H3C S12500-X & S12500X-AF 系列以太网交换机

VXLAN 配置指导

杭州华三通信技术有限公司
http://www.h3c.com.cn

资料版本：6W101-20151130
产品版本：Release 1135 及以上

前 言

H3C S12500-X & S12500X-AF 系列以太网交换机 配置指导（R113x）共分为十七本手册，介绍了S12500-X & S12500X-AF 系列以太网交换机 Release 1135 及以上版本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》主要介绍了 VXLAN 技术的原理及具体配置方法。

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. 端口编号示例约定

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。

资料获取方式

您可以通过H3C网站（www.h3c.com.cn）获取最新的产品资料：

H3C 网站与产品资料相关的主要栏目介绍如下：

- [服务支持/文档中心]：可以获取硬件安装类、软件升级类、配置类或维护类等产品资料。
- [产品技术]：可以获取产品介绍和技术介绍的文档，包括产品相关介绍、技术介绍、技术白皮书等。
- [解决方案]：可以获取解决方案类资料。
- [服务支持/软件下载]：可以获取与软件版本配套的资料。

技术支持

用户支持邮箱：service@h3c.com

技术支持热线电话：400-810-0504（手机、固话均可拨打）

网址：http://www.h3c.com.cn

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简介

说明

- 仅 FC、FE、FX 类型的单板支持配置 VXLAN 功能。
- FC 类型的单板不支持用作 VXLAN IP 网关。
- FE、FX 类型的单板用作集中式 VXLAN IP 网关时，不能同时用作接入本地站点的 VTEP 设备。
- VXLAN 功能受设备的工作模式限制，在使用 VXLAN 功能前，请在系统视图下使用“**system-working-mode standard**”命令将设备设置为标准模式，需要注意的是，使用该命令修改设备的工作模式时，需要保存设备当前配置文件，再删除文件名后缀为“.mdb”的二进制类型的配置文件，然后重启设备后才能生效。有关设备工作模式的详细介绍请参见“基础配置指导”中的“设备管理”。

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。

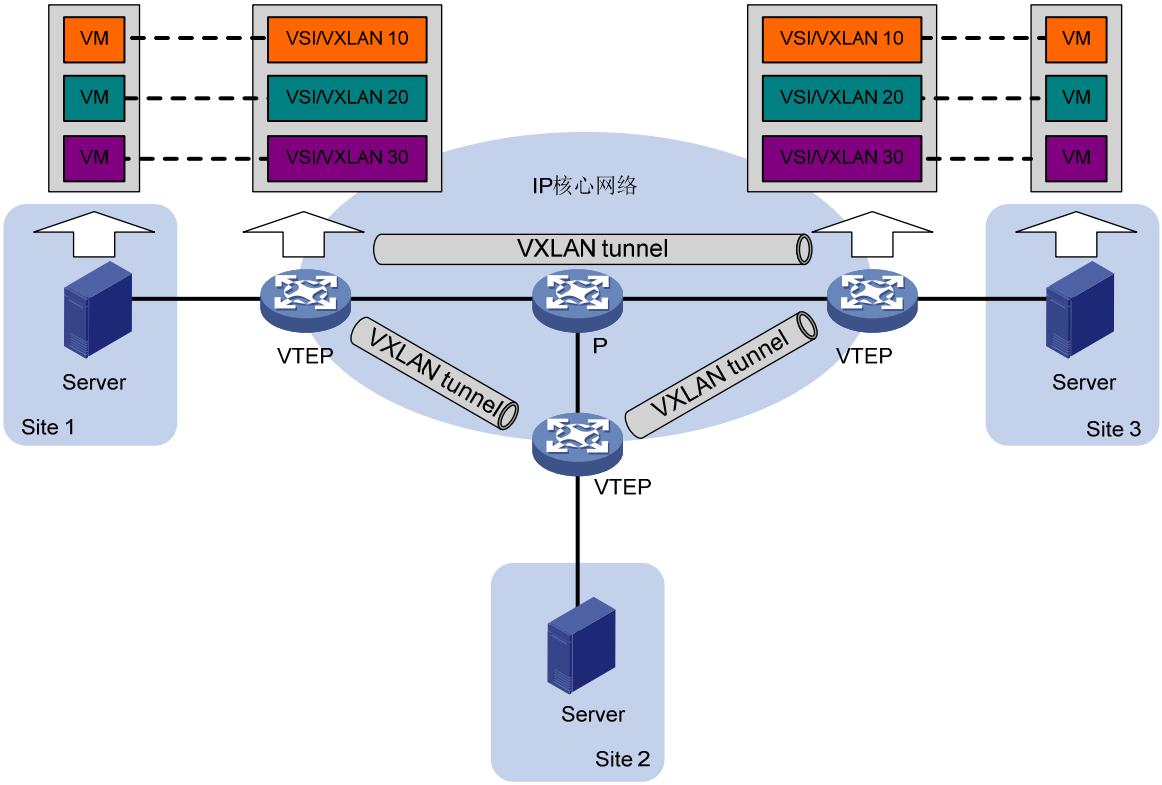
VXLAN 具有如下特点：

- 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，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。

目前，设备只支持基于 IPv4 网络的 VXLAN 技术。

1.1 VXLAN网络模型

图1-1 VXLAN 网络模型示意图

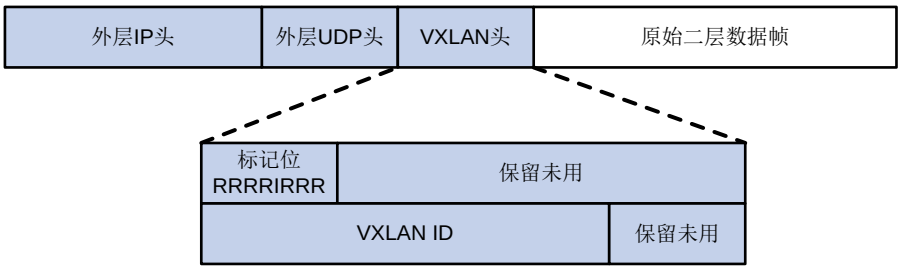


如 图 1-1 所示，VXLAN的典型网络模型中包括如下几部分：

- VM（Virtual Machine，虚拟机）：在一台服务器上可以创建多台虚拟机，不同的虚拟机可以属于不同的 VXLAN。属于相同 VXLAN 的虚拟机处于同一个逻辑二层网络，彼此之间二层互通；属于不同 VXLAN 的虚拟机之间二层隔离。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。
- 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。VTEP 可以是一台独立的物理设备，也可以是虚拟机所在的服务器。
- 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、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 Switching 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 可以看做是 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。

1.2 VXLAN报文封装格式

图1-2 VXLAN 报文封装示意图



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。

1.3 VXLAN运行机制

VXLAN 运行机制可以概括为：

(1) 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。

(2) 学习虚拟机的 MAC 地址。

(3) 根据学习到的 MAC 地址表项转发报文。

1.3.2 识别报文所属的VXLAN

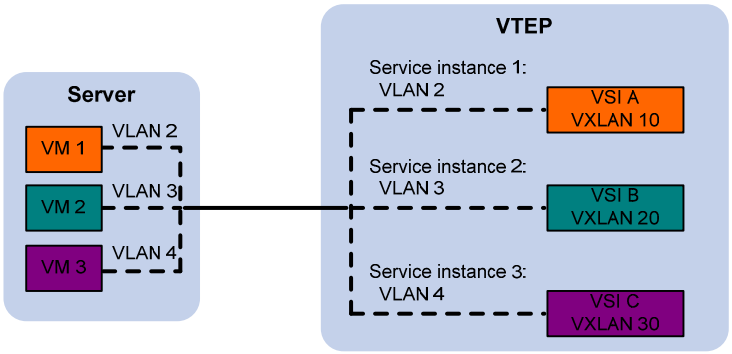
1. 本地站点内接收到数据帧的识别

VTEP 将连接本地站点的以太网服务实例（Service Instance）与 VSI 关联。VTEP 从以太网服务实例接收到数据帧后，查找与其关联的 VSI，VSI 内创建的 VXLAN 即为该数据帧所属的 VXLAN。

在 VXLAN 中，与 VSI 关联的以太网服务实例统称为 AC（Attachment Circuit，接入电路）。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。

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。

1.3.3 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。

1. 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。

2. 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。

静态配置的远端 MAC 地址表项优先级高于源 MAC 地址动态学习的表项。

1.3.4 接入模式

接入模式分为以下两种：

- VLAN 接入模式：从本地站点接收到的、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 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。
- Ethernet 接入模式：从本地站点接收到的、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。

缺省情况下，接入模式为 VLAN 模式。下文对于流量转发过程的介绍均以 VLAN 模式为例。

1.3.5 VXLAN隧道工作模式

VXLAN 隧道支持如下两种工作模式：

- 三层转发模式：VTEP 设备通过查找 ARP 表项对流量进行转发。
- 二层转发模式：VTEP 通过查找 MAC 地址表项对流量进行转发。

本章主要介绍VXLAN工作在二层转发模式时的工作原理，当VXLAN隧道工作在三层转发模式时，可以用作VXLAN IP网关，具体介绍请参见“3 VXLAN IP网关”。

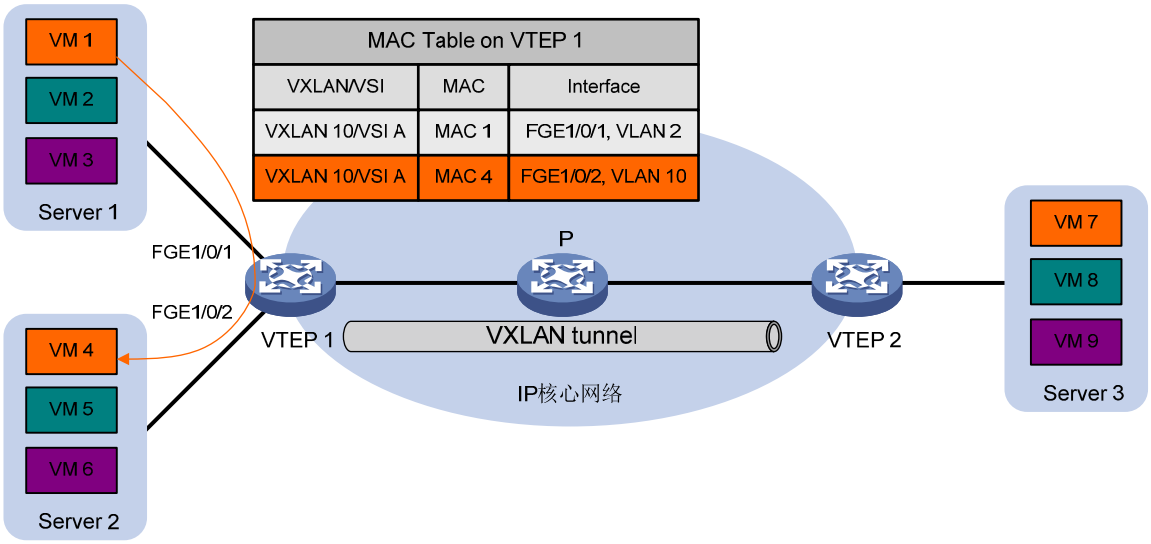
1.3.6 转发单播流量

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。

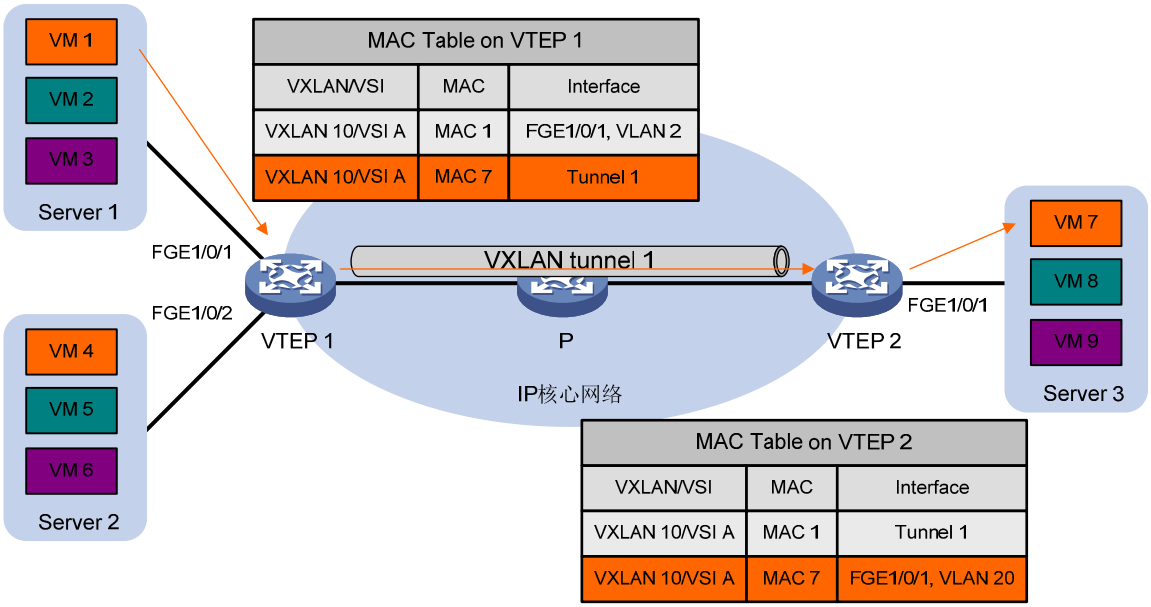
图1-4 站点内单播流量转发



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FortyGigE1/0/1 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FortyGigE1/0/2，所在VLAN为VLAN 10，则将以太网帧从接口FortyGigE1/0/2 的VLAN 10 内发送给VM 4。

2. 站点间流量

图1-5 站点间单播流量转发



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 tag 为 2。

(2) VTEP 1 从接口 FortyGigE1/0/1 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 FortyGigE1/0/1，所在 VLAN 为 VLAN 20。

(6) VTEP 2 从接口 FortyGigE1/0/1 的 VLAN 20 内将数据帧发送给 VM 7。

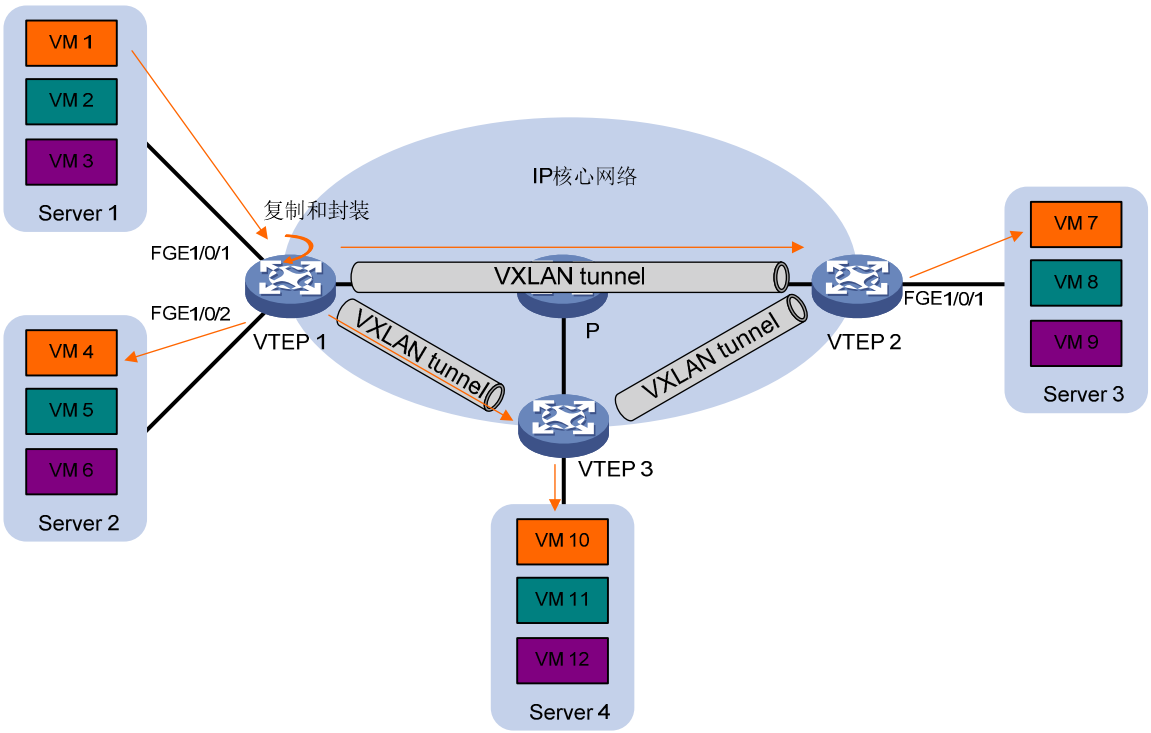
1.3.7 转发泛洪流量

泛洪流量包括组播、广播和未知单播流量。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）和组播路由方式（核心复制）和泛洪代理方式（服务器复制）三种。

1. 单播路由方式（头端复制）

在单播路由方式下，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



如 图 1-6 所示，单播路由方式的泛洪流量转发过程为：

(1) VTEP 1 接收到本地虚拟机发送的组播、广播和未知单播数据帧后，判断数据帧所属的 VXLAN，通过该 VXLAN 内除接收接口外的所有本地接口和 VXLAN 隧道转发该数据帧。通过 VXLAN 隧道转发数据帧时，需要为其封装 VXLAN 头、UDP 头和 IP 头，将泛洪流量封装在多个单播报文中，发送到 VXLAN 内的所有远端 VTEP。

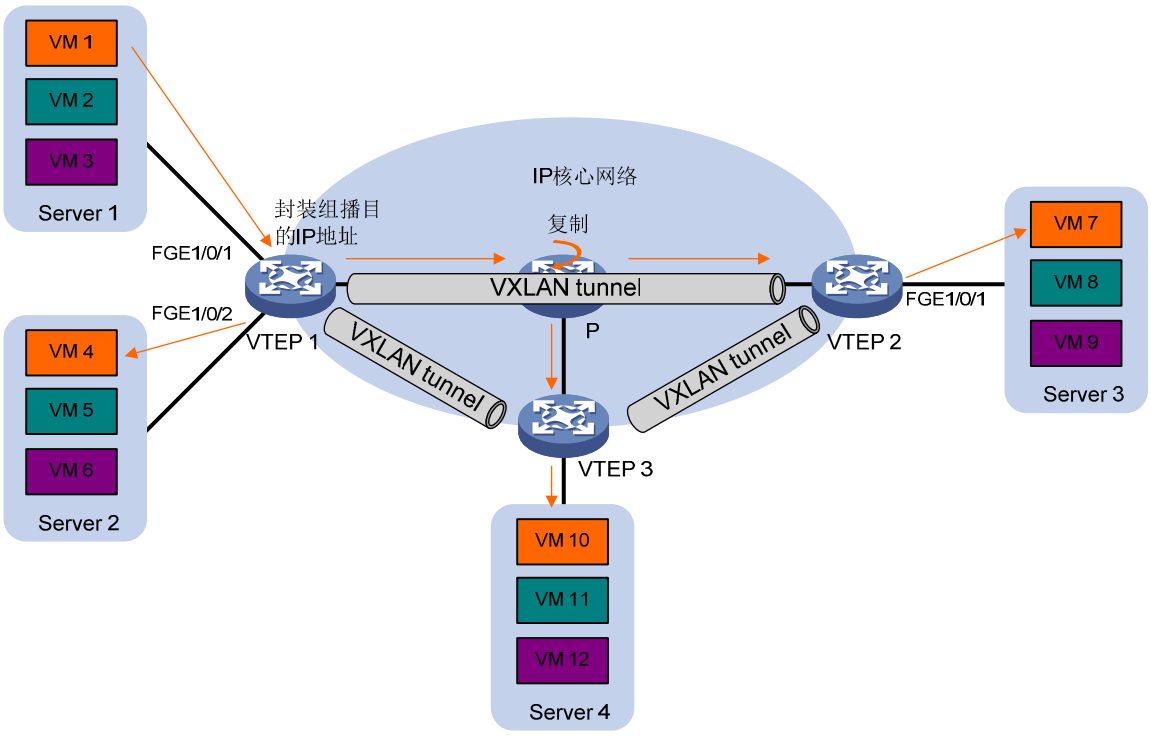
(2) 远端 VTEP（VTEP 2 和 VTEP 3）接收到 VXLAN 报文后，解封装报文，将原始的数据帧在本地站点的指定 VXLAN 内泛洪。为了避免环路，远端 VTEP 从 VXLAN 隧道上接收到报文后，不会再将其泛洪到其他的 VXLAN 隧道。

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。

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



如 图 1-7 所示，组播路由方式的泛洪流量转发过程为：

(1) VTEP 1 接收到本地虚拟机发送的组播、广播和未知单播数据帧后，判断数据帧所属的 VXLAN，不仅通过该 VXLAN 内除接收接口外的所有本地接口将数据帧转发到本地站点，还会为其封装 VXLAN 头、UDP 头和 IP 头（目的 IP 地址为组播地址）通过组播转发表项将其发送到远端 VTEP。

(2) 在 IP 核心网内，P 设备根据已经建立的组播转发表项复制并转发该组播报文。

(3) 远端 VTEP（VTEP 2 和 VTEP 3）接收到 VXLAN 报文后，解封装报文，将原始的数据帧在本地站点的指定 VXLAN 内泛洪。为了避免环路，远端 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。

3. 泛洪代理方式（服务器复制）

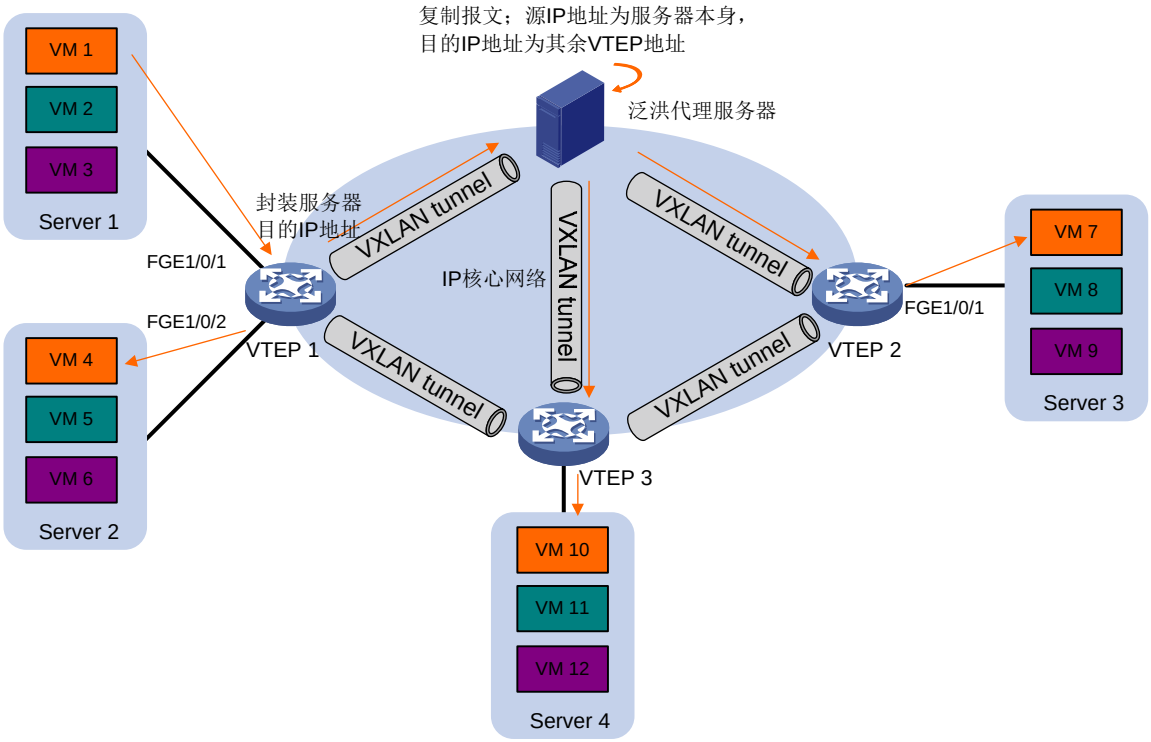
说明

仅 Release 1138P01 及以上版本支持此方式。

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。

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1-8 泛洪代理方式转发示意图



如 图 1-8 所示，泛洪代理方式的流量转发过程为：

(1) VTEP 1 接收到本地虚拟机发送的组播、广播和未知单播数据帧后，判断数据帧所属的 VXLAN，不仅通过该 VXLAN 内除接收接口外的所有本地接口将数据帧转发到本地站点，还会为其封装 VXLAN 头、UDP 头和 IP 头（目的 IP 地址为泛洪代理服务器地址）通过指定的隧道将其发送到泛洪代理服务器。

(2) 泛洪代理服务器收到报文后，修改报文的 IP 头，源地址为服务器本身，目的 IP 地址为其余 VTEP 地址，通过不同的隧道发送到远端 VTEP。

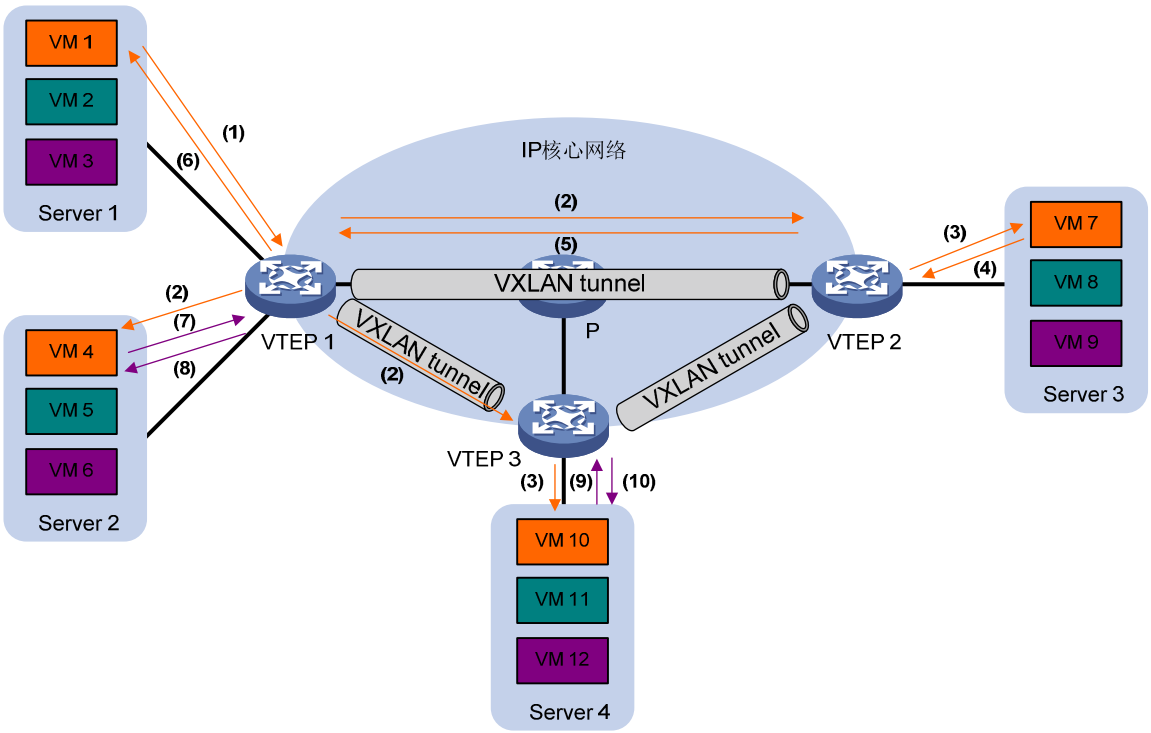
(3) 远端 VTEP（VTEP 2 和 VTEP 3）接收到 VXLAN 报文后，解封装报文，将原始的数据帧在本地站点的指定 VXLAN 内泛洪。为了避免环路，远端 VTEP 从 VXLAN 隧道上接收到报文后，不会再将其泛洪到其他的 VXLAN 隧道。

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。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。

1.4 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、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。

1.5 协议规范

与 VXLAN 相关的协议规范有：

- IETF 草案：draft-mahalingam-dutt-dcops-vxlan-04

2 配置VXLAN

2.1 VXLAN配置任务简介

在 VXLAN 组网中，IP 核心网络中的设备只需要配置路由协议，确保 VTEP 之间路由可达。VXLAN 相关配置都在 VTEP 上进行。

当 VXLAN 隧道工作在三层转发模式时，在 VSI 视图创建 VXLAN 前，需要先配置全局类型 VLAN 接口资源预留。每创建一个 VSI 虚接口，需要预留一个全局类型 VLAN 接口资源。关于全局类型 VLAN 接口资源预留的详细介绍，请参见“二层技术-以太网交换配置指导”中的“VLAN”。

表2-1 VXLAN 配置任务简介

配置任务	说明	详细配置
配置VXLAN隧道工作模式	必选	2.2
创建VSI和VXLAN	必选	2.3
创建VXLAN隧道	必选	2.4
关联VXLAN与VXLAN隧道	必选	2.5
配置AC与VSI关联	必选	2.6
管理本地和远端MAC地址	可选	2.7
配置VSI泛洪抑制	可选	2.9
配置VXLAN报文的目的UDP端口号	可选	2.10
配置VXLAN报文检查功能	可选	2.11
配置ARP泛洪抑制	可选	2.12
关闭VXLAN远端ARP自动学习功能	可选 仅Release 1138P01及以上版本支持此功能	2.13
配置VXLAN流量统计	可选	2.14
配置VXLAN的硬件资源模式	可选 仅Release 1138P01及以上版本支持此功能	2.15

2.2 配置VXLAN隧道工作模式

表2-2 配置 VXLAN 隧道工作模式

操作	命令	说明
进入系统视图	**system-view**	-

操作	命令	说明
配置VXLAN隧道工作在二层转发模式	**undo vxlan ip-forwarding**	二者选其一 缺省情况下，VXLAN隧道工作在三层转发模式 当设备作为VTEP时，需要配置VXLAN隧道工作在二层转发模式；当设备作为VXLAN IP网关时，需要配置VXLAN隧道工作在三层转发模式。有关VXLAN IP网关的详细介绍请参见“3 VXLAN IP网关” 修改本配置前必须先删除设备上的所有VSI、VSI虚接口和VXLAN隧道，否则将配置失败
配置VXLAN隧道工作在三层转发模式	**vxlan ip-forwarding** [**tagged** | **untagged**]	

2.3 创建VSI和VXLAN

表2-3 创建 VSI 和 VXLAN

操作	命令	说明
进入系统视图	**system-view**	-
使能L2VPN功能	**l2vpn enable**	缺省情况下，L2VPN功能处于关闭状态
创建VSI，并进入VSI视图	**vsi** *vsi-name*	缺省情况下，设备上不存在任何VSI
（可选）设置VSI的描述信息	**description** *text*	缺省情况下，未配置VSI的描述信息
开启当前的VSI	**undo shutdown**	缺省情况下，VSI处于开启状态
创建VXLAN，并进入VXLAN视图	**vxlan** *vxlan-id*	缺省情况下，设备上不存在任何VXLAN 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

2.4 创建VXLAN隧道

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。

请不要将 VXLAN 隧道出接口和 AC 口配置为同一个接口，以避免流量转发失败。

表2-4 手工创建 VXLAN 隧道

操作	命令	说明
进入系统视图	**system-view**	-

操作	命令	说明
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 仅Release 1138P01及以上版本支持此命令
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**	缺省情况下，设备上不存在任何Tunnel接口 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 只有*tunnel-number*为0～511的VXLAN隧道支持ECMP功能
配置隧道的源端地址或源接口	**source** { *ipv4-address* | *interface-type interface-number* }	缺省情况下，没有设置VXLAN隧道的源端地址和源接口 如果设置的是隧道的源端地址，则该地址将作为封装后VXLAN报文的源IP地址；如果设置的是隧道的源接口，则该接口的主IP地址将作为封装后VXLAN报文的源IP地址 采用VXLAN组播路由泛洪方式时，VXLAN隧道的源接口不能是Loopback接口、源端地址不能是Loopback接口的地址
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 关联VXLAN与VXLAN隧道

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。

表2-5 手工关联 VXLAN 与 VXLAN 隧道

操作	命令	说明
进入系统视图	**system-view**	-
进入VSI视图	**vsi** *vsi-name*	-
进入VXLAN视图	**vxlan** *vxlan-id*	-

操作	命令	说明
配置VXLAN与VXLAN隧道关联	**tunnel** *tunnel-number* [**flooding-proxy**]	缺省情况下，VXLAN没有与任何VXLAN隧道关联 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 仅Release 1138P01及以上版本支持**flooding-proxy**参数

2.6 配置AC与VSI关联

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。

表2-6 配置以太网服务实例与 VSI 关联

操作		命令	说明
进入系统视图		**system-view**	-
进入二层以太网接口视图或二层聚合接口视图	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*	
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*	缺省情况下，不存在任何以太网服务实例
配置以太网服务实例的报文匹配规则	匹配当前端口接收的所有报文	**encapsulation default**	请用户选择其中一种匹配方式进行配置 缺省情况下，未配置任何报文匹配规则 本系列设备暂不支持匹配携带任意VLAN标签的报文 以太网服务实例所匹配的VLAN必须在设备上已经创建并配置当前接口允许以太网服务实例所匹配VLAN的报文通过
	匹配携带任意VLAN标签或不携带VLAN标签的报文	**encapsulation** { **tagged** | **untagged** }	
	匹配携带指定外层VLAN标签的报文	**encapsulation s-vid** *vlan-id* [**only-tagged**]	
	匹配携带指定外层VLAN标签和内层VLAN标签的报文	**encapsulation s-vid** *vlan-id* **c-vid** *vlan-id*	
将以太网服务实例与VSI关联		**xconnect vsi** *vsi-name* [**access-mode** { **ethernet** | **vlan** }]	缺省情况下，以太网服务实例没有与VSI关联

注意

- 在接口上创建以太网服务实例匹配当前接口接收的部分 VLAN 的报文后，以太网服务实例未匹配的其它 VLAN 内的组播流量将无法正常转发，请注意避免在上述情况下使用组播业务。
- 接入层 VTEP 设备需要开启 ARP 泛洪抑制功能时，需要在设备上创建与以太网服务实例所匹配 VLAN 对应的 VLAN 虚接口，该 VLAN 虚接口可以不配置 IP 地址，但是需保证允许该 VLAN 通过的端口至少有一个处于 Up 状态。

2.7 管理本地和远端MAC地址

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 地址表项可以静态添加，也可以根据接收到的 VXLAN 报文内封装的源 MAC 地址自动学习。

2.7.1 配置本端MAC地址添加/删除的日志功能

执行本配置后，VXLAN 添加、删除本地 MAC 地址时，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）。

表2-7 配置本端 MAC 地址添加/删除的日志功能

操作	命令	说明
进入系统视图	**system-view**	-
开启VXLAN本地MAC地址添加/删除的日志功能	**vxlan local-mac report**	缺省情况下，VXLAN添加/删除本地MAC地址时不会记录日志信息

2.7.2 添加静态远端MAC地址

表2-8 添加静态远端 MAC 地址

操作	命令	说明
进入系统视图	**system-view**	-
添加静态远端MAC地址表项	**mac-address static** *mac-address* **interface tunnel** *tunnel-number* **vsi** *vsi-name*	缺省情况下，设备上不存在任何静态的远端MAC地址表项 **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，且该VXLAN必须已经创建，否则配置将失败

2.7.3 开启远端MAC地址自动学习功能

缺省情况下，设备可以自动学习远端 MAC 地址。如果网络中存在攻击，为了避免学习到错误的远端 MAC 地址，也可以手工关闭远端 MAC 地址自动学习功能。

表2-9 开启远端 MAC 地址自动学习功能

操作	命令	说明
进入系统视图	**system-view**	-
开启远端MAC地址自动学习功能	**undo 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

2.8 配置VXLAN组播路由泛洪方式

说明

配置 VXLAN 采用组播路由方式转发泛洪流量后，如果组播路由隧道 down 掉，即使还存在单播路由隧道，泛洪流量也不会沿单播路由隧道转发。

配置 VXLAN 组播路由泛洪方式时，需要完成以下配置任务：

- 在 VTEP 和核心设备上启动三层组播路由功能。
- 在核心设备上配置 IGMP 和组播路由协议。

表2-10 配置 VXLAN 组播路由泛洪方式

操作	命令	说明
进入系统视图	**system-view**	-
进入VSI视图	**vsi** *vsi-name*	-
进入VXLAN视图	**vxlan** *vxlan-id*	-
配置VXLAN泛洪的组播地址和组播报文的源IP地址	**group** *group-address* **source** *source-address*	缺省情况下，未指定VXLAN泛洪的组播地址和组播报文的源IP地址，VXLAN采用单播路由方式泛洪 执行本命令后，VTEP将加入指定的组播组。同一VXLAN的所有VTEP都要加入相同的组播组 为确保组播报文转发正常，VXLAN组播报文的源IP地址（*source-address*）应指定为一个已创建且处于up状态的VXLAN隧道的源端地址
进入组播报文源IP地址所在接口的接口视图	**interface** *interface-type interface-number*	组播报文源IP地址是指通过**group**命令中的**source**参数指定的地址

操作	命令	说明
在接口上使能IGMP协议的主机功能	**igmp host enable**	缺省情况下，接口上IGMP协议的主机功能处于关闭状态 执行本命令后，当前接口将作为IGMP主机，即从该接口收到IGMP查询报文后，通过该接口发送组播组的报告报文，以便接收该组播组的报文 只有通过**multicast routing**命令使能IP组播路由后，本命令才会生效

2.9 配置VSI泛洪抑制

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。

表2-11 配置 VSI 泛洪抑制

操作	命令	说明
进入系统视图	**system-view**	-
进入VSI视图	**vsi** *vsi-name*	-
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
配置VSI选择性泛洪的MAC地址	**selective-flooding mac-addres** *mac-address*	缺省情况下，设备上不存在任何VSI选择性泛洪MAC地址 如果用户只希望某些目的MAC地址的报文可以泛洪到其它站点，可以先通过**flooding disable**命令关闭泛洪功能，再通过本命令配置选择性泛洪的MAC地址

2.10 配置VXLAN报文的目的UDP端口号

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

表2-12 配置 VXLAN 报文的目的 UDP 端口号

操作	命令	说明
进入系统视图	**system-view**	-
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

2.11 配置VXLAN报文检查功能

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。

需要注意的是：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，在这种情况下建议不要执行 **vxlan invalid-vlan-tag discard** 命令，以免错误地丢弃报文。

表2-13 配置 VXLAN 报文检查功能

操作	命令	说明
进入系统视图	**system-view**	-
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
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

2.12 配置ARP泛洪抑制

配置 ARP 泛洪抑制时需要注意：如果同时执行 **flooding disable** 命令关闭了 VSI 的泛洪功能，则建议通过 **mac-address timer** 命令配置动态 MAC 地址的老化时间大于 25 分钟（ARP 泛洪抑制表项的老化时间），以免 MAC 地址在 ARP 泛洪抑制表项老化之前老化，产生黑洞 MAC 地址。

注意

采用组播路由（核心复制）方式转发泛洪流量时：

- 若需要使用 ARP 泛洪抑制功能，必须保证所有 VTEP 设备均开启 ARP 泛洪抑制功能；
- 当需要和其它厂商的 VTEP 设备互通时，不能使用 ARP 泛洪抑制功能。

表2-14 配置 ARP 泛洪抑制

操作	命令	说明
进入系统视图	**system-view**	-
进入VSI视图	**vsi** *vsi-name*	-
开启ARP泛洪抑制功能	**arp suppression enable**	缺省情况下，ARP泛洪抑制功能处于关闭状态

2.13 关闭VXLAN远端ARP自动学习功能

说明

仅 Release 1138P01 及以上版本支持此功能。

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 **vxlan tunnel arp-learning disable** 命令暂时关闭远端 ARP 自动学习功能，以节省占用的设备资源。同步完成后，再执行 **undo vxlan tunnel arp-learning disable** 命令开启远端 ARP 自动学习功能。

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。

表2-15 关闭远端 ARP 自动学习功能

操作	命令	说明
进入系统视图	**system-view**	-
关闭远端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

2.14 配置VXLAN流量统计

2.14.1 配置VSI的报文统计功能

本配置用来开启 VSI 的报文统计功能，用户可以使用 **display l2vpn vsi verbose** 命令查看 VSI 的报文统计信息，使用 **reset l2vpn statistics vsi** 命令清除 VSI 的报文统计信息。

表2-16 配置 VSI 的报文统计功能

操作	命令	说明
进入系统视图	**system-view**	-
配置报文的统计模式为VSI模式	**statistic mode vsi**	缺省情况下，报文的统计模式为VSI模式 多次执行**statistic mode**命令，最后一次执行的命令生效
进入VXLAN所在VSI视图	**vsi** *vsi-name*	-
开启VSI的报文统计功能	**statistics enable**	缺省情况下，VSI的报文统计功能处于关闭状态

2.14.2 配置以太网服务实例的报文统计功能

说明

仅 Release 1138P01 及以上版本支持此功能。

执行本配置后，用户可以使用 **display l2vpn service-instance verbose** 命令查看以太网服务实例的报文统计信息，使用 **reset l2vpn statistics ac** 命令清除以太网服务实例的报文统计信息。

只有为以太网服务实例配置了报文匹配方式并绑定了 VSI 实例，报文统计功能才会生效。如果在报文统计过程中修改报文匹配方式或绑定的 VSI 实例，则重新进行报文统计计数。

表2-17 配置以太网服务实例的报文统计功能

操作		命令	说明
进入系统视图		**system-view**	-
配置报文的统计模式为AC模式		**statistic mode ac**	缺省情况下，报文的统计模式为VSI模式 多次执行**statistic mode queue**命令、**statistic mode ac**命令和**statistic mode vsi**命令，最后一次执行的命令生效 **statistic mode queue**命令的详细介绍，请参见“ACL和QoS命令参考”中的“QoS”
进入二层以太网接口视图或二层聚合接口视图	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*	
进入以太网服务实例视图		**service-instance** *instance-id*	-
开启以太网服务实例的报文统计功能		**statistics enable**	缺省情况下，以太网服务实例的报文统计功能处于关闭状态

2.15 配置VXLAN的硬件资源模式

说明

仅 Release 1138P01 及以上版本支持此功能。

建立 VXLAN 隧道、生成 MAC 地址表项都会占用设备的硬件资源。设备上的硬件资源有限，通过本配置，可以指定硬件资源的分配模式：

- MAC 地址模式：允许建立的 VXLAN 隧道数目较少；允许生成的 MAC 地址表项数目较多。
- Normal 模式：允许建立的 VXLAN 隧道数目较多；允许生成的 MAC 地址表项数目较少。

表2-18 配置 VXLAN 的硬件资源模式

操作	命令	说明
进入系统视图	**system-view**	-
配置VXLAN的硬件资源模式	**hardware-resource vxlan** { **mac** | **normal** }	缺省情况下，VXLAN的硬件资源模式为Normal模式

2.16 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2-19 VXLAN 显示和维护

操作	命令
显示VSI的ARP泛洪抑制表项信息（独立运行模式）	**display arp suppression vsi** [**name** *vsi-name*] [**slot** *slot-number*] [**count**]
显示VSI的ARP泛洪抑制表项信息（IRF模式）	**display arp suppression vsi** [**name vs***i-name*] [**chassis** *chassis-number* **slot** *slot-number*] [**count**]
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IGMP执行主机行为的所有组播组信息	**display igmp host group** [*group-address* | **interface** *interface-type interface-number*] [**verbose**]
显示Tunnel接口信息	**display interface** [**tunnel** [*number*]] [**brief** [**description** | **down**]]
显示VXLAN关联的VXLAN隧道信息	**display vxlan tunnel** [**vxlan-id** *vxlan-id*]
显示设备的报文统计模式	**display statistic mode**
清除VSI的ARP泛洪抑制表项	**reset arp suppression vsi** [**name** *vsi-name*]
清除VSI动态学习的MAC地址表项	**reset l2vpn mac-address** [**vsi** *vsi-name*]
清除VSI的报文统计信息	**reset l2vpn statistics vsi** [**name** *vsi-name*]
清除以太网服务实例的报文统计信息（仅Release 1138P01及以上版本支持此命令）	**reset l2vpn statistics ac** [**interface** *interface-type interface-number* **service-instance** *instance-id*]

2.17 VXLAN典型配置举例

2.17.1 VXLAN头端复制配置举例

1. 组网需求

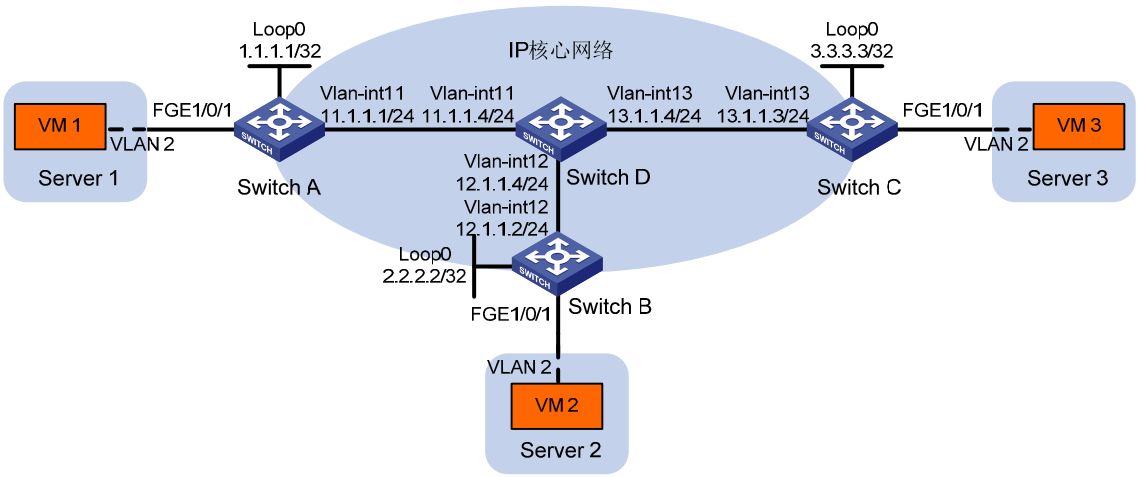
Switch A、Switch B、Switch C 为与服务器连接的 VTEP 设备。虚拟机 VM 1、VM 2 和 VM 3 同属于 VXLAN 10。通过 VXLAN 实现不同站点间的二层互联，确保虚拟机在站点之间进行迁移时用户的访问流量不会中断。

具体需求为：

- 不同 VTEP 之间手工建立 VXLAN 隧道。
- 手工关联 VXLAN 和 VXLAN 隧道。
- 通过源 MAC 地址动态学习远端 MAC 地址表项。
- 站点之间的泛洪流量采用头端复制的方式转发。

2. 组网图

图2-1 VXLAN 头端复制组网图



3. 配置步骤

(1) 配置 IP 地址和单播路由协议

请按照 图 2-1 配置各接口的IP地址和子网掩码，并在IP核心网络内配置OSPF协议，具体配置过程略。

(2) 配置 Switch A

开启 L2VPN 能力。

```
<SwitchA> system-view
[SwitchA] l2vpn enable
```

配置 VXLAN 隧道工作在二层转发模式。

```
[SwitchA] undo vxlan ip-forwarding
```

创建 VSI 实例 vpna 和 VXLAN 10。

```
[SwitchA] vsi vpna
[SwitchA-vsi-vpna] vxlan 10
```

```
[SwitchA-vsi-vpna-vxlan10] quit
[SwitchA-vsi-vpna] quit
```
配置接口 Loopback0 的 IP 地址，作为隧道的源端地址。
```
[SwitchA] interface loopback0
[SwitchA-Loopback0] ip address 1.1.1.1 255.255.255.255
[SwitchA-Loopback0] quit
```
在 Switch A 和 Switch B 之间建立 VXLAN 隧道：

- 创建模式为 VXLAN 的隧道接口 Tunnel1。
- 指定隧道的源端地址为本地接口 Loopback0 的地址 1.1.1.1。
- 指定隧道的目的端地址为 Switch B 上接口 Loopback0 的地址 2.2.2.2。

```
[SwitchA] interface tunnel 1 mode vxlan
[SwitchA-Tunnel1] source 1.1.1.1
[SwitchA-Tunnel1] destination 2.2.2.2
[SwitchA-Tunnel1] quit
```
在 Switch A 和 Switch C 之间建立 VXLAN 隧道。
```
[SwitchA] interface tunnel 2 mode vxlan
[SwitchA-Tunnel2] source 1.1.1.1
[SwitchA-Tunnel2] destination 3.3.3.3
[SwitchA-Tunnel2] quit
```
配置 Tunnel1 和 Tunnel2 与 VXLAN 10 关联。
```
[SwitchA] vsi vpna
[SwitchA-vsi-vpna] vxlan 10
[SwitchA-vsi-vpna-vxlan10] tunnel 1
[SwitchA-vsi-vpna-vxlan10] tunnel 2
[SwitchA-vsi-vpna-vxlan10] quit
[SwitchA-vsi-vpna] quit
```
创建 VLAN2，将 FortyGigE1/0/1 端口加入 VLAN2。
```
[SwitchA] vlan 2
[SwitchA–vlan2] port fortygige 1/0/1
[SwitchA–vlan2] quit
```
在接入服务器的接口 FortyGigE1/0/1 上创建以太网服务实例 1000，该实例用来匹配 VLAN 2 的数据帧。
```
[SwitchA] interface fortygige 1/0/1
[SwitchA-FortyGigE1/0/1] service-instance 1000
[SwitchA-FortyGigE1/0/1-srv1000] encapsulation s-vid 2
```
配置以太网服务实例 1000 与 VSI 实例 vpna 关联。
```
[SwitchA-FortyGigE1/0/1-srv1000] xconnect vsi vpna
[SwitchA-FortyGigE1/0/1-srv1000] quit
[SwitchA-FortyGigE1/0/1] quit
```
(3) 配置 Switch B

开启 L2VPN 能力。
```
<SwitchB> system-view
[SwitchB] l2vpn enable
```
配置 VXLAN 隧道工作在二层转发模式。
```
[SwitchB] undo vxlan ip-forwarding
```
创建 VSI 实例 vpna 和 VXLAN 10。

```
[SwitchB] vsi vpna
[SwitchB-vsi-vpna] vxlan 10
[SwitchB-vsi-vpna-vxlan10] quit
[SwitchB-vsi-vpna] quit
```
配置接口 Loopback0 的 IP 地址，作为隧道的源端地址。
```
[SwitchB] interface loopback0
[SwitchB-Loopback0] ip address 2.2.2.2 255.255.255.255
[SwitchB-Loopback0] quit
```
在 Switch A 和 Switch B 之间建立 VXLAN 隧道。
```
[SwitchB] interface tunnel 2 mode vxlan
[SwitchB-Tunnel2] source 2.2.2.2
[SwitchB-Tunnel2] destination 1.1.1.1
[SwitchB-Tunnel2] quit
```
在 Switch B 和 Switch C 之间建立 VXLAN 隧道。
```
[SwitchB] interface tunnel 3 mode vxlan
[SwitchB-Tunnel3] source 2.2.2.2
[SwitchB-Tunnel3] destination 3.3.3.3
[SwitchB-Tunnel3] quit
```
配置 Tunnel2 和 Tunnel3 与 VXLAN10 关联。
```
[SwitchB] vsi vpna
[SwitchB-vsi-vpna] vxlan 10
[SwitchB-vsi-vpna-vxlan10] tunnel 2
[SwitchB-vsi-vpna-vxlan10] tunnel 3
[SwitchB-vsi-vpna-vxlan10] quit
[SwitchB-vsi-vpna] quit
```
创建 VLAN2，将 FortyGigE1/0/1 端口加入 VLAN2。
```
[SwitchB] vlan 2
[SwitchB–vlan2] port fortygige 1/0/1
[SwitchB–vlan2] quit
```
在接入服务器的接口 FortyGigE1/0/1 上创建以太网服务实例 1000，该实例用来匹配 VLAN 2 的数据帧。
```
[SwitchB] interface fortygige 1/0/1
[SwitchB-FortyGigE1/0/1] service-instance 1000
[SwitchB-FortyGigE1/0/1-srv1000] encapsulation s-vid 2
```
配置以太网服务实例 1000 与 VSI 实例 vpna 关联。
```
[SwitchB-FortyGigE1/0/1-srv1000] xconnect vsi vpna
[SwitchB-FortyGigE1/0/1-srv1000] quit
[SwitchB-FortyGigE1/0/1] quit
```
(4)　配置 Switch C

开启 L2VPN 能力。
```
<SwitchC> system-view
[SwitchC] l2vpn enable
```
配置 VXLAN 隧道工作在二层转发模式。
```
[SwitchC] undo vxlan ip-forwarding
```
创建 VSI 实例 vpna 和 VXLAN 10。
```
[SwitchC] vsi vpna
[SwitchC-vsi-vpna] vxlan 10
```

```
[SwitchC-vsi-vpna-vxlan10] quit
[SwitchC-vsi-vpna] quit
```
配置接口 Loopback0 的 IP 地址，作为隧道的源端地址。
```
[SwitchC] interface loopback0
[SwitchC-Loopback0] ip address 3.3.3.3 255.255.255.255
[SwitchC-Loopback0] quit
```
在 Switch A 和 Switch C 之间建立 VXLAN 隧道。
```
[SwitchC] interface tunnel 1 mode vxlan
[SwitchC-Tunnel1] source 3.3.3.3
[SwitchC-Tunnel1] destination 1.1.1.1
[SwitchC-Tunnel1] quit
```
在 Switch B 和 Switch C 之间建立 VXLAN 隧道。
```
[SwitchC] interface tunnel 3 mode vxlan
[SwitchC-Tunnel3] source 3.3.3.3
[SwitchC-Tunnel3] destination 2.2.2.2
[SwitchC-Tunnel3] quit
```
配置 Tunnel1 和 Tunnel3 与 VXLAN 10 关联。
```
[SwitchC] vsi vpna
[SwitchC-vsi-vpna] vxlan 10
[SwitchC-vsi-vpna-vxlan10] tunnel 1
[SwitchC-vsi-vpna-vxlan10] tunnel 3
[SwitchC-vsi-vpna-vxlan10] quit
[SwitchC-vsi-vpna] quit
```
创建 VLAN2，将 FortyGigE1/0/1 端口加入 VLAN2。
```
[SwitchC] vlan 2
[SwitchC–vlan2] port fortygige 1/0/1
[SwitchC–vlan2] quit
```
在接入服务器的接口 FortyGigE1/0/1 上创建以太网服务实例 1000，该实例用来匹配 VLAN 2 的数据帧。
```
[SwitchC] interface fortygige 1/0/1
[SwitchC-FortyGigE1/0/1] service-instance 1000
[SwitchC-FortyGigE1/0/1-srv1000] encapsulation s-vid 2
```
配置以太网服务实例 1000 与 VSI 实例 vpna 关联。
```
[SwitchC-FortyGigE1/0/1-srv1000] xconnect vsi vpna
[SwitchC-FortyGigE1/0/1-srv1000] quit
[SwitchC-FortyGigE1/0/1] quit
```
4. 验证配置

(1)　验证 VTEP 设备（下文以 Switch A 为例，其它设备验证方法与此类似）

查看 Switch A 上的 Tunnel 接口信息，可以看到 VXLAN 模式的 Tunnel 接口处于 up 状态。
```
[SwitchA] display interface tunnel 1
Tunnel1
Current state: UP
Line protocol state: UP
Description: Tunnel1 Interface
Bandwidth: 64 kbps
Maximum transmission unit: 64000
```

```
Internet protocol processing: Disabled
Last clearing of counters: Never
Tunnel source 1.1.1.1, destination 2.2.2.2
Tunnel protocol/transport UDP_VXLAN/IP
```

查看 Switch A 上的 VSI 信息，可以看到 VSI 内创建的 VXLAN、与 VXLAN 关联的 VXLAN 隧道、与 VSI 关联的以太网服务实例等信息。

```
[SwitchA] display l2vpn vsi verbose
VSI Name: vpna
  VSI Index               : 0
  VSI State               : Up
  MTU                     : 1500
  Bandwidth               : -
  Broadcast Restrain      : -
  Multicast Restrain      : -
  Unknown Unicast Restrain: -
  MAC Learning            : Enabled
  MAC Table Limit         : -
  Drop Unknown            : -
  Flooding                : Enabled
  VXLAN ID                : 10
  Tunnels:
    Tunnel Name          Link ID    State    Type        Flooding proxy
    Tunnel1              0x5000001  Up       Manual      Disabled
    Tunnel2              0x5000002  Up       Manual      Disabled
  ACs:
    AC                              Link ID    State
    FGE1/0/1 srv1000                0          Up
```

查看 Switch A 上 VSI 的 MAC 地址表项信息，可以看到已学习到的 MAC 地址信息。

```
<SwitchA> display l2vpn mac-address
MAC Address      State    VSI Name                        Link ID/Name  Aging
cc3e-5f9c-6cdb   Dynamic  vpna                            Tunnel1       Aging
cc3e-5f9c-23dc   Dynamic  vpna                            Tunnel2       Aging
--- 2 mac address(es) found  ---
```

(2) 验证主机

虚拟机 VM 1、VM 2、VM 3 之间可以互访。

2.17.2 VXLAN核心复制配置举例

1. 组网需求

Switch A、Switch B、Switch C 为与服务器连接的 VTEP 设备。虚拟机 VM 1、VM 2 和 VM 3 同属于 VXLAN 10。通过 VXLAN 实现不同站点间的二层互联，确保虚拟机在站点之间进行迁移时用户的访问流量不会中断。

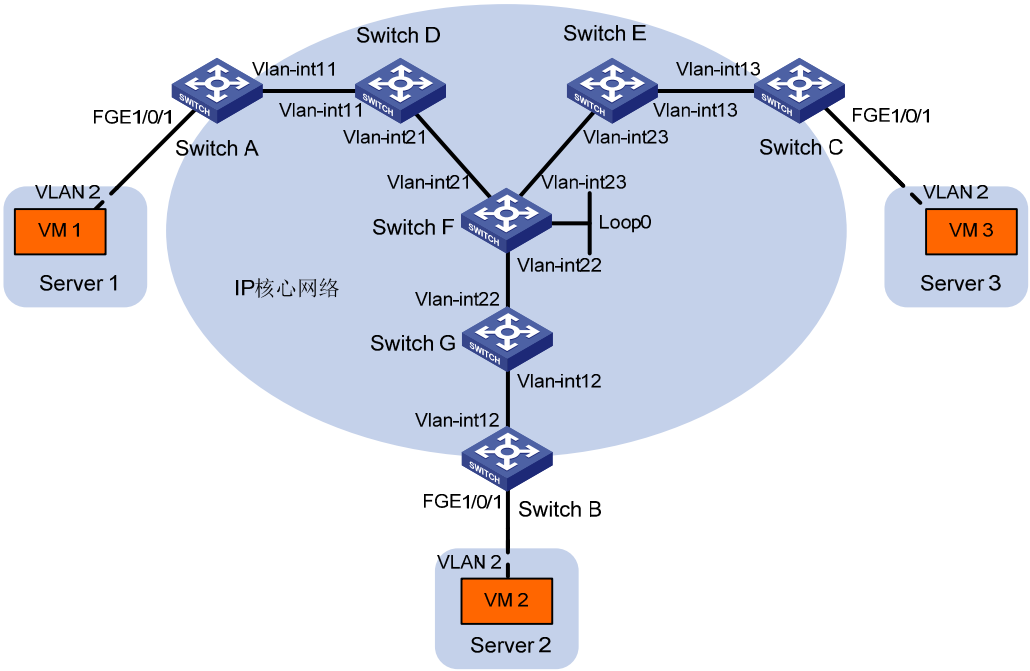
具体需求为：

- 不同 VTEP 之间手工建立 VXLAN 隧道。
- 手工关联 VXLAN 和 VXLAN 隧道。
- 通过源 MAC 地址动态学习远端 MAC 地址表项。

- 站点之间的泛洪流量采用核心复制的方式转发。

2. 组网图

图2-2 VXLAN 核心复制组网图



设备	接口	IP地址	设备	接口	IP地址
Switch A	Vlan-int11	11.1.1.1/24	Switch C	Vlan-int13	13.1.1.3/24
Switch D	Vlan-int11	11.1.1.4/24	Switch E	Vlan-int13	13.1.1.5/24
	Vlan-int21	21.1.1.4/24		Vlan-int23	23.1.1.5/24
Switch F	Vlan-int21	21.1.1.6/24	Switch G	Vlan-int12	12.1.1.7/24
	Vlan-int22	22.1.1.6/24		Vlan-int22	22.1.1.7/24
	Vlan-int23	23.1.1.6/24	Switch B	Vlan-int12	12.1.1.2/24
	Loop0	6.6.6.6/32			

3. 配置步骤

(1) 配置 IP 地址和单播路由协议

请按照 图 2-2 配置各接口的IP地址和子网掩码，并在IP核心网络内配置OSPF协议，具体配置过程略。

(2) 配置 Switch A

开启 L2VPN 能力。

```
<SwitchA> system-view
[SwitchA] l2vpn enable
```

配置 VXLAN 隧道工作在二层转发模式。

```
[SwitchA] undo vxlan ip-forwarding
```

使能 IP 组播路由。

```
[SwitchA] multicast routing
[SwitchA-mrib] quit
```

创建 VSI 实例 vpna 和 VXLAN 10。

```
[SwitchA] vsi vpna
[SwitchA-vsi-vpna] vxlan 10
```

```
[SwitchA-vsi-vpna-vxlan10] quit
[SwitchA-vsi-vpna] quit
```

配置接口 Vlan-interface11 的 IP 地址，并在该接口上使能 IGMP 协议的主机功能。

```
[SwitchA] interface vlan-interface 11
[SwitchA-Vlan-interface11] ip address 11.1.1.1 24
[SwitchA-Vlan-interface11] igmp host enable
[SwitchA-Vlan-interface11] quit
```

在 Switch A 和 Switch B 之间建立 VXLAN 隧道：

- 创建模式为 VXLAN 的隧道接口 Tunnel1。
- 指定隧道的源端地址为本地接口 Vlan-interface11 的地址 11.1.1.1。
- 指定隧道的目的端地址为 Switch B 上接口 Vlan-interface12 的地址 12.1.1.2。

```
[SwitchA] interface tunnel 1 mode vxlan
[SwitchA-Tunnel1] source 11.1.1.1
[SwitchA-Tunnel1] destination 12.1.1.2
[SwitchA-Tunnel1] quit
```

在 Switch A 和 Switch C 之间建立 VXLAN 隧道。

```
[SwitchA] interface tunnel 2 mode vxlan
[SwitchA-Tunnel2] source 11.1.1.1
[SwitchA-Tunnel2] destination 13.1.1.3
[SwitchA-Tunnel2] quit
```

配置 Tunnel1 和 Tunnel2 与 VXLAN 10 关联。

```
[SwitchA] vsi vpna
[SwitchA-vsi-vpna] vxlan 10
[SwitchA-vsi-vpna-vxlan10] tunnel 1
[SwitchA-vsi-vpna-vxlan10] tunnel 2
```

配置 VXLAN 泛洪的组播地址为 225.1.1.1，组播报文的源 IP 地址为 11.1.1.1。

```
[SwitchA-vsi-vpna-vxlan10] group 225.1.1.1 source 11.1.1.1
[SwitchA-vsi-vpna-vxlan10] quit
[SwitchA-vsi-vpna] quit
```

创建 VLAN2，将 FortyGigE1/0/1 端口加入 VLAN2。

```
[SwitchA] vlan 2
[SwitchA–vlan2] port fortygige 1/0/1
[SwitchA–vlan2] quit
```

在接入服务器的接口 FortyGigE1/0/1 上创建以太网服务实例 1000，该实例用来匹配 VLAN 2 的数据帧。

```
[SwitchA] interface fortygige 1/0/1
[SwitchA-FortyGigE1/0/1] service-instance 1000
[SwitchA-FortyGigE1/0/1-srv1000] encapsulation s-vid 2
```

配置以太网服务实例 1000 与 VSI 实例 vpna 关联。

```
[SwitchA-FortyGigE1/0/1-srv1000] xconnect vsi vpna
[SwitchA-FortyGigE1/0/1-srv1000] quit
[SwitchA-FortyGigE1/0/1] quit
```

(3) 配置 Switch B

开启 L2VPN 能力。

```
<SwitchB> system-view
```

```
[SwitchB] l2vpn enable
```
配置 VXLAN 隧道工作在二层转发模式。
```
[SwitchB] undo vxlan ip-forwarding
```
使能 IP 组播路由。
```
[SwitchB] multicast routing
[SwitchB-mrib] quit
```
创建 VSI 实例 vpna 和 VXLAN 10。
```
[SwitchB] vsi vpna
[SwitchB-vsi-vpna] vxlan 10
[SwitchB-vsi-vpna-vxlan10] quit
[SwitchB-vsi-vpna] quit
```
配置接口 Vlan-interface12 的 IP 地址，并在该接口上使能 IGMP 协议的主机功能。
```
[SwitchB] interface vlan-interface 12
[SwitchB-Vlan-interface12] ip address 12.1.1.2 24
[SwitchB-Vlan-interface12] igmp host enable
[SwitchB-Vlan-interface12] quit
```
在 Switch A 和 Switch B 之间建立 VXLAN 隧道。
```
[SwitchB] interface tunnel 2 mode vxlan
[SwitchB-Tunnel2] source 12.1.1.2
[SwitchB-Tunnel2] destination 11.1.1.1
[SwitchB-Tunnel2] quit
```
在 Switch B 和 Switch C 之间建立 VXLAN 隧道。
```
[SwitchB] interface tunnel 3 mode vxlan
[SwitchB-Tunnel3] source 12.1.1.2
[SwitchB-Tunnel3] destination 13.1.1.3
[SwitchB-Tunnel3] quit
```
配置 Tunnel2 和 Tunnel3 与 VXLAN10 关联。
```
[SwitchB] vsi vpna
[SwitchB-vsi-vpna] vxlan 10
[SwitchB-vsi-vpna-vxlan10] tunnel 2
[SwitchB-vsi-vpna-vxlan10] tunnel 3
```
配置 VXLAN 泛洪的组播地址为 225.1.1.1，组播报文的源 IP 地址为 12.1.1.2。
```
[SwitchB-vsi-vpna-vxlan10] group 225.1.1.1 source 12.1.1.2
[SwitchB-vsi-vpna-vxlan10] quit
[SwitchB-vsi-vpna] quit
```
创建 VLAN2，将 FortyGigE1/0/1 端口加入 VLAN2。
```
[SwitchB] vlan 2
[SwitchB–vlan2] port fortygige 1/0/1
[SwitchB–vlan2] quit
```
在接入服务器的接口 FortyGigE1/0/1 上创建以太网服务实例 1000，该实例用来匹配 VLAN 2 的数据帧。
```
[SwitchB] interface fortygige 1/0/1
[SwitchB-FortyGigE1/0/1] service-instance 1000
[SwitchB-FortyGigE1/0/1-srv1000] encapsulation s-vid 2
```
配置以太网服务实例 1000 与 VSI 实例 vpna 关联。

```
[SwitchB-FortyGigE1/0/1-srv1000] xconnect vsi vpna
[SwitchB-FortyGigE1/0/1-srv1000] quit
[SwitchB-FortyGigE1/0/1] quit
```

(4) 配置 Switch C

开启 L2VPN 能力。

```
<SwitchC> system-view
[SwitchC] l2vpn enable
```

配置 VXLAN 隧道工作在二层转发模式。

```
[SwitchC] undo vxlan ip-forwarding
```

使能 IP 组播路由。

```
[SwitchC] multicast routing
[SwitchC-mrib] quit
```

创建 VSI 实例 vpna 和 VXLAN 10。

```
[SwitchC] vsi vpna
[SwitchC-vsi-vpna] vxlan 10
[SwitchC-vsi-vpna-vxlan10] quit
[SwitchC-vsi-vpna] quit
```

配置接口 Vlan-interface13 的 IP 地址，并在该接口上使能 IGMP 协议的主机功能。

```
[SwitchC] interface vlan-interface 13
[SwitchC-Vlan-interface13] ip address 13.1.1.3 24
[SwitchC-Vlan-interface13] igmp host enable
[SwitchC-Vlan-interface13] quit
```

在 Switch A 和 Switch C 之间建立 VXLAN 隧道。

```
[SwitchC] interface tunnel 1 mode vxlan
[SwitchC-Tunnel1] source 13.1.1.3
[SwitchC-Tunnel1] destination 11.1.1.1
[SwitchC-Tunnel1] quit
```

在 Switch B 和 Switch C 之间建立 VXLAN 隧道。

```
[SwitchC] interface tunnel 3 mode vxlan
[SwitchC-Tunnel3] source 13.1.1.3
[SwitchC-Tunnel3] destination 12.1.1.2
[SwitchC-Tunnel3] quit
```

配置 Tunnel1 和 Tunnel3 与 VXLAN 10 关联。

```
[SwitchC] vsi vpna
[SwitchC-vsi-vpna] vxlan 10
[SwitchC-vsi-vpna-vxlan10] tunnel 1
[SwitchC-vsi-vpna-vxlan10] tunnel 3
```

配置 VXLAN 泛洪的组播地址为 225.1.1.1，组播报文的源 IP 地址为 13.1.1.3。

```
[SwitchC-vsi-vpna-vxlan10] group 225.1.1.1 source 13.1.1.3
[SwitchC-vsi-vpna-vxlan10] quit
[SwitchC-vsi-vpna] quit
```

创建 VLAN2，将 FortyGigE1/0/1 端口加入 VLAN2。

```
[SwitchC] vlan 2
[SwitchC–vlan2] port fortygige 1/0/1
[SwitchC–vlan2] quit
```

在接入服务器的接口 FortyGigE1/0/1 上创建以太网服务实例 1000，该实例用来匹配 VLAN 2 的数据帧。

```
[SwitchC] interface fortygige 1/0/1
[SwitchC-FortyGigE1/0/1] service-instance 1000
[SwitchC-FortyGigE1/0/1-srv1000] encapsulation s-vid 2
```

配置以太网服务实例 1000 与 VSI 实例 vpna 关联。

```
[SwitchC-FortyGigE1/0/1-srv1000] xconnect vsi vpna
[SwitchC-FortyGigE1/0/1-srv1000] quit
[SwitchC-FortyGigE1/0/1] quit
```

(5) 配置 Switch D

使能 IP 组播路由。

```
<SwitchD> system-view
[SwitchD] multicast routing
[SwitchD-mrib] quit
```

在接口 Vlan-interface11 上使能 IGMP 和 PIM-SM。

```
[SwitchD] interface vlan-interface 11
[SwitchD-Vlan-interface11] igmp enable
[SwitchD-Vlan-interface11] pim sm
[SwitchD-Vlan-interface11] quit
```

在接口 Vlan-interface21 上使能 PIM-SM。

```
[SwitchD] interface vlan-interface 21
[SwitchD-Vlan-interface21] pim sm
[SwitchD-Vlan-interface21] quit
```

(6) 配置 Switch E

使能 IP 组播路由。

```
<SwitchE> system-view
[SwitchE] multicast routing
[SwitchE-mrib] quit
```

在接口 Vlan-interface13 上使能 IGMP 和 PIM-SM。

```
[SwitchE] interface vlan-interface 13
[SwitchE-Vlan-interface13] igmp enable
[SwitchE-Vlan-interface13] pim sm
[SwitchE-Vlan-interface13] quit
```

在接口 Vlan-interface23 上使能 PIM-SM。

```
[SwitchE] interface vlan-interface 23
[SwitchE-Vlan-interface23] pim sm
[SwitchE-Vlan-interface23] quit
```

(7) 配置 Switch F

使能 IP 组播路由。

```
<SwitchF> system-view
[SwitchF] multicast routing
[SwitchF-mrib] quit
```

在各接口上使能 PIM-SM。

```
[SwitchF] interface vlan-interface 21
[SwitchF-Vlan-interface21] pim sm
```

```
[SwitchF-Vlan-interface21] quit
[SwitchF] interface vlan-interface 22
[SwitchF-Vlan-interface22] pim sm
[SwitchF-Vlan-interface22] quit
[SwitchF] interface vlan-interface 23
[SwitchF-Vlan-interface23] pim sm
[SwitchF-Vlan-interface23] quit
```

(8) 配置 Switch G

使能 IP 组播路由。

```
<SwitchG> system-view
[SwitchG] multicast routing
[SwitchG-mrib] quit
```

在接口 Vlan-interface12 上使能 IGMP 和 PIM-SM。

```
[SwitchG] interface vlan-interface 12
[SwitchG-Vlan-interface12] igmp enable
[SwitchG-Vlan-interface12] pim sm
[SwitchG-Vlan-interface12] quit
```

在接口 Vlan-interface22 上使能 PIM-SM。

```
[SwitchG] interface vlan-interface 22
[SwitchG-Vlan-interface22] pim sm
[SwitchG-Vlan-interface22] quit
```

4. 验证配置

(1) 验证 VTEP 设备（下文以 Switch A 为例，其它设备验证方法与此类似）

查看 Switch A 上的 Tunnel 接口信息，可以看到 VXLAN 模式的 Tunnel 接口处于 up 状态。

```
[SwitchA] display interface tunnel 1
Tunnel1
Current state: UP
Line protocol state: UP
Description: Tunnel1 Interface
Bandwidth: 64 kbps
Maximum transmission unit: 64000
Internet protocol processing: Disabled
Last clearing of counters: Never
Tunnel source 11.1.1.1, destination 12.1.1.2
Tunnel protocol/transport UDP_VXLAN/IP
```

查看 Switch A 上的 VSI 信息，可以看到 VSI 内创建的 VXLAN、与 VXLAN 关联的 VXLAN 隧道、与 VSI 关联的以太网服务实例等信息。

```
[SwitchA] display l2vpn vsi verbose
VSI Name: vpna
  VSI Index               : 0
  VSI State               : Up
  MTU                     : 1500
  Bandwidth               : -
  Broadcast Restrain      : -
  Multicast Restrain      : -
  Unknown Unicast Restrain: -
```

```
  MAC Learning            : Enabled
  MAC Table Limit         : -
  Drop Unknown            : -
  Flooding                : Enabled
  VXLAN ID                : 10
  Tunnels:
    Tunnel Name          Link ID    State    Type        Flooding proxy
    Tunnel1              0x5000001  Up       Manual      Disabled
    Tunnel2              0x5000002  Up       Manual      Disabled
    MTunnel0             0x6000000  Up       Auto        Disabled
  ACs:
    AC                               Link ID    State
    FGE1/0/1 srv1000                 0          Up
```

查看 Switch A 上 VSI 的 MAC 地址表项信息，可以看到已学习到的 MAC 地址信息。

```
<SwitchA> display l2vpn mac-address
MAC Address      State    VSI Name                        Link ID/Name  Aging
cc3e-5f9c-6cdb   Dynamic  vpna                            Tunnel1       Aging
cc3e-5f9c-23dc   Dynamic  vpna                            Tunnel2       Aging
--- 2 mac address(es) found  ---
```

查看 Switch A 上 IGMP 执行主机行为的所有组播组信息，可以看到接口 Vlan-interface11 下存在组播组 225.1.1.1 的信息。

```
<SwitchA> display igmp host group
IGMP host groups in total: 1
 Vlan-interface11(11.1.1.1):
  IGMP host groups in total: 1
   Group address      Member state      Expires
   225.1.1.1          Idle              Off
```

(2) 验证主机

虚拟机 VM 1、VM 2、VM 3 之间可以互访。

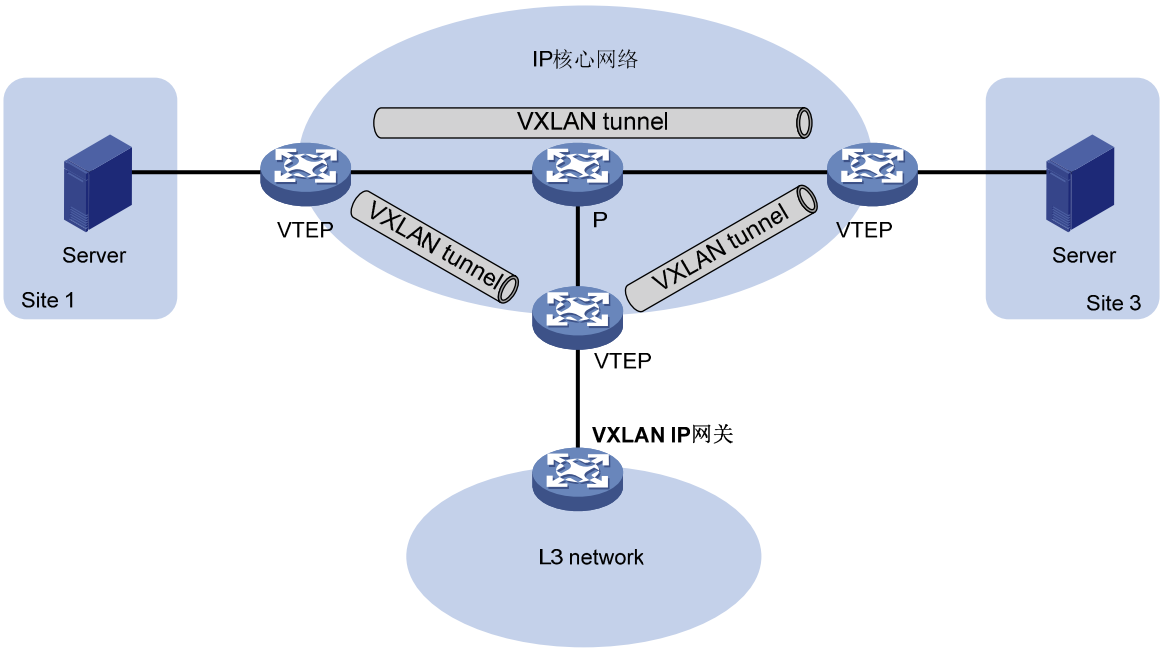
3 VXLAN IP网关

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。

3.1.1 独立的VXLAN IP网关

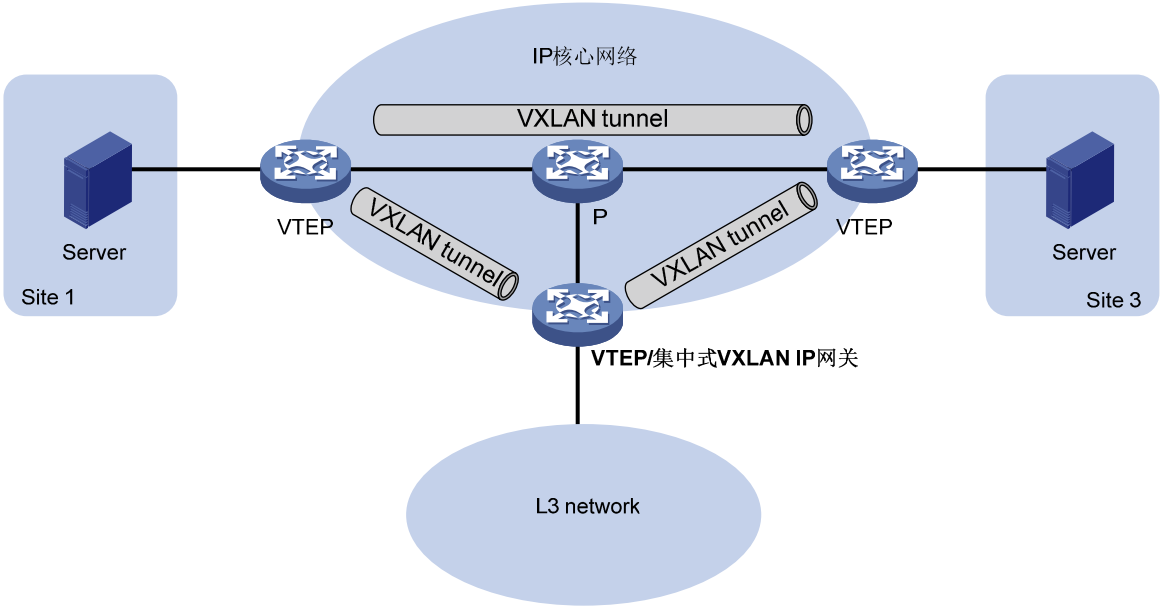
图3-1 独立的 VXLAN IP 网关示意图



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。

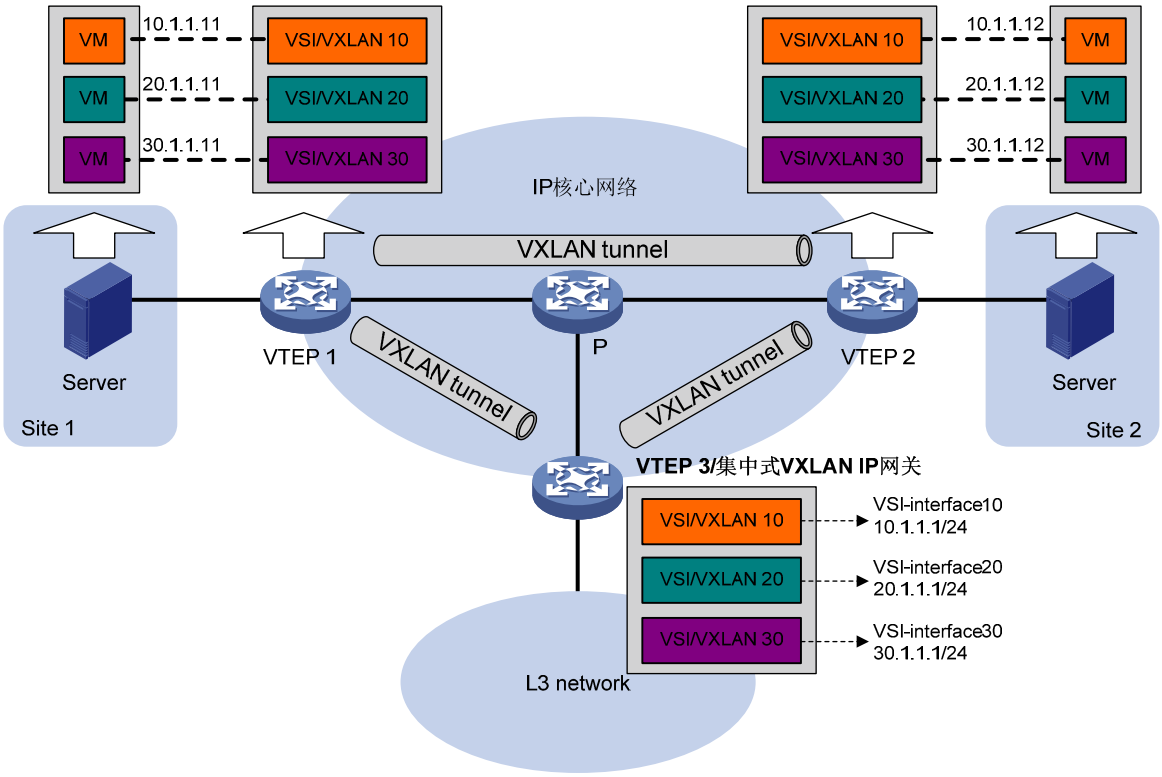
3.1.2 集中式VXLAN IP网关

图3-2 集中式 VXLAN IP 网关示意图



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3-3 集中式 VXLAN IP 网关的三层通信过程



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。

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。

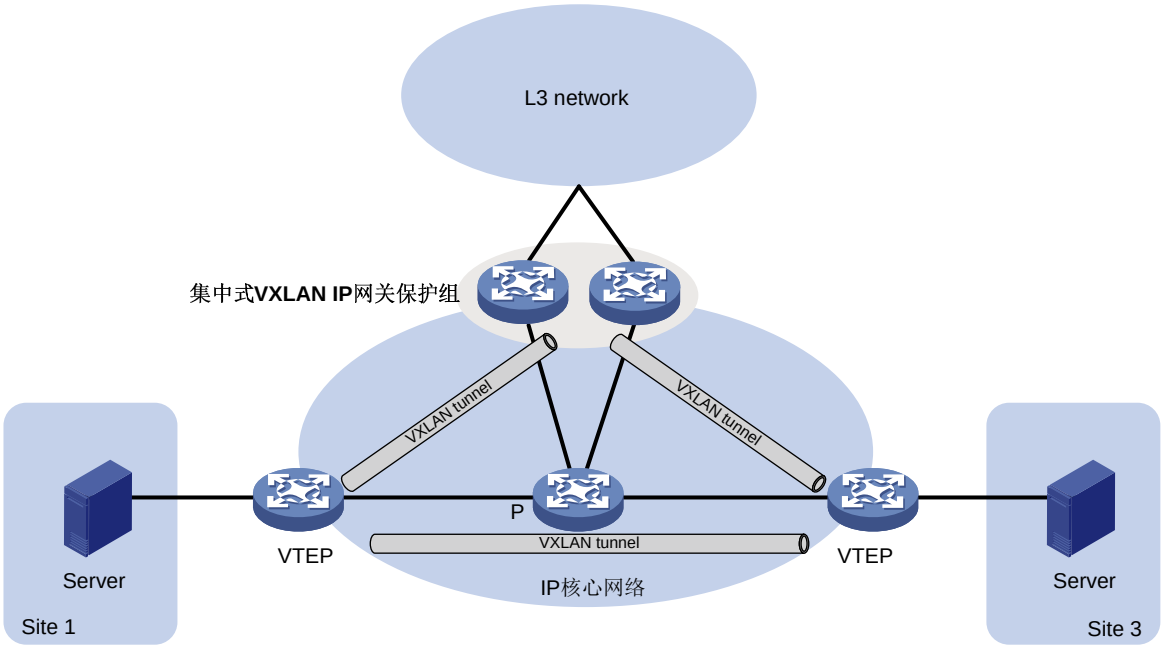
3.1.3 集中式VXLAN IP网关保护组

说明

仅 Release 1138P01 及以上版本支持此功能。

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。

3.2 配置限制和指导

配置集中式 VXLAN IP 网关时，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。
- 配置 VXLAN 隧道，并将 VXLAN 与 VXLAN 隧道关联。

3.3 配置集中式VXLAN IP网关

表3-1 配置集中式 VXLAN IP 网关

操作	命令	说明
进入系统视图	**system-view**	-
创建VSI虚接口，并进入VSI虚接口视图	**interface vsi-interface** *vsi-interface-id*	缺省情况下，设备上不存在任何VSI虚接口 如果VSI虚接口已经存在，则直接进入该VSI虚接口视图
配置VSI虚接口的IP地址	**ip address** *ip-address* { *mask* | *mask-length* } [**sub**]	缺省情况下，未配置VSI虚接口的IP地址
退回系统视图	**quit**	-
进入VXLAN所在VSI视图	**vsi** *vsi-name*	-
为VSI指定网关接口	**gateway vsi-interface** *vsi-interface-id*	缺省情况下，没有为VSI指定网关接口

3.4 配置集中式VXLAN IP网关保护组

说明

仅 Release 1138P01 及以上版本支持此功能。

3.4.1 VXLAN IP网关上的配置

保护组中所有网关上的 VXLAN 配置需要保证完全一致。

表3-2 配置集中式 VXLAN IP 网关保护组

操作	命令	说明
进入系统视图	**system-view**	-
创建VSI虚接口，并进入VSI虚接口视图	**interface vsi-interface** *vsi-interface-id*	缺省情况下，设备上不存在任何VSI虚接口 如果VSI虚接口已经存在，则直接进入该VSI虚接口视图 请在保护组中的每台网关上创建相同的VSI虚接口
配置VSI虚接口的IP地址	**ip address** *ip-address* { *mask* | *mask-length* } [**sub**]	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

操作	命令	说明
配置VSI虚接口的MAC地址	**mac-address** *mac-address*	缺省情况下，VSI虚接口的MAC地址为设备上VLAN接口的MAC地址 保护组中所有网关上配置的MAC地址必须相同
退回系统视图	**quit**	-
进入VXLAN所在VSI视图	**vsi** *vsi-name*	-
为VSI指定网关接口	**gateway vsi-interface** *vsi-interface-id*	缺省情况下，没有为VSI指定网关接口
退回系统视图	**quit**	-
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，设备没有加入VTEP IP网关保护组 在保护组中每台VXLAN IP网关中执行此配置，*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
配置VTEP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，设备上没有配置VTEP保护组的成员地址列表 在保护组中每台VXLAN IP网关中执行此配置，必须输入保护组中所有其它成员的成员地址

3.4.2 接入层VTEP上的配置

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。

表3-3 配置接入层 VTEP 识别 VXLAN IP 网关保护组

操作	命令	说明
进入系统视图	**system-view**	-
配置VTEP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，设备上没有配置VTEP保护组的成员地址列表 必须输入保护组中所有成员的成员地址

3.5 配置VSI虚接口

通过本配置，可以根据需要调整 VSI 虚接口的参数及状态。

表3-4 配置 VSI 虚接口

操作	命令	说明
进入系统视图	**system-view**	-
进入VSI虚接口视图	**interface vsi** *vsi-interface-id*	-
配置VSI虚接口的MAC地址	**mac-address** *mac-address*	缺省情况下，VSI虚接口的MAC地址为设备上VLAN接口的MAC地址 如果VSI虚接口MAC地址的高36位与VXLAN IP网关桥MAC地址的高36位不同，则VXLAN IP网关发送报文的源MAC地址是VSI虚接口的缺省MAC地址（VLAN接口的MAC地址）；如果二者相同，则源MAC地址为VSI虚接口的MAC地址
开启ARP报文发送限速功能	**arp send-rate** *pps*	缺省情况下，ARP报文发送限速功能处于关闭状态 仅Release 1138P01及以上版本支持此命令
（可选）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
（可选）配置接口的MTU	**mtu** *mtu-value*	缺省情况下，接口的MTU值为1500字节
（可选）配置接口的期望带宽	**bandwidth** *bandwidth-value*	缺省情况下，接口的期望带宽为1000000kbps
恢复当前接口的缺省配置	**default**	-
开启当前接口	**undo shutdown**	缺省情况下，接口处于开启状态

3.6 开启VSI虚接口的报文统计功能

执行本配置后，设备会开启与 VSI 关联的 VSI 虚接口的报文统计功能。通过 **display interface vsi-interface** 命令的 input 和 output 字段可以查看统计信息。

需要注意的是：

- 只有 VSI 虚接口仅与一个 VSI 关联时，才能开启该 VSI 虚接口的报文统计功能。
- 多次执行 **statistic mode** 命令，最后一次执行的命令生效。

表3-5 开启 VSI 虚接口的报文统计功能

操作	命令	说明
进入系统视图	**system-view**	-
配置报文的统计模式为VSI模式	**statistic mode vsi**	缺省情况下，报文的统计模式为VSI模式
进入VXLAN所在VSI视图	**vsi** *vsi-name*	-
开启VSI的报文统计功能	**statistics enable**	缺省情况下，VSI的报文统计功能处于关闭状态

3.7 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3-6 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** [*vsi-interface-id*]] [**brief** [**description**]]
清除接口的统计信息	**reset counters interface** [**vsi-interface** [*vsi-interface-id*]]

3.8 VXLAN IP网关典型配置举例

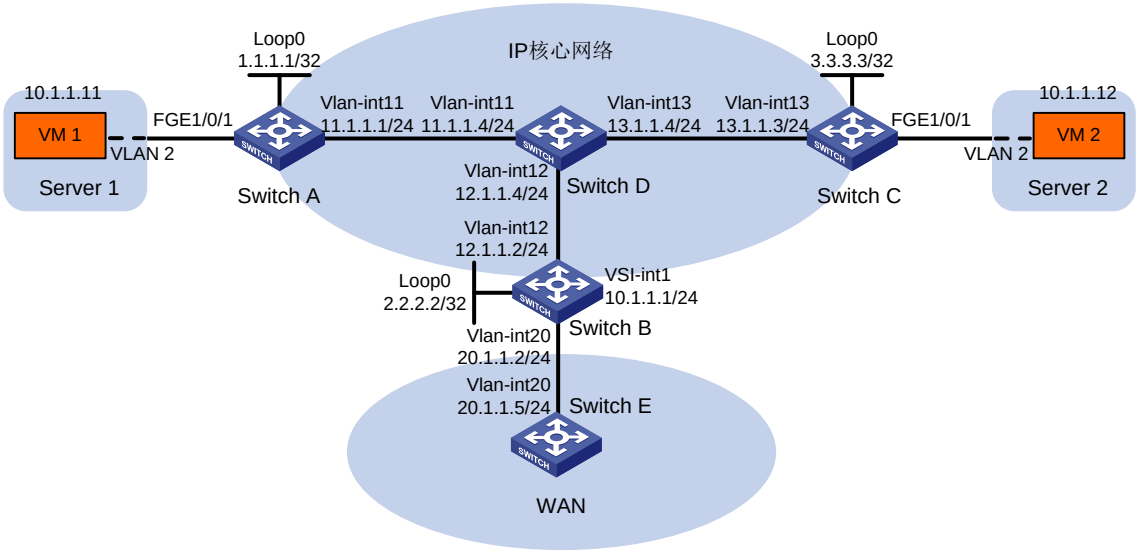
3.8.1 集中式VXLAN IP网关配置举例

1. 组网需求

Switch A、Switch C 为与服务器连接的 VTEP 设备，Switch B 为与广域网连接的集中式 VXLAN IP 网关设备，Switch E 为广域网内的三层交换机。虚拟机 VM 1、VM 2 同属于 VXLAN 10，通过 VXLAN 实现不同站点间的二层互联，并通过 VXLAN IP 网关与广域网三层互联。

2. 组网图

图3-5 集中式 VXLAN IP 网关配置组网图



3. 配置步骤

(1) 配置 IP 地址和单播路由协议

请按照 图 3-5 配置各接口的IP地址和子网掩码，并在IP核心网络内配置OSPF协议，具体配置过程略。

Switch B 和 Switch E 上配置 OSPF 协议，发布 10.1.1.0/24 和 20.1.1.0/24 网段的路由，具体配置过程略。

(2) 配置 Switch A

开启 L2VPN 能力。

```
<SwitchA> system-view
[SwitchA] l2vpn enable
```

配置 VXLAN 隧道工作在二层转发模式。

```
[SwitchA] undo vxlan ip-forwarding
```

创建 VSI 实例 vpna 和 VXLAN 10。

```
[SwitchA] vsi vpna
[SwitchA-vsi-vpna] vxlan 10
[SwitchA-vsi-vpna-vxlan10] quit
[SwitchA-vsi-vpna] quit
```

配置接口 Loopback0 的 IP 地址，作为隧道的源端地址。

```
[SwitchA] interface loopback0
[SwitchA-Loopback0] ip address 1.1.1.1 255.255.255.255
[SwitchA-Loopback0] quit
```

在 Switch A 和 Switch B 之间建立 VXLAN 隧道：

- 创建模式为 VXLAN 的隧道接口 Tunnel1。
- 指定隧道的源端地址为本地接口 Loopback0 的地址 1.1.1.1。
- 指定隧道的目的端地址为 Switch B 上接口 Loopback0 的地址 2.2.2.2。

```
[SwitchA] interface tunnel 1 mode vxlan
[SwitchA-Tunnel1] source 1.1.1.1
[SwitchA-Tunnel1] destination 2.2.2.2
[SwitchA-Tunnel1] quit
```

在 Switch A 和 Switch C 之间建立 VXLAN 隧道。

```
[SwitchA] interface tunnel 2 mode vxlan
[SwitchA-Tunnel2] source 1.1.1.1
[SwitchA-Tunnel2] destination 3.3.3.3
[SwitchA-Tunnel2] quit
```

配置 Tunnel1 和 Tunnel2 与 VXLAN 10 关联。

```
[SwitchA] vsi vpna
[SwitchA-vsi-vpna] vxlan 10
[SwitchA-vsi-vpna-vxlan10] tunnel 1
[SwitchA-vsi-vpna-vxlan10] tunnel 2
[SwitchA-vsi-vpna-vxlan10] quit
[SwitchA-vsi-vpna] quit
```

创建 VLAN2，将 FortyGigE1/0/1 端口加入 VLAN2。

```
[SwitchA] vlan 2
[SwitchA–vlan2] port fortygige 1/0/1
[SwitchA–vlan2] quit
```

在接入服务器的接口 FortyGigE1/0/1 上创建以太网服务实例 1000，该实例用来匹配 VLAN 2 的数据帧。

```
[SwitchA] interface fortygige 1/0/1
[SwitchA-FortyGigE1/0/1] service-instance 1000
[SwitchA-FortyGigE1/0/1-srv1000] encapsulation s-vid 2
```
配置以太网服务实例 1000 与 VSI 实例 vpna 关联。
```
[SwitchA-FortyGigE1/0/1-srv1000] xconnect vsi vpna
[SwitchA-FortyGigE1/0/1-srv1000] quit
[SwitchA-FortyGigE1/0/1] quit
```
(3) 配置 Switch B

开启 L2VPN 能力。
```
<SwitchB> system-view
[SwitchB] l2vpn enable
```
配置预留 VLAN 3000 接口资源（全局类型）。
```
[SwitchB] reserve-vlan-interface 3000 global
```
创建 VSI 实例 vpna 和 VXLAN 10。
```
[SwitchB] vsi vpna
[SwitchB-vsi-vpna] vxlan 10
[SwitchB-vsi-vpna-vxlan10] quit
[SwitchB-vsi-vpna] quit
```
配置接口 Loopback0 的 IP 地址，作为隧道的源端地址。
```
[SwitchB] interface loopback0
[SwitchB-Loopback0] ip address 2.2.2.2 255.255.255.255
[SwitchB-Loopback0] quit
```
在 Switch A 和 Switch B 之间建立 VXLAN 隧道。
```
[SwitchB] interface tunnel 2 mode vxlan
[SwitchB-Tunnel2] source 2.2.2.2
[SwitchB-Tunnel2] destination 1.1.1.1
[SwitchB-Tunnel2] quit
```
在 Switch B 和 Switch C 之间建立 VXLAN 隧道。
```
[SwitchB] interface tunnel 3 mode vxlan
[SwitchB-Tunnel3] source 2.2.2.2
[SwitchB-Tunnel3] destination 3.3.3.3
[SwitchB-Tunnel3] quit
```
配置 Tunnel2 和 Tunnel3 与 VXLAN10 关联。
```
[SwitchB] vsi vpna
[SwitchB-vsi-vpna] vxlan 10
[SwitchB-vsi-vpna-vxlan10] tunnel 2
[SwitchB-vsi-vpna-vxlan10] tunnel 3
[SwitchB-vsi-vpna-vxlan10] quit
[SwitchB-vsi-vpna] quit
```
创建 VSI 虚接口 VSI-interface1，并为其配置 IP 地址，该 IP 地址作为 VXLAN 10 内虚拟机的网关地址。
```
[SwitchB] interface vsi-interface 1
[SwitchB-Vsi-interface1] ip address 10.1.1.1 255.255.255.0
[SwitchB-Vsi-interface1] quit
```
配置 VXLAN 10 所在的 VSI 实例和接口 VSI-interface1 关联。
```
[SwitchB] vsi vpna
[SwitchB-vsi-vpna] gateway vsi-interface 1
```

```
[SwitchB-vsi-vpna] quit
```

(4) 配置 Switch C

开启 L2VPN 能力。

```
<SwitchC> system-view
[SwitchC] l2vpn enable
```

配置 VXLAN 隧道工作在二层转发模式。

```
[SwitchC] undo vxlan ip-forwarding
```

创建 VSI 实例 vpna 和 VXLAN 10。

```
[SwitchC] vsi vpna
[SwitchC-vsi-vpna] vxlan 10
[SwitchC-vsi-vpna-vxlan10] quit
[SwitchC-vsi-vpna] quit
```

配置接口 Loopback0 的 IP 地址，作为隧道的源端地址。

```
[SwitchC] interface loopback0
[SwitchC-Loopback0] ip address 3.3.3.3 255.255.255.255
[SwitchC-Loopback0] quit
```

在 Switch A 和 Switch C 之间建立 VXLAN 隧道。

```
[SwitchC] interface tunnel 1 mode vxlan
[SwitchC-Tunnel1] source 3.3.3.3
[SwitchC-Tunnel1] destination 1.1.1.1
[SwitchC-Tunnel1] quit
```

在 Switch B 和 Switch C 之间建立 VXLAN 隧道。

```
[SwitchC] interface tunnel 3 mode vxlan
[SwitchC-Tunnel3] source 3.3.3.3
[SwitchC-Tunnel3] destination 2.2.2.2
[SwitchC-Tunnel3] quit
```

配置 Tunnel1 和 Tunnel3 与 VXLAN 10 关联。

```
[SwitchC] vsi vpna
[SwitchC-vsi-vpna] vxlan 10
[SwitchC-vsi-vpna-vxlan10] tunnel 1
[SwitchC-vsi-vpna-vxlan10] tunnel 3
[SwitchC-vsi-vpna-vxlan10] quit
[SwitchC-vsi-vpna] quit
```

创建 VLAN2，将 FortyGigE1/0/1 端口加入 VLAN2。

```
[SwitchC] vlan 2
[SwitchC–vlan2] port fortygige 1/0/1
[SwitchC–vlan2] quit
```

在接入服务器的接口 FortyGigE1/0/1 上创建以太网服务实例 1000，该实例用来匹配 VLAN 2 的数据帧。

```
[SwitchC] interface fortygige 1/0/1
[SwitchC-FortyGigE1/0/1] service-instance 1000
[SwitchC-FortyGigE1/0/1-srv1000] encapsulation s-vid 2
```

配置以太网服务实例 1000 与 VSI 实例 vpna 关联。

```
[SwitchC-FortyGigE1/0/1-srv1000] xconnect vsi vpna
[SwitchC-FortyGigE1/0/1-srv1000] quit
[SwitchC-FortyGigE1/0/1] quit
```

4. 验证配置

(1) 验证 VXLAN IP 网关设备 Switch B

查看 Switch B 上的 Tunnel 接口信息，可以看到 VXLAN 模式的 Tunnel 接口处于 up 状态。

```
[SwitchB] display interface tunnel 2
Tunnel2
Current state: UP
Line protocol state: UP
Description: Tunnel1 Interface
Bandwidth: 64 kbps
Maximum transmission unit: 64000
Internet protocol processing: Disabled
Last clearing of counters: Never
Tunnel source 2.2.2.2, destination 1.1.1.1
Tunnel protocol/transport UDP_VXLAN/IP
```

查看 Switch B 上的 VSI 虚接口信息，可以看到 VSI 虚接口处于 up 状态。

```
[SwitchB] display interface vsi-interface 1
Vsi-interface1
Current state: UP
Line protocol state: UP
Description: Vsi-interface100 Interface
Bandwidth: 1000000kbps
Maximum transmission unit: 1500
Internet Address is 10.1.1.1/24 Primary
IP Packet Frame Type:PKTFMT_ETHNT_2, Hardware Address: 0011-2200-0102
IPv6 Packet Frame Type:PKTFMT_ETHNT_2, Hardware Address: 0011-2200-0102
Physical: Unknown, baudrate: 1000000 kbps
Last clearing of counters: Never
Last 300 seconds input rate: 0 bytes/sec, 0 bits/sec, 0 packets/sec
Last 300 seconds output rate: 0 bytes/sec, 0 bits/sec, 0 packets/sec
Input: 0 packets, 0 bytes, 0 drops
Output: 0 packets, 0 bytes, 0 drops
```

查看 Switch B 上的 VSI 信息，可以看到 VSI 内创建的 VXLAN、与 VXLAN 关联的 VXLAN 隧道、与 VSI 关联的 VSI 虚接口等信息。

```
[SwitchB] display l2vpn vsi verbose
VSI Name: vpna
  VSI Index               : 0
  VSI State               : Up
  MTU                     : 1500
  Bandwidth               : -
  Broadcast Restrain      : -
  Multicast Restrain      : -
  Unknown Unicast Restrain: -
  MAC Learning            : Enabled
  MAC Table Limit         : -
  Drop Unknown            : -
  Flooding                : Enabled
  Gateway interface       : VSI-interface 1
```

```
  VXLAN ID                : 10
  Tunnels:
    Tunnel Name           Link ID    State    Type        Flooding proxy
    Tunnel2               0x5000002  Up       Manual      Disabled
    Tunnel3               0x5000003  Up       Manual      Disabled
```

查看 Switch B 上 VSI 的 ARP 表项信息，可以看到已学习到了虚拟机的 ARP 信息。

```
[SwitchB] display arp
  Type: S-Static   D-Dynamic   O-Openflow   R-Rule   M-Multiport  I-Invalid
IP Address       MAC Address     VID     Interface/Link ID      Aging Type
20.1.1.5         000c-29c1-5e46  N/A     Vlan20                 19    D
10.1.1.11        0000-1234-0001  N/A     Vsi10                  20    D
10.1.1.12        0000-1234-0002  N/A     Vsi10                  19    D
```

查看 Switch B 上 FIB 表项信息，可以看到已学习到了虚拟机的转发表项信息。

```
[SwitchB] display fib 10.1.1.11
Destination count: 1 FIB entry count: 1
Flag:
  U:Useable   G:Gateway   H:Host   B:Blackhole   D:Dynamic   S:Static
  R:Relay     F:FRR
Destination/Mask   Nexthop         Flag     OutInterface/Token       Label
10.1.1.11/32       10.1.1.11       UH       Vsi10                    Null
```

(2) 验证主机和广域网互访

虚拟机 VM 1、VM 2 之间可以互访，VM 1、VM 2 和 Switch E 上接口 Vlan-interface20 的地址 20.1.1.5 之间可以互访。

3.8.2 集中式VXLAN IP网关保护组配置举例

说明

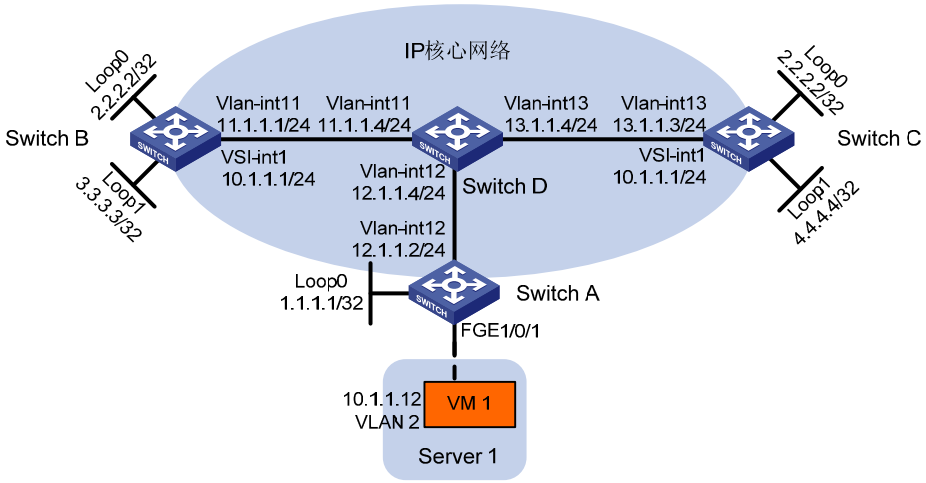
仅 Release 1138P01 及以上版本支持此配置举例。

1. 组网需求

Switch A 为与服务器连接的 VTEP 设备，Switch B 和 Switch C 为与广域网连接的集中式 VXLAN IP 网关设备。虚拟机 VM 1 属于 VXLAN 10，通过 VXLAN IP 网关保护组实现 Switch B 和 Switch C 能够同时为 VM 1 的跨网络报文进行三层转发，同时实现网关设备的备份。

2. 组网图

图3-6 集中式 VXLAN IP 网关保护组配置组网图



3. 配置步骤

(1) 配置 IP 地址和单播路由协议

请按照 图 3-6 配置各接口的IP地址和子网掩码，并在IP核心网络内配置OSPF协议，具体配置过程略。

(2) 配置 Switch A

开启 L2VPN 能力。

```
<SwitchA> system-view
[SwitchA] l2vpn enable
```

配置 VXLAN 隧道工作在二层转发模式。

```
[SwitchA] undo vxlan ip-forwarding
```

创建 VSI 实例 vpna 和 VXLAN 10。

```
[SwitchA] vsi vpna
[SwitchA-vsi-vpna] vxlan 10
[SwitchA-vsi-vpna-vxlan10] quit
[SwitchA-vsi-vpna] quit
```

配置接口 Loopback0 的 IP 地址，作为隧道的源端地址。

```
[SwitchA] interface loopback 0
[SwitchA-Loopback0] ip address 1.1.1.1 255.255.255.255
[SwitchA-Loopback0] quit
```

在 Switch A 和 VXLAN IP 保护组之间建立 VXLAN 隧道：

- 创建模式为 VXLAN 的隧道接口 Tunnel1。
- 指定隧道的源端地址为本地接口 Loopback0 的地址 1.1.1.1。
- 指定隧道的目的端地址为 Switch B 和 Switch C 上同时存在的接口 Loopback0 的地址 2.2.2.2。

```
[SwitchA] interface tunnel 1 mode vxlan
[SwitchA-Tunnel1] source 1.1.1.1
[SwitchA-Tunnel1] destination 2.2.2.2
```

```
[SwitchA-Tunnel1] quit
```
配置 Tunnel1 与 VXLAN 10 关联。
```
[SwitchA] vsi vpna
[SwitchA-vsi-vpna] vxlan 10
[SwitchA-vsi-vpna-vxlan10] tunnel 1
[SwitchA-vsi-vpna-vxlan10] quit
[SwitchA-vsi-vpna] quit
```
创建 VLAN2，将 FortyGigE1/0/1 端口加入 VLAN2。
```
[SwitchA] vlan 2
[SwitchA–vlan2] port fortygige 1/0/1
[SwitchA–vlan2] quit
```
在接入服务器的接口 FortyGigE1/0/1 上创建以太网服务实例 1000，该实例用来匹配 VLAN 2 的数据帧。
```
[SwitchA] interface fortygige 1/0/1
[SwitchA-FortyGigE1/0/1] service-instance 1000
[SwitchA-FortyGigE1/0/1-srv1000] encapsulation s-vid 2
```
配置以太网服务实例 1000 与 VSI 实例 vpna 关联。
```
[SwitchA-FortyGigE1/0/1-srv1000] xconnect vsi vpna
[SwitchA-FortyGigE1/0/1-srv1000] quit
[SwitchA-FortyGigE1/0/1] quit
```
指定 VXLAN IP 网关保护组以及成员。
```
[SwitchA] vtep group 2.2.2.2 member remote 3.3.3.3 4.4.4.4
```
(3) 配置 Switch B

开启 L2VPN 能力。
```
<SwitchB> system-view
[SwitchB] l2vpn enable
```
配置预留 VLAN 3000 接口资源（全局类型）。
```
[SwitchB] reserve-vlan-interface 3000 global
```
创建 VSI 实例 vpna 和 VXLAN 10。
```
[SwitchB] vsi vpna
[SwitchB-vsi-vpna] vxlan 10
[SwitchB-vsi-vpna-vxlan10] quit
[SwitchB-vsi-vpna] quit
```
配置接口 Loopback0 的 IP 地址，作为保护组的 VTEP IP 地址。
```
[SwitchB] interface loopback 0
[SwitchB-Loopback0] ip address 2.2.2.2 255.255.255.255
[SwitchB-Loopback0] quit
```
配置接口 Loopback1 的 IP 地址，作为保护组的成员地址。
```
[SwitchB] interface loopback 1
[SwitchB-Loopback1] ip address 3.3.3.3 255.255.255.255
[SwitchB-Loopback1] quit
```
在 VXLAN IP 网关保护组和 Switch A 之间建立 VXLAN 隧道。
```
[SwitchB] interface tunnel 2 mode vxlan
[SwitchB-Tunnel2] source 2.2.2.2
[SwitchB-Tunnel2] destination 1.1.1.1
[SwitchB-Tunnel2] quit
```

配置 Tunnel2 与 VXLAN10 关联。

```
[SwitchB] vsi vpna
[SwitchB-vsi-vpna] vxlan 10
[SwitchB-vsi-vpna-vxlan10] tunnel 2
[SwitchB-vsi-vpna-vxlan10] quit
[SwitchB-vsi-vpna] quit
```

创建 VSI 虚接口 VSI-interface1，为其配置 IP 地址，该 IP 地址作为 VXLAN 10 内虚拟机的网关地址，并指定该接口的 MAC 地址。

```
[SwitchB] interface vsi-interface 1
[SwitchB-Vsi-interface1] ip address 10.1.1.1 255.255.255.0
[SwitchB-Vsi-interface1] mac-address 2-2-2
[SwitchB-Vsi-interface1] quit
```

配置 VXLAN 10 所在的 VSI 实例和接口 VSI-interface1 关联。

```
[SwitchB] vsi vpna
[SwitchB-vsi-vpna] gateway vsi-interface 1
[SwitchB-vsi-vpna] quit
```

配置 VXLAN IP 网关保护组，并配置本地成员地址。

```
[SwitchB] vtep group 2.2.2.2 member local 3.3.3.3
```

配置 VXLAN IP 网关保护组的其它成员地址。

```
[SwitchB] vtep group 2.2.2.2 member remote 4.4.4.4
```

(4) 配置 Switch C

开启 L2VPN 能力。

```
<SwitchC> system-view
[SwitchC] l2vpn enable
```

配置预留 VLAN 3000 接口资源（全局类型）。

```
[SwitchC] reserve-vlan-interface 3000 global
```

创建 VSI 实例 vpna 和 VXLAN 10。

```
[SwitchC] vsi vpna
[SwitchC-vsi-vpna] vxlan 10
[SwitchC-vsi-vpna-vxlan10] quit
[SwitchC-vsi-vpna] quit
```

配置接口 Loopback0 的 IP 地址，作为数据隧道的源端地址。

```
[SwitchC] interface loopback 0
[SwitchC-Loopback0] ip address 2.2.2.2 255.255.255.255
[SwitchC-Loopback0] quit
```

配置接口 Loopback1 的 IP 地址，作为保护组的成员地址。

```
[SwitchC] interface loopback 1
[SwitchC-Loopback1] ip address 4.4.4.4 255.255.255.255
[SwitchC-Loopback1] quit
```

在 VXLAN IP 网关保护组和 Switch A 之间建立 VXLAN 隧道。

```
[SwitchC] interface tunnel 2 mode vxlan
[SwitchC-Tunnel2] source 2.2.2.2
[SwitchC-Tunnel2] destination 1.1.1.1
[SwitchC-Tunnel2] quit
```

配置 Tunnel2 与 VXLAN10 关联。

```
[SwitchC] vsi vpna
```

```
[SwitchC-vsi-vpna] vxlan 10
[SwitchC-vsi-vpna-vxlan10] tunnel 2
[SwitchC-vsi-vpna-vxlan10] quit
[SwitchC-vsi-vpna] quit
```

创建 VSI 虚接口 VSI-interface1，为其配置 IP 地址，该 IP 地址作为 VXLAN 10 内虚拟机的网关地址，并指定该接口的 MAC 地址。

```
[SwitchC] interface vsi-interface 1
[SwitchC-Vsi-interface1] ip address 10.1.1.1 255.255.255.0
[SwitchC-Vsi-interface1] mac-address 2-2-2
[SwitchC-Vsi-interface1] quit
```

配置 VXLAN 10 所在的 VSI 实例和接口 VSI-interface1 关联。

```
[SwitchC] vsi vpna
[SwitchC-vsi-vpna] gateway vsi-interface 1
[SwitchC-vsi-vpna] quit
```

配置 VXLAN IP 网关保护组，并配置本地成员地址。

```
[SwitchC] vtep group 2.2.2.2 member local 4.4.4.4
```

配置 VXLAN IP 网关保护组的其它成员地址。

```
[SwitchC] vtep group 2.2.2.2 member remote 3.3.3.3
```

4 OVSDB-VTEP

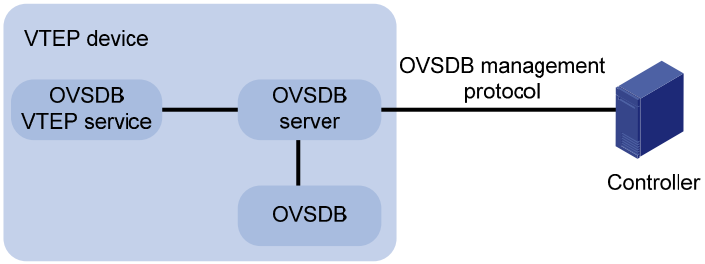
说明

仅 Release 1138P01 及以上版本支持此特性。

4.1 简介

OVSDB（Open vSwitch Database，开源虚拟交换机数据库）控制协议用来实现 NVC（Network Virtualization Controller，网络虚拟化控制器）对网络中 VTEP 设备的管理和部署。

图4-1 OVSDB-VTEP 示意图



如 图 4-1 所示，VTEP设备上维护OVSDB数据库，VXLAN相关配置以表项的形式保存在该数据库中。控制器与VTEP设备上的OVSDB服务器建立连接，二者采用OVSDB控制协议进行交互并操作OVSDB数据库中的数据。OVSDB VTEP服务从OVSDB服务器获取数据库中的数据，将其转变为VXLAN相关配置（例如创建或删除VXLAN、创建或删除VXLAN隧道）下发到设备上。同时，OVSDB VTEP服务也会通过OVSDB服务器，将本地的用户侧接入端口和VXLAN隧道全局源地址信息添加到数据库中，并上报给控制器。

提示

用户可以同时通过命令行和控制器配置 VTEP 设备。建议不要在 VTEP 设备上通过命令行删除控制器下发的配置。

4.2 协议规范

与 OVSDB 相关的协议规范有：

- RFC 7047：The Open vSwitch Database Management Protocol

4.3 OVSDB-VTEP配置任务简介

要实现控制器对VTEP设备的部署，需要在VTEP设备上完成 表 4-1 所示配置。

表4-1 OVSDB-VTEP 配置任务简介

配置任务		说明	详细配置
与控制器建立OVSDB连接	与控制器建立主动SSL连接	必选 OVSDB服务器支持同时与多个控制器建立连接，且支持同时建立多种类型的连接 在开启OVSDB服务器之前，必须先进行本配置。如果在开启OVSDB服务器之后修改本配置，那么需要关闭OVSDB服务器后再重新开启，新的连接配置才能生效	4.5.1
	与控制器建立被动SSL连接		4.5.2
	与控制器建立主动TCP连接		4.5.3
	与控制器建立被动TCP连接		4.5.4
开启OVSDB服务器		必选	4.6
开启OVSDB VTEP服务		必选	4.7
配置VXLAN隧道的全局源地址		必选	4.8
指定用户侧的接入端口		必选	4.9
开启组播隧道泛洪代理功能		可选	4.10

4.4 配置准备

在进行 OVSDB-VTEP 相关配置前，需要首先通过 **l2vpn enable** 命令使能 L2VPN 功能。

如果 OVSDB 服务器与控制器之间建立 SSL 连接，则还需要完成 SSL 相关配置。

4.5 与控制器建立OVSDB连接

OVSDB 服务器和控制器之间可以建立多种类型的 OVSDB 连接，设备支持的 OVSDB 连接类型包括：

- 主动 SSL 连接：OVSDB 服务器主动向控制器发起 SSL 连接。该连接方式必须指定 SSL 使用的 PKI 域。OVSDB 服务器最多可以同时与 8 个控制器建立主动 SSL 连接。
- 被动 SSL 连接：OVSDB 服务器监听并接收来自控制器的 SSL 连接请求。该连接方式必须指定 SSL 使用的 PKI 域。OVSDB 服务器只能监听 1 个端口的 SSL 连接请求。
- 主动 TCP 连接：OVSDB 服务器主动向控制器发起 TCP 连接。OVSDB 服务器最多可以同时与 8 个控制器建立主动 TCP 连接。
- 被动 TCP 连接：OVSDB 服务器监听并接收来自控制器的 TCP 连接请求。OVSDB 服务器只能监听 1 个端口的 TCP 连接请求。

说明

所有 SSL 连接，包括主动 SSL 连接和被动 SSL 连接，使用相同的 PKI 域和 CA 证书文件。

4.5.1 与控制器建立主动SSL连接

表4-2 与控制器建立主动 SSL 连接

操作	命令	说明
进入系统视图	**system-view**	-
指定与控制器进行SSL通信时使用的PKI域	**ovsdb server pki domain** *domain-name*	缺省情况下，未指定与控制器进行SSL通信时使用的PKI域
（可选）设置SSL通信时使用的CA证书文件	**ovsdb server bootstrap ca-certificate** *ca-filename*	缺省情况下，与控制器进行SSL通信时使用PKI域中的CA证书文件
与控制器建立主动SSL连接	**ovsdb server ssl ip** *ip-address* **port** *port-number*	缺省情况下，不会与控制器建立主动SSL连接

4.5.2 与控制器建立被动SSL连接

表4-3 与控制器建立被动 SSL 连接

操作	命令	说明
进入系统视图	**system-view**	-
指定与控制器进行SSL通信时使用的PKI域	**ovsdb server pki domain** *domain-name*	缺省情况下，未指定与控制器进行SSL通信时使用的PKI域
（可选）设置SSL通信时使用的CA证书文件	**ovsdb server bootstrap ca-certificate** *ca-filename*	缺省情况下，与控制器进行SSL通信时使用PKI域中的CA证书文件
与控制器建立被动SSL连接	**ovsdb server pssl** [**port** *port-number*]	缺省情况下，不会与控制器建立被动SSL连接

4.5.3 与控制器建立主动TCP连接

表4-4 与控制器建立主动 TCP 连接

操作	命令	说明
进入系统视图	**system-view**	-
与控制器建立主动TCP连接	**ovsdb server tcp ip** *ip-address* **port** *port-number*	缺省情况下，不会与控制器建立主动TCP连接

4.5.4 与控制器建立被动TCP连接

表4-5 与控制器建立被动 TCP 连接

操作	命令	说明
进入系统视图	**system-view**	-

操作	命令	说明
与控制器建立被动TCP连接	**ovsdb server ptcp** [**port** *port-number*]	缺省情况下，不会与控制器建立被动TCP连接

4.6 开启OVSDB服务器

需要注意的是，为了保证 OVSDB 服务器能够与控制器建立连接，在开启 OVSDB 服务器前，必须先配置至少一条与控制器的连接。

表4-6 开启 OVSDB 服务器

操作	命令	说明
进入系统视图	**system-view**	-
开启OVSDB服务器	**ovsdb server enable**	缺省情况下，OVSDB服务器处于关闭状态

4.7 开启OVSDB VTEP服务

表4-7 开启 OVSDB VTEP 服务

操作	命令	说明
进入系统视图	**system-view**	-
开启OVSDB VTEP服务	**vtep enable**	缺省情况下，OVSDB VTEP服务处于关闭状态

4.8 配置VXLAN隧道的全局源地址

用户需要在 VTEP 设备上配置 VXLAN 隧道的全局源地址，该地址会通过 OVSDB 协议上报给控制器，用于控制器对 VTEP 设备进行部署和控制。

采用 OVSDB 对 VTEP 设备进行部署和控制时，用户不能在 VXLAN 隧道的 Tunnel 接口下手工指定源地址，否则会影响控制器对 VTEP 设备的管理。

表4-8 配置 VXLAN 隧道的全局源地址

操作	命令	说明
进入系统视图	**system-view**	-
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

4.9 指定用户侧的接入端口

表4-9 指定接入侧端口

操作		命令	说明
进入系统视图		**system-view**	-
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*	
指定当前接口为用户侧的接入端口		**vtep access port**	缺省情况下，接口不是用户侧的接入端口

4.10 开启组播隧道泛洪代理功能

开启组播隧道泛洪代理功能后，系统会将控制器下发的组播隧道转换为具有泛洪代理功能的隧道。VXLAN 内的广播、组播和未知单播流量将通过具有泛洪代理功能的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。采用泛洪代理（服务器复制）方式转发站点间的泛洪流量时，必须开启该功能。

表4-10 开启组播隧道泛洪代理功能

操作	命令	说明
进入系统视图	**system-view**	-
开启组播隧道泛洪代理功能	**vxlan tunnel service node**	缺省情况下，组播隧道泛洪代理功能处于关闭状态

4.11 OVSDB-VTEP典型配置举例

4.11.1 OVSDB-VTEP头端复制配置举例

1. 组网需求

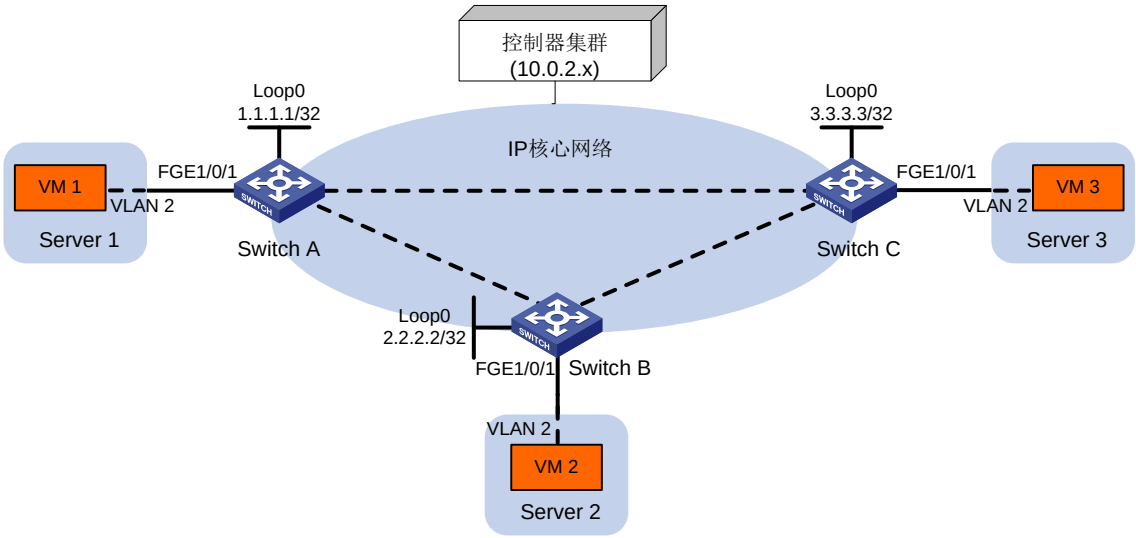
Switch A、Switch B、Switch C 为与服务器连接的 VTEP 设备。虚拟机 VM 1、VM 2 和 VM 3 同属于 VXLAN 10。通过 VXLAN 实现不同站点间的二层互联，确保虚拟机在站点之间进行迁移时用户的访问流量不会中断。

具体需求为：

- 通过控制器下发配置，在不同 VTEP 之间建立 VXLAN 隧道。
- 站点之间的泛洪流量采用头端复制的方式转发。

2. 组网图

图4-2 OVSDB-VTEP 头端复制组网图



3. 配置步骤

(1) 配置 IP 地址、单播路由协议、控制器

为 图 4-2 中的控制器和各台设备配置IP地址，并配置单播路由协议，实现不同设备之间的互通。在IP核心网中还需要部署VXLAN IP网关，以实现VXLAN 10 内的虚拟机与外界网络通信。具体配置过程略。

(2) 配置 Switch A

开启 L2VPN 能力。

```
<SwitchA> system-view
[SwitchA] l2vpn enable
```

配置 VXLAN 隧道工作在二层转发模式。

```
[SwitchA] undo vxlan ip-forwarding
```

配置与控制器建立主动 TCP 连接，连接的目的地址为 10.0.2.15（控制器的地址），目的端口号为 6632。

```
[SwitchA] ovsdb server tcp ip 10.0.2.15 port 6632
```

开启 OVSDB 服务器。

```
[SwitchA] ovsdb server enable
```

开启 OVSDB VTEP 服务。

```
[SwitchA] vtep enable
```

指定接入服务器的接口 FortyGigE1/0/1 上为用户侧的接入端口。

```
[SwitchA] interface fortygige 1/0/1
[SwitchA-FortyGigE1/0/1] vtep access port
[SwitchA-FortyGigE1/0/1] quit
```

(3) 配置 Switch B

开启 L2VPN 功能。

```
<SwitchB> system-view
[SwitchB] l2vpn enable
```

配置 VXLAN 隧道工作在二层转发模式。

```
[SwitchB] undo vxlan ip-forwarding
```

配置与控制器建立主动 TCP 连接，连接的目的地址为 10.0.2.15（控制器的地址），目的端口号为 6632。

```
[SwitchB] ovsdb server tcp ip 10.0.2.15 port 6632
```

开启 OVSDB 服务器。

```
[SwitchB] ovsdb server enable
```

开启 OVSDB VTEP 服务。

```
[SwitchB] vtep enable
```

指定接入服务器的接口 FortyGigE1/0/1 上为用户侧的接入端口。

```
[SwitchB] interface fortygige 1/0/1
[SwitchB-FortyGigE1/0/1] vtep access port
[SwitchB-FortyGigE1/0/1] quit
```

(4) 配置 Switch C

开启 L2VPN 功能。

```
<SwitchC> system-view
[SwitchC] l2vpn enable
```

配置 VXLAN 隧道工作在二层转发模式。

```
[SwitchC] undo vxlan ip-forwarding
```

配置与控制器建立主动 TCP 连接，连接的目的地址为 10.0.2.15（控制器的地址），目的端口号为 6632。

```
[SwitchC] ovsdb server tcp ip 10.0.2.15 port 6632
```

开启 OVSDB 服务器。

```
[SwitchC] ovsdb server enable
```

开启 OVSDB VTEP 服务。

```
[SwitchC] vtep enable
```

指定接入服务器的接口 FortyGigE1/0/1 上为用户侧的接入端口。

```
[SwitchC] interface fortygige 1/0/1
[SwitchC-FortyGigE1/0/1] vtep access port
[SwitchC-FortyGigE1/0/1] quit
```

(5) 控制器上进行 VXLAN 配置（略）

4. 验证配置

(1) 验证 VTEP 设备（下文以 Switch A 为例，其它设备验证方法与此类似）

查看 Switch A 上的 Tunnel 接口信息，可以看到 VXLAN 模式的 Tunnel 接口处于 up 状态。

```
[SwitchA] display interface tunnel
Tunnel1
Current state: UP
Line protocol state: UP
Description: Tunnel1 Interface
Bandwidth: 64kbps
Maximum transmission unit: 1464
Internet protocol processing: Disabled
Last clearing of counters: Never
Tunnel source 1.1.1.1, destination 2.2.2.2
```

```
Tunnel protocol/transport UDP_VXLAN/IP
Last 300 seconds input rate: 0 bytes/sec, 0 bits/sec, 0 packets/sec
Last 300 seconds output rate: 0 bytes/sec, 0 bits/sec, 0 packets/sec
Input: 0 packets, 0 bytes, 0 drops
Output: 0 packets, 0 bytes, 0 drops
```

查看 Switch A 上的 VSI 信息，可以看到 VSI 内创建的 VXLAN、与 VXLAN 关联的 VXLAN 隧道、与 VSI 关联的以太网服务实例等信息。

```
[SwitchA] display l2vpn vsi verbose
VSI Name: evpn2014
  VSI Index               : 0
  VSI State               : Up
  MTU                     : 1500
  Bandwidth               : -
  Broadcast Restrain      : -
  Multicast Restrain      : -
  Unknown Unicast Restrain: -
  MAC Learning            : Enabled
  MAC Table Limit         : -
  Drop Unknown            : -
  Flooding                : Enabled
  VXLAN ID                : 10
  Tunnels:
    Tunnel Name          Link ID    State  Type        Flooding proxy
    Tunnel1              0x5000001  Up     Manual      Disabled
    Tunnel2              0x5000002  Up     Manual      Disabled
  ACs:
    AC                              Link ID    State
    FGE1/0/1 srv2                    0          Up
```

查看 Switch A 上 VSI 的 MAC 地址表项信息，可以看到已学习到的 MAC 地址信息。

```
<SwitchA> display l2vpn mac-address
MAC Address      State    VSI Name                        Link ID/Name  Aging
00ea-1100-0001   Dynamic  SDN_VSI_8008                    0             Aging
0023-89aa-2f0a   Dynamic  SDN_VSI_8008                    Tunnel257     Aging
3c8c-404e-dd46   Dynamic  SDN_VSI_8008                    Tunnel257     Aging
--- 3 mac address(es) found  ---
```

(2)　验证主机

虚拟机 VM 1、VM 2、VM 3 之间可以互访。

4.11.2 OVSDB-VTEP泛洪代理配置举例

1. 组网需求

Switch A、Switch B、Switch C 为与服务器连接的 VTEP 设备。虚拟机 VM 1、VM 2 和 VM 3 同属于 VXLAN 10。通过 VXLAN 实现不同站点间的二层互联，确保虚拟机在站点之间进行迁移时用户的访问流量不会中断。

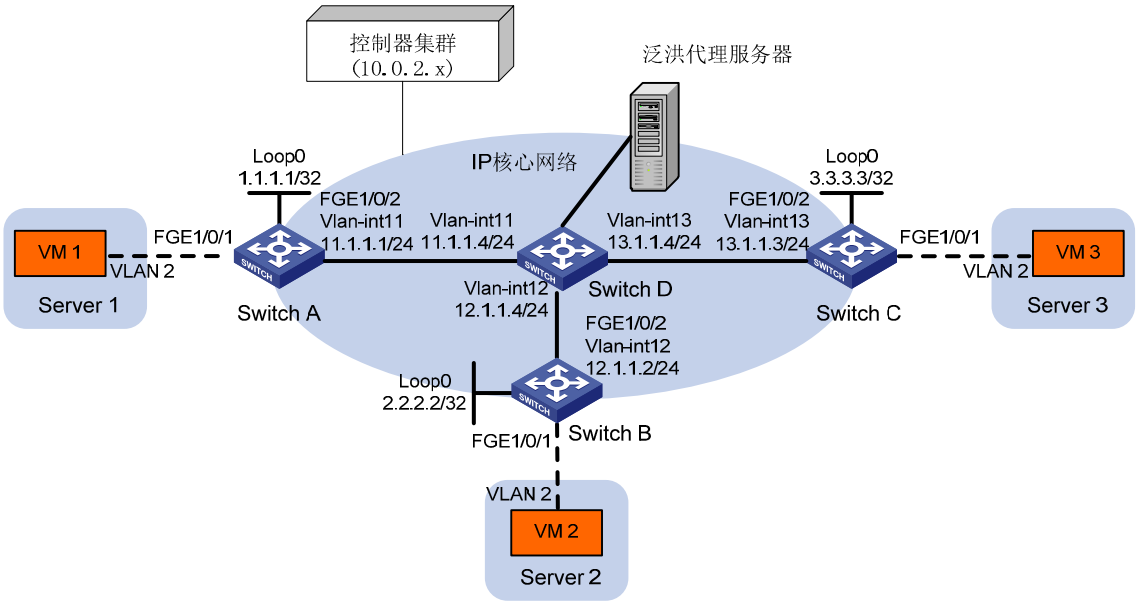
具体需求为：

- 不同 VTEP 之间通过控制器下发配置建立 VXLAN 隧道。

- 站点之间的泛洪流量采用泛洪代理（服务器复制）的方式转发。
- VTEP 采用控制器下发的 MAC 地址表项进行流量转发。

2. 组网图

图4-3 OVSDB-VTEP 泛洪代理组网图



3. 配置步骤

(1) 配置 IP 地址、单播路由协议、控制器和服务器

请按照 图 4-3 配置各接口的IP地址和子网掩码，并在IP核心网络内配置OSPF协议，具体配置过程略。

(2) 配置 Switch A

开启 L2VPN 功能。

```
<SwitchA> system-view
[SwitchA] l2vpn enable
```

与控制器建立连接并开启 OVSDB 服务器。

```
[SwitchA] ovsdb server tcp ip 10.0.2.15 port 6632
[SwitchA] ovsdb server enable
```

开启 OVSDB VTEP 服务。

```
[SwitchA] vtep enable
```

配置接口 Loopback0 的 IP 地址，作为 VXLAN 隧道的全局源地址。

```
[SwitchA] interface loopback 0
[SwitchA-LoopBack0] ip address 1.1.1.1 255.255.255.255
[SwitchA-LoopBack0] quit
[SwitchA] tunnel global source-address 1.1.1.1
```

在接入服务器的接口 FortyGigE1/0/1 上指定其为接入侧端口。

```
[SwitchA] interface fortygige 1/0/1
[SwitchA-FortyGigE1/0/1] vtep access port
[SwitchA-FortyGigE1/0/1] quit
```

关闭远端 MAC 地址自动学习功能。

```
[SwitchA] vxlan tunnel mac-learning disable
```

开启组播隧道泛洪代理功能。

```
[SwitchA] vxlan tunnel service node
```

(3) 配置 Switch B

开启 L2VPN 功能。

```
<SwitchB> system-view
[SwitchB] l2vpn enable
```

与控制器建立连接并开启 OVSDB 服务器。

```
[SwitchB] ovsdb server tcp ip 10.0.2.15 port 6632
[SwitchB] ovsdb server enable
```

开启 OVSDB VTEP 服务。

```
[SwitchB] vtep enable
```

配置接口 Loopback0 的 IP 地址，作为 VXLAN 隧道的全局源地址。

```
[SwitchB] interface loopback0
[SwitchB-LoopBack0] ip address 2.2.2.2 255.255.255.255
[SwitchB-LoopBack0] quit
[SwitchB] tunnel global source-address 2.2.2.2
```

在接入服务器的接口 FortyGigE1/0/1 上指定其为接入侧端口。

```
[SwitchB] interface fortygige 1/0/1
[SwitchB-FortyGigE1/0/1] vtep access port
[SwitchB-FortyGigE1/0/1] quit
```

关闭远端 MAC 地址自动学习功能。

```
[SwitchB] vxlan tunnel mac-learning disable
```

开启组播隧道泛洪代理功能。

```
[SwitchB] vxlan tunnel service node
```

(4) 配置 Switch C

开启 L2VPN 功能。

```
<SwitchC> system-view
[SwitchC] l2vpn enable
```

与控制器建立连接并开启 OVSDB 服务器。

```
[SwitchC] ovsdb server tcp ip 10.0.2.15 port 6632
[SwitchC] ovsdb server enable
```

开启 OVSDB VTEP 服务。

```
[SwitchC] vtep enable
```

配置接口 Loopback0 的 IP 地址，作为 VXLAN 隧道的全局源地址。

```
[SwitchC] interface loopback0
[SwitchC-LoopBack0] ip address 3.3.3.3 255.255.255.255
[SwitchC-LoopBack0] quit
[SwitchC] tunnel global source-address 3.3.3.3
```

在接入服务器的接口 FortyGigE1/0/1 上指定其为接入侧端口。

```
[SwitchC] interface fortygige 1/0/1
[SwitchC-FortyGigE1/0/1] vtep access port
[SwitchC-FortyGigE1/0/1] quit
```

关闭远端 MAC 地址自动学习功能。

```
[SwitchC] vxlan tunnel mac-learning disable
```

开启组播隧道泛洪代理功能。

```
[SwitchC] vxlan tunnel service node
```

(5) 控制器上进行 VXLAN 配置（略）

4. 验证配置

(1) 验证 VTEP 设备（下文以 Switch A 为例，其它设备验证方法与此类似）

查看 Switch A 上的 Tunnel 接口信息，可以看到 VXLAN 模式的 Tunnel 接口处于 up 状态。

```
[SwitchA] display interface tunnel 1
Tunnel1
Current state: UP
Line protocol state: UP
Description: Tunnel1 Interface
Bandwidth: 64 kbps
Maximum transmission unit: 1464
Internet protocol processing: Disabled
Last clearing of counters: Never
Tunnel source 1.1.1.1, destination 2.2.2.2
Tunnel protocol/transport UDP_VXLAN/IP
```

查看 Switch A 上的 VSI 信息，可以看到 VSI 内创建的 VXLAN、与 VXLAN 关联的 VXLAN 隧道、与 VSI 关联的以太网服务实例等信息，其中 Tunnel1 和 Tunnel2 为去往 SwitchB、SwitchC 的隧道，Tunnel3 为去往代理服务器的隧道。

```
[SwitchA] display l2vpn vsi verbose
VSI Name: evpn2014
  VSI Index               : 0
  VSI State               : Up
  MTU                     : 1500
  Bandwidth               : -
  Broadcast Restrain      : -
  Multicast Restrain      : -
  Unknown Unicast Restrain: -
  MAC Learning            : Enabled
  MAC Table Limit         : -
  Drop Unknown            : -
  Flooding                : Enabled
  VXLAN ID                : 10
  Tunnels:
    Tunnel Name          Link ID    State  Type        Flooding proxy
    Tunnel1              0x5000001  Up     Manual      Disabled
    Tunnel2              0x5000002  Up     Manual      Disabled
    Tunnel3              0x5000003  Up     Manual      Enabled
  ACs:
    AC                               Link ID    State
    FGE1/0/1 srv2                    0          Up
```

查看 Switch A 上 VSI 的 MAC 地址表项信息，可以看到控制器下发的 MAC 地址信息。

```
<SwitchA> display l2vpn mac-address
MAC Address      State    VSI Name                        Link ID/Name  Aging
```

```
cc3e-5f9c-6cdb   Static   evpn2014                        Tunnel1       NotAging
cc3e-5f9c-23dc   Static   evpn2014                        Tunnel2       NotAging
--- 2 mac address(es) found  ---
```

(2) 验证主机

虚拟机 VM 1、VM 2、VM 3 之间可以互访。