

vPCを使用した高度なVXLAN:L2VNIおよびL3VNIの設定と検証

内容

[はじめに](#)

[前提条件](#)

[要件](#)

[使用するコンポーネント](#)

[背景説明](#)

[設定](#)

[ネットワーク図](#)

[コンフィギュレーション](#)

[検証](#)

はじめに

このドキュメントでは、Virtual Port-Channel(vPC)でAdvanced Virtual eXtensible Local Area Network(VXLAN)を使用して、Nexus 9Kvスイッチを使用するラボをセットアップする方法について説明します。

前提条件

要件

次の項目に関する知識があることが推奨されます。

- ルーティングとスイッチング、およびマルチプロトコルラベルスイッチング(MPLS)テクノロジーの理解
- ランデブーポイント(RP)やPlatform Independent Multicast(PIM)などのマルチキャストルーティングの原則に関する経験
- Border Gateway Protocol(BGP)アドレスファミリインジケータ(AFI)/後続のアドレスファミリインジケータ(SAFI)の理解

使用するコンポーネント

このドキュメントの内容は、特定のソフトウェアやハードウェアのバージョンに限定されるものではありません。

このドキュメントの情報は、特定のラボ環境にあるデバイスに基づいて作成されました。このドキュメントで使用するすべてのデバイスは、クリアな（デフォルト）設定で作業を開始しています。本稼働中のネットワークでは、各コマンドによって起こる可能性がある影響を十分確認してください。

背景説明

また、ラボの導入に関するガイダンス、および設定と動作の検証についても説明します。

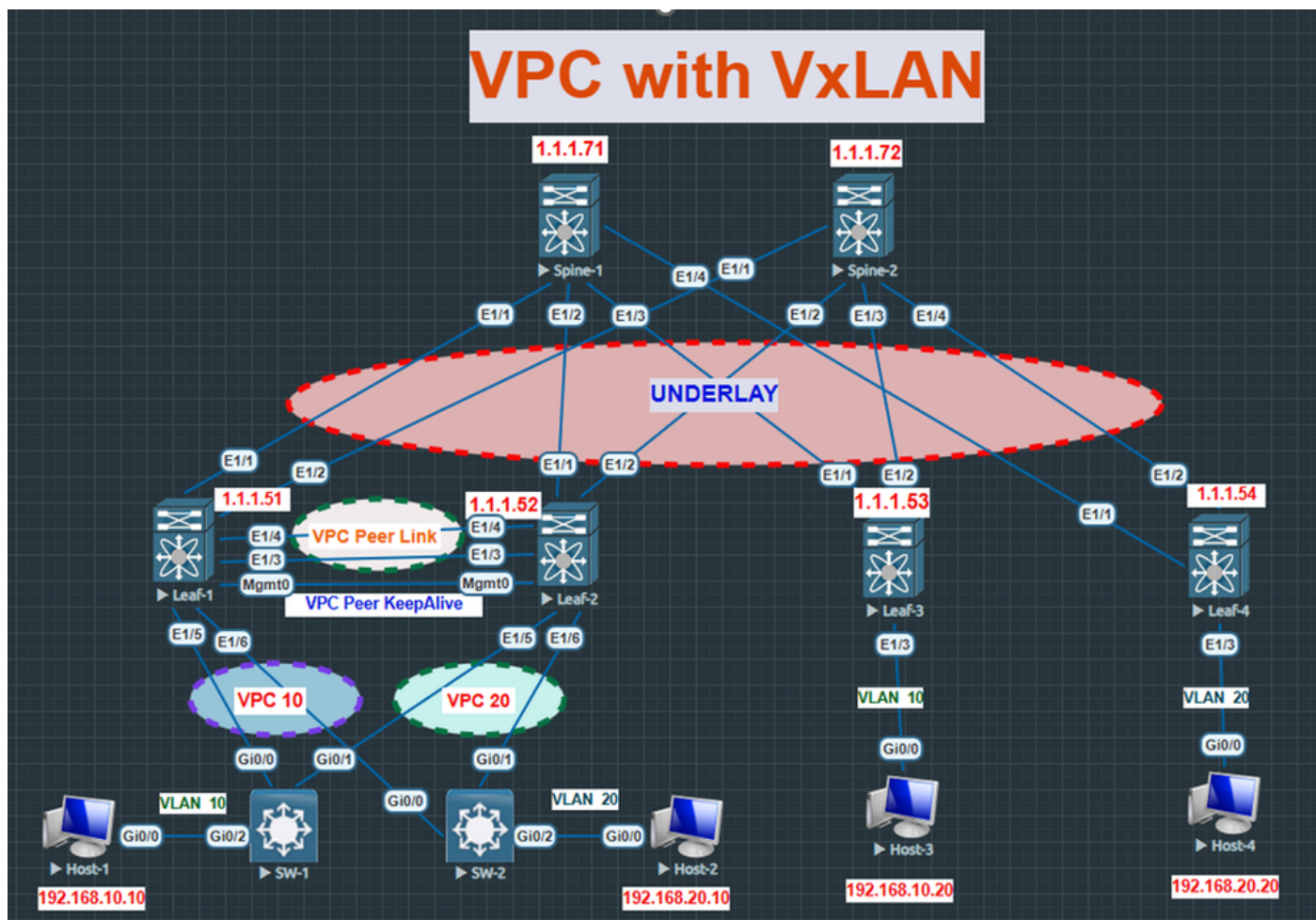
このラボでは、リーフとスパインの両方にNexus 9000Vスイッチを使用したEveNgを使用します。

仮想トンネルエンドポイント(VTEP)	リーフ1、リーフ2、リーフ3、リーフ4
vPC	LEAF1およびLEAF2
LEAF1プライマリおよびセカンダリループバックIP	ループバック0 - 1.1.1.51、ループバック1 - 10.1.1.100
LEAF2プライマリおよびセカンダリループバックIP	ループバック0 - 1.1.1.52、ループバック1 - 10.1.1.100
リーフ3ループバックIP	1.1.1.53
リーフ4ループバックIP	1.1.1.54
SPINE1ループバックおよびエニーキャストRP	Loopback0 - 1.1.1.71、Loopback1 - 10.1.2.10 (エニーキャストRP)
SPINE2ループバックおよびエニーキャストRP	Loopback0 - 1.1.1.72、Loopback1 - 10.1.2.10 (エニーキャストRP)
ホスト1	192.168.10.10(0000.0000.aaaa)(VLAN 10)
ホスト2	192.168.20.10(0000.0000.bbbb)(VLAN 20)
ホスト3	192.168.10.20(0000.0000.cccc)(VLAN 10)
ホスト4	192.168.20.20(0000.0000.dddd)(VLAN 20)
VLAN 10	L2VNI 100010

VLAN 20	L2VNI 100020
VLAN 500	L3VNI 50000

設定

ネットワーク図



コンフィギュレーション

- アンダーレイとPIMネイバーシップは既に確立されています。

リーフスイッチ：

```
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
```

リーフスイッチでのOpen Shortest Path First(OSPF)の有効化

```
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
```

リーフスイッチでのPIMの有効化

```
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
                  Interface      Uptime     Expires     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#
```

OSPFネイバー

スパインスイッチ：

```
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
```

スパインスイッチでのPIMの有効化

- ・ アンダーレイとPIMネイバーシップは既に確立されています。
- ・ 両方のスパインスイッチは、マルチキャストグループ224.0.0.0/4全体で同一のユニキャストRPになります。
- ・ リーフ/スパイン型スイッチ間のインターフェイスでは、最大伝送ユニット(MTU)が9000/9216に設定されています。

まず、Leaf1とLeaf2の間にvPCを設定します。

ステップ1:vPC機能とドメインの有効化。

- ・ vPCおよびLink Aggregation Control Protocol(LACP)機能を有効にします。
- ・ vPCドメインを設定します。
- ・ mgmt 0インターフェイスはピアキーブアライブリンクとして使用され、Eth1/3およびEth1/4はvPCピアリンク (ポートチャネル1) の一部になります。
- ・ peer-switchコマンドが、降順スイッチと共通のMACアドレスを共有するように設定されていることを確認します。

feature 1 acp feature vpc

リーフスイッチの機能の有効化

```
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
```

リーフスイッチ1のvPCの有効化

```
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
```

リーフスイッチ2のvPCの有効化

ステップ 2：ポートメンバーの割り当て。

- ポートメンバーをチャネルグループに割り当て、vPCに含めます。この例では、2台のvPC (vPC 20とvPC 10) が使用されています。

```
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#
```

リーフスイッチ1のポートチャネルの割り当て

```
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#
```

リーフスイッチ2でのポートチャネルの割り当て

- ここでvPCが作成され、ピアは可用性を確認するためにキープアライブメッセージの交換を開始します。

```

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#

```

リーフスイッチ1のvPCステータス

```

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# █

```

リーフスイッチ2のvPCステータス

- VLAN 10、20、500はすでに設定されており、vPCメンバーポートおよびvPCピアリンク経由で渡されています。

ステップ 3：セカンダリIPアドレスを設定します。

- vPCがVXLANファブリックに含まれている場合、両方のvPC VTEPピアは、物理IPアドレス(PIP)ではなく、送信元アドレスとして仮想IP(VIP)アドレスの使用を開始します。また、BGPイーサネットVPN(EVPN)がルートタイプ2 (MAC/IPアドバタイズメント) と5 (IPプレフィックスルート) をデフォルトでアドバタイズする際に、VIPがネクストホップとして使用されることも意味します。この例のLoopback 0インターフェイスは、セカンダリIPとして10.1.1.100/32(VIP)とプライマリIPとして1.1.1.51/32(PIP)の2つのIPアドレスで設定されています。
- ここでは、共通のIPアドレスが、ループバック0インターフェイスの下でセカンダリIPアドレスとして設定されています。

```
LEAF-1# sh run int 10
```

```
!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#
```

リーフスイッチ1のセカンダリIP

```
LEAF-2# sh run int 10
```

```
!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#
```

リーフスイッチ2のセカンダリIP

ステップ 4 : VXLANおよび関連機能を有効にします。

- ネットワーク仮想化(nV)オーバーレイ : VXLANを有効にします。
- 機能nVオーバーレイEVPN:EVPNコントロールプレーンを有効化
- ファブリックフォワーディング機能 : Host Mobility Managerを有効にします。
- 仮想ネットワーク(VN)セグメントVLANベースの機能 – VLANベースのVXLANを実現

```
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#
```

リーフスイッチの機能

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

スパインスイッチの機能

- スパインはクライアントのVLAN情報を必要としないため、VNセグメントおよびファブリック機能を有効にする必要はありません。

ステップ 5： BGPネイバーシップを確立します。

- リーフスイッチとスパインスイッチ間のBGPを有効にする必要があります。スパインは、ラボ内でルートリフレクタとして機能します。
- ルートリフレクタ(RR)の設定はオプションですが、スケーラビリティの観点から、RRを使用することをお勧めします。

```
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
```

リーフスイッチでのBGPの有効化

```

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# █

```

スパインスイッチでのBGPの有効化

```

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#

```

リーフスイッチのBGPステータス

```

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#

```

スパインスイッチのBGPステータス

手順 6：リーフスイッチでVRFコンテキストを有効にします。VRFは顧客トラフィックを分離し、L3VNIを介した2つの異なるL2VNI間の通信を促進します。

- VRF TENANT1の下にL3VNI 50000を割り当てます。

```

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
router-id 1.1.1.1

```

L3VNI割り当て

手順 7：ネットワーク仮想インターフェイス(NVE)、VXLAN識別子(VNI)、およびVLAN設定。

- Loopback 0を送信元として使用して、NVEインターフェイスを設定します。レイヤ2ブロードキャスト、不明なユニキャスト、およびマルチキャスト(BUM)トラフィックが配信される各VNIのマルチキャストグループを定義してから、VNI 100010と100020 IDをNVEインターフェイスに接続します。VXLANヘッダーには、VNIが属するVXLANセグメントを識別するために使用する情報が含まれています。
- L3VNI 50000はVRFインスタンスにリンクされています(スパインスイッチに送信する際に、VNI 50000がVRFテーブルにアタッチされています)。
- host-reachability protocol BGPコマンドは、VXLANトンネルのEVPNアドレスファミリをアクティブにします。これは、MACアドレスとIPアドレスがデータプレーンではなく、コントロールプレーンのBGPプロトコルによって学習されることを意味します。
- NVEインターフェイスでsuppress-arpを設定します。
- 関連するVNIにレイヤ2およびレイヤ3 VLANを接続します。

アドレス解決プロトコル(ARP)の抑制：

マルチプロトコル(MP)-BGP EVPNコントロールプレーンでは、ARP要求からのブロードキャストトラフィックによって引き起こされるネットワークフラッディングを軽減するために、ARP抑制と呼ばれる改良が行われています。各VNI VTEPは、既知のIPホストに関するARP抑制キャッシュテーブルと、そのVNIでARP抑制が有効になっている場合にVNIセグメントでそれらに対応するMACアドレスを保持します。そのローカルVTEPは、ARP要求を代行受信し、VNIのエンドホストが別のエンドホストIPアドレスのARP要求を送信するたびに、ARP抑制キャッシュテーブルでARPによって解決されたIPアドレスを探します。リモートエンドホストの代わりに、ローカルVTEPは、一致を検出した場合にARP応答を送信します。ARP応答は、ローカルホストにリモートホストのMACアドレスを提供します。ローカルVTEPのARP抑制テーブルにARP解決IPアドレスがない場合、ARP要求はVNI内の他のVTEPにフラッディングされます。サイレントネットワークホストへの最初のARP要求では、このARPフラッディングが発生する可能性があります。

```
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# █
```

NVEインターフェイス

```

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#

```

VLANからVNセグメントへのマッピング

- ・ スパインにPIM joinメッセージを送信することで、NVEインターフェイスはブートアップするとすぐに、マルチキャストグループ239.0.0.10と239.0.0.20にそれぞれ参加します。
- ・ イメージ内に他の(S, G)テーブル(1.1.1.54,239.0.0.20)と(10.1.1.100, 239.0.0.10/239.0.0.20)も表示され、これらのテーブルは別のリーフスイッチのスパインにすでに登録されています。

```

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#

```

Mrouteテーブル

ステップ 8 : EVPNインスタンスを有効にします。

- ・ BGPでEVPNおよびVRFのアドレスファミリとともにEVPNインスタンスを有効にします。

```

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 12
    rd auto
    route-target import auto
    route-target export auto
  vni 100020 12
    rd auto
    route-target import auto
    route-target export auto
vrf context TENANT1

```

EVPNインスタンス

- ルートマップREDISTの唯一の目的は、すべてを許可することです。
- redistribute directコマンドを使用すると、接続されたVRF対応ルートがMP-BGP (タイプ5ルート) にプロモートされます。
- 上記で表示されたEVPN設定は、BGPがMACルート (タイプ2ルート) をアドバタイズするために使用するnetwork文と同じです。

ステップ 9 : VRFの下で、エンドホストの各VLANにスイッチ仮想インターフェイス(SVI)を設定します。

- 対称ルーティング情報ベース(RIB)を実現するために、各リーフスイッチで、ローカルに設定されたVLANに対してSVIが、L3VNI VLANに対して1つのSVIが設定されています。

対称リブ :

- エンドホストが別のネットワークにデータパケットを送信し、リーフスイッチに受信すると、最初にL2VNIに処理されてから、VRFを使用してL3VNIに配置され、リモートリーフに送信されます。
- リモートリーフは、まずルーティングを使用してVRFテーブル内のパケットを受信し、次にL2VNIへのブリッジングを行い、それをエンドホストに送信します。
- これにより、対称ルーティング(B-R-R-B)が実現されます。

```

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# █

```

VLAN インターフェイス

- VLAN 500でIP forwardコマンドを使用すると、すべてのVXLANに対してレイヤ3転送が有効になります。L2VNIテーブルからL3VNIテーブルへのパケットを処理するだけなので、IPアドレスを設定する必要はありません。

```

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                0           100           0 i
*>r192.168.20.0/24        0.0.0.0                0           100         32768 ?
*>i192.168.20.20/32       1.1.1.54                0           100           0 i

LEAF-1# █

```

VRF TENANT1のBGP VPNv4ルートの学習

- 各VLANのIPアドレスは、すべてのリーフスイッチのすべてのSVIに共通です。これは工二

ーキャストIPと呼ばれ、エンドが中断なく別のホストとシームレスに通信できるモビリティ管理で使用されます。

ステップ 10：エンドホストのファブリックフォワーディングエニーキャストゲートウェイMACを有効にします。

- これにより、シームレスなレイヤ3ゲートウェイの冗長性が確保され、ファブリックに接続されたデバイスの転送が最適化されます。
- エニーキャストゲートウェイMACアドレスは、ファブリック内のすべてのレイヤ3ゲートウェイに使用される、グローバルに一貫性のあるMACアドレスです。
- この概念は、First Hop Redundancy Protocol(FHRP)で採用されている概念と同じです。FHRPでは、各グループに仮想MACが発行されます。

```
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#
```

ファブリック転送の有効化

ステップ 11メンバーポートへのアクセス/トランクVLANを有効にします。

vPCスイッチ：

```
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#
```

vPCメンバーインターフェイスへのトランクポートの有効化

非vPCスイッチ：

```
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#
```

非vPCメンバーインターフェイスへのトランクポートの有効化

検証

- ARPおよびMACアドレステーブルをチェックします。

```
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 Addr
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#
```

リーフスイッチ1のARPおよびMACテーブル

```
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#
```

リーフスイッチ2のARPおよびMACテーブル

- 両方のピアがARPエントリを維持します。
- ネットワーク仮想インターフェイス(NVI)のステータスを確認します。

vPCスイッチ：

```
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#
```

vPCスイッチ上のNVEピア

非vPCスイッチ：

```
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#
```

非vPCスイッチ上のNVEピア

- ここでは、ピアIPがプライマリループバックIPアドレスではなく10.1.1.100であることに気付きます。そのため、戻りのパケットはこのIP用にvPCスイッチのいずれかにルーティングされます。
- BGP EVPNルートをチェックします。

```
LEAF-1# show iproute 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 I2route 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#
```

BGP EVPNの概要

```

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                        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#

```

BGP EVPNルート

- リーフスイッチがリモートホストのMACエントリを取得する方法を疑問に思われがちです。このプロセスは、Gratuitous ARPによって促進されます。ネットワークポートがアクティブになると、IPアドレスが一意であることを確認するために、ただちにARP要求が送信されます。各リーフスイッチはMACアドレスを記録し、BGPアップデートパケットに含めます。これにより、他のリーフスイッチはそれぞれのMACアドレステーブルを適宜更新できます。ただし、エンドホストがGratuitous ARP (サイレントホスト) を生成しない場合があります。その場合、ARP要求はリーフにブロードキャストされ、これがブロードキャスト要求であるため、リーフスイッチは特定のVNIの各グループにマルチキャスト要求を生成します。この例では、239.0.0.10と239.0.0.20です。
- 同じVNI内でHost-1からHost-3にpingを実行し、キャプチャを確認します。

```

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#

```

ホスト1からホスト3へのping

VXLAN上のインターネット制御メッセージプロトコル(ICMP)パケット：

```

> 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
✓ 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)
✓ 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)
✓ 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

```

L2VNI 10010を通過するICMP要求パケットを示すWiresharkキャプチャ

- 送信元IPは10.1.1.100で、UDPの宛先はポート4789であることがわかります。
- VNI内通信であるため、VLAN 10はVNI 100010を使用し、VLAN 20はVNI 1000を使用します。
- 異なるVNIを使用してHost-1からHost-4にpingを実行し、キャプチャを確認します。

```

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#

```

ホスト1からホスト4へのping

VXLAN上のICMPパケット：

```

> 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
▼ 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)
▼ 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

```

L3VNI 50000を通過するICMP要求パケットを示すWiresharkキャプチャ

- これはVNI間通信であるため、L3VNI 50000が使用されます。
- エンドホストのARPテーブルをチェックします。

```

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#

```

ホスト1 ARPエントリ

```

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#

```

ホスト2 ARPエントリ

```

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#

```

ホスト3 ARPエントリ

```

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 ARPエントリ

```
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)#
```

HOST-4から他のすべてのエンドホストへのping

翻訳について

シスコは世界中のユーザにそれぞれの言語でサポート コンテンツを提供するために、機械と人による翻訳を組み合わせて、本ドキュメントを翻訳しています。ただし、最高度の機械翻訳であっても、専門家による翻訳のような正確性は確保されません。シスコは、これら翻訳の正確性について法的責任を負いません。原典である英語版（リンクからアクセス可能）もあわせて参照することを推奨します。