AMD Pensando Policy and Services Manager for Aruba CX 10000: User Guide

August 2023

Disclaimer

The information presented in this document is for informational purposes only and may contain technical inaccuracies, omissions, and typographical errors. The information contained herein is subject to change and may be rendered inaccurate for many reasons, including but not limited to product and roadmap changes, component and motherboard version changes, new model and/or product releases, product differences between differing manufacturers, software changes, BIOS flashes, firmware upgrades, or the like. Any computer system has risks of security vulnerabilities that cannot be completely prevented or mitigated. AMD assumes no obligation to update or otherwise correct or revise this information. However, AMD reserves the right to revise this information and to make changes from time to time to the content hereof without obligation of AMD to notify any person of such revisions or changes.

THIS INFORMATION IS PROVIDED 'AS IS." AMD MAKES NO REPRESENTATIONS OR WARRANTIES WITH RESPECT TO THE CONTENTS HEREOF AND ASSUMES NO RESPONSIBILITY FOR ANY INACCURACIES, ERRORS, OR OMISSIONS THAT MAY APPEAR IN THIS INFORMATION. AMD SPECIFICALLY DISCLAIMS ANY IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY, OR FITNESS FOR ANY PARTICULAR PURPOSE. IN NO EVENT WILL AMD BE LIABLE TO ANY PERSON FOR ANY RELIANCE, DIRECT, INDIRECT, SPECIAL, OR OTHER CONSEQUENTIAL DAMAGES ARISING FROM THE USE OF ANY INFORMATION CONTAINED HEREIN, EVEN IF AMD IS EXPRESSLY ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

AMD, the AMD Arrow logo, Pensando and combinations thereof are trademarks of Advanced Micro Devices, Inc. VMware ESXi™, VMware vSphere® vMotion® and VMware vCenter® are trademarks of VMware. Other product names used in this publication are for identification purposes only and may be trademarks of their respective companies.

amd.com/pensando

PPD22002

Revision History

Version	Description	Date
1.0	First release	February 9, 2022
1.1	Miscellaneous updates	February 25, 2022
1.3	1.49.1-T	June 13, 2022
1.5	1.49.2-T	August 11, 2022
1.6	1.49.3-T: minor corrections, caveat on disabling vSphere DRS	September 2022
1.7	1.54.1-T: new features (see Release Notes)	December 2022
1.7.1	1.54.2-T: minor errata fixes to guide; no new functionality	February 2023
1.9	1.62.1-T: new features, including policy distribution targets, network address translation (NAT), IP collections.	May 2023
1.11	1.62.2-T Release: IPsec VPN Tunnels	August 2023

Contents

Revision History 3

Introduction 11

  Key Features 11

  Related Documentation 12

Glossary 13

PSM Overview 15

Initial Deployment Workflow: High-Level Overview 18

PSM Object Model 18

  Firewall Objects 18

  Apps, Network Security Policy 19

  Key PSM Objects 20

  Labels 22

PSM Installation 23

    Storage Considerations 23

    Data Retention 24

  PSM Installation on ESXi 25

  PSM Installation on KVM 26

  Bootstrap the PSM Cluster 29

The PSM Graphical User Interface 32

  Online Help 32

  Searching 33

  Global Icons 34

  Server Certificate 36

  API Capture 36

Create PSM User Authentication Policy and Users 37

  User Authentication Policy 37

  Local User Lockout Policy 40

  Role-Based Access Control (RBAC) 40

    Roles 41

Role Binding 43
System Upgrade 44
AFC Upgrade 44
PSM Upgrade 44
Upload PSM Upgrade Bundle 44
Create Rollout 46
AOS-CX Upgrade 48
Configuration Snapshots 49
Associating a DSS with the PSM 49
AOS-CX CLI 50
ZTP 51
AFC 52
Verification 52
Decommissioning a DSS 53
Associating the PSM to Aruba Fabric Composer 57
Firewall Policy Functionality and Configuration 58
Considerations 58
VSX and Firewall High Availability 60
Supported Topologies 61
Connection Tracking 62
Firewall Policy Configuration 63
Create a VRF 64
Create a Network Security Policy 65
Rule Overlap Detection 68
Attach Policy to a Network 71
Create VRF on AOS-CX 72
Create VLAN on AOS-CX 72
Understanding Firewall Policy Scaling Profiles 73
Switch Policy Scaling Profile 73
Verifying the Number of Rules Consumed in the Data Plane 77

Understanding Hierarchical Security Policy 84
Policy Enforcement 86
Configuration and Verification 89
Enable/Disable Individual Firewall Rules 90
Configuring Firewall Log Export 93
Firewall Log Export Policy Configuration 93
Bind Export Policy to DSS 94
Firewall Log Record Format 97
Firewall Syslog Message Examples 99
Deduplication for Firewall Logs 101
Flow Export (IPFIX) 104
Step 1a: Enable Flow Export feature globally on the switch: 104
Step 1b: Set the source IP address of the exported IPFIX packet: 104
Step 2a: Configure the Flow Export policies under the “Tenants” -> “Flow Export” menu in PSM. 105
Step 2b: Applying a flow export policy at the DSS level 105
Step 2c: Apply the defined IPFIX flow-export-policy at the VRF level 107
Guidelines 107
Considerations 108
IP Flow Information Export (IE) Entities (1/2) 109
IP Flow Information Export (IE) Entities (2/2) 110
Configuring Apps 112
Protocol And Ports 113
ALG 114
DDoS Detection and Alerting: Maximum Sessions and CPS Limits 116
Maximum Session Limit 116
Maximum CPS 117
Min and Max Values 118
Configuring the Maximum Sessions / CPS on a VRF via the PSM UI 118
Configuring the Maximum Sessions/CPS on a Network via the PSM UI 119
API Examples 120

Behavior on Reaching the Maximum Session Limit 122
Behavior on Reaching the Maximum CPS Limit 123
Implication of Configuring `session-limit` on an Active System 125
VSX Implications 125
Multiple ALG Types/Apps/Protocols in Firewall Policy Rules 127
Steps To Configure Rules With Multiple Proto-Ports Via the PSM UI 127
Steps to configure rules with multiple ALG types via the PSM UI: 128
API Examples: 128
API Example: Policy with Rule Referencing Multiple ALG Types: 130
IP Protocols Support for Firewall Policy 130
UI Examples 131
API Examples 132
Policy Distribution Targets 134
Adding a Switch to a PDT 135
IP Collections 136
Defining an IP Collection 136
Using an IP Collection 136
Caveats 138
Network Address Translation (NAT) 139
NAT in Data Center Design 139
NAT Policy Direction 139
Ingress NAT Policy 140
Egress NAT policy 140
Supported NAT Operations 140
Supported Static One-to-One NAT Types 140
Source NAT 141
Destination NAT 142
Twice NAT 144
Configure NAT Policy on the PSM 145
Switch-Side Configuration: 145

Step 1: Preparing the DSS to be in border leaf mode 145

Step 2: Configuring interface persona 146

Step 3: Advertise the post NATed addresses in IGP 146

PSM Side Configuration: 147

Step 1: Defining a policy distribution target (PDT) containing the border-leaf DSS devices 147

On the PSM UI: 147

Step 2: Define and apply source NAT policies 148

Step 2a: Define a required NAT policy 148

Step 2b: Apply the NAT policy on the VRF 149

Considerations 149

Example: SNAT Flow on DSM 150

Example: NAT Rule Statistics 150

Step 3: Defining and applying destination NAT 151

Step 3a: Defining a destination NAT policy 151

Step 3b: Apply the NAT policy on the VRF 152

Step 3c: Defining a destination NAT policy with both DIP and DPORT translation 153

Step 3d: Apply the defined DNAT policy on the VRF for enforcement 153

Step 3e: Defining a twice NAT policy with SIP, DIP and DPORT translation 154

Caveats 155

Stateful Firewall Flow Migration with vMotion 156

Considerations for Multi-Homed ESX Servers 157

Behavior with Flow Logs and Flow Statistics 157

Configuration 158

DSS Required AOS-CX CLI Configuration 159

Caveats 160

Configuring IPsec VPN Tunnels 161

Overview 161

IPsec Functionality Overview 161

IPsec Topology Designs 162

IPsec Active/Active with VSX: 162

IPsec Active/Standby with VSX 163

IPsec Supported Modes and Crypto Options 164

Order of Operation with Service Chaining of FW, NAT and IPsec services 166

IPsec Configuration Workflow 167

Prerequisite Configuration On the Switch 167

Configure IPsec VPN Tunnel on standalone switch (no-HA mode) 169

VSX Requirement For IPsec 174

Configure IPsec VPN With Active Standby Failover 174

Configure IPsec VPN With Active/Active Failover 180

QoS Support over IPsec VPN Tunnels 183

QoS classification 183

Inner DSCP to outer DSCP copy 184

QoS Policing 184

QoS Shaping 185

Monitor IPsec VPN Tunnels with the PSM 186

Troubleshoot IPsec VPN Tunnels With the PSM 191

Duplicate SAs Scenarios 193

Caveats 193

Monitoring the DSM via the PSM UI 196

Metrics Charts 199

PSM Automation 204

Python Language Bindings 204

Ansible Modules 204

REST API 206

Tech Support Collection 207

Appendix A: PSM Quorum High Availability 209

Appendix B: PSM Operational Network Ports 210

Appendix C: Configuring Microsegmentation in Non-AFC Environments 213

Topology 213

Configuration on the DSS 214

Global Config Mapping the Primary and Secondary VLAN 214

Host-Facing Interface Configured as Regular Trunk, Allowing Both Primary and Secondary VLAN 214

SVI Config on Primary with Local Proxy ARP 215

Configuring VMware (ESXi) 215

Configuration on the PSM 215

Appendix D: Saving the PSM Recovery Key 216

Saving the Key 217

Recovering the Cluster 217

Appendix E: Using the PSM Network Graph to Create Security Policies 219

First Method 219

Second Method 229

Introduction

This guide describes how to install and operate the *AMD Pensando Policy and Services Manager* (*PSM*) to manage the stateful services of the *Aruba CX 10000 with AMD Pensando* distributed services switch (abbreviated as either *CX 10000* or *DSS*).

The PSM can be accessed via the IP address or host name of any of the PSM cluster nodes or, if a load balancer is being used, the IP address or host name presented by the load balancer. In this document, the PSM address will be referenced as either `$PSMaddr` when used in the context of shell commands or scripts, or as *`PSMaddr`* in other examples.

The PSM is managed through either its browser-based GUI or its secure RESTful API. Most examples in this document show the GUI, which is accessible at the URI `https://`*`PSMaddr`* .

Key Features

Core functionality supported includes:

- Distributed stateful firewall
- Microsegmentation (using PVLAN)
- DDoS detection and alerting
- Firewall logging and metrics
- Network address translation (NAT)
- IPsec VPN Tunnels
- IPFIX
- Full AOS-CX routing and switching feature set
  (see AOS-CX documentation for further details)[1]
- Fabric and services orchestration with AFC and PSM

[1] https://www.arubanetworks.com/techdocs/AOS-CX/help_portal/Content/ArubaTopics/Switches/10000.htm

Related Documentation

- *Aruba CX 10000 Switch Series Installation and Getting Started Guide*
- *PSM Release Notes*
- Release notes for AOS-CX and Aruba Fabric Composer (AFC)
- *Aruba Fabric Composer User Guide*
- AOS-CX Feature Guides
- Aruba transceiver data sheet

Aruba documentation can be found at the Aruba 10000 Switch Series documentation portal.

See the Aruba Support Portal for details on feature support.

Review the PSM release notes for details and information about new features, known issues, fixed bugs, and supported servers, cables, and switches.

Glossary

Name	Description
AFC	Aruba Fabric Composer
AOS-CX	The Aruba switch operating system, providing network services functions and management
CoPP	Control Plane Policing
Data traffic	(aka *data plane traffic*) the actual network data being processed by the DSS environment
DSE	Distributed Services Entity: collectively describes the services and monitorability provided by the two DSMs in a DSS
DSM	AMD Pensando Distributed Services Module (two per DSS): the stateful services execution engine of the DSS
DSS	Aruba Distributed Services Switch with AMD Pensando
Egress	Traffic leaving a host to fabric, in reference to security policies
Ingress	Traffic entering a host from fabric, in reference to security policies
ISL	Inter-Switch Link, a layer 2 interface between two VSX peer switches
Management and control traffic	(processed by the management and *control plane*) network communication related to the interoperability, reporting, and policy management of the DSS environment

Table 1: Glossary of terms (1/2)

Name	Description
Persona	A configuration type that can be set for a port, determining if it is connected to workloads or to the network. Can be either `access` for host-facing ports, or `uplink` for fabric-facing ports.
PVLAN	Private virtual LAN
PSM	AMD Pensando Policy and Services Manager
VRF	Virtual Routing and Forwarding instance
VSX	Aruba Virtual Switching Extension, providing high availability and redundancy capabilities
ZTP	Zero-Touch Provisioning: automated network configuration and deployment of managed devices

Table 1: Glossary of terms (2/2)

PSM Overview

The AMD Pensando Policy and Services Manager is a programmable, secure, highly available, centralized system for managing infrastructure policy, with capabilities for:

- Deploying and controlling distributed firewall security, IPsec, NAT and other functions
- Telemetry and analytics
- Troubleshooting
- Operations and maintenance: events, alerts, technical support
- Authentication, authorization, and accounting (AAA)

The PSM is designed to establish and manage consistent policies for a number of Distributed Services Switches. (Refer to the *Aruba CX 10000 Release Notes* for current support limits.)

The PSM operates as a 3-node quorum-based cluster running on virtual machines (VMs) hosted on multiple servers for fault tolerance. A PSM cluster can tolerate the loss of one controller node and continue to maintain full service. The PSM cluster is not involved in datapath operations; if it becomes unreachable or multiple nodes fail, there will be no impact on data traffic and stateful services on the DSSes it manages.

Figure 1 is a diagram of the interconnection between the PSM and the switches it manages; interactions take place through an IP network.

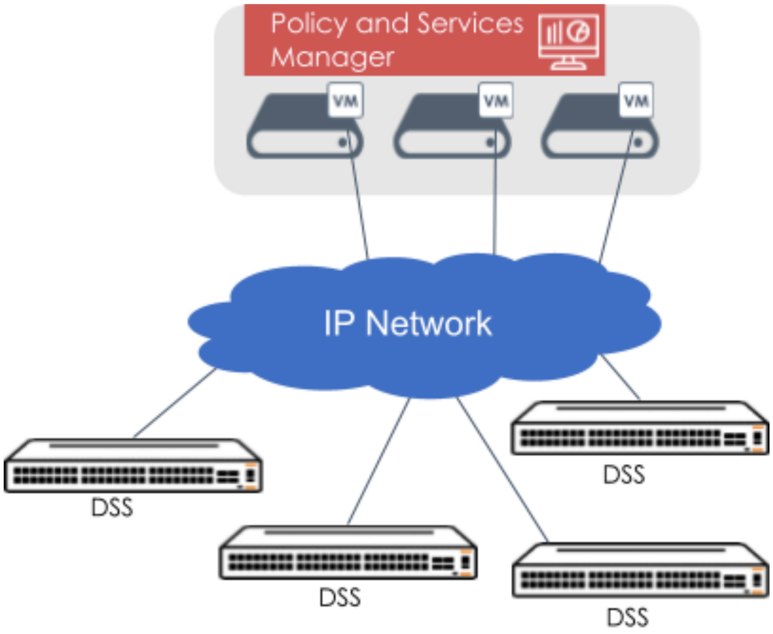


Figure 1. PSM/DSS management plane

Each DSS is configured with an IP address that is used for communication with its associated PSM over any IP network. This is referred to as its *management address*.

Each DSS runs an agent which constantly watches for incoming configuration changes upon which it must take action.

The PSM employs an *intent-based* configuration management structure, similar to Kubernetes. Any configuration changes are continuously monitored within the PSM until it has been confirmed that the changes have been propagated to all DSSes. The PSM resends configuration requests until the desired state is reported back from each DSS, as shown in Figure 2:

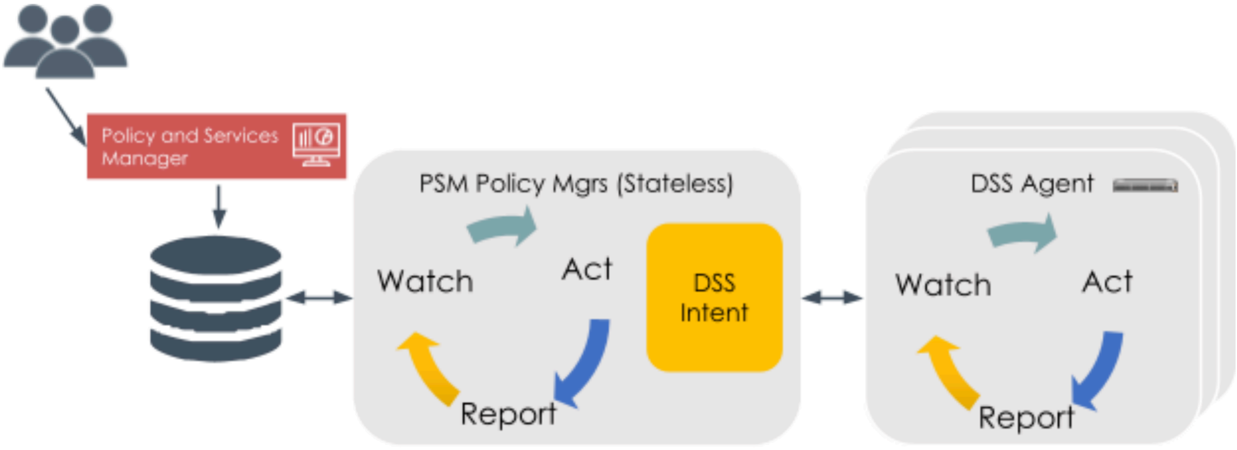


Figure 2. Working principles of intent-based configuration

Intent is expressed in terms of *policies* established for firewall and flow telemetry.

Initial Deployment Workflow: High-Level Overview

This is an outline of the steps necessary for initial deployment of a PSM cluster and its associated DSSes. Detailed steps are provided further below in this document.

- **Install the PSM**
    - Install the PSM software on either an ESX-based or KVM-based 3-node cluster.
    - Configure the PSM using the `bootstrap_PSM.py` utility.
    - Save a copy of the PSM recovery key on a different server from PSM, in case the PSM needs to be rebuilt later as part of disaster recovery. (Refer to Appendix D for more details)
    - Set the PSM user authentication policy, and create PSM users with appropriate roles.

- **DSS Configuration**
  For each DSS:
    - Plan for one additional IP address allocated to each DSS as a management interface, configured either from its host or via DHCP.
    - Associate each DSS to the PSM
    - Admit the DSS into its PSM cluster. The PSM can be configured to do this automatically.

Installation of the PSM cluster is a one-time activity; other procedures may be performed during initial installation, but will also be part of the standard operation of the PSM, performed as more DSSes are added.

PSM Object Model

The PSM's intent-based paradigm relies on the PSM object model described in this section.

Firewall Objects

The primary firewall objects are illustrated in Figure 3. The *NSPRule* (Network Security Policy Rule) specifies the firewall behavior, but is not a managed object itself. Instead, the *NetworkSecurityPolicy*[2] is the managed object that contains an array of NSPRule specifications.

[2] Refer to the Release Notes for the number of NetworkSecurityPolicy objects supported.

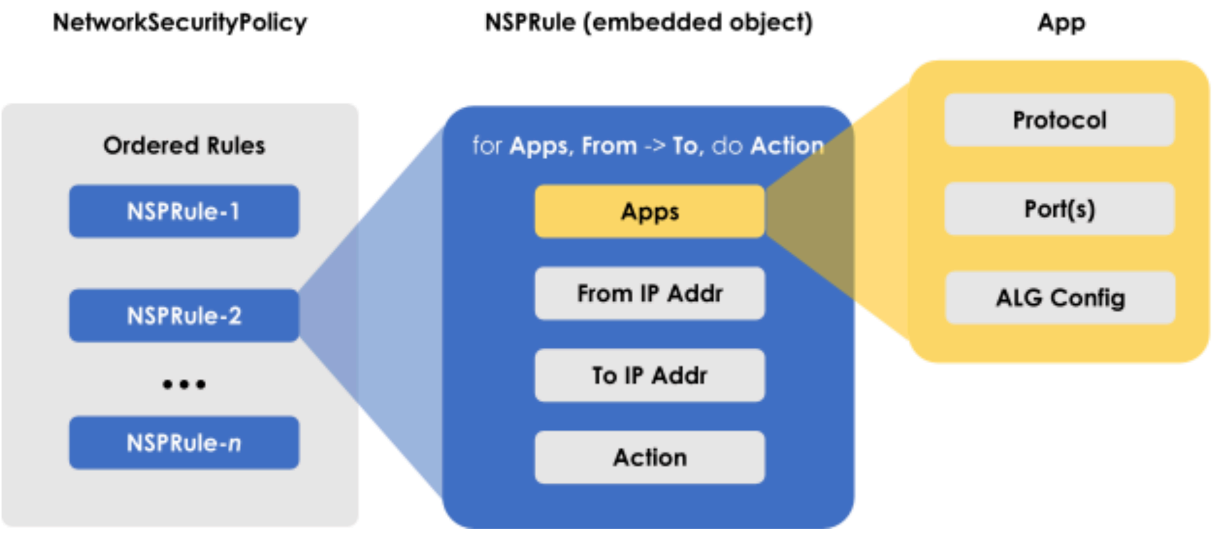


Figure 3. PSM primary firewall objects: NetworkSecurityPolicy, NSPRule, and App

Apps, Network Security Policy

In PSM terminology, an *App* is a service defined either by a protocol/port pair, or by an application level gateway (ALG) for any of several predefined apps. A *Network Security Policy* is a collection of firewall rules governing App connectivity.

Key PSM Objects

Table 3 contains sample key PSM objects. For a complete list, refer to the REST API online help available through the PSM GUI.

Object	Description
Distributed Services Entity	● Entity is a synonym for switch identified by switchname, automatically assigned when a DSS is admitted. Each DSS contains two Distributed Services Modules (DSM).
Tenant	● Individual tenant, supporting future multi-tenancy. Currently only “default” is supported.
VRF	● Virtual Routing and Forwarding. Collection of subnets. ● Identified by name
Network	● A Network object represents a subnet for which security policy is defined and enforced ● Identified by a name ● Contains a VRF, a VLAN ID and the associated ingress/egress security policies to be enforced.
PolicyDistributionTarget	● Is defined by a list of switches to which a policy can be selectively distributed to
Security Policy	● Stateful firewall security policies, defined between network endpoints (IP:port:protocol or apps) ● Identified by policy name ● Contains one or more firewall rules ● Firewall rules can contain IP collections
NAT Policy	● Policy containing one or more NAT rules for SNAT, DNAT or twice NAT defined in a static 1:1 format ● Identified by policy name ● Attached to VRF object with ingress/egress VRF attachment options
IPSec Policy	● Policy containing HA options for active-active and active-standby ● Identified by pair of tunnel definitions ● Each tunnel contains crypto parameters for IKE and IPSec
App	● Describes the networking specification of an application, service or traffic ● Identified by name ● Contains either:

    - List of Ports/Protocols
    - Application Layer Gateway (ALG)
    - Both
- An App object is subsequently referred to by a SecurityPolicy object

Table 3. PSM objects (part 1/2)

Object	Description
Firewall Export Policy	● Syslog destination for firewall logs
FlowExport Policy	● Provides the ability to export FW logs in IPFIX format to external destination/IPFIX collector
Firewall Profiles	● Provides the ability to modify session idle timeout (stateless) and TCP timeout (stateful)
Alert Policies	● A collection of conditions that trigger operator alerts of a given severity type. Triggered alerts are then sent to designated syslog destinations
Event Policies	● The Event stream captures all configuration changes as well as system state changes. All Events can be streamed to designated syslog destinations.

Table 3. PSM objects (part 2/2)

Labels

Each object can be associated with one or more *labels* that can be used to refer to a group of objects, which is a very effective way to enable “administration at scale”.

⚠ ***Note**: Labels that begin with "io.pensando." are reserved for system use, and cannot be created or modified by the user. if the user attempts to create or modify an object's labels with a system label, the label will be silently removed from the user configuration.*

PSM Installation

⚠ ***Note**: Before installation, see the* Release Notes *for the minimum resource requirements needed to operate the PSM, as well as the minimum supported versions of virtualization platforms, AOS-CX and Aruba Fabric Composer in conjunction with a given PSM release.*

The AMD Pensando PSM software is installed on a virtualized compute infrastructure based on VMware ESXi™ or KVM QEMU emulator. The PSM is delivered as an OVA package for ESXi, or a QCOW2 image for KVM. The recommended configuration to ensure high availability is to install three instances of the PSM software on three physical servers; a single-node PSM cluster configuration is also supported for smaller deployments. See the *Release Notes* for the maximum number of DSSes supported per PSM configuration.

Storage Considerations

The storage requirement when firewall logs are exported to the PSM shown in Tables 3 and 3a is based on the assumption that the PSM is receiving 1k logs per second.

With increased firewall logs ingestion rates, it is recommended to take into consideration future requirements when initially sizing necessary disk space. However, extra space can be added to the PSM VMs by adding an additional disk. Follow the specific instructions for the hypervisor the PSM is deployed on.

To estimate current and future storage requirements over time for maximum flow log ingestion rates for either a 3-node or 1-node PSM cluster, use Table 4:

Cluster Size	Ingestion Rate	Total Storage			
		3 Days	7 Days	14 Days	30 Days
3 nodes	Sustained 10k LPS	1.04 TB	1.66 TB	2.76 TB	5.28 TB
1 node	Sustained 2k LPS	206.16 GB	320 GB	516.4 GB	967.6 GB

Table 4. Flow log storage estimator

Data Retention

The PSM has the following retention policy, which is currently not configurable:

- **Events:** retained for 10 days
- **Audit Logs:** retained for 30 days
- **Metrics:** retained for a variable period of time based on roll-up
    - 1 day with 30 second granularity
    - 5 days with 5 minute aggregation
    - 30 days with 1 hour aggregation
    - 1 year with 1 day aggregation
- **Firewall Logs:** retained for up to 30 days (subject to disk space availability, otherwise oldest logs are overwritten).

PSM Installation on ESXi

Notes:

- *Reserve the amount of memory indicated in the "Installation Requirements" section of the* Release Notes *for the VM.*
- *Dedicate a VMFS partition to the VM.*
- *Live migration (VMware vSphere® vMotion®) is supported.*
- *VMware HA is supported.*

The PSM installs as a virtual appliance (OVA format file), deployed through VMware Virtual Center (vCenter) or using the VMware OVF tool. The PSM deployment depends on vApp and requires vCenter or OVF tool for installation."

1. Log in to vCenter. Locate the ESXi host to install a PSM node on and select "Deploy OVF Template" from the Action button.
2. Specify the URI or Local File name of the PSM OVA file `psm.ova`.
3. Specify the PSM VM name.
4. Under the storage section, select Thick Provision.
5. Specify the OVA properties: hostname, IP address, etc.

   ⚠ ***Note**: changing PSM cluster node IP addresses after bootstrapping is not supported.*

   a. If using DHCP, leave the IP address blank, and configure a static MAC address-to-IP binding (reservation) for this host in the DHCP server.
   b. It is strongly recommended that static IP addresses be used.
   c. Under Password, specify the SSH/console password.

6. Review details. Click "Next" to accept the warnings about advanced configuration and lack of Publisher certificate[3].
7. Start the VM in vCenter once the OVA deployment status shows "Completed". The boot process will untar and install the PSM distribution from a read-only partition.
8. When the VM comes up, verify that the hostname has changed to what was specified in the OVA properties above and is not "localhost". If "localhost" appears, then contact Technical Support, as this indicates that the initialization did not complete successfully.
9. Login to the PSM as user `root`, with the password specified in the OVA properties above (if one was not defined in the OVA properties, the default password is `centos`). If a non-default root password is configured, it may take 1 to 2 minutes for the password

[3] In this release, certificates are self-signed, triggering a warning. This may be changed to authority-signed in a future release.

to take effect after the login prompt becomes available. If this is a concern, make sure network access to the VM is disabled until the password has been reset.

> ***Note:*** *Deploying a 3-node cluster involves importing the* `psm.ova` *file once for each VM instance. The number of imports can be reduced by cloning the first VM as a template, and then deploying subsequent VMs from the template. If taking this approach, follow these steps:*

10. Create the first VM from the `psm.ova` file.
11. In vCenter, choose “Clone as Template to Library” to save the VM as a Template (.vmtx) file. Be sure to give the VM a unique name.
12. Select the new VM in vCenter. Select the “Configure” Menu item. Expand “Settings” and select “vApp Options”. Scroll down to “Properties”. Click the radio button for “hostname” and the “Set Value” action to change the hostname to a unique value (typically corresponding to the VM unique name).
13. If applicable, apply any network-specific settings that may have been used in deploying the original VM from the `psm.ova` file.
14. Start the new VM and verify that the VM name and network setting are as intended.
15. Although single-node clusters are supported for smaller production deployments, it is recommended to configure a three node PSM cluster—see the *Release Notes* for the number of DSSes supported in each configuration.
    Repeat the deployment of the OVA for the second and third nodes and assign the IP addresses before proceeding further.
    By default, the PSM’s `autoadmit` parameter is set to `True`. This means that once a DSS is pointed to a PSM to register, it will be immediately admitted into the cluster. Setting its value to `False` provides additional security, as it requires an operator to manually admit each DSS.

PSM Installation on KVM

Notes:

- *Ensure no memory oversubscription of VMs running on the same KVM host.*
- *Dedicate a disk or a Logical Volume to the VM.*

1. Verify the available vCPU, memory, and disk resources:

```
$ virsh nodecpumap
CPUs present:   72
CPUs online:    72
$ virsh nodememstats
```

```
total   :               131520292 KiB
free    :               40785484 KiB
$ df -h
Filesystem                 Size  Used Avail Use% Mounted on
/dev/mapper/centos-root     50G  1.8G   49G   4% /
devtmpfs                    32G     0   32G   0% /dev
tmpfs                       32G     0   32G   0% /dev/shm
tmpfs                       32G  8.9M   32G   1% /run
tmpfs                       32G     0   32G   0% /sys/fs/cgroup
/dev/sda1                 1014M  145M  870M  15% /boot
/dev/mapper/centos-home     57G   33M   57G   1% /home
tmpfs                      6.3G     0  6.3G   0% /run/user/1000
```

2. Verify that the necessary system packages are already installed:

```
$ sudo yum install qemu-kvm qemu-img qemu-system virt-manager virt-
install libvirt libvirt-python libvirt-client bridge-utils -y
```

3. Obtain the PSM qcow2 image.

   Note: *the qcow2 image file will be modified by KVM; make a backup copy of the original image if you need to have a master copy.*

4. Use the following command to deploy the PSM node onto KVM, substituting the actual path to the qcow2 image for the highlighted portion:

```
$ virt-install --import --name PSM1 --virt-type kvm --cpu host-
passthrough --os-variant rhel7.6 --ram 32768 --vcpus 12 --
network=bridge:br0,model=virtio --disk
path=/home/user/PSM.qcow2,format=qcow2,bus=scsi --controller
scsi,model=virtio-scsi --nographics
Starting install...
Connected to domain PSM1
Escape character is ^]
[    0.000000] Initializing cgroup subsys cpuset
[    0.000000] Initializing cgroup subsys cpu
[    0.000000] Initializing cgroup subsys cpuacct

—snip—
```

5. A console connection will activate while the VM is booting. Wait for the login prompt.
6. Log in with user ID `root` and password `centos`.
7. Run the `config_PSM_networking.py` command to set the hostname and password, and configure the static IP for this PSM VM. (Use `config_PSM_networking.py -h` to see all available parameters.)

   Example: set the hostname to `PSM1`, root password to `pensando123`, and configure static IP, gateway, and DNS addresses:

   ```
   [root@localhost ~]# config_PSM_networking.py -m PSM1 -p pensando123 -
   a static -i 203.0.113.49 -n 255.255.252.0 -g 203.0.113.1 -d
   203.0.113,8.8.8.8
   ```

8. Use `ip addr` to verify the newly configured static IP address. Test logging out and logging back in to verify that the new hostname and password have taken effect.
9. Once everything has been verified, power off the VM:

   ```
   $ poweroff
   ```

10. To deploy a PSM with more than one interface (as in this example), use the `virsh` command to add the second interface. The below example assumes the VM name is `PSM1` and the second interface is attached to bridge `br54`.

    ```
    # virsh attach-interface PSM1 bridge br54 --model virtio --config
    ```

11. Power on the VM and log in with the default username `root` and password `centos`.

    ```
    # virsh start PSM1
    ```

12. The second interface will be detected but unconfigured. The first interface should be configured already as part of step 1. Proceed to configure the second interface (if required by your topology). Follow the steps in the next section to finish bootstrapping the PSM.

Bootstrap the PSM Cluster

Before the PSM cluster can be administered, it must be initialized via the `bootstrap_PSM.py` utility. (This utility is in `/usr/local/bin`, which is already in the root user's command path.) Below are some usage examples. The command `bootstrap_PSM.py -h` will show all parameters that can be specified.

Determine the IP address assigned to each PSM VM that has been deployed. In ESXi deployments this can be obtained from vCenter, or from within a CentOS VM with the command `ip addr`. This address is required when launching the `bootstrap_PSM.py` utility and is provided through the `-v` option.

> ***Note:*** *A PSM VM should have a single L3 interface that it uses to communicate to other PSM VMs as well as its DSSes. The IP address of this interface should be used in bootstrapping the cluster. The default IP route should point to this interface.*

The `bootstrap_PSM.py` utility can be used to provide configuration information to the PSM. The following example provides a cluster name, a domain name, and the address of an NTP server, and activates automatic DSS admission. (If a PSM node has multiple interfaces, use the IP address of the PSM interface that will be used to communicate to DSSes during the bootstrap process.)

Example: Bootstrap a 3-node PSM cluster for production, specifying the three nodes' IP addresses. Before executing this command, make sure that all three PSM VMs are already running. The PSM VMs can be deployed using the same OVA, but must have unique IP addresses. The bootstrap script only needs to be executed on one of the nodes.

```
# bootstrap_PSM.py -distributed_services_switch -clustername Demo -domain
training.local -ntpservers 10.29.5.5 -autoadmit True 192.168.71.134
192.168.68.179 192.168.71.49

2021-12-03 00:22:42.786803: * all messages printed to the console will
also be logged to the file: /var/log/pensando/pen-bootstrap-psm.log
2021-12-03 00:22:42.786834: * start PSM bootstrapping process
2021-12-03 00:22:42.786851: * - list of PSM ips: ['192.168.71.134',
'192.168.68.179', '192.168.71.49']
2021-12-03 00:22:42.786860: * - list of ntp servers: ['10.29.5.5']
2021-12-03 00:22:42.786867: * - using domain name: training.local
2021-12-03 00:22:42.786874: * - auto-admit dse: True
2021-12-03 00:22:42.786881: * - cluster-recovery-keys file path: None
2021-12-03 00:22:42.786888: * - accept EULA: False


2021-12-03 00:22:42.786897: * checking for mandatory
distributed_services_switch option
2021-12-03 00:22:42.786904: * checking for reachability
2021-12-03 00:22:46.790891: * connectivity check to 192.168.71.134 passed
2021-12-03 00:22:50.794810: * connectivity check to 192.168.68.179 passed
2021-12-03 00:22:54.798926: * connectivity check to 192.168.71.49 passed
—snip—
```

Notes:

1. *The cluster name must start and end with an alphanumeric character, and can only contain alphanumeric characters, dash, underscore and period.*
2. *If the PSM is not connected to the Internet or can't resolve IP names for other reasons, specify the IP addresses of the NTP servers (as shown in this example) instead of their FQDNs.*

If everything completes successfully, the message below will be seen in the log:

```
—snip—
2021-12-03 00:26:17.707248: * PSM bootstrap completed successfully
2021-12-03 00:26:17.707280: * you may access PSM at https://192.168.71.134
```

The PSM browser GUI and REST API should now be available at any of the PSM addresses. Note that there is no virtual IP address for the 3-node cluster. If this is desired, a load balancer should be installed in front of the cluster.

If an `admin` password is not specified when bootstrapping the PSM cluster, the default is `Pensando0$`.

Once this process completes (which may take a few minutes), open a browser connection to the PSM cluster as described in the next section, allowing you to verify that the PSM is healthy.

> ***Note:*** *After the PSM is bootstrapped, a copy of its recovery key should be generated and stored in a safe location, in case the PSM needs to be rebuilt later as part of disaster recovery. See Appendix D.*

The PSM Graphical User Interface

The PSM Graphical User Interface (GUI) is accessible via a web browser at `https://`*`PSMaddr`*, where *PSMaddr* corresponds to the IP address of any of the PSM cluster nodes. Figure 4 shows a configurable dashboard that offers an overview of the status of the AMD Pensando Distributed Services Platform; the main menu on the left side provides access to the configuration of all supported features.

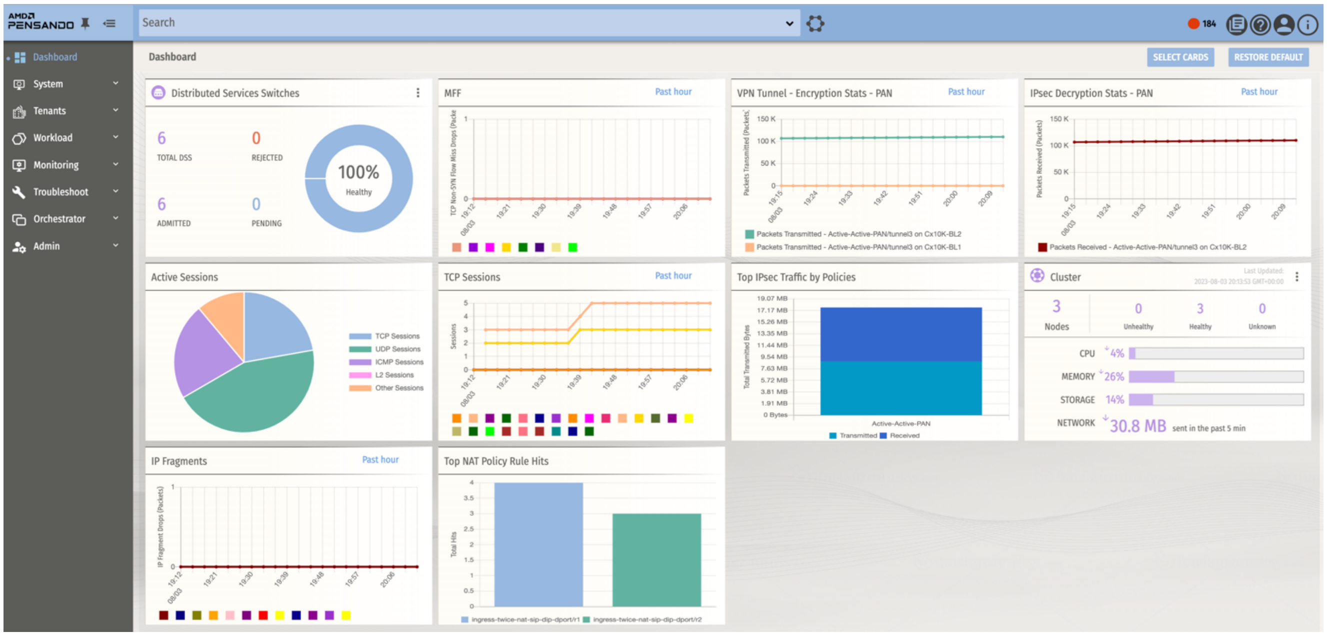
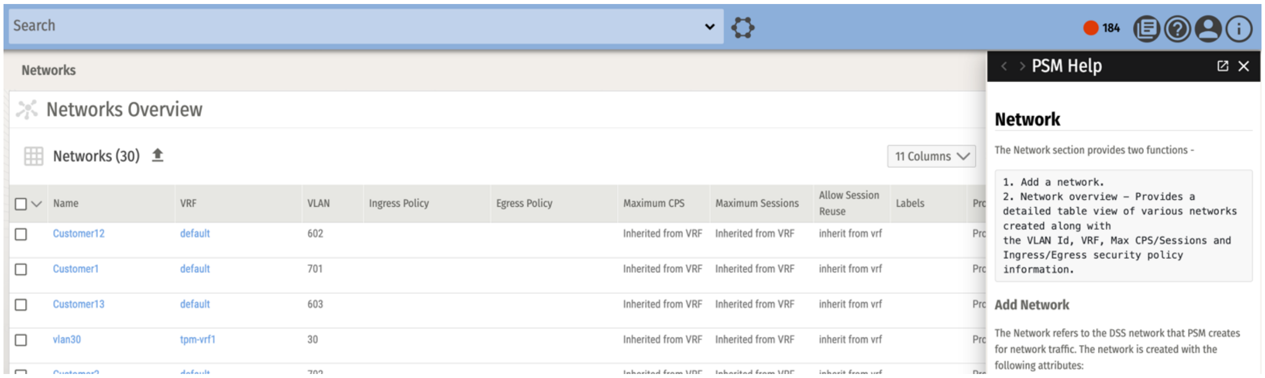


Figure 4. AMD Pensando PSM interface, showing the dashboard, with the “System” section of the main menu open; its two subsections, “Cluster” and “DSS”, are visible and selectable.

Online Help

Detailed and comprehensive online help is offered in a context-sensitive manner for each page in the PSM GUI, accessible through the help icon in the upper right-hand corner: ?

The help icon is context sensitive, showing information related to the currently displayed GUI elements. For example, clicking the help icon while in the Networks overview will display descriptive help and examples on how to create a Network object, as shown in Figure 5. Similarly, clicking the help icon while in the Monitoring -> Alerts & Events view will show help on configuring Alert Policies.



Search

Networks

Networks Overview

Networks (30)

11 Columns

	Name	VRF	VLAN	Ingress Policy	Egress Policy	Maximum CPS	Maximum Sessions	Allow Session Reuse	Labels	Pro
□	Customer12	default	602			Inherited from VRF	Inherited from VRF	inherit from vrf		Pro
□	Customer1	default	701			Inherited from VRF	Inherited from VRF	inherit from vrf		Pro
□	Customer13	default	603			Inherited from VRF	Inherited from VRF	inherit from vrf		Pro
□	vlan30	tpm-vrf1	30			Inherited from VRF	Inherited from VRF	inherit from vrf		Pro

PSM Help

Network

The Network section provides two functions -

```
1. Add a network.
2. Network overview – Provides a
detailed table view of various networks
created along with
the VLAN Id, VRF, Max CPS/Sessions and
Ingress/Egress security policy
information.
```

Add Network

The Network refers to the DSS network that PSM creates for network traffic. The network is created with the following attributes:

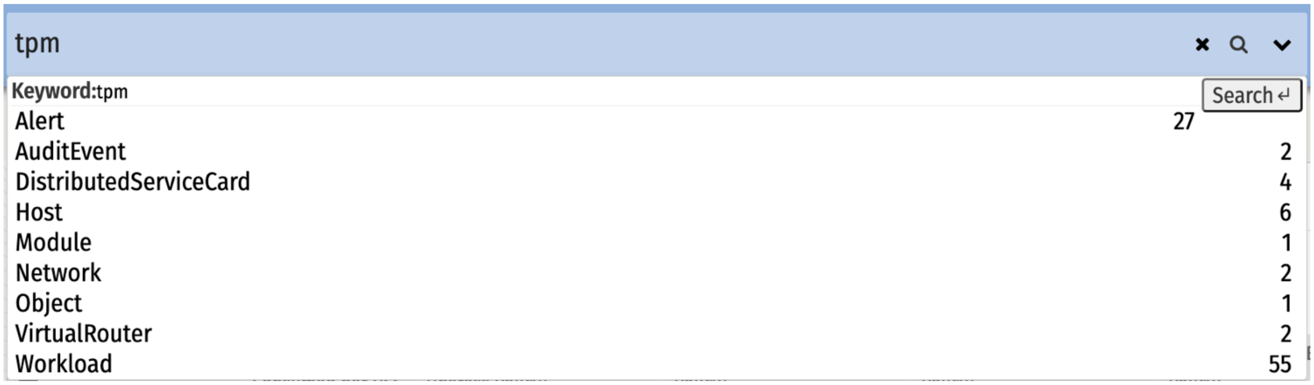
Figure 5. Example of PSM help

Online Help windows can be easily undocked, redocked, resized, or closed.

Online Help has its own presentation context, so it does not need to be closed prior to subsequent operations; selecting different items from the left-hand-side Navigation pane will automatically display the corresponding Online Help information.

Searching

The easy-to-use search facility is accessed from the search bar at the top of the screen.



tpm

Keyword:tpm

Alert	27
AuditEvent	2
DistributedServiceCard	4
Host	6
Module	1
Network	2
Object	1
VirtualRouter	2
Workload	55

Figure 6. PSM search facility

In the example in Figure 6 above, doing a free form text search for the string “ae-s7” shows a summary of the various objects where that string appears, along with a count of the number of occurrences for each object type.

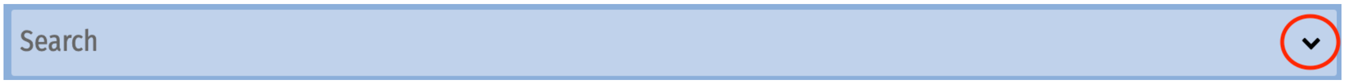
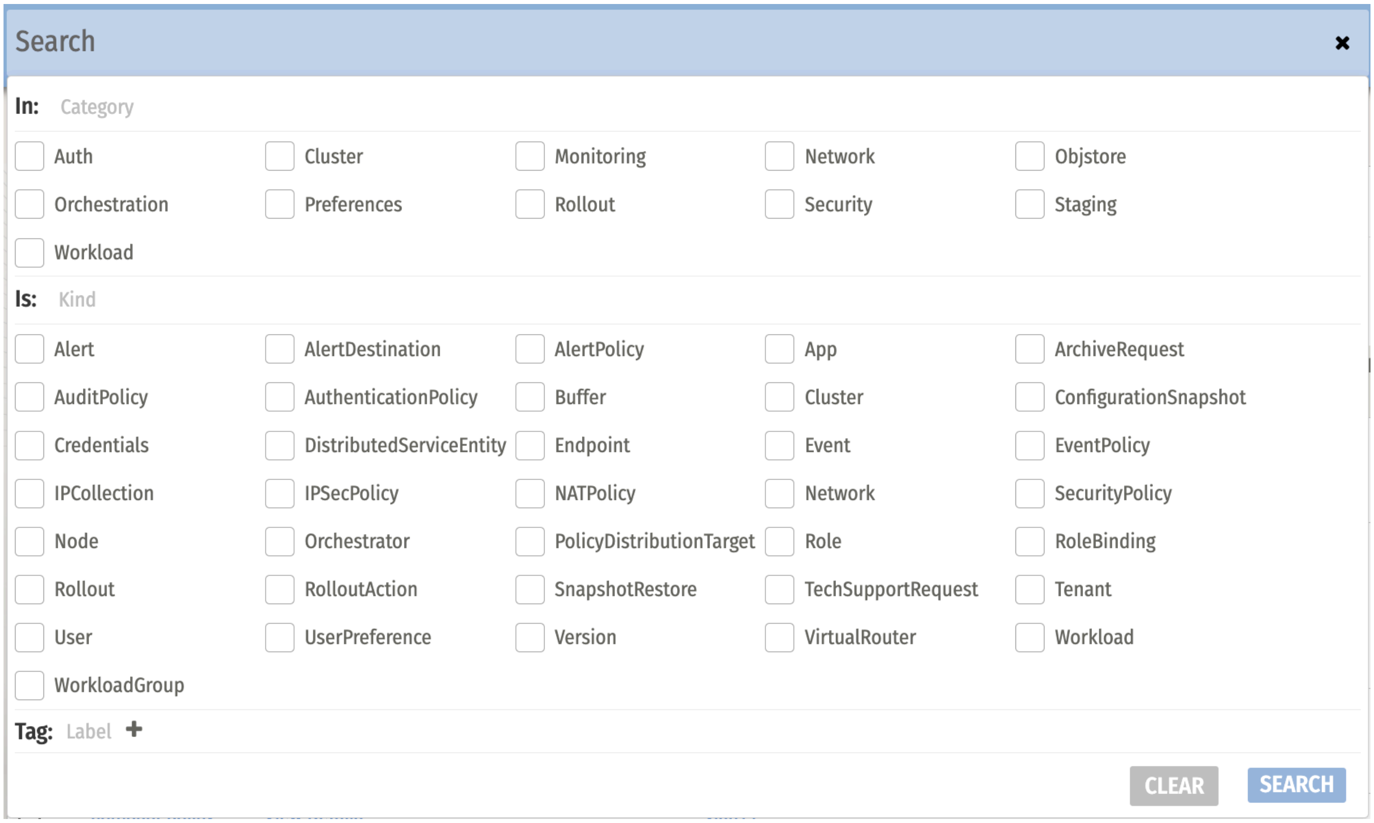


Figure 7. Accessing Advanced Search

Clicking on the downward arrow on the right hand side of the text box (shown in Figure 7 above) gives access to the Advanced Search capability shown in Figure 8, where users can search based on object Category, Kind or Tag (arbitrary labels associated to objects):



Search ✖

In: Category

□ Auth □ Cluster □ Monitoring □ Network □ Objstore
□ Orchestration □ Preferences □ Rollout □ Security □ Staging
□ Workload

Is: Kind

□ Alert □ AlertDestination □ AlertPolicy □ App □ ArchiveRequest
□ AuditPolicy □ AuthenticationPolicy □ Buffer □ Cluster □ ConfigurationSnapshot
□ Credentials □ DistributedServiceEntity □ Endpoint □ Event □ EventPolicy
□ IPCollection □ IPSecPolicy □ NATPolicy □ Network □ SecurityPolicy
□ Node □ Orchestrator □ PolicyDistributionTarget □ Role □ RoleBinding
□ Rollout □ RolloutAction □ SnapshotRestore □ TechSupportRequest □ Tenant
□ User □ UserPreference □ Version □ VirtualRouter □ Workload
□ WorkloadGroup

Tag: Label +

CLEAR SEARCH

Figure 8. Advanced Search

All the keywords used in Advanced Search can also be typed directly into the search bar to avoid having to bring up the Advanced Search tab.

Global Icons

The GUI makes use of common/global icons for many actions, regardless of context, such as “Edit” or “Delete”, which can be used to edit fields, add labels, or delete objects such as “VRF”, “Networks”, and “Security Policies”, as shown in Figure 9 :

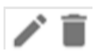
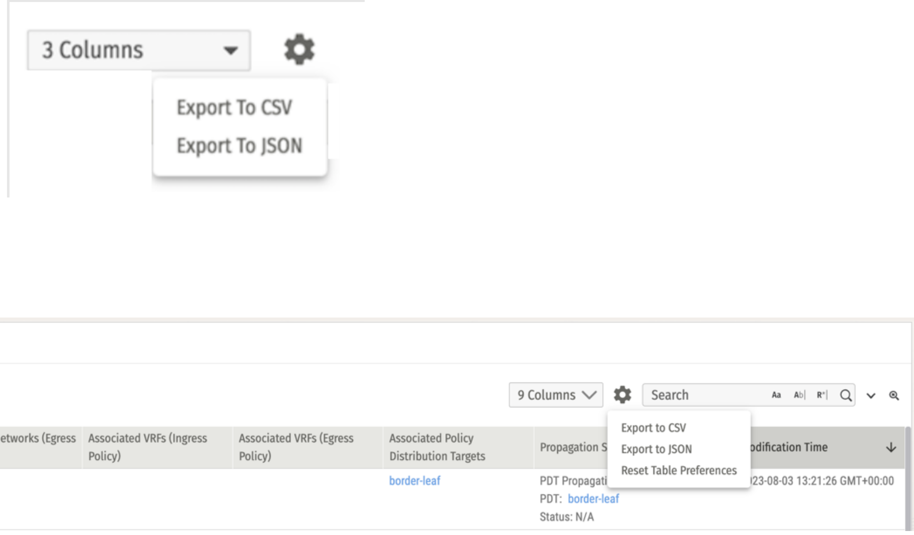


Figure 9. Edit and Delete icons

Many of the tables displayed in the GUI can be exported as CSV or JSON text files, as shown in Figure 10:



☐	Policy Name	Policy/Rule Entries Consumed per DSS	Associated Networks (Ingress Policy)	Associated Networks (Egress Policy)	Associated VRFs (Ingress Policy)	Associated VRFs (Egress Policy)	Associated Policy Distribution Targets	Propagation S	odification Time
☐	FR-FRRHAFCMFSS0	View Details					border-leaf	PDT Propagati PDT: border-leaf Status: N/A	23-08-03 13:21:26 GMT+00:00

Figure 10. Example of how a table can be exported in CSV or JSON format

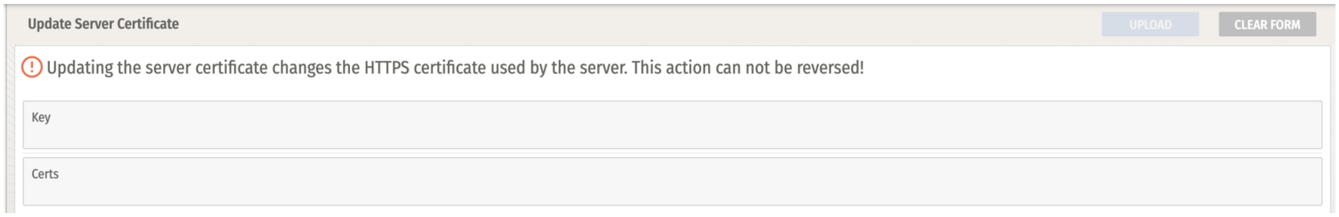
Server Certificate

By default, the PSM uses a self-signed certificate to authenticate itself to browser-based GUI or REST clients.

Sites may instead provide a custom key and certificate for the PSM to use. If the root certificate authority (CA) of the custom certificate is included in either the browser or the client hosts' trusted root CA certificate list, warning messages related to certificate validity will no longer be shown when accessing the PSM cluster login page.

The two supported encoded key formats are RSA and ECDSA. To change the PSM certificate, click "Admin" --> "Server Certificate". On the top right hand side, click "UPDATE". Enter the key and certificate in Privacy Enhanced Mail (PEM) format and then click "Upload" to apply the change, as shown in Figure 11.

> ***Note:*** *this action will not disrupt existing connections, even if they were established with a previous certificate.*



Update Server Certificate

Updating the server certificate changes the HTTPS certificate used by the server. This action can not be reversed!

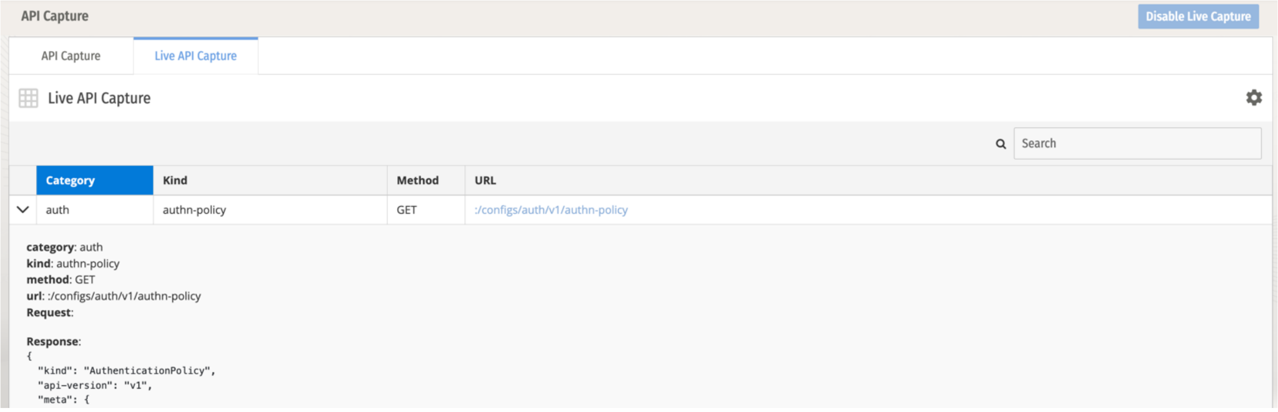
Figure 11. Changing the PSM server certificate

API Capture

Users of the PSM's REST API for external integration can take advantage of the fact that the PSM GUI itself uses this API, to see examples of it in use. REST API calls sent from the GUI to the PSM as it implements the configurations created by the user can be examined using the API Capture feature.

When the API Capture menu item is selected, a view as shown in Figure 12 appears. Use this screen to browse sample API calls (in the API Capture tab) or a live capture of APIs generated while navigating the GUI (in the Live API Capture tab).

The scope of the live capture tool results is per GUI session; large responses are trimmed down to two records to present the look and feel of the response, rather than the entire response.



	Category	Kind	Method	URL
⌄	auth	authn-policy	GET	:/configs/auth/v1/authn-policy

category: auth
kind: authn-policy
method: GET
url: :/configs/auth/v1/authn-policy
Request:

Response:

```
{
  "kind": "AuthenticationPolicy",
  "api-version": "v1",
  "meta": {
```

Figure 12. Example of captured REST API calls

Create PSM User Authentication Policy and Users

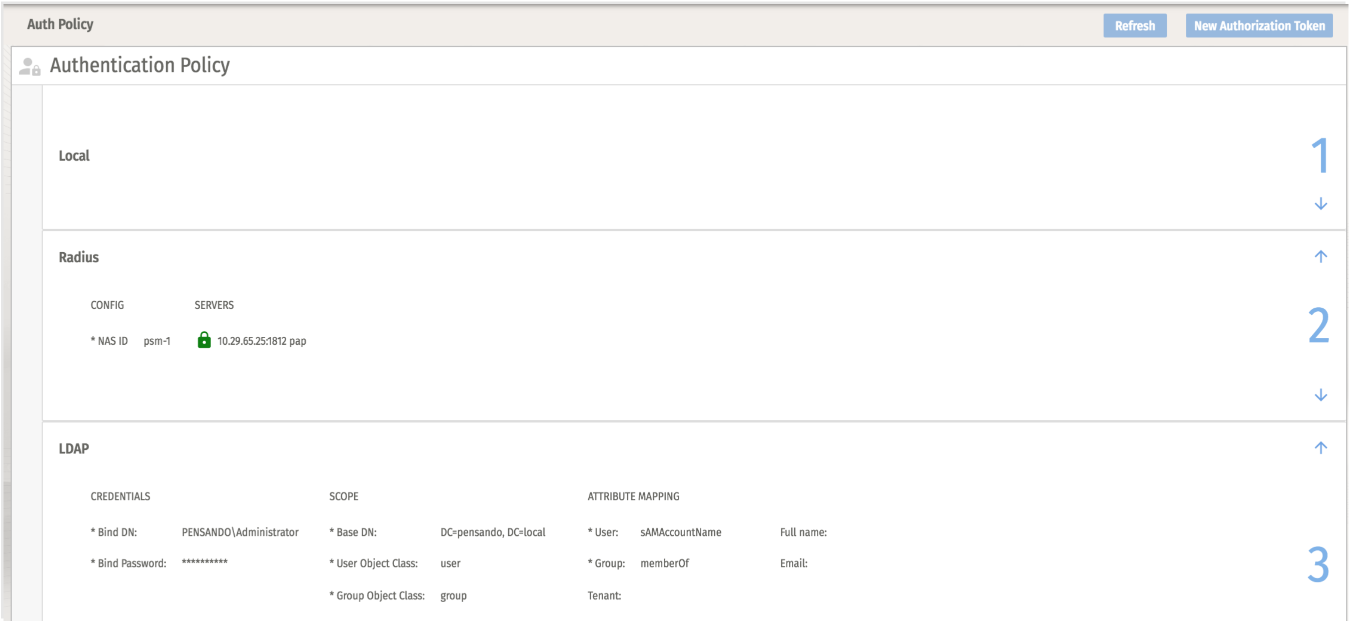
Each PSM will have one or more users defined. Users are assigned roles granting them privileges depending on the tasks they need to perform; during the installation process, an initial user, named `admin` with the password `Pensando0$` , is created with full administrator privileges.

A different name for the initial user as well as a custom password can be provided as parameters to the `bootstrap_PSM.py` utility used to initialize the cluster, as described in the section “Bootstrap The PSM Cluster”.

User Authentication Policy

The PSM supports Local (i.e., username and password), LDAP and RADIUS authenticators, as shown in Figure 13. Creation of authenticators should be done early in the system setup process. Once two or more authenticators are created, they can be re-ordered dynamically to specify the priority with which they should be applied.

See the *PSM LDAP Configuration Guide* for specific configuration details.



Local

Radius

CONFIG

* NAS ID psm-1

SERVERS

🔒 10.29.65.25:1812 pap

LDAP

CREDENTIALS

* Bind DN: PENSANDO\Administrator

* Bind Password: **********

SCOPE

* Base DN: DC=pensando, DC=local

* User Object Class: user

* Group Object Class: group

ATTRIBUTE MAPPING

* User: sAMAccountName

* Group: memberOf

Tenant:

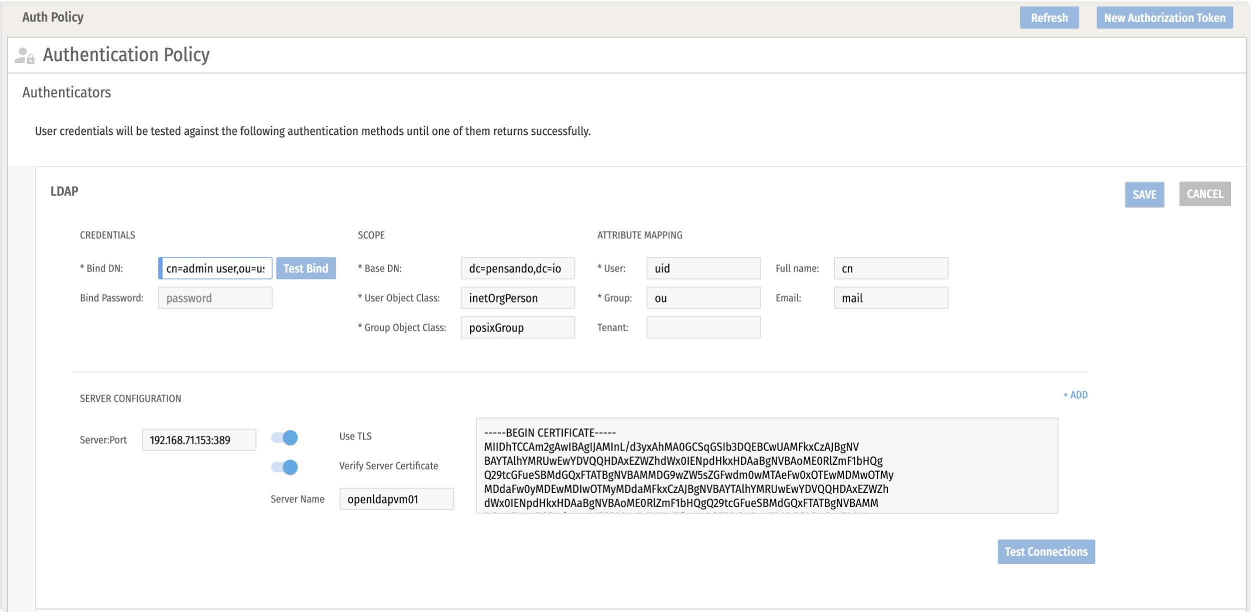
Full name:

Email:

Figure 13. Managing user authentication policy

To create an LDAP Authenticator, click the “CREATE LDAP AUTHENTICATOR” button. Active Directory (AD) and OpenLDAP providers are supported.

Configure *Credentials*, *Scope* (which controls user and group entry search) and *Attribute* mapping (which contains the name of the attributes that should be extracted from the LDAP entry of the user, such as full name and email address) as appropriate, ensuring all required (*) fields are properly filled, as in Figure 14:



User credentials will be tested against the following authentication methods until one of them returns successfully.

CREDENTIALS

* Bind DN: cn=admin user,ou=u:

Bind Password: password

SCOPE

* Base DN: dc=pensando,dc=io

* User Object Class: inetOrgPerson

* Group Object Class: posixGroup

ATTRIBUTE MAPPING

* User: uid

* Group: ou

Tenant:

Full name: cn

Email: mail

SERVER CONFIGURATION

Server:Port 192.168.71.153:389

Use TLS

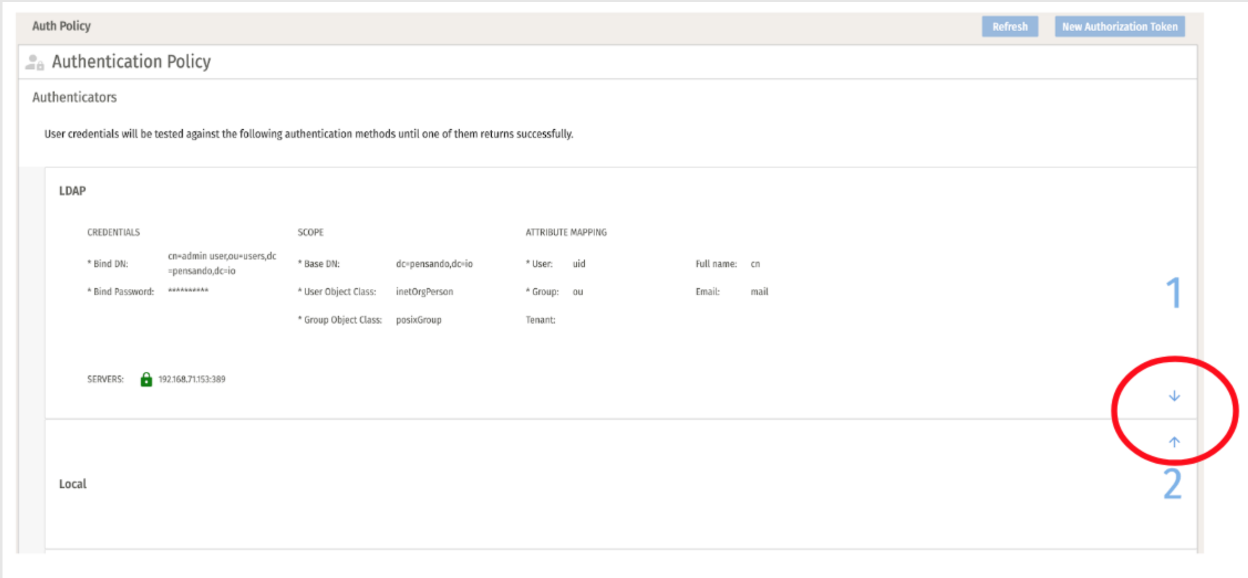
Verify Server Certificate

Server Name openldapvm01

```
-----BEGIN CERTIFICATE-----
MIIDhTCCAm2gAwIBAgIJAMInL/d3yxAhMA0GCSqGSIb3DQEBCwUAMFkxCzAJBgNV
BAYTAlhYMRUwEwYDVQQHDAxEZWZhdWx0IENpdHkxHDAaBgNVBAoME0RlZmF1bHQg
Q29tcGFueSBMdGQxFTATBgNVBAMMDG9wZW5sZGFwdm0wMTAeFw0xOTEwMDMwOTMy
MDdaFw0yMDEwMDIwOTMyMDdaMFkxCzAJBgNVBAYTAlhYMRUwEwYDVQQHDAxEZWZh
dWx0IENpdHkxHDAaBgNVBAoME0RlZmF1bHQgQ29tcGFueSBMdGQxFTATBgNVBAMM
```

Figure 14. LDAP configuration

Once saved, the values should be visible, as shown in Figure 15. The order of the various authenticators can be changed (using the small arrows on the right-hand side).



User credentials will be tested against the following authentication methods until one of them returns successfully.

CREDENTIALS

* Bind DN: cn=admin user,ou=users,dc =pensando,dc=io

* Bind Password: **********

SCOPE

* Base DN: dc=pensando,dc=io

* User Object Class: inetOrgPerson

* Group Object Class: posixGroup

ATTRIBUTE MAPPING

* User: uid

* Group: ou

Tenant:

Full name: cn

Email: mail

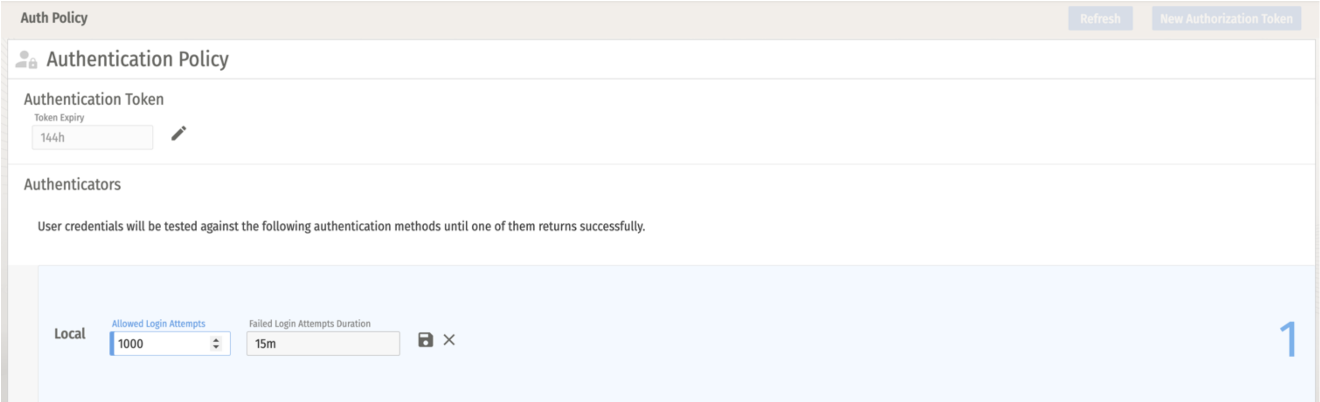
SERVERS: 192.168.71.153:389

Figure 15. Changing authentication order

Local User Lockout Policy

By default, any local account on the PSM will be locked out after 10 incorrect login attempts within a 15 minute interval. This policy applies to all PSM local accounts, including admin accounts. Locked accounts can be unlocked by an admin account.

To change the values for the maximum number of attempts, navigate to *Admin → Auth Policy*. Click the Edit icon (✎) on the Local Authenticators line and set a new value. Click on the Save (💾) icon when done.



Authentication Token

Token Expiry: 144h

Authenticators

User credentials will be tested against the following authentication methods until one of them returns successfully.

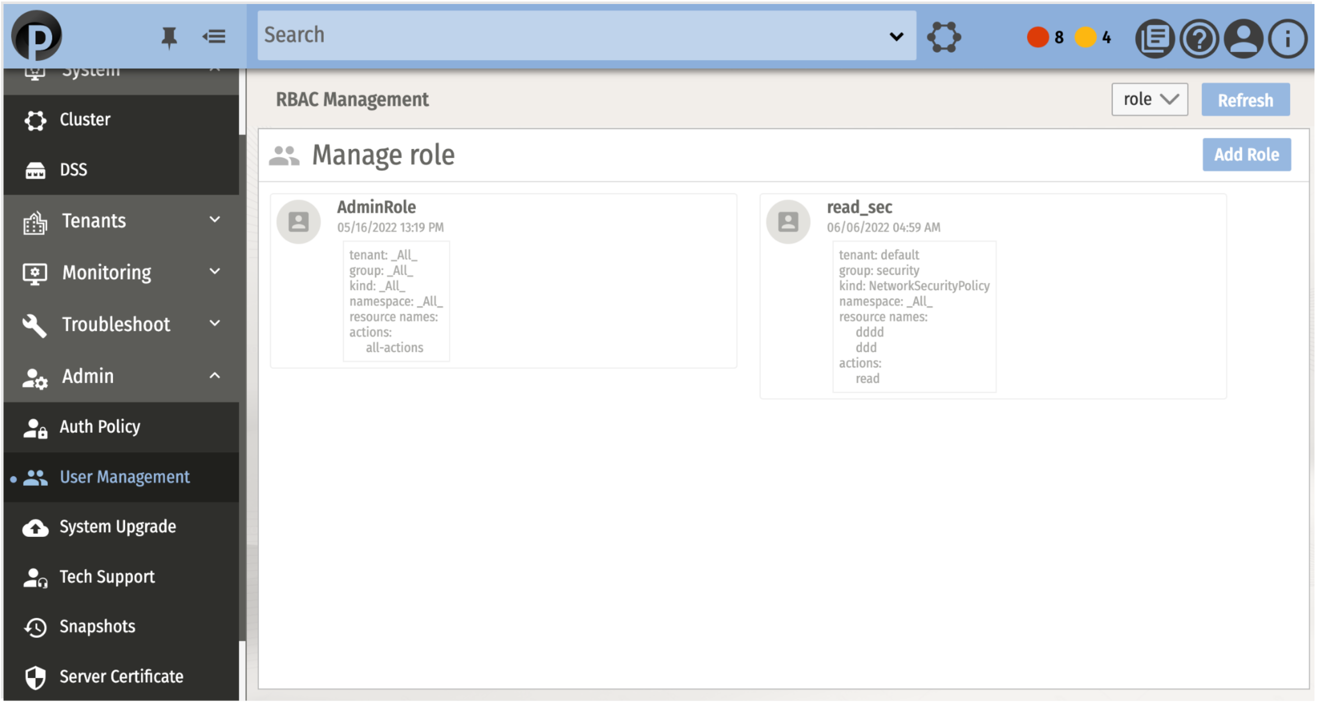
Local — Allowed Login Attempts: 1000 — Failed Login Attempts Duration: 15m

Figure 16. Changing lockout policy

It is strongly recommended that sites create a backup admin local account on the PSM in case all other admin accounts become locked.

Role-Based Access Control (RBAC)

The User Management menu gives access to the RBAC Management screen, shown in Figure 17, which allows management of users, roles, and the association of users to roles. Each action can be selected with the drop-down menu on the top right corner. It is recommended to first create one or more Roles, followed by one or more Users, and then create the corresponding associations between Roles and Users (rolebinding).



AdminRole
05/16/2022 13:19 PM

```
tenant: _All_
group: _All_
kind: _All_
namespace: _All_
resource names:
actions:
    all-actions
```

read_sec
06/06/2022 04:59 AM

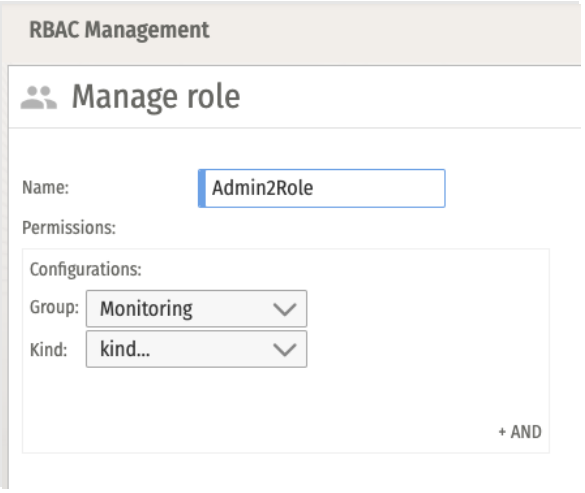
```
tenant: default
group: security
kind: NetworkSecurityPolicy
namespace: _All_
resource names:
    dddd
    ddd
actions:
    read
```

Figure 17. RBAC management

Roles

Roles are created to control access to classes of features by sets of users. Roles can have scope over various objects, which are grouped by the PSM in the categories:

- Auth
- Cluster
- Diagnostics
- Monitoring
- Network
- Objstore
- Preferences
- Rollout
- Security
- Staging



RBAC Management

Manage role

Name: Admin2Role

Permissions:

Configurations:

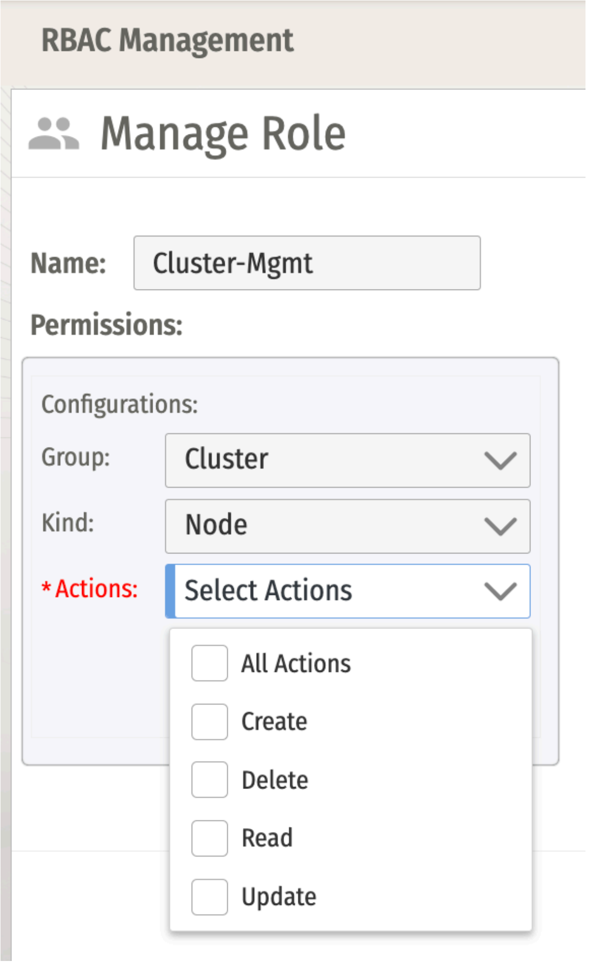
Group: Monitoring

Kind: kind...

Figure 18. RBAC roles

As shown in Figure 18 above, various kinds of management aspects are available for a given Group. Once one is selected, access to actions can be added or removed, as shown in Figure 19:

- Create
- Read
- Update
- Delete



RBAC Management

Manage Role

Name: Cluster-Mgmt

Permissions:

Configurations:

Group: Cluster

Kind: Node

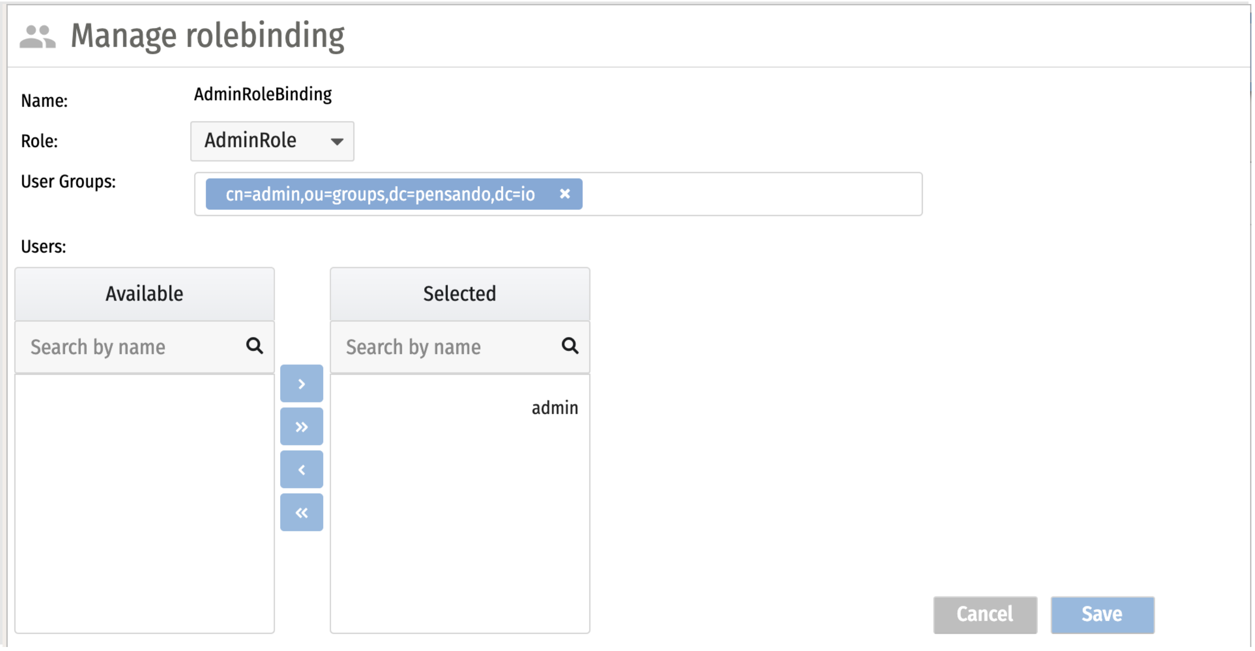
*Actions: Select Actions

Figure 19. Assigning actions to a group

Role Binding

Once a Role is created, a corresponding rolebinding is automatically created. Rolebindings allow users to be flexibly mapped to various sets of roles. Figure 20 shows the view to modify a “rolebinding” that allows to associate any of the users defined in the system (in the left list titled Available) with the Role specified in the form. Users successfully associated with the Role appear in the right list titled Selected.

The rolebinding can be also specified using the “Group” attribute value configured in the LDAP authenticator “Attribute Mapping” section and retrieved from the LDAP user entry. This is the distinguished name of the LDAP group entry to which a user belongs.


Manage rolebinding

Name: AdminRoleBinding

Role: AdminRole

User Groups: cn=admin,ou=groups,dc=pensando,dc=io ×

Users:

Figure 20. Rolebinding

System Upgrade

The PSM platform nodes can be upgraded using the “System Upgrade” option, available in the Admin menu. The process is broken into two tasks:

- Download the upgrade bundle from your vendor’s support site.
- Upgrade the PSM cluster (approximately 10 minutes per cluster node).

When upgrading a fabric, make sure that AFC (if present) is upgraded first, then the PSM cluster nodes, and finally the CX 10000 switches.

See the PSM release notes for supported upgrade paths for AOS-CX, PSM and AFC.

AFC Upgrade

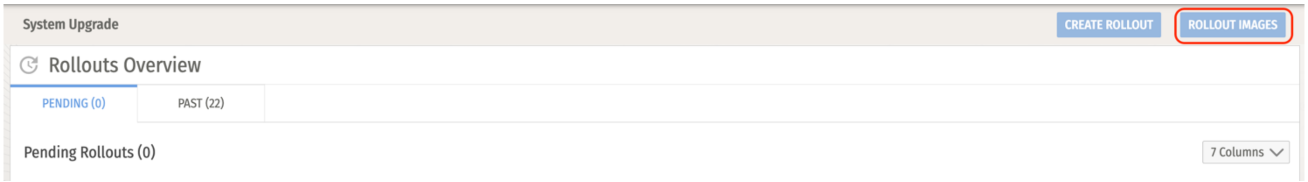
Upgrading AFC is documented in the *Aruba Fabric Composer Installation Guide*.

PSM Upgrade

Upload PSM Upgrade Bundle

1. Download the PSM upgrade bundle. This bundle is part of the software release package, and will have a file name like

`psm_upgrade_bundle_`*`release_number`*`.tar` . Once this file is downloaded, upload it to the PSM. Click on the ROLLOUT IMAGES button in the Admin > System Upgrade view.



System Upgrade

Rollouts Overview

Pending Rollouts (0)

Figure 21. Rollouts Overview screen

2. Click on “Upload Image File”:

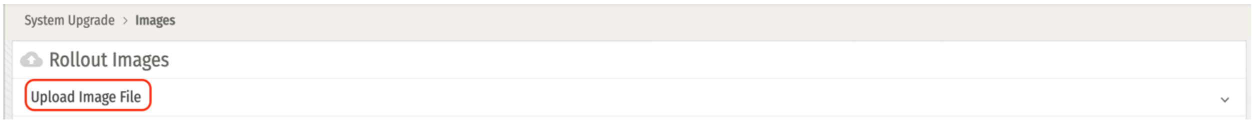
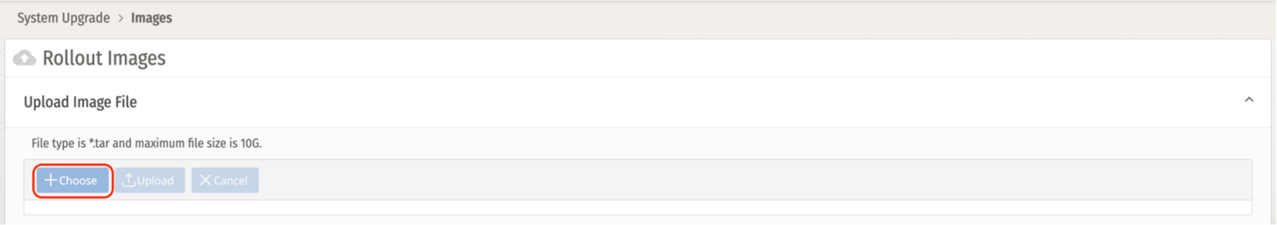


Figure 22. Bring up image file selector

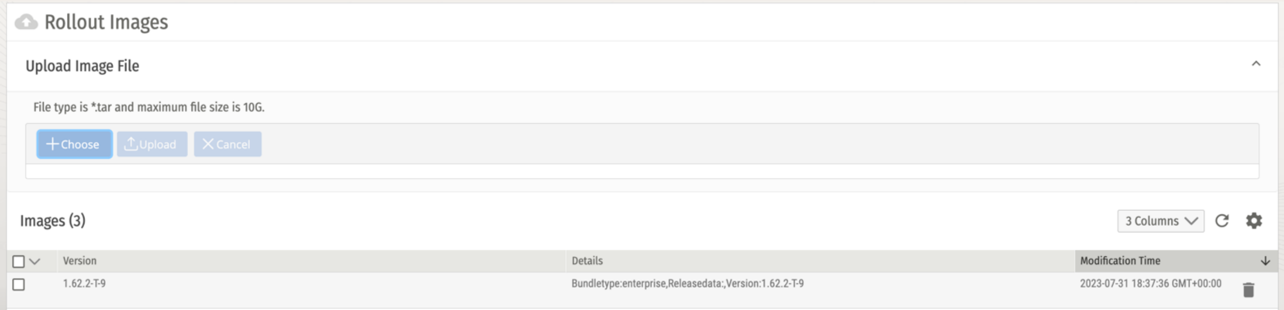
3. Click on “Choose” and select the PSM upgrade image to be uploaded.



File type is *.tar and maximum file size is 10G.

Figure 23. Choose rollout image file

Once uploaded, the new bundle will show in the Images repository, and will now be available to be included in a rollout configuration.



File type is *.tar and maximum file size is 10G.

Images (3)

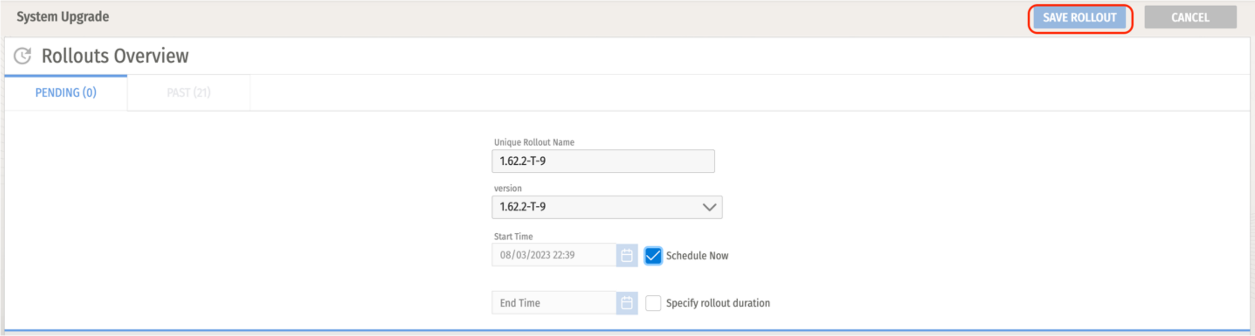
□	Version	Details	Modification Time
□	1.62.2-T-9	Bundletype:enterprise,Releasedata:,Version:1.62.2-T-9	2023-07-31 18:37:36 GMT+00:00

Figure 24. New upgrade image is available to select.

Create Rollout

Once the upgrade bundle is uploaded to PSM, create a rollout. The rollout can be scheduled to start at a particular date/time or immediately.

1. Return to the Admin -> System Upgrade screen (as in Figure 21), and click “Create Rollout”. The rollout creation form will appear, as shown in Figure 25.



Unique Rollout Name: 1.62.2-T-9

version: 1.62.2-T-9

Start Time: 08/03/2023 22:39 ☑ Schedule Now

End Time □ Specify rollout duration

Figure 25. Specify name, version, and time to begin upgrade.

- **Unique Rollout Name:** Name of the rollout; can be any arbitrary name
- **Version:** Choose the version to upgrade to. The choices are auto-populated from the image(s) uploaded to the PSM.
- **Start time**: Time when the rollout should start.
    - **Schedule Now**: Select this to start the upgrade as soon as “Save Rollout” is applied.

2. Click on SAVE ROLLOUT to start the upgrade. If “Schedule Now” was selected, the PSM upgrade will start immediately.

AOS-CX Upgrade

For upgrading AOS-CX on the switches being managed, consult Aruba documentation at https://asp.arubanetworks.com/. High level instructions are provided here for quick reference.

1. Copy image to non-current boot bank on the DSS using the command

   ```
   switch# copy scp://server/path/switch_image.swi
   (primary|secondary) (vrf mgmt)
   ```

   a. The existing image in the primary or secondary location will be overwritten.
   b. Use the appropriate VRF.

2. Verify that the image has been copied to the DSS:

   ```
   switch# show images
   ```

3. Use the `boot system` command as a trial check to see whether any device component updates are required:

   ```
   switch# boot system (primary|secondary)
   ```

   Enter `n` when prompted to continue.

   If any non-failsafe device updates are required, enter configuration mode and enter `allow-unsafe-updates 30`.

   If AFC is used to upgrade the switches, the next step can be skipped.

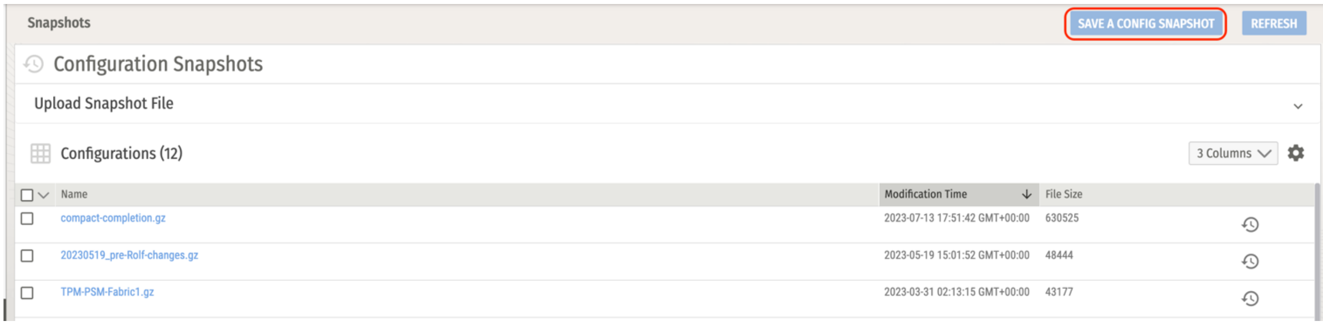
4. Start the upgrade using the command

   ```
   switch# boot system (primary|secondary)
   ```

Consult Aruba AOS-CX documentation for information on related topics such as VSX Live Upgrade.

Configuration Snapshots

The PSM configuration can be saved and later restored via *Snapshots*. A PSM configuration Snapshot can be created by going to `Admin->Snapshots`, and clicking on SAVE A CONFIG SNAPSHOT, as shown in Figure 26.

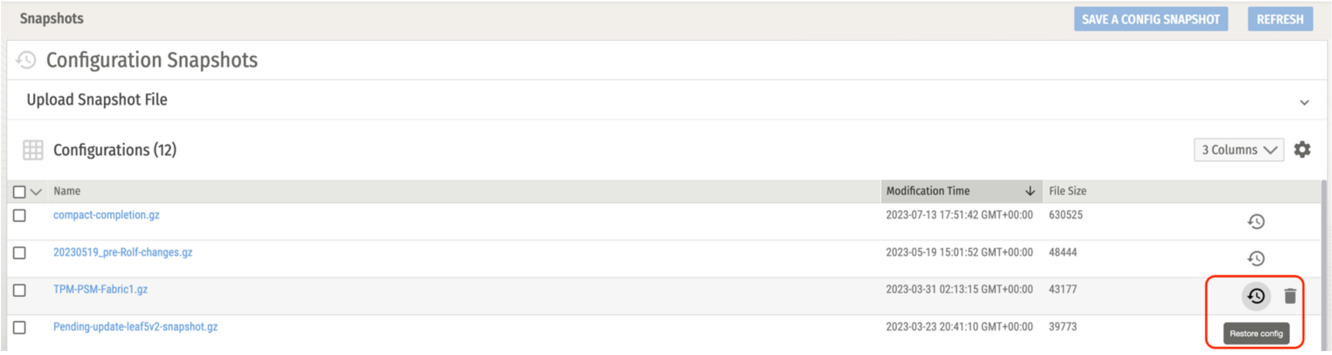


□	Name	Modification Time	File Size
□	compact-completion.gz	2023-07-13 17:51:42 GMT+00:00	630525
□	20230519_pre-Rolf-changes.gz	2023-05-19 15:01:52 GMT+00:00	48444
□	TPM-PSM-Fabric1.gz	2023-03-31 02:13:15 GMT+00:00	43177

Figure 26. Configuration Snapshots screen

You may specify a snapshot name, in which case the snapshot will be saved in the local file *`name`*`.gz`; otherwise the snapshot will be saved with a filename based on the current day and time.

To restore the state of the PSM to an earlier configuration, click the “Restore config” icon corresponding to the snapshot, as shown in Figure 27.



□	Name	Modification Time	File Size
□	compact-completion.gz	2023-07-13 17:51:42 GMT+00:00	630525
□	20230519_pre-Rolf-changes.gz	2023-05-19 15:01:52 GMT+00:00	48444
□	TPM-PSM-Fabric1.gz	2023-03-31 02:13:15 GMT+00:00	43177
□	Pending-update-leaf5v2-snapshot.gz	2023-03-23 20:41:10 GMT+00:00	39773

Figure 27. Accessing the Restore config icon

The PSM will be unavailable during the configuration restore process.

Configuration backups are backward compatible: a configuration snapshot taken with an older version of the PSM software can be restored on a later version of the PSM. Features that were not supported when the snapshot was taken will not be configured upon restore.

Associating a DSS with the PSM

A DSS can be associated with a PSM using:

- The Aruba AOS-CX CLI
- ZTP (Zero Touch Provisioning), by passing PSM network addresses
- Aruba Fabric Composer (AFC) for day 0 provisioning

Note: Before associating a DSS with a PSM:

1. Ensure that there will be Layer 3 connectivity between the DSS and the PSM through the management or inband interface. Make sure that the PSM service ports are open for any firewall in-between; see the list of PSM-to-DSS ports in Appendix B.
2. Make sure that the DSS is not currently associated with another PSM. If you need to move a DSS from one PSM cluster to another PSM cluster, follow the steps in the “Decommissioning a DSS” section.

AOS-CX CLI

***Note**: Verify that the time is set correctly on the switch.*

1. Enter configuration mode on the switch and type `psm` to get into the PSM configuration section. To show the commands that you can execute, enter the ? symbol (the ? will not display on the screen when you enter it).

```
switch# conf t
switch(config)# psm
switch (config-psm)# ?
  end   End current mode and change to enable mode.
  exit  Exit current mode and change to previous mode
  host  Configure the Policy and Services Manager address
  list  Print command list
  no    Negate a command or set its defaults
  show  Show running system information
```

2. Use the `host` command to provide the IP addresses for the three nodes of the PSM to associate the DSS with.

```
switch(config-psm)# host 203.0.113.49 203.0.113.50 203.0.113.51
vrf default
```

3. Each switch is typically managed via the out-of-band management port. In order for the PSM can be managed over the in-band (data) network, it must have either the IP address configured on the front-end port or be enabled by the following command.:

```
switch# conf t
switch(config)# ip source-interface psm ip address vrf vrf name
```

ZTP

Add the below options to `dhcp.conf`; see the Aruba Using ZTP to Provision a Managed Device documentation for further information.

```
// Global options:

# Need this for CX 10000 ZTP
option local-proxy-config code 252 = text;
option space Aruba code width 1 length width 1 hash size 4;
option Aruba.image-file code 145                = text;
option Aruba.conf-file code 144                 = text;
option Aruba.cop-location code 146              = text;
option Aruba.http-proxy-location code 148       = text;

host cx10000-xyz {
    #### DSS
    hardware ethernet 04:90:81:00:0a:ab;
    fixed-address 10.30.2.49;
    option tftp-server-name        "10.30.4.188";
    vendor-option-space Aruba;
    option Aruba.conf-file         "p0-29-running-config.cfg";
}
```

Add the lines highlighted in the example above to specify the TFTP server IP address and configuration file to use (10.30.4.188 and p0-29-running-config.cfg in the example).

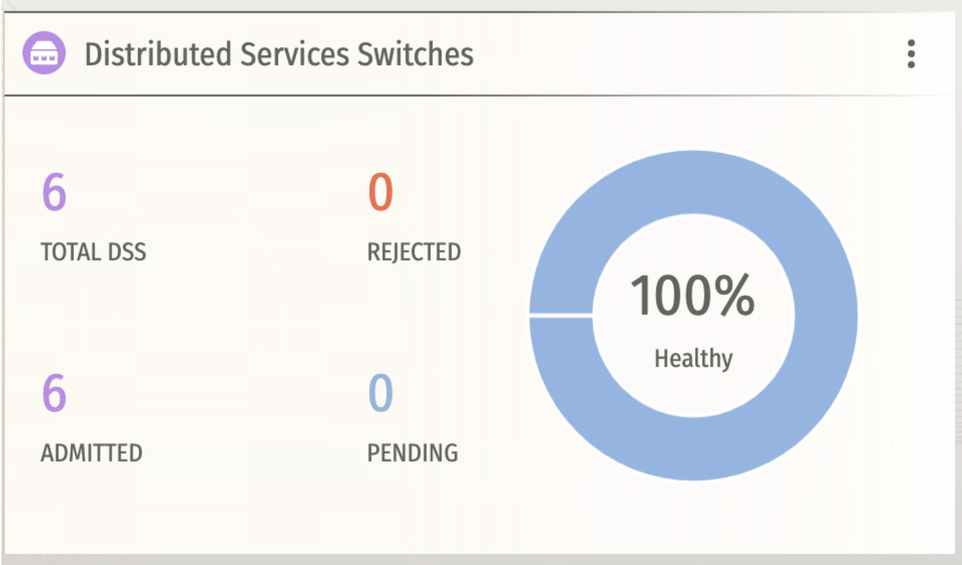
The configuration file is a regular ASCII file containing the PSM IP addresses.

AFC

Refer to the “PSM Integration” section of the *Aruba Fabric Composer User Guide* for details.

Verification

Once the association is completed, the DSS should show up in the PSM GUI.[4]



Distributed Services Switches

6 TOTAL DSS

0 REJECTED

6 ADMITTED

0 PENDING

100% Healthy

Figure 28. The PSM dashboard DSS status card, showing the number of admitted and pending DSSes. See also the System->DSS view, for a list of DSSes associated with this PSM.

[4] The DSS may need to be manually admitted, depending on the `autoadmit` admission policy selected when bootstrapping the PSM quorum, as discussed above.

The AOS-CX CLI can also be used to verify the existing configuration:

```
switch# show psm

Policy and Services Manager Information

Operational Status      : admitted
Host Addresses          : 10.30.28.197, 10.30.28.198, 10.30.28.199
VRF                     : default
PSM Identifier          : 10.30.18.166
Version                 : 1.62.2-T-9
Operational Addresses : 10.30.28.197, 10.30.28.198, 10.30.28.199
switch#
```

Decommissioning a DSS

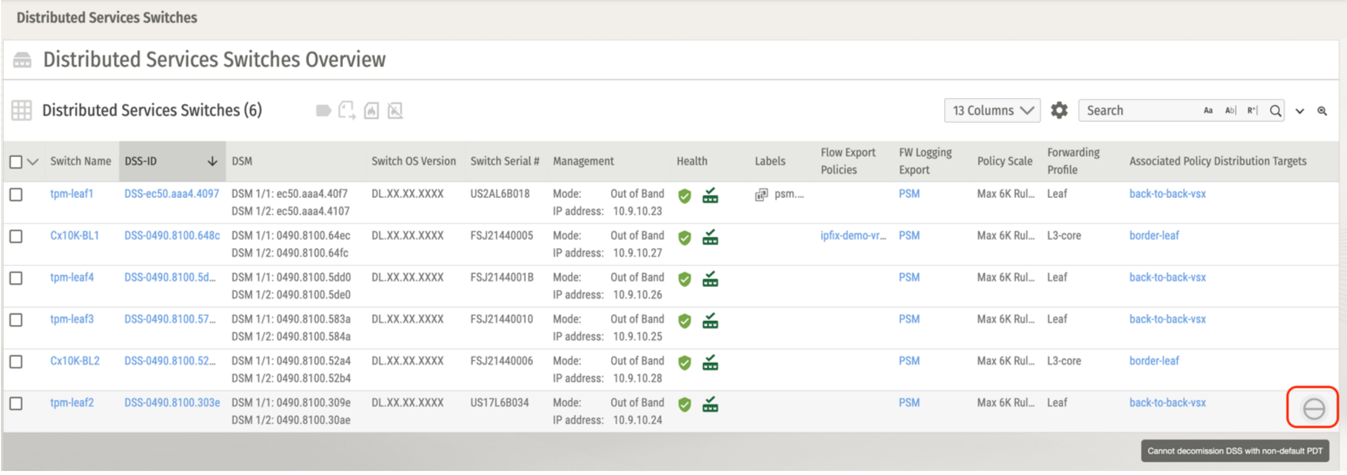
Note: *Once a DSS is decommissioned, all stateful policies on the DSS will be purged and there will be no stateful inspection/services performed until it is readmitted to a PSM.*

This procedure can be used to move a DSS from one PSM cluster to another, by first removing the DSS from the PSM, and then removing the configuration from AOS-CX.

Notes:

1. *If the DSS is in a non-default PDT , it must be moved to the default PDT before decommission.*
2. *To avoid traffic loss, all active ports (uplink / host-facing link / ISL link) needs to be shut before decommission.*

- On the PSM, Navigate to `System → DSS` .
- For the DSS to be decommissioned, move the cursor over to the right end of the row corresponding to the DSS. A mouseover option will appear that reads “Decommission DSS”:



Distributed Services Switches

Distributed Services Switches Overview

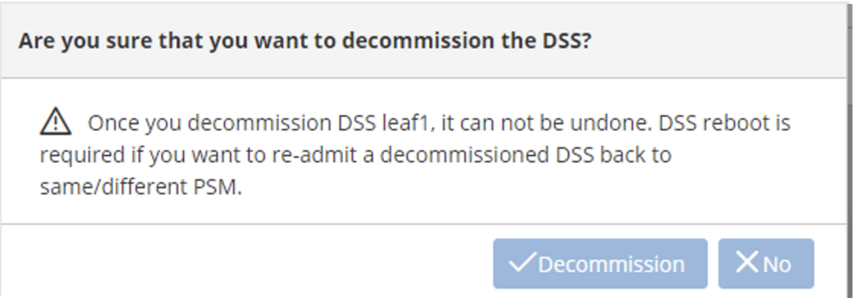
Distributed Services Switches (6)

□	Switch Name	DSS-ID	DSM	Switch OS Version	Switch Serial #	Management	Health	Labels	Flow Export Policies	FW Logging Export	Policy Scale	Forwarding Profile	Associated Policy Distribution Targets
□	tpm-leaf1	DSS-ec50.aaa4.4097	DSM 1/1: ec50.aaa4.40f7 DSM 1/2: ec50.aaa4.4107	DL.XX.XX.XXXX	US2AL6B018	Mode: Out of Band IP address: 10.9.10.23	✓	psm....		PSM	Max 6K Rul...	Leaf	back-to-back-vsx
□	Cx10K-BL1	DSS-0490.8100.648c	DSM 1/1: 0490.8100.64ec DSM 1/2: 0490.8100.64fc	DL.XX.XX.XXXX	FSJ21440005	Mode: Out of Band IP address: 10.9.10.27	✓		ipfix-demo-vr...	PSM	Max 6K Rul...	L3-core	border-leaf
□	tpm-leaf4	DSS-0490.8100.5d...	DSM 1/1: 0490.8100.5dd0 DSM 1/2: 0490.8100.5de0	DL.XX.XX.XXXX	FSJ2144001B	Mode: Out of Band IP address: 10.9.10.26	✓			PSM	Max 6K Rul...	Leaf	back-to-back-vsx
□	tpm-leaf3	DSS-0490.8100.57...	DSM 1/1: 0490.8100.583a DSM 1/2: 0490.8100.584a	DL.XX.XX.XXXX	FSJ21440010	Mode: Out of Band IP address: 10.9.10.25	✓			PSM	Max 6K Rul...	Leaf	back-to-back-vsx
□	Cx10K-BL2	DSS-0490.8100.52...	DSM 1/1: 0490.8100.52a4 DSM 1/2: 0490.8100.52b4	DL.XX.XX.XXXX	FSJ21440006	Mode: Out of Band IP address: 10.9.10.28	✓			PSM	Max 6K Rul...	L3-core	border-leaf
□	tpm-leaf2	DSS-0490.8100.303e	DSM 1/1: 0490.8100.309e DSM 1/2: 0490.8100.30ae	DL.XX.XX.XXXX	US17L6B034	Mode: Out of Band IP address: 10.9.10.24	✓			PSM	Max 6K Rul...	Leaf	back-to-back-vsx

Cannot decomission DSS with non-default PDT

Figure 29. The Decommission icon

- Click on this icon, and confirm that you want to decommission the DSS:

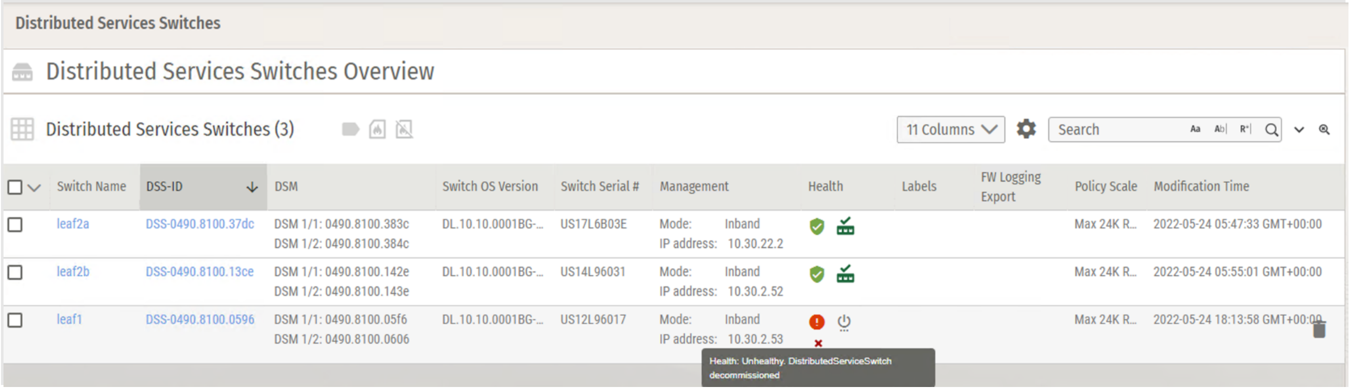


Are you sure that you want to decommission the DSS?

⚠ Once you decommission DSS leaf1, it can not be undone. DSS reboot is required if you want to re-admit a decommissioned DSS back to same/different PSM.

Figure 30. Confirm decommissioning; stateful services will no longer be active on this switch until it is readmitted to a PSM.

- Once the DSS is decommissioned, the Health icons will be updated as shown in Figure 31:



Distributed Services Switches

Distributed Services Switches Overview

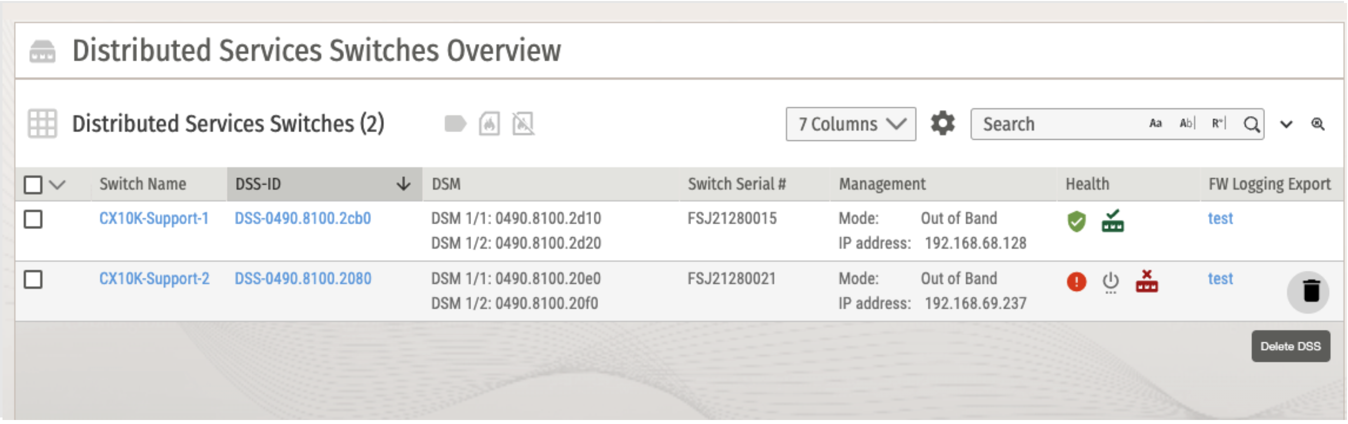
Distributed Services Switches (3)

□	Switch Name	DSS-ID	DSM	Switch OS Version	Switch Serial #	Management	Health	Labels	FW Logging Export	Policy Scale	Modification Time
□	leaf2a	DSS-0490.8100.37dc	DSM 1/1: 0490.8100.383c DSM 1/2: 0490.8100.384c	DL.10.10.0001BG-...	US17L6B03E	Mode: Inband IP address: 10.30.22.2	✓			Max 24K R...	2022-05-24 05:47:33 GMT+00:00
□	leaf2b	DSS-0490.8100.13ce	DSM 1/1: 0490.8100.142e DSM 1/2: 0490.8100.143e	DL.10.10.0001BG-...	US14L96031	Mode: Inband IP address: 10.30.2.52	✓			Max 24K R...	2022-05-24 05:55:01 GMT+00:00
□	leaf1	DSS-0490.8100.0596	DSM 1/1: 0490.8100.05f6 DSM 1/2: 0490.8100.0606	DL.10.10.0001BG-...	US12L96017	Mode: Inband IP address: 10.30.2.53	! ✕			Max 24K R...	2022-05-24 18:13:58 GMT+00:00

Health: Unhealthy. DistributedServiceSwitch decommissioned

Figure 31. Health icons show that the DSS is no longer admitted.

- Click on the Delete DSS (🗑) icon to delete the DSS object:



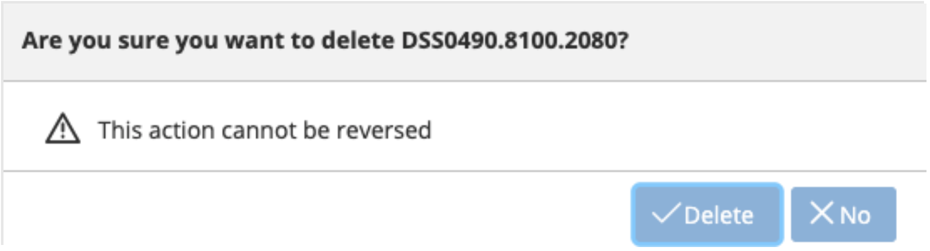
Distributed Services Switches Overview

Distributed Services Switches (2)

☐	Switch Name	DSS-ID	DSM	Switch Serial #	Management	Health	FW Logging Export
☐	CX10K-Support-1	DSS-0490.8100.2cb0	DSM 1/1: 0490.8100.2d10 DSM 1/2: 0490.8100.2d20	FSJ21280015	Mode: Out of Band IP address: 192.168.68.128	✓	test
☐	CX10K-Support-2	DSS-0490.8100.2080	DSM 1/1: 0490.8100.20e0 DSM 1/2: 0490.8100.20f0	FSJ21280021	Mode: Out of Band IP address: 192.168.69.237	!	test

Figure 32. The Delete option becomes available when an entry is hovered over.

- Confirm that you want to delete the DSS object:

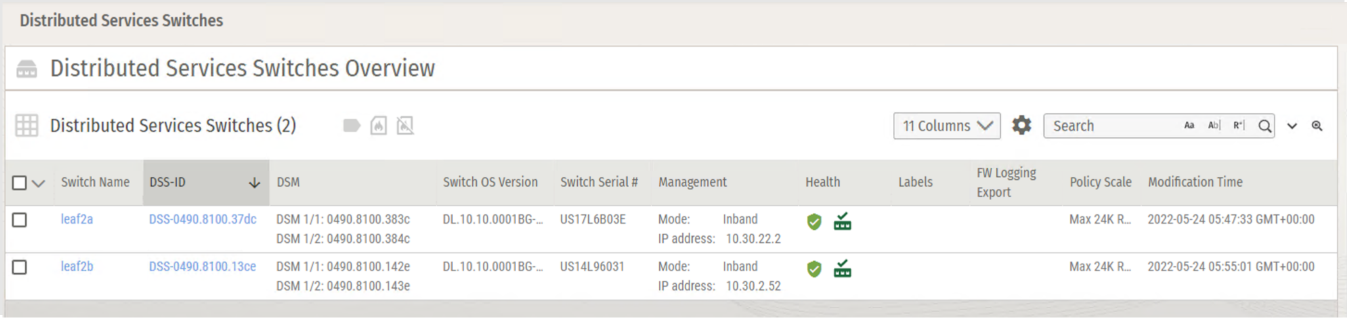


Are you sure you want to delete DSS0490.8100.2080?

⚠ This action cannot be reversed

Figure 33. Confirm removing the DSS from this PSM.

Once this is done, the DSS will be removed from the PSM cluster. Once you have confirmed the DSS does not show up in the PSM GUI, proceed to the next step.



Distributed Services Switches

Distributed Services Switches Overview

Distributed Services Switches (2)

☐	Switch Name	DSS-ID	DSM	Switch OS Version	Switch Serial #	Management	Health	Labels	FW Logging Export	Policy Scale	Modification Time
☐	leaf2a	DSS-0490.8100.37dc	DSM 1/1: 0490.8100.383c DSM 1/2: 0490.8100.384c	DL.10.10.0001BG-...	US17L6B03E	Mode: Inband IP address: 10.30.22.2	✓			Max 24K R...	2022-05-24 05:47:33 GMT+00:00
☐	leaf2b	DSS-0490.8100.13ce	DSM 1/1: 0490.8100.142e DSM 1/2: 0490.8100.143e	DL.10.10.0001BG-...	US14L96031	Mode: Inband IP address: 10.30.2.52	✓			Max 24K R...	2022-05-24 05:55:01 GMT+00:00

Figure 34. The removed DSS is no longer listed.

- After removing the DSS from PSM, log in to AOS-CX and remove the PSM configuration. (If you want to switch the DSS to a new PSM, you can also enter the new PSM addresses at this point.) Save the configuration and reload the switch.

```
switch# show running-config psm
psm
    host 20.3.0.37 20.3.0.38 20.3.0.39 vrf default

switch (config)# no psm
Warning: Once the PSM is removed, a reboot is required to
reconnect to any PSM.

Continue (y/n)? Y

switch (config)# show running-config psm
switch (config)#

switch (config)# copy running-config startup-config
Copying configuration: [Success]

switch# boot system primary
```

Note: *If the switch being decommissioned is also being managed as part of an AFC fabric, it may also need to be removed from AFC management.*

Associating the PSM to Aruba Fabric Composer

Aruba Fabric Composer (AFC) automates CX switch provisioning and management, including coordination with stateful services via the PSM. It can be used to create and provision leaf-spine networks, simplify troubleshooting, and perform life cycle management.

To synchronize these capabilities with the PSM's stateful services, AFC must be configured to be associated with the PSM by providing AFC with the PSM instance's network address. Refer to the AFC documentation for details on this process.

Firewall Policy Functionality and Configuration

The CX 10000 extends the capabilities of the leaf-spine fabric to natively provide a distributed stateful firewall for east-west traffic, zero trust segmentation, and pervasive telemetry. These built-in security services help minimize attack surface within the data center and block lateral movement in attack attempts.

Functionality supported by the stateful L4 firewall includes:

- Connection tracking:
    - Initial TCP handshake validation
    - Session closing sequence validation
    - TCP reset checks
- Application-level gateways (TFTP, DNS, FTP, MSRPC, SUNRPC, RTSP)
    - The DNS ALG, for example, tracks and correlates DNS queries with corresponding responses; once a response is received and forwarded, the ALG will accelerate aging/closing of the DNS UDP session, resulting in better DNS session scale.
    - Other ALGs follow standard protocol implementation.
- Firewall Log Export to the PSM
- Firewall Log Export to external collectors, in the default VRF context.

Network Security Policies are created, then attached to either networks, VRFs or both. When a policy is attached to a VRF it is internally applied to all networks in that VRF. *Ingress* security policies are applied to traffic entering a host and *egress* security policies are applied to traffic leaving a host. The same policy can be used for both ingress and egress, or different policies for ingress and egress can be used for more granular control.

Policy direction is from a workload/host perspective.

Considerations

- All traffic leaving the workload/host and entering the switch in host-facing ports (“access” persona) is subject to the egress policy on the switch.
- All traffic destined for the workload/host, entering the switch from the fabric side (“uplink” persona) ports, is subject to ingress policy on the switch.
- The `persona` CLI config option on a port is used to set the appropriate policy (ingress/egress) on the DSM.
- By default, all 10G/25G ports are configured as workload/host facing ports (“access” persona) and the 100G/40G ports are configured as fabric ports (“uplink” persona).

    - If a 40G/100G port needs to be used as a workload/host facing port, then the “access” persona needs to be explicitly configured on it.
    - If a 10G/25G port needs to be used as a fabric port, then the “uplink” persona needs to be explicitly configured on it.
    - Inter-switch links (ISLs) should be configured to “No Persona”
- For locally-switched flows on the switch, the traffic from the host is subject to policy processing only once and only egress policy is enforced.
- For locally-routed flows on the switch, the traffic from the host is subject to egress policy in source VLAN (pre-routing) first and then subject to ingress policy in destination VLAN (post-routing). In this case, the firewall log is generated for the ingress policy applied on the post-routed destination VLAN. No firewall log will be generated for the egress policy applied on the source VLAN.
- For PVLAN flows that are locally switched on the switch, the traffic from the host is subject to egress policy in the source VLAN (which is the secondary PVLAN) first and then subject to ingress policy in destination VLAN (which is the primary PVLAN). A firewall log record will be generated for the ingress policy applied to the destination VLAN.
- Users can choose to export Firewall Log records, which are generated for flows subject to policy evaluation.
    - If PSM is chosen as an export target for Firewall Log records, the supported sustained ingestion rate is 10K logs per second (LPS), with a burst rate of 20K LPS.

- If a security policy attached to a network has no rules, then there is an implicit default deny rule to drop all the traffic. In this scenario, there are no rule statistics available for denied flows. Flow logs generated in this scenario will not contain the policy and rule information.
- If a VLAN or VNI belonging to a given VRF needs a security policy, users *must* define network objects on the PSM for all VLANs in that VRF.
- If rules are edited within an existing in-use Firewall Policy, all existing rule statistics within the Firewall Policy will be reset.
- If a Firewall Policy is updated, existing active flows will be re-evaluated and the new updated policy will be enforced. The new policy will be enforced only after the initiator of the flow sends traffic or at least 1 packet. This dynamic flow refresh feature is not supported for data-flows of ALG type of traffic including TFTP, RTSP,SUNRPC and MSRPC ALG. The policy for these types of active ALG data flows will not be updated dynamically. For FTP, the policy for active data flows will be updated only after there is a control packet exchange from the client.
- With flow log export, there are no periodic updates for long-lived flows.

- For Leaf deployments, if route leaking is enabled on the DSS and if the VRFs involved in the route leaking are pinned to different DSMs, then stateful services on traffic between VRFs is not supported. It is recommended to use the VRF pinning CLI for statically pinning both the VRFs to the same DSM instance.
- When a DSS is acting as an L2 access switch, having different actions (PERMIT/DENY) for the same flow across ingress and egress policies on different VLANs is not permitted.
- Broadcast (ARP), L2 multicast, IP multicast and IPv6 traffic are not redirected to the DSE and therefore are not subject to policy evaluation.
- Firewall policy enforcement on VXLAN centralized gateways is not supported; this is only supported on VXLAN distributed anycast gateway designs.
- On non-AFC controlled setups, for a given VRF, the same VRF name *must* be used on the PSM and the DSS.
- For Edge Gateway deployment, enforcement is only supported at the VRF level.
- Bidirectional Forwarding Detection (BFD) echo packets via “persona access” ports are redirected to DSM, hence BFD should be permitted in the policy configured.
- BFD echo packets via “persona uplink” ports require “uplink-uplink” configuration under the dsm mode for correct functionality.

VSX and Firewall High Availability

Aruba VSX can be used to pair switches, increasing firewall availability by preventing application connections from being dropped when a switch fails or is upgraded.

Note**: In order for VSX firewall high availability to function effectively,* `linkup-delay-timer` *and* `inter-switch-link peer-detect-interval` ***must be configured.

- Configure the required `linkup-delay-timer` and `inter-switch-link peer-detect-interval` to 600 seconds under VSX in AOS-CX

```
switch# conf t
switch(config)# vsx
switch(config-vsx)# linkup-delay-timer 600
switch(config-vsx)# inter-switch-link peer-detect-
interval 600
```

- Configure OSPF `max-metric` using the following commands:

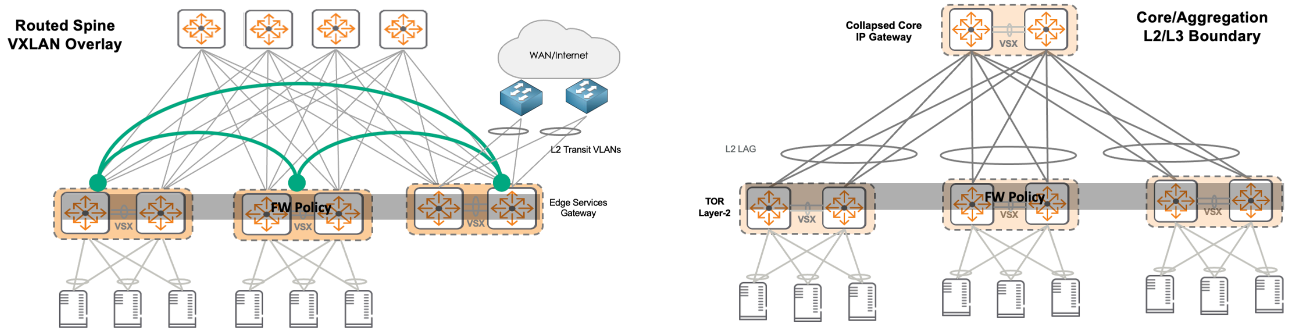
  ```
  switch# conf t
  switch (config)# router ospf <>
  switch (config-ospf-1)# max-metric router-lsa include-stub
  on-startup
  ```

  This command advertises a high metric for 600s (default) post reload.

- When adding ports as members of an ISL LAG for VSX, make sure to configure `no persona` for each one.
- Connection tracking with VSX is fully supported.
- Using ALG protocols in security policies with VSX deployments is fully supported.
- Logging considerations:
  - Only the VSX node that sees the first packet in a flow will increment rule statistics.
  - With Firewall Log export, an open record will be generated by either the primary or secondary switch that receives the first flow miss, and a “close” record will be generated by both primary or secondary switches.
- The maximum session limit on VRF/network will apply to both local learn flows and to flows that get flow-synced from the VSX peer.

Supported Topologies

CX 10000 firewall can be supported in standard switching deployment architectures, including traditional 3-tier architecture with VSX (core, aggregation, and access), and Clos-based spine-leaf fabrics with VXLAN overlays. Starting with the 10.12 release, CX 10000 can be deployed as an Edge Gateway to provide NAT functionality. With the 10.12.1000 release, IPsec can also be enabled on the Edge Gateway.



CX 10000 Leaf
- Virtual Switching Extension (VSX) for HA
- L2/L3 Boundary with Distributed Anycast Gateway
- Support for single and dual attached servers
- Active/Active 1.6T Stateful FW with VSX

CX 10000 Access
- Virtual Switching Extension (VSX) for HA
- Support for single and dual attached servers
- Active/Active 1.6T Stateful FW with VSX

Figure 35. Deployment architectures

Connection Tracking

TCP connection tracking is always enabled on the CX 10000 platform; there is no need to explicitly enable this on PSM as part of the security profile. This feature is supported on both the standalone single-switch CX 10000 as well as the VSX pair. TCP connection tracking is a key feature as part of the stateful firewall implementation, providing detailed per-flow validation of TCP control packets during the initial 3-way handshake sequence (SYN, SYN-ACK, ACK) and also the closing sequence involving the FIN/RST packets.

As part of connection-tracking, invalid TCP control packet sequences, invalid TCP packets with incorrect flag combinations, incorrect sequence numbers in TCP control packets are detected and dropped providing a more robust and security stateful firewall.

Connection tracking on VSX provides at high level several categories of checks:

- Checks based on existing state of flow and direction of flow (i-flow/r-flow)
- Checks for invalid TCP packets based on current state of flow
- Checks for invalid sequences in each direction of flow
- Sequence number based checks for different types of TCP packets, including SYN/ACK/RST/FIN during the initial 3-way handshake and session closures due to FIN/RST

Packets are dropped for these invalid conditions and will not result in changes of the per-flow state machine.

Firewall Policy Configuration

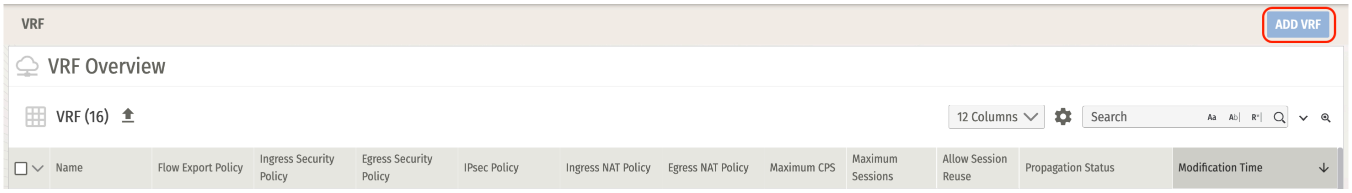
To enable stateful L4 firewalling, follow these steps:

- Create a VRF on the PSM
- Create a Security Policy on the PSM
- Attach the Policy to a Network on the PSM
- Create VRF and VLAN on AOS-CX

  The VRF and VLAN must also be defined on the switch using AOS-CX, either before or after configuring security policy for those networks in PSM. If this is not done, traffic will not be redirected within the switch to the DSM, and there will be no stateful security evaluation of flows for the specific VRF or VLAN/network.

Create a VRF

On the PSM, navigate to `Tenants` → `VRF` and click ADD VRF.



VRF

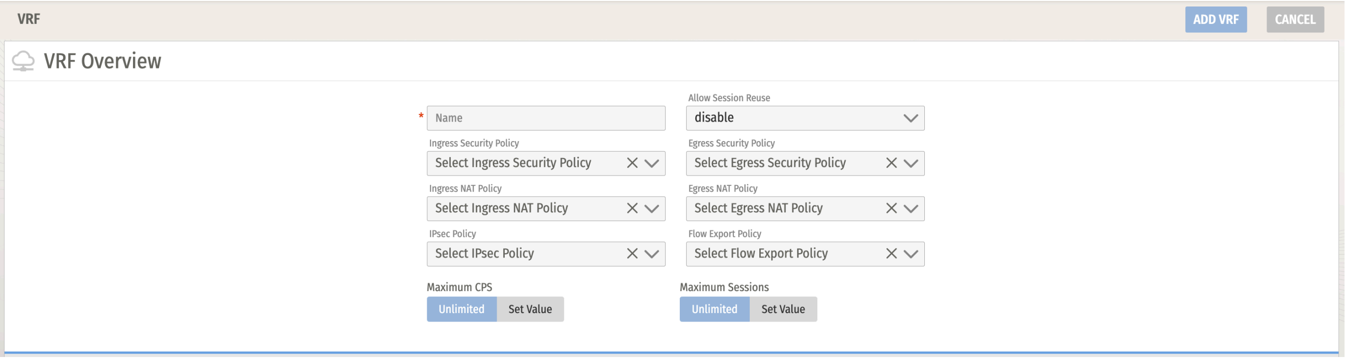
VRF Overview

VRF (16)

□	Name	Flow Export Policy	Ingress Security Policy	Egress Security Policy	IPsec Policy	Ingress NAT Policy	Egress NAT Policy	Maximum CPS	Maximum Sessions	Allow Session Reuse	Propagation Status	Modification Time ↓

Figure 36. PSM VRF screen detail showing ADD VRF button

Provide a name for this VRF. To save, click on ADD VRF on this screen as well.



VRF

ADD VRF CANCEL

VRF Overview

* Name	Allow Session Reuse: disable
Ingress Security Policy: Select Ingress Security Policy	Egress Security Policy: Select Egress Security Policy
Ingress NAT Policy: Select Ingress NAT Policy	Egress NAT Policy: Select Egress NAT Policy
IPsec Policy: Select IPsec Policy	Flow Export Policy: Select Flow Export Policy
Maximum CPS: Unlimited / Set Value	Maximum Sessions: Unlimited / Set Value

Figure 37. VRF creation screen

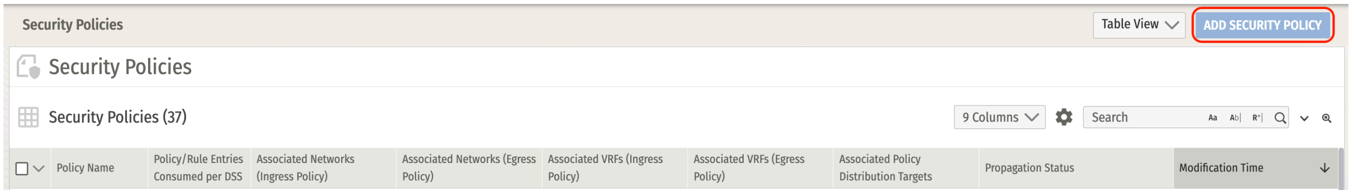
***Note**:The VRF names in AOS-CX and the PSM* must *match.*

Create a Network Security Policy

There are two ways to create a network security policy: manually entering a predetermined policy, or by using the PSM UI network graph feature to help construct policies.

If the intended policy has already been determined, use the *table view* within the UI to define it. (An API call is available as well.) The beta feature *network graph* feature can be used instead, which aids in discovering what flows should be allowed and designing a zero trust network policy to only allow those flows. The network graph method is covered in detail in Appendix E.

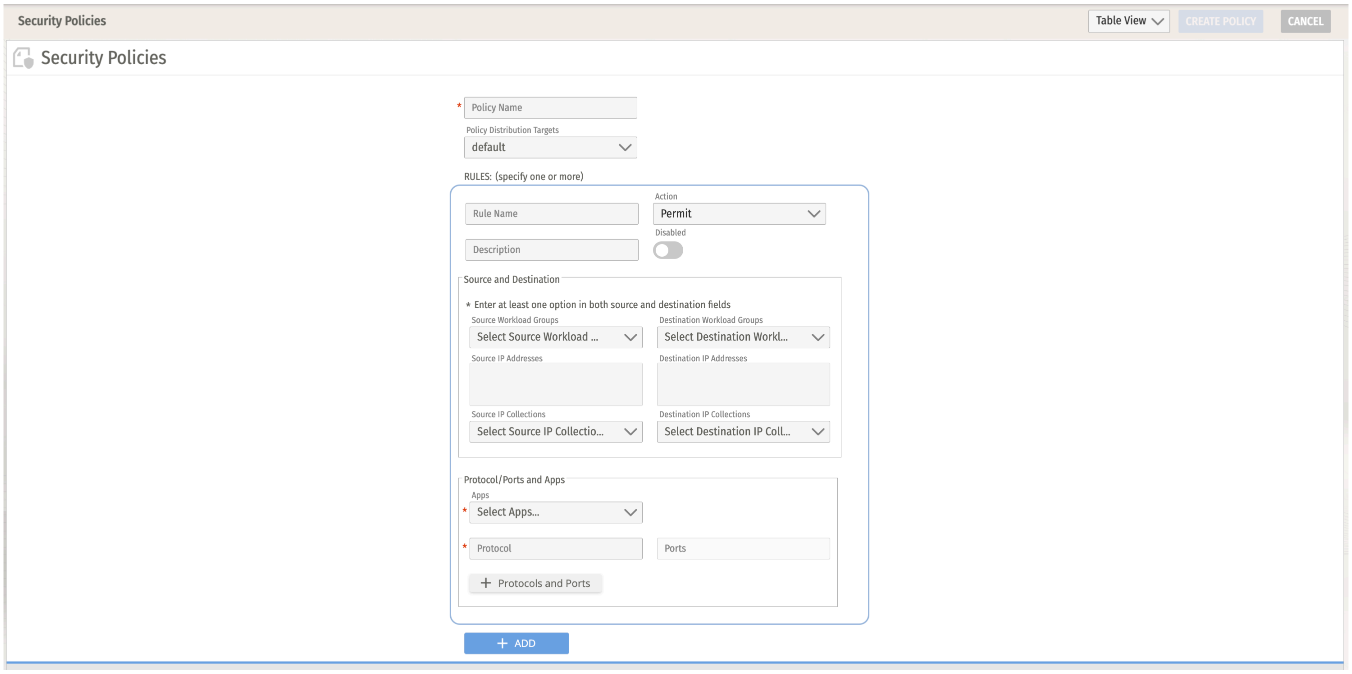
To use table view, navigate to `Tenants → Security Policy` and click on ADD SECURITY POLICY :



□	Policy Name	Policy/Rule Entries Consumed per DSS	Associated Networks (Ingress Policy)	Associated Networks (Egress Policy)	Associated VRFs (Ingress Policy)	Associated VRFs (Egress Policy)	Associated Policy Distribution Targets	Propagation Status	Modification Time

Figure 38. PSM Security Policy screen detail showing ADD SECURITY POLICY button

Enter the relevant information and click CREATE POLICY to save the new policy.



Policy Distribution Targets
default

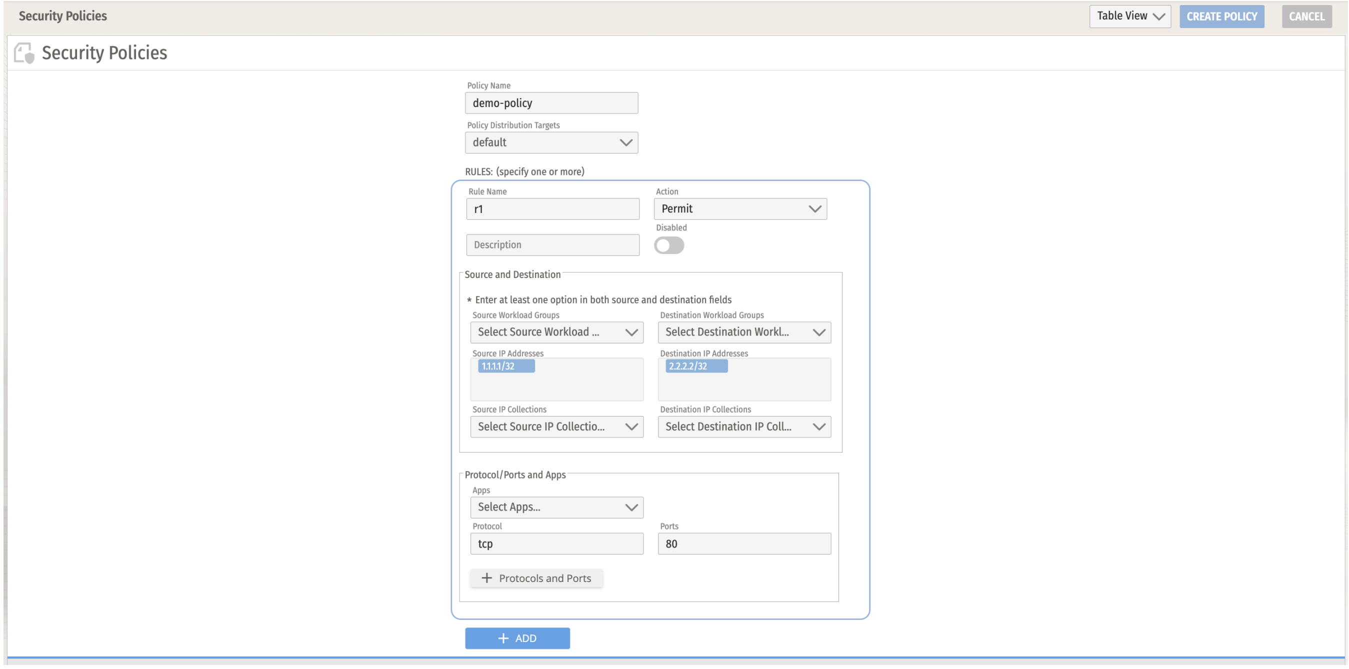
Action
Permit

Figure 39. Creating a new security policy

Perform the following steps to create a security policy:

- Policy Name: Define a unique name for the policy
- Policy Distribution Targets: If applicable for the DSS, select the user-defined PDT this switch is already part of, or the default PDT
- Rule Name: Unique rule name for this specific rule
- Action: Depending on desired policy action or behavior for this rule, select Permit or Deny
  - Note, because of dynamic flow refresh when a rule is changed from permit to deny, the session is immediately aged out in the dataplane.

  - This doesn’t impact any other flows / connections active at this time that were acctive, just flows that match the rule whose action was changed to deny
- Description: If needed, apply an additional description for context about this rule
- Disabled: Default (off), which means rule will be activated once the policy is created and attached to a network or VRF. If disabled is (on) then the rule will not be active once policy is created
- Source / Destination IP Addresses: IP addresses or subnets, in CIDR format for IP based policy rules
- Source/ Destination IP Collections: IP collections allow the creation of groups in policy, for IP subnets in CIDR form making it easier to apply in policy and reference in multiple policies. See section on IP collections for further details.
- Proto-Ports: Specific protocols (udp/tcp/icmp/gre/esp/ah) or protocol reference numbers (0 - 254), along with ports (comma separated or dash for range) to use in policy rules.
- Apps: Used to reference apps defined in PSM for well-known applications instead of protocols and ports in policy rules. See section on Apps definition for further details.



Policy Name: demo-policy

Policy Distribution Targets: default

RULES: (specify one or more)

Rule Name: r1

Action: Permit

Source and Destination

* Enter at least one option in both source and destination fields

Source Workload Groups
Select Source Workload ...
Destination Workload Groups
Select Destination Workl...

Source IP Addresses: 1.1.1.1/32

Destination IP Addresses: 2.2.2.2/32

Source IP Collections
Select Source IP Collectio...
Destination IP Collections
Select Destination IP Coll...
Protocol/Ports and Apps
Apps
Select Apps...

Protocol: tcp

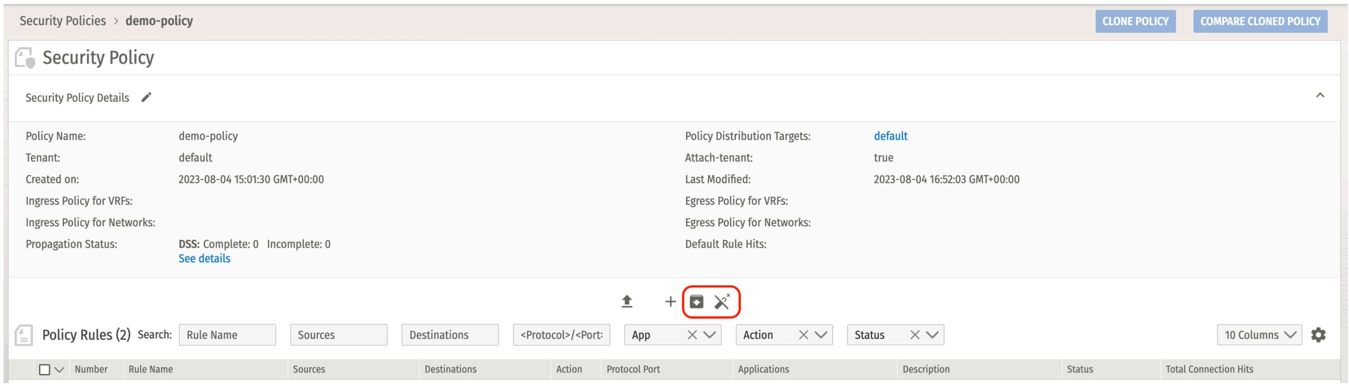
Ports: 80

Figure 40. Security Policy creation screen

The Network Graph UI feature helps discover the relationships between endpoints within a VRF, within and between VLANs, and within or between endpoint groups. The relationships help better understand common connection occurrences within the data center that should be permitted. See Appendix E for a description of two methods for determining policies.

Rule Overlap Detection

During policy creation, if there are rules in a policy that overlap, the PSM will flag them to prevent a possible misconfiguration when configured to detect, as shown in Figures 41 and 41a.

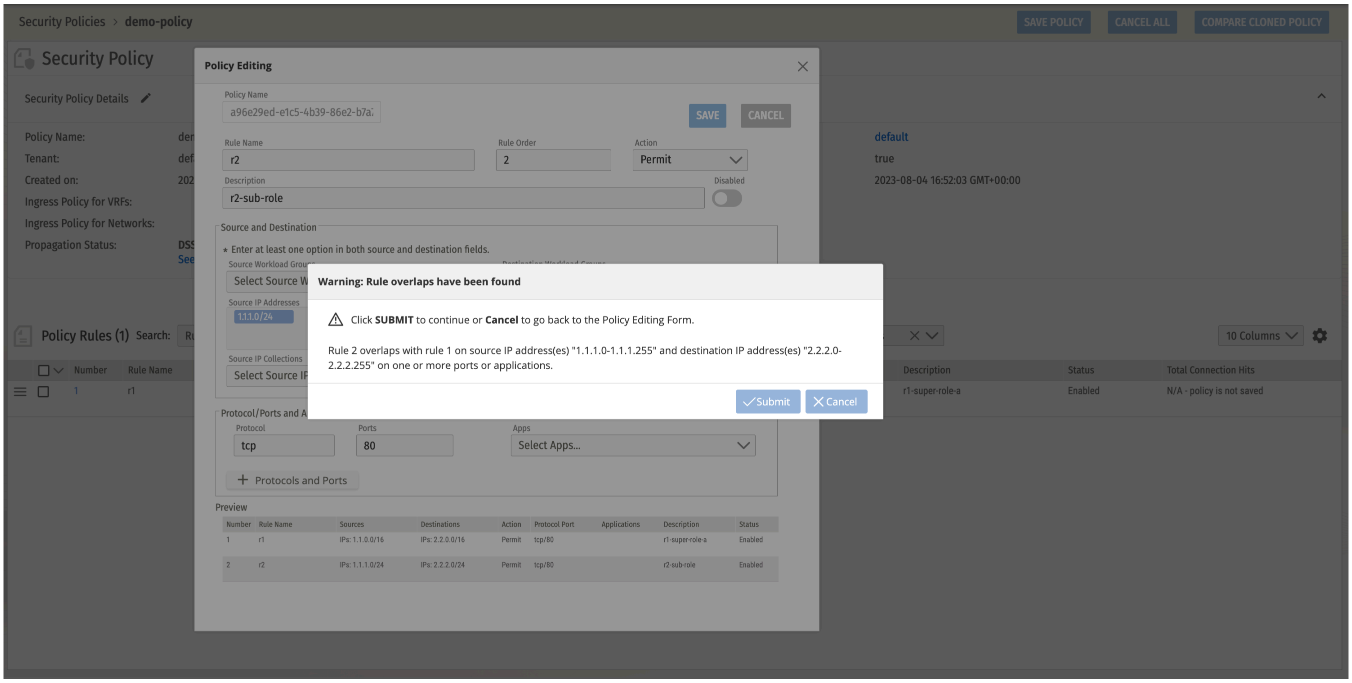


Policy Name: demo-policy
Tenant: default
Created on: 2023-08-04 15:01:30 GMT+00:00
Ingress Policy for VRFs:
Ingress Policy for Networks:
Propagation Status: DSS: Complete: 0 Incomplete: 0
See details

Policy Distribution Targets: default
Attach-tenant: true
Last Modified: 2023-08-04 16:52:03 GMT+00:00
Egress Policy for VRFs:
Egress Policy for Networks:
Default Rule Hits:

Figure 41. Enable overlapping rule detection

In the following example, the user has created 2 rules “r1” and “r2”. Rules “r1” and ”r2” are overlapping rules with “r1” matching HTTP traffic from “1.1/16” to “2.2/16”, whereas “r2” matches HTTP traffic from a smaller subnet “1.1.1/24” to “2.2.2/24”. In this config, the rule “r2” will never get matched and hence the PSM is flagging this as an overlapping rule:



Policy Name: a96e29ed-e1c5-4b39-86e2-b7a

Rule Name: r2

Rule Order: 2

Action: Permit

Description: r2-sub-role

Disabled

* Enter at least one option in both source and destination fields.

Protocol: tcp

Ports: 80

Apps: Select Apps...

Warning: Rule overlaps have been found

⚠ Click **SUBMIT** to continue or **Cancel** to go back to the Policy Editing Form.

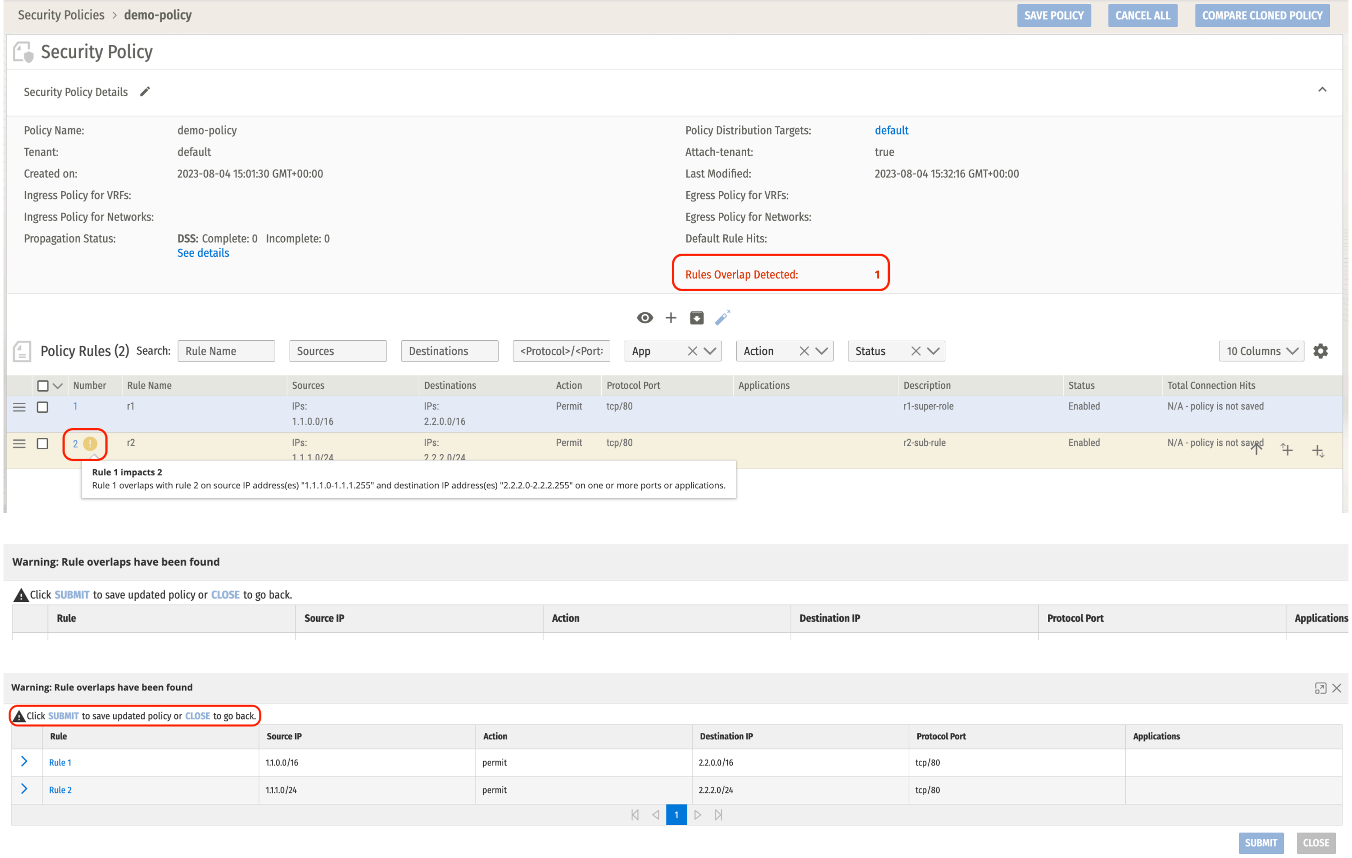
Rule 2 overlaps with rule 1 on source IP address(es) "1.1.1.0-1.1.1.255" and destination IP address(es) "2.2.2.0-2.2.2.255" on one or more ports or applications.

Preview

Number	Rule Name	Sources	Destinations	Action	Protocol Port	Applications	Description	Status
1	r1	IPs: 1.1.0.0/16	IPs: 2.2.0.0/16	Permit	tcp/80		r1-super-role-a	Enabled
2	r2	IPs: 1.1.1.0/24	IPs: 2.2.2.0/24	Permit	tcp/80		r2-sub-role	Enabled

Figure 41a. Rule overlap warning

Overlapping rule detection is only a warning. You can still proceed and save a policy with overlapping rules. If the rule detection function wasn’t enabled when creating a new rule, it can be enabled to scan the policy for detecting rule overlaps as shown in the following example:



Security Policy

Security Policy Details

Policy Name:	demo-policy	Policy Distribution Targets:	default
Tenant:	default	Attach-tenant:	true
Created on:	2023-08-04 15:01:30 GMT+00:00	Last Modified:	2023-08-04 15:32:16 GMT+00:00
Ingress Policy for VRFs:		Egress Policy for VRFs:	
Ingress Policy for Networks:		Egress Policy for Networks:	
Propagation Status:	DSS: Complete: 0 Incomplete: 0 See details	Default Rule Hits:	
		Rules Overlap Detected:	1

Number	Rule Name	Sources	Destinations	Action	Protocol Port	Applications	Description	Status	Total Connection Hits
1	r1	IPs: 1.1.0.0/16	IPs: 2.2.0.0/16	Permit	tcp/80		r1-super-role	Enabled	N/A - policy is not saved
2	r2	IPs: 1.1.1.0/24	IPs: 2.2.2.0/24	Permit	tcp/80		r2-sub-rule	Enabled	N/A - policy is not saved

Rule 1 impacts 2
Rule 1 overlaps with rule 2 on source IP address(es) "1.1.1.0-1.1.1.255" and destination IP address(es) "2.2.2.0-2.2.2.255" on one or more ports or applications.

Warning: Rule overlaps have been found

Click SUBMIT to save updated policy or CLOSE to go back.

Rule	Source IP	Action	Destination IP	Protocol Port	Applications

Warning: Rule overlaps have been found

Click SUBMIT to save updated policy or CLOSE to go back.

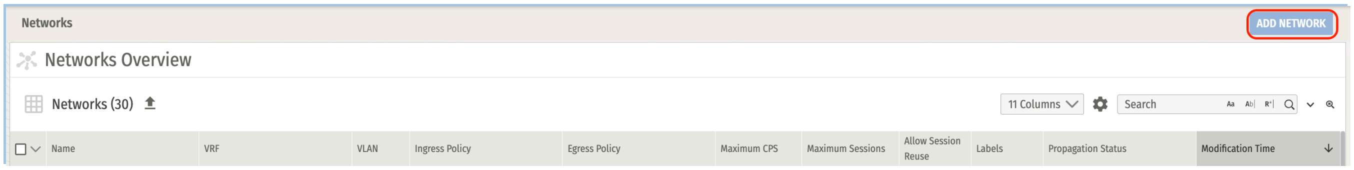
Rule	Source IP	Action	Destination IP	Protocol Port	Applications
Rule 1	1.1.0.0/16	permit	2.2.0.0/16	tcp/80	
Rule 2	1.1.1.0/24	permit	2.2.2.0/24	tcp/80	

Figure 41b. Rule overlap warning

Note: *It is not recommended to use overlapping rule detection for policies with large rule scale.*

Attach Policy to a Network

Navigate to `Tenants → Networks` and click on ADD NETWORK.

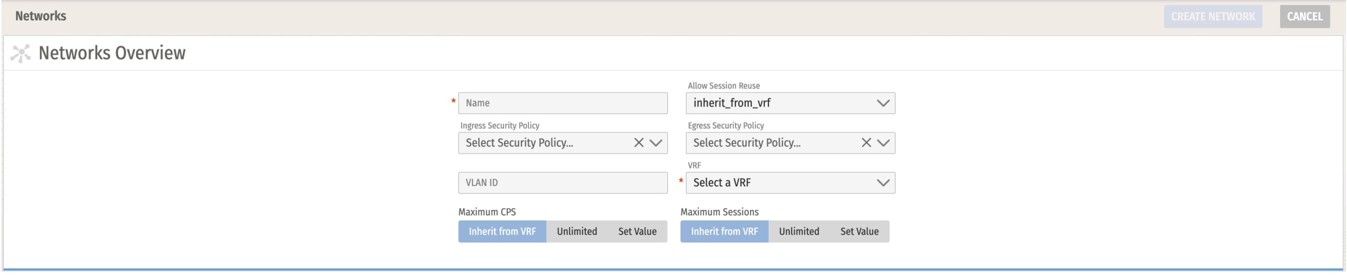


Name	VRF	VLAN	Ingress Policy	Egress Policy	Maximum CPS	Maximum Sessions	Allow Session Reuse	Labels	Propagation Status	Modification Time

Figure 42. PSM Network screen detail showing ADD button

Enter the relevant information and click CREATE NETWORK.

- Select the VRF to be used.
- Select the ingress and egress security policies.
- Specify the VLAN where this policy is to be applied. This does not create the AOS-CX VLAN.



* Name

Allow Session Reuse: inherit_from_vrf

Ingress Security Policy: Select Security Policy...

Egress Security Policy: Select Security Policy...

VLAN ID

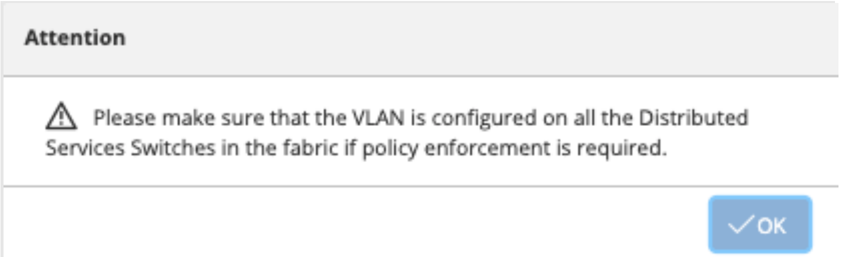
VRF: * Select a VRF

Maximum CPS: Inherit from VRF | Unlimited | Set Value

Maximum Sessions: Inherit from VRF | Unlimited | Set Value

Figure 43. Network creation screen

Once the network is created, you will be prompted to create the VLAN in AOS-CX. Creating the policy on the PSM does not automatically create the VLAN on AOS-CX.



Attention

⚠ Please make sure that the VLAN is configured on all the Distributed Services Switches in the fabric if policy enforcement is required.

Figure 44. Warning notice for required AOS-CX action

Create VRF on AOS-CX

If the VRF does not already exist in AOS-CX, create the VRF in AOS-CX using the AOS-CX CLI, AFC UI, or via the AOS-CX REST API. The example shown below uses the CLI.

```
switch# configure
switch (config)# vrf test
switch (config-vrf)# exit
switch (config)#
```

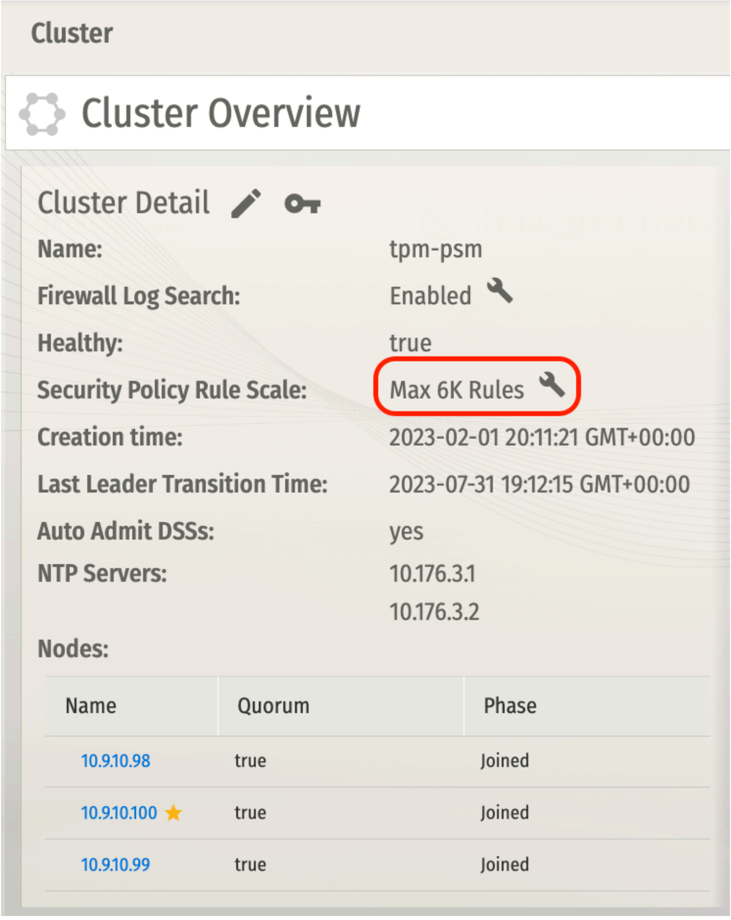
Create VLAN on AOS-CX

Create the VLAN via the AOS-CX CLI, AFC UI, or programmatically. The example shown below uses the CLI.

```
switch# configure
switch (config)# vlan 199
switch (config-vlan-199)# name demo-vrf
switch (config-vlan-199)# exit
switch (config)# exit
```

Understanding Firewall Policy Scaling Profiles

By default, the CX 10000 firewall can scale up to 6K rules per policy. This can be seen in the PSM UI; navigate to the cluster object and the policy scale is shown as in Figure 45:



Cluster

Cluster Overview

Cluster Detail

Name:	tpm-psm
Firewall Log Search:	Enabled
Healthy:	true
Security Policy Rule Scale:	Max 6K Rules
Creation time:	2023-02-01 20:11:21 GMT+00:00
Last Leader Transition Time:	2023-07-31 19:12:15 GMT+00:00
Auto Admit DSSs:	yes
NTP Servers:	10.176.3.1 10.176.3.2

Nodes:

Name	Quorum	Phase
10.9.10.98	true	Joined
10.9.10.100 ★	true	Joined
10.9.10.99	true	Joined

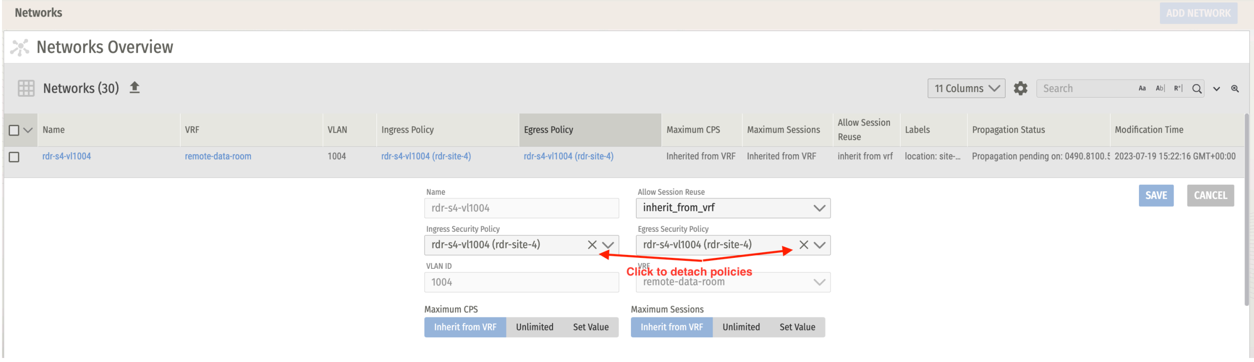
Figure 45. PSM cluster settings, including policy scale limit.

While this scale should satisfy the majority of enterprise site needs, it is not uncommon for some sites to require firewall policy scaling beyond the default. CX 10000 firewall offers the ability to increase policy scaling by switching to a different policy scaling profile, which will allow up to 24K rules per policy.

Switch Policy Scaling Profile

To change the profile, use the following workflow:

- Navigate to Tenant | Networks and click on the network name. Detach all the attached policies from the Network or VRF. If one or more policies remain attached, the policy profile change will fail.



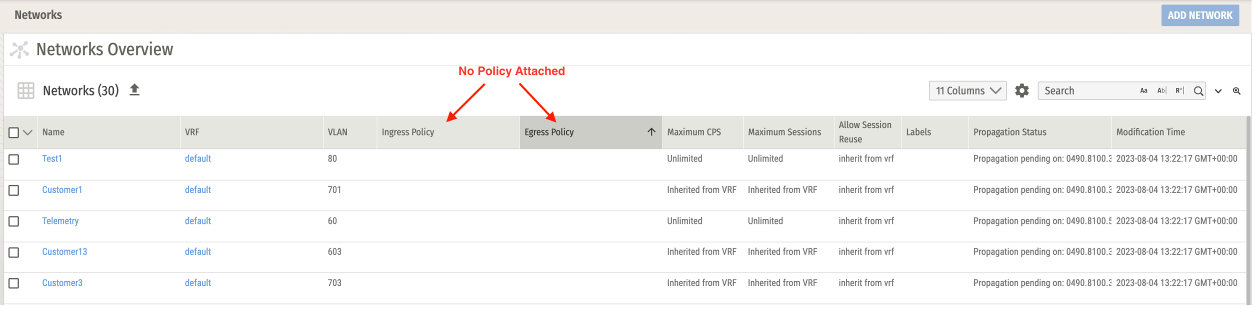
Name	VRF	VLAN	Ingress Policy	Egress Policy	Maximum CPS	Maximum Sessions	Allow Session Reuse	Labels	Propagation Status	Modification Time
rdr-s4-vl1004	remote-data-room	1004	rdr-s4-vl1004 (rdr-site-4)	rdr-s4-vl1004 (rdr-site-4)	Inherited from VRF	Inherited from VRF	inherit from vrf	location: site-...	Propagation pending on: 0490.8100.5	2023-07-19 15:22:16 GMT+00:00

Name: rdr-s4-vl1004
Allow Session Reuse: inherit_from_vrf
Ingress Security Policy: rdr-s4-vl1004 (rdr-site-4)
Egress Security Policy: rdr-s4-vl1004 (rdr-site-4)
VLAN ID: 1004
VRF: remote-data-room
Maximum CPS: Inherit from VRF | Unlimited | Set Value
Maximum Sessions: Inherit from VRF | Unlimited | Set Value

Click to detach policies

Figure 46. Detach policies

- There must be no attached policies from networks or VRF. Verify this by navigating to `Tenants | Networks` or `VRF` to confirm there is no policy attached.

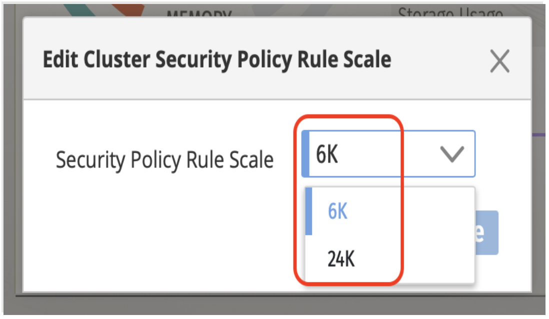


No Policy Attached

Name	VRF	VLAN	Ingress Policy	Egress Policy	Maximum CPS	Maximum Sessions	Allow Session Reuse	Labels	Propagation Status	Modification Time
Test1	default	80			Unlimited	Unlimited	inherit from vrf		Propagation pending on: 0490.8100.3	2023-08-04 13:22:17 GMT+00:00
Customer1	default	701			Inherited from VRF	Inherited from VRF	inherit from vrf		Propagation pending on: 0490.8100.3	2023-08-04 13:22:17 GMT+00:00
Telemetry	default	60			Unlimited	Unlimited	inherit from vrf		Propagation pending on: 0490.8100.5	2023-08-04 13:22:17 GMT+00:00
Customer13	default	603			Inherited from VRF	Inherited from VRF	inherit from vrf		Propagation pending on: 0490.8100.3	2023-08-04 13:22:17 GMT+00:00
Customer3	default	703			Inherited from VRF	Inherited from VRF	inherit from vrf		Propagation pending on: 0490.8100.3	2023-08-04 13:22:17 GMT+00:00

Figure 47. No policies are listed for Ingress or Egress.

- Switch the profile to "24K" under `System | Cluster` in the PSM UI:

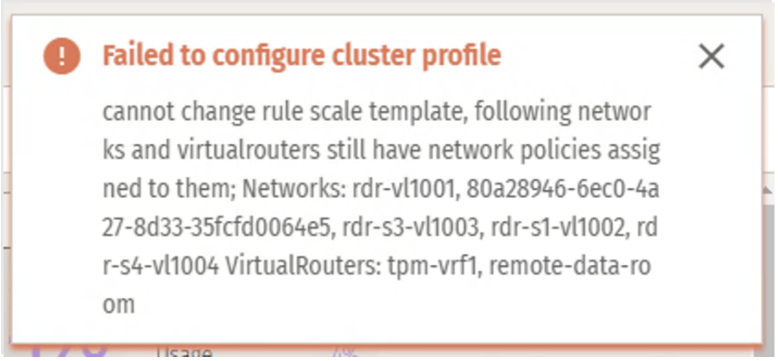


Edit Cluster Security Policy Rule Scale

Security Policy Rule Scale: 6K

Figure 48. Changing the policy scale limit

If any policies are attached to the VRF/network, the cluster-wide scale profile change is not allowed and a pop-up message indicating the problem will appear at the top right corner:

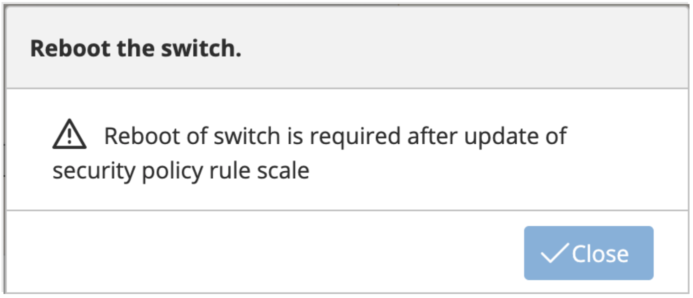


Failed to configure cluster profile ✕

cannot change rule scale template, following networks and virtualrouters still have network policies assigned to them; Networks: rdr-vl1001, 80a28946-6ec0-4a27-8d33-35fcfd0064e5, rdr-s3-vl1003, rdr-s1-vl1002, rdr-s4-vl1004 VirtualRouters: tpm-vrf1, remote-data-room

Figure 49. Failure notification due to attached policies, which must be detached in order to proceed.

- After the change is saved, a pop-up message will appear, noting that a reboot is required to activate the profile on each switch.

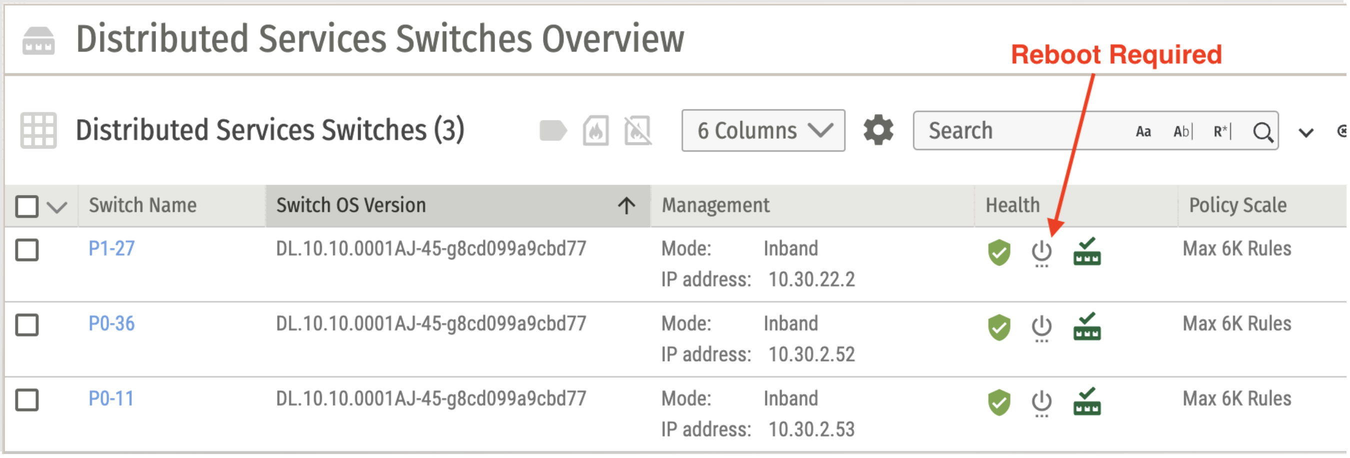


Reboot the switch.

⚠ Reboot of switch is required after update of security policy rule scale

Figure 50. Alert that rebooting the switch is needed

- Before the reboot, under `System | DSS`, "Distributed Services Switches Overview" will show the previous scale setting of "6k" and the "Health" column will indicate a reboot is required:



Distributed Services Switches Overview

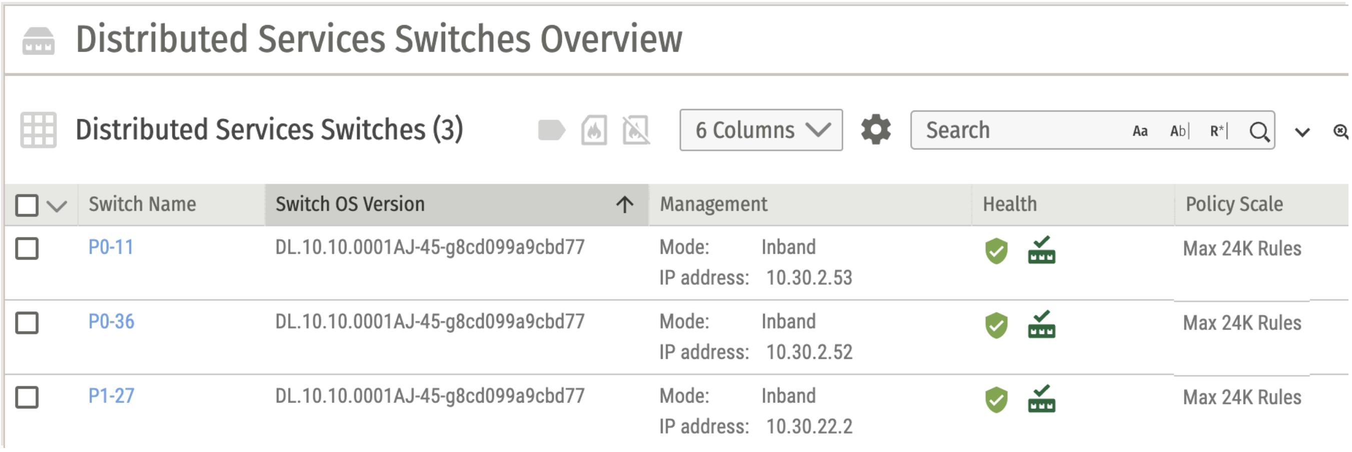
Reboot Required

Distributed Services Switches (3)

□	Switch Name	Switch OS Version ↑	Management	Health	Policy Scale
□	P1-27	DL.10.10.0001AJ-45-g8cd099a9cbd77	Mode: Inband IP address: 10.30.22.2	✓ ⏻	Max 6K Rules
□	P0-36	DL.10.10.0001AJ-45-g8cd099a9cbd77	Mode: Inband IP address: 10.30.2.52	✓ ⏻	Max 6K Rules
□	P0-11	DL.10.10.0001AJ-45-g8cd099a9cbd77	Mode: Inband IP address: 10.30.2.53	✓ ⏻	Max 6K Rules

Figure 51. The power icon's presence indicates which switches need rebooting.

- After the reboot, `System | DSS` now shows the new scale setting of “24k”



Distributed Services Switches Overview

Distributed Services Switches (3)

□	Switch Name	Switch OS Version ↑	Management	Health	Policy Scale
□	P0-11	DL.10.10.0001AJ-45-g8cd099a9cbd77	Mode: Inband IP address: 10.30.2.53	✓ ✓	Max 24K Rules
□	P0-36	DL.10.10.0001AJ-45-g8cd099a9cbd77	Mode: Inband IP address: 10.30.2.52	✓ ✓	Max 24K Rules
□	P1-27	DL.10.10.0001AJ-45-g8cd099a9cbd77	Mode: Inband IP address: 10.30.22.2	✓ ✓	Max 24K Rules

Figure 52. Switch status after the switch is rebooted and new profile is applied

> ***NOTE:*** *Policy Rule Scale is a cluster-wide setting; all switches managed by the same PSM cluster will have the same scale setting. Mixing different rule scale settings for different switches within the same cluster is not supported.*

When a new switch joins a PSM cluster, if the default profile value on the switch does not match the value configured in the PSM, then the switch will need another reboot to get adjusted automatically.

__NOTE__: With a 6K policy profile:

- *The max-dsm-rules per policy is set to 6138.*
- *Up to 80 policies can be defined if each policy is of size 6k.*
- *Up to 499 policies can be defined if each policy is of size <=1k.*

With a 24K policy profile:

- *The max-dsm-rules per policy is set to 24,570.*
- *Up to 17 unique policies can be defined if each policy is of size 24k.*
- *Up to 102 unique policies can be defined if each policy is of size <=4k*

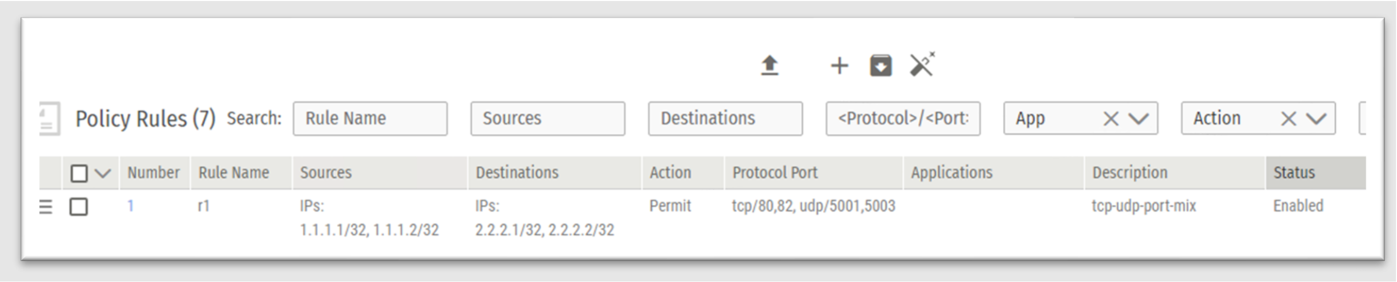
NOTE: The number of rules seen in a PSM policy is not the same as the number of rules seen by the DSM. The next section talks about the rule expansion logic between a given rule in PSM and corresponding rule in DS

Verifying the Number of Rules Consumed in the Data Plane

Before release 1.62.2-T, a given rule in PSM may get expanded into multiple internal rules on DSM. This expansion is based on how the rule is constructed in PSM. In general, the number of expanded rules in DSM is based on the formula:-

*number of DSM rules = number_of_sources_in_rule * number_of_destinations_in_rule * number_of_proto_port_pairs_in_rule*

For example, a rule definition like this on PSM would lead to 16 rules in DSM



Number	Rule Name	Sources	Destinations	Action	Protocol Port	Applications	Description	Status
1	r1	IPs: 1.1.1.1/32, 1.1.1.2/32	IPs: 2.2.2.1/32, 2.2.2.2/32	Permit	tcp/80,82, udp/5001,5003		tcp-udp-port-mix	Enabled

Number of sources: 2

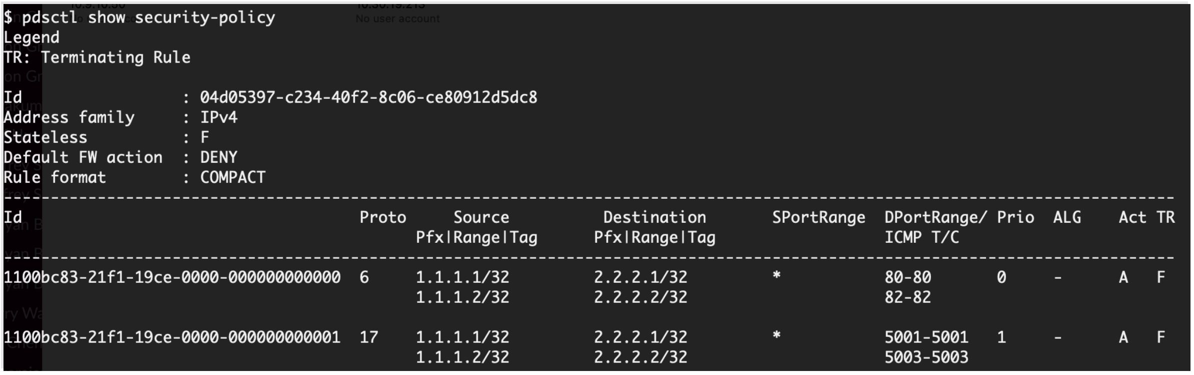
Number of destinations: 2

Number of proto-port-pairs: 4 (2 TCP port-pairs, 2 UDP port-pairs)

Overall expanded rules: 2*2*4 = 16 rules on DSM

Now with release 1.62.2-T, there is optimization done using compact rules, wherein multiple source IPs/prefixes, multiple destination IPs/prefixes, multiple ports (continuous/discontinous) for a given protocol type (TCP/UDP) can be compacted with a same rule in DSM. This exponentially helps in optimizing the number of rules consumed in DSM and helps with overall rule scale.

The above PSM rule with the compact rule implementation will look like this in DSM:



```
$ pdsctl show security-policy
Legend
TR: Terminating Rule

Id                  : 04d05397-c234-40f2-8c06-ce80912d5dc8
Address family      : IPv4
Stateless           : F
Default FW action   : DENY
Rule format         : COMPACT
----------------------------------------------------------------------------------------------------------------------
Id                                   Proto      Source            Destination       SPortRange  DPortRange/ Prio  ALG    Act TR
                                             Pfx|Range|Tag     Pfx|Range|Tag                    ICMP T/C
----------------------------------------------------------------------------------------------------------------------
1100bc83-21f1-19ce-0000-000000000000  6      1.1.1.1/32        2.2.2.1/32        *           80-80       0     -      A   F
                                             1.1.1.2/32        2.2.2.2/32                    82-82

1100bc83-21f1-19ce-0000-000000000001  17     1.1.1.1/32        2.2.2.1/32        *           5001-5001   1     -      A   F
                                             1.1.1.2/32        2.2.2.2/32                    5003-5003
```

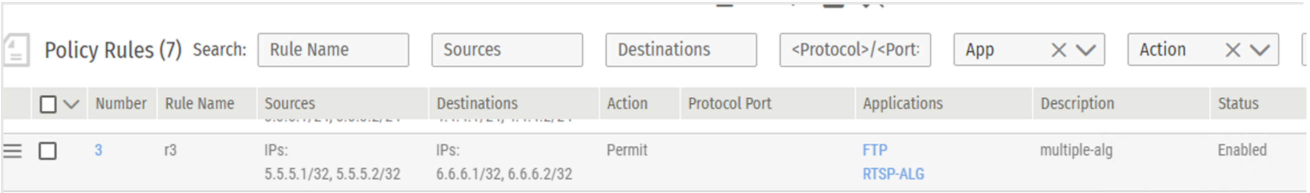
Effectively it gets programmed as 2 rules in DSM, 1 rule for TCP and 1 rule for UDP.

As shown in the above example, multiple source-prefixes, multiple destination-prefixes, multiple port numbers for a given protocol (TCP/UDP) can be compacted in a single DSM rule.

Exceptions to rule compaction:

The following are the exception to compaction of rules:

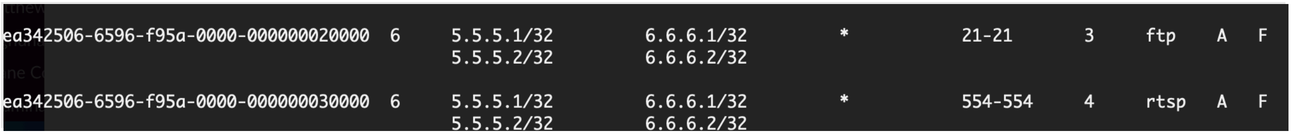
- Mix of TCP ports, UDP ports (or any other protocol) cannot be compacted. It will result in 1 unique rule per protocol in DSM.
- When multiple ALGs are used in the same PSM rule, it is split into 1 rule per ALG on PSM



Policy Rules (7) Search: Rule Name Sources Destinations <Protocol>/<Port: App Action

	Number	Rule Name	Sources	Destinations	Action	Protocol Port	Applications	Description	Status
☰ □	3	r3	IPs: 5.5.5.1/32, 5.5.5.2/32	IPs: 6.6.6.1/32, 6.6.6.2/32	Permit		FTP RTSP-ALG	multiple-alg	Enabled

In the above example, rule “r3” on PSM contains 2 ALG types, FTP and RTSP. Even though both ALGs are of the same protocol type (TCP), it does not get compacted in DSM but rather gets installed as 1 rule per ALG type.



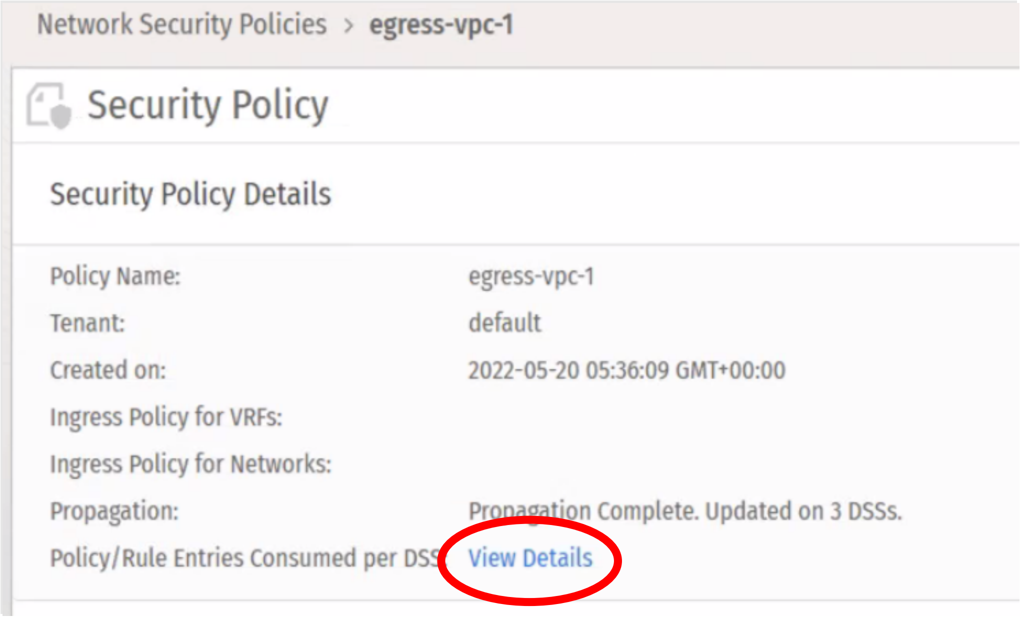
```
ea342506-6596-f95a-0000-000000020000  6     5.5.5.1/32        6.6.6.1/32        *           21-21       3     ftp    A   F
                                            5.5.5.2/32        6.6.6.2/32

ea342506-6596-f95a-0000-000000030000  6     5.5.5.1/32        6.6.6.1/32        *           554-554     4     rtsp   A   F
                                            5.5.5.2/32        6.6.6.2/32
```

With the 6K policy profile, if the number of expanded rules is >=1023, then the DSM internally combines multiple policies to accommodate them. For example, if a policy with 6K rules is created, then the DSM uses 6 internal policies of 1K rules each to support it.

With the 24K policy profile, if the number of expanded rules is >=4095, then the DSM internally combines multiple policies to accommodate them. For example, if a policy with 24K rules is created, then the DSM uses 6 internal policies of 4K rules each to support it.

In order to avoid exceeding the switch’s available internal resources, it is important to be aware of how many internal policies and internal expanded rules are actually consumed in the data plane.

To verify internal policies and rules, navigate to `Tenants | Security Policies`, then click on the name of the policy you wish to verify. Under Security Policy Details, click on "View Details" .



Network Security Policies > egress-vpc-1

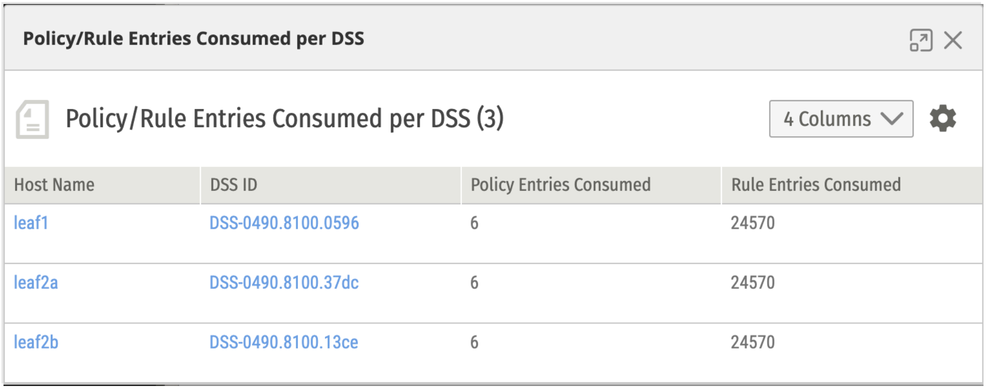
Security Policy

Security Policy Details

Policy Name:	egress-vpc-1
Tenant:	default
Created on:	2022-05-20 05:36:09 GMT+00:00
Ingress Policy for VRFs:	
Ingress Policy for Networks:	
Propagation:	Propagation Complete. Updated on 3 DSSs.
Policy/Rule Entries Consumed per DSS:	View Details

Figure 53. Select View Details to see Policy/Rule Entries Consumed per DSS.

1. A new window, as seen in figure Figure 54, shows the actual number of internal policies and actual number of rule entries consumed in hardware.



Policy/Rule Entries Consumed per DSS

Policy/Rule Entries Consumed per DSS (3)

Host Name	DSS ID	Policy Entries Consumed	Rule Entries Consumed
leaf1	DSS-0490.8100.0596	6	24570
leaf2a	DSS-0490.8100.37dc	6	24570
leaf2b	DSS-0490.8100.13ce	6	24570

Figure 54. Detail view: Policy/Rule Entries Consumed per DSS

Alerts can be set to notify if the number of policies and rules in the data plane have reached the resource limit:

2. Navigate to `Monitoring | Alerts & Events` , then add a new Alert Policy

3. Click on the “OBJECT BASED ALERT POLICIES” tab:

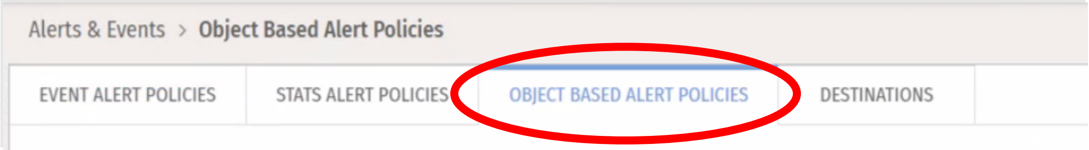
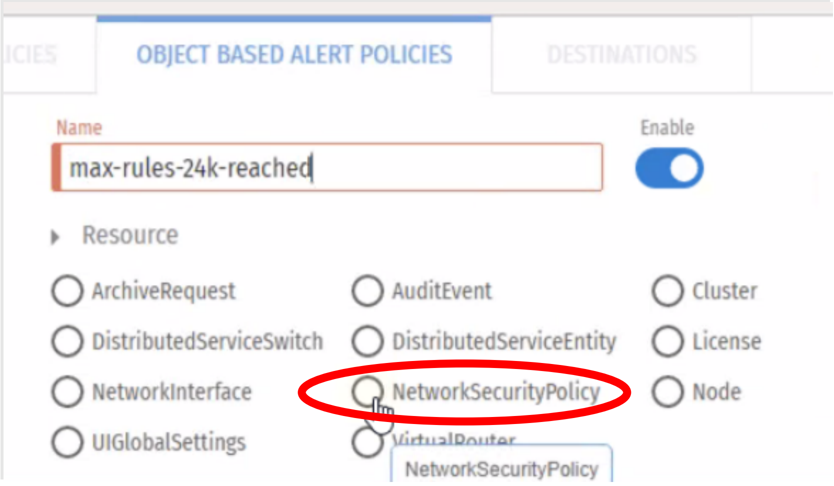

Alerts & Events > Object Based Alert Policies
EVENT ALERT POLICIES
STATS ALERT POLICIES
OBJECT BASED ALERT POLICIES
DESTINATIONS

Figure 55. Object Based Alert Policies tab selector

- Type in the policy name and choose “NetworkSecurityPolicy”

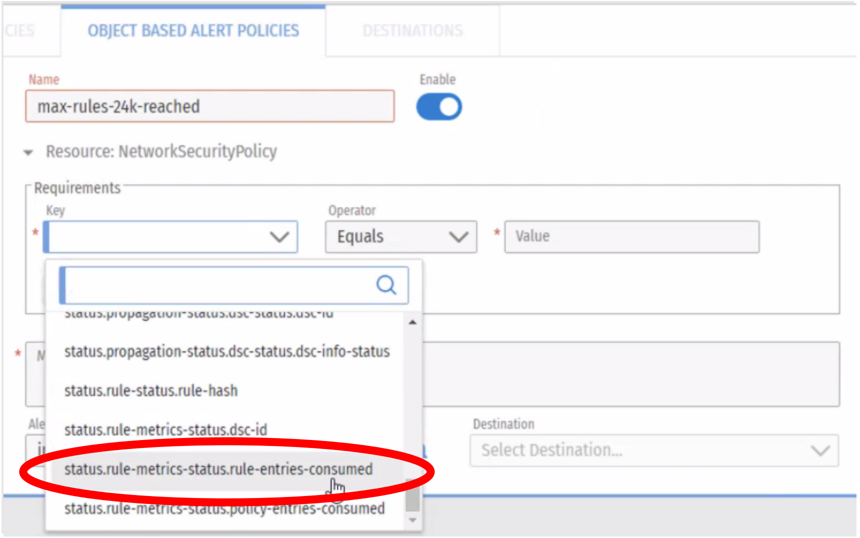


Name: max-rules-24k-reached

Enable: ●

Figure 56. Object Based Alert Policy for Network Security Policy

4. In the next screen, from the “Key” dropdown menu select “status.rule-metrics-status.rule-entries-consumed”

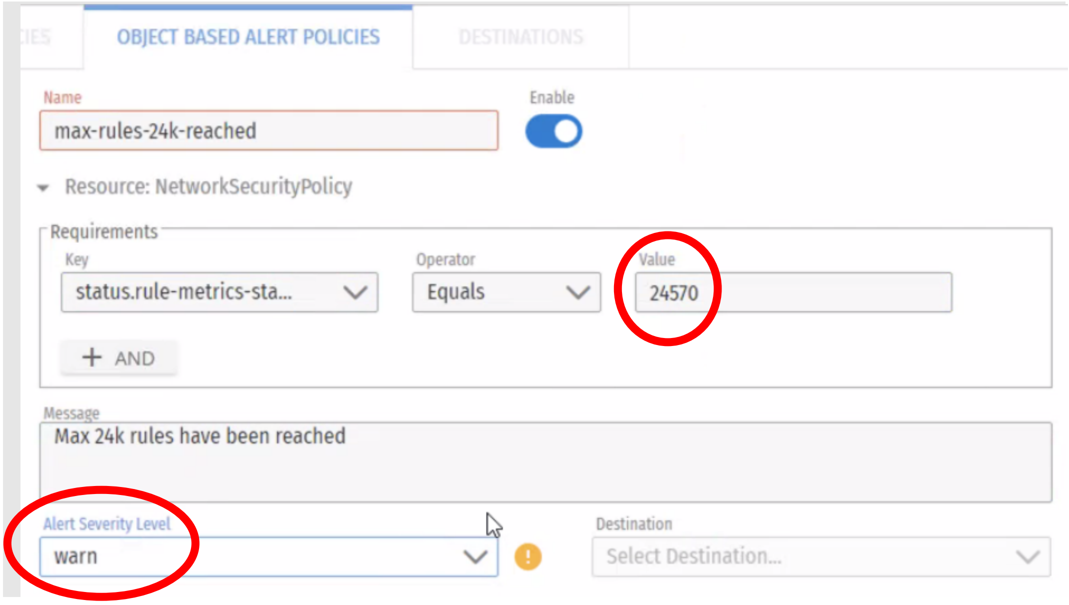


CIES
OBJECT BASED ALERT POLICIES
DESTINATIONS
Enable
Resource: NetworkSecurityPolicy
Requirements
Key
Operator
Equals
Value
status.propagation-status.dsc-status.dsc-info-status
status.rule-status.rule-hash
status.rule-metrics-status.dsc-id
status.rule-metrics-status.rule-entries-consumed
status.rule-metrics-status.policy-entries-consumed
Destination
Select Destination...

Name: max-rules-24k-reached

Figure 57. Object Based Alert Policy for Rule-Entries Consumed in Policy

5. Put the max allowed value of 24570 in the “Value” field and change the severity level to “warn”



Name: max-rules-24k-reached

Key: status.rule-metrics-sta...

Operator: Equals

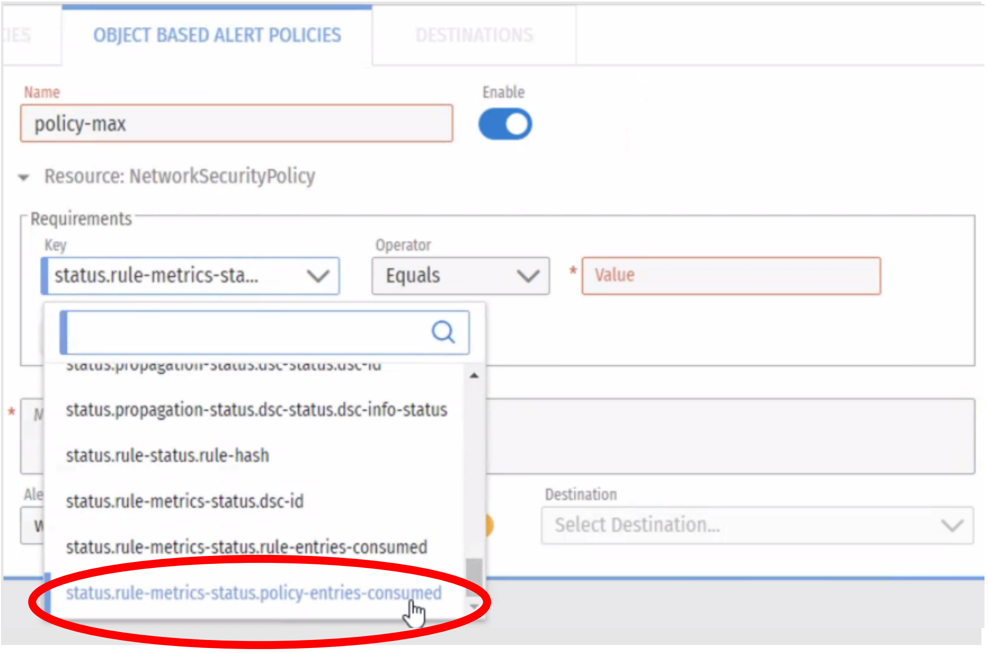
Value: 24570

Message: Max 24k rules have been reached

Alert Severity Level: warn

Figure 58. Selecting values and Alert Security Level

6. Similarly, create another alert policy once the internal policies number reaches 6:


IES
OBJECT BASED ALERT POLICIES
DESTINATIONS
Enable
status.propagation-status.dsc-status.dsc-id
status.propagation-status.dsc-status.dsc-info-status
status.rule-status.rule-hash
status.rule-metrics-status.dsc-id
status.rule-metrics-status.rule-entries-consumed
status.rule-metrics-status.policy-entries-consumed
Select Destination...

Name: policy-max

Resource: NetworkSecurityPolicy

Requirements

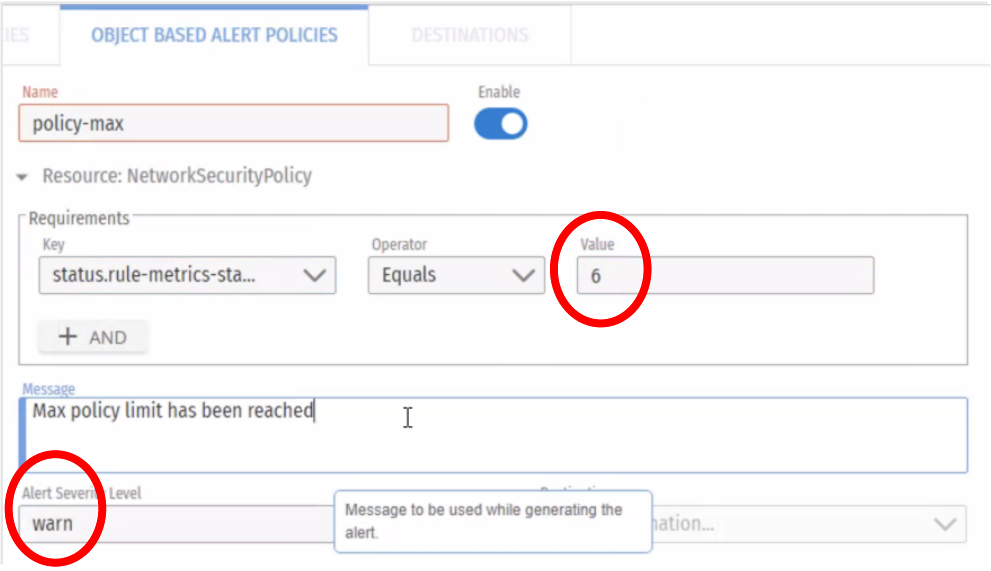
Key: status.rule-metrics-sta...

Operator: Equals

* Value

Destination

Figure 59. Object Based Alert Policy for Policy Entries Consumed


IES
OBJECT BASED ALERT POLICIES
DESTINATIONS
Enable
+ AND
Message to be used while generating the alert.
ation...

Name: policy-max

Resource: NetworkSecurityPolicy

Requirements

Key: status.rule-metrics-sta...

Operator: Equals

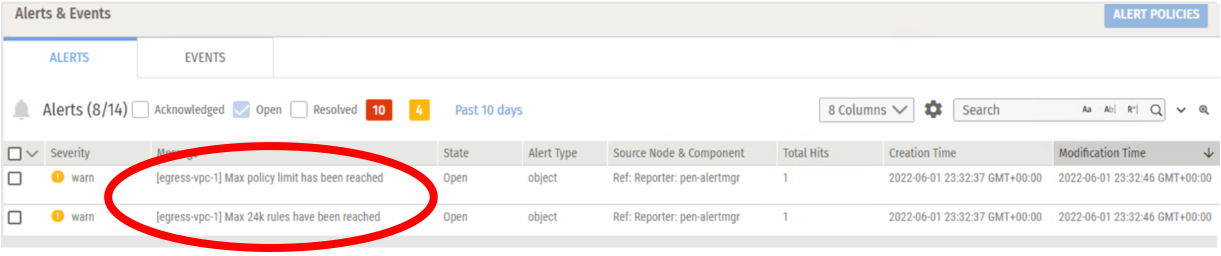
Value: 6

Message: Max policy limit has been reached

Alert Severity Level: warn

Figure 60. Selecting values and Alert Security Level for Policy Entries Consumed

With these two alert policies configured, once the internal policy number or the internal expanded rule number reaches the configured limit, alert messages will be raised to warn the administrator.



Alerts & Events
ALERT POLICIES
ALERTS
EVENTS
Alerts (8/14) □ Acknowledged ☑ Open □ Resolved 10 4 Past 10 days
8 Columns
Search

□	Severity	Message	State	Alert Type	Source Node & Component	Total Hits	Creation Time	Modification Time
□	warn	[egress-vpc-1] Max policy limit has been reached	Open	object	Ref: Reporter: pen-alertmgr	1	2022-06-01 23:32:37 GMT+00:00	2022-06-01 23:32:46 GMT+00:00
□	warn	[egress-vpc-1] Max 24k rules have been reached	Open	object	Ref: Reporter: pen-alertmgr	1	2022-06-01 23:32:37 GMT+00:00	2022-06-01 23:32:46 GMT+00:00

Figure 61. Warnings that the policy limit has been reached

***NOTE**: Profile change is a significantly disruptive operation, as it involves policy detachment and restarting switches. It is highly recommended to plan and implement the appropriate scale profile for your site during the initial design to be established when the PSM cluster is bootstrapped.*

Understanding Hierarchical Security Policy

Data center networking is usually implemented as a hierarchical design: a VRF represents the higher level network, and each VRF network consists of multiple VLAN networks which represent the lower level network, as in figure 64:

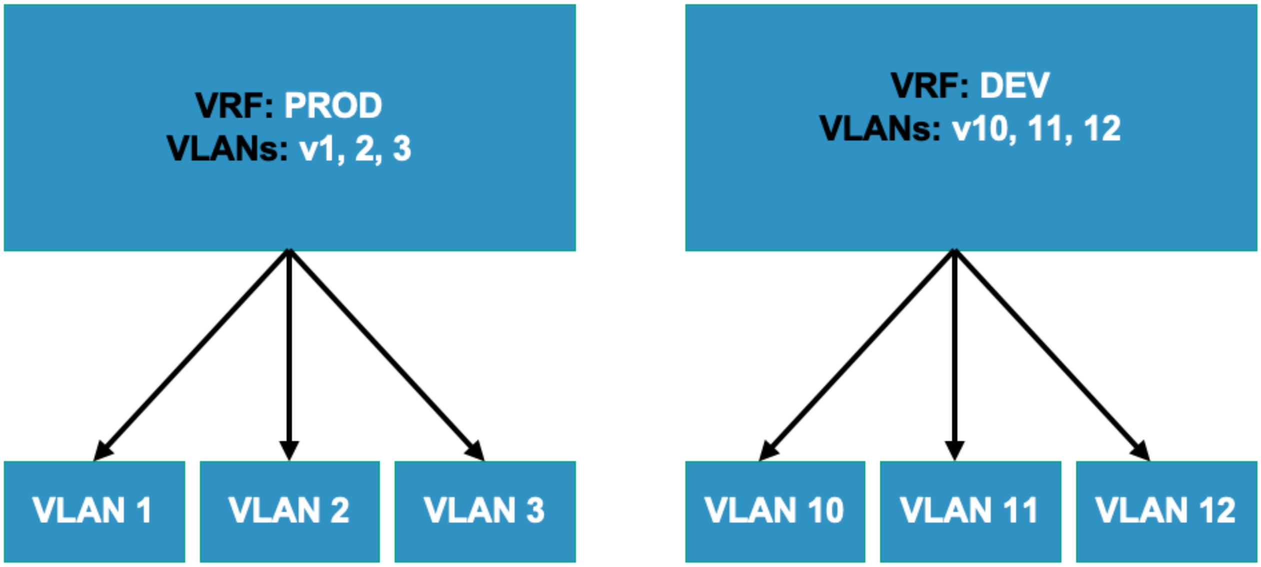


Figure 62. A typical data center hierarchical network design

Attaching and enforcing security policies at only the VLAN level may not be sufficient for some enterprise sites for any of several reasons:

- For simplicity purposes, sites may want to have a single unified security policy enforced across all VLANs within the VRF. Attaching the same policy explicitly to each and every VLAN can be cumbersome.
- In a large enterprise data center deployment, common scenarios include:
    - A set of commonly shared services are provided to all workloads in the data center, such as DNS, DHCP, or NTP.
    - Different VLAN networks provide different services such as production vs development, depending on what apps are running on them.
    - The commonly shared services are usually stable or fixed over time. Administers commonly would prefer not to change security policies on a regular basis. However, app-specific services may change frequently. Applications can be brought up and down regularly in a dynamic environment. Their related policies would need to be modified in turn.

The DSS hierarchical security policy model is designed to handle the above scenarios by supporting the security policy to be attached and enforced at the VRF level:

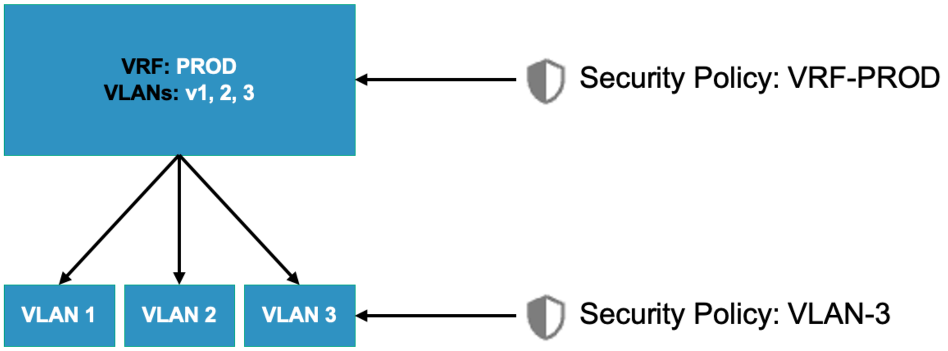


Figure 63. Attaching Security Policy

Policy Enforcement

- Two levels of policy enforcement will happen in both directions when traffic is between DSSes
- Traffic going out of the host will be subject to egress policy enforcement, first at Network level, followed by VRF level. Traffic coming into the host will be subject to ingress policy enforcement first at VRF level followed by and Network level.

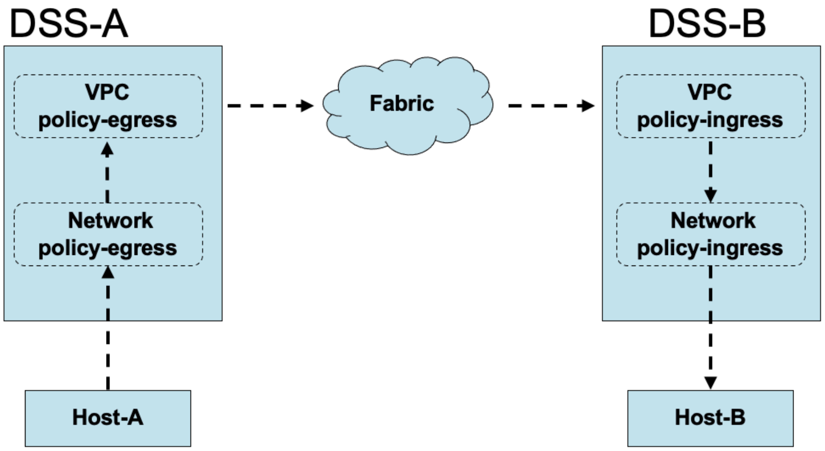


Figure 64. Egress/Ingress Policy Enforcement at VRF/Network levels

- For *local switched traffic* (between hosts connected to the same DSS):
  - Two levels of policy enforcement will happen in both directions for routed traffic and one direction (egress) for switched traffic
  - Traffic going out of the host will be subject to egress policy enforcement, first at Network level, followed by VRF level. Traffic coming into the host will be subject to ingress policy enforcement first at VRF level followed by and Network level.

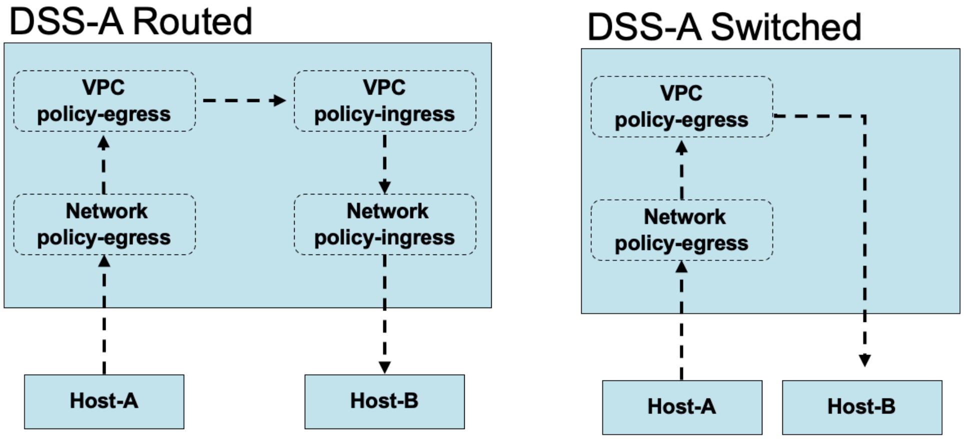


Figure 65. Egress/Ingress Policy Enforcement at VRF/Network levels

- If there is a “Deny” at any level (VRF/Network) the traffic will be denied (both ingress and egress)
- If no policy is configured, that will be considered as “Allow” and the traffic will pass
- Tables 5 and 6 list the final result for egress/ingress policy evaluation:

Network→	VRF	Final result
Allow	Allow*	**Allow**
Allow	Deny*	**Deny**
Deny*	Allow	**Deny**
Deny*	Deny	**Deny**
No policy	Allow*	**Allow**
No policy	Deny*	**Deny**
Allow*	No policy	**Allow**
Deny*	No policy	**Deny**
No policy	No policy	**Allow**

Table 6. Egress policy enforcement rules. For Egress, network policy is evaluated first followed by VRF policy.
** indicates rule-stats / FW log hits*

VRF→	Network	Final result
Allow	Allow*	**Allow**
Allow	Deny*	**Deny**
Deny*	Allow	**Deny**
Deny*	Deny	**Deny**
No policy	Allow*	**Allow**
No policy	Deny*	**Deny**
Allow*	No policy	**Allow**
Deny*	No policy	**Deny**
No policy	No policy	**Allow**

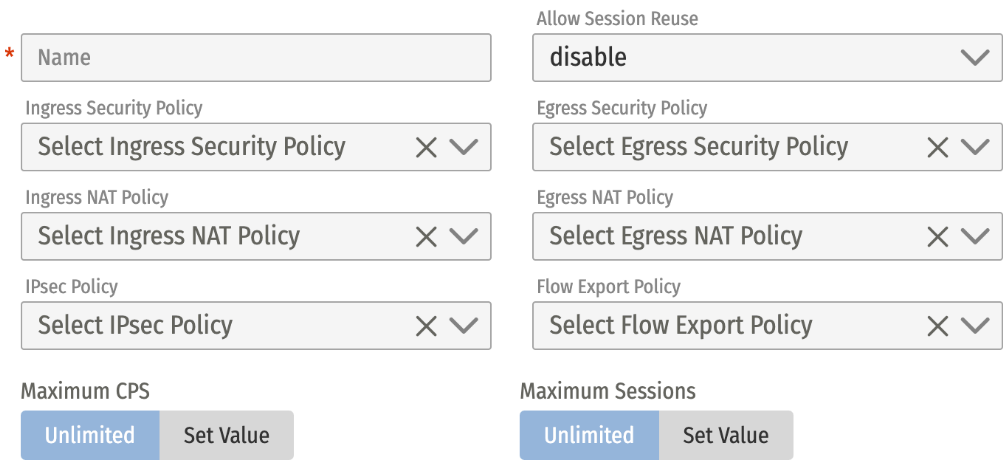
Table 7. Ingress policy enforcement rules. For Ingress, VRF policy is evaluated first followed by network policy.
** indicates rule-stats / FW log hits*

Rule-stats: Statistics that are maintained by DSM at a per-rule level for each policy accounting for number of packets hitting a rule during flow-create. These statistics are also exported to PSM and are available to view as "Connection Hits" in the policy UI page against each rule.

FW Log hits: A FW log is created and exported to configured destination (PSM/external) for each rule-hit within policy at the time of flow-create / flow-delete

Configuration and Verification

To apply Stateful L4 Security Policy to a VRF, navigate to `Tenants → VRF` in the PSM GUI, and click on ADD VRF.

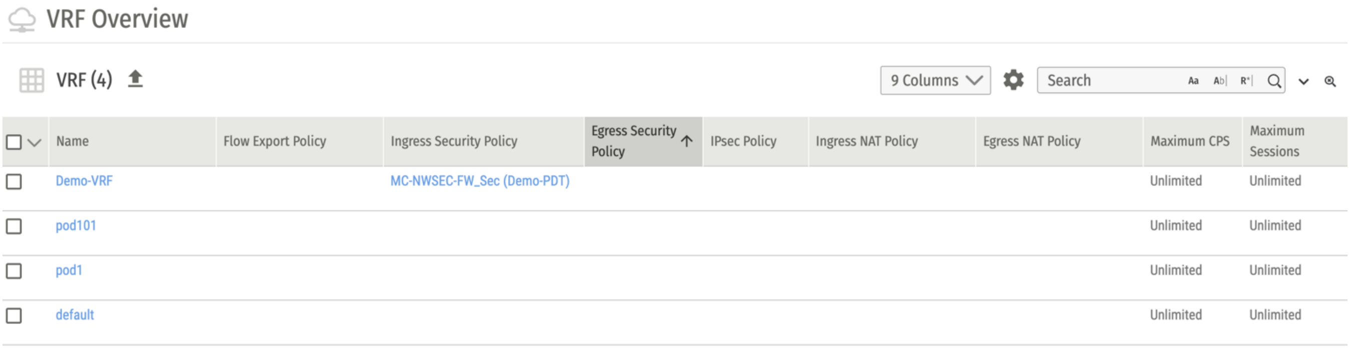


* Name	Allow Session Reuse: disable
Ingress Security Policy: Select Ingress Security Policy	Egress Security Policy: Select Egress Security Policy
Ingress NAT Policy: Select Ingress NAT Policy	Egress NAT Policy: Select Egress NAT Policy
IPsec Policy: Select IPsec Policy	Flow Export Policy: Select Flow Export Policy
Maximum CPS: Unlimited / Set Value	Maximum Sessions: Unlimited / Set Value

Figure 66. Hierarchical design

- **Name** of the VRF (string).
- **Ingress Security Policy:** choose a previously configured security policy.
- **Egress Security Policy:** choose a previously configured security policy.
- **Flow export policies:** choose a previously configured policy.
- The Maximum CPS and Maximum Sessions for a VRF can be set to Unlimited or to a value.
- **Maximum CPS per DSE**, which can be up to the maximum supported connections per Second (range 1000-1000,000) for the entire system.
- **Maximum Sessions per DSE**, which can be up to the maximum supported sessions (range 10,000 - 5000,000) for the entire system.

The VRF policies can be verified in the VRF Overview under `Tenants → VRF`



VRF Overview

VRF (4)

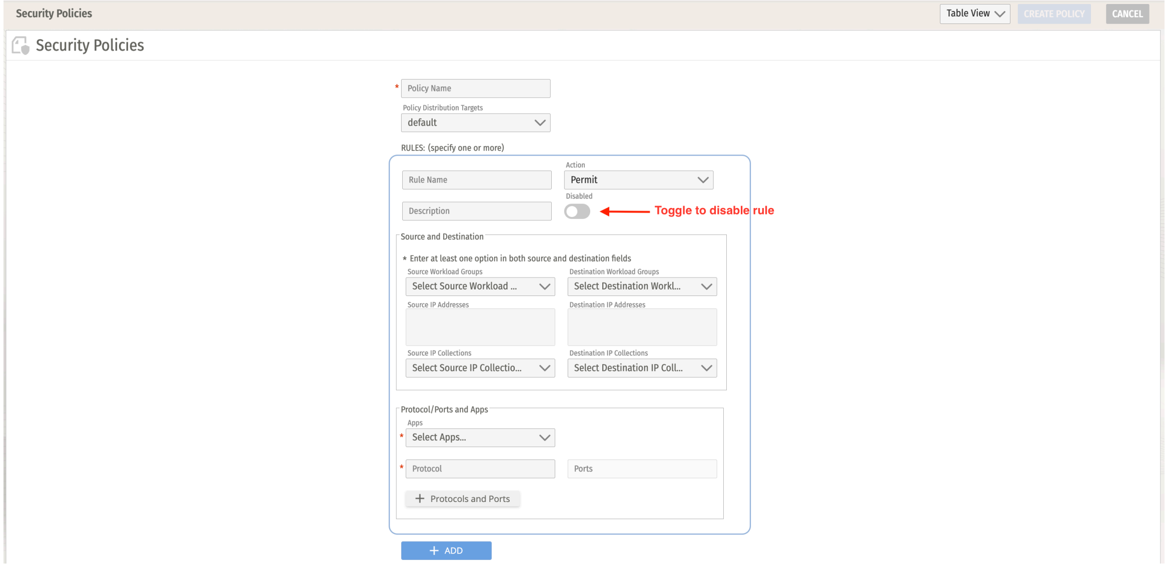
□	Name	Flow Export Policy	Ingress Security Policy	Egress Security Policy	IPsec Policy	Ingress NAT Policy	Egress NAT Policy	Maximum CPS	Maximum Sessions
□	Demo-VRF		MC-NWSEC-FW_Sec (Demo-PDT)					Unlimited	Unlimited
□	pod101							Unlimited	Unlimited
□	pod1							Unlimited	Unlimited
□	default							Unlimited	Unlimited

Figure 67. Hierarchical design

Enable/Disable Individual Firewall Rules

Firewall rules can be enabled or disabled individually. There are two ways to use this feature:

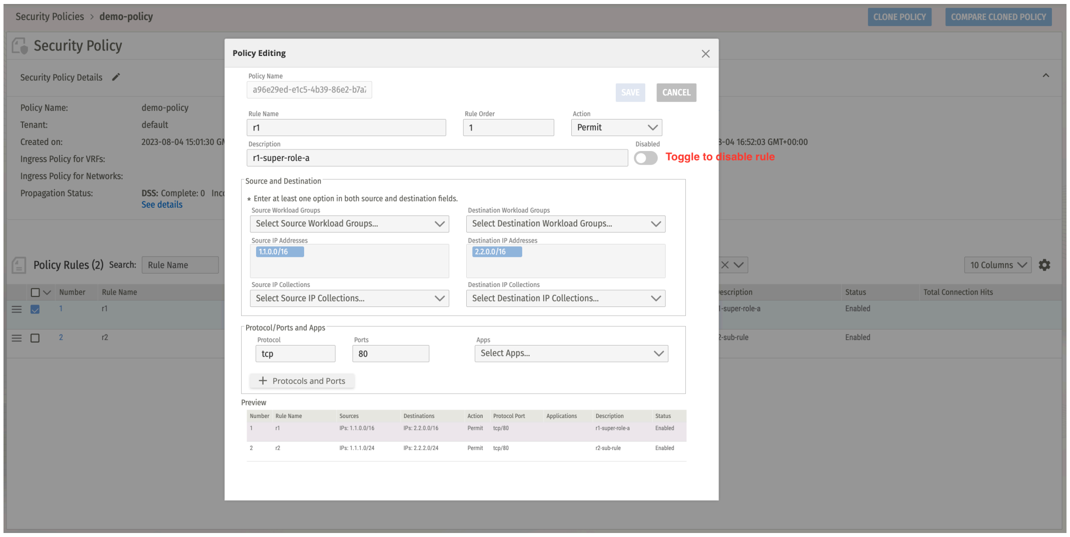
1. When creating a new security policy, each rule can be added with the option to disable it (rules are enabled by default):



Toggle to disable rule

Figure 68. Specify a new rule should be disabled

2. For existing security policies, go to the Policy Editing page, where each rule can be enabled or disabled individually:



Policy Name: demo-policy
Tenant: default
Created on: 2023-08-04 15:01:30 GM
Ingress Policy for VRFs:
Ingress Policy for Networks:
Propagation Status: DSS: Complete: 0 Inc
See details

Policy Name: a96e29ed-e1c5-4b39-86e2-b7a7

Rule Name: r1
Rule Order: 1
Action: Permit
Description: r1-super-role-a
Disabled

Source and Destination

* Enter at least one option in both source and destination fields.

Source Workload Groups: Select Source Workload Groups...
Destination Workload Groups: Select Destination Workload Groups...
Source IP Addresses: 1.1.0.0/16
Destination IP Addresses: 2.2.0.0/16
Source IP Collections: Select Source IP Collections...
Destination IP Collections: Select Destination IP Collections...

Protocol/Ports and Apps

Protocol: tcp
Ports: 80
Apps: Select Apps...

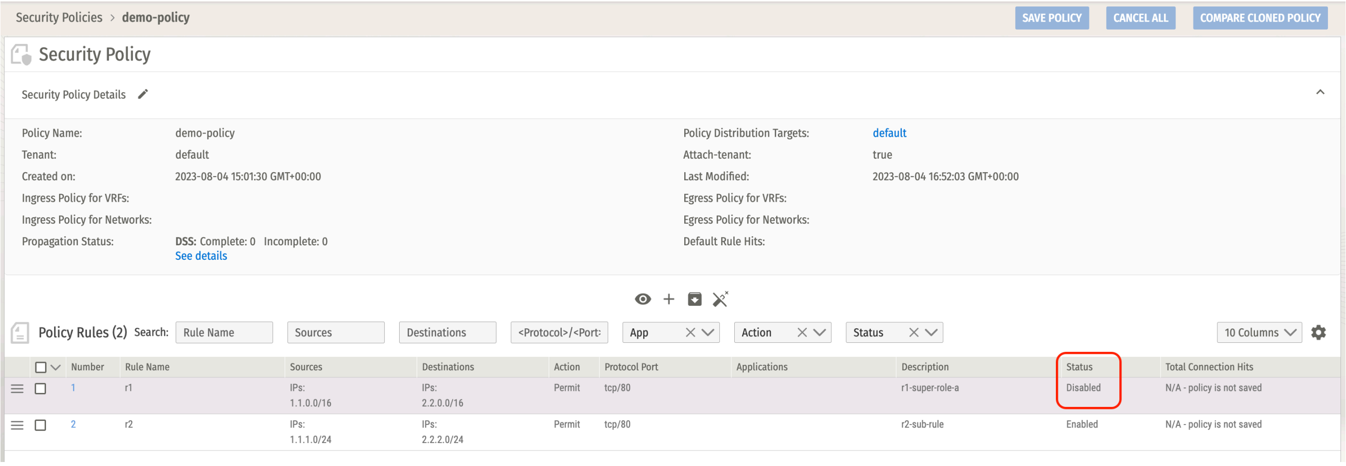
Preview

Number	Rule Name	Sources	Destinations	Action	Protocol Port	Applications	Description	Status
1	r1	IPs: 1.1.0.0/16	IPs: 2.2.0.0/16	Permit	tcp/80		r1-super-role-a	Enabled
2	r2	IPs: 1.1.1.0/24	IPs: 2.2.2.0/24	Permit	tcp/80		r2-sub-rule	Enabled

Figure 69. Disable an existing rule

NOTE: *Multiple rules can be disabled or enabled at once from the security policy details screen, by selecting those rules and then using the Enable/Disable selected rules icon ( ) that appears at the top of the rules list.*

A disabled rule is indicated in the “Status” field on the Security Policy page:



Security Policies > demo-policy

SAVE POLICY   CANCEL ALL   COMPARE CLONED POLICY

Security Policy

Security Policy Details

Policy Name:	demo-policy	Policy Distribution Targets:	default
Tenant:	default	Attach-tenant:	true
Created on:	2023-08-04 15:01:30 GMT+00:00	Last Modified:	2023-08-04 16:52:03 GMT+00:00
Ingress Policy for VRFs:		Egress Policy for VRFs:	
Ingress Policy for Networks:		Egress Policy for Networks:	
Propagation Status:	DSS: Complete: 0 Incomplete: 0 See details	Default Rule Hits:	

Policy Rules (2) Search: Rule Name Sources Destinations <Protocol>/<Port: App Action Status 10 Columns

Number	Rule Name	Sources	Destinations	Action	Protocol Port	Applications	Description	Status	Total Connection Hits
1	r1	IPs: 1.1.0.0/16	IPs: 2.2.0.0/16	Permit	tcp/80		r1-super-role-a	Disabled	N/A - policy is not saved
2	r2	IPs: 1.1.1.0/24	IPs: 2.2.2.0/24	Permit	tcp/80		r2-sub-rule	Enabled	N/A - policy is not saved

Figure 70. Identify disabled rules

Configuring Firewall Log Export

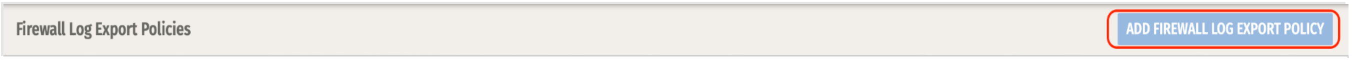
Firewall logs from the DSS can be exported to an external destination/collector, or to the PSM. Export to external collector destinations is only supported in the default VRF with inband IP.

Enabling firewall log export requires two steps:

- Defining Firewall Log export policy configuration
- Binding policy to the DSS

Firewall Log Export Policy Configuration

Navigate to `Tenants → Firewall Export Policies` and click on ADD FIREWALL LOG EXPORT POLICY:

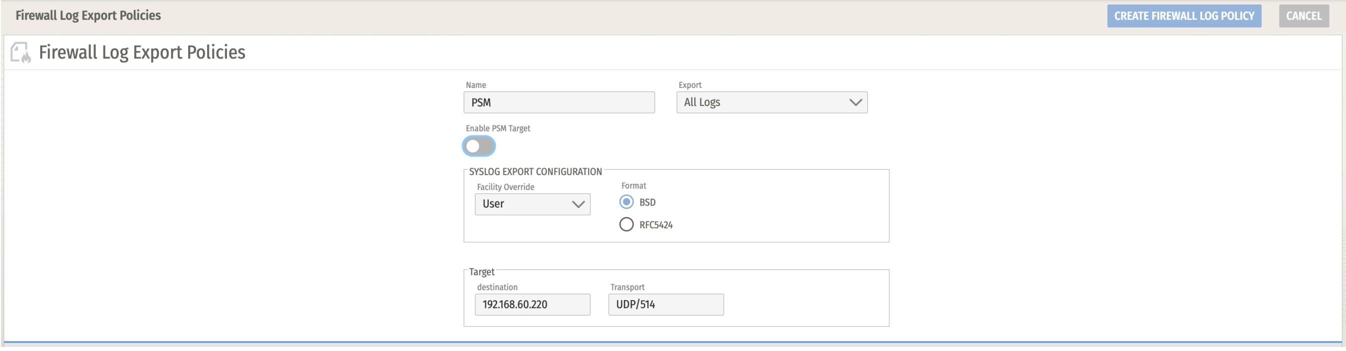


Firewall Log Export Policies

Figure 71. Firewall Log Policies screen detail, showing ADD FIREWALL button

Enter the relevant information and click CREATE FIREWALL LOG POLICY.

- Fill out the mandatory fields: `name`, `destination`, `transport`.
- Select the format and facility override.



Firewall Log Export Policies

Firewall Log Export Policies

Name: PSM

Export: All Logs

Enable PSM Target

SYSLOG EXPORT CONFIGURATION

Facility Override: User

Format: ● BSD ○ RFC5424

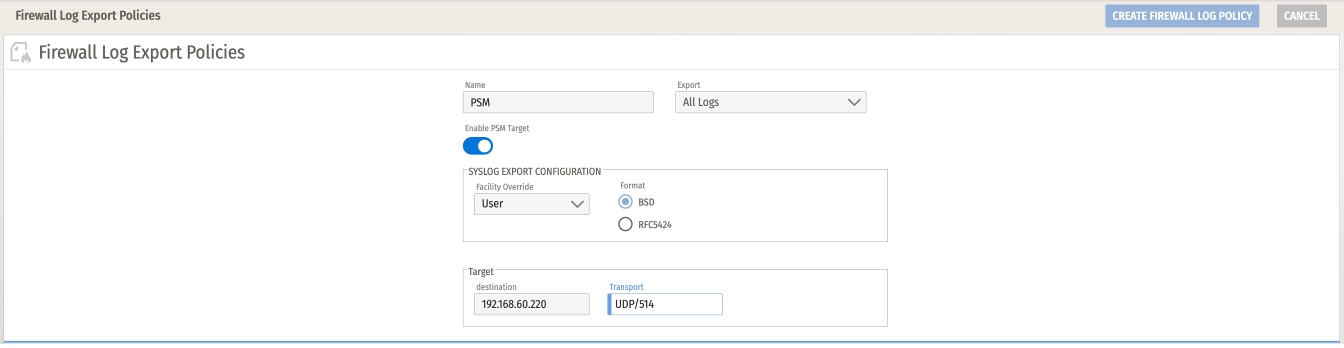
Target

destination: 192.168.60.220

Transport: UDP/514

Figure 72. Firewall log policy creation screen

Figure 73 shows an example of defining a firewall log policy to export to the PSM:



Firewall Log Export Policies

Firewall Log Export Policies

Name: PSM

Export: All Logs

Enable PSM Target

SYSLOG EXPORT CONFIGURATION

Facility Override: User

Format: ● BSD ○ RFC5424

Target

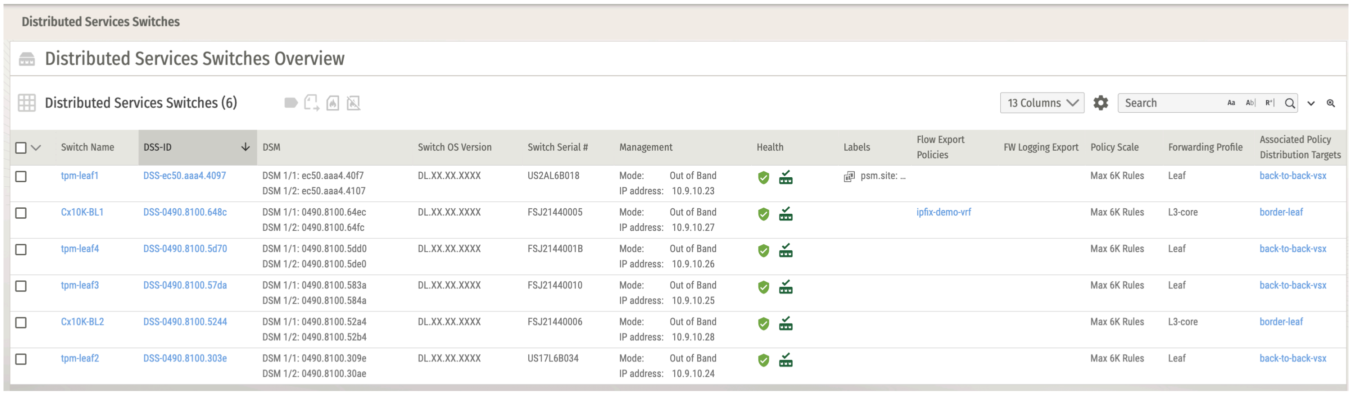
destination: 192.168.60.220

Transport: UDP/514

Figure 73. Firewall Log Export to PSM

Bind Export Policy to DSS

Navigate to “`System → DSS`” and click on the DSS-ID of the switch to apply the policy to.



Distributed Services Switches

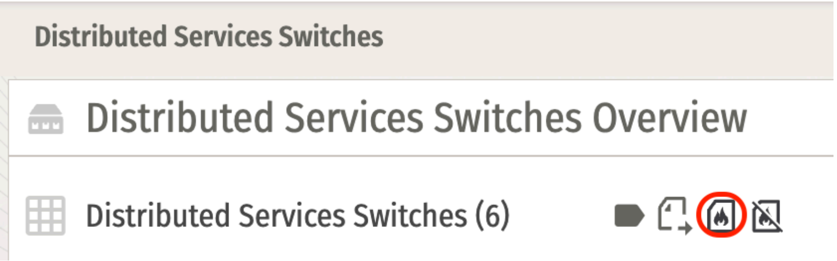
Distributed Services Switches Overview

Distributed Services Switches (6)

□	Switch Name	DSS-ID	DSM	Switch OS Version	Switch Serial #	Management	Health	Labels	Flow Export Policies	FW Logging Export	Policy Scale	Forwarding Profile	Associated Policy Distribution Targets
□	tpm-leaf1	DSS-ec50.aaa4.4097	DSM 1/1: ec50.aaa4.40f7 DSM 1/2: ec50.aaa4.4107	DL.XX.XX.XXXX	US2AL6B018	Mode: Out of Band IP address: 10.9.10.23	✓ ✓	psm.site: ...			Max 6K Rules	Leaf	back-to-back-vsx
□	Cx10K-BL1	DSS-0490.8100.648c	DSM 1/1: 0490.8100.64ec DSM 1/2: 0490.8100.64fc	DL.XX.XX.XXXX	FSJ21440005	Mode: Out of Band IP address: 10.9.10.27	✓ ✓		ipfix-demo-vrf		Max 6K Rules	L3-core	border-leaf
□	tpm-leaf4	DSS-0490.8100.5d70	DSM 1/1: 0490.8100.5dd0 DSM 1/2: 0490.8100.5de0	DL.XX.XX.XXXX	FSJ2144001B	Mode: Out of Band IP address: 10.9.10.26	✓ ✓				Max 6K Rules	Leaf	back-to-back-vsx
□	tpm-leaf3	DSS-0490.8100.57da	DSM 1/1: 0490.8100.583a DSM 1/2: 0490.8100.584a	DL.XX.XX.XXXX	FSJ21440010	Mode: Out of Band IP address: 10.9.10.25	✓ ✓				Max 6K Rules	Leaf	back-to-back-vsx
□	Cx10K-BL2	DSS-0490.8100.5244	DSM 1/1: 0490.8100.52a4 DSM 1/2: 0490.8100.52b4	DL.XX.XX.XXXX	FSJ21440006	Mode: Out of Band IP address: 10.9.10.28	✓ ✓				Max 6K Rules	L3-core	border-leaf
□	tpm-leaf2	DSS-0490.8100.303e	DSM 1/1: 0490.8100.309e DSM 1/2: 0490.8100.30ae	DL.XX.XX.XXXX	US17L6B034	Mode: Out of Band IP address: 10.9.10.24	✓ ✓				Max 6K Rules	Leaf	back-to-back-vsx

Figure 74. DSS list screen

Click on the icon to apply Firewall Log Policy:



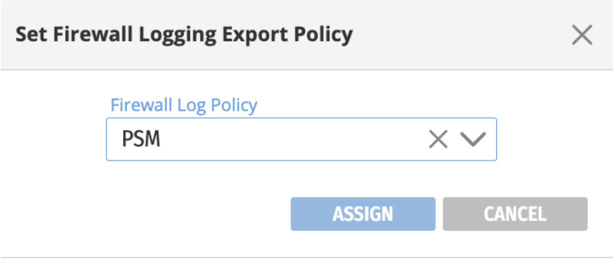
Distributed Services Switches

Distributed Services Switches Overview

Distributed Services Switches (6)

Figure 75. Select the "assign" action.

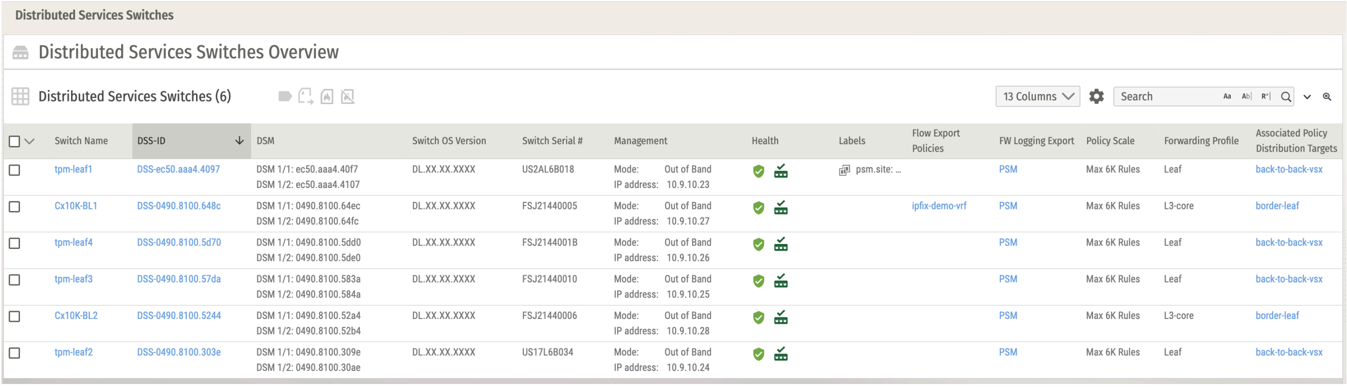
Select the previously-defined policy, then click ASSIGN:



Set Firewall Logging Export Policy

Firewall Log Policy: PSM

Figure 76. Policy assignment window



Distributed Services Switches

Distributed Services Switches Overview

Distributed Services Switches (6)

□	Switch Name	DSS-ID	DSM	Switch OS Version	Switch Serial #	Management	Health	Labels	Flow Export Policies	FW Logging Export	Policy Scale	Forwarding Profile	Associated Policy Distribution Targets
□	tpm-leaf1	DSS-ec50.aaa4.4097	DSM 1/1: ec50.aaa4.40f7 DSM 1/2: ec50.aaa4.4107	DL.XX.XX.XXXX	US2AL6B018	Mode: Out of Band IP address: 10.9.10.23	✓	psm.site: ...		PSM	Max 6K Rules	Leaf	back-to-back-vsx
□	Cx10K-BL1	DSS-0490.8100.648c	DSM 1/1: 0490.8100.64ec DSM 1/2: 0490.8100.64fc	DL.XX.XX.XXXX	FSJ21440005	Mode: Out of Band IP address: 10.9.10.27	✓		ipfix-demo-vrf	PSM	Max 6K Rules	L3-core	border-leaf
□	tpm-leaf4	DSS-0490.8100.5d70	DSM 1/1: 0490.8100.5dd0 DSM 1/2: 0490.8100.5de0	DL.XX.XX.XXXX	FSJ2144001B	Mode: Out of Band IP address: 10.9.10.26	✓			PSM	Max 6K Rules	Leaf	back-to-back-vsx
□	tpm-leaf3	DSS-0490.8100.57da	DSM 1/1: 0490.8100.583a DSM 1/2: 0490.8100.584a	DL.XX.XX.XXXX	FSJ21440010	Mode: Out of Band IP address: 10.9.10.25	✓			PSM	Max 6K Rules	Leaf	back-to-back-vsx
□	Cx10K-BL2	DSS-0490.8100.5244	DSM 1/1: 0490.8100.52a4 DSM 1/2: 0490.8100.52b4	DL.XX.XX.XXXX	FSJ21440006	Mode: Out of Band IP address: 10.9.10.28	✓			PSM	Max 6K Rules	L3-core	border-leaf
□	tpm-leaf2	DSS-0490.8100.303e	DSM 1/1: 0490.8100.309e DSM 1/2: 0490.8100.30ae	DL.XX.XX.XXXX	US17L6B034	Mode: Out of Band IP address: 10.9.10.24	✓			PSM	Max 6K Rules	Leaf	back-to-back-vsx

Figure 77. List of switches, shown with an attached firewall log policy

Once a firewall log policy is attached to a DSS, the list of switches should indicate the attachment, as shown in Figure 80; it will also be listed on the switch's detail page.

Firewall Log Record Format

NOTE: The order of the fields described in the table also indicates the order of the fields exported in the syslog message:

Field Name	Description
`ts`	Flow record timestamp, RFC3339
`flowaction`	Can be `flow_create` or `flow_delete`
`act`	Allow or Deny action applied for the session
`vpcid`	Source VRF UUID for the flow record
`sip`	Source IPv4 address of the flow
`sport`	Source TCP/UDP port of the flow
`dip`	Destination IPv4 address of the flow
`dport`	Destination TCP/UDP port of the flow
`proto`	IP protocol number
`sessionid`	Flow session identifier
`securitypolicyid`	UUID of the security policy
`ruleid`	Hash of the rule against which the flow was evaluated
`rulename`	User assigned name of the rule within the policy
`iflowpkts`	Packet count sent from flow Initiator to flow Responder
`iflowbytes`	Bytes count sent from flow Initiator to flow Responder
`rflowpkts`	Packet count sent from flow Responder to flow Initiator
`rflowbytes`	Bytes count sent from flow Responder to flow Initiator

Table 8. Firewall record fields (1/2)

Field Name	Description
`vlan`	VLAN-ID configured in Network Object
`producttype`	DSS
`softwareversion`	AOS-CX version
`serialnumber`	Serial Number of DSS
`devicename`	MAC address of DSS
`unitid`	Unit ID for DSM. Can be 1 or 2, since there are two DSMs per DSS.
`version`	V3
`policyname`	Name of security policy used to evaluate this flow
`policydisplayname`	Display name of security policy if policy is renamed, else null
`nattranslatedsrcip`	post-NAT translated source IP
`nattranslateddestip`	post-NAT translated destination IP
`nattranslateddestport`	post-NAT translated destination port
`encrypted`	Indicates if flow is subject to IPsec encryption/decryption
`direction`	Indicates if I-flow is from host/uplink
`createreason`	Indicates how the flow was created - flow_miss/flow_sync/vmotion/full_sync

Table 8. Firewall record fields (2/2)

Firewall Syslog Message Examples

Example of syslog message:

```
2023-04-25T05:38:51Z,flow_create,allow,4e78cf93-cf28-46eb-b6a3-
4b8111f2e3e3,163.2.1.99,34695,63.3.1.88,4000,17,1048832,33602eed-3dd3-45c5-91c5-
4358f95a69ca,6462113369068598675,allow-
udp,10,100,10,100,3000,DSS,DL.XX.XX.XXXX,US17L6B03D,0490.8100.32c8,2,v3,Ingress-SG-Policy-
1,,63.2.1.99,65.2.1.99,100,false,from-uplink,flow_miss
```

The following table gives example values for each exported field for the message displayed above:

Field	Example Value from Syslog
ts	2023-04-25T05:38:51Z
flowaction	flow_create
act	allow
vpcid	4e78cf93-cf28-46eb-b6a3-4b8111f2e3e3
sip	163.2.1.99
sport	34695
dip	63.3.1.88
dport	4000
protocol	17
sessionid	1048832
securitypolicyid	33602eed-3dd3-45c5-91c5-4358f95a69ca
ruleid	6462113369068598675
rulename	allow-udp
iflowpkts	10
iflowbytes	100
rflowpkts	10
rflowbytes	100
vlan	3000

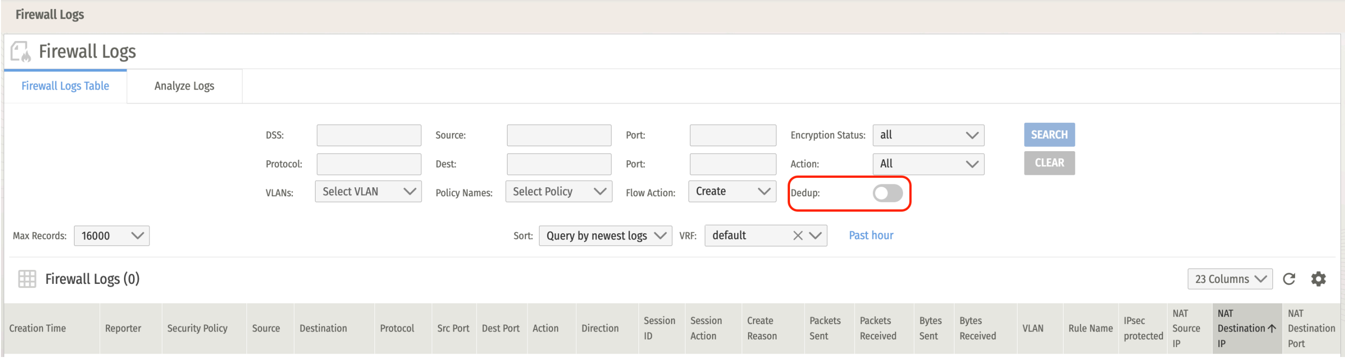
Table 9. Values shown in deny rule example (1/2)

Field	Example Value from Syslog
policyname	deny-tcp-policy
producttype	DSS
AOS-CX version	DL.XX.XX.XXXX
serialnumber	US17L6B03D
devicename	0490.8100.32c8
unitid	2
version	v3
nattranslatedsrcip	63.2.1.99
nattranslatedsrcip	65.2.1.99
nattranslateddestport	100
encrypted	false
direction	from-uplink
createreason	flow_miss

Table 9. Values shown in deny rule example (2/2)

Deduplication for Firewall Logs

The PSM backend supports a deduplication option for the API flowlog query option. This option can be enabled via the PSM, under the `Tenants | Firewall Logs` page, using the `Dedup` toggle:



DSS:
Source:
Port:
Encryption Status: all
Protocol:
Dest:
Port:
Action: All
VLANs: Select VLAN
Policy Names: Select Policy
Flow Action: Create
Dedup:
Max Records: 16000
Sort: Query by newest logs
VRF: default

Firewall Logs (0)

Creation Time	Reporter	Security Policy	Source	Destination	Protocol	Src Port	Dest Port	Action	Direction	Session ID	Session Action	Create Reason	Packets Sent	Packets Received	Bytes Sent	Bytes Received	VLAN	Rule Name	IPsec protected	NAT Source IP	NAT Destination IP ↑	NAT Destination Port

Figure 78. Enabling deduplication

When `Dedup` is enabled, duplicate flow records for a given 4-tuple are removed, When the same 4-tuple flow exists on a given switch over different time instances, only one instance, which is the latest instance of the flow record, is returned. When the same 4-tuple flow record is exported from multiple switches in the fabric, only one instance is maintained.

API users can enable flow log de-duplication with a specific type of aggregation.

```
message FwLogQuery {
     "start-time": xxx
     "end-time": xxx
     "source-ips": xxx
     "destination-ips": xxx
—snip—
     filter-out-duplicates: "do-not-filter" | "srcip-destip-
destport-proto"
}
```

When submitting a JSON format request, users can either enable 4-tuple based deduplication by specifying `srcip-destip-destport-proto`, or disable the deduplication by specifying `do-not-filter`.

Notes:

- By default the query will be with `do-not-filter` and all the logs in response will not be deduplicated.
- `srcip-destip-destport-proto` will make each combination of the four tuples only have one log entry in the response.
- The `filter-out-duplicates` option only supports the latest 24 hours of entries, including queries from any to any IP or queries with a specific IP address.
- The `filter-out-duplicates` option will not work with the `countOnly` type of `FwLogQuery.`

Flow Export (IPFIX)

IP Flow export, implementing the IETF IPFIX standard, is used to gain visibility into network traffic usage patterns within data center networks, which can then be used as a basis for creating firewall policies.

Configuring IPFIX flow export on the CX 10000 switches involves the following::

- Step 1: Switch-side configuration
    - Enable flow export globally
    - Specify source IP for exported IPFIX packets
- Step 2: PSM-side configuration
    - Configure the export policies in the PSM GUI
    - Apply the configured flow export policy to DSS/VRF

Step 1a: Enable Flow Export feature globally on the switch:

```
switch(config)# dsm
switch (config-dsm)# ipfix
```

Step 1b: Set the source IP address of the exported IPFIX packet:

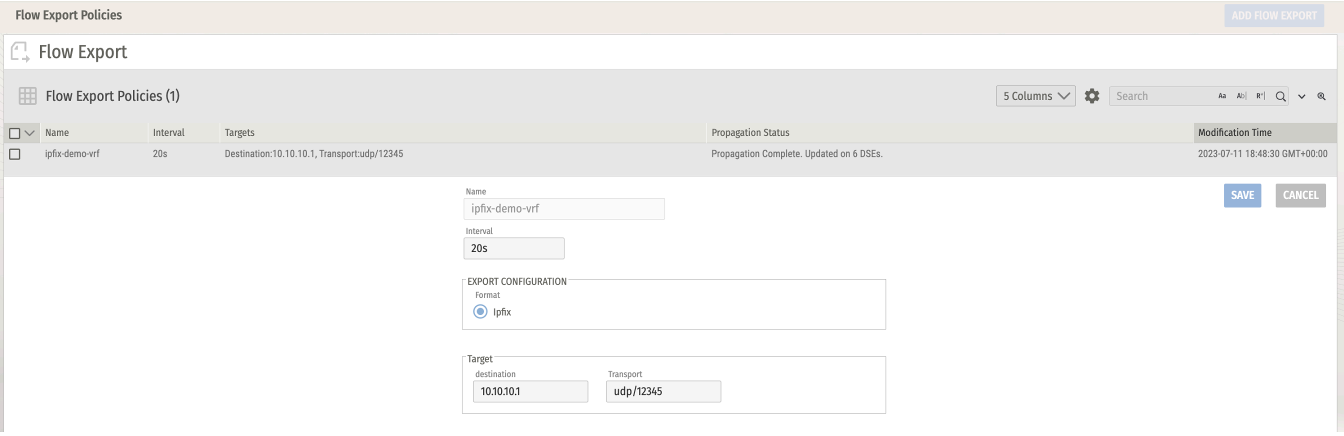
```
switch(config)# ip source-interface ipfix 1.1.1.1
```

or

```
switch(config)# ip source-interface ipfix interface loopback1
```

For production environments, although specifying the `source-interface` CLI option is not mandatory, it is highly recommended. Without the `ip source-interface` configuration, IPFIX packets will be exported with internal private IP addresses, and may get dropped in the underlay network due to unicast reverse path forwarding checks.

Step 2a: Configure the Flow Export policies under the “Tenants” -> “Flow Export” menu in PSM.



☐	Name	Interval	Targets	Propagation Status	Modification Time
☐	ipfix-demo-vrf	20s	Destination:10.10.10.1, Transport:udp/12345	Propagation Complete. Updated on 6 DSEs.	2023-07-11 18:48:30 GMT+00:00

Name: ipfix-demo-vrf

Interval: 20s

EXPORT CONFIGURATION

Format: ● Ipfix

Target

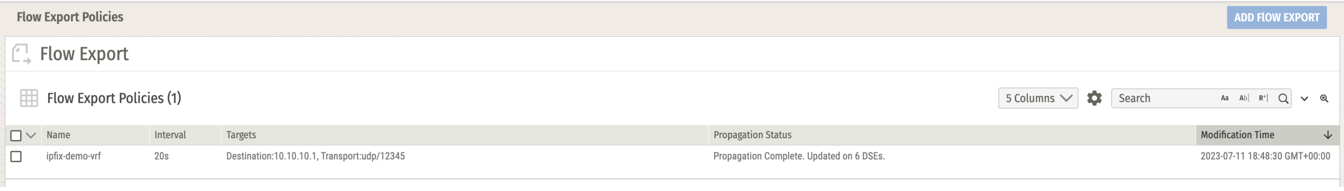
destination: 10.10.10.1

Transport: udp/12345

Figure 79. Flow Export Policies

In the above example, a policy called “ipfix-demo” is created with flow-export-interval of 20s with destination IP as “10.10.10.1” using UDP port of “12345”. Define all the fields and then click on “Create Flow Export Policy”.

Once the IPFIX policy is defined, it is propagated immediately to all registered DSSes. Each DSS receiving this IPFIX policy would start exporting the IPFIX template records (not the flow-records) to the defined external collector at the defined export frequency. This is done so that the collector can receive the IPFIX template records before receiving the actual IPFIX flow records.



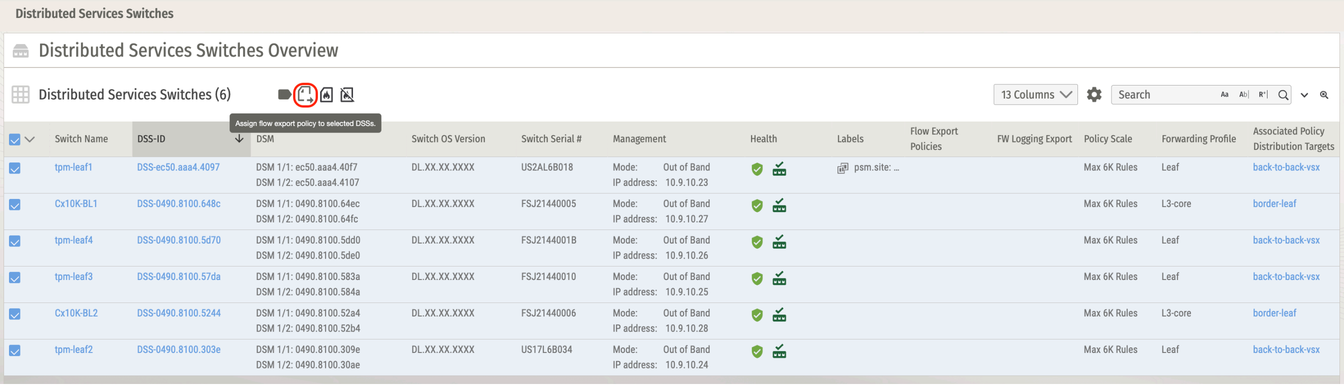
☐	Name	Interval	Targets	Propagation Status	Modification Time
☐	ipfix-demo-vrf	20s	Destination:10.10.10.1, Transport:udp/12345	Propagation Complete. Updated on 6 DSEs.	2023-07-11 18:48:30 GMT+00:00

Figure 80. Flow Export Policies

The next step involves applying this policy either at a DSS level or at a VRF level

Step 2b: Applying a flow export policy at the DSS level

When an IPFIX policy is applied at a DSS level, all VRFs in that DSS inherit that policy.



Distributed Services Switches

Distributed Services Switches Overview

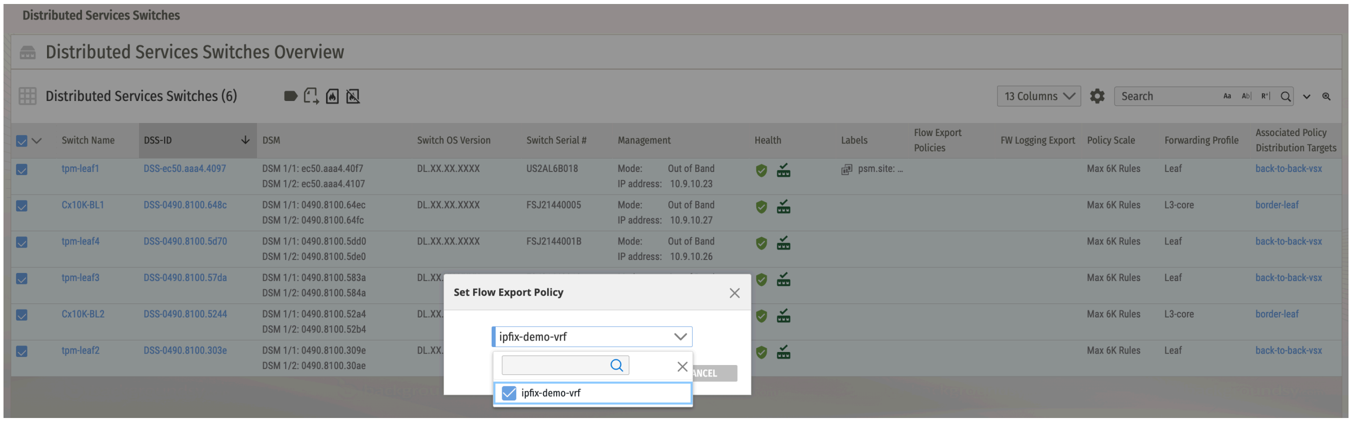
Distributed Services Switches (6)

Assign flow export policy to selected DSSs.

13 Columns | Search

☑	Switch Name	DSS-ID	DSM	Switch OS Version	Switch Serial #	Management	Health	Labels	Flow Export Policies	FW Logging Export	Policy Scale	Forwarding Profile	Associated Policy Distribution Targets
☑	tpm-leaf1	DSS-ec50.aaa4.4097	DSM 1/1: ec50.aaa4.40f7 DSM 1/2: ec50.aaa4.4107	DL.XX.XX.XXXX	US2AL6B018	Mode: Out of Band IP address: 10.9.10.23	✓	psm.site: ...			Max 6K Rules	Leaf	back-to-back-vsx
☑	Cx10K-BL1	DSS-0490.8100.648c	DSM 1/1: 0490.8100.64ec DSM 1/2: 0490.8100.64fc	DL.XX.XX.XXXX	FSJ21440005	Mode: Out of Band IP address: 10.9.10.27	✓				Max 6K Rules	L3-core	border-leaf
☑	tpm-leaf4	DSS-0490.8100.5d70	DSM 1/1: 0490.8100.5dd0 DSM 1/2: 0490.8100.5de0	DL.XX.XX.XXXX	FSJ2144001B	Mode: Out of Band IP address: 10.9.10.26	✓				Max 6K Rules	Leaf	back-to-back-vsx
☑	tpm-leaf3	DSS-0490.8100.57da	DSM 1/1: 0490.8100.583a DSM 1/2: 0490.8100.584a	DL.XX.XX.XXXX	FSJ21440010	Mode: Out of Band IP address: 10.9.10.25	✓				Max 6K Rules	Leaf	back-to-back-vsx
☑	Cx10K-BL2	DSS-0490.8100.5244	DSM 1/1: 0490.8100.52a4 DSM 1/2: 0490.8100.52b4	DL.XX.XX.XXXX	FSJ21440006	Mode: Out of Band IP address: 10.9.10.28	✓				Max 6K Rules	L3-core	border-leaf
☑	tpm-leaf2	DSS-0490.8100.303e	DSM 1/1: 0490.8100.309e DSM 1/2: 0490.8100.30ae	DL.XX.XX.XXXX	US17L6B034	Mode: Out of Band IP address: 10.9.10.24	✓				Max 6K Rules	Leaf	back-to-back-vsx

Figure 81. Click on "DSS" tab -> Select the DSS -> Click on icon "Assign flow export policy to selected DSSs"



Distributed Services Switches

Distributed Services Switches Overview

Distributed Services Switches (6)

Set Flow Export Policy

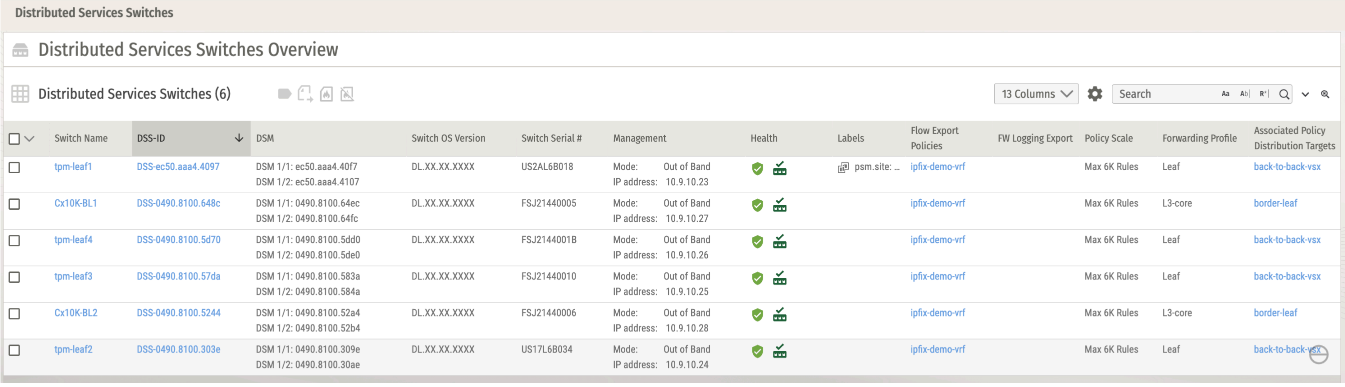
ipfix-demo-vrf

☑ ipfix-demo-vrf

ANCEL

Figure 82. Selected the required IPFIX policy from the list.

This step will apply the policy `ipfix-demo-vrf` at the DSS level.



Distributed Services Switches

Distributed Services Switches Overview

Distributed Services Switches (6)

13 Columns | Search

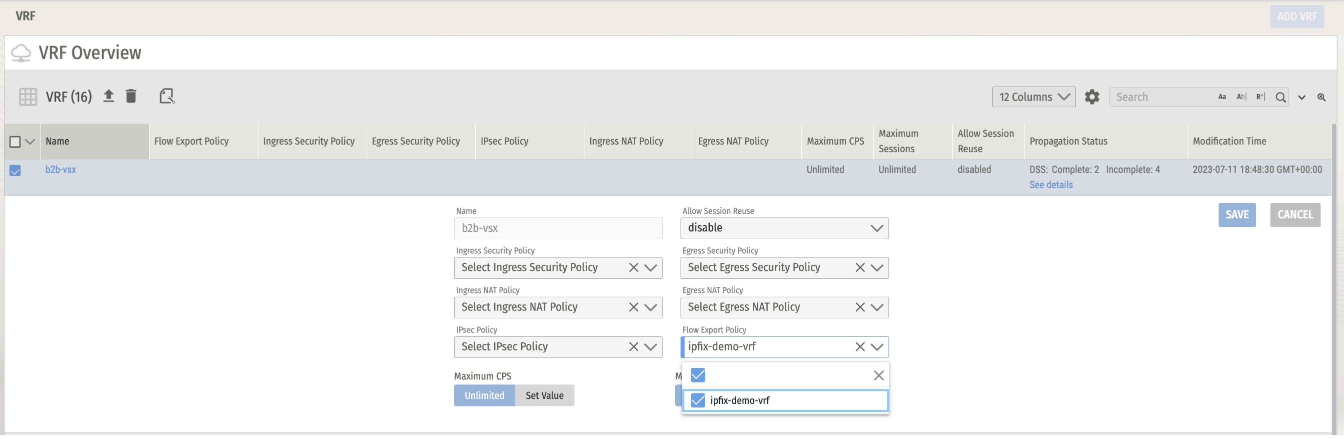
□	Switch Name	DSS-ID	DSM	Switch OS Version	Switch Serial #	Management	Health	Labels	Flow Export Policies	FW Logging Export	Policy Scale	Forwarding Profile	Associated Policy Distribution Targets
□	tpm-leaf1	DSS-ec50.aaa4.4097	DSM 1/1: ec50.aaa4.40f7 DSM 1/2: ec50.aaa4.4107	DL.XX.XX.XXXX	US2AL6B018	Mode: Out of Band IP address: 10.9.10.23	✓	psm.site: ...	ipfix-demo-vrf		Max 6K Rules	Leaf	back-to-back-vsx
□	Cx10K-BL1	DSS-0490.8100.648c	DSM 1/1: 0490.8100.64ec DSM 1/2: 0490.8100.64fc	DL.XX.XX.XXXX	FSJ21440005	Mode: Out of Band IP address: 10.9.10.27	✓		ipfix-demo-vrf		Max 6K Rules	L3-core	border-leaf
□	tpm-leaf4	DSS-0490.8100.5d70	DSM 1/1: 0490.8100.5dd0 DSM 1/2: 0490.8100.5de0	DL.XX.XX.XXXX	FSJ2144001B	Mode: Out of Band IP address: 10.9.10.26	✓		ipfix-demo-vrf		Max 6K Rules	Leaf	back-to-back-vsx
□	tpm-leaf3	DSS-0490.8100.57da	DSM 1/1: 0490.8100.583a DSM 1/2: 0490.8100.584a	DL.XX.XX.XXXX	FSJ21440010	Mode: Out of Band IP address: 10.9.10.25	✓		ipfix-demo-vrf		Max 6K Rules	Leaf	back-to-back-vsx
□	Cx10K-BL2	DSS-0490.8100.5244	DSM 1/1: 0490.8100.52a4 DSM 1/2: 0490.8100.52b4	DL.XX.XX.XXXX	FSJ21440006	Mode: Out of Band IP address: 10.9.10.28	✓		ipfix-demo-vrf		Max 6K Rules	L3-core	border-leaf
□	tpm-leaf2	DSS-0490.8100.303e	DSM 1/1: 0490.8100.309e DSM 1/2: 0490.8100.30ae	DL.XX.XX.XXXX	US17L6B034	Mode: Out of Band IP address: 10.9.10.24	✓		ipfix-demo-vrf		Max 6K Rules	Leaf	back-to-back-vsx

Figure 83.

Step 2c: Apply the defined IPFIX flow-export-policy at the VRF level

When IPFIX policy is applied at the VRF level, all DSSs containing that VRF will inherit that policy on that VRF

To apply the IPFIX policy on the VRF, click on “VRF” tab -> Select/Edit VRF settings -> Select policy from IPFIX policy drop-down menu



VRF

VRF Overview

VRF (16)

☐	Name	Flow Export Policy	Ingress Security Policy	Egress Security Policy	IPsec Policy	Ingress NAT Policy	Egress NAT Policy	Maximum CPS	Maximum Sessions	Allow Session Reuse	Propagation Status	Modification Time
☑	b2b-vsx							Unlimited	Unlimited	disabled	DSS: Complete: 2 Incomplete: 4 See details	2023-07-11 18:48:30 GMT+00:00

Name: b2b-vsx

Allow Session Reuse: disable

Ingress Security Policy: Select Ingress Security Policy

Egress Security Policy: Select Egress Security Policy

Ingress NAT Policy: Select Ingress NAT Policy

Egress NAT Policy: Select Egress NAT Policy

IPsec Policy: Select IPsec Policy

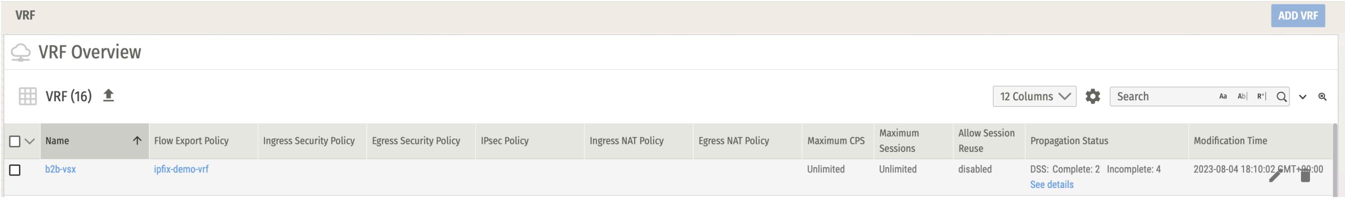
Flow Export Policy: ipfix-demo-vrf

☑ ipfix-demo-vrf

Maximum CPS: Unlimited / Set Value

Figure 84.

Once applied, the VRF page will indicate that policy is applied on that VRF:



VRF

VRF Overview

VRF (16)

☐	Name	Flow Export Policy	Ingress Security Policy	Egress Security Policy	IPsec Policy	Ingress NAT Policy	Egress NAT Policy	Maximum CPS	Maximum Sessions	Allow Session Reuse	Propagation Status	Modification Time
☐	b2b-vsx	ipfix-demo-vrf						Unlimited	Unlimited	disabled	DSS: Complete: 2 Incomplete: 4 See details	2023-08-04 18:10:02 GMT+00:00

Figure 85.

Flow records with flow information (including source and destination IP addresses, source and destination ports, and permitted and dropped packets) are exported to the external collector(s), based on the configured interval timers (the default is 10 seconds, as shown in the screenshot below). Flows that are cleaned up due to aging or session close are exported at the next export interval, hence they can linger in the flow-table from anywhere between (0 - export-interval) seconds.

Guidelines

- The export interval can be configured from a minimum of 10 seconds (default) to a maximum of 15 minutes.
- The collector target must be reachable via IPv4 in the default VRF.

    - The target can be a Virtual IP of a load balancer for a cluster of collectors, as long as it is IPv4 reachable.
    - The target cannot be in a non-default VRF.
- Each export policy can have only one collector target configured.
- If the export policy is applied at the DSS level, it then applies to all VRFs in the DSS.
- A maximum of 16 policies and 16 collectors can be configured.
    - A maximum of 2 policies and 2 collectors are allowed for any given VRF.

Considerations

1. If one collector has already been applied at the DSS level, then all VRFs in that DSS will only be able to add one additional collector.
2. If two collectors have already been applied at the DSS level, then no more collectors can be allowed on a VRF.
3. If no collectors are configured at the DSS level, each VRF can have up to 2 collectors. Users can either configure the same 2 collectors on all VRFs, or choose any 2 collectors for any VRF (to a maximum of 16 for the DSS).

IP Flow Information Export (IE) Entities (1/2)

Description	Name	Purpose / Value	Type
Permit-Packets	packetTotalCount	Provides visibility on permitted packets during flow life-span	Standard
Permit-Bytes	octetTotalCount	Provides visibility on permitted bytes during flow life-span	Standard
Drop-Packets	droppedPacketTotalCount	Provides visibility on dropped packets during flow life-span	Standard
Drop-Bytes	droppedOctetTotalCount	Provides visibility on dropped bytes during flow life-span	Standard
Delta Permit-Packets	packetDeltaCount	Provides visibility on permitted PPS during flow life-span	Standard
Delta Permit-Bytes	octetDeltaCount	Provides visibility on permitted bandwidth during flow life-span	Standard
Delta Drop-Packets	droppedPacketDeltaCount	Provides visibility on dropped PPS during flow life-span	Standard
Delta Drop-Bytes	droppedOctetDeltaCount	Provides visibility on dropped bandwidth during flow life-span	Standard
Start-Timestamp in msecs	flowStartMilliseconds	Facilitates visibility on long-lived vs short-lived flows	Standard
Last-seen Timestamp in msecs	flowLastMilliseconds	Facilitates visibility on long-lived vs short-lived flows	AMD
End-Timestamp in msecs	flowEndMilliseconds	Facilitates visibility on long-lived vs short-lived flows	Standard

IP Flow Information Export (IE) Entities (2/2)

Description	Name	Purpose / Value	Type
Flow-ID	flowId	unique flow-identifier within an observation domain	Standard
Last-seen TCP-Sequence-Number	tcpSequenceNumber	provides visibility on sender's transmit-capability in relation to tcpacknowledgement IE	Standard
Last-seen TCP-ACK-Number	tcpAcknowledgementNumber	provides visibility on recipient's receive-capability in relation to tcpsequencenumber IE	Standard
TCP-Retransmit Instances Count	tcpRetransmitCount	provides visibility on packet loss in TCP transport	AMD
TCP-Options Bitmap	tcpOptions	provides visibility on TCP transport options usage in the flow life-span	Standard
TCP-Flags	tcpControlBits	provides visibility on TCP-controls seen in the flow life-span	Standard
ICMP-Type/Code	icmpTypeCodeIPV4	provides visibility on abnormal IP-events relayed via ICMP	Standard
Dot1Q-Vlan-id	dot1qVlanId	provides visibility on the virtual-LAN where flow-traffic is seen	Standard
IPv4-Diffserv / IPv6-TrafficClass	ipClassOfService	provides visibility on the class-of-service experienced by flow-packets	Standard
Reason for Flow-state cleanup	flowEndReason	provides visibility on flow-cleanups: https://www.iana.org/assignments/ipfix/ipfix.xhtml#ipfix-flow-end-reason	Standard
Flow-Role	flowRole	provides visibility on initiator/responder flow-role	Standard
Source VPC/VRF-UUID	sourceVPCUUID	Provides Visibility on Source VPC/VRF	AMD

Destination VPC/VRF-UUID	destinationVPCUUID	Provides Visibility on Destination VPC/VRF	AMD
Source Subnet-UUID	sourceSubnetUUD	Provides Visibility on Source Subnet	AMD
Destination Subnet-UUID	destinationSubnetUUID	Provides Visibility on Destination Subnet	AMD
Vlan PCP Value	vlanPriority	Provides Visibility on VLAN-priority	Standard
Forwarding Status	forwardingStatus	Provides Visibility on Forwarding Status of packets https://www.rfc-editor.org/rfc/rfc7270.html#section-4.12	Standard

IPv4-TCP/UDP-FLOW-KEYS IEs

4B	IPv4-SIP	sourceIPV4Address
4B	IPv4-DIP	destinationIPV4Address
1B	IPv4-Protocol	protocolIdentifier
2B	L4-Source-Portnum	sourceTransportPort
2B	L4-Dest-Portnum	destinationTransportPort

IPv6-TCP/UDP-FLOW-KEYS IEs

16B	IPV6-SIP	sourceIPV6Address	Standard
16B	IPV6-DIP	destinationIPV6Address	Standard
1B	IPV6-Protocol	protocolIdentifier	Standard
2B	L4-Source-Portnum	sourceTransportPort	Standard
2B	L4-Dest-Portnum	destinationTransportPort	Standard

IPv4-ICMP-FLOW-KEYS IEs

4B	IPv4-SIP	sourceIPV4Address
4B	IPv4-DIP	destinationIPV4Address
1B	IPv4-Protocol	protocolIdentifier
1B	L4-ICMP-Type	ICMP-Type
1B	L4-ICMP-Code	ICMP-Code

IPv6-ICMP-FLOW-KEYS IEs

16B	IPV6-SIP	sourceIPV6Address	Standard
16B	IPV6-DIP	destinationIPV6Address	Standard
1B	IPV6-Protocol	protocolIdentifier	Standard
1B	L4-ICMP-Type	ICMP-Type	Standard
1B	L4-ICMP-Code	ICMP-Code	Standard

IPv4-OTHER-FLOW-KEYS IEs

4B	IPv4-SIP	sourceIPV4Address
4B	IPv4-DIP	destinationIPV4Address
1B	IPv4-Protocol	protocolIdentifier

IPv6-OTHER-FLOW-KEYS IEs

16B	IPV6-SIP	sourceIPV6Address	Standard
16B	IPV6-DIP	destinationIPV6Address	Standard
1B	IPV6-Protocol	protocolIdentifier	Standard

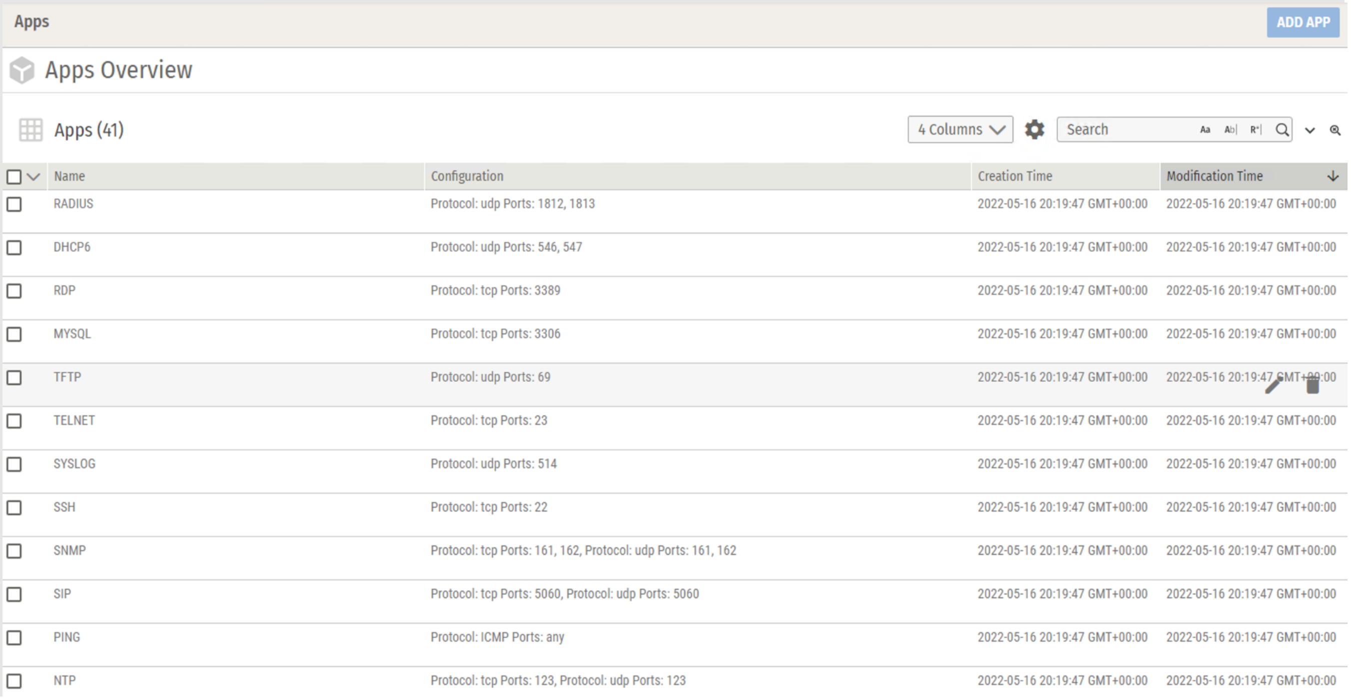
NON-IP-FLOW-KEYS IEs

6B	MAC-SA	sourceMacAddress	Standard
6B	MAC-DA	destinationMacAddress	Standard
2B	ETHER-TYPE	ethernetType	Standard

Configuring Apps

An App is a service defined either by a protocol/port pair, or by an application-level gateway (ALG). It allows a user to define commonly-deployed applications' protocol/ports, allowing the app's name to be used in policies instead of repeatedly entering the protocol/port information. App objects also allow for the specification of ALG configurations that define application-specific firewall behavior.

Supported ALGs include FTP, TFTP, RTSP, DNS, SUNRPC, and MSRPC. A set of 41 pre-defined Apps are available.



Name	Configuration	Creation Time	Modification Time
RADIUS	Protocol: udp Ports: 1812, 1813	2022-05-16 20:19:47 GMT+00:00	2022-05-16 20:19:47 GMT+00:00
DHCP6	Protocol: udp Ports: 546, 547	2022-05-16 20:19:47 GMT+00:00	2022-05-16 20:19:47 GMT+00:00
RDP	Protocol: tcp Ports: 3389	2022-05-16 20:19:47 GMT+00:00	2022-05-16 20:19:47 GMT+00:00
MYSQL	Protocol: tcp Ports: 3306	2022-05-16 20:19:47 GMT+00:00	2022-05-16 20:19:47 GMT+00:00
TFTP	Protocol: udp Ports: 69	2022-05-16 20:19:47 GMT+00:00	2022-05-16 20:19:47 GMT+00:00
TELNET	Protocol: tcp Ports: 23	2022-05-16 20:19:47 GMT+00:00	2022-05-16 20:19:47 GMT+00:00
SYSLOG	Protocol: udp Ports: 514	2022-05-16 20:19:47 GMT+00:00	2022-05-16 20:19:47 GMT+00:00
SSH	Protocol: tcp Ports: 22	2022-05-16 20:19:47 GMT+00:00	2022-05-16 20:19:47 GMT+00:00
SNMP	Protocol: tcp Ports: 161, 162, Protocol: udp Ports: 161, 162	2022-05-16 20:19:47 GMT+00:00	2022-05-16 20:19:47 GMT+00:00
SIP	Protocol: tcp Ports: 5060, Protocol: udp Ports: 5060	2022-05-16 20:19:47 GMT+00:00	2022-05-16 20:19:47 GMT+00:00
PING	Protocol: ICMP Ports: any	2022-05-16 20:19:47 GMT+00:00	2022-05-16 20:19:47 GMT+00:00
NTP	Protocol: tcp Ports: 123, Protocol: udp Ports: 123	2022-05-16 20:19:47 GMT+00:00	2022-05-16 20:19:47 GMT+00:00

Figure 86. The Apps Overview screen

Figure 86 shows some of the predefined Apps which can be used in policy configuration. Additional apps can be created manually, by navigating to `Tenants → Apps` and selecting ADD APP, as shown in Figure 87.

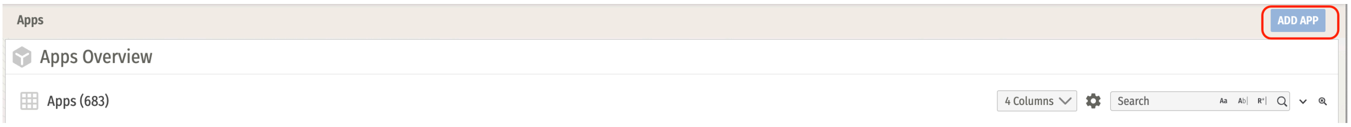
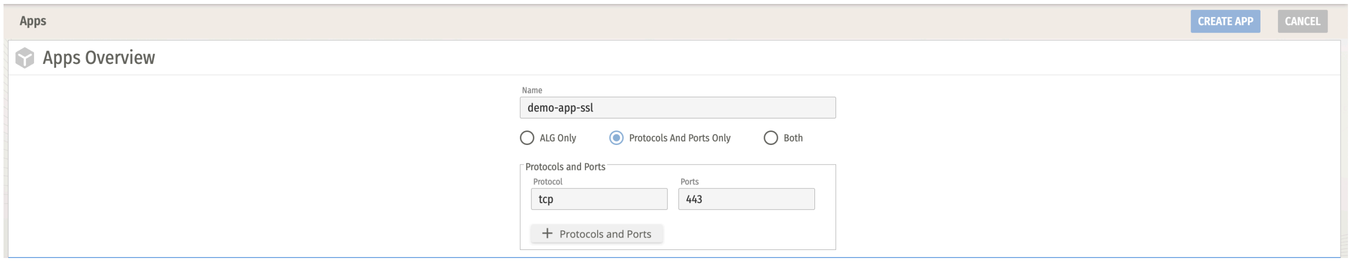


Figure 87. Apps screen detail, showing ADD button

Protocol And Ports

In this example, an App is created for the protocol/port pair of `TCP` and `443`. Click on CREATE APP when complete.



Name: demo-app-ssl

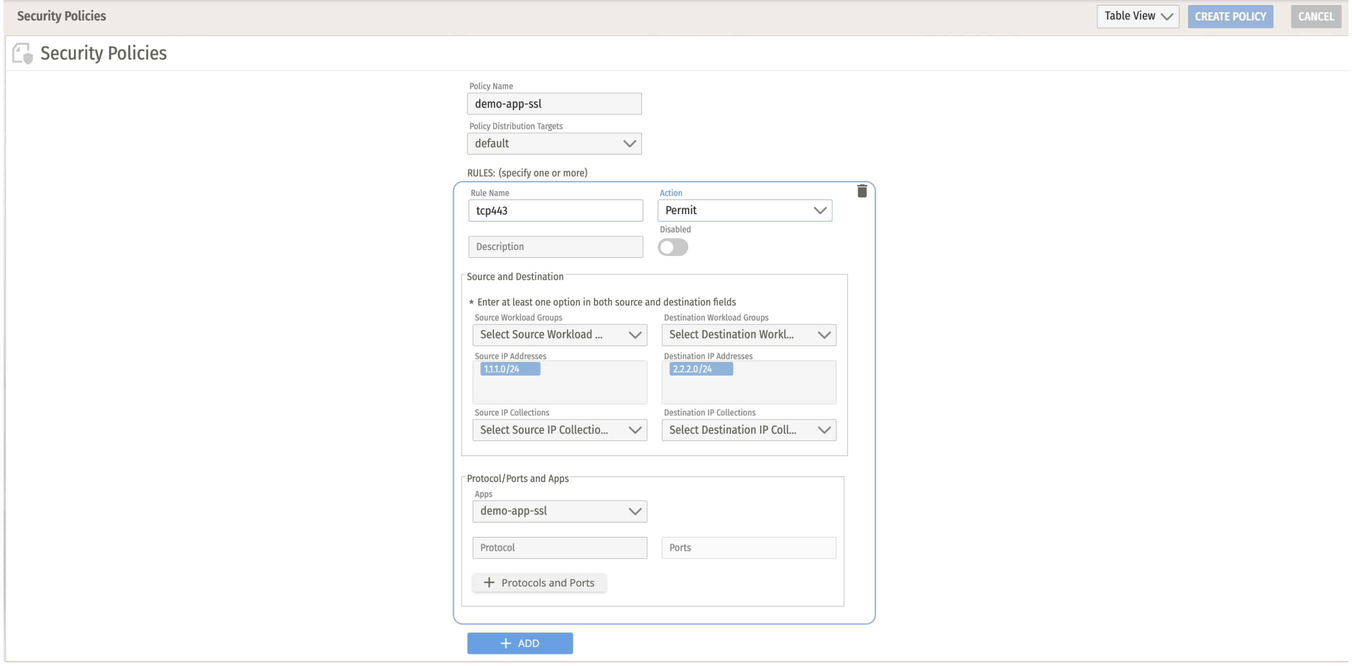
○ ALG Only ● Protocols And Ports Only ○ Both

Protocols and Ports

Protocol: tcp
Ports: 443

Figure 88. App definition window

This App can now be used with a security policy. When creating a rule within a policy, select the APPS radio button and choose the name of an App that was created previously.



Policy Name: demo-app-ssl

Policy Distribution Targets: default

RULES: (specify one or more)

Rule Name: tcp443

Action: Permit

Description

Disabled

Source and Destination

* Enter at least one option in both source and destination fields

Source Workload Groups: Select Source Workload ...

Destination Workload Groups: Select Destination Workl...

Source IP Addresses: 1.1.1.0/24

Destination IP Addresses: 2.2.2.0/24

Source IP Collections: Select Source IP Collectio...

Destination IP Collections: Select Destination IP Coll...

Protocol/Ports and Apps

Apps: demo-app-ssl

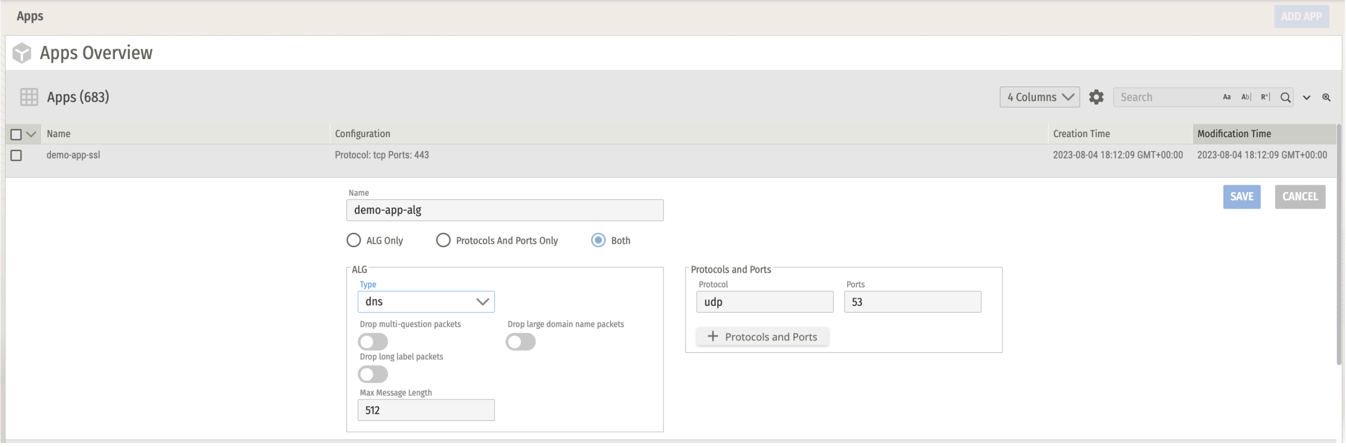
Protocol

Ports

Figure 89. Adding an App to a security policy

ALG

In this example, an App is created for DNS ALG.



Name	Configuration	Creation Time	Modification Time
demo-app-ssl	Protocol: tcp Ports: 443	2023-08-04 18:12:09 GMT+00:00	2023-08-04 18:12:09 GMT+00:00

Name: demo-app-alg

○ ALG Only ○ Protocols And Ports Only ● Both

ALG

Type: dns

Drop multi-question packets

Drop large domain name packets

Drop long label packets

Max Message Length: 512

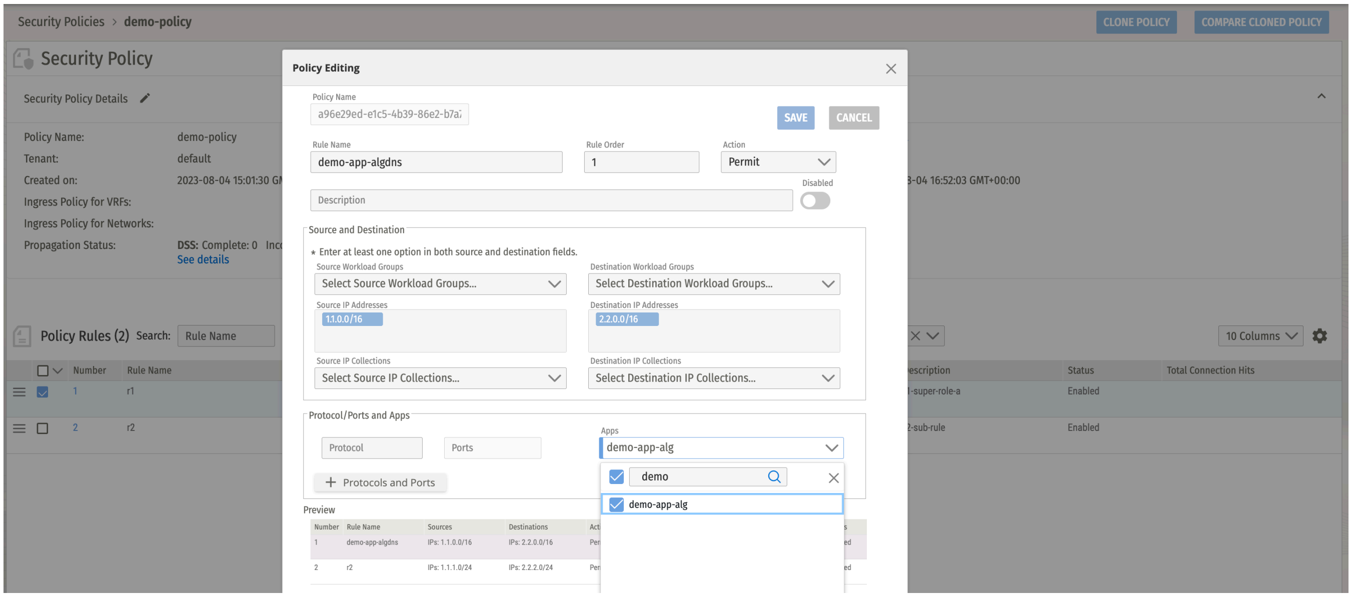
Protocols and Ports

Protocol: udp

Ports: 53

Figure 90. Creating an ALG App (compare to Figure 84)

This App can now be used with a security policy. When creating a rule within a policy, select the APPS radio button and select the name of the App that was created previously.



Security Policies > demo-policy

Security Policy

Security Policy Details

Policy Name: demo-policy
Tenant: default
Created on: 2023-08-04 15:01:30 GM
Propagation Status: DSS: Complete: 0 Inc
See details

Policy Rules (2) Search: Rule Name

Policy Editing

Policy Name: a96e29ed-e1c5-4b39-86e2-b7a
Rule Name: demo-app-algdns
Rule Order: 1
Action: Permit

Source and Destination

* Enter at least one option in both source and destination fields.

Source IP Addresses: 1.1.0.0/16
Destination IP Addresses: 2.2.0.0/16

Protocol/Ports and Apps

Apps: demo-app-alg

☑ demo
☑ demo-app-alg

Preview

Number	Rule Name	Sources	Destinations
1	demo-app-algdns	IPs: 1.1.0.0/16	IPs: 2.2.0.0/16
2	r2	IPs: 1.1.1.0/24	IPs: 2.2.2.0/24

Figure 91. Adding the DNS ALG app to the policy

DDoS Detection and Alerting: Maximum Sessions and CPS Limits

Two limits can be configured for distributed denial-of-service (DDoS) protection: *maximum session limit* and *maximum connections per second* (*CPS*). The PSM is used to configure these limits, and the DSM enforces these limits in the data plane.

Configuring these limits can help to detect unusual network activities which might indicate a DDoS attack underway in the network, and mitigate their effect. When the session limit is enforced and reached, the switch will stop learning any new sessions, preventing the switch's flow table from being exhausted by a DDoS attack. When CPS limits are enforced, a high-rate DDoS attack (SYN attack) will be prevented from overwhelming the switch with new connection requests. All new connections will be dropped when a CPS limit is reached.

These are not per-protocol (TCP/UDP/ICMP) limits; rather, they can be either applied at VRF level or at a network level or both. This section explains details about these limits, including how and where they can be enforced and used.

Maximum Session Limit

In the CX 10000 stateful firewall implementation, a *session* is defined as a pair of flows: an *I-flow* (initiator flow) and an *R-flow* (responder flow). Note that session and CPS limits apply to total session count and not flow count. This limit can be either applied at VRF level, or at network level, or both. When applied at the VRF level, all networks that are part of that VRF will inherit the configuration. The maximum session limit is not enforced based on strict values; there is a leeway of about 512 sessions between the configured value and the enforced value.

Maximum CPS

CPS (connections per second) is essentially the rate of incoming new connections that can be successfully processed and installed in the switch's flow table. A DDoS attack with a high rate of incoming SYN packets can potentially exhaust the switch's DSMs and cause it to start dropping valid connections, which is highly undesirable.

Configuring maximum CPS enables a limit to throttle the maximum number of new connections that can be set up. From the data plane perspective this limit is enforced by the DSM's P4 engine with the help of a policer that is applied at a per-VLAN level. The PSM user has options to set this limit as per-VRF, per-network or both.

In general, the best practice is to set general VRF-wide (applied to all networks in VRF) max session-limits and max CPS to values slightly greater than the general operating load in steady-state production. Specific network-level CPS and session limits can be applied to only those networks where more specific values (fine-tuned) are required.

Table 10 describes the behavior of max session limit and max CPS with different configurations. (UI) in parentheses refers to what is seen in the UI, and (API) in parentheses refers to the value sent in the REST payload.

VRF-Level	Network-Level	Expected Behavior
Not-configured (API)	(API) - Not configured	No CPS/max-session limits enforced. This is the default behavior.
Not-configured (API)	Inherit from VRF (UI) "-1" (API)	No CPS/max-session limits enforced
Not-configured (API)	Unlimited (UI) "0" (API)	No CPS/max-session limits enforced
Not-configured (API)	Set Value (UI) "x" sessions "y" CPS	Limit to max of "x" sessions and "y" CPS rate for the given network
Configured "x" sessions "y" CPS	Not-configured (API)	Not-configured defaults to inherit. The configured value of "x" sessions and "y" CPS will be enforced.

Configured "x" sessions "y" CPS	Inherit from VRF(UI) "-1" (API)	The network in the VRF inherits the configured value of the max session limit/rate under the VRF. User will get max of "x" sessions and "y" CPS for each network in VRF
Configured "x" sessions "y" CPS	Unlimited (UI) "0" (API)	No CPS/max-session limits enforced. ***NOTE**: The configured value of "x" sessions and "y" CPS **will not** be enforced.*
Configured "x1" sessions "y1" CPS	Set Value (UI) "x2" sessions "y2" CPS	Maximum session limit of "x2" sessions and "y2" CPS gets enforced for the given network

Table 10. Expected results for different options for VRF and Network-level configs

Min and Max Values

The minimum value of the maximum session limit field is 10,000. The maximum value of this field is 5M (5,000,000), well beyond the supported limit of max sessions per DSM, which is 2M sessions.

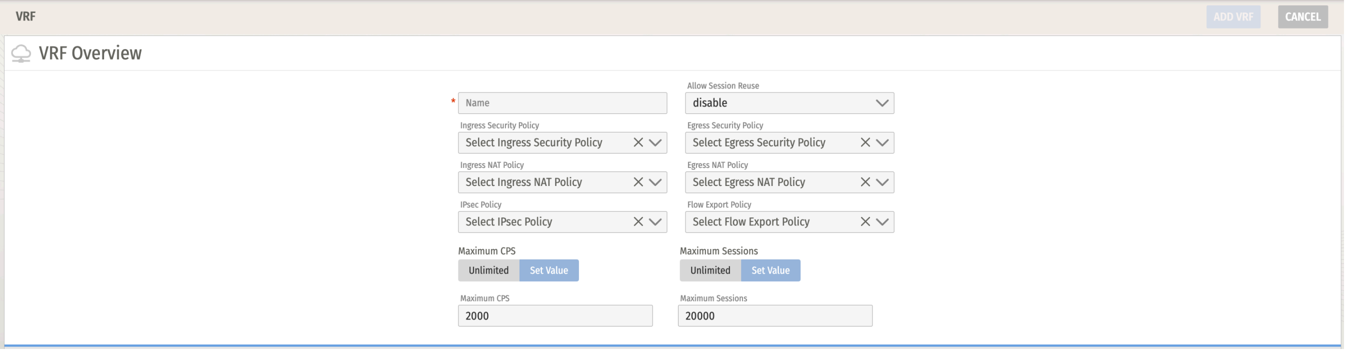
The minimum value for the max CPS limit field is 1,000. The maximum value of this field is 1,000,000 (1M), well beyond the supported limit of 800K CPS (single-switch / non-VSX).

The following section describes and gives examples of how the maximum sessions and CPS features can be configured via the PSM UI.

Configuring the Maximum Sessions / CPS on a VRF via the PSM UI

For new VRFs:

1. Select Tenants -> VRF from side menu
2. Select ADD VRF
3. Enter a name for the VRF
4. Select Set Value (under Maximum CPS), and enter value
5. Select Set Value (under Maximum Sessions), and enter a value



* Name

Allow Session Reuse: disable

Ingress Security Policy: Select Ingress Security Policy

Egress Security Policy: Select Egress Security Policy

Ingress NAT Policy: Select Ingress NAT Policy

Egress NAT Policy: Select Egress NAT Policy

IPsec Policy: Select IPsec Policy

Flow Export Policy: Select Flow Export Policy

Maximum CPS: Unlimited | Set Value

Maximum Sessions: Unlimited | Set Value

Maximum CPS: 2000

Maximum Sessions: 20000

Figure 92. Adding a VRF

For existing VRFs:

1. Select Tenants -> VRF from side menu
2. Hover over the VRF you wish to modify, and click on the Update VRF (pencil) icon
3. Update the Maximum CPS value
4. Update the Maximum Sessions value

Configuring the Maximum Sessions/CPS on a Network via the PSM UI

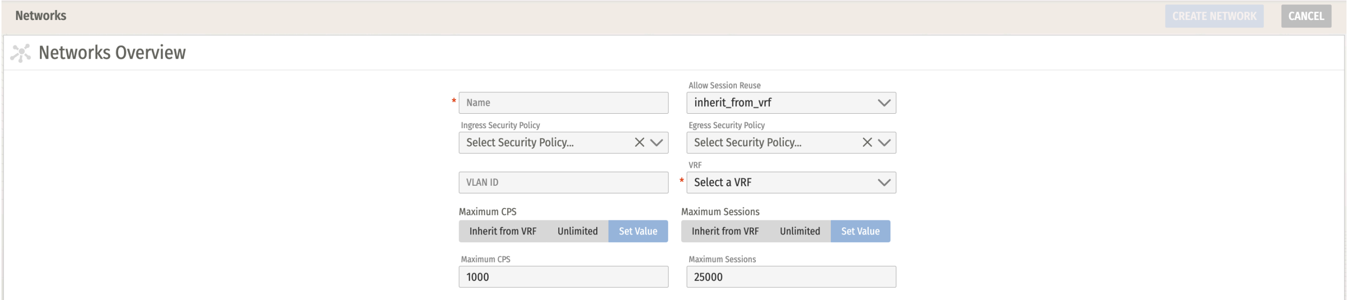
For New Networks:

1. Select Network from panel (left)
2. Add Network
3. Select from one of the options (“Inherit from VRF” or “Unlimited” or “Set Value”)
4. Create Network

For Existing Networks:

1. Select Network from panel (left)
2. Select Network
3. Update Network
4. Select from one of the options (“Inherit from VRF” or “Unlimited” or “Set Value”)
5. Save

When “Inherit from VRF” is used, the Maximum CPS and Maximum Sessions will inherit the value from the value configured in the corresponding VRF. When Unlimited is used, no limits are enforced for Maximum sessions/ Max CPS. When “Set value” is used, the configured values will be applied and will also override the values configured at the corresponding VRF level.



Networks Overview

* Name
Allow Session Reuse: inherit_from_vrf
Ingress Security Policy: Select Security Policy...
Egress Security Policy: Select Security Policy...
VLAN ID
VRF: * Select a VRF
Maximum CPS: Inherit from VRF | Unlimited | Set Value
Maximum Sessions: Inherit from VRF | Unlimited | Set Value
Maximum CPS: 1000
Maximum Sessions: 25000

Figure 93. Configuring session and CPS limits

API Examples

An example REST API call to set maximum session limit/CPS for a given VRF:

PUT: `https://$PSMaddr/configs/network/v1/tenant/default/virtualrouters/Test-VRF-1`

Body:

```
{
    "kind": "VirtualRouter",
    "api-version": "v1",
    "meta": {
        "name": "Test-VRF-1",
        "tenant": "default"
    },
    "spec": {
        "maximum-sessions-per-network-per-distributed-services-entity": 400000,
        "maximum-cps-per-network-per-distributed-services-entity": 40000
    }
}
```

In the above example, a max session-limit of 400000 and max-CPS of 40000 is enforced on VRF `Test-VRF-1`.

An example REST API call to set maximum session limit/CPS for a given network:

PUT: `https://$PSMaddr/configs/network/v1/tenant/default/networks/Test-Subnet-1`

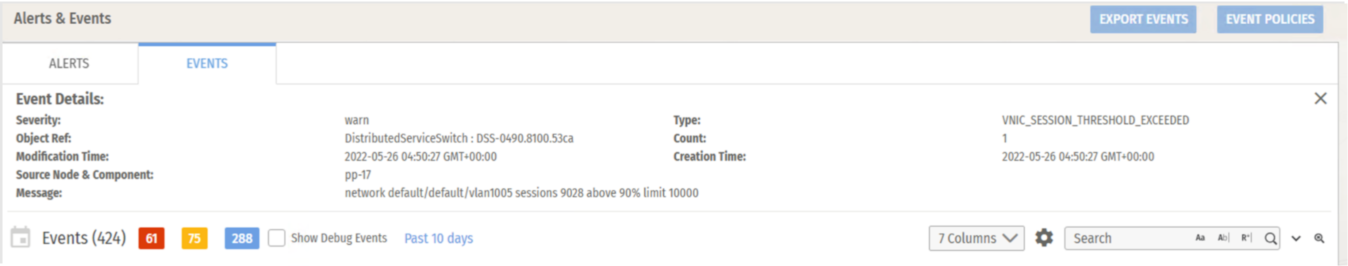
Body:

```
{
    "kind": "Network",
    "api-version": "v1",
    "meta": {
        "name": "Test-Subnet-1",
        "tenant": "default"
    },
    "spec": {
        "vlan-id": 5,
        "virtual-router": "Test-VRF-1",
"firewall-profile": {
            "maximum-cps-per-distributed-services-entity": 10000,
            "maximum-sessions-per-distributed-services-entity": 25000
        }
    }
}
```

In the above example, a per-network limit of 25000 sessions and 10000 CPS is set.Events/Alerts Related To Maximum Configured Session-Limit

- As the flow table is populated, a warning event will be generated when 90% of the configured max limit is reached.
- A critical Alert will be generated for a VRF/network when it reaches 100% of its configured limit.
- When the session count drops below 80% of the configured limit, an INFO event is generated.

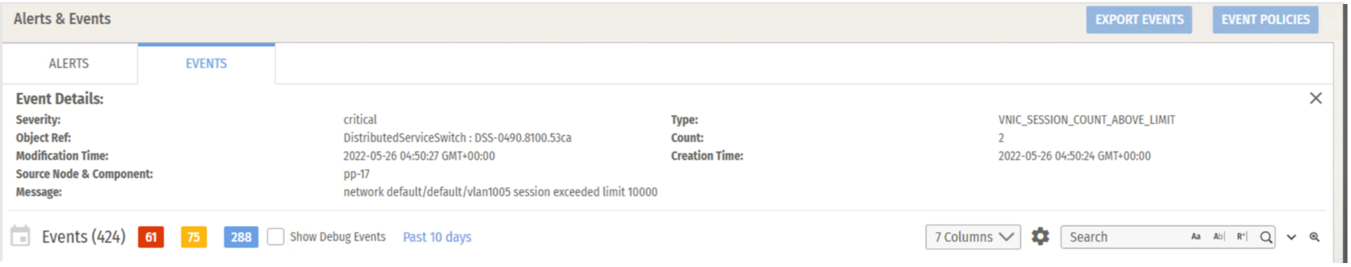
Examples:



Event Details:

Severity: warn
Object Ref: DistributedServiceSwitch : DSS-0490.8100.53ca
Modification Time: 2022-05-26 04:50:27 GMT+00:00
Source Node & Component: pp-17
Message: network default/default/vlan1005 sessions 9028 above 90% limit 10000
Type: VNIC_SESSION_THRESHOLD_EXCEEDED
Count: 1
Creation Time: 2022-05-26 04:50:27 GMT+00:00

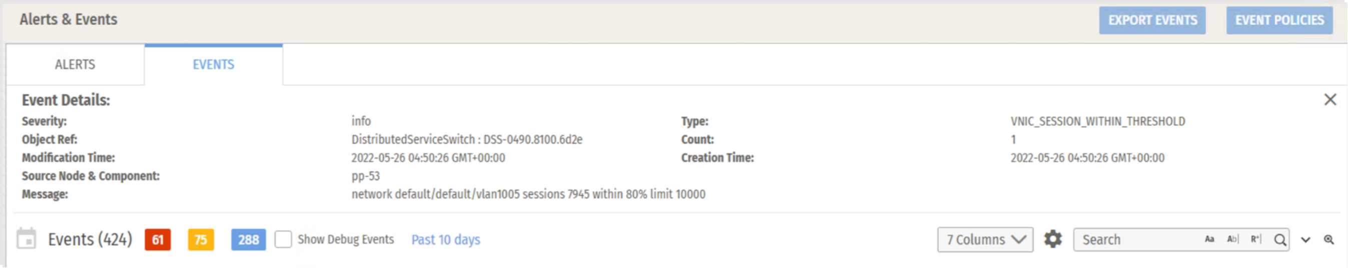
Figure 94. A warning event generated when 90% of the configured limit is reached



Event Details:

Severity: critical
Object Ref: DistributedServiceSwitch : DSS-0490.8100.53ca
Modification Time: 2022-05-26 04:50:27 GMT+00:00
Source Node & Component: pp-17
Message: network default/default/vlan1005 session exceeded limit 10000
Type: VNIC_SESSION_COUNT_ABOVE_LIMIT
Count: 2
Creation Time: 2022-05-26 04:50:24 GMT+00:00

Figure 95. A critical event generated when a limit is reached



Event Details:

Severity: info
Object Ref: DistributedServiceSwitch : DSS-0490.8100.6d2e
Modification Time: 2022-05-26 04:50:26 GMT+00:00
Source Node & Component: pp-53
Message: network default/default/vlan1005 sessions 7945 within 80% limit 10000
Type: VNIC_SESSION_WITHIN_THRESHOLD
Count: 1
Creation Time: 2022-05-26 04:50:26 GMT+00:00

Figure 96. An INFO event generated when the count returns to below 80%.

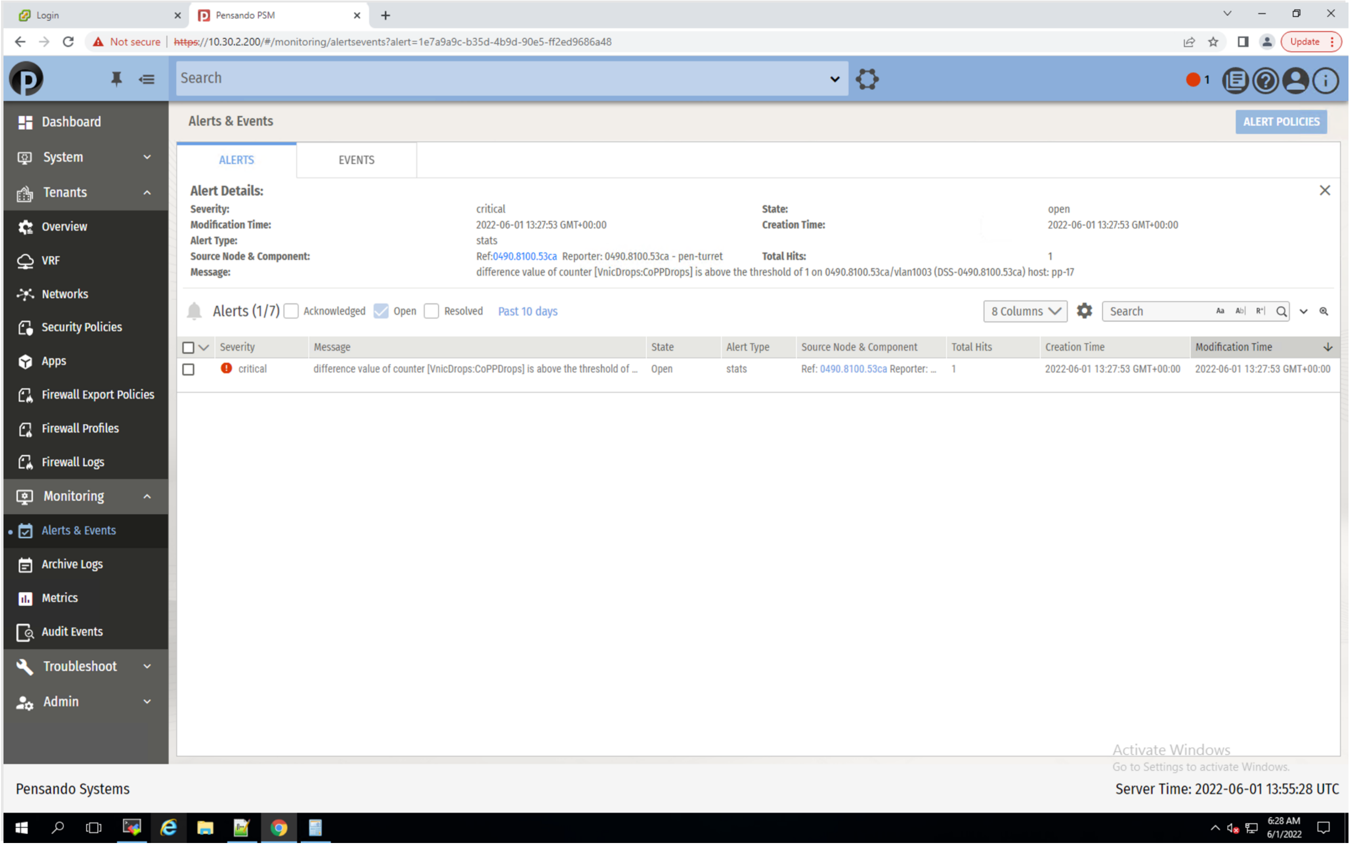
Behavior on Reaching the Maximum Session Limit

- Flow learning (learning of new flows) will not be stopped at exactly 100% of the configured limit. An excess of 5% is allowed, so flow learning is completely stopped at 105% of the configured limit value and a critical alert will be reported to the PSM.

Behavior on Reaching the Maximum CPS Limit

When the incoming CPS/flow-create/SYN rate is higher than the configured rate, then incoming connections will start getting dropped (or policed) at the rate defined by the configuration. This policing action is done by the P4 engine itself. However, the event that gets generated will get generated only on exceeding the configured limit. No warning events for 90% of configured CPS are reported to the PSM. Also, note that it could take up to 2 minutes for the MAX CPS exceed event to get reported to the PSM.

Example of event reported at the PSM:



Alerts & Events

Alert Details:

Severity:	critical	**State:**	open
Modification Time:	2022-06-01 13:27:53 GMT+00:00	**Creation Time:**	2022-06-01 13:27:53 GMT+00:00
Alert Type:	stats		
Source Node & Component:	Ref:0490.8100.53ca Reporter: 0490.8100.53ca - pen-turret	**Total Hits:**	1
Message:	difference value of counter [VnicDrops:CoPPDrops] is above the threshold of 1 on 0490.8100.53ca/vlan1003 (DSS-0490.8100.53ca) host: pp-17		

Alerts (1/7) □ Acknowledged ☑ Open □ Resolved Past 10 days

□	Severity	Message	State	Alert Type	Source Node & Component	Total Hits	Creation Time	Modification Time
□	critical	difference value of counter [VnicDrops:CoPPDrops] is above the threshold of ...	Open	stats	Ref: 0490.8100.53ca Reporter: ...	1	2022-06-01 13:27:53 GMT+00:00	2022-06-01 13:27:53 GMT+00:00

Server Time: 2022-06-01 13:55:28 UTC

Figure 97. Example of an event generated upon reaching MAX CPS limit

In addition to checking for above alert, PSM users can create a metric chart and look for CoPP drops (see Figure 98):

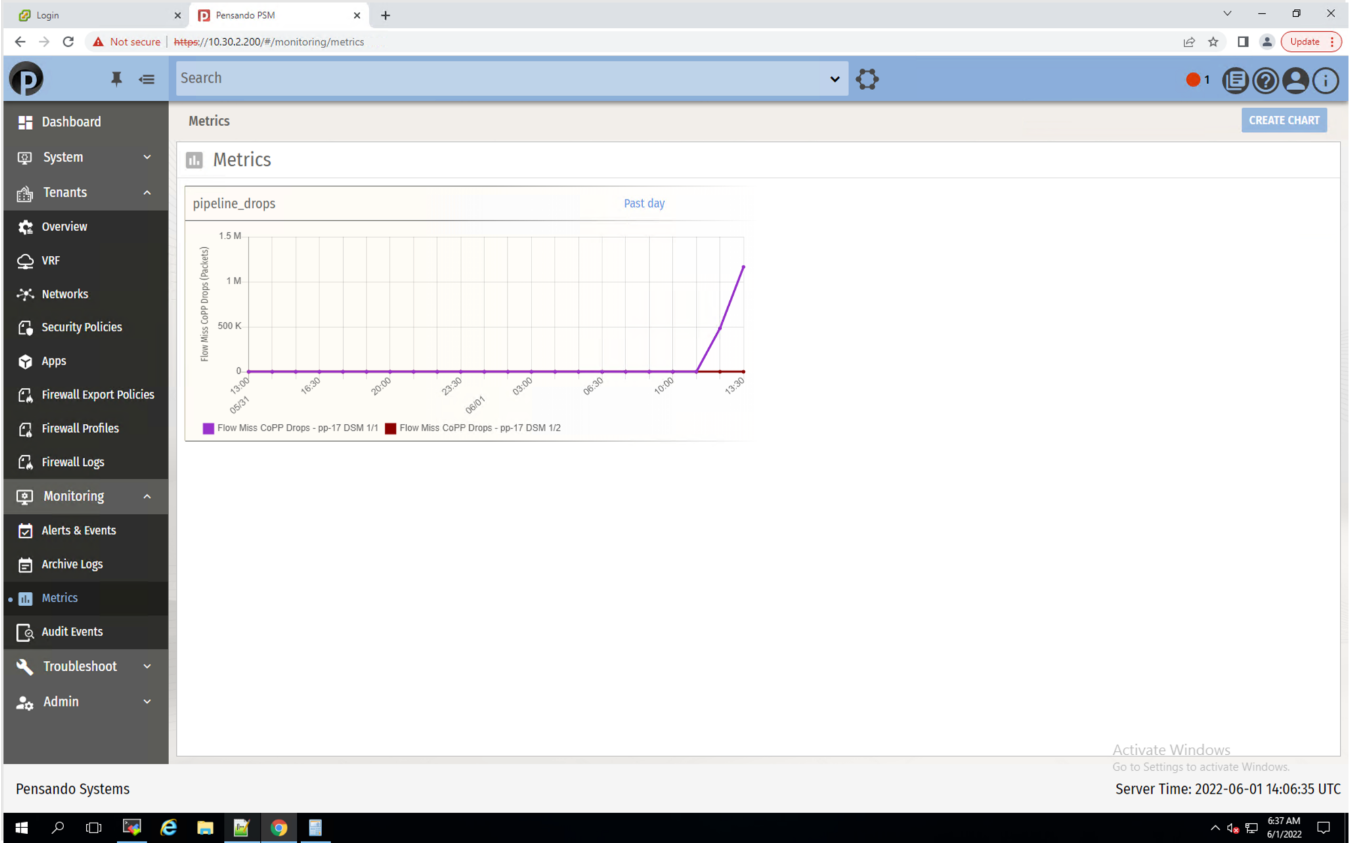


Figure 94. Example of Flow Miss CoPP Drops graph; Max CPS exceeded

Implication of Configuring `session-limit` on an Active System

Configuring the `session-limit` value on a live system (a switch with active flows) for any given VRF/network does not have any traffic impact or disruption of existing active flows.

Active flows are not purged if the new configured value is less than the total number of active sessions for that VRF/network. The change or the new limit will only apply to new flows created after the last config change.

VSX Implications

Important considerations in VSX topology with regard to this configuration:

- Each switch in VSX independently applies the configuration at a per-VRF /network level. The PSM will push identical VRF/network configurations to both switches in VSX, so configuration consistency is expected and independently enforced without the need to manually coordinate between peers.

- Both flows locally learned by VSX switch and flows learned through the process of flow-sync from VSX peer are accounted towards the calculation of max sessions before enforcing the max session-limit.

  Consider the following examples:

  CASE 1: A and B in VSX pair and switch A gets all the new flows, no flow hashes to switch B.

  Configured limit is 10000 sessions on network/VRF.

  Switch A can only learn a max of 10000 sessions. It will sync 10000 sessions to B. Both A and B are at max session limit of 10000 and cannot accept any more flows.

  CASE 2: A and B are in VSX pair and there is equal distribution of flows from host to switch

  Configured limit is 10,000 sessions on network/VRF.

  In this case, each switch Switch A and Switch B can learn only up to 5K new flows from hosts since the remaining 5k flows are learnt from flow-sync from the respective VSX peer.

- The max-CPS limit is independently applied on each switch in a VSX pair. This is illustrated in the following examples:

  CASE 1: A and B in VSX pair and switch A gets all the new flows, no flow hashes to switch B.

  Application sends traffic at a rate of 20,000 CPS. A limit of 5,000 CPS is applied on both switches A and B.

  All traffic hashes to switch A and gets throttled to 5KCPS. This 5KCPS will be flow-synced to switch B without further throttling on switch B.

  Effective CPS achieved: 5K CPS.

  CASE 2: A and B are in VSX pair, flows hash evenly to switch A and switch B.

  Application sends traffic at a rate of 20K CPS. A limit of 5K CPS is applied to both switches A and B.

  The 10K CPS traffic that hashes to A will get throttled to 5K CPS and the 10K CPS traffic that gets hashed to switch B will also get throttled to 5K CPS. Each switch will flow sync its 5K CPS to the other peer. No further throttling takes place when it is synced to peer.

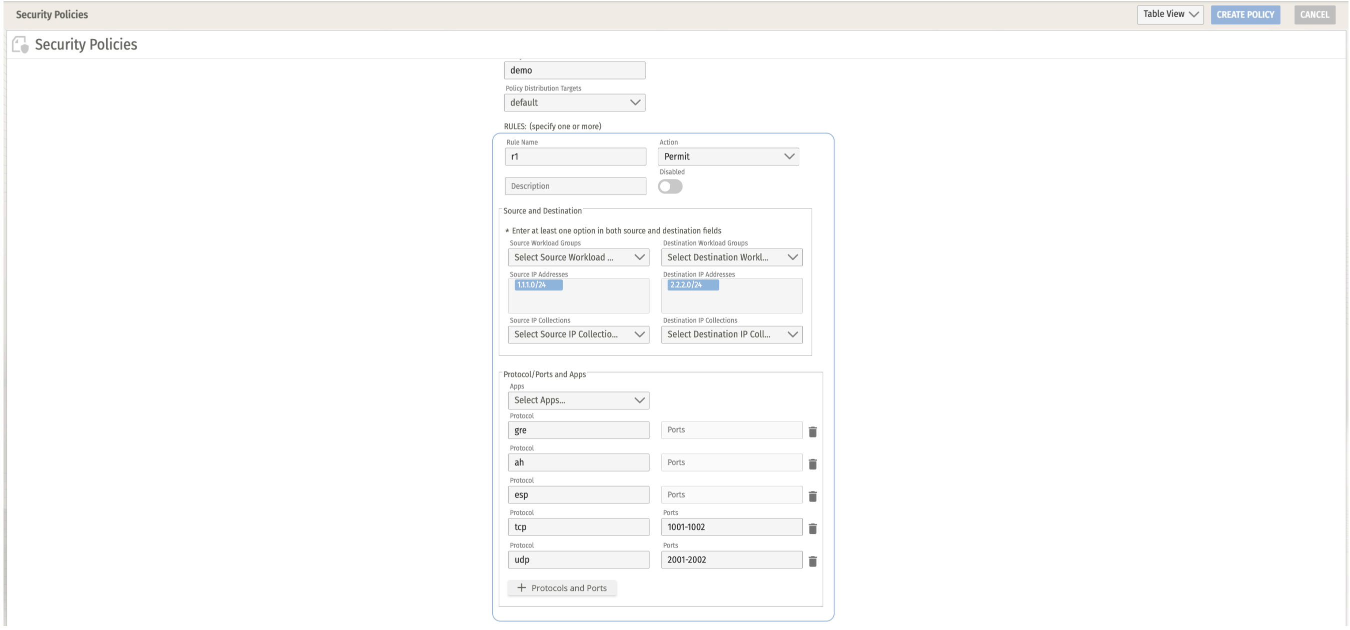
Effective CPS achieved: 10K CPS.

Multiple ALG Types/Apps/Protocols in Firewall Policy Rules

It is possible to match on multiple ALG types / multiple apps / multiple IP protocol types (TCP/UDP/ICMP/GRE/AH/ESP) within a single rule inside a PSM policy. One or more of the defined ALG types/apps/IP protocols can be used. This section provides examples of defining such a rule via the PSM UI, as well as API examples.

Steps To Configure Rules With Multiple Proto-Ports Via the PSM UI

1. Security Policies (left panel)
2. Add Security Policy
3. Select Proto-Ports



demo

Policy Distribution Targets: default

RULES: (specify one or more)

Rule Name: r1

Action: Permit

Description

Disabled

Source and Destination

* Enter at least one option in both source and destination fields

Source Workload Groups: Select Source Workload ...

Destination Workload Groups: Select Destination Workl...

Source IP Addresses: 1.1.1.0/24

Destination IP Addresses: 2.2.2.0/24

Source IP Collections: Select Source IP Collectio...

Destination IP Collections: Select Destination IP Coll...

Protocol/Ports and Apps

Apps: Select Apps...

Protocol: gre — Ports

Protocol: ah — Ports

Protocol: esp — Ports

Protocol: tcp — Ports: 1001-1002

Protocol: udp — Ports: 2001-2002

+ Protocols and Ports

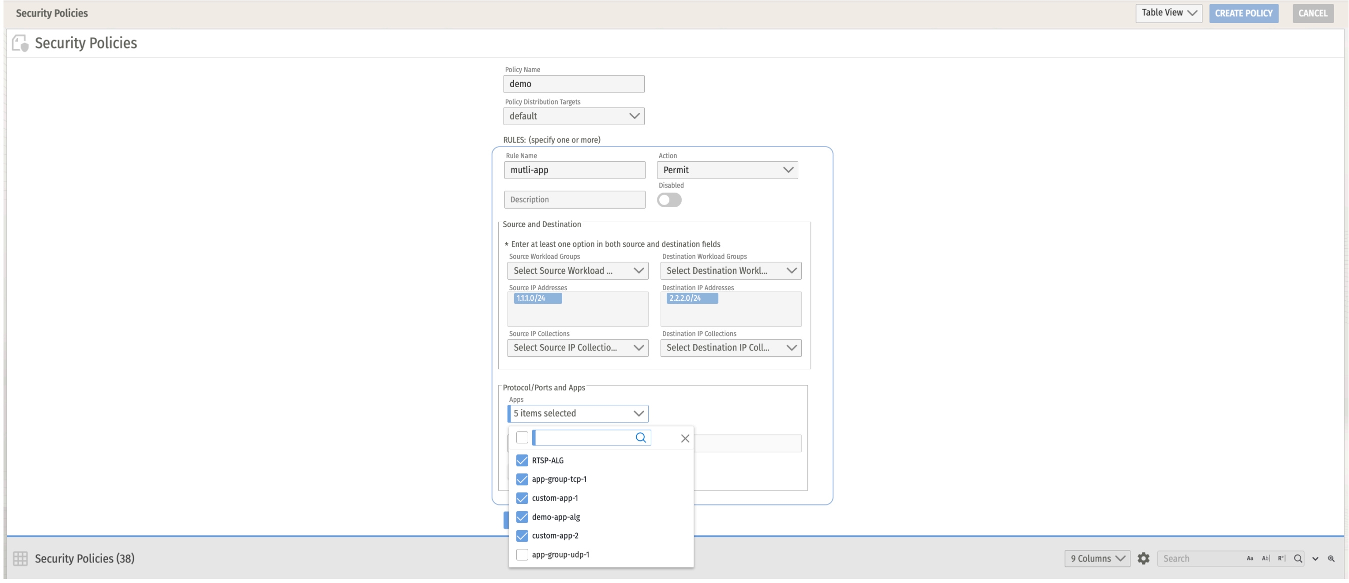
Figure 99. UI example of rules with multiple proto-ports

As shown above, a mix of TCP ports, ID ports and protocol types of ICMP, GRE, IPsec (AH/ESP) can be referenced within a single rule.

A rule can also reference multiple ALG types, as shown below:

Steps to configure rules with multiple ALG types via the PSM UI:

1. Security Policies (left panel)
2. Add Security Policy
3. Select Apps
4. From drop-down menu choose multiple defined ALG types from list



Policy Name: demo

Policy Distribution Targets: default

RULES: (specify one or more)

Rule Name: mutli-app

Action: Permit

Disabled: ○

Description

Source and Destination

* Enter at least one option in both source and destination fields

Source Workload Groups: Select Source Workload ...

Destination Workload Groups: Select Destination Workl...

Source IP Addresses: 1.1.1.0/24

Destination IP Addresses: 2.2.2.0/24

Source IP Collections: Select Source IP Collectio...

Destination IP Collections: Select Destination IP Coll...

Protocol/Ports and Apps

Apps: 5 items selected

- ☑ RTSP-ALG
- ☑ app-group-tcp-1
- ☑ custom-app-1
- ☑ demo-app-alg
- ☑ custom-app-2
- ☐ app-group-udp-1

Figure 100. UI example of rules with multiple ALG types

***NOTE:** A policy rule can either reference multiple proto-ports (list of TCP and list of UDP ports) or multiple apps/ALG types but NOT a mix of both proto-ports and app/ALG types*

API Examples:

This section provides API examples to create a policy with rules referencing either multiple protocol ports or multiple apps.

Example:

POST: `https://$PSMADDR/configs/security/v1/tenant/default/networksecuritypolicies`

Request body:

```
{
    "kind": "NetworkSecurityPolicy",
    "api-version": "v1",
    "meta": {
        "name": "rule-with-multiple-protocols",
        "tenant": "default"
    },
    "spec": {
        "attach-tenant": true,
        "rules": [
            {
                "proto-ports": [
                    {
                        "protocol": "tcp",
                        "ports": "1001-1002,2001,3001"
                    },
                    {
                        "protocol": "udp",
                        "ports": "2001-2002,3003,4004"
                    },
                    {
                        "protocol": "icmp"
                    },
                    {
                        "protocol": "gre"
                    },
                    {
                        "protocol": "ah"
                    },
                    {
                        "protocol": "esp"
                    }
                ],
                "action": "permit",
                "from-ip-addresses": [
                    "1.1.1.1/32"
                ],
                "to-ip-addresses": [
                    "2.2.2.2/32"
                ]
            }
        ]
    }
}
```

API Example: Policy with Rule Referencing Multiple ALG Types:

```
{
    "kind": "NetworkSecurityPolicy",
    "api-version": "v1",
    "meta": {
        "name": "rule-with-multiple-alg-types",
        "tenant": "default"
    },
    "spec": {
        "attach-tenant": true,
        "rules": [
            {
                "apps": [
                    "FTP",
                    "DNS"
                ],
                "action": "permit",
                "from-ip-addresses": [
                    "1.1.1.1/32"
                ],
                "to-ip-addresses": [
                    "2.2.2.2/32"
                ]
            }
        ]
    }
}
```

IP Protocols Support for Firewall Policy

Any IP protocol can be matched inside of a firewall policy rule. The rule can either match on reserved keywords like “gre”, “ah” and “esp” for GRE, IPsec-AH and IPsec-ESP protocols respectively or it can match on any valid numeric protocol number in the range of 0-254.

The following section gives examples for doing this using the PSM UI as well as the REST API.

UI Examples

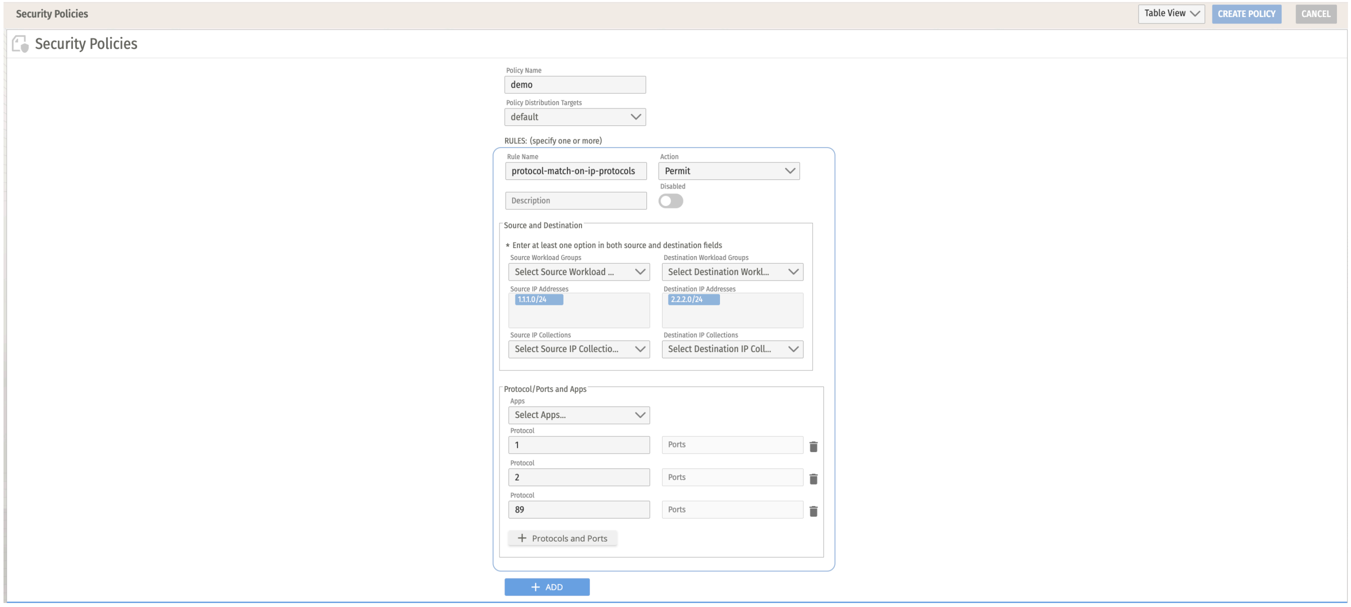

Security Policies

Table View | CREATE POLICY | CANCEL

Security Policies

Policy Name: demo

Policy Distribution Targets: default

RULES: (specify one or more)

Rule Name: protocol-match-on-ip-protocols

Action: Permit

Disabled

Description

Source and Destination

* Enter at least one option in both source and destination fields

Source Workload Groups: Select Source Workload ...

Destination Workload Groups: Select Destination Workl...

Source IP Addresses: 1.1.1.0/24

Destination IP Addresses: 2.2.2.0/24

Source IP Collections: Select Source IP Collectio...

Destination IP Collections: Select Destination IP Coll...

Protocol/Ports and Apps

Apps: Select Apps...

Protocol: 1 — Ports

Protocol: 2 — Ports

Protocol: 89 — Ports

+ Protocols and Ports

+ ADD

Figure 101. UI example of rules with multiple IP protocol numbers

Figure 101 shows an example of a policy created with a rule referencing a combination of multiple IP protocol numbers.

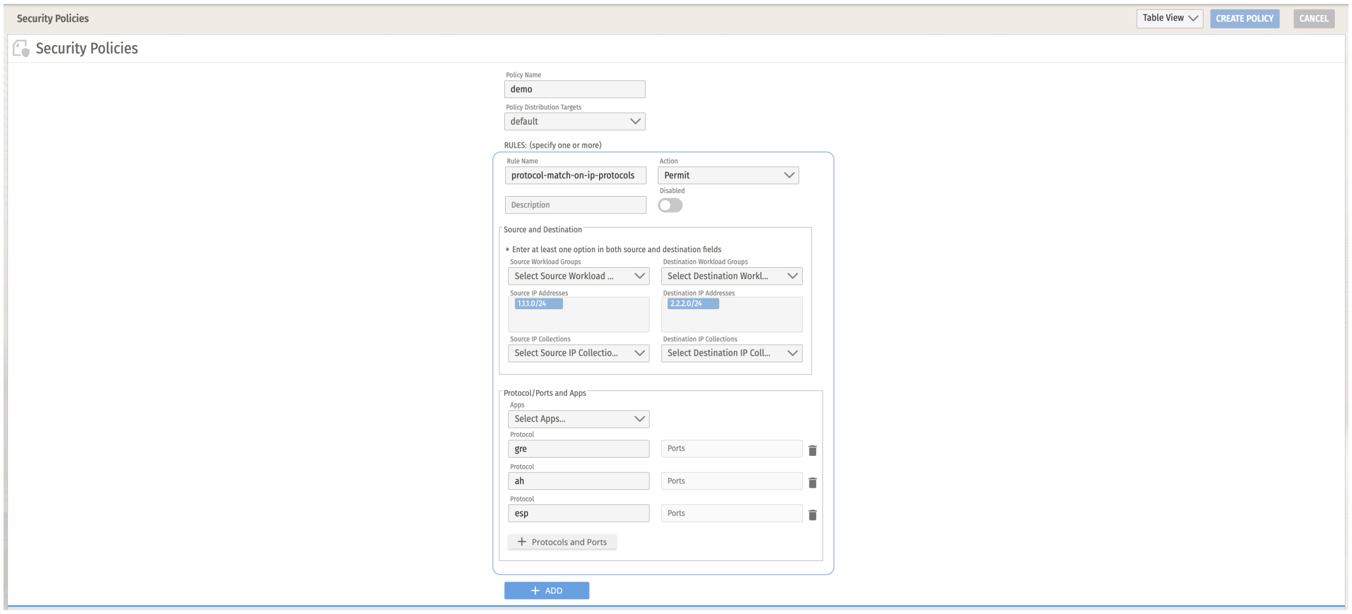

Security Policies

Table View | CREATE POLICY | CANCEL

Security Policies

Policy Name: demo

Policy Distribution Targets: default

RULES: (specify one or more)

Rule Name: protocol-match-on-ip-protocols

Action: Permit

Disabled

Description

Source and Destination

* Enter at least one option in both source and destination fields

Source Workload Groups: Select Source Workload ...

Destination Workload Groups: Select Destination Workl...

Source IP Addresses: 1.1.1.0/24

Destination IP Addresses: 2.2.2.0/24

Source IP Collections: Select Source IP Collectio...

Destination IP Collections: Select Destination IP Coll...

Protocol/Ports and Apps

Apps: Select Apps...

Protocol: gre — Ports

Protocol: ah — Ports

Protocol: esp — Ports

+ Protocols and Ports

+ ADD

Figure 102. UI example of rules with multiple named IP protocols

Figure 102 shows an example where policy is created with a rule referencing a combination of named IP protocols such as GRE, AH or ESP (IPsec protocols).

API Examples

The following creates a policy with a rule referencing multiple IP protocol numbers:

```
{
    "kind": "NetworkSecurityPolicy",
    "api-version": "v1",
    "meta": {
        "name": "rule-with-proto-numbers",
        "tenant": "default"
    },
    "spec": {
        "attach-tenant": true,
        "rules": [
            {
                "proto-ports": [
                    {
                        "protocol": "10"
                    },
                    {
                        "protocol": "254"
                    },
                    {
                        "protocol": "67"
                    },
                    {
                        "protocol": "202"
                    },
                    {
                        "protocol": "143"
                    },
                    {
                        "protocol": "199"
                    }
                ],
                "action": "permit",
                "from-ip-addresses": [
                    "1.1.1.1/32"
                ],
                "to-ip-addresses": [
                    "2.2.2.2/32"
                ]
            }
        ]
        }
        }
}
```

Policy Distribution Targets

A Policy Distribution Target (PDT) is a collection of DSSes that can be used to specify which DSSes inherit a given policy. In earlier releases, policies were always deployed on all registered switches. With PDTs, administrators have more control over the deployment and distribution of policies.

In this release, the scope of PDTs is limited to VRF-based firewall, NAT and IPsec policies. (The use of PDTs for network-based firewall policies is not currently supported.) NAT and IPsec policies make use of PDTs in their definition. In typical deployments, it is common to have a pair of VSX switches deployed as a border leaf pair. In such cases, the pair of switches in VSX *must* be defined as part of the same PDT.

The following rules apply to PDTs:

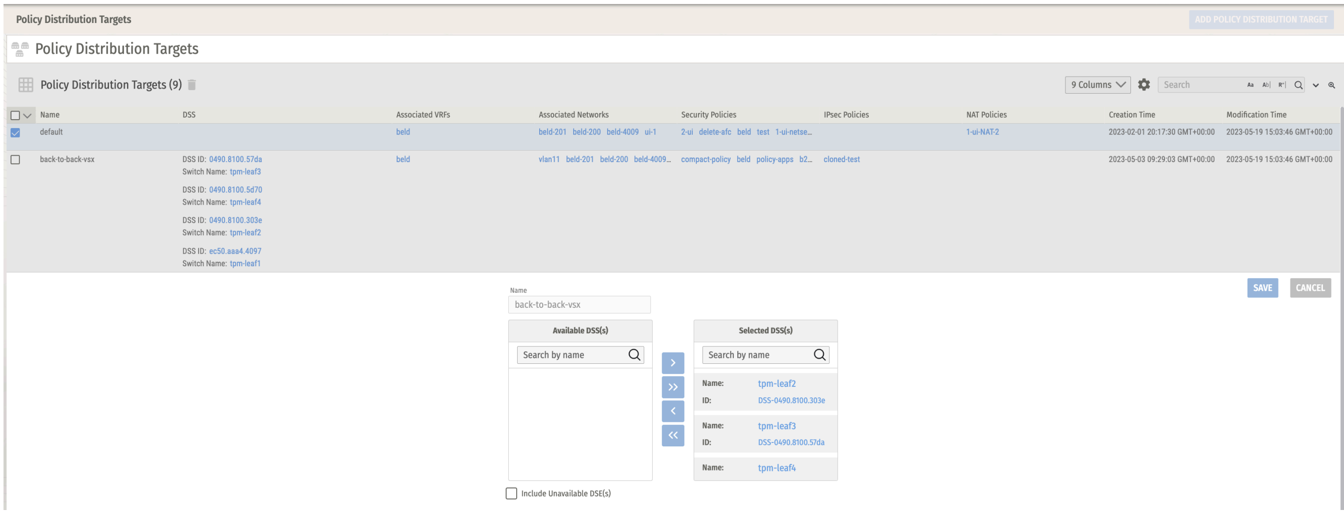
- A PDT is defined as a collection of one or more switches.
- All switches by default are part of the “default” PDT.
- A given switch can only be part of a single PDT.
- Policies can reference one or more PDTs.

NOTE: *VRF default will not be redirected if switch is in L3-core or Spine profile.*

Adding a Switch to a PDT

Follow these steps on the UI to add a switch or list of switches to a PDT:

1. On the left panel, click on “Tenants” → “Policy Distribution Targets” → “Add Policy Distribution Target”. That will open a page as shown below:



Policy Distribution Targets

ADD POLICY DISTRIBUTION TARGET

Policy Distribution Targets

Policy Distribution Targets (9)

9 Columns

☐	Name	DSS	Associated VRFs	Associated Networks	Security Policies	IPsec Policies	NAT Policies	Creation Time	Modification Time
☑	default		beld	beld-201 beld-200 beld-4009 ui-1	2-ui delete-afc beld test 1-ui-netse…		1-ui-NAT-2	2023-02-01 20:17:30 GMT+00:00	2023-05-19 15:03:46 GMT+00:00
☐	back-to-back-vsx	DSS ID: 0490.8100.57da Switch Name: tpm-leaf3 DSS ID: 0490.8100.5d70 Switch Name: tpm-leaf4 DSS ID: 0490.8100.303e Switch Name: tpm-leaf2 DSS ID: ec50.aaa4.4097 Switch Name: tpm-leaf1	beld	vlan11 beld-201 beld-200 beld-4009…	compact-policy beld policy-apps b2…	cloned-test		2023-05-03 09:29:03 GMT+00:00	2023-05-19 15:03:46 GMT+00:00

Name: back-to-back-vsx

SAVE CANCEL

Available DSS(s)

Selected DSS(s)

Name: tpm-leaf2
ID: DSS-0490.8100.303e

Name: tpm-leaf3
ID: DSS-0490.8100.57da

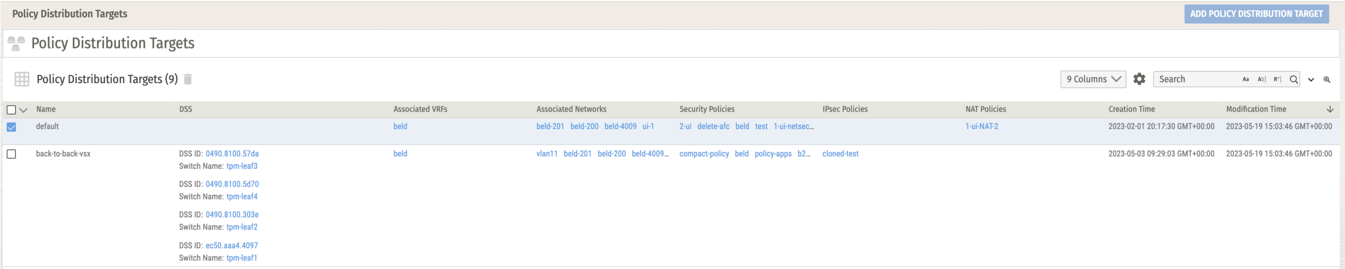
Name: tpm-leaf4

☐ Include Unavailable DSE(s)

Figure 103.

2. Create a PDT with a given name (e.g.“Border-Leaf-Devices”) and select the given DSSes that need to be added as part of this PDT, then click “CREATE POLICY DISTRIBUTION TARGET”.

If this is successful, the PDT UI page should list the created PDT and its associated switches, as shown below:



Policy Distribution Targets

ADD POLICY DISTRIBUTION TARGET

Policy Distribution Targets

Policy Distribution Targets (9)

9 Columns

☐	Name	DSS	Associated VRFs	Associated Networks	Security Policies	IPsec Policies	NAT Policies	Creation Time	Modification Time
☑	default		beld	beld-201 beld-200 beld-4009 ui-1	2-ui delete-afc beld test 1-ui-netsec…		1-ui-NAT-2	2023-02-01 20:17:30 GMT+00:00	2023-05-19 15:03:46 GMT+00:00
☐	back-to-back-vsx	DSS ID: 0490.8100.57da Switch Name: tpm-leaf3 DSS ID: 0490.8100.5d70 Switch Name: tpm-leaf4 DSS ID: 0490.8100.303e Switch Name: tpm-leaf2 DSS ID: ec50.aaa4.4097 Switch Name: tpm-leaf1	beld	vlan11 beld-201 beld-200 beld-4009…	compact-policy beld policy-apps b2…	cloned-test		2023-05-03 09:29:03 GMT+00:00	2023-05-19 15:03:46 GMT+00:00

Figure 104.

Once a switch is added to a PDT, it will be moved from the default PDT to the user-defined PDT, as shown above.

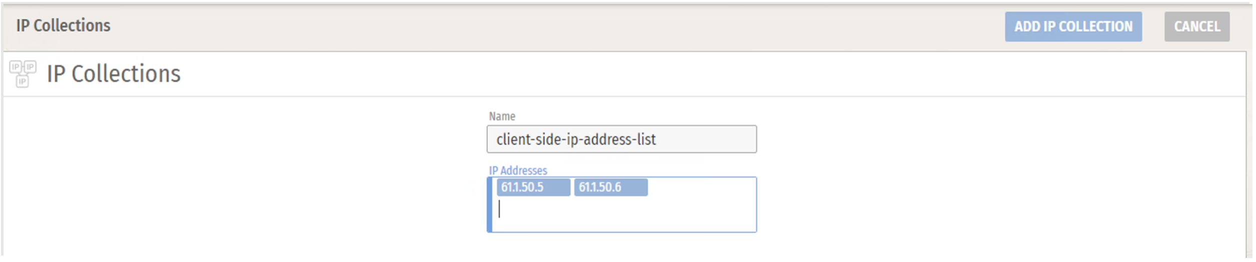
IP Collections

An *IP collection* is a list of IP addresses that can be referenced later in a firewall policy or a NAT policy.

Defining an IP Collection

On the left menu panel, click on IP Collections, which opens up the IP Collections UI page.

Click on “ADD IP COLLECTION” and define a new IP collection. The IP addresses in a collection can be defined as a space-delimited list or a range with delimiter “-“.

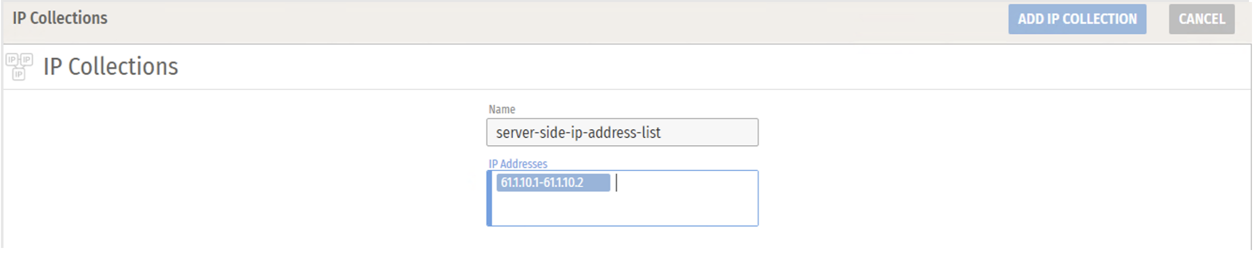


Name: client-side-ip-address-list

IP Addresses: 61.1.50.5 61.1.50.6

Figure 105.

In the above example, the IP addresses “61.1.50.5” and “61.1.50.6” are part of an IP Collection called `client-side-ip-address-list`.



Name: server-side-ip-address-list

IP Addresses: 61.1.10.1-61.1.10.2

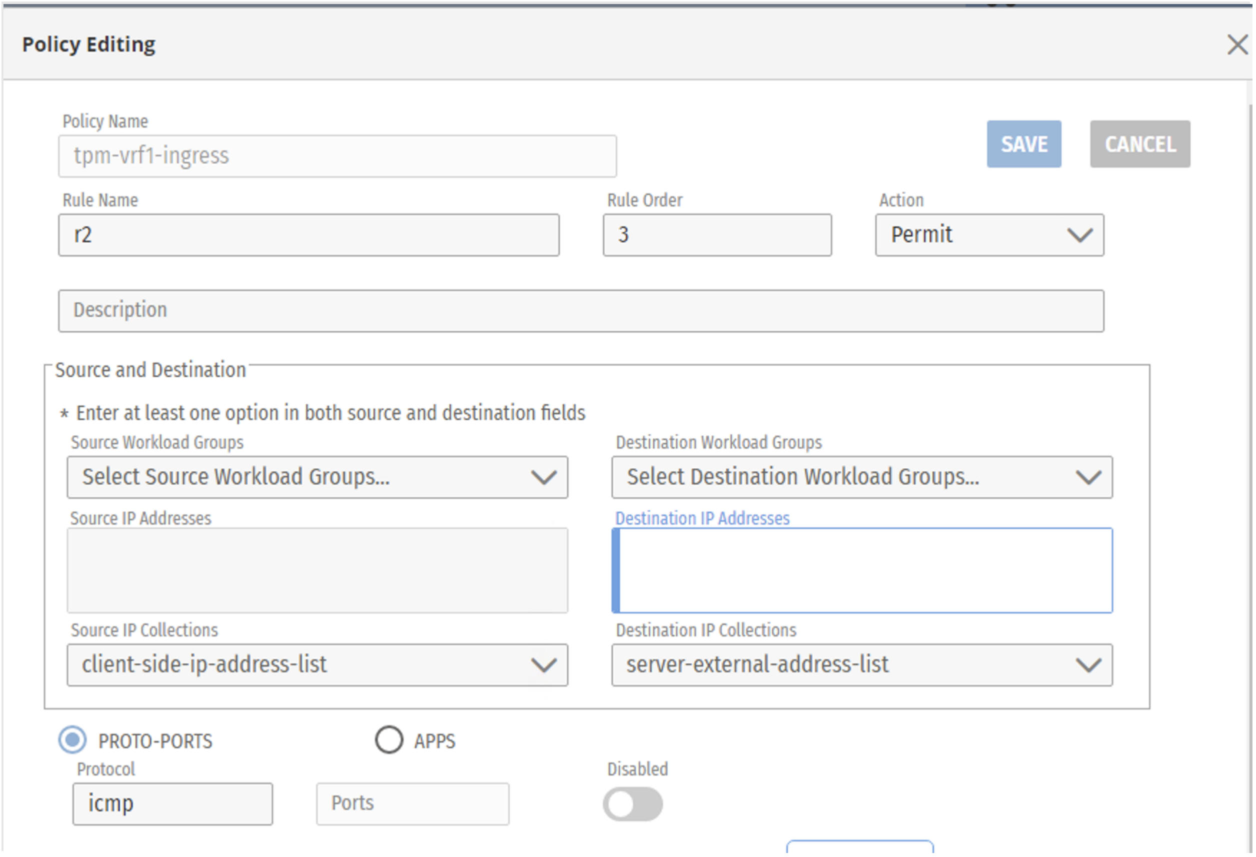
Figure 106.

In the above example, an IP collection is defined as consisting of a range denoted by “-“ .

Using an IP Collection

IP Collections can be referenced and used in a firewall policy or a NAT policy.

The following example shows an IP collection being referenced and used by a firewall policy:



Policy Editing

Policy Name: tpm-vrf1-ingress

Rule Name: r2

Rule Order: 3

Action: Permit

Description

Source and Destination

* Enter at least one option in both source and destination fields

Source Workload Groups: Select Source Workload Groups...

Destination Workload Groups: Select Destination Workload Groups...

Source IP Addresses

Destination IP Addresses

Source IP Collections: client-side-ip-address-list

Destination IP Collections: server-external-address-list

● PROTO-PORTS ○ APPS

Protocol: icmp

Ports

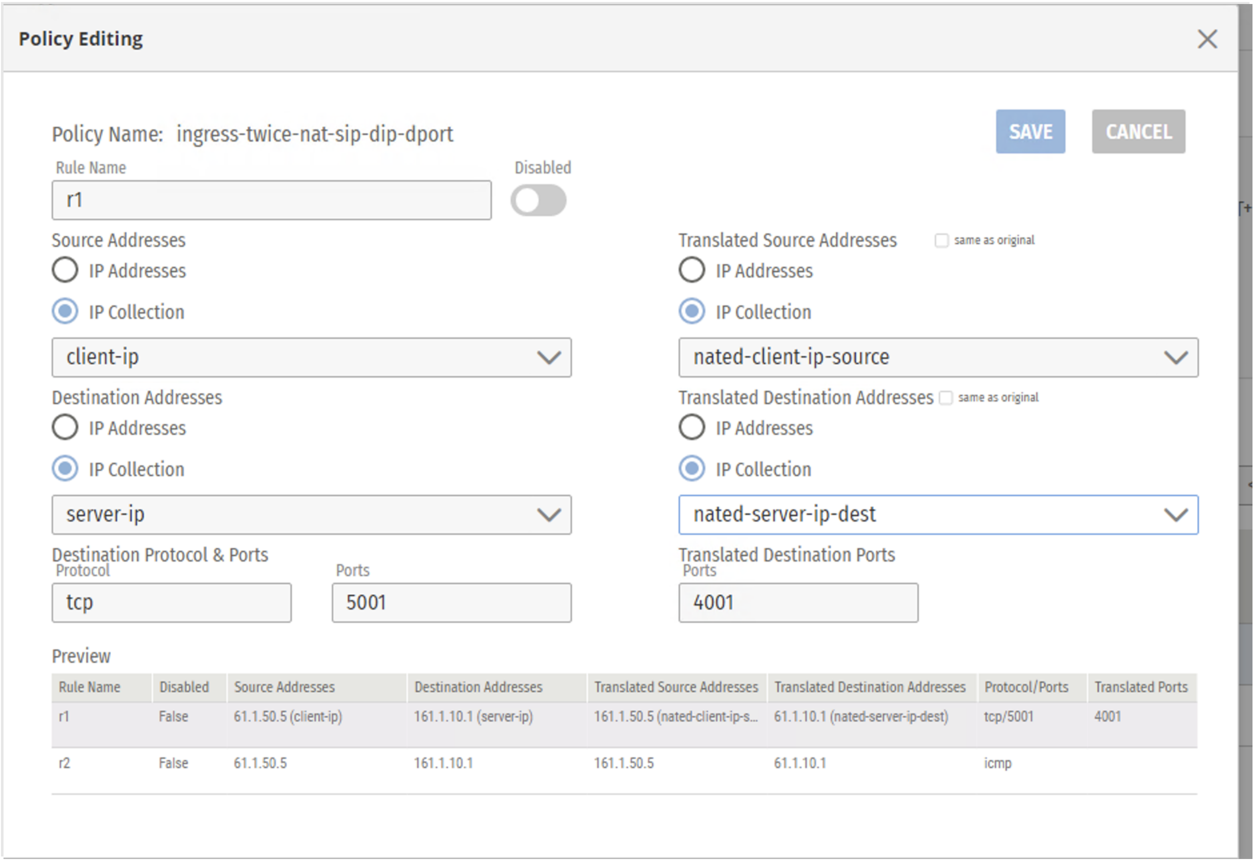
Disabled

Figure 107.

In the policy rule definition UI, either the source IP address range or the destination IP address range or both can be defined as an IP collection. This facilitates the easy definition of rules that allows for logical grouping.

An IP collection can also be referenced inside of a NAT policy.

An example of a NAT policy referencing IP collections:



Policy Editing

Policy Name: ingress-twice-nat-sip-dip-dport

Rule Name: r1

Disabled

Source Addresses
○ IP Addresses
● IP Collection
client-ip

Translated Source Addresses □ same as original
○ IP Addresses
● IP Collection
nated-client-ip-source

Destination Addresses
○ IP Addresses
● IP Collection
server-ip

Translated Destination Addresses □ same as original
○ IP Addresses
● IP Collection
nated-server-ip-dest

Destination Protocol & Ports
Protocol: tcp
Ports: 5001

Translated Destination Ports
Ports: 4001

Preview

Rule Name	Disabled	Source Addresses	Destination Addresses	Translated Source Addresses	Translated Destination Addresses	Protocol/Ports	Translated Ports
r1	False	61.1.50.5 (client-ip)	161.1.10.1 (server-ip)	161.1.50.5 (nated-client-ip-s...	61.1.10.1 (nated-server-ip-dest)	tcp/5001	4001
r2	False	61.1.50.5	161.1.10.1	161.1.50.5	61.1.10.1	icmp	

Figure 108.

Caveats

- When referenced inside of a NAT policy, an IP collection can have only one address/prefix.
- Nested IP collections are not supported; an IP collection cannot have another IP collection object as a list item.

Network Address Translation (NAT)

Network address translation (*NAT*) modifies a packet's source IP address, destination IP address/port, or in some cases both (called *twice NAT*). Typically, the translation is from an address in a private IP range to an address in a public (globally routable) IP range. The DSS can perform this operation as one of its distributed services.

NAT in Data Center Design

As one example, a company offering hosting services uses a data center fabric and host services inside the data center. These services are accessible internally using private IP addresses inside the data center fabric, but where these services need to be offered to clients outside of the data center, there is a need to convert these IP addresses from private space to global public space. Companies can advertise internal addresses externally either with a one-to-one static relationship (called static NAT) or with a many-to-one relationship (called dynamic NAT or many-to-one NAT/PAT).

Some organizations also want customer/client IP addresses to be translated inside of their data center fabric in their own data center fabric IP addressing scheme.

The other common use case is to design a data center fabric to accommodate hosting multiple customers by assigning a VRF to each customer. In this case, customer applications in each VRF are likely to have subnets or IP addresses that overlap with other customers' uses. Customers may need to access shared services inside of the data center in a shared VRF; one way to avoid conflict with overlapping customer IP addresses is to translate the customer-specific addresses to unique IP addresses.

NAT Policy Direction

The data center border leaf positioning of the DSS is well suited for offering NAT services, as it is able to perform NAT operation for connections coming into the data center (incoming connections) by enforcing what is called an *ingress NAT policy*, and also perform the NAT operation for outbound connections from the data center by enforcing what is called an *egress NAT policy*.

Ingress NAT Policy

All traffic entering the data center (incoming connections, as seen on the border leaf) are subject to ingress NAT policy evaluation. The upstream (external facing) ports on the border leaf devices need “persona uplink” configuration on the upstream facing external interfaces for enforcing the ingress NAT policy. Only SVI uplinks are supported in this release. L3 interfaces and L3 sub-interfaces cannot be used as uplinks.

Egress NAT policy

All traffic leaving the data center (outgoing connections, as seen on the border leaf), are subject to egress NAT policy evaluation. The internal data center fabric-facing ports must be configured with “persona access” on the DC facing internal interfaces on the border leaf. These data center fabric-facing ports also need to be SVI interfaces. L3 interfaces or L3 sub-interfaces cannot be used as data center fabric-facing ports.

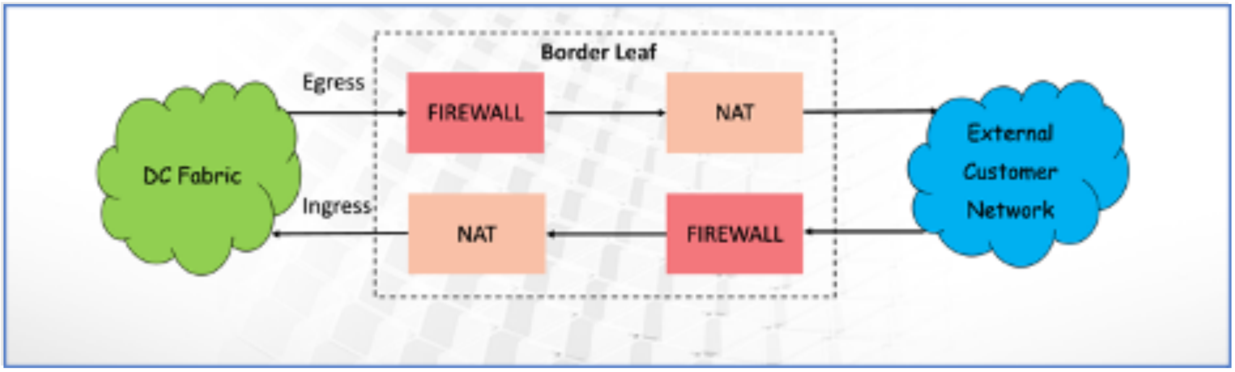


Figure 109.

In both directions, firewall policy is always applied first. In other words, firewall policy evaluation precedes NAT policy evaluation.

Supported NAT Operations

In this release, only static one-to-one NAT is supported. The term *static* refers to pre-configured mapping between pre-NAT IPv4 address/prefix/collection to post-NAT IPv4 address/prefix/collection. (One-to-many or many-to-one mapping is not allowed.) This configuration is done on the PSM.

The other form of NAT, called *dynamic NAT* or *many-to-one NAT/PAT*, is not supported in this release, but is planned for a future release.

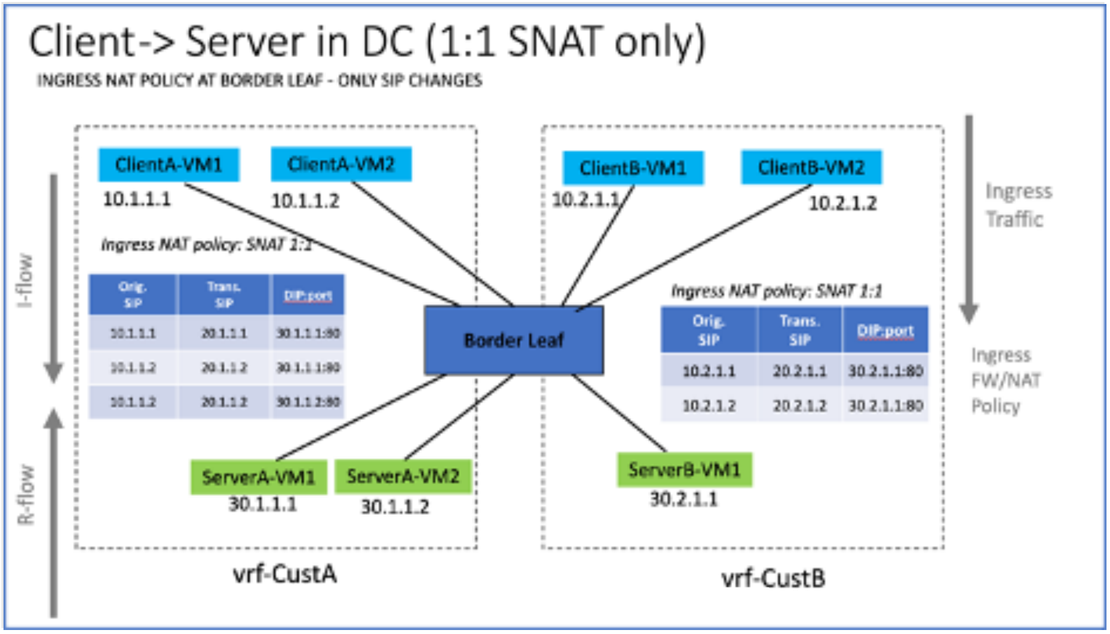
Supported Static One-to-One NAT Types

Source NAT

In the context of a stateful firewall implementation, the term source NAT refers to when the source IPv4 address of the I-flow (initiator flow) packet is translated. For the R-flow (responder flow), the destination IPv4 address is implicitly translated by the DSS. The NAT rule itself can be defined as a one-to-one mapping for a given address (/32) or a given prefix range (/24 for example) or a given IP collection (list of IPs).

The following describes examples of source NAT as applied as ingress and egress policies:

Source NAT: Ingress Policy



Ingress NAT policy: SNAT 1:1

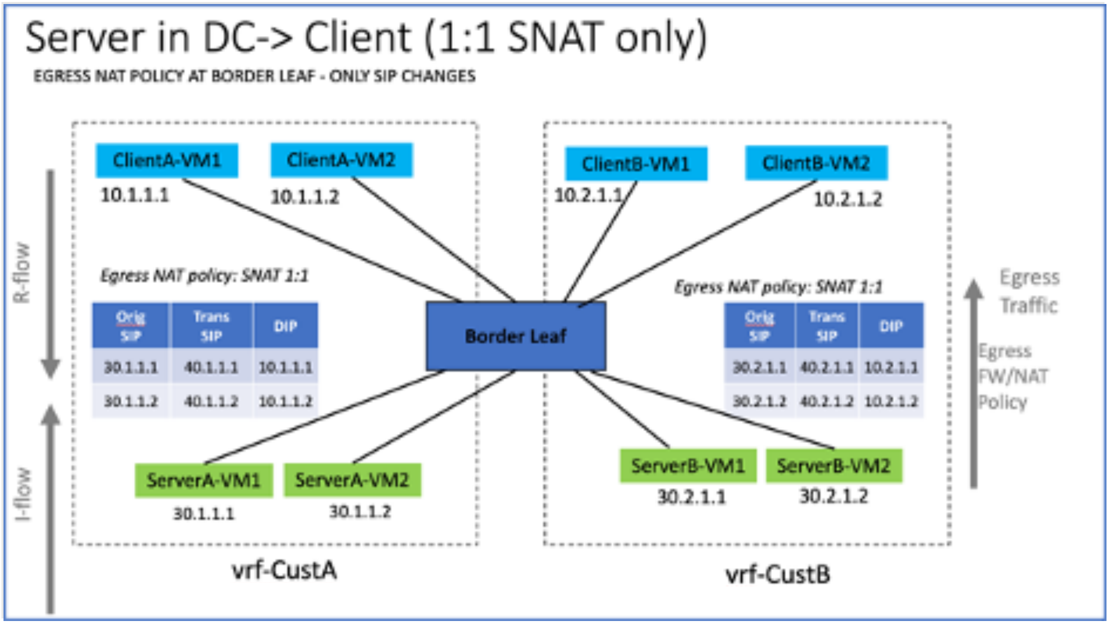
Orig. SIP	Trans. SIP	DIP:port
10.1.1.1	20.1.1.1	30.1.1.1:80
10.1.1.2	20.1.1.2	30.1.1.1:80
10.1.1.2	20.1.1.2	30.1.1.2:80

Ingress NAT policy: SNAT 1:1

Orig. SIP	Trans. SIP	DIP:port
10.2.1.1	20.2.1.1	30.2.1.1:80
10.2.1.2	20.2.1.2	30.2.1.1:80

Figure 110.

Source NAT – Egress Policy:



Server in DC-> Client (1:1 SNAT only)

EGRESS NAT POLICY AT BORDER LEAF - ONLY SIP CHANGES

Egress NAT policy: SNAT 1:1

Orig SIP	Trans SIP	DIP
30.1.1.1	40.1.1.1	10.1.1.1
30.1.1.2	40.1.1.2	10.1.1.2

Egress NAT policy: SNAT 1:1

Orig SIP	Trans SIP	DIP
30.2.1.1	40.2.1.1	10.2.1.1
30.2.1.2	40.2.1.2	10.2.1.2

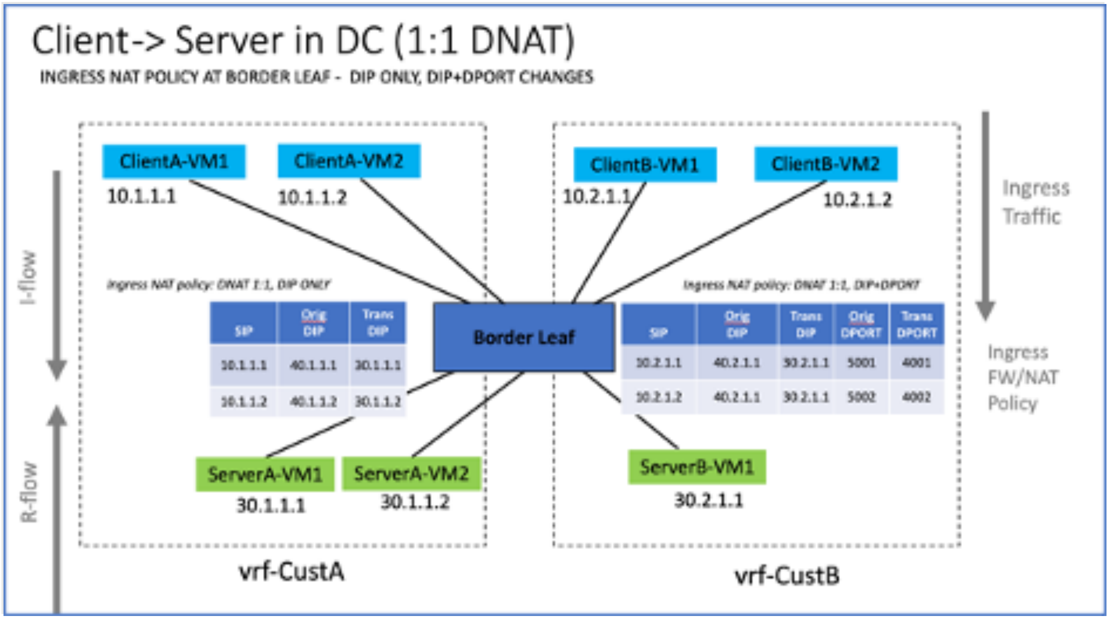
Figure 111.

Destination NAT

In the context of a stateful firewall implementation, the term *destination NAT* refers to when either only the destination IPv4 address of the I-flow packet is translated or when the destination IPv4 address and the destination port of the I-flow are translated. For the R-flow, the source IPv4 address or source IPv4 and source port is implicitly translated by the DSS. The NAT rule itself can be defined as a one-to-one mapping for a given address (/32) or a given prefix range (/24 for example) or a given IP collection (list of IPs) and destination port (if applicable).

The following describes examples of destination NAT as applied as ingress and egress policies:

Destination NAT: Ingress NAT Policy



Ingress NAT policy: DNAT 1:1, DIP ONLY

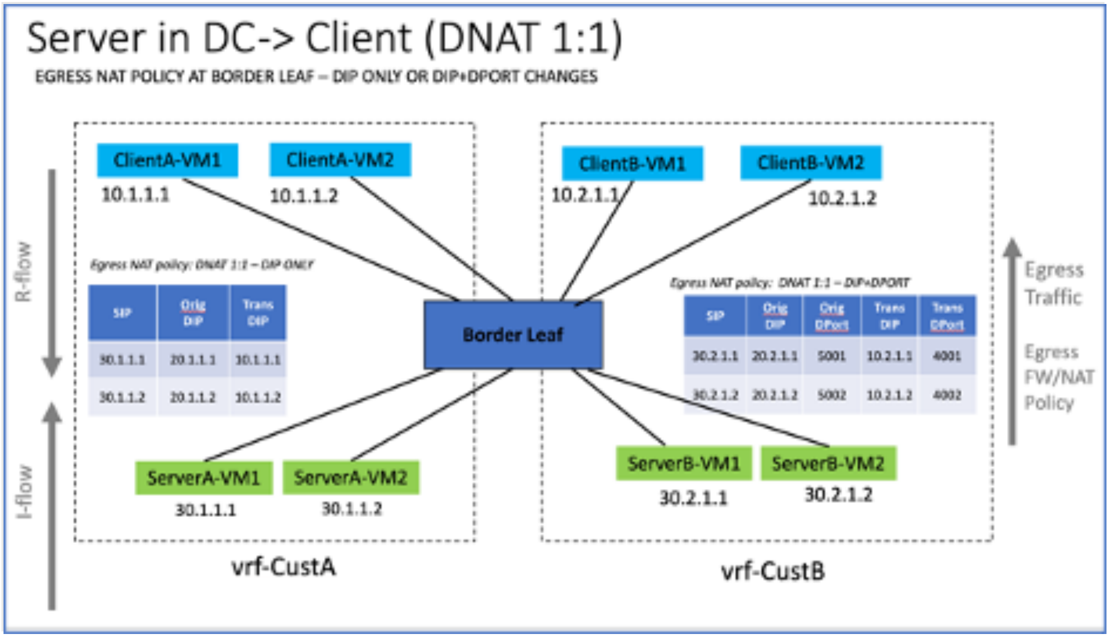
SIP	Orig DIP	Trans DIP
10.1.1.1	40.1.1.1	30.1.1.1
10.1.1.2	40.1.1.2	30.1.1.2

Ingress NAT policy: DNAT 1:1, DIP+DPORT

SIP	Orig DIP	Trans DIP	Orig DPORT	Trans DPORT
10.2.1.1	40.2.1.1	30.2.1.1	5001	4001
10.2.1.2	40.2.1.1	30.2.1.1	5002	4002

Figure 112.

Destination NAT: Egress NAT Policy



Egress NAT policy: DNAT 1:1 – DIP ONLY

SIP	Orig DIP	Trans DIP
30.1.1.1	20.1.1.1	10.1.1.1
30.1.1.2	20.1.1.2	10.1.1.2

Egress NAT policy: DNAT 1:1 – DIP+DPORT

SIP	Orig DIP	Orig DPort	Trans DIP	Trans DPort
30.2.1.1	20.2.1.1	5001	10.2.1.1	4001
30.2.1.2	20.2.1.2	5002	10.2.1.2	4002

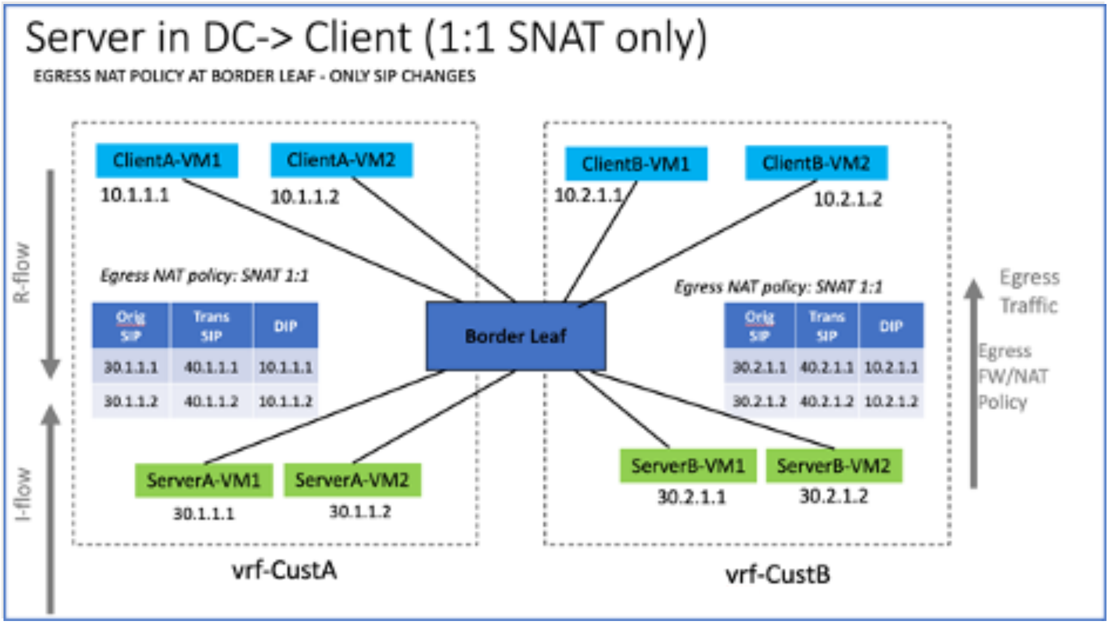
Figure 113.

Twice NAT

In the context of a stateful firewall implementation, the term *twice NAT* refers to the scenario where either both the source IPv4 address and destination IPv4 address of the I-flow get translated at the same time or the source IPv4, destination IPv4 and destination port all get translated for the same packet. For the R-flow, both the source IPv4, destination IPv4, source-port (if applicable) is NATed implicitly by the DSM. The NAT rule itself can be defined as a one-to-one mapping for a given address (/32) or a given prefix range (/24 for example) or a given IP collection for both source and destination address/prefix and destination port (if applicable).

The following describes examples of twice NAT as applied as ingress and egress policies:

Twice NAT: Ingress NAT Policy



Server in DC-> Client (1:1 SNAT only)

EGRESS NAT POLICY AT BORDER LEAF - ONLY SIP CHANGES

Egress NAT policy: SNAT 1:1

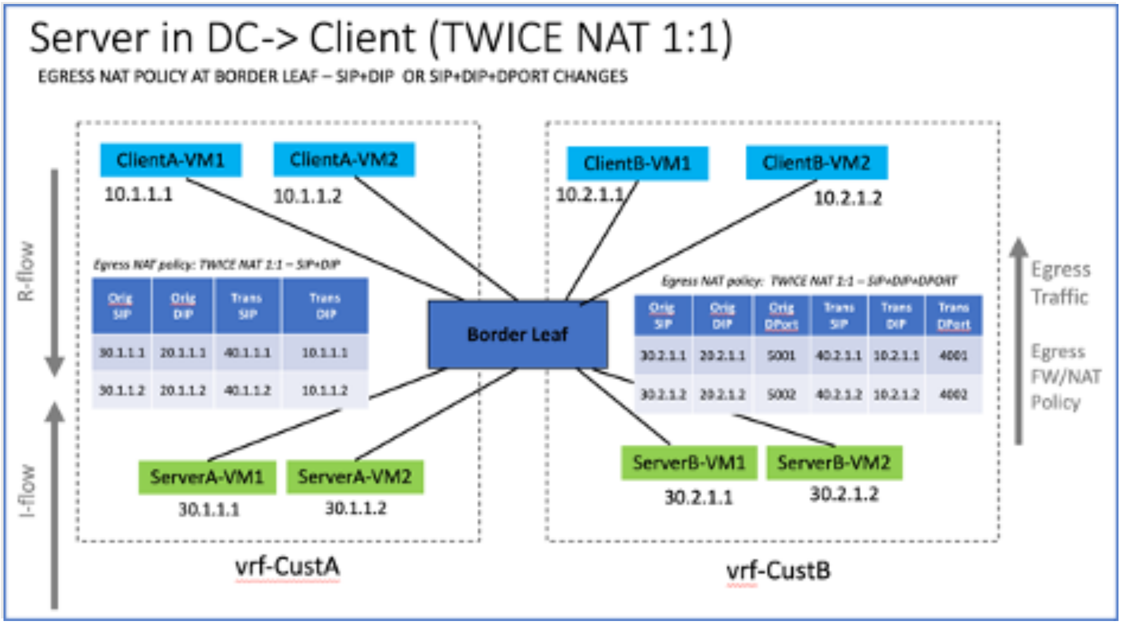
Orig SIP	Trans SIP	DIP
30.1.1.1	40.1.1.1	10.1.1.1
30.1.1.2	40.1.1.2	10.1.1.2

Egress NAT policy: SNAT 1:1

Orig SIP	Trans SIP	DIP
30.2.1.1	40.2.1.1	10.2.1.1
30.2.1.2	40.2.1.2	10.2.1.2

Figure 114.

Twice NAT: Egress NAT Policy



Egress NAT policy: TWICE NAT 1:1 – SIP+DIP

Orig SIP	Orig DIP	Trans SIP	Trans DIP
30.1.1.1	20.1.1.1	40.1.1.1	10.1.1.1
30.1.1.2	20.1.1.2	40.1.1.2	10.1.1.2

Egress NAT policy: TWICE NAT 1:1 – SIP+DIP+DPORT

Orig SIP	Orig DIP	Orig DPort	Trans SIP	Trans DIP	Trans DPort
30.2.1.1	20.2.1.1	5001	40.2.1.1	10.2.1.1	4001
30.2.1.2	20.2.1.2	5002	40.2.1.2	10.2.1.2	4002

Figure 115.

Configure NAT Policy on the PSM

Before configuring NAT on the PSM, there are a few pre-requisite conditions to be met and configuration to be done on the switch. Here are the steps:

Switch-Side Configuration:

Step 1: Preparing the DSS to be in border leaf mode

Issue the following AOS-CX command on the DSS:

```
tpm-leaf5(config)# profile l3-core
Save this config and reboot the switch for the
changes to take effect
tpm-leaf5# copy running-config startup-config
Copying configuration: [Success]
```

Reboot the switch. After the switch comes online, make sure it is in the "L3-core" profile by issuing the following command:

```
tpm-leaf5(config)# show profiles current
Current Profile
---------------
L3-core
```

Step 2: Configuring interface persona

On interfaces facing externally, configure “persona uplink”:

```
tpm-leaf5(config)# int 1/1/3
tpm-leaf5(config-if)# persona uplink
```

On interfaces facing the internal DC fabric, configure “persona access”:

```
tpm-leaf5(config)# int 1/1/4
tpm-leaf5(config-if)# persona access
```

Step 3: Advertise the post NATed addresses in IGP

The post-NAT address must always be advertised, either within the data center fabric, or externally, based on the direction of NAT and type of policy. If the post-NAT addresses are known, this step can be done prior to defining the NAT policy on the PSM. If not, this step can be defined after NAT policy on PSM is complete.

Example of static route definition on AOS-CX:

```
tpm-leaf5(config)# ip route 161.1.10.0/24 nullroute
vrf tpm-vrf1
```

Here 161.1.10/24 is the post NAT prefix that will be used in the NAT policy.

Advertise static routes in BGP using the following command:

```
tpm-leaf5(config)# router bgp 65003
     vrf tpm-vrf1
     address-family ipv4 unicast
          redistribute static
```

This should complete the required switch side config on the PSM.

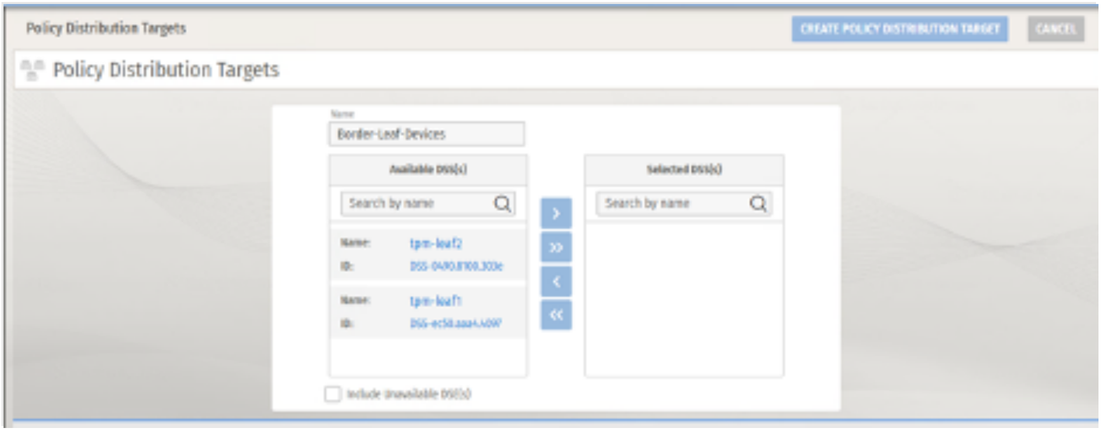
PSM Side Configuration:

Step 1: Defining a policy distribution target (PDT) containing the border-leaf DSS devices

NAT policies are pushed to only those DSS devices that are in border leaf (BL) mode. It is typical to have a pair of DSS devices in VSX act as a BL pair for redundancy and load balancing. The DSS devices in the BL-VSX pair must be configured and added in a PDT.

On the PSM UI:

On the left-panel, click on “Tenants” -> “Policy Distribution Targets” -> “Add Policy Distribution Target”. That will open a page as shown below:


Policy Distribution Targets
CREATE POLICY DISTRIBUTION TARGET
CANCEL
Policy Distribution Targets
Available DSS(s)
Selected DSS(s)
Search by name
Search by name
›
»
‹
«
□ Include Unavailable DSS(s)

Name: Border-Leaf-Devices

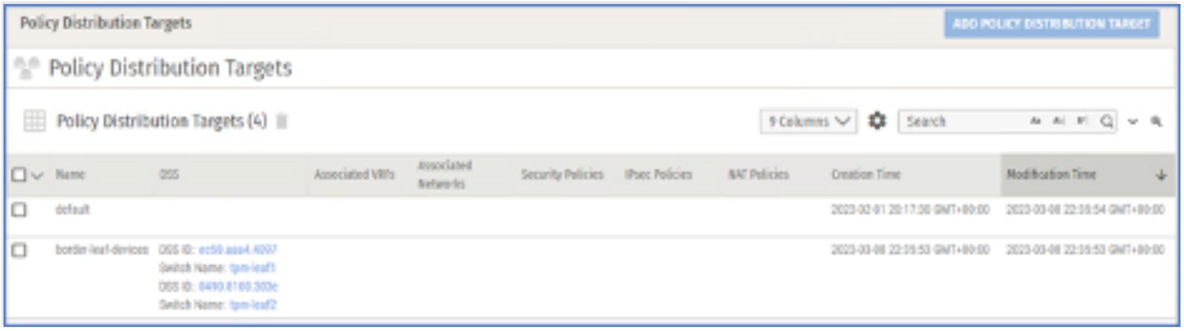
Name: tpm-leaf2
ID: DSS-0490.8100.303e

Name: tpm-leaf1
ID: DSS-ec58.aaa4.a0f7

Figure 116.

Create a PDT (in this example, named `border-leaf-devices`), select the DSSes that need to be added to the PDT and then click “Create Policy Distribution Target”.

If this is successful, the PSM “Policy Distribution Targets” page should list the created PDT and its associated switches, as shown below:


Policy Distribution Targets
ADD POLICY DISTRIBUTION TARGET
Policy Distribution Targets
Policy Distribution Targets (4)
9 Columns
Search

Name	DSS	Associated VRFs	Associated Networks	Security Policies	IPsec Policies	NAT Policies	Creation Time	Modification Time
default							2023-02-01 20:17:30 GMT+00:00	2023-03-08 22:35:54 GMT+00:00
border-leaf-devices	DSS ID: ec50.aaa4.4097 Switch Name: tpm-leaf1 DSS ID: 0490.8100.300e Switch Name: tpm-leaf2						2023-03-08 22:35:53 GMT+00:00	2023-03-08 22:35:53 GMT+00:00

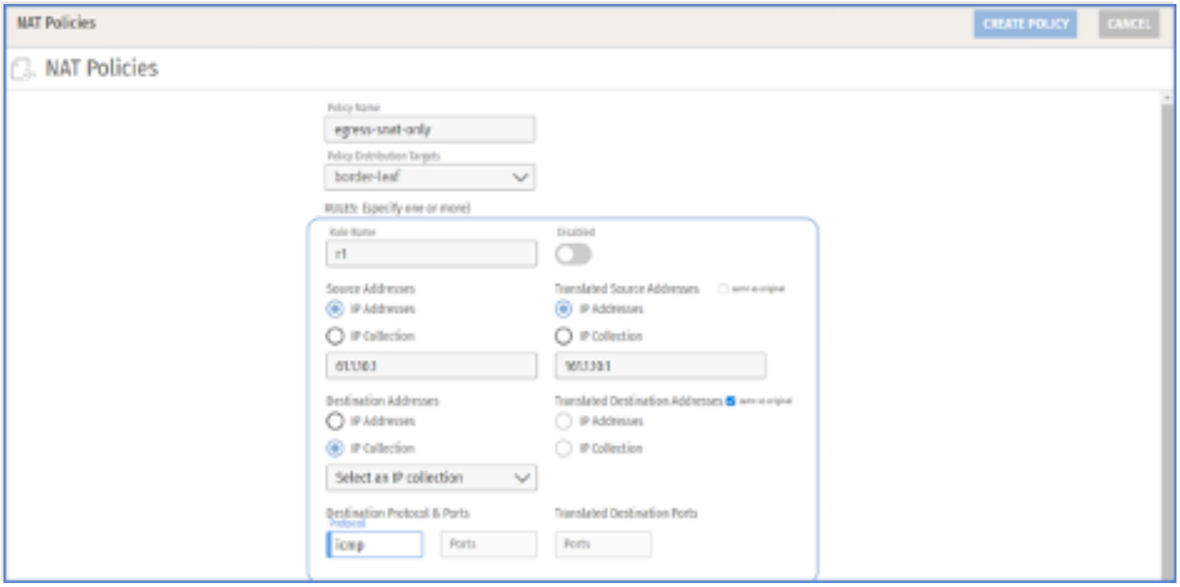
Figure 117.

Step 2: Define and apply source NAT policies

Step 2a: Define a required NAT policy

This example demonstrates the definition of an egress NAT policy that does one-to-one SNAT and applies it to VRF `tpm-vrf1`.

Navigate to “Tenants” -> “NAT Policies” and then select “Add NAT Policy”. This will open the NAT Policy definition page:



Policy Name: egress-snat-only

Policy Distribution Targets: border-leaf

RULES: (specify one or more)

Rule Name: r1

Disabled

Source Addresses: ● IP Addresses ○ IP Collection

61.1.10.1

Translated Source Addresses □ same as original: ● IP Addresses ○ IP Collection

161.1.10.1

Destination Addresses: ○ IP Addresses ● IP Collection

Select an IP collection

Translated Destination Addresses ☑ same as original: ○ IP Addresses ○ IP Collection

Destination Protocol & Ports: Protocol: icmp, Ports

Translated Destination Ports: Ports

Figure 118.

Specify the following details:

Policy Name: “egress-snat-only”

Policy Distribution Targets: select the required PDT “border-leaf”. Make sure this PDT is previously created and available

Rule Name: <name of the rule>: specified as “r1” here.

Source Addresses: <this is the original address/prefix> - specified as “61.1.10.1” here

Translated Address: <this is the post NAT IPv4 address/prefix> - specified as “161.1.10.1” here

Destination Protocol <icmp/tcp/udp> - specified as “icmp” here

Once all the required fields are entered, click on “Create Policy”. This will create a policy called “egress-nat”.

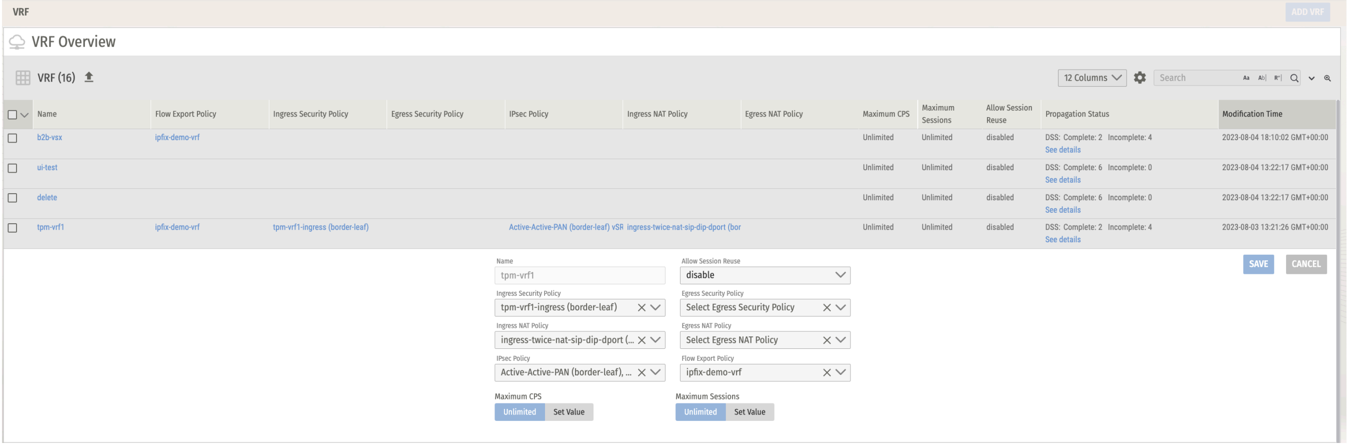
Step 2b: Apply the NAT policy on the VRF

In this step, we apply the configured NAT policy on the VRF.

In this example, will apply this policy as an egress NAT policy on the VRF

Click on the left menu panel and open the VRF page. Select the required VRF.

In the Edit VRF UI, select the Egress NAT policy as “egress-snat-only” as shown below:



Name	Flow Export Policy	Ingress Security Policy	Egress Security Policy	IPsec Policy	Ingress NAT Policy	Egress NAT Policy	Maximum CPS	Maximum Sessions	Allow Session Reuse	Propagation Status	Modification Time
b2b-vsx	ipfix-demo-vrf						Unlimited	Unlimited	disabled	DSS: Complete: 2 Incomplete: 4 See details	2023-08-04 18:10:02 GMT+00:00
ui-test							Unlimited	Unlimited	disabled	DSS: Complete: 6 Incomplete: 0 See details	2023-08-04 13:22:17 GMT+00:00
delete							Unlimited	Unlimited	disabled	DSS: Complete: 6 Incomplete: 0 See details	2023-08-04 13:22:17 GMT+00:00
tpm-vrf1	ipfix-demo-vrf	tpm-vrf1-ingress (border-leaf)		Active-Active-PAN (border-leaf) vSR	ingress-twice-nat-sip-dip-dport (bor		Unlimited	Unlimited	disabled	DSS: Complete: 2 Incomplete: 4 See details	2023-08-03 13:21:26 GMT+00:00

Name: tpm-vrf1

Allow Session Reuse: disable

Ingress Security Policy: tpm-vrf1-ingress (border-leaf)

Egress Security Policy: Select Egress Security Policy

Ingress NAT Policy: ingress-twice-nat-sip-dip-dport (...

Egress NAT Policy: Select Egress NAT Policy

IPsec Policy: Active-Active-PAN (border-leaf), ...

Flow Export Policy: ipfix-demo-vrf

Maximum CPS: Unlimited | Set Value

Maximum Sessions: Unlimited | Set Value

Figure 119.

That completes the UI workflow for configuring the egress SNAT policy on the VRF

Considerations

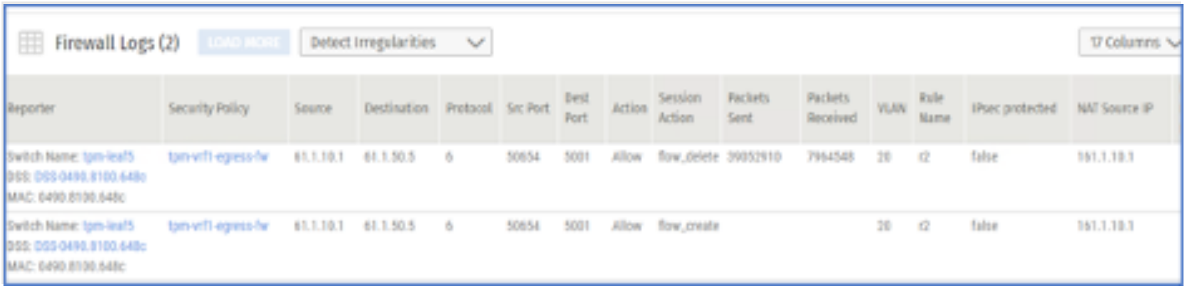
- Since firewall policy evaluation happens before NAT policy evaluation, make sure that firewall policy permits the traffic flow for NAT policy to get evaluated. Else, the packets will be dropped by the firewall policy.

- There is no need to create a NAT policy/rule for the R-flow (responder flow). The R-flow is implicitly installed. In the above example where SNAT changes the source IP of the I-flow source, the R-flow will implicit NAT the DIP
- In the case of VSX, flow-sync applies to NATed flows as well.

Example: SNAT Flow on DSM

In this example, endpoint source 61.1.10.1 initiates a TCP connection to EP destination 61.1.50.5 with SNAT translating the source IP to 161.1.10.1

Example of a firewall log as seen on the PSM for a SNAT-enabled flow:



Reporter	Security Policy	Source	Destination	Protocol	Src Port	Dest Port	Action	Session Action	Packets Sent	Packets Received	VLAN	Rule Name	IPsec protected	NAT Source IP
Switch Name: tpm-leaf5 DSS: DSS-0490.8100.648c MAC: 0490.8100.648c	tpm-vrf1-egress-fw	61.1.10.1	61.1.50.5	6	50654	5001	Allow	flow_delete	39052910	7964548	20	r2	false	161.1.10.1
Switch Name: tpm-leaf5 DSS: DSS-0490.8100.648c MAC: 0490.8100.648c	tpm-vrf1-egress-fw	61.1.10.1	61.1.50.5	6	50654	5001	Allow	flow_create			20	r2	false	161.1.10.1

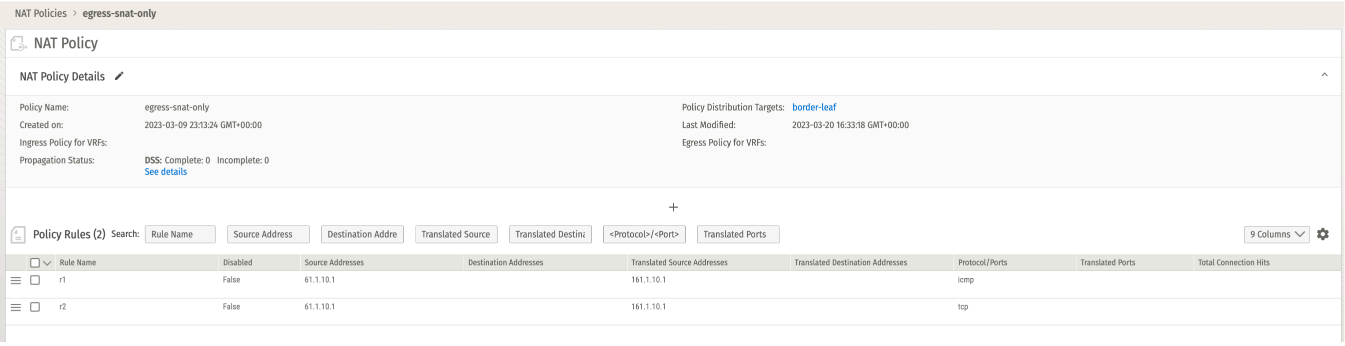
Figure 120.

As shown above, the translated source address is shown in both `firewall flow_create` and `flow_delete` logs.

Example: NAT Rule Statistics

As shown in the below output, rule-based statistics are also available for NAT rules.

In the following example, the rule “r2” in NAT policy “egress-snat-only” has received 26 connection hits on “DSM1/2”



Policy Name: egress-snat-only
Created on: 2023-03-09 23:13:24 GMT+00:00
Ingress Policy for VRFs:
Propagation Status: DSS: Complete: 0 Incomplete: 0
See details

Policy Distribution Targets: border-leaf
Last Modified: 2023-03-20 16:33:18 GMT+00:00
Egress Policy for VRFs:

Rule Name	Disabled	Source Addresses	Destination Addresses	Translated Source Addresses	Translated Destination Addresses	Protocol/Ports	Translated Ports	Total Connection Hits
r1	False	61.1.10.1		161.1.10.1		icmp		
r2	False	61.1.10.1		161.1.10.1		tcp		

Figure 121.

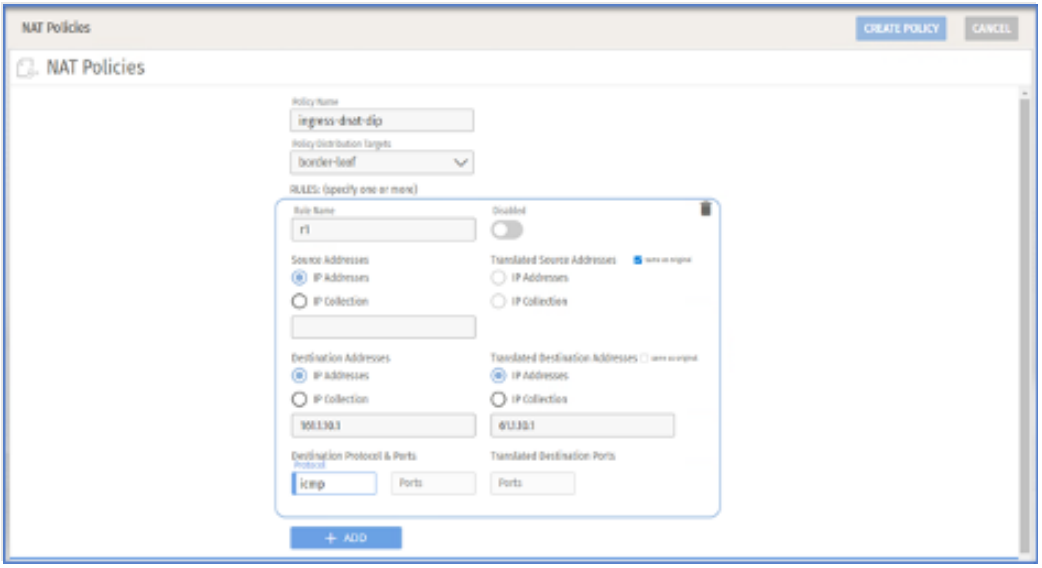
The following examples show defining and applying destination NAT and twice NAT options.

Step 3: Defining and applying destination NAT

Step 3a: Defining a destination NAT policy

In this example shows the definition of an ingress NAT policy that does one-to-one DNAT with destination IP address only translation, and applies it to a VRF “tpm-vrf1”

Click on “Tenants” -> “NAT Policies” -> “Add NAT Policy”. This will open up a UI page that looks like this:



Policy Name: ingress-dnat-dip
Policy Distribution Targets: border-leaf
RULES: (specify one or more)
Rule Name: r1
Disabled
Source Addresses: ● IP Addresses ○ IP Collection
Translated Source Addresses ☑ same as original: ○ IP Addresses ○ IP Collection
Destination Addresses: ● IP Addresses ○ IP Collection
161.1.10.1
Translated Destination Addresses □ same as original: ● IP Addresses ○ IP Collection
61.1.10.1
Destination Protocol & Ports: Protocol: icmp Ports
Translated Destination Ports: Ports

Figure 122.

Specify the following details:

- Policy Name: “ingress-dnat-dip”
- Policy Distribution Targets: select the required PDT “border-leaf”. Make sure this PDT is previously created and available
- Rule Name: specified as “r1” here.
- Destination Addresses: specified as “161.1.10.1” here
- Translated Address: specified as “61.1.10.1” here
- Destination Protocol: specified as “icmp” here

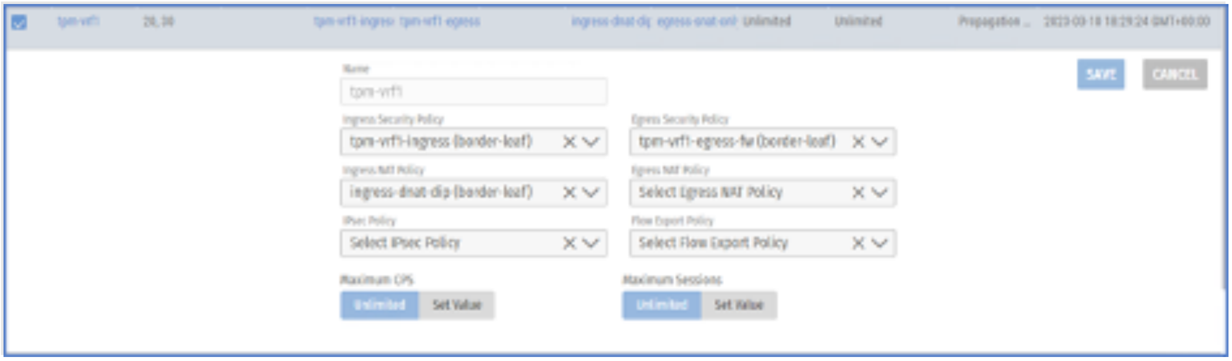
Once all the required fields are entered, click on “Create Policy”. This will create a policy called “ingress-dnat-dip”.

Step 3b: Apply the NAT policy on the VRF

In this step, we apply the configured NAT policy on the VRF. We will apply this policy as an ingress NAT policy on the VRF.

Click on the left menu panel and open the VRF page. Select the required VRF.

In the Edit VRF UI, select the Ingress NAT policy as “ingress-dnat-ip” as shown below:



☑ tpm-vrf1 20, 30 tpm-vrf1-ingress tpm-vrf1-egress ingress-dnat-dip egress-dnat-snt Unlimited Unlimited Propagation ... 2023-03-18 18:29:24 GMT+00:00

Name: tpm-vrf1

Ingress Security Policy: tpm-vrf1-ingress (border-leaf)

Egress Security Policy: tpm-vrf1-egress-fw (border-leaf)

Ingress NAT Policy: ingress-dnat-dip (border-leaf)

Egress NAT Policy: Select Egress NAT Policy

IPsec Policy: Select IPsec Policy

Flow Export Policy: Select Flow Export Policy

Maximum CPS: Unlimited | Set Value

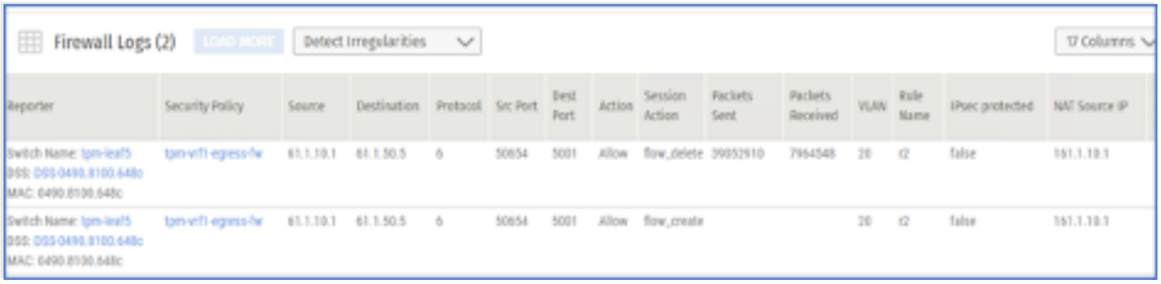
Maximum Sessions: Unlimited | Set Value

Figure 123.

This completes the UI workflow for configuring the ingress DNAT policy on the VRF

> ***NOTE:*** *There is no need to create a NAT policy/rule for the R-flow (responder flow). The R-flow is implicitly installed. In the above example where DNAT changes the destination IP of the I-flow source, the R-flow will implicitly NAT the SIP*

Example of a firewall log as seen on the PSM for a DNAT enabled flow:



Firewall Logs (2)

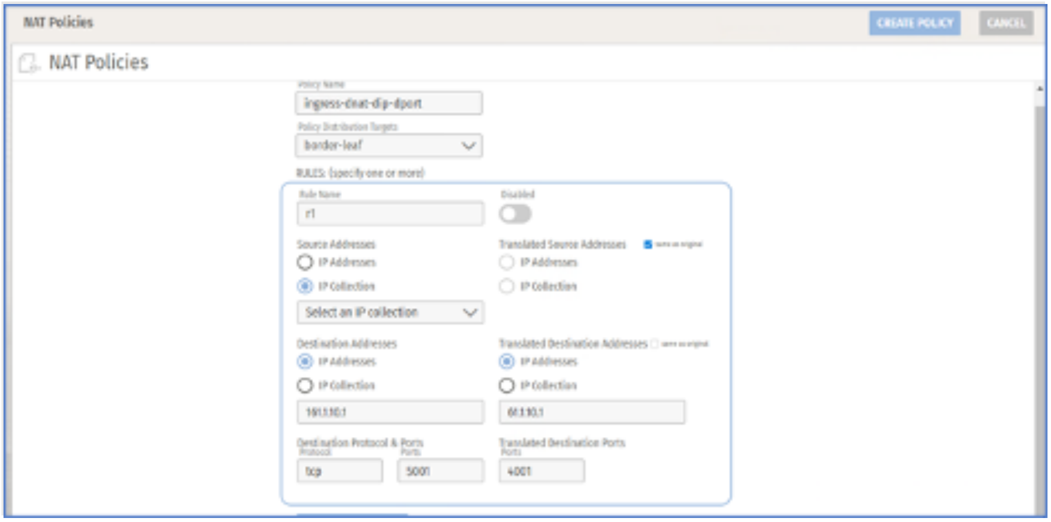
Reporter	Security Policy	Source	Destination	Protocol	Src Port	Dest Port	Action	Session Action	Packets Sent	Packets Received	VLAN	Rule Name	IPsec protected	NAT Source IP
Switch Name: tpm-leaf5 DSS: DSS-0490.8100.648c MAC: 0490.8100.648c	tpm-vrf1-egress-fw	61.1.10.1	61.1.50.5	6	50654	5001	Allow	flow_delete	39052910	7964548	20	r2	false	161.1.10.1
Switch Name: tpm-leaf5 DSS: DSS-0490.8100.648c MAC: 0490.8100.648c	tpm-vrf1-egress-fw	61.1.10.1	61.1.50.5	6	50654	5001	Allow	flow_create			20	r2	false	161.1.10.1

Figure 124.

As shown above, the translated destination address is shown in the firewall log as “61.1.10.1” in the flow_create and flow_delete logs for the original flow with SIP of “61.1.50.5” and DIP of “161.1.10.1”.

In this section, we show an example of configuring a DNAT policy with both DIP and DPORT translation

Step 3c: Defining a destination NAT policy with both DIP and DPORT translation



Policy Name: ingress-dnat-dip-dport

Policy Distribution Targets: border-leaf

RULES: (specify one or more)

Rule Name: r1

Disabled

Source Addresses: ○ IP Addresses ● IP Collection

Select an IP collection

Translated Source Addresses ☑ same as original: ○ IP Addresses ○ IP Collection

Destination Addresses: ● IP Addresses ○ IP Collection

161.1.10.1

Translated Destination Addresses □ same as original: ● IP Addresses ○ IP Collection

61.1.10.1

Destination Protocol & Ports: Protocol: tcp Ports: 5001

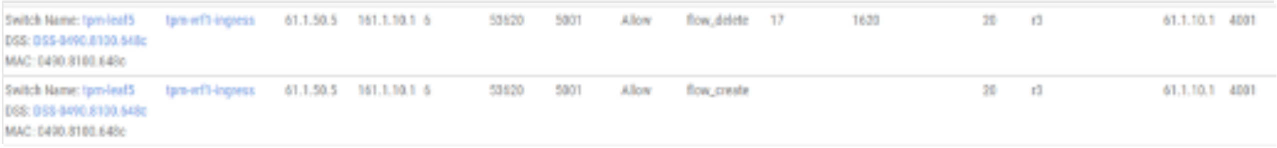
Translated Destination Ports: Ports: 4001

Figure 125.

Here the policy `ingress-dnat-dip-port` is applied to PDT `border-leaf`, and the rule `r1` translates both the DIP and DPORT. The DIP is translated from “161.1.10.1” to “61.1.10.1” and the DPORT is translated from “5001” to “4001”.

Step 3d: Apply the defined DNAT policy on the VRF for enforcement

Example of a firewall log on the PSM that shows both DIP and DPORT translated:



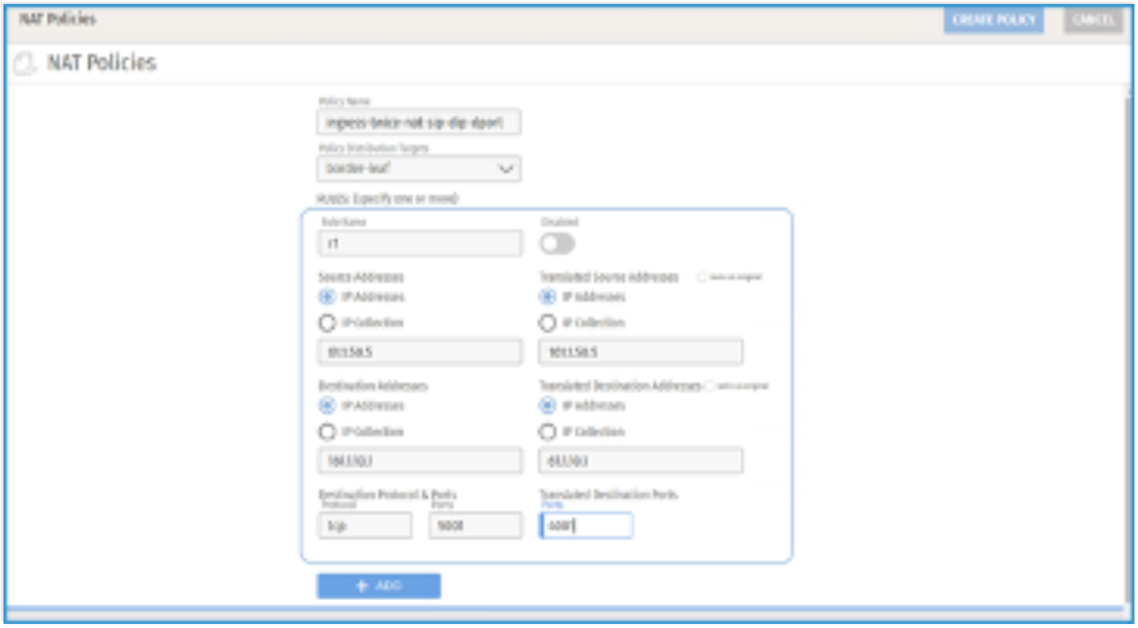
```
Switch Name: tpm-leaf5    tpm-vrf1-ingress   61.1.50.5   161.1.10.1  6   53620   5001   Allow   flow_delete   17   1620   20   r3   61.1.10.1   4001
DSS: DSS-0490.8100.648c
MAC: 0490.8100.648c
Switch Name: tpm-leaf5    tpm-vrf1-ingress   61.1.50.5   161.1.10.1  6   53620   5001   Allow   flow_create             20   r3   61.1.10.1   4001
DSS: DSS-0490.8100.648c
MAC: 0490.8100.648c
```

Figure 126.

As shown above, the translated destination address is shown in the firewall log as “61.1.10.1” and the translated destination port is shown as “4001” in the `flow_create` and `flow_delete` logs for the original flow with SIP of “61.1.50.5” and DIP of “161.1.10.1” and DPORT of “5001”

An example of configuring an ingress twice NAT policy with SIP, DIP and DPORT translation:

Step 3e: Defining a twice NAT policy with SIP, DIP and DPORT translation



NAT Policies

Policy Name: ingress-twice-nat-sip-dip-dport

Policy Distribution Targets: border-leaf

Rules (Specify one or more)

Rule Name: r1

Disabled

Source Addresses: ● IP Addresses ○ IP Collection — 61.1.50.5

Translated Source Addresses: ● IP Addresses ○ IP Collection — 161.1.50.5

Destination Addresses: ● IP Addresses ○ IP Collection — 161.1.10.1

Translated Destination Addresses: ● IP Addresses ○ IP Collection — 61.1.10.1

Destination Protocol & Ports: Protocol: tcp Ports: 5001

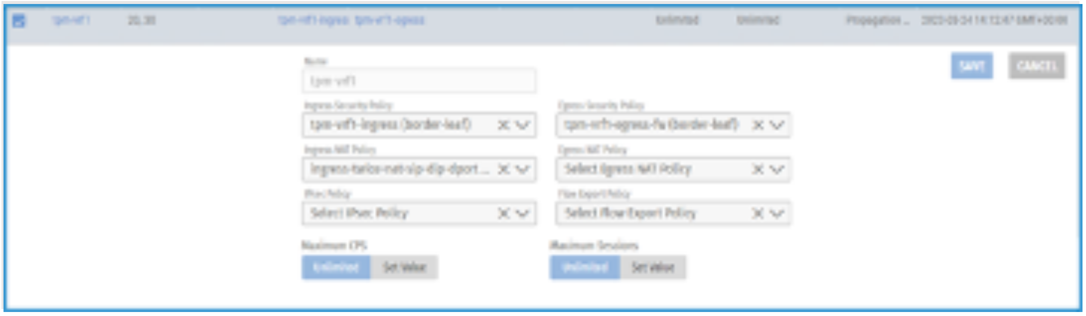
Translated Destination Ports: Ports: 4001

Figure 127.

The above twice NAT policy has a rule “r1” that translates the SIP from “61.1.50.5” to “161.1.50.5”, the DIP from “161.1.10.1” to “61.1.10.1” and the DPORT from “5001” to “4001” for protocol “tcp”. This policy will be pushed to PDT “border-leaf”.

Step3f: Apply the defined twice NAT policy as an ingress policy on the VRF for enforcement

In this example, the twice NAT policy `ingress-twice-nat-sip-dip-dport` is applied as an ingress NAT policy on the vrf “tpm-vrf1”



tpm-vrf1 | tpm-vrf1-ingress tpm-vrf1-egress | Unlimited | Unlimited | Propagation ...

Name: tpm-vrf1

Ingress Security Policy: tpm-vrf1-ingress (border-leaf)

Egress Security Policy: tpm-vrf1-egress-fw (border-leaf)

Ingress NAT Policy: ingress-twice-nat-sip-dip-dport...

Egress NAT Policy: Select Egress NAT Policy

IPsec Policy: Select IPsec Policy

Flow Export Policy: Select Flow Export Policy

Maximum CPS: Unlimited | Set Value

Maximum Sessions: Unlimited | Set Value

Figure 128.

Example of a firewall log on the PSM that shows SIP, DIP and DPORT translated:

As shown above, the NATed source IP is shown as “161.1.50.5”, NATed destination IP is shown as “61.1.10.1” and the NATed destination port is shown as “4001” in the `flow_create` and `flow_delete` logs for the original flow with SIP of “61.1.50.5” and DIP of “161.1.10.1” and DPORT of “5001”

Caveats

The following are known caveats for the NAT feature in this release:

- Applicable only to DSSes in border leaf mode. NAT policies are not supported on host leaf DSS devices.
- NAT policies can only be applied at VRF level; they cannot be applied to network level.
- NAT policy cannot be applied on default VRF, only applicable for user-defined VRFs.
- ALG traffic cannot be subject to NAT policies.
- vMotion flow migration is not applicable for flows subject to NAT policies.
- Firewall logs will not contain the NAT policy and NAT rule name.
- IPv4 only. IPv6 NAT is not supported.
- Static one-to-one NAT only; no support for Dynamic NAT or many-to-one NAT/PAT options.

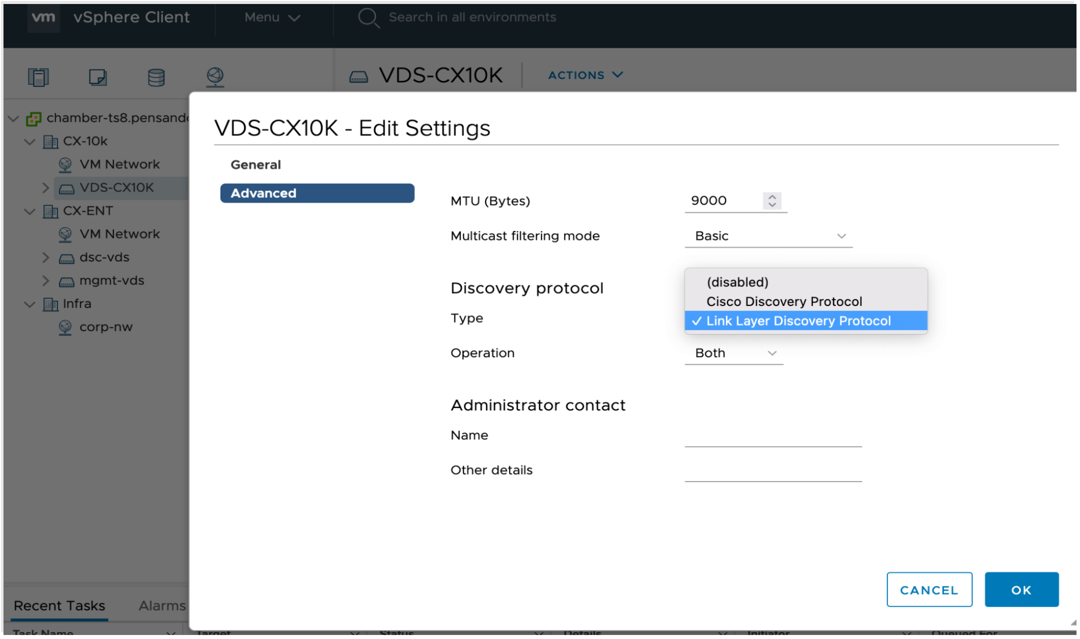
Stateful Firewall Flow Migration with vMotion

In the event of a vMotion migration of VMs between two ESX hosts directly connected to a DSS (or a VSX pair of DSSes), PSM integration with vCenter allows firewall flows to be statefully migrated from the original connected DSS to the new connected DSS.

The vMotion migration itself is transparent to the hosted application, as the stateful nature of flows are preserved, with applications experiencing very minimal loss during the event. The intent for flow migration is to allow connection tracking to be enabled and to preserve associated features for flows, such as flow statistics and flow logs, following a vMotion event. To support vMotion of VMs between ESX hosts connected to DSSes in the cluster at steady state, some additional information is collected from vCenter and AOS-CX.

Prerequisites:

1. All ESX hosts **must** be directly connected to a DSS, as this feature relies on LLDP for locality resolution.
2. vMotion will be supported only within the same PSM cluster: The source DSS and the destination DSS must be controlled by the same PSM cluster.
3. All ESX hosts **must** have LLDP enabled and LLDP **must** be enabled on all DSS ports.



VDS-CX10K - Edit Settings

MTU (Bytes): 9000

Multicast filtering mode: Basic

Discovery protocol

Type: ✓ Link Layer Discovery Protocol

Operation: Both

Administrator contact

Name:

Other details:

Figure 129. Enabling LLDP on a vDS

4. All DSSes **must** be reachable by each other via the default VRF in-band.

5. Each DSM is available on a different TCP port. They share the same common inband IP of the DSS.
    a. NAT rules are programmed in the DSS for each DSM using the reserved L4 ports (11400-11450 for DSM1 and 11451-11502 for DSM2)
6. LLDP is not supported on standard vSwitches. A workload's network interfaces **must** be attached to distributed virtual switch (DVS) if stateful firewall vMotion support is needed.
7. vCenter versions 6.7 and 7.0 are supported.

Considerations for Multi-Homed ESX Servers

The DSS can be deployed as VSX pairs where the flows are synchronized between the pair. Hosts can then be dual-homed to the pair with the multi-chassis LAG. These hosts can be running ESX. VMs can then move from any combination of VSX pair and single homed host; for example, a single homed host to a VSX pair and back, or a VSX pair to a VSX pair. All such scenarios are supported. If the source is a VSX pair, then the flows are pulled from any one of the source DSS switches and then locally deleted on both. If the destination is a VSX pair, then the primary switch is used for the flow synch; the flows will then be synched to the secondary switch. vMotion for VMs in a PVLAN is now supported.

Behavior with Flow Logs and Flow Statistics

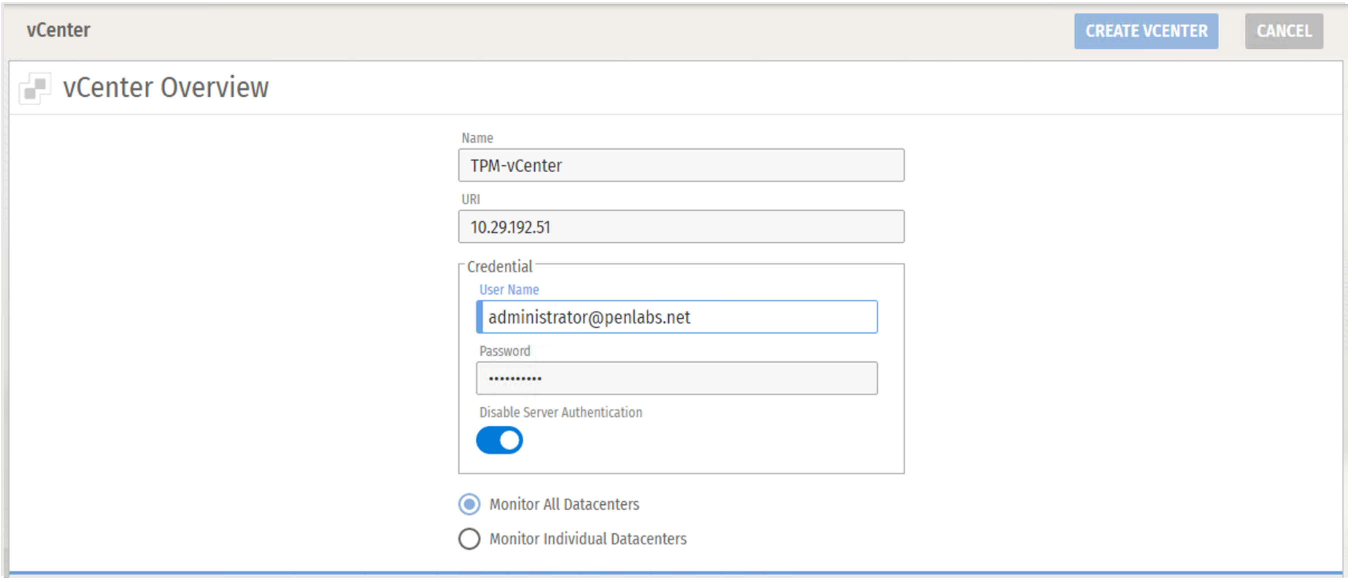
When a flow is migrated from the source DSS to the destination DSS:

1. A flow log OPEN packet is sent from the destination DSS once flow migration completes and the flow is installed. The statistics of the migrated flow are reset to 0—statistics are not carried over during flow migration. The flow log OPEN message will carry the value of “vmotion” in the “createreason” key field to indicate that the flow got inserted due to a vMotion event.
2. A flow log CLOSE message is sent from the source DSS once flow migration completes, and will include the flow's I-flow and R-flow packet/byte statistics.

When the flow eventually is removed/aged out on its new location on the destination DSS, another CLOSE packet is sent with I-flow and R-flow statistics that correspond to statistics on the destination DSS. The flow log CLOSE packet that is sent from the destination DSS post-vMotion will not have the original policy-name and the rule-name contained in the logs.

Configuration

1. The PSM connects to vCenter via user-provided credentials. There is an option to monitor either all or specific datacenters on the vCenter and receive notifications as appropriate.



vCenter Overview

Name: TPM-vCenter

URI: 10.29.192.51

Credential

User Name: administrator@penlabs.net

Password: ••••••••••

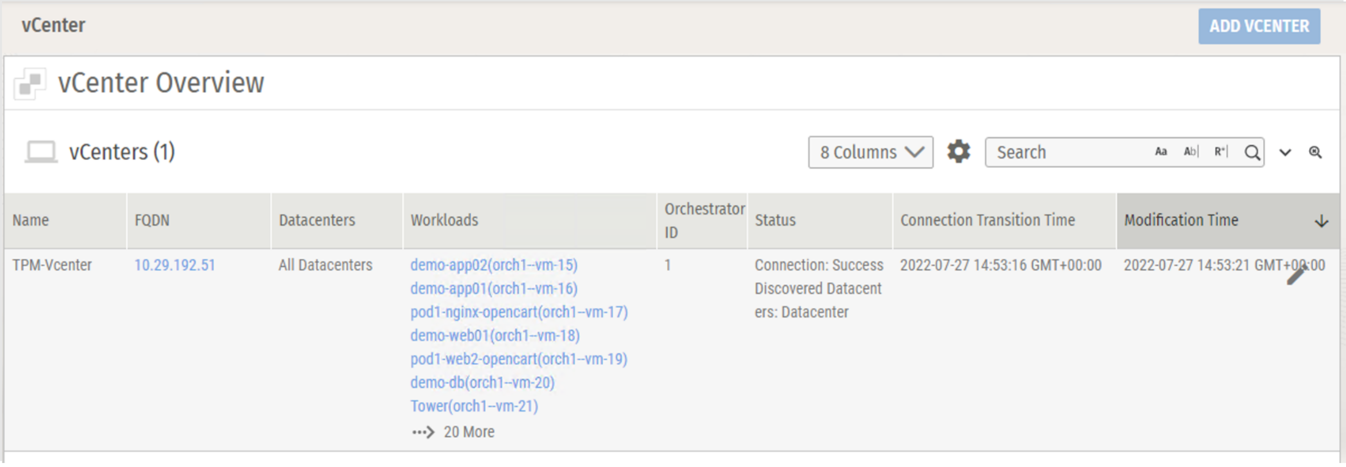
Disable Server Authentication: ●

● Monitor All Datacenters

○ Monitor Individual Datacenters

Figure 130. Add a connection to vCenter

2. The PSM reads the Host, Workloads and LLDP information from vCenter via the VMware VIM API, and is notified by vCenter of vMotion-related events.



vCenter Overview

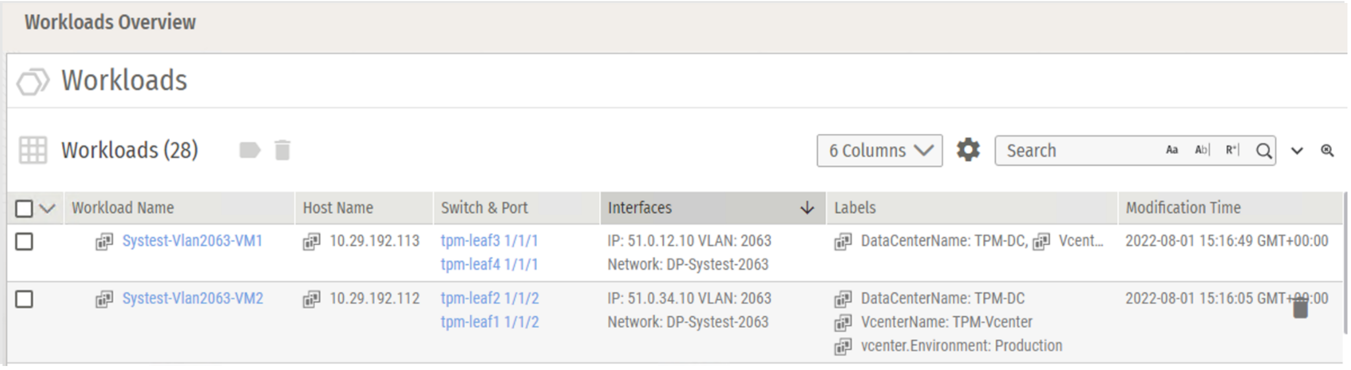
vCenters (1)

Name	FQDN	Datacenters	Workloads	Orchestrator ID	Status	Connection Transition Time	Modification Time

TPM-Vcenter	10.29.192.51	All Datacenters	demo-app02(orch1--vm-15)
			demo-app01(orch1--vm-16)
			pod1-nginx-opencart(orch1--vm-17)
			demo-web01(orch1--vm-18)
			pod1-web2-opencart(orch1--vm-19)
			demo-db(orch1--vm-20)
			Tower(orch1--vm-21)
			···> 20 More
				1	Connection: Success
					Discovered Datacenters: Datacenter
						2022-07-27 14:53:16 GMT+00:00	2022-07-27 14:53:21 GMT+00:00

Figure 131. All the workloads imported by the PSM from vCenter

3. Information received from vCenter is used to create the ESXi Host to local DSS mapping.



Workloads Overview

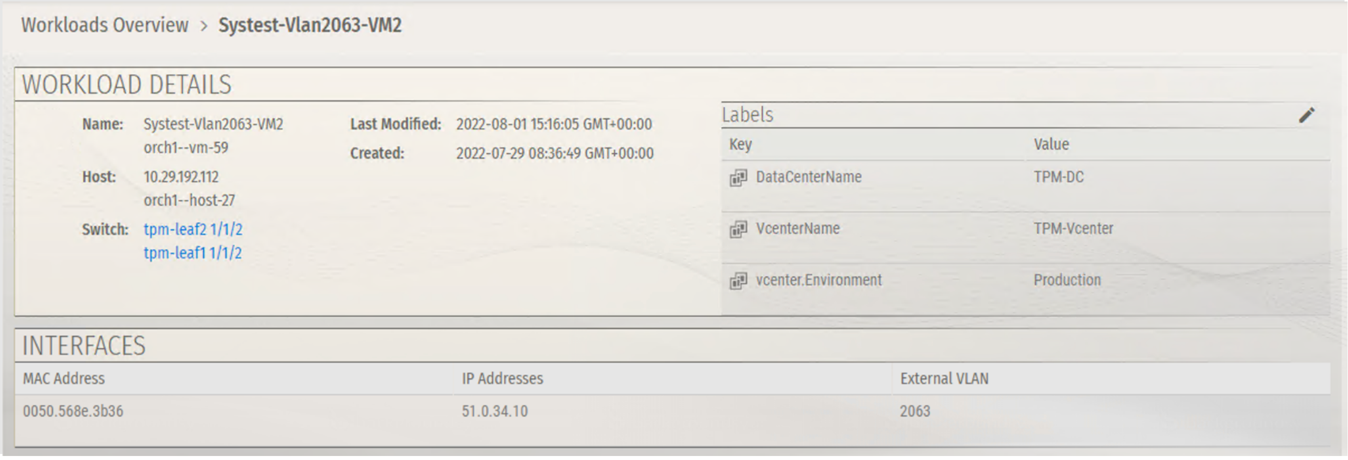
Workloads

Workloads (28)

☐	Workload Name	Host Name	Switch & Port	Interfaces	Labels	Modification Time
☐	Systest-Vlan2063-VM1	10.29.192.113	tpm-leaf3 1/1/1 tpm-leaf4 1/1/1	IP: 51.0.12.10 VLAN: 2063 Network: DP-Systest-2063	DataCenterName: TPM-DC, Vcent...	2022-08-01 15:16:49 GMT+00:00
☐	Systest-Vlan2063-VM2	10.29.192.112	tpm-leaf2 1/1/2 tpm-leaf1 1/1/2	IP: 51.0.34.10 VLAN: 2063 Network: DP-Systest-2063	DataCenterName: TPM-DC VcenterName: TPM-Vcenter vcenter.Environment: Production	2022-08-01 15:16:05 GMT+00:00

Figure 132. Two workloads and their local mappings and tags/labels

4. This database of ESC and workload/VM locality is maintained in the PSM by correlating the Host Object (ESXi Node) information and the LLDP information populated in the DSS Object.



Workloads Overview > Systest-Vlan2063-VM2

WORKLOAD DETAILS

Name: Systest-Vlan2063-VM2
orch1--vm-59

Host: 10.29.192.112
orch1--host-27

Switch: tpm-leaf2 1/1/2
tpm-leaf1 1/1/2

Last Modified: 2022-08-01 15:16:05 GMT+00:00

Created: 2022-07-29 08:36:49 GMT+00:00

Labels

Key	Value
DataCenterName	TPM-DC
VcenterName	TPM-Vcenter
vcenter.Environment	Production

INTERFACES

MAC Address	IP Addresses	External VLAN
0050.568e.3b36	51.0.34.10	2063

Figure 133. A workload and its local specific mappings

DSS Required AOS-CX CLI Configuration

Neighbor resolution must be enabled using the following AOS-CX commands on each DSS:

- MAC-IP Bindings
  - New CLI to trigger the ARP snooping on DSS.

```
switch(config)# dsm
switch(config-dsm)# workload-migration
```

- DSM-DSM communication for vMotion flow migration
  - DSM will communicate with DSS via bond interfaces created during init in vlan 4093.

***Note**: The following CLI configuration is* required *on each DSS to create the NAT rules to allow sessions establishment between the source and destination DSMs:*

```
switch(config)# ip source-interface workload_migration
interface vlanX (or routeable loopback X)
```

Caveats

- Even though vCenter can be configured to support up to 8 parallel vMotion migrations, a DSS will process one flow migration at a time; all the vMotion move-in and move-out requests received are queued while processing the existing request, and then processed sequentially.
- Port Group VLAN changes between the source DVS and the destination DVS are not supported for the Migrating VM, even though vCenter allows it. In other words, the DPG on the source and destination DVS must have the same VLAN-ID.
- When a vMotion migration occurs for a workload from behind a non-DSS connected host to a DSS connected host, existing TCP applications may experience packet loss and re-transmissions as all mid-stream non-SYN data packets will be dropped. The existing TCP connections will time out and connections need to be re-established.
- The PSM can be configured with up to four vCenters. However, flow migration is only supported when vMotion occurs between hosts managed by the same vCenter.

Configuring IPsec VPN Tunnels

Overview

IPsec Standards: IPsec is a framework of IETF standards that provides data confidentiality, data integrity, and data authentication between participating peers. IPsec provides these security services at the IP layer; IPsec uses IKE to handle negotiation of protocols and algorithms based on the local policy, and generate the encryption and authentication keys to be used by IPsec. IPsec can be used to protect one or more data flows between a pair of hosts, between a pair of security gateways, or between a security gateway and a host.

IKE Standards: IKE (IKEv1 and IKEv2)—A hybrid protocol that implements Oakley and SKEME key exchanges inside the Internet Security Association and Key Management Protocol (ISAKMP) framework. While IKE is used with other protocols, its initial implementation is with the IPsec protocol. IKE provides authentication of IPsec peers, negotiates IPsec security associations, and establishes IPsec keys.

IPsec Functionality Overview

IPsec provides the following network security services. (In general, the local security policy dictates the use of one or more of these services.)

- Data confidentiality—The IPsec sender can encrypt packets before transmitting them across a network.
- Data integrity—The IPsec receiver can authenticate packets sent by the IPsec sender to ensure that the data has not been altered during transmission.
- Data origin authentication—The IPsec receiver can authenticate the source of the sent IPsec packets. This service is dependent upon the data integrity service.

IPsec provides secure tunnels between two peers, such as two routers. You define which packets are considered sensitive and should be sent through these secure tunnels, and you define the parameters that should be used to protect these sensitive packets by specifying the characteristics of these tunnels.

CX 10000 switches support IPsec VPN encryption to address customer requirements for site-to-site and site-to-cloud end-to-end encryption for data privacy and compliance concerns.

Data encryption and decryption is hardware-accelerated in the CX 10000 DSM dataplane to deliver high-rates of encryption throughput at rates of up to 400Gbps (200Gbps full-duplex) per switch.

IPsec Topology Designs

The following 3 types of IPsec deployment designs are supported:

- IPsec Active/Active with VSX
- IPsec Active/Standby with VSX
- IPsec no-HA mode (no VSX)

The high-level design for each of these modes:

IPsec Active/Active with VSX:

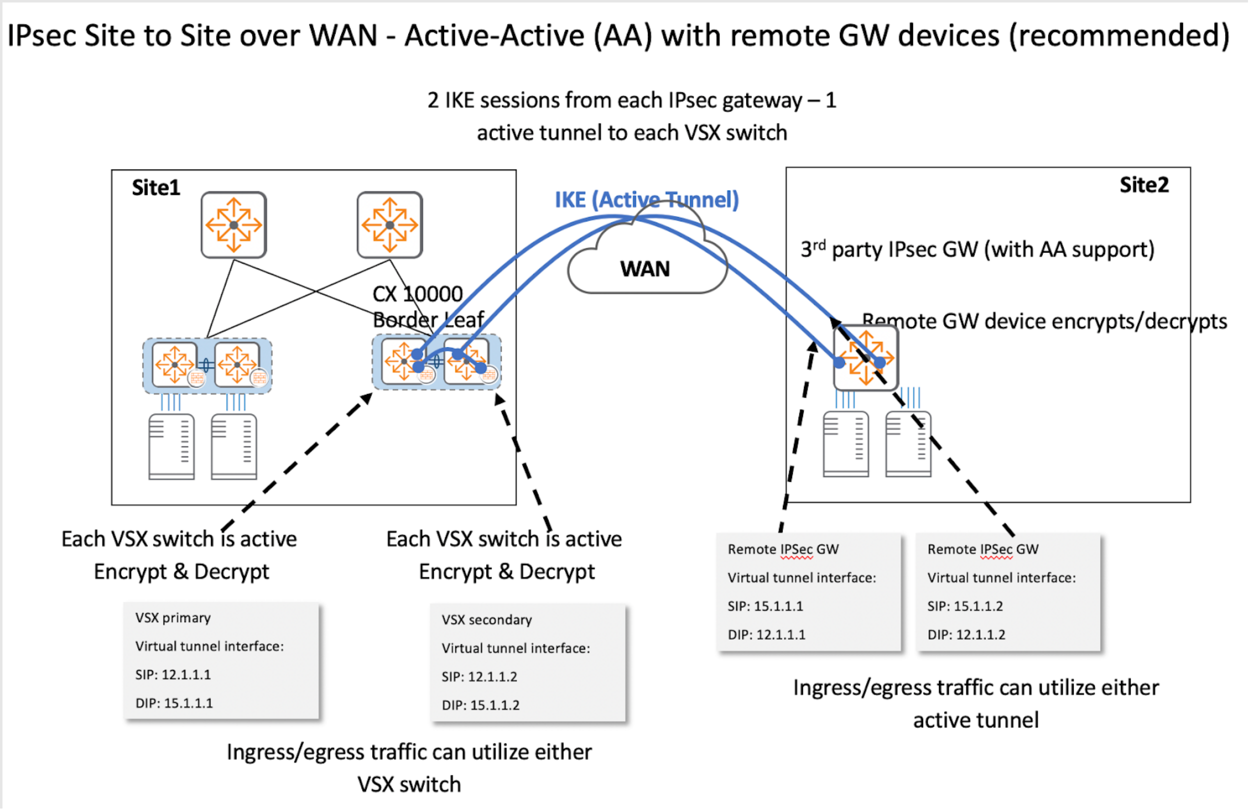
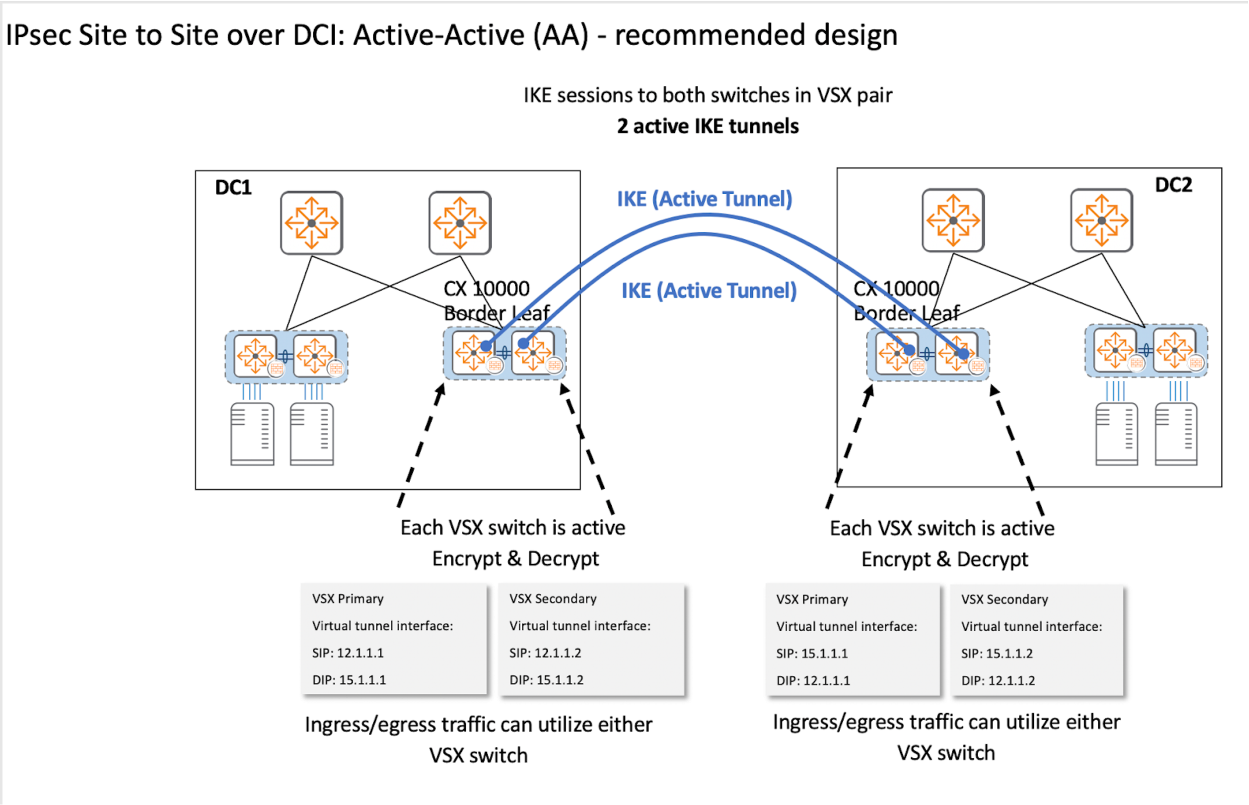

IPsec Site to Site over WAN - Active-Active (AA) with remote GW devices (recommended)
2 IKE sessions from each IPsec gateway – 1 active tunnel to each VSX switch
Site1
CX 10000 Border Leaf
IKE (Active Tunnel)
WAN
Site2
3rd party IPsec GW (with AA support)
Remote GW device encrypts/decrypts
Each VSX switch is active Encrypt & Decrypt
Each VSX switch is active Encrypt & Decrypt
VSX primary
Virtual tunnel interface:
SIP: 12.1.1.1
DIP: 15.1.1.1
VSX secondary
Virtual tunnel interface:
SIP: 12.1.1.2
DIP: 15.1.1.2
Remote IPSec GW
Virtual tunnel interface:
SIP: 15.1.1.1
DIP: 12.1.1.1
Remote IPSec GW
Virtual tunnel interface:
SIP: 15.1.1.2
DIP: 12.1.1.2
Ingress/egress traffic can utilize either active tunnel
Ingress/egress traffic can utilize either VSX switch

Figure 134.

This mode is the recommended design option with VSX. Here the remote IPsec GW device has 2 active IKE sessions with each VSX switch. Each VSX switch is capable of independently encrypting/decrypting the traffic providing a true Active/Active design. When one VSX node fails, IPsec traffic immediately switches over to the other switch in the VSX pair with minimal convergence time. This mode can only be used only when the remote IPsec GW device supports 2 active tunnels. Effectively, this mode is ideal as it provides both HA redundancy and traffic load balancing.


IPsec Site to Site over DCI: Active-Active (AA) - recommended design

IKE sessions to both switches in VSX pair
2 active IKE tunnels

Each VSX switch is active
Encrypt & Decrypt

Each VSX switch is active
Encrypt & Decrypt

VSX Primary
Virtual tunnel interface:
SIP: 12.1.1.1
DIP: 15.1.1.1

VSX Secondary
Virtual tunnel interface:
SIP: 12.1.1.2
DIP: 15.1.1.2

VSX Primary
Virtual tunnel interface:
SIP: 15.1.1.1
DIP: 12.1.1.1

VSX Secondary
Virtual tunnel interface:
SIP: 15.1.1.2
DIP: 12.1.1.2

Ingress/egress traffic can utilize either
VSX switch

Ingress/egress traffic can utilize either
VSX switch

Figure 135.

In the above example, IPsec can be used to encrypt traffic over the DCI across 2 sites.

In the current release, each switch deployed as a VSX border-leaf pair at one site can have one active tunnel to each border-leaf switch in the VSX pair on the other site across the DCI link. However, a given remote network prefix is reachable only across a single IPsec tunnel. Multiple active ECMP tunnels for the same remote network prefix on each switch are currently not supported and planned for future release.

IPsec Active/Standby with VSX

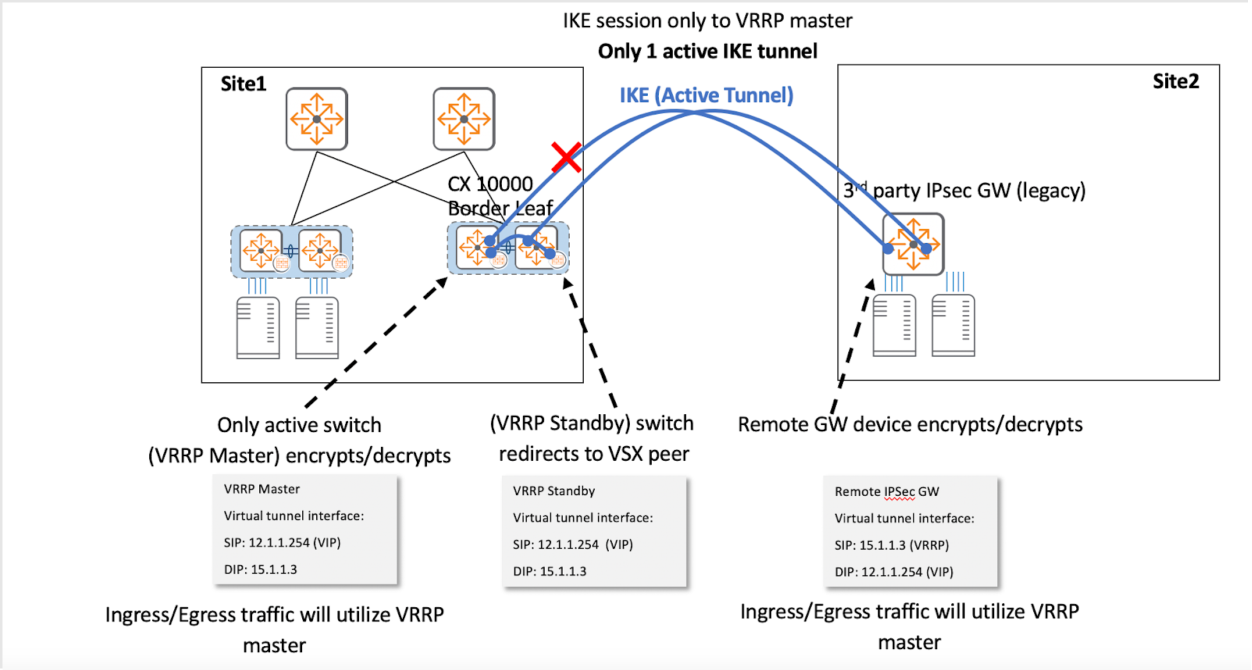

IKE session only to VRRP master
Only 1 active IKE tunnel
Site1
Site2
IKE (Active Tunnel)
CX 10000
Border Leaf
3rd party IPsec GW (legacy)
Only active switch
(VRRP Master) encrypts/decrypts
(VRRP Standby) switch
redirects to VSX peer
Remote GW device encrypts/decrypts
VRRP Master
Virtual tunnel interface:
SIP: 12.1.1.254 (VIP)
DIP: 15.1.1.3
VRRP Standby
Virtual tunnel interface:
SIP: 12.1.1.254 (VIP)
DIP: 15.1.1.3
Remote IPSec GW
Virtual tunnel interface:
SIP: 15.1.1.3 (VRRP)
DIP: 12.1.1.254 (VIP)
Ingress/Egress traffic will utilize VRRP master
Ingress/Egress traffic will utilize VRRP master

Figure 136. IPsec Site to Site: Active/Standby, for legacy remote gateway devices

This mode can be used with legacy remote IPsec GW devices that support only 1 active tunnel. In this mode, the remote IPsec GW device has 1 active IKE session with only one of the VSX nodes in the pair (the VRRP master). There is a dependency on the underlay network config to enable VRRP on the northbound uplinks SVI interface. In this mode, only one VSX node encrypts/decrypts the traffic. When the active node fails, the IPsec data-traffic will switchover to the new active node after IKE establishment. In this mode if the traffic that is subject to IPsec encryption/decryption lands on the VRRP standby switch, it is forwarded to the VRRP master switch over the ISL.

IPsec Supported Modes and Crypto Options

The following IPsec modes and crypto options are supported:

- Encapsulation Mode: Tunnel Mode
- Protocol: ESP
- IKE versions: v1, v2
- Types of Identifiers - The default identifier type is ip, which can be empty as well. If empty then the system picks up the VTI tunnel's source & destination from the Part 1 configuration as the ID. The rest of the ID types (fqdn, email, keyid) cannot be empty if specified.
    - ip
    - keyid

- fqdn
  - email
- *IKE Encryption Algorithms*:
  - aes_128
  - aes_256
  - 3des
  - cast_128
  - aes_gcm_128
  - aes_gcm_256
- *IKE Hash Algorithms*:
  - sha_256
  - sha_512
  - sha_384
- *IKE DH Group*:
  - group1
  - group2
  - group5
  - group14
  - group15
  - group19
  - group20
- *IKE Authentication Method:*
  - Only Pre-shared-key authentication is supported in this release
- *IPsec Encryption Algorithms* - Supported options are:
  - aes_gcm_128
  - aes_gcm_256
- *IPsec DH Group (To enable Perfect Forward Secrecy):*
  - group1
  - group2
  - group5
  - group14
  - group15
  - group19
  - group20

Order of Operation with Service Chaining of FW, NAT and IPsec services

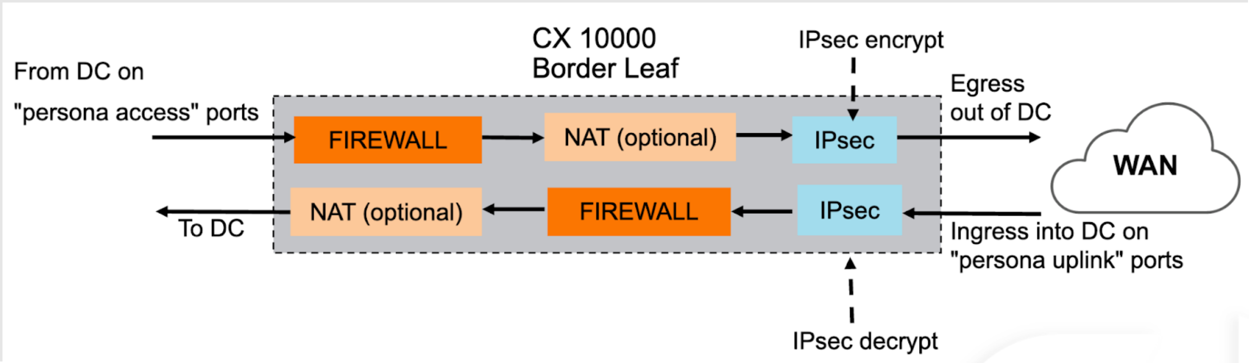


Figure 137.

As shown in the diagram above:

- For traffic entering the switch from inside the DC fabric on the “persona access” ports, here is the order of operations:
  - Firewall inspection
  - NAT translation (optional)
  - IPsec encryption
- For traffic entering the switch from outside the DC on the “persona uplink” ports, here is the order of operations:
  - IPsec decryption
  - Firewall inspection
  - NAT translation (optional)

For both directions:

- Explicit FW policy is NOT mandatory, implicit FW policy is "permit all if no explicit policy applied
- NAT is optional

IPsec Configuration Workflow

At the high level, configuring IPsec VPN tunnels on the CX 10000 switches involves two configuration operations:

- Part 1: Switch-side configuration
    - Boot the switch in L3-core profile
    - Configure the persona for the interfaces on the switch
    - Configure the required VRF, and SVI interfaces in that VRF
    - Configure the IPsec Virtual Tunnel Interface (VTI) with the correct tunnel source, tunnel destination address and VRF
- Part 2: PSM-side configuration
    - Configure the IPsec policy
    - Attach the configured policy to required VRFs

AFC can be used to automate the workflow end to end, and provision the required configuration on both switch and PSM.

Prerequisite Configuration On the Switch

Step 1: Boot the switch in the L3-core profile
Issue the following AOS-CX command on the DSS:

```
switch(config)# profile l3-core
Save this config and reboot the switch for the
changes to take effect
switch# copy running-config startup-config
Copying configuration: [Success]
```

Reboot the switch. After the switch comes online, make sure it is in the “L3-core” profile by issuing the following command:

```
switch(config)# show profiles current
Current Profile
---------------
L3-core
```

Step 2: Configuring interface persona

On external facing interfaces (WAN), configure “persona uplink”:

```
switch(config)# int 1/1/3
switch(config-if)# persona uplink
```

On interfaces facing the internal DC fabric, configure “persona access”:

```
switch(config)# int 1/1/4
switch(config-if)# persona access
```

Step 3: Configuring the tenant VRF

```
switch(config)# vrf tpm-vrf1
switch(config-vrf)#
```

Step 4a: Configure the external upstream L3 SVI interface per VRF

```
switch(config)# int vlan 30
switch(config-if)# vrf attach tpm-vrf1
switch(config-if)# ip address 61.1.30.5/24
```

⚠ ***Note**: Only L3 SVI interfaces can be used as uplinks in this release. L3 sub-interface uplink support is planned for a future release.*

Step 4b: Configure the DC internal interface

For non-VxLAN/legacy DC fabric, configure L3 SVI interface:

```
switch(config)# int vlan 20
switch(config-if)# vrf attach tpm-vrf1
switch(config-if)# ip address 61.1.20.5/24
```

For VxLAN DC fabric, configure the native L3 interface:

```
switch(config)# int 1/1/4
switch(config-if)# ip address 61.1.20.5/24
```

Configure IPsec VPN Tunnel on standalone switch (no-HA mode)

Follow the steps below to configure an IPsec tunnel on a standalone switch, also called “IPsec no-HA mode”, with the remote peer.

Notes:

1. *IPsec without HA is only recommended for POC/test purposes, not for production use.*
2. *Prior to configuring IPsec tunnels, basic networking, such as VRFs and VLAN interfaces, needs to be properly configured and functional.*

Step 1: Configure the IPsec VTI tunnel interface in AOS-CX on the CX 10000 switch. Please note:

- The tunnel source IP needs to be the public-facing interface’s IP. This is the IP of the upstream L3 SVI interface
- Both the tunnel source and destination can NOT be in the default VRF. They need to be reachable in the same VRF as the tunnel interface itself, which is *tpm-vrf1* in the following example.

```
switch(config)# interface tunnel 1 mode ipsec ipv4
vrf attach tpm-vrfl
source ip 61.1.30.5
destination ip 2.2.2.3
no shutdown
```

Step 2: Configure the static route for the remote network pointing to the VTI interface

```
switch(config)# ip route 3.3.3.0/24 tunnel 1 vrf
tpm-vrf1
```

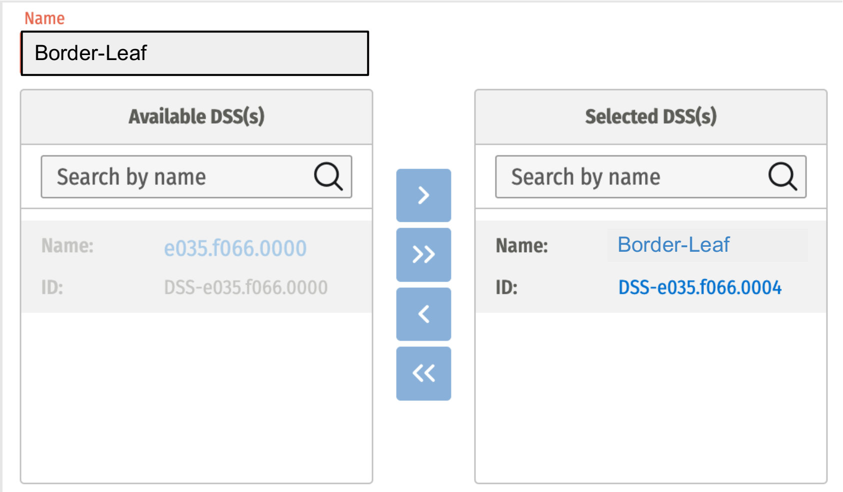
In the above example, this static route needs to be advertised into the DC fabric using IGP (redistribute static). The traffic destined to prefix “3.3.3.0/24” will be encrypted and sent over the tunnel “tunnel1”. In this release reachability to routes behind a remote IPsec GW needs to be defined statically. Support for dynamically learning these tunnel routes over BGP is planned for a future release.

PSM Side Configuration

Step 1: Defining a policy distribution target (PDT) containing the border-leaf DSS switch

IPsec policies are pushed to only those DSS devices that are in “L3-core” profile. A PDT needs to be defined containing the DSS border-leaf switch in L3-core profile.

Step 2: In the PSM, create the PDT and add the appropriate DSS. Note this DSS will terminate the VPN tunnel:



Name

Border-Leaf

Available DSS(s)	
Name:	e035.f066.0000
ID:	DSS-e035.f066.0000

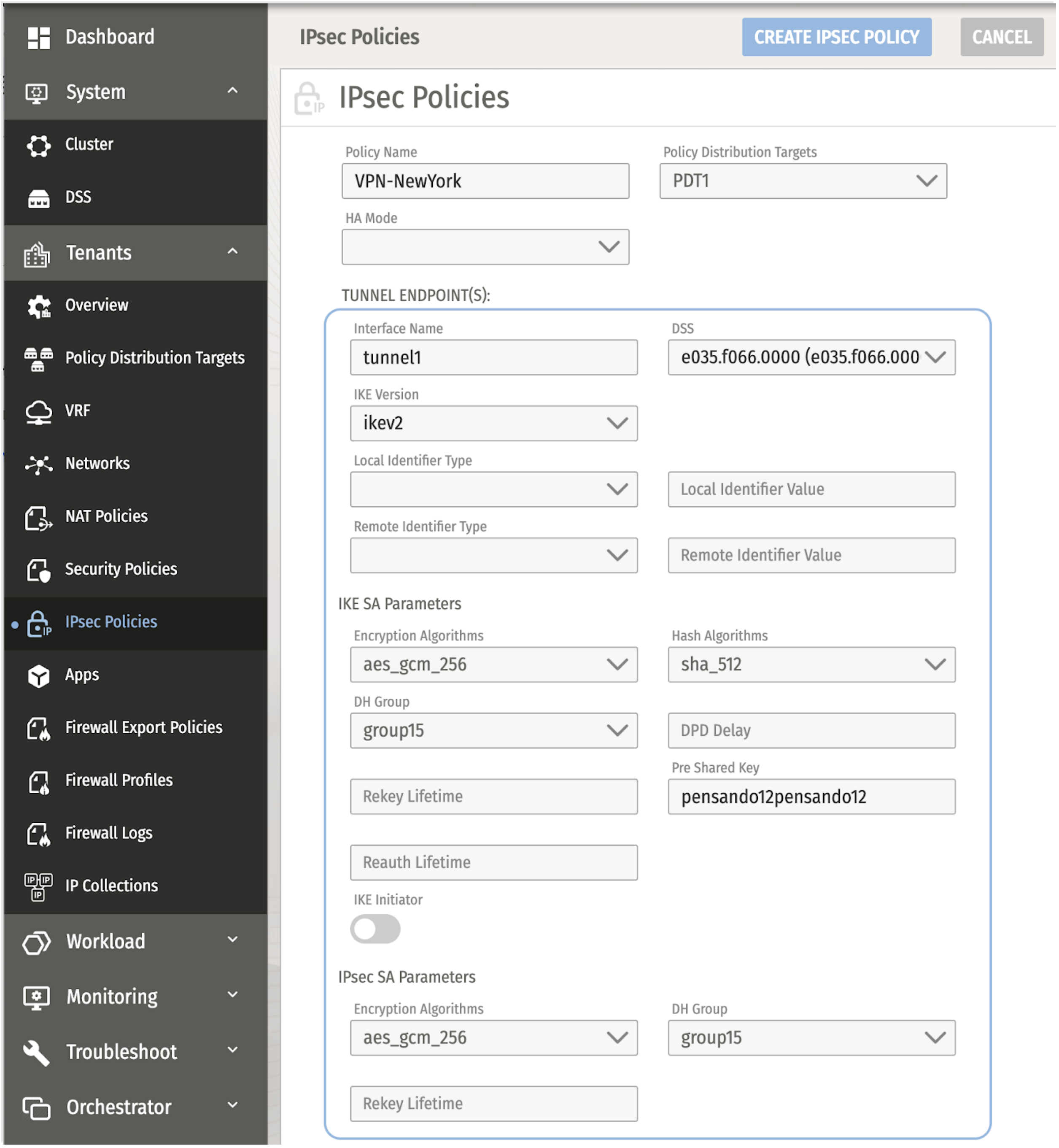
Selected DSS(s)	
Name:	Border-Leaf
ID:	DSS-e035.f066.0004

Figure 138.

Step 3: In the PSM, create the IPsec policy with desired crypto parameters

Select `no_ha` for the HA Mode drop down menu.

The *Local Identifier* and *Remote Identifier* represent the parameters IDi and IDr in the IKE RFC.



IPsec Policies

Policy Name: VPN-NewYork

Policy Distribution Targets: PDT1

HA Mode:

TUNNEL ENDPOINT(S):

Interface Name: tunnel1

DSS: e035.f066.0000 (e035.f066.000

IKE Version: ikev2

Local Identifier Type:

Local Identifier Value

Remote Identifier Type:

Remote Identifier Value

IKE SA Parameters

Encryption Algorithms: aes_gcm_256

Hash Algorithms: sha_512

DH Group: group15

DPD Delay

Rekey Lifetime

Pre Shared Key: pensando12pensando12

Reauth Lifetime

IKE Initiator: ○

IPsec SA Parameters

Encryption Algorithms: aes_gcm_256

DH Group: group15

Rekey Lifetime

Figure 139.

NOTE:

- *The tunnel interface name given in the IPsec policy needs to match the tunnel name configured on the switch. Mismatching tunnel names will fail the distribution of the IPsec policy to the desired DSS.*
- *IPsec VPN for each remote site requires a unique IPsec policy to be configured, so from the operation perspective, it is recommended to have the policy name match the remote site that the tunnel is meant to connect to.*
- *Only Pre-shared key is supported for IKE authentication in this release. Authentication based on digital certificates are planned for a future release.*
- *If "prefer_ikev2_support_ikev1" is selected for IKE version setting,*
  - *If "IKE Initiator" is disabled, meaning responder only, CX 10000 will accept either IKE v1 or v2 negotiation.*
  - *If "IKE Initiator" is enabled, CX 10000 will only initiate IKEv2 negotiation. If the remote end does not support IKEv2, the negotiation will fail*

Step 4: In the PSM, attach the configured IPsec policy to the right VRF from step 1.

> ***Note***: *Multiple IPsec policies can be attached to the same VRF.*

This completes the configuration workflow for setting up an IPsec tunnel on the DSS in standalone mode (no HA mode) with remote IPsec GW device. Follow the required configuration steps on the remote IPsec GW to bring up IPsec configurations.

After this setup, use the following commands to verify the IPsec VTI tunnel status

```
switch# show interface tunnel 1


Interface tunnel1 is up
 Admin state is up
 tunnel type  IPsec IPv4
 tunnel source IPv4 address 61.1.30.5
 tunnel destination IPv4 address 2.2.2.3
```

The next section will cover the HA options available to deploy IPsec.

VSX Requirement For IPsec

The IPsec implementation on DSS is built on top of the existing stateful firewall implementation of the DSS. Packets that are subject to encryption are evaluated by stateful FW prior to encryption. Packets that are subject to decryption are evaluated by stateful FW after decryption. For both Active/Active and Active/Standby, it is required that the two switches are configured to be part of a VSX pair. Since stateful firewall is always involved and TCP traffic flows can be asymmetric in nature (a SYN packet that needs to be decrypted can land on switch A and a SYN-ACK packet that needs to be encrypted lands on switch B, the 2 switches A and B need to be in a VSX pair so that TCP stateful flows can be flow-synced between the 2 switches in VSX pair and a true Active/Active stateful FW functionality can be provided with both IPsec encryption and decryption support. In summary, VSX provides the Active/Active stateful FW in conjunction with redundancy and load-balancing options for IPsec encrypt/decrypt traffic.

On the CX 10000 switch, VSX is always required in both IPsec HA modes: Active/Active or Active Standby, for:

- *VSX Flow synchronization* - needed for asymmetric TCP traffic, when deploying across two border leaf device. VSX is needed to synchronize all TCP stateful traffic flows between the two switches in a VSX pair. In case one device fails, the other device can take over all flows seamlessly.
- *Traffic Redirection* - In the data center deployments, it is very common for outbound traffic to land at the CX 10000 switch that doesn’t possess the active IPsec tunnel to reach the remote site. VSX is then needed to redirect the traffic through the interconnecting ISL links to the other CX 10000 switch that does have the active IPsec tunnel for encryption.

Configure IPsec VPN With Active Standby Failover

Many legacy remote VPN devices support only one active IKE/IPsec tunnel. IPsec Active/Standby) is the recommended option for such remote devices. IPsec Active/Standby failover is expected to have a higher data-path convergence times as compared to IPsec Active/Active. This is because the new IKE negotiation is initiated after the failover event and IKE needs to re-establish itself. Note: prior to configuring IPsec tunnels, basic networking needs to be configured properly and functioning, such as VRFs, VLAN interfaces and VSX redundancy, etc. In this mode it is mandatory to have VRRP enabled on the upstream SVI interface. An IPsec Active/Standby tunnel will only be UP on the VRRP master switch. A switch will only advertise IPsec tunnel routes if the next-hop tunnel interface is UP. Because of this, the prefixes behind the remote IPsec GW will be advertised only by the VRRP master.

In steady state, the IPsec active switch is the VRRP master and the IPsec standby switch is the VRRP standby switch. VRRP states dictate the IPsec active/standby switch roles. IPsec failover from the active switch to the standby switch can happen under the following scenarios:

- *VRRP Status Change* - On a CX 10000 switch, IPsec failover is implemented to monitor the VRRP status change. If the VRRP status changes for any reason, such as link flapping, priority configuration change, etc, IPsec failover will be triggered.
- *Active Device Reboot/Power-down* - If the active IPsec device reboot or power down for any reason, IPsec will automatically fail over to the standby device.
- *ISL Link down* - The ISL link is configured to interconnect the two CX 10000 switches for VSX pairing. If the ISL link goes down for any reason, IPsec needs to go with the VSX primary switch:
  - If the VSX primary switch happens to be the VRRP active device, IPsec tunnel will not be affected.
  - If the VSX primary switch happens to be the VRRP standby device, IPsec tunnel will fail over to the standby device and this device will also become VRRP active.

NOTE: *Tunnel interface admin shut will not trigger an IPsec Active/Standby failover, since there is no VRRP state change associated with that trigger.*

Upon failover, the IPsec tunnel will be re-negotiated with the standby (now active) device. Once the new IPsec tunnel is up and running, the standby (now active) device will start advertising the remote network into the data center fabric, as shown in the diagram below.

Follow the steps below to configure IPsec VPN with active standby failover.

Step 1: Configure VRRP on the upstream SVI interface of both CX 10000 switches

Switch 1:

```
switch-1 (config)# interface vlan 30
vrf attach tpm-vrf1
vsx active-forwarding
ip address 61.1.30.5/24
vrrp 30 address-family ipv4
   address 61.1.30.254 primary
   no shutdown
   exit
vrrp dual-active-forwarding
```

Switch 2:

```
switch-2 (config)# interface vlan 30
description NB SVI
vrf attach tpm-vrf1
vsx active-forwarding
ip address 61.1.30.6/24
vrrp 30 address-family ipv4
   address 61.1.30.254 primary
   no shutdown
   exit
vrrp dual-active-forwarding
```

Step 2: Configure the IPsec VTI tunnel interface on both switches

Switch 1:

Note that the source address of the IPsec VTI interface is the VRRP virtual IP address configured in step 1.

```
switch-1 (config)# interface tunnel 1 mode ipsec
ipv4
vrf attach tpm-vrf1
source ip 61.1.30.254
destination ip 2.2.2.3
no shutdown
```

Switch 2:

Note that the source address of the IPsec VTI interface is the VRRP virtual IP address configured in step 1.

```
switch-2 (config)# interface tunnel 1 mode ipsec
ipv4
vrf attach tpm-vrf1
source ip 61.1.30.254
destination ip 2.2.2.3
no shutdown
```

Step 3: Configure the static route for the remote network pointing to the VTI interface on both switches

Switch 1:

```
switch-1 (config)# ip route 3.3.3.0/24 tunnel1
vrf tpm-vrf1
```

Switch 2:

```
switch-2 (config)# ip route 3.3.3.0/24 tunnel1
vrf tpm-vrf1
```

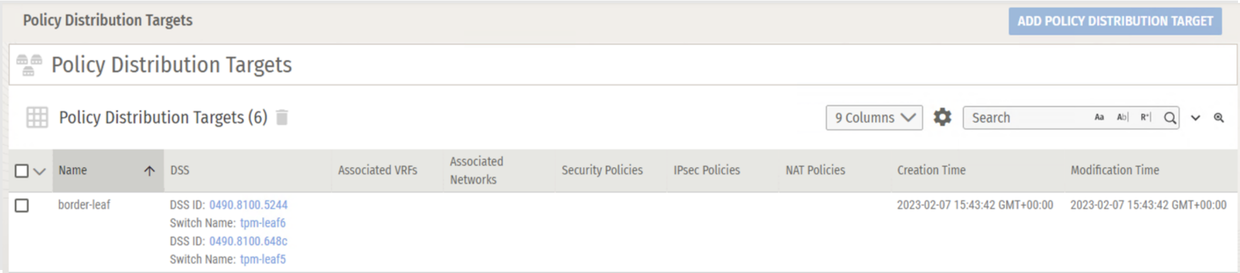
Step 4: Configure the dynamic routing protocol (such as OSPF) on both switches and advertise the static routes for the remote network into the data center fabric

```
router ospf 1 vrf tpm-vrf1
    router-id 16.16.16.16
    max-metric router-lsa include-stub on-startup
    bfd all-interfaces
    maximum-paths 32
    redistribute static
    area 0.0.0.0
```

As described previously, only the VRRP active switch will be advertising the remote network into the data center fabric. Preventing the VRRP standby switch from advertising the same remote network is achieved by keeping its VTI tunnel status down, as shown in the following example:

```
switch-2# show interface tunnel 1


Interface tunnel1 is down
 Admin state is up
 Description: A/S VPN Tunnel
 tunnel type  IPsec IPv4
 tunnel source IPv4 address 61.1.30.254
 tunnel destination IPv4 address 2.2.2.3
```

Step 5: in the PSM, configure the Policy Distribution Target as described in the previous chapter.



Policy Distribution Targets

Policy Distribution Targets

Policy Distribution Targets (6)

Name	DSS	Associated VRFs	Associated Networks	Security Policies	IPsec Policies	NAT Policies	Creation Time	Modification Time
border-leaf	DSS ID: 0490.8100.5244 Switch Name: tpm-leaf6 DSS ID: 0490.8100.648c Switch Name: tpm-leaf5						2023-02-07 15:43:42 GMT+00:00	2023-02-07 15:43:42 GMT+00:00

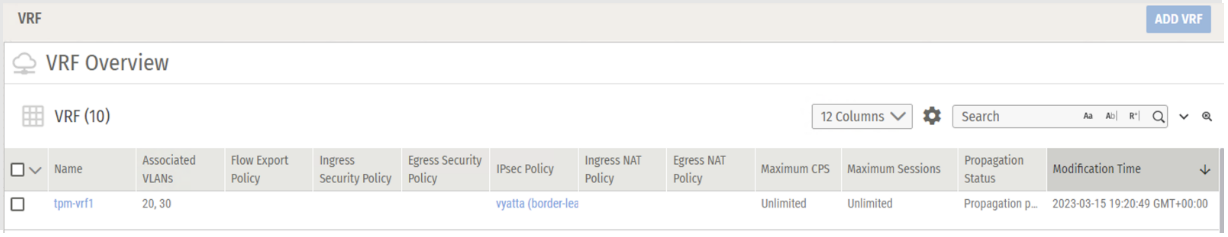
Figure 140.

Step 6: in the PSM, configure the IPsec policy as shown in screenshots below

Note: once active_passive is selected for the HA mode, two tunnel configurations are required for the policy, one for each switch.

In the case of Active/Standby deployment, always enable the "IKE Initiator" setting in the policy configuration for faster convergence in the event of failover. Without enabling the "IKE Initiator" setting, when failover happens, the remote peer has to wait for the DPD timer to expire (10+ seconds delay) to renegotiate the tunnel.

Step 7: In the PSM, attach the configured IPsec Policy to the VRF as described in the previous chapter.



VRF

VRF Overview

VRF (10)

Name	Associated VLANs	Flow Export Policy	Ingress Security Policy	Egress Security Policy	IPsec Policy	Ingress NAT Policy	Egress NAT Policy	Maximum CPS	Maximum Sessions	Propagation Status	Modification Time
tpm-vrf1	20, 30				vyatta (border-lea			Unlimited	Unlimited	Propagation p...	2023-03-15 19:20:49 GMT+00:00

Figure 141.

Configure IPsec VPN With Active/Active Failover

This mode is the recommended mode if the remote IPsec gateway supports 2 active IKE/IPsec tunnels. With IPsec Active/Active, each switch in VSX pair can encrypt/decrypt traffic. This provides both redundancy/HA as well as traffic load-balancing at the same time.

IPsec Active/Active failover has better convergence times compared to IPsec Active/Standby failover. This is because there is no need to set up new IKE sessions after the failover event. Follow the steps below to configure IPsec VPN with Active/Active failover.

Note: prior to configuring IPsec tunnels, basic networking needs to be configured properly and functioning, such as VRFs, VLAN interfaces and VSX redundancy.

As shown in the following diagram, both CX 10000 switches will be advertising the remote network into the data center fabric.

Any of the CX 10000 switches may stop propagating the remote network into the fabric due to the following reasons:

- *Device Reboot/Power-down* - If the CX 10000 switch reboots or powers down for any reason, its IPsec will go down as well.
- The uplink from CX 10000 to the Gateway fails, which causes the IPsec tunnel to go down
- The ISL link between the VSX pairs goes down - the VSX secondary CX 10000 will shutdown all its ports and the IPsec tunnel. The VSX primary CX 10000 switch takes over all traffic

When the failover happens, only the remaining active CX 10000 switch will be advertising the remote network into the data center fabric, as shown in the diagram below.

Follow the steps below to configure IPsec VPN with Active/Active failover.

Step 1: Configure the VLAN interfaces of both CX 10000 switches

> ***Note:*** *VRRP is not needed for Active/Active failover.*

Switch 1:

```
switch-1 (config)# interface vlan 330
vrf attach tpm-vrf1
vsx active-forwarding
ip address 61.1.33.5/24
```

Switch 2:

```
switch-2 (config)# interface vlan 330
vrf attach tpm-vrf1
vsx active-forwarding
ip address 61.1.33.6/24
```

Step 2: Configure the IPsec VTI tunnel interface on both switches

Note that the source address of the IPsec VTI interface is the physical IP address of the sourcing VLAN interface configured in step 1. Only the SVI's primary IP address is supported.

Switch 1:

```
switch-1 (config)# interface tunnel 2 mode ipsec
ipv4
vrf attach tpm-vrf1
source ip 61.1.33.5
destination ip 2.2.2.4
no shutdown
```

Switch 2:

Note that the source address of the IPsec VTI interface is the physical IP address of the sourcing VLAN interface configured in step 1.

```
switch-2 (config)# interface tunnel 2 mode ipsec
ipv4
vrf attach tpm-vrf1
source ip 61.1.33.6
destination ip 2.2.2.4
no shutdown
```

Step 3: Configure the static route for the remote network pointing to the VTI interface on both switches

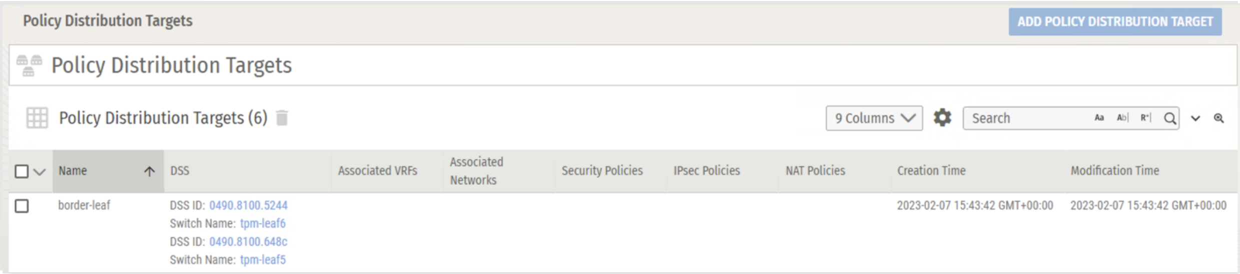
Switch 1:

```
switch-1 (config)# ip route 3.3.3.3/32 tunnel2
vrf tpm-vrf1
```

Switch 2:

```
switch-2 (config)# ip route 3.3.3.3/32 tunnel2
vrf tpm-vrf1
```

Step 4: in the PSM, configure the Policy Distribution Target as described earlier in this guide.

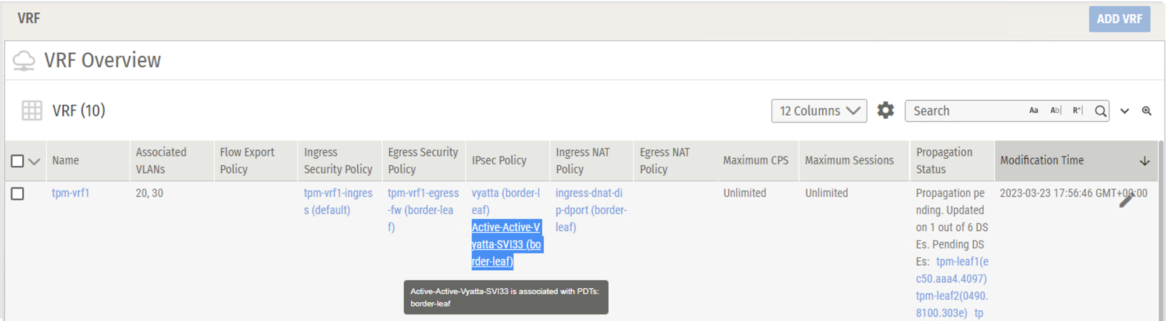

Policy Distribution Targets
ADD POLICY DISTRIBUTION TARGET
Policy Distribution Targets
Policy Distribution Targets (6)
9 Columns
Search

□	Name	DSS	Associated VRFs	Associated Networks	Security Policies	IPsec Policies	NAT Policies	Creation Time	Modification Time
□	border-leaf	DSS ID: 0490.8100.5244 Switch Name: tpm-leaf6 DSS ID: 0490.8100.648c Switch Name: tpm-leaf5						2023-02-07 15:43:42 GMT+00:00	2023-02-07 15:43:42 GMT+00:00

Step 5: in the PSM, configure the IPsec policy as shown below

> ***Note**: once `active_active` is selected for the HA mode, two tunnel configurations are required for the policy, one for each switch.*

Step 6: In the PSM, attach the configured IPsec policy to the VRF as described in the previous chapter.



VRF Overview

VRF (10)

Name	Associated VLANs	Flow Export Policy	Ingress Security Policy	Egress Security Policy	IPsec Policy	Ingress NAT Policy	Egress NAT Policy	Maximum CPS	Maximum Sessions	Propagation Status	Modification Time
tpm-vrf1	20, 30		tpm-vrf1-ingress (default)	tpm-vrf1-egress-fw (border-leaf)	vyatta (border-leaf) Active-Active-Vyatta-SVI33 (border-leaf)	ingress-dnat-dip-dport (border-leaf)		Unlimited	Unlimited	Propagation pending. Updated on 1 out of 6 DSEs. Pending DSEs: tpm-leaf1(ec50.aaa4.4097) tpm-leaf2(0490.8100.303e) tp	2023-03-23 17:56:46 GMT+00:00

Active-Active-Vyatta-SVI33 is associated with PDTs: border-leaf

QoS Support over IPsec VPN Tunnels

QoS classification

QoS classification offers the ability to classify data packets into non-default queues in the switch for prioritized treatment if necessary. This classification always happens at the first ingress port.

Encrypt direction:

The traffic entering the switch from inside the DC, can be classified on the incoming downstream interface into non-default queues and post-encrypted packets take this queue when egressing out on the uplink port.

Decrypt direction:

The traffic entering the switch from outside the data center can be classified on the incoming upstream SVI interface into non-default queues and post-decrypted packets take the same queue when egressing out on the downlink SVI interface.

Define an ACL based class to match on SIP, DIP, IPsec ESP IP protocol type packets: and classify them into queue 7:

```
switch# class ip ipsec1
     10 match esp 50.1.28.0/255.255.255.0 50.1.35.0/255.255.255.0

switch# policy ipsec1
     10 class ip ipsec1 action local-priority 7
```

By default, local-priority 7 is mapped to Q7.

Inner DSCP to outer DSCP copy

In the encrypt path, the DSCP from the inner IP header is copied over to the DSCP field in the outer IP header (tunnel DSCP). In the decrypt path, there is no copy, the inner DSCP value is retained post-decryption. This is turned on by default.

QoS Policing

Incoming traffic on each tunnel can be policed, by applying an ingress policer on the upstream SVI interface. An ACL must be defined matching on tunnel SIP, DIP and IPsec ESP protocol. The policer uses the ACL for match and it is applied on the SVI. This offers a way to police on a per-tunnel basis.

Policing is applicable only for ingress IPsec traffic and is not supported for egress traffic.

```
switch# policy pol_ip_ipsec1 10 class ip ipsec1 action cir kbps
5000000 cbs 30000 exceed drop
```

The above policer config polices traffic at the rate of 5Gbps for all traffic matching the ACL-based class “ipsec1” which is defined by the following ACL:

```
switch# class ip ipsec1
     10 match esp 50.1.28.0/255.255.255.0 50.1.35.0/255.255.255.0
```

Apply the policer on the uplink SVI interface:

```
switch# int vlan1001 >>> Uplink SVI
switch# apply policy pol_ipsec1 routed-in
```

QoS Shaping

IPsec traffic can be classified into non-default queues and shapers can be used on a per-port per-queue level to shape IPsec encrypted traffic.

```
// Define the shaping profile
qos schedule-profile ipsec1
     strict queue 0
     strict queue 1
     strict queue 2
     strict queue 3
     strict queue 4
     strict queue 5
     strict queue 6
     strict queue 7 max-bandwidth 10000000 >>> Shaped to 10 Gbps
```

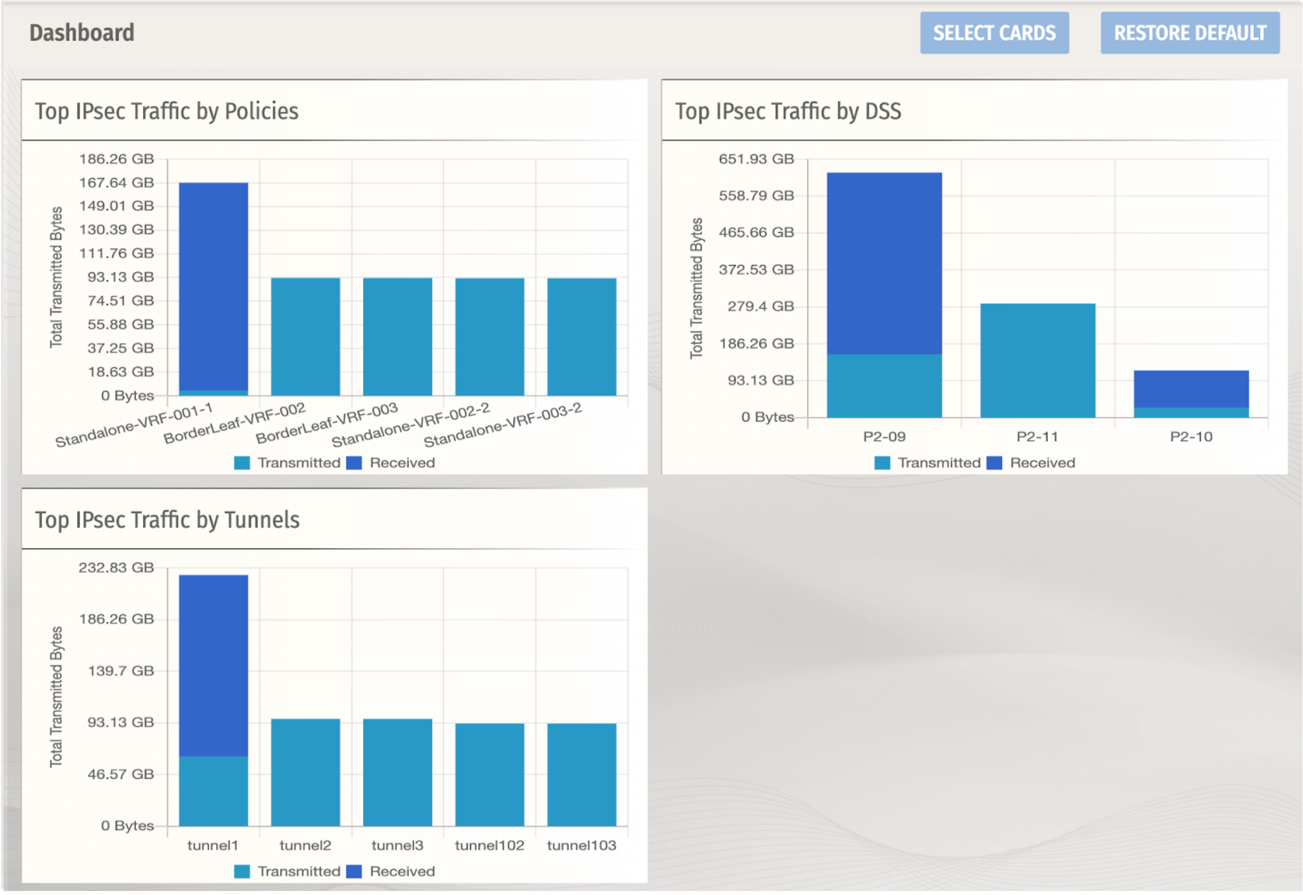
The above config defines a shaper on queue 7 to shape traffic to max BW of 10Gbps.

```
// Apply the defined shaper on the uplink port:

int 1/1/49 >>> Schedule profile on physical uplink port.
     apply qos schedule-profile ipsec1
```

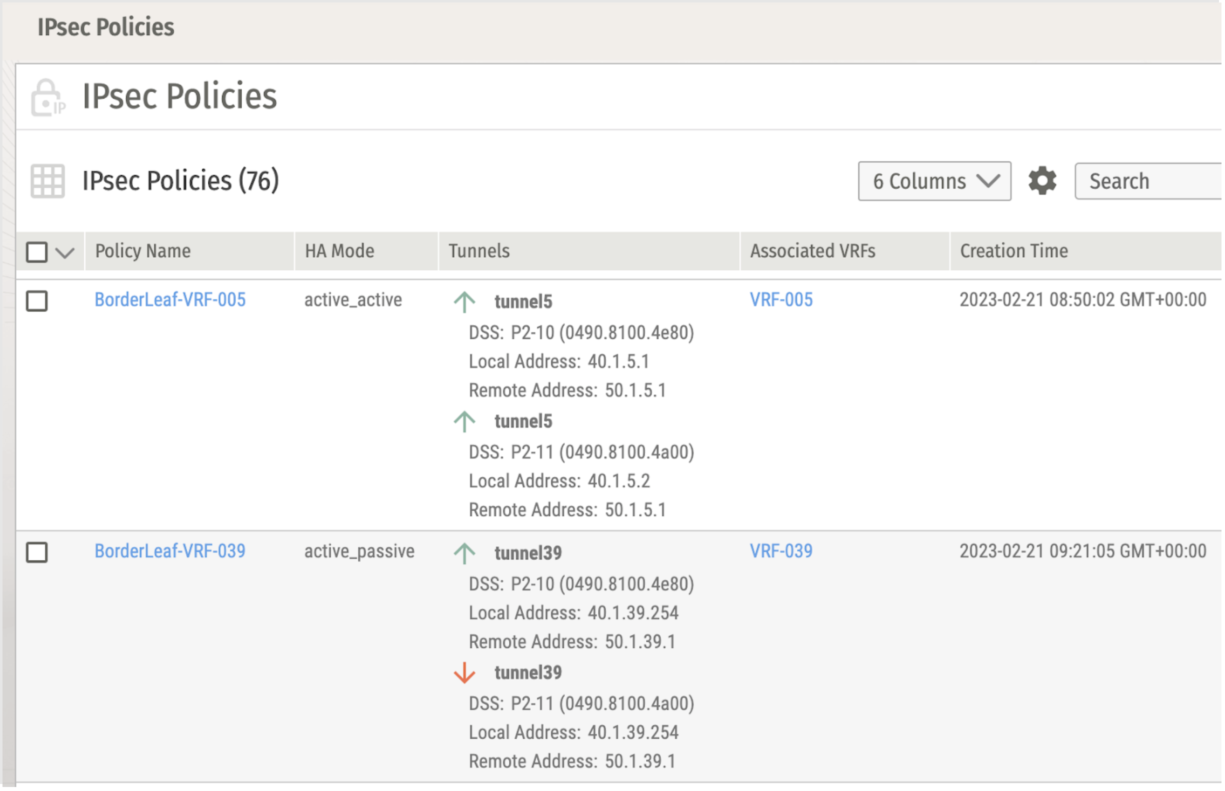
Monitor IPsec VPN Tunnels with the PSM

The PSM Dashboard can report on top IPsec stats, such as:

- Top 5 policies by IPsec traffic volume in total bytes
- Top 5 switches by IPsec traffic volume in total bytes
- Top 5 tunnels by IPsec traffic volume in total bytes



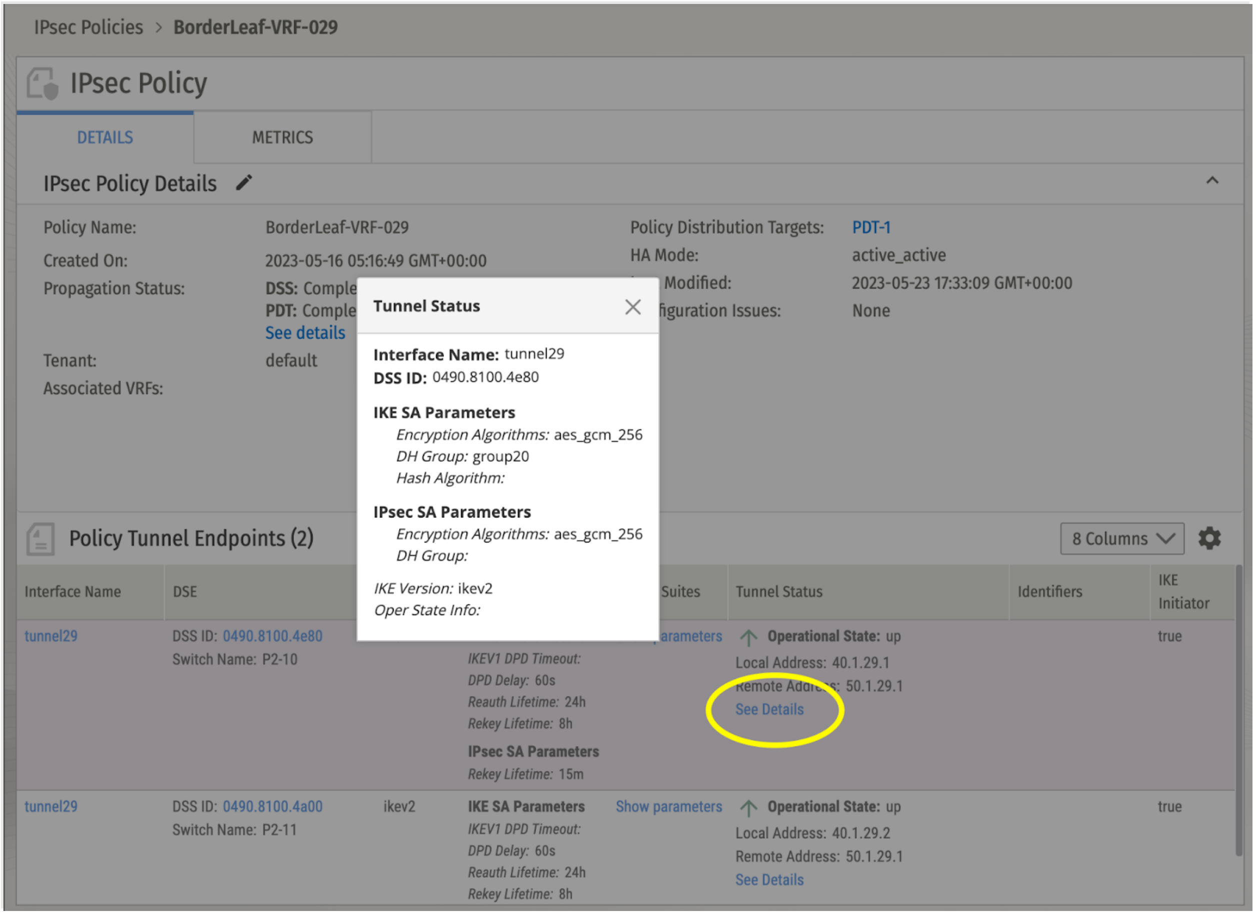
Admins can monitor the tunnel operational status under the IPsec policy page:



□	Policy Name	HA Mode	Tunnels	Associated VRFs	Creation Time
□	BorderLeaf-VRF-005	active_active	↑ **tunnel5** DSS: P2-10 (0490.8100.4e80) Local Address: 40.1.5.1 Remote Address: 50.1.5.1 ↑ **tunnel5** DSS: P2-11 (0490.8100.4a00) Local Address: 40.1.5.2 Remote Address: 50.1.5.1	VRF-005	2023-02-21 08:50:02 GMT+00:00
□	BorderLeaf-VRF-039	active_passive	↑ **tunnel39** DSS: P2-10 (0490.8100.4e80) Local Address: 40.1.39.254 Remote Address: 50.1.39.1 ↓ **tunnel39** DSS: P2-11 (0490.8100.4a00) Local Address: 40.1.39.254 Remote Address: 50.1.39.1	VRF-039	2023-02-21 09:21:05 GMT+00:00

***Note**: In the case of Active/Active, both tunnels should be UP and in the case of Active/Standby, only the tunnel on the VRRP master node will be UP and the VRRP standby will be DOWN.*

Admins may look for the individual tunnel status information in the IPsec Policy Details page, by clicking on “Tunnel Status Details”:



Policy Name: BorderLeaf-VRF-029
Created On: 2023-05-16 05:16:49 GMT+00:00
Propagation Status: **DSS:** Comple
PDT: Comple
See details
Tenant: default
Associated VRFs:

Policy Distribution Targets: PDT-1
HA Mode: active_active
Modified: 2023-05-23 17:33:09 GMT+00:00
figuration Issues: None

Tunnel Status

Interface Name: tunnel29
DSS ID: 0490.8100.4e80

IKE SA Parameters
Encryption Algorithms: aes_gcm_256
DH Group: group20
Hash Algorithm:

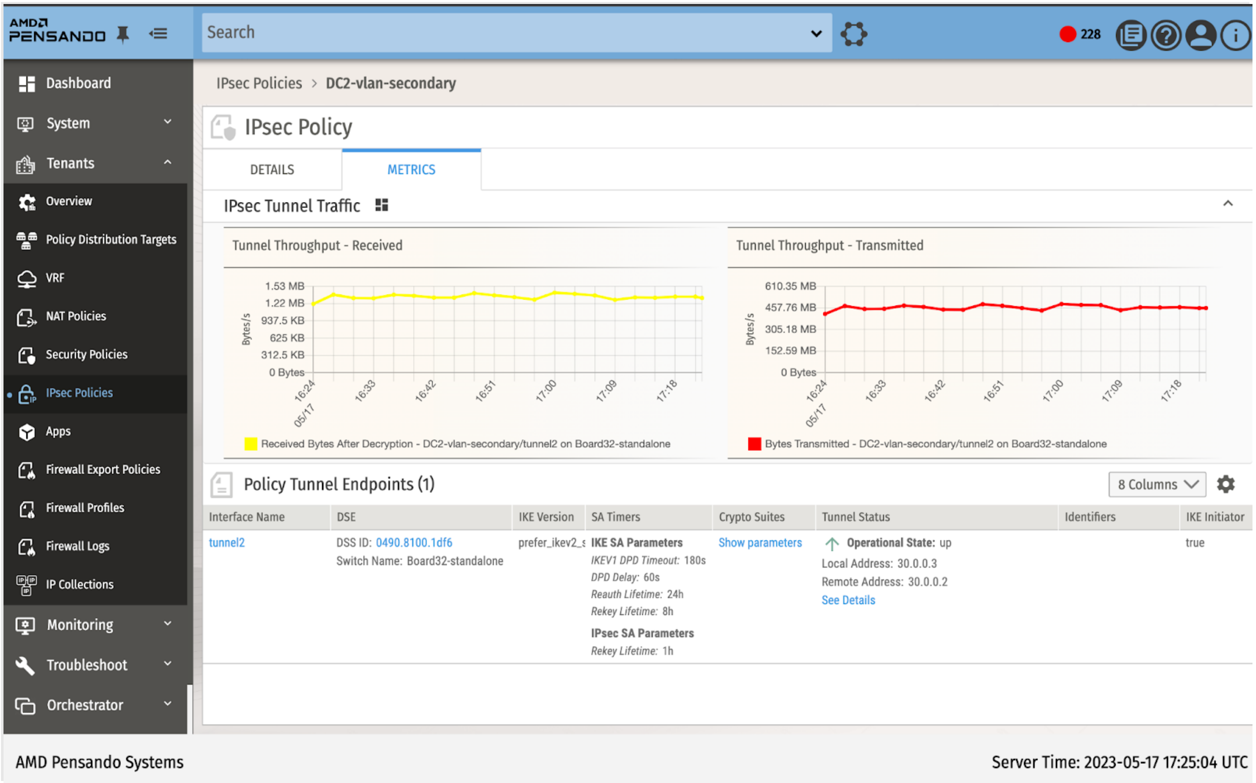
IPsec SA Parameters
Encryption Algorithms: aes_gcm_256
DH Group:

IKE Version: ikev2
Oper State Info:

Policy Tunnel Endpoints (2)

Interface Name	DSE			Suites	Tunnel Status	Identifiers	IKE Initiator
tunnel29	DSS ID: 0490.8100.4e80 Switch Name: P2-10		*IKEV1 DPD Timeout:* *DPD Delay:* 60s *Reauth Lifetime:* 24h *Rekey Lifetime:* 8h **IPsec SA Parameters** *Rekey Lifetime:* 15m	arameters	**Operational State:** up Local Address: 40.1.29.1 Remote Addr s: 50.1.29.1 See Details		true
tunnel29	DSS ID: 0490.8100.4a00 Switch Name: P2-11	ikev2	**IKE SA Parameters** *IKEV1 DPD Timeout:* *DPD Delay:* 60s *Reauth Lifetime:* 24h *Rekey Lifetime:* 8h	Show parameters	**Operational State:** up Local Address: 40.1.29.2 Remote Address: 50.1.29.1 See Details		true

For an individual tunnel, admins can monitor its health information such as throughput under the IPsec policy detail page:

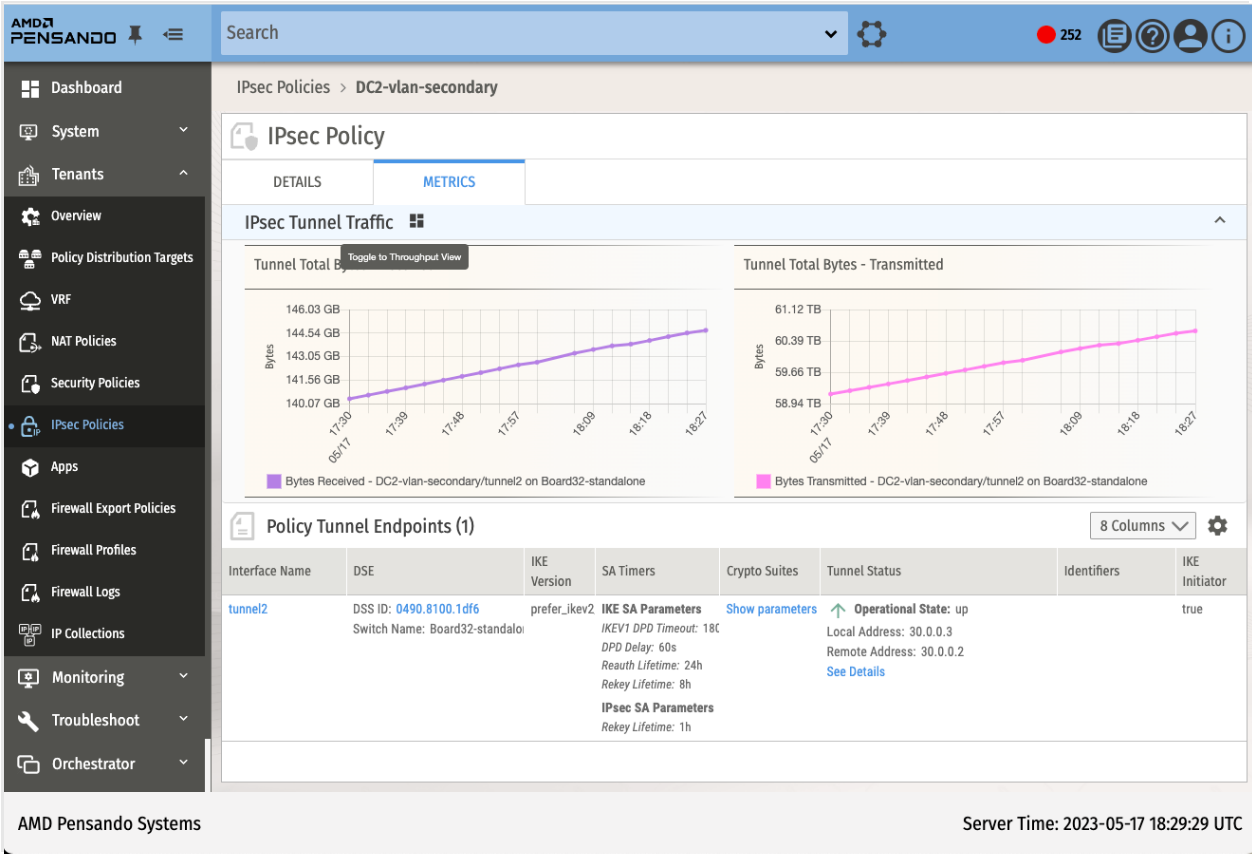


Interface Name	DSE	IKE Version	SA Timers	Crypto Suites	Tunnel Status	Identifiers	IKE Initiator
tunnel2	DSS ID: 0490.8100.1df6 Switch Name: Board32-standalone	prefer_ikev2_s	**IKE SA Parameters** *IKEV1 DPD Timeout:* 180s *DPD Delay:* 60s *Reauth Lifetime:* 24h *Rekey Lifetime:* 8h **IPsec SA Parameters** *Rekey Lifetime:* 1h	Show parameters	**Operational State:** up Local Address: 30.0.0.3 Remote Address: 30.0.0.2 See Details		true

AMD Pensando Systems

Server Time: 2023-05-17 17:25:04 UTC

Admins may also toggle to the view of total bytes by switching from the DETAILS to the METRICS tab:



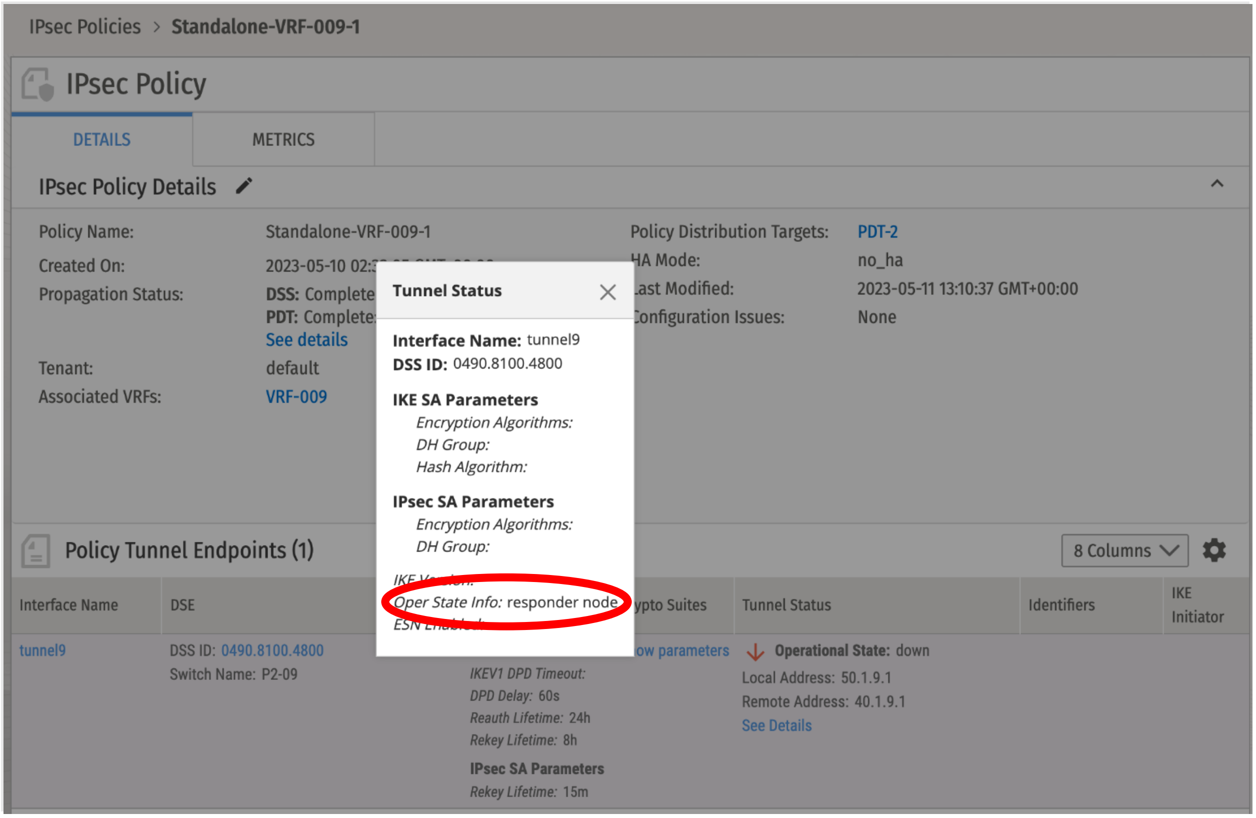
Interface Name	DSE	IKE Version	SA Timers	Crypto Suites	Tunnel Status	Identifiers	IKE Initiator
tunnel2	DSS ID: 0490.8100.1df6 Switch Name: Board32-standalo	prefer_ikev2	**IKE SA Parameters** *IKEV1 DPD Timeout:* 18(*DPD Delay:* 60s *Reauth Lifetime:* 24h *Rekey Lifetime:* 8h **IPsec SA Parameters** *Rekey Lifetime:* 1h	Show parameters	**Operational State:** up Local Address: 30.0.0.3 Remote Address: 30.0.0.2 See Details		true

AMD Pensando Systems

Server Time: 2023-05-17 18:29:29 UTC

Troubleshoot IPsec VPN Tunnels With the PSM

If an IPsec tunnel is not coming up, the associated IKE control plane error can be found in the `OperStateInfo` field of the tunnel status structure inside the IPsec policy detail page:



IPsec Policies > **Standalone-VRF-009-1**

IPsec Policy

IPsec Policy Details

Policy Name:	Standalone-VRF-009-1	Policy Distribution Targets:	PDT-2
Created On:	2023-05-10 02:	HA Mode:	no_ha
Propagation Status:	**DSS:** Complete **PDT:** Complete See details	Last Modified:	2023-05-11 13:10:37 GMT+00:00
		Configuration Issues:	None
Tenant:	default		
Associated VRFs:	VRF-009		

Tunnel Status

Interface Name: tunnel9
DSS ID: 0490.8100.4800

IKE SA Parameters
Encryption Algorithms:
DH Group:
Hash Algorithm:

IPsec SA Parameters
Encryption Algorithms:
DH Group:

IKE Version:
Oper State Info: responder node
ESN Enabled:

Policy Tunnel Endpoints (1)

Interface Name	DSE		ypto Suites	Tunnel Status	Identifiers	IKE Initiator
tunnel9	DSS ID: 0490.8100.4800 Switch Name: P2-09	*IKEV1 DPD Timeout:* *DPD Delay:* 60s *Reauth Lifetime:* 24h *Rekey Lifetime:* 8h **IPsec SA Parameters** *Rekey Lifetime:* 15m	ow parameters	↓ **Operational State:** down Local Address: 50.1.9.1 Remote Address: 40.1.9.1 See Details		

In the above example, the tunnel is down because it is configured as a responder only. So if the remote peer is not initiating the IKE exchange, the tunnel won't come up.

In the cases of tunnel down due to IPsec protocol negotiation failure with the remote peer, the "OperStateInfo" field will display the IKE control plane logs. The last few log messages usually give a good indication what the issue may be. For example:

Last Few Log Messages	Possible Cause
● *Err:vici: command failed: CHILD_SA 'IPsec-default/default/testipsec' not established* ***after 4000ms***	Remote peer not reachable
● *group:IKE level:1 ikesa-name:default/default/testipsec ikesa-uniqueid:26 msg:received* ***AUTHENTICATION_FAILED*** *notify error* ● *Err:vici: command failed: establishing CHILD_SA 'IPsec-default/default/testipsec' failed*	Mismatching preshared key or identifier
● *group:IKE level:1 ikesa-name:default/default/testipsec ikesa-uniqueid:11 msg:received* ***NO_PROPOSAL_CHOSEN*** *notify error* ● *Err:vici: command failed: establishing CHILD_SA 'IPsec-default/default/testipsec' failed*	IKE proposal mismatch
● *group:IKE level:1 ikesa-name:default/default/testipsec ikesa-uniqueid:74 msg:received* ***NO_PROPOSAL_CHOSEN*** *notify, no CHILD_SA built* ● *group:IKE level:1 ikesa-name:default/default/testipsec ikesa-uniqueid:74 msg:failed to establish CHILD_SA,* ***keeping IKE_SA*** ● *Err:vici: command failed: establishing CHILD_SA 'IPsec-default/default/testipsec' failed*	IPsec proposal mismatch

Duplicate SAs Scenarios

In certain circumstances, duplicate security associations (SAs) may be generated between peers over the lifetime of an SA, bypassing the uniqueness checks built into the software, leading to duplicate IKE SAs, and in turn duplicate Child SAs. Possible scenarios include:

- If both peers simultaneously reauthenticate their SA, or
- If child connections are initiated simultaneously.

These child duplicates will be tracked and reported in the PSM, and the latest SA will be used to encrypt/decrypt data traffic.

While the situation is always reported in the PSM `OperStateInfo` field, here is how admins can remove duplicate SAs manually:

- The duplicate tunnels can be torn down on the remote end.
- If the "active" tunnel on the CX 10000 end is torn down accidentally, its software will pick up another duplicate tunnel and use it to encrypt/decrypt the IPsec traffic.
- An operator can also choose to shut down and bring up (shut/no shut) the tunnel interface on AOS-CX, which would force termination of all existing IPsec tunnels between the peers and renegotiate a new tunnel. (**Note:** This would lead to some traffic being dropped, and the underlying condition that caused the duplicates could still act on the new tunnel.

To minimize the chance of this happening, AMD recommends using configurations that minimize the circumstances where duplicate connections may occur. This includes:

- Setting reauthentication timers of at least 24 hours for greater randomization of authentication times.
- Alternatively, disabling reauthentication on one end, or just disabling authentication in favor of IKE rekey.
- Avoid Start and Close actions on the child SAs, which may lead to duplicate SAs if not properly applied.

Caveats

The following are known caveats for the IPsec feature in this release:

- Supported only on switches in “l3-core”profile. IPsec policies are not supported on host leaf switches (“leaf” profile).
- IPsec policies can only be applied at VRF level; they cannot be applied to network level.

- IPsec policy cannot be applied on default VRF, only applicable for user-defined non-default VRFs.
- No IPsec encrypt/decrypt support for ALG traffic .
- vMotion flow migration is not applicable for flows subject to IPsec policies.
- Firewall logs will not contain the IPsec policy name.

The following IKE/crypto functions are not supported in this release:

- IPsec Anti-Replay
- IPsec Transport mode
- IPsec NAT Traversal
- Extended Sequence Numbers (ESN)
- Digital Certificate for IKE Authentication

For IP routing over the IPsec tunnel, the following are not supported in this release:

- Ping/traceroute applications are not supported over the IPsec tunnel
- Dynamic routing protocols, such as BGP, OSPF etc, over the IPsec tunnels
- ECMP load balancing over multiple IPsec tunnels

Config Change related:

- Modification of tunnel parameters like SIP, DIP and VRF are not allowed. The tunnel needs to be deleted and re-configured with new SIP, DIP and VRF parameters. Static routes need to be deleted and re-configured as well if the tunnel parameters change.

Data-path related:

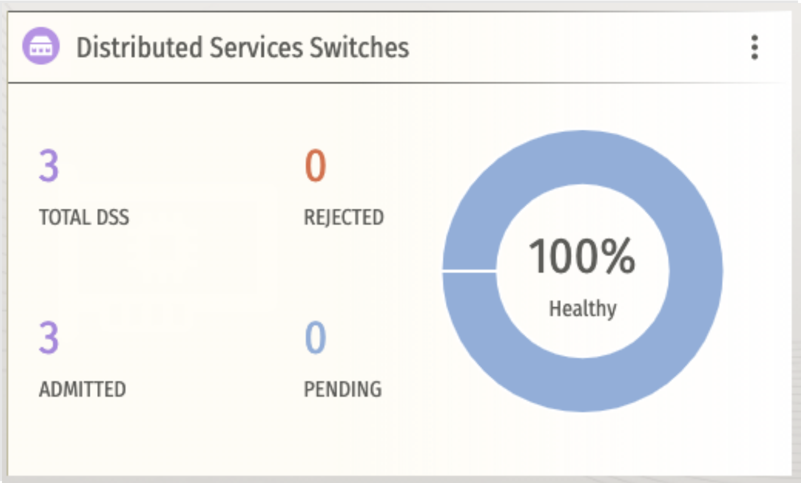
- No support for IPv6. IKE sessions and IPsec tunnels are Pv4 only.IPsec IPv4 tunnel can only encrypt IPv4 payloads, cannot encrypt IPv6 payloadsVRF route-leaking is not supported.
- In the encrypt direction, native un-encrypted traffic and post-encrypted traffic needs to be in same VRF. In the decrypt direction, the received encrypted traffic and post-decrypted native traffic needs to be in the same VRF.
- VxLAN packets cannot be encrypted over IPsec tunnel.
- IPsec cannot be used to encrypt inter-site VxLAN traffic over DCI in multi-site VxLAN topologies
- No IP fragmentation support for IPsec. After IPsec encryption, if egress MTU is less than packet-size, the packets will be dropped. The recommendation is to do path MTU discovery and adjust MTU end-to-end to avoid fragmentation post-encryption.
- IP fragments are not supported by the stateful firewall. The native packets received from inside the DC fabric or encrypted packets that the switch receives over the IPsec tunnel

cannot be IP fragments. End-to-end MTU needs to be considered and designed to make sure that the switch does not see IP fragments.

- No support for interface statistics and SNMP MIBs on tunnel interface
- IPsec packets cannot be received from a tunnel “tunnel1” (from spoke1) and re-encrypted back in another tunnel “tunnel2” (towards spoke2) with the switch acting as a hub node.
- Transit IPsec deployment is not supported with Active-Standby Failover mode.

Monitoring the DSM via the PSM UI

The PSM dashboard page provides an overview of the total number of CX 10000 switches in the fabric, the number pending admission, number rejected and a health readout.



Distributed Services Switches

3 TOTAL DSS

0 REJECTED

3 ADMITTED

0 PENDING

100% Healthy

Figure 140. Summary card for CX 10000 switches from the PSM dashboard:

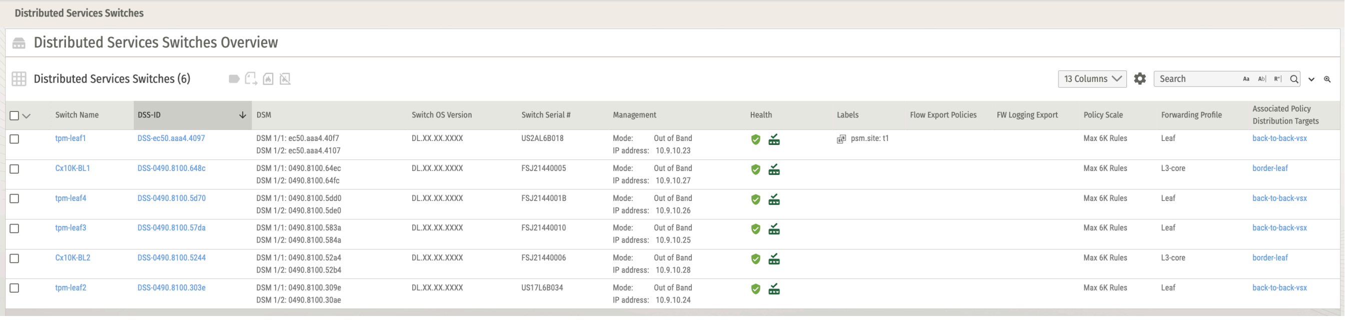
TOTAL DSS: *Number of DSSes attached to this PSM*

ADMITTED: *DSSes admitted to this PSM*

REJECTED: *DSSes that have been rejected from joining this PSM quorum*

PENDING: *DSSes pending admission*

The Distributed Services Switches Overview page, which can be reached by clicking on the menu symbol (⋮) and choosing `Navigate to DSSes`, lists top-level information, including the health of CX 10000 switches admitted to the cluster. Health status reflects the health and reachability of the DSM present in the DSS.



Distributed Services Switches

Distributed Services Switches Overview

Distributed Services Switches (6)

□	Switch Name	DSS-ID	DSM	Switch OS Version	Switch Serial #	Management	Health	Labels	Flow Export Policies	FW Logging Export	Policy Scale	Forwarding Profile	Associated Policy Distribution Targets
□	tpm-leaf1	DSS-ec50.aaa4.4097	DSM 1/1: ec50.aaa4.40f7 DSM 1/2: ec50.aaa4.4107	DL.XX.XX.XXXX	US2AL6B018	Mode: Out of Band IP address: 10.9.10.23	✓ ✓	psm.site: t1			Max 6K Rules	Leaf	back-to-back-vsx
□	Cx10K-BL1	DSS-0490.8100.648c	DSM 1/1: 0490.8100.64ec DSM 1/2: 0490.8100.64fc	DL.XX.XX.XXXX	FSJ21440005	Mode: Out of Band IP address: 10.9.10.27	✓ ✓				Max 6K Rules	L3-core	border-leaf
□	tpm-leaf4	DSS-0490.8100.5d70	DSM 1/1: 0490.8100.5dd0 DSM 1/2: 0490.8100.5de0	DL.XX.XX.XXXX	FSJ2144001B	Mode: Out of Band IP address: 10.9.10.26	✓ ✓				Max 6K Rules	Leaf	back-to-back-vsx
□	tpm-leaf3	DSS-0490.8100.57da	DSM 1/1: 0490.8100.583a DSM 1/2: 0490.8100.584a	DL.XX.XX.XXXX	FSJ21440010	Mode: Out of Band IP address: 10.9.10.25	✓ ✓				Max 6K Rules	Leaf	back-to-back-vsx
□	Cx10K-BL2	DSS-0490.8100.5244	DSM 1/1: 0490.8100.52a4 DSM 1/2: 0490.8100.52b4	DL.XX.XX.XXXX	FSJ21440006	Mode: Out of Band IP address: 10.9.10.28	✓ ✓				Max 6K Rules	L3-core	border-leaf
□	tpm-leaf2	DSS-0490.8100.303e	DSM 1/1: 0490.8100.309e DSM 1/2: 0490.8100.30ae	DL.XX.XX.XXXX	US17L6B034	Mode: Out of Band IP address: 10.9.10.24	✓ ✓				Max 6K Rules	Leaf	back-to-back-vsx

Figure 141. Distributed Services Switches Overview list

Switch Name: *Name name of the DSS. Clicking on this will open a link to that DSS's Web UI.*

DSS-ID: *ID of the DSS, used in metrics and logs; derived from switch MAC address. Clicking on this will switch to this DSS's detail page.*

DSM: *Unique MAC address IDs of the 2 DSMs in the DSS*

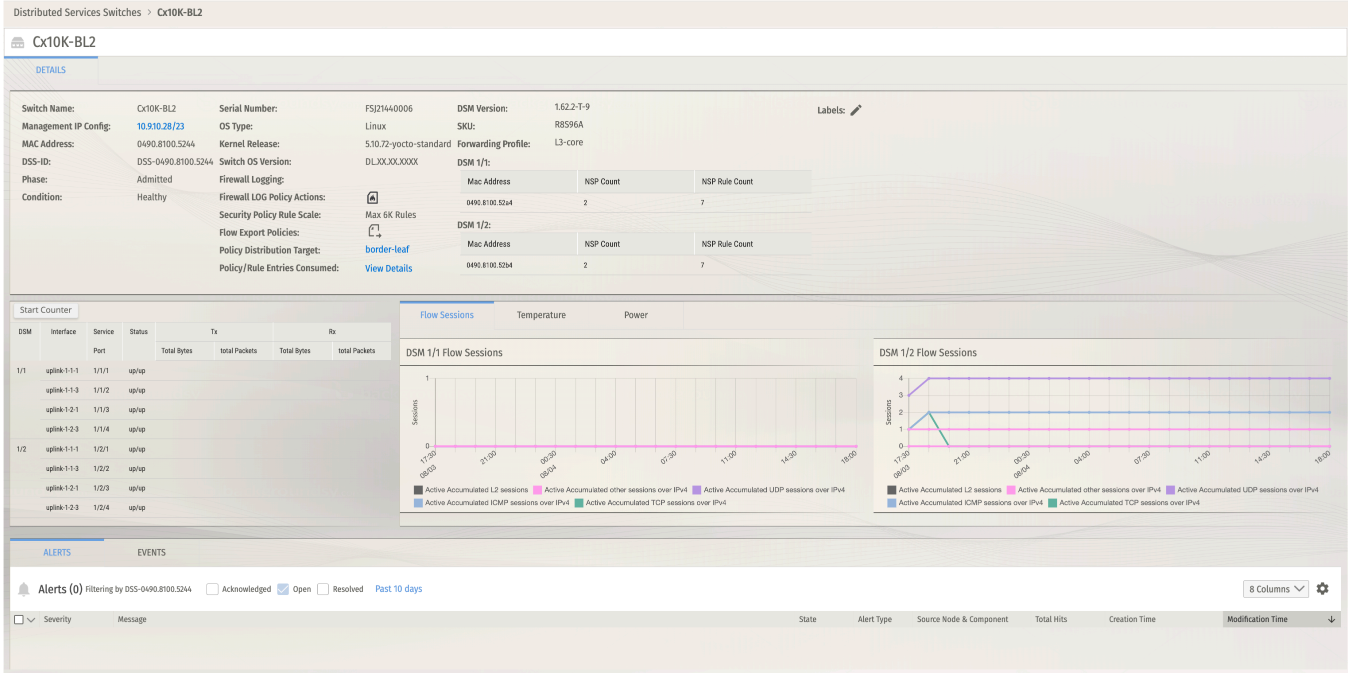
Switch OS Version: *AOS-CX version the switch is running*

Management: *Mode used to communicate between DSM and PSM*

Health: *Health and admission status of the DSS*

Labels: *Admin-definable label field*

Clicking on the DSS-ID of any admitted CX 10000 in this table will switch to its Distributed Services Switch detail page, as shown in Figure 142, which provides an overview of the two DSMs in the CX 10000, and any specific Events/Alerts related to the DSMs.



Distributed Services Switches > Cx10K-BL2

Cx10K-BL2

DETAILS

Switch Name:	Cx10K-BL2	Serial Number:	FSJ21440006	DSM Version:	1.62.2-T-9
Management IP Config:	10.9.10.28/23	OS Type:	Linux	SKU:	R8S96A
MAC Address:	0490.8100.5244	Kernel Release:	5.10.72-yocto-standard	Forwarding Profile:	L3-core
DSS-ID:	DSS-0490.8100.5244	Switch OS Version:	DL.XX.XX.XXXX		
Phase:	Admitted	Firewall Logging:			
Condition:	Healthy	Firewall LOG Policy Actions:			
		Security Policy Rule Scale:	Max 6K Rules		
		Flow Export Policies:			
		Policy Distribution Target:	border-leaf		
		Policy/Rule Entries Consumed:	View Details		

Labels:

DSM 1/1:

Mac Address	NSP Count	NSP Rule Count
0490.8100.52a4	2	7

DSM 1/2:

Mac Address	NSP Count	NSP Rule Count
0490.8100.52b4	2	7

Start Counter

DSM	Interface	Service Port	Status	Tx Total Bytes	Tx total Packets	Rx Total Bytes	Rx total Packets
1/1	uplink-1-1-1	1/1/1	up/up				
	uplink-1-1-3	1/1/2	up/up				
	uplink-1-2-1	1/1/3	up/up				
	uplink-1-2-3	1/1/4	up/up				
1/2	uplink-1-1-1	1/2/1	up/up				
	uplink-1-1-3	1/2/2	up/up				
	uplink-1-2-1	1/2/3	up/up				
	uplink-1-2-3	1/2/4	up/up				

Flow Sessions
Temperature
Power
DSM 1/1 Flow Sessions
Sessions
1
0
17:30 08/03
21:00
00:30 08/04
04:00
07:30
11:00
14:30
18:00
Active Accumulated L2 sessions
Active Accumulated other sessions over IPv4
Active Accumulated UDP sessions over IPv4
Active Accumulated ICMP sessions over IPv4
Active Accumulated TCP sessions over IPv4
DSM 1/2 Flow Sessions
Sessions
4
3
2
1
0
17:30 08/03
21:00
00:30 08/04
04:00
07:30
11:00
14:30
18:00
Active Accumulated L2 sessions
Active Accumulated other sessions over IPv4
Active Accumulated UDP sessions over IPv4
Active Accumulated ICMP sessions over IPv4
Active Accumulated TCP sessions over IPv4

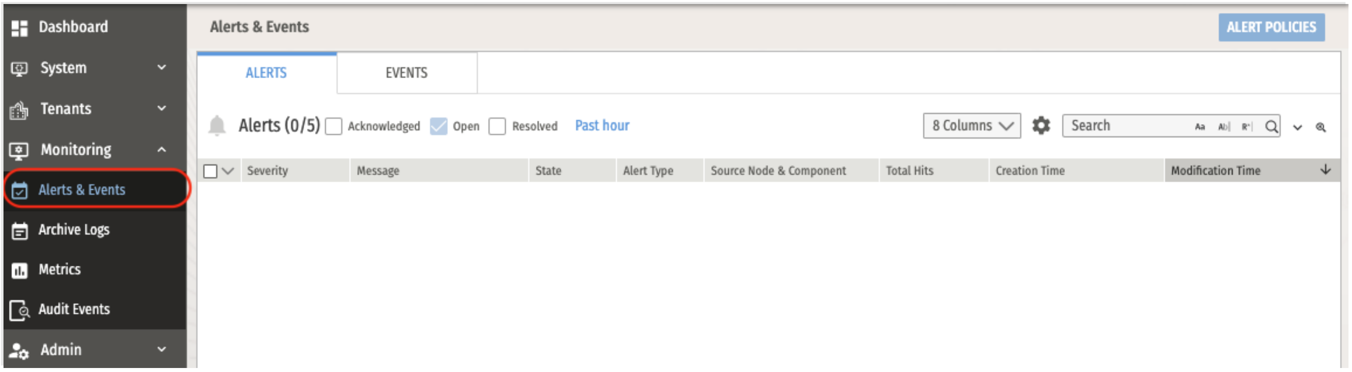
ALERTS EVENTS

Alerts (0) Filtering by DSS-0490.8100.5244 □ Acknowledged ☑ Open □ Resolved Past 10 days 8 Columns

□	Severity	Message	State	Alert Type	Source Node & Component	Total Hits	Creation Time	Modification Time

Figure 142. CX 10000 detail page

Alerts and events for all DSSes admitted to the PSM can be viewed from the Monitoring tab.



Alerts & Events

ALERT POLICIES

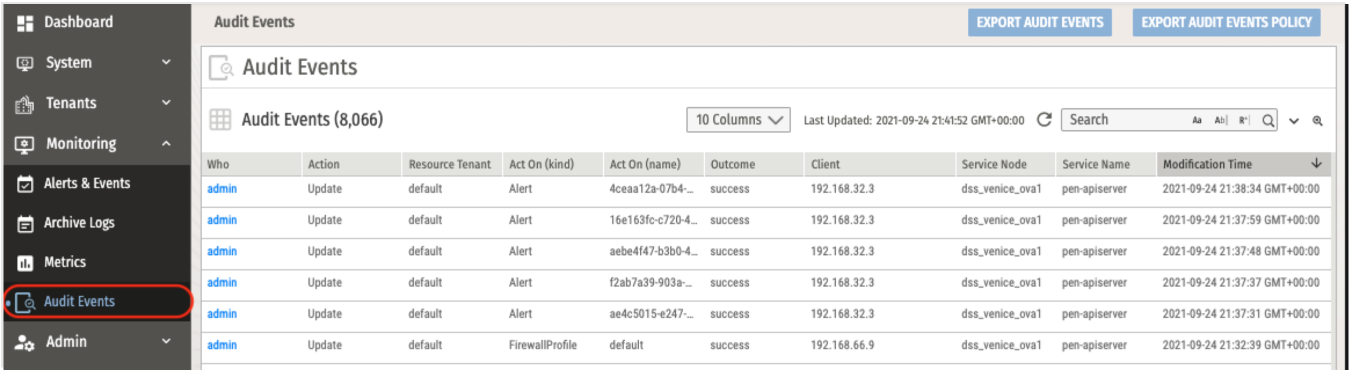
ALERTS EVENTS

Alerts (0/5) □ Acknowledged ☑ Open □ Resolved Past hour 8 Columns Search

□	Severity	Message	State	Alert Type	Source Node & Component	Total Hits	Creation Time	Modification Time

Figure 143. Monitoring -> Alerts & Events screen

There is also an option for viewing audit logs related to the DSMs and the PSM. Audit logs are a history of configuration changes related to the PSM and the DSM.



Audit Events

Audit Events (8,066)

Last Updated: 2021-09-24 21:41:52 GMT+00:00

Who	Action	Resource Tenant	Act On (kind)	Act On (name)	Outcome	Client	Service Node	Service Name	Modification Time
admin	Update	default	Alert	4ceaa12a-07b4-...	success	192.168.32.3	dss_venice_ova1	pen-apiserver	2021-09-24 21:38:34 GMT+00:00
admin	Update	default	Alert	16e163fc-c720-4...	success	192.168.32.3	dss_venice_ova1	pen-apiserver	2021-09-24 21:37:59 GMT+00:00
admin	Update	default	Alert	aebe4f47-b3b0-4...	success	192.168.32.3	dss_venice_ova1	pen-apiserver	2021-09-24 21:37:48 GMT+00:00
admin	Update	default	Alert	f2ab7a39-903a-...	success	192.168.32.3	dss_venice_ova1	pen-apiserver	2021-09-24 21:37:37 GMT+00:00
admin	Update	default	Alert	ae4c5015-e247-...	success	192.168.32.3	dss_venice_ova1	pen-apiserver	2021-09-24 21:37:31 GMT+00:00
admin	Update	default	FirewallProfile	default	success	192.168.66.9	dss_venice_ova1	pen-apiserver	2021-09-24 21:32:39 GMT+00:00

Figure 144. Monitoring -> Audit Events screen

Metrics Charts

Metrics charts can be created for various switch statistics. Select Metrics from the left menu. On the Metrics screen, select CREATE CHART:

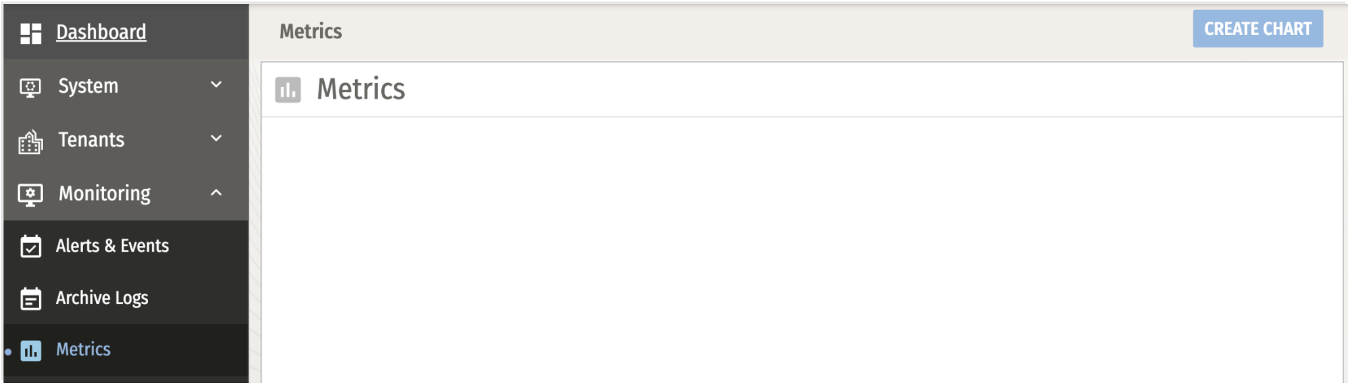
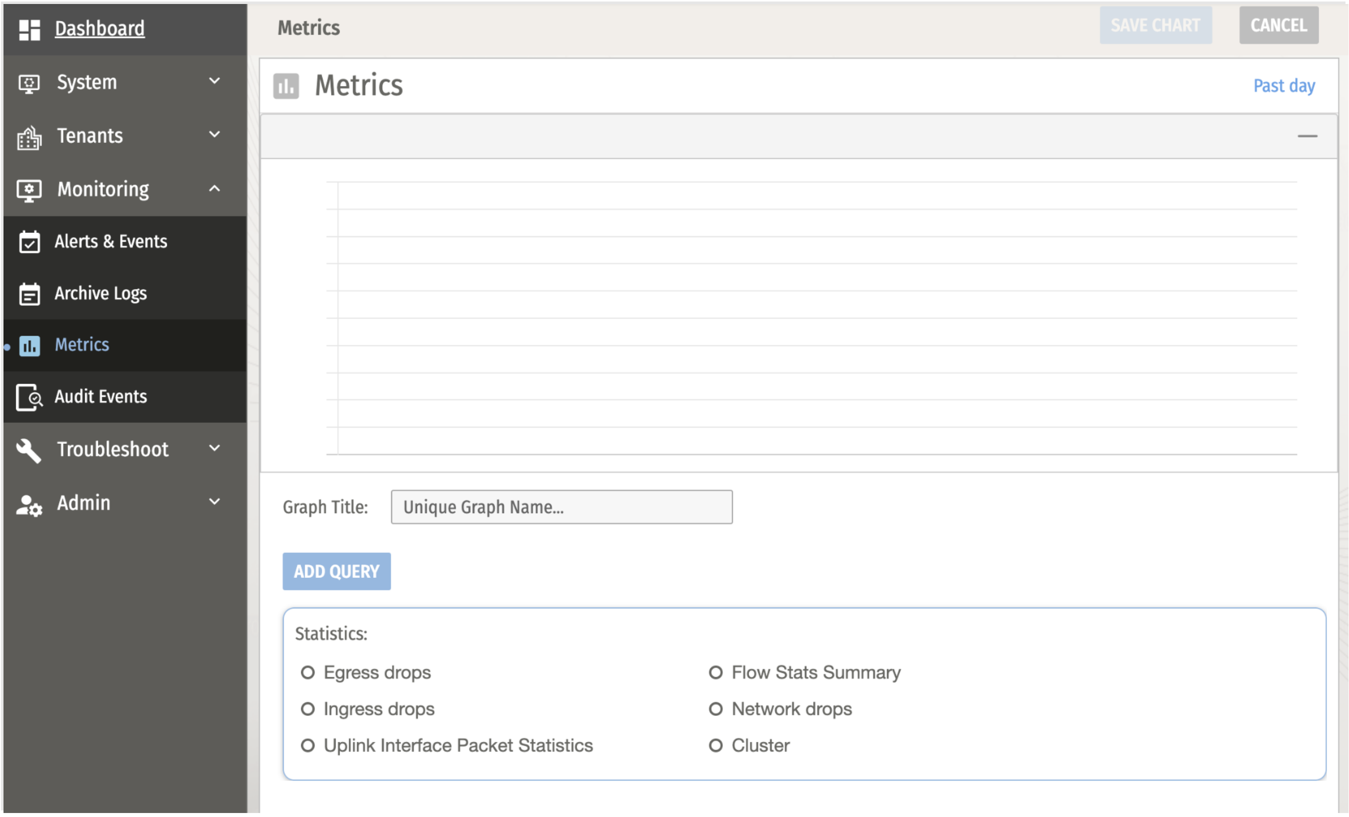


Figure 145. Metrics screen showing CREATE CHART button



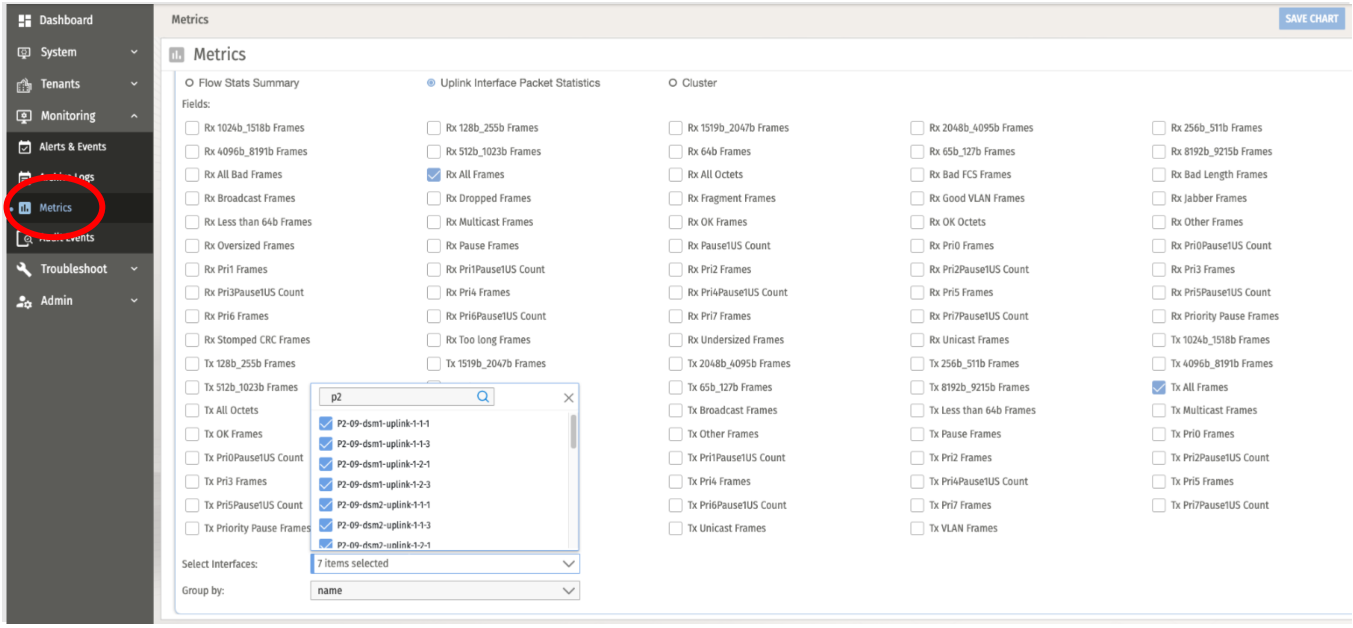
Graph Title: Unique Graph Name...

Statistics:

○ Egress drops
○ Ingress drops
○ Uplink Interface Packet Statistics
○ Flow Stats Summary
○ Network drops
○ Cluster

Figure 146. Select the appropriate statistics class for this chart

In Figure 147, the Tx/Rx frames are selected:



Metrics

○ Flow Stats Summary ● Uplink Interface Packet Statistics ○ Cluster

Fields:

□ Rx 1024b_1518b Frames □ Rx 128b_255b Frames □ Rx 1519b_2047b Frames □ Rx 2048b_4095b Frames □ Rx 256b_511b Frames
□ Rx 4096b_8191b Frames □ Rx 512b_1023b Frames □ Rx 64b Frames □ Rx 65b_127b Frames □ Rx 8192b_9215b Frames
□ Rx All Bad Frames ☑ Rx All Frames □ Rx All Octets □ Rx Bad FCS Frames □ Rx Bad Length Frames
□ Rx Broadcast Frames □ Rx Dropped Frames □ Rx Fragment Frames □ Rx Good VLAN Frames □ Rx Jabber Frames
□ Rx Less than 64b Frames □ Rx Multicast Frames □ Rx OK Frames □ Rx OK Octets □ Rx Other Frames
□ Rx Oversized Frames □ Rx Pause Frames □ Rx Pause1US Count □ Rx Pri0 Frames □ Rx Pri0Pause1US Count
□ Rx Pri1 Frames □ Rx Pri1Pause1US Count □ Rx Pri2 Frames □ Rx Pri2Pause1US Count □ Rx Pri3 Frames
□ Rx Pri3Pause1US Count □ Rx Pri4 Frames □ Rx Pri4Pause1US Count □ Rx Pri5 Frames □ Rx Pri5Pause1US Count
□ Rx Pri6 Frames □ Rx Pri6Pause1US Count □ Rx Pri7 Frames □ Rx Pri7Pause1US Count □ Rx Priority Pause Frames
□ Rx Stomped CRC Frames □ Rx Too long Frames □ Rx Undersized Frames □ Rx Unicast Frames □ Tx 1024b_1518b Frames
□ Tx 128b_255b Frames □ Tx 1519b_2047b Frames □ Tx 2048b_4095b Frames □ Tx 256b_511b Frames □ Tx 4096b_8191b Frames
□ Tx 512b_1023b Frames □ Tx 65b_127b Frames □ Tx 8192b_9215b Frames ☑ Tx All Frames
□ Tx All Octets □ Tx Broadcast Frames □ Tx Less than 64b Frames □ Tx Multicast Frames
□ Tx OK Frames □ Tx Other Frames □ Tx Pause Frames □ Tx Pri0 Frames
□ Tx Pri0Pause1US Count □ Tx Pri1Pause1US Count □ Tx Pri2 Frames □ Tx Pri2Pause1US Count
□ Tx Pri3 Frames □ Tx Pri4 Frames □ Tx Pri4Pause1US Count □ Tx Pri5 Frames
□ Tx Pri5Pause1US Count □ Tx Pri6Pause1US Count □ Tx Pri7 Frames □ Tx Pri7Pause1US Count
□ Tx Priority Pause Frames □ Tx Unicast Frames □ Tx VLAN Frames

p2

☑ P2-09-dsm1-uplink-1-1-1
☑ P2-09-dsm1-uplink-1-1-3
☑ P2-09-dsm1-uplink-1-2-1
☑ P2-09-dsm1-uplink-1-2-3
☑ P2-09-dsm2-uplink-1-1-1
☑ P2-09-dsm2-uplink-1-1-3
☑ P2-09-dsm2-uplink-1-2-1

Select Interfaces: 7 items selected

Group by: name

Figure 147. Creating a graph

Once the chart definition is saved, a graph is created:

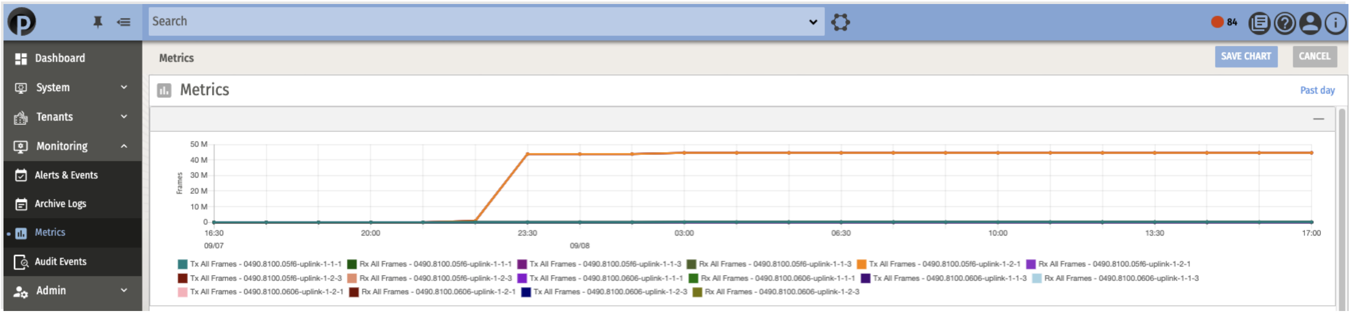

Search
84
Dashboard
System
Tenants
Monitoring
Alerts & Events
Archive Logs
Metrics
Audit Events
Admin
Metrics
SAVE CHART
CANCEL
Metrics
Past day
50 M
40 M
30 M
20 M
10 M
0
Frames
16:30
09/07
20:00
23:30
09/08
03:00
06:30
10:00
13:30
17:00
Tx All Frames - 0490.8100.05f6-uplink-1-1-1
Rx All Frames - 0490.8100.05f6-uplink-1-1-1
Tx All Frames - 0490.8100.05f6-uplink-1-1-3
Rx All Frames - 0490.8100.05f6-uplink-1-1-3
Tx All Frames - 0490.8100.05f6-uplink-1-2-1
Rx All Frames - 0490.8100.05f6-uplink-1-2-1
Tx All Frames - 0490.8100.05f6-uplink-1-2-3
Rx All Frames - 0490.8100.05f6-uplink-1-2-3
Tx All Frames - 0490.8100.0606-uplink-1-1-1
Rx All Frames - 0490.8100.0606-uplink-1-1-1
Tx All Frames - 0490.8100.0606-uplink-1-1-3
Rx All Frames - 0490.8100.0606-uplink-1-1-3
Tx All Frames - 0490.8100.0606-uplink-1-2-1
Rx All Frames - 0490.8100.0606-uplink-1-2-1
Tx All Frames - 0490.8100.0606-uplink-1-2-3
Rx All Frames - 0490.8100.0606-uplink-1-2-3

Figure 148. The ID used for selecting interfaces is the CX 10000 MAC address. The ID shown in the result is the two DSMs that are part of the CX 10000.

First DSM MAC (reporterID) ⇒ CX 10000 system MAC + 96.

Second DSM MAC (reporterID) ⇒ CX 10000 system MAC + 112.

Any chart can be “pinned”, which will place it on the PSM dashboard. To pin a chart, hover over it, and a selection of action icons will appear.

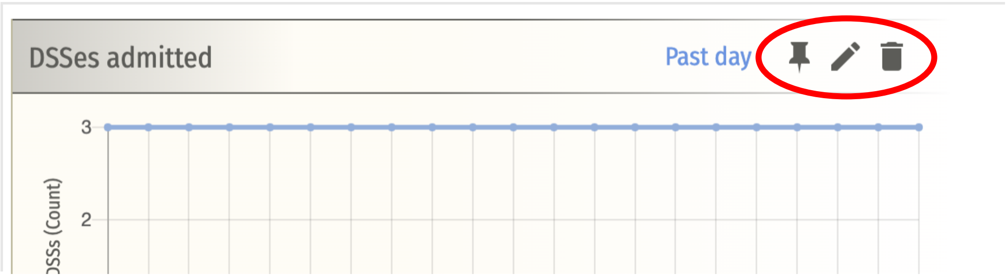
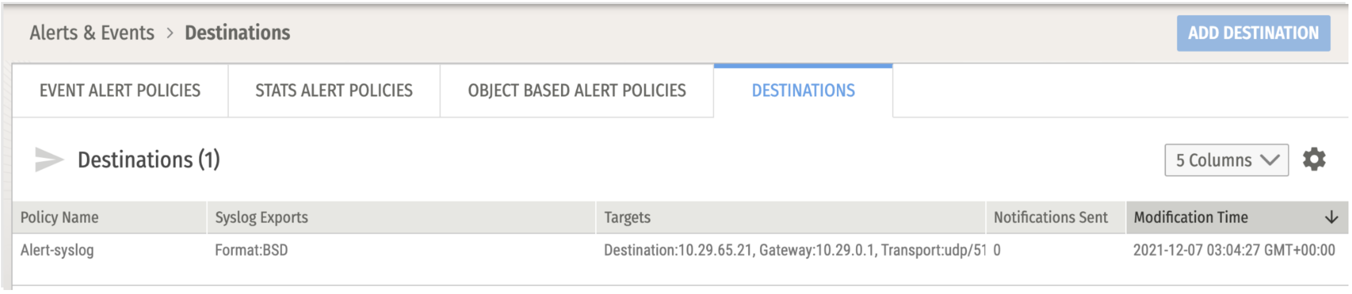


Figure 149. An example chart, showing the pin, edit, and delete controls

Select the pin (📌) icon to pin this chart. The icon will change to a white outlined pin (📌) to indicate its pinned status; you can unpin it by selecting it again.

Alerts and Events

There are two types of alerts that may be managed by policy. Stats Alerts are triggered based on min/max thresholds for given statistics/metrics. Object-Based Alerts are triggered based on attribute values of a given object. Once policies are created for Stats and Object-Based Alerts, the resulting alerts get sent to one of the syslog-based destinations that have been defined.

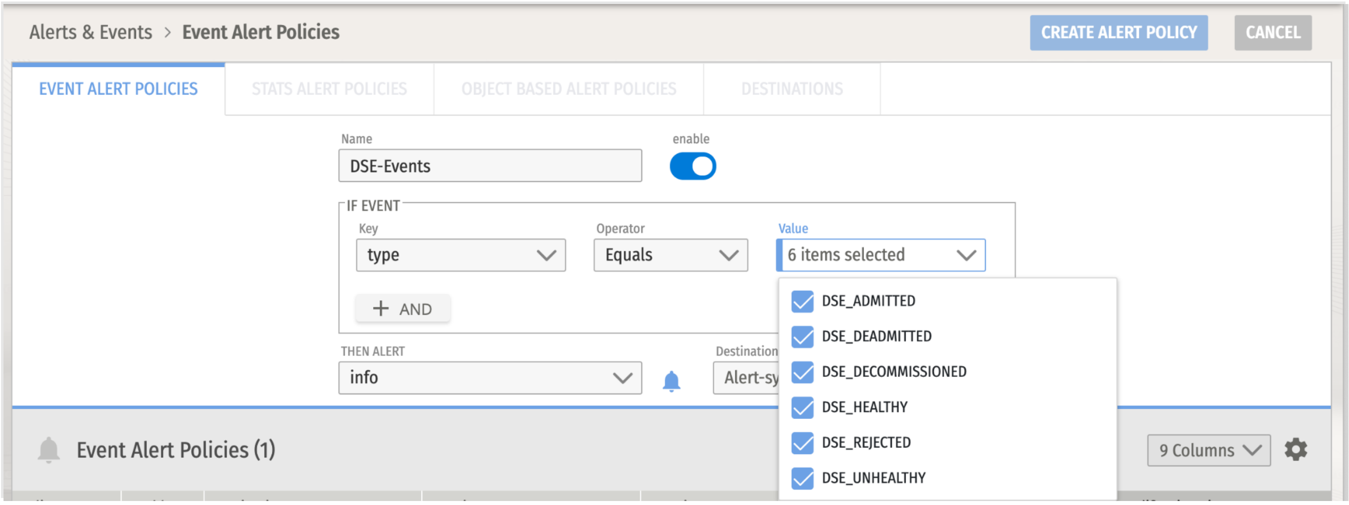


Policy Name	Syslog Exports	Targets	Notifications Sent	Modification Time
Alert-syslog	Format:BSD	Destination:10.29.65.21, Gateway:10.29.0.1, Transport:udp/51	0	2021-12-07 03:04:27 GMT+00:00

Figure 150. Viewing the list of defined alert destinations

First, select the “Destinations” tab and click “Add Destination” to configure a syslog collector that should receive the Alerts and Events.

Next, select the “Event Alert Policies” tab and click “Add Alert Policy” to provide criteria for an event to be captured. In the example below, an event is sent to the destination for any of the DSE actions. Event alert conditions can be combined into a single Alert Policy (“+AND”), or be created as individual Alert Policies.



Name: DSE-Events

enable: ●

IF EVENT

Key: type

Operator: Equals

Value: 6 items selected

- ☑ DSE_ADMITTED
- ☑ DSE_DEADMITTED
- ☑ DSE_DECOMMISSIONED
- ☑ DSE_HEALTHY
- ☑ DSE_REJECTED
- ☑ DSE_UNHEALTHY

THEN ALERT: info

Destination: Alert-sy

Figure 151. Adding an Alert Policy

PSM Automation

PSM automation and programmability can be controlled via the following methods:

- Python language bindings
- Ansible modules
- REST API

Python Language Bindings

Python language bindings are available through `pypi.org` and can be installed easily via the `pip` utility.

The pypi.org link is : https://pypi.org/project/pensando-dss/

The Python bindings can be easily installed via:

```
# pip3 install pensando-dss
```

where “dss” indicates the bindings specific to distributed services switches.

Documentation for the Python bindings is provided in the GitHub repo: https://github.com/pensando/pypi

IPlease refer to the top-level README file.

Documentation for the objects is located at : https://github.com/pensando/pypi/tree/main/src_dss/pensando_dss and includes working code examples.

All the code in the AMD Pensando PyPi repo is automatically code-generated from the PSM’s Swagger specification.

Ansible Modules

A number of Ansible modules are available at :

https://gitlab.com/pensando/tbd/netops/ansible/dss-ansible-collection

Unlike the Python bindings, the Ansible modules are not auto-generated.

Examples for the corresponding playbooks are included in the module sources.   The following example corresponds to creating a network security policy:

```
---
- hosts: localhost
  connection: local
  gather_facts: no

  vars:
    psm_ip: 203.0.113.42
    psm_username: '{{ lookup("env", "PSM_USER") }}'
    psm_passwd: '{{ lookup("env", "PSM_PASSWORD") }}'

  tasks:

    - name: Test Network Security Policy
      pnso_network_security_policy:
        state: present
        policy_name: test_policy_new
        tenant: default
        psm_user: '{{ psm_username }}'
        psm_password: '{{ psm_passwd }}'
        psm_host: '{{ psm_ip }}'
        attach_tenant: true
        rules:
          - proto_ports:
              - protocol: udp
                ports: '111-222'
            action: permit
            from_ip_addresses: [ 1.1.1.1, 2.2.2.2, 3.3.3.3 ]
            to_ip_addresses: [ 5.5.5.5, 6.6.6.6, 7.7.7.7, 8.8.8.8 ]
          - proto_ports:
              - protocol: tcp
                ports: '22'
            action: permit
            from_ip_addresses:
              - 10.10.10.10
            to_ip_addresses:
              - 20.20.20.20
          - proto_ports:
              - protocol: udp
                ports: '169'
              - protocol: tcp
                ports: '80-92,1010-1100'
              - protocol: tcp
                ports: '9000,9100,9200'
            action: deny
            from_ip_addresses:
              - 12.12.12.12
              - 13.13.13.13
              - 14.14.14.14
            to_ip_addresses:
              - 42.24.42.24
```

REST API

The PSM REST API documentation can be accessed at `https://`*`PSMaddr`*`/docs` , where *`PSMaddr`* corresponds to the PSM cluster address at your site.

A sample Postman collection is available at :
`https://`*`PSMaddr`*`/docs/examples/Sample.postman_collection.json`

The PSM Swagger specification is available at:
`https://`*`PSMaddr`*`/docs/generated/swaggeruri.html`

Tech Support Collection

The Tech Support feature collects various logs and troubleshooting information needed by technical support teams in case of an issue.

Tech support files are collected for both the PSM and the specified CX 10000 DSSes (including their DSMs).

Individual PSM nodes and DSSes can be selected. Technical Support will provide information on what components of the PSM/DSS fabric (including selecting specific DSSes) may need to be collected.

Tech Support Requests can be created in the PSM views shown below. First, navigate to Admin -> Tech Support, and click on ADD TECH-SUPPORT REQUEST.

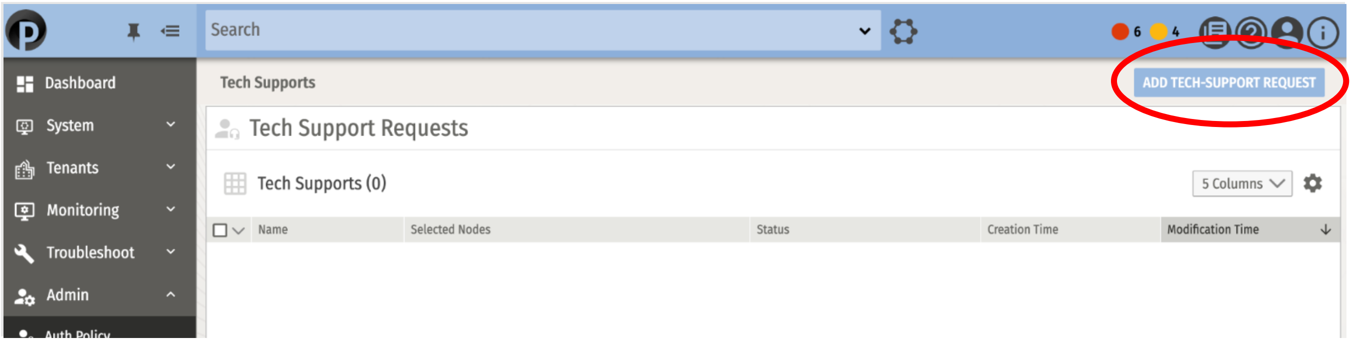
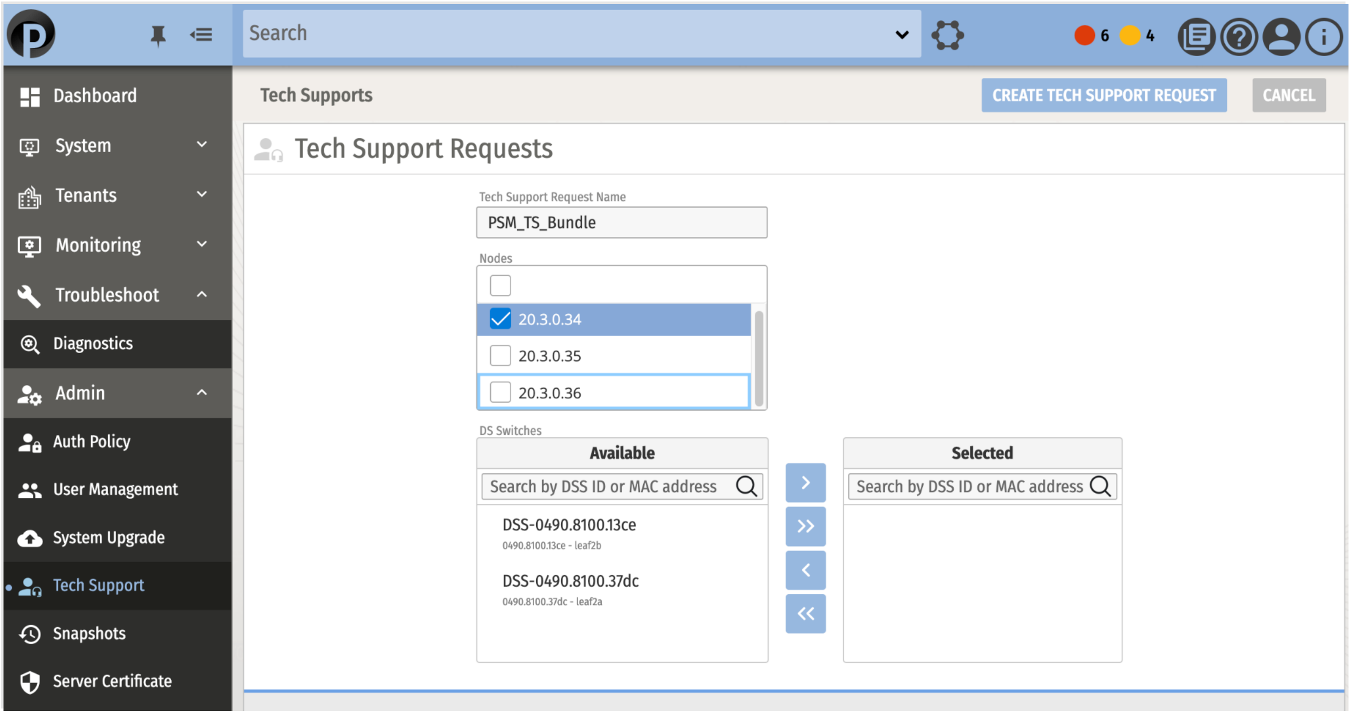


Figure 152. Tech Support screen

Specify a name for this request. A list of available PSM nodes and a list of available DSSes are shown; a search mechanism is provided to quickly find the specific switches to be included.

Once you have selected the nodes and switches to be included, click CREATE TECH SUPPORT REQUEST. It may take some time for the creation process to complete.



Tech Support Request Name: PSM_TS_Bundle

Nodes

□

☑ 20.3.0.34

□ 20.3.0.35

□ 20.3.0.36

DS Switches

Available:

DSS-0490.8100.13ce
0490.8100.13ce - leaf2b

DSS-0490.8100.37dc
0490.8100.37dc - leaf2a

Figure 153. PSM Tech Support Requests screen

With the tech-support request completed, the new request should appear on the list of tech supports. Click on "Download all files". The zip file will be downloaded to the default location of your browser.

Appendix A: PSM Quorum High Availability

The management plane is the communication channel between DSSes and the PSM cluster (gRPC over HTTPS). This traffic includes configuration, metrics, and logs. It is functionally distinguished from the data plane: the actual network traffic being managed by the DSS environment.

The data plane, including all active stateful services, is never affected by any full or partial PSM outage. In addition, since there are three PSM VMs in a cluster, the management plane remains fully operational unless two or more PSM VMs are down or unreachable.

Table 11 describes the various PSM failure conditions and how the PSM and DSSes can tolerate different failure types. It explains the various failure conditions and the expected impact to the management (configuration, logs) and data plane (traffic forwarding) from the DSS perspective.

Scenario	Management Plane Impact	Data Plane Impact
Management plane is down between a DSS and the PSM	Configuration changes are accepted by the PSM, but will not take effect until the DSS and PSM can communicate. Metrics and logs will be buffered at the DSS as local memory availability permits	No impact to stateful services Configuration changes are not possible
One PSM node fails or becomes unreachable	Configuration changes are accepted by the PSM and applied to DSSes Metrics and logs are regularly collected	No impact to stateful services No impact on configuration changes
Two PSM nodes fail or become unreachable	The PSM does not allow new user logins and does not accept configuration changes, to avoid a “split brain” scenario. Metrics and logs are not collected by the PSM, but will be buffered at the DSS subject to local storage availability.	No impact to stateful services Configuration changes are not possible

Table 11. PSM failure conditions and impacts

Appendix B: PSM Operational Network Ports

These ports must be opened in each direction in order for the PSM cluster to function correctly.

TCP Port	Service
From user station to PSM node	
22	sshd (for node management)
80	redirects to 443
443	ApiGw HTTPS
9001	Initial POST for cluster bootstrap
From PSM node to PSM node	
5001	etcd (peer)
5002	etcd (client)
6443	Kubernetes APIServer
7000	Citadel
7087	Citadel Query
9002	Cluster management
9003	ApiServer
9004	Orchestrator Hub
9009	Resolver

Table 12. Required port availability (part 1/3)

TCP Port	Service
From PSM node to PSM node (cont'd)	
9010	EventsManager
9011	Spyglass
9012	EventsProxy
9014	CMD Leader Services
9015	Rollout
9020	TPM
9030	TSM
9051	VOS
9200	Elastic (client)
9300	Elastic (peer)
10250	Kubelet
10257	Kubernetes Controller Manager
10259	Kubernetes Scheduler
10777	Citadel Collector
19001	Minio

Table 12. Required port availability (part 2/3)

TCP Port	Service
From DSS to PSM	
9005	NPM
9009	Resolver
9010	EventsManager
9012	EventsProxy
9014	CMD Health Updates
9015	Rollout
9019	NIC Registration
9020	TPM
9030	TSM
9051	VOS
10777	Citadel Collector
19001	Minio

Table 12. Required port availability (part 3/3)

Appendix C: Configuring Microsegmentation in Non-AFC Environments

Topology

Figure 154 shows an example topology for PVLAN setup.

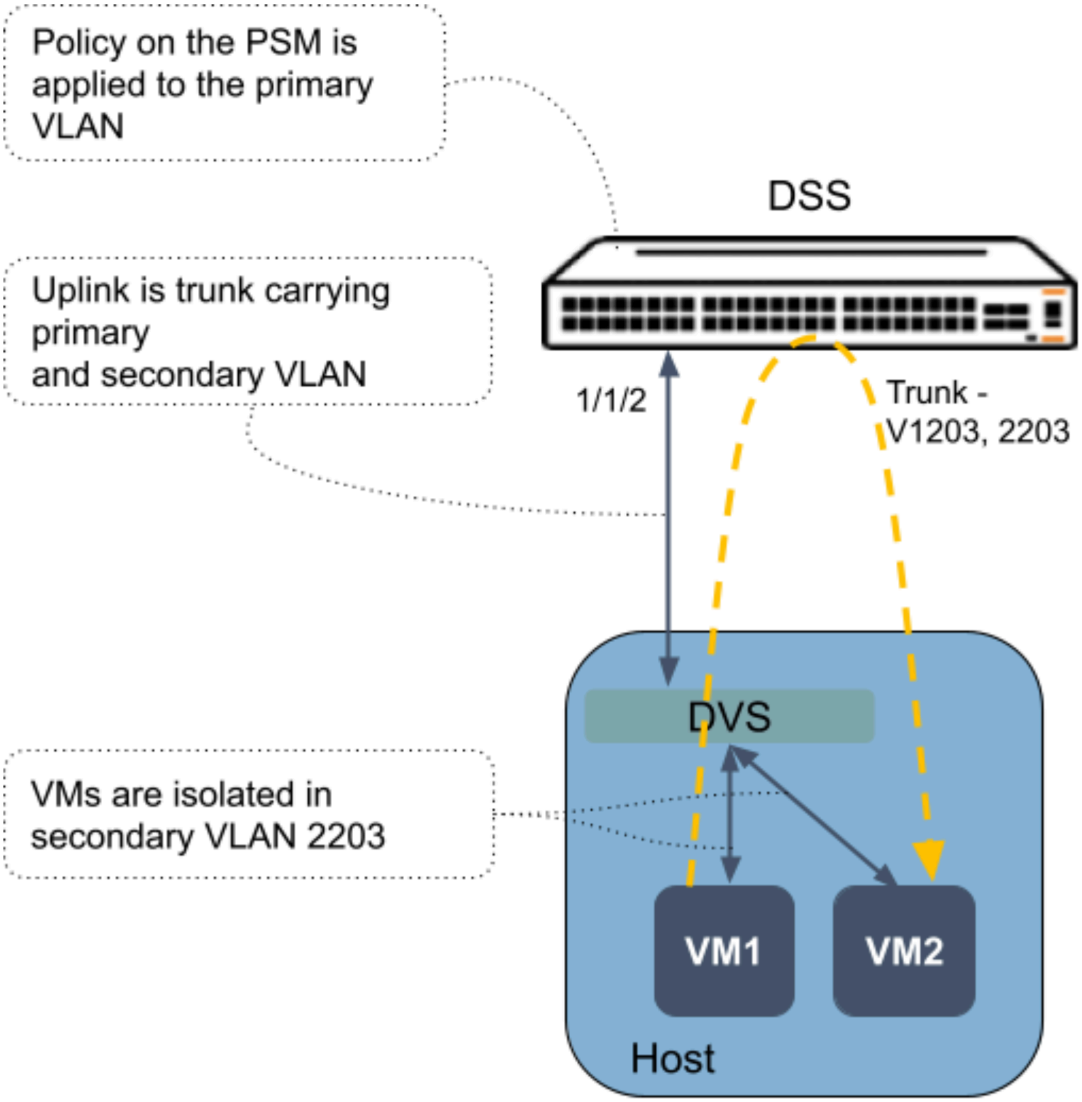

Policy on the PSM is applied to the primary VLAN
DSS
Uplink is trunk carrying primary and secondary VLAN
1/1/2
Trunk - V1203, 2203
DVS
VMs are isolated in secondary VLAN 2203
VM1
VM2
Host

Figure 154. PVLAN setup and configuration between VMs

Configuration on the DSS

In the topology in Figure 118, traffic between VM1 and VM2 would normally be switched using the DVS. By using PVLANs, policy can be enforced on the traffic between VM1 and VM2. This stateful policy enforcement is done on the DSS.

Consider 2 VMs, VM1 and VM2 which are part of the same DVS. To enable stateful policy enforcement for traffic between VM1 and VM2:

1. Configure PVLAN on the ESXi server
2. Configure PVLAN on AOS-CX
3. Configure Security Policy for the primary VLAN on the PSM

The example below shows VLAN 1203 as primary VLAN and VLAN 2203 as secondary/isolated VLAN in AOS-CX.

Global Config Mapping the Primary and Secondary VLAN

```
vlan 1203
private-vlan primary
vlan 2203
private-vlan isolated primary-vlan 1203
```

Host-Facing Interface Configured as Regular Trunk, Allowing Both Primary and Secondary VLAN

```
interface 1/1/2
no shutdown
persona access
mtu 9198
no routing
vlan trunk native 1
vlan trunk allowed 1,1203,2203
```

SVI Config on Primary with Local Proxy ARP

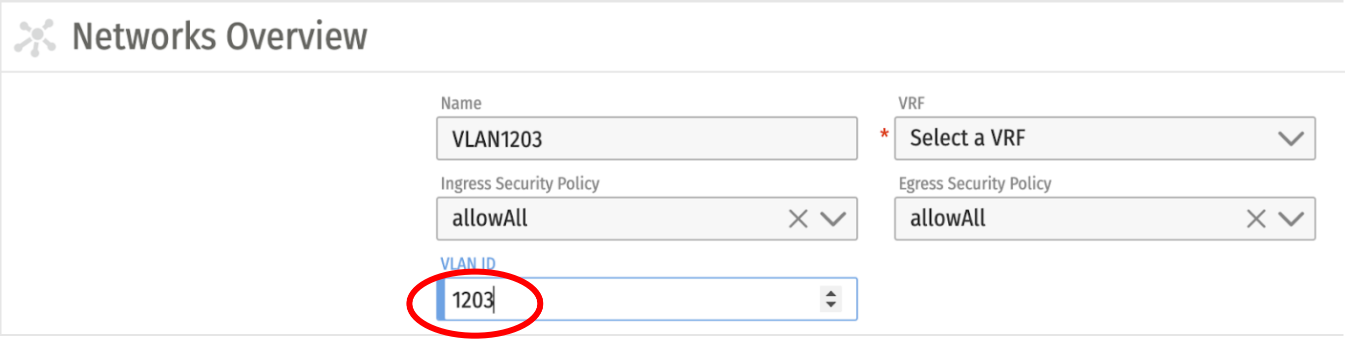
```
# show running-config interface vlan 1203
interface vlan1203
ip address 10.6.203.2/24
active-gateway ip mac 00:00:00:00:01:00
active-gateway ip 10.6.203.1
ip mtu 9198
ip local-proxy-arp
Exit
```

Configuring VMware (ESXi)

Refer to VMware documentation for the appropriate procedures for configuring a PVLAN.

Configuration on the PSM

On the PSM, create the primary private VLAN and apply the appropriate policy to the particular object (ingress and/or egress):



Name: VLAN1203

VRF: * Select a VRF

Ingress Security Policy: allowAll

Egress Security Policy: allowAll

VLAN ID: 1203

Figure 155. Setting up the VLAN

Appendix D: Saving the PSM Recovery Key

When a PSM cluster is created, a private/public key pair and corresponding self-signed certificate is generated (the cluster credentials). These credentials are used to secure communication for different PSM and DSS functions. Whenever a failed PSM node is replaced or a new PSM node is added to the cluster, these credentials are automatically passed along to the new node.

In a catastrophic situation where all of the PSM nodes are lost and a new cluster must be created, it will have new credentials, and DSSes that were admitted to the old cluster will not be able to connect to it, requiring all DSSes to be decommissioned and readmitted, even if the cluster is reloaded with a snapshot of the old cluster.

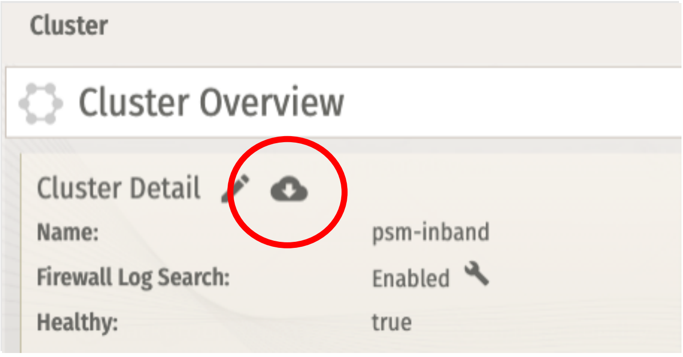
To simplify recovery from this type of loss, the PSM allows for a one-time download of its credentials (its recovery key), which can then be used with `bootstrap_PSM.py` when the replacement cluster is created.

After a first-time installation of a cluster, a privileged user may use either the PSM browser interface or a REST request to save the credentials to a file. Once this is done, the new cluster can be created using the original credentials, allowing the existing DSSes to remain admitted.

Saving the Key

From the System menu, go to the Cluster view. The “Download Cluster Recovery Key” icon will be present if the key has not already been downloaded; select this option, and save the file in a secure location.

> ***NOTE:*** *Once the key is obtained, the icon will no longer appear in the UI.*



Name: psm-inband
Firewall Log Search: Enabled
Healthy: true

Figure 156. System/Cluster screen, showing the position of the download button

Alternatively, the key may be obtained from a REST GET call to the `sysruntime/v1/cluster/recoverykeys` endpoint. After either method is used, the icon will no longer appear in the UI, and a REST call to the endpoint will fail.

> ***Note:*** *The recovery key should be stored on a system other than those hosting the PSM.*

Recovering the Cluster

To restore the cluster, three components are needed:

1. The recovery key file
2. The IP addresses previously used for the cluster
3. A snapshot of the cluster configuration

Follow the same procedure used to initially install the cluster, as described above, with one additional element: when running `bootstrap_PSM.py`, include the option

```
-recovery_keys RECOVERY_KEYS
```

where *RECOVERY_KEYS* is replaced by the path to the file containing the saved credentials.

Note that the new cluster must use the same IP addresses as the old cluster. Restore the cluster snapshot to finish the recovery.

Appendix E: Using the PSM Network Graph to Create Security Policies

In addition to manually adding predetermined policies via table view in the UI or via the equivalent API call, the PSM UI *network graph* feature can be used instead, which aids in discovering what flows should be allowed and designing a zero trust network policy to only allow those flows. This Appendix describes two approaches to using the network graph.

First Method

The Network Graph method helps discover the relationships between endpoints within a VRF, within and between VLANs, and within or between endpoint groups. The relationships help better understand common connection occurrences within the data center that should be permitted. Through an iterative process of analyzing the behavior and connections of your endpoints and applications over a time domain, the network graph method will help discover a zero trust network policy.

Navigate to `Tenants → Security Policy`; click on the drop-down menu, which by default has `Table View` selected, and change it to `Network Graph`.

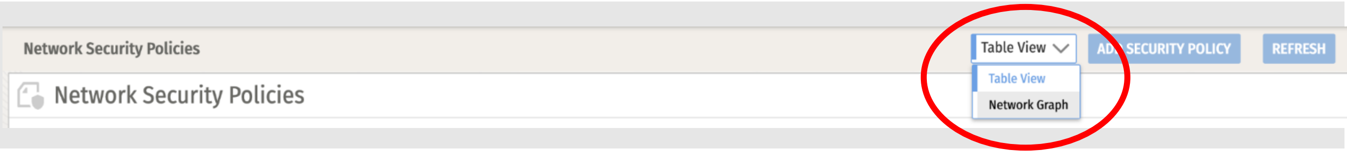
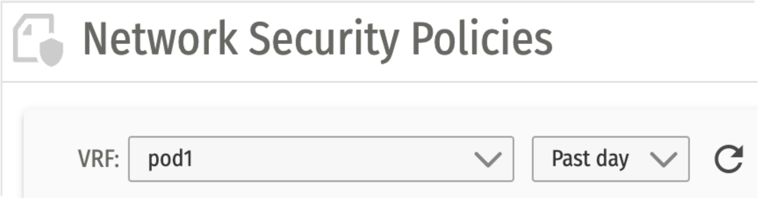


Figure 157. Network Security Policy screen, showing the policy selection method drop-down menu; change to `Network Graph`.

Use the VRF drop-down menu to select the VRF whose firewall logs you wish to evaluate. Select `Past hour` or `Past day` as the range of data to evaluate.



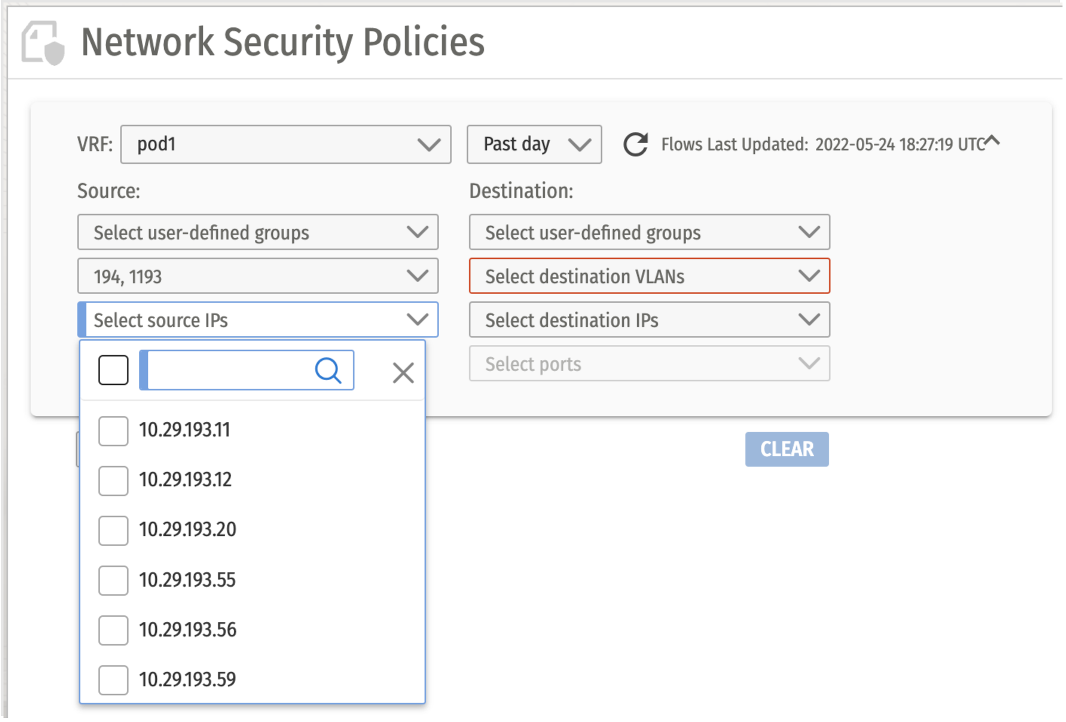
Network Security Policies

VRF: pod1

Past day

Figure 158. Choose the VRF and the range of data to evaluate.

The firewall logs can be filtered even more narrowly by selecting VLAN, IP, port and policy, as shown in Figure 123:



Network Security Policies

VRF: pod1

Past day

Flows Last Updated: 2022-05-24 18:27:19 UTC

Source:

Select user-defined groups

194, 1193

Select source IPs

□ 10.29.193.11

□ 10.29.193.12

□ 10.29.193.20

□ 10.29.193.55

□ 10.29.193.56

□ 10.29.193.59

Destination:

Select user-defined groups

Select destination VLANs

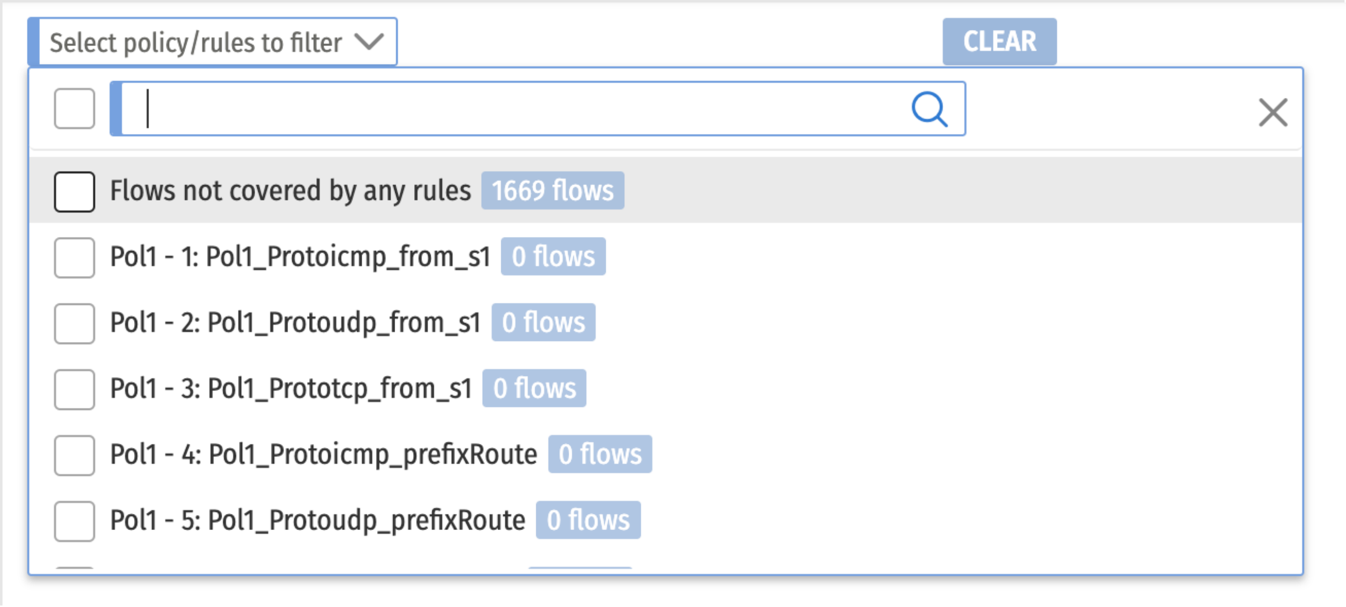
Select destination IPs

Select ports

CLEAR

Figure 159. Restrict the flows evaluated by user group, VLAN, IP, port and policy

Select `policy/rules to filter` to see that the flows for the VRF selected currently have no flows protected by any policy/rules.



□ Flows not covered by any rules 1669 flows

□ Pol1 - 1: Pol1_Protoicmp_from_s1 0 flows

□ Pol1 - 2: Pol1_Protoudp_from_s1 0 flows

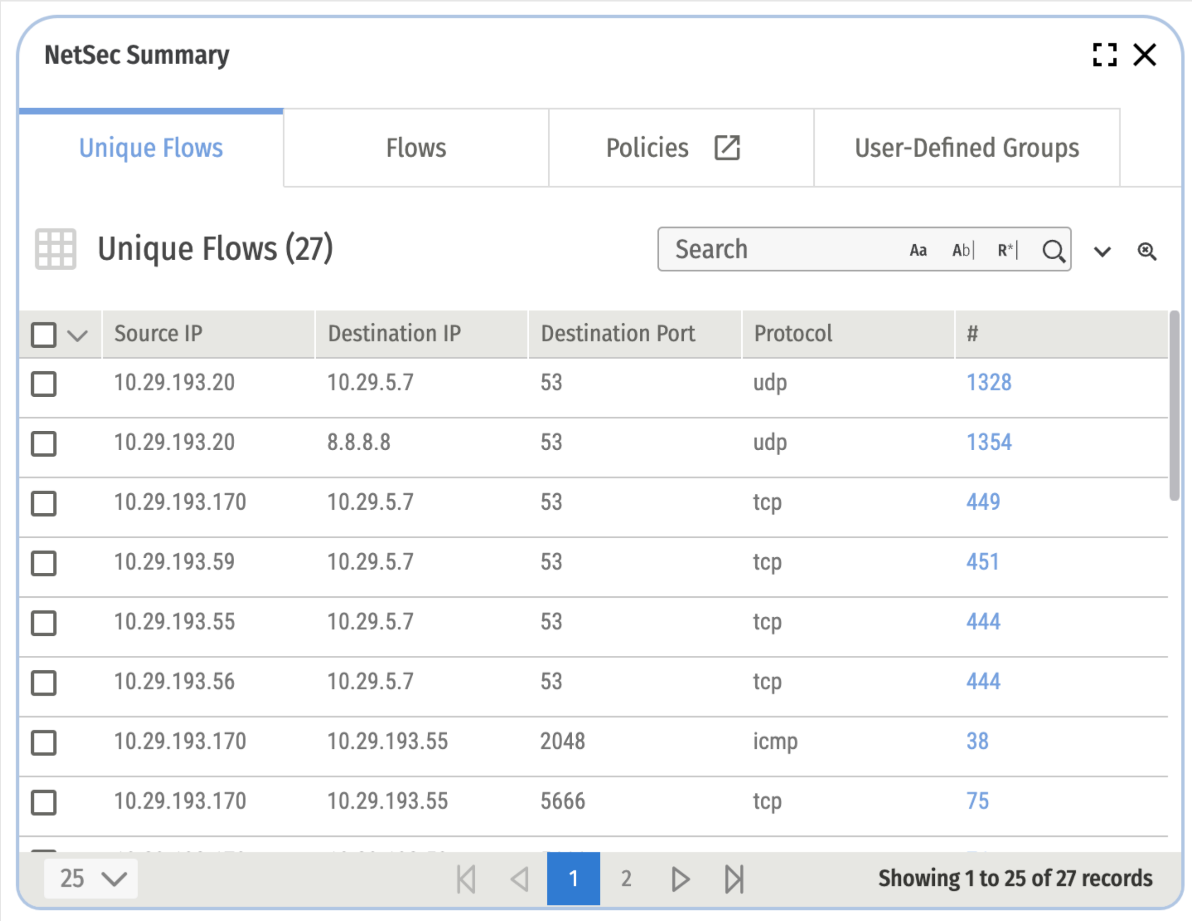
□ Pol1 - 3: Pol1_Prototcp_from_s1 0 flows

□ Pol1 - 4: Pol1_Protoicmp_prefixRoute 0 flows

□ Pol1 - 5: Pol1_Protoudp_prefixRoute 0 flows

Figure 160. Check for any flows protected by policy/rules

On the right of the screen, NetSec Summary will display flows captured for the VRF, time frame, and any other filters selected:



NetSec Summary

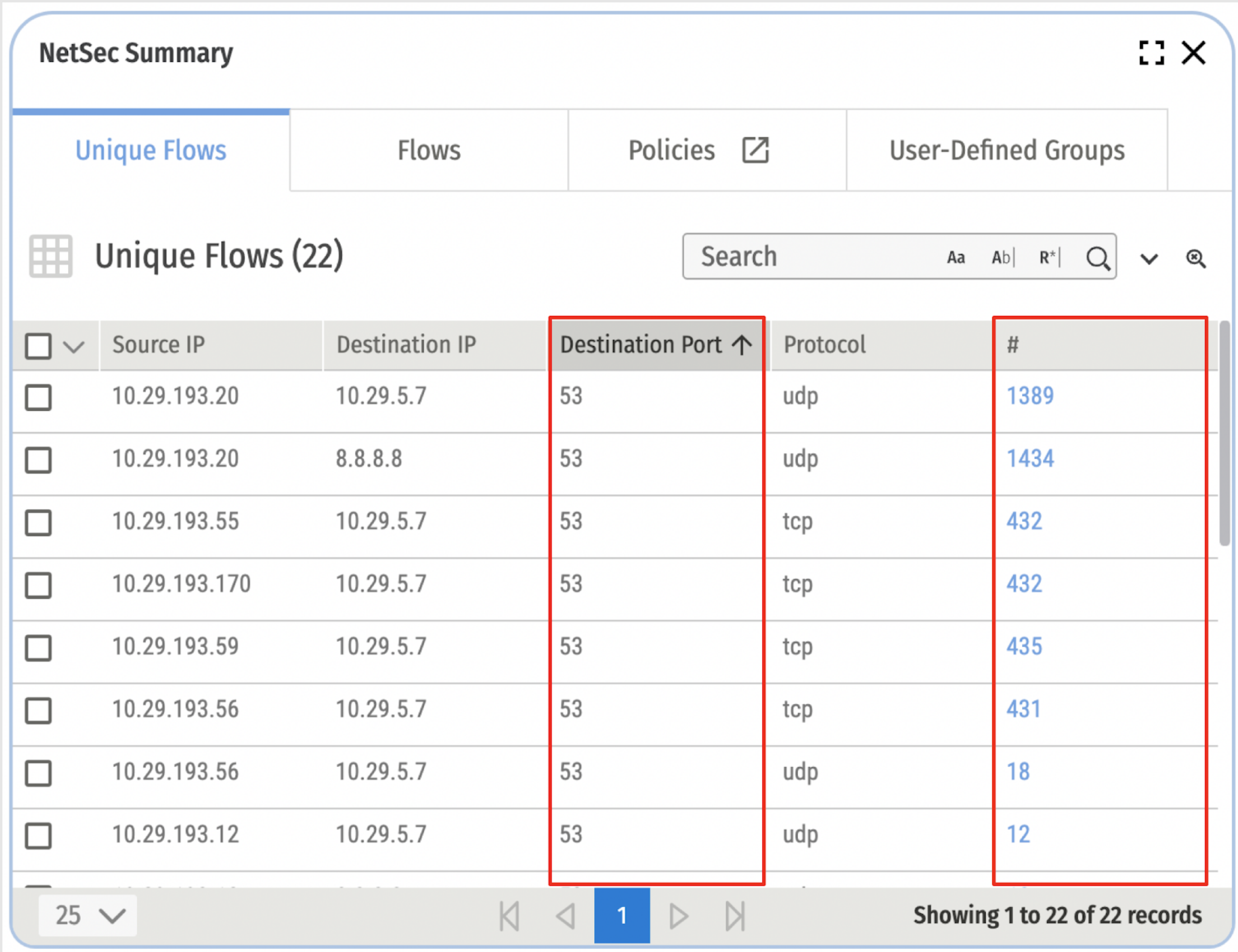
Unique Flows (27)

□	Source IP	Destination IP	Destination Port	Protocol	#
□	10.29.193.20	10.29.5.7	53	udp	1328
□	10.29.193.20	8.8.8.8	53	udp	1354
□	10.29.193.170	10.29.5.7	53	tcp	449
□	10.29.193.59	10.29.5.7	53	tcp	451
□	10.29.193.55	10.29.5.7	53	tcp	444
□	10.29.193.56	10.29.5.7	53	tcp	444
□	10.29.193.170	10.29.193.55	2048	icmp	38
□	10.29.193.170	10.29.193.55	5666	tcp	75

Showing 1 to 25 of 27 records

Figure 161. NetSec Summary shows the flows captured.

Sort by unique flows by clicking on the # column to filter the most commonly occurring flows from the top down; next, sort by clicking on the destination port column to filter by service and include the number of unique occurrences. The approach allows a focus on building policy based on a destination service and frequency of connections.



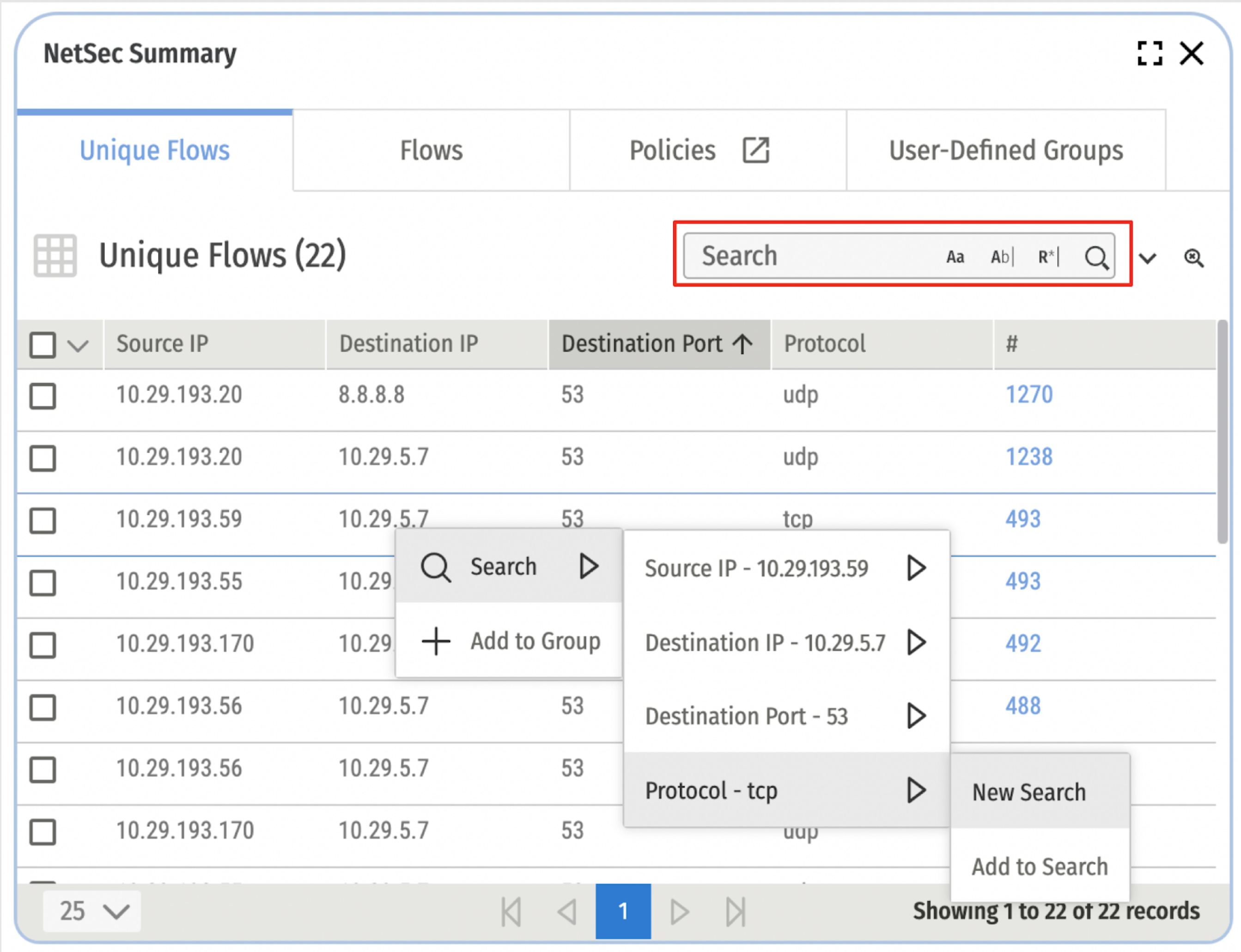
Unique Flows (22)

□	Source IP	Destination IP	Destination Port ↑	Protocol	#
□	10.29.193.20	10.29.5.7	53	udp	1389
□	10.29.193.20	8.8.8.8	53	udp	1434
□	10.29.193.55	10.29.5.7	53	tcp	432
□	10.29.193.170	10.29.5.7	53	tcp	432
□	10.29.193.59	10.29.5.7	53	tcp	435
□	10.29.193.56	10.29.5.7	53	tcp	431
□	10.29.193.56	10.29.5.7	53	udp	18
□	10.29.193.12	10.29.5.7	53	udp	12

Showing 1 to 22 of 22 records

Figure 162. The summary list can be sorted by selecting columns.

Optionally, instead of sorting by column you use the search bar to filter for a specific service or IP. The search bar can be manually entered or auto-populated by right-clicking a row of interest from the unique flow table and selecting your search criteria.


NetSec Summary

Unique Flows | Flows | Policies | User-Defined Groups

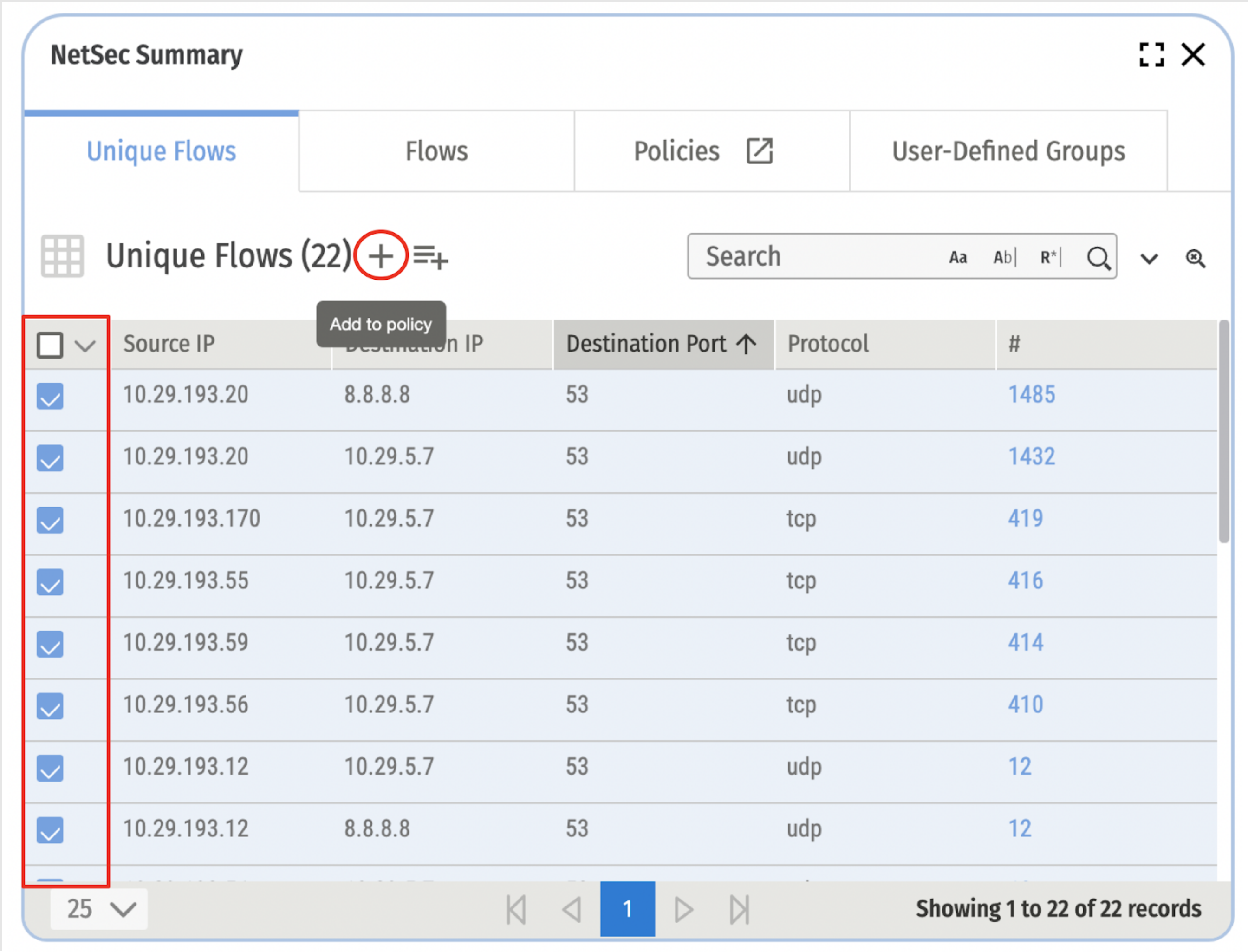
Unique Flows (22)

☐	Source IP	Destination IP	Destination Port ↑	Protocol	#
☐	10.29.193.20	8.8.8.8	53	udp	1270
☐	10.29.193.20	10.29.5.7	53	udp	1238
☐	10.29.193.59	10.29.5.7	53	tcp	493
☐	10.29.193.55	10.29			493
☐	10.29.193.170	10.29			492
☐	10.29.193.56	10.29.5.7	53		488
☐	10.29.193.56	10.29.5.7	53		
☐	10.29.193.170	10.29.5.7	53	udp	

Showing 1 to 22 of 22 records

Figure 163. Use the Search bar to filter flows

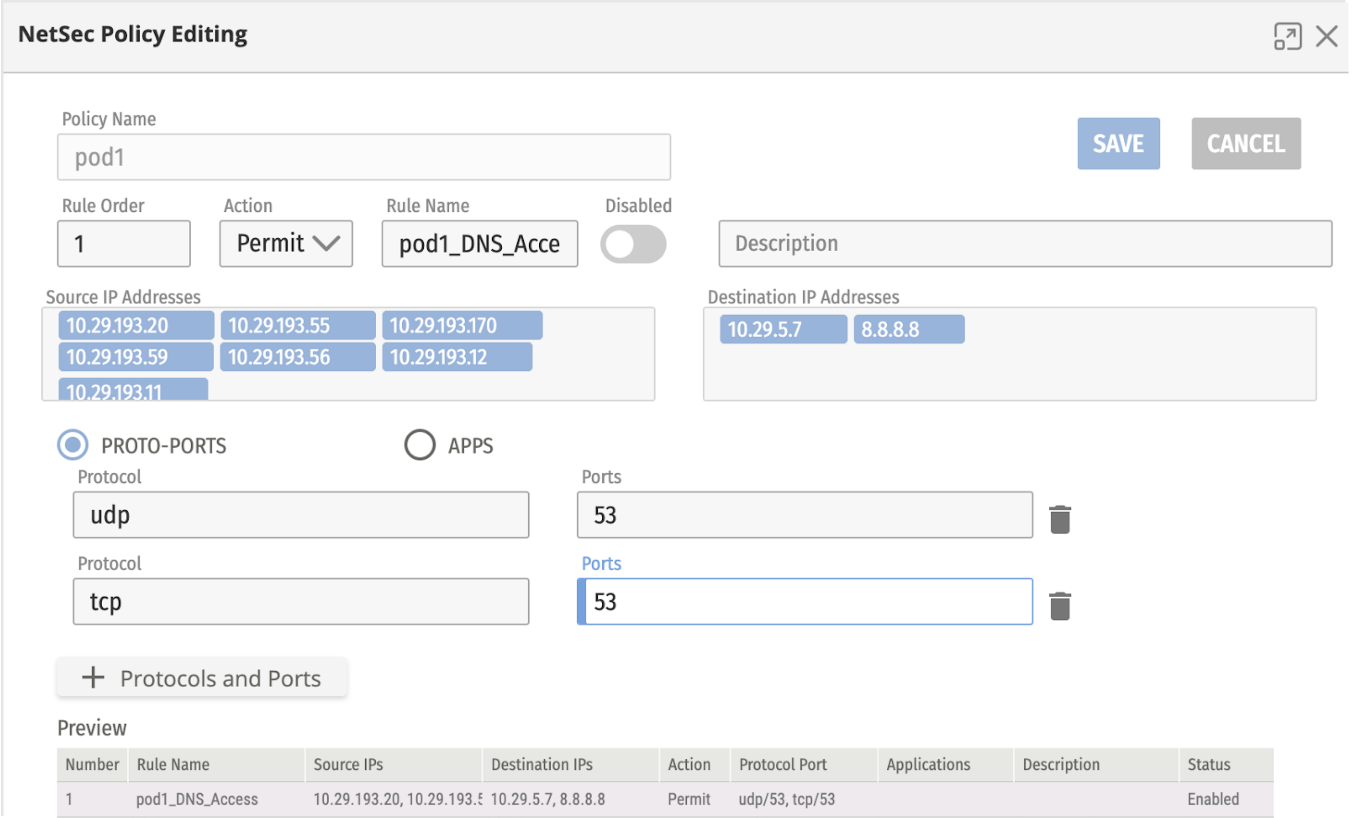
Select the filtered flows to build rules for by clicking the left column box:



☐	Source IP	Destination IP	Destination Port ↑	Protocol	#
☑	10.29.193.20	8.8.8.8	53	udp	1485
☑	10.29.193.20	10.29.5.7	53	udp	1432
☑	10.29.193.170	10.29.5.7	53	tcp	419
☑	10.29.193.55	10.29.5.7	53	tcp	416
☑	10.29.193.59	10.29.5.7	53	tcp	414
☑	10.29.193.56	10.29.5.7	53	tcp	410
☑	10.29.193.12	10.29.5.7	53	udp	12
☑	10.29.193.12	8.8.8.8	53	udp	12

Figure 164. Check the filtered flows to build rules for, and click "Add to policy".

Click "Add to policy" (the + sign, as shown in Figure 128). Based on the flows selected, the predefined source, destination, ports and protocols are pre-populated to form a rule, which can be added to an existing or new policy. Provide a policy and rule name. If a new policy, the policy is not attached to a network or VRF automatically. Remember to click save.



NetSec Policy Editing

Policy Name: pod1

Rule Order: 1
Action: Permit
Rule Name: pod1_DNS_Acce
Disabled: ○
Description

Source IP Addresses: 10.29.193.20, 10.29.193.55, 10.29.193.170, 10.29.193.59, 10.29.193.56, 10.29.193.12, 10.29.193.11

Destination IP Addresses: 10.29.5.7, 8.8.8.8

● PROTO-PORTS ○ APPS

Protocol: udp Ports: 53

Protocol: tcp Ports: 53

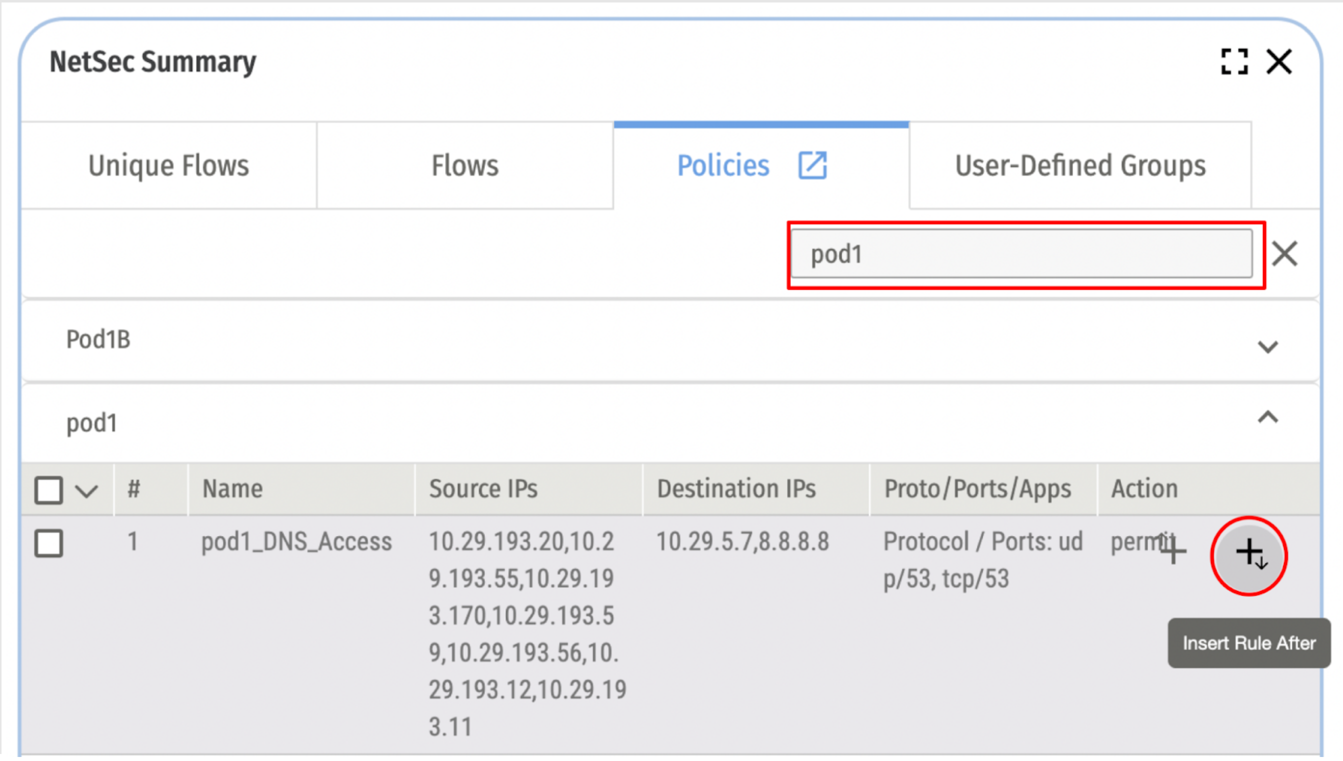
Preview

Number	Rule Name	Source IPs	Destination IPs	Action	Protocol Port	Applications	Description	Status
1	pod1_DNS_Access	10.29.193.20, 10.29.193.5	10.29.5.7, 8.8.8.8	Permit	udp/53, tcp/53			Enabled

Figure 165. Rule table populated with objects from the flow table selected

To use this method for policy discovery it is essential to add a “permit any any” rule, to collect which flows would have been denied. Using a “permit any any” rule at the end of the policy assists with policy evaluation while iterating through the discovery of new flows and evaluating new time dimensions of the firewall logs within the VRF and other selected filters.

Select the Policies tab, search for your policy name, and click `Insert Rule After`, as shown in Figure 130. Enter a rule as shown in Figure 131, to catch all traffic that did not match rule 1.

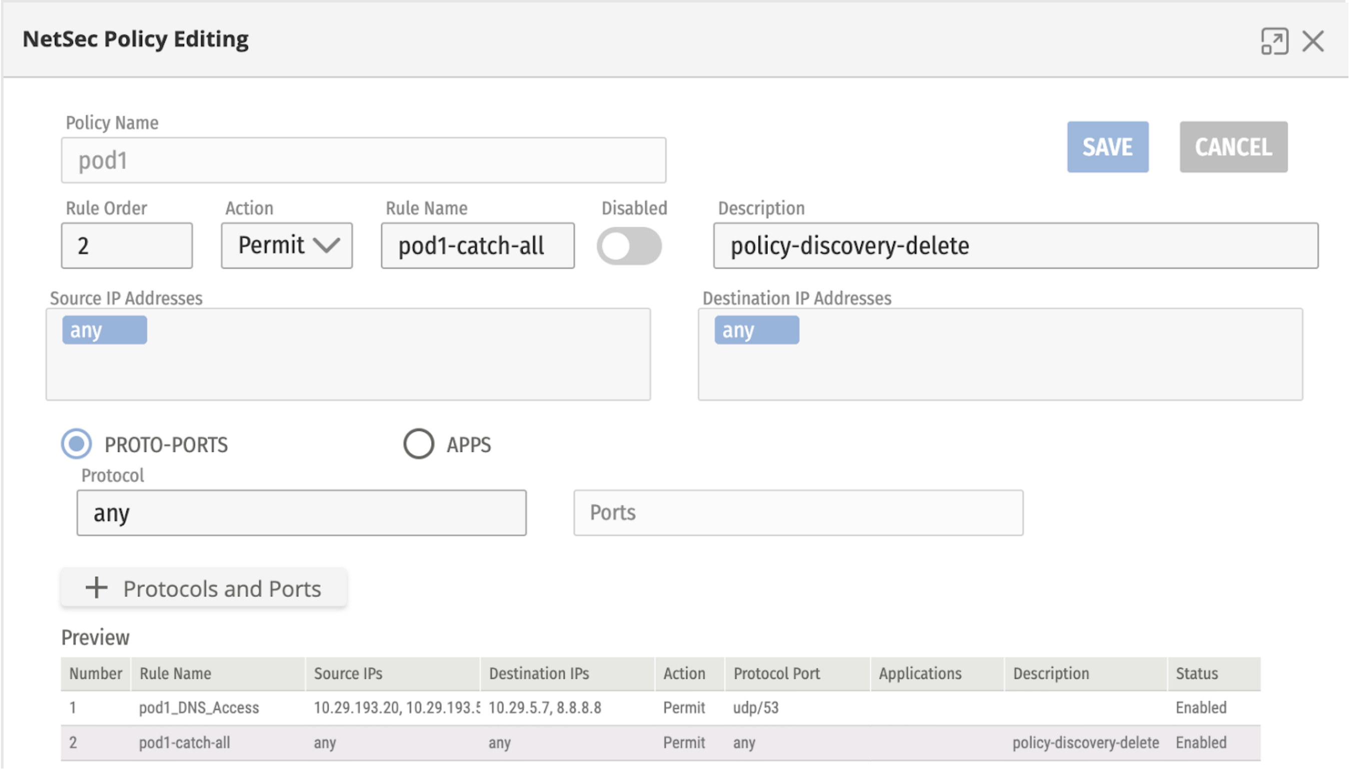


NetSec Summary

pod1

□	#	Name	Source IPs	Destination IPs	Proto/Ports/Apps	Action
□	1	pod1_DNS_Access	10.29.193.20,10.29.193.55,10.29.193.170,10.29.193.59,10.29.193.56,10.29.193.12,10.29.193.11	10.29.5.7,8.8.8.8	Protocol / Ports: udp/53, tcp/53	permit

Figure 166. Inserting a rule as shown to catch all traffic that didn't match any of the specific permit rules.



NetSec Policy Editing

Policy Name: pod1

Rule Order	Action	Rule Name	Disabled	Description
2	Permit	pod1-catch-all	○	policy-discovery-delete

Source IP Addresses: any

Destination IP Addresses: any

● PROTO-PORTS ○ APPS

Protocol: any

Ports

Preview

Number	Rule Name	Source IPs	Destination IPs	Action	Protocol Port	Applications	Description	Status
1	pod1_DNS_Access	10.29.193.20, 10.29.193.5	10.29.5.7, 8.8.8.8	Permit	udp/53			Enabled
2	pod1-catch-all	any	any	Permit	any		policy-discovery-delete	Enabled

Figure 167. Creation of the catch-all rule for policy discovery. Remember to click Save.

Remember to click SUBMIT CHANGES in the top right of the screen.

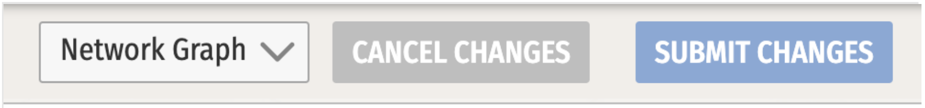
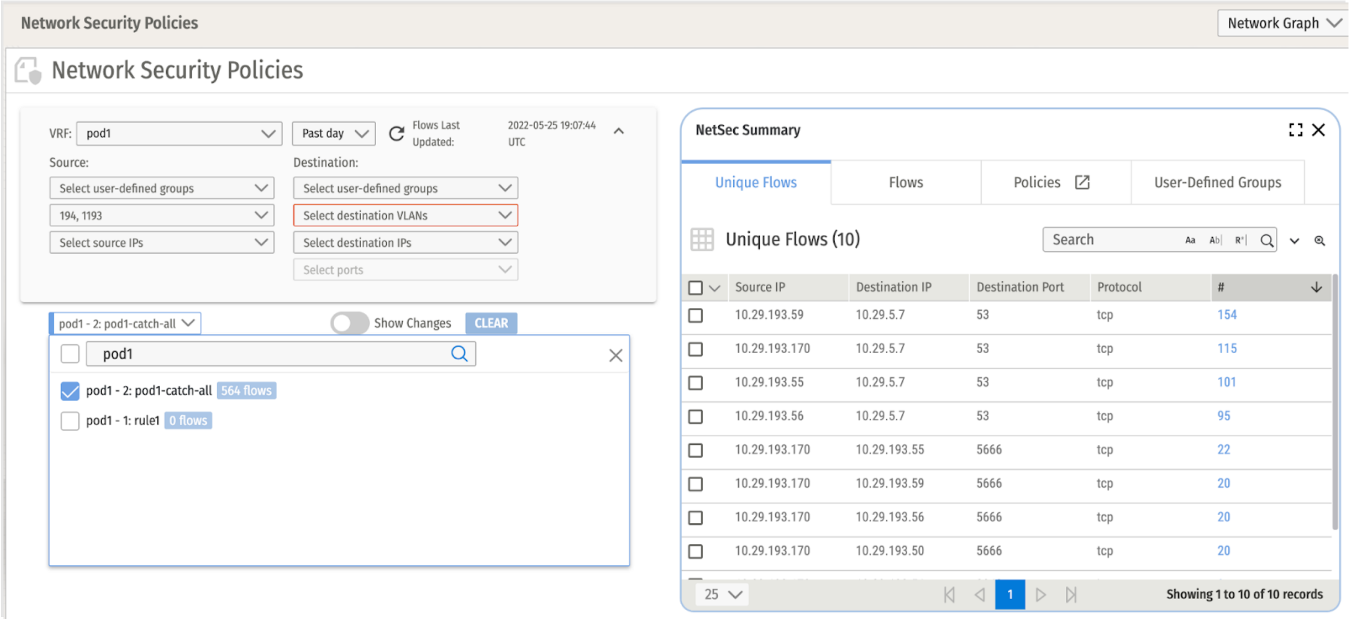


Figure 168. Click SUBMIT CHANGES

At the next time domain of collected firewall logs for the VRF policy discovery, follow the steps below to filter flows to analyze. Click on `Select policy/rules` to filter the policy/rules to analyze what flows match the specified permit rules and what flows match the catch-all “permit any” rule. Any such flows that did not match a specific permit rule and matched the catch-all rule require further analysis to determine if these flows should be permitted with a specific rule. Through several iterations this process will help create a network zero trust policy for a micro, macro, or zone based segmentation strategy.

The example in Figure 133 shows that 24 hours later, 564 new flows are net new and matched the catch-all rule. These flows require evaluation to determine whether they should be considered permitted flows. The NetSec Summary, on the right of Figure 133, displays the unique flows that would have been implicitly denied if not for the catch-all permit. Iterate through this process until you are satisfied that a zero trust implicit deny at the end of the policy will not block any needed flows.



VRF: pod1

Past day

Flows Last Updated: 2022-05-25 19:07:44 UTC

194, 1193

pod1 - 2: pod1-catch-all

pod1

☑ pod1 - 2: pod1-catch-all 564 flows

□ pod1 - 1: rule1 0 flows

Unique Flows (10)

Source IP	Destination IP	Destination Port	Protocol	#
10.29.193.59	10.29.5.7	53	tcp	154
10.29.193.170	10.29.5.7	53	tcp	115
10.29.193.55	10.29.5.7	53	tcp	101
10.29.193.56	10.29.5.7	53	tcp	95
10.29.193.170	10.29.193.55	5666	tcp	22
10.29.193.170	10.29.193.59	5666	tcp	20
10.29.193.170	10.29.193.56	5666	tcp	20
10.29.193.170	10.29.193.50	5666	tcp	20

25

Showing 1 to 10 of 10 records

Figure 169. Analysis of catch-all rule and related flows and occurrences

Second Method

This Network Graph method helps discover the relationships between endpoints within a VRF, within and between VLANs, and within or between user-defined groups. End users who prefer to evaluate connections and relationships rendered as a graph and build policy will choose this approach. For example, you can manually create user-defined groups to analyze their application's flows (Eg: Web->App, App->DB, etc.)

Navigate to `Tenants → Security Policy`; click on the drop-down menu, which by default has `Table View` selected, and change it to `Network Graph`.

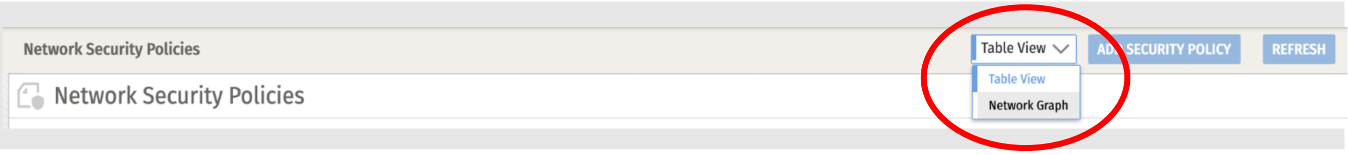
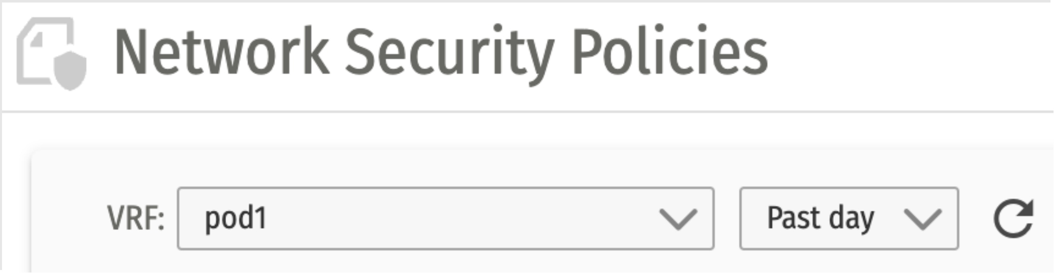


Figure 170. Network Security Policy screen, showing the policy selection method drop-down menu; change to Network Graph.

Use the VRF drop-down menu to select the VRF whose firewall logs you wish to evaluate. Select `Past hour` or `Past day` as the range of data to evaluate.

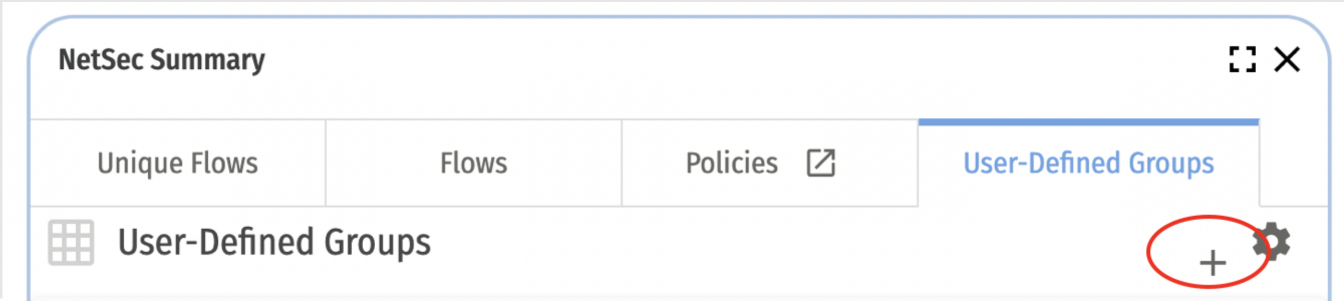


Network Security Policies

VRF: pod1 Past day

Figure 171. Choose the VRF and the range of data to evaluate.

User-defined groups (UDGs) can be manually created by clicking on the `User-Defined Groups` tab and manually defining each group based on known IPs and/or VLANs. Click the "Add to group" button (the "+" sign highlighted in Figure 172); the Create Group popup will appear as seen in Figure 173.



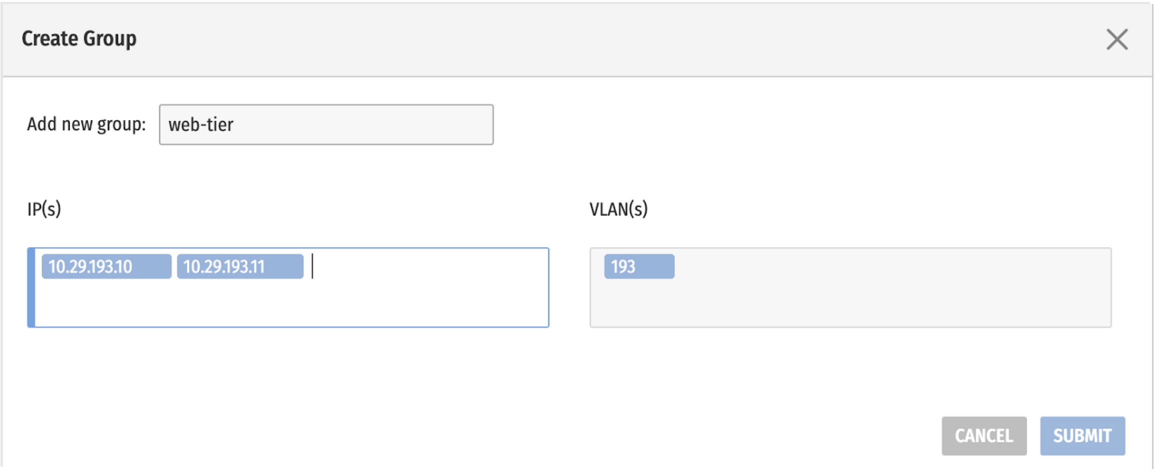
NetSec Summary

Unique Flows | Flows | Policies | User-Defined Groups

User-Defined Groups

Figure 172. The User-Defined Groups tab, as shown on the right side of the Network Security Policies screen

Example of a UDG based on IP addresses and/or VLAN:



Create Group

Add new group: web-tier

IP(s): 10.29.193.10, 10.29.193.11

VLAN(s): 193

Figure 173. Create Group popup

Optionally, known groups can be uploaded from a JSON file. Select the gear icon, and choose "Upload Groups":

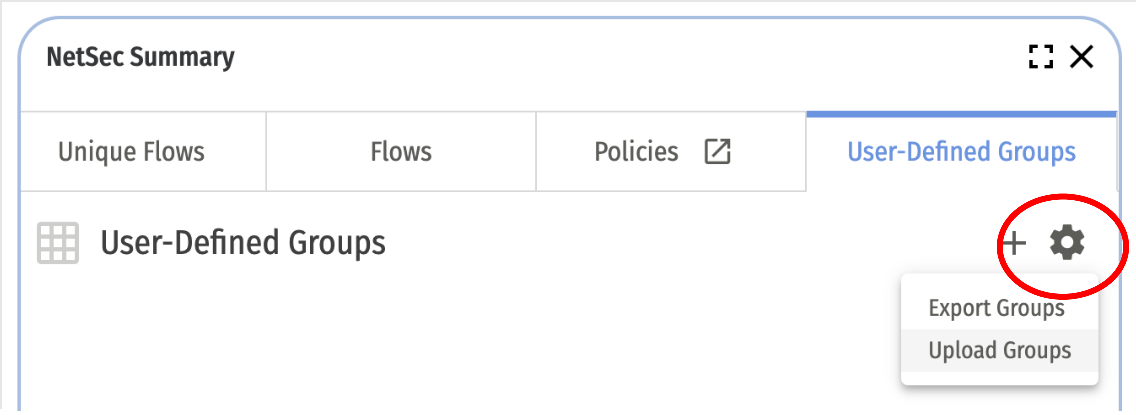
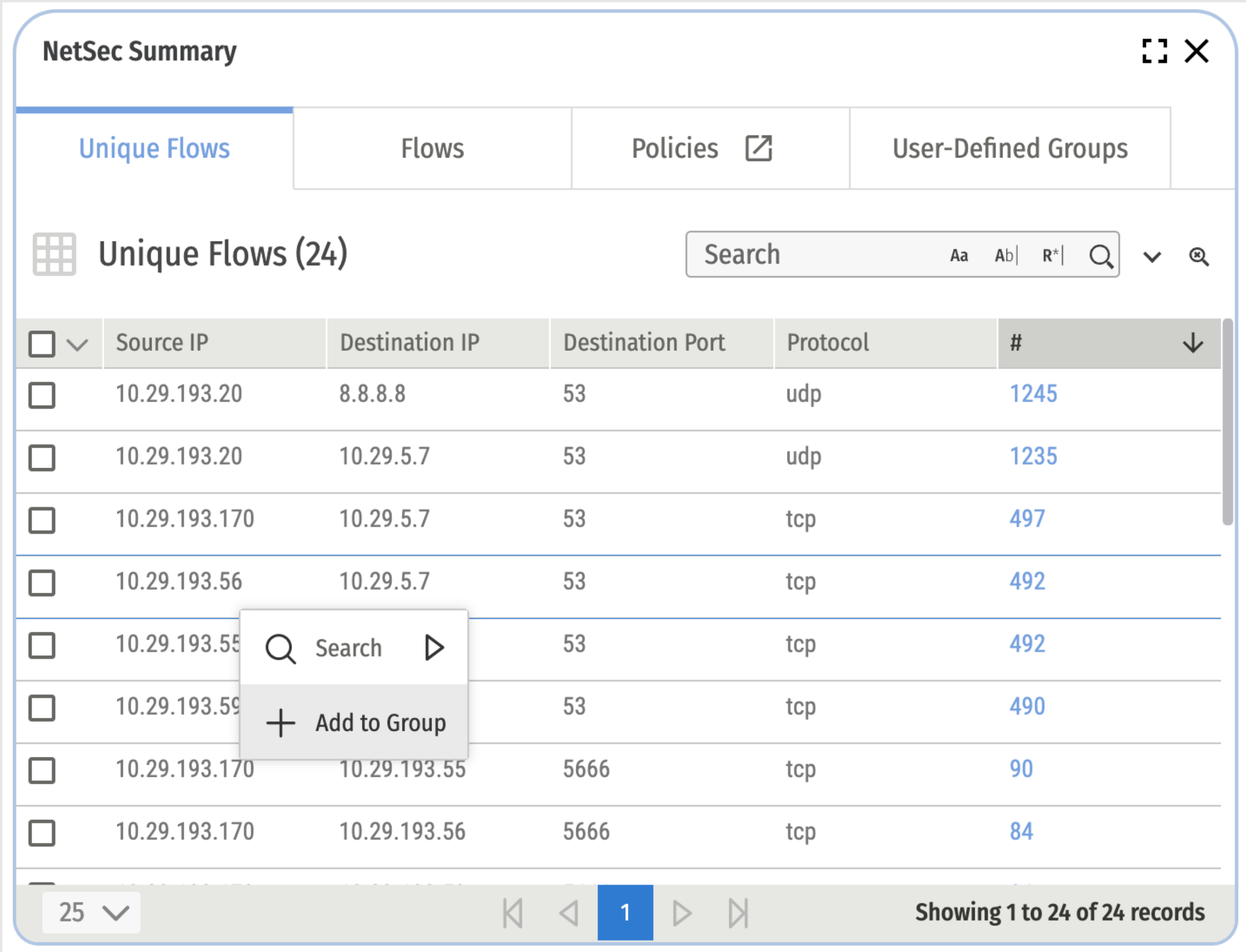


Figure 174. The gear icon accesses the Export/Upload Groups menu.

Exam`ple file:

```
{
   "App-Tier": [
     {
       "type"; “VLAN",
       "value": "193"
     },
     {
       "type": "IP",
       "value": "10.29.193.56"
     }
   ]
}
```

A final option for creating a UDG is to select the `Flows` tab, right-click on a row, and select `Add to Group` (which can be new or existing):

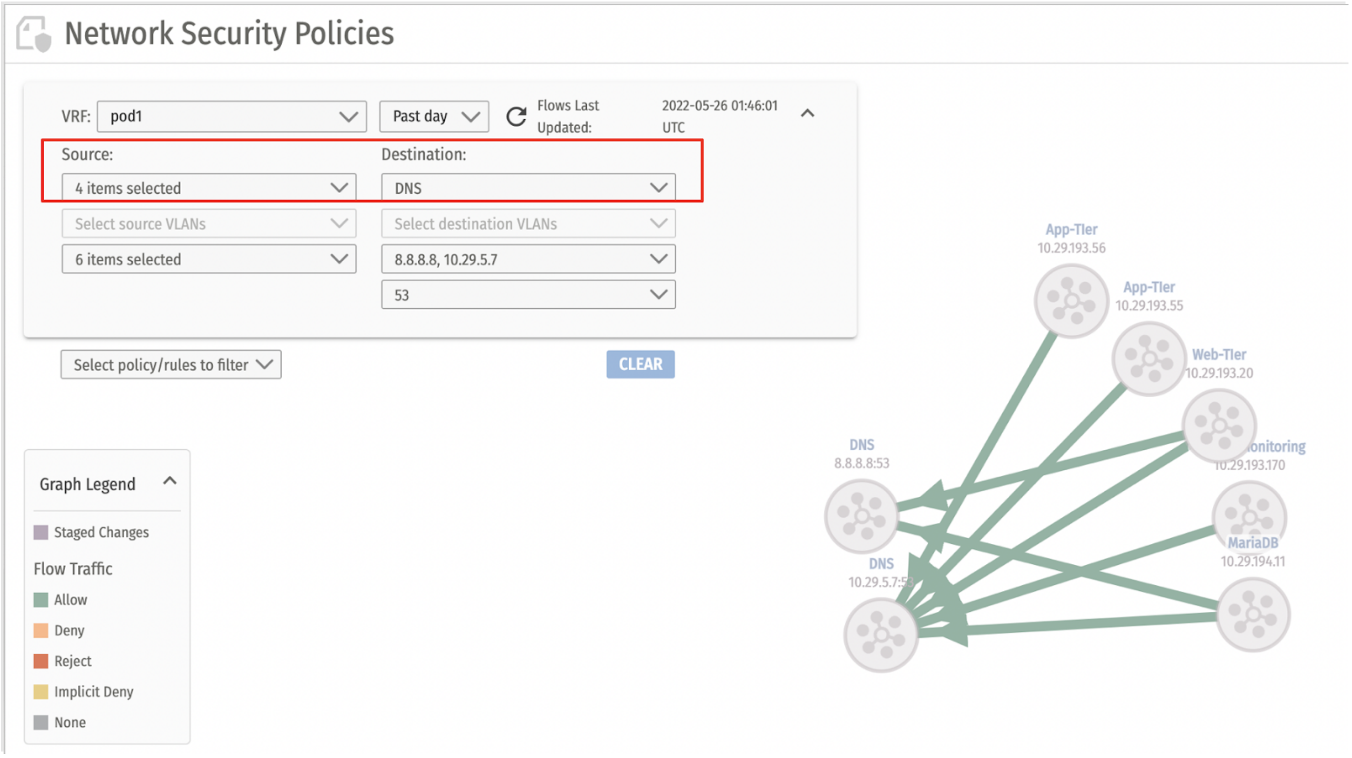


Unique Flows (24)

☐	Source IP	Destination IP	Destination Port	Protocol	#
☐	10.29.193.20	8.8.8.8	53	udp	1245
☐	10.29.193.20	10.29.5.7	53	udp	1235
☐	10.29.193.170	10.29.5.7	53	tcp	497
☐	10.29.193.56	10.29.5.7	53	tcp	492
☐	10.29.193.55		53	tcp	492
☐	10.29.193.59		53	tcp	490
☐	10.29.193.170	10.29.193.55	5666	tcp	90
☐	10.29.193.170	10.29.193.56	5666	tcp	84

Figure 175. Creating/Updating a user-defined group by right-clicking on a row and selecting Add to Group

After creating user-defined groups, filter the flows under analysis by clicking on `User-Defined Groups` for the desired source and destination. The result is a rendered graph of the relationships between the groups of endpoints, as shown in Figure 176:



Network Security Policies

VRF: pod1 | Past day | Flows Last Updated: 2022-05-26 01:46:01 UTC

Source: 4 items selected

Destination: DNS

Select source VLANs

Select destination VLANs

6 items selected

8.8.8.8, 10.29.5.7

53

Select policy/rules to filter

CLEAR

Graph Legend

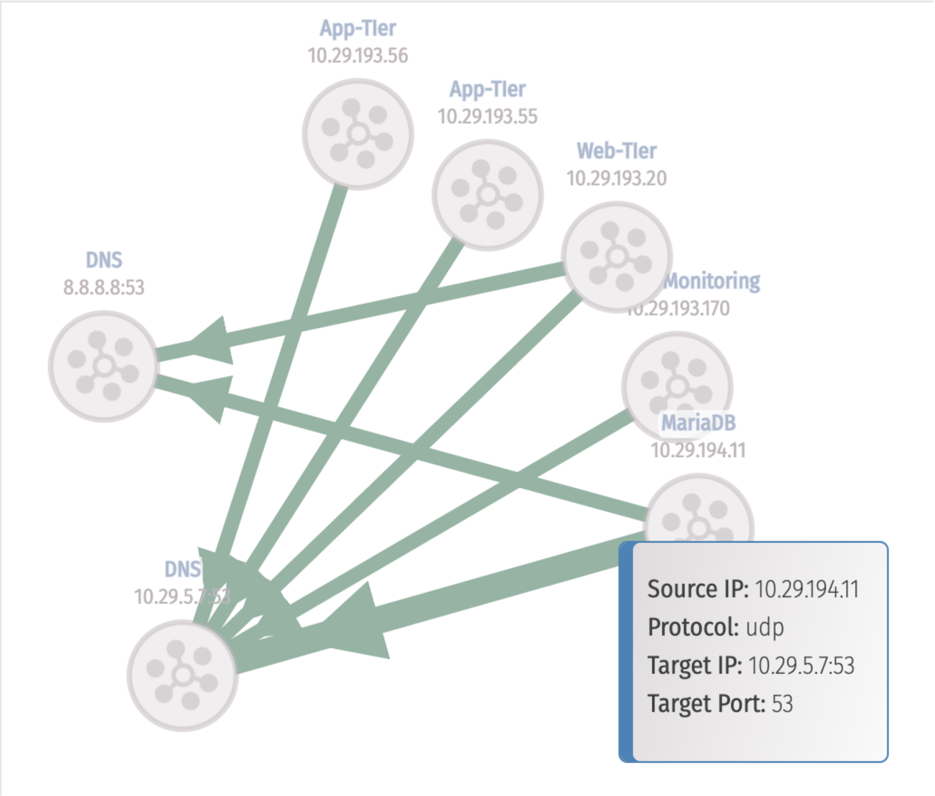
- ▨ Staged Changes

Flow Traffic

- 🟩 Allow
- 🟧 Deny
- 🟥 Reject
- 🟨 Implicit Deny
- ⬜ None

Figure 177. Selecting user-defined groups to render a graph representing UDG’s relationships

Hover over a directional line to highlight the 4-tuple information of the connection as shown in Figure 178:



Source IP: 10.29.194.11
Protocol: udp
Target IP: 10.29.5.7:53
Target Port: 53

Figure 178. Hovering over directional line between MariaDB and DNS displays the 4-tuple information for the connection

To create rules based on the connections rendered in the graph, click on the directional lines of interest and right click on any of the directional lines selected as shown in Figure 179. The result is the connections are added as a new rule, as shown in Figure 180. The predefined source, destination, ports and protocols are pre-populated to form a rule, which can be added to an existing or new policy. Provide a policy and rule name. If a new policy, the policy is not attached to a network or VRF automatically. Remember to click save.

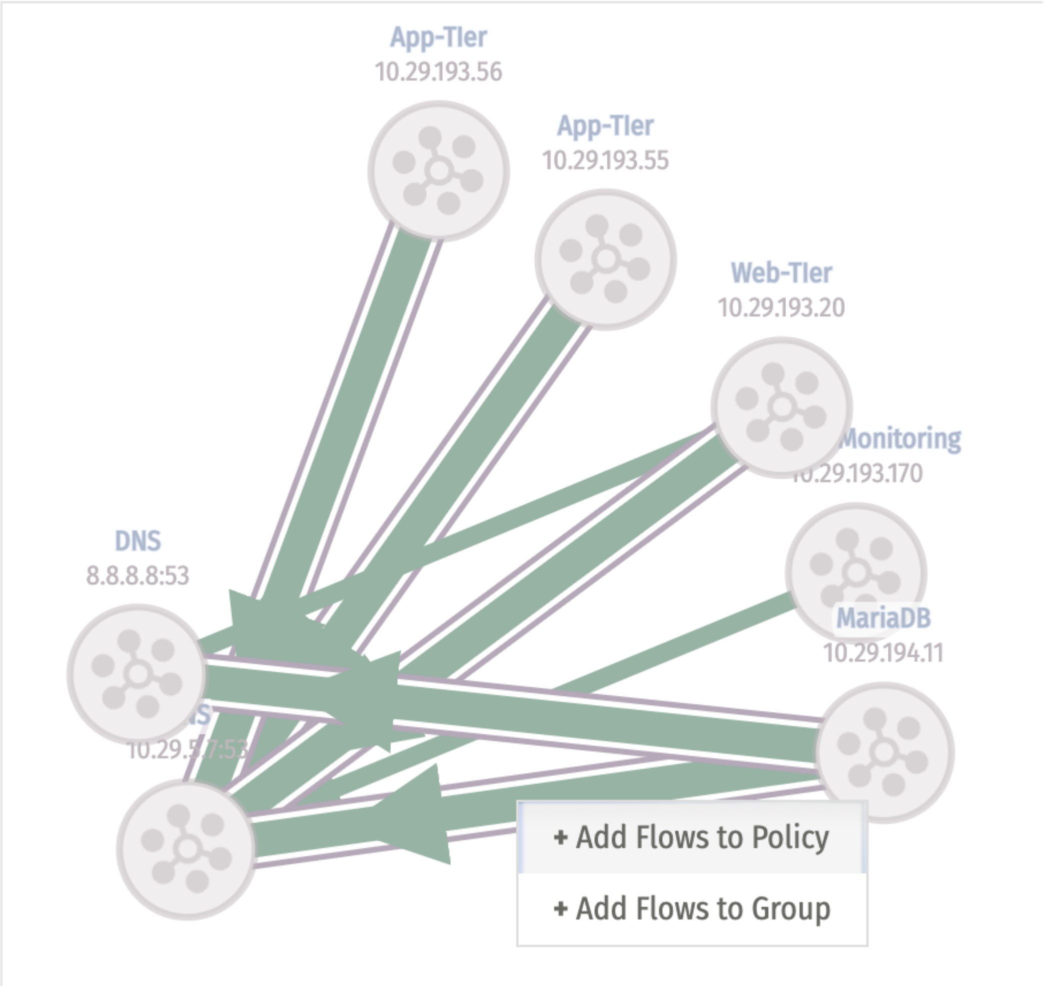
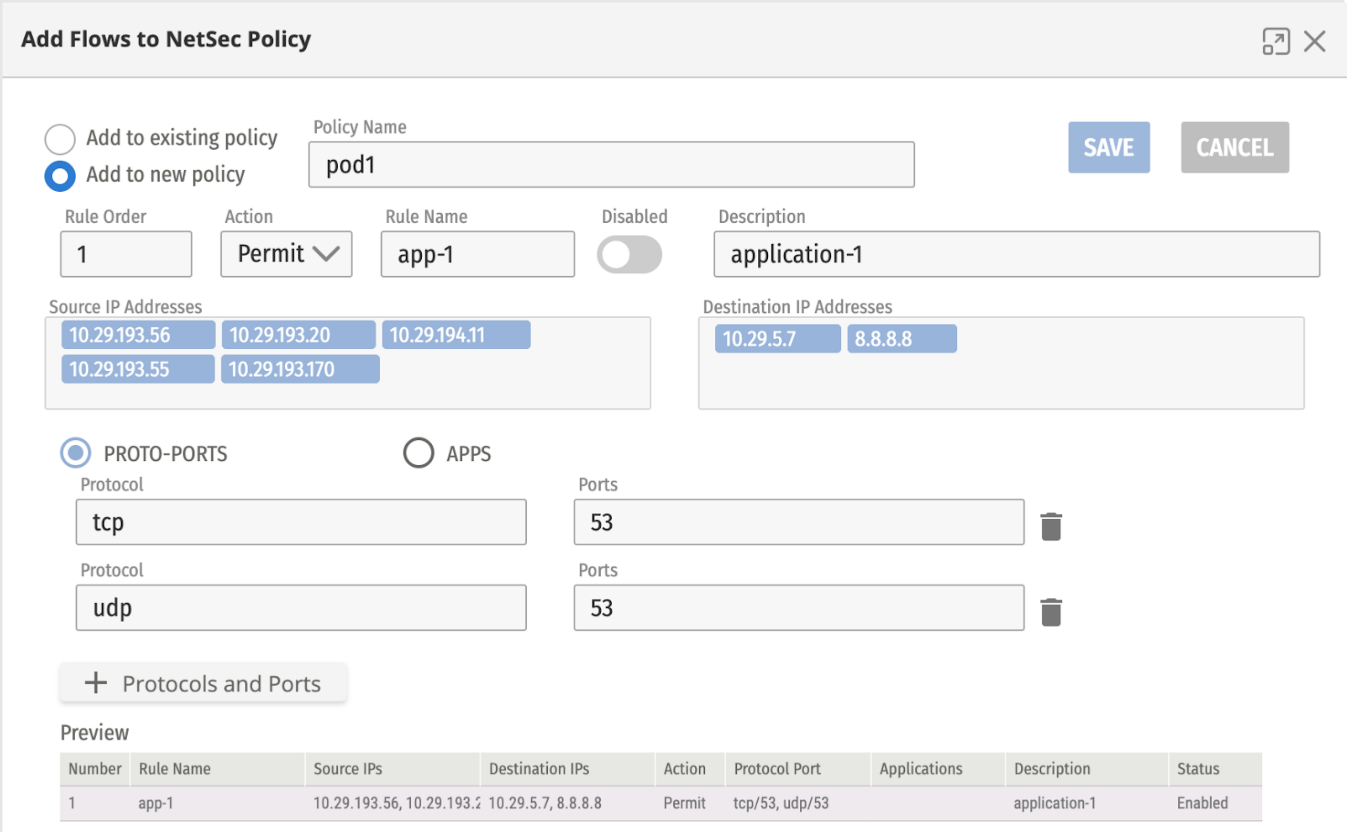


Figure 179. Click on the green directional lines to highlight and right click on a highlighted line to add to policy



Add Flows to NetSec Policy

○ Add to existing policy
● Add to new policy

Policy Name: pod1

Rule Order	Action	Rule Name	Disabled	Description
1	Permit	app-1	○	application-1

Source IP Addresses: 10.29.193.56, 10.29.193.20, 10.29.194.11, 10.29.193.55, 10.29.193.170

Destination IP Addresses: 10.29.5.7, 8.8.8.8

● PROTO-PORTS ○ APPS

Protocol	Ports
tcp	53
udp	53

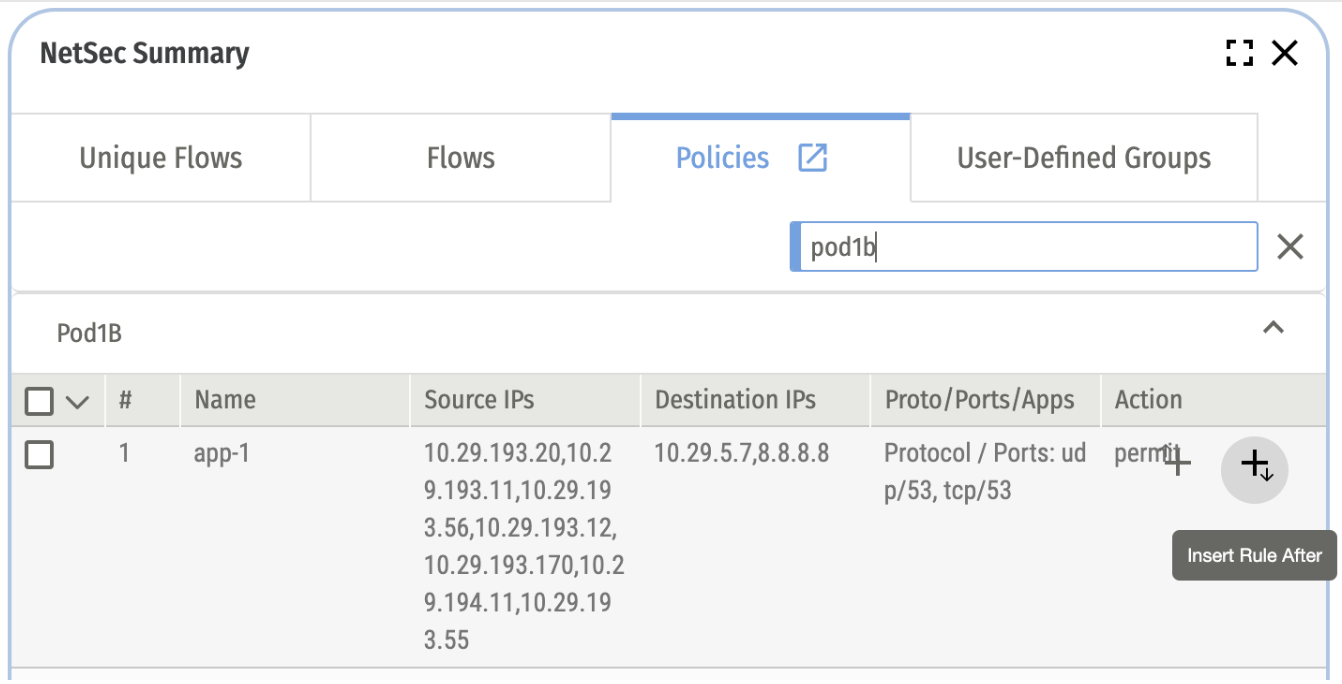
Preview

Number	Rule Name	Source IPs	Destination IPs	Action	Protocol Port	Applications	Description	Status
1	app-1	10.29.193.56, 10.29.193.2	10.29.5.7, 8.8.8.8	Permit	tcp/53, udp/53		application-1	Enabled

Figure 180. The highlighted connections (Figure 142) are pre-populated as a rule. Provide a name, description, select a policy and click SAVE.

To continue with policy discovery, it's also essential to add a “permit any any” to identify what flows would have been denied. Using a “permit any any” rule at the end of the policy assists with policy evaluation while iterating through the discovery of new flows and evaluating new time dimensions of the firewall logs within the VRF and other selected filters.

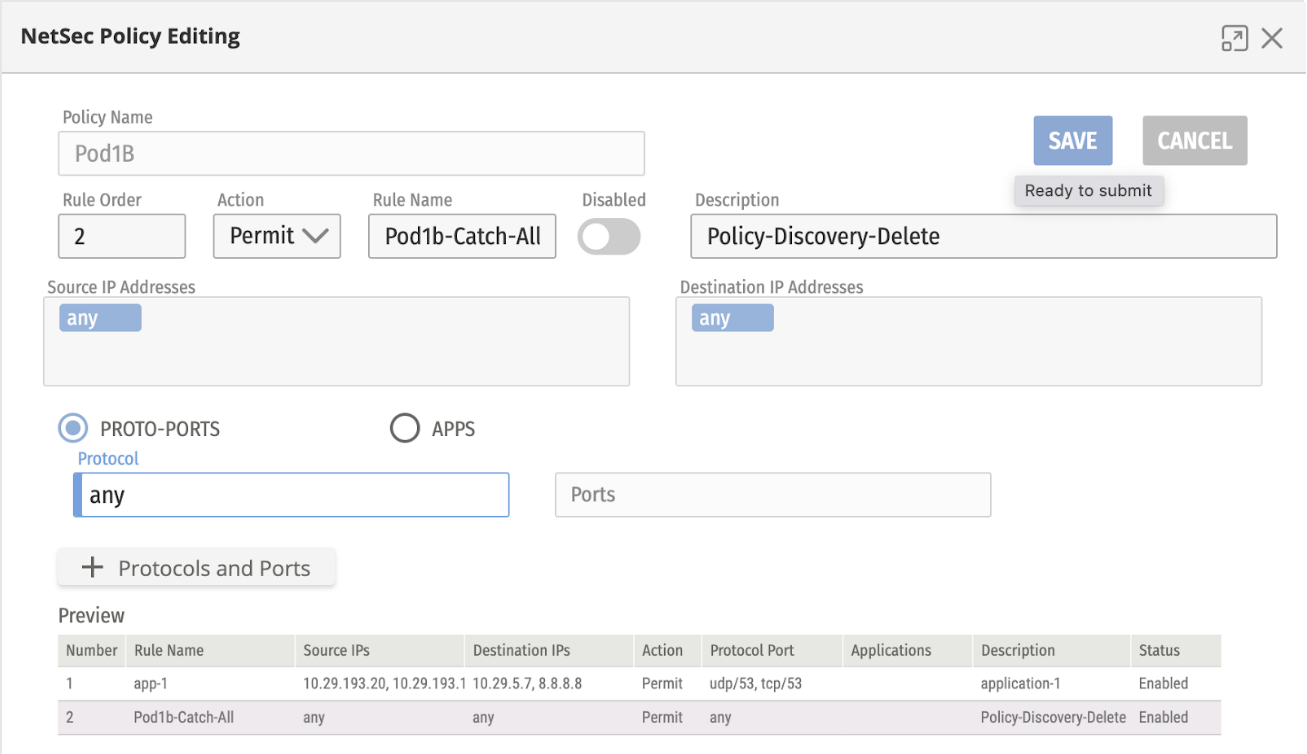
Select the Policies tab, search for the policy name, and click `Insert Rule After` as shown in figure 144. Enter a rule as shown below to catch-all traffic that didn't match rule 1.



□	#	Name	Source IPs	Destination IPs	Proto/Ports/Apps	Action
□	1	app-1	10.29.193.20,10.29.193.11,10.29.193.56,10.29.193.12,10.29.193.170,10.29.194.11,10.29.193.55	10.29.5.7,8.8.8.8	Protocol / Ports: udp/53, tcp/53	permit

Figure 181. Filter for policy name and Insert a catch-all rule after specific permit rules

Enter a rule as shown in Figure 182 to catch all traffic that didn't match rule 1. Remember to click SAVE.



NetSec Policy Editing

Policy Name: Pod1B

Rule Order: 2

Action: Permit

Rule Name: Pod1b-Catch-All

Disabled: ○

Description: Policy-Discovery-Delete

Source IP Addresses: any

Destination IP Addresses: any

● PROTO-PORTS ○ APPS

Protocol: any

Ports

+ Protocols and Ports

Preview

Number	Rule Name	Source IPs	Destination IPs	Action	Protocol Port	Applications	Description	Status
1	app-1	10.29.193.20, 10.29.193.1	10.29.5.7, 8.8.8.8	Permit	udp/53, tcp/53		application-1	Enabled
2	Pod1b-Catch-All	any	any	Permit	any		Policy-Discovery-Delete	Enabled

Figure 182. Create a "permit any any" rule as a catch-all rule to assist with policy discovery.

Remember to click SUBMIT CHANGES in the top right of the screen.

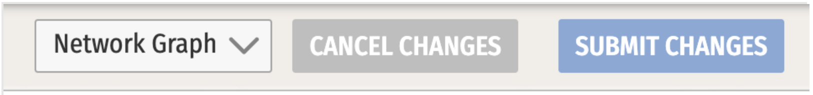
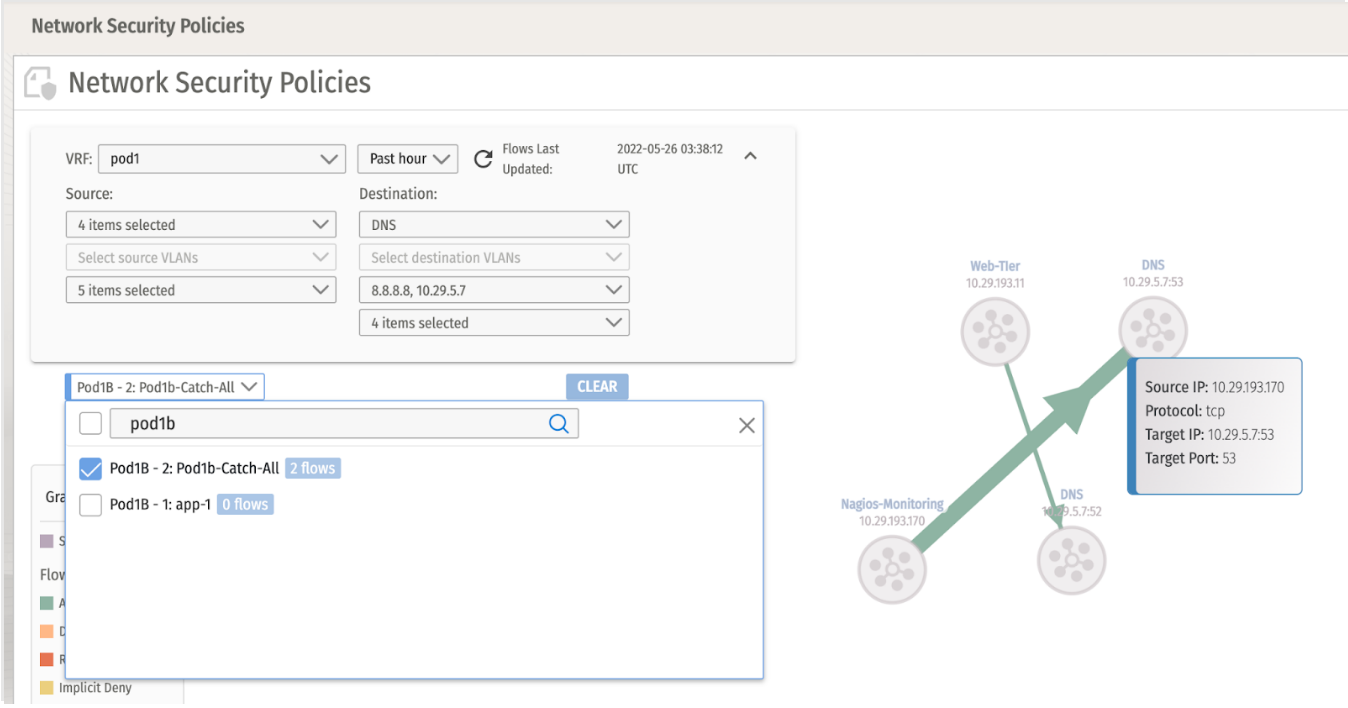


Figure 183. SUBMIT CHANGES button

At the next time domain of collected firewall logs for the VRF discovering policy, follow the steps below to filter flows to analyze. Click on Select policy/rules to filter the policy/rules to analyze what flows match the specified permit rules and what flows match the catch-all "permit any" rule. Any such flows that did not match a specific permit rule and matched the catch-all rule require further analysis to determine if these flows should be permitted with a specific rule. Through several iterations this process will help create a network zero trust policy for a micro, macro, or zone based segmentation strategy. In the example in Figure 184, one hour later, two new unique flows are net new and matched on the catch-all. The catch-all flows must be evaluated whether they should be permitted flows or not. The rendered graph displays the unique flows that would have been implicitly denied if not for the catch-all permit. Iterate through this process until you are satisfied with a zero trust implicit deny at the end of the policy.



Network Security Policies

Network Security Policies

VRF: pod1

Past hour

Flows Last Updated: 2022-05-26 03:38:12 UTC

Source:

4 items selected

Select source VLANs

5 items selected

Destination:

DNS

Select destination VLANs

8.8.8.8, 10.29.5.7

4 items selected

Pod1B - 2: Pod1b-Catch-All

CLEAR

□ pod1b

☑ Pod1B - 2: Pod1b-Catch-All 2 flows

□ Pod1B - 1: app-1 0 flows

Source IP: 10.29.193.170
Protocol: tcp
Target IP: 10.29.5.7:53
Target Port: 53

Figure 184. Select policy/rules to filter and search for a policy name. Click on the catch-all rule to render a graph showing new flows discovered.