LightReading
www.lightleading.com

REPORT
FEBRUARY 2016

IN PARTNERSHIP WITH: EANTC

Validating Cisco's Threat-Centric Security Solutions

Contents

INTRODUCTION

With 2016 set to be the first significant year for the commercial deployment of virtual network functions (VNFs), it's time to address one of the biggest challenges facing network operators in the vanguard of virtualization — network security in next-generation networks (virtualized and hybrid).

The challenge is not insignificant and not just focused on building a top-class defense against security threats: In addition to figuring out how to protect networks that incorporate virtualized functions and cloud environments, operators also need to determine how they can use their next-generation network security tools to develop new revenue-generating services.

With that in mind, Light Reading commissioned its independent test lab partner European Advanced Networking Test Center AG (EANTC) to evaluate a range of security tools and functions on offer from Cisco Systems Inc. (Nasdaq: CSCO) that traverses the virtual and physical worlds.

The results make for fascinating reading, as the EANTC team's report, which you can read over the course of the next ten pages, tells the story of a group of experienced technicians keen to examine just how emerging network topologies can be secured, defended and recovered before, during and after attacks of various kinds.

The evaluations are numerous and varied and simulate real-world scenarios, including: threat detection; attack visibility and mitigation; and security platform performance.

Of particular interest to service providers seeking to build a business case around next-generation security functionality is the test case dedicated to the verification of the provisioning process for security-as-a-service VNFs.

The EANTC team found Cisco's suite of capabilities more than capable of meeting the needs of today's progressive enterprises and service providers, whether in a virtualized environment or when a hardware-based solution is needed to deliver certain levels of performance and scale.

So let's get to the heart of the report, which is presented over the course of the following pages:

— The Light Reading team and Carsten Rossenhövel, managing director, European Advanced Networking Test Center AG (EANTC) (http://www.eantc.de/), an independent test lab in Berlin. EANTC offers vendor-neutral network test facilities for manufacturers, service providers, and enterprises.

THE EANTC PERSPECTIVE

Communications service providers and enterprises are increasingly conscious of network security issues for a number of reasons: Their complex network and IT infrastructures are becoming more sensitive to an increasing range of malicious threats; their business continuity depends more than ever on the ability of their networks to perform while under attack; and their networks are becoming increasingly distributed as cloud services play a greater role in day-to-day operations.

New threats are increasingly more sophisticated and are exploiting the growing attack options presented by new services, expanded network connections and device proliferation. To counter such threats, service providers need to quickly detect and mitigate threats as close to the source as possible across their networks.

This is one of the reasons why Light Reading commissioned EANTC to validate the functionality, performance and manageability of Cisco's virtualized security products line-up and ask the question: What is the state of the art in the functionality and performance of (telco) cloud-ready network security solutions?

Another key reason to undertake such an evaluation is the emergence of network functions virtualization (NFV), which opens up new opportunities for more fine-grained, precisely-placed, adaptable security functions. NFV permits the stitching together of network security components and enables the management of those components from a common platform based on SDN principles. In effect, network security could evolve from the traditional perimeter-style approach to a web of functions located close to assets exposed to potential threats, wherever needed across the cloud.

With this in mind, Light Reading asked EANTC to evaluate how prepared Cisco's virtualized security portfolio is for the new challenges (and opportunities).

But it's not all about virtualization: There is also the need to test the traditional network perimeter security functions that are still so important to enterprises and service providers alike. There is still the need for what insiders jokingly call a “BAF” (a big **** firewall). So does Cisco have a modern product to meet such needs, one that is ready to serve 100 Gbit/s and more?

Finally, complex IT solutions require superior orchestration, so that the operator understands what is going on at any time: Element management, network-wide management and fault and performance management aggregation must all work together with orchestration to provide a timely insight into any current threats and their mitigation options.

In summary, there are new risks as a result of more complex telco cloud technology and new types of threats — and there are new security infrastructure opportunities surprisingly enabled by that same complex technology. This provided EANTC with a great opportunity to dive into practical, independent performance testing and functional evaluation of what Cisco has ready for current commercial production.

TEST COVERAGE

Service provider security is a multi-dimensional challenge. This test validates the function and performance of virtual security services, also known as VNFs (virtual network functions), which, in this case, deliver security services.

The security controls validated in this test protect the trust boundaries at critical points in the service provider network, data center and cloud. Security is an in-depth process requiring the mitigation of threats as close to the source as possible so as to minimize collateral damage. In the course of our evaluation we looked at threats in the context of:

Before — things that can be done before the attack happens

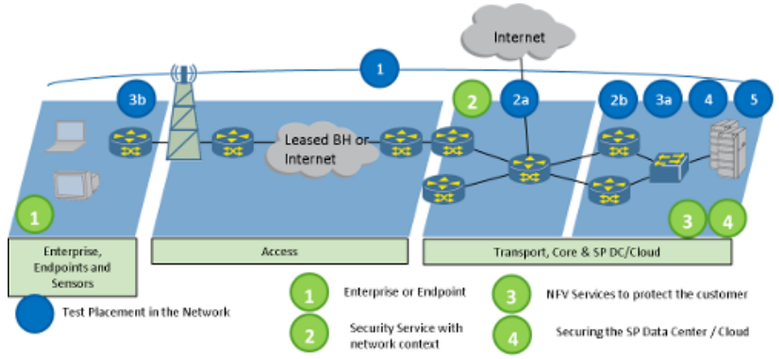
During — things that need to be done while the attack is happening, and

After — things that should take place post-attack, so the network operator is better prepared to deal with it the next time.Cisco calls this the "threat-centric security model."

CISCO'S 'THREAT-CENTRIC SECURITY MODEL'

There is no single "box" that secures everything. As the diagram below shows, virtual network functions are delivered differently by Cisco in different form factors dependent on the use case. Each use case leverages the network to deliver augmented security capabilities. Just having a security function is not sufficient: It must be placed into the right network context at the right time and at the right place in the network to minimize a threat as close to the source as possible and so minimize collateral damage.

PRODUCT & TEST AREAS



The tests run in this validation highlight use cases where the chain of security functions is purchased as a managed service focusing on service agility and use cases that apply the virtual security functions in a purpose-built appliance that delivers them with the performance and scale required to protect the service provider data center and cloud.

That appliance is the Cisco Firepower 9300. It takes the capabilities of a typical NFV system (orchestration, VM Lifecycle Management and other functions) and brings it all into an appliance showing the delivery of a catalog of security functions in a highly scalable, high-performance appliance.

CISCO NFV SECURITY SOLUTION SCOPE

In Cisco's view, there are five key areas that must be addressed for cloud security solutions, including:

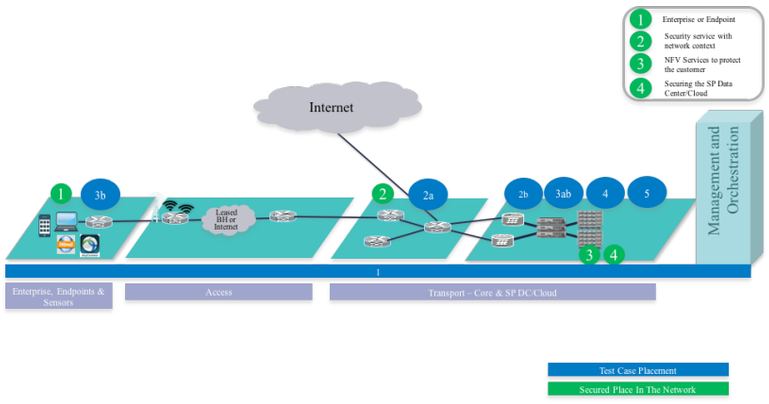
1. Security Effectiveness. One can't stop invisible threats: Does the solution quickly and accurate detect a threat? Does the operator have the ability to quickly detect and mitigate against sophisticated attacks that are designed to evade traditional defenses?
2. Service Chaining and Stitching. Different security functions must be linked (chained) together in proper order to provide proper protection, such as ASA Firewall/Next-Generation Firewall (NGFW), Distributed Denial of Service (DDoS), Next-Generation Intrusion Preven-ion (NGIPS), and Advanced Malware Protection (AMP). Ideally a solution is capable of supporting best-in-class virtual functions from third parties, since no single vendor has all of the technologies required for "defense-in-depth."
3. Orchestrating Security in SDN & NFV. The dynamic nature of cloud-delivered services means that security must "keep up" and be capable of being orchestrated and instantiated "on the fly." Manual processes must be minimized, if not eliminated.
4. Security as a Service in a Virtualized Multi-Tenant Environment. Security is a business enabler that can help carriers develop their cloud & NFV business transformation initiatives. They can extend security capabilities that they typically use to protect their own network infrastructure into revenue-generating offerings that also protect their customers from cyber attacks.
5. Carrier-class Performance, Scalability & Resilience. Any service provider solution must be "carrier-class" in terms of meeting the performance demands of such networks, including high throughput and line rate security processing, easy scaling as the network demands grow to address high bursts of network traffic and millions of subscribers and devices, while maintaining resiliency to minimize network or service disruptions.

TEST CASES

To secure cloud-based services, we sought to validate five key areas. We validated that, no matter where in the service provider network the security function runs, it is a combination of performance, network context and security effectiveness that enables the service provider to deliver secure business outcomes to protect the network infrastructure and/or to provide services to their customers.

Here is a landscape of the security tests and their purposes, followed by a more detailed tabular description:

TEST LANDSCAPE



(Note that each test case was run in a different set-up; Cisco did not set up all of the infrastructure above in a joint scenario during the EANTC test.)

TABLE 1: LIST OF TEST CASES

ID	TITLE	PURPOSE
1	Threat Detection Effectiveness	Verify the ability of the virtualized next-generation intrusion prevention systems (NGIPSv) solutions to detect and report concealed attacks.
2a	Service Chaining/Stitching – Firepower 9300 + FTD	Review the capabilities of the Firepower 9300 platform.
2b	Service Chaining/Stitching – Radware DDoS	DDoS attack visibility and mitigation of application and volumetric attacks using the FP9300 + vDP (Radware) – Time to see the attack, time to mitigate the attack.
3a	Orchestrating Security in SDN – ACI APIC with ASA.	Review the provisioning process of ASA appliances in a multi-tenant environment within ACI fabric.
3b	Orchestrating Security in SDN – CSR, ASAv and WSAv with Tail-F	Review the provisioning process of security-as-a-service VNFs in vMS environment.
4	Security-as-a-service in a virtualized multi-tenant environment	Verify the function of security services for the tenants in a virtualized data center environment. Verify the ability of ASAv to act as a VPN gateway and apply security policies to VPN users. Validate day-two move/add/change requests for virtual security services.
5a	Performance of the Firepower 9300 platform	Measure throughput, connection setup rate and connection capacity of the Firepower 9300 platform.
5b	ASA Firewall Clustering	Review ASA firewall clustering capabilities

THREAT DETECTION EFFECTIVENESS

SUMMARY

EANTC verified the ability of Cisco's virtualized next-generation intrusion prevention system (NGIPSv) solution to detect and report concealed attacks (exploit code in heavily obfuscated form).

TEST DESCRIPTION

An application-layer attack on a protected system can be concealed using manipulations on the lower layer protocols, which can severely impact the threat detection efficiency. So, for example, the attacker may utilize incorrect IP and TCP fragmentation when transmitting a malicious payload. Such payload can be successfully reassembled at the target host but can be difficult to detect at the firewalls located in the network.

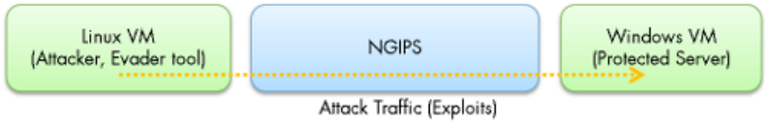
With this test case, we verified the ability of the NGIPSv to detect and block simulated attacks and also its ability to correctly reassemble and analyze obfuscated attack traffic. We simulated the attacks in combination with various obfuscation techniques using the McAfee Evader tool against a Windows PC located in a protected network. We compared how successful the attacks were when: a) the PC was unprotected; b) protected by NGIPSv and; c) when obfuscation was used.

We also evaluated the ability of the NGISPv platform to be managed uniformly across all functions and provide a comprehensive report on the detected threats from the information collected from different functions.

TEST SETUP

We set up a Windows host located in the network, protected by NGIPSv, and set up a Linux host located in the external network and equipped with McAfee Evader software.

SCHEMATIC TEST SETUP



With this test setup, we demonstrated the ability of the NGIPSv to detect and block malicious attacks with exploit payload. We also demonstrated that NGIPSv is successfully able to analyze traffic obfuscated by a number of manipulations on the IP, TCP and application protocols.

1. Normal attack – firewall not enabled
In the first step we deactivated a "no protection policy" on the firewall and sent legitimate traffic. With this step we wanted just to be sure the host is up and able to receive the incoming traffic.

We started the first exploit, using McAfee's Evader tool, verified that the attack was successful and the attacker gained control over the victim. The attacker was able to open a bind shell and create a file on the victim host.

2. Obfuscated attack – firewall not enabled
We then applied obfuscation to the exploit, while still keeping the firewall inactive. The goal of this step was to calibrate the attack to be sure that the client was able to understand the obfuscated traffic and properly reassemble it.

We set the following parameters:

- IPv4 Fragmentation (Size = 64 bytes, reverse fragment order)
- TCP Fragmentation (Size = 32 bytes, random fragment order)
- MSRPC options (big endian byte order, request segmentation 2048 bytes)
- SMB chaff (100% probability, WriteAndX with invalid payload and an invalid write flag, fill payload with random bytes)

We verified again that the attacker succeeded to gain control over the victim host.

We activated the option:

- SMB decoy trees (8 SMB trees, 8 write per tree, each write 2048 bytes, random bytes)

This time the attack caused the victim host to crash.

3. Regular attack – Firewall enabled
Next, we enabled the default out-of-the-box policy on the firewall, called 'Maximum Detection.'

We started with clean traffic and verified that it could reach the victim. After that we sent malicious traffic without obfuscation and verified that the firewall properly identified and blocked it.

4. Obfuscated attack – firewall enabled
Lastly, we activated the obfuscation techniques used in step 2 previously and started to send legitimate traffic.

We observed that the firewall blocked the legitimate traffic when TCP Segmentation was set to 64 bytes, marking that as "potentially bad traffic." Without this option set, the traffic was able to reach the client.

Now, we started the exploit of step 2 once more. To make the attack more difficult to identify, we set the following options:

- Obfuscate the shellcode encoder
- Bannerless bind shell

With TCP segmentation set to 64 bytes, the firewall blocked the attack and marked it as potentially bad traffic. After deactivating this option, the firewall properly decoded the attack and recognized the attack signature. After activating the decoys trees option, we verified that the firewall was also able to block only the malicious packets and allow the legitimate ones.

TABLE 2: HARDWARE & SOFTWARE VERSIONS

Role	Hardware	Software
NFVI	Cisco UCS C240M3	VMware ESX 6
Security Platform	VM	Cisco vFirepower v3D64 v5.4.0 - Using stock "Maximum Detection" policy
MGMT	VM	Cisco vDefenseCenter64 v5.4.1.1-33, SRU 2016 01 27 001 vrt
Attacker	VM	Ubuntu Linux 14.04LTS vm with McAfee Evader 2013-4_954
Victim Host	VM	Windows XP SP 2 vm

SERVICE CHAINING/STITCHING

Test case 2a: Firepower 9300 with FTD, NGIPS, AMP.

SUMMARY

We reviewed the Cisco Firepower 9300 platform architecture in combination with Firepower Threat Defense (TFD), Next-Generation Intrusion Prevention System (NGIPS) and Advanced Malware Protection (AMP), focusing on functionality, configuration and manageability aspects.

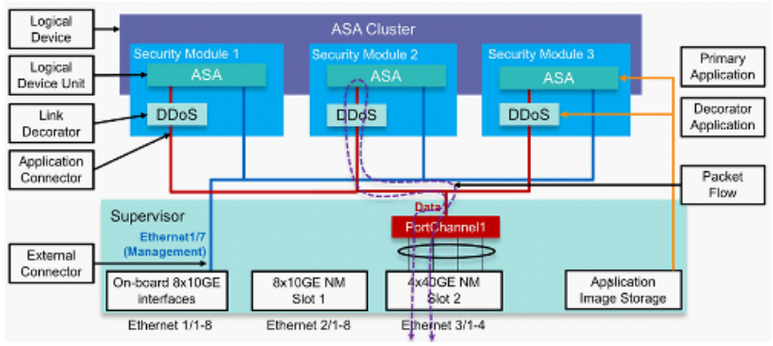
1. Overview of Firepower 9300

Cisco's security architecture allows the implementation of the same security applications (eg. ASA, NGIPS) on top of a range of physical or virtualized platforms. At some positions in the network, where a high performance is required, Firepower 9300 provides a suitable hardware platform, equipped with accelerator cards for encryption and potentially other functions. Meanwhile, at the edge, or at a customer's premises, where low costs are more important than high performance, common server/virtualization platforms can be used to host the same security functions at lower scale.

Firepower 9300 is based on the Cisco UCS chassis, and augmented with mezzanine acceleration cards for encryption, packet classifiers, and so on. Cisco claims its architecture provides up to 960Gbit/s internal fabric capacity, with 2x40Gbit/s backplane connection to each module. Currently Cisco offers SM-36 and SM-24 blades with respective 36 and 24 physical CPU cores, providing, respectively, an estimated 80 Gbit/s or 60 Gbit/s of firewall throughput performance. We tested Firepower 9300 performance aspects in test case 5 (documented later).

From the software perspective, the Firepower 9300 platform differs from the generic server or NFVi. The operating system basis is FXOS, with a central Supervisor and modules that run on each of the three blades. On top of FXOS, the orchestration can deploy a single software package, the “main application” such as ASAv and optionally a “decorator application,” such as DDoS protection. Third-party software not specifically designed for use with the Firepower platform can be instantiated within generic KVM supervisor running on top of FXOS.

SECURITY SERVICES ARCHITECTURE ON FIREPOWER 9300



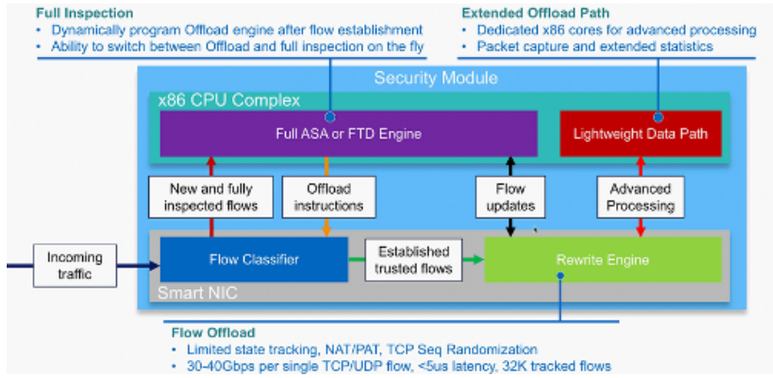
The software is staged through the management module. Cisco explained that all packages would be digitally signed (Cisco Secure Package — CSP), while the software integrity would be verified via secure boot and would be tamper-proof. EANTC did not test supply chain security aspects.

Cisco explained that the platform is designed for clustering the individual security modules (three within a single chassis, up to 16 total), to achieve a higher throughput capacity if necessary. The clustering is possible within a chassis, but also between multiple chassis.

By defining a cluster, one can automatically deploy the same software stack across all security modules. The clusters can be defined within a single chassis (intra-chassis), or between multiple chassis (inter-chassis), up to a five-chassis setup. Cisco mentioned that active-standby and active-active cluster configurations would be supported; EANTC did not evaluate resiliency aspects.

The interconnection between chassis can be provided by setting up Virtual PortChannel (vPC) from the redundant Cisco Nexus switches.

FLOW OFFLOADING CAPABILITY



The Firepower 9300 platform supports flow offloading. Cisco explained that the flows are processed by the flow classifier acceleration module, and the newly detected flows are first redirected to the software module for full inspection. Once the software achieves a sufficient classification of the flow, it can offload further processing by supplying offloading instructions to the flow classifier. Cisco explained that Firepower 9300 can redirect trusted flows to the lightweight data path, freeing the module's performance for other processing. Vendors use a range of mechanisms to optimize performance of virtualized network functions; this Cisco feature looks exciting and EANTC looks forward to testing it in a future project.

2. Demo of the Firepower management

We reviewed the Firesight Manager, the management and monitoring application for the Firepower and reviewer the process of security services provisioning.

In the initial configuration of the test bed, a single Firepower 9300 chassis was available, equipped with two security modules. In the first step, only one security module was provisioned with Cisco's Virtual Firepower Threat Defense (FTDv) function. In the main UI view, we were able to view the status of the modules and the network ports.

As the next step, we added a new logical device — a Cisco FTD to be assigned to the as yet unused second security module on the chassis and assigned physical network interfaces to it. The chassis management proceeded with the installation of the FTD software package on the second module.

3. Statistics reporting

Next, we reviewed example statistics provided by the Firepower 9300 platform and the functions were instantiated. The main dashboard provides an overview of the traffic volume by detected applications, network areas and behavior.

SUMMARY DASHBOARD

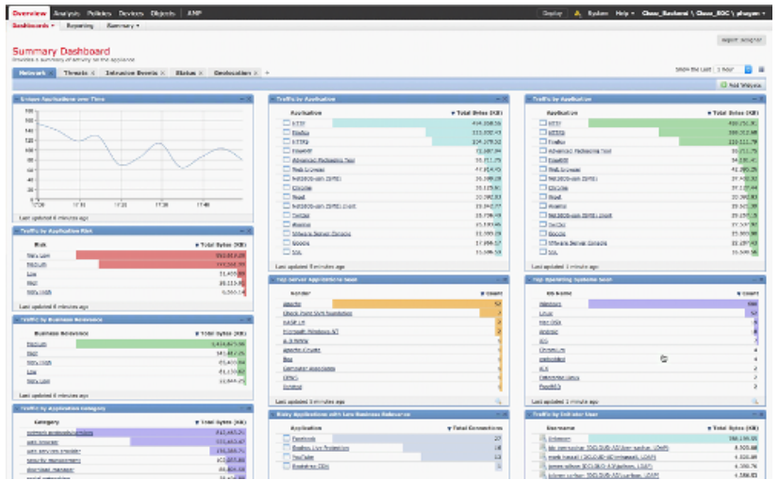


TABLE 11: HARDWARE & SOFTWARE VERSIONS

Role	Hardware	Software
Security Platform	Firepower 9300	Not disclosed

SERVICE CHAINING/STITCHING

Test case 2b: Radware DefensePro on Firepower 9300

SUMMARY

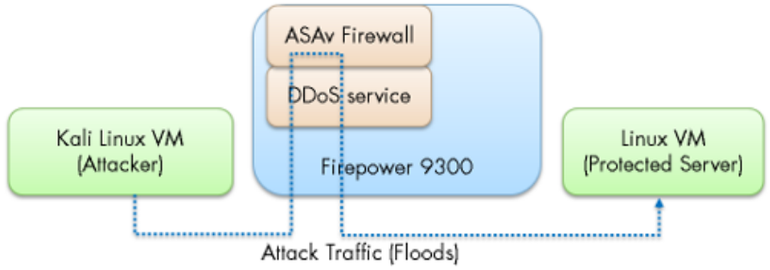
We reviewed the Radware DefensePro DDoS protection technology running on the Firepower 9300 platform, verifying the solution's ability to quickly detect and mitigate DDoS attacks on services and infrastructure through behavioral traffic analysis.

TEST DESCRIPTION

Radware DefensePro is a third-party solution for DDoS protection that is capable of running on the Cisco Firepower 9300 platform as a decorator application.

In this test case, we reviewed the functionality and management of DefensePro, as well as use cases for its utilization on the Firepower platform. Subsequently, we verified the DDoS protection function by simulating three widespread types of attacks — SYN Flood, NTP Amplification Attack (i.e. UDP Flood) and the DNS Flood. As the source of attacks, we used a Linux PC running Kali Linux, which is equipped with various tools for network security testing.

TEST SETUP FOR DDOS ATTACK SIMULATION



TEST RESULTS

The Radware Defense Pro is designed to mitigate DDoS attacks and provide an additional line of defense for the protected networks and services. It is available as a standalone appliance, a KVM image for generic virtualization, or as a Firepower 9300 application, which we evaluated in this test series.

The recommended location for the DDoS protection is in front of the firewall. This way, DefensePro is able to detect the attacks and analyze their behavior before they can be affected by other security infrastructure and at the same time protect the security infrastructure from the attacks.

When integrated to the Firepower 9300 platform, the DDoS protection function is placed in the similar way and inserted as a decorator application on top of the main application of the security module, e.g. the ASAv firewall.

DDoS protection can be utilized in various ways and locations. On one hand, it can be used as a part of the security service provided to the end customers in order to protect their cloud-based server infrastructure against attacks originating from the Internet. The security policies, DDoS mitigation techniques and other parameters can be defined and applied on a per-tenant basis.

Alternatively, it can provide protection for the security infrastructure itself, improving performance and reliability of other security products by deflecting some of the malicious traffic before it reaches them.

Finally, it can be used to detect and mitigate attacks originating from the local, protected networks, for example, from a protected office environment where a host may have become infected.

With the three attacks used in our test, we verified the functionality of the three main DDoS mitigation engines available in DefensePro.

The attacks based on traffic volume, but based on valid communication processes, can be detected and mitigated by statistical analysis of the traffic over time, a technology called Behavioral DoS by Radware. When such DDoS protection policy is applied, the BDoS engine will monitor traffic and learn the typical load profile of the service over a period of time. Strong deviations from the expected behavior can be then identified as ongoing attacks on the protected network.

The analysis of the traffic goes beyond layer 4/7 analysis. DefensePro collects a large number of parameters known about the flow, such as packet interval distribution, packet sizes, TCP Window size and so on. Abnormalities and difference from the values observed in the legitimate traffic may indicate an attack and trigger countermeasures.

The DNS Flood protection works on similar principles and can detect non-typical DNS query type distribution, a common characteristic for attack traffic. In case of a SYN Flood, DefensePro can apply a series of challenges to the incoming connections, in order to recognize legitimate and malicious clients. Unlike the BDoS engine, SYN challenge does not lead to packet drop or rate limiting. Instead, the engine answers the connection itself and runs a series of tests to distinguish the legitimate clients from malicious attack tools.

On the TCP layer, the protection against SYN Flood is achieved through SYN cookies, but even a legitimate connection can be additionally verified on ´the higher layers. For example, DefensePro can intercept the connection attempt to a web server and serve the untrusted client a HTTP redirect or a login/captcha page. Malicious tools designed for attacks and unable to process returned content will not be able to gain access to the protected web server.

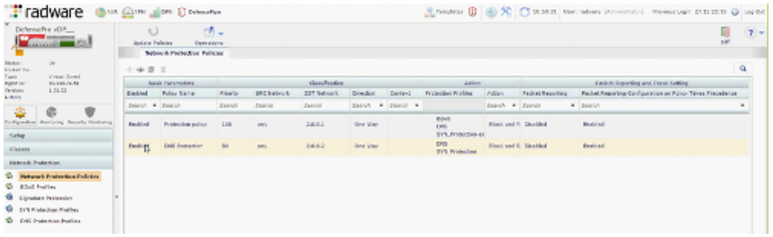
The SYN Flood protection can be triggered when a certain threshold of unanswered SYN packet rate is reached, with rules defined individually for each network address and port. The threshold is not based on the total packet rate, but rather on the difference in observed SYN and ACK packets. Therefore, occasional spikes in the legitimate traffic will not trigger the protection.

Once the protection was triggered for an engine, it can apply dynamically generated rules (based on signatures) to drop or limit traffic classified as malicious. Once the attack traffic returns below the threshold, the rule will be removed.

We reviewed the configuration process of DefensePro using Radware's configuration and monitoring tool APSolute Vision.

Configuration of the protection involves creation of a network protection policy that defines which of the DefensePro's detection engines will be applied, and with which parameters. Then, the policy and the required action are applied to a specific traffic direction. Each instance of a learning engine works independently and is able to learn the unique traffic profile specific to a network or customer.

APPLICATION OF THE NETWORK PROTECTION POLICIES



1. Attack mitigation – SYN Flood

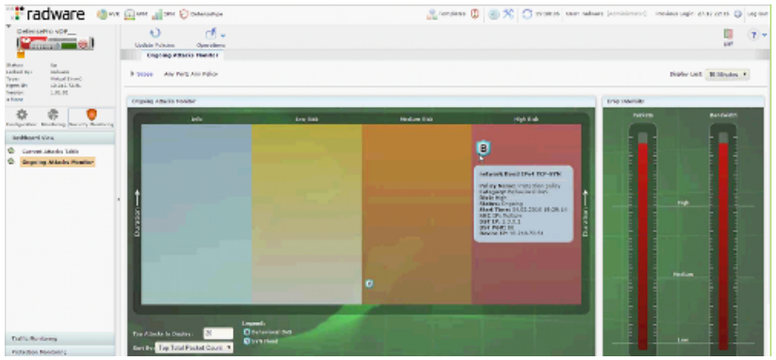
We verified the attack detection and mitigation functionality of DefensePro by transmitting three types of attacks from the Kali Linux VM. We initiated a SYN Flood attack by issuing the following command on the attacker machine:

hping3 –S –p 80 –rand-source –flood 2.0.0.1

In the APSolute Vision GUI, we observed that the SYN Flood protection engine recognized an ongoing SYN Flood and applied a new rule to challenge incoming connections.

At the same time, the BDoS engine detected the same ongoing attack, although it was less specifically recognized as network flood IP.

SYN Flood attack has triggered the BDoS and SYN Flood engines

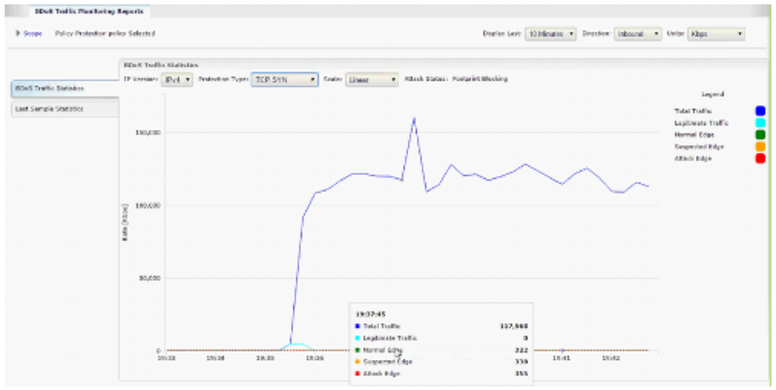


Within the attack details, we could also display the dynamically generated classification rule to identify the attack flow. In case of an ongoing attack, this rule can be examined in order to better understand the source of the attack and identify the traffic being dropped.

After stopping the attack on the attacker machine, we verified that the attack status has been cleared and the classification rule was removed.

We could observe the behavior of the traffic after the BDoS mitigation was triggered. We see only a brief increase in the rate of SYN packets, which is then quickly blocked.

MITIGATION OF THE SYN FLOOD ATTACK



2. Attack Mitigation

NTP Amplification attack (UDP Flood) In this test we simulated the flood of NTP response packets generated by a NTP Amplification attack. From the perspective of the attacked network, this attack can be seen as a UDP Flood. The attack traffic was generated by hping tool.

We observed that BDoS engine reported a new attack and generated a new dynamic rule to classify the attack traffic flow. Under BDoS Traffic Monitoring Reports, we observed that the attack traffic was quickly mitigated after detection.

While the attack was ongoing, we accessed the web server located in the protected network and verified that there was no observable impact on the legitimate traffic.

3. Attack Mitigation – DNS Flood

In this case, we simulated a flood of DNS packets directed against a DNS server located within the protected network. The attack traffic was generated by the dnsflood tool.

We observed that both BDoS and DNS engines were triggered by the attack. A new dynamic rule was generated to classify the attack flows.

MITIGATION OF THE SYN FLOOD ATTACK

Similarly to the other test steps, the attack traffic was blocked within seconds after detection.

TABLE 3: HARDWARE AND SOFTWARE VERSIONS

Role	Hardware	Software
Firewall	Firepower 9300	APSolute Vision v. 3.30.00

ORCHESTRATING SECURITY IN SDN

Test Case 3a: Application Centric Infrastructure (ACI) Application Policy Infrastructure Controller (APIC) with ASA firewalls.

SUMMARY

We reviewed the provisioning process of ASA appliances in a multi-tenant environment within ACI fabric, using both manual interface and scripting.

TEST DESCRIPTION

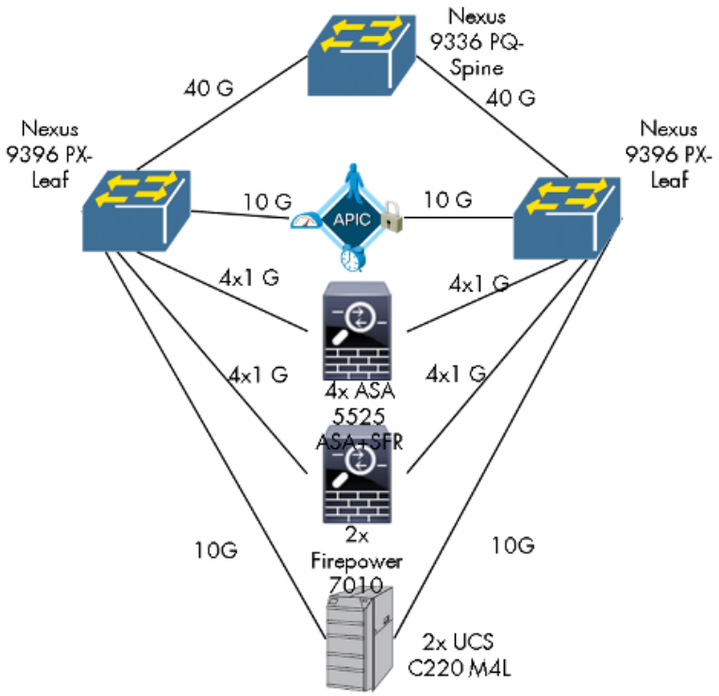
Cisco explained that Application Centric Infrastructure (ACI) is a new offering by Cisco, based on SDN and next-generation switching fabric and controlled by the Application Policy Infrastructure Controller (APIC). According to Cisco, the ACI architecture intends to provide a simple, flexible, scalable and resilient platform for data centers.

In this test case, Cisco aimed to demonstrate the orchestration process of physical and virtual security appliances on the ACI platform — ASA firewalls and NGIPS and malware detection platform Firepower.

As our test bed, we used a test setup available for security engineer training at the Cisco labs. The basis of this setup is an ACI fabric consisting of one Nexus 9336 PQ as the spine switch, two Nexus 9396 PX as leaf switches and 2x UCS 220 M4L compute nodes. APIC is a software component responsible for the provisioning and control of the data plane within ACI, including the provisioning of the security functions.

Attached to the ASA fabric, the test bed contains 4x ASA 5525 and 2x Firepower 7010 appliances. The 4 ASA devices were used to demonstrate their setup as 2 resilient clusters using different model of operation — a load-sharing cluster and an active-standby cluster as described in the following steps.

PHYSICAL TEST SETUP



TEST RESULTS

Cisco demonstrated the multi-context capability of the ASA firewalls within the ACI fabric, allowing a single ASA appliance to maintain many independent contexts for different clients, and for different locations in their service chain. The provisioning process of the security services was observed in an example scenario using the APIC (Application Policy Infrastructure Controller) to orchestrate security functions into the service chain.

The security concept of the ACI defines separate contexts for each tenant, and multiple security zones, so-called EPGs (End Point Groups) — network areas containing network elements with specific function and security status. The administrator of the ACI can flexibly define what set of configuration abilities can be granted to each tenant — this way, tenants may administrate the security policies within their context on their own, or delegate the administration to the ACI provider.

The communication between the EPGs is established through so called "Contracts" — a service chain connection that also has a security policy associated with it. APIC manipulates the service graph to redirect the traffic through the security solutions. In our case, we tested a physical ASA appliance; however virtualized ASAv solution is supported in exactly same way, as well as other security solutions from Cisco or other vendors.

In the scenario used for this test case, we had several such areas representing different functions in a typical web application accessible from the Internet. Step by step, we provisioned security services in the service chain, using different security policies and resiliency settings, as described in the steps below.

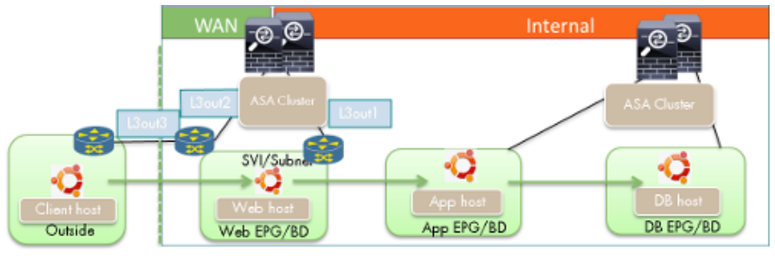
1. Review of the ACI Tenant Structure

Prior to our test, the test bed already contained a set of provisioned tenants ('pod1' through 'pod20') used for the training purposes, as well as several Linux VMs representing the network functions (web server, application server and database).

As the first step, we reviewed the existing structure of the service chain. The example service architecture for a single tenant contains four security areas ('EPGs') — Outside, Webserver, Application and Database, as presented in the diagram below. The goal of the provisioning steps was to insert and configure the security functions (ASA clusters) in the service chain between Outside and Webserver areas (load-sharing ASA cluster) and between Application and Database areas (active-standby ASA cluster).

At the beginning, the service already had provisioned contracts web-to-app (interconnecting Webserver and Application EPGs) and app-to-db (between Application and Database EPGs).

SERVICE GRAPH OF A SINGLE TENANT



2. Review of the ASA cluster

The ASA cluster to be added to our setup is an external ASA appliance (the same procedure can be applied to a virtualized ASAv solution). Inside APIC, this cluster can be registered as a 'L4-L7 Device' and later associated with a specific tenant content and contract.

APIC supports management and configuration of a variety of external devices through a set of plugin-like packages. There is support for Cisco ASA, ASAv, Firepower platforms, but also for third-party vendors such as Radware. Cisco explained that the device-specific package is a set of scripts making it possible for APIC to manage and configure them via management connection.

We verified that the device is in fact recognized by APIC and also recognized as a cluster setup and ready to be used in a tenant context.

3. Insertion of the ASA cluster and dynamic route peering

We inserted the load-sharing ASA cluster into the service chain — however, without a service policy applied to it. In this configuration, the ASA cluster only performs the routing between the two segments (Outside and Webserver) and applies a basic ACL-based filtering. This test step demonstrated the dynamic route peering feature of the ACI.

The ACI supports two types of adjacency for the devices (or VNFs) connected to the fabric: L2 for the direct connection to a VLAN; and L3 adjacency for the routed connection, where the fabric simulates a router instance between two network segments. In our case, the ASA cluster had a L3 adjacency to the outside network and to the webserver network. In addition, ACI acts as an OSPF neighbor to these instances, and is able to dynamically supply routes to them. We verified that the attached ASA cluster automatically obtained routes necessary to provide the connectivity between the external hosts and the webserver segment. From now on, the ASA cluster acted as a router between the EPGs. We verified the connectivity by sending ICMP pings in both directions.

4. Dynamic VLAN Allocation

Within the ACI fabric, the dataplane paths are dynamically established as the service graph setup requires. On the context-aware endpoint devices, APIC dynamically allocates VLAN interfaces for the data paths from the pool of available IDs. We verified this function by removing and reinserting the ASA cluster from the tenant context and monitoring the IDs assigned to VLAN interfaces within the ASA:

DYNAMIC VLAN ALLOCATION

```
Current IP Addresses:
Interface              Name                 IP address      Subnet mask     Method
Management0/0          management           10.10.10.82     255.255.0.0     CONFIG
GigabitEthernet0/0.809 internalIf           10.2.0.1        255.255.255.0   manual
GigabitEthernet0/1.702 externalIf           10.1.0.1        255.255.255.0   manual
```

```
Current IP Addresses:                                       IP address      Subnet mask     Method
Interface              Name                                 10.10.10.82     255.255.0.0     CONFIG
Management0/0          management                           10.2.0.1        255.255.255.0   manual
GigabitEthernet0/0.915 internalIf                           10.1.0.1        255.255.255.0   manual
GigabitEthernet0/1.819 externalIf
```

5. Modifying ASA ACLs via APIC GUI

We applied a simple security policy to the provisioned ASA cluster by modifying the ACL rules configured on it via the APIC GUI. In order to apply a different set of ACLs, we modified the function profile defined for the contract of the ASA in the APIC GUI, then applied changes to the ASA cluster. The change in the ACL rules was to deny, and later to permit, ICMP traffic again.

ASA device package for APIC translated the necessary configuration changes to the low-level configuration suitable for the ASA devices. Cisco explained that APIC does not completely rebuild the configuration, but is capable to applying exactly the changed fragment of it, thus the operation of the device would not be disrupted.

We observed that the changes we made were applied in less than one (1) second and verified the application of the new ACL rules by running ping between the Outside and Webserver areas.

6. Orchestration via Scripting

APIC provides a Python-based API that allows users to perform orchestration and configuration tasks otherwise possible with the APIC GUI from scripts.

We verified the functionality by running a series of scripts to delete, and then to completely recreate a tenant context and the associated service chain that included ASA cluster.

7. Active-Standby ASA Cluster Configuration

In addition to the load-sharing ASA cluster inserted between the Outside and Webserver EPGs, we also reviewed the second ASA cluster inserted between the Application and Database EPGs, and configured to operate in Active-Standby mode. Although EANTC did not perform actual resiliency testing, we watched the configuration of the ASA cluster using the management platform.

TABLE 4: HARDWARE & SOFTWARE VERSIONS

Role	Hardware	Software
ACI Spine switch	Nexus N9336PQ	v. 11.1(1r)
ACI Leaf switch	Nexus N9396Px	v. 11.1(1r)
Firewall 1	ASA5525	v. 9.5.1
	CPU: 1x Lynnfield 2393 MHz	ASA device package v. 1.2.3.4
	RAM: 8G	
Firewall 2	ASAv30	v. 9.5.1
	CPU: 1x Lynnfield 2393 MHz	ASA device package v. 1.2.3.4
	RAM: 4G	
IPS	Virtual NGIPS	v. 5.4.1
Firewall (L2 mode)	Firepower 7710	v.5.4.1
		Firepower device package v. 1.0.1.13
		Virtual NGIPS v.5.4.1

ORCHESTRATING SECURITY IN SDN

Test Case 3b: CSR, ASAv and WSAv with Tail-F

TEST PURPOSE

We reviewed the provisioning process of cloud-based NFV security services for service provider customers called 'Cloud VPN,' observing the integration of VPN functionality on the CSR1000v virtual router, firewall functionality on the ASAv, and URL filtering functionality on the WSAv.

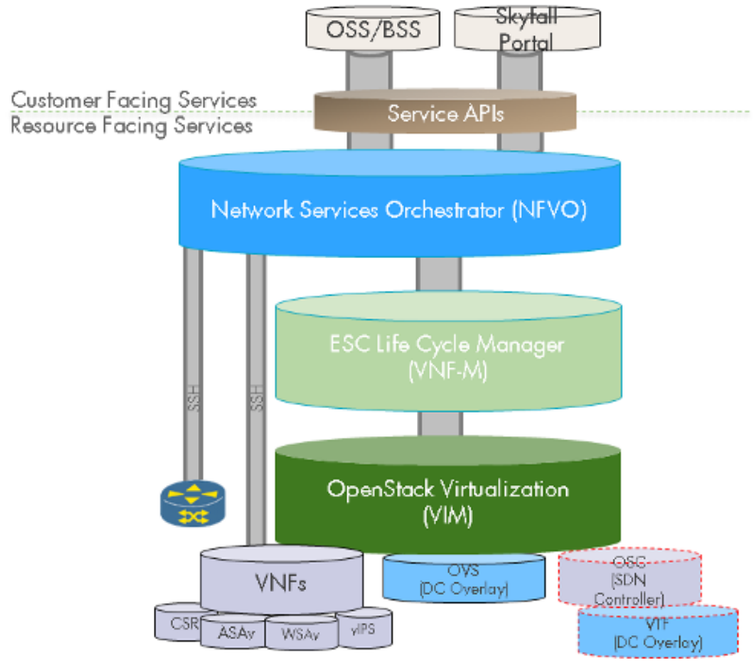
TEST DESCRIPTION

In this test, EANTC observed the ability of the Cisco Virtualized Managed Services (vMS) solution to provide rapid provisioning of cloud-based security services to customers. We also watched so-called "zero-touch" deployment of the customer CPEs (in their branch locations) and secure connectivity to the cloud security service. All the provisioning was done from the vMS self-service portal. The vMS infrastructure (management stack) and service chains (CSR1kv, ASAv, WSAv) used for this test are hosted on the Cisco Intercloud Services (CIS).

This test was conducted within Cisco Intercloud Services. The general architecture of vMS follows the ETSI NFV architecture and is presented schematically in the diagram below. vMS uses the Tail-F

Network Services Orchestrator (NSO) for Yang/NETCONF-based service orchestration, and the Elastic Services Controller (ESC) for VNF lifecycle management. The VNF service chain deployed consists of the CSR1000v (for IPsec VPN hub, routing), ASAv (Internet firewall) and WSAv (web content filtering), which are deployed in an Openstack-based cloud. The Skyfall Portal provides the customer interface for ordering and orchestration of services.

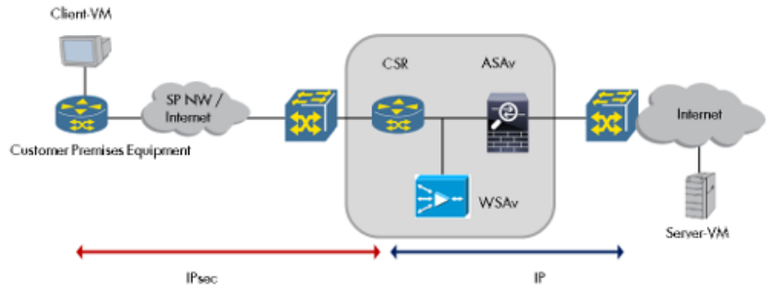
VMS ARCHITECTURE



The use case presented in this test case involves a service chain composed of security service VNFs, provisioned as a service to customers. The customers use this security service to protect their access to the Internet.

The diagram below depicts the specific network topology used in the test. The clients accessed the vMS platform via an IPsec tunnel established between the CPE and the CSR. The provisioned service chain included CSR, ASAv and WSAv. In addition, we provisioned Linux VMs at the client and server side as traffic endpoints.

SERVICE TOPOLOGY FOR A SINGLE CPE



TEST RESULTS

In this test case we observed the deployment process of the Cisco Virtual managed Services (vMS).

The goal of the vMS platform is to provide managed, integrated, NFV-based security solutions-as-a-service to customers. The presented solution consisted of the following VNFs from Cisco:

- CSR – as a general-purpose router and an IPsec gateway
- ASAv – firewall
- WSAv – web and email filtering

The demonstration ran in the Cisco Cloud environment, thus the administration capabilities would be as limited as for a customer of a vMS provider.

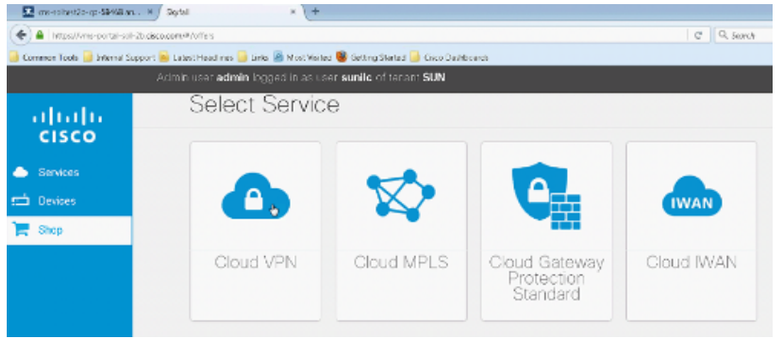
The general procedure of the vMS deployment involves the customer ordering a service from the provider via a customer portal. The orchestration procedure prepares all necessary configurations for the cloud, the VNFs and the CPE device. The provider only needs to apply the generated initial configuration to the CPE. The customer can immediately deploy the service by connecting the CPE to the available WAN connection and allowing it to complete the auto configuration process.

In our test, we simulated this process by placing an order for a security service on the provider's portal, applying the configuration to a CPE device and verifying the connectivity and security functions of the provisioned service.

1. Service Ordering by Customer

We initiated the provisioning process by accessing the vMS portal as a customer user and selecting the 'Cloud VPN' service from the shop. This service provides access to the vMS via IPsec tunnel. Other alternatives presented in the shop provide alternative access methods, such as MPLS, but these were not part of the test.

SELECTING THE SERVICE



The next pages provide more specific selection of the functions to be available in the service and the service parameters. We selected the 'Cloud VPN Advanced with Web Security,' which adds the VPN-based Internet and remote access (CSR), VPN-based firewall (ASAv) and the web security (WSAv) function.

On the next web page, we configured the service in more detail by specifying the number of CPE locations, desired bandwidth and capacity, CPE hardware, type of resiliency and options related to the security functions.

2. Monitoring the Service Orchestration Process

After reviewing the configuration of the service, we placed an order, triggering the orchestration process. In order to monitor the progress of the service provisioning, we accessed the NSO and ESC shell and logs (as admin user).

We monitored the provisioning debug output from NSO and identified events listed in the following table. (See yangesc.log for source information.)

TABLE 5: EVENTS OBSERVED IN NSO DEBUG OUTPUT

Timestamp	Event
21:07:10	Deployment initiated
21:07:18	Networks and subnets created
21:10:08	CSR VM deployed
21:11:02	ASAv VM deployed
21:12:46	ASAv VM operational
21:13:33	CSR VM operational
21:18:30	WSAv VM deployed
21:19:30	WSAv VM operational
21:19:30	Service chain operational

The total time from placing the order in the portal for the service to become operational was approximately 12.5 minutes. According to Cisco, the deployment of the WSA function requires a significantly longer time compared to other functions. The reason is that WSA performs an update of the website categorization and malware signatures database upon deployment.

We verified that all functions were instantiated by logging into their management console and reviewing the configuration on the VNFs. We verified that the CSR configuration does not yet include IPsec. This is expected behavior, as the CPE was not provisioned yet.

3. CPE Provisioning

As the next step, we obtained the serial number of the CPE device to be configured for our new service. This action is normally done by the customer, when receiving and installing a CPE from the provider, pre-provisioned with the Day0 configuration. In our case, the CPE was already installed in the test bed and connected to the WAN interface, but since the CPE was not yet provisioned in the NSO, it was not able to establish a tunnel and proceed with the configuration process.

Logged in as a customer, we submitted the serial number to the Skyfall portal to facilitate the further provisioning of the service.

4. CPE Configuration Process

We reviewed the configuration on the CPE. It included the configuration for the IPsec tunnel to the management VRF ("tunnel0") that is used to communicate with NSO and continue the provisioning process, and a tunnel to the data VRF ("tunnel1") that will be used for service data. We analyzed the logs available on the CPE and identified following events:

TABLE 6: EVENTS OBSERVED IN CPE DEBUG OUTPUT

Timestamp	Event
21:53:39	Submitting the CPE's serial number to the portal
21:57:04	Day1 configuration was applied to CPE
22:10:34	Tunnel1 established
22:10:41	Routes updated via BGP, data VRF operational

In total, the provisioning process on the CPE side took approximately 17 minutes from the submitting the CPE's serial number to the portal and until successful data path establishment.

We also reviewed the current configuration on the CSR and found that it was updated with IPsec configuration. This update is performed by the orchestration processes once the CPE serial number is submitted.

Finally, we verified that the CSR contained correct routes for the service data path and that an IPsec tunnel on the CSR was open and exchanged data.

5. Verifying connectivity

After the provisioning process for the service was complete, we verified the connectivity to the Internet by sending pings from a Linux VM attached to the LAN side of the CPE and directed to common online sites (such as www.google.com).

We verified that the ICMP requests from the client are indeed redirected through the IPsec tunnel, CSR and ASA modules, by comparing the packet statistics on these functions and by displaying connection table on ASA.

6. Verifying web filtering

As the next step, we verified that access to Internet websites worked. We confirmed that the access to a site poker.com (blocked by the category rules of the WSA function, as you might expect...) was indeed blocked and a notification page was returned instead.

TABLE 7: HARDWARE & SOFTWARE VERSIONS

Role	Hardware	Software
vMS	Not disclosed	Network Services Orchestrator (NSO) v. 4.0.3
		Elastic Services Controller (ESC) v. 2.2
		Openstack Red Hat Icehouse
		vMS Portal v. 2.2 build 1122
Router	2vCPU, 4G RAM, 8G disk	Cloud Services Router CSR1000v, v. IOS-XE 3.16.1a
Firewall	2vCPU, 4G RAM, 8G disk	Virtual Adaptive Security Appliance (ASAv), v. 9.5.1
Web Filter	2vCPU, 6G RAM, 8G disk	Virtual Web Security Appliance (WSAv), v. 9.0
CPE	Integrated Services Router ISR 891G	IOS 15.5(3)M

SECURITY AS A SERVICE IN A VIRTUALIZED MULTI-TENANT ENVIRONMENT

SUMMARY

We reviewed the web and email security functions available for the Cisco Cloud VPN service and verified their ability to work in a multi-tenant environment.

TEST DESCRIPTION

In this test, we looked at the ability of Cisco's security solutions to provide versatile security services — firewall, web filtering and VPN access to the tenants in a virtualized data center environment. In our scenario, multiple tenants share the same virtualized data center environment, where the access from the tenant network to the Internet can be secured by the virtualized Adaptive Security Appliance (ASAv), Web Security Virtual Appliance (WSAv), and the E-mail Security Appliance (ESAv).

We verified the capability of the CSR1000v to serve as a VPN gateway for the external users that is able to apply identical security policies as for users in the tenant's internal network.

The WSAv provided security for outgoing web access — Domain- and URL-based access control to the Internet — while ESAv is responsible for the email scanning for malicious content or spam.

We observed the multi-tenancy capabilities of these functions by defining multiple tenants with different rule sets. We also demonstrated the capability of applying identical security policies to both local and VPN users.

TEST RESULTS

In this test case we used a set of virtualized security functions from Cisco to verify their basic functionality in a multi-tenant security-as-a-service environment.

While the test case 3b covered the provisioning process of Cisco's Hosted Security Solution, in this test we verified the basic functionality.

We provisioned two tenants on the platform, each with one client in the customer's network (connected to the service via MPLS VPN), and one remote client connected via VPN access. For each client, we provisioned them with a web browser, to test the application of the web filtering policies, and a mail client.

On the network representing the Internet, we configured multiple web servers with a set of URLs that would be affected by the web filtering policies of these tenants.

1. Web Filtering

WSAv provides functions for fine-grained filtering for the Internet access.

Cisco explained that WSA has access to a large database of categorized websites and is able to restrict the access to the Internet, with the policy defined on basis of allowed or restricted categories. Additionally, a more fine-grained policy can be defined on a per-URL basis.

We configured WSA profiles for Tenant A and B using different rulesets. While some websites should be accessible for both tenants, URLs on one site should be accessible to Tenant A but not B, or vice versa.

We verified the correct function of the URL filtering defined for both Tenants by logging into the Windows VM simulating the clients and using web browser to access all URLs we defined on the web servers instantiated in the network representing the Internet.

We also verified that the web access was restricted identically, when accessing the service as VPN remote user.

2. E-Mail Filtering

As an example for the e-mail filtering function, we defined a simple policy to react on specific keywords in the incoming messages. As the action, we defined that detection should trigger a notification email.

We verified that when clients receive an email message containing a keyword, they also received a notification message, if the keyword matched the email policy settings defined for this tenant. The ESAv correctly evaluated the per-tenant policy even when a message with a specific keyword was sent to all clients in a single mail.

We also verified that the e-mail filtering works equally well for the VPN-based clients.

TABLE 8: HARDWARE & SOFTWARE VERSIONS

Role	Hardware	Software
NFVI	Cisco UCS B200-M4	VMware ESXi 5.5.0 Build 1618071
	CPU: 2x Intel Xeon E5-2695v3, 2.3GHz	
	RAM: 192GB	
Virtual Router	VM	CSR1000v 3.15.0.S (medium)
Firewall	VM	ASAv 5.9.4
Web Filtering	VM	WSA S100V 8-8-0-085
E-mail Filtering	VM	ESA C100V 9-1-0-032

PERFORMANCE, SCALABILITY AND RESILIENCE

Test Case 5a: Performance of the Firepower 9300 platform

SUMMARY

We tested the performance of the Firepower 9300 platform, yielding up to 155 Gbit/s throughput, 60 million concurrent connections and 2 million connection setups per second.

TEST DESCRIPTION

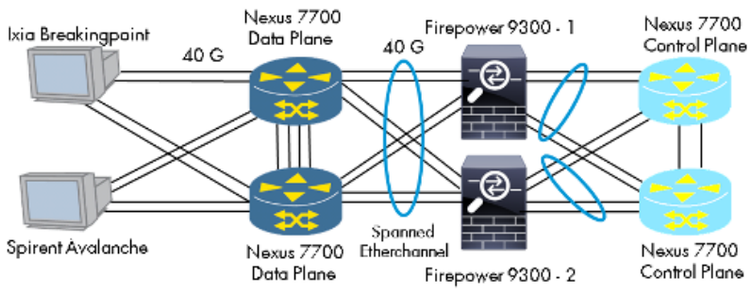
We measured the application-layer throughput of the Firepower 9300 platform by transmitting high rate application traffic. The test consisted of three separate groups:

- Throughput
- Connection setup rate
- Connection capacity

In the test, we used two Firepower 9300 chassis in an active-active (load sharing) cluster configuration. Each chassis was equipped with three SM36 blades. Two Nexus 7700 switches provided a virtual PortChannel connection for traffic to the cluster. A separate VLAN provided a peer-link connection between the two Firepower chassis for session synchronization and overflow traffic.

The traffic generators — Spirent Avalanche (throughput test only) and Ixia Breakingpoint — were attached to the Nexus switches with 40G links.

TEST SETUP FOR THE PERFORMANCE TESTS



We configured the two tests systems to jointly emulate 2,000 clients and 1,080 servers. All tests were run with IPv4 traffic. Usually, it is EANTC's policy to request mixed IPv4 and IPv6 scenarios; we have run tests of some of Cisco's virtualized solutions with IPv6 previously and look forward to expanding the virtualized firewall tests with additional traffic configurations, including IPv6, in the future.

TEST RESULTS – THROUGHPUT PERFORMANCE

First, we baselined the maximum forwarding performance of the hardware platform and the network infrastructure — excluding the added complexity of firewall feature configurations. To confirm Cisco's claims of the maximum performance, EANTC performed this test without any firewall rulesets, using a simplified, HTTP-only protocol mix.

We used large object sizes of 20,000 and 100,000 bytes per object, with ten transactions per HTTP connection. A total number of 8 million concurrent connections was targeted and connections were set up with a medium establishment rate of 151,800 connections per second.

The Cisco Firepower 9300 platform yielded a total throughput of 155 Gbit/s, equivalent to 17.6 million IPv4 packets/s. In fact, the throughput was limited by the amount of test equipment available, not by the device under test, which might have potentially supported even higher throughput.

Separately, we measured throughput with a more realistic traffic mix including Citrix, storage access, email, database access, secure shell and secure HTTP, voice-over-IP and other services as shown in the diagram below.

REALISTIC TRAFFIC MIX FOR THROUGHPUT TEST

Name	Weight	Seed	Sessions	% Bandwidth	% Flows	# Bytes
AppMix01_Citrix	6	13	1	6.00	0.22	683,111
AppMix01_HTTP_300K	28	13	1	28.00	2.07	335,808
AppMix01_HTTPS_300K	6	13	1	6.00	0.66	225,353
AppMix01_LDAP	2	13	1	2.00	26.98	1,838
AppMix01_NETBIOS	1	13	1	1.00	42.76	580
AppMix01_NFS_1M	3	13	1	3.00	0.05	1,434,364
AppMix01_RTP	3	13	1	3.00	0.14	511,406
AppMix01_SIP	2	13	1	2.00	12.38	4,005
AppMix01_SMBv2_1M	34	13	1	34.00	1.32	640,870
AppMix01_SMTP_300K	3	13	1	3.00	0.07	1,032,643
AppMix01_SNMPv1	2	13	2	2.00	7.62	6,512
AppMix01_SQL	7	13	1	7.00	5.22	33,244
AppMix01_SSH_1M	3	13	1	3.00	0.50	147,559

With otherwise identical parameters, the Firepower 9300 reached 63.6 Gbit/s throughput. In our experience, firewalls are usually sensitive to the types of protocols and services so it is important to use relevant and realistic traffic mixes in firewall testing.

TEST RESULTS – CONNECTION ESTABLISHMENT RATE

The second important firewall performance metric is the number of new stateful TCP/IP connections that can be established per time interval. This metric characterizes the control plane performance in terms of memory (connection table) management. It is crucial in environments with a lot of short-lived connections, such as in web services.

We chose minimally-sized objects for this test to increase the number of new connections as much as possible, avoiding having to wait for large transactions before closing down established sessions. There were ten transactions per HTTP session configured as before. Since connections were closed down very fast, there was a maximum number of only 9,617 concurrent sessions.

The Cisco Firepower 9300 managed to establish 2,000,000 connections per second — without session loss or other failures. That's a lot of connections albeit in an optimized configuration!

TEST RESULTS – CONNECTION CAPACITY

For scenarios with longer-lived connections, the total connection capacity is an important figure. If the firewall secures a lot of TCP-sessions — for example for remote desktop, e-mail access, storage services — the number of concurrent connections grows and the firewall needs to maintain all these connections in parallel.

In this test case, we used a moderate creation rate of 110,000 connections per second, and a moderate throughput of 0.55 Gbit/s.

Firepower 9300 reached a total of 60 million concurrent sessions in this scenario.

TABLE 9: HARDWARE & SOFTWARE VERSIONS

Role	Hardware	Software
Firewall	2x Firepower 9300 + 6x SM-36 modules	FX-OS 9.1.1.(3.84) + ASA 9.5.2.1

PERFORMANCE, SCALABILITY AND RESILIENCE

Test Case 5b: ASA Firewall Clustering

SUMMARY: We reviewed the ASA clustering architecture and followed a step-by-step process of creating a spanned cluster. We verified the failover and recovery capability of the cluster in case of a link and unit failure.

TEST DESCRIPTION

ASA Firewalls provide several methods of clustering for resiliency and load balancing, with a possibility to combine up to 16 units into a single cluster.

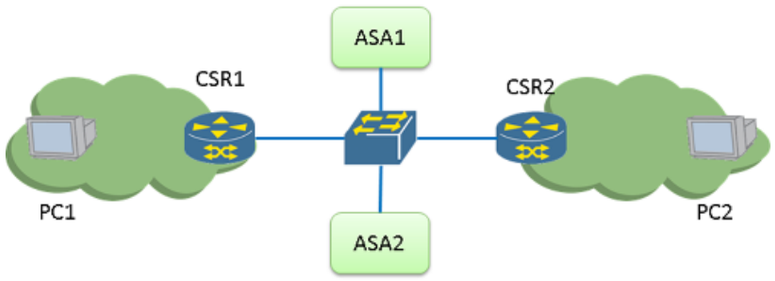
In this test case, we reviewed the different types of ASA cluster configurations and the process of setting up a cluster. Afterwards, we tested the failover functionality by simulating a link and a chassis failure and measuring the impact on the traffic.

The test bed consisted of two ASA appliances and two CSR routers, interconnected by a switch. Initially, the ASA units are not configured as a cluster and operate independently. For the connectivity verification, we use two Linux PCs attached at the routers.

The ASA units were configured to allow all traffic to pass, but require TCP and ICMP to be statefully inspected.

The routing was established by configuring OSPF peering between CSR and ASA units with both possible paths as equal cost multipath (ECMP). Due to the function principle of CEF routing (Cisco Express Forwarding), this configuration causes the traffic to take a different path in the returning direction.

PHYSICAL SETUP



TEST RESULTS

1. Initial state - cluster is not configured

We reviewed the initial status of the ASA units without cluster configuration via ASA CLI. We also reviewed the route table on both routers in the test bed and verified that the routing was configured to use different forward and return paths for the traffic. The traffic sent from PC1 to PC2 would be routed through ASA1, and the returning traffic from PC2 to PC1 via ASA2.

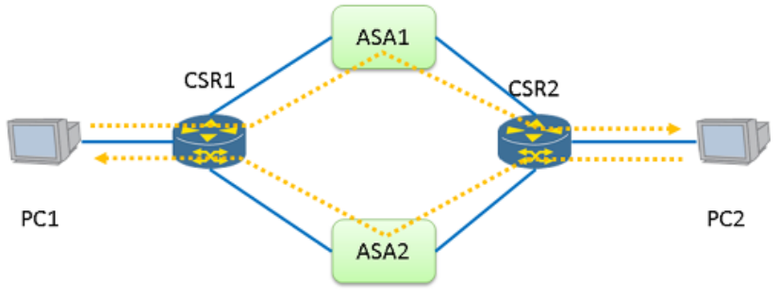
Due to stateful inspection on the ASA for TCP and ICMP traffic, only stateless UDP traffic is expected to pass the firewall. Both returning TCP packets and ICMP replies would flow through ASA2, which did not have the corresponding entry in the connection table and therefore will drop these packets.

We verified this behavior by sending different types of traffic from PC1 to PC2:

- A SSH connection from PC1 to PC2 failed to establish
- We sent ping from PC1 to PC2 and received no replies.
- We sent unidirectional UDP traffic from PC1 to PC2 using iperf and observed no loss.

This step demonstrates one of the main issues when attempting to build a scalable solution for traffic processing (such as a firewall) spanning multiple devices. In a setup where the traffic needs to be evenly distributed across multiple routing paths, the solution needs to ensure that the forward and return paths lead through the same device in order to allow stateful processing.

ASYMMETRIC ROUTING SETUP



2. Configuring a spanned cluster and the Master unit

On the ASA1, we erased the current configuration and switched the network interfaces to spanned mode, followed by loading a firewall-specific configuration from a saved file. After the configuration was loaded and applied, we queried the cluster information, which now indicated the ASA1 unit as being in a spanned cluster and in state Master.

We reviewed the cluster configuration on ASA1. It defines a new IP address on the spanned network interface, which becomes the cluster's IP shared by all members. Through the use of CLACP (Cluster Link Aggregation Control Protocol), the members of the cluster will present themselves to the attached switch as if they were individual links of an LAG connected to a single device. This way, a LACP-enabled switch will automatically perform load distribution across the cluster and the cluster itself appears as a single device to the network.

In addition, the cluster configuration defines the parameters for the health checks that will be performed within the cluster to detect failures.

3. Adding the secondary unit to the cluster

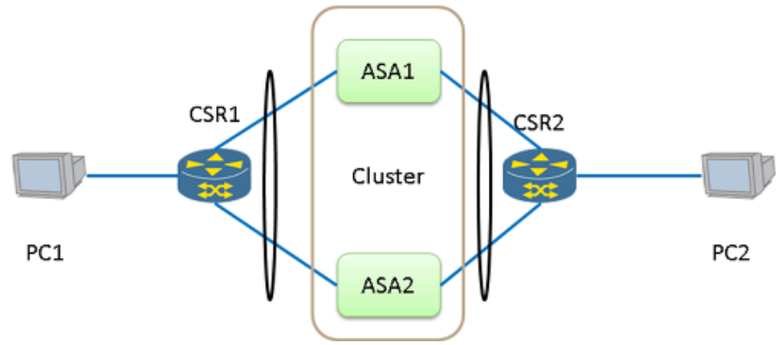
On the ASA2, we erased the current configuration and switched into clustered mode. Then, we added cluster configuration referencing ASA1 as Master. After application of the cluster configuration on ASA2, we observed in its CLI that the unit has contacted the Master unit via common cluster address and replicated the entire remaining configuration from it. The cluster configuration therefore is the only fragment that is needed to be configured on the slave units. The firewall configuration is automatically distributed and updated from the Master unit. The administrator may also use the master unit to execute commands on each unit in the cluster.

We observed a message on ASA2 CLI indicating that the unit is now active. We queried the cluster information, which now showed the cluster with two units, ASA1 being in Master role and ASA2 in Slave role.

4. Verifying connectivity in the cluster mode

We proceeded to review the routing table on the routers. We observed that there were no longer redundant paths with multiple gateways. Instead, the cluster appeared as a single gateway attached via Etherchannel interface.

ROUTING SETUP IN A SPANNED CLUSTER



At this point the cluster operated in an active-active state. All members of the cluster are active and will process traffic. This operation mode provides a performance advantage over the active-standby clusters without sacrificing the resiliency.

The distribution of traffic occurs via Etherchannel, and each flow can be assigned to any of the ASA units in the cluster randomly.

We verified the connectivity through the cluster by sending traffic between PC1 and PC2 the same way as it was done in Step 1. Each type of traffic — SSH, ping and iPerf — was forwarded correctly.

We observed that the active connections/flows are known on each unit, and are marked as active on the unit carrying that flow, while on the other unit it was marked as backup.

The connection state synchronization concept of the ASA is to synchronize the connection state only to one other member. In our case, since the cluster only has two units, each of them will have state of every flow (active or backup). In a cluster with more than two members, the backup flow state will appear only once.

With this concept, the state synchronization effort is limited, allowing for high scalability — synchronization of the state to all units would consume the same amount of memory on every unit as needed to hold the entire flow table, effectively limiting the scalability to one unit.

5. Failover on link failure

We initiated an iPerf test between PC1 and PC2. Using the connection table on the ASA units, we determined that this particular flow was handled on the ASA2 unit.

Using the management interface to the switch interconnecting the ASA and CSR units, we shut down the network interfaces leading to the ASA2 unit. In the iPerf output, we observed packet loss for approximately 10 seconds. Afterwards, the test traffic continued without loss. SSH connection between PC1 and PC2 was still open.

The CLI of the ASA1 unit indicated that ASA2 unit is no longer reachable and is suspended from the cluster. The cluster information showed that the cluster now only contains one unit, ASA1, as a master. In the flow table, we saw that all active flows have been migrated to ASA1 unit.

We restored the network connectivity to the ASA2 unit and restored its cluster state to 'enabled' in order to force it to rejoin the cluster.

We verified that the ASA2 unit appeared again as a part of the cluster. We also verified that all active connections were automatically synchronized to it. All active connections remained on ASA1 unit and the traffic was not affected.

6. Failover on software fault

As the alternative method to cause a failover, we initiated a software crash on the ASA1 (Master) unit through a debug command "crashinfo." The ASA1 unit produced a crash dump and proceeded to reboot. We observed packet loss in a running iPerf test for less than two seconds. The test traffic continued to run and the active SSH connection was still open. We confirmed that ASA2 unit now obtained the Master state in the cluster, and the ASA1 unit was excluded. All flows were migrated to ASA2. After the ASA1 unit was rebooted, it rejoined the cluster, but obtained the Slave state.

We then set the cluster mode on ASA2 to disabled, which forced it to be removed from the cluster. All flows were migrated to ASA1 and it obtained the Master state. We re-added ASA2 to the cluster as a Slave unit. We did not observe any traffic loss.

TABLE 10: HARDWARE & SOFTWARE VERSIONS

Role	Hardware	Software
Router	To be defined	Cloud Services Router (version to be defined)
Firewall	2x ASA5525-X	ASA 9.5.2