

CISCO *Live!*



#CiscoLive



The bridge to possible

# How to Talk to the Rest of the World

External Connectivity for VXLAN EVPN Fabrics

Lukas Krattiger, Distinguished Engineer

@CCIE21921

BRKDCN-2267



#CiscoLive

# 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

Webex spaces will be moderated by the speaker until June 17, 2022.



<https://ciscolive.ciscoevents.com/ciscolivebot/#BRKDCN-2267>

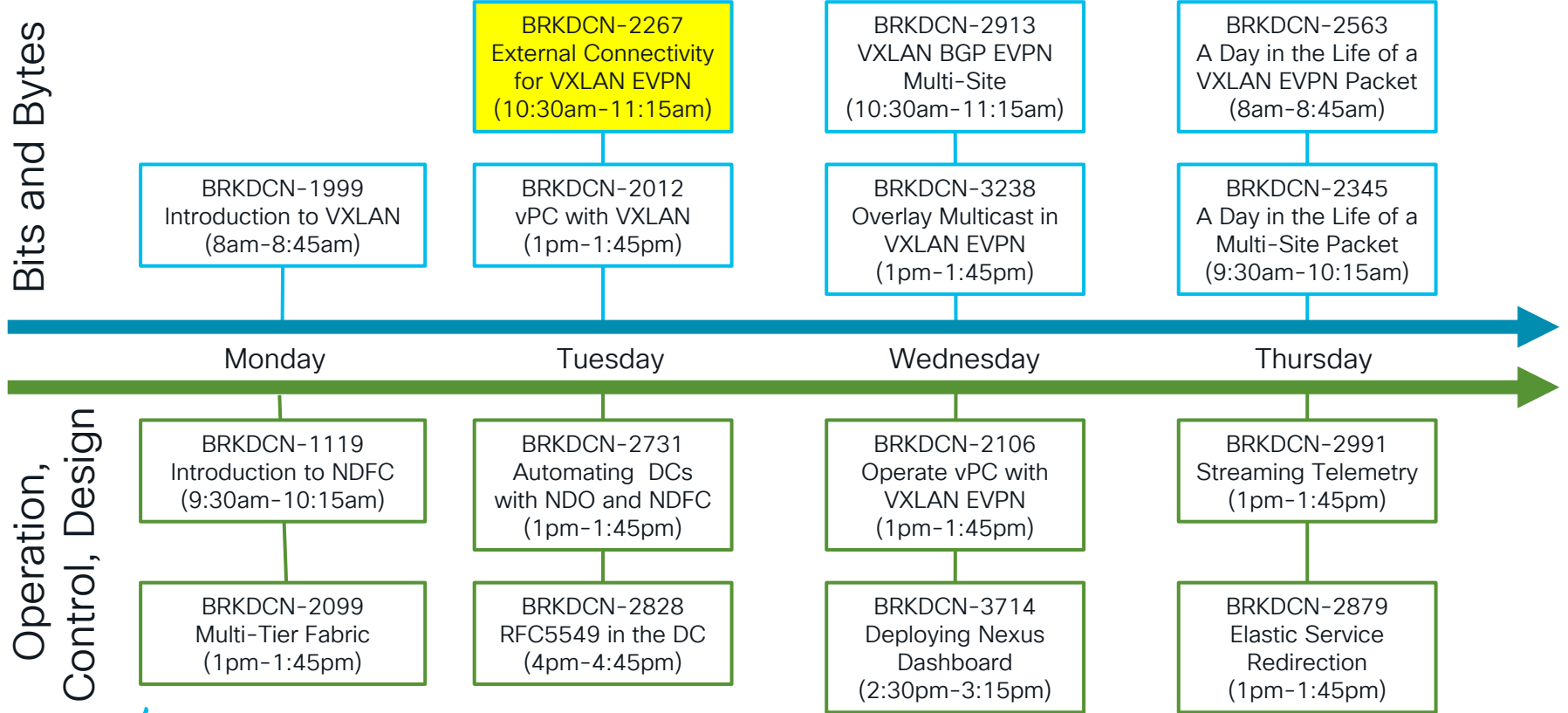
# Abstract

A super long set of acronyms VXLAN EVPN, we want to get you the details to get you started. The design and attachment for your VXLAN EVPN Fabric is paramount. How to best design that attachment and benefit from redundancy, scalability and wide ECMP is key. We are going to talk about an Inter-AS Option A (aka VRF-lite) approach but also expand into the Gateway capabilities with MPLS (LDP) and Segment Routing.

# Introduction

- A brief Overview on how VXLAN EVPN works
- Looking at the Control- and Data-Plane protocols that are being used
  - We focus on the Fabric internal part
- Best Practices on External Connectivity (Inter-AS Option A)
  - Some people call it L3Out, others VRF-lite
  - Details and Challenges
- What else do we have, other than “Inter-AS Option A”?
  - Multi-Site aka DCI Overlay
  - MPLS, MPLS-SR, SRv6 Gateways

# Companion Sessions – Week at a Glance





# Agenda

- Introduction on VXLAN EVPN
- External Connectivity Scenarios
- Conclusion

# Introduction



# What is VXLAN?

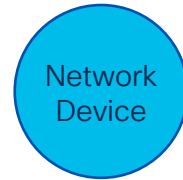
- Standards based Encapsulation
- RFC 7348
- Uses UDP-Encapsulation
- Transport Independent
- Layer-3 Transport (Underlay)
- Flexible Namespace
- 24-bit field (VNID) provides ~16M unique identifier
- Allows Segmentations

# What is EVPN?

- Standards based Control-Plane
- RFC 8365 (and RFC 7432)
- Uses Multiprotocol BGP
- Uses Various Data-Planes
- VXLAN (EVPN-Overlay), MPLS, Provider Backbone (PBB)
- Many Use-Cases Covered
- Bridging, MAC Mobility, First-Hop & Prefix Routing, Multi-Tenancy (VPN)

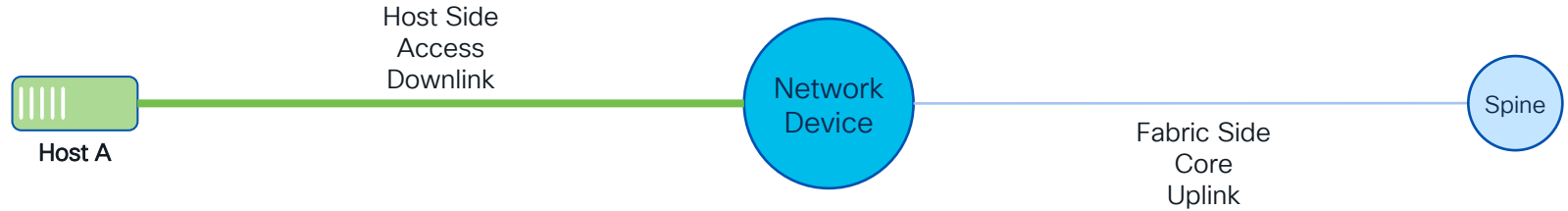
# It all starts with a Network Device

The Dating Network - When Control- meets Data-Plane



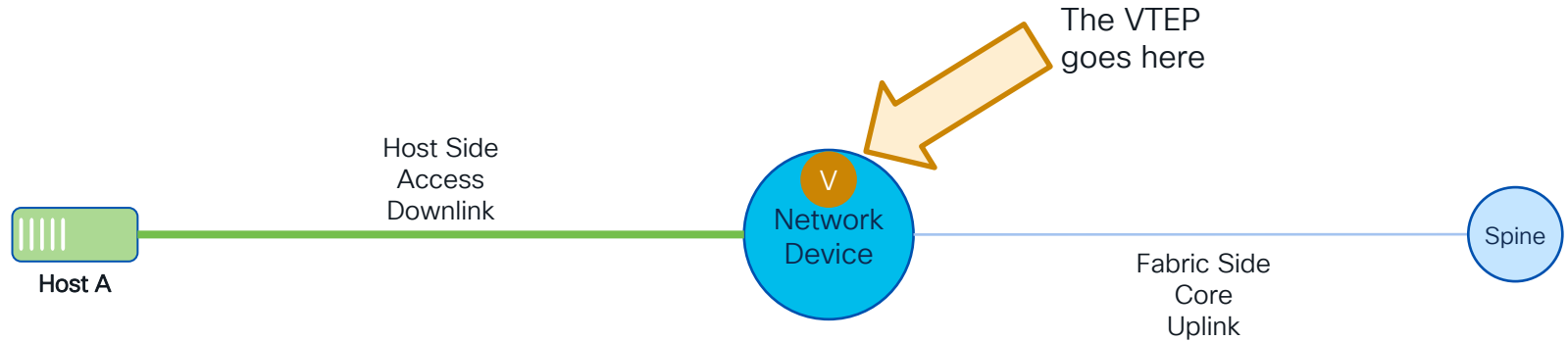
# It all starts with a Network Device

The Dating Network - When Control- meets Data-Plane



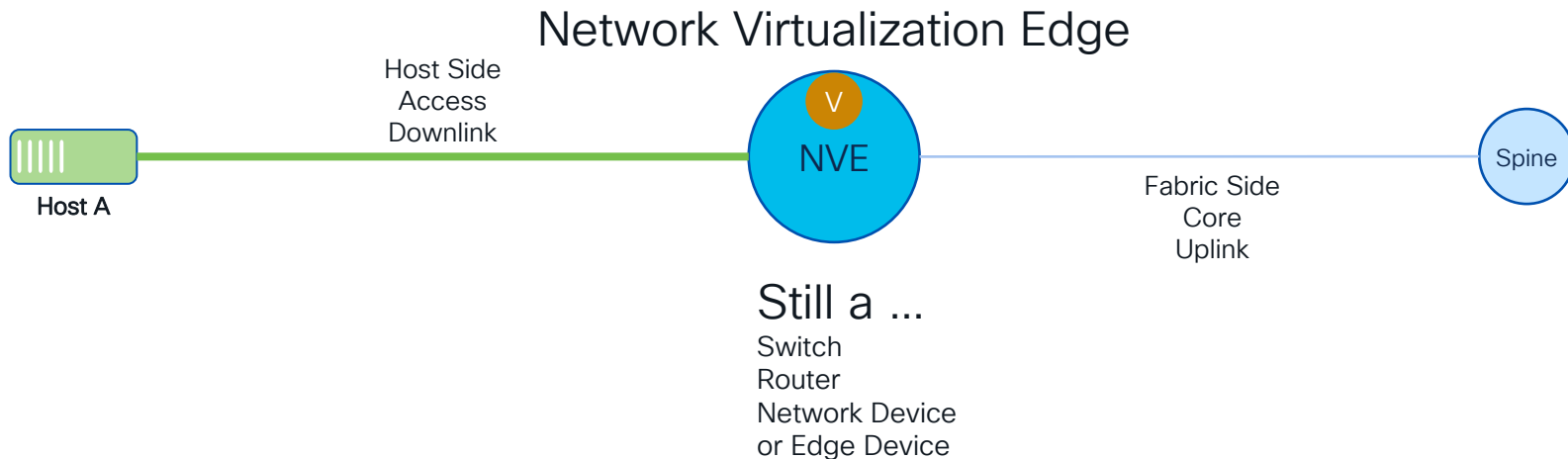
# It all starts with a Network Device

## The Dating Network - When Control- meets Data-Plane



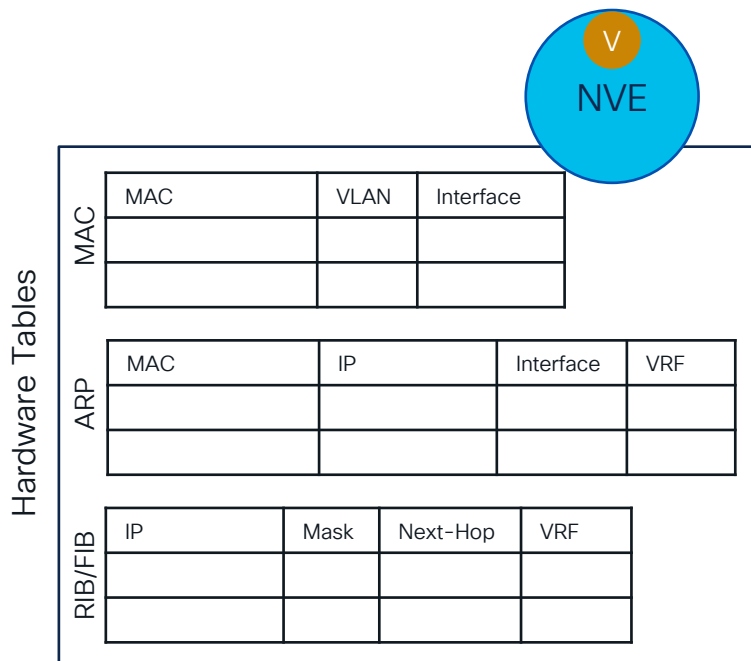
# Making the Network Device an NVE

## The Dating Network - When Control- meets Data-Plane



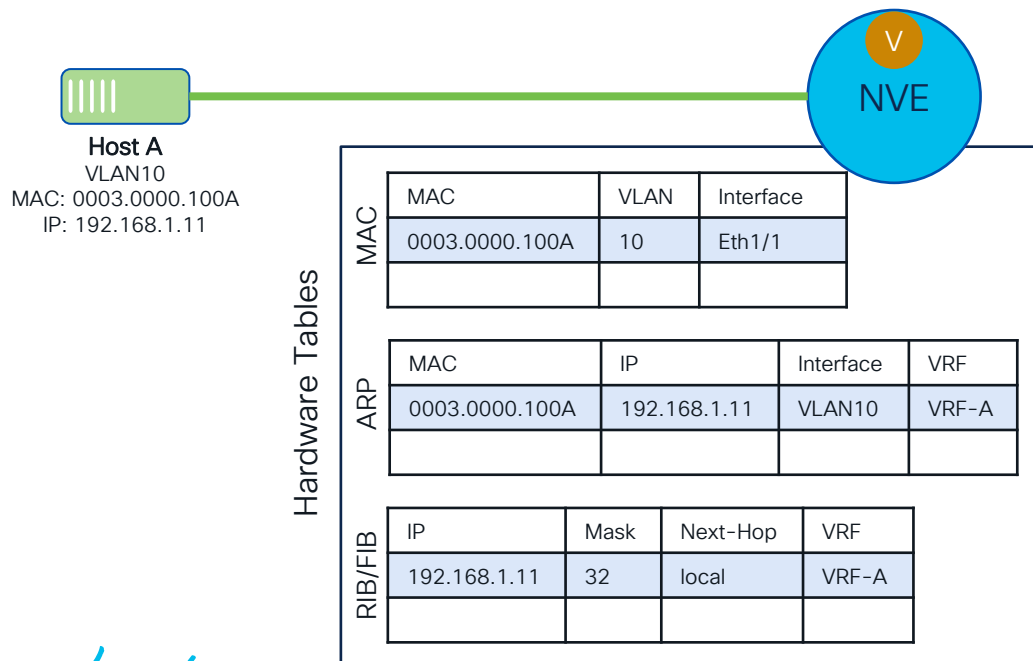
# The NVE and Some Important Table

## The Dating Network - When Control- meets Data-Plane



# Local Learning on the NVE

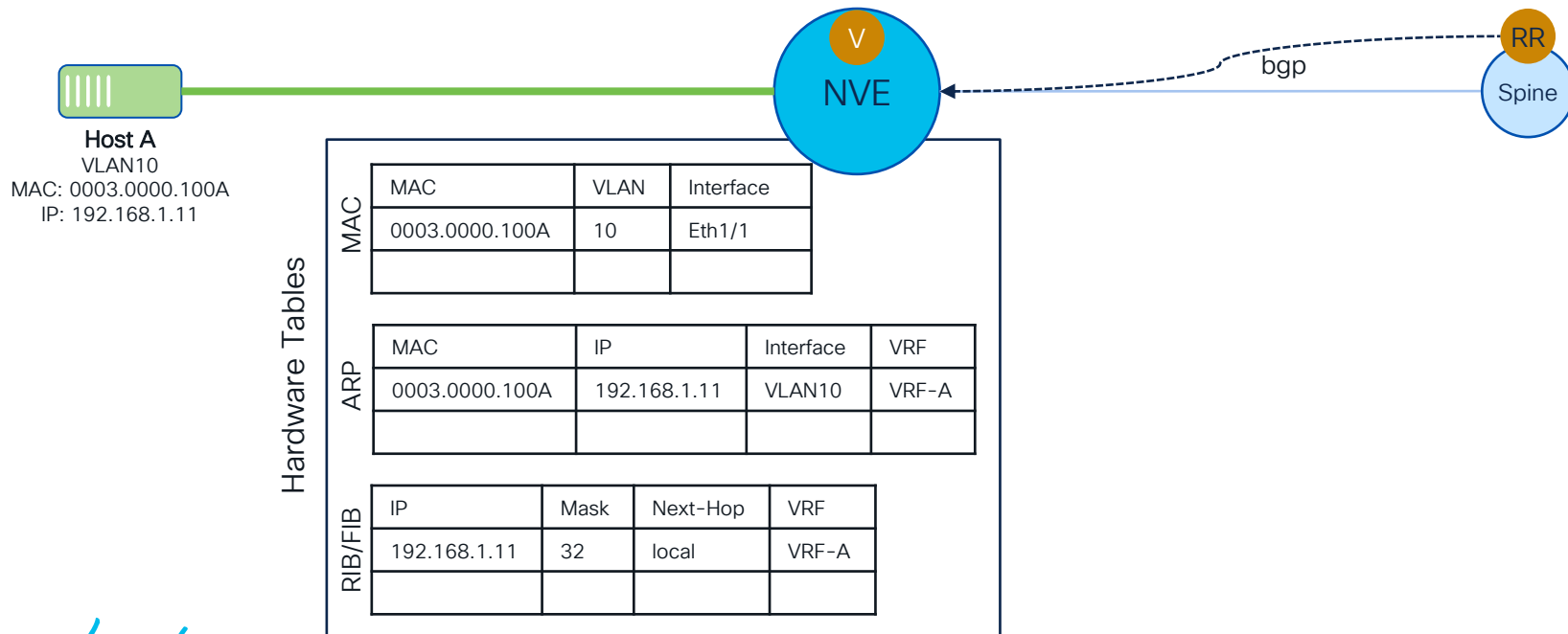
## The Dating Network - When Control- meets Data-Plane





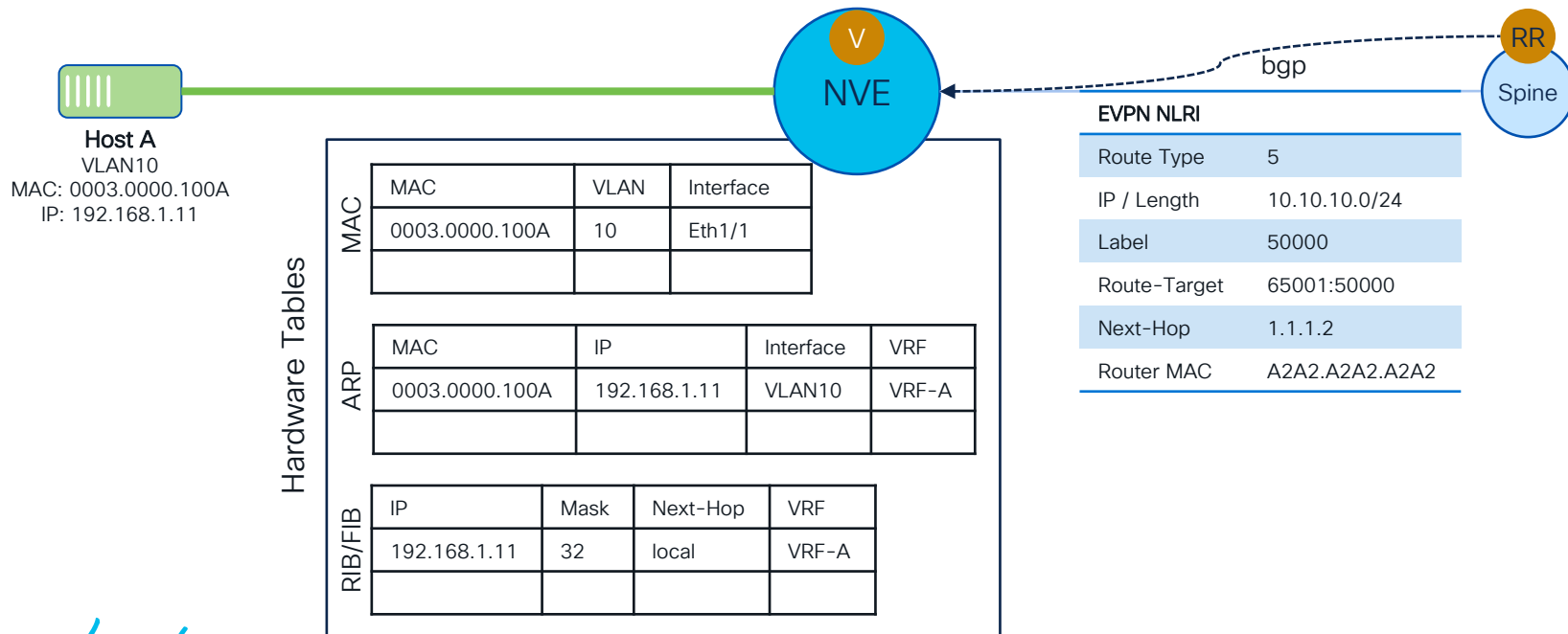
# Remote Learning from other NVE

## The Dating Network - When Control- meets Data-Plane



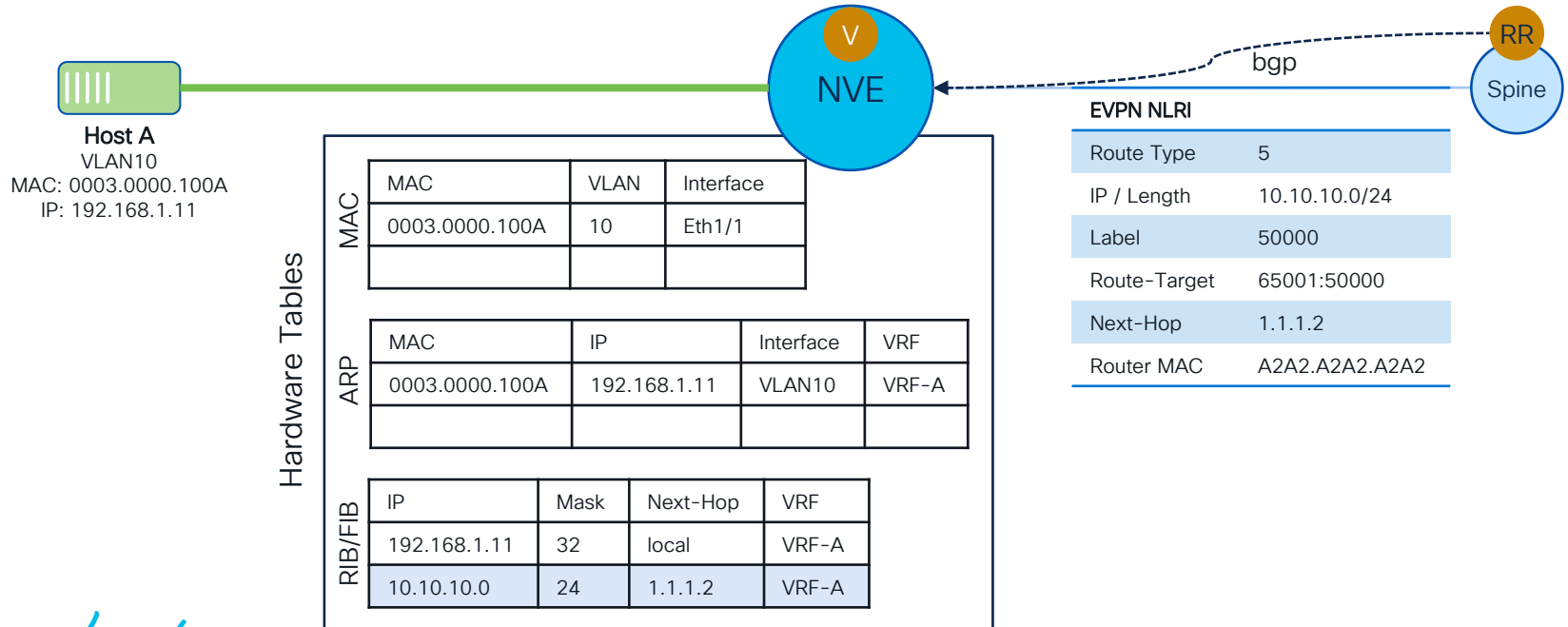
# Remote Learning from other NVE

## The Dating Network - When Control- meets Data-Plane



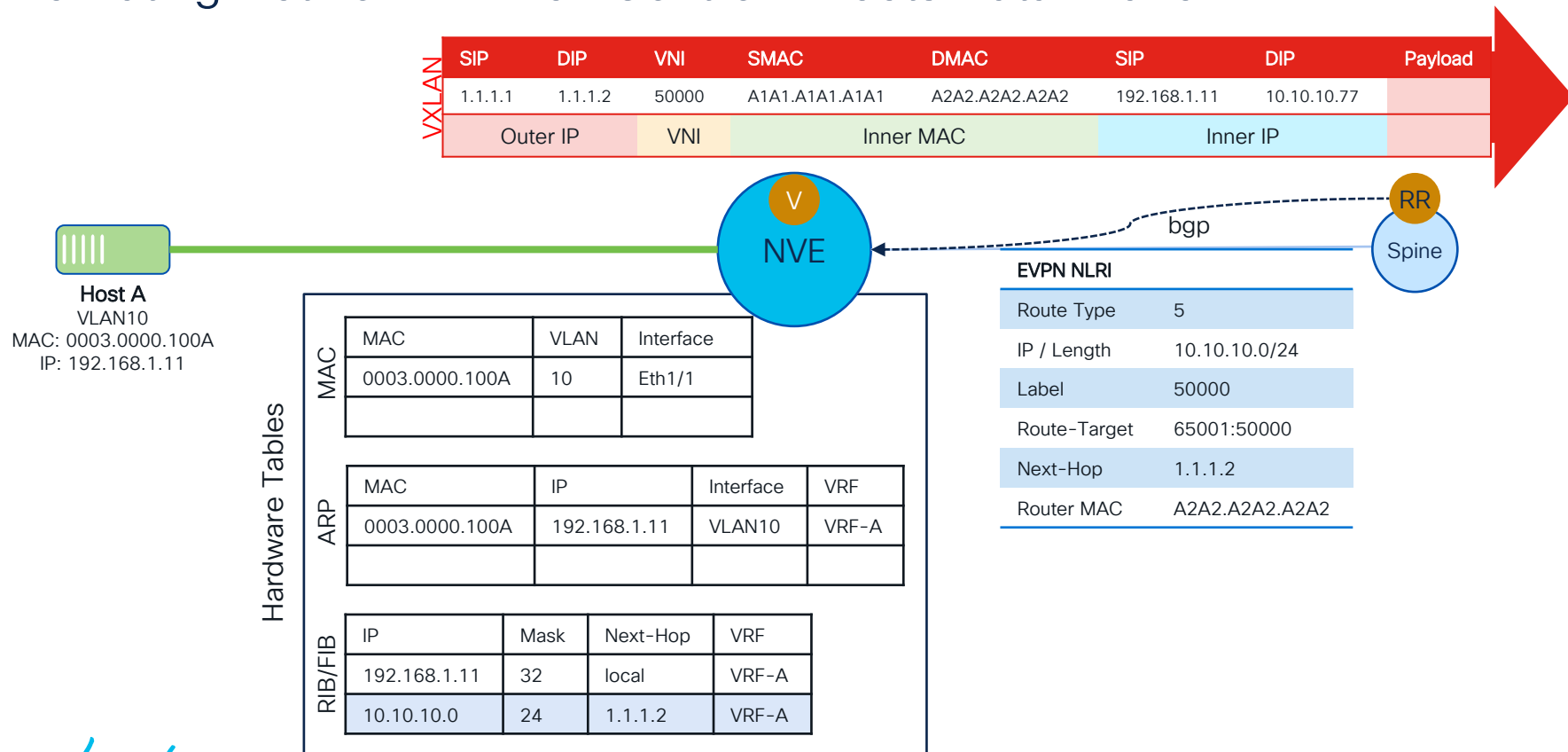
# Remote Learning from other NVE

## The Dating Network - When Control- meets Data-Plane



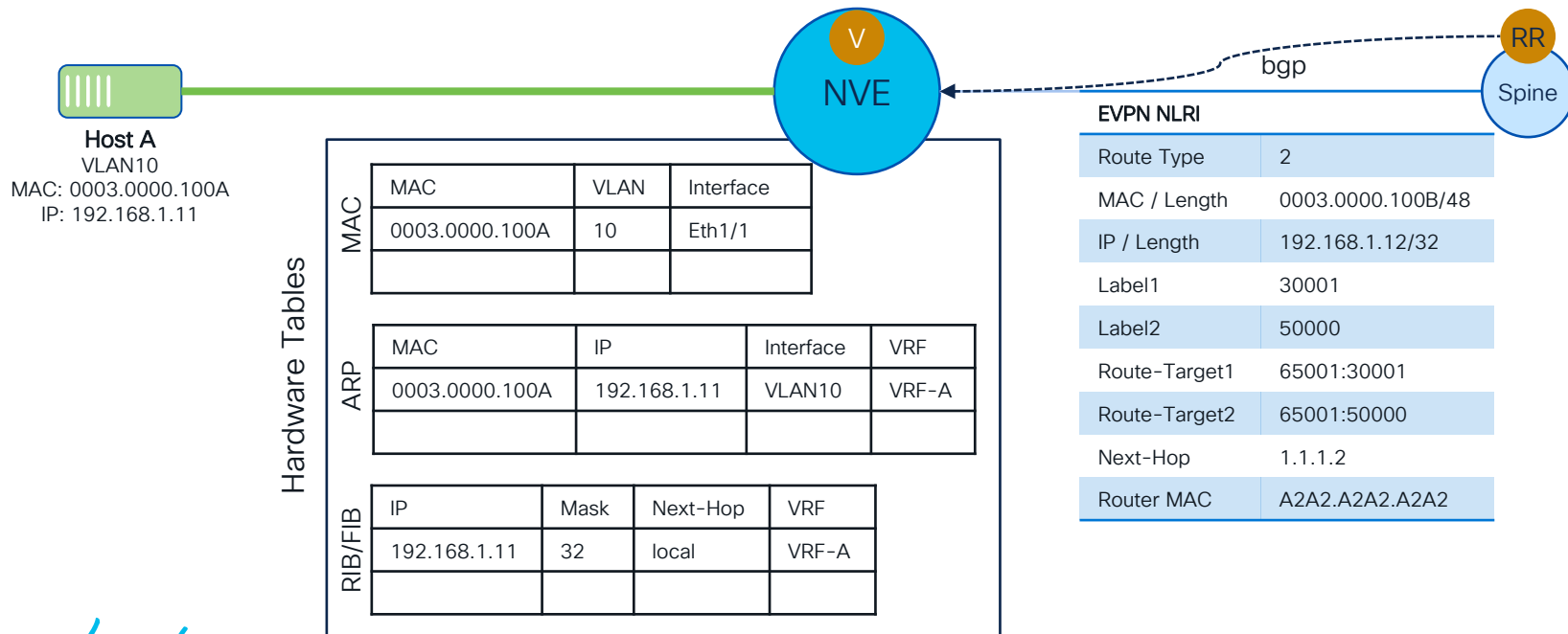
# Routing between NVE (based on VXLAN EVPN)

## The Dating Network - When Control- meets Data-Plane



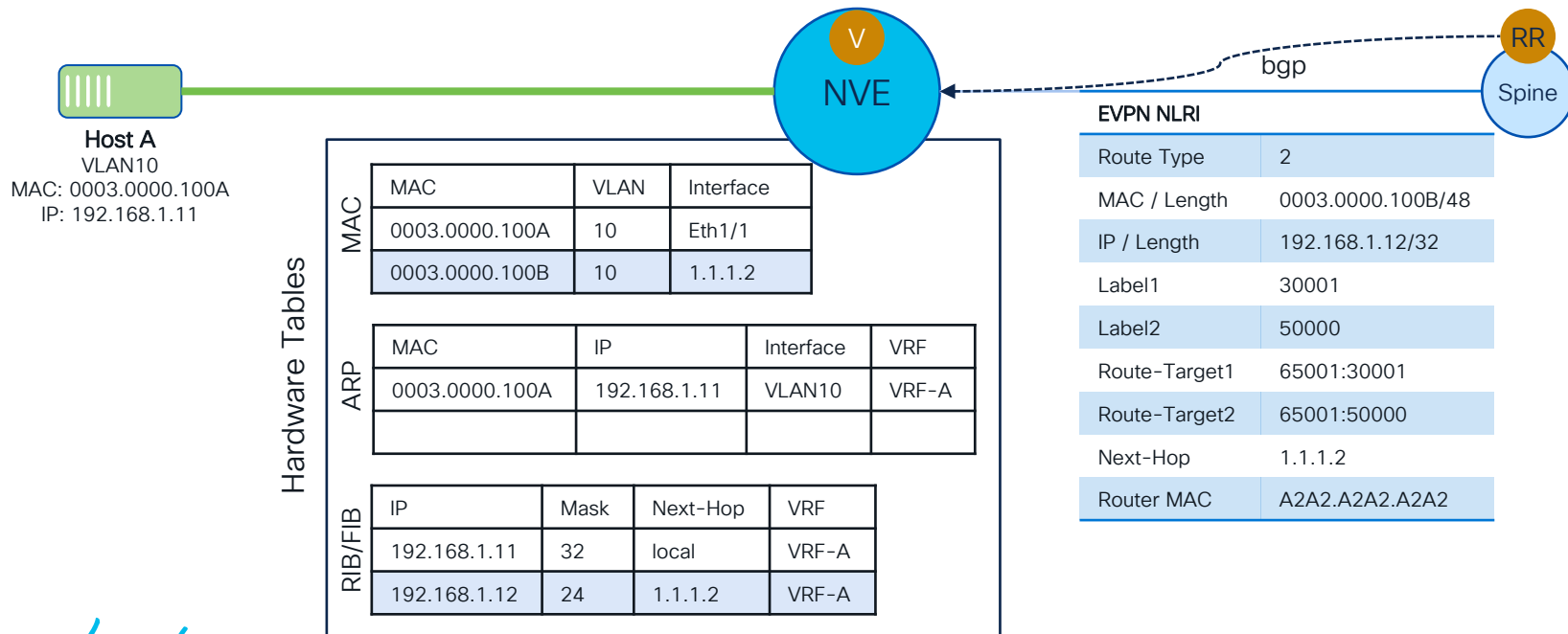
# Remote Learning from other NVE

## The Dating Network - When Control- meets Data-Plane



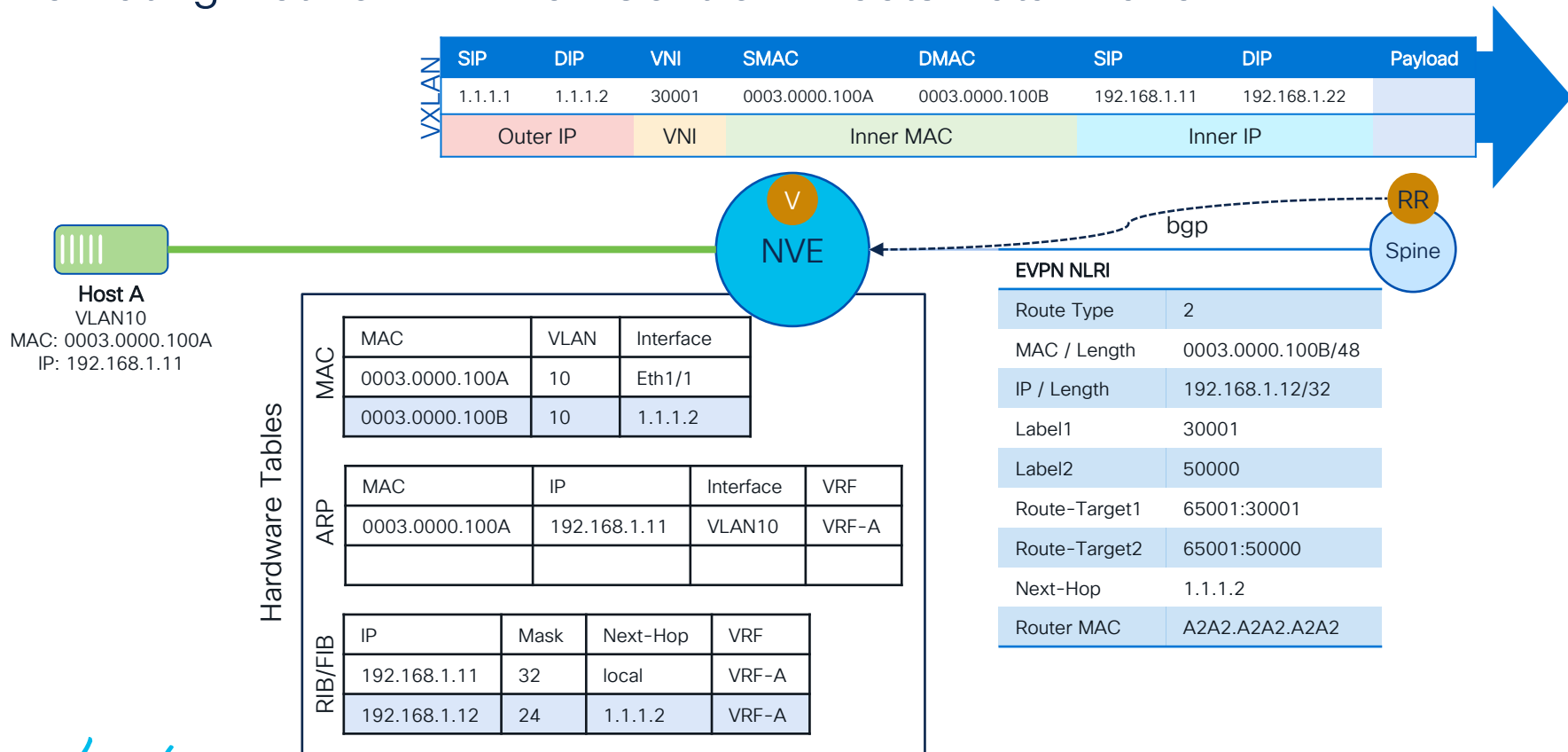
# Remote Learning from other NVE

## The Dating Network - When Control- meets Data-Plane



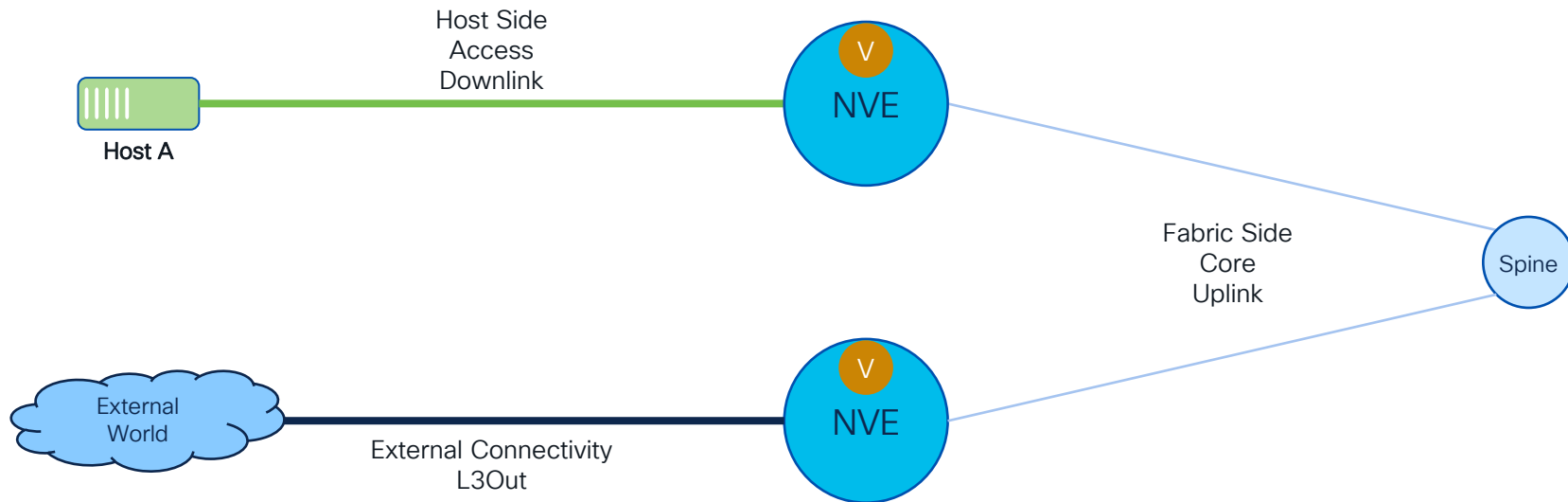
# Bridging between NVE (based on VXLAN EVPN)

## The Dating Network - When Control- meets Data-Plane



# Expanding to the Border – A “special” NVE

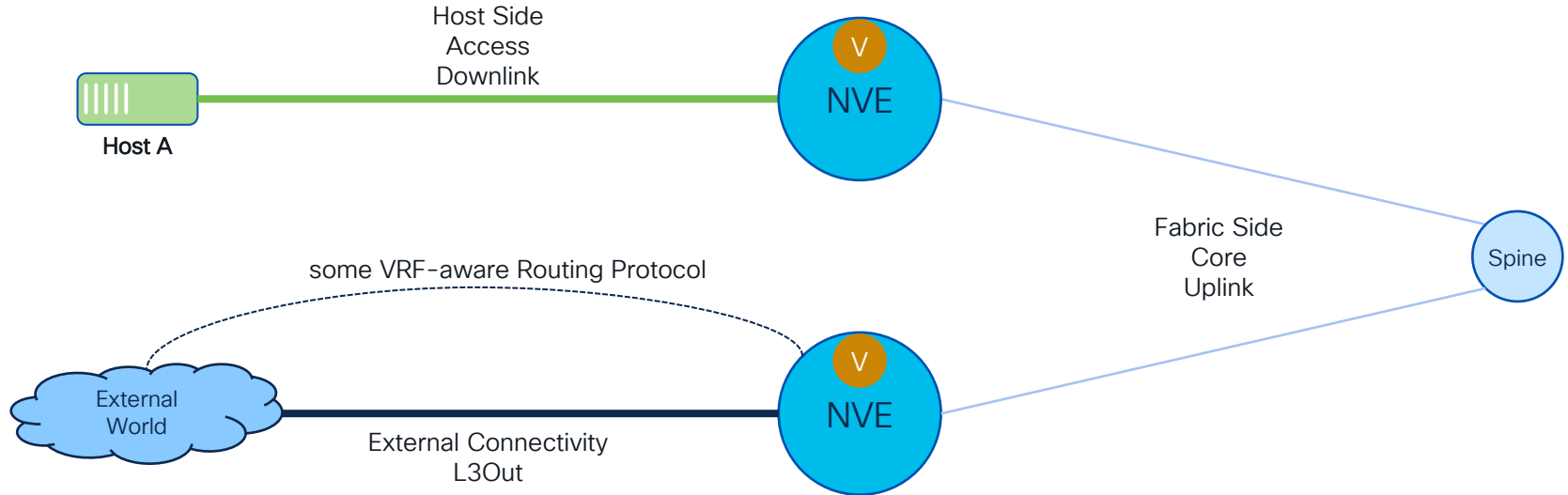
## The Dating Network – When Control- meets Data-Plane





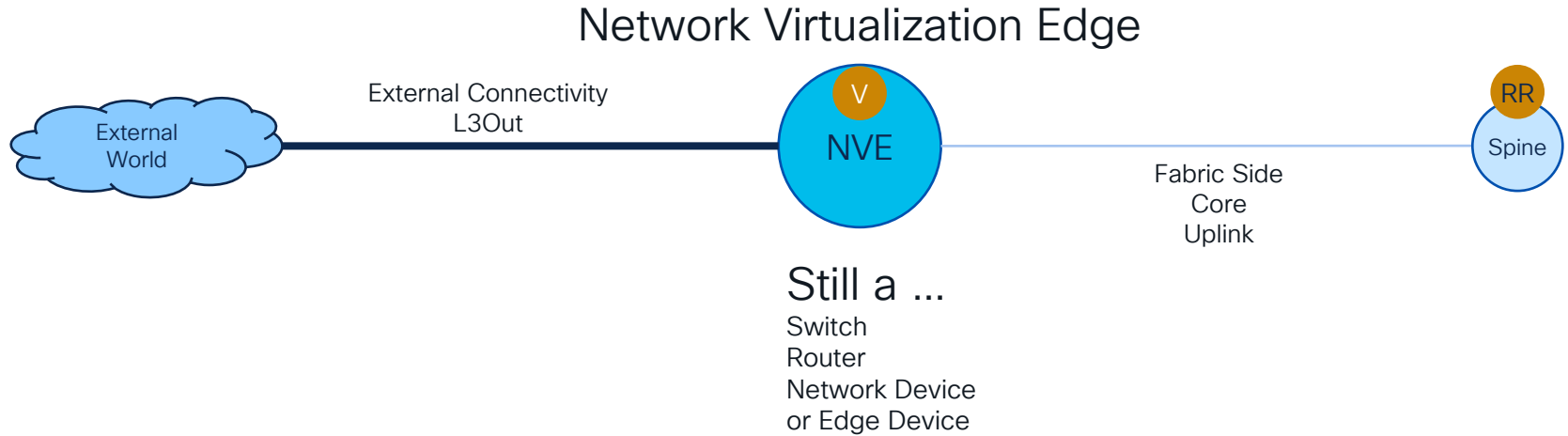
# Expanding to the Border – A “special” NVE

## The Dating Network – When Control- meets Data-Plane



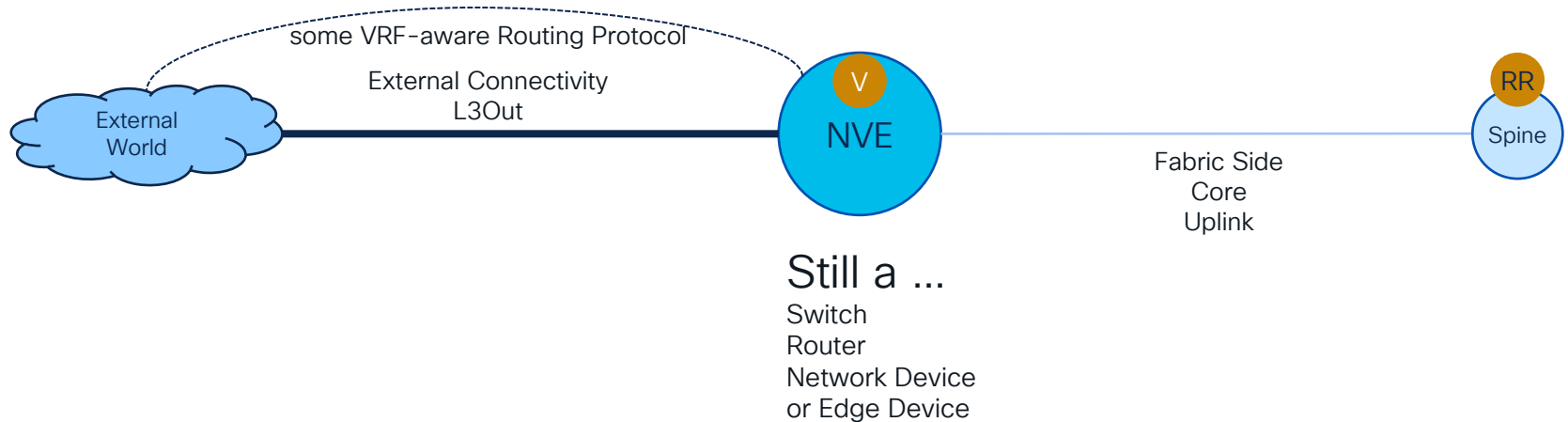
# Expanding to the Border – A “special” NVE

## The Dating Network – When Control- meets Data-Plane



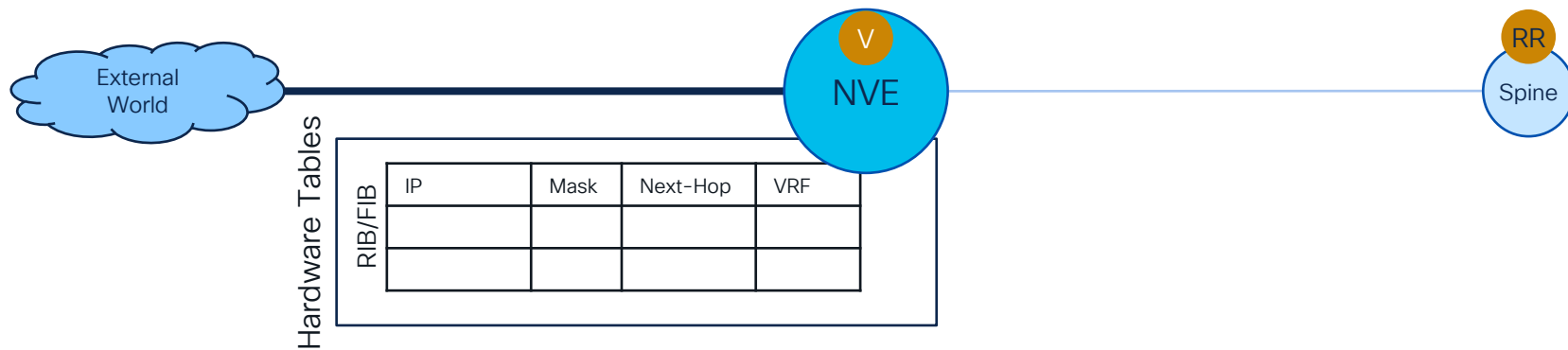
# Expanding to the Border – A “special” NVE

## The Dating Network – When Control- meets Data-Plane



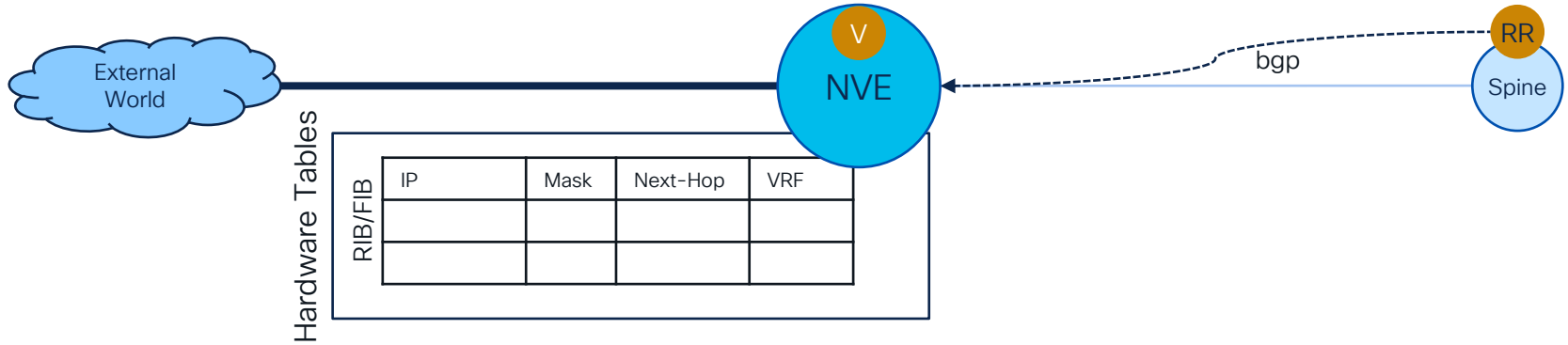
# The Border NVE and Some Important Table

## The Dating Network - When Control- meets Data-Plane



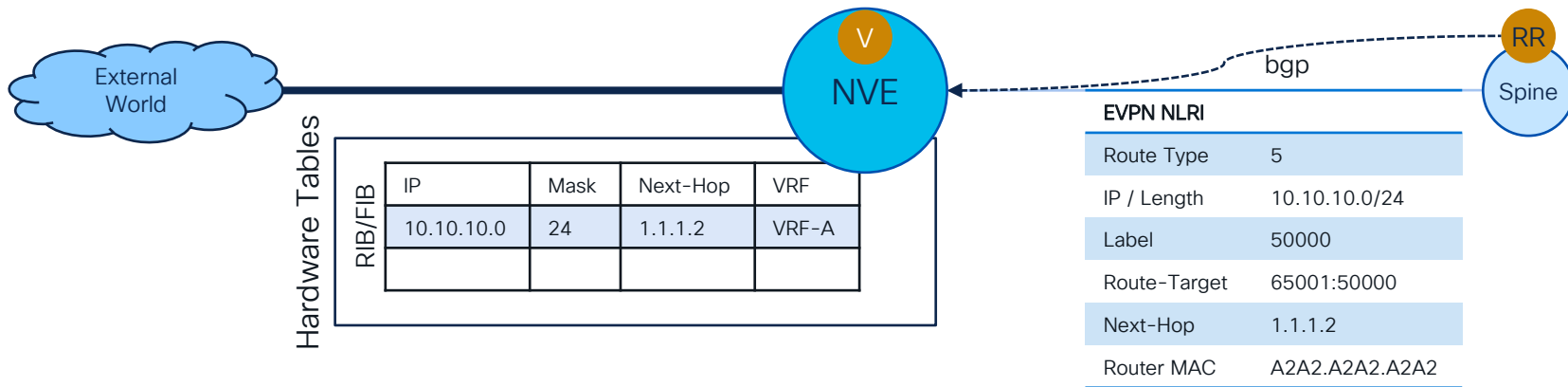
# Always Remote Learning on a Border NVE

## The Dating Network - When Control- meets Data-Plane



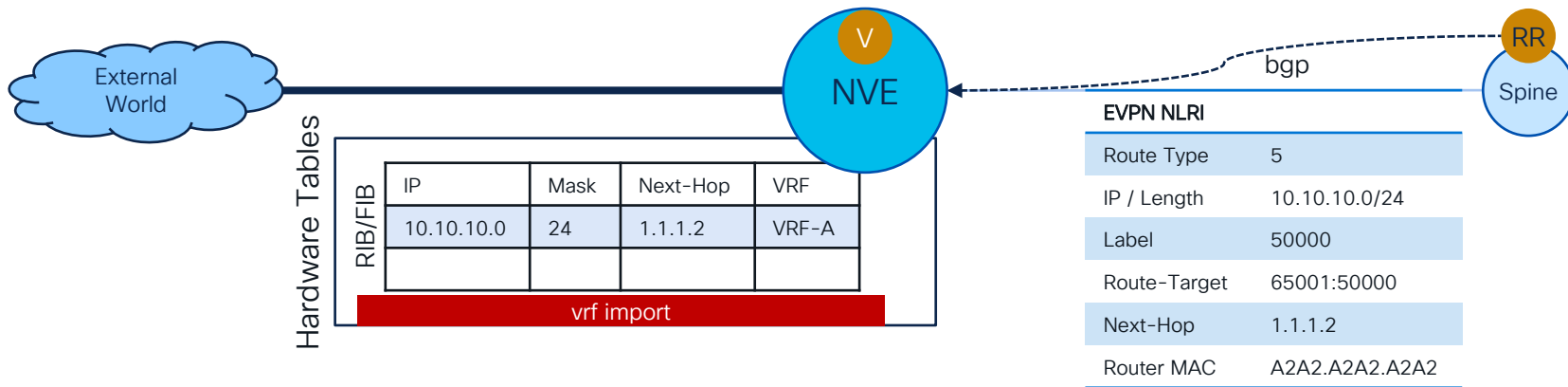
# Remote Learning from Fabric Internal (NVE)

## The Dating Network - When Control- meets Data-Plane



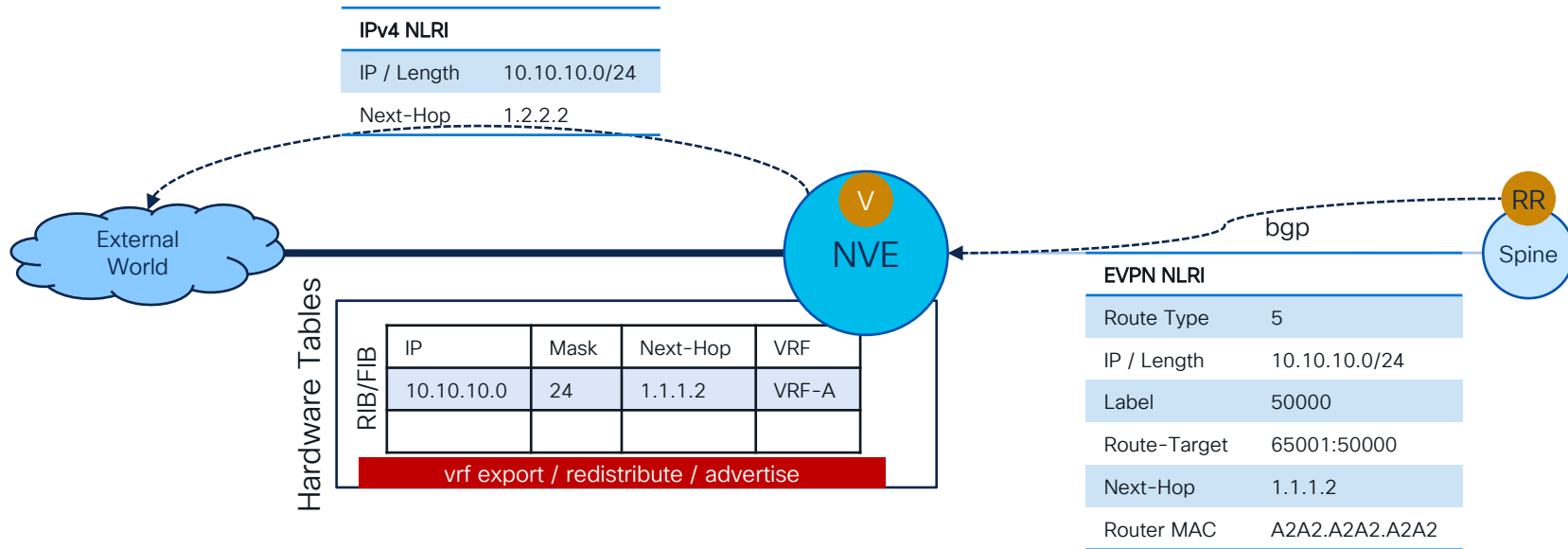
# Remote Learning from Fabric Internal (NVE)

## The Dating Network - When Control- meets Data-Plane



# Remote Learning from Fabric Internal (NVE)

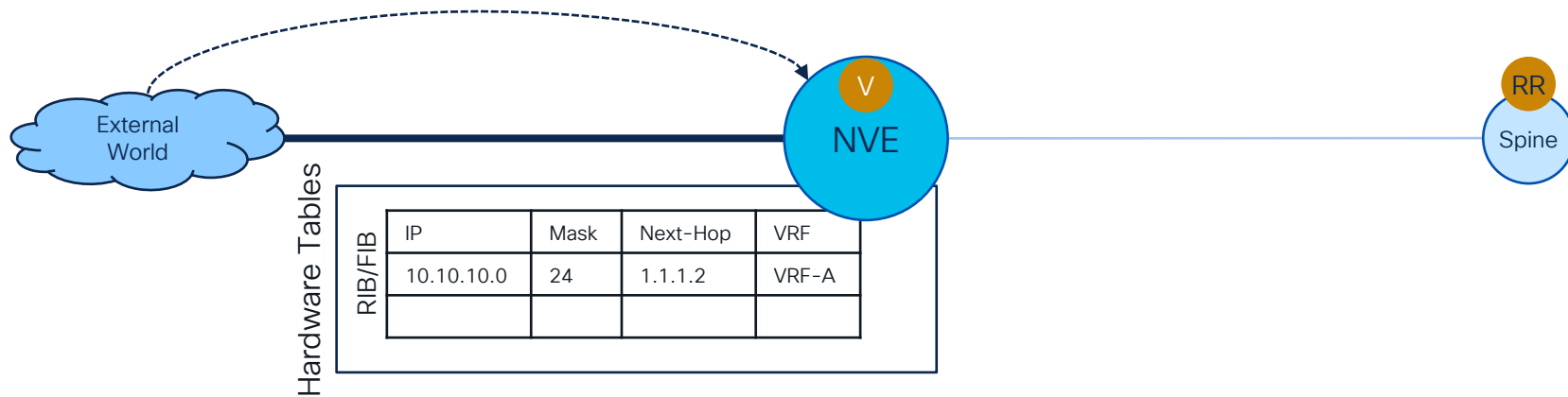
## The Dating Network - When Control- meets Data-Plane





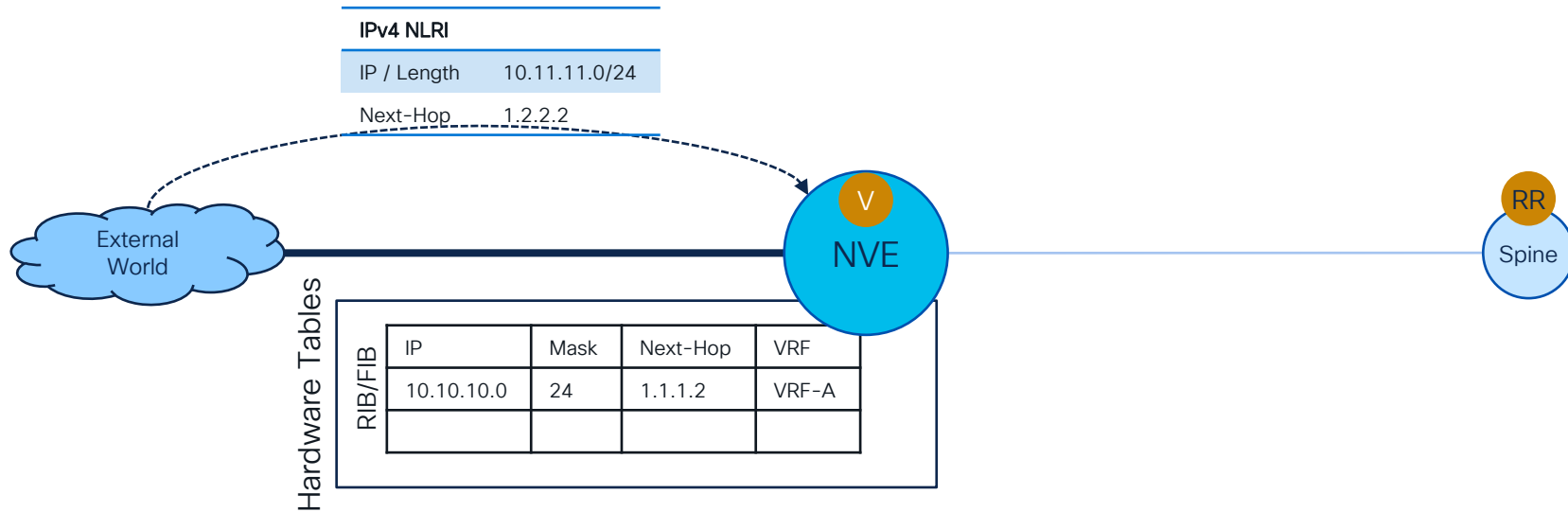
# Remote Learning from Fabric External (Outside)

## The Dating Network - When Control- meets Data-Plane



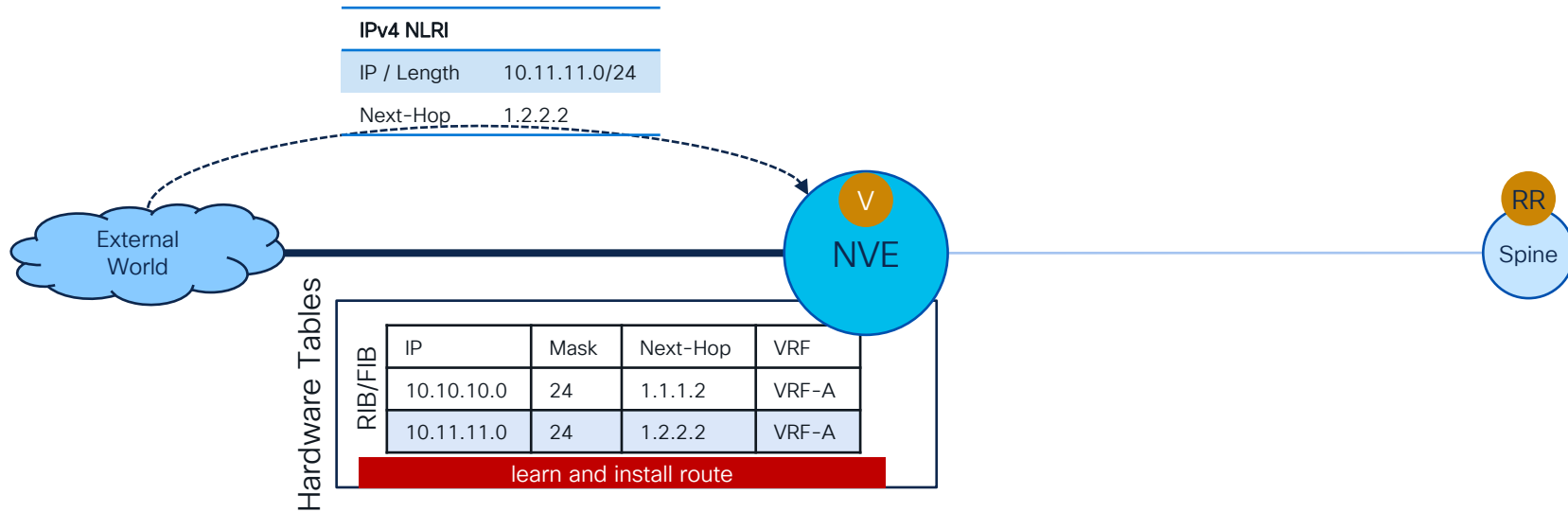
# Remote Learning from Fabric External (Outside)

## The Dating Network - When Control- meets Data-Plane



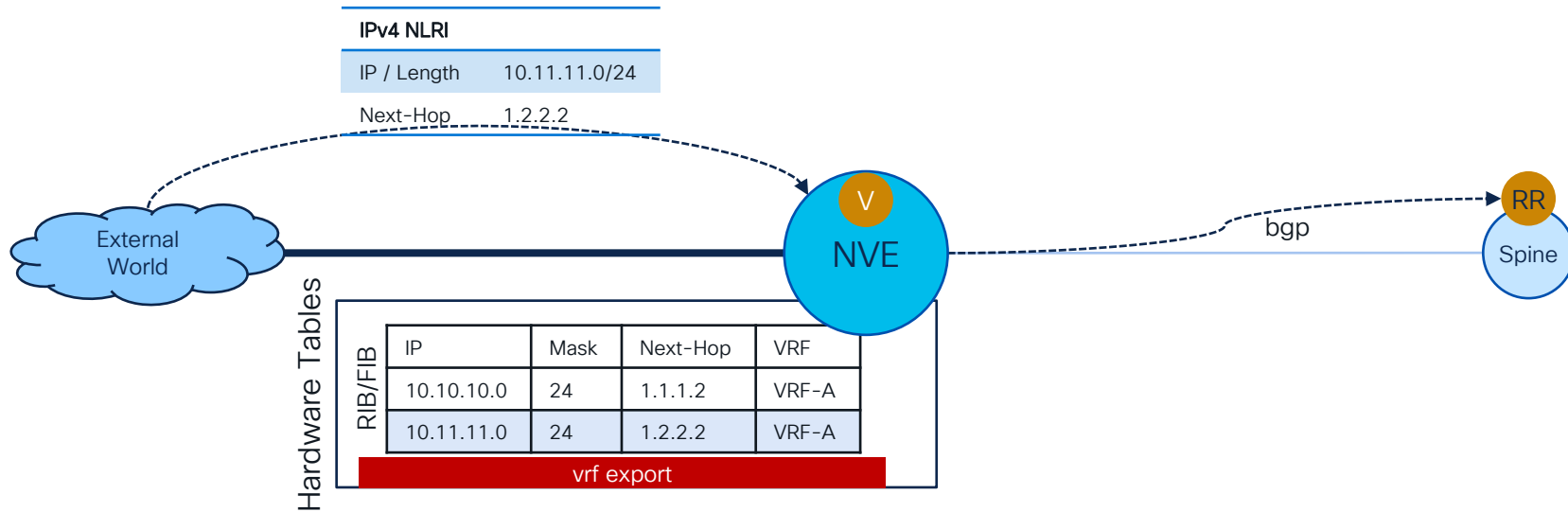
# Remote Learning from Fabric External (Outside)

## The Dating Network - When Control- meets Data-Plane



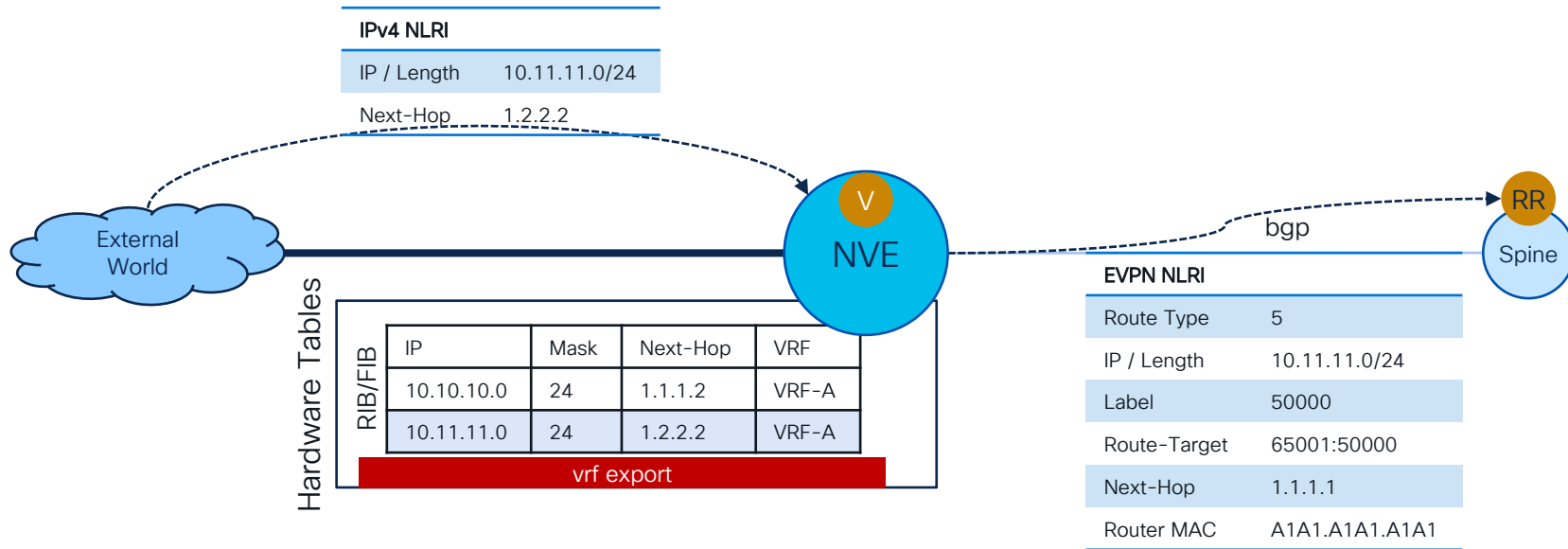
# Remote Learning from Fabric External (Outside)

## The Dating Network - When Control- meets Data-Plane



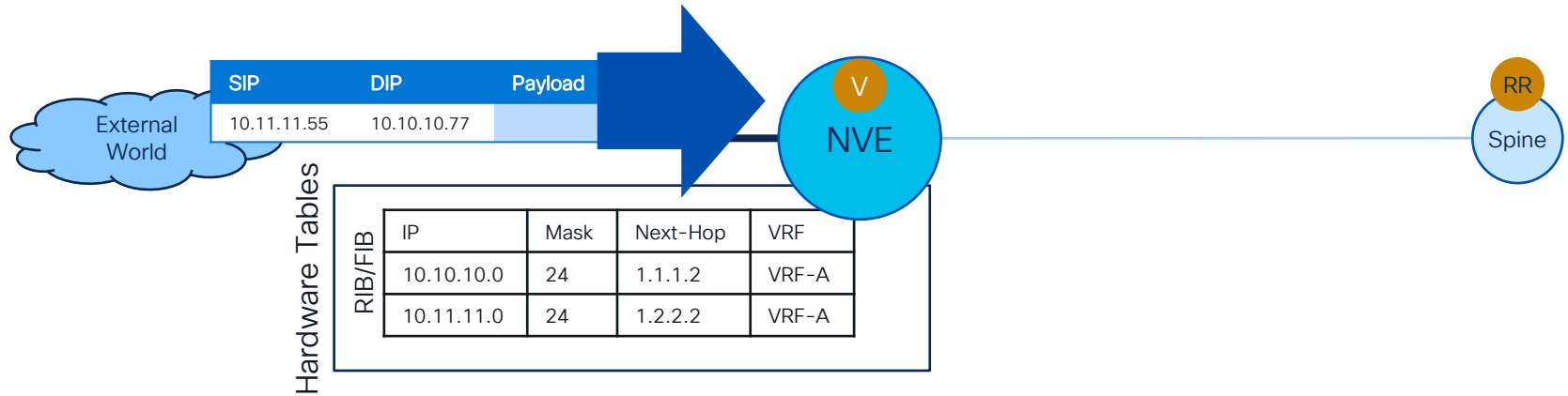
# Remote Learning from Fabric External (Outside)

## The Dating Network - When Control- meets Data-Plane



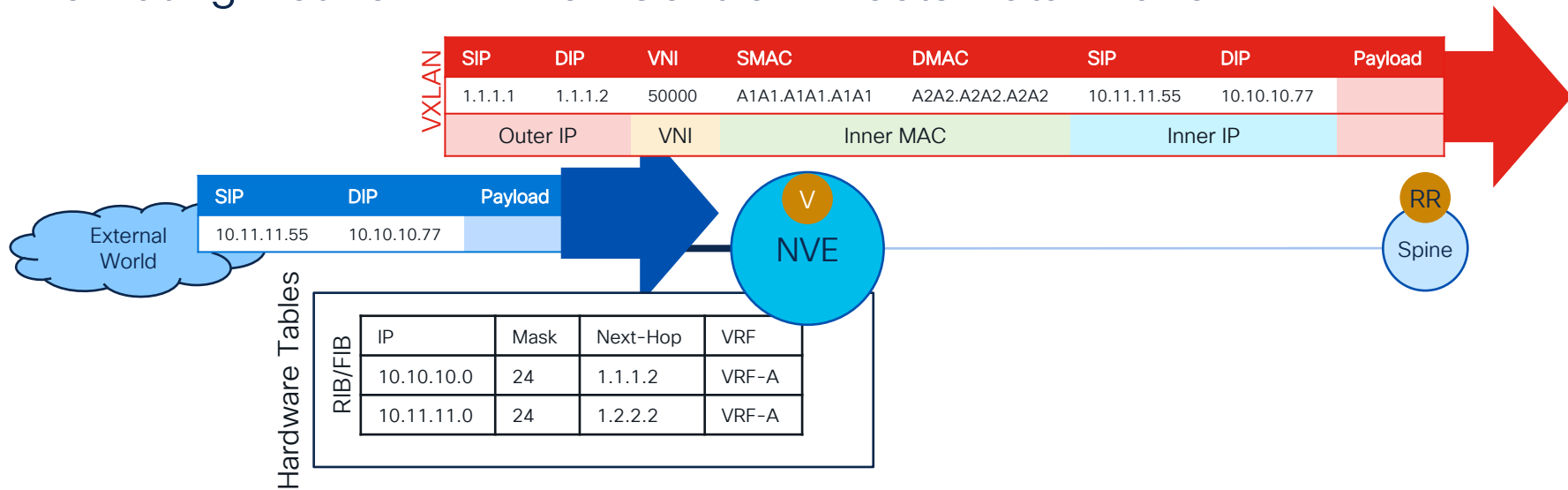
# Routing from External to Internal

## The Dating Network - When Control- meets Data-Plane



# Routing from External to Internal

## The Dating Network - When Control- meets Data-Plane





# *A Day in the Life of a VXLAN EVPN Packet*

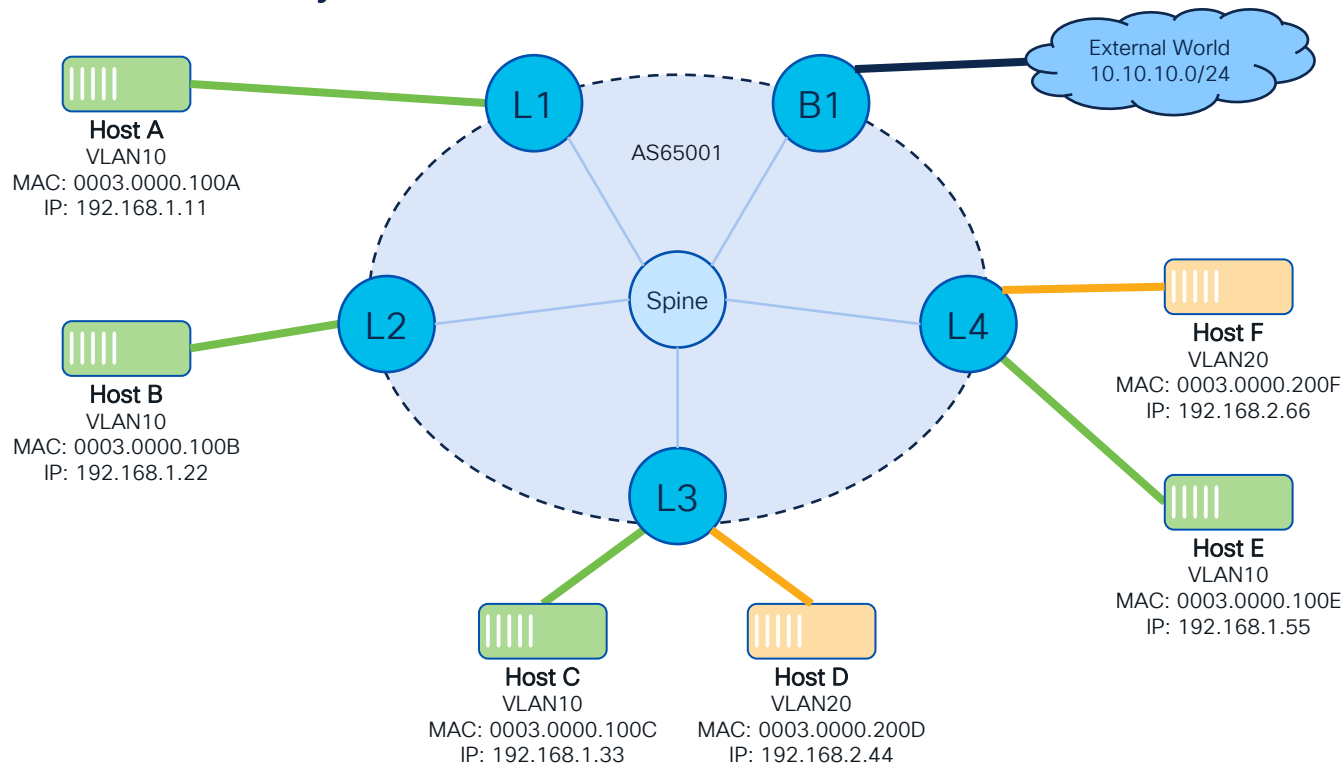
BRKDCN-2563



# External Connectivity Scenarios

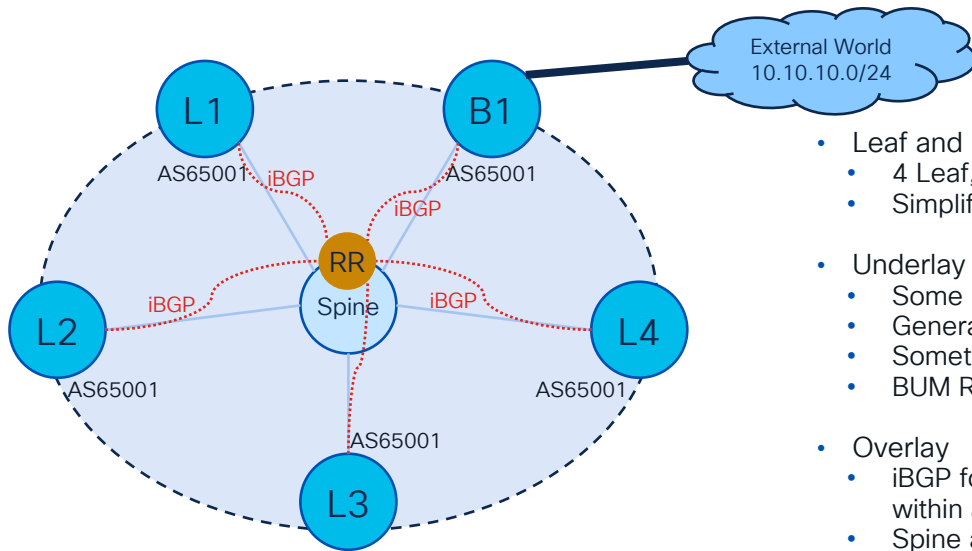
# Topology Overview

## External Connectivity



# Topology Overview – How is the Fabric built?

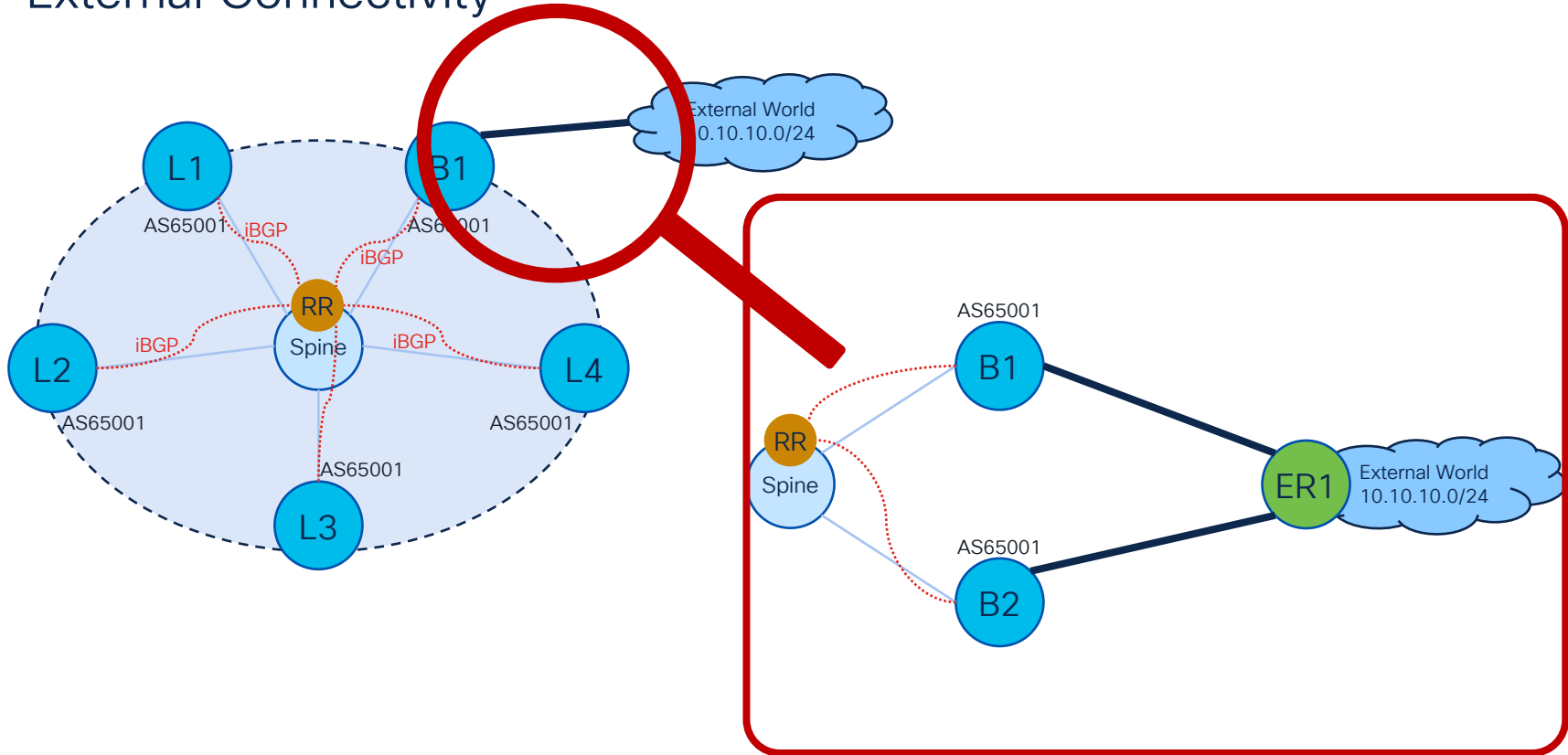
## External Connectivity



- Leaf and Spine Topology
  - 4 Leaf, 1 Spine
  - Simplified for readability
- Underlay
  - Some Routing Protocol for VTEP reachability
  - Generally, a IGP like OSPF (AD 110) or IS-IS (AD 115) for simplicity
  - Sometimes eBGP (AD 20) for some reasons\*
  - BUM Replication can be Multicast or Ingress/Head-End Replication
- Overlay
  - iBGP for EVPN Address-Family – There is no Inter-AS use case within a Fabric
  - Spine acts as BGP Route Reflector
  - iBGP peering goes Loopback to Loopback; no change on topology failures
  - All EVPN Routes have Administrative Distance 200

# Let's Focus Here

## External Connectivity

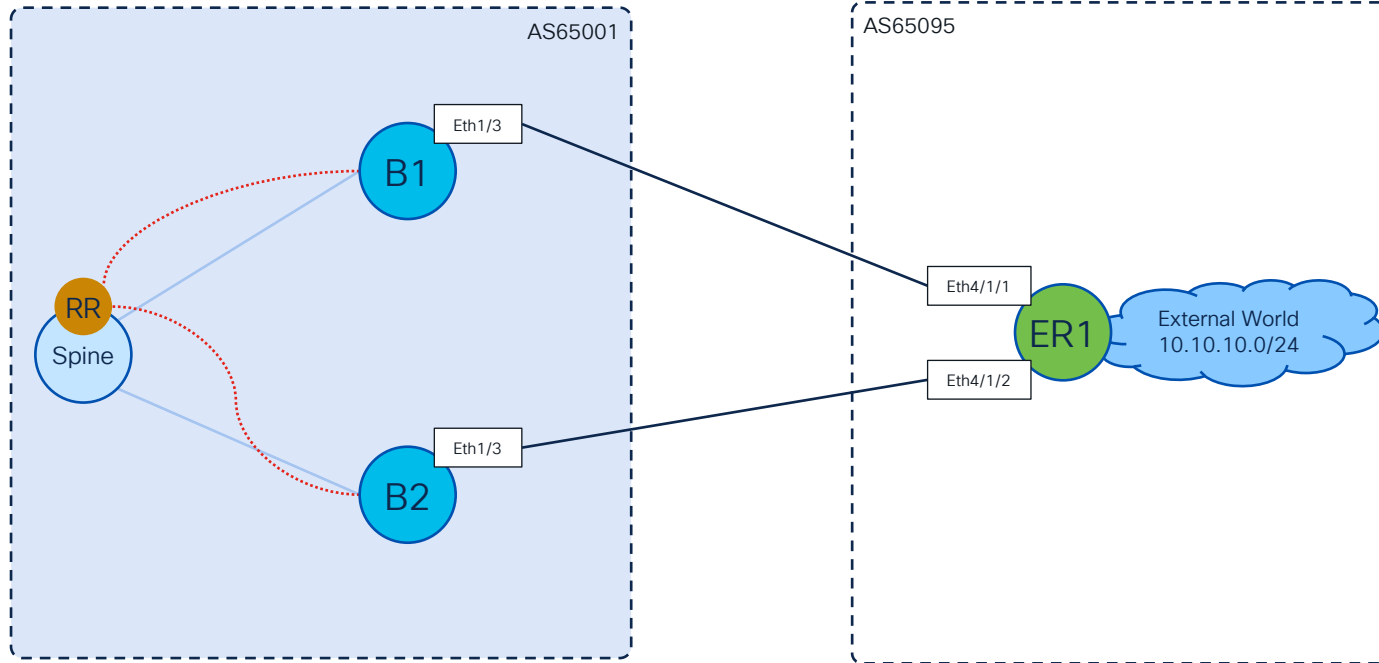


# Scenario #1 eBGP and Sub- Interface



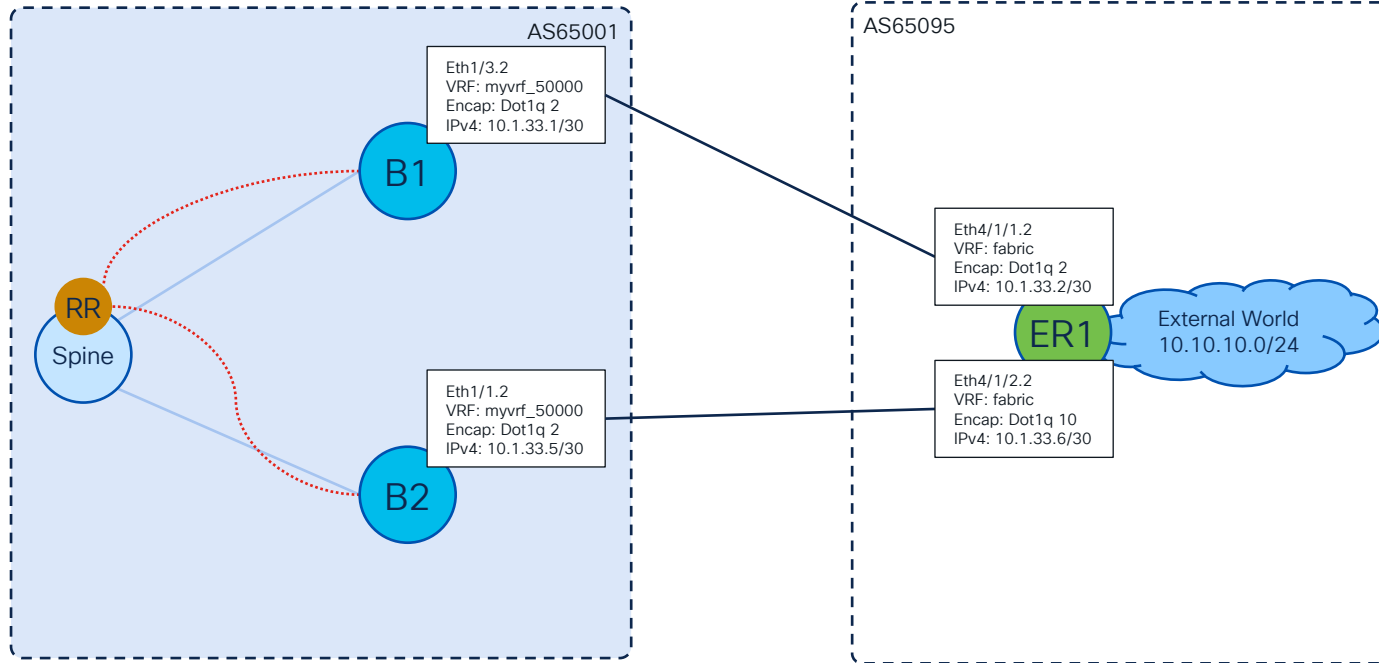
# Scenario #1 – eBGP and Sub-Interface

## External Connectivity



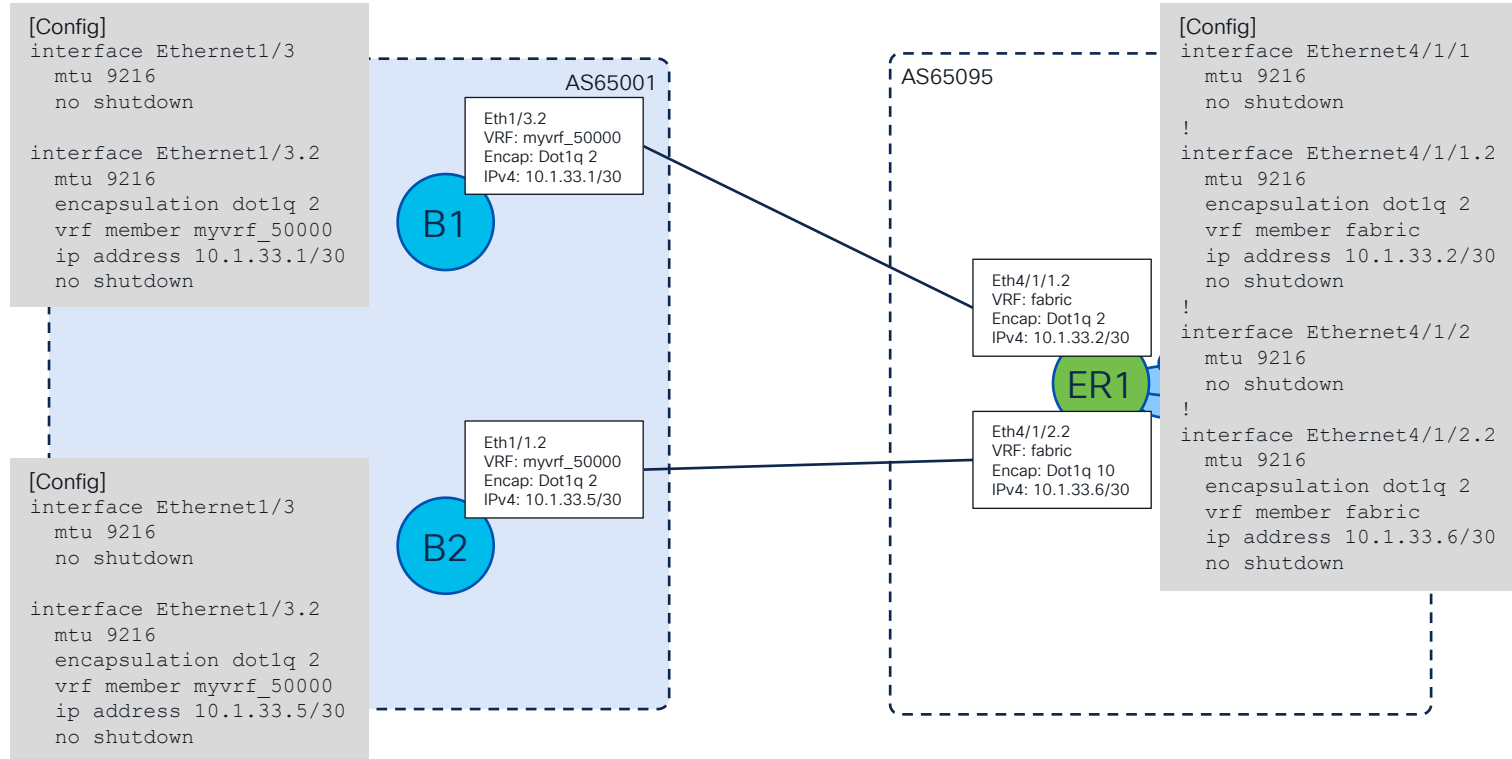
# Scenario #1 – eBGP and Sub-Interface

## External Connectivity



# Scenario #1 – eBGP and Sub-Interface

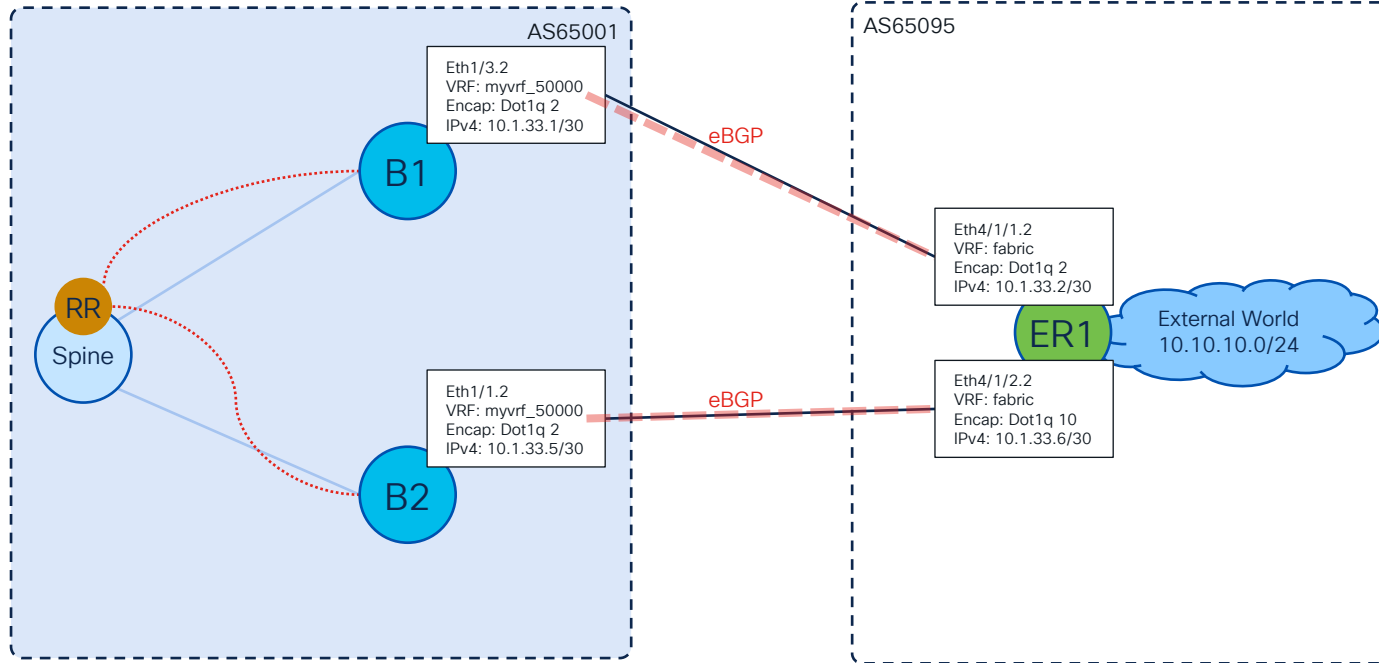
## External Connectivity





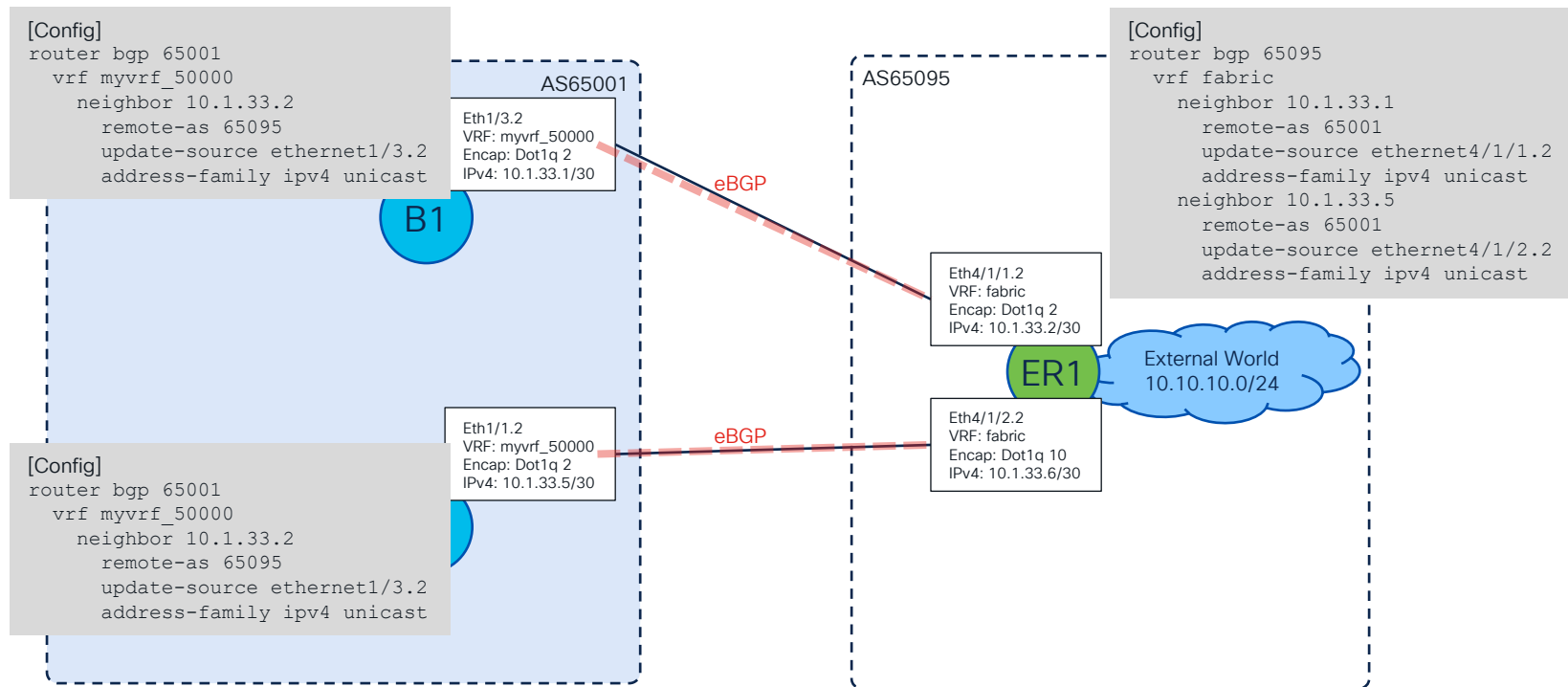
# Scenario #1 – eBGP and Sub-Interface

## External Connectivity



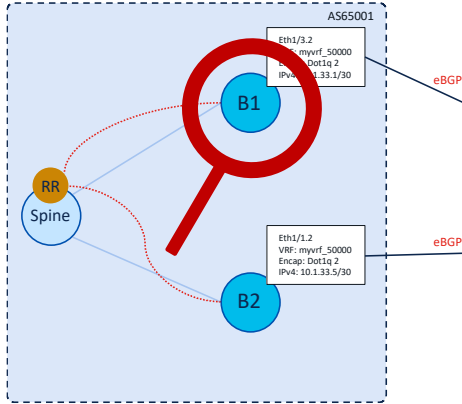
# Scenario #1 – eBGP and Sub-Interface

## External Connectivity



# Scenario #1 – eBGP and Sub-Interface

## External Connectivity



[Output]

```
B1# show ip route vrf myvrf_50000
```

IP Route Table for VRF "myvrf\_50000"

'\*' denotes best ucast next-hop

'\*\*' denotes best mcast next-hop

'[x/y]' denotes [preference/metric]

'%<string>' in via output denotes VRF <string>

```
10.10.10.0/24, ubest/mbest: 1/0
```

```
*via 10.1.33.2, [20/0], 1w1d, bgp-65001, external, tag 65095
```

```
192.168.1.0/24, ubest/mbest: 1/0
```

```
*via 10.1.1.1%default, [200/0], 00:30:42, bgp-65001, internal, tag 65001, segid: 50000 tunnelid: 0xa
```

[Output]

```
B1# show ip bgp vrf myvrf_50000
```

BGP routing table information for VRF myvrf\_50000, address family IPv4 Unicast

BGP table version is 59, Local Router ID is 10.1.33.1

Status: s-suppressed, x-deleted, S-stale, d-dampened, h-history, \*-valid, >-best

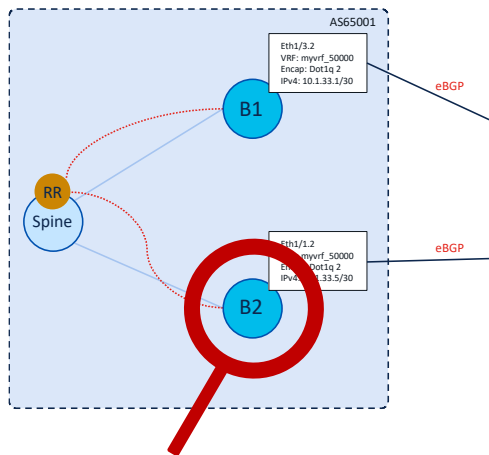
Path type: i-internal, e-external, c-confed, l-local, a-aggregate, r-redist, I-injected

Origin codes: i - IGP, e - EGP, ? - incomplete, | - multipath, & - backup, 2 - best2

Network	Next Hop	Metric	LocPrf	Weight	Path
*>e10.10.10.0/24	10.1.33.2			0	65095 i
* i	10.1.1.2		100	0	65095 i
* i192.168.1.0/24	10.1.1.1	0	100	0	?

# Scenario #1 – eBGP and Sub-Interface

## External Connectivity



[Output]

```
B2# show ip route vrf myvrf_50000
```

IP Route Table for VRF "myvrf\_50000"

'\*' denotes best ucast next-hop

'\*\*' denotes best mcast next-hop

'[x/y]' denotes [preference/metric]

'%<string>' in via output denotes VRF <string>

```
10.10.10.0/24, ubest/mbest: 1/0
```

```
*via 10.1.33.6, [20/0], 1w1d, bgp-65001, external, tag 65095
```

```
192.168.1.0/24, ubest/mbest: 1/0
```

```
*via 10.1.1.1%default, [200/0], 00:30:42, bgp-65001, internal, tag 65001, segid: 50000 tunnelid: 0xa
```

[Output]

```
B2# show ip bgp vrf myvrf_50000
```

BGP routing table information for VRF myvrf\_50000, address family IPv4 Unicast

BGP table version is 59, Local Router ID is 10.1.33.5

Status: s-suppressed, x-deleted, S-stale, d-dampened, h-history, \*-valid, >-best

Path type: i-internal, e-external, c-confed, l-local, a-aggregate, r-redist, I-injected

Origin codes: i - IGP, e - EGP, ? - incomplete, | - multipath, & - backup, 2 - best2

Network	Next Hop	Metric	LocPrf	Weight	Path
*>e10.10.10.0/24	10.1.33.6			0	65095 i
* i	10.1.1.3		100	0	65095 i
* i192.168.1.0/24	10.1.1.1	0	100	0	?

# Scenario #1 – eBGP and Sub-Interface

## External Connectivity

[Output]

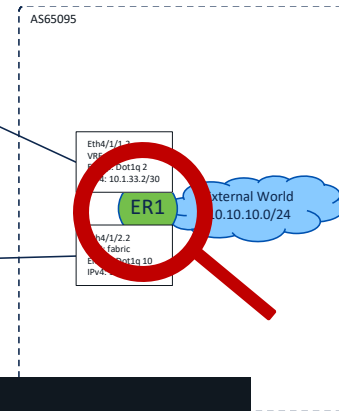
```
ER1# show ip route vrf fabric
IP Route Table for VRF "fabric"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

10.10.10.0/24, ubest/mbest: 1/0, attached
  *via Null10, [254/0], 1wld, static
192.168.1.0/24, ubest/mbest: 1/0 time
  *via 10.1.33.1, [20/0], 6d20h, bgp-65095, external, tag 65001
  *via 10.1.33.5, [20/0], 6d20h, bgp-65095, external, tag 65001
```

[Output]

```
ER1# show ip bgp vrf fabric
BGP routing table information for VRF fabric, address family IPv4 Unicast
BGP table version is 51, Local Router ID is 10.1.33.2
Status: s-suppressed, x-deleted, S-stale, d-dampened, h-history, *-valid, >-best
Path type: i-internal, e-external, c-confed, l-local, a-aggregate, r-redist, I-injected
Origin codes: i - IGP, e - EGP, ? - incomplete, | - multipath, & - backup, 2 - best2
```

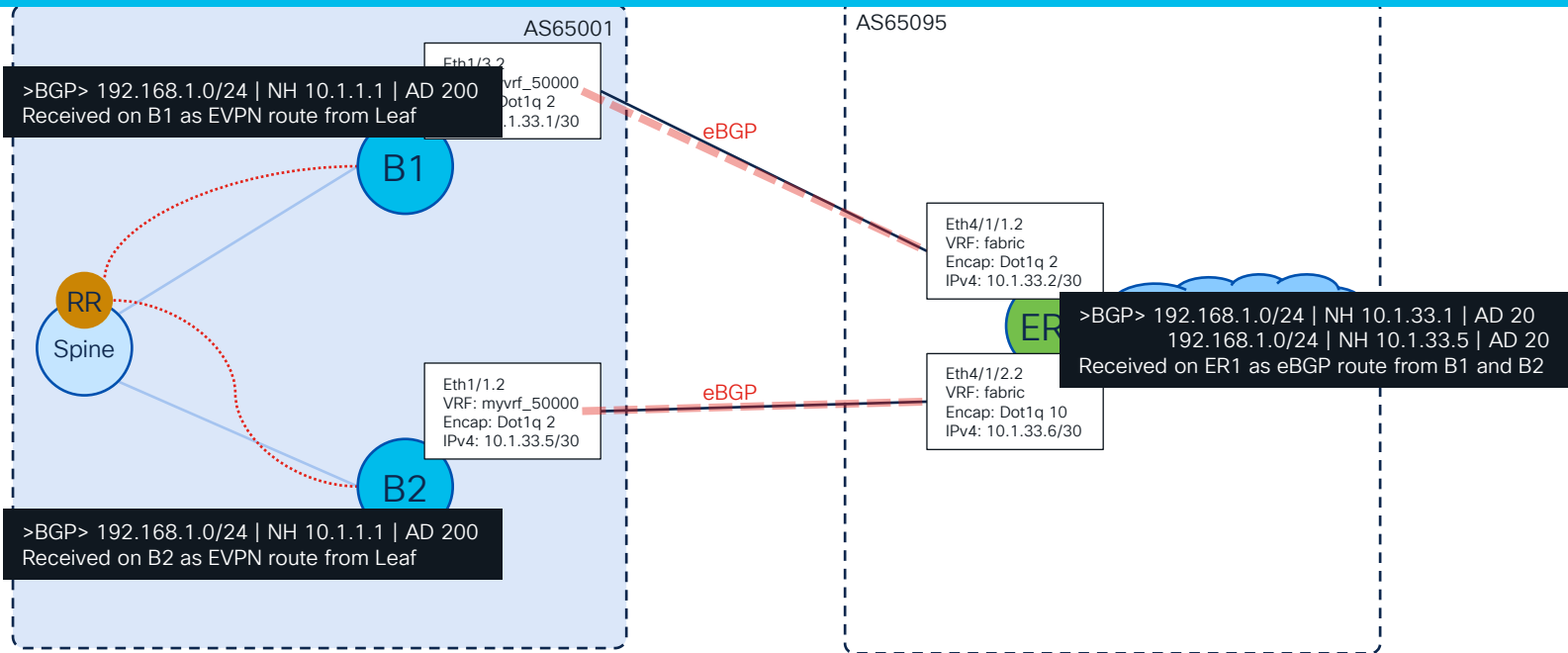
Network	Next Hop	Metric	LocPrf	Weight	Path
*>10.10.10.0/24	0.0.0.0		100	32768	i
* e192.168.1.0/24	10.1.33.1			0	65001 ?
*>e	10.1.33.5			0	65001 ?



# Scenario #1 – eBGP and Sub-Interface

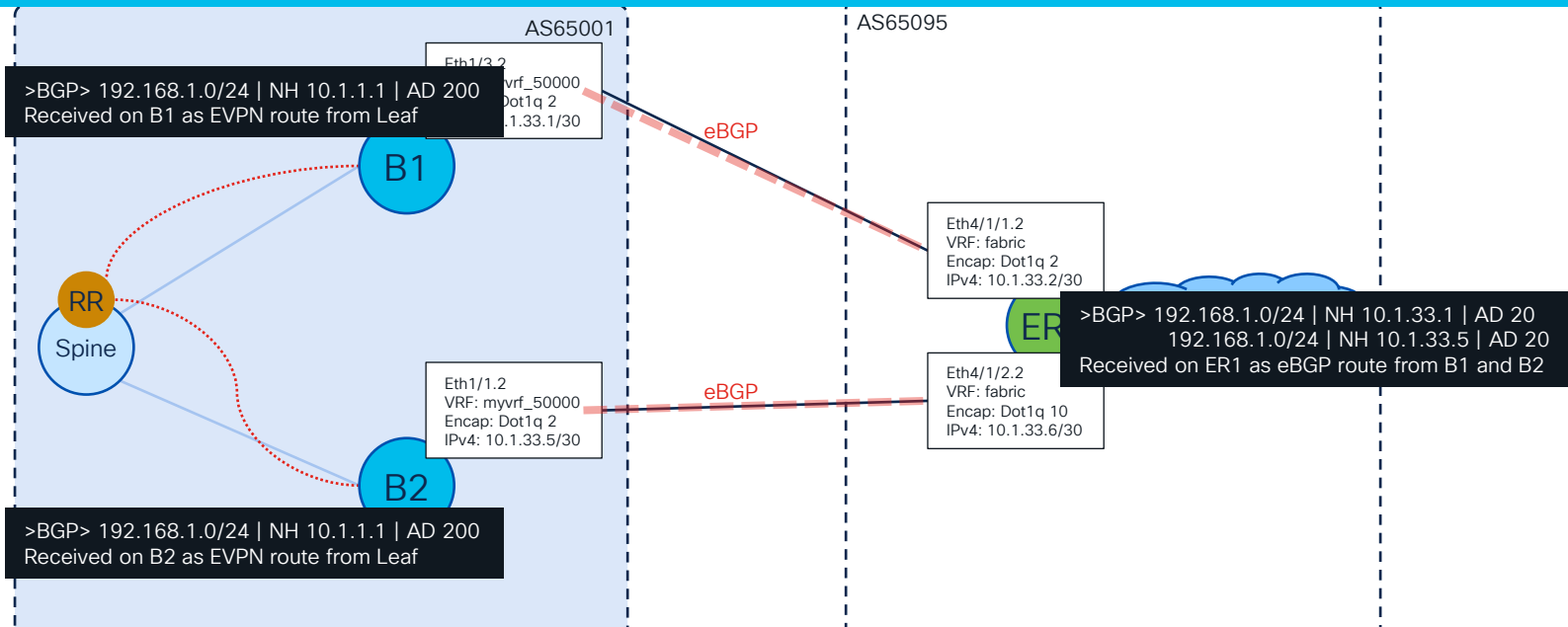
## External Connectivity

Does Everything Look Good?



# Scenario #1 – eBGP and Sub-Interface External Connectivity

# Does Everything Look Good?

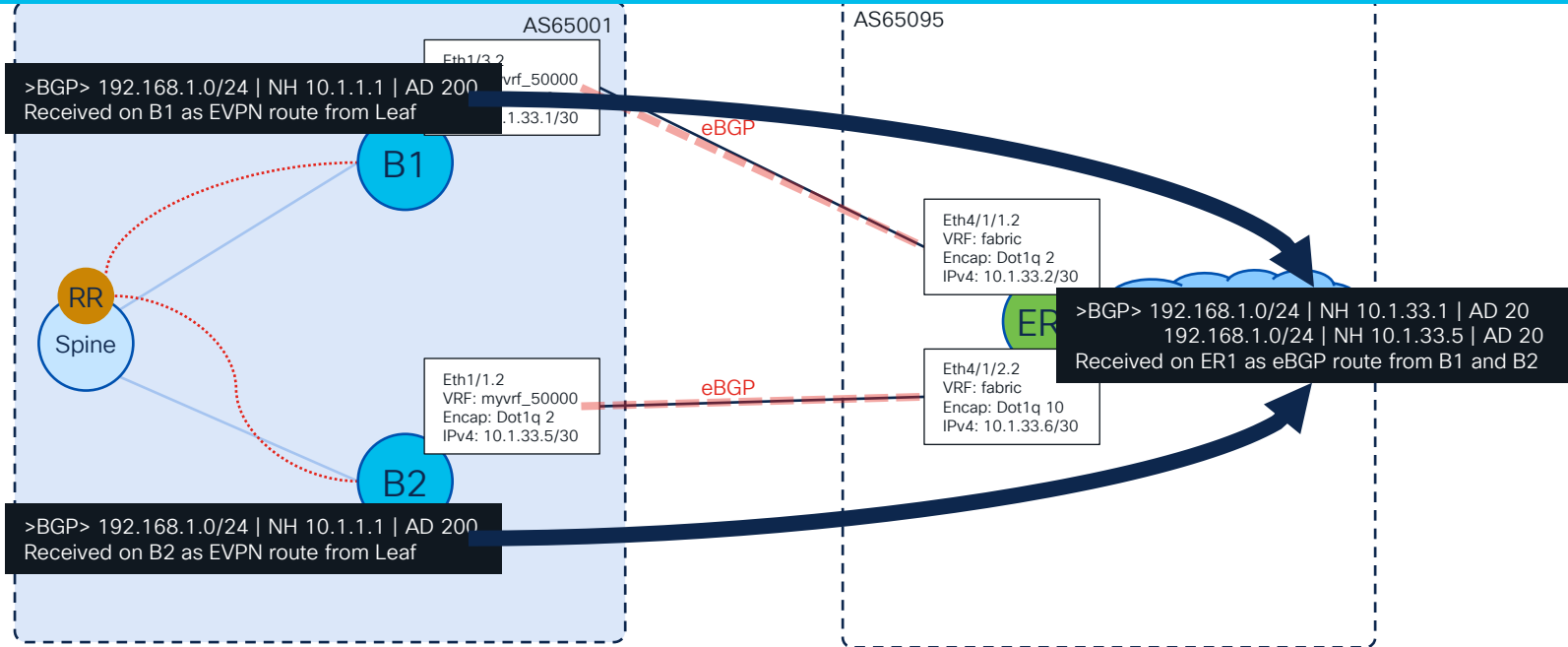


## ECMP into and out of the EVPN Fabric

# Scenario #1 – eBGP and Sub-Interface

## External Connectivity

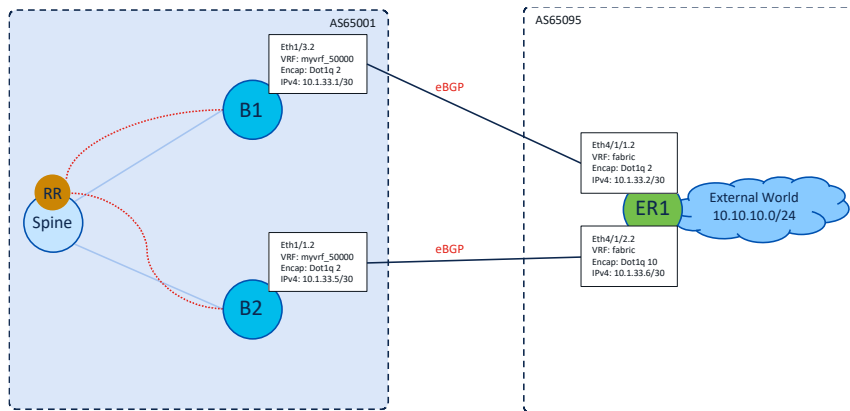
Does Everything Look Good?





# Scenario #1 – eBGP and Sub-Interface

## External Connectivity



- Simple and Straight Forward
  - From EVPN address-family to IPv4/IPv6 address-family
  - Re-Origination at the Border
  - No need for Redistribution
  - Per-VRF Peering for each VRF
  - BGP route filtering from Border to External Router applies
  - Natural protection from learning your own Fabric routes

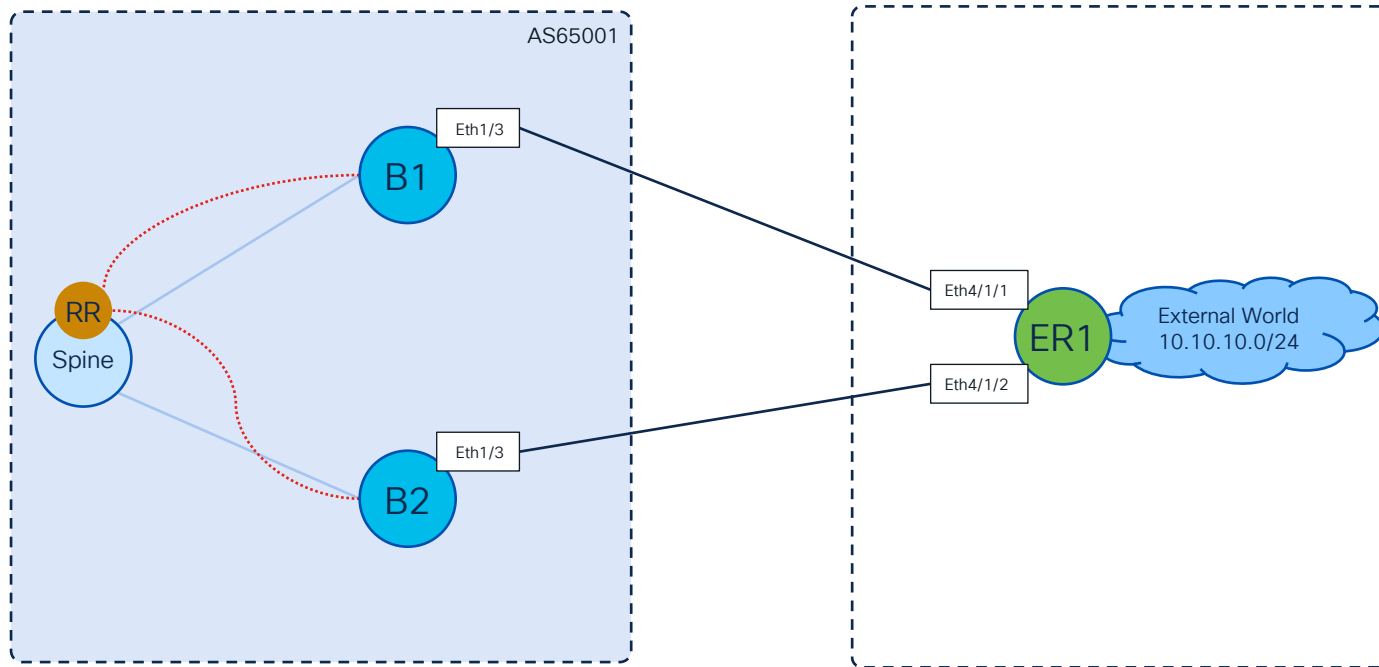
# Scenario #2

## IGP (OSPF) and Sub-Interface



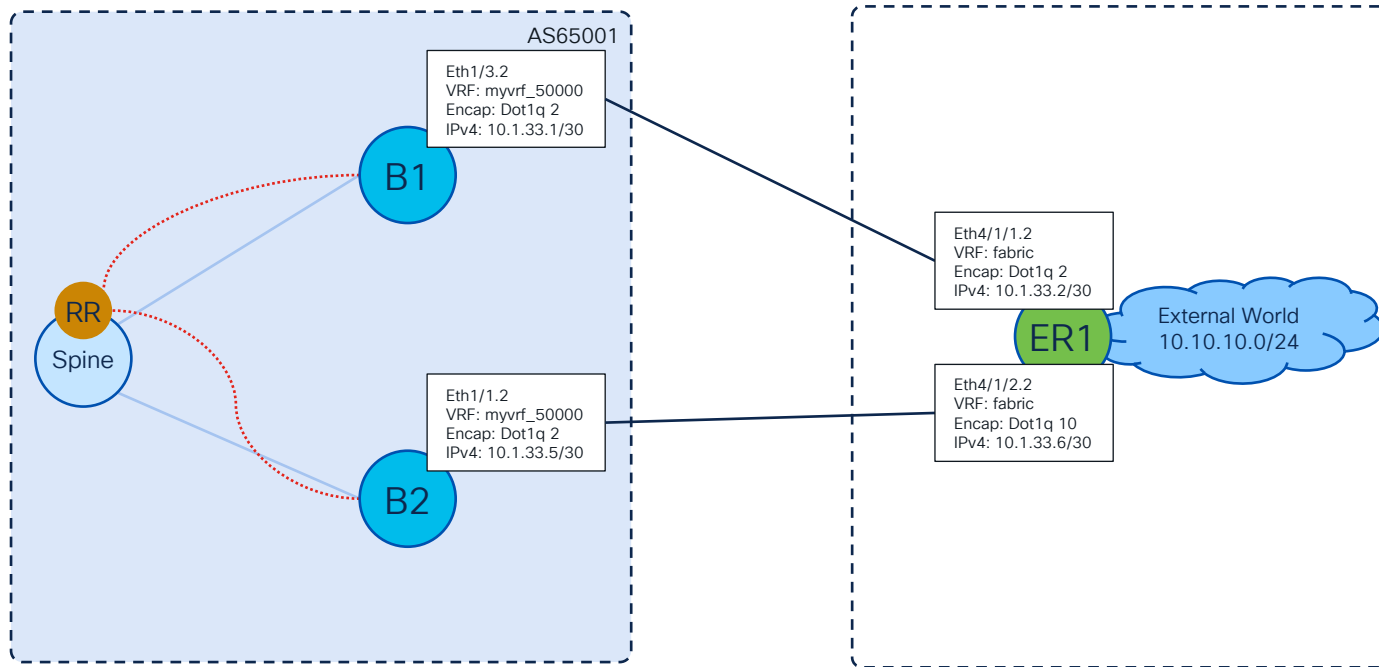
# Scenario #2 – IGP (OSPF) and Sub-Interface

## External Connectivity

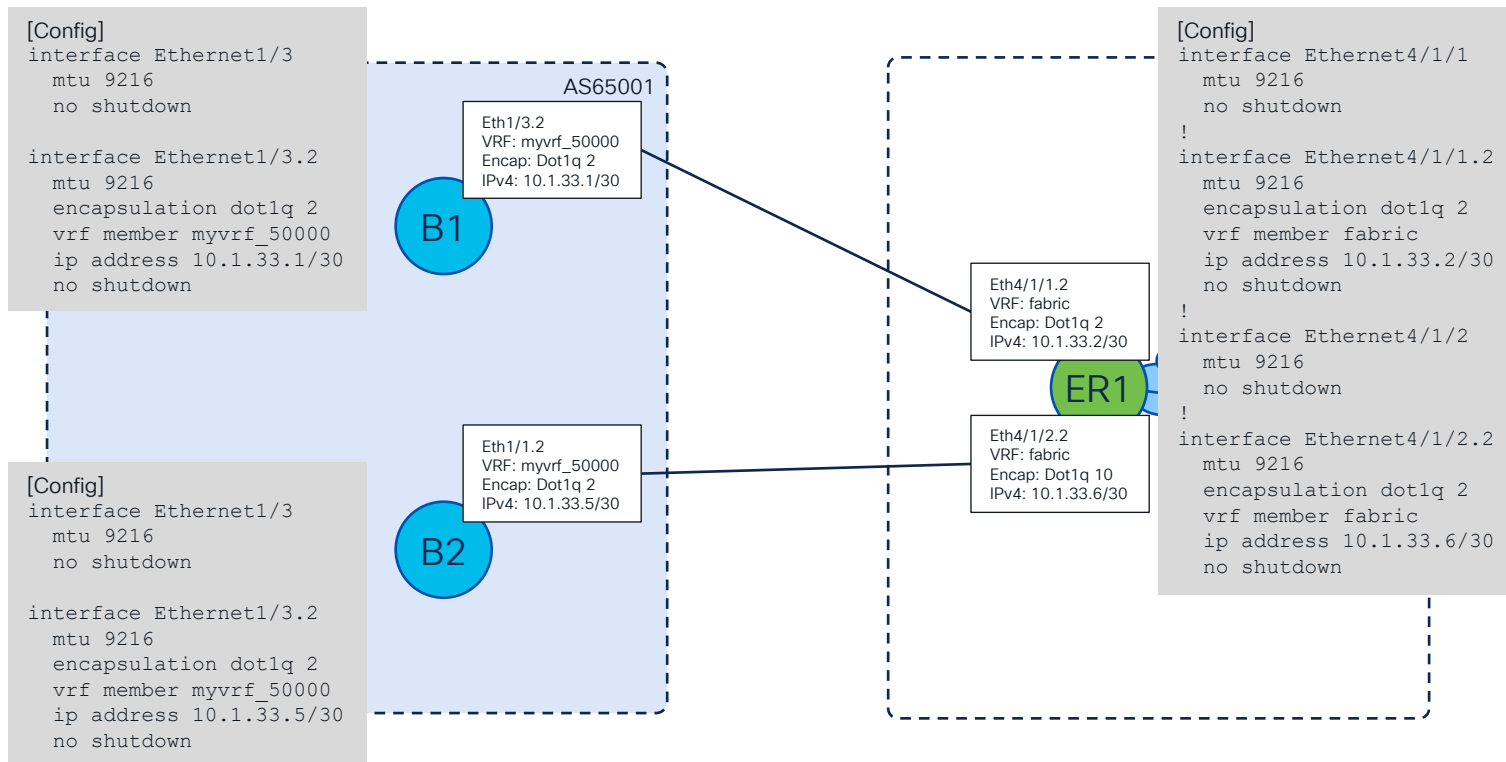


# Scenario #2 – IGP (OSPF) and Sub-Interface

## External Connectivity

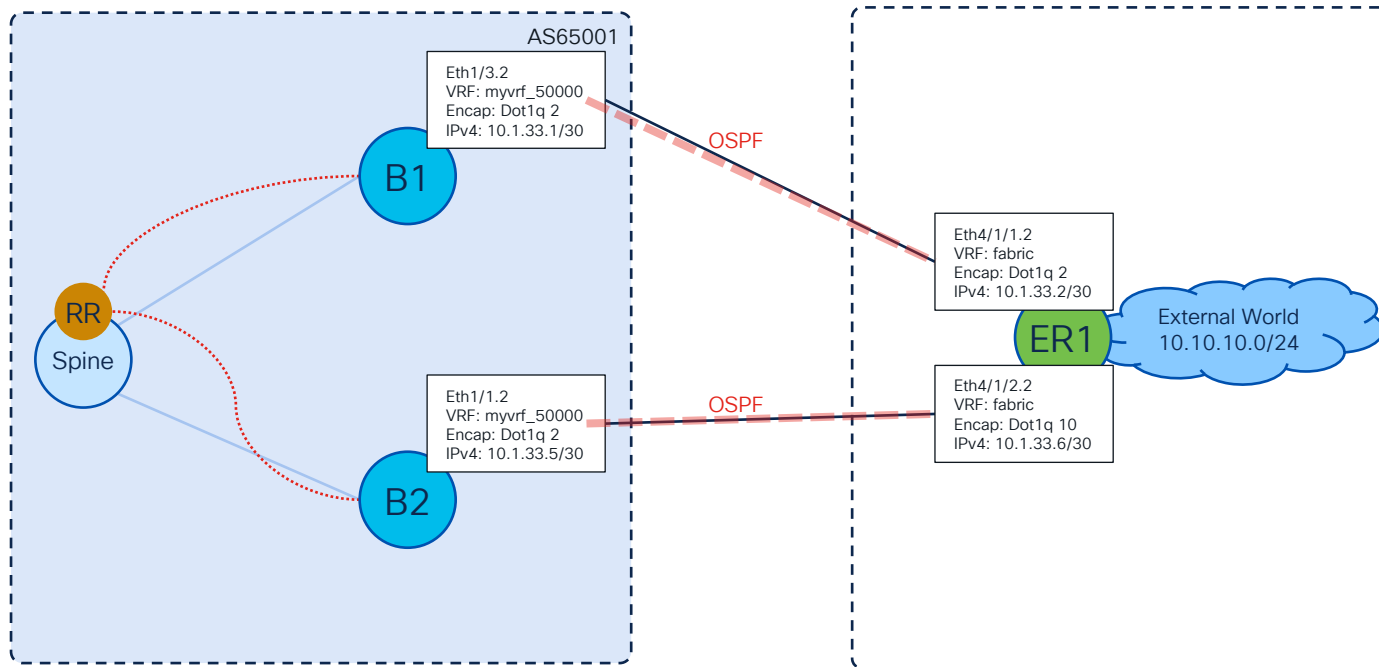


# Scenario #2 – IGP (OSPF) and Sub-Interface External Connectivity

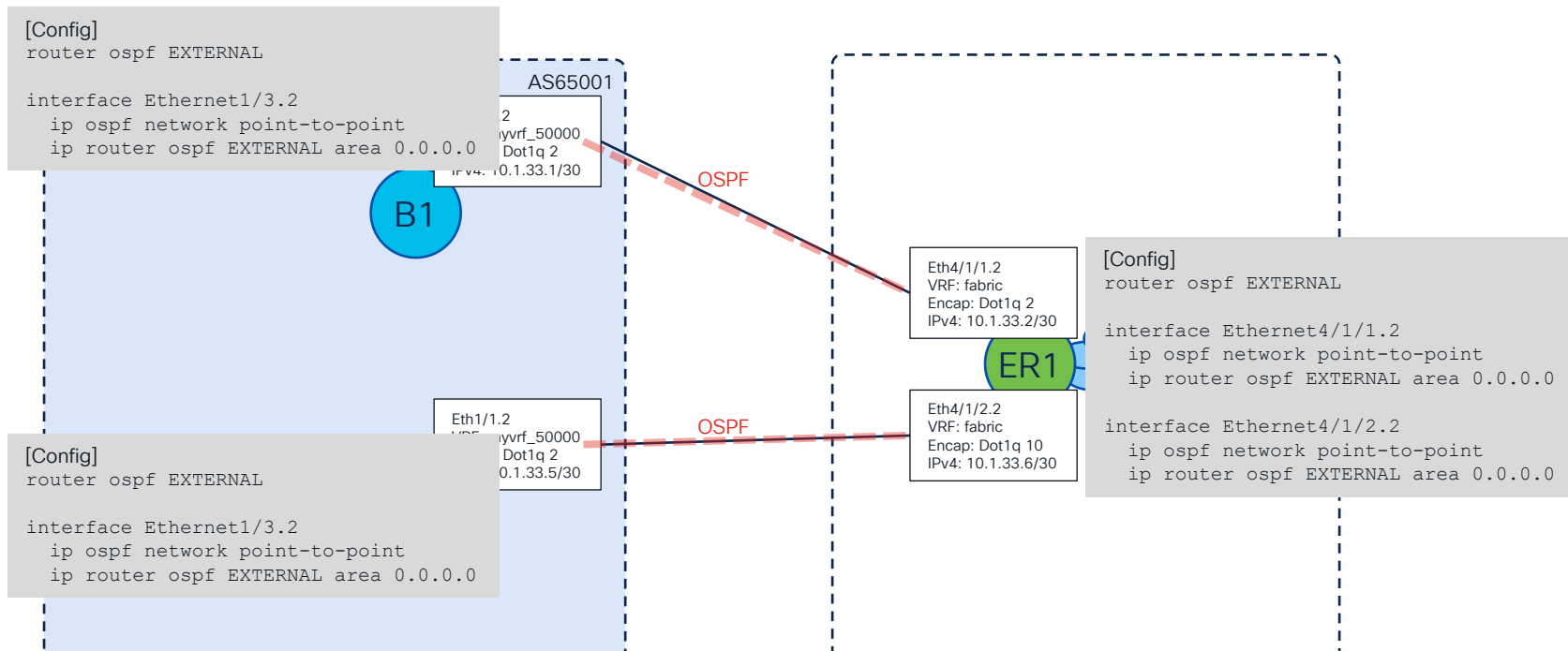


# Scenario #2 – IGP (OSPF) and Sub-Interface

## External Connectivity

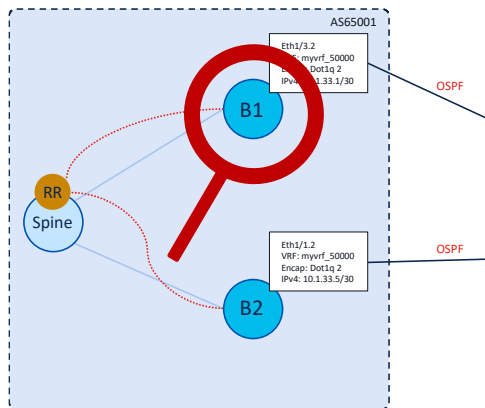


# Scenario #2 – IGP (OSPF) and Sub-Interface External Connectivity



Note: At this stage no BGP to OSPF interaction exists. EVPN routes are not seen in OSPF

# Scenario #2 – IGP (OSPF) and Sub-Interface External Connectivity



[Output]

```
B1# show ip route vrf myvrf_50000
```

```
IP Route Table for VRF "myvrf_50000"
```

```
'*' denotes best ucast next-hop
```

```
'**' denotes best mcast next-hop
```

```
'[x/y]' denotes [preference/metric]
```

```
'%<string>' in via output denotes VRF <string>
```

```
10.1.33.4/30, ubest/mbest: 1/0
```

```
*via 10.1.33.2, Eth1/3.2, [110/8], 00:18:22, ospf-EXTERNAL, intra
```

```
10.10.10.0/24, ubest/mbest: 1/0
```

```
*via 10.1.33.2, Eth1/3.2, [110/20], 00:18:22, ospf-EXTERNAL, type-2
```

```
192.168.1.0/24, ubest/mbest: 1/0
```

```
*via 10.1.1.1%default, [200/0], 00:30:42, bgp-65001, internal, tag 65001, segid: 50000 tunnelid: 0xa0
```

[Output]

```
B1# show ip ospf database vrf myvrf_50000
```

```
OSPF Router with ID (10.1.33.1) (Process ID EXTERNAL VRF myvrf_50000)
```

Router Link States (Area 0.0.0.0)

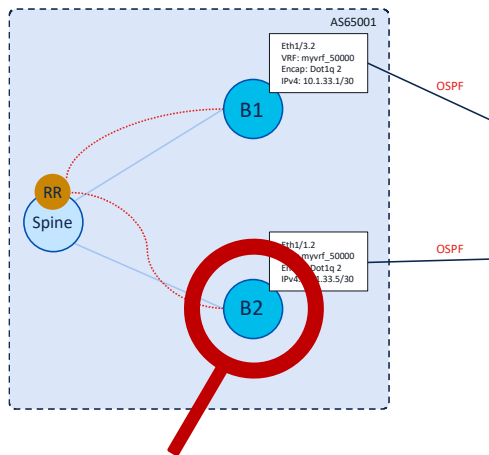
Link ID	ADV Router	Age	Seq#	Checksum	Link Count
10.1.33.1	10.1.33.1	842	0x80000005	0x85d2	2
10.1.33.2	10.1.33.2	1146	0x80000007	0x3e5f	4
10.1.33.5	10.1.33.5	656	0x80000005	0xd572	2

Type-5 AS External Link States

Link ID	ADV Router	Age	Seq#	Checksum	Tag
10.10.10.0	10.1.33.2	1263	0x80000002	0xc20f	0



# Scenario #2 – IGP (OSPF) and Sub-Interface External Connectivity



[Output]

```
B2# show ip route vrf myvrf_50000
```

IP Route Table for VRF "myvrf\_50000"

'\*' denotes best ucast next-hop

'\*\*' denotes best mcast next-hop

'[x/y]' denotes [preference/metric]

'%<string>' in via output denotes VRF <string>

```
10.1.33.0/30, ubest/mbest: 1/0
```

```
*via 10.1.33.2, Eth1/3.2, [110/8], 00:58:41, ospf-EXTERNAL, intra
```

```
10.10.10.0/24, ubest/mbest: 1/0
```

```
*via 10.1.33.2, Eth1/3.2, [110/20], 00:58:41, ospf-EXTERNAL, type-2
```

```
192.168.1.0/24, ubest/mbest: 1/0
```

```
*via 10.1.1.1%default, [200/0], 00:54:38, bgp-65001, internal, tag 65001, segid: 50000 tunnelid: 0xa0
```

[Output]

```
B1# show ip ospf database vrf myvrf_50000
```

OSPF Router with ID (10.1.33.5) (Process ID EXTERNAL VRF myvrf\_50000)

Router Link States (Area 0.0.0.0)

Link ID	ADV Router	Age	Seq#	Checksum	Link Count
10.1.33.1	10.1.33.1	842	0x80000005	0x85d2	2
10.1.33.2	10.1.33.2	1146	0x80000007	0x3e5f	4
10.1.33.5	10.1.33.5	656	0x80000005	0xd572	2

Type-5 AS External Link States

Link ID	ADV Router	Age	Seq#	Checksum	Tag
10.10.10.0	10.1.33.2	1263	0x80000002	0xc20f	0

# Scenario #2 – IGP (OSPF) and Sub-Interface

## External Connectivity

```
[Output]
ER1# show ip route vrf fabric
IP Route Table for VRF "fabric"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

10.10.10.0/24, ubest/mbest: 1/0, attached
    *via Null0, [254/0], 1wld, static
```

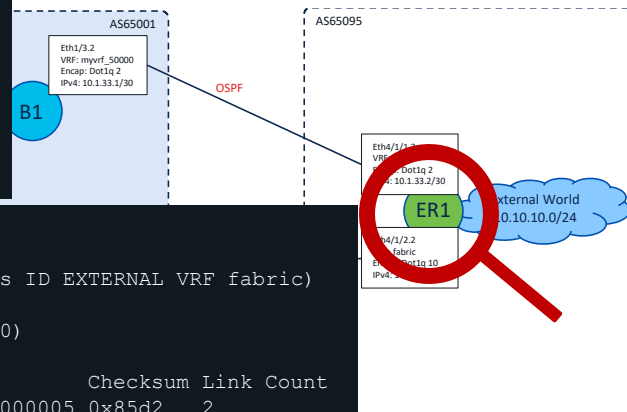
```
[Output]
ER1# show ip ospf database vrf fabric
      OSPF Router with ID (10.1.33.2) (Process ID EXTERNAL VRF fabric)
```

### Router Link States (Area 0.0.0.0)

Link ID	ADV Router	Age	Seq#	Checksum	Link Count
10.1.33.1	10.1.33.1	1159	0x80000005	0x85d2	2
10.1.33.2	10.1.33.2	1461	0x80000007	0x3e5f	4
10.1.33.5	10.1.33.5	971	0x80000005	0xd572	2

### Type-5 AS External Link States

Link ID	ADV Router	Age	Seq#	Checksum	Tag
10.10.10.0	10.1.33.2	1579	0x80000002	0xc20f	0



# Scenario #2 – IGP (OSPF) and Sub-Interface

## External Connectivity

```
[Config]
router bgp 65001
  vrf myvrf_50000
    address-family ipv4 unicast
      redistribute ospf EXTERNAL route-map ALL

router ospf EXTERNAL
  vrf myvrf_50000
    redistribute bgp 65001 route-map OSPF-internal

interface Ethernet1/3.2
  ip ospf network point-to-point
  ip router ospf EXTERNAL area 0.0.0.0
```

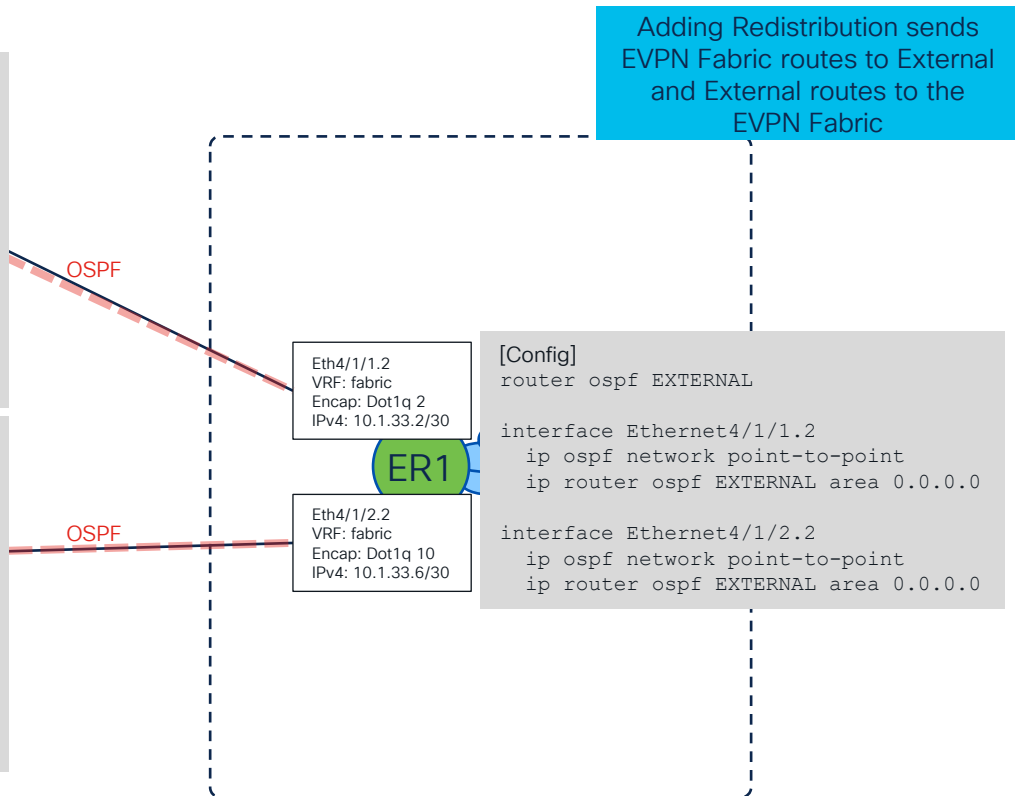
```
[Config]
router bgp 65001
  vrf myvrf_50000
    address-family ipv4 unicast
      redistribute ospf EXTERNAL route-map ALL

router ospf EXTERNAL
  vrf myvrf_50000
    redistribute bgp 65001 route-map OSPF-internal

interface Ethernet1/3.2
  ip ospf network point-to-point
  ip router ospf EXTERNAL area 0.0.0.0
```

```
[Config]
route-map OSPF-internal permit 10
  match route-type internal

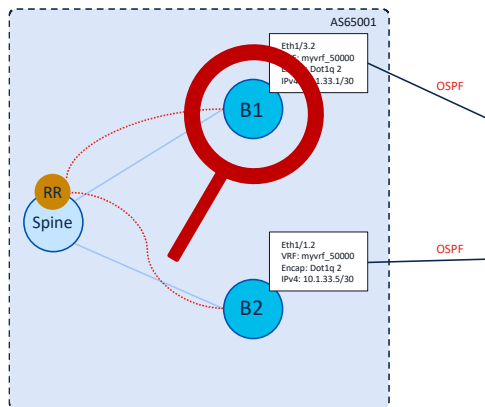
route-map ALL permit 10
```



# Scenario #2 – IGP (OSPF) and Sub-Interface

## External Connectivity

Adding Redistribution sends  
EVPN Fabric routes to External  
and External routes to the  
EVPN Fabric



[Output]

B1# show ip route vrf myvrf 50000

IP Route Table for VRF "myvrf\_50000"

'\*' denotes best ucast next-hop

\*\*\* denotes best mcast next-hop

'[x/y]' denotes [preference/metric]

'%<string>' in via output denotes VRF <string>

10.1.33.4/30, ubest/mbest: 1/0

\*via 10.1.33.2, Eth1/3.2, [110/8], 00:58:41, ospf-EXTERNAL, intra

10.10.10.0/24, ubest/mbest: 1/0

\*via 10.1.33.2, Eth1/3.2, [110/20], 00:58:41, ospf-EXTERNAL, type-2

192.168.1.0/24, ubest/mbest: 1/0

\*via 10.1.1.1%default, [200/0], 00:54:38, bgp-65001, internal, tag 65001, segid: 50000 tunnelid: 0xa0

[Output]

B1# show ip ospf database vrf myvrf\_50000

OSPF Router with ID (10.1.33.1) (Process ID EXTERNAL VRF myvrf\_50000)

Router Link States (Area 0.0.0.0)

Link ID	ADV Router	Age	Seq#	Checksum	Link Count
10.1.33.1	10.1.33.1	1446	0x80000006	0x83d3	2
10.1.33.2	10.1.33.2	1754	0x80000008	0x3c60	4
10.1.33.5	10.1.33.5	1264	0x80000006	0xd373	2

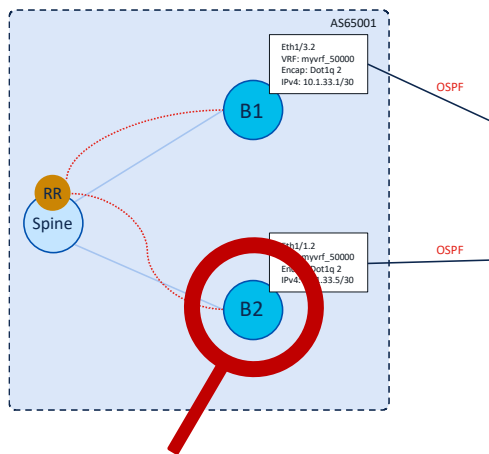
Type-5 AS External Link States

Link ID	ADV Router	Age	Seq#	Checksum	Tag
10.10.10.0	10.1.33.2	42	0x80000004	0xbe11	0
192.168.1.0	10.1.33.1	1356	0x80000003	0x6d74	65001

# Scenario #2 – IGP (OSPF) and Sub-Interface

## External Connectivity

Adding Redistribution sends  
EVPN Fabric routes to External  
and External routes to the  
EVPN Fabric



[Output]

B2# show ip route vrf myvrf 50000

IP Route Table for VRF "myvrf\_50000"

'\*' denotes best ucast next-hop

'\*\*' denotes best mcast next-hop

'[x/y]' denotes [preference/metric]

'%<string>' in via output denotes VRF <string>

10.1.33.0/30, ubest/mbest: 1/0

\*via 10.1.33.6, Eth1/3.2, [110/8], 01:01:12, ospf-EXTERNAL, intra

10.10.10.0/24, ubest/mbest: 1/0

\*via 10.1.33.6, Eth1/3.2, [110/20], 01:01:12, ospf-EXTERNAL, type-2

192.168.1.0/24, ubest/mbest: 1/0

\*via 10.1.33.6, Eth1/3.2, [110/1], 00:54:40, ospf-EXTERNAL, type-2, tag 65001

[Output]

B2# show ip ospf database vrf myvrf\_50000

OSPF Router with ID (10.1.33.5) (Process ID EXTERNAL VRF myvrf\_50000)

Router Link States (Area 0.0.0.0)

Link ID	ADV Router	Age	Seq#	Checksum	Link Count
10.1.33.1	10.1.33.1	1623	0x80000006	0x83d3	2
10.1.33.2	10.1.33.2	107	0x80000009	0x3a61	4
10.1.33.5	10.1.33.5	1437	0x80000006	0xd373	2

Type-5 AS External Link States

Link ID	ADV Router	Age	Seq#	Checksum	Tag
10.10.10.0	10.1.33.2	217	0x80000004	0xbe11	0
192.168.1.0	10.1.33.1	1533	0x80000003	0x6d74	65001

# Scenario #2 – IGP (OSPF) and Sub-Interface

## External Connectivity

```
[Output]
ER1# sh ip route vrf fabric
IP Route Table for VRF "fabric"
'*' denotes best ucast next-hop
 '**' denotes best mcast next-hop
 '[x/y]' denotes [preference/metric]
 '%<string>' in via output denotes VRF <string>

10.10.10.0/24, ubest/mbest: 1/0 time
    *via Null0, [254/0], 1w1d, static
192.168.1.0/24, ubest/mbest: 1/0 time
    *via 10.1.33.1, Eth4/1/1.2, [110/1], 00:56:37, ospf-EXTERNAL, type-2, tag 65001
```

```
[Output]
ER1# show ip ospf database vrf fabric
      OSPF Router with ID (10.1.33.2) (Process ID EXTERNAL VRF fabric)

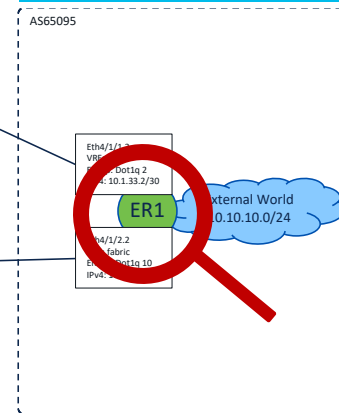
      Router Link States (Area 0.0.0.0)

Link ID      ADV Router   Age         Seq#         Checksum Link Count
10.1.33.1    10.1.33.1    1743        0x80000006  0x83d3    2
10.1.33.2    10.1.33.2    227         0x80000009  0x3a61    4
10.1.33.5    10.1.33.5    1560        0x80000006  0xd373    2

      Type-5 AS External Link States

Link ID      ADV Router   Age         Seq#         Checksum Tag
10.10.10.0   10.1.33.2    338         0x80000004  0xbe11    0
192.168.111.0 10.1.33.1    1653        0x80000003  0x6d74    65001
```

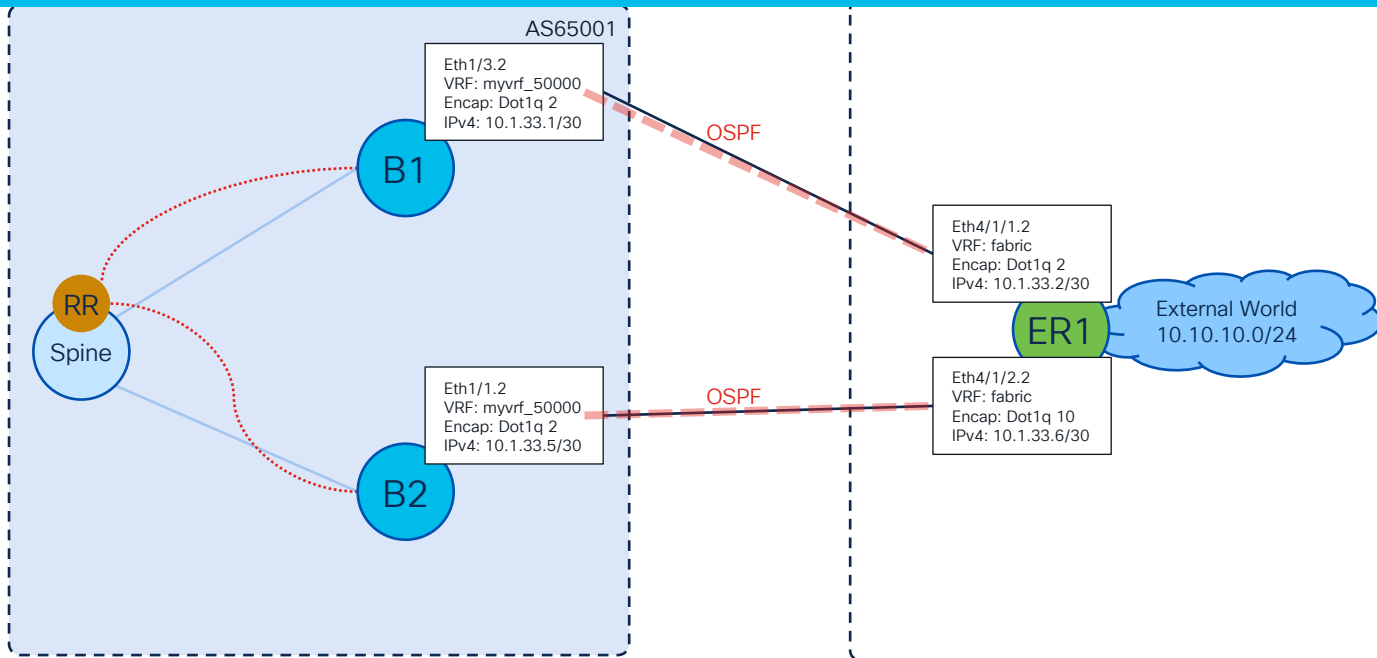
Adding Redistribution sends  
EVPN Fabric routes to External  
and External routes to the  
EVPN Fabric



# Scenario #2 – IGP (OSPF) and Sub-Interface

## External Connectivity

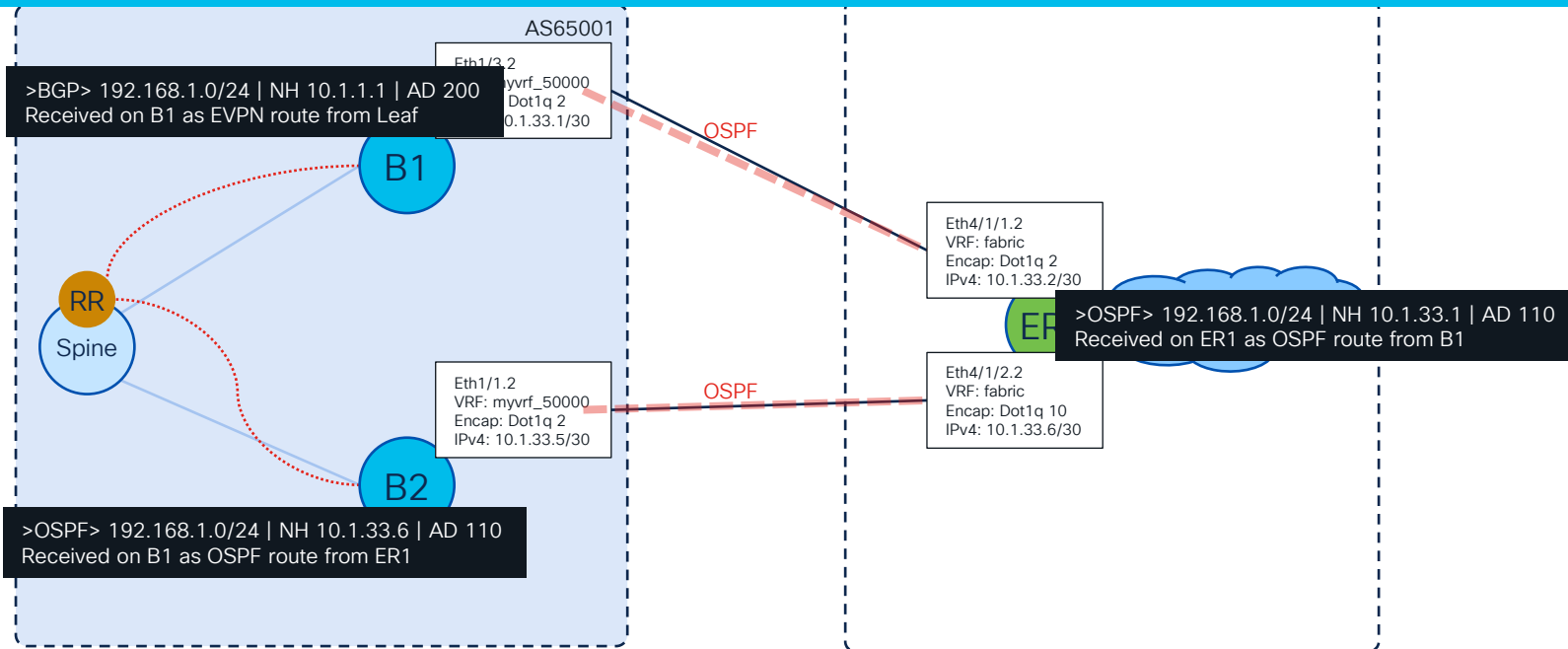
Does Everything Look Good?



# Scenario #2 – IGP (OSPF) and Sub-Interface

## External Connectivity

Does Everything Look Good?

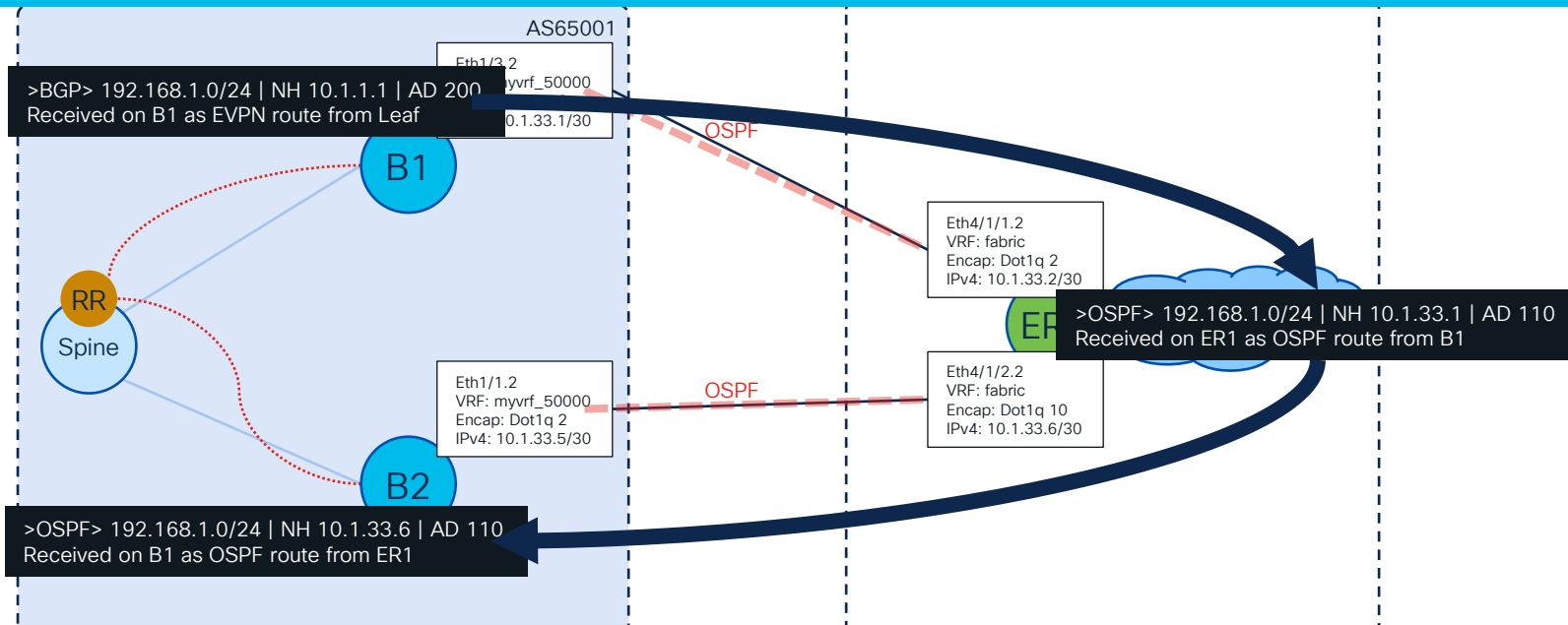




# Scenario #2 – IGP (OSPF) and Sub-Interface

## External Connectivity

Does Everything Look Good?

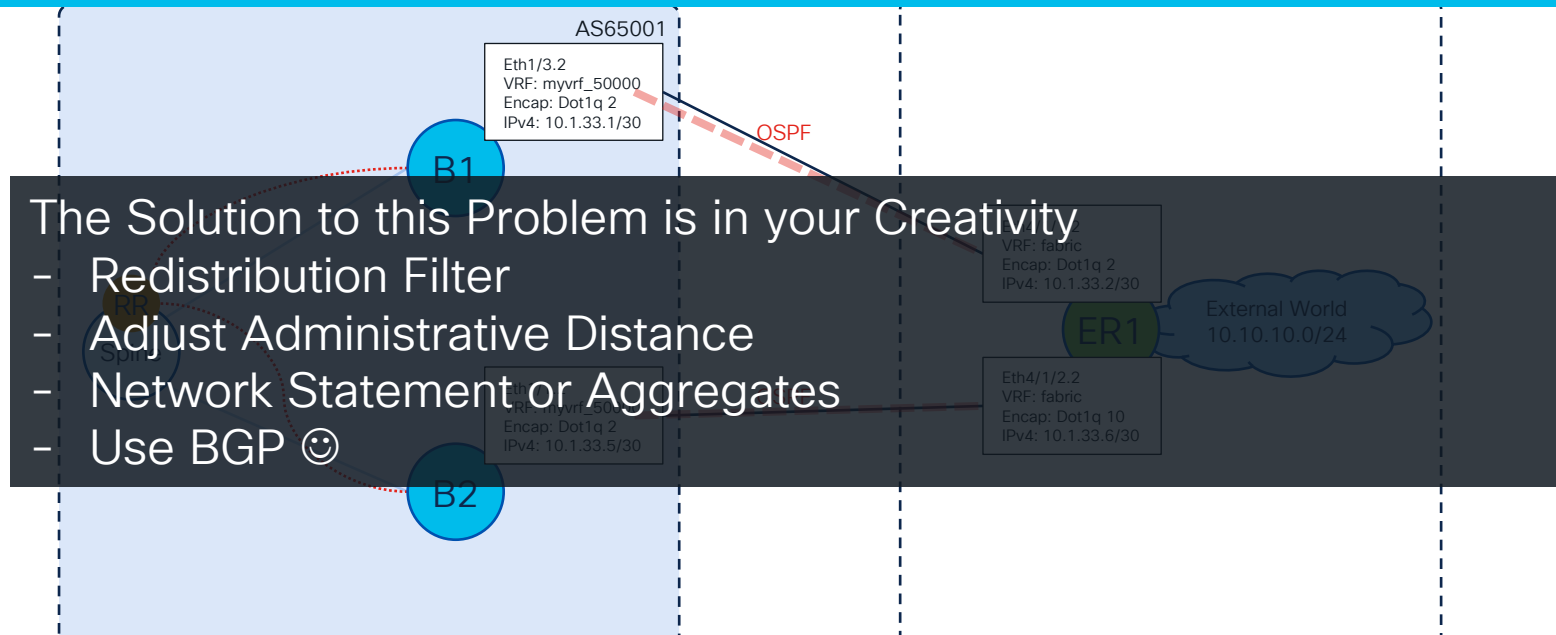


No ECMP into the EVPN Fabric

# Scenario #2 – IGP (OSPF) and Sub-Interface

## External Connectivity

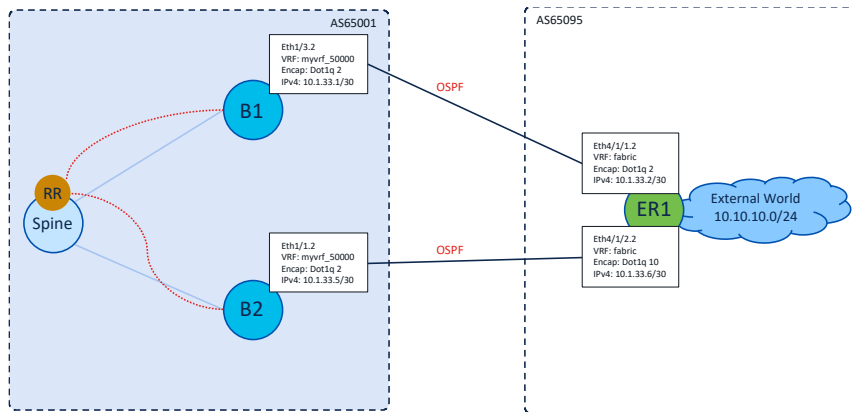
Does Everything Look Good?



No ECMP into the EVPN Fabric

# Scenario #2 – IGP (OSPF) and Sub-Interface

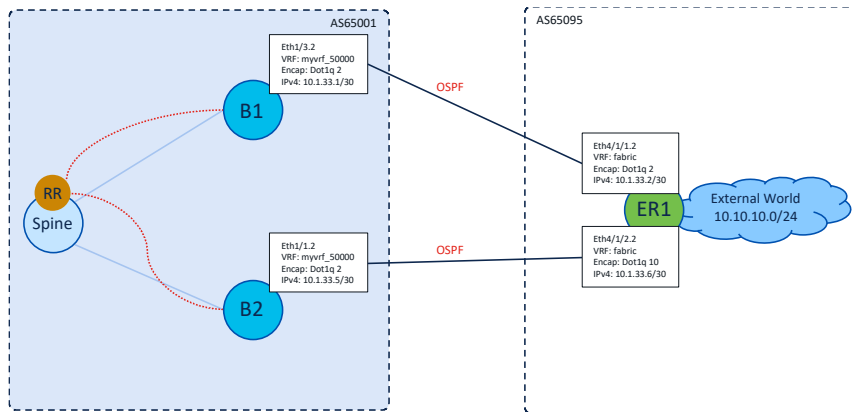
## External Connectivity



- Solvable but Not Simple or Straight Forward
  - From BGP EVPN to OSPF
  - Redistribution or Network/Aggregate statement needed
  - iBGP (internal) doesn't redistribute automatically to OSPF
  - OSPF routes (External) are preferred over EVPN routes (internal)
    - OSPF Administrative Distance 110, iBGP Administrative Distance 200
  - Filtering and Aggregation in IGP's (intra Area is a nightmare)
  - No Natural protection from learning your own Fabric routes
    - Filter configuration needed special with mutual redistribution

# Scenario #2 – IGP (OSPF) and Sub-Interface

## External Connectivity



- Solvable but Not Simple or Straight Forward
  - From BGP EVPN to OSPF
  - Redistribution or Network/Aggregate statement needed
  - iBGP (internal) doesn't redistribute automatically to OSPF
  - OSPF routes (External) are preferred over EVPN routes (internal)
    - OSPF Administrative Distance 110, iBGP Administrative Distance 200
  - Filtering and Aggregation in IGPs (intra Area is a nightmare)
  - No Natural protection from learning your own Fabric routes
    - Filter configuration needed special with mutual redistribution

Remember: Doing BGP to the ER doesn't mean you can't use IGPs beyond the ER

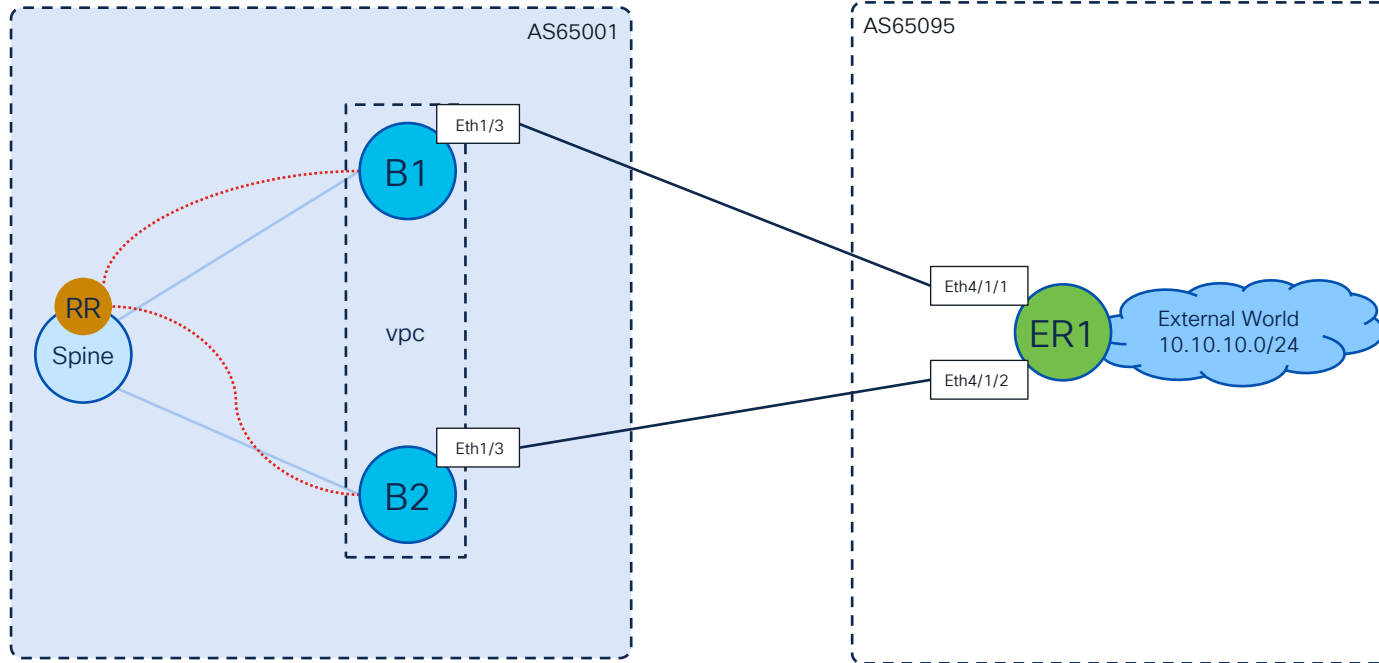
# Scenario #3

## eBGP, vPC and SVIs



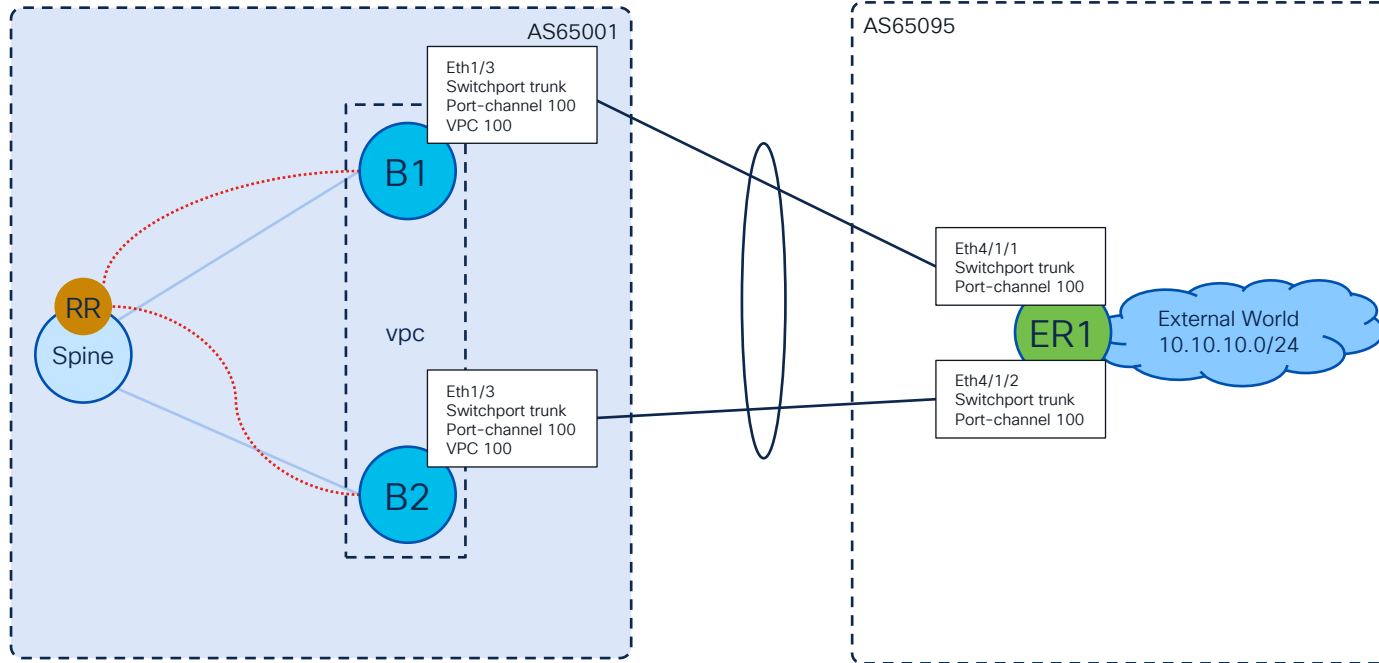
# Scenario #3 – eBGP with SVI and vPC

## External Connectivity



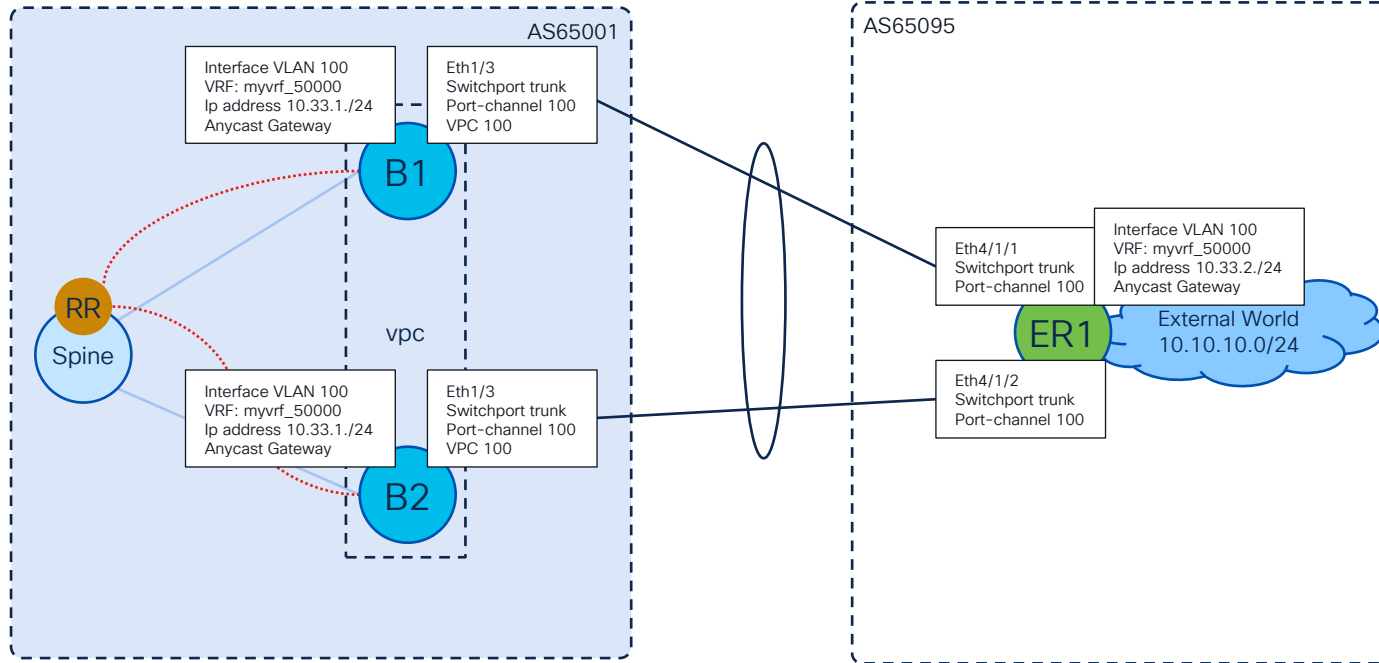
# Scenario #3 – eBGP with SVI and vPC

## External Connectivity



# Scenario #3 – eBGP with SVI and vPC

## External Connectivity





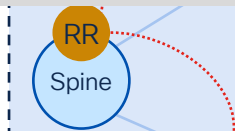
# Scenario #3 – eBGP with SVI and vPC

## External Connectivity

[Config]

```
interface Ethernet1/3
  switchport
  switchport mode trunk
  channel-group 100 mode active
  no shutdown

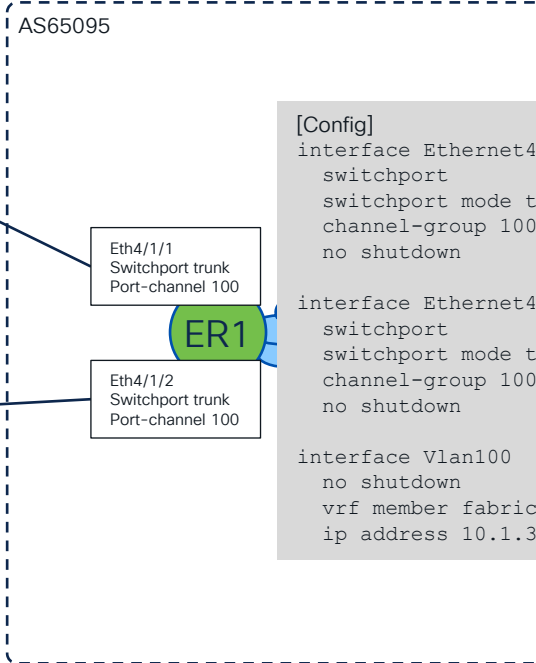
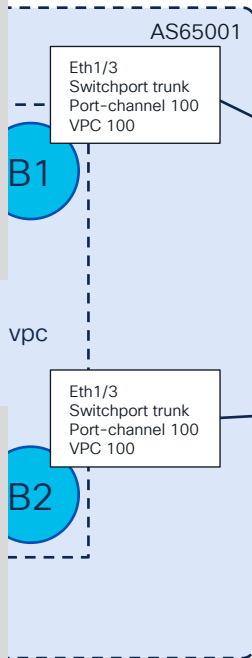
interface Vlan100
  no shutdown
  vrf member myvrf_50000
  ip address 10.1.33.1/24 tag 12345
  fabric forwarding mode anycast-gateway
```



[Config]

```
interface Ethernet1/3
  switchport
  switchport mode trunk
  channel-group 100 mode active
  no shutdown

interface Vlan100
  no shutdown
  vrf member myvrf_50000
  ip address 10.1.33.1/24 tag 12345
  fabric forwarding mode anycast-gateway
```



[Config]

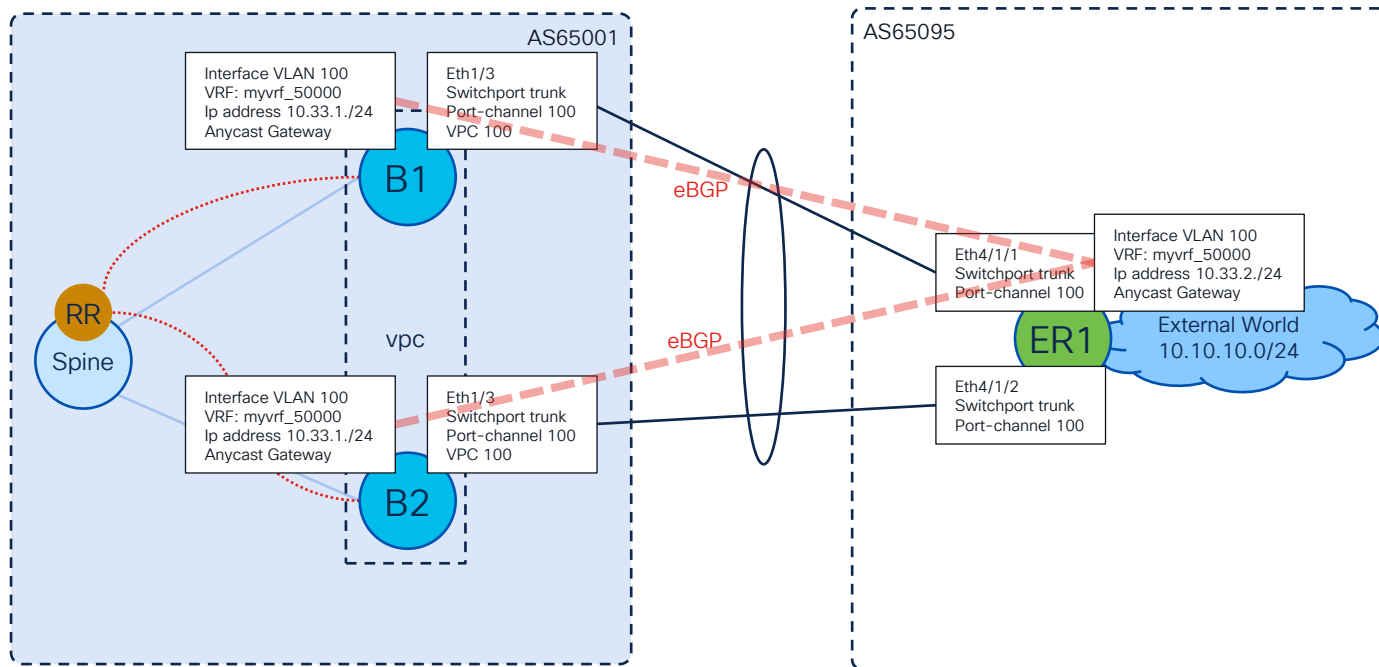
```
interface Ethernet4/1/1
  switchport
  switchport mode trunk
  channel-group 100 mode active
  no shutdown

interface Ethernet4/1/2
  switchport
  switchport mode trunk
  channel-group 100 mode active
  no shutdown

interface Vlan100
  no shutdown
  vrf member fabric
  ip address 10.1.33.2/24 tag 12345
```

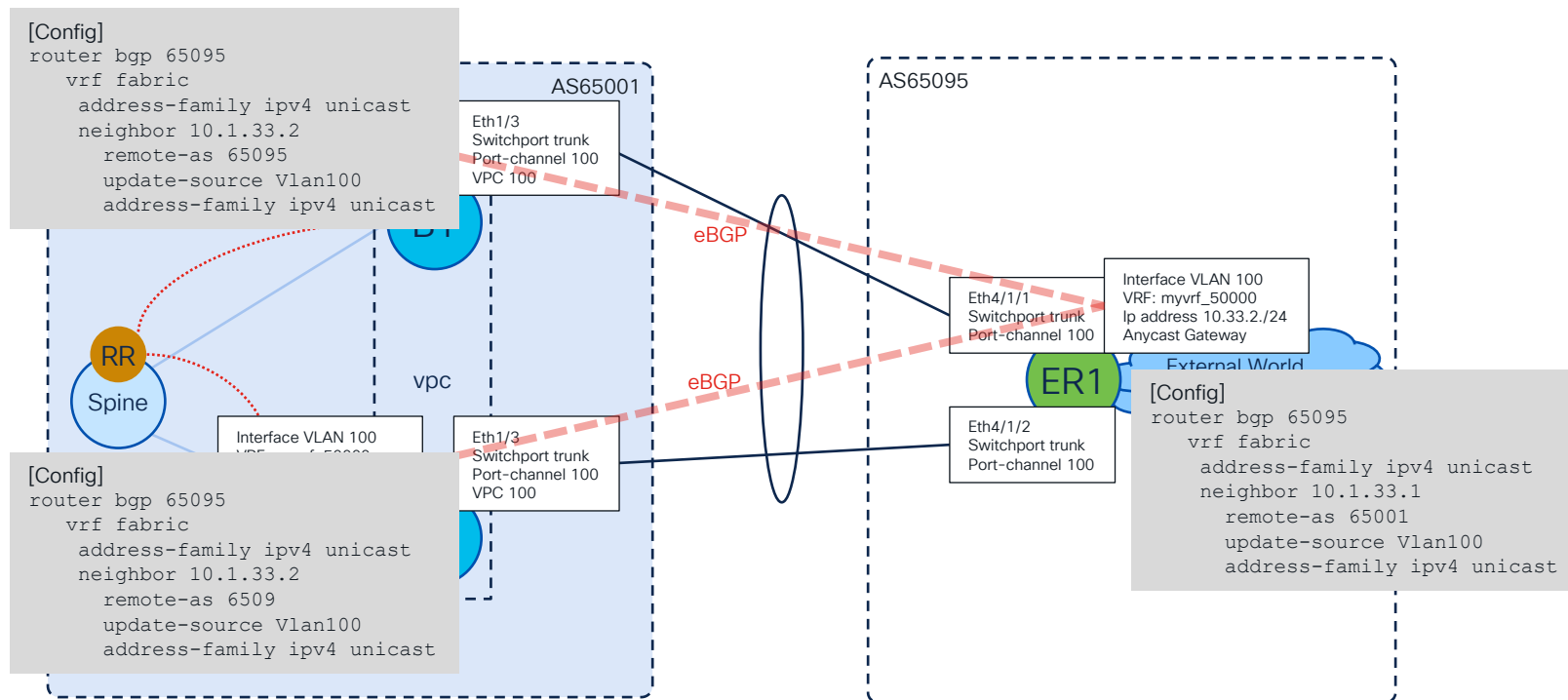
# Scenario #3 – eBGP with SVI and vPC

## External Connectivity



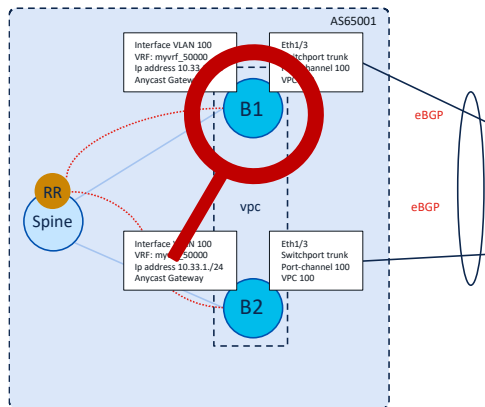
# Scenario #3 – eBGP with SVI and vPC

## External Connectivity



# Scenario #3 – eBGP with SVI and vPC

## External Connectivity



[Output]

```
B1# show ip route vrf myvrf_50000
```

IP Route Table for VRF "myvrf\_50000"

'\*' denotes best unicast next-hop

'\*\*' denotes best mcast next-hop

'[x/y]' denotes [preference/metric]

'%<string>' in via output denotes VRF <string>

```
192.168.1.0/24, ubest/mbest: 1/0
```

```
*via 10.1.1.1%default, [200/0], 00:28:18, bgp-65001, internal, tag 65001, segid: 50000 tunnelid: 0xa
```

[Output]

```
B1# show ip bgp vrf myvrf_50000
```

BGP routing table information for VRF myvrf\_50000, address family IPv4 Unicast

BGP table version is 30, Local Router ID is 10.1.33.1

Status: s-suppressed, x-deleted, S-stale, d-dampened, h-history, \*-valid, >-best

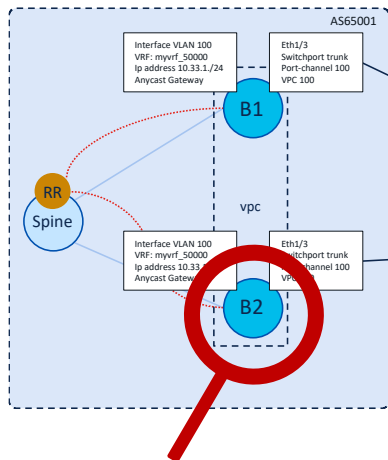
Path type: i-internal, e-external, c-confed, l-local, a-aggregate, r-redist, I-injected

Origin codes: i - IGP, e - EGP, ? - incomplete, | - multipath, & - backup, 2 - best2

Network	Next Hop	Metric	LocPrf	Weight	Path
*>i192.168.1.0/24	10.1.1.1	0	100	0	?

# Scenario #3 – eBGP with SVI and vPC

## External Connectivity



[Output]

B2# show ip route vrf myvrf\_50000

IP Route Table for VRF "myvrf\_50000"

'\*' denotes best unicast next-hop

'\*\*' denotes best multicast next-hop

'[x/y]' denotes [preference/metric]

'%<string>' in via output denotes VRF <string>

10.10.10.0/24, ubest/mbest: 1/0

\*via 10.1.33.2, [20/0], 00:22:59, bgp-65001, external, tag 65095

192.168.1.0/24, ubest/mbest: 1/0

\*via 10.1.1.1%default, [200/0], 00:29:35, bgp-65001, internal, tag 65001, segid: 50000 tunnelid: 0xa

[Output]

B2# show ip bgp vrf myvrf\_50000

BGP routing table information for VRF myvrf\_50000, address family IPv4 Unicast

BGP table version is 42, Local Router ID is 10.1.33.1

Status: s-suppressed, x-deleted, S-stale, d-dampened, h-history, \*-valid, >-best

Path type: i-internal, e-external, c-confed, l-local, a-aggregate, r-redist, I-injected

Origin codes: i - IGP, e - EGP, ? - incomplete, | - multipath, & - backup, 2 - best2

Network	Next Hop	Metric	LocPrf	Weight	Path
*>e10.10.0.0/24	10.1.33.2			0	65095 i
*>i192.168.1.0/24	10.1.1.1	0	100	0	?

# Scenario #3 – eBGP with SVI and vPC

## External Connectivity

[Output]

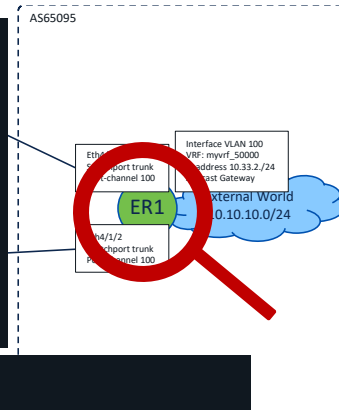
```
ER1# show ip route vrf fabric
IP Route Table for VRF "fabric"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

10.10.10.0/24, ubest/mbest: 1/0 time
    *via Null0, [254/0], 00:34:43, static
192.168.111.0/24, ubest/mbest: 1/0 time
    *via 10.1.33.1, [20/0], 00:26:10, bgp-65095, external, tag 65001
```

[Output]

```
ER1# show ip bgp vrf fabric
BGP routing table information for VRF fabric, address family IPv4 Unicast
BGP table version is 64, Local Router ID is 10.2.33.2
Status: s-suppressed, x-deleted, S-stale, d-dampened, h-history, *-valid, >-best
Path type: i-internal, e-external, c-confed, l-local, a-aggregate, r-redist, I-injected
Origin codes: i - IGP, e - EGP, ? - incomplete, | - multipath, & - backup, 2 - best2
```

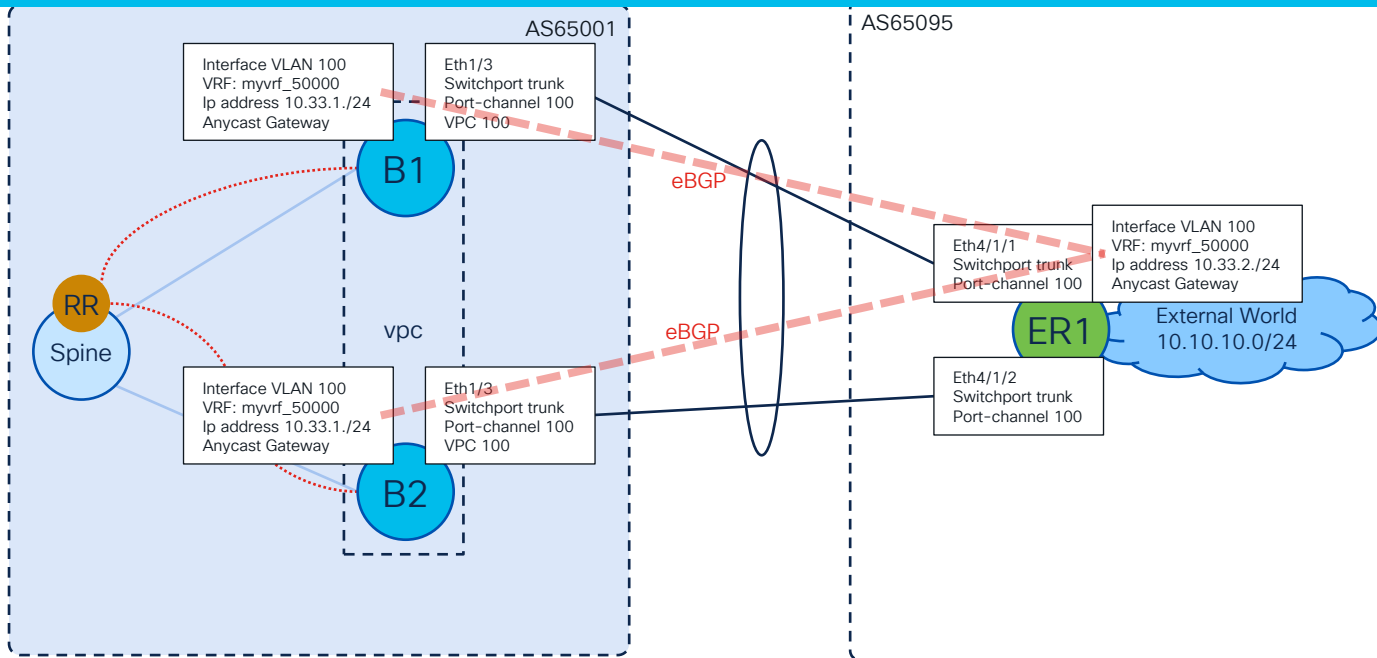
Network	Next Hop	Metric	LocPrf	Weight	Path
*>l10.10.10.0/24	0.0.0.0		100	32768	i
*>e192.168.1.0/24	10.1.33.1			0	65001 ?



# Scenario #3 – eBGP with SVI and vPC

## External Connectivity

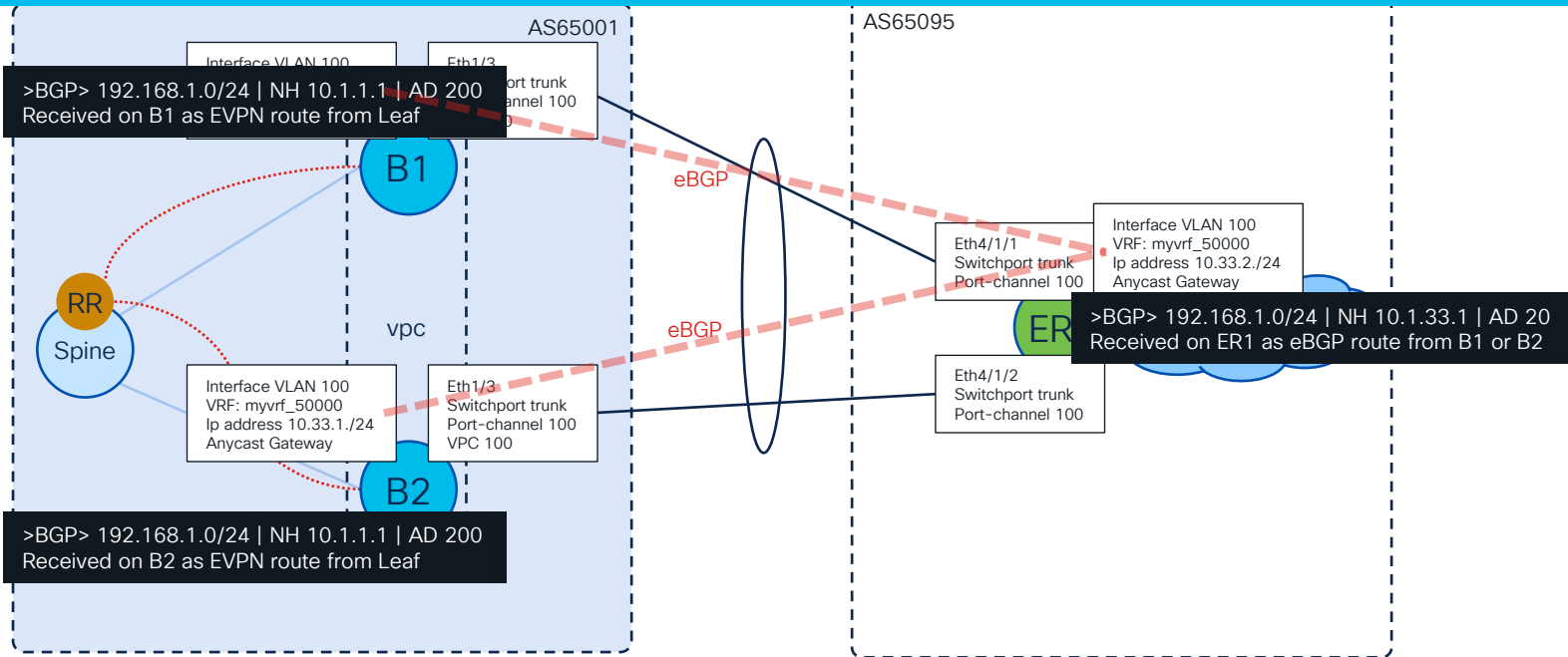
Does Everything Look Good?



# Scenario #3 – eBGP with SVI and vPC

## External Connectivity

Does Everything Look Good?

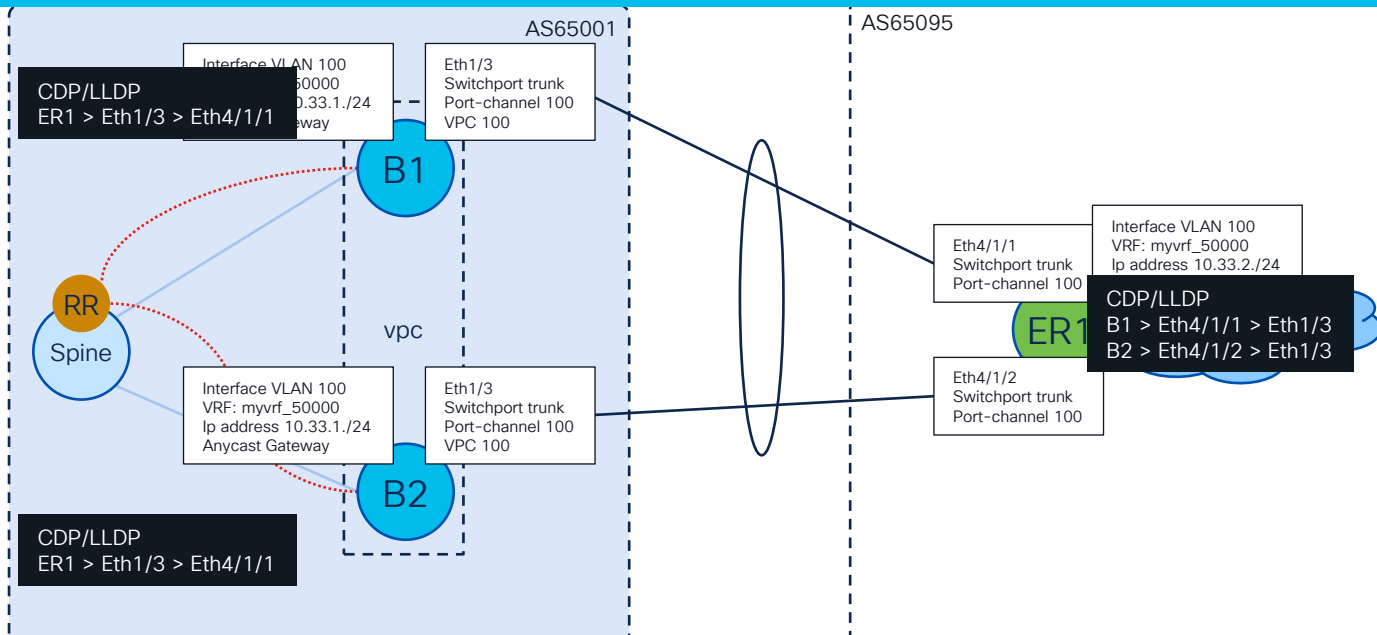




# Scenario #3 – eBGP with SVI and vPC

## External Connectivity

Does Everything Look Good?

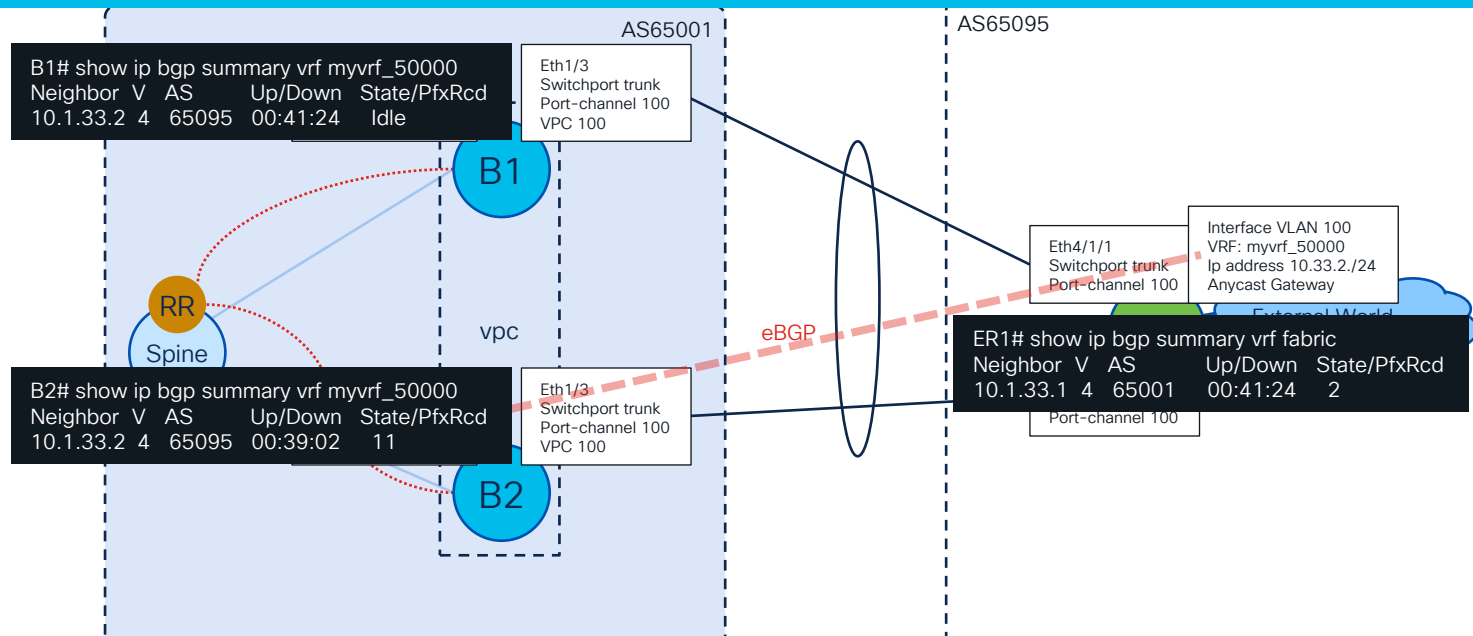


We have Link Redundancy!

# Scenario #3 – eBGP with SVI and vPC

## External Connectivity

Does Everything Look Good?

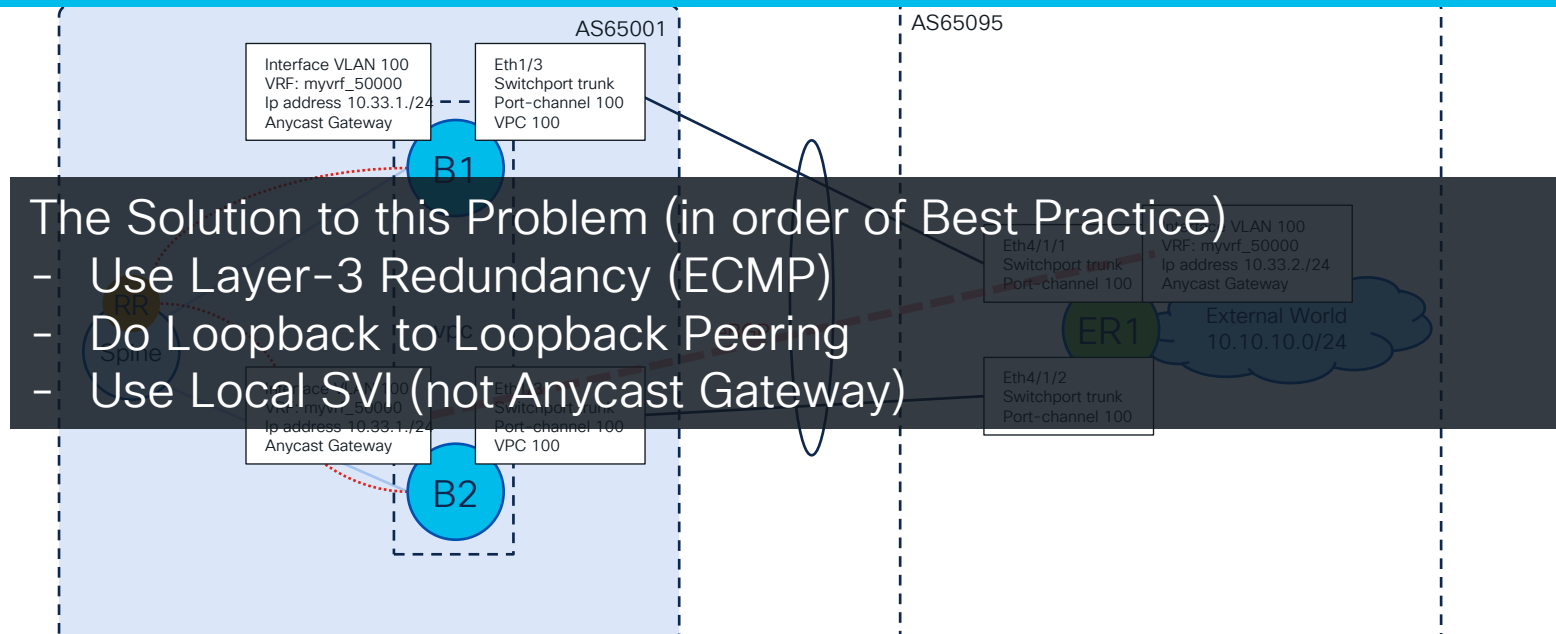


We have NO Routing Protocol Redundancy!

# Scenario #3 – eBGP with SVI and vPC

## External Connectivity

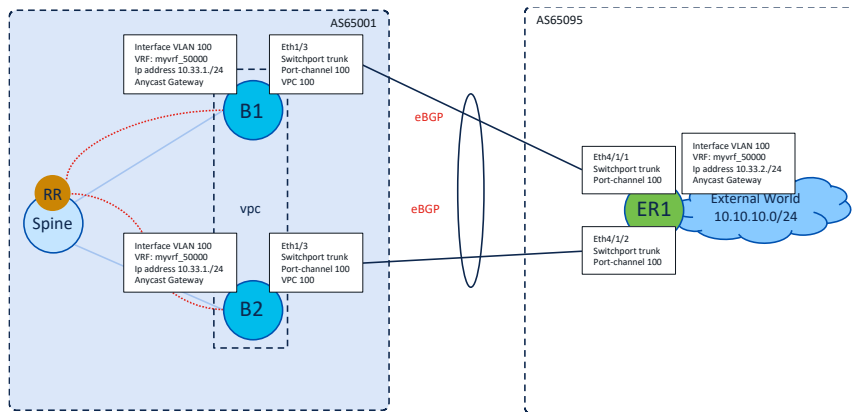
Does Everything Look Good?



We have NO Routing Protocol Redundancy!

# Scenario #3 – eBGP with SVI and vPC

## External Connectivity

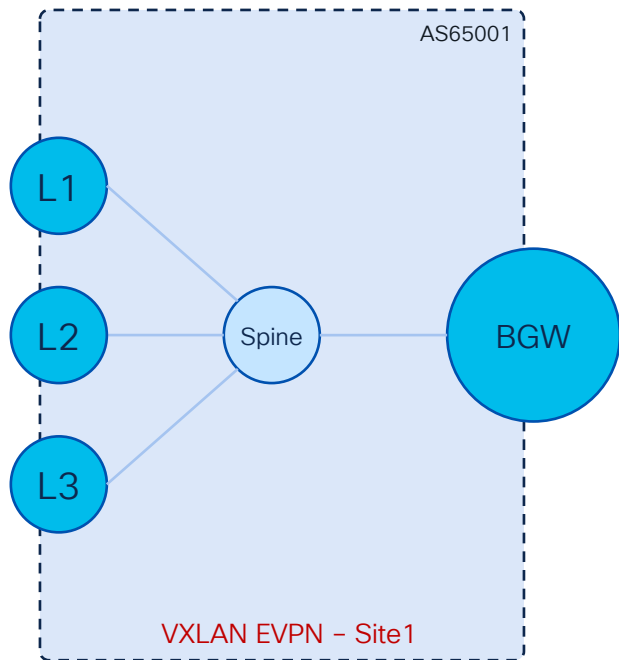


- Sounds Simple but Not Straight Forward
  - How many items to consider for Convergence?
  - Use Loopback to Loopback Peering if you can't avoid this
    - Or use local VLANs (non VXLAN VLANs)
  - Worst case, use a IGP to avoid single session BUT
    - See scenario #2
    - Dynamic Routing over vPC needs lots of considerations
  - Maybe we just do Static Routing!?
  - Avoid Multi-Chassis Link Redundancy with Dynamic Routing
    - Convert it to ECMP

# Gateway Modes

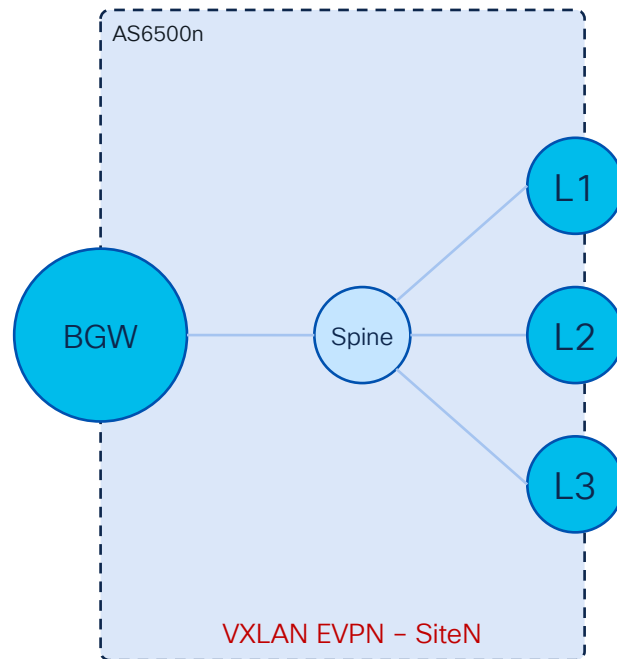
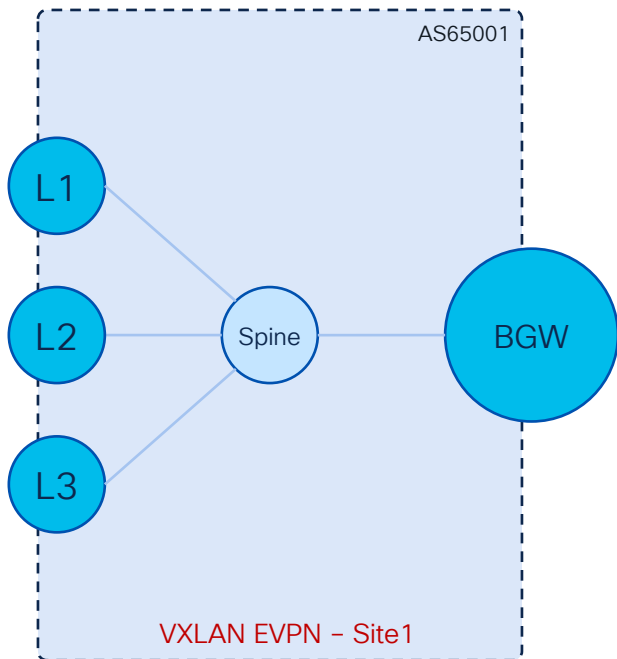
# VXLAN EVPN Multi-Site Border Gateway (BGW)

## Layer-2 and Layer-3 Gateway



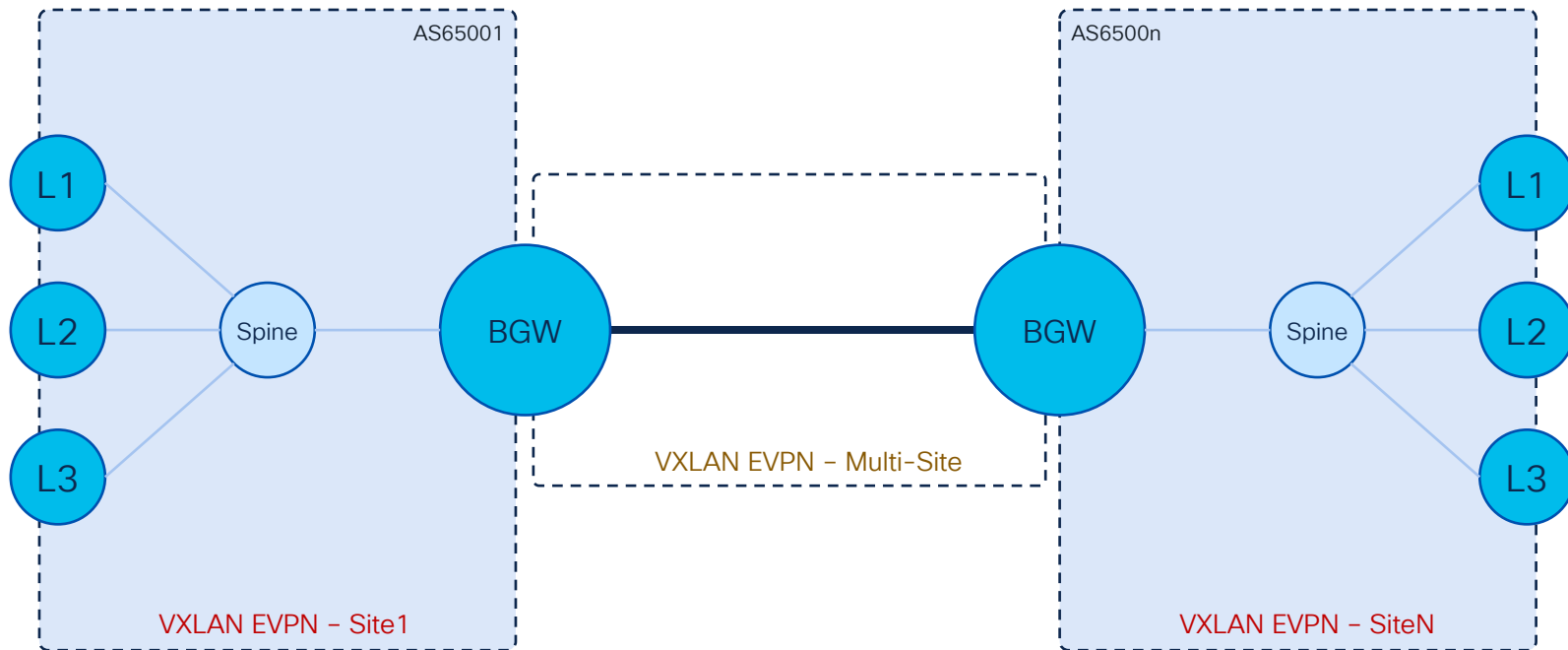
# VXLAN EVPN Multi-Site Border Gateway (BGW)

## Layer-2 and Layer-3 Gateway



# VXLAN EVPN Multi-Site Border Gateway (BGW)

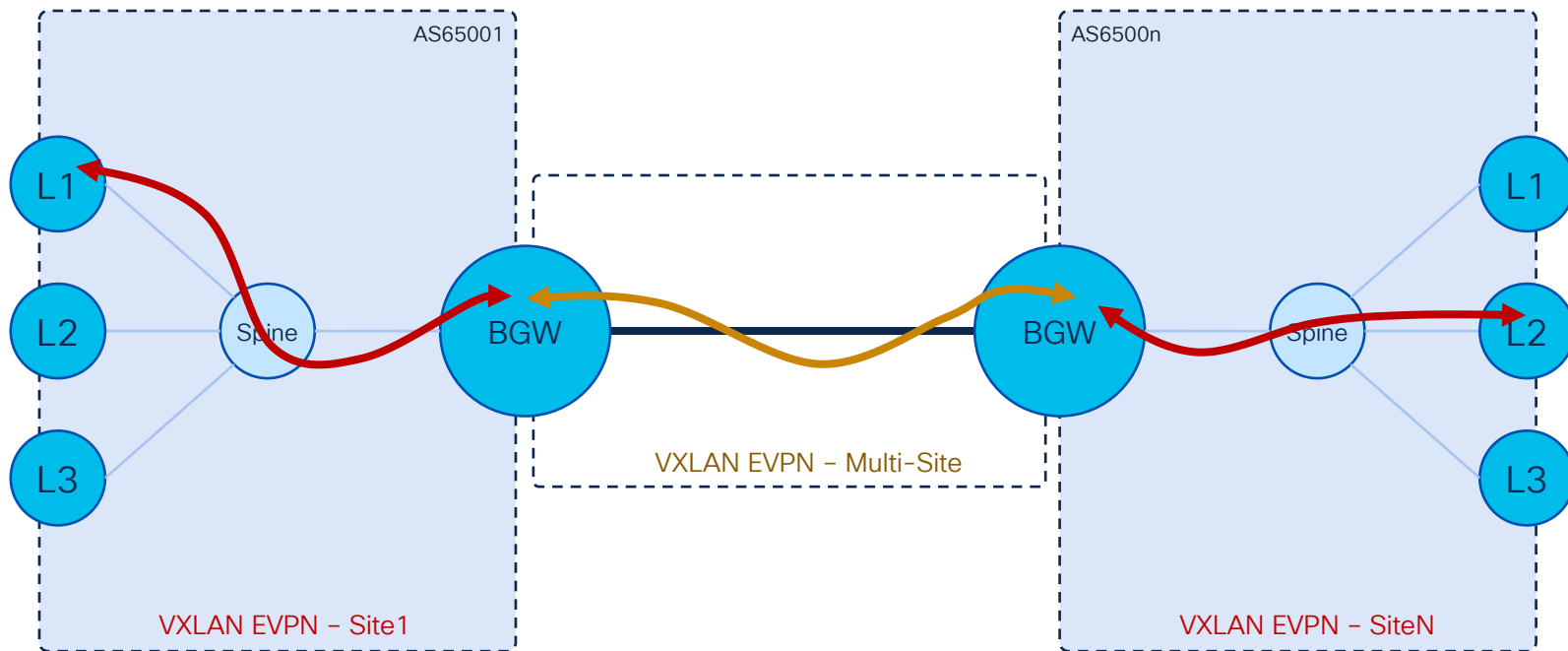
## Layer-2 and Layer-3 Gateway





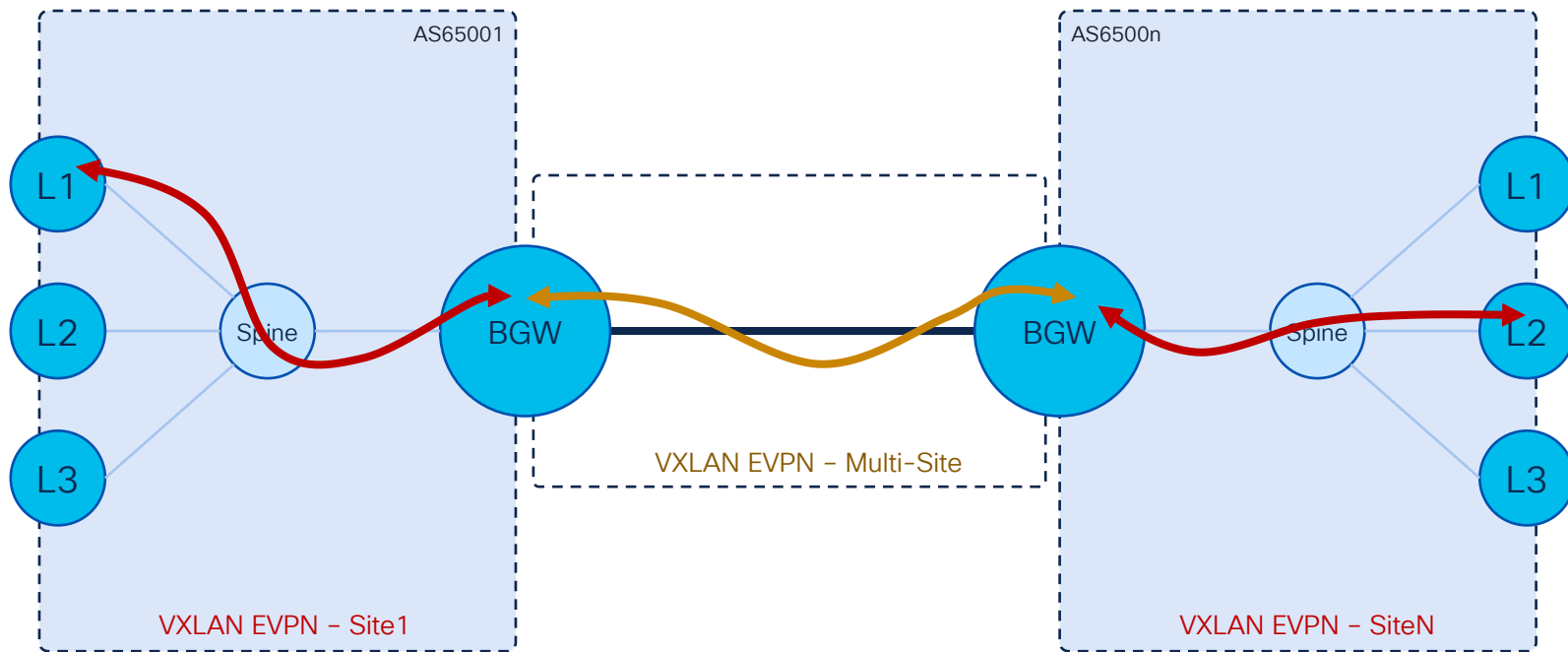
# VXLAN EVPN Multi-Site Border Gateway (BGW)

## Layer-2 and Layer-3 Gateway



# VXLAN EVPN Multi-Site Border Gateway (BGW)

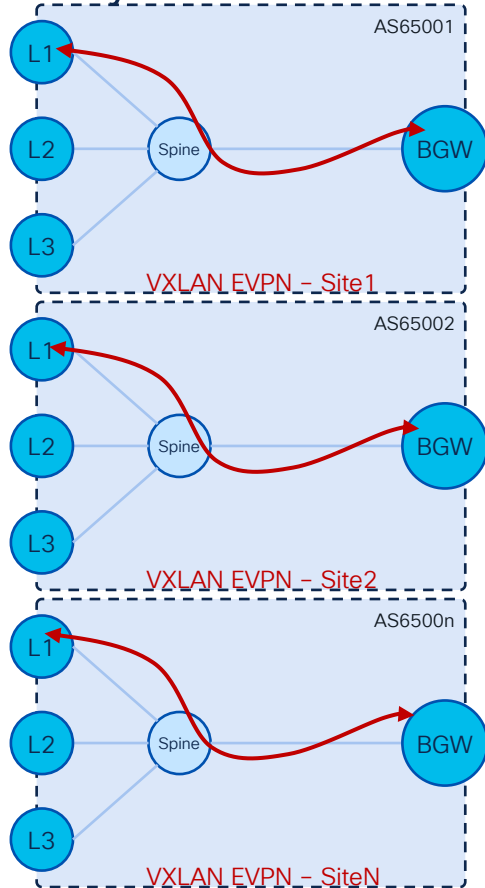
## Layer-2 and Layer-3 Gateway



Extending Layer-2 and Layer-3 with VXLAN EVPN to VXLAN EVPN Gateways (Multi-Site)

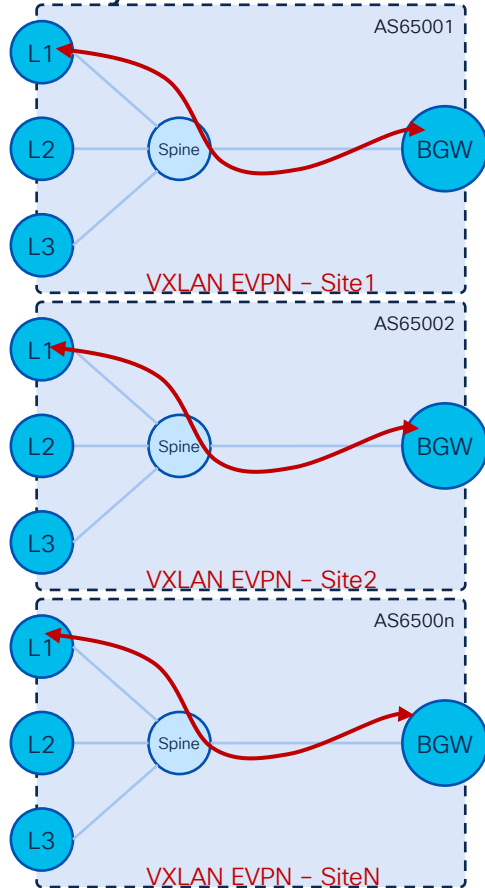
# VXLAN EVPN Gateways

## Layer-3



# VXLAN EVPN Gateways

## Layer-3



Shared  
Border

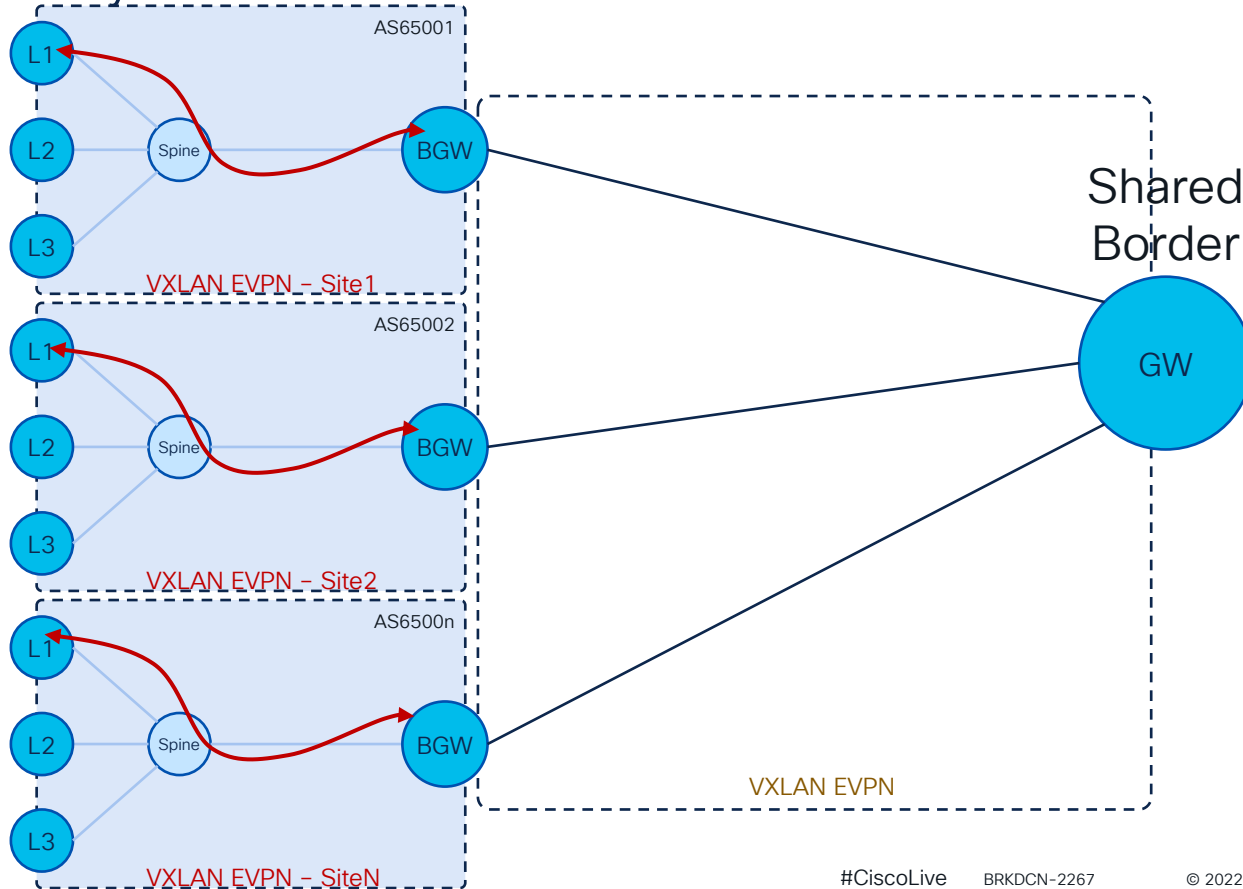


Still a ...

NVE  
VTEP  
Gateway  
or Border

# VXLAN EVPN Gateways

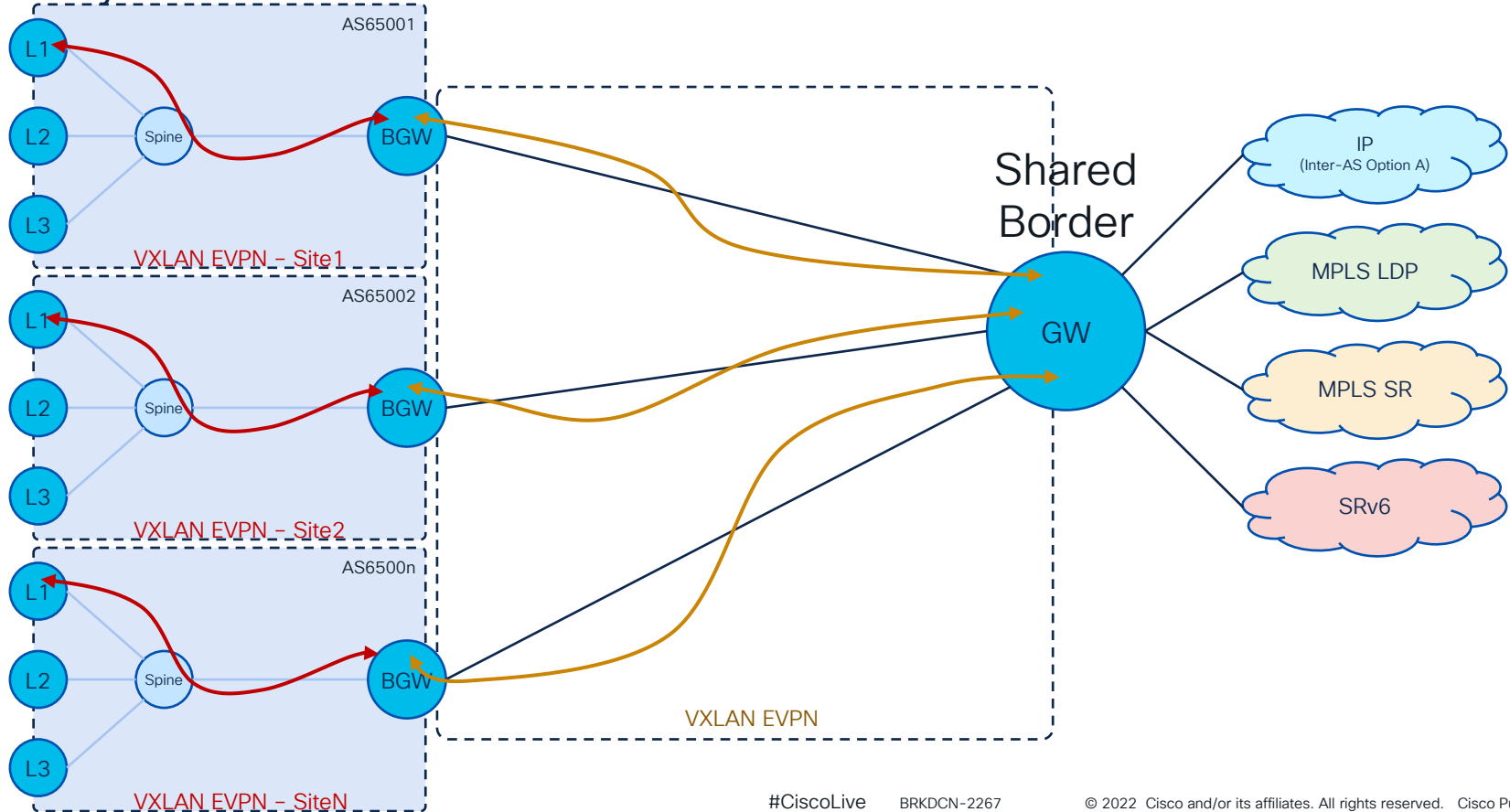
## Layer-3





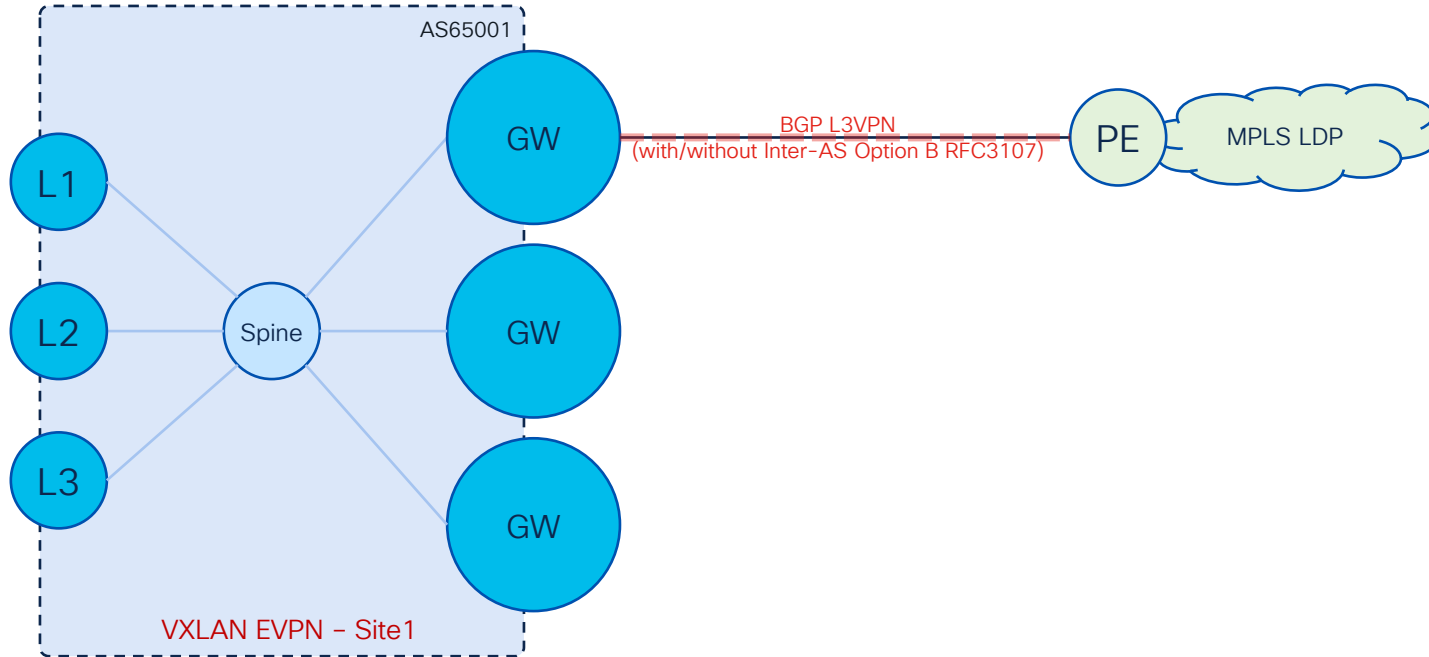
# VXLAN EVPN Gateways

## Layer-3



# VXLAN EVPN Gateways to other Encapsulations

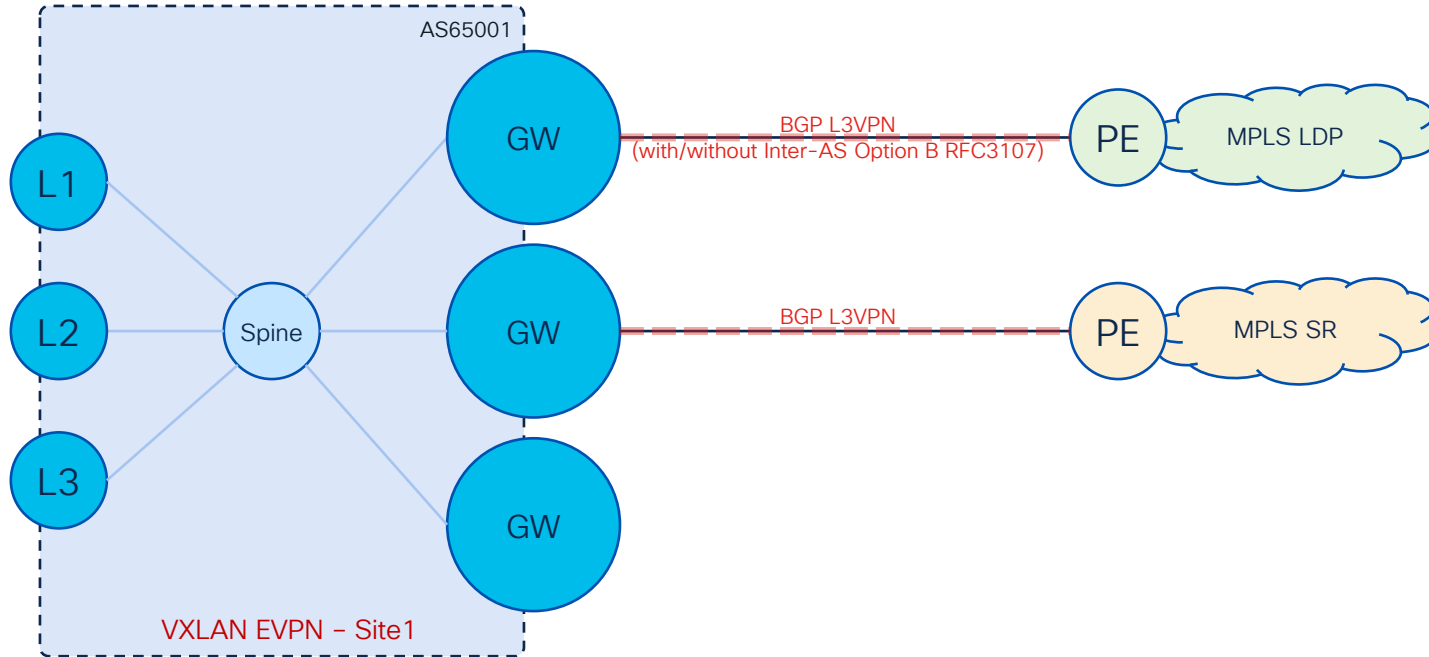
## Layer-3





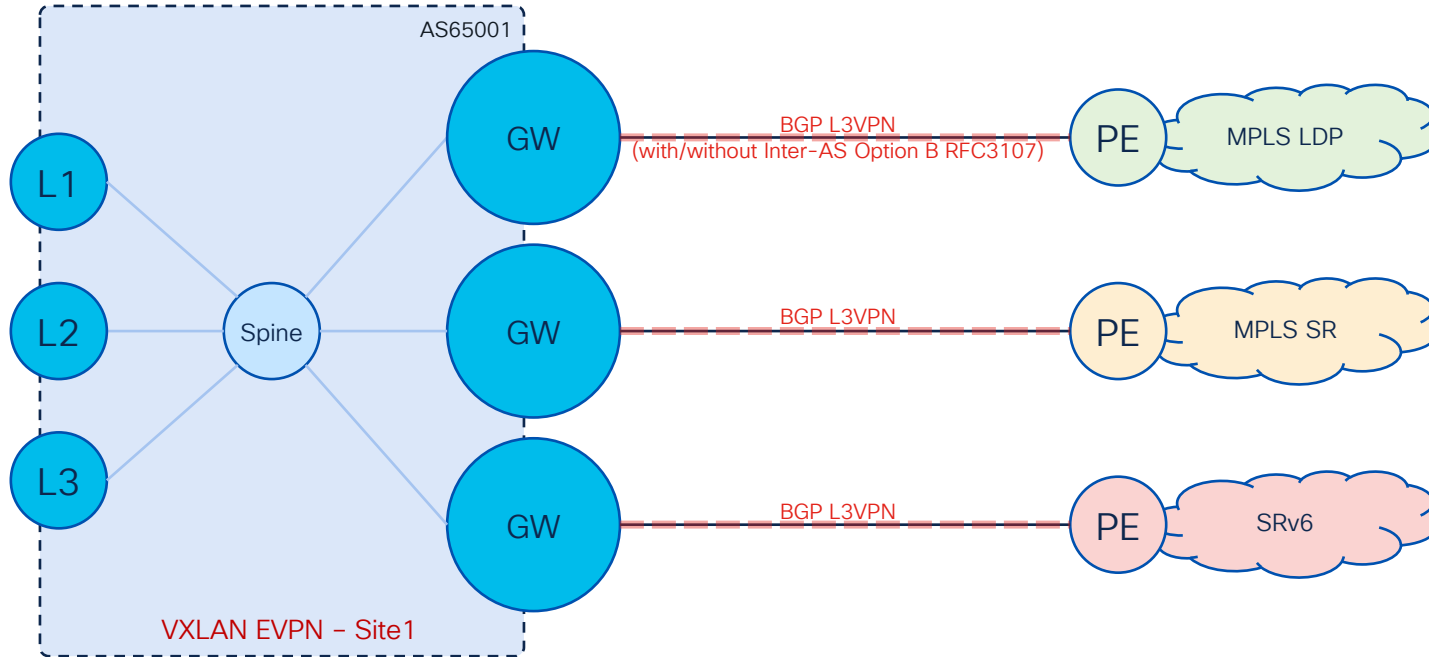
# VXLAN EVPN Gateways to other Encapsulations

## Layer-3



# VXLAN EVPN Gateways to other Encapsulations

## Layer-3



# Conclusion

# Conclusion

- A Border is yet another VTEP
- Border connect to the Outside World of the Fabric
- Many methods to attach, from simple to complicated
- Gateways are essentials to interconnect multiple Domains
- VXLAN to VXLAN Gateways (BGW) are essential for VXLAN EVPN Multi-Site
- Plenty of Options to integrated other Encapsulations
  - From VXLAN (EVPN) to MPLS (LDP), MPLS with RFC3107, MPLS SR or SRv6 (all L3VPN)

# Technical Session Surveys

- Attendees who fill out a minimum of four session surveys and the overall event survey will get Cisco Live branded socks!
- Attendees will also earn 100 points in the Cisco Live Game for every survey completed.
- These points help you get on the leaderboard and increase your chances of winning daily and grand prizes.



# Cisco Learning and Certifications

From technology training and team development to Cisco certifications and learning plans, let us help you empower your business and career. [www.cisco.com/go/certs](https://www.cisco.com/go/certs)

## Pay for Learning with Cisco Learning Credits

(CLCs) are prepaid training vouchers redeemed directly with Cisco.



## Learn

### Cisco U.

IT learning hub that guides teams and learners toward their goals

### Cisco Digital Learning

Subscription-based product, technology, and certification training

### Cisco Modeling Labs

Network simulation platform for design, testing, and troubleshooting

### Cisco Learning Network

Resource community portal for certifications and learning



## Train

### Cisco Training Bootcamps

Intensive team & individual automation and technology training programs

### Cisco Learning Partner Program

Authorized training partners supporting Cisco technology and career certifications

### Cisco Instructor-led and Virtual Instructor-led training

Accelerated curriculum of product, technology, and certification courses



## Certify

### Cisco Certifications and Specialist Certifications

Award-winning certification program empowers students and IT Professionals to advance their technical careers

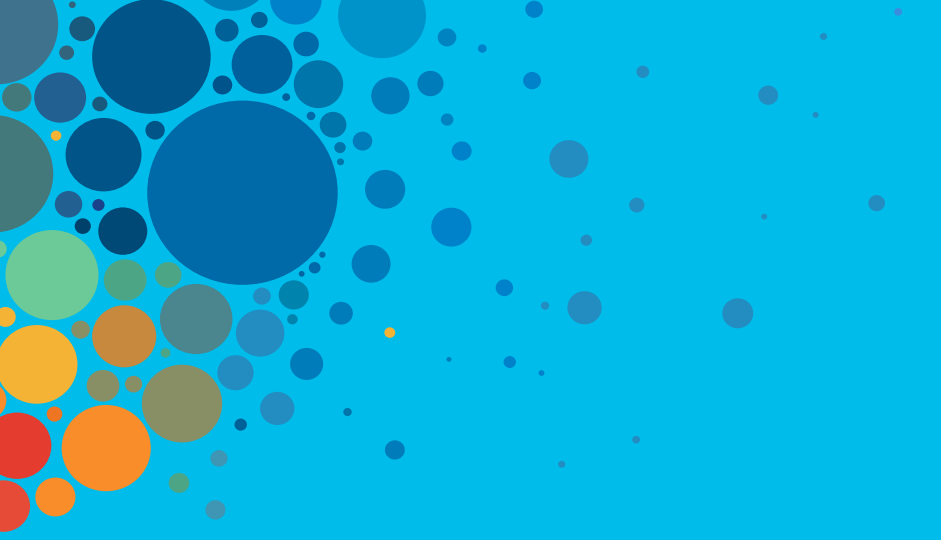
### Cisco Guided Study Groups

180-day certification prep program with learning and support

### Cisco Continuing Education Program

Recertification training options for Cisco certified individuals

Here at the event? Visit us at **The Learning and Certifications lounge at the World of Solutions**



# Continue your education

- Visit the Cisco Showcase for related demos
- Book your one-on-one Meet the Engineer meeting
- Attend the interactive education with DevNet, Capture the Flag, and Walk-in Labs
- Visit the On-Demand Library for more sessions at [www.CiscoLive.com/on-demand](https://www.CiscoLive.com/on-demand)



The bridge to possible

# Thank you



CISCO *Live!*



#CiscoLive