



# VXLAN L2-EVPN Guide for SONiC in GNS3

# Revision History

Revision No.	Description	Editor	Date
1.0	VXLAN L2-EVPN Guide for SONiC in GNS3	Ghulam Bahoo Hafiz Mati Ur Rehman	Jun 9, 2023

# Table of Contents

Introduction	2
Intended Audience	3
VXLAN EVPN	3
Testbed	4
Network Topology	4
Configurations	5
Spine-1	5
Sonic Native Configurations	5
FRR Configurations	6
Leaf-1	8
SONiC Native Configurations	8
FRR configurations	10
Results & Verification	12
FRR Verifications for Leaf-1	12
Check IP routes	12
Check bgp neighborship	13
Check VXLAN VNI Status	14
Check the evpn mac learning	14
SONiC Native Configurations	15
Check vxlan interface configuration	15
Check vxlan and VLAN mapping	15
Check the status for Vxlan tunneling	16
Check the Mac learning	16
Ping from PC1 to PC3 (Within Vlan10)	17
References	18

# Introduction

This comprehensive document serves as a guide for configuring VXLAN L2-EVPN using GNS3, a powerful network simulation tool. By leveraging GNS3, users can create virtual instances of SONiC, enabling thorough testing and evaluation of its diverse features. Through practical demonstrations and detailed instructions, the document aims to equip readers with the knowledge and insights required to successfully configure VXLAN in their network environments. The step-by-step procedure outlined in this guide provides a comprehensive approach for deployment and verification of VXLAN L2 EVPN in a Spine-Leaf topology.

## Intended Audience

This document is tailored for data center experts, system integrators, and network engineers who are interested in implementing **VXLAN L2-EVPN** in SONiC. It is designed for individuals with a solid understanding of networking principles, including L2 and L3 protocols. Whether you are a hyperscaler, network operator, or vendor, this document aims to provide you with practical, step-by-step guidance, and best practices for deploying, configuring, and conducting behavioral testing of SONiC's L2 and L3 features using the GNS3 network simulation tool.

## VXLAN EVPN

**VXLAN EVPN** leverages VXLAN encapsulation to extend Layer 2 Ethernet networks over Layer 3 IP networks. It uses a network overlay approach, where tenant-specific Layer 2 segments are encapsulated within VXLAN packets and transported over the underlying IP infrastructure. This allows for the creation of virtualized Layer 2 networks that can span across multiple physical switches, enabling seamless workload mobility and network scalability.

**EVPN**, on the other hand, provides a control plane mechanism for distributing MAC (Media Access Control) and IP routing information across the VXLAN overlay network. It uses the Border Gateway Protocol (BGP) as the control plane protocol to exchange this information between different network devices, facilitating efficient forwarding and enabling features like multi-tenancy, end-to-end Layer 2 and Layer 3 connectivity, and MAC/IP address mobility.

VXLAN in SONiC provides the foundation for building scalable, flexible, and virtualized networks in data center environments. Its applications range from network virtualization and workload mobility to traffic segmentation and integration with other network overlays, enabling administrators to create efficient and secure network architectures.

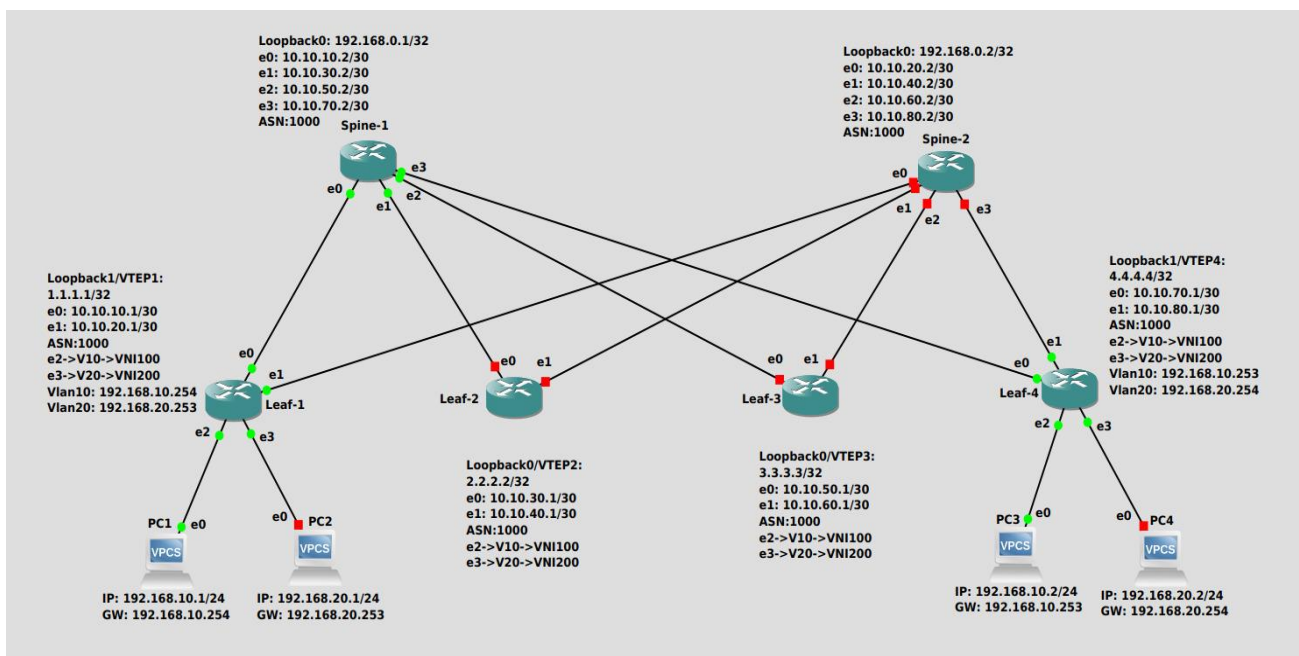
Overall, **VXLAN EVPN** offers a flexible and scalable solution for building large-scale data center networks with virtualized Layer 2 networks. It provides the necessary capabilities to meet the demands of modern data center architectures, including workload mobility, scalability, multi-tenancy, and efficient traffic forwarding.

## Testbed

To set up the testbed for VXLAN configuration please refer to the document [Installation of GNS3 and vTestbed setup for SONiC](#).

## Network Topology

After importing the image, create a Spine-Leaf topology in GNS3 using SONiC devices and hosts.



In the above topology, 6 switches (Leaf-1, Leaf-2, Leaf-3, Leaf-4, Spine-1, & Spine-2) and four hosts (PC1, PC2, PC3 and PC4) are used. PC1 & PC3 belong to Vlan10 and PC2 & PC4 belong to Vlan20. We want to enable intra-vlan communication between hosts over the Spine-Leaf underlay network.

# Configurations

The configuration consists of two parts i.e. SONiC Native Configurations and FRR Configuration. All hosts and switches are configured for the above topology. We have shown Spine-1 & Leaf-1 configurations only other Spine & Leaf switches can be configured similarly. A [command reference](#) guide is also available on GitHub for SONiC.

## Spine-1

The details of the subnets between Spine-1 and Leaf Switches are as follows:

1. Spine-1 - Leaf 1 -> 10.10.10.0/30
2. Spine-1 - Leaf 2 -> 10.10.30.0/30
3. Spine-1 - Leaf 3 -> 10.10.50.0/30
4. Spine-1 - Leaf 4 -> 10.10.70.0/30

## Sonic Native Configurations

Remove default IP's from Ethernet0, Ethernet4, Ethernet8 and Ethernet12.

- **sudo config interface ip remove Ethernet0 10.0.0.0/31**
- **sudo config interface ip remove Ethernet4 10.0.0.2/31**
- **sudo config interface ip remove Ethernet8 10.0.0.4/31**
- **sudo config interface ip remove Ethernet12 10.0.0.6/31**

Configure IP's on Ethernet0, Ethernet4, Ethernet8 and Ethernet12.

- **sudo config interface ip add Ethernet0 10.10.10.2/30**
- **sudo config interface ip add Ethernet4 10.10.30.2/30**
- **sudo config interface ip add Ethernet8 10.10.50.2/30**
- **sudo config interface ip add Ethernet12 10.10.70.2/30**

```
Spine-1
admin@sonic:~$ sudo config interface ip remove Ethernet0 10.0.0.0/31
admin@sonic:~$ sudo config interface ip remove Ethernet4 10.0.0.2/31
admin@sonic:~$ sudo config interface ip remove Ethernet8 10.0.0.4/31
admin@sonic:~$ sudo config interface ip remove Ethernet12 10.0.0.6/31
admin@sonic:~$ sudo config interface ip add Ethernet0 10.10.10.2/30
admin@sonic:~$ sudo config interface ip add Ethernet4 10.10.30.2/30
admin@sonic:~$ sudo config interface ip add Ethernet8 10.10.50.2/30
admin@sonic:~$ sudo config interface ip add Ethernet12 10.10.70.2/30
admin@sonic:~$
```

## FRR Configurations

We have configured Spine-1 as bgp route-reflector to reflect routes between Leaf switches.

Enter FRR and remove default bgp instance

- **vtysh**
- **no router bgp 65100**

Configure BGP neighbors for ASN-1000

- **router bgp 1000**
- **bgp router-id 192.168.0.1**
- **neighbor 10.10.10.1 remote-as 1000**
- **neighbor 10.10.30.1 remote-as 1000**
- **neighbor 10.10.50.1 remote-as 1000**
- **neighbor 10.10.70.1 remote-as 1000**

Advertise connected networks

- **address-family ipv4**
- **network 192.168.0.1/32**
- **network 10.10.10.2/30**
- **network 10.10.30.2/30**
- **network 10.10.50.2/30**
- **network 10.10.70.2/30**

configure route-reflector for Leaf Switches.

- **neighbor 10.10.10.1 route-reflector-client**
- **neighbor 10.10.70.1 route-reflector-client**

```
Spine-1
File Edit View Search Terminal Help
sonic# con
sonic(config)# no router bgp 65100
sonic(config)# router bgp 1000
sonic(config-router)# bgp router-id 192.168.0.1
sonic(config-router)# neighbor 10.10.10.1 remote-as 1000
sonic(config-router)# neighbor 10.10.30.1 remote-as 1000
sonic(config-router)# neighbor 10.10.50.1 remote-as 1000
sonic(config-router)# neighbor 10.10.70.1 remote-as 1000
sonic(config-router)# address-family ipv4
sonic(config-router-af)# network 192.168.0.1/32
sonic(config-router-af)# network 10.10.10.2/30
sonic(config-router-af)# network 10.10.30.2/30
sonic(config-router-af)# network 10.10.50.2/30
sonic(config-router-af)# neighbor 10.10.10.1 route-reflector-client
sonic(config-router-af)# neighbor 10.10.70.1 route-reflector-client
sonic(config-router-af)# exit
sonic(config-router)# address-family l2vpn evpn
sonic(config-router-af)# neighbor 10.10.10.1 route-reflector-client
sonic(config-router-af)# neighbor 10.10.70.1 route-reflector-client
sonic(config-router-af)# exit
sonic(config-router)# exit
sonic(config)# exit
sonic# wr
Note: this version of vtysh never writes vtysh.conf
Building Configuration...
Configuration saved to /etc/frr/zebra.conf
Configuration saved to /etc/frr/bgpd.conf
Configuration saved to /etc/frr/staticd.conf
sonic#
```



# Leaf-1

## SONiC Native Configurations

Remove default IP's from Ethernet0, Ethernet4, Ethernet8 and Ethernet12.

- **sudo config interface ip remove Ethernet0 10.0.0.0/31**
- **sudo config interface ip remove Ethernet4 10.0.0.2/31**
- **sudo config interface ip remove Ethernet8 10.0.0.4/31**
- **sudo config interface ip remove Ethernet12 10.0.0.6/31**

Configure IP's on Ethernet0, Ethernet4 and Loopback1.

- **sudo config interface ip add Ethernet0 10.10.10.1/30**
- **sudo config interface ip add Ethernet4 10.10.20.1/30**
- **sudo config interface ip add Loopback1 1.1.1.1/32**

```
Unauthorized access and/or use are prohibited.
All access and/or use are subject to monitoring.

Help:   https://sonic-net.github.io/SONiC/

admin@sonic:~$
admin@sonic:~$
admin@sonic:~$
admin@sonic:~$
admin@sonic:~$
admin@sonic:~$
admin@sonic:~$
admin@sonic:~$ sudo config interface ip remove Ethernet0 10.0.0.0/31
admin@sonic:~$ sudo config interface ip remove Ethernet4 10.0.0.2/31
admin@sonic:~$ sudo config interface ip remove Ethernet8 10.0.0.4/31
admin@sonic:~$ sudo config interface ip remove Ethernet12 10.0.0.6/31
admin@sonic:~$ sudo config interface ip add Ethernet0 10.10.10.1/30
admin@sonic:~$ sudo config interface ip add Ethernet4 10.10.20.1/30
admin@sonic:~$ sudo config interface ip add Loopback1 1.1.1.1/32
```

According to the Network Topology diagram, VLAN 10 is configured on Ethernet8 while VLAN 20 is configured on Ethernet12.

- **sudo config vlan add 10**
- **sudo config add vlan 20**
- **sudo config vlan mem add -u 10 Ethernet8**
- **sudo config vlan mem add -u 20 Ethernet12**

Create VXLAN Interface

- **sudo config vxlan add vtep 1.1.1.1**

Create a NVO and bind to VXLAN interface.

- **sudo config vxlan evpn\_nvo add nvo vtep**

VLAN10 is mapped to VNI100 and VLAN20 is mapped to VNI200.

- **sudo config vclan map add vtep 10 100**
- **sudo config vxlan map add vtep 20 200**

```
admin@sonic:~$ sudo config vlan add 10
admin@sonic:~$ sudo config vlan add 20
admin@sonic:~$ sudo config vlan mem add -u 10 Ethernet8
admin@sonic:~$ sudo config vlan mem add -u 20 Ethernet12
admin@sonic:~$ sudo config vxlan add vtep 1.1.1.1
admin@sonic:~$ sudo config vxlan evpn_nvo add nvo vtep
admin@sonic:~$ sudo config vxlan map add vtep 10 100
admin@sonic:~$ sudo config vxlan map add vtep 20 200
admin@sonic:~$ sudo config save -y
Running command: /usr/local/bin/sonic-cfggen -d --print-data > /etc/sonic/config_db.json
admin@sonic:~$
```

## FRR configurations

Enter FRR and remove default bgp instance

- **vtys**
- **configure terminal**
- **no router bgp 65100**

Configure BGP neighbors for ASN-1000

- **router bgp 1000**
- **bgp router-id 1.1.1.1**
- **neighbor 10.10.10.2 remote-as 1000**
- **neighbor 10.10.20.2 remote-as 1000**
- **neighbor 10.10.70.1 remote-as 1000**
- **neighbor 10.10.70.1 update-source 10.10.10.1**

Advertise connected networks

- **address-family ipv4**
- **network 1.1.1.1/32**
- **exit**

Create a L2vpn neighbor

- **address-family l2vpn evpn**
- **neighbor 10.10.70.1**
- **advertise-all-vni**

```
sonic# conf t
sonic(config)# no router bgp 65100
sonic(config)# router bgp 1000
sonic(config-router)# bgp router-id 1.1.1.1
sonic(config-router)# neighbor 10.10.10.2 remote-as 1000
sonic(config-router)# neighbor 10.10.20.2 remote-as 1000
sonic(config-router)# neighbor 10.10.70.1 remote-as 1000
sonic(config-router)# neighbor 10.10.70.1 update-source 10.10.10.1
sonic(config-router)# address-family ipv4
sonic(config-router-af)# network 1.1.1.1/32
sonic(config-router-af)# exit
sonic(config-router)# address-family l2vpn evpn
sonic(config-router-af)# neighbor 10.10.70.1 activate
sonic(config-router-af)# advertise-all-vni
sonic(config-router-af)# exit
sonic(config-router)# exit
sonic(config)# exit
sonic# wr
Note: this version of vtysh never writes vtysh.conf
Building Configuration...
Configuration saved to /etc/frr/zebra.conf
Configuration saved to /etc/frr/bgpd.conf
Configuration saved to /etc/frr/staticd.conf
sonic#
```

# Results & Verification

## FRR Verifications for Leaf-1

Check IP routes

- Show ip bgp

```
admin@sonic:~$ vtysh
Hello, this is FRRouting (version 8.2.2).
Copyright 1996-2005 Kunihiro Ishiguro, et al.

sonic# show ip bgp
BGP table version is 8, local router ID is 1.1.1.1, vrf id 0
Default local pref 100, local AS 1000
Status codes:  s suppressed, d damped, h history, * valid, > best, = multipath,
                i internal, r RIB-failure, S Stale, R Removed
Nexthop codes: @NNN nexthop's vrf id, < announce-nh-self
Origin codes:  i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found

   Network          Next Hop        Metric LocPrf Weight Path
*> 1.1.1.1/32       0.0.0.0         0      32768  i
*>i4.4.4.4/32       10.10.70.1      0      100    0  i
* i                 10.10.70.1      0      100    0  i
*>i10.10.10.0/30    10.10.10.2      0      100    0  i
*>i10.10.30.0/30    10.10.10.2      0      100    0  i
*>i10.10.50.0/30    10.10.10.2      0      100    0  i
*>i10.10.70.0/30    10.10.10.2      0      100    0  i
*>i192.168.0.1/32  10.10.10.2      0      100    0  i

Displayed 7 routes and 8 total paths
sonic#
```

We can see routes from Leaf-4 and Spine-1 in the above screenshot.

## Check bgp neighborship

- **show bgp summary**

```
Leaf-1
admin@sonic:~$ vtysh
Hello, this is FRRouting (version 8.2.2).
Copyright 1996-2005 Kunihiro Ishiguro, et al.

sonic# show bgp summary

IPv4 Unicast Summary (VRF default):
BGP router identifier 1.1.1.1, local AS number 1000 vrf-id 0
BGP table version 8
RIB entries 13, using 2392 bytes of memory
Peers 3, using 2170 KiB of memory

Neighbor      V      AS  MsgRcvd  MsgSent  TblVer  InQ  OutQ  Up/Down  Sta
te/PfxRcd    PfxSnt Desc
10.10.10.2    4      1000     11       9        0     0     0 00:05:03
      6      1 N/A
10.10.20.2    4      1000      0       0        0     0     0  never
      Active  0 N/A
10.10.70.1    4      1000     16      17        0     0     0 00:05:02
      1      1 N/A

Total number of neighbors 3

L2VPN EVPN Summary (VRF default):
BGP router identifier 1.1.1.1, local AS number 1000 vrf-id 0
BGP table version 0
RIB entries 7, using 1288 bytes of memory
Peers 1, using 723 KiB of memory

Neighbor      V      AS  MsgRcvd  MsgSent  TblVer  InQ  OutQ  Up/Down  Sta
te/PfxRcd    PfxSnt Desc
10.10.70.1    4      1000     16      17        0     0     0 00:05:02
      2      3 N/A

Total number of neighbors 1
sonic#
```

We can see two active bgp neighbors Leaf-4 and Spine-1. Other neighbors are not active as devices are shut-down.

## Check VXLAN VNI Status

- **show evpn vni**

```
Leaf-1
admin@sonic:~$ vtysh
Hello, this is FRRouting (version 8.2.2).
Copyright 1996-2005 Kunihiro Ishiguro, et al.

sonic# show evpn vni
VNI      Type VxLAN IF      # MACs  # ARPs  # Remote VTEPs  Tenant V
RF
200      L2   vtep-20      0        1        1                default
100      L2   vtep-10      1        1        1                default
```

## Check the evpn mac learning

- **show evpn mac vni all**

```
Leaf-1
admin@sonic:~$ vtysh
Hello, this is FRRouting (version 8.2.2).
Copyright 1996-2005 Kunihiro Ishiguro, et al.

sonic# show evpn mac vni all
VNI 100 #MACs (local and remote) 3

Flags: N=sync-neighs, I=local-inactive, P=peer-active, X=peer-proxy
MAC      Type  Flags Intf/Remote ES/VTEP      VLAN  Seq #'s
00:50:79:66:68:02 local   Ethernet8      10    0/0
00:50:79:66:68:00 remote  4.4.4.4        10    0/0
33:33:00:00:00:01 local   Ethernet8      10    0/0
sonic#
```

We can see local and remote hosts' mac addresses.

## SONiC Native Configurations

### Check vxlan interface configuration

- **show vxlan interface**

```
Leaf-1
admin@sonic:~$ show vxlan interface
VTEP Information:
    VTEP Name : vtep, SIP : 1.1.1.1
    NVO Name  : nvo,  VTEP : vtep
    Source interface : Loopback1
admin@sonic:~$
```

### Check vxlan and VLAN mapping

- **show vxlan vlanvni**

```
Leaf-1
admin@sonic:~$ show vxlan vlanvni
+-----+-----+
| VLAN  |  VNI  |
+=====+=====+
| Vlan10 |  100  |
+-----+-----+
| Vlan20 |  200  |
+-----+-----+
Total count : 2
admin@sonic:~$
```



## Check the status for Vxlan tunneling

- **show vxlan remotevtep**

```
Leaf-1
admin@sonic:~$ show vxlan remotevtep
+-----+-----+-----+-----+
| SIP    | DIP    | Creation Source | OperStatus |
+=====+=====+=====+=====+
| 1.1.1.1 | 4.4.4.4 | EVPN            | oper_down  |
+-----+-----+-----+-----+
Total count : 1
admin@sonic:~$
```

- **show vxlan tunnel**

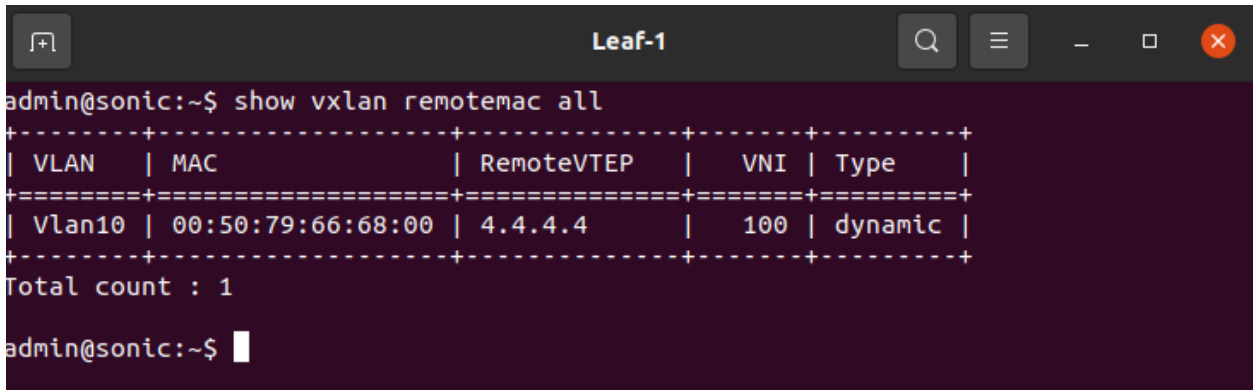
```
Leaf-1
admin@sonic:~$ show vxlan tunnel
vxlan tunnel name      source ip      destination ip  tunnel map name  tunnel ma
p mapping(vni -> vlan)
-----
vtep
an10                   1.1.1.1      map_100_Vlan10 100 -> Vl
an20                   map_200_Vlan20 200 -> Vl
admin@sonic:~$
```

## Check the Mac learning

- **show mac**

```
Leaf-1
admin@sonic:~$ show mac
No.   Vlan  MacAddress      Port      Type
-----
1     10   00:50:79:66:68:02 Ethernet8  Dynamic
2     10   33:33:00:00:00:01 Ethernet8  Dynamic
Total number of entries 2
admin@sonic:~$
```

- `show vxlan remotemac all`

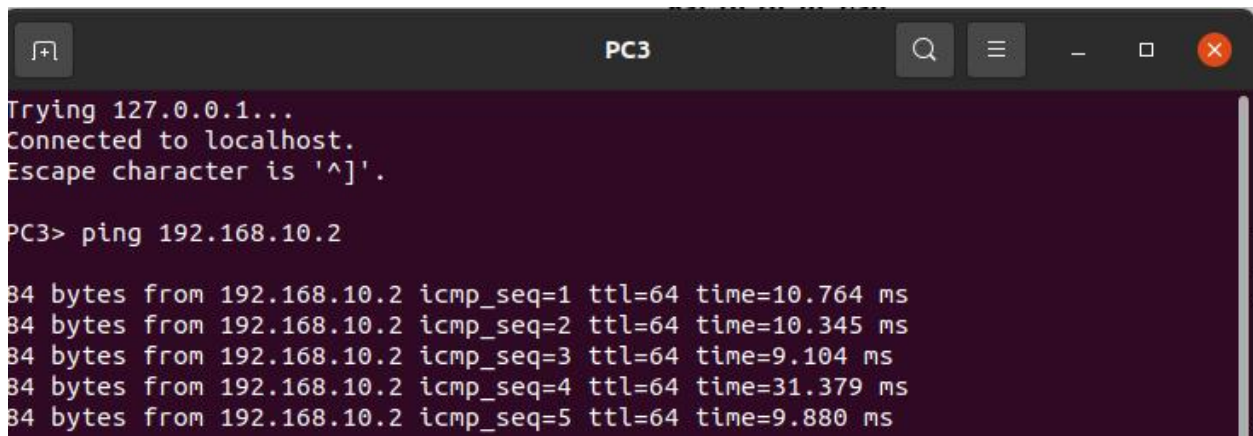


```
admin@sonic:~$ show vxlan remotemac all
+-----+-----+-----+-----+-----+
| VLAN  | MAC                | RemoteVTEP | VNI  | Type   |
+=====+=====+=====+=====+=====+
| Vlan10 | 00:50:79:66:68:00 | 4.4.4.4    | 100  | dynamic |
+-----+-----+-----+-----+-----+
Total count : 1
admin@sonic:~$
```

We see the MAC of the remote host (PC3) on Leaf-1.

### Ping from PC1 to PC3 (Within Vlan10)

The ping result displayed in the image illustrates the communication between host PC1 and host PC3, both of which belong to VLAN-10. At Leaf-1 and Leaf-4, VLAN-10 is associated with VNI-100, and both devices act as VTEPs (Virtual Tunnel Endpoints). When a packet is generated from PC3, it is first received by Leaf-1. Leaf-1 maps VLAN-10 to VNI-100 and forwards the packet accordingly. Subsequently, the packet traverses the Spine-Leaf Underlay network via VXLAN L2 EVPN tunnel with VNI-100 and reaches Leaf-4. At Leaf-4, the VNI-to-VLAN mapping lookup is performed, and the packet is forwarded to its intended destination Vlan i.e. PC3 in Vlan10.



```
Trying 127.0.0.1...
Connected to localhost.
Escape character is '^]'.
PC3> ping 192.168.10.2

84 bytes from 192.168.10.2 icmp_seq=1 ttl=64 time=10.764 ms
84 bytes from 192.168.10.2 icmp_seq=2 ttl=64 time=10.345 ms
84 bytes from 192.168.10.2 icmp_seq=3 ttl=64 time=9.104 ms
84 bytes from 192.168.10.2 icmp_seq=4 ttl=64 time=31.379 ms
84 bytes from 192.168.10.2 icmp_seq=5 ttl=64 time=9.880 ms
```

## References

- <https://docs.frrouting.org/en/latest/bgp.html>
- <https://support.edge-core.com/hc/en-us/articles/13292640639385-Advanced-Setting-EV-PN-VXLAN>