# H3C SecPath M9000 系列 多业务安全网关

## VXLAN 配置指导(V7)

新华三技术有限公司
http://www.h3c.com

资料版本：6W303-20201210
产品版本：

| | |
|---|---|
| M9006/M9010/M9010-GM/M9014 | R9141 |
| M9008-S/M9008-S-6GW/M9012-S | R9712 |

# 前 言

本配置指导介绍了各软件特性的原理及其配置方法，包含原理简介、配置任务描述和配置举例。《VXLAN 配置指导》主要介绍 VXLAN 相关的特性。

前言部分包含如下内容：



## 读者对象

本手册主要适用于如下工程师：

- 网络规划人员
- 现场技术支持与维护人员
- 负责网络配置和维护的网络管理员

## 本书约定

### 1. 命令行格式约定

| 格　式 | 意　义 |
| --- | --- |
| **粗体** | 命令行关键字（命令中保持不变、必须照输的部分）采用**加粗**字体表示。 |
| *斜体* | 命令行参数（命令中必须由实际值进行替代的部分）采用*斜体*表示。 |
| [ ] | 表示用“[ ]”括起来的部分在命令配置时是可选的。 |
| { x \| y \| ... } | 表示从多个选项中仅选取一个。 |
| [ x \| y \| ... ] | 表示从多个选项中选取一个或者不选。 |
| { x \| y \| ... }* | 表示从多个选项中至少选取一个。 |
| [ x \| y \| ... ]* | 表示从多个选项中选取一个、多个或者不选。 |
| &<1-n> | 表示符号&前面的参数可以重复输入1～n次。 |
| # | 由“#”号开始的行表示为注释行。 |

### 2. 图形界面格式约定

| 格　式 | 意　义 |
| --- | --- |
| < > | 带尖括号“< >”表示按钮名，如“单击<确定>按钮”。 |
| [ ] | 带方括号“[ ]”表示窗口名、菜单名和数据表，如“弹出[新建用户]窗口”。 |
| / | 多级菜单用“/”隔开。如[文件/新建/文件夹]多级菜单表示[文件]菜单下的[新建]子菜单下的[文件夹]菜单项。 |

### 3. 各类标志

本书还采用各种醒目标志来表示在操作过程中应该特别注意的地方，这些标志的意义如下：

| | |
|---|---|
|  警告 | 该标志后的注释需给予格外关注，不当的操作可能会对人身造成伤害。 |
|  注意 | 提醒操作中应注意的事项，不当的操作可能会导致数据丢失或者设备损坏。 |
|  提示 | 为确保设备配置成功或者正常工作而需要特别关注的操作或信息。 |
|  说明 | 对操作内容的描述进行必要的补充和说明。 |
|  窍门 | 配置、操作、或使用设备的技巧、小窍门。 |

### 4. 图标约定

本书使用的图标及其含义如下：

| | |
|---|---|
|  | 该图标及其相关描述文字代表一般网络设备，如路由器、交换机、防火墙等。 |
|  | 该图标及其相关描述文字代表一般意义下的路由器，以及其他运行了路由协议的设备。 |
|  | 该图标及其相关描述文字代表二、三层以太网交换机，以及运行了二层协议的设备。 |
|  | 该图标及其相关描述文字代表无线控制器、无线控制器业务板和有线无线一体化交换机的无线控制引擎设备。 |
|  | 该图标及其相关描述文字代表无线接入点设备。 |
|  | 该图标及其相关描述文字代表无线终结单元。 |
|  | 该图标及其相关描述文字代表无线终结者。 |
|  | 该图标及其相关描述文字代表无线Mesh设备。 |
|  | 该图标代表发散的无线射频信号。 |
|  | 该图标代表点到点的无线射频信号。 |
|  | 该图标及其相关描述文字代表防火墙、UTM、多业务安全网关、负载均衡等安全设备。 |
|  | 该图标及其相关描述文字代表防火墙插卡、负载均衡插卡、NetStream插卡、SSL VPN插卡、IPS插卡、ACG插卡等安全插卡。 |

### 5. 示例约定

由于设备型号不同、配置不同、版本升级等原因，可能造成本手册中的内容与用户使用的设备显示信息不一致。实际使用中请以设备显示的内容为准。

本手册中出现的端口编号仅作示例，并不代表设备上实际具有此编号的端口，实际使用中请以设备上存在的端口编号为准。

## 资料意见反馈

如果您在使用过程中发现产品资料的任何问题，可以通过以下方式反馈：

E-mail：info@h3c.com

感谢您的反馈，让我们做得更好！

# 目 录

# 1 VXLAN 概述

VXLAN（Virtual eXtensible LAN，可扩展虚拟局域网络）是基于 IP 网络、采用“MAC in UDP”封装形式的二层 VPN 技术。VXLAN 可以基于已有的服务提供商或企业 IP 网络，为分散的物理站点提供二层互联，并能够为不同的租户提供业务隔离。VXLAN 主要应用于数据中心网络。

目前，设备只支持基于 IPv4 网络的 VXLAN 技术，不支持基于 IPv6 网络的 VXLAN 技术。

## 1.1 VXLAN的优点

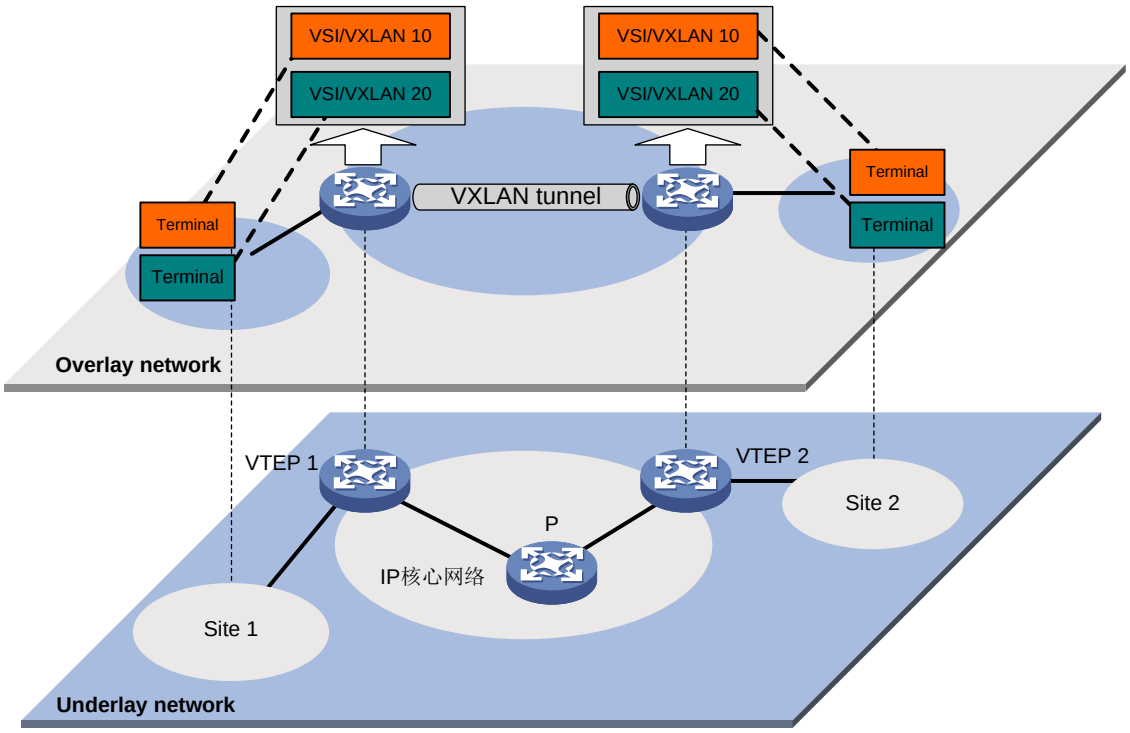
VXLAN 具有如下优点：

- 支持大量的租户：使用 24 位的标识符，最多可支持 2 的 24 次方（16777216）个 VXLAN，使支持的租户数目大规模增加，解决了传统二层网络 VLAN 资源不足的问题。
- 易于维护：基于 IP 网络组建大二层网络，使得网络部署和维护更加容易，并且可以充分地利用现有的 IP 网络技术，例如利用等价路由进行负载分担等；只有 IP 核心网络的边缘设备需要进行 VXLAN 处理，网络中间设备只需根据 IP 头转发报文，降低了网络部署的难度和费用。

## 1.2 VXLAN网络模型

VXLAN 技术将已有的三层物理网络作为 Underlay 网络，在其上构建出虚拟的二层网络，即 Overlay 网络。Overlay 网络通过封装技术、利用 Underlay 网络提供的三层转发路径，实现租户二层报文跨越三层网络在不同站点间传递。对于租户来说，Underlay 网络是透明的，同一租户的不同站点就像工作在一个局域网中。

图1-1 VXLAN 网络模型示意图




如图 1-1 所示，VXLAN 的典型网络模型中包括如下几部分：

- 用户终端（Terminal）：用户终端设备可以是 PC 机、无线终端设备、服务器上创建的 VM（Virtual Machine，虚拟机）等。不同的用户终端可以属于不同的 VXLAN。属于相同 VXLAN 的用户终端处于同一个逻辑二层网络，彼此之间二层互通；属于不同 VXLAN 的用户终端之间二层隔离。VXLAN 通过 VXLAN ID 来标识，VXLAN ID 又称 VNI（VXLAN Network Identifier，VXLAN 网络标识符），其长度为 24 比特。

  说明

  本文档中如无特殊说明，均以 VM 为例介绍 VXLAN 工作机制。采用其它类型用户终端时，VXLAN 工作机制与 VM 相同，不再赘述。

- VTEP（VXLAN Tunnel End Point，VXLAN 隧道端点）：VXLAN 的边缘设备。VXLAN 的相关处理都在 VTEP 上进行，例如识别以太网数据帧所属的 VXLAN、基于 VXLAN 对数据帧进行二层转发、封装/解封装报文等。
- VXLAN 隧道：两个 VTEP 之间的点到点逻辑隧道。VTEP 为数据帧封装 VXLAN 头、UDP 头和 IP 头后，通过 VXLAN 隧道将封装后的报文转发给远端 VTEP，远端 VTEP 对其进行解封装。
- 核心设备：IP 核心网络中的设备（如图 1-1 中的 P 设备）。核心设备不参与 VXLAN 处理，仅需要根据封装后报文的目的 IP 地址对报文进行三层转发。
- VSI（Virtual Switch Instance，虚拟交换实例）：VTEP 上为一个 VXLAN 提供二层交换服务的虚拟交换实例。VSI 可以看作是 VTEP 上的一台基于 VXLAN 进行二层转发的虚拟交换机，它

具有传统以太网交换机的所有功能，包括源 MAC 地址学习、MAC 地址老化、泛洪等。VSI 与 VXLAN 一一对应。

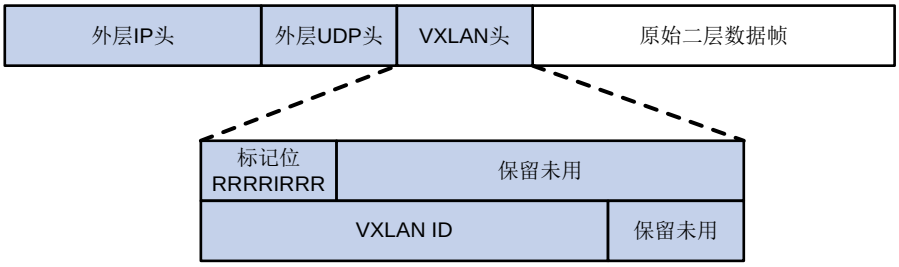
- AC（Attachment Circuit，接入电路）：VTEP 连接本地站点的物理电路或虚拟电路。在 VTEP 上，与 VSI 关联的三层接口或以太网服务实例（service instance）称为 AC。其中，以太网服务实例在二层以太网接口上创建，它定义了一系列匹配规则，用来匹配从该二层以太网接口上接收到的数据帧。

## 1.3 VXLAN报文封装格式

如图 1-2 所示，VXLAN 报文的封装格式为：在原始二层数据帧外添加 8 字节 VXLAN 头、8 字节 UDP 头和 20 字节 IP 头。其中，UDP 头的目的端口号为 VXLAN UDP 端口号（缺省为 4789）。VXLAN 头主要包括两部分：

- 标记位：“I”位为 1 时，表示 VXLAN 头中的 VXLAN ID 有效；为 0，表示 VXLAN ID 无效。其它位保留未用，设置为 0。
- VXLAN ID：用来标识一个 VXLAN 网络，长度为 24 比特。

图1-2 VXLAN 报文封装示意图




## 1.4 VXLAN运行机制

### 1.4.1 运行机制概述

VXLAN 运行机制可以概括为：

(1) 发现远端 VTEP，在 VTEP 之间建立 VXLAN 隧道，并将 VXLAN 隧道与 VXLAN 关联。

(2) 识别接收到的报文所属的 VXLAN，以便将报文的源 MAC 地址学习到 VXLAN 对应的 VSI，并在该 VSI 内转发该报文。

(3) 学习虚拟机的 MAC 地址。

(4) 根据学习到的 MAC 地址表项转发报文。

### 1.4.2 建立 VXLAN 隧道并将其与 VXLAN 关联

为了将 VXLAN 报文传递到远端 VTEP，需要创建 VXLAN 隧道，并将 VXLAN 隧道与 VXLAN 关联。

#### 1. 创建 VXLAN 隧道

通过手工方式创建 VXLAN 隧道，手工配置 Tunnel 接口，并指定隧道的源和目的 IP 地址分别为本端和远端 VTEP 的 IP 地址。

### 2. 关联 VXLAN 隧道与 VXLAN

通过手工方式将 VXLAN 隧道与 VXLAN 关联。

## 1.4.3 识别报文所属的 VXLAN

### 1. 本地站点内接收到数据帧的识别

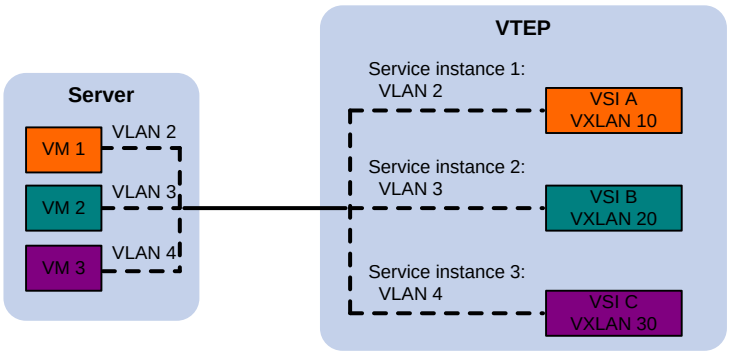
VTEP 采用如下几种方式在数据帧和 VXLAN 之间建立关联：

- 将三层接口与 VSI 关联：从该三层接口接收到的数据帧均属于指定的 VSI。VSI 内创建的 VXLAN 即为该数据帧所属的 VXLAN。
- 将以太网服务实例与 VSI 关联：以太网服务实例定义了一系列匹配规则，如匹配指定 VLAN 的报文、匹配接口接收到的所有报文等。从二层以太网接口上接收到的、与规则匹配的数据帧均属于指定的 VSI/VXLAN。

VTEP 从三层接口或以太网服务实例接收到数据帧后，根据关联方式判断报文所属的 VXLAN。

如图 1-3 所示，VM 1 属于 VLAN 2，在 VTEP 上配置以太网服务实例 1 匹配 VLAN 2 的报文，将以太网服务实例 1 与 VSI A 绑定，并在 VSI A 内创建 VXLAN 10，则 VTEP 接收到 VM 1 发送的数据帧后，可以判定该数据帧属于 VXLAN 10。

图1-3 二层数据帧所属 VXLAN 识别




### 2. VXLAN 隧道上接收报文的识别

对于从 VXLAN 隧道上接收到的 VXLAN 报文，VTEP 根据报文中携带的 VXLAN ID 判断该报文所属的 VXLAN。

## 1.4.4 学习 MAC 地址

MAC 地址学习分为本地 MAC 地址学习和远端 MAC 地址学习两部分：

- 本地 MAC 地址学习

  是指 VTEP 对本地站点内虚拟机 MAC 地址的学习。VTEP 接收到本地虚拟机发送的数据帧后，判断该数据帧所属的 VSI，并将数据帧中的源 MAC 地址（本地虚拟机的 MAC 地址）添加到该 VSI 的 MAC 地址表中，该 MAC 地址对应的接口为接收到数据帧的接口。

  VXLAN 不支持静态配置本地 MAC 地址。

- 远端 MAC 地址学习

  是指 VTEP 对远端站点内虚拟机 MAC 地址的学习。远端 MAC 地址的学习方式有如下几种：

- 静态配置：手工指定远端 MAC 地址所属的 VSI（VXLAN），及其对应的 VXLAN 隧道接口。
- 通过报文中的源 MAC 地址动态学习：VTEP 从 VXLAN 隧道上接收到远端 VTEP 发送的 VXLAN 报文后，根据 VXLAN ID 判断报文所属的 VXLAN，对报文进行解封装，还原二层数据帧，并将数据帧中的源 MAC 地址（远端虚拟机的 MAC 地址）添加到所属 VXLAN 对应 VSI 的 MAC 地址表中，该 MAC 地址对应的接口为 VXLAN 隧道接口。

通过不同方式学习到的远端 MAC 地址优先级由高到低依次为：

a. 静态配置的 MAC 地址优优先级最高。

b. 动态学习的 MAC 地址优先级最低。

## 1.4.5 转发单播流量

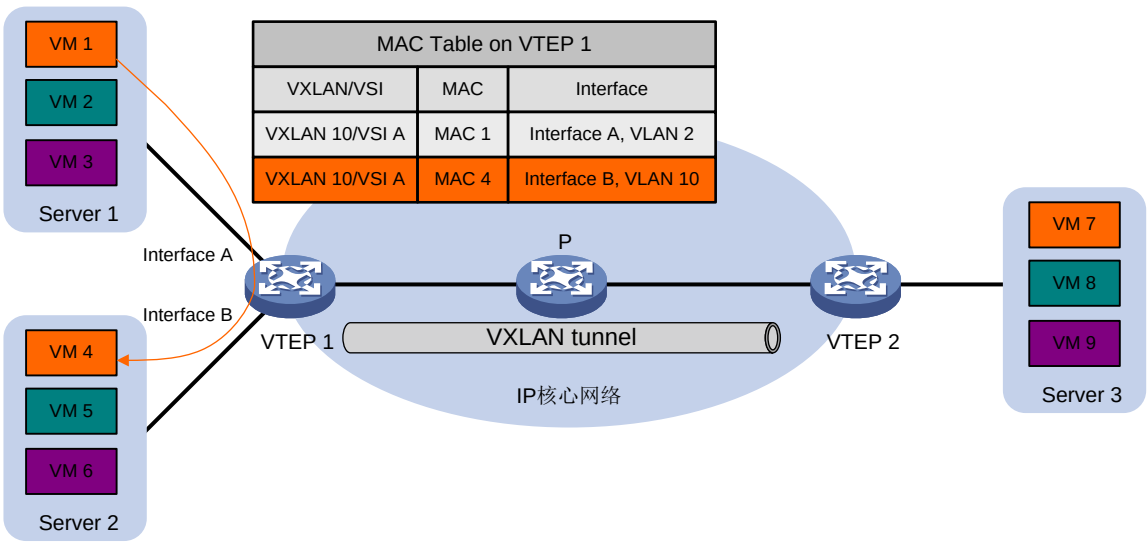
完成本地和远端 MAC 地址学习后，VTEP 在 VXLAN 内转发单播流量的过程如下所述。

### 1. 站点内流量

对于站点内流量，VTEP 判断出报文所属的 VSI 后，根据目的 MAC 地址查找该 VSI 的 MAC 地址表，从相应的本地接口转发给目的 VM。

如图 1-4 所示，VM 1（MAC 地址为 MAC 1）发送以太网帧到 VM 4（MAC 地址为 MAC 4）时，VTEP 1 从接口 Interface A 收到该以太网帧后，判断该数据帧属于 VSI A（VXLAN 10），查找 VSI A 的 MAC 地址表，得到 MAC 4 的出接口为 Interface B，所在 VLAN 为 VLAN 10，则将以太网帧从接口 Interface B 的 VLAN 10 内发送给 VM 4。

图1-4 站点内单播流量转发




### 2. 站点间流量

如图 1-5 所示，以 VM 1（MAC 地址为 MAC 1）发送以太网帧给 VM 7（MAC 地址为 MAC 7）为例，站点间单播流量的转发过程为：

(1) VM 1 发送以太网数据帧给 VM 7，数据帧的源 MAC 地址为 MAC 1，目的 MAC 为 MAC 7，VLAN ID 为 2。

(2) VTEP 1 从接口 Interface A（所在 VLAN 为 VLAN 2）收到该数据帧后，判断该数据帧属于 VSI A（VXLAN 10），查找 VSI A 的 MAC 地址表，得到 MAC 7 的出端口为 Tunnel1。

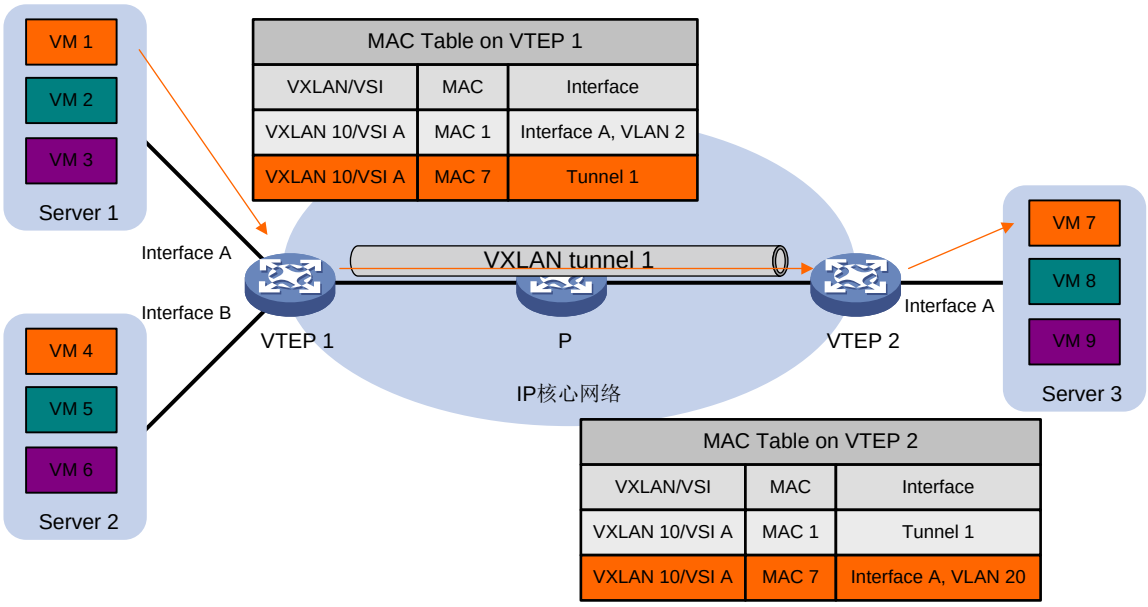
(3) VTEP 1 为数据帧封装 VXLAN 头、UDP 头和 IP 头后，将封装好的报文通过 VXLAN 隧道 Tunnel1、经由 P 设备发送给 VTEP 2。

(4) VTEP 2 接收到报文后，根据报文中的 VXLAN ID 判断该报文属于 VXLAN 10，并剥离 VXLAN 头、UDP 头和 IP 头，还原出原始的数据帧。

(5) VTEP 2 查找与 VXLAN 10 对应的 VSI A 的 MAC 地址表，得到 MAC 7 的出端口为 Interface A（所在 VLAN 为 VLAN 20）。

(6) VTEP 2 从接口 Interface A 的 VLAN 20 内将数据帧发送给 VM 7。

图1-5 站点间单播流量转发




## 1.4.6 转发泛洪流量

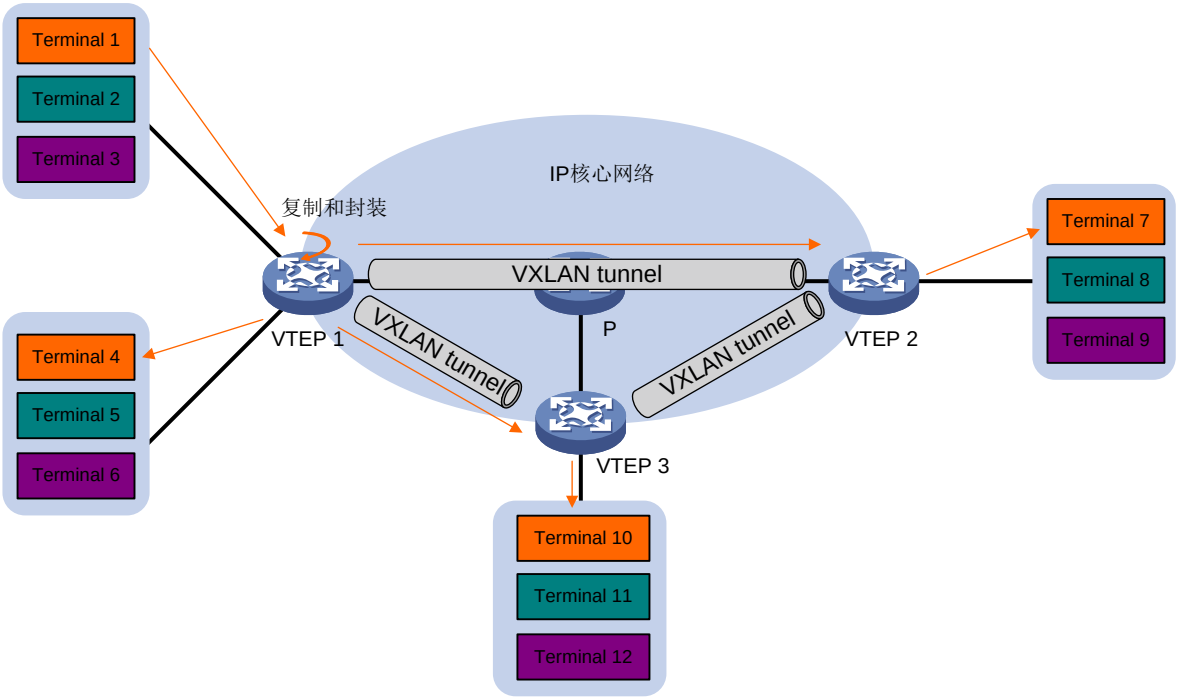
VTEP 从本地站点接收到泛洪流量（组播、广播和未知单播流量）后，将其转发给除接收接口外的所有本地接口和 VXLAN 隧道。为了避免环路，VTEP 从 VXLAN 隧道上接收到报文后，不会再将其泛洪到其它的 VXLAN 隧道，只会转发给所有本地接口。

根据复制方式的不同，流量泛洪方式分为单播路由方式（头端复制）、组播路由方式（核心复制）和泛洪代理方式（服务器复制）。设备暂不支持组播路由方式（核心复制）。

### 1. 单播路由方式（头端复制）

如图 1-6 所示，VTEP 负责复制报文，采用单播方式将复制后的报文通过本地接口发送给本地站点，并通过 VXLAN 隧道发送给 VXLAN 内的所有远端 VTEP。

图1-6 单播路由方式转发示意图

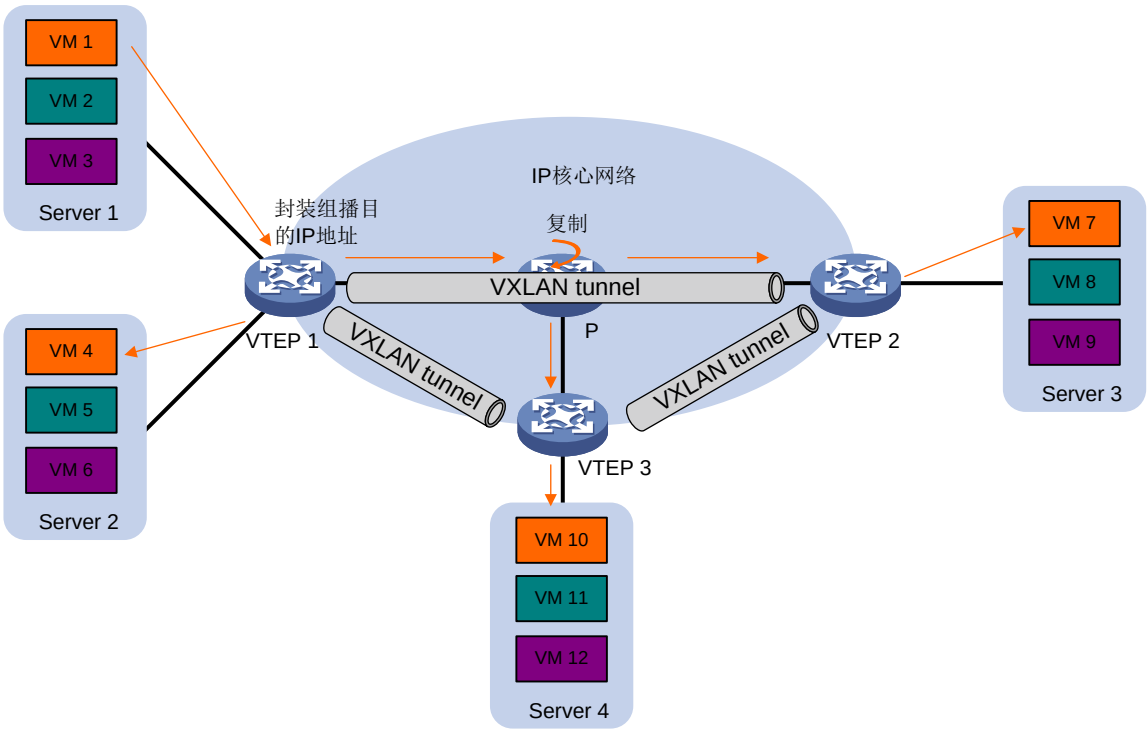



## 2. 组播路由方式（核心复制）

数据中心网络中需要通过 IP 核心网络进行二层互联的站点较多时，采用组播路由方式可以节省泛洪流量对核心网络带宽资源的占用。

如图 1-7 所示，在组播路由方式下，同一个 VXLAN 内的所有 VTEP 都加入同一个组播组，利用组播路由协议（如 PIM）在 IP 核心网上为该组播组建立组播转发表项。VTEP 接收到泛洪流量后，不仅在本地站点内泛洪，还会为其封装组播目的 IP 地址，封装后的报文根据已建立的组播转发表项转发到远端 VTEP。

图1-7 组播路由方式转发示意图

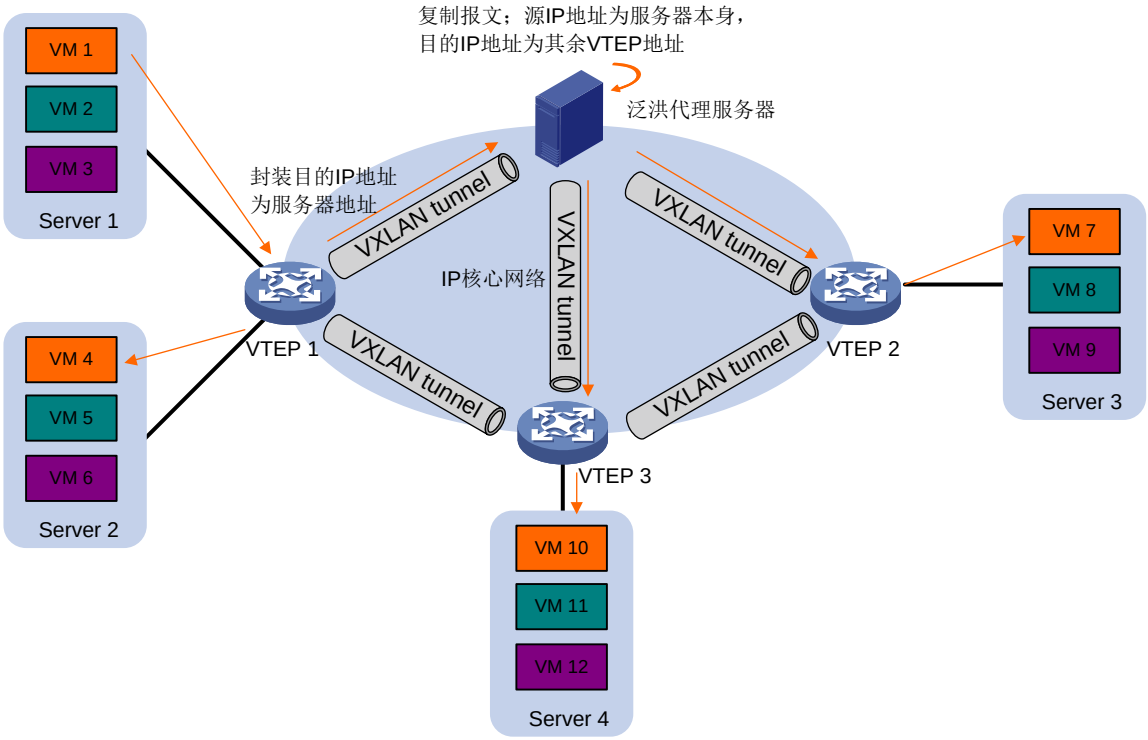



### 3. 泛洪代理方式（服务器复制）

数据中心网络中需要通过 IP 核心网络进行二层互联的站点较多时，采用泛洪代理方式可以在没有组播协议参与的情况下，节省泛洪流量对核心网络带宽资源的占用。

如图 1-8 所示，在泛洪代理方式下，同一个 VXLAN 内的所有 VTEP 都通过手工方式与代理服务器建立隧道。VTEP 接收到泛洪流量后，不仅在本地站点内泛洪，还会将其发送到代理服务器，由代理服务器转发到其它远端 VTEP。

图1-8 泛洪代理方式转发示意图




目前泛洪代理方式主要用于 SDN 网络，使用虚拟服务器作为泛洪代理服务器。

## 1.4.7 接入模式

接入模式分为 VLAN 接入模式和 Ethernet 接入模式两种。

### 1. VLAN 接入模式

在该模式下，从本地站点接收到的和发送给本地站点的以太网帧必须带有 VLAN Tag。

- VTEP 从本地站点接收到以太网帧后，删除该帧的所有 VLAN Tag，再转发该数据帧；
- VTEP 发送以太网帧到本地站点时，为其添加本地站点的 VLAN Tag。

采用该模式时，VTEP 不会传递 VLAN Tag 信息，不同站点可以独立地规划自己的 VLAN，不同站点的不同 VLAN 之间可以互通。

### 2. Ethernet 接入模式

在该模式下，从本地站点接收到的和发送给本地站点的以太网帧可以携带 VLAN Tag，也可以不携带 VLAN Tag。

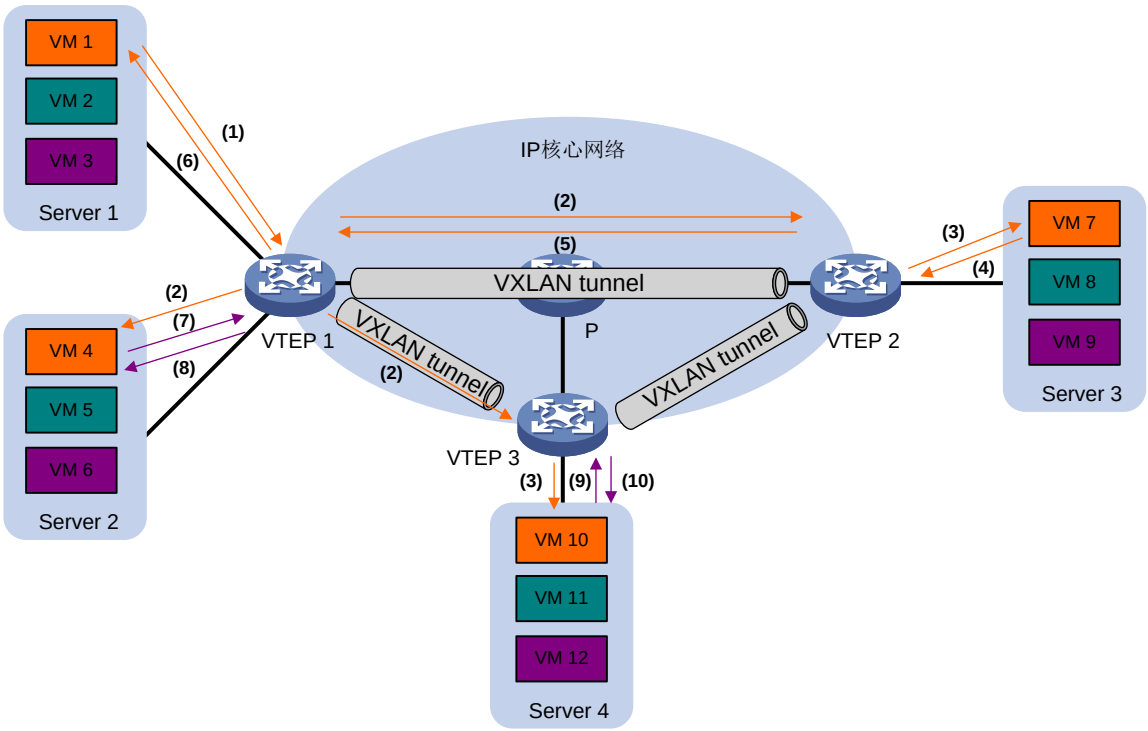
- VTEP 从本地站点接收到以太网帧后，保持该帧的 VLAN Tag 信息不变，转发该数据帧；
- VTEP 发送以太网帧到本地站点时，不会为其添加 VLAN Tag。

采用该模式时，VTEP 会在不同站点间传递 VLAN Tag 信息，不同站点的 VLAN 需要统一规划，否则无法互通。

## 1.5 ARP泛洪抑制

为了避免广播发送的 ARP 请求报文占用核心网络带宽，VTEP 从本地站点或 VXLAN 隧道接收到 ARP 请求和 ARP 应答报文后，根据该报文在本地建立 ARP 泛洪抑制表项。后续当 VTEP 收到本站点内虚拟机请求其它虚拟机 MAC 地址的 ARP 请求时，优先根据 ARP 泛洪抑制表项进行代答。如果没有对应的表项，则将 ARP 请求泛洪到核心网。ARP 泛洪抑制功能可以大大减少 ARP 泛洪的次数。

图1-9 ARP 泛洪抑制示意图




如图 1-9 所示，ARP 泛洪抑制的处理过程如下：

(1) 虚拟机 VM 1 发送 ARP 请求，获取 VM 7 的 MAC 地址。

(2) VTEP 1 根据接收到的 ARP 请求，建立 VM 1 的 ARP 泛洪抑制表项，并在 VXLAN 内泛洪该 ARP 请求（图 1-9 以单播路由泛洪方式为例）。

(3) 远端 VTEP（VTEP 2 和 VTEP 3）解封装 VXLAN 报文，获取原始的 ARP 请求报文后，建立 VM 1 的 ARP 泛洪抑制表项，并在本地站点的指定 VXLAN 内泛洪该 ARP 请求。

(4) VM 7 接收到 ARP 请求后，回复 ARP 应答报文。

(5) VTEP 2 接收到 ARP 应答后，建立 VM 7 的 ARP 泛洪抑制表项，并通过 VXLAN 隧道将 ARP 应答发送给 VTEP 1。

(6) VTEP 1 解封装 VXLAN 报文，获取原始的 ARP 应答，并根据该应答建立 VM 7 的 ARP 泛洪抑制表项，之后将 ARP 应答报文发送给 VM 1。

(7) 在 VTEP 1 上建立 ARP 泛洪抑制表项后，虚拟机 VM 4 发送 ARP 请求，获取 VM 1 或 VM 7 的 MAC 地址。

(8) VTEP 1 接收到 ARP 请求后，建立 VM 4 的 ARP 泛洪抑制表项，并查找本地 ARP 泛洪抑制表项，根据已有的表项回复 ARP 应答报文，不会对 ARP 请求进行泛洪。

(9) 在 VTEP 3 上建立 ARP 泛洪抑制表项后，虚拟机 VM 10 发送 ARP 请求，获取 VM 1 的 MAC 地址。

(10) VTEP 3 接收到 ARP 请求后，建立 VM 10 的 ARP 泛洪抑制表项，并查找本地 ARP 泛洪抑制表项，根据已有的表项回复 ARP 应答报文，不会对 ARP 请求进行泛洪。

## 1.6 协议规范

与 VXLAN 相关的协议规范有：

- RFC 7348：Virtual eXtensible Local Area Network (VXLAN): A Framework for Overlaying Virtualized Layer 2 Networks over Layer 3 Networks

# 2 配置 VXLAN

## 2.1 VXLAN配置任务简介

VXLAN 组网中，需要在 VTEP 上进行如下配置：

(1) 创建 VSI 和 VXLAN

(2) 配置 VXLAN 隧道

(3) 手工关联 VXLAN 与 VXLAN 隧道

(4) 建立数据帧与 VSI 的关联

(5) （可选）配置 VXLAN 报文相关功能

- 配置 VXLAN 报文的目的 UDP 端口号
- 配置 VXLAN 报文检查功能

(6) （可选）开启 VXLAN 软件快速转发功能

## 2.2 VXLAN配置准备

在 VXLAN 组网中，IP 核心网络中的设备上需要配置路由协议，确保 VTEP 之间路由可达。

## 2.3 创建VSI和VXLAN

(1) 进入系统视图。

**system-view**

(2) 开启 L2VPN 功能。

**l2vpn enable**

缺省情况下，L2VPN 功能处于关闭状态。

(3) 创建 VSI，并进入 VSI 视图。

**vsi** *vsi-name*

(4) 开启 VSI。

**undo shutdown**

缺省情况下，VSI 处于开启状态。

(5) 创建 VXLAN，并进入 VXLAN 视图。

**vxlan** *vxlan-id*

在一个 VSI 下只能创建一个 VXLAN。

不同 VSI 下创建的 VXLAN，其 VXLAN ID 不能相同。

(6) （可选）配置 VSI 相关参数。

a. 退回 VSI 视图。

**quit**

b. 配置 VSI 的描述信息。

**description** *text*

缺省情况下，未配置 VSI 的描述信息。

c. 配置 VSI 的 MTU 值。

**mtu** *mtu*

缺省情况下，VSI 的 MTU 值为 1500 字节。

VSI 的 MTU 值是指从 AC 上接收且通过 VXLAN 隧道转发的用户报文的最大长度。VSI 内的其它报文不受该 MTU 值的限制。

d. 开启 VSI 的 MAC 地址学习功能。

**mac-learning enable**

缺省情况下，VSI 的 MAC 地址学习功能处于开启状态。

# 2.4 配置VXLAN隧道

## 2.4.1 手工创建 VXLAN 隧道

### 1. 功能简介

手工创建 VXLAN 隧道时，隧道的源端地址和目的端地址需要分别手工指定为本地和远端 VTEP 的接口地址。

### 2. 配置限制和指导

在同一台设备上，VXLAN 隧道模式的不同 Tunnel 接口建议不要同时配置完全相同的源端地址和目的端地址。

关于隧道的详细介绍及 Tunnel 接口下的更多配置命令，请参见“三层技术-IP 业务配置指导”中的“隧道”。

### 3. 配置步骤

(1) 进入系统视图。

**system-view**

(2) （可选）配置 VXLAN 隧道的全局源地址。

**tunnel global source-address** *ip-address*

缺省情况下，未配置 VXLAN 隧道的全局源地址。

如果隧道下未配置源地址或源接口，则隧道会使用全局源地址作为隧道的源地址。

(3) 创建模式为 VXLAN 隧道的 Tunnel 接口，并进入 Tunnel 接口视图。

**interface tunnel** *tunnel-number* **mode vxlan**

在隧道的两端应配置相同的隧道模式，否则会造成报文传输失败。

(4) 配置隧道的源端地址。请选择其中一项进行配置。

- 直接指定隧道的源端地址。

  **source** *ipv4-address*

  指定的地址将作为封装后 VXLAN 报文的源 IP 地址。

- 指定隧道的源接口。

  **source** *interface-type interface-number*

指定接口的主 IP 地址将作为封装后 VXLAN 报文的源 IP 地址。

缺省情况下，未设置 VXLAN 隧道的源端地址。

(5) 配置隧道的目的端地址。

**destination** *ipv4-address*

缺省情况下，未指定隧道的目的端地址。

隧道的目的端地址是对端设备上接口的 IP 地址，该地址将作为封装后 VXLAN 报文的目的地址。

## 2.5 手工关联VXLAN与VXLAN隧道

### 1. 功能简介

一个 VXLAN 可以关联多条 VXLAN 隧道。一条 VXLAN 隧道可以关联多个 VXLAN，这些 VXLAN 共用该VXLAN隧道，VTEP根据VXLAN报文中的VXLAN ID来识别隧道传递的报文所属的VXLAN。VTEP 接收到某个 VXLAN 的泛洪流量后，如果采用单播路由泛洪方式，则 VTEP 将在与该 VXLAN 关联的所有 VXLAN 隧道上发送该流量，以便将流量转发给所有的远端 VTEP；如果采用泛洪代理方式，则 VTEP 通过与该 VXLAN 关联、通过 **flooding-proxy** 参数开启了泛洪代理功能的 VXLAN 隧道将泛洪流量发送给泛洪代理服务器。

### 2. 配置限制和指导

VTEP 必须与相同 VXLAN 内的其它 VTEP 建立 VXLAN 隧道，并将该隧道与 VXLAN 关联。

### 3. 配置步骤

(1) 进入系统视图。

**system-view**

(2) 进入 VSI 视图。

**vsi** *vsi-name*

(3) 进入 VXLAN 视图。

**vxlan** *vxlan-id*

(4) 配置 VXLAN 与 VXLAN 隧道关联。

**tunnel** *tunnel-number* [ **flooding-proxy** ]

缺省情况下，VXLAN 未关联 VXLAN 隧道。

| 参数 | 说明 |
| --- | --- |
| **flooding-proxy** | 如果指定了本参数，则VXLAN内的广播、组播和未知单播流量将通过该隧道发送到泛洪代理服务器，由代理服务器进行复制并转发到其它远端VTEP |

## 2.6 建立数据帧与VSI的关联

### 2.6.1 配置三层接口与 VSI 关联

#### 1. 功能简介

将三层接口与 VSI 关联后，从该接口接收到的报文，将通过查找关联 VSI 的 MAC 地址表进行转发。

#### 2. 配置步骤

(1) 进入系统视图。

**system-view**

(2) 进入三层接口视图。

**interface** *interface-type interface-number*

(3) 将三层接口与 VSI 关联。

**xconnect vsi** *vsi-name* [ **track** *track-entry-number*&<1-3> ]

缺省情况下，三层接口未关联 VSI。

### 2.6.2 配置以太网服务实例与 VSI 关联

#### 1. 功能简介

手工创建以太网服务实例，并将以太网服务实例与 VSI 关联后，从该接口接收到的、符合以太网服务实例报文匹配规则的报文，将通过查找关联 VSI 的 MAC 地址表进行转发。以太网服务实例提供了多种报文匹配规则（包括接口接收到的所有报文、所有携带 VLAN Tag 的报文和所有不携带 VLAN Tag 的报文等），为报文关联 VSI 提供了更加灵活的方式。

#### 2. 配置限制和指导

不能通过重复执行本命令修改报文匹配规则。如需修改报文匹配规则，请先通过 **undo encapsulation** 命令删除报文匹配规则，再执行 **encapsulation** 命令。

删除以太网服务实例下的报文匹配规则后，会自动取消以太网服务实例与 VSI 的关联。

#### 3. 配置步骤

(1) 进入系统视图。

**system-view**

(2) 进入接口视图。

- 进入二层以太网接口视图。

  **interface** *interface-type interface-number*

- 进入二层聚合接口视图。

  **interface bridge-aggregation** *interface-number*

(3) 手工创建以太网服务实例，并进入以太网服务实例视图。

**service-instance** *instance-id*

(4) 配置以太网服务实例的报文匹配规则。请选择其中一项进行配置。

- 匹配报文的外层 VLAN tag。

  **encapsulation s-vid** *vlan-id-list* [ **only-tagged** ]

- 匹配携带 VLAN tag 或不携带 VLAN tag 的所有报文。

  **encapsulation { tagged | untagged }**
- 匹配未匹配到接口上其它以太网服务实例的所有报文。

  **encapsulation default**

  同一个接口上最多只能有一个服务实例的报文匹配规则为 **encapsulation default**。

  如果接口上只存在一个配置了 **encapsulation default** 规则的以太网服务实例，则该接口上的所有报文都匹配该以太网服务实例。

缺省情况下，未配置报文匹配规则。

(5) 将以太网服务实例与 VSI 关联。

**xconnect vsi** *vsi-name* **[ access-mode { ethernet | vlan } ] [ track** *track-entry-number*&<1-3> **]**

缺省情况下，以太网服务实例未关联 VSI。

## 2.7 配置VXLAN报文的目的UDP端口号

(1) 进入系统视图。

**system-view**

(2) 配置 VXLAN 报文的目的 UDP 端口号。

**vxlan udp-port** *port-number*

缺省情况下，VXLAN 报文的目的 UDP 端口号为 4789。

属于同一个 VXLAN 的 VTEP 设备上需要配置相同的 UDP 端口号。

## 2.8 配置VXLAN报文检查功能

### 1. 功能简介

通过本配置可以实现对接收到的 VXLAN 报文的 UDP 校验和、内层封装的以太网数据帧是否携带 VLAN Tag 进行检查：

- UDP 校验和检查：VTEP 接收到 VXLAN 报文后，检查该报文的 UDP 校验和是否为 0。若 UDP 校验和为 0，则接收该报文；若 UDP 校验和不为 0，则检查 UDP 检验和是否正确，正确则接收该报文；否则，丢弃该报文。
- VLAN Tag 检查：VTEP 接收到 VXLAN 报文并对其解封装后，若内层以太网数据帧带有 VLAN Tag，则丢弃该 VXLAN 报文。

### 2. 配置限制和指导

远端 VTEP 上通过 **xconnect vsi** 命令的 **access-mode** 参数配置接入模式为 **ethernet** 时，VXLAN 报文可能携带 VLAN Tag。这种情况下建议不要在本端 VTEP 上执行 **vxlan invalid-vlan-tag discard** 命令，以免错误地丢弃报文。

### 3. 配置步骤

(1) 进入系统视图。

**system-view**

(2) 配置丢弃 UDP 校验和检查失败的 VXLAN 报文。

**vxlan invalid-udp-checksum discard**

缺省情况下，不会检查 VXLAN 报文的 UDP 校验和。

(3) 配置丢弃内层数据帧含有 VLAN Tag 的 VXLAN 报文。

**vxlan invalid-vlan-tag discard**

缺省情况下，不会检查 VXLAN 报文内层封装的以太网数据帧是否携带 VLAN Tag。

## 2.9 开启VXLAN软件快速转发功能

### 1. 功能简介

开启本功能后，数据报文通过 VXLAN 隧道进行软件转发时，不会进行 QoS、安全等业务处理，直接进行转发，以提高处理性能。建议仅在 VSI 虚接口和 VXLAN 隧道对应的报文出接口上没有配置 QoS、安全等业务，且需要加快 VXLAN 软件转发速度的场景下，开启本功能。

### 2. 配置限制和指导

开启本功能后，如果到达 VXLAN 隧道目的端地址存在多条等价路由，只会从中选择一条路由转发 VXLAN 报文，不能在多条路由之间进行负载分担。

### 3. 配置步骤

(1) 进入系统视图。

**system-view**

(2) 开启 VXLAN 软件快速转发功能。

**vxlan fast-forwarding enable**

缺省情况下，VXLAN 软件快速转发功能处于关闭状态。

## 2.10 VXLAN显示和维护

在完成上述配置后，在任意视图下执行 **display** 命令可以显示配置后 VXLAN 的运行情况，通过查看显示信息验证配置的效果。

在用户视图下，用户可以执行 **reset** 命令来清除 VXLAN 的相关信息。

表2-1 VXLAN 显示和维护

| 操作 | 命令 |
|---|---|
| 显示Tunnel接口信息 | **display interface** [ **tunnel** [ *number* ] ] [ **brief** [ **description** \| **down** ] ] |
| 显示与VSI关联的三层接口的L2VPN信息 | **display l2vpn interface** [ **vsi** *vsi-name* \| *interface-type interface-number* ] [ **verbose** ] |
| 显示VSI的MAC地址表信息 | **display l2vpn mac-address** [ **vsi** *vsi-name* ] [ **dynamic** ] [ **count** ] |
| 显示以太网服务实例的信息 | **display l2vpn service-instance** [ **interface** *interface-type interface-number* [ **service-instance** *instance-id* ] ] [ **verbose** ] |
| 显示VSI的信息 | **display l2vpn vsi** [ **name** *vsi-name* ] [ **verbose** ] |
| 显示VXLAN关联的VXLAN隧道信息 | **display vxlan tunnel** [ **vxlan-id** *vxlan-id* ] |

| 操作 | 命令 |
| --- | --- |
| 清除VSI动态学习的MAC地址表项 | **reset l2vpn mac-address** [ **vsi** *vsi-name* ] |

说明

**display interface tunnel** 命令的详细介绍，请参见“三层技术-IP 业务命令参考”中的“隧道”。