SD-WAN Topologies

For Releases 20.2 and later.

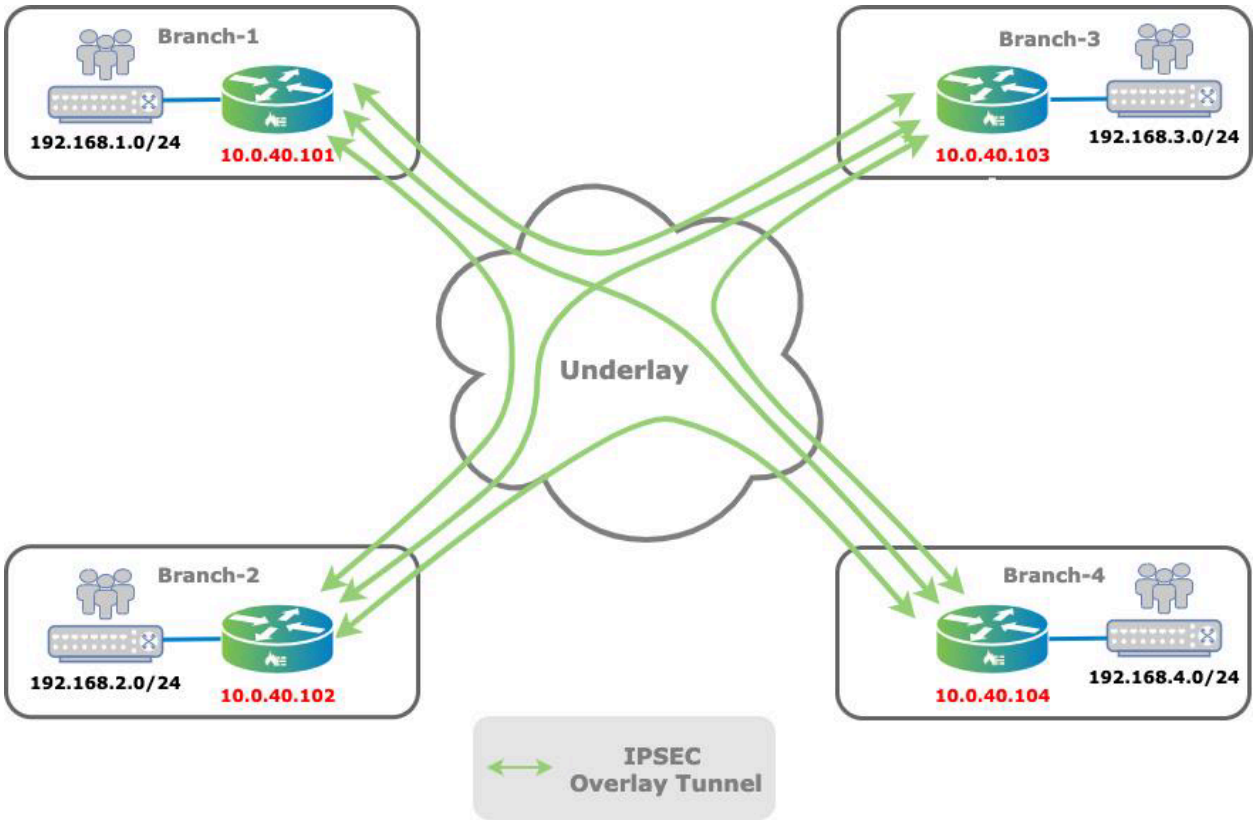
The Versa Networks solution supports the following SD-WAN overlay topologies:

- Full mesh
- Hub and spoke
- Regional mesh
- Multi-VRF, or multitenancy

These topologies are established by using that well-known routing techniques that have been used for a long time in MPLS Layer 3 VPN networks. They use MP-BGP communities to achieve fine-grained route control and to provide flexible options for manipulating and fine-tuning routes. You can use Director Workflows to create these topologies, thus simplifying these complex configurations.

Full-Mesh Topology

You use a full-mesh topology for any-to-any communication. In this type of topology, branches communicate directly using overlay tunnels, and traffic does not need to transit through a hub or centralized site. The following figure illustrates a full-mesh topology.



A full-mesh topology is generally the preferred topology when branches must communicate directly with each other. Typically, you choose a full-mesh topology over a hub-and-spoke topology for voice applications, because a hub-and-spoke topology introduces delay when the hub is distant from the branches. Another instance in which a full-mesh topology is preferred over hub and spoke is a distributed security architecture, where policy enforcement is performed at the branch. Here, the full-mesh topology avoids the need to funnel traffic to hub sites for inspection.

In Director Workflows, the full-mesh topology is the default option.

In a full-mesh topology SLA monitoring probes are sent to every remote branch on every available transport. SLA monitoring probes are used to track reachability and to measure link metrics for each access circuit towards any given remote site. You can use SLA optimization features such as adaptive SLA and data-driven SLA to optimize the SLA load in large deployments.To verify SLA monitoring status, issue the following CLI command:

```
admin@Branch-1-cli> show orgs org Tenant1 sd-wan sla-monitor status
                                    LOCAL  REMOTE
                                    WAN    WAN
          PATH     FWD   LOCAL     REMOTE   LINK   LINK    ADAPTIVE    DAMP    DAMP   CONN
LAST
SITE NAME    HANDLE    CLASS  WAN LINK   WAN LINK  ID     ID      MONITORING   STATE    FLAPS
STATE  FLAPS  FLAPPED
-------------------------------------------------------------------------------------------------------
Branch-2     6689028   fc_ef  MPLS       MPLS      1      1       active       disable  0      up     1      00:06:26
             6693380   fc_ef  Internet   Internet  2      2       active       disable  0      up     1      00:06:26
Branch-3     6754564   fc_ef  MPLS       MPLS      1      1       active       disable  0      up     1      00:06:32
```

```
          6758916  fc_ef  Internet  Internet  2     2     active    disable  0    up   1    00:06:31
Branch-4  6820100  fc_ef  MPLS      MPLS      1     1     active    disable  0    up   2    00:05:45
          6824452  fc_ef  Internet  Internet  2     2     active    disable  0    up   2    00:05:44
Controller-1 69888 fc_nc  MPLS      MPLS      1     1     disable   disable  0    up   1    00:16:38
          74240    fc_nc  Internet  Internet  2     2     disable   disable  0    up   1    00:16:38
```

The output above shows the SLA monitoring view from Branch-1, which has internet and MPLS transports towards all branches and towards the Controller node.

In a full-mesh topology, you must determine the proper scaling of the maximum number of branches. To dimension the deployment, you must consider many variables, including the following:

- Number of WAN links
- Number of tenants
- Forwarding classes being monitored
- SLA monitor interval
- Branch hardware
- Bandwidth to the branch

For example, in a full-mesh topology with 1000 branches that have one tenant per site and two WAN links in different transport domains, if you use the Versa Operating System™ (VOS™) device default SLA monitoring configuration, the SLA probe traffic consumes 6.25 Mbps of bandwidth at each site. Increasing the number of branch CPE devices increases both the bandwidth usage and the CPU overhead to perform SLA monitoring. You can limit the SLA monitoring traffic to lower link utilization, for instance, on high-cost links such as LTE connections. For more information, see Configure SLA Monitoring for SD-WAN Traffic Steering.

In a full-mesh topology, there is direct reachability to prefixes in remote branches. Traffic is routed to those prefixes using the next hop of the remote branch loopback (TVI) interfaces. If there is an underlay cut or the SLA probing cannot declare the remote branch to be reachable, the SLA monitoring session is down and therefore the next hop is not reachable. The result is that this prefix is withdrawn from the routing table. The route table below illustrates that for the Branch-1 VRF, three routes have been withdrawn. These routes have the interface name "indirect."

```
admin@Branch-1-cli> show route routing-instance Tenant1-LAN-VR
Routes for Routing instance : Tenant1-LAN-VR AFI: ipv4 SAFI: unicast
Codes: E1 - OSPF external type 1, E2 - OSPF external type 2
IA - inter area, iA - intra area,
L1 - IS-IS level-1, L2 - IS-IS level-2
N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
RTI - Learnt from another routing-instance
+ - Active Route
Prot   Type    Dest Address/Mask   Next-hop       Age      Interface name
----   ----    ----------------    --------       ---      --------------
BGP    N/A     +0.0.0.0/0          169.254.0.2    1w6d20h  tvi-0/603.0
conn   N/A     +169.254.0.2/31     0.0.0.0        1w6d20h  tvi-0/603.0
local  N/A     +169.254.0.3/32     0.0.0.0        1w6d20h  directly connected
conn   N/A     +192.168.1.0/24     0.0.0.0        1w6d20h  vni-0/2.0
local  N/A     +192.168.1.1/32     0.0.0.0        1w6d20h  directly connected
BGP    N/A     +192.168.2.0/24     10.0.40.102    1w6d20h  Indirect
```

```
BGP    N/A    +192.168.3.0/24    10.0.40.103    1w6d20h  Indirect
BGP    N/A    +192.168.4.0/24    10.0.40.104    00:05:58 Indirect
```

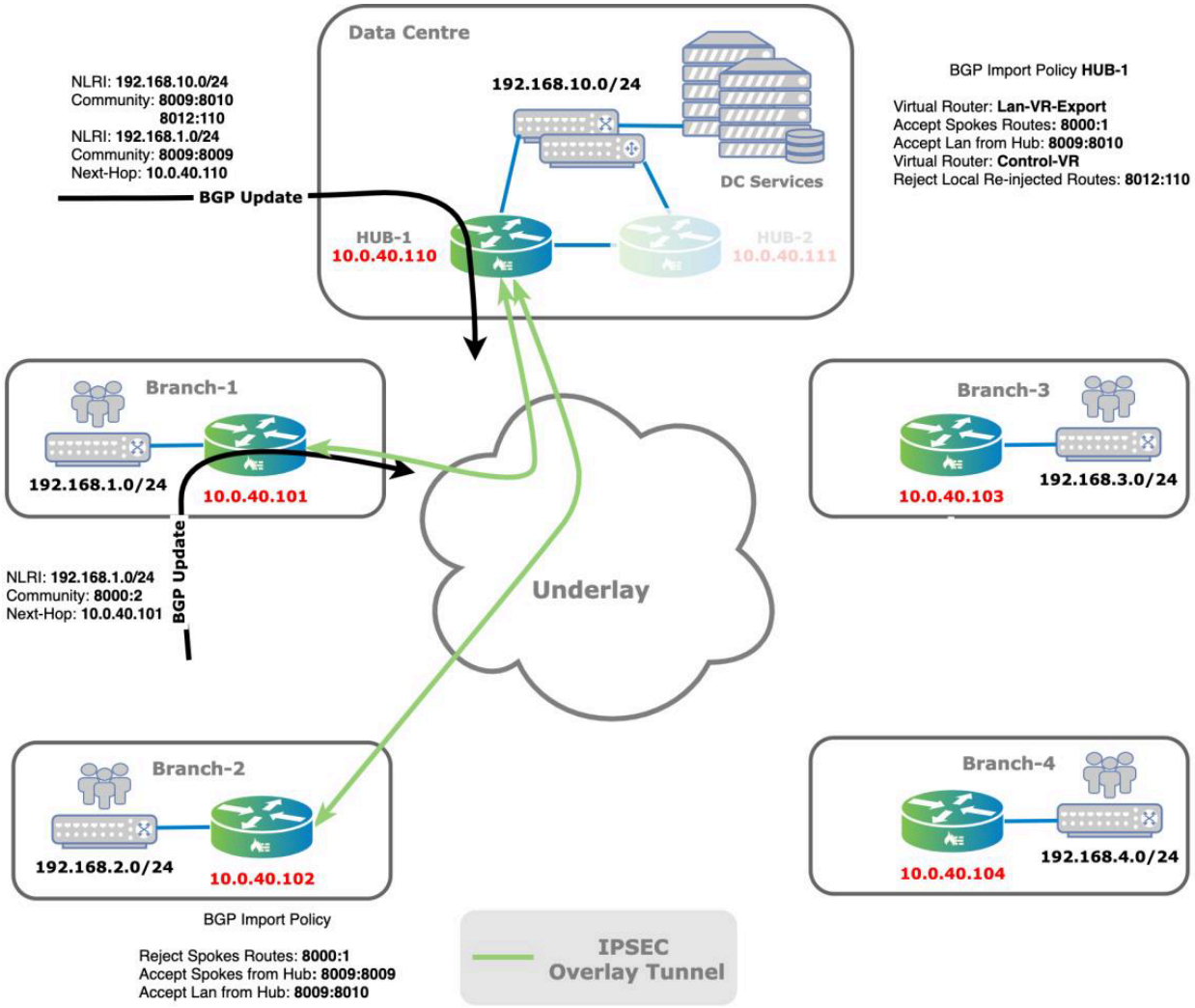
Hub-and-Spoke Topology

The Versa Networks SD-WAN solution supports different types of hub-and-spoke topologies:

- Spoke to hub only
- Spoke to spoke through a hub
- Spoke to spoke direct
- Spoke-hub-hub-spoke

Spoke-to-Hub Only

In a spoke-to-hub-only topology, the only prefixes advertised, by default, are hub routes, and spokes routes are not re-advertised by the hub branch. You use this topology when spokes do not have to communicate with each other. A good example is a network of ATM cash machines in which devices communicate exclusively with resources in the customer data center. The following figures shows that spoke prefixes are accepted only by the hub and that they are rejected by other spokes based on the BGP community configuration.



The following CLI output shows that the spoke Branch-1 VRF route table contains routes only from the hub:

```
admin@Branch-1-cli> show route routing-instance Tenant1-LAN-VR
Routes for Routing instance : Tenant1-LAN-VR AFI: ipv4 SAFI: unicast
Codes: E1 - OSPF external type 1, E2 - OSPF external type 2
IA - inter area, iA - intra area,
L1 - IS-IS level-1, L2 - IS-IS level-2
N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
RTI - Learnt from another routing-instance
+ - Active Route

Prot   Type   Dest Address/Mask   Next-hop       Age      Interface name
----   ----   -----------------   --------       ---      --------------
conn   N/A   +192.168.1.0/24      0.0.0.0        2d20h04m  vni-0/2.0
local  N/A   +192.168.1.1/32      0.0.0.0        2d20h04m  directly connected
BGP    N/A   +192.168.10.0/24     10.0.40.110    2d20h04m  Indirect
BGP    N/A    192.168.10.0/24     10.0.40.111    2d20h04m  Indirect
```

The route table on spoke Branch-1 shows only destinations behind hubs, again with Hub-1 being preferred, and the

table shows no spokes routes. The following output shows the prefixes advertised by spoke Branch-1:

```
admin@Branch-1-cli> show route table l3vpn.ipv4.unicast advertising-protocol bgp
Routes for Routing instance : Tenant1-Control-VR AFI: ipv4 SAFI: unicast

Routing entry for 192.168.1.0/24
   Peer Address          : 10.0.40.1
   Route Distinguisher   : 2L:2
   Next-hop              : 10.0.40.101
   VPN Label             : 24704
   Local Preference      : 110
   AS Path               : N/A
   Origin                : Igp
   MED                   : 0
   Community             : [ 8000:2 8001:110 8002:111 ]
   Extended community    : [ target:2L:2 ]
```

You use BGP import policies to filter spoke routes. For example, using spoke community 8000:2 filters out spoke routes on hubs in the LAN-VR-Export VR, and these routes are not advertised back to the spokes. Therefore, the hub route tables contains all spoke prefixes:

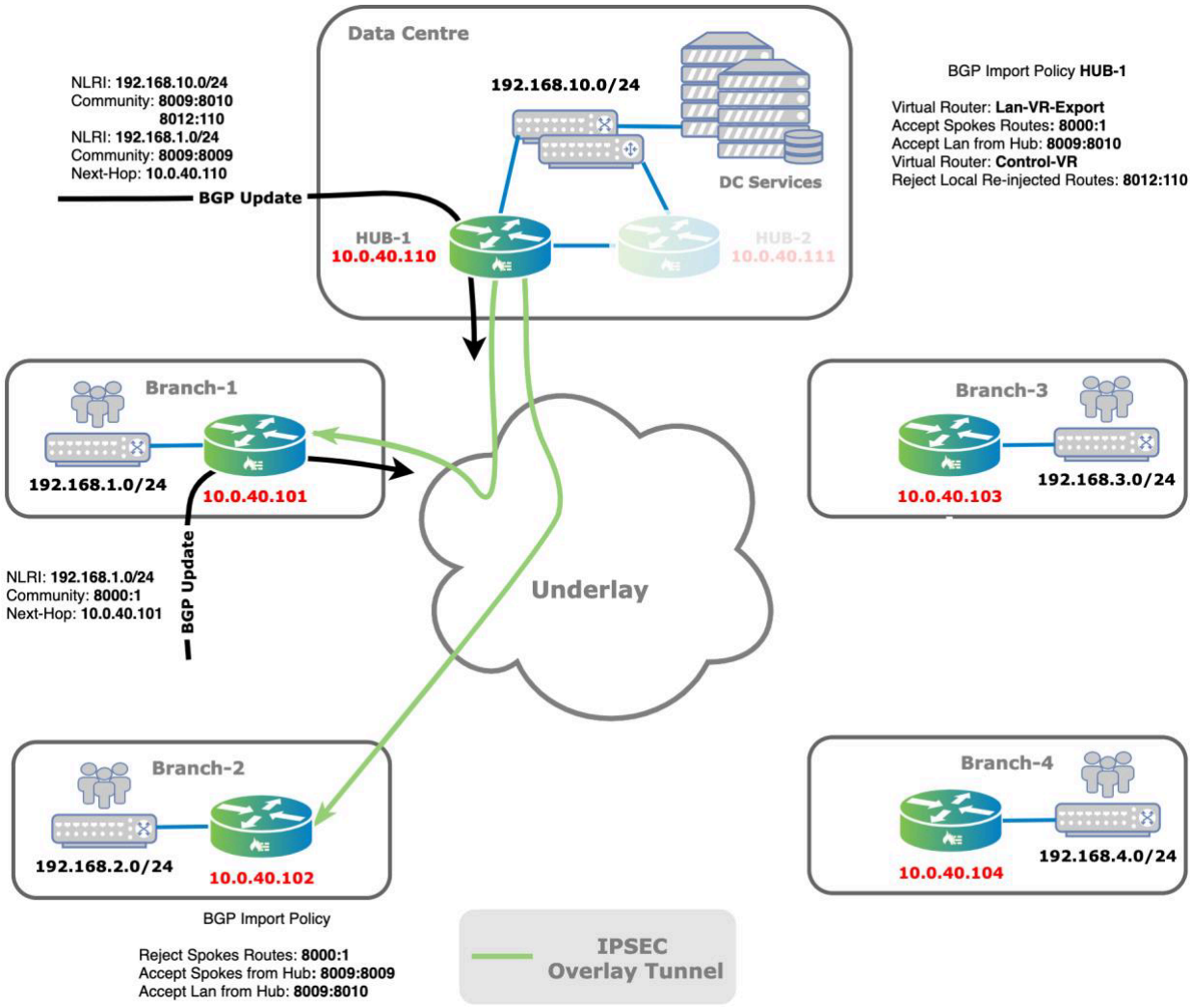
```
admin@Hub-1-cli> show route routing-instance Tenant1-LAN-VR
Routes for Routing instance : Tenant1-LAN-VR AFI: ipv4 SAFI: unicast
Codes: E1 - OSPF external type 1, E2 - OSPF external type 2
IA - inter area, iA - intra area,
L1 - IS-IS level-1, L2 - IS-IS level-2
N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
RTI - Learnt from another routing-instance
+ - Active Route

Prot   Type  Dest Address/Mask   Next-hop       Age       Interface name
----   ----  -----------------   --------       ---       --------------
BGP    N/A  +192.168.1.0/24      10.0.40.101     00:21:24  Indirect
BGP    N/A  +192.168.2.0/24      10.0.40.102     00:21:27  Indirect
BGP    N/A  +192.168.3.0/24      10.0.40.103     00:21:23  Indirect
BGP    N/A  +192.168.4.0/24      10.0.40.104     00:21:26  Indirect
BGP    N/A   192.168.10.0/24     10.0.40.111     00:36:45  Indirect
conn   N/A  +192.168.10.0/24     0.0.0.0         00:51:13  vni-0/2.0
local  N/A  +192.168.10.1/32     0.0.0.0         00:51:13  directly connected
```

On hubs, all spokes prefixes are installed in the corresponding VRF. You can implement redistribution policy on hubs to perform route summarization or to generate default a static route, for instance, to attract traffic from spokes.

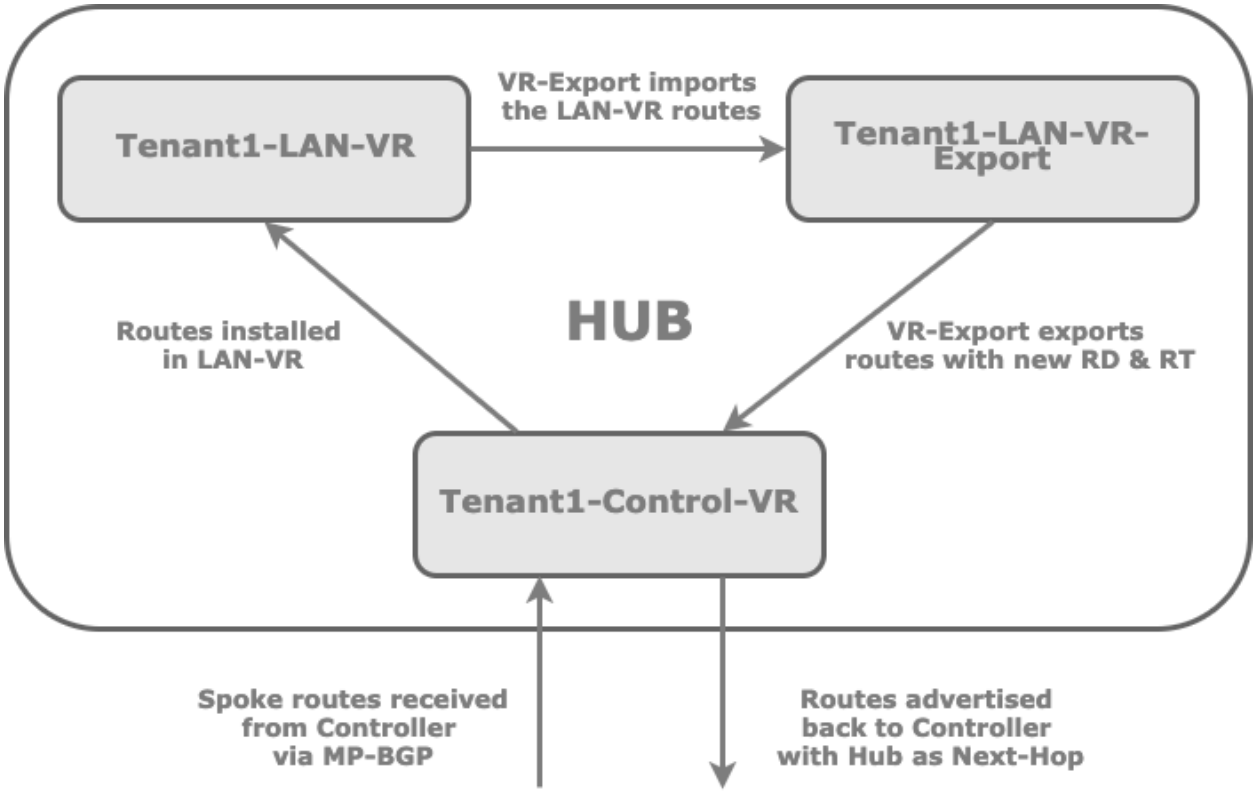
Spoke-to-Spoke Through a Hub

In a spoke-to-spoke through a hub topology, spoke sites are connected to each other through hub site. The data path between two spokes travels through the hub. The following figure illustrates this topology.



You can use the spoke-to-spoke through a hub topology when communication between branches is not required, for example, when security and other services are centralized at the hub site or when the cost of ownership of WAN links dictates it.

The following figure shows the method that hub sites use manipulate VRF spoke routes. With this method, you change the route distinguisher on the hub for a set of VRF routes that you are advertising so that the Controller nodes can separate the routes and accept them during BGP route selection. The Controller nodes have the original routes from the spokes and the spoke routes advertised by the hubs, and they reflects them to the branches. You use route-target filtering on the spokes to perform the remainder of the route selection. With route-target filter, you import the hub-advertised spoke routes, and the Controller nodes use these routes to select the hub as the next hop towards the remote sites.



In a spoke-to-spoke via hub topology, the branches communicate only through the hub. The IP prefixes of remote branches always have the hub as the next hop.

SLA monitoring is active only on paths towards hub sites and Controller nodes, and spoke sites are not monitored. This reduces the amount of SLA probe traffic compared to a full-mesh topology and addresses the concerns of scalability in deployments that have a large number of branches.

The following CLI output shows the SLA monitoring view on Branch-1:

```
admin@Branch-1-cli> show orgs org Tenant1 sd-wan sla-monitor status
                                  LOCAL  REMOTE
                                  WAN    WAN
            PATH    FWD   LOCAL    REMOTE   LINK  LINK  ADAPTIVE    DAMP    DAMP  CONN
LAST
SITE NAME   HANDLE  CLASS WAN LINK WAN LINK ID    ID    MONITORING  STATE   FLAPS
STATE FLAPS FLAPPED
-------------------------------------------------------------------------------------------------------------------------------------------
Controller-1 69888   fc_nc MPLS     MPLS     1     1     disable     disable 0     up    1     3d01h39m
             74240   fc_nc Internet Internet 2     2     disable     disable 0     up    1     3d01h39m
Hub-1        7213316 fc_ef MPLS     MPLS     1     1     suspend     disable 0     up    1     2d21h02m
             7217668 fc_ef Internet Internet 2     2     suspend     disable 0     up    1     2d21h02m
Hub-2        7278852 fc_ef MPLS     MPLS     1     1     suspend     disable 0     up    1     2d21h00m
             7283204 fc_ef Internet Internet 2     2     suspend     disable 0     up    1     2d21h00m
```

Because the hub nodes re-advertise the spoke branch prefixes, the spoke branches learn all the spoke prefixes. The

next-hop IP address for the spoke branches is the hub's loopback TVI address.

The following output from the Branch-1 VRF routes table shows a deployment with two hubs. The hub that you configure with a higher priority is the one that maintains the active route

```
admin@Branch-1-cli> show route routing-instance Tenant1-LAN-VR
Routes for Routing instance : Tenant1-LAN-VR AFI: ipv4 SAFI: unicast
Codes: E1 - OSPF external type 1, E2 - OSPF external type 2
IA - inter area, iA - intra area,
L1 - IS-IS level-1, L2 - IS-IS level-2
N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
RTI - Learnt from another routing-instance
+ - Active Route

Prot   Type   Dest Address/Mask   Next-hop     Age        Interface name
----   ----   -----------------   --------     ---        --------------
conn   N/A   +192.168.1.0/24      0.0.0.0      1d23h16m   vni-0/2.0
local  N/A   +192.168.1.1/32      0.0.0.0      1d23h16m   directly connected
BGP    N/A   +192.168.2.0/24      10.0.40.110  03:26:28   Indirect
BGP    N/A    192.168.2.0/24      10.0.40.111  03:26:28   Indirect
BGP    N/A   +192.168.3.0/24      10.0.40.110  1d23h16m   Indirect
BGP    N/A    192.168.3.0/24      10.0.40.111  1d23h16m   Indirect
BGP    N/A   +192.168.4.0/24      10.0.40.110  1d23h16m   Indirect
BGP    N/A    192.168.4.0/24      10.0.40.111  1d23h16m   Indirect
BGP    N/A   +192.168.10.0/24     10.0.40.110  1d23h16m   Indirect
BGP    N/A    192.168.10.0/24     10.0.40.111  1d23h16m   Indirect
```

The following CLI output shows an example of the spoke routes advertised by Branch-1:

```
admin@Branch-1-cli> show route table l3vpn.ipv4.unicast advertising-protocol bgp
Routes for Routing instance : Tenant1-Control-VR AFI: ipv4 SAFI: unicast

Routing entry for 192.168.1.0/24
   Peer Address         : 10.0.40.1
   Route Distinguisher  : 2L:2
   Next-hop             : 10.0.40.101
   VPN Label            : 24704
   Local Preference     : 110
   AS Path              : N/A
   Origin               : Igp
   MED                  : 0
   Community            : [ 8000:1 8001:110 8002:111 ]
   Extended community   : [ target:2L:2 ]
```

The community string 8000:1 marks the spokes routes so that the BGP import policy on the spokes can identify them, and in this case, it rejects routes starting with this community string. The following CLI output is an example of a spoke route advertised by Hub-1:

```
admin@Hub-1-cli> show route table l3vpn.ipv4.unicast advertising-protocol bgp

Routing entry for 192.168.2.0/24
   Peer Address         : 10.0.40.1
```

```
Route Distinguisher  : 16002L:110
Next-hop             : 10.0.40.110
VPN Label            : 24705
Local Preference     : 100
AS Path              : N/A
Origin               : Incomplete
MED                  : 0
Community            : [ 8000:0 8000:1 8001:110 8002:111 8009:8009 ]
Extended community   : [ target:16002L:0 target:16002L:110 ]
```

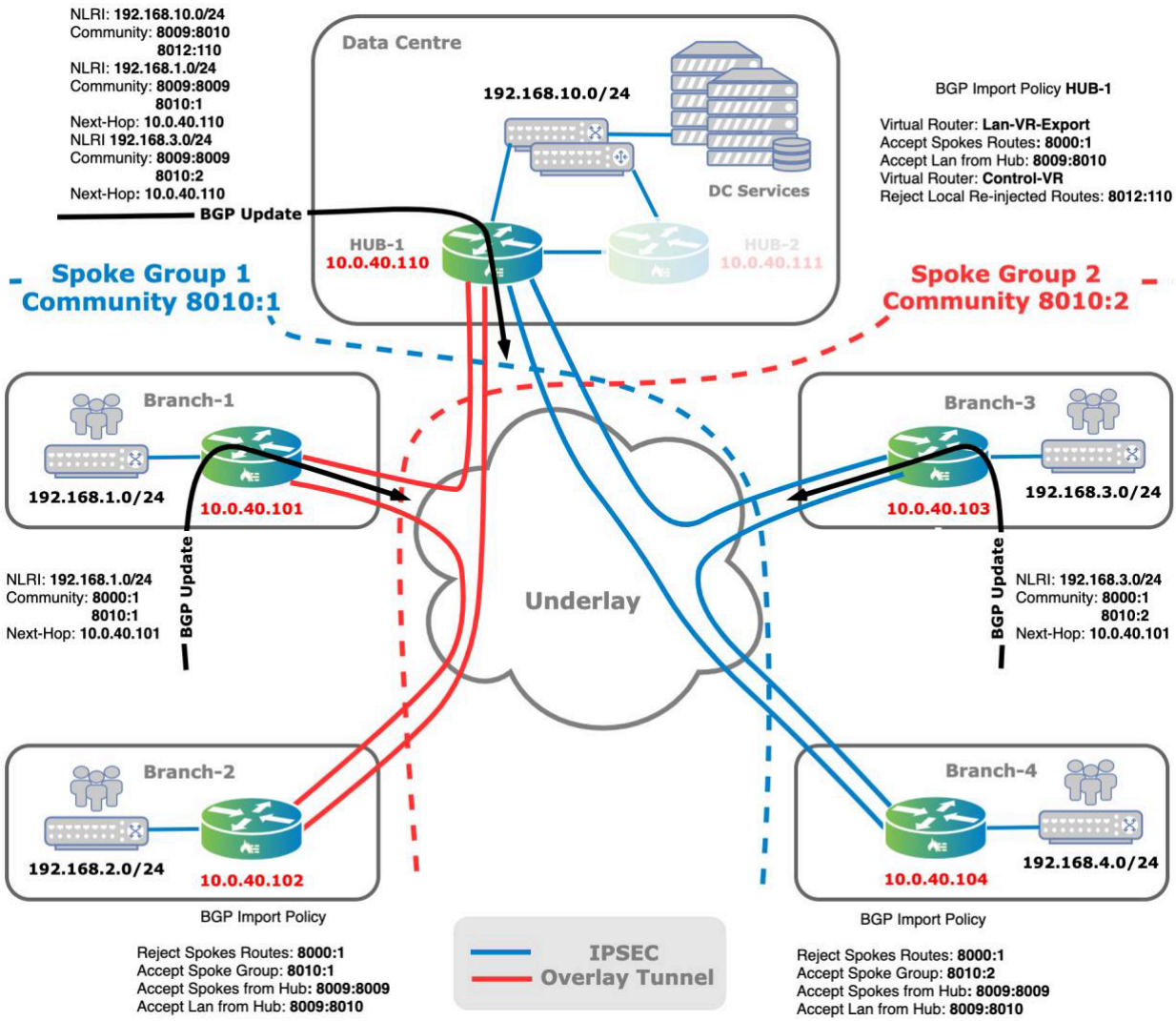
The community string 8009:8009 marks the spokes routes advertised by hubs so that the BGP import policy can identify them, and in this case, it accepts these routes, which have the hub as the next hop. The community string 8009:8010 marks direct LAN route from hubs, which the BGP import policy also accepts.

In this topology, Hub-1 is configured with a higher priority than Hub-2. This configuration explains why there are two entries in the route table for each prefix and why Hub-1 is the preferred next hop. Having two hubs provides redundancy, because Hub-2 is used when Hub-1 is not reachable.

The BGP import policy uses the extended-community attribute to accept routes from the hubs and to set a higher local preference for Hub-1. The extended target string 16002L:110 is derived from site ID 110, which is Hub-1.

Spoke-to-Spoke Direct and Partial Mesh

In a partial-mesh topology, some nodes are directly attached to each other, while other nodes are attached only to one or two nodes. You can select this topology when there are geographically dispersed sites in the same region that you want to communicate directly with each other, and when you want inter-regional traffic to transit through hub branches or when there is a high level of traffic exchanged between specific sites. The following figure illustrates a partial-mesh topology.



For a spoke-to-spoke–direct topology, you use spoke groups. Branches within the same spoke group can communicate directly with each other, and they use hubs to reach branches in different spoke groups. In the topology shown above, Branch-1 and Branch-2 communicate with each other directly, but to reach Branch-3 the next hop is Hub-1. The hubs are connected using a full-mesh topology.

It is recommended that you deploy a spoke-to-spoke–direct topology whenever feasible, so that you can use spoke groups to provide redundancy and flexible meshing of branches.

The following CLI output shows the SLA monitoring view on Branch-1:

```
admin@Branch-1-cli> show orgs org Tenant1 sd-wan sla-monitor status
                                    LOCAL REMOTE
                                    WAN   WAN
          PATH    FWD    LOCAL     REMOTE    LINK  LINK    ADAPTIVE    DAMP     DAMP   CONN
LAST
SITE NAME     HANDLE   CLASS  WAN LINK  WAN LINK  ID    ID      MONITORING  STATE    FLAPS
STATE  FLAPS  FLAPPED
```

```
-------------------------------------------------------------------------------------------------------------------
Branch-2     6689028 fc_ef MPLS     MPLS     1    1       active    disable 0     up     1     00:04:59
             6693380 fc_ef Internet Internet 2    2       active    disable 0     up     1     00:04:59
Controller-1 69888   fc_nc MPLS     MPLS     1    1       disable   disable 0     up     5     1d22h41m
             74240   fc_nc Internet Internet 2    2       disable   disable 0     up     1     1d22h45m
Hub-1        7213316 fc_ef MPLS     MPLS     1    1       suspend   disable 0     up     7     02:48:59
             7217668 fc_ef Internet Internet 2    2       suspend   disable 0     up     3     02:48:58
Hub-2        7278852 fc_ef MPLS     MPLS     1    1       suspend   disable 0     up     5     02:48:41
             7283204 fc_ef Internet Internet 2    2       suspend   disable 0     up     3     02:48:41
```

SLA monitoring is performed on the paths towards Hub-1, Hub-2, and Branch-2, because these belong to the same spoke group. SLA monitoring is not performed on branches in different spoke groups.

The following CLI output shows the Branch-1 VRF route table:

```
admin@Branch-1-cli> show route routing-instance Tenant1-LAN-VR
Routes for Routing instance : Tenant1-LAN-VR AFI: ipv4 SAFI: unicast
Codes: E1 - OSPF external type 1, E2 - OSPF external type 2
IA - inter area, iA - intra area,
L1 - IS-IS level-1, L2 - IS-IS level-2
N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
RTI - Learnt from another routing-instance
+ - Active Route

Prot   Type  Dest Address/Mask   Next-hop        Age     Interface name
----   ----  ----------------    --------        ---     --------------
conn   N/A  +192.168.1.0/24      0.0.0.0         3d18h59m vni-0/2.0
local  N/A  +192.168.1.1/32      0.0.0.0         3d18h59m directly connected
BGP    N/A  +192.168.2.0/24      10.0.40.102     00:25:29 Indirect
BGP    N/A   192.168.2.0/24      10.0.40.110     00:25:30 Indirect
BGP    N/A   192.168.2.0/24      10.0.40.111     00:25:30 Indirect
BGP    N/A  +192.168.3.0/24      10.0.40.110     00:20:49 Indirect
BGP    N/A   192.168.3.0/24      10.0.40.111     00:20:48 Indirect
BGP    N/A  +192.168.4.0/24      10.0.40.110     00:24:56 Indirect
BGP    N/A   192.168.4.0/24      10.0.40.111     00:24:56 Indirect
BGP    N/A  +192.168.10.0/24     10.0.40.110     03:09:26 Indirect
BGP    N/A   192.168.10.0/24     10.0.40.111     02:55:29 Indirect
```

A spoke-to-spoke direct topology incorporates additional redundancy. The CLI output above shows three route table entries for 192.168.2.0/24. These routes are advertised by Branch-2 as well as by Hub-1 and Hub-2. When there are underlay connectivity issues between Branch-1 and Branch-2, the hubs provide a redundant path to reach the Branch-2 prefixes.

Routes for branches in different spoke groups are only reachable through the hubs.

To create the expected topology, you use BGP community strings and import policies to accept or reject routes.

The following CLI output shows the prefixes advertised by Branch-2:

```
admin@Branch-2-cli> show route table l3vpn.ipv4.unicast advertising-protocol bgp

Routes for Routing instance : Internet-Transport-VR AFI: ipv4 SAFI: unicast
```

```
Routes for Routing instance : MPLS-Transport-VR AFI: ipv4 SAFI: unicast
Routes for Routing instance : Tenant1-Control-VR AFI: ipv4 SAFI: unicast

Routing entry for 192.168.2.0/24
   Peer Address         : 10.0.40.1
   Route Distinguisher  : 2L:2
   Next-hop             : 10.0.40.102
   VPN Label            : 24704
   Local Preference     : 110
   AS Path              : N/A
   Origin               : Igp
   MED                  : 0
   Community            : [ 8000:1 8001:110 8002:111 8010:1 ]
   Extended community   : [ target:2L:2 ]
```

Here, the community string 8010:1 corresponds to spoke Group-1. The community string is a unique BGP community that is associated with the spoke group and that you assign during the Workflow configuration. Based on this community string, the import policy based is pushed to the spokes that are in the same spoke group.

In the topology shown in the figure above, Hub-1 has a higher priority. You configure an import policy that manipulates the Local-Pref attribute (Branch-2 > Hub-1 > Hub-2) to prefer the routes advertised by Hub-1. The following output shows the Local Preference configuration on Branch-1:

```
admin@Branch-1-cli> show route table l3vpn.ipv4.unicast receive-protocol bgp 192.168.2.0

Routes for Routing instance : Internet-Transport-VR AFI: ipv4 SAFI: unicast

Routes for Routing instance : MPLS-Transport-VR AFI: ipv4 SAFI: unicast

Routes for Routing instance : Tenant1-Control-VR AFI: ipv4 SAFI: unicast

Routing entry for 192.168.2.0/24
   Peer Address         : 10.0.40.1
   Route Distinguisher  : 2L:2
   Next-hop             : 10.0.40.102
   VPN Label            : 24704
   Local Preference     : 110
   AS Path              : N/A
   Origin               : Igp
   MED                  : 0
   Community            : [ 8000:1 8001:110 8002:111 8009:8009 8010:1 ]
   Extended community   : [ target:2L:2 ]
   Preference           : Default

Routing entry for 192.168.2.0/24
   Peer Address         : 10.0.40.1
   Route Distinguisher  : 16002L:110
   Next-hop             : 10.0.40.110
   VPN Label            : 24705
   Local Preference     : 102
   AS Path              : N/A
```

```
    Origin              : Incomplete
    MED                 : 0
    Community           : [ 8000:0 8000:1 8001:110 8002:111 8009:8009 8009:8009 8010:1 ]
    Extended community  : [ target:16002L:0 target:16002L:110 ]
    Preference          : Default

Routing entry for 192.168.2.0/24
    Peer Address        : 10.0.40.1
    Route Distinguisher : 16002L:111
    Next-hop            : 10.0.40.111
    VPN Label           : 24705
    Local Preference    : 101
    AS Path             : N/A
    Origin              : Incomplete
    MED                 : 0
    Community           : [ 8000:0 8000:1 8001:110 8002:111 8009:8009 8009:8009 8010:1 ]
    Extended community  : [ target:16002L:0 target:16002L:111 ]
    Preference          : Default
```

The following CLI output shows the entries in the Branch-1 route table:

```
admin@Hub-1-cli> show route routing-instance Tenant1-LAN-VR

Routes for Routing instance : Tenant1-LAN-VR AFI: ipv4 SAFI: unicast

Codes: E1 - OSPF external type 1, E2 - OSPF external type 2
IA - inter area, iA - intra area,
L1 - IS-IS level-1, L2 - IS-IS level-2
N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
RTI - Learnt from another routing-instance
+ - Active Route

Prot   Type  Dest Address/Mask  Next-hop       Age        Interface name
----   ----  -----------------  --------       ---        --------------
BGP    N/A  +192.168.1.0/24     10.0.40.101     1d03h20m   Indirect
BGP    N/A   192.168.1.0/24     10.0.40.111     1d03h20m   Indirect
BGP    N/A  +192.168.2.0/24     10.0.40.102     1d03h20m   Indirect
BGP    N/A   192.168.2.0/24     10.0.40.111     1d03h20m   Indirect
BGP    N/A  +192.168.3.0/24     10.0.40.103     1d03h19m   Indirect
BGP    N/A   192.168.3.0/24     10.0.40.111     1d03h19m   Indirect
BGP    N/A  +192.168.4.0/24     10.0.40.104     1d03h19m   Indirect
BGP    N/A   192.168.4.0/24     10.0.40.111     1d03h19m   Indirect
BGP    N/A   192.168.10.0/24    10.0.40.111     1w4d03h    Indirect
conn   N/A  +192.168.10.0/24    0.0.0.0         1w4d03h    vni-0/2.0
local  N/A  +192.168.10.1/32    0.0.0.0         1w4d03h    directly connected
```

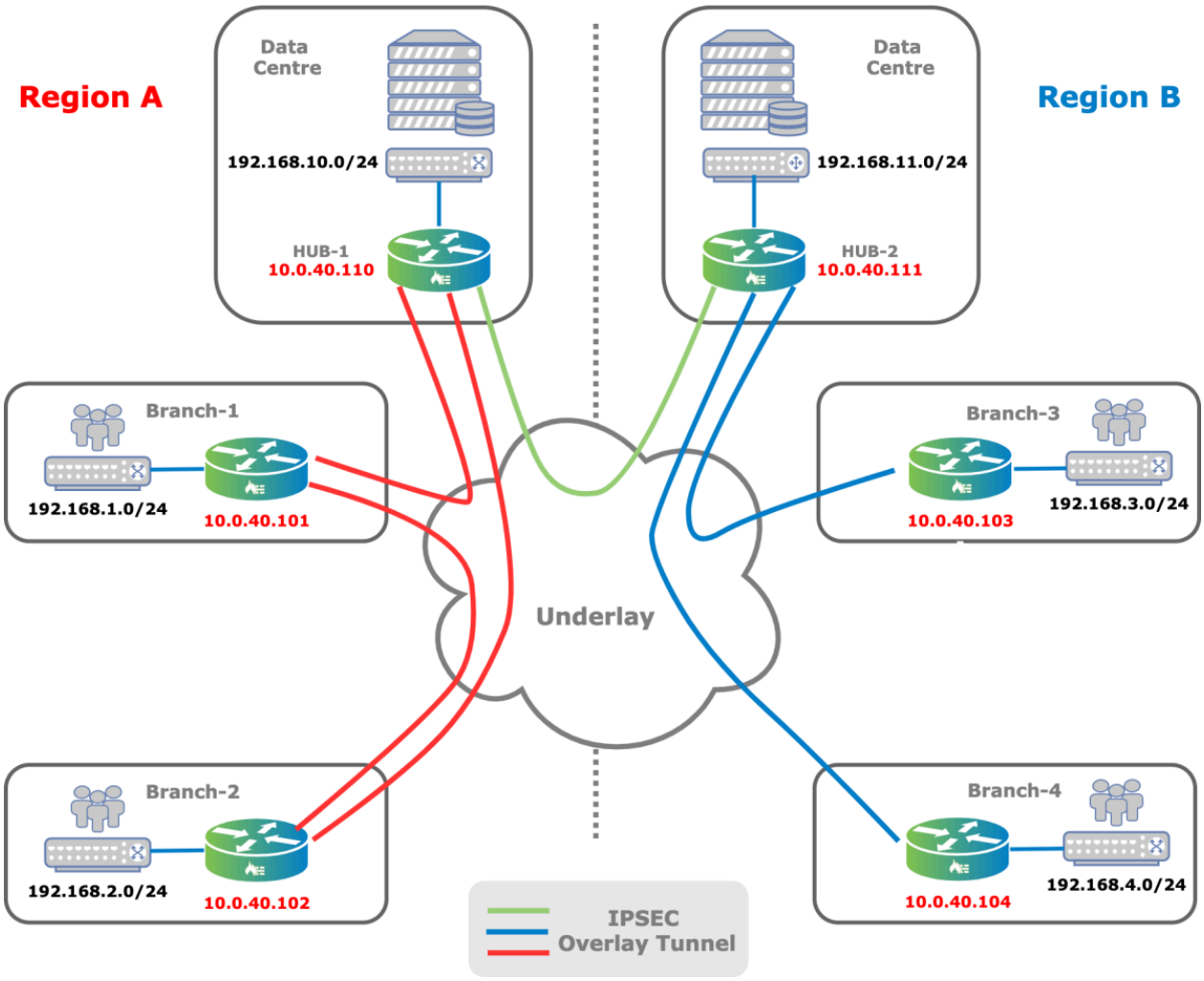
Spokes routes are directly reachable from hubs, and the backup is to use the remote hubs. For example, in the output above, the prefixes learned from Branch-1 (the first two entries in the roue table) show that the direct route is active.

Spoke-Hub-Hub-Spoke (Regional-Mesh) Topology

In a spoke-hub-hub-spoke (SHHS) topology, also called a regional-mesh topology, you group hub and branch devices

by region. Within a particular region, the topology can be anything—full mesh, partial mesh, or hub and spoke—and branches communicate based on the selected topology. When a branch in one region wants to communicate with a branch in another region, the communication transits through regional hubs.

You can select this topology when there is geographical separation and when regional WAN transport networks are available and hubs between regions use the company backbone or high-bandwidth WAN links. The following figures shows two regional networks with different topologies: Region A uses a spoke-to-spoke–direct topology, and Region B uses a spoke-to-spoke topology through a hub. In this example, communication between Branch-1 and Branch-3 uses Hub-1 and Hub-2, but the branches can use any regional hubs in the local region. This topology demonstrates how you can use the SHHS topology in regional networks.



The following CLI output shows the SD-WAN topology view from Branch-1. Because Region A uses a full-mesh topology, the output shows only regional hubs and branches.

```
admin@Branch-1-cli> show orgs org Tenant1 sd-wan brief
          SITE  MANAGEMENT                    CONNECTIVITY  IS
SITE NAME    ID   IP       TYPE    UP TIME       STATUS      CTRLR
```

```
--------------------------------------------------------------------------------
Branch-1      101   10.0.40.101 local   12d:19h:48m:37s  -            no
Branch-2      102   10.0.40.102 remote  22m:34s          Connected    no
Controller-1  1     10.0.40.1   remote  6d:21h:54m:14s   Connected    yes
Hub-1         110   10.0.40.110 remote  12d:19h:47m:28s  Connected    no
```

The following CLI output shows the SD-WAN topology view from Branch-3. Branch-3 is the only hub branch, which is normal in a spoke-to-spoke through a hub topology:

```
admin@Branch-3-cli> show orgs org Tenant1 sd-wan brief
                SITE        MANAGEMENT                      CONNECTIVITY     IS
SITE NAME       ID          IP          TYPE     UP TIME        STATUS         CTRLR
--------------------------------------------------------------------------------
Branch-3        103         10.0.40.103 local    12d:22h:41m:34s -              no
Controller-1    1           10.0.40.1   remote   7d:21h:59m:35s  Connected      yes
Hub-2           111         10.0.40.111 remote   3h:14m:8s       Connected
```

The following CLI output shows the entries in the Branch-1 VRF route table. The two prefixes in both regions are highlighted. The prefix 192.168.2.0/24, for Branch-2, is a direct route towards Branch-2 through Hub-1 for redundancy, and it is the less preferred path. The preferred route is the direct route that uses Branch-2 as the next hop and that is the active route, as indicated by the plus sign (+). The prefix 192.168.3.0/24 is for Branch-3, and the route table entries points to Hub-1.

```
admin@Branch-1-cli> show route routing-instance Tenant1-LAN-VR

Routes for Routing instance : Tenant1-LAN-VR AFI: ipv4 SAFI: unicast
Codes: E1 - OSPF external type 1, E2 - OSPF external type 2
IA - inter area, iA - intra area,
L1 - IS-IS level-1, L2 - IS-IS level-2
N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
RTI - Learnt from another routing-instance
+ - Active Route

Prot    Type    Dest Address/Mask     Next-hop     Age     Interface name
----    ----    -----------------     --------     ---     --------------
conn    N/A     +192.168.1.0/24       0.0.0.0      1w5d22h     vni-0/2.0
local   N/A     +192.168.1.1/32       0.0.0.0      1w5d22h     directly connected
BGP     N/A     +192.168.2.0/24       10.0.40.102 03:15:57     Indirect
BGP     N/A      192.168.2.0/24       10.0.40.110 03:08:36     Indirect
BGP     N/A     +192.168.3.0/24       10.0.40.110 00:01:06     Indirect
BGP     N/A     +192.168.4.0/24       10.0.40.110 00:01:13     Indirect
BGP     N/A     +192.168.10.0/24      10.0.40.110 03:08:35     Indirect
BGP     N/A     +192.168.11.0/24      10.0.40.110 03:08:36     Indirect
```

The following CLI output shows the entries in the Branch-3 VRF route table. The highlighted output shows two prefixes in both regions and that there is no difference in the treatment between regions. Region B uses a spoke-to-spoke through a hub topology. The prefix 192.168.1.0/24 is behind Branch-1, and the prefix 192.168.4.0/24 prefix is behind Branch-4.

```
admin@Branch-3-cli> show route routing-instance Tenant1-LAN-VR
```

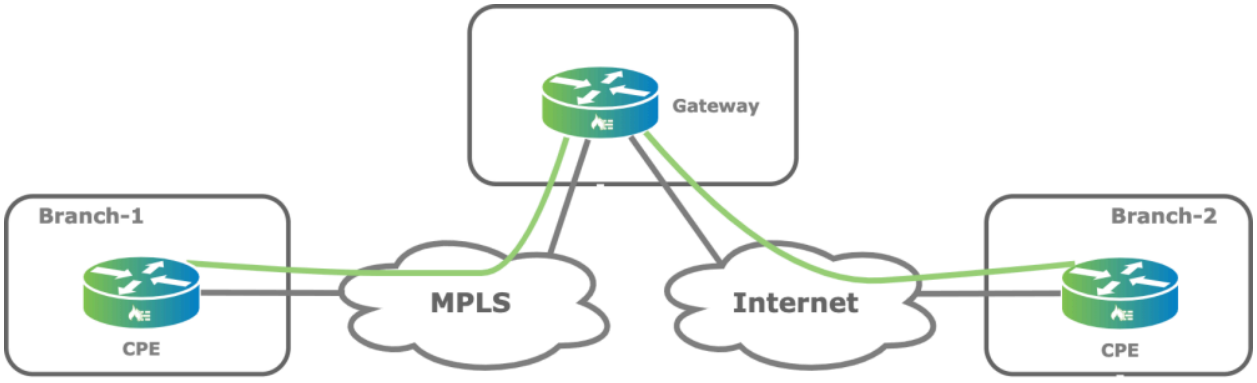
```
Routes for Routing instance : Tenant1-LAN-VR AFI: ipv4 SAFI: unicast

Codes: E1 - OSPF external type 1, E2 - OSPF external type 2
IA - inter area, iA - intra area,
L1 - IS-IS level-1, L2 - IS-IS level-2
N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
RTI - Learnt from another routing-instance
+ - Active Route

Prot   Type  Dest Address/Mask   Next-hop        Age     Interface name
----   ----  ----------------    --------        ---     --------------
BGP    N/A  +192.168.1.0/24      10.0.40.111     02:46:42 Indirect
BGP    N/A  +192.168.2.0/24      10.0.40.111     02:46:42 Indirect
conn   N/A  +192.168.3.0/24      0.0.0.0         1w5d22h  vni-0/2.0
local  N/A  +192.168.3.1/32      0.0.0.0         1w5d22h  directly connected
BGP    N/A  +192.168.4.0/24      10.0.40.111     00:04:31 Indirect
BGP    N/A  +192.168.10.0/24     10.0.40.111     02:46:43 Indirect
BGP    N/A  +192.168.11.0/24     10.0.40.111     02:46:43 Indirect
```

Connecting Sites over Disjointed Underlay Networks

You can use gateways to connect to sites over disjointed underlay networks. In a disjointed underlay, two sites do not have a common underlay that allows them to communicate directly. An example is one site that has only Internet connectivity and a seconnd site that has only MPLS connectivity. Another example is one in which a site provisioned with internet and MPLS links communicates with a site that has only an MPLS link. When the internet link on the first site becomes unavailable, then without a gateway, that site loses connectivity to the MPLS-only site. Another common scenario for using a disjointed underlay network is when NAT traversal issues do not allow two internet-connected branches to connect directly. A third device, which is the gateway, can interconnect the two branches, as shown in the following figure.



You can connect the sites by configuring one of the following in on the gateway device:

- Configure the branch as the gateway and configure the branches in the spoke-to-spoke-direct topology in the

Workflow templates

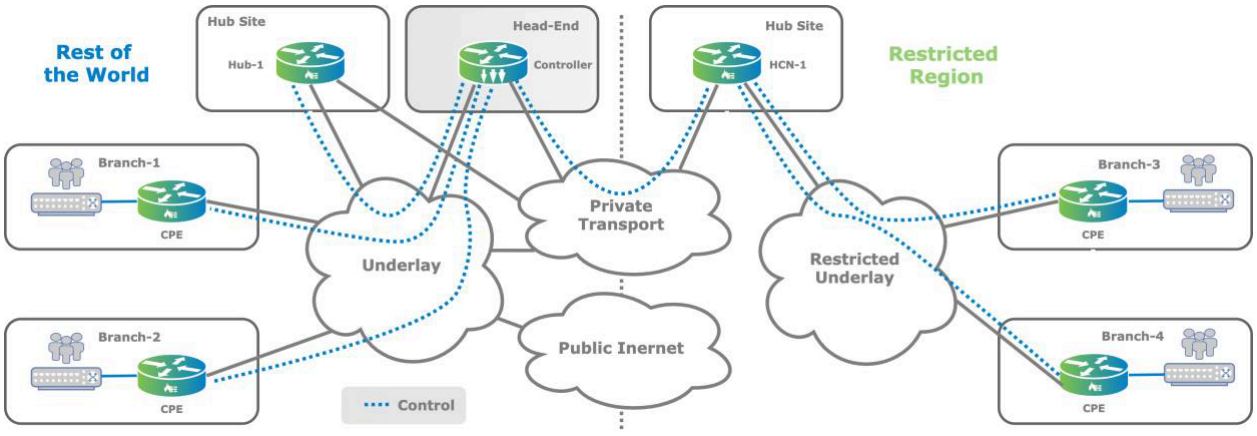
- Advertise a summary route from the gateway. The gateway can be any branch site, and you do not need to configure the branch as a gateway in the Workflow templates.
- Advertise a default route from the gateway. The gateway can be any branch site, and you do not need to configure the branch as a gateway in the Workflow templates.

SD-WAN Topologies for Geographically Isolated Regions

In some regions of the world, government place restriction on encrypted (IPsec) traffic to block it from going in and out of the country. Inside the country, traffic is typically not affected, but blocking the IPsec traffic can isolate the SD-WAN regional network. For these situations, you can create separated SD-WAN islands in which each SD-WAN domain operates within itself and has no visibility into or information about other domains. To interconnect these SD-WAN islands, you use IPsec or other connection options, which you must provision and manage outside the SD-WAN domain. You must consider the stitching complexities on the boundary node, which acts like a network-to-network interface (NNI) point.

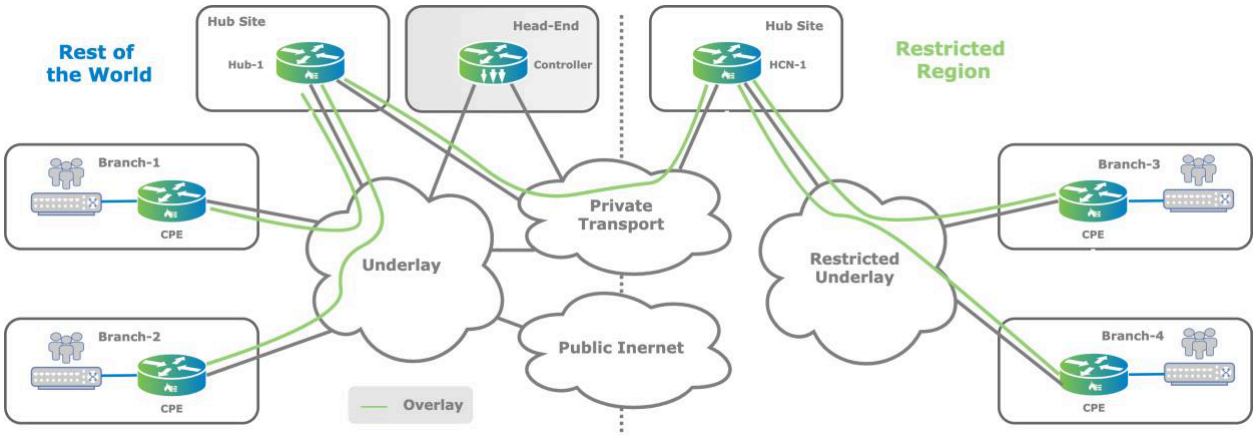
The separation prevents the control planes of each SD-WAN island from exchanging information with each other, and there is no connectivity between domains. Creating tenants and service templates across SD-WAN islands become a complex task. The SHHS topology provides a solution that is highly scalable, compartmentalized, fully automated, and easy to deploy and operate.

The following figure shows a control plane connection for an isolated region. An IBGP session is established between the headend Controller node and the hub Controller node (HCN-1) site in the restricted region and is used to exchange SD-WAN route information over a suitable transport, in this case a private transport, that allows IPsec traffic. The data plane flow between a branch in the restricted region and a branch with no restriction traverses the hub sites using IPsec tunnels. Note that the headend Controller node is not part of the data plane communication.



You can apply the SD-WAN network topology shown in the following figure to restricted regions. Branches in this region create an SD-WAN island in the form of a restricted underlay that has no connectivity with the rest of the world. An approved transport must be provided by a government-approved service provider to allow IPsec traffic, which can be a

Layer 2 service (such as Layer 2 VPN and VPLS) or a Layer 3 service. This service is represented by the private transport.



For such a topology, the VOS software provides a hierarchical controller structure called a hub-controller node (HCN) device type. An HCN can interconnect the control planes of SD-WAN islands that may be using their own Controller instances. The MP-BGP–based control plane assigns separate community values to represent each SD-WAN island, and IBGP sessions are established between the regional HCNs and the Versa headend Controller node, unifying the control plane. HCNs can consolidate Controller and hub functionalities into one node to preserve resources on the control side of the network. HCN nodes exchange information with spokes and implement the data plane functions of hub nodes.

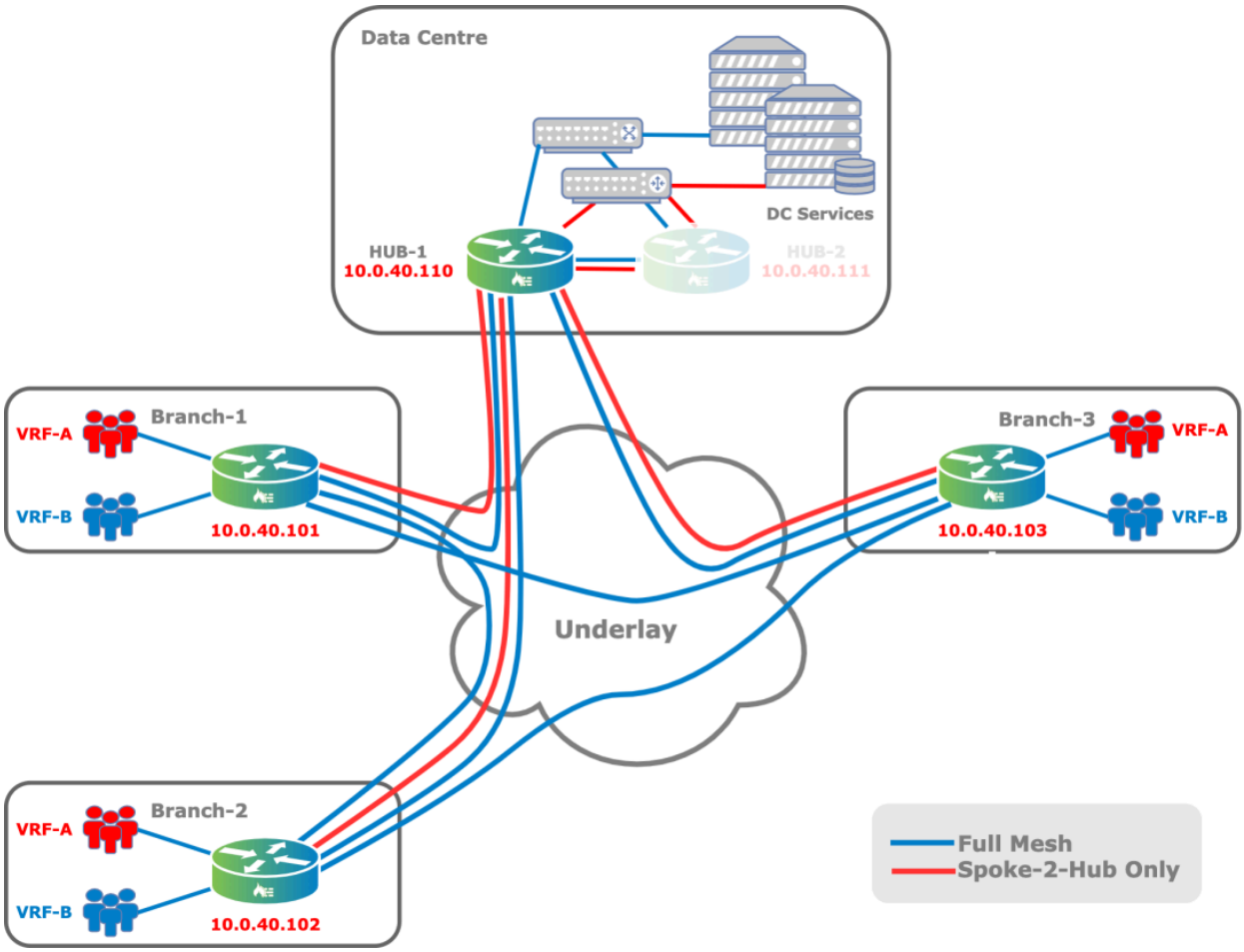
You can deploy HCNs in active–active mode for the control plane to maximize uptime and ease of serviceability. In active–active mode, multitenancy is preserved across the topology using the organization and sub-organization structure. Data plane redundancy is provided by BGP next-hop route resolution. You then provision tenants and services, starting from the spokes of one SD-WAN island to the other SD-WAN island, using the expanded templates for SHHS deployment. After the control plane is fully functional, SD-WAN paths are automatically established between Branch-1 and Hub-1, Hub-1 and HCN-1, and HCN-1 and Branch-3 in a hop-by-hop manner, in both directions, as shown in the figure above. These paths can be separate SD-WAN tunnels or shared SD-WAN tunnels that may already be present between spokes and hub routers. All data plane operations are handled automatically and require no manual configuration. After the end-to-end data paths are set up, users can communicate seamlessly between the spokes of separate SD-WAN islands.

Multi-VRF (Multitenant) Topologies

The examples discussed thus far in this article use a single VRF (a single tenant) to explain the differences between topologies. The Versa Networks SD-WAN solution is highly flexible and allows you to define and establish different topologies at the VRF, or tenant, level. The default Versa SD-WAN model provisions a full forwarding mesh using IP prefix advertisement in MP-BGP. You can configure the VPN topologies discussed in previous sections using Director Workflows.

The following figure shows an SD-WAN topology with two VRFs in one organization. VRF-A (red) is configured for

spoke-to-hub only, and VRF-B (blue) uses a full-mesh topology. This principle can apply to multitenant branches in which each organization's topology may be different.



Best Practices for SD-WAN Topologies

The following are a few best practices for various SD-WAN topologies:

- The full-mesh topology is the default option in Director Workflows, and you can deploy a full mesh without restrictions for VPNs with up to 100 sites. If there are more sites, SLA monitoring consumes a significant amount bandwidth. This means that branches with low-bandwidth connections must assign a relatively high proportion of their available bandwidth for SLA monitoring traffic. For low-bandwidth branches, it is recommended that you deploy a spoke-to-spoke through a hub topology so that SLA monitoring is performed only towards the hubs.
- Spoke-to-spoke-direct topologies are recommended over full-mesh topologies. For many use cases, having a hub node is preferable in the topology, for example, to avoid disjointed underlays caused by circuit failures or NAT traversal issues. Also, spoke-to-spoke–direct topology provides better support for automatically importing routes from LAN adjacent networks, which avoids manual configuration of redistribution policies.
- Use a distinct WAN network name on the hub sites.

Software Release Information

Releases 20.2 and later support all content described in this article.

Additional Information

Configure Regional Hub-and-Controller Nodes for SHHS Topologies
Configure SLA Monitoring for SD-WAN Traffic Steering