
cisco Live!
Let's go

The bridge to possible

A Day in the Life of a Packet

VXLAN BGP EVPN Fabrics

Lukas Krattiger, Cisco Fellow @CCIE21921

cisco Live!

BRKDCN-3966

Introduction

- We are going to cover all different kind of Packet Walks
- Route, Bridge, BUM and Silent Host Discovery
- A brief intro to VXLAN with EVPN
  - Sorry, not a VXLAN or VXLAN EVPN Intro Session

Agenda

- Introduction to VXLAN EVPN
- Layer-3 Packet Walk
- Layer-2 Packet Walk
- BUM Packet Walk
- Silent Host Discovery
- 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

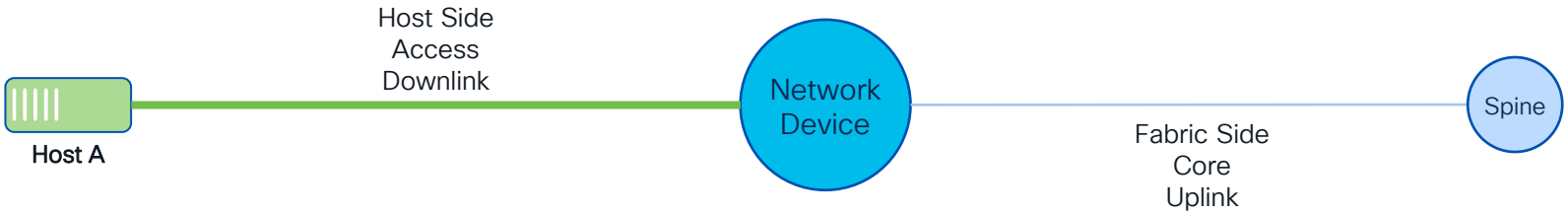
What is EVPN?

- 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)

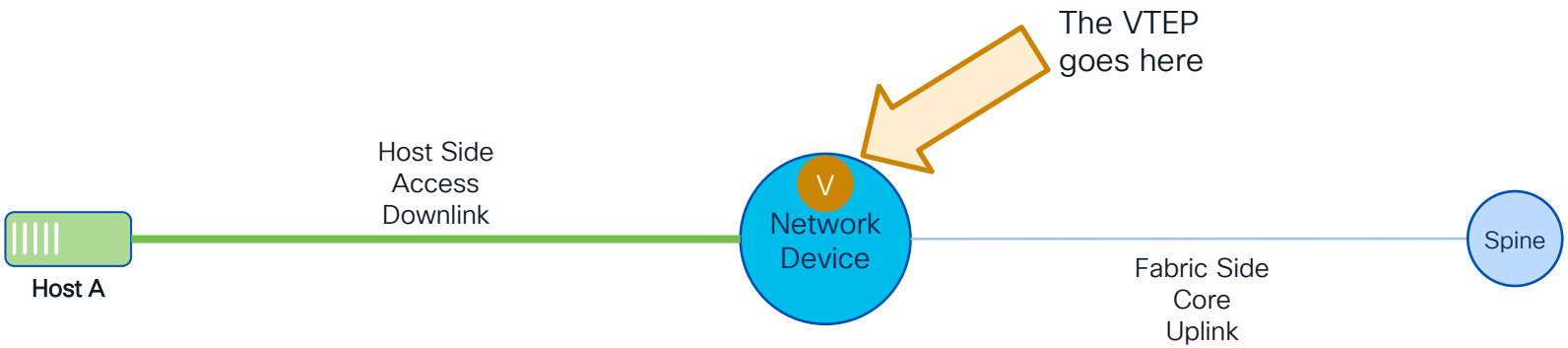
It all starts with a Network Device
The Dating Network - When Control- meets Data-Plane

It all starts with a Network Device

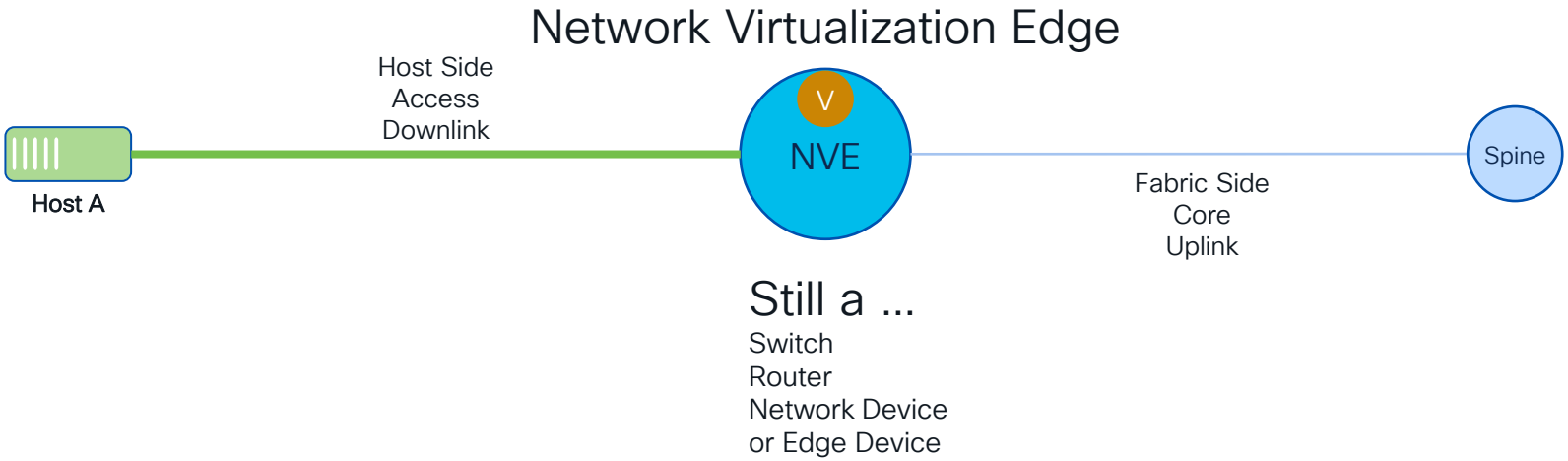
The Dating Network - When Control- meets Data-Plane



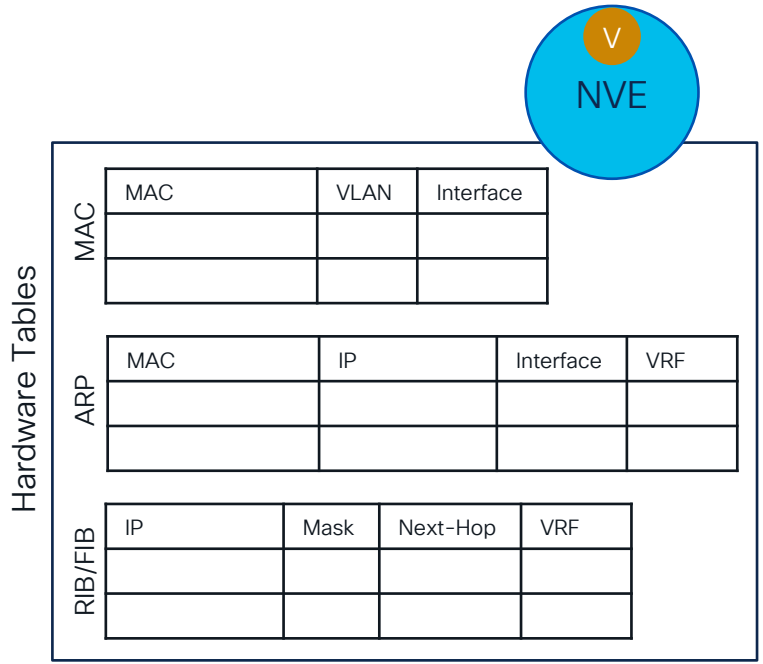
It all starts with a Network Device
The Dating Network - When Control- meets Data-Plane



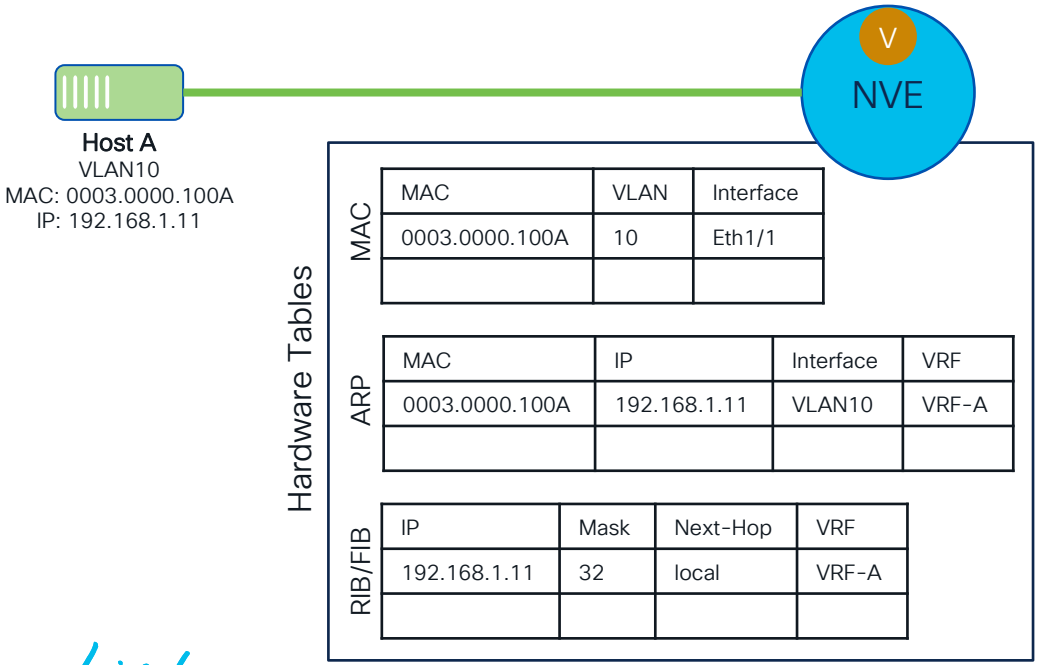
Making the Network Device an NVE
The Dating Network - When Control- meets Data-Plane



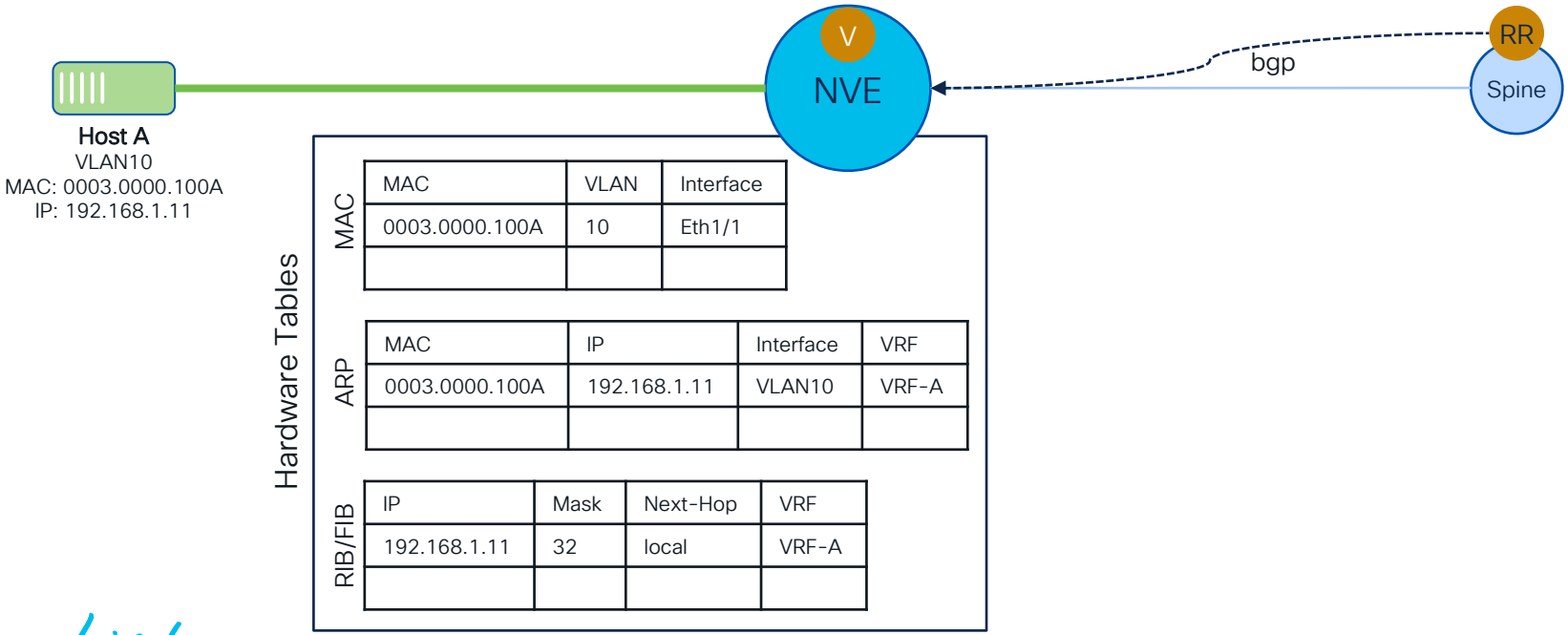
The NVE and Some Important Table
The Dating Network - When Control- meets Data-Plane



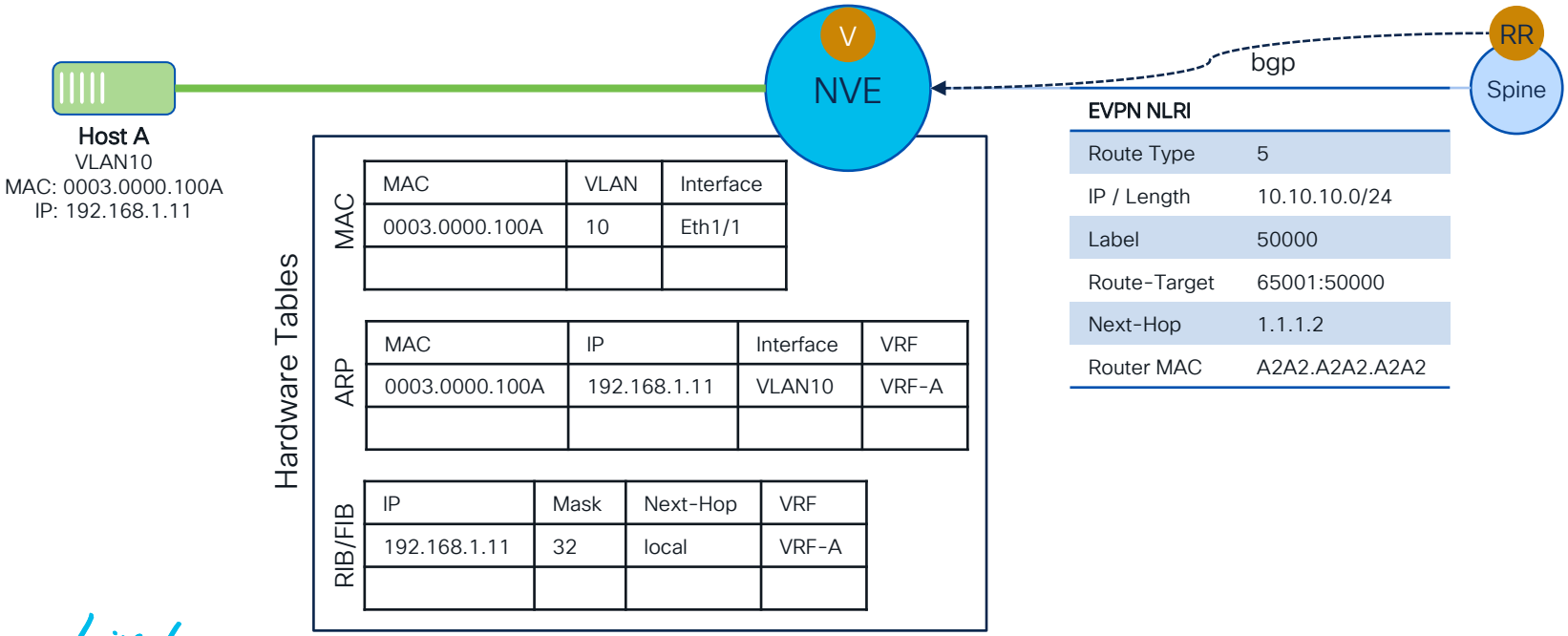
Local Learning on the NVE
The Dating Network - When Control- meets Data-Plane



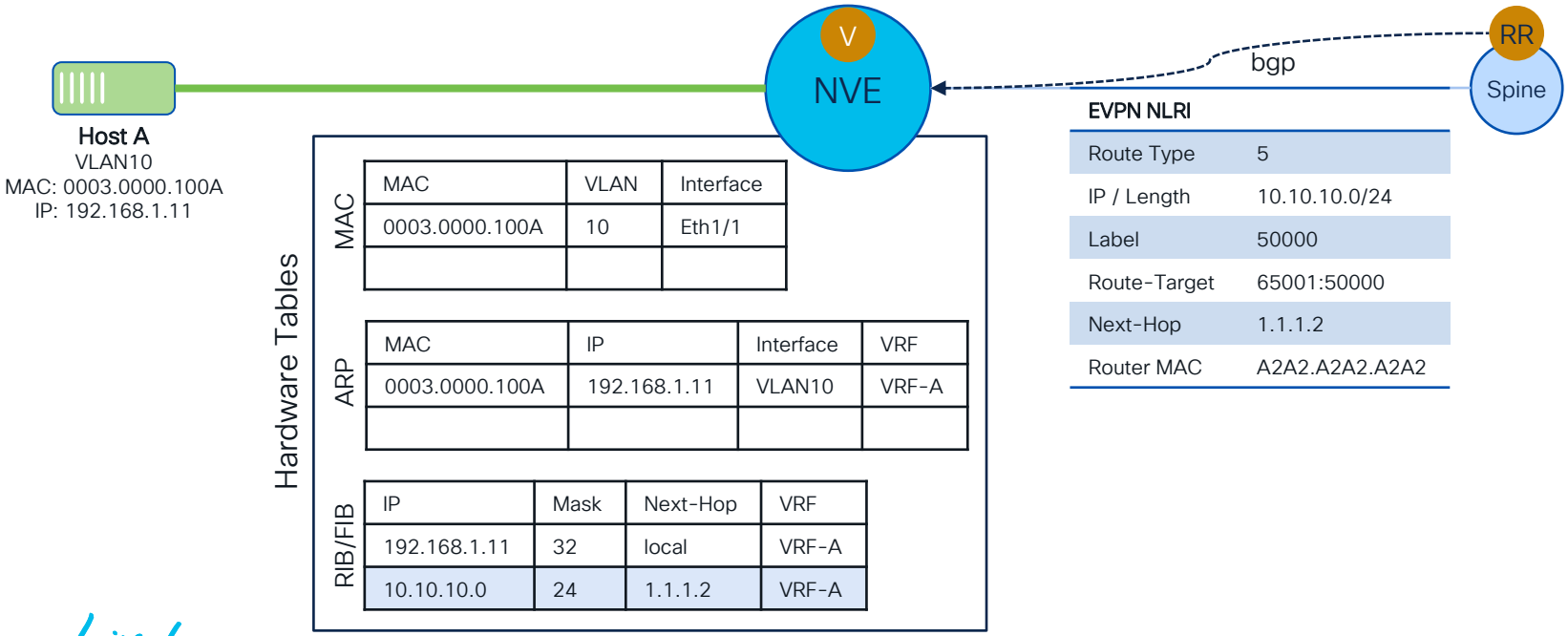
Remote Learning from other NVE
The Dating Network - When Control- meets Data-Plane



Remote Learning from other NVE
The Dating Network - When Control- meets Data-Plane

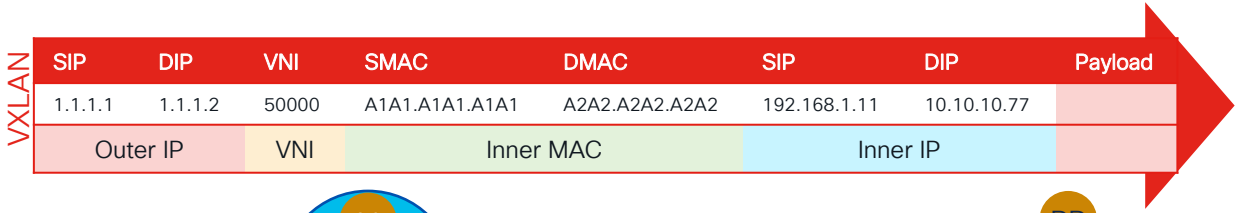


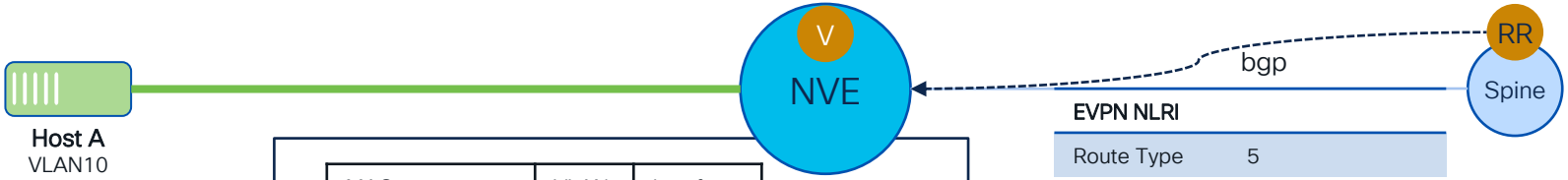
Remote Learning from other NVE
The Dating Network - When Control- meets Data-Plane



Routing between NVE (based on VXLAN EVPN)

The Dating Network - When Control- meets Data-Plane





Host A
VLAN10
MAC: 0003.0000.100A
IP: 192.168.1.11

Hardware Tables

MAC

MAC	VLAN	Interface
0003.0000.100A	10	Eth1/1

ARP

MAC	IP	Interface	VRF
0003.0000.100A	192.168.1.11	VLAN10	VRF-A

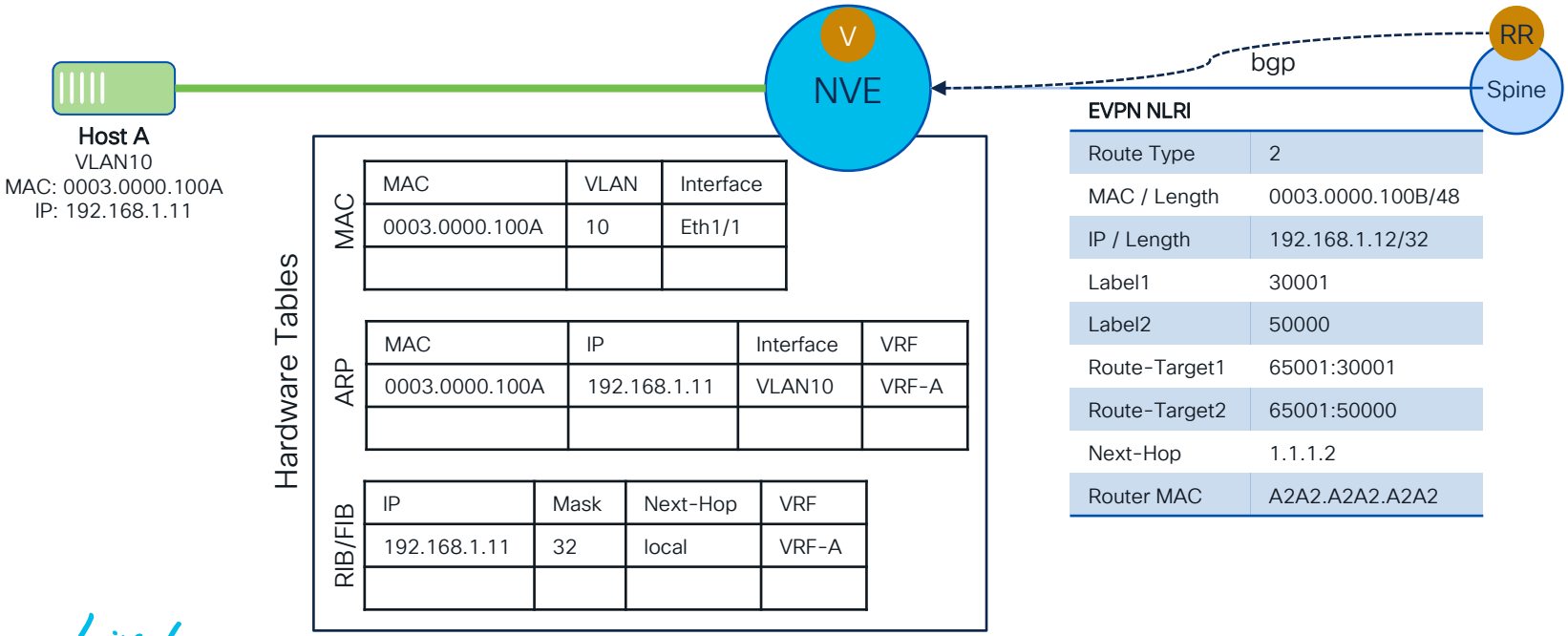
RIB/FIB

IP	Mask	Next-Hop	VRF
192.168.1.11	32	local	VRF-A
10.10.10.0	24	1.1.1.2	VRF-A

EVPN NLRI

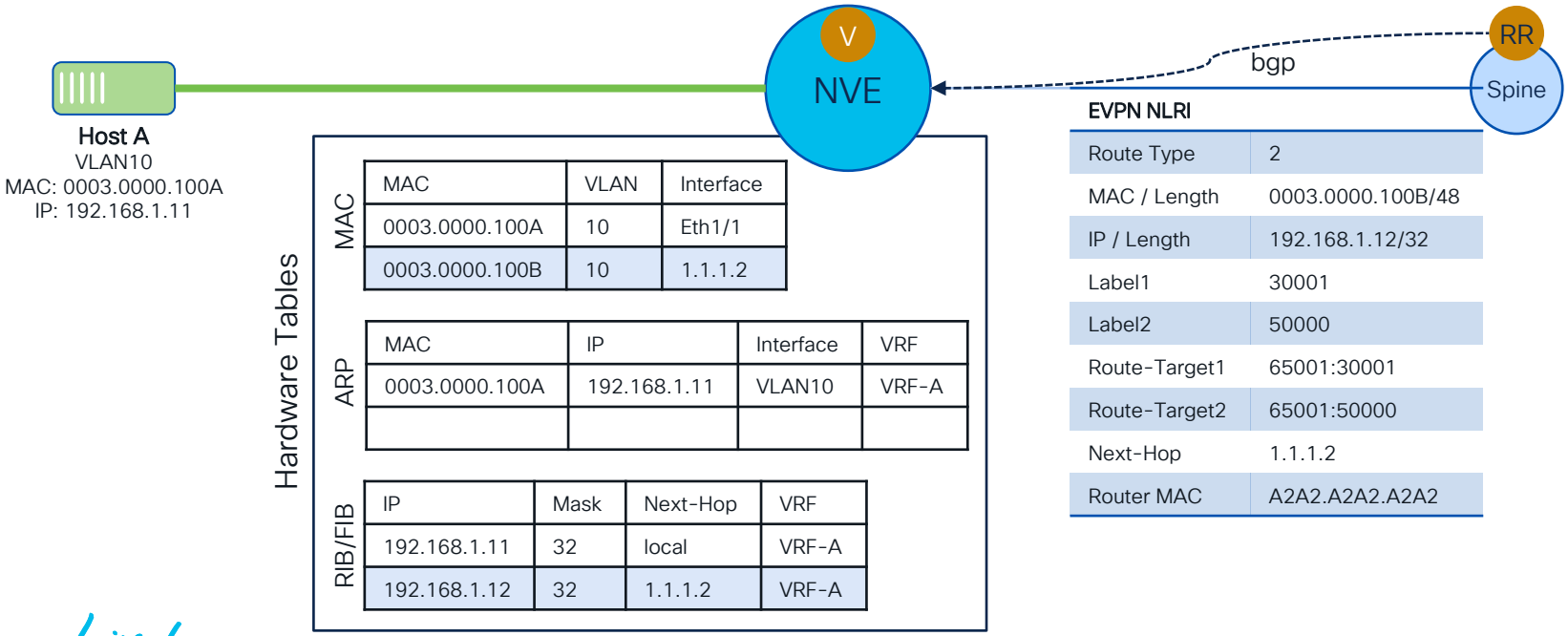
Route Type	5
IP / Length	10.10.10.0/24
Label	50000
Route-Target	65001:50000
Next-Hop	1.1.1.2
Router MAC	A2A2.A2A2.A2A2

Remote Learning from other NVE
The Dating Network - When Control- meets Data-Plane



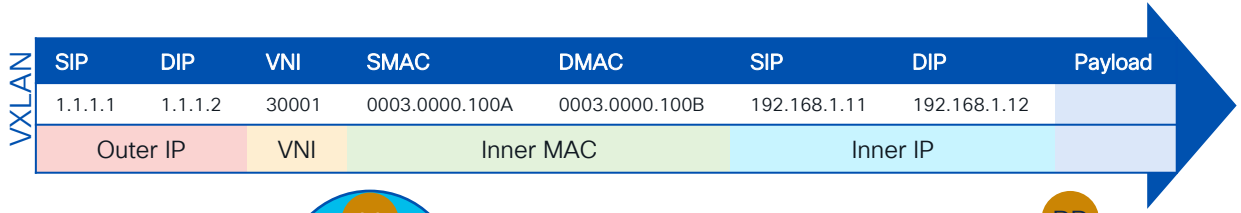
EVPN NLRI	
Route Type	2
MAC / Length	0003.0000.100B/48
IP / Length	192.168.1.12/32
Label1	30001
Label2	50000
Route-Target1	65001:30001
Route-Target2	65001:50000
Next-Hop	1.1.1.2
Router MAC	A2A2.A2A2.A2A2

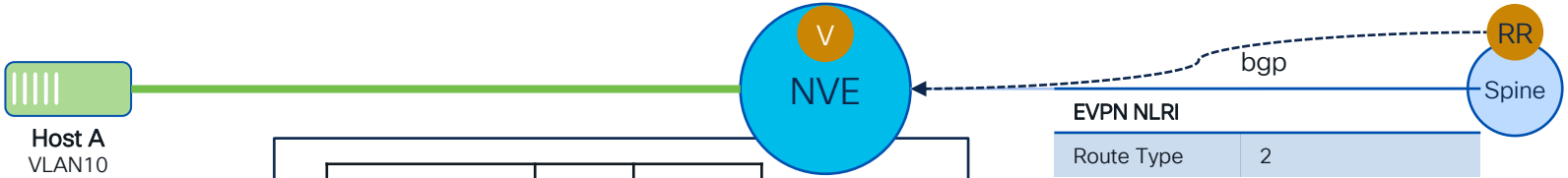
Remote Learning from other NVE
The Dating Network - When Control- meets Data-Plane



Bridging between NVE (based on VXLAN EVPN)

The Dating Network - When Control- meets Data-Plane





Host A
VLAN10
MAC: 0003.0000.100A
IP: 192.168.1.11

Hardware Tables

MAC

MAC	VLAN	Interface
0003.0000.100A	10	Eth1/1
0003.0000.100B	10	1.1.1.2

ARP

MAC	IP	Interface	VRF
0003.0000.100A	192.168.1.11	VLAN10	VRF-A

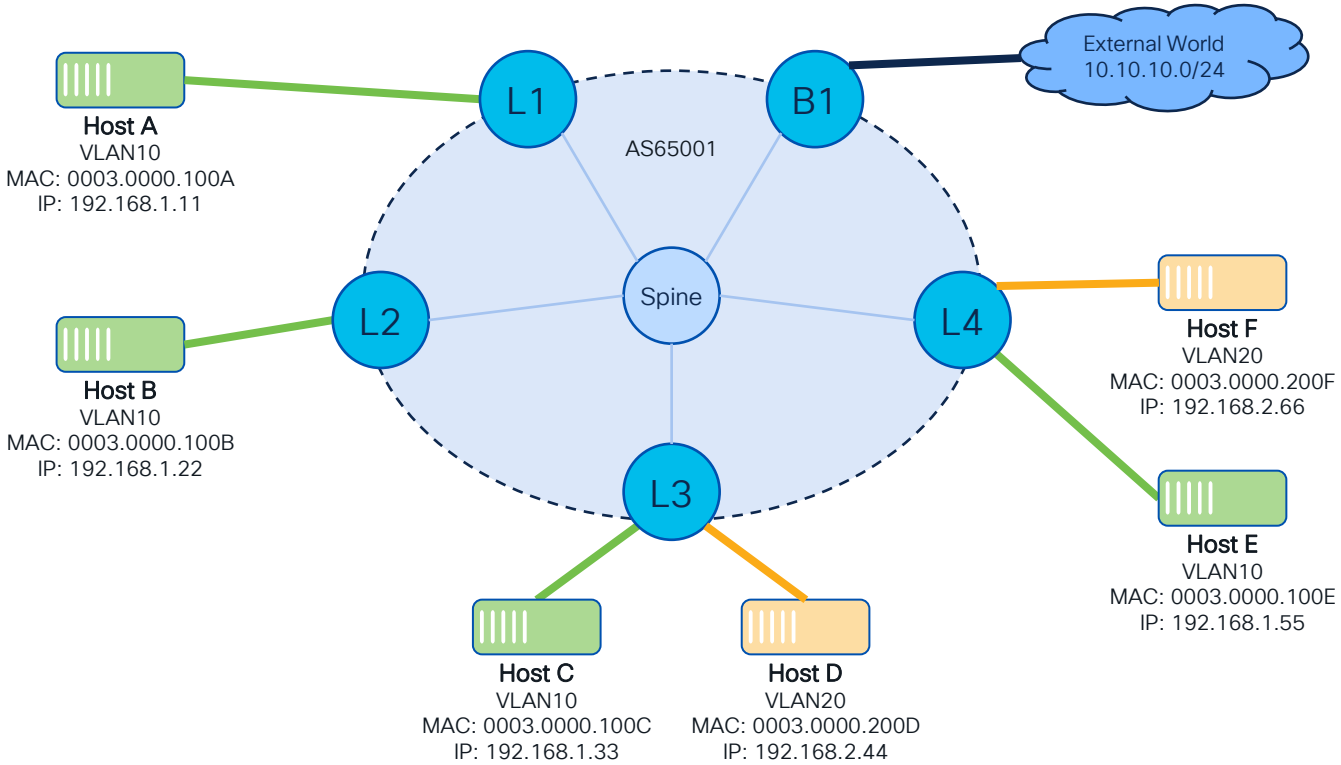
RIB/FIB

IP	Mask	Next-Hop	VRF
192.168.1.11	32	local	VRF-A
192.168.1.12	24	1.1.1.2	VRF-A

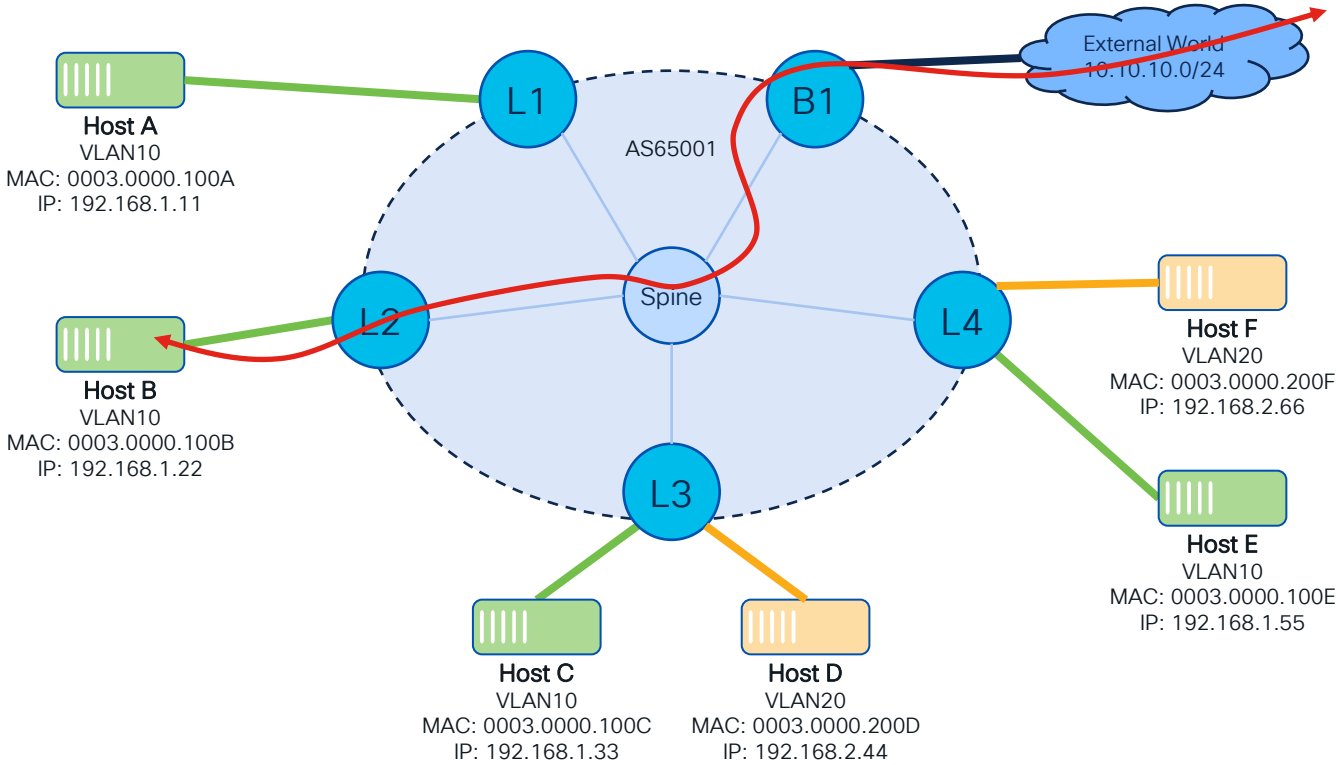
EVPN NLRI	
Route Type	2
MAC / Length	0003.0000.100B/48
IP / Length	192.168.1.12/32
Label1	30001
Label2	50000
Route-Target1	65001:30001
Route-Target2	65001:50000
Next-Hop	1.1.1.2
Router MAC	A2A2.A2A2.A2A2

Packet Walk: Layer-3 – Host to External World

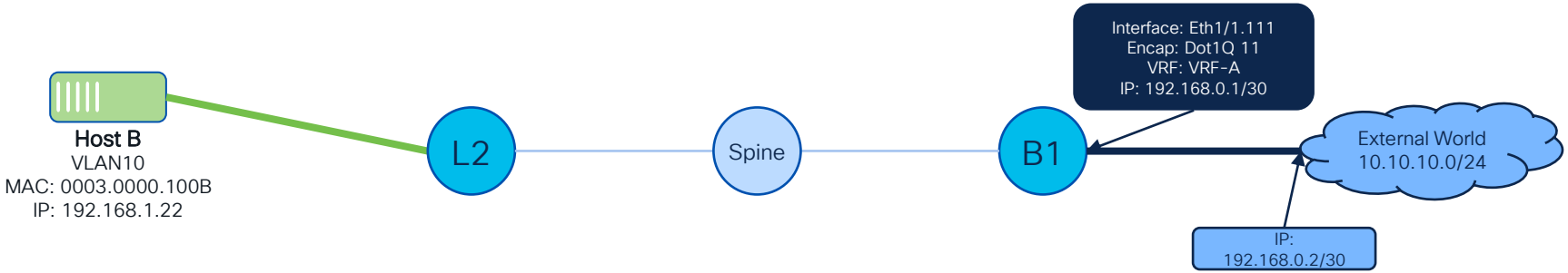
Topology Overview
Layer-3 Packet Walk



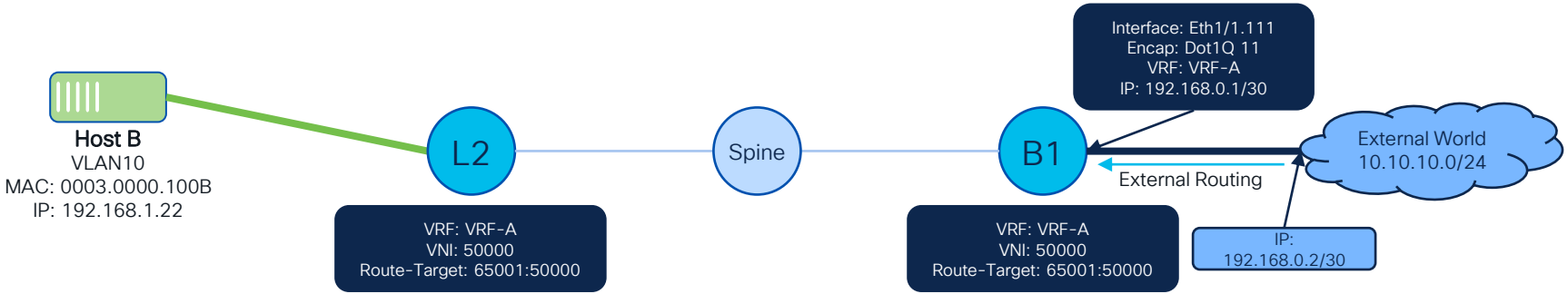
Topology Overview
Layer-3 Packet Walk



Learning: External World to Leaf2

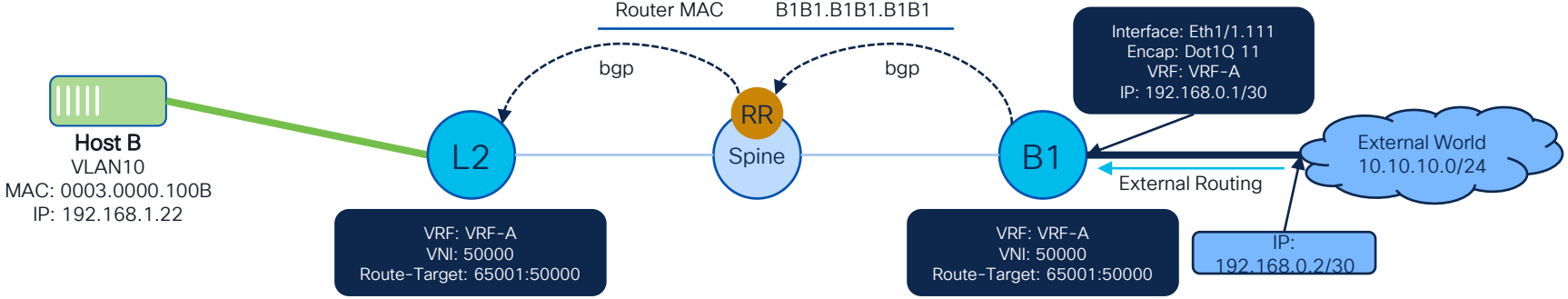


Learning: External World to Leaf2

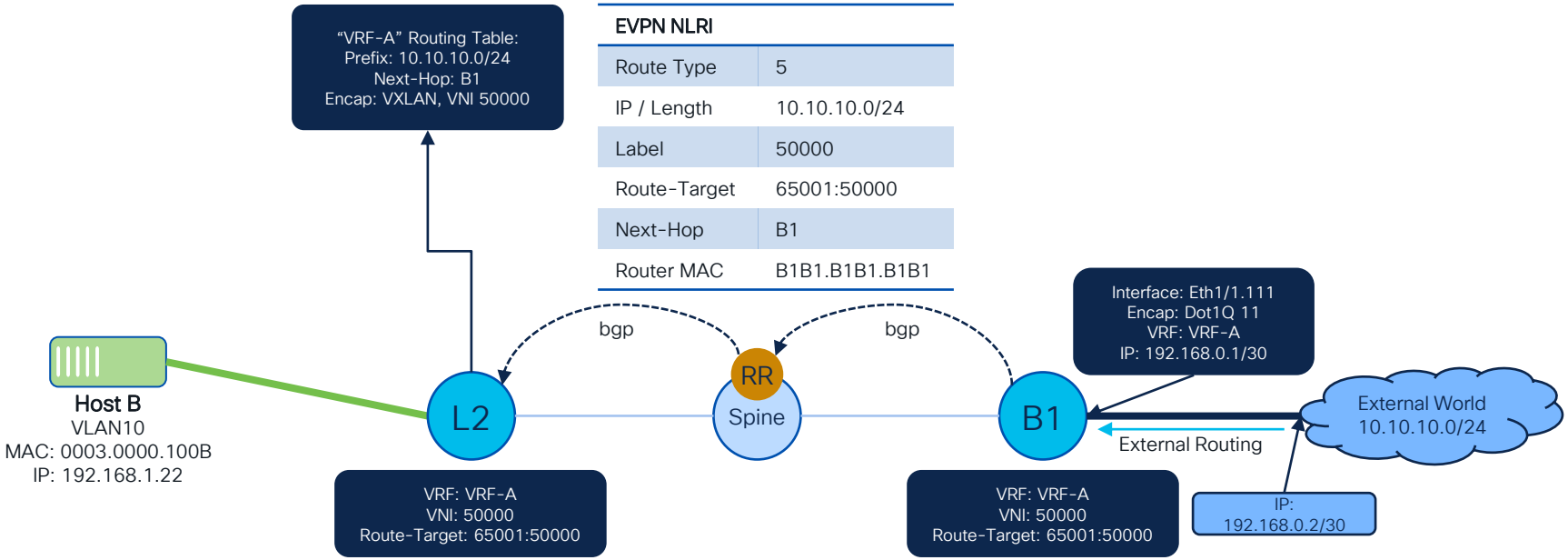


Learning: External World to Leaf2

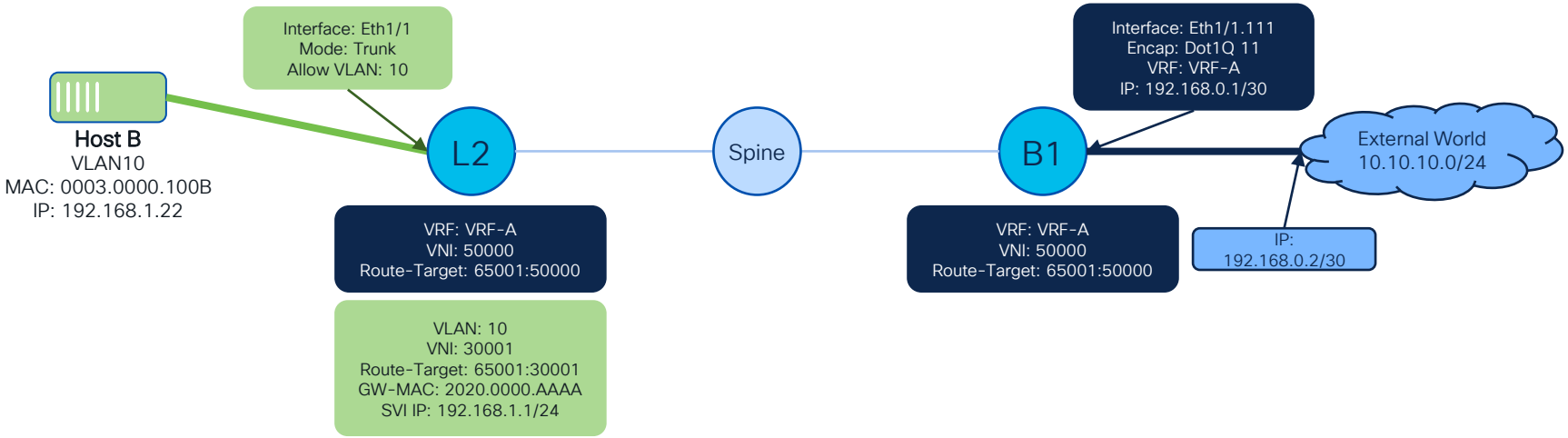
EVPN NLRI	
Route Type	5
IP / Length	10.10.10.0/24
Label	50000
Route-Target	65001:50000
Next-Hop	B1
Router MAC	B1B1.B1B1.B1B1



Learning: External World to Leaf2

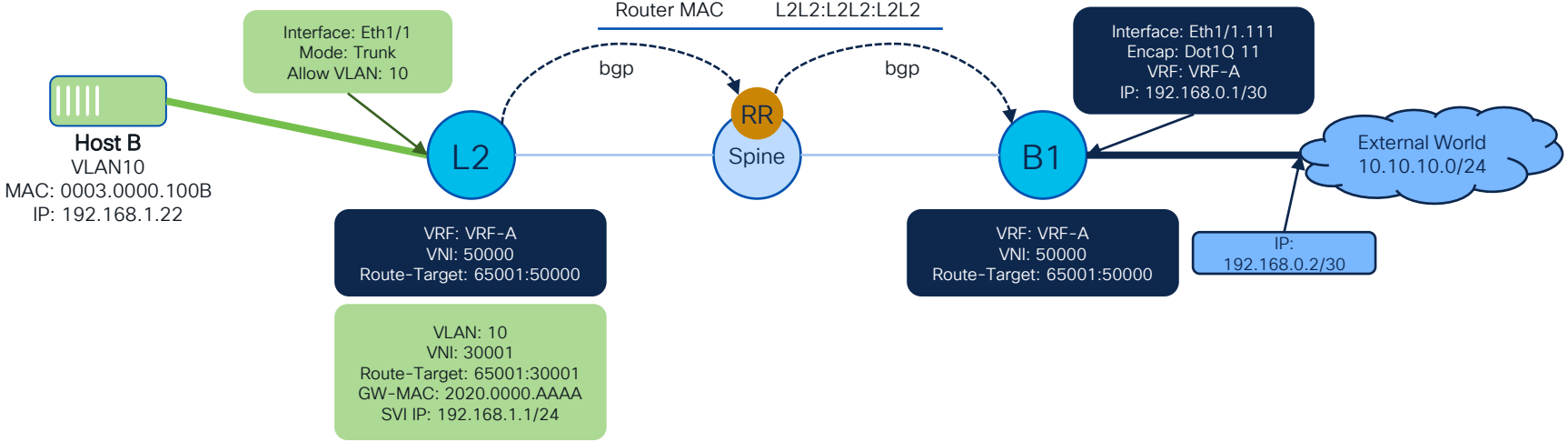


Learning: HostB to External World

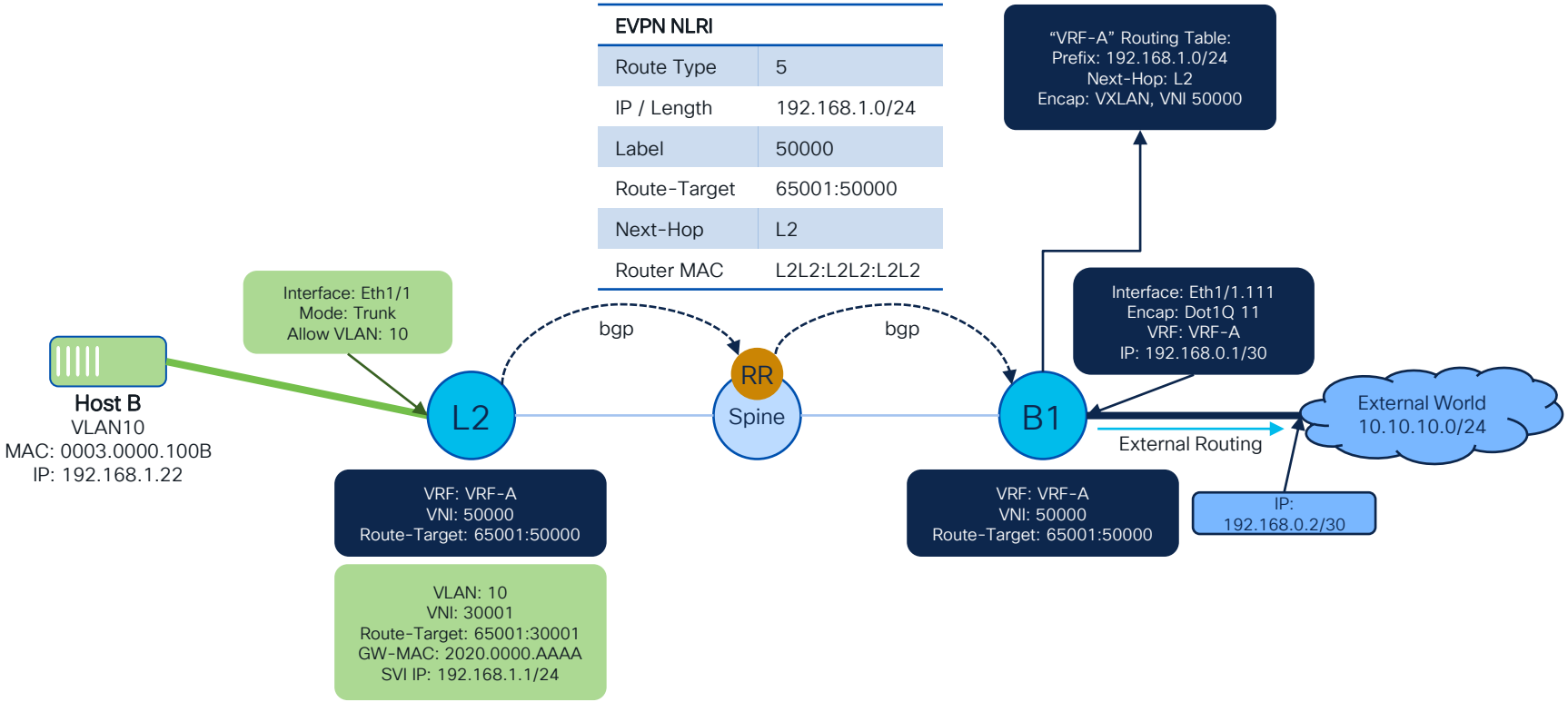


Learning: HostB to External World

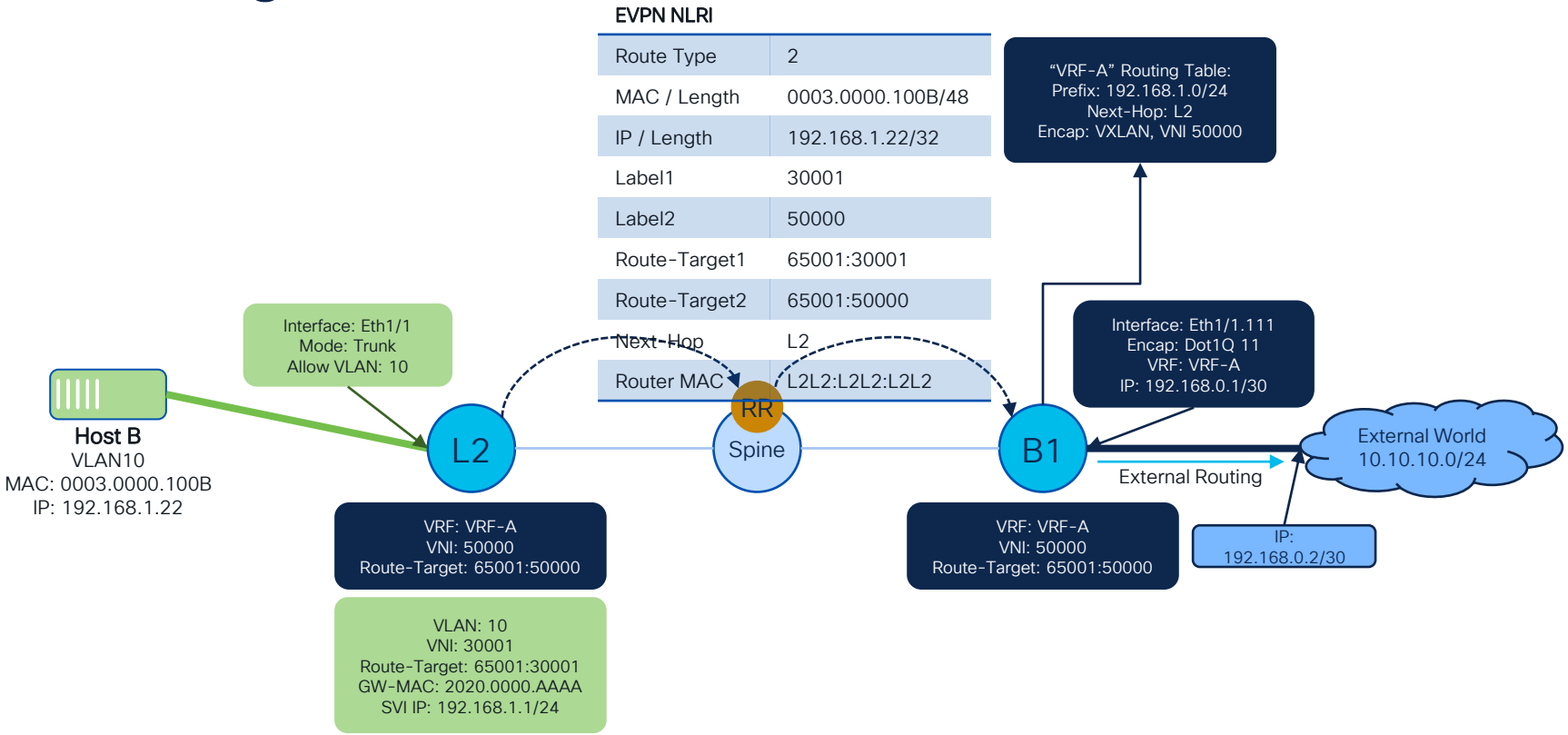
EVPN NLRI	
Route Type	5
IP / Length	192.168.1.0/24
Label	50000
Route-Target	65001:50000
Next-Hop	L2
Router MAC	L2L2:L2L2:L2L2



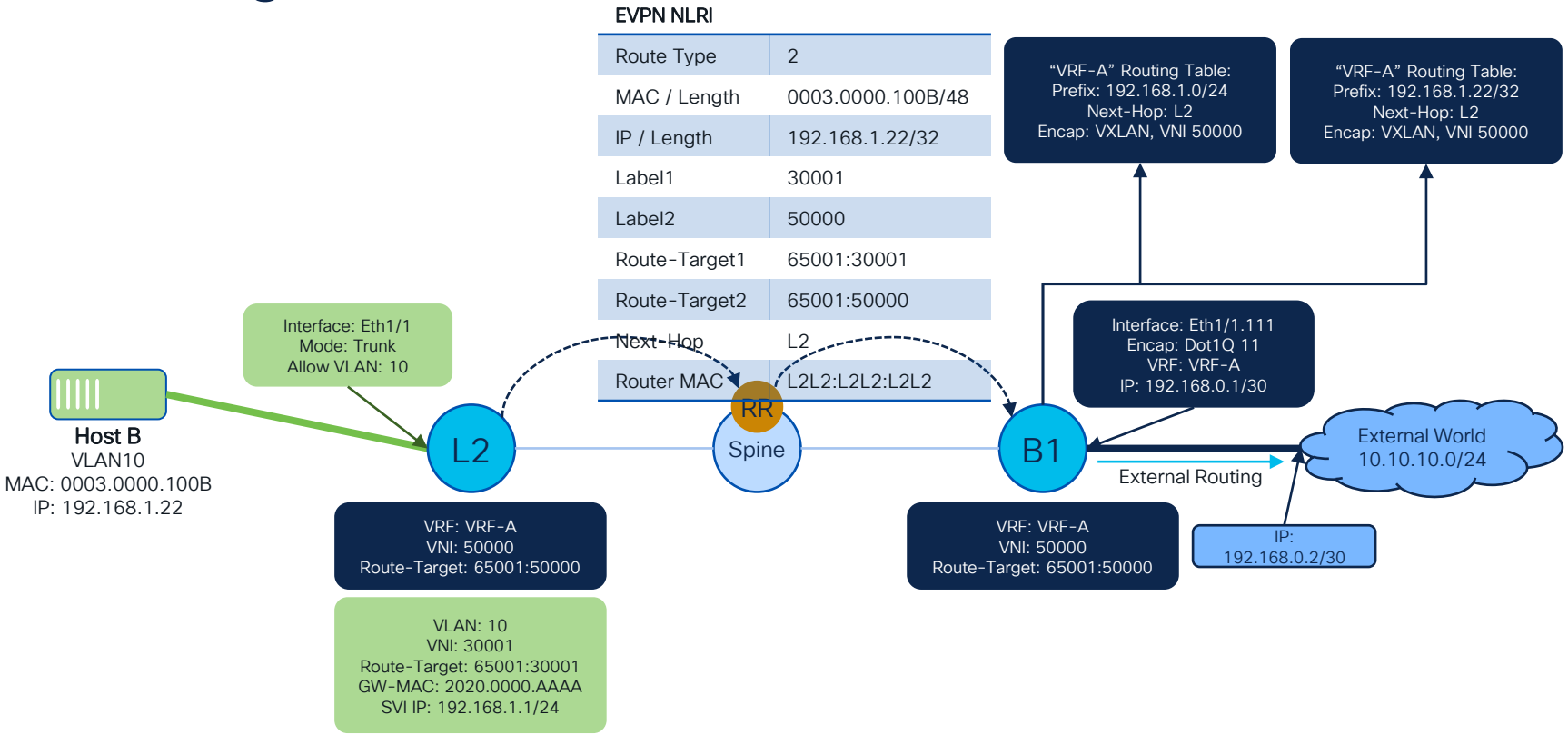
Learning: HostB to External World



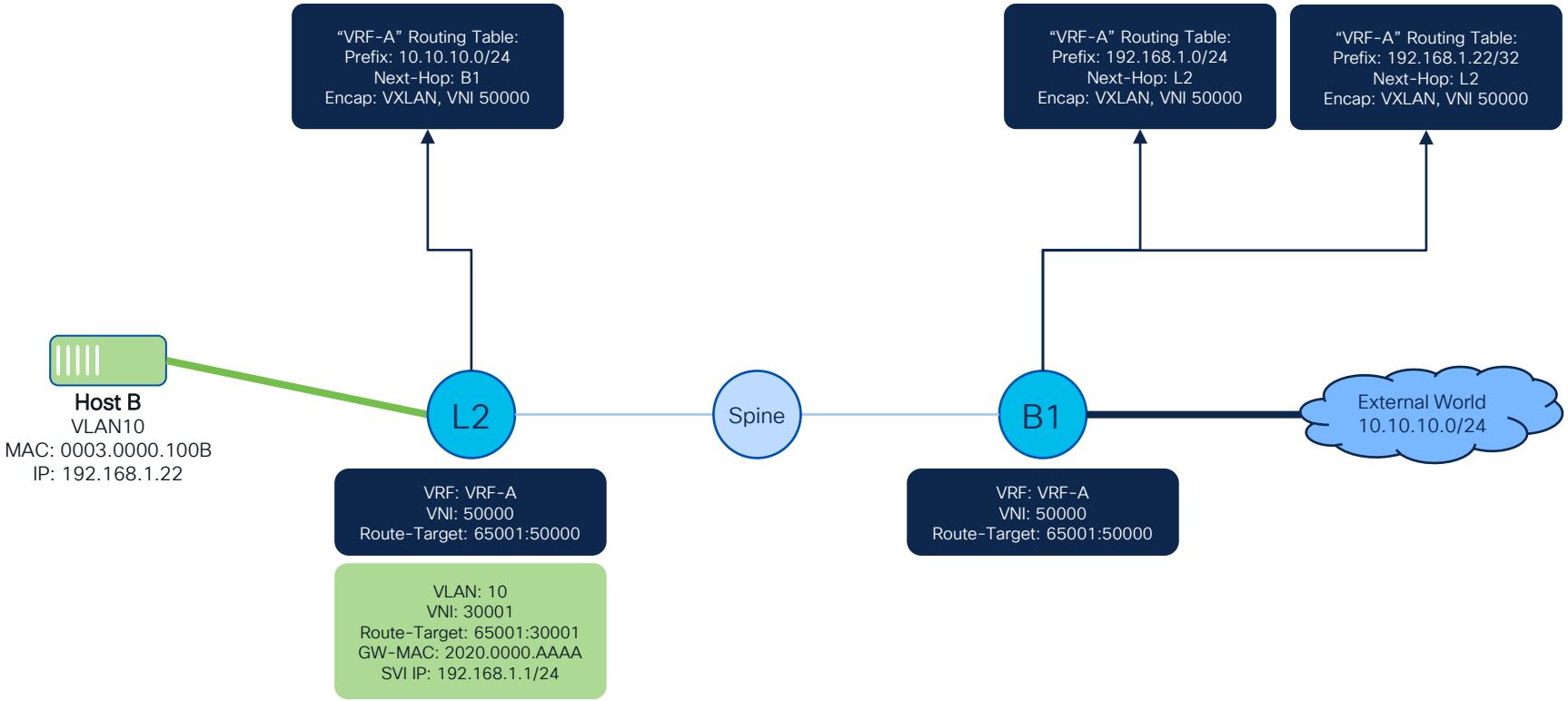
Learning: HostB to External World



Learning: HostB to External World

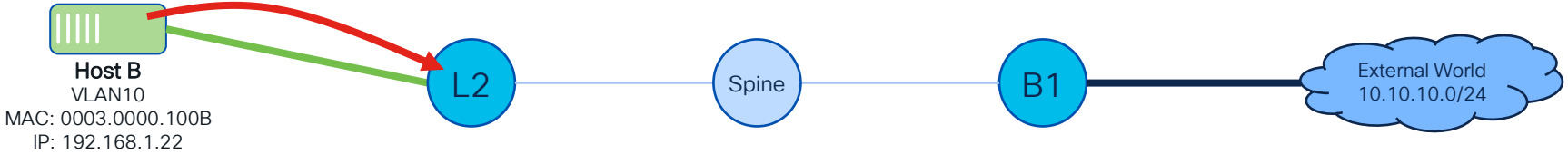


Overview: Forwarding Tables



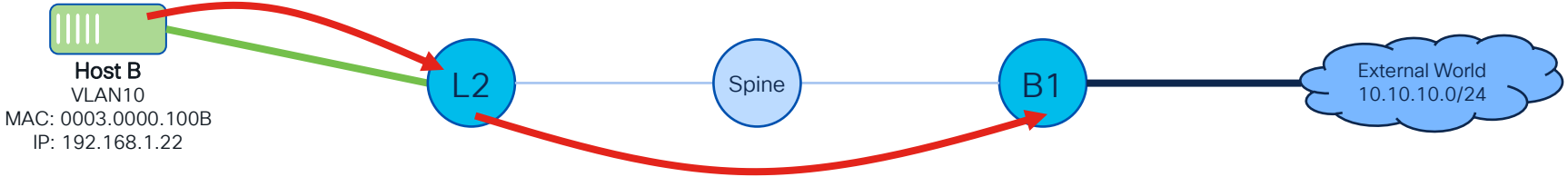
HostB to External World

SMAC	DMAC	VLAN	SIP	DIP	Payload
0003.0000.100B	2020.0000.AAAA	10	192.168.1.22	10.10.10.77	



HostB to External World

SMAC	DMAC	VLAN	SIP	DIP	Payload
0003.0000.100B	2020.0000.AAAA	10	192.168.1.22	10.10.10.77	

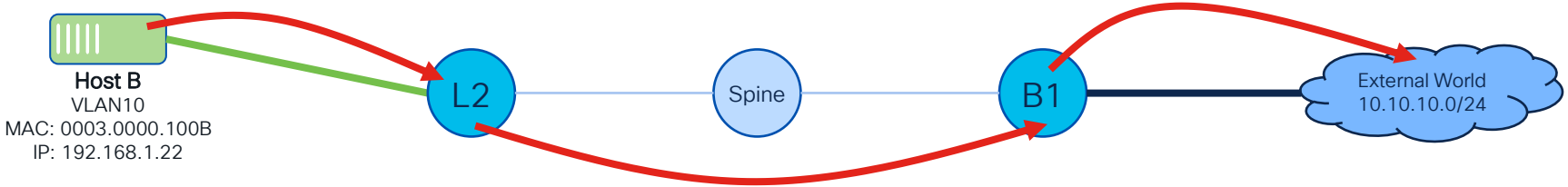


SIP	DIP	VNI	SMAC	DMAC	SIP	DIP	Payload
L2-IP	B1-IP	50000	L2-RMAC	B1-RMAC	192.168.1.22	10.10.10.77	

HostB to External World

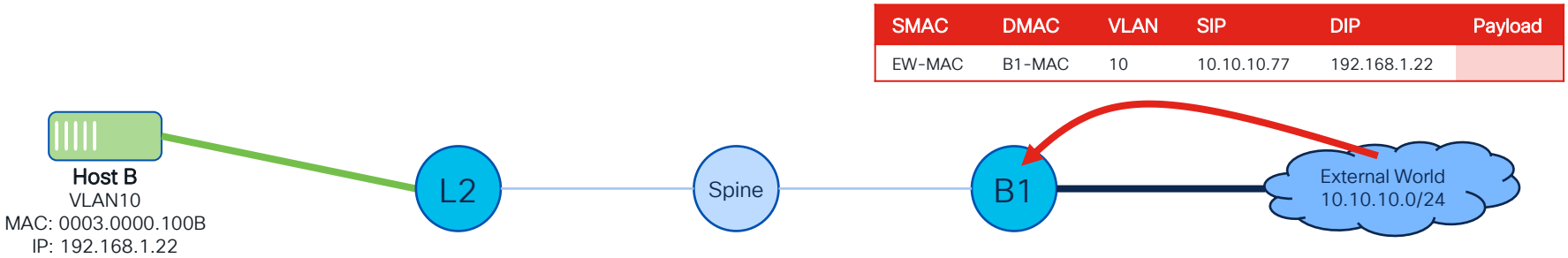
SMAC	DMAC	VLAN	SIP	DIP	Payload
0003.0000.100B	2020.0000.AAAA	10	192.168.1.22	10.10.10.77	

SMAC	DMAC	VLAN	SIP	DIP	Payload
B1-MAC	EW-MAC	10	192.168.1.22	10.10.10.77	

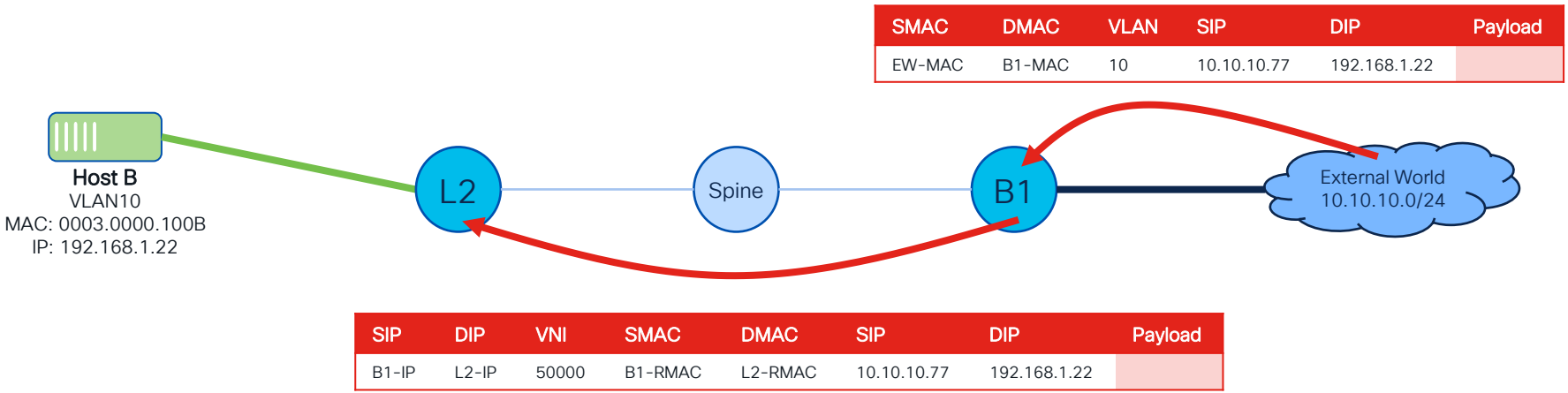


SIP	DIP	VNI	SMAC	DMAC	SIP	DIP	Payload
L2-IP	B1-IP	50000	L2-RMAC	B1-RMAC	192.168.1.22	10.10.10.77	

External World to HostB



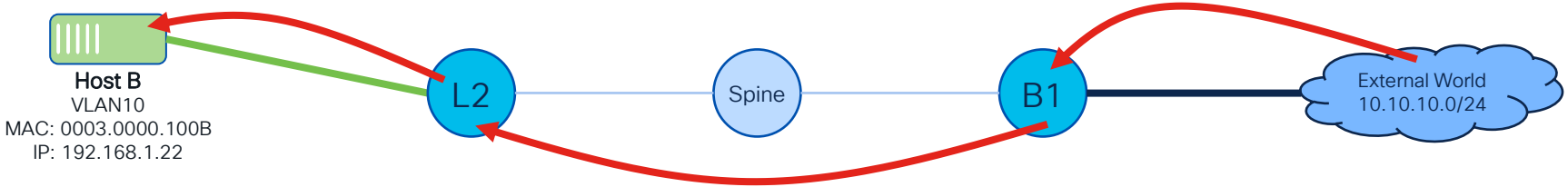
External World to HostB



External World to HostB

SMAC	DMAC	VLAN	SIP	DIP	Payload
2020.0000.AAAA	0003.0000.100B	10	10.10.10.77	192.168.1.22	

SMAC	DMAC	VLAN	SIP	DIP	Payload
EW-MAC	B1-MAC	10	10.10.10.77	192.168.1.22	



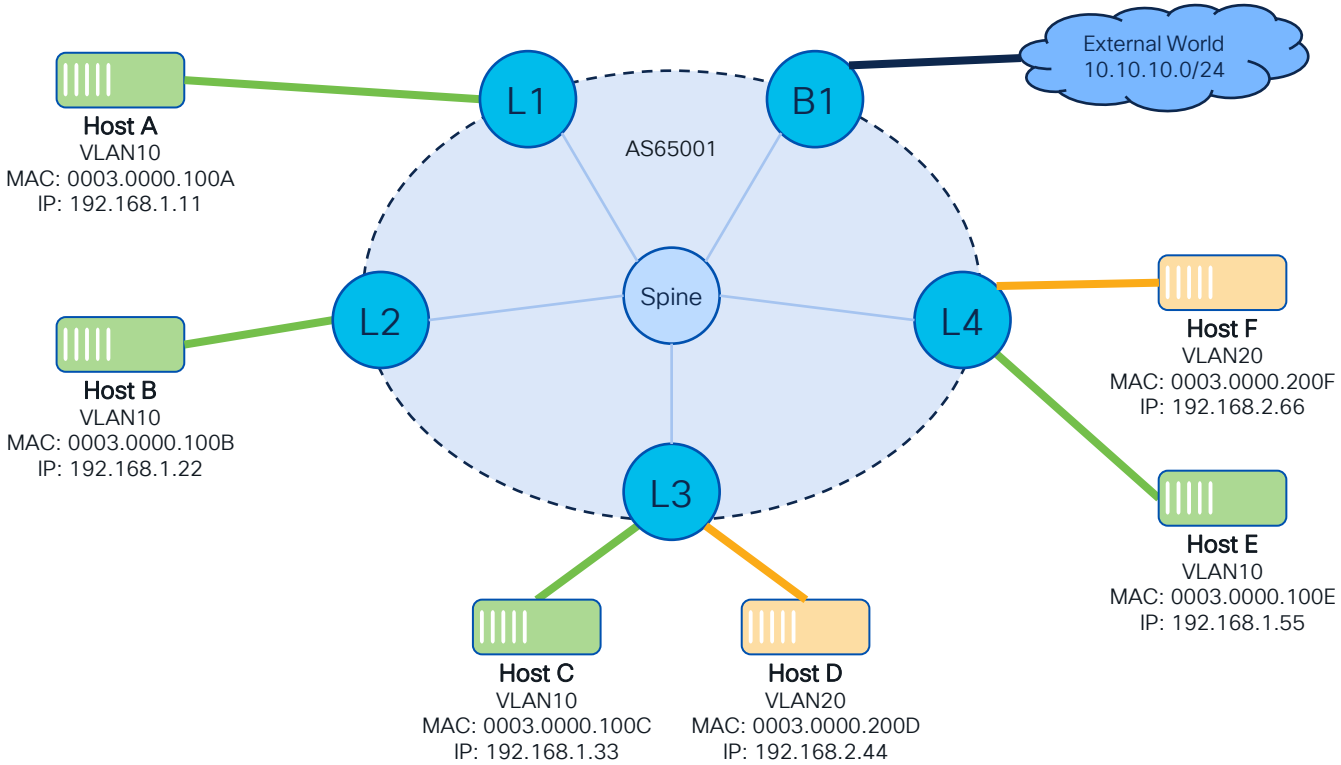
SIP	DIP	VNI	SMAC	DMAC	SIP	DIP	Payload
B1-IP	L2-IP	50000	B1-RMAC	L2-RMAC	10.10.10.77	192.168.1.22	

How to Talk to the Rest of the World - External Connectivity for VXLAN EVPN Fabrics

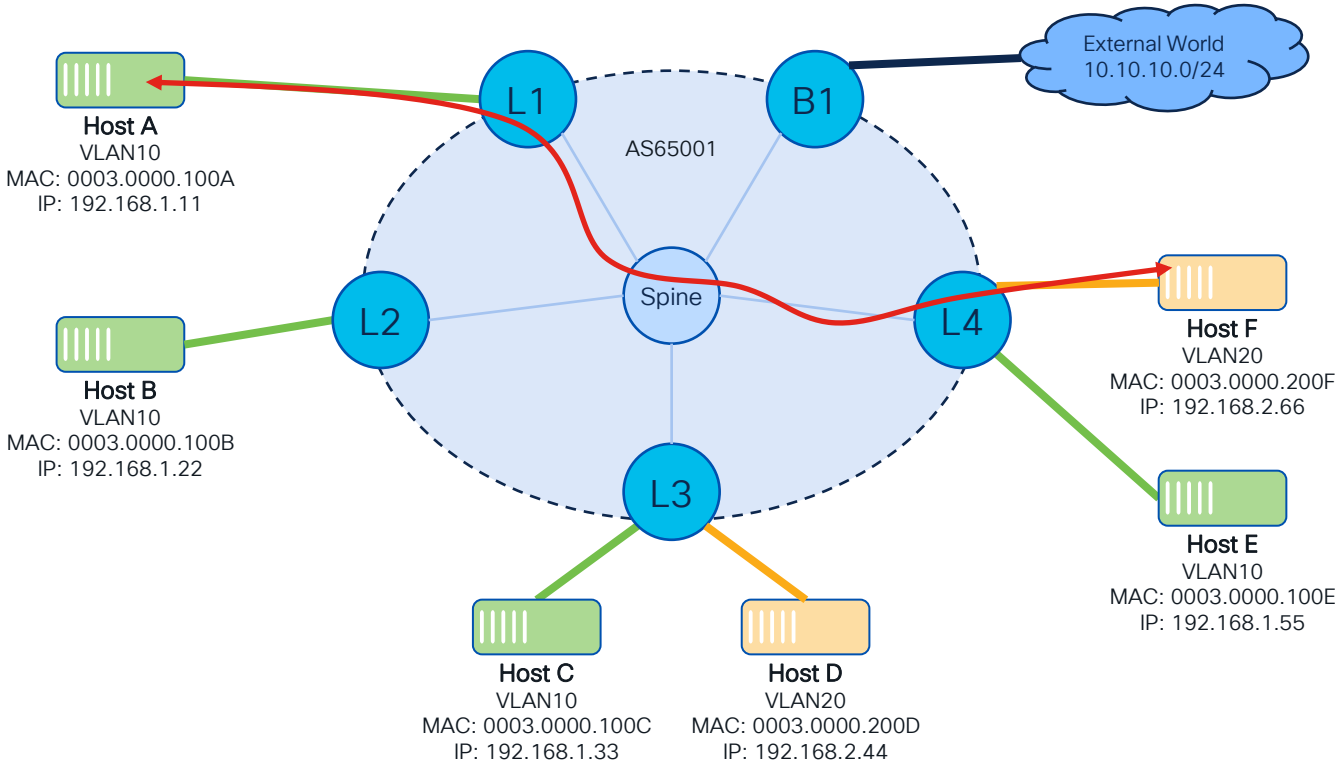
BRKDCN-2267

Packet Walk: Layer-3 – Host to Host

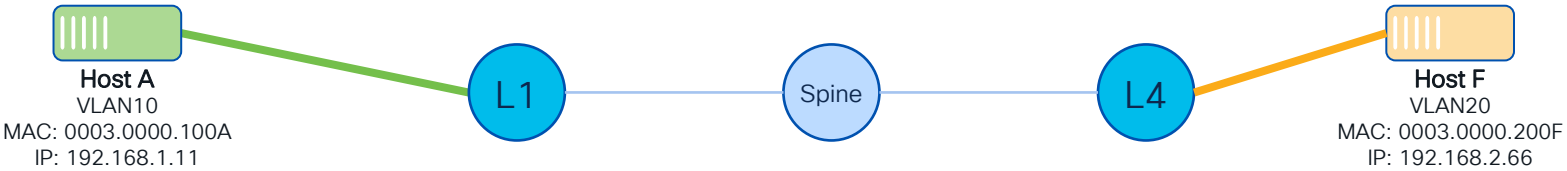
Topology Overview
Layer-3 Packet Walk



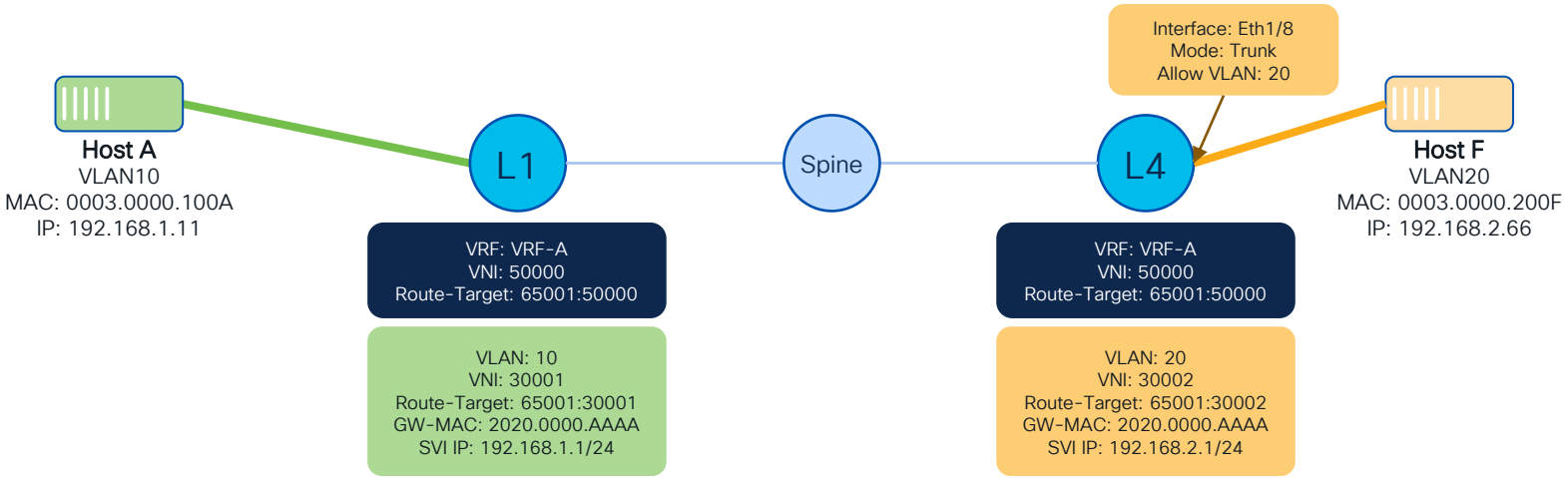
Topology Overview
Layer-3 Packet Walk



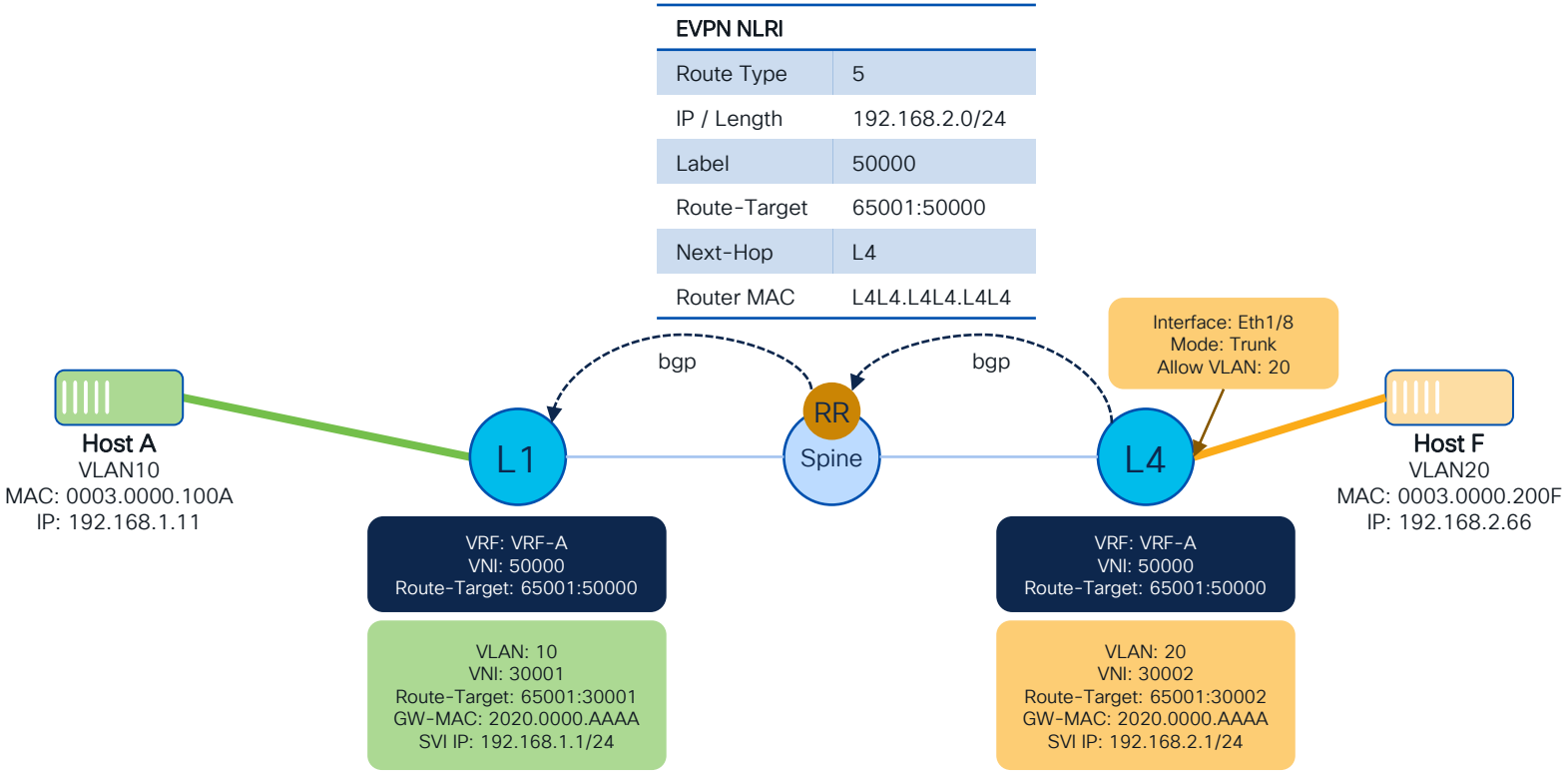
Learning: HostF to Leaf1



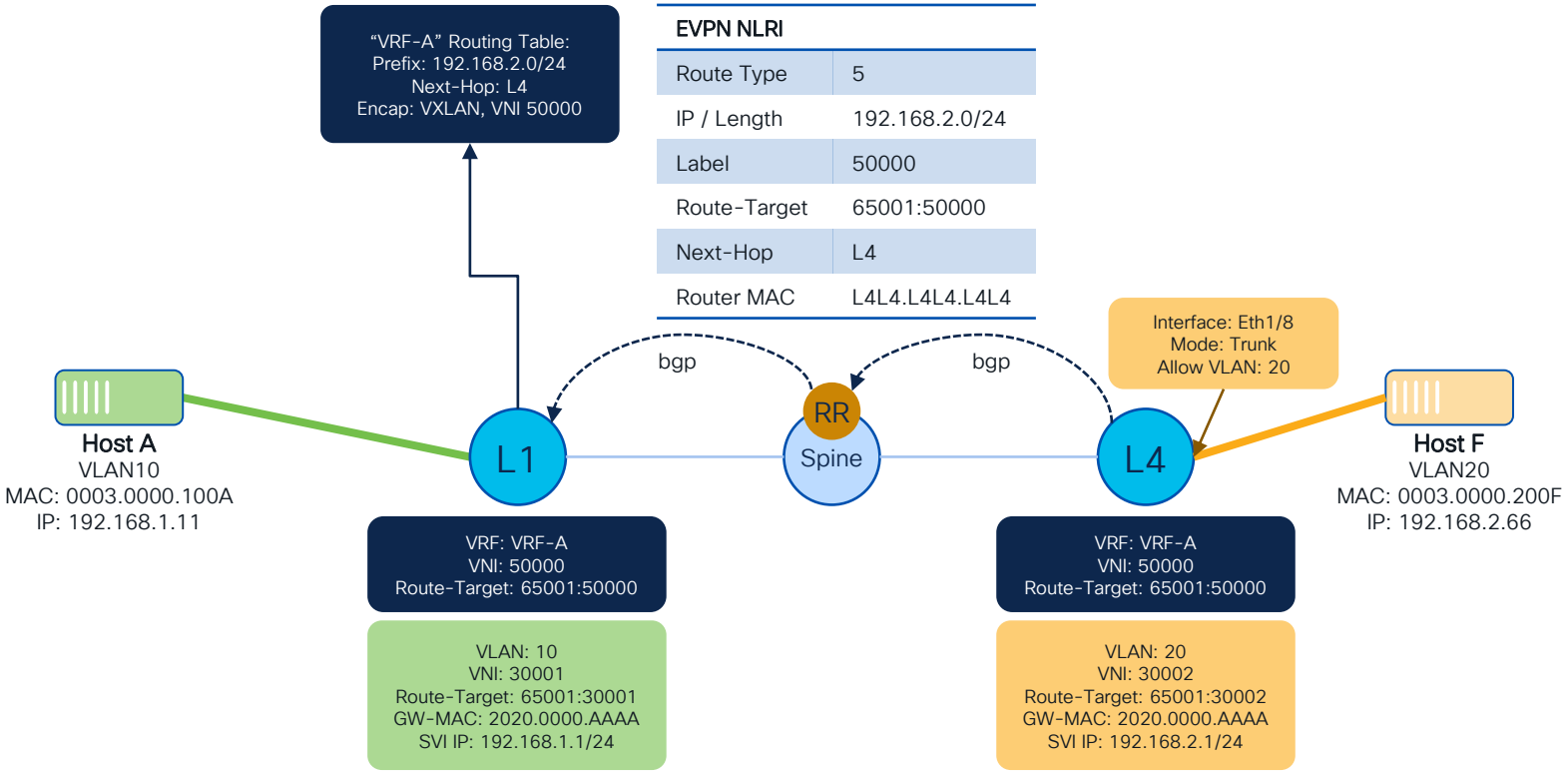
Learning: HostF to Leaf1



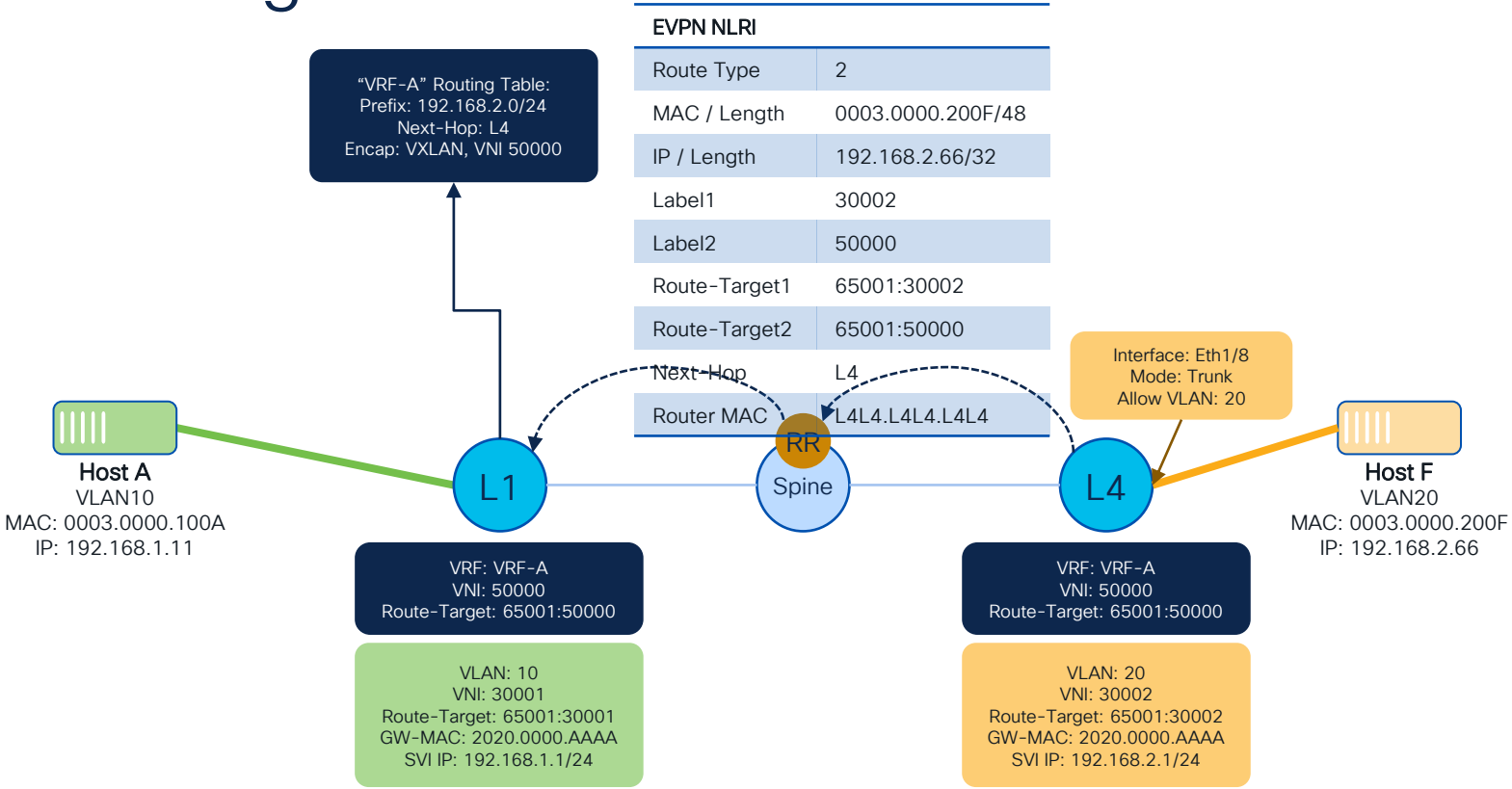
Learning: HostF to Leaf1



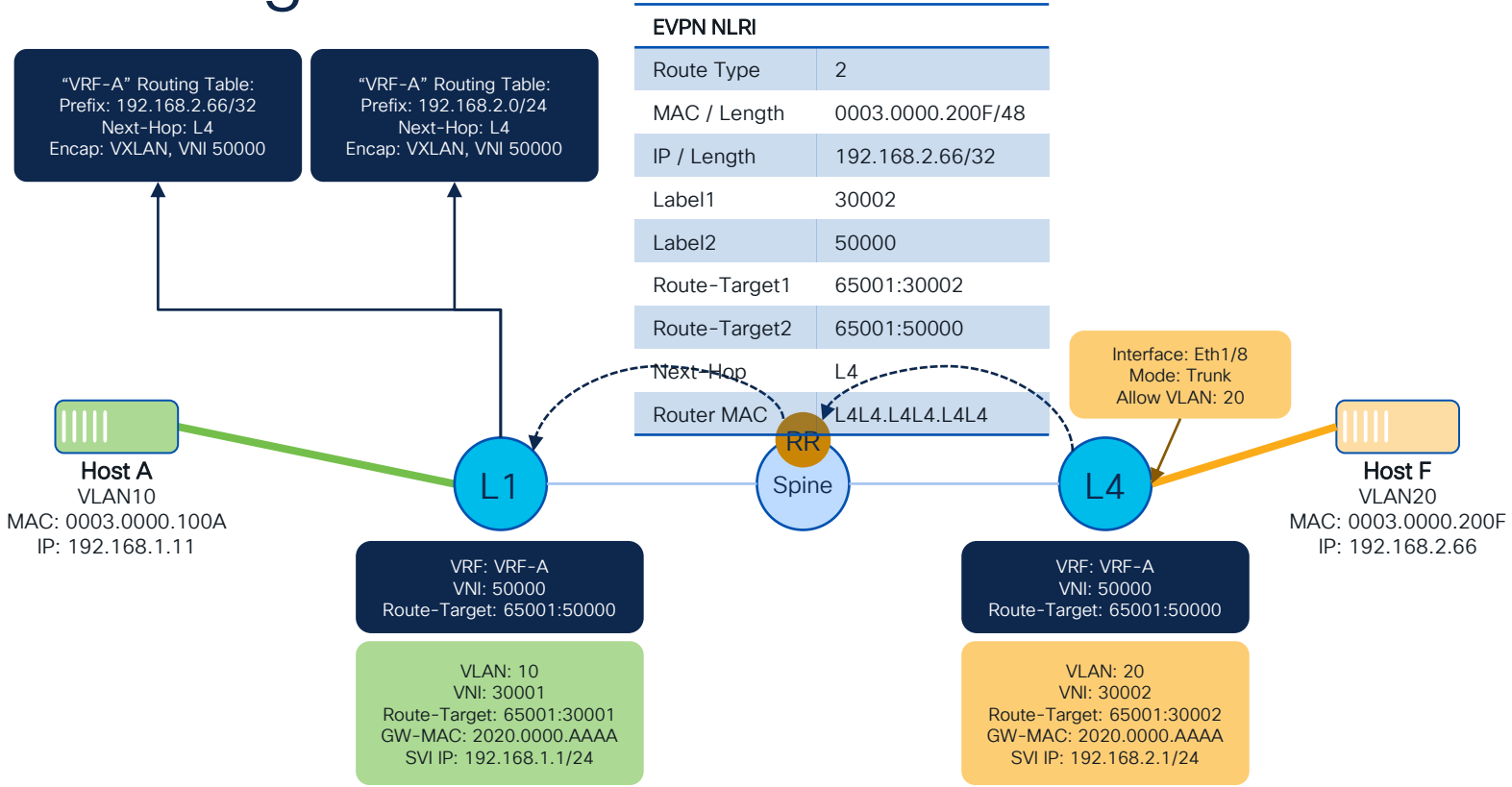
Learning: HostF to Leaf1



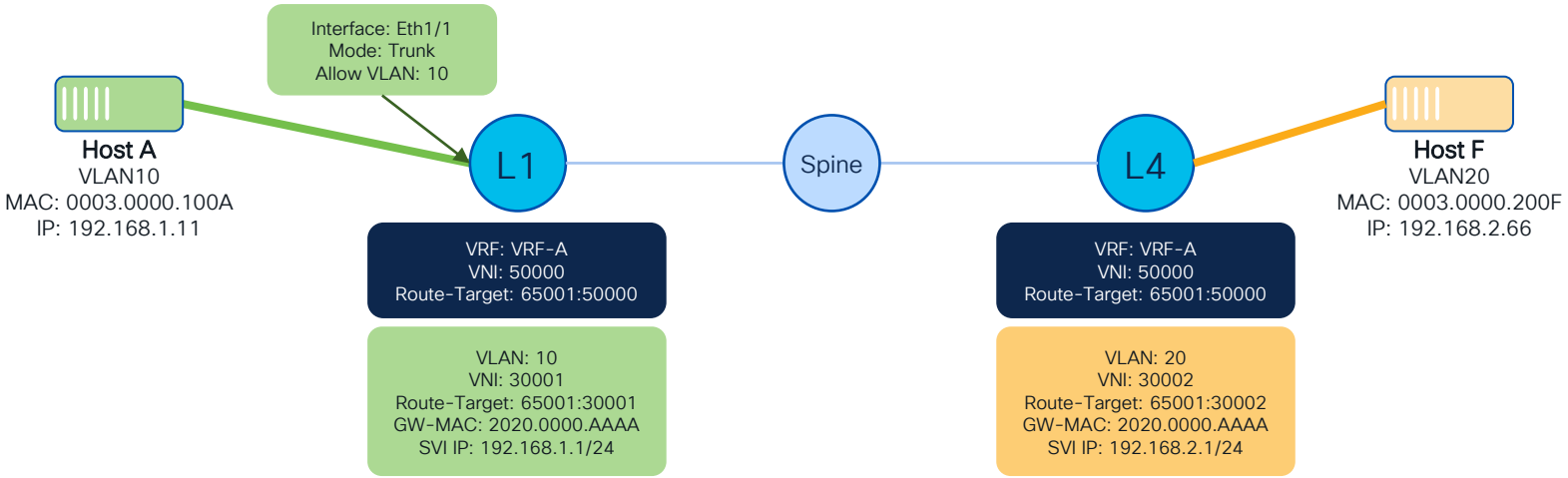
Learning: HostF to Leaf1



Learning: HostF to Leaf1

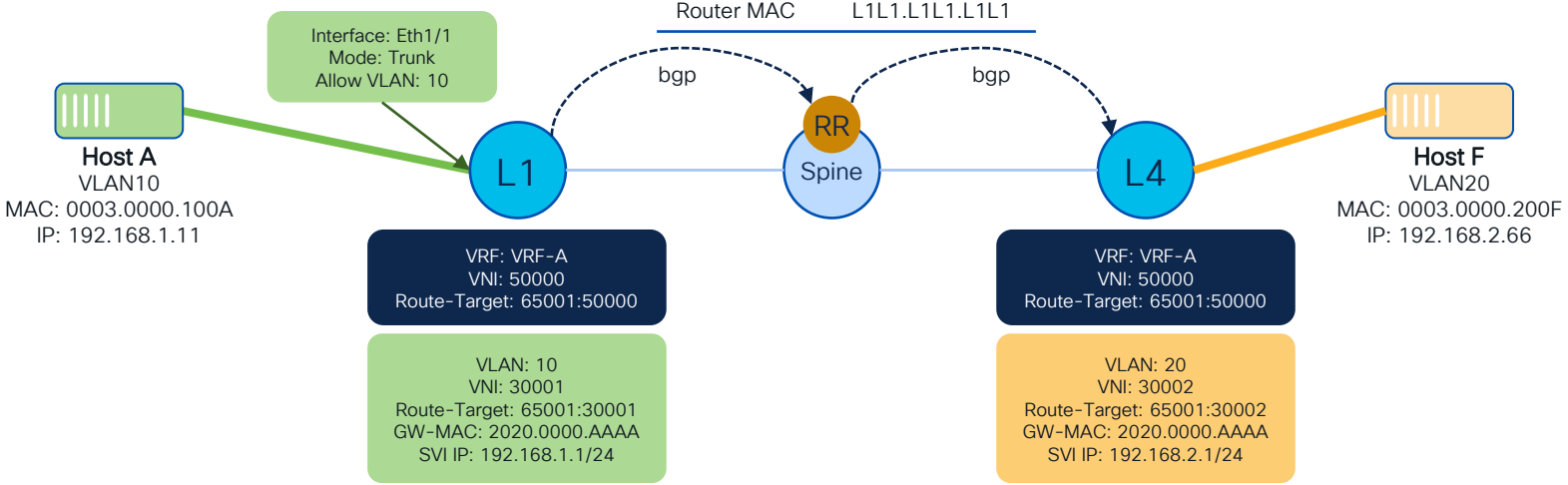


Learning: HostA to Leaf4

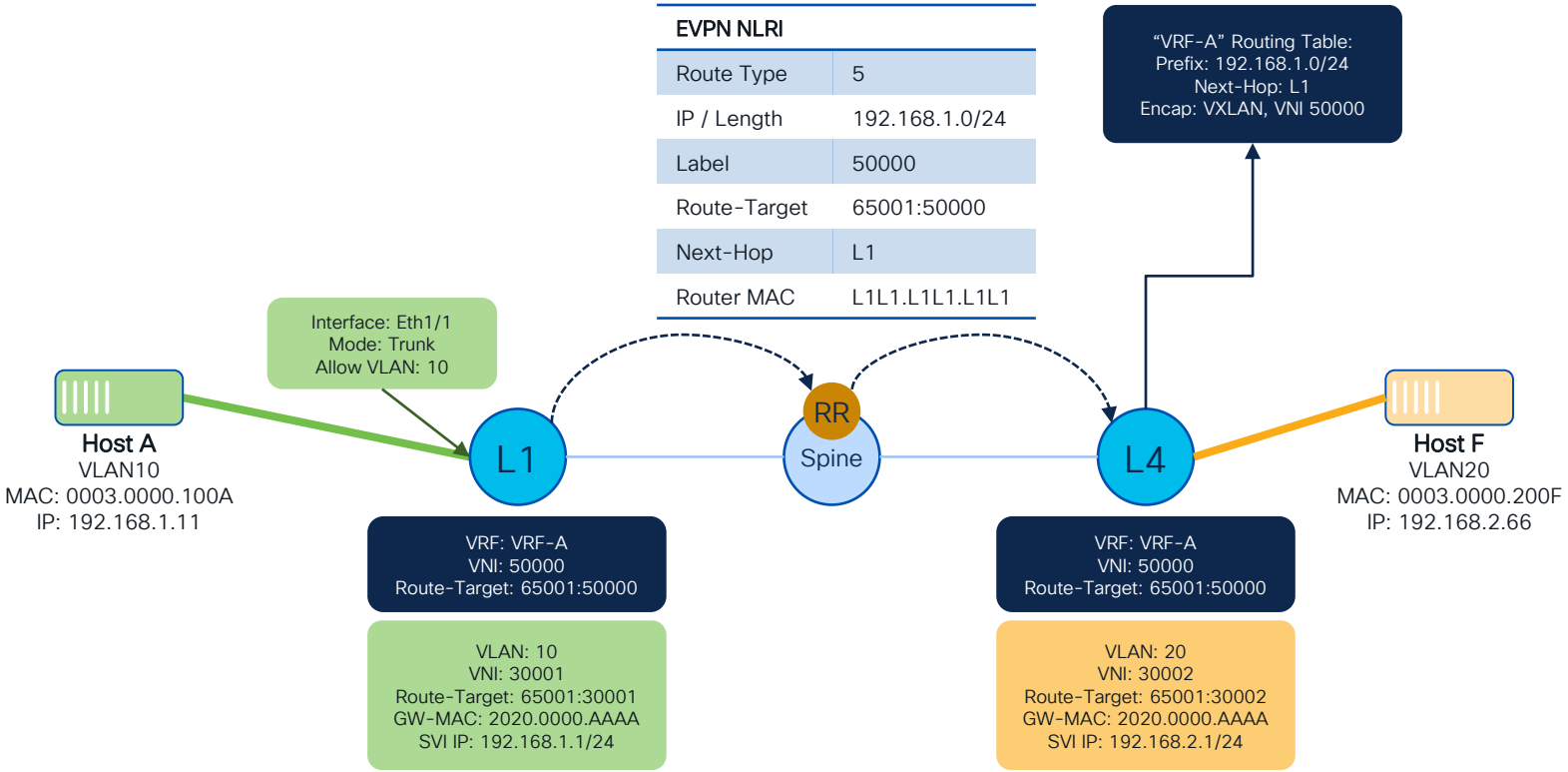


Learning: HostA to Leaf4

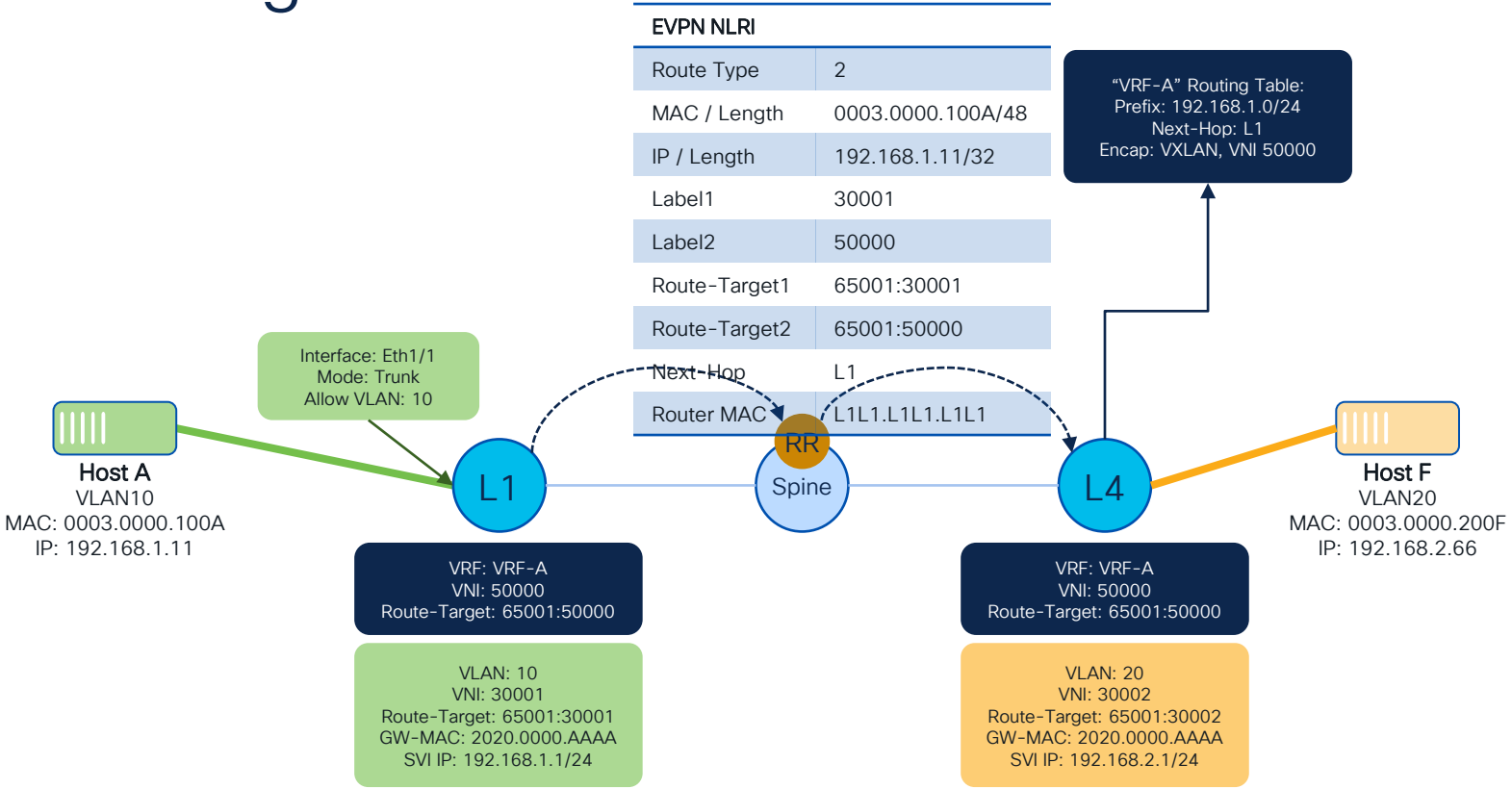
EVPN NLRI	
Route Type	5
IP / Length	192.168.1.0/24
Label	50000
Route-Target	65001:50000
Next-Hop	L1
Router MAC	L1L1.L1L1.L1L1



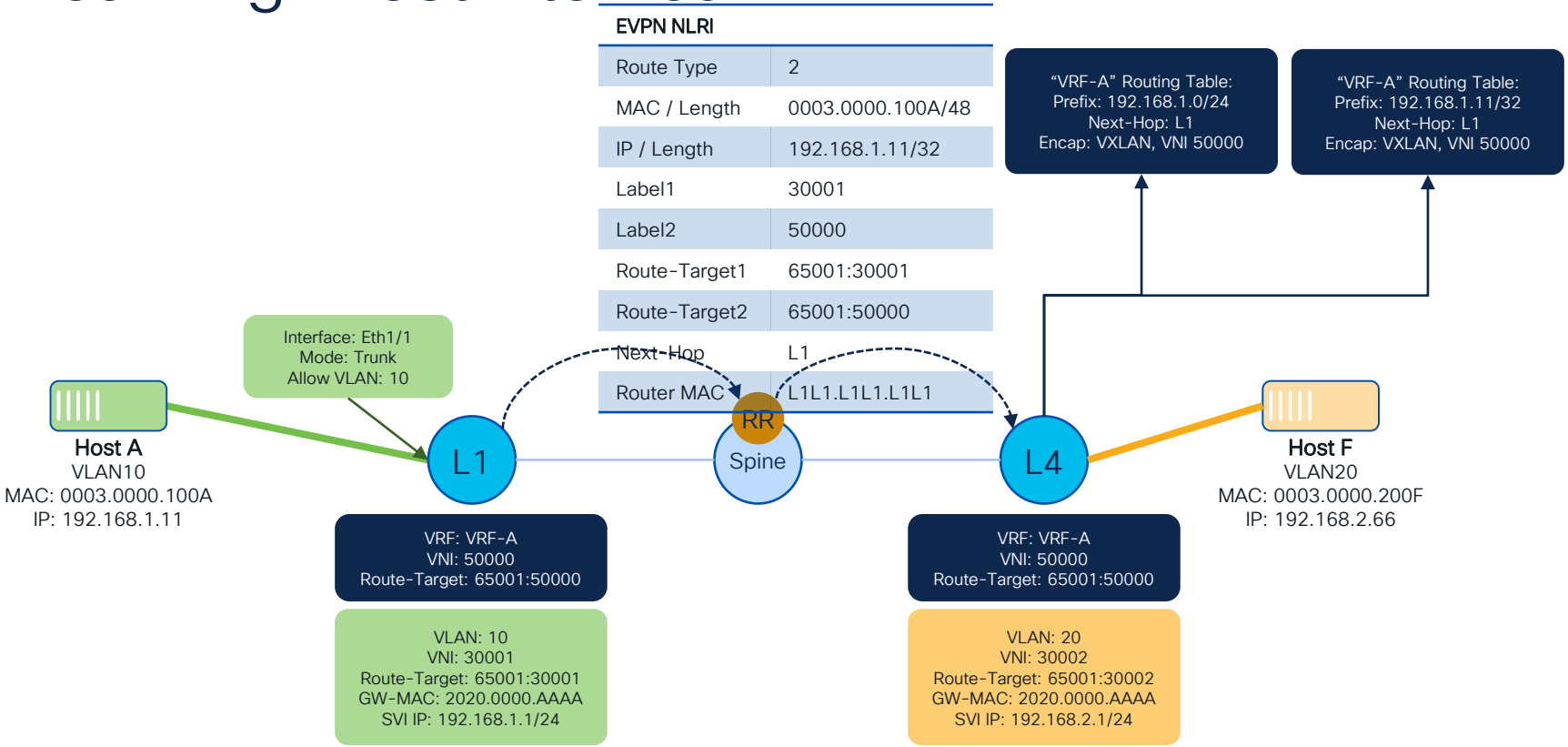
Learning: HostA to Leaf4



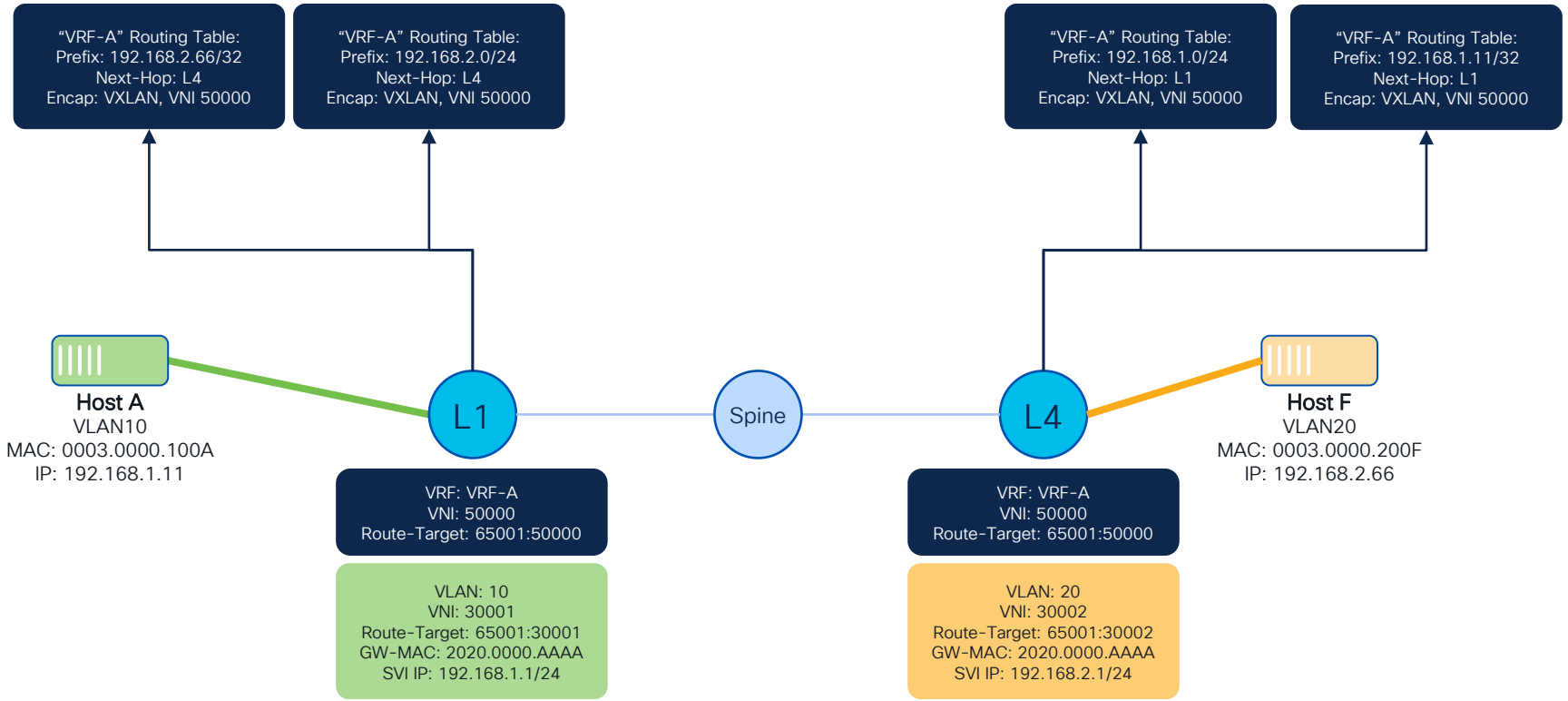
Learning: HostA to Leaf4



Learning: HostA to Leaf4

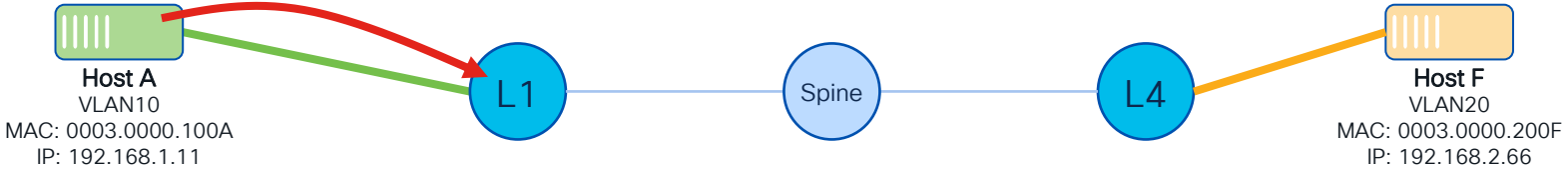


Overview: Forwarding Tables



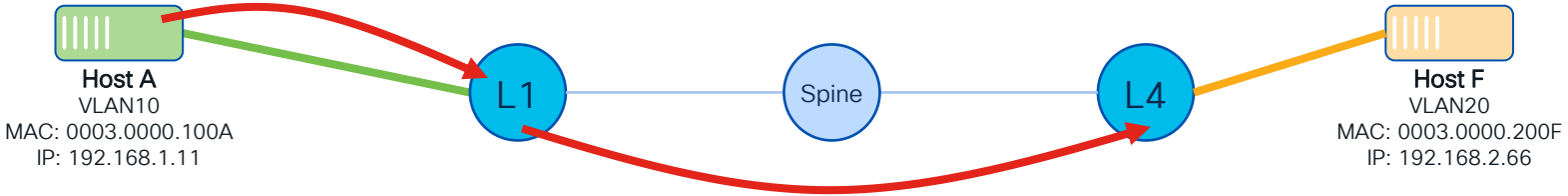
HostA to HostF

SMAC	DMAC	VLAN	SIP	DIP	Payload
0003.0000.100A	2020.0000.AAAA	10	192.168.1.11	192.168.2.66	



HostA to HostF

SMAC	DMAC	VLAN	SIP	DIP	Payload
0003.0000.100A	2020.0000.AAAA	10	192.168.1.11	192.168.2.66	

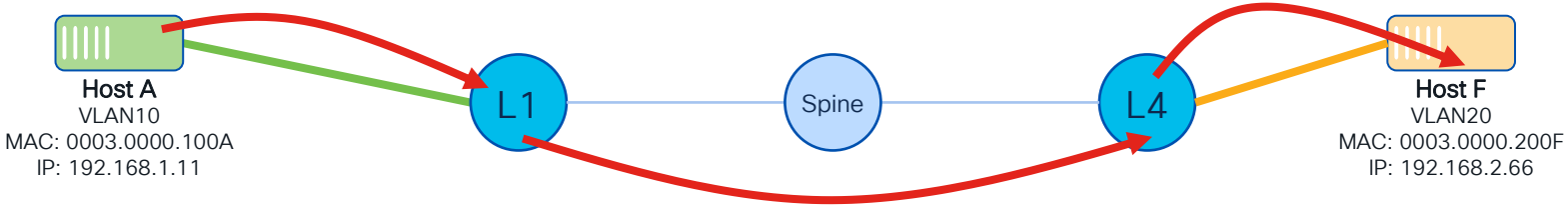


SIP	DIP	VNI	SMAC	DMAC	SIP	DIP	Payload
L1-IP	L4-IP	50000	L1-RMAC	L4-RMAC	192.168.1.11	192.168.2.66	

HostA to HostF

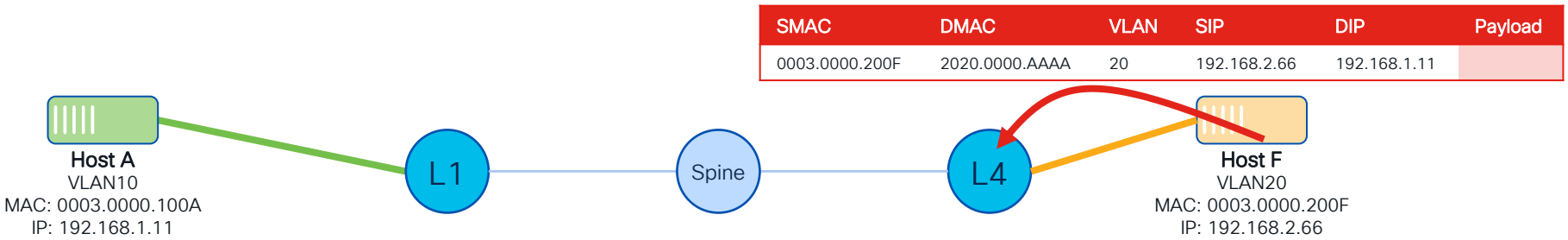
SMAC	DMAC	VLAN	SIP	DIP	Payload
0003.0000.100A	2020.0000.AAAA	10	192.168.1.11	192.168.2.66	

SMAC	DMAC	VLAN	SIP	DIP	Payload
2020.0000.AAAA	0003.0000.200F	20	192.168.1.11	192.168.2.66	

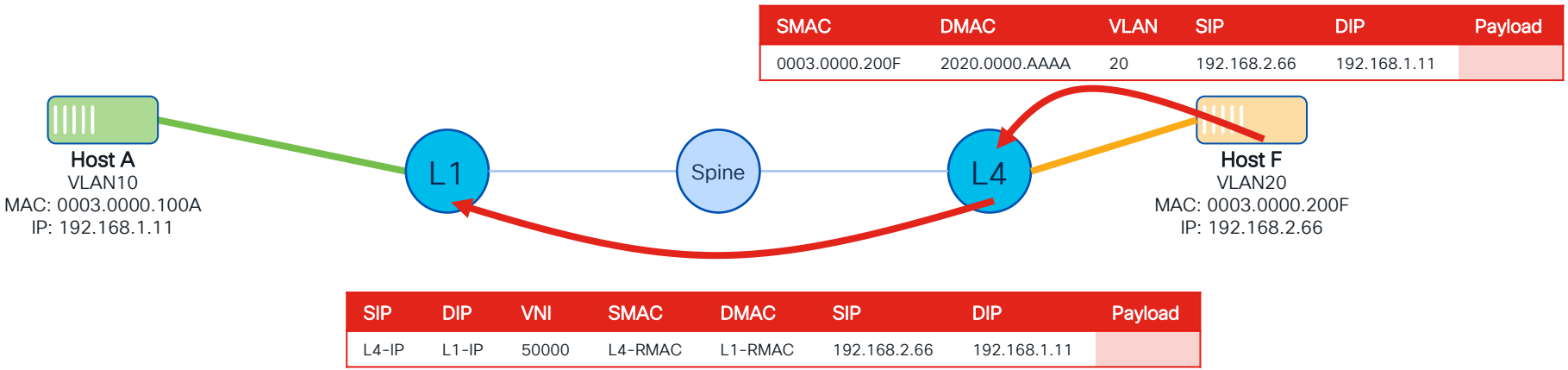


SIP	DIP	VNI	SMAC	DMAC	SIP	DIP	Payload
L1-IP	L4-IP	50000	L1-RMAC	L4-RMAC	192.168.1.11	192.168.2.66	

HostF to HostA



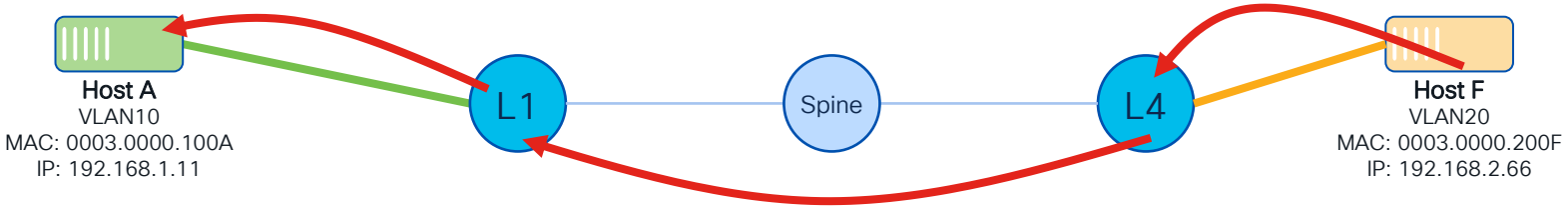
HostF to HostA



HostF to HostA

SMAC	DMAC	VLAN	SIP	DIP	Payload
2020.0000.AAAA	0003.0000.100A	10	192.168.2.66	192.168.1.11	

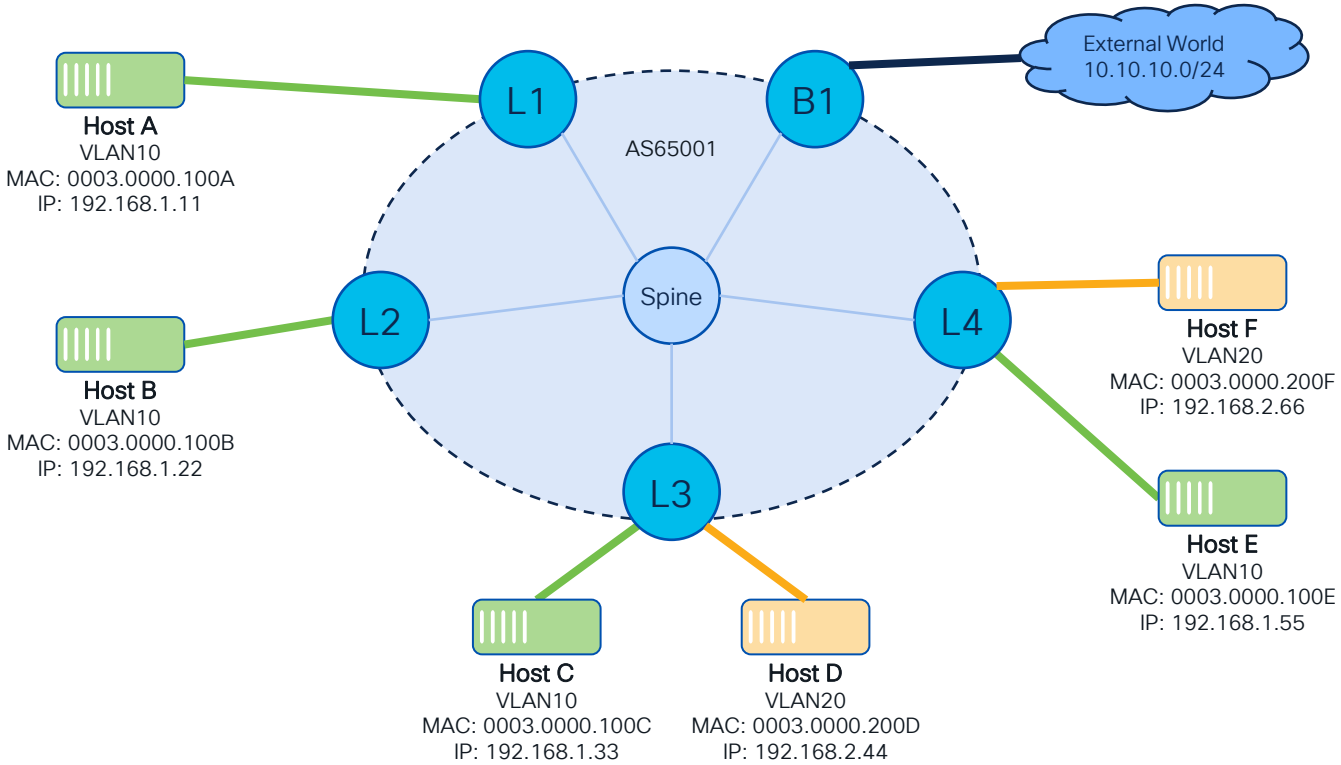
SMAC	DMAC	VLAN	SIP	DIP	Payload
0003.0000.200F	2020.0000.AAAA	20	192.168.2.66	192.168.1.11	



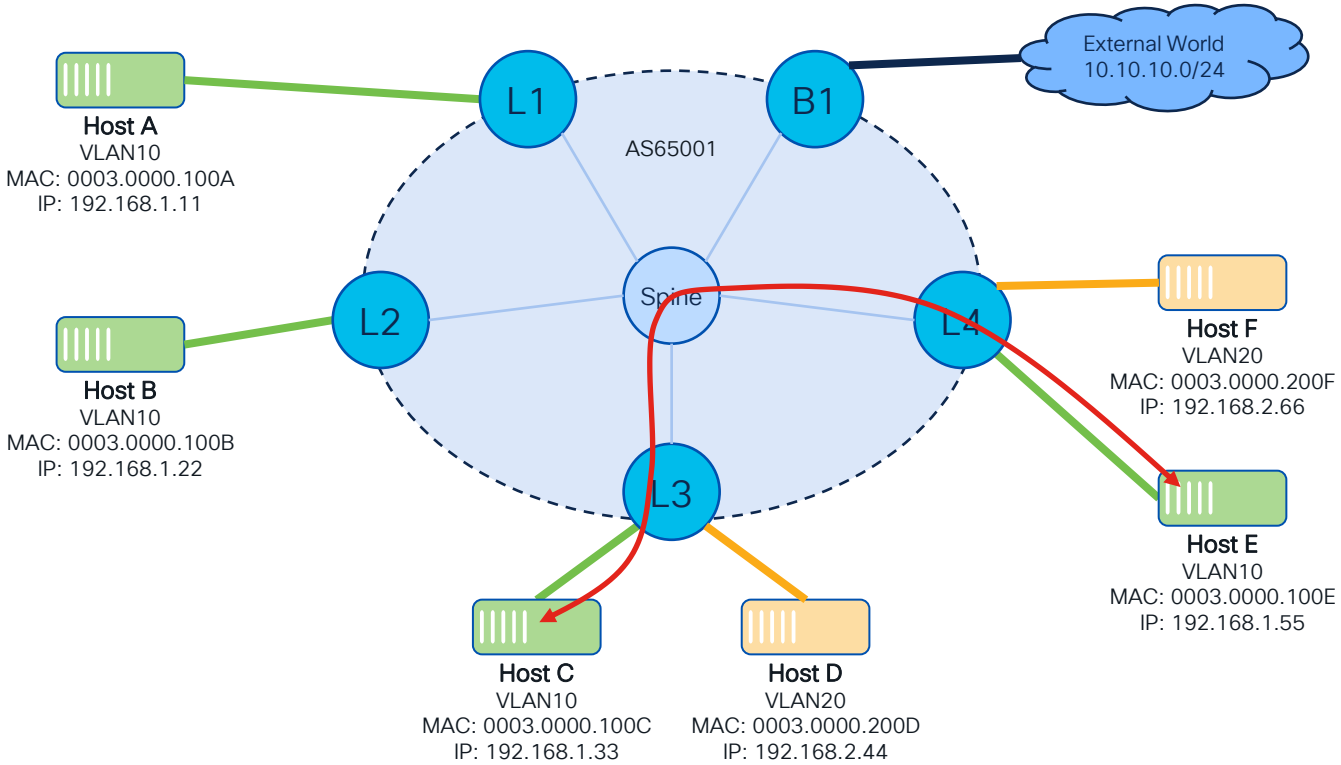
SIP	DIP	VNI	SMAC	DMAC	SIP	DIP	Payload
L4-IP	L1-IP	50000	L4-RMAC	L1-RMAC	192.168.2.66	192.168.1.11	

Packet Walk: Layer-2

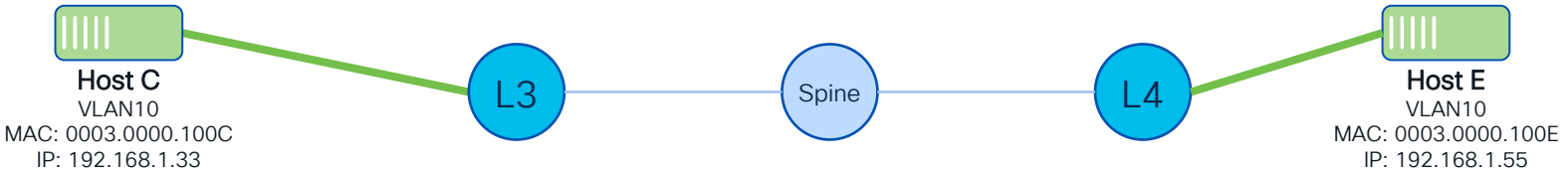
Topology Overview
Layer-2 Packet Walk



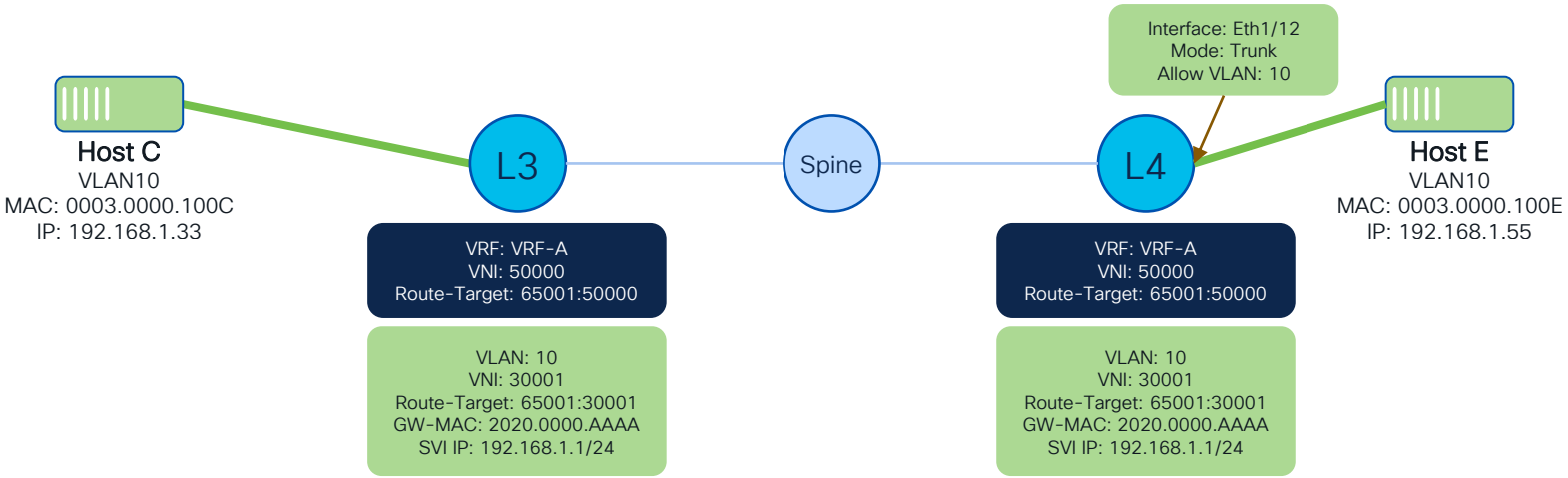
Topology Overview
Layer-2 Packet Walk



Learning: HostE to Leaf3

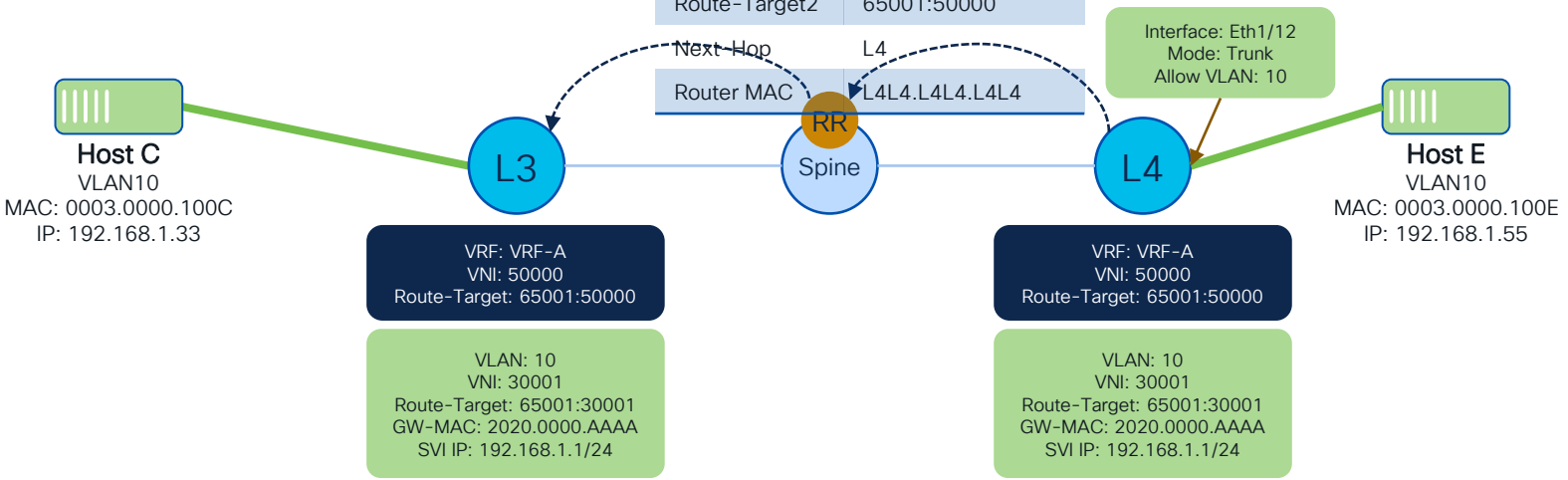


Learning: HostE to Leaf3

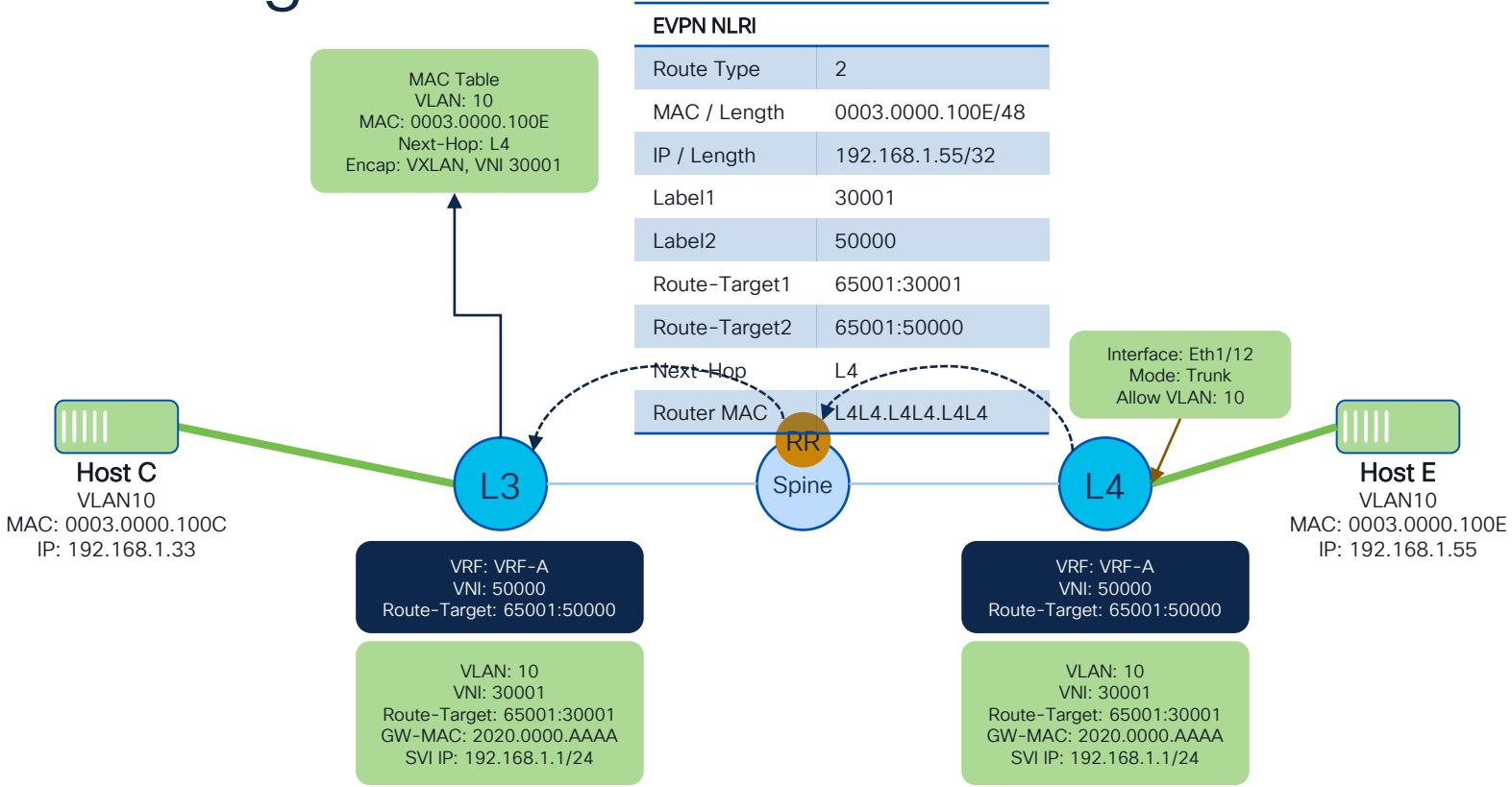


Learning: HostE to Leaf3

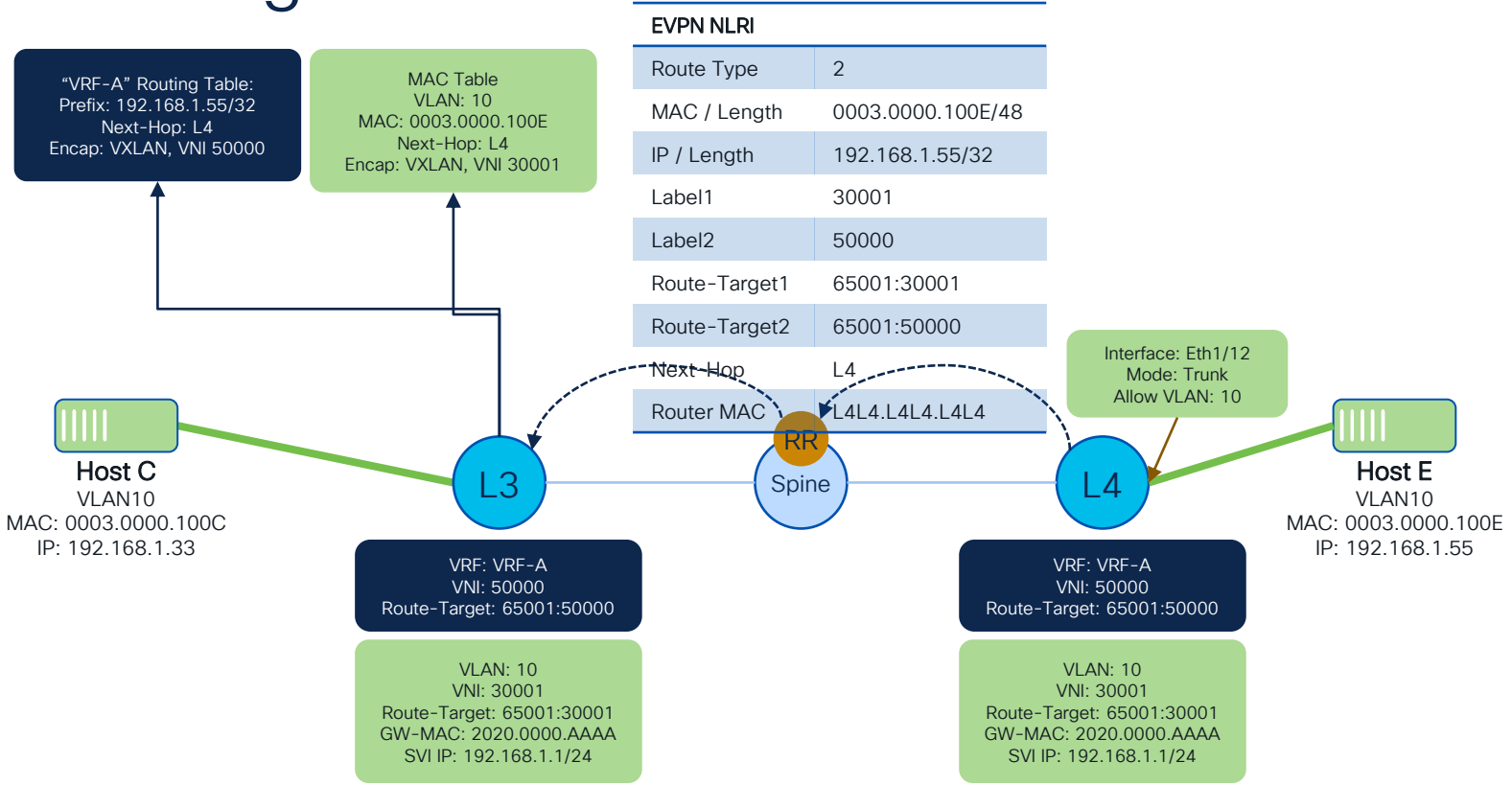
EVPN NLRI	
Route Type	2
MAC / Length	0003.0000.100E/48
IP / Length	192.168.1.55/32
Label1	30001
Label2	50000
Route-Target1	65001:30001
Route-Target2	65001:50000
Next-Hop	L4
Router MAC	L4L4.L4L4.L4L4



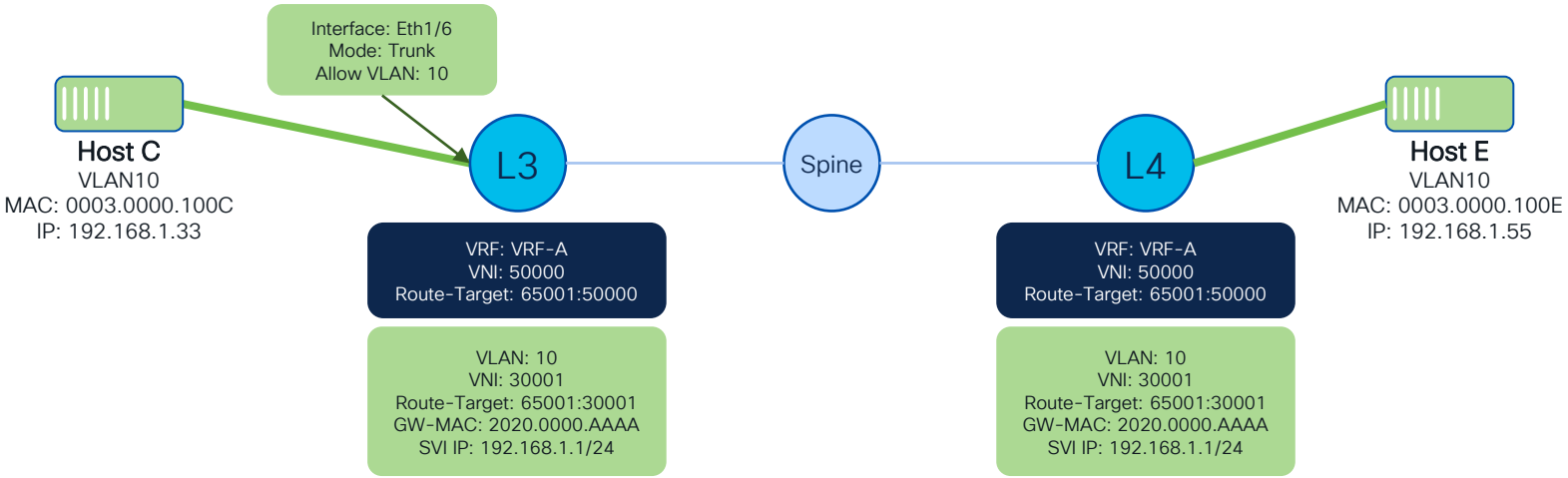
Learning: HostE to Leaf3



Learning: HostE to Leaf3

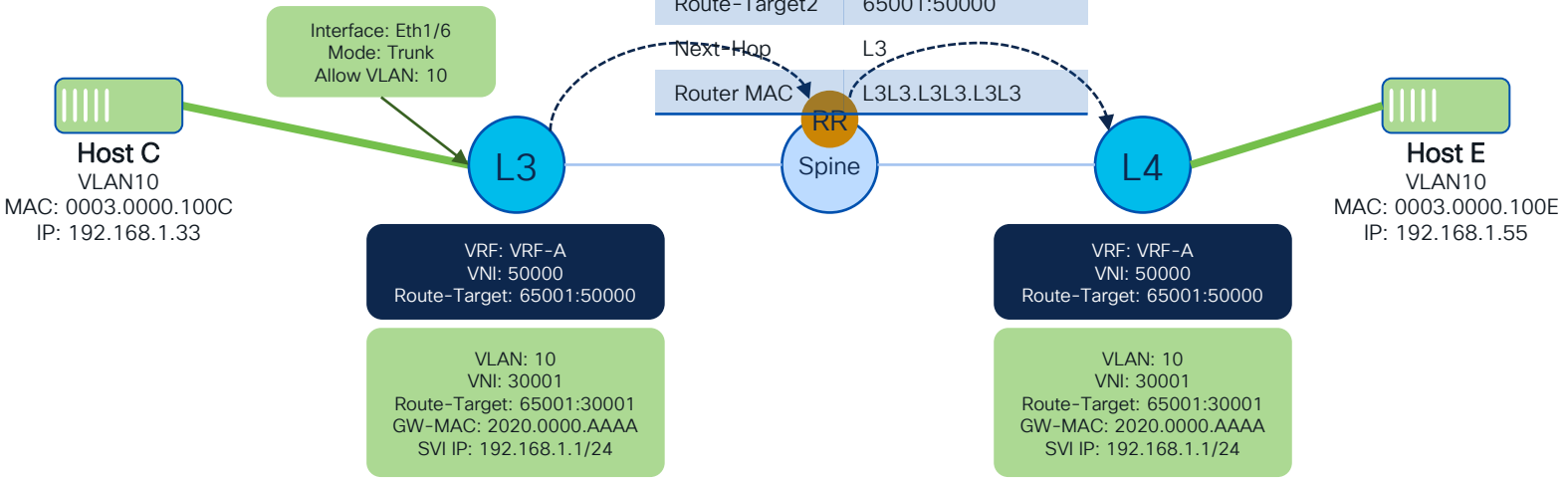


Learning: HostC to Leaf4

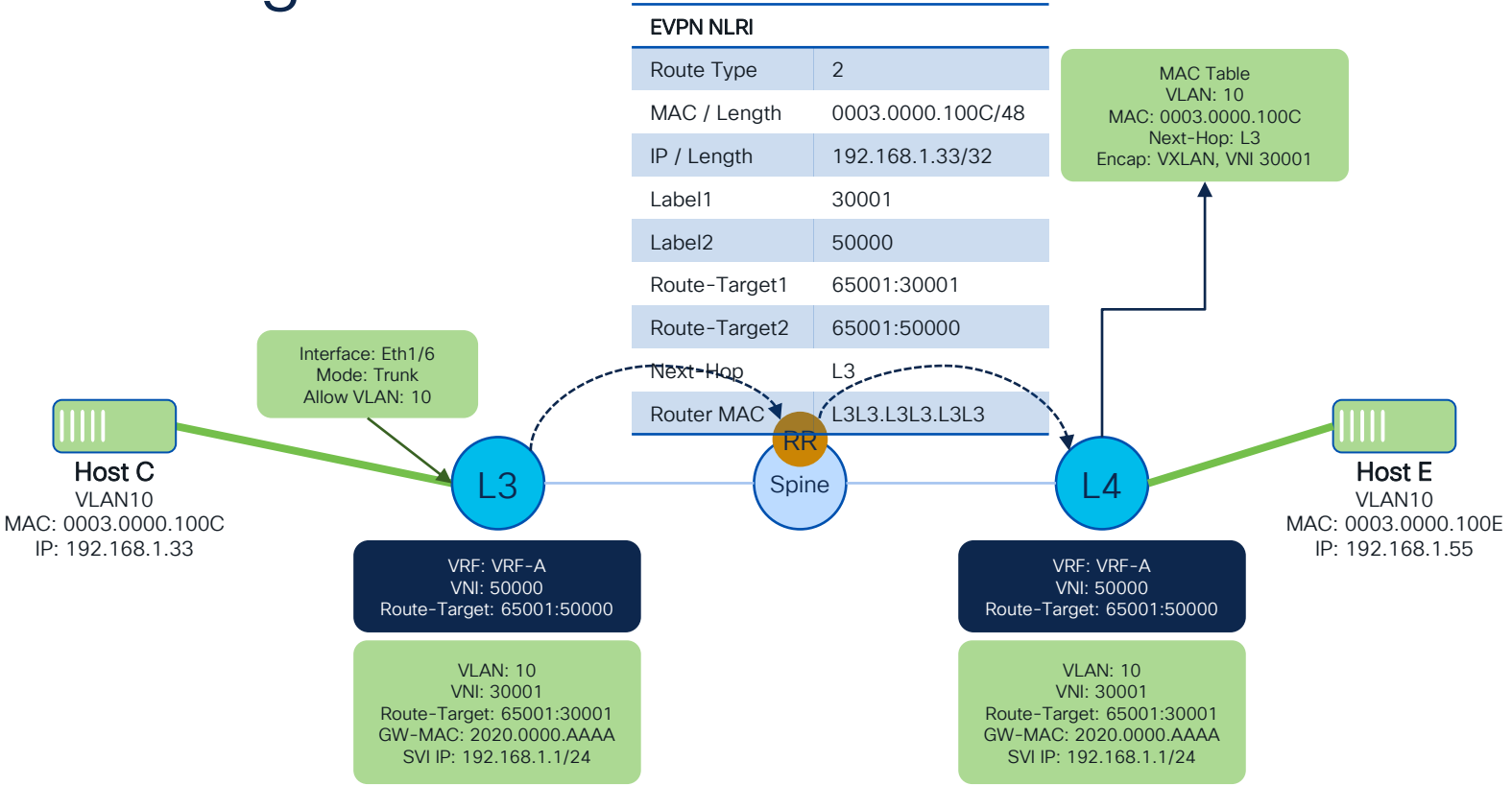


Learning: HostC to Leaf4

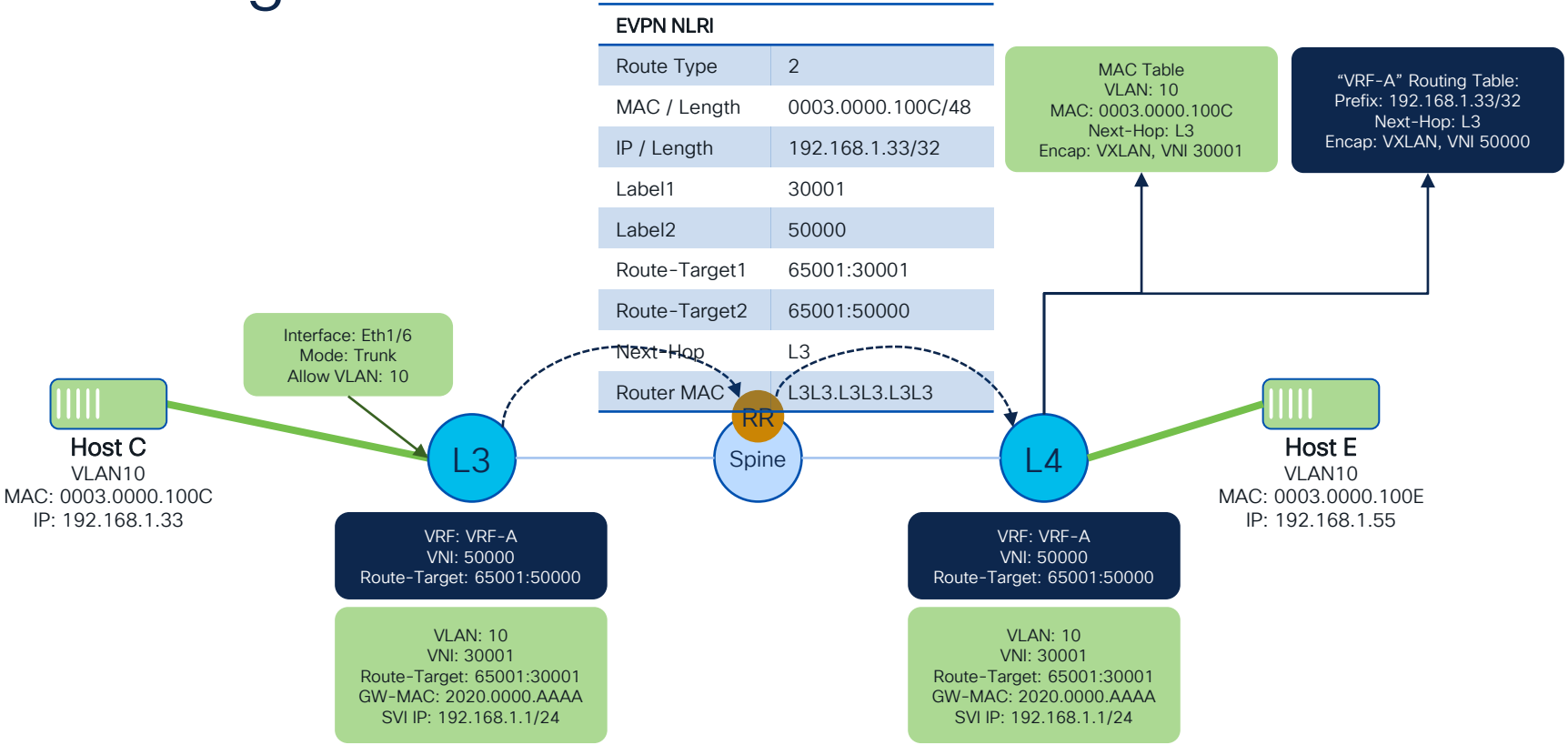
EVPN NLRI	
Route Type	2
MAC / Length	0003.0000.100C/48
IP / Length	192.168.1.33/32
Label1	30001
Label2	50000
Route-Target1	65001:30001
Route-Target2	65001:50000
Next-Hop	L3
Router MAC	L3L3.L3L3.L3L3



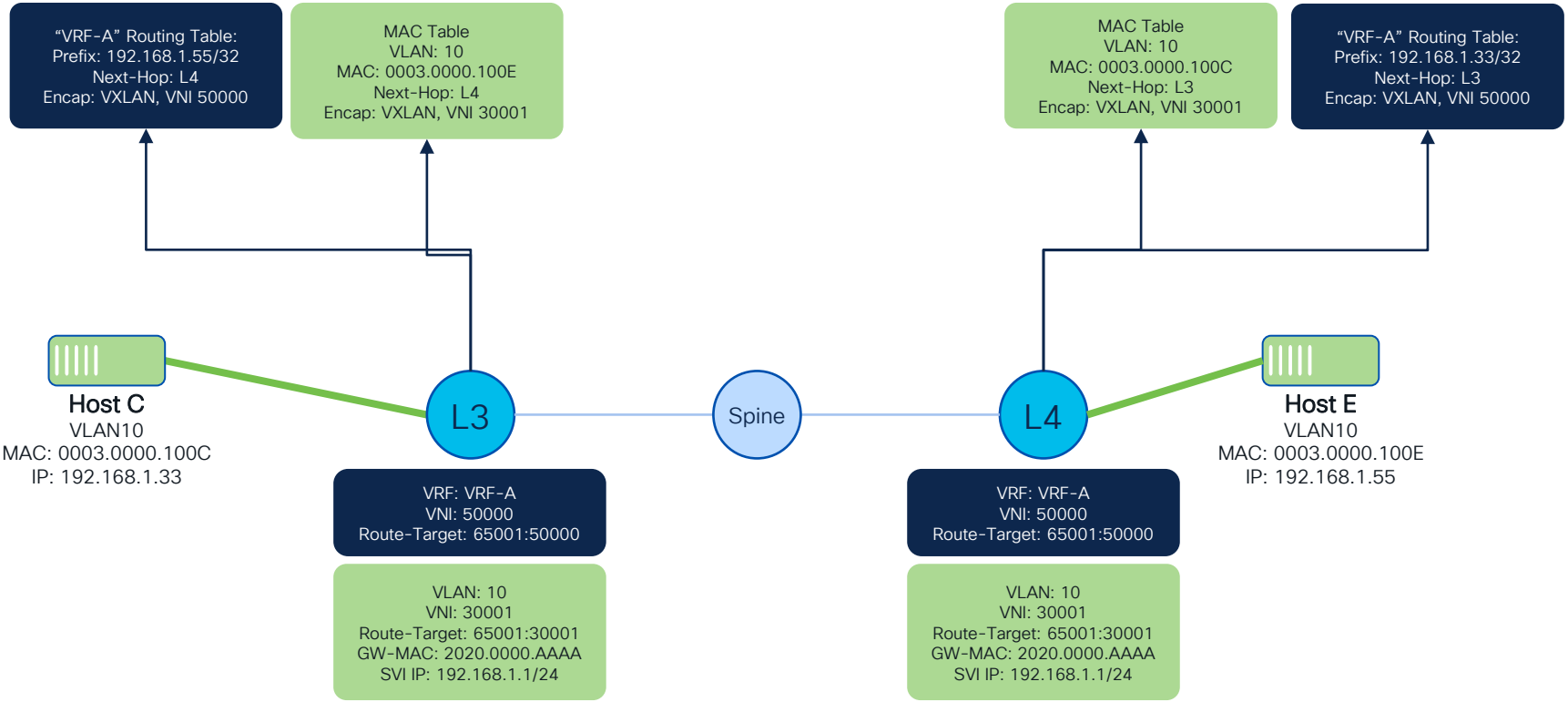
Learning: HostC to Leaf4



Learning: HostC to Leaf4

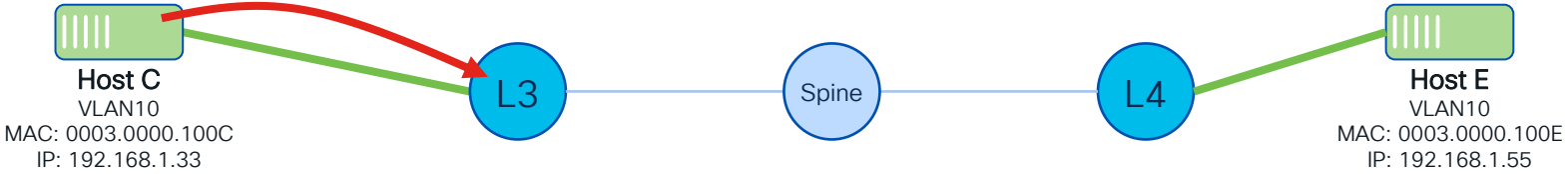


Forwarding Tables



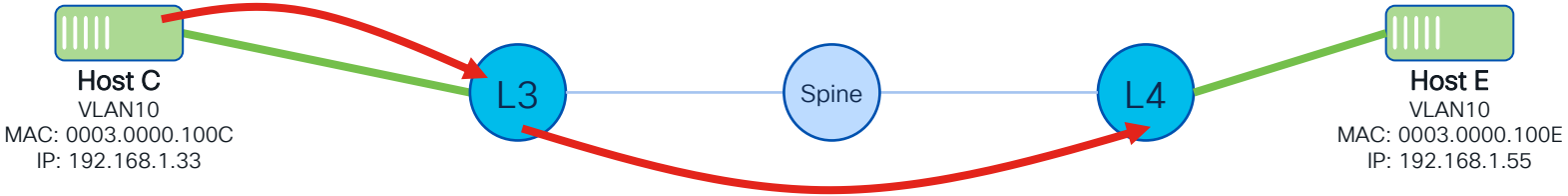
HostC to HostE

SMAC	DMAC	VLAN	SIP	DIP	Payload
0003.0000.100C	0003.0000.100E	10	192.168.1.33	192.168.1.55	



HostC to HostE

SMAC	DMAC	VLAN	SIP	DIP	Payload
0003.0000.100C	0003.0000.100E	10	192.168.1.33	192.168.1.55	

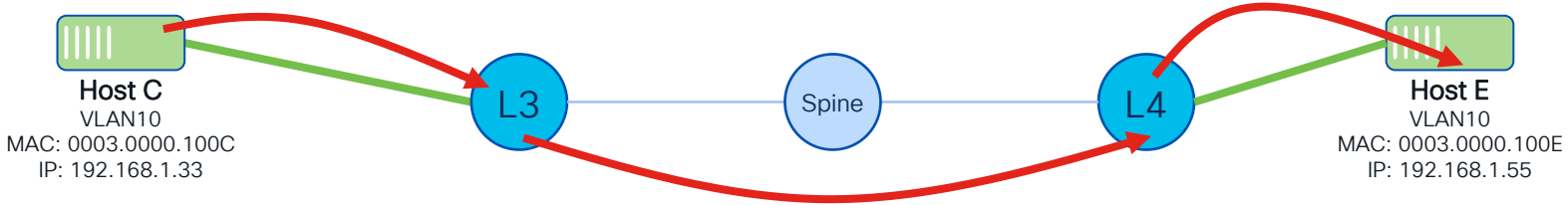


SIP	DIP	VNI	SMAC	DMAC	SIP	DIP	Payload
L3-IP	L4-IP	30001	0003.0000.100C	0003.0000.100E	192.168.1.33	192.168.1.55	

HostC to HostE

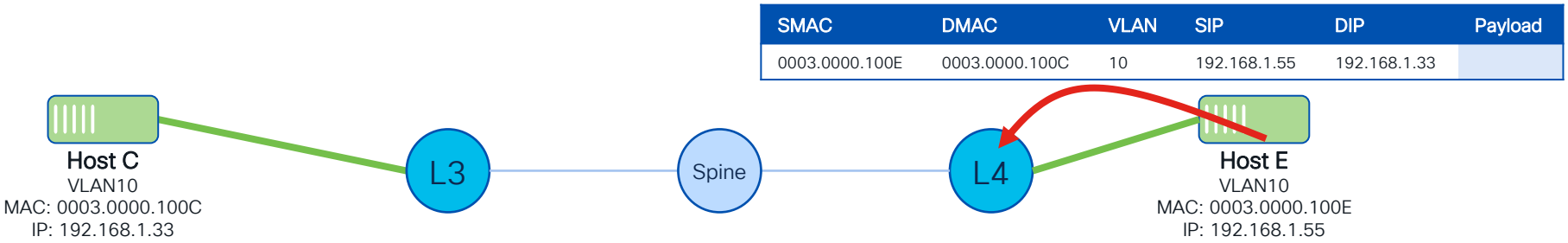
SMAC	DMAC	VLAN	SIP	DIP	Payload
0003.0000.100C	0003.0000.100E	10	192.168.1.33	192.168.1.55	

SMAC	DMAC	VLAN	SIP	DIP	Payload
0003.0000.100C	0003.0000.100E	10	192.168.1.33	192.168.1.55	

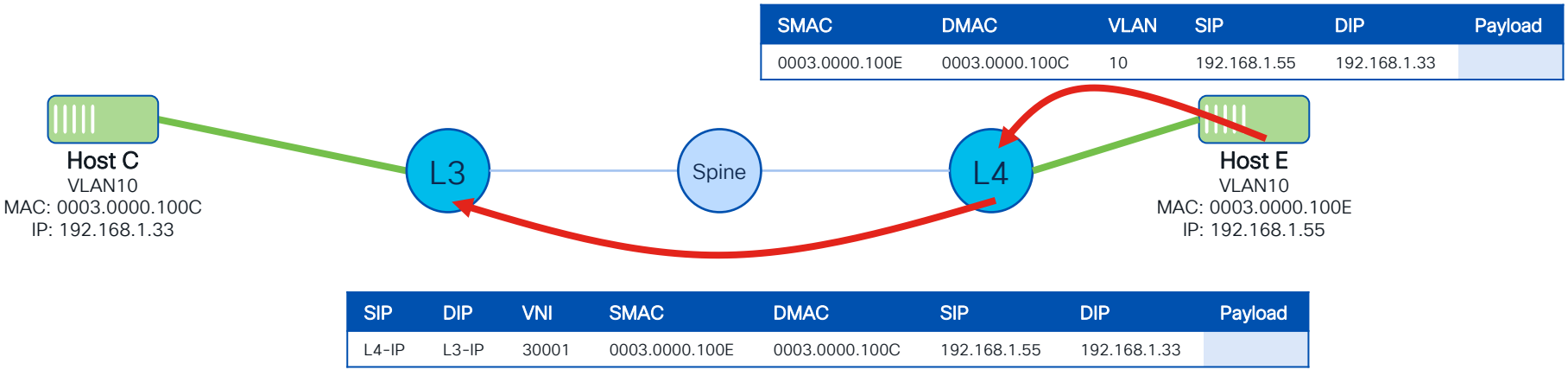


SIP	DIP	VNI	SMAC	DMAC	SIP	DIP	Payload
L3-IP	L4-IP	30001	0003.0000.100C	0003.0000.100E	192.168.1.33	192.168.1.55	

HostE to HostC



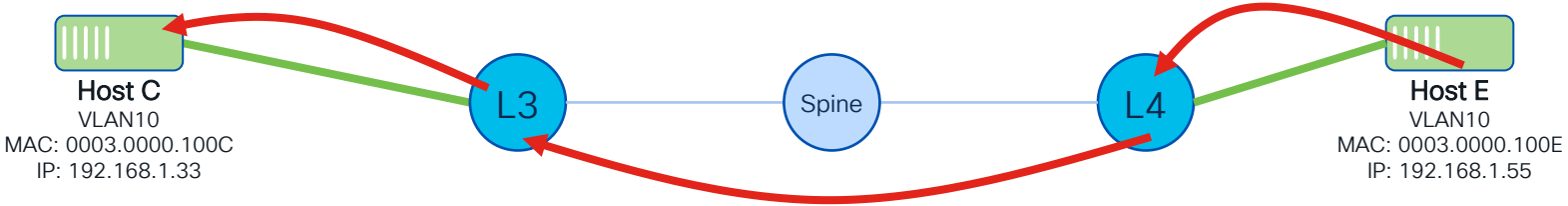
HostE to HostC



HostE to HostC

SMAC	DMAC	VLAN	SIP	DIP	Payload
0003.0000.100E	0003.0000.100C	10	192.168.1.55	192.168.1.33	

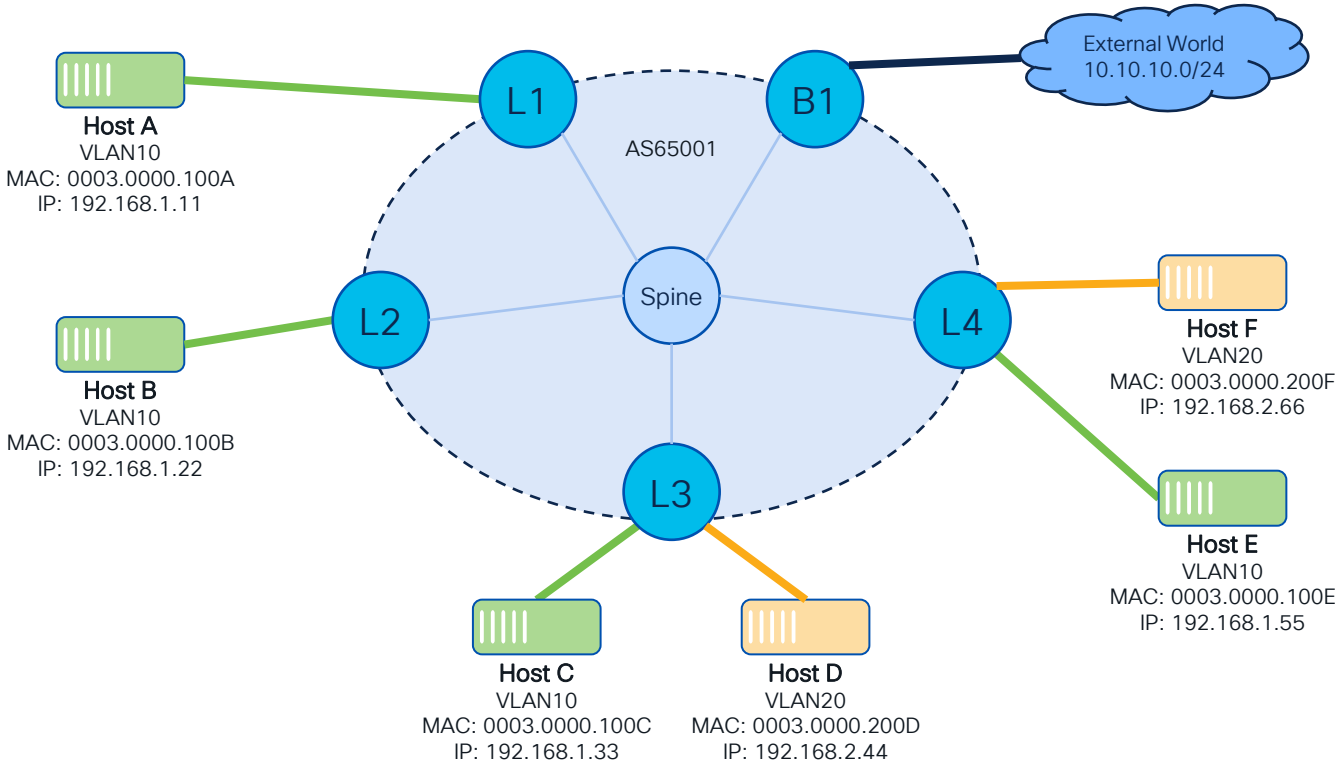
SMAC	DMAC	VLAN	SIP	DIP	Payload
0003.0000.100E	0003.0000.100C	10	192.168.1.55	192.168.1.33	



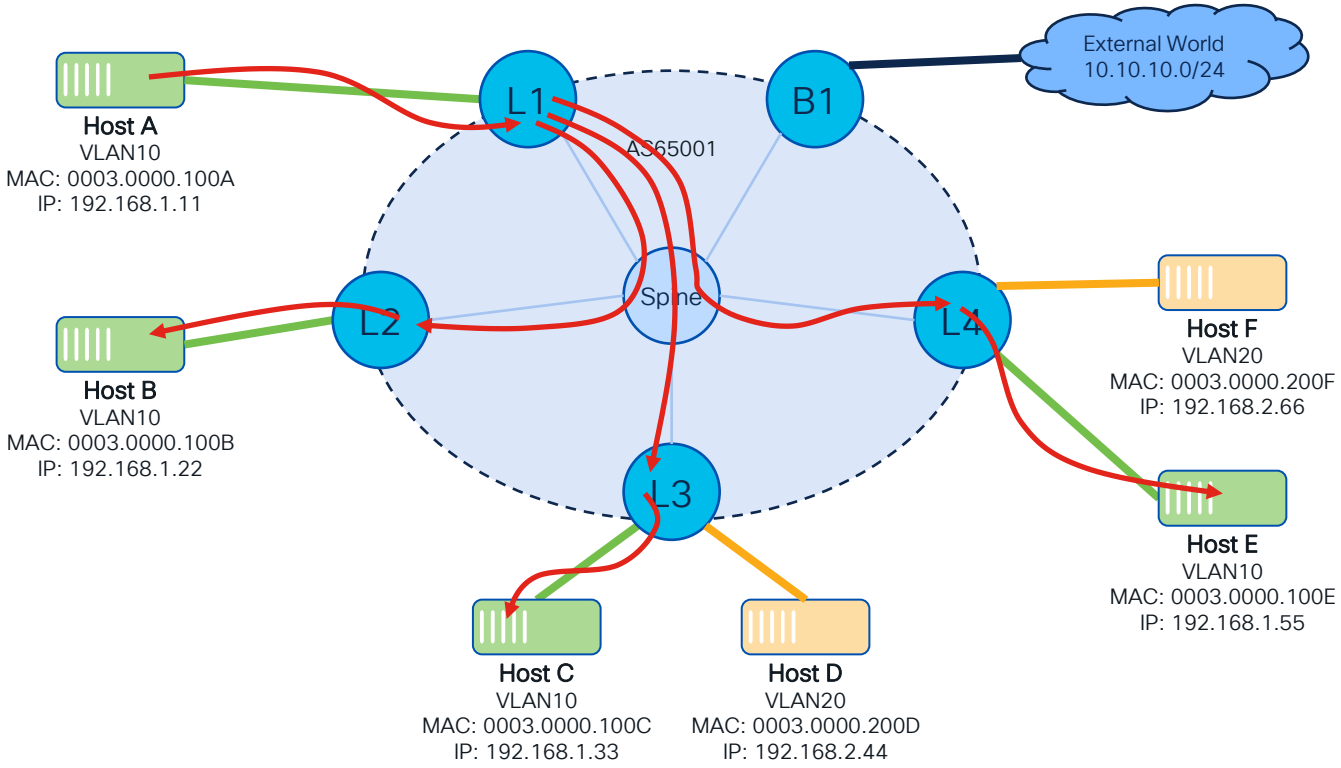
SIP	DIP	VNI	SMAC	DMAC	SIP	DIP	Payload
L4-IP	L3-IP	30001	0003.0000.100E	0003.0000.100C	192.168.1.55	192.168.1.33	

Packet Walk: BUM

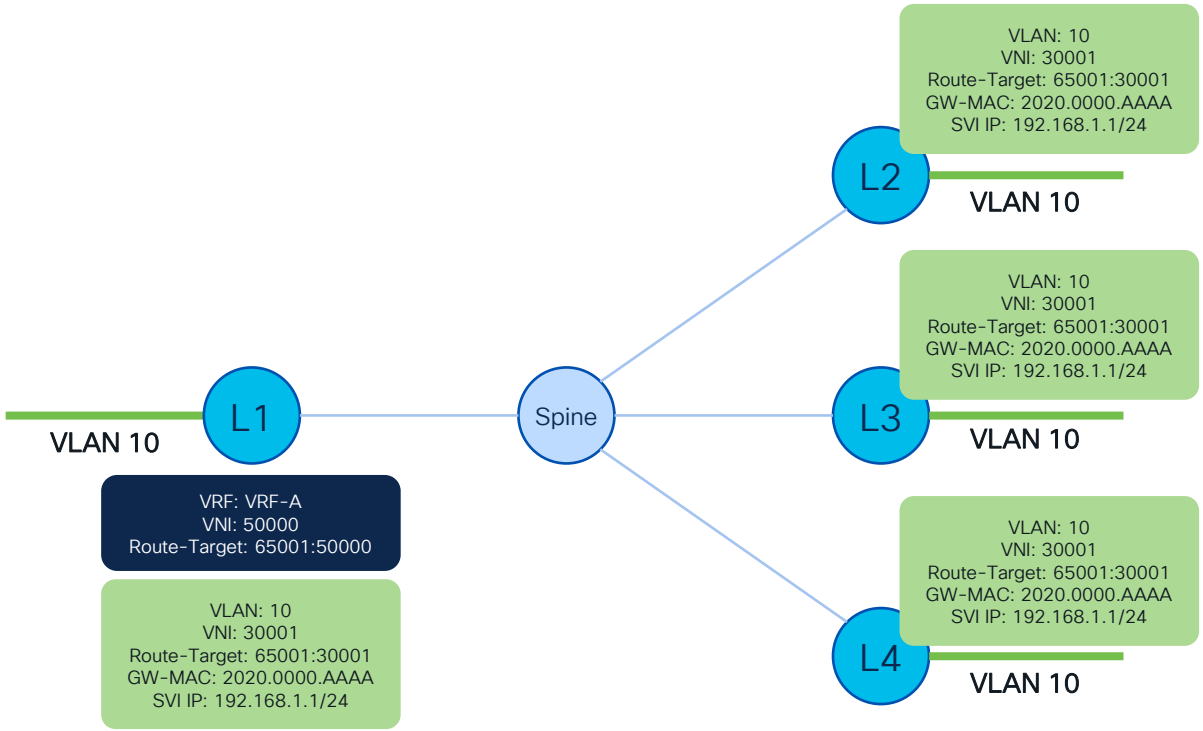
Topology Overview
BUM Packet Walk



Topology Overview
BUM Packet Walk (Ingress Replication)



Learning: VNI 30001 Participation (IR)

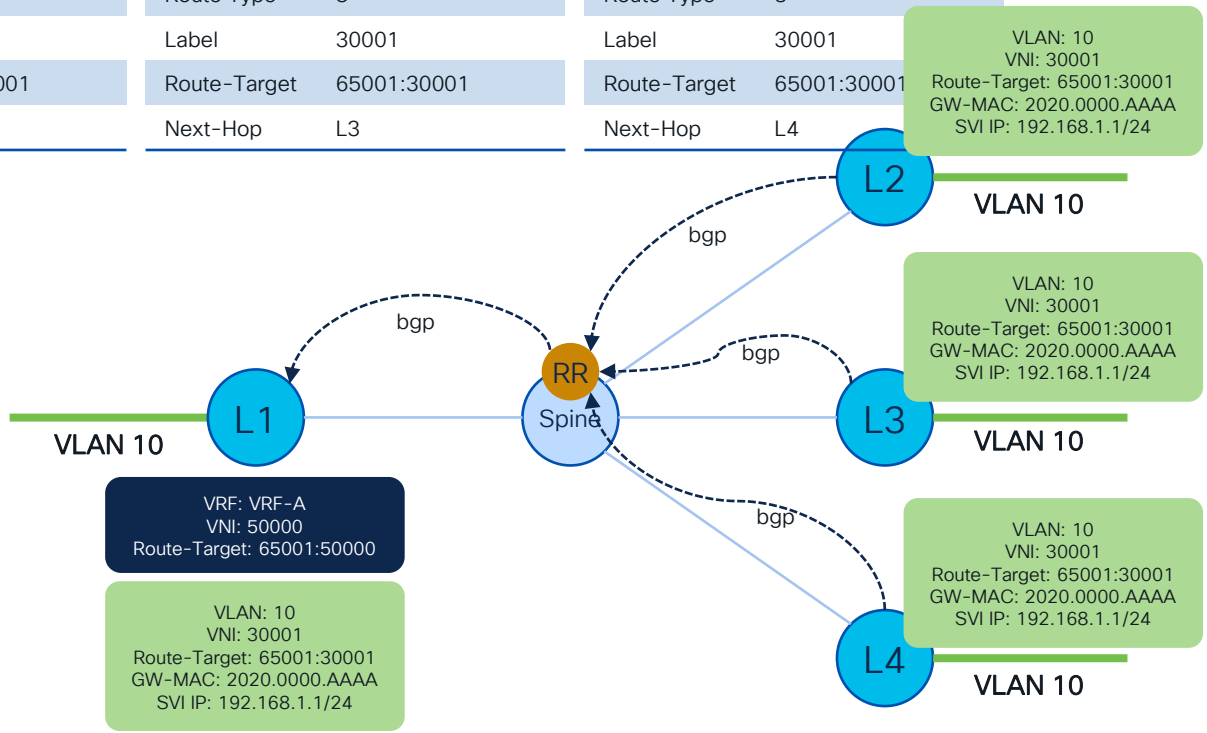


Learning: VNI 30001 Participation (IR)

EVPN NLRI	
Route Type	3
Label	30001
Route-Target	65001:30001
Next-Hop	L2

EVPN NLRI	
Route Type	3
Label	30001
Route-Target	65001:30001
Next-Hop	L3

EVPN NLRI	
Route Type	3
Label	30001
Route-Target	65001:30001
Next-Hop	L4

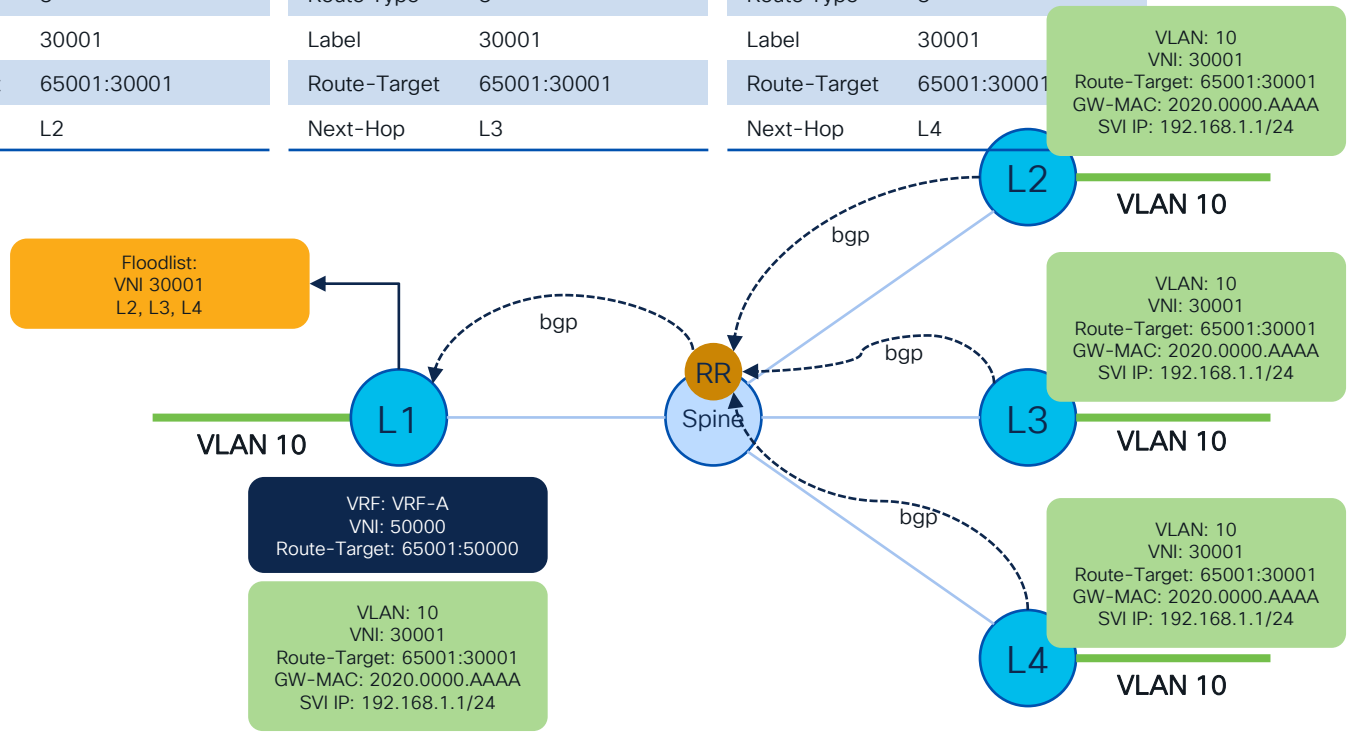


Learning: VNI 30001 Participation (IR)

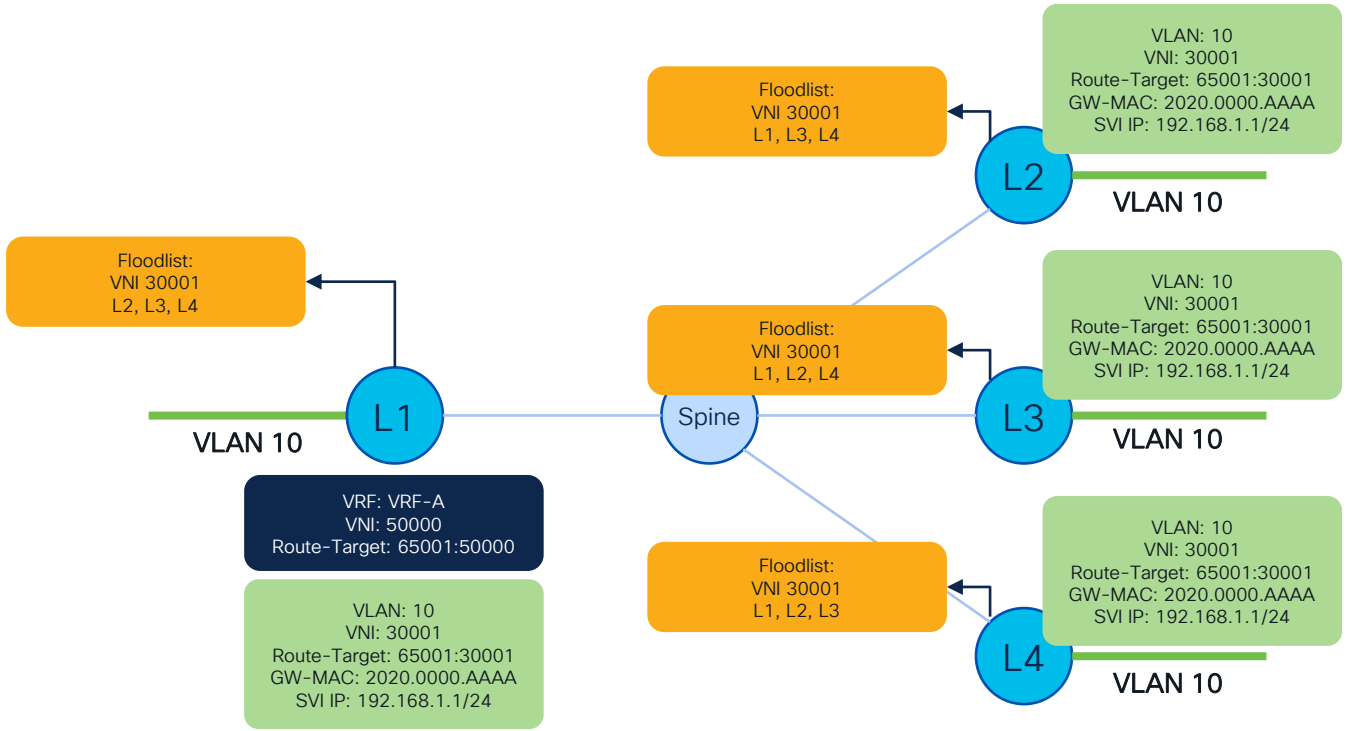
EVPN NLRI	
Route Type	3
Label	30001
Route-Target	65001:30001
Next-Hop	L2

EVPN NLRI	
Route Type	3
Label	30001
Route-Target	65001:30001
Next-Hop	L3

EVPN NLRI	
Route Type	3
Label	30001
Route-Target	65001:30001
Next-Hop	L4

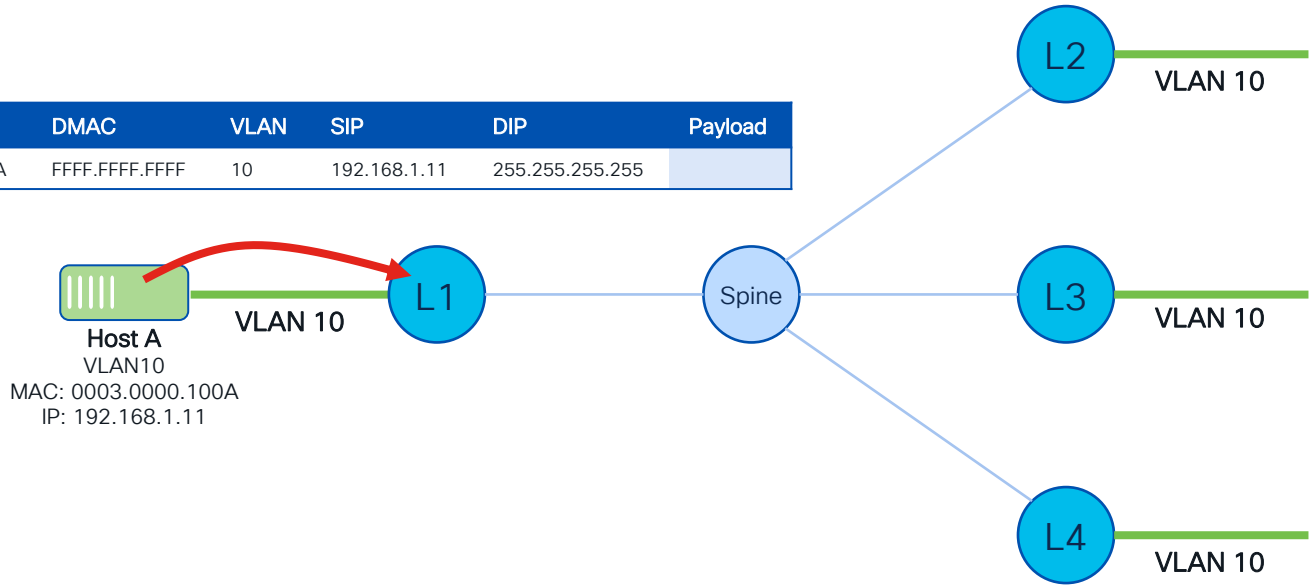


Forwarding Tables (IR)

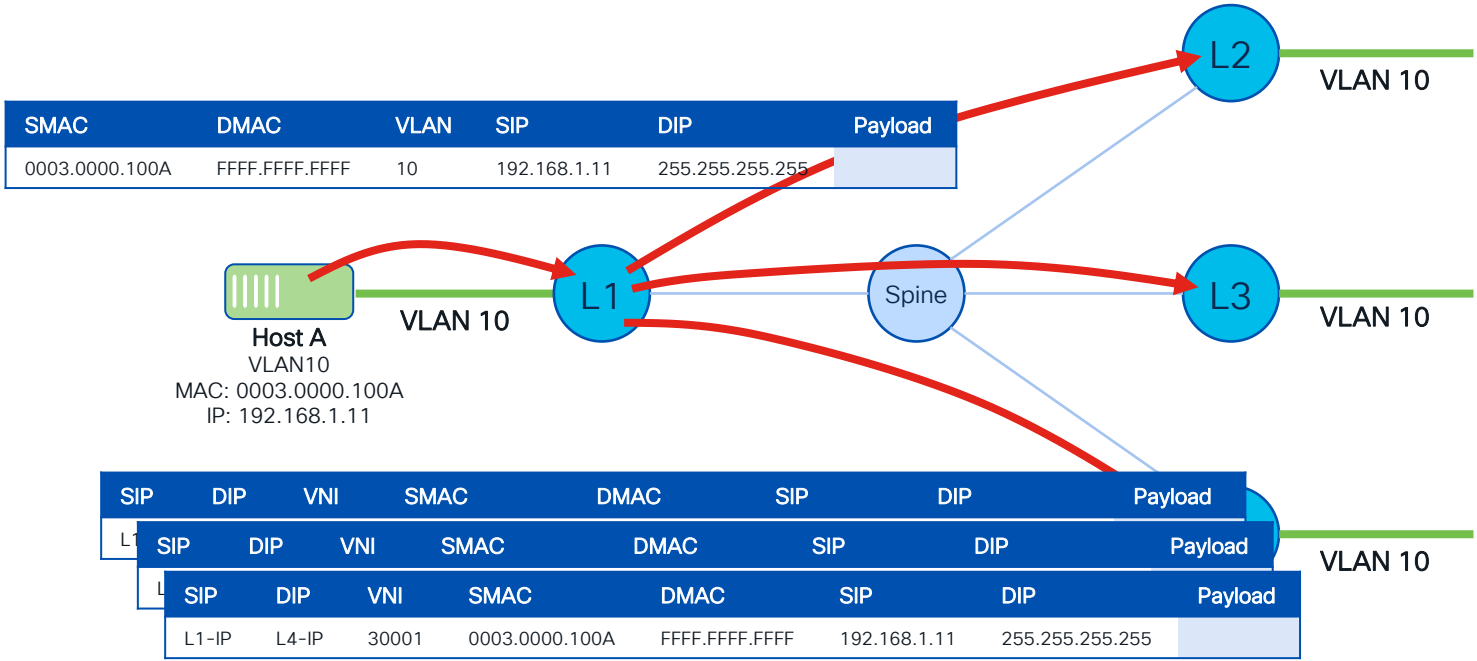


HostA sends BUM (IR)

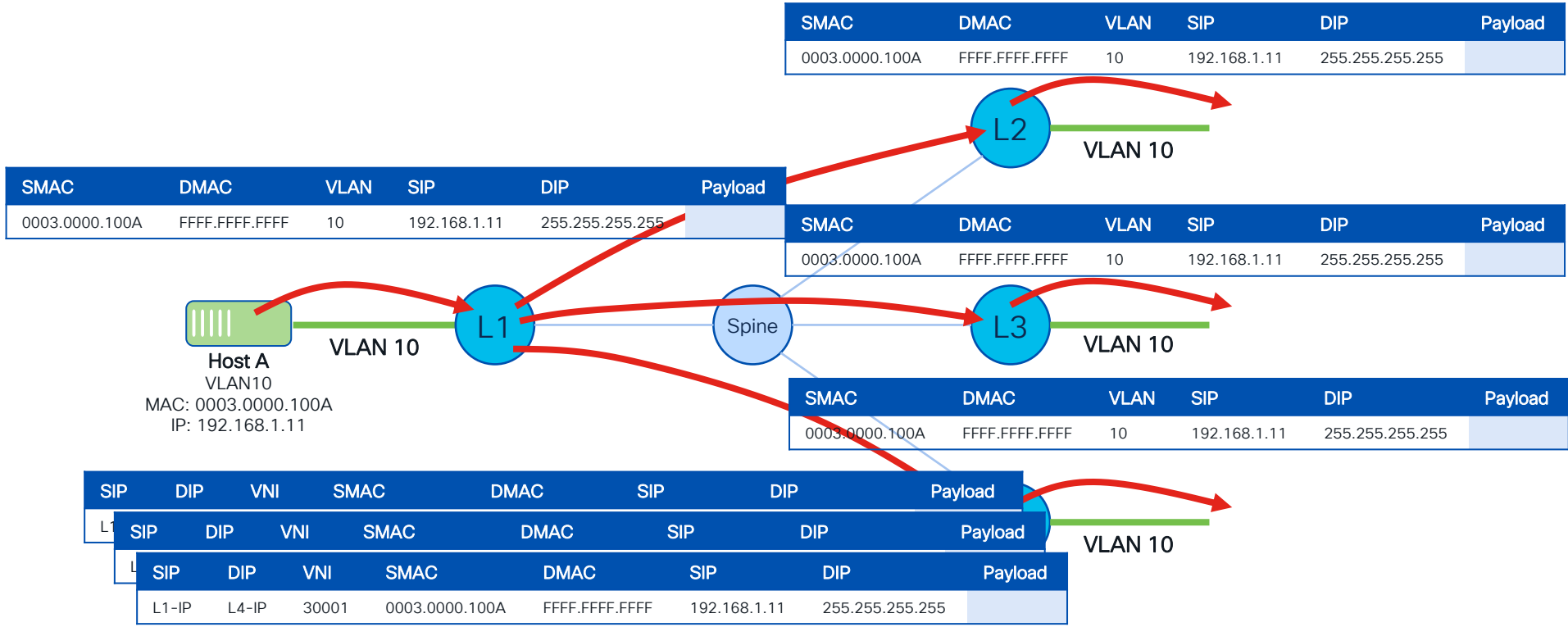
SMAC	DMAC	VLAN	SIP	DIP	Payload
0003.0000.100A	FFFF.FFFF.FFFF	10	192.168.1.11	255.255.255.255	



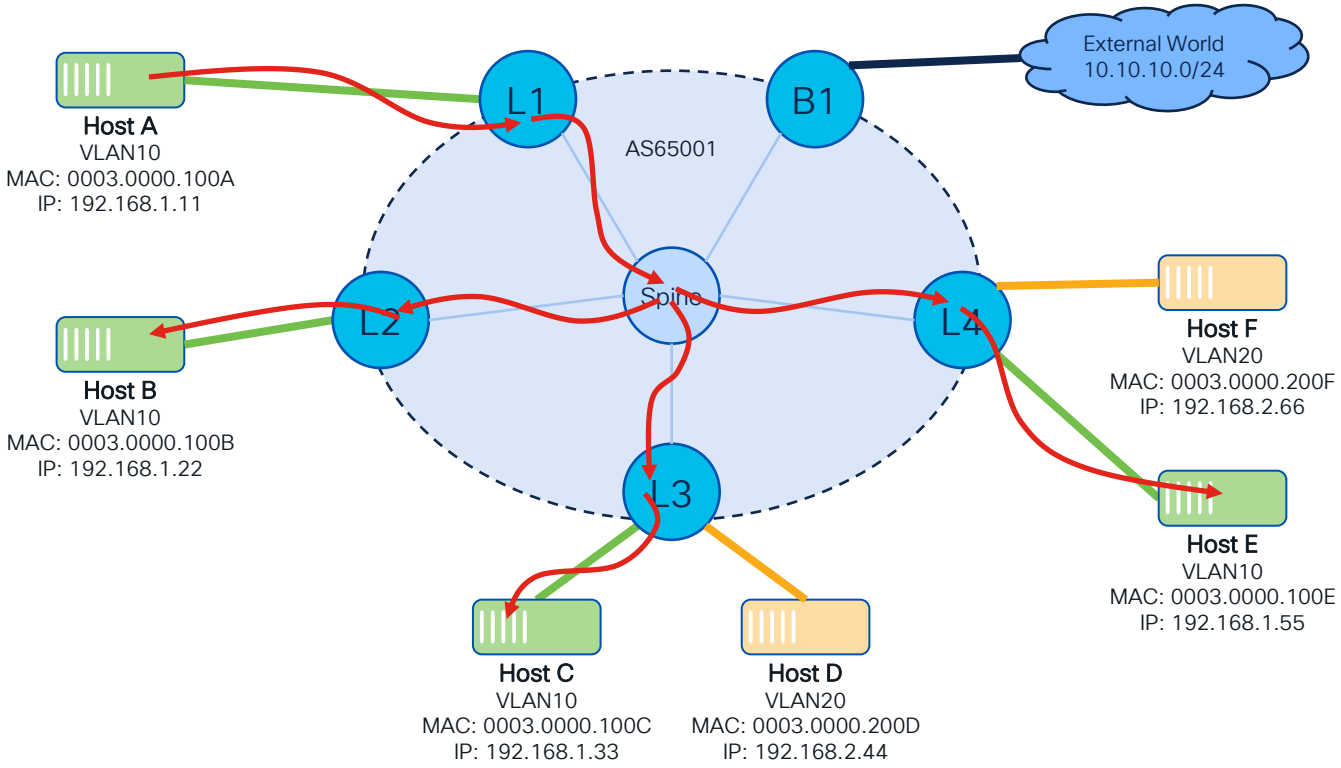
HostA sends BUM (IR)



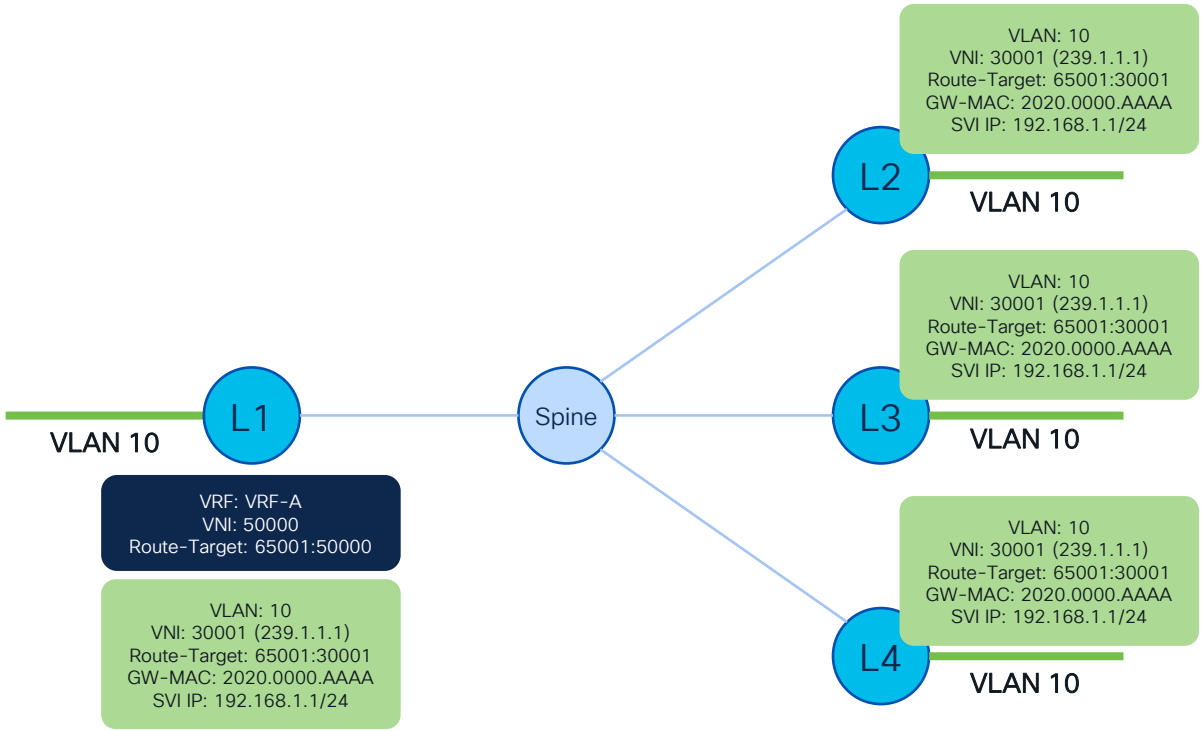
HostA sends BUM (IR)



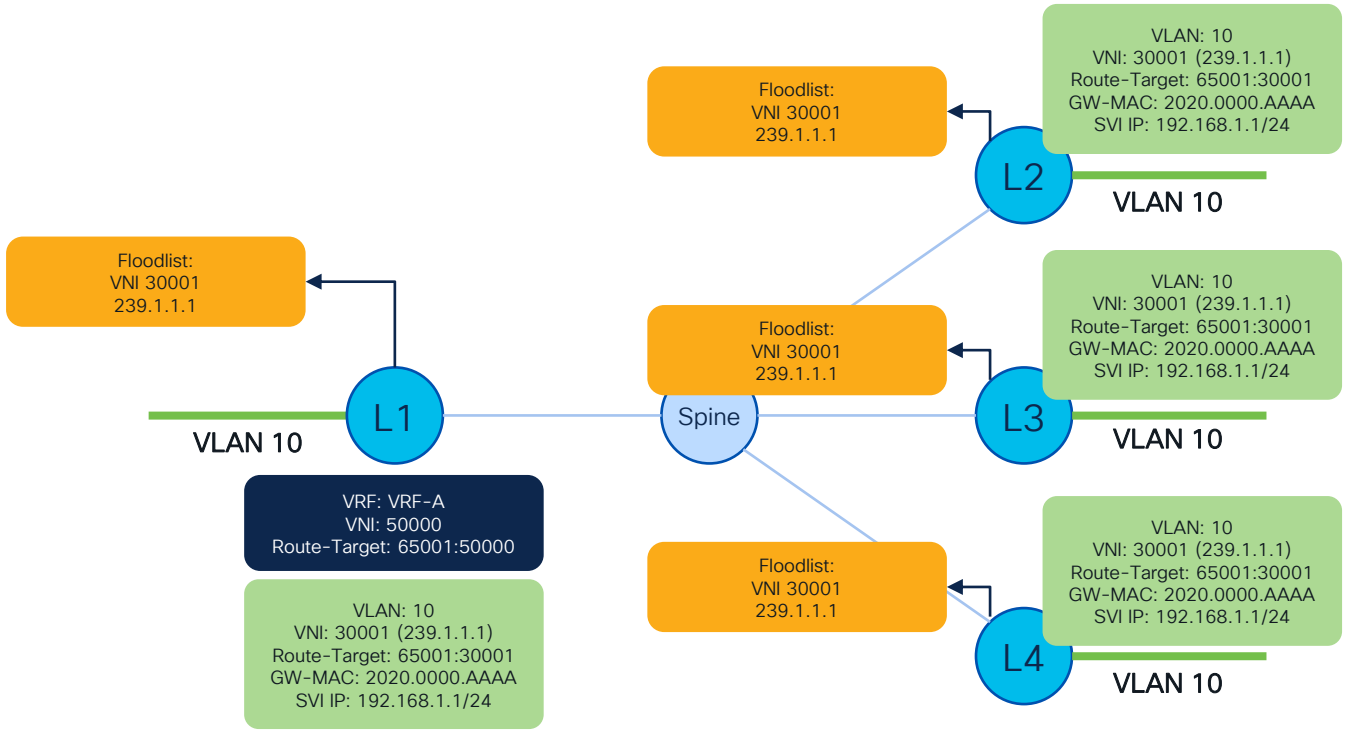
Topology Overview
BUM Packet Walk (Multicast)



Learning: VNI 30001 Participation (MCAST)

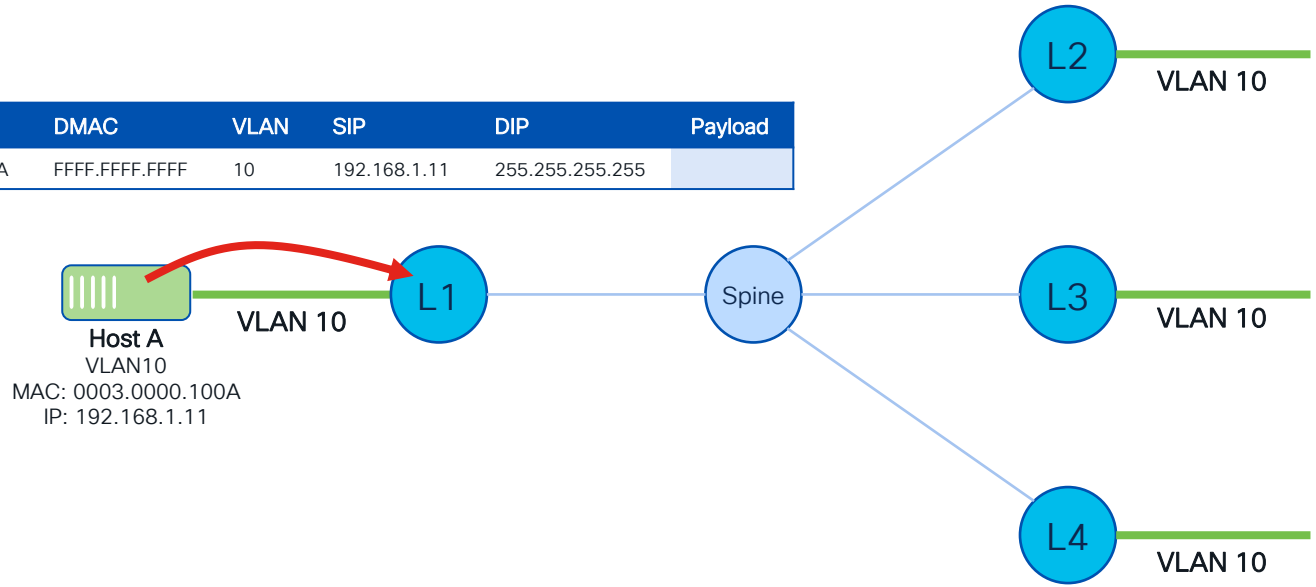


Forwarding Tables (MCAST)



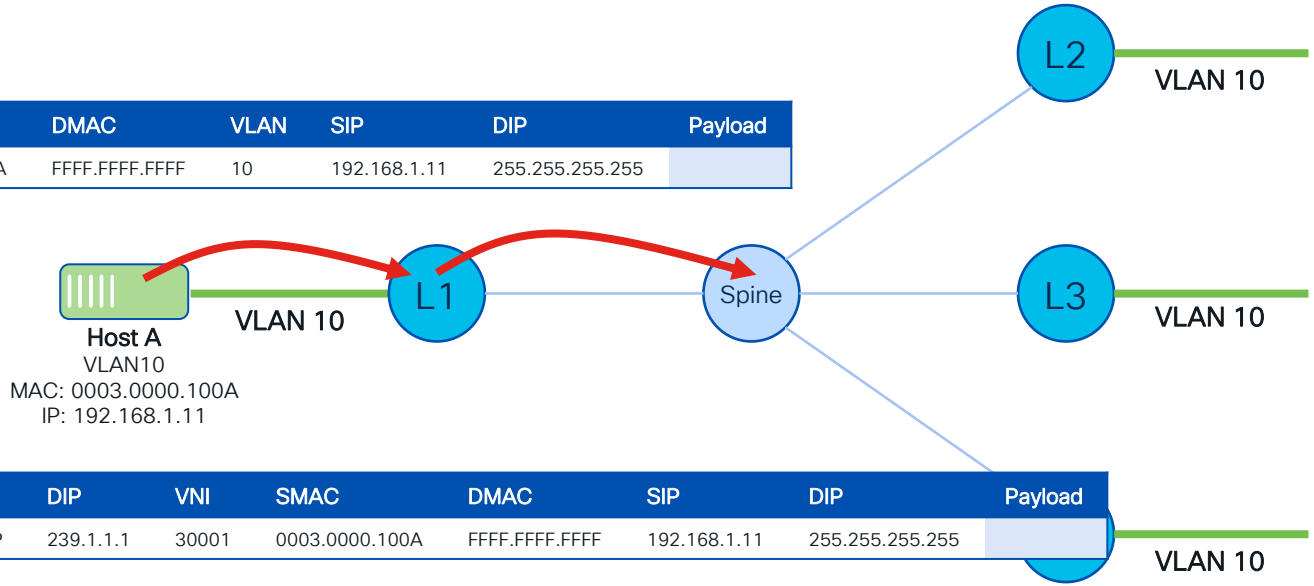
HostA sends BUM (MCAST)

SMAC	DMAC	VLAN	SIP	DIP	Payload
0003.0000.100A	FFFF.FFFF.FFFF	10	192.168.1.11	255.255.255.255	



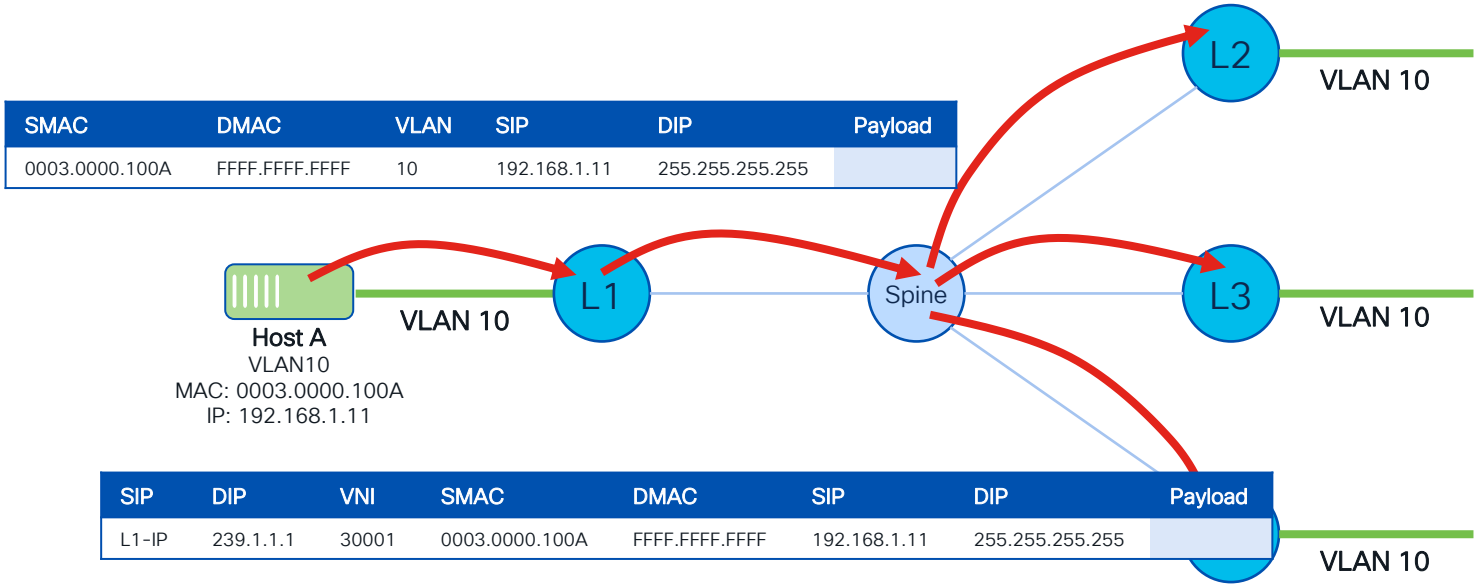
HostA sends BUM (MCAST)

SMAC	DMAC	VLAN	SIP	DIP	Payload
0003.0000.100A	FFFF.FFFF.FFFF	10	192.168.1.11	255.255.255.255	

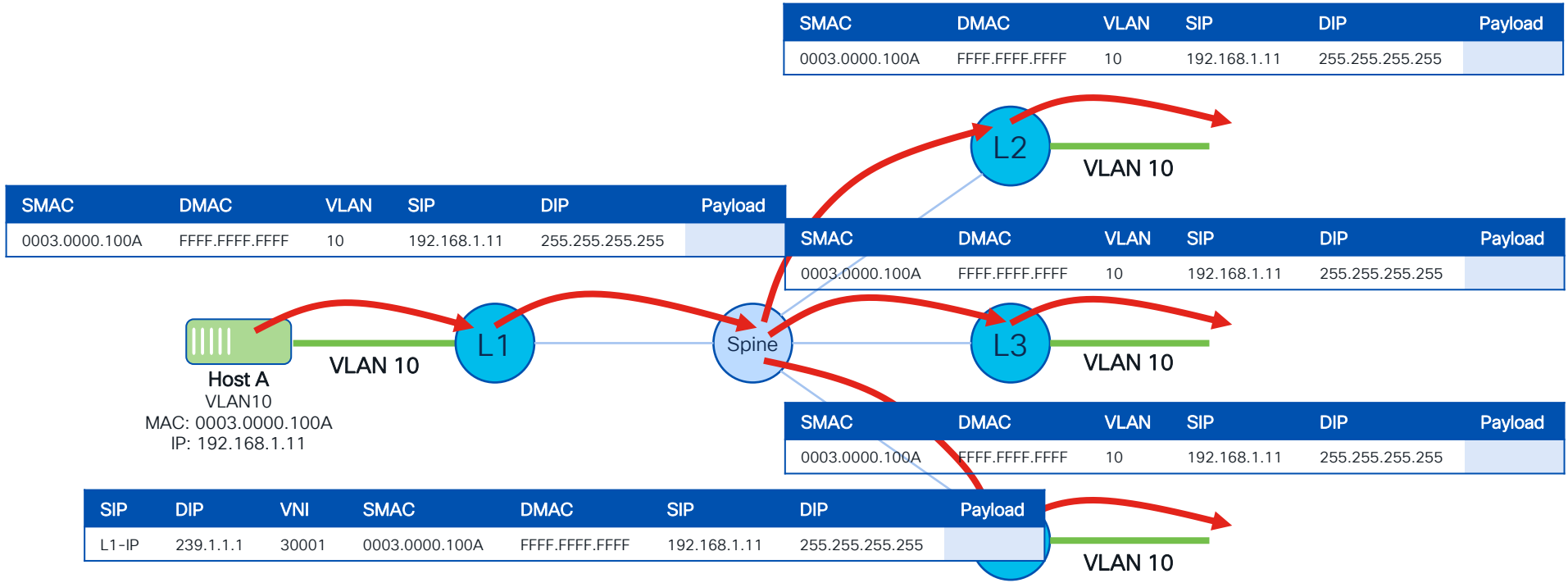


SIP	DIP	VNI	SMAC	DMAC	SIP	DIP	Payload
L1-IP	239.1.1.1	30001	0003.0000.100A	FFFF.FFFF.FFFF	192.168.1.11	255.255.255.255	

HostA sends BUM (MCAST)

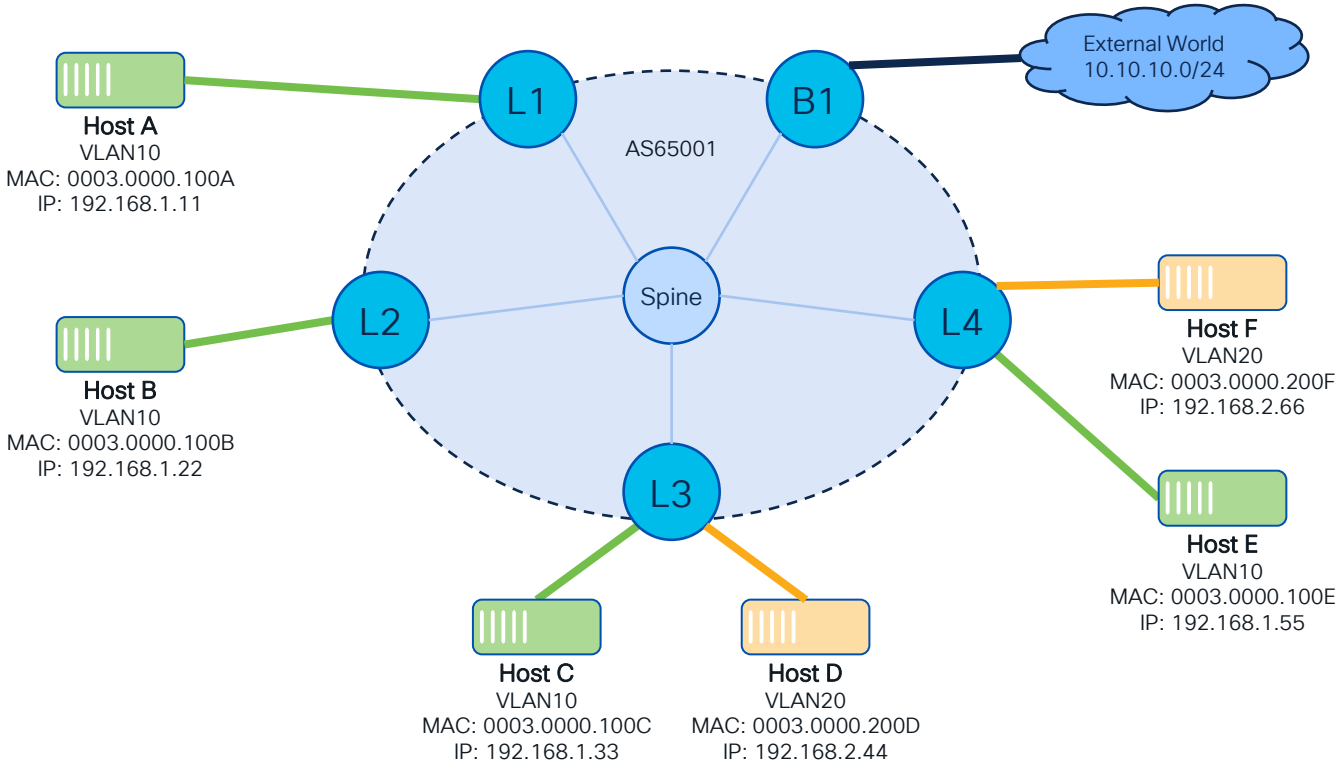


HostA sends BUM (MCAST)

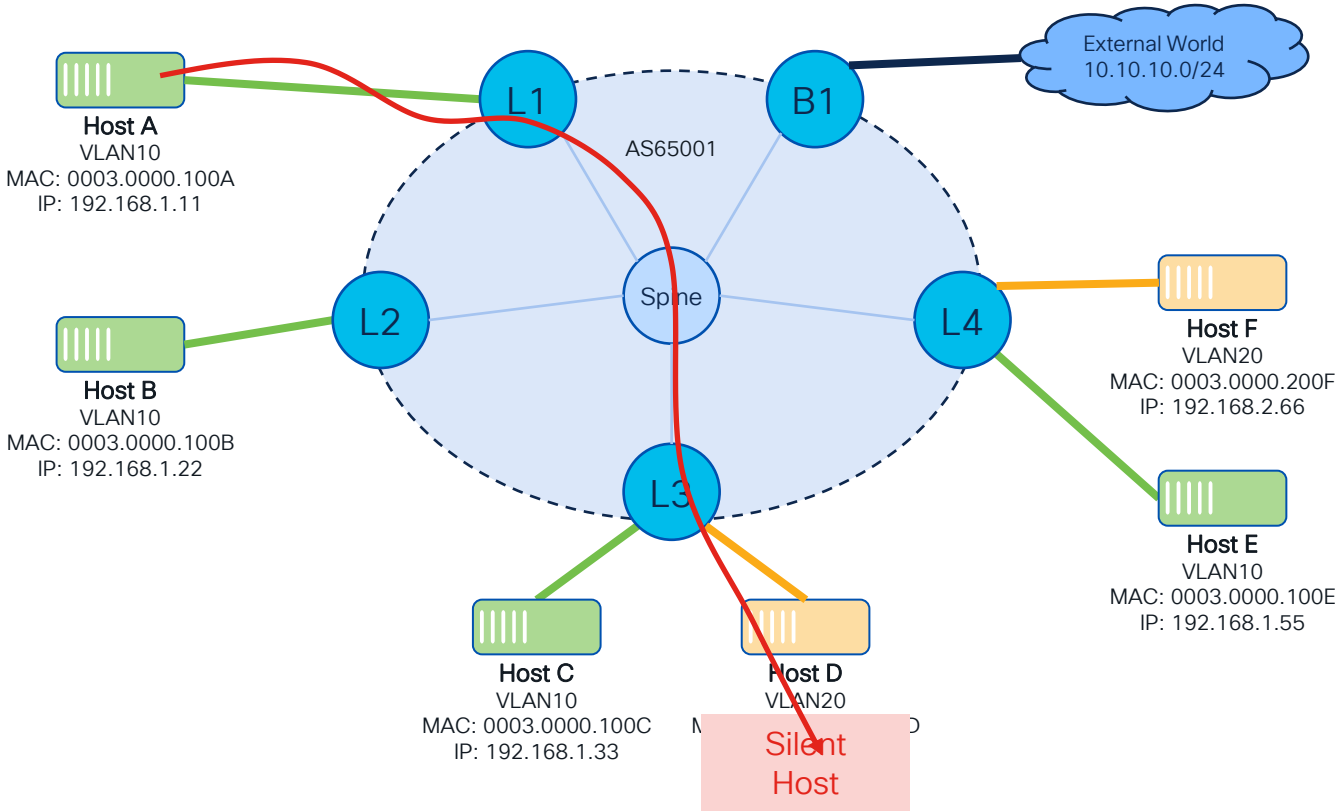


Silent Host Discovery

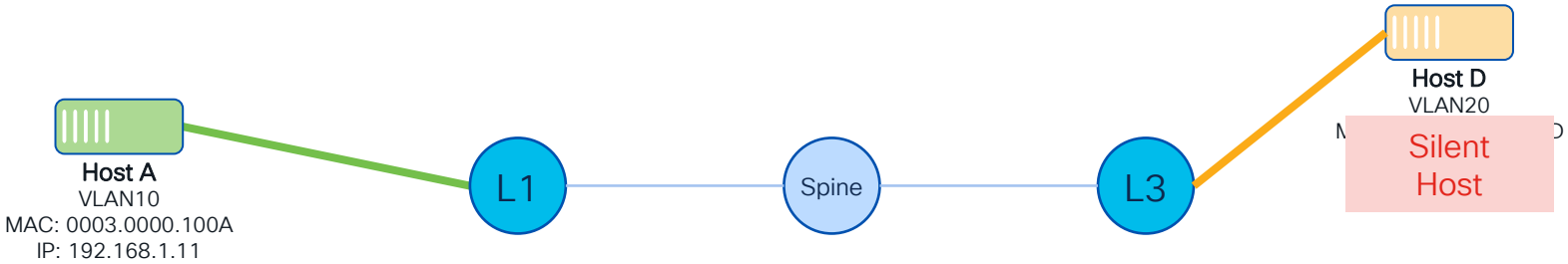
Topology Overview
Silent Host Discovery



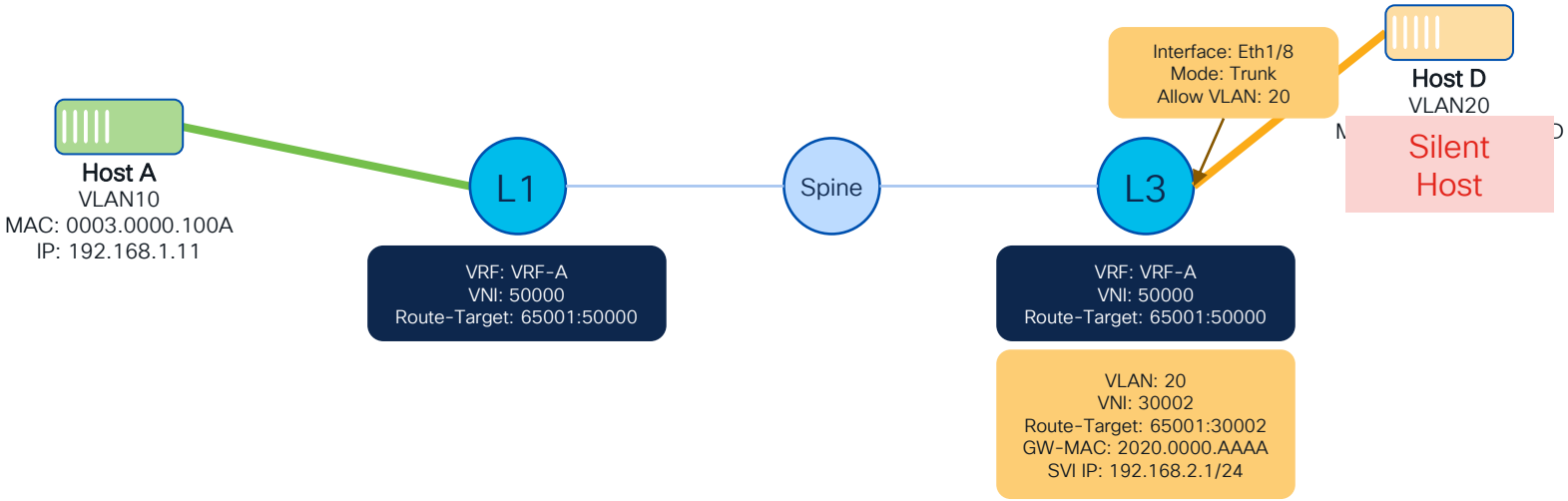
Topology Overview
Silent Host Discovery



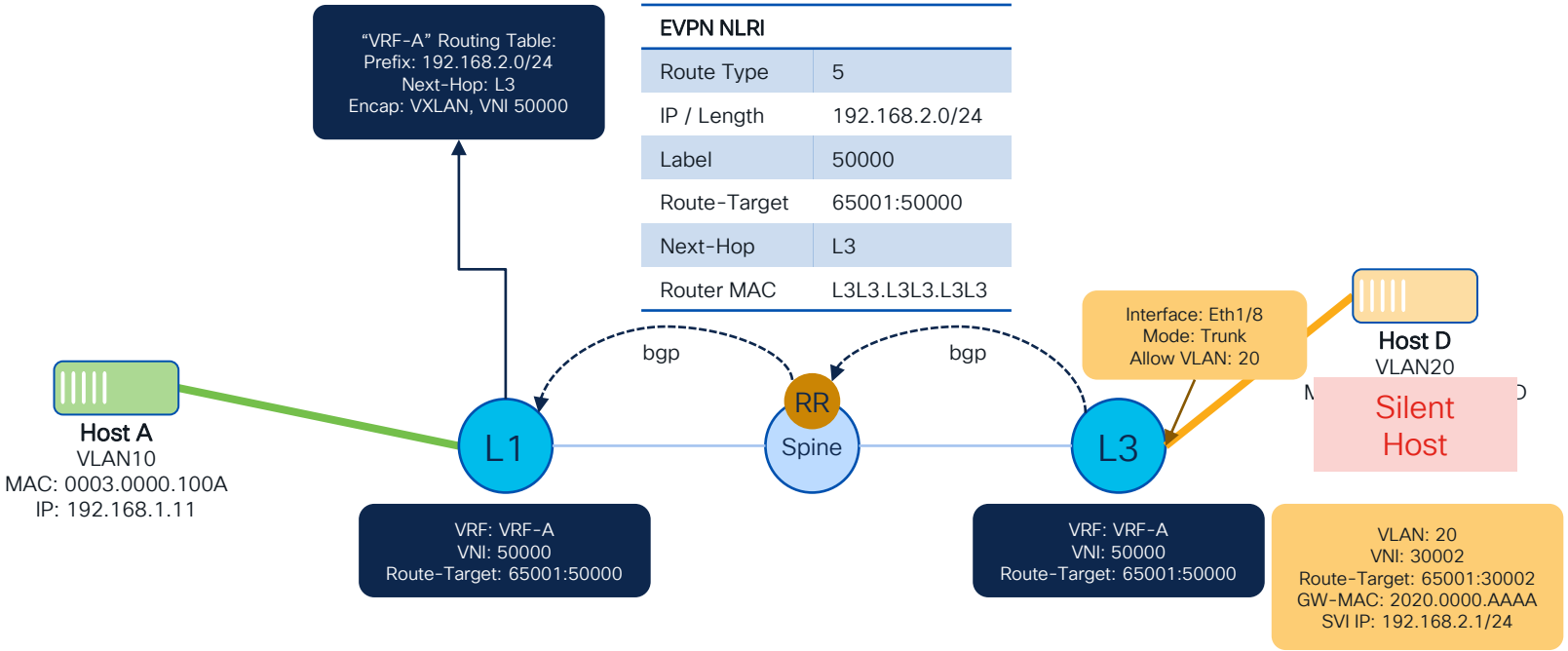
Learning: HostD to Leaf1



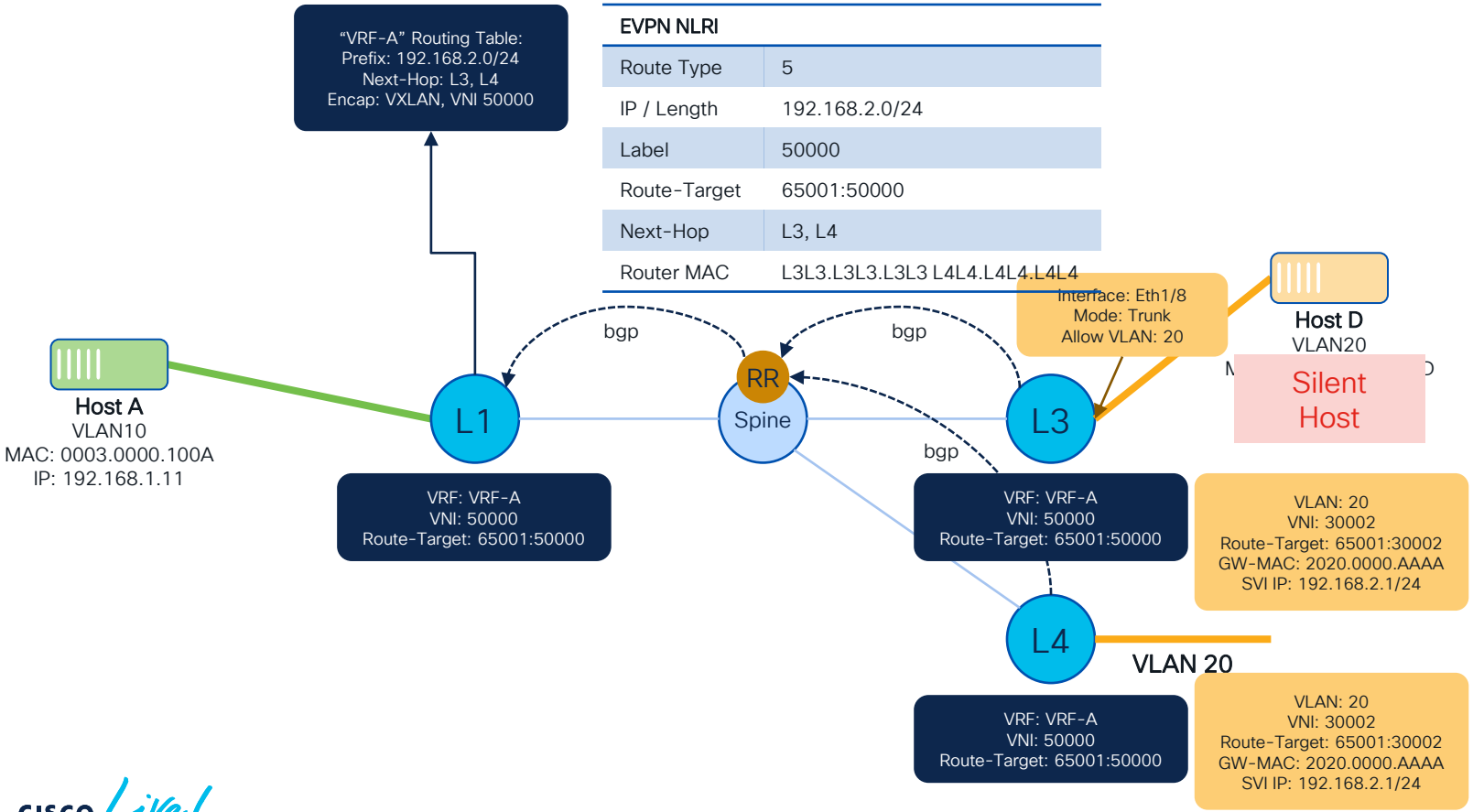
Learning: HostD to Leaf1



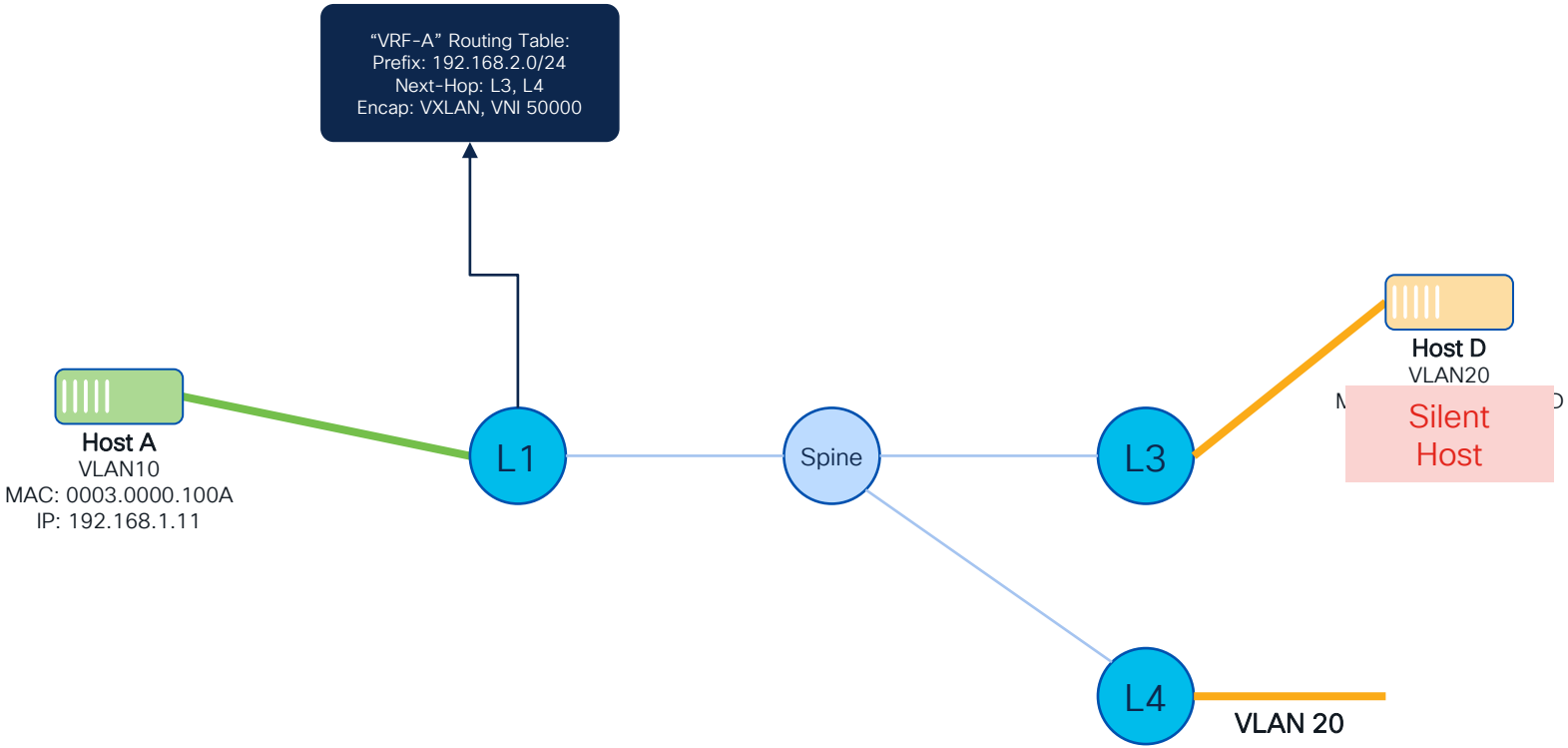
Learning: HostD to Leaf1



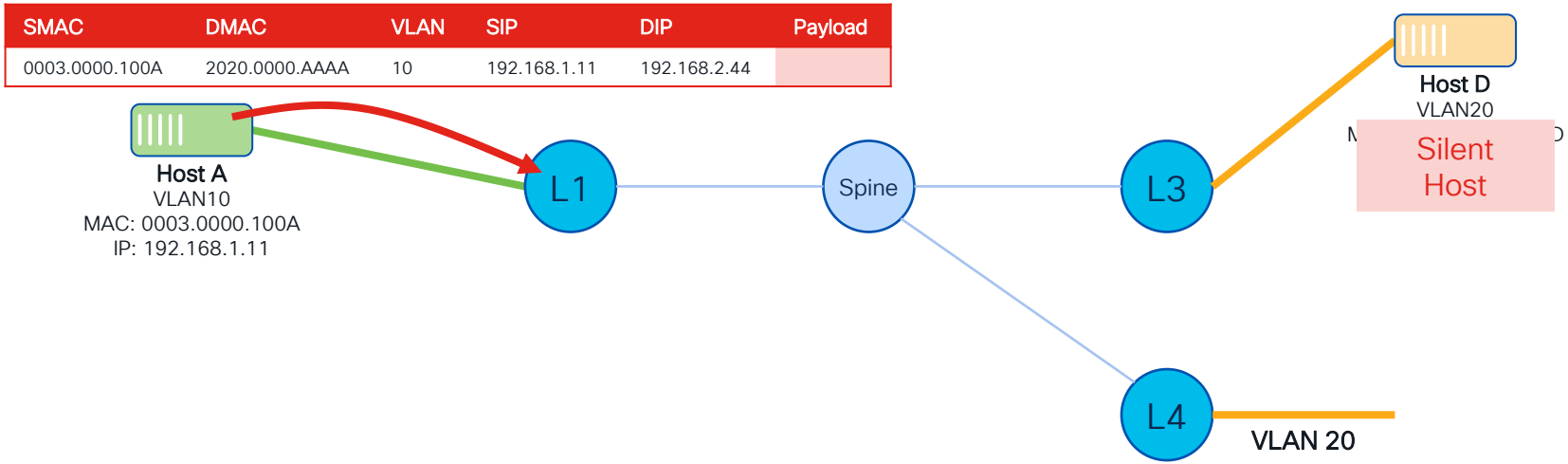
Learning: HostD to Leaf1



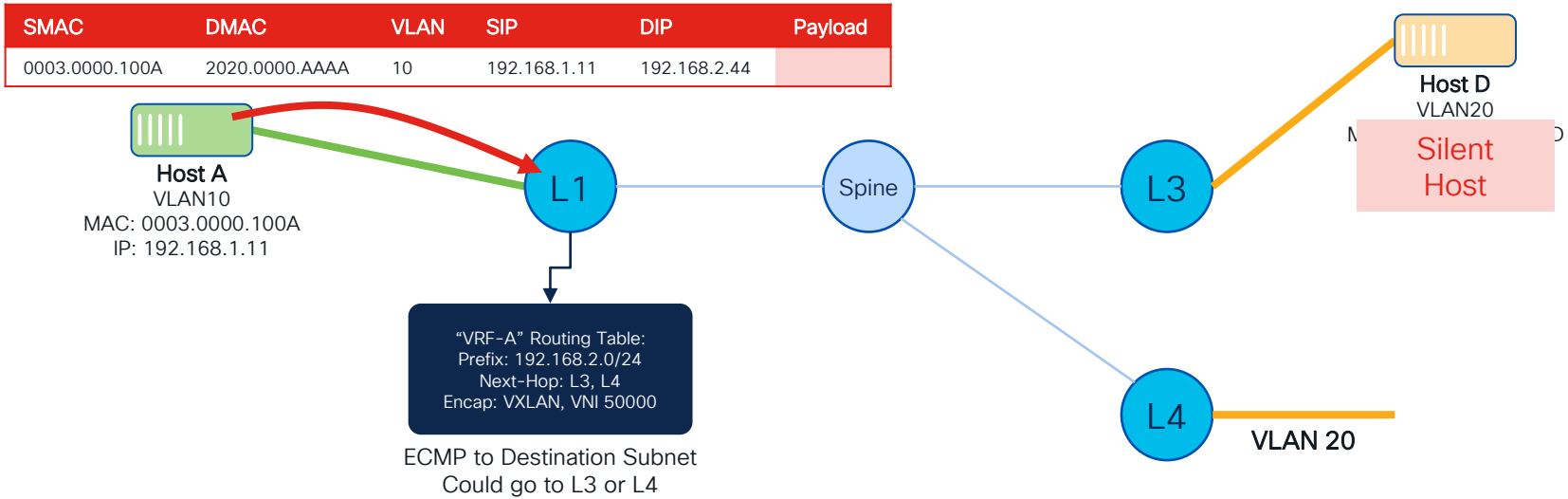
Forwarding Tables



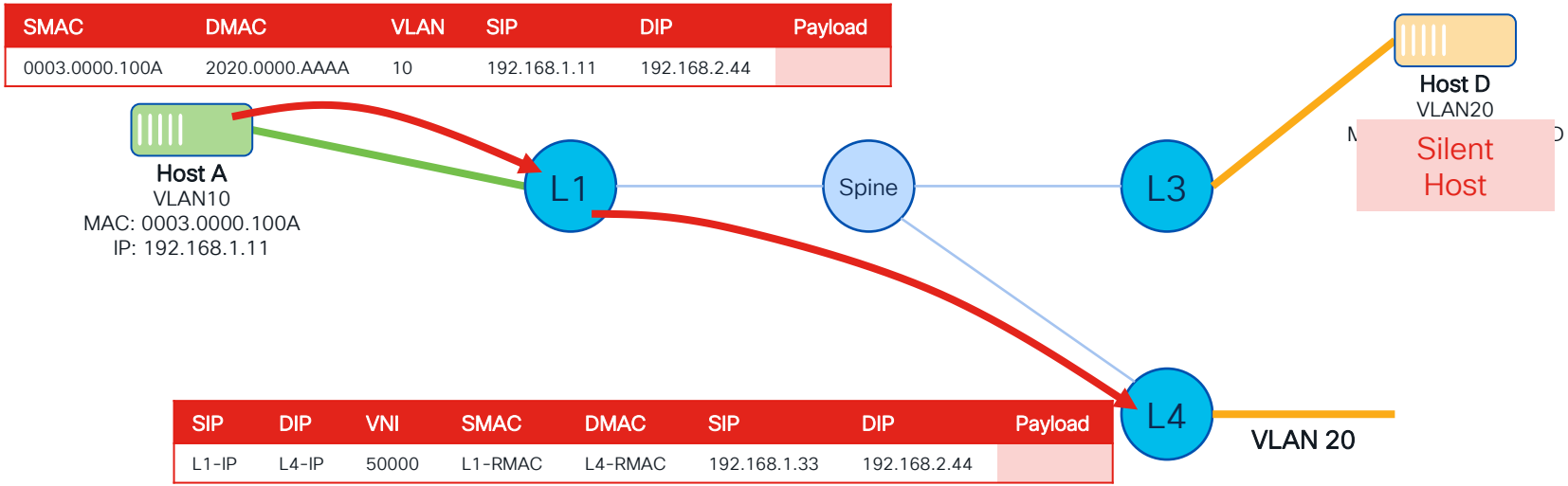
HostA to HostD



HostA to HostD

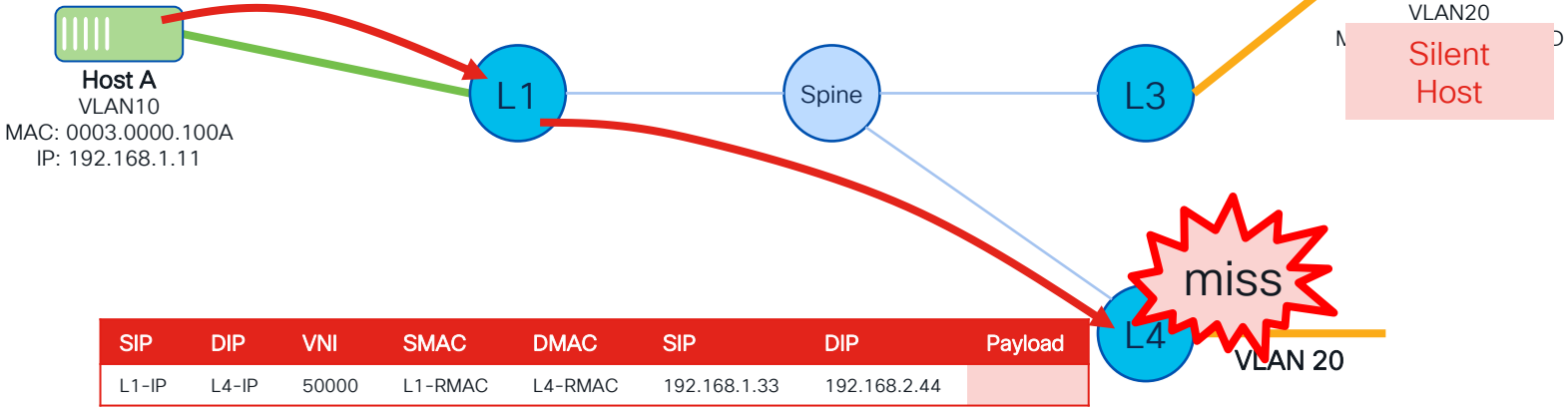


HostA to HostD



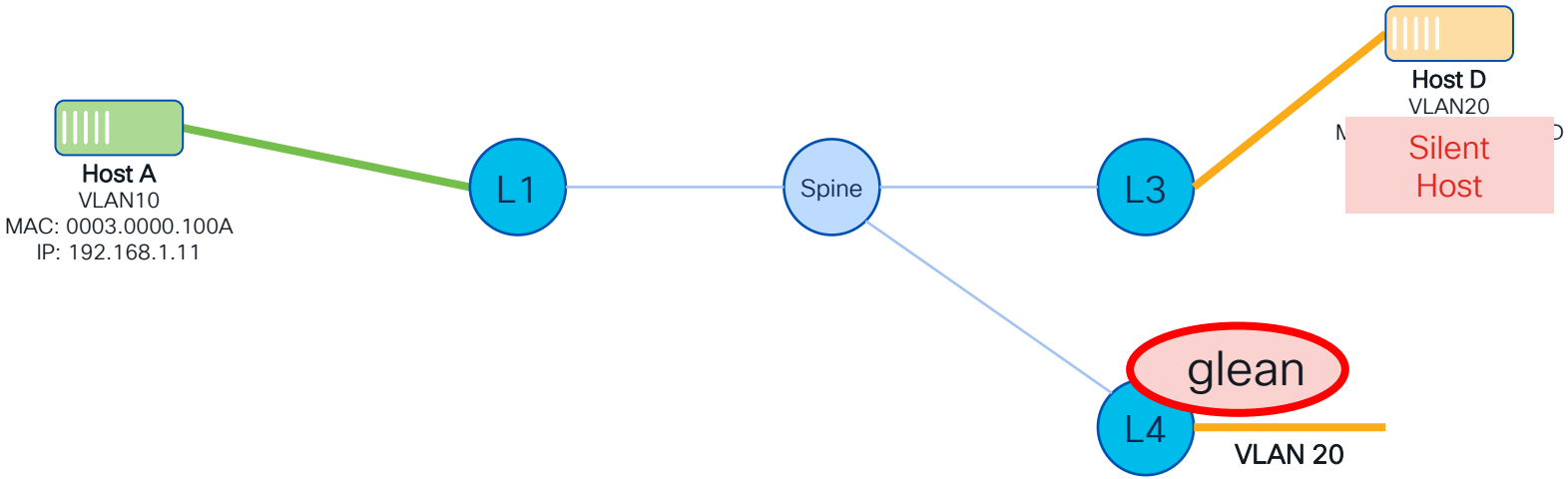
HostA to HostD

SMAC	DMAC	VLAN	SIP	DIP	Payload
0003.0000.100A	2020.0000.AAAA	10	192.168.1.11	192.168.2.44	

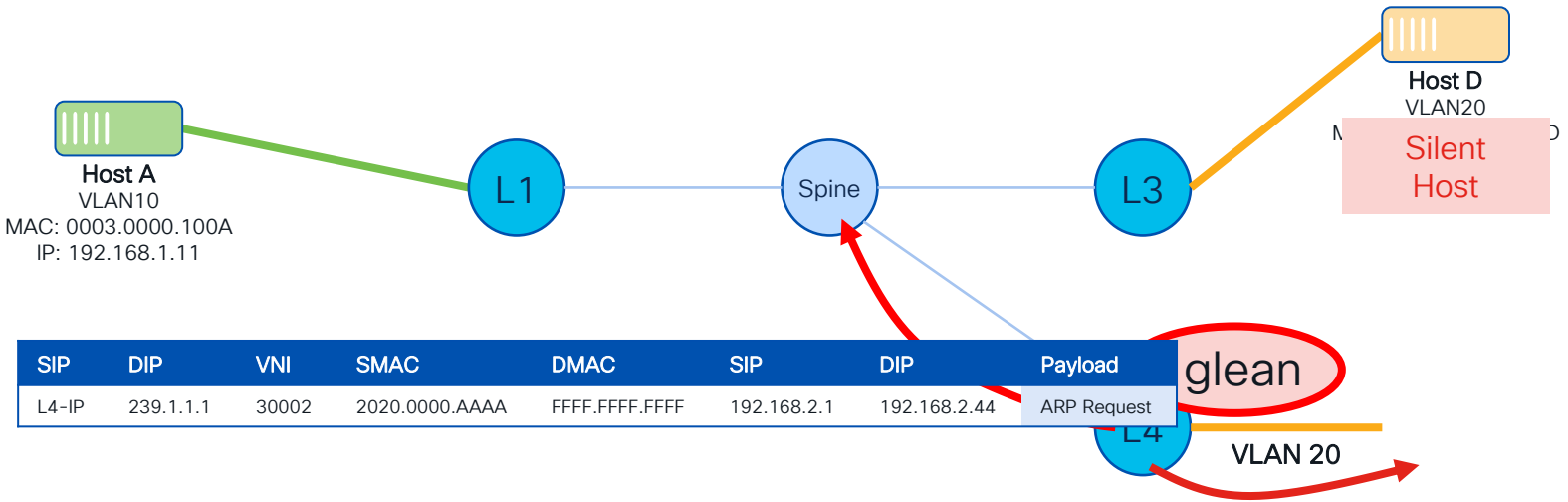


SIP	DIP	VNI	SMAC	DMAC	SIP	DIP	Payload
L1-IP	L4-IP	50000	L1-RMAC	L4-RMAC	192.168.1.33	192.168.2.44	

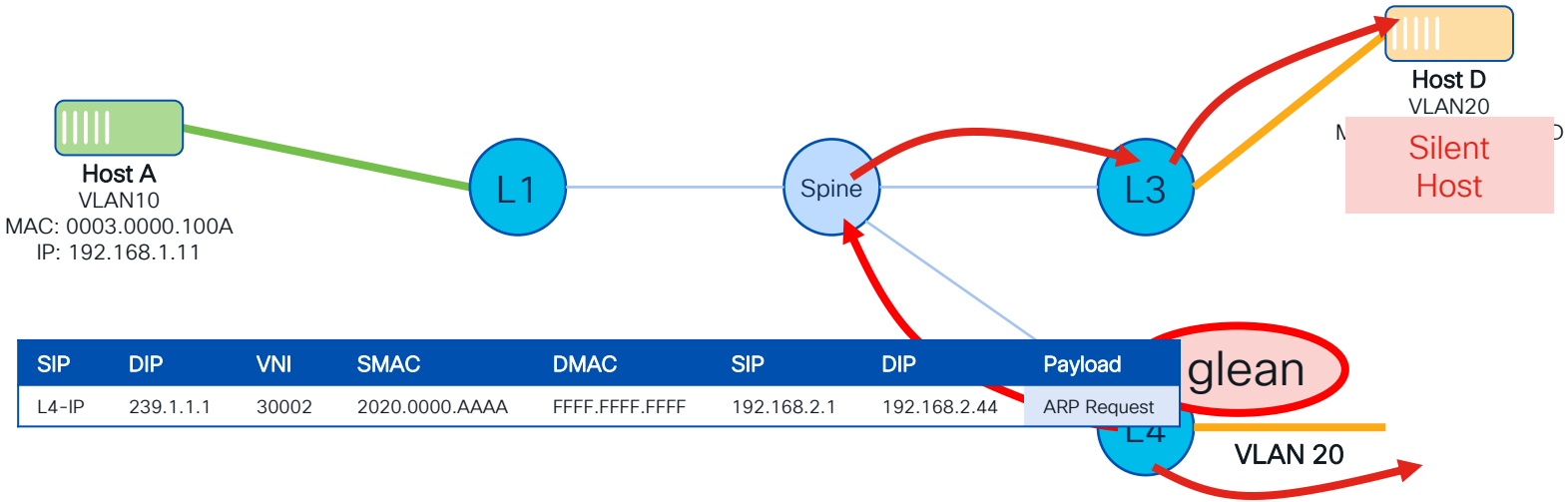
HostA to HostD



HostA to HostD



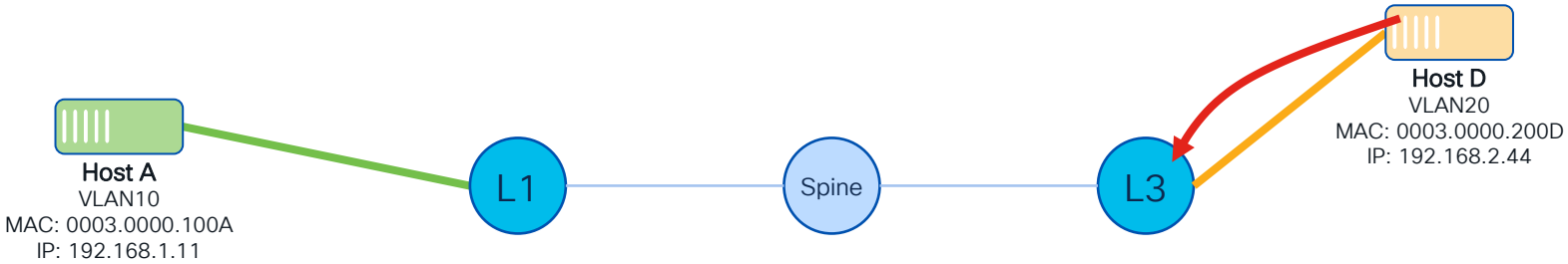
HostA to HostD



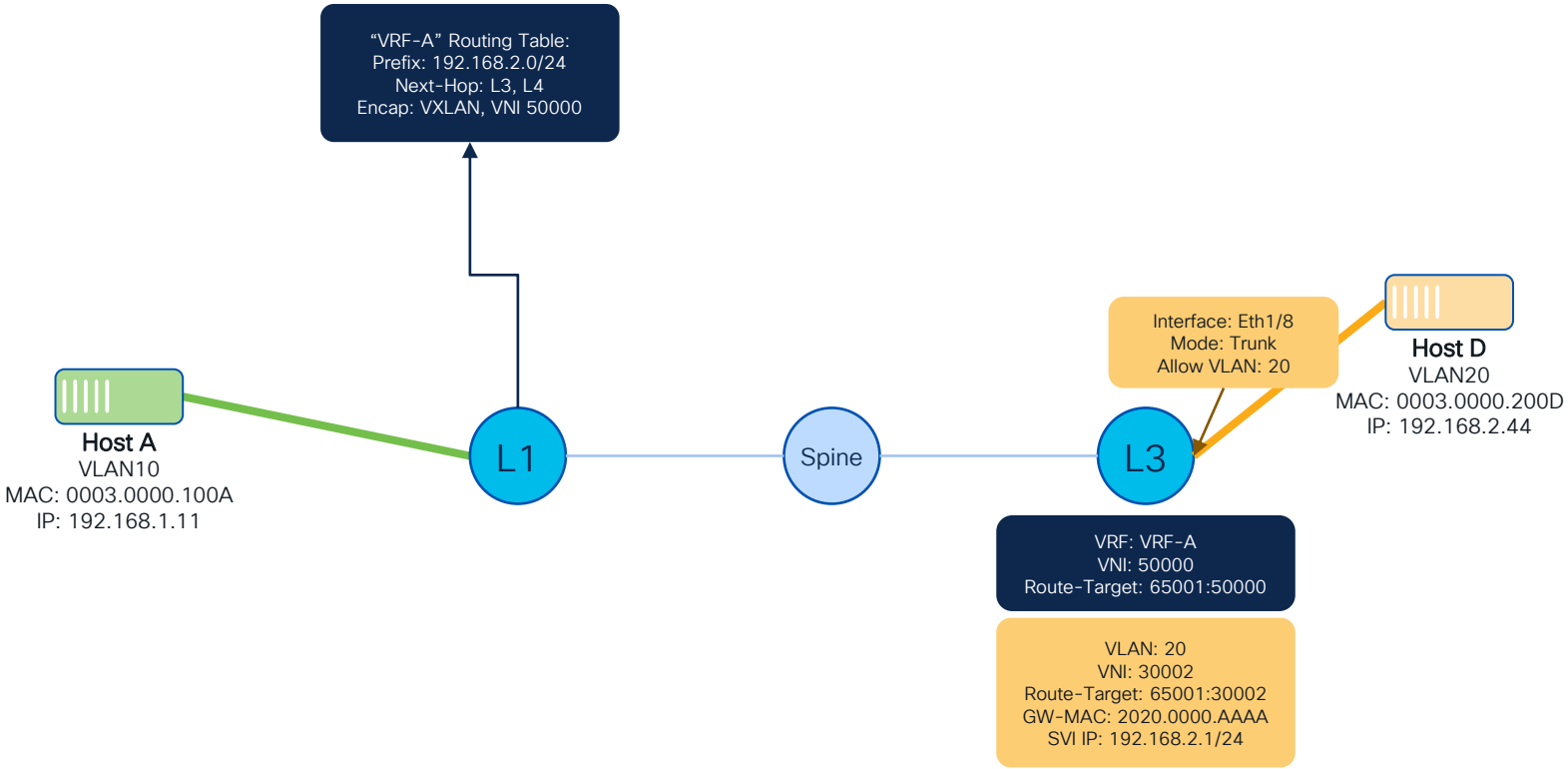
SIP	DIP	VNI	SMAC	DMAC	SIP	DIP	Payload
L4-IP	239.1.1.1	30002	2020.0000.AAAA	FFFF.FFFF.FFFF	192.168.2.1	192.168.2.44	ARP Request

HostA to HostD

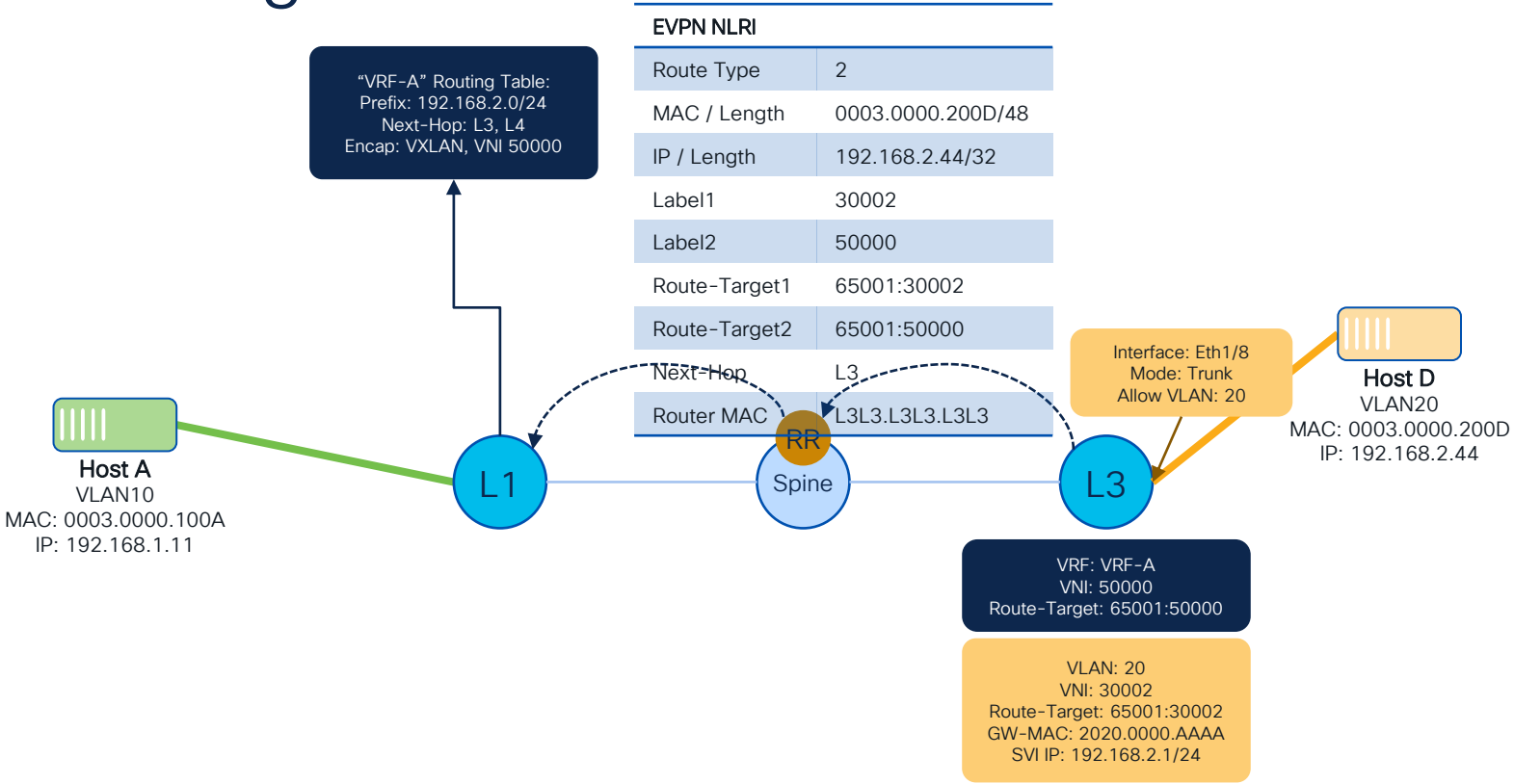
SMAC	DMAC	VLAN	SIP	DIP	Payload
0003.0000.200D	2020.0000.AAAA	20	192.168.2.44	192.168.2.1	ARP Response



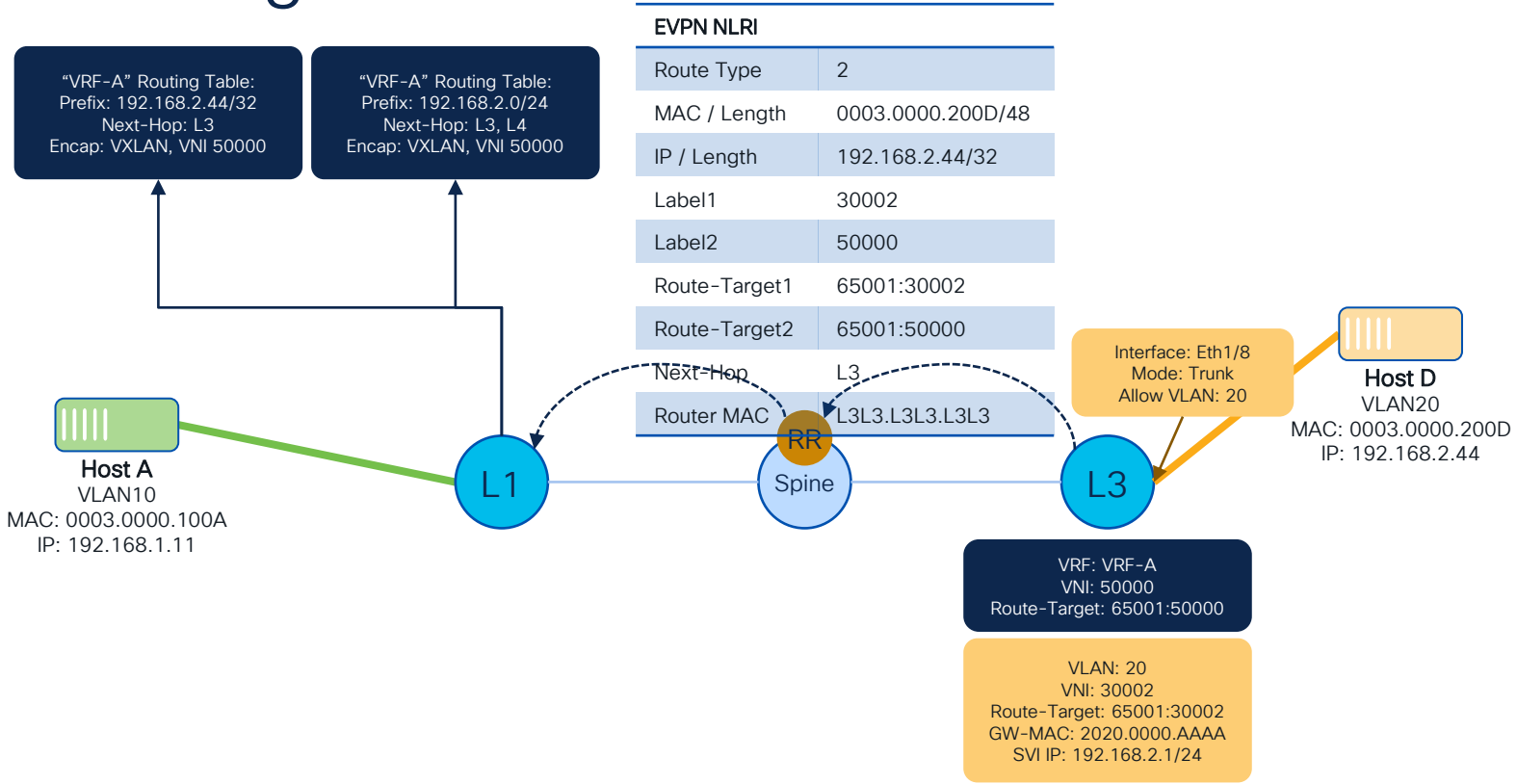
Learning: HostD to Leaf1



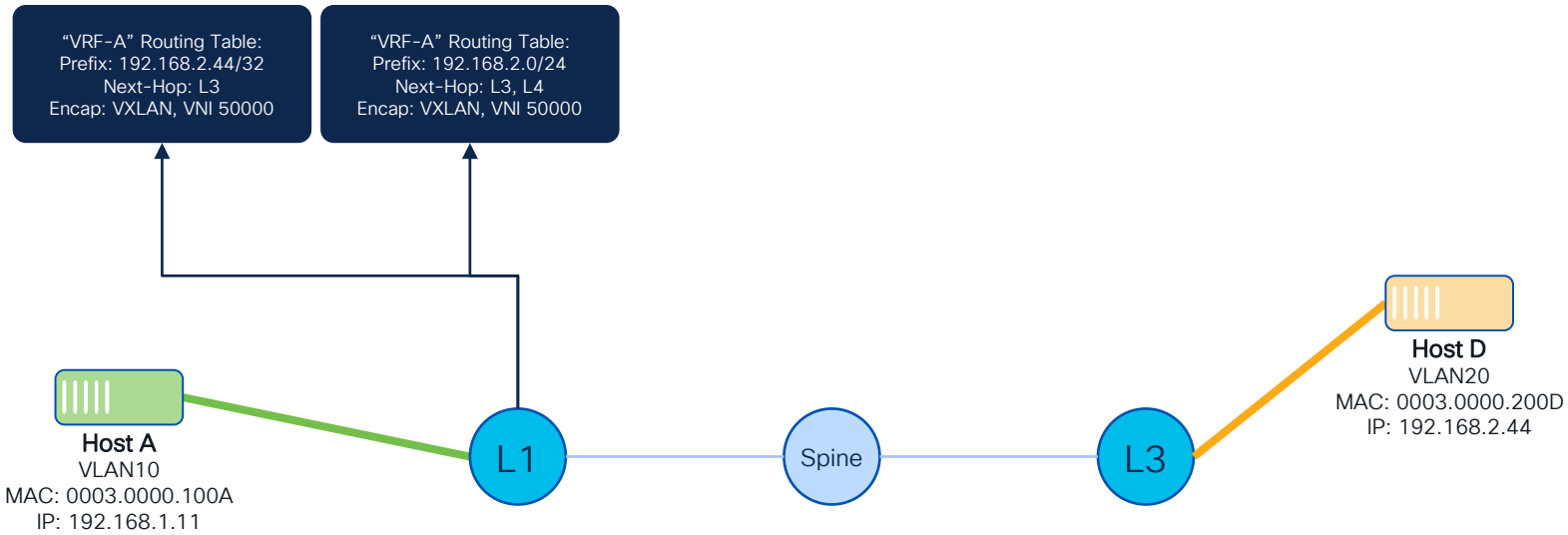
Learning: HostD to Leaf1



Learning: HostD to Leaf1

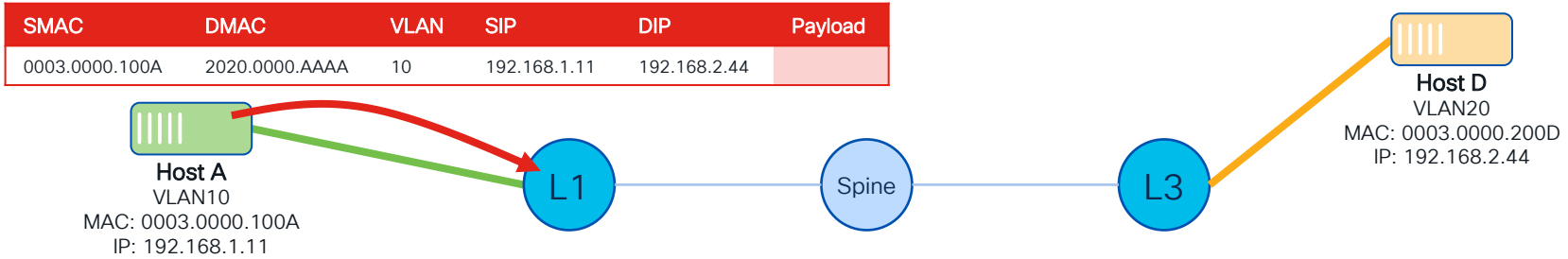


Forwarding Tables



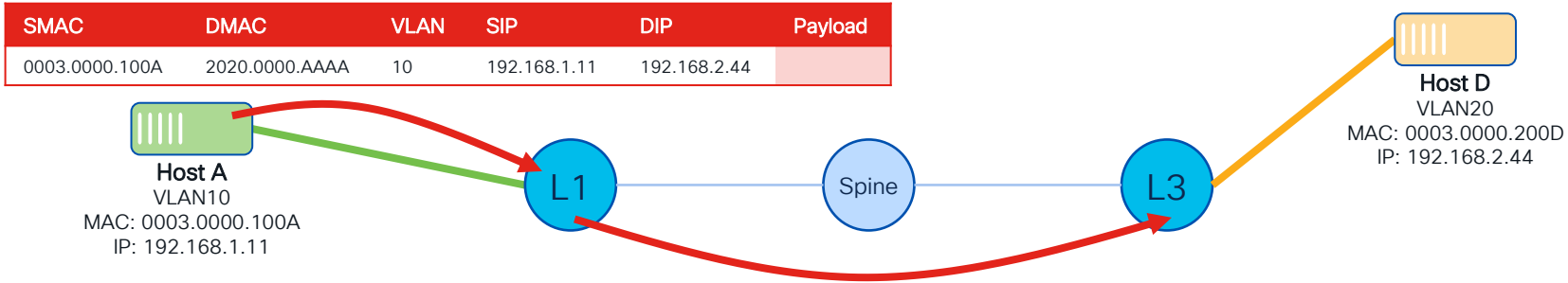
HostA to HostD

SMAC	DMAC	VLAN	SIP	DIP	Payload
0003.0000.100A	2020.0000.AAAA	10	192.168.1.11	192.168.2.44	



HostA to HostD

SMAC	DMAC	VLAN	SIP	DIP	Payload
0003.0000.100A	2020.0000.AAAA	10	192.168.1.11	192.168.2.44	

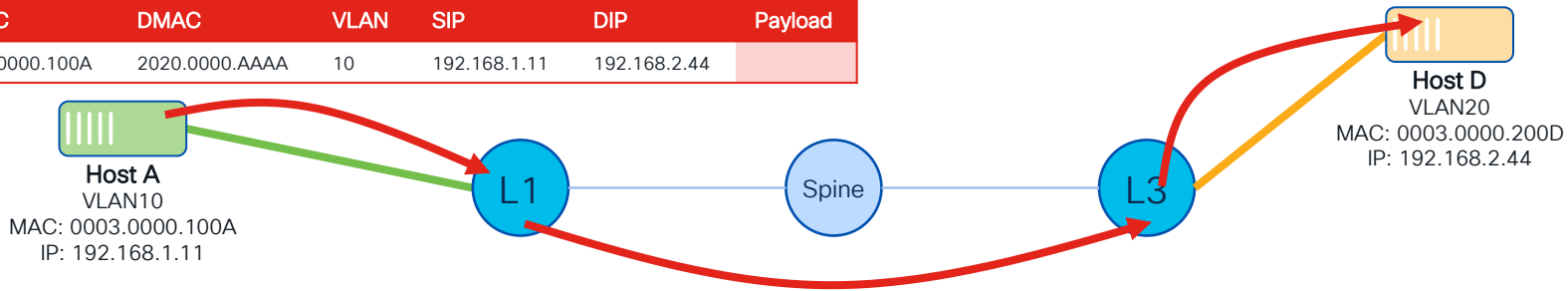


SIP	DIP	VNI	SMAC	DMAC	SIP	DIP	Payload
L1-IP	L3-IP	50000	L1-RMAC	L3-RMAC	192.168.1.11	192.168.2.44	

HostA to HostD

SMAC	DMAC	VLAN	SIP	DIP	Payload
0003.0000.100A	2020.0000.AAAA	10	192.168.1.11	192.168.2.44	

SMAC	DMAC	VLAN	SIP	DIP	Payload
2020.0000.AAAA	0003.0000.200D	20	192.168.1.11	192.168.2.44	



SIP	DIP	VNI	SMAC	DMAC	SIP	DIP	Payload
L1-IP	L3-IP	50000	L1-RMAC	L3-RMAC	192.168.1.11	192.168.2.44	

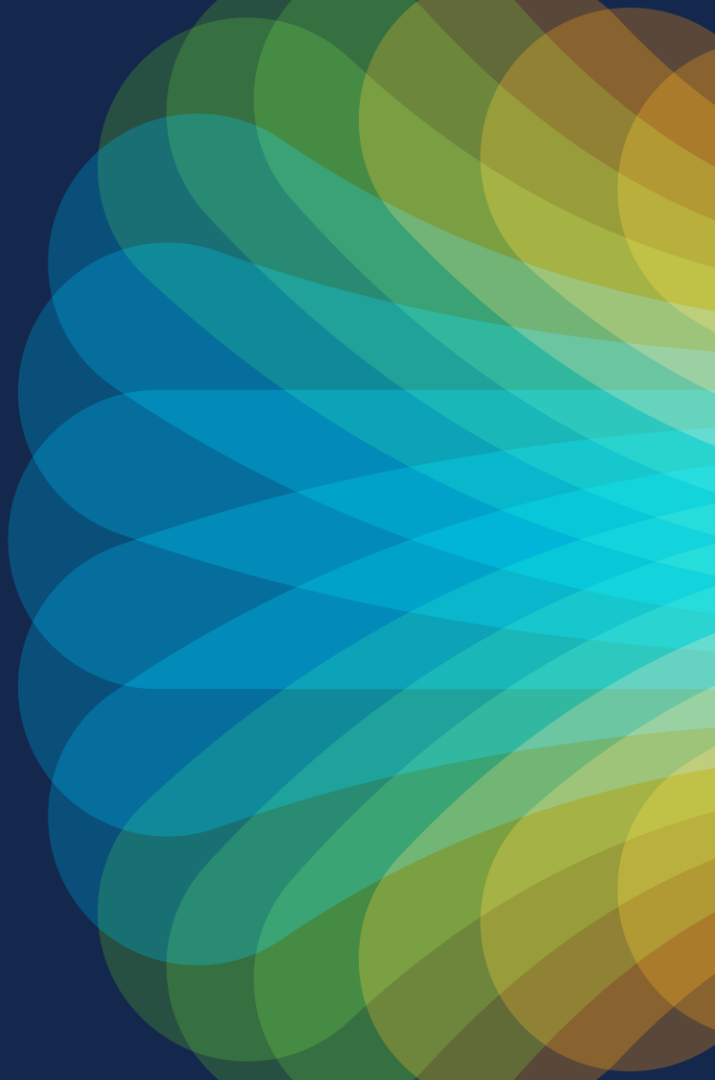
Conclusion

Conclusion

- Did you have enough Packet Walks?
- We covered
  - Host to External Network (RT-5 based routing)
  - Host to Host in different Subnet (RT-2 based routing)
  - Host to Host in same Subnet (RT-2 based bridging)
  - BUM – Broadcast, Unknown Unicast and Multicast (bridged)
    - We looked at Ingress / Head-End Replication and Multicast
    - Note: EVPN works well with BUM forwarding in Multicast (efficiency)
  - Silent Host Discovery (integrated forward and learn)

The bridge to possible

Thank you



cisco Live!


cisco Live!
Let's go