

Network Optimization and High Availability Configuration Guide for vEdge Routers, Cisco SD-WAN Releases 19.1, 19.2, and 19.3

First Published: 2019-07-02

Last Modified: 2019-12-22

Americas Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
http://www.cisco.com
Tel: 408 526-4000
800 553-NETS (6387)
Fax: 408 527-0883

© 2019 Cisco Systems, Inc. All rights reserved.



CONTENTS

CHAPTER 1 **What's New for Cisco SD-WAN** 1

What's New for Cisco SD-WAN Release 19.2.x 1

CHAPTER 2 **Network Optimization Overview** 3

Cloud OnRamp for IaaS 3

Provision vManage for Cloud OnRamp for IaaS 4

Configure Cloud OnRamp for IaaS for AWS 8

Configure Cloud OnRamp for IaaS for Azure 14

Troubleshoot Cloud OnRamp for IaaS 19

Cloud OnRamp for SaaS 22

Enable Cloud OnRamp for SaaS 23

Configure Cloud OnRamp for SaaS 24

Monitor Performance of Cloud OnRamp for SaaS 26

Cloud OnRamp for Colocation Solution Overview 27

Manage Clusters 28

Provision and Configure Cluster 29

Create and Activate Clusters 30

Cluster Settings 34

View Cluster 36

Edit Cluster 37

Remove Cluster 38

Reactivate Cluster 38

Create Service Chain in a Service Group 39

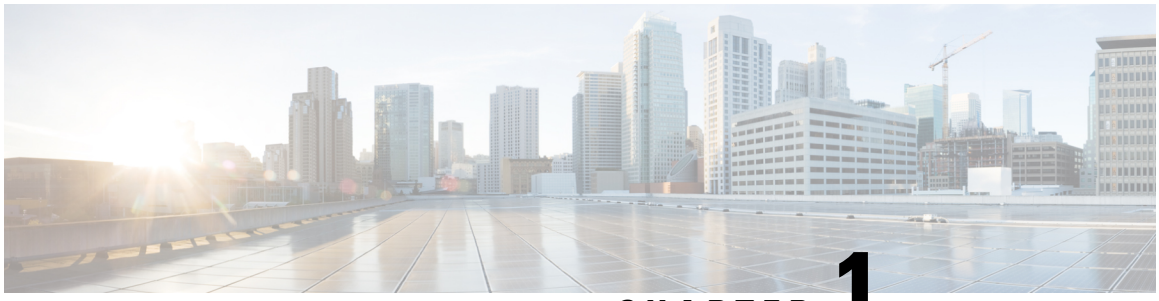
Create Custom Service Chain 43

Custom Service Chain with Shared PNF Devices 44

Configure PNF and Catalyst 9500 47

Custom Service Chain with Shared VNF Devices 48
Shared VNF Use Cases 49
View Service Groups 56
Edit Service Group 56
Attach and Detach Service Group with Cluster 57
View Information About VNFs 58
View Cisco Colo Manager Health from vManage 59
Monitor Cloud OnRamp Colocation Clusters 61
Manage VM Catalog and Repository 64
Upload VNF Images 65
Create Customized VNF Image 66
View VNF Images 71
Delete VNF Images 71
Upgrade NFVIS Software Through vManage 72
Upload NFVIS Upgrade Image 72
Upgrade CSP Device with NFVIS Upgrade Image 72
TCP Optimization: Cisco vEdge Devices 74
Configure TCP Optimization 74
Monitor TCP Optimization 76

CHAPTER 3 **High Availability Overview** 79
vBond Orchestrator Redundancy 83
vManage NMS Redundancy 84
vSmart Controller Redundancy 91
Cisco vEdge Device Redundancy 92
Configure Affinity between vSmart and Cisco vEdge Device 92
Configure Control Plane and Data Plane High Availability Parameters 96
Configure Disaster Recovery 97
Disaster Recovery with Manual Switchover 101
High Availability CLI Reference 104
High Availability Configuration Examples 105



CHAPTER 1

What's New for Cisco SD-WAN

Note The documentation set for this product strives to use bias-free language. For purposes of this documentation set, bias-free is defined as language that does not imply discrimination based on age, disability, gender, racial identity, ethnic identity, sexual orientation, socioeconomic status, and intersectionality. Exceptions may be present in the documentation due to language that is hardcoded in the user interfaces of the product software, language used based on RFP documentation, or language that is used by a referenced third-party product.

This chapter describes what's new in Cisco SD-WAN for each release.

- What's New for Cisco SD-WAN Release 19.2.x, on page 1

What's New for Cisco SD-WAN Release 19.2.x

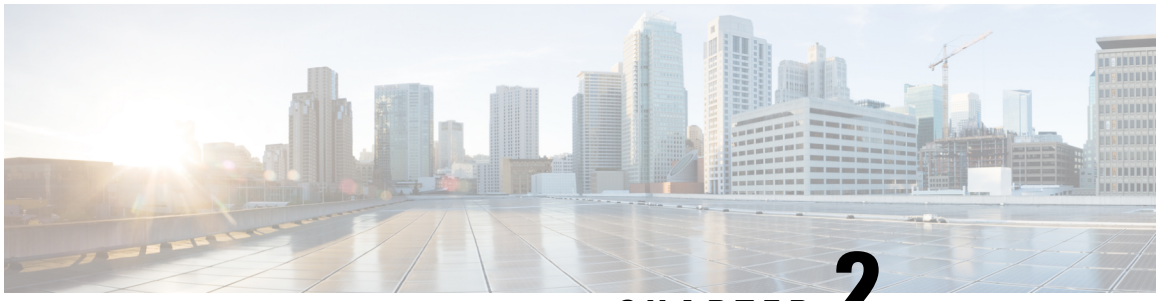
This section applies to Cisco vEdge devices.

Cisco is constantly enhancing the SD-WAN solution with every release and we try and keep the content in line with the latest enhancements. The following table lists new and modified features we documented in the Configuration, Command Reference, and Hardware Installation guides. For information on additional features and fixes that were committed to the SD-WAN solution, see the *Resolved and Open Bugs* section in the Release Notes.

Table 1: What's New for Cisco vEdge Device

Feature	Description
Getting Started	
API Cross-Site Request Forgery Prevention	This feature adds protection against Cross-Site Request Forgery (CSRF) that occurs when using Cisco SD-WAN REST APIs. This protection is provided by including a CSRF token with API requests. You can put requests on an allowed list so that they do not require protection if needed. See Cross-Site Request Forgery Prevention.
Systems and Interfaces	

Feature	Description
Secure Shell Authentication Using RSA Keys	This feature helps configure RSA keys by securing communication between a client and a Cisco SD-WAN server. See SSH Authentication using vManage on Cisco XE SD-WAN Devices. See Configure SSH Authentication.
Policies	
Packet Duplication for Noisy Channels	This feature helps mitigate packet loss over noisy channels, thereby maintaining high application QoE for voice and video. See Configure and Monitor Packet Duplication.
Control Traffic Flow Using Class of Service Values	This feature lets you control the flow of traffic into and out of a Cisco device's interface based on the conditions defined in the quality of service (QoS) map. A priority field and a layer 2 class of service (CoS) were added for configuring the re-write rule. See Configure Localized Data Policy for IPv4 Using Cisco vManage.
Security	
Secure Communication Using Pairwise IPsec Keys	This feature allows private pairwise IPSec session keys to be created and installed for secure communication between IPSec devices and its peers. For related information, see IPSec Pairwise Keys Overview.
Network Optimization and High Availability	
Disaster Recovery for vManage	This feature helps you configure Cisco vManage in an active or standby mode to counteract hardware or software failures that may occur due to unforeseen circumstances. See Configure Disaster Recovery.
Share VNF Devices Across Service Chains	This feature lets you share Virtual Network Function (VNF) devices across service chains to improve resource utilisation and reduce resource fragmentation. See Share VNF Devices Across Service Chains.
Monitor Service Chain Health	This feature lets you configure periodic checks on the service chain data path and reports the overall status. To enable service chain health monitoring, NFVIS version 3.12.1 or later should be installed on all CSP devices in a cluster. See Monitor Service Chain Health.
Manage PNF Devices in Service Chains	This feature lets you add Physical Network Function (PNF) devices to a network, in addition to the Virtual Network function (VNF) devices. These PNF devices can be added to service chains and shared across service chains, service groups, and a cluster. Inclusion of PNF devices in the service chain can overcome the performance and scaling issues caused by using only VNF devices in a service chain. See Manage PNF Devices in Service Chains.



CHAPTER 2

Network Optimization Overview

- Cloud OnRamp for IaaS, on page 3
- Cloud OnRamp for SaaS, on page 22
- Cloud OnRamp for Colocation Solution Overview, on page 27
- TCP Optimization: Cisco vEdge Devices, on page 74

Cloud OnRamp for IaaS

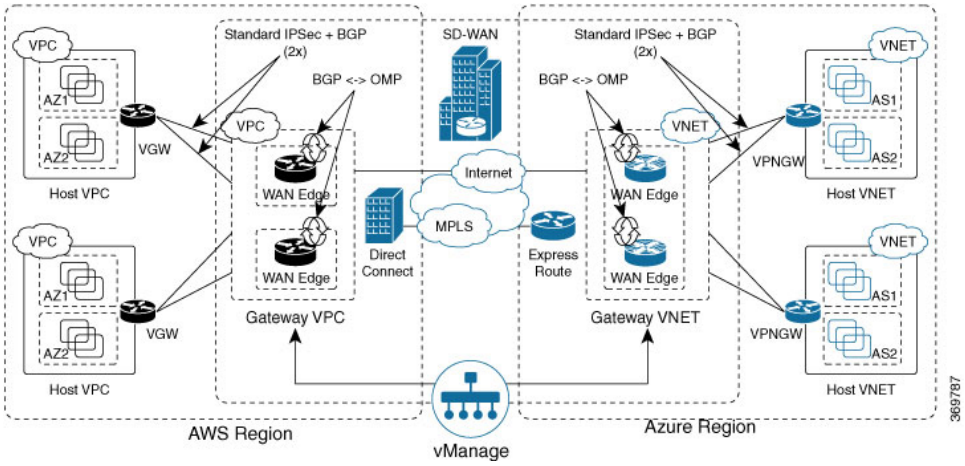
Cloud OnRamp for IaaS extends the fabric of the Cisco SD-WAN overlay network into public clouds, allowing branches with vEdge routers to connect directly to public-cloud application providers. By eliminating the need for a physical data center, Cloud OnRamp for IaaS improves the performance of IaaS applications.

The connection between the overlay network and a public-cloud application is provided by two or four pairs of redundant vEdge Cloud routers for AWS, which act together as a transit between the overlay network and the application. By using redundant routers to form the transit offers path resiliency to the public cloud. In addition, having redundant routers improves the availability of public-cloud applications. Together, the two routers can remediate in the event of link degradation. You create these routers as part of the Cloud OnRamp workflow.

Cloud OnRamp for IaaS discovers any already existing private cloud instances in geographical cloud regions and allows you to select which of them to make available for the overlay network. In such a scenario, Cloud OnRamp for IaaS allows simple integration between legacy public-cloud connections and the Cisco SD-WAN overlay network.

You configure and manage Cloud OnRamp for IaaS through the vManage NMS server. A configuration wizard in the vManage NMS automates the bring-up of the transit to a your public cloud account and automates the connections between public-cloud applications and the users of those applications at branches in the overlay network.

The Cloud OnRamp for IaaS works in conjunction with AWS virtual private clouds (VPCs) and Azure virtual networks (VNets). The following image provides a high level overview of multi-cloud onRamp for IaaS.



Supported Routers

Cloud OnRamp for IaaS is supported on Cisco Cloud vEdge and Cisco Cloud Services Routers (CSRs). In this topic, supported routers are referred to collectively as *cloud routers*.

Provision vManage for Cloud OnRamp for IaaS

Before you configure Cloud OnRamp for IaaS, ensure that you provision the vManage NMS, AWS, and Azure.

vManage NMS Prerequisites

Before you can configure Cloud OnRamp for IaaS, you must properly provision the vManage NMS.

- Ensure that your vManage server has access to the internet and that it has a DNS server configured so that it can reach AWS. To configure a DNS server, in the vManage VPN feature configuration template, enter the IP address of a DNS server, and then reattach the configuration template to the vManage server.
- Ensure that two cloud routers that are to be used to bring up the Cloud OnRamp for IaaS have been added to the vManage NMS and have been attached to the appropriate configuration template. (These two routers are deployed in AWS in their own VPC, and together they form the transit VPC, which is the bridge between the overlay network and AWS cloud applications.) Ensure that the configuration for these routers includes the following:
    - Hostname
    - IP address of vBond orchestrator
    - Site ID
    - Organization name
    - Tunnel interface configuration on the eth1 interface
- Ensure that the vManage NMS is synchronized with the current time. To check the current time, click the Help (?) icon in the top bar of any vManage screen. The Timestamp field shows the current time. If the time is not correct, configure the vManage server's time to point to an NTP time server, such as the Google NTP server. To do this, in the vManage NTP feature configuration template, enter the hostname

of an NTP server, and then reattach the configuration template to the vManage server. The Google NTP servers are time.google.com, time2.google.com, time3.google.com, and time4.google.com

AWS Prerequisites

Before you can configure Cloud OnRamp for IaaS, ensure that you provision AWS properly.

- Ensure that you have subscribed to the Viptela marketplace Amazon machine images (AMIs) and the Cisco CSR AMIs in your AWS account. See *Subscribe to Cisco SD-WAN AMIs*.
- Ensure that at least one user who has administrative privileges has the AWS API keys for your AWS account. For Cloud OnRamp for IaaS, these keys are used to authenticate the vManage server with AWS and to bring up the VPC and Elastic Compute Cloud (EC2) instances.
- Check the AWS limits associated with your account (in the Trusted Advisor section of AWS) to ensure that the following resources can be created in your account:
  - 1 VPC, which is required for creating the transit VPC
  - 6 Elastic IP addresses associated with each pair of transit vEdge Cloud routers
  - 1 AWS virtual transit (VGW) for each host VPC
  - 4 VPN connections for mapping each host VPC



Note vEdge Cloud Routers support C3 and C4 compute-intensive families.

Subscribe to Cisco SD-WAN AMIs

To use the Cloud OnRamp for IaaS and other Cisco SD-WAN services, you must subscribe to the Amazon Machine Image (AMI) for your router in AWS. When you subscribe, you can complete the following tasks:

- Launch a cloud router AMI instance
- Generate a key pair to use for the instance
- Use the key pair to subscribe to the cloud router instance.

You subscribe to the vEdge Cloud router AMI only once, when you first create a Viptela AMI instance.

To create a new AMI subscription and generate a key pair:

1. In AWS, search to locate a cloud router AMI for your devices.
2. Select and launch an EC2 instance with the AMI instance. For more information, see *Create Cisco IOS XE SD-WAN Cloud VM Instance on AWS*.
3. Generate a key pair. For full instructions, see *Set Up the Cisco SD-WAN Cloud VM Instance*.
4. Click **Download Key Pair**. The key pair then downloads to your local computer as a .pem file.
5. Click **Launch Instance**. A failure message displays, because you now need to upload the key pair to complete the subscription process.

To upload the key pair:

1. In AWS Marketplace, search for your router AMI.
2. Click **Continue**.
3. Click **Key Pair** to bring up a vEdge Cloud router instance. In the option to enter the key pair, upload the .pem file from your local computer. This is the file that you had generated in Step 3 when creating a new AMI subscription.

Azure Prerequisites

Before you can configure Cloud OnRamp for IaaS, you must properly provision Azure.

- Ensure that you have accepted the terms and conditions for the Cisco Cloud vEdge Router in the Azure Marketplace. See *Accept the Azure Terms of Service*.
- Ensure that you create an App Registration in Azure and retrieve the credentials for your Azure account. For Cloud OnRamp for IaaS, these credentials are used to authenticate the vManage server with Azure and bring up the VNet and the Virtual Machine instances. See *Create and Retrieve Azure Credentials*.
- Check the Azure limits associated with your account (by going to your subscription in the portal and checking Usage + Quotas) to ensure that the following resources can be created in your account:
  - 1 VNet, which is required for creating the transit VNet
  - 1 Availability set, required for Virtual Machine distribution in the transit VNet
  - 6 Static Public IP addresses associated with the transit cloud routers
  - 1 Azure Virtual Network Gateway and 2 Static Public IP Addresses for each host VNet
  - 4 VPN connections for mapping each host VNet
- F-series VMs (F4 and F8) are supported on the cloud routers.

Accept the Azure Terms of Service

To use a Cisco cloud router as part of the Cloud OnRamp workflow, you must accept marketplace terms for using a virtual machine (VM). You can do this in one of the following ways:

- Spin up the cloud router on the portal manually, and accept the terms as part of the final page of the bringup wizard.
- In the Azure APIs or Powershell/Cloud Shell, use the Set-AzureRmMarketplaceTerms command.

Create and Retrieve Azure Credentials

To create and retrieve Azure credentials, you must create an App Registration in Azure with Contributor privileges:

1. Launch the Microsoft Azure portal.
2. Create an application ID:
   a. In the left pane of the Azure portal, click **Azure Active Directory**.
   b. In the sub-menu, click **App registrations**.

c. Click **New application registration**. The system displays the Create screen.

d. In the **Name** field, enter a descriptive name such as CloudOnRampApp.

e. In the **Application Type** field, select **Web app / API**

f. In the **Sign-on URL** field, enter any valid sign-on URL; this URL is not used in Cloud OnRamp.

g. Click **Create**. The system displays a summary screen with the Application ID.

3. Create a secret key for the Cloud OnRamp application:

   a. In the summary screen, click **Settings** in the upper-left corner.

   b. In the right pane, click **Keys**. The system displays the **Keys** > **Password** screen.

   c. On the Passwords screen:

      1. In the **Description** column, enter a description for your secret key.

      2. In the **Expires** column, from the **Duration** drop-down, select the duration for your secret key.

      3. Click **Save** in the upper-left corner of the screen. The system displays the secret key in the Value column but then hides it permanently, so be sure to copy and save the password in a separate location.

4. In the left pane of the Azure portal, click **Subscriptions** to view the subscription ID. If you have multiple subscriptions, copy and save the subscription ID which you are planning to use for configuring the Cloud OnRamp application.

5. View the Tenant ID:

   a. In the left pane of the Azure portal, click **Azure Active Directory**.

   b. Click **Properties**. The system displays the directory ID which is equivalent to the tenant ID.

6. Assign Contributor privileges to the application:

   a. In the left pane of the Azure portal, click **Subscriptions**.

   b. Click the subscription that you will be using for the Cloud OnRamp application.

   c. In the subscription pane, navigate to Access Control (IAM).

   d. Click **Add**. The system displays the Add Permissions screen.

   e. From the **Role** drop-down menu, select **Contributor**.

   f. From the **Assign Access To** drop-down, select the default value **Azure AD user**, **group**, or **application**.

   g. From the **Select** drop-down, select the application you just created for Cloud OnRamp.

   h. Click **Save**.

You can now log into the Cloud OnRamp application with the Azure credentials you just created and saved.

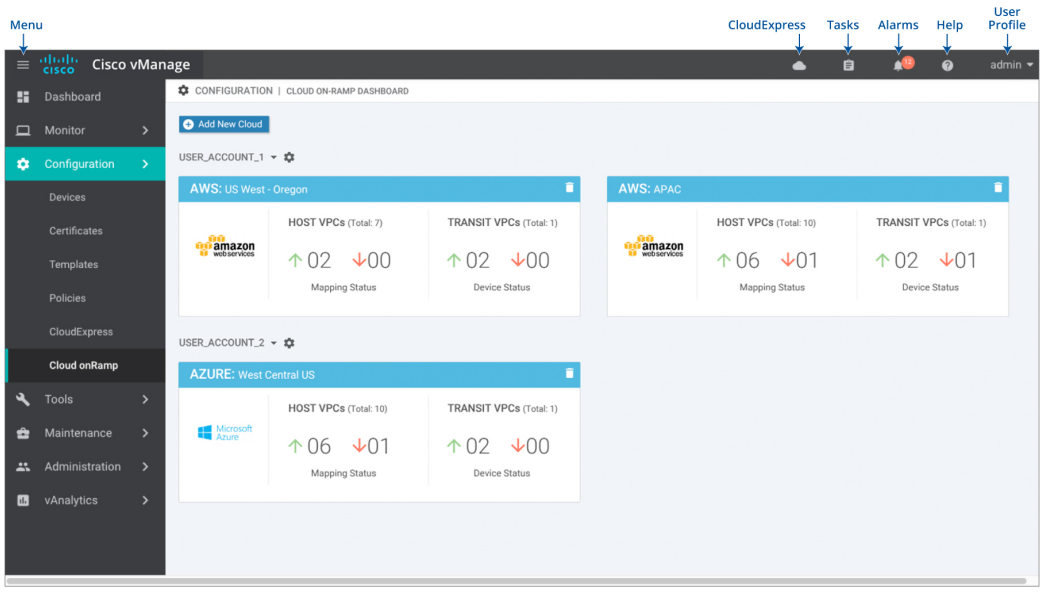
Configure Cloud OnRamp for IaaS for AWS

Configure Cloud OnRamp for IaaS for AWS

To configure Cloud OnRamp for IaaS for AWS, you create AWS transit VPCs, each of which consists of up to four pairs of Cisco vEdge devices. You then map the transit virtual private clouds (VPC)s to host VPCs that already exist in the AWS cloud.

- Transit VPCs provide the connection between the Cisco overlay network and the cloud-based applications running on host VPCs. Each transit VPC consists of up to four pairs of cloud routers that reside in their own VPC. Multiple routers are used to provide redundancy for the connection between the overlay network and cloud-based applications. On each of these two cloud routers, the transport VPN (VPN 0) connects to a branch router, and the service-side VPNs (any VPN except for VPN 0 and VPN 512) connect to applications and application providers in the public cloud.
- Cloud OnRamp supports auto-scale for AWS. To use auto-scale, ensure that you associate two to four pairs of cloud routers to a transit VPC. Each of the devices that are associated with the transit VPC for auto-scale should have a device template attached to it.
- Host VPCs are virtual private clouds in which your cloud-based applications reside. When a transit VPC connects to an application or application provider, it is simply connecting to a host VPC.
- All host VPCs can belong to the same account, or each host VPC can belong to a different account. A host that belongs one account can be mapped to a transit VPC that belongs to a completely different account. You configure cloud instances by using a configuration wizard.

1. In vManage NMS, select the **Configuration** > **Cloud onRamp for IaaS** screen.



2. Click **Add New Cloud Instance**.
3. In the Add Cloud Instance – log in to a Cloud Server popup:
   a. In the **Cloud** drop-down, select the **Amazon Web Services** radio button.

b. Click **IAM Role** or **Key** to log in to the cloud server. It is recommended that you use IAM Role.

c. If you select **IAM Role**:

1. In the **Role ARN** field, enter the role ARN of the IAM role.
2. In the **External ID** field, enter external ID created for the role ARN. It is recommended that the external ID include 10 to 20 characters in random order. To authenticate to the vManage NMS using an IAM role, vManage NMS must be hosted by Cisco on AWS and have the following attributes:
   - Trusts the AWS account, 200235630647, that hosts the vManage NMS.
   - Have all permissions for EC2 and VPC resources.
   - A default timeout of at least one hour.

   If vManage NMS is not hosted by Cisco on AWS, assign an IAM role with permissions to AssumeRole to the vManage server running the Cloud OnRamp process. Refer to the AWS documentation for details.

d. If you select **Key**:

1. In the **API Key** field, enter your Amazon API key.
2. In the **Secret Key** field, enter the password associated with the API key.

4. Click **Login** to log in to the cloud server.

   The cloud instance configuration wizard opens. This wizard consists of three screens that you use to select a region and discover host VPCs, add transit VPC, and map host VPCs to transit VPCs. A graphic on the right side of each wizard screen illustrates the steps in the cloud instance configuration process. The steps that are not yet completed are shown in light gray. The current step is highlighted within a blue box. Completed steps are indicated with a green checkmark and are shown in light orange.

5. Select a region:

   a. In the **Choose Region** drop-down, choose a geographical region.

   b. Click **Save and Finish** to create a transit VPC or optionally click **Proceed to Discovery and Mapping** to continue with the wizard.

6. Add a transit VPC:

   a. In the **Transit VPC Name** field, type a name for the transit VPC.

   The name can be up to 128 characters and can contain only uppercase and lowercase letters, the digits 0 through 9, hyphens (–), and underscores (_). It cannot contain spaces or any other characters.

   b. Under **Device Information**, enter information about the transit VPC:

   1. In the **WAN Edge Version** drop-down, select the software version of the cloud router to run on the transit VPC.
   2. In the **Size of Transit WAN Edge** drop-down, choose how much memory and how many CPUs to use for each of the cloud routers that run on the transit VPC.

3. In the **Max. Host VPCs per Device Pair** field, select the maximum number of host VPCs that can be mapped to each device pair for the transit VPC. Valid values are 1 through 32.
4. In the **Device Pair 1#** fields, select the serial numbers of each device in the pair. To remove a device serial number, click the **X** that appears in a field.

   The devices that appear in this field have been associated with a configuration template and support the WAN Edge Version that you selected.
5. To add additional device pairs, click .

   To remove a device pair, click .

   A transit VPC can be associated with one to four device pairs. To enable the Cloud onRamp auto-scale feature for AWS, you must associate at least two device pairs with the transit VPC.
6. Click **Save and Finish** to complete the transit VPC configuration or optionally click **Proceed to Discovery and Mapping** to continue with the wizard.
7. Click **Advanced** if you wish to enter more specific configuration options:
   a. In the **Transit VPC CIDR** field, enter a custom CIDR that has a network mask in the range of 16 to 25. If you choose to leave this field empty, the Transit VPC is created with a default CIDR of 10.0.0.0/16.
   b. In the **SSH PEM Key** drop-down, select a PEM key pair to log in to an instance. Note that the key pairs are region-specific. Refer to the AWS documentation for instructions on creating key pairs.
8. Click **Save and Finish** to complete the transit VPC configuration, or optionally click **Proceed to Discovery and Mapping** to continue with the wizard.
9. Select hosts to discover:
   a. In the **Select an account to discover** field, select a host to map to this transit VPC.
   b. Click **Discover Host VPCs**.
   c. In the table that displays, choose one or more hosts to map to this transit VPC.

      You can use the search field and options to display only host VPCs that mention specific search criteria.

      You can click the **Refresh** icon to update the table with current information.

      You can click the **Show Table Columns** icon to specify which columns display in the table.
   d. Click **Next**.

7. Map the host VPCs to transit VPCs:
   a. In the table of host VPCs, select the desired host VPCs.
   b. Click **Map VPCs**. The Map Host VPCs popup opens.
   c. In the **Transit VPC** drop-down, select the transit VPC to map to the host VPCs.
   d. In the **VPN** drop-down, select the VPN in the overlay network in which to place the mapping.

e. Enable the **Route Propagation** option if you want vManage to automatically propagate routes to the host VPC routes table.

f. Click **Map VPCs**.

g. Click **Save and Complete**.

In the VPN feature configuration template for VPN 0, when configuring the two cloud routers that form the transit VPC, ensure that the color you assign to the tunnel interface is a public color, not a private color. Public colors are **3g**, **biz-internet**, **blue**, **bronze**, **custom1**, **custom2**, **custom3**, **default**, **gold**, **green**, **lte**, **metro-ethernet**, **mpls**, **public-internet**, **red**, and **silver**.

Display Host VPCs

1. In the Cloud OnRamp Dashboard, click the pane for the desired VPC. The Host VPCs/Transit VPCs screen opens, and Host VPCs is selected by default. In the bar below this, Mapped Host VPCs is selected by default, and the table on the screen lists the mapping between host and transit VPCs, the state of the transit VPC, and the VPN ID.
2. To list unmapped host VPCs, click **Unmapped Host VPCs**. Then click **Discover Host VPCs**.
3. To display the transit VPCs, click **Transit VPCs**.

Map Host VPCs to a Transit VPC

1. In the Cloud OnRamp Dashboard, click the pane for the desired VPC. The Host VPCs/Transit VPCs screen opens.
2. Click **Un-Mapped Host VPCs**.
3. Click **Discover Host VPCs**.
4. From the list of discovered host VPCs, select the desired host VPCs
5. Click **Map VPCs**. The Map Host VPCs popup opens.
6. In the **Transit VPC** drop-down, choose the desired transit VPC.
7. In the **VPN** drop-down, choose the VPN in the overlay network in which to place the mapping.
8. Click **Map VPCs**.

Unmap Host VPCs

1. In the Cloud OnRamp Dashboard, click the pane for the desired VPC. The Host VPCs/Transit VPCs screen opens.
2. Click **Mapped Host VPCs**.
3. From the list of VPCs, select the desired host VPCs.
4. Click **Unmap VPCs**.
5. Click **OK** to confirm the unmapping.

Unmapping host VPCs deletes all VPN connections to the VPN gateway in the host VPC, and then deletes the VPN gateway. When you make additional VPN connections to a mapped host VPC, they will be terminated as part of the unmapping process.

Display Transit VPCs

1. In the Cloud OnRamp Dashboard, click the pane for the desired VPC. The Host VPCs/Transit VPCs screen opens, and Host VPCs is selected by default.
2. Click **Transit VPCs**.

The table at the bottom of the screen lists the transit VPCs.

Add Transit VPC

1. In the Cloud OnRamp Dashboard, click the pane for the desired VPC. The Host VPCs/Transit VPCs screen opens, and Host VPCs is selected by default.
2. Click **Transit VPCs**.
3. Click **Add Transit VPC**.

   To add a transit VPC, perform operations from step 6 of Configure Cloud OnRamp for IaaS for AWS, on page 8.

Delete Device Pair

The device pair must be offline.

1. In the Cloud OnRamp Dashboard,
2. Click a device pair ID.
3. Verify that the status of the device pair is offline.
4. To descale the device pairs, click the trash can icon in the Action column or click the **Trigger Autoscale** option.

Delete Transit VPC

Prerequisite: Delete the device pairs that are associated with the transit VPC.



Note To delete the last pair of online device pairs, you must delete a transit VPC.

1. In the Cloud OnRamp Dashboard, click the pane for the desired VPC. The Host VPCs/Transit VPCs screen opens, and Host VPCs is selected by default.
2. Click **Host VPCs**.
3. Select all host VPCs, and click **Unmap VPCs**.

   Ensure that all host mappings with transit VPCs are unmapped.
4. Click **OK** to confirm the unmapping.

5. Click **Transit VPCs**.
6. Click the trash icon to the left of the row for the transit VPC.



Note The trash icon is not available for the last device pair of transit VPC. Hence, to delete the last device pair, click **Delete Transit** drop-down list at the right corner. The trash icon is only available from the second device pair onwards.

7. Click **OK** to confirm.

Add Device Pairs

1. Click **Add Device Pair**.

   Ensure that the devices you are adding are already associated with a device template.
2. In the box, select a device pair.
3. Click the **Add** icon to add more device pairs.

   You can add up to a total of four device pairs to the transit VPC.
4. Click **Save**.

History of Device Pairs for Transit VPCs

To display the Transit VPC Connection History page with all its corresponding events, click **History for a device pair**.

In this view, by default, a histogram of events that have occurred in the previous one hour is displayed and a table of all events for the selected transit VPC. The table lists all the events generated in the transit VPC. The events can be one of the following:

- Device Pair Added
- Device Pair Spun Up
- Device Pair Spun Down
- Device Pair Removed
- Host Vpc Mapped
- Host Vpc Unmapped
- Host Vpc Moved
- Transit Vpc Created
- Transit Vpc Removed

Edit Transit VPC

You can change the maximum number of host VPCs that can be mapped to a device pair.

1. Click **Edit Transit Details**. Provide a value for the maximum number of host VPCs per device pair to which the transit VPC can be mapped.
2. Click **OK**.

This operation can trigger auto-scale.

Configure Cloud OnRamp for IaaS for Azure

To configure Cloud OnRamp for IaaS for Azure, you create Azure transit VNets, each of which consist of a pair of routers. You then map the host vNets to transit VNets that already exist in the Azure cloud. All VNets reside in the same resource group.

- Transit VNets provide the connection between the overlay network and the cloud-based applications running on host VNet. Each transit VNet consists of two routers that reside in their own VNet. Two routers are used to provide redundancy for the connection between the overlay network and cloud-based applications. On each of these two cloud routers, the transport VPN (VPN 0) connects to a branch router, and the service-side VPNs (any VPN except for VPN 0 and VPN 512) connect to applications and application providers in the public cloud.
- Host VNets are virtual private clouds in which your cloud-based applications reside. When a transit VNet connects to an application or application provider, it is simply connecting to a host VNet.

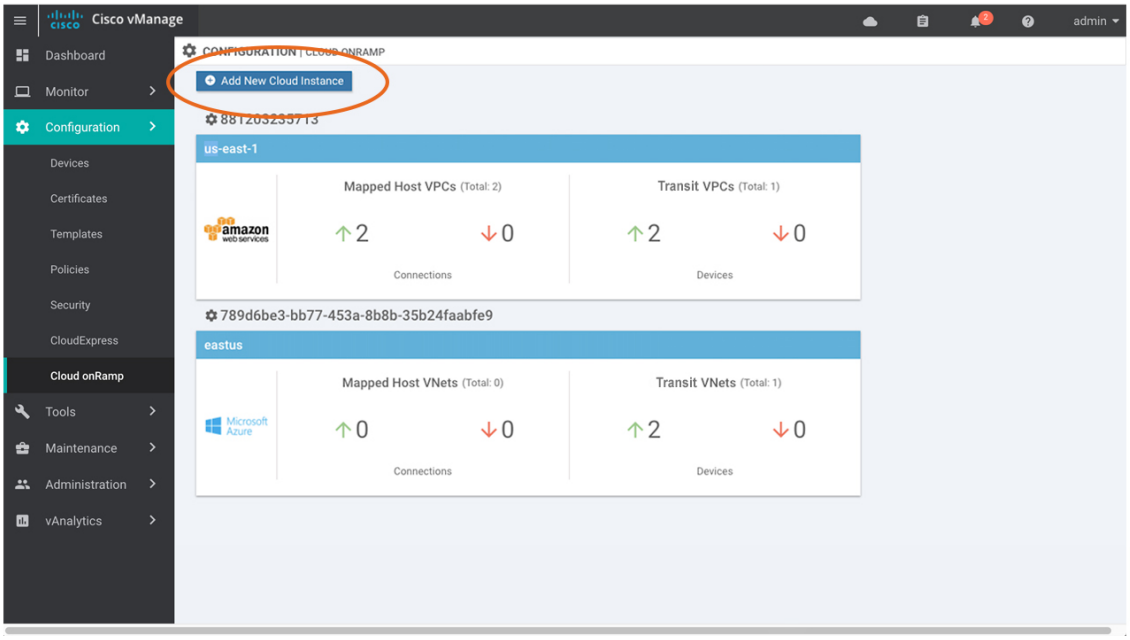
In the Cloud OnRamp configuration process, you map one or more host VPCs or host VNets to a single transit VPC or transit VNet. In doing this, you are configuring the cloud-based applications that branch users are able to access.

The mapping process establishes IPsec and BGP connections between the transit VPC or transit VNet and each host VPC or host VNet. The IPsec tunnel that connects the transit and host VPC or VNet runs IKE to provide security for the connection. For AWS, the IPsec tunnel runs IKE Version 1. For Azure, the IPsec tunnel runs IKE version 2. The BGP connection that is established over the secure IPsec tunnel allows the transit and host VPC or VNet to exchange routes so that the transit VPC or VNet can direct traffic from the branch to the proper host VPC or VNet, and hence to the proper cloud-based application.

During the mapping process, the IPsec tunnels and BGP peering sessions are configured and established automatically. After you establish the mappings, you can view the IPsec and BGP configurations, in the VPN Interface IPsec and BGP feature configuration templates, respectively, and you can modify them as necessary. You can configure Cloud OnRamp for IaaS for Azure by using the configuration wizard:

Create a Cloud Instance

1. In vManage NMS, select the **Configuration** > **Cloud onRamp for IaaS** screen.
2. Click **Add New Cloud Instance**:

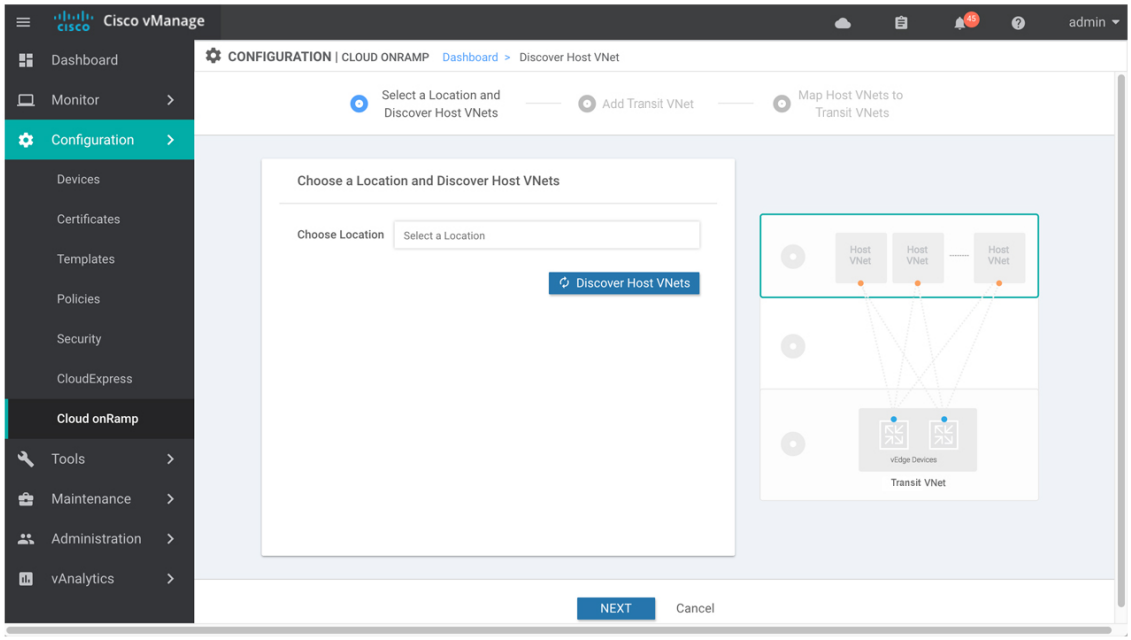


368405

3. In the Add Cloud Instance–Log In to a Cloud Server popup:
   a. In the **Cloud** drop-down, select **Azure** as the cloud type.
   b. To give vManage programmatic access to your Azure Subscription, log in to the cloud server:
      1. In the **Subscription ID** field, enter the ID of the Azure subscription you want to use as part of the Cloud OnRamp workflow.
      2. In the **Client ID** field, enter the ID of an existing application or create a new application in Azure. To create a new application, go to your **Azure Active Directory** > **App Registrations** > **New Application Registration**.
      3. In the **Tenant ID** field, enter the ID of your Azure account. To find the tenant ID, go to your Azure Active Directory and click **Properties**.
      4. In the **Secret Key** field, enter the password associated with the client ID.

4. Click **Log In**. The cloud instance configuration wizard opens.

   This wizard consists of three screens that you use to select a location and discover host VNets, add transit VNet, and map host VNets to transit VNets. A graphic on the right side of each wizard screen illustrates the steps in the cloud instance configuration process. Steps not yet completed are shown in light gray. The current step is highlighted within a blue box. Completed steps are indicated with a green checkmark and are shown in light orange.



5. Select a location and discover host VNets:
   a. In the **Choose Location** drop-down, select a geographical location.
   b. Click **Save and Finish** to create a transit VNet or optionally click **Proceed to Discovery and Mapping** to continue with the wizard.

6. Add a transit VNet:
   a. In the **Transit VNet Name** field, type a name for the transit VNet.

      The name can be up to 32 characters and can contain only uppercase and lowercase letters, the digits 0 through 9, hyphens (–), and underscores (_). It cannot contain spaces or any other characters.

   b. Under **Device Information**, enter information about the transit VNet:
      1. In the **WAN Edge Version** drop-down, select the software version to run on the VNet transit. The drop-down lists the published versions of the Viptela software in the Azure marketplace.
      2. In the **Size of Transit VNet** drop-down, select how much memory and how many CPUs to create on the VNet transit.
      3. In the **Device 1** drop-down, select the serial number to use.
      4. In the **Device 2** drop-down, select the serial number to use.
      5. To add additional device pairs, click .

         To remove a device pair, click .
      6. Click **Save and Finish** to complete the transit VNet configuration or optionally click **Proceed to Discovery and Mapping** to continue with the wizard.
      7. Click **Advanced** if you wish to enter more specific configuration options.

8. In the **Transit VNet CIDR** field, enter a custom CIDR that has a network mask in the range of 16 to 25. If you choose to leave this field empty, the Transit VPC is created with a default CIDR of 10.0.0.0/16.

c. Click **Save and Finish** to complete the transit VPC configuration, or optionally click **Proceed to Discovery and Mapping** to continue with the wizard.

7. Map the host VNets to transit VNets:

a. In the table of host VNets, select the desired host VNet.

b. Click **Map VNets**. The Map Host VNets popup opens.

c. In the **Transit VNet** drop-down, choose the transit VNet to map to the host VNets.

d. In the **VPN** drop-down, choose the VPN in the overlay network in which to place the mapping.

e. In the IPSec Tunnel CIDR section, enter two pairs of interface IP addresses for each Cisco vEdge device and a pair of loopback IP adresses to configure IPSec tunnels to reach the Azure virtual network transit. The IP addresses must be network addresses in the /30 subnet, be unique across the overlay network, and not be a part of the host VNet CIDR. If they are part of the host VNet CIDR, Azure will return an error while attempting to create VPN connections to the transit VNet.

As Azure supports single Virtual Private Gateway (VGW) configuration over IPSec tunnels with redundancy provided over a single tunnel, Cloud OnRamp for IaaS supports two VGWs for redundancy. During a planned maintenance or an unplanned event of a VGW, the IPSec tunnel from the VGW to the cloud routers get disconnected. This loss of connectivity causes the cloud routers lose BGP peering with vManage over IPSec tunnel. To enable BGP peering with cloud routers rather than IP address of the IPSec tunnel, provide the loopback addresses for each cloud router.



Note The loopback option for BGP peering supports single and multiple Virtual Gateway or Customer Gateway configuration or both on Azure cloud. This option applies only to the new host VNets mapped to transit VNets and not on the existing VNets.

f. In the Azure Information section:

1. In the **BGP ASN** field, enter the ASN that will be configured on the Azure Virtual Network Gateway that is spun up within the host VNet. Use an ASN that is not part of an existing configuration on Azure. For acceptable ASN values, refer to Azure documentation.

2. In the **Host VNet Gateway Subnet** field, enter a host VNet subnet in which the Virtual Network Gateway can reside. It is recommended you use a /28 subnet or higher. You must not provide a subnet that is already created in the VNet.

g. Click **Map VNets**.

h. Click **Save** and **Complete**.

In the VPN feature configuration template for VPN 0, when configuring the two cloud routers that form the transit VNet, ensure that the color you assign to the tunnel interface is a public color, not a private color. Public colors are **3g**, **biz-internet**, **blue**, **bronze**, **custom1**, **custom2**, **custom3**, **default**, **gold**, **green**, **lte**, **metro-ethernet**, **mpls**, **public-internet**, **red**, and **silver**.

Display Host VNets

1. In the Cloud OnRamp Dashboard, click the pane for the desired VNet. The Host VNets/Transit VNets screen opens, and Host VNets is selected by default. In the bar below this, Mapped Host VNets is selected by default, and the table on the screen lists the mapping between host and transit VNets, the state of the transit VNet, and the VPN ID.
2. To list unmapped host VNets, click **Unmapped Host VNets**.
3. To display the transit VNets, click **Transit** VNets.

Map Host VNets to an Existing Transit VNet

1. In the Cloud OnRamp Dashboard, click the pane for the desired location of the required account. The Host VNets/Transit VNets screen opens.
2. Click **Unmapped Host VNets**.
3. Click **Discover Host VNets**.
4. From the list of discovered host VNets, select the desired host VNet.
5. Click **Map VNets**. The Map Host VNets popup opens.
6. In the **Transit VNet** drop-down, select the desired transit VNet.
7. In the **VPN** drop-down, choose the VPN in the overlay network in which to place the mapping.
8. Click **Map VNets**.

Unmap Host VNets

1. In the Cloud OnRamp Dashboard, click the pane for the desired VNet. The Host VNets/Transit VNets screen opens.
2. Click **Mapped Host VNets**.
3. From the list of VNets, select the desired host VNets. It is recommended that you unmap one vNet at a time. If you want to unmap multiple vNets, do not select more than three in a single unmapping operation.
4. Click **Unmap VNets**.
5. Click **OK** to confirm the unmapping.

Display Transit VNets

1. In the Cloud OnRamp Dashboard, click the pane for the desired VNets. The Host VNets/Transit VNets screen opens, and Host VNets is selected by default.
2. Click **Transit VNets**.

The table at the bottom of the screen lists the transit VNets.

Add a Transit VNet

1. In the Cloud OnRamp Dashboard, click the pane for the desired VNet. The Host VNets/Transit VNets screen opens, and Host VNets is selected by default.

2. Click **Transit VNets**.
3. Click **Add Transit VNet**.

Delete a Transit VNet

1. In the Cloud OnRamp Dashboard, click the pane for the desired VNet. The Host VNets/Transit VNets screen opens, and Host VNets is selected by default.
2. Click **Mapped Host VNets**.
3. Select the desired host VNet, and click **Unmap VNets**.
4. Click **OK** to confirm the unmapping.
5. Click **Transit VNets**.
6. Click the trash icon to the left of the row for the transit VNet.
7. Click **OK** to confirm.

Troubleshoot Cloud OnRamp for IaaS

This section describes how to troubleshoot common problems with Cloud OnRamp for IaaS.

Two vEdge Routers are Not Available

Problem Statement

In vManage NMS, when you select the **Configuration** > **Cloud OnRamp** screen and click **Add New Cloud instance**, you see an error message indicating that two vEdge routers are not available.

Resolve the Problem

The vManage NMS does not have two vEdge Cloud routers that are running licensed Cisco SD-WAN software. Contact your operations team so that they can create the necessary vEdge Cloud routers.

If the vEdge routers are present and the error message persists, the two vEdge Cloud routers are not attached to configuration templates. Attach these templates in the vManage **Configuration** > **Templates** Device screen. Select the vEdge Cloud router, and then select **Attach Devices** from the More Actions icon to the right of the row.

Required Permissions for API

Problem Statement

When you enter your API keys, you get an error message indicating that this user does not have the required permissions.

Resolve the Problem

Ensure that the vManage server can reach the internet and has a DNS server configured so that it can reach AWS or Azure. To configure a DNS server, in the vManage VPN feature configuration template, enter the IP address of a DNS server, and then reattach the configuration template to the vManage server.

For AWS, check the API keys belonging to your AWS account. If you think you have the wrong keys, generate another pair of keys.

For AWS, if you are entering the correct keys and the error message persists, the keys do not have the required permissions. Check the user permissions associated with the key. Give the user the necessary permissions to create and edit VPCs and EC2 instances.

If the error message persists, check the time of the vManage server to ensure that it is set to the current time. If it is not, configure the vManage server's time to point to the Google NTP server. In the vManage NTP feature configuration template, enter a hostname of time.google.com, time2.google.com, time3.google.com, or time4.google.com. Then reattach the configuration template to the vManage server.

No vEdge Software Versions Appear in the Drop-Down

Problem Statement

When you are trying to configure transit VPC parameters for the transit VPC, no vEdge Cloud software versions are listed in the drop-down.

Resolve the Problem

Ensure that your customer account has subscribed to the Cisco SD-WAN vEdge Cloud routers.

Ensure that the vEdge cloud router is running software Release 17.2.0 or later.

No VPNs Appear in Drop-Down

Problem Statement

When you select the host VPCs or VNets to map, no VPNs are listed in the drop-down.

Resolve the Problem

This problem occurs when the device configuration template attached to the cloud router includes no service-side VPNs. Service-side VPNs (VPNs other than VPN 0 and VPN 512) are required to configure the IPsec connection between the two cloud routers selected for the transit and host VPCs or VNets.

This problem can also occur if the two cloud routers selected for the transit VPC or VNet have no overlapping service-side VPNs. Because the two vEdge routers form and active–active pair, the same service-side VPNs must be configured on both of them.

To configure service-side VPNs, in the vManage VPN feature configuration template, configure at least one service-side VPN. Ensure that at least one of the service-side VPNs is the same on both routers. Then reattach the configuration template to the routers.

Cloud OnRamp Task Fails

Problem Statement

After you have completed mapping the host VPCs to the transit VPCs, or host VNets to transit VNets, the Cloud OnRamp tasks fails.

Resolve the Problem

Review the displayed task information that is displayed on the screen to determine why the task failed. If the errors are related to AWS or Azure resources, ensure that all required resources are in place.

Cloud OnRamp Task Succeeds, But Routers Are Down

Problem Statement

The Cloud OnRamp task was successful, but the cloud routers are still in the Down state.

Resolve the Problem

Check the configuration templates:

- Check that all portions of the cloud router configuration, including policies, are valid and correct. If the configuration are invalid, they are not applied to the router, so the router never comes up.
- Check that the configuration for the vBond orchestrator is correct. If the DNS name or IP address configured of the vBond orchestrator is wrong, the vEdge router is unable to reach it and hence is unable to join the overlay network.

After you have determined what the configuration issues are:

1. Delete the Cloud OnRamp components:
   a. Unmap the host VPNs and the transit VPCs or VNets.
   b. Delete the transit vEdge routers.
2. Edit the configuration templates and reattach them to the cloud routers.
3. Repeat the Cloud OnRamp configuration process.

Desired Routes Not Exchanged

Problem Statement

The Cloud OnRamp configuration workflow is successful, the Cloud vEdge routers are up and running, but the desired routes are not getting exchanged.

Resolve the Problem

In vManage NMS, check the BGP configuration on the transit cloud routers. During the mapping process when you configure Cloud OnRamp service, BGP is configured to advertise the network 0.0.0.0/0. Make sure that the service-side VPN contains an IP route that points to 0.0.0.0/0. If necessary, add a static route in the VPN feature configuration template, and then reattach the configuration to the two cloud routers that you selected for the transit VPC or VNet.

On AWS, go to the host VPC and check its route table. In the route table, click the option **Enable route propagation** to ensure that the VPC receives the routes.

End-to-End Ping Is Unsuccessful

Problem Statement

Routing is working properly, but an end-to-end ping is not working.

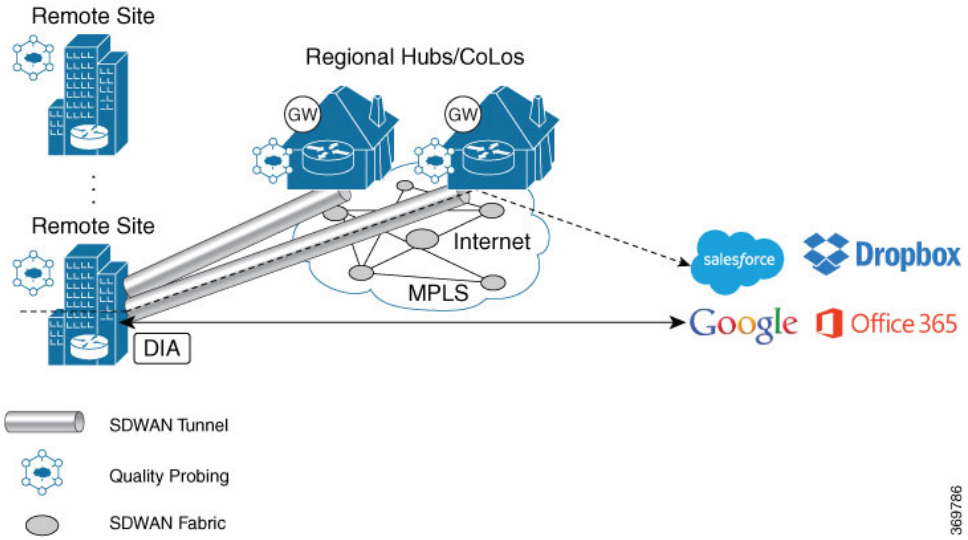
Resolve the Problem

On AWS, check the security group rules of the host VPC. The security group rules must allow the source IP address range subnets of the on-premises or branch-side devices, to allow traffic from the branch to reach AWS.

Cloud OnRamp for SaaS

Enterprise software providers deliver many applications as Software as a Service (SaaS) cloud applications, such as Dropbox, Microsoft Office365, and Salesforce. Latency and packet loss impact the performance of these applications, but in legacy networks, network administrators have little visibility into network characteristics between end users and SaaS applications. When a path is impaired in a legacy network, the manual process of shifting application traffic to an alternate path is complex, time consuming, and error prone.

Cloud OnRamp for SaaS (formerly called CloudExpress service) addresses these issues by optimizing performance for SaaS applications in the Cisco SD-WAN overlay network. From a central dashboard, Cloud OnRamp for SaaS provides clear visibility into the performance of individual cloud applications and automatically chooses the best path for each one. It responds to changes in network performance in real-time, intelligently re-routing cloud application traffic onto the best available path. The following image provdes a high level overview of OnRamp for SaaS.



Cloud OnRamp for SaaS calculates a value called the Viptela Quality of Experience (vQoE). The vQoE value weighs loss and latency using a formula customized for each application. For example, email applications tolerate latency better than video applications, and video applications tolerate loss better than email applications. The vQoE value ranges from zero to ten, with zero being the worst quality and ten being the best. Cloud OnRamp for SaaS computes vQoE values for applications and paths, then assigns applications to the paths that best match their vQoE value. Cloud OnRamp for SaaS periodically recalculates vQoE values for paths to ensure ongoing optimal application performance.

Cloud OnRamp for SaaS supports the following enterprise applications:

- Amazon Web Service (AWS)
- Box
- Concur
- Dropbox
- Google Apps
- GoToMeeting

- Intuit
- Microsoft Office 365
- Oracle
- Salesforce
- SugarCRM
- Zendesk
- Zoho CRM

Cisco XE SD-WAN devices use VRFs in place of VPNs. So, when you complete the configuration, the system automatically maps the VPN configurations to VRF configurations.

Enable Cloud OnRamp for SaaS

You can enable Cloud OnRamp for SaaS in your Cisco SD-WAN overlay network on sites with Direct Internet Access (DIA) and on DIA sites that access the internet through a secure web gateway such as Zscaler or iboss. You can also enable Cloud OnRamp for SaaS on client sites that access the internet through another site in the overlay network, called a gateway site. Gateway sites can include regional data centers or carrier-neutral facilities. When you enable Cloud OnRamp for SaaS on a client site that accesses the internet through a gateway, you also enable Cloud OnRamp for SaaS on the gateway site.

All Cisco SD-WAN devices configured for Cloud OnRamp for SaaS must meet the following requirements:

- The devices must run Cisco SD-WAN Software Release 16.3 or higher.
- The devices must run in vManage mode.
- You must configure a DNS server address in VPN 0.
- You must configure local exit interfaces in VPN 0:
  - If the local interface list contains only physical interfaces, you must enable NAT on those interfaces. You can use normal default IP routes for next hops.
  - If the local interface list contains only GRE interfaces, you do not need to enable NAT on those interfaces. You can add default routes to the IP address of the GRE tunnel to the destination.

Enable Cloud OnRamp for SaaS

1. In vManage NMS, click **Administration** > **Settings**.
2. Click the **Edit** button to the right of the **Cloud onRamp for SaaS** bar.
3. In the **Cloud onRamp for SaaS** field, click **Enabled**.
4. Click **Save**.

Configure Cloud OnRamp for SaaS

Add Applications

1. In vManage NMS, select the **Configuration** > **Cloud onRamp for SaaS** screen. The Cloud OnRamp for SaaS Dashboard screen appears.

   To edit the VPN configured for an application, click the Edit icon for that application, then enter the new VPN. You can enter any VPN other than 0, which is the transport VPN, or 512, which is the management VPN.

2. To add applications, from the **Manage Cloud OnRamp for SaaS** drop-down, located to the right of the title bar, select **Applications** to add applications to the cloud onRamp configuration.
3. Click the **Add Applications and VPN** button. The Add Applications & VPN pop-up window appears.
4. In the **Applications** field, select an application.
5. In the **VPN** field, enter the service VPN in which that application runs. You can enter any VPN other than 0 and 512.
6. Click **Add**.
7. Repeat Steps 3 through 6 for each application you want to add.
8. Click **Save Changes**.

Configure Client Sites

To configure Cloud OnRamp for SaaS on client sites that access the internet through gateways, you must configure Cloud OnRamp for SaaS both on the client sites and on the gateway sites.

Client sites in Cloud onRamp service choose the best gateway site for each application to use for accessing the internet.

1. In vManage NMS, select the **Configuration** > **Cloud onRamp for SaaS** screen. The Cloud OnRamp for SaaS Dashboard screen opens.
2. From the **Manage Cloud OnRamp for SaaS** drop-down, located to the right of the title bar, select **Client Sites**. The screen changes and displays the following elements:
   - Attach Sites—Add client sites to Cloud onRamp for SaaS service.
   - Detach Sites—Remove client sites from Cloud onRamp for SaaS service.
   - Client sites table—Display client sites configured for Cloud onRamp for SaaS service.
3. In the Manage Sites screen, click the **Attach Sites** button. The Attach Sites screen displays all sites in the overlay network with available sites highlighted. For a site to be available, all devices at that site must be running in vManage mode.
4. In the Available Sites pane, select a client site to attach and click the right arrow. To remove a site, select it in the Selected Sites pane and click the left arrow.
5. Click **Attach**. vManage NMS pushes the feature template configuration to the devices. The Task View window displays a Validation Success message.

6. Select **Configuration** > **Cloud onRamp for SaaS** to return to the Cloud OnRamp for SaaS Dashboard screen.
7. From the **Manage Cloud OnRamp for SaaS** drop-down, located to the right of the title bar, choose **Gateways**. The screen changes and displays the following elements:
   - Attach Gateways—Attach gateway sites.
   - Detach Sites—Remove gateway sites from Cloud onRamp service.
   - Edit Sites—Edit interfaces on gateway sites.
   - Gateways table—Display gateway sites configured for Cloud onRamp service.
8. In the Manage Gateways screen, click the **Attach Gateways** button. The Attach Gateways popup window displays all sites in your overlay network with available sites highlighted. For a site to be available, all devices at that site must be running in vManage mode.
9. In the Available Gateways pane, select a gateway site to attach and click the right arrow. To remove a site, select it in the Selected Sites pane and click the left arrow.
10. If you do not specify interfaces for Cloud OnRamp for SaaS to use, the system selects a NAT-enabled physical interface from VPN 0. To specify GRE interfaces for Cloud OnRamp for SaaS to use:
    a. Click the link **Add interfaces** to selected sites (optional), located in the bottom right corner of the window.
    b. In the **Select Interfaces** drop-down, select GRE interfaces to add.
    c. Click **Save Changes**.
11. Click **Attach**. vManage NMS pushes the feature template configuration to the devices. The Task View window displays a Validation Success message.
12. To return to the Cloud OnRamp for SaaS Dashboard, select **Configuration** > **Cloud onRamp for SaaS**.

To edit Cloud OnRamp for SaaS interfaces on gateway sites:

1. Select the sites you want to edit and click **Edit Gateways**.
2. In the **Edit Interfaces** of Selected Sites screen, select a site to edit.
   - To add interfaces, click the **Interfaces** field to select available interfaces.
   - To remove an interface, click the **X** beside its name.
3. Click **Save Changes** to push the new template to the vEdge routers.

Configure DIA Sites

1. In vManage NMS, select the **Configuration** > **Cloud onRamp for SaaS** screen. The Cloud OnRamp for SaaS Dashboard screen opens.

   In the title bar, choose **Manage Cloud OnRamp for SaaS** > **DIA**. The screen changes and displays the following elements:

   - Attach DIA Sites—Attach DIA sites.

- Detach DIA Sites—Remove DIA sites.
- Edit DIA Sites—Edit interfaces on DIA sites.
- Sites table—Display sites configured for Cloud onRamp service.

2. From the **Manage Cloud OnRamp for SaaS** drop-down, located to the right of the title bar, choose **Direct Internet Access (DIA) Sites**.
3. In the Manage DIA screen, click **Attach DIA Sites**. The Attach DIA Sites popup window displays all sites in your overlay network with available sites highlighted. For a site to be available, all devices at that site must be running in vManage mode.
4. In the Available Sites pane, select a site to attach and click the right arrow. To remove a site, select it in the Selected Sites pane and click the left arrow.
5. If you do not specify interfaces for Cloud OnRamp for SaaS to use, the system will select a NAT-enabled physical interface from VPN 0. If you would like to specify GRE interfaces for Cloud OnRamp for SaaS to use:
   a. Click the link, **Add interfaces** to selected sites (optional), located in the bottom right corner of the window.
   b. In the **Select Interfaces** drop-down, choose GRE interfaces to add.
   c. Click **Save Changes**.
6. Click **Attach**. vManage NMS pushes the feature template configuration to the devices. The Task View window displays a Validation Success message.
7. To return to the Cloud OnRamp for SaaS Dashboard, choose **Configuration** > **Cloud onRamp for SaaS**.

To edit Cloud onRamp interfaces on DIA sites:

1. Select the sites you want to edit and click Edit DIA Sites.
2. In the Edit Interfaces of Selected Sites screen, select a site to edit.
   - To add interfaces, click the **Interfaces** field to select available interfaces.
   - To remove an interface, click the **X** beside its name.
3. Click **Save Changes** to push the new template to the Cisco vEdge devices.

You have now completed configuring the Cloud OnRamp for SaaS. To return to the Cloud OnRamp for SaaS Dashboard, choose the **Configuration** > **Cloud onRamp for SaaS** screen.

Monitor Performance of Cloud OnRamp for SaaS

View Application Performance

In vManage NMS, select the **Configuration** > **Cloud OnRamp for SaaS** screen. The Cloud OnRamp for SaaS Dashboard displays the performance of each cloud application in a separate pane.

Each application pane displays the number of Cisco vEdge devices accessing the application and the quality of the connection:

- The bottom status bar displays green for devices experiencing good quality.
- The middle status bar displays yellow for devices experiencing average quality.
- The top status bar displays red for devices experiencing bad quality.

The number to the right of each status bar indicates how many devices are experiencing that quality of connection.

View Application Details

1. In vManage NMS, choose the **Configuration** > **Cloud OnRamp for SaaS** screen. The Cloud OnRamp for SaaS Dashboard displays each cloud application in a separate pane.
2. Click in an application's pane. vManage NMS displays a list of sites accessing the application.
3. Click a graph icon in the vQoE Score column to display vQoE history for that site:
   - Click a predefined or custom time period for which to display data.
   - Hover over a point on the chart to display vQoE details for that point in time.

Cloud OnRamp for Colocation Solution Overview

Digitization is placing high demands on IT to increase their speed of services and products that are delivered to customers, partners, and employees, while maintaining a high level of security. The interconnectivity between users and applications is becoming complex digital business architecture. This means network must be fast and flexible to meet the expanding changes and demand. At the same time, users want to increase the speed and reduce complexity of deployment without compromising the security.

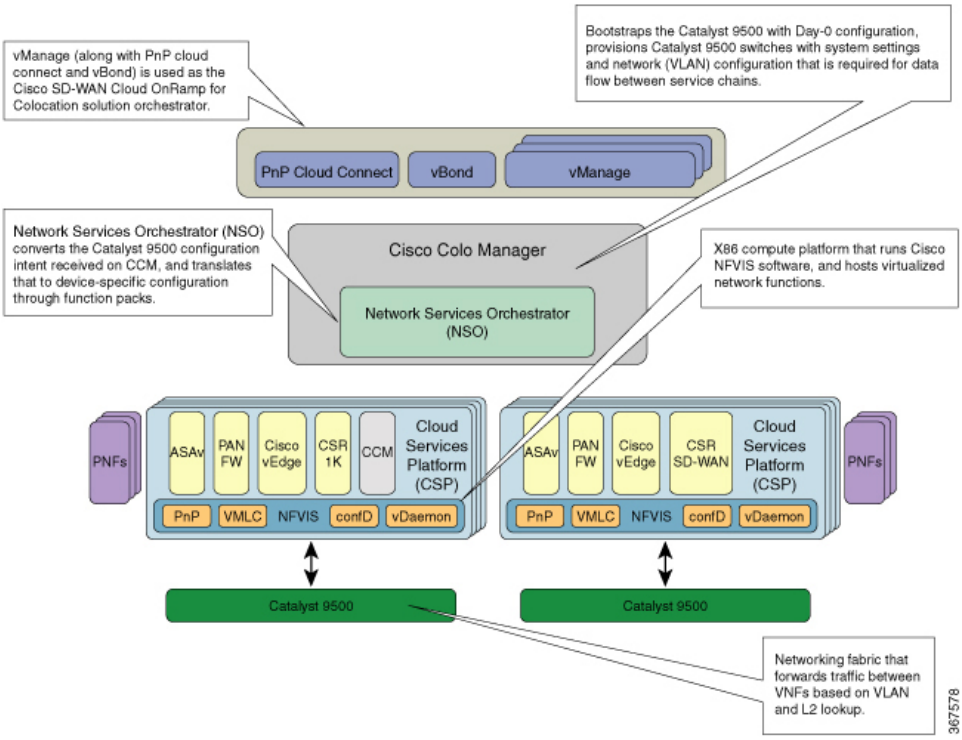
A Cloud OnRamp for Colocation is a campus, large branch, or a colocation, where the traffic gets aggregated. This solution is a flexible architecture that securely connects to enterprise applications that are hosted in the enterprise data center, public cloud, private or hybrid cloud to its endpoints such as, employees, devices, customers, or partners. This functionality is achieved by using Cloud Services Platform 5000 (CSP 5444) as the base Network Function Virtualization (NFV) platform that securely connects endpoints of an enterprise to applications. By deploying Cloud OnRamp for Colocation solution in colocation centers, customers can virtualize network services and other applications, and consolidate them into a single platform. The primary goal of the solution is to facilitate secure multicloud connectivity for Enterprise customers.

The Cloud OnRamp for Colocation solution offers the following benefits:

- Performance—Enterprises can optimize application performance by strategically placing the solution in colocation centers that are closest to the SaaS and public IaaS cloud providers.
- Agility—By virtualizing network services, enterprises can simplify their operations. Scaling up and down, and adding new services can now be done remotely. The Cisco Network Function Virtualization Infrastructure Software (NFVIS) on CSP 5444 negates the need to order, cable, rack, and stack dedicated hardware appliances when capacity must be increased or changes are required.
- Security—The centralization of communication patterns between employees, customers, partners, and applications allows for better and more consistent implementation of security policies.

- Cost savings—By having a central location to connect to various clouds (including private clouds), enterprises can optimize the cost of circuits to connect their users to applications. The circuit costs for a colocation facility are less than in a private data center.

Figure 1: Solution Architectural Overview



The Cloud OnRamp for Colocation solution can be deployed in multiple colocations. A colocation is a stack of compute and networking fabric that brings up multiple virtual networking functions and multiple service chains on them. This stack connects branch users, endpoints to a hybrid cloud or data center. vManage is used as the orchestrator to provision the devices in a colocation. Each colocation does not have visibility of other colocations in the same site or across sites.

Manage Clusters

Use the Cloud OnRamp for Colocation screen to configure a Cloud OnRamp for Colocation cluster and service groups that can be used with the cluster.

The three steps to configure Cloud OnRamp for Colocation devices are:

- Create a cluster. See Create and Activate Clusters, on page 30.
- Create a service group. See Create Service Chain in a Service Group, on page 39.
- Attach a cluster with a service group. See Attach and Detach Service Group with Cluster, on page 57.

A Cloud OnRamp for Colocation cluster is a collection of two to eight CSP devices and two switches. The supported cluster templates are:

- Small cluster—2 Catalyst 9500+2 CSP

- Medium Cluster—2 Catalyst 9500+4 CSP
- Large Cluster—2 Catalyst 9500+6 CSP
- X-Large Cluster—2 Catalyst 9500+8 CSP

Note Ensure that you add a minimum of two CSP devices one-by-one to a cluster. You can keep adding three, four, and so on, up to a maximum of eight CSP devices. You can edit a Day-N configuration of any cluster, and add pairs of CSP devices to each site up to a maximum of eight CSP devices.

Ensure that all devices that you bring into a cluster have the same software version.

Following are the cluster states:

- Incomplete—When a cluster is created from the vManage interface without providing the minimum requirement of two CSP devices and two switches. Also, cluster activation is not yet triggered.
- Inactive—When a cluster is created from the vManage interface after providing the minimum requirement of two CSP devices and two Switches, and cluster activation is not yet triggered.
- Init—When the cluster activation is triggered from the vManage interface and Day-0 configuration push to the end devices is pending.
- Inprogress—When one of the CSP devices within a cluster comes up with control connections, the cluster moves to this state.
- Pending—When the Day-0 configuration push is pending or VNF install is pending.
- Active—When a cluster is activated successfully and NCS has pushed the configuration to the end device.
- Failure—If Cisco Colo Manager has not been brought up or if any of the CSP devices that failed to receive an UP event.

A cluster transitioning to an active state or failure state is as follows:

- **Inactive** > **Init** > **Inprogress** > **Pending** > **Active**—Success
- **Inactive** > **Init** > **Inprogress** > **Pending** > **Failure**—Failure

Provision and Configure Cluster

This topic describes about activating a cluster that enable deployment of service chains.

To provision and configure a cluster, perform the following:

1. Create a cluster by adding two to eight CSP devices and two switches.

   CSP devices can be added to a cluster and configured through vManage before bringing them up. You can configure CSP devices and Catalyst 9K switches with the global features such as, AAA, default user (admin) password, NTP, syslog, and more.

2. Configure cluster parameters including IP address pool input such as, service chain VLAN pool, VNF management IP address pool, management gateway, VNF data plane IP pool, and system IP address pool.
3. Configure a service group.

A service group consists of one or more service chains.



Note You can add a service chain by selecting one of the predefined or validated service chain template, or create a custom one. For each service chain, configure input and output VLAN handoff and service chain throughput or bandwidth, as mentioned.

4. Configure each service chain by selecting each VNF from the service template. Choose a VNF image that is already uploaded to the VNF repository to bring up the VM along with required resources (CPU, memory, and disk). Provide the following information for each VNF in a service chain:
   - The specific VM instance behavior such as, HA, shared VM can be shared across service chains.
   - Day-0 configuration values for tokenized keys and not part of the VLAN pool, management IP address, or data HA IP address. The first and last VMs handoff-related information such as peering IP and autonomous system values must be provided. The internal parameters of a service chain are automatically filled by the orchestrator from the VLAN or Management or Data Plane IP address pool provided.
5. Add the required number of service chains for each service group and create the required number of service groups for a cluster.
6. To attach a cluster to a site or location, activate the cluster after all configuration has been completed.

   You can watch the cluster status change from in progress to active or error.

To edit a cluster, perform the following:

1. Modify the activated cluster by adding or deleting service groups or service chains.
2. Modify the global features configuration such as, AAA, system setting, and more.

You can predesign a service group and service chain before creating a cluster. They can be attached with a cluster after the cluster is active.

Create and Activate Clusters

This topic provide the steps about how a cluster can be formed with CSP devices, Catalyst 9500 switches as single unit, and provision the cluster with cluster-specific configuration.

Before you begin

Ensure that the clock on Cisco vManage and CSP devices are synchronized.

Procedure

Step 1 In vManage NMS, choose **Configuration** > **Cloud OnRamp for Colocation**. The CLOUD ONRAMP FOR COLOCATION screen appears, and the **Configure & Provision Cluster** button is highlighted. In the CLOUD ONRAMP FOR COLOCATION screen, perform the following tasks:

a) In the **Cluster** tab, click the **Configure & Provision Cluster** button.

A graphical representation of the default cluster, which consists of two switches each connected to two Cloud Services Platform (CSP) devices is displayed in the design view window.

b) Provide cluster name, description, site id, and location information.

Table 2: Cluster Information

Field	Description
Cluster Name	The cluster name can be up to 128 characters and can contain only alphanumeric characters.
Description	The description can be up to 2048 characters and can contain only alphanumeric characters.
Site ID	Specifies overlay network site identifier. This entry can be a value from 1 through 4294967295 (2^{32}– 1).
Location	The location can be up to 128 characters and can contain only alphanumeric characters.

c) From the graphical representation, to configure a switch, click a switch icon, the **Edit Switch** dialog box is displayed. Provide a name and choose the switch serial number. Click **Save**.

The switch name can be up to 128 characters and can contain only alphanumeric characters.

When you order Cisco SD-WAN Cloud OnRamp for Colocation solution PID on CCW and buy the Catalyst 9500 switches, a serial number is assigned for the switches. These serial numbers are integrated with vManage through PNP.

Note
You can keep the serial number field blank, design your cluster, and edit the cluster later to include the serial number after you have bought the switches.

d) To configure another switch, repeat the previous step.

e) From the graphical representation, to configure CSP, click a CSP icon in the CSP box. The **Edit CSP** dialog box is displayed. Provide a hostname and choose the CSP serial number. Click **Save**.

The hostname can be up to 128 characters and can contain only alphanumeric characters.

Note
You can keep the serial number field blank, design your cluster, and edit the cluster later to include the serial number after you have bought CSP devices. However, you cannot activate a cluster, where the serial number of CSP devices are not being included.

Note
Ensure that you configure the OTP for the CSP devices to bring them up. See Bring Up Cloud Services Platform in Cisco SD-WAN Cloud OnRamp for Colocation Solution Guide.

f) To add remaining CSP devices, repeat step e.
After you design a cluster, an ellipsis that is enclosed in a yellow circle next to the device appears if a serial number has not been assigned for a device.

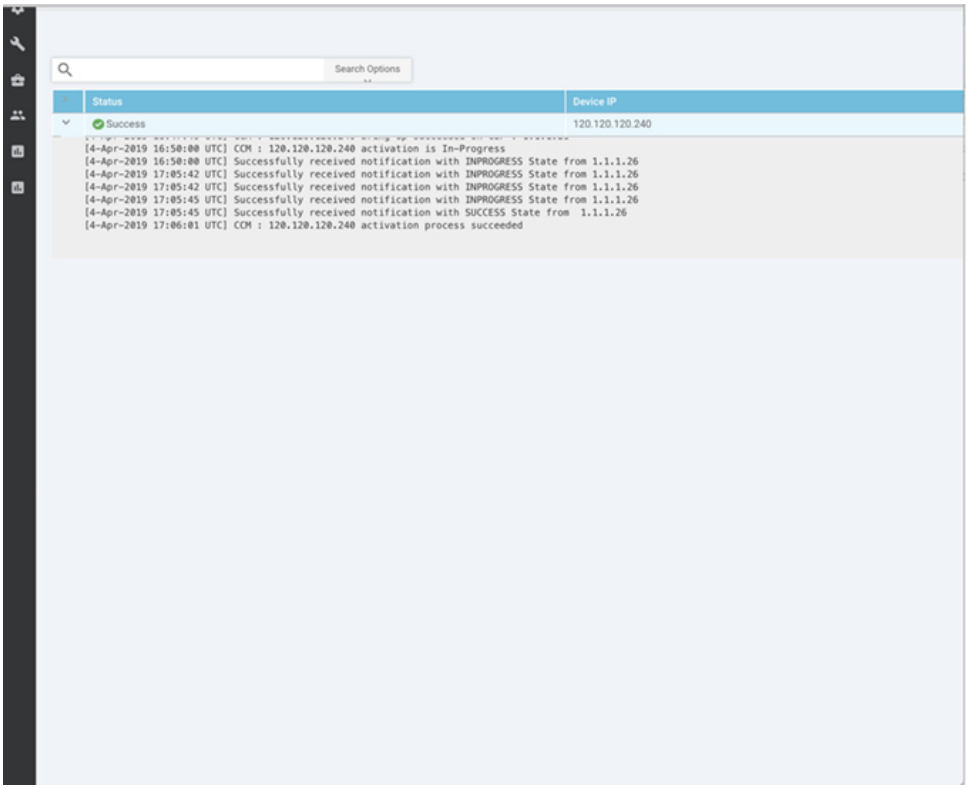
g) To edit a CSP device configuration, click a CSP from the graphical representation, and follow the process that is mentioned in substep e.

h) For mandatory and optional global parameters to be set for a cluster, click and choose from **Cluster Settings** drop-down. The dialog boxes for each of the global parameters are displayed. Enter values for the cluster settings parameters and click **Save**. See Cluster Settings, on page 34.

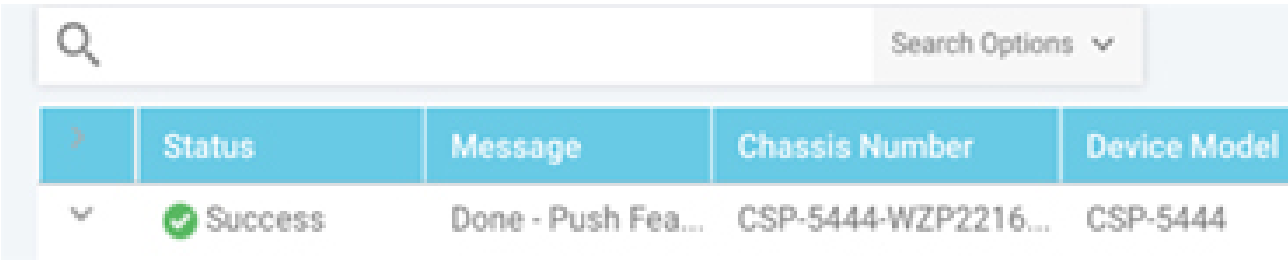
i) Click the **Save Cluster** button.

Step 2 In the **Cluster** tab, to activate a cluster, click a cluster, click the **More Actions** icon to the right of its row, click **Activate** against the cluster.

When you click Activate, vManage establishes a DTLS tunnel with CSP devices in the cluster where it connects with the switches through Cisco Colo Manager. After the DTLS connection is running, a CSP device in the cluster is chosen to host the Cisco Colo Manager. Cisco Colo Manager is brought up and vManage sends global parameter configurations to the CSP devices and switches. To verify if a cluster has been activated, you can view the task progress as shown.



To verify if cluster has been activated from the CSP end, you can view the task progress as shown.



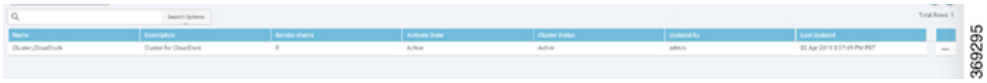
If the Cisco Colo Manager status does not go to "HEALTHY" after "STARTING", see the "Troubleshoot Cisco Colo Manager Issues" topic in the Cisco SD-WAN Cloud OnRamp for Colocation Solution Guide .

If the status of Cisco Colo Manager goes to "HEALTHY" after "STARTING" but the status of Cisco Colo Manager shows IN-PROGRESS for more than 20 minutes after the switch configurations are already complete, see the "Switch devices are not calling home to PNP or Cisco Colo Manager" topic in the Cisco SD-WAN Cloud OnRamp for Colocation Solution Guide.

If the status of the tasks running on a CSP device does not show success for more than five minutes after the activation through OTP, see the "Troubleshoot Cloud Services Platform Issues" topic in the Cisco SD-WAN Cloud OnRamp for Colocation Solution Guide.

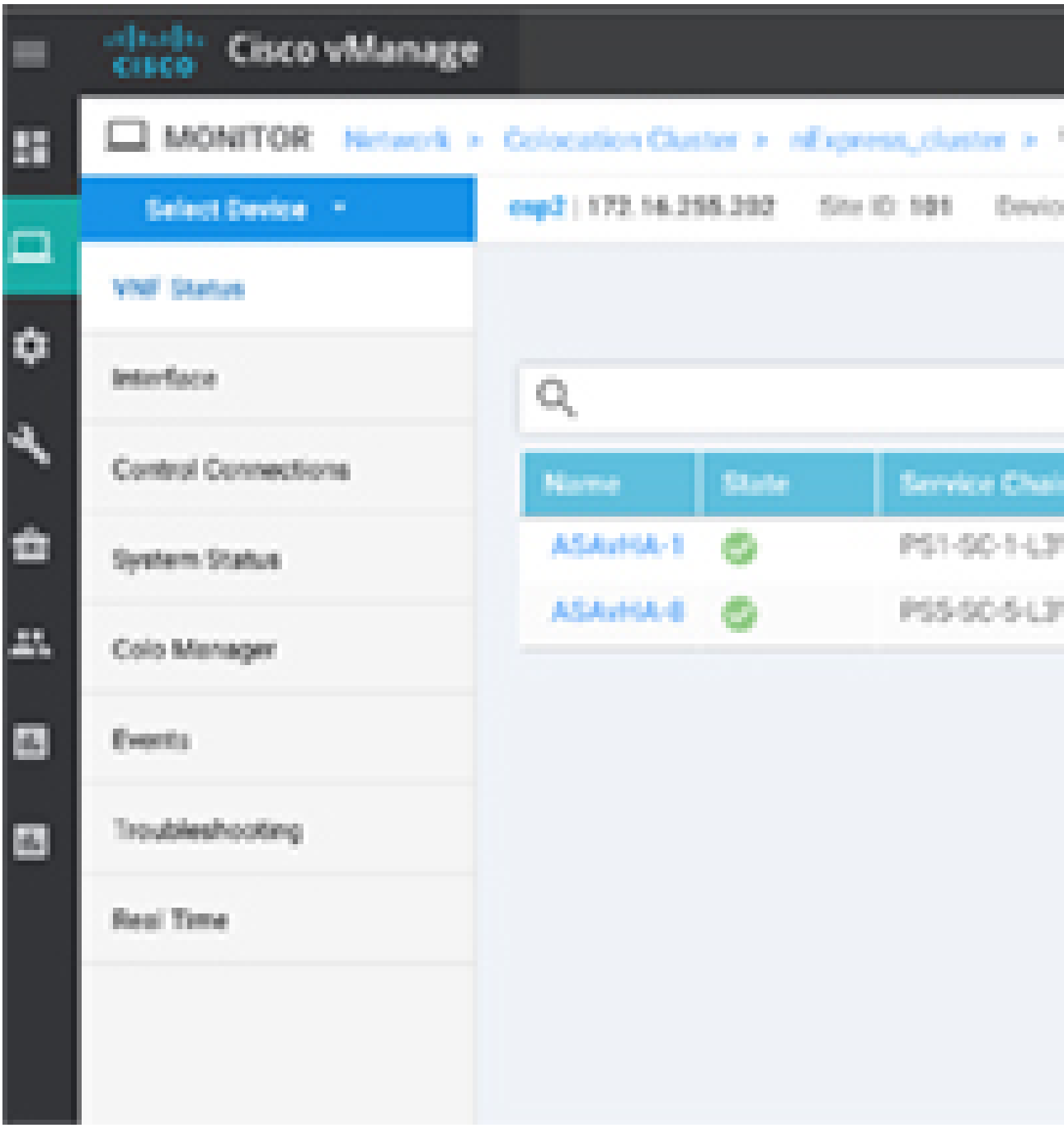
Note If a cluster goes into a "PENDING" state, click the **More Actions** icon to the right of its row, and then click the **Sync** button. This action moves a cluster back to an "ACTIVE" state.

To view if a cluster moves back to an "ACTIVE" state, you can view the successful activation as shown.



To determine the service groups present on CSP devices, navigate to **Monitor** > **Network** > **Colocation Cluster**.

Choose a cluster and then choose a CSP device as shown in the following image. You can choose and view other CSP devices.



Cluster Settings

The cluster settings parameters are:

- Configure login credentials for the cluster:

  1. In the Cluster Settings drop-down, click **Credentials**. The Credentials dialog box is displayed. Enter the values for the following fields:

(Mandatory) Template Name: The template name can be up to 128 characters and can contain only alphanumeric characters.

(Optional) Description: The description can be up to 2048 characters and can contain only alphanumeric characters.

2. Click **New User**.

   Provide name, password, and role of a user.

- Configure the Resource pool for the cluster:

1. In the Cluster Settings drop-down, click **Resource Pool**. The Resource Pool dialog box is displayed. Enter the values for the following fields:

   (Mandatory) Name: Name of the IP address pool. The name can be up to 128 characters and can contain only alphanumeric characters.

   (Optional) Description: IP address pool description. The description can be up to 2048 characters and can contain only alphanumeric characters.

   (Mandatory) DTLS Tunnel IP: IP addresses to be used for the DTLS tunnel. To enter multiple IP addresses, separate them by commas. To enter a range, separate the IP addresses with a hyphen (for example, 172.16.0.180-172.16.255.190).

   (Mandatory) Service Chain VLAN Pool: Numbers of the VLAN to be used for service chains. To enter multiple numbers, separate them by commas. To enter a numeric range, separate the numbers with a hyphen (for example, 1021-2021).



Note A VLAN range brings up VNFs, so that each circuit has VLAN configured when it comes up. The VLAN pool can only start from 1021 as switch reserves the VLANs until 1021. We recommend you to enter VLAN pools between 1021-2021.

   (Mandatory) VNF Data Plane IP Pool: IP addresses to be used for auto configuring data plane on a VNF interface. To enter multiple IP addresses, separate them by commas. To enter a range, separate the IP addresses with a hyphen (for example, 10.0.0.1-10.0.0.100).

   (Mandatory) VNF Management IP Pool: IP addresses to be used for theVNF. To enter multiple IP addresses, separate them by commas. To enter a range, separate the IP addresses with a hyphen (for example, 192.168.30.99-192.168.30.150).



Note These addresses are IP addresses for secure interfaces.

   (Mandatory) Management Subnet Gateway: IP address of the gateway to the management network. It enables DNS to exit the cluster.

   (Mandatory) Management Mask: Mask value for the failover cluster. For example, /24 and not 255.255.255.0

   (Mandatory) Switch PNP Server IP: IP address of the switch device.

Note The IP address of the switch is automatically picked from the management pool, which is the first IP address. You can change it if a different IP is configured in the DHCP server for the switch.

- Optionally, configure NTP servers for the cluster:
  1. In the Cluster Settings drop-down, select NTP. The NTP configuration box is displayed. Enter the values for the following fields:

     Template Name: Name of the NTP template. The name can be up to 128 characters and can contain only alphanumeric characters.

     Description: The description can be up to 2048 characters and can contain only alphanumeric characters.

     Preferred server: IP address of the primary NTP server.

     Backup server: IP address of the secondary NTP server.

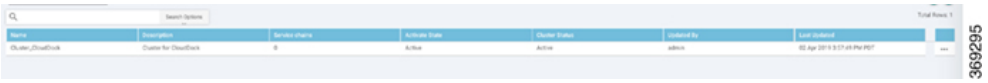
- Optionally, configure syslog parameters for the cluster:
  1. In the Cluster Settings drop-down, select Syslog. The System Log configuration box is displayed. Enter the values for the following fields:

     Template Name: Name of the System Log template. The name can be up to 128 characters and can contain only alphanumeric characters.

     Description: The description can be up to 2048 characters and can contain only alphanumeric characters.

     Severity drop-down: Select the severity of syslog messages to be logged.
  2. To configure a syslog server, click **New Server**.
  3. Type the IP address of a syslog server.

If all global parameters are set through cluster settings, you can verify if the cluster has been activated successfully, as shown.



View Cluster

To view a cluster configuration, perform the following steps:

Procedure

Step 1 In vManage, choose **Configuration** > **Cloud OnRamp for Colocation**. The CLOUD ONRAMP FOR COLOCATION Cluster screen appears, and the **Configure & Provision Cluster** button is highlighted.

Step 2 In the **Cluster** tab, click a cluster, click the **More Actions** icon to the right of its row, and click **View** against the cluster.

The Cluster window opens, displaying the switches and CSP devices in the cluster and showing which cluster settings have been configured.

Step 3 You can only view the global parameters being set, configuration of switches and CSP devices.

Step 4 Click the **Cancel** button to return to the CLOUD ONRAMP FOR COLOCATION Cluster screen.

Edit Cluster

To modify any existing cluster configuration such as global parameters, perform the following steps:

Procedure

Step 1 In vManage, select **Configuration** > **Cloud OnRamp for Colocation**. The CLOUD ONRAMP FOR COLOCATION Cluster screen appears, and the **Configure & Provision Cluster** button is highlighted.

Step 2 In the **Cluster** tab, click a cluster, click the **More Actions** icon to the right of its row, and click **Edit** against the cluster.

The Cluster window opens, displaying the switches and CSP devices in the cluster and showing which cluster settings have been configured.

Step 3 In the cluster design window, you can modify some of the global parameters. Based on whether a cluster is in active or inactive state, following are the restrictions for editing a cluster:

a. Inactive state.

- Edit all global parameters, and the Resource pool parameter.
- Add more CSP devices (up to eight).
- Cannot edit the name or serial number of a switch or CSP device. Instead, delete the CSP or switch and add another switch or CSP with a different name and serial number.
- Delete an entire cluster configuration.

b. Activate state.

- Edit all global parameters, except the Resource pool parameter.

  Note
  The Resource pool parameter cannot be changed when the cluster is activated. However, the only way to change the Resource pool parameter is to delete the cluster and recreate it again with the correct Resource pool parameter.

- Cannot edit the name or serial number of a switch or CSP device. Instead, delete the CSP or switch and add another switch or CSP with a different name and serial number.
- Cannot delete a cluster in active state.

Step 4 Click the **Save Cluster** button.

Remove Cluster

To decommission an entire cluster , perform the following steps:

Procedure

Step 1 In Cisco vManage, in the **Configuration** > **Certificates** screen, locate and verify status of devices to be deleted, and click **Invalid** against the devices.

Step 2 In the **Configuration|Ceritificates** screen, click **Send to Controllers**.

Step 3 In vManage, click **Configuration** > **Cloud OnRamp for Colocation**. The CLOUD ONRAMP FOR COLOCATION Cluster screen appears, and the **Configure & Provision Cluster** button is highlighted.

Step 4 In the **Cluster** tab, locate the cluster that has invalid devices, click the **More Actions** icon to the right of its row, and click **Deactivate** against the cluster.

If the cluster is attached to one or more service groups, you are prompted with a message that service chains hosting the VMs are running on this device and whether you can continue with the cluster deletion. However, although you confirm deletion of a cluster, you are not allowed to remove the cluster without detaching the service groups that are hosted on this device. If the cluster is not attached to any service group, you are prompted with a message to confirm the cluster deletion.

Note
You can delete the cluster, if necessary, or can keep it in deactivated state.

Step 5 To delete the cluster, select **Delete**.

Step 6 Click the **Cancel** button to return to the CLOUD ONRAMP FOR COLOCATION Cluster screen without deleting the cluster.

Step 7 To decommission invalid devices, in vManage, click **Configuration** > **Devices**.

Step 8 Locate the devices that are in the deactivated cluster, click the **More Actions** icon to the right of the device row, and click **Decommission WAN Edge**.

This action provides new tokens to your devices.

Step 9 Reset the devices to the factory default by using the command:

factory-default-reset all

Step 10 Log into NFVIS by using **admin** as the login name and **Admin123#** as the default password.

Step 11 Reset switch configuration and reboot switches. See the troubleshooting chapter in Cisco SD-WAN Cloud OnRamp for Colocation Solution Guide.

Reactivate Cluster

To add new CSP devices or when CSP devices are considered for RMA process, perform the following steps:

Procedure

Step 1 In Cisco vManage, in the **Configuration** > **Devices** screen, locate the devices that are in the deactivated cluster.

Step 2 Get new token from vManage for the devices.

Step 3 Log into NFVIS by using **admin** as the login name and **Admin123#** as the default password.

Step 4 Use the **request activate chassis-number** *chassis-serial-number* **token** *token-number* command.

Step 5 From vManage, configure the system configuration and then activate the cluster. See Create and Activate Clusters, on page 30.

If the cluster has been deleted, recreate and then activate it.

Step 6 In Cisco vManage, in the **Configuration** > **Certificates** screen, locate, and verify status of devices.

Step 7 To validate the devices, click **Valid** if it is invalid.

Step 8 In the **Configuration|Ceritificates** screen, click **Send to Controllers**.

Create Service Chain in a Service Group

A service group consists of one or more service chains.

Table 3: Feature History

Feature Name	Release Information	Feature Description
Monitor Service Chain Health	Cisco SD-WAN Release 19.2.1	This feature lets you configure periodic checks on the service chain data path and reports the overall status. To enable service chain health monitoring, NFVIS version 3.12.1 or later should be installed on all CSP devices in a cluster.

Procedure

In vManage, click **Configuration** > **Cloud OnRamp for Colocation**. The CLOUD ONRAMP FOR COLOCATION Cluster screen appears, and the **Configure & Provision Cluster** button is highlighted. In the CLOUD ONRAMP FOR COLCATION Cluster screen, perform the following tasks:

a) Click the **Service Group** tab, and then click the **Create Service Group** button. Provide the service group name and description.

The service group name can be up to 128 characters and can contain only alphanumeric characters.

The service group description can be up to 2048 characters and can contain only alphanumeric characters.

b) Click **Add Service Chain**.

c) In the Add Service Chain dialog box, provide the following information:

Table 4: Add Service Chain Information

Field	Description
Name	The service chain name can be up to 128 characters and can contain only alphanumeric characters.

Field	Description
Description	The service chain description can be up to 2048 characters and can contain only alphanumeric characters.
Bandwidth	The service chain bandwidth is in MBPS. The default bandwidth is 10 MB and you can configure a maximum bandwidth of 5G.
Input Handoff VLANS and Output Handoff VLANS	The Input VLAN handoff and output VLAN handoff can be comma separated values (10, 20) or a range from 10 through 20.
Monitoring	A toggle button that allows you to enable or disable service chain health monitoring. The service chain health monitoring is a periodic monitoring service that checks health of a service chain data path and reports the overall service chain health status. By default, the monitoring service is disabled. A service chain with subinterfaces such as, SCHM (Service Chain Health Monitoring Service) can only monitor the service chain including the first VLAN from subinterface VLAN list. The service chain monitoring reports status based on end-to-end connectivity. Hence, ensure that you take care of the routing and return traffic path, especially with SD-WAN chains for better results. **Note** • Ensure that you provide input and output monitoring IP addresses from input and output handoff subnets respectively. However, if the first and last VNF devices are VPN terminated, you do not need to provide an input and output monitoring IP addresses. For example, if the network function isn't VPN terminated, the input monitoring IP can be 192.0.2.1/24 from the inbound subnet, 192.0.2.0/24. The inbound subnet connects to the first network function and the output monitoring IP can be 203.0.113.11/24 that comes from outbound subnet, 203.0.113.0/24 of the last network function of a service chain. • If the first or last VNF firewall in a service chain is in transparent mode, those service chains can't be monitored.
Service Chain	Choose a topology from the service chain drop-down. For a service chain topology, you can choose any of the four validated service chains such as, Router - Firewall - Router, Firewall, Firewall - Router. See the topic "Validated Service Chains" in Cisco SD-WAN Cloud OnRamp Colocation Solution Guide.You can also create a customized service chain. See Create Custom Service Chain, on page 43.

d) In the Add Service Chain definition box, click **Add**.
Based on the service chain configuration information, a graphical representation of the service group with all the service chains and its VNFs automatically appear in the design view window. A VNF appears with a "V" or "P" around its circumference specifying that it is a virtual network function. It shows all the configured service chains within each service group. A check against the service chain indicates that all configuration information for the service chain has been completed.

After a cluster is activated, attached with the service group, and monitoring service is enabled for the service chain, when the CSP device is brought up where CCM is running, vManage chooses the same CSP device to start the monitoring service. The monitoring service monitors all service chains periodically in a round robin fashion by setting the monitoring interval to 30 minutes. See Monitor Cloud OnRamp Colocation Clusters , on page 61.

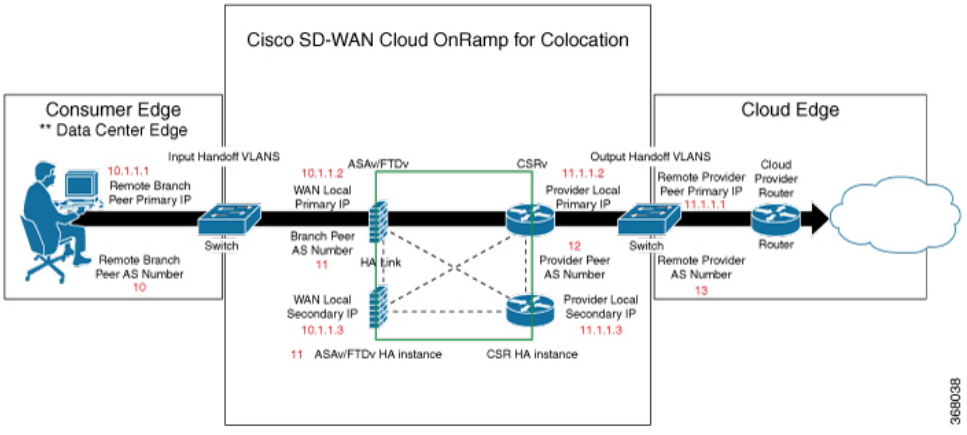
e) In the design view window, to configure a VNF, click a VNF in the service chain.
The Configure VNF dialog box appears.

f) Configure the VNF with the following information and perform the actions, as appropriate:

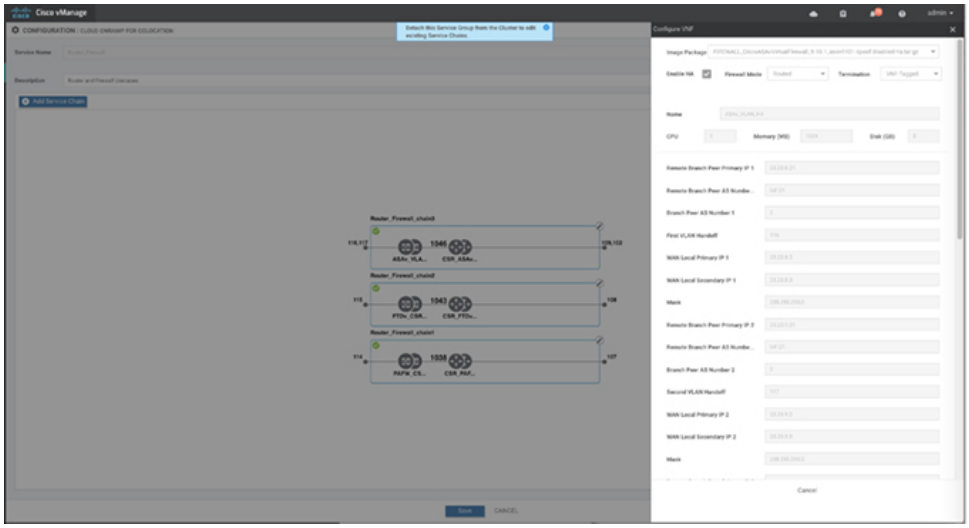
Table 5: VNF Properties of Router and Firewall

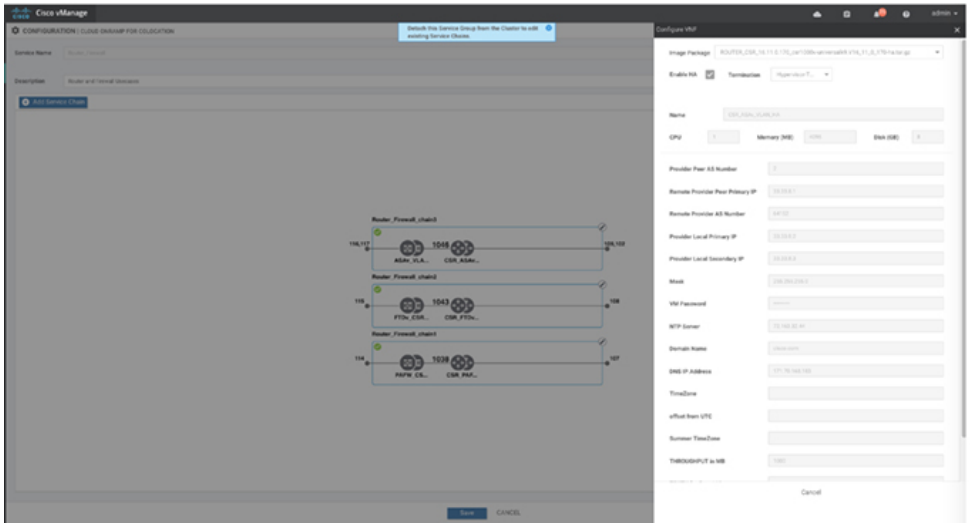
Field	Mandatory or Optional	Description
Image Package	Mandatory	Choose a router or firewall package.
Click **Fetch VNF Properties**. The available information for the image package is displayed in the Configure VNF dialog box.		
Name	Mandatory	VNF image name
CPU	Optional If you do not enter, the default value is considered, which is 1 vCpu.	Specifies the number of virtual CPUs that are required for a VNF.
Memory	Optional If you do not enter, the default value is considered, which is 1024 MB.	Specifies the maximum primary memory in MB that the VNF can use.
Disk	Optional If you do not enter, the default value is considered, which is 8 GB.	Specifies disk in GB required for the VM.
You are prompted with any custom tokenized variables from Day-0 that requires your input. Provide the values.		

In the following image, all IP addresses, VLAN, and autonomous system within the green box are system that is generated (from the VLAN, IP pools provided for the cluster) and automatically populated into Day-0 configurations of VMs.



The following images provide an example of the configuration for VNF IP addresses and autonomous system numbers in vManage.





For edge VMs such as first and last VM in a service chain, user must provide the following addresses as they peer with a branch and provider.

Table 6: VNF Options for First VM in Service Chain

Field	Mandatory or Optional	Description
Firewall Mode	Mandatory	Choose Routed or Transparent mode. **Note** Firewall mode is applicable only for firewall VMs and not other VMs.
Enable HA	Optional	HA enabled or not for VNF.

Field	Mandatory or Optional	Description
Termination mode	Mandatory	Specifies the following modes: • L3 mode selection with subinterfaces that are trunked. **<type>selection</type> <val help="L3 Mode With Sub-interfaces(Trunked)" display="VNF-Tagged">vlan</val>** • L3 mode with IPSEC termination from a consumer and routed to a provider gateway. **<val help="L3 Mode With IPSEC Termination From Consumer and Routed to Provider GW" display="Tunneled">vpn</val>** • L3 mode with access mode (nontrunked) **<val help="L3 Mode In Access Mode (Non-Trunked)" display="Hypervisor-Tagged">routed</val>**

g) Click **Configure**. The service chain is configured with the VNF configuration.
h) To add another service chain, repeat from step b.
i) Click **Save**.

The new service group is listed in a table on the **Service Group** tab. To view the status of the service chains that are monitored, use the task view page that displays a list of all running tasks along with the total number of successes and failures. On the CSP device where service chain health monotioring is enabled, to determine the service chain health status, use the **show system:system status** command.

Create Custom Service Chain

You can customize service chains,

- By including extra VNFs or add other VNF types.
- By creating new VNF sequence that is not part of the predefined service chains.

Procedure

Step 1 Create a service group and service chains within the service group. See Create Service Chain in a Service Group, on page 39.

Step 2 In the **Add Service Chain** dialog box, provide the service chain name, description, bandwidth, input VLAN handoff, output VLAN handoff, monitoring health information of a service chain, and service chain configuration. Click **Add**.

For the service chain configuration, choose **Create Custom** from the drop-down. An empty service chain in the design view window is available.

Step 3 To add a VNF such as a router, load balancer, firewall, and others, click a VNF icon on the left panel, and drag the icon to its proper location within the service group box. After adding all required VNFs and forming the VNF service chain,

configure each of the VNFs. Click a VNF in the service group box. The Configure VNF dialog box appears. Enter the following parameters:

a) Choose the software image to load from the **Image Package** drop-down.
b) Click **Fetch VNF Properties**.
c) Enter a name of the VNF in the **Name** field.
d) Enter the number of virtual CPUs required for the VNF in the **CPU** field.
e) Enter the amount of memory in megabytes to be allocated for the VNF in the **Memory** field.
f) Enter the amount of memory for storage in gigabytes to be allocated for the VNF in the **Disk** field.
g) Enter VNF-specific parameters, as required.

Note
These VNF details are the custom variables that are required for Day-0 operations of the VNF.

h) Click **Configure**.
i) To delete the VNF or cancel the VNF configuration, click **Delete** or **Cancel** respectively.

The customized service chains are added to a service group.



Note You can customize a VNF sequence with only up to four VNFs in a service chain.

Custom Service Chain with Shared PNF Devices

You can customize service chains by including supported PNF devices.



Caution Ensure that you do not share PNF devices across clusters. A PNF device can be shared across service chains, or across service groups. However, a PNF device can now be shared only across a single cluster.

Table 7: Feature History

Feature Name	Release Information	Feature Description
Manage PNF Devices in Service Chains	Cisco SD-WAN Release 19.2.1	This feature lets you add Physical Network Function (PNF) devices to a network, in addition to the Virtual Network function (VNF) devices. These PNF devices can be added to service chains and shared across service chains, service groups, and a cluster. Inclusion of PNF devices in the service chain can overcome the performance and scaling issues caused by using only VNF devices in a service chain.

Before you begin

For more information on validated physical network functions, see the "Validated Physical Network Functions" topic in the Cisco SD-WAN Cloud OnRamp for Colocation Solution Guide, Release 19.2 book.

To create a customized service chain by adding a router or firewall to an existing service chain, perform the following steps:

- If a PNF device needs to be managed by vManage, ensure that the serial number is already available in the vManage, which can then be available for selection during PNF configuration.
- The FTD device can be in any position in a service chain.
- An ASR 1000 Series Aggregation Services Routers can only be in the first and last position in a service chain.
- You can add PNF devices across service chains and service groups.
- You can share PNF devices across service groups. They can be shared across service groups by entering the same serial numbers.
- You can share PNF devices across a single cluster and cannot share across multiple clusters.

Procedure

Step 1 Create a service group and service chains within the service group. See Create Service Chain in a Service Group, on page 39.

Step 2 In the **Add Service Chain** dialog box, provide the service chain name, description, bandwidth, input VLAN handoff, output VLAN handoff, monitoring health information of a service chain, and service chain configuration. Click **Add**.

For the service chain configuration, choose **Create Custom** from the drop-down. An empty service chain in the design view window is available. In the left panel, the set of VNF devices and PNF devices that you can add into the service chain appears. The 'V' in the circumference of VNF devicess represents a VNF and 'P' in the circumference of PNF devices represent a PNF.

Note
Ensure that you choose the **Create Custom** option for creating service chains by sharing PNF devices.

Step 3 To add a PNF such as physical routers, physical firewalls in a service chain, click the required PNF icon on the left panel, and drag the icon to its proper location within the service chain box.

After adding all required PNF devices, configure each of them.

a) Click a PNF device in the service chain box.

The Configure PNF dialog box appears. To configure a PNF, enter the following parameters:

b) Check **HA Enabled** if HA is enabled for the PNF device.
c) If the PNF is HA enabled, ensure that you include the HA serial number in **HA Serial**.

If the PNF device is FTD, enter the following information.

1. Enter a name of the PNF in the **Name** field.
2. Choose Routed or Transparent mode as the **Firewall Mode**.
3. Enter the serial number of the PNF device in the **PNF Serial** field.

If the PNF device is ASR 1000 Series Aggregation Services Routers, enter the following information.

1. Check **vManaged** if the device is managed by vManage.

2. Click **Fetch Properties**.
3. Enter a name of the PNF in the **Name** field.
4. Enter the serial number of the PNF device in the **PNF Serial** field.

d) Click **Configure**.

Step 4 To add service chains and share PNF devices, repeat from step 2.

Step 5 Edit an existing PNF configuration by clicking it.

Step 6 In **Share NF To**, choose the service chains with which the PNF should be shared.

After a PNF is shared, if you hover on a PNF, the respective shared PNF devices are highlighted in blue color. However, the PNFs from different service groups are not highlighted in blue color. After you choose a NF to be shared, a blue color rim appears on it. If the same PNF is shared across multiple service chains, it can be used in different positions by dragging and placing the PNF icons in a specific positon.

Figure 2: Single PNF in a Service Chain

Here a service chain consists of a single PNF, Ftd_Pnf (not shared with other service chains).



Figure 3: Two PNF Devices in Service Chains

Here service chains consist of two PNFs, FTdv_PNF shared across SC1 and SC2 and ASR_PNF (non shred).

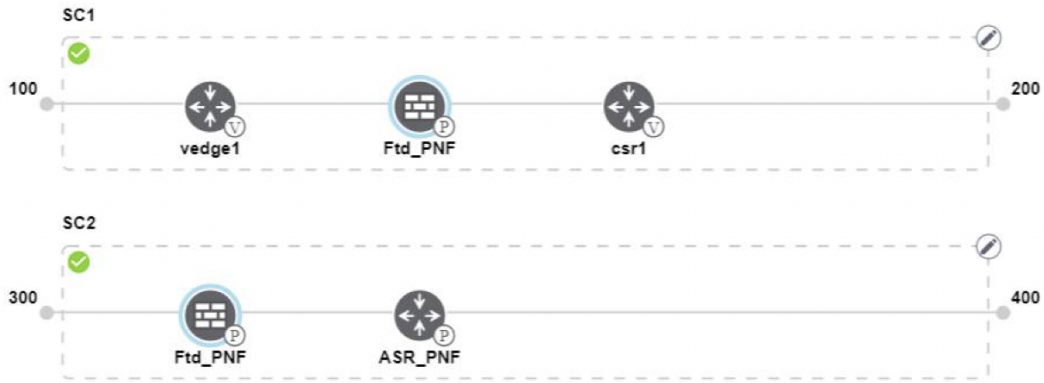
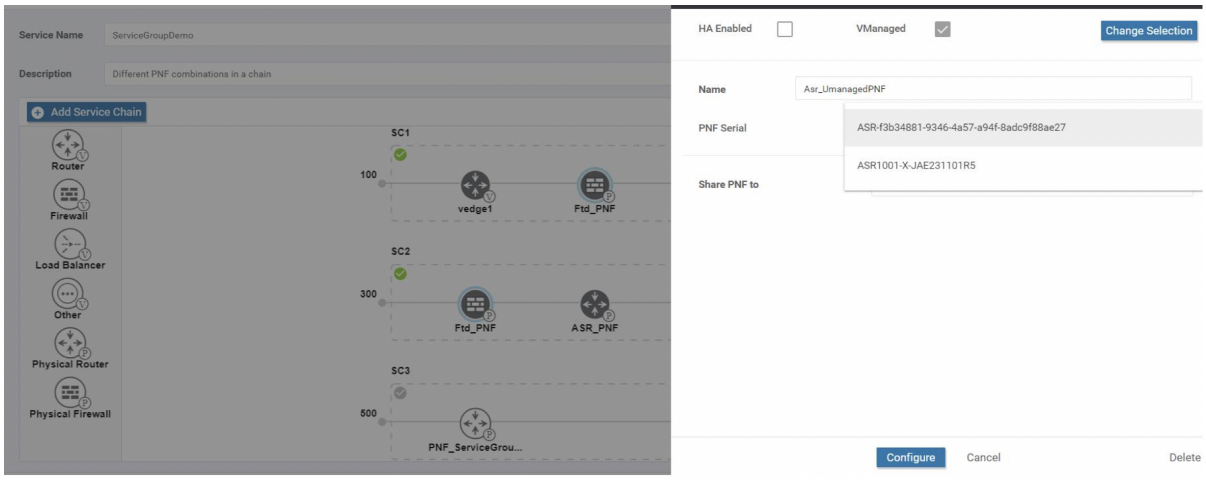


Figure 4: Three PNF Devices in Service Chains

Here service chains consist of three PNF devices in two different positions along with the vManage configuration.



Step 7 To delete a NF or cancel the NF configuration, click **Delete** or **Cancel** respectively.

You must attach the service groups to a cluster. After attaching service groups containing PNF devices with a cluster, the PNF configuration is not automatically pushed to the device unlike VNF devices. Instead, you must manually configure the PNF device by noting configuration that is generated on the Monitor Cloud OnRamp Colocation Clusters screen. The VLANs must be also configured on the Catalyst 9500 interfaces. See the ASR 1000 Series Aggregation Services Routers Configuration Guides and Cisco Firepower Threat Defense Configuration Guides for more information about the specific PNF configuration.

Configure PNF and Catalyst 9500

Procedure

Step 1 Identify ports from the switches where the PNF devices should be added, which are part of a service chain. To verify the availability of the ports, see "Service Chains and Port Connectivity Details" topic in Cisco SD-WAN Cloud OnRamp for Colocation Solution Guide.

Step 2 Connect with Catalyst 9500 by using either the terminal server of any of the Catalyst 9500 switches or use the **vty session** command with the IP address of the active switch.

Step 3 Configure VLANs from the generated configuration parameters on Catalyst 9500 with interfaces that are connected the PNF. See the Monitor Cloud OnRamp Colocation Clusters screen for the generated VLAN configuration.

Step 4 To configure FTD or ASR 1000 Series on a device, note the configuration from the Monitor screen and then manually configure it on the device.

Custom Service Chain with Shared VNF Devices

You can customize service chains by including supported VNF devices.

Table 8: Feature History

Feature Name	Release Information	Feature Description
Share VNF Devices Across Service Chains	Cisco SD-WAN Release 19.2.1	This feature lets you share Virtual Network Function (VNF) devices across service chains to improve resource utilisation and reduce resource fragmentation.

Before you begin

Ensure that you note the following points about sharing VNF devices:

- You can share only the first, last, or both first and last VNF devices in a service chain.
- You can share a VNF with a minimum of one more service chain and maximum up to five service chains.
- Each service chain can have a maximum of up to four VNF devices in a service chain.
- You can share VNF devices only in the same service group

Procedure

Step 1 Create a service group and service chains within the service group. See Create Service Chain in a Service Group, on page 39.

Step 2 In the **Add Service Chain** dialog box, provide the service chain name, description, bandwidth, input VLAN handoff, output VLAN handoff, monitoring health information of a service chain, and service chain configuration. Click **Add**.

For the service chain configuration, choose **Create Custom** from the drop-down. An empty service chain in the design view window is available. In the left panel, the set of VNF devices and PNF devices that you can add into the service chain appears. The 'V' in the circumference of VNF devices represents a VNF and 'P' in the circumference of PNF devices represent a PNF.

Note
Ensure that you choose the **Create Custom** option for creating a shared VNF package.

Step 3 To add a VNF such as a router, load balancer, firewall, and others, click a VNF icon on the left panel, and drag the icon to its proper location within the service chain box.

After adding all required VNF devices, configure each of them.

a) Click a VNF in the service chain box.

The Configure VNF dialog box appears. To configure VNF, enter the following parameters:

b) Choose the software image to load from the **Image Package** drop-down.

To create a customized VNF package from vManage, see Create Customized VNF Image, on page 66.

c) Click **Fetch VNF Properties**.

d) Enter a name of the VNF in the **Name** field.
e) Enter the number of virtual CPUs required for the VNF in the **CPU** field.
f) Enter the amount of memory in megabytes to be allocated for the VNF in the **Memory** field.
g) Enter the amount of memory for storage in gigabytes to be allocated for the VNF in the **Disk** field.
h) Enter VNF-specific parameters, as required. See Create Service Chain in a Service Group, on page 39 for more information about VNF-specific properties.

   These VNF-specific parameters are the custom user variables that are required for Day-0 operations of the VNF.

   For a complete information about the list of user and system variables for different VNF types such as vEdge, ASAv, CSR1000v when located at various positions, see .

   Note
   Ensure that you provide the values of the user variables if they are defined as mandatory, and for the system variables, vManage automatically sets the values for them.

i) Click **Configure**.

Step 4 To share VNF devices, repeat from step 2.

Step 5 Edit an existing VNF configuration by clicking it.

Step 6 Scroll down the VNF configuration slider to find the **Share NF To** field. Select the service chains from the **Share NF To** drop-down list with which the VNF should be shared.

After a VNF is shared, if you hover on a VNF, the respective shared VNF devices are highlighted in blue color. After you choose a NF to be shared, a blue rim appears on it.

Step 7 To delete a VNF or cancel the VNF configuration, click **Delete** or **Cancel** respectively.

You must attach service groups to a cluster.

Shared VNF Use Cases

The following images depict some of the shared VNF use cases and their predefined variable list:

Figure 5: Shared First vEdge VNF

The vEdge VNF in the first position is shared with the second service chain in the first position. The input to the first VNF is in access mode (hypervisor-tagged) and the neighbor (ASAv firewall) is in HA mode. To view and use the variable list that is associated for this scenario and various other scenarios, see the "vEdge Variable List" topic in the Cisco SD-WAN Cloud OnRamp for Colocation Solution Guide.



Figure 6: Shared First vEdge VNF

The vEdge VNF in the first position is shared with the second service chain in the first position. The input to the first VNF is in access mode (hypervisor-tagged) and the neighbor is in StandAlone mode. To view and use the variable list that is associated for this scenario and various other scenarios, see the "vEdge Variable List" topic in the Cisco SD-WAN Cloud OnRamp for Colocation Solution Guide.

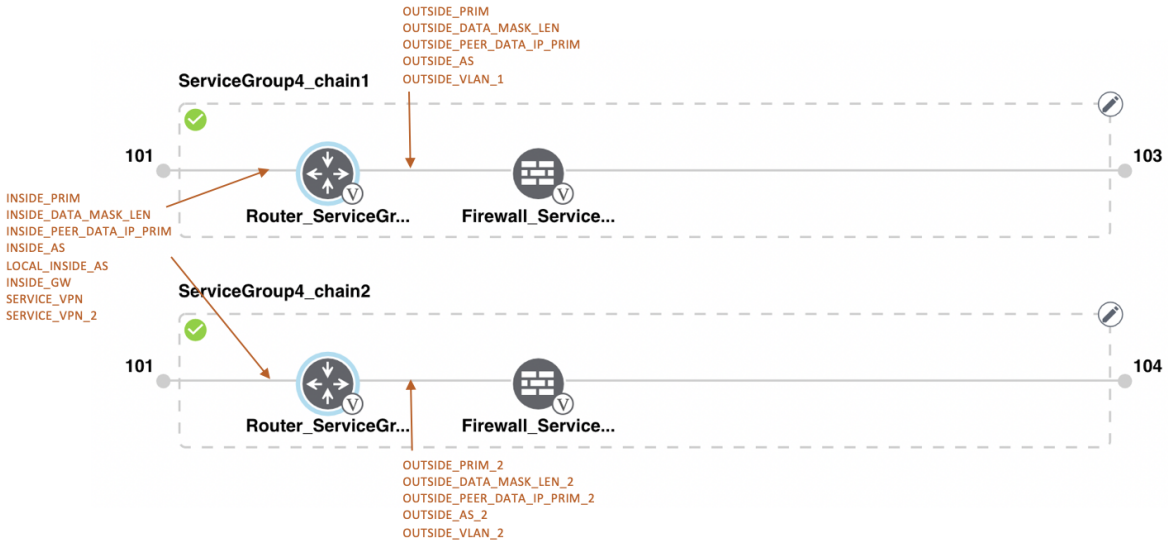


Figure 7: Shared First vEdge VNF

The vEdge VNF in the first position is shared with the second service chain in the first position. The input to the first VNF is in trunk mode (VNF-tagged) and the neighbor is in StandAlone mode. To view and use the variable list that is associated for this scenario and various other scenarios, see the "vEdge Variable List" topic in the Cisco SD-WAN Cloud OnRamp for Colocation Solution Guide.

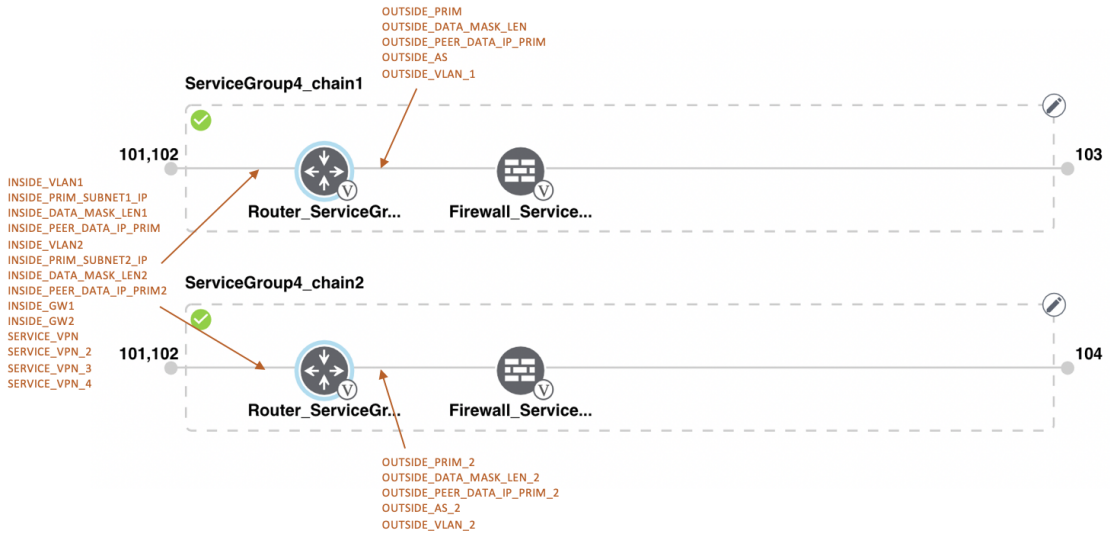


Figure 8: Shared First vEdge VNF

The vEdge VNF in the first position is shared with the second service chain in the first position. The input to the first VNF is in trunk mode (VNF-tagged) and the neighbor is in HA mode. To view and use the variable list that is associated for this scenario and various other scenarios, see the "vEdge Variable List" topic in the Cisco SD-WAN Cloud OnRamp for Colocation Solution Guide.

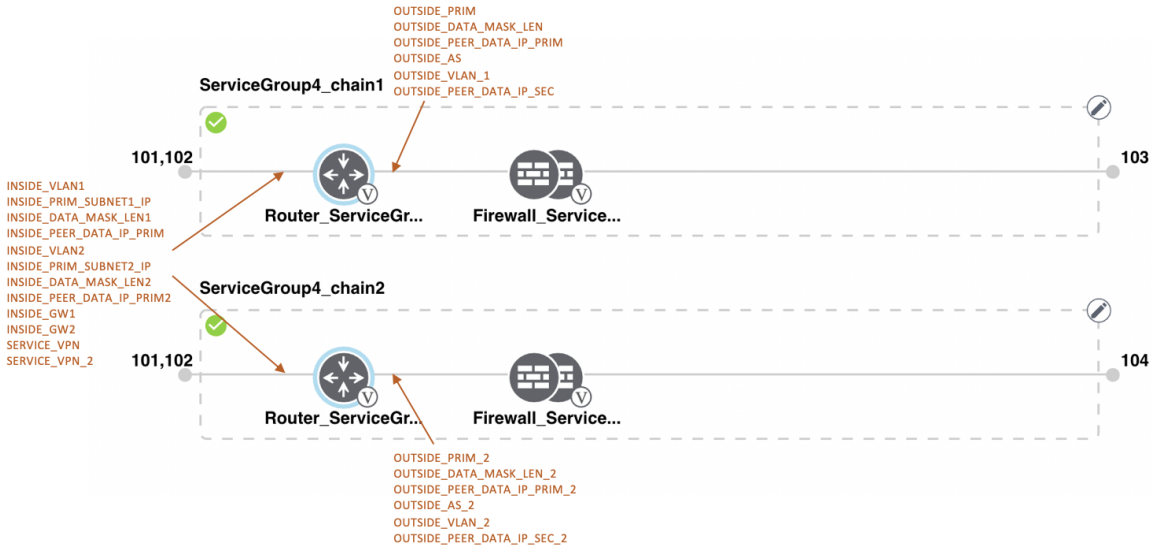


Figure 9: Shared Last CSR VNF

The CSR VNF in the last position is shared with the second service chain in the second position. The output from the last VNF is in access mode (hypervisor-tagged) and the neighbor (ASAv firewall) is in StandAlone mode. To view and use the variable list that is associated for this scenario and various other scenarios, see the "CSR Variable List" topic in the Cisco SD-WAN Cloud OnRamp for Colocation Solution Guide.

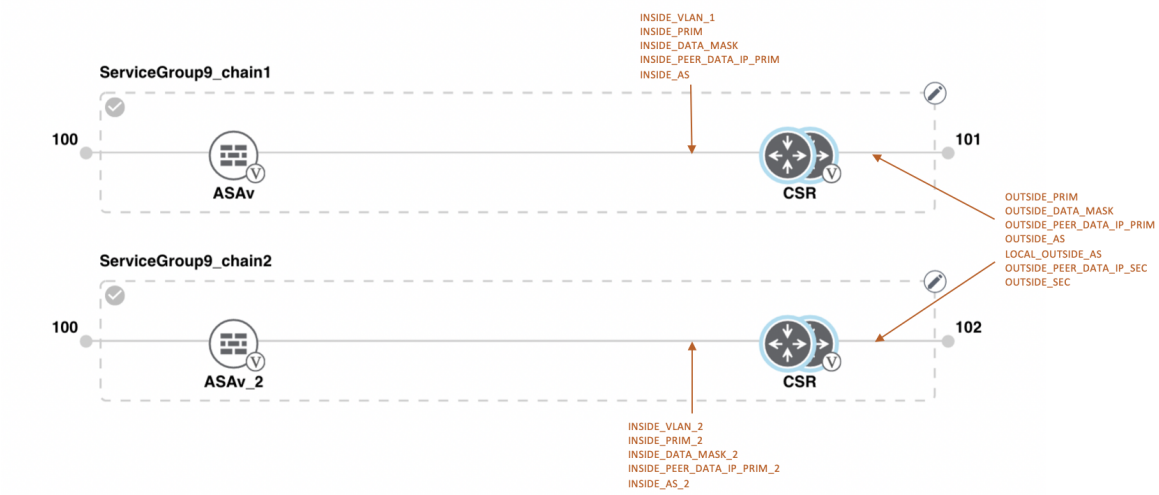


Figure 10: Shared Last CSR VNF

The CSR VNF in the last position is shared with the second service chain in the second position. The output from the last VNF is in access mode (hypervisor-tagged) and the neighbor is in StandAlone mode. To view and use the variable list that is associated for this scenario and various other scenarios, see the "CSR Variable List" topic in the Cisco SD-WAN Cloud OnRamp for Colocation Solution Guide.

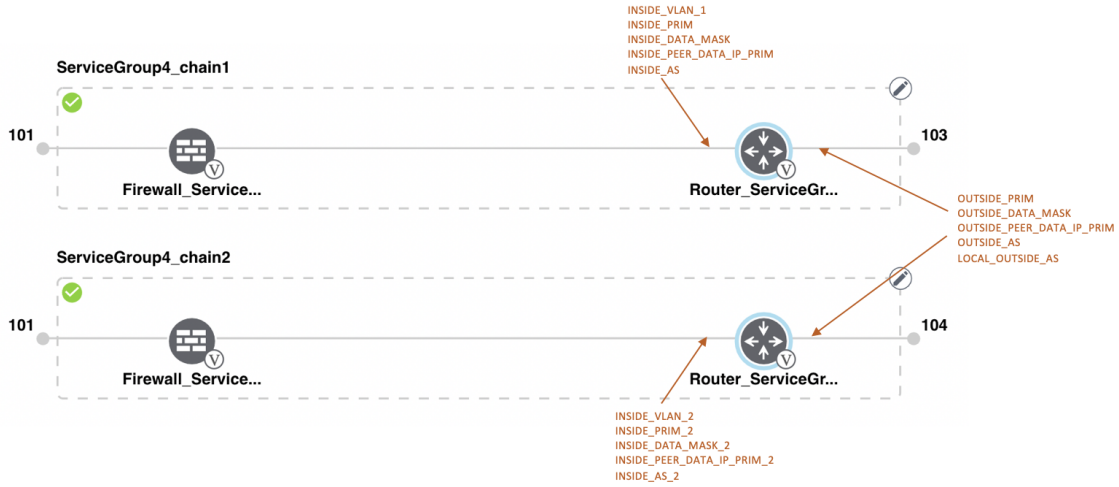


Figure 11: Shared Last CSR VNF

The CSR VNF in the last position is shared with the second service chain in the second position. The output from the last VNF is in access mode (hypervisor-tagged) and the neighbor (Firewall_Service) is in HA mode. To view and use the variable list that is associated for this scenario and various other scenarios, see the "CSR Variable List" topic in the Cisco SD-WAN Cloud OnRamp for Colocation Solution Guide.

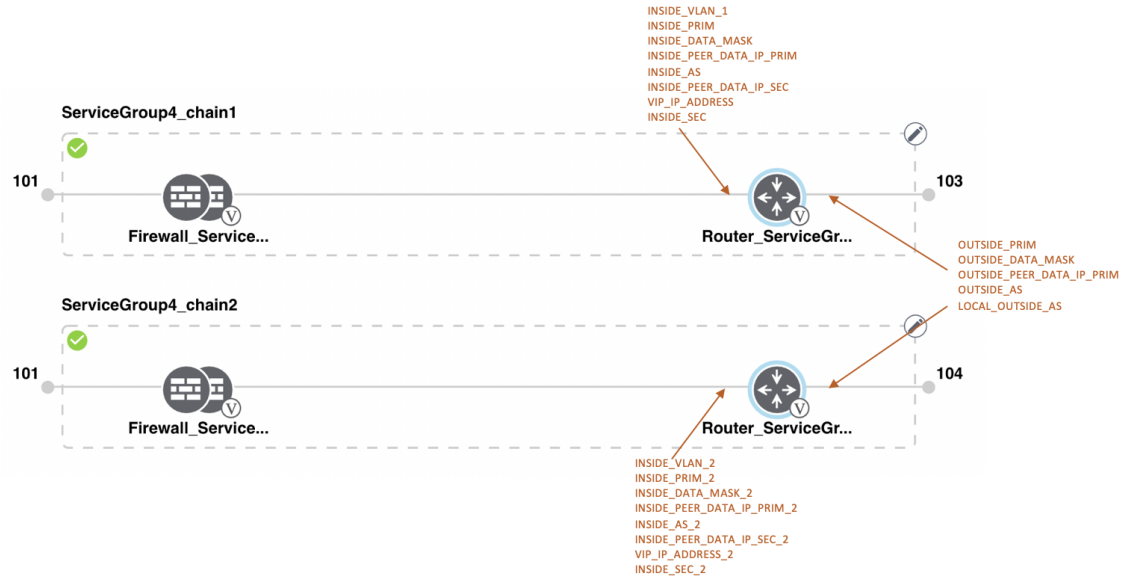


Figure 12: Shared Last CSR VNF

The CSR VNF in the last position is shared with the second service chain in the second position. The output from the last VNF is in access mode (hypervisor-tagged) and the neighbor (Firewall_Service) is in HA mode. To view and use the variable list that is associated for this scenario and various other scenarios, see the "CSR Variable List" topic in the Cisco SD-WAN Cloud OnRamp for Colocation Solution Guide .

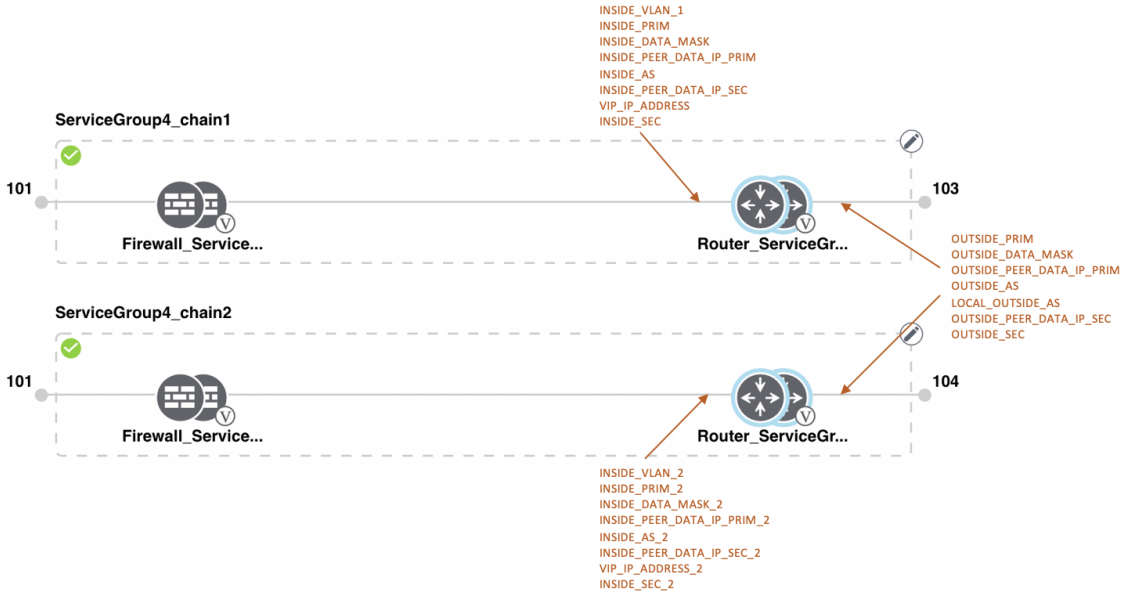


Figure 13: Shared First ASAv VNF

The ASAv VNF in the first position is shared with the second service chain in the first position. The input to the first VNF is in access mode (hypervisor-tagged) and the neighbor (CSR) is in redundant mode. To view

and use the variable list that is associated for this scenario and various other scenarios, see the "ASAv Variable List" topic in the Cisco SD-WAN Cloud OnRamp for Colocation Solution Guide.

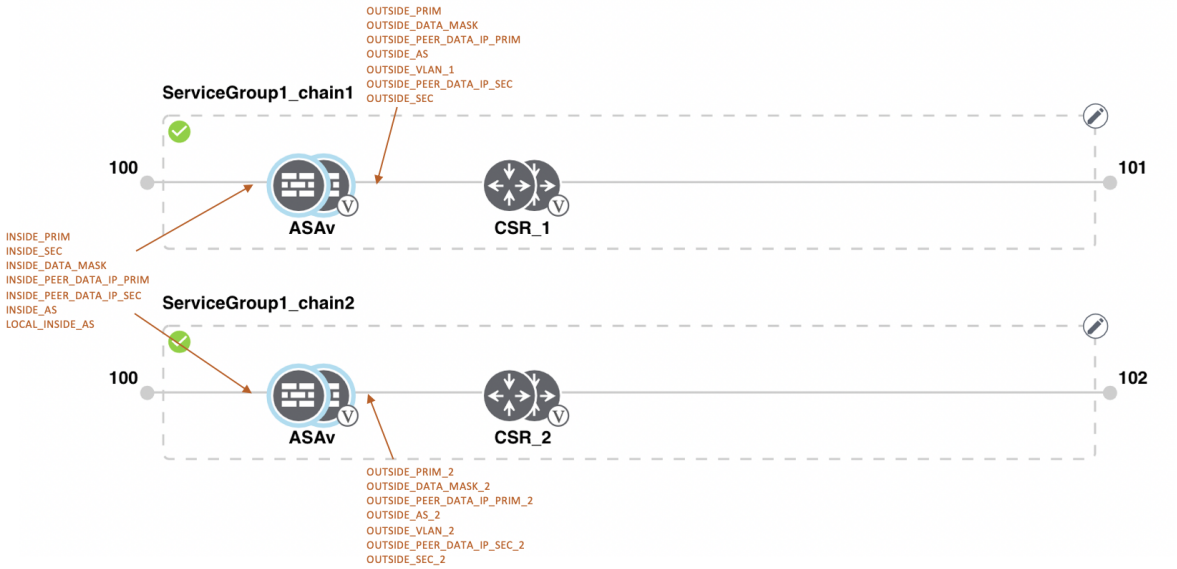


Figure 14: Shared First ASAv VNF

The ASAv (Firewall_Service) VNF in the first position is shared with the second service chain in the first position. The input to the first VNF is in access mode (hypervisor-tagged) and the neighbor is in StandAlone mode. To view and use the variable list that is associated for this scenario and various other scenarios, see the "ASAv Variable List" topic in the Cisco SD-WAN Cloud OnRamp for Colocation Solution Guide.

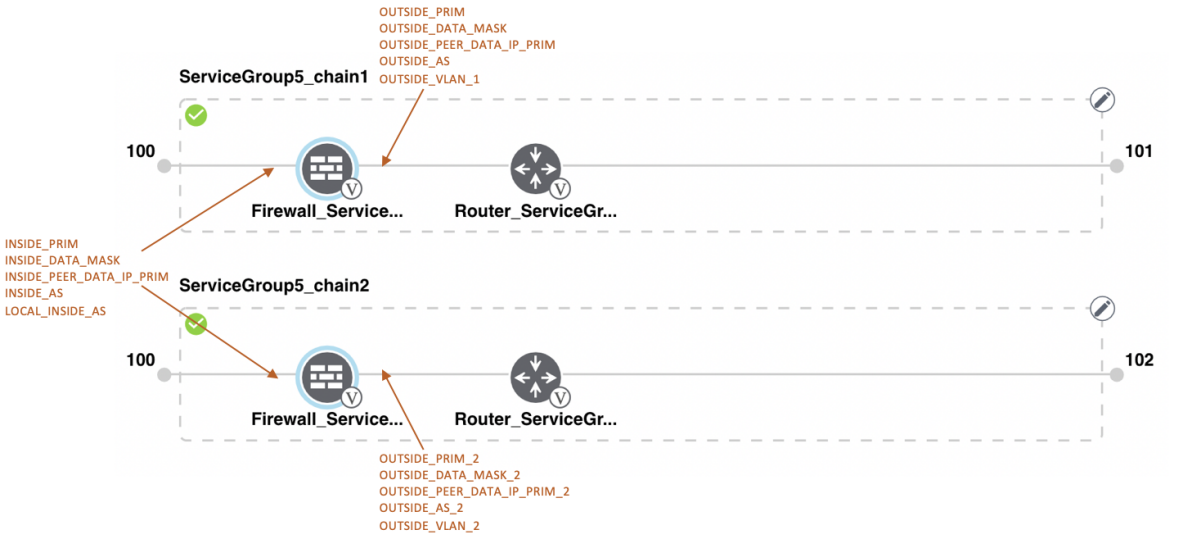


Figure 15: Shared First ASAv VNF

The ASAv (Firewall_Service) VNF in the first position is shared with the second service chain in the first position. The input to the first VNF is in access mode (hypervisor-tagged) and the neighbor, which is a router is in redundant mode. To view and use the variable list that is associated for this scenario and various other scenarios, see the "ASAv Variable List" topic in the Cisco SD-WAN Cloud OnRamp for Colocation Solution Guide.

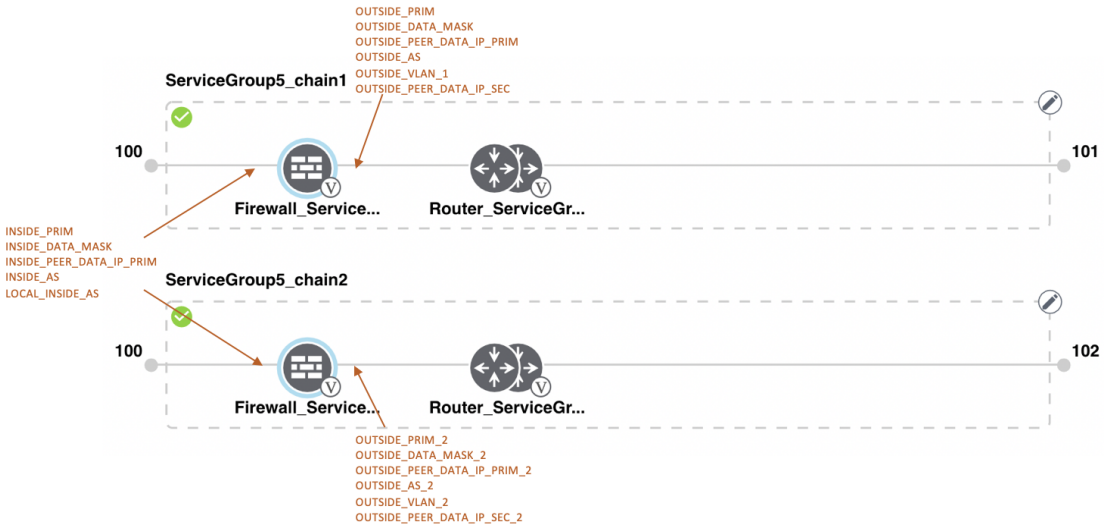
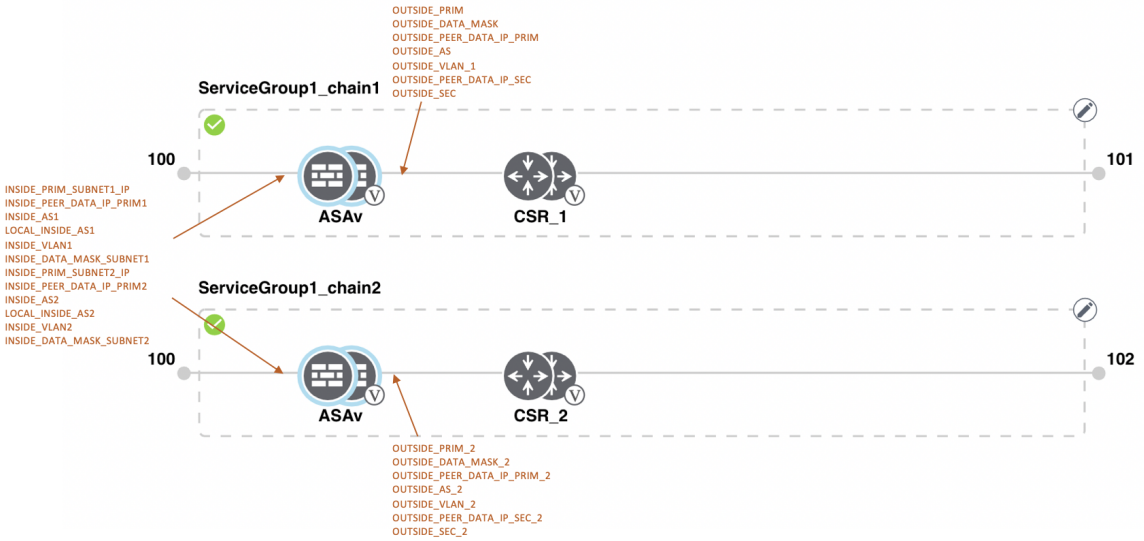


Figure 16: Shared First ASAv VNF

The ASAv VNF in the first position in HA mode is shared with the second service chain in the first position. The input to the first VNF is in trunk mode (vnf-tagged) and the neighbor (CSR) is in redundant mode. To view and use the variable list that is associated for this scenario and various other scenarios, see the "ASAv Variable List" topic in the Cisco SD-WAN Cloud OnRamp for Colocation Solution Guide.



View Service Groups

To view service groups, perform the following steps:

Procedure

In vManage, click **Configuration** > **Cloud OnRamp for Colocation**. The CLOUD ONRAMP FOR COLOCATION Cluster screen appears, and the **Configure & Provision Cluster** button is highlighted. In the CLOUD ONRAMP FOR COLOCATION Cluster screen, perform the following tasks:

a) Click the **Service Group** tab.

b) To view the service chains in the design view window, click a service chain box.

Edit Service Group

Before attaching a service group with a cluster, you can edit all parameters. After attaching a service group with a cluster, you can only edit monitoring configuration parameters. Also, after attaching a service group, you can only add new service chains but not edit or attach a service chain. Hence, ensure that you detach a service group from a cluster before editing an existing service chain. To edit and delete a service group, perform the following steps:

Procedure

In vManage, click **Configuration** > **Cloud OnRamp for Colocation**. The CLOUD ONRAMP FOR COLOCATION Cluster screen appears, and the **Configure & Provision Cluster** button is highlighted. In the CLOUD ONRAMP FOR COLOCATION Cluster screen, perform the following tasks:

a) Click the **Service Group** tab.
b) To modify either service chain configuration or modify a VNF configuration, click a router or firewall VNF icon.
c) To add new service chains, click a service chain button.

Attach and Detach Service Group with Cluster

To complete the Cloud OnRamp for Colocation configuration, you must attach service groups to a cluster. To attach or detach a service group from a cluster, perform the following steps:

Procedure

Step 1 In vManage, click **Configuration** > **Cloud OnRamp for Colocation**. The CLOUD ONRAMP FOR COLOCATION Cluster screen appears, and the **Configure & Provision Cluster** button is highlighted. To attach a service group with a cluster, perform the following steps:

a) In the **Cluster** tab, click a cluster from the table, click the **More Actions** icon to the right of its row, and click **Attach Service Groups**.

Step 2 In the **Attach Service Groups** dialog box, select a service group from the available service groups.

Step 3 Click the right arrow to move the chosen service groups to the selected box.

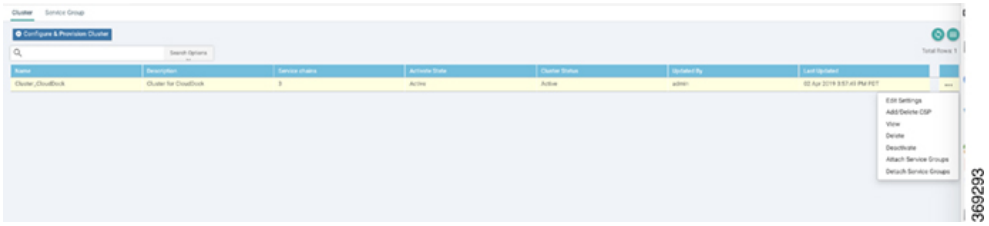
Step 4 Click **Attach**.

Step 5 To detach a service group from a cluster, perform the following action:

a) In the **Cluster** tab, click a cluster from the table, click the **More Actions** icon to the right of its row.
b) Click **Detach Service Groups**.

You cannot attach or detach an individual service chain within a group.

Step 6 To verify if service groups have been attached and detached, you can view from the following vManage screen:

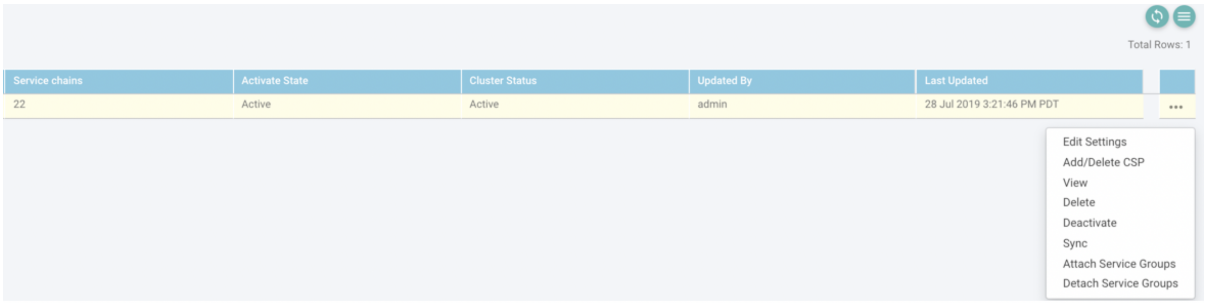


If the statuses of the tasks are "FAILURE" or in "PENDING" state for long duration, see the topic, "Troubleshoot Service Chain Issues" in the Cisco SD-WAN Cloud OnRamp for Colocation Solution Guide.

If a Cisco Colo Manager task fails, see the topic, "Troubleshoot Cisco Colo Manager Issues" in the Cisco SD-WAN Cloud OnRamp for Colocation Solution Guide.

Note If a cluster goes into a "PENDING" state, click the **More Actions** icon to the right of its row and then click the **Sync** button. This action moves the cluster back to an "ACTIVE" state. The sync button keeps the vManage synched with the devices and is visible when a cluster is active.

Figure 17: Sync Button for a Cluster



View Information About VNFs

You can view performance specifications and required resources for each VNF. Reviewing this information can help you to determine which VNF to use when you are designing a network service. To view information about VNFs, perform the following steps:

Procedure

Step 1 In Cisco vManage, click **Monitor** > **Network**.

The right pane displays VNF information in a tabular format. The table includes information such as CPU use, memory consumption, and disk, and other core parameters that define performance of a network service.

Step 2 Click a CSP device from the table.

Step 3 From the left pane, click **VNF Status**.

Step 4 From the table, click the VNF name. The right pane displays information about the specific VNF. You can click the network utilization, CPU utilization, memory utilization, and disk utilization to monitor the resources utilization of a VNF.

The primary part of the right pane contains the following VNF information:

Table 9: VNF Information

Chart options bar	VNF information in graphical format	VNF information in color coded format
• Chart Options drop-down—Click Chart Options drop-down list to select the type of data to display. • Time periods—Click either a predefined time period, or a custom time period for which to display data.	Choose a VNF from the **Select Device** drop-down list to display information for the VNF.	The VNFs are assigned a state based on the following operational status of VNF life cycle: • Inert (Green)—VNF is deployed but not alive. • Error (Red)—VNF deployment or any other operation fails, or VNF stops. • Active (Green)—VNF is healthy, deployed, and successfully booted up. An active state is also referred as alive. • Undefined (Yellow)—VNF is transitioning from one state to another.

The detail part of the right pane contains:

- Filter criteria
- VNF table that lists information about all VNFs or VMs. By default, the first six VNFs are checked. The network utilization charts for VNICs attached to SR-IOVand OVS are displayed.

  The graphical display plots information for the checked VNFs

  - Click the checkbox at the left to select and deselect VNFs. You can select and display information for a maximum of six VNFs at one time.
  - To change the sort order of a column, click the column title.

View Cisco Colo Manager Health from vManage

You can view Cisco Colo Manager health for a device, Cisco Colo Manager host system IP, Cisco Colo Manager IP, and Cisco Colo Manager state. Reviewing this information can help you to determine which VNF to use when you are designing a network service. To view information about VNFs, perform the following steps:

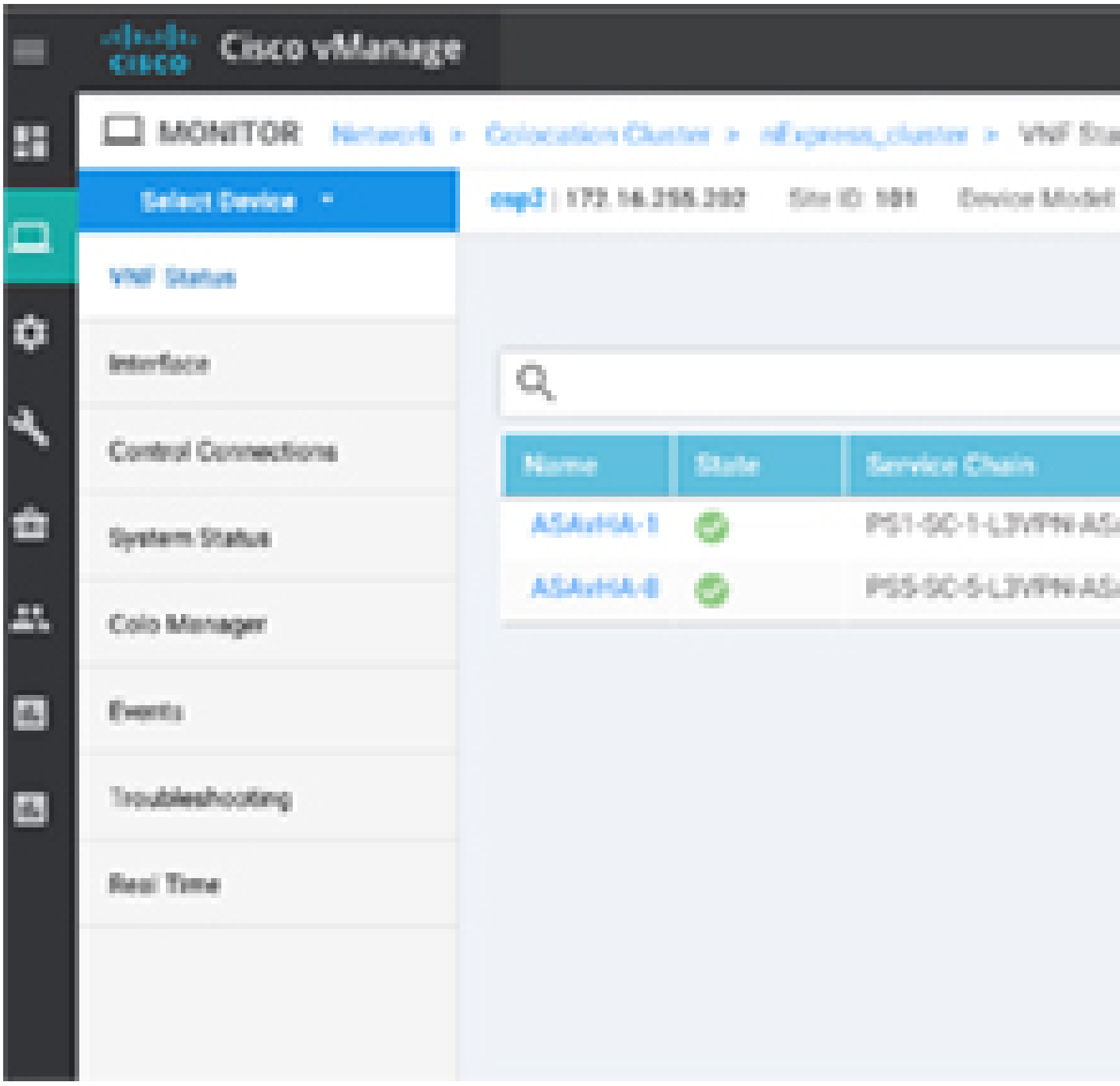
Procedure

Step 1 In vManage, click **Monitor** > **Network**.

The right pane displays VNF information in a tabular format. The table includes information such as CPU use, memory consumption, and disk, and other core parameters that define performance of a network service.

Step 2 Click a CSP device from the table.

Step 3 From the left pane, click **Colo Manager**.



The right pane displays information about the memory usage, CPU usage, uptime, and so on, of the colo manager.

Monitor Cloud OnRamp Colocation Clusters

You can view the cluster information and their health states. Reviewing this information can help you to determine which CSP device is responsible for hosting each VNF in a service chain. To view information about a cluster, perform the following steps:

Procedure

Step 1 In vManage, click **Monitor** > **Network**.

Step 2 To monitor clusters, click the **Colocation Clusters** tab.

All clusters with its relevant information are displayed in tabular format. Click a cluster name.

In the primary part of the left pane, you can view the PNF devices in a service group that are attached to a cluster along with the switches. In the right pane, you can view the cluster information such as the available and total CPU resources, available and allocated memory, and so on, based on Cloud OnRamp for Colocation size.

The detail part of the left pane contains:

- Filter criteria: Select the fields to be displayed from the search options drop-down.
- A table that lists information about all devices in the cluster (CSP devices, PNFs, and switches).

  Click a CSP cluster. VNF information is displayed in tabular format. The table includes information such as VNF name, service chains, CPU use, memory consumption, disk, management IP, and other core parameters that define performance of network service. See View Information About VNFs , on page 58.

Step 3 Click the **Services** tab.

In this tab, you can view:

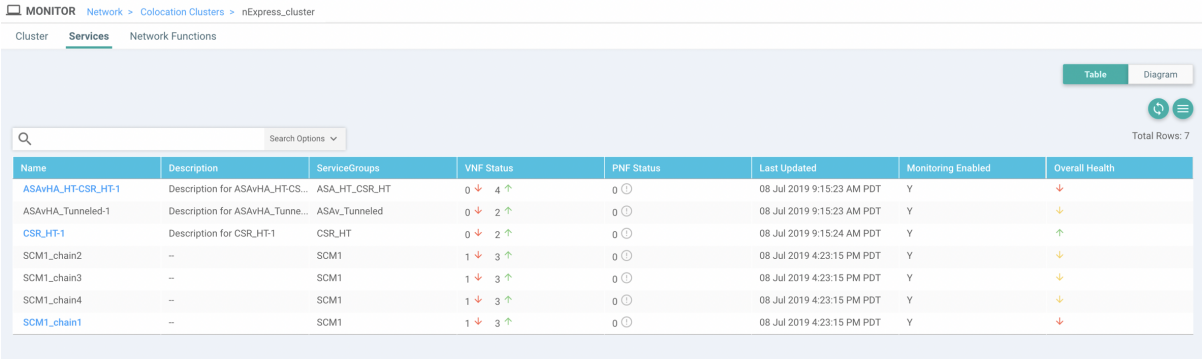
- The monitoring information of a service chain can be viewed in tabular format. The first two columns display the name and description of the service chain within the service group and the remaining columns mention about the VNF, PNF statuses, monitoring service enablemement, and the overall health of a service chain. The various health statuses and their representations are:
  - Healthy—Up arrow in green. A service chain is in 'Healthy' status when all the VNF, PNF devices are running and are in healthy state. Ensure that you configure the routing and policy correctly.
  - Unhealthy—Down arrow in red. If one of the VNFs or PNFs are in unhealthy state, the service chain is reported to be in 'Unhealthy' status. For example, after deploying a service chain, if one of the network function IP address changes on the WAN or LAN side, or the firewall policy is not configured to let the traffic pass through, then unnhealthy state is reported. This is because the network function or overall service chain is Unhealthy or both are in Unhealthy state.
  - Undetermined—Down arrow in yellow. This is a third state that is reported when the health of the service chain cannot be determined. This state is also reported when there is no status such as healthy or unhealthy available for the monitored service chain over a time period. You cannot query or search a service chain with undetermined status.

    If a service chain consists of a single PNF and PNF is orchestrated outside of vManage, then it cannot be monitored. If a service chain consists of a single network function, firewall that has VPN termination on both sides which cannot be monitored, then it is reported as Undetermined.

    Note

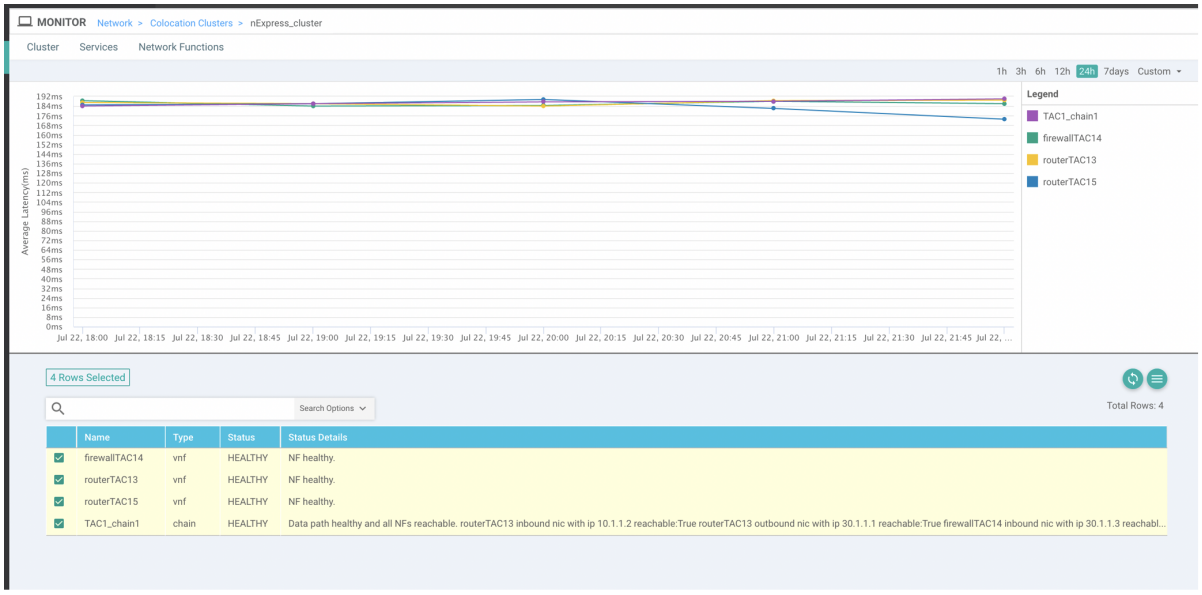
If the status of a service chain is undetermined, you cannot choose the service chain to view the detailed monitoring information.

Figure 18: Service Chain Health Monitoring Results



- Click a service group that is in Healthy or Unhealthy state. The primary part of the service chain monitoring in the right pane contains the following elements:

Figure 19: Service Chain Health Monitoring Status



Graphical display that plots the latency information of the service chain, VNFs, PNFs.

The detail part of the right pane contains:

- Filter criteria
- A table that lists information about all service chains, VNFs, PNFs, their health status, and types.
  - Check the checkbox at the left of a row to select and deselect a service chain, VNF, PNF.
  - To change the sort order of a column, click the column title.

In the following image, the status details column indicate the monitored data path and it provides the per hop analysis.

- Click the **Diagram** button and view the service group with all its service chains and VNFs in the design view window.
- Click a VNF. You can view CPU, memory, and disk allocated to the VNF in a dialog box.
- Select a service group from the **Service Groups** drop-down. The design view displays the selected service group with all its service chains and VNFs.

Step 4 Click the **Network Functions** tab.

In this tab, you can view:

- All the virtual or physical network functions in tabular format. From the **Show** button, you can choose to display either a VNF or PNF.

  VNF information is displayed in tabular format. The table includes information such as VNF name, service chains, CPU use, memory consumption, disk, management IP, Share NF column, and other core parameters that define performance of network service. Click a VNF to view more information about the VNF. See View Information About VNFs , on page 58.

- PNF information is displayed in tabular format. The table includes information such as the serial number and PNF type. To view and note configuration of a specific PNF, click the desired PNF serial number. Ensure that you manually note all the configuration of the PNFs and then configure the PNF devices. For example, the following are some of the PNF configuration where you position the PNF at various locations in the service chain. See Custom Service Chain with Shared PNF Devices, on page 44 to create services chains by adding PNFs. Also, see the ASR 1000 Series Aggregation Services Routers Configuration Guides and Cisco Firepower Threat Defense Configuration Guides to configure the PNFs manually.

Figure 20: PNF in the First Position with Service Chain Side Parameters

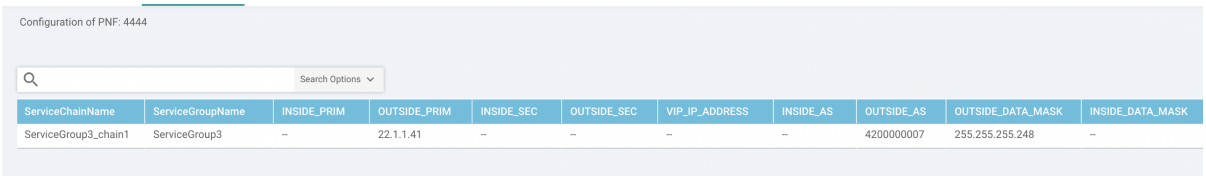


Figure 21: PNF in the First Position with Outside Neighbor Information

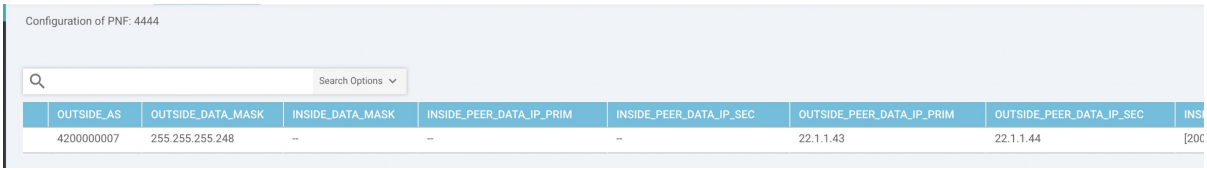


Figure 22: PNF Shared Across Two Service Chains

The ServiceGroup2_chain3 is a PNF-only service chain and therefore no configuration gets generated. The PNF is in the last position of the ServiceGroup2_chain1, so only INSIDE variables gets generated.

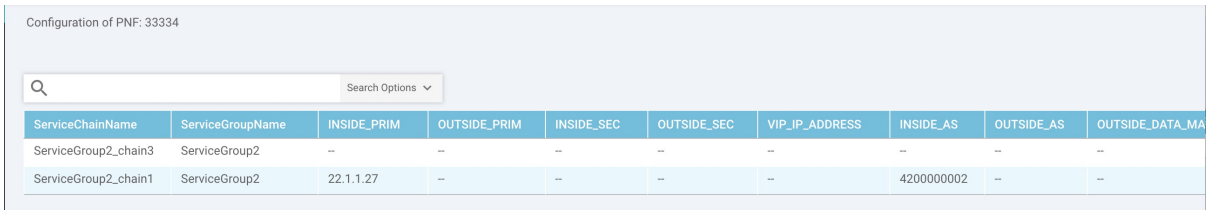


Figure 23: PNF Shared Across Two Service Chains with Outside Neighbor Information



Manage VM Catalog and Repository

vManage supports uploading a prepackaged Cisco VM image, tar.gz in this phase. Alternatively, you can package the VM image by providing a root disk image in any of the supported formats (qcow2). Use the linux command-line NFVIS VM packaging tool, **nfvpt.py** to package the qcow2 or alternatively create a customized VM image from vManage. See Create Customized VNF Image, on page 66.

If VM is SR-IOV capable, which means sriov_supported is set to true in image_properties.xml in the vm package *.tar.gz. Also, the service chain network is automatically connected to SR-IOV network. If sriov_supported is set to false, OVS network is created on the data port channel. It is attached to VM VNICs for service chaining, which is done by using the OVS network. For the Cloud OnRamp for Colocation solution, service chaining a VM uses homogeneous type of network. This type of network means it is either OVS or SR-IOV, and not a combination of SR-IOV and OVS.

Only two data VNICs are attached to any VM–one for inbound traffic and the other for outbound traffic. If more than two data interfaces are required, use subinterfaces configuration within the VM. The VM packages are stored in the VM catalog.

Note Each VM type such as firewall can have multiple VM images that are uploaded to vManage from same or different vendors being added to the catalog. Also, different versions that are based on the release of the same VM can be added to the catalog. However, ensure that the VM name is unique.

The Cisco VM image format can be bundled as *.tar.gz and can include:

- Root disk images to boot the VM.
- Package manifest for checksum validation of the file listing in the package.
- Image properties file in XML format that lists the VM meta data.
- (Optional) Day-0 configuration, other files that are required to bootstrap the VM.
- (Optional) HA Day-0 configuration if VM supports stateful HA.

- System generated properties file in XML format that lists the VM system properties

VM images can be hosted on both HTTP server local repository that vManage hosts or the remote server.

If VM is in NFVIS supported VM package format such as, tar.gz, vManage performs all the processing and you can provide variable key and values during VNF provisioning.



Note vManage only manages the Cisco VNFs, whereas Day-1 and Day-N configurations within VNF are not supported for other VNFs. See the NFVIS Configuration Guide, VM Image Packaging for more information about VM package format and content, and samples on image_properties.xml and manifest (package.mf).

To upload multiple packages for the same VM, same version, Communication Manager (CM) type, ensure that one of the three values (name, version, VNF type) are different. Then, you can repackage the VM *.tar.gz to be uploaded.

Upload VNF Images

The VNF images are stored in software respository. These VNF images are referenced during service chain deployment, and then they are pushed to NFVIS during service chain attachment.

Procedure

In vManage, click **Maintenance** > **Software Repository**. The Maintenance|Software Repository screen appears, and the **Add New Software** button is highlighted. To upload VNF images, use the **Virtual Images** tab. In the Maintenance|Software Repository screen, perform the following tasks:

a) To add a prepackaged VNF image, click the **Virtual Images** tab, and then click the **Upload Virtual Images** button.

b) Choose the location to store the virtual image.

- To store the virtual image on the local vManage server and then get it downloaded to CSP devices over a control plane connection, click **vManage**. The **Upload Software to vManage** dialog box appears.

  1. Drag and drop the virtual image file to the dialog box or click **Browse** to choose the virtual image from the local vManage server. For example, CSR.tar.gz, ASAv.tar.gz.
  2. Click **Upload** to add the image to the virtual image repository. The virtual image repository table displays the added virtual image, and it is available for installing on the CSP devices.

- To store the image on a remote vManage server and then get it downloaded to CSP devices over an out-of-band management connection, click **Remote Server - vManage**. The **Upload Virtual Image to Remote Server - vManage** dialog box appears.

  1. In **vManage Hostname/IP Address**, enter the IP address of an interface on the vManage server that is in a management VPN (typically, VPN 512).
  2. Drag and drop the virtual image file to the dialog box, or click **Browse** to choose the virtual image from the local vManage server.
  3. Click **Upload** to add the image to the virtual image repository. The virtual image repository table displays the added virtual image, and it is available for installing on the CSP devices.

c) Click **Submit**.

You can have multiple VNF entries such as a firewall from same or different vendors. Also, different versions of VNF that are based on the release of the same VNF can be added. However, ensure that the VNF name is unique.

Create Customized VNF Image

Before you begin

You can upload one or more qcow2 images in addition to a root disk image as an input file along with VM-specific properties, bootstrap configuration files (if any), and generate a compressed TAR file. Through custom packaging, you can:

- Create a custom VM package along with image properties and bootstrap files (if needed) into a TAR archive file.
- Tokenize custom variables and apply system variables that are passed with the bootstrap configuration files.

Ensure that the following custom packaging requirements are met:

- Root disk image for a VNF–qcow2
- Day-0 configuration files–system and tokenized custom variables
- VM configuration–CPU, memory, disk, NICs
- HA mode–If a VNF supports HA, specify Day-0 primary and secondary files, NICs for a HA link
- Additional Storage–If additional storage is required, specify predefined disks (qcow2), storage volumes (NFVIS layer)

Procedure

Step 1 In the **Maintenance** > **Software Repository** screen, click the **Add Custom VNF Package** button from the **Virtual Images** tab.

Step 2 Configure the VNF with the following VNF package properties and click **Save**.

Table 10: VNF Package Properties

Field	Mandatory or Optional	Description
Package Name	Mandatory	Specifies the filename of the target VNF package. It is the NFVIS image name with .tar or .gz extensions.
App Vendor	Mandatory	Specifies whether Cisco VNFs or third-party VNFs.
Name	Mandatory	Specifies name of the VNF image.

Field	Mandatory or Optional	Description
Version	Optional	Specifies version number of the program.
Type	Mandatory	Choose VNF type. Supported VNF types are: Router, Firewall, Load Balancer, and Other.

Step 3 To package a VM qcow2 image, click **File Upload** under **Image**, and browse to choose a qcow2 image file.

Step 4 To choose a bootstrap configuration file for VNF, if any, click the **Bootstrap Files** button under **Day 0 Configuration**, click **File Upload**, and then browse to choose a bootstrap file.

Include the following Day-0 configuration properties:

Table 11: Day-0 Configuration

Field	Mandatory or Optional	Description
Mount	Mandatory	Specifies the path where the bootstrap file gets mounted.
Parseable	Mandatory	Specifies whether a Day-0 configuration file can be parsed or not. Options are: true or false. By default, it is true.
High Availability	Mandatory	Choose high availability of a Day-0 configuration file. Supported values are: Standalone, HA Primary, HA Secondary.

Note
If any bootstrap configuration is required for a VNF, you must create *bootstrap-config* or *day0-config*.

Step 5 To add a Day-0 configuration, click **Add**, and then click **Save**. The Day-0 configuration appears in the **Day 0 Config File** table. You can tokenize the bootstrap configuration variables with system and custom variables. To tokenize variables of a Day-0 configuration file, click **View Configuration File** against the configuration file. In the **Day 0 configuration file** dialog box, perform the following tasks:

Note
The bootstrap configuration file is an XML or a text file, and contains properties specific to a VNF and the environment. For a shared VNF, see the topic, Additional References in Cisco SD-WAN Cloud OnRamp for Colocation Solution Guide for the list of system variables that must be added for different VNF types..

a) To add a system variable, in the **CLI configuration** dialog box, select and highlight a property from the text fields. Click **System Variable**. The **Create System Variable** dialog box appears.
b) Choose a system variable from the **Variable Name** drop-down, and click **Done**. The highlighted property is replaced by the system variable name.
c) To add a custom variable, in the **CLI configuration** dialog box, select and highlight a custom variable attribute from the text fields. Click **Custom Variable**. The **Create Custom Variable** dialog box appears.
d) Enter custom variable name and choose a type from **Type** drop-down.

e) To set the custom variable attribute, do the following:

- To ensure that the custom variable is mandatory when creating a service chain, check the **Type** check box against **Mandatory**.
- To ensure that a VNF includes both primary and secondary Day-0 files, check the **Type** check box against **Common**.

f) Click **Done**, and then click **Save**. The highlighted custom variable attribute is replaced by the custom variable name.

Step 6 To upload extra VM images, expand **Advance Options**, click **Upload Image**, and then browse to choose an additional qcow2 image file. Choose the root disk, Ephemeral disk 1, or Ephemeral disk 2, and click **Add**. The newly added VM image appears in the **Upload Image** table.

Note
Ensure that you do not combine ephemeral disks and storage volumes when uploading extra VM images.

Step 7 To add the storage information, expand **Add Storage**, and click **Add volume**. Provide the following storage information and click **Add**. The added storage details appear in the **Add Storage** table.

Table 12: Storage Properties

Field	Mandatory or Optional	Description
Size	Mandatory	Specifies the disk size that is required for the VM operation. The maximum disk size can be 256 if the size unit is GiB.
Size Unit	Mandatory	Choose size unit. Supported units are: MIB, GiB, TiB.
Device Type	Optional	Choose a disk or CD-ROM. Default is a disk.
Location	Optional	Specifies location of the disk or CD-ROM. By default, it is local.
Format	Optional	Choose a disk image format. Supported formats are: qcow2, raw, and vmdk. Buy default, it is raw.
Bus	Optional	Choose a value from the drop-down. Supported values for a bus are: virtio, scsi, and ide. By default, it is virtio.

Step 8 To add VNF image properties, expand **Image Properties** and provide the following image information.

Table 13: VNF Image Properties

Field	Mandatory or Optional	Description
SR-IOV Mode	Mandatory	Specifies enabling or disabling SR-IOV support. By default, it is enabled.
Monitored	Mandatory	VM health monitoring for those VMs that can be bootstrapped. Options are: enable or disable. By default, it is enabled.
Bootup Time	Mandatory	Specifies monitoring timeout period for a monitored VM. By default, it is 600 seconds.
Serial Console	Optional	Specifies serial console that is supported or not. Options are: enable or disable. By default, it is disabled.
Privileged Mode	Optional	Allows special features like promiscuous mode and snooping. Options are: enable or disable. By default, it is disabled.
Dedicate Cores	Mandatory	Facilitates allocation of a dedicated resource (CPU) to supplement a VM's low latency (for example, router and firewall). Otherwise, shared resources are used. Options are: enable or disable. By default, it is enabled.

Step 9 To add VM resource requirements, expand **Resource Requirements** and provide the following information.

Table 14: VM Resource Requirements

Field	Mandatory or Optional	Description
Default CPU	Mandatory	Specifies CPUs supported by a VM. The maximum numbers of CPUs supported are 8.
Default RAM	Mandatory	Specifies RAM supported by a VM. The RAM can range from 2–32.
Disk Size	Mandatory	Specifies disk size in GB supported by a VM. The disk size can range from 4–256.

Field	Mandatory or Optional	Description
Max number of VNICs	Optional	Specifies maximum number of VNICs allowed for the VM. The number of VNICs can range from 8–32 and the default value is 8.
Management VNIC ID	Mandatory	Specifies the management VNIC ID corresponding to the management interface. Valid range is from 0 to maximum number of VNICs.
Number of Management VNICs ID	Mandatory	Specifies number of VNICs.
High Availability VNIC ID	Mandatory	Specifies VNIC IDs where high availability is enabled. Valid range is from 0–maximum number of VNICs. It should not conflict with management VNIC Id. The default value is 1.
Number of High Availability VNICs ID	Mandatory	Specifies maximum number of VNIC IDs where high availability is enabled. Valid range is 0–(maximum number of VNICs-number of management VNICs-2) and default value is 1.

Step 10 To add Day-0 configuration drive options, expand **Day0 Configuration Drive options** and provide the following information.

Table 15: Day-0 Configuration Drive Options

Field	Mandatory or Optional	Description
Volume Label	Mandatory	Displays the volume label of the Day-0 configuration drive. Options are: V1 or V2. By default, it is V2. V2 is the config-drive label config-2. V1 is config-drive label cidata.
Init Drive	Optional	Mounts the Day-0 configuration file as a disk. The default drive is CD-ROM.
Init Bus	Optional	Choose an init bus. Supported values for a bus are: virtio, scsi, and ide. By default, it is ide.

The Software Repository table displays the customized VNF image, and it is available for choosing while creating a custom service chain.

View VNF Images

Procedure

In vManage, click **Maintenance** > **Software Repository**. The Maintenance|Software Repository screen appears, and the **Add New Software** button is highlighted. To view VNF images, use the **Virtual Images** tab. In the Maintenance|Software Repository screen, perform the following tasks:

a) To view VNF images, click the **Virtual Images** tab. The images in the repository are displayed in the table.
b) To filter the list, search or type a string in the Search box.

   The Software Version column provides the version of the software image.

   The Software Location column indicates where the software images are stored. It can be stored either in the repository on the vManage server or in a repository in a remote location.

   The Version Type Name column provides the type of firewall.

   The Available Files column lists the names of the VNF image files.

   The Update On column displays when the software image was added to the repository.

c) To view details of a VNF image, click a VNF image, click the **More Actions** icon, and click **Show Info** against the VNF image.

Delete VNF Images

Procedure

In vManage, click **Maintenance** > **Software Repository**. The Maintenance|Software Repository screen appears, and the **Add New Software** button is highlighted. To upload VM images, use the **Virtual Images** tab. In the Maintenance|Software Repository screen, perform the following tasks:

a) To delete a VM image, click the **Virtual Images** tab. The images in the repository are displayed in the table.
b) In the repository table, click a VM image.
c) Click the **More Actions** icon to the right of its row, and click **Delete** against the VM image.

Note
If a VNF image is being download to a router, you cannot delete the VNF image until the download process completes.



Note If the VNF image is referenced by a service chain, it cannot be deleted.

Upgrade NFVIS Software Through vManage

To upload and upgrade NFVIS, the upgrade image must be available as an archive file that can be uploaded to vManage repository through vManage. After you upload the NFVIS image, the upgraded image can be applied to a CSP device by using the Software Upgrade screen in vManage. You can perform the following tasks during upgrading NFVIS software through vManage:

- Upload NFVIS upgrade image. See Upload NFVIS Upgrade Image, on page 72.
- Upgrade a CSP device with the uploaded image. See Upgrade CSP Device with NFVIS Upgrade Image, on page 72.
- View the upgrade status in the CSP device. See the "View Log of Software Upgrade Activities" in the Cisco SD-WAN Configuration Guide.

Upload NFVIS Upgrade Image

Procedure

Step 1 Download the NFVIS upgrade image from a prescribed location to your local system. You can also download the software image to an FTP server in your network.

Step 2 Log into Cisco vManage, at the URL HTTPS: //vManage-ip-address/, as a user with "admin" credentials.

Step 3 In the **Maintenance** > **Software Repository** screen, click the **Add New Software** > **Remote Server/Remote Server - vManage** button.

You can either store the software image on a remote file server, on a remote vManage server, or on a vManage server.

Note
The vManage server is available in the current version.

vManage server–saves software images on a local vManage server.

Remote server–saves the URL pointing to the location of the software image and can be accessed through an FTP or HTTP URL.

Remote vManage server–saves software images on a remote vManager server and location of the remote vManage server is stored in the local vManage server.

Step 4 To add the image to the software repository, browse and choose the NFVIS upgrade image that you had downloaded in step1.

Step 5 Click **Add|Upload**.

The Software Repository table displays the added NFVIS upgrade image, and it is available for installing on the CSP devices. See the "Software Repository" topic in the Cisco SD-WAN Configuraion Guides.

Upgrade CSP Device with NFVIS Upgrade Image

Before you begin

Ensure that the NFVIS software versions are the files that have `.nfvispkg` extension.

Procedure

Step 1 In the **Maintenance** > **Software Upgrade** > **WAN Edge** screen, view the list of all CSP devices along with their current and available versions.

Step 2 Select one or more devices, and click **Upgrade**.

Step 3 Choose a CSP device on which to upgrade the NFVIS software image.

Step 4 Click the **Upgrade** button. The **Software Upgrade** dialog box appears.

Step 5 Choose the NFVIS software version to install on the CSP device. If software is located on a remote server, choose the appropriate remote version.

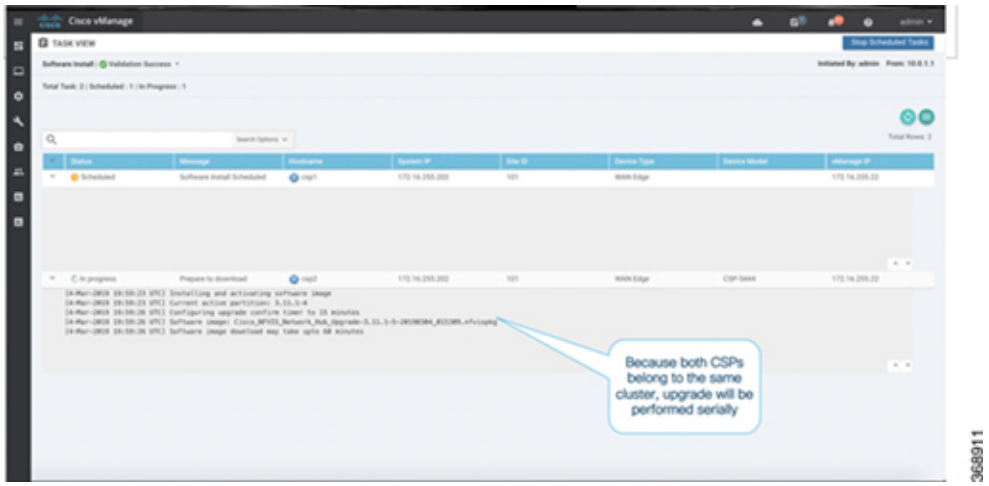
Step 6 To automatically upgrade and activate with the new NFVIS software version and reboot the CSP device, check the **Activate and Reboot** checkbox.

If you do not check the **Activate and Reboot** checkbox, the CSP device downloads and verifies the software image. However, the CSP device continues to run the old or current version of the software image. To enable the CSP device to run the new software image, you must manually activate the new NFVIS software version by selecting the device again and clicking the **Activate** button on the **Software Upgrade** page. For more information about activation, see the "Activate a New Software Image" topic in the Cisco SD-WAN Configuration Guides.

Step 7 Click **Upgrade**.

To view the status of software upgrades, the task view page displays a list of all running tasks along with total number of successes and failures. The page periodically refreshes and displays messages to indicate the progress or status of the upgrade. You can easily access the software upgrade status page by clicking the Tasks icon located in the vManage toolbar.

Note
If two or more CSP devices belonging to the same cluster are upgraded, the software upgrade for the CSP devices happen in a sequence.



Note
The **Set the Default Software Version** option is not available for NFVIS images.

The CSP device reboots and the new NFVIS version is activated on it. This reboot happens during the **Activate** phase. The activation can either happen immediately after upgrade if you check the **Activate and Reboot** check box, or by manually selecting the activate button after selecting the device again.

To verify if CSP device has rebooted and is running, vManage polls your entire network every 90 seconds up to 30 times.



Note You can delete an NFVIS software image from a CSP device if the image version is not the active version that is running on the device.

TCP Optimization: Cisco vEdge Devices

TCP optimization fine-tunes the processing of TCP data traffic to decrease round-trip latency and improve throughput. Optimizing TCP traffic is especially useful for improving TCP traffic performance on long-latency links, such as transcontinental links and the high-latency transport links used by VSAT satellite communications systems. TCP optimization can also improve the performance of SaaS applications.

Configure TCP Optimization

Supported Platforms

TCP Optimization is supported on the following platforms:

- Cisco vEdge 1000
- Cisco vEdge 2000

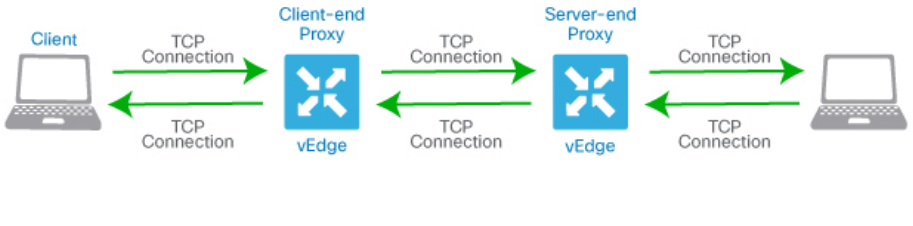
TCP Optimization is not supported on the following platforms:

- Cisco vEdge 100
- Cisco vEdge 100b
- Cisco vEdge 100m
- Cisco vEdge 100mw

- Cisco vEdge 5000
- Cisco ISR1100 4G
- Cisco ISR1100 6G
- Cisco ISR1100 4GLTEGB
- Cisco ISR1100 4GLTENA

TCP optimization fine-tunes the processing of TCP data traffic to decrease round-trip latency and improve throughput. You can optimize TCP traffic in service-side VPNs on vEdge routers. Optimizing TCP traffic is especially useful for improving TCP traffic performance on long-latency links, such as transcontinental links and the high-latency transport links used by VSAT satellite communications systems. TCP optimization can also improve the performance of SaaS applications.

With TCP optimization, a vEdge router acts as a TCP proxy between a client that is initiating a TCP flow and a server that is listening for a TCP flow, as illustrated in the following figure:



This figure shows two vEdge routers acting as proxies. vEdge-1 is the proxy for the client, and is called the client proxy. vEdge-2 is the proxy for the server, called the server proxy. Without TCP optimization, the client establishes a TCP connection directly to the server. When you enable TCP optimization on the two routers, vEdge-1 terminates the TCP connection from client and establishes a TCP connection with vEdge-2. The vEdge-2 router then establishes a TCP connection to the server. The two vEdge routers cache the TCP traffic in their buffers to ensure that the traffic from the client reaches the server without the TCP connection timing out.

It is recommended that you configure TCP optimization on both the vEdge routers, the router closer to the client and the router closer to the server. This configuration is sometimes referred to as a dual-ended proxy. While it is possible to configure TCP optimization only on the vEdge router closer to the client, called single-ended proxy, this configuration is not recommended because the TCP optimization process is compromised. TCP is a bidirectional protocol and operates only when connection-initiation messages (SYNs) are acknowledged by ACK messages in a timely fashion.

If both the client and the server are connected to the same vEdge router, no TCP optimization is performed.

To use TCP optimization, you first enable the feature on the vEdge router. Then you define which TCP traffic to optimize.

1. To use TCP optimization on a vEdge router, you must enable it:

   ```
   vEdge(config-system)#  tcp-optimization-enabled
   ```

   On vEdge 1000 and vEdge 2000 routers, enabling TCP optimization carves out a separate CPU core to use for performing TCP optimization, which is a CPU-intensive process.

2. To enable TCP optimization for all TCP traffic in a VPN, include the following command when configuring service-side VPNs on a vEdge router:

   ```
   vEdge(config-vpn)# tcp-optimization
   ```

For example, to enable TCP optimization in VPN 1:

```
vEdge# show running-config vpn 1
vpn 1
interface ge0/4.1
  ip address 10.20.24.15/24
  no shutdown
!
tcp-optimization
!
```

To display information about the TCP flows that the vEdge router is optimizing, use the **show app tcp-opt active-flows** command to view flows that are currently active and the **show app tcp-opt expired-flows** command to view flows that were active but have expired.

3. Reboot the vEdge router.

```
vEdge# reboot
```

To enable TCP optimization for a specific flow of TCP traffic in a VPN, create a centralized data policy that includes the **tcp-optimization** action. Use the match conditions in the data policy to define the VPN in which to optimize the TCP traffic and the traffic properties to match.

The following example enables TCP optimization on all TCP traffic destined to port 22, which is used for SSH sessions:

```
vSmart# show running-config policy data-policy tcp_optimization_data_policy
policy
 data-policy tcp_optimization_data_policy
  vpn-list vpn_2
   sequence 100
    match
     destination-port 22
    !
    action accept
     count            sample_count
     tcp-optimization
    !
   !
   default-action accept
  !
 !
!

vSmart# show running-config apply-policy
apply-policy site-list tcp_optimization_sites data-policy tcp_optimization_data_policy
all
```

Note The TCP optimization feature is enabled based on the data policy. The data policy match needs to be created for the TCP flows in both directions. For example, if the TCP flow is from 10.20.24.17:30000 -> 10.20.25.18:5001, the data policy to enable TCP optimization must be created to match the reverse direction also. That is from 10.20.25.18:5001 -> 10.20.24.17:30000.

Monitor TCP Optimization

If TCP optimization is enabled on a router, you can view information about how the optimization is affecting the processing and throughput of TCP data traffic on the router:

1. From the **Monitor** > **Network** screen, select a vEdge router.
2. Click TCP Optimization–WAN Throughput in the left page. The right pane displays the WAN throughput, in megabits per second.

The upper part of the right pane contains the following elements:

- Chart Options bar—Located directly under the device name, this bar includes the Filter Options drop-down and time periods. Click Filter to limit the data to display based on VPN, local TLOC color, destination IP address, remote TLOC color, and remote system IP address. Click a predefined or custom time period for which to display data.
- Average optimized throughput information in graphical format.
- WAN graph legend—Identifies non-optimized and TCP optimized packet throughput.

The lower part of the right pane shows the hourly average throughput and the total optimized throughput, both in megabits per second.

Click TCP Optimization–Flows in the left pane to display information about TCP-optimized traffic flows. The upper part of the right pane contains the following elements:

- Chart Options bar—Located directly under the device name, this bar includes the Filter drop-down and time periods. Click Filter to limit the data to display based on VPN, local TLOC color, destination IP address, remote TLOC color, and remote system IP address. Click a predefined or custom time period for which to display data.
- Average optimized throughput information in graphical format.
- Flows graph legend—Identifies traffic flows.

The lower part of the right pane contains the following elements:

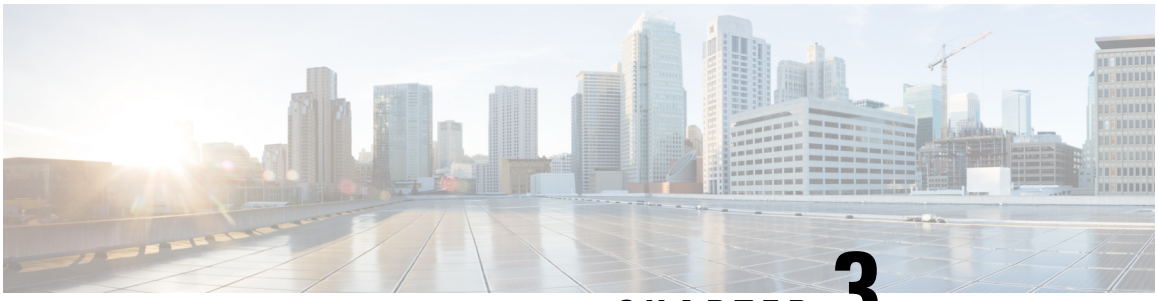
- Set perspective—Select the flow direction.
- Search box—Includes the Search Options drop-down, for a Contains or Match string.
- Flow table that lists the flow destination, usage, and percentage of total traffic for all TCP-optimized flows. By default, the first six flows are selected. Click the checkbox to the left to select and deselect flows to display. The graphical display in the upper part of the right pane plots information for the selected flows.

Click TCP Optimization–Connections in the left pane to display status information about all the tunnels over which the most TCP-optimized traffic is flowing. The upper part of the right pane contains the following elements:

- TCP Optimization Connections in graphical format
- Connection State boxes—Select the connection state or states to display TCP optimization information about.

The lower part of the right pane contains the following elements:

- Filter criteria.
- Flow table that lists information about each of the tunnels, including the tunnel's connection state.



CHAPTER 3

High Availability Overview

The goal of any high availability solution is to ensure that all network services are resilient to failure. Such a solution aims to provide continuous access to network resources by addressing the potential causes of downtime through functionality, design, and best practices. The core of the Cisco SD-WAN high availability solution is achieved through a combination of three factors:

- Functional hardware device redundancy. The basic strategy consists of installing and provisioning redundant hardware devices and redundant components on the hardware. These devices are connected by a secure control plane mesh of Datagram Transport Layer Security (DTLS) connections among themselves, which allows for rapid failover should a device fail or otherwise become unavailable. A key feature of the Cisco SD-WAN control plane is that it is established and maintained automatically, by the Cisco vEdge devices and software themselves.
- Robust network design.
- Software mechanisms ensure rapid recovery from a failure. To provide a resilient control plane, the Cisco SD-WAN Overlay Management Protocol (OMP) regularly monitors the status of all Cisco vEdge devices in the network and automatically adjusts to changes in the topology as devices join and leave the network. For data plane resiliency, the Cisco SD-WAN software implements standard protocol mechanisms, specifically Bidirectional Forwarding Detection (BFD), which runs on the secure IPsec tunnels between routers.

Recovery from a failure is a function of the time it takes to detect the failure and then repair or recover from it. TheCisco SD-WAN solution provides the ability to control the amount of time to detect a failure in the network. In most cases, repair of the failure is fairly instantaneous.

Hardware Support of High Availability

A standard best practice in any network setup is to install redundant hardware at all levels, including duplicate parallel routers and other systems, redundant fans, power supplies and other hardware components within these devices, and backup network connections. Providing high availability in the Cisco SD-WAN solution is no different. A network design that is resilient in the face of hardware failure should include redundant vBond orchestrators, vSmart controllers, and routers and any available redundant hardware components.

Recovery from the total failure of a hardware component in the Cisco SD-WAN overlay network happens in basically the same way as in any other network. A backup component has been preconfigured, and it is able to perform all necessary functions by itself.

Robust Network Design

In addition to simple duplication of hardware components, the high availability of a Cisco SD-WAN network can be enhanced by following best practices to design a network that is robust in the face of failure. In one such network design, redundant components are spread around the network as much as possible. Design practices include situating redundant vBond orchestrators and vSmart controllers at dispersed geographical locations and connecting them to different transport networks. Similarly, the routers at a local site can connect to different transport networks and can reach these networks through different NATs and DMZs.

Software Support of High Availability

The Cisco SD-WAN software support for high availability and resiliency in the face of failure is provided both in the control plane, using the standard DTLS protocol and the proprietary Cisco SD-WAN Overlay Management Protocol (OMP), and in the data plane, using the industry-standard protocols BFD, BGP, OSPF, and VRRP.

Control Plane Software Support of High Availability

The Cisco SD-WAN control plane operates in conjunction with redundant components to ensure that the overlay network remains resilient if one of the components fails. The control plane is built on top of DTLS connections between the Cisco devices, and it is monitored by the Cisco SD-WAN OMP protocol, which establishes peering sessions (similar to BGP peering sessions) between pairs of vSmart controllers and routers, and between pairs of vSmart controllers. These peering sessions allow OMP to monitor the status of the Cisco devices and to share the information among them so that each device in the network has a consistent view of the overlay network. The exchange of control plane information over OMP peering sessions is a key piece in the Cisco SD-WAN high availability solution:

- vSmart controllers quickly and automatically learn when a vBond orchestrator or a router joins or leaves the network. They can then rapidly make the necessary modifications in the route information that they send to the routers.
- vBond orchestrators quickly and automatically learn when a device joins the network and when a vSmart controller leaves the network. They can then rapidly make the necessary changes to the list of vSmart controller IP addresses that they send to routers joining the network.
- vBond orchestrators learn when a domain has multiple vSmart controllers and can then provide multiple vSmart controller addresses to routers joining the network.
- vSmart controllers learn about the presence of other vSmart controllers, and they all automatically synchronize their route tables. If one vSmart controller fails, the remaining systems take over management of the control plane, simply and automatically, and all routers in the network continue to receive current, consistent routing and TLOC updates from the remaining vSmart controllers.

Let's look at the redundancy provided by each of the Cisco SD-WAN hardware devices in support of network high availability.

Recovering from a Failure in the Control Plane

The combination of hardware component redundancy with the architecture of the Cisco SD-WAN control plane results in a highly available network, one that continues to operate normally and without interruption when a failure occurs in one of the redundant control plane components. Recovery from the total failure of a vSmart controller, vBond orchestrator, or router in the Cisco SD-WAN overlay network happens in basically the same way as the recovery from the failure of a regular router or server on the network: A preconfigured backup component is able to perform all necessary functions by itself.

In the Cisco SD-WAN solution, when a network device fails and a redundant device is present, network operation continues without interruption. This is true for all Cisco devices—vBond orchestrators, vSmart controllers, and routers. No user configuration is required to implement this behavior; it happens automatically. The OMP peering sessions running between Cisco devices ensure that all the devices have a current and accurate view of the network topology.

Let's examine failure recovery device by device.

Data Plane Software Support for High Availability

For data plane resiliency, the Cisco SD-WAN software implements the standard BFD protocol, which runs automatically on the secure IPsec connections between routers. These IPsec connections are used for the data plane, and for data traffic, and are independent of the DTLS tunnels used by the control plane. BFD is used to detect connection failures between the routers. It measures data loss and latency on the data tunnel to determine the status of the devices at either end of the connection.

BFD is enabled, by default, on connections between Cisco XE SD-WAN devices and Cisco vEdge devices. BFD sends Hello packets periodically (by default, every 1 second) to determine whether the session is still operational. If a certain number of the Hello packets are not received, BFD considers that the link has failed and brings the BFD session down (the default dead time is 3 seconds). When BFD sessions goes down, any route that points to a next hop over that IPsec tunnel is removed from the forwarding table (FIB), but it is still present in the route table (RIB).

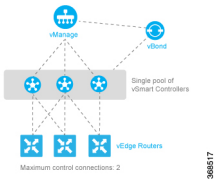
In the Cisco SD-WAN software, you can adjust the Hello packet and dead time intervals. If the timers on the two ends of a BFD link are different, BFD negotiates to use the lower value.

Using Affinity To Manage Network Scaling

In the Cisco SD-WAN overlay network, all Cisco XE SD-WAN devices and Cisco vEdge devices establish control connections to all vSmart controllers, to ensure that the routers are always able to properly route data traffic across the network. As networks increase in size, with routers at thousands of sites and with vSmart controllers in multiple data centers managing the flow of control and data traffic among routers, network operation can be improved by limiting the number of vSmart controllers that a router can connect to. When data centers are distributed across a broad geography, network operation can also be better managed by having routers establish control connections only with the vSmart controllers collocated in the same geographic region.

Establishing affinity between vSmart controllers and Cisco vEdge devices allow you to control the scaling of the overlay network, by limiting the number of vSmart controllers that a Cisco vEdge device can establish control connections (and form TLOCs) with. When you have redundant routers in a single data center, affinity allows you to distribute the vEdge control connections across the vSmart controllers. Similarly, when you have multiple data centers in the overlay network, affinity allows you to distribute the vEdge control connections across the data centers. With affinity, you can also define primary and backup control connections, to maintain overlay network operation in case the connection to a single vSmart controller or to a single data center fails.

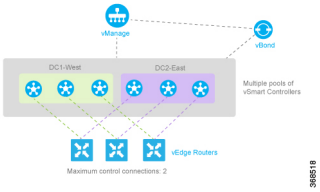
A simple case for using affinity is when redundant vSmart controllers are collocated in a single data center. Together, these vSmart controllers service all the Cisco vEdge devices in the overlay network. The figure below illustrates this situation, showing a scenario with three vSmart controllers in the data center and, for simplicity, showing just three of the many Cisco vEdge devices in the network.



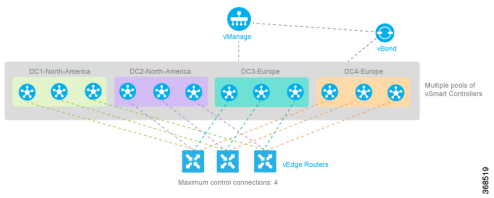
If you do not enable affinity, each Cisco vEdge device establishes a control connection—that is, a TLOC—to each of the three vSmart controllers in the data center. Thus, a total of nine TLOCs are established, and each router exchanges OMP updates with each controller. Having this many TLOCs can strain the resources of both the vSmart controllers and the Cisco vEdge devices , and the strain increases in networks with larger numbers of Cisco vEdge devices .

Enabling affinity allows each Cisco vEdge device to establish TLOC connections with only a subset of the vSmart controllers. The figure above shows each router connecting to just two of the three vSmart controllers, thus reducing the total number of TLOCs from nine to six. Both TLOC connections can be active, for a total of six control connections. It is also possible for one of the TLOC connections be the primary, or preferred, and the other to be a backup, to be used as an alternate only when the primary is unavailable, thus reducing the number of active TLOCs to three.

Affinity also enables redundancy among data centers, for a scenario in which multiple vSmart controllers are collocated in two or more data centers. Then, if the link between a Cisco vEdge device and one of the data centers goes down, the vSmart controllers in the second data center are available to continue servicing the overlay network. The figure below illustrates this scenario, showing three vSmart controllers in each of two data centers. Each of the three Cisco vEdge devices establishes a TLOC connection to one controller in the West data center and one in the East data center.



You might think of the scenario in the figure above as one where there are redundant data centers in the same region of the world, such as in the same city, province, or country. For an overlay network that spans a larger geography, such as across a continent or across multiple continents, you can use affinity to limit the network scale either by restricting Cisco vEdge devices so that they connect only to local vSmart controllers or by having Cisco vEdge devices preferentially establish control connections with data centers that are in their geographic region. With geographic affinity, Cisco vEdge devices establish their only or their primary TLOC connection or connections with vSmart controllers in more local data centers, but they have a backup available to a more distant region to provide redundancy in case the closer data centers become unavailable. The figure below illustrates this scenario, Here, the Cisco vEdge devices in Europe have their primary TLOC connections to the two European data centers and alternate connections to the data centers in North America. Similarly, for the Cisco vEdge devices in North America, the primary connections are to the two North American data centers, and the backup connections are to the two European data centers.



As is the case with any overlay network that has multiple vSmart controllers, all policy configurations on all the vSmart controllers must be the same.

Before you cofigure High availability, to start with the configuration transaction, you can use the following command such as,

```
ntp server 198.51.241.229 source GigabitEthernet1 version 4
```


- vBond Orchestrator Redundancy, on page 83
- vManage NMS Redundancy, on page 84
- vSmart Controller Redundancy, on page 91
- Cisco vEdge Device Redundancy, on page 92
- Configure Affinity between vSmart and Cisco vEdge Device , on page 92
- Configure Control Plane and Data Plane High Availability Parameters, on page 96
- Configure Disaster Recovery, on page 97
- High Availability CLI Reference, on page 104
- High Availability Configuration Examples, on page 105

vBond Orchestrator Redundancy

The vBond orchestrator performs two key functions in the Cisco SD-WAN overlay network:

- Authenticates and validates all vSmart controllers and routers that attempt to join the Cisco SD-WAN network.
- Orchestrates the control plane connections between the vSmart controllers and routers, thus enabling vSmart controllers and routers to connect to each other in the Cisco SD-WAN network.

The vBond orchestrator runs as a VM on a network server. The vBond orchestrator can also run on a router that is configured to be a vBond orchestrator, however this is not recommended, and it limits the number of router control connections to 50. If using running the vBond daemon on a router, note that only one vBond daemon can run at a time on a router, so to provide redundancy and high availability, the network must have two or more routers that function as vBond orchestrators. (Note also that it is not recommended to use a router acting as a vBond orchestrator as a regular router.)

Having multiple vBond orchestrators ensures that one of them is always available whenever a Cisco device such as a router or a vSmart controller is attempting to join the network.

Configuration of Redundant vBond Orchestrators

A router learns that it is acting as a vBond orchestrator from its configuration. In the **system vbond** configuration command, which defines the IP address (or addresses) of the vBond orchestrator (or orchestrators) in the Cisco SD-WAN overlay network, you include the **local** option. In this command, you also include the local public IP address of the vBond orchestrator. (Even though on Cisco vEdge device and vSmart controllers you can specify an IP address of vBond orchestrator as a DNS name, on the vBond orchestrator itself, you must specify it as an IP address.)

On vSmart controllers and Cisco vEdge devices, when the network has only a single vBond orchestrator, you can configure the location of the vBond system either as an IP address or as the name of a DNS server (such as vbond.cisco.com). (Again, you configure this in the **system vbond** command.) When the network has two or more vBond orchestrators and they must all be reachable, you should use the name of a DNS server. The DNS server then resolves the name to a single IP address that the vBond orchestrator returns to the Cisco

vEdge device. If the DNS name resolves to multiple IP addresses, the vBond orchestrator returns them all to the Cisco vEdge device, and the router tries each address sequentially until it forms a successful connection.

Note that even if your Cisco SD-WAN network has only a single vBond orchestrator, it is recommended as a best practice that you specify a DNS name rather than an IP address in the **system vbond** configuration command, because this results in a scalable configuration. Then, if you add additional vBond orchestrators to your network, you do not need to change the configurations on any of the routers or vSmart controllers in your network.

Recovering from a vBond Orchestrator Failure

In a network with multiple vBond orchestrators, if one of them fails, the other vBond orchestrators simply continue operating and are able to handle all requests by Cisco devices to join the network. From a control plane point of view, each vBond orchestrator maintains a permanent DTLS connections to each of the vSmart controllers in the network. (Note however, that there are no connections between the vBond orchestrators themselves.) As long as one vBond orchestrator is present in the domain, the Cisco SD-WAN network is able to continue operating without interruption, because vSmart controllers and routers are still able to locate each other and join the network.

Because vBond orchestrators never participate in the data plane of the overlay network, the failure of any vBond orchestrator has no impact on data traffic. vBond orchestrators communicate with routers only when the routers are first joining the network. The joining router establishes a transient DTLS connection with a vBond orchestrator to learn the IP address of a vSmart controller. When the Cisco vEdge device configuration lists the vBond address as a DNS name, the router tries each of the vBond orchestrators in the list, one by one, until it is able to establish a DTLS connection. This mechanism allows a router to always be able to join the network, even after one of a group of vBond orchestrators has failed.

vManage NMS Redundancy

The vManage NMSs comprise a centralized network management system that enables configuration and management of the Cisco devices in the overlay network. It also provides a real-time dashboard into the status of the network and network devices. The vManage NMSs maintain permanent communication channels with all Cisco vEdge devices in the network. Over these channels, the vManage NMSs push out files that list the serial numbers of all valid devices, they push out each device's configuration, and they push out new software images as part of a software upgrade process. From each network device, the vManage NMSs receive various status information that is displayed on the vManage Dashboard and other screens.

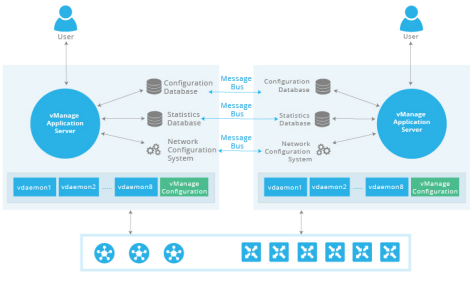
A highly available Cisco SD-WAN network contains three or more vManage NMSs in each domain. This scenario is referred to as a cluster of vManage NMSs, and each vManage NMS in a cluster is referred to as a vManage instance. Each vManage instance or device in a cluster can manage approximately 2000 devices, so a cluster of three vManage devices can manage up to 6000 devices. The vManage devices automatically load-balances the devices that they manage. With three devices, the vManage cluster remains operational if one of the devices in that cluster fail. A vManage cluster consists of the following architectural components:

- Application server—This provides a web server for user sessions. Through these sessions, a logged-in user can view a high-level dashboard summary of networks events and status, and can drill down to view details of these events. A user can also manage network serial number files, certificates, software upgrades, device reboots, and configuration of the vManage cluster itself from the vManage application server.
- Configuration database—Stores the inventory and state and the configurations for all Cisco vEdge devices.
- Network configuration system—Stores all configuration information, policies, templates, certificates, and more.

- Statistics database—Stores the statistics information collected from all Cisco devices in the overlay network.
- Message bus—Communication bus among the different vManage instances. This bus is used to share data and coordinate operations among the vManage instances in the cluster.

The Statistics database and Configuration database services must run on an odd number of vManage instances, with a minimum of three. For these databases to be writeable, there must be a quorum of vManage instances running and they should be in sync. A quorum is a simple majority. For example, if you have a cluster of three vManage devices running these databases, then two must be running and in sync. Initially, all vManage devices run the same services. However, you can choose not to run some services on some devices. From the Cluster Management screen, you can select the services that can run on each vManage. You can add a fourth vManage device to load-balance more Cisco vEdge device s. In such a case, disable the Statistics database and Configuration database on one of the vManage devices because those services need to run on an odd number of devices. Optionally, you can run the Configuration database on a single vManage device to reduce the amount of information shared between the devices and reduce load.

The following figure shows the interaction between vManage devices in a cluster, although a minimum of three devices are required. The figure illustrates the NMS services that synchronize between the vManage devices. Also in this figure, you see that each vManage instance resides on a virtual machine (VM). The VM can have from one to eight cores, with a Cisco SD-WAN software process (vdaemon) running on each core. In addition, the VM stores the actual configuration for the vManage NMS itself.



The vManage cluster implements an active-active architecture in the following way:

- Each vManage instance in the cluster is an independent processing node.
- All vManage instances are active simultaneously.
- All user sessions to the application server are load-balanced by using an external load balancer.
- All control sessions between the vManage application servers and the routers are load-balanced. A single vManage instance can manage a maximum of about 2000 Cisco vEdge devices . However, all the controller sessions—the sessions between the vManage instances and the vSmart controllers, and the sessions between the vManage instances and the vBond orchestrators—are arranged in a full-mesh topology.
- The configuration and statistics databases can be replicated across vManage instances, and these databases can be accessed and used by all the vManage instances.
- If one of the vManage instances in the cluster fails or otherwise becomes unavailable, the network management services that are provided by the vManage NMS are still fully available across the network.

The message bus among the vManage instances in the cluster allows all the instances to communicate using an out-of-band network. This design, which leverages a third vNIC on the vManage VM, avoids using WAN bandwidth for management traffic.

You configure the vManage cluster from the vManage web application server. During the configuration process, you can configure each vManage instance that can run the following services:

- Application server—Each vManage server runs an application server instance.
- Configuration database—Within the vManage cluster, no more than three iterations of the configuration database can run.
- Load balancer—The vManage cluster requires a load balancer to distribute user login sessions among the vManage instances in the cluster. As mentioned above, a single vManage instance can manage a maximum of 2000 Cisco vEdge devices.
- Messaging server—It is recommended that you configure each vManage instance to run the message bus so that all the instances in the cluster can communicate with each other.
- Statistics database—Within the vManage cluster, no more than three iterations of the statistics database can run.
- Coordination server: It is used internally by the Messaging server.

The following are the design considerations for a vManage cluster:

- A vManage cluster should consist of a minimum of three vManage instances.
- The application server and message bus should run on all vManage instances.
- Within a cluster, a maximum of three instances of the configuration database and three instances of the statistics database can run. Note, however, that any individual vManage instance can run both, one, or none of these two databases.
- To provide the greatest availability, it is recommended that you run the configuration and statistics databases on three vManage instances.

Deploy vManage Cluster

Ensure that you have a minimum of three vManage devices in a vManage cluster.



Note This process requires multiple reboots and should be performed during a scheduled maintenance window.

1. Back up the vManage database. Use the following command to back up.

   request nms configuration-db backup path /home/admin/<db _ backup _ filename>

2. If the current vManage device has only two network interface cards (NICs), add a third NIC. Do not use Dynamic Host Configuration Protocol (DHCP) for addressing. This third NIC is used for cluster messaging between the vManage devices, within vpn 0. For the vManage device to detect the new interface, the device must be rebooted. Configure the interface and verify that it has connectivity.
3. In vManage GUI, click **Administration** > **Cluster Management**, edit the IP address to change localhost to the IP address of the third NIC, which is used for cluster messaging.

4. Restart the vManage NMS services. This may take some time. You can view the */var/log/nms/vmanage-server.log* for the log output to stabilize, and then use the **request nms all status** command to determine for how long the processes have been running. When it comes up, verify that vManage is operational and stable.

5. Provision the two additional vManage VMs in your virtual environment with the appropriate disk size and third NIC.

6. Configure the two additional vManage VMs with minimal system configuration and addressing for the NICs. Configure the admin user password to match that is configured on the original vManage device. If you are using enterprise certificates, ensure that you install the root certificate chain on the new vManage device as you did with the first vManage device. Also, ensure the clocks of the new vManage devices are in sync with the original vManage device.

   The following is a sample of a minimal configuration:

```
system
            host-name             vManage3
            system-ip             10.0.1.102
            site-id               1
            organization-name     cisco-sdwan1
            vbond vbond!
            vpn 0
                host vbond ip 198.51.100.103 192.51.100.104
                interface eth0
                ip address 198.51.100.102/24
                tunnel-interface
                no shutdown
                !
                interface eth1
                ip address 198.51.100.102/24
                no shutdown !
                ip route 10.0.0.1/0 10.0.1.254
                ip route 10.0.0.1/0 10.0.1.254
                !
                vpn 512
                interface eth2
                    ip address 198.56.101.102/24
                    no shutdown
                !
```



Note While a default gateway is given for the out of band vManage cluster interfaces, it is not required if the vManage cluster nodes are using addresses in the same subnet and are adjacent.

7. In vManage GUI, click **Administration** > **Cluster Management**, to add one of the new vManage VMs to the cluster by adding the IP address of the third NIC for database replication, click **Add vManage**. Select all services.

   vManage2 NMS processes restart. This process might take some time. View the */var/log/nms/vmanage-server.log* and then use the **request nms all status** command to determine process completion time.

8. View th new vManage device on the **Configuration** > **Certificates** > **Controllers** page.

9. Generate a certificate signing request (CSR), get the device signed, and install the signed device certificate for this new vManage device. See Cluster Management for more information.

The cluster shows that the new vManage device is rebalancing and that NMS services are restarting on the previous vManage device. A new task appears in the task bar for the **Cluster Sync**. Although the task appears as Complete, view the */var/log/nms/vmanage-server.log* to resolve errors, if any.

vManage1 also restarts NMS services, which eventually resets the GUI session. It might take several minutes for the GUI to become available after the NMS services restarts. View the */var/log/nms/vmanage-server.log*, and then use the **request nms all status** command.

10. Wait for **Cluster Sync** to complete. View the *vmanage-server.log* to resolve errors, if any.
11. After the Cluster Sync is completed, add the second of the new vManage VMs to the cluster by adding the IP address of the third NIC for database replication. Select all services.

    vManage3 restarts NMS services. This might take some time. A new task appears in the task bar for the **Cluster Sync**. vManage1 and vManage2 also restarts NMS services, which eventually resets the GUI session. It may take several minutes for the GUI to become available after the NMS services restart. The new vManage device appears on the **Configuration** > **Certificates** > **Controllers** page. Perform steps 9 and 10 for this vManage device.

Upgrade vManage Cluster

To upgrade a cluster, ensure that the services start in an orderly manner. After the upgrade steps, use the steps in the Restarting the NMS Processes section to start all the services in an orderly manner.

To get the software partitions prepared to be activated, upgrade the devices (without activating).

1. Collect an NMS backup. This can take a while. If Cisco hosts the controllers and there is a recent snapshot, this step can be skipped.

   ```
   request nms configuration-db backup path/home/admin/<db _ backup _ filename>
   ```

2. Stop NMS services on all vManage devices in the cluster by using the following command on each device.

   ```
   request nms all stop
   ```

3. Activate the new version on each device. This activation causes each device to reload.

   ```
   request software activate <version>
   ```

4. If you do the upgrade from CLI, ensure that you manually confirm the upgrade from the CLI after the reload and before it reverts to the previous version.

   ```
   request software upgrade-confirm
   ```

5. After the vManage devices reboot, stop NMS services on all vManage devices in the cluster.

   ```
   request nms all stop
   ```

Next, ensure that you perform the steps of restarting the NMS process manually.

Restarting the NMS Processes Manually

When the cluster is in a bad state as part of the upgrade, you should manually restart the NMS processes. Restart the processes one at a time in an orderly manner instead of using **request nms all restart** or a similar command. The following manual restart order might vary for your cluster, depending on what services you are running on the vManage devices in the cluster. The following order is based on a basic cluster with three vManage devices.

Note Consider bringing up the services manually as mentioned in the following method whenever you have to reboot a vManage device or after an upgrade.

1. On each vManage device, stop all NMS services.

   ```
   request nms all stop
   ```

2. Verify that all services have stopped. It is normal for the above command to give some message about failing to stop a service if it takes too long, so use the following command to verify that everything is stopped before proceeding.

   ```
   request nms all status
   ```

3. Start the Statistics database on each device that is configured to run it. Wait for the service to start each time before proceeding to the next vManage device.

   ```
   request nms statistics-db start
   ```

4. Verify that the service is started before proceeding to start it on the next vManage. After service starts, perform step 3 to start the Statistics database on the next vManage device. Once all the vManage devices have the Statistics database running, proceed to the next step.

   ```
   request nms statistics-db status
   ```

5. Start the Configuration database on each device that is configured to run it. Wait for the service to start each time before proceeding to the next vManage device.

   ```
   request nms configuration-db start
   ```

6. Verify that the service has started before proceeding to start it on the next vManage device. Go to vshell and tail a log file to look for a database is online message. When confirmed, go to step 5 to start the Configuration database on the next vManage device. After all vManage devices have the Configuration database running, proceed to the next step.

   ```
   tail -f -n 100 /var/log/nms/debug.log
   ```

Note For versions prior to 18.2, use this log file

```
tail -f -n 100 /var/log/nms/vmanage-orientdb-database.log
```

7. Start the Coordination server on each device. Wait for the service to start each time before proceeding to the next vManage device.

   ```
   request nms coordination-server start
   ```

8. Verify that the service is started before proceeding to start it on the next vManage device. After verifying, go to step 7 to start the Coordination server on the next vManage device. After the Coordination server runs on all the vManage devices, proceed to the next step.

   ```
   request nms coordination-server status
   ```

9. Start the Messaging server on each device. Wait for the service to start each time before proceeding to the next vManage device.

   ```
   request nms messaging-server start
   ```

10. Verify that the service has started before proceeding to start it on the next vManage device. After verifying, go to step 9 to start the Messaging server on the next vManage device. After the Messaging server runs on all vManage devices, proceed to the next step.

    ```
    request nms messaging-server status
    ```

11. Start the Application server on each device. Wait for the service to start each time before proceeding to the next vManage device.

    ```
    request nms application-server start
    ```

12. Verify that the service has started before proceeding to start it on the next vManage device. To verify if the service is fully started, open the GUI of that vManage device. After the GUI is fully loaded and you are able to log in, go to step 11 to start the Application server on the next vManage device.

13. Restart the NMS cloud services on each device. Wait for the service to start each time before proceeding to the next vManage device.

    ```
     request nms cloud-agent start
    ```

14. Verify that the service has started before proceeding to start it on the next vManage device. After verifying, go to step 12 to start the cloud services on the next vManage device. After the cloud services run on all vManage devices, proceed to the next step.

    ```
    request nms cloud-agent status
    ```

15. To verify that there are no errors and everything has loaded cleanly, tail the log files.

Check the vManage GUI to verify that all devices appear as online and reachable, and that the statistics exist.

vManage Backups

Cisco manages vManage by taking regular snapshots of the vManage devices for the purpose of recovery due to a catastrophic failure or corruption. The frequency and retention of these snapshots are set for each overlay. Generally, the snapshots are taken daily and retained for up to 10 days. For certain scheduled maintenance activities, such as the upgrade of the vManage devices, another snapshot can be taken before the scheduled activity. In all other cases, it is your responsibility to take regular backups of the vManage configuration database and snapshots of the vManage virtual machine, and follow the example of frequency and retention that is followed by Cisco.

vManage Database Backup

Although the vManage cluster provides high availability and a level of fault tolerance, regular backup of the configuration database should be taken and stored securely off-site. vManage does not have a mechanism of automating the collection of a configuration database backup on a schedule and copying it to another server. The greater the time between the backup and when it is needed for a recovery, the greater the risk that data might be lost. Perform configuration database backups often. Use the following command to create a configuration database backup file.

```
request nms configuration-db backup path <path>
```

vSmart Controller Redundancy

vSmart Controller Redundancy

The vSmart controllers are the central orchestrators of the control plane. They have permanent communication channels with all the Cisco devices in the network. Over the DTLS connections between the vSmart controllers and vBond orchestrators and between pairs of vSmart controllers, the devices regularly exchange their views of the network, to ensure that their route tables remain synchronized. The vSmart controllers pass accurate and timely route information over DTLS connections to Cisco vEdge device .

A highly available Cisco SD-WAN network contains two or more vSmart controllers in each domain. A Cisco SD-WAN domain can have up to 20 vSmart controllers, and each router, by default, connects to two of them. When the number of vSmart controllers in a domain is greater than the maximum number of controllers that a domain's routers are allowed to connect to, the Cisco SD-WAN software load-balances the connections among the available vSmart controllers.

While the configurations on all the vSmart controllers must be functionally similar, the control policies must be identical. This is required to ensure that, at any time, all Cisco vEdge devices receive consistent views of the network. If the control policies are not absolutely identical, different vSmart controllers might give different information to a Cisco vEdge device , and the likely result will be network connectivity issues.



Note To reiterate, the Cisco SD-WAN overlay network works properly only when the control policies on all vSmart controllers are identical. Even the slightest difference in the policies will result in issues with the functioning of the network.

To remain synchronized with each other, the vSmart controllers establish a full mesh of DTLS control connections, as well as a full mesh of OMP sessions, between themselves. Over the OMP sessions, the vSmart controllers advertise routes, TLOCs, services, policies, and encryption keys. It is this exchange of information that allows the vSmart controllers to remain synchronized.

You can place vSmart controllers anywhere in the network. For availability, it is highly recommended that the vSmart controllers be geographically dispersed.

Each vSmart controller establishes a permanent DTLS connection to each of the vBond orchestrators. These connections allow the vBond orchestrators to track which vSmart controllers are present and operational. So, if one of the vSmart controller fails, the vBond orchestrator does not provide the address of the unavailable vSmart controller to a router that is just joining the network.

To reiterate, the Cisco SD-WAN overlay network works properly only when the control policies on all vSmart controllers are identical. Even the slightest difference in the policies result in issues with the functioning of the network.

Recovering from a vSmart Controller Failure

The vSmart controllers are the primary controllers of the network. To maintain this control, they maintain permanent DTLS connections to all the vBond orchestrators and Cisco XE SD-WAN devices and Cisco vEdge devices. These connections allow the vSmart controllers to be constantly aware of any changes in the network topology. When a network has multiple vSmart controllers:

- There is a full mesh of OMP sessions among the vSmart controllers.

- Each vSmart controller has a permanent DTLS connection to each vBond orchestrator.
- The vSmart controllers have permanent DTLS connections to the Cisco XE SD-WAN devices and Cisco vEdge devices. More specifically, each router has a DTLS connection to one of the vSmart controllers.

If one of the vSmart controllers fails, the other vSmart controllers seamlessly take over handling control of the network. The remaining vSmart controllers are able to work with Cisco XE SD-WAN devices and Cisco vEdge devices joining the network and are able to continue sending route updates to the routers. As long as one vSmart controller is present and operating in the domain, the Cisco SD-WAN network can continue operating without interruption.

For more information about configuring graceful restart for OMP, see graceful restart.

Cisco vEdge Device Redundancy

Cisco vEdge Device Redundancy

The Cisco vEdge device s are commonly used in two ways in the Cisco SD-WAN network: to be the Cisco SD-WAN routers at a branch site, and to create a hub site that branch routers connect to.

A branch site can have two Cisco vEdge device s in a branch site for redundancy. Each of the router maintains:

- A secure control plane connection, via a DTLS connection, with each vSmart controller in its domain
- A secure data plane connection with the other routers at the site

Because both the routers receive the same routing information from the vSmart controllers, each one is able to continue to route traffic if one fails, even if they are connected to different transport providers.

When using Cisco XE SD-WAN devices and Cisco vEdge devices in a hub site, you can provide redundancy by installing two Cisco vEdge device s. The branch routers need to connect to each of the hub routers by using separate DTLS connections.

You can also have Cisco vEdge device provide redundancy by configuring up to tunnel interfaces on a single router. Each tunnel interface can go through the same or different firewalls, service providers, and network clouds, and each maintains a secure control plane connection, by means of a DTLS tunnel, with the vSmart controllers in its domain.

Recovering from a Cisco vEdge Device Failure

The route tables on Cisco XE SD-WAN devices and Cisco vEdge devices are populated by OMP routes received from the vSmart controllers. For a site or branch with redundant routers, the route tables on both routers remain synchronized, so if either of the routers fail, the other one continues to be able to route data traffic to its destination.

Configure Affinity between vSmart and Cisco vEdge Device

One way to manage network scale is to configure affinity between vSmart controllers and Cisco vEdge device s. To do this, you place each vSmart controller into a controller group, and then you configure which group or groups a Cisco vEdge device can establish control connections with. The controller groups are what establishes the affinity between vSmart controllers and Cisco vEdge device s.

Configure the Controller Group Identifier on vSmart Controllers

To participate in affinity, each vSmart controller must be assigned a controller group identifier:

```
vSmart(config)#system controller-group-id number
```

The identifier number can be from 0 through 100.

When vSmart controllers are in different data centers, it is recommended that you assign different controller group identifiers to the vSmart controllers. Doing this provides redundancy among data centers, in case a data center becomes unreachable.

For vSmart controllers in the same a data center, they can have the same controller group identifier or different identifiers:

- If the vSmart controllers have the same controller group identifier, a Cisco vEdge device establishes a control connection to any one of them. If that vSmart controller becomes unreachable, the router simply establishes a control connection with another one of the controllers in the data center. As an example of how this might work, if one vSmart controller becomes unavailable during a software upgrade, the Cisco vEdge device immediately establishes a new TLOC with another vSmart controller, and the router's network operation is not interrupted. This network design provides redundancy among vSmart controllers in a data center.
- If the vSmart controllers have different controller group identifiers, a Cisco vEdge device can use one controller as the preferred and the other as backup. As an example of how this might work, if you are upgrading the vSmart controller software, you can upgrade one controller group at a time. If a problem occurs with the upgrade, a Cisco vEdge device establishes TLOCs with the vSmart controllers in the second, backup controller group, and the router's network operation is not interrupted. When the vSmart controller in the first group again becomes available, the Cisco vEdge device switches its TLOCs back to that controller. This network design, while offerring redundancy among the vSmart controllers in a data center, also provides additional fault isolation.

Configure Affinity on Cisco vEdge Device s

For a Cisco vEdge device to participate in affinity, you configure the vSmart controllers that the router is allowed to establish control connections with, and you configure the maximum number of control connections (or TLOCs) that the Cisco vEdge device itself, and that an individual tunnel on the router, is allowed to establish.

Configure a Controller Group List

Configuring the vSmart controllers that the router is allowed to establish control connections is a two-part process:

- At the system level, configure a single list of all the controller group identifiers that are present in the overlay network.
- For each tunnel interface in VPN 0 , you can choose to restrict which controller group identifiers the tunnel interface can establish control connections with. To do this, configure an exclusion list.

At a system level, configure the identifiers of the vSmart controller groups:

```
vEdge(config)#system controller-group-list numbers
```

List the vSmart controller group identifiers that any of the tunnel connections on the Cisco vEdge device might want to establish control connections with. It is recommended that this list contain the identifiers for all the vSmart controller groups in the overlay network.

If, for a specific tunnel interface in VPN 0 , you want it to establish control connections to only a subset of all the vSmart controller groups, configure the group identifiers to exclude:

```
vEdge(config-vpn-0-interface)#tunnel-interface exclude-controller-group-list numbers
```

Or

In this command, list the identifiers of the vSmart controller groups that this particular tunnel interface should never establish control connections with. The controller groups in this list must be a subset of the controller groups configured with the **system controller-group-list** command.

To display the controller groups configured on a Cisco vEdge device, use the **show control connections** command.

Configure the Maximum Number of Control Connections

Configuring the maximum number of control connections for the Cisco vEdge device is a two-part process:

- At the system level, configure the maximum number of control connections that the Cisco vEdge device can establish to vSmart controllers.
- For each tunnel interface in VPN 0 , configure the maximum number of control connections that the tunnel can establish to vSmart controllers.

By default, a Cisco vEdge device can establish two OMP sessions for control connections to vSmart controllers. To modify the maximum number of OMP sessions:

```
vEdge(config)#system max-omp-sessions number
```

The number of OMP sessions can be from 0 through 100.

A Cisco vEdge device establishes OMP sessions as follows:

- Each DTLS and and each TLS control plane tunnel creates a separate OMP session.
- It is the Cisco vEdge device as a whole, not the individual tunnel interfaces in VPN 0 , that establishes OMP sessions with vSmart controllers. When different tunnel interfaces on the router have affinity with the same vSmart controller group, the Cisco vEdge device creates a single OMP session to one of the vSmart controllers in that group, and the different tunnel interfaces use this single OMP session.

By default, each tunnel interface in VPN 0 can establish two control connections. To change this:

```
vEdge(config-vpn-0-interface)#vpn 0 interface interface-name tunnel-interface
max-control-connections number
```

The number of control connections can be from 0 through 100. The default value is the maximum number of OMP sessions configured with the **system max-omp-sessions** command.

When a Cisco vEdge device s has multiple WAN transport connections, and hence has multiple tunnel interfaces in VPN 0 , the sum of the maximum number of control connections that all the tunnels can establish cannot exceed the maximum number allowed on the router itself.

To display the maximum number of control connections configured on an interface, use the **show control local-properties** command.

To display the actual number of control connections for each tunnel interface, use the **show control affinity config** command.

To display a list of the vSmart controllers that each tunnel interface has established control connections with, use the **show control affinity status** command.

Best Practices for Configuring Affinity

- In the **system controller-group-list** command on the Cisco vEdge device , list all the controller groups available in the overlay network. Doing so ensures that all the vSmart controllers in the overlay network are available for the affinity configuration, and it provides additional redundancy in case connectivity to the preferred group or groups is lost. You manipulate the number of control connections and their priority based on the maximum number of OMP sessions for the router, the maximum number of control connections for the tunnel, the controller groups a tunnel should not use. A case in which listing all the controller groups in the **system controller-group-list** command provides additional redundancy is when the Cisco vEdge device site is having connectivity issues in reaching the vSmart controllers in the controller group list. To illustrate this, suppose, in a network with three controller groups (1, 2, and 3), the controller group list on a Cisco vEdge device contains only groups 1 and 2, because these are the preferred groups. If the router learns from the vBond controller that the vSmart controllers in groups 1 and 2 are up, but the router is having connectivity issues to both sites, the router loses its connectivity to the overlay network. However, if the controller group list contains all three controller groups, even though group 3 is not a preferred group, if the router is unable to connect to the vSmart controllers in group 1 or group 2, it is able to fall back and connect to the controllers in group 3. Configuring affinity and the order in which to connect to vSmart controllers is only a preference. The preference is honored whenever possible. However, the overarching rule in enforcing high availability on the overlay network is to use any operational vSmart controller. The network ceases to function only when no vSmart controllers are operational. So it might happen that the least preferred vSmart controller is used if it is the only controller operational in the network at a particular time. When a Cisco vEdge device boots, it learns about all the vSmart controllers in the overlay network, and the vBond orchestrator is continuously communicating to the router which vSmart controllers are up. So, if a Cisco vEdge device cannot reach any of the preferred vSmart controllers in the configured controller group and another vSmart controller is up, the router connects to that controller. Put another way, in a network with multiple vSmart controllers, as a last resort, a Cisco vEdge device connects to any of the controllers, to ensure that the overlay network remains operational, whether or not these controllers are configured in the router's controller group list.
- The controller groups listed in the **exclude-controller-group-list** command must be a subset of the controller groups configured for the entire router, in the **system controller-group-list** command.
- When a data center has multiple vSmart controllers that use the same controller group identifier, and when the overlay network has two or more data centers, it is recommended that the number of vSmart controllers in each of the controller groups be the same. For example, if Data Center 1 has three vSmart controllers, all with the same group identifier (let's say, 1), Data Center 2 should also have three vSmart controllers, all with the same group identifier (let's say, 2), and any additional data centers should also have three vSmart controllers.
- When a data center has vSmart controllers in the same controller group, the hardware capabilities—specifically, the memory and CPU—on all the vSmart controllers should be identical. More broadly, all the vSmart controllers in the overlay network, whether in one data center or in many, should have the same hardware capabilities. Each vSmart controller should have equal capacity and capability to handle a control connection from any of the Cisco vEdge device s in the network.
- When a router has two tunnel connections and the network has two (or more) data centers, it is recommended that you configure one of the tunnel interfaces to go to one of the data centers and the other to go to the second. This configuration provides vSmart redundancy with the minimum number of OMP sessions.
- Whenever possible in your network design, you should leverage affinity configurations to create fault-isolation domains.

Configure Control Plane and Data Plane High Availability Parameters

This topic discusses the configurable high availability parameters for the control plane and the data plane.

Control Plane High Availability

A highly available Cisco SD-WAN network contains two or more vSmart controllers in each domain. A Cisco SD-WAN domain can have up to 20 vSmart controllers, and each Cisco vEdge device , by default, connects to two of them. You change this value on a per-tunnel basis:

```
vEdge(config-tunnel-interface)# max-control-connections number
```



Note The **no-control-connections** command is not applicable. Instead, use the **max-control-connections** command.

When the number of vSmart controllers in a domain is greater than the maximum number of controllers that a domain's Cisco vEdge devices are allowed to connect to, the SD-WAN software load-balances the connections among the available vSmart controllers.



Note To maximize the efficiency of the load-balancing among vSmart controllers, use sequential numbers when assigning system IP addresses to the Cisco vEdge devices in the domain. One example of a sequential numbering schemes is 172.1.1.1, 172.1.1.2, 172.1.1.3, and so forth. Another is 172.1.1.1, 172.1.2.1, 172.1.3.1, and so forth.

Data Plane High Availability

BFD, which detects link failures as part of the Cisco SD-WAN high availability solution, is enabled by default on all Cisco devices. BFD runs automatically on all IPsec data tunnels between Cisco vEdge devices. It does not run on the control plane (DTLS or TLS) tunnels that vSmart controllers establish with all Cisco devices in the network.

You can modify the BFD Hello packet interval and the number of missed Hello packets (the BFD interval multiplier) before BFD declares that a link has failed.

Change the BFD Hello Packet Interval

BFD sends Hello packets periodically to detect faults on the IPsec data tunnel between two Cisco vEdge devices. By default, BFD sends these packets every 1000 milliseconds (that is, once per second). To change this interval on one or more traffic flow, use the **hello-interval** command:

```
vEdge(config)#bfd color color hello-interval milliseconds
```

The interval can be a value from 100 to 300000 milliseconds (5 minutes).

Configure the interval for each tunnel connection, which is identified by a color. The color can be **3g**, **biz-internet**, **blue**, **bronze**, **custom1**, **custom2**, **custom3**, **default**, **gold**, **green**, **lte**, **metro-ethernet**, **mpls**, **private1**, **private2**, **public-internet**, **red**, or **silver**.

Change the BFD Packet Interval Multiplier

After BFD has not received a certain number of Hello packets on a link, it declares that the link has failed. This number of packets is a multiplier of the Hello packet interval time. By default, the multiplier is 7 for hardware routers and 20 for Cloud software routers. This means that if BFD has not received a Hello packet after 7 seconds, it considers that the link has failed and implements its redundancy plan.

To change the BFD packet interval multiplier, use the **multiplier** command:

```
vEdge(config)#bfd color color multiplier integer
```

Multiplier range: 1 to 60 (integer)

You configure the multiplier for each tunnel connection, which is represented by a color.

Control PMTU Discovery

On each transport connection (that is, for each TLOC, or color), the Cisco SD-WAN BFD software performs path MTU (PMTU) discovery, which automatically negotiates the MTU size in an effort to minimize or eliminate packet fragmentation on the connection. BFD PMTU discovery is enabled by default, and it is recommended that you use BFD PMTU discovery and not disable it. To explicitly enable it:

```
vEdge(config)#bfd color color pmtu-discovery
```

With PMTU discovery enabled, the path MTU for the tunnel connection is checked periodically, about once per minute, and it is updated dynamically. With PMTU discovery enabled, 16 bytes might be required by PMTU discovery, so the effective tunnel MTU might be as low as 1452 bytes. From an encapsulation point of view, the default IP MTU for GRE is 1468 bytes, and for IPsec it is 1442 bytes because of the larger overhead. Enabling PMTU discovery adds to the overhead of the BFD packets that are sent between the Cisco vEdge devices, but does not add any overhead to normal data traffic.

If PMTU discovery is disabled, the expected tunnel MTU is 1472 bytes (tunnel MTU of 1500 bytes less 4 bytes for the GRE header, 20 bytes for the outer IP header, and 4 bytes for the MPLS header). However, the effective tunnel MTU might be 1468 bytes, because the software might sometimes erroneously add 4 bytes to the header.

Configure Disaster Recovery

Table 16: Feature History

Feature Name	Release Information	Feature Description
Disaster Recovery for Cisco vManage	Cisco SD-WAN Release 19.2.1 Cisco vManage Release 19.2.1	This feature helps you configure Cisco vManage in an active or standby mode to counteract hardware or software failures that may occur due to unforeseen circumstances.

You want to deploy the Cisco SD-WAN controllers across two data centers, and if a data center goes down due to a disaster, you want the network to be available. Out of the three controllers that make up theCisco SD-WAN solution, vManage is the only one that is stateful and cannot be deployed in an active/active mode. The goal of the disaster recovery solution is to deploy vManage across two data centers in some sort of primary/secondary mode.

The disaster recovery option provides automatic failover of the primary cluster to the secondary cluster. Data is replicated from the primary cluster to the secondary cluster.

There are two available disaster recovery options:

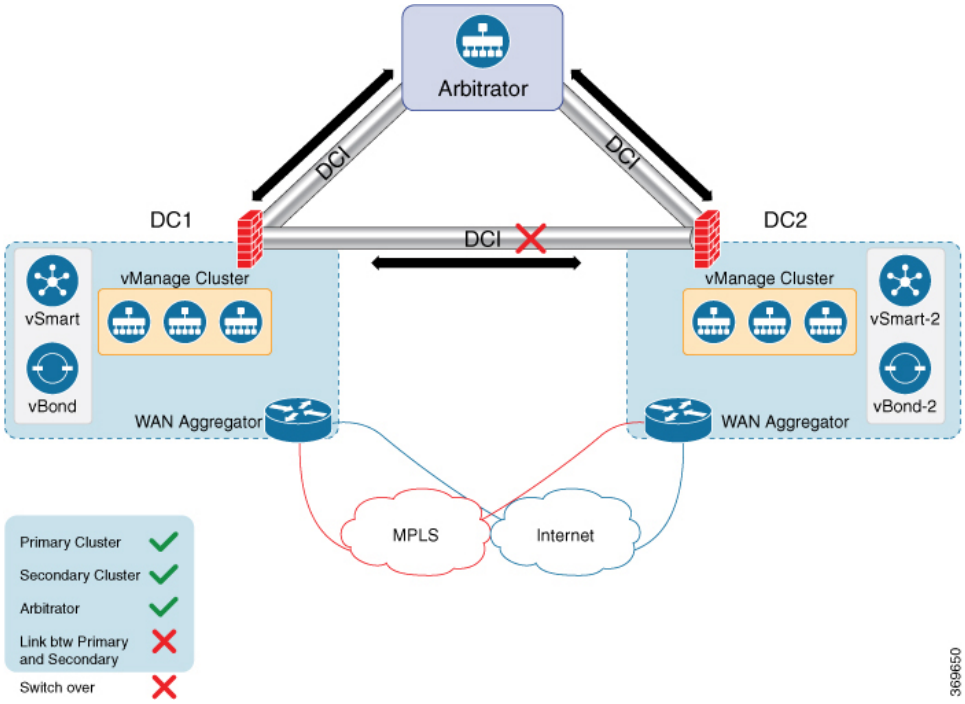
- Manual—If you want to make the clusters active, you can do it manually rather than having the arbitrator do the switchover. You can specify the switchover threshold.
- Automated —Arbitrator does the monitoring of the cluster and performs the necessary action.

  A highly available Cisco SD-WAN network contains three or more vManage NMSs in each domain. This scenario is referred to as a cluster of vManage NMSs, and each vManage NMS in a cluster is referred to as a vManage instance.

Architecture Overview

The following diagram describes the high-level architecture of the disaster recovery solution.

The arbitrator is an additional vManage cluster that runs in arbitrator mode. The arbitrator monitors the health of the primary and the secondary clusters and performs the necessary actions.



Prerequisites

Prior to configuring disaster recovery, make sure you have met the following requirements:

- You must have two vManage clusters with three nodes in each cluster. If automated recovery option is selected, then another vManage node is required.
- You must be able to reach the primary and the secondary cluster using HTTPS on a transport VPN (VPN 0).
- Make sure that vSmart and vBond devices on the secondary cluster are connected to the primary cluster.

Best Practices and Recommendations

- Ensure that you use a netadmin user privilege for Disaster Recovery registration. We recommend that you modify the factory-default password, admin before you start the registration process.
- To change user credentials, we recommend that you use the Cisco vManage GUI, and not use the CLI of a Cisco SD-WAN device.
- If Cisco SD-WAN devices are configured using feature templates, ensure that you create separate feature templates for both primary data center and secondary data center.
- When primary cluster is switched over to the secondary cluster, Cisco vManage detaches the Cisco SD-WAN devices from the feature templates. Therefore, ensure that you reattach the devices to the specific feature templates.
- For an on-premises deployment, ensure that you regularly take backup of the Configuration database from the active Cisco vManage instance.

Changing the Cisco vManage or Cisco vBond Orchestrator Administrator Password

If you need to change the administrator password for Cisco vManage or Cisco vBond Orchestrator, first change the password, then deregister disaster recovery from the Cisco vManage cluster, and then re-register disaster recovery on the cluster.

Enable Disaster Recovery on Day-0:

You need to bring up two separate clusters with no devices being shared, which means do not share any vSmart, vBond, or vManage devices.

On both clusters, configure the following:

Item	Action
Secondary cluster	Bring up the secondary vManage cluster with three vManage clusters.
Arbitrator	To assign an IP address for the OOB network, navigate to **Administration > Cluster Management**.
	Ensure reachability between the primary, secondary clusters, and arbitrator on VPN (0) using HTTPS.
	Ensure reachability between the primary cluster, secondary cluster, and vBond orchestrators.

Verify after Registering for Disaster Recovery on Day-1

- Replication from the primary cluster to the secondary cluster happens at the configured intervals.
- Status check: **Administration > Disaster Recovery**.
- Arbitrator:
  - First health check after 15 minutes. This check provides enough time for all the nodes to be up and running with the configured disaster recovery processes.

- Health check of the primary cluster, secondary cluster, and the arbitrator every five minutes.
- Check the */var/log/nms/vmanage-server.log* for the status information on the arbitrator cluster.

Configure Disaster Recovery

1. From the Cisco vManage dashboard, select **Administration > Disaster Recovery**.
2. On the **Administration > Disaster Recovery** page, select **Manage Disaster Recovery**.
3. To configure primary and secondary cluster, on the vManage Disaster Recovery screen, select an IP address for any vManage node within the respective cluster.

   If a cluster is behind a load balancer, specify the IP address of the load balancer.
4. Specify the following: **Start Time**, **Replication Interval**, and **Delay Threshold** for replicating data from the primary to the secondary cluster.

   The default value for **Delay Threshold** is 30 minutes.

   The default value for **Replication Interval** is 15 minutes.
5. Click **Administration > Disaster Recovery**, and for Cluster 2 (Secondary), click **Make Primary**.

   It can take 10 to 15 minutes to push all changes from all the devices.
6. You can also decide to pause disaster recovery, pause replication, or delete your disaster recovery configuration.

   After disaster recovery is configured and you have replicated data, you can view the following:
   - when your data was last replicated, how long it took to replicate, and the size of the data that was replicated.
   - when the primary cluster was switched over to the secondary cluster and the reason for the switchover.
   - the replication schedule and the delay threshold.

Disaster Recovery Striking the Primary Data Center

- Switchover happens only when all the nodes in the primary data center are lost.
- The arbitrator detects the loss of all the primary data center members and initiates switchover to the secondary data center.
- Secondary data center updates the vBond:
  - Invalidates old Cisco vManage systems.
  - New Cisco vManage systems from the secondary data center are updated, as valid.
  - Routers reach vBond after losing control connections.
  - Routers start forming control connections with the new valid Cisco vManage systems.

Troubleshooting Tips

If disaster recovery registration fails, verify the following:

- Reachability to the vBond orchestrator from all cluster members on the secondary cluster.
- Reachability between the secondary cluster, primary cluster, and the arbitrator on the transport interface (VPN 0).
- Check that you have the correct username and password.

If disaster recovery registration fails due to arbitrator reachability, check the following:

- You must configure the arbitrator in cluster mode. Navigate to **Administration > Cluster Management**, and add a Cisco vManage as the arbitrator.
- If the IP address is not assigned to the correct arbitrator, log on to the arbitrator cluster and do the following:
  - Navigate to **Administration > Cluster Management**.
  - Edit the Cisco vManage.
  - Select the correct IP address from the drop-down list and save the configuration.

  The disaster recovery consul process uses this IP address for disaster recovery communication. This is set once you configure the Cisco vManage in cluster mode.

Disaster Recovery with Manual Switchover

This section provides information about setting up and registering disaster recovery in a Cisco SD-WANdeployment, and about performing a manual switchover. Disaster recovery has been validated for a three-node cluster

Prerequisites for Setting up Disaster Recovery with Manaual Switchover

Before you set up disaster recovery for your SD-WAN deployment, perform the following tasks:

- Configure an out-of-band or cluster interface on the VPN 0 of each Cisco vManage node that is to be used for disaster recovery. This interface is the same one that Cisco vManage uses to communicate with its peers in a cluster.
- Make sure that all Cisco vManage nodes can reach each other through the out-of-band interface.
- Make sure that all services (application-server, configuration-db, messaging server, coordination server, and statistics-db) are enabled on all Cisco vManage nodes in the cluster.
- Make sure that all Cisco vManage nodes in a cluster reside on the same LAN segment.
- Make sure that all Cisco vManage nodes are running same Cisco vManage software version.
- To allow Cisco vManage clusters to communicate with each other across data centers, enable TCP ports 8443 and 830 on your data center firewalls.
- Spin all controllers, including Cisco vBond Orchestrators, across both primary and secondary data centers. Ensure that these controllers are reachable by Cisco vManage nodes that are spun across these data centers. The controllers connect only to the primary Cisco vManagecluster.

- Distribute each Cisco vManage VM on a separate physical server so that a single physical server outage does not affect the Cisco vManage cluster in a data center.

Disaster Recovery Registration

Disaster Recovery must be registered on the primary Cisco vManage cluster. Before you start the registration process, make sure that no other operations in process in the active (primary) and the standby (secondary)Cisco vManage cluster. For example, make sure that no servers are in the process of upgrading or no templates are in the process of attaching templates to devices.

You can use the out-of-band IP address of a reachable Cisco vManage node in the cluster for disaster recovery registration.

Before you start the registration process, go to the **Tools** > **Rediscover Network** page on the primary Cisco vManage node and rediscover the Cisco vBond Orchestrators.

Disaster recovery registration is a day 1 operation. The registration can take up to 30 minutes to complete. After the registration starts, the message "No Data Available" may display for a short time in the Disaster Registration Task View. During the registration process, the message "In-progress" displays.

If you see the message "Error occurred retrieving status for action disaster_recovery_registration," click the **Reload** button in your browser after the last active Cisco vManage node restarts.

If you need to upgrade your Cisco vManage software in the future, pause disaster recovery, perform the upgrade, and then resume disaster recovery. When upgrading Cisco vManage, follow the best practices as described in Cisco SD-WAN vManageCluster Creation and Troubleshooting.

Scheduled Disaster Recovery

Performing a manually scheduled switchover let you test the operation of disaster recovery.

Detach templates from Cisco vManage devices in the primary cluster before you perform a switchover.

To manually perform a scheduled switchover, follow these steps:

1. Shut off the tunnel interfaces on the primary Cisco vManage cluster to prevent devices from toggling during the switchover.
2. From a Cisco vManage system on the secondary cluster, select **Administration** > **Disaster Recovery**.
3. Wait for data replication to complete, then click **Make Primary**.

   Devices and controllers converge to the secondary cluster and that cluster assumes the role of the primary cluster. When this process completes, the original primary cluster assumes the role of the secondary cluster. Then data replicates from the new primary cluster to the new secondary cluster.

To move back to the original primary cluster, repeat these steps.

Disaster Recovery Operations

This sections explains how to perform disaster recovery in a variety of situations.

Loss of Primary Cisco vManage Cluster

If your primary Cisco vManage cluster goes down, follow these steps for disaster recovery:

1. From a Cisco vManage system on the secondary cluster, select **Administration** > **Disaster Recovery**.
2. Click **Make Primary**.

Devices and controllers converge to the secondary cluster and that cluster assumes the role of the primary cluster.

When the original primary cluster recovers and is back on line, it assumes the role of the secondary cluster and begins to receive data from the primary cluster.

Loss of Primary Data Center

If your primary data center cluster goes down, follow these steps for disaster recovery:

1. From a Cisco vManage system on the secondary cluster, select **Administration** > **Disaster Recovery**.
2. Click **Make Primary**.

   The switchover process begins. During the process, only the Cisco vBond Orchestrators in the secondary data center are updated with a new valid Cisco vManage list. Devices and controllers that are on line converge to the secondary cluster and that cluster assumes the role of the primary cluster.

   After the original primary data center recovers and all VMs, including controllers, are back on line, the controllers are updated with a new valid Cisco vManage and converge to the new primary Cisco vManage cluster. The original primary cluster assumes the role of secondary cluster and begins to receive data from the primary cluster.

Partial Loss of Primary Cisco vManage Cluster

If you experience a partial loss of the primary Cisco vManage cluster, we recommend that you try to recover that cluster instead of switching over to the secondary cluster.

A cluster with N nodes is considered to be operational if (N/2)+1 nodes are operational.

A cluster with N nodes becomes read only if (N/2)+1 or more nodes are lost.

Loss of Enterprise Network Between Data Centers

If a link failure occurs between your data centers but the WAN in the primary data center is operational, data replication fails. In this situation, attempt to recover the link so that data replication can resume.

To avoid a possible split brain scenario, do not perform a switchover operation.

Delete Disaster Recovery

If you want to delete disaster recovery, we recommend that you initiate the delete operation on the primary cluster. Before deleting, make sure that there is no data replication session in pending state, and make sure that the secondary cluster is not importing data.

If the primary Cisco vManage cluster is down, you can perform the delete operation on the secondaryCisco vManage cluster.

If any Cisco vManage cluster that was offline during the disaster recovery delete operation come on line, execute the following POST request on that cluster to complete the delete disaster recovery operation:

POST /dataservice/disasterrecovery/deleteLocalDC

After you delete disaster recovery, makes sure that the primary and secondary clusters are operating correctly. To do so, go to the **Administration** > **Cluster Management** page and make sure that all Cisco vManage nodes are present in the cluster. If the nodes are not present, restart the application server. Also go to the **Administration** > **Disaster Recovery** page and make sure that no nodes appear.

Data centers must be deleted from disaster recovery before you can reregister disaster recovery for the data centers.

Guidelines and Best Practices

- The disaster recovery functionality does not replace the best practices of taking database backups and VM snapshots on the primary Cisco vManage cluster. We recommend that you to take these backups and snapshots.
- In some situations, The **Administration** > **Disaster Recovery** page displays the message "Disaster recovery not configured." This message represents the transient issue data replication occurring on the secondary cluster.
- The time that it takes for devices to converge to new primary Cisco vManage cluster after a switchover depends on the number of devices that are involved in the switchover. For example, a switchover of 10 devices might take less than 30 seconds, but the switchover of 100 devices can take few minutes.
- After a switchover, the old primary cluster requires reachability to the new primary cluster to update the states of both clusters. It can take up to 5 minutes for the update to complete.
- After you complete the registration process, wait for the first data replication cycle between clusters to complete before you perform any action on the primary Cisco vManage cluster. You can verify the replication status in the **Administration** > **Disaster Recovery** page.
- After switching over, do not update disaster recovery settings until the first data replication cycle between clusters completes. You can verify the replication status in the **Administration** > **Disaster Recovery** page.
- To avoid a split brain scenario, do not perform a make-primary operation from a secondary data center when the tunnel interfaces of the Cisco vManage nodes in the primary data center are up and accessible from other controllers and devices. Bring down the tunnel interfaces for the primary Cisco vManage cluster before you perform the make primary operation for the secondary cluster.
- If the **Make Primary** button in the **Administration** > **Disaster Recovery** page becomes dim after you click it, click the **Reload** button in your browser.
- In a deployment in which Cisco vManage acts as the certificate authority (CA) for Cisco edge devices, devices that you add to the overlay network after performing a switchover from one data center to another do not join the overlay network after you switch back. In this situation, manually synchronize root certificates between the data centers after you perform the first switchover.
- Periodically check the Cisco vBond Orchestrator information on the Cisco vManage Dashboard in the secondary cluster. If you see an issue, go to the **Tools** > **Rediscover Network** page on the primary Cisco vManage node and rediscover the Cisco vBond Orchestrators. The updated discovery information replicates to the secondary Cisco vManage cluster during the next replication cycle.
- After a switchover completes, review the replication information in the **Administration** > **Disaster Recovery** page to ensure that data is transferred from the primary cluster to the secondary cluster.
- If replication fails, verify that the primary cluster can reach the secondary cluster.
- For related disaster recovery troubleshooting tips and information, see the "High Availability Overview" chapter in Network Optimization and High Availability Configuration Guide.

High Availability CLI Reference

CLI commands for configuring and monitoring high availability.

High Availability Configuration Commands

Use the following commands to configure high availability on a Cisco vEdge device:

```
bfd
  app-route
    multiplier number
    poll-interval milliseconds
  color color
    hello-interval milliseconds
    multiplier number
    pmtu-discovery
```

High Availability Monitoring Commands

show bfd sessions—Display information about the BFD sessions running on the local Cisco vEdge device.

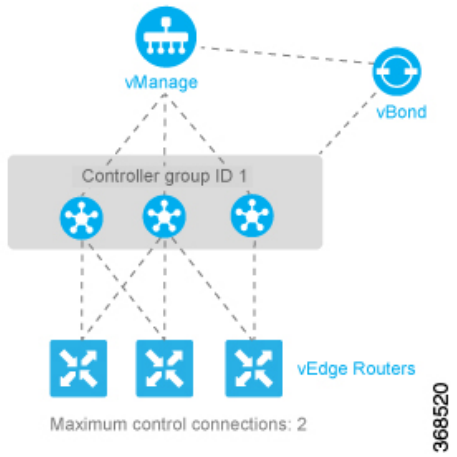
High Availability Configuration Examples

This topic provides examples of configuring high availability, specifically, of configuring affinity between vSmart controllers and Cisco vEdge device .

Configure Affinity to vSmart Controllers in a Single Data Center

In an overlay network with a single data center that has multiple vSmart controllers, if you want the Cisco vEdge device to establish a single control connection to one of vSmart controllers, there is no need to configure affinity because this situation is the default behavior.

However, if you want the Cisco vEdge device to establish control connections to more than one vSmart controllers, to provide redundancy in case one of the controllers becomes unavailable, you configure affinity. You generally place the vSmart controllers in the same controller group.



Let's say that all the vSmart controllers use the same controller group identifier, 1. You configure the identifier on all three controllers as follows:

```
vSmart(config)# system controller-group-id 1
```

To verify the configuration, use the **show running-config** command:

```
vSmart# show running-config system
system
```

```
 description         "vSmart in data center 1"
 host-name           vSmart
 gps-location latitude 37.368140
 gps-location longitude -121.913658
 system-ip           172.16.255.19
 site-id             100
 controller-group-id 1
 organization-name   "Cisco"
 clock timezone America/Los_Angeles
```

We want the three Cisco vEdge device s to establish two control connections to two of the three vSmart controllers. We do this for purposes of redundancy, in case one of the controllers becomes available. Because all the vSmart controllers are in the same controller group, we cannot specify or influence which of the two controllers the Cisco vEdge device s connect to. The configurations on all three routers are effectively identical. We show here the configuration for router vEdge -1.

First, configure the available vSmart controller groups. This scenario has just one group:

```
vEdge-1(config)# system controller-group-list 1
```

By default, a Cisco vEdge device can establish two control connections. Because we want each Cisco vEdge device and each tunnel interface to connect to two vSmart controllers, no configuration is required here. However, if you want to explicitly configure these parameters, you configure the maximum number of OMP sessions at the system level and the maximum number of control connections per tunnel:

```
vEdge-1(config)# system max-omp-sessions 2
vEdge-1(config)# vpn 0 interface ge0/2 tunnel-interface
vEdge-1(config-tunnel-interface)# max-control-connections 2
```

Here are the relevant configuration snippets from Cisco vEdge device-1:

```
vEdge-1# show running-config system
system
 host-name             vEdge-1
 gps-location latitude 43.0
 gps-location longitude -75.0
 system-ip             172.16.255.11
 site-id               100
 max-omp-sessions      2
 controller-group-list 1
 organization-name     "Cisco"
 ...
vEdge-1# show running-config vpn 0
...
 interface ge0/2
  ip address 10.0.5.11/24
  tunnel-interface
   encapsulation ipsec
   color lte
   max-control-connections 1
   no allow-service bgp
   allow-service dhcp
   allow-service dns
   allow-service icmp
   no allow-service sshd
   no allow-service netconf
   no allow-service ntp
   no allow-service ospf
   no allow-service stun
  !
  no shutdown
...
```

To display the control connections with the vSmart controllers, use the **show control connections** command. The last column, Controller Group ID, lists the vSmart controller group that a router is in.

```
vEdge-1# show control connections

                                                                    PEER         PEER                                          CONTROLLER
PEER     PEER     PEER            SITE    DOMAIN   PEER           PRIVATE  PEER         PUBLIC                                GROUP
TYPE     PROTOCOL SYSTEM IP       ID      ID       PRIVATE IP     PORT     PUBLIC IP    PORT    LOCAL COLOR   STATE    UPTIME     ID
-------------------------------------------------------------------------------------------------------------------------------------
vsmart   dtls     172.16.255.19    100     1        10.0.5.19      12446    10.0.5.19    12446   lte           up       0:00:00:53 1
vsmart   dtls     172.16.255.20    200     1        10.0.12.20     12446    10.0.12.20   12446   lte           up       0:00:00:22 1
```

To display the maximum number of control connections allowed on the router, use the **show control local-properties** command. The last line of the output lists the maximum controllers. The following is the abbreviated output for this command:

```
vEdge-1# show control local-properties

personality                 vedge
organization-name           Cisco
certificate-status          Installed
root-ca-chain-status        Installed

certificate-validity        Valid
certificate-not-valid-before Mar 10 19:50:04 2016 GMT
certificate-not-valid-after  Mar 10 19:50:04 2017 GMT
...

                       PUBLIC          PUBLIC PRIVATE         PRIVATE                                PRIVATE                              MAX
 RESTRICT/           LAST         SPI
TIME    NAT  VM
INTERFACE              IPv4            PORT   IPv4            IPv6                                   PORT    VS/VM COLOR            STATE CNTRL
 CONTROL/     LR/LB  CONNECTION
REMAINING   TYPE CON

 STUN
             PRF
-------------------------------------------------------------------------------------------------------------------------------------------------------------
-------------------
ge0/0                  2.1.1.11        12346  2.1.1.11        ::                                     12346    2/1  default          up     2
   no/yes/no   No/No  2:19:35:21
0:07:58:46  N    5
```

Two commands display information about the control connections established by the affinity configuration. To see, for each interface, which controller groups are configured and which the interface is connected to, use the **show control affinity config** command:

```
vEdge-1# show control affinity config

EFFECTIVE CONTROLLER LIST FORMAT - G(C),...  - Where G is the Controller Group ID
                                                 C is the Required vSmart Count
CURRENT CONTROLLER LIST FORMAT - G(c)s,...   - Where G is the Controller Group ID
                                                 c is the current vSmart count
                                                 s Status □ when matches, χ when does not match
                 EFFECTIVE
                 REQUIRED
INDEX INTERFACE  VS COUNT   EFFECTIVE CONTROLLER LIST    CURRENT CONTROLLER LIST  EQUILIBRIUM
-----------------------------------------------------------------------------------------------
0     ge0/2      2           1(2)                       1(2)□                    Yes
```

The command output above shows that affinity is configured on interface ge0/2.

- The **Effective Required** and **Count** column shows that the interface is configured to create two control connections, and, in fact, two control connections have been established. You configure the number of control connections for the tunnel interface with the **max-control-connections** command.
- The Effective Controller List column shows that affinity on the interface is configured to use Cisco vSmart Controller identifier 1 and that the router supports two OMP sessions. You configure the affinity controller identifiers with the **controller-group-list** command (at the **system** level) and, for the tunnel interface, the **exclude-controller-group-list** command.

- The Current Controller List column lists the actual affinity configuration for the interface. The output here shows that the interface has two control connections with Cisco vSmart Controllers in group 1. The check mark indicates that the current and effective controller lists match each other. If, for example, the tunnel had established only one TLOC connection to a vSmart controller, this column would show "1(1)X".
- The Equilibrium column indicates that the current controller lists matches what is expected from the affinity configuration for that tunnel interface.

To determine the exact Cisco vSmart Controllers that the tunnel interface has established control connections with, use the **show control affinity status** command:

```
vEdge-1# show control affinity status

ASSIGNED CONNECTED CONTROLLERS - System IP( G),..       - System IP of the assigned vSmart
                                                             G is the group ID to which the vSmart belongs to

UNASSIGNED CONNECTED CONTROLLERS - System IP( G),..     - System IP of the unassigned vSmart
                                                             G is the group ID to which the vSmart belongs to

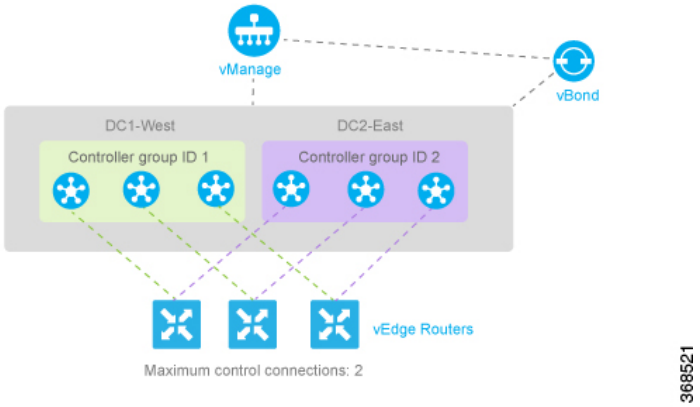
INDEX INTERFACE ASSIGNED CONNECTED  CONTROLLERS        UNASSIGNED CONNECTED CONTROLLERS
-----------------------------------------------------------------------------------------------------------------
0     ge0/2     172.16.255.19(  1), 172.16.255.20(  1)
```

The command output above shows that interface **ge0/2** has control connections to two vSmart controllers, 172.16.255.19 and 172.16.255.20, that both controllers are in group 1, and that both controllers are in one of the groups configured in the controller group list. If the interface were connected to a vSmart controller not in the controller group list, it would be listed in the Unassigned Connected Controllers column.

When a data center has multiple vSmart controllers, you can configure them to be in different controller groups. For example, if you configure them to be in two different controller groups, each Cisco vEdge device can establish two control connections, one to each of the groups. While this configuration design is similar to what we discussed in the previous section, providing redundant control connections to the vSmart controllers, on subtle difference is that it provides fault isolation between the twoCisco vSmart Controller groups in the data center. The configuration for this scenario is almost identical to the configuration when Cisco vSmart Controllers are two data centers. The only difference is that here, two Cisco vSmart Controller groups are collocated in the same data center. See the configuration example in the next section.

Configure Affinity to vSmart Controllers in Two Data Centers

You can use affinity to enable redundancy among data centers, for a network design in which multiple Cisco vSmart Controllers are spread across two or more data centers. Then, if the link between a Cisco vEdge device and one of the data centers goes down, the Cisco vSmart Controllers in the second data center are available to continue servicing the overlay network. The figure below illustrates this scenario, showing three Cisco vSmart Controllers in each of two data centers. Each of the three Cisco vEdge device s establishes a TLOC connection to one controller in the West data center and one in the East data center.



You configure the three vSmart controllers in DC1-West with controller group identifier 1:

```
vSmart-DC1(config)# system controller-group-id 1
```

The three vSmart controllers in DC2-East are in controller group 2:

```
vSmart-DC2(config)# system controller-group-id 2
```

We want all the Cisco vEdge device s to have a maximum of two OMP sessions, and we want each tunnel interface to have a maximum of two control connections and to not exclude any controller groups. So the only configuration that needs to be done on the routers is to set the controller group list. We want Cisco vEdge device s in the west to prefer Cisco vSmart Controllers in DC1-West over DC2-East:

```
vEdge-West(config)# system controller-group-list 1 2
```

Similarly, we want Cisco vEdge device s in the east to prefer DC2-East:

```
vEdge-East(config)# system controller-group-list 2 1
```

The software evaluates the controller group list in order, so with this configuration, the vEdge-West routers prefer vSmart controller group 1 (which is the West data center), and the vEdge-East routers prefer Cisco vSmart Controller group 2.

You can fine-tune the controller group preference in other ways:

- Set the maximum number of OMP sessions allowed on the router to 1 (**system max-omp-sessions 1**). To illustrate how this works, let's look at a vEdge-West router. The router has only one tunnel interface, and that interface creates one control connection to Cisco vSmart Controller list 1. If all the Cisco vSmart Controllers in this group become unavailable, or if the connection between the router that the DC1-West data center goes down, the tunnel interface establishes one control connection to Cisco vSmart Controller list 2, because this group is listed in the **system controller-group-list** command. If all Cisco vSmart Controllers in both controller groups, or the connections to them, become unavailable, and if the vBond orchestrator also indicates that all these vSmart controllers are unreachable, the tunnel interface establishes a control connection to any other Cisco vSmart Controller in the overlay network if other controllers are present.

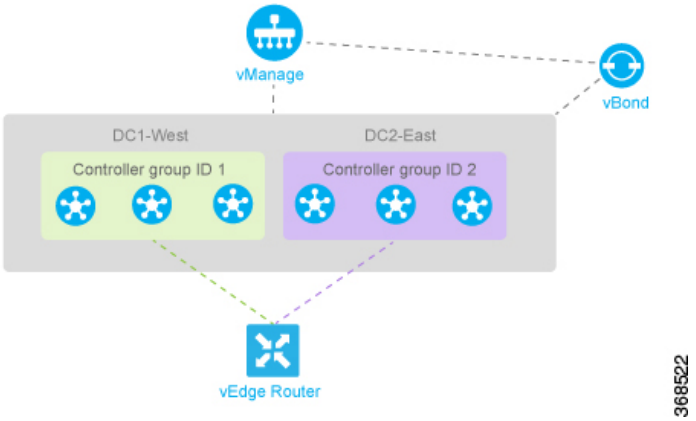
- Set the maximum number of control connections that the tunnel interface can establish to 1 (**vpn 0 sdwan interface tunnel-interface max-control-connections 1**). Because the software evaluates the controller group list in order, for a vEdge-West router, this configuration forces the tunnel interface to establish a control connection to Cisco vSmart Controller group 1. Again, if this controller group or data center becomes unreachable, the tunnel establishes a control connection with controller group 2, because this group is configured in the **system controller-group-list** command. And if neither controller group 1

or 2 is available, and if another vSmart controller is present in the network, the tunnel interface establishes a control connection with that controller.

- Exclude the non-preferred Cisco vSmart Controller group for a particular tunnel. For example, for a vEdge-West router to prefer controller group 1, you configure **vpn 0 interface tunnel-interface exclude-controller-group-list 2** . As with the above configurations, if this controller group or data center becomes unreachable, the tunnel establishes a control connection with controller group 2, because this group is configured in the **system controller-group-list** command. And if neither controller group 1 or 2 is available, and if another Cisco vSmart Controller is present in the network, the tunnel interface establishes a control connection with that controller.

Configure Redundant Control Connections on One Cisco vEdge Device

When a router has two tunnel connections and the network has two (or more) data centers, you can configure redundant control connections from the Cisco vEdge device to Cisco vSmart Controllers in two of the data centers. It is recommended that do this using the minimum number of OMP sessions—in this case, two. To do this, you configure one of the tunnel interfaces to go only to one of the data centers and the other to go only to the second. This configuration provides vSmart redundancy with the minimum number of OMP sessions.



On the Cisco vEdge device router, define the controller group list and configure the maximum number of OMP sessions to be 2:

```
vEdge(config-system)# controller-group-list 1 2
vEdge(config-system)# max-omp-sessions 2
```

For one of the tunnels, you can use the default affinity configuration (that is, there is nothing to configure) to have this tunnel prefer a Cisco vSmart Controller in group 1. You can also explicitly force this tunnel to prefer Cisco vSmart Controller group 1:

```
vEdge(config-tunnel-interface-1)# max-control-connections 1
```

You do not need to configure **exclude-controller-group-list 2** , because the software evaluates the controller group list in order, starting with group 1. However, you could choose to explicitly exclude vSmart controller group 2.

Then, on the second tunnel, configure it to prefer a vSmart controller in group 2. As with the other tunnel, you limit the maximum number of control connections to 1. In addition, you have to exclude controller group 1 for this tunnel.

```
vEdge(config-tunnel-interface-2)# max-control-connections 1
vEdge(config-tunnel-interface-2)# exclude-controller-group-list 1
```