TEST METHODOLOGY

Software-Defined Wide Area Network

January 3, 2023

V2.0

Table of Contents

Overview

1.1 Software-Defined Wide Area Network (SD-WAN)

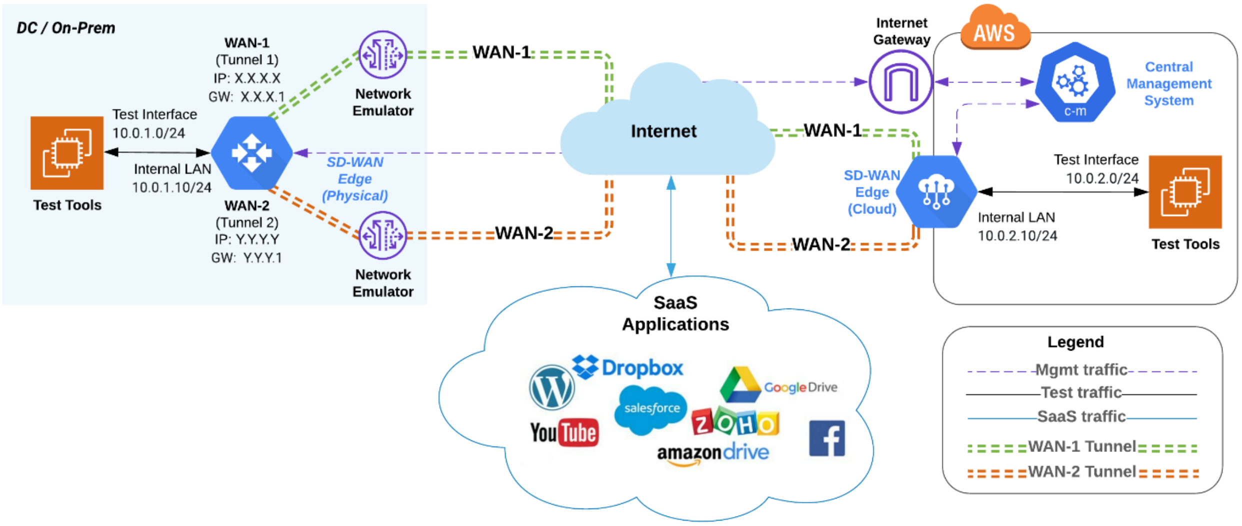
Modern networks are complex and difficult to manage. Software-defined networking (SDN) enables dynamic network performance monitoring and configuration by separating the network control plane from the forwarding plane. This abstraction enables the use of logical objects composed of multiple physical objects as opposed to having to manage individual physical objects with complex policies.

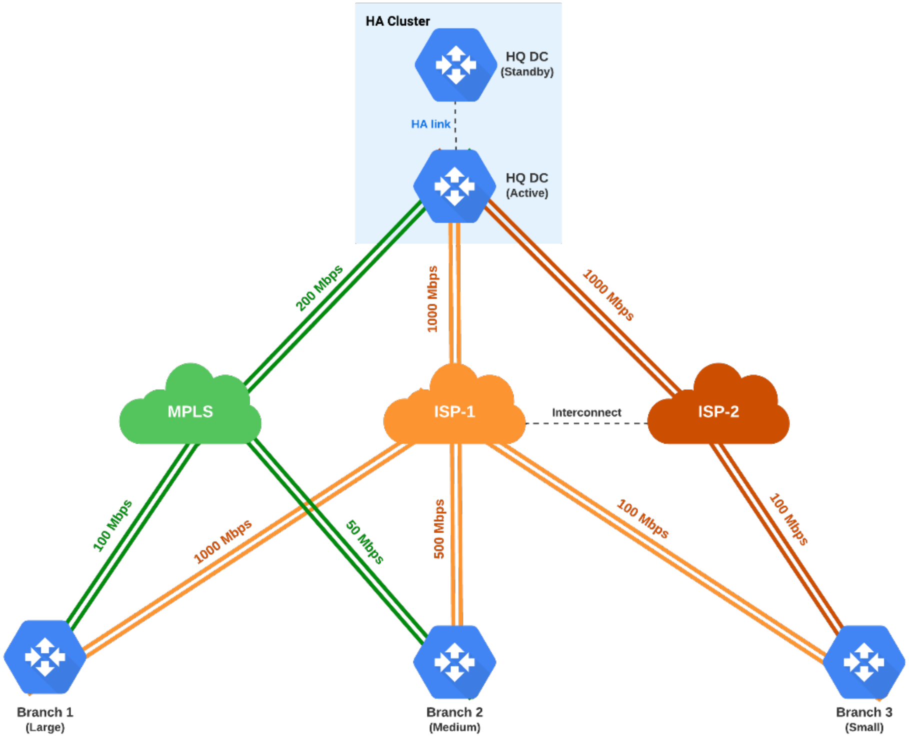
Wide area networks (WANs) have undergone several stages of development, from the early days of dial-up to dedicated circuits such as T1 / E1, to frame relay and ATM that then implemented multi-protocol label switching (MPLS) as an overlay technique to improve performance. Throughout their history the goal of WANs has been to connect networks across long distances.

The marriage of software-defined networking (SDN) with wide area network (WAN) technology enables efficient management of complex diverse deployments using common virtual private network (VPN) capabilities and the separation of data and control planes within SDN. With SD-WAN, 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 manage traffic 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). SD-WAN options are part router, part WAN optimization and traffic shaping, part access control, and part VPN. 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 Topology

The topology includes a redundant HA cluster at the HQ DC (On-Prem) location and an edge device (virtual SDWAN edge) located at a public cloud. The WAN environment is provisioned with behavioral characteristics like those typically encountered over WAN links. The test harness baseline is recorded to ensure consistent behavior, then the vendor solution is deployed, and each test case is measured against the baseline.





1.3 What Will Be Tested?

CyberRatings.org test reports are designed to address the challenges faced by enterprise security and IT professionals in selecting and managing SD-WAN products. Since SD-WAN technology manages WAN 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 technology and 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 applications
- Remote configuration capabilities
- Predictable performance experience for users

Location	Minimum tunnel throughput	Minimum Number of Devices
HQ in (DC/On-Prem)	4 Gbps	2 (1 x Active, 1 x Standby)
Branch (Cloud)	1 Gbps	1
Branch (Large)	2 Gbps	1
Branch (Medium)	1 Gbps	1
Branch (Small)	200 Mbps	1

1.4 Inclusion Criteria

CyberRatings invites all leading SD-WAN suppliers to include their products for testing at no cost. Products with significant market share as well as challengers with innovative technologies will be considered for inclusion. Decisions regarding inclusion of a vendor in a group test are based on an analysis of the market and an understanding of the criteria important to customers. Some of the elements considered are:

- Market presence
- Identification by industry analysts covering the specific technology area
- Consumer requests
- Innovative technology/offering or marketing claims that receive significant market attention (requires internal vetting for emerging vendors)

1.5 Product Guidance

CyberRatings tests products based on evaluation criteria that are relevant to enterprise networking and security professionals. These criteria are as follows:

- **Network policy effectiveness** – The purpose of an SD-WAN is to connect multiple networks over traditional commercial broadband and internet links and manage applications and performance between sites.
- **Resistance to impairments** – How effectively does the product handle impairments that would otherwise degrade network performance?
- **Stability and reliability** – Long-term stability is particularly important for an inline device, where failure can produce network outages and business disruption.
- **Performance** – Correctly sizing an SD-WAN is essential.
- **Management** – What level of effort is required to install, configure, operate, and maintain the product?
- **Cost** – What are the licensing option models for the product and the associated operational costs over time?
- **Value** – Customers desire low TCO, high effectiveness and predictable performance.

Routing & Access Control

This section verifies that the SD-WAN is capable of consistently and effectively enforcing network configuration policies. Testing of SD-WAN network policy effectiveness 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, and applications. Exception events that occur must also be accompanied by detailed log information to enable network forensics. Additionally, administrative visibility is critical. To facilitate analysis and troubleshooting, CyberRatings requires the logging of any function or event that results in dropped traffic.

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.

Every test scenario employs routing. The SD-WAN must be capable of handling common enterprise routing protocol configurations (e.g., BGP and OSPF). Each of the following test cases intends to measure and record an SD-WAN's ability to perform according to its parameters. In each case, the scenario is built to mimic real-world deployments and traffic experiences in addition to common configurations to reveal how well the configuration policy performs.

2.1 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.

This test determines whether the SD-WAN devices can dynamically establish and route traffic across multiple endpoints using VPN tunnels. Passing this test is a basic requirement for all SD-WAN devices.

2.2 Internet Breakout (Direct Internet Access)

One of the promising capabilities of SD-WAN is to enable branch offices to connect key services directly via the internet versus routing traffic back through the HQ DMZ. This offers both performance benefits to the user and reduces link loads back to the HQ.

In this test, the branches will be configured with Direct Internet Access (DIA) connection, in addition to the VPN tunnels from the branch. The solution is expected to provide Direct Cloud Access, i.e., direct connectivity to the SaaS applications with real-time optimization, such that the user experience (UX) isn’t affected/delayed, while still enforcing performance and application prioritization policies for VoIP and video. It is expected that the SD-WAN can choose the best SaaS hosting location for the applications and provide low-latency experience for the user. Traffic will be passed through the tunnel to up to 70% of its total supported capacity. MOS scores for voice and video should remain consistent with the baseline.

2.3 Simple Policy

A baseline policy is a routed configuration with an “allow all” policy treating all traffic as equal. This is to ensure that the SD-WAN can pass traffic between all sites over VPN tunnels without incident.

2.4 Complex Policies

Multiple outbound and inbound policies would be required to configure various scenarios such as allowing basic browsing and email access so internal clients can access untrusted and allow external networks access without giving external clients the ability to access internal network(s). The SD-WAN should provide the ability to control which applications and protocols are passed based on test case criteria, SLAs, scenario requirements, etc. A combination of policies will be tested to simulate the following list of enterprise application prioritization:

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

- VoIP
- H.323
- RTP
- RTCP
- RTSP

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

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

WAN Impairment

A critical function of any SD-WAN is the identification and routing of traffic based on policy prioritization (autonomous or configured), which is influenced by network performance characteristics (e.g., packet loss, variability, latency, jitter, etc.). Link impairment tests subject links to a representative set of real-world conditions encountered by enterprises today. Latency, jitter, and variability are all commonly encountered on public link technologies. Products are 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 SDWAN 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 to assess the adaptability of the SD-WAN (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 product 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 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 when the product is subjected to various impairments.

Different permutations of the following link impairments will be used to verify all SD-WAN performance features. Additionally, the impairment order and severity will be variable, representing the real-world experience of line quality, service degradation, and recovery.

3.1 Quality of Experience (QoE)

While throughput is important in SD-WAN, so is the user’s QoE. 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., congestion, packet loss, latency, packet delay variation, etc.). Link impairment tests subject connected links to testing that represents real-world conditions encountered by enterprises today. Congestion, packet loss, latency, and packet delay variation are all commonly encountered on public links.

3.1.1 Mean Opinion Score (Video & Voice)

Relative (video) MOS is an estimated perceptual quality score that considers the effects of codec; the impact of IP impairments (such as packet loss) on the group of pictures (GoP) structure and video content; and the effectiveness of loss concealment methods. The encoding specifications for video codec are used as guidelines and conformance, and vendors are free to design encoders to improve video quality and reduce the number of transmission bits. Simply put, MOS for video (relative MOS) can vary based on different advancements in the video estimation or encoding techniques.

3.1.2 Packet Loss (Video & Voice)

Packet loss for both Voice and Video plays a significant role in affecting MOS. Minimal or zero packet loss signals a reliable and effective solution for real-time communication.

3.1.3 Out-of-Order Packets (Video & Voice)

Misordered packets received due to improper application of error prevention techniques can impact MOS. Minimal or zero out-of-order packets signals proper sequencing and processing of received packets.

3.1.4 Duplicate Packets (Video & Voice)

Incorrect duplication techniques cause packets to be duplicated unnecessarily, which impacts bandwidth utilization rather than providing redundancy/reliability. Minimal or zero duplicate packets indicates effective processing of received packets.

3.1.5 RTP One-Way Delay (Voice)

The RTP one-way delay is used to assess the perceived quality of the link for Voice applications. High value indicates a possible lag in real-time applications.

3.1.6 RTP Jitter (Voice)

Variation in packet delay indicates the frequency of the received packets. High value indicates serious issues for realtime applications.

3.1.7 Connect Time (HTTP)

Time taken to establish an HTTP connection between client and server via the TCP handshake. Low value means faster connectivity.

3.1.8 Time to First Byte and Time to Last Byte (HTTP)

Time to first byte is the time it takes for a client to receive the first byte of a response to a request that it sends. Time to last byte is the time it takes for the client to receive all content in response. Low value means faster response.

3.1.9 Connection Latency (FTP)

Time taken for data to travel between source and destination. Low value means faster response.

3.1.10 Downloads Requested per Second & Downloads Successful per Second (FTP)

These values indicate the success rate of requests and responses for file downloads between the client and server. Close to equal values for both parameters indicate a very high success rate.

3.1.11 Mail Sent per Second (SMTP)

Indicates the total number of emails sent out per second. Higher value indicates better performance.

3.1.12 Sessions Requested per Second & Sessions Established per Second (SMTP)

These values indicate the success rate of requests and responses for session establishment between the client and server. Close to equal values for both parameters indicate a very high success rate.

3.2 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, an SD-WAN should support path decisions on a per-flow basis according to available links and according to the conditions that exist on those links.

3.2.1 Packet Loss

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

Any sustained packet loss should be identified by the SD-WAN and the links should be managed accordingly based on application or policy.

3.2.2 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 and Internet delay with minimum and maximum values.

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

3.3 Path Conditioning

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

3.3.1 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 Poisson and Gaussian distribution of selected packets.

Products should 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.

3.3.2 Packet Duplication

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

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

3.4 Link Saturation and Congestion

Global awareness of quality of service (QoS) 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.

3.4.1 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 stresses network buffering capacity. Sustained burst behaviors reveal that there is link congestion or other issues, and SD-WANs should alter paths based on known good alternate paths.

3.4.2 “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. The SD-WAN is expected to utilize bandwidth appropriately.

To replicate congestion in the “first mile,” impairments will be applied to the links from the HQ DC to the emulated aggregation point (ISP 1 and ISP 2).

3.4.3 “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. It is expected that the product will utilize bandwidth appropriately.

To replicate congestion in the “last mile,” impairments will be applied to the links from the emulated aggregation point (ISP 1 and ISP 2) to the branch SD WAN sites.

3.5 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 measures QoS using voice traffic and video stream. The test will include MOS scores for video and call measurements for VoIP.

3.5.1 All Impairments

This test applies to all the above impairments. The SD-WAN should manage traffic according to configured QoS classification settings.

3.6 Application-Aware Traffic Steering

This test will assess how the product 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.

3.6.1 Application Control Policies

These complex outbound and inbound policies consist of many rules, objects, and applications that verify whether the SD-WAN 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, a product's ability to perform the following functions will be tested:

3.6.1.1 Steer

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

3.6.1.2 Block Specific Action (Depends on Application)

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

3.6.1.3 Drop Low-Priority Application During Congestion Event

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

3.7 High Availability (HA)

In a redundant network, the SD-WAN deployed in HA mode is expected to provide uninterrupted network service during any system failures, while delivering an acceptable user experience.

3.7.1 Hardware Power Fail (Master-Slave Node Negotiation)

In this test, the HQ-Master device will be subjected to a device power loss event. It is expected that the HQ-Slave device will become active and ensure that communication persists with minimal or zero loss.

3.7.1.1 Persistence of Data

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

3.7.2 CMS offline or Orchestrator Outage

In this test, the CMS or orchestration service will be disconnected or unreachable. During such an event, enterprises would expect that this would not impact established links. Two random impairment tests will be run with MOS scoring to validate that the performance matches earlier tests.

3.7.3 WAN Link Failure

In this test, an established WAN link between sites is interrupted at the HQ DC 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 traffic should be redirected without loss or interruption to the applications using the available links based on prioritization schema. 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.

An SD-WAN must be able to operate resiliently despite link outages. As the total traffic sent through the links will be less than the available bandwidth, it is expected that the sessions or applications should continue without interruption, and there should be no noticeable user impact.

Performance

This section measures the performance of an SD-WAN 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 product 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 a product 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 will indicate a product's ability to withstand punishing performance scenarios. In addition to these impairment scenarios, standard traffic performance will be recorded, the results of which will be included in test reports and scorecards.

4.1 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 SD-WAN. 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 SD-WAN, as well as its effectiveness at forwarding packets quickly.

L2 Ethernet Frame Size	L1 Ethernet Packet Overhead			L1 Ethernet Packet Size	L2 Ethernet Frames / Second
	Preamble	Start Frame	Interpacket Gap		
Bytes	Bytes	Bytes	Bytes	Bytes	
64	7	1	12	84	1,488,095
128	7	1	12	148	844,595
256	7	1	12	276	452,899
512	7	1	12	532	234,962
1,024	7	1	12	1,044	119,732
1,280	7	1	12	1,300	96,154
1,518	7	1	12	1,538	81,274

4.1.1 64-Byte Packets

Maximum 1,488,095 frames per second per Gigabit of traffic. This test determines the ability of a device to process packets from the wire under the most challenging packet processing conditions.

4.1.2 128-Byte Packets

Maximum 844,595 frames per second per Gigabit of traffic

4.1.3 256-Byte Packets

Maximum 452,899 frames per second per Gigabit of traffic

4.1.4 512-Byte Packets

Maximum 234,962 frames per second per Gigabit of traffic. This test provides a reasonable indication of the ability of a device to process packets from the wire on an “average” network.

4.1.5 1024-Byte Packets

Maximum 119,732 frames per second per Gigabit of traffic. Some chipsets have difficulty with uncommon packet sizes. This test is designed to determine whether or not the SD-WAN handles uncommon packet sizes appropriately.

4.1.6 1280-Byte Packets

Maximum 96,154 frames per second per Gigabit of traffic.

4.1.7 1518-Byte Packets

Maximum 81,274 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.

4.2 Latency

In addition, the latency and user response time will be recorded to determine the effect the device has on traffic passing through it under various load conditions. Test traffic is passed across the infrastructure switches and through all inline port pairs of the device simultaneously. Packet loss and average latency are recorded for each packet size (64, 128, 256, 512, 1024, 1280, and 1518 bytes) at a load level of 90% of the maximum throughput with zero packet loss, as previously determined in section 1.17.

4.3 Maximum Capacity

The use of traffic generation appliances allows test 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 SD-WAN is causing an unacceptable increase in open connections.
- Excessive concurrent HTTP connections – Latency within the SD-WAN 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 SD-WAN is causing connections to time out.

4.3.1 Theoretical Maximum Concurrent TCP Connections

This test is designed to determine the maximum concurrent TCP connections of the SD-WAN 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 SD-WAN. 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.

4.3.2 Maximum TCP Connections per Second

This test is designed to determine the maximum TCP connection rate of the SD-WAN 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 SD-WAN and 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.

4.3.3 Maximum HTTP Connections per Second

This test is designed to determine the maximum TCP connection rate of the SD-WAN 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 SD-WAN introduces on the network. Load is increased until one or more of the breaking points defined earlier is reached.

4.3.4 Maximum HTTP Transactions per Second

This test is designed to determine the maximum HTTP transaction rate of the SD-WAN 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.

4.4 HTTP Capacity

The purpose of this test is to stress the detection engine to see how the device copes with HTTP network loads of varying average packet size and varying connections per second. By creating session-based traffic with varying session lengths, the device is forced to track valid TCP sessions, ensuring a higher workload than simple packetbased background traffic. The http test traffic characteristics are shown in the table below.

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.

Connections per Second (per Gigabit)	HTML Response Size (bytes)	Total Response Size (bytes)
1,000	115,570	129,738
2,000	57,388	64,824
4,000	28,048	32,136
8,000	13,512	15,920
16,000	6,353	7,916
32,000	2,667	3,903

4.4.1 1,000 Connections per Second

Maximum 1000 new connections per second per gigabit of traffic with corresponding HTML response size.

4.4.2 2,000 Connections per Second

Maximum 2,000 new connections per second per gigabit of traffic with corresponding HTML response size.

4.4.3 4,000 Connections per Second

Maximum 4,000 new connections per second per gigabit of traffic with corresponding HTML response size.

4.4.4 8,000 Connections per Second

Maximum 8,000 new connections per second per gigabit of traffic with corresponding HTML response size.

4.4.5 16,000 Connections per Second

Maximum 16,000 new connections per second per gigabit of traffic with corresponding HTML response size.

4.4.6 32,000 Connections per Second

Maximum 32,000 new connections per second per gigabit of traffic with corresponding HTML response size.

4.5 Application Average Response Time: HTTP

Test traffic is passed across the infrastructure switches and through all inline port pairs of the SD-WAN simultaneously (the latency of the basic infrastructure is known and is constant throughout the tests). The results are recorded at each HTTP response size at a load level of 95% of the maximum throughput with zero packet loss.

4.6 HTTPS Capacity

The purpose of these tests is to determine the performance curve and identify potential bottlenecks. We stress the HTTPS detection engine to see how the device copes with network loads of varying average packet size and varying connections per second. By creating session-based traffic with varying session lengths, the device is forced to track valid TCP sessions, ensuring a higher workload for simple packet-based background traffic.

Each transaction consists of a single HTTP(S) GET request, and there are no transaction delays (i.e., the web server responds immediately to all requests). All packets contain a valid payload (a mix of binary and ASCII objects) and address data.

Protocol	Cipher Suite Description	(Value)	Frequency Ranking	Security Classification
TLS 1.3	TLS_AES_256_GCM_SHA384	(0x13, 0x02)	1	Recommended
TLS 1.2	TLS_ECDHE_RSA_WITH_AES_256_GCM_SHA384	(0xC0, 0x30)	2	Secure

4.6.1 1,000 Connections per Second

Maximum 1000 new connections per second per gigabit of traffic with corresponding HTML response size.

4.6.2 2,000 Connections per Second

Maximum 2,000 new connections per second per gigabit of traffic with corresponding HTML response size.

4.6.3 4,000 Connections per Second

Maximum 4,000 new connections per second per gigabit of traffic with corresponding HTML response size.

4.6.4 8,000 Connections per Second

Maximum 8,000 new connections per second per gigabit of traffic with corresponding HTML response size.

4.6.5 16,000 Connections per Second

Maximum 16,000 new connections per second per gigabit of traffic with corresponding HTML response size.

4.6.6 32,000 Connections per Second

Maximum 32,000 new connections per second per gigabit of traffic with corresponding HTML response size.

TLS_AES_256_GCM_SHA384 (0x13, 0x02)		
Connections per Second (per Gigabit)	HTML Response Size (bytes)	Total Response Size (bytes)
1,000	113,430	127,666
2,000	54,917	62,455
4,000	25,700	29,710
8,000	11,170	13,483
16,000	3,870	5,358
32,000	150	1,227

TLS_ECDHE_RSA_WITH_AES_256_GCM_SHA384 (0xC0, 0x30)		
Connections per Second (per Gigabit)	HTML Response Size (bytes)	Total Response Size (bytes)
1,000	115,000	129,360
2,000	56,257	63,945
4,000	26,970	31,047
8,000	12,394	14,808
16,000	5,047	6,738
32,000	1,365	2,605

4.7 Application Average Response Time: HTTPS

Test traffic is passed across the virtual infrastructure and through the device simultaneously (the latency of the basic virtual infrastructure is known and is constant throughout the tests). The results are recorded at each response size at a load level of 90% of the maximum throughput with zero packet loss as previously determined in section 1.22.

4.8 File Download/Copy Time/Speed

Files from each of the following types are downloaded from the following locations to a local folder:

- MS Word files
- MS Excel files
- PDFs
- Zipped files/folders

This test is first performed without the SD-WAN offering to establish a baseline. The SD-WAN offering is then deployed, and the test is run again. Thus, results are relative to the baseline. The net increase in time to copy clean files of the various sizes is determined.

4.8.1 Microsoft OneDrive

4.8.1.1 Net increase in time to copy clean file – 100KB

4.8.1.2 Net increase in time to copy clean file – 1MB

4.8.1.3 Net increase in time to copy clean file – 10MB

4.8.1.4 Net increase in time to copy clean file – 100MB

4.8.1.5 Net increase in time to copy clean file – 1GB

4.8.2 Dropbox folder

4.8.2.1 Net increase in time to copy clean file – 100KB

4.8.2.2 Net increase in time to copy clean file – 1MB

4.8.2.3 Net increase in time to copy clean file – 10MB

4.8.2.4 Net increase in time to copy clean file – 100MB

4.8.2.5 Net increase in time to copy clean file – 1GB

4.8.3 Google Drive

4.8.3.1 Net increase in time to copy clean file – 100KB

4.8.3.2 Net increase in time to copy clean file – 1MB

4.8.3.3 Net increase in time to copy clean file – 10MB

4.8.3.4 Net increase in time to copy clean file – 100MB

4.8.3.5 Net increase in time to copy clean file – 1GB

4.8.4 HTTP web server

4.8.4.1 Net increase in time to copy clean file – 100KB

4.8.4.2 Net increase in time to copy clean file – 1MB

4.8.4.3 Net increase in time to copy clean file – 10MB

4.8.4.4 Net increase in time to copy clean file – 100MB

4.8.4.5 Net increase in time to copy clean file – 1GB

Stability and Reliability

Long-term stability is particularly important for an inline device, where failure can produce network outages. The product is required to remain operational and stable throughout these tests.

5.1 Behavior of the State Engine Under Load

This test determines whether the product 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 product is still capable of verifying and blocking traffic that is in violation of the currently applied access control policy, while confirming that legitimate traffic is not blocked (perhaps as a result of exhaustion of the resources allocated to state tables). The product must be able to apply policy decisions effectively based on inspected traffic at all load levels.

5.1.1 Passing Legitimate Traffic – Normal Load

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

5.1.2 State Preservation – Maximum Exceeded

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

5.1.3 Drop Legitimate Traffic – Maximum Exceeded

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

Management

6.1 Remote Initial Configuration

SD-WAN technology helps organizations achieve operational savings by enabling remote configuration of new locations rather than requiring engineers to be on-site. Many vendors offer ZTP, where onsite engineering expertise is not required other than the ability to connect a device to the appropriate internal and external links and power up the device. Once online, the device will call “home” (whether that is the HQ or a cloud configuration service) to gather and download the operational configuration information. This can also be achieved via a CMS (where applicable, this will be noted in test reports). The SD-WAN is expected to be remotely configurable to receive a pass for this remote configuration test case. For this test, the time to configure and deploy the site will be recorded.

Time to create configuration: Time taken to create a new configuration or clone from an existing configuration, apply updates where necessary, and add a new site to the existing network topology. The metric includes creation of a template configuration, updating IP addresses for management interfaces, creating VPN tunnels for WAN links, setting up thresholds and traffic policies to allow or deny traffic, and installing the required security packages wherever required.

Time to deploy configuration: Time taken to deploy the configuration and includes connecting to the CMS, selecting the appropriate configuration to be deployed, validating for errors/issues, and provisioning the device for a desired site.

Both metrics will be components of the operational cost model used to calculate an SD-WAN's TCO.

6.2 Centralized Management System (CMS)

The CMS must provide enterprises with the ability to centrally manage, configure, and monitor devices via a graphical user interface (GUI). CMS capabilities include but are not limited to monitoring, reporting, configuration changes, and the ability to update software on the SD-WAN. We will evaluate centralized management solutions across features that highlight how difficult it is to configure, maintain, and operate (i.e., find information).

Products are expected to offer robust and standardized logging and reporting formats along with several predefined and customizable dashboards and report generators, enabling administrators to create custom reports for outputs in a range of standard formats. Furthermore, support for role-based access control (RBAC) and comprehensive third-party authentication, two-factor authentication, and token/time-based authentication will be evaluated. Products are also expected to provide the administrator the ability to define and save multiple policies. Features such as Inheritance (nested rules), version control, and revision history should be fully supported.

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 the following should be considered over the course of the useful life of the SD-WAN:

- **Product Purchase** – The cost of acquisition
- Product Maintenance – The fees paid to the vendor, including software and hardware support, maintenance, and other updates
- **Installation** – The time required to take a solution out of the box, configure it, deploy 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
- **Operational Management** – Time commitment for day-to-day operations within a production environment, including support for monitoring logs, updating policies, and supporting incident investigations

Contact Information

CyberRatings.org

2303 Ranch Road 620 South

Suite 160, #501

Austin, TX 78734

info@cyberratings.org

www.cyberratings.org

Please read the disclaimer in this box because it contains important information that binds you. If you do not agree to these conditions, you should not read the rest of this report but should instead return the report immediately to us. "You" or "your" means the person who accesses this report and any entity on whose behalf he/she has obtained this report.

1. The information in this report is subject to change by us without notice, and we disclaim any obligation to update it.

2. The information in this report is believed by us to be accurate and reliable at the time of publication but is not guaranteed. All use of and reliance on this report are at your sole risk. We are not liable or responsible for any damages, losses, or expenses of any nature whatsoever arising from any error or omission in this report.

3. NO WARRANTIES, EXPRESS OR IMPLIED ARE GIVEN BY US. ALL IMPLIED WARRANTIES, INCLUDING IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, AND NON-INFRINGEMENT, ARE HEREBY DISCLAIMED AND EXCLUDED BY US. IN NO EVENT SHALL WE BE LIABLE FOR ANY DIRECT, CONSEQUENTIAL, INCIDENTAL, PUNITIVE, EXEMPLARY, OR INDIRECT DAMAGES, OR FOR ANY LOSS OF PROFIT, REVENUE, DATA, COMPUTER PROGRAMS, OR OTHER ASSETS, EVEN IF ADVISED OF THE POSSIBILITY THEREOF.

4. This report does not constitute an endorsement, recommendation, or guarantee of any of the products (hardware or software) tested or the hardware and/or software used in testing the products. The testing does not guarantee that there are no errors or defects in the products or that the products will meet your expectations, requirements, needs, or specifications, or that they will operate without interruption.

5. This report does not imply any endorsement, sponsorship, affiliation, or verification by or with any organizations mentioned in this report.

6. All trademarks, service marks, and trade names used in this report are the trademarks, service marks, and trade names of their respective owners.