VXLAN avançado com vPC: Configuração e verificação de L2VNI e L3VNI

Contents

Introdução

Este documento descreve como configurar um laboratório com switches Nexus 9Kv usando a Virtual eXtensible Local Area Network (VXLAN) com Virtual Port-Channel (vPC).

Pré-requisitos

Requisitos

A Cisco recomenda que você tenha conhecimento destes tópicos:

- Compreensão de roteamento e comutação, bem como a tecnologia Multiprotocol Label Switching (MPLS)
- Experiência com princípios de roteamento multicast, como Ponto de reunião (RP) e Multicast independente de plataforma (PIM)
- Compreensão do Indicador de Família de Endereços (AFI)/Indicador de Família de Endereços Subsequente (SAFI) do Protocolo de Gateway de Borda (BGP)

Componentes Utilizados

Este documento não se restringe a versões de software e hardware específicas.

As informações neste documento foram criadas a partir de dispositivos em um ambiente de laboratório específico. Todos os dispositivos utilizados neste documento foram iniciados com uma configuração (padrão) inicial. Se a rede estiver ativa, certifique-se de que você entenda o impacto potencial de qualquer comando.

Informações de Apoio

O documento também fornece orientação sobre a implantação do laboratório, bem como a verificação de configurações e operações.

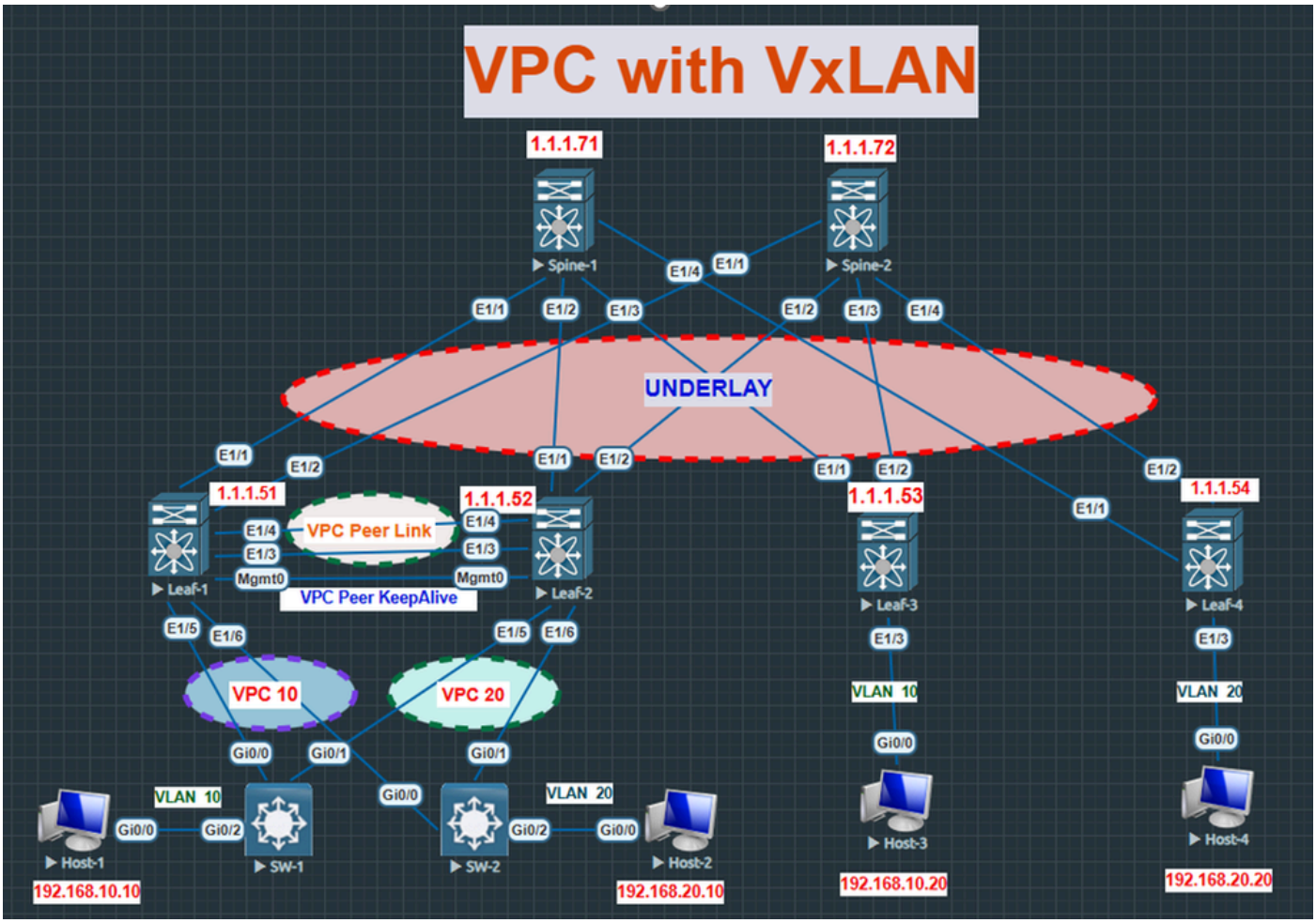
Para este laboratório, o EveNg com switches Nexus 9000V é utilizado para ambos, Leaf e Spine.

Ponto de Extremidade de Túnel Virtual (VTEP)	FOLHA1, FOLHA2, FOLHA3, FOLHA4
vPC	LEAF1 e LEAF2
LEAF1 IP de loopback primário e secundário	Loopback0 - 1.1.1.51, Loopback1 - 10.1.1.100
LEAF2 IP de loopback primário e secundário	Loopback0 - 1.1.1.52, Loopback1 - 10.1.1.100
IP de loopback LEAF3	1.1.1.53
IP de loopback LEAF4	1.1.1.54
Loopback SPINE1 e RP Anycast	Loopback0 - 1.1.1.71, Loopback1 - 10.1.2.10 (RP Anycast)
Loopback SPINE2 e RP Anycast	Loopback0 - 1.1.1.72, Loopback1 - 10.1.2.10 (RP Anycast)
HOST 1	192.168.10.10 (0000.000.aaaa) (VLAN 10)
HOST 2	192.168.20.10 (0000. 0000.bbb) (VLAN 20)
HOST 3	192.168.10.20 (0000. 0000.cccc) (VLAN 10)
HOST 4	192.168.20.20 (0000. 0000.dddd) (VLAN 20)
VLAN 10	100010 L2VNI
VLAN 20	100020 L2VNI

VLAN 500	50000 L3VNI

Configurar

Diagrama de Rede



Configurações

- As vizinhanças Underlay e PIM já estão estabelecidas.

Switch LEAF:

```
feature ospf

router ospf UNDERLAY
  log-adjacency-changes

interface loopback0
  ip router ospf UNDERLAY area 0.0.0.0

interface Ethernet1/1
  ip ospf cost 4
  ip ospf network point-to-point
  ip router ospf UNDERLAY area 0.0.0.0

interface Ethernet1/2
  ip ospf cost 4
  ip ospf network point-to-point
  ip router ospf UNDERLAY area 0.0.0.0
```

Ativação do Open Shortest Path First (OSPF) no Switch Leaf

```
feature pim

ip pim rp-address 10.1.2.10 group-list 224.0.0.0/4
ip pim ssm range 232.0.0.0/8

vrf context TENANT1
  ip pim ssm range 232.0.0.0/8

interface Vlan10
  ip pim sparse-mode

interface Vlan20
  ip pim sparse-mode

interface loopback0
  ip pim sparse-mode

interface Ethernet1/1
  ip pim sparse-mode

interface Ethernet1/2
  ip pim sparse-mode
```

Habilitando o PIM no Switch Leaf

```
LEAF-1# show ip ospf neighbors
 OSPF Process ID UNDERLAY VRF default
 Total number of neighbors: 2
 Neighbor ID     Pri State            Up Time  Address         Interface
 1.1.1.71          1 FULL/ -          04:32:03 192.168.11.1    Eth1/1
 1.1.1.72          1 FULL/ -          04:17:47 192.168.21.2    Eth1/2
LEAF-1# sh ip pim neighbor
PIM Neighbor Status for VRF "default"
Neighbor        Interface            Uptime    Expires   DR       Bidir-  BFD    ECMP Redirect
                                                          Priority Capable State     Capable
192.168.11.1    Ethernet1/1          04:32:14  00:01:30  1        yes     n/a     no
192.168.21.2    Ethernet1/2          04:17:58  00:01:44  1        yes     n/a     no
LEAF-1#
```

Vizinho OSPF

Switch Spine:

```
feature pim

ip pim rp-address 10.1.2.10 group-list 224.0.0.0/4
ip pim ssm range 232.0.0.0/8
ip pim anycast-rp 10.1.2.10 1.1.1.71
ip pim anycast-rp 10.1.2.10 1.1.1.72
```

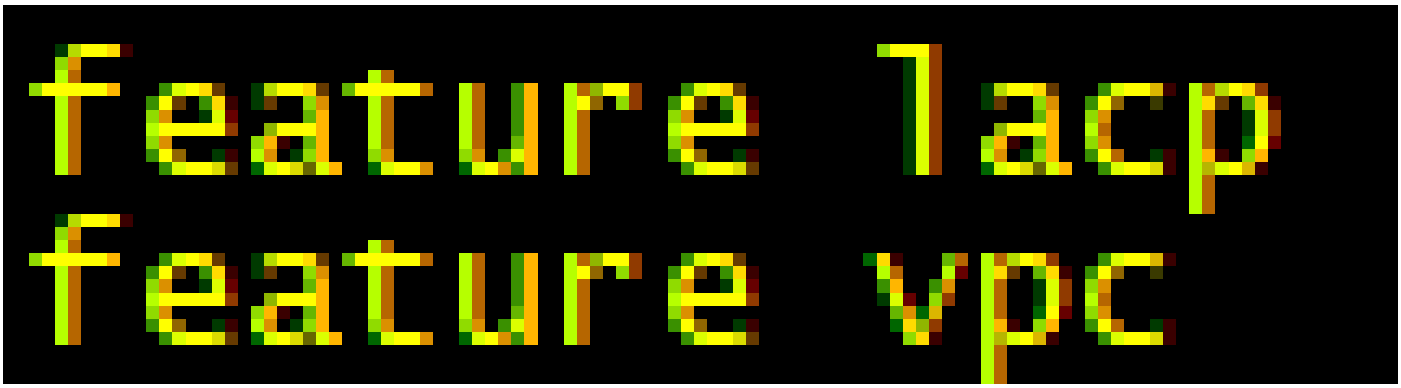
Habilitando o PIM no Switch Spine

- As vizinhanças Underlay e PIM já estão estabelecidas.
- Ambos os switches Spine serão o RP Anycast idêntico para todo o grupo multicast 224.0.0.0/4.
- Maximum Transmission Unit (MTU) é definido como 9000/9216 nas interfaces entre os switches Leaf e Spine.

Primeiro, vamos configurar um vPC entre Leaf1 e Leaf2.

Etapa 1. Ativação de recursos e domínios do vPC.

- Ative o recurso vPC e o LACP (Link Aggregation Control Protocol).
- Configure o domínio do vPC.
- As interfaces mgmt 0 são usadas como um link keepalive par e Eth1/3 e Eth1/4 serão a parte do link peer vPC (Port-Channel 1).
- Certifique-se de que o comando peer-switch esteja configurado para compartilhar um endereço MAC comum com switches descendentes.



Ativando o recurso no switch leaf

```
LEAF-1# sh run vpc

!Command: show running-config vpc
!Running configuration last done at: Sat Dec 28 07:17:18 2024
!Time: Sat Dec 28 07:39:48 2024

version 7.0(3)I7(9) Bios:version
feature vpc

vpc domain 1
  peer-switch
  role priority 100
  peer-keepalive destination 192.168.0.52
  peer-gateway

interface port-channel1
  vpc peer-link
```

Ativação do vPC no switch leaf 1

```
LEAF-2# sh run vpc

!Command: show running-config vpc
!Running configuration last done at: Sat Dec 28 07:17:14 2024
!Time: Sat Dec 28 07:40:20 2024

version 7.0(3)I7(9) Bios:version
feature vpc

vpc domain 1
  peer-switch
  role priority 200
  peer-keepalive destination 192.168.0.51
  peer-gateway

interface port-channel1
  vpc peer-link
```

Ativação do vPC no switch leaf 2

Etapa 2. Atribuição de membro de porta.

- Atribua o membro da porta ao grupo de canais e inclua-o no vPC. Nesse caso, dois vPCs estão sendo usados. vPC 20 e vPC 10.

```
LEAF-1# sh run int port-channel 10, port-channel 20 membership

!Command: show running-config interface port-channel10, port-channel20 membership
!Running configuration last done at: Sat Dec 28 07:17:18 2024
!Time: Sat Dec 28 07:42:44 2024

version 7.0(3)I7(9) Bios:version

interface port-channel10
  switchport mode trunk
  vpc 10

interface Ethernet1/5

  switchport mode trunk
  channel-group 10 mode active

interface port-channel20
  switchport mode trunk
  vpc 20

interface Ethernet1/6

  switchport mode trunk
  channel-group 20 mode active

LEAF-1#
```

Atribuindo o Port Channel no Switch Leaf 1

```
LEAF-2# sh run int port-channel 10, port-channel 20 membership

!Command: show running-config interface port-channel10, port-channel20 membership
!Running configuration last done at: Sat Dec 28 07:17:14 2024
!Time: Sat Dec 28 07:43:16 2024

version 7.0(3)I7(9) Bios:version

interface port-channel10
  switchport mode trunk
  vpc 10

interface Ethernet1/5

  switchport mode trunk
  channel-group 10 mode active

interface port-channel20
  switchport mode trunk
  vpc 20

interface Ethernet1/6

  switchport mode trunk
  channel-group 20 mode active

LEAF-2#
```

Atribuindo o Port Channel no Switch Leaf 2

- Aqui, um vPC é criado e os correspondentes começam a trocar mensagens de keepalive para verificar a disponibilidade.

```
LEAF-1# show vpc
Legend:
                (*) - local vPC is down, forwarding via vPC peer-link

vPC domain id                     : 1
Peer status                       : peer adjacency formed ok
vPC keep-alive status             : peer is alive
Configuration consistency status  : success
Per-vlan consistency status       : success
Type-2 consistency status         : success
vPC role                          : primary
Number of vPCs configured         : 2
Peer Gateway                      : Enabled
Dual-active excluded VLANs        : -
Graceful Consistency Check        : Enabled
Auto-recovery status              : Disabled
Delay-restore status              : Timer is off.(timeout = 30s)
Delay-restore SVI status          : Timer is off.(timeout = 10s)
Operational Layer3 Peer-router    : Disabled

vPC Peer-link status
---------------------------------------------------------------------
id    Port   Status Active vlans
--    ----   ------ --------------------------------------------------
1     Po1    up     1,10,20,500

vPC status
----------------------------------------------------------------------------
Id    Port          Status Consistency Reason                Active vlans
--    ------------  ------ ----------- ------                ---------------
10    Po10          up     success     success               1,10,20,500

20    Po20          up     success     success               1,10,20,500


Please check "show vpc consistency-parameters vpc <vpc-num>" for the
consistency reason of down vpc and for type-2 consistency reasons for
any vpc.

LEAF-1#
```

Status do vPC no Switch Leaf 1

```
LEAF-2# sh vpc
Legend:
                (*) - local vPC is down, forwarding via vPC peer-link

vPC domain id                     : 1
Peer status                       : peer adjacency formed ok
vPC keep-alive status             : peer is alive
Configuration consistency status  : success
Per-vlan consistency status       : success
Type-2 consistency status         : success
vPC role                          : secondary
Number of vPCs configured         : 2
Peer Gateway                      : Enabled
Dual-active excluded VLANs        : -
Graceful Consistency Check        : Enabled
Auto-recovery status              : Disabled
Delay-restore status              : Timer is off.(timeout = 30s)
Delay-restore SVI status          : Timer is off.(timeout = 10s)
Operational Layer3 Peer-router    : Disabled

vPC Peer-link status
---------------------------------------------------------------------
id     Port   Status Active vlans
--     ----   ------ --------------------------------------------------
1      Po1    up     1,10,20,500

vPC status
----------------------------------------------------------------------------
Id     Port          Status Consistency Reason                Active vlans
--     ------------  ------ ----------- ------                ---------------
10     Po10          up     success     success               1,10,20,500

20     Po20          up     success     success               1,10,20,500


Please check "show vpc consistency-parameters vpc <vpc-num>" for the
consistency reason of down vpc and for type-2 consistency reasons for
any vpc.

LEAF-2#
```

Status do vPC no Switch Leaf 2

- As VLANs 10, 20, 500 já estão configuradas e passaram pelas portas membro do vPC e pelo link par do vPC.

Etapa 3. Configurar o endereço IP secundário.

- Quando o vPC é incluído na estrutura VXLAN, ambos os pares VTEP do vPC começam a usar endereços IP virtuais (VIP) como endereços de origem em vez de seus endereços IP físicos (PIP). Isso também significa que quando o BGP Ethernet VPN (EVPN) anuncia os tipos de rota 2 (anúncio MAC/IP) e 5 (rota de prefixo IP) por padrão, o VIP é usado como o próximo salto. A interface Loopback 0 em nosso exemplo é configurada com dois endereços IP: 10.1.1.100/32 (VIP) como o IP secundário e 1.1.1.51/32 (PIP) como o IP primário.
- Aqui um endereço IP comum é configurado como secundário na interface de loopback 0.

```
LEAF-1# sh run int l0

!Command: show running-config interface loopback0
!Running configuration last done at: Sat Dec 28 07:51:58 2024
!Time: Sat Dec 28 07:55:26 2024

version 7.0(3)I7(9) Bios:version

interface loopback0
  ip address 1.1.1.51/32
  ip address 10.1.1.100/32 secondary
  ip router ospf UNDERLAY area 0.0.0.0
  ip pim sparse-mode

LEAF-1#
```

IP secundário no switch leaf 1

```
LEAF-2# sh run int l0

!Command: show running-config interface loopback0
!Running configuration last done at: Sat Dec 28 07:52:05 2024
!Time: Sat Dec 28 07:55:37 2024

version 7.0(3)I7(9) Bios:version

interface loopback0
  ip address 1.1.1.52/32
  ip address 10.1.1.100/32 secondary
  ip router ospf UNDERLAY area 0.0.0.0
  ip pim sparse-mode

LEAF-2#
```

IP secundário no switch leaf 2

Etapa 4. Ativar a VXLAN e recursos relacionados.

- Sobreposição de virtualização de rede (nV) - permite VXLAN
- EVPN de sobreposição de nV de recurso - permite plano de controle de EVPN
- Encaminhamento de malha de recursos - habilita o Host Mobility Manager
- Feature Virtual Network (VN)-Segment-VLAN-Based - habilita VXLAN baseada em VLAN

```
LEAF-1# sh run | sec "feature|nv over"
nv overlay evpn
feature ospf
feature bgp
feature pim
feature fabric forwarding
feature interface-vlan
feature vn-segment-vlan-based
feature lacp
feature vpc
feature nv overlay
LEAF-1#
```

Recursos no Switch Leaf

```
SPINE-1# sh run | sec "feature|nv over"
nv overlay evpn
feature ospf
feature bgp
feature pim
feature nv overlay
SPINE-1#
```

Recursos no switch spine

- Como o spine não exige conhecimento das informações de VLAN do cliente, os recursos de estrutura e do segmento de VLAN não precisam ser ativados.

Etapa 5. Ative a vizinhança BGP.

- O BGP entre os switches Leaf e Spine deve ser habilitado. A lombada servirá como um refletor de rota no laboratório.
- Embora seja opcional configurar o refletor de rota (RR), para fins de escalabilidade, a Cisco recomenda o RR.

```
LEAF-1# sh run bgp

!Command: show running-config bgp
!Running configuration last done at: Sat Dec 28 07:51:58 2024
!Time: Sat Dec 28 08:07:35 2024

version 7.0(3)I7(9) Bios:version
feature bgp

router bgp 65000
  router-id 1.1.1.51
  neighbor 1.1.1.71
    remote-as 65000
    update-source loopback0
    address-family l2vpn evpn
      send-community extended
  neighbor 1.1.1.72
    remote-as 65000
    update-source loopback0
    address-family l2vpn evpn
      send-community extended
```

Habilitando BGP em Switch Leaf

```
SPINE-1# sh run bgp

!Command: show running-config bgp
!Running configuration last done at: Sat Dec 28 07:16:33 2024
!Time: Sat Dec 28 08:08:21 2024

version 7.0(3)I7(9) Bios:version
feature bgp

router bgp 65000
  router-id 1.1.1.71
  neighbor 1.1.1.51
    remote-as 65000
    update-source loopback0
    address-family l2vpn evpn
      send-community extended
      route-reflector-client
  neighbor 1.1.1.52
    remote-as 65000
    update-source loopback0
    address-family l2vpn evpn
      send-community extended
      route-reflector-client
  neighbor 1.1.1.53
    remote-as 65000
    update-source loopback0
    address-family l2vpn evpn
      send-community extended
      route-reflector-client
  neighbor 1.1.1.54
    remote-as 65000
    update-source loopback0
    address-family l2vpn evpn
      send-community extended
      route-reflector-client


SPINE-1#
```

Ativando o BGP no switch spine

```
LEAF-1# show bgp l2vpn evpn summary
BGP summary information for VRF default, address family L2VPN EVPN
BGP router identifier 1.1.1.51, local AS number 65000
BGP table version is 62, L2VPN EVPN config peers 2, capable peers 2
10 network entries and 13 paths using 2228 bytes of memory
BGP attribute entries [10/1600], BGP AS path entries [0/0]
BGP community entries [0/0], BGP clusterlist entries [4/16]

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
1.1.1.71        4 65000     146     121       62    0    0 01:45:52 3
1.1.1.72        4 65000     141     114       62    0    0 01:39:12 3
LEAF-1#
```

Status de BGP em Switch Leaf

```
SPINE-1# show bgp l2vpn evpn summary
BGP summary information for VRF default, address family L2VPN EVPN
BGP router identifier 1.1.1.71, local AS number 65000
BGP table version is 98, L2VPN EVPN config peers 4, capable peers 4
9 network entries and 9 paths using 2124 bytes of memory
BGP attribute entries [7/1120], BGP AS path entries [0/0]
BGP community entries [0/0], BGP clusterlist entries [0/0]

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
1.1.1.51        4 65000     147     124       98    0    0 01:46:29 2
1.1.1.52        4 65000     147     124       98    0    0 01:46:30 2
1.1.1.53        4 65000     128     155       98    0    0 02:01:15 1
1.1.1.54        4 65000     191     225       98    0    0 03:03:08 2
SPINE-1#
```

Status de BGP no Switch Spine

Etapa 6. Ativar o contexto VRF em switches leaf. O VRF separa o tráfego do cliente e facilita a comunicação entre dois L2VNIs distintos via L3VNI.

- Aloque 50000 L3VNI em VRF TENANT1.

```
vrf context TENANT1
  vni 50000
  ip pim ssm range 232.0.0.0/8
  rd auto
  address-family ipv4 unicast
    route-target both auto
    route-target both auto evpn
```

Alocação L3VNI

Etapa 7. Interface Virtual de Rede (NVE - Network Virtual Interface), identificador de VXLAN (VNI - VXLAN Identifier) e configuração de VLAN.

- Configure a interface NVE, usando Loopback 0 como origem. Defina o grupo Multicast para cada VNI, em que o tráfego de broadcast da camada 2, unicast desconhecido e multicast (BUM) será entregue e, em seguida, conecte os 100010 VNI e os IDs de 100020 à interface NVE. O cabeçalho de VXLAN contém as informações que o VNI usa para identificar a quais segmentos de VXLAN ele pertence.
- O 50000 L3VNI é vinculado à instância VRF (ao enviá-lo ao switch spine, o VNI 50000 foi conectado à tabela VRF).
- O comando host-reachability protocol BGP ativa a família de endereços EVPN no túnel VXLAN, o que significa que os endereços MAC e os endereços IP são aprendidos através do protocolo BGP no plano de controle e não no plano de dados.
- Configure suppress-arp na interface NVE.
- Conecte VLAN de Camada 2 e Camada 3 ao VNI relevante.

Protocolo de Resolução de Endereço (ARP - Suppress-Address Resolution Protocol):

O plano de controle do EVPN Multiprotocolo (MP - Multi-Protocol)-BGP oferece uma melhoria chamada supressão ARP para reduzir a inundação de rede provocada pelo tráfego de broadcast das solicitações ARP. Cada um dos VTEPs de um VNIs mantém uma tabela de cache de supressão ARP para hosts IP conhecidos e os endereços MAC que correspondem a eles no segmento VNI quando a supressão ARP é ativada para esse VNI. Seu VTEP local intercepta a solicitação ARP e procura o endereço IP resolvido ARP em sua tabela de cache de supressão ARP sempre que um host final no VNI envia uma solicitação ARP para outro endereço IP de host final. Em nome do host final remoto, o VTEP local envia uma resposta ARP se descobrir uma correspondência. A resposta ARP então fornece ao host local o endereço MAC dos hosts remotos. A solicitação ARP é inundada para os outros VTEPs no VNI se o VTEP local não tiver o endereço IP resolvido ARP em sua tabela de supressão ARP. Para a primeira solicitação ARP para um host de rede silencioso, essa inundação ARP pode ocorrer.

```
LEAF-1# sh run interface nve 1

!Command: show running-config interface nve1
!Running configuration last done at: Sat Dec 28 07:51:58 2024
!Time: Sat Dec 28 08:44:44 2024

version 7.0(3)I7(9) Bios:version

interface nve1
  no shutdown
  host-reachability protocol bgp
  source-interface loopback0
  member vni 50000 associate-vrf
  member vni 100010
    suppress-arp
    mcast-group 239.0.0.10
  member vni 100020
    suppress-arp
    mcast-group 239.0.0.20

LEAF-1#
```

Interface NVE

```
LEAF-1# sh run vlan

!Command: show running-config vlan
!Running configuration last done at: Sat Dec 28 07:51:58 2024
!Time: Sat Dec 28 08:46:44 2024

version 7.0(3)I7(9) Bios:version
vlan 1,10,20,500
vlan 10
  vn-segment 100010
vlan 20
  vn-segment 100020
vlan 500
  vn-segment 50000


LEAF-1#
```

Mapeamento de VLAN para segmento VLAN

- Enviando à Spine uma mensagem de união PIM, a interface NVE se juntará aos grupos de multicast 239.0.0.10 e 239.0.0.20, respectivamente, assim que for inicializada.
- Você pode ver outras tabelas (S, G) também (1.1.1.54,239.0.0.20) e (10.1.1.100, 239.0.0.10/239.0.0.20) na imagem e elas já estão registradas com Spine de diferentes Switches Leaf.

```
LEAF-1# sh ip mroute summary
IP Multicast Routing Table for VRF "default"
Route Statistics unavailable - only liveness detected

Total number of routes: 7
Total number of (*,G) routes: 2
Total number of (S,G) routes: 4
Total number of (*,G-prefix) routes: 1
Group count: 2, rough average sources per group: 2.0

Group: 232.0.0.0/8, Source count: 0
Source          packets      bytes          aps   pps       bit-rate     oifs
(*,G)           0            0              0     0         0.000   bps  0

Group: 239.0.0.10/32, Source count: 2
Source          packets      bytes          aps   pps       bit-rate     oifs
(*,G)           1            100            100   0         0.000   bps  1
1.1.1.53        48           4644           96    0         78.267  bps  1
10.1.1.100      1124         113514         100   0         131.467 bps  1

Group: 239.0.0.20/32, Source count: 2
Source          packets      bytes          aps   pps       bit-rate     oifs
(*,G)           1            100            100   0         0.000   bps  1
1.1.1.54        51           4944           96    0         63.200  bps  1
10.1.1.100      1116         112729         101   0         70.667  bps  1
LEAF-1#
```

Tabela Mroute

Etapa 8. Ativar a instância EVPN.

- Ative a instância de EVPN junto com a família de endereços para EVPN e VRF sob BGP.

```
LEAF-1# sh run bgp

!Command: show running-config bgp
!Running configuration last done at: Sat Dec 28 09:22:19 2024
!Time: Sat Dec 28 09:43:07 2024

version 7.0(3)I7(9) Bios:version
feature bgp

router bgp 65000
  router-id 1.1.1.51
  neighbor 1.1.1.71
    remote-as 65000
    update-source loopback0
    address-family l2vpn evpn
      send-community extended
  neighbor 1.1.1.72
    remote-as 65000
    update-source loopback0
    address-family l2vpn evpn
      send-community extended
  vrf TENANT1
    address-family ipv4 unicast
      redistribute direct route-map REDIST
evpn
  vni 100010 l2
    rd auto
    route-target import auto
    route-target export auto
  vni 100020 l2
    rd auto
    route-target import auto
    route-target export auto
vrf context TENANT1
```

Instância de EVPN

- A única finalidade do mapa de rota REDIST é permitir tudo.
- Usando o comando redistribute direct, as rotas conectadas com reconhecimento de VRF são promovidas em MP-BGP (rotas tipo 5).
- A configuração de EVPN exibida acima é idêntica à instrução de rede usada pelo BGP para anunciar rotas MAC (rotas tipo 2).

Etapa 9. Configurar a interface virtual do switch (SVI) para cada VLAN para o host final em VRF.

- Em cada switch leaf, o SVI é configurado para VLAN configurada localmente e um SVI para VLAN L3VNI para obter a Base de Informações de Roteamento Simétrico (RIB).

RIB simétrico:

- Quando o host final envia o pacote de dados para uma rede diferente e recebe para o Switch Leaf, ele será processado primeiro no L2VNI e, em seguida, será colocado no L3VNI usando VRF e enviado para o Leaf remoto.
- A Folha Remota recebe primeiro os pacotes na tabela VRF usando Roteamento e depois fazendo a ponte para o L2VNI e envia para o host final.
- Dessa forma, o roteamento simétrico (B-R-R-B) é alcançado.

```
LEAF-1# sh run interface vlan 10,vlan 20,vlan 500

!Command: show running-config interface Vlan10, Vlan20, Vlan500
!Running configuration last done at: Sat Dec 28 09:22:19 2024
!Time: Sat Dec 28 10:00:26 2024

version 7.0(3)I7(9) Bios:version

interface Vlan10
  no shutdown
  mtu 9216
  vrf member TENANT1
  no ip redirects
  ip address 192.168.10.254/24
  no ipv6 redirects
  ip pim sparse-mode
  fabric forwarding mode anycast-gateway

interface Vlan20
  no shutdown
  mtu 9216
  vrf member TENANT1
  no ip redirects
  ip address 192.168.20.254/24
  no ipv6 redirects
  ip pim sparse-mode
  fabric forwarding mode anycast-gateway

interface Vlan500
  no shutdown
  vrf member TENANT1
  no ip redirects
  ip forward
  no ipv6 redirects

LEAF-1#
```

Interfaces VLAN

- O comando IP forward na VLAN 500 é usado para permitir o encaminhamento de Camada 3 para todas as VXLANs. Não há necessidade de configurar o endereço IP, pois ele apenas processa o pacote da tabela L2VNI para a tabela L3VNI.

```
LEAF-1# show bgp vpnv4 unicast vrf TENANT1
BGP routing table information for VRF default, address family VPNv4 Unicast
BGP table version is 15, Local Router ID is 1.1.1.51
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

   Network            Next Hop            Metric     LocPrf     Weight Path
Route Distinguisher: 1.1.1.51:3    (VRF TENANT1)
*>r192.168.10.0/24    0.0.0.0                  0        100      32768 ?
*>i192.168.10.20/32   1.1.1.53                          100          0 i
*>r192.168.20.0/24    0.0.0.0                  0        100      32768 ?
*>i192.168.20.20/32   1.1.1.54                          100          0 i

LEAF-1#
```

Aprendendo rotas BGP VPNv4 para VRF TENANT1

- O endereço IP de cada VLAN será comum para todas as SVIs em todos os switches leaf.

Isso é chamado de IP anycast e é usado no Gerenciamento de mobilidade, onde o usuário final pode se comunicar com outro host perfeitamente, sem qualquer interrupção.

Etapa 10. Habilitar o MAC do gateway anycast de encaminhamento de estrutura para o host final.

- Ele garante redundância transparente de gateway de Camada 3 e encaminhamento otimizado para dispositivos conectados à estrutura.
- O endereço MAC do gateway Anycast é um endereço MAC globalmente consistente usado para todos os gateways de Camada 3 em uma malha.
- O conceito é idêntico ao empregado no First Hop Redundancy Protocol (FHRP), em que cada grupo recebe um MAC virtual.

```
LEAF-1# show running-config fabric forwarding

!Command: show running-config fabric forwarding
!Running configuration last done at: Sat Dec 28 09:22:19 2024
!Time: Sat Dec 28 10:08:08 2024

version 7.0(3)I7(9) Bios:version
nv overlay evpn
feature fabric forwarding

fabric forwarding anycast-gateway-mac 0000.1234.5678


interface Vlan10
  fabric forwarding mode anycast-gateway

interface Vlan20
  fabric forwarding mode anycast-gateway

LEAF-1#
```

Ativando o encaminhamento de estrutura

Etapa 11. Ative a VLAN de acesso/tronco para as portas membro.

Switch vPC:

```
LEAF-1# sh run int po10 membership

!Command: show running-config interface port-channel10 membership
!Running configuration last done at: Sat Dec 28 09:22:19 2024
!Time: Sat Dec 28 10:13:19 2024

version 7.0(3)I7(9) Bios:version

interface port-channel10
  switchport mode trunk
  vpc 10

interface Ethernet1/5

  switchport mode trunk
  channel-group 10 mode active

LEAF-1#
```

Ativação de portas de tronco para a interface de membro do vPC

Switch não vPC:

```
LEAF-3# show running-config interface e1/3

!Command: show running-config interface Ethernet1/3
!Running configuration last done at: Sat Dec 28 09:28:18 2024
!Time: Sat Dec 28 10:14:42 2024

version 7.0(3)I7(9) Bios:version

interface Ethernet1/3
  switchport access vlan 10
  spanning-tree port type edge

LEAF-3#
```

Ativação de portas de tronco para a interface de membro não vPC

Verificação

- Verifique o ARP e a tabela de endereços MAC.

```
LEAF-1# sh ip arp vrf TENANT1

Flags: * - Adjacencies learnt on non-active FHRP router
       + - Adjacencies synced via CFSoE
       # - Adjacencies Throttled for Glean
       CP - Added via L2RIB, Control plane Adjacencies
       PS - Added via L2RIB, Peer Sync
       RO - Re-Originated Peer Sync Entry
       D - Static Adjacencies attached to down interface

IP ARP Table for context TENANT1
Total number of entries: 2
Address         Age       MAC Address     Interface       Flags
192.168.20.10   00:00:36  0000.0000.bbbb  Vlan20
192.168.10.10   00:04:19  0000.0000.aaaa  Vlan10
LEAF-1#  sh ip arp suppression-cache deta

Flags: + - Adjacencies synced via CFSoE
       L - Local Adjacency
       R - Remote Adjacency
       L2 - Learnt over L2 interface
       PS - Added via L2RIB, Peer Sync
       RO - Dervied from L2RIB Peer Sync Entry

Ip Address      Age      Mac Address    Vlan Physical-ifindex    Flags    Remote Vtep Addrs

192.168.10.10   00:04:33 0000.0000.aaaa   10 port-channel10      L
192.168.10.20   00:55:53 0000.0000.cccc   10 (null)              R        1.1.1.53
192.168.20.10   00:00:50 0000.0000.bbbb   20 port-channel20      L
192.168.20.20   03:26:04 0000.0000.dddd   20 (null)              R        1.1.1.54
LEAF-1#
```

Tabela ARP e MAC no Switch LEAF 1

```
LEAF-2# show ip arp vrf TENANT1

Flags: * - Adjacencies learnt on non-active FHRP router
       + - Adjacencies synced via CFSoE
       # - Adjacencies Throttled for Glean
       CP - Added via L2RIB, Control plane Adjacencies
       PS - Added via L2RIB, Peer Sync
       RO - Re-Originated Peer Sync Entry
       D - Static Adjacencies attached to down interface

IP ARP Table for context TENANT1
Total number of entries: 2
Address         Age       MAC Address     Interface       Flags
192.168.20.10   00:01:28  0000.0000.bbbb  vlan20          +
192.168.10.10   00:00:11  0000.0000.aaaa  vlan10          +
LEAF-2#
```

Tabela ARP e MAC no Switch LEAF 2

- Ambos os peers mantêm as entradas ARP.
- Verifique o status da NVI (Network Virtual Interface, interface virtual de rede).

Switch vPC:

```
LEAF-1# show nve peers
Interface Peer-IP          State LearnType Uptime   Router-Mac
--------- ---------------  ----- --------- -------- -----------------
nve1      1.1.1.53         Up    CP        01:09:04 5000.0003.0007
nve1      1.1.1.54         Up    CP        03:39:16 5000.0004.0007

LEAF-1# show nve vni
Codes: CP - Control Plane        DP - Data Plane
       UC - Unconfigured         SA - Suppress ARP
       SU - Suppress Unknown Unicast
       Xconn - Crossconnect
       MS-IR - Multisite Ingress Replication

Interface VNI      Multicast-group   State Mode Type [BD/VRF]      Flags
--------- -------- ----------------- ----- ---- ------------------ -----
nve1      50000    n/a               Up    CP   L3 [TENANT1]
nve1      100010   239.0.0.10        Up    CP   L2 [10]           SA
nve1      100020   239.0.0.20        Up    CP   L2 [20]           SA

LEAF-1#
```

Pares NVE no Switch vPC

Switch não vPC:

```
LEAF-3# show nve peers
Interface Peer-IP          State LearnType Uptime   Router-Mac
--------- ---------------  ----- --------- -------- -----------------
nve1      1.1.1.54         Up    CP        01:14:00 5000.0004.0007
nve1      10.1.1.100       Up    CP        01:14:16 5000.0001.0007

LEAF-3#
```

Pares NVE em Switch não vPC

- Aqui, você observará que o IP do peer é 10.1.1.100 em vez do endereço IP de loopback primário, de modo que o pacote de retorno será roteado para esse IP para qualquer um dos switches vPC.
- Verifique as rotas BGP EVPN.

```
LEAF-1# show l2route evpn mac-ip all
Flags -(Rmac):Router MAC (Stt):Static (L):Local (R):Remote (V):vPC link
(Dup):Duplicate (Spl):Split (Rcv):Recv(D):Del Pending (S):Stale (C):Clear
(Ps):Peer Sync (Ro):Re-Originated
Topology    Mac Address    Prod   Flags         Seq No     Host IP         Next-Hops
----------- -------------- ------ ----------- ---------- --------------- --------------
10          0000.0000.aaaa HMM    --            0          192.168.10.10   Local
10          0000.0000.cccc BGP    --            0          192.168.10.20   1.1.1.53
20          0000.0000.bbbb HMM    --            0          192.168.20.10   Local
20          0000.0000.dddd BGP    --            0          192.168.20.20   1.1.1.54
LEAF-1#
```

BGP l2route EVPN MAC-IP

```
LEAF-1# show l2route evpn mac all

Flags -(Rmac):Router MAC (Stt):Static (L):Local (R):Remote (V):vPC link
(Dup):Duplicate (Spl):Split (Rcv):Recv (AD):Auto-Delete (D):Del Pending
(S):Stale (C):Clear, (Ps):Peer Sync (O):Re-Originated (Nho):NH-Override
(Pf):Permanently-Frozen

Topology     Mac Address    Prod   Flags         Seq No     Next-Hops
-----------  -------------  -----  ------------  ---------  ---------------
10           0000.0000.aaaa Local  L,            0          Po10
10           0000.0000.cccc BGP    Spl           0          1.1.1.53
20           0000.0000.bbbb Local  L,            0          Po20
20           0000.0000.dddd BGP    SplRcv        0          1.1.1.54
500          5000.0003.0007 VXLAN  Rmac          0          1.1.1.53
500          5000.0004.0007 VXLAN  Rmac          0          1.1.1.54
LEAF-1#
```

BGP l2route EVPN MAC

```
LEAF-1# show bgp l2vpn evpn summary
BGP summary information for VRF default, address family L2VPN EVPN
BGP router identifier 1.1.1.51, local AS number 65000
BGP table version is 134, L2VPN EVPN config peers 2, capable peers 2
12 network entries and 15 paths using 2568 bytes of memory
BGP attribute entries [12/1920], BGP AS path entries [0/0]
BGP community entries [0/0], BGP clusterlist entries [4/16]

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
1.1.1.71        4 65000     312     263      134    0    0 03:46:01 3
1.1.1.72        4 65000     307     256      134    0    0 03:39:21 3
LEAF-1#
```

Resumo do BGP EVPN

```
LEAF-1# show bgp l2vpn evpn
BGP routing table information for VRF default, address family L2VPN EVPN
BGP table version is 146, Local Router ID is 1.1.1.51
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

   Network            Next Hop            Metric     LocPrf     Weight Path
Route Distinguisher: 1.1.1.51:32777    (L2VNI 100010)
*>l[2]:[0]:[0]:[48]:[0000.0000.aaaa]:[0]:[0.0.0.0]/216
                      10.1.1.100                             100      32768 i
*>l[2]:[0]:[0]:[48]:[0000.0000.aaaa]:[32]:[192.168.10.10]/272
                      10.1.1.100                             100      32768 i
*>i[2]:[0]:[0]:[48]:[0000.0000.cccc]:[32]:[192.168.10.20]/272
                      1.1.1.53                               100          0 i

Route Distinguisher: 1.1.1.51:32787    (L2VNI 100020)
*>l[2]:[0]:[0]:[48]:[0000.0000.bbbb]:[0]:[0.0.0.0]/216
                      10.1.1.100                             100      32768 i
*>i[2]:[0]:[0]:[48]:[0000.0000.dddd]:[0]:[0.0.0.0]/216
                      1.1.1.54                               100          0 i
*>l[2]:[0]:[0]:[48]:[0000.0000.bbbb]:[32]:[192.168.20.10]/272
                      10.1.1.100                             100      32768 i
*>i[2]:[0]:[0]:[48]:[0000.0000.dddd]:[32]:[192.168.20.20]/272
                      1.1.1.54                               100          0 i

Route Distinguisher: 1.1.1.53:32777
*>i[2]:[0]:[0]:[48]:[0000.0000.cccc]:[32]:[192.168.10.20]/272
                      1.1.1.53                               100          0 i
* i                   1.1.1.53                               100          0 i

Route Distinguisher: 1.1.1.54:32787
* i[2]:[0]:[0]:[48]:[0000.0000.dddd]:[0]:[0.0.0.0]/216
                      1.1.1.54                               100          0 i
*>i                   1.1.1.54                               100          0 i
* i[2]:[0]:[0]:[48]:[0000.0000.dddd]:[32]:[192.168.20.20]/272
                      1.1.1.54                               100          0 i
*>i                   1.1.1.54                               100          0 i

Route Distinguisher: 1.1.1.51:3    (L3VNI 50000)
*>i[2]:[0]:[0]:[48]:[0000.0000.cccc]:[32]:[192.168.10.20]/272
                      1.1.1.53                               100          0 i
*>i[2]:[0]:[0]:[48]:[0000.0000.dddd]:[32]:[192.168.20.20]/272
                      1.1.1.54                               100          0 i

LEAF-1#
```

Rotas BGP EVPN

- É comum questionar como os Switches Leaf adquirem entradas MAC para hosts remotos. Esse processo é facilitado pelo Gratuitous ARP. Quando uma porta de rede é ativada, ela imediatamente envia uma solicitação ARP para verificar a exclusividade do endereço IP. Em seguida, cada Switch Leaf registra o endereço MAC e o inclui em um Pacote de Atualização BGP. Isso permite que outros Switches Leaf atualizem suas respectivas tabelas de endereços MAC de acordo. Mas pode haver um caso em que o host final não gere o Gratuitous ARP (host silencioso) e, nesse caso, a solicitação ARP será transmitida para a folha e, como é uma solicitação de transmissão, o switch Leaf gerará a solicitação de transmissão múltipla para o respectivo grupo para o VNI específico. Nesse caso, é 239.0.0.10 e 239.0.0.20.

- Permite fazer ping do Host-1 para o Host-3 dentro do mesmo VNI e examinar a captura.

```
HOST-1#ping 192.168.10.20 rep 2
Type escape sequence to abort.
Sending 2, 100-byte ICMP Echos to 192.168.10.20, timeout is 2 seconds:
!!
Success rate is 100 percent (2/2), round-trip min/avg/max = 11/11/12 ms
HOST-1#
```

Ping do HOST-1 para o HOST-3

Pacote ICMP (Internet Control Message Protocol) sobre VXLAN:

```
> Frame 213: 164 bytes on wire (1312 bits), 164 bytes captured (1312 bits) on interface -, id 0
> Ethernet II, Src: 50:00:00:06:00:07 (50:00:00:06:00:07), Dst: 50:00:00:03:00:07 (50:00:00:03:00:07)
> Internet Protocol Version 4, Src: 10.1.1.100, Dst: 1.1.1.53
v User Datagram Protocol, Src Port: 50413, Dst Port: 4789
    Source Port: 50413
    Destination Port: 4789
    Length: 130
  > Checksum: 0x0000 [zero-value ignored]
    [Stream index: 24]
    [Stream Packet Number: 1]
  > [Timestamps]
    UDP payload (122 bytes)
v Virtual eXtensible Local Area Network
  > Flags: 0x0800, VXLAN Network ID (VNI)
    Group Policy ID: 0
    VXLAN Network Identifier (VNI): 100010
    Reserved: 0
> Ethernet II, Src: 00:00:00_00:aa:aa (00:00:00:00:aa:aa), Dst: 00:00:00_00:cc:cc (00:00:00:00:cc:cc)
v Internet Protocol Version 4, Src: 192.168.10.10, Dst: 192.168.10.20
    0100 .... = Version: 4
    .... 0101 = Header Length: 20 bytes (5)
  > Differentiated Services Field: 0x00 (DSCP: CS0, ECN: Not-ECT)
    Total Length: 100
    Identification: 0x0000 (0)
  > 000. .... = Flags: 0x0
    ...0 0000 0000 0000 = Fragment Offset: 0
    Time to Live: 255
    Protocol: ICMP (1)
    Header Checksum: 0x262a [validation disabled]
    [Header checksum status: Unverified]
    Source Address: 192.168.10.10
    Destination Address: 192.168.10.20
    [Stream index: 11]
> Internet Control Message Protocol
```

Captura do Wireshark mostrando o pacote de solicitação ICMP viajando através do 10010 L2VNI

- Como você pode ver, o IP de origem é 10.1.1.100 com a porta 4789 como destino UDP.
- Como é uma comunicação intraVNI, a VLAN 10 usará 100010 VNI e a VLAN 20 usará o VNI 1000.
- Permite fazer ping do Host-1 para o Host-4 com VNI diferentes e observar a captura.

```
HOST-1#ping 192.168.20.20
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 192.168.20.20, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 11/13/21 ms
HOST-1#
```

Ping do HOST-1 para o HOST-4

Pacote ICMP sobre VXLAN:

```
> Frame 27: 164 bytes on wire (1312 bits), 164 bytes captured (1312 bits) on interface -, id 0
> Ethernet II, Src: 50:00:00:05:00:07 (50:00:00:05:00:07), Dst: 50:00:00:04:00:07 (50:00:00:04:00:07)
> Internet Protocol Version 4, Src: 10.1.1.100, Dst: 1.1.1.54
v User Datagram Protocol, Src Port: 54712, Dst Port: 4789
    Source Port: 54712
    Destination Port: 4789
    Length: 130
  > Checksum: 0x0000 [zero-value ignored]
    [Stream index: 3]
    [Stream Packet Number: 1]
  > [Timestamps]
    UDP payload (122 bytes)
v Virtual eXtensible Local Area Network
  > Flags: 0x0800, VXLAN Network ID (VNI)
    Group Policy ID: 0
    VXLAN Network Identifier (VNI): 50000
    Reserved: 0
> Ethernet II, Src: 50:00:00:01:00:07 (50:00:00:01:00:07), Dst: 50:00:00:04:00:07 (50:00:00:04:00:07)
> Internet Protocol Version 4, Src: 192.168.10.10, Dst: 192.168.20.20
> Internet Control Message Protocol
```

Captura do Wireshark mostrando o pacote de solicitação ICMP viajando através do 50000 L3VNI

- Como é uma comunicação entre VNIs, o 50000 L3VNI será usado.
- Verifique a tabela ARP para o host final.

```
HOST-1#sh ip arp
Protocol  Address          Age (min)  Hardware Addr   Type   Interface
Internet  192.168.10.10           -   0000.0000.aaaa  ARPA   GigabitEthernet0/0
Internet  192.168.10.20          18   0000.0000.cccc  ARPA   GigabitEthernet0/0
Internet  192.168.10.254          3   0000.1234.5678  ARPA   GigabitEthernet0/0
HOST-1#
```

Entradas ARP HOST-1

```
HOST-2#sh ip arp
Protocol  Address          Age (min)  Hardware Addr   Type   Interface
Internet  192.168.20.10           -   0000.0000.bbbb  ARPA   GigabitEthernet0/0
Internet  192.168.20.20          44   0000.0000.dddd  ARPA   GigabitEthernet0/0
Internet  192.168.20.254          4   0000.1234.5678  ARPA   GigabitEthernet0/0
HOST-2#
```

Entradas ARP HOST-2

```
HOST-3#sh ip arp
Protocol  Address          Age (min)  Hardware Addr   Type   Interface
Internet  192.168.10.10         103   0000.0000.aaaa  ARPA   GigabitEthernet0/0
Internet  192.168.10.20           -   0000.0000.cccc  ARPA   GigabitEthernet0/0
Internet  192.168.10.254         10   0000.1234.5678  ARPA   GigabitEthernet0/0
HOST-3#
```

Entradas ARP HOST-3

```
HOST-4#sh ip arp
Protocol  Address          Age (min)  Hardware Addr   Type   Interface
Internet  192.168.20.10          43   0000.0000.bbbb  ARPA   GigabitEthernet0/0
Internet  192.168.20.20           -   0000.0000.dddd  ARPA   GigabitEthernet0/0
Internet  192.168.20.254          6   0000.1234.5678  ARPA   GigabitEthernet0/0
HOST-4#
```

Entradas ARP HOST-4

```
HOST-4#tclsh
HOST-4(tcl)#set ip_list {192.168.10.10 192.168.10.20 192.168.20.10 192.168.20.20}
192.168.10.10 192.168.10.20 192.168.20.10 192.168.20.20
HOST-4(tcl)#foreach ip $ip_list {
HOST-4(tcl)#foreach ip $ip_list {
+>            puts "Pinging $ip rep 50 size 1500"
+>            set result [exec ping $ip]
+>            puts $result
+>        }
Pinging 192.168.10.10 rep 50 size 1500

Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 192.168.10.10, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 12/14/16 ms
Pinging 192.168.10.20 rep 50 size 1500

Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 192.168.10.20, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 10/12/15 ms
Pinging 192.168.20.10 rep 50 size 1500

Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 192.168.20.10, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 8/11/17 ms
Pinging 192.168.20.20 rep 50 size 1500

Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 192.168.20.20, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 1/1/3 ms

HOST-4(tcl)#
```

Ping do HOST-4 para todos os outros hosts finais

Sobre esta tradução
A Cisco traduziu este documento com a ajuda de tecnologias de tradução automática e humana para oferecer conteúdo de suporte aos seus usuários no seu próprio idioma, independentemente da localização.

Observe que mesmo a melhor tradução automática não será tão precisa quanto as realizadas por um tradutor profissional.

A Cisco Systems, Inc. não se responsabiliza pela precisão destas traduções e recomenda que o documento original em inglês (link fornecido) seja sempre consultado.