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

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.

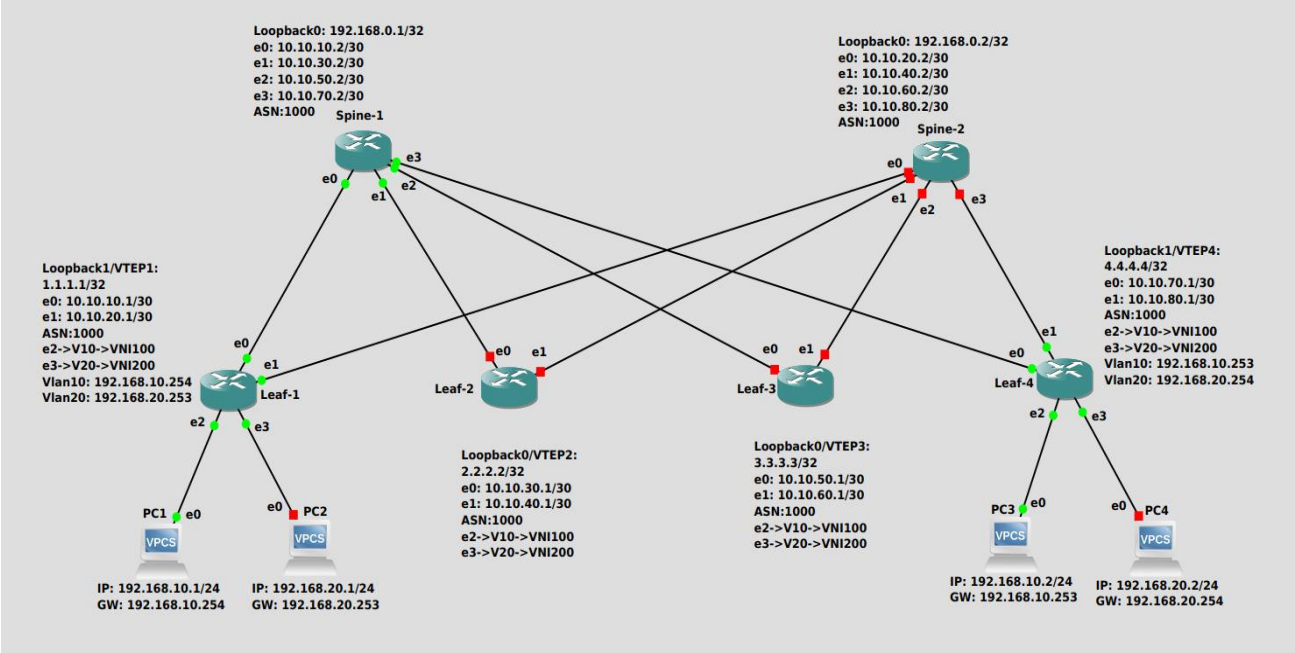
Overall, **VXLAN EVPN** offers a flexible and scalable solution for building large-scale data center networks with virtualized Layer 2 networks. It provides the necessary capabilities to meet the demands of modern data center architectures, including workload mobility, scalability, multi-tenancy, and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

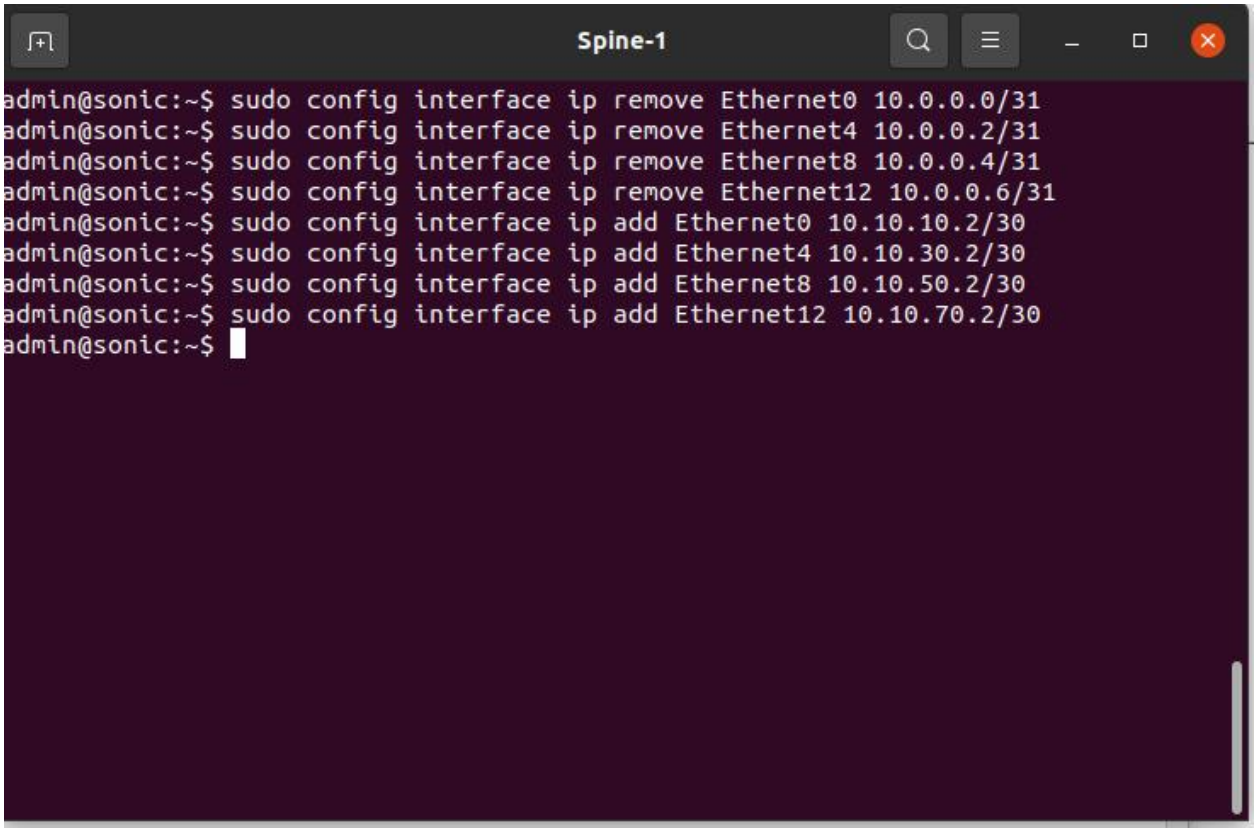
Sonic Native Configurations

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**



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

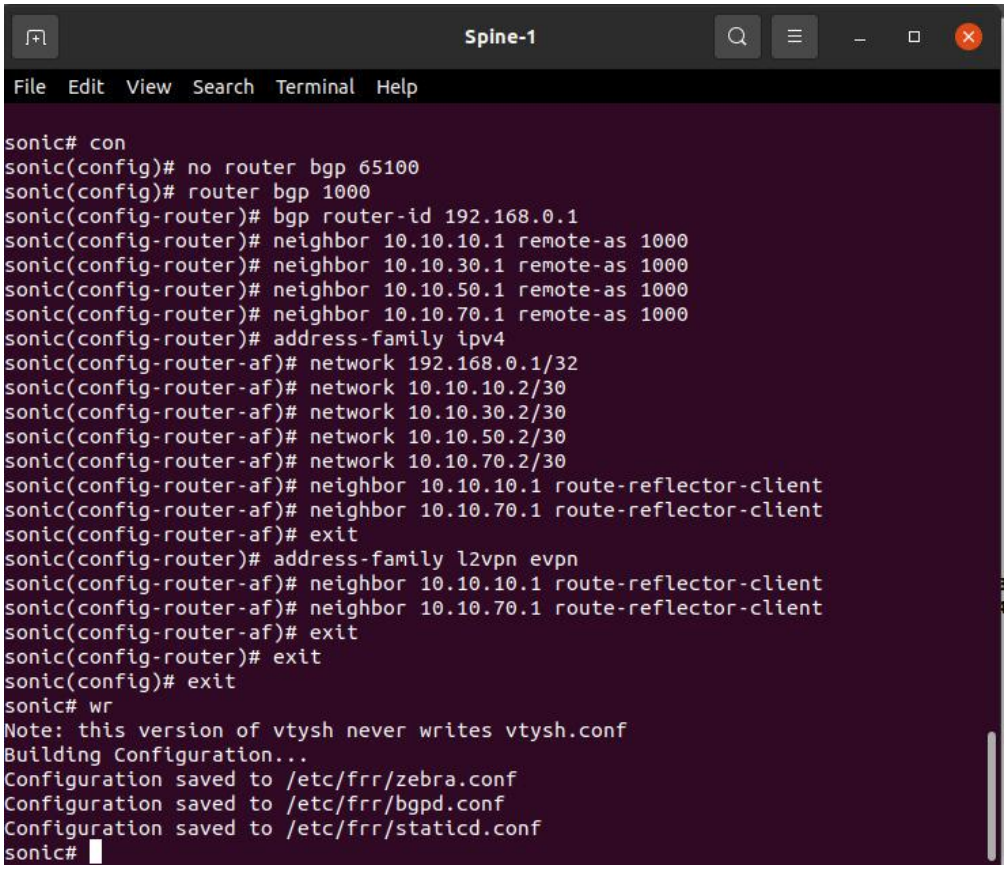
- **router bgp 1000**
- **bgp router-id 192.168.0.1**
- **neighbor 10.10.10.1 remote-as 1000**
- **neighbor 10.10.30.1 remote-as 1000**
- **neighbor 10.10.50.1 remote-as 1000**
- **neighbor 10.10.70.1 remote-as 1000**

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 router-reflector-client**
- **neighbor 10.10.70.1 router-reflector-client**



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:~$ 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

- **vtysh**
- **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**

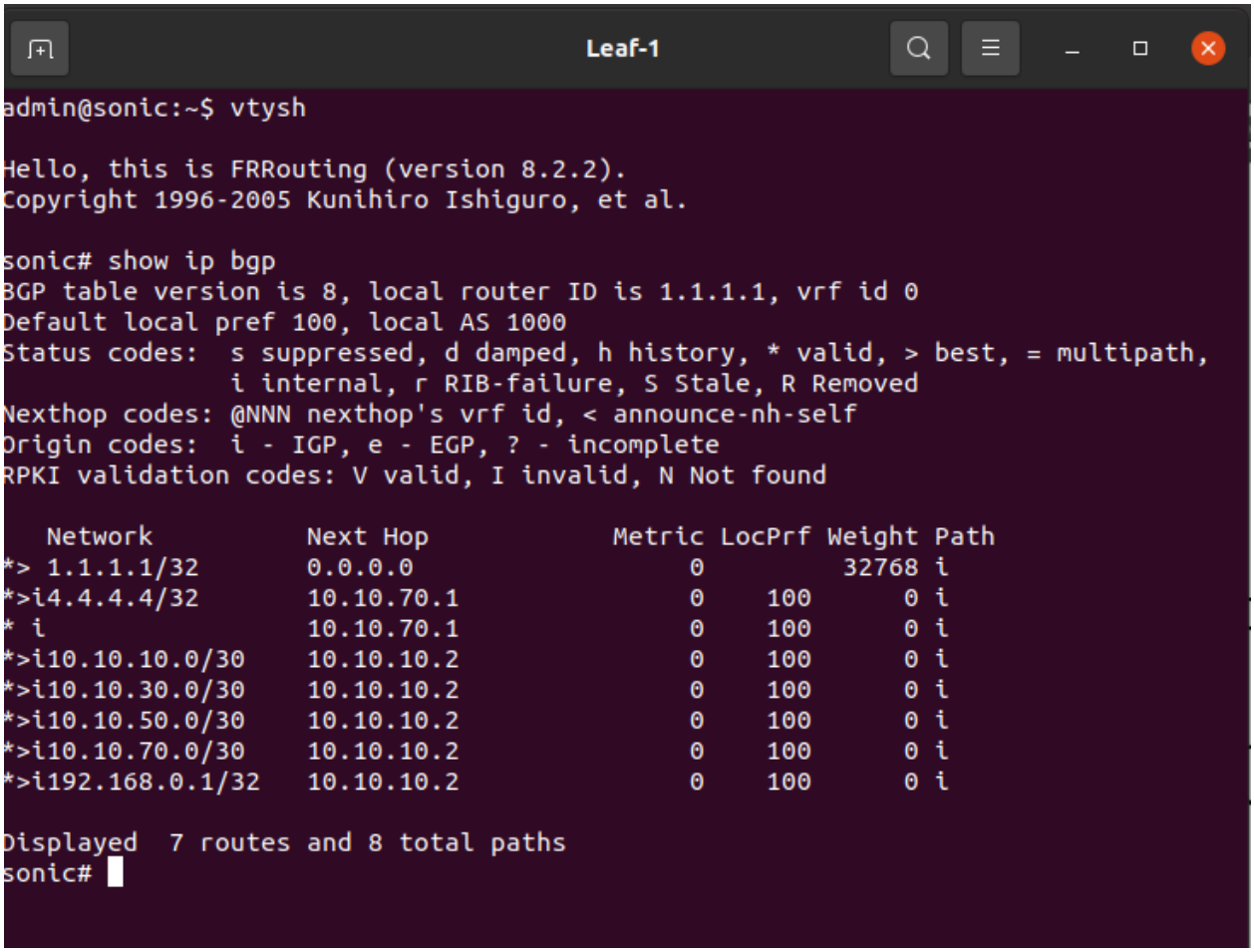
```
sonic# conf t
sonic(config)# no router bgp 65100
sonic(config)# router bgp 1000
sonic(config-router)# bgp router-id 1.1.1.1
sonic(config-router)# neighbor 10.10.10.2 remote-as 1000
sonic(config-router)# neighbor 10.10.20.2 remote-as 1000
sonic(config-router)# neighbor 10.10.70.1 remote-as 1000
sonic(config-router)# 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

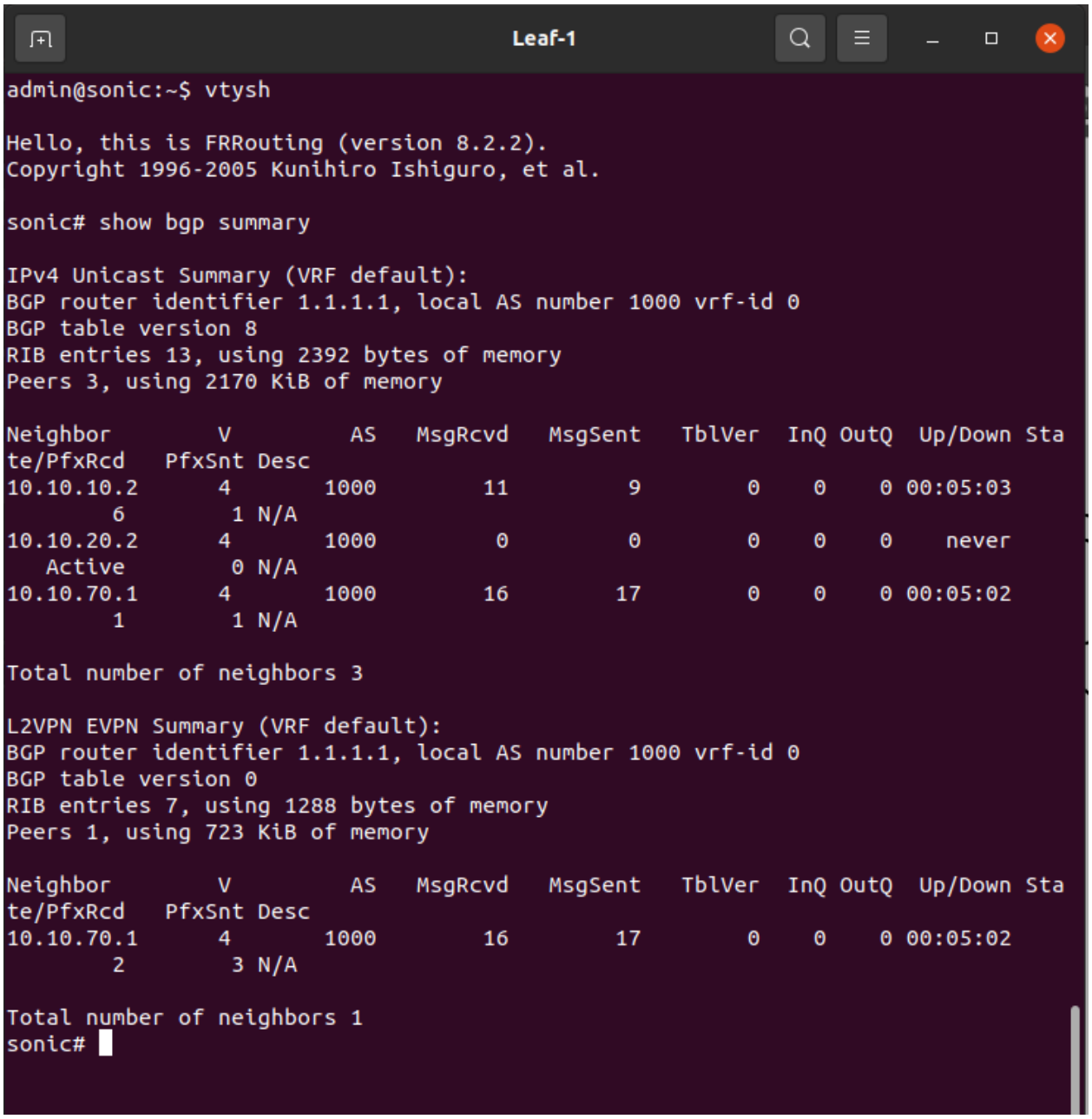
- **Show ip bgp**



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

Check bgp neighborship

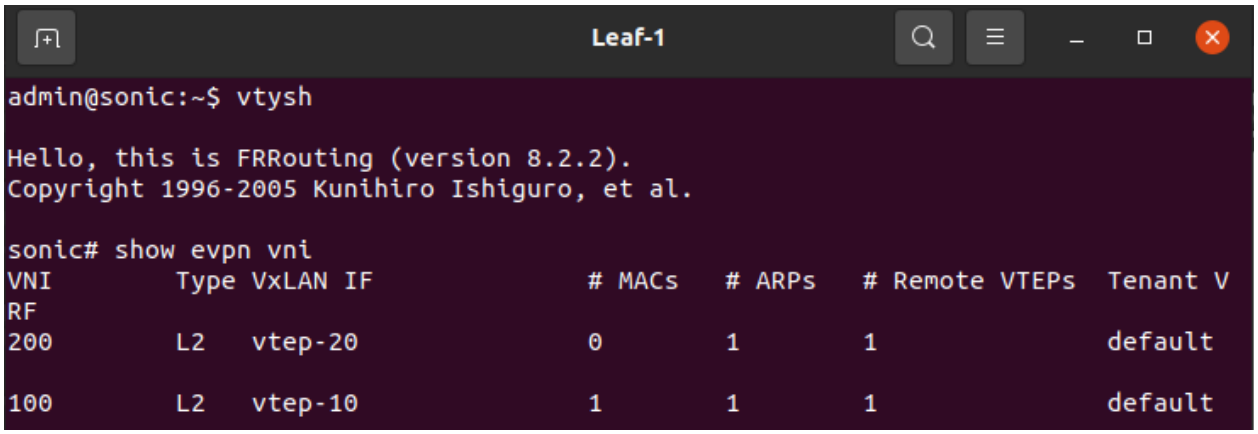
- **show bgp summary**



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

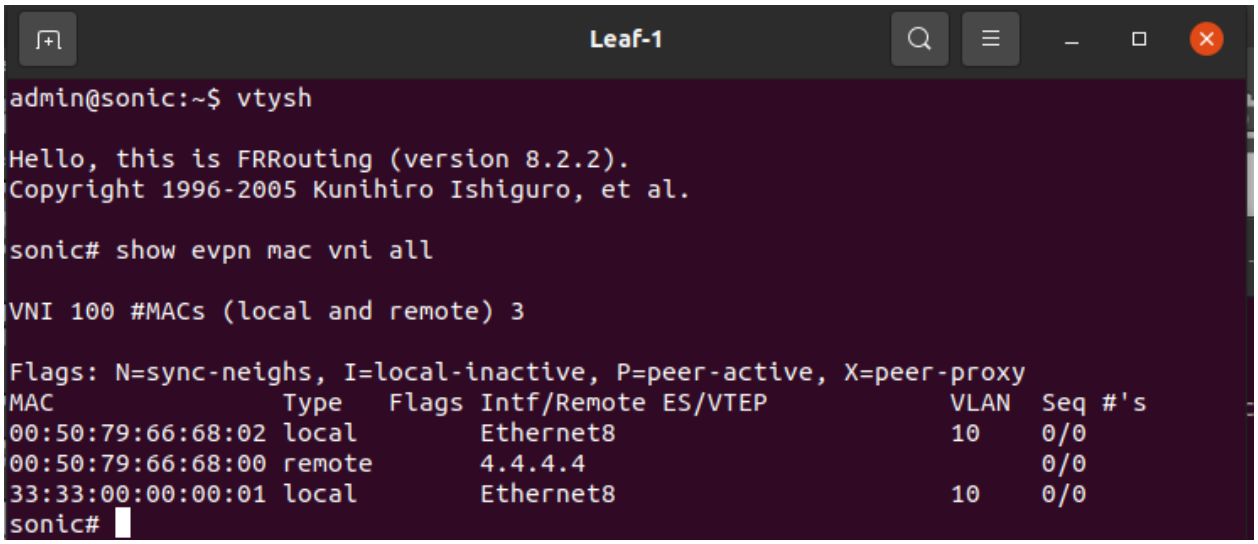
Check VXLAN VNI Status

- **show evpn vni**



Check the evpn mac learning

- **show evpn mac vni all**

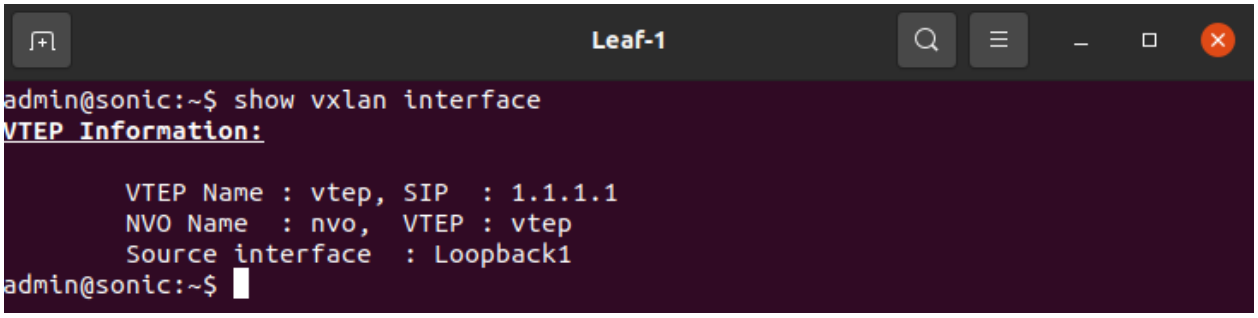


We can see local and remote hosts' mac addresses.

SONiC Native Configurations

Check vxlan interface configuration

- **show vxlan interface**

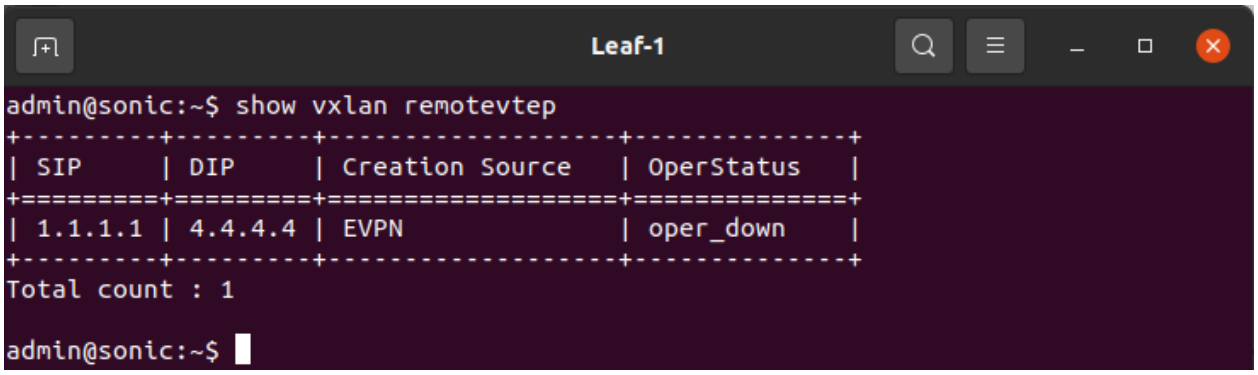


Check vxlan and VLAN mapping

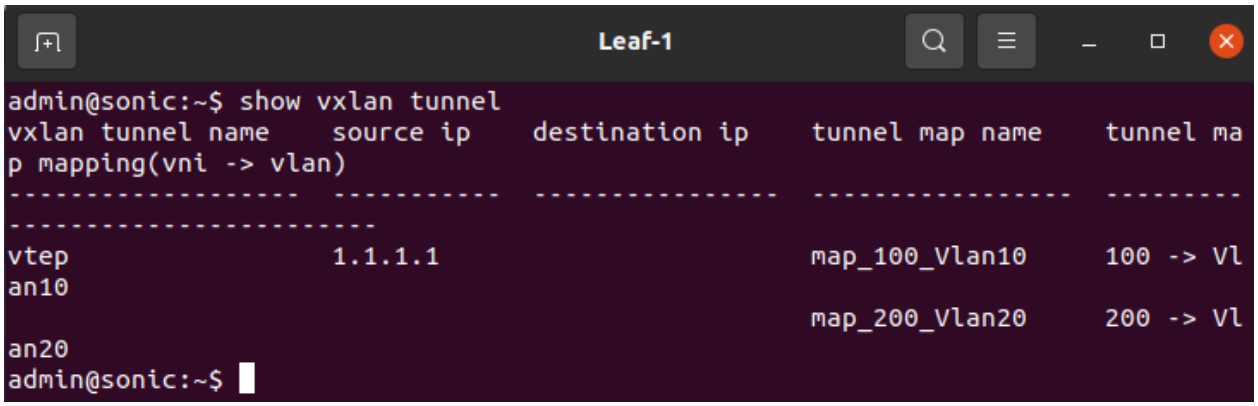
- **show vxlan vlanvnimap**

Check the status for Vxlan tunneling

- **show vxlan remotevtep**

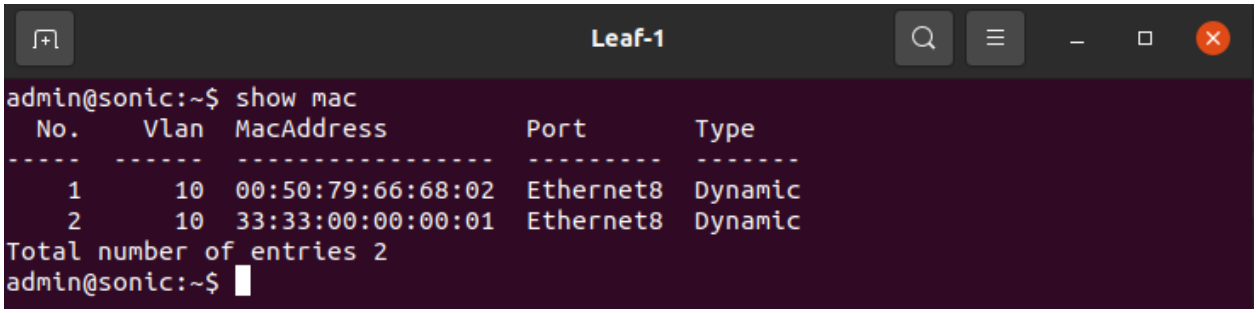


- **show vxlan tunnel**

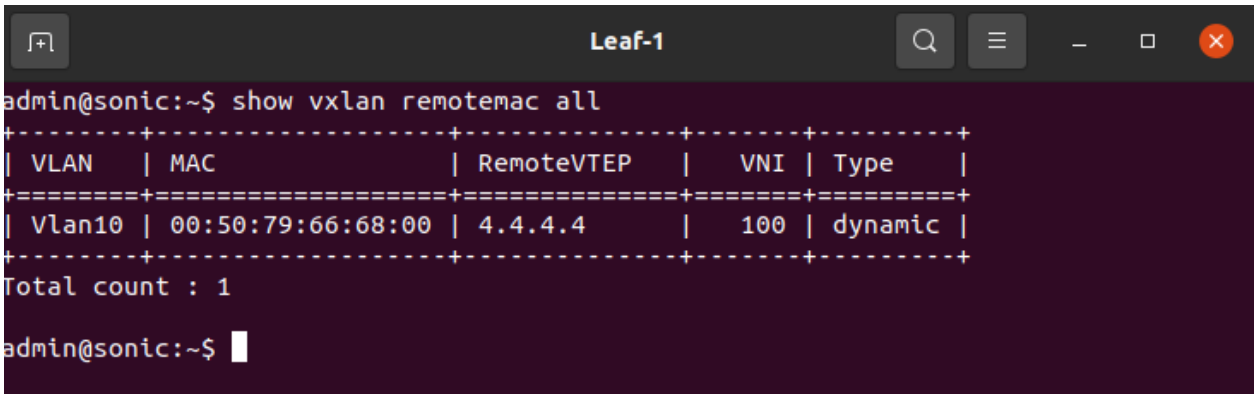


Check the Mac learning

- **show mac**



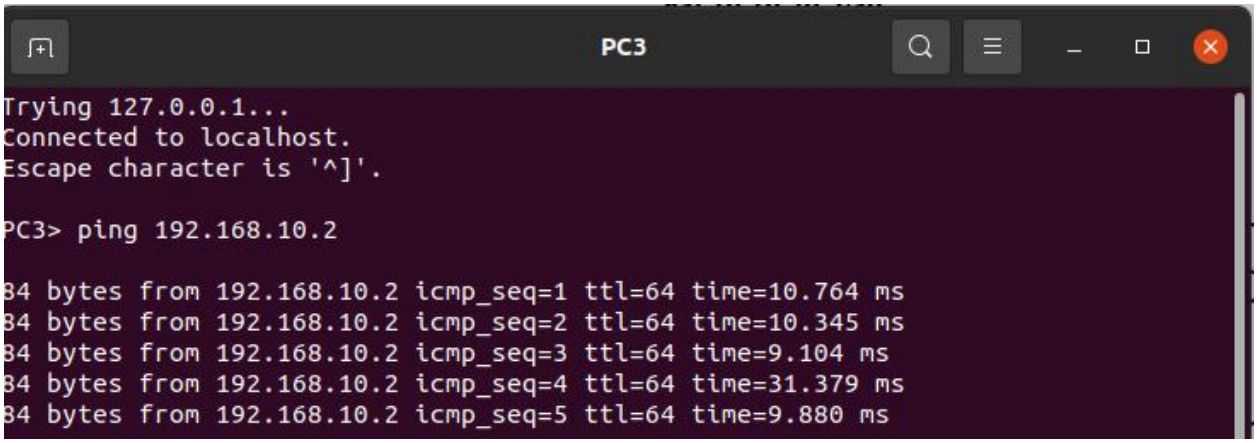
- **show vxlan remotemac all**



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.



References

- *https://docs.frrouting.org/en/latest/bgp.html*
- *https://support.edge-core.com/hc/en-us/articles/13292640639385-Advanced-Setting-EVPN-VXLAN*