VXLAN avanzata con vPC: Configurazione e verifica di L2VNI e L3VNI

Sommario

Introduzione

In questo documento viene descritto come configurare un laboratorio con switch Nexus 9Kv utilizzando una VXLAN (Virtual eXtensible Local Area Network) avanzata con vPC (Virtual Port-Channel).

Prerequisiti

Requisiti

Cisco raccomanda la conoscenza dei seguenti argomenti:

- Informazioni su routing e switching, nonché sulla tecnologia Multiprotocol Label Switching (MPLS)
- Esperienza con i principi di routing multicast, ad esempio Rendezvous Point (RP) e Platform Independent Multicast (PIM)
- Descrizione di Border Gateway Protocol (BGP) Address Family Indicator (AFI)/Successive Address Family Indicator (SAFI)

Componenti usati

Il documento può essere consultato per tutte le versioni software o hardware.

Le informazioni discusse in questo documento fanno riferimento a dispositivi usati in uno specifico ambiente di emulazione. Su tutti i dispositivi menzionati nel documento la configurazione è stata ripristinata ai valori predefiniti. Se la rete è operativa, valutare attentamente eventuali conseguenze derivanti dall'uso dei comandi.

Premesse

Il documento fornisce anche linee guida per la distribuzione del lab, nonché per la verifica delle configurazioni e delle operazioni.

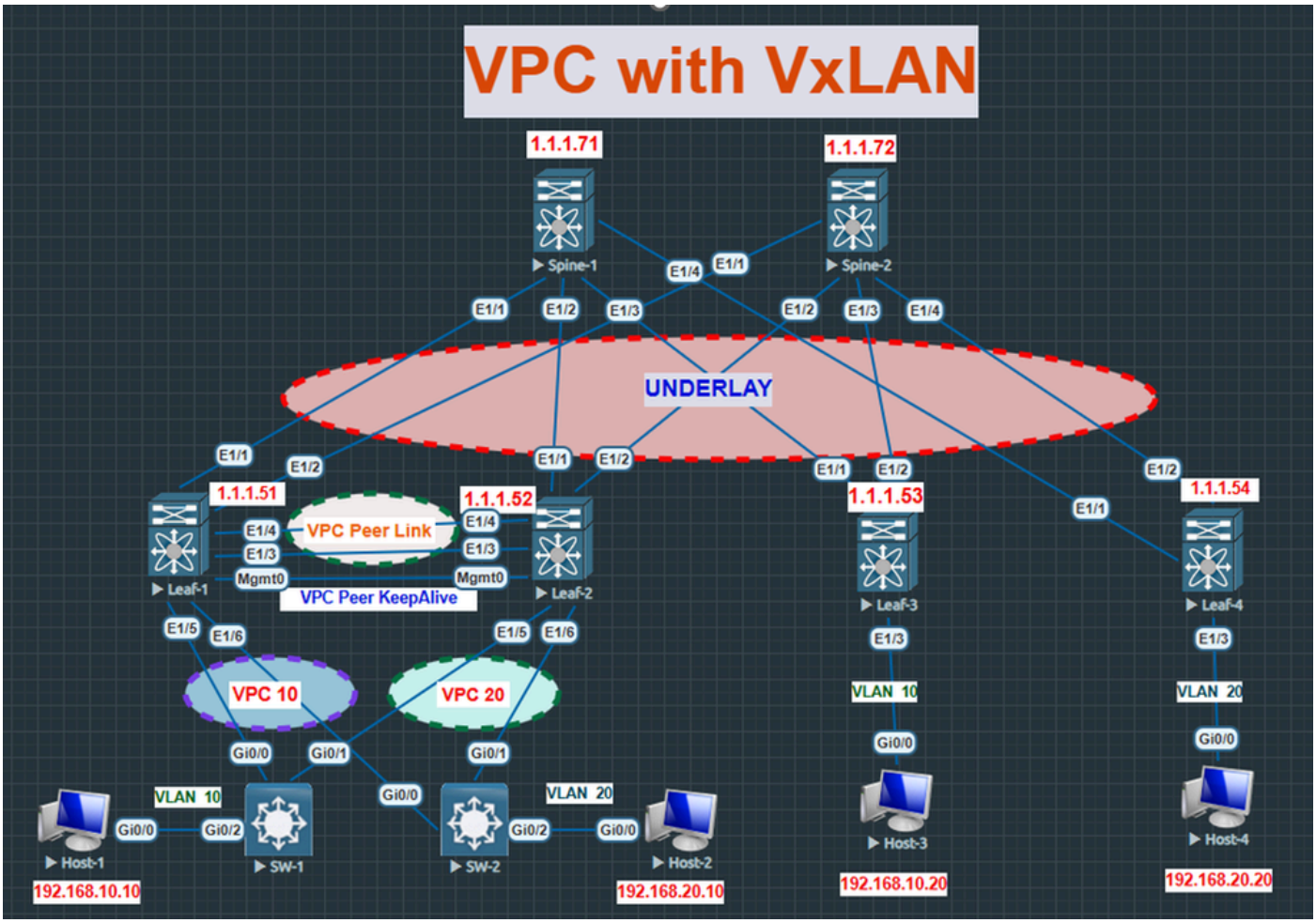
In questo laboratorio, gli switch EveNg con Nexus 9000V vengono utilizzati sia per Leaf che per Spine.

VTEP (Virtual Tunnel Endpoint)	FOGLIA1, FOGLIA2, FOGLIA3, FOGLIA4
vPC	FOGLIA1 e FOGLIA2
IP di loopback primario e secondario LEAF1	Loopback0 - 1.1.1.51, Loopback1 - 10.1.1.100
IP di loopback primario e secondario LEAF2	Loopback0 - 1.1.1.52, Loopback1 - 10.1.1.100
IP loopback LEAF3	1.1.1.53
IP loopback LEAF4	1.1.1.54
LOOPBACK SPINE1 e Anycast RP	Loopback0 - 1.1.1.71, Loopback1 - 10.1.2.10 (Anycast RP)
LOOPBACK SPINE2 e Anycast RP	Loopback0 - 1.1.1.72, Loopback1 - 10.1.2.10 (Anycast RP)
HOST 1	192.168.10.10 (0000. 0000.aaa) (VLAN 10)
HOST 2	192.168.20.10 (0000. 0000.bbb) (VLAN 20)
HOST 3	192.168.10.20 (0000. 0000.ccc) (VLAN 10)
HOST 4	192.168.20.20 (0000. 0000.ddd) (VLAN 20)
VLAN 10	L2VNI 100010
VLAN 20	L2VNI 100020

VLAN 500	L3VNI 50000

Configurazione

Esempio di rete



Configurazioni

- I quartieri di Underlay e PIM sono già stati creati.

Interruttore FOGLIA:

```
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
```

Abilitazione di Open Shortest Path First (OSPF) sullo switch foglia

```
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
```

Abilitazione di PIM sullo switch foglia

```
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#
```

Router adiacente OSPF

Interruttore dorso:

```
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
```

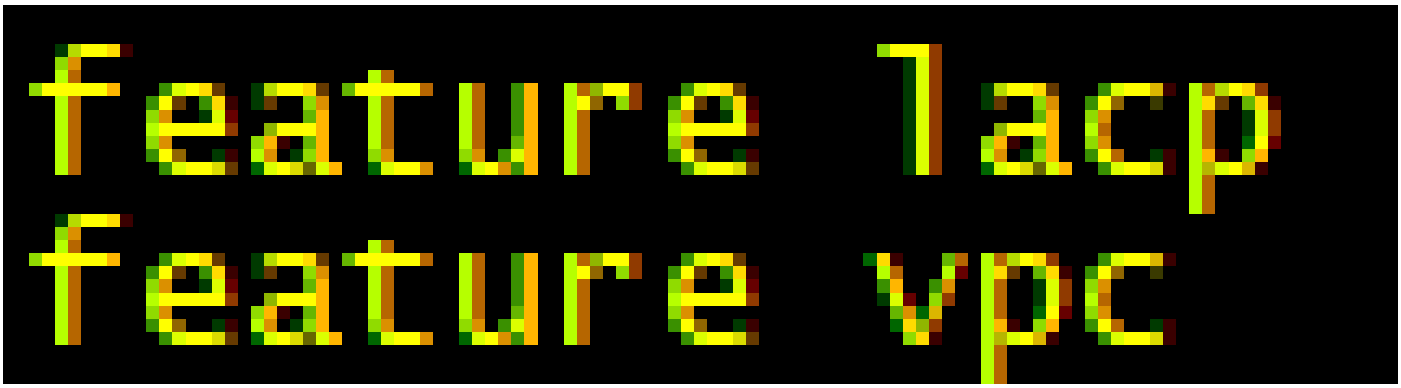
Abilitazione di PIM su Spine Switch

- I quartieri di Underlay e PIM sono già stati creati.
- Entrambi gli switch Spine saranno identici all'Anycast RP per l'intero gruppo multicast 24.0.0.0/4.
- L'MTU (Maximum Transmission Unit) è impostata su 9000/9216 sulle interfacce tra gli switch Leaf e Spine.

Innanzitutto, è necessario configurare un vPC tra Foglia1 e Foglia2.

Passaggio 1. Abilitazione delle funzionalità e dei domini vPC.

- Abilitare la funzionalità vPC e Link Aggregation Control Protocol (LACP).
- Configurare il dominio vPC.
- Le interfacce di gestione 0 vengono usate come collegamento peer keepalive e Eth1/3 ed Eth1/4 saranno la parte del collegamento peer vPC (Port-Channel 1).
- Verificare che il comando peer-switch sia configurato per condividere un indirizzo MAC comune con switch decrescenti.



Abilitazione della funzionalità sullo switch foglia

```
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
```

Abilitazione di vPC sullo switch foglia 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
```

Abilitazione di vPC sullo switch foglia 2

Passaggio 2. Assegnazione dei membri della porta.

- Assegnare il membro della porta al gruppo di canali e includerlo nel vPC. In questo caso, vengono utilizzati due vPC. 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#
```

Assegnazione del canale della porta sullo switch foglia 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#
```

Assegnazione del canale della porta sullo switch foglia 2

- Qui viene creato un vPC e i peer iniziano a scambiarsi messaggi keepalive per verificare la disponibilità.

```
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#
```

Stato vPC sullo switch foglia 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#
```

Stato vPC sullo switch foglia 2

- Le VLAN 10, 20 e 500 sono già configurate e passate sulle porte membro vPC e sul collegamento peer vPC.

Passaggio 3. Configurare l'indirizzo IP secondario.

- Quando vPC è incluso nell'infrastruttura VXLAN, entrambi i peer VTEP vPC iniziano a utilizzare indirizzi IP virtuali (VIP) come indirizzi di origine anziché gli indirizzi IP fisici (PIP). Ciò significa anche che quando BGP Ethernet VPN (EVPN) annuncia i tipi di route 2 (annuncio MAC/IP) e 5 (prefisso IP-route) per impostazione predefinita, l'indirizzo VIP viene utilizzato come hop successivo. Nell'esempio riportato, l'interfaccia di loopback 0 è configurata con due indirizzi IP: 10.1.1.100/32 (VIP) come IP secondario e 1.1.1.51/32 (PIP) come IP primario.
- In questo caso, un indirizzo IP comune viene configurato come secondario nell'interfaccia 0 di loopback.

```
LEAF-1# sh run int lo

!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 secondario sullo switch foglia 1

```
LEAF-2# sh run int lo

!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 secondario sullo switch foglia 2

Passaggio 4. Abilitare la VXLAN e le funzionalità correlate.

- Sovrapposizione Network Virtualization (nV): abilita VXLAN
- Sovrapposizione nV feature EVPN- abilita il piano di controllo EVPN
- Inoltro fabric funzionalità - abilita Host Mobility Manager
- Funzionalità VLAN (Virtual Network)-segment-VLAN-based: abilita VXLAN basata su 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#
```

Funzioni su Leaf Switch

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

Funzioni di Spine Switch

- Poiché il dorso non richiede la conoscenza delle informazioni sulla VLAN del client, non è necessario abilitare le funzionalità del segmento VN e della struttura.

Passaggio 5. Avviare il vicinato BGP.

- È necessario abilitare il protocollo BGP tra gli switch Leaf e Spine. La direttrice fungerà da riflettore di percorso in laboratorio.
- Sebbene sia facoltativo configurare Route Reflector (RR), per motivi di scalabilità Cisco consiglia di utilizzare 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
```

Abilitazione di BGP su switch foglia

```
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#
```

Abilitazione di BGP su Spine Switch

```
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#
```

Stato BGP su switch foglia

```
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#
```

Stato BGP su Spine Switch

Passaggio 6. Abilitare il contesto VRF sugli switch foglia. VRF separa il traffico del cliente e facilita la comunicazione tra due L2VNI distinti tramite L3VNI.

- Allocare L3VNI 50000 in 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
```

Allocazione L3VNI

Passaggio 7. Configurazione di NVE (Network Virtual Interface), VNI (VXLAN Identifier) e VLAN.

- Impostare l'interfaccia NVE, utilizzando Loopback 0 come origine. Definire il gruppo multicast per ciascun VNI, in cui verrà consegnato il traffico broadcast di layer 2, unicast sconosciuto e multicast (BUM), quindi collegare gli ID VNI 100010 e 100020 all'interfaccia NVE. L'intestazione VXLAN contiene le informazioni usate dal VNI per identificare i segmenti VXLAN a cui appartiene.
- L3VNI 50000 è collegato all'istanza VRF (quando viene inviato allo switch spine, VNI 50000 è stato collegato alla tabella VRF).
- Il comando host-reachability protocol BGP attiva la famiglia di indirizzi EVPN nel tunnel VXLAN, ossia gli indirizzi MAC e gli indirizzi IP vengono appresi dal protocollo BGP nel control plane e non nel data plane.
- Configurare suppress-arp nell'interfaccia NVE.
- Collegare la VLAN di layer 2 e layer 3 alla VNI pertinente.

Protocollo ARP (Suppress-Address Resolution Protocol):

Il control plane EVPN Multi-Protocol (MP)-BGP offre un miglioramento denominato soppressione ARP per ridurre il sovraccarico della rete causato dal traffico broadcast proveniente dalle richieste ARP. Ogni VTEP di un VNI mantiene una tabella della cache di soppressione ARP per gli host IP noti e gli indirizzi MAC corrispondenti nel segmento VNI quando la soppressione ARP è abilitata per quel VNI. Il VTEP locale intercetta la richiesta ARP e cerca l'indirizzo IP con risoluzione ARP nella relativa tabella della cache di eliminazione ARP ogni volta che un host finale nel VNI invia una richiesta ARP per un altro indirizzo IP dell'host finale. Per conto dell'host remoto, il VTEP locale invia una risposta ARP se rileva una corrispondenza. La risposta ARP fornisce quindi all'host locale l'indirizzo MAC degli host remoti. La richiesta ARP viene inviata agli altri VTEP nel VNI se il VTEP locale non ha l'indirizzo IP risolto in ARP nella relativa tabella di soppressione ARP. Per la prima richiesta ARP a un host di rete silenzioso, questa inondazione ARP può avvenire.

```
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#
```

Interfaccia 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#
```

Mapping da VLAN a segmento VN

- Inviando a Spine un messaggio di unione PIM, l'interfaccia NVE si unirà ai gruppi multicast 239.0.0.10 e 239.0.0.20, rispettivamente, non appena viene avviata.
- Si possono vedere anche altre tabelle (S, G) (1.1.1.54,239.0.0.20) e (10.1.1.100, 239.0.0.10/239.0.0.20) nell'immagine e quelle sono già registrate con Spine da diversi Leaf Switch.

```
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#
```

Tabella Mroute

Passaggio 8. Abilitare l'istanza EVPN.

- Abilitare l'istanza EVPN insieme alla famiglia di indirizzi per EVPN e VRF in 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
```

Istanza EVPN

- L'unico scopo della mappa del percorso REDIST è permettere tutto.
- Utilizzando il comando redistribute direct, le route connesse compatibili con VRF vengono promosse a MP-BGP (route di tipo 5).
- La configurazione EVPN mostrata sopra è identica all'istruzione network utilizzata da BGP per annunciare le route MAC (route di tipo 2).

Passaggio 9. Configurare lo switch con interfaccia virtuale (SVI) per ciascuna VLAN per l'host terminale in VRF.

- Su ciascuno switch foglia, la SVI è configurata per la VLAN configurata localmente e una SVI per la VLAN L3VNI al fine di ottenere il database RIB (Symmetric Routing Information Base).

NERVATURA simmetrica:

- Quando l'host finale invia il pacchetto di dati a una rete diversa e lo riceve allo switch foglia, viene prima elaborato in L2VNI, quindi viene posizionato in L3VNI utilizzando VRF e inviato alla foglia remota.
- Remote Leaf riceve prima i pacchetti nella tabella VRF utilizzando Routing, quindi esegue il bridging a L2VNI e li invia all'host finale.
- In questo modo, viene ottenuto il routing simmetrico (B-R-R-B).

```
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#
```

Interfacce VLAN

- il comando IP forward sulla VLAN 500 viene usato per abilitare l'inoltro di layer 3 per tutte le VXLAN. Non è necessario configurare l'indirizzo IP, in quanto elabora semplicemente il pacchetto dalla tabella L2VNI alla tabella 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#
```

Apprendimento delle route BGP VPNv4 per VRF TENANT1

- L'indirizzo IP di ciascuna VLAN sarà comune a tutte le SVI su tutti gli switch foglia. Si chiama

anycast IP e viene usato nella gestione della mobilità, dove l'estremità può comunicare con un altro host senza interruzioni.

Passaggio 10. Abilitare l'indirizzo MAC del gateway anycast di inoltro dell'infrastruttura per l'host finale.

- Assicura la ridondanza del gateway di layer 3 e l'inoltro ottimizzato per i dispositivi collegati al fabric.
- L'indirizzo MAC di Anycast Gateway è un indirizzo MAC uniforme a livello globale utilizzato per tutti i gateway di layer 3 di un'infrastruttura.
- Il concetto è identico a quello utilizzato nel protocollo FHRP (First Hop Redundancy Protocol), in cui a ogni gruppo viene assegnato un MAC virtuale.

```
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#
```

Abilitazione dell'inoltro fabric

Passaggio 11. Abilitare la VLAN di accesso/trunk sulle porte 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#
```

Abilitazione delle porte trunk all'interfaccia membro vPC

Switch non 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#
```

Abilitazione delle porte trunk sull'interfaccia membro non vPC

Verifica

- Controllare la tabella degli indirizzi ARP e 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#
```

Tabella ARP e MAC su LEAF Switch 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#
```

Tabella ARP e MAC su LEAF Switch 2

- Entrambi i peer mantengono le voci ARP.
- Controllare lo stato dell'interfaccia virtuale di rete (NVI).

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#
```

Peer NVE su switch vPC

Switch non 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#
```

Peer NVE su switch non vPC

- In questo caso, l'IP del peer è 10.1.1.100 invece dell'indirizzo IP di loopback primario, quindi il pacchetto restituito verrà instradato per quell'IP a uno degli switch vPC.
- Controllare le route VPN BGP.

```
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 VPN 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#
```

MAC EVPN BGP l2route

```
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#
```

Riepilogo VPN BGP

```
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#
```

Route VPN BGP

- È normale chiedersi in che modo gli switch foglia acquisiscano le voci MAC per gli host remoti. Questo processo è facilitato dall'ARP Gratuito. Quando una porta di rete è attivata, invia immediatamente una richiesta ARP per verificare l'univocità dell'indirizzo IP. Ogni switch foglia registra quindi l'indirizzo MAC e lo include in un pacchetto di aggiornamento BGP. Questo consente ad altri switch foglia di aggiornare le rispettive tabelle degli indirizzi MAC di conseguenza. Tuttavia, è possibile che l'host finale non generi un ARP Gratuito (host silenzioso) e, in questo caso, la richiesta ARP venga trasmessa alla foglia e, trattandosi di una richiesta broadcast, lo switch Leaf genera la richiesta multicast al rispettivo gruppo per il VNI in questione. In questo caso, sono 239.0.0.10 e 239.0.0.20.

- Consente di eseguire il ping tra l'host 1 e l'host 3 all'interno dello stesso VNI e di esaminare l'acquisizione.

```
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 da HOST-1 a HOST-3

Pacchetto Internet Control Message Protocol (ICMP) sulla 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
```

Acquisizione di Wireshark con richiesta ICMP su un pacchetto in viaggio tramite L2VNI 1010

- Come si può vedere, l'IP di origine è 10.1.1.100 con la porta 4789 come destinazione UDP.
- Poiché si tratta di una comunicazione intra VNI, la VLAN 10 utilizzerà il VNI 100010 e la VLAN 20 il VNI 1000.
- Consente di eseguire il ping tra l'host 1 e l'host 4 con VNI diversi e di esaminare l'acquisizione.

```
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 da HOST-1 a HOST-4

Pacchetto ICMP sulla 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
```

Acquisizione di Wireshark con richiesta ICMP su un pacchetto in viaggio tramite L3VNI 5000

- Poiché si tratta di una comunicazione inter-VNI, verrà utilizzato L3VNI 50000.
- Controllare la tabella ARP per l'host finale.

```
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#
```

Voci 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#
```

Voci 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#
```

HOST-3 Voci ARP

```
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#
```

HOST-4 Voci ARP

```
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 da HOST-4 a tutti gli altri host terminali

Informazioni su questa traduzione

Cisco ha tradotto questo documento utilizzando una combinazione di tecnologie automatiche e umane per offrire ai nostri utenti in tutto il mondo contenuti di supporto nella propria lingua. Si noti che anche la migliore traduzione automatica non sarà mai accurata come quella fornita da un traduttore professionista. Cisco Systems, Inc. non si assume alcuna responsabilità per l’accuratezza di queste traduzioni e consiglia di consultare sempre il documento originale in inglese (disponibile al link fornito).