Cisco SD-WAN Multi-Region Fabric

Table 1: Feature History

Feature Name	Release Information	Description
Multi-Region Fabric (also Hierarchical SD-WAN)	Cisco vManage Release 20.7.1	Cisco SD-WAN Multi-Region Fabric provides the ability to divide the architecture of the Cisco SD-WAN overlay network into multiple regional networks that operate distinctly from one another, and a central core-region network for managing inter-regional traffic. The hierarchical architecture enables you to use different traffic transport service providers for each region, and for the central core-region network, to optimize cost and traffic performance. It also simplifies traffic configuration for some scenarios, and provides a robust, adaptive topology that can help prevent routing failures in specific network scenarios.
Re-Origination Dampening	Cisco IOS XE Release 17.9.1a	In networks experiencing instability, TLOCs and bidirectional forwarding detection (BFD) tunnels may alternate repeatedly between being available and unavailable. This causes the overlay management protocol (OMP) to repeatedly withdraw and re-originate routes. This churn adversely affects Cisco vSmart controller performance. Adding a delay before re-originating routes that have gone down repeatedly prevents excessive churn, and prevents this type of network instability from diminishing Cisco vSmart controller performance.

Feature Name	Release Information	Description
Cisco vSmart Controller Optimizations	Cisco SD-WAN Controllers Release 20.10.1	There are two optimizations of Cisco vSmart controller performance: • Cisco vSmart controller optimization of outbound control policy: This feature helps to optimize Cisco vSmart controller performance by streamlining the evaluation of outbound control policies. The controller evaluates the policy only once for all peers rather than reevaluating for each peer. • Cisco vSmart controller resistance to TLOC flapping: When TLOCs cycle between unavailable and available, called flapping, they cause Cisco vSmart controllers to continually readvertise the list of routes to devices in the network. This degrades the performance of Cisco vSmart controllers and devices in the network. To address this and improve performance, Cisco vSmart controllers isolate the disruption to devices that use the same control policy, leaving other devices unaffected.

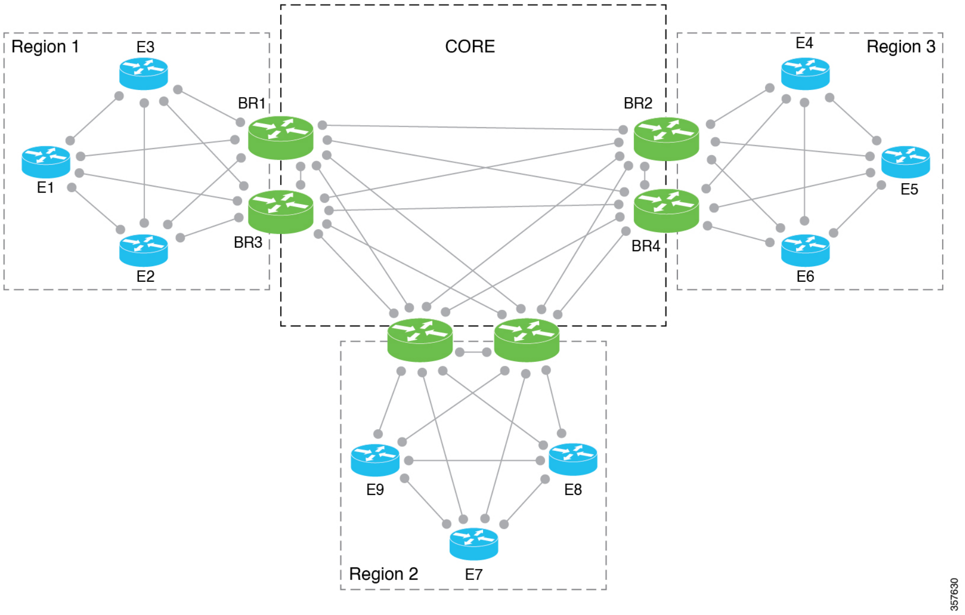
Information About Multi-Region Fabric

Multi-Region Fabric (formerly Hierarchical SD-WAN) provides the option to divide the architecture of the Cisco SD-WAN overlay network into the following:

- A core overlay network: This network, called region 0, consists of border routers (BR in the illustration below) that connect to regional overlays and connect to each other.
- One or more regional overlay networks: Each regional network consists of edge routers that connect to other edge routers within the same region, and can connect to core region border routers that are assigned to the region.

The following figure shows a core overlay network with six border routers (BR1 to BR6), two assigned to each of three regions. In the three regional overlay networks, edge routers connect only to other edge routers within the same region or to core border router assigned to the region.

Figure 1: Multi-Region Fabric Architecture



Intra-Region and Inter-Region Traffic

The division into regions creates a distinction between intra-region traffic and inter-region traffic.

- Intra-region traffic: Edge routers connect directly to other edge routers within the region.

  The traffic traverses direct tunnels between source devices and destination devices.

- Inter-region traffic: Edge routers in one region do not connect directly to edge routers in a different region. For inter-region traffic, the edge routers connect to core border routers, which forward the traffic to the core border routers assigned to the target region, and those border routers forward the traffic to the edge routers within the target region.

  The traffic traverses three tunnels between the source device and the destination device.

Disaggregated Transport

An important principle in Multi-Region Fabric is that after you define regions and a core-region network, you can arrange to use different traffic transport services for each region and for the core-region network.

In a common use case, the core region is used for traffic between distant geographic regions. In this scenario, the core region uses a premium transport service to provide the required level of performance and cost effectiveness for long-distance connectivity.

Network Topology

Multi-Region Fabric provides the flexibility to use different network topologies in different regions. For example, region 1 can use a full mesh of Cisco SD-WAN tunnels, while region 2 can use a hub-and-spoke topology, and Region3 can use a full mesh topology with dynamic tunnels.

We recommend using a full mesh of tunnels for the overlay topology of the core region (region 0). This means that each border router in the core region requires a tunnel to each other border router in the core. These direct tunnels provide optimal connectivity for forwarding traffic from one region to another.

The implementation of a full mesh topology minimizes the complexity of routing within the core overlay network. By contrast, partial mesh topology would require topology-aware routing to compute inter-region paths. For scaling limitations, see Restrictions for Multi-Region Fabric, on page 8.

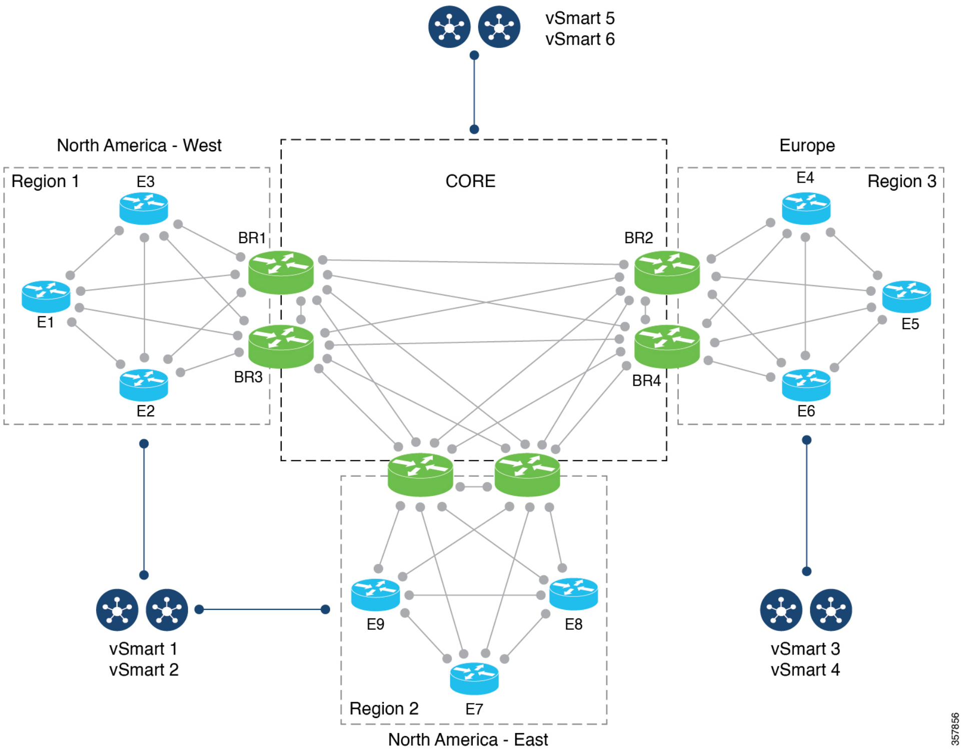
Distributed Cisco vSmart Controllers

Multi-Region Fabric enables you to assign Cisco vSmart controllers to serve specific regions. If your organization’s network contains only a small number of devices, a single Cisco vSmart controller, or typically a pair of Cisco vSmart controllers, can serve all regions in the network. For larger numbers of devices, we recommend that you assign Cisco vSmart controllers to serve specific regions.

Note the following for the example below:

- Cisco vSmart controllers 1 and 2 serve regions 1 and 2.
- Cisco vSmart controllers 3 and 4 serve region 3.
- Cisco vSmart controllers 5 and 6 serve the core region (region 0).

Figure 2: Cisco vSmart Controllers Serving Different Regions



Note For Cisco vSmart controller restrictions, see Restrictions for Multi-Region Fabric, on page 8.

Re-Origination Dampening

Minimum release: Cisco IOS XE Release 17.9.1a

In networks experiencing instability, TLOCs and bidirectional forwarding detection (BFD) tunnels may alternate repeatedly between being available and unavailable. This type of network stability may have various causes, including the following:

- Malfunctioning physical connections
- Network issues that interfere with connectivity
- Weak signals in a cellular network

The alternating between availability and unavailability can cause the overlay management protocol (OMP) operating on border routers and transport gateways to repeatedly withdraw routes that become unavailable and then re-originate the routes when they become available again. This churn propagates to the Cisco vSmart controllers managing the network, creating unnecessary demands on Cisco vSmart controller resources and diminishing performance.

To prevent network instability from diminishing Cisco vSmart controller performance, from Cisco IOS XE Release 17.9.1a, when border routers and transport gateways detect repeated problems with network stability, they introduce a delay before re-originating routes after the routes become available. This reduces unnecessary load on the Cisco vSmart controllers and keeps the control plane stable.

Re-origination dampening is enabled by default and does not require any configuration.

Cisco vSmart Controller Optimization of Outbound Control Policy

Beginning with Cisco SD-WAN Controllers Release 20.10.x, Cisco vSmart controllers use a caching feature to optimize performance when applying a control policy to multiple peers.

When a Cisco vSmart controller applies an outbound control policy to a peer, it evaluates each sequence (each of which specifies a match condition and an action) in the policy. For each action in the policy, the controller creates what is called an attribute, which represents the action. For example, if the action in a sequence is to set an OMP tag to 100, the Cisco vSmart controller generates an attribute for setting the OMP tag of a route to 100.

When the Cisco vSmart controller applies the policy to a peer in the outbound direction, for each path that is matched, the controller saves the action attribute to a cache. When the Cisco vSmart controller applies the same control policy to another peer, it does not have to evaluate the policy again. It can use the cached attributes. Minimizing the number of times the Cisco vSmart controller must evaluate a policy improves the performance of the controller.

You can confirm that this feature is operating on a Cisco vSmart controller by running the **show running-config omp** command on the controller. The output includes the following line:

```
outbound-policy-caching
```

On a Cisco vSmart controller, to view the attributes for a path (VPN and prefix), resulting from its evaluating a control policy, run the **show support omp rib vroute** *vpn*:*prefix* **detail** command, and view the RIB-CACHE sections of the output, as shown in the following example:

```
vsmart#show support omp rib vroute 1:192.168.30.0/24 detail | begin RIB-CACHE
  RIB-CACHE-ENTRY: (0xc733cb0), Policy-name: sc_test, Policy-seq-num: 100, RI-ID: 64, attr:
 0xc77bb40
    Attribute: (0xc77bb40), ROUTE-IPV4, Length: 1160, Ref: 6
      Flags: (0x8000c25) WEIGHT TLOC SITE-ID OVERLAY-ID ORIGIN ORIGINATOR
      Region-id: 65534, Secondary-Region-id: 65535, Orig-Access-Region-id: 65534,
Sub-Region-ID: 0, Pref: 0, Weight: 1, Tag: 0, Stale: 0 Version: 0, Restrict: 0, on-Demand:
 0, Domain: 0, BR-Preference: 0, Affinity:0, MRF-Route-Originator:None
      Distance: 0, Site-ID: 300, Carrier: 0, Query: 0, Gen-ID: 0x0, Border: 0 Overlay: 1
Site-Type: 0 0 0 0
      Originator: 172.16.255.30
      Origin: Protocol: connected[1], Sub-Type: none[0], Metric: 0
      TLOC: ((nil)) 172.16.255.30 : biz-internet : ipsec
    TE count 2
     TE: TLOC: 172.16.255.40 : mpls : ipsec, Label: 8389618, Pref: 0, Affinity: 0
     TE: TLOC: 172.16.255.40 : biz-internet : ipsec, Label: 8389618, Pref: 0, Affinity: 0
  RIB-CACHE-ENTRY: (0xc7deb20), Policy-name: sc_test, Policy-seq-num: 100, RI-ID: 70, attr:
 0xc7de3b0
    Attribute: (0xc7de3b0), ROUTE-IPV4, Length: 1160, Ref: 6
      Flags: (0x8000c25) WEIGHT TLOC SITE-ID OVERLAY-ID ORIGIN ORIGINATOR
      Region-id: 65534, Secondary-Region-id: 65535, Orig-Access-Region-id: 65534,
Sub-Region-ID: 0, Pref: 0, Weight: 1, Tag: 0, Stale: 0 Version: 0, Restrict: 0, on-Demand:
 0, Domain: 0, BR-Preference: 0, Affinity:0, MRF-Route-Originator:None
      Distance: 0, Site-ID: 300, Carrier: 0, Query: 0, Gen-ID: 0x0, Border: 0 Overlay: 1
Site-Type: 0 0 0 0
      Originator: 172.16.255.30
      Origin: Protocol: connected[1], Sub-Type: none[0], Metric: 0
      TLOC: ((nil)) 172.16.255.30 : mpls : ipsec
    TE count 2
     TE: TLOC: 172.16.255.40 : mpls : ipsec, Label: 8389618, Pref: 0, Affinity: 0
     TE: TLOC: 172.16.255.40 : biz-internet : ipsec, Label: 8389618, Pref: 0, Affinity: 0
```

Cisco vSmart Controller Resistance to TLOC Flapping

Sometimes TLOCs cycle between unavailable and available—this is called flapping. This flapping can degrade the performance of the Cisco vSmart controllers that advertise routes based on available TLOCs by causing the Cisco vSmart controllers to review and readvertise the routes repeatedly.

Beginning with Cisco SD-WAN Controllers Release 20.9.x, Cisco vSmart controllers minimize wasting resources when TLOCs in the network flap by creating an interest list of all of the TLOCs used in all control policies, cumulatively. The Cisco vSmart controller ignores flapping of TLOCs that are not on the interest list, meaning that if a TLOC that is not on the interest list experiences flapping, the Cisco vSmart controller does not have to readvertise the routes based on available TLOCs.

To further optimize Cisco vSmart controller performance, beginning with Cisco SD-WAN Controllers Release 20.10.x, the controllers maintain a separate TLOC interest list for each control policy, limiting the disruption caused by TLOC flapping. If a TLOC used by a specific control policy experiences flapping, it affects only the Cisco vSmart controller instances that make use of that control policy. This minimizes the performance impact of TLOC flapping on Cisco vSmart controller instances in the network.

You can use the **show support policy route-policy** command on a Cisco vSmart controller to show the TLOCs of interest for each control policy.

Note This strategy, introduced with Cisco SD-WAN Controllers Release 20.10.1, limits the number of TLOCs that you can include in a control policy to 64.

Benefits of Multi-Region Fabric

- Simplified policy design

- Prevention of certain traffic routing failures caused by policy—specifically, routing failures that can occur when a device responsible for one of the hops between the source and destination of a traffic flow is unavailable
- End-to-end encryption of inter-region traffic
- Flexibility to select the best transport for each region

  This flexibility can provide better performance for traffic across geographical regions. In the typical use case, an organization arranges to use premium traffic transport for the core region, providing better traffic performance across distant geographical regions.
- Better control over traffic paths between domains

  In some scenarios, it is advantageous to control how traffic is routed between domains, such as between geographical regions. The Multi-Region Fabric architecture simplifies this.

  For an example of how this is useful, see "Control over traffic paths between domains" in Use Cases for Multi-Region Fabric, on page 11.
- Enabling site-to-site traffic paths between disjoint providers

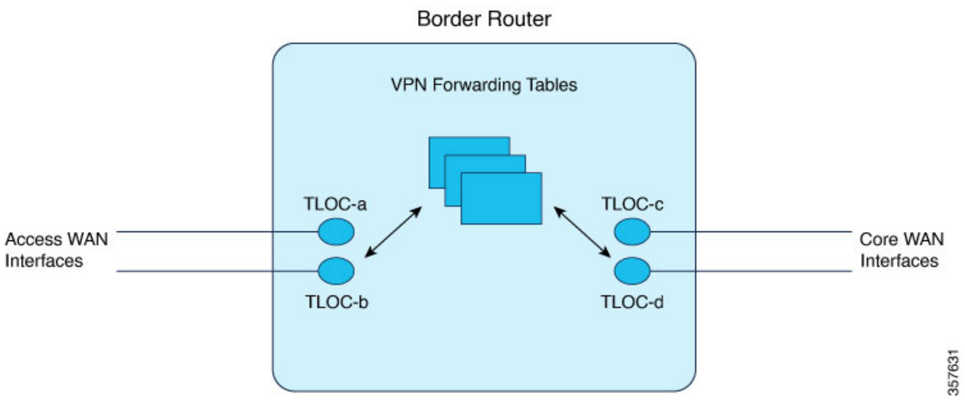
  The architecture of Multi-Region Fabric separates between edge routers and border routers. This enables you to establish site-to-site traffic paths between disjoint providers, which are two providers that cannot provide direct IP routing reachability between them. If each site connects to a core-region border router, then the core-region network can provide connectivity between the two sites.

  The core-region network can provide this connectivity because each border router has the following:

  - A set of (one or more) WAN interfaces to connect to regional edge routers
  - A separate set of WAN interfaces for connectivity within the core region

  The border router uses VPN forwarding tables to route traffic flows between its two sets of WAN interfaces.

  Figure 3: Disjoint Providers



- Optimized tunnel encapsulation

  You can use different types of tunnel encapsulation for the core region and for regional networks.

  For example, you might use IPsec tunnel encapsulation, which is encrypted, between a regional edge router and a core border router. If the core-region infrastructure does not require encryption, you might use generic routing encapsulation (GRE) for tunnels within the core region to provide better throughput.

The advantage of selecting the optimal tunnel encapsulation method for each region is better performance for inter-regional traffic.

Supported Devices for Multi-Region Fabric

- Edge router role: All Cisco IOS XE SD-WAN devices, all Cisco vEdge devices
- Border router role: All Cisco IOS XE SD-WAN devices

Prerequisites for Multi-Region Fabric

- Minimum software version for Cisco IOS XE SD-WAN devices: Cisco IOS XE Release 17.7.1a
- Minimum software version for Cisco vEdge devices: Cisco SD-WAN Release 20.7.1

Restrictions for Multi-Region Fabric

General Restrictions

- If you configure the devices in a network to use Multi-Region Fabric (assigning a region to each device), then all devices in the network must be configured to use Multi-Region Fabric. A device that is not configured for Multi-Region Fabric cannot connect to a device that is configured forMulti-Region Fabric.

Note Because of this restriction, the process of enabling Multi-Region Fabric for an existing network may temporarily disrupt connectivity between devices within the network.

- We recommend that you use a full mesh topology for the Multi-Region Fabric core-region network, with tunnels from each border router in the core region to each other border router in the core. While this has the advantage of simpler configuration, it limits the ability to scale number of border routers in the core region.
- Only Cisco IOS XE SD-WAN devices can have the border router role.

Note For an explanation of edge router and border router terminology, see Information About Multi-Region Fabric, on page 2.

- A border router can serve only one access region. (Regions other than the core region are called access regions.)

Routing Restrictions

Multi-Region Fabric does not support the following routing features:

- End-to-end SLA aware routing
- Multi-tenancy support for edge routers and border routers
- Overlay management protocol (OMP) route aggregation on border routers
- IP multicast on overlay support
- Per-region SLA policies. A border router always applies its region's SLA policy to traffic to and from other regions, irrespective of the SLA configurations in the other regions.
- Fast convergence by backup path selection in border routers

Cisco vSmart Controller Restrictions

- Region 0 restriction: If you assign a Cisco vSmart controller to the core-region (region 0) network, you cannot assign it to any other region.
- Region parity: Cisco vSmart controllers can serve multiple regions. If you configure two Cisco vSmart controllers to serve any one region in common, then those controllers must serve all of the same regions. They cannot have only partial overlap in their coverage of regions.

  The following examples show valid and invalid Cisco vSmart controller scenarios:

  - Valid (non-overlapping):

    Controller A serves region 1.
    Controller B serves region 2.

  - Valid (overlapping single region):

    Controller A serves region 1.
    Controller B serves region 1.

  - Valid (overlapping multiple regions):

    Controller A serves regions 1, 2, and 3.
    Controller B serves regions 1, 2, and 3.

  - Invalid (partially overlapping regions):

    Controller A serves regions 1, 2, and 3.
    Controller B serves only regions 1 and 2.

Scale Limitations

Note The scale limitations described here are for the Multi-Region Fabric feature. Other limitations may apply to your network configuration.

Multi-Region Fabric has the following scale limitations:

Item	Supported Scale
Regions and Routers	
Maximum number of regions	8 From Cisco IOS XE Release 17.8.1a and Cisco SD-WAN Release 20.8.1: 12
Maximum number of edge routers per region	1,000
Maximum number of edge routers across all regions in the overlay	5,500 From Cisco IOS XE Release 17.8.1a and Cisco SD-WAN Release 20.8.1: 6,800
Maximum number of border routers per region	4
Unique unicast prefixes in the overlay	50,000 From Cisco IOS XE Release 17.8.1a and Cisco SD-WAN Release 20.8.1: 100,000
Interface Limitations	
For a border router, maximum number of TLOCs in the core region	2
For a border router, number of TLOCs for access traffic (traffic between a border router and an edge router)	2 or more
Controller Limitations	
Maximum number of Cisco vSmart controllers that can be assigned to a region	2
Maximum number of Cisco vManage instances	3 From Cisco IOS XE Release 17.8.1a and Cisco SD-WAN Release 20.8.1: 6
Maximum number of Cisco vBond Orchestrator instances	2 From Cisco IOS XE Release 17.8.1a and Cisco SD-WAN Release 20.8.1: 4
Maximum Cisco vSmart controllers assigned to the core region (region 0)	2
Maximum number of regions that a Cisco vSmart controller can serve	7

Use Cases for Multi-Region Fabric

Control of Traffic Paths Between Domains

One advantage of Multi-Region Fabric is the separation between individual regional networks and the core region. Each of these component networks can employ a different type of routing infrastructure, different service providers, and a different set of traffic policies.

In some scenarios, it is advantageous to use different types of traffic transport for intra-regional traffic and for inter-regional traffic. For example, you might use a specific transport service only for inter-regional traffic, to provide the performance that you need at a reasonable cost. The separation of component networks in Multi-Region Fabric architecture simplifies the configuration required to accomplish this.

For example, an organization operating in North America with offices and network infrastructure on the West Coast, and offices and network infrastructure on the East Coast might use different service providers in those two regions to support traffic within the region. Those service providers might not offer the optimal cost or performance for inter-regional traffic between the West Coast and the East Coast.

Without Multi-Region Fabric, one approach has been the following:

- Create a cloud service gateway in the West Coast region.
- Create another cloud service gateway in the East Coast region.
- For traffic between the two regions, configure edge devices to route the traffic to the West Coast gateway or the East Coast gateway, whichever is closest.
- Rely on the cloud services provider for transport between the two gateways.

With Multi-Region Fabric, you can use the core region to manage all traffic between the West Coast and the East Coast, and you can choose the optimal type of backbone infrastructure specifically for the core region to meet your cost and performance requirements. For example, the organization might use the following:

- A West Coast regional service provider for intra-regional West Coast traffic
- An East Coast regional service provider for intra-regional East Coast traffic
- A cloud services provider, or Cisco SD-WAN Cloud Interconnect, for the backbone infrastructure

Using Multi-Region Fabric in this scenario offers the following advantages:

- The routing configuration is far simpler.
- The Multi-Region Fabric method prevents certain routing failures—specifically, routing failures that can occur when a device responsible for one of the hops between the source and destination of a traffic flow is unavailable. These failures can occur if you use one of the more complex configuration methods for accomplishing a similar result. The Multi-Region Fabric core region that manages these intermediate hops is more responsive than other methods (such as configuring traffic to use regional gateways, as described above) to device failure and reroutes such traffic to avoid the routing failure.

In general, this disaggregation of transport providers enables you to optimize the cost and performance of operating each regional segment of the organization’s network.