow legitimate traffic while maintaining protection against threats and exploits is just as important as its ability to protect against malicious content. This test will include a varied sample of legitimate application traffic, which should be identified and allowed, or blocked, based on policy rules.

5.3 Evasions

Attackers can modify basic attacks to evade detection in a number of ways. If a solution fails to detect a single form of evasion, any exploit can pass through the solution, rendering it ineffective. NSS verifies that the solution is capable of detecting and blocking basic exploits when subjected to varying common evasion techniques. Wherever possible, the solution is expected to successfully decode the obfuscated traffic to provide an accurate alert relating to the original exploit, rather than alerting purely on anomalous traffic detected as a result of the evasion technique itself.

Evasions are considered highly sophisticated or advanced attacker capabilities; therefore, in order to perform well in this testing, a device must already possess strong detection and prevention technologies prior to evasion techniques being applied.

For more details, please refer to the NSS Labs Evasions Test Methodology v1.0.

6 Security Stability and Reliability (Optional)

Long-term stability is particularly important for an inline device, where failure can produce network outages. These tests verify the stability of the solution along with its ability to maintain security effectiveness while under normal load and while passing malicious traffic. Products that are not able to sustain legitimate traffic (or that crash) while under hostile attack will not pass.

The solution is required to remain operational and stable throughout these tests, and to block 100% of previously blocked traffic, raising an alert for each. If any policy-forbidden traffic passes, caused by either the volume of traffic or by the solution failing open for any reason, this will result in a fail.

6.1 Blocking Under Extended Attack

The solution is exposed to a constant stream of policy or protocol violations over an extended period of time. The solution is configured to block and alert, and thus this test provides an indication of the effectiveness of both the flow management and alert handling mechanisms.

A continuous stream of policy or protocol violations mixed with legitimate traffic is transmitted through the solution for eight hours at 10 Mbps, with no additional background traffic. This is not intended as a stress test in terms of traffic load (covered in the previous section); it is merely a reliability test in terms of consistency of performance with regard to policy handling.

The solution is expected to remain operational and stable throughout this test and to correctly handle 100% of recognizable policy or protocol requests, raising an alert for each. If any recognizable policy violations are passed, caused by either the volume of traffic or by the solution failing open for any reason, this will result in a fail.

6.2 Passing Legitimate Traffic under Extended Attack

This test is identical to the stability test run previously where the external interface of the solution is exposed to a constant stream of policy or protocol violations over an extended period of time.

The solution is expected to remain operational and stable throughout this test, and to pass most or all of the legitimate traffic. If an excessive amount of legitimate traffic is blocked throughout this test, caused by either the volume of traffic or by the solution failing for any reason, this will result in a fail.

6.3 Behavior of the State Engine Under Load

This test determines whether the solution is capable of preserving state across a large number of open connections over an extended time period.

At various points throughout the test (including after the maximum has been reached), it is confirmed that the solution is still capable of inspecting and blocking traffic that is in violation of the currently applied network control policy, whilst confirming that legitimate traffic is not blocked (perhaps as a result of exhaustion of the resources allocated to state tables). The solution must be able to apply policy decisions effectively based on inspected traffic at all load levels.

6.3.1 Passing Legitimate Traffic – Normal Load

This test ensures that the solution continues to pass legitimate traffic as the number of open sessions reaches 75% of the maximum determined previously in performance testing.

6.3.2 State Preservation – Maximum Exceeded

This test determines whether the solution maintains the state of pre-existing sessions as the number of open sessions exceeds the maximum determined previously in performance testing.

6.3.3 Drop Legitimate Traffic – Maximum Exceeded

This test ensures that the solution continues to drop all traffic as the number of open sessions exceeds the maximum determined previously in performance testing.

Note: If a solution allows traffic to "leak" due to the way it expires old connections, this will result in an automatic fail for the entire test.

6.4 Persistence of Data

The solution should retain all configuration data, policy data, and locally logged data once it has been restored to operation following power failure.

7 Total Cost of Ownership and Value

Implementation of infrastructure and security solutions can be complex, with several factors affecting the overall cost of deployment, maintenance, and upkeep. All of the following should be considered over the course of the useful life of the solution:

- **Product Purchase** – The cost of acquisition
- **Operational Benefits –** The "zero-touch" provisioning concept for SD-WAN cites considerably reduced deployment requirements, specifically regarding configuration and tuning, for example, time to add a new site is measured in hours rather than days or weeks. These reduced configuration requirements contribute to operational savings for the enterprise.
- **ROI Assessment –** There are savings associated with moving from high-cost, service-assured links (e.g., MPLS) to commercial broadband. There is value both in aggregating multiple low-cost links to support demand as well as in the ease of deployment and recurring service cost reductions that are associated with moving from expensive, service-assured links to less expensive options.
- **Product Maintenance** – The fees paid to the vendor, including software and hardware support, maintenance, and other updates
- **Installation** – The time required to take the solution out of the box, configure it, put it into the network, apply updates and patches, and set up desired logging and reporting
- **Upkeep** – The time required to apply periodic updates and patches from vendors, including hardware, software, and other updates

Appendix A: Change Log

Version 1.1 – 1 February, 2018

- Section 4 (Network Policy & Performance) moved prior to section 3 (optional security section) in order to emphasize performance nature of testing
- Added Section 3.1.3: Remote Configuration. Clarification on what will be assessed and measured.

Version 1.2 – 6 April, 2018

- Section 1.3: Revised topology to reflect logical configuration
- Section 4.1:
    - Added saturation test case
    - Clarified test cases
    - Added section 4.1.6: First Mile Network Behavior
    - Added section 4.1.7: Last Mile Network Behavior

Contact Information

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