Service insertion

Service insertion enables network administrators to intercept and redirect traffic flows through intermediate service functions such as firewalls, load balancers, intrusion detection systems, or other network appliances without requiring changes to the underlying network topology.

- Feature history for service insertion, on page 1
- Service insertion, on page 2
- Restrictions for service insertion, on page 4
- Use cases for service insertion, on page 4
- Service insertion configuration methods, on page 5
- Configure service chain actions in a data policy, on page 7
- Traffic steering to a service chain, on page 9
- Configure path preference, on page 13
- How service chains share across user VPNs, on page 14
- How separate interfaces for transmitted and received traffic work, on page 15
- Service chaining trusted and untrusted traffic, on page 16
- How service chains work between two routers, on page 17
- Fall back and restrict behavior configuration for service chain traffic, on page 18
- How service chain interfaces work, on page 18
- Service chaining with software defined cloud interconnect bring your own service, on page 19

Feature history for service insertion

This reference provides a chronological overview of service insertion features introduced in Cisco Catalyst SD-WAN releases, documenting when capabilities were added and their key functionality.

Table 1: Feature History

Feature Name	Release Information	Description
Service Insertion Using Workflows	Cisco IOS XE Catalyst SD-WAN Release 17.13.1a Cisco Catalyst SD-WAN Manager Release 20.13.1	With this feature, you can create a service chain from the **Workflow Library** and configure a service chain action for a policy. A service chain inserts a set of services in the flow of traffic and can be designed to affect traffic according to your needs.

Feature Name	Release Information	Description
Trusted and Untrusted Postures	Cisco IOS XE Catalyst SD-WAN Release 17.14.1a Cisco Catalyst SD-WAN Manager Release 20.14.1	With this feature, you can configure trusted traffic to flow to a trusted high availability pair in a service chain.

Service insertion

Service insertion, also known as *service chaining*, is a network capability that

- places one or more network or security services into the path of specific data traffic within the Cisco Catalyst SD-WAN overlay fabric
- routes traffic through services defined in a service chain according to service chain actions configured for a data policy, and
- can be deployed in any device and topology, including full mesh, HUB-spoke, and Cisco Catalyst SD-WAN Multi-Region Fabric (MRF).

Service insertion capabilities

Cisco Catalyst SD-WAN service chaining is flexible, fully automated, and can be deployed on a per VPN basis. Service chaining includes the following key features:

- Service chaining can be used for overlay, local ingress and egress, inter- and intra-VPN, transit, branch-to-branch, branch-to-internet, branch-to-cloud, and cloud-to-cloud traffic
- Automatic forwarding of traffic through all services in a chain
- Services attachment methods of IPv4, IPv6, dual stack, and tunneled
- Configurable high availability across instances of a single service
- Built-in load balancing across instances of a single service, which supports equal cost multipath routing (ECMP) across high availability pairs
- Advanced service tracking
- Service chain sharing across multiple user VPNS, which can be different or the same as user traffic VPNS
- Traffic steering methods using control policy, data policy, interface ACL, and supported match conditions
- Fall back and restrict behavior
- Path preference and symmetric routing
- Security services to and from service transports
- Trusted and untrusted high availability pairs and traffic marking (from Cisco Catalyst SD-WAN Manager Release 20.14.1)
- Periodic on demand state notifications for serviceability

- Cisco Catalyst SD-WAN Manager orchestration: Workflow based service chaining and traffic policy configuration

The following table provides information about the capabilities of the service chaining feature in releases before and after Cisco Catalyst SD-WAN Manager Release 20.13.1.

Table 2: Service insertion capabilities

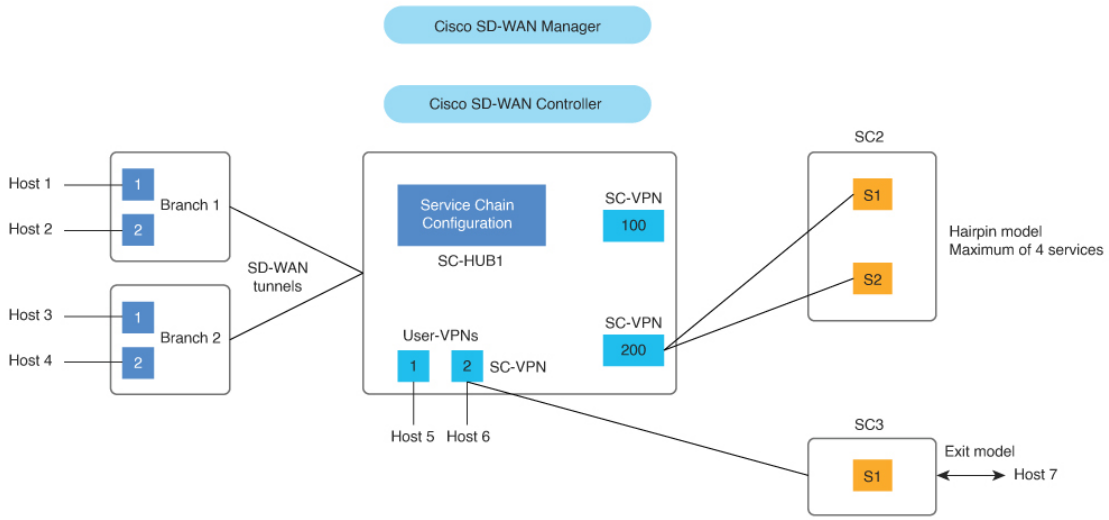
Capability	Releases before Cisco Cisco Catalyst SD-WAN Manager Release 20.13.1	Releases from Cisco Cisco Catalyst SD-WAN Manager Release 20.13.1
Multiple services in a chain	No native support	Native support
Traffic steering	Control policy	Control policy, data policy, interface ACL
Policy binding	Remote	Remote and local
Traffic type	IPv4	IPv4, IPv6, dual stack, tunnel
Load balancing	Across 4 IP addresses that serve as service endpoints	Across 4 instances of active-backup pairs for every traffic type
High availability	As provided by load balancing	Active and backup pairs
Supported deployments	• Single-arm deployment • Single-arm deployment with multiple interfaces • Bump-in-the-wire deployment	• Single-arm deployment • Single-arm deployment with multiple interfaces • Bump-in-the-wire deployment

The following figure illustrates the key elements in service insertion. In this figure, SC-HUB1 is the router to which the service chain is attached.

In the hairpin model, traffic is sent by SC-HUB1 to a service in the service chain, and the service returns the traffic to SC-HUB1. SC-HUB1 then either forwards the traffic to the next service in the service chain or to the destination if the traffic is returning from the last service in the service chain.

In the exit Model, traffic is sent by SC-HUB1 to a service in a service chain, and the service forwards the traffic to the destination. Traffic may return from the destination to the service, which returns it to SC-HUB1.

Figure 1: Service insertion key elements



Restrictions for service insertion

Adhere to the following limitations when configuring service insertion to ensure proper functionality and avoid configuration conflicts.

- A service chain can include up to four service types. Each service type can have multiple instances of the service, either as high availability pairs that are load-balanced by the feature, or behind a third-party load balancer.
- The services in a service chain must be in a single VPN.
- If you are using a dual stack service in a service chain, every service in that service chain must have a dual stack high-availability pair.
- A specific device interface should not be used for more than one service in a particular service chain.
- A specific interface can be used in different service chains only if it is used for the same service type in each of the service chains.
- All the interfaces and tunnels for the services in a service chain should be part of the VPN in which the service chain is defined.
- More than one tracker should not be associated with a given interface. For example, if endpoint-tracker tracker1 is associated with GigabitEthernet1, a different tracker cannot be associated with GigabitEthernet1.

Use cases for service insertion

Service insertion use cases demonstrate how traffic can be directed through security services like firewalls to enhance network protection, ensure inter-VPN security, and meet regulatory compliance requirements.

- Service chaining can be used when traffic from a less secure region of a network should pass through a firewall to ensure that it has not been tampered with.
- Service chaining can be used in a network that consists of multiple VPNs, where each represents a different function or organization, to ensure that traffic between VPNs flows through a firewall. For example, in a campus, interdepartmental traffic might go through a firewall, while intradepartmental traffic might be routed directly.
- Service chaining can be used to ensure regulatory compliance, such as Payment Card Industry Data Security Standard (PCI DSS), where PCI traffic should flow through firewalls in a centralized data center or regional hub.

Service insertion configuration methods

Use one of these methods to configure service insertion:

- Configure service insertion using a CLI template
- Configure service insertion using the workflow library

Configure service insertion using a CLI template

This task configures service insertion using CLI templates to create service chains that define the sequence of services through which traffic is processed.

For more information about using CLI templates, see CLI Add-On Feature Templates and CLI Templates.

Note By default, CLI templates execute commands in global config mode.

Procedure

Step 1 Create a service chain:

Example:

```
service-chain chain-number
```

Step 2 Configure a description for the service chain:

Example:

```
service-chain-description description
```

Step 3 Specify the services that are in the service chain and configure related options:

Example:

```
service service-type service-parameters
```

Step 4 (Optional, from Cisco Catalyst SD-WAN Manager Release 20.14.1) Configure the trust posture for the services that are in the service chain:

Example:

```
service service-type service-transport-ha-pair value attribute trust-posture {trusted | untrusted}
```

Step 5 (Optional) Configure all Cisco Catalyst SD-WAN bidirectional forwarding (BFD) sessions to be brought down:

Example:

```
service-chain-affect-bfd
```

Step 6 Specify the name of the VPN that hosts all services in the service chain:

Example:

```
service-chain-vrf vrf
```

Step 7 (Optional, enabled by default) Enable endpoint tracking for services in the service chain:

Example:

```
track-enable
```

Step 8 (Optional, enabled by default) Enable the service chain, which makes it active on devices:

Example:

```
service-chain-enable
```

The service insertion is configured using the CLI template with the specified service chain parameters and settings.

Configure service insertion using the workflow library

Service insertion allows you to configure network services in a specific order to process traffic before it reaches its destination. Using the workflow library provides a guided approach to create and manage service chains effectively.

Beginning with Cisco Catalyst SD-WAN Manager Release 20.13.1, you can configure service insertion by using the **Workflow Library**. From the **Workflow Library**, you can create a new service chain or modify an existing one. A service chain can contain UP to four service types.

The workflow guides you through configuring several steps, including:

- Configuring the name and description of the service chain
- Specifying the services in the service chain and the order of the services in the chain
- Provide attachment parameters for the services in the chain, which are used when you attach the service chain to routers
- For each service type, specify the VPN and configure options such as load balancing, high availability and tracking

Procedure

Step 1 From the Cisco SD-WAN Manager menu, choose **Workflows** > **Workflow Library**.

Step 2 Click **Define and Configure Service Chain**.

Step 3 Follow the prompts in the workflow.

Ensure that you define a tacker. Tracker configuration is critical to avoid blackhauling. Defining a tracker ensures that the service chain is determined to be in the UP state and is used. If the IP address of a service chain firewall is used with an ICMP-based tracker, ensure that the firewall allows ICMP on the appropriate interface.

Ensure that the service chain can route returning traffic back into the Cisco Catalyst SD-WAN fabric. To do so, use dynamic routing protocols between the service chain and Cisco Catalyst SD-WAN router (service chain hub) or use static routes.

Attach the service chain to the appropriate Cisco Catalyst SD-WAN SC-Hub router. The service chain does not need to be attached to branch routers.

The service chain is configured and ready to be used with data policies and traffic steering configurations.

What to do next

After you configure service insertion, perform the following actions as needed:

- Configure service chain actions for a data policy to route traffic through a service chain. See Configure Service Chain Actions in a Data Policy.
- Use a control policy, data policy, or interface access control list to direct traffic to a service chain. See Traffic Steering to a Service Chain.
- Configure TLOC preference or affinity preference to choose the preferred path for traffic to a service chain. See Path Preference.
- Configure separate interfaces for transmitted and received traffic. See Separate Interfaces for Transmitted and Received Traffic.
- Configure trusted traffic to flow to a trusted high availability pair. See Service Chaining Trusted and Untrusted Traffic.
- Configure fall back or restrict behavior for traffic that travels through a service chain. See Configure Fall Back and Restrict Behavior for Traffic Through a Service Chain.

Configure service chain actions in a data policy

This task enables you to route traffic through a service chain by configuring service chain actions for a data policy.

Beginning with Cisco Catalyst SD-WAN Manager Release 20.13.1, you can route traffic through a service chain by configuring service chain actions for a data policy.

Procedure

Step 1 From the Cisco SD-WAN Manager menu, choose **Configuration** > **Policies**.

Step 2 Click **Custom Options** then click **Traffic Policy** under **Centralized Policy**.

Step 3 Click the **Traffic Data** tab.

Step 4 Click **Add Policy** and click **Create New**.

Step 5 Click **Sequence Type** and choose **Service Chaining** from the **Add Data Policy** dialog box.

Step 6 Click the **Actions** tab.

Step 7 Click **Service**.

Step 8 Configure the fields that the following table describes.

Table 3: Service Chain Action Fields

Field	Description
Service: Type	Choose a type of service for the service chain.
Service: VPN	VPN in which the service chain is hosted. Range: 1 to 65525, excluding 512. For details see the VRF range behavior change described here.
Service: TLOC IP	Enter the IP address of the Transport Locator (TLOC) for applying the services in the service chain.
Color	Choose a color for the TLOC.
Encapsulation	Choose the encapsulation type for the TLOC.
Service: TLOC List	Choose a predefined TLOC list to use for applying services to branch traffic.
Local	Check the **Local** check box if the service chain is hosted locally. If you do not check this check box, the service chain is hosted remotely.
Restrict	Check this option to cause packets to be dropped if the service chain goes down. f you configure this policy with the **Local** option, packets are dropped locally. If you configure this policy with the **Remote** option, packets are dropped on the remote host. This option is unchecked by default (traffic falls back to routing).

Service chain actions are configured in the data policy, enabling traffic to be routed through the specified service chain.

Traffic steering to a service chain

You can direct traffic to a service chain by using a control policy, data policy, or interface access control list.

How traffic steering using a control policy works

You can use a control policy to modify Cisco Overlay Management Routes, also referred to as vRoutes, to direct traffic to a service chain instead of the original destination.

Summary

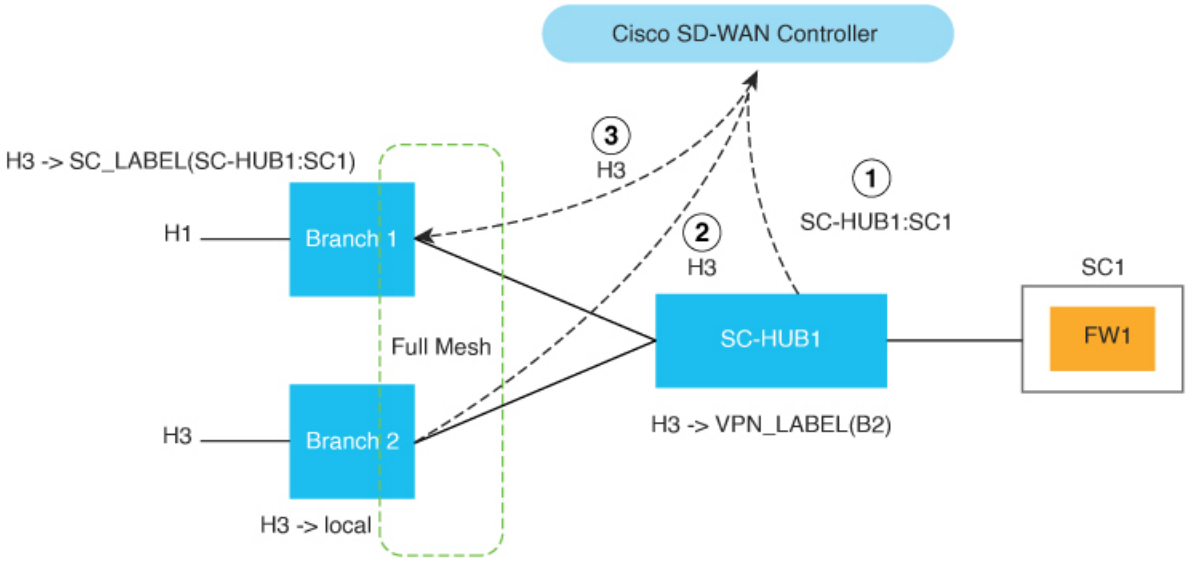
The key components involved in traffic steering using a control policy are:

- Control policy: Modifies vRoutes to redirect traffic to service chains
- Service chain hub (SC-HUB1): Advertises service chain routes
- Branch routers (B1, B2): Source and destination points for traffic flow
- Cisco SD-WAN Controller: Processes and advertises modified routes based on control policy
- Service chain (SC1): Traffic processing path applied to matching flows

The policy causes service chain 1 (SC1) to be applied to traffic that flows between H1 (host 1) and H3 (host 3) by setting SC1 as the next hop for H1 and H3 traffic routes.

Workflow

Figure 2: Traffic Steering with a Control Policy



These stages describe how traffic steering using a control policy works:

1. SC-HUB1 advertises the SC1 route to make the service chain available for traffic steering.

2. B2 advertises the H3 route to the Cisco SD-WAN Controller to establish the original destination path.
3. The control policy results in overriding the H3 route next hop to SC1 and the Cisco SD-WAN Controller advertises the modified H3 route to B1.

   Before the policy is in effect, traffic flows from B2 (branch 2) to B1 (branch 1). After the policy is in effect, traffic flows from B2 to SC-HUB1:SC1 to B1.

Note Example configuration:

```
Control-policy name
    sequence number
        match route
        action accept
   set service-chain sc_name [tloc|tloc-list name] [vpn vpn]
apply-policy site-list site_list control-policy name out
```

How data policies steer traffic

You can use a data policy to match traffic and operate in the context of source VPNs during forwarding.

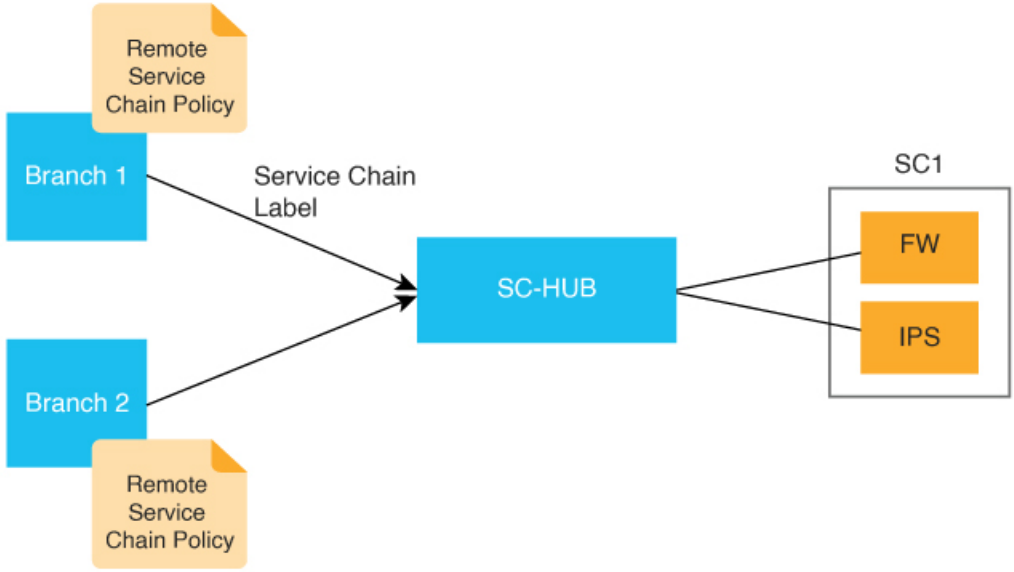
Summary

The key components involved in traffic steering using data policies are:

- Data policy: Matches traffic and defines forwarding actions based on specified criteria
- Match criteria: Specifies applications to be matched to source and destination IP address combinations
- Service chain: Defines the path for traffic processing through network services
- TLOC ranking: Specifies traffic path preference using TLOC ranking

Workflow

Figure 3: Traffic Steering with Traffic Service Chaining Intent Specified at a Remote Branch



These stages describe how data policies steer traffic:

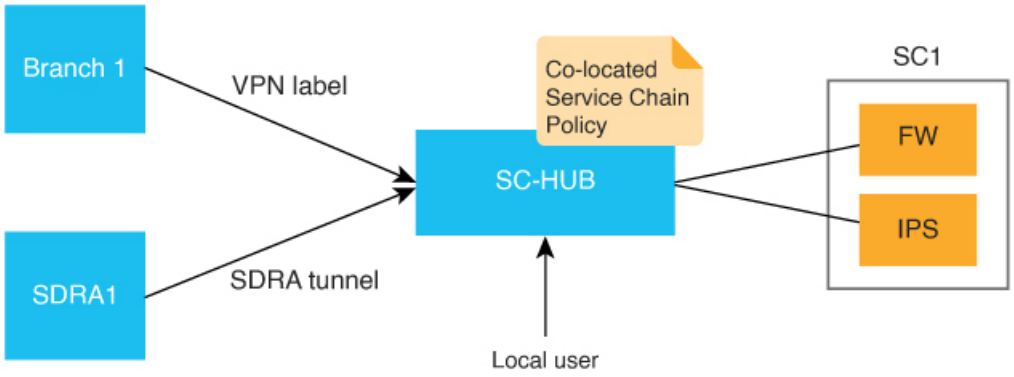
1. The system configures the data policy with match criteria and service chaining parameters for remote branch traffic steering.
   - **match criteria** specifies applications to be matched to source and destination IP address combinations
   - **restrict|fallback** configures restrict or fall back
   - **TLOC|TLOC-list** list specifies the traffic path preference using TLOC ranking

Note **set attribute trust-posture** is available from Cisco Catalyst SD-WAN Manager Release 20.14.1.

```
policy
 data-policy name
  vpn-list name
   sequence 100
    match criteria
    action accept
     set service-chain sc_name vpn vpn {restrict|fallback} [tloc|tloc-list list]
set attribute trust-posture {trusted | untrusted}
apply-policy site-list remote-sites data-policy name from-service
```

2. The system applies service chaining intent locally on the device to which the service chain is attached.

Figure 4: Traffic Steering with Traffic Service Chaining Intent Specified on a Local Device



In this configuration, **local** indicates that traffic needs to be directed to a service chain locally.

```
set service-chain SC1 [vpn vpn] local [restrict|fallback]
apply-policy site-list SC-HUB-sites data-policy policy {from-service|from
tunnel}|from-tunnel}
```

Traffic steering using an interface access control list

You can use an interface access control list (ACL) to service chain traffic that is incoming or outgoing on a specified interface. In some situations, the traffic forwarding decision may need to come from a prior routing lookup or data policy.

This approach is useful when all traffic from an interface should be directed through a service chain.

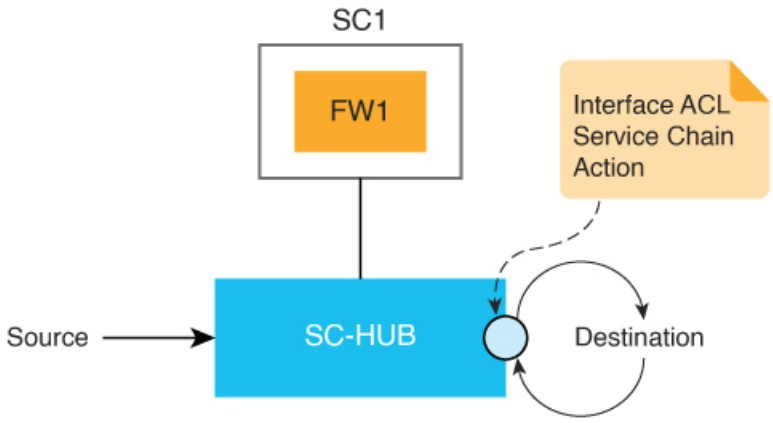
Summary

The key components involved in traffic steering using an interface ACL are:

- **Interface ACL**: Controls traffic flow on specified interfaces and directs traffic through service chains
- **Service chain**: Processes traffic according to configured policies and routing decisions
- **Traffic matching criteria**: Determines which traffic flows are subject to service chaining

Workflow

Figure 5: Traffic Steering with an ACL



Traffic steering using an interface ACL involves the following stages:

1. Configure the access control list with matching criteria and service chain action.

```
access-list list
  sequence number
    match criteria
    action accept
     set service-chain SC1 [vpn vpn] {restrict|fallback}
```

2. Apply the access list to the interface for incoming or outgoing traffic.

```
interface interface
 access-list list {in|out}
```

3. The system processes traffic according to the configured ACL rules and directs matching traffic through the specified service chain.

Note Service chaining can be used in dual-site setups using Virtual Router Redundancy Protocol (VRRP) for redundancy. Ensure your service chaining policy does not block VRRP control traffic. Service chaining policies often have a default action to **drop** traffic that does not meet the match criteria. This can accidentally block VRRP control packets. To prevent this, set the default action in your service chaining policy to **accept** VRRP control traffic. This will ensure VRRP works correctly for redundancy.

Configure path preference

You configure path preference to control how traffic flows through service chains by directing traffic over specific TLOCs or by setting affinity preferences between branch sites and service chain hubs.

You can use TLOC preference or affinity preference to choose the preferred path for traffic to a service chain. If both TLOC preference and affinity preference are configured, the affinity preference is evaluated first, then the TLOC preference is evaluated.

Procedure

Step 1 Configure a TLOC list to direct traffic only over certain TLOCs or to prefer certain TLOCs over others.

The TLOC list can be specified with **TLOC-list** as part of a service chain action in a data policy or a control policy.

Step 2 To configure affinity preference, use **affinity-group preference** in branch sites to set the affinities of branches, and use **affinity-group** in service chain hubs to set the affinities of VPNs.

The data policy **set service chain** action is compliant with affinity by default.

Step 3 (Optional) Configure the following command to disable consideration of affinity in a data policy: **data-policy-ignore-affinity-metric**

Traffic to the service chain follows the configured path preferences, with affinity preference taking priority over TLOC preference when both are configured.

How service chains share across user VPNs

Service chain sharing allows a single service chain VPN to be utilized by traffic from multiple user VPNs without requiring additional configuration. When source and destination VPNs differ, route leaking is required between the source and destination VPN.

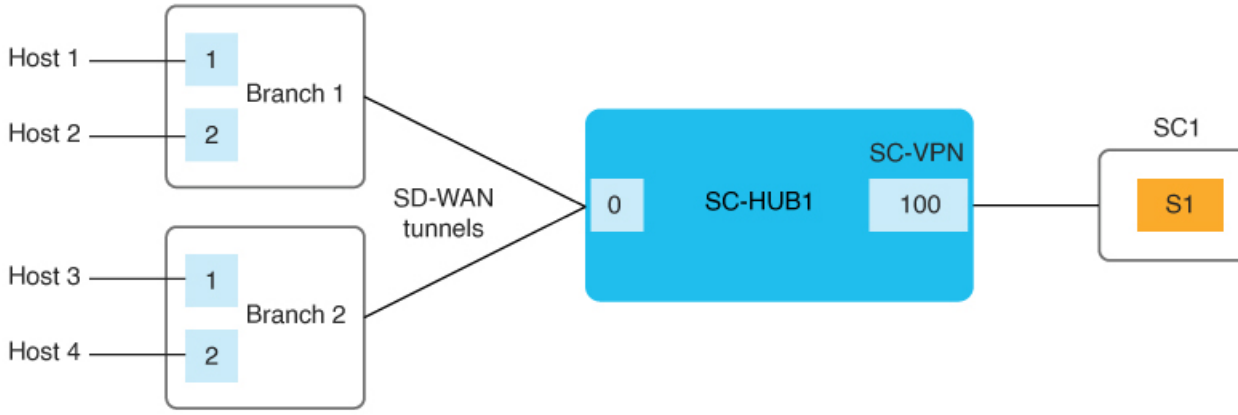
Summary

The key components involved in the service chain sharing process are:

- Service chain VPN: A dedicated VPN that provides service chaining functionality and can be shared across multiple user VPNs
- User VPNs: Individual VPNs that can utilize shared service chains for traffic processing
- Route leaking: Required mechanism when source and destination VPNs are different to enable traffic flow

Workflow

Figure 6: Service Chain Sharing Across VPNs



These stages describe how service chains share across user VPNs:

1. The service chain VPN is automatically made available to traffic from multiple user VPNs.
   - SC1 (service chain 1) is attached to VPN100 can be automatically shared by traffic in VPN1 (H1) and VPN2 (H4)
2. Traffic between different VPNs can be service chained through various VPN locations.
   - Traffic between VPN1 (H1) and VPN2 (H4) can be service chained in VPN1 or VPN2 or in a shared service chain (VPN100)

How separate interfaces for transmitted and received traffic work

You can use the **service** command to configure separate interfaces for transmitted and received traffic through a service chain. In this situation, transmitted and received traffic are tracked independently. For more information, see service.

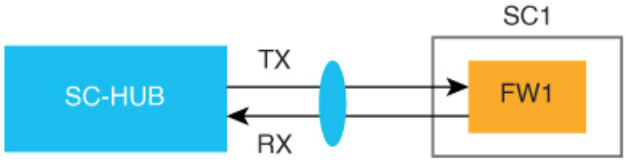
Summary

The key components involved in the separate interface configuration for transmitted and received traffic are:

- Service chain: Routes traffic through multiple network services with independent interfaces for each direction
- Transmitted traffic interface: Handles outgoing traffic flow through the service chain
- Received traffic interface: Handles incoming traffic flow through the service chain
- Traffic tracking system: Monitors transmitted and received traffic independently

Workflow

Figure 7: Separate Interfaces for Transmitted and Received Traffic



Service chaining trusted and untrusted traffic

Minimum supported releases: Cisco IOS XE Catalyst SD-WAN Release 17.14.1a, Cisco Catalyst SD-WAN Manager Release 20.14.1

In service chaining configurations, the system can separate trusted and untrusted traffic flows to maintain security and performance isolation. The default trust-posture of a packet is trusted unless explicitly configured otherwise.

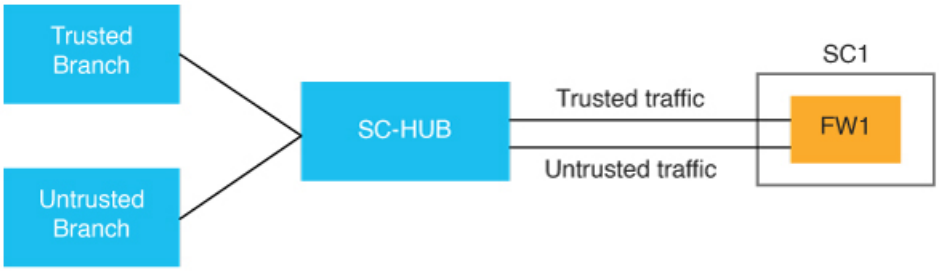
Summary

The key components involved in service chaining trusted and untrusted traffic are:

- Data policy: Uses the **set attribute trust-posture untrusted action** command to mark packets as trusted or untrusted
- Trusted high availability pair: Receives and processes trusted traffic flows
- Untrusted high availability pair: Receives and processes untrusted traffic flows
- Service chain configuration: Defines the service transport HA pairs and their trust-posture attributes

Workflow

Figure 8: Trusted and Untrusted Traffic



These stages describe how the system processes and routes trusted and untrusted traffic through service chains:

1. The system evaluates incoming packets and determines their trust-posture based on data policy configuration. By default, packets are marked as trusted unless explicitly configured otherwise.
2. The system routes trusted traffic to the configured trusted high availability pair within the service chain.

3. The system routes untrusted traffic to the configured untrusted high availability pair within the service chain.
4. The service chain configuration defines the trust-posture attribute for each service transport HA pair using the following example configuration:

```
service-chain SC1
    service netsvc1
        sequence 10
        service-transport-ha-pair 1
            attribute trust-posture {trusted|untrusted}
```

How service chains work between two routers

When the router that transmits traffic to a service chain differs from the router that receives traffic from the service chain, a specific configuration approach is required. This scenario applies to intra-VPN traffic only and involves service chains with a single service.

Summary

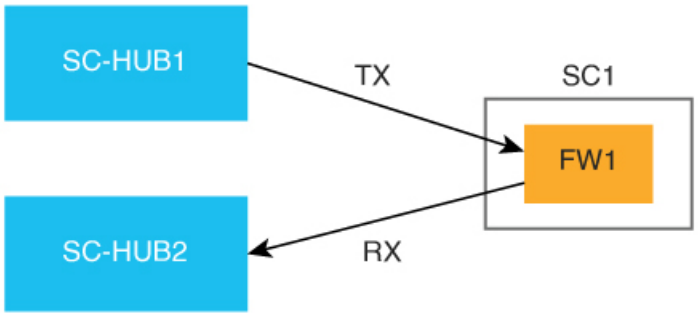
The key components involved in the service chain process between two routers are:

- Transmitting router: Sends traffic to the service chain for processing
- Receiving router: Receives processed traffic from the service chain
- Service chain: Contains only one service and processes intra-VPN traffic

The process requires identical service chain configuration on both routers to ensure proper traffic flow and processing.

Workflow

Figure 9: Service Chain Between Two Router



These stages describe how service chains operate between two routers:

1. The transmitting router sends intra-VPN traffic to the configured service chain for processing.
2. The service chain processes the traffic using its single configured service.
3. The receiving router receives the processed traffic from the service chain and continues routing operations.

Fall back and restrict behavior configuration for service chain traffic

You can configure fall back or restrict behavior for traffic that travels through a service chain.

When **fallback** is configured in the **set service-chain** action, traffic falls back to routing if a service chain goes down or if the TLOCs that are specified in a policy are not available.

When **restrict** is configured in the **set service-chain** action, packets are dropped if a service chain goes down or if the TLOCs that are specified in a policy are not available. The restrict behavior is suitable for security services such as a firewall.

Fall back and restrict can be specified in a centralized data policy (remote or collocated) and an interface ACL.

Note If an egress ACL is used to direct traffic to a service chain, all packets continue to the destination even if the restrict behavior is configured because the forwarding decision is made before the state of the service chain is detected.

How service chain interfaces work

The services in a service chain must be in a single VPN, called a *service chain VPN*, or *SC-VPN*.

Summary

The key components involved in attaching services in a service chain to a router are:

- Cisco Catalyst SD-WAN router: Acts as the attachment point for service chain services
- Service chain VPN (SC-VPN): Contains all services in the service chain
- Interface types: Provide various attachment methods including IPv4, IPv6, dual stack, or tunnel interfaces

Workflow

Figure 10: Attaching Services to a Router

These stages describe how services in a service chain attach to a router through different interface types:

1. The services can be attached through any combination of the following interface types:

1	IPv4 attachment
2	IPv6 attachment
3	Dual stack attachment
4	Tunnel attachment

Service chaining with software defined cloud interconnect bring your own service

Service chaining with software defined cloud interconnect bring your own service is a functionality that

- establishes a centralized location for service inspection by connecting a service chain to the Cisco Catalyst 8000v Edge Software (Catalyst 8000v) SDCI gateways that are deployed in the middle mile network
- enables the seamless integration of external services with the SDCI infrastructure, and
- enforces colocated data policies, also known as centralized data policies, on these gateways within the middle mile network for selective data traffic inspection.

Network architecture

The Software Defined Cloud Interconnect (SDCI) establishes connections between branch sites and the cloud through network service providers, including Megaport and Equinix. In this context, a branch site represents the first mile, a service provider acts as the middle mile, and the cloud serves as the last mile.

The BYOS service inspection for SDCI allows service chaining in the following situations:

- Connecting branch sites to cloud workloads through middle mile providers using the C8000v SDCI gateway.

- Interconnecting branch site through the middle mile provider using the Catalyst 8000v SDCI gateway.
- Facilitating intercloud traffic connectivity by the middle mile provider through the Catalyst 8000v SDCI gateway.