H3C SecPath 防火墙产品

VXLAN 配置指导(V7)

新华三技术有限公司
http://www.h3c.com

资料版本：6W403-20210107

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（Virtual eXtensible LAN，可扩展虚拟局域网络）工作原理及相关配置。

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

格　式	意　义
	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与硬件适配关系

本特性的支持情况与设备型号有关，请以设备的实际情况为准。

型号	说明
F5010、F5020、F5020-GM、F5030、F5030-6GW、F5040、F5060、F5080、F5000-AI-20、F5000-AI-40、F5000-V30、F5000-C、F5000-S、F5000-M、F5000-A	支持
F1000-AI-20、F1000-AI-30、F1000-AI-50、F1000-AI-60、F1000-AI-70、F1000-AI-80、F1000-AI-90	不支持
F1003-L、F1005-L、F1010-L	不支持
F1005、F1010	不支持
F1020、F1020-GM、F1030、F1030-GM、F1050、F1060、F1070、F1070-GM、F1070-GM-L、F1080、F1090、F1000-V70	不支持
F1000-AK1110、F1000-AK1120、F1000-AK1130、F1000-AK1140	不支持
F1000-AK1212、F1000-AK1222、F1000-AK1232、F1000-AK1312、F1000-AK1322、F1000-AK1332	不支持
F1000-AK1414、F1000-AK1424、F1000-AK1434、F1000-AK1514、F1000-AK1524、F1000-AK1534、F1000-AK1614	不支持
F1000-AK108、F1000-AK109、F1000-AK110、F1000-AK115、F1000-AK120、F1000-AK125、F1000-AK710	不支持
F1000-AK130、F1000-AK135、F1000-AK140、F1000-AK145、F1000-AK150、F1000-AK155、F1000-AK160、F1000-AK165、F1000-AK170、F1000-AK175、F1000-AK180、F1000-AK185、F1000-GM-AK370、F1000-GM-AK380、F1000-AK711	不支持
LSU3FWCEA0、LSUM1FWCEAB0、LSX1FWCEA1	不支持
LSXM1FWDF1、LSUM1FWDEC0、IM-NGFWX-IV、LSQM1FWDSC0、LSWM1FWD0、LSPM6FWD、LSQM2FWDSC0	不支持
vFW1000、vFW2000	支持

1.2 VXLAN的优点

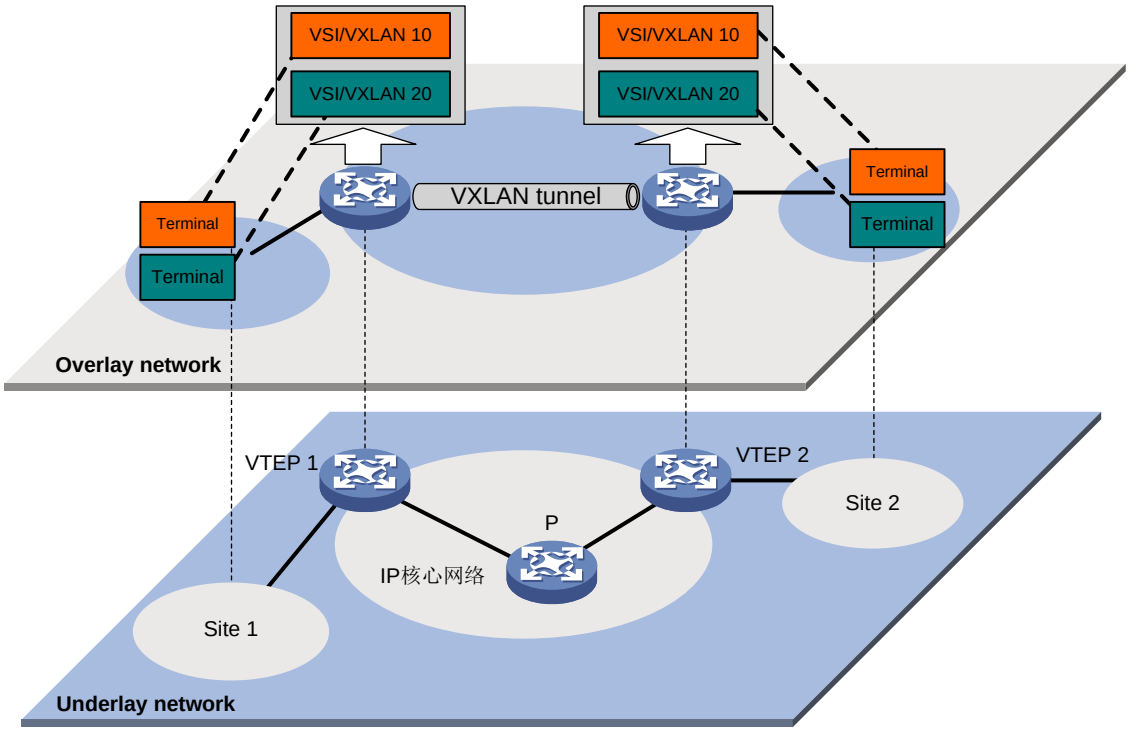
VXLAN 具有如下优点：

- 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。

1.3 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。采用其他类型用户终端时，VXLAN 工作机制与 VM 相同，不再赘述。

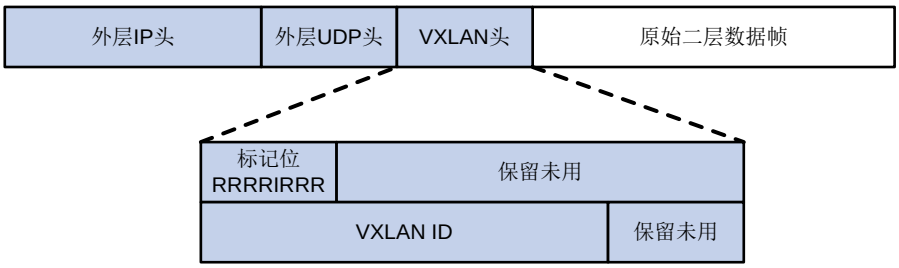
- VTEP（VXLAN Tunnel End Point，VXLAN 隧道端点）：VXLAN 的边缘设备。VXLAN 的相关处理都在 VTEP 上进行，例如识别以太网数据帧所属的 VXLAN、基于 VXLAN 对数据帧进行二层转发、封装/解封装报文等。
- VXLAN 隧道：两个 VTEP 之间的点到点逻辑隧道。VTEP 为数据帧封装 VXLAN 头、UDP 头和 IP 头后，通过 VXLAN 隧道将封装后的报文转发给远端 VTEP，远端 VTEP 对其进行解封装。
- 核心设备：IP 核心网络中的设备（如图 1-1 中的 P 设备）。核心设备不参与 VXLAN 处理，仅需要根据封装后报文的目的 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 关联的三层接口称为 AC。

1.4 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。其他位保留未用，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



1.5 VXLAN运行机制

1.5.1 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。

1.5.2 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。

1.5.3 识别报文所属的 VXLAN

1. 本地站点内接收到数据帧的识别

VTEP 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。

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。

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。

1.5.4 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。

1.5.5 转发单播流量

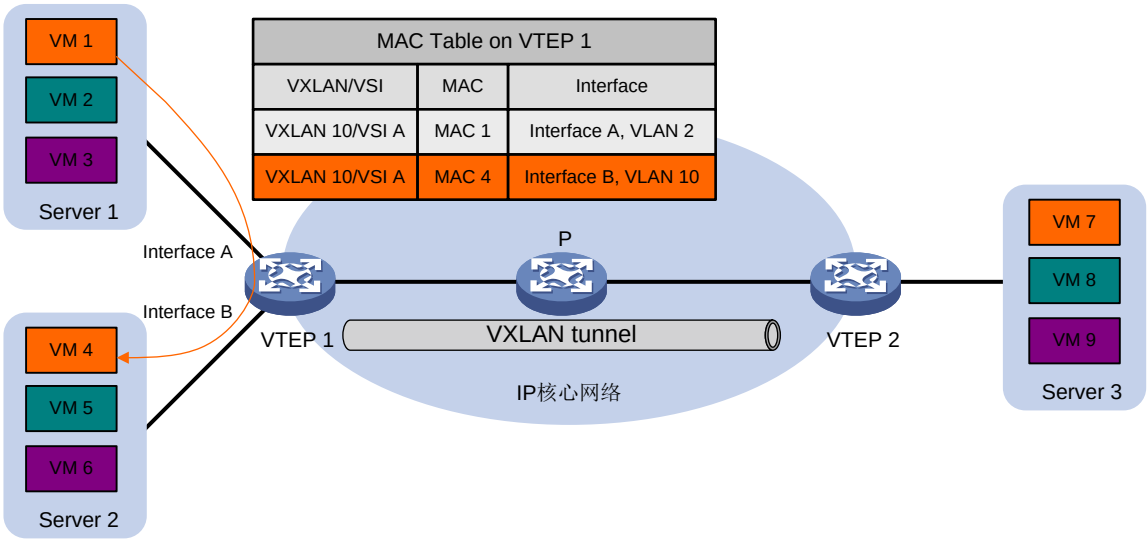
完成本地和远端 MAC 地址学习后，VTEP 在 VXLAN 内转发单播流量的过程如下所述。

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 1-3 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1-3 站点内单播流量转发



2. 站点间流量

如图 1-4 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。

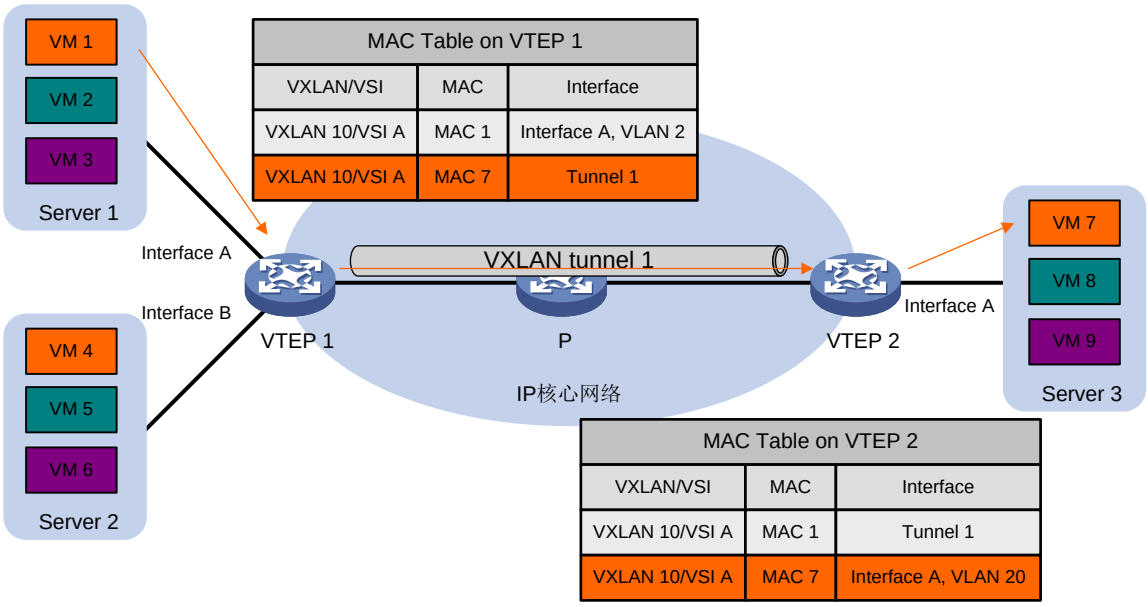
(3) VTEP 1 为数据帧封装 VXLAN 头、UDP 头和 IP 头后，将封装好的报文通过 VXLAN 隧道 Tunnel1、经由 P 设备发送给 VTEP 2。

(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1-4 站点间单播流量转发



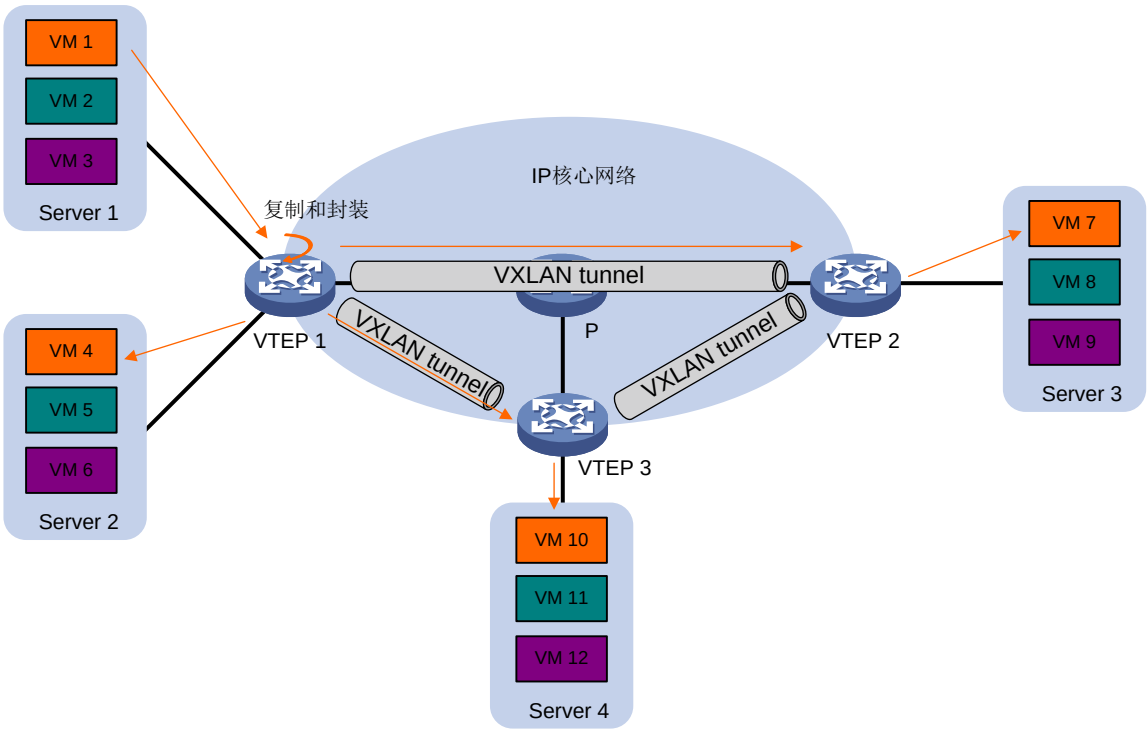
1.5.6 转发泛洪流量

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他的 VXLAN 隧道，只会转发给所有本地接口。

1. 单播路由方式（头端复制）

如图 1-5 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1-5 单播路由方式转发示意图

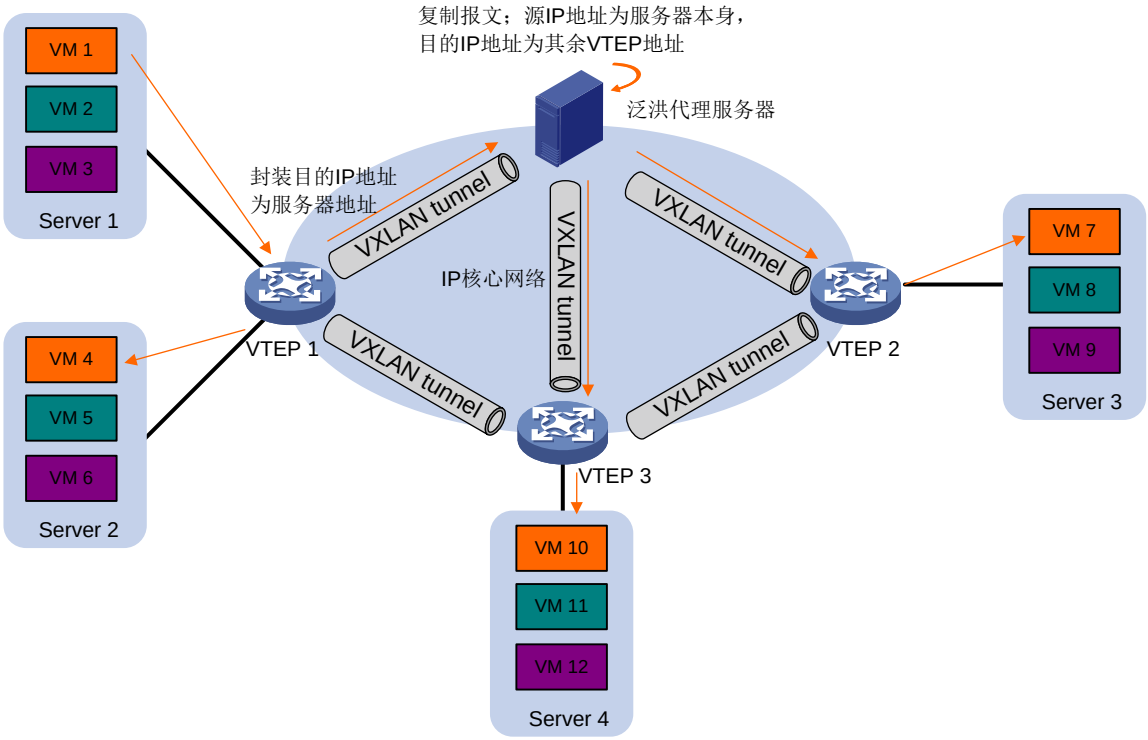


2. 泛洪代理方式（服务器复制）

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 1-6 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他远端 VTEP。

图1-6 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，使用服务器作为泛洪代理服务器。采用泛洪代理方式时，需要在 VTEP 上使用 **vxlan tunnel mac-learning disable** 命令关闭远端 MAC 地址自动学习功能，采用 SDN 控制器下发的 MAC 地址表项进行流量转发。

1.5.7 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。

1.6 VXLAN IP网关

VXLAN 可以为分散的物理站点提供二层互联。如果要为 VXLAN 站点内的虚拟机提供三层业务，则需要在网络中部署 VXLAN IP 网关，以便站点内的虚拟机通过 VXLAN IP 网关与外界网络或其他 VXLAN 网络内的虚拟机进行三层通信。

VXLAN IP 网关的详细介绍，请参见“3 VXLAN IP 网关”。

1.7 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）管理本地和远端 MAC 地址

- 添加静态远端 MAC 地址
- 关闭远端 MAC 地址自动学习功能
- 开启本地 MAC 地址的日志记录功能

(6) （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

(7) （可选）减少发送到核心网的泛洪流量

- 配置 VSI 泛洪抑制

(8) （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 内的其他报文不受该 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。

e. 配置允许 VSI 学习到的最大 MAC 地址数。

mac-table limit *mac-limit*

缺省情况下，不对 VSI 学习到的最大 MAC 地址数进行限制。

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“VPN 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。

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。

配置 VXLAN 与 VXLAN 隧道关联时，如果指定了 **no-split-horizon** 参数，则该 VXLAN 内不能存在去往同一个 VTEP 的其他 VXLAN 隧道。

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** | **no-split-horizon**] *

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他远端VTEP
no-split-horizon	如果指定了本参数，则VXLAN内从AC和VXLAN隧道接收到的报文均可以通过该隧道转发

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。

2.7 管理本地和远端MAC地址

2.7.1 功能简介

本地 MAC 地址只能动态学习，不能静态配置。在动态添加、删除本地 MAC 地址时，可以记录日志信息。

远端 MAC 地址表项的产生方法包括静态添加、根据接收到的 VXLAN 报文内封装的源 MAC 地址自动学习等。

2.7.2 添加静态远端 MAC 地址

(1) 进入系统视图。

system-view

(2) 添加静态远端 MAC 地址表项。

mac-address static *mac-address* **interface tunnel** *tunnel-number* **vsi** *vsi-name*

interface tunnel *interface-number* 参数指定的隧道接口必须与 **vsi** *vsi-name* 参数指定的 VSI 对应的 VXLAN 关联，否则配置将失败。

2.7.3 关闭远端 MAC 地址自动学习功能

1. 功能简介

如果网络中存在攻击，为了避免学习到错误的远端 MAC 地址，可以手工关闭远端 MAC 地址自动学习功能，手动添加静态的远端 MAC 地址。

2. 配置步骤

(1) 进入系统视图。

system-view

(2) 关闭远端 MAC 地址自动学习功能。

vxlan tunnel mac-learning disable

缺省情况下，远端 MAC 地址自动学习功能处于开启状态。

2.7.4 开启本地 MAC 地址的日志记录功能

1. 功能简介

开启本地 MAC 地址的日志记录功能后，VXLAN 会立即根据已经学习到的本地 MAC 地址表项生成日志信息，之后在增加或删除本地 MAC 地址时也将产生日志信息。生成的日志信息将被发送到设备的信息中心，通过设置信息中心的参数，决定日志信息的输出规则（即是否允许输出以及输出方向）。

2. 配置步骤

(1) 进入系统视图。

system-view

(2) 开启本地 MAC 地址的日志记录功能。

vxlan local-mac report

缺省情况下，本地 MAC 地址的日志记录功能处于关闭状态。

2.8 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。

2.9 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。

2.10 配置VSI泛洪抑制

1. 功能简介

缺省情况下，VTEP 从本地站点内接收到目的 MAC 地址未知的单播数据帧后，会在该 VXLAN 内除接收接口外的所有本地接口和 VXLAN 隧道上泛洪该数据帧，将该数据帧发送给 VXLAN 内的所有站点。如果用户希望把该类数据帧限制在本地站点内，不通过 VXLAN 隧道将其转发到远端站点，则可以通过本命令手工禁止 VXLAN 对应 VSI 的泛洪功能。

禁止泛洪功能后，为了将某些 MAC 地址的数据帧泛洪到远端站点以保证某些业务的流量在站点间互通，可以配置选择性泛洪的 MAC 地址，当数据帧的目的 MAC 地址匹配该 MAC 地址时，该数据帧可以泛洪到远端站点。

2. 配置步骤

(1) 进入系统视图。

system-view

(2) 进入 VSI 视图。

vsi *vsi-name*

(3) 关闭 VSI 的泛洪功能。

flooding disable

缺省情况下，VSI 泛洪功能处于开启状态。

(4) （可选）配置 VSI 选择性泛洪的 MAC 地址。

selective-flooding mac-address *mac-address*

2.11 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。

2.12 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* [**tunnel** *tunnel-number*]]
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“VPN 命令参考”中的“隧道”。

3 VXLAN IP 网关

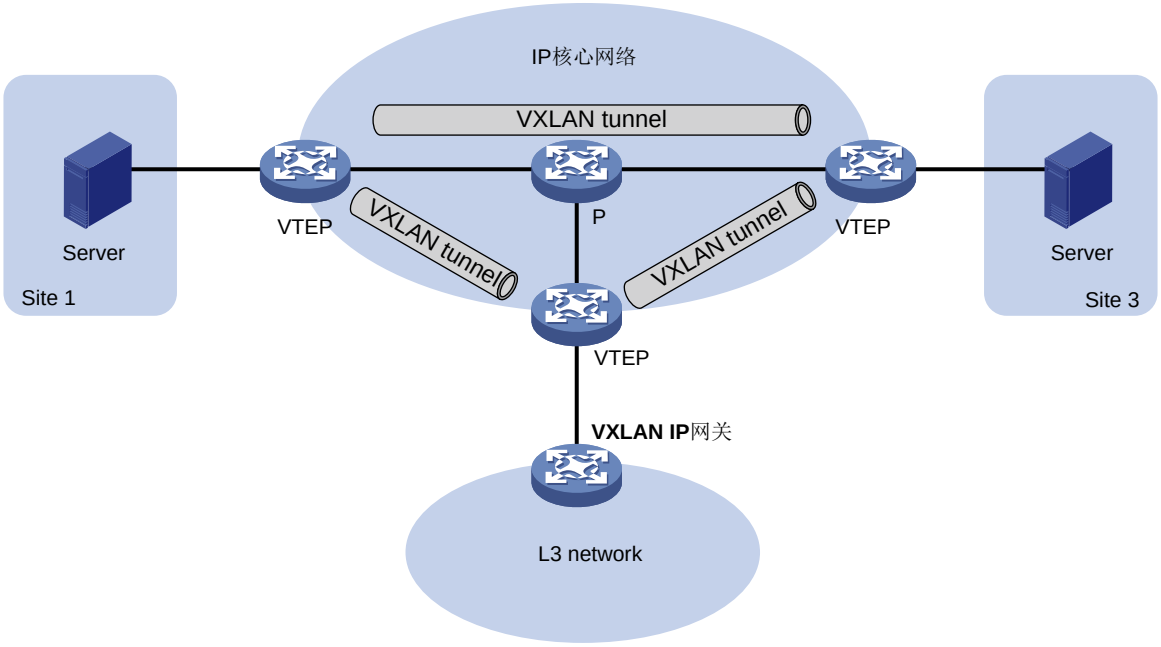
3.1 VXLAN IP网关简介

VXLAN 可以为分散的物理站点提供二层互联。如果要为 VXLAN 站点内的虚拟机提供三层业务，则需要在网络中部署 VXLAN IP 网关，以便站点内的虚拟机通过 VXLAN IP 网关与外界网络或其他 VXLAN 网络内的虚拟机进行三层通信。VXLAN IP 网关既可以部署在独立的物理设备上，也可以部署在 VTEP 设备上。VXLAN IP 网关部署在 VTEP 设备上时，又分为集中式 VXLAN IP 网关和分布式 VXLAN IP 网关两种方式。

3.1.1 独立的 VXLAN IP 网关

如图 3-1 所示，VXLAN IP 网关部署在独立的物理设备上时，VXLAN IP 网关作为物理站点接入 VTEP，VXLAN 业务对于网关设备透明。虚拟机通过 VXLAN IP 网关与三层网络中的节点通信时，虚拟机将三层报文封装成二层数据帧发送给 VXLAN IP 网关。VTEP 对该数据帧进行 VXLAN 封装，并在 IP 核心网络上将其转发给远端 VTEP（连接 VXLAN IP 网关的 VTEP）。远端 VTEP 对 VXLAN 报文进行解封装，并将原始的二层数据帧转发给 VXLAN IP 网关。VXLAN IP 网关去掉链路层封装后，对报文进行三层转发。

图3-1 独立的 VXLAN IP 网关示意图

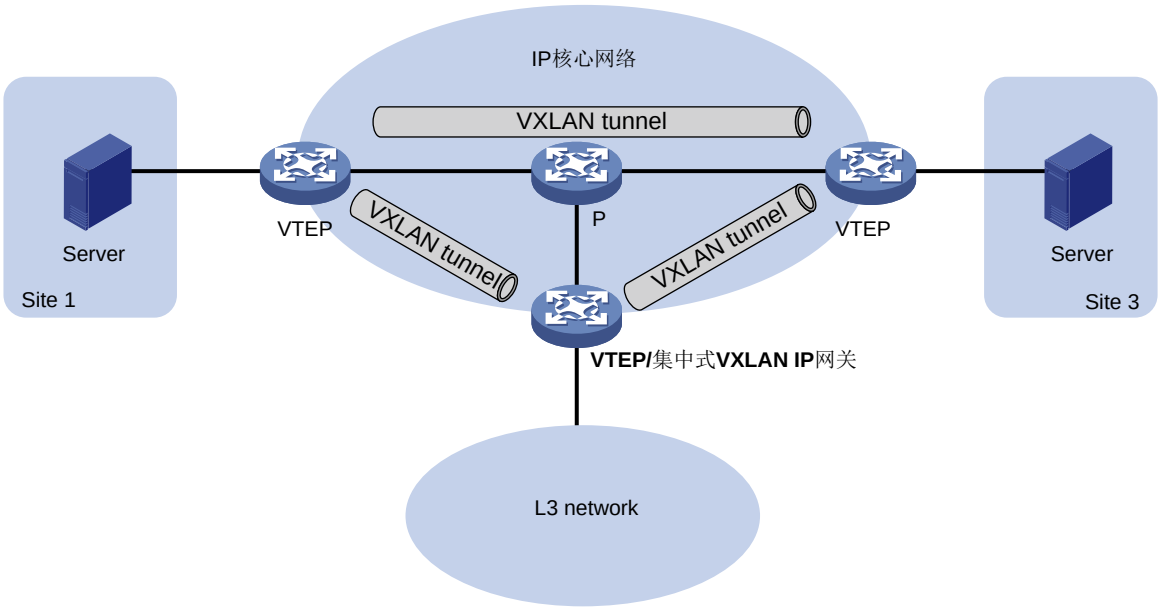


3.1.2 集中式 VXLAN IP 网关

如图 3-2 所示，集中式 VXLAN IP 网关进行二层 VXLAN 业务终结的同时，还对内层封装的 IP 报文进行三层转发处理。与独立的 VXLAN IP 网关相比，该方式除了能够节省设备资源外，VXLAN IP

网关功能由 VXLAN 对应的三层虚接口（VSI 虚接口）承担，三层业务的部署和控制也更加灵活和方便。

图3-2 集中式 VXLAN IP 网关示意图



如图 3-3 所示，以地址为 10.1.1.11 的虚拟机为例，虚拟机与外界网络进行三层通信的过程为：

(1) 虚拟机（10.1.1.11）跨网段进行三层通信时，先广播发送 ARP 请求消息，解析 VXLAN IP 网关（10.1.1.1）的 MAC 地址。

(2) VTEP 1 收到 ARP 请求消息后，添加 VXLAN 封装并发送给所有的远端 VTEP。

(3) VTEP 3 解封装 VXLAN 报文后，发现 ARP 请求的目的 IP 为 VXLAN 对应的本地网关 IP 地址，即与 VXLAN 关联的 VSI 虚接口的 IP 地址，则学习 10.1.1.11 的 ARP 信息，并向虚拟机回应 ARP 应答消息。

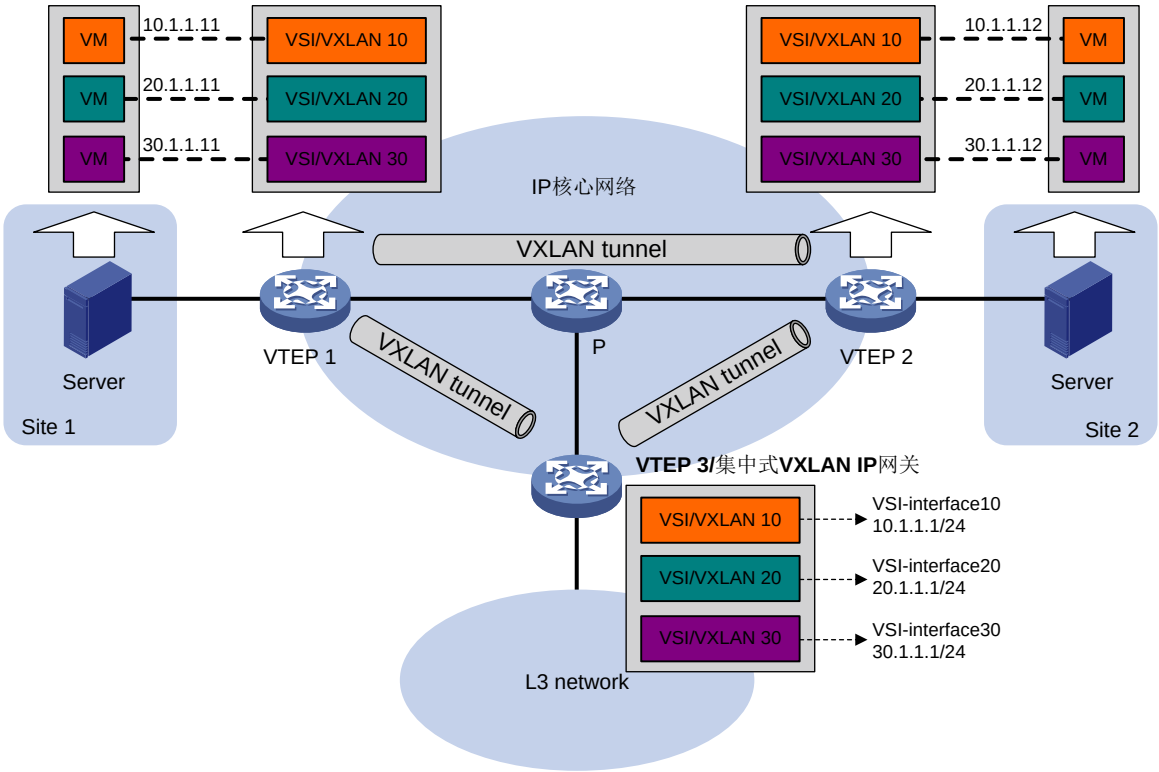
(4) VTEP 1 收到 ARP 应答消息后，将该消息转发给虚拟机。

(5) 虚拟机获取到网关的 MAC 地址后，为三层报文添加网关的 MAC 地址，通过 VXLAN 网络将二层数据帧发送给 VTEP 3。

(6) VTEP 3 解封装 VXLAN 报文，并去掉链路层头后，对内层封装的 IP 报文进行三层转发，将其发送给最终的目的节点。

(7) 目的节点回复的报文到达网关后，网关根据已经学习到的 ARP 表项，为报文封装链路层头，并通过 VXLAN 网络将其发送给虚拟机。

图3-3 集中式 VXLAN IP 网关的三层通信过程

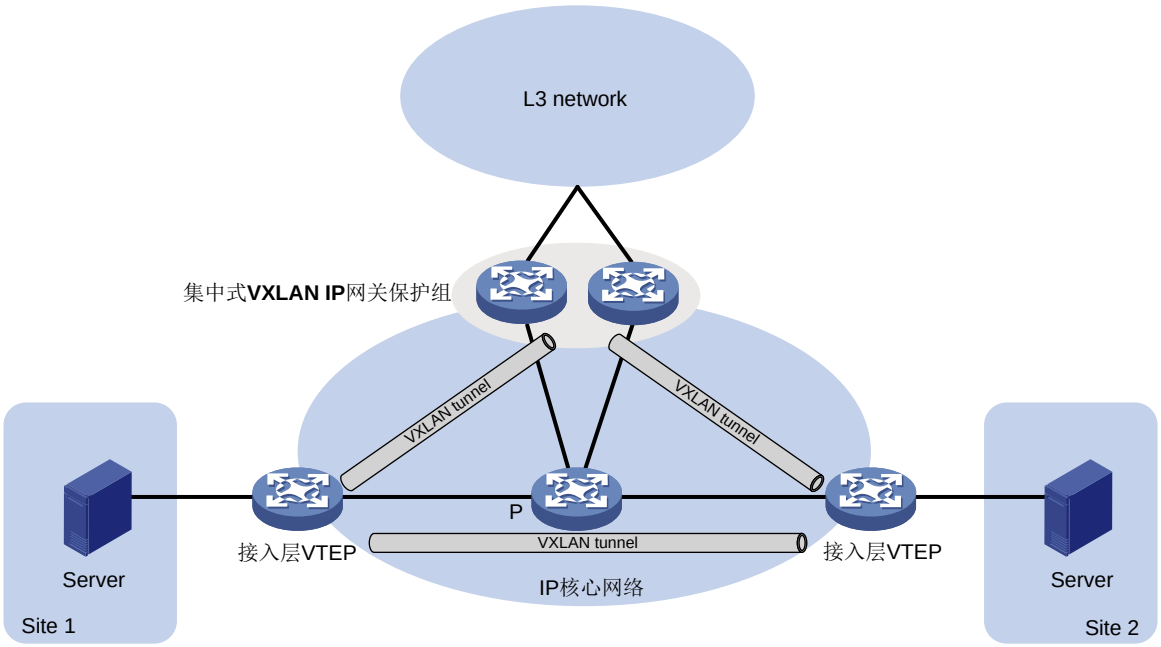


属于不同 VXLAN 网络的虚拟机之间的通信过程与上述过程类似，不同之处在于一个 VXLAN 网络的集中式网关需要将报文转发给另一个 VXLAN 网络的集中式网关，再由该集中式网关将报文转发给本 VXLAN 内对应的虚拟机。

3.1.3 集中式 VXLAN IP 网关保护组

由单台设备承担站点内大量虚拟机的集中式 VXLAN IP 网关功能，对设备的处理资源占用较高，并且对于网关的单点故障没有保护措施。通过集中式 VXLAN IP 网关保护组，可以实现多台设备同时承担网关功能，在提供单点故障保护机制的同时，还可以实现上下行流量的负载分担。

图3-4 集中式 VXLAN IP 网关保护组示意图



如图 3-4 所示，两台集中式 VXLAN IP 网关形成保护组，两台设备上存在相同的 VTEP IP，称为保护组的 VTEP IP。接入层 VTEP 与保护组的 VTEP IP 建立 VXLAN 隧道，将虚拟机发送至其它网络的报文转发至保护组，保护组中的两台网关设备均可以接收并处理虚拟机发往其它网络的流量。

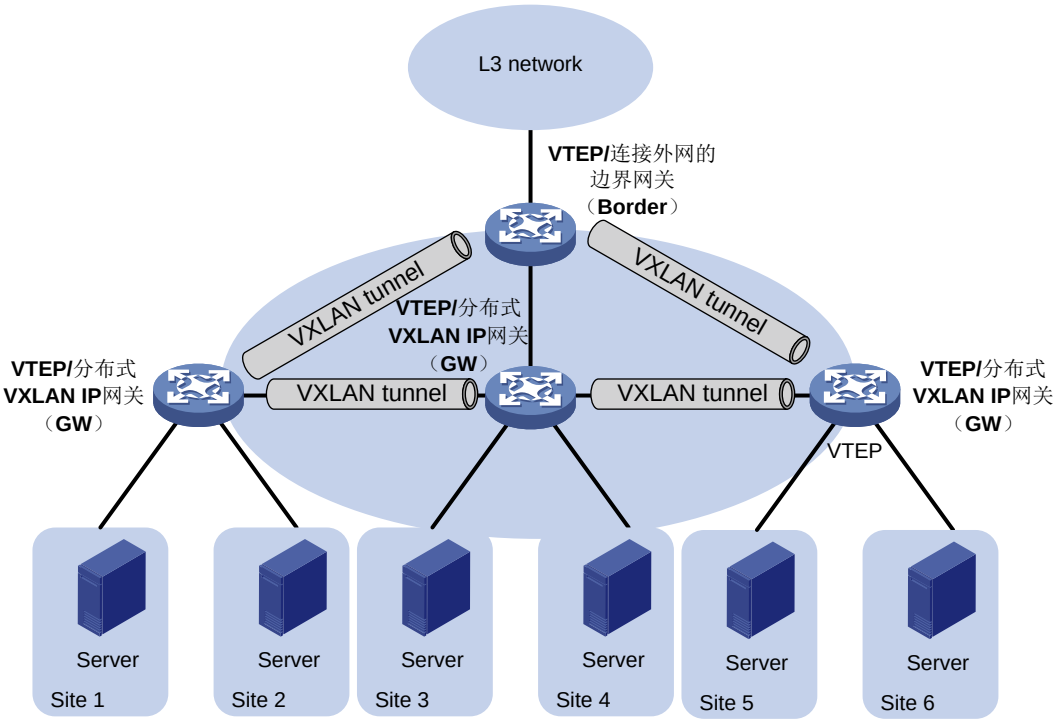
保护组中的成员 VTEP 之间、每个成员 VTEP 与接入层 VTEP 之间还会采用成员自身的 IP 地址建立 VXLAN 隧道，以便进行协议通信和表项同步。

3.1.4 分布式 VXLAN IP 网关

1. 简介

采用集中式 VXLAN IP 网关方案时，不同 VXLAN 之间的流量以及 VXLAN 访问外界网络的流量全部由集中式 VXLAN IP 网关处理，网关压力较大，并加剧了网络带宽资源的消耗。如图 3-5 所示，在分布式 VXLAN IP 网关方案中，每台 VTEP 设备都可以作为 VXLAN IP 网关，对本地站点的流量进行三层转发，很好地缓解了网关的压力。

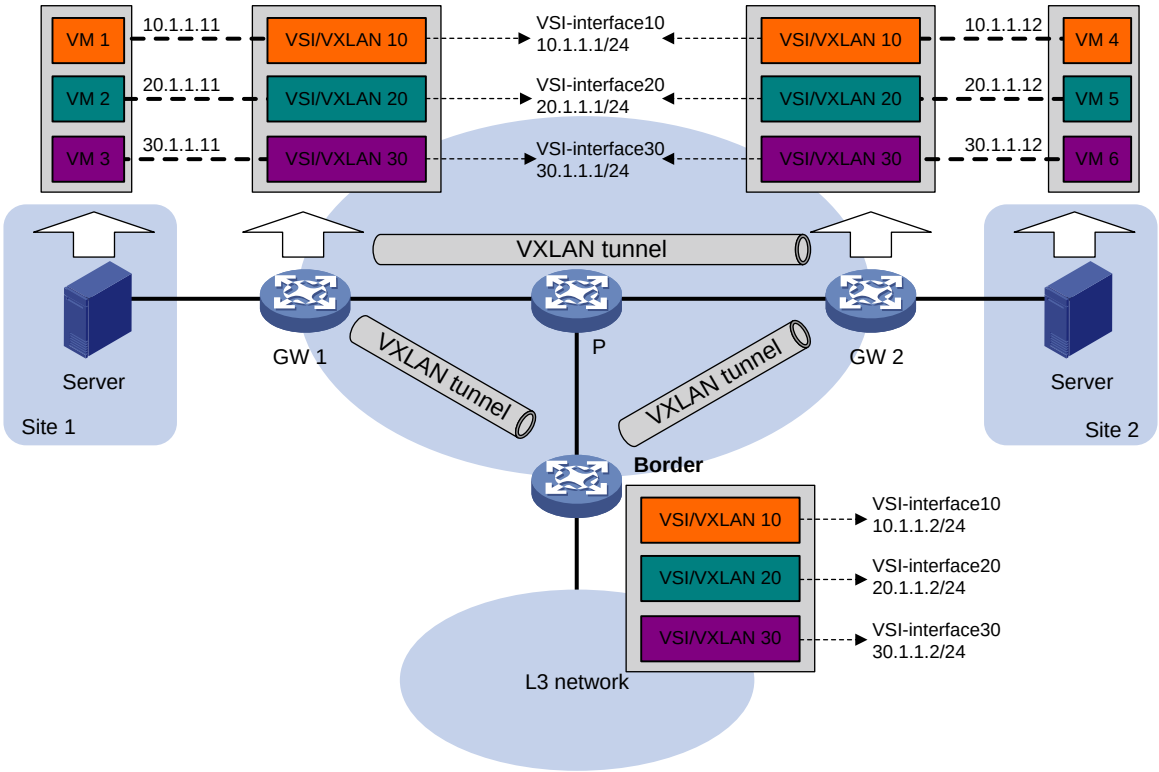
图3-5 分布式 VXLAN IP 网关示意图



如图 3-6 所示，在分布式 VXLAN IP 网关组网中，所有的分布式 VXLAN IP 网关（GW）上都需要创建 VSI 虚接口，并为不同 GW 上的相同 VSI 虚接口配置相同的 IP 地址，作为 VXLAN 内虚拟机的网关地址。边界网关（Border）上也需要创建 VSI 虚接口，并配置 IP 地址。在分布式 VXLAN IP 网关上还需要开启本地代理 ARP 功能或本地 ND 代理功能，开启后所有流量都通过查找 ARP 表项或 ND 表项进行三层转发。下文均以此功能为例，介绍分布式 VXLAN IP 网关中的通信过程。

网关可以通过多种方式生成 ARP 表项和 ND 表项，下文以根据 ARP 协议和 ND 协议动态学习表项来介绍分布式 VXLAN IP 网关中的通信过程。

图3-6 分布式 VXLAN IP 网关部署示意图



2. 相同 VXLAN 内不同站点的虚拟机通信过程

如图 3-6 所示，以 VM 1 访问 VM 4 为例，相同 VXLAN 内不同站点的虚拟机的通信过程为：

(1) VM 1 广播发送 ARP 请求消息，获取 VM 4 的 MAC 地址。

(2) GW 1 收到 ARP 请求消息后，学习 VM 1 的 ARP 信息，并代理应答该 ARP 请求，即：向 VM 1 发送 ARP 应答消息，应答的 MAC 地址为 VSI 虚接口 10 的 MAC 地址。

(3) VM 1 学习到 VM 4 的 MAC 地址为 GW 1 上 VSI 虚接口 10 的 MAC 地址。

(4) GW 1 将接收到的 ARP 请求消息中的源 MAC 地址修改为 VSI 虚接口 10 的 MAC 地址，在 VXLAN 10 内向本地站点和远端站点广播发送该 ARP 请求。

(5) GW 2 对 VXLAN 报文进行解封装后，学习 VM 1 的 ARP 信息（IP 为 10.1.1.11、MAC 为 GW 1 上 VSI 虚接口 10 的 MAC、出接口为接收该 VXLAN 报文的 Tunnel 接口），并将 ARP 请求消息中的源 MAC 修改为本地 VSI 虚接口 10 的 MAC 地址，在 VXLAN 10 的本地站点内进行广播。

(6) VM 4 收到 ARP 请求后，学习 VM 1 的 ARP 信息（IP 为 10.1.1.11、MAC 为 GW 2 上 VSI 虚接口 10 的 MAC），并发送 ARP 应答消息给本地网关 GW 2。

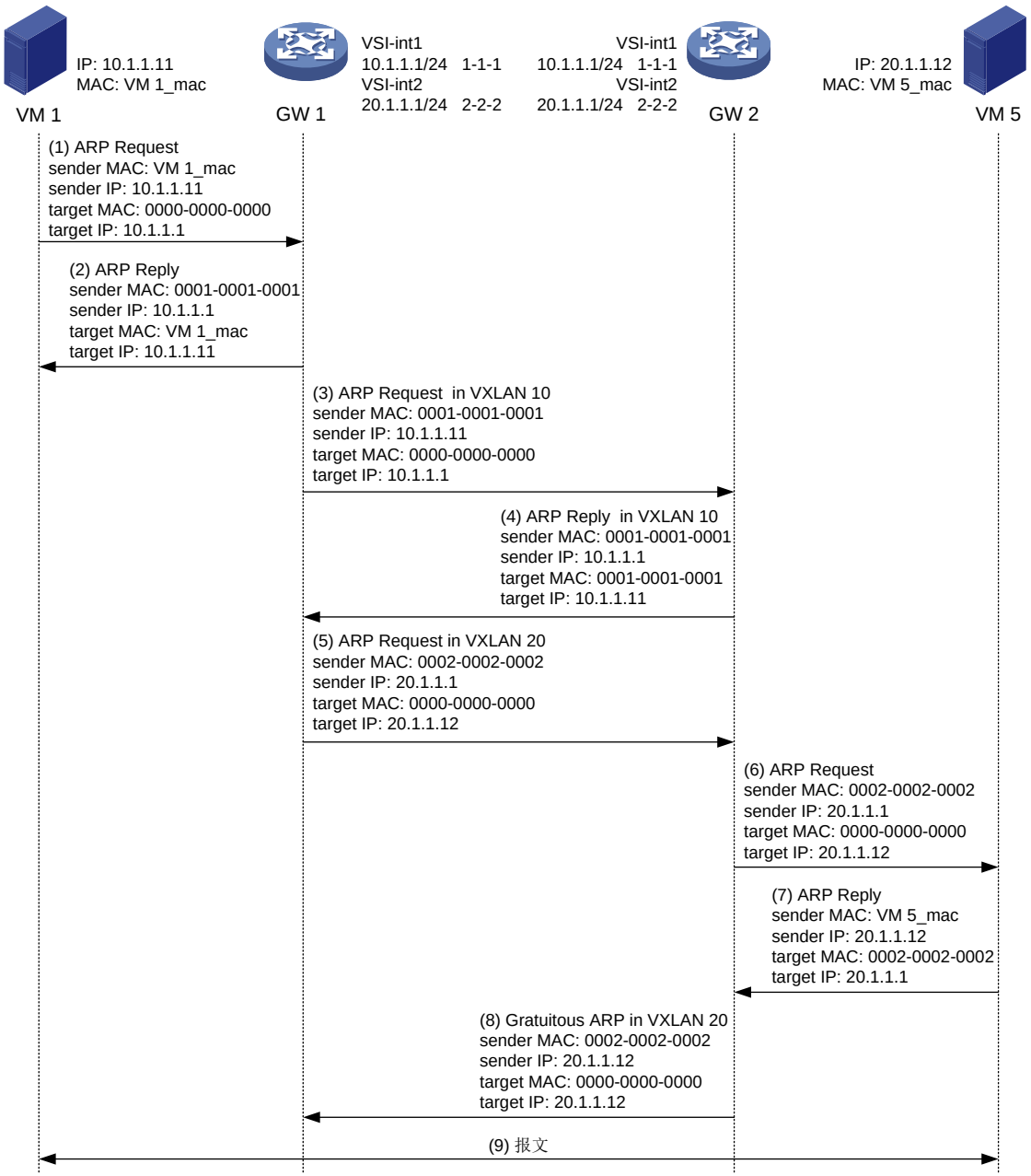
(7) GW 2 从 VM 4 收到 ARP 应答消息后，学习 VM 4 的 ARP 信息，将 ARP 应答消息中的源 MAC 修改为本地 VSI 虚接口 10 的 MAC 地址，并根据已经学习到的 ARP 表项，为 ARP 应答消息添加 VXLAN 封装后发送给 GW 1。

(8) GW 1 对 VXLAN 报文进行解封装后，根据收到的 ARP 应答消息学习 VM 4 的 ARP 信息（IP 为 10.1.1.12、MAC 为 GW 2 上 VSI 虚接口 10 的 MAC、出接口为接收该 VXLAN 报文的 Tunnel 接口）。

(9) 通过上述步骤完成 ARP 信息的学习后，VM 1 发送给 VM 4 的报文，根据已经学习到的 ARP 信息进行转发：首先发送给 GW 1；GW 1 对其进行 VXLAN 封装后，将其发送给 GW 2；GW 2 解封装后，将其发送给 VM 4。

3. 不同 VXLAN 间不同站点的虚拟机通信过程

图3-7 不同 VXLAN 间不同站点的虚拟机通信过程示意图



如图 3-7 所示，以 VM 1（VXLAN 10）访问 VM 5（VXLAN 20）为例，不同 VXLAN 的虚拟机的通信过程为：

VM 1 广播发送 ARP 请求消息，获取网关 10.1.1.1 的 MAC 地址。

(2) GW 1 收到 ARP 请求消息后，学习 VM 1 的 ARP 信息，并向 VM 1 发送 ARP 应答消息，应答的 MAC 地址为 VSI 虚接口 10 的 MAC 地址。这样，VM 1 会将访问 VM 5 的报文发送给 GW 1。

(3) GW 1 在 VXLAN 10 内向本地站点和远端站点广播发送 ARP 请求。ARP 请求消息中的源 IP 地址为 10.1.1.11、源 MAC 地址为本地 VSI 虚接口 10 的 MAC 地址。

(4) GW 2 从 VXLAN 隧道上接收到 VXLAN 报文，对其进行解封装后，学习 VM 1 的 ARP 信息（IP 为 10.1.1.11、MAC 为 GW 1 上 VSI 虚接口 10 的 MAC、出接口为接收该 VXLAN 报文的 Tunnel 接口），并将 ARP 请求消息中的源 MAC 修改为本地 VSI 虚接口 10 的 MAC 地址，在 VXLAN 10 的本地站点内广播该 ARP 请求消息。GW 2 发送 ARP 应答消息（IP 为 10.1.1.1、MAC 为 GW 2 上 VSI 虚接口 10 的 MAC）给 GW 1。

(5) GW 1 在 VXLAN 10 内发送 ARP 请求的同时，也会在 VXLAN 20 内向本地站点和远端站点广播发送 ARP 请求，获取 VM 5 的 MAC 地址。ARP 请求消息中的源 IP 地址为 20.1.1.1、源 MAC 地址为本地 VSI 虚接口 20 的 MAC 地址。

(6) GW 2 从 VXLAN 20 内收到 ARP 请求后，将 ARP 请求消息中的源 MAC 修改为本地 VSI 虚接口 20 的 MAC 地址，在 VXLAN 20 的本地站点内广播该 ARP 请求消息。

(7) VM 5 收到 ARP 请求后，学习 GW 2 的 ARP 信息（IP 为 20.1.1.1、MAC 为 GW 2 上 VSI 虚接口 20 的 MAC），并发送 ARP 应答消息给本地网关 GW 2。

(8) GW 2 从 VM 5 收到 ARP 应答消息后，学习 VM 5 的 ARP 信息，并向本地站点和远端站点发送免费 ARP。免费 ARP 消息中的源 IP 地址为 20.1.1.12、源 MAC 地址为本地 VSI 虚接口 20 的 MAC 地址。GW 1 从 VXLAN 隧道上接收到 VXLAN 报文，对其进行解封装后，根据收到的免费 ARP 消息学习 VM 5 的 ARP 信息（IP 为 20.1.1.12、MAC 为 GW 2 上 VSI 虚接口 20 的 MAC、出接口为接收该 VXLAN 报文的 Tunnel 接口）。

(9) 通过上述步骤完成 ARP 信息的学习后，VM 1 发送给 VM 5 的报文，根据已经学习到的 ARP 信息进行转发：首先发送给 GW 1；GW 1 对其进行 VXLAN 封装后，将其发送给 GW 2；GW 2 解封装后，将其发送给 VM 5。

4. 虚拟机与外部网络的三层通信过程

虚拟机要想与外部网络进行三层通信，需要在接入虚拟机的本地分布式 VXLAN IP 网关上指定流量的下一跳为 Border，可以通过如下方式来实现：

- 在本地分布式 VXLAN IP 网关上配置静态路由，指定路由下一跳为 Border 上同一个 VXLAN 对应 VSI 虚接口的 IP 地址。
- 在本地分布式 VXLAN IP 网关上配置策略路由，设置报文的下一跳为 Border 上同一个 VXLAN 对应 VSI 虚接口的 IP 地址。

如图 3-6 所示，以 VM 1 访问外部网络内的主机 50.1.1.1 为例，虚拟机访问外部网络的三层通信过程为：

(1) VM 1 广播发送 ARP 请求消息，获取网关 10.1.1.1 的 MAC 地址。

(2) GW 1 收到 ARP 请求消息后，学习 VM 1 的 ARP 信息，并向 VM 1 发送 ARP 应答消息，应答的 MAC 地址为 VSI 虚接口 10 的 MAC 地址。

(3) VM 1 将访问外部网络的报文发送给 GW 1。

(4) GW 1 接收到报文后，根据策略路由判断报文的下一跳地址为 10.1.1.2。GW 1 在 VXLAN 10 内向本地站点和远端站点广播发送 ARP 请求消息，获取 10.1.1.2 对应的 MAC 地址。

(5) Border 对 VXLAN 报文进行解封装，学习 GW 1 的 ARP 信息，并通过 VXLAN 隧道回复 ARP 应答消息。

(6) GW 1 对 VXLAN 报文进行解封装，并获取到 10.1.1.2 的 ARP 信息。

(7) GW 1 根据获取到的信息为 VM 1 发送的报文封装链路层地址（10.1.1.2 对应的 MAC 地址），并通过 VXLAN 隧道将报文发送给 Border。

(8) Border 对接收到的报文进行解封装后，对报文进行三层转发。

3.2 VXLAN IP网关配置限制和指导

建议不要在同一台设备上同时配置集中式 VXLAN IP 网关和集中式 VXLAN IP 网关保护组功能。

建议在 VXLAN IP 网关上为 VXLAN 隧道的出接口配置较大的 MTU 值，以免 VXLAN 报文分片导致转发失败。

3.3 VXLAN IP网关配置任务简介

VXLAN IP 网关配置任务如下：

(1) 配置 VXLAN IP 网关

请根据实际组网，选择以下一项任务进行配置：

- 配置集中式 VXLAN IP 网关
- 配置集中式 VXLAN IP 网关保护组
- 配置分布式 VXLAN IP 网关

(2) （可选）关闭 VXLAN 远端 ARP 自动学习功能

(3) （可选）配置 VSI 虚接口

3.4 VXLAN IP网关配置准备

配置集中式 VXLAN IP 网关和分布式 VXLAN IP 网关设备前，需要完成以下配置任务：

- 配置 VXLAN 隧道工作在三层转发模式。
- 创建 VSI 和 VXLAN。

3.5 配置集中式VXLAN IP网关

3.5.1 配置限制和指导

在集中式 VXLAN IP 网关组网中，请不要执行 **local-proxy-arp enable** 命令开启本地代理 ARP 功能。

3.5.2 配置集中式网关的网关接口

(1) 进入系统视图。

system-view

(2) 创建 VSI 虚接口，并进入 VSI 虚接口视图。

interface vsi-interface *vsi-interface-id*

(3) 配置 VSI 虚接口的 IP 地址。

ip address *ip-address* { *mask* | *mask-length* }

缺省情况下，未配置 VSI 虚接口的 IP 地址。

(4) 退回系统视图。

quit

(5) 进入 VXLAN 所在 VSI 视图。

vsi *vsi-name*

(6) 为 VSI 指定网关接口。

gateway vsi-interface *vsi-interface-id*

缺省情况下，未指定 VSI 的网关接口。

3.6 配置集中式VXLAN IP网关保护组

3.6.1 VXLAN IP 网关上的配置

1. 配置限制和指导

保护组中所有网关上的 VXLAN 配置需要保证完全一致。

2. 配置步骤

(1) 进入系统视图。

system-view

(2) 创建 VSI 虚接口，并进入 VSI 虚接口视图。

interface vsi-interface *vsi-interface-id*

(3) 配置 VSI 虚接口的 IP 地址。

ip address *ip-address* { *mask* | *mask-length* }

缺省情况下，未配置 VSI 虚接口的 IP 地址。

请在保护组中的每台网关上配置相同的 VSI 虚接口 IP 地址。

(4) 配置 VSI 虚接口的 MAC 地址。

mac-address *mac-address*

保护组中所有网关上配置的 MAC 地址必须相同。

(5) 退回系统视图。

quit

(6) 进入 VXLAN 所在 VSI 视图。

vsi *vsi-name*

(7) 为 VSI 指定网关接口。

gateway vsi-interface *vsi-interface-id*

缺省情况下，未指定 VSI 的网关接口。

(8) 退回系统视图。

quit

(9) 将本设备加入 VXLAN IP 网关保护组，并配置本设备的成员地址。

vtep group *group-ip* **member local** *member-ip*

缺省情况下，设备未加入 VXLAN IP 网关保护组。

member-ip 为本设备的成员地址，该地址必须是设备上已经存在的 IP 地址，并且需要通过路由协议发布到 IP 网络。同一个保护组中不同成员 VTEP 的地址不能相同。

(10) 配置 VXLAN IP 网关保护组的成员地址列表。

vtep group *group-ip* **member remote** *member-ip*&<1-8>

缺省情况下，未配置 VXLAN IP 网关保护组的成员地址列表。

执行本命令时，必须输入保护组中所有其它成员的成员地址。

3.6.2 接入层 VTEP 上的配置

1. 配置准备

执行本配置时，需要完成以下配置任务：

- 配置 VXLAN 隧道工作在二层转发模式。
- 创建 VSI 和 VXLAN。
- 配置 VXLAN 隧道，并将 VXLAN 与 VXLAN 隧道关联。

2. 配置步骤

(1) 进入系统视图。

system-view

(2) 配置 VXLAN IP 网关保护组的成员地址列表。

vtep group *group-ip* **member remote** *member-ip*&<1-8>

缺省情况下，未配置 VXLAN IP 网关保护组的成员地址列表。

执行本命令时，必须输入保护组中所有成员的成员地址。

3.7 配置分布式VXLAN IP网关

3.7.1 配置限制和指导

如果虚拟机要想与外部网络进行三层通信，除本配置外，还需要在接入虚拟机的本地分布式 VXLAN IP 网关上配置静态路由或策略路由：

- 配置静态路由：指定路由的下一跳为 Border 上同一个 VXLAN 对应 VSI 虚接口的 IP 地址。
- 配置策略路由：通过 **apply default-next-hop** 命令设置报文的缺省下一跳为 Border 上同一个 VXLAN 对应 VSI 虚接口的 IP 地址。策略路由的配置方法，请参见“三层技术-IP 路由配置指导”中的“策略路由”。

分布式 VXLAN IP 网关连接 IPv4 站点网络时，所有网关上都需要为相同 VSI 虚接口配置相同的 MAC 地址。如果网关同时连接 IPv4 站点网络和 IPv6 站点网络，则不同分布式 VXLAN IP 网关上需要为相同 VSI 虚接口配置不同的链路本地地址。

3.7.2 配置分布式网关的网关接口

(1) 进入系统视图。

system-view

(2) 创建 VSI 虚接口，并进入 VSI 虚接口视图。

interface vsi-interface *vsi-interface-id*

(3) 配置 VSI 虚接口的 IP 地址。

（IPv4 网络）

ip address *ip-address* { *mask* | *mask-length* } [**sub**]

（IPv6 网络）

IPv6 地址的配置方法，请参见“三层技术-IP 业务配置指导”中的“IPv6 基础”。

缺省情况下，未配置 VSI 虚接口的 IP 地址。

(4) 配置 VSI 虚接口为分布式网关接口。

distributed-gateway local

缺省情况下，VSI 虚接口不是分布式本地网关接口。

(5) 开启本地代理 ARP 或本地 ND 代理功能。

（IPv4 网络）

local-proxy-arp enable [**ip-range** *startIP* **to** *endIP*]

缺省情况下，本地代理 ARP 功能处于关闭状态。

本命令的详细介绍，请参见“三层技术-IP 业务命令参考”中的“代理 ARP”。

（IPv6 网络）

local-proxy-nd enable

缺省情况下，本地 ND 代理功能处于关闭状态。

本命令的详细介绍，请参见“三层技术-IP 业务命令参考”中的“IPv6 基础”。

(6) 开启 VSI 虚接口。

undo shutdown

缺省情况下，VSI 虚接口处于开启状态。

(7) 退回系统视图。

quit

(8) 进入 VXLAN 所在 VSI 视图。

vsi *vsi-name*

(9) 为 VSI 指定网关接口。

gateway vsi-interface *vsi-interface-id*

缺省情况下，未指定 VSI 的网关接口。

3.7.3 开启分布式网关的动态 ARP 表项同步功能

1. 功能简介

分布式 VXLAN IP 网关上开启本地代理 ARP 功能时，本地网关不会将目标 IP 地址为分布式网关 VSI 虚接口的 ARP 报文转发给其他网关，只有本地网关能够学习到 ARP 报文发送者的 ARP 表项。如果希望所有网关都能学习到该 ARP 表项，需要开启分布式网关的动态 ARP 表项同步功能。分布式 VXLAN IP 网关之间也可以通过控制器在彼此之间同步 ARP 表项，此时无需开启该功能。

2. 配置步骤

(1) 进入系统视图。

system-view

(2) 开启分布式网关的动态 ARP 表项同步功能。

arp distributed-gateway dynamic-entry synchronize

缺省情况下，分布式网关的动态 ARP 表项同步功能处于关闭状态。

3.8 关闭VXLAN远端ARP自动学习功能

1. 功能简介

缺省情况下，设备从 VXLAN 隧道接收到报文后可以自动学习远端虚拟机的 ARP 信息，即远端 ARP 信息。在 SDN 控制器组网下，当控制器和设备间进行表项同步时，可以通过本配置暂时关闭远端 ARP 自动学习功能，以节省占用的设备资源。同步完成后，再开启远端 ARP 自动学习功能。

2. 配置限制和指导

建议用户只在控制器和设备间同步表项的情况下执行本配置。

3. 配置步骤

(1) 进入系统视图。

system-view

(2) 关闭 VXLAN 远端 ARP 自动学习功能。

vxlan tunnel arp-learning disable

缺省情况下，远端 ARP 自动学习功能处于开启状态。

3.9 配置VSI虚接口

3.9.1 配置 VSI 虚接口的可选参数

(1) 进入系统视图。

system-view

(2) 进入 VSI 虚接口视图。

interface vsi-interface *vsi-interface-id*

(3) 配置接口的 MAC 地址。

mac-address *mac-address*

缺省情况下，VSI 虚接口的 MAC 地址为设备的桥 MAC 地址。

(4) 配置接口的描述信息。

description *text*

缺省情况下，接口的描述信息为“*接口名* Interface”，例如：Vsi-interface100 Interface。

(5) 配置接口的 MTU。

mtu *mtu-value*

缺省情况下，VSI 虚接口的 MTU 值为 1500。

(6) 配置接口的期望带宽。

bandwidth *bandwidth-value*

缺省情况下，接口的期望带宽＝接口的波特率÷1000（kbps）。

期望带宽供业务模块使用，不会对接口实际带宽造成影响。

3.9.2 恢复 VSI 虚接口的缺省配置

1. 配置限制和指导

注意

接口下的某些配置恢复到缺省情况后，会对设备上当前运行的业务产生影响。建议您在执行本配置前，完全了解其对网络产生的影响。

您可以在执行 **default** 命令后通过 **display this** 命令确认执行效果。对于未能成功恢复缺省的配置，建议您查阅相关功能的命令手册，手工执行恢复该配置缺省情况的命令。如果操作仍然不能成功，您可以通过设备的提示信息定位原因。

2. 配置步骤

(1) 进入系统视图。

system-view

(2) 进入 VSI 虚接口视图。

interface vsi-interface *vsi-interface-id*

(3) 恢复接口的缺省配置。

default

3.10 VXLAN IP网关显示和维护

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 IP 网关的运行情况，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 VSI 虚接口的统计信息。

表3-1 VXLAN IP 网关显示和维护

操作	命令
显示VSI虚接口信息	**display interface** [**vsi-interface**] [**brief** [**description** | **down**]]

操作	命令
清除VSI虚接口的统计信息	**reset counters interface** [**vsi-interface** [*vsi-interface-id*]]