目 录

1 VXLAN 简介

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和园区接入网络。

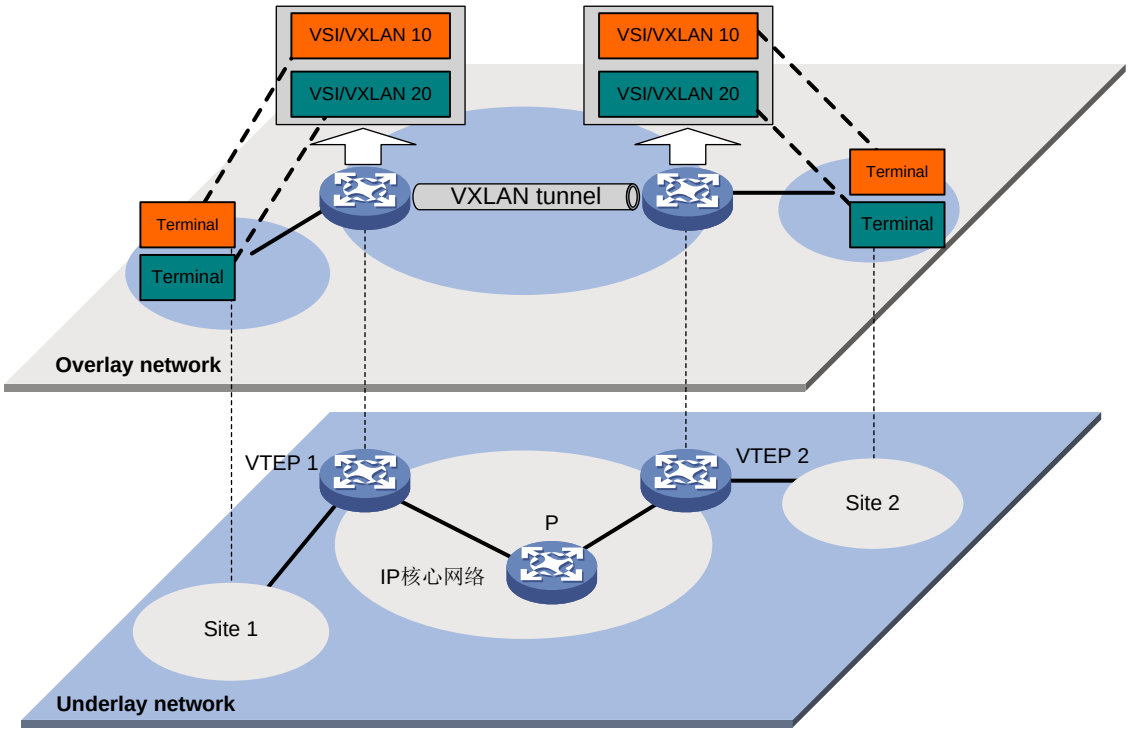
VXLAN 具有如下特点：

- 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，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。

1.1 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

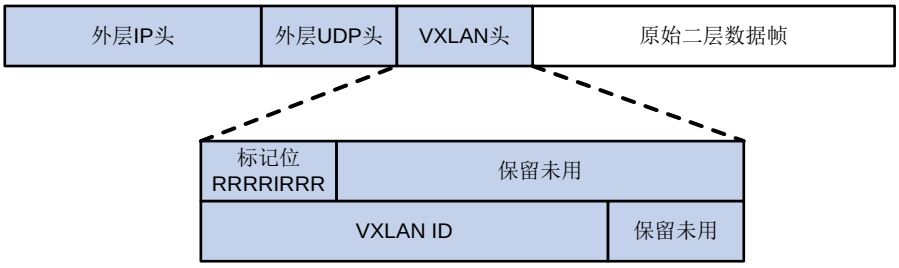


如图 1-1 所示，VXLAN 的典型网络模型中包括如下几部分：

- 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。
- 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。

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) 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) 学习用户终端的 MAC 地址。

(4) 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 采用如下几种方式在数据帧和 VXLAN 之间建立关联：

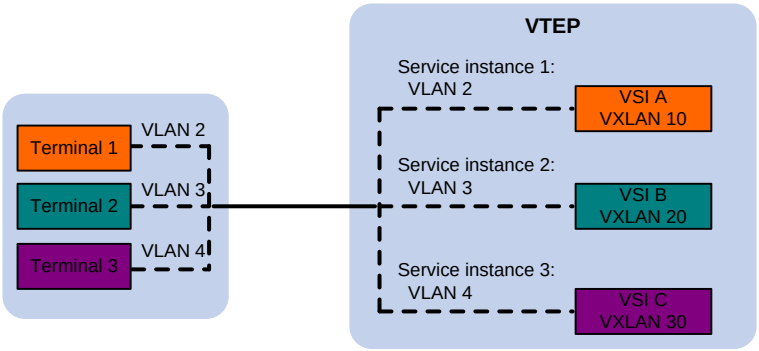
- 将以太网服务实例与 VSI 关联：以太网服务实例在二层以太网接口上创建，它定义了一系列匹配规则（匹配接口接收到的所有报文、匹配所有携带 VLAN Tag 的报文、匹配所有不携带 VLAN Tag 的报文等）。从二层以太网接口上接收到的、与规则匹配的数据帧均属于指定的 VSI/VXLAN。
- 将 VLAN 与 VXLAN 关联：VTEP 接收到的该 VLAN 的数据帧均属于指定的 VXLAN。

VTEP 从指定 VLAN 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。

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

如图 1-3 所示，Terminal 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 Terminal 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。

- 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 对远端站点内用户终端 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（远端用户终端的 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 BGP EVPN 学习：在 VTEP 上运行 BGP EVPN，通过 BGP EVPN 将本地 MAC 地址及其所属的 VXLAN 信息通告给远端 VTEP。远端 VTEP 接收到该信息后，在 VXLAN 对应 VSI 的 MAC 地址表中添加 MAC 地址表项。EVPN 的详细介绍请参见“EVPN 配置指导”。
  - 通过 OpenFlow 下发：OpenFlow 控制器以流表的形式向 VTEP 设备下发远端 MAC 地址表项。OpenFlow 的详细介绍请参见“OpenFlow 配置指导”。
  - 通过 OVSDB 下发：控制器通过 OVSDB 协议向 VTEP 设备下发远端 MAC 地址表项。

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：

- 静态配置、OpenFlow 下发、OVSDB 下发的 MAC 地址优先级相同，且优先级最高。
- 通过 BGP EVPN 学习的 MAC 地址优先级，且优先级次之。
- 动态学习的 MAC 地址优先级最低。

1.3.4 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

如果以太网服务实例上的匹配规则为 **encapsulation untagged**，则该模式下，VTEP 从本地站点接收到不携带 VLAN tag 的数据帧后，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。

如果以太网服务实例上配置了其他匹配规则，则在该模式下，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 VLAN 接入模式时，以太网服务实例上的匹配规则不能为 **encapsulation default**。

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。

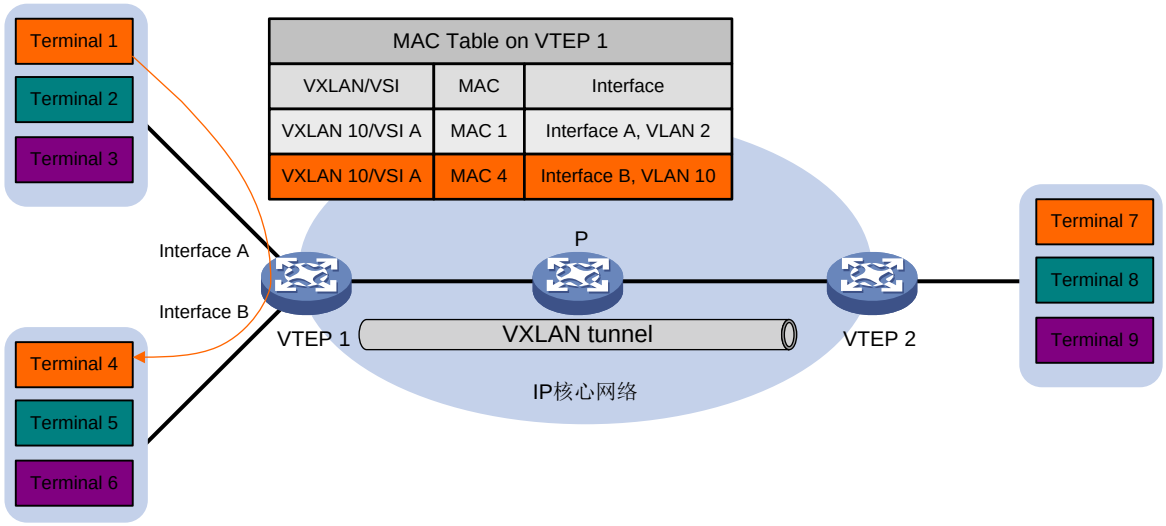
1.3.5 转发单播流量

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用户终端。

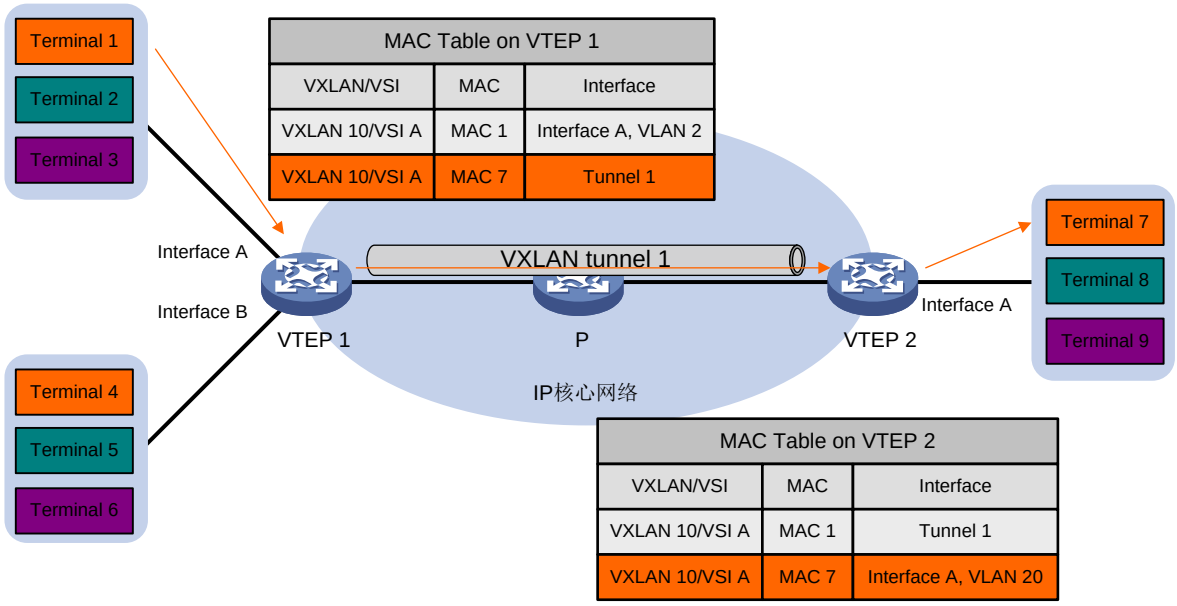
图1-4 站点内单播流量转发



如图 1-4 所示，Terminal 1（MAC 地址为 MAC 1）发送以太网帧到 Terminal 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 Terminal 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 Terminal 1（MAC 地址为 MAC 1）发送以太网帧给 Terminal 7（MAC 地址为 MAC 7）为例，站点间单播流量的转发过程为：

(1) Terminal 1 发送以太网数据帧给 Terminal 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 Terminal 7。

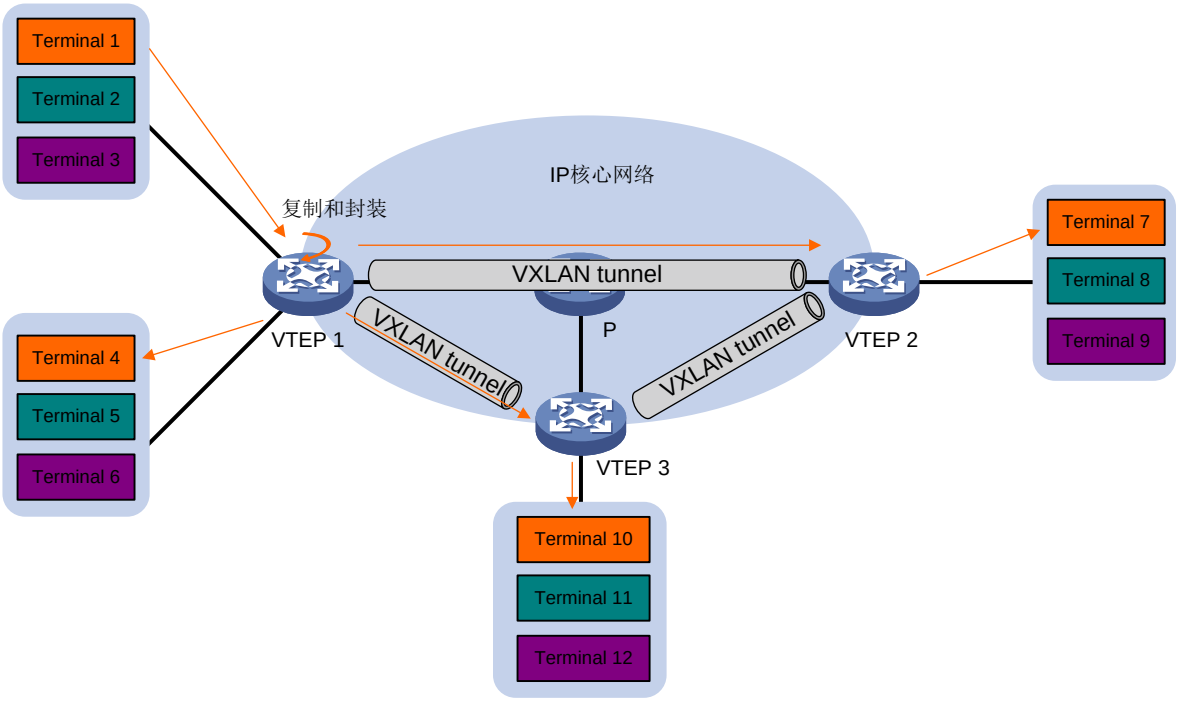
1.3.6 转发泛洪流量

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）。

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 接收到本地用户终端发送的组播、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。

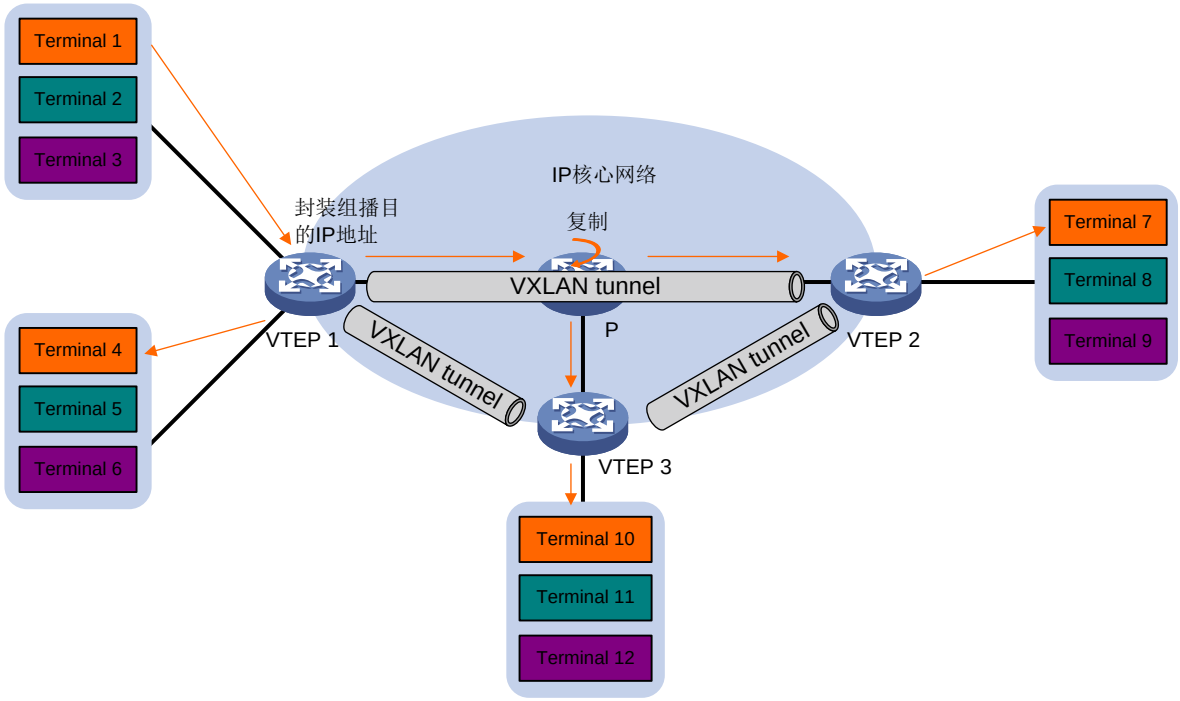
(2) 远端 VTEP（VTEP 2 和 VTEP 3）接收到 VXLAN 报文后，解封装报文，将原始的数据帧在本地站点的指定 VXLAN 内泛洪。为了避免环路，远端 VTEP 从 VXLAN 隧道上接收到报文后，不会再将其泛洪到其他的 VXLAN 隧道。

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 接收到本地用户终端发送的组播、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。

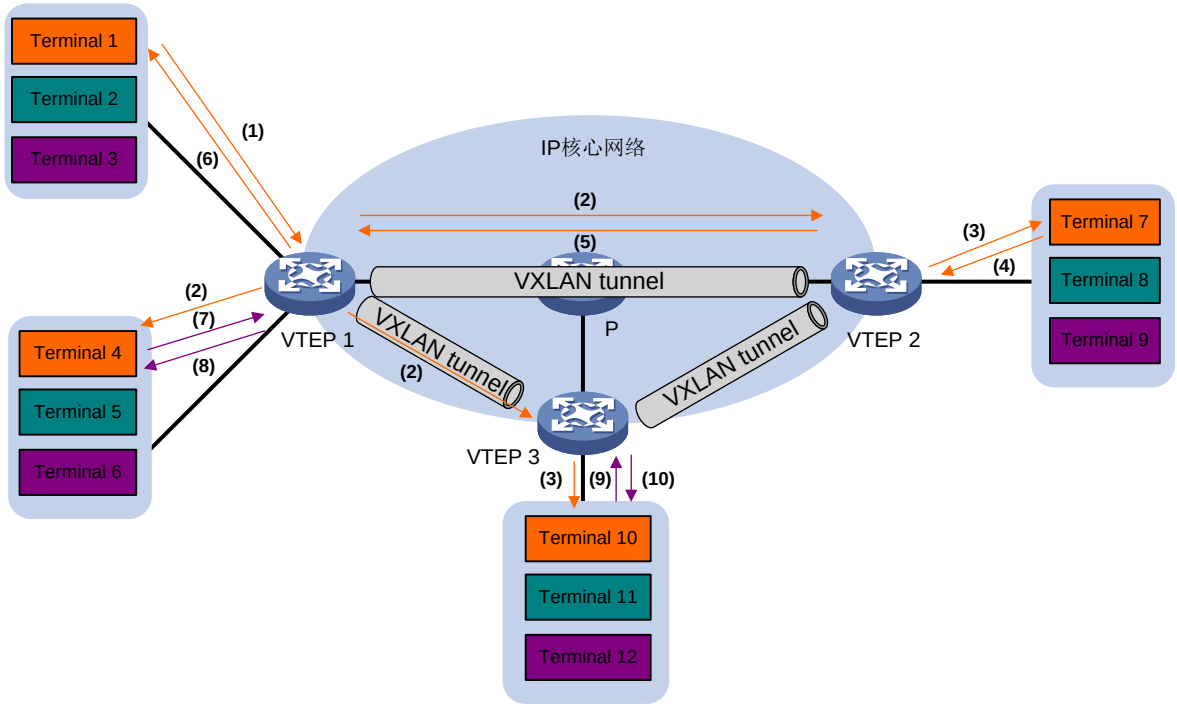
(2) 在 IP 核心网内，P 设备根据已经建立的组播转发表项复制并转发该组播报文。

(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。

1.4 ARP/ND泛洪抑制

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

图1-8 ARP 泛洪抑制示意图



如图 1-8 所示，以 ARP 为例，泛洪抑制的处理过程如下：

(1) 用户终端 Terminal 1 发送 ARP 请求，获取 Terminal 7 的 MAC 地址。

(2) VTEP 1 根据接收到的 ARP 请求，建立 Terminal 1 的 ARP 泛洪抑制表项，并在 VXLAN 内泛洪该 ARP 请求（图 1-8 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 Terminal 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) Terminal 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 Terminal 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 Terminal 7 的 ARP 泛洪抑制表项，之后将 ARP 应答报文发送给 Terminal 1。

(7) 在 VTEP 1 上建立 ARP 泛洪抑制表项后，用户终端 Terminal 4 发送 ARP 请求，获取 Terminal 1 或 Terminal 7 的 MAC 地址。

(8) VTEP 1 接收到 ARP 请求后，建立 Terminal 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，用户终端 Terminal 10 发送 ARP 请求，获取 Terminal 1 的 MAC 地址。

(10) VTEP 3 接收到 ARP 请求后，建立 Terminal 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：

- 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配置限制和指导

2.1.1 VXLAN 硬件限制

- VXLAN 公网侧端口和 IRF 物理端口必须位于 FD/SG/SH 系列接口板或 LSUM1CQGS32SF0-Z 接口板上。
- VXLAN 用户侧端口必须位于 FD/SG/SH 系列接口板或以下接口板上：
  - SE 系列接口板
  - LSUM1CQGS32SF0-Z 接口板
- 当多个 VXLAN 隧道共用一个公网侧端口，且该端口位于 LSUM1CQGS32SF0-Z 接口板 上时，这些隧道必须使用设备的同一个 VLAN 接口来转发报文。

2.1.2 VXLAN 软件限制

- 请不要在 VXLAN 网络中配置 MPLS 相关功能。有关 MPLS 的详细介绍，请参见“MPLS 配置指导”。
- 请不要同时配置 VXLAN 功能和 MAC 地址认证请求中携带用户 IP 地址功能，关于 MAC 地址认证请求中携带用户 IP 地址功能的详细介绍，请参见“安全配置指导”中的“MAC 地址认证”。
- 端口使能了 EVB 功能后，不支持配置 VXLAN 功能，否则二者均将无法正常工作。有关 EVB 的介绍，请参见“EVB 配置指导”。
- 配置 AC（以太网服务实例）时，需要注意：
  - 对于 VTEP 上配置了 AC 的端口，如果收到的流量不符合以太网服务实例报文匹配规则，则该流量不支持组播。
  - VTEP 设备的同一端口下不支持同时配置 AC 和 QinQ。有关 QinQ 的介绍，请参见“二层技术-以太网交换配置指导”中的“QinQ”。
  - 为保证 VTEP 与本地站点的正常对接，应在配置了 AC 的端口上关闭生成树协议（**undo stp enable**）。
  - 如果端口上配置了 **mac-based ac** 命令，或者该端口的不同 AC 配置了以下任意两种报文匹配规则命令，则该端口不支持透传 VXLAN 报文。
    - **encapsulation s-vid** *vlan-id-list* [**only-tagged**]
    - **encapsulation s-vid** *vlan-id-list* **c-vid** { *vlan-id-list* | **all** }
    - **encapsulation c-vid** *vlan-id-list*
  - 有关 AC 的其他配置限制，请参见“2.6.2 配置手工创建的以太网服务实例与 VSI 关联”
- VXLAN 网络的公网侧端口不支持配置端口优先级信任模式为 DSCP。关于优先级信任模式的配置，请参见“ACL 和 QoS 配置指导”中的“QoS”。
- 在组播路由方式下，VTEP 通过 AC 从本地站点接收到的协议报文（如 ARP/ND 报文、组播协议报文）不能透传到远端站点。

- 在组播路由方式下，不允许在核心设备上创建 VXLAN 或 VXLAN 隧道，否则会导致 VXLAN 报文转发不通。
- VTEP 不支持通过多条等价路由对组播、广播和未知单播流量进行负载分担。
- VXLAN 网络不支持处理带两层以上 VLAN 标签的报文。

2.2 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。

表2-1 VXLAN 配置任务简介

配置任务	说明	详细配置
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
建立数据帧与VSI的关联	必选	2.6
管理本地和远端MAC地址	可选	2.7
配置VXLAN组播路由泛洪方式	可选	2.8
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
配置ND泛洪抑制	可选	2.13
关闭VXLAN远端ARP/ND自动学习功能	可选	2.14
配置VXLAN流量统计	可选	2.15

2.3 创建VSI和VXLAN

表2-2 创建 VSI 和 VXLAN

操作	命令	说明
进入系统视图	**system-view**	-
开启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，不存在VSI
（可选）配置VSI的描述信息	**description** *text*	缺省情况下，未配置VSI的描述信息
开启VSI	**undo shutdown**	缺省情况下，VSI处于开启状态
（可选）配置VSI的MTU值	**mtu** *size*	缺省情况下，VSI的MTU值为1500字节

操作	命令	说明
（可选）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 对于VXLAN，设备不支持关闭VSI的MAC地址学习功能，执行**undo mac-learning enable**命令后不生效
创建VXLAN，并进入VXLAN视图	**vxlan** *vxlan-id*	缺省情况下，不存在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。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。

如果设备上配置了通过 EVPN 自动建立并关联 VXLAN 隧道，则隧道目的地址相同的 EVPN 自动创建隧道和手工创建隧道不能关联同一个 VXLAN。EVPN 的详细介绍请参见“EVPN 配置指导”。

关于隧道的详细介绍及 Tunnel 接口下的更多配置命令，请参见“三层技术-IP 业务配置指导”中的“隧道”。关于 **interface tunnel**、**source** 和 **destination** 命令的详细介绍，请参见“三层技术-IP 业务命令参考”中的“隧道”。

表2-3 手工创建 VXLAN 隧道

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 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
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，不存在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

操作	命令	说明
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，未设置VXLAN隧道的源端地址和源接口 采用OVSDB对VTEP设备进行部署和控制时，不能执行本配置 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
（可选）开启隧道的BFD检测功能	**tunnel bfd enable destination-mac** *mac-address*	缺省情况下，隧道的BFD检测功能处于关闭状态 执行本命令的同时，需要在系统视图下执行**reserved vxlan**命令配置保留VXLAN。否则，BFD会话无法up 本命令不能与uRPF功能同时配置，否则，BFD会话无法up。关于uRPF功能的详细介绍请参见“安全配置指导”中的“uRPF”
（可选）退回系统视图	**quit**	-
（可选）配置保留VXLAN	**reserved vxlan** *vxlan-id*	缺省情况下，未指定保留VXLAN 只能在系统视图下配置一个全局保留VXLAN，该VXLAN不能与VSI下创建的VXLAN相同 配置的保留VXLAN不能与**mapping vni**命令配置的映射远端VXLAN相同。**mapping vni**命令的详细介绍，请参见“EVPN命令参考”中的“EVPN”

2.5 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

配置 VXLAN 与 VXLAN 隧道关联时：

- 如果指定了 **backup-tunnel** *tunnel-number* 参数，则该参数指定的隧道作为备用 VXLAN 隧道，为主用 VXLAN 隧道提供保护。当主用 VXLAN 隧道 down 时，VXLAN 将启用备用 VXLAN 隧道。
- 如果指定了 **relay-agent ipoe** 参数，则在隧道上开启 IPoE 报文的代理透传功能，即从 VXLAN 内接收到的 IPoE 报文均通过该隧道转发给远端 VTEP。

  隧道的代理透传功能通常应用在转发控制分离组网中：在 DP 上为 DP 与 CP 之间的 VXLAN 隧道开启代理透传功能，以便将控制流量（IPoE 报文）发送给 CP。有关转发控制分离的介绍，请参见《vBRAS 系列虚拟宽带远程接入服务器配置指导手册》。

  转控分离特性不能和 ARP/ND 攻击防御以及 DHCP/DHCPv6 攻击防御功能（例如 DHCP/DHCPv6 Snooping 信任端口特性）一起使用。

  开启 IPoE 报文的代理透传功能时，具有如下配置限制：
  - SH 系列接口板不支持该功能。
  - 该功能不支持 VXLAN-DCI 隧道。
  - 在 CP 和 DP 上必须为关联同一个 VXLAN 的 VSI 虚接口配置相同的 MAC 地址。
  - 设备的所有公网侧端口上都需要配置 **undo mac-address static source-check enable** 命令。
  - 设备将不支持 DHCP/DHCPv6 相关功能。
  - VSI 将由 CP 下发 ARP/ND 表项，本身不再支持 ARP/ND 学习（ARP/ND 泛洪抑制、ARP/ND Detection、本地代理 ARP/本地 ND Proxy 等相关功能也不支持）。不建议在用户所属的 VXLAN 部署应用服务器。
  - 如果 VSI 对应的 VXLAN 关联了多个 VXLAN 隧道，那么该 VSI 转发 ARP/ND 广播报文或 DHCP/DHCPv6 广播报文时，会转发到所有这些隧道。
  - 不支持出方向的 VSI 报文统计。
  - VSI 关联的 AC 仅支持如下报文匹配规则：
    - 匹配单个外层 VLAN 标签（**encapsulation s-vid** *vlan-id* [**only-tagged**]）
    - 匹配外层 VLAN 标签+内层 VLAN 标签（**encapsulation s-vid** { *vlan-id* | *vlan-id-list* } **c-vid** { *vlan-id-list* | **all** }）

3. 配置步骤

表2-4 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*	-
配置VXLAN与VXLAN隧道关联	**tunnel** { *tunnel-number* [**backup-tunnel** *tunnel-number* | **relay-agent ipoe**] | **all** }	缺省情况下，VXLAN未关联VXLAN隧道 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

2.6 建立数据帧与VSI的关联

2.6.1 配置限制和指导

开启 VLAN 关联 VXLAN 功能后，不能再手工或动态创建以太网服务实例，如需创建以太网服务实例，请先执行 **undo vxlan vlan-based** 命令关闭 VLAN 关联 VXLAN 功能后再创建；反之，手工或动态创建以太网服务实例后，不能再开启 VLAN 关联 VXLAN 功能，如需开启该功能，请先删除所有的以太网服务实例后再开启。

2.6.2 配置手工创建的以太网服务实例与 VSI 关联

1. 功能简介

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

对于使用 Ethernet 接入模式的以太网服务实例，可以修改该以太网服务实例上入方向和出方向报文的 VLAN 标签。修改方式包括添加标签、映射标签和剥离标签。关于修改入方向/出方向报文 VLAN 标签的配置限制和指导，请参见“VXLAN 命令参考”中的 **rewrite inbound tag** 和 **rewrite outbound tag** 命令。

2. 配置限制和指导

注意

- 请不要在同一端口下同时配置 VLAN 映射和配置以太网服务实例匹配多个 VLAN 标签（**encapsulation** 命令中 *vlan-id-list* 参数指定多个 VLAN）；否则会导致报文转发异常，必须删除冲突配置并重启端口所在接口板才能恢复业务。有关 VLAN 映射的介绍，请参见“二层技术-以太网交换配置指导”中的“VLAN 映射”。
- **encapsulation** 命令中匹配的外层 VLAN 不能再配置其他业务（包括二层、三层业务）。

为确保转发正常，端口上以太网服务实例的报文匹配规则需要与该端口上允许通过的 VLAN、VLAN 报文是否带 Tag 配置保持一致。当端口上以太网服务实例的报文匹配规则为 **encapsulation { default | tagged | untagged }**时，该端口需要允许缺省 VLAN 通过。

当接入模式为 VLAN 时，如果端口接收到的报文不带 Tag，需要配置报文匹配规则为 **encapsulation untagged**。

存在以下情况时，请用 **xconnect vsi** 命令指定接入模式为 Ethernet。

- 以太网服务实例匹配了多个外层或内层 VLAN 标签（**encapsulation** 命令中 *vlan-id-list* 参数指定了多个 VLAN）。
- 以太网服务实例采用缺省的报文匹配规则（**encapsulation default**）。
- 以太网服务实例匹配携带 VLAN 标签的报文（**encapsulation tagged**）。

当以太网服务实例匹配了多个 VLAN 标签时（**encapsulation** 命令中 *vlan-id-list* 参数指定了多个 VLAN），需要注意：

- 该以太网服务实例只匹配携带 VLAN 标签的报文。
- 该以太网服务实例所在端口不支持 802.1X、MAC 地址认证或端口安全功能。有关 802.1X、MAC 地址认证和端口安全的介绍，请参见“安全配置指导”。
- 如果在以太网服务实例上修改入方向/出方向报文的 VLAN 标签（**rewrite inbound tag** 或 **rewrite outbound tag** 命令），请不要在该以太网服务实例或它的所在端口上配置 DHCP 相关功能，否则会导致功能异常。有关 DHCP 的介绍，请参见“三层技术-IP 业务配置指导”中的“DHCP 概述”。

仅当以太网服务实例的报文匹配规则为 **encapsulation s-vid** *vlan-id* [**only-tagged**]时，它的所在端口支持 MAC 地址认证功能。

3. 配置步骤

表2-5 配置手工创建的以太网服务实例与 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*	-
将二层接口加入本地站点VLAN	配置端口的链路类型	**port link-type** { **access** | **trunk** | **hybrid** }	缺省情况下，所有端口的链路类型均为Access类型
	将当前端口加入本地站点VLAN	• **port access vlan** *vlan-id* • **port trunk permit vlan** { *vlan-id-list* | **all** } • **port hybrid vlan** *vlan-id-list* { **tagged** | **untagged** }	三者选其一 本地站点VLAN必须是设备上已创建的VLAN
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，不存在以太网服务实例
配置以太网服务实例的报文匹配规则		• **encapsulation s-vid** { *vlan-id* | *vlan-id-list* } [**only-tagged**] • **encapsulation s-vid** { *vlan-id* | *vlan-id-list* } **c-vid** { *vlan-id-list* | **all** } • **encapsulation c-vid** { *vlan-id* | *vlan-id-list* } • **encapsulation** { **default** | **tagged** | **untagged** }	缺省情况下，未配置报文匹配规则
（可选）配置入方向报文的处理规则		**rewrite inbound tag** { **nest** { **c-vid** *vlan-id* | **s-vid** *vlan-id* [**c-vid** *vlan-id*] } | **remark** { { **1-to-1** | **2-to-1** } { **c-vid** *vlan-id* | **s-vid** *vlan-id* } | { **1-to-2** | **2-to-2** } **s-vid** *vlan-id* **c-vid** *vlan-id* } | **strip** { **c-vid** | **s-vid** [**c-vid**] } } [**symmetric**]	缺省情况下，不对入方向报文进行处理
（可选）配置出方向报文的处理规则		**rewrite outbound tag** { **nest** { **c-vid** *vlan-id* | **s-vid** *vlan-id* [**c-vid** *vlan-id*] } | **remark** { { **1-to-1** | **2-to-1** } { **c-vid** *vlan-id* | **s-vid** *vlan-id* } | { **1-to-2** | **2-to-2** } **s-vid** *vlan-id* **c-vid** *vlan-id* } | **strip** { **c-vid** | **s-vid** [**c-vid**] } }	缺省情况下，不对出方向报文进行处理

操作	命令	说明
将以太网服务实例与VSI关联	**xconnect vsi** *vsi-name* [**access-mode** { **ethernet** | **vlan** }] [**track** *track-entry-number*&<1-3>]	缺省情况下，以太网服务实例未关联VSI 配置本功能时，如果不指定**access-mode**参数，将采用缺省接入模式VLAN

2.6.3 配置动态创建的以太网服务实例与 VSI 关联

1. 功能简介

802.1X 或 MAC 地址认证为用户下发授权 VSI、Guest VSI、Auth-Fail VSI 或 Critical VSI 后，将用户信息（接入端口、所属 VLAN、MAC 地址）及 VSI 信息通知给 VXLAN。VXLAN 根据用户信息动态创建以太网服务实例，并将其与 VSI 关联。802.1X 和 MAC 地址认证的详细介绍，请参见“安全配置指导”中的“802.1X”和“MAC 地址认证”。

动态创建的以太网服务实例仅支持通过 MAC 地址方式判断接口接收到的报文是否属于该 AC：检查报文携带的源 MAC 地址是否与以太网服务实例匹配的 MAC 地址相同。只有二者相同，报文才属于该 AC。

2. 配置限制和指导

配置本功能时，必须采用 MAC 地址认证或基于 MAC 接入控制的 802.1X 认证，并开启动态创建的以太网服务实例匹配 MAC 地址功能。

配置本功能时，需要注意：

- 用户侧端口必须位于 FD/SG/SH 系列接口板或 LSUM1CQGS32SF0-Z 接口板上。
- 配置了 **mac-based ac** 命令的端口不支持源 MAC 地址相同、VLAN 不同的用户同时接入。
- 用户侧端口仅支持链路类型为 Access 或 Trunk。
- 如果用户侧端口接收到的报文不带 Tag 或带的 Tag 等于端口缺省 VLAN，则该端口仅支持转发不带 Tag 的报文。
- 如果要同时配置以太网服务实例的报文匹配规则，仅支持 **encapsulation s-vid** *vlan-id* [**only-tagged**]命令。
- 同一接口上源 MAC 地址相同的报文只能匹配到一个动态创建的以太网服务实例。
- 二层聚合接口的成员端口上无法动态创建以太网服务实例。

3. 配置步骤

802.1X 或 MAC 地址认证中，接入认证设备上配置了 Guest VSI、Auth-Fail VSI、Critical VSI，或远程 AAA 服务器为认证成功用户下发了授权 VSI，则接入认证设备上会自动地创建以太网服务实例，并将其与 Guest VSI、Auth-Fail VSI、Critical VSI 或授权 VSI 关联。

在接入认证设备上开启动态创建的以太网服务实例匹配 MAC 地址功能。

表2-6 开启动态创建的以太网服务实例匹配 MAC 地址功能

操作	命令	说明
进入系统视图	**system-view**	-

操作		命令	说明
进入用户侧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*	
开启动态创建的以太网服务实例匹配MAC地址功能		**mac-based ac**	缺省情况下，动态创建的以太网服务实例匹配MAC地址功能处于关闭状态，即动态创建的以太网服务实例只匹配VLAN

2.6.4 配置 VLAN 与 VXLAN 关联

1. 功能简介

开启 VLAN 关联 VXLAN 功能，并在 VLAN 视图下配置与该 VLAN 关联的 VXLAN 后，如果存在属于该 VLAN 的接口，则自动在该接口上创建编号为当前 VLAN ID、匹配外层 VLAN tag 为当前 VLAN ID 的以太网服务实例，并将该以太网服务实例与指定 VXLAN 对应的 VSI 关联，从而确保属于该 VLAN 的数据帧均通过指定的 VSI 转发。

2. 配置限制和指导

将 VLAN 与 VXLAN 关联后，该 VLAN 内将不能进行普通的二层转发，该 VLAN 对应的 VLAN 接口也不能进行三层转发。

与 VXLAN 关联的 VLAN 数目、允许这些 VLAN 通过的 Trunk 类型的端口数目较多时，AC 创建和删除过程可能会耗费一定的时间，VXLAN、EVPN 等相关操作需要等待 AC 创建、删除完成后才会响应。

3. 配置准备

本配置中指定的与 VLAN 关联的 VXLAN 需要通过 **vxlan** 命令创建。

4. 配置步骤

操作	命令	说明
进入系统视图	**system-view**	-
开启VLAN关联VXLAN功能	**vxlan vlan-based**	缺省情况下，VLAN关联VXLAN功能处于关闭状态
进入VLAN视图	**vlan** *vlan-id*	本配置中指定的VLAN不能为VLAN 1
配置VLAN与指定的VXLAN关联	**vxlan vni** *vxlan-id*	缺省情况下，未指定与VLAN关联的VXLAN 本配置中指定的VXLAN ID不能为EVPN组网中的L3VNI

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。

2.7.1 开启本地 MAC 地址的日志记录功能

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）。

表2-7 配置增删本地 MAC 地址时记录日志

操作	命令	说明
进入系统视图	**system-view**	-
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

2.7.2 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，不存在静态的远端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，否则配置将失败

说明

请不要为 EVPN 动态创建的隧道配置静态远端 MAC 地址表项，避免出现如下问题：如果公网侧接口 down，设备将删除已创建的隧道，同时删除为该隧道配置的静态远端 MAC 地址表项，公网侧接口重新 up 后会自动重新建立隧道，但是无法恢复静态远端 MAC 地址表项；如果执行了配置回滚操作，设备会重新创建隧道，新创建的隧道编号可能发生变化，造成配置回滚失败。

2.7.3 关闭远端 MAC 地址自动学习功能

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。

表2-9 关闭远端 MAC 地址自动学习功能

操作	命令	说明
进入系统视图	**system-view**	-
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 配置接口的 MAC 地址软件学习功能

本功能适用于 SDN（Software Defined Network，软件定义网络）组网。

在 SDN 组网中，设备将接口学习到的 MAC 地址上传给控制器，控制器把收到的 MAC 地址下发给其它远端设备，以减少不必要的广播流量。

接口的 MAC 地址学习方式包括：

- 硬件学习：接口通过硬件学习 MAC 地址。软件周期性地检查硬件是否学习到新的 MAC 地址，把学到的地址上传控制器处理。硬件学习方式下需等待软件检查周期的到来，控制器获取 MAC 地址的速度较慢。
- 软件学习：接口通过软件学习 MAC 地址。软件把学到的 MAC 地址下发给硬件，同时上传控制器处理。软件学习方式下不需等待软件检查周期的到来，控制器获取 MAC 地址的速度较快。

需要注意的是，开启接口的MAC地址软件学习功能后，大量的MAC地址学习可能对系统造成冲击，不建议用户在大量 MAC 地址频繁变化的情况下开启本功能。

表2-10 配置接口的 MAC 地址软件学习功能

操作	命令	说明
进入系统视图	**system-view**	-
进入二层以太网接口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*	-
开启接口的MAC地址软件学习功能	**l2vpn mac-address software-learning enable**	缺省情况下，接口的MAC地址软件学习功能处于关闭状态，设备采用硬件方式学习MAC地址

2.8 配置VXLAN组播路由泛洪方式

组播路由泛洪方式支持如下两种实现模式：

- PIM 模式：在 VTEP 和核心设备上运行 PIM 协议，以建立组播转发表项。采用该模式时，可以使用 Loopback 接口地址作为组播报文的源 IP 地址。当 VTEP 存在多个网络侧接口时，PIM 协议可以动态选择报文的出接口。
- IGMP 主机模式：在 VTEP 上开启 IGMP 协议的主机功能、在连接 VTEP 的核心设备上配置 IGMP、在所有核心设备上运行 PIM 协议，以建立组播转发表项。当 VTEP 存在多个网络侧接口时，IGMP 主机模式只能采用组播报文的源 IP 地址所在的接口作为报文的出接口。

同一 VXLAN 网络中的不同 VTEP 可以采用不同的实现模式。

2.8.1 配置准备

- 在 VTEP 和核心设备上使能 IP 组播路由功能。
- 在核心设备上配置组播路由协议。由于 VTEP 同时作为组播源和组播接收者，因此推荐使用双向 PIM 作为组播路由协议。
- VXLAN 网络中存在采用 IGMP 主机模式的 VTEP 时，需要在连接该 VTEP 的核心设备上配置 IGMP。

2.8.2 配置 PIM 模式

表2-11 配置 PIM 模式

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 可以使用Loopback接口地址作为组播报文的源IP地址 为确保组播报文转发正常，VXLAN组播报文的源IP地址（*source-address*）需要指定为一个已创建且处于up状态的VXLAN隧道的源端地址
进入接口视图	**interface** *interface-type interface-number*	Loopback接口和与核心设备相连的接口上均需要使能PIM协议
在接口上使能PIM协议	**pim sm**	二者选其一 缺省情况下，接口上PIM协议处于关闭状态
	pim dm	

2.8.3 配置 IGMP 主机模式

采用该模式时，必须使用 VTEP 上网络侧接口的 IP 地址作为组播报文的源 IP 地址，并在该接口上开启 IGMP 协议的主机功能。

表2-12 配置 IGMP 主机模式

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 必须使用VTEP上网络侧接口的IP地址作为组播报文的源IP地址
进入与核心设备相连接口的接口视图	**interface** *interface-type interface-number*	-
在接口上开启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 地址为广播、未知单播和未知组播的数据帧后，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；VTEP 从 VXLAN 隧道接收到目的 MAC 地址为广播、未知单播和未知组播的数据帧后，会在该 VXLAN 内的所有本地接口上泛洪该数据帧。通过本配置可以手工禁止某类数据帧在 VXLAN 内泛洪，以减少网络中的泛洪流量。

禁止过 VXLAN 隧道向远端站点泛洪后，为了将某些单播或组播 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2-13 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** { **all** | { **broadcast** | **unknown-multicast** | **unknown-unicast** } * } [**all-direction** | **dci**]	缺省情况下，VSI泛洪功能处于开启状态 数据中心互联场景中，如果只想禁止向数据中心间的VXLAN-DCI隧道泛洪，而数据中心内部的VXLAN隧道可以泛洪，请指定**dci**参数

操作	命令	说明
（可选）配置VSI选择性泛洪的MAC地址	**selective-flooding mac-address** *mac-address*	缺省情况下，不存在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2-14 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 报文内层封装的以太网数据帧是否携带 VLAN Tag 进行检查：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：远端 VTEP 上通过 **xconnect vsi** 命令的 **access-mode** 参数配置接入模式为 **ethernet** 时，VXLAN 报文可能携带 VLAN Tag。这种情况下建议不要在本端 VTEP 上执行 **vxlan invalid-vlan-tag discard** 命令，以免错误地丢弃报文。

表2-15 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：

- 当同时执行 **flooding disable** 命令关闭了 VSI 的泛洪功能时：
  - 如果要与远端站点互通，则两端 VTEP 都需要为对端站点添加静态远端 MAC 地址表项（相关命令为 **mac-address static**）。

- 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。
- 如果用户终端通过 DHCP 获取 IP 地址，为使 DHCP 请求报文能在 VSI 内泛洪，需要在 VTEP 上配置 VXLAN 的 AC 链路为 DHCP Snooping 信任端口（**dhcp snooping trust**），并配置 VXLAN 隧道接口为 DHCP Snooping 信任接口（**dhcp snooping trust tunnel**）。相关配置请参见“三层技术-IP 业务配置指导”中的“DHCP Snooping”。

- VLAN 接入模式下，如果以太网服务实例的报文匹配规则为 **encapsulation s-vid** *vlan-id*，且匹配的 VLAN 为接口的缺省 VLAN，则 VTEP 对匹配 ARP 泛洪抑制表项的 ARP 请求进行应答时，ARP 应答报文的 VLAN tag 将被删除后再转发给用户终端。如果用户终端要求接收到的 ARP 应答报文携带 VLAN tag，则会导致该用户终端无法学习到 ARP 表项。在这种情况下，建议不要将以太网服务实例匹配的 VLAN 指定为接口的缺省 VLAN。
- 当 VXLAN 网络采用组播路由（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；
  - 如果需要和其他厂商的 VTEP 设备互通，则不能使用 ARP 泛洪抑制功能。

表2-16 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 配置ND泛洪抑制

1. 配置步骤

表2-17 配置 ND 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ND泛洪抑制功能	**ipv6 nd suppression enable**	缺省情况下，ND泛洪抑制功能处于关闭状态

2.14 关闭VXLAN远端ARP/ND自动学习功能

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用户终端的 ARP/ND 信息，即远端 ARP/ND 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/ND 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/ND 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2-18 关闭远端 ARP/ND 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
关闭远端ND自动学习功能	**vxlan tunnel nd-learning disable**	缺省情况下，远端ND自动学习功能处于开启状态

2.15 配置VXLAN流量统计

2.15.1 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2-19 配置 VSI 的报文统计功能

操作	命令	说明
进入系统视图	**system-view**	-
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

2.15.2 配置 AC 的报文统计功能

1. 配置限制和指导

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，则报文统计重新开始。

2. 配置以太网服务实例的报文统计功能

表2-20 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**	-
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*	-

操作	命令	说明
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

3. 配置 VLAN 下对应 AC 的报文统计功能

操作	命令	说明
进入系统视图	**system-view**	-
进入VLAN视图	**vlan** *vlan-id*	-
开启VLAN下对应AC的报文统计功能	**ac statistics enable**	缺省情况下，VLAN下对应AC的报文统计功能处于关闭状态 本功能用来对VLAN与VXLAN关联方式下自动生成的AC进行报文统计。开启本功能前，必须先执行**vxlan vlan-based**命令开启VLAN关联VXLAN功能

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2-21 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** | **verbose**]
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VSI的ND泛洪抑制表项信息（独立运行模式）	**display ipv6 nd suppression vsi** [**name** *vsi-name*] [**slot** *slot-number*] [**count**]
显示VSI的ND泛洪抑制表项信息（IRF模式）	**display ipv6 nd suppression vsi** [**name** *vsi-name*] [**chassis** *chassis-number* **slot** *slot-number*] [**count**]
显示VXLAN关联的VXLAN隧道信息	**display vxlan tunnel** [**vxlan-id** *vxlan-id*]

操作	命令
清除VSI的ARP泛洪抑制表项	**reset arp suppression vsi** [**name** *vsi-name*]
清除VSI的ND泛洪抑制表项	**reset ipv6 nd 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AC的报文统计信息	**reset l2vpn statistics ac** [**interface** *interface-type interface-number* **service-instance** *instance-id*]


说明

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

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

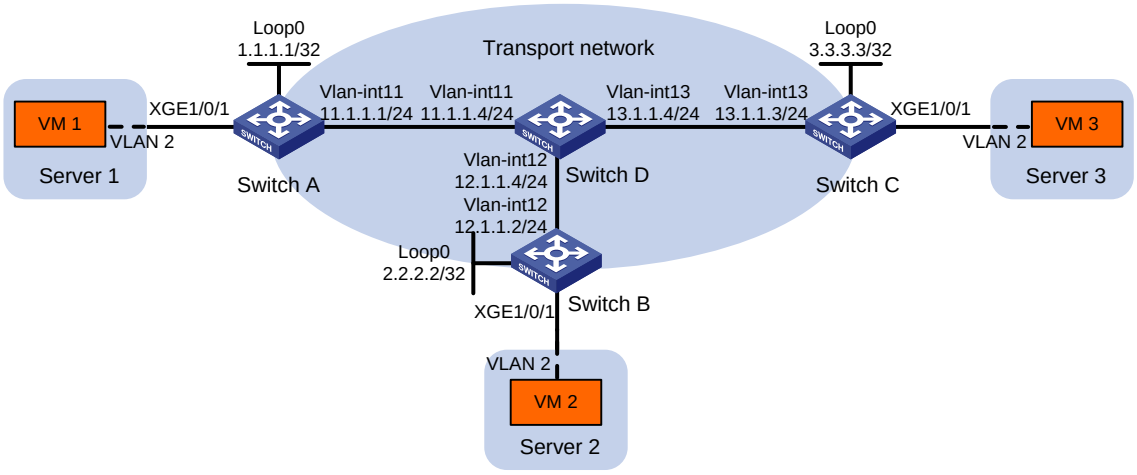
Switch A、Switch B、Switch C 为与服务器连接的 VTEP 设备。虚拟机 VM 1、VM 2 和 VM 3 同属于 VXLAN 10。通过 VXLAN 实现不同站点间的二层互联，确保虚拟机在站点之间进行迁移时用户的访问流量不会中断。

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
```

创建 VSI 实例 vpna 和 VXLAN 10。

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

在接入服务器的接口 Ten-GigabitEthernet1/0/1 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 ten-gigabitethernet 1/0/1
[SwitchA-Ten-GigabitEthernet1/0/1] port link-type trunk
[SwitchA-Ten-GigabitEthernet1/0/1] port trunk permit vlan 2
[SwitchA-Ten-GigabitEthernet1/0/1] service-instance 1000
[SwitchA-Ten-GigabitEthernet1/0/1-srv1000] encapsulation s-vid 2
```

配置以太网服务实例 1000 与 VSI 实例 vpna 关联。

```
[SwitchA-Ten-GigabitEthernet1/0/1-srv1000] xconnect vsi vpna
[SwitchA-Ten-GigabitEthernet1/0/1-srv1000] quit
[SwitchA-Ten-GigabitEthernet1/0/1] quit
```

(3) 配置 Switch B

开启 L2VPN 能力。

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

创建 VSI 实例 vpna 和 VXLAN 10。

```
[SwitchB] vsi vpna
[SwitchB-vsi-vpna] vxlan 10
[SwitchB-vsi-vpna-vxlan-10] 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 loopback 0
[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-vxlan-10] tunnel 2
[SwitchB-vsi-vpna-vxlan-10] tunnel 3
[SwitchB-vsi-vpna-vxlan-10] quit
[SwitchB-vsi-vpna] quit
```

在接入服务器的接口 Ten-GigabitEthernet1/0/1 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 ten-gigabitethernet 1/0/1
[SwitchB-Ten-GigabitEthernet1/0/1] port link-type trunk
[SwitchB-Ten-GigabitEthernet1/0/1] port trunk permit vlan 2
[SwitchB-Ten-GigabitEthernet1/0/1] service-instance 1000
[SwitchB-Ten-GigabitEthernet1/0/1-srv1000] encapsulation s-vid 2
```

配置以太网服务实例 1000 与 VSI 实例 vpna 关联。

```
[SwitchB-Ten-GigabitEthernet1/0/1-srv1000] xconnect vsi vpna
[SwitchB-Ten-GigabitEthernet1/0/1-srv1000] quit
[SwitchB-Ten-GigabitEthernet1/0/1] quit
```

(4) 配置 Switch C

开启 L2VPN 能力。

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

创建 VSI 实例 vpna 和 VXLAN 10。

```
[SwitchC] vsi vpna
[SwitchC-vsi-vpna] vxlan 10
[SwitchC-vsi-vpna-vxlan-10] 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 loopback 0
[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-vxlan-10] tunnel 1
[SwitchC-vsi-vpna-vxlan-10] tunnel 3
[SwitchC-vsi-vpna-vxlan-10] quit
[SwitchC-vsi-vpna] quit
```

在接入服务器的接口 Ten-GigabitEthernet1/0/1 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 ten-gigabitethernet 1/0/1
[SwitchC-Ten-GigabitEthernet1/0/1] port link-type trunk
[SwitchC-Ten-GigabitEthernet1/0/1] port trunk permit vlan 2
[SwitchC-Ten-GigabitEthernet1/0/1] service-instance 1000
[SwitchC-Ten-GigabitEthernet1/0/1-srv1000] encapsulation s-vid 2
```

配置以太网服务实例 1000 与 VSI 实例 vpna 关联。

```
[SwitchC-Ten-GigabitEthernet1/0/1-srv1000] xconnect vsi vpna
[SwitchC-Ten-GigabitEthernet1/0/1-srv1000] quit
[SwitchC-Ten-GigabitEthernet1/0/1] 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: 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: vpna
  VSI Index               : 0
  VSI State               : Up
  MTU                     : 1500
  Bandwidth               : -
  Broadcast Restrain      : -
  Multicast Restrain      : -
  Unknown Unicast Restrain: -
  MAC Learning            : Enabled
  MAC Table Limit         : -
  MAC Learning rate       : -
  Drop Unknown            : -
  Flooding                : Enabled
  VXLAN ID                : 10
```

```
  Tunnels:
    Tunnel Name          Link ID    State  Type        Flood proxy
    Tunnel1              0x5000001  Up     Manual      Disabled
    Tunnel2              0x5000002  Up     Manual      Disabled
  ACs:
    AC                               Link ID  State       Type
    XGE1/0/1 srv1000                 0        Up          Manual
```

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
dc2d-cb9c-6cdb   Dynamic  vpna                            Tunnel1       Aging
dc2d-cb9c-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

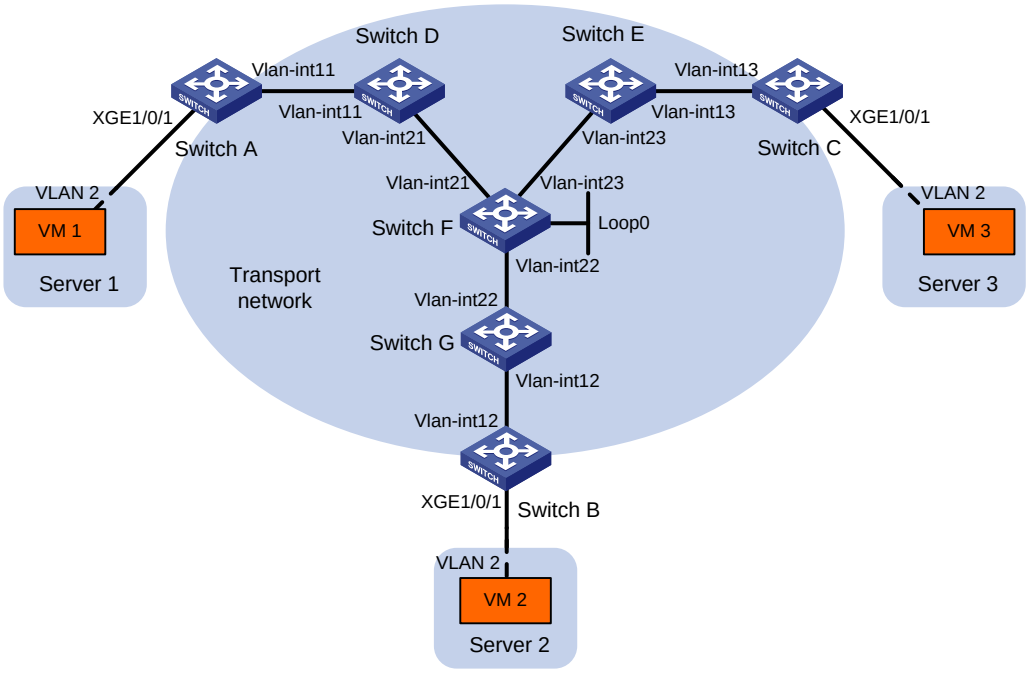
Switch A、Switch B、Switch C 为与服务器连接的 VTEP 设备。虚拟机 VM 1、VM 2 和 VM 3 同属于 VXLAN 10。通过 VXLAN 实现不同站点间的二层互联，确保虚拟机在站点之间进行迁移时用户的访问流量不会中断。

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
```

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

配置接口 Vlan-interface11 的 IP 地址，并在该接口上开启 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-vxlan-10] tunnel 1
[SwitchA-vsi-vpna-vxlan-10] 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-vxlan-10] group 225.1.1.1 source 11.1.1.1
[SwitchA-vsi-vpna-vxlan-10] quit
[SwitchA-vsi-vpna] quit
```

在接入服务器的接口 Ten-GigabitEthernet1/0/1 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 ten-gigabitethernet 1/0/1
[SwitchA-Ten-GigabitEthernet1/0/1] port link-type trunk
[SwitchA-Ten-GigabitEthernet1/0/1] port trunk permit vlan 2
[SwitchA-Ten-GigabitEthernet1/0/1] service-instance 1000
[SwitchA-Ten-GigabitEthernet1/0/1-srv1000] encapsulation s-vid 2
```

配置以太网服务实例 1000 与 VSI 实例 vpna 关联。

```
[SwitchA-Ten-GigabitEthernet1/0/1-srv1000] xconnect vsi vpna
[SwitchA-Ten-GigabitEthernet1/0/1-srv1000] quit
[SwitchA-Ten-GigabitEthernet1/0/1] quit
```

(3) 配置 Switch B

开启 L2VPN 能力。

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

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-vxlan-10] quit
[SwitchB-vsi-vpna] quit
```
配置接口 Vlan-interface12 的 IP 地址，并在该接口上开启 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-vxlan-10] tunnel 2
[SwitchB-vsi-vpna-vxlan-10] 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-vxlan-10] group 225.1.1.1 source 12.1.1.2
[SwitchB-vsi-vpna-vxlan-10] quit
[SwitchB-vsi-vpna] quit
```
在接入服务器的接口 Ten-GigabitEthernet1/0/1 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 ten-gigabitethernet 1/0/1
[SwitchB-Ten-GigabitEthernet1/0/1] port link-type trunk
[SwitchB-Ten-GigabitEthernet1/0/1] port trunk permit vlan 2
[SwitchB-Ten-GigabitEthernet1/0/1] service-instance 1000
[SwitchB-Ten-GigabitEthernet1/0/1-srv1000] encapsulation s-vid 2
```
配置以太网服务实例 1000 与 VSI 实例 vpna 关联。
```
[SwitchB-Ten-GigabitEthernet1/0/1-srv1000] xconnect vsi vpna
[SwitchB-Ten-GigabitEthernet1/0/1-srv1000] quit
[SwitchB-Ten-GigabitEthernet1/0/1] quit
```
(4) 配置 Switch C

开启 L2VPN 能力。
```
<SwitchC> system-view
[SwitchC] l2vpn enable
```

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

配置接口 Vlan-interface13 的 IP 地址，并在该接口上开启 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-vxlan-10] tunnel 1
[SwitchC-vsi-vpna-vxlan-10] 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-vxlan-10] group 225.1.1.1 source 13.1.1.3
[SwitchC-vsi-vpna-vxlan-10] quit
[SwitchC-vsi-vpna] quit
```

在接入服务器的接口 Ten-GigabitEthernet1/0/1 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 ten-gigabitethernet 1/0/1
[SwitchC-Ten-GigabitEthernet1/0/1] port link-type trunk
[SwitchC-Ten-GigabitEthernet1/0/1] port trunk permit vlan 2
[SwitchC-Ten-GigabitEthernet1/0/1] service-instance 1000
[SwitchC-Ten-GigabitEthernet1/0/1-srv1000] encapsulation s-vid 2
```

配置以太网服务实例 1000 与 VSI 实例 vpna 关联。

```
[SwitchC-Ten-GigabitEthernet1/0/1-srv1000] xconnect vsi vpna
[SwitchC-Ten-GigabitEthernet1/0/1-srv1000] quit
[SwitchC-Ten-GigabitEthernet1/0/1] 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
```

使能双向 PIM。

```
[SwitchD] pim
[SwitchD-pim] bidir-pim enable
[SwitchD-pim] 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
```

使能双向 PIM。

```
[SwitchE] pim
[SwitchE-pim] bidir-pim enable
[SwitchE-pim] 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
[SwitchF] interface loopback 0
[SwitchF-LoopBack0] pim sm
[SwitchF-LoopBack0] quit
```

使能双向 PIM。

```
[SwitchF] pim
[SwitchF-pim] bidir-pim enable
```

将接口 Vlan-interface22 配置为 C-BSR，并将接口 Loopback0 配置为服务于双向 PIM 的 C-RP。

```
[SwitchF-pim] c-bsr 22.1.1.6
[SwitchF-pim] c-rp 6.6.6.6 bidir
[SwitchF-pim] 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
```

使能双向 PIM。

```
[SwitchG] pim
[SwitchG-pim] bidir-pim enable
[SwitchG-pim] 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: 64kbps
Maximum transmission unit: 1464
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
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: vpna
  VSI Index               : 0
  VSI State               : Up
  MTU                     : 1500
  Bandwidth               : -
  Broadcast Restrain      : -
  Multicast Restrain      : -
  Unknown Unicast Restrain: -
  MAC Learning            : Enabled
  MAC Table Limit         : -
  MAC Learning rate       : -
  Drop Unknown            : -
  Flooding                : Enabled
  VXLAN ID                : 10
  Tunnels:
    Tunnel Name          Link ID    State  Type        Flood proxy
    Tunnel1              0x5000001  Up     Manual      Disabled
    Tunnel2              0x5000002  Up     Manual      Disabled
    MTunnel0             0x6000000  Up     Auto        Disabled
  ACs:
    AC                              Link ID  State       Type
    XGE1/0/1 srv1000                0        Up          Manual
```

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
dc2d-cb9c-6cdb   Dynamic  vpna                            Tunnel1   Aging
dc2d-cb9c-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。

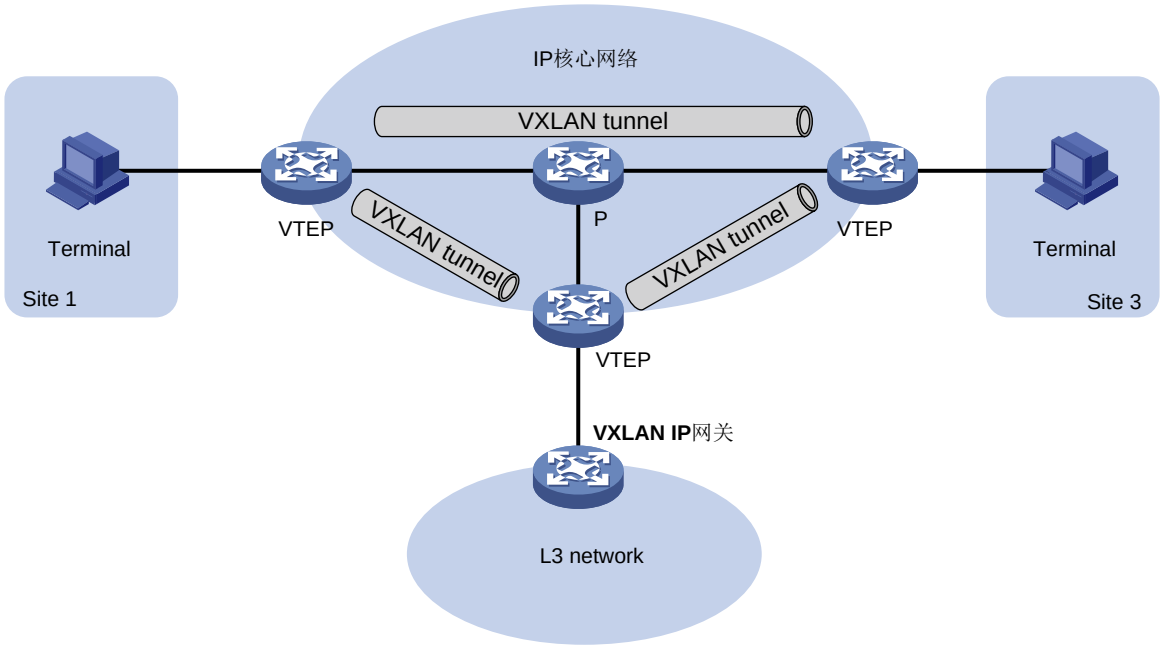
3 VXLAN IP 网关

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 站点内的用户终端提供三层业务，则需要在网络中部署 VXLAN IP 网关，以便站点内的用户终端通过 VXLAN IP 网关与外界网络或其他 VXLAN 网络内的用户终端进行三层通信。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

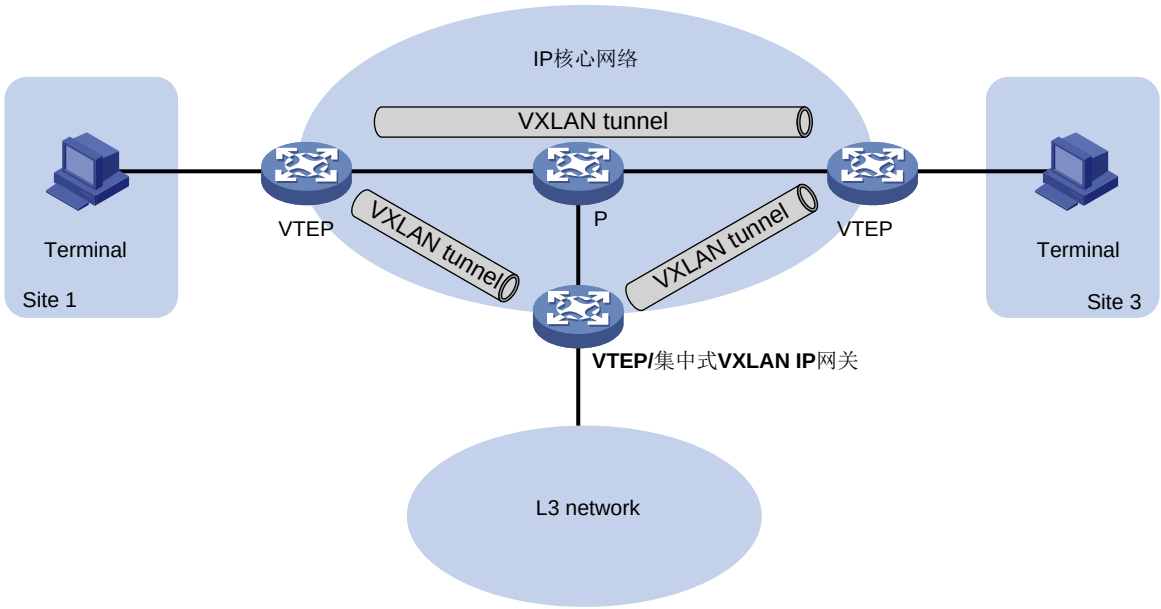
图3-1 独立的 VXLAN IP 网关示意图



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。用户终端通过 VXLAN IP 网关与三层网络中的节点通信时，用户终端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。

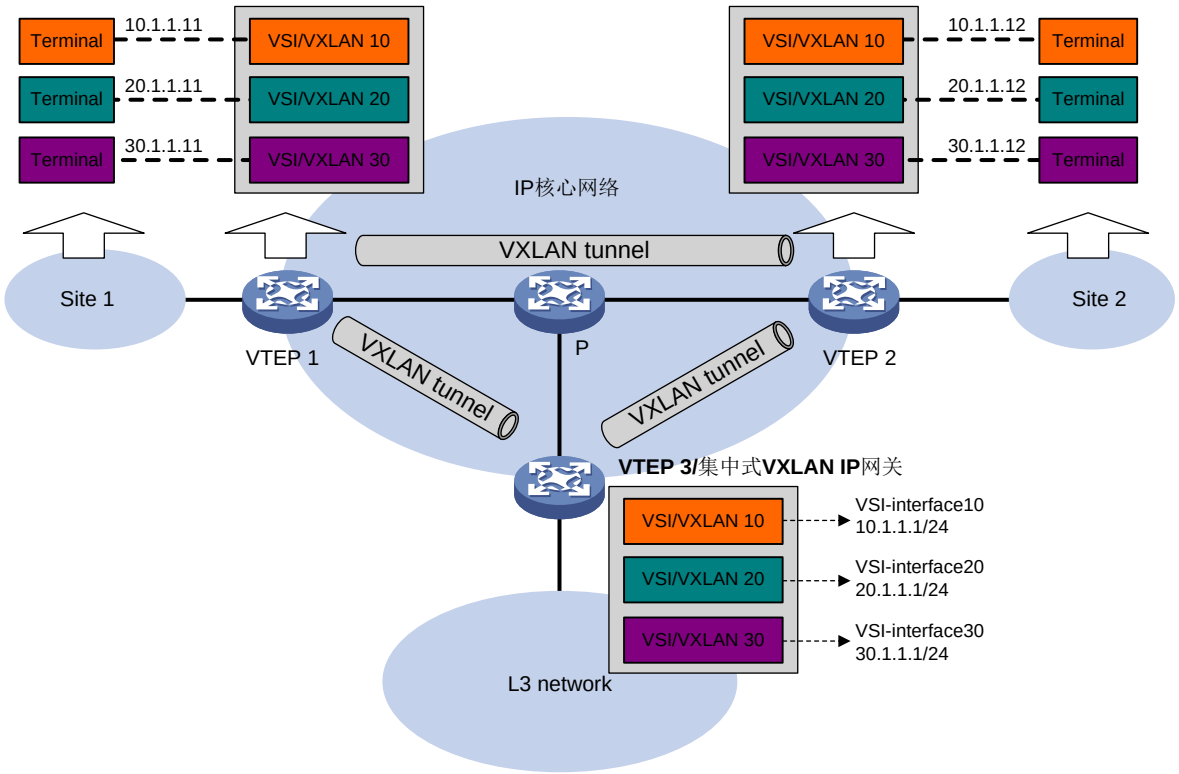
3.1.2 集中式 VXLAN IP 网关

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 的用户终端为例，用户终端与外界网络进行三层通信的过程为：

(1) 用户终端（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，并向用户终端回应 ARP 应答消息。

(4) VTEP 1 收到 ARP 应答消息后，将该消息转发给用户终端。

(5) 用户终端获取到网关的 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。

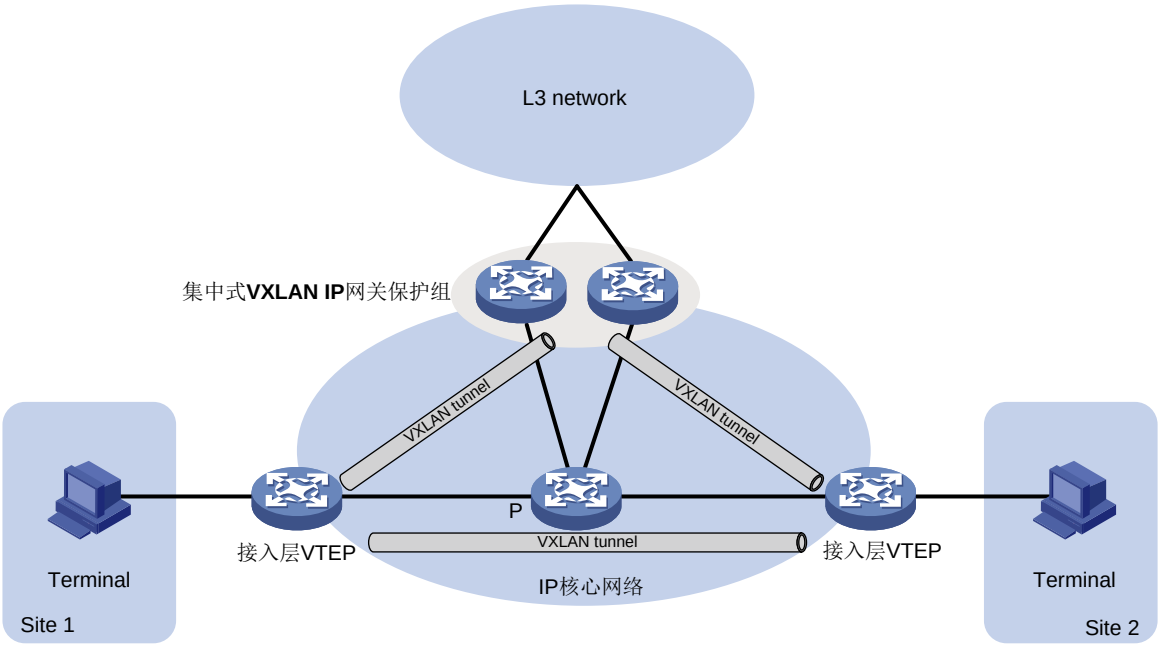
(7) 目的节点回复的报文到达网关后，网关根据已经学习到的 ARP 表项，为报文封装链路层头，并通过 VXLAN 网络将其发送给用户终端。

属于不同 VXLAN 网络的用户终端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 内对应的用户终端。

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用户终端的集中式 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

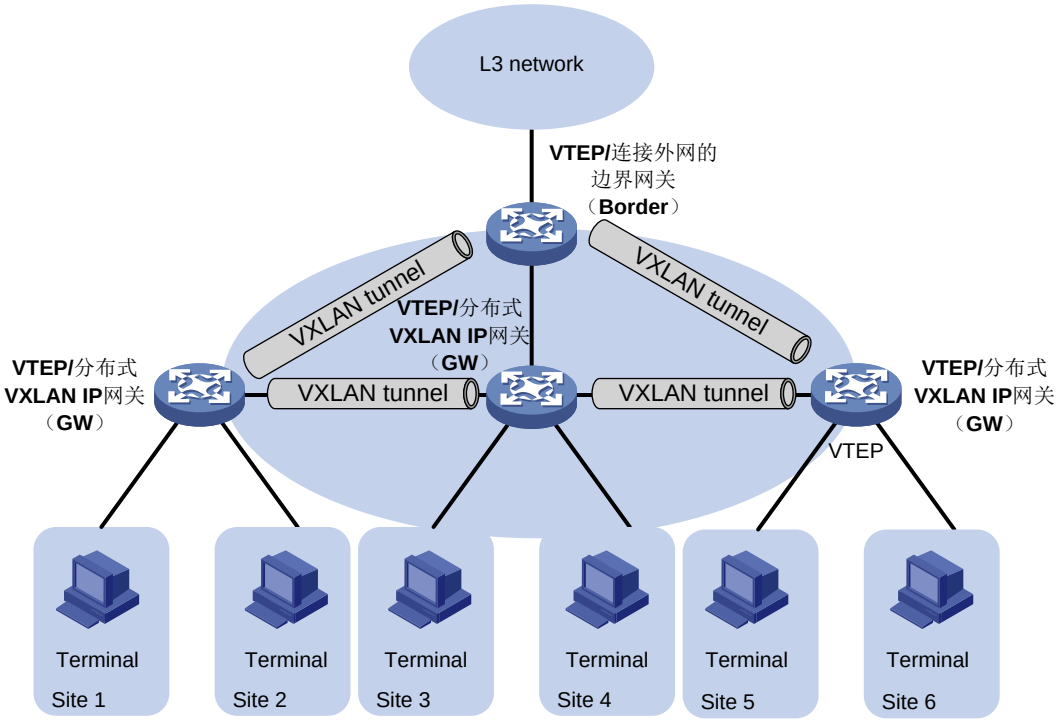


如图 3-4 所示，两台集中式 VXLAN IP 网关形成保护组，两台设备上存在相同的 VTEP IP，称为保护组的 VTEP IP。接入层 VTEP 与保护组的 VTEP IP 建立 VXLAN 隧道，将用户终端发送至其它网络的报文转发至保护组，保护组中的两台网关设备均可以接收并处理用户终端发往其它网络的流量。保护组中的成员 VTEP 之间、每个成员 VTEP 与接入层 VTEP 之间还会采用成员自身的 IP 地址建立 VXLAN 隧道，以便进行协议通信和表项同步。

3.1.4 分布式 VXLAN IP 网关

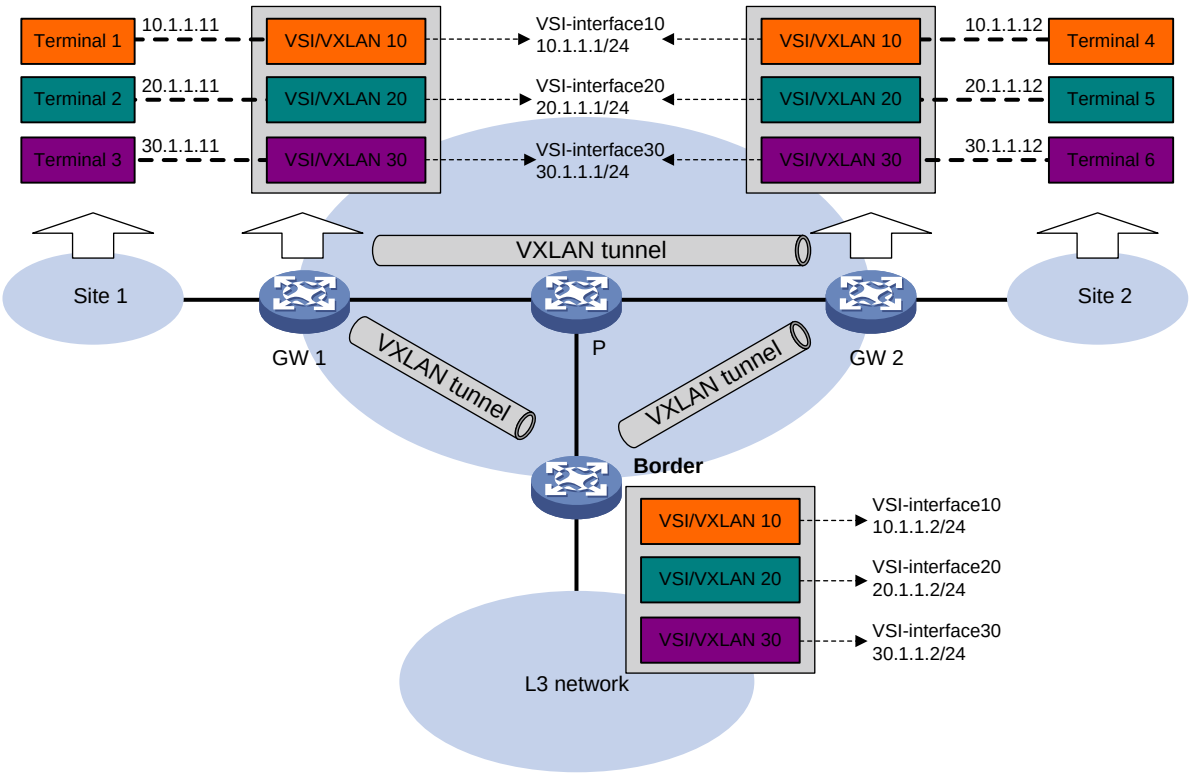
1. 简介

图3-5 分布式 VXLAN IP 网关示意图



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3-6 分布式 VXLAN IP 网关部署示意图



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 内用户终端的网关地址。在分布式 VXLAN IP 网关上还需要开启本地代理 ARP 功能（IPv4 网络）或本地 ND 代理功能（IPv6 网络）。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 网关上还需要开启以下功能中的一种：

- ARP/ND 泛洪抑制功能：开启本功能后，二层流量查找 MAC 地址表进行转发，三层流量查找 ARP/ND 表项进行转发。
- 本地代理 ARP 功能或本地 ND 代理功能：开启本功能后，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。

2. 相同 VXLAN 内不同站点的用户终端通信过程

如图 3-6 所示，以 Terminal 1 访问 Terminal 4 为例，相同 VXLAN 内不同站点的用户终端的通信过程为：

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

(2) GW 1 收到 ARP 请求消息后，学习 Terminal 1 的 ARP 信息，并代理应答该 ARP 请求，即：向 Terminal 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) Terminal 1 学习到 Terminal 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，对该消息进行 VXLAN 封装后，将其发送给 VXLAN 内的所有远端 VTEP。

(5) GW 2 对 VXLAN 报文进行解封装后，学习 Terminal 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) Terminal 4 收到 ARP 请求后，学习 Terminal 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 Terminal 4 收到 ARP 应答消息后，学习 Terminal 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 Terminal 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，Terminal 1 发送给 Terminal 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 Terminal 4。

3. 不同 VXLAN 间不同站点的用户终端通信过程

如图 3-6 所示，以 Terminal 1 访问 Terminal 5 为例，不同 VXLAN 的用户终端的通信过程为：

(1) Terminal 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 Terminal 1 的 ARP 信息，并向 Terminal 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) Terminal 1 将访问 Terminal 5 的报文发送给 GW 1。

(4) GW 1 在所有 VXLAN 内向本地站点和远端站点广播发送 ARP 请求，获取 Terminal 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。

(5) GW 2 从 VXLAN 隧道上接收到 VXLAN 报文，对其进行解封装后，学习 GW 1 的 ARP 信息（IP 为 20.1.1.1、MAC 为 GW 1 上 VSI 虚接口 20 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 20 的 MAC 地址，在 VXLAN 20 的本地站点内广播该 ARP 请求消息。

(6) Terminal 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。

(7) GW 2 从 Terminal 5 收到 ARP 应答消息后，学习 Terminal 5 的 ARP 信息，将 ARP 应答消息中的源 MAC 修改为本地 VSI 虚接口 20 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 Terminal 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，Terminal 1 发送给 Terminal 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 Terminal 5。

4. 用户终端与外部网络的三层通信过程

用户终端要想与外部网络进行三层通信，需要在接入用户终端的本地分布式 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 Terminal 1 访问外部网络内的主机 50.1.1.1 为例，用户终端访问外部网络的三层通信过程为：

(1) Terminal 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 Terminal 1 的 ARP 信息，并向 Terminal 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) Terminal 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 Terminal 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

3.2.1 VXLAN IP 网关硬件限制

配置 VXLAN IP 网关时，VXLAN 公网侧端口必须位于以下接口板上：

- FD 系列接口板
- 下列 SG 系列接口板：LSUM2QGS12SG0、LSUM2TGS32QSSG0、LSUM2TGS48SG0
- LSUM1CQGS32SF0-Z 接口板
- SH 系列接口板

配置 VXLAN IP 网关时，VXLAN 用户侧端口除了不允许位于 LSUM1CQGS32SF0-Z 之外，允许的所属接口板与“2.1.1 VXLAN 硬件限制”一致。

3.2.2 VXLAN IP 网关软件限制

当使用 LSUM1CQGS32SF0-Z 接口板的端口作为 VXLAN IP 网关端口时，有如下限制：

- 系统工作模式必须切换到 **expert**。有关系统工作模式的介绍，请参见“基础配置配置指导”中的“设备管理”。
- 该接口板的端口不支持作为 VXLAN IP 网关的用户侧端口使用。
- 每个网关接口（VSI 虚接口）仅支持被一个 VSI 指定。

- 不支持基于 EVPN 的服务链策略路由配置。有关策略路由的介绍，请参见“三层技术-IP 路由配置指导”中的“策略路由”。
- 该接口板的端口所属 VLAN 的报文不支持 VXLAN 三层转发。

VXLAN IP 网关功能不支持与以下特性组合使用：Super VLAN、Private VLAN。有关 Super VLAN 和 Private VLAN 的介绍，请参见“二层技术-以太网交换配置指导”中的“Super VLAN”和“Private VLAN”。

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。

VXLAN IP 网关功能仅支持以下几种以太网服务实例：

- 报文匹配规则为 **encapsulation s-vid** *vlan-id* 或 **encapsulation untagged**，且接入模式为 VLAN 的以太网服务实例。
- 报文匹配规则为 **encapsulation s-vid** *vlan-id* **c-vid** *vlan-id*，且接入模式为 VLAN 的以太网服务实例。该 AC 仅支持与配置相同的本地 AC 三层互通；与远端 VTEP 三层通信时，VXLAN 隧道报文内层封装的以太网数据帧会携带 **c-vid**（Customer VLAN ID）标签。

VXLAN IP 网关接口（VSI 虚接口）仅支持如下协议类型：Telnet、TFTP、DHCP、ARP、ND、RADIUS、SSH、NTP、SNMP、NETCONF。有关 DHCP、ARP 和 ND 协议的介绍，请分别参见“三层技术-IP 业务配置指导”中的“DHCP”、“ARP”和“IPv6 基础”；有关 RADIUS 和 SSH 协议的介绍，请分别参见“安全配置指导”中的“AAA”和“SSH”；有关 NTP、SNMP 和 NETCONF 协议的介绍，请分别参见“网络管理和监控配置指导”中的“NTP”、“SNMP”和“NETCONF”。

3.2.3 VXLAN IP 网关用户侧接口板使用限制

以下接口板的端口上，如果以太网服务实例匹配两层 VLAN 标签的报文（**encapsulation** 命令配置了 **s-vid** 和 **c-vid** 参数），则该以太网服务实例关联的 VSI 不支持 VXLAN IP 网关功能。

- SE 系列接口板

对于以上接口板，还有如下使用限制：

- 当流量从这些单板进入时，不能通过 SH 系列接口板对该流量进行三层转发。
- 对于这些接口板的端口上以太网实例关联的 VSI，当不同的 VSI 指定同一网关接口时，该接口不支持应用 ACL 进行报文过滤。有关应用 ACL 进行报文过滤的介绍，请参见“ACL 和 QoS 配置指导”中的“ACL”。
- 请不要将这些接口板上的三层以太网接口/三层聚合接口关联到 VPN 实例（**ip binding vpn-instance**），否则从这些接口上进入的流量不能进行 VXLAN 三层转发。有关 VPN 实例的配置，请参见“MPLS 配置指导”中的“MPLS L3VPN”。
- 从这些接口板的端口进入的流量进行 VXLAN 三层转发时，入接口和出接口不能是同一个 AC。
- 当用户 VLAN 流量从这些单板进入，并通过 FD 系列接口板或 SG 系列接口板 LSUM2QGS12SG0、LSUM2TGS32QSSG0、LSUM2TGS48SG0 的出接口进行 VXLAN 三层转发时，这些 VXLAN 报文内层封装的以太网数据帧会携带用户 VLAN Tag。

3.3 VXLAN IP网关配置准备

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

- 创建 VSI 和 VXLAN。

- 配置 VXLAN 隧道，并将 VXLAN 与 VXLAN 隧道关联。

3.4 配置集中式VXLAN IP网关

3.4.1 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。

3.4.2 配置步骤

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，不存在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* }	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，未指定VSI的网关接口

3.5 配置集中式VXLAN IP网关保护组

3.5.1 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

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，不存在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* }	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地址为桥MAC地址+1 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，未指定VSI的网关接口
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，设备未加入VXLAN IP网关保护组 保护组中的每台VXLAN IP网关上都要执行此配置。*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 保护组中每台VXLAN IP网关上都要执行此配置，且必须输入保护组中所有其它成员的成员地址

3.6 配置分布式VXLAN IP网关

3.6.1 配置限制和指导

分布式 VXLAN IP 网关上所有公网侧端口都需要配置 **undo mac-address static source-check enable** 命令。

分布式网关上不能 ping 通该网关学习到的远端主机的 ARP/ND 表项对应的 IPv4/IPv6 地址。

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。

在分布式 VXLAN IP 网关设备上，如果开启了 ARP 泛洪抑制功能，并在 VSI 虚接口上开启了本地代理 ARP 功能，则只有本地代理 ARP 功能生效。建议不要在分布式 VXLAN IP 网关设备上同时开启这两个功能。有关 ARP 泛洪抑制功能的详细介绍请参见“2.12 配置 ARP 泛洪抑制”。

3.6.2 配置准备

如果用户终端要想与外部网络进行三层通信，那么需要在接入用户终端的本地分布式 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** 命令或 **apply next-hop** 命令设置报文的缺省下一跳或下一跳为 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”。

3.6.3 配置步骤

表3-3 配置分布式 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，不存在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地址或IPv6地址	• 配置 VSI 虚接口的 IP 地址：**ip address** *ip-address* { *mask* | *mask-length* } [**sub**] • 配置 VSI 虚接口的 IPv6 地址： 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地址和IPv6地址
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
开启本地代理ARP功能	**local-proxy-arp enable** [**ip-range** *startIP* **to** *endIP*]	对于IPv4网络，必选 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”
开启本地ND代理功能	**local-proxy-nd enable**	对于IPv6网络，必选 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”
退回系统视图	**quit**	-
（可选）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 分布式VXLAN IP网关上开启本地代理ARP功能时，需要开启本功能，以保证所有网关都能学习到ARP表项 分布式VXLAN IP网关之间也可以通过控制器或EVPN等在彼此之间同步ARP表项，此时无需开启本功能

操作	命令	说明
（可选）开启分布式网关的动态IPv6 ND表项同步功能	**ipv6 nd distributed-gateway dynamic-entry synchronize**	缺省情况下，分布式网关的动态IPv6 ND表项同步功能处于关闭状态 分布式VXLAN IP网关上开启本地ND代理功能时，本地网关不会将目标IP地址为分布式网关VSI虚接口的IPv6 ND报文转发给其他网关，只有本地网关能够学习到IPv6 ND报文发送者的ND表项。如果希望所有网关都能学习到该ND表项，需要开启分布式网关的动态IPv6 ND表项同步功能 分布式VXLAN IP网关之间也可以通过控制器或EVPN等在彼此之间同步IPv6 ND表项，此时无需开启本功能
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，未指定VSI的网关接口
配置当前VSI所属的子网网段	**gateway subnet** { *ipv4-address wildcard-mask* | *ipv6-address prefix-length* }	缺省情况下，未指定VSI所属的子网网段 为了节省分布式VXLAN IP网关设备上的三层接口资源，在网关设备上多个VXLAN可以共用一个VSI虚接口，为VSI虚接口配置一个主IPv4地址和多个从IPv4地址、或多个IPv6地址，分别作为不同VXLAN内用户终端的网关地址 多个VXLAN共用一个VSI虚接口时，网关设备无法判断从VSI虚接口接收到的报文属于哪个VXLAN。为了解决该问题，需要在VSI视图下通过本命令指定VSI所属的子网网段，通过子网网段判断报文所属的VSI，并在该VSI内转发报文，从而限制广播报文范围，有效地节省带宽资源

3.7 静态配置ARP表项

VXLAN IP 网关既可以动态学习 ARP 表项，也可以通过本配置静态创建 ARP 表项。

表3-4 静态配置 ARP 表项

操作	命令	说明
进入系统视图	**system-view**	-
静态配置本地ARP表项	**arp static** *ip-address mac-address* **vsi-interface** *vsi-interface-id interface-type interface-number* **service-instance** *instance-id* **vsi** *vsi-name* [**vpn-instance** *vpn-instance-name*]	缺省情况下，不存在本地ARP表项 本命令的详细介绍，请参见“三层技术-IP业务命令参考”中的“ARP”
静态配置远端ARP表项	**arp static** *ip-address mac-address* **vsi-interface** *vsi-interface-id* **tunnel** *number* **vsi** *vsi-name* [**vpn-instance** *vpn-instance-name*]	缺省情况下，不存在远端ARP表项 本命令的详细介绍，请参见“三层技术-IP业务命令参考”中的“ARP”

3.8 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3-5 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-interface** *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地址为桥MAC地址+1
（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** *size*	缺省情况下，VSI虚接口的MTU值为1444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＝接口的波特率÷1000（kbps）
（可选）恢复接口的缺省配置	**default**	-
（可选）开启VSI虚接口的ARP报文发送限速功能	**arp send-rate** *pps*	缺省情况下，VSI虚接口的ARP报文发送限速功能处于关闭状态
开启VSI虚接口	**undo shutdown**	缺省情况下，VSI虚接口处于开启状态

3.9 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。

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** | **down**]]

3.10 VXLAN IP网关典型配置举例

3.10.1 集中式 VXLAN IP 网关配置举例

1. 组网需求

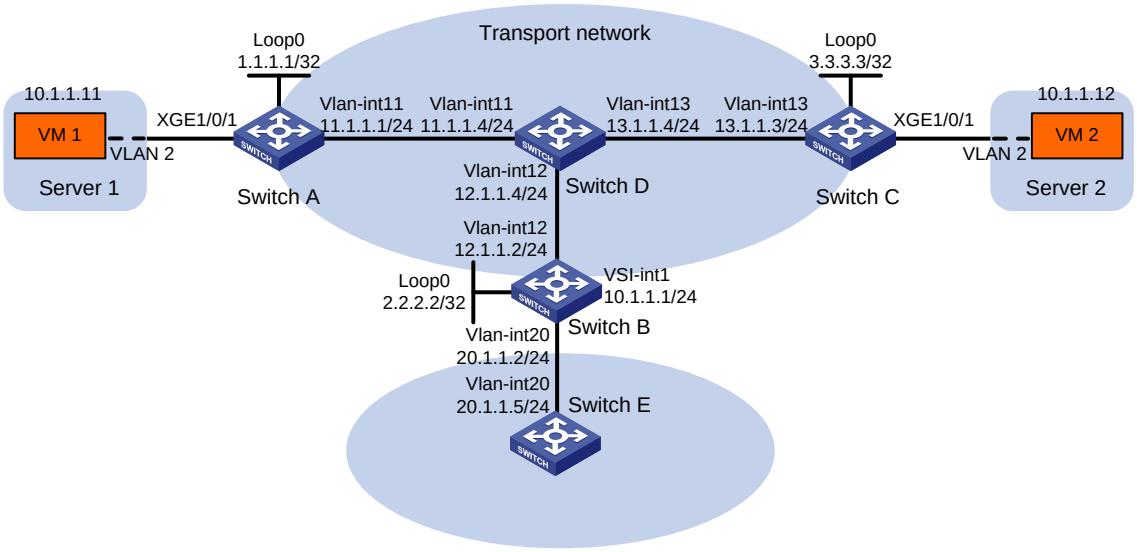
Switch A、Switch C 为与服务器连接的 VTEP 设备，Switch B 为与广域网连接的集中式 VXLAN IP 网关设备，Switch E 为广域网内的三层交换机。虚拟机 VM 1、VM 2 同属于 VXLAN 10，通过 VXLAN 实现不同站点间的二层互联，并通过 VXLAN IP 网关与广域网三层互联。

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3-7 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

在 VM 1 和 VM 2 上指定网关地址为 10.1.1.1。（具体配置过程略）

请按照图 3-7 配置各接口的 IP 地址和子网掩码；在 IP 核心网络内配置 OSPF 协议，确保交换机之间路由可达；配置 Switch B 和 Switch E 上发布 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
```

创建 VSI 实例 vpna 和 VXLAN 10。

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

在接入服务器的接口 Ten-GigabitEthernet1/0/1 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 ten-gigabitethernet 1/0/1
[SwitchA-Ten-GigabitEthernet1/0/1] port link-type trunk
[SwitchA-Ten-GigabitEthernet1/0/1] port trunk permit vlan 2
[SwitchA-Ten-GigabitEthernet1/0/1] service-instance 1000
[SwitchA-Ten-GigabitEthernet1/0/1-srv1000] encapsulation s-vid 2
```

配置以太网服务实例 1000 与 VSI 实例 vpna 关联。

```
[SwitchA-Ten-GigabitEthernet1/0/1-srv1000] xconnect vsi vpna
[SwitchA-Ten-GigabitEthernet1/0/1-srv1000] quit
[SwitchA-Ten-GigabitEthernet1/0/1] quit
```

(3) 配置 Switch B

开启 L2VPN 能力。

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

创建 VSI 实例 vpna 和 VXLAN 10。

```
[SwitchB] vsi vpna
[SwitchB-vsi-vpna] vxlan 10
[SwitchB-vsi-vpna-vxlan-10] 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 loopback 0
[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-vxlan-10] tunnel 2
[SwitchB-vsi-vpna-vxlan-10] tunnel 3
[SwitchB-vsi-vpna-vxlan-10] 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
```

创建 VSI 实例 vpna 和 VXLAN 10。

```
[SwitchC] vsi vpna
[SwitchC-vsi-vpna] vxlan 10
[SwitchC-vsi-vpna-vxlan-10] 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 loopback 0
[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-vxlan-10] tunnel 1
[SwitchC-vsi-vpna-vxlan-10] tunnel 3
[SwitchC-vsi-vpna-vxlan-10] quit
[SwitchC-vsi-vpna] quit
```

在接入服务器的接口 Ten-GigabitEthernet1/0/1 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 ten-gigabitethernet 1/0/1
[SwitchC-Ten-GigabitEthernet1/0/1] port link-type trunk
[SwitchC-Ten-GigabitEthernet1/0/1] port trunk permit vlan 2
[SwitchC-Ten-GigabitEthernet1/0/1] service-instance 1000
[SwitchC-Ten-GigabitEthernet1/0/1-srv1000] encapsulation s-vid 2
```

配置以太网服务实例 1000 与 VSI 实例 vpna 关联。

```
[SwitchC-Ten-GigabitEthernet1/0/1-srv1000] xconnect vsi vpna
[SwitchC-Ten-GigabitEthernet1/0/1-srv1000] quit
[SwitchC-Ten-GigabitEthernet1/0/1] 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: 64kbps
Maximum transmission unit: 1464
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
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 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: 10.1.1.1/24 (primary)
IP packet frame type: Ethernet II, hardware address: 0011-2200-0102
IPv6 packet frame type: Ethernet II, 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         : -
  MAC Learning rate       : -
  Drop Unknown            : -
  Flooding                : Enabled
  Gateway interface       : VSI-interface 1
  VXLAN ID                : 10
  Tunnels:
    Tunnel Name          Link ID    State  Type        Flood 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    VLAN/VSI   Interface/Link ID        Aging Type
20.1.1.5        000c-29c1-5e46 --         Vlan20                   19    D
10.1.1.11       0000-1234-0001 0          Tunnel2                  20    D
10.1.1.12       0000-1234-0002 0          Tunnel3                  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:Usable     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        Vsi100                   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。

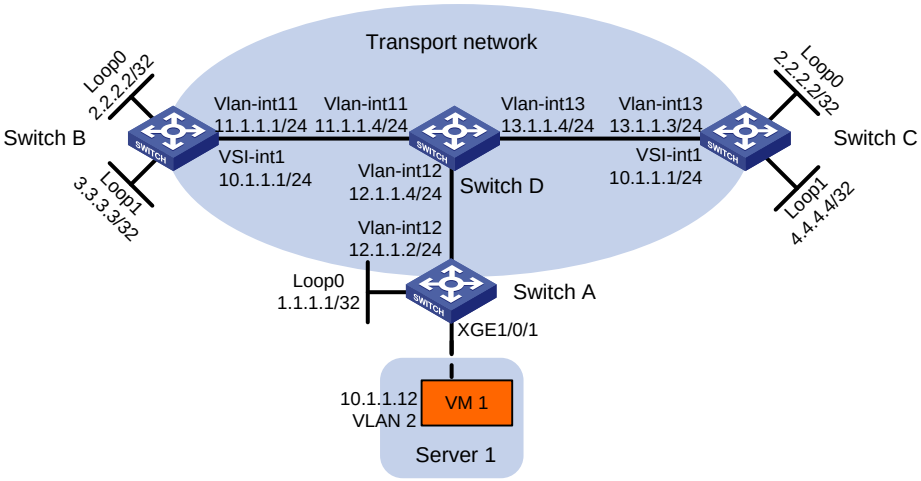
3.10.2 集中式 VXLAN IP 网关保护组配置举例

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3-8 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

在 VM 1 上指定网关地址为 10.1.1.1。（具体配置过程略）

请按照图 3-8 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
```

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
```

在接入服务器的接口 Ten-GigabitEthernet1/0/1 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 ten-gigabitethernet 1/0/1
[SwitchA-Ten-GigabitEthernet1/0/1] port link-type trunk
[SwitchA-Ten-GigabitEthernet1/0/1] port trunk permit vlan 2
[SwitchA-Ten-GigabitEthernet1/0/1] service-instance 1000
[SwitchA-Ten-GigabitEthernet1/0/1-srv1000] encapsulation s-vid 2
```

配置以太网服务实例 1000 与 VSI 实例 vpna 关联。

```
[SwitchA-Ten-GigabitEthernet1/0/1-srv1000] xconnect vsi vpna
[SwitchA-Ten-GigabitEthernet1/0/1-srv1000] quit
[SwitchA-Ten-GigabitEthernet1/0/1] quit
```

(3) 配置 Switch B

开启 L2VPN 能力。

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

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 地址和 MAC 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
```

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 地址和 MAC 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
```

3.10.3 分布式 VXLAN IP 网关连接 IPv4 网络配置举例

1. 组网需求

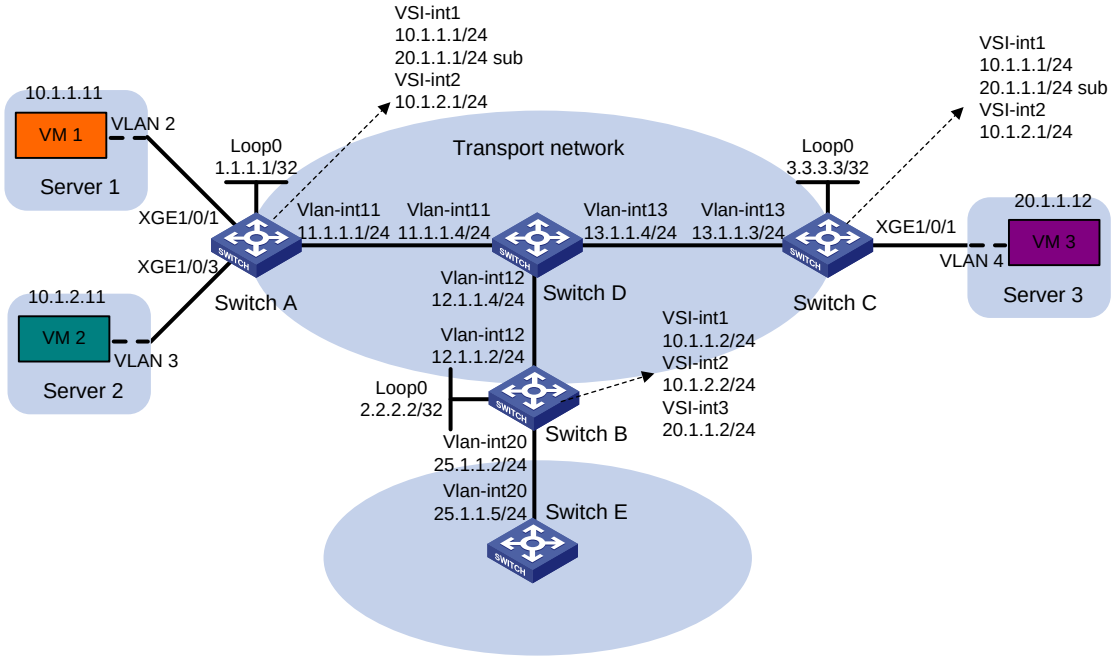
Switch A、Switch C 为分布式 VXLAN IP 网关设备，Switch B 为与广域网连接的边界网关设备，Switch E 为广域网内的三层交换机。虚拟机 VM 1 属于 VXLAN 10，VM 2 属于 VXLAN 20，VM 3 属于 VXLAN 30。通过分布式 VXLAN IP 网关实现不同 VXLAN 网络的三层互联，并通过边界网关实现与广域网的三层互联。

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。
- 站点之间的泛洪流量采用头端复制的方式转发。
- VM 1、VM 2、VM 3 之间可以互访，且 VM 1、VM 2 和 VM 3 都可以访问外部网络。

2. 组网图

图3-9 分布式 VXLAN IP 网关连接 IPv4 网络配置组网图



3. 配置步骤

(1) 配置 IP 地址和单播路由协议

在 VM 1、VM 2 和 VM 3 上分别指定网关地址为 10.1.1.1、10.1.2.1、20.1.1.1。（具体配置过程略）

请按照图 3-9 配置各接口的 IP 地址和子网掩码；在 IP 核心网络内配置 OSPF 协议，确保交换机之间路由可达；配置 Switch B 和 Switch E 发布 10.1.1.0/24、10.1.2.0/24、20.1.1.0/24 和 25.1.1.0/24 网段的路由。（具体配置过程略）

在 Switch A、Switch C 的公网侧端口上配置 **undo mac-address static source-check enable** 命令。（具体配置过程略）

(2) 配置 Switch A

开启 L2VPN 能力。

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

创建 VSI 实例 vpna 和 VXLAN 10。

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

创建 VSI 实例 vpnb 和 VXLAN 20。

```
[SwitchA] vsi vpnb
[SwitchA-vsi-vpnb] vxlan 20
[SwitchA-vsi-vpnb-vxlan-20] quit
[SwitchA-vsi-vpnb] quit
```

创建 VSI 实例 vpnc 和 VXLAN 30。

```
[SwitchA] vsi vpnc
[SwitchA-vsi-vpnc] vxlan 30
[SwitchA-vsi-vpnc-vxlan-30] quit
[SwitchA-vsi-vpnc] 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 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-vxlan-10] tunnel 1
[SwitchA-vsi-vpna-vxlan-10] tunnel 2
[SwitchA-vsi-vpna-vxlan-10] quit
[SwitchA-vsi-vpna] quit
```

配置 Tunnel1 和 Tunnel2 与 VXLAN 20 关联。

```
[SwitchA] vsi vpnb
[SwitchA-vsi-vpnb] vxlan 20
[SwitchA-vsi-vpnb-vxlan-20] tunnel 1
[SwitchA-vsi-vpnb-vxlan-20] tunnel 2
[SwitchA-vsi-vpnb-vxlan-20] quit
[SwitchA-vsi-vpnb] quit
```

配置 Tunnel2 与 VXLAN 30 关联。

```
[SwitchA] vsi vpnc
[SwitchA-vsi-vpnc] vxlan 30
[SwitchA-vsi-vpnc-vxlan-30] tunnel 2
[SwitchA-vsi-vpnc-vxlan-30] quit
[SwitchA-vsi-vpnc] quit
```

在接入 VM 1 的接口 Ten-GigabitEthernet1/0/1 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 ten-gigabitethernet 1/0/1
[SwitchA-Ten-GigabitEthernet1/0/1] port link-type trunk
[SwitchA-Ten-GigabitEthernet1/0/1] port trunk permit vlan 2
```

```
[SwitchA-Ten-GigabitEthernet1/0/1] service-instance 1000
[SwitchA-Ten-GigabitEthernet1/0/1-srv1000] encapsulation s-vid 2
```
配置以太网服务实例 1000 与 VSI 实例 vpna 关联。
```
[SwitchA-Ten-GigabitEthernet1/0/1-srv1000] xconnect vsi vpna
[SwitchA-Ten-GigabitEthernet1/0/1-srv1000] quit
[SwitchA-Ten-GigabitEthernet1/0/1] quit
```
在接入 VM 2 的接口 Ten-GigabitEthernet1/0/3 上创建以太网服务实例 1000，该实例用来匹配 VLAN 3 的数据帧。
```
[SwitchA] interface ten-gigabitethernet 1/0/3
[SwitchA-Ten-GigabitEthernet1/0/3] port link-type trunk
[SwitchA-Ten-GigabitEthernet1/0/3] port trunk permit vlan 3
[SwitchA-Ten-GigabitEthernet1/0/3] service-instance 1000
[SwitchA-Ten-GigabitEthernet1/0/3-srv1000] encapsulation s-vid 3
```
配置以太网服务实例 1000 与 VSI 实例 vpnb 关联。
```
[SwitchA-Ten-GigabitEthernet1/0/3-srv1000] xconnect vsi vpnb
[SwitchA-Ten-GigabitEthernet1/0/3-srv1000] quit
[SwitchA-Ten-GigabitEthernet1/0/3] quit
```
创建 VSI 虚接口 VSI-interface1，并为其配置 IP 地址和 MAC 地址，该 IP 地址作为 VXLAN 10 内虚拟机的网关地址，指定该 VSI 虚接口为分布式本地网关接口，并开启本地代理 ARP 功能。
```
[SwitchA] interface vsi-interface 1
[SwitchA-Vsi-interface1] ip address 10.1.1.1 255.255.255.0
[SwitchA-Vsi-interface1] mac-address 1-1-1
[SwitchA-Vsi-interface1] distributed-gateway local
[SwitchA-Vsi-interface1] local-proxy-arp enable
[SwitchA-Vsi-interface1] quit
```
创建 VSI 虚接口 VSI-interface2，并为其配置 IP 地址和 MAC 地址，该 IP 地址作为 VXLAN 20 内虚拟机的网关地址，指定该 VSI 虚接口为分布式本地网关接口，并开启本地代理 ARP 功能。
```
[SwitchA] interface vsi-interface 2
[SwitchA-Vsi-interface2] ip address 10.1.2.1 255.255.255.0
[SwitchA-Vsi-interface2] mac-address 2-2-2
[SwitchA-Vsi-interface2] distributed-gateway local
[SwitchA-Vsi-interface2] local-proxy-arp enable
[SwitchA-Vsi-interface2] quit
```
开启分布式网关的动态 ARP 表项同步功能。
```
[SwitchA] arp distributed-gateway dynamic-entry synchronize
```
配置 VXLAN 10 所在的 VSI 实例和接口 VSI-interface1 关联，并配置该 VSI 实例的子网网段为 10.1.1.0/24。
```
[SwitchA] vsi vpna
[SwitchA-vsi-vpna] gateway vsi-interface 1
[SwitchA-vsi-vpna] gateway subnet 10.1.1.0 0.0.0.255
[SwitchA-vsi-vpna] quit
```
配置 VXLAN 20 所在的 VSI 实例和接口 VSI-interface2 关联。
```
[SwitchA] vsi vpnb
[SwitchA-vsi-vpnb] gateway vsi-interface 2
[SwitchA-vsi-vpnb] quit
```
为 VSI 虚接口 VSI-interface1 配置从 IP 地址，该从 IP 地址作为 VXLAN 30 内虚拟机的网关地址。

```
[SwitchA] interface vsi-interface 1
[SwitchA-Vsi-interface1] ip address 20.1.1.1 255.255.255.0 sub
[SwitchA-Vsi-interface1] quit
```

配置 VXLAN 30 所在的 VSI 实例和接口 VSI-interface1 关联，并配置该 VSI 实例的子网网段为 20.1.1.0/24。

```
[SwitchA] vsi vpnc
[SwitchA-vsi-vpnc] gateway vsi-interface 1
[SwitchA-vsi-vpnc] gateway subnet 20.1.1.0 0.0.0.255
[SwitchA-vsi-vpnc] quit
```

配置策略路由，指定 IPv4 报文如果未找到匹配的路由表项，则设置报文的下一跳为 Switch B 上接口 VSI-interface1 的 IP 地址 10.1.1.2。

```
[SwitchA] acl advanced 3000
[SwitchA-acl-ipv4-adv-3000] rule 0 permit ip
[SwitchA-acl-ipv4-adv-3000] quit
[SwitchA] policy-based-route vxlan10 permit node 5
[SwitchA-pbr-vxlan10-5] if-match acl 3000
[SwitchA-pbr-vxlan10-5] apply default-next-hop 10.1.1.2
[SwitchA-pbr-vxlan10-5] quit
```

配置策略路由，指定 IPv4 报文如果未找到匹配的路由表项，则设置报文的下一跳为 Switch B 上接口 VSI-interface2 的 IP 地址 10.1.2.2。

```
[SwitchA] policy-based-route vxlan20 permit node 5
[SwitchA-pbr-vxlan20-5] if-match acl 3000
[SwitchA-pbr-vxlan20-5] apply default-next-hop 10.1.2.2
[SwitchA-pbr-vxlan20-5] quit
```

在 VSI 虚接口 VSI-interface1 和 VSI-interface2 上应用策略路由。

```
[SwitchA] interface vsi-interface 1
[SwitchA-Vsi-interface1] ip policy-based-route vxlan10
[SwitchA-Vsi-interface1] quit
[SwitchA] interface vsi-interface 2
[SwitchA-Vsi-interface2] ip policy-based-route vxlan20
[SwitchA-Vsi-interface2] quit
```

(3) 配置 Switch B

开启 L2VPN 能力。

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

创建 VSI 实例 vpna 和 VXLAN 10。

```
[SwitchB] vsi vpna
[SwitchB-vsi-vpna] vxlan 10
[SwitchB-vsi-vpna-vxlan-10] quit
[SwitchB-vsi-vpna] quit
```

创建 VSI 实例 vpnb 和 VXLAN 20。

```
[SwitchB] vsi vpnb
[SwitchB-vsi-vpnb] vxlan 20
[SwitchB-vsi-vpnb-vxlan-20] quit
[SwitchB-vsi-vpnb] quit
```

创建 VSI 实例 vpnc 和 VXLAN 30。

```
[SwitchB] vsi vpnc
[SwitchB-vsi-vpnc] vxlan 30
[SwitchB-vsi-vpnc-vxlan-30] quit
[SwitchB-vsi-vpnc] quit
```

配置接口 Loopback0 的 IP 地址，作为隧道的源端地址。

```
[SwitchB] interface loopback 0
[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 与 VXLAN 10 关联。

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

配置 Tunnel2 与 VXLAN 20 关联。

```
[SwitchB] vsi vpnb
[SwitchB-vsi-vpnb] vxlan 20
[SwitchB-vsi-vpnb-vxlan-20] tunnel 2
[SwitchB-vsi-vpnb-vxlan-20] quit
[SwitchB-vsi-vpnb] quit
```

配置 Tunnel3 与 VXLAN 30 关联。

```
[SwitchB] vsi vpnc
[SwitchB-vsi-vpnc] vxlan 30
[SwitchB-vsi-vpnc-vxlan-30] tunnel 3
[SwitchB-vsi-vpnc-vxlan-30] quit
[SwitchB-vsi-vpnc] quit
```

创建 VSI 虚接口 VSI-interface1，并为其配置 IP 地址。

```
[SwitchB] interface vsi-interface 1
[SwitchB-Vsi-interface1] ip address 10.1.1.2 255.255.255.0
[SwitchB-Vsi-interface1] quit
```

创建 VSI 虚接口 VSI-interface2，并为其配置 IP 地址。

```
[SwitchB] interface vsi-interface 2
[SwitchB-Vsi-interface2] ip address 10.1.2.2 255.255.255.0
[SwitchB-Vsi-interface2] quit
```

创建 VSI 虚接口 VSI-interface3，并为其配置 IP 地址。

```
[SwitchB] interface vsi-interface 3
```

```
[SwitchB-Vsi-interface3] ip address 20.1.1.2 255.255.255.0
[SwitchB-Vsi-interface3] 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 20 所在的 VSI 实例和接口 VSI-interface2 关联。
```
[SwitchB] vsi vpnb
[SwitchB-vsi-vpnb] gateway vsi-interface 2
[SwitchB-vsi-vpnb] quit
```
配置 VXLAN 30 所在的 VSI 实例和接口 VSI-interface3 关联。
```
[SwitchB] vsi vpnc
[SwitchB-vsi-vpnc] gateway vsi-interface 3
[SwitchB-vsi-vpnc] quit
```
(4)　配置 Switch C

开启 L2VPN 能力。
```
<SwitchC> system-view
[SwitchC] l2vpn enable
```
创建 VSI 实例 vpna 和 VXLAN 10。
```
[SwitchC] vsi vpna
[SwitchC-vsi-vpna] vxlan 10
[SwitchC-vsi-vpna-vxlan-10] quit
[SwitchC-vsi-vpna] quit
```
创建 VSI 实例 vpnb 和 VXLAN 20。
```
[SwitchC] vsi vpnb
[SwitchC-vsi-vpnb] vxlan 20
[SwitchC-vsi-vpnb-vxlan-20] quit
[SwitchC-vsi-vpnb] quit
```
创建 VSI 实例 vpnc 和 VXLAN 30。
```
[SwitchC] vsi vpnc
[SwitchC-vsi-vpnc] vxlan 30
[SwitchC-vsi-vpnc-vxlan-30] quit
[SwitchC-vsi-vpnc] quit
```
配置接口 Loopback0 的 IP 地址，作为隧道的源端地址。
```
[SwitchC] interface loopback 0
[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 与 VXLAN 10 关联。

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

配置 Tunnel1 与 VXLAN 20 关联。

```
[SwitchC] vsi vpnb
[SwitchC-vsi-vpnb] vxlan 20
[SwitchC-vsi-vpnb-vxlan-20] tunnel 1
[SwitchC-vsi-vpnb-vxlan-20] quit
[SwitchC-vsi-vpnb] quit
```

配置 Tunnel1 和 Tunnel3 与 VXLAN 30 关联。

```
[SwitchC] vsi vpnc
[SwitchC-vsi-vpnc] vxlan 30
[SwitchC-vsi-vpnc-vxlan-30] tunnel 1
[SwitchC-vsi-vpnc-vxlan-30] tunnel 3
[SwitchC-vsi-vpnc-vxlan-30] quit
[SwitchC-vsi-vpnc] quit
```

在接入 VM 3 的接口 Ten-GigabitEthernet1/0/1 上创建以太网服务实例 1000，该实例用来匹配 VLAN 4 的数据帧。

```
[SwitchC] interface ten-gigabitethernet 1/0/1
[SwitchC-Ten-GigabitEthernet1/0/1] port link-type trunk
[SwitchC-Ten-GigabitEthernet1/0/1] port trunk permit vlan 4
[SwitchC-Ten-GigabitEthernet1/0/1] service-instance 1000
[SwitchC-Ten-GigabitEthernet1/0/1-srv1000] encapsulation s-vid 4
```

配置以太网服务实例 1000 与 VSI 实例 vpnc 关联。

```
[SwitchC-Ten-GigabitEthernet1/0/1-srv1000] xconnect vsi vpnc
[SwitchC-Ten-GigabitEthernet1/0/1-srv1000] quit
[SwitchC-Ten-GigabitEthernet1/0/1] quit
```

创建 VSI 虚接口 VSI-interface1，并为其配置 IP 地址和 MAC 地址，该 IP 地址作为 VXLAN 10 内虚拟机的网关地址，指定该 VSI 虚接口为分布式本地网关接口，并开启本地代理 ARP 功能。

```
[SwitchC] interface vsi-interface 1
[SwitchC-Vsi-interface1] ip address 10.1.1.1 255.255.255.0
[SwitchC-Vsi-interface1] mac-address 1-1-1
[SwitchC-Vsi-interface1] distributed-gateway local
[SwitchC-Vsi-interface1] local-proxy-arp enable
[SwitchC-Vsi-interface1] quit
```

创建 VSI 虚接口 VSI-interface2，并为其配置 IP 地址和 MAC 地址，该 IP 地址作为 VXLAN 20 内虚拟机的网关地址，指定该 VSI 虚接口为分布式本地网关接口，并开启本地代理 ARP 功能。

```
[SwitchC] interface vsi-interface 2
[SwitchC-Vsi-interface2] ip address 10.1.2.1 255.255.255.0
[SwitchC-Vsi-interface2] mac-address 2-2-2
[SwitchC-Vsi-interface2] distributed-gateway local
```

```
[SwitchC-Vsi-interface2] local-proxy-arp enable
[SwitchC-Vsi-interface2] quit
```
开启分布式网关的动态 ARP 表项同步功能。
```
[SwitchC] arp distributed-gateway dynamic-entry synchronize
```
配置 VXLAN 10 所在的 VSI 实例和接口 VSI-interface1 关联，并配置该 VSI 实例的子网网段为 10.1.1.0/24。
```
[SwitchC] vsi vpna
[SwitchC-vsi-vpna] gateway vsi-interface 1
[SwitchC-vsi-vpna] gateway subnet 10.1.1.0 0.0.0.255
[SwitchC-vsi-vpna] quit
```
配置 VXLAN 20 所在的 VSI 实例和接口 VSI-interface2 关联。
```
[SwitchC] vsi vpnb
[SwitchC-vsi-vpnb] gateway vsi-interface 2
[SwitchC-vsi-vpnb] quit
```
为 VSI 虚接口 VSI-interface1 配置从 IP 地址，该从 IP 地址作为 VXLAN 30 内虚拟机的网关地址。
```
[SwitchC] interface vsi-interface 1
[SwitchC-Vsi-interface1] ip address 20.1.1.1 255.255.255.0 sub
[SwitchC-Vsi-interface1] quit
```
配置 VXLAN 30 所在的 VSI 实例和接口 VSI-interface1 关联，并配置该 VSI 实例的子网网段为 20.1.1.0/24。
```
[SwitchC] vsi vpnc
[SwitchC-vsi-vpnc] gateway vsi-interface 1
[SwitchC-vsi-vpnc] gateway subnet 20.1.1.0 0.0.0.255
[SwitchC-vsi-vpnc] quit
```
配置策略路由，指定 IPv4 报文如果未找到匹配的路由表项，则设置报文的下一跳为 Switch B 上接口 VSI-interface1 的 IP 地址 20.1.1.2。
```
[SwitchC] acl advanced 3000
[SwitchC-acl-ipv4-adv-3000] rule 0 permit ip
[SwitchC-acl-ipv4-adv-3000] quit
[SwitchC] policy-based-route vxlan permit node 5
[SwitchC-pbr-vxlan-5] if-match acl 3000
[SwitchC-pbr-vxlan-5] apply default-next-hop 20.1.1.2
[SwitchC-pbr-vxlan-5] quit
```
在 VSI 虚接口 VSI-interface1 上应用策略路由。
```
[SwitchC] interface vsi-interface1
[SwitchC-Vsi-interface1] ip policy-based-route vxlan
[SwitchC-Vsi-interface1] quit
```
4. 验证配置

(1) 验证分布式 VXLAN IP 网关设备 Switch A

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 2
Tunnel2
Current state: UP
Line protocol state: UP
Description: Tunnel2 Interface
```

```
Bandwidth: 64 kbps
Maximum transmission unit: 1464
Internet protocol processing: Disabled
Last clearing of counters: Never
Tunnel source 1.1.1.1, destination 3.3.3.3
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 虚接口信息，可以看到 VSI 虚接口处于 up 状态。

```
[SwitchA] display interface vsi-interface 1
Vsi-interface1
Current state: UP
Line protocol state: UP
Description: Vsi-interface1 Interface
Bandwidth: 1000000 kbps
Maximum transmission unit: 1500
Internet address: 10.1.1.1/24 (primary)
IP packet frame type: Ethernet II, hardware address: 0001-0001-0001
IPv6 packet frame type: Ethernet II, hardware address: 0001-0001-0001
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 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 VSI 虚接口等信息。

```
[SwitchA] display l2vpn vsi name vpna verbose
VSI Name: vpna
  VSI Index               : 0
  VSI State               : Up
  MTU                     : 1500
  Bandwidth               : Unlimited
  Broadcast Restrain      : 5120 kbps
  Multicast Restrain      : 5120 kbps
  Unknown Unicast Restrain: 5120 kbps
  MAC Learning            : Enabled
  MAC Table Limit         : -
  MAC Learning rate       : -
  Drop Unknown            : -
  Flooding                : Enabled
  Gateway Interface       : VSI-interface 1
  VXLAN ID                : 10
  Tunnels:
    Tunnel Name          Link ID    State  Type
    Tunnel1              0x5000001  Up     Manual
```

```
    Tunnel2              0x5000002  Up     Manual
  ACs:
    AC                               Link ID  State       Type
    XGE1/0/1 srv1000                 0        Up          Manual
```

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

```
[SwitchA] display arp
  Type: S-Static   D-Dynamic   O-Openflow   R-Rule   M-Multiport  I-Invalid
IP address      MAC address    VLAN/VSI   Interface/Link ID        Aging Type
11.1.1.4        000c-29c1-5e46 11         Vlan11                   19    D
10.1.1.2        dc2d-cb0d-867a 0          Tunnel1                  20    D
10.1.1.11       dc2d-cbb5-cf09 0          0                        20    D
10.1.2.2        dc2d-cb0d-867a 1          Tunnel1                  20    D
10.1.2.11       dc2d-cbb5-cf89 1          0                        20    D
20.1.1.12       0001-0001-0001 2          Tunnel2                  19    D
```

(2) 验证边界网关设备 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: Tunnel2 Interface
Bandwidth: 64 kbps
Maximum transmission unit: 1464
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
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 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-interface1 Interface
Bandwidth: 1000000 kbps
Maximum transmission unit: 1500
Internet address: 10.1.1.2/24 (primary)
IP packet frame type: Ethernet II, hardware address: 0011-2200-0102
IPv6 packet frame type: Ethernet II, 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 name vpna 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         : -
  MAC Learning rate       : -
  Drop Unknown            : -
  Flooding                : Enabled
  Gateway interface       : VSI-interface 1
  VXLAN ID                : 10
  Tunnels:
    Tunnel Name          Link ID    State  Type
    Tunnel1              0x5000001  Up     Manual
    Tunnel2              0x5000002  Up     Manual
```

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    VLAN/VSI   Interface/Link ID        Aging Type
12.1.1.4        0000-fc00-00ab 12         Vlan12                   14    D
25.1.1.5        dc2d-cb34-24bb 20         Vlan20                   17    D
10.1.1.1        0001-0001-0001 0          Tunnel2                  17    D
10.1.1.11       0001-0001-0001 0          Tunnel2                  20    D
20.1.1.1        0002-0002-0002 1          Tunnel3                  17    D
20.1.1.12       0002-0002-0002 1          Tunnel3                  20    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:Usable    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       Vsi1                     Null
[SwitchB] display fib 20.1.1.12
Destination count: 1 FIB entry count: 1
Flag:
  U:Usable    G:Gateway   H:Host   B:Blackhole   D:Dynamic   S:Static
  R:Relay     F:FRR
Destination/Mask   Nexthop         Flag     OutInterface/Token       Label
```

```
20.1.1.12/32        20.1.1.12        UH        Vsi3                    Null
```

(3) 验证主机和广域网互访

虚拟机 VM 1、VM 2、VM 3 之间可以互访；VM 1、VM 2 和 VM 3 可以与 Switch E 上接口 Vlan-interface20 的地址 25.1.1.5 之间互访。

3.10.4 分布式 VXLAN IP 网关连接 IPv6 网络配置举例

1. 组网需求

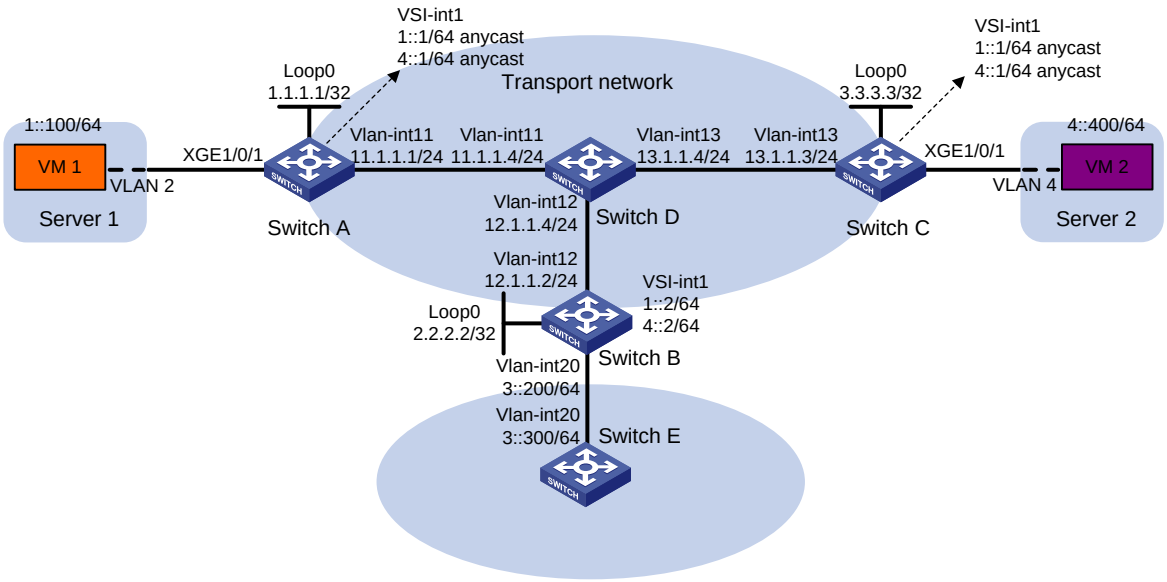
Switch A、Switch C 为分布式 VXLAN IP 网关设备，Switch B 为与广域网连接的边界网关设备，Switch E 为广域网内的三层交换机。虚拟机 VM 1 属于 VXLAN 10，VM 2 属于 VXLAN 20。通过分布式 VXLAN IP 网关实现不同 VXLAN 网络的三层互联，并通过边界网关实现与广域网的三层互联。

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。
- 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3-10 分布式 VXLAN IP 网关连接 IPv6 网络配置组网图



3. 配置步骤

(1) 配置 IPv6 地址和单播路由协议

在 VM 1 和 VM 2 上分别指定网关地址为 1::1、4::1。（具体配置过程略）

请按照图 3-10 配置各接口的地址；在 IP 核心网络内配置 OSPF 协议，确保交换机之间路由可达；配置 Switch B 和 Switch E 发布 1::/64、4::/64 和 3::/64 网段的路由。（具体配置过程略）

在 Switch A、Switch C 的公网侧端口上配置 **undo mac-address static source-check enable** 命令。（具体配置过程略）

(2) 配置 Switch A

开启 L2VPN 能力。

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

创建 VSI 实例 vpna 和 VXLAN 10。

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

创建 VSI 实例 vpnb 和 VXLAN 20。

```
[SwitchA] vsi vpnb
[SwitchA-vsi-vpnb] vxlan 20
[SwitchA-vsi-vpnb-vxlan-20] quit
[SwitchA-vsi-vpnb] 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 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-vxlan-10] tunnel 1
[SwitchA-vsi-vpna-vxlan-10] tunnel 2
[SwitchA-vsi-vpna-vxlan-10] quit
[SwitchA-vsi-vpna] quit
```

配置 Tunnel1 和 Tunnel2 与 VXLAN 20 关联。

```
[SwitchA] vsi vpnb
[SwitchA-vsi-vpnb] vxlan 20
[SwitchA-vsi-vpnb-vxlan-20] tunnel 1
[SwitchA-vsi-vpnb-vxlan-20] tunnel 2
[SwitchA-vsi-vpnb-vxlan-20] quit
[SwitchA-vsi-vpnb] quit
```

在接入 VM 1 的接口 Ten-GigabitEthernet1/0/1 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 ten-gigabitethernet 1/0/1
[SwitchA-Ten-GigabitEthernet1/0/1] port link-type trunk
[SwitchA-Ten-GigabitEthernet1/0/1] port trunk permit vlan 2
[SwitchA-Ten-GigabitEthernet1/0/1] service-instance 1000
[SwitchA-Ten-GigabitEthernet1/0/1-srv1000] encapsulation s-vid 2
```

配置以太网服务实例 1000 与 VSI 实例 vpna 关联。

```
[SwitchA-Ten-GigabitEthernet1/0/1-srv1000] xconnect vsi vpna
[SwitchA-Ten-GigabitEthernet1/0/1-srv1000] quit
[SwitchA-Ten-GigabitEthernet1/0/1] quit
```

创建 VSI 虚接口 VSI-interface1，并为其配置 IPv6 任播地址，其中 1::1/64 地址作为 VXLAN 10 内虚拟机的网关地址、4::1/64 作为 VXLAN 20 内虚拟机的网关地址，指定该 VSI 虚接口为分布式本地网关接口，并开启本地 ND 代理功能。

```
[SwitchA] interface vsi-interface 1
[SwitchA-Vsi-interface1] ipv6 address 1::1/64 anycast
[SwitchA-Vsi-interface1] ipv6 address 4::1/64 anycast
[SwitchA-Vsi-interface1] distributed-gateway local
[SwitchA-Vsi-interface1] local-proxy-nd enable
[SwitchA-Vsi-interface1] quit
```

开启分布式网关的动态 IPv6 ND 表项同步功能。

```
[SwitchA] ipv6 nd distributed-gateway dynamic-entry synchronize
```

配置 VXLAN 10 所在的 VSI 实例和接口 VSI-interface1 关联，并配置该 VSI 实例的子网网段为 1::1/64。

```
[SwitchA] vsi vpna
[SwitchA-vsi-vpna] gateway vsi-interface 1
[SwitchA-vsi-vpna] gateway subnet 1::1 64
[SwitchA-vsi-vpna] quit
```

配置 VXLAN 20 所在的 VSI 实例和接口 VSI-interface1 关联，并配置该 VSI 实例的子网网段为 4::1/64。

```
[SwitchA] vsi vpnb
[SwitchA-vsi-vpnb] gateway vsi-interface 1
[SwitchA-vsi-vpnb] gateway subnet 4::1 64
[SwitchA-vsi-vpnb] quit
```

配置静态路由，指定到达网络 3::/64 网络的路由下一跳为 Switch B 的 IPv6 地址 1::2。

```
[SwitchA] ipv6 route-static 3:: 64 1::2
```

(3) 配置 Switch B

开启 L2VPN 能力。

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

创建 VSI 实例 vpna 和 VXLAN 10。

```
[SwitchB] vsi vpna
[SwitchB-vsi-vpna] vxlan 10
[SwitchB-vsi-vpna-vxlan-10] quit
[SwitchB-vsi-vpna] quit
```

创建 VSI 实例 vpnb 和 VXLAN 20。

```
[SwitchB] vsi vpnb
[SwitchB-vsi-vpnb] vxlan 20
[SwitchB-vsi-vpnb-vxlan-20] quit
[SwitchB-vsi-vpnb] quit
```

配置接口 Loopback0 的 IP 地址，作为隧道的源端地址。

```
[SwitchB] interface loopback 0
[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-vxlan-10] tunnel 2
[SwitchB-vsi-vpna-vxlan-10] tunnel 3
[SwitchB-vsi-vpna-vxlan-10] quit
[SwitchB-vsi-vpna] quit
```

配置 Tunnel2 和 Tunnel3 与 VXLAN20 关联。

```
[SwitchB] vsi vpnb
[SwitchB-vsi-vpnb] vxlan 20
[SwitchB-vsi-vpnb-vxlan-20] tunnel 2
[SwitchB-vsi-vpnb-vxlan-20] tunnel 3
[SwitchB-vsi-vpnb-vxlan-20] quit
[SwitchB-vsi-vpnb] quit
```

创建 VSI 虚接口 VSI-interface1，并为其配置 IPv6 地址。

```
[SwitchB] interface vsi-interface 1
[SwitchB-Vsi-interface1] ipv6 address 1::2/64
[SwitchB-Vsi-interface1] ipv6 address 4::2/64
[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 20 所在的 VSI 实例和接口 VSI-interface1 关联。

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

(4) 配置 Switch C

开启 L2VPN 能力。

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

创建 VSI 实例 vpna 和 VXLAN 10。

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

创建 VSI 实例 vpnb 和 VXLAN 20。

```
[SwitchC] vsi vpnb
[SwitchC-vsi-vpnb] vxlan 20
[SwitchC-vsi-vpnb-vxlan-20] quit
[SwitchC-vsi-vpnb] quit
```

配置接口 Loopback0 的 IP 地址，作为隧道的源端地址。

```
[SwitchC] interface loopback 0
[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-vxlan-10] tunnel 1
[SwitchC-vsi-vpna-vxlan-10] tunnel 3
[SwitchC-vsi-vpna-vxlan-10] quit
[SwitchC-vsi-vpna] quit
```

配置 Tunnel1 和 Tunnel3 与 VXLAN 20 关联。

```
[SwitchC] vsi vpnb
[SwitchC-vsi-vpnb] vxlan 20
[SwitchC-vsi-vpnb-vxlan-20] tunnel 1
[SwitchC-vsi-vpnb-vxlan-20] tunnel 3
[SwitchC-vsi-vpnb-vxlan-20] quit
[SwitchC-vsi-vpnb] quit
```

在接入 VM 2 的接口 Ten-GigabitEthernet1/0/1 上创建以太网服务实例 1000，该实例用来匹配 VLAN 4 的数据帧。

```
[SwitchC] interface ten-gigabitethernet 1/0/1
[SwitchC-Ten-GigabitEthernet1/0/1] port link-type trunk
```

```
[SwitchC-Ten-GigabitEthernet1/0/1] port trunk permit vlan 4
[SwitchC-Ten-GigabitEthernet1/0/1] service-instance 1000
[SwitchC-Ten-GigabitEthernet1/0/1-srv1000] encapsulation s-vid 4
```

配置以太网服务实例 1000 与 VSI 实例 vpnb 关联。

```
[SwitchC-Ten-GigabitEthernet1/0/1-srv1000] xconnect vsi vpnb
[SwitchC-Ten-GigabitEthernet1/0/1-srv1000] quit
[SwitchC-Ten-GigabitEthernet1/0/1] quit
```

创建 VSI 虚接口 VSI-interface1，并为其配置 IPv6 地址，其中 1::1/64 地址作为 VXLAN 10 内虚拟机的网关地址，4::1/64 地址作为 VXLAN 20 内虚拟机的网关地址，指定该 VSI 虚接口为分布式本地网关接口，并开启本地 ND 代理功能。

```
[SwitchC] interface vsi-interface 1
[SwitchC-Vsi-interface1] ipv6 address 1::1/64 anycast
[SwitchC-Vsi-interface1] ipv6 address 4::1/64 anycast
[SwitchC-Vsi-interface1] distributed-gateway local
[SwitchC-Vsi-interface1] local-proxy-nd enable
[SwitchC-Vsi-interface1] quit
```

开启分布式网关的动态 IPv6 ND 表项同步功能。

```
[SwitchC] ipv6 nd distributed-gateway dynamic-entry synchronize
```

配置 VXLAN 10 所在的 VSI 实例和接口 VSI-interface1 关联，并配置该 VSI 实例的子网网段为 1::1/64。

```
[SwitchC] vsi vpna
[SwitchC-vsi-vpna] gateway vsi-interface 1
[SwitchC-vsi-vpna] gateway subnet 1::1 64
[SwitchC-vsi-vpna] quit
```

配置 VXLAN 20 所在的 VSI 实例和接口 VSI-interface1 关联，并配置该 VSI 实例的子网网段为 4::1/64。

```
[SwitchC] vsi vpnb
[SwitchC-vsi-vpnb] gateway vsi-interface 1
[SwitchC-vsi-vpnb] gateway subnet 4::1 64
[SwitchC-vsi-vpnb] quit
```

配置静态路由，指定到达网络 3::/64 网络的路由下一跳为 Switch B 的 IPv6 地址 4::2。

```
[SwitchC] ipv6 route-static 3:: 64 4::2
```

4. 验证配置

(1) 验证分布式 VXLAN IP 网关设备 Switch A

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 2
Tunnel2
Current state: UP
Line protocol state: UP
Description: Tunnel2 Interface
Bandwidth: 64kbps
Maximum transmission unit: 1464
Internet protocol processing: Disabled
Last clearing of counters: Never
Tunnel source 1.1.1.1, destination 3.3.3.3
```

```
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 虚接口信息，可以看到 VSI 虚接口处于 up 状态。

```
[SwitchA] display interface vsi-interface 1
Vsi-interface1
Current state: UP
Line protocol state: UP
Description: Vsi-interface1 Interface
Bandwidth: 1000000kbps
Maximum transmission unit: 1500
Internet protocol processing: Disabled
IP packet frame type: Ethernet II, hardware address: 0011-2200-0102
IPv6 packet frame type: Ethernet II, 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 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 VSI 虚接口等信息。

```
[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         : -
  MAC Learning rate       : -
  Drop Unknown            : -
  Flooding                : Enabled
  Gateway Interface       : VSI-interface 1
  VXLAN ID                : 10
  Tunnels:
    Tunnel Name          Link ID    State  Type        Flood proxy
    Tunnel1              0x5000001  Up     Manual      Disabled
    Tunnel2              0x5000002  Up     Manual      Disabled
  ACs:
    AC                              Link ID  State       Type
    XGE1/0/1 srv1000                0        Up          Manual
```

```
VSI Name: vpnb
  VSI Index               : 0
  VSI State               : Up
  MTU                     : 1500
  Bandwidth               : Unlimited
  Broadcast Restrain      : 5120 kbps
  Multicast Restrain      : 5120 kbps
  Unknown Unicast Restrain: 5120 kbps
  MAC Learning            : Enabled
  MAC Table Limit         : -
  Drop Unknown            : Disabled
  Flooding                : Enabled
  Gateway Interface       : VSI-interface 1
  VXLAN ID                : 20
  Tunnels:
    Tunnel Name          Link ID    State  Type
    Tunnel1              0x5000001  Up     Manual
    Tunnel2              0x5000002  Up     Manual
```

查看 Switch A 上 IPv6 neighbors 表项信息，可以看到已经建立的邻居信息。

```
[SwitchA] display ipv6 neighbors all
Type: S-Static    D-Dynamic    O-Openflow     R-Rule    I-Invalid
IPv6 address                  Link layer     VID  Interface      State T  Age
1::2                          dc2d-cb0d-867a N/A  Vsi1           STALE D  7
1::100                        0001-0000-0047 N/A  Vsi1           STALE D  22
4::400                        0002-0000-0047 N/A  Vsi1           REACH D  5
FE80::201:FF:FE00:47          0001-0000-0047 N/A  Vsi1           REACH D  30
FE80::202:FF:FE00:0           0002-0000-0000 N/A  Vsi1           REACH D  27
FE80::202:FF:FE00:47          0002-0000-0047 N/A  Vsi1           DELAY D  5
```

查看 Switch A 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。

```
[SwitchA] display ipv6 fib 4::400
Destination count: 1 FIB entry count: 1
Flag:
  U:Usable    G:Gateway   H:Host   B:Blackhole   D:Dynamic   S:Static
  R:Relay     F:FRR
Destination: 4::400                                    Prefix length: 128
Nexthop    : 4::400                                    Flags: UH
Time stamp : 0x2c                                      Label: Null
Interface  : Vsi1                                      Token: Invalid
[SwitchA] display ipv6 fib 3::300
Destination count: 1 FIB entry count: 1
Flag:
  U:Usable    G:Gateway   H:Host   B:Blackhole   D:Dynamic   S:Static
  R:Relay     F:FRR
Destination: 3::                                       Prefix length: 40
Nexthop    : 1::2                                      Flags: USGR
Time stamp : 0x23                                      Label: Null
Interface  : Vsi1                                      Token: Invalid
```

(2) 验证 VXLAN IP 边界网关设备 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: Tunnel2 Interface
Bandwidth: 64kbps
Maximum transmission unit: 1464
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
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 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-interface1 Interface
Bandwidth: 1000000kbps
Maximum transmission unit: 1500
Internet protocol processing: Disabled
IP packet frame type: Ethernet II, hardware address: 0011-2200-0102
IPv6 packet frame type: Ethernet II, 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 name vpna 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         : -
  MAC Learning rate       : -
```

```
  Drop Unknown            : -
  Flooding                : Enabled
  Gateway interface       : VSI-interface 1
  VXLAN ID                : 10
  Tunnels:
    Tunnel Name           Link ID    State  Type
    Tunnel1               0x5000001  Up     Manual
    Tunnel2               0x5000002  Up     Manual
```

查看 Switch B 上 IPv6 neighbors 表项信息，可以看到已经建立的邻居信息。

```
[SwitchB] display ipv6 neighbors all
Type: S-Static    D-Dynamic    O-Openflow     R-Rule    I-Invalid
IPv6 address                  Link layer     VID  Interface      State T  Age
3::300                        0003-0000-0047 N/A  FGE2/0/24      DELAY D  3
FE80::203:FF:FE00:47          0003-0000-0047 N/A  FGE2/0/24      STALE D  222
1::100                        0001-0000-0047 N/A  Vsi1           STALE D  232
4::400                        0002-0000-0047 N/A  Vsi1           REACH D  3
FE80::201:FF:FE00:0           0001-0000-0000 N/A  Vsi1           STALE D  237
FE80::201:FF:FE00:47          0001-0000-0047 N/A  Vsi1           STALE D  222
FE80::202:FF:FE00:0           0002-0000-0000 N/A  Vsi1           STALE D  345
```

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 ipv6 fib 1::100
Destination count: 1 FIB entry count: 1
Flag:
  U:Usable    G:Gateway   H:Host   B:Blackhole   D:Dynamic   S:Static
  R:Relay     F:FRR
Destination: 1::100                                    Prefix length: 128
Nexthop    : 1::100                                    Flags: UH
Time stamp : 0x21                                      Label: Null
Interface  : Vsi1                                      Token: Invalid
[SwitchB] display ipv6 fib 4::400
Destination count: 1 FIB entry count: 1
Flag:
  U:Usable    G:Gateway   H:Host   B:Blackhole   D:Dynamic   S:Static
  R:Relay     F:FRR
Destination: 4::                                       Prefix length: 64
Nexthop    : ::                                        Flags: U
Time stamp : 0x19                                      Label: Null
Interface  : Vsi1                                      Token: Invalid
```

(3) 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 3::300 之间可以互访。

4 VXLAN 数据中心互联

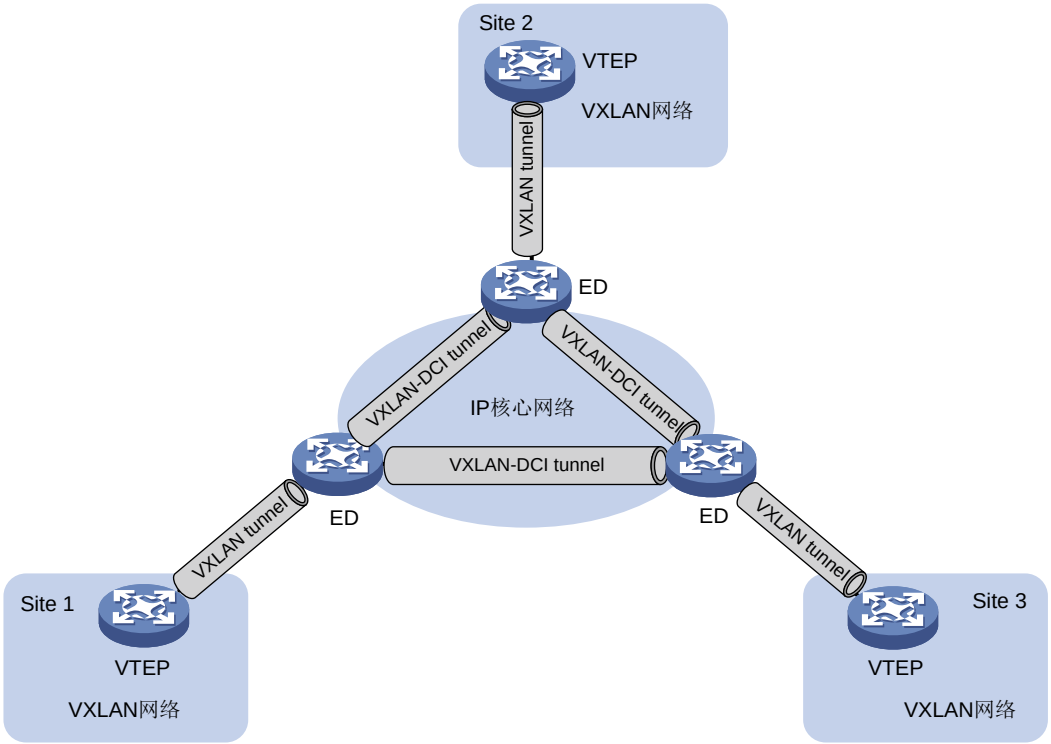
4.1 VXLAN数据中心互联简介

VXLAN 隧道只能用于数据中心内部，实现数据中心内部用户终端的互联。VXLAN-DCI（VXLAN Data Center Interconnection，VXLAN 数据中心互联）隧道可以用来实现数据中心之间的互联。

4.1.1 VXLAN 数据中心互联典型组网

VXLAN 数据中心互联典型组网如图 4-1 所示。VXLAN-DCI 隧道采用 VXLAN 封装格式，该隧道的端点称为 ED（Edge Device，边缘设备）。ED 与数据中心内部的 VTEP 建立 VXLAN 隧道。ED 从 VXLAN 隧道或 VXLAN-DCI 隧道上接收到报文后，解除 VXLAN 封装，根据目的网络重新对报文进行 VXLAN 封装，并将其转发到 VXLAN-DCI 隧道或 VXLAN 隧道，从而实现 VXLAN 跨数据中心之间的互通。

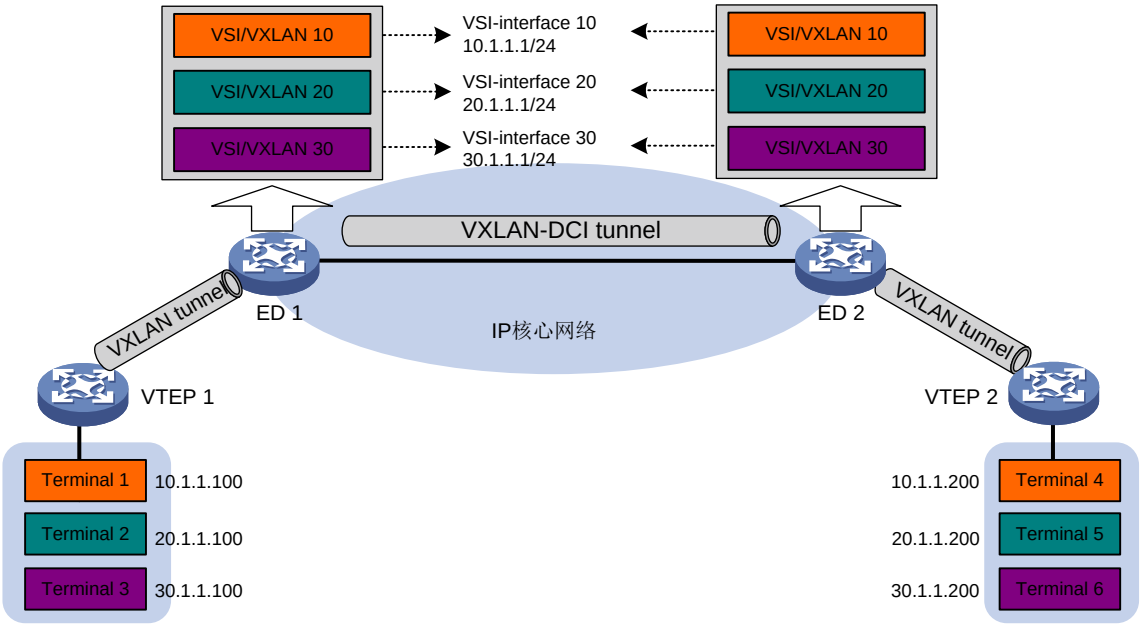
图4-1 VXLAN 数据中心互联典型组网图



4.1.2 VXLAN 数据中心互联工作机制

VXLAN 数据中心互联组网中，VTEP 根据 MAC 地址表项对流量进行二层转发。ED 根据 ARP 表项或 ND 表项对流量进行三层转发。

图4-2 VXLAN 数据中心互联原理图



如图 4-2 所示，所有的 ED 上都需要创建 VSI 虚接口，并为不同 ED 上的相同 VSI 虚接口配置相同的 IP 地址，作为 VXLAN 网络内用户终端的网关地址。在 ED 上还需要开启本地代理 ARP 功能（IPv4 网络）或本地 ND 代理功能（IPv6 网络）。

2. 相同 VXLAN 内不同站点的用户终端通信过程

以 Terminal 1 访问 Terminal 4 为例，相同 VXLAN 内不同站点的用户终端的通信过程为：

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

(2) VTEP 1 收到 ARP 请求消息后，学习 Terminal 1 的 MAC 地址，并在 Terminal 1 所属的 VXLAN 内广播该 ARP 请求。

(3) ED 1 接收到 ARP 请求后，解除 VXLAN 封装，学习 Terminal 1 的 ARP 信息，并代理应答该 ARP 请求，即：向 Terminal 1 发送 ARP 应答消息，应答的 MAC 地址为 VSI 虚接口 10 的 MAC 地址。ARP 应答消息通过 VXLAN 隧道发送给 VTEP 1。

(4) VTEP 1 解除 VXLAN 封装，学习 ED 1 的 MAC 地址，并将 ARP 应答消息转发给 Terminal 1。

(5) Terminal 1 学习 Terminal 4 的 MAC 地址，该地址为 ED 1 上 VSI 虚接口 10 的 MAC 地址。

(6) ED 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，对该消息进行 VXLAN 封装后，将其发送给 VXLAN 10 内的所有远端 ED。

(7) ED 2 对 VXLAN 报文进行解封装后，学习 Terminal 1 的 ARP 信息（IP 为 10.1.1.100、MAC 为 ED 1 上 VSI 虚接口 10 的 MAC、出接口为接收该 VXLAN 报文的 VXLAN-DCI 模式 Tunnel 接口），将 ARP 请求消息中的源 MAC 修改为本地 VSI 虚接口 10 的 MAC 地址，并在 VXLAN 10 的所有 VXLAN 隧道上进行广播。

(8) VTEP 2 收到 ARP 请求后，解除 VXLAN 封装，学习 ED 2 的 MAC 地址，并向本地站点广播该 ARP 请求。

(9) Terminal 4 收到 ARP 请求后，学习 Terminal 1 的 ARP 信息（IP 为 10.1.1.100、MAC 为 ED 2 上 VSI 虚接口 10 的 MAC），并发送 ARP 应答消息给 VTEP 2。

(10) VTEP 2 收到 ARP 应答消息后，查找 MAC 地址表，对报文进行 VXLAN 封装后发送给 ED 2。

(11) ED 2 根据接收到的 ARP 应答消息学习 Terminal 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 ED 1。

(12) ED 1 对 VXLAN 报文进行解封装后，根据收到的 ARP 应答消息学习 Terminal 4 的 ARP 信息（IP 为 10.1.1.200、MAC 为 ED 2 上 VSI 虚接口 10 的 MAC、出接口为接收该 VXLAN 报文的 VXLAN-DCI 模式 Tunnel 接口）。

(13) 通过上述步骤完成 MAC 地址表项和 ARP 表项的学习后，Terminal 1 发送给 Terminal 4 的报文，根据已经学习到的表项进行转发：首先 VTEP 1 添加 VXLAN 封装，发送给 ED 1；ED 1 重新进行 VXLAN 封装后，通过 VXLAN-DCI 隧道将其发送给 ED 2；ED 2 重新封装后，通过 VXLAN 隧道将其发送给 VTEP 2；VTEP 2 将报文转发给 Terminal 4。

3. 不同 VXLAN 间不同站点的用户终端通信过程

以 Terminal 1 访问 Terminal 5 为例，不同 VXLAN 内用户终端的通信过程为：

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

(2) VTEP 1 收到 ARP 请求消息后，学习 Terminal 1 的 MAC 地址，并在 Terminal 1 所属的 VXLAN 内广播该 ARP 请求。

(3) ED 1 接收到 ARP 请求后，解除 VXLAN 封装，学习 Terminal 1 的 ARP 信息，并向 Terminal 1 发送 ARP 应答消息，应答的 MAC 地址为 VSI 虚接口 10 的 MAC 地址。ARP 应答消息通过 VXLAN 隧道发送给 VTEP 1。

(4) VTEP 1 解除 VXLAN 封装，学习 ED 1 的 MAC 地址，并将 ARP 应答消息转发给 Terminal 1。

(5) Terminal 1 学习网关的 ARP 信息，并将访问 Terminal 5 的报文发送给 VTEP 1。

(6) VTEP 1 查找 MAC 地址表项，填加 VXLAN 封装后，将报文发送给 ED 1。

(7) ED 1 收到数据报文后，解除 VXLAN 封装，并根据报文的目的 IP 地址查找路由表。由于目的 IP 地址与 VSI 虚接口 20 的接口 IP 地址在同一网段，ED 1 在 VXLAN 20 内向所有 VTEP 和 ED 广播发送 ARP 请求，获取 Terminal 5 的 MAC 地址。ARP 请求的源 IP 地址为 20.1.1.1、目标 IP 地址为 20.1.1.200、源 MAC 为 ED 1 上 VSI 虚接口 20 的 MAC 地址。

(8) ED 2 接收到 VXLAN 报文后，对其进行解封装，将 ARP 请求中的源 MAC 修改为本地 VSI 虚接口 20 的 MAC 地址，并在 VXLAN 20 的所有 VXLAN 隧道上广播该 ARP 请求。

(9) VTEP 2 收到 ARP 请求后，解除 VXLAN 封装，学习 ED 2 的 MAC 地址，并向本地站点广播该 ARP 请求。

(10) Terminal 5 收到 ARP 请求后，学习 ED 2 的 ARP 信息（IP 为 20.1.1.1、MAC 为 ED 2 上 VSI 虚接口 20 的 MAC 地址），并发送 ARP 应答消息给 VTEP 2。

(11) VTEP 2 收到 ARP 应答消息后，查找 MAC 地址表项，对报文进行 VXLAN 封装后发送给 ED 2。

(12) ED 2 根据接收到的 ARP 应答消息学习 Terminal 5 的 ARP 信息，并通过 VXLAN-DCI 隧道向 ED 1 发送免费 ARP 消息（源 IP 为 20.1.1.200、目标 IP 为 20.1.1.200、源 MAC 为 ED 2 上 VSI 虚接口 20 的 MAC 地址）。

(13) ED 1 对 VXLAN 报文进行解封装后，根据收到的免费 ARP 消息学习 Terminal 5 的 ARP 信息（IP 为 20.1.1.200、 MAC 为 ED 2 上 VSI 虚接口 20 的 MAC 地址、出接口为接收该 VXLAN 报文的 VXLAN-DCI 模式 Tunnel 接口）。

(14) 通过上述步骤 ED 1 学习到了 Terminal 5 的 ARP 信息，ED 2 学习 Terminal 1 的 ARP 信息的过程同上面过程类似。ED 1 和 ED 2 均学习到 Terminal 1 和 Terminal 5 的 ARP 信息后，Terminal 1 发送给 Terminal 5 的报文根据已经学习到的表项进行转发。

4.2 VXLAN数据中心互联配置限制和指导

VXLAN-DCI 隧道和 VXLAN 隧道建议不要共用一个公网侧端口，否则会导致报文丢失。

4.3 VXLAN数据中心互联配置任务简介

在 VXLAN 数据中心互联组网中，各设备上需要进行如下配置：

- IP 核心网络中的设备配置路由协议，确保 ED 之间路由可达。
- ED 和 VTEP 上配置路由协议，确保二者之间路由可达。
- ED 和 VTEP 上配置 VXLAN，在二者之间建立 VXLAN 隧道。
- ED 上配置 VXLAN 数据中心互联，在 ED 之间建立 VXLAN-DCI 隧道。

VXLAN 数据中心互联的大部分配置与 VXLAN 配置、VXLAN IP 网关配置相同。表 4-1 罗列了 VXLAN 数据中心互联支持的所有配置。本文只介绍与 VXLAN、VXLAN IP 网关不同的配置，二者相同的配置请分别参见“2 配置 VXLAN”和“3 VXLAN IP 网关”。

表4-1 VXLAN 数据中心互联配置任务简介

配置任务	说明	详细配置
创建VSI和VXLAN	必选 配置方式与VXLAN相同	2.3
创建VXLAN-DCI隧道	必选	4.4
关联VXLAN与VXLAN-DCI隧道	必选	4.5
配置VSI虚接口	必选	4.6
为VSI指定网关接口	必选	4.7
配置VXLAN报文的目的UDP端口号	可选 配置方式与VXLAN相同	2.10
配置VXLAN报文检查功能	可选 配置方式与VXLAN相同	2.11
配置VXLAN流量统计	可选 配置方式与VXLAN相同，且仅支持VSI的报文统计功能	2.15

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

关于隧道的详细介绍及 Tunnel 接口下的更多配置命令，请参见“三层技术-IP 业务配置指导”中的“隧道”。关于 **interface tunnel**、**source** 和 **destination** 命令的详细介绍，请参见“三层技术-IP 业务命令参考”中的“隧道”。

表4-2 手工创建 VXLAN-DCI 隧道

操作	命令	说明
进入系统视图	**system-view**	-
创建模式为VXLAN-DCI隧道的Tunnel接口，并进入Tunnel接口视图	**interface tunnel** *tunnel-number* **mode vxlan-dci**	缺省情况下，不存在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
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，未设置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
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

4.5 关联VXLAN与VXLAN-DCI隧道

一个 VXLAN 可以关联多条 VXLAN-DCI 隧道。一条 VXLAN-DCI 隧道可以关联多个 VXLAN，这些 VXLAN 共用该 VXLAN-DCI 隧道，ED 根据 VXLAN 报文中的 VXLAN ID 来识别隧道传递的报文所属的 VXLAN。ED 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 ED 将在与该 VXLAN 关联的所有 VXLAN-DCI 隧道上发送该流量，以便将流量转发给所有的远端 VTEP。

表4-3 手工关联 VXLAN 与 VXLAN-DCI 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*	-
配置VXLAN与VXLAN-DCI隧道关联	**tunnel** { *tunnel-number* | **all** }	缺省情况下，VXLAN未关联VXLAN-DCI隧道 ED必须与相同VXLAN内的其它ED建立VXLAN-DCI隧道，并将该隧道与VXLAN关联

4.6 配置VSI虚接口

表4-4 配置 VSI 虚接口

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IPv4地址或IPv6地址	配置VSI虚接口的IPv4地址	**ip address** *ip-address* { *mask* | *mask-length* } [**sub**]	缺省情况下，未配置VSI虚接口的IPv4地址和IPv6地址
	配置VSI虚接口的IPv6地址	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”	
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
开启本地代理ARP功能		**local-proxy-arp enable** [**ip-range** *startIP* **to** *endIP*]	对于IPv4网络，必选 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”
开启本地ND代理功能		**local-proxy-nd enable**	对于IPv6网络，必选 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”
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地址为设备的桥MAC地址+1
（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为1444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＝接口的波特率÷1000（kbps）
（可选）恢复当前接口的缺省配置		**default**	-
（可选）开启VSI虚接口的ARP报文发送限速功能		**arp send-rate** *pps*	缺省情况下，VSI虚接口的ARP报文发送限速功能处于关闭状态
开启当前接口		**undo shutdown**	缺省情况下，接口处于开启状态

4.7 为VSI指定网关接口

表4-5 为 VSI 指定网关接口

操作	命令	说明
进入系统视图	**system-view**	-
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

4.8 VXLAN数据中心互联显示和维护

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 数据中心互联的运行情况，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 数据中心互联的相关信息。

表4-6 VXLAN 数据中心互联显示和维护

操作	命令
显示VSI的信息	**display l2vpn vsi** [**name** *vsi-name*] [**verbose**]
显示Tunnel接口信息	**display interface** [**tunnel** [*number*]] [**brief** [**description** | **down**]]
显示VXLAN关联的VXLAN-DCI隧道信息	**display vxlan tunnel** [**vxlan-id** *vxlan-id*]
清除VSI的报文统计信息	**reset l2vpn statistics vsi** [**name** *vsi-name*]

说明

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

4.9 VXLAN数据中心互联典型配置举例

1. 组网需求

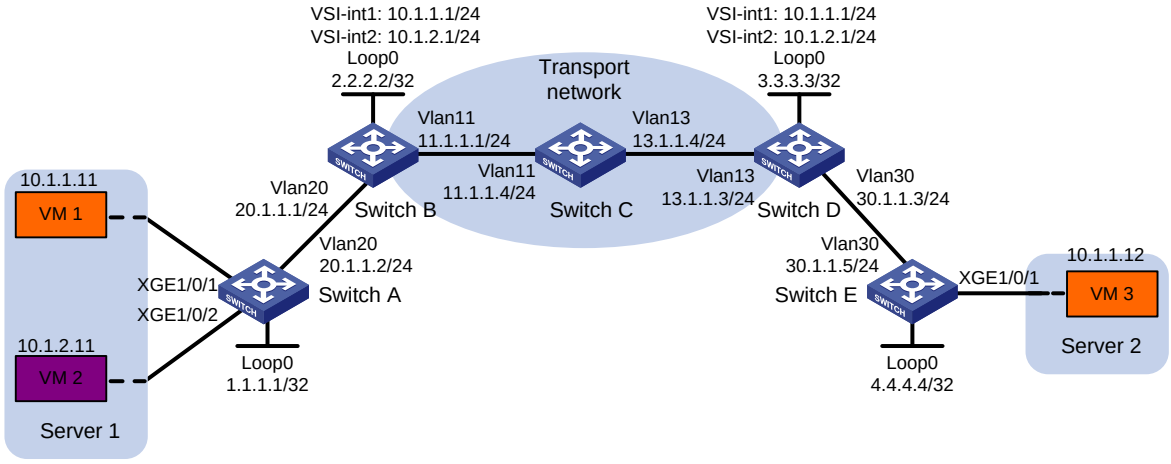
Switch B、Switch D 为 ED。虚拟机 VM 1、VM 3 属于 VXLAN 10，VM 2 属于 VXLAN 20。通过 VXLAN 数据中心互联实现不同数据中心、不同 VXLAN 网络的互联。

具体需求为：

- 手工建立 VXLAN 隧道和 VXLAN-DCI 隧道。
- 手工关联 VXLAN 和 VXLAN 隧道、VXLAN-DCI 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4-3 VXLAN 数据中心互联配置组网图



3. 配置步骤

(1) 配置 IP 地址和单播路由协议

请按照图 4-3 配置各接口的 IP 地址和子网掩码；在各台交换机上配置 OSPF 协议，确保交换机之间路由可达；配置 Switch B 和 Switch D 发布 10.1.1.0/24、10.1.2.0/24 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
```

创建 VSI 实例 vpna 和 VXLAN 10。

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

创建 VSI 实例 vpnb 和 VXLAN 20。

```
[SwitchA] vsi vpnb
[SwitchA-vsi-vpnb] vxlan 20
[SwitchA-vsi-vpnb-vxlan-20] quit
[SwitchA-vsi-vpnb] 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 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
```

配置 Tunnel1 与 VXLAN 10 关联。

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

配置 Tunnel1 与 VXLAN 20 关联。

```
[SwitchA] vsi vpnb
[SwitchA-vsi-vpnb] vxlan 20
[SwitchA-vsi-vpnb-vxlan-20] tunnel 1
[SwitchA-vsi-vpnb-vxlan-20] quit
[SwitchA-vsi-vpnb] quit
```

在接入服务器的接口 Ten-GigabitEthernet1/0/1 上创建以太网服务实例 1000，该实例用来匹配 VLAN 100 的数据帧，并将该服务实例与 VSI 实例 vpna 关联。

```
[SwitchA] interface ten-gigabitethernet 1/0/1
[SwitchA-Ten-GigabitEthernet1/0/1] port link-type trunk
[SwitchA-Ten-GigabitEthernet1/0/1] port trunk permit vlan 100
[SwitchA-Ten-GigabitEthernet1/0/1] service-instance 1000
[SwitchA-Ten-GigabitEthernet1/0/1-srv1000] encapsulation s-vid 100
[SwitchA-Ten-GigabitEthernet1/0/1-srv1000] xconnect vsi vpna
[SwitchA-Ten-GigabitEthernet1/0/1-srv1000] quit
[SwitchA-Ten-GigabitEthernet1/0/1] quit
```

在接入服务器的接口 Ten-GigabitEthernet1/0/2 上创建以太网服务实例 1000，该实例用来匹配 VLAN 200 的数据帧，并将该服务实例与 VSI 实例 vpnb 关联。

```
[SwitchA] interface ten-gigabitethernet 1/0/2
[SwitchA-Ten-GigabitEthernet1/0/2] port link-type trunk
[SwitchA-Ten-GigabitEthernet1/0/2] port trunk permit vlan 200
[SwitchA-Ten-GigabitEthernet1/0/2] service-instance 1000
[SwitchA-Ten-GigabitEthernet1/0/2-srv1000] encapsulation s-vid 200
[SwitchA-Ten-GigabitEthernet1/0/2-srv1000] xconnect vsi vpnb
[SwitchA-Ten-GigabitEthernet1/0/2-srv1000] quit
[SwitchA-Ten-GigabitEthernet1/0/2] quit
```

(3) 配置 Switch B

开启 L2VPN 能力。

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

创建 VSI 实例 vpna 和 VXLAN 10。

```
[SwitchB] vsi vpna
[SwitchB-vsi-vpna] vxlan 10
[SwitchB-vsi-vpna-vxlan-10] quit
[SwitchB-vsi-vpna] quit
```

创建 VSI 实例 vpnb 和 VXLAN 20。

```
[SwitchB] vsi vpnb
[SwitchB-vsi-vpnb] vxlan 20
[SwitchB-vsi-vpnb-vxlan-20] quit
[SwitchB-vsi-vpnb] quit
```

配置接口 Loopback0 的 IP 地址，作为隧道的源端地址。

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

在 Switch B 和 Switch A 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 2.2.2.2。
- 指定隧道的目的端地址为 Switch A 上接口 Loopback0 的地址 1.1.1.1。

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

在 Switch B 和 Switch D 之间建立 VXLAN-DCI 隧道：

- 创建模式为 VXLAN-DCI 的隧道接口 Tunnel2。
- 指定隧道的源端地址为本地接口 Loopback0 的地址 2.2.2.2。
- 指定隧道的目的端地址为 Switch D 上接口 Loopback0 的地址 3.3.3.3。

```
[SwitchB] interface tunnel 2 mode vxlan-dci
[SwitchB-Tunnel2] source 2.2.2.2
[SwitchB-Tunnel2] destination 3.3.3.3
[SwitchB-Tunnel2] quit
```

配置 Tunnel1、Tunnel2 与 VXLAN 10 关联。

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

配置 Tunnel1、Tunnel2 与 VXLAN 20 关联。

```
[SwitchB] vsi vpnb
[SwitchB-vsi-vpnb] vxlan 20
[SwitchB-vsi-vpnb-vxlan-20] tunnel 1
[SwitchB-vsi-vpnb-vxlan-20] tunnel 2
[SwitchB-vsi-vpnb-vxlan-20] quit
[SwitchB-vsi-vpnb] 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，指定该 VSI 虚接口为分布式本地网关接口，并开启本地代理 ARP 功能。

```
[SwitchB] interface vsi-interface 1
[SwitchB-Vsi-interface1] ip address 10.1.1.1 255.255.255.0
[SwitchB-Vsi-interface1] distributed-gateway local
[SwitchB-Vsi-interface1] local-proxy-arp enable
[SwitchB-Vsi-interface1] quit
```

创建 VSI 虚接口 VSI-interface2，并为其配置 IP 地址，该 IP 地址作为 VXLAN 20 内虚拟机的网关地址，指定该 VSI 虚接口为分布式本地网关接口，并开启本地代理 ARP 功能。

```
[SwitchB] interface vsi-interface 2
[SwitchB-Vsi-interface2] ip address 10.1.2.1 255.255.255.0
[SwitchB-Vsi-interface2] distributed-gateway local
[SwitchB-Vsi-interface2] local-proxy-arp enable
[SwitchB-Vsi-interface2] quit
```

开启分布式网关的动态 ARP 表项同步功能。

```
[SwitchB] arp distributed-gateway dynamic-entry synchronize
```

为 VXLAN 10 所在的 VSI 实例 vpna 指定 VSI 虚接口 VSI-interface1。

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

配置 VXLAN 20 所在的 VSI 实例和接口 VSI-interface2 关联。

```
[SwitchB] vsi vpnb
[SwitchB-vsi-vpnb] gateway vsi-interface 2
[SwitchB-vsi-vpnb] quit
```

(4) 配置 Switch D

开启 L2VPN 能力。

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

创建 VSI 实例 vpna 和 VXLAN 10。

```
[SwitchD] vsi vpna
[SwitchD-vsi-vpna] vxlan 10
[SwitchD-vsi-vpna-vxlan-10] quit
[SwitchD-vsi-vpna] quit
```

创建 VSI 实例 vpnb 和 VXLAN 20。

```
[SwitchD] vsi vpnb
[SwitchD-vsi-vpnb] vxlan 20
[SwitchD-vsi-vpnb-vxlan-20] quit
[SwitchD-vsi-vpnb] quit
```

配置接口 Loopback0 的 IP 地址，作为隧道的源端地址。

```
[SwitchD] interface loopback 0
[SwitchD-Loopback0] ip address 3.3.3.3 255.255.255.255
[SwitchD-Loopback0] quit
```

在 Switch D 和 Switch E 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 3.3.3.3。
- 指定隧道的目的端地址为 Switch E 上接口 Loopback0 的地址 4.4.4.4。

```
[SwitchD] interface tunnel 1 mode vxlan
[SwitchD-Tunnel1] source 3.3.3.3
[SwitchD-Tunnel1] destination 4.4.4.4
[SwitchD-Tunnel1] quit
```

在 Switch D 和 Switch B 之间建立 VXLAN-DCI 隧道：

- 创建模式为 VXLAN-DCI 的隧道接口 Tunnel2。
- 指定隧道的源端地址为本地接口 Loopback0 的地址 3.3.3.。
- 指定隧道的目的端地址为 Switch B 上接口 Loopback0 的地址 2.2.2.2。

```
[SwitchD] interface tunnel 2 mode vxlan-dci
[SwitchD-Tunnel2] source 3.3.3.3
[SwitchD-Tunnel2] destination 2.2.2.2
[SwitchD-Tunnel2] quit
```

配置 Tunnel1、Tunnel2 与 VXLAN 10 关联。

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

配置 Tunnel2 与 VXLAN 20 关联。

```
[SwitchD] vsi vpnb
[SwitchD-vsi-vpnb] vxlan 20
[SwitchD-vsi-vpnb-vxlan-20] tunnel 2
[SwitchD-vsi-vpnb-vxlan-20] quit
[SwitchD-vsi-vpnb] 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，指定该 VSI 虚接口为分布式本地网关接口，并开启本地代理 ARP 功能。

```
[SwitchD] interface vsi-interface 1
[SwitchD-Vsi-interface1] ip address 10.1.1.1 255.255.255.0
[SwitchD-Vsi-interface1] distributed-gateway local
[SwitchD-Vsi-interface1] local-proxy-arp enable
[SwitchD-Vsi-interface1] quit
```

创建 VSI 虚接口 VSI-interface2，并为其配置 IP 地址，该 IP 地址作为 VXLAN 20 内虚拟机的网关地址，指定该 VSI 虚接口为分布式本地网关接口，并开启本地代理 ARP 功能。

```
[SwitchD] interface vsi-interface 2
[SwitchD-Vsi-interface2] ip address 10.1.2.1 255.255.255.0
[SwitchD-Vsi-interface2] distributed-gateway local
[SwitchD-Vsi-interface2] local-proxy-arp enable
[SwitchD-Vsi-interface2] quit
```

开启分布式网关的动态 ARP 表项同步功能。

```
[SwitchD] arp distributed-gateway dynamic-entry synchronize
```

为 VXLAN 10 所在的 VSI 实例 vpna 指定 VSI 虚接口 VSI-interface1。

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

配置 VXLAN 20 所在的 VSI 实例和接口 VSI-interface2 关联。

```
[SwitchD] vsi vpnb
[SwitchD-vsi-vpnb] gateway vsi-interface 2
[SwitchD-vsi-vpnb] quit
```

(5) 配置 Switch E

开启 L2VPN 能力。

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

创建 VSI 实例 vpna 和 VXLAN 10。

```
[SwitchE] vsi vpna
[SwitchE-vsi-vpna] vxlan 10
[SwitchE-vsi-vpna-vxlan-10] quit
[SwitchE-vsi-vpna] quit
```

配置接口 Loopback0 的 IP 地址，作为隧道的源端地址。

```
[SwitchE] interface loopback 0
[SwitchE-Loopback0] ip address 4.4.4.4 255.255.255.255
[SwitchE-Loopback0] quit
```

在 Switch E 和 Switch D 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 4.4.4.4。
- 指定隧道的目的端地址为 Switch D 上接口 Loopback0 的地址 3.3.3.3。

```
[SwitchE] interface tunnel 1 mode vxlan
[SwitchE-Tunnel1] source 4.4.4.4
[SwitchE-Tunnel1] destination 3.3.3.3
[SwitchE-Tunnel1] quit
```

配置 Tunnel1 与 VXLAN 10 关联。

```
[SwitchE] vsi vpna
[SwitchE-vsi-vpna] vxlan 10
[SwitchE-vsi-vpna-vxlan-10] tunnel 1
[SwitchE-vsi-vpna-vxlan-10] quit
[SwitchE-vsi-vpna] quit
```

在接入服务器的接口 Ten-GigabitEthernet1/0/1 上创建以太网服务实例 1000，该实例用来匹配 VLAN 100 的数据帧。

```
[SwitchA] interface ten-gigabitethernet 1/0/1
[SwitchA-Ten-GigabitEthernet1/0/1] port link-type trunk
[SwitchA-Ten-GigabitEthernet1/0/1] port trunk permit vlan 100
[SwitchA-Ten-GigabitEthernet1/0/1] service-instance 1000
[SwitchA-Ten-GigabitEthernet1/0/1-srv1000] encapsulation s-vid 100
```

配置以太网服务实例 1000 与 VSI 实例 vpna 关联。

```
[SwitchA-Ten-GigabitEthernet1/0/1-srv1000] xconnect vsi vpna
[SwitchA-Ten-GigabitEthernet1/0/1-srv1000] quit
[SwitchA-Ten-GigabitEthernet1/0/1] quit
```

4. 验证配置

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

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

```
[SwitchB] display interface tunnel
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 2.2.2.2, destination 1.1.1.1
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

Tunnel2
Current state: UP
Line protocol state: UP
Description: Tunnel2 Interface
Bandwidth: 64 kbps
Maximum transmission unit: 1464
Internet protocol processing: Disabled
Last clearing of counters: Never
Tunnel source 2.2.2.2, destination 3.3.3.3
Tunnel protocol/transport UDP_VXLAN_DCI/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 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
Vsi-interface1
Current state: UP
Line protocol state: UP
Description: Vsi-interface1 Interface
Bandwidth: 1000000 kbps
Maximum transmission unit: 1500
Internet address: 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

Vsi-interface2
Current state: UP
Line protocol state: UP
```

```
Description: Vsi-interface2 Interface
Bandwidth: 1000000 kbps
Maximum transmission unit: 1500
Internet address: 10.1.2.1/24 (primary)
IP packet frame type:PKTFMT_ETHNT_2, hardware address: 0011-3300-0102
IPv6 packet frame type:PKTFMT_ETHNT_2, hardware address: 0011-3300-0102
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 VSI 关联的 VSI 虚接口，以及 VXLAN 关联的 VXLAN 隧道、VXLAN-DCI 隧道等信息。

```
[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         : -
  MAC Learning rate       : -
  Drop Unknown            : -
  Flooding                : Enabled
  Gateway interface       : VSI-interface 1
  VXLAN ID                : 10
  Tunnels:
    Tunnel Name           Link ID    State  Type        Flood proxy
    Tunnel1               0x5000001  Up     Manual      Disabled
    Tunnel2               0x5000002  Up     Manual      Disabled

VSI Name: vpnb
  VSI Index               : 0
  VSI State               : Up
  MTU                     : 1500
  Bandwidth               : -
  Broadcast Restrain      : -
  Multicast Restrain      : -
  Unknown Unicast Restrain: -
  MAC Learning            : Enabled
  MAC Table Limit         : -
  MAC Learning rate       : -
  Drop Unknown            : -
  Flooding                : Enabled
```

```
  Gateway interface       : VSI-interface 2
  VXLAN ID                : 20
  Tunnels:
    Tunnel Name           Link ID    State  Type        Flood proxy
    Tunnel1               0x5000001  Up     Manual      Disabled
    Tunnel2               0x5000002  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    VLAN/VSI   Interface/Link ID        Aging Type
11.1.1.4        000c-29c1-5e46 N/A        Vlan11                   19    D
10.1.1.11       dc2d-cbb5-cf09 0          Tunnel1                  20    D
10.1.1.12       0011-4400-0102 0          Tunnel2                  20    D
10.1.2.11       dc2d-cbb5-cf89 1          Tunnel1                  20    D
```

(2) 验证主机

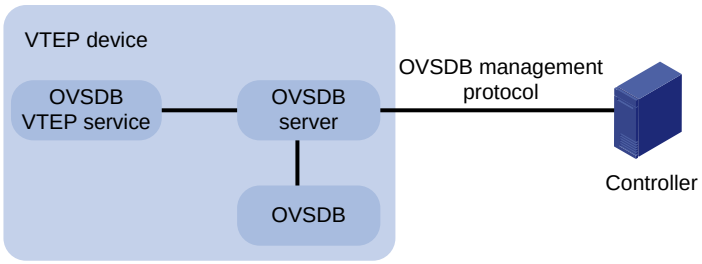
虚拟机 VM 1、VM 2、VM 3 之间可以互访。

5 OVSDB-VTEP

5.1 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5-1 OVSDB-VTEP 示意图



如图 5-1 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。

5.2 协议规范

与 OVSDB 相关的协议规范有：

- RFC 7047：The Open vSwitch Database Management Protocol

5.3 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 5-1 所示配置。

表5-1 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	5.5.1
	与控制器建立被动SSL连接		5.5.2
	与控制器建立主动TCP连接		5.5.3
	与控制器建立被动TCP连接		5.5.4
开启OVSDB服务器		必选	5.6
开启OVSDB VTEP服务		必选	5.7
配置VXLAN隧道的全局源地址		必选	5.8
指定用户侧的接入端口		必选	5.9
开启禁止控制器下发的ACL在VTEP上生效功能		可选	5.10

5.4 配置准备

在进行 OVSDB-VTEP 相关配置前，需要首先通过 **l2vpn enable** 命令开启 L2VPN 功能，该命令的详细介绍请参见“MPLS 命令参考”中的“MPLS L2VPN”。

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，详细配置方法请参见“安全配置指导”中的“SSL”。

5.5 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。
- 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。
- 主动 TCP 连接：OVSDB 服务器主动向控制器发起 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。

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。

5.5.1 与控制器建立主动 SSL 连接

表5-2 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 PKI域需提前配置，具体方法请参见“安全配置指导”中的“PKI”
（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 如果指定的CA证书文件不存在，则使用开启OVSDB服务器时通过SSL连接获取的自签名证书，并通过本命令指定证书文件名
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 OVSDB服务器最多可以同时与8个控制器建立主动SSL连接

5.5.2 与控制器建立被动 SSL 连接

表5-3 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 PKI域需提前配置，具体方法请参见“安全配置指导”中的“PKI”
（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 如果指定的CA证书文件不存在，则使用开启OVSDB服务器时通过SSL连接获取的自签名证书，并通过本命令指定证书文件名
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 OVSDB服务器只能监听1个端口的SSL连接请求

5.5.3 与控制器建立主动 TCP 连接

表5-4 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 OVSDB服务器最多可以同时与8个控制器建立主动TCP连接

5.5.4 与控制器建立被动 TCP 连接

表5-5 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**	-
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 OVSDB服务器只能监听1个端口的TCP连接请求

5.6 开启OVSDB服务器

在开启 OVSDB 服务器之前，必须先建立 OVSDB 连接。如果在开启 OVSDB 服务器之后修改 OVSDB 连接，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。

表5-6 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

5.7 开启OVSDB VTEP服务

表5-7 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

5.8 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5-8 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

5.9 指定用户侧的接入端口

为了在控制器上显示 VTEP 上的端口并对其进行控制，必须在 VTEP 上将该端口配置为用户侧的接入端口。

将接口配置为接入侧端口后，建议不要在该接口下进行手工配置，避免影响控制器对接入侧端口的管理。

表5-9 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，当前接口不是用户侧的接入端口

5.10 开启禁止控制器下发的ACL在VTEP上生效功能

1. 功能简介

在 OVSDB-VTEP 组网中，控制器通过 OVSDB 控制协议下发 ACL 到 VTEP，占用 VTEP 上的 ACL 资源。通过配置本命令，用户可以禁止控制器下发的 ACL 在 VTEP 上生效，以便节约设备上的 ACL 资源。

2. 配置步骤

操作	命令	说明
进入系统视图	**system-view**	-
开启禁止控制器下发的ACL在VTEP上生效功能	**vtep acl disable**	缺省情况下，控制器下发的ACL在VTEP上生效

5.11 OVSDB-VTEP典型配置举例

5.11.1 OVSDB-VTEP 头端复制配置举例

1. 组网需求

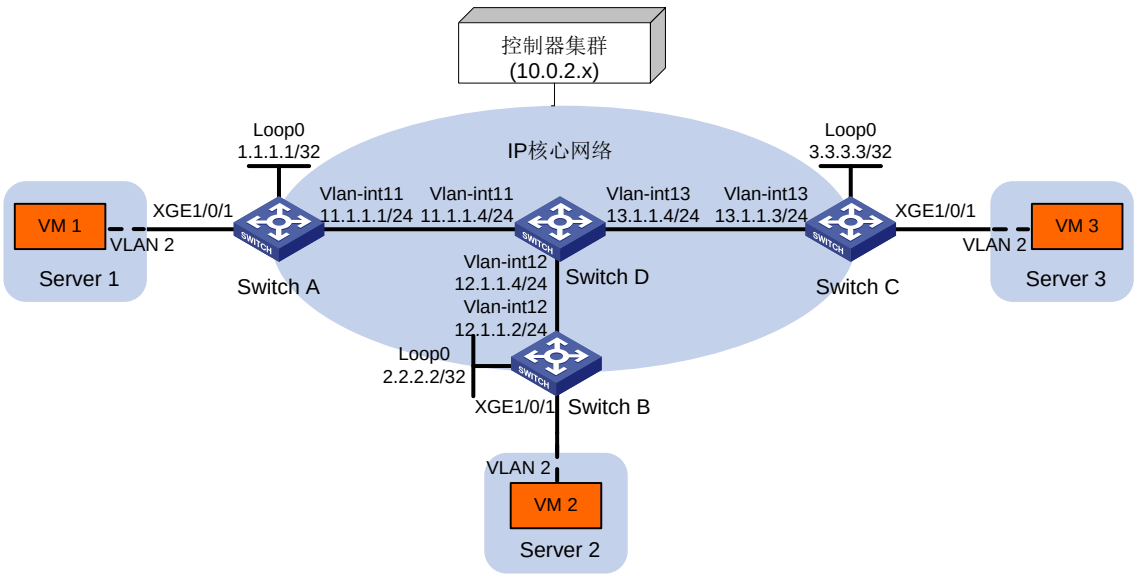
Switch A、Switch B、Switch C 为与服务器连接的 VTEP 设备。虚拟机 VM 1、VM 2 和 VM 3 同属于 VXLAN 10。通过 VXLAN 实现不同站点间的二层互联，确保虚拟机在站点之间进行迁移时用户的访问流量不会中断。

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5-2 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

请按照图 5-2 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
```

配置与控制器建立主动 SSL 连接，SSL 连接使用的 PKI 域为 a，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 pki domain a
[SwitchA] ovsdb server ssl 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
```

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
```

指定接入服务器的接口 Ten-GigabitEthernet1/0/1 上为用户侧的接入端口。

```
[SwitchA] interface ten-gigabitethernet 1/0/1
[SwitchA-Ten-GigabitEthernet1/0/1] vtep access port
[SwitchA-Ten-GigabitEthernet1/0/1] quit
```

(3) 配置 Switch B

开启 L2VPN 功能。

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

配置与控制器建立主动 SSL 连接，SSL 连接使用的 PKI 域为 a，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 pki domain a
[SwitchB] ovsdb server ssl 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
```

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

```
[SwitchB] interface loopback 0
[SwitchB-LoopBack0] ip address 2.2.2.2 255.255.255.255
[SwitchB-LoopBack0] quit
[SwitchB] tunnel global source-address 2.2.2.2
```

指定接入服务器的接口 Ten-GigabitEthernet1/0/1 上为用户侧的接入端口。

```
[SwitchB] interface ten-gigabitethernet 1/0/1
[SwitchB-Ten-GigabitEthernet1/0/1] vtep access port
[SwitchB-Ten-GigabitEthernet1/0/1] quit
```

(4) 配置 Switch C

开启 L2VPN 功能。

```
<SwitchC> system-view
```

```
[SwitchC] l2vpn enable
```
配置与控制器建立主动 SSL 连接，SSL 连接使用的 PKI 域为 a，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 pki domain a
[SwitchC] ovsdb server ssl 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
```
配置接口 Loopback0 的 IP 地址，作为 VXLAN 隧道的全局源地址。
```
[SwitchC] interface loopback 0
[SwitchC-LoopBack0] ip address 3.3.3.3 255.255.255.255
[SwitchC-LoopBack0] quit
[SwitchC] tunnel global source-address 3.3.3.3
```
指定接入服务器的接口 Ten-GigabitEthernet1/0/1 上为用户侧的接入端口。
```
[SwitchC] interface ten-gigabitethernet 1/0/1
[SwitchC-Ten-GigabitEthernet1/0/1] vtep access port
[SwitchC-Ten-GigabitEthernet1/0/1] 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         : -
  MAC Learning rate       : -
  Drop Unknown            : -
  Flooding                : Enabled
  VXLAN ID                : 10
  Tunnels:
    Tunnel Name          Link ID    State  Type        Flood proxy
    Tunnel1              0x5000001  Up     Manual      Disabled
    Tunnel2              0x5000002  Up     Manual      Disabled
  ACs:
    AC                              Link ID  State       Type
    XGE1/0/1 srv2                   0        Up          Manual
```

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
dc2d-cb9c-6cdb   Dynamic  evpn2014                        Tunnel1       Aging
dc2d-cb9c-23dc   Dynamic  evpn2014                        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。