The bridge to possible

Solving Global WAN Challenges with Multi-Region Fabric

Jean-Marc Barozet, Principal Engineer
@jbarozet

BRKENT-2609

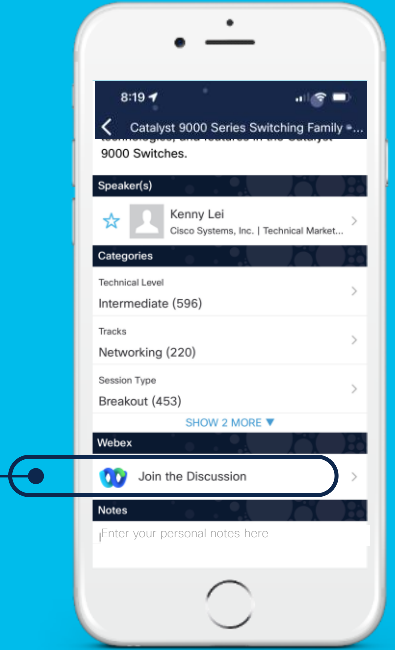
Cisco Webex App

Questions?

Use Cisco Webex App to chat
with the speaker after the session

How

1. Find this session in the Cisco Live Mobile App
2. Click “Join the Discussion”
3. Install the Webex App or go directly to the Webex space
4. Enter messages/questions in the Webex space



8:19

Catalyst 9000 Series Switching Family ...

9000 Switches.

Speaker(s)

Kenny Lei
Cisco Systems, Inc. | Technical Market...

Categories

Technical Level
Intermediate (596)

Tracks
Networking (220)

Session Type
Breakout (453)

SHOW 2 MORE ▼

Webex

Join the Discussion

Notes

Enter your personal notes here

**Webex spaces will be moderated
until February 24, 2023.**

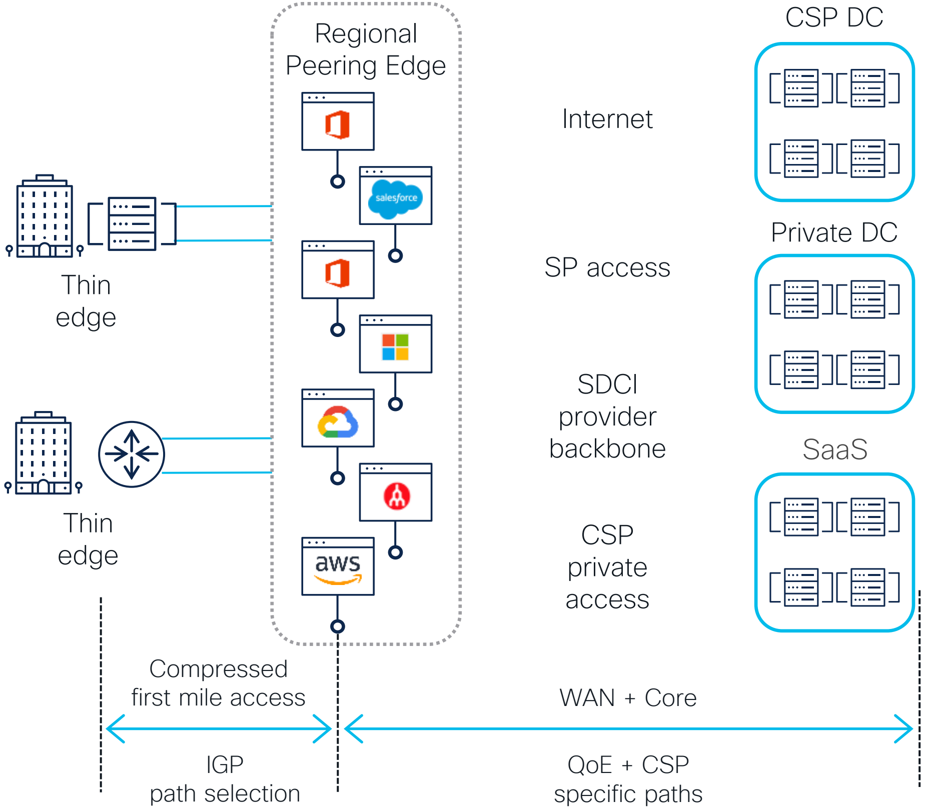
Agenda

- Introduction
- Network Design with Multi Region Fabric
- Leveraging SDCI Backbone
- Connecting Disjoint Transports
- Horizontal Scaling at your regional Hubs/Colos/PoPs
- Using Secondary Regions and Sub Regions
- Conclusion

Introduction

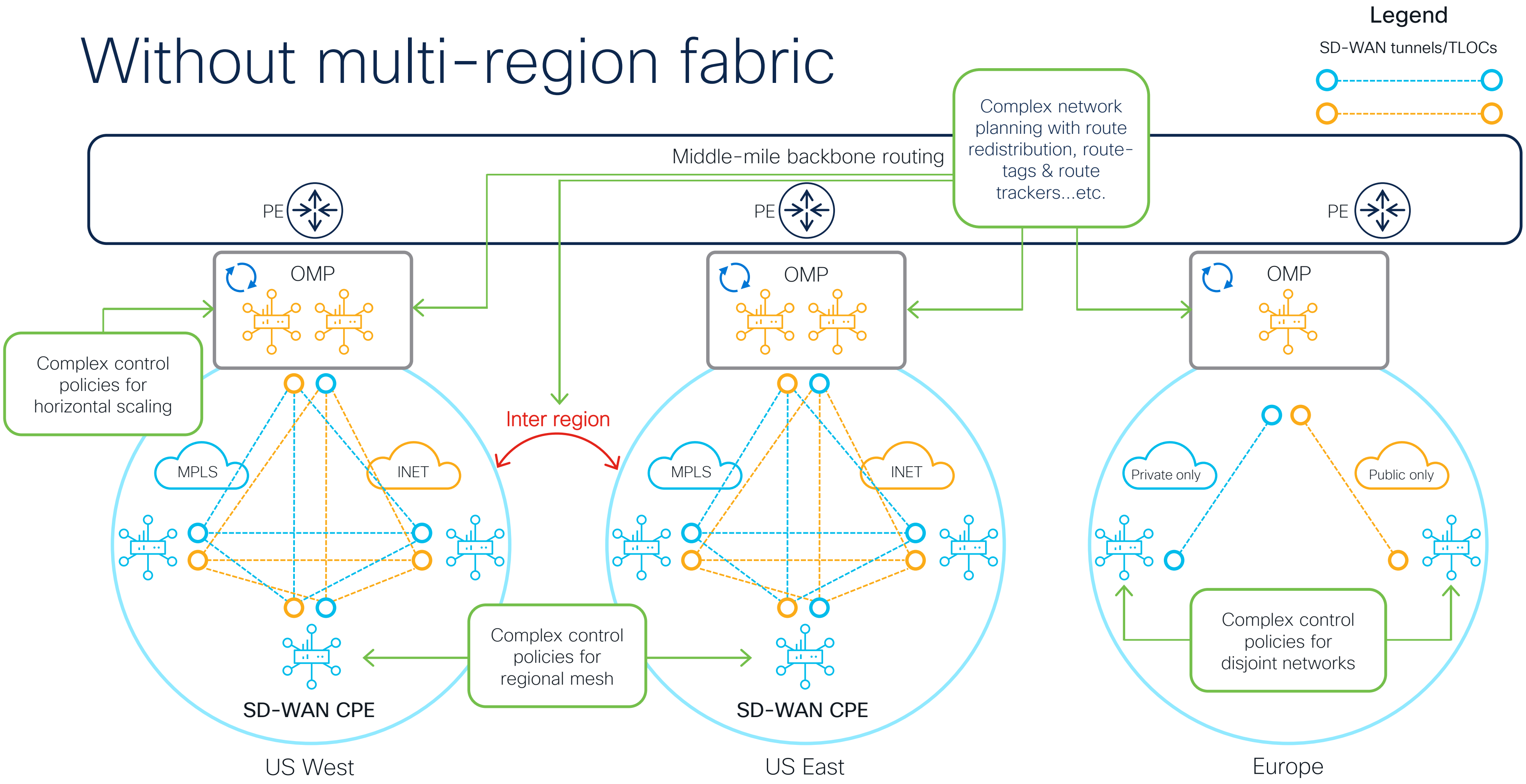
WAN is evolving to a service exchange

- The internet is changing from a network-of-networks to a network of data centers
- SDCI* and multiple provider backbones
- Large POP and Colo footprint
- Short-term contracts, usage-based
- Trending toward single ISP first-mile access
- On demand



*SDCI – Software Defined Cloud Interconnect

Without multi-region fabric



Legend

SD-WAN tunnels/TLOCs

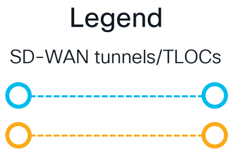
Complex network planning with route redistribution, route-tags & route trackers...etc.

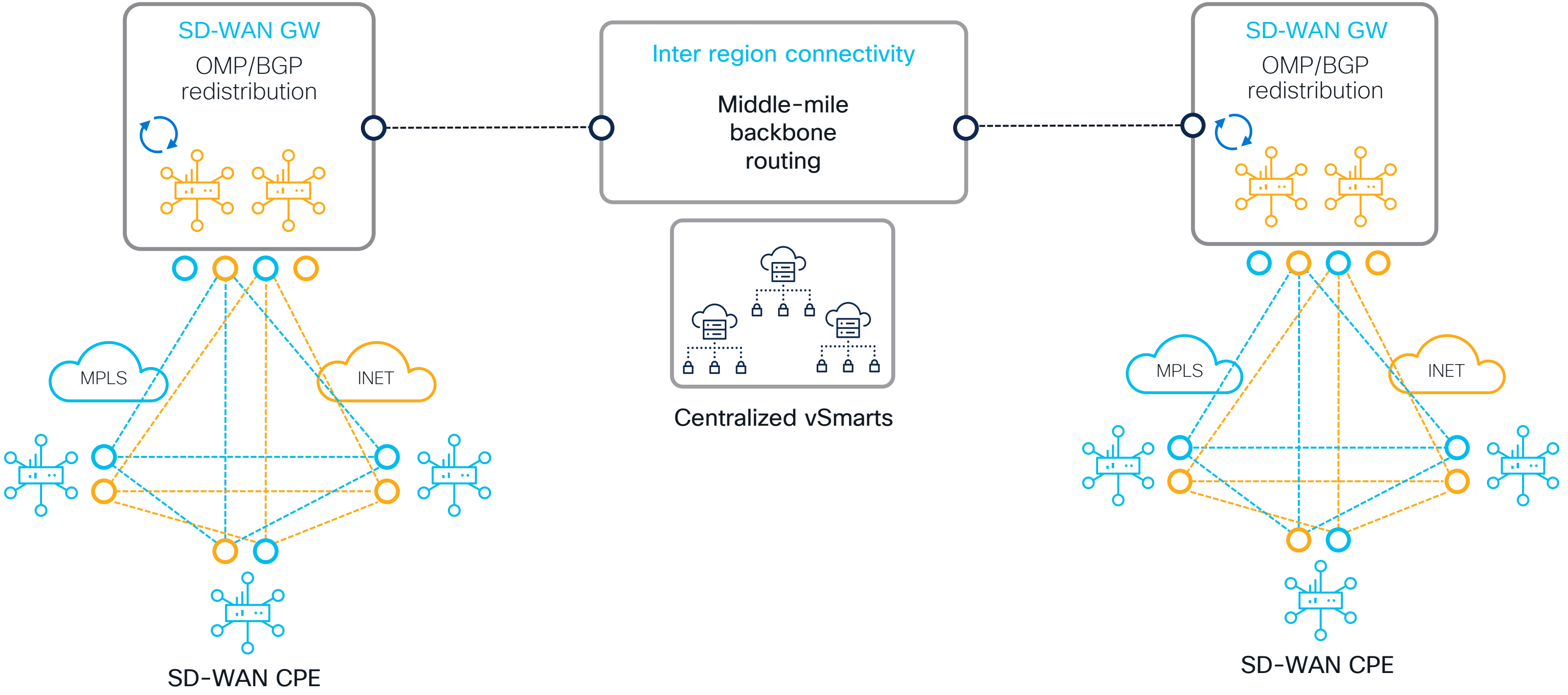
Complex control policies for horizontal scaling

Complex control policies for regional mesh

Complex control policies for disjoint networks

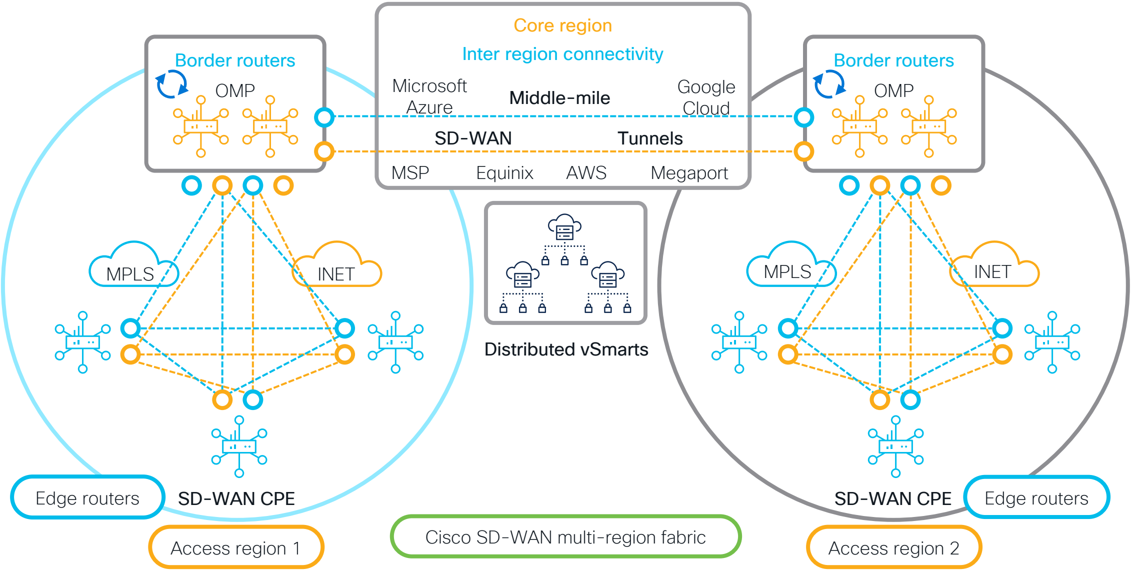
Transform from...





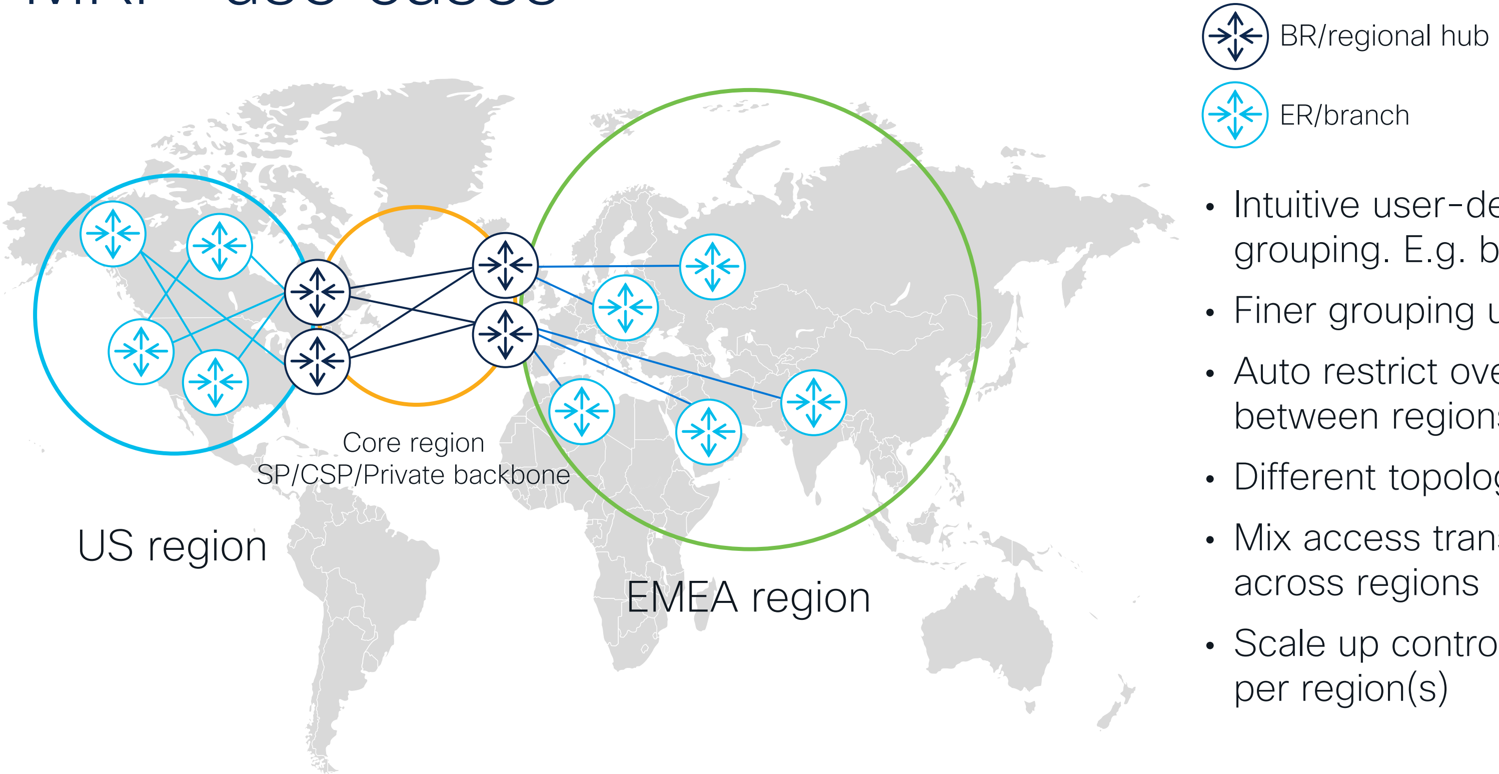
Cisco SD-WAN multi-region fabric

SD-WAN evolved for any middle-mile topology



- Eliminate lengthy global network policies
- Automatic hop-by-hop inter-region routing
- Scalable design
- Simpler redundancy planning
- Flexible architecture to cater to dynamic network needs
- Operationally easier to deploy and manage

MRF—use cases



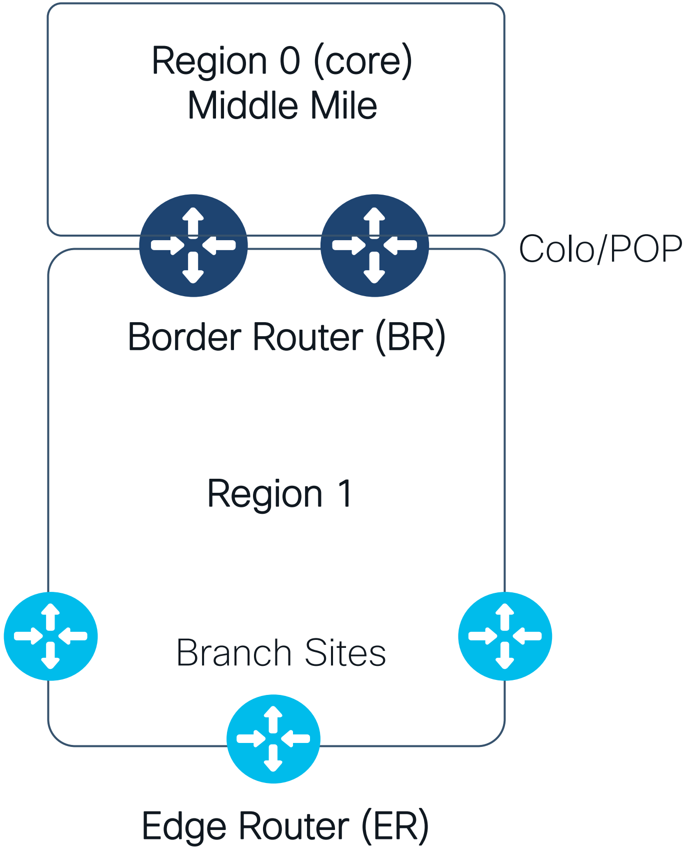
- Intuitive user-defined site grouping. E.g. based on geo
- Finer grouping using sub-regions
- Auto restrict overlay tunnels between regions
- Different topologies per region
- Mix access transports across regions
- Scale up control-plane per region(s)

Network Design with Multi Region Fabric

Regions and Roles

- Break down the network into groups, based on geo/nature of access needed at sites/nature of services needed by sites or other such parameters
- Regions – Access and Core
- Core must be fully meshed (IP reachability)
- Roles – Edge and Border

- Tunnels contained within regions – potentially use smaller branch routers with lower tunnel capabilities
- Flexibility and Scaling – mesh/partial mesh/hub and spoke within a region
- Global reachability via multiple Border Routers in every Region



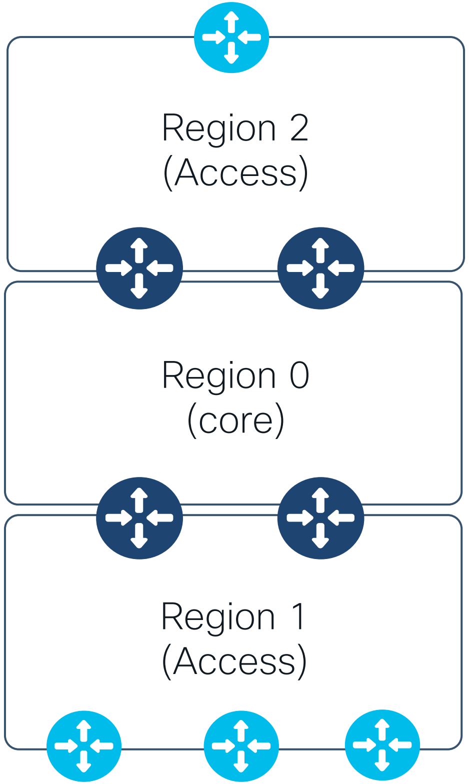
Roles – Border Router

Border Router

- Configured with Region IDs in which they operate
- Provides inter-region connectivity by connecting regional overlay to a common core or backbone overlay
- Platform – cEdge only (HW or VNF)

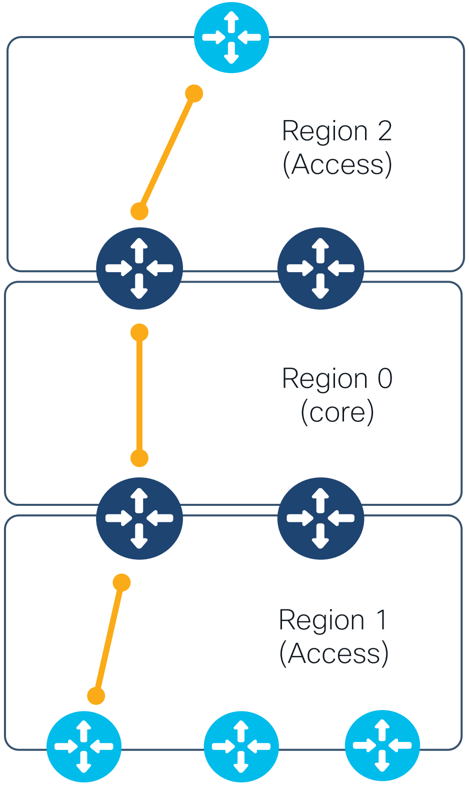
Edge Router

- Configured with Region IDs in which they operate
- Default role
- Platform: cEdge or vEdge
- Use Border Routers as next hop for inter region prefixes



Topology – IP Forwarding

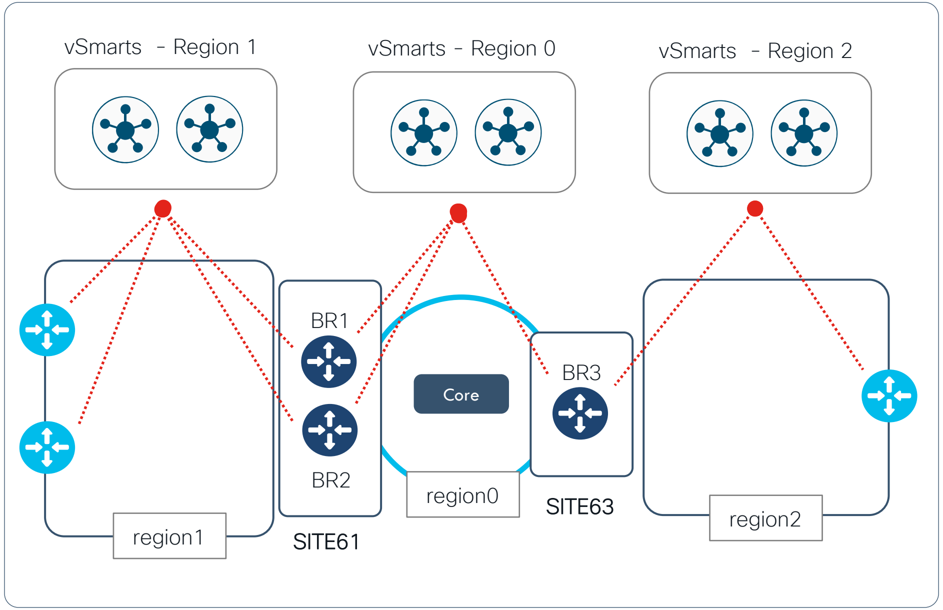
- 2-Layer Architecture
- SDWAN tunnels limited to regions
- Hop by Hop tunnels
- Decrypt/Encrypt on all nodes along the path
- IP Lookup and Forwarding per node
- Requires Service VPN on intermediate nodes (Border Routers)
- Mix of encapsulation is possible GRE in core/access
  Example: IPsec on access region and GRE on core



Distributed vSmarts

- vSmart controllers become regional
- No full mesh between region vSmarts
- vSmart for region0 cannot be shared with any access region
- Edge Routers connected to region vSmarts
- Border Routers connected to Region 0 vSmarts and Access Region vSmarts

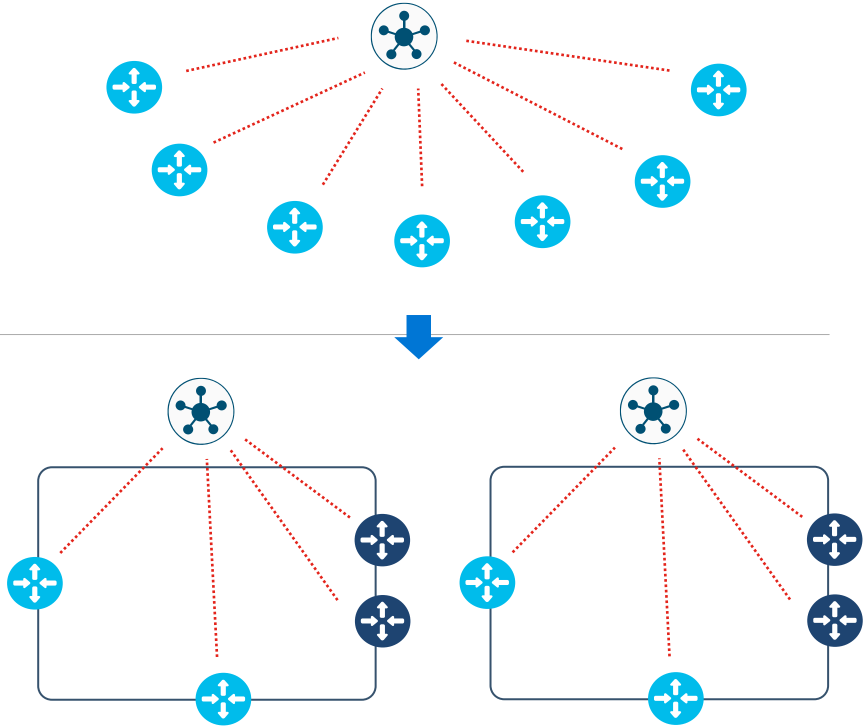
- Allow for reasonably horizontal growth in number of edge routers and mitigate the path scale requirements



vSmart Scaling based on Regions

Flat

- All devices connected
- Rib-out – replicate prefixes to all routers



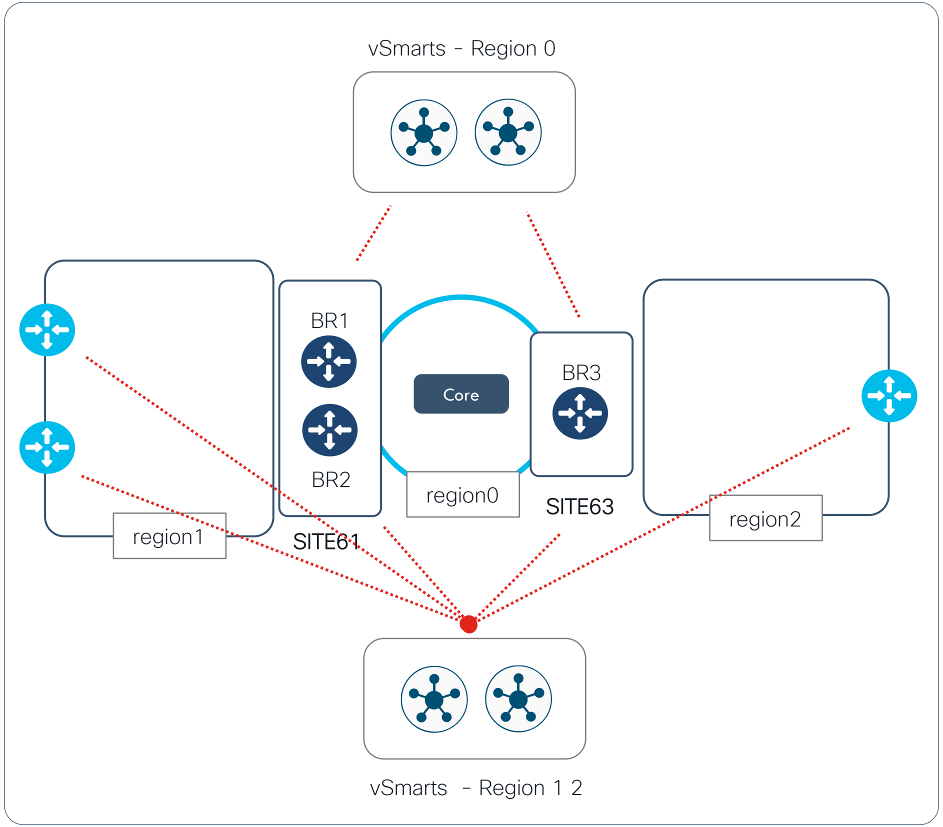
Per region vSmart

- Lower number of devices connected per region, per vSmart
- Lower number of next-hops/paths per prefix per VPN

- Number of prefixes in
- Replicated to devices connected to region vSmarts

Distributed vSmarts

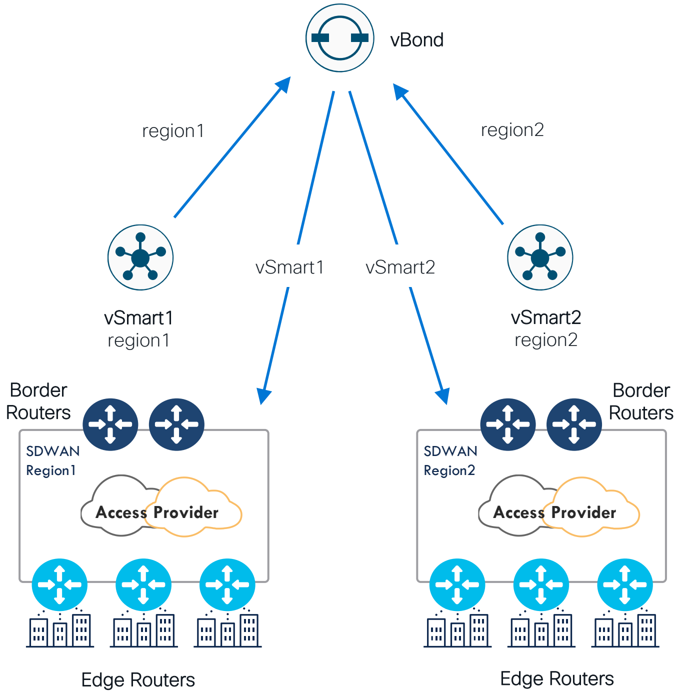
- Same vSmart can serve multiple access regions
- vSmart for region0 cannot be shared with any access region
- Avoid vSmarts with some partial overlapping regions
  - ~~vs1: [1, 2, 3], vs2: [1, 2, 4]~~



vBond remains global

- vSmarts are configured with Region IDs in which they operate
- vSmarts register their configured region IDs with vBond orchestrator. Thus vBond orchestrator is aware of list of vSmart instances that are responsible for a given region(s).
- vBond responds to ER/BR Register requests with list of vSmarts that is filtered by match of Region ID between ER/BR and the vSmarts. Edge routers and border routers peer only with vSmart controllers in their matching region
- Edge Router
  - The Edge router requests vBond about vSmarts that are in the region-id across all its tlocs
  - vBond responds to the edge with only the filtered list of vSmarts.

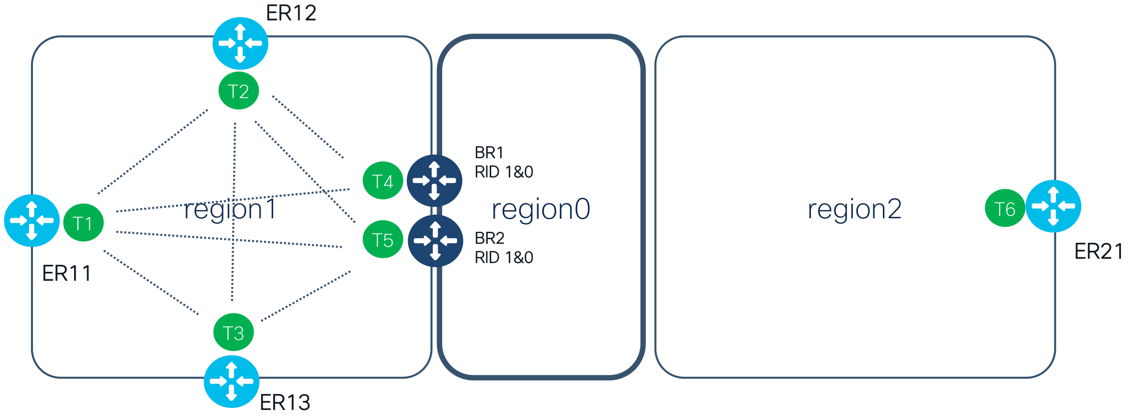
region	vSmart
1	vSmart1
2	vSmart2



Building the topology - Tunnels

ER11 OMP TLOCS

TLOC	FROM	REGION
T2	ER12	1
T3	ER13	1
T4	BR1	1
T5	BR2	1

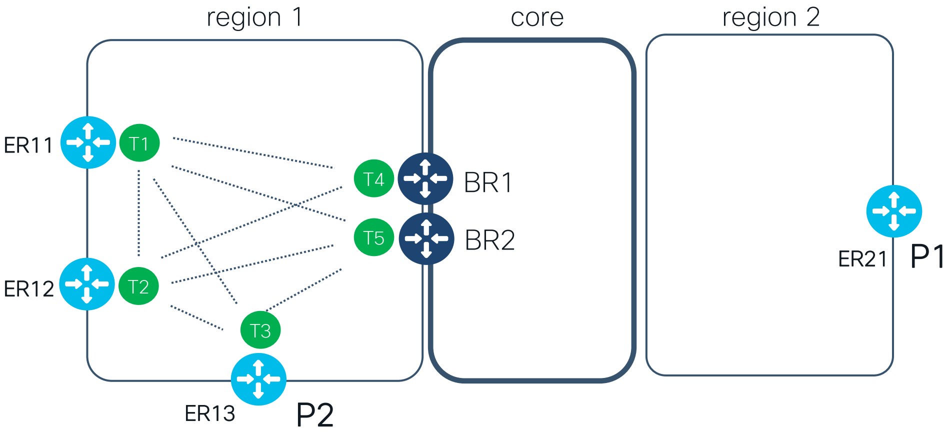


- vSmart advertises only intra-region TLOCs to WAN Edge
  - Spoke has only TLOCs from the same region
  - Border Node has TLOCs from edge region and core
- Region-id used to restrict tunnels between WAN Edge devices in the same region
- Full mesh within region

Building the topology - Routes

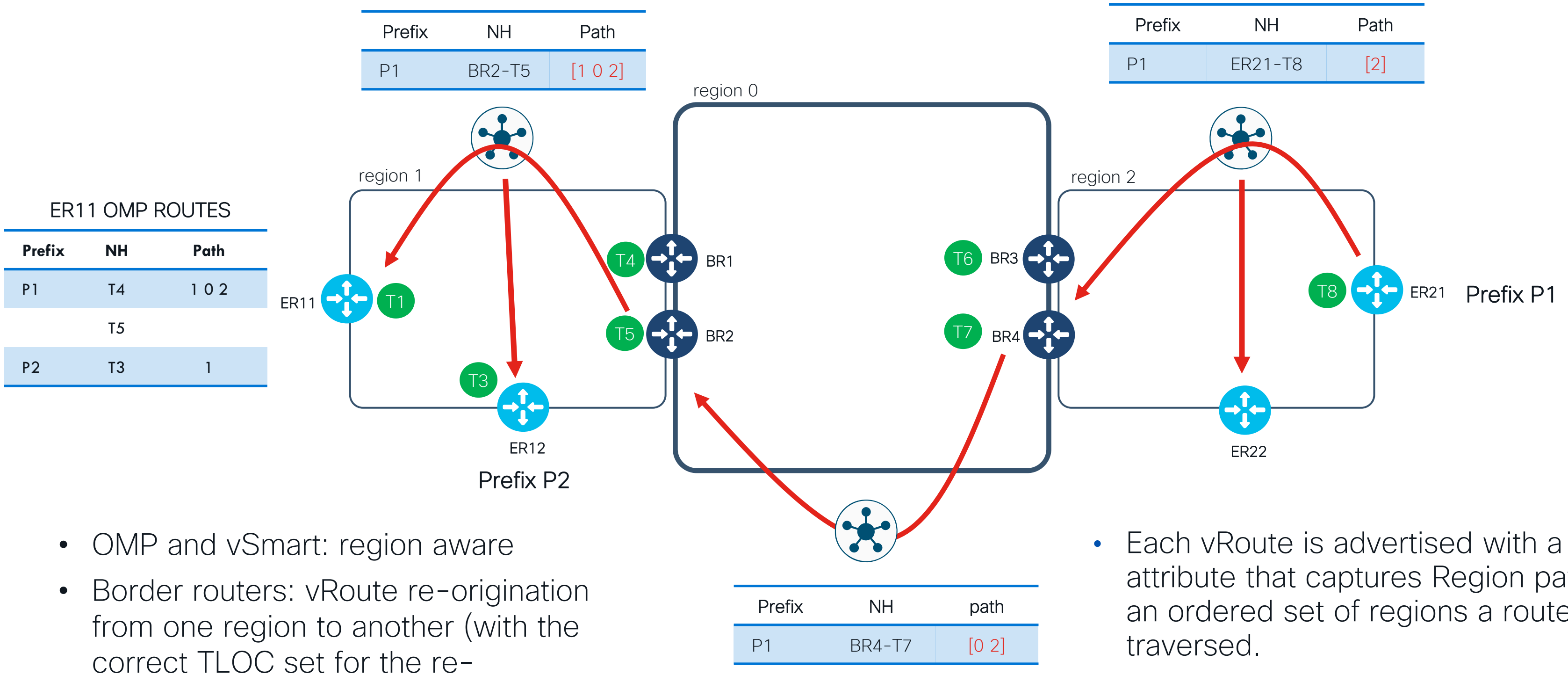
ER11 OMP ROUTES

Prefix	NH	Path
P1	T4	1 0 2
	T5	
P2	T3	1



- vSmart advertises **intra-region** routes unchanged to Edge and Border Routers
- Border Routers re-advertises **inter-region** routes to local access region
  - check inter-region prefixes reachability
  - Borders Routers must have VPN configured (vpn routing updates)
- Edge Routers
  - Intra-region prefix reachability using direct tunnels
  - Inter-region prefix reachability via Border Nodes - Default load Balancing

Building the topology - Routes



ER11 OMP ROUTES

Prefix	NH	Path
P1	T4	1 0 2
	T5	
P2	T3	1

Prefix	NH	Path
P1	BR2-T5	[1 0 2]

Prefix	NH	Path
P1	ER21-T8	[2]

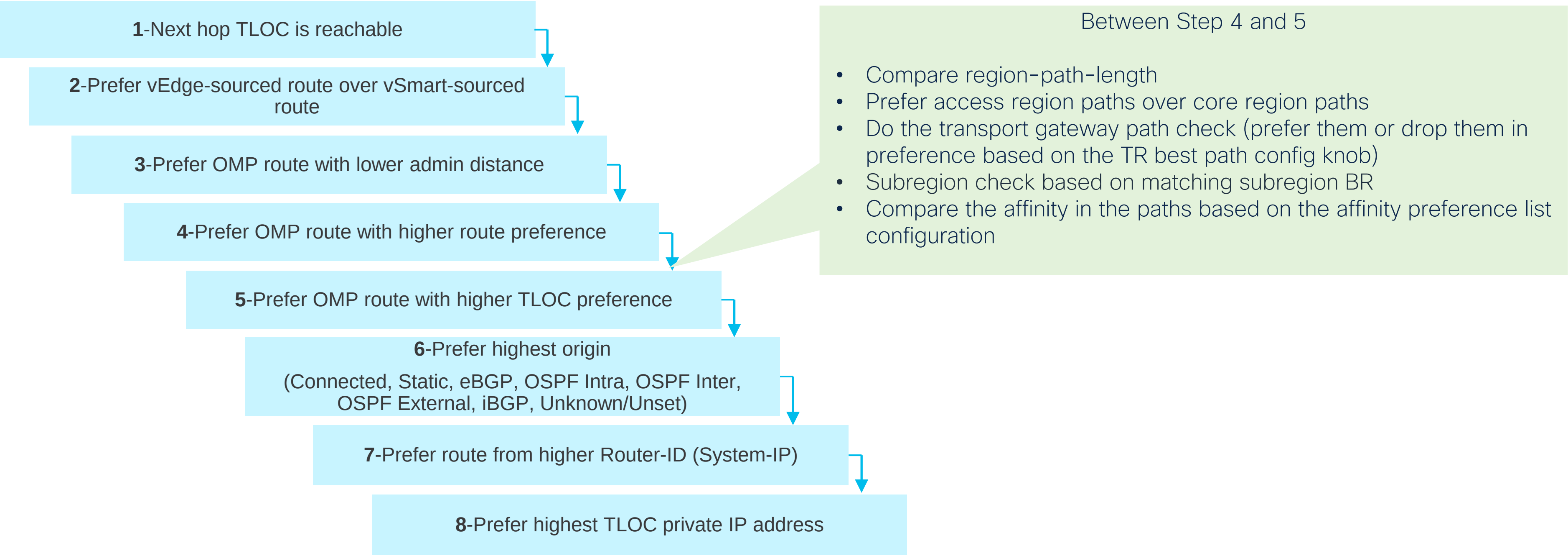
Prefix	NH	path
P1	BR4-T7	[0 2]

- OMP and vSmart: region aware
- Border routers: vRoute re-origination from one region to another (with the correct TLOC set for the re-originated route)

- Each vRoute is advertised with a new attribute that captures Region path- which is an ordered set of regions a route has traversed.
- Re-originated routes are withdrawn if the connectivity goes down. This helps prevent blackholing scenarios.

For Information

OMP Best-Path Algorithm (New)



1-Next hop TLOC is reachable

2-Prefer vEdge-sourced route over vSmart-sourced route

3-Prefer OMP route with lower admin distance

4-Prefer OMP route with higher route preference

Between Step 4 and 5

- Compare region-path-length
- Prefer access region paths over core region paths
- Do the transport gateway path check (prefer them or drop them in preference based on the TR best path config knob)
- Subregion check based on matching subregion BR
- Compare the affinity in the paths based on the affinity preference list configuration

5-Prefer OMP route with higher TLOC preference

6-Prefer highest origin

(Connected, Static, eBGP, OSPF Intra, OSPF Inter, OSPF External, iBGP, Unknown/Unset)

7-Prefer route from higher Router-ID (System-IP)

8-Prefer highest TLOC private IP address

Configuration

Configuration can’t be easier …

Core - vSmart1

```
system
 region 0
```

Access – vSmart2

```
system
 region 1 2
```

- Dedicated vSmart for core
- OMP session will not come up if core vsmart shared with access region

Edge Router - ER11

```
system
 system-ip                1.1.1.11
 overlay-id               1
 site-id                  11
 region 1
  secondary-region 4
 !
 role edge-router
!
sdwan
 interface GigabitEthernet1
  tunnel-interface
   color biz-internet
   region secondary-shared
  exit
 exit
```

- All Transport Interfaces in access region

* secondary-region explained later

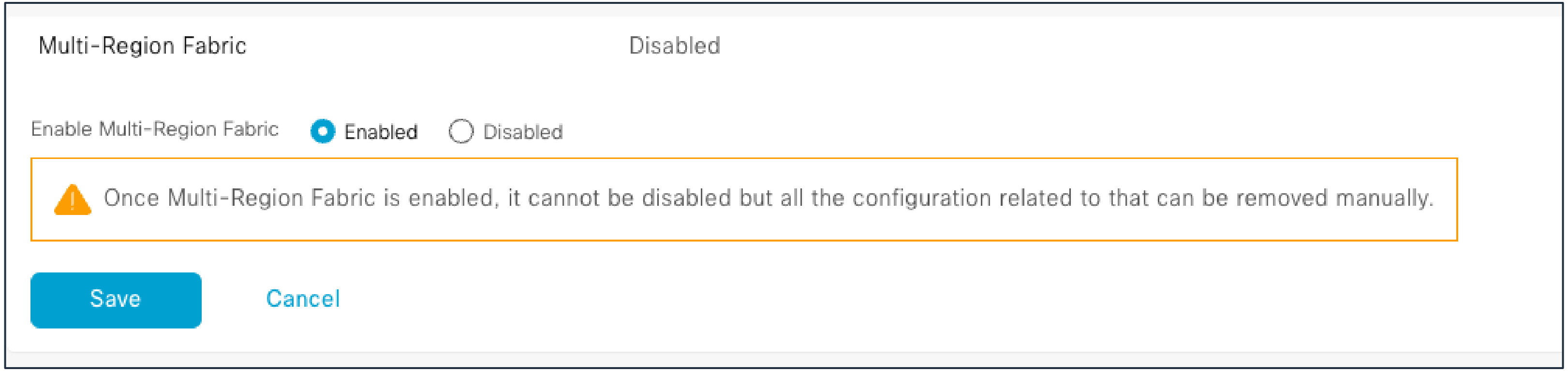
Border Router - BR1

```
system
 region 1
 role border-router
!
sdwan
 interface GigabitEthernet1
  tunnel-interface
   encapsulation ipsec
   region core
   color private1
!
 interface GigabitEthernet2
  tunnel-interface
   encapsulation ipsec
   color biz-internet
!
```

- Do not forget to configure service VPNs (w/ loopback interface for example) if you want to receive and re-advertise routes

Global Settings

Enable MRF under Administration >> settings



Multi-Region Fabric Disabled

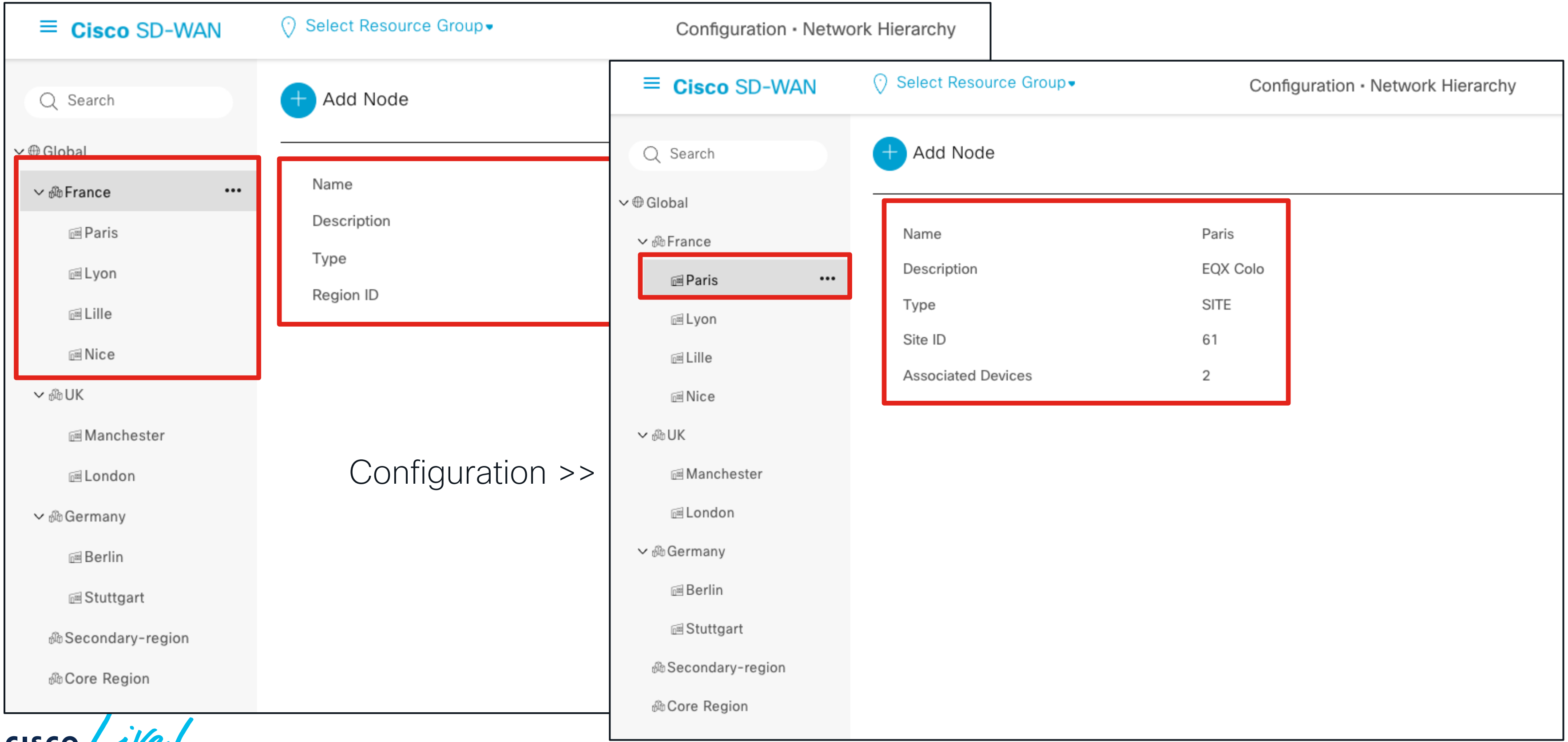
Enable Multi-Region Fabric ● Enabled ○ Disabled

⚠ Once Multi-Region Fabric is enabled, it cannot be disabled but all the configuration related to that can be removed manually.

Pre-requisite to see MRF template features in vManage

Note: MRF was previously called Hierarchical SD-WAN (H-SDWAN)

Network Hierarchy



Cisco SD-WAN | Select Resource Group | Configuration · Network Hierarchy

Search

- Global
  - France
    - Paris
    - Lyon
    - Lille
    - Nice
  - UK
    - Manchester
    - London
  - Germany
    - Berlin
    - Stuttgart
  - Secondary-region
  - Core Region

Add Node

Name

Description

Type

Region ID

Configuration >>

Cisco SD-WAN | Select Resource Group | Configuration · Network Hierarchy

Search

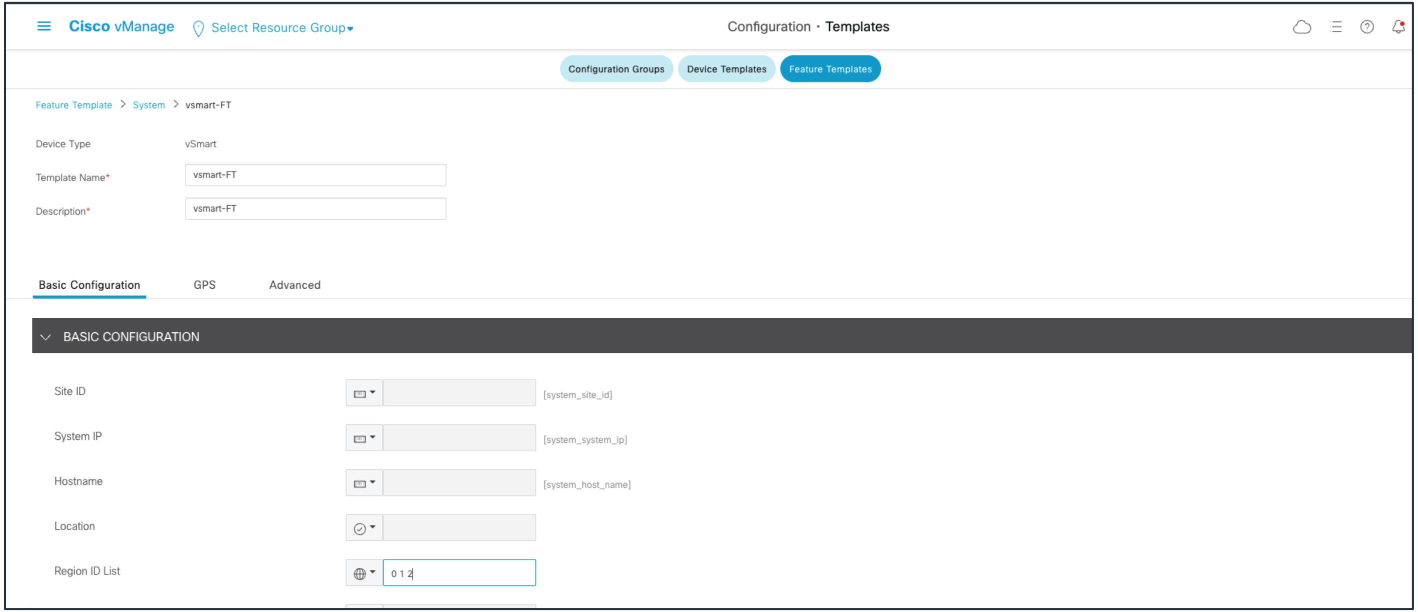
- Global
  - France
    - Paris
    - Lyon
    - Lille
    - Nice
  - UK
    - Manchester
    - London
  - Germany
    - Berlin
    - Stuttgart
  - Secondary-region
  - Core Region

Add Node

Name	Paris
Description	EQX Colo
Type	SITE
Site ID	61
Associated Devices	2

vSmart – Region

Provide region information under system template of vSmart



Device Type vSmart

Template Name* vsmart-FT

Description* vsmart-FT

Site ID [system_site_id]

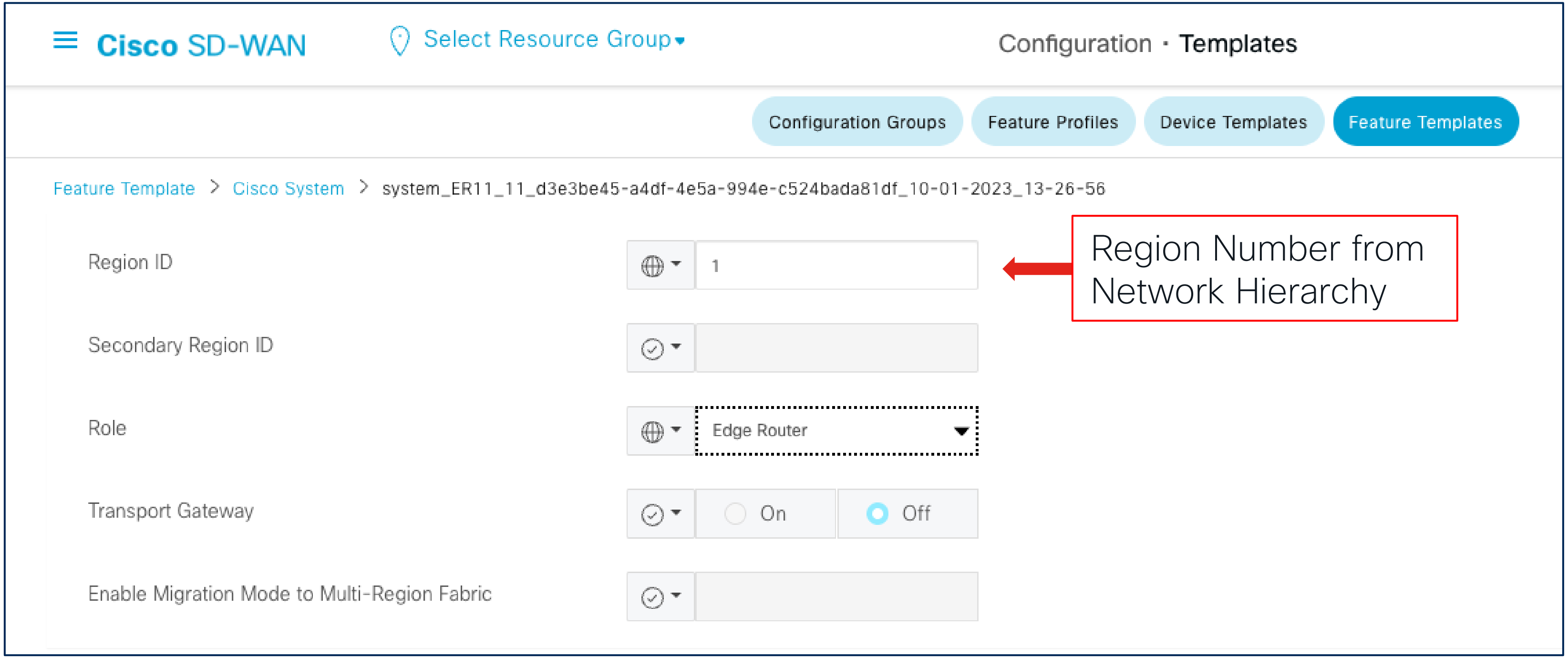
System IP [system_system_ip]

Hostname [system_host_name]

Location

Region ID List 0 1 2

Router – Region

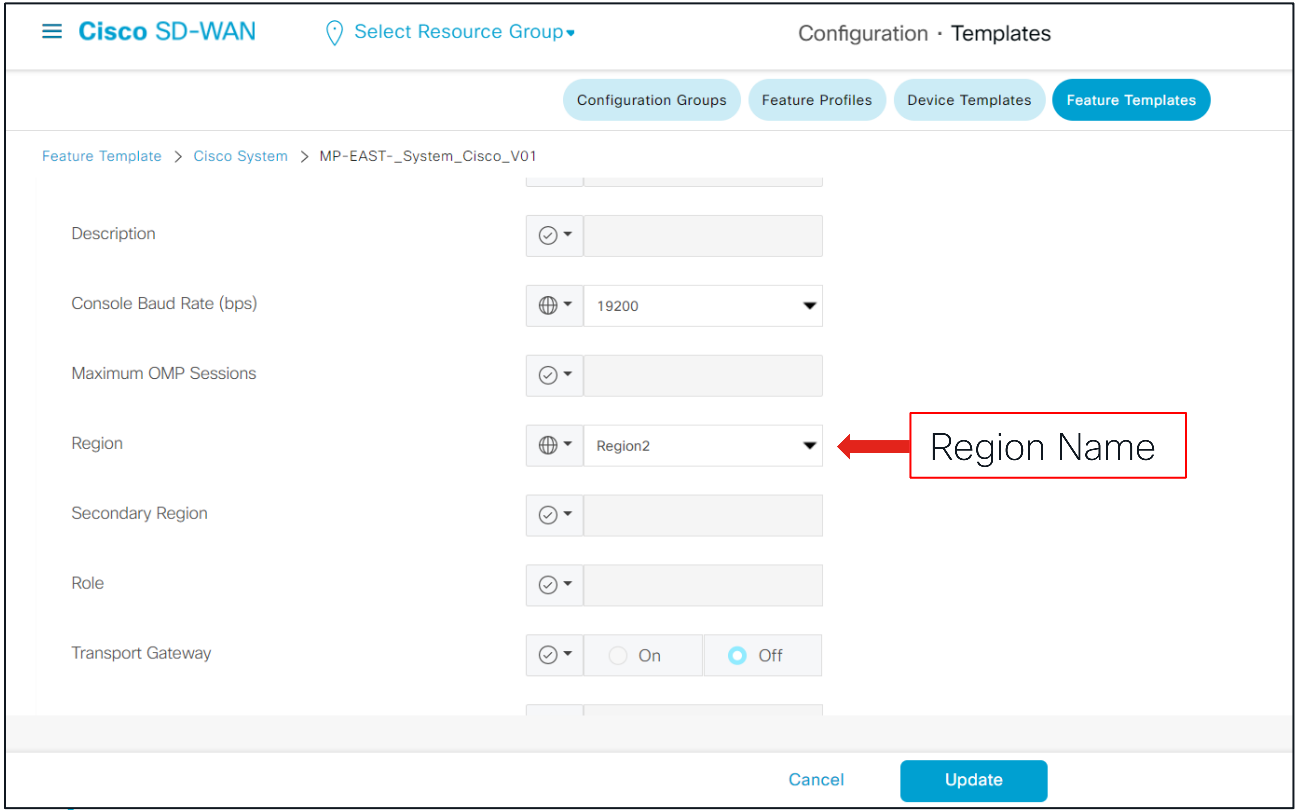


Feature Template > Cisco System > system_ER11_11_d3e3be45-a4df-4e5a-994e-c524bada81df_10-01-2023_13-26-56

Region ID	1
Secondary Region ID	
Role	Edge Router
Transport Gateway	○ On ● Off
Enable Migration Mode to Multi-Region Fabric	

Region Number from Network Hierarchy

Router – Region Name (ID)



Feature Template > Cisco System > MP-EAST-_System_Cisco_V01

Description:

Console Baud Rate (bps): 19200

Maximum OMP Sessions:

Region: Region2

Secondary Region:

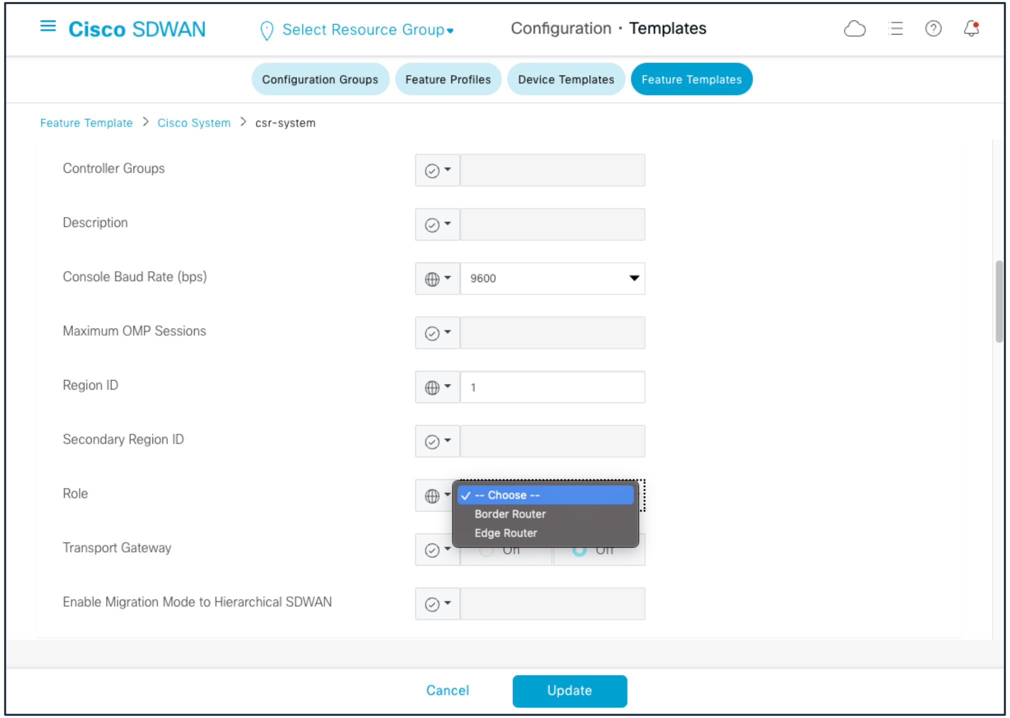
Role:

Transport Gateway: ○ On ● Off

Region Name

Router – Device Role

Enable role under System feature template of WAN edges



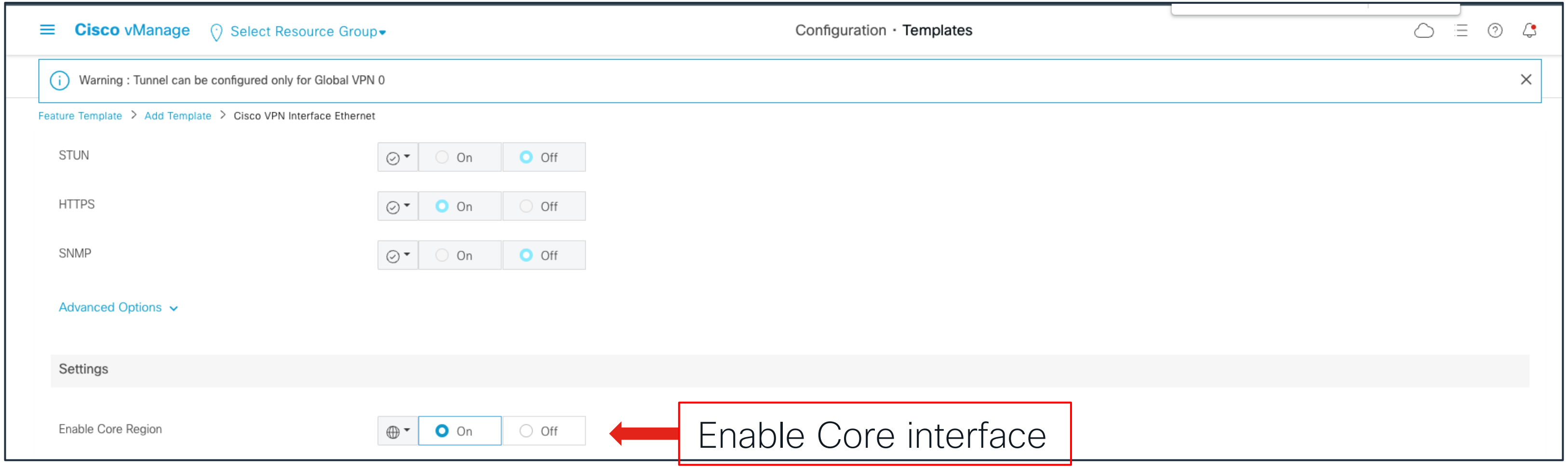
Feature Template > Cisco System > csr-system

Controller Groups	
Description	
Console Baud Rate (bps)	9600
Maximum OMP Sessions	
Region ID	1
Secondary Region ID	
Role	✓ -- Choose -- Border Router Edge Router
Transport Gateway	○ On ● Off
Enable Migration Mode to Hierarchical SDWAN	

Border Router – Core Interface

VPN Interface feature template

Enable Core tlocs under vpn interface template with region-core enabled for border routers



Warning : Tunnel can be configured only for Global VPN 0

Feature Template > Add Template > Cisco VPN Interface Ethernet

STUN ○ On ● Off

HTTPS ● On ○ Off

SNMP ○ On ● Off

Advanced Options

Settings

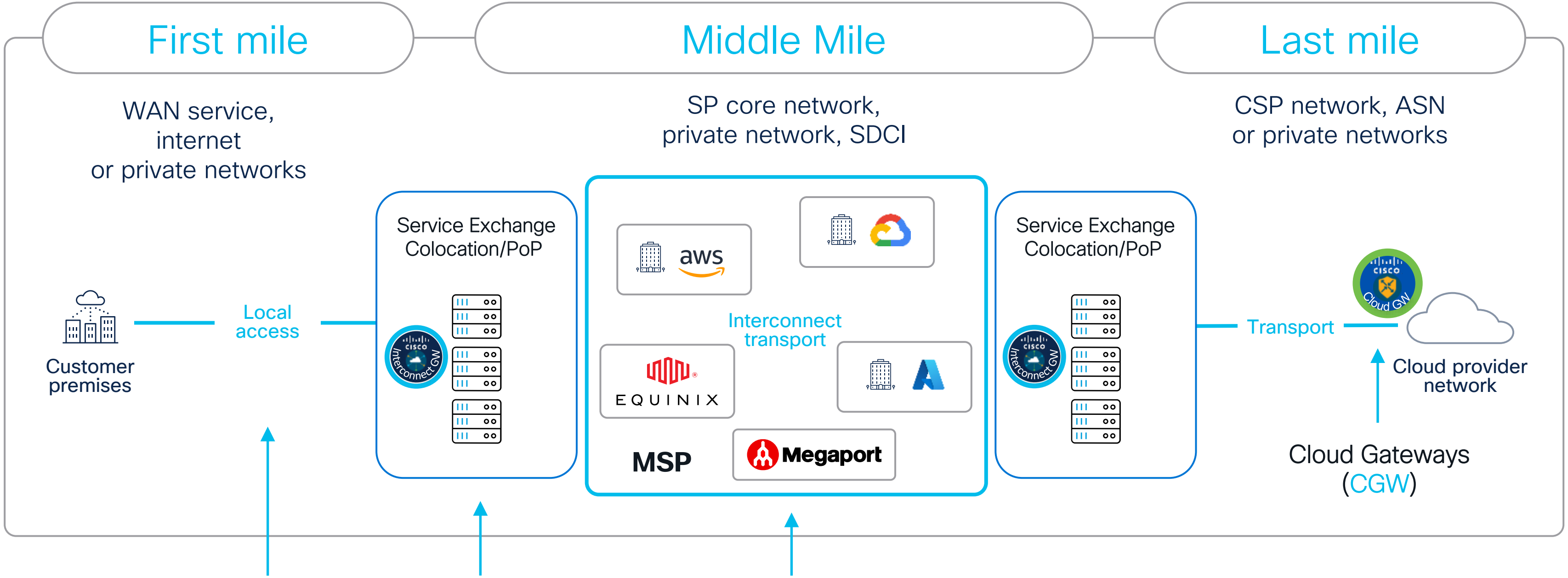
Enable Core Region ● On ○ Off

Enable Core interface

Leverage an SDCI backbone

SDCI: Software Defined Cloud Interconnect

MRF and Middle Mile Optimization



First mile
Middle Mile
Last mile
WAN service, internet or private networks
SP core network, private network, SDCI
CSP network, ASN or private networks
Customer premises
Local access
Service Exchange Colocation/PoP
Interconnect GW
aws
Interconnect transport
EQUINIX
MSP
Megaport
Service Exchange Colocation/PoP
Transport
Cloud GW
Cloud provider network
Cloud Gateways (CGW)

Compressed First Mile Access

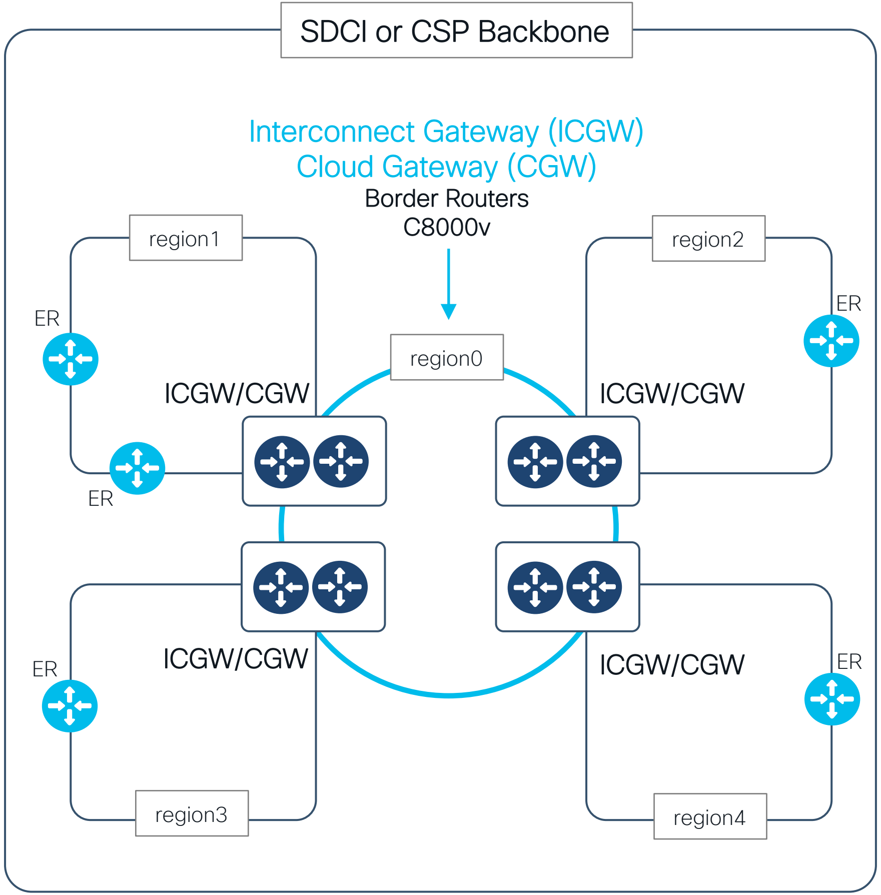
Interconnect Gateways (ICGW)

On-demand Cross-connect

New consumption model:
- Short-term Contracts, Usage-based
- On Demand

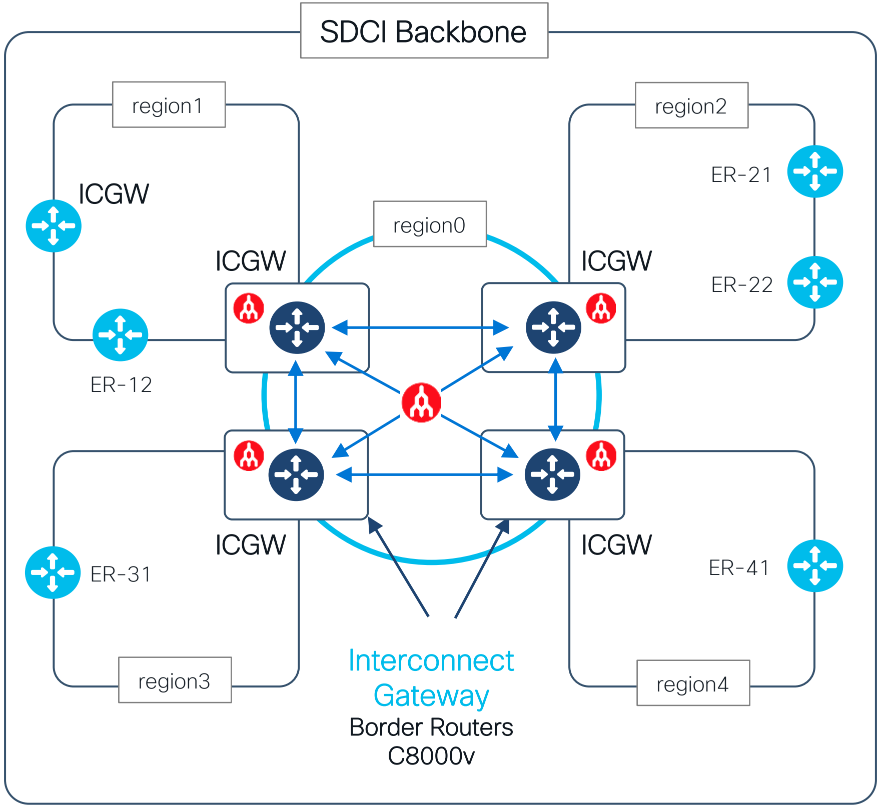
MRF with SDCI and CGW

- Split the network into MRF Regions with Cloud Gateways (CGWs) or Interconnect Gateways (ICGWs) acting as Border or Edge Routers
- Use SDCI or CSP backbone to form MRF Core Region
- Global CSP/SDCI PoPs offers on-demand connectivity
- Does NOT require underlay full mesh
  - In overlay we can establish full mesh over a partially meshed underlay with appropriate IGP routing
- Add Region vSmarts
- Integration with Network Hierarchy Feature of vManage



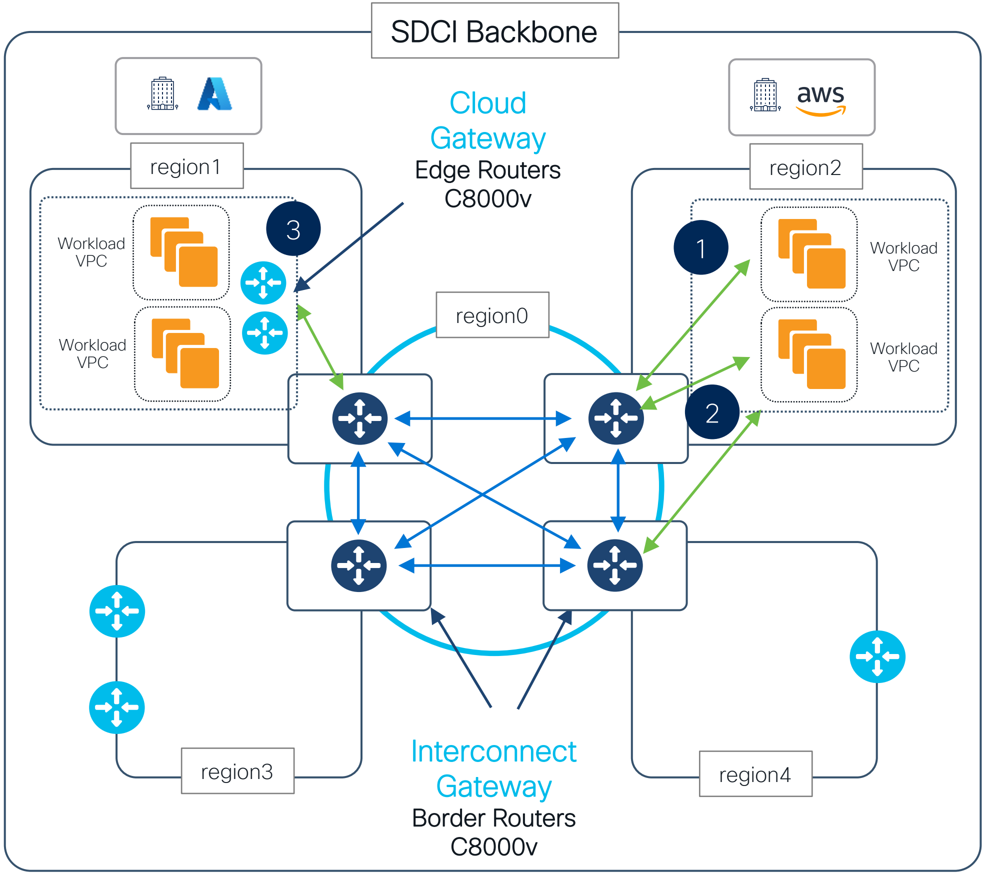
MRF with SDCI – Megaport

- MRF Support with Cloud OnRamp for MultiCloud workflow (20.10)
- Create one or more ICGW as BR for a Region.
- Full-Mesh connectivity between the Border-Router ICGWs is recommended (but not required)
- Appropriate ICGW instance license and VXC licenses, supplemental licenses should be available.
- ICGW can be BR or ER role in a topology.
- The ICGW c8kv version should be 17.8 and higher for MRF support
- Equinix not (yet) supported (Roadmap)



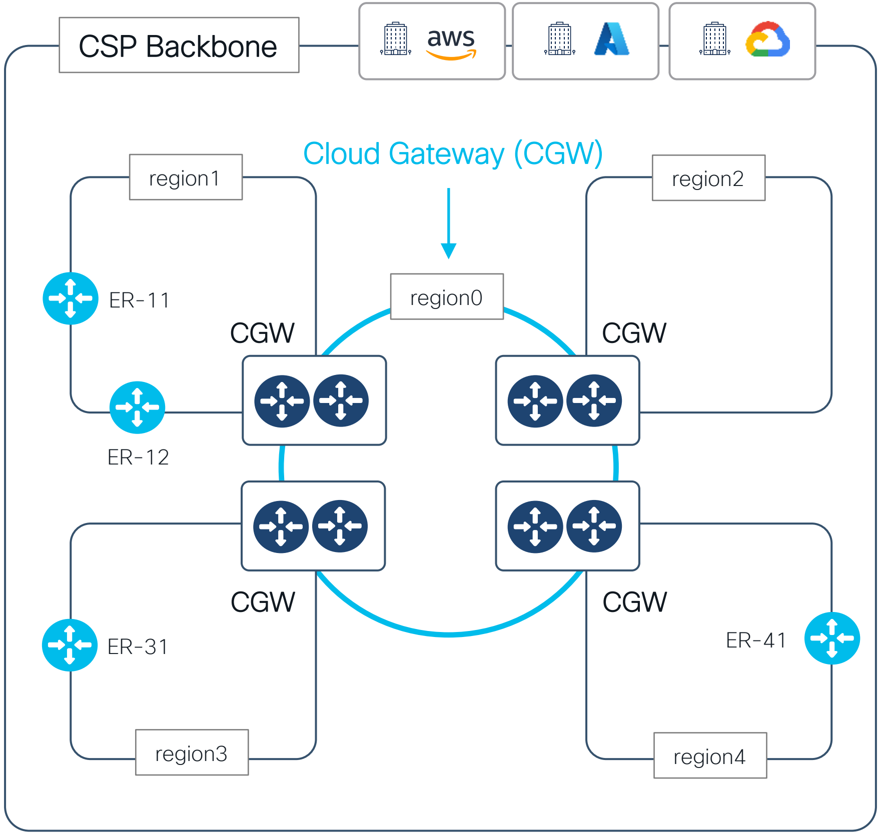
Connectivity to Cloud

- Backbone based on SDCI
- Provides connectivity to Cloud using private connections
- Site to Cloud Use Cases Supported:
  - Direct Peering to Workload VPC/Vnet.
  - Primary/Secondary Direct Peering to Workload VPC/Vnet.
  - ICGW connectivity to CGW (17.9 feature)

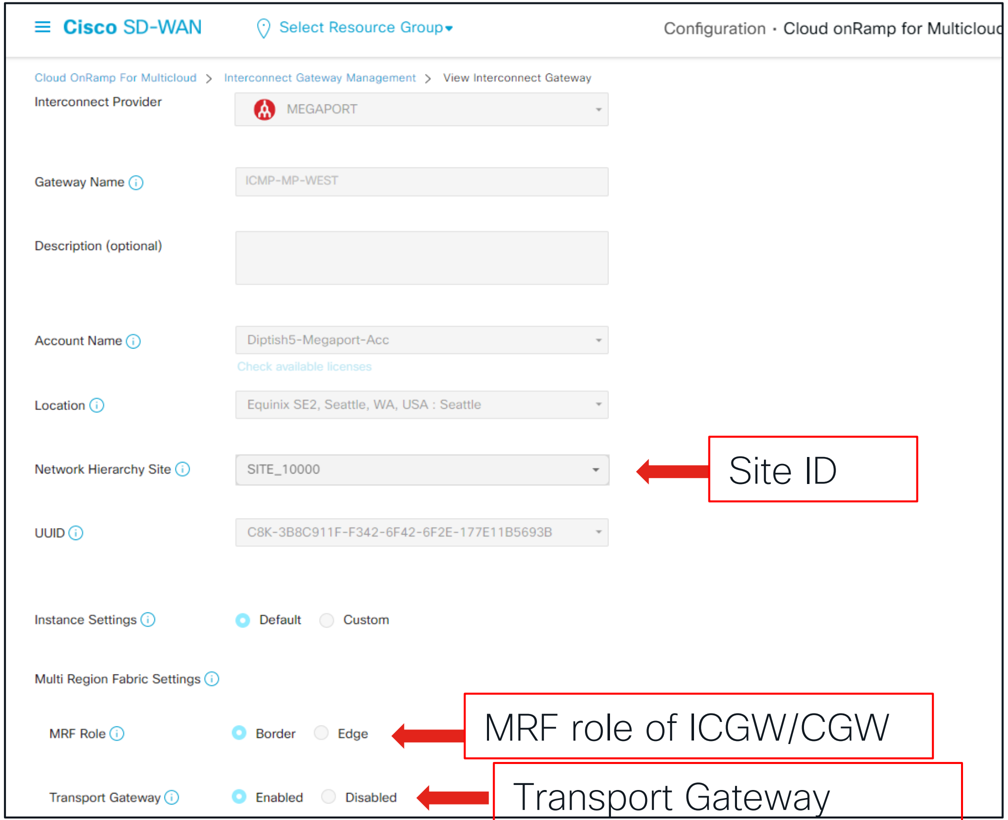


MRF with MultiCloud

- Enable CSP-Specific requirement for full-mesh S2S (Core) connectivity.
- Create one or more CGW as BR for a Region
- Both the SD-WAN router instances in the CGW belong to the same region.
- Supports AWS, Azure, GCP, AWS GovCloud, Azure GovCloud
- The CGW c8kv version should be 17.8 and higher for MRF support



Configuration : Create ICGW/CGW



Cloud OnRamp For Multicloud > Interconnect Gateway Management > View Interconnect Gateway

Interconnect Provider	MEGAPORT
Gateway Name ⓘ	ICMP-MP-WEST
Description (optional)	
Account Name ⓘ	Diptish5-Megaport-Acc Check available licenses
Location ⓘ	Equinix SE2, Seattle, WA, USA : Seattle
Network Hierarchy Site ⓘ	SITE_10000
UUID ⓘ	C8K-3B8C911F-F342-6F42-6F2E-177E11B5693B
Instance Settings ⓘ	● Default ○ Custom

Multi Region Fabric Settings ⓘ

MRF Role ⓘ	● Border ○ Edge
Transport Gateway ⓘ	● Enabled ○ Disabled

Site ID

MRF role of ICGW/CGW

Transport Gateway option for ICGW only.

MRF with Multicloud: CSP-specific Consideration

- AWS, AWS GovCloud:
  - When Site-to-Site (S2S) is enabled at global setting level, all the CGWs should have a role of Border Router
  - All CGW will be BR.
- Azure, Azure GovCloud:
  - TLOC needs to be a 'shared' TLOC
  - CGW can either be in BR or ER mode
- GCP:
  - The role that could be defined for CGW is dependent on S2S enablement for the CGW
    - S2S Enabled: Border
    - S2S Disabled: Edge
  - When S2S is enabled at global setting level, only the S2S enabled border routers participate in backbone/core routing.
  - CGW can either be in BR or ER mode.

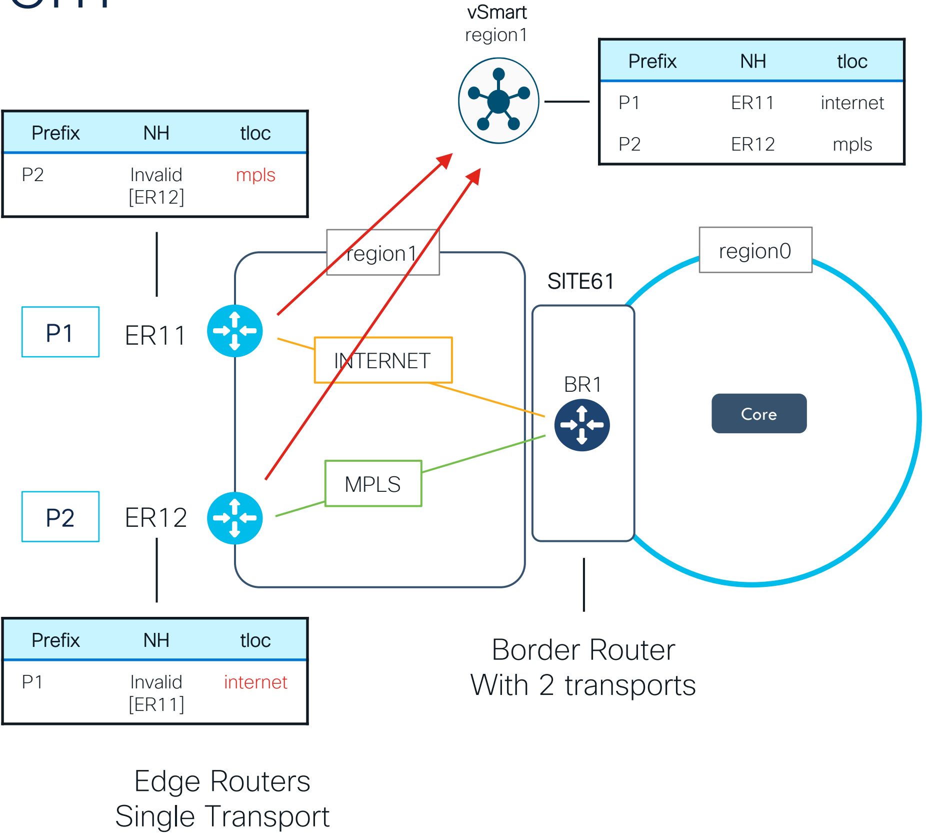
Connecting Disjoint WAN Transports in a Given Region

Dis-Joint Transport Problem

Use Case

- Single Branch sites connected to separate transports
- Connect them using a Hub or Gateway
- Dynamic / Automated

- By default, vSmart reflects route with next hop tloc of originator
- Ability to change this behavior by modifying tloc to set it for central hub (BR11)



Prefix	NH	tloc
P1	ER11	internet
P2	ER12	mpls

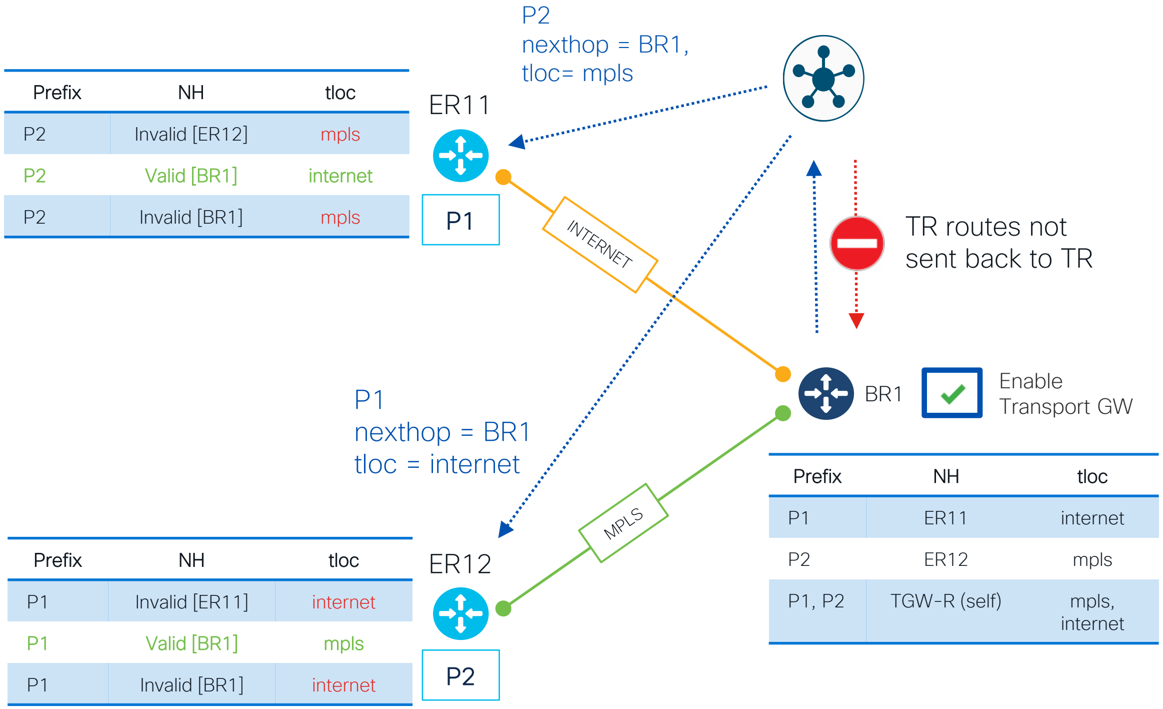
Prefix	NH	tloc
P2	Invalid [ER12]	mpls

Prefix	NH	tloc
P1	Invalid [ER11]	internet

Introducing Transport Router (TR)

vSmart learns via capability-exchange which node is working as a transit-router

- Simple easy check knob
- No need of control policies
- Works both on ER and BR
- Automatically withdraw routes, avoids blackholing
- ECMP with Multiple TRs within region
- Access region only, i.e no re-origination to/from core
- XE-SDWAN only



P2
nexthop = BR1,
tloc= mpls

P1
nexthop = BR1
tloc = internet

TR routes not sent back to TR

☑ Enable Transport GW

Prefix	NH	tloc
P2	Invalid [ER12]	mpls
P2	Valid [BR1]	internet
P2	Invalid [BR1]	mpls

Prefix	NH	tloc
P1	Invalid [ER11]	internet
P1	Valid [BR1]	mpls
P1	Invalid [BR1]	internet

Prefix	NH	tloc
P1	ER11	internet
P2	ER12	mpls
P1, P2	TGW-R (self)	mpls, internet

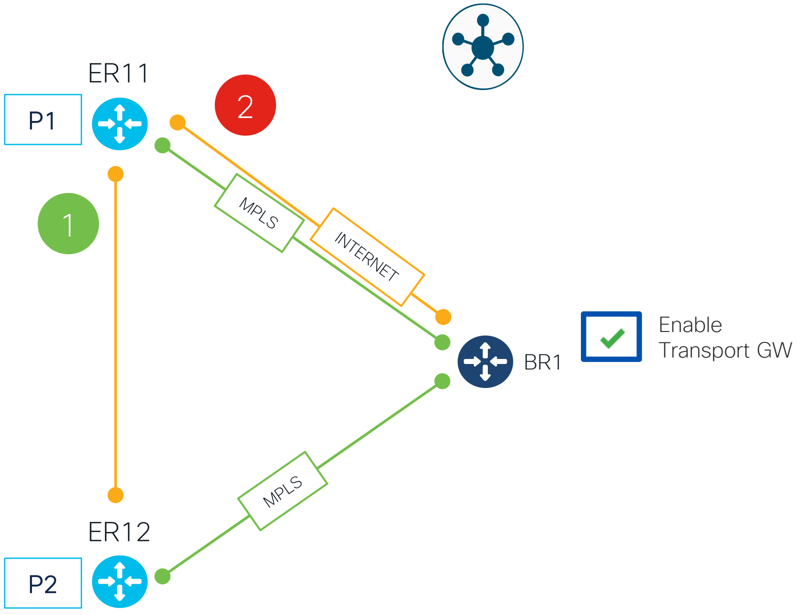
Transport Router

Direct Tunnels vs Transport Router

- Default routing
- ER11 dual transport (INET + MPLS)
- ER12 single transport (MPLS)

- ER11 Routing Table
  - DIRECT = P2 nexthop **ER12** tloc **mpls**
  - INDIRECT = P2 nexthop **BR1** tloc **internet**

▶ PREFER DIRECT (Default)



Transport Gateway
Direct Tunnels vs Transport Router

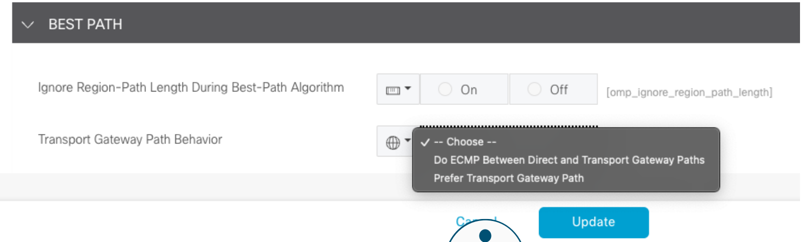
- ER11 dual transport (INET + MPLS)
- ER12 single transport (MPLS)
- OMP KNOB:
  - ECMP Direct and TR
  - Prefer TR

- ER11 Routing Table
  - DIRECT = P2 nexthop **ER12** tloc **mpls**
  - INDIRECT = P2 nexthop **BR1** tloc **internet**

▶ PREFER DIRECT (Default)

▶ OR ECMP

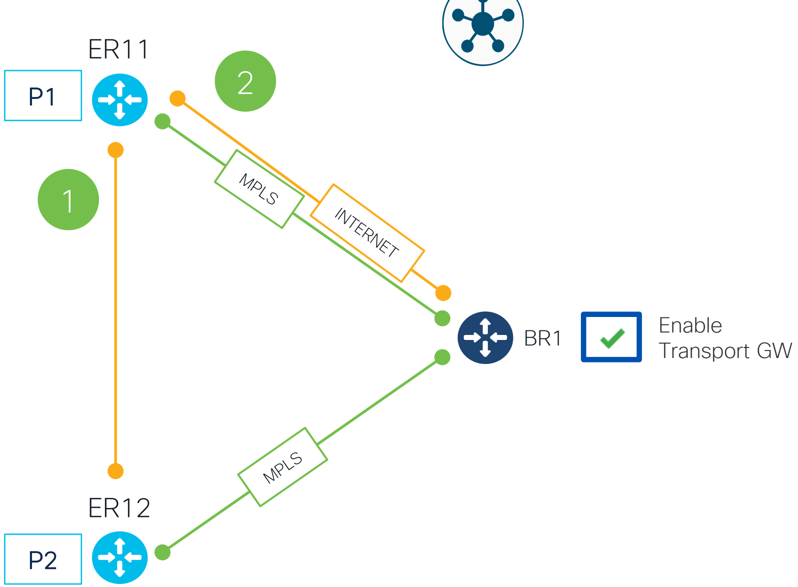
▶ OR PREFER TR



BEST PATH

Ignore Region-Path Length During Best-Path Algorithm ○ On ○ Off [omp_ignore_region_path_length]

Transport Gateway Path Behavior
✓ -- Choose --
Do ECMP Between Direct and Transport Gateway Paths
Prefer Transport Gateway Path

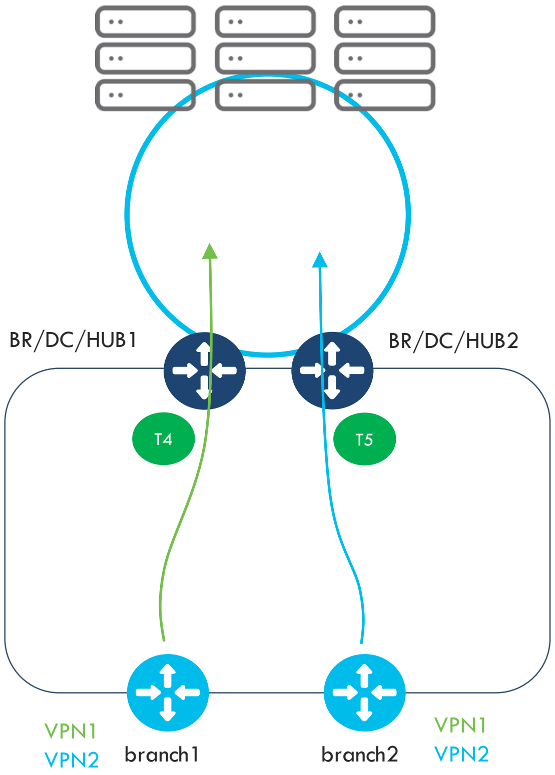


Horizontal Scaling at your regional hubs/Colo/PoPs

Border Routers Horizontal Scaling- Use-case

Use Case

- Horizontally scale Border Routers
- Automated based on intent configuration
- vSmart intelligently pin branches to relevant BR/DC/HUB without using policies

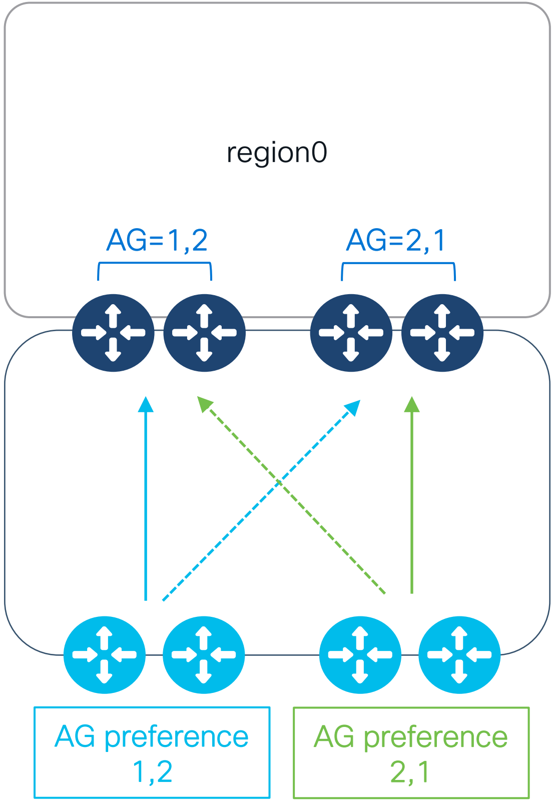


cEdge 17.8

Introducing Affinity Groups

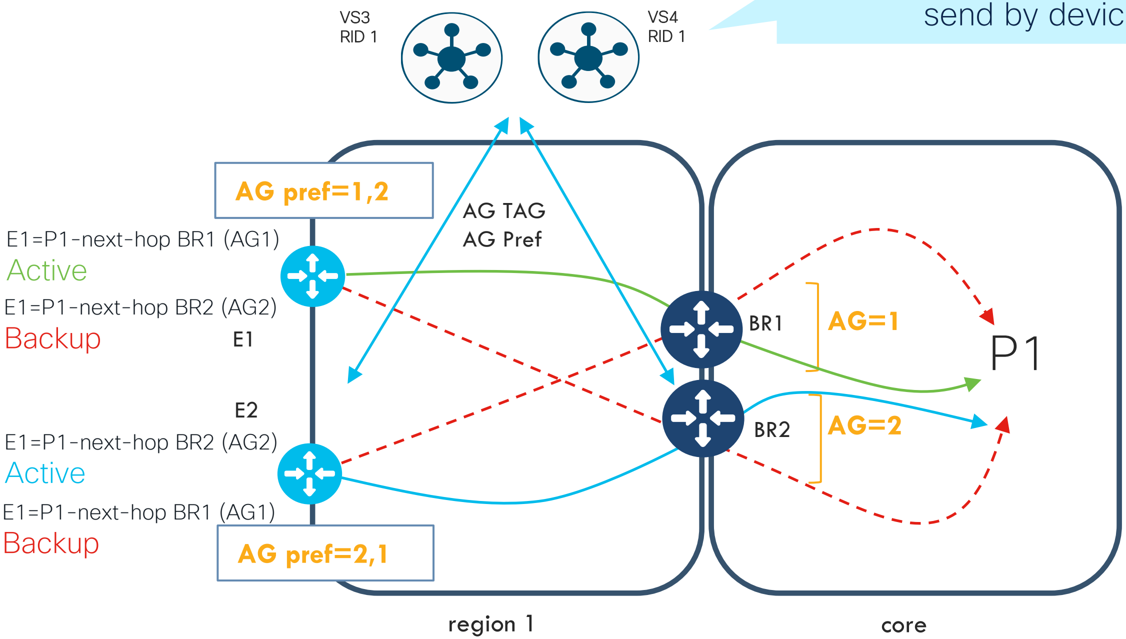
- A way to achieve this is to introduce the notion of **affinity-groups**

- Similar to tunnel groups, affinity groups (AG) can be a list of numbers configured under the system settings on the Edge Router or per TLOC
- Edge Routers with AG preference=1,2 will prefer to build tunnels and forward traffic to BRs with AG=1,2
- If BRs serving AG=1,2 go down, then branches fallback to BRs serving AG=2,1
- Control policy should allow matching based on AG or AG-list



Border Routers Horizontal Scaling

Affinity Groups

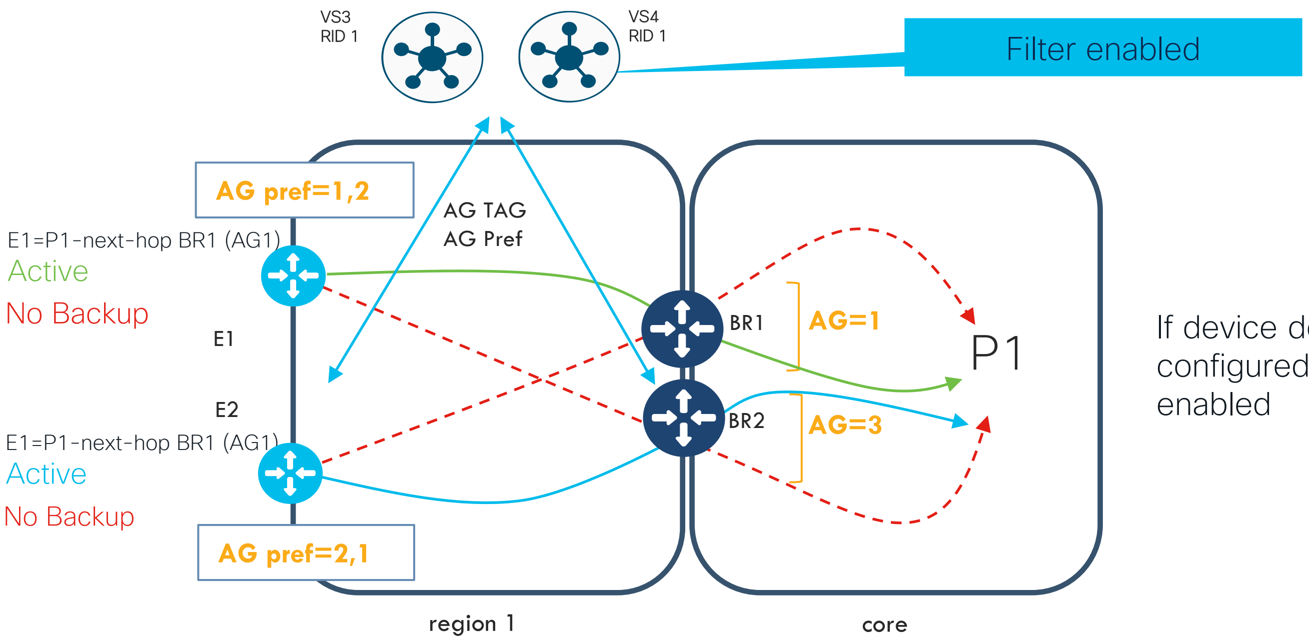


Optional configuration to filter and send routes based on affinity Pref send by devices

- Device communicates Affinity group and preference to vSmart during omp peering.
- vSmart by default ignores affinity values and propagates the routes to all the edges with affinity tags
- **Device gets all the routes based on configured affinity Pref and installs them in order of preference for forwarding.**
- **Backup path is installed only when all the path with primary affinity groups are gone**
- **Routes with no affinity configured will also be allowed if its not competing with any affinity for the same routes.**
- **Can configure multiple affinity preference**
- Optional knob on vSmart to filter only routes with affinity which devices is configured with.

Border Routers Horizontal Scaling

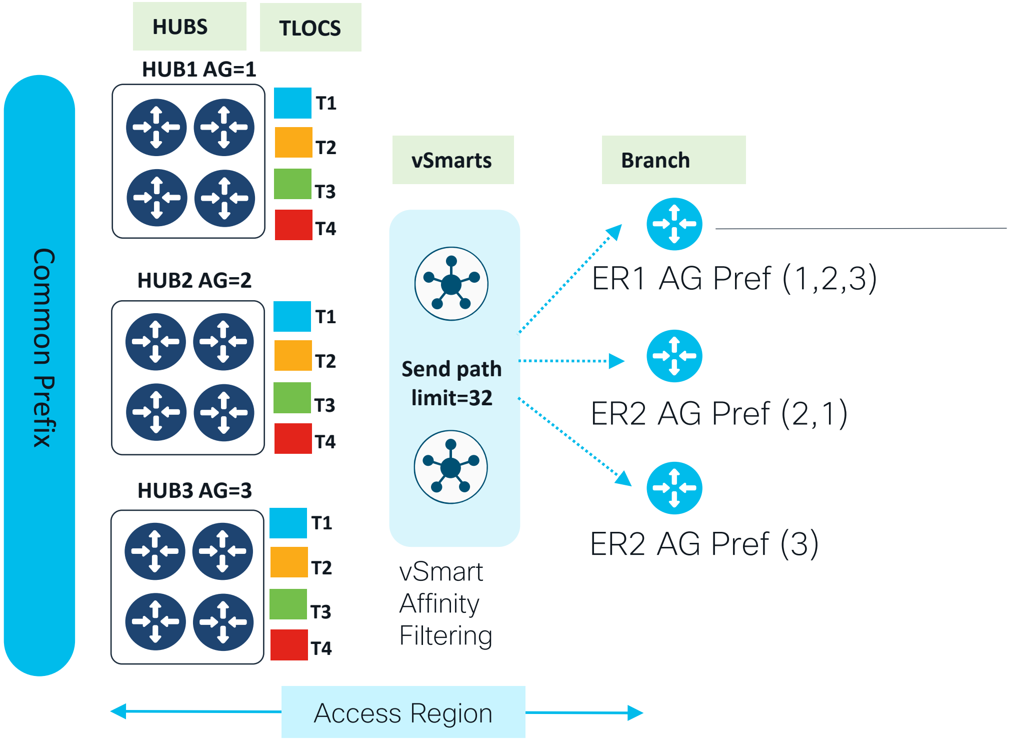
Affinity Groups



If device doesn’t have affinity preference configured it will be ignored with filtering enabled

Border Routers Horizontal Scaling

Affinity Groups Filtering outbound paths



HUBS

TLOCS

HUB1 AG=1

T1
T2
T3
T4

HUB2 AG=2

T1
T2
T3
T4

HUB3 AG=3

T1
T2
T3
T4

Common Prefix

vSmarts

Send path limit=32

vSmart Affinity Filtering

Branch

ER1 AG Pref (1,2,3)

ER2 AG Pref (2,1)

ER2 AG Pref (3)

Access Region

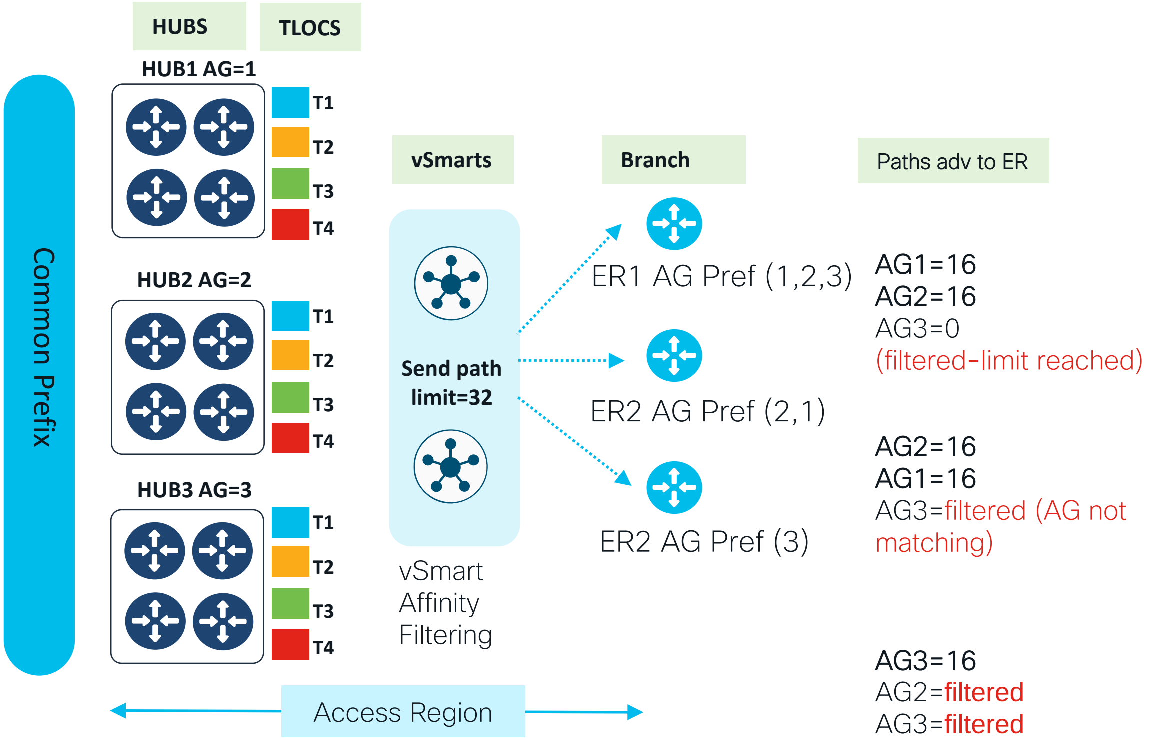
ER1 wants
HUB1=Preferred
HUB2=Backup
HUB3= Tertiary

Potentially
vSmart Best 32 list =HUB2 (16) + HUB3 (16)

There is a chance ER1 cannot get HUB1 routes in best 32

In 17.9 Border Routers Horizontal Scaling

Affinity Groups Filtering outbound paths



HUBS

TLOCS

Common Prefix

HUB1 AG=1

T1
T2
T3
T4

HUB2 AG=2

T1
T2
T3
T4

HUB3 AG=3

T1
T2
T3
T4

vSmarts

Send path limit=32

vSmart Affinity Filtering

Access Region

Branch

ER1 AG Pref (1,2,3)

ER2 AG Pref (2,1)

ER2 AG Pref (3)

Paths adv to ER

AG1=16
AG2=16
AG3=0
(filtered-limit reached)

AG2=16
AG1=16
AG3=filtered (AG not matching)

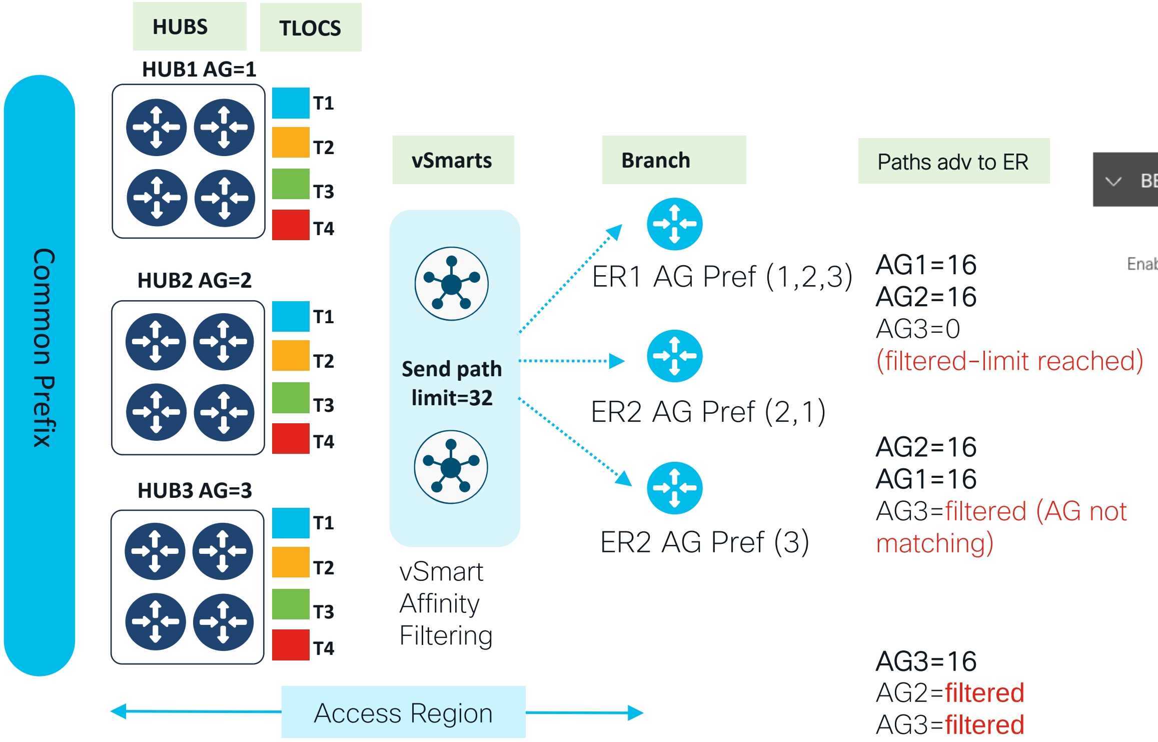
AG3=16
AG2=**filtered**
AG3=**filtered**

Outbound path filtering with vSmart

- If filtering is enabled on vSmart, it will sort the outbound paths with a sorted list based on the requested device affinity preference and send to devices
- Ensures devices gets paths filtered and based on pref priorities
- Example ER1: vSmarts sends AG1 paths first , AG2 paths second until the send-path limit is reached
- Send-backup-path (if configured) will also be prioritized based on affinity
- Works with cedge/vedge
- Path with no affinity configured for same prefix has least preference
- No new command required, just the vSmart filtering enabled

In 17.9 Border Routers Horizontal Scaling

Affinity Groups Filtering outbound paths

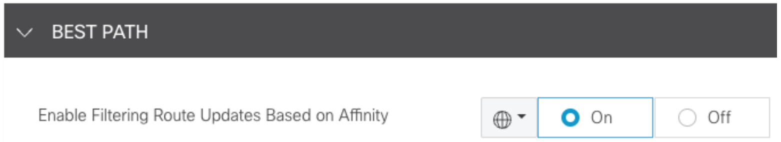


AG1=16
AG2=16
AG3=0
(filtered-limit reached)

AG2=16
AG1=16
AG3=filtered (AG not matching)

AG3=16
AG2=filtered
AG3=filtered

Outbound path filtering with vSmart



BEST PATH

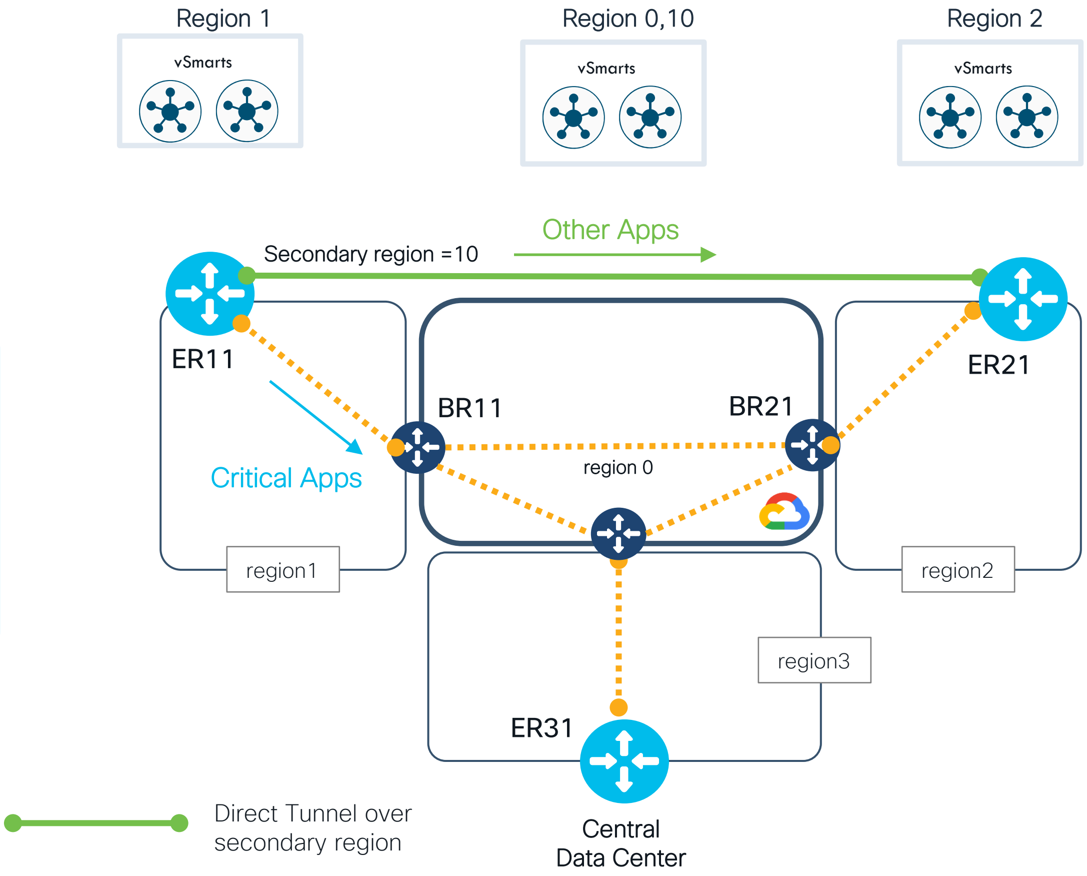
Enable Filtering Route Updates Based on Affinity ● On ○ Off

Secondary Regions

Use-case1

Use Case

Send non-critical traffic using cheap links rather than using optimal Middle-mile bandwidth or PAYG links

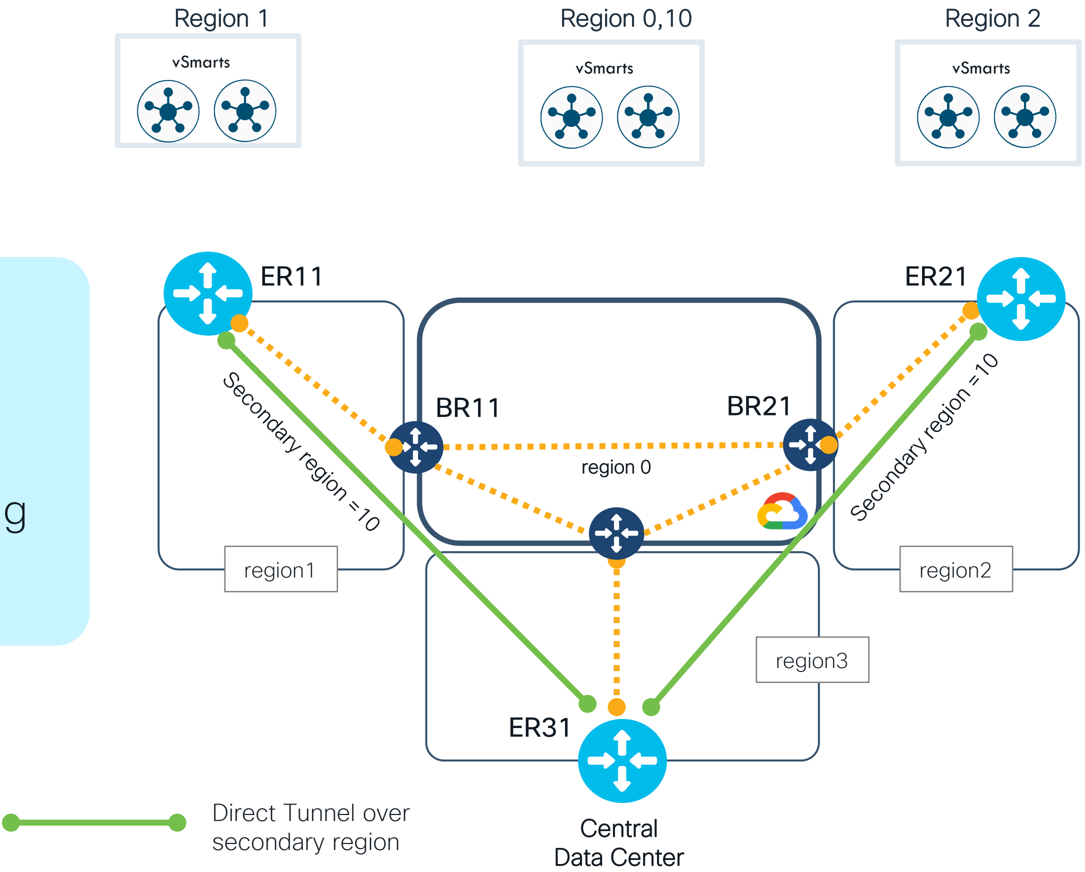


Use-case2

Use Case

Connect to central site from all regions or specific regions

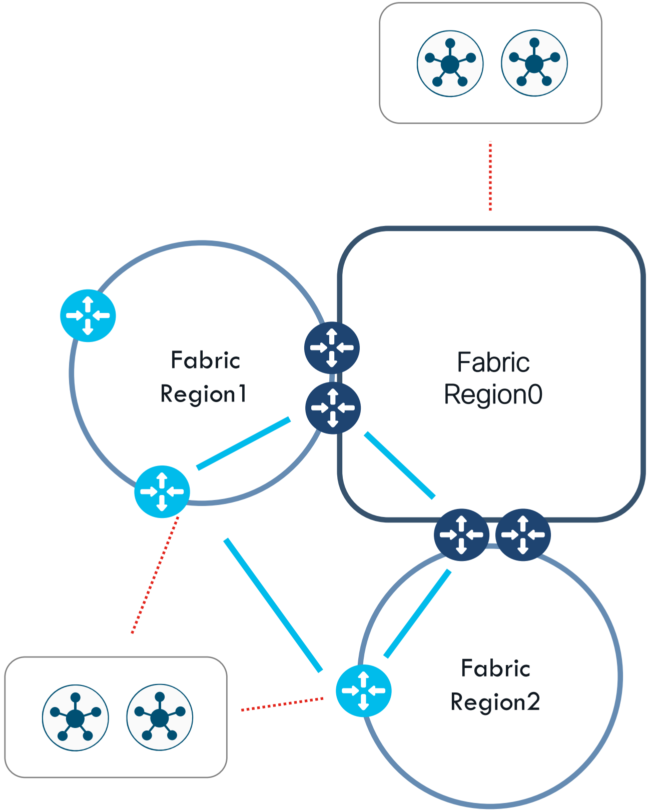
Help BRs scale better for critical traffic and reduce horizontal scaling cost at PoP/COLO



vEdge 20.8
cEdge 17.8

Introducing Secondary Region

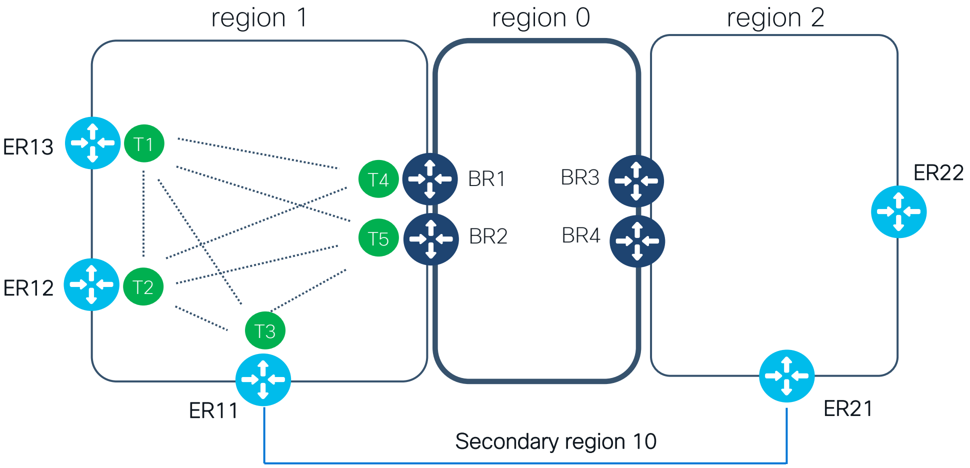
- In the most basic Multi-Region Fabric architecture, each device belongs to a single region
- Connections from an edge router in one region to an edge router in another region are routed through border routers and region 0 and therefore require multiple hops
- A secondary region contains only edge router and it enables direct tunnel connections between edge routers in different primary regions
- You can create multiple secondary regions, but an edge router cannot belong to more than one secondary region



Notes on Deployment

- vSmart for secondary region cannot be shared with primary region vSmarts. Either separate or share with core region vSmarts
- Edge forms control connections and OMP peering using primary and secondary region tlocs.
- Region numbers cannot be same for primary and secondary
- Only one secondary region per device, multiple secondary region throughout network for other set of devices
- By default, direct tunnel will be preferred over indirect tunnel due to shorter path
    - Ability to change this behavior in OMP by ignoring path length. In that case, ECMP is used

Secondary Region Configuration



ER11 spoke

```
system
 region 1
  secondary-region 10
 role edge-router
!
vpn 0
 interface ge0/0
  ip dhcp-client
  ipv6 dhcp-client
  tunnel-interface
   encapsulation ipsec
   color mpls
   region secondary-shared
```

ER21 Spoke

```
system
 region 2
  secondary-region 10
 role edge-router
!
vpn 0
 interface ge0/1
  ip dhcp-client
  tunnel-interface
   encapsulation ipsec
   color mpls
   region secondary-only
```

Primary region vSmart

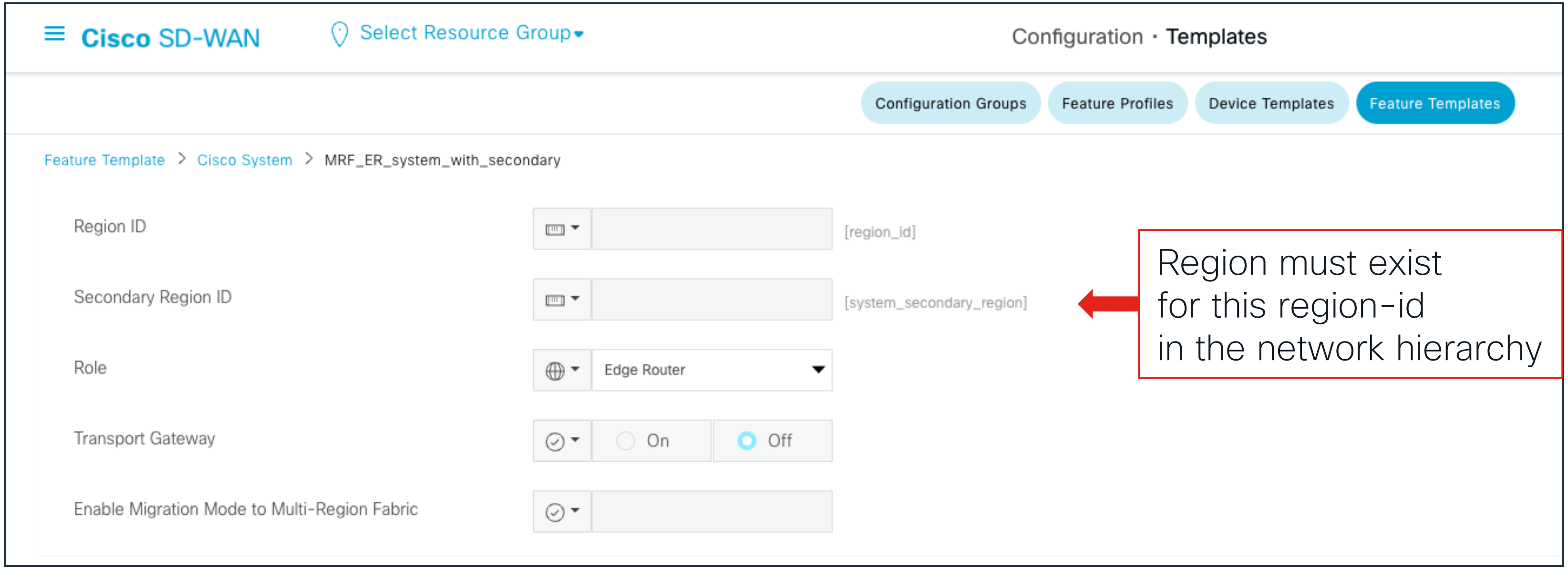
```
system
 host-name      vsmart1-r1
 system-ip      1.1.1.4
 site-id        1
 region 1,2
```

Secondary region vSmart

```
system
 host-name      vsmart1-r0
 system-ip      1.1.1.254
 site-id        1
 region 0 10
```

For Information

System Feature Template



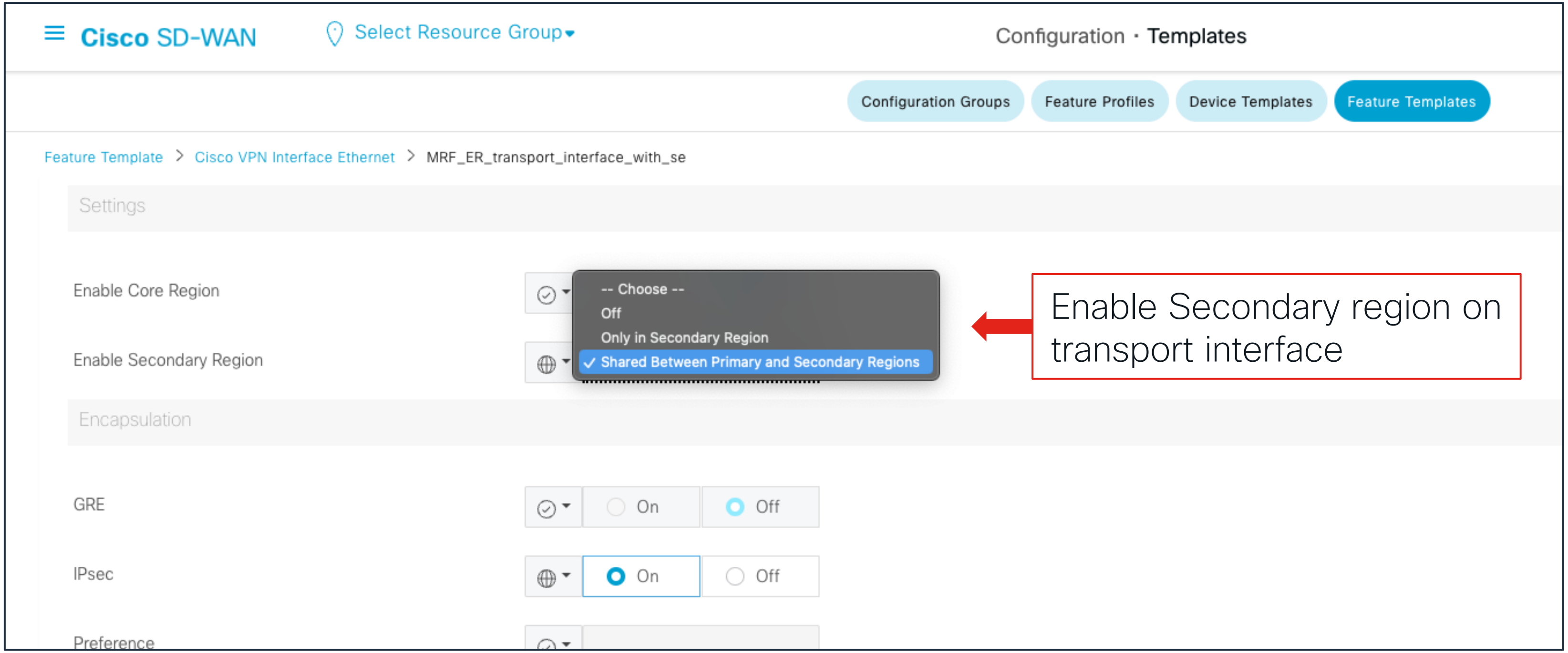
Feature Template > Cisco System > MRF_ER_system_with_secondary

Region ID		[region_id]
Secondary Region ID		[system_secondary_region]
Role	Edge Router	
Transport Gateway	○ On ● Off	
Enable Migration Mode to Multi-Region Fabric		

Region must exist for this region-id in the network hierarchy

For Information

Transport Interface Feature Template



Feature Template > Cisco VPN Interface Ethernet > MRF_ER_transport_interface_with_se

Settings

Enable Core Region

Enable Secondary Region: ✓ Shared Between Primary and Secondary Regions

Encapsulation

GRE: ○ On ● Off

IPsec: ● On ○ Off

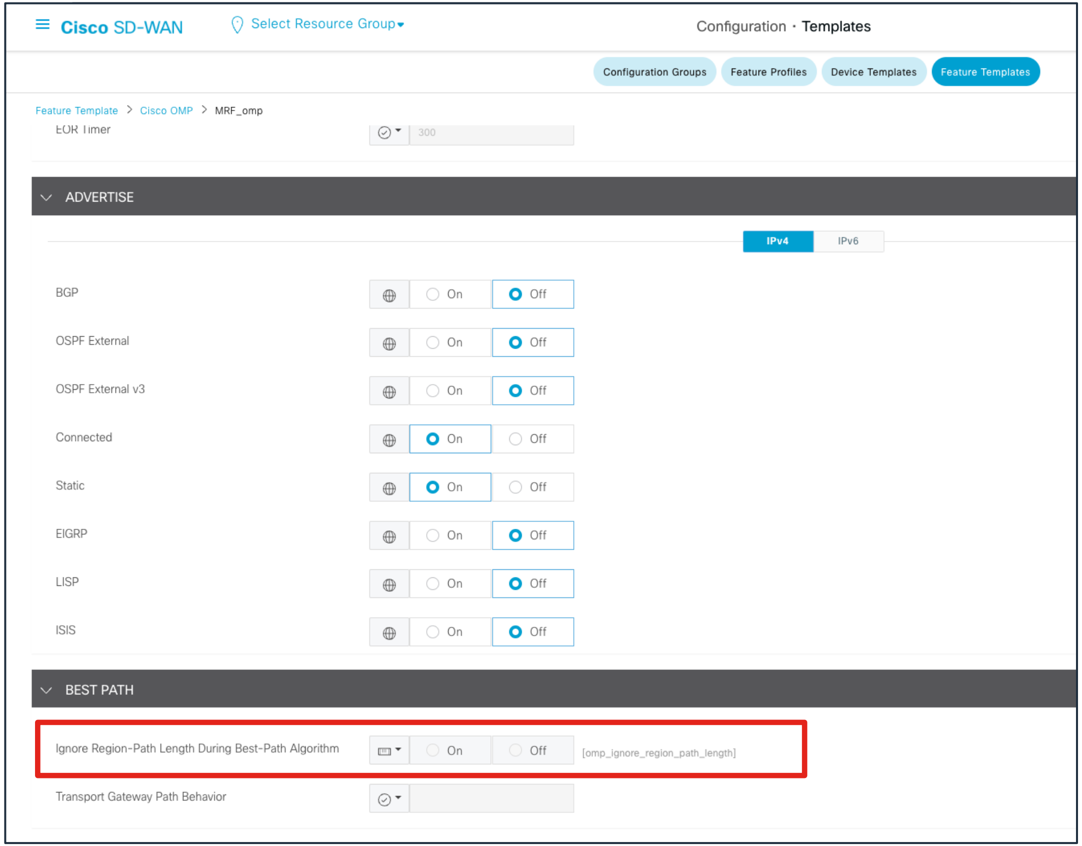
Preference

Enable Secondary region on transport interface

For Information

Configure OMP Path Selection

- By default, direct tunnel will be preferred over indirect tunnel due to shorter path
- Ability to change this behavior in OMP by ignoring path length
  - In that case, ECMP is used



EOR Timer 300

ADVERTISE

IPv4 | IPv6

BGP	○ On	● Off
OSPF External	○ On	● Off
OSPF External v3	○ On	● Off
Connected	● On	○ Off
Static	● On	○ Off
EIGRP	○ On	● Off
LISP	○ On	● Off
ISIS	○ On	● Off

BEST PATH

Ignore Region-Path Length During Best-Path Algorithm ○ On ○ Off [omp_ignore_region_path_length]

Transport Gateway Path Behavior

Summary

- Secondary Tunnel provides additional flexibility in a hierarchical SD - WAN network to connect regions directly (if possible)
- Available for both cEdge and vEdge
- Only valid between Edge routers not between ER-BR or BR-BR
- OMP option to ignore direct path over regional path and do ECMP
- Control Policy option to easily select Hierarchical path vs direct path.

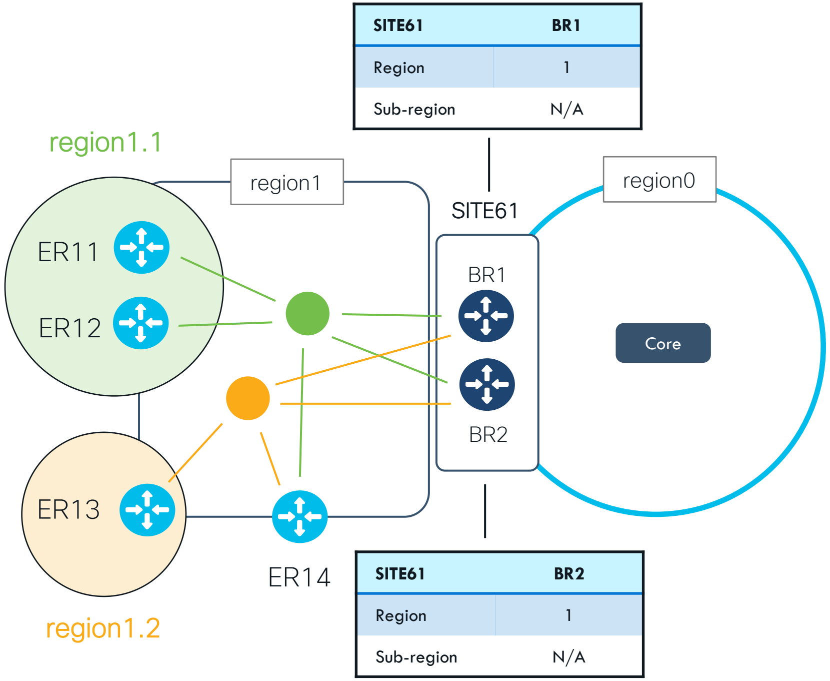
Sub Regions

Sub-regions

- Currently in MRF, the BR devices are dedicated to a region and ERs form full mesh tunnel within a primary region.
- However, some customers wants
  - Capability to share smaller regions on the same BRs to avoid cost implication of having dedicated BRs per region.
  - Also, they want BR failover in next geographical region to act as backup for their region.
- Sub-region are more granular form of regions under the same primary access region optionally configured on BR, ER and TR

Use Case 1 – Shared Border Routers

- Use case1: Shared Border between small regions
- Avoids cost implication for having dedicated BRs per region specially for smaller regions in a geographical area,
- Here BR1 and BR2 will be shared between all sub-regions and doesn't have any sub-region configs on them, just the primary region configured
- Notes:
  - Transport Router (TR) also supports sub-region
  - Router capabilities are exchanged in OMP
  - Not configured on vSmarts. But filtering of route and tlocs



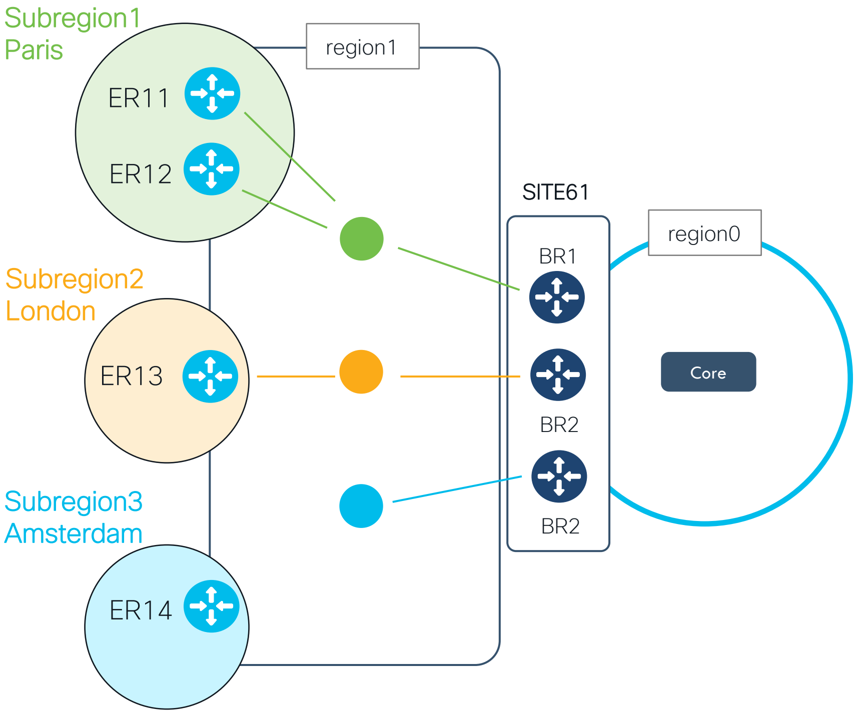
SITE61	BR1
Region	1
Sub-region	N/A

SITE61	BR2
Region	1
Sub-region	N/A

Shared BRs

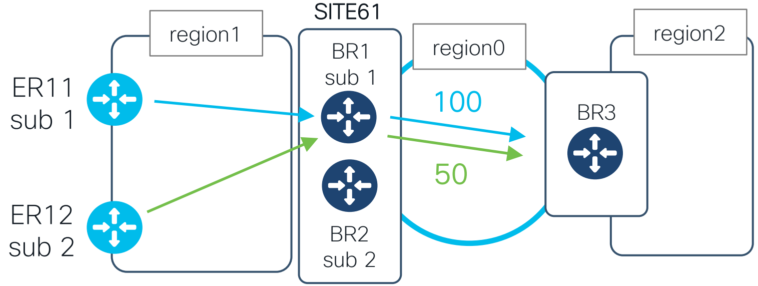
Use Case 2 – Dedicated BRs with Failover

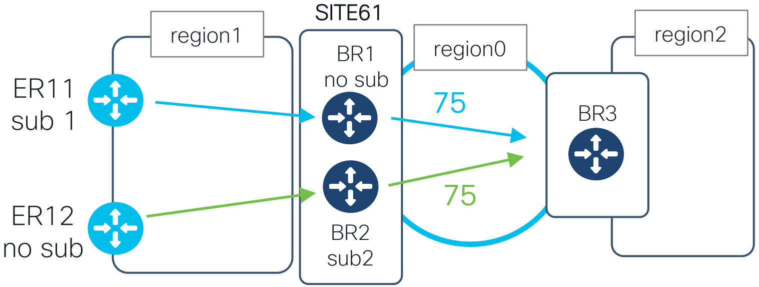
- Use case 2: Active/Backup Borders for neighboring regions (Dedicated BRs)
- Dedicated BR
- BR failover in next geographical region to act as backup for their region.
- Here BR1 is primary for Paris, while BR2 is primary for London. Also, BR1 act as backup for London and Amsterdam. BR2 backup for Paris and Amsterdam and so on.
- BRs will be configured with sub-regions in this dedicated BR model



BR Preference

- With the introduction of subregions, border routers add a new attribute, called br-preference, to access region routes that they re-originate to the core region.
- The br-preference attribute ensures that other border routers in the core region choose the optimal path to devices in subregions when more than one path is available for the return traffic. This is relevant to core-region only
- 3 types:
  - If the route has a subregion ID, and the BR has the same subregion ID. BR will advertise 100 (best)
  - If the route has a subregion ID, and the BR has a different subregion ID, BR will advertise 50 (worse)
  - If the route from ER has a NO subregion ID or BR has no sub-region ID configured, BR advertises preference of 75.





For Information

Configuration

```
system
 system-ip 1.1.1.11
 site-id 11
 region 1
  subregion 5
 role edge-router
```

Device is configured to be in Subregion 5 of Region 1
No vManage UI support in 20.10

For Information

Rules on ER

- Forms tunnels with all devices matching its sub-region and blocks tunnel to any devices not matching its sub-region or primary region
- Form tunnels with all BRs configured with same primary region id with or without sub-region configured
- Form tunnels with ER configured with only primary region E.g ER14 doesn’t have sub-region forms tunnels with all devices in the primary region configured with specific sub-regions (migration support)
- Secondary Region (direct tunnel) can also be formed between sub-region devices
- Routes:
  - If ER has sub-region id configured it prefer routes which has matching subregion-id over mis-matching or no sub-region configured
  - If ER has doesn’t have subregion id, it prefer routes with no-subregion over any sub-region configured routes
  - ER prefers matching sub-region for BRs over BR affinity, which comes after sub-region id check in the omp best-path-algo

For Information

Rules on BR

- Forms tunnels with all devices matching its primary region including ER with or without sub-regions
- Act as active for devices matching its sub-region
- Act as backup for devices matching its primary region but not sub-region
- If more than 1 backup BR available traffic gets ECMp’ed on the remaining BR with non-matching sub-region or without sub-region configured equally
- Routes:
  - ER prefers routes from BR matching its sub-region over affinity, which comes next in the omp best-path-algo
  - BR resets sub-region attribute when re-originating routes from access to core
  - BR sets BR-preference when re-originating routes from access to core (more in later section)
  - BR sets its own sub-region when re-originating routes from core to access (if configured)

Key Takeaways

Multi-Region Fabric – the journey

Cisco SD-WAN

17.7 Phase 0

- Segment fabric into **multiple regions**, including a special region-0
- vSmart vRoute filtering based on region IDs (**no control policies needed!**)
- Flexibility to have different topologies in different regions
- Notion of ‘**roles**’
- **Hierarchical hop-by-hop routing** for IPv4 and IPv6 in overlay with automatic route re-origination and withdrawal
  (No more traffic blackholes)
- IPsec/GRE encapsulation per region
- Control policies based on region-id
- **Distributed vSmart for scaling**
- Simplification of control policies

17.8 Phase 1

- Support for direct **inter-region tunnels**
- SD-WAN **policy evolution**
  BRs to support policies for traffic in/out from/to core/access
- **Connect discontiguous WAN sites via one-click** - Transport Gateway capability
- **Simplify dataplane horizontal scale-out** (throughput/tunnel scale) Affinity-groups
- Large vSmart memory optimizations
- **OMP** ECMP **send-path limit** from vS to Edges enhanced from **16 to 32**

17.9 Phase 2

- Support for **brownfield** network (flat SD-WAN overlay) migration to H-SDWAN
- **SD-WAN policy evolution** Extend ability to use centralized policies to control/steer traffic based on user intent, even for direct inter-region tunnels
- **vSmart memory optimization** *(scale)*
- Intelligent outbound path filtering on vSmart using device affinity-group preference *(scale)*
- OMP route re-origination dampening *(scale)*

17.10 Phase 3

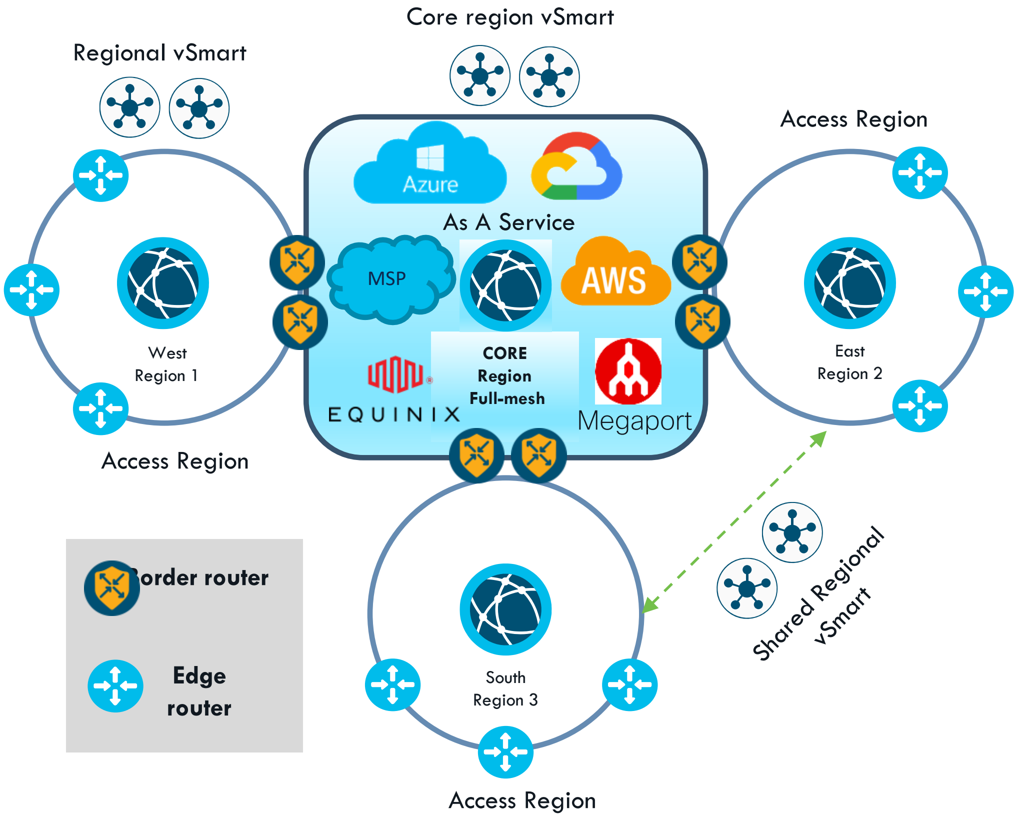
- Support for **sub-regions**
  - Ability to share BRs
  - Ability to failover across BRs
- **OMP enhancements**
  - Transport Gateway best path logic (introduced ‘site-type’)
  - OMP RIBout policy caching (scale)

17.11 Phase 4

- OMP **route aggregation** support at BR and TR
- **Smart filtering of paths** between vSmart and Edges (color-based)
- **Affinity-groups** support for **service-insertion**/chain
- Ability to set **affinity-groups** dynamically **via policy**
- OMP RIBout scale ENH

Hierarchical SD-WAN- Key takeaway

- Multi Region Fabric is the core enabler for architectures involving a middle-mile and lays the foundation for very large deployment
    - For Managed Services SD-WAN
    - Large Enterprise deployments using MSP/Cloud/SDCI backbone
- Lots of new features improving Cisco SD-WAN and bringing flexibility to support custom deployments



Complete your Session Survey

- Please complete your session survey after each session. Your feedback is important.
- Complete a minimum of 4 session surveys and the Overall Conference survey (open from Thursday) to receive your Cisco Live t-shirt.

- All surveys can be taken in the Cisco Events Mobile App or by logging in to the Session Catalog and clicking the "Attendee Dashboard” at https://www.ciscolive.com/emea/learn/sessions/session-catalog.html

Continue Your Education

- Visit the Cisco Showcase for related demos.
- Book your one-on-one Meet the Engineer meeting.
- Attend any of the related sessions at the DevNet, Capture the Flag, and Walk-in Labs zones.
- Visit the On-Demand Library for more sessions at ciscolive.com/on-demand.

cisco The bridge to possible

Thank you

cisco Live!

cisco Live!
ALL IN