

# 配置VXLAN泛洪并使用组播核心学习

## 目录

[简介](#)

[先决条件](#)

[要求](#)

[使用的组件](#)

[背景信息](#)

[VXLAN的数据包格式](#)

[远程VTEP发现](#)

[配置](#)

[网络图](#)

[9396-A配置](#)

[9396-B配置](#)

[9508-A配置](#)

[9396-C配置](#)

[验证](#)

[流量在对等体之间启动后的状态](#)

[故障排除](#)

## 简介

本文档介绍如何配置和验证虚拟可扩展局域网(VXLAN)泛洪和IPv4组播传输的学习模式。

## 先决条件

### 要求

思科建议您了解基本IP组播。

### 使用的组件

本文档中的信息基于Nexus平台。

本文档中的信息都是基于特定实验室环境中的设备编写的。本文档中使用的所有设备最初均采用原始(默认)配置。如果您的网络处于活动状态,请确保您了解所有命令的潜在影响。

## 背景信息

VXLAN的设计目的是提供与VLAN相同的以太网第2层网络服务。VXLAN封装MAC地址通过UDP数据包,使第2层数据包通过第3层网络传输。因此,它基本上是MAC-in-UDP报头。

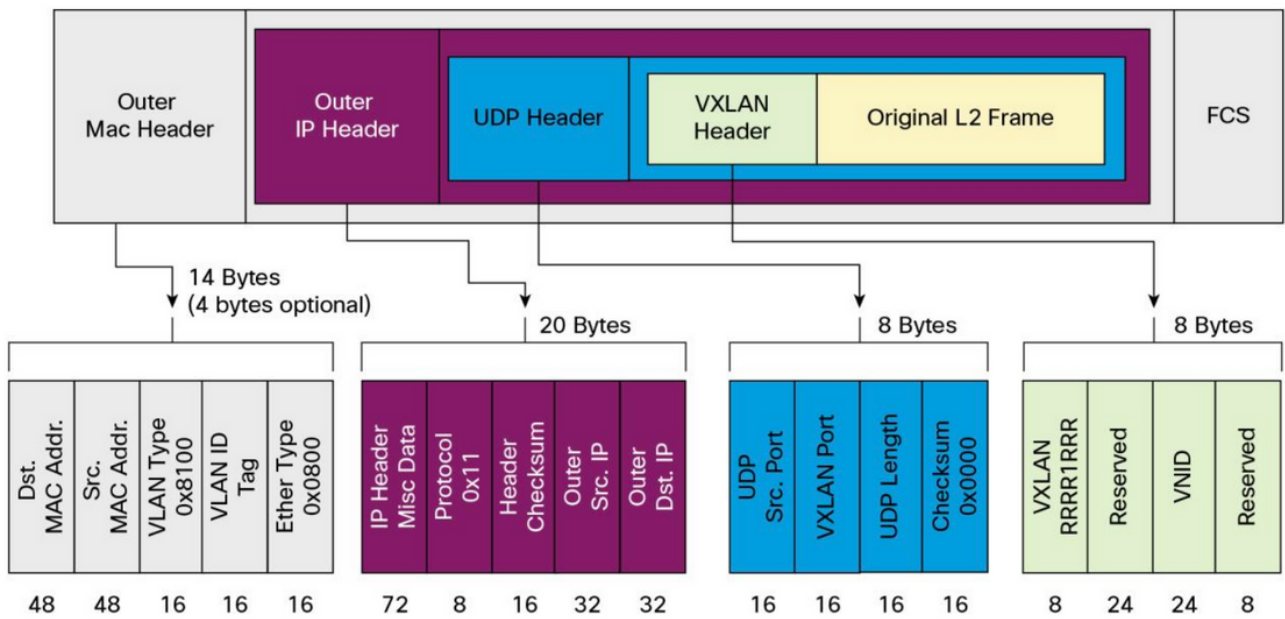
VXLAN引入了一个8字节的VXLAN报头,该报头由24位VXLAN网络标识符(VNID)和几个保留位组成。VXLAN报头与原始以太网帧一起进入UDP负载中。24位VNID用于识别第2层网段并保持网段之间

的第2层隔离。VXLAN在VNID中全部包含24位，可支持1600万个LAN网段。从而解决了VLAN限制的问题。没有VxLAN，您只能拥有4094个VLAN，随着需求的增加，现代网络需要更多的VLAN，而VXLAN是解决该问题的解决方案。

由于它使用以太网帧来封装数据包，因此以太网属性需要像广播、未知单播和组播一样保持完整。为了处理这些类型的流量，使用组播。本文档介绍VXLAN泛洪和学习。因为名称指定它泛洪数据包并学习远程端。这意味着数据平面在流量数据平面建立后并在MAC地址到期后立即过期时并非始终处于打开状态。

## VXLAN的数据包格式

Figure 1. VXLAN Packet Format



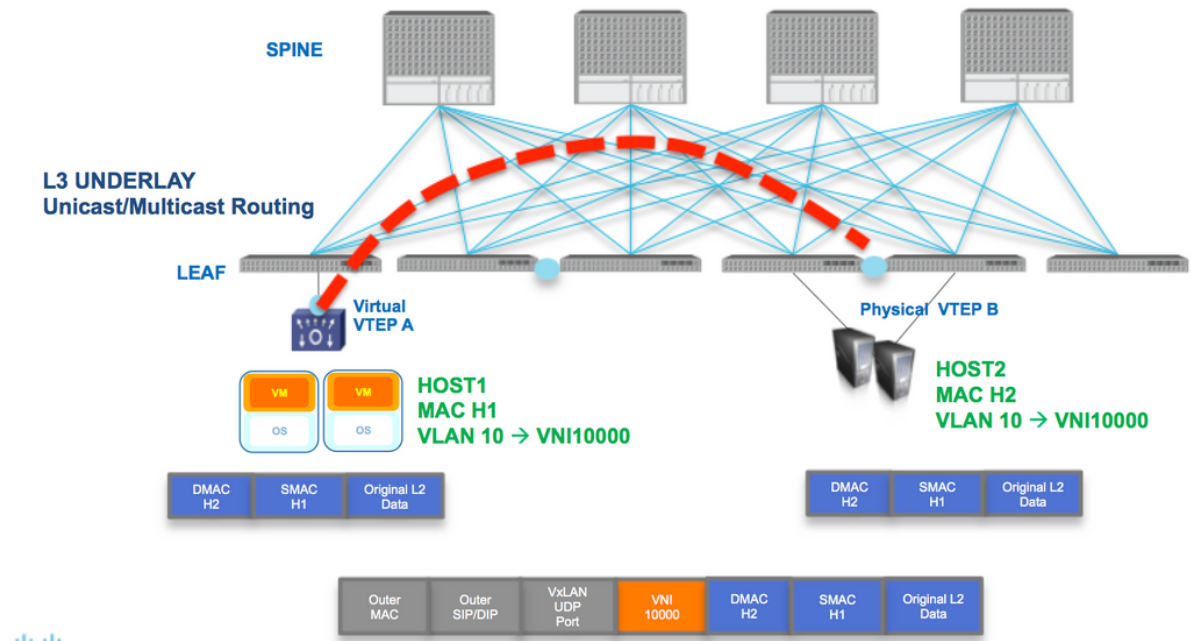
如图所示，原始帧封装在VXLAN报头中，该报头为8字节，VNID为24位。UDP报头中进一步封装，外部报头是IP报头。

源IP地址是封装虚拟终端终端(VTEP)的IP地址，目的IP可以是组播或单播IP。VXLAN使用VXLAN隧道终端(VTEP)设备将租户的终端设备映射到VXLAN网段，并执行VXLAN封装和解封。每个VTEP有两个接口：一个是本地LAN网段上的交换机接口，用于通过桥接支持本地终端通信，另一个是到传输IP网络的IP接口。

### 远程VTEP发现

当主机开始发送流量时，遵循的过程如下所述。此时，VTEP不知道远程主机的MAC地址。

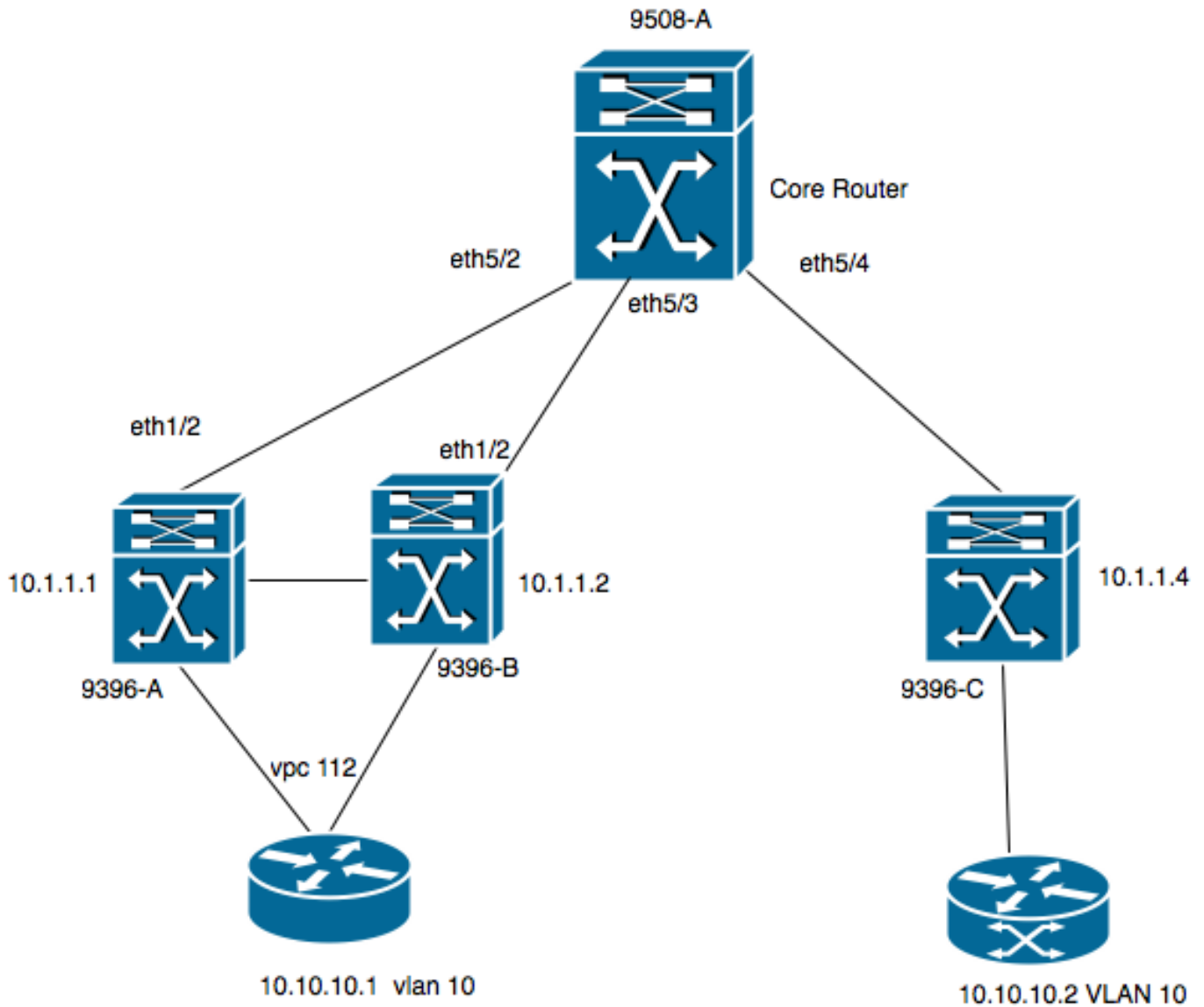
1. 终端站向远程终端站发送地址解析协议(ARP)数据包。
2. 数据包到达VTEP-A，由于VTEP-A不知道VTEP-B，因此它将数据包封装在VXLAN报头中。它将组播IP地址作为目的IP地址。由于所有VTEP都使用相同的组播地址，因此所有VTEP都加入同一组播组。
3. 此数据包到达所有VTEP并解封，这样远程VTEP就能获知其他VTEP。由于解封的VTEP具有VNID，因此它会在配置了相同VNID的VLAN中转发。
4. 现在，远程端发送ARP应答数据包并到达VTEP-B，因为现在VTEP-B知道VTEP-A，它再次封装原始帧，但现在目的IP地址是VTEP-B，它是单播IP地址。
5. ARP应答到达VTEP-A，VTEP-A现在知道VTEP-B，它与VTEP-B形成邻居关系。



如图所示，主机H1属于VLAN 10，封装在VNID 10000中。 如图所示，带H1的SMAC和带H2的DMAC封装在VNI 10000中，源IP和目标IP可以是组播或单播，如本节所述。

## 配置

## 网络图



- 9396-A和9396-B是被视为VTEP-1的VPC对等体
- 9396-C是VTEP-2
- 该图在VLAN 10中有两台主机，即10.10.10.1和10.10.10.2
- VLAN 10与VNID一起使用，为10010
- 230.1.1.1用作组播组

要在Nexus上启用VXLAN，您需要启用此功能。

### 9396-A配置

```

!
feature vn-segment-vlan-based
feature nv overlay
!
vlan 10
  vn-segment 10010  -----> 10010 is VNID
!
interface nve1
  no shutdown
  source-interface loopback0
  member vni 10010 mcast-group 230.1.1.1
!
interface eth1/2

```

```

!
ip pim sparse-mode
!
interface loopback0
 ip address 10.1.1.1/32
 ip address 10.1.1.10/32 secondary
 ip router ospf 9k area 0.0.0.0
 ip pim sparse-mode
!

```

**注意：**10.1.1.10用作辅助IP地址，并且环回必须具有辅助IP地址，仅在vPC的情况下。两个vPC对等体必须具有相同的辅助IP地址，而不同的主IP地址。

```

!
feature vpc
!
vpc domain 1
 peer-switch
 peer-keepalive destination 10.31.113.41 source 10.31.113.40
 peer-gateway
!
interface port-channel1
 vpc peer-link
!
interface port-channel112
 vpc 112
!

```

## 9396-B配置

```

!
vlan 10
 vn-segment 10010 -----> 10010 is VNID
!
interface nve1
 no shutdown
 source-interface loopback0
 member vni 10010 mcast-group 230.1.1.1
!
interface eth1/2
 ip pim sparse-mode
!
interface loopback0
 ip address 10.1.1.2/32
 ip address 10.1.1.10/32 secondary
 ip router ospf 9k area 0.0.0.0
 ip pim sparse-mode
!
feature vpc
!
vpc domain 1
 peer-switch
 peer-keepalive destination 10.31.113.40 source 10.31.113.41
 peer-gateway
!
interface port-channel1
 vpc peer-link
!
interface port-channel112
 vpc 112

```

!

## 9508-A配置

```
feature pim

ip pim rp-address 10.1.1.5 group-list 224.0.0.0/4
ip pim ssm range 232.0.0.0/8

interface loopback0
 ip pim sparse-mode

interface Ethernet5/2
 ip pim sparse-mode

interface Ethernet5/3
 ip pim sparse-mode

interface Ethernet5/4
 ip pim sparse-mode
```

**注意：**在9508上，它只需启用pim。因为这是VTEP，所以它不需要VXLAN的任何功能。

## 9396-C配置

```
!
vlan 10
 vn-segment 10010
!
interface loopback0
 ip address 10.1.1.3/32
 ip router ospf 9k area 0.0.0.0
 ip pim sparse-mode
!
interface nve1
 no shutdown
 source-interface loopback0
 member vni 10010 mcast-group 230.1.1.1
!
int eth1/2
 ip pim sparse-mode
!
```

## 验证

使用本部分可确认配置能否正常运行。

到目前为止，主机尚未开始发送数据包流。由于9396-A是VPC保持设备，因此它从辅助IP地址发起流量源，并充当组播流的源IP地址。

```
9396-A# sh nve interface
```

```
Interface: nve1, State: Up, encapsulation: VXLAN
VPC Capability: VPC-VIP-Only [notified]
Local Router MAC: d8b1.9076.9053
Host Learning Mode: Data-Plane
```

Source-Interface: loopback0 (primary: 10.1.1.1, secondary: 10.1.1.10)

9396-A# sh ip mroute 230.1.1.1

IP Multicast Routing Table for VRF "default"

(\*, 230.1.1.1/32), uptime: 01:09:34, ip pim nve  
Incoming interface: Ethernet1/2, RPF nbr: 192.168.10.2  
Outgoing interface list: (count: 1)  
nve1, uptime: 00:11:20, nve

(10.1.1.3/32, 230.1.1.1/32), uptime: 00:12:19, ip mrrib pim nve  
Incoming interface: Ethernet1/2, RPF nbr: 192.168.10.2  
Outgoing interface list: (count: 1)  
nve1, uptime: 00:11:20, nve

(10.1.1.10/32, 230.1.1.1/32), uptime: 00:11:20, nve ip mrrib pim  
Incoming interface: loopback0, RPF nbr: 10.1.1.10  
Outgoing interface list: (count: 1)  
Ethernet1/2, uptime: 00:11:20, pim

在\*中，G入口接口填充在传出接口列表(OIL)中。如图所示，10.1.1.10是组播流的源，nve接口是组播流的最后一跳路由器，面向核心的eth1/2是传出接口。

由于没有来自主机的流量，因此没有对等体：

9396-A# show mac address-table vlan 10

Legend:

\* - primary entry, G - Gateway MAC, (R) - Routed MAC, O - Overlay MAC  
age - seconds since last seen,+ - primary entry using vPC Peer-Link,  
(T) - True, (F) - False

VLAN	MAC Address	Type	age	Secure	NTFY	Ports
* 10	8c60.4f93.5ffc	dynamic	0	F	F	Pol12 >> This mac is for host 10.10.10.1

9396-A# sh nve peers

Interface	Peer-IP	State	LearnType	Uptime	Router-Mac
-----------	---------	-------	-----------	--------	------------

此输出显示了vPC输出的外观：

9396-A# sh vpc brief

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 : 1  
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)

vPC Peer-link status

```

-----
id   Port   Status Active vlans
--   ----   -
1    Po1     up     1-10

```

vPC status

```

-----
id   Port   Status Consistency Reason           Active vlans
--   ----   -
112  Po112  up     success    success           1-10

```

**9396-A# sh vpc consistency-parameters global**

Legend:

Type 1 : vPC will be suspended in case of mismatch

Name	Type	Local Value	Peer Value
Vlan to Vn-segment Map	1	1 Relevant Map(s)	1 Relevant Map(s)
STP Mode	1	Rapid-PVST	Rapid-PVST
STP Disabled	1	None	None
STP MST Region Name	1	" "	" "
STP MST Region Revision	1	0	0
STP MST Region Instance to VLAN Mapping	1		
STP Loopguard	1	Disabled	Disabled
STP Bridge Assurance	1	Enabled	Enabled
STP Port Type, Edge	1	Normal, Disabled,	Normal, Disabled,
BPDUFILTER, Edge BPDUGuard		Disabled	Disabled
STP MST Simulate PVST	1	Enabled	Enabled
<b>Nve Admin State, Src Admin State, Secondary IP, Host Reach Mode</b>	<b>1</b>	<b>Up, Up, 10.1.1.10, DP</b>	<b>Up, Up, 10.1.1.10, DP</b>
<b>Nve Vni Configuration</b>	<b>1</b>	<b>10010</b>	<b>10010</b>
<b>Nve encap Configuration</b>	<b>1</b>	<b>vxlan</b>	<b>vxlan</b>
Interface-vlan admin up	2		
Interface-vlan routing capability	2	1	1
Allowed VLANs	-	1-10	1-10
Local suspended VLANs	-	-	-

**9508-A**

由于9508-A路由是核心路由器，它不知道VXLAN，它只知道mroute条目，如下所示：

**9508-A# sh ip mroute 230.1.1.1**

IP Multicast Routing Table for VRF "default"

```

(*, 230.1.1.1/32), uptime: 01:30:06, pim ip
  Incoming interface: loopback0, RPF nbr: 10.1.1.5, uptime: 01:30:06
  Outgoing interface list: (count: 3)
    Ethernet5/3, uptime: 00:14:11, pim
    Ethernet5/2, uptime: 00:14:31, pim
    Ethernet5/4, uptime: 00:16:22, pim

```

```

(10.1.1.3/32, 230.1.1.1/32), uptime: 00:15:44, pim mrrib ip
  Incoming interface: Ethernet5/4, RPF nbr: 192.168.10.10, uptime: 00:15:44, internal

```



```

Outgoing interface list: (count: 2)
  Ethernet5/3, uptime: 00:14:11, pim
  Ethernet5/2, uptime: 00:14:31, pim
(10.1.1.10/32, 230.1.1.1/32), uptime: 00:14:31, pim mrib ip
Incoming interface: Ethernet5/2, RPF nbr: 192.168.10.1, uptime: 00:14:31, internal
Outgoing interface list: (count: 1)
  Ethernet5/4, uptime: 00:14:31, pim

```

## 9396-C

### 9396-C# show ip mroute

IP Multicast Routing Table for VRF "default"

```

(*, 230.1.1.1/32), uptime: 01:07:34, ip pim nve
Incoming interface: Ethernet1/2, RPF nbr: 192.168.10.9
Outgoing interface list: (count: 1)
  nve1, uptime: 00:10:38, nve

```

```

(10.1.1.3/32, 230.1.1.1/32), uptime: 00:10:38, nve ip mrib pim
Incoming interface: loopback0, RPF nbr: 10.1.1.3
Outgoing interface list: (count: 1)
  Ethernet1/2, uptime: 00:09:49, pim

```

```

(10.1.1.10/32, 230.1.1.1/32), uptime: 00:08:05, ip mrib pim nve
Incoming interface: Ethernet1/2, RPF nbr: 192.168.10.9
Outgoing interface list: (count: 1)
  nve1, uptime: 00:08:05, nve

```

## 流量在对等体之间启动后的状态

当主机1 ( 即10.10.10.1 ) 开始向10.10.10.2 NVE对等体发送流量时 :

### 9396-A# sh mac address-table dynamic

Legend:

\* - primary entry, G - Gateway MAC, (R) - Routed MAC, O - Overlay MAC  
age - seconds since last seen,+ - primary entry using vPC Peer-Link,  
(T) - True, (F) - False

VLAN	MAC Address	Type	age	Secure	NTFY	Ports
* 10	8c60.4f93.5ffc	dynamic	0	F	F	Pol12
+ 10	8c60.4f93.647c	dynamic	0	F	F	nve1(10.1.1.3)

### 9396-A# sh nve peers

Interface	Peer-IP	State	LearnType	Uptime	Router-Mac
nve1	10.1.1.3	Up	DP	00:00:14	n/a

### 9396-A# sh nve peers detail

Details of nve Peers:

```

-----
Peer-IP: 10.1.1.3
  NVE Interface      : nve1
  Peer State         : Up
  Peer Uptime        : 00:04:49

```

```
Router-Mac          : n/a
Peer First VNI     : 10010
Time since Create  : 00:04:49
Configured VNIs   : 10010
Provision State    : add-complete
Route-Update       : Yes
Peer Flags         : None
Learnt CP VNIs    : --
Peer-ifindex-resp : Yes
```

-----

**9396-A sh nve vni 10010 detail**

VNI: 10010

```
NVE-Interface      : nve1
Mcast-Addr         : 230.1.1.1
VNI State          : Up
Mode               : data-plane
VNI Type           : L2 [10]
VNI Flags          :
Provision State    : add-complete
Vlan-BD            : 10
SVI State          : n/a
```

9396-A# sh nve internal vni 10010

VNI 10010

```
Ready-State        : Ready [L2-vni-flood-learn-ready]
```

同样，在9396-C NVE对等体上必须启用：

**9396-C# show mac address-table dynamic**

Legend:

```
* - primary entry, G - Gateway MAC, (R) - Routed MAC, O - Overlay MAC
age - seconds since last seen,+ - primary entry using vPC Peer-Link,
(T) - True, (F) - False
```

VLAN	MAC Address	Type	age	Secure	NTFY	Ports
* 10	8c60.4f93.5ffc	dynamic	0	F	F	nve1(10.1.1.10)
* 10	8c60.4f93.647c	dynamic	0	F	F	Eth1/13

**9396-C# sh nve peers**

Interface	Peer-IP	State	LearnType	Uptime	Router-Mac
nve1	10.1.1.10	Up	DP	00:08:28	n/a

**9396-C# sh nve peers detail**

Details of nve Peers:

Peer-IP: 10.1.1.10

```
NVE Interface      : nve1
Peer State         : Up
Peer Uptime        : 00:08:32
Router-Mac         : n/a
Peer First VNI     : 10010
Time since Create  : 00:08:32
Configured VNIs   : 10010
Provision State   : add-complete
Route-Update       : Yes
Peer Flags         : None
Learnt CP VNIs    : --
Peer-ifindex-resp : Yes
```

-----

**9396-C sh nve vni 10010 detail**

```
VNI: 10010
NVE-Interface      : nve1
Mcast-Addr        : 230.1.1.1
VNI State          : Up
Mode               : data-plane
VNI Type           : L2 [10]
VNI Flags          :
Provision State    : add-complete
Vlan-BD           : 10
SVI State          : n/a
```

```
9396-C# sh nve internal vni 10010
```

```
VNI 10010
Ready-State       : Ready [L2-vni-flood-learn-ready]
```

如图所示，新的对等体基于数据平面学习，它使用泛洪和学习机制。如果MAC地址超时，则无对等体断开。

## 故障排除

目前没有针对此配置的故障排除信息。