CloudEngine 12800, 12800E, 8800, 7800, 6800, 5800 系列交换机

VXLAN 最佳实践

文档版本 03

发布日期 2021-10-21





版权所有 © 华为技术有限公司 2021。 保留一切权利。

非经本公司书面许可,任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部,并不得以任何形式传播。

商标声明



nuawe和其他华为商标均为华为技术有限公司的商标。 本文档提及的其他所有商标或注册商标,由各自的所有人拥有。

注意

您购买的产品、服务或特性等应受华为公司商业合同和条款的约束,本文档中描述的全部或部分产品、服务或 特性可能不在您的购买或使用范围之内。除非合同另有约定,华为公司对本文档内容不做任何明示或暗示的声 明或保证。

由于产品版本升级或其他原因,本文档内容会不定期进行更新。除非另有约定,本文档仅作为使用指导,本文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。

华为技术有限公司

地址: 深圳市龙岗区坂田华为总部办公楼 邮编: 518129

网址: https://e.huawei.com

目录

1 VXLAN 最佳实践概述	1
1.1 VXLAN 简介	1
1.2 VXLAN 基本概念	3
1.3 VXLAN 优势	6
1.4 VXLAN 原理介绍	6
1.4.1 网关划分	6
1.4.2 报文格式	8
1.4.3 EVPN 基本原理	10
1.4.4 BGP EVPN 方式部署集中式网关	14
1.4.5 BGP EVPN 方式部署分布式网关	24
2 VXLAN 网络规划设计	38
2.1 硬件集中式 VXLAN	38
2.1.1 GW/Spine/Leaf 三层架构的硬件集中式 VXLAN 设计	38
2.1.2 Spine/Leaf 二层架构的硬件集中式 VXLAN 设计	39
2.1.3 防火墙接入设计	41
2.1.4 负载均衡器接入设计	41
2.1.5 服务器接入设计	42
2.1.6 逻辑隔离设计	44
2.1.7 流量转发模型设计	45
2.1.8 路由协议设计	46
2.1.9 可靠性设计	50
2.1.10 约束限制	52
2.2 硬件分布式 VXLAN	52
2.2.1 GW/Spine/Leaf 三层架构的硬件分布式 VXLAN 设计	52
2.2.2 Spine/Leaf 二层架构的硬件分布式 VXLAN 设计	54
2.2.3 防火墙接入设计	55
2.2.4 负载均衡器接入设计	56
2.2.5 服务器接入设计	56
2.2.6 逻辑隔离设计	58
2.2.7 流量转发模型设计	59
2.2.8 路由协议设计	61
2.2.9 可靠性设计	65
2.2.10 约束限制	67

3 VXLAN 部署建议	68
3.1 硬件集中式 VXLAN	68
3.1.1 GW/Spine/Leaf 三层架构的硬件集中式 VXLAN	68
3.1.1.1 配置 Leaf 堆叠工作组	78
3.1.1.2 配置 Leaf M-LAG 工作组	82
3.1.1.3 配置 Leaf SVF 工作组	87
3.1.1.4 配置 Spine 节点	92
3.1.1.5 配置网关工作组	97
3.1.1.6 配置增强功能	105
3.1.2 Spine/Leaf 二层架构的硬件集中式 VXLAN	108
3.1.2.1 配置 Gateway 对接 Leaf 的接口地址	118
3.1.2.2 配置 Leaf 上的路由	119
3.1.2.3 配置网关上的路由	120
3.1.2.4 配置 BGP EVPN	122
3.2 硬件分布式 VXLAN	122
3.2.1 GW/Spine/Leaf 三层架构的硬件分布式 VXLAN	123
3.2.1.1 配置 Leaf 堆叠工作组	134
3.2.1.2 配置 Leaf M-LAG 工作组	139
3.2.1.3 配置 Leaf SVF 工作组	144
3.2.1.4 配置 Spine 节点	149
3.2.1.5 配置网关工作组	154
3.2.1.6 配置增强功能	162
3.2.2 Spine/Leaf 二层架构的硬件分布式 VXLAN	165
3.2.2.1 配置 Gateway 对接 Leaf 的接口地址	176
3.2.2.2 配置 Leaf 上的路由	177
3.2.2.3 配置网关上的路由	178
3224 配置 RGP FVPN	180

VXLAN 最佳实践概述

1.1 VXLAN 简介

介绍VXLAN的定义、目的和收益。

定义

RFC7348定义了VLAN扩展方案VXLAN(Virtual eXtensible Local Area Network)。 VXLAN采用MAC in UDP(User Datagram Protocol)封装方式,是NVO3(Network Virtualization over Layer 3)中的一种网络虚拟化技术。

目的

作为云计算的核心技术之一,服务器虚拟化凭借其大幅降低IT成本、提高业务部署灵活性、降低运维成本等优势已经得到越来越多的认可和部署。

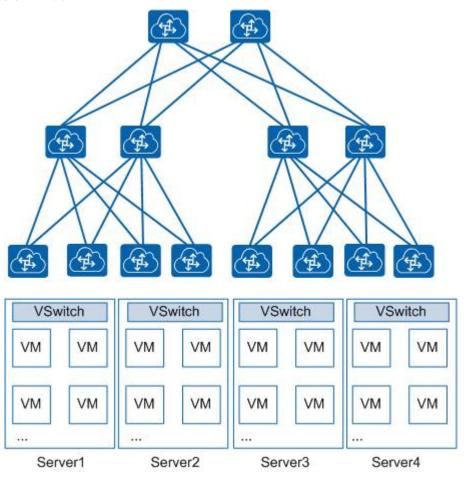


图 1-1 服务器虚拟化示意图

如<mark>图</mark>1所示,一台服务器可虚拟多台虚拟机,而一台虚拟机相当于一台主机。主机的数量发生了数量级的变化,这也为虚拟网络带来了如下问题:

• 虚拟机规模受网络规格限制

在传统二层网络环境下,数据报文是通过查询MAC地址表进行二层转发,而MAC 地址表的容量限制了虚拟机的数量。

网络隔离能力限制

当前主流的网络隔离技术是VLAN,在大规模的虚拟化网络中部署存在如下限制:

- 由于IEEE 802.1Q中定义的VLAN Tag域只有12比特,仅能表示4096个 VLAN,无法满足大二层网络中标识大量租户或租户群的需求。
- 传统二层网络中的VLAN无法满足网络动态调整的需求。
- 虚拟机迁移范围受网络架构限制

虚拟机启动后,可能由于服务器资源等问题(如CPU过高,内存不够等),需要将虚拟机迁移到新的服务器上。为了保证虚拟机迁移过程中业务不中断,则需要保证虚拟机的IP地址、MAC地址等参数保持不变,这就要求业务网络是一个二层网络,且要求网络本身具备多路径的冗余备份和可靠性。

针对大二层网络, VXLAN的提出很好地解决了上述问题:

针对虚拟机规模受网络规格限制

VXLAN将虚拟机发出的数据包封装在UDP中,并使用物理网络的IP、MAC地址作为外层头进行封装,对网络只表现为封装后的参数。因此,极大降低了大二层网络对MAC地址规格的需求。

• 针对网络隔离能力限制

VXLAN引入了类似VLAN ID的用户标识,称为VXLAN网络标识VNI(VXLAN Network Identifier),由24比特组成,支持多达16M的VXLAN段,从而满足了大量的用户标识。

• 针对虚拟机迁移范围受网络架构限制

VXLAN通过采用MAC in UDP封装来延伸二层网络,将以太报文封装在IP报文之上,通过路由在网络中传输,无需关注虚拟机的MAC地址。且路由网络无网络结构限制,具备大规模扩展能力、故障自愈能力、负载均衡能力。通过路由网络,虚拟机迁移不受网络架构限制。

收益

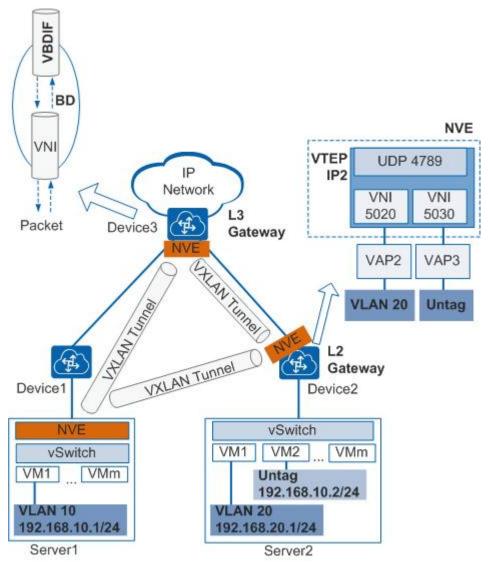
随着数据中心在物理网络基础设施上实施服务器虚拟化的快速发展,作为NVO3技术之一的VXLAN:

- 通过24比特的VNI可以支持多达16M的VXLAN段的网络隔离,对用户进行隔离和标识不再受到限制,可满足海量租户。
- 除VXLAN网络边缘设备,网络中的其他设备不需要识别虚拟机的MAC地址,减轻 了设备的MAC地址学习压力,提升了设备性能。
- 通过采用MAC in UDP封装来延伸二层网络,实现了物理网络和虚拟网络解耦,租户可以规划自己的虚拟网络,不需要考虑物理网络IP地址和广播域的限制,大大降低了网络管理的难度。

1.2 VXLAN 基本概念

VXLAN是NVO3中的一种网络虚拟化技术,通过将虚拟机发出的数据包封装在UDP中,并使用物理网络的IP、MAC作为outer-header进行封装,然后在IP网络上传输,到达目的地后由隧道终结点解封装并将数据发送给目标虚拟机。

图 1-2 VXLAN 结构示意图



通过VXLAN,虚拟网络可接入大量租户,且租户可以规划自己的虚拟网络,不需要考虑物理网络IP地址和广播域的限制,降低了网络管理的难度。如<mark>表1</mark>介绍VXLAN相关概念。

表 1-1 VXLAN 基本概念

概念	描述
Underlay网络 和Overlay网络	VXLAN技术将已有的物理网络作为Underlay网络,在其上构建出虚拟的二层或三层网络,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,实现租户报文在不同站点间传递。对于租户来说,Underlay网络是透明的,只能感知到Overlay网络。

概念	描述
NVE (Network Virtualization Edge)	网络虚拟边缘节点NVE,实现网络虚拟化功能的网络实体。报文 经过NVE封装转换后,NVE间就可基于三层基础网络建立二层虚 拟化网络。 说明 设备和服务器上的虚拟交换机VSwitch都可以作为NVE。
	按照NVE部署位置的不同,可以分为以下三种模式:
	● 硬件模式:所有的NVE都部署在支持NVE的设备上,所有的 VXLAN报文封装与解封装都在设备上进行。
	软件模式:所有的NVE都部署在vSwitch上,所有的VXLAN报 文封装与解封装都在vSwitch上进行。
	 混合模式:部分NVE部署在vSwitch上,部分NVE部署在支持 NVE的设备上,在vSwitch和设备上都有可能会进行VXLAN报 文封装与解封装。
VTEP (VXLAN Tunnel	VTEP是VXLAN隧道端点,封装在NVE中,用于VXLAN报文的封装 和解封装。
Endpoints)	VTEP与物理网络相连,分配有物理网络的IP地址,该地址与虚拟 网络无关。
	VXLAN报文中源IP地址为本节点的VTEP地址,VXLAN报文中目的 IP地址为对端节点的VTEP地址,一对VTEP地址就对应着一个 VXLAN隧道。
VNI(VXLAN Network	VXLAN网络标识VNI类似VLAN ID,用于区分VXLAN段,不同 VXLAN段的虚拟机不能直接二层相互通信。
Identifier)	一个VNI表示一个租户,即使多个终端用户属于同一个VNI,也表示一个租户。VNI由24比特组成,支持多达16M的租户。
	在分布式网关部署场景下,VNI分为二层VNI和三层VNI。
	● 二层VNI是普通的VNI,以1:1方式映射到广播域BD,实现 VXLAN报文同子网的转发。
	● 三层VNI和VPN实例进行关联,用于VXLAN报文跨子网的转 发。
BD (Bridge	BD是VXLAN网络中转发数据报文的二层广播域。
Domain)	在VXLAN网络中,将VNI以1:1方式映射到广播域BD,BD成为 VXLAN网络转发数据报文的实体。
VBDIF接口	基于BD创建的三层逻辑接口。通过VBDIF接口配置IP地址可实现不同网段的VXLAN间,及VXLAN和非VXLAN的通信,也可实现二层网络接入三层网络。
VAP (Virtual Access Point)	虚拟接入点VAP,即VXLAN业务接入点,可以是二层子接口或 VLAN:
	当接入节点是二层子接口时,通过在二层子接口上配置流封装 类型实现不同的接口接入不同的数据报文,将二层子接口关联 广播域BD后,可实现数据报文通过BD转发。
	当业务接入点是VLAN时,需要将VLAN绑定到广播域BD,也可以实现数据报文通过BD转发。

概念	描述
网关 (Gateway)	和VLAN类似,不同VNI之间的VXLAN,及VXLAN和非VXLAN之间 不能直接相互通信。为了使VXLAN之间,以及VXLAN和非VXLAN 之间能够进行通信,VXLAN引入了VXLAN网关。
	VXLAN网关分为:
	● 二层网关:用于解决租户接入VXLAN虚拟网络的问题,也可用于同一VXLAN虚拟网络的子网通信。
	● 三层网关:用于VXLAN虚拟网络的跨子网通信以及外部网络的 访问。

1.3 VXLAN 优势

随着数据中心在物理网络基础设施上实施服务器虚拟化的快速发展,作为NVO3技术之一的VXLAN:

- 通过24比特的VNI可以支持多达16M的VXLAN段的网络隔离,对用户进行隔离和标识不再受到限制,可满足海量租户。
- VNI可以分为二层VNI和三层VNI,既可以和广播域BD关联,也可以和VPN实例关 联。因此VXLAN可以支持L2 VPN、L3 VPN等复杂业务。
- 除VXLAN网络边缘设备,网络中的其他设备不需要识别虚拟机的MAC地址,减轻 了设备的MAC地址学习压力,提升了设备性能。
- 通过采用MAC in UDP封装来延伸二层网络,实现了物理网络和虚拟网络解耦,租户可以规划自己的虚拟网络,不需要考虑物理网络IP地址和广播域的限制,大大降低了网络管理的难度。
- VXLAN封装的UDP源端口,由内层的流信息Hash而来,Underlay网络不需要解析 内层报文就可负载分担,实现网络高吞吐量。

1.4 VXLAN 原理介绍

1.4.1 网关划分

和VLAN类似,不同VNI之间的VXLAN,及VXLAN和非VXLAN之间不能直接相互通信。 为了使VXLAN之间,以及VXLAN和非VXLAN之间能够进行通信,VXLAN引入了VXLAN 网关。

VXLAN网关分为:

- 二层网关:用于解决租户接入VXLAN虚拟网络的问题,也可用于同一VXLAN虚拟网络的子网通信。
- 三层网关:用于VXLAN虚拟网络的跨子网通信以及外部网络的访问。

根据三层网关部署方式的不同,VXLAN三层网关又可以分为集中式网关和分布式网 关。

集中式网关部署

集中式网关是指将三层网关集中部署在一台设备上,如<mark>图</mark>1所示,所有跨子网的流量都 经过三层网关进行转发,实现流量的集中管理。

L2 Gateway
Leaf1
Leaf2
Server1 Server2
10.1.1.1/24 10.10.1.1/24
B子网流量

图 1-3 VXLAN 集中式网关组网图

部署集中式网关的优点和缺点如下:

- 优点:对跨子网流量进行集中管理,网关的部署和管理比较简单。
- 缺点:
 - 转发路径不是最优:同一二层网关下跨子网的数据中心三层流量都需要经过 集中三层网关转发。
 - ARP表项规格瓶颈:由于采用集中三层网关,通过三层网关转发的终端租户 的ARP表项都需要在三层网关上生成,而三层网关上的ARP表项规格有限,这 不利于数据中心网络的扩展。

分布式网关部署

• 产生原因

通过部署分布式网关可以解决集中式网关部署的缺点。VXLAN分布式网关是指在典型的"Spine-Leaf"组网结构下,将Leaf节点作为VXLAN隧道端点VTEP,每个Leaf节点都可作为VXLAN三层网关,Spine节点不感知VXLAN隧道,只作为VXLAN报文的转发节点。如图2所示,Server1和Server2不在同一个网段,但是都连接到一个Leaf节点。Server1和Server2通信时,流量只需要在Leaf1节点进行转发,不再需要经过Spine节点。

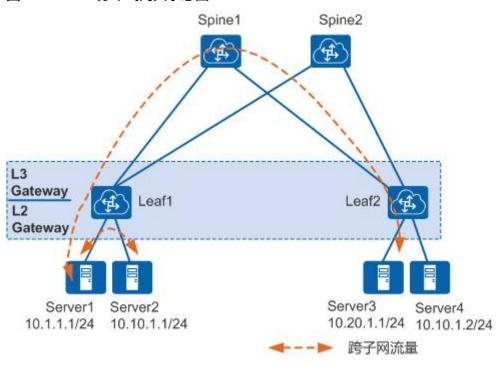


图 1-4 VXLAN 分布式网关示意图

Spine节点:关注于高速IP转发,强调的是设备的高速转发能力。 Leaf节点:

- 作为VXLAN网络中的二层网关设备,与物理服务器或VM对接,用于解决终端 租户接入VXLAN虚拟网络的问题。
- 作为VXLAN网络中的三层网关设备,进行VXLAN报文封装/解封装,实现跨子网的终端租户通信,以及外部网络的访问。

• 分布式网关的特点

VXLAN分布式网关具有如下特点:

- 同一个Leaf节点既可以做VXLAN二层网关,也可以做VXLAN三层网关,部署 灵活。
- Leaf节点只需要学习自身连接服务器的ARP表项,而不必像集中三层网关一样,需要学习所有服务器的ARP表项,解决了集中式三层网关带来的ARP表项 瓶颈问题,网络规模扩展能力强。

1.4.2 报文格式

VXLAN是MAC in UDP的网络虚拟化技术,所以其报文封装是在原始以太报文之前添加了一个UDP封装及VXLAN头封装。具体报文格式如<mark>图</mark>1所示。

图 1-5 VXLAN 报文格式

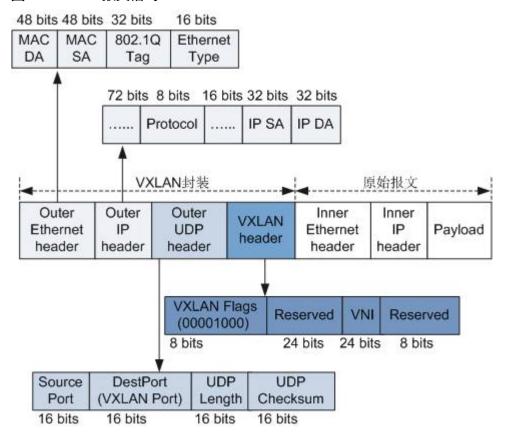


表 1-2 VXLAN 报文格式说明

字段	描述	
VXLAN header	VXLAN Flags: 8比特,取值为00001000。VNI: VXLAN网络标识,24比特,用于区分VXLAN段。Reserved: 24比特和8比特,必须设置为0。	
Outer UDP header	DestPort: 目的UDP端口号是4789。Source Port: 源端口号是内层以太报文头通过哈希算法计算后的值。	
Outer IP header	IP SA:源IP地址是VXLAN隧道本端VTEP的IP地址。IP DA:目的IP地址是VXLAN隧道远端VTEP的IP地址。	

字段	描述
Outer Ethernet header	MAC DA: 在发送报文的虚拟机所属VTEP上根据目的 VTEP地址查找路由表,路由表中下一跳IP地址对应的 MAC地址。
	● MAC SA:发送报文的虚拟机所属VTEP的MAC地址。
	● 802.1Q Tag:可选字段,该字段为报文中携带的VLAN Tag。
	● Ethernet Type:以太报文类型,IP协议报文该字段取值 为0x0800。

1.4.3 EVPN 基本原理

EVPN(Ethernet Virtual Private Network)是一种用于二层网络互联的VPN技术。EVPN技术采用类似于BGP/MPLS IP VPN的机制,在BGP协议的基础上定义了一种新的NLRI(Network Layer Reachability Information,网络层可达信息)即EVPN NLRI,EVPN NLRI定义了几种新的BGP EVPN路由类型,用于处在二层网络的不同站点之间的MAC地址学习和发布。

原有的VXLAN实现方案没有控制平面,是通过数据平面的流量泛洪进行VTEP发现和主机信息(包括IP地址、MAC地址、VNI、网关VTEP IP地址)学习的,这种方式导致数据中心网络存在很多泛洪流量。为了解决这一问题,VXLAN引入了EVPN作为控制平面,通过在VTEP之间交换BGP EVPN路由实现VTEP的自动发现、主机信息相互通告等特性,从而避免了不必要的数据流量泛洪。

综上所述,EVPN通过扩展BGP协议新定义了几种BGP EVPN路由,这些BGP EVPN路由可以用于传递VTEP地址和主机信息,因此EVPN应用于VXLAN网络中,可以使VTEP发现和主机信息学习从数据平面转移到控制平面。

BGP EVPN 路由

在EVPN NLRI中定义了如下几种应用于VXLAN控制平面的BGP EVPN路由类型:

Type2路由——MAC/IP路由

该类型路由的报文格式如下图所示:

图 1-6 MAC/IP 路由的报文格式

Route Distinguisher (8字	节)
Ethernet Segment Identifie	er (10字节)
Ethernet Tag ID(4 字节)	
MAC Address Length (1 :	字节)
MAC Address (6字节)	
IP Address Length (1字	5)
IP Address (0或4或16字	节)
MPLS Label1 (3字节)	
MPLS Label2 (0或3字节)

各字段的解释如下表所示:

字段	说明
Route Distinguisher	该字段为EVPN实例下设置的RD(Route Distinguisher)值。
Ethernet Segment Identifier	该字段为当前设备与对端连接定义的唯 一标识。
Ethernet Tag ID	该字段为当前设备上实际配置的VLAN ID。
MAC Address Length	该字段为此路由携带的主机MAC地址的 长度。
MAC Address	该字段为此路由携带的主机MAC地址。
IP Address Length	该字段为此路由携带的主机IP地址的掩码 长度。
IP Address	该字段为此路由携带的主机IP地址。
MPLS Label1	该字段为此路由携带的二层VNI。
MPLS Label2	该字段为此路由携带的三层VNI。

该类型路由在VXLAN控制平面中的作用包括:

● 主机MAC地址通告

要实现同子网主机的二层互访,两端VTEP需要相互学习主机MAC。作为BGP EVPN对等体的VTEP之间通过交换MAC/IP路由,可以相互通告已经获取到的主机 MAC。其中,MAC Address Length和MAC Address字段为主机MAC地址。

● 主机ARP通告

MAC/IP路由可以同时携带主机MAC地址+主机IP地址,因此该路由可以用来在VTEP之间传递主机ARP表项,实现主机ARP通告。其中,MAC Address和MAC Address Length字段为主机MAC地址,IP Address和IP Address Length字段为主机IP地址。此时的MAC/IP路由也称为ARP类型路由。主机ARP通告主要用于以下两种场景:

- a. ARP广播抑制。当三层网关学习到其子网下的主机ARP时,生成主机信息(包含主机IP地址、主机MAC地址、二层VNI、网关VTEP IP地址),然后通过传递ARP类型路由将主机信息同步到二层网关上。这样当二层网关再收到ARP请求时,先查找是否存在目的IP地址对应的主机信息,如果存在,则直接将ARP请求报文中的广播MAC地址替换为目的单播MAC地址,实现广播变单播,达到ARP广播抑制的目的。
- b. 分布式网关场景下的虚拟机迁移。当一台虚拟机从当前网关迁移到另一个网关下之后,新网关学习到该虚拟机的ARP(一般通过虚拟机发送免费ARP实现),并生成主机信息(包含主机IP地址、主机MAC地址、二层VNI、网关VTEP IP地址),然后通过传递ARP类型路由将主机信息发送给虚拟机的原网关。原网关收到后,感知到虚拟机的位置发生变化,触发ARP探测,当探测不到原位置的虚拟机时,撤销原位置虚拟机的ARP和主机路由。

● 主机IP路由通告

在分布式网关场景中,要实现跨子网主机的三层互访,两端VTEP(作为三层网 关)需要互相学习主机IP路由。作为BGP EVPN对等体的VTEP之间通过交换 MAC/IP路由,可以相互通告已经获取到的主机IP路由。其中,IP Address Length和IP Address字段为主机IP路由的目的地址,同时MPLS Label2字段必须携带三层VNI。此时的MAC/IP路由也称为IRB(Intergrate Routing and Bridge)类型路由。

□ 说明

ARP类型路由携带的有效信息有: 主机MAC地址+主机IP地址+二层VNI; IRB类型路由携带的有效信息有: 主机MAC地址+主机IP地址+二层VNI+三层VNI。因此,IRB类型路由包含着ARP类型路由,不仅可以用于主机IP路由通告,也能用于主机ARP通告。

Type3路由——Inclusive Multicast路由

该类型路由是由前缀和PMSI属性组成,报文格式如下图所示:

图 1-7 Inclusive Multicast 路由的报文格式

前缀

Route Distinguisher(8 字节)	
Ethernet Tag ID(4 字节)	
IP Address Length (1字节)	
Originating Router's IP Address	(4或16字节)

PMSI属性

Flags(1字节)
Tunnel Type(1字节)
MPLS Label(3 字节)
Tunnel Identifier (variable)

各字段的解释如下表所示:

字段	说明
Route Distinguisher	该字段为EVPN实例下设置的RD(Route Distinguisher)值。
Ethernet Tag ID	该字段为当前设备上的VLAN ID。在此路由中为全0。
IP Address Length	该字段为此路由携带的本端VTEP IP地址 的掩码长度。
Originating Router's IP Address	该字段为此路由携带的本端VTEP IP地址。
Flags	该字段为标志位,标识当前隧道是否需 要叶子节点信息。
	在VXLAN场景中,该字段没有实际意 义。

字段	说明
Tunnel Type	该字段为此路由携带的隧道类型。目前,在VXLAN场景中,支持的类型只有 "6: Ingress Replication",即头端复制,用于BUM报文转发。
MPLS Label	该字段为此路由携带的二层VNI。
Tunnel Identifier	该字段为此路由携带的隧道信息。目 前,在VXLAN场景中,该字段也是本端 VTEP IP地址。

该类型路由在VXLAN控制平面中主要用于VTEP的自动发现和VXLAN隧道的动态建立。 作为BGP EVPN对等体的VTEP,通过Inclusive Multicast路由互相传递二层VNI和VTEP IP地址信息。其中,Originating Router's IP Address字段为本端VTEP IP地址,MPLS Label字段为二层VNI。如果对端VTEP IP地址是三层路由可达的,则建立一条到对端的 VXLAN隧道。同时,如果对端VNI与本端相同,则创建一个头端复制表,用于后续 BUM报文转发。

Type5路由——IP前缀路由

该类型路由的报文格式如下图所示:

图 1-8 Figure 1-3 IP 前缀路由的报文格式

Route Dist	tinguisher (8字节)
Ethernet S	Segment Identifier (10字节)
Ethernet T	ag ID (4字节)
IP Prefix L	ength (1字节)
IP Prefix (4或16 字节)
GW IP Ad	dress (4或16 字节)
MPLS Lab	el (3字节)

各字段的解释如下表所示:

字段	说明
Route Distinguisher	该字段为EVPN实例下设置的RD(Route Distinguisher)值。
Ethernet Segment Identifier	该字段为当前设备与对端连接定义的唯 一标识。
Ethernet Tag ID	该字段为当前设备上实际配置的VLAN ID。
IP Prefix Length	该字段为此路由携带的IP前缀掩码长度。
IP Prefix	该字段为此路由携带的IP前缀。

字段	说明
GW IP Address	该字段为默认网关地址。该字段在 VXLAN场景中没有实际意义。
MPLS Label	该字段为此路由携带的三层VNI。

该类型路由的IP Prefix Length和IP Prefix字段既可以携带主机IP地址,也可以携带网段地址:

- 当携带主机IP地址时,该类型路由在VXLAN控制平面中的作用与IRB类型路由是一样的,主要用于分布式网关场景中的主机IP路由通告。
- 当携带网段地址时,通过传递该类型路由,可以实现VXLAN网络中的主机访问外部网络。

1.4.4 BGP EVPN 方式部署集中式网关

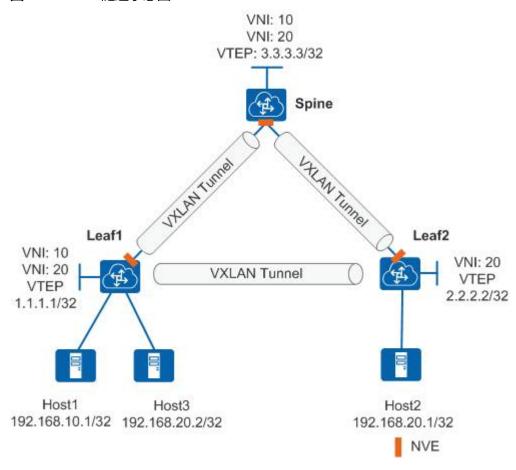
在BGP EVPN方式部署集中式网关的场景中,控制平面的流程包括VXLAN隧道建立、MAC地址动态学习;转发平面的流程包括同子网已知单播报文转发、同子网BUM(Broadcast&Unknown-unicast&Multicast)报文转发、跨子网报文转发。该方式通过部署EVPN协议实现VTEP自动发现和VXLAN隧道的动态创建,灵活性高,适合大规模的VXLAN组网场景,如果在VXLAN网络中采用集中式网关,推荐使用该方式。

VXLAN 隧道建立

VXLAN隧道由一对VTEP IP地址确定,创建VXLAN隧道实际上是两端VTEP获取对端 VTEP IP地址的过程,只要对端VTEP IP地址是三层路由可达的,VXLAN隧道就可以建立成功。通过BGP EVPN方式动态建立VXLAN隧道,就是在两端VTEP之间建立BGP EVPN对等体,然后对等体之间利用BGP EVPN路由来互相传递VNI和VTEP IP地址信息,从而实现动态建立VXLAN隧道。

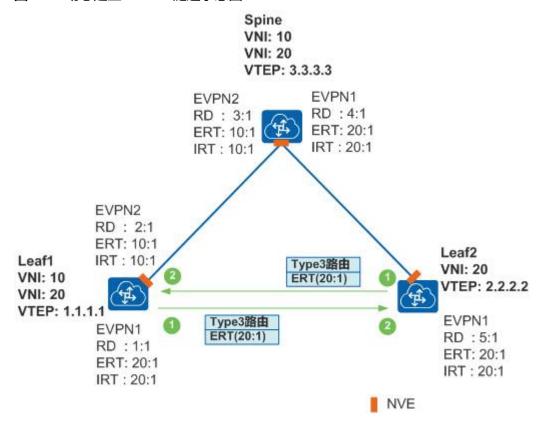
如<mark>图1</mark>所示,Leaf1上部署了两个Host,Leaf2上部署了一个Host,Spine上部署三层网关。为了实现Host3和Host2之间的通信,需要在Leaf1和Leaf2之间创建VXLAN隧道;为了实现Host1和Host2之间的通信,需要在Leaf1和Spine之间以及Spine和Leaf2之间创建VXLAN隧道。对于Host1和Host3之间的通信,虽然都属于Leaf1,但由于属于不同子网,需要经过三层网关Spine,因此也需要在Leaf1和Spine之间创建VXLAN隧道。

图 1-9 VXLAN 隧道示意图



下面结合<mark>图2</mark>,以Leaf1和Leaf2为例,介绍一下通过BGP EVPN方式动态建立VXLAN隧道的过程:

图 1-10 动态建立 VXLAN 隧道示意图



1. 首先在Leaf1和Leaf2之间建立BGP EVPN对等体。然后,在Leaf1和Leaf2上分别创建二层广播域,并在二层广播域下配置关联的VNI。接下来在二层广播域下创建EVPN实例,配置本端EVPN实例的RD、出方向VPN-Target(ERT)、入方向VPN-Target(IRT)。在配置完本端VTEP IP地址后,Leaf1和Leaf2会生成BGP EVPN路由并发送给对端,该路由携带本端EVPN实例的出方向VPN-Target和BGP EVPN协议新定义的Type3路由即Inclusive Multicast路由。其中,Inclusive Multicast路由如图3所示,由前缀和PMSI属性组成,VTEP IP地址存放在前缀的Originating Router's IP Address字段中,VNI存放在PMSI属性的MPLS Label字段中。

图 1-11 Inclusive Multicast 路由

前缀

Route Distinguisher(8 字节)
Ethernet Tag ID(4 字节)
IP Address Length(1 字节)
Originating Router's IP Address(4或16 字节)

PMSI属性

Flags(1 字节)
Tunnel Type(1 字节)
MPLS Label(3字节)
Tunnel Identifier(variable)

2. Leaf1和Leaf2在收到对端发来的BGP EVPN路由后,首先检查该路由携带的EVPN 实例的出方向VPN-Target,如果与本端EVPN实例的入方向VPN-Target相等,则接收该路由,否则丢弃该路由。在接收该路由后,Leaf1和Leaf2将获取其中携带的对端VTEP IP地址和VNI,如果对端VTEP IP地址是三层路由可达的,则建立一条到对端的VXLAN隧道;同时,如果对端VNI与本端相同,则创建一个头端复制表,用于后续BUM报文转发。

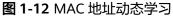
Leaf1和Spine之间、Leaf2和Spine之间通过BGP EVPN方式动态建立VXLAN隧道的过程与上述相同,这里不再赘述。

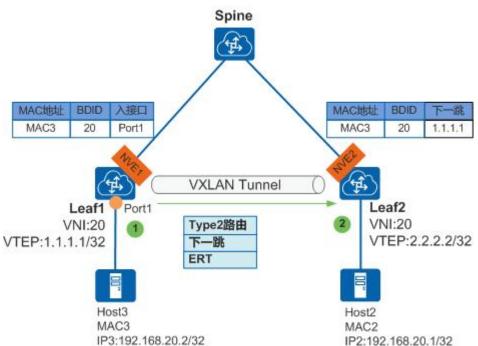
□ 说明

VPN-Target是一种32位的BGP扩展团体属性,一个EVPN实例可以配置出方向和入方向两类VPN-Target,两端EVPN实例的VPN-Target要相互匹配(即本端EVPN实例配置的出方向VPN-target值需要与对端EVPN实例配置的入方向VPN-target值相等),才能相互交换EVPN路由,否则VXLAN隧道无法建立成功。如果仅有一端匹配成功可以接收路由,则此端设备上可以建立通往另一端设备的隧道,但是无法传输数据报文,因为另一端设备在收到报文后,将会检查本端是否有通往对端的VXLAN隧道,如果没有的话,将会丢弃报文。

MAC 地址动态学习

在VXLAN网络中,为了实现终端租户的互通,支持MAC地址动态学习,不需要网络管理员手工维护,大大减少了维护工作量。下面结合<mark>图4</mark>,详细介绍一下同子网主机互通时,MAC地址动态学习的过程:





1. Host3首次与Leaf1通信时,通过动态ARP报文,Leaf1学习到Host3的MAC地址、BDID(二层广播域标识)和报文入接口(即二层子接口对应的物理接口Port1)的对应关系,记录在本地MAC表中。同时,Leaf1根据Host3的ARP表项生成BGPEVPN路由并发送给对等体Leaf2,该路由携带本端EVPN实例的出方向VPN-Target、路由下一跳属性以及BGPEVPN协议新定义的Type2路由即MAC/IP路由。其中,路由下一跳属性携带的是本端VTEPIP地址;MAC/IP路由如图5所示,Host3的MAC地址存放在MAC Address Length和MAC Address字段中,二层VNI存放在MPLS Label1字段中。

图 1-13 MAC/IP 路由



2. Leaf2收到Leaf1发来的BGP EVPN路由后,首先检查该路由携带的EVPN实例的出方向VPN-Target,如果与本端EVPN实例的入方向VPN-Target相等,则接收该路由,否则丢弃该路由。在接收该路由后,Leaf2获得Host3的MAC地址、BDID和Leaf1上VTEP IP地址(下一跳属性)的对应关系,并保存到本地的MAC表中。

Leaf1学习Host2的主机MAC的过程与上述过程相同,这里不再赘述。

Host3初次与Host2通信时,首先发送目的MAC为全F、目的IP为IP2的ARP请求报文,请求Host2的MAC地址。缺省情况下,Leaf1收到该ARP请求后将在本网段进行广播,

为了减少广播报文,此时可以在Leaf1上使能ARP广播抑制功能。这样当Leaf1收到该ARP请求报文时,先根据目的IP检查本地是否有Host2的MAC地址,如果有则将目的MAC替换为Host2的MAC地址,将ARP请求的广播报文变为单播报文,然后通过VXLAN隧道发给Leaf2。Leaf2收到后转发给Host2,Host2收到该ARP请求后学习到Host3的MAC地址,并以单播形式进行ARP应答。Host3收到ARP应答报文后学习到Host2的MAC地址。至此,Host3和Host2互相学习到对方的MAC地址,后续双方将采用单播通信。

□ 说明

- 在跨子网主机互通时,只需在主机和三层网关之间进行MAC地址动态学习,与上述过程相同。
- Leaf节点也可以通过数据转发过程学习主机MAC地址,但是依赖于设备从数据报文中学习MAC地址的能力。在BGP EVPN方式建立VXLAN隧道的场景中,Leaf节点之间可以通过BGP EVPN路由动态学习主机MAC地址,不再依赖数据转发过程。此时,可以关闭Leaf节点从数据报文中学习MAC地址的能力,提升转发效率。

同子网已知单播报文转发

同子网已知单播报文转发只在VXLAN二层网关之间进行,三层网关无需感知。报文转发流程如图6所示。

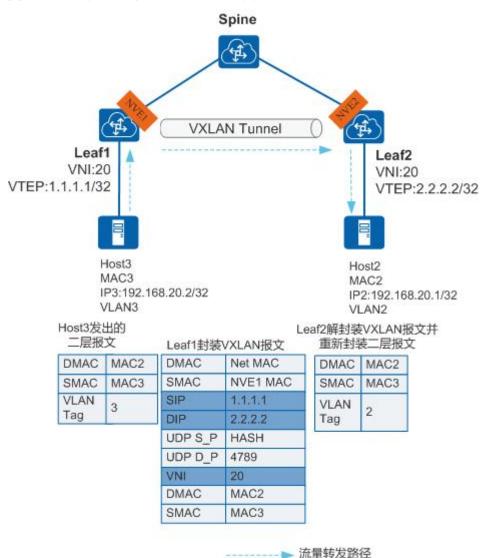


图 1-14 同子网已知单播报文转发示意图

- 1. Leaf1收到来自Host3的报文,根据报文中接入的端口和VLAN信息获取对应的二层 广播域,并在该二层广播域内查找出接口和封装信息。
- 2. Leaf1上VTEP根据查找到的封装信息对数据报文进行VXLAN封装,然后根据查找到的出接口进行报文转发。
- 3. Leaf2上VTEP收到VXLAN报文后,根据UDP目的端口号、源/目的IP地址、VNI判断VXLAN报文的合法有效性。然后依据VNI获取对应的二层广播域,进行VXLAN解封装,获取内层的二层报文。
- 4. Leaf2根据内层二层报文的目的MAC,从本地MAC表找到对应的出接口和封装信息,为报文添加VLAN Tag,转发给对应的主机Host2。

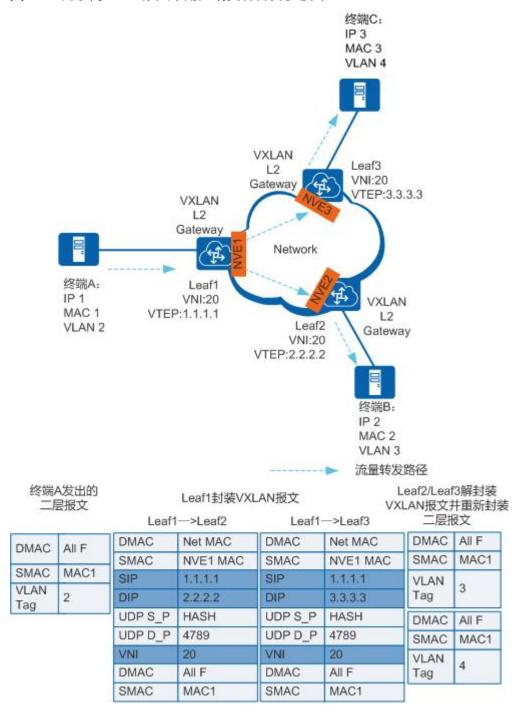
Host2向Host3发送报文的过程类似,这里不再赘述。

同子网 BUM 报文转发

同子网BUM报文转发只在VXLAN二层网关之间进行,三层网关无需感知。同子网BUM报文转发可以采用头端复制方式。

头端复制是指,当BUM报文进入VXLAN隧道时,接入端VTEP根据头端复制列表进行报文的VXLAN封装,并将报文发送给头端复制列表中的所有出端口VTEP。BUM报文出VXLAN隧道时,出口端VTEP对报文解封装。BUM报文采用头端复制的转发流程如图7所示。

图 1-15 同子网 BUM 报文采用头端复制转发示意图

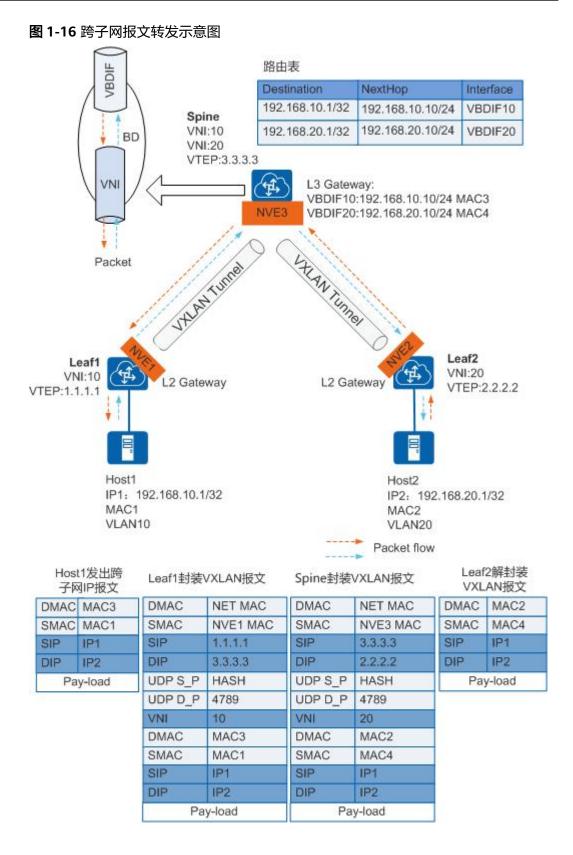


1. Leaf1收到来自终端A的报文,根据报文中接入的端口和VLAN信息获取对应的二层 广播域。

- 2. Leaf1上VTEP根据对应的二层广播域获取对应VNI的头端复制隧道列表,依据获取的隧道列表进行报文复制,并进行VXLAN封装。然后将封装后的报文从出接口转发出去。
- 3. Leaf2/Leaf3上VTEP收到VXLAN报文后,根据UDP目的端口号、源/目的IP地址、VNI判断VXLAN报文的合法有效性。然后依据VNI获取对应的二层广播域,进行VXLAN解封装,获取内层二层报文。
- 4. Leaf2/Leaf3检查内层二层报文的目的MAC,发现是BUM MAC,在对应的二层广播域内的非VXLAN隧道侧进行广播处理,即: Leaf2/Leaf3分别从本地MAC表中找到非VXLAN隧道侧的所有出接口和封装信息,为报文添加VLAN Tag,转发给对应的终端B/C。

跨子网报文转发

跨子网报文转发需要通过三层网关实现。在集中式网关场景中,跨子网报文转发的流程如<mark>图8</mark>所示。



1. Leaf1收到来自Host1的报文,根据报文中接入的端口和VLAN信息获取对应的二层 广播域,在对应的二层广播域内查找出接口和封装信息。

- 2. Leaf1上VTEP根据查找到的出接口和封装信息进行VXLAN封装,向Spine转发报文。
- 3. Spine收到VXLAN报文后进行解封装,发现内层报文中的目的MAC是三层网关接口VBDIF10的MAC地址MAC3,判断需要进行三层转发。
- 4. Spine剥除内层报文的以太封装,解析目的IP。根据目的IP查找路由表,找到目的IP的下一跳地址,再根据下一跳地址查找ARP表项,获取目的MAC、VXLAN隧道出接口及VNI等信息。
- 5. Spine重新封装VXLAN报文,向Leaf2转发。其中内层报文以太头中的源MAC是三层网关接口VBDIF20的MAC地址MAC4。
- 6. Leaf2上VTEP收到VXLAN报文后,根据UDP目的端口号、源/目的IP地址、VNI判断VXLAN报文的合法有效性。依据VNI获取对应的二层广播域,然后进行VXLAN解封装,获取内层二层报文,并在对应的二层广播域内查找出接口和封装信息。
- 7. Leaf2根据查找到的出接口和封装信息,为报文添加VLAN Tag,转发给对应的 Host2。

Host2向Host1发送报文的过程类似,这里不再赘述。

1.4.5 BGP EVPN 方式部署分布式网关

在BGP EVPN方式部署分布式网关的场景中,控制平面的流程包括VXLAN隧道建立、MAC地址动态学习;转发平面的流程包括同子网已知单播报文转发、同子网BUM报文转发、跨子网报文转发。该方式实现的功能全面,支持主机IP路由通告、主机MAC地址通告、主机ARP通告,可以直接使能ARP广播抑制功能。如果在VXLAN网络中采用分布式网关,推荐使用此方式。

VXLAN 隧道建立

VXLAN隧道由一对VTEP IP地址确定,创建VXLAN隧道实际上是两端VTEP获取对端VTEP IP地址的过程,只要对端VTEP IP地址是三层路由可达的,VXLAN隧道就可以建立成功。通过BGP EVPN方式动态建立VXLAN隧道,就是在两端VTEP之间建立BGP EVPN对等体,然后对等体之间利用BGP EVPN路由来互相传递VNI和VTEP IP地址信息,从而实现动态建立VXLAN隧道。

在分布式网关的场景中,Leaf同时作为二层网关和三层网关,Spine节点不感知VXLAN隧道,只作为VXLAN报文的转发节点。控制平面只需在Leaf之间建立VXLAN隧道,如图2所示,Leaf1与Leaf2之间建立VXLAN隧道用于Host1和Host2、Host3和Host2之间的通信。对于Host1和Host3之间的通信,由于都属于Leaf1,所以互访的流量只需在Leaf1上处理,无需通过VXLAN隧道转发。

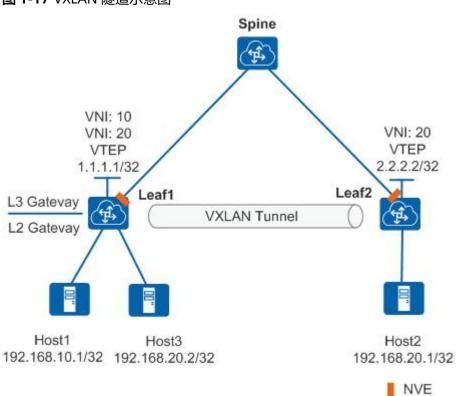


图 1-17 VXLAN 隧道示意图

在分布式网关场景中,通过BGP EVPN方式动态建立VXLAN隧道有以下两种情况:

同子网互通

如<mark>图2</mark>所示,当处于同一子网的Host3与Host2互通时,只需要进行二层转发。通过BGP EVPN方式建立VXLAN隧道的过程如下:

Spine 书 Leaf1 Leaf2 VNI: 20 VNI: 20 VTEP: 1.1.1.1 VTEP: 2.2.2.2 Type3路由 EVPN1 ERT(20:1) EVPN1 RD: 2:1 RD: 1:1 ERT: 20:1 ERT: 20:1 Type3路由 IRT: 20:1 IRT: 20:1 ERT(20:1) NVE

图 1-18 动态建立 VXLAN 隧道示意图-1

1. 首先在Leaf1和Leaf2之间建立BGP EVPN对等体。然后,在Leaf1和Leaf2上分别创建二层广播域,并在二层广播域下配置关联的二层VNI。接下来在二层广播域下创建EVPN实例,配置本端EVPN实例的RD、出方向VPN-Target(ERT)、入方向VPN-Target(IRT)。在配置完本端VTEP IP地址后,Leaf1和Leaf2会生成BGP EVPN路由并发送给对端,该路由携带本端EVPN实例的出方向VPN-Target和BGP EVPN协议新定义的Type3路由即Inclusive Multicast路由。其中,Inclusive Multicast路由如图3所示,由前缀和PMSI属性组成,VTEP IP地址存放在前缀的Originating Router's IP Address字段中,二层VNI存放在PMSI属性的MPLS Label字段中。

图 1-19 Inclusive Multicast 路由

前缀

Route Distinguisher(8 字节)
Ethernet Tag ID(4 字节)
IP Address Length(1 字节)
Originating Router's IP Address(4或16 字节)

PMSI属性

Flags(1 字节)
Tunnel Type(1 字节)
MPLS Label(3字节)
Tunnel Identifier(variable)

2. Leaf1和Leaf2在收到对端发来的BGP EVPN路由后,首先检查该路由携带的EVPN 实例的出方向VPN-Target,如果与本端EVPN实例的入方向VPN-Target相等,则 接收该路由,否则丢弃该路由。在接收该路由后,Leaf1和Leaf2将获取其中携带 的对端VTEP IP地址和二层VNI,如果对端VTEP IP地址是三层路由可达的,则建立一条到对端的VXLAN隧道;同时,如果对端二层VNI与本端相同,则创建一个头端复制表,用于后续BUM(Broadcast&Unknown-unicast&Multicast)报文转发。

□ 说明

VPN-Target是一种32位的BGP扩展团体属性,一个EVPN实例可以配置出方向和入方向两类VPN-Target,两端EVPN实例的VPN-Target要相互匹配(即本端EVPN实例配置的出方向VPN-target值需要与对端EVPN实例配置的入方向VPN-target值相等),才能相互交换EVPN路由,否则VXLAN隧道无法建立成功。如果仅有一端匹配成功可以接收路由,则此端设备上可以建立通往另一端设备的隧道,但是无法传输数据报文,因为另一端设备在收到报文后,将会检查本端是否有通往对端的VXLAN隧道,如果没有的话,将会丢弃报文。

跨子网互通

当处于不同子网的Host1与Host2互通时,需要进行三层转发,因此在通过BGP EVPN 方式建立VXLAN隧道的过程中,网关Leaf1和Leaf2需要发布下属主机的IP路由。一般情况下,这里发布的是32位主机IP路由,因为在VXLAN网络中,不同的Leaf节点可能连接着相同的网段,所以如果Leaf节点发布的是下属主机IP所在的网段路由,则可能与其他Leaf节点发布的网段路由冲突,进而导致某些Leaf节点的下属主机不可达。只有在如下两种场景中,Leaf节点可发布网段路由:

- Leaf节点连接的网段在整个VXLAN网络中是唯一的,而且有效的主机明细路由数量较大,此时可发布主机IP所在的网段路由,从而减轻Leaf节点上路由存储的压力。
- VXLAN网络中的主机需要访问外部网络,此时Leaf节点可在VXLAN网络中发布其连接的外部网段路由,从而使其他Leaf节点学习到去往外部网络的路由。

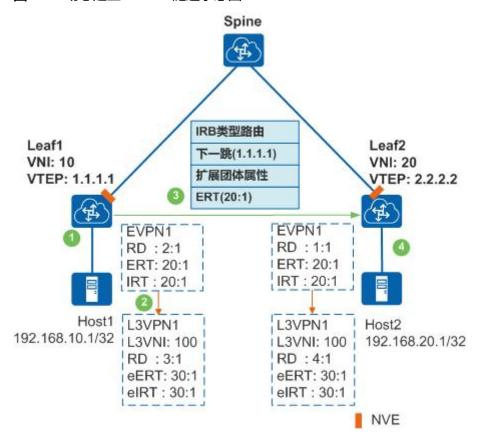
建立VXLAN隧道之前,需要在Leaf1和Leaf2上进行如下配置准备:

配置任务	作用
创建二层广播域(BD),并在二层广播 域下配置关联的二层VNI。	创建VXLAN网络转发数据报文的实体。
在Leaf1和Leaf2之间建立BGP EVPN对等体。	用于交换BGP EVPN路由。
在二层广播域下创建EVPN实例,配置 EVPN实例的RD、出方向VPN-Target (ERT)、入方向VPN-Target(IRT)。	用于生成BGP EVPN路由。
为不同租户创建L3VPN实例,将L3VPN 实例绑定到指定二层广播域的VBDIF接口 上。	用于区分和隔离不同租户的IP路由表。
为L3VPN实例指定关联的三层VNI。	用于Leaf节点在收到数据报文时判断使 用哪个L3VPN实例的路由表指导转发。
配置L3VPN实例到EVPN实例的出方向 VPN-Target(eERT),以及从EVPN实 例到L3VPN实例的入方向VPN-Target (eIRT)。	用于控制本端L3VPN实例与对端EVPN实例之间BGP EVPN路由的发布和接收。
配置Leaf1和Leaf2之间发布的路由类型。	用于发布Host1和Host2的主机IP路由。 这里有IRB(Intergrate Routing and Bridge)和IP前缀两种路由类型,可根据 如下原则进行选择:
	● IRB类型路由只能发布32位主机IP路由。由于IRB类型路由包含着ARP类型路由,因此选择发布IRB类型路由后,可直接在Leaf节点上使能ARP广播抑制功能,无需其他配置,同时也利于虚拟机的迁移。如果只需发布32位主机IP路由,建议选择发布IRB类型路由。
	• IP前缀类型路由既能发布32位主机IP路由又能发布网段路由。但是,在发布32位主机IP路由时,需先将主机IP地址生成直连路由,这会影响虚拟机的迁移。因此,如果只需发布32位主机IP路由,不建议选择IP前缀类型路由。只有需要发布网段路由时(场景见上文),才需选择IP前缀类型路由。

根据主机IP路由发布方式的不同,动态建立VXLAN隧道的过程有以下两种:

● 通过IRB类型路由发布主机IP路由(如图4所示)

图 1-20 动态建立 VXLAN 隧道示意图-2



1. Host1首次与Leaf1通信时,通过动态ARP报文,Leaf1学习到Host1的ARP表项。同时,Leaf1根据Host1所在的二层广播域找到绑定VBDIF接口的L3VPN实例,获取关联的三层VNI。然后Leaf1上的EVPN实例根据上述信息生成IRB类型路由,如图5所示。其中,主机IP地址存放在IP Address Length和IP Address字段中,三层VNI存放在MPLS Label2字段中。

图 1-21 IRB 类型路由



2. Leaf1上的EVPN实例从IRB类型路由中获取Host1的主机IP地址、三层VNI,发给本端L3VPN实例,L3VPN实例在其路由表中保存Host1的主机IP路由,如图6所示。

图 1-22 本端主机 IP 路由

目的地址	三层VNI	下一跳	出接口
192.168.10.1/32	100	网关接口地址	VBDIF

- 3. Leaf1向Leaf2发送BGP EVPN路由,该路由携带本端EVPN实例的ERT、扩展团体属性、路由下一跳属性以及IRB类型路由。其中,扩展团体属性携带的是隧道类型(取值是VXLAN隧道)、本端VTEP的MAC地址;路由下一跳属性携带的是本端的VTEP IP地址
- 4. Leaf2收到Leaf1发来的BGP EVPN路由后,同时进行如下处理:
 - 检查该路由携带的ERT,如果与本端EVPN实例的IRT相同,则接收该路由,否则丢弃该路由。EVPN实例获取到IRB类型路由后,还能提取到其中包含的ARP类型路由,用于主机ARP通告。
 - 检查该路由携带的ERT,如果与本端L3VPN实例的eIRT相同,则接收该路由,否则丢弃该路由。然后,L3VPN实例获取到该路由携带的IRB类型路由,从中提取Host1的主机IP地址、三层VNI,在其路由表中保存Host1的主机IP路由,并将路由的下一跳迭代出接口设置为VXLAN隧道接口,如图7所示。

图 1-23 远端主机 IP 路由



- 在通过EVPN实例或L3VPN实例接收该路由后,Leaf2通过下一跳属性获取 Leaf1的VTEP IP地址,如果该VTEP IP地址是三层路由可达的,则建立一条到 Leaf1的VXLAN隧道。

Leaf1建立到Leaf2的VXLAN隧道的过程与上述相同,这里不再赘述。

● 通过IP前缀类型路由发布主机IP路由(如图8所示)

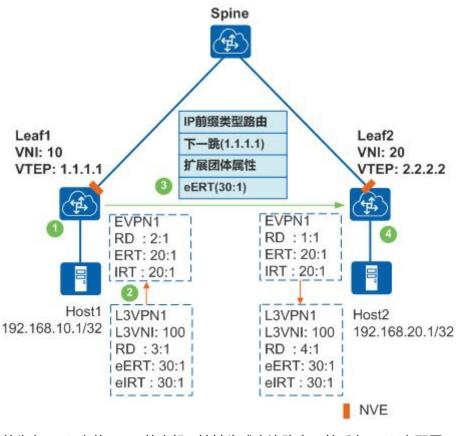


图 1-24 动态建立 VXLAN 隧道示意图-3

1. 首先在Leaf1上将Host1的主机IP地址生成直连路由,然后在Leaf1上配置L3VPN实例引入直连路由,这样Host1的主机IP路由就保存到L3VPN实例的路由表中,并添加L3VPN实例关联的三层VNI,如图9所示。

图 1-25 本端主机 IP 路由

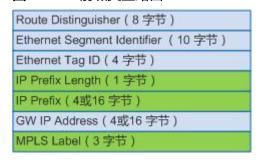
目的地址	三层VNI	下一跳	出接口
192.168.10.1/32	100	网关接口地址	VBDIF

□ 说明

如果要发布网段路由,这里需要先利用动态路由协议(如OSPF等)发布该网段路由,再配置L3VPN实例引入对应的动态协议路由。

2. 在Leaf1上配置L3VPN实例向EVPN实例发布IP路由后,L3VPN实例下Host1的主机IP路由将发布给EVPN实例。然后由EVPN实例生成IP前缀类型路由,如<mark>图10</mark>所示。其中,主机IP地址存放在IP Prefix Length和IP Prefix字段中,三层VNI存放在MPLS Label字段中。

图 1-26 IP 前缀类型路由



- 3. Leaf1向Leaf2发送BGP EVPN路由,该路由携带本端L3VPN实例的eERT、扩展团体属性、路由下一跳属性以及IP前缀类型路由。其中,扩展团体属性携带的是隧道类型(取值是VXLAN隧道)、本端VTEP的MAC地址;路由下一跳属性携带的是本端的VTEP IP地址。
- 4. Leaf2收到Leaf1发来的BGP EVPN路由后,同时进行如下处理:
 - 检查该路由携带的eERT,如果与本端EVPN实例的IRT相同,则接收该路由, 否则丢弃该路由。
 - 检查该路由携带的eERT,如果与本端L3VPN实例的eIRT相同,则接收该路由,否则丢弃该路由。然后,L3VPN实例获取到该路由携带的IP前缀类型路由,从中提取Host1的主机IP地址、三层VNI,在其路由表中保存Host1的主机IP路由,并将路由的下一跳迭代出接口设置为VXLAN隧道接口,如图11所示。

图 1-27 远端主机 IP 路由



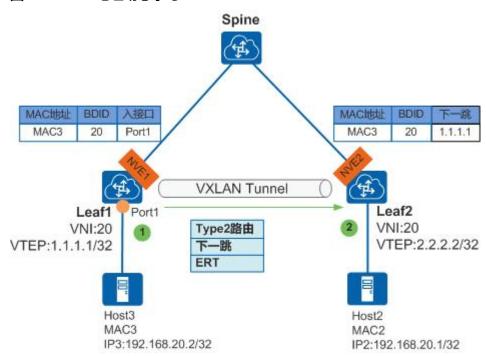
- 在通过EVPN实例或L3VPN实例接收该路由后,Leaf2通过下一跳属性获取 Leaf1的VTEP IP地址,如果该VTEP IP地址是三层路由可达的,则建立一条到 Leaf1的VXLAN隧道。

Leaf1建立到Leaf2的VXLAN隧道的过程与上述相同,这里不再赘述。

MAC 地址动态学习

在VXLAN网络中,为了实现终端租户的互通,支持MAC地址动态学习,不需要网络管理员手工维护,大大减少了维护工作量。在分布式网关场景中,跨子网互通需要进行三层转发,MAC地址学习只在本端主机和网关之间通过动态ARP报文实现,这里不再详述。下面结合图12,详细介绍一下同子网主机互通时,MAC地址动态学习的过程:

图 1-28 MAC 地址动态学习



1. Host3首次与Leaf1通信时,通过动态ARP报文,Leaf1学习到Host3的MAC地址、BDID(二层广播域标识)和报文入接口(即二层子接口对应的物理接口Port1)的对应关系,记录在本地MAC表中。同时,Leaf1根据Host3的ARP表项生成BGPEVPN路由并发送给对等体Leaf2,该路由携带本端EVPN实例的出方向VPN-Target、路由下一跳属性以及BGPEVPN协议新定义的Type2路由即MAC/IP路由。其中,路由下一跳属性携带的是本端VTEMIP地址;MAC/IP路由如图13所示,Host3的MAC地址存放在MACAddressLength和MACAddress字段中,二层VNI存放在MPLSLabel1字段中。

图 1-29 MAC/IP 路由



2. Leaf2收到Leaf1发来的BGP EVPN路由后,首先检查该路由携带的EVPN实例的出方向VPN-Target,如果与本端EVPN实例的入方向VPN-Target相等,则接收该路由,否则丢弃该路由。在接收该路由后,Leaf2获得Host3的MAC地址、BDID和Leaf1上VTEP IP地址(下一跳属性)的对应关系,并保存到本地的MAC表中。

Leaf1学习Host2的主机MAC的过程与上述过程相同,这里不再赘述。

Host3初次与Host2通信时,首先发送目的MAC为全F、目的IP为IP2的ARP请求报文,请求Host2的MAC地址。缺省情况下,Leaf1收到该ARP请求后将在本网段进行广播,

为了减少广播报文,此时可以在Leaf1上使能ARP广播抑制功能。这样当Leaf1收到该ARP请求报文时,先根据目的IP检查本地是否有Host2的MAC地址,如果有则将目的MAC替换为Host2的MAC地址,将ARP请求的广播报文变为单播报文,然后通过VXLAN隧道发给Leaf2。Leaf2收到后转发给Host2,Host2收到该ARP请求后学习到Host3的MAC地址,并以单播形式进行ARP应答。Host3收到ARP应答报文后学习到Host2的MAC地址。至此,Host3和Host2互相学习到对方的MAC地址,后续双方将采用单播通信。

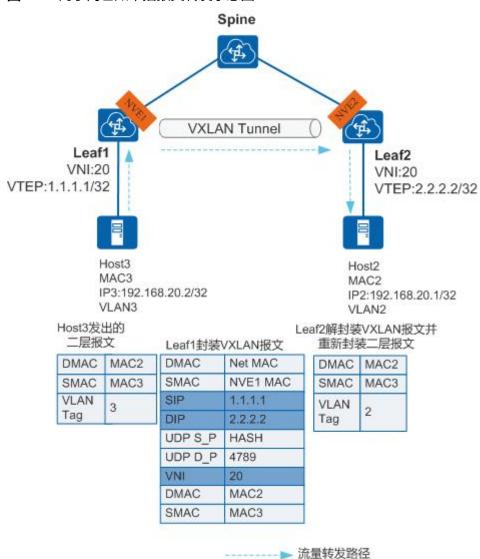
□ 说明

Leaf节点也可以通过数据转发过程学习主机MAC地址,但是依赖于设备从数据报文中学习MAC地址的能力。在BGP EVPN方式建立VXLAN隧道的场景中,Leaf节点之间可以通过BGP EVPN路由动态学习主机MAC地址,不再依赖数据转发过程。此时,可以关闭Leaf节点从数据报文中学习MAC地址的能力,提升转发效率。

同子网已知单播报文转发

同子网已知单播报文转发只在VXLAN二层网关之间进行,三层网关无需感知。报文转发流程如图14所示。

图 1-30 同子网已知单播报文转发示意图



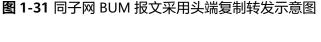
- 1. Leaf1收到来自Host3的报文,根据报文中接入的端口和VLAN信息获取对应的二层 广播域,并在该二层广播域内查找出接口和封装信息。
- 2. Leaf1上VTEP根据查找到的封装信息对数据报文进行VXLAN封装,然后根据查找到的出接口进行报文转发。
- 3. Leaf2上VTEP收到VXLAN报文后,根据UDP目的端口号、源/目的IP地址、VNI判断VXLAN报文的合法有效性。然后依据VNI获取对应的二层广播域,进行VXLAN解封装,获取内层的二层报文。
- 4. Leaf2根据内层二层报文的目的MAC,从本地MAC表找到对应的出接口和封装信息,为报文添加VLAN Tag,转发给对应的主机Host2。

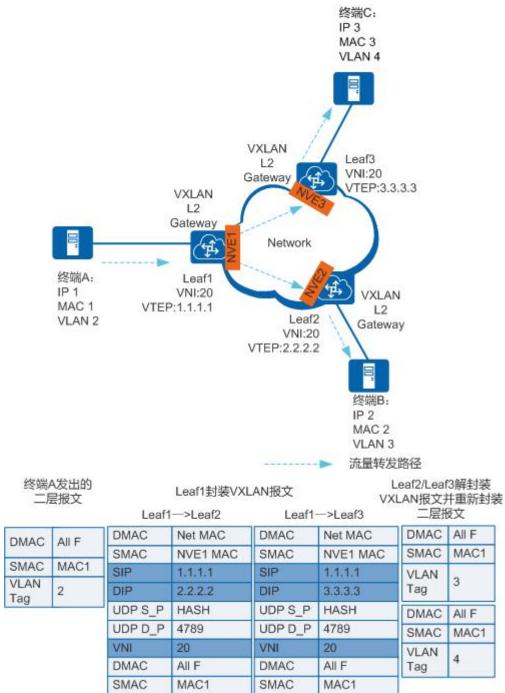
Host2向Host3发送报文的过程类似,这里不再赘述。

同子网 BUM 报文转发

同子网BUM报文转发只在VXLAN二层网关之间进行,三层网关无需感知。同子网BUM 报文转发可以采用头端复制方式。

头端复制是指,当BUM报文进入VXLAN隧道时,接入端VTEP根据头端复制列表进行报文的VXLAN封装,并将报文发送给头端复制列表中的所有出端口VTEP。BUM报文出VXLAN隧道时,出口端VTEP对报文解封装。BUM报文采用头端复制的转发流程如图15所示。





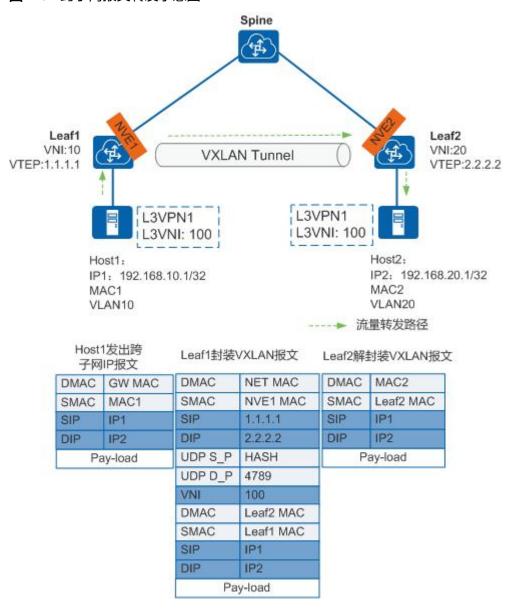
- Leaf1收到来自终端A的报文,根据报文中接入的端口和VLAN信息获取对应的二层 1.
- Leaf1上VTEP根据对应的二层广播域获取对应VNI的头端复制隧道列表,依据获取 的隧道列表进行报文复制,并进行VXLAN封装。然后将封装后的报文从出接口转 发出去。
- Leaf2/Leaf3上VTEP收到VXLAN报文后,根据UDP目的端口号、源/目的IP地址、 VNI判断VXLAN报文的合法有效性。然后依据VNI获取对应的二层广播域,进行 VXLAN解封装,获取内层二层报文。

4. Leaf2/Leaf3检查内层二层报文的目的MAC,发现是BUM MAC,在对应的二层广播域内的非VXLAN隧道侧进行广播处理,即: Leaf2/Leaf3分别从本地MAC表中找到非VXLAN隧道侧的所有出接口和封装信息,为报文添加VLAN Tag,转发给对应的终端B/C。

跨子网报文转发

跨子网报文转发需要通过三层网关实现。在分布式网关场景中,跨子网报文转发的流程如<mark>图16</mark>所示。

图 1-32 跨子网报文转发示意图



- 1. Leaf1收到来自Host1的报文,检测到报文的目的MAC是网关接口MAC,判断该报文需要进行三层转发。
- 2. Leaf1根据报文的入接口找到对应的二层广播域,然后找到绑定该广播域VBDIF接口的L3VPN实例。根据报文的目的IP地址,查找该L3VPN实例下的路由表(如图

17所示),获取该路由对应的三层VNI,以及下一跳地址。再根据下一跳的迭代出接口是VXLAN隧道接口,判断需要进行VXLAN封装:

- 根据VXLAN隧道的目的IP和源IP地址,获取对应的MAC地址,并将内层目的MAC和源MAC替换。
- 将三层VNI封装到报文中。
- 外层封装VXLAN隧道的目的IP和源IP地址,源MAC地址为Leaf1的NVE1接口 MAC地址,目的MAC地址为网络下一跳的MAC地址。

图 1-33 L3VPN 实例下的主机路由

L3VPN1:



- 3. 封装后的报文根据外层MAC和IP信息在IP网络中传输,送达Leaf2。
- 4. Leaf2收到VXLAN报文后进行解封装,检测到报文的目的MAC是自己的MAC地址,判断该报文需要进行三层转发。
- 5. Leaf2根据报文携带的三层VNI找到对应的L3VPN实例,通过查找该L3VPN实例下的路由表(如<mark>图18</mark>所示),获取报文的下一跳是网关接口地址,然后将目的MAC地址替换为Host2的MAC地址,源MAC地址替换为Leaf2的MAC地址,转发给Host2。

图 1-34 L3VPN 实例下的主机路由



Host2向Host1发送报文的过程类似,这里不再赘述。

2 VXLAN 网络规划设计

2.1 硬件集中式 VXLAN

硬件集中式中的"硬件"是指硬件交换机作NVE节点,"集中式"是指VXLAN L3网关集中在一组(或多组)硬件交换机上。硬件集中式VXLAN根据是否有Border Leaf交换机作为网关,硬件集中式VXLAN的架构可以分为Spine/Leaf二层架构和GW/Spine/Leaf三层架构。其中Spine/Leaf二层架构方案流量模型简单,维护简便,但扩展比较困难,适用于中小型网络场景;GW/Spine/Leaf三层架构方案架构灵活,网络扩展能力强,但流量路径和维护相对复杂,适用于大中型网络场景。

2.1.1 GW/Spine/Leaf 三层架构的硬件集中式 VXLAN 设计

如<mark>图1</mark>所示,在GW/Spine/Leaf三层架构的硬件集中式VXLAN中,Border Leaf作为VXLAN三层网关,Spine作为网络东西向流量的汇聚设备,Leaf作为VXLAN二层网关

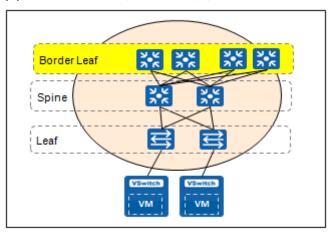
集中式双活 Router-1 Border Leaf Exit-Gateway 12808-2 🔝 LB-1 FW-USG9560-1 🗼 LB-2 FW-USG9560-2 Spine-12804-2 Spine-12804-1 Leaf Leaf-CE6851HI-1 Leaf-CE6851HI-5 Leaf-CE6851HI-3 Leaf-CE6851HI-4 Leaf-CE6851HI-6 Leaf-CE6851HI-2 Server

图 2-1 GW/Spine/Leaf 三层架构的硬件集中式 VXLAN 组网图

整体设计:

- 根据业务类型和规模灵活扩展Border Leaf网关组,如<mark>图</mark>2所示。
- 全局部署ARP广播抑制和端口下部署流量抑制功能,优化广播流量的泛洪。

图 2-2 Border Leaf 网关组扩展图



Border Leaf:

- Border Leaf的部署可以采用集中式双活网关或堆叠。堆叠部署和维护相对简便但是版本升级时业务中断时间长,推荐使用集中式双活网关。
- Router-Border Leaf-Spine之间三层全连接形成ECMP等价路径, Router之间、Border Leaf之间三层互连作为路由备份路径。
- VAS设备旁挂Border Leaf(集中式双活场景使用M-LAG,堆叠场景使用跨框LAG)。
- Border Leaf节点之间规划两条路径:一条路径作为M-LAG的peer-link,另一 条路径作为上行链路的备份。

• Spine:

- Spine作为汇聚网络东西向流量的设备,需要大容量交换能力,推荐选择 CE12800系列交换机,便于接入Leaf的规模扩展。
- Spine以单机形式部署,Spine-Leaf间使用三层路由口直连,整网形成ECMP等价路径,流量负载分担,无阻塞转发,故障快速收敛。

• Leaf:

- 服务器网卡负载分担方式接入时,Leaf节点支持多种组网如堆叠、M-LAG、 盒盒SVF,推荐M-LAG方式因为其接入可靠性更高。服务器网卡主备方式接 入时Leaf采用单机方式部署。
- 当接入Leaf交换机使用M-LAG时,需要部署monitor-link,上行链路与所有下行链路绑定联动,防止当上行链路全部故障时出现流量中断的现象。

2.1.2 Spine/Leaf 二层架构的硬件集中式 VXLAN 设计

如<mark>图</mark>1所示,在Spine/Leaf二层架构的硬件集中式VXLAN中,Spine和GW融合作为VXLAN的三层网关,Leaf作为VXLAN二层网关。

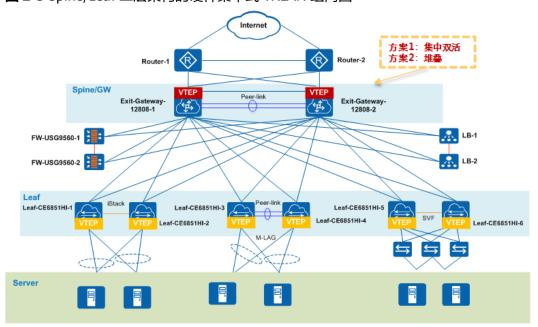


图 2-3 Spine/Leaf 二层架构的硬件集中式 VXLAN 组网图

● 整体设计:

- 根据DC的规模灵活部署Spine和Leaf节点的数量,如<mark>图</mark>2所示。
- Spine与Leaf之间三层互联,整网配置ECMP等价路径,实现流量负载分担, 无阻塞转发以及故障的快速收敛。
- 全局部署ARP广播抑制和端口下部署流量抑制功能,优化广播流量的泛洪。

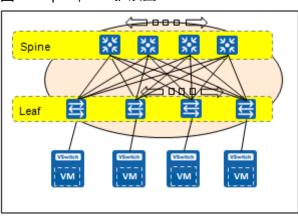


图 2-4 Spine/Leaf 扩展图

Router:

- Router-Spine之间全连接,Spine到Router之间形成ECMP等价路径,Router-Router之间、Spine-Spine之间三层互连作为路由备份链路。
- Router作为网络出口设备,分别与各个外联区互连。

Spine:

- Spine的部署可以采用集中式双活网关或堆叠。堆叠部署和维护相对简便但是版本升级时业务中断时间长,因此推荐使用集中式双活网关。
- Spine节点之间规划两条路径:一条路径作为M-LAG的peer-link,另一条路径 作为上行链路的备份。

- Spine与Leaf直接采用三层路由ECMP(包含堆叠网关场景),VAS设备旁挂Spine(集中式双活场景使用M-LAG,堆叠场景使用跨框LAG)。
- 为了满足未来网络的扩展需求,Spine设备推荐使用CE12800系列交换机。

• Leaf:

- 服务器网卡负载分担方式接入时,Leaf节点支持多种组网如堆叠、M-LAG、 盒盒SVF,推荐M-LAG方式因为其接入可靠性更高。服务器网卡主备方式接 入时Leaf采用单机方式部署。
- 当接入Leaf交换机使用M-LAG时,需要部署monitor-link,上行链路与所有下行链路绑定联动,防止当上行链路全部故障时出现流量中断的现象。

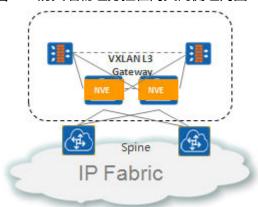
2.1.3 防火墙接入设计

防火墙接入方案可以分为下面三种:

- 防火墙物理旁挂在网关两侧(推荐)
- 防火墙物理连接Service Leaf
- 防火墙直挂在网关和PE之间

如<mark>图</mark>1所示,在防火墙物理旁挂在网关两侧的接入方案中,组网拓扑简单,可以简化配置部署,而且防火墙可以伴随网关组扩展同步扩容,支持灵活部署安全策略,所以推荐部署此方案。

图 2-5 防火墙物理旁挂在网关两侧组网图



2.1.4 负载均衡器接入设计

如<mark>图</mark>1所示,LB旁挂在Border Leaf两侧,适用于LB同时做SNAT和DNAT的方式,GW使用VLANIF或者三层路由口对接LB。LB接入网关设备方式拓扑简单,可以伴随网关组扩展同步扩容。当需要灵活部署负载均衡业务,且LB以三层方式接入,LB的浮动IP与服务器不在同一网段时推荐使用此方案。

图 2-6 LB 物理旁挂在 Border Leaf 两侧组网图



2.1.5 服务器接入设计

服务器接入VXLAN的方案可以分为如下四种方式:

- 1. (推荐)服务器主备接入Leaf单机,如图1所示
- 2. 服务器Eth-Trunk接入Leaf堆叠工作组,如<mark>图</mark>2所示
- 3. (推荐)服务器Eth-Trunk接入Leaf M-LAG工作组,如图3所示
- 4. 服务器Eth-Trunk接入Leaf SVF工作组,如图4所示

图 2-7 服务器主备接入 Leaf 单机组网图

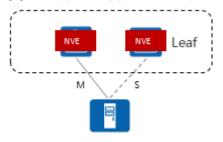


图 2-8 服务器 Eth-Trunk 接入 Leaf 堆叠工作组组网图

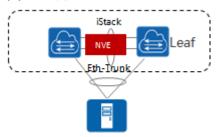


图 2-9 服务器 Eth-Trunk 接入 Leaf M-LAG 工作组组网图

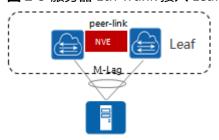
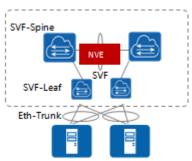


图 2-10 服务器 Eth-Trunk 接入 Leaf SVF 工作组组网图



部署 方式	特点	管理	可靠 性	接入成本
服器备入 Leaf 上eaf 中(荐	Leaf独立部署,服务器双网卡绑定以主备模式接入两台Leaf设备,同一时间只有一个网卡收发报文,带宽利用率低。主备网卡切换时接收流量的VTEP IP变化,依赖于发生切换的服务器发送免费ARP报文重新引流。		個	中
服务 Eth- Trunk 接合 止eaf 堆作 组	两台Leaf利用iStack技术堆叠为一台逻辑设备,共用一个控制面,简化管理。服务器双网卡BOND绑定运行在主备/负载分担模式,带宽利用率高。但逻辑设备升级维护流程较为复杂。	低		中

部署 方式	特点	管理	可靠 性	接入成本
服器 Eth- Trunk Leaf M- LAG 作 推)	两台Leaf设备通过peer-link互联并建立DFS Group,对外表现为一台逻辑设备,但又各自 有独立的控制面,服务器以负载分担方式接入 两台Leaf设备升级维护简单,运行可靠性高。 下行口配置M-LAG特性双归接入服务器,服务 器双网卡运行在主备/负载分担模式。因设备有 独立控制面,故部署配置相对复杂。	。	高	中
服务 Eth- Trunk 接入 Leaf SVF 工作 组	两台高性能具备VXLAN接入功能的Leaf部署iStack堆叠系统,下挂低成本Leaf设备配置SVF,提供低成本1G/10G带宽VXLAN网络接入能力。服务器网卡双归接入两台SVF-Leaf叶子节点,运行在主备/负载分担模式。	中	低	低

2.1.6 逻辑隔离设计

通过VXLAN技术在同一个物理网络上构建虚拟网络然后通过VRF(VPN实例)技术进行物理网络的访问隔离,整个网络通过VRF可以划分三个逻辑平面(业务平面、存储平面、管理平面),默认情况下三个平面不互通,当需要互通时可以配置静态路由交叉,或者通过RT互引实现VRF路由交叉。

- 存储VRF:承载NAS镜像流量、共享文件等存储上的流量。
- 业务VRF:承载业务服务器的业务流量
- 管理VRF:承载虚拟机和宿主机管理流量,如SSH登录,备份,虚拟机的创建、迁 移等

处理这三个逻辑平面外,还需要配置一个管理网络,管理网络可以分为带外管理和带内管理两种:

- 带内管理:使用设备的业务口进行设备管理,此方式的弊端是如果业务网络出现 故障,可能会影响设备登录。
- 带外管理(推荐):采用设备专用的管理口进行设备管理,此方式管控分离。
 以带外管理为例,客户管理VLAN ID为20,需要配置带外管理的网络设备主要有交换机、防火墙、LB。将网络设备的管理口接入客户的管理交换机,根据规划好的管理地址分别配置交换机的管理网口,方便后续远程登录。管理口配置及各网元的登录设置请参考各产品的产品文档。

2.1.7 流量转发模型设计

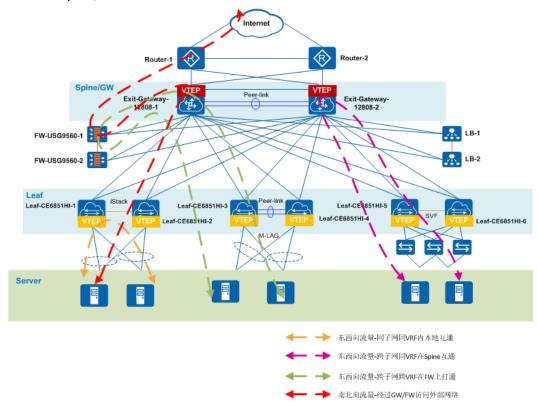
在VXLAN中的流量可以分为两种:

- 南北向流量: VXLAN网络中的服务器和外部网络之间相关访问的流量。
- 东西向流量: VXLAN网络中服务器之间相互访问的流量。

Spine/Leaf 二层架构的硬件集中式 VXLAN 流量转发模型

在Spine/Leaf二层架构的硬件集中式VXLAN中,部分流量转发如图1所示。

图 2-11 Spine/Leaf 二层架构的硬件集中式 VXLAN 部分流量转发示意图



南北向流量,根据业务模型分为如下三种:

- 流量经过FW和LB后,绕回GW转发到Router。
- 流量不经过FW和LB直接访问外部网络。
- 流量经过FW不经过LB, 绕回GW转发至Router。

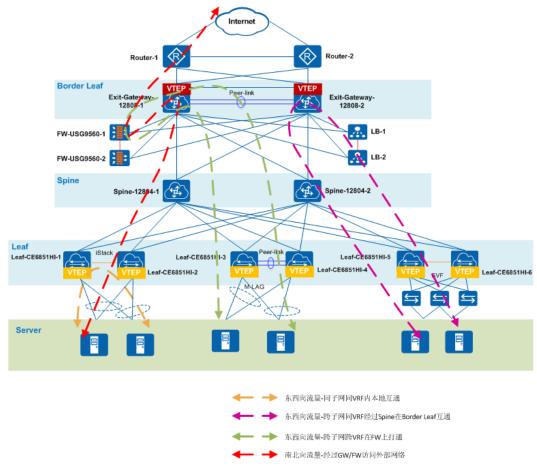
东西向流量,根据业务模型分为如下三种:

- 同子网同VRF流量:直接在Leaf上本地转发,或通过在Leaf上封装绕过Spine转发 到另外的Leaf上然后解封装。
- 跨子网同VRF流量:流量在Leaf设备上封装VXLAN后,在Spine设备上解封装后查 找直连路由三层转发流量。
- 跨子网跨VRF流量:流量在Leaf设备上封装VXLAN后,在Spine设备上解封装后查 找静态流量转发至FW,由FW统一安全控制跨VRF的流量,也可以在Spine本地上 直接通过静态路由或RT互引方式完成路由交叉转发流量。

GW/Spine/Leaf 三层架构的硬件集中式 VXLAN 流量转发模型

在GW/Spine/Leaf三层架构的硬件集中式VXLAN中,部分流量转发如图2所示。

图 2-12 GW/Spine/Leaf 三层架构的硬件集中式 VXLAN 部分流量转发示意图



南北向流量,根据业务模型分为如下三种:

- 流量经过FW和LB后,绕回GW转发到Router。
- 流量不经过FW和LB直接访问外部网络。
- 流量经过FW不经过LB,绕回GW转发至Router。

东西向流量,根据业务模型分为如下三种:

- 同子网同VRF流量: 直接在Leaf上本地转发,或通过在Leaf上封装绕过Spine转发到另外的Leaf上然后解封装。
- 跨子网同VRF流量: 流量在Leaf设备上封装VXLAN后,经过Spine在Border Leaf 上解封装后查找直连路由三层转发流量。
- 跨子网跨VRF流量:流量在Leaf设备上封装VXLAN后,经过Spine在Border Leaf上解封装后查找静态路由转发至FW,由FW统一安全控制跨VRF的流量,也可以在Border Leaf本地直接通过静态路由或MCE方式完成路由交叉转发流量。

2.1.8 路由协议设计

VXLAN网络中,需要同时打通Underlay网络和Overlay网络的路由。在Underlay网络中可以采用BGP、OSPF或者IS-IS路由协议并使用Loopback地址作为VTEP IP。Overlay

网络中均使用EVPN且配置Spine设备作为路由反射器。各协议的优缺点和适用场景如表1所示。

表 2-1 路由选择

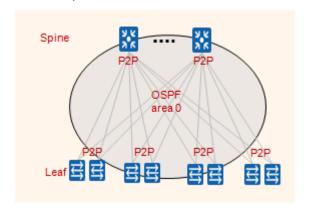
Underlay路由	Overl ay路 由	优缺点	推荐使用场景
OSPF	IBGP- EVPN	优点: OSPF路由 协议部署简单, 快速收敛 缺点: OSPF路由 域规模受限	中小型网络(小于100个路由邻居)
IS-IS	IBGP- EVPN	优点: IS-IS路由 协议部署简单, 占用网络资源较 小,规模易扩 充,收敛速度快 缺点: 客户群接 受度受限	中大型网络(小于500个路由邻居)
EBGP	IBGP- EVPN	优点:每个分区 路由域独立,故 障域可控,路由 控制最为灵活, 可灵活扩展规模 缺点:配置复杂	中大型网络(小于1000个路由邻居)

Underlay 路由

OSPF路由

如<mark>图1</mark>和<mark>图2</mark>所示,在Spine/Leaf二层架构的硬件集中式VXLAN架构和GW/Spine/Leaf三层架构的硬件集中式VXLAN的架构中,所有设备统一规划在OSPF Area0,使用三层路由口地址建立OSPF邻居,打通Underlay路由,network类型建议为P2P。

图 2-13 Spine/Leaf 二层架构的硬件集中式 VXLAN 的 Underlay 采用 OSPF 路由



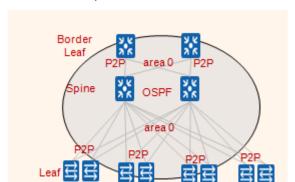


图 2-14 GW/Spine/Leaf 三层架构的硬件集中式 VXLAN 的 Underlay 采用 OSPF 路由

EBGP路由

如<mark>图3</mark>所示,在Spine/Leaf二层架构的硬件集中式VXLAN中,两个Spine节点规划在一个AS中,每组Leaf(M-LAG组or堆叠组)规划在一个AS中,使用三层路由口地址建立EBGP邻居,network发布本地VTEP IP地址,打通Underlay路由。

如<mark>图4</mark>所示,两个Border Leaf节点规划在一个AS中,两个Spine节点规划在一个AS中,每组Leaf(M-LAG组or堆叠组)规划在一个AS中,使用三层路由口地址建立EBGP邻居,network发布本地VTEP IP地址,打通Underlay路由。

图 2-15 Spine/Leaf 二层架构的硬件集中式 VXLAN 的 Underlay 采用 EBGP 路由

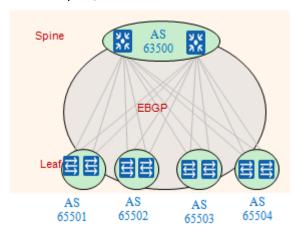
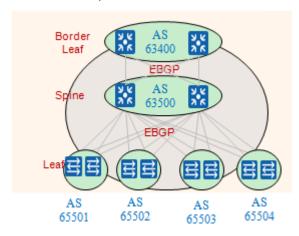


图 2-16 GW/Spine/Leaf 三层架构的硬件集中式 VXLAN 的 Underlay 采用 EBGP 路由



IS-IS路由

如图5和图6所示,在Spine/Leaf二层架构的硬件集中式VXLAN架构和GW/Spine/Leaf三层架构的硬件集中式VXLAN中,使用三层路由口地址建立ISIS邻居,所有设备统一规划在Level-2,当本fabric规模扩大时,可以在Level-1层面进行拓展,如果后续fabric规模不大,但多地市间多fabric建设联合超大规模网络时,则fabric规划为Leavl-1,联合多fabric的核心区规划为Level-2。

图 2-17 Spine/Leaf 二层架构的硬件集中式 VXLAN 的 Underlay 采用 IS-IS 路由

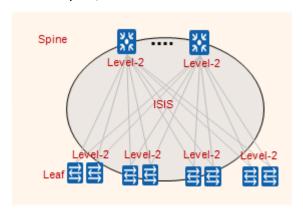
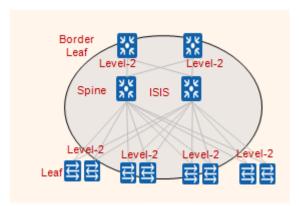


图 2-18 GW/Spine/Leaf 三层架构的硬件集中式 VXLAN 的 Underlay 采用 IS-IS 路由



Overlay 路由

Overlay路由采用IBGP EVPN路由协议。

如图7和图8所示,在Spine/Leaf二层架构的硬件集中式VXLAN架构和GW/Spine/Leaf三层架构的硬件集中式VXLAN中,所有的Leaf(包含Border Leaf)和Spine设备规划独立的AS号,且与Underlay路由AS号不能重叠。Spine作为BGP EVPN的RR,Spine与Leaf设备使用独立的Loopback地址建立EVPN对等体。

当采用M-LAG接入时,DFS的loopback地址建议与建立EVPN对等体的loopback地址共用,但不可以与VTEP IP地址共用。

GW与出口路由器之间推荐采用静态路由,如果两者间需要传递很多的明细路由可以使用动态路由协议,推荐如下: Underlay路由协议为OSPF时,推荐选择OSPF,规划独立的OSPF进程,BGP协议作为次选方案,简化配置;当Underlay路由协议为EBGP时,推荐选择OSPF,次选与EBGP共用AS号的BGP,简化配置复杂度。

图 2-19 Spine/Leaf 二层架构的硬件集中式 VXLAN 的 Underlay 采用 IBGP EVPN 路由

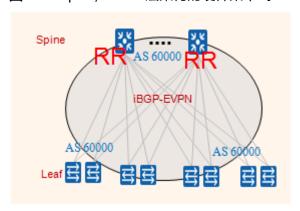
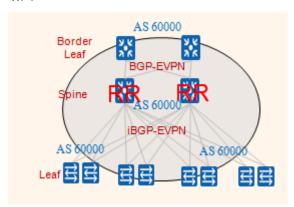


图 2-20 GW/Spine/Leaf 三层架构的硬件集中式 VXLAN 的 Underlay 采用 IBGP EVPN 路由



2.1.9 可靠性设计

配置网络冗余备份

在实际网络中,各种因素造成的故障难以避免,因此能够让网络从故障中快速恢复的技术就显得非常重要。通过设备冗余备份提高网络的可靠性是重要手段。

如<mark>图</mark>1所示,以GW/Spine/Leaf三层架构的硬件集中式VXLAN为例,通过设备冗余备份解决网络故障。

- 1. 服务器链路故障:服务器双归接入,网卡负载分担/主备,当服务器一条链路故障时,业务倒换到冗余/备份链路。
- 2. Leaf/Border Leaf设备故障: Leaf/Border Leaf配置M-LAG(推荐)/堆叠/SVF, 当一台Leaf/Border Leaf故障时,业务倒换到另外一台Leaf/Border Leaf上。
- 3. Leaf上行链路故障: Leaf和Spine间通过多条链路实现ECMP,当一条上行链路故障后,业务倒换到冗余链路。
- 4. Spine设备故障:一台Spine故障后,流量从另外一台Spine设备转发。
- 5. LB和FW故障: LB和FW配置主备备份,当LB和FW故障后,流量切换到备份LB和FW设备。
- 6. 堆叠链路/peer-link故障: 当堆叠链路或者peer-link故障时,通过双主检测,触发状态为备的设备上除管理网口、peer-link接口和堆叠口以外的接口处于Error-Down状态,避免网络出现双主,提高可靠性。

Router-1 Border Leaf Exit-Gateway12808-1 FW-USG9560-2 Spine-12804-1 Leaf-CE6851HI-1 Leaf-CE6851HI-1 Leaf-CE6851HI-1 Leaf-CE6851HI-1 Spine-12804-2 Leaf-CE6851HI-1 Leaf-CE6851HI-1 Spine-12804-2 Leaf-CE6851HI-1 Spine-12804-2 Server

图 2-21 网络冗余备份解决网络故障组网图

配置网络环路防护

当网络中出现环路时,会消耗大量的网络资源导致业务中断,网络中需要部署防护措施提高可靠性。

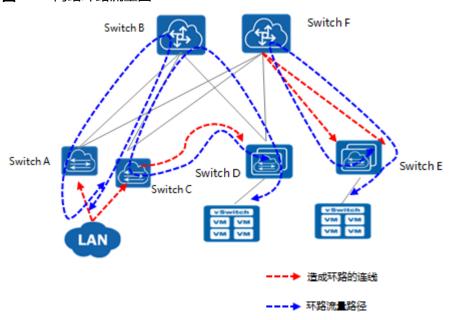
如<mark>图</mark>2所示,当网络中下挂具有二层转发功能的设备(含服务器的vSwitch)时,如果 人为配置错误或者误接线,就会与下挂设备组成环路,出现环路流量。

推荐的防护措施如下:

配置基于端口的流量抑制功能,推荐配置广播报文抑制为1000pps,组播报文抑制为200pps,未知单播报文抑制为1%,ARP报文端口限速200pps。

● 设备默认配置MAC偏移检测功能,基于BD检测到MAC漂移次数周期内达到一定数值时会产生告警。

图 2-22 网络环路流量图



2.1.10 约束限制

VXLAN的约束很多,具体可以参见"CloudEngine 系列交换机产品文档-配置-VXLAN配置指南-配置注意事项"。

2.2 硬件分布式 VXLAN

硬件分布式VXLAN是VXLAN组网的一种类型,"分布式"是指VXLAN L3网关分散部署在所有Leaf节点交换机上,这样可使接入网络计算节点的VXLAN L3网关部署在就近的Leaf节点上;"硬件"是指硬件交换机作NVE节点。硬件分布式VXLAN根据是否有Border Leaf交换机作为网关,VXLAN的架构可以分为Spine/Leaf二层架构和GW/Spine/Leaf三层架构。其中Spine/Leaf二层架构方案流量模型简单,维护简便,但扩展比较困难,适用于中小型网络场景;GW/Spine/Leaf三层架构方案架构灵活,网络扩展能力强,但流量路径和维护相对复杂,适用于大中型网络场景。

2.2.1 GW/Spine/Leaf 三层架构的硬件分布式 VXLAN 设计

如<mark>图</mark>1所示,在GW/Spine/Leaf三层架构的硬件分布式VXLAN中,Border Leaf作为VXLAN出口网关-Exit Gateway,也称南北网关。Leaf作为VXLAN分布式网关,也称为东西网关,Spine作为网络东西向流量的汇聚设备。

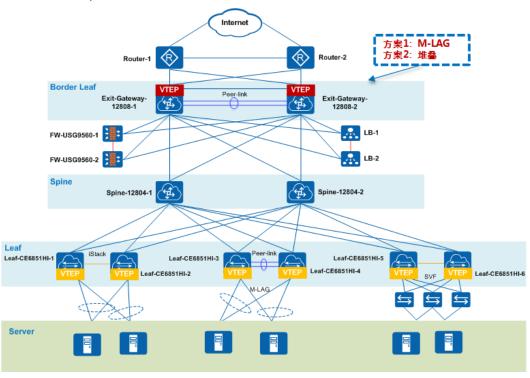


图 2-23 GW/Spine/Leaf 三层架构的硬件分布式 VXLAN 组网图

● 整体设计:

- 根据业务类型和规模灵活扩展Border Leaf网关组,如图2所示。
- 全局部署ARP广播抑制和端口下部署流量抑制功能,优化广播流量的泛洪。 也可以配置ARP代理采用网关代答功能作为次选方案,可以将流量引向对应 的网关,在网关上实现对二层流量的监控。

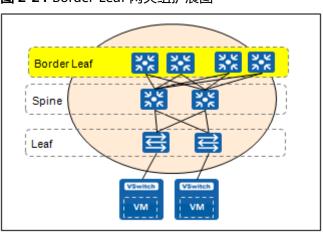


图 2-24 Border Leaf 网关组扩展图

Border Leaf:

- Border Leaf的部署可以采用M-LAG或堆叠。堆叠部署和维护相对简便但是版本升级时业务中断时间长,推荐使用M-LAG方式。
- Router-Border Leaf-Spine间使用三层路由口直连,整网形成ECMP等价路径,实现流量负载分担,无阻塞转发以及故障快速收敛。

Spine:

- Spine以单机形式部署作为汇聚网络东西向流量的设备,需要大容量交换能力,推荐选择CE12800系列交换机,便于接入Leaf的规模扩展。
- Spine作为BGP EVPN的路由反射器。

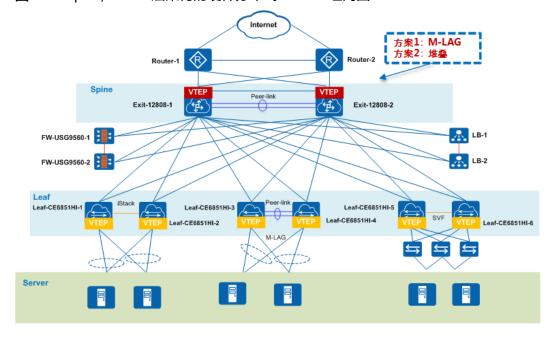
• Leaf:

- 服务器网卡负载分担方式接入时,Leaf节点支持多种组网如堆叠、M-LAG、 盒盒SVF,推荐M-LAG方式因为其接入可靠性更高。服务器网卡主备方式接 入时Leaf采用单机方式部署。
- 当Server Leaf交换机使用M-LAG时,需要部署monitor-link,上行链路与所有下行链路绑定联动,防止当上行链路故障时出现流量中断的现象。

2.2.2 Spine/Leaf 二层架构的硬件分布式 VXLAN 设计

如<mark>图1</mark>所示,在Spine/Leaf二层架构的硬件分布式VXLAN中,GW和Spine融合作为VXLAN网络出口,Leaf作为VXLAN分布式网关。

图 2-25 Spine/Leaf 二层架构的硬件分布式 VXLAN 组网图



该架构的具体部署如下:

● 整体设计:

- 根据DC的规模灵活部署Spine和Leaf节点的数量,如<mark>图</mark>2所示。
- Spine与Leaf之间三层互联,整网配置ECMP等价路径,实现流量负载分担, 无阻塞转发以及故障的快速收敛。
- 全局部署ARP广播抑制和端口下部署流量抑制功能,优化广播流量的泛洪。 也可以配置ARP代理采用网关代答功能作为次选方案,可以将流量引向对应 的网关,在网关上实现对二层流量的监控。

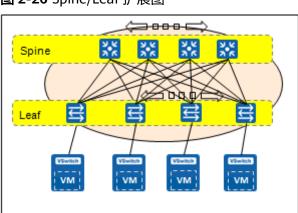


图 2-26 Spine/Leaf 扩展图

• Spine:

- Spine的部署可以采用M-LAG或堆叠。堆叠部署和维护相对简便但是版本升级时业务中断时间长,因此推荐使用M-LAG方式。
- Spine作为BGP EVPN的路由反射器。
- 为了满足未来网络的扩展需求,Spine设备推荐使用CE12800系列交换机。

• Leaf:

- 服务器网卡负载分担方式接入时,Leaf节点支持多种组网如堆叠、M-LAG、 盒盒SVF,推荐M-LAG方式因为其接入可靠性更高。服务器网卡主备方式接 入时Leaf采用单机方式部署。
- 当Server Leaf交换机使用M-LAG时,需要部署monitor-link,上行链路与所有下行链路绑定联动,防止当上行链路故障时出现流量中断的现象。

Router:

- Router-Spine之间全连接,Spine到Router之间形成ECMP等价路径,Router-Router之间、Spine-Spine之间三层互连作为路由备份链路。
- Router作为网络出口设备,分别与各个外联区互连。

2.2.3 防火墙接入设计

防火墙接入方案可以分为下面三种:

- (推荐)防火墙物理旁挂在Border/Server Leaf两侧
- 防火墙直挂在Border Leaf和PE之间

如<mark>图</mark>1所示,防火墙物理旁挂在Border/Server Leaf两侧(Border Leaf与Spine融合场景时,防火墙物理旁挂在Spine/Server Leaf两侧)方式拓扑简单,可以简化配置部署,防火墙可以伴随网关组扩展同步扩容,支持灵活部署安全策略。

在防火墙接入的两种方案中推荐部署此方案。

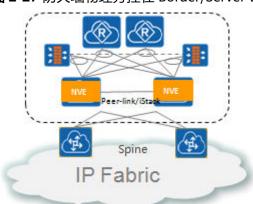


图 2-27 防火墙物理旁挂在 Border/Server Leaf 两侧组网图

2.2.4 负载均衡器接入设计

如<mark>图1</mark>所示,LB旁挂在Border Leaf两侧(Border Leaf与Spine融合场景时,LB旁挂在Spine两侧),适用于LB同时做SNAT和DNAT的方式,GW使用VLANIF或者三层路由口对接LB。LB接入网关设备方式拓扑简单,可以伴随网关组扩展同步扩容。当需要灵活部署负载均衡业务, 且LB以三层方式接入,LB的浮动IP与服务器不在同一网段时推荐使用此方案。

图 2-28 LB 物理旁挂在 Border Leaf 两侧组网图



2.2.5 服务器接入设计

服务器接入VXLAN的方案可以分为如下四种方式:

- 1. (推荐)服务器主备接入Leaf单机,如<mark>图</mark>1所示
- 2. 服务器Eth-Trunk接入Leaf堆叠工作组,如图2所示
- 3. (推荐)服务器Eth-Trunk接入Leaf M-LAG工作组,如图3所示
- 4. 服务器Eth-Trunk接入Leaf SVF工作组,如图4所示

图 2-29 服务器主备接入 Leaf 单机组网图

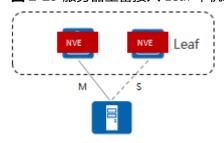


图 2-30 服务器 Eth-Trunk 接入 Leaf 堆叠工作组组网图

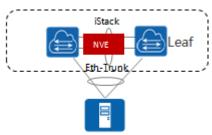


图 2-31 服务器 Eth-Trunk 接入 Leaf M-LAG 工作组组网图

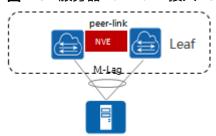
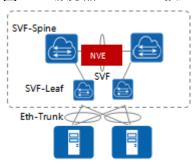


图 2-32 服务器 Eth-Trunk 接入 Leaf SVF 工作组组网图



部署 方式	特点	管理	可靠性	接入成本
服器备入 Leaf 上单(荐	Leaf独立部署,服务器双网卡BOND绑定以主备模式接入两台Leaf设备,同一时间只有一个网卡收发报文,带宽利用率低。主备网卡切换时接收流量的VTEP IP变化,依赖于发生切换的服务器发送免费ARP报文重新引流。	中	讵	中
服器 Eth- Trunk Leaf 生工组	两台Leaf利用iStack技术堆叠为一台逻辑设备,共用一个控制面,简化管理。服务器双网卡BOND绑定运行在主备/负载分担模式,带宽利用率高。但逻辑设备升级维护流程较为复杂。	低	中	中
服器 Eth- Tru入 Leaf LA工组(荐 LA工组(荐	两台Leaf设备通过peer-link互联并建立DFS Group,对外表现为一台逻辑设备,但又各自 有独立的控制面,服务器以负载分担方式接入 两台Leaf设备升级维护简单,运行可靠性高。 下行口配置M-LAG特性双归接入服务器,服务 器双网卡运行在主备/负载分担模式。因设备有 独立控制面,故部署配置相对复杂。	個	追	中
服务 器 Eth- Trunk 接入 Leaf SVF 工作 组	两台高性能具备VXLAN接入功能的Leaf部署iStack堆叠系统,下挂低成本Leaf设备配置SVF,提供低成本1G/10G带宽VXLAN网络接入能力。服务器网卡双归接入两台SVF-Leaf叶子节点,运行在主备/负载分担模式。	中	低	低

2.2.6 逻辑隔离设计

通过VXLAN技术在同一个物理网络上构建虚拟网络然后通过VRF(VPN实例)技术进行物理网络的访问隔离,整个网络通过VRF可以划分三个逻辑平面(业务平面、存储平面、管理平面),默认情况下三个平面不互通,当需要互通时可以配置静态路由交叉,或者通过RT互引实现VRF路由交叉。

- 存储VRF:承载NAS镜像流量、共享文件等存储上的流量。
- 业务VRF:承载业务服务器的业务流量
- 管理VRF:承载虚拟机和宿主机管理流量,如SSH登录,备份,虚拟机的创建、迁 移等

处理这三个逻辑平面外,还需要配置一个管理网络,管理网络可以分为带外管理和带内管理两种:

- 带内管理:使用设备的业务口进行设备管理,此方式的弊端是如果业务网络出现 故障,可能会影响设备登录。
- 带外管理(推荐):采用设备专用的管理口进行设备管理,此方式管控分离。以带外管理为例,客户管理VLAN ID为20,需要配置带外管理的网络设备主要有交换机、防火墙、LB。将网络设备的管理口接入客户的管理交换机,根据规划好的管理地址分别配置交换机的管理网口,方便后续远程登录。管理口配置及各网元的登录设置请参考各产品的产品文档。

2.2.7 流量转发模型设计

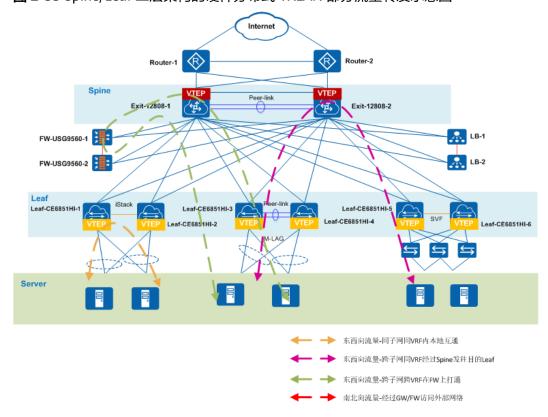
在VXLAN中的流量可以分为两种:

- 南北向流量: VXLAN网络中的服务器和外部网络之间相关访问的流量。
- 东西向流量: VXLAN网络中服务器之间相互访问的流量。

Spine/Leaf 二层架构的硬件分布式 VXLAN 流量转发模型

在Spine/Leaf二层架构的硬件分布式VXLAN中,部分流量转发如图1所示。

图 2-33 Spine/Leaf 二层架构的硬件分布式 VXLAN 部分流量转发示意图



南北向流量,根据业务模型分为如下三种:

- 流量经过FW和LB后,绕回GW转发到Router。
- 流量不经过FW和LB直接访问外部网络。
- 流量经过FW不经过LB,绕回GW转发至Router。

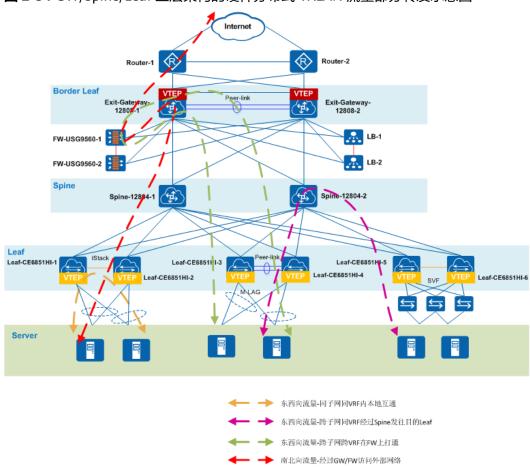
东西向流量,根据业务模型分为如下三种:

- 同子网同VRF流量: 直接在Leaf上本地转发,或通过在Leaf上封装绕过Spine转发到另外的Leaf上然后解封装。
- 跨子网同VRF流量: 在Leaf设备上查找32位主机路由表封装VXLAN后,发送到目的主机所在的Leaf,Spine三层路由转发流量。
- 跨子网跨VRF流量: 当需要FW安全控制时按照a的流量转发,当不需要FW安全控制时按照b转发。
 - a. 流量在Leaf设备上封装VXLAN后,查询默认路由到达Border Leaf上解封装,流量引致FW,由FW统一安全策略控制跨VRF的流量。
 - b. 跨VRF的流量在各Leaf上完成路由交叉(VPN的vpn-target在EVPN进程中互引),查找交叉过来的32位主机路由直接转发至目的Leaf。

GW/Spine/Leaf 三层架构的硬件分布式 VXLAN 流量转发模型

在GW/Spine/Leaf三层架构的硬件分布式VXLAN中,部分流量转发如图2所示。

图 2-34 GW/Spine/Leaf 三层架构的硬件分布式 VXLAN 流量部分转发示意图



南北向流量,根据业务模型分为如下三种:

- 流量经过FW和LB后,绕回GW转发到Router。
- 流量不经过FW和LB直接访问外部网络。
- 流量经过FW不经过LB,绕回GW转发至Router。

东西向流量,根据业务模型分为如下三种:

- 同子网同VRF流量: 直接在Leaf上本地转发,或通过在Leaf上封装绕过Spine转发到另外的Leaf上然后解封装。
- 跨子网同VRF流量: 在Leaf设备上查找32位主机路由表封装VXLAN后,发送到目的主机所在的Leaf,Spine三层路由转发流量。
- 跨子网跨VRF流量: 当需要FW安全控制时按照a的流量转发,当不需要FW安全控制时按照b转发。
 - a. 流量在Leaf设备上封装VXLAN后,查询默认路由到达Border Leaf上解封装,流量引致FW,由FW统一安全策略控制跨VRF的流量。
 - b. 跨VRF的流量在各Leaf上完成路由交叉(VPN的vpn-target在EVPN进程中互引),查找交叉过来的32位主机路由直接转发至目的Leaf。

2.2.8 路由协议设计

VXLAN网络中,需要同时打通Underlay网络和Overlay网络的路由。在Underlay网络中可以采用BGP、OSPF或者IS-IS路由协议并使用Loopback地址作为VTEP IP。Overlay网络中均使用EVPN且配置Spine设备作为路由反射器。各协议的优缺点和适用场景如表1所示。

表 2-2 路由选择

Underlay路由	Overl ay路 由	优缺点	推荐使用场景
OSPF	IBGP- EVPN	优点: OSPF路由 协议部署简单, 快速收敛 缺点: OSPF路由 域规模受限	中小型网络(小于100个路由邻居)
IS-IS	IBGP- EVPN	优点: IS-IS路由 协议部署简单, 占用网络资源较 小,规模易扩 充,收敛速度快 缺点: 客户群接 受度受限	中大型网络(小于500个路由邻居)

Underlay路由	Overl ay路 由	优缺点	推荐使用场景
EBGP	IBGP- EVPN	优点:每个分区 路由域独立,故 障域可控,路由 控制最为灵活, 可灵活扩展规模 缺点:配置复杂	中大型网络(小于1000个路由邻居)

Underlay 路由

OSPF路由

如<mark>图1</mark>和<mark>图2</mark>所示,在Spine/Leaf二层架构的硬件分布式VXLAN架构和GW/Spine/Leaf三层架构的硬件分布式VXLAN的架构中,所有设备统一规划在OSPF Area0,使用三层路由口地址建立OSPF邻居,打通Underlay路由,network类型建议为P2P。

图 2-35 Spine/Leaf 二层架构的硬件分布式 VXLAN 的 Underlay 采用 OSPF 路由

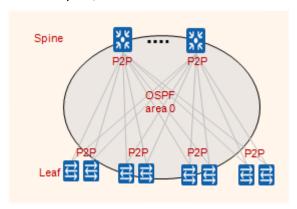
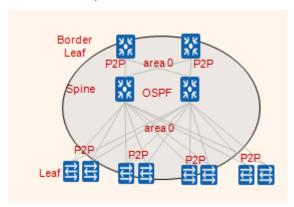


图 2-36 GW/Spine/Leaf 三层架构的硬件分布式 VXLAN 的 Underlay 采用 OSPF 路由



EBGP路由

如图3所示,在Spine/Leaf二层架构的硬件分布式VXLAN中,两个Spine节点规划在一个AS中,每组Leaf(M-LAG组or堆叠组)规划在一个AS中,使用三层路由口地址建立EBGP邻居,network发布本地VTEP IP地址,打通Underlay路由。

如<mark>图4</mark>所示,两个Border Leaf节点规划在一个AS中,两个Spine节点规划在一个AS中,每组Leaf(M-LAG组or堆叠组)规划在一个AS中,使用三层路由口地址建立EBGP邻居,network发布本地VTEP IP地址,打通Underlay路由。

图 2-37 Spine/Leaf 二层架构的硬件分布式 VXLAN 的 Underlay 采用 EBGP 路由

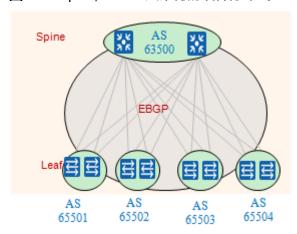
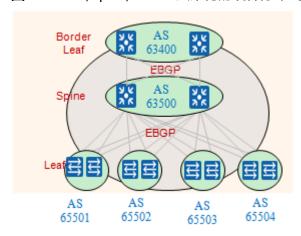


图 2-38 GW/Spine/Leaf 三层架构的硬件分布式 VXLAN 的 Underlay 采用 EBGP 路由



IS-IS路由

如图5和图6所示,在Spine/Leaf二层架构的硬件分布式VXLAN架构和GW/Spine/Leaf三层架构的硬件分布式VXLAN中,使用三层路由口地址建立ISIS邻居,所有设备统一规划在Level-2,当本fabric规模扩大时,可以在Level-1层面进行拓展,如果后续fabric规模不大,但多地市间多fabric建设联合超大规模网络时,则fabric规划为Leavl-1,联合多fabric的核心区规划为Level-2。

图 2-39 Spine/Leaf 二层架构的硬件分布式 VXLAN 的 Underlay 采用 IS-IS 路由

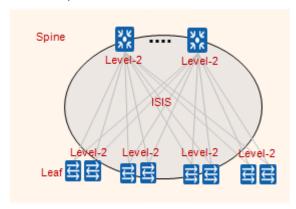
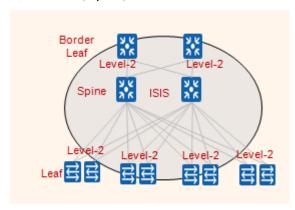


图 2-40 GW/Spine/Leaf 三层架构的硬件分布式 VXLAN 的 Underlay 采用 IS-IS 路由



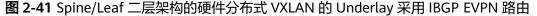
Overlay 路由

Overlay路由采用IBGP EVPN路由协议。

如<mark>图7和图8</mark>所示,在Spine/Leaf二层架构的硬件分布式VXLAN架构和GW/Spine/Leaf三层架构的硬件分布式VXLAN中,所有的Leaf(包含Border Leaf)和Spine设备规划独立的AS号,且与Underlay路由AS号不能重叠。Spine作为BGP EVPN的RR,Spine与Leaf设备使用独立的Loopback地址建立EVPN对等体。

当采用M-LAG接入时,DFS的loopback地址建议与建立EVPN对等体的loopback地址共用,但不可以与VTEP IP地址共用。

GW与出口路由器之间推荐采用静态路由,如果两者间需要传递很多的明细路由可以使用动态路由协议,推荐如下: Underlay路由协议为OSPF时,推荐选择OSPF,规划独立的OSPF进程,BGP协议作为次选方案,简化配置; 当Underlay路由协议为EBGP时,推荐选择OSPF,次选与EBGP共用AS号的BGP,简化配置复杂度。



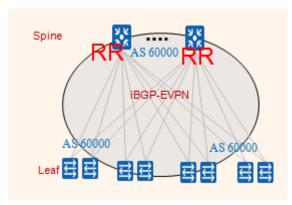
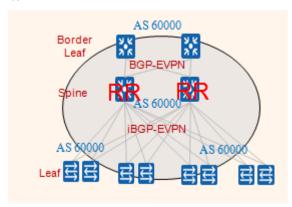


图 2-42 GW/Spine/Leaf 三层架构的硬件分布式 VXLAN 的 Underlay 采用 IBGP EVPN 路由



2.2.9 可靠性设计

配置网络冗余备份

在实际网络中,各种因素造成的故障难以避免,因此能够让网络从故障中快速恢复的技术就显得非常重要。通过设备冗余备份提高网络的可靠性是重要手段。

如<mark>图</mark>1所示,以GW/Spine/Leaf三层架构的硬件分布式VXLAN为例,通过设备冗余备份解决网络故障。

- 1. 服务器链路故障:服务器双归接入,网卡负载分担/主备,当服务器一条链路故障时,业务倒换到冗余/备份链路。
- 2. Leaf/Border Leaf设备故障: Leaf/Border Leaf配置M-LAG(推荐)/堆叠/SVF,当一台Leaf/Border Leaf故障时,业务倒换到另外一台Leaf/Border Leaf上。
- 3. Leaf上行链路故障: Leaf和Spine间通过多条链路实现ECMP,当一条上行链路故障后,业务倒换到冗余链路。
- 4. Spine设备故障:一台Spine故障后,流量从另外一台Spine设备转发。
- 5. LB和FW故障: LB和FW配置主备备份,当LB和FW故障后,流量切换到备份LB和FW设备。
- 6. 堆叠链路/peer-link故障:当堆叠链路或者peer-link故障时,通过双主检测,触发状态为备的设备上除管理网口、peer-link接口和堆叠口以外的接口处于Error-Down状态,避免网络出现双主,提高可靠性。

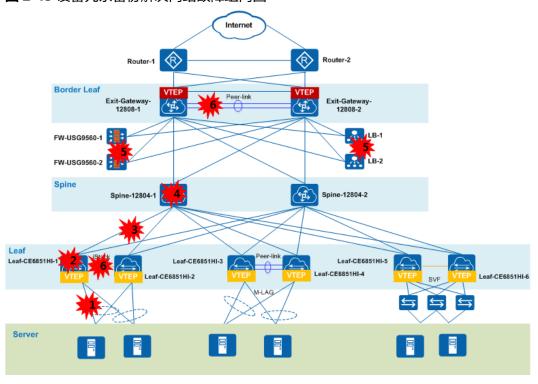


图 2-43 设备冗余备份解决网络故障组网图

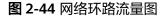
配置网络环路防护

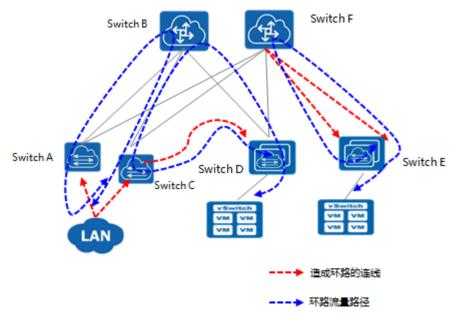
当网络中出现环路时,会消耗大量的链路带宽资源导致业务中断,网络中需要提供破 环技术提高可靠性。

如<mark>图</mark>2所示,当网络中下挂具有二层转发功能的(含服务器的vSwitch)时,如果人为配置错误或者误接线,就会与下挂设备组成环路,出现环路流量。

推荐的防护措施如下:

- 配置基于端口的流量抑制功能,推荐配置广播报文抑制为1000pps,组播报文抑制为200pps,未知单播报文抑制为1%,ARP报文端口限速200pps。
- 设备默认配置MAC偏移检测功能,基于BD检测到MAC漂移次数周期内达到一定数值时会产生告警。





2.2.10 约束限制

VXLAN的约束很多,具体可以参见"CloudEngine 系列交换机产品文档-配置-VXLAN配置指南-配置注意事项"。

3 vxlan 部署建议

3.1 硬件集中式 VXLAN

硬件集中式中的"硬件"是指硬件交换机作NVE节点,"集中式"是指VXLAN L3网关集中在一组(或多组)硬件交换机上。硬件集中式VXLAN根据是否有Border Leaf交换机作为网关,硬件集中式VXLAN的架构可以分为Spine/Leaf二层架构和GW/Spine/Leaf三层架构。其中Spine/Leaf二层架构方案流量模型简单,维护简便,但扩展比较困难,适用于中小型网络场景;GW/Spine/Leaf三层架构方案架构灵活,网络扩展能力强,但流量路径和维护相对复杂,适用于大中型网络场景。

3.1.1 GW/Spine/Leaf 三层架构的硬件集中式 VXLAN

组网需求

GW/Spine/Leaf三层架构的硬件集中式VXLAN组网图如图1所示,

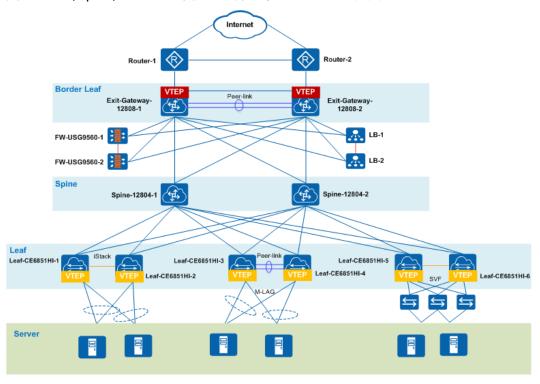


图 3-1 GW/Spine/Leaf 三层架构的硬件集中式 VXLAN 组网图

- 服务器层:通过二层子接口方式接入VXLAN网络。
- Leaf:采用堆叠方式、M-LAG或SVF方式连接服务器。Leaf节点和Spine节点通过 三层互联。Leaf的堆叠工作组、M-LAG工作组或者SVF工作组作为VTEP实现服务 器流量接入VXLAN网络。
- Spine: 分别与Leaf节点和GW互联,运行路由协议保证Underlay三层可达。Spine 节点不作为VTEP。
- Border Leaf: 两台Border Leaf之间组成M-LAG,即双活网关。Border Leaf与Spine之间三层互联。Border Leaf对外连接外部路由器Router-1和Router-2。
- FW: 两台FW分别旁挂在GW, FW之间配置为主备镜像模式工作。
- LB:由厂商部署。

端口连线规划

Leaf-CE6851HI-1和Leaf-CE6851HI-2部署堆叠场景,端口规划参见下表。

规划描述	本端设备名称	物理端口	对端设备名称	物理端口
规划Leaf间互连端口	Leaf-	40GE1/0/1	Leaf-	40GE1/0/1
作为堆叠口,堆叠链	CE6851HI-1	~2	CE6851HI-2	~2
路至少2条	Leaf-	40GE1/0/1	Leaf-	40GE1/0/1
	CE6851HI-2	~2	CE6851HI-1	~2
规划Leaf与Spine互连 端口	Leaf- CE6851HI-1	40GE1/0/3	Spine- CE12804-1	40GE1/0/0

规划描述	本端设备名称	物理端口	对端设备名称	物理端口
	Leaf- CE6851HI-1	40GE1/0/4	Spine- CE12804-2	40GE1/0/1
	Leaf- CE6851HI-2	40GE1/0/3	Spine- CE12804-2	40GE1/0/0
	Leaf- CE6851HI-2	40GE1/0/4	Spine- CE12804-1	40GE1/0/1
规划Leaf与服务器的 互连端口,服务器双	Leaf- CE6851HI-1	10GE1/0/1 ~2	服务器	Eth0
网卡以负载分担方式 上连,保证可靠性的 同时负载分担转发, 提高链路利用率	Leaf- CE6851HI-2	10GE1/0/1 ~2	服务器	Eth1

Leaf-CE6851HI-3和Leaf-CE6851HI-4部署M-LAG,端口规划参见下表。

规划描述	本端设备名 称	物理端口	对端设备名称	物理端口
规划M-LAG的peer- link端口,用于承载	Leaf- CE6851HI-3	40GE1/0/1~2	Leaf- CE6851HI-4	40GE1/0/1 ~2
M-LAG的协议报文和 故障场景下的数据报 文,为保证可靠性, 建议至少2条peer-link 链路	Leaf- CE6851HI-4	40GE1/0/1~2	Leaf- CE6851HI-3	40GE1/0/1 ~2
规划Leaf连接Spine互 连端口	Leaf- CE6851HI-3	40GE1/0/3	Spine- CE12804-1	40GE1/0/2
	Leaf- CE6851HI-3	40GE1/0/4	Spine- CE12804-2	40GE1/0/3
	Leaf- CE6851HI-4	40GE1/0/3	Spine- CE12804-2	40GE1/0/2
	Leaf- CE6851HI-4	40GE1/0/4	Spine- CE12804-1	40GE1/0/3
规划Leaf与服务器的 互连端口,服务器双	Leaf- CE6851HI-3	10GE1/0/1	服务器	Eth0
网卡以负载分担方式 上连,保证可靠性的 同时负载分担转发, 提高链路利用率	Leaf- CE6851HI-4	10GE1/0/1	服务器	Eth1

Leaf-CE6851HI-5和Leaf-CE6851HI-6部署SVF场景,下连三台CE5810交换机,端口规划参见下表。

规划描述	本端设备名 称	物理端口	对端设备名称	物理端口	
规划Leaf间互连端口 作为堆叠口,堆叠链	Leaf- CE6851HI-5	40GE1/0/1~ 2	Leaf- CE6851HI-6	40GE1/0/1 ~2	
路至少2条	Leaf- CE6851HI-6	40GE1/0/1~ 2	Leaf- CE6851HI-5	40GE1/0/1 ~2	
规划Leaf与Spine互连 端口	Leaf- CE6851HI-5	40GE1/0/3	Spine- CE12804-1	40GE1/0/6	
	Leaf- CE6851HI-5	40GE1/0/4	Spine- CE12804-2	40GE1/0/7	
	Leaf- CE6851HI-6	40GE1/0/3	Spine- CE12804-1	40GE1/0/6	
	Leaf- CE6851HI-6	40GE1/0/4	Spine- CE12804-2	40GE1/0/7	
规划SVF父节点与叶	Leaf-	10GE1/0/1	CE5800-1	GE1/0/1	
子节点互联端口 	CE6851HI-5	CE003 HI-3	10GE1/0/2	CE5800-2	GE1/0/1
		10GE1/0/3	CE5800-3	GE1/0/1	
	Leaf-	10GE1/0/1	CE5800-1	GE1/0/2	
	CE6851HI-6	10GE1/0/2	CE5800-2	GE1/0/2	
		10GE1/0/3	CE5800-3	GE1/0/2	
规划叶子交换机与服	CE5810-1	GE1/0/1	服务器	Eth0	
务器的互连端口,服 务器双网卡以负载分 担方式上连,保证链 路可靠性	CE5810-2	GE1/0/1	服务器	Eth1	

Spine-CE12804-1& Spine-CE12804-2端口规划参见下表。

规划描述	本端设备名 称	物理端口	对端设备名称	物理端口
规划Spine与Leaf的互 连端口	Spine- CE12804-1	40GE1/0/0~3 40GE1/0/6~7	Leaf- CE6851HI-1 Leaf- CE6851HI-2	40GE1/0/3 ~4
			Leaf- CE6851HI-3	
			Leaf- CE6851HI-4	
			Leaf- CE6851HI-5	
			Leaf- CE6851HI-6	
	Spine- CE12804-2	40GE1/0/0~3 40GE1/0/6~7	Leaf- CE6851HI-1	40GE1/0/3 ~4
			Leaf- CE6851HI-2	
			Leaf- CE6851HI-3	
			Leaf- CE6851HI-4	
			Leaf- CE6851HI-5	
			Leaf- CE6851HI-6	
规划Spine与Gateway 的互连端口	Spine- CE12804-1	40GE1/0/4~5	Gateway- CE12808-1	40GE1/0/0
			Gateway- CE12808-2	
	Spine- CE12804-2	40GE1/0/4~5	Gateway- CE12808-1	40GE1/0/1
			Gateway- CE12808-2	

Gateway-CE12808-1& Gateway-CE12808-2端口规划参见下表。

规划描述	本端设备名称	物理端口	对端设备名称	物理端口
规划Gateway与Spine 互连端口	Gateway- CE12808-1	40GE1/0/0~1	Spine- CE12804-1 Spine-	40GE1/0/ 4
			CE12804-2	
	Gateway- CE12808-2	40GE1/0/0~1	Spine- CE12804-1	40GE1/0/ 5
			Spine- CE12804-2	
规划M-LAG的peer- link端口,用于承载M- LAG的协议报文和故障 场景下的数据报文,为 保证可靠性,建议至少	Gateway- CE12808-1	40GE1/0/23 40GE2/0/23	Gateway- CE12808-2	40GE1/0/ 23 40GE2/0/ 23
2条peer-link链路,且 跨板部署	Gateway- CE12808-2	40GE1/0/23 40GE2/0/23	Gateway- CE12808-1	40GE1/0/ 23 40GE2/0/ 23
规划Gateway与防火墙的互连端口 说明 两台主备镜像模式防火 墙必须使用同样的接口	Gateway- CE12808-1	10GE3/0/0~1	FW- USG9560-1 FW- USG9560-2	GE1/0/1
连接到同一个GW设备。 例如FW-1使用GE 1/0/1 接口连接GW-1,那么 FW-2也需要使用GE 1/0/1接口连接GW-1。	Gateway- CE12808-1	10GE3/0/2~3	FW- USG9560-1 FW- USG9560-2	GE1/0/3
	Gateway- CE12808-2	10GE3/0/0~1	FW- USG9560-1 FW- USG9560-2	GE1/0/2
	Gateway- CE12808-2	10GE3/0/2~3	FW- USG9560-1 FW- USG9560-2	GE1/0/4
规划Gateway与出口路 由器的互连端口	Gateway- CE12808-1	10GE3/0/4	Router-1	GE1/0/0
	Gateway- CE12808-1	10GE3/0/5	Router-2	GE1/0/1
	Gateway- CE12808-2	10GE3/0/4	Router-1	GE1/0/1

规划描述	本端设备名 称	物理端口	对端设备名称	物理端口
	Gateway- CE12808-2	10GE3/0/5	Router-2	GE1/0/0
规划Gateway之间的三 层互联接口,用于和出 口路由器组成"口"字 型和交叉联合出口网络	Gateway- CE12808-1	10GE3/0/6	Gateway- CE12808-2	10GE3/0/ 6

VLAN 规划

VLAN规划参见下表。

规划描述	规划建议	本示例使用ID
与防火墙 互连 VLAN	三层互连建议使用三层主接口,与FW因涉 及多个网段对接,需要规划VLAN,达到共 用物理链路目的	11 12(创建VLANIF 12绑定 VPN)
租户 VLAN	租户使用VXLAN较多,需要考虑后续的业 务扩容	10

□ 说明

对于CE6855HI和CE7855EI,在切换三层接口之前,要先配置**vlan reserved for maininterface** *startvlanid* **to** *endvlanid*,配置三层主接口专用的保留VLAN。

BD、VNI 规划

BD、VNI的规划参见下表。

规划描述	规划建议	本示例使用ID
BD	规划数量与VLAN相同,建议BD的ID与 VLAN的ID相同	本例统一使用BD 10
VNI	规划数量与BD相同,建议VNI的ID规划为 "BD+10000"(BD与VNI——对应,且 VNI的ID需要大于4096)	本例统一使用VNI 10010

RD 和 RT 值规划

RD的规划参见下表。

规划描述	规划建议	本示例使用ID
CE6851HI- 1、 CE6851HI- 2的RD值		11: 1
CE6851HI- 3的RD值	_	13: 1
CE6851HI- 4的RD值	_	14: 1
CE6851HI- 5、 CE6851HI- 6的RD值		15: 1
Gateway- CE12808-1 的RD值	_	16: 1
Gateway- CE12808-2 的RD值	_	17: 1

RT的规划参见下表。

规划描述	规划建议	本示例使用ID
CE6851HI- 1、 CE6851HI- 2的RT值		1: 1
CE6851HI- 3的RT值	_	1: 1
CE6851HI- 4的RT值	_	1: 1
CE6851HI- 5、 CE6851HI- 6的RT值		1: 1
Gateway- CE12808-1 的RT值	_	1: 1
Gateway- CE12808-2 的RT值	_	1: 1

IP 地址规划

网元接口地址规划参见下表,涉及互连地址、VTEP地址、BGP Router-ID、BGP建邻居 loopback地址、M-LAG心跳地址、业务网段地址。

端口互连IP地址规划参见下表。

规划描述	本端设备	对端设备	IP/MASK
规划CE6851HI-1、 CE6851HI-2与 Spine互连端口地	Leaf- CE6851HI-1&	Spine-CE12804-1	192.168.40.156/30 192.168.40.168/30
址	CE6851HI-2	Spine-CE12804-2	192.168.40.164/30 192.168.40.160/30
规划CE6851HI-3、 CE6851HI-4与 Spine互连端口地	Leaf- CE6851HI-3&	Spine-CE12804-1	192.168.41.156/30 192.168.41.168/30
址	CE6851HI-4	Spine-CE12804-2	192.168.41.164/30 192.168.41.160/30
规划CE6851HI-5、 CE6851HI-6与 Spine互连端口地	Leaf- CE6851HI-5&	Spine-CE12804-1	192.168.46.156/30 192.168.46.168/30
址	CE6851HI-6	Spine-CE12804-2	192.168.46.164/30 192.168.46.160/30
规划Spine- CE12804-1与 Gateway-	Spine-CE12804-1	Gateway- CE12808-1	192.168.42.156/30
CE12808-1/ Gateway- CE12808-2互连端 口地址		Gateway- CE12808-2	192.168.42.160/30
规划Spine- CE12804-2与	Spine-CE12804-2	Gateway- CE12808-1	192.168.43.156/30
Gateway- CE12808-1/ Gateway- CE12808-2互连端 口地址		Gateway- CE12808-2	192.168.43.160/30
规划Gateway- CE12808-1与 Router-1互连端口 地址	Gateway- CE12808-1	Router-1	192.168.44.156/30
规划Gateway- CE12808-2与 Router-2互连端口 地址	Gateway- CE12808-2	Router-2	192.168.44.160/30

规划描述	本端设备	对端设备	IP/MASK
规划Gateway- CE12808-1和 Gateway- CE12808-2互联	Gateway- CE12808-1	Gateway- CE12808-2	192.168.44.164/30
规划Gateway- CE12808-1&2与 FW互连端口地址	Gateway- CE12808-1 Gateway- CE12808-2	FW-USG9560-1 FW-USG9560-2	192.168.45.152/29 Virtual IP: 192.168.45.153 192.168.45.160/29 Virtual IP: 192.168.45.161

Loopback地址规划参见下表。

规划描述	设备	IP/MASK
规划Loopback0地址,作 为VTEP地址	Leaf-CE6851HI-1&	11.11.11.11/32
说明	CE6851HI-2	
Leaf-CE6851HI-3和Leaf-	Leaf-CE6851HI-3	11.11.11.12/32
CE6851HI-4组成M-LAG用于服务器双活接入,因此它们的VTEP IP地址相同。	Leaf-CE6851HI-4	同Leaf-CE6851HI-3: 11.11.11.12/32
Gateway-CE12808-1和 Gateway-CE12808-2组成 M-LAG用于防火墙双活接 入,因此它们的VTEP IP地址 相同。	Leaf-CE6851HI-5&	11.11.11.17/32
	CE6851HI-6	
	Gateway-CE12808-1	11.11.11.16/32
	Gateway-CE12808-2	同Gateway-CE12808-1: 11.11.11.16/32
规划Spine的Loopback0地	Spine-CE12804-1	11.11.11.14/32
址,作为Router ID	Spine-CE12804-2	11.11.11.15/32
规划Loopback1地址,作	Leaf-CE6851HI-3	13.13.13.13/32
为M-LAG心跳检测地址 	Leaf-CE6851HI-4	14.14.14.14/32
	Gateway-CE12808-1	18.18.18.18/32
	Gateway-CE12808-2	19.19.19.19/32
规划Loopback2地址,作	Gateway-CE12808-1	21.21.21.21/32
为对接Router的EBGP邻居 地址	Gateway-CE12808-2	22.22.22/32

业务地址规划参见下表。

租户	租户IP/MASK	网关VBDIF	VRF
服务器1	192.168.10.2/24	192.168.10.1/24	VPN 1
服务器2	192.168.10.3/24	192.168.10.1/24	VPN 1
服务器3	192.168.10.4/24	192.168.10.1/24	VPN 1

路由规划

Underlay网络使用的路由协议常见为EBGP和OSPF两种,BGP从多方面保证了网络的安全性、灵活性、稳定性、可靠性和高效性:

- BGP采用认证和GTSM的方式,保证了网络的安全性。
- BGP提供了丰富的路由策略,能够灵活的进行路由选路。
- BGP提供了路由聚合和路由衰减功能用于防止路由振荡,有效提高了网络的稳定性。
- BGP使用TCP作为其传输层协议(端口号为179),并支持BGP与BFD联动、BGP Auto FRR和BGP GR和NSR,提高了网络的可靠性。

从网络演进方面考虑,EBGP适用于大型网络,而OSPF适用于中小型网络,本文档使用EBGP作为示例。

路由相关的规划参见下表。

网元	AS域号	Router ID
Leaf- CE6851HI-1&CE6851HI-2	65021	Loopback0地址
Leaf-CE6851HI-3	65022	Loopback1地址
Leaf-CE6851HI-4	65022	Loopback1地址
Leaf- CE6851HI-5&CE6851HI-6	65024	Loopback0地址
Spine-CE12804-1	65010	Loopback0地址
Spine-CE12804-2	65010	Loopback0地址
Gateway-CE12808-1	65000	Loopback1地址
Gateway-CE12808-2	65000	Loopback1地址
Router-1	65047	Loopback0地址
Router-2	65048	Loopback0地址

3.1.1.1 配置 Leaf 堆叠工作组

两台Leaf交换机Leaf-CE6851HI-1和Leaf-CE6851HI-2组成堆叠组,分别上联Spine-1和Spine-2,下联服务器。

配置思路

- 1. 组建堆叠组:配置堆叠和双主检测,并重启设备、连接线缆,使堆叠生效。
- 2. 配置IP地址:配置Leaf与Spine的三层互联接口地址、Loopback0地址(作为Router-ID和VTEP IP)。
- 3. 配置服务器接入:介绍服务器接入堆叠组时,堆叠组交换机上的配置。
- 4. 配置路由:在堆叠组上配置BGP动态路由,邻居为两个Spine设备,使堆叠组上的网段与Spine上的网段三层可达。
- 5. 配置BGP EVPN:配置BGP EVPN作为VXLAN的控制平面、并配置BGP EVPN对等 体、EVPN实例以及头端复制功能。

组建堆叠组

步骤1 配置Leaf-CE6851HI-1的堆叠成员ID为1,优先级为150,Domain ID为10。

<HUAWEI> system-view

[~HUAWEI] sysname Leaf-CE6851HI-1

[*HUAWEI] commit

[~Leaf-CE6851HI-1] stack

[~Leaf-CE6851HI-1-stack] stack member 1 priority 150

[*Leaf-CE6851HI-1-stack] stack member 1 domain 10

[*Leaf-CE6851HI-1-stack] quit

[*Leaf-CE6851HI-1] commit

步骤2 配置Leaf-CE6851HI-2的Domain ID为10。

<HUAWEI> system-view

[~HUAWEI] sysname Leaf-CE6851HI-2

[*HUAWEI] commit

[~Leaf-CE6851HI-2] stack

[~Leaf-CE6851HI-2-stack] stack member 1 renumber 2 inherit-config

[*Leaf-CE6851HI-2-stack] stack member 1 domain 10

[*Leaf-CE6851HI-2-stack] quit

[*Leaf-CE6851HI-2] commit

步骤3 配置堆叠端口。

将Leaf-CE6851HI-1的业务口40GE1/0/1~40GE1/0/2加入堆叠端口1/1。

[~Leaf-CE6851HI-1] interface stack-port 1/1

[*Leaf-CE6851HI-1-Stack-Port1/1] port member-group interface 40ge 1/0/1 to 1/0/2

Warning: After the configuration is complete,1.The interface(s) (40GE1/0/1-1/0/2) will be converted to stack mode and be configured with the port crc-statistics trigger error-down command if the configuration does not exist.2.The interface(s) may go Error-Down (crc-statistics) because there is no shutdown configuration on the interfaces.Continue? [Y/N]: y

[*Leaf-CE6851HI-1-Stack-Port1/1] quit

[*Leaf-CE6851HI-1] commit

[~Leaf-CE6851HI-1] quit

将Leaf-CE6851HI-2的业务口40GE1/0/1~40GE1/0/2加入堆叠端口1/1。

[~Leaf-CE6851HI-2] interface stack-port 1/1

[*Leaf-CE6851HI-2-Stack-Port1/1] port member-group interface 40ge 1/0/1 to 1/0/2

Warning: After the configuration is complete,1.The interface(s) (40GE1/0/1-1/0/2) will be converted to stack mode and be configured with the port crc-statistics trigger error-down command if the configuration does not exist.2.The interface(s) may go Error-Down (crc-statistics) because there is no shutdown configuration on the interfaces.Continue? [Y/N]: y

[*Leaf-CE6851HI-2-Stack-Port1/1] quit

[*Leaf-CE6851HI-2] commit

[~Leaf-CE6851HI-2] quit

步骤4 保存配置后重启设备。

保存Leaf-CE6851HI-1的配置,然后重启设备。Leaf-CE6851HI-2的与之类似,不再赘述。

<Leaf-CE6851HI-1> save

Warning: The current configuration will be written to the device. Continue? [Y/N]: y

<Leaf-CE6851HI-1> reboot

Warning: The system will reboot. Continue? [Y/N]:v

步骤5 连接堆叠线缆,建立堆叠。

步骤6 堆叠组建后执行命令save保存配置。

步骤7 堆叠建立后,配置堆叠组的双主检测,防止堆叠组分裂后网络中存在两个配置冲突的网络设备。

采用业务口直连的方式配置双主检测,建议规划的直连链路至少2条,保证可靠性, 配置示例如下。

```
[~Leaf-CE6851HI-1] sysname Leaf-CE6851HI-1&CE6851HI-2
[*Leaf-CE6851HI-1] commit
[~Leaf-CE6851HI-1&CE6851HI-2] interface 10ge 1/0/30
[~Leaf-CE6851HI-1&CE6851HI-2-10GE1/0/30] description "for DAD"
[*Leaf-CE6851HI-1&CE6851HI-2-10GE1/0/30] dual-active detect mode direct
Warning: The interface will block common data packets, except BPDU packets. Continue? [Y/N]: y
[*Leaf-CE6851HI-1&CE6851HI-2-10GE1/0/30] quit
[*Leaf-CE6851HI-1&CE6851HI-2] interface 10ge 1/0/31
[*Leaf-CE6851HI-1&CE6851HI-2-10GE1/0/31] description "for DAD"
[*Leaf-CE6851HI-1&CE6851HI-2-10GE1/0/31] dual-active detect mode direct
Warning: The interface will block common data packets, except BPDU packets. Continue? [Y/N]: y
[*Leaf-CE6851HI-1&CE6851HI-2-10GE1/0/31] quit
[*Leaf-CE6851HI-1&CE6851HI-2] interface 10ge 2/0/30
[*Leaf-CE6851HI-1&CE6851HI-2-10GE2/0/30] description "for DAD"
[*Leaf-CE6851HI-1&CE6851HI-2-10GE2/0/30] dual-active detect mode direct
Warning: The interface will block common data packets, except BPDU packets. Continue? [Y/N]: y
[*Leaf-CE6851HI-1&CE6851HI-2-10GE2/0/30] quit
[*Leaf-CE6851HI-1&CE6851HI-2] interface 10ge 2/0/31
[*Leaf-CE6851HI-1&CE6851HI-2-10GE2/0/31] description "for DAD"
[*Leaf-CE6851HI-1&CE6851HI-2-10GE2/0/31] dual-active detect mode direct
Warning: The interface will block common data packets, except BPDU packets. Continue? [Y/N]: v
[*Leaf-CE6851HI-1&CE6851HI-2-10GE2/0/31] commit
[~Leaf-CE6851HI-1&CE6851HI-2-10GE2/0/31] quit
```

步骤8 配置作为Server Leaf的堆叠系统的MAC地址为成员设备1的MAC地址。

```
[~Leaf-CE6851HI-1&CE6851HI-2] stack
[~Leaf-CE6851HI-1&CE6851HI-2-stack] set system mac-address slot 1 //如果CE12800系列交换机作为
Server Leaf时,此处配置命令为set system mac-address chassis 1
```

----结束

配置 IP 地址

步骤1 配置互连接口地址。

□说明

对于CE6855HI和CE7855EI,在切换三层接口之前,要先配置vlan reserved for maininterface startylanid to endylanid,配置三层主接口专用的保留VLAN。

```
[~Leaf-CE6851HI-1&CE6851HI-2] interface 40ge 1/0/3
[~Leaf-CE6851HI-1&CE6851HI-2] description "to_Spine-CE12804-1-40GE1/0/0"
[*Leaf-CE6851HI-1&CE6851HI-2-40GE1/0/3] undo portswitch
[*Leaf-CE6851HI-1&CE6851HI-2-40GE1/0/3] ip address 192.168.40.157 30
[*Leaf-CE6851HI-1&CE6851HI-2-40GE1/0/3] commit
[~Leaf-CE6851HI-1&CE6851HI-2-40GE1/0/3] quit
[~Leaf-CE6851HI-1&CE6851HI-2] interface 40ge 1/0/4
[~Leaf-CE6851HI-1&CE6851HI-2] description "to_Spine-CE12804-2-40GE1/0/1"
[*Leaf-CE6851HI-1&CE6851HI-2-40GE1/0/4] undo portswitch
[*Leaf-CE6851HI-1&CE6851HI-2-40GE1/0/4] ip address 192.168.40.165 30
[*Leaf-CE6851HI-1&CE6851HI-2-40GE1/0/4] commit
[~Leaf-CE6851HI-1&CE6851HI-2-40GE1/0/4] quit
```

```
[~Leaf-CE6851HI-1&CE6851HI-2] interface 40ge 2/0/3
[~Leaf-CE6851HI-1&CE6851HI-2] description "to_Spine-CE12804-2-40GE1/0/0"
[*Leaf-CE6851HI-1&CE6851HI-2-40GE2/0/3] undo portswitch
[*Leaf-CE6851HI-1&CE6851HI-2-40GE2/0/3] ip address 192.168.40.161 30
[*Leaf-CE6851HI-1&CE6851HI-2-40GE2/0/3] commit
[~Leaf-CE6851HI-1&CE6851HI-2-40GE2/0/3] quit
[~Leaf-CE6851HI-1&CE6851HI-2] interface 40ge 2/0/4
[~Leaf-CE6851HI-1&CE6851HI-2] description "to_Spine-CE12804-1-40GE1/0/1"
[*Leaf-CE6851HI-1&CE6851HI-2-40GE2/0/4] undo portswitch
[*Leaf-CE6851HI-1&CE6851HI-2-40GE2/0/4] ip address 192.168.40.169 30
[*Leaf-CE6851HI-1&CE6851HI-2-40GE2/0/4] commit
[~Leaf-CE6851HI-1&CE6851HI-2-40GE2/0/4] quit
```

步骤2 配置Loopback口地址。

```
[~Leaf-CE6851HI-1&CE6851HI-2] interface loopback 0
[~Leaf-CE6851HI-1&CE6851HI-2] description VTEP&Router-ID
[*Leaf-CE6851HI-1&CE6851HI-2-LoopBack0] ip address 11.11.11.11 32
[*Leaf-CE6851HI-1&CE6851HI-2-LoopBack0] commit
[~Leaf-CE6851HI-1&CE6851HI-2-LoopBack0] quit
```

----结束

配置服务器接入

介绍服务器接入堆叠组时,堆叠组交换机上的配置。

步骤1 在Leaf-CE6851HI-1与Leaf-CE6851HI-2上配置以负载均衡模式接入堆叠组。

```
[~Leaf-CE6851HI-1&CE6851HI-2] interface eth-trunk 1
[*Leaf-CE6851HI-1&CE6851HI-2-Eth-Trunk1] mode lacp-static
[*Leaf-CE6851HI-1&CE6851HI-2-Eth-Trunk1] port link-type trunk
[*Leaf-CE6851HI-1&CE6851HI-2-Eth-Trunk1] undo port trunk allow-pass vlan 1
[*Leaf-CE6851HI-1&CE6851HI-2-Eth-Trunk1] trunkport 10ge 1/0/1
[*Leaf-CE6851HI-1&CE6851HI-2-Eth-Trunk1] trunkport 10ge 2/0/1
[*Leaf-CE6851HI-1&CE6851HI-2-Eth-Trunk1] quit
[*Leaf-CE6851HI-1&CE6851HI-2] commit
```

步骤2 在Leaf-CE6851HI-1与Leaf-CE6851HI-2上配置VXLAN业务接入点。

```
[~Leaf-CE6851HI-1&CE6851HI-2] bridge-domain 10
[~Leaf-CE6851HI-1&CE6851HI-2-bd10] vxlan vni 10010
[*Leaf-CE6851HI-1&CE6851HI-2-bd10] quit
[*Leaf-CE6851HI-1&CE6851HI-2] commit

[~Leaf-CE6851HI-1&CE6851HI-2] interface eth-trunk 1.1 mode l2
[*Leaf-CE6851HI-1&CE6851HI-2-Eth-Trunk1.1] encapsulation dot1q vid 30
[*Leaf-CE6851HI-1&CE6851HI-2-Eth-Trunk1.1] bridge-domain 10
[*Leaf-CE6851HI-1&CE6851HI-2-Eth-Trunk1.1] quit
[*Leaf-CE6851HI-1&CE6851HI-2] commit
```

----结束

配置路由

执行以下步骤配置Underlay的路由(本例以BGP为例,也可以使用OSPF)。

步骤1 在堆叠组上配置BGP路由,对接Spine。

```
[~Leaf-CE6851HI-1&CE6851HI-2] bgp 65021

[*Leaf-CE6851HI-1&CE6851HI-2-bgp] router-id 11.11.11.11

[*Leaf-CE6851HI-1&CE6851HI-2-bgp] timer keepalive 10 hold 30

[*Leaf-CE6851HI-1&CE6851HI-2-bgp] group Spine-CE12804 external

[*Leaf-CE6851HI-1&CE6851HI-2-bgp] peer Spine-CE12804 as-number 65010

[*Leaf-CE6851HI-1&CE6851HI-2-bgp] peer 192.168.40.158 as-number 65010

[*Leaf-CE6851HI-1&CE6851HI-2-bgp] peer 192.168.40.158 group Spine-CE12804

[*Leaf-CE6851HI-1&CE6851HI-2-bgp] peer 192.168.40.170 as-number 65010

[*Leaf-CE6851HI-1&CE6851HI-2-bgp] peer 192.168.40.170 group Spine-CE12804

[*Leaf-CE6851HI-1&CE6851HI-2-bgp] peer 192.168.40.166 as-number 65010
```

```
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] peer 192.168.40.166 group Spine-CE12804
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] peer 192.168.40.162 as-number 65010
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] peer 192.168.40.162 group Spine-CE12804
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] ipv4-family unicast
[*Leaf-CE6851HI-1&CE6851HI-2-bgp-af-ipv4] preference 20 200 10
[*Leaf-CE6851HI-1&CE6851HI-2-bgp-af-ipv4] network 11.11.11.11 255.255.255.255
[*Leaf-CE6851HI-1&CE6851HI-2-bgp-af-ipv4] maximum load-balancing 32
[*Leaf-CE6851HI-1&CE6851HI-2-bgp-af-ipv4] quit
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] quit
[*Leaf-CE6851HI-1&CE6851HI-2] commit
```

----结束

配置 BGP EVPN

使用BGP EVPN作为VXLAN的控制平面时,需执行以下步骤将Leaf作为BGP路由反射器客户端,Spine作为BGP路由器反射器,在Leaf与Spine之间建立IBGP EVPN对等体关系,配置EVPN实例和头端复制功能。

步骤1 在堆叠组上使能EVPN作为VXLAN的控制平面功能。

```
[~Leaf-CE6851HI-1&CE6851HI-2] evpn-overlay enable
[*Leaf-CE6851HI-1&CE6851HI-2] commit
```

步骤2 配置堆叠组与Spine节点分别建立IBGP EVPN对等体关系,Spine作为路由反射器RR。

```
[~Leaf-CE6851HI-1&CE6851HI-2] bgp 100 instance evpn1
[*Leaf-CE6851HI-1&CE6851HI-2-bgp-instance-evpn1] router-id 11.11.11.11
[*Leaf-CE6851HI-1&CE6851HI-2-bgp-instance-evpn1] peer 11.11.11.14 as-number 100
//与网关Spine-CE12804-1建立BGP EVPN对等体关系
[*Leaf-CE6851HI-1&CE6851HI-2-bgp-instance-evpn1] peer 11.11.11.14 connect-interface loopback 0
[*Leaf-CE6851HI-1&CE6851HI-2-bgp-instance-evpn1] peer 11.11.11.15 as-number 100
//与网关Spine-CE12804-2建立BGP EVPN对等体关系
[*Leaf-CE6851HI-1&CE6851HI-2-bgp-instance-evpn1] peer 11.11.11.15 connect-interface loopback 0
[*Leaf-CE6851HI-1&CE6851HI-2-bgp-instance-evpn1] l2vpn-family evpn
[*Leaf-CE6851HI-1&CE6851HI-2-bgp-instance-evpn1-af-evpn] peer 11.11.11.14 enable
[*Leaf-CE6851HI-1&CE6851HI-2-bgp-instance-evpn1-af-evpn] peer 11.11.11.14 advertise arp
[*Leaf-CE6851HI-1&CE6851HI-2-bgp-instance-evpn1-af-evpn] peer 11.11.11.15 enable
[*Leaf-CE6851HI-1&CE6851HI-2-bgp-instance-evpn1-af-evpn] peer 11.11.11.15 advertise arp
[*Leaf-CE6851HI-1&CE6851HI-2-bgp-instance-evpn1-af-evpn] quit
[*Leaf-CE6851HI-1&CE6851HI-2-bgp-instance-evpn1] quit
[*Leaf-CE6851HI-1&CE6851HI-2] commit
```

步骤3 在堆叠组上配置EVPN实例。

```
[~Leaf-CE6851HI-1&CE6851HI-2] bridge-domain 10

[~Leaf-CE6851HI-1&CE6851HI-2-bd10] evpn

[*Leaf-CE6851HI-1&CE6851HI-2-bd10-evpn] route-distinguisher 11:1

[*Leaf-CE6851HI-1&CE6851HI-2-bd10-evpn] vpn-target 1:1

[*Leaf-CE6851HI-1&CE6851HI-2-bd10-evpn] quit

[*Leaf-CE6851HI-1&CE6851HI-2-bd10] quit

[*Leaf-CE6851HI-1&CE6851HI-2] commit
```

步骤4 在堆叠组上配置头端复制功能。

```
[~Leaf-CE6851HI-1&CE6851HI-2] interface nve 1

[*Leaf-CE6851HI-1&CE6851HI-2-Nve1] source 11.11.11.11

[*Leaf-CE6851HI-1&CE6851HI-2-Nve1] vni 10010 head-end peer-list protocol bgp

[*Leaf-CE6851HI-1&CE6851HI-2-Nve1] quit

[*Leaf-CE6851HI-1&CE6851HI-2] commit
```

----结束

3.1.1.2 配置 Leaf M-LAG 工作组

两台Leaf交换机Leaf-CE6851HI-3和Leaf-CE6851HI-4组成M-LAG,上联Spine-CE12804-1和Spine-CE12804-2,下联服务器。

配置思路

- 1. 配置IP地址:配置Leaf交换机与Spine的三层互联接口地址、Loopback0地址(作为VTEP IP)、Loopback1地址(作为Router-ID)。
- 2. 配置M-LAG: 在交换机上配置M-LAG全局模式、DFS组、Peer-Link,并配置服务器接入M-LAG,配置Monitor Link关联上行接口和下行接口。
- 3. 配置路由:在M-LAG组上配置BGP动态路由,邻居为两个Spine设备,使M-LAG组 上的网段与Spine上的网段三层可达。
- 4. 配置BGP EVPN: 配置BGP EVPN作为VXLAN的控制平面、并配置BGP EVPN对等体、EVPN实例以及头端复制功能。

配置 IP 地址

步骤1 配置Leaf-CE6851HI-3与Leaf-CE6851HI-4直连接口地址。

□□说明

对于CE6855HI和CE7855EI,在切换三层接口之前,要先配置命令vlan reserved for maininterface startvlanid to endvlanid,配置三层主接口专用的保留VLAN。

```
[~HUAWEI] sysname Leaf-CE6851HI-3
[~HUAWEI] commit
[~Leaf-CE6851HI-3] interface 40ge 1/0/3
[~Leaf-CE6851HI-3-40GE1/0/3] description "to_Spine-CE12804-1-40GE1/0/2"
[~Leaf-CE6851HI-3-40GE1/0/3] undo portswitch
[*Leaf-CE6851HI-3-40GE1/0/3] ip address 192.168.41.157 30
[*Leaf-CE6851HI-3-40GE1/0/3] commit
[~Leaf-CE6851HI-3-40GE1/0/3] quit
[~Leaf-CE6851HI-3] interface 40ge 1/0/4
[*Leaf-CE6851HI-3-40GE1/0/4] description "to_Spine-CE12804-2-40GE1/0/3"
[~Leaf-CE6851HI-3-40GE1/0/4] undo portswitch
[*Leaf-CE6851HI-3-40GE1/0/4] ip address 192.168.41.165 30
[*Leaf-CE6851HI-3-40GE1/0/4] commit
[~Leaf-CE6851HI-3-40GE1/0/4] quit
[~HUAWEI] sysname Leaf-CE6851HI-4
[~HUAWEI] commit
[~Leaf-CE6851HI-4] interface 40ge 1/0/3
[~Leaf-CE6851HI-4-40GE1/0/3] description "to_Spine-CE12804-2-40GE1/0/2"
[*Leaf-CE6851HI-4-40GE1/0/3] undo portswitch
[*Leaf-CE6851HI-4-40GE1/0/3] ip address 192.168.41.169 30
[*Leaf-CE6851HI-4-40GE1/0/3] commit
[~Leaf-CE6851HI-4-40GE1/0/3] quit
[~Leaf-CE6851HI-4] interface 40ge 1/0/4
[~Leaf-CE6851HI-4-40GE1/0/4] description "to_Spine-CE12804-1-40GE1/0/3"
[*Leaf-CE6851HI-4-40GE1/0/4] undo portswitch
[*Leaf-CE6851HI-4-40GE1/0/4] ip address 192.168.41.161 30
[*Leaf-CE6851HI-4-40GE1/0/4] commit
[~Leaf-CE6851HI-4-40GE1/0/4] quit
```

步骤2 配Loopback口地址,规划用Loopback0地址作为VTEP,用Loopback1地址作为Router-ID。

```
[~Leaf-CE6851HI-3] interface loopback 0 //作为VTEP地址
[*Leaf-CE6851HI-3-LoopBack0] ip address 11.11.11.12 32
[*Leaf-CE6851HI-3-LoopBack0] commit
[~Leaf-CE6851HI-3-LoopBack0] quit
[~Leaf-CE6851HI-3] interface loopback 1
[*Leaf-CE6851HI-3-LoopBack1] ip address 13.13.13.13 32
[*Leaf-CE6851HI-3-LoopBack1] commit
[~Leaf-CE6851HI-3-LoopBack1] quit

[~Leaf-CE6851HI-4] interface loopback 0 //作为VTEP地址
[*Leaf-CE6851HI-4-LoopBack0] ip address 11.11.11.12 32
[*Leaf-CE6851HI-4-LoopBack0] commit
```

```
[~Leaf-CE6851HI-4-LoopBack0] quit
[~Leaf-CE6851HI-4] interface loopback 1
[*Leaf-CE6851HI-4-LoopBack1] ip address 14.14.14.14 32
[*Leaf-CE6851HI-4-LoopBack1] commit
[~Leaf-CE6851HI-4-LoopBack1] quit
```

----结束

配置路由

执行以下步骤配置Underlay的路由(本例以BGP为例,也可以使用OSPF)。

步骤1 在Leaf-CE6851HI-3上配置BGP路由,对接Spine。

```
[~Leaf-CE6851HI-3] bqp 65022
[*Leaf-CE6851HI-3-bgp] router-id 13.13.13.13
[*Leaf-CE6851HI-3-bgp] timer keepalive 10 hold 30
[*Leaf-CE6851HI-3-bgp] group Spine-CE12804 external
[*Leaf-CE6851HI-3-bgp] peer Spine-CE12804 as-number 65010
[*Leaf-CE6851HI-3-bgp] peer 192.168.41.158 as-number 65010
[*Leaf-CE6851HI-3-bgp] peer 192.168.41.158 group Spine-CE12804
[*Leaf-CE6851HI-3-bgp] peer 192.168.41.166 as-number 65010
[*Leaf-CE6851HI-3-bgp] peer 192.168.41.166 group Spine-CE12804
[*Leaf-CE6851HI-3-bgp] ipv4-family unicast
[*Leaf-CE6851HI-3-bgp-af-ipv4] preference 20 200 10
[*Leaf-CE6851HI-3-bgp-af-ipv4] network 11.11.11.12 255.255.255.255
[*Leaf-CE6851HI-3-bgp-af-ipv4] network 13.13.13.13 255.255.255.255
[*Leaf-CE6851HI-3-bgp-af-ipv4] maximum load-balancing 32
[*Leaf-CE6851HI-3-bgp-af-ipv4] quit
[*Leaf-CE6851HI-3-bgp] quit
[*Leaf-CE6851HI-3] commit
```

步骤2 在Leaf-CE6851HI-4上配置BGP路由,对接Spine。

```
[~Leaf-CE6851HI-4] bgp 65022
[*Leaf-CE6851HI-4-bgp] router-id 14.14.14.14
[*Leaf-CE6851HI-4-bgp] timer keepalive 10 hold 30
[*Leaf-CE6851HI-4-bgp] group Spine-CE12804 external
[*Leaf-CE6851HI-4-bgp] peer Spine-CE12804 as-number 65010
[*Leaf-CE6851HI-4-bgp] peer 192.168.41.170 as-number 65010
[*Leaf-CE6851HI-4-bqp] peer 192.168.41.170 group Spine-CE12804
[*Leaf-CE6851HI-4-bgp] peer 192.168.41.162 as-number 65010
[*Leaf-CE6851HI-4-bgp] peer 192.168.41.162 group Spine-CE12804
[*Leaf-CE6851HI-4-bgp] ipv4-family unicast
[*Leaf-CE6851HI-4-bgp-af-ipv4] preference 20 200 10
[*Leaf-CE6851HI-4-bgp-af-ipv4] network 11.11.11.12 255.255.255.255
[*Leaf-CE6851HI-4-bgp-af-ipv4] network 14.14.14.14 255.255.255.255
[*Leaf-CE6851HI-4-bgp-af-ipv4] maximum load-balancing 32
[*Leaf-CE6851HI-4-bgp-af-ipv4] quit
[*Leaf-CE6851HI-4-bgp] quit
[*Leaf-CE6851HI-4] commit
```

----结束

配置 M-LAG

步骤1 配置M-LAG模式。

```
<Leaf-CE6851HI-3> system-view
[~Leaf-CE6851HI-3] stp mode rstp
[*Leaf-CE6851HI-3] stp v-stp enable
[*Leaf-CE6851HI-3] lacp m-lag system-id 00e0-fc00-0001
[*Leaf-CE6851HI-3] commit

<Leaf-CE6851HI-4> system-view
[~Leaf-CE6851HI-4] stp mode rstp
```

```
[*Leaf-CE6851HI-4] stp v-stp enable
[*Leaf-CE6851HI-4] lacp m-lag system-id 00e0-fc00-0001
[*Leaf-CE6851HI-4] commit
```

步骤2 在Leaf-CE6851HI-3与Leaf-CE6851HI-4上配置M-LAG的DFS Group。

```
[~Leaf-CE6851HI-3] dfs-group 1
[*Leaf-CE6851HI-3-dfs-group-1] source ip 13.13.13.13
[*Leaf-CE6851HI-3-dfs-group-1] priority 150
[*Leaf-CE6851HI-3-dfs-group-1] quit
[*Leaf-CE6851HI-3] commit

[~Leaf-CE6851HI-4] dfs-group 1
[*Leaf-CE6851HI-4-dfs-group-1] source ip 14.14.14.14
[*Leaf-CE6851HI-4-dfs-group-1] priority 120
[*Leaf-CE6851HI-4-dfs-group-1] quit
[*Leaf-CE6851HI-4] commit
```

步骤3 在Leaf-CE6851HI-3与Leaf-CE6851HI-4上配置M-LAG的Peer-Link。

```
[~Leaf-CE6851HI-3] interface eth-trunk 0
[*Leaf-CE6851HI-3-Eth-Trunk0] trunkport 40ge 1/0/1
[*Leaf-CE6851HI-3-Eth-Trunk0] trunkport 40ge 1/0/2
[*Leaf-CE6851HI-3-Eth-Trunk0] mode lacp-static
[*Leaf-CE6851HI-3-Eth-Trunk0] peer-link 1
[*Leaf-CE6851HI-3-Eth-Trunk0] quit
[*Leaf-CE6851HI-3] commit

[*Leaf-CE6851HI-4] interface eth-trunk 0
[*Leaf-CE6851HI-4-Eth-Trunk0] trunkport 40ge 1/0/1
[*Leaf-CE6851HI-4-Eth-Trunk0] trunkport 40ge 1/0/2
[*Leaf-CE6851HI-4-Eth-Trunk0] mode lacp-static
[*Leaf-CE6851HI-4-Eth-Trunk0] peer-link 1
[*Leaf-CE6851HI-4-Eth-Trunk0] quit
[*Leaf-CE6851HI-4-Eth-Trunk0] quit
```

步骤4 配置服务器接入M-LAG。

在Leaf-CE6851HI-3与Leaf-CE6851HI-4上配置M-LAG的成员口(服务器以负载均衡模式接入M-LAG)。

```
[~Leaf-CE6851HI-3] interface eth-trunk 1
[*Leaf-CE6851HI-3-Eth-Trunk1] mode lacp-static
[*Leaf-CE6851HI-3-Eth-Trunk1] port link-type trunk
[*Leaf-CE6851HI-3-Eth-Trunk1] undo port trunk allow-pass vlan 1
[*Leaf-CE6851HI-3-Eth-Trunk1] trunkport 10ge 1/0/1
[*Leaf-CE6851HI-3-Eth-Trunk1] dfs-group 1 m-lag 1
[*Leaf-CE6851HI-3-Eth-Trunk1] quit
[*Leaf-CE6851HI-3] commit
[~Leaf-CE6851HI-4] interface eth-trunk 1
[*Leaf-CE6851HI-4-Eth-Trunk1] mode lacp-static
[*Leaf-CE6851HI-4-Eth-Trunk1] port link-type trunk
[*Leaf-CE6851HI-4-Eth-Trunk1] trunkport 10ge 1/0/1
[*Leaf-CE6851HI-4-Eth-Trunk1] undo port trunk allow-pass vlan 1
[*Leaf-CE6851HI-4-Eth-Trunk1] dfs-group 1 m-lag 1
[*Leaf-CE6851HI-4-Eth-Trunk1] quit
[*Leaf-CE6851HI-4] commit
```

步骤5 在Leaf-CE6851HI-3与Leaf-CE6851HI-4上配置VXLAN业务接入点。

```
[~Leaf-CE6851HI-3] bridge-domain 10
[*Leaf-CE6851HI-3-bd10] vxlan vni 10010
[*Leaf-CE6851HI-3-bd10] quit
[*Leaf-CE6851HI-3] commit

[~Leaf-CE6851HI-3] interface eth-trunk 1.1 mode l2
[*Leaf-CE6851HI-3-Eth-Trunk1.1] encapsulation dot1q vid 30
[*Leaf-CE6851HI-3-Eth-Trunk1.1] bridge-domain 10
[*Leaf-CE6851HI-3-Eth-Trunk1.1] quit
```

```
[*Leaf-CE6851HI-3] commit

[~Leaf-CE6851HI-4] bridge-domain 10
[*Leaf-CE6851HI-4-bd10] vxlan vni 10010
[*Leaf-CE6851HI-4-bd10] quit
[*Leaf-CE6851HI-4] commit

[~Leaf-CE6851HI-4] interface eth-trunk 1.1 mode l2
[*Leaf-CE6851HI-4-Eth-Trunk1.1] encapsulation dot1q vid 30
[*Leaf-CE6851HI-4-Eth-Trunk1.1] bridge-domain 10
[*Leaf-CE6851HI-4-Eth-Trunk1.1] quit
[*Leaf-CE6851HI-4] commit
```

步骤6 分别在Leaf-CE6851HI-3与Leaf-CE6851HI-4上配置Monitor Link关联上行接口和下行接口,避免单台设备的所有上行链路都发生故障,本台设备用户侧流量无法转发。

```
[~Leaf-CE6851HI-3] monitor-link group 1
[*Leaf-CE6851HI-3-mtlk-group1] port 40ge 1/0/3 uplink
[*Leaf-CE6851HI-3-mtlk-group1] port 40ge 1/0/4 uplink
[*Leaf-CE6851HI-3-mtlk-group1] port 10ge 1/0/1 downlink 1
[*Leaf-CE6851HI-3-mtlk-group1] commit

[~Leaf-CE6851HI-4] monitor-link group 1
[*Leaf-CE6851HI-4-mtlk-group1] port 40ge 1/0/3 uplink
[*Leaf-CE6851HI-4-mtlk-group1] port 40ge 1/0/4 uplink
[*Leaf-CE6851HI-4-mtlk-group1] port 10ge 1/0/1 downlink 1
[*Leaf-CE6851HI-4-mtlk-group1] commit
```

----结束

配置 BGP-EVPN

使用BGP EVPN作为VXLAN的控制平面时,需执行以下步骤将Leaf作为BGP路由反射器客户端,Spine作为BGP路由器反射器,在Leaf与Spine之间建立IBGP EVPN对等体关系,配置EVPN实例和头端复制功能。

步骤1 在Leaf-CE6851HI-3和Leaf-CE6851HI-4上使能EVPN作为VXLAN的控制平面功能。

```
[~Leaf-CE6851HI-3] evpn-overlay enable

[*Leaf-CE6851HI-3] commit

[~Leaf-CE6851HI-4] evpn-overlay enable

[*Leaf-CE6851HI-4] commit
```

步骤2 配置Leaf-CE6851HI-3和Leaf-CE6851HI-4上分别与Spine节点建立BGP EVPN对等体关系,Spine作为路由反射器RR。

在Leaf-CE6851HI-3上配置与Spine建立IBGP EVPN对等体关系。

```
[~Leaf-CE6851HI-3] bgp 100 instance evpn1
[*Leaf-CE6851HI-3-bgp-instance-evpn1] router-id 13.13.13.13
[*Leaf-CE6851HI-3-bgp-instance-evpn1] peer 11.11.11.14 as-number 100
//与网关Spine-CE12804-1建立BGP EVPN对等体关系
[*Leaf-CE6851HI-3-bgp-instance-evpn1] peer 11.11.11.14 connect-interface loopback 1
[*Leaf-CE6851HI-3-bgp-instance-evpn1] peer 11.11.11.15 as-number 100
//与网关Spine-CE12804-2建立BGP EVPN对等体关系
[*Leaf-CE6851HI-3-bgp-instance-evpn1] peer 11.11.11.15 connect-interface loopback 1
[*Leaf-CE6851HI-3-bgp-instance-evpn1] l2vpn-family evpn
[*Leaf-CE6851HI-3-bgp-instance-evpn1-af-evpn] peer 11.11.11.14 enable
[*Leaf-CE6851HI-3-bgp-instance-evpn1-af-evpn] peer 11.11.11.14 advertise arp
[*Leaf-CE6851HI-3-bgp-instance-evpn1-af-evpn] peer 11.11.11.15 enable
[*Leaf-CE6851HI-3-bgp-instance-evpn1-af-evpn] peer 11.11.11.15 advertise arp
[*Leaf-CE6851HI-3-bgp-instance-evpn1-af-evpn] quit
[*Leaf-CE6851HI-3-bgp-instance-evpn1] quit
[*Leaf-CE6851HI-3] commit
```

在Leaf-CE6851HI-4上配置与Spine建立IBGP EVPN对等体关系。

```
[~Leaf-CE6851HI-4] bgp 100 instance evpn1
[*Leaf-CE6851HI-4-bgp-instance-evpn1] router-id 14.14.14.14
[*Leaf-CE6851HI-4-bgp-instance-evpn1] peer 11.11.11.14 as-number 100
-
//与网关Spine-CE12804-1建立BGP EVPN对等体关系
[*Leaf-CE6851HI-4-bgp-instance-evpn1] peer 11.11.11.14 connect-interface loopback 1
[*Leaf-CE6851HI-4-bgp-instance-evpn1] peer 11.11.11.15 as-number 100
[*Leaf-CE6851HI-4-bgp-instance-evpn1] peer 11.11.11.15 connect-interface loopback 1
[*Leaf-CE6851HI-4-bgp-instance-evpn1] [2vpn-family evpn
[*Leaf-CE6851HI-4-bgp-instance-evpn1-af-evpn] peer 11.11.11.14 enable
[*Leaf-CE6851HI-4-bgp-instance-evpn1-af-evpn] peer 11.11.11.14 advertise arp
[*Leaf-CE6851HI-4-bgp-instance-evpn1-af-evpn] peer 11.11.11.15 enable
[*Leaf-CE6851HI-4-bgp-instance-evpn1-af-evpn] peer 11.11.11.15 advertise arp
[*Leaf-CE6851HI-4-bgp-instance-evpn1-af-evpn] quit
[*Leaf-CE6851HI-4-bgp-instance-evpn1] quit
[*Leaf-CE6851HI-4] commit
```

步骤3 # 在Leaf-CE6851HI-3上配置EVPN实例。

```
[~Leaf-CE6851HI-3] bridge-domain 10
[~Leaf-CE6851HI-3-bd10] evpn
[*Leaf-CE6851HI-3-bd10-evpn] route-distinguisher 13:1
[*Leaf-CE6851HI-3-bd10-evpn] vpn-target 1:1
[*Leaf-CE6851HI-3-bd10-evpn] quit
[*Leaf-CE6851HI-3-bd10] quit
[*Leaf-CE6851HI-3-bd10] quit
```

在Leaf-CE6851HI-4上配置EVPN实例。

```
[~Leaf-CE6851HI-4] bridge-domain 10
[~Leaf-CE6851HI-4-bd10] evpn
[*Leaf-CE6851HI-4-bd10-evpn] route-distinguisher 14:1
[*Leaf-CE6851HI-4-bd10-evpn] vpn-target 1:1
[*Leaf-CE6851HI-4-bd10-evpn] quit
[*Leaf-CE6851HI-4-bd10] quit
[*Leaf-CE6851HI-4-bd10] quit
[*Leaf-CE6851HI-4] commit
```

步骤4 配置头端复制功能。

在Leaf-CE6851HI-3配置头端复制功能。

```
[~Leaf-CE6851HI-3] interface nve 1
[*Leaf-CE6851HI-3-Nve1] source 11.11.11.12
[*Leaf-CE6851HI-3-Nve1] mac-address 0000-5e00-0102
[*Leaf-CE6851HI-3-Nve1] vni 10010 head-end peer-list protocol bgp
[*Leaf-CE6851HI-3-Nve1] quit
[*Leaf-CE6851HI-3] commit
```

#在Leaf-CE6851HI-4配置头端复制功能。

```
[~Leaf-CE6851HI-4] interface nve 1
[*Leaf-CE6851HI-4-Nve1] source 11.11.11.12
[*Leaf-CE6851HI-4-Nve1] mac-address 0000-5e00-0102
[*Leaf-CE6851HI-4-Nve1] vni 10010 head-end peer-list protocol bgp
[*Leaf-CE6851HI-4-Nve1] quit
[*Leaf-CE6851HI-4] commit
```

----结束

3.1.1.3 配置 Leaf SVF 工作组

两台Leaf交换机Leaf-CE6851HI-5和Leaf-CE6851HI-6组成堆叠组,下连多台低成本CE交换机,以扩展接口接入数量。

配置思路

1. 两台CE6851HI作为父节点来组建堆叠组:配置堆叠和双主检测,并重启设备、连接线缆,使堆叠生效。

- 2. 配置SVF:配置父节点的Fabric-port,用于连接叶子节点交换机CE5810,并连接 父节点和叶子节点之间的SVF链路。
- 3. 配置IP地址:配置Leaf与Spi**n**e的三层互联接口地址、Loopback0地址(作为Router-ID和VTEP IP)。
- 4. 配置服务器接入:介绍业务服务器接入堆叠组时,堆叠组交换机上的配置。
- 5. 配置路由:在堆叠组上配置BGP动态路由,邻居为两个Spine设备,使SVF组上的网段与Spine上的网段三层可达。
- 6. 配置BGP EVPN:配置BGP EVPN作为VXLAN的控制平面、并配置BGP EVPN对等 体、EVPN实例以及头端复制功能。

组建堆叠组

步骤1 配置Leaf-CE6851HI-5的堆叠成员ID为1,优先级为150,Domain ID为10。

<HUAWEI> system-view

[~HUAWEI] sysname Leaf-CE6851HI-5

[*HUAWEI] commit

[~Leaf-CE6851HI-5] stack

[~Leaf-CE6851HI-5-stack] stack member 1 priority 150

[*Leaf-CE6851HI-5-stack] stack member 1 domain 10

[*Leaf-CE6851HI-5-stack] quit

[*Leaf-CE6851HI-5] commit

步骤2 配置Leaf-CE6851HI-6的Domain ID为10。

<HUAWEI> system-view

[~HUAWEI] sysname Leaf-CE6851HI-6

[*HUAWEI] commit

[~Leaf-CE6851HI-6] stack

[~Leaf-CE6851HI-6-stack] stack member 1 renumber 2 inherit-config

[*Leaf-CE6851HI-6-stack] stack member 1 domain 10

[*Leaf-CE6851HI-6-stack] quit

[*Leaf-CE6851HI-6] commit

□ 说明

组建SVF时要求父交换机的堆叠成员ID必须小于或等于4。

步骤3 配置堆叠端口。

将Leaf-CE6851HI-5的业务口40GE1/0/1~40GE1/0/2加入堆叠端口1/1。

[~Leaf-CE6851HI-5] interface stack-port 1/1

[*Leaf-CE6851HI-5-Stack-Port1/1] port member-group interface 40ge 1/0/1 to 1/0/2

Warning: After the configuration is complete,1.The interface(s) (40GE1/0/1-1/0/2) will be converted to stack mode and be configured with the port crc-statistics trigger error-down command if the configuration does not exist.2.The interface(s) may go Error-Down (crc-statistics) because there is no shutdown configuration on the interfaces.Continue? [Y/N]: y

[*Leaf-CE6851HI-5-Stack-Port1/1] quit

[*Leaf-CE6851HI-5] commit

[~Leaf-CE6851HI-5] quit

将Leaf-CE6851HI-6的业务口40GE1/0/1~40GE1/0/2加入堆叠端口1/1。

[~Leaf-CE6851HI-6] interface stack-port 1/1

[*Leaf-CE6851HI-6-Stack-Port1/1] port member-group interface 40ge 1/0/1 to 1/0/2

Warning: After the configuration is complete,1.The interface(s) (40GE1/0/1-1/0/2) will be converted to stack mode and be configured with the port crc-statistics trigger error-down command if the configuration does not exist.2.The interface(s) may go Error-Down (crc-statistics) because there is no shutdown configuration on the interfaces.Continue? [Y/N]: y

[*Leaf-CE6851HI-6-Stack-Port1/1] quit

[*Leaf-CE6851HI-6] commit

[~Leaf-CE6851HI-6] quit

步骤4 保存配置后重启设备。

保存Leaf-CE6851HI-5的配置,然后重启设备。Leaf-CE6851HI-6的操作与之类似,不再赘述。

<Leaf-CE6851HI-5> save

Warning: The current configuration will be written to the device. Continue? [Y/N]: y

<Leaf-CE6851HI-5> reboot

Warning: The system will reboot. Continue? [Y/N]:y

步骤5 连接堆叠线缆,建立堆叠。

步骤6 堆叠组建后执行命令save保存配置。

步骤7 堆叠建立后,配置堆叠组的双主检测,防止堆叠组分裂后网络中存在两个配置冲突的网络设备。

可采用业务口直连的方式配置双主检测,建议规划的直连链路至少2条,保证可靠性, 配置示例如下。

[~Leaf-CE6851HI-5] sysname Leaf-CE6851HI-5&CE6851HI-6

[*Leaf-CE6851HI-5] commit

[~Leaf-CE6851HI-5&CE6851HI-6] interface 10ge 1/0/30

[~Leaf-CE6851HI-5&CE6851HI-6-10GE1/0/30] description "for DAD"

[*Leaf-CE6851HI-5&CE6851HI-6-10GE1/0/30] dual-active detect mode direct

Warning: The interface will block common data packets, except BPDU packets. Continue? [Y/N]: y

[*Leaf-CE6851HI-5&CE6851HI-6-10GE1/0/30] quit

[*Leaf-CE6851HI-5&CE6851HI-6] interface 10ge 1/0/31

[*Leaf-CE6851HI-5&CE6851HI-6-10GE1/0/31] description "for DAD"

[*Leaf-CE6851HI-5&CE6851HI-6-10GE1/0/31] dual-active detect mode direct

Warning: The interface will block common data packets, except BPDU packets. Continue? [Y/N]: y

[*Leaf-CE6851HI-5&CE6851HI-6-10GE1/0/31] **quit**

[*Leaf-CE6851HI-5&CE6851HI-6] interface 10ge 2/0/30

[*Leaf-CE6851HI-5&CE6851HI-6-10GE2/0/30] description "for DAD"

[*Leaf-CE6851HI-5&CE6851HI-6-10GE2/0/30] dual-active detect mode direct

Warning: The interface will block common data packets, except BPDU packets. Continue? [Y/N]: y

[*Leaf-CE6851HI-5&CE6851HI-6-10GE2/0/30] quit

[*Leaf-CE6851HI-5&CE6851HI-6] interface 10ge 2/0/31

[*Leaf-CE6851HI-5&CE6851HI-6-10GE2/0/31] description "for DAD"

[*Leaf-CE6851HI-5&CE6851HI-6-10GE2/0/31] dual-active detect mode direct

Warning: The interface will block common data packets, except BPDU packets. Continue? [Y/N]: y

[*Leaf-CE6851HI-5&CE6851HI-6-10GE2/0/31] commit

[~Leaf-CE6851HI-5&CE6851HI-6-10GE2/0/31] quit

----结束

配置 SVF

步骤1 配置Fabric-port1,并加入成员端口10GE1/0/1和10GE2/0/1,绑定的叶子ID为101。

[~Leaf-CE6851HI-5&CE6851HI-6] interface fabric-port 1

[*Leaf-CE6851HI-5&CE6851HI-6-Fabric-Port1] port bind member 101

*Leaf-CE6851HI-5&CE6851HI-6-Fabric-Port1 port member-group interface 10ge 1/0/1 2/0/1

Warning: The interface(s) (10GE1/0/1-1/0/4,10GE2/0/1-2/0/4) will be converted to stack mode and have "port crc-statistics trigger error-down" configured. [Y/N]: y

[*Leaf-CE6851HI-5&CE6851HI-6-Fabric-Port1] quit

[*Leaf-CE6851HI-5&CE6851HI-6] commit

步骤2 配置Fabric-port2,并加入成员端口10GE1/0/2和10GE2/0/2,绑定的叶子ID为102。

[~Leaf-CE6851HI-5&CE6851HI-6] interface fabric-port 2

[*Leaf-CE6851HI-5&CE6851HI-6-Fabric-Port2] port bind member 102

[*Leaf-CE6851HI-5&CE6851HI-6-Fabric-Port2] port member-group interface 10ge 1/0/2 2/0/2

[*Leaf-CE6851HI-5&CE6851HI-6-Fabric-Port2] quit

[*Leaf-CE6851HI-5&CE6851HI-6] commit

步骤3 配置Fabric-port3,并加入成员端口10GE1/0/3和10GE2/0/3,绑定的叶子ID为103。

```
[~Leaf-CE6851HI-5&CE6851HI-6] interface fabric-port 3
[*Leaf-CE6851HI-5&CE6851HI-6-Fabric-Port3] port bind member 103
[*Leaf-CE6851HI-5&CE6851HI-6-Fabric-Port3] port member-group interface 10ge 1/0/3 2/0/3
[*Leaf-CE6851HI-5&CE6851HI-6-Fabric-Port3] commit
[~Leaf-CE6851HI-5&CE6851HI-6-Fabric-Port3] quit
```

步骤4 配置SVF系统转发模式为集中式,使叶子交换机的流量全部上送至父交换机进行转发。

```
[~Leaf-CE6851HI-5&CE6851HI-6] stack
[*Leaf-CE6851HI-5&CE6851HI-6-stack] stack forwarding-model centralized
[*Leaf-CE6851HI-5&CE6851HI-6-stack] commit
[*Leaf-CE6851HI-5&CE6851HI-6-stack] return
<Leaf-CE6851HI-5&CE6851HI-6> save
<Leaf-CE6851HI-5&CE6851HI-6> reboot //重启父交换机使转发模式生效
```

步骤5 连接叶子交换机与父交换机,并将叶子交换机上电。叶子交换机使用上行10GE端口与 父交换机相连。

此处默认叶子交换机为空配置文件启动且为自协商模式,可通过自协商加入SVF,无需任何配置。

如果叶子交换机有启动配置文件,则需要清空叶子交换机的下次启动配置文件后重启(自协商模式下),或者执行以下命令将叶子交换机强制配置为Leaf模式。

```
[~Leaf-CE6851HI-5&CE6851HI-6] stack
[~Leaf-CE6851HI-5&CE6851HI-6-stack] switch mode leaf member all //配置交换机的工作模式后,需要重启该设备配置才能生效
Warning: The switch will run in leaf mode after reboot. Continue? [Y/N]: y
[*Leaf-CE6851HI-5&CE6851HI-6-stack] commit
[*Leaf-CE6851HI-5&CE6851HI-6-stack] return
<*Leaf-CE6851HI-5&CE6851HI-6> save
<*Leaf-CE6851HI-5&CE6851HI-6> reboot
```

----结束

配置 IP 地址

步骤1 配置互连接口地址。

□ 说明

对于CE6855HI和CE7855EI,在切换三层接口之前,要先配置vlan reserved for maininterface startvlanid to endvlanid,配置三层主接口专用的保留VLAN。

```
[~Leaf-CE6851HI-5&CE6851HI-6] interface 40ge 1/0/3
[~Leaf-CE6851HI-5&CE6851HI-6-40GE1/0/3] description "to_Spine-CE12804-1-40GE1/0/6"
[*Leaf-CE6851HI-5&CE6851HI-6-40GE1/0/3] undo portswitch
*Leaf-CE6851HI-5&CE6851HI-6-40GE1/0/3] ip address 192.168.46.157 30
[*Leaf-CE6851HI-5&CE6851HI-6-40GE1/0/3] commit
-
[~Leaf-CE6851HI-5&CE6851HI-6-40GE1/0/3] quit
[~Leaf-CE6851HI-5&CE6851HI-6] interface 40ge 1/0/4
[~Leaf-CE6851HI-5&CE6851HI-6-40GE1/0/4] description "to Spine-CE12804-2-40GE1/0/6"
[*Leaf-CE6851HI-5&CE6851HI-6-40GE1/0/4] undo portswitch
[*Leaf-CE6851HI-5&CE6851HI-6-40GE1/0/4] ip address 192.168.46.165 30
[*Leaf-CE6851HI-5&CE6851HI-6-40GE1/0/4] commit
[~Leaf-CE6851HI-5&CE6851HI-6-40GE1/0/4] quit
[~Leaf-CE6851HI-5&CE6851HI-6] interface 40ge 2/0/3
[~Leaf-CE6851HI-5&CE6851HI-6-40GE2/0/3] description "to_Spine-CE12804-1-40GE1/0/7"
[*Leaf-CE6851HI-5&CE6851HI-6-40GE2/0/3] undo portswitch
*Leaf-CE6851HI-5&CE6851HI-6-40GE2/0/3] ip address 192.168.46.169 30
[*Leaf-CE6851HI-5&CE6851HI-6-40GE2/0/3] commit
[~Leaf-CE6851HI-5&CE6851HI-6-40GE2/0/3] quit
[~Leaf-CE6851HI-5&CE6851HI-6] interface 40ge 2/0/4
[~Leaf-CE6851HI-5&CE6851HI-6-40GE2/0/4] description "to_Spine-CE12804-2-40GE1/0/7"
[*Leaf-CE6851HI-5&CE6851HI-6-40GE2/0/4] undo portswitch
[*Leaf-CE6851HI-5&CE6851HI-6-40GE2/0/4] ip address 192.168.46.161 30
[*Leaf-CE6851HI-5&CE6851HI-6-40GE2/0/4] commit
[~Leaf-CE6851HI-5&CE6851HI-6-40GE2/0/4] quit
```

步骤2 配置Loopback口地址。

```
[~Leaf-CE6851HI-5&CE6851HI-6] interface loopback 0
[*Leaf-CE6851HI-5&CE6851HI-6-LoopBack0] description VTEP&Router-ID
[*Leaf-CE6851HI-5&CE6851HI-6-LoopBack0] ip address 11.11.11.17 32
[*Leaf-CE6851HI-5&CE6851HI-6-LoopBack0] commit
[~Leaf-CE6851HI-5&CE6851HI-6-LoopBack0] quit
```

----结束

配置服务器接入

介绍业务服务器接入SVF组时,在SVF交换机上的配置。

步骤1 配置服务器以负载均衡模式接入SVF。

```
[*Leaf-CE6851HI-5&CE6851HI-6] interface eth-trunk 1
[*Leaf-CE6851HI-5&CE6851HI-6-Eth-Trunk1] mode lacp-static
[*Leaf-CE6851HI-5&CE6851HI-6-Eth-Trunk1] port link-type trunk
[*Leaf-CE6851HI-5&CE6851HI-6-Eth-Trunk1] undo port trunk allow-pass vlan 1
[*Leaf-CE6851HI-5&CE6851HI-6-Eth-Trunk1] trunkport 10ge 101/0/1
[*Leaf-CE6851HI-5&CE6851HI-6-Eth-Trunk1] trunkport 10ge 101/0/2
[*Leaf-CE6851HI-5&CE6851HI-6-Eth-Trunk1] quit
[*Leaf-CE6851HI-5&CE6851HI-6] commit
```

步骤2 配置VXLAN业务接入点。

```
[~Leaf-CE6851HI-5&CE6851HI-6] bridge-domain 10
[*Leaf-CE6851HI-5&CE6851HI-6-bd10] vxlan vni 10010
[*Leaf-CE6851HI-5&CE6851HI-6-bd10] quit
[*Leaf-CE6851HI-5&CE6851HI-6] commit

[~Leaf-CE6851HI-5&CE6851HI-6] interface eth-trunk 1.1 mode l2
[*Leaf-CE6851HI-5&CE6851HI-6-Eth-Trunk1.1] encapsulation dot1q vid 30
[*Leaf-CE6851HI-5&CE6851HI-6-Eth-Trunk1.1] bridge-domain 10
[*Leaf-CE6851HI-5&CE6851HI-6-Eth-Trunk1.1] quit
[*Leaf-CE6851HI-5&CE6851HI-6] commit
```

----结束

配置路由

执行以下步骤配置Underlay的路由(本例以BGP为例,也可以使用OSPF)。

步骤1 在SVF组上配置BGP路由,对接Spine。

```
[~Leaf-CE6851HI-5&CE6851HI-6] bgp 65024
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] router-id 11.11.11.17
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] timer keepalive 10 hold 30
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] group Spine-CE12804 external
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] peer Spine-CE12804 as-number 65010
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] peer 192.168.40.158 as-number 65010
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] peer 192.168.40.158 group Spine-CE12804
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] peer 192.168.40.170 as-number 65010
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] peer 192.168.40.170 group Spine-CE12804
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] peer 192.168.40.166 as-number 65010
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] peer 192.168.40.166 group Spine-CE12804
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] peer 192.168.40.162 as-number 65010
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] peer 192.168.40.162 group Spine-CE12804
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] ipv4-family unicast
[*Leaf-CE6851HI-5&CE6851HI-6-bgp-af-ipv4] preference 20 200 10
[*Leaf-CE6851HI-5&CE6851HI-6-bgp-af-ipv4] network 11.11.11.17 255.255.255.255
[*Leaf-CE6851HI-5&CE6851HI-6-bgp-af-ipv4] maximum load-balancing 32
[*Leaf-CE6851HI-5&CE6851HI-6-bgp-af-ipv4] quit
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] quit
[*Leaf-CE6851HI-5&CE6851HI-6] commit
```

----结束

配置 BGP EVPN

使用BGP EVPN作为VXLAN的控制平面时,需执行以下步骤将Leaf作为BGP路由反射器客户端,Spine作为BGP路由器反射器,在Leaf与Spine之间建立IBGP EVPN对等体关系,配置EVPN实例和头端复制功能。

步骤1 在SVF上使能EVPN作为VXLAN的控制平面功能。

```
[~Leaf-CE6851HI-5&CE6851HI-6] evpn-overlay enable
[*Leaf-CE6851HI-5&CE6851HI-6] commit
```

步骤2 配置堆叠组与Spine节点分别建立IBGP EVPN对等体关系,Spine作为路由反射器RR。

```
[~Leaf-CE6851HI-5&CE6851HI-6] bgp 100 instance evpn1
[*Leaf-CE6851HI-5&CE6851HI-6-bgp-instance-evpn1] router-id 11.11.11.17
[*Leaf-CE6851HI-5&CE6851HI-6-bgp-instance-evpn1] peer 11.11.11.14 as-number 100
//与网关Spine-CE12804-1建立BGP EVPN对等体关系
[*Leaf-CE6851HI-5&CE6851HI-6-bgp-instance-evpn1] peer 11.11.11.14 connect-interface loopback 0
[*Leaf-CE6851HI-5&CE6851HI-6-bgp-instance-evpn1] peer 11.11.11.15 as-number 100
//与网关Spine-CE12804-2建立BGP EVPN对等体关系
[*Leaf-CE6851HI-5&CE6851HI-6-bgp-instance-evpn1] peer 11.11.11.15 connect-interface loopback 0
[*Leaf-CE6851HI-5&CE6851HI-6-bgp-instance-evpn1] l2vpn-family evpn
[*Leaf-CE6851HI-5&CE6851HI-6-bgp-instance-evpn1-af-evpn] peer 11.11.11.14 enable
[*Leaf-CE6851HI-5&CE6851HI-6-bgp-instance-evpn1-af-evpn] peer 11.11.11.14 advertise arp
[*Leaf-CE6851HI-5&CE6851HI-6-bgp-instance-evpn1-af-evpn] peer 11.11.11.15 enable
[*Leaf-CE6851HI-5&CE6851HI-6-bgp-instance-evpn1-af-evpn] peer 11.11.11.15 advertise arp
[*Leaf-CE6851HI-5&CE6851HI-6-bgp-instance-evpn1-af-evpn] quit
[*Leaf-CE6851HI-5&CE6851HI-6-bgp-instance-evpn1] quit
[*Leaf-CE6851HI-5&CE6851HI-6] commit
```

步骤3 在SVF上配置EVPN实例。

```
[~Leaf-CE6851HI-5&CE6851HI-6] bridge-domain 10
[~Leaf-CE6851HI-5&CE6851HI-6-bd10] evpn
[*Leaf-CE6851HI-5&CE6851HI-6-bd10-evpn] route-distinguisher 15:1
[*Leaf-CE6851HI-5&CE6851HI-6-bd10-evpn] vpn-target 1:1
[*Leaf-CE6851HI-5&CE6851HI-6-bd10-evpn] quit
[*Leaf-CE6851HI-5&CE6851HI-6-bd10] quit
[*Leaf-CE6851HI-5&CE6851HI-6] commit
```

步骤4 在SVF上配置头端复制功能。

```
[~Leaf-CE6851HI-5&CE6851HI-6] interface nve 1

[*Leaf-CE6851HI-5&CE6851HI-6-Nve1] source 11.11.11.17

[*Leaf-CE6851HI-5&CE6851HI-6-Nve1] vni 10010 head-end peer-list protocol bgp

[*Leaf-CE6851HI-5&CE6851HI-6-Nve1] quit

[*Leaf-CE6851HI-5&CE6851HI-6] commit
```

----结束

3.1.1.4 配置 Spine 节点

在两台Spine节点上配置上、下联的互联接口地址和路由,使Underlay网络三层互通。

配置思路

- 1. 配置IP地址:分别配置Spine节点与Leaf和GW设备的三层互联地址、Loopback地址(作为Router-ID)。
- 2. 配置路由:在两台Spine节点上分别配置BGP动态路由,邻居为Leaf堆叠组、LeafM-LAG组两台设备、GW M-LAG两台设备,使Spine上的网段与Leaf和GW上的网段三层可达。
- 配置BGP EVPN: 配置BGP EVPN作为VXLAN的控制平面、并配置BGP EVPN对等 体。将Spine节点作为BGP路由反射器,Leaf/网关作为反射器客户端。

配置 IP 地址

步骤1 配置互连接口IP地址。

在Spine-CE12804-1上配置接口IP地址。

```
[~HUAWEI] sysname Spine-CE12804-1
[*HUAWEI] commit
[~Spine-CE12804-1] interface 40ge 1/0/0
[~Spine-CE12804-1-40GE1/0/0] undo portswitch
[*Spine-CE12804-1-40GE1/0/0] ip address 192.168.40.158 30
[*Spine-CE12804-1-40GE1/0/0] commit
[~Spine-CE12804-1-40GE1/0/0] quit
[~Spine-CE12804-1] interface 40ge 1/0/1
[~Spine-CE12804-1-40GE1/0/1] undo portswitch
[*Spine-CE12804-1-40GE1/0/1] ip address 192.168.40.170 30
[*Spine-CE12804-1-40GE1/0/1] commit
[~Spine-CE12804-1-40GE1/0/1] quit
[~Spine-CE12804-1] interface 40ge 1/0/2
[~Spine-CE12804-1-40GE1/0/2] description "to-Leaf-CE6851-3-40GE1/0/3"
[*Spine-CE12804-1-40GE1/0/2] undo portswitch
[*Spine-CE12804-1-40GE1/0/2] ip address 192.168.41.158 30
[*Spine-CE12804-1-40GE1/0/2] commit
[~Spine-CE12804-1-40GE1/0/2] quit
[~Spine-CE12804-1] interface 40ge 1/0/3
[~Spine-CE12804-1-40GE1/0/3] description "to-Leaf-CE6851-4-40GE1/0/4"
[*Spine-CE12804-1-40GE1/0/3] undo portswitch
[*Spine-CE12804-1-40GE1/0/3] ip address 192.168.41.170 30
[*Spine-CE12804-1-40GE1/0/3] commit
[~Spine-CE12804-1-40GE1/0/3] quit
[~Spine-CE12804-1] interface 40ge 1/0/4
[~Spine-CE12804-1-40GE1/0/4] undo portswitch
[*Spine-CE12804-1-40GE1/0/4] ip address 192.168.42.157 30
[*Spine-CE12804-1-40GE1/0/4] commit
[~Spine-CE12804-1-40GE1/0/4] quit
[~Spine-CE12804-1] interface 40ge 1/0/5
[~Spine-CE12804-1-40GE1/0/5] undo portswitch
[*Spine-CE12804-1-40GE1/0/5] ip address 192.168.42.161 30
[*Spine-CE12804-1-40GE1/0/5] commit
[~Spine-CE12804-1-40GE1/0/5] quit
[~Spine-CE12804-1] interface 40ge 1/0/6
[~Spine-CE12804-1-40GE1/0/6] undo portswitch
[*Spine-CE12804-1-40GE1/0/6] ip address 192.168.46.158 30
[*Spine-CE12804-1-40GE1/0/6] commit
[~Spine-CE12804-1-40GE1/0/6] quit
[~Spine-CE12804-1] interface 40ge 1/0/7
[~Spine-CE12804-1-40GE1/0/7] undo portswitch
[*Spine-CE12804-1-40GE1/0/7] ip address 192.168.46.170 30
[*Spine-CE12804-1-40GE1/0/7] commit
[~Spine-CE12804-1-40GE1/0/7] quit
```

#在Spine-CE12804-2节点上配置接口IP地址。

```
[~HUAWEI] sysname Spine-CE12804-2
[*HUAWEI] commit
[~Spine-CE12804-2] interface 40ge 1/0/0
[~Spine-CE12804-2-40GE1/0/0] undo portswitch
[*Spine-CE12804-2-40GE1/0/0] ip address 192.168.40.162 30
[*Spine-CE12804-2-40GE1/0/0] commit
[~Spine-CE12804-2-40GE1/0/0] quit
[~Spine-CE12804-2] interface 40ge 1/0/1
[~Spine-CE12804-2-40GE1/0/1] undo portswitch
[*Spine-CE12804-2-40GE1/0/1] ip address 192.168.40.166 30
[*Spine-CE12804-2-40GE1/0/1] commit
[~Spine-CE12804-2-40GE1/0/1] quit
[~Spine-CE12804-2] interface 40ge 1/0/2
[~Spine-CE12804-2-40GE1/0/2] description "to-Leaf-CE6851-4-40GE1/0/3"
[*Spine-CE12804-2-40GE1/0/2] undo portswitch
[*Spine-CE12804-2-40GE1/0/2] ip address 192.168.41.162 30
```

```
[*Spine-CE12804-2-40GE1/0/2] commit
[~Spine-CE12804-2-40GE1/0/2] quit
[~Spine-CE12804-2] interface 40ge 1/0/3
[~Spine-CE12804-2-40GE1/0/3] description "to-Leaf-CE6851-3-40GE1/0/4"
[*Spine-CE12804-2-40GE1/0/3] undo portswitch
[*Spine-CE12804-2-40GE1/0/3] ip address 192.168.41.166 30
[*Spine-CE12804-2-40GE1/0/3] commit
[~Spine-CE12804-2-40GE1/0/3] quit
[~Spine-CE12804-2] interface 40ge 1/0/4
[~Spine-CE12804-2-40GE1/0/4] undo portswitch
[*Spine-CE12804-2-40GE1/0/4] ip address 192.168.43.157 30 [*Spine-CE12804-2-40GE1/0/4] commit
[~Spine-CE12804-2-40GE1/0/4] quit
[~Spine-CE12804-2] interface 40ge 1/0/5
[~Spine-CE12804-2-40GE1/0/5] undo portswitch
[*Spine-CE12804-2-40GE1/0/5] ip address 192.168.43.161 30
[*Spine-CE12804-2-40GE1/0/5] commit
[~Spine-CE12804-2-40GE1/0/5] quit
[~Spine-CE12804-2] interface 40ge 1/0/6
[~Spine-CE12804-2-40GE1/0/6] undo portswitch
[*Spine-CE12804-2-40GE1/0/6] ip address 192.168.46.166 30
[*Spine-CE12804-2-40GE1/0/6] commit
[~Spine-CE12804-2-40GE1/0/6] quit
[~Spine-CE12804-2] interface 40ge 1/0/7
[~Spine-CE12804-2-40GE1/0/7] undo portswitch
[*Spine-CE12804-2-40GE1/0/7] ip address 192.168.46.162 30
[*Spine-CE12804-2-40GE1/0/7] commit
[~Spine-CE12804-2-40GE1/0/7] quit
```

步骤2 配置Loopback地址。

```
[~Spine-CE12804-1] interface loopback 0
[*Spine-CE12804-1-LoopBack0] ip address 11.11.11.14 32
[*Spine-CE12804-1-LoopBack0] commit
[~Spine-CE12804-1-LoopBack0] quit

[~Spine-CE12804-2] interface loopback 0
[*Spine-CE12804-2-LoopBack0] ip address 11.11.11.15 32
[*Spine-CE12804-2-LoopBack0] commit
[~Spine-CE12804-2-LoopBack0] quit
```

----结束

配置路由

执行以下配置配置Underlay的路由(本例以BGP为例,也可以使用OSPF)。

步骤1 在Spine-CE12804-1上配置BGP路由。

```
[~Spine-CE12804-1] bgp 65010
[*Spine-CE12804-1-bgp] router-id 11.11.11.14
[*Spine-CE12804-1-bgp] timer keepalive 10 hold 30
[*Spine-CE12804-1-bgp] group Leaf-CE6851HI-1&CE6851HI-2 external
[*Spine-CE12804-1-bgp] peer Leaf-CE6851HI-1&CE6851HI-2 as-number 65021
[*Spine-CE12804-1-bgp] peer 192.168.40.157 as-number 65021
[*Spine-CE12804-1-bgp] peer 192.168.40.157 group Leaf-CE6851HI-1&CE6851HI-2
[*Spine-CE12804-1-bgp] peer 192.168.40.169 as-number 65021
[*Spine-CE12804-1-bgp] peer 192.168.40.169 group Leaf-CE6851HI-1&CE6851HI-2
[*Spine-CE12804-1-bgp] group Leaf-CE6851HI-3&4 external
[*Spine-CE12804-1-bgp] peer Leaf-CE6851HI-3&4 as-number 65022
[*Spine-CE12804-1-bgp] peer 192.168.41.157 as-number 65022
[*Spine-CE12804-1-bqp] peer 192.168.41.157 group Leaf-CE6851HI-3&4
[*Spine-CE12804-1-bgp] peer 192.168.41.169 as-number 65022
[*Spine-CE12804-1-bgp] peer 192.168.41.169 group Leaf-CE6851HI-3&4
[*Spine-CE12804-1-bgp] group Leaf-CE6851HI-5&CE6851HI-6 external
[*Spine-CE12804-1-bgp] peer Leaf-CE6851HI-5&CE6851HI-6 as-number 65024
[*Spine-CE12804-1-bgp] peer 192.168.46.157 as-number 65024
[*Spine-CE12804-1-bgp] peer 192.168.46.157 group Leaf-CE6851HI-5&CE6851HI-6
```

```
[*Spine-CE12804-1-bgp] peer 192.168.46.169 as-number 65024
[*Spine-CE12804-1-bgp] peer 192.168.46.169 group Leaf-CE6851HI-5&CE6851HI-6
[*Spine-CE12804-1-bgp] group Gateway-CE12808 external
[*Spine-CE12804-1-bgp] peer Gateway-CE12808 as-number 65000
[*Spine-CE12804-1-bgp] peer 192.168.42.158 as-number 65000
[*Spine-CE12804-1-bgp] peer 192.168.42.158 group Gateway-CE12808
[*Spine-CE12804-1-bgp] peer 192.168.42.162 as-number 65000
[*Spine-CE12804-1-bgp] peer 192.168.42.162 group Gateway-CE12808

[*Spine-CE12804-1-bgp] ipv4-family unicast
[*Spine-CE12804-1-bgp-af-ipv4] preference 20 200 10
[*Spine-CE12804-1-bgp-af-ipv4] network 11.11.11.14 255.255.255.255
[*Spine-CE12804-1-bgp-af-ipv4] maximum load-balancing 32
[*Spine-CE12804-1-bgp-af-ipv4] quit
[*Spine-CE12804-1-bgp] quit
[*Spine-CE12804-1] commit
```

步骤2 在Spine-CE12804-2上配置BGP路由。

```
[~Spine-CE12804-2] bgp 65010
[*Spine-CE12804-2-bgp] router-id 11.11.11.15
[*Spine-CE12804-2-bgp] timer keepalive 10 hold 30
[*Spine-CE12804-2-bgp] group Leaf-CE6851HI-1&CE6851HI-2 external
[*Spine-CE12804-2-bgp] peer Leaf-CE6851HI-1&CE6851HI-2 as-number 65021
[*Spine-CE12804-2-bgp] peer 192.168.40.165 as-number 65021
[*Spine-CE12804-2-bgp] peer 192.168.40.165 group Leaf-CE6851HI-1&CE6851HI-2
[*Spine-CE12804-2-bgp] peer 192.168.40.161 as-number 65021
[*Spine-CE12804-2-bgp] peer 192.168.40.161 group Leaf-CE6851HI-1&CE6851HI-2
[*Spine-CE12804-2-bgp] group Leaf-CE6851HI-3&4 external
[*Spine-CE12804-2-bgp] peer Leaf-CE6851HI-3&4 as-number 65022
[*Spine-CE12804-2-bgp] peer 192.168.41.165 as-number 65022
[*Spine-CE12804-2-bgp] peer 192.168.41.165 group Leaf-CE6851HI-3&4
[*Spine-CE12804-2-bgp] peer 192.168.41.161 as-number 65022
[*Spine-CE12804-2-bgp] peer 192.168.41.161 group Leaf-CE6851HI-3&4
[*Spine-CE12804-2-bgp] group Gateway-CE12808 external
[*Spine-CE12804-2-bgp] peer Gateway-CE12808 as-number 65000
[*Spine-CE12804-2-bgp] peer 192.168.43.162 as-number 65000
[*Spine-CE12804-2-bgp] peer 192.168.43.162 group Gateway-CE12808
[*Spine-CE12804-2-bgp] peer 192.168.43.158 as-number 65000
[*Spine-CE12804-2-bgp] peer 192.168.43.158 group Gateway-CE12808
[*Spine-CE12804-2-bgp] group Leaf-CE6851HI-5&CE6851HI-6 external
[*Spine-CE12804-2-bgp] peer Leaf-CE6851HI-5&CE6851HI-6 as-number 65024
[*Spine-CE12804-2-bgp] peer 192.168.46.165 as-number 65024
[*Spine-CE12804-2-bgp] peer 192.168.46.165 group Leaf-CE6851HI-5&CE6851HI-6
[*Spine-CE12804-2-bgp] peer 192.168.46.161 as-number 65024
[*Spine-CE12804-2-bgp] peer 192.168.46.161 group Leaf-CE6851HI-5&CE6851HI-6
[*Spine-CE12804-2-bgp] ipv4-family unicast
[*Spine-CE12804-2-bgp-af-ipv4] preference 20 200 10
[*Spine-CE12804-2-bgp-af-ipv4] network 11.11.11.15 255.255.255.255
[*Spine-CE12804-2-bgp-af-ipv4] maximum load-balancing 32
[*Spine-CE12804-2-bgp-af-ipv4] quit
[*Spine-CE12804-2-bgp] quit
[*Spine-CE12804-2] commit
```

----结束

配置 BGP-EVPN

使用BGP EVPN作为VXLAN的控制平面时,需执行以下步骤将Leaf/网关作为BGP路由反射器客户端,Spine作为BGP路由器反射器,在Leaf/网关与Spine之间建立IBGP EVPN邻居。

步骤1 在Spine-CE12804-1和Spine-CE12804-2上使能EVPN作为VXLAN的控制平面功能。

```
[~Spine-CE12804-1] evpn-overlay enable [*Spine-CE12804-1] commit
```

```
[~Spine-CE12804-2] evpn-overlay enable
[*Spine-CE12804-2] commit
```

步骤2 配置Spine-CE12804-1和Spine-CE12804-2上分别与Leaf/网关节点建立BGP EVPN对等体关系,Spine作为路由反射器RR。

在Spine-CE12804-1上配置与Leaf/网关建立IBGP EVPN对等体关系,Spine-CE12804-1作为路由反射器,Leaf/网关作为反射器客户端。

```
[~Spine-CE12804-1] bgp 100 instance evpn1
[*Spine-CE12804-1-bgp-instance-evpn1] router-id 11.11.11.14
[*Spine-CE12804-1-bgp-instance-evpn1] peer 11.11.11.11 as-number 100
//与堆叠组建立BGP EVPN对等体关系
[*Spine-CE12804-1-bgp-instance-evpn1] peer 11.11.11 connect-interface loopback 0
[*Spine-CE12804-1-bgp-instance-evpn1] peer 13.13.13.13 as-number 100
//与Leaf-CE6851HI-3建立BGP EVPN对等体关系
[*Spine-CE12804-1-bqp-instance-evpn1] peer 13.13.13 connect-interface loopback 0
[*Spine-CE12804-1-bgp-instance-evpn1] peer 14.14.14.14 as-number 100
//与Leaf-CE6851HI-4建立BGP EVPN对等体关系
[*Spine-CE12804-1-bgp-instance-evpn1] peer 14.14.14 connect-interface loopback 0
[*Spine-CE12804-1-bgp-instance-evpn1] peer 11.11.11.17 as-number 100
//与SVF组建立BGP EVPN对等体关系
[*Spine-CE12804-1-bgp-instance-evpn1] peer 11.11.11.17 connect-interface loopback 0
[*Spine-CE12804-1-bgp-instance-evpn1] peer 18.18.18 as-number 100
//与Gateway-CE12808-1建立BGP EVPN对等体关系
[*Spine-CE12804-1-bgp-instance-evpn1] peer 18.18.18 connect-interface loopback 0
[*Spine-CE12804-1-bgp-instance-evpn1] peer 19.19.19.19 as-number 100
//与Gateway-CE12808-2建立BGP EVPN对等体关系
[*Spine-CE12804-1-bgp-instance-evpn1] peer 19.19.19.19 connect-interface loopback 0 [*Spine-CE12804-1-bgp-instance-evpn1] l2vpn-family evpn
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] undo policy vpn-target
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 11.11.11.11 enable
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 11.11.11.11 reflect-client
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 11.11.11.11 next-hop-invariable
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 11.11.11.11 advertise arp
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 13.13.13.13 enable
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 13.13.13.13 reflect-client
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 13.13.13.13 next-hop-invariable
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 13.13.13.13 advertise arp
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 14.14.14.14 enable
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 14.14.14.14 reflect-client
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 14.14.14.14 next-hop-invariable
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 14.14.14.14 advertise arp
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 11.11.11.17 enable
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 11.11.11.17 reflect-client
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 11.11.11.17 next-hop-invariable
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 11.11.11.17 advertise arp
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 18.18.18.18 enable
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 18.18.18.18 reflect-client
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 18.18.18.18 next-hop-invariable
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 18.18.18.18 advertise arp
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 19.19.19.19 enable
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 19.19.19.19 reflect-client
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 19.19.19.19 next-hop-invariable
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 19.19.19.19 advertise arp
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] quit
[*Spine-CE12804-1-bgp-instance-evpn1] quit
[*Spine-CE12804-1] commit
```

在Spine-CE12804-2上配置与Leaf建立IBGP EVPN对等体关系,Spine-CE12804-2作为路由反射器,Leaf作为反射器客户端。

```
[~Spine-CE12804-2] bgp 100 instance evpn1
[*Spine-CE12804-2-bgp-instance-evpn1] router-id 11.11.11.15
[*Spine-CE12804-2-bgp-instance-evpn1] peer 11.11.11.11 as-number 100
//与堆叠组建立BGP EVPN对等体关系
[*Spine-CE12804-2-bgp-instance-evpn1] peer 11.11.11.11 connect-interface loopback 0
[*Spine-CE12804-2-bgp-instance-evpn1] peer 13.13.13.13 as-number 100
//与Leaf-CE6851HI-3建立BGP EVPN对等体关系
```

```
[*Spine-CE12804-2-bgp-instance-evpn1] peer 13.13.13 connect-interface loopback 0
[*Spine-CE12804-2-bgp-instance-evpn1] peer 14.14.14 as-number 100
//与Leaf-CE6851HI-4建立BGP EVPN对等体关系
[*Spine-CE12804-2-bgp-instance-evpn1] peer 14.14.14.14 connect-interface loopback 0
[*Spine-CE12804-2-bgp-instance-evpn1] peer 11.11.11.17 as-number 100
//与SVF组建立BGP EVPN对等体关系
[*Spine-CE12804-2-bgp-instance-evpn1] peer 11.11.11.17 connect-interface loopback 0
[*Spine-CE12804-2-bgp-instance-evpn1] peer 18.18.18.18 as-number 100
-
//与Gateway-CE12808-1建立BGP EVPN对等体关系
[*Spine-CE12804-2-bgp-instance-evpn1] peer 18.18.18.connect-interface loopback 0
[*Spine-CE12804-2-bgp-instance-evpn1] peer 19.19.19.19 as-number 100
//与Gateway-CE12808-2建立BGP EVPN对等体关系
[*Spine-CE12804-2-bqp-instance-evpn1] peer 19.19.19 connect-interface loopback 0
[*Spine-CE12804-2-bgp-instance-evpn1] l2vpn-family evpn
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] undo policy vpn-target
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 11.11.11.11 enable
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 11.11.11.11 reflect-client
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 11.11.11.11 next-hop-invariable
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 11.11.11.11 advertise arp
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 13.13.13.13 enable
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 13.13.13.13 reflect-client
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 13.13.13.13 next-hop-invariable
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 13.13.13.13 advertise arp
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 14.14.14.14 enable
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 14.14.14.14 reflect-client
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 14.14.14.14 next-hop-invariable
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 14.14.14.14 advertise arp
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 11.11.11.17 enable
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 11.11.11.17 reflect-client
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 11.11.11.17 next-hop-invariable
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 11.11.11.17 advertise arp
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 18.18.18.18 enable
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 18.18.18.18 reflect-client
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 18.18.18.18 next-hop-invariable
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 18.18.18.18 advertise arp
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 19.19.19.19 enable
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 19.19.19.19 reflect-client
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 19.19.19.19 next-hop-invariable
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 19.19.19.19 advertise arp
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] quit
[*Spine-CE12804-2-bgp-instance-evpn1] quit
[*Spine-CE12804-2] commit
```

----结束

3.1.1.5 配置网关工作组

VXLAN集中式网关由两台CE交换机Gateway-CE12808-1和Gateway-CE12808-2承担。可以是堆叠模式也可以是M-LAG。本节以M-LAG形式为例进行介绍。

配置思路

- 1. 配置IP地址:配置与Spine互连的接口IP地址,配置与防火墙互连的管理VLAN地址,配置Loopback0地址(作为VTEP地址)、Loopback1地址(作为Router-ID)。
- 2. 配置M-LAG:在GW交换机上配置M-LAG全局模式、DFS组、Peer-Link,并分别配置与两台防火墙的互联管理VLAN和业务链路。
- 3. 配置路由:在GW上分别配置BGP动态路由,邻居为两个Spine设备,使GW上的网段与Spine上的网段三层可达;配置与外部路由器的路由。
- 4. 配置BGP EVPN:配置BGP EVPN作为VXLAN的控制平面、并配置BGP EVPN对等体、EVPN实例以及头端复制功能、三层网关。
- 5. 配置MAC漂移白名单:在GW设备上将Spine设备与GW互联的三层口MAC地址配置为白名单,不再检测MAC漂移。

6. 配置网关设备与外部路由器的互联,使网关与外部路由器形成"口"字型和交叉结合的出口网络:分别配置网关与外部路由器之间的互联接口地址、配置网关设备之间的互联接口地址、配置网关设备上的默认路由。

配置 IP 地址

步骤1 配置接口IP地址。

#配置Gateway-CE12808-1的接口IP地址。

```
[~Gateway-CE12808-1] interface 40ge 1/0/0 //连接Spine节点
[~Gateway-CE12808-1-40GE1/0/0] undo portswitch
[*Gateway-CE12808-1-40GE1/0/0] ip address 192.168.42.158 30
[*Gateway-CE12808-1-40GE1/0/0] commit
[~Gateway-CE12808-1-40GE1/0/0] quit
[~Gateway-CE12808-1] interface 40ge 1/0/1 //连接Spine节点
[~Gateway-CE12808-1-40GE1/0/1] undo portswitch
[*Gateway-CE12808-1-40GE1/0/1] ip address 192.168.43.162 30
[*Gateway-CE12808-1-40GE1/0/1] commit
[~Gateway-CE12808-1-40GE1/0/1] quit
[~Gateway-CE12808-1] vlan batch 11
[*Gateway-CE12808-1] interface vlanif 11 //网关设备与防火墙的Public互联VLAN [*Gateway-CE12808-1-vlanif11] description "to firewall-1~2"
[*Gateway-CE12808-1-vlanif11] ip address 192.168.45.154 29
[*Gateway-CE12808-1-vlanif11] vrrp vrid 1 virtual-ip 192.168.45.153
[*Gateway-CE12808-1-vlanif11] commit
[~Gateway-CE12808-1-vlanif11] quit
[~Gateway-CE12808-1] vlan batch 12
[*Gateway-CE12808-1] interface vlanif 12 //网关设备与防火墙的业务互联VLAN [*Gateway-CE12808-1-vlanif12] description "to firewall-1~2"
[*Gateway-CE12808-1-vlanif12] ip address 192.168.45.162 29
[*Gateway-CE12808-1-vlanif12] vrrp vrid 2 virtual-ip 192.168.45.161
[*Gateway-CE12808-1-vlanif12] commit
[~Gateway-CE12808-1-vlanif12] quit
```

#配置Gateway-CE12808-2的接口IP地址。

```
[~Gateway-CE12808-2] interface 40ge 1/0/0 //连接Spine节点
[~Gateway-CE12808-2-40GE1/0/0] undo portswitch
[*Gateway-CE12808-2-40GE1/0/0] ip address 192.168.42.162 30
[*Gateway-CE12808-2-40GE1/0/0] commit
[~Gateway-CE12808-2-40GE1/0/0] quit
[~Gateway-CE12808-2] interface 40ge 1/0/1 //连接Spine节点
[~Gateway-CE12808-2-40GE1/0/1] undo portswitch
[*Gateway-CE12808-2-40GE1/0/1] ip address 192.168.43.158 30
[*Gateway-CE12808-2-40GE1/0/1] commit
[~Gateway-CE12808-2-40GE1/0/1] quit
[~Gateway-CE12808-2] vlan batch 11
[*Gateway-CE12808-2] interface vlanif 11 //网关设备与防火墙的Public互联VLAN
[*Gateway-CE12808-2-Vlanif11] description "to firewall-1-2"
[*Gateway-CE12808-2-Vlanif11] ip address 192.168.45.155 29
[*Gateway-CE12808-2-Vlanif11] vrrp vrid 1 virtual-ip 192.168.45.153
[*Gateway-CE12808-2-Vlanif11] commit
[~Gateway-CE12808-2-Vlanif11] quit
[~Gateway-CE12808-2] vlan batch 12
[*Gateway-CE12808-2] interface vlanif 12 //网关设备与防火墙的业务互联VLAN
[*Gateway-CE12808-2-Vlanif12] description "to firewall-1-2"
[*Gateway-CE12808-2-Vlanif12] ip address 192.168.45.163 29
[*Gateway-CE12808-2-Vlanif12] vrrp vrid 2 virtual-ip 192.168.45.161
[*Gateway-CE12808-2-Vlanif12] commit
[~Gateway-CE12808-2-Vlanif12] quit
```

步骤2 配置Loopback口地址。

```
[~Gateway-CE12808-1] interface loopback 0 //作为VTEP地址
[*Gateway-CE12808-1-LoopBack0] ip address 11.11.11.16 32
[*Gateway-CE12808-1-LoopBack0] commit
```

```
[~Gateway-CE12808-1-LoopBack0] quit
[~Gateway-CE12808-1] interface loopback 1
[*Gateway-CE12808-1-LoopBack1] ip address 18.18.18.18 32
[*Gateway-CE12808-1-LoopBack1] commit
[~Gateway-CE12808-1-LoopBack1] quit
[~Gateway-CE12808-1] interface loopback 2
[*Gateway-CE12808-1-LoopBack2] ip address 21.21.21.21 32
[*Gateway-CE12808-1-LoopBack2] commit
[~Gateway-CE12808-1-LoopBack2] quit
[~Gateway-CE12808-2] interface loopback 0 //作为VTEP地址
[*Gateway-CE12808-2-LoopBack0] ip address 11.11.11.16 32
[*Gateway-CE12808-2-LoopBack0] commit
[~Gateway-CE12808-2-LoopBack0] quit
[~Gateway-CE12808-2] interface loopback 1
[*Gateway-CE12808-2-LoopBack1] ip address 19.19.19.19 32
[*Gateway-CE12808-2-LoopBack1] commit
[~Gateway-CE12808-2-LoopBack1] quit
[~Gateway-CE12808-2] interface loopback 2
[*Gateway-CE12808-2-LoopBack2] ip address 22.22.22.22 32
[*Gateway-CE12808-2-LoopBack2] commit
[~Gateway-CE12808-2-LoopBack2] quit
```

----结束

配置路由

执行以下步骤配置Underlay的路由(本例以BGP为例,也可以使用OSPF)。

步骤1 在Gateway-CE12808-1上配置BGP路由用于打通Underlay层面的路由。

```
[~Gateway-CE12808-1] bgp 65000
[*Gateway-CE12808-1-bgp] router-id 18.18.18.18
[*Gateway-CE12808-1-bgp] timer keepalive 10 hold 30
[*Gateway-CE12808-1-bgp] group Spine-CE12804 external //对接Spine
[*Gateway-CE12808-1-bgp] peer Spine-CE12804 as-number 65010
[*Gateway-CE12808-1-bgp] peer 192.168.42.157 as-number 65010
[*Gateway-CE12808-1-bgp] peer 192.168.42.157 group Spine-CE12804
[*Gateway-CE12808-1-bgp] peer 192.168.43.161 as-number 65010
[*Gateway-CE12808-1-bgp] peer 192.168.43.161 group Spine-CE12804
[*Gateway-CE12808-1-bgp] ipv4-family unicast
[*Gateway-CE12808-1-bgp-af-ipv4] preference 20 200 10
[*Gateway-CE12808-1-bgp-af-ipv4] network 11.11.11.16 255.255.255.255
[*Gateway-CE12808-1-bgp-af-ipv4] network 18.18.18.18 255.255.255.255
[*Gateway-CE12808-1-bgp-af-ipv4] network 192.168.45.152 255.255.255.248
[*Gateway-CE12808-1-bgp-af-ipv4] maximum load-balancing 32
[*Gateway-CE12808-1-bgp-af-ipv4] quit
[*Gateway-CE12808-1-bgp] quit
[*Gateway-CE12808-1] commit
```

步骤2 在Gateway-CE12808-2上配置BGP路由用于打通Underlay层面的路由。

```
[~Gateway-CE12808-2] bgp 65000
[*Gateway-CE12808-2-bgp] router-id 19.19.19.19
[*Gateway-CE12808-2-bgp] timer keepalive 10 hold 30
[*Gateway-CE12808-2-bgp] group Spine-CE12804 external //对接Spine
[*Gateway-CE12808-2-bgp] peer Spine-CE12804 as-number 65010
[*Gateway-CE12808-2-bgp] peer 192.168.42.161 as-number 65010
[*Gateway-CE12808-2-bgp] peer 192.168.42.161 group Spine-CE12804
[*Gateway-CE12808-2-bgp] peer 192.168.43.157 as-number 65010
[*Gateway-CE12808-2-bgp] peer 192.168.43.157 group Spine-CE12804

[*Gateway-CE12808-2-bgp] ipv4-family unicast
[*Gateway-CE12808-2-bgp af-ipv4] preference 20 200 10
[*Gateway-CE12808-2-bgp-af-ipv4] network 11.11.11.16 255.255.255.255
[*Gateway-CE12808-2-bgp-af-ipv4] network 19.19.19.19 255.255.255.255
[*Gateway-CE12808-2-bgp-af-ipv4] network 19.19.168.45.152 255.255.255.248
[*Gateway-CE12808-2-bgp-af-ipv4] maximum load-balancing 32
```

```
[*Gateway-CE12808-2-bgp-af-ipv4] quit
[*Gateway-CE12808-2-bgp] quit
[*Gateway-CE12808-2] commit
```

----结束

组建 M-LAG 组

步骤1 配置M-LAG模式。

```
<Gateway-CE12808-1> system-view
[~Gateway-CE12808-1] stp mode rstp
[*Gateway-CE12808-1] stp v-stp enable
[*Gateway-CE12808-1] lacp m-lag priority 10
[~Gateway-CE12808-1] lacp m-lag system-id 00e0-fc00-0101 //建议使用M-LAG中Master的系统MAC作为 system-id, 在对端设备上system-id的配置要保持一致。可以使用命令display system mac-address来查看系统MAC。
[*Gateway-CE12808-1] commit

<Gateway-CE12808-2> system-view
[~Gateway-CE12808-2] stp mode rstp
[*Gateway-CE12808-2] stp v-stp enable
[*Gateway-CE12808-2] commit
[*Gateway-CE12808-2] lacp m-lag priority 10
[~Gateway-CE12808-2] lacp m-lag system-id 00e0-fc00-0101
[*Gateway-CE12808-2] commit
```

步骤2 在Gateway-CE12808-1和Gateway-CE12808-2上配置M-LAG的DFS Group及网关双活。

```
[~Gateway-CE12808-1] dfs-group 1
[*Gateway-CE12808-1-dfs-group-1] source ip 18.18.18.18
[*Gateway-CE12808-1-dfs-group-1] priority 150
[*Gateway-CE12808-1-dfs-group-1] active-active-gateway
[*Gateway-CE12808-1-dfs-group-1-active-active-gateway] peer 19.19.19.19
[*Gateway-CE12808-1-dfs-group-1-active-active-gateway] quit
[*Gateway-CE12808-1-dfs-group-1] quit
[*Gateway-CE12808-1] commit
[~Gateway-CE12808-2] dfs-group 1
[*Gateway-CE12808-2-dfs-group-1] source ip 19.19.19.19
[*Gateway-CE12808-2-dfs-group-1] priority 120
[*Gateway-CE12808-2-dfs-group-1] active-active-gateway
[*Gateway-CE12808-2-dfs-group-1-active-active-gateway] peer 18.18.18.18
[*Gateway-CE12808-2-dfs-group-1-active-active-gateway] quit
[*Gateway-CE12808-2-dfs-group-1] quit
[*Gateway-CE12808-2] commit
```

步骤3 在Gateway-CE12808-1和Gateway-CE12808-2上配置M-LAG的Peer-Link。

```
[~Gateway-CE12808-1] interface eth-trunk 0
[*Gateway-CE12808-1-Eth-Trunk0] trunkport 40ge 1/0/23
[*Gateway-CE12808-1-Eth-Trunk0] trunkport 40ge 2/0/23
[*Gateway-CE12808-1-Eth-Trunk0] mode lacp-static
[*Gateway-CE12808-1-Eth-Trunk0] peer-link 1
[*Gateway-CE12808-1-Eth-Trunk0] quit
[*Gateway-CE12808-1] commit

[~Gateway-CE12808-2] interface eth-trunk 0
[*Gateway-CE12808-2-Eth-Trunk0] trunkport 40ge 1/0/23
[*Gateway-CE12808-2-Eth-Trunk0] trunkport 40ge 2/0/23
[*Gateway-CE12808-2-Eth-Trunk0] mode lacp-static
[*Gateway-CE12808-2-Eth-Trunk0] peer-link 1
[*Gateway-CE12808-2-Eth-Trunk0] quit
[*Gateway-CE12808-2] commit
```

步骤4 在Gateway-CE12808-1和Gateway-CE12808-2上配置M-LAG成员口(对接防火墙示例)。

#配置Gateway-CE12808-1与防火墙的互联。

```
[~Gateway-CE12808-1] interface eth-trunk 21
[*Gateway-CE12808-1-Eth-Trunk21] description "to-FW-USG9560-1"
[*Gateway-CE12808-1-Eth-Trunk21] port link-type trunk
[*Gateway-CE12808-1-Eth-Trunk21] undo port trunk allow-pass vlan 1
[*Gateway-CE12808-1-Eth-Trunk21] port trunk allow-pass vlan 11
[*Gateway-CE12808-1-Eth-Trunk21] port trunk allow-pass vlan 12
[*Gateway-CE12808-1-Eth-Trunk21] trunkport 10ge 3/0/0
[*Gateway-CE12808-1-Eth-Trunk21] trunkport 10ge 3/0/1
[*Gateway-CE12808-1-Eth-Trunk21] dfs-group 1 m-lag 1
[*Gateway-CE12808-1-Eth-Trunk21] quit
[*Gateway-CE12808-1] interface eth-trunk 31
[*Gateway-CE12808-1-Eth-Trunk31] description "to-FW-USG9560-2"
[*Gateway-CE12808-1-Eth-Trunk31] port link-type trunk
[*Gateway-CE12808-1-Eth-Trunk31] undo port trunk allow-pass vlan 1
[*Gateway-CE12808-1-Eth-Trunk31] port trunk allow-pass vlan 11
[*Gateway-CE12808-1-Eth-Trunk31] port trunk allow-pass vlan 12
[*Gateway-CE12808-1-Eth-Trunk31] trunkport 10ge 3/0/2
[*Gateway-CE12808-1-Eth-Trunk31] trunkport 10ge 3/0/3
[*Gateway-CE12808-1-Eth-Trunk31] dfs-group 1 m-lag 2
[*Gateway-CE12808-1-Eth-Trunk31] quit
[*Gateway-CE12808-1] commit
```

配置Gateway-CE12808-2与防火墙的互联。

```
[~Gateway-CE12808-2] interface eth-trunk 21
[*Gateway-CE12808-2-Eth-Trunk21] description "to-FW-USG9560-1"
[*Gateway-CE12808-2-Eth-Trunk21] port link-type trunk
[*Gateway-CE12808-2-Eth-Trunk21] undo port trunk allow-pass vlan 1
[*Gateway-CE12808-2-Eth-Trunk21] port trunk allow-pass vlan 11
[*Gateway-CE12808-2-Eth-Trunk21] port trunk allow-pass vlan 12
[*Gateway-CE12808-2-Eth-Trunk21] trunkport 10ge 3/0/0
[*Gateway-CE12808-2-Eth-Trunk21] trunkport 10ge 3/0/1
[*Gateway-CE12808-2-Eth-Trunk21] dfs-group 1 m-lag 1
[*Gateway-CE12808-2-Eth-Trunk21] quit
[*Gateway-CE12808-2] interface eth-trunk 31
[*Gateway-CE12808-2-Eth-Trunk31] description "to-FW-USG9560-2"
[*Gateway-CE12808-2-Eth-Trunk31] port link-type trunk
[*Gateway-CE12808-2-Eth-Trunk31] undo port trunk allow-pass vlan 1
*Gateway-CE12808-2-Eth-Trunk31 port trunk allow-pass vlan 11
[*Gateway-CE12808-2-Eth-Trunk31] port trunk allow-pass vlan 12
[*Gateway-CE12808-2-Eth-Trunk31] trunkport 10ge 3/0/2
[*Gateway-CE12808-2-Eth-Trunk31] trunkport 10ge 3/0/3
[*Gateway-CE12808-2-Eth-Trunk31] dfs-group 1 m-lag 2
[*Gateway-CE12808-2-Eth-Trunk31] quit
[*Gateway-CE12808-2] commit
```

□ 说明

防火墙对接网关采用M-LAG时,两端的LAG模式仅能为手工负载分担模式。

----结束

配置 BGP EVPN

配置BGP EVPN作为VXLAN的控制平面,将Leaf作为BGP路由反射器客户端,Spine作为BGP路由器反射器,在Leaf与Spine之间建立IBGP EVPN邻居。

步骤1 在Gateway-CE12808-1和Gateway-CE12808-2上使能EVPN作为VXLAN的控制平面功能。

```
[~Gateway-CE12808-1] evpn-overlay enable
[*Gateway-CE12808-1] commit
[~Gateway-CE12808-2] evpn-overlay enable
[*Gateway-CE12808-2] commit
```

步骤2 配置Gateway-CE12808-1和Gateway-CE12808-2上分别与Spine节点建立BGP EVPN对等体关系,Spine作为路由反射器RR。

#在Gateway-CE12808-1上配置与Spine建立BGP EVPN对等体关系。

```
[~Gateway-CE12808-1] bgp 100 instance evpn1
[*Gateway-CE12808-1-bgp-instance-evpn1] router-id 18.18.18.18
[*Gateway-CE12808-1-bgp-instance-evpn1] peer 11.11.11.14 as-number 100
//与Spine-CE12804-1建立BGP EVPN对等体关系
[*Gateway-CE12808-1-bgp-instance-evpn1] peer 11.11.11.14 connect-interface loopback 1
[*Gateway-CE12808-1-bgp-instance-evpn1] peer 11.11.11.15 as-number 100
//与Spine-CE12804-2建立BGP EVPN对等体关系
[*Gateway-CE12808-1-bgp-instance-evpn1] peer 11.11.11.15 connect-interface loopback 1
[*Gateway-CE12808-1-bgp-instance-evpn1] l2vpn-family evpn
[*Gateway-CE12808-1-bgp-instance-evpn1-af-evpn] peer 11.11.11.14 enable
[*Gateway-CE12808-1-bgp-instance-evpn1-af-evpn] peer 11.11.11.14 advertise arp
[*Gateway-CE12808-1-bgp-instance-evpn1-af-evpn] peer 11.11.11.15 enable
[*Gateway-CE12808-1-bgp-instance-evpn1-af-evpn] peer 11.11.11.15 advertise arp
[*Gateway-CE12808-1-bgp-instance-evpn1-af-evpn] quit
[*Gateway-CE12808-1-bgp-instance-evpn1] quit
[*Gateway-CE12808-1] commit
```

#在Gateway-CE12808-2上配置与Spine建立BGP EVPN对等体关系。

```
[~Gateway-CE12808-2] bgp 100 instance evpn1
[*Gateway-CE12808-2-bgp-instance-evpn1] router-id 19.19.19.19
[*Gateway-CE12808-2-bgp-instance-evpn1] peer 11.11.11.14 as-number 100
//与Spine-CE12804-1建立BGP EVPN对等体关系
[*Gateway-CE12808-2-bgp-instance-evpn1] peer 11.11.11.14 connect-interface loopback 1
[*Gateway-CE12808-2-bgp-instance-evpn1] peer 11.11.11.15 as-number 100
//与Spine-CE12804-2建立BGP EVPN对等体关系
[*Gateway-CE12808-2-bgp-instance-evpn1] peer 11.11.11.15 connect-interface loopback 1
[*Gateway-CE12808-2-bgp-instance-evpn1] l2vpn-family evpn
[*Gateway-CE12808-2-bgp-instance-evpn1-af-evpn] peer 11.11.11.14 enable
[*Gateway-CE12808-2-bgp-instance-evpn1-af-evpn] peer 11.11.11.14 advertise arp
[*Gateway-CE12808-2-bgp-instance-evpn1-af-evpn] peer 11.11.11.15 enable
[*Gateway-CE12808-2-bgp-instance-evpn1-af-evpn] peer 11.11.11.15 advertise arp
[*Gateway-CE12808-2-bgp-instance-evpn1-af-evpn] quit
[*Gateway-CE12808-2-bgp-instance-evpn1] quit
[*Gateway-CE12808-2] commit
```

步骤3 配置EVPN实例。

在Gateway-CE12808-1上配置EVPN实例。

```
[~Gateway-CE12808-1] bridge-domain 10
[*Gateway-CE12808-1-bd10] vxlan vni 10010
[*Gateway-CE12808-1-bd10] evpn
[*Gateway-CE12808-1-bd10-evpn] route-distinguisher 16:1
[*Gateway-CE12808-1-bd10-evpn] vpn-target 1:1
[*Gateway-CE12808-1-bd10-evpn] quit
[*Gateway-CE12808-1-bd10] quit
[*Gateway-CE12808-1] commit
```

在Gateway-CE12808-2上配置EVPN实例。

```
[~Gateway-CE12808-2] bridge-domain 10
[*Gateway-CE12808-2-bd10] vxlan vni 10010
[*Gateway-CE12808-2-bd10] evpn
[*Gateway-CE12808-2-bd10-evpn] route-distinguisher 17:1
[*Gateway-CE12808-2-bd10-evpn] vpn-target 1:1
[*Gateway-CE12808-2-bd10-evpn] quit
[*Gateway-CE12808-2-bd10] quit
[*Gateway-CE12808-2] commit
```

步骤4 配置头端复制功能。

```
# 在Gateway-CE12808-1配置头端复制功能。
```

```
[*Gateway-CE12808-1] interface nve 1

[*Gateway-CE12808-1-Nve1] source 11.11.11.16

[*Gateway-CE12808-1-Nve1] mac-address 0000-5e00-0106

[*Gateway-CE12808-1-Nve1] vni 10010 head-end peer-list protocol bgp

[*Gateway-CE12808-1-Nve1] quit

[*Gateway-CE12808-1] commit
```

在Gateway-CE12808-2配置头端复制功能。

[~Gateway-CE12808-2] interface nve 1

[*Gateway-CE12808-2-Nve1] source 11.11.11.16

[*Gateway-CE12808-2-Nve1] mac-address 0000-5e00-0106

[*Gateway-CE12808-2-Nve1] vni 10010 head-end peer-list protocol bgp

[*Gateway-CE12808-2-Nve1] quit

[*Gateway-CE12808-2] commit

步骤5 配置VPN实例绑定网关。

```
[~Gateway-CE12808-1] ip vpn-instance vpn1
```

[*Gateway-CE12808-1-vpn-instance-vpn1] route-distinguisher 11:1

[*Gateway-CE12808-1-vpn-instance-vpn1-af-ipv4] quit

[*Gateway-CE12808-1-vpn-instance-vpn1] quit

[*Gateway-CE12808-1] commit

[~Gateway-CE12808-2] ip vpn-instance vpn1

[*Gateway-CE12808-2-vpn-instance-vpn1] route-distinguisher 11:2

[*Gateway-CE12808-2-vpn-instance-vpn1-af-ipv4] quit

[*Gateway-CE12808-2-vpn-instance-vpn1] quit

[*Gateway-CE12808-2] commit

[~Gateway-CE12808-1] **interface vlanif 12** //GW的VRF与防火墙的业务互联VLAN [*Gateway-CE12808-1-Vlanif12] **description** "**to firewall-1~2**"

[*Gateway-CE12808-1-Vlanif12] ip binding vpn-instance vpn1

[*Gateway-CE12808-1-Vlanif12] commit

[~Gateway-CE12808-1-Vlanif12] quit

[-Gateway-CE12808-1] ip route-static vpn-instance vpn1 0.0.0.0 0.0.0.0 192.168.45.164 //GW的VRF vpn1 的默认路由指向防火墙

[*Gateway-CE12808-1] commit

[~Gateway-CE12808-2] interface vlanif 12 //网关设备与防火墙的业务互联VLAN

[*Gateway-CE12808-2-Vlanif12] description "to firewall-1~2 Vsys"

[*Gateway-CE12808-2-Vlanif12] ip binding vpn-instance vpn1

[*Gateway-CE12808-2-Vlanif12] commit

[~Gateway-CE12808-2-Vlanif12] quit

[~Gateway-CE12808-2] ip route-static vpn-instance vpn1 0.0.0.0 192.168.45.164 //GW的VRF

vpn1的默认路由指向防火墙

[*Gateway-CE12808-1] commit

步骤6 配置三层网关。

在Gateway-CE12808-1配置三层网关。

[~Gateway-CE12808-1] interface vbdif 10

[*Gateway-CE12808-1-Vbdif10] ip address 192.168.10.1 24 //服务器所在网段

[*Gateway-CE12808-1-Vbdif10] mac-address 0000-5e00-0101

[~Gateway-CE12808-1-Vbdif10] ip binding vpn-instance vpn1 //连接防火墙时区分流量引入防火墙

[*Gateway-CE12808-1-Vbdif10] quit

[*Gateway-CE12808-1] commit

#在Gateway-CE12808-2配置三层网关。

[~Gateway-CE12808-2] interface vbdif 10

[*Gateway-CE12808-2-Vbdif10] ip address 192.168.10.1 24 //服务器所在网段

[*Gateway-CE12808-2-Vbdif10] mac-address 0000-5e00-0101

[~Gateway-CE12808-2-Vbdif10] ip binding vpn-instance vpn1

[*Gateway-CE12808-2-Vbdif10] quit

[*Gateway-CE12808-2] commit

----结束

配置 MAC 漂移白名单

三层架构时,VXLAN流量经过Spine设备到达GW时,可能会由于产品约束限制在GW 上学习了错误的MAC地址导致MAC漂移告警,需要在GW设备上基于MAC地址配置 MAC漂移白名单,将Spine设备与GW互联的三层口MAC地址配置为白名单不再检测 MAC漂移,正常情况下此MAC是作为隧道报文的外层源MAC,不应该被学习,所以调 整后对业务功能不会有任何影响。

步骤1 查看两台Spine设备与GW互连的三层口MAC地址。

```
[~Spine-CE12804-1] interface 40ge 1/0/4 //获取Spine-1设备与GW互连的三层口MAC地址
[*Spine-CE12804-1-40GE1/0/4] undo portswitch
[~Spine-CE12804-1-40GE1/0/4] commit
[~Spine-CE12804-1-40GE1/0/4] display this interface | include Hardware address
IP Sending Frames' Format is PKTFMT_ETHNT_2, Hardware address is 200b-c732-d202
[~Spine-CE12804-1-40GE1/0/4] quit
[~Spine-CE12804-1] interface 40ge 1/0/5
[*Spine-CE12804-1-40GE1/0/5] undo portswitch
[~Spine-CE12804-1-40GE1/0/5] commit
[~Spine-CE12804-1-40GE1/0/5] display this interface | include Hardware address
IP Sending Frames' Format is PKTFMT_ETHNT_2, Hardware address is 200b-c732-d202
[~Spine-CE12804-1-40GE1/0/5] quit
[~Spine-CE12804-2] interface 40ge 1/0/4 //获取Spine-2设备与GW互连的三层口MAC地址
[*Spine-CE12804-2-40GE1/0/4] undo portswitch
[~Spine-CE12804-2-40GE1/0/4] commit
[~Spine-CE12804-2-40GE1/0/4] display this interface | include Hardware address
IP Sending Frames' Format is PKTFMT_ETHNT_2, Hardware address is 346a-c246-be01
[~Spine-CE12804-2-40GE1/0/4] quit
[~Spine-CE12804-2] interface 40ge 1/0/5
[*Spine-CE12804-2-40GE1/0/5] undo portswitch
[~Spine-CE12804-2-40GE1/0/5] commit
[~Spine-CE12804-2-40GE1/0/5] display this interface | include Hardware address
IP Sending Frames' Format is PKTFMT_ETHNT_2, Hardware address is 346a-c246-be01
[~Spine-CE12804-2-40GE1/0/5] quit
```

步骤2 根据步骤一中获取到的MAC地址配置白名单。

```
[~Gateway-CE12808-1] mac-address flapping detection exclude 200b-c732-d202 48
[~Gateway-CE12808-1] mac-address flapping detection exclude 346a-c246-be01 48
[~Gateway-CE12808-2] mac-address flapping detection exclude 200b-c732-d202 48
[~Gateway-CE12808-2] mac-address flapping detection exclude 346a-c246-be01 48
```

----结束

配置网关与外部路由器互联

步骤1 配置网关与外部路由器的互联端口。

#在Gateway-CE12808-1配置与Router的互联端口。

```
[~Gateway-CE12808-1] interface 10ge 3/0/4 //连接外部PE设备(Router-1)
[~Gateway-CE12808-1-10GE3/0/4] undo portswitch
[*Gateway-CE12808-1-10GE3/0/4] ip address 192.168.44.157 30
[*Gateway-CE12808-1-10GE3/0/4] set up-delay 180 //配置接口上报状态变化事件的延时时间
[*Gateway-CE12808-1-10GE3/0/4] quit
[*Gateway-CE12808-1] interface 10ge 3/0/5 //连接外部PE设备(Router-2)
[*Gateway-CE12808-1-10GE3/0/5] undo portswitch
[*Gateway-CE12808-1-10GE3/0/5] ip address 192.168.44.169 30
[*Gateway-CE12808-1-10GE3/0/5] set up-delay 180
[*Gateway-CE12808-1-10GE3/0/5] quit
[*Gateway-CE12808-1] commit
```

#在Gateway-CE12808-2配置与Router的互联端口。

```
[~Gateway-CE12808-2] interface 10ge 3/0/4 //连接外部PE设备(Router-1)
[~Gateway-CE12808-2-10GE3/0/4] undo portswitch
[*Gateway-CE12808-2-10GE3/0/4] ip address 192.168.44.173 30
[*Gateway-CE12808-2-10GE3/0/4] set up-delay 180
[*Gateway-CE12808-2-10GE3/0/4] quit
[*Gateway-CE12808-2] interface 10ge 3/0/5 //连接外部PE设备(Router-2)
[*Gateway-CE12808-2-10GE3/0/5] undo portswitch
[*Gateway-CE12808-2-10GE3/0/5] ip address 192.168.44.161 30
[*Gateway-CE12808-2-10GE3/0/5] set up-delay 180
[*Gateway-CE12808-2-10GE3/0/5] quit
[*Gateway-CE12808-2] commit
```

步骤2 配置网关之间的三层互联端口,以组成"口"字型和交叉字型结合的组网。

在Gateway-CE12808-1上配置与Gateway-CE12808-2的互联端口。

[--Gateway-CE12808-1] **interface 10ge 3/0/6** //连接Gateway-CE12808-2 [--Gateway-CE12808-1-10GE3/0/6] **undo portswitch** [*Gateway-CE12808-1-10GE3/0/6] **ip address 192.168.44.165 30** [*Gateway-CE12808-1-10GE3/0/6] **commit** [--Gateway-CE12808-1-10GE3/0/6] **quit**

在Gateway-CE12808-2上配置与Gateway-CE12808-1的互联端口。

[~Gateway-CE12808-2] interface 10ge 3/0/6 //连接Gateway-CE12808-1 [~Gateway-CE12808-2-10GE3/0/6] undo portswitch [*Gateway-CE12808-2-10GE3/0/6] ip address 192.168.44.166 30 [*Gateway-CE12808-2-10GE3/0/6] commit [~Gateway-CE12808-2-10GE3/0/6] quit

步骤3 网关设备上配置指向外部路由器的默认路由。

在Gateway-CE12808-1上配置默认路由。

```
[~Gateway-CE12808-1] ip route-static 0.0.0.0 0.0.0.0 192.168.44.158
[*Gateway-CE12808-1] ip route-static 0.0.0.0 0.0.0.0 192.168.44.170
[*Gateway-CE12808-1] ip route-static 0.0.0.0 0.0.0 192.168.44.166 preference 80 //指向Gateway-CE12808-2,作为备份。
[*Gateway-CE12808-1] commit
[~Gateway-CE12808-1] quit
```

#在Gateway-CE12808-2上配置默认路由。

```
[~Gateway-CE12808-2] ip route-static 0.0.0.0 0.0.0.0 192.168.44.174
[*Gateway-CE12808-2] ip route-static 0.0.0.0 0.0.0.0 192.168.44.162
[*Gateway-CE12808-2] ip route-static 0.0.0.0 0.0.0 192.168.44.165 preference 80 //指向Gateway-CE12808-1, 作为备份。
[*Gateway-CE12808-2] commit
[~Gateway-CE12808-2] quit
```

----结束

3.1.1.6 配置增强功能

CE12800系列交换机需要配置步骤1/2/3中的哪些命令行取决于设备所选择的单板,具体可以参照下表(CE12800E系列交换机无需配置)。

表 3-1 CE12800 系列交换机优化配置

单板	配置命令
全部插入E系列接口板	1. assign forward nvo3 service extend enable 2. assign forward nvo3 acl extend enable (重启
	生效)
	3. assign forward nvo3-gateway enhanced L3
全部插入F系列的FD/FDA类型接口板	1. assign forward nvo3 acl extend enable(重启 生效)
(仅允许插入F/G类型交换	2. set serdes capability enhanced (重启生效)
网板)	3. set forward capability enhanced (重启生效)

单板	配置命令
E系列接口板和F系列接口板 混插 (不推荐) (仅允许插入F/G类型交换 网板)	 assign forward nvo3 service extend enable assign forward nvo3 acl extend enable (重启 生效) assign forward nvo3-gateway enhanced L3 assign forward nvo3 f-linecard compatibility enable

步骤1 配置网关设备的NVO3扩展功能。

当CE12800设备作为网关时,默认情况下,NVO3的扩展功能未使能。此时在设备上部署NVO3业务后,若再叠加其他需要使用ACL的业务(如MQC、简化ACL、流量监管、BD流量统计、DHCP等),会较大概率出现叠加失败的情况。

在部署NVO3的设备上可以使用以下方法以降低其他业务叠加失败的风险。

[~Gateway-CE12808-1] assign forward nvo3 service extend enable

[*Gateway-CE12808-1] assign forward nvo3 acl extend enable

[*Gateway-CE12808-1] commit

[~Gateway-CE12808-1] quit

<~Gateway-CE12808-1> save

<~Gateway-CE12808-1> reboot

[~Gateway-CE12808-2] assign forward nvo3 service extend enable

[*Gateway-CE12808-2] assign forward nvo3 acl extend enable

[*Gateway-CE12808-2] commit

[~Gateway-CE12808-2] quit

<~Gateway-CE12808-2> save

<~Gateway-CE12808-2> reboot

上述两条命令的使用说明如下:

● 在GW设备系统视图下,使用**assign forward nvo3 service extend enable**命令,使能NVO3业务扩展功能。使能此命令后,在非CE-L48GT-EC、CE-L48GS-EA、CE-L48GS-EC、CE-L24XS-BA、CE-L24XS-EA、CE-L48XS-BA、CE-L48XS-EA和CE-L24LQ-EA单板上,能减少这种业务叠加失败的问题。

山 说明

此命令对CE-L48GT-EA、CE-L48GT-EC、CE-L48GS-EA、CE-L48GS-EC、CE-L24XS-BA、CE-L24XS-EA、CE-L48XS-BA、CE-L48XS-EA和CE-L24LQ-EA单板不生效。

使能此命令后,非上述单板上通过NVO3隧道传输的、长度在230~294字节的报文不能发送至上述单板。

配置此命令后,才能在除CE-L48GT-EA、CE-L48GT-EC、CE-L48GS-EA、CE-L48GS-EC、CE-L24XS-BA、CE-L24XS-EA、CE-L48XS-BA、CE-L48XS-EA和CE-L24LQ-EA之外的单板上使用路径探测功能。

 在GW设备系统视图下,使用assign forward nvo3 acl extend enable命令,使 能NVO3的ACL扩展功能。

山 说明

此命令仅支持在Admin-VS中配置,配置后对所有VS生效。 执行此命令使能NVO3的ACL扩展功能后,需重启设备使配置生效。

步骤2 配置NVO3网关的增强模式。

CE12800缺省情况下,未配置NVO3网关的增强模式,设备采用的是环回模式,即经过NVO3封装后的报文需要在设备内部先环回后才能继续转发。此时,GW设备上的VXLAN三层转发封装/解封装的流量,超过线卡总转发性能的50%时,有较大可能出现丢包。目前,可以根据实际需要配置NVO3的网关增强模式,解决此问题。

在GW设备的系统视图下,使用**assign forward nvo3-gateway enhanced l3**命令行配置NVO3网关的增强模式。

[~Gateway-CE12808-1] assign forward nvo3-gateway enhanced l3

[~Gateway-CE12808-2] assign forward nvo3-gateway enhanced l3

□ 说明

需要先执行**assign forward nvo3 service extend enable**命令,使能NVO3业务扩展功能。此时请确保设备上不存在CE-L48GT-EA、CE-L48GT-EC、CE-L48GS-EA、CE-L48GS-EC、CE-L24XS-BA、CE-L24XS-BA、CE-L48XS-BA、CE-L48XS-EA和CE-L24LQ-EA单板,或者这些单板上没有承载VXLAN相关业务。

若承载VXLAN业务的单板为CE-L24XS-EC、CE-L48XS-EC、CE-L24LQ-EC、CE-L48XT-EC、CE-L24LQ-EC1、CE-L08CC-EC、CE-L02LQ-EC、CE-L06LQ-EC时,仅支持通过查询ARP主机表进行VXLAN隧道封装,不支持通过查询最长匹配路由表进行VXLAN隧道封装。

服务器网卡主备方式接入的两台接入交换机需要组建为堆叠系统,或者M-LAG方式将服务器双 归连接到接入交换机,此时接入交换机连接服务器的端口不要作为M-LAG的成员接口。

步骤3 配置CE12800系列的增强模式(CE12800E系列不支持配置本步骤命令)。

1. 配置单板的互通模式为增强模式。

[~Gateway-CE12808-1] set forward capability enhanced

Warning: Current configuration should be committed and saved, and it will take effect after reboot. [Y/N]: **y**

□ 说明

增强模式是指:设备仅允许插入FD/FDA类型接口板,且仅允许插入F/G类型交换网板。

当单板互通模式为增强模式时,不可以插入EA/EC/ED/EF/EG/BA/CE-FWA/CE-IPSA类型接口板和A/B/C类型交换网板,否则单板会被下电。设备交换网板是A/B/C类型情况下,不能插入FD/FDA类接口板;设备插入了FD/FDA类接口板的情况下,不能插入A/B/C类型交换网板。

配置该命令后,需要保存配置并重启设备后配置才会生效(保存的配置文件要作为下次启动文件)。设备重启后需要再次执行命令save保存配置。

配置该命令后,建议同时配置命令set serdes capability enhanced,设置接口板与交换网板之间Serdes的速率模式为增强模式,这样可以达到最大转发性能。

该命令与以下命令互斥,配置本命令前,需要删除以下配置:

- assign forward nvo3-gateway enhanced { l2 | l3 }
- assign forward nvo3 eth-trunk hash disable

当设备上有EA/EC/ED/EF/EG/BA/CE-FWA/CE-IPSA类型接口板和FD/FDA类型接口板混插时,需要配置单板的互通模式为非增强模式,此时为了防止VXLAN流量成环,需要执行assign forward nvo3 f-linecard compatibility enable命令。

2. 配置Serdes速率模式为增强模式。

[~Gateway-CE12808-1] set serdes capability enhanced

Warning: Current configuration should be committed and saved, and it will take effect after reboot. [Y/N]: **y**

🗀 说明

配置该命令后,需要重启设备配置才会生效。

设备交换网板是A/B/C类型情况下,不能插入FD/FDA类接口板;设备插入了FD/FDA类接口板的情况下,不能插入A/B/C类型交换网板。

步骤4 配置ARP广播报文抑制。

前面已经在网关工作组和Leaf工作组上执行命令**peer** { *ipv4-address* | *group-name* } **advertise arp**,配置BGP EVPN发布路由功能,用来发布VXLAN学习到的主机信息。为了实现ARP广播报文抑制的功能,还需要在网关工作组和Leaf工作组上配置如下步骤。

1. 在网关工作组上使能BGP EVPN进行主机信息搜集的功能。

[~Gateway-CE12808-1] interface vbdif 10

[~Gateway-CE12808-1-Vbdif10] arp collect host enable

[*Gateway-CE12808-1-Vbdif10] commit

[~Gateway-CE12808-1-Vbdif10] quit

[~Gateway-CE12808-2] interface vbdif 10

[~Gateway-CE12808-2-Vbdif10] arp collect host enable

[*Gateway-CE12808-2-Vbdif10] commit

[~Gateway-CE12808-2-Vbdif10] quit

2. 在所有Leaf工作组上使能ARP广播报文抑制功能(以Leaf堆叠工作组为例)。

[~Leaf-CE6851HI-1&CE6851HI-2] bridge-domain 10

[~Leaf-CE6851HI-1&CE6851HI-2-bd10] arp broadcast-suppress enable

[*Leaf-CE6851HI-1&CE6851HI-2-bd10] commit

[~Leaf-CE6851HI-1&CE6851HI-2-bd10] quit

步骤5 在所有Leaf工作组上配置BUM报文的流量抑制功能(以Leaf堆叠工作组为例)。

[~Leaf-CE6851HI-1&CE6851HI-2] interface 10ge 1/0/1

[~Leaf-CE6851HI-1&CE6851HI-2-10GE1/0/1] storm suppression unknown-unicast 1

[*Leaf-CE6851HI-1&CE6851HI-2-10GE1/0/1] storm suppression multicast packets 200

[*Leaf-CE6851HI-1&CE6851HI-2-10GE1/0/1] storm suppression broadcast packets 1000

[*Leaf-CE6851HI-1&CE6851HI-2-10GE1/0/1] commit

[~Leaf-CE6851HI-1&CE6851HI-2-10GE1/0/1] quit

[~Leaf-CE6851HI-1&CE6851HI-2] interface 10ge 1/0/2

[~Leaf-CE6851HI-1&CE6851HI-2-10GE1/0/2] storm suppression unknown-unicast 1

[*Leaf-CE6851HI-1&CE6851HI-2-10GE1/0/2] storm suppression multicast packets 200

[*Leaf-CE6851HI-1&CE6851HI-2-10GE1/0/2] storm suppression broadcast packets 1000

[*Leaf-CE6851HI-1&CE6851HI-2-10GE1/0/2] commit

[~Leaf-CE6851HI-1&CE6851HI-2-10GE1/0/2] quit

步骤6 在网关工作组上配置ARP资源分配模式为扩展模式。

[~Gateway-CE12808-1] arp resource-mode extend

[~Gateway-CE12808-2] arp resource-mode extend

----结束

3.1.2 Spine/Leaf 二层架构的硬件集中式 VXLAN

□ 说明

本节只重点描述本方案与"GW/Spine/Leaf三层架构"方案中不同的配置内容,其余内容与"GW/Spine/Leaf三层架构"一致。

主要配置差异内容如下:

- Gateway与Leaf互联的接口配置
- Leaf上BGP路由的配置
- Gateway上BGP路由的配置
- Gateway和Leaf上的BGP EVPN功能配置

3 VXLAN 部署建议

组网需求

Spine/Leaf二层架构的硬件集中式VXLAN组网图如图 1 Spine/Leaf二层架构的硬件集中式VXLAN组网图所示,

Internet **R**> **(R)** Router-2 Spine/GW Exit-Gateway-12808-1 Exit-Gateway-12808-2 🔝 LB-1 FW-USG9560-1 RB-2 FW-USG9560-2 Leaf Leaf-CE6851HI-1 Leaf-CE6851HI-5 Leaf-CE6851HI-3 Leaf-CE6851HI-4 Leaf-CE6851HI-6 Leaf-CE6851HI-2 Server

图 3-2 Spine/Leaf 二层架构的硬件集中式 VXLAN 组网图

- 服务器层:通过二层子接口方式接入VXLAN网络。
- Leaf:采用堆叠方式、M-LAG或SVF方式连接服务器。Leaf节点和Spine节点通过 三层互联。Leaf的堆叠工作组、M-LAG工作组或者SVF工作组作为VTEP实现服务 器流量接入VXLAN网络。
- Spine/GW:两台GW之间组成M-LAG,即双活网关。GW与Leaf之间三层互联。GW对外连接外部路由器Router-1和Router-2。
- FW: 两台FW分别旁挂在GW, FW之间配置为主备镜像模式工作。
- LB: 由厂商部署。

端口连线规划

Leaf-CE6851HI-1和Leaf-CE6851HI-2部署堆叠场景,端口规划参见下表。

规划描述	本端设备名称	物理端口	对端设备名称	物理端口
规划Leaf间互连端口	Leaf-	40GE1/0/1	Leaf-	40GE1/0/1
作为堆叠口,堆叠链	CE6851HI-1	~2	CE6851HI-2	~2
路至少2条	Leaf-	40GE1/0/1	Leaf-	40GE1/0/1
	CE6851HI-2	~2	CE6851HI-1	~2
规划Leaf与Spine互连 端口	Leaf- CE6851HI-1	40GE1/0/3	Gateway- CE12808-1	40GE1/0/0

规划描述	本端设备名称	物理端口	对端设备名称	物理端口
	Leaf- CE6851HI-1	40GE1/0/4	Gateway- CE12808-2	40GE1/0/1
	Leaf- CE6851HI-2	40GE1/0/3	Gateway- CE12808-2	40GE1/0/0
	Leaf- CE6851HI-2	40GE1/0/4	Gateway- CE12808-1	40GE1/0/1
规划Leaf与服务器的 互连端口,服务器双	Leaf- CE6851HI-1	10GE1/0/1 ~2	服务器	Eth0
网卡以负载分担方式 上连,保证可靠性的 同时负载分担转发, 提高链路利用率	Leaf- CE6851HI-2	10GE1/0/1 ~2	服务器	Eth1

Leaf-CE6851HI-3和Leaf-CE6851HI-4部署M-LAG,端口规划参见下表。

规划描述	本端设备名 称	物理端口	对端设备名称	物理端口
规划M-LAG的peer- link端口,用于承载	Leaf- CE6851HI-3	40GE1/0/1~2	Leaf- CE6851HI-4	40GE1/0/1 ~2
M-LAG的协议报文和 故障场景下的数据报 文,为保证可靠性, 建议至少2条peer-link 链路	Leaf- CE6851HI-4	40GE1/0/1~2	Leaf- CE6851HI-3	40GE1/0/1 ~2
规划Leaf连接Spine互 连端口	Leaf- CE6851HI-3	40GE1/0/3	Gateway- CE12808-1	40GE1/0/2
	Leaf- CE6851HI-3	40GE1/0/4	Gateway- CE12808-2	40GE1/0/3
	Leaf- CE6851HI-4	40GE1/0/3	Gateway- CE12808-2	40GE1/0/2
	Leaf- CE6851HI-4	40GE1/0/4	Gateway- CE12808-1	40GE1/0/3
规划Leaf与服务器的 互连端口,服务器双	Leaf- CE6851HI-3	10GE1/0/1	服务器	Eth0
网卡以负载分担方式 上连,保证可靠性的 同时负载分担转发, 提高链路利用率	Leaf- CE6851HI-4	10GE1/0/1	服务器	Eth1

Leaf-CE6851HI-5和Leaf-CE6851HI-6部署SVF场景,下连三台CE5810交换机,端口规划参见下表。

规划描述	本端设备名 称	物理端口	对端设备名称	物理端口
规划Leaf间互连端口 作为堆叠口,堆叠链	Leaf- CE6851HI-5	40GE1/0/1~ 2	Leaf- CE6851HI-6	40GE1/0/1 ~2
路至少2条	Leaf- CE6851HI-6	40GE1/0/1~ 2	Leaf- CE6851HI-5	40GE1/0/1 ~2
规划Leaf与Spine互连 端口	Leaf- CE6851HI-5	40GE1/0/3	Gateway- CE12808-1	40GE1/0/6
	Leaf- CE6851HI-5	40GE1/0/4	Gateway- CE12808-2	40GE1/0/7
	Leaf- CE6851HI-6	40GE1/0/3	Gateway- CE12808-1	40GE1/0/6
	Leaf- CE6851HI-6	40GE1/0/4	Gateway- CE12808-2	40GE1/0/7
规划SVF父节点与叶	Leaf-	10GE1/0/1	CE5800-1	GE1/0/1
子节点互联端口 	CE6851HI-5	10GE1/0/2	CE5800-2	GE1/0/1
		10GE1/0/3	CE5800-3	GE1/0/1
	Leaf-	10GE1/0/1	CE5800-1	GE1/0/2
	CE6851HI-6	10GE1/0/2	CE5800-2	GE1/0/2
		10GE1/0/3	CE5800-3	GE1/0/2
规划叶子交换机与服	CE5810-1	GE1/0/1	服务器	Eth0
务器的互连端口,服 务器双网卡以负载分 担方式上连,保证链 路可靠性	CE5810-2	GE1/0/1	服务器	Eth1

Gateway-CE12808-1和Gateway-CE12808-2端口规划参见下表。

规划描述	本端设备名称	物理端口	对端设备名 称	物理端口
规划 Gateway与 Leaf的互连 端口	Gateway-CE12808-1	40GE1/0/0 ~3 40GE1/0/6 ~7	Leaf- CE6851HI- 1 Leaf- CE6851HI- 2 Leaf- CE6851HI- 3 Leaf- CE6851HI- 4 Leaf- CE6851HI- 5 Leaf- CE6851HI- 5	40GE1/0/3 ~4
	Gateway-CE12808-2	40GE1/0/0 ~3 40GE1/0/6 ~7	Leaf- CE6851HI- 1 Leaf- CE6851HI- 2 Leaf- CE6851HI- 3 Leaf- CE6851HI- 4 Leaf- CE6851HI- 5 Leaf- CE6851HI- 5	40GE1/0/3 ~4
规划M-LAG 的peer-link 端口,用于 承载M-LAG 的协议报员 和故障场景 下的数据报 文,靠性,建	Gateway-CE12808-1	40GE1/0/2 3 40GE2/0/2 3	Gateway- CE12808-2	40GE1/0/2 3 40GE2/0/2 3

规划描述	本端设备名称	物理端口	对端设备名 称	物理端口
议至少2条 peer-link链 路,且跨板 部署	Gateway-CE12808-2	40GE1/0/2 3 40GE2/0/2 3	Gateway- CE12808-1	40GE1/0/2 3 40GE2/0/2 3
规划 Gateway与 防火墙的互 连端口 说明	Gateway-CE12808-1	10GE3/0/0 ~1	FW- USG9560-1 FW- USG9560-2	GE1/0/1
两台主备镜 像模式防火 墙必须使用 同样的接口 连接到同一	Gateway-CE12808-1	10GE3/0/2 ~3	FW- USG9560-1 FW- USG9560-2	GE1/0/3
个GW设 备。例如 FW-1使用 GE 1/0/1接 口连接 GW-1,那	Gateway-CE12808-2	10GE3/0/0 ~1	FW- USG9560-1 FW- USG9560-2	GE1/0/2
么FW-2也 需要使用 GE 1/0/1接 口连接 GW-1。	Gateway-CE12808-2	10GE3/0/2 ~3	FW- USG9560-1 FW- USG9560-2	GE1/0/4
规划	Gateway-CE12808-1	10GE3/0/4	Router-1	GE1/0/0
Gateway与 出口路由器	Gateway-CE12808-1	10GE3/0/5	Router-2	GE1/0/0
的互连端口	Gateway-CE12808-2	10GE3/0/4	Router-1	GE1/0/1
	Gateway-CE12808-2	10GE3/0/5	Router-2	GE1/0/1
规划 Gateway之 间联接三, 时接和出组。 "口是是是是一个。" "可是是是一个。" "可是是是一个。 "可是是是是一个。" "可是是是是一个。" "可是是是是一个。" "可是是是是一个。" "可是是是是一个。" "可是是是是一个。" "可是是是是一个。" "可是是是一个。" "可是是是一个。" "可是是是一个。" "可是是是一个。" "可是是是一个。" "可是是是一个。" "可是是是一个。" "可是是是一个。" "可是是是一个。" "可是是是一个。" "可是是是一个。" "可是是是一个。" "可是是是一个。" "可是是是一个。" "可是是是一个。" "可是是是一个。" "可是是是一个。" "可是是是一个。" "可是是是是是一个。" "可是是是一个。" "可是是是一个。" "可是是是一个。" "可是是是一个。" "可是是是一个。" "可是是是一个。" "可是是是一个。" "可是是是一个。" "可是是是一个。" "可是是是一个。" "可是是是一个。" "可是是是一个。" "可是是是一个。" "可是是是一个。" "可是是是一个。" "可是是是一个。" "可是是是一个。" "可是是是一个。" "可是是是是是是是是是是是是是是是是是是是是是是是是是是是是是是是是是是是是	Gateway-CE12808-1	10GE3/0/6	Gateway- CE12808-2	10GE3/0/6

VLAN 规划

VLAN规划参见下表。

规划描述	规划建议	本示例使用ID
与防火墙 互连 VLAN	三层互连建议使用三层主接口,与FW因涉 及多个网段对接,需要规划VLAN,达到共 用物理链路目的	11 12(创建VLANIF 12绑定 VPN)
租户 VLAN	租户/业务使用VXLAN较多,需要考虑后续 的业务扩容	10

山 说明

对于CE6855HI和CE7855EI,在切换三层接口之前,要先配置vlan reserved for maininterface startvlanid to endvlanid,配置三层主接口专用的保留VLAN。

BD、VNI 规划

BD、VNI的规划参见下表。

规划描述	规划建议	本示例使用ID
BD	规划数量与VLAN相同,建议BD的ID与 VLAN的ID相同	本例统一使用BD 10
VNI	规划数量与BD相同,建议VNI的ID规划为 "BD+10000"(BD与VNI——对应,且 VNI的ID需要大于4096)	本例统一使用VNI 10010

RD 和 RT 值规划

RD的规划参见下表。

规划描述	规划建议	本示例使用ID
CE6851HI- 1、 CE6851HI- 2的RD值		11: 1
CE6851HI- 3的RD值	_	13: 1
CE6851HI- 4的RD值	_	14: 1
CE6851HI- 5、 CE6851HI- 6的RD值	_	15: 1

规划描述	规划建议	本示例使用ID
Gateway- CE12808-1 的RD值		16: 1
Gateway- CE12808-2 的RD值		17: 1

RT的规划参见下表。

规划描述	规划建议	本示例使用ID
CE6851HI- 1、 CE6851HI- 2的RT值	_	1: 1
CE6851HI- 3的RT值	_	1: 1
CE6851HI- 4的RT值	_	1: 1
CE6851HI- 5、 CE6851HI- 6的RT值		1: 1
Gateway- CE12808-1 的RT值	_	1: 1
Gateway- CE12808-2 的RT值	_	1: 1

IP 地址规划

网元接口地址规划参见下表,涉及互连地址、VTEP地址、BGP Router-ID、BGP建邻居 loopback地址、M-LAG心跳地址、业务网段地址。

端口互连IP地址规划参见下表。

规划描述	本端设备	对端设备	IP/MASK
规划CE6851HI-1、 CE6851HI-2与 Gateway互连端口 地址	Leaf- CE6851HI-1& CE6851HI-2	Gateway- CE12808-1	192.168.40.156/30 192.168.40.168/30

规划描述	本端设备	对端设备	IP/MASK
		Gateway- CE12808-2	192.168.40.164/30 192.168.40.160/30
规划CE6851HI-3、 CE6851HI-4与 Gateway互连端口	Leaf- CE6851HI-3& CE6851HI-4	Gateway- CE12808-1	192.168.41.156/30 192.168.41.168/30
地址	62000	Gateway- CE12808-2	192.168.41.164/30 192.168.41.160/30
规划CE6851HI-5、 CE6851HI-6与 Gateway互连端口	Leaf- CE6851HI-5&	Gateway- CE12808-1	192.168.46.156/30 192.168.46.168/30
地址	CE6851HI-6	Gateway- CE12808-2	192.168.46.164/30 192.168.46.160/30
规划Gateway- CE12808-1与 Router-1互连端口 地址	Gateway- CE12808-1	Router-1	192.168.44.156/30
规划Gateway- CE12808-2与 Router-2互连端口 地址	Gateway- CE12808-2	Router-2	192.168.44.160/30
规划Gateway- CE12808-1和 Gateway- CE12808-2互联	Gateway- CE12808-1	Gateway- CE12808-2	192.168.44.164/30
规划Gateway- CE12808-1&2与 FW互连端口地址	Gateway- CE12808-1 Gateway- CE12808-2	FW-USG9560-1 FW-USG9560-2	192.168.45.152/29 Virtual IP: 192.168.45.153 192.168.45.160/29 Virtual IP: 192.168.45.161

Loopback地址规划参见下表。

规划描述	设备	IP/MASK
规划Loopback0地址,作 为VTEP地址	Leaf-CE6851HI-1& CE6851HI-2	11.11.11.11/32
	Leaf-CE6851HI-3	11.11.11.12/32
	Leaf-CE6851HI-4	同Leaf-CE6851HI-3: 11.11.11.12/32

规划描述	设备	IP/MASK
	Leaf-CE6851HI-5& CE6851HI-6	11.11.11.17/32
	Gateway-CE12808-1	11.11.11.16/32
	Gateway-CE12808-2	同Gateway-CE12808-1: 11.11.11.16/32
规划Loopback1地址,作	Leaf-CE6851HI-3	13.13.13.13/32
为M-LAG心跳检测地址	Leaf-CE6851HI-4	14.14.14.14/32
	Gateway-CE12808-1	18.18.18.18/32
	Gateway-CE12808-2	19.19.19.19/32
规划Loopback2地址,作	Gateway-CE12808-1	21.21.21.21/32
为对接Router的EBGP邻居 地址	Gateway-CE12808-2	22.22.22.22/32

业务地址规划参见下表。

租户	租户IP/MASK	网关VBDIF	VRF
服务器1	192.168.10.2/24	192.168.10.1/24	VPN 1
服务器2	192.168.10.3/24	192.168.10.1/24	VPN 1
服务器3	192.168.10.4/24	192.168.10.1/24	VPN 1

路由规划

Underlay网络使用的路由协议常见为EBGP和OSPF两种,BGP从多方面保证了网络的安全性、灵活性、稳定性、可靠性和高效性:

- BGP采用认证和GTSM的方式,保证了网络的安全性。
- BGP提供了丰富的路由策略,能够灵活的进行路由选路。
- BGP提供了路由聚合和路由衰减功能用于防止路由振荡,有效提高了网络的稳定性。
- BGP使用TCP作为其传输层协议(端口号为179),并支持BGP与BFD联动、BGP Auto FRR和BGP GR和NSR,提高了网络的可靠性。

从网络演进方面考虑,EBGP适用于大型网络,而OSPF适用于中小型网络,本文档使用EBGP作为示例。

路由相关的规划参见下表。

网元	AS域号	Router ID
Leaf- CE6851HI-1&CE6851-2	65021	Loopback0地址
Leaf-CE6851HI-3	65022	Loopback1地址
Leaf-CE6851HI-4	65022	Loopback1地址
Leaf- CE6851HI-5&CE6851HI-6	65024	Loopback0地址
Gateway-CE12808-1	65000	Loopback1地址
Gateway-CE12808-2	65000	Loopback1地址
Router-1	65047	Loopback0地址
Router-2	65048	Loopback0地址

3.1.2.1 配置 Gateway 对接 Leaf 的接口地址

在本方案中,Spine节点和Gateway融合部署,因此Gateway直接与Leaf互联。

步骤1 在Gateway-CE12808-1上配置接口IP地址。

```
[~Huawei] sysname Gateway-CE12808-1
[*Huawei] commit
[~Gateway-CE12808-1] interface 40ge 1/0/0
[~Gateway-CE12808-1-40GE1/0/0] description "to-Leaf-CE6851HI-1& CE6851HI-2"
[*Gateway-CE12808-1-40GE1/0/0] undo portswitch
[*Gateway-CE12808-1-40GE1/0/0] ip address 192.168.40.158 30
[*Gateway-CE12808-1-40GE1/0/0] commit
[~Gateway-CE12808-1-40GE1/0/0] quit
[~Gateway-CE12808-1] interface 40ge 1/0/1
[~Gateway-CE12808-1-40GE1/0/1] description "to-Leaf-CE6851HI-1& CE6851HI-2"
[*Gateway-CE12808-1-40GE1/0/1] undo portswitch
[*Gateway-CE12808-1-40GE1/0/1] ip address 192.168.40.170 30
[*Gateway-CE12808-1-40GE1/0/1] commit
[~Gateway-CE12808-1-40GE1/0/1] quit
[~Gateway-CE12808-1] interface 40ge 1/0/2
[~Gateway-CE12808-1-40GE1/0/2] description "to-Leaf-CE6851-3-40GE1/0/3"
[*Gateway-CE12808-1-40GE1/0/2] undo portswitch
[*Gateway-CE12808-1-40GE1/0/2] ip address 192.168.41.158 30
[*Gateway-CE12808-1-40GE1/0/2] commit
[~Gateway-CE12808-1-40GE1/0/2] quit
[~Gateway-CE12808-1] interface 40ge 1/0/3
[~Gateway-CE12808-1-40GE1/0/3] description "to-Leaf-CE6851-4-40GE1/0/4"
[*Gateway-CE12808-1-40GE1/0/3] undo portswitch
[*Gateway-CE12808-1-40GE1/0/3] ip address 192.168.41.170 30
[*Gateway-CE12808-1-40GE1/0/3] commit
[~Gateway-CE12808-1-40GE1/0/3] quit
[~Gateway-CE12808-1] interface 40ge 1/0/6
[~Gateway-CE12808-1-40GE1/0/6] undo portswitch
[*Gateway-CE12808-1-40GE1/0/6] ip address 192.168.46.158 30
[*Gateway-CE12808-1-40GE1/0/6] commit
[~Gateway-CE12808-1-40GE1/0/6] quit
[~Gateway-CE12808-1] interface 40ge 1/0/7
[~Gateway-CE12808-1-40GE1/0/7] undo portswitch
[*Gateway-CE12808-1-40GE1/0/7] ip address 192.168.46.170 30
[*Gateway-CE12808-1-40GE1/0/7] commit
[~Gateway-CE12808-1-40GE1/0/7] quit
```

步骤2 在Gateway-CE12808-2节点上配置接口IP地址。

```
[~Huawei] sysname Gateway-CE12808-2
[*Huawei] commit
[~Gateway-CE12808-2] interface 40ge 1/0/0
[~Gateway-CE12808-2-40GE1/0/0] undo portswitch
[*Gateway-CE12808-2-40GE1/0/0] ip address 192.168.40.162 30
[*Gateway-CE12808-2-40GE1/0/0] commit
[~Gateway-CE12808-2-40GE1/0/0] quit
[~Gateway-CE12808-2] interface 40ge 1/0/1
[~Gateway-CE12808-2-40GE1/0/1] undo portswitch
[*Gateway-CE12808-2-40GE1/0/1] ip address 192.168.40.166 30
[*Gateway-CE12808-2-40GE1/0/1] commit
[~Gateway-CE12808-2-40GE1/0/1] quit
[~Gateway-CE12808-2] interface 40ge 1/0/2
[~Gateway-CE12808-2-40GE1/0/2] description "to-Leaf-CE6851-4-40GE1/0/3"
[*Gateway-CE12808-2-40GE1/0/2] undo portswitch
[*Gateway-CE12808-2-40GE1/0/2] ip address 192.168.41.162 30
[*Gateway-CE12808-2-40GE1/0/2] commit
[~Gateway-CE12808-2-40GE1/0/2] quit
[~Gateway-CE12808-2] interface 40ge 1/0/3
[~Gateway-CE12808-2-40GE1/0/3] description "to-Leaf-CE6851-3-40GE1/0/4"
[*Gateway-CE12808-2-40GE1/0/3] undo portswitch
[*Gateway-CE12808-2-40GE1/0/3] ip address 192.168.41.166 30
[*Gateway-CE12808-2-40GE1/0/3] commit
[~Gateway-CE12808-2-40GE1/0/3] quit
[~Gateway-CE12808-2] interface 40ge 1/0/6
[~Gateway-CE12808-2-40GE1/0/6] undo portswitch
[*Gateway-CE12808-2-40GE1/0/6] ip address 192.168.46.166 30
[*Gateway-CE12808-2-40GE1/0/6] commit
[~Gateway-CE12808-2-40GE1/0/6] quit
[~Gateway-CE12808-2] interface 40ge 1/0/7
[~Gateway-CE12808-2-40GE1/0/7] undo portswitch
[*Gateway-CE12808-2-40GE1/0/7] ip address 192.168.46.162 30
[*Gateway-CE12808-2-40GE1/0/7] commit
[~Gateway-CE12808-2-40GE1/0/7] quit
```

----结束

3.1.2.2 配置 Leaf 上的路由

在本方案中,Spine节点和Gateway融合部署,因此Leaf上的BGP路由配置与三层架构组网所有不同。

步骤1 在堆叠组上配置BGP路由,对接Gateway。

```
[~Leaf-CE6851HI-1&CE6851HI-2] bqp 65021
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] router-id 11.11.11.11
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] timer keepalive 10 hold 30
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] group Gateway-CE12808 external
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] peer Gateway-CE12808 as-number 65000
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] peer 192.168.40.158 as-number 65000
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] peer 192.168.40.158 group Gateway-CE12808
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] peer 192.168.40.170 as-number 65000
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] peer 192.168.40.170 group Gateway-CE12808
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] peer 192.168.40.166 as-number 65000
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] peer 192.168.40.166 group Gateway-CE12808
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] peer 192.168.40.162 as-number 65000
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] peer 192.168.40.162 group Gateway-CE12808
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] ipv4-family unicast
[*Leaf-CE6851HI-1&CE6851HI-2-bgp-af-ipv4] preference 20 200 10
[*Leaf-CE6851HI-1&CE6851HI-2-bgp-af-ipv4] network 11.11.11.11 255.255.255.255
[*Leaf-CE6851HI-1&CE6851HI-2-bgp-af-ipv4] maximum load-balancing 32
[*Leaf-CE6851HI-1&CE6851HI-2-bgp-af-ipv4] quit
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] quit
[*Leaf-CE6851HI-1&CE6851HI-2] commit
```

步骤2 在Leaf-CE6851HI-3上配置BGP路由,对接Gateway。

```
[~Leaf-CE6851HI-3] bgp 65022
[*Leaf-CE6851HI-3-bgp] router-id 13.13.13.13
[*Leaf-CE6851HI-3-bgp] timer keepalive 10 hold 30
[*Leaf-CE6851HI-3-bgp] group Gateway-CE12808 external
[*Leaf-CE6851HI-3-bgp] peer Gateway-CE12808 as-number 65000
[*Leaf-CE6851HI-3-bgp] peer 192.168.41.158 as-number 65000
[*Leaf-CE6851HI-3-bgp] peer 192.168.41.158 group Gateway-CE12808
[*Leaf-CE6851HI-3-bgp] peer 192.168.41.166 as-number 65000
[*Leaf-CE6851HI-3-bgp] peer 192.168.41.166 group Gateway-CE12808
[*Leaf-CE6851HI-3-bgp] ipv4-family unicast
[*Leaf-CE6851HI-3-bgp-af-ipv4] preference 20 200 10
[*Leaf-CE6851HI-3-bgp-af-ipv4] network 11.11.11.12 255.255.255.255
[*Leaf-CE6851HI-3-bgp-af-ipv4] network 10.10.100.0 255.255.255.0
[*Leaf-CE6851HI-3-bgp-af-ipv4] network 13.13.13.13 255.255.255.255
[*Leaf-CE6851HI-3-bgp-af-ipv4] maximum load-balancing 32
[*Leaf-CE6851HI-3-bgp-af-ipv4] quit
[*Leaf-CE6851HI-3-bgp] quit
[*Leaf-CE6851HI-3] commit
```

步骤3 在Leaf-CE6851HI-4上配置BGP路由,对接Gateway。

```
[~Leaf-CE6851HI-4] bgp 65022
[*Leaf-CE6851HI-4-bgp] router-id 14.14.14.14
[*Leaf-CE6851HI-4-bgp] timer keepalive 10 hold 30
[*Leaf-CE6851HI-4-bgp] group Gateway-CE12808 external
[*Leaf-CE6851HI-4-bgp] peer Gateway-CE12808 as-number 65000
[*Leaf-CE6851HI-4-bgp] peer 192.168.41.170 as-number 65000
[*Leaf-CE6851HI-4-bgp] peer 192.168.41.170 group Gateway-CE12808
[*Leaf-CE6851HI-4-bgp] peer 192.168.41.162 as-number 65000
[*Leaf-CE6851HI-4-bgp] peer 192.168.41.162 group Gateway-CE12808
[*Leaf-CE6851HI-4-bgp] ipv4-family unicast
[*Leaf-CE6851HI-4-bgp-af-ipv4] preference 20 200 10
[*Leaf-CE6851HI-4-bgp-af-ipv4] network 11.11.11.12 255.255.255.255
[*Leaf-CE6851HI-4-bgp-af-ipv4] network 10.10.100.0 255.255.255.0
[*Leaf-CE6851HI-4-bgp-af-ipv4] network 14.14.14.14 255.255.255.255
[*Leaf-CE6851HI-4-bgp-af-ipv4] maximum load-balancing 32
[*Leaf-CE6851HI-4-bgp-af-ipv4] quit
[*Leaf-CE6851HI-4-bgp] quit
[*Leaf-CE6851HI-4] commit
```

步骤4 在SVF组上配置BGP路由,对接Gateway。

```
[~Leaf-CE6851HI-5&CE6851HI-6] bgp 65024
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] router-id 11.11.11.17
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] timer keepalive 10 hold 30
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] group Gateway-CE12808 external
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] peer Gateway-CE12808 as-number 65000
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] peer 192.168.46.158 as-number 65000
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] peer 192.168.46.158 group Gateway-CE12808
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] peer 192.168.46.170 as-number 65000
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] peer 192.168.46.170 group Gateway-CE12808
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] peer 192.168.46.166 as-number 65000
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] peer 192.168.46.166 group Gateway-CE12808
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] peer 192.168.46.162 as-number 65000
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] peer 192.168.46.162 group Gateway-CE12808
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] ipv4-family unicast
[*Leaf-CE6851HI-5&CE6851HI-6-bgp-af-ipv4] preference 20 200 10
[*Leaf-CE6851HI-5&CE6851HI-6-bgp-af-ipv4] network 11.11.11.17 255.255.255.255
[*Leaf-CE6851HI-5&CE6851HI-6-bgp-af-ipv4] maximum load-balancing 32
[*Leaf-CE6851HI-5&CE6851HI-6-bgp-af-ipv4] quit
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] quit
[*Leaf-CE6851HI-5&CE6851HI-6] commit
```

----结束

3.1.2.3 配置网关上的路由

在本方案中,Spine节点和Gateway融合部署,因此Gateway上的BGP路由配置与三层架构组网所有不同。

步骤1 在Gateway-CE12808-1上配置BGP路由。

```
[~Gateway-CE12808-1] bgp 65000
[*Gateway-CE12808-1-bgp] router-id 18.18.18.18
[*Gateway-CE12808-1-bgp] timer keepalive 10 hold 30
                                                    //对接出口路由器
[*Gateway-CE12808-1-bgp] group Router-1 external
[*Gateway-CE12808-1-bgp] peer Router-1 as-number 65047
[*Gateway-CE12808-1-bgp] peer Router-1 ebgp-max-hop 10
[*Gateway-CE12808-1-bgp] peer Router-1 connect-interface loopback 2
[*Gateway-CE12808-1-bgp] peer 21.21.21.22 as-number 65047
[*Gateway-CE12808-1-bgp] peer 21.21.21.22 group Router-1
[*Gateway-CE12808-1-bgp] group Leaf-CE6851HI-1&CE6851HI-2 external //对接Leaf
[*Gateway-CE12808-1-bgp] peer Leaf-CE6851HI-1&CE6851HI-2 as-number 65021
[*Gateway-CE12808-1-bgp] peer 192.168.40.157 as-number 65021
[*Gateway-CE12808-1-bgp] peer 192.168.40.157 group Leaf-CE6851HI-1&CE6851HI-2
[*Gateway-CE12808-1-bgp] peer 192.168.40.169 as-number 65021
[*Gateway-CE12808-1-bgp] peer 192.168.40.169 group Leaf-CE6851HI-1&CE6851HI-2
[*Gateway-CE12808-1-bgp] group Leaf-CE6851HI-3&4 external
[*Gateway-CE12808-1-bgp] peer Leaf-CE6851HI-3&4 as-number 65022
[*Gateway-CE12808-1-bgp] peer 192.168.41.157 as-number 65022
[*Gateway-CE12808-1-bgp] peer 192.168.41.157 group Leaf-CE6851HI-3&4
[*Gateway-CE12808-1-bgp] peer 192.168.41.169 as-number 65022
[*Gateway-CE12808-1-bgp] peer 192.168.41.169 group Leaf-CE6851HI-3&4
[*Gateway-CE12808-1-bgp] group Leaf-CE6851HI-5&CE6851HI-6 external
[*Gateway-CE12808-1-bgp] peer Leaf-CE6851HI-5&CE6851HI-6 as-number 65024
[*Gateway-CE12808-1-bgp] peer 192.168.46.157 as-number 65024
[*Gateway-CE12808-1-bgp] peer 192.168.46.157 group Leaf-CE6851HI-5&CE6851HI-6
[*Gateway-CE12808-1-bgp] peer 192.168.46.169 as-number 65024
[*Gateway-CE12808-1-bgp] peer 192.168.46.169 group Leaf-CE6851HI-5&CE6851HI-6
[*Gateway-CE12808-1-bgp] ipv4-family unicast
[*Gateway-CE12808-1-bgp-af-ipv4] preference 20 200 10
[*Gateway-CE12808-1-bgp-af-ipv4] network 11.11.11.16 255.255.255.255
[*Gateway-CE12808-1-bgp-af-ipv4] network 18.18.18.18 255.255.255.255
[*Gateway-CE12808-1-bgp-af-ipv4] network 21.21.21.21 255.255.255.255
[*Gateway-CE12808-1-bgp-af-ipv4] network 192.168.44.156 255.255.255.252
[*Gateway-CE12808-1-bgp-af-ipv4] network 192.168.45.152 255.255.255.248
[*Gateway-CE12808-1-bgp-af-ipv4] maximum load-balancing 32
[*Gateway-CE12808-1-bgp-af-ipv4] quit
[*Gateway-CE12808-1-bgp] quit
[*Gateway-CE12808-1] commit
```

步骤2 在Gateway-CE12808-2上配置BGP路由。

```
[~Gateway-CE12808-2] bgp 65000
[*Gateway-CE12808-2-bgp] router-id 19.19.19.19
[*Gateway-CE12808-2-bgp] timer keepalive 10 hold 30
[*Gateway-CE12808-2-bgp] group Router-2 external //对接出口路由器
[*Gateway-CE12808-2-bgp] peer Router-2 as-number 65048
[*Gateway-CE12808-2-bgp] peer Router-2 ebgp-max-hop 10
[*Gateway-CE12808-2-bgp] peer Router-2 connect-interface loopback 2
[*Gateway-CE12808-2-bgp] peer 22.22.23 as-number 65048
[*Gateway-CE12808-2-bgp] peer 22.22.23 group Router-2
[*Gateway-CE12808-2-bgp] group Leaf-CE6851HI-1&CE6851HI-2 external
[*Gateway-CE12808-2-bgp] peer Leaf-CE6851HI-1&CE6851HI-2 as-number 65021
[*Gateway-CE12808-2-bgp] peer 192.168.40.165 as-number 65021
[*Gateway-CE12808-2-bgp] peer 192.168.40.165 group Leaf-CE6851HI-1&CE6851HI-2
[*Gateway-CE12808-2-bgp] peer 192.168.40.161 as-number 65021
[*Gateway-CE12808-2-bgp] peer 192.168.40.161 group Leaf-CE6851HI-1&CE6851HI-2
[*Gateway-CE12808-2-bgp] group Leaf-CE6851HI-3&4 external
[*Gateway-CE12808-2-bgp] peer Leaf-CE6851HI-3&4 as-number 65022
[*Gateway-CE12808-2-bgp] peer 192.168.41.165 as-number 65022
[*Gateway-CE12808-2-bgp] peer 192.168.41.165 group Leaf-CE6851HI-3&4
[*Gateway-CE12808-2-bgp] peer 192.168.41.161 as-number 65022
[*Gateway-CE12808-2-bgp] peer 192.168.41.161 group Leaf-CE6851HI-3&4
[*Gateway-CE12808-2-bgp] group Leaf-CE6851HI-5&CE6851HI-6 external
[*Gateway-CE12808-2-bgp] peer Leaf-CE6851HI-5&CE6851HI-6 as-number 65024
[*Gateway-CE12808-2-bgp] peer 192.168.46.165 as-number 65024
[*Gateway-CE12808-2-bgp] peer 192.168.46.165 group Leaf-CE6851HI-5&CE6851HI-6
[*Gateway-CE12808-2-bgp] peer 192.168.46.161 as-number 65024
[*Gateway-CE12808-2-bgp] peer 192.168.46.161 group Leaf-CE6851HI-5&CE6851HI-6
```

```
[*Gateway-CE12808-2-bgp] ipv4-family unicast
[*Gateway-CE12808-2-bgp-af-ipv4] preference 20 200 10
[*Gateway-CE12808-2-bgp-af-ipv4] network 11.11.11.16 255.255.255.255
[*Gateway-CE12808-2-bgp-af-ipv4] network 19.19.19.19 255.255.255.255
[*Gateway-CE12808-2-bgp-af-ipv4] network 22.22.22.22 255.255.255
[*Gateway-CE12808-2-bgp-af-ipv4] network 192.168.44.160 255.255.255.255
[*Gateway-CE12808-2-bgp-af-ipv4] network 192.168.45.152 255.255.255.258
[*Gateway-CE12808-2-bgp-af-ipv4] maximum load-balancing 32
[*Gateway-CE12808-2-bgp-af-ipv4] quit
[*Gateway-CE12808-2-bgp] quit
[*Gateway-CE12808-2] commit
```

----结束

3.1.2.4 配置 BGP EVPN

在本方案中,Spine节点和Gateway融合部署,当使用BGP EVPN作为VXLAN的控制平面时,不再需要网关与Spine之间建立对等体关系,而是网关和Leaf间建立对等体。

步骤1 在集中式三层网关上使能EVPN作为VXLAN的控制平面功能,其他设备上的配置与之类似,这里不再赘述。

```
[~Gateway-CE12808-1] evpn-overlay enable
[~Gateway-CE12808-1] commit
```

步骤2 配置集中式三层网关分别与Leaf节点建立BGP EVPN对等体关系。

在Gateway-CE12808-1上配置与Leaf建立IBGP EVPN对等体关系,其他设备与之类似,不再赘述。

```
[~Gateway-CE12808-1] bgp 100 instance evpn1
[*Gateway-CE12808-1-bgp-instance-evpn1] router-id 18.18.18.18
[*Gateway-CE12808-1-bgp-instance-evpn1] peer 11.11.11.11 as-number 100
//与堆叠组建立BGP EVPN对等体关系
[*Gateway-CE12808-1-bgp-instance-evpn1] peer 11.11.11.11 connect-interface loopback 0
[*Gateway-CE12808-1-bgp-instance-evpn1] peer 13.13.13.13 as-number 100
//与Leaf-CE6851HI-3建立BGP EVPN对等体关系
[*Gateway-CE12808-1-bgp-instance-evpn1] peer 13.13.13.13 connect-interface loopback 0
[*Gateway-CE12808-1-bgp-instance-evpn1] peer 14.14.14.14 as-number 100
//与Leaf-CE6851HI-4建立BGP EVPN对等体关系
[*Gateway-CE12808-1-bgp-instance-evpn1] peer 14.14.14 connect-interface loopback 0
[*Gateway-CE12808-1-bgp-instance-evpn1] peer 11.11.11.17 as-number 100
//与SVF组建立BGP EVPN对等体关系
[*Gateway-CE12808-1-bgp-instance-evpn1] peer 11.11.11.17 connect-interface loopback 0
[*Gateway-CE12808-1-bgp-instance-evpn1] l2vpn-family evpn
[*Gateway-CE12808-1-bgp-instance-evpn1-af-evpn] peer 11.11.11.11 enable
[*Gateway-CE12808-1-bgp-instance-evpn1-af-evpn] peer 13.13.13.13 enable
[*Gateway-CE12808-1-bgp-instance-evpn1-af-evpn] peer 14.14.14.14 enable
[*Gateway-CE12808-1-bgp-instance-evpn1-af-evpn] peer 11.11.11.17 enable
[*Gateway-CE12808-1-bgp-instance-evpn1-af-evpn] quit
[*Gateway-CE12808-1-bgp-instance-evpn1] quit
[*Gateway-CE12808-1] commit
```

----结束

3.2 硬件分布式 VXLAN

硬件分布式VXLAN是VXLAN组网的一种类型,"分布式"是指VXLAN L3网关分散部署在所有Leaf节点交换机上,这样可使接入网络计算节点的VXLAN L3网关部署在就近的Leaf节点上;"硬件"是指硬件交换机作NVE节点。硬件分布式VXLAN根据是否有Border Leaf交换机作为网关,VXLAN的架构可以分为Spine/Leaf二层架构和GW/Spine/Leaf三层架构。其中Spine/Leaf二层架构方案流量模型简单,维护简便,但扩展比较困难,适用于中小型网络场景;GW/Spine/Leaf三层架构方案架构灵活,网络扩展能力强,但流量路径和维护相对复杂,适用于大中型网络场景。

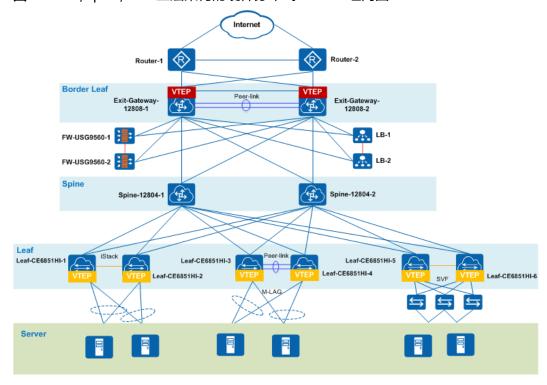
3 VXLAN 部署建议

3.2.1 GW/Spine/Leaf 三层架构的硬件分布式 VXLAN

组网需求

GW/Spine/Leaf三层架构的硬件分布式VXLAN组网图如图1所示,

图 3-3 GW/Spine/Leaf 三层架构的硬件分布式 VXLAN 组网图



- 服务器层:通过二层子接口方式接入VXLAN网络。
- Leaf(分布式网关,也称东西网关):采用堆叠方式、M-LAG或SVF方式连接服务器。Leaf节点和Spine节点通过三层互联。Leaf的堆叠工作组、M-LAG工作组或者SVF工作组作为VTEP实现服务器流量接入VXLAN网络。
- Spine: 分别与Leaf节点和GW互联,运行路由协议保证Underlay三层可达。Spine 节点不作为VTEP。
- Border Leaf(出口三层网关/Exit Gateway,也称南北网关):两台Border Leaf 之间组成M-LAG,即双活网关。Border Leaf与Spine之间三层互联,对外连接外 部路由器Router-1和Router-2。
- FW: 两台FW分别旁挂在GW, FW之间配置为主备镜像模式工作。
- LB:由厂商部署。

端口连线规划

Leaf-CE6851HI-1和Leaf-CE6851HI-2部署堆叠场景,端口规划参见下表。

规划描述	本端设备名 称	物理端口	对端设备名称	物理端口
规划Leaf间互连端口 作为堆叠口,堆叠链	Leaf- CE6851HI-1	40GE1/0/1~ 2	Leaf- CE6851HI-2	40GE1/0/1 ~2
路至少2条	Leaf- CE6851HI-2	40GE1/0/1~ 2	Leaf- CE6851HI-1	40GE1/0/1 ~2
规划Leaf与Spine互连 端口	Leaf- CE6851HI-1	40GE1/0/3	Spine- CE12804-1	40GE1/0/0
	Leaf- CE6851HI-1	40GE1/0/4	Spine- CE12804-2	40GE1/0/1
	Leaf- CE6851HI-2	40GE1/0/3	Spine- CE12804-2	40GE1/0/0
	Leaf- CE6851HI-2	40GE1/0/4	Spine- CE12804-1	40GE1/0/1
规划Leaf与服务器的 互连端口,服务器双	Leaf- CE6851HI-1	10GE1/0/1~ 2	服务器	Eth0
网卡以负载分担方式 上连,保证可靠性的 同时负载分担转发, 提高链路利用率	Leaf- CE6851HI-2	10GE1/0/1~ 2	服务器	Eth1

Leaf-CE6851HI-3和Leaf-CE6851HI-4部署M-LAG,端口规划参见下表。

规划描述	本端设备名 称	物理端口	对端设备名称	物理端口
规划M-LAG的peer-link端口,用于承载	Leaf- CE6851HI-3	40GE1/0/1~2	Leaf- CE6851HI-4	40GE1/0/1 ~2
M-LAG的协议报文和 故障场景下的数据报 文,为保证可靠性, 建议至少2条peer-link 链路	Leaf- CE6851HI-4	40GE1/0/1~2	Leaf- CE6851HI-3	40GE1/0/1 ~2
规划Leaf连接Spine互 连端口	Leaf- CE6851HI-3	40GE1/0/3	Spine- CE12804-1	40GE1/0/2
	Leaf- CE6851HI-3	40GE1/0/4	Spine- CE12804-2	40GE1/0/3
	Leaf- CE6851HI-4	40GE1/0/3	Spine- CE12804-2	40GE1/0/2
	Leaf- CE6851HI-4	40GE1/0/4	Spine- CE12804-1	40GE1/0/3

规划描述	本端设备名 称	物理端口	对端设备名称	物理端口
规划Leaf与服务器的 互连端口,服务器双	Leaf- CE6851HI-3	10GE1/0/1	服务器	Eth0
网卡以负载分担方式 上连,保证可靠性的 同时负载分担转发, 提高链路利用率	Leaf- CE6851HI-4	10GE1/0/1	服务器	Eth1

Leaf-CE6851HI-5和Leaf-CE6851HI-6部署SVF场景,下连三台CE5810交换机,端口规划参见下表。

规划描述	本端设备名 称	物理端口	对端设备名称	物理端口
规划Leaf间互连端口 作为堆叠口,堆叠链	Leaf- CE6851HI-5	40GE1/0/1~ 2	Leaf- CE6851HI-6	40GE1/0/1 ~2
路至少2条	Leaf- CE6851HI-6	40GE1/0/1~ 2	Leaf- CE6851HI-5	40GE1/0/1 ~2
规划Leaf与Spine互连 端口	Leaf- CE6851HI-5	40GE1/0/3	Spine- CE12804-1	40GE1/0/6
	Leaf- CE6851HI-5	40GE1/0/4	Spine- CE12804-2	40GE1/0/7
	Leaf- CE6851HI-6	40GE1/0/3	Spine- CE12804-1	40GE1/0/6
	Leaf- CE6851HI-6	40GE1/0/4	Spine- CE12804-2	40GE1/0/7
规划SVF父节点与叶	Leaf- CE6851HI-5	10GE1/0/1	CE5800-1	GE1/0/1
子节点互联端口 		10GE1/0/2	CE5800-2	GE1/0/1
		10GE1/0/3	CE5800-3	GE1/0/1
	Leaf-	10GE1/0/1	CE5800-1	GE1/0/2
	CE6851HI-6	10GE1/0/2	CE5800-2	GE1/0/2
		10GE1/0/3	CE5800-3	GE1/0/2
规划Leaf与服务器的	CE5810-1	GE1/0/1	服务器	Eth0
互连端口,服务器双 网卡以负载分担方式 上连,保证可靠性的 同时负载分担转发, 提高链路利用率	CE5810-2	GE1/0/1	服务器	Eth1

Spine-CE12804-1& Spine-CE12804-2端口规划参见下表。

规划描述	本端设备名 称	物理端口	对端设备名称	物理端口
规划Spine与Leaf的互 连端口	Spine- CE12804-1	40GE1/0/0~3 40GE1/0/6~7	Leaf- CE6851HI-1 Leaf- CE6851HI-2 Leaf- CE6851HI-3 Leaf- CE6851HI-4 Leaf- CE6851HI-5 Leaf- CE6851HI-5	40GE1/0/3 ~4
	Spine- CE12804-2	40GE1/0/0~3 40GE1/0/6~7	Leaf- CE6851HI-1 Leaf- CE6851HI-2 Leaf- CE6851HI-3 Leaf- CE6851HI-4 Leaf- CE6851HI-5 Leaf- CE6851HI-5	40GE1/0/3 ~4
规划Spine与Exit- Gateway的互连端口	Spine- CE12804-1	40GE1/0/4~5	Exit- Gateway- CE12808-1 Exit- Gateway- CE12808-2	40GE1/0/0
	Spine- CE12804-2	40GE1/0/4~5	Exit- Gateway- CE12808-1 Exit- Gateway- CE12808-2	40GE1/0/1

Exit-Gateway-CE12808-1& Exit-Gateway-CE12808-2端口规划参见下表。

规划描述	本端设备名 称	物理端口	对端设备名称	物理端口
规划Exit-Gateway与 Spine互连端口	Exit- Gateway- CE12808-1	40GE1/0/0~1	Spine- CE12804-1 Spine- CE12804-2	40GE1/0/ 4
	Exit- Gateway- CE12808-2	40GE1/0/0~1	Spine- CE12804-1 Spine- CE12804-2	40GE1/0/ 5
规划M-LAG的peer- link端口,用于承载M- LAG的协议报文和故障 场景下的数据报文,为 保证可靠性,建议至少	Exit- Gateway- CE12808-1	40GE1/0/23 40GE2/0/23	Exit- Gateway- CE12808-2	40GE1/0/ 23 40GE2/0/ 23
操证可靠性,建议主义 2条peer-link链路,且 跨板部署	Exit- Gateway- CE12808-2	40GE1/0/23 40GE2/0/23	Exit- Gateway- CE12808-1	40GE1/0/ 23 40GE2/0/ 23
规划Exit-Gateway与防 火墙的互连端口 说明 两台主备镜像模式防火 墙必须使用同样的接口	Exit- Gateway- CE12808-1	10GE3/0/0~1	FW- USG9560-1 FW- USG9560-2	GE1/0/1
墙必须使用同样的接口 连接到同一个GW设备。 例如FW-1使用GE 1/0/1 接口连接GW-1,那么 FW-2也需要使用GE 1/0/1接口连接GW-1。	Exit- Gateway- CE12808-1	10GE3/0/2~3	FW- USG9560-1 FW- USG9560-2	GE1/0/3
	Exit- Gateway- CE12808-2	10GE3/0/0~1	FW- USG9560-1 FW- USG9560-2	GE1/0/2
	Exit- Gateway- CE12808-2	10GE3/0/2~3	FW- USG9560-1 FW- USG9560-2	GE1/0/4
规划Exit-Gateway与出 口路由器的互连端口	Exit- Gateway- CE12808-1	10GE3/0/4	Router-1	GE1/0/0
	Exit- Gateway- CE12808-1	10GE3/0/5	Router-2	GE1/0/0

规划描述	本端设备名 称	物理端口	对端设备名称	物理端口
	Exit- Gateway- CE12808-2	10GE3/0/4	Router-1	GE1/0/1
	Exit- Gateway- CE12808-2	10GE3/0/5	Router-2	GE1/0/1
规划Exit-Gateway之间 的三层互联接口,用于 和出口路由器组成 "口"字型和交叉联合 出口网络	Exit- Gateway- CE12808-1	10GE3/0/6	Exit- Gateway- CE12808-2	10GE3/0/ 6

VLAN 规划

VLAN规划参见下表。

规划描述	规划建议	本示例使用ID
与防火墙 互连 VLAN	三层互连建议使用三层主接口,与FW因涉 及多个网段对接,需要规划VLAN,达到共 用物理链路目的	11 12(创建VLANIF 12绑定 VPN)
租户 VLAN	租户/业务使用VXLAN较多,需要考虑后续 的业务扩容	10

□ 说明

对于CE6855HI和CE7855EI,在切换三层接口之前,要先配置vlan reserved for maininterface startvlanid to endvlanid,配置三层主接口专用的保留VLAN。

BD、VNI 规划

BD、VNI的规划参见下表。

规划描述	规划建议	本示例使用ID
BD	规划数量与VLAN相同,建议BD的ID与 VLAN的ID相同	本例统一使用BD 10
VNI	二层VNI规划数量与BD相同,建议VNI的ID规划为"BD+10000"(BD与VNI——对应) 三层VNI数量和L3VPN相同。	本例统一使用二层VNI 10010,三层VNI为10

RD 和 RT 值规划

RD的规划参见下表。

规划描述	规划建议	本示例使用ID
CE6851HI- 1、 CE6851HI- 2的RD值	分布式网关需要配置EVPN实例和VPN实 例的RD值,且需要保持全局唯一。	11: 1 12: 1 (VPN实例)
CE6851HI- 3的RD值	分布式网关需要配置EVPN实例和VPN实 例的RD值,且需要保持全局唯一。	13: 1 12: 2 (VPN实例)
CE6851HI- 4的RD值	分布式网关需要配置EVPN实例和VPN实 例的RD值,且需要保持全局唯一。	14: 1 12: 3 (VPN实例)
CE6851HI- 5、 CE6851HI- 6的RD值	分布式网关需要配置EVPN实例和VPN实 例的RD值,且需要保持全局唯一。	15: 1 12: 4(VPN实例)
Exit- Gateway- CE12808-1 的RD值	分布式网关需要配置EVPN实例和VPN实 例的RD值,且需要保持全局唯一。	16: 1
Exit- Gateway- CE12808-2 的RD值	分布式网关需要配置EVPN实例和VPN实 例的RD值,且需要保持全局唯一。	17: 1

RT的规划参见下表。

设备	规划建议	本示例使用ID
CE6851HI- 1、 CE6851HI- 2的RT值	分布式网关需要配置EVPN实例和VPN实例的RT值。	1: 1 10: 1 11: 1 (EVPN实例出方向 RT值,用于与VPN实例交 叉) 11: 1 (VPN实例的RT 值,用于与EVPN实例交 叉)

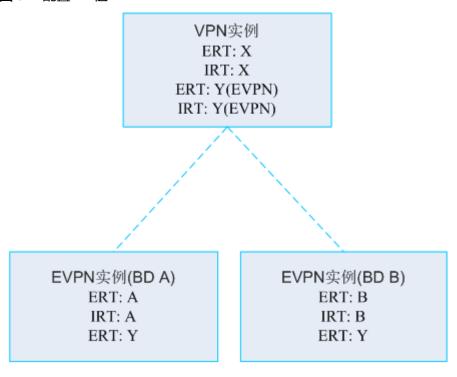
设备	规划建议	本示例使用ID
CE6851HI- 3的RT值	分布式网关需要配置EVPN实例和VPN实例的RT值。	1: 1 10: 1 11: 1 (EVPN实例出方向 RT值,用于与VPN实例交 叉) 11: 1 (VPN实例的RT 值,用于与EVPN实例交 叉)
CE6851HI- 4的RT值	分布式网关需要配置EVPN实例和VPN实例的RT值。	1: 1 10: 1 11: 1 (EVPN实例出方向 RT值,用于与VPN实例交 叉) 11: 1 (VPN实例的RT 值,用于与EVPN实例交 叉)
CE6851HI- 5、 CE6851HI- 6的RT值	分布式网关需要配置EVPN实例和VPN实例的RT值。	1: 1 10: 1 11: 1 (EVPN实例出方向 RT值,用于与VPN实例交 叉) 11: 1 (VPN实例的RT 值,用于与EVPN实例交 叉)
Exit- Gateway- CE12808-1 的RT值	分布式网关需要配置EVPN实例和VPN实 例的RT值。	1: 1 11: 1 (EVPN实例出方向 RT值,用于与VPN实例交 叉) 11: 1 (VPN实例的RT 值,用于与EVPN实例交 叉)
Exit- Gateway- CE12808-2 的RT值	分布式网关需要配置EVPN实例和VPN实例的RT值。	1: 1 11: 1 (EVPN实例出方向 RT值,用于与VPN实例交 叉) 11: 1 (VPN实例的RT 值,用于与EVPN实例交 叉)

VPN实例和EVPN实例的RT值配置原则如图2所示,其中:

● 在VPN实例中,除了配置本地VPN实例的ERT X和IRT X外,还需要配置带EVPN参数的ERT Y和IRT Y,用于和EVPN实例交叉,生成主机路由。

● 在EVPN实例中,除了针对不同的BD配置不同的ERT A、ERT B和IRT A、IRT B外,还必须配置用来和VPN实例交叉的ERT Y。一般情况下不需要配置IRT Y,否则会导致不同BD下的EVPN实例相互扩散MAC地址。

图 3-4 配置 RT 值



IP 地址规划

网元接口地址规划参见下表,涉及互连地址、VTEP地址、BGP Route-ID、BGP建邻居 loopback地址、M-LAG心跳地址、业务网段地址。

端口互连IP地址规划参见下表。

规划描述	本端设备	对端设备	IP/MASK
CE6851HI-2与 C	Leaf- CE6851HI-1&	Spine-CE12804-1	192.168.40.156/30 192.168.40.168/30
Spine互连端口地 址 	3连端口地 CE6851HI-2	Spine-CE12804-2	192.168.40.164/30 192.168.40.160/30
规划CE6851HI-3、 Leaf- CE6851HI-4与 CE6851HI-3& Spine互连端口地 CE6851HI-4	CE6851HI-3&	Spine-CE12804-1	192.168.41.156/30 192.168.41.168/30
址	CE6851HI-4	Spine-CE12804-2	192.168.41.164/30 192.168.41.160/30
规划CE6851HI-5、 CE6851HI-6与 Spine互连端口地 址	Leaf- CE6851HI-5& CE6851HI-6	Spine-CE12804-1	192.168.46.156/30 192.168.46.168/30

规划描述	本端设备	对端设备	IP/MASK
		Spine-CE12804-2	192.168.46.164/30
			192.168.46.160/30
规划Spine- CE12804-1与Exit-	Spine-CE12804-1	Exit-Gateway- CE12808-1	192.168.42.156/30
Gateway- CE12808-1/Exit- Gateway- CE12808-2互连端 口地址		Exit-Gateway- CE12808-2	192.168.42.160/30
规划Spine- CE12804-2与Exit-	Spine-CE12804-2	Exit-Gateway- CE12808-1	192.168.43.156/30
Gateway- CE12808-1/Exit- Gateway- CE12808-2互连端 口地址		Exit-Gateway- CE12808-2	192.168.43.160/30
规划Exit-Gateway- CE12808-1与 Router-1互连端口 地址	Exit-Gateway- CE12808-1	Router-1	192.168.44.156/30
规划Exit-Gateway- CE12808-2与 Router-2互连端口 地址	Exit-Gateway- CE12808-2	Router-2	192.168.44.160/30
规划Exit-Gateway- CE12808-1和Exit- Gateway- CE12808-2互联	Exit-Gateway- CE12808-1	Exit-Gateway- CE12808-2	192.168.44.164/30
规划Exit-Gateway- CE12808-1&2与 FW互连端口地址	Exit-Gateway- CE12808-1 Exit-Gateway- CE12808-2	FW-USG9560-1 FW-USG9560-2	192.168.45.152/29 Virtual IP: 192.168.45.153 192.168.45.160/29
			Virtual IP: 192.168.45.161

Loopback地址规划参见下表。

规划描述	设备	IP/MASK
规划Loopback0地址,作 为VTEP地址	Leaf-CE6851HI-1& CE6851HI-2	11.11.11.11/32
	Leaf-CE6851HI-3	11.11.11.12/32

规划描述	设备	IP/MASK
说明 Leaf-CE6851HI-3和Leaf- CE6851HI-4组成M-LAG用	Leaf-CE6851HI-4	同Leaf-CE6851HI-3: 11.11.11.12/32
于服务器双活接入,因此它 们的VTEP IP地址相同。 Gateway-CE12808-1和	Leaf-CE6851HI-5& CE6851HI-6	11.11.11.17/32
Gateway-CE12808-2组成 M-LAG用于防火墙双活接	Exit-Gateway-CE12808-1	11.11.11.16/32
入,因此它们的VTEP IP地址相同。	Exit-Gateway-CE12808-2	同Exit-Gateway- CE12808-1: 11.11.11.16/32
规划Spine的Loopback0地	Spine-CE12804-1	11.11.11.14/32
址,作为Router ID	Spine-CE12804-2	11.11.11.15/32
规划Loopback1地址,作	Leaf-CE6851HI-3	13.13.13.13/32
为M-LAG心跳检测地址 	Leaf-CE6851HI-4	14.14.14.14/32
	Exit-Gateway-CE12808-1	18.18.18.18/32
	Exit-Gateway-CE12808-2	19.19.19.19/32
规划Loopback2地址,作	Exit-Gateway-CE12808-1	21.21.21.21/32
为对接Router的EBGP邻居 地址	Exit-Gateway-CE12808-2	22.22.22.22/32

业务地址规划参见下表。

租户	租户IP/MASK	网关VBDIF	VRF
服务器1	192.168.10.2/24	192.168.10.1/24	VPN 1
服务器2	192.168.10.3/24	192.168.10.1/24	VPN 1
服务器3	192.168.10.4/24	192.168.10.1/24	VPN 1

路由规划

Underlay网络使用的路由协议常见为EBGP和OSPF两种,BGP从多方面保证了网络的安全性、灵活性、稳定性、可靠性和高效性:

- BGP采用认证和GTSM的方式,保证了网络的安全性。
- BGP提供了丰富的路由策略,能够灵活的进行路由选路。
- BGP提供了路由聚合和路由衰减功能用于防止路由振荡,有效提高了网络的稳定性。
- BGP使用TCP作为其传输层协议(端口号为179),并支持BGP与BFD联动、BGP Auto FRR和BGP GR和NSR,提高了网络的可靠性。

从网络演进方面考虑,EBGP适用于大型网络,而OSPF适用于中小型网络,本文档使用EBGP作为示例,其中Gateway对接Router的IGP使用OSPF路由作为示例(此处使用ISIS或者IBGP均可以)。

路由相关的规划参见下表。

网元	AS域号	Router ID
Leaf- CE6851HI-1&CE6851-2	65021	Loopback0地址
Leaf-CE6851HI-3	65022	Loopback1地址
Leaf-CE6851HI-4	65022	Loopback1地址
Leaf- CE6851HI-5&CE6851HI-6	65024	Loopback0地址
Spine-CE12804-1	65010	Loopback0地址
Spine-CE12804-2	65010	Loopback0地址
Exit-Gateway-CE12808-1	65000	Loopback1地址
Exit-Gateway-CE12808-2	65000	Loopback1地址
Router-1	65047	Loopback0地址
Router-2	65048	Loopback0地址

3.2.1.1 配置 Leaf 堆叠工作组

两台Leaf交换机Leaf-CE6851HI-1和Leaf-CE6851HI-2组成堆叠组,分别上联Spine-1和Spine-2,下联服务器。

配置思路

- 1. 组建堆叠组:配置堆叠和双主检测,并重启设备、连接线缆,使堆叠生效。
- 2. 配置IP地址:配置Leaf与Spine的三层互联接口地址、Loopback0地址(作为Router-ID和VTEP IP)。
- 3. 配置服务器接入:介绍服务器接入堆叠组时,堆叠组交换机上的配置。
- 4. 配置路由:在堆叠组上配置BGP动态路由,邻居为两个Spine设备,使堆叠组上的网段与Spine上的网段三层可达。
- 5. 配置BGP EVPN:配置BGP EVPN作为VXLAN的控制平面、并配置BGP EVPN对等体、EVPN实例和VPN实例、头端复制功能和三层网关。

组建堆叠组

步骤1 配置Leaf-CE6851HI-1的堆叠成员ID为1,优先级为150,Domain ID为10。

<HUAWEI> system-view
[~HUAWEI] sysname Leaf-CE6851HI-1
[*HUAWEI] commit

[~Leaf-CE6851HI-1] stack

[~Leaf-CE6851HI-1-stack] stack member 1 priority 150 [*Leaf-CE6851HI-1-stack] stack member 1 domain 10

[*Leaf-CE6851HI-1-stack] quit [*Leaf-CE6851HI-1] commit

步骤2 配置Leaf-CE6851HI-2的Domain ID为10。

<HUAWEI> system-view

[~HUAWEI] sysname Leaf-CE6851HI-2

[*HUAWEI] commit

[~Leaf-CE6851HI-2] stack

[~Leaf-CE6851HI-2-stack] stack member 1 renumber 2 inherit-config

[*Leaf-CE6851HI-2-stack] stack member 1 domain 10

[*Leaf-CE6851HI-2-stack] quit

[*Leaf-CE6851HI-2] commit

步骤3 配置堆叠端口。

将Leaf-CE6851HI-1的业务口40GE1/0/1~40GE1/0/2加入堆叠端口1/1。

[~Leaf-CE6851HI-1] interface stack-port 1/1

[*Leaf-CE6851HI-1-Stack-Port1/1] port member-group interface 40ge 1/0/1 to 1/0/2

Warning: After the configuration is complete,1.The interface(s) (40GE1/0/1-1/0/2) will be converted to stack mode and be configured with the port crc-statistics trigger error-down command if the configuration does not exist.2.The interface(s) may go Error-Down (crc-statistics) because there is no shutdown configuration on the interfaces.Continue? [Y/N]: y

[*Leaf-CE6851HI-1-Stack-Port1/1] quit

[*Leaf-CE6851HI-1] commit

[~Leaf-CE6851HI-1] quit

将Leaf-CE6851HI-2的业务口40GE1/0/1~40GE1/0/2加入堆叠端口1/1。

[~Leaf-CE6851HI-2] interface stack-port 1/1

[*Leaf-CE6851HI-2-Stack-Port1/1] port member-group interface 40ge 1/0/1 to 1/0/2

Warning: After the configuration is complete,1.The interface(s) (40GE1/0/1-1/0/2) will be converted to stack mode and be configured with the port crc-statistics trigger error-down command if the configuration does not exist.2.The interface(s) may go Error-Down (crc-statistics) because there is no shutdown configuration on the interfaces.Continue? [Y/N]: y

[*Leaf-CE6851HI-2-Stack-Port1/1] quit

[*Leaf-CE6851HI-2] commit

[~Leaf-CE6851HI-2] quit

步骤4 保存配置后重启设备。

#保存Leaf-CE6851HI-1的配置,然后重启设备。Leaf-CE6851HI-2的操作与之类似,不再赘述。

<Leaf-CE6851HI-1> save

Warning: The current configuration will be written to the device. Continue? [Y/N]: y

<Leaf-CE6851HI-1> reboot

Warning: The system will reboot. Continue? [Y/N]:y

步骤5 连接堆叠线缆,建立堆叠。

步骤6 堆叠组建后执行命令save保存配置。

步骤7 堆叠建立后,配置堆叠组的双主检测,防止堆叠组分裂后网络中存在两个配置冲突的网络设备。

采用业务口直连的方式配置双主检测,建议规划的直连链路至少2条,保证可靠性, 配置示例如下。

[~Leaf-CE6851HI-1] sysname Leaf-CE6851HI-1&CE6851HI-2

[*Leaf-CE6851HI-1] commit

[~Leaf-CE6851HI-1&CE6851HI-2] interface 10ge 1/0/30

[~Leaf-CE6851HI-1&CE6851HI-2-10GE1/0/30] description "for DAD"

[*Leaf-CE6851HI-1&CE6851HI-2-10GE1/0/30] dual-active detect mode direct

Warning: The interface will block common data packets, except BPDU packets. Continue? [Y/N]: y

[*Leaf-CE6851HI-1&CE6851HI-2-10GE1/0/30] quit

[*Leaf-CE6851HI-1&CE6851HI-2] interface 10ge 1/0/31

 $[$^{\text{Leaf-CE6851HI-}1\&CE6851HI-}2-10GE1/0/31]$ \ \textbf{description} \ "for \ \textbf{DAD"}"$

```
[*Leaf-CE6851HI-1&CE6851HI-2-10GE1/0/31] dual-active detect mode direct
Warning: The interface will block common data packets, except BPDU packets. Continue? [Y/N]: y
[*Leaf-CE6851HI-1&CE6851HI-2-10GE1/0/31] quit
[*Leaf-CE6851HI-1&CE6851HI-2] interface 10ge 2/0/30
[*Leaf-CE6851HI-1&CE6851HI-2-10GE2/0/30] description "for DAD"
[*Leaf-CE6851HI-1&CE6851HI-2-10GE2/0/30] dual-active detect mode direct
Warning: The interface will block common data packets, except BPDU packets. Continue? [Y/N]: y
[*Leaf-CE6851HI-1&CE6851HI-2-10GE2/0/30] quit
[*Leaf-CE6851HI-1&CE6851HI-2-10GE2/0/31] description "for DAD"
[*Leaf-CE6851HI-1&CE6851HI-2-10GE2/0/31] dual-active detect mode direct
Warning: The interface will block common data packets, except BPDU packets. Continue? [Y/N]: y
[*Leaf-CE6851HI-1&CE6851HI-2-10GE2/0/31] commit
[-Leaf-CE6851HI-1&CE6851HI-2-10GE2/0/31] quit
```

步骤8 配置作为Server Leaf的堆叠系统的MAC地址为成员设备1的MAC地址。

```
[~Leaf-CE6851HI-1&CE6851HI-2] stack
[~Leaf-CE6851HI-1&CE6851HI-2-stack] set system mac-address slot 1 //如果CE12800系列作为Server
Leaf时,此处配置命令为set system mac-address chassis 1
```

----结束

配置 IP 地址

步骤1 配置互连接口地址。

□ 说明

对于CE6855HI和CE7855EI,在切换三层接口之前,要先配置vlan reserved for maininterface startvlanid to endvlanid,配置三层主接口专用的保留VLAN。

```
[~Leaf-CE6851HI-1&CE6851HI-2] interface 40ge 1/0/3
[~Leaf-CE6851HI-1&CE6851HI-2-40GE1/0/3] description "to_Spine-CE12804-1-40GE1/0/0"
[*Leaf-CE6851HI-1&CE6851HI-2-40GE1/0/3] undo portswitch
[*Leaf-CE6851HI-1&CE6851HI-2-40GE1/0/3] ip address 192.168.40.157 30
[*Leaf-CE6851HI-1&CE6851HI-2-40GE1/0/3] commit
[~Leaf-CE6851HI-1&CE6851HI-2-40GE1/0/3] quit
[~Leaf-CE6851HI-1&CE6851HI-2] interface 40ge 1/0/4
[~Leaf-CE6851HI-1&CE6851HI-2-40GE1/0/4] description "to_Spine-CE12804-2-40GE1/0/1"
[*Leaf-CE6851HI-1&CE6851HI-2-40GE1/0/4] undo portswitch
*Leaf-CE6851HI-1&CE6851HI-2-40GE1/0/4] ip address 192.168.40.165 30
[*Leaf-CE6851HI-1&CE6851HI-2-40GE1/0/4] commit
[~Leaf-CE6851HI-1&CE6851HI-2-40GE1/0/4] quit
[~Leaf-CE6851HI-1&CE6851HI-2] interface 40ge 2/0/3
[~Leaf-CE6851HI-1&CE6851HI-2-40GE2/0/3] description "to_Spine-CE12804-2-40GE1/0/0"
[*Leaf-CE6851HI-1&CE6851HI-2-40GE2/0/3] undo portswitch
[*Leaf-CE6851HI-1&CE6851HI-2-40GE2/0/3] ip address 192.168.40.161 30
[*Leaf-CE6851HI-1&CE6851HI-2-40GE2/0/3] commit
[~Leaf-CE6851HI-1&CE6851HI-2-40GE2/0/3] quit
[~Leaf-CE6851HI-1&CE6851HI-2] interface 40ge 2/0/4
[~Leaf-CE6851HI-1&CE6851HI-2-40GE2/0/4] description "to_Spine-CE12804-1-40GE1/0/1"
[*Leaf-CE6851HI-1&CE6851HI-2-40GE2/0/4] undo portswitch
[*Leaf-CE6851HI-1&CE6851HI-2-40GE2/0/4] ip address 192.168.40.169 30
[*Leaf-CE6851HI-1&CE6851HI-2-40GE2/0/4] commit
[~Leaf-CE6851HI-1&CE6851HI-2-40GE2/0/4] quit
```

步骤2 配置Loopback口地址。

```
[~Leaf-CE6851HI-1&CE6851HI-2] interface loopback 0

[*Leaf-CE6851HI-1&CE6851HI-2-LoopBack0] description VTEP&Router-ID

[*Leaf-CE6851HI-1&CE6851HI-2-LoopBack0] ip address 11.11.11.11 32

[*Leaf-CE6851HI-1&CE6851HI-2-LoopBack0] commit

[~Leaf-CE6851HI-1&CE6851HI-2-LoopBack0] quit
```

----结束

配置服务器接入

介绍服务器接入堆叠组时,堆叠组交换机上的配置。

步骤1 在Leaf-CE6851HI-1与Leaf-CE6851HI-2上配置以负载均衡模式接入堆叠组。

```
[~Leaf-CE6851HI-1&CE6851HI-2] interface eth-trunk 1
[*Leaf-CE6851HI-1&CE6851HI-2-Eth-Trunk1] mode lacp-static
[*Leaf-CE6851HI-1&CE6851HI-2-Eth-Trunk1] port link-type trunk
[*Leaf-CE6851HI-1&CE6851HI-2-Eth-Trunk1] undo port trunk allow-pass vlan 1
[*Leaf-CE6851HI-1&CE6851HI-2-Eth-Trunk1] trunkport 10ge 1/0/1
[*Leaf-CE6851HI-1&CE6851HI-2-Eth-Trunk1] trunkport 10ge 2/0/1
[*Leaf-CE6851HI-1&CE6851HI-2-Eth-Trunk1] quit
[*Leaf-CE6851HI-1&CE6851HI-2] commit
```

步骤2 在Leaf-CE6851HI-1与Leaf-CE6851HI-2上配置VXLAN业务接入点。

```
[~Leaf-CE6851HI-1&CE6851HI-2] bridge-domain 10
[~Leaf-CE6851HI-1&CE6851HI-2-bd10] vxlan vni 10010
[*Leaf-CE6851HI-1&CE6851HI-2-bd10] quit
[*Leaf-CE6851HI-1&CE6851HI-2] commit

[~Leaf-CE6851HI-1&CE6851HI-2] interface eth-trunk 1.1 mode l2
[*Leaf-CE6851HI-1&CE6851HI-2-Eth-Trunk1.1] encapsulation dot1q vid 30
[*Leaf-CE6851HI-1&CE6851HI-2-Eth-Trunk1.1] bridge-domain 10
[*Leaf-CE6851HI-1&CE6851HI-2-Eth-Trunk1.1] quit
[*Leaf-CE6851HI-1&CE6851HI-2-Eth-Trunk1.1] commit
```

----结束

配置路由

执行以下步骤配置Underlay的路由(本例以BGP为例,也可以使用OSPF)。

步骤1 在堆叠组上配置BGP路由,对接Spine。

```
[~Leaf-CE6851HI-1&CE6851HI-2] bgp 65021
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] router-id 11.11.11.11
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] timer keepalive 10 hold 30
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] group Spine-CE12804 external
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] peer Spine-CE12804 as-number 65010
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] peer 192.168.40.158 as-number 65010
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] peer 192.168.40.158 group Spine-CE12804
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] peer 192.168.40.170 as-number 65010
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] peer 192.168.40.170 group Spine-CE12804
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] peer 192.168.40.166 as-number 65010
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] peer 192.168.40.166 group Spine-CE12804
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] peer 192.168.40.162 as-number 65010
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] peer 192.168.40.162 group Spine-CE12804
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] ipv4-family unicast
[*Leaf-CE6851HI-1&CE6851HI-2-bgp-af-ipv4] preference 20 200 10
[*Leaf-CE6851HI-1&CE6851HI-2-bgp-af-ipv4] network 11.11.11.11 255.255.255.255
[*Leaf-CE6851HI-1&CE6851HI-2-bgp-af-ipv4] maximum load-balancing 32
[*Leaf-CE6851HI-1&CE6851HI-2-bgp-af-ipv4] quit
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] quit
[*Leaf-CE6851HI-1&CE6851HI-2] commit
```

----结束

配置 BGP EVPN

使用BGP EVPN作为VXLAN的控制平面时,需执行以下步骤将Leaf作为BGP路由反射器客户端,Spine作为BGP路由器反射器,在Leaf与Spine之间建立IBGP EVPN对等体关系,配置EVPN实例和VPN实例、头端复制功能和三层网关。

步骤1 在堆叠组上使能EVPN作为VXLAN的控制平面功能。

```
[~Leaf-CE6851HI-1&CE6851HI-2] evpn-overlay enable
[*Leaf-CE6851HI-1&CE6851HI-2] commit
```

步骤2 配置堆叠组与Spine节点分别建立IBGP EVPN对等体关系,Spine作为路由反射器RR。

```
[~Leaf-CE6851HI-1&CE6851HI-2] bgp 100 instance evpn1
*Leaf-CE6851HI-1&CE6851HI-2-bgp-instance-evpn1] router-id 11.11.11.11
[*Leaf-CE6851HI-1&CE6851HI-2-bgp-instance-evpn1] peer 11.11.11.14 as-number 100
//与网关Spine-CE12804-1建立BGP EVPN对等体关系
[*Leaf-CE6851HI-1&CE6851HI-2-bgp-instance-evpn1] peer 11.11.11.14 connect-interface loopback 0
[*Leaf-CE6851HI-1&CE6851HI-2-bgp-instance-evpn1] peer 11.11.11.15 as-number 100
//与网关Spine-CE12804-2建立BGP EVPN对等体关系
[*Leaf-CE6851HI-1&CE6851HI-2-bgp-instance-evpn1] peer 11.11.11.15 connect-interface loopback 0
[*Leaf-CE6851HI-1&CE6851HI-2-bgp-instance-evpn1] l2vpn-family evpn
[*Leaf-CE6851HI-1&CE6851HI-2-bgp-instance-evpn1-af-evpn] peer 11.11.11.14 enable
[*Leaf-CE6851HI-1&CE6851HI-2-bgp-instance-evpn1-af-evpn] peer 11.11.11.14 advertise irb
[*Leaf-CE6851HI-1&CE6851HI-2-bgp-instance-evpn1-af-evpn] peer 11.11.11.15 enable
[*Leaf-CE6851HI-1&CE6851HI-2-bgp-instance-evpn1-af-evpn] peer 11.11.11.15 advertise irb
[*Leaf-CE6851HI-1&CE6851HI-2-bgp-instance-evpn1-af-evpn] quit
[*Leaf-CE6851HI-1&CE6851HI-2-bgp-instance-evpn1] quit
[*Leaf-CE6851HI-1&CE6851HI-2] commit
```

步骤3 在堆叠组上配置EVPN实例和VPN实例。

在堆叠组上配置EVPN实例。

```
# 11年宣出工品直LVFIN表例。
[~Leaf-CE6851HI-1&CE6851HI-2-bd10] evpn
[~Leaf-CE6851HI-1&CE6851HI-2-bd10-evpn] route-distinguisher 11:1
[*Leaf-CE6851HI-1&CE6851HI-2-bd10-evpn] vpn-target 10:1
[*Leaf-CE6851HI-1&CE6851HI-2-bd10-evpn] vpn-target 11:1 export-extcommunity //配置和VPN实例交叉的RT值
[*Leaf-CE6851HI-1&CE6851HI-2-bd10-evpn] quit
[*Leaf-CE6851HI-1&CE6851HI-2-bd10] quit
[*Leaf-CE6851HI-1&CE6851HI-2-bd10] quit
[*Leaf-CE6851HI-1&CE6851HI-2-bd10] quit
```

#在堆叠组上配置VPN实例。

```
[~Leaf-CE6851HI-1&CE6851HI-2] ip vpn-instance vpn2
[*Leaf-CE6851HI-1&CE6851HI-2-vpn-instance-vpn2] vxlan vni 10
[*Leaf-CE6851HI-1&CE6851HI-2-vpn-instance-vpn2] ipv4-family
[*Leaf-CE6851HI-1&CE6851HI-2-vpn-instance-vpn2-af-ipv4] route-distinguisher 12:1
[*Leaf-CE6851HI-1&CE6851HI-2-vpn-instance-vpn2-af-ipv4] vpn-target 1:1
[*Leaf-CE6851HI-1&CE6851HI-2-vpn-instance-vpn2-af-ipv4] vpn-target 11:1 evpn //需要携带EVPN参数,用于和EVPN实例交叉
[*Leaf-CE6851HI-1&CE6851HI-2-vpn-instance-vpn2-af-ipv4] quit
[*Leaf-CE6851HI-1&CE6851HI-2-vpn-instance-vpn2] quit
[*Leaf-CE6851HI-1&CE6851HI-2-vpn-instance-vpn2] quit
```

步骤4 在堆叠组上配置头端复制功能。

```
[~Leaf-CE6851HI-1&CE6851HI-2] interface nve 1
[*Leaf-CE6851HI-1&CE6851HI-2-Nve1] source 11.11.11.11
[*Leaf-CE6851HI-1&CE6851HI-2-Nve1] vni 10010 head-end peer-list protocol bgp
[*Leaf-CE6851HI-1&CE6851HI-2-Nve1] quit
[*Leaf-CE6851HI-1&CE6851HI-2] commit
```

步骤5 在堆叠组上配三层网关。

```
[~Leaf-CE6851HI-1&CE6851HI-2] interface vbdif 10

[*Leaf-CE6851HI-1&CE6851HI-2-Vbdif10] ip binding vpn-instance vpn2

[*Leaf-CE6851HI-1&CE6851HI-2-Vbdif10] ip address 192.168.10.5 255.255.255.0

[*Leaf-CE6851HI-1&CE6851HI-2-Vbdif10] vxlan anycast-gateway enable

[*Leaf-CE6851HI-1&CE6851HI-2-Vbdif10] arp collect host enable

[*Leaf-CE6851HI-1&CE6851HI-2-Vbdif10] quit

[*Leaf-CE6851HI-1&CE6851HI-2] commit
```

----结束

3.2.1.2 配置 Leaf M-LAG 工作组

两台Leaf交换机Leaf-CE6851HI-3和Leaf-CE6851HI-4组成M-LAG,上联Spine-1和 Spine-2,下联服务器。

配置思路

- 1. 配置IP地址:配置Leaf交换机与Spine的三层互联接口地址、Loopback0地址(作为VTEP IP)、Loopback1地址(作为Router-ID)。
- 2. 配置M-LAG:在交换机上配置M-LAG全局模式、DFS组、Peer-Link,并配置服务器接入M-LAG,配置Monitor Link关联上行接口和下行接口。
- 3. 配置路由:在M-LAG组上配置BGP动态路由,邻居为两个Spine设备,使M-LAG组 上的网段与Spine上的网段三层可达。
- 4. 配置BGP EVPN: 配置BGP EVPN作为VXLAN的控制平面、并配置BGP EVPN对等体、EVPN实例和VPN实例、头端复制功能和三层网关。

配置 IP 地址

步骤1 配置Leaf-CE6851HI-3与Leaf-CE6851HI-4直连接口地址。

山 说明

对于CE6855HI和CE7855EI,在切换三层接口之前,要先配置命令vlan reserved for maininterface startvlanid to endvlanid,配置三层主接口专用的保留VLAN。

```
[~HUAWEI] sysname Leaf-CE6851HI-3
[*HUAWEI] commit
[~Leaf-CE6851HI-3] interface 40ge 1/0/3
[~Leaf-CE6851HI-3-40GE1/0/3] description "to_Spine-CE12804-1-40GE1/0/2"
[*Leaf-CE6851HI-3-40GE1/0/3] undo portswitch
[*Leaf-CE6851HI-3-40GE1/0/3] ip address 192.168.41.157 30
[*Leaf-CE6851HI-3-40GE1/0/3] commit
[~Leaf-CE6851HI-3-40GE1/0/3] quit
[~Leaf-CE6851HI-3] interface 40ge 1/0/4
[~Leaf-CE6851HI-3-40GE1/0/4] description "to_Spine-CE12804-2-40GE1/0/3"
[*Leaf-CE6851HI-3-40GE1/0/4] undo portswitch
[*Leaf-CE6851HI-3-40GE1/0/4] ip address 192.168.41.165 30
[*Leaf-CE6851HI-3-40GE1/0/4] commit
[~Leaf-CE6851HI-3-40GE1/0/4] quit
[~HUAWEI] sysname Leaf-CE6851HI-4
[*HUAWEI] commit
[~Leaf-CE6851HI-4] interface 40ge 1/0/3
[~Leaf-CE6851HI-4-40GE1/0/3] description "to_Spine-CE12804-2-40GE1/0/2"
[*Leaf-CE6851HI-4-40GE1/0/3] undo portswitch
[*Leaf-CE6851HI-4-40GE1/0/3] ip address 192.168.41.169 30
[*Leaf-CE6851HI-4-40GE1/0/3] commit
[~Leaf-CE6851HI-4-40GE1/0/3] quit
[~Leaf-CE6851HI-4] interface 40ge 1/0/4
[~Leaf-CE6851HI-4-40GE1/0/4] description "to_Spine-CE12804-1-40GE1/0/3"
[*Leaf-CE6851HI-4-40GE1/0/4] undo portswitch
[*Leaf-CE6851HI-4-40GE1/0/4] ip address 192.168.41.161 30
[*Leaf-CE6851HI-4-40GE1/0/4] commit
[~Leaf-CE6851HI-4-40GE1/0/4] quit
```

步骤2 配Loopback口地址,规划用Loopback0地址作为VTEP,用Loopback1地址作为Router-ID。

```
[~Leaf-CE6851HI-3] interface loopback 0 //作为VTEP地址
[*Leaf-CE6851HI-3-LoopBack0] ip address 11.11.11.12 32
[*Leaf-CE6851HI-3-LoopBack0] commit
[~Leaf-CE6851HI-3-LoopBack0] quit
[~Leaf-CE6851HI-3] interface loopback 1
```

VXLAN 最佳实践 3 VXLAN 部署建议

```
[*Leaf-CE6851HI-3-LoopBack1] ip address 13.13.13.13 32
[*Leaf-CE6851HI-3-LoopBack1] commit
[~Leaf-CE6851HI-3-LoopBack1] quit

[~Leaf-CE6851HI-4] interface loopback 0 //作为VTEP地址
[*Leaf-CE6851HI-4-LoopBack0] ip address 11.11.11.12 32
[*Leaf-CE6851HI-4-LoopBack0] commit
[~Leaf-CE6851HI-4-LoopBack0] quit
[~Leaf-CE6851HI-4-LoopBack0] ip address 14.14.14.14 32
[*Leaf-CE6851HI-4-LoopBack1] ip address 14.14.14.14 32
[*Leaf-CE6851HI-4-LoopBack1] commit
[~Leaf-CE6851HI-4-LoopBack1] quit
```

----结束

配置路由

执行以下步骤配置Underlay的路由(本例以BGP为例,也可以使用OSPF)。

步骤1 在Leaf-CE6851HI-3上配置BGP路由,对接Spine。

```
[~Leaf-CE6851HI-3] bgp 65022
[*Leaf-CE6851HI-3-bgp] router-id 13.13.13.13
[*Leaf-CE6851HI-3-bgp] timer keepalive 10 hold 30
[*Leaf-CE6851HI-3-bgp] group Spine-CE12804 external
[*Leaf-CE6851HI-3-bgp] peer Spine-CE12804 as-number 65010
[*Leaf-CE6851HI-3-bgp] peer 192.168.41.158 as-number 65010
[*Leaf-CE6851HI-3-bgp] peer 192.168.41.158 group Spine-CE12804
[*Leaf-CE6851HI-3-bgp] peer 192.168.41.166 as-number 65010
[*Leaf-CE6851HI-3-bgp] peer 192.168.41.166 group Spine-CE12804
[*Leaf-CE6851HI-3-bgp] ipv4-family unicast
[*Leaf-CE6851HI-3-bgp-af-ipv4] preference 20 200 10
[*Leaf-CE6851HI-3-bgp-af-ipv4] network 11.11.11.12 255.255.255.255
[*Leaf-CE6851HI-3-bgp-af-ipv4] network 13.13.13.13 255.255.255.255
[*Leaf-CE6851HI-3-bgp-af-ipv4] maximum load-balancing 32
[*Leaf-CE6851HI-3-bgp-af-ipv4] quit
[*Leaf-CE6851HI-3-bgp] quit
[*Leaf-CE6851HI-3] commit
```

步骤2 在Leaf-CE6851HI-4上配置BGP路由,对接Spine。

```
[~Leaf-CE6851HI-4] bgp 65022
[*Leaf-CE6851HI-4-bgp] router-id 14.14.14.14
[*Leaf-CE6851HI-4-bgp] timer keepalive 10 hold 30
[*Leaf-CE6851HI-4-bgp] group Spine-CE12804 external
[*Leaf-CE6851HI-4-bgp] peer Spine-CE12804 as-number 65010
[*Leaf-CE6851HI-4-bgp] peer 192.168.41.170 as-number 65010
[*Leaf-CE6851HI-4-bgp] peer 192.168.41.170 group Spine-CE12804
[*Leaf-CE6851HI-4-bgp] peer 192.168.41.162 as-number 65010
[*Leaf-CE6851HI-4-bgp] peer 192.168.41.162 group Spine-CE12804
[*Leaf-CE6851HI-4-bgp] ipv4-family unicast
[*Leaf-CE6851HI-4-bgp-af-ipv4] preference 20 200 10
[*Leaf-CE6851HI-4-bgp-af-ipv4] network 11.11.11.12 255.255.255.255
[*Leaf-CE6851HI-4-bgp-af-ipv4] network 14.14.14.14 255.255.255.255
[*Leaf-CE6851HI-4-bgp-af-ipv4] maximum load-balancing 32
[*Leaf-CE6851HI-4-bgp-af-ipv4] quit
[*Leaf-CE6851HI-4-bgp] quit
[*Leaf-CE6851HI-4] commit
```

----结束

配置 M-LAG

步骤1 配置M-LAG模式。

<Leaf-CE6851HI-3> system-view [~Leaf-CE6851HI-3] stp mode rstp

```
[*Leaf-CE6851HI-3] stp v-stp enable
[*Leaf-CE6851HI-3] commit
[~Leaf-CE6851HI-3] lacp m-lag system-id 00e0-fc00-0001
[*Leaf-CE6851HI-3] commit

<Leaf-CE6851HI-4> system-view
[~Leaf-CE6851HI-4] stp mode rstp
[*Leaf-CE6851HI-4] stp v-stp enable
[*Leaf-CE6851HI-4] commit
[~Leaf-CE6851HI-4] commit
[~Leaf-CE6851HI-4] commit
```

步骤2 在Leaf-CE6851HI-3与Leaf-CE6851HI-4上配置M-LAG的DFS Group。

```
[~Leaf-CE6851HI-3] dfs-group 1
[*Leaf-CE6851HI-3-dfs-group-1] source ip 13.13.13.13
[*Leaf-CE6851HI-3-dfs-group-1] priority 150
[*Leaf-CE6851HI-3-dfs-group-1] quit
[*Leaf-CE6851HI-3] commit

[~Leaf-CE6851HI-4] dfs-group 1
[*Leaf-CE6851HI-4-dfs-group-1] source ip 14.14.14.14
[*Leaf-CE6851HI-4-dfs-group-1] priority 120
[*Leaf-CE6851HI-4-dfs-group-1] quit
[*Leaf-CE6851HI-4] commit
```

步骤3 在Leaf-CE6851HI-3与Leaf-CE6851HI-4上配置M-LAG的Peer-Link。

```
[~Leaf-CE6851HI-3] interface eth-trunk 0
[*Leaf-CE6851HI-3-Eth-Trunk0] trunkport 40ge 1/0/1
[*Leaf-CE6851HI-3-Eth-Trunk0] trunkport 40ge 1/0/2
[*Leaf-CE6851HI-3-Eth-Trunk0] mode lacp-static
[*Leaf-CE6851HI-3-Eth-Trunk0] peer-link 1
[*Leaf-CE6851HI-3-Eth-Trunk0] quit
[*Leaf-CE6851HI-3] commit

[~Leaf-CE6851HI-4] interface eth-trunk 0
[*Leaf-CE6851HI-4-Eth-Trunk0] trunkport 40ge 1/0/1
[*Leaf-CE6851HI-4-Eth-Trunk0] trunkport 40ge 1/0/2
[*Leaf-CE6851HI-4-Eth-Trunk0] mode lacp-static
[*Leaf-CE6851HI-4-Eth-Trunk0] peer-link 1
[*Leaf-CE6851HI-4-Eth-Trunk0] quit
[*Leaf-CE6851HI-4] commit
```

步骤4 配置服务器接入M-LAG。

在Leaf-CE6851HI-3与Leaf-CE6851HI-4上配置M-LAG的成员口(服务器以负载均衡模式接入M-LAG)。

```
[~Leaf-CE6851HI-3] interface eth-trunk 1
[*Leaf-CE6851HI-3-Eth-Trunk1] mode lacp-static
[*Leaf-CE6851HI-3-Eth-Trunk1] port link-type trunk
[*Leaf-CE6851HI-3-Eth-Trunk1] undo port trunk allow-pass vlan 1
[*Leaf-CE6851HI-3-Eth-Trunk1] trunkport 10ge 1/0/1
[*Leaf-CE6851HI-3-Eth-Trunk1] dfs-group 1 m-lag 1
[*Leaf-CE6851HI-3-Eth-Trunk1] quit
[*Leaf-CE6851HI-3] commit
[*Leaf-CE6851HI-4] interface eth-trunk 1
[*Leaf-CE6851HI-4-Eth-Trunk1] mode lacp-static
[*Leaf-CE6851HI-4-Eth-Trunk1] port link-type trunk
[*Leaf-CE6851HI-4-Eth-Trunk1] trunkport 10ge 1/0/1
[*Leaf-CE6851HI-4-Eth-Trunk1] undo port trunk allow-pass vlan 1
[*Leaf-CE6851HI-4-Eth-Trunk1] dfs-group 1 m-lag 1
[*Leaf-CE6851HI-4-Eth-Trunk1] quit
[*Leaf-CE6851HI-4] commit
```

步骤5 在Leaf-CE6851HI-3与Leaf-CE6851HI-4上配置VXLAN业务接入点。

```
[~Leaf-CE6851HI-3] bridge-domain 10
[*Leaf-CE6851HI-3-bd10] vxlan vni 10010
```

```
[*Leaf-CE6851HI-3-bd10] quit
[*Leaf-CE6851HI-3] commit
[~Leaf-CE6851HI-3] interface eth-trunk 1.1 mode l2
[*Leaf-CE6851HI-3-Eth-Trunk1.1] encapsulation dot1q vid 30
[*Leaf-CE6851HI-3-Eth-Trunk1.1] bridge-domain 10
[*Leaf-CE6851HI-3-Eth-Trunk1.1] quit
[*Leaf-CE6851HI-3] commit
[~Leaf-CE6851HI-4] bridge-domain 10
[*Leaf-CE6851HI-4-bd10] vxlan vni 10010
[*Leaf-CE6851HI-4-bd10] quit
[*Leaf-CE6851HI-4] commit
[~Leaf-CE6851HI-4] interface eth-trunk 1.1 mode l2
[*Leaf-CE6851HI-4-Eth-Trunk1.1] encapsulation dot1q vid 30
[*Leaf-CE6851HI-4-Eth-Trunk1.1] bridge-domain 10
[*Leaf-CE6851HI-4-Eth-Trunk1.1] quit
[*Leaf-CE6851HI-4] commit
```

步骤6 分别在Leaf-CE6851HI-3与Leaf-CE6851HI-4上配置Monitor Link关联上行接口和下行接口,避免单台设备的所有上行链路都发生故障,本台设备用户侧流量无法转发。

```
[~Leaf-CE6851HI-3] monitor-link group 1
[*Leaf-CE6851HI-3-mtlk-group1] port 40ge 1/0/3 uplink
[*Leaf-CE6851HI-3-mtlk-group1] port 40ge 1/0/4 uplink
[*Leaf-CE6851HI-3-mtlk-group1] port 10ge 1/0/1 downlink 1
[*Leaf-CE6851HI-3-mtlk-group1] commit

[~Leaf-CE6851HI-4] monitor-link group 1
[*Leaf-CE6851HI-4-mtlk-group1] port 40ge 1/0/3 uplink
[*Leaf-CE6851HI-4-mtlk-group1] port 40ge 1/0/4 uplink
[*Leaf-CE6851HI-4-mtlk-group1] port 10ge 1/0/1 downlink 1
[*Leaf-CE6851HI-4-mtlk-group1] commit
```

----结束

配置 BGP-EVPN

使用BGP EVPN作为VXLAN的控制平面时,需执行以下步骤将Leaf作为BGP路由反射器客户端,Spine作为BGP路由器反射器,在Leaf与Spine之间建立IBGP EVPN对等体关系,配置EVPN实例和VPN实例、头端复制功能和三层网关。

步骤1 在Leaf-CE6851HI-3和Leaf-CE6851HI-4上使能EVPN作为VXLAN的控制平面功能。

```
[~Leaf-CE6851HI-3] evpn-overlay enable

[*Leaf-CE6851HI-3] commit

[~Leaf-CE6851HI-4] evpn-overlay enable

[*Leaf-CE6851HI-4] commit
```

步骤2 配置Leaf-CE6851HI-3和Leaf-CE6851HI-4上分别与Spine节点建立BGP EVPN对等体关系,Spine作为路由反射器RR。

#在Leaf-CE6851HI-3上配置与Spine建立IBGP EVPN对等体关系。

```
[~Leaf-CE6851HI-3] bgp 100 instance evpn1
[*Leaf-CE6851HI-3-bgp-instance-evpn1] router-id 13.13.13.13
[*Leaf-CE6851HI-3-bgp-instance-evpn1] peer 11.11.11.14 as-number 100
//与网关Spine-CE12804-1建立BGP EVPN对等体关系
[*Leaf-CE6851HI-3-bgp-instance-evpn1] peer 11.11.11.14 connect-interface loopback 1
[*Leaf-CE6851HI-3-bgp-instance-evpn1] peer 11.11.11.15 as-number 100
//与网关Spine-CE12804-2建立BGP EVPN对等体关系
[*Leaf-CE6851HI-3-bgp-instance-evpn1] peer 11.11.11.15 connect-interface loopback 1
[*Leaf-CE6851HI-3-bgp-instance-evpn1] l2vpn-family evpn
[*Leaf-CE6851HI-3-bgp-instance-evpn1-af-evpn] peer 11.11.11.14 enable
[*Leaf-CE6851HI-3-bgp-instance-evpn1-af-evpn] peer 11.11.11.15 enable
```

```
[*Leaf-CE6851HI-3-bgp-instance-evpn1-af-evpn] peer 11.11.11.15 advertise irb
[*Leaf-CE6851HI-3-bgp-instance-evpn1-af-evpn] quit
[*Leaf-CE6851HI-3-bgp-instance-evpn1] quit
[*Leaf-CE6851HI-3] commit
```

在Leaf-CE6851HI-4上配置与Spine建立IBGP EVPN对等体关系。

```
[~Leaf-CE6851HI-4] bgp 100 instance evpn1
[*Leaf-CE6851HI-4-bgp-instance-evpn1] router-id 14.14.14.14
[*Leaf-CE6851HI-4-bgp-instance-evpn1] peer 11.11.11.14 as-number 100
//与网关Spine-CE12804-1建立BGP EVPN对等体关系
[*Leaf-CE6851HI-4-bgp-instance-evpn1] peer 11.11.11.14 connect-interface loopback 1
[*Leaf-CE6851HI-4-bgp-instance-evpn1] peer 11.11.11.15 as-number 100
//与网关Spine-CE12804-2建立BGP EVPN对等体关系
[*Leaf-CE6851HI-4-bgp-instance-evpn1] peer 11.11.11.15 connect-interface loopback 1
[*Leaf-CE6851HI-4-bgp-instance-evpn1] l2vpn-family evpn
[*Leaf-CE6851HI-4-bgp-instance-evpn1-af-evpn] peer 11.11.11.14 enable
[*Leaf-CE6851HI-4-bgp-instance-evpn1-af-evpn] peer 11.11.11.14 advertise irb
[*Leaf-CE6851HI-4-bgp-instance-evpn1-af-evpn] peer 11.11.11.15 enable
[*Leaf-CE6851HI-4-bgp-instance-evpn1-af-evpn] peer 11.11.11.15 advertise irb
[*Leaf-CE6851HI-4-bgp-instance-evpn1-af-evpn] quit
[*Leaf-CE6851HI-4-bgp-instance-evpn1] quit
[*Leaf-CE6851HI-4] commit
```

步骤3 # 在Leaf-CE6851HI-3上配置EVPN实例。

```
[~Leaf-CE6851HI-3] bridge-domain 10
[~Leaf-CE6851HI-3-bd10] evpn
[*Leaf-CE6851HI-3-bd10-evpn] route-distinguisher 13:1
[*Leaf-CE6851HI-3-bd10-evpn] vpn-target 1:1
[*Leaf-CE6851HI-3-bd10-evpn] vpn-target 11:1 export-extcommunity //配置和VPN实例交叉的RT值
[*Leaf-CE6851HI-3-bd10-evpn] quit
[*Leaf-CE6851HI-3-bd10] quit
[*Leaf-CE6851HI-3-bd10] commit
```

在Leaf-CE6851HI-3上配置VPN实例。

```
[~Leaf-CE6851HI-3] ip vpn-instance vpn2
[*Leaf-CE6851HI-3-vpn-instance-vpn2] vxlan vni 10
[*Leaf-CE6851HI-3-vpn-instance-vpn2] ipv4-family
[*Leaf-CE6851HI-3-vpn-instance-vpn2-af-ipv4] route-distinguisher 12:2
[*Leaf-CE6851HI-3-vpn-instance-vpn2-af-ipv4] vpn-target 11:1 evpn //需要携带EVPN参数,用于和EVPN 实例交叉
[*Leaf-CE6851HI-3-vpn-instance-vpn2-af-ipv4] quit
[*Leaf-CE6851HI-3-vpn-instance-vpn2] quit
[*Leaf-CE6851HI-3] commit
```

在Leaf-CE6851HI-4上配置EVPN实例。

```
[~Leaf-CE6851HI-4] bridge-domain 10
[~Leaf-CE6851HI-4-bd10] evpn
[*Leaf-CE6851HI-4-bd10-evpn] route-distinguisher 13:1
[*Leaf-CE6851HI-4-bd10-evpn] vpn-target 10:1
[*Leaf-CE6851HI-4-bd10-evpn] vpn-target 11:1 export-extcommunity //配置和VPN实例交叉的RT值
[*Leaf-CE6851HI-4-bd10-evpn] quit
[*Leaf-CE6851HI-4-bd10] quit
[*Leaf-CE6851HI-4] commit
```

在Leaf-CE6851HI-4上配置VPN实例。

```
[~Leaf-CE6851HI-4] ip vpn-instance vpn2
[*Leaf-CE6851HI-4-vpn-instance-vpn2] vxlan vni 10
[*Leaf-CE6851HI-4-vpn-instance-vpn2] ipv4-family
[*Leaf-CE6851HI-4-vpn-instance-vpn2-af-ipv4] route-distinguisher 12:3
[*Leaf-CE6851HI-4-vpn-instance-vpn2-af-ipv4] vpn-target 1:1
[*Leaf-CE6851HI-4-vpn-instance-vpn2-af-ipv4] vpn-target 11:1 evpn //需要携带EVPN参数,用于和EVPN实例交叉
[*Leaf-CE6851HI-4-vpn-instance-vpn2-af-ipv4] quit
[*Leaf-CE6851HI-4-vpn-instance-vpn2] quit
[*Leaf-CE6851HI-4-vpn-instance-vpn2] quit
```

步骤4 配置头端复制功能。

在Leaf-CE6851HI-3配置头端复制功能。

[~Leaf-CE6851HI-3] interface nve 1

[*Leaf-CE6851HI-3-Nve1] source 11.11.11.12

[*Leaf-CE6851HI-3-Nve1] source mac-address 0000-5e00-0102

[*Leaf-CE6851HI-3-Nve1] vni 10010 head-end peer-list protocol bgp

[*Leaf-CE6851HI-3-Nve1] quit

[*Leaf-CE6851HI-3] commit

在Leaf-CE6851HI-4配置头端复制功能。

[~Leaf-CE6851HI-4] interface nve 1

[*Leaf-CE6851HI-4-Nve1] source 11.11.11.12

[*Leaf-CE6851HI-4-Nve1] source mac-address 0000-5e00-0102

[*Leaf-CE6851HI-4-Nve1] vni 10010 head-end peer-list protocol bgp

[*Leaf-CE6851HI-4-Nve1] quit

[*Leaf-CE6851HI-4] commit

步骤5 在Leaf-CE6851HI-3上配三层网关。

[~Leaf-CE6851HI-3] interface vbdif 10

[*Leaf-CE6851HI-3-Vbdif10] ip binding vpn-instance vpn2

[*Leaf-CE6851HI-3-Vbdif10] ip address 192.168.10.5 255.255.255.0

[*Leaf-CE6851HI-3-Vbdif10] source mac-address 0000-5e00-0115

[*Leaf-CE6851HI-3-Vbdif10] vxlan anycast-gateway enable

[*Leaf-CE6851HI-3-Vbdif10] arp collect host enable

[*Leaf-CE6851HI-3-Vbdif10] quit

[*Leaf-CE6851HI-3] commit

在Leaf-CE6851HI-4上配三层网关。

[~Leaf-CE6851HI-4] interface vbdif 10

[*Leaf-CE6851HI-4-Vbdif10] ip binding vpn-instance vpn2

[*Leaf-CE6851HI-4-Vbdif10] ip address 192.168.10.5 255.255.255.0

[*Leaf-CE6851HI-4-Vbdif10] source mac-address 0000-5e00-0115

[*Leaf-CE6851HI-4-Vbdif10] vxlan anycast-gateway enable

[*Leaf-CE6851HI-4-Vbdif10] arp collect host enable

[*Leaf-CE6851HI-4-Vbdif10] quit

[*Leaf-CE6851HI-4] commit

----结束

3.2.1.3 配置 Leaf SVF 工作组

两台Leaf交换机Leaf-CE6851HI-5和Leaf-CE6851HI-6组成堆叠组,下连多台低成本CE交换机,以扩展接口接入数量。

配置思路

- 1. 两台CE6851HI作为父节点来组建堆叠组:配置堆叠和双主检测,并重启设备、连接线缆,使堆叠生效。
- 2. 配置SVF:配置父节点的Fabric-port,用于连接叶子节点交换机CE5810,并连接 父节点和叶子节点之间的SVF链路。
- 3. 配置IP地址:配置Leaf与Spine的三层互联接口地址、Loopback0地址(作为Router-ID和VTEP IP)。
- 4. 配置服务器接入:介绍业务服务器接入堆叠组时,堆叠组交换机上的配置。
- 5. 配置路由:在堆叠组上配置BGP动态路由,邻居为两个Spine设备,使SVF组上的网段与Spine上的网段三层可达。
- 6. 配置BGP EVPN: 配置BGP EVPN作为VXLAN的控制平面、并配置BGP EVPN对等体、EVPN实例和VPN实例、头端复制功能和三层网关。

组建堆叠组

步骤1 配置Leaf-CE6851HI-5的堆叠成员ID为1,优先级为150,Domain ID为10。

<HUAWEI> system-view

[~HUAWEI] sysname Leaf-CE6851HI-5

[*HUAWEI] commit

[~Leaf-CE6851HI-5] stack

[~Leaf-CE6851HI-5-stack] stack member 1 priority 150

[*Leaf-CE6851HI-5-stack] stack member 1 domain 10

[*Leaf-CE6851HI-5-stack] quit

[*Leaf-CE6851HI-5] commit

步骤2 配置Leaf-CE6851HI-6的Domain ID为10。

<HUAWEI> system-view

[~HUAWEI] sysname Leaf-CE6851HI-6

[*HUAWEI] commit

[~Leaf-CE6851HI-6] stack

[~Leaf-CE6851HI-6-stack] stack member 1 renumber 2 inherit-config

[*Leaf-CE6851HI-6-stack] stack member 1 domain 10

[*Leaf-CE6851HI-6-stack] quit

[*Leaf-CE6851HI-6] commit

□ 说明

组建SVF时要求父交换机的堆叠成员ID必须小于或等于4。

步骤3 配置堆叠端口。

将Leaf-CE6851HI-5的业务口40GE1/0/1~40GE1/0/2加入堆叠端口1/1。

[~Leaf-CE6851HI-5] interface stack-port 1/1

[*Leaf-CE6851HI-5-Stack-Port1/1] port member-group interface 40ge 1/0/1 to 1/0/2

Warning: After the configuration is complete,1.The interface(s) (40GE1/0/1-1/0/2) will be converted to stack mode and be configured with the port crc-statistics trigger error-down command if the configuration does not exist.2.The interface(s) may go Error-Down (crc-statistics) because there is no shutdown configuration on the interfaces.Continue? [Y/N]: y

[*Leaf-CE6851HI-5-Stack-Port1/1] quit

[*Leaf-CE6851HI-5] commit

[~Leaf-CE6851HI-5] quit

将Leaf-CE6851HI-6的业务口40GE1/0/1~40GE1/0/2加入堆叠端口1/1。

[~Leaf-CE6851HI-6] interface stack-port 1/1

*Leaf-CE6851HI-6-Stack-Port1/1] port member-group interface 40ge 1/0/1 to 1/0/2

Warning: After the configuration is complete,1.The interface(s) (40GE1/0/1-1/0/2) will be converted to stack mode and be configured with the port crc-statistics trigger error-down command if the configuration does not exist.2.The interface(s) may go Error-Down (crc-statistics) because there is no shutdown configuration on the interfaces.Continue? [Y/N]: y

[*Leaf-CE6851HI-6-Stack-Port1/1] quit

[*Leaf-CE6851HI-6] commit

[~Leaf-CE6851HI-6] quit

步骤4 保存配置后重启设备。

保存Leaf-CE6851HI-5的配置,然后重启设备。Leaf-CE6851HI-6的操作与之类似,不再赘述。

<Leaf-CE6851HI-5> save

Warning: The current configuration will be written to the device. Continue? [Y/N]: y

<Leaf-CE6851HI-5> reboot

Warning: The system will reboot. Continue? [Y/N]:y

步骤5 连接堆叠线缆,建立堆叠。

步骤6 堆叠组建后执行命令save保存配置。

步骤7 堆叠建立后,配置堆叠组的双主检测,防止堆叠组分裂后网络中存在两个配置冲突的网络设备。

可采用业务口直连的方式配置双主检测,建议规划的直连链路至少2条,保证可靠性,配置示例如下。

```
[~Leaf-CE6851HI-5] sysname Leaf-CE6851HI-5&CE6851HI-6
[*Leaf-CE6851HI-5] commit
[~Leaf-CE6851HI-5&CE6851HI-6] interface 10ge 1/0/30
[~Leaf-CE6851HI-5&CE6851HI-6-10GE1/0/30] description "for DAD"
[*Leaf-CE6851HI-5&CE6851HI-6-10GE1/0/30] dual-active detect mode direct
Warning: The interface will block common data packets, except BPDU packets. Continue? [Y/N]: y
[*Leaf-CE6851HI-5&CE6851HI-6-10GE1/0/30] quit
[*Leaf-CE6851HI-5&CE6851HI-6] interface 10ge 1/0/31
[*Leaf-CE6851HI-5&CE6851HI-6-10GE1/0/31] description "for DAD"
[*Leaf-CE6851HI-5&CE6851HI-6-10GE1/0/31] dual-active detect mode direct
Warning: The interface will block common data packets, except BPDU packets. Continue? [Y/N]: y
[*Leaf-CE6851HI-5&CE6851HI-6-10GE1/0/31] quit
[*Leaf-CE6851HI-5&CE6851HI-6] interface 10ge 2/0/30
[*Leaf-CE6851HI-5&CE6851HI-6-10GE2/0/30] description "for DAD"
[*Leaf-CE6851HI-5&CE6851HI-6-10GE2/0/30] dual-active detect mode direct
Warning: The interface will block common data packets, except BPDU packets. Continue? [Y/N]: y
[*Leaf-CE6851HI-5&CE6851HI-6-10GE2/0/30] quit
[*Leaf-CE6851HI-5&CE6851HI-6] interface 10ge 2/0/31
[*Leaf-CE6851HI-5&CE6851HI-6-10GE2/0/31] description "for DAD"
[*Leaf-CE6851HI-5&CE6851HI-6-10GE2/0/31] dual-active detect mode direct
Warning: The interface will block common data packets, except BPDU packets. Continue? [Y/N]: y
[*Leaf-CE6851HI-5&CE6851HI-6-10GE2/0/31] commit
[~Leaf-CE6851HI-5&CE6851HI-6-10GE2/0/31] quit
```

----结束

配置 SVF

步骤1 配置Fabric-port1,并加入成员端口10GE1/0/1和10GE2/0/1,绑定的叶子ID为101。

```
[~Leaf-CE6851HI-5&CE6851HI-6] interface fabric-port 1
[*Leaf-CE6851HI-5&CE6851HI-6-Fabric-Port1] port bind member 101
[*Leaf-CE6851HI-5&CE6851HI-6-Fabric-Port1] port member-group interface 10ge 1/0/1 2/0/1
Warning: The interface(s) (10GE1/0/1-1/0/4,10GE2/0/1-2/0/4) will be converted to stack mode and have "port crc-statistics trigger error-down" configured. [Y/N]: y
[*Leaf-CE6851HI-5&CE6851HI-6-Fabric-Port1] quit
[*Leaf-CE6851HI-5&CE6851HI-6] commit
```

步骤2 配置Fabric-port2,并加入成员端口10GE1/0/2和10GE2/0/2,绑定的叶子ID为102。

```
[~Leaf-CE6851HI-5&CE6851HI-6] interface fabric-port 2

[*Leaf-CE6851HI-5&CE6851HI-6-Fabric-Port2] port bind member 102

[*Leaf-CE6851HI-5&CE6851HI-6-Fabric-Port2] port member-group interface 10ge 1/0/2 2/0/2

[*Leaf-CE6851HI-5&CE6851HI-6-Fabric-Port2] quit

[*Leaf-CE6851HI-5&CE6851HI-6] commit
```

步骤3 配置Fabric-port3,并加入成员端口10GE1/0/3和10GE2/0/3,绑定的叶子ID为103。

```
[~Leaf-CE6851HI-5&CE6851HI-6] interface fabric-port 3
[*Leaf-CE6851HI-5&CE6851HI-6-Fabric-Port3] port bind member 103
[*Leaf-CE6851HI-5&CE6851HI-6-Fabric-Port3] port member-group interface 10ge 1/0/3 2/0/3
[*Leaf-CE6851HI-5&CE6851HI-6-Fabric-Port3] commit
[~Leaf-CE6851HI-5&CE6851HI-6-Fabric-Port3] quit
```

步骤4 配置SVF系统转发模式为集中式,使叶子交换机的流量全部上送至父交换机进行转发。

```
[~Leaf-CE6851HI-5&CE6851HI-6] stack
[*Leaf-CE6851HI-5&CE6851HI-6-stack] stack forwarding-model centralized
[*Leaf-CE6851HI-5&CE6851HI-6-stack] commit
[*Leaf-CE6851HI-5&CE6851HI-6-stack] return
<Leaf-CE6851HI-5&CE6851HI-6> save
<Leaf-CE6851HI-5&CE6851HI-6> reboot //重启父交换机使转发模式生效
```

步骤5 连接叶子交换机与父交换机,并将叶子交换机上电。叶子交换机使用上行10GE端口与 父交换机相连。

此处默认叶子交换机为空配置文件启动且为自协商模式,可通过自协商加入SVF,无需任何配置。

如果叶子交换机有启动配置文件,则需要清空叶子交换机的下次启动配置文件后重启(自协商模式下),或者执行以下命令将叶子交换机强制配置为Leaf模式。

```
[~Leaf-CE6851HI-5&CE6851HI-6] stack
[~Leaf-CE6851HI-5&CE6851HI-6-stack] switch mode leaf member all //配置交换机的工作模式后,需要重启该设备配置才能生效
[*Leaf-CE6851HI-5&CE6851HI-6-stack] commit
[*Leaf-CE6851HI-5&CE6851HI-6-stack] return
<*Leaf-CE6851HI-5&CE6851HI-6> save
```

----结束

<*Leaf-CE6851HI-5&CE6851HI-6> reboot

配置 IP 地址

步骤1 配置互连接口地址。

□ 说明

对于CE6855HI和CE7855EI,在切换三层接口之前,要先配置vlan reserved for maininterface startvlanid to endvlanid,配置三层主接口专用的保留VLAN。

```
[~Leaf-CE6851HI-5&CE6851HI-6] interface 40ge 1/0/3
[~Leaf-CE6851HI-5&CE6851HI-6-40GE1/0/3] description "to_Spine-CE12804-1-40GE1/0/6"
[*Leaf-CE6851HI-5&CE6851HI-6-40GE1/0/3] undo portswitch
[*Leaf-CE6851HI-5&CE6851HI-6-40GE1/0/3] ip address 192.168.46.157 30
[*Leaf-CE6851HI-5&CE6851HI-6-40GE1/0/3] commit
[~Leaf-CE6851HI-5&CE6851HI-6-40GE1/0/3] quit
[~Leaf-CE6851HI-5&CE6851HI-6] interface 40ge 1/0/4
[~Leaf-CE6851HI-5&CE6851HI-6-40GE1/0/4] description "to_Spine-CE12804-2-40GE1/0/6"
[*Leaf-CE6851HI-5&CE6851HI-6-40GE1/0/4] undo portswitch
[*Leaf-CE6851HI-5&CE6851HI-6-40GE1/0/4] ip address 192.168.46.165 30
[*Leaf-CE6851HI-5&CE6851HI-6-40GE1/0/4] commit
[~Leaf-CE6851HI-5&CE6851HI-6-40GE1/0/4] quit
[~Leaf-CE6851HI-5&CE6851HI-6] interface 40ge 2/0/3
[~Leaf-CE6851HI-5&CE6851HI-6-40GE2/0/3] description "to_Spine-CE12804-1-40GE1/0/7"
[*Leaf-CE6851HI-5&CE6851HI-6-40GE2/0/3] undo portswitch
[*Leaf-CE6851HI-5&CE6851HI-6-40GE2/0/3] ip address 192.168.46.169 30
[*Leaf-CE6851HI-5&CE6851HI-6-40GE2/0/3] commit
[~Leaf-CE6851HI-5&CE6851HI-6-40GE2/0/3] quit
[~Leaf-CE6851HI-5&CE6851HI-6] interface 40ge 2/0/4
[~Leaf-CE6851HI-5&CE6851HI-6-40GE2/0/4] description "to_Spine-CE12804-2-40GE1/0/7"
[*Leaf-CE6851HI-5&CE6851HI-6-40GE2/0/4] undo portswitch
[*Leaf-CE6851HI-5&CE6851HI-6-40GE2/0/4] ip address 192.168.46.161 30
[*Leaf-CE6851HI-5&CE6851HI-6-40GE2/0/4] commit
[~Leaf-CE6851HI-5&CE6851HI-6-40GE2/0/4] quit
```

步骤2 配置Loopback口地址。

```
[~Leaf-CE6851HI-5&CE6851HI-6] interface loopback 0

[*Leaf-CE6851HI-5&CE6851HI-6] description VTEP&Router-ID

[*Leaf-CE6851HI-5&CE6851HI-6-LoopBack0] ip address 11.11.11.17 32

[*Leaf-CE6851HI-5&CE6851HI-6-LoopBack0] commit

[~Leaf-CE6851HI-5&CE6851HI-6-LoopBack0] quit
```

----结束

配置服务器接入

介绍业务服务器接入SVF组时,在SVF交换机上的配置。

步骤1 配置服务器以负载均衡模式接入SVF。

[*Leaf-CE6851HI-5&CE6851HI-6] interface eth-trunk 1 [*Leaf-CE6851HI-5&CE6851HI-6-Eth-Trunk1] mode lacp-static

```
[*Leaf-CE6851HI-5&CE6851HI-6-Eth-Trunk1] port link-type trunk
[*Leaf-CE6851HI-5&CE6851HI-6-Eth-Trunk1] undo port trunk allow-pass vlan 1
[*Leaf-CE6851HI-5&CE6851HI-6-Eth-Trunk1] trunkport 10ge 101/0/1
[*Leaf-CE6851HI-5&CE6851HI-6-Eth-Trunk1] trunkport 10ge 101/0/2
[*Leaf-CE6851HI-5&CE6851HI-6-Eth-Trunk1] quit
[*Leaf-CE6851HI-5&CE6851HI-6] commit
```

步骤2 配置VXLAN业务接入点。

```
[~Leaf-CE6851HI-5&CE6851HI-6] bridge-domain 10
[*Leaf-CE6851HI-5&CE6851HI-6-bd10] vxlan vni 10010
[*Leaf-CE6851HI-5&CE6851HI-6-bd10] quit
[*Leaf-CE6851HI-5&CE6851HI-6] commit

[~Leaf-CE6851HI-5&CE6851HI-6] interface eth-trunk 1.1 mode l2
[*Leaf-CE6851HI-5&CE6851HI-6-Eth-Trunk1.1] encapsulation dot1q vid 30
[*Leaf-CE6851HI-5&CE6851HI-6-Eth-Trunk1.1] bridge-domain 10
[*Leaf-CE6851HI-5&CE6851HI-6-Eth-Trunk1.1] quit
[*Leaf-CE6851HI-5&CE6851HI-6-CE0HI-1] commit
```

----结束

配置路由

执行以下步骤配置Underlay的路由(本例以BGP为例,也可以使用OSPF)。

步骤1 在SVF组上配置BGP路由,对接Spine。

```
[~Leaf-CE6851HI-5&CE6851HI-6] bgp 65024
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] router-id 11.11.11.17
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] timer keepalive 10 hold 30
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] group Spine-CE12804 external
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] peer Spine-CE12804 as-number 65010
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] peer 192.168.40.158 as-number 65010
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] peer 192.168.40.158 group Spine-CE12804
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] peer 192.168.40.170 as-number 65010
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] peer 192.168.40.170 group Spine-CE12804
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] peer 192.168.40.166 as-number 65010
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] peer 192.168.40.166 group Spine-CE12804
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] peer 192.168.40.162 as-number 65010
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] peer 192.168.40.162 group Spine-CE12804
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] ipv4-family unicast
[*Leaf-CE6851HI-5&CE6851HI-6-bgp-af-ipv4] preference 20 200 10
[*Leaf-CE6851HI-5&CE6851HI-6-bgp-af-ipv4] network 11.11.11.17 255.255.255.255
[*Leaf-CE6851HI-5&CE6851HI-6-bgp-af-ipv4] maximum load-balancing 32
[*Leaf-CE6851HI-5&CE6851HI-6-bgp-af-ipv4] quit
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] quit
[*Leaf-CE6851HI-5&CE6851HI-6] commit
```

----结束

配置 BGP EVPN

使用BGP EVPN作为VXLAN的控制平面时,需执行以下步骤将Leaf作为BGP路由反射器客户端,Spine作为BGP路由器反射器,在Leaf与Spine之间建立IBGP EVPN对等体关系,配置EVPN实例和VPN实例、头端复制功能和三层网关。

步骤1 在SVF上使能EVPN作为VXLAN的控制平面功能。

```
[~Leaf-CE6851HI-5&CE6851HI-6] evpn-overlay enable
[*Leaf-CE6851HI-5&CE6851HI-6] commit
```

步骤2 配置堆叠组与Spine节点分别建立IBGP EVPN对等体关系,Spine作为路由反射器RR。

```
[~Leaf-CE6851HI-5&CE6851HI-6] bgp 100 instance evpn1
[*Leaf-CE6851HI-5&CE6851HI-6-bgp-instance-evpn1] router-id 11.11.11.17
[*Leaf-CE6851HI-5&CE6851HI-6-bgp-instance-evpn1] peer 11.11.11.14 as-number 100
```

```
//与网关Spine-CE12804-1建立BGP EVPN对等体关系
[*Leaf-CE6851HI-5&CE6851HI-6-bgp-instance-evpn1] peer 11.11.11.14 connect-interface LoopBack0
[*Leaf-CE6851HI-5&CE6851HI-6-bgp-instance-evpn1] peer 11.11.11.15 as-number 100
//与网关Spine-CE12804-2建立BGP EVPN对等体关系
[*Leaf-CE6851HI-5&CE6851HI-6-bgp-instance-evpn1] peer 11.11.11.15 connect-interface LoopBack0
[*Leaf-CE6851HI-5&CE6851HI-6-bgp-instance-evpn1] l2vpn-family evpn
[*Leaf-CE6851HI-5&CE6851HI-6-bgp-instance-evpn1] peer 11.11.11.14 enable
[*Leaf-CE6851HI-5&CE6851HI-6-bgp-instance-evpn1-af-evpn] peer 11.11.11.15 enable
[*Leaf-CE6851HI-5&CE6851HI-6-bgp-instance-evpn1-af-evpn] peer 11.11.11.15 enable
[*Leaf-CE6851HI-5&CE6851HI-6-bgp-instance-evpn1-af-evpn] peer 11.11.11.15 advertise irb
[*Leaf-CE6851HI-5&CE6851HI-6-bgp-instance-evpn1-af-evpn] quit
[*Leaf-CE6851HI-5&CE6851HI-6-bgp-instance-evpn1] quit
[*Leaf-CE6851HI-5&CE6851HI-6-bgp-instance-evpn1] quit
```

步骤3 在SVF上配置EVPN实例。

#在SVF上配置EVPN实例。

```
[~Leaf-CE6851HI-5&CE6851HI-6] bridge-domain 10
[~Leaf-CE6851HI-5&CE6851HI-6-bd10] evpn
[*Leaf-CE6851HI-5&CE6851HI-6-bd10-evpn] route-distinguisher 15:1
[*Leaf-CE6851HI-5&CE6851HI-6-bd10-evpn] vpn-target 10:1
[*Leaf-CE6851HI-5&CE6851HI-6-bd10-evpn] vpn-target 11:1 export-extcommunity //配置和VPN实例交叉的RT值
[*Leaf-CE6851HI-5&CE6851HI-6-bd10-evpn] quit
[*Leaf-CE6851HI-5&CE6851HI-6-bd10] quit
[*Leaf-CE6851HI-5&CE6851HI-6-bd10] commit
```

SVF上配置VPN实例。

```
[~Leaf-CE6851HI-5&CE6851HI-3] ip vpn-instance vpn2
[*Leaf-CE6851HI-5&CE6851HI-3-vpn-instance-vpn2] vxlan vni 10
[*Leaf-CE6851HI-5&CE6851HI-3-vpn-instance-vpn2] ipv4-family
[*Leaf-CE6851HI-5&CE6851HI-3-vpn-instance-vpn2-af-ipv4] route-distinguisher 12:4
[*Leaf-CE6851HI-5&CE6851HI-3-vpn-instance-vpn2-af-ipv4] vpn-target 11:1
[*Leaf-CE6851HI-5&CE6851HI-3-vpn-instance-vpn2-af-ipv4] vpn-target 11:1 evpn //需要携带EVPN参数,用于和EVPN实例交叉
[*Leaf-CE6851HI-5&CE6851HI-3-vpn-instance-vpn2-af-ipv4] quit
[*Leaf-CE6851HI-5&CE6851HI-3-vpn-instance-vpn2] quit
[*Leaf-CE6851HI-5&CE6851HI-3] commit
```

步骤4 在SVF上配置头端复制功能。

```
[~Leaf-CE6851HI-5&CE6851HI-6] interface nve 1

[*Leaf-CE6851HI-5&CE6851HI-6-Nve1] source 11.11.11.17

[*Leaf-CE6851HI-5&CE6851HI-6-Nve1] vni 10010 head-end peer-list protocol bgp

[*Leaf-CE6851HI-5&CE6851HI-6-Nve1] quit

[*Leaf-CE6851HI-5&CE6851HI-6] commit
```

步骤5 在SVF上配三层网关。

```
[~Leaf-CE6851HI-5&CE6851HI-6] interface vbdif 10
[*Leaf-CE6851HI-5&CE6851HI-6-Vbdif10] ip binding vpn-instance vpn2
[*Leaf-CE6851HI-5&CE6851HI-6-Vbdif10] ip address 192.168.10.5 255.255.255.0
[*Leaf-CE6851HI-5&CE6851HI-6-Vbdif10] vxlan anycast-gateway enable
[*Leaf-CE6851HI-5&CE6851HI-6-Vbdif10] arp collect host enable
[*Leaf-CE6851HI-5&CE6851HI-6-Vbdif10] quit
[*Leaf-CE6851HI-5&CE6851HI-6] commit
```

----结束

3.2.1.4 配置 Spine 节点

在两台Spine节点上配置上、下联的互联接口地址和路由,使Underlay网络三层互通。

配置思路

1. 配置IP地址:分别配置Spine节点与Leaf和GW设备的三层互联地址、Loopback地址(作为Router-ID)。

- 2. 配置路由:在两台Spine节点上分别配置BGP动态路由,邻居为Leaf堆叠组、LeafM-LAG组两台设备、GW M-LAG两台设备,使Spine上的网段与Leaf和GW上的网段三层可达。
- 3. 配置BGP EVPN: 配置BGP EVPN作为VXLAN的控制平面、并配置BGP EVPN对等体。将Spine节点作为BGP路由反射器,Leaf/网关作为反射器客户端。

配置 IP 地址

步骤1 配置互连接口IP地址。

在Spine-CE12804-1上配置接口IP地址。

```
[~HUAWEI] sysname Spine-CE12804-1
[*SHUAWEI] commit
[~Spine-CE12804-1] interface 40ge 1/0/0
[~Spine-CE12804-1-40GE1/0/0] undo portswitch
[*Spine-CE12804-1-40GE1/0/0] ip address 192.168.40.158 30
[*Spine-CE12804-1-40GE1/0/0] commit
[~Spine-CE12804-1-40GE1/0/0] quit
[~Spine-CE12804-1] interface 40ge 1/0/1
[~Spine-CE12804-1-40GE1/0/1] undo portswitch
[*Spine-CE12804-1-40GE1/0/1] ip address 192.168.40.170 30
[*Spine-CE12804-1-40GE1/0/1] commit
[~Spine-CE12804-1-40GE1/0/1] quit
[~Spine-CE12804-1] interface 40ge 1/0/2
[~Spine-CE12804-1-40GE1/0/2] description "to-Leaf-CE6851-3-40GE1/0/3"
[*Spine-CE12804-1-40GE1/0/2] undo portswitch
[*Spine-CE12804-1-40GE1/0/2] ip address 192.168.41.158 30
[*Spine-CE12804-1-40GE1/0/2] commit
[~Spine-CE12804-1-40GE1/0/2] quit
[~Spine-CE12804-1] interface 40ge 1/0/3
[*Spine-CE12804-1-40GE1/0/3] description "to-Leaf-CE6851-4-40GE1/0/4"
[*Spine-CE12804-1-40GE1/0/3] undo portswitch
[*Spine-CE12804-1-40GE1/0/3] ip address 192.168.41.170 30
[*Spine-CE12804-1-40GE1/0/3] commit
[~Spine-CE12804-1-40GE1/0/3] quit
[~Spine-CE12804-1] interface 40ge 1/0/4
[~Spine-CE12804-1-40GE1/0/4] undo portswitch
[*Spine-CE12804-1-40GE1/0/4] ip address 192.168.42.157 30
[*Spine-CE12804-1-40GE1/0/4] commit
[~Spine-CE12804-1-40GE1/0/4] quit
[~Spine-CE12804-1] interface 40ge 1/0/5
[~Spine-CE12804-1-40GE1/0/5] undo portswitch
[*Spine-CE12804-1-40GE1/0/5] ip address 192.168.42.161 30
[*Spine-CE12804-1-40GE1/0/5] commit
[~Spine-CE12804-1-40GE1/0/5] quit
[~Spine-CE12804-1] interface 40ge 1/0/6
[~Spine-CE12804-1-40GE1/0/6] undo portswitch
[*Spine-CE12804-1-40GE1/0/6] ip address 192.168.46.158 30
[*Spine-CE12804-1-40GE1/0/6] commit
[~Spine-CE12804-1-40GE1/0/6] quit
[~Spine-CE12804-1] interface 40ge 1/0/7
[~Spine-CE12804-1-40GE1/0/7] undo portswitch
[*Spine-CE12804-1-40GE1/0/7] ip address 192.168.46.170 30
[*Spine-CE12804-1-40GE1/0/7] commit
[~Spine-CE12804-1-40GE1/0/7] quit
```

在Spine-CE12804-2节点上配置接口IP地址。

```
[-HUAWEI] sysname Spine-CE12804-2
[*HUAWEI] commit
[-Spine-CE12804-2] interface 40ge 1/0/0
[-Spine-CE12804-2-40GE1/0/0] undo portswitch
[*Spine-CE12804-2-40GE1/0/0] ip address 192.168.40.162 30
[*Spine-CE12804-2-40GE1/0/0] commit
[-Spine-CE12804-2-40GE1/0/0] quit
[-Spine-CE12804-2] interface 40ge 1/0/1
```

```
[~Spine-CE12804-2-40GE1/0/1] undo portswitch
[*Spine-CE12804-2-40GE1/0/1] ip address 192.168.40.166 30
[*Spine-CE12804-2-40GE1/0/1] commit
[~Spine-CE12804-2-40GE1/0/1] quit
[~Spine-CE12804-2] interface 40ge 1/0/2
[*Spine-CE12804-2-40GE1/0/2] description "to-Leaf-CE6851-4-40GE1/0/3"
[*Spine-CE12804-2-40GE1/0/2] undo portswitch
[*Spine-CE12804-2-40GE1/0/2] ip address 192.168.41.162 30
[*Spine-CE12804-2-40GE1/0/2] commit
[~Spine-CE12804-2-40GE1/0/2] quit
[~Spine-CE12804-2] interface 40ge 1/0/3
[*Spine-CE12804-2-40GE1/0/3] description "to-Leaf-CE6851-3-40GE1/0/4"
[*Spine-CE12804-2-40GE1/0/3] undo portswitch
[*Spine-CE12804-2-40GE1/0/3] ip address 192.168.41.166 30
[*Spine-CE12804-2-40GE1/0/3] commit
[~Spine-CE12804-2-40GE1/0/3] quit
[~Spine-CE12804-2] interface 40ge 1/0/4
[~Spine-CE12804-2-40GE1/0/4] undo portswitch
[*Spine-CE12804-2-40GE1/0/4] ip address 192.168.43.157 30
[*Spine-CE12804-2-40GE1/0/4] commit
[~Spine-CE12804-2-40GE1/0/4] quit
[~Spine-CE12804-2] interface 40ge 1/0/5
[~Spine-CE12804-2-40GE1/0/5] undo portswitch
[*Spine-CE12804-2-40GE1/0/5] ip address 192.168.43.161 30
[*Spine-CE12804-2-40GE1/0/5] commit
[~Spine-CE12804-2-40GE1/0/5] quit
[~Spine-CE12804-2] interface 40ge 1/0/6
[~Spine-CE12804-2-40GE1/0/6] undo portswitch
[*Spine-CE12804-2-40GE1/0/6] ip address 192.168.46.166 30
[*Spine-CE12804-2-40GE1/0/6] commit
[~Spine-CE12804-2-40GE1/0/6] quit
[~Spine-CE12804-2] interface 40ge 1/0/7
[~Spine-CE12804-2-40GE1/0/7] undo portswitch
[*Spine-CE12804-2-40GE1/0/7] ip address 192.168.46.162 30
[*Spine-CE12804-2-40GE1/0/7] commit
[~Spine-CE12804-2-40GE1/0/7] quit
```

步骤2 配置Loopback地址。

```
[~Spine-CE12804-1] interface loopback 0
[*Spine-CE12804-1-LoopBack0] ip address 11.11.11.14 32
[*Spine-CE12804-1-LoopBack0] commit
[~Spine-CE12804-1-LoopBack0] quit

[~Spine-CE12804-2] interface loopback 0
[*Spine-CE12804-2-LoopBack0] ip address 11.11.11.15 32
[*Spine-CE12804-2-LoopBack0] commit
[~Spine-CE12804-2-LoopBack0] quit
```

----结束

配置路由

执行以下配置Underlay的路由(本例以BGP为例,也可以使用OSPF)。

步骤1 在Spine-CE12804-1上配置BGP路由。

```
[~Spine-CE12804-1] bgp 65010
[*Spine-CE12804-1-bgp] router-id 11.11.11.14
[*Spine-CE12804-1-bgp] timer keepalive 10 hold 30
[*Spine-CE12804-1-bgp] group Leaf-CE6851HI-1&CE6851HI-2 external
[*Spine-CE12804-1-bgp] peer Leaf-CE6851HI-1&CE6851HI-2 as-number 65021
[*Spine-CE12804-1-bgp] peer 192.168.40.157 as-number 65021
[*Spine-CE12804-1-bgp] peer 192.168.40.157 group Leaf-CE6851HI-1&CE6851HI-2
[*Spine-CE12804-1-bgp] peer 192.168.40.169 as-number 65021
[*Spine-CE12804-1-bgp] peer 192.168.40.169 group Leaf-CE6851HI-1&CE6851HI-2
[*Spine-CE12804-1-bgp] peer 192.168.40.169 group Leaf-CE6851HI-1&CE6851HI-2
[*Spine-CE12804-1-bgp] peer Leaf-CE6851HI-3&4 external
[*Spine-CE12804-1-bgp] peer Leaf-CE6851HI-3&4 as-number 65022
[*Spine-CE12804-1-bgp] peer 192.168.41.157 as-number 65022
```

```
[*Spine-CE12804-1-bgp] peer 192.168.41.157 group Leaf-CE6851HI-3&4
[*Spine-CE12804-1-bgp] peer 192.168.41.169 as-number 65022
[*Spine-CE12804-1-bgp] peer 192.168.41.169 group Leaf-CE6851HI-3&4
[*Spine-CE12804-1-bgp] group Leaf-CE6851HI-5&CE6851HI-6 external
[*Spine-CE12804-1-bgp] peer Leaf-CE6851HI-5&CE6851HI-6 as-number 65024
[*Spine-CE12804-1-bgp] peer 192.168.46.157 as-number 65024
[*Spine-CE12804-1-bgp] peer 192.168.46.157 group Leaf-CE6851HI-5&CE6851HI-6
[*Spine-CE12804-1-bgp] peer 192.168.46.169 as-number 65024
[*Spine-CE12804-1-bgp] peer 192.168.46.169 group Leaf-CE6851HI-5&CE6851HI-6
[*Spine-CE12804-1-bgp] group Gateway-CE12808 external
[*Spine-CE12804-1-bgp] peer Gateway-CE12808 as-number 65000
[*Spine-CE12804-1-bgp] peer 192.168.42.158 as-number 65000
[*Spine-CE12804-1-bgp] peer 192.168.42.158 group Gateway-CE12808
[*Spine-CE12804-1-bgp] peer 192.168.42.162 as-number 65000
[*Spine-CE12804-1-bgp] peer 192.168.42.162 group Gateway-CE12808
[*Spine-CE12804-1-bgp] ipv4-family unicast
[*Spine-CE12804-1-bgp-af-ipv4] preference 20 200 10
[*Spine-CE12804-1-bgp-af-ipv4] network 11.11.11.14 255.255.255.255
[*Spine-CE12804-1-bgp-af-ipv4] maximum load-balancing 32
[*Spine-CE12804-1-bgp-af-ipv4] quit
[*Spine-CE12804-1-bgp] quit
[*Spine-CE12804-1] commit
```

步骤2 在Spine-CE12804-2上配置BGP路由。

```
[~Spine-CE12804-2] bgp 65010
[*Spine-CE12804-2-bgp] router-id 11.11.11.15
[*Spine-CE12804-2-bgp] timer keepalive 10 hold 30
[*Spine-CE12804-2-bgp] group Leaf-CE6851HI-1&CE6851HI-2 external
[*Spine-CE12804-2-bgp] peer Leaf-CE6851HI-1&CE6851HI-2 as-number 65021
[*Spine-CE12804-2-bgp] peer 192.168.40.165 as-number 65021
[*Spine-CE12804-2-bqp] peer 192.168.40.165 group Leaf-CE6851HI-1&CE6851HI-2
[*Spine-CE12804-2-bgp] peer 192.168.40.161 as-number 65021
[*Spine-CE12804-2-bgp] peer 192.168.40.161 group Leaf-CE6851HI-1&CE6851HI-2
[*Spine-CE12804-2-bgp] group Leaf-CE6851HI-3&4 external
[*Spine-CE12804-2-bgp] peer Leaf-CE6851HI-3&4 as-number 65022
[*Spine-CE12804-2-bgp] peer 192.168.41.165 as-number 65022
[*Spine-CE12804-2-bgp] peer 192.168.41.165 group Leaf-CE6851HI-3&4
[*Spine-CE12804-2-bgp] peer 192.168.41.161 as-number 65022
[*Spine-CE12804-2-bgp] peer 192.168.41.161 group Leaf-CE6851HI-3&4
[*Spine-CE12804-2-bgp] group Gateway-CE12808 external
[*Spine-CE12804-2-bgp] peer Gateway-CE12808 as-number 65000
[*Spine-CE12804-2-bgp] peer 192.168.43.162 as-number 65000
[*Spine-CE12804-2-bgp] peer 192.168.43.162 group Gateway-CE12808
[*Spine-CE12804-2-bgp] peer 192.168.43.158 as-number 65000
[*Spine-CE12804-2-bgp] peer 192.168.43.158 group Gateway-CE12808
[*Spine-CE12804-2-bgp] group Leaf-CE6851HI-5&CE6851HI-6 external
[*Spine-CE12804-2-bgp] peer Leaf-CE6851HI-5&CE6851HI-6 as-number 65024
[*Spine-CE12804-2-bgp] peer 192.168.46.165 as-number 65024
[*Spine-CE12804-2-bgp] peer 192.168.46.165 group Leaf-CE6851HI-5&CE6851HI-6
[*Spine-CE12804-2-bgp] peer 192.168.46.161 as-number 65024
[*Spine-CE12804-2-bgp] peer 192.168.46.161 group Leaf-CE6851HI-5&CE6851HI-6
[*Spine-CE12804-2-bgp] ipv4-family unicast
[*Spine-CE12804-2-bgp-af-ipv4] preference 20 200 10
[*Spine-CE12804-2-bgp-af-ipv4] network 11.11.11.15 255.255.255.255
[*Spine-CE12804-2-bgp-af-ipv4] maximum load-balancing 32
[*Spine-CE12804-2-bgp-af-ipv4] quit
[*Spine-CE12804-2-bgp] quit
[*Spine-CE12804-2] commit
```

----结束

配置 BGP-EVPN

使用BGP EVPN作为VXLAN的控制平面时,需执行以下步骤将Leaf作为BGP路由反射器客户端,Spine作为BGP路由器反射器,在Leaf/网关与Spine之间建立IBGP EVPN邻居。

步骤1 在Spine-CE12804-1和Spine-CE12804-2上使能EVPN作为VXLAN的控制平面功能。

```
[~Spine-CE12804-1] evpn-overlay enable
[*Spine-CE12804-1] commit

[~Spine-CE12804-2] evpn-overlay enable
[*Spine-CE12804-2] commit
```

步骤2 配置Spine-CE12804-1和Spine-CE12804-2上分别与Leaf/网关节点建立BGP EVPN对等体关系,Spine作为路由反射器RR。

在Spine-CE12804-1上配置与Leaf/网关建立IBGP EVPN对等体关系,Spine-CE12804-1作为路由反射器,Leaf/网关作为反射器客户端。

```
[~Spine-CE12804-1] bgp 100 instance evpn1
[*Spine-CE12804-1-bgp-instance-evpn1] router-id 11.11.11.14
[*Spine-CE12804-1-bgp-instance-evpn1] peer 11.11.11.11 as-number 100
//与堆叠组建立BGP EVPN对等体关系
[*Spine-CE12804-1-bgp-instance-evpn1] peer 11.11.11.11 connect-interface loopback 0
[*Spine-CE12804-1-bgp-instance-evpn1] peer 13.13.13.13 as-number 100
//与Leaf-CE6851HI-3建立BGP EVPN对等体关系
[*Spine-CE12804-1-bgp-instance-evpn1] peer 13.13.13 connect-interface loopback 0
[*Spine-CE12804-1-bgp-instance-evpn1] peer 14.14.14.14 as-number 100
//与Leaf-CE6851HI-4建立BGP EVPN对等体关系
[*Spine-CE12804-1-bgp-instance-evpn1] peer 14.14.14 connect-interface loopback 0
[*Spine-CE12804-1-bgp-instance-evpn1] peer 11.11.11.17 as-number 100
//与SVF组建立BGP EVPN对等体关系
[*Spine-CE12804-1-bgp-instance-evpn1] peer 11.11.11.17 connect-interface loopback 0
[*Spine-CE12804-1-bgp-instance-evpn1] peer 18.18.18.18 as-number 100
//与Exit-Gateway-CE12808-1建立BGP EVPN对等体关系
[*Spine-CE12804-1-bgp-instance-evpn1] peer 18.18.18 connect-interface loopback 0
[*Spine-CE12804-1-bgp-instance-evpn1] peer 19.19.19.19 as-number 100
//与Exit-Gateway-CE12808-2建立BGP EVPN对等体关系
[*Spine-CE12804-1-bgp-instance-evpn1] peer 19.19.19.19 connect-interface loopback 0
[*Spine-CE12804-1-bgp-instance-evpn1] l2vpn-family evpn
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] undo policy vpn-target
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 11.11.11.11 enable
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 11.11.11.11 reflect-client
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 11.11.11.11 next-hop-invariable
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 11.11.11.11 advertise irb
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 13.13.13.13 enable
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 13.13.13.13 reflect-client
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 13.13.13.13 next-hop-invariable
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 13.13.13.13 advertise irb
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 14.14.14.14 enable
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 14.14.14.14 reflect-client
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 14.14.14.14 next-hop-invariable
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 14.14.14.14 advertise irb
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 11.11.11.17 enable
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 11.11.11.17 reflect-client
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 11.11.11.17 next-hop-invariable
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 11.11.11.17 advertise irb
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 18.18.18.18 enable
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 18.18.18.18 reflect-client
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 18.18.18.18 next-hop-invariable
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 18.18.18.18 advertise irb
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 19.19.19.19 enable
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 19.19.19.19 reflect-client
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 19.19.19.19 next-hop-invariable
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] peer 19.19.19.19 advertise irb
[*Spine-CE12804-1-bgp-instance-evpn1-af-evpn] quit
[*Spine-CE12804-1-bgp-instance-evpn1] quit
[*Spine-CE12804-1] commit
```

在Spine-CE12804-2上配置与Leaf建立IBGP EVPN对等体关系,Spine-CE12804-2作为路由反射器,Leaf作为反射器客户端。

```
[~Spine-CE12804-2] bgp 100 instance evpn1
[*Spine-CE12804-2-bgp-instance-evpn1] router-id 11.11.11.15
[*Spine-CE12804-2-bgp-instance-evpn1] peer 11.11.11.11 as-number 100
//与堆叠组建立BGP EVPN对等体关系
[*Spine-CE12804-2-bgp-instance-evpn1] peer 11.11.11 connect-interface loopback 0
[*Spine-CE12804-2-bgp-instance-evpn1] peer 13.13.13.13 as-number 100
//与Leaf-CE6851HI-3建立BGP EVPN对等体关系
[*Spine-CE12804-2-bgp-instance-evpn1] peer 13.13.13 connect-interface loopback 0
[*Spine-CE12804-2-bgp-instance-evpn1] peer 14.14.14.14 as-number 100
//与Leaf-CE6851HI-4建立BGP EVPN对等体关系
[*Spine-CE12804-2-bgp-instance-evpn1] peer 14.14.14.14 connect-interface loopback 0
[*Spine-CE12804-2-bgp-instance-evpn1] peer 11.11.11.17 as-number 100
//与SVF组建立BGP EVPN对等体关系
[*Spine-CE12804-2-bgp-instance-evpn1] peer 11.11.11.17 connect-interface loopback 0
[*Spine-CE12804-2-bgp-instance-evpn1] peer 18.18.18.18 as-number 100
[*Spine-CE12804-2-bgp-instance-evpn1] peer 18.18.18 connect-interface loopback 0
[*Spine-CE12804-2-bgp-instance-evpn1] peer 19.19.19.19 as-number 100
//与Exit-Gateway-CE12808-2建立BGP EVPN对等体关系
[*Spine-CE12804-2-bgp-instance-evpn1] peer 19.19.19.19 connect-interface loopback 0
[*Spine-CE12804-2-bgp-instance-evpn1] l2vpn-family evpn
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] undo policy vpn-target
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 11.11.11.11 enable
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 11.11.11.11 reflect-client
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 11.11.11.11 next-hop-invariable
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 11.11.11.11 advertise irb
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 13.13.13.13 enable
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 13.13.13.13 reflect-client
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 13.13.13.13 next-hop-invariable
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 13.13.13.13 advertise irb
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 14.14.14.14 enable
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 14.14.14.14 reflect-client
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 14.14.14.14 next-hop-invariable
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 14.14.14.14 advertise irb
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 11.11.11.17 enable
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 11.11.11.17 reflect-client
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 11.11.11.17 next-hop-invariable
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 11.11.11.17 advertise irb
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 18.18.18.18 enable
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 18.18.18.18 reflect-client
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 18.18.18.18 next-hop-invariable
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 18.18.18.18 advertise irb
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 19.19.19.19 enable
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 19.19.19.19 reflect-client
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 19.19.19.19 next-hop-invariable
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] peer 19.19.19.19 advertise irb
[*Spine-CE12804-2-bgp-instance-evpn1-af-evpn] quit
[*Spine-CE12804-2-bgp-instance-evpn1] quit
[*Spine-CE12804-2] commit
```

----结束

3.2.1.5 配置网关工作组

VXLAN分布式出口网关由两台CE交换机Exit-Gateway-CE12808-1和Exit-Gateway-CE12808-2承担。可以是堆叠模式也可以是M-LAG。本节以M-LAG形式为例进行介绍。

配置思路

1. 配置IP地址:分别配置与Spine节点,并配置防火墙的管理VLAN地址;配置Loopback0地址(作为VTEP地址)、Loopback1地址(作为Router-ID)。

- 2. 配置M-LAG:在GW交换机上配置M-LAG全局模式、DFS组、Peer-Link,并分别配置与两台防火墙的互联管理VLAN和业务链路。
- 3. 配置路由: Exit-Gateway上分别配置BGP动态路由,邻居为两个Spine设备,使 Exit-Gateway上的网段与Spine上的网段三层可达;配置与外部路由器的路由。
- 4. 配置BGP EVPN:配置BGP EVPN作为VXLAN的控制平面、并配置BGP EVPN对等 体。
- 5. 配置MAC漂移白名单:在Exit-Gateway设备上将Spine设备与Exit-Gateway互联的三层口MAC地址配置为白名单,不再检测MAC漂移。
- 6. 配置网关设备与外部路由器的互联,使网关与外部路由器形成"口"字型和交叉结合的出口网络:分别配置网关与外部路由器之间的互联接口地址、配置网关设备之间的互联接口地址、配置网关设备上的默认路由。

配置 IP 地址

步骤1 配置接口IP地址。

#配置Exit-Gateway-CE12808-1的接口IP地址。

```
[~Exit-Gateway-CE12808-1] interface 40ge 1/0/0 //连接Spine节点
[~Exit-Gateway-CE12808-1-40GE1/0/0] undo portswitch
[*Exit-Gateway-CE12808-1-40GE1/0/0] ip address 192.168.42.158 30
[*Exit-Gateway-CE12808-1-40GE1/0/0] commit
[~Exit-Gateway-CE12808-1-40GE1/0/0] quit
[~Exit-Gateway-CE12808-1] interface 40ge 1/0/1 //连接Spine节点
[~Exit-Gateway-CE12808-1-40GE1/0/1] undo portswitch
[*Exit-Gateway-CE12808-1-40GE1/0/1] ip address 192.168.43.162 30
[*Exit-Gateway-CE12808-1-40GE1/0/1] commit
[~Exit-Gateway-CE12808-1-40GE1/0/1] quit
[~Exit-Gateway-CE12808-1] vlan batch 11
[*Exit-Gateway-CE12808-1] interface vlanif 11 //网关设备与防火墙的Public互联VLAN
[*Exit-Gateway-CE12808-1-Vlanif11] description "to firewall-1~2"
*Exit-Gateway-CE12808-1-Vlanif11 ip address 192.168.45.154 29
[*Exit-Gateway-CE12808-1-Vlanif11] vrrp vrid 1 virtual-ip 192.168.45.153
[*Exit-Gateway-CE12808-1-Vlanif11] commit
[~Exit-Gateway-CE12808-1-Vlanif11] quit
[~Exit-Gateway-CE12808-1] vlan batch 12
[*Exit-Gateway-CE12808-1] interface vlanif 12 //网关设备与防火墙的业务互联VLAN
[*Exit-Gateway-CE12808-1-Vlanif12] description "to firewall-1~2"
[*Exit-Gateway-CE12808-1-Vlanif12] ip address 192.168.45.162 29
[*Exit-Gateway-CE12808-1-Vlanif12] vrrp vrid 2 virtual-ip 192.168.45.161
[*Exit-Gateway-CE12808-1-Vlanif12] commit
[~Exit-Gateway-CE12808-1-Vlanif12] quit
```

#配置Exit-Gateway-CE12808-2的接口IP地址。

```
[~Exit-Gateway-CE12808-2] interface 40ge 1/0/0 //连接Spine节点
[~Exit-Gateway-CE12808-2-40GE1/0/0] undo portswitch
[*Exit-Gateway-CE12808-2-40GE1/0/0] ip address 192.168.42.162 30
[*Exit-Gateway-CE12808-2-40GE1/0/0] commit
[~Exit-Gateway-CE12808-2-40GE1/0/0] quit
[~Exit-Gateway-CE12808-2] interface 40ge 1/0/1 //连接Spine节点
[~Exit-Gateway-CE12808-2-40GE1/0/1] undo portswitch
[*Exit-Gateway-CE12808-2-40GE1/0/1] ip address 192.168.43.158 30
[*Exit-Gateway-CE12808-2-40GE1/0/1] commit
[~Exit-Gateway-CE12808-2-40GE1/0/1] quit
[~Exit-Gateway-CE12808-2] vlan batch 11
[*Exit-Gateway-CE12808-2] interface vlanif 11 //网关设备与防火墙的Public互联VLAN
[*Exit-Gateway-CE12808-2-Vlanif11] description "to firewall-1-2"
[*Exit-Gateway-CE12808-2-Vlanif11] ip address 192.168.45.155 29
[*Exit-Gateway-CE12808-2-Vlanif11] vrrp vrid 1 virtual-ip 192.168.45.153
[*Exit-Gateway-CE12808-2-Vlanif11] commit
[~Exit-Gateway-CE12808-2-Vlanif11] quit
```

```
[~Exit-Gateway-CE12808-2] vlan batch 12
[*Exit-Gateway-CE12808-2] interface vlanif 12 //网关设备与防火墙的业务互联VLAN
[*Exit-Gateway-CE12808-2-Vlanif12] description "to firewall-1-2"
[*Exit-Gateway-CE12808-2-Vlanif12] ip address 192.168.45.163 29
[*Exit-Gateway-CE12808-2-Vlanif12] vrrp vrid 2 virtual-ip 192.168.45.161
[*Exit-Gateway-CE12808-2-Vlanif12] commit
[~Exit-Gateway-CE12808-2-Vlanif12] quit
```

步骤2 配置Loopback口地址。

```
[~Exit-Gateway-CE12808-1] interface loopback 0 //作为VTEP地址
[*Exit-Gateway-CE12808-1-LoopBack0] ip address 11.11.11.16 32
[*Exit-Gateway-CE12808-1-LoopBack0] commit
[~Exit-Gateway-CE12808-1-LoopBack0] quit
[~Exit-Gateway-CE12808-1] interface loopback 1
[*Exit-Gateway-CE12808-1-LoopBack1] ip address 18.18.18.18 32
[*Exit-Gateway-CE12808-1-LoopBack1] commit
[~Exit-Gateway-CE12808-1-LoopBack1] quit
[~Exit-Gateway-CE12808-1] interface loopback 2
[*Exit-Gateway-CE12808-1-LoopBack2] ip address 21.21.21.21 32
[*Exit-Gateway-CE12808-1-LoopBack2] commit
[~Exit-Gateway-CE12808-1-LoopBack2] quit
[~Exit-Gateway-CE12808-2] interface loopback 0 //作为VTEP地址
[*Exit-Gateway-CE12808-2-LoopBack0] ip address 11.11.11.16 32
[*Exit-Gateway-CE12808-2-LoopBack0] commit
[~Exit-Gateway-CE12808-2-LoopBack0] quit
[~Exit-Gateway-CE12808-2] interface loopback 1
[*Exit-Gateway-CE12808-2-LoopBack1] ip address 19.19.19.19 32
[*Exit-Gateway-CE12808-2-LoopBack1] commit
[~Exit-Gateway-CE12808-2-LoopBack1] quit
[~Exit-Gateway-CE12808-2] interface loopback 2
[*Exit-Gateway-CE12808-2-LoopBack2] ip address 22.22.22.22 32
[*Exit-Gateway-CE12808-2-LoopBack2] commit
[~Exit-Gateway-CE12808-2-LoopBack2] quit
```

----结束

配置路由

执行以下步骤配置Underlay的路由(本例以BGP为例,也可以使用OSPF)。

步骤1 在Exit-Gateway-CE12808-1上配置BGP路由用于打通Underlay层面的路由。

```
[~Exit-Gateway-CE12808-1] bgp 65000
[*Exit-Gateway-CE12808-1-bgp] router-id 18.18.18.18
[*Exit-Gateway-CE12808-1-bgp] timer keepalive 10 hold 30
[*Exit-Gateway-CE12808-1-bgp] group Spine-CE12804 external //对接Spine
[*Exit-Gateway-CE12808-1-bgp] peer Spine-CE12804 as-number 65010 [*Exit-Gateway-CE12808-1-bgp] peer 192.168.42.157 as-number 65010
[*Exit-Gateway-CE12808-1-bgp] peer 192.168.42.157 group Spine-CE12804
[*Exit-Gateway-CE12808-1-bgp] peer 192.168.43.161 as-number 65010
[*Exit-Gateway-CE12808-1-bgp] peer 192.168.43.161 group Spine-CE12804
[*Exit-Gateway-CE12808-1-bgp] ipv4-family unicast
[*Exit-Gateway-CE12808-1-bgp-af-ipv4] preference 20 200 10
[*Exit-Gateway-CE12808-1-bgp-af-ipv4] network 11.11.11.16 255.255.255.255
[*Exit-Gateway-CE12808-1-bgp-af-ipv4] network 18.18.18 255.255.255.255
[*Exit-Gateway-CE12808-1-bgp-af-ipv4] network 192.168.45.152 255.255.255.248
[*Exit-Gateway-CE12808-1-bgp-af-ipv4] maximum load-balancing 32
[*Exit-Gateway-CE12808-1-bgp-af-ipv4] quit
[*Exit-Gateway-CE12808-1-bgp] quit
[*Exit-Gateway-CE12808-1] commit
```

步骤2 在Exit-Gateway-CE12808-2上配置BGP路由用于打通Underlay层面的路由。

```
[~Exit-Gateway-CE12808-2] bgp 65000
[*Exit-Gateway-CE12808-2-bgp] router-id 19.19.19.19
[*Exit-Gateway-CE12808-2-bgp] timer keepalive 10 hold 30
[*Exit-Gateway-CE12808-2-bgp] group Spine-CE12804 external //对接Spine
```

```
[*Exit-Gateway-CE12808-2-bgp] peer Spine-CE12804 as-number 65010
[*Exit-Gateway-CE12808-2-bgp] peer 192.168.42.161 as-number 65010
[*Exit-Gateway-CE12808-2-bgp] peer 192.168.42.161 group Spine-CE12804
[*Exit-Gateway-CE12808-2-bgp] peer 192.168.43.157 as-number 65010
[*Exit-Gateway-CE12808-2-bgp] peer 192.168.43.157 group Spine-CE12804

[*Exit-Gateway-CE12808-2-bgp] ipv4-family unicast
[*Exit-Gateway-CE12808-2-bgp-af-ipv4] preference 20 200 10
[*Exit-Gateway-CE12808-2-bgp-af-ipv4] network 11.11.11.16 255.255.255.255
[*Exit-Gateway-CE12808-2-bgp-af-ipv4] network 19.19.19 255.255.255.255
[*Exit-Gateway-CE12808-2-bgp-af-ipv4] network 192.168.45.152 255.255.255.255
[*Exit-Gateway-CE12808-2-bgp-af-ipv4] maximum load-balancing 32
[*Exit-Gateway-CE12808-2-bgp-af-ipv4] quit
[*Exit-Gateway-CE12808-2-bgp] quit
[*Exit-Gateway-CE12808-2-bgp] quit
[*Exit-Gateway-CE12808-2] commit
```

----结束

组建 M-LAG 组

步骤1 配置M-LAG模式。

步骤2 在Exit-Gateway-CE12808-1和Exit-Gateway-CE12808-2上配置M-LAG的DFS Group及 网关双活。

```
[~Exit-Gateway-CE12808-1] dfs-group 1
[*Exit-Gateway-CE12808-1-dfs-group-1] source ip 18.18.18.18
[*Exit-Gateway-CE12808-1-dfs-group-1] priority 150
[*Exit-Gateway-CE12808-1-dfs-group-1] quit
[*Exit-Gateway-CE12808-1] commit

[~Exit-Gateway-CE12808-2] dfs-group 1
[*Exit-Gateway-CE12808-2-dfs-group-1] source ip 19.19.19.19
[*Exit-Gateway-CE12808-2-dfs-group-1] priority 120
[*Exit-Gateway-CE12808-2-dfs-group-1] quit
[*Exit-Gateway-CE12808-2] commit
```

步骤3 在Exit-Gateway-CE12808-1和Exit-Gateway-CE12808-2上配置M-LAG的Peer-Link。

```
[~Exit-Gateway-CE12808-1] interface eth-trunk 0
[*Exit-Gateway-CE12808-1-Eth-Trunk0] trunkport 40ge 1/0/23
[*Exit-Gateway-CE12808-1-Eth-Trunk0] trunkport 40ge 2/0/23
[*Exit-Gateway-CE12808-1-Eth-Trunk0] mode lacp-static
[*Exit-Gateway-CE12808-1-Eth-Trunk0] peer-link 1
[*Exit-Gateway-CE12808-1-Eth-Trunk0] quit
[*Exit-Gateway-CE12808-1] commit

[~Exit-Gateway-CE12808-2] interface eth-trunk 0
[*Exit-Gateway-CE12808-2-Eth-Trunk0] trunkport 40ge 1/0/23
[*Exit-Gateway-CE12808-2-Eth-Trunk0] trunkport 40ge 2/0/23
[*Exit-Gateway-CE12808-2-Eth-Trunk0] mode lacp-static
[*Exit-Gateway-CE12808-2-Eth-Trunk0] peer-link 1
```

```
[*Exit-Gateway-CE12808-2-Eth-Trunk0] quit
[*Exit-Gateway-CE12808-2] commit
```

步骤4 在Exit-Gateway-CE12808-1和Exit-Gateway-CE12808-2上配置M-LAG成员口(对接防火墙示例)。

配置Exit-Gateway-CE12808-1与防火墙的互联。

```
[~Exit-Gateway-CE12808-1] interface eth-trunk 21
[*Exit-Gateway-CE12808-1-Eth-Trunk21] description "to-FW-USG9560-1-GE1/0/3"
[*Exit-Gateway-CE12808-1-Eth-Trunk21] port link-type trunk
[*Exit-Gateway-CE12808-1-Eth-Trunk21] undo port trunk allow-pass vlan 1
[*Exit-Gateway-CE12808-1-Eth-Trunk21] port trunk allow-pass vlan 11
[*Exit-Gateway-CE12808-1-Eth-Trunk21] port trunk allow-pass vlan 12
[*Exit-Gateway-CE12808-1-Eth-Trunk21] trunkport 10ge 3/0/0
[*Exit-Gateway-CE12808-1-Eth-Trunk21] trunkport 10ge 3/0/1
[*Exit-Gateway-CE12808-1-Eth-Trunk21] dfs-group 1 m-lag 1
[*Exit-Gateway-CE12808-1-Eth-Trunk21] quit
[*Exit-Gateway-CE12808-1] interface eth-trunk 31
[*Exit-Gateway-CE12808-1-Eth-Trunk31] description "to-FW-USG9560-2-GE1/0/3"
[*Exit-Gateway-CE12808-1-Eth-Trunk31] port link-type trunk
[*Exit-Gateway-CE12808-1-Eth-Trunk31] undo port trunk allow-pass vlan 1
[*Exit-Gateway-CE12808-1-Eth-Trunk31] port trunk allow-pass vlan 11
[*Exit-Gateway-CE12808-1-Eth-Trunk31] port trunk allow-pass vlan 12
[*Exit-Gateway-CE12808-1-Eth-Trunk31] trunkport 10ge 3/0/2
[*Exit-Gateway-CE12808-1-Eth-Trunk31] trunkport 10ge 3/0/3
[*Exit-Gateway-CE12808-1-Eth-Trunk31] dfs-group 1 m-lag 2
[*Exit-Gateway-CE12808-1-Eth-Trunk31] quit
[*Exit-Gateway-CE12808-1] commit
```

#配置Gateway-CE12808-2与防火墙的互联。

```
[~Exit-Gateway-CE12808-2] interface eth-trunk 21
[*Exit-Gateway-CE12808-2-Eth-Trunk21] description "to-FW-USG9560-1-GE1/0/4"
[*Exit-Gateway-CE12808-2-Eth-Trunk21] port link-type trunk
[*Exit-Gateway-CE12808-2-Eth-Trunk21] undo port trunk allow-pass vlan 1
[*Exit-Gateway-CE12808-2-Eth-Trunk21] port trunk allow-pass vlan 11
[*Exit-Gateway-CE12808-2-Eth-Trunk21] port trunk allow-pass vlan 12
[*Exit-Gateway-CE12808-2-Eth-Trunk21] trunkport 10ge 3/0/0
[*Exit-Gateway-CE12808-2-Eth-Trunk21] trunkport 10ge 3/0/1
[*Exit-Gateway-CE12808-2-Eth-Trunk21] dfs-group 1 m-lag 1
[*Exit-Gateway-CE12808-2-Eth-Trunk21] quit
[*Exit-Gateway-CE12808-2] interface eth-trunk 31
[*Exit-Gateway-CE12808-2-Eth-Trunk31] description "to-FW-USG9560-2-GE1/0/4"
[*Exit-Gateway-CE12808-2-Eth-Trunk31] port link-type trunk
[*Exit-Gateway-CE12808-2-Eth-Trunk31] undo port trunk allow-pass vlan 1
[*Exit-Gateway-CE12808-2-Eth-Trunk31] port trunk allow-pass vlan 11
[*Exit-Gateway-CE12808-2-Eth-Trunk31] port trunk allow-pass vlan 12
[*Exit-Gateway-CE12808-2-Eth-Trunk31] trunkport 10ge 3/0/2
[*Exit-Gateway-CE12808-2-Eth-Trunk31] trunkport 10ge 3/0/3
[*Exit-Gateway-CE12808-2-Eth-Trunk31] dfs-group 1 m-lag 2
[*Exit-Gateway-CE12808-2-Eth-Trunk31] quit
[*Exit-Gateway-CE12808-2] commit
```

□ 说明

防火墙对接网关采用M-LAG时,两端的LAG模式仅能为手工负载分担模式。

----结束

配置 BGP EVPN

当使用BGP EVPN作为VXLAN的控制平面时,还需执行以下步骤将网关作为BGP路由反射器客户端,Spine作为BGP路由器反射器,在网关与Spine之间建立IBGP EVPN邻居。

步骤1 在Exit-Gateway-CE12808-1和Gateway-CE12808-2上使能EVPN作为VXLAN的控制平面功能。

```
[~Exit-Gateway-CE12808-1] evpn-overlay enable
[*Exit-Gateway-CE12808-1] commit

[~Exit-Gateway-CE12808-2] evpn-overlay enable
[*Exit-Gateway-CE12808-2] commit
```

步骤2 配置Gateway-CE12808-1和GatewExit-Gateway-CE12808与Spine节点建立BGP EVPN 对等体关系,Spine作为路由反射器RR。

在Exit-Gateway-CE12808-1上配置与Spine建立BGP EVPN对等体关系。

```
[~Exit-Gateway-CE12808-1] bgp 100 instance evpn1
[*Exit-Gateway-CE12808-1-bgp-instance-evpn1] router-id 18.18.18.18
[*Exit-Gateway-CE12808-1-bgp-instance-evpn1] peer 11.11.11.14 as-number 100
//与Spine-CE12804-1建立BGP EVPN对等体关系
[*Exit-Gateway-CE12808-1-bgp-instance-evpn1] peer 11.11.11.14 connect-interface loopback 1
[*Exit-Gateway-CE12808-1-bgp-instance-evpn1] peer 11.11.11.15 as-number 100
//与Spine-CE12804-2建立BGP EVPN对等体关系
[*Exit-Gateway-CE12808-1-bgp-instance-evpn1] peer 11.11.11.15 connect-interface loopback 1
[*Exit-Gateway-CE12808-1-bgp-instance-evpn1] l2vpn-family evpn
[*Exit-Gateway-CE12808-1-bgp-instance-evpn1-af-evpn] peer 11.11.11.14 enable
[*Exit-Gateway-CE12808-1-bgp-instance-evpn1-af-evpn] peer 11.11.11.14 advertise irb
[*Exit-Gateway-CE12808-1-bgp-instance-evpn1-af-evpn] peer 11.11.11.15 enable
[*Exit-Gateway-CE12808-1-bgp-instance-evpn1-af-evpn] peer 11.11.11.15 advertise irb
[*Exit-Gateway-CE12808-1-bgp-instance-evpn1-af-evpn] quit
[*Exit-Gateway-CE12808-1-bgp-instance-evpn1] quit
[*Exit-Gateway-CE12808-1] commit
```

在Exit-Gateway-CE12808-2上配置与Spine建立BGP EVPN对等体关系。

```
[~Exit-Gateway-CE12808-2] bgp 100 instance evpn1
[*Exit-Gateway-CE12808-2-bgp-instance-evpn1] router-id 19.19.19.19
[*Exit-Gateway-CE12808-2-bgp-instance-evpn1] peer 11.11.11.14 as-number 100
//与Spine-CE12804-1建立BGP EVPN对等体关系
[*Exit-Gateway-CE12808-2-bgp-instance-evpn1] peer 11.11.11.14 connect-interface loopback 1
[*Exit-Gateway-CE12808-2-bgp-instance-evpn1] peer 11.11.11.15 as-number 100
//与Spine-CE12804-2建立BGP EVPN对等体关系
[*Exit-Gateway-CE12808-2-bgp-instance-evpn1] peer 11.11.11.15 connect-interface loopback 1
[*Exit-Gateway-CE12808-2-bgp-instance-evpn1] l2vpn-family evpn
[*Exit-Gateway-CE12808-2-bgp-instance-evpn1-af-evpn] peer 11.11.11.14 enable
[*Exit-Gateway-CE12808-2-bgp-instance-evpn1-af-evpn] peer 11.11.11.14 advertise irb
[*Exit-Gateway-CE12808-2-bgp-instance-evpn1-af-evpn] peer 11.11.11.15 enable
[*Exit-Gateway-CE12808-2-bgp-instance-evpn1-af-evpn] peer 11.11.11.15 advertise irb
[*Exit-Gateway-CE12808-2-bgp-instance-evpn1-af-evpn] quit
[*Exit-Gateway-CE12808-2-bgp-instance-evpn1] quit
[*Exit-Gateway-CE12808-2] commit
```

步骤3 配置VPN实例绑定网关。

```
[~Exit-Gateway-CE12808-1] ip vpn-instance vpn1
[*Exit-Gateway-CE12808-1-vpn-instance-vpn1] vxlan vni 10
[*Exit-Gateway-CE12808-1-vpn-instance-vpn1] route-distinguisher 11:1
[*Exit-Gateway-CE12808-1-vpn-instance-vpn1-af-ipv4] vpn-target 1:1
[*Exit-Gateway-CE12808-1-vpn-instance-vpn1-af-ipv4] vpn-target 11:1 evpn
[*Exit-Gateway-CE12808-1-vpn-instance-vpn1-af-ipv4] quit
[*Exit-Gateway-CE12808-1-vpn-instance-vpn1] quit
[*Exit-Gateway-CE12808-1] commit
[~Exit-Gateway-CE12808-2] ip vpn-instance vpn1
[*Exit-Gateway-CE12808-2-vpn-instance-vpn1] vxlan vni 10
[*Exit-Gateway-CE12808-2-vpn-instance-vpn1] route-distinguisher 11:2
[*Exit-Gateway-CE12808-2-vpn-instance-vpn1-af-ipv4] vpn-target 1:1
[*Exit-Gateway-CE12808-2-vpn-instance-vpn1-af-ipv4] vpn-target 11:1 evpn
[*Exit-Gateway-CE12808-2-vpn-instance-vpn1-af-ipv4] quit
[*Exit-Gateway-CE12808-2-vpn-instance-vpn1] quit
[*Exit-Gateway-CE12808-2] commit
```

```
[~Exit-Gateway-CE12808-1] interface vlanif 12 //GW的VRF与防火墙的业务互联VLAN
[~Exit-Gateway-CE12808-1-Vlanif12] description "to firewall-1~2Vsys"
[*Exit-Gateway-CE12808-1-Vlanif12] ip binding vpn-instance vpn1
[*Exit-Gateway-CE12808-1-Vlanif12] commit
[~Exit-Gateway-CE12808-1-Vlanif12] quit
[~Exit-Gateway-CE12808-1] ip route-static vpn-instance vpn1 0.0.0.0 0.0.0.0 192.168.45.164 //GW的VRF
vpn1的默认路由指向防火墙
[*Exit-Gateway-CE12808-1] commit
[~Exit-Gateway-CE12808-2] interface vlanif 12 //网关设备与防火墙的业务互联VLAN
[~Exit-Gateway-CE12808-2-Vlanif12] description "to firewall-1~2Vsys"
[*Exit-Gateway-CE12808-2-Vlanif12] ip binding vpn-instance vpn1
[*Exit-Gateway-CE12808-2-Vlanif12] commit
[~Exit-Gateway-CE12808-2-Vlanif12] quit
[~Exit-Gateway-CE12808-2] ip route-static vpn-instance vpn1 0.0.0.0 0.0.0.0 192.168.45.164 //GW的VRF
vpn1的默认路由指向防火墙
[*Exit-Gateway-CE12808-2] commit
```

步骤4 将VPN实例与IPv4地址族进行关联。

```
[~Exit-Gateway-CE12808-1] bgp 100 instance evpn1
[*Exit-Gateway-CE12808-1-bgp-instance-evpn1] ipv4-family vpn-instance vpn1
[*Exit-Gateway-CE12808-1-bgp-instance-evpn1-vpn1] default-route imported
[*Exit-Gateway-CE12808-1-bgp-instance-evpn1-vpn1] import-route direct
[*Exit-Gateway-CE12808-1-bgp-instance-evpn1-vpn1] import-route static
[*Exit-Gateway-CE12808-1-bgp-instance-evpn1-vpn1] maximum load-balancing 32 //设置形成负载分担的
等价路由的最大条数为32
[*Exit-Gateway-CE12808-1-bgp-instance-evpn1-vpn1] advertise l2vpn evpn
                                                                       //使能VPN实例向EVPN实例
发布IP路由功能。
[*Exit-Gateway-CE12808-1-bgp-instance-evpn1-vpn1] quit
[*Exit-Gateway-CE12808-1-bgp-instance-evpn1] quit
[*Exit-Gateway-CE12808-1] commit
[~Exit-Gateway-CE12808-2] bgp 100 instance evpn1
[*Exit-Gateway-CE12808-2-bgp-instance-evpn1] ipv4-family vpn-instance vpn1
[*Exit-Gateway-CE12808-2-bgp-instance-evpn1-vpn1] default-route imported
[*Exit-Gateway-CE12808-2-bgp-instance-evpn1-vpn1] import-route direct
[*Exit-Gateway-CE12808-2-bgp-instance-evpn1-vpn1] import-route static
[*Exit-Gateway-CE12808-2-bgp-instance-evpn1-vpn1] maximum load-balancing 32 //设置形成负载分担的
等价路由的最大条数为32
[*Exit-Gateway-CE12808-2-bqp-instance-evpn1-vpn1] advertise l2vpn evpn
                                                                       //使能VPN实例向EVPN实例
发布IP路由功能。
[*Exit-Gateway-CE12808-2-bgp-instance-evpn1-vpn1] quit
[*Exit-Gateway-CE12808-2-bgp-instance-evpn1] quit
[*Exit-Gateway-CE12808-2] commit
```

----结束

配置 MAC 漂移白名单

三层架构时,VXLAN流量经过Spine设备到达GW时,可能会由于产品约束限制在GW上学习了错误的MAC地址导致MAC漂移告警,需要在GW设备上基于MAC地址配置MAC漂移白名单,将Spine设备与GW互联的三层口MAC地址配置为白名单不再检测MAC漂移,正常情况下此MAC是作为隧道报文的外层源MAC,不应该被学习,所以调整后对业务功能不会有任何影响。

步骤1 查看两台Spine设备与GW互连的三层口MAC地址。

```
[~Spine-CE12804-1] interface 40ge 1/0/4 //获取Spine-1设备与GW互连的三层口MAC地址 [*Spine-CE12804-1-40GE1/0/4] undo portswitch [*Spine-CE12804-1-40GE1/0/4] commit [~Spine-CE12804-1-40GE1/0/4] display this interface | include Hardware address IP Sending Frames' Format is PKTFMT_ETHNT_2, Hardware address is 200b-c732-d202 [~Spine-CE12804-1-40GE1/0/4] quit [~Spine-CE12804-1] interface 40ge 1/0/5 [~Spine-CE12804-1-40GE1/0/5] undo portswitch [*Spine-CE12804-1-40GE1/0/5] commit [~Spine-CE12804-1-40GE1/0/5] display this interface | include Hardware address
```

```
IP Sending Frames' Format is PKTFMT_ETHNT_2, Hardware address is 200b-c732-d202 [~Spine-CE12804-1-40GE1/0/5] quit

[~Spine-CE12804-2] interface 40ge 1/0/4 //获取Spine-2设备与GW互连的三层口MAC地址 [~Spine-CE12804-2-40GE1/0/4] undo portswitch [*Spine-CE12804-2-40GE1/0/4] commit [~Spine-CE12804-2-40GE1/0/4] display this interface | include Hardware address IP Sending Frames' Format is PKTFMT_ETHNT_2, Hardware address is 346a-c246-be01 [~Spine-CE12804-2-40GE1/0/4] quit [~Spine-CE12804-2] interface 40ge 1/0/5 [~Spine-CE12804-2-40GE1/0/5] undo portswitch [*Spine-CE12804-2-40GE1/0/5] commit [~Spine-CE12804-2-40GE1/0/5] display this interface | include Hardware address IP Sending Frames' Format is PKTFMT_ETHNT_2, Hardware address is 346a-c246-be01 [~Spine-CE12804-2-40GE1/0/5] quit
```

步骤2 根据步骤1中获取到的MAC地址配置白名单。

```
[~Exit-Gateway-CE12808-1] mac-address flapping detection exclude 200b-c732-d202 48
[~Exit-Gateway-CE12808-1] mac-address flapping detection exclude 346a-c246-be01 48

[~Exit-Gateway-CE12808-2] mac-address flapping detection exclude 200b-c732-d202 48
[~Exit-Gateway-CE12808-2] mac-address flapping detection exclude 346a-c246-be01 48
```

----结束

配置网关与外部路由器互联

步骤1 配置网关与外部路由器的互联端口。

#在Gateway-CE12808-1配置与Router的互联端口。

```
[~Exit-Gateway-CE12808-1] interface 10ge 3/0/4 //连接外部PE设备(Router-1)
[~Exit-Gateway-CE12808-1-10GE3/0/4] undo portswitch
[*Exit-Gateway-CE12808-1-10GE3/0/4] ip address 192.168.44.157 30
[*Exit-Gateway-CE12808-1-10GE3/0/4] set up-delay 180 //配置接口上报状态变化事件的延时时间
[*Exit-Gateway-CE12808-1-10GE3/0/4] quit
[*Exit-Gateway-CE12808-1] interface 10ge 3/0/5 //连接外部PE设备(Router-2)
[*Exit-Gateway-CE12808-1-10GE3/0/5] undo portswitch
[*Exit-Gateway-CE12808-1-10GE3/0/5] ip address 192.168.44.169 30
[*Exit-Gateway-CE12808-1-10GE3/0/5] set up-delay 180 //配置接口上报状态变化事件的延时时间
[*Exit-Gateway-CE12808-1-10GE3/0/5] quit
[*Exit-Gateway-CE12808-1] commit
```

在Exit-Gateway-CE12808-2配置与Router的互联端口。

```
[~Exit-Gateway-CE12808-2] interface 10ge 3/0/4 //连接外部PE设备(Router-1)
[~Exit-Gateway-CE12808-2-10GE3/0/4] undo portswitch
[*Exit-Gateway-CE12808-2-10GE3/0/4] ip address 192.168.44.173 30
[*Exit-Gateway-CE12808-2-10GE3/0/4] set up-delay 180
[*Exit-Gateway-CE12808-2-10GE3/0/4] quit
[*Exit-Gateway-CE12808-2] interface 10ge 3/0/5 //连接外部PE设备(Router-2)
[*Exit-Gateway-CE12808-2-10GE3/0/5] undo portswitch
[*Exit-Gateway-CE12808-2-10GE3/0/5] ip address 192.168.44.161 30
[*Exit-Gateway-CE12808-2-10GE3/0/5] set up-delay 180
[*Exit-Gateway-CE12808-2-10GE3/0/5] quit
[*Exit-Gateway-CE12808-2] commit
```

步骤2 配置网关之间的三层互联端口,以组成"口"字型和交叉字型结合的组网。

在Exit-Gateway-CE12808-1配置与Exit-Gateway-CE12808-2的互联端口。

```
[~Exit-Gateway-CE12808-1] interface 10ge 3/0/6 //连接Exit-Gateway-CE12808-2 [~Exit-Gateway-CE12808-1-10GE3/0/6] undo portswitch [*Exit-Gateway-CE12808-1-10GE3/0/6] ip address 192.168.44.165 30 [*Exit-Gateway-CE12808-1-10GE3/0/6] commit [~Exit-Gateway-CE12808-1-10GE3/0/6] quit
```

在Exit-Gateway-CE12808-2配置与Exit-Gateway-CE12808-1的互联端口。

[~Exit-Gateway-CE12808-2] interface 10ge 3/0/6 //连接Exit-Gateway-CE12808-1

[~Exit-Gateway-CE12808-2-10GE3/0/6] undo portswitch

[*Exit-Gateway-CE12808-2-10GE3/0/6] ip address 192.168.44.166 30

[*Exit-Gateway-CE12808-2-10GE3/0/6] commit

[~Exit-Gateway-CE12808-2-10GE3/0/6] quit

步骤3 网关设备上配置指向外部路由器的默认路由。

在Gateway-CE12808-1上配置默认路由。

[~Exit-Gateway-CE12808-1] ip route-static 0.0.0.0 0.0.0.0 192.168.44.158

[*Exit-Gateway-CE12808-1] ip route-static 0.0.0.0 0.0.0.0 192.168.44.170

[*Exit-Gateway-CE12808-1] ip route-static 0.0.0.0 0.0.0.0 192.168.44.166 preference 80 //指向Exit-

Gateway-CE12808-2,作为备份。

[*Exit-Gateway-CE12808-1] **commit**

[~Exit-Gateway-CE12808-1] quit

#在Gateway-CE12808-2上配置默认路由。

[~Exit-Gateway-CE12808-2] ip route-static 0.0.0.0 0.0.0.0 192.168.44.174

[*Exit-Gateway-CE12808-2] ip route-static 0.0.0.0 0.0.0.0 192.168.44.162

[*Exit-Gateway-CE12808-2] ip route-static 0.0.0.0 0.0.0.0 192.168.44.165 preference 80 //指向Exit-

Gateway-CE12808-1,作为备份。

[*Exit-Gateway-CE12808-2] commit

[~Exit-Gateway-CE12808-2] quit

----结束

3.2.1.6 配置增强功能

CE12800系列交换机需要配置步骤1/2/3中的哪些命令行取决于设备所选择的单板,具体可以参照下表(CE12800E系列交换机无需配置)。

表 3-2 CE12800 系列交换机优化配置

单板	配置命令
全部插入E系列接口板	 assign forward nvo3 service extend enable assign forward nvo3 acl extend enable (重启 生效) assign forward nvo3-gateway enhanced L3
全部插入F系列的FD/FDA类型接口板(仅允许插入F/G类型交换网板)	 assign forward nvo3 acl extend enable (重启生效) set serdes capability enhanced (重启生效) set forward capability enhanced (重启生效)
E系列接口板和F系列接口板 混插 (不推荐) (仅允许插入F/G类型交换 网板)	 assign forward nvo3 service extend enable assign forward nvo3 acl extend enable (重启 生效) assign forward nvo3-gateway enhanced L3 assign forward nvo3 f-linecard compatibility enable

步骤1 配置网关设备的NVO3的扩展功能。

当CE12800设备作为网关时,默认情况下,NVO3的扩展功能未使能。此时在设备上部署NVO3业务后,若再叠加其他需要使用ACL的业务(如MQC、简化ACL、流量监管、BD流量统计、DHCP等),会较大概率出现叠加失败的情况。

在部署NVO3的设备上可以使用以下方法以降低其他业务叠加失败的风险。

```
[~Exit-Gateway-CE12808-1] assign forward nvo3 service extend enable
[*Exit-Gateway-CE12808-1] commit
[~Exit-Gateway-CE12808-1] quit
<~Exit-Gateway-CE12808-1> save
<~Exit-Gateway-CE12808-1> reboot

[~Exit-Gateway-CE12808-2] assign forward nvo3 service extend enable
[*Exit-Gateway-CE12808-2] assign forward nvo3 service extend enable
[*Exit-Gateway-CE12808-2] commit
[~Exit-Gateway-CE12808-2] commit
[~Exit-Gateway-CE12808-2] quit
<~Exit-Gateway-CE12808-2> save
<~Exit-Gateway-CE12808-2> reboot
```

上述两条命令的使用说明如下。

在Exit-Gateway设备系统视图下,使用assign forward nvo3 service extend enable命令,使能NVO3业务扩展功能。使能此命令后,在非CE-L48GT-EA、CE-L48GT-EC、CE-L48GS-EA、CE-L48GS-EC、CE-L24XS-BA、CE-L24XS-EA、CE-L48XS-BA、CE-L48XS-EA和CE-L24LQ-EA单板上,能减少这种业务叠加失败的问题。

山 说明

此命令对CE-L48GT-EA、CE-L48GT-EC、CE-L48GS-EA、CE-L48GS-EC、CE-L24XS-BA、CE-L24XS-EA、CE-L48XS-BA、CE-L48XS-EA和CE-L24LQ-EA单板不生效。

使能此命令后,非上述单板上通过NVO3隧道传输的、长度在230~294字节的报文不能发送至上述单板。

配置此命令后,才能在除CE-L48GT-EA、CE-L48GT-EC、CE-L48GS-EA、CE-L48GS-EC、CE-L24XS-BA、CE-L24XS-EA、CE-L48XS-BA、CE-L48XS-EA和CE-L24LQ-EA之外的单板上使用路径探测功能。

● 在Exit-Gateway设备系统视图下,使用**assign forward nvo3 acl extend enable** 命令,使能NVO3的ACL扩展功能。

山 说明

此命令仅支持在Admin-VS中配置,配置后对所有VS生效。 执行此命令使能NVO3的ACL扩展功能后,需重启设备使配置生效。

步骤2 配置NVO3网关的增强模式。

CE12800缺省情况下,未配置NVO3网关的增强模式,设备采用的是环回模式,即经过NVO3封装后的报文需要在设备内部先环回后才能继续转发。此时,GW设备上的VXLAN三层转发封装/解封装的流量,超过线卡总转发性能的50%时,有较大可能出现丢包。目前,可以根据实际需要配置NVO3的网关增强模式,解决此问题。

在Exit-Gateway设备的系统视图下,使用**assign forward nvo3-gateway enhanced l3**命令行配置NVO3网关的增强模式。

[~Exit-Gateway-CE12808-1] assign forward nvo3-gateway enhanced l3

[~Exit-Gateway-CE12808-2] assign forward nvo3-gateway enhanced l3

3 VXLAN 部署建议

□ 说明

需要先执行**assign forward nvo3 service extend enable**命令,使能NVO3业务扩展功能。此时请确保设备上不存在CE-L48GT-EA、CE-L48GT-EC、CE-L48GS-EA、CE-L48GS-EC、CE-L24XS-BA、CE-L24XS-BA、CE-L48XS-BA、CE-L48XS-EA和CE-L24LQ-EA单板,或者这些单板上没有承载VXLAN相关业务。

若承载VXLAN业务的单板为CE-L24XS-EC、CE-L48XS-EC、CE-L24LQ-EC、CE-L48XT-EC、CE-L24LQ-EC1、CE-L08CC-EC、CE-L02LQ-EC、CE-L06LQ-EC时,仅支持通过查询ARP主机表进行VXLAN隧道封装,不支持通过查询最长匹配路由表进行VXLAN隧道封装。

服务器网卡主备方式接入的两台接入交换机需要组建为堆叠系统,或者M-LAG方式将服务器双归连接到接入交换机,此时接入交换机连接服务器的端口不要作为M-LAG的成员接口。

步骤3 配置CE12800系列的增强模式(CE12800E系列不支持配置本步骤命令)。

1. 配置单板的互通模式为增强模式。

[~Exit-Gateway-CE12808-1] **set forward capability enhanced** Warning: Current configuration should be committed and saved, and it will take effect after reboot. [Y/N]: **y**

□ 说明

增强模式是指:设备仅允许插入FD/FDA类型接口板,且仅允许插入F/G类型交换网板。

当单板互通模式为增强模式时,不可以插入EA/EC/ED/EF/EG/BA/CE-FWA/CE-IPSA类型接口板和A/B/C类型交换网板,否则单板会被下电。设备交换网板是A/B/C类型情况下,不能插入FD/FDA类接口板;设备插入了FD/FDA类接口板的情况下,不能插入A/B/C类型交换网板。

配置该命令后,需要保存配置并重启设备后配置才会生效(保存的配置文件要作为下次启动文件)。设备重启后需要再次执行命令save保存配置。

配置该命令后,建议同时配置命令set serdes capability enhanced,设置接口板与交换网板之间Serdes的速率模式为增强模式,这样可以达到最大转发性能。

该命令与以下命令互斥,配置本命令前,需要删除以下配置:

- assign forward nvo3-gateway enhanced { l2 | l3 }
- assign forward nvo3 eth-trunk hash disable

当设备上有EA/EC/ED/EF/EG/BA/CE-FWA/CE-IPSA类型接口板和FD/FDA类型接口板混插时,需要配置单板的互通模式为非增强模式,此时为了防止VXLAN流量成环,需要执行assign forward nvo3 f-linecard compatibility enable命令。

2. 配置Serdes速率模式为增强模式。

[~Exit-Gateway-CE12808-1] set serdes capability enhanced

Warning: Current configuration should be committed and saved, and it will take effect after reboot. [Y/N]: **y**

□ 说明

配置该命令后,需要重启设备配置才会生效。

设备交换网板是A/B/C类型情况下,不能插入FD/FDA类接口板;设备插入了FD/FDA类接口板的情况下,不能插入A/B/C类型交换网板。

步骤4 配置ARP广播报文抑制。

前面已经在网关工作组和Leaf工作组上执行命令peer { ipv4-address | group-name } advertise irb,配置BGP EVPN发布路由功能,以及执行arp collect host enable使能了BGP EVPN进行主机信息搜集的功能。为了实现ARP广播报文抑制的功能,还需要配置使能ARP广播报文抑制功能即可。

在所有Leaf工作组上使能ARP广播报文抑制功能(以Leaf堆叠工作组为例)。

[~Leaf-CE6851HI-1&CE6851HI-2] bridge-domain 10

[~Leaf-CE6851HI-1&CE6851HI-2-bd10] arp broadcast-suppress enable

[*Leaf-CE6851HI-1&CE6851HI-2-bd10] **commit**

[~Leaf-CE6851HI-1&CE6851HI-2-bd10] quit

步骤5 在所有Leaf工作组上配置BUM报文的流量抑制功能(以Leaf堆叠工作组为例)。

[~Leaf-CE6851HI-1&CE6851HI-2] interface 10ge 1/0/1
[~Leaf-CE6851HI-1&CE6851HI-2-10GE1/0/1] storm suppression unknown-unicast 1
[*Leaf-CE6851HI-1&CE6851HI-2-10GE1/0/1] storm suppression multicast packets 200
[*Leaf-CE6851HI-1&CE6851HI-2-10GE1/0/1] storm suppression broadcast packets 1000
[*Leaf-CE6851HI-1&CE6851HI-2-10GE1/0/1] commit
[~Leaf-CE6851HI-1&CE6851HI-2-10GE1/0/1] quit

[~Leaf-CE6851HI-1&CE6851HI-2-10GE1/0/2] interface 10ge 1/0/2
[~Leaf-CE6851HI-1&CE6851HI-2-10GE1/0/2] storm suppression unknown-unicast 1
[*Leaf-CE6851HI-1&CE6851HI-2-10GE1/0/2] storm suppression multicast packets 200
[*Leaf-CE6851HI-1&CE6851HI-2-10GE1/0/2] commit
[~Leaf-CE6851HI-1&CE6851HI-2-10GE1/0/2] commit
[~Leaf-CE6851HI-1&CE6851HI-2-10GE1/0/2] quit

步骤6 在所有Leaf上配置系统资源模式为大路由模式(以Leaf堆叠工作组为例)。

[~Leaf-CE6851HI-1&CE6851HI-2] system resource large-route

----结束

3.2.2 Spine/Leaf 二层架构的硬件分布式 VXLAN

□ 说明

本节只重点描述本方案与"GW/Spine/Leaf三层架构"方案中不同的配置内容,其余内容与"GW/Spine/Leaf三层架构"一致。

主要配置差异内容如下:

- Gateway与Leaf互联的接口配置
- Leaf上BGP路由的配置
- Gateway上BGP路由的配置
- Gateway和Leaf上的BGP EVPN功能配置

组网需求

Spine/Leaf二层架构的硬件分布式VXLAN组网图如图 1 Spine/Leaf二层架构的硬件分布式VXLAN组网图所示,

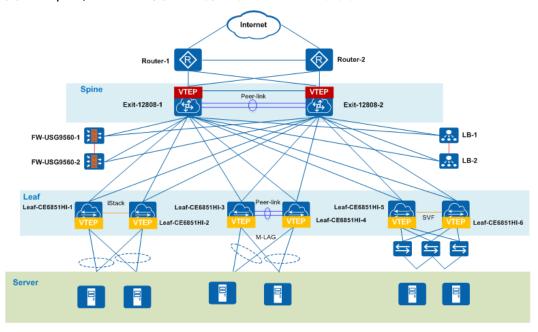


图 3-5 Spine/Leaf 二层架构的硬件分布式 VXLAN 组网图

- 服务器层:通过二层子接口方式接入VXLAN网络。
- Leaf(分布式网关,也称东西网关):采用堆叠方式、M-LAG或SVF方式连接服务器。Leaf节点和Spine节点通过三层互联。Leaf的堆叠工作组、M-LAG工作组或者SVF工作组作为VTEP实现服务器流量接入VXLAN网络。
- Spine/GW:两台GW之间组成M-LAG,即双活网关。GW与Leaf之间三层互联。GW对外连接外部路由器Router-1和Router-2。
- FW: 两台FW分别旁挂在GW, FW之间配置为主备镜像模式工作。
- LB: 由厂商部署。

端口连线规划

Leaf-CE6851HI-1和Leaf-CE6851HI-2部署堆叠场景,端口规划参见下表。

规划描述	本端设备名 称	物理端口	对端设备名称	物理端口
规划Leaf间互连端口作 为堆叠口,堆叠链路至	Leaf- CE6851HI-1	40GE1/0/1~ 2	Leaf- CE6851HI-2	40GE1/0/1 ~2
少2条	Leaf- CE6851HI-2	40GE1/0/1~ 2	Leaf- CE6851HI-1	40GE1/0/1 ~2
规划Leaf与Spine互连 端口	Leaf- CE6851HI-1	40GE1/0/3	Exit- Gateway- CE12808-1	40GE1/0/0
	Leaf- CE6851HI-1	40GE1/0/4	Exit- Gateway- CE12808-2	40GE1/0/1

规划描述	本端设备名 称	物理端口	对端设备名称	物理端口
	Leaf- CE6851HI-2	40GE1/0/3	Exit- Gateway- CE12808-2	40GE1/0/0
	Leaf- CE6851HI-2	40GE1/0/4	Exit- Gateway- CE12808-1	40GE1/0/1
规划Leaf与服务器的互连端口,服务器双网卡	Leaf- CE6851HI-1	10GE1/0/1~ 2	服务器	Eth0
以负载分担方式上连, 保证链路可靠性 	Leaf- CE6851HI-2	10GE1/0/1~ 2	服务器	Eth1

Leaf-CE6851HI-3和Leaf-CE6851HI-4部署M-LAG,端口规划参见下表。

规划描述	本端设备名 称	物理端口	对端设备名称	物理端口
规划M-LAG的peer-link端口,用于承载	Leaf- CE6851HI-3	40GE1/0/1~2	Leaf- CE6851HI-4	40GE1/0/1 ~2
M-LAG的协议报文和 故障场景下的数据报 文,为保证可靠性, 建议至少2条peer-link 链路	Leaf- CE6851HI-4	40GE1/0/1~2	Leaf- CE6851HI-3	40GE1/0/1 ~2
规划Leaf连接Spine互 连端口	Leaf- CE6851HI-3	40GE1/0/3	Exit- Gateway- CE12808-1	40GE1/0/2
	Leaf- CE6851HI-3	40GE1/0/4	Exit- Gateway- CE12808-2	40GE1/0/3
	Leaf- CE6851HI-4	40GE1/0/3	Exit- Gateway- CE12808-2	40GE1/0/2
	Leaf- CE6851HI-4	40GE1/0/4	Exit- Gateway- CE12808-1	40GE1/0/3
规划Leaf与服务器的 互连端口,服务器双 网卡以负载分担方式 上连,保证链路可靠 性	Leaf- CE6851HI-3	10GE1/0/1	服务器	Eth0
	Leaf- CE6851HI-4	10GE1/0/1	服务器	Eth1

Leaf-CE6851HI-5和Leaf-CE6851HI-6部署SVF场景,下连三台CE5810交换机,端口规划参见下表。

规划描述	本端设备名称	物理端口	对端设备名称	物理端口
规划Leaf间互连端 口作为堆叠口,堆	Leaf- CE6851HI-5	40GE1/0/1~ 2	Leaf- CE6851HI-6	40GE1/0/1 ~2
叠链路至少2条	Leaf- CE6851HI-6	40GE1/0/1~ 2	Leaf- CE6851HI-5	40GE1/0/1 ~2
规划Leaf与Spine互 连端口	Leaf- CE6851HI-5	40GE1/0/3	Exit- Gateway- CE12808-1	40GE1/0/6
	Leaf- CE6851HI-5	40GE1/0/4	Exit- Gateway- CE12808-2	40GE1/0/7
	Leaf- CE6851HI-6	40GE1/0/3	Exit- Gateway- CE12808-1	40GE1/0/6
	Leaf- CE6851HI-6	40GE1/0/4	Exit- Gateway- CE12808-2	40GE1/0/7
规划SVF父节点与叶	Leaf- CE6851HI-5	10GE1/0/1	CE5800-1	GE1/0/1
子节点互联端口 		10GE1/0/2	CE5800-2	GE1/0/1
		10GE1/0/3	CE5800-3	GE1/0/1
	Leaf-	10GE1/0/1	CE5800-1	GE1/0/2
	CE6851HI-6	10GE1/0/2	CE5800-2	GE1/0/2
		10GE1/0/3	CE5800-3	GE1/0/2
规划叶子交换机与	CE5810-1	GE1/0/1	服务器3	Eth0
服务器的互连端 口,服务器双网卡 以负载分担方式上 连,保证链路可靠 性	CE5810-2	GE1/0/1	服务器3	Eth1

Exit-Gateway-CE12808-1和Exit-Gateway-CE12808-2端口规划参见下表。

规划描述	本端设备名 称	物理端口	对端设备名称	物理端口
规划Exit-Gateway与 Leaf的互连端口	Exit- Gateway- CE12808-1	40GE1/0/0~3 40GE1/0/6~7	Leaf- CE6851HI-1 Leaf- CE6851HI-2 Leaf- CE6851HI-3 Leaf- CE6851HI-4 Leaf- CE6851HI-5 Leaf- CE6851HI-5	40GE1/0/ 3~4
	Exit- Gateway- CE12808-2	40GE1/0/0~3 40GE1/0/6~7	Leaf- CE6851HI-1 Leaf- CE6851HI-2 Leaf- CE6851HI-3 Leaf- CE6851HI-4 Leaf- CE6851HI-5 Leaf- CE6851HI-5	40GE1/0/ 3~4
规划M-LAG的peer- link端口,用于承载M- LAG的协议报文和故障 场景下的数据报文,为 保证可靠性,建议至少	Exit- Gateway- CE12808-1	40GE1/0/23 40GE2/0/23	Exit- Gateway- CE12808-2	40GE1/0/ 23 40GE2/0/ 23
2条peer-link链路,且 跨板部署	Exit- Gateway- CE12808-2	40GE1/0/23 40GE2/0/23	Exit- Gateway- CE12808-1	40GE1/0/ 23 40GE2/0/ 23
规划Exit-Gateway与防 火墙的互连端口 说明 两台主备镜像模式防火 墙必须使用同样的接口	Exit- Gateway- CE12808-1	10GE3/0/0~1	FW- USG9560-1 FW- USG9560-2	GE1/0/1
连接到同一个GW设备。 例如FW-1使用GE 1/0/1 接口连接GW-1,那么 FW-2也需要使用GE 1/0/1接口连接GW-1。	Exit- Gateway- CE12808-1	10GE3/0/2~3	FW- USG9560-1 FW- USG9560-2	GE1/0/3

规划描述	本端设备名 称	物理端口	对端设备名称	物理端口
	Exit- Gateway- CE12808-2	10GE3/0/0~1	FW- USG9560-1 FW- USG9560-2	GE1/0/2
	Exit- Gateway- CE12808-2	10GE3/0/2~3	FW- USG9560-1 FW- USG9560-2	GE1/0/4
规划Exit-Gateway与出口路由器的互连端口	Exit- Gateway- CE12808-1	10GE3/0/4	Router-1	GE1/0/0
	Exit- Gateway- CE12808-1	10GE3/0/5	Router-2	GE1/0/0
	Exit- Gateway- CE12808-2	10GE3/0/4	Router-1	GE1/0/1
	Exit- Gateway- CE12808-2	10GE3/0/5	Router-2	GE1/0/1
规划Exit-Gateway之间 的三层互联接口,用于 和出口路由器组成 "口"字型和交叉联合 出口网络	Exit- Gateway- CE12808-1	10GE3/0/6	Exit- Gateway- CE12808-2	10GE3/0/ 6

VLAN 规划

VLAN规划参见下表。

规划描述	规划建议	本示例使用ID
与防火墙 互连 VLAN	三层互连建议使用三层主接口,与FW因涉 及多个网段对接,需要规划VLAN,达到共 用物理链路目的	11 12(创建VLANIF 12绑定 VPN)
租户 VLAN	租户/业务使用VXLAN较多,需要考虑后续 的业务扩容	10

□ 说明

对于CE6855HI和CE7855EI,在切换三层接口之前,要先配置**vlan reserved for maininterface** *startvlanid* **to** *endvlanid*,配置三层主接口专用的保留VLAN。

BD、VNI 规划

BD、VNI的规划参见下表。

规划描述	规划建议	本示例使用ID
BD	规划数量与VLAN相同,建议BD的ID与 VLAN的ID相同	本例统一使用BD 10
VNI	二层VNI规划数量与BD相同,建议VNI的ID规划为"BD+10000"(BD与VNI——对应) 三层VNI数量和L3VPN相同。	本例统一使用二层VNI 10010,三层VNI为10

RD 和 RT 值规划

RD的规划参见下表。

规划描述	规划建议	本示例使用ID
CE6851HI- 1、 CE6851HI- 2的RD值	分布式网关需要配置EVPN实例和VPN实 例的RD值,且需要保持全局唯一。	11: 1 12: 1 (VPN实例)
CE6851HI- 3的RD值	分布式网关需要配置EVPN实例和VPN实 例的RD值,且需要保持全局唯一。	13: 1 12: 2(VPN实例)
CE6851HI- 4的RD值	分布式网关需要配置EVPN实例和VPN实 例的RD值,且需要保持全局唯一。	14: 1 12: 3(VPN实例)
CE6851HI- 5、 CE6851HI- 6的RD值	分布式网关需要配置EVPN实例和VPN实 例的RD值,且需要保持全局唯一。	15: 1 12: 4(VPN实例)
Exit- Gateway- CE12808-1 的RD值	分布式网关需要配置EVPN实例和VPN实 例的RD值,且需要保持全局唯一。	16: 1
Exit- Gateway- CE12808-2 的RD值	分布式网关需要配置EVPN实例和VPN实 例的RD值,且需要保持全局唯一。	17: 1

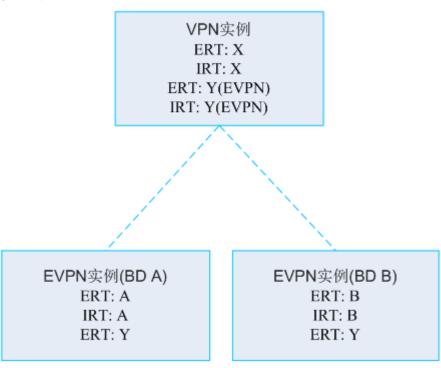
RT的规划参见下表。

设备	规划建议	本示例使用ID
CE6851HI- 1、 CE6851HI- 2的RT值	分布式网关需要配置EVPN实例和VPN实 例的RT值。	1: 1 11: 1 (EVPN实例出方向 RT值,用于与VPN实例交 叉) 11: 1 (VPN实例的RT 值,用于与EVPN实例交 叉)
CE6851HI- 3的RT值	分布式网关需要配置EVPN实例和VPN实 例的RT值。	1: 1 11: 1 (EVPN实例出方向 RT值,用于与VPN实例交 叉) 11: 1 (VPN实例的RT 值,用于与EVPN实例交 叉)
CE6851HI- 4的RT值	分布式网关需要配置EVPN实例和VPN实 例的RT值。	1: 1 11: 1 (EVPN实例出方向 RT值,用于与VPN实例交 叉) 11: 1 (VPN实例的RT 值,用于与EVPN实例交 叉)
CE6851HI- 5、 CE6851HI- 6的RT值	分布式网关需要配置EVPN实例和VPN实 例的RT值。	1: 1 11: 1 (EVPN实例出方向 RT值,用于与VPN实例交 叉) 11: 1 (VPN实例的RT 值,用于与EVPN实例交 叉)
Exit- Gateway- CE12808-1 的RT值	分布式网关需要配置EVPN实例和VPN实例的RT值。	1: 1 11: 1 (EVPN实例出方向 RT值,用于与VPN实例交 叉) 11: 1 (VPN实例的RT 值,用于与EVPN实例交 叉)
Exit- Gateway- CE12808-2 的RT值	分布式网关需要配置EVPN实例和VPN实 例的RT值。	1: 1 11: 1 (EVPN实例出方向 RT值,用于与VPN实例交 叉) 11: 1 (VPN实例的RT 值,用于与EVPN实例交 叉)

VPN实例和EVPN实例的RT值配置原则如图 2 配置RT值所示,其中:

- 在VPN实例中,除了配置本地VPN实例的ERT X和IRT X外,还需要配置带EVPN参数的ERT Y和IRT Y,用于和EVPN实例交叉,生成主机路由。
- 在EVPN实例中,除了针对不同的BD配置不同的ERT A、ERT B和IRT A、IRT B 外,还必须配置用来和VPN实例交叉的ERT Y。一般情况下不需要配置IRT Y,否则会导致不同BD下的EVPN实例相互扩散MAC地址。

图 3-6 配置 RT 值



IP 地址规划

网元接口地址规划参见下表,涉及互连地址、VTEP地址、BGP Route-ID、BGP建邻居 loopback地址、M-LAG心跳地址、业务网段地址。

端口互连IP地址规划参见下表。

规划描述	本端设备	对端设备	IP/MASK
规划CE6851HI-1、 CE6851HI-2与出口 网关互连端口地址		Exit-Gateway- CE12808-1	192.168.40.156/30 192.168.40.168/30
		Exit-Gateway- CE12808-2	192.168.40.164/30 192.168.40.160/30
规划CE6851HI-3、 CE6851HI-4与出口 网关互连端口地址	Leaf- CE6851HI-3& CE6851HI-4	Exit-Gateway- CE12808-1	192.168.41.156/30 192.168.41.168/30

规划描述	本端设备	对端设备	IP/MASK
		Exit-Gateway- CE12808-2	192.168.41.164/30
		CL12000-2	192.168.41.160/30
规划CE6851HI-5、 CE6851HI-6与出口	Leaf- CE6851HI-5& CE6851HI-6	Exit-Gateway- CE12808-1	192.168.46.156/30
网关互连端口地址			192.168.46.168/30
7 35 (===================================		Exit-Gateway- CE12808-2	192.168.46.164/30
			192.168.46.160/30
规划Exit-Gateway- CE12808-1与 Router-1互连端口 地址	Exit-Gateway- CE12808-1	Router-1	192.168.44.156/30
规划Exit-Gateway- CE12808-2与 Router-2互连端口 地址	Exit-Gateway- CE12808-2	Router-2	192.168.44.160/30
规划Exit-Gateway- CE12808-1和Exit- Gateway- CE12808-2互联	Exit-Gateway- CE12808-1	Exit-Gateway- CE12808-2	192.168.44.164/30
规划Exit-Gateway- CE12808-1&2与 FW互连端口地址	Exit-Gateway- CE12808-1 Exit-Gateway- CE12808-2	FW-USG9560-1	192.168.45.152/29
		FW-USG9560-2	Virtual IP: 192.168.45.153
			192.168.45.160/29
			Virtual IP: 192.168.45.161

Loopback地址规划参见下表。

规划描述	设备	IP/MASK
规划Loopback0地址,作 为VTEP地址	Leaf-CE6851HI-1& CE6851HI-2	11.11.11.11/32
说明 Leaf-CE6851HI-3和Leaf- CE6851HI-4组成M-LAG用 于服务器双活接入,因此它们的VTEP IP地址相同。	Leaf-CE6851HI-3	11.11.11.12/32
	Leaf-CE6851HI-4	同Leaf-CE6851HI-3: 11.11.11.12/32
Exit-Gateway-CE12808-1和 Gateway-CE12808-2组成 M-LAG用于防火墙双活接 入,因此它们的VTEP IP地址 相同。	Leaf-CE6851HI-5& CE6851HI-6	11.11.11.17/32
	Exit-Gateway-CE12808-1	11.11.11.16/32

规划描述	设备	IP/MASK
	Exit-Gateway-CE12808-2	同Exit-Gateway- CE12808-1: 11.11.11.16/32
规划Loopback1地址,作 为M-LAG心跳检测地址	Leaf-CE6851HI-3	13.13.13.13/32
	Leaf-CE6851HI-4	14.14.14.14/32
	Exit-Gateway-CE12808-1	18.18.18.18/32
	Exit-Gateway-CE12808-2	19.19.19.19/32
规划Loopback2地址,作 为对接Router的EBGP邻居 地址	Exit-Gateway-CE12808-1	21.21.21.21/32
	Exit-Gateway-CE12808-2	22.22.22.22/32

业务地址规划参见下表。

租户	租户IP/MASK	网关VBDIF	VRF
服务器1	192.168.10.2/24	192.168.10.1/24	VPN 1
服务器2	192.168.10.3/24	192.168.10.1/24	VPN 1
服务器3	192.168.10.4/24	192.168.10.1/24	VPN 1

路由规划

Underlay网络使用的路由协议常见为EBGP和OSPF两种,BGP从多方面保证了网络的安全性、灵活性、稳定性、可靠性和高效性:

- BGP采用认证和GTSM的方式,保证了网络的安全性。
- BGP提供了丰富的路由策略,能够灵活的进行路由选路。
- BGP提供了路由聚合和路由衰减功能用于防止路由振荡,有效提高了网络的稳定性。
- BGP使用TCP作为其传输层协议(端口号为179),并支持BGP与BFD联动、BGP Auto FRR和BGP GR和NSR,提高了网络的可靠性。

从网络演进方面考虑,EBGP适用于大型网络,而OSPF适用于中小型网络,本文档使用EBGP作为示例,其中Gateway对接Router的IGP使用OSPF路由作为示例(此处使用ISIS或者IBGP均可以)。

路由相关的规划参见下表。

网元	AS域号	Router ID
Leaf- CE6851HI-1&CE6851-2	65021	Loopback0地址

网元	AS域号	Router ID
Leaf-CE6851HI-3	65022	Loopback1地址
Leaf-CE6851HI-4	65022	Loopback1地址
Leaf- CE6851HI-5&CE6851HI-6	65024	Loopback0地址
Exit-Gateway-CE12808-1	65000	Loopback1地址
Exit-Gateway-CE12808-2	65000	Loopback1地址
Router-1	65047	Loopback0地址
Router-2	65048	Loopback0地址

3.2.2.1 配置 Gateway 对接 Leaf 的接口地址

在本方案中,Spine节点和Gateway融合部署,因此Gateway直接与Leaf互联。

步骤1 在Exit-Gateway-CE12808-1上配置接口IP地址。

```
[~Huawei] sysname Exit-Gateway-CE12808-1
[*Huawei] commit
[~Exit-Gateway-CE12808-1] interface 40ge 1/0/0
[*Exit-Gateway-CE12808-1-40GE1/0/0] description "to-Leaf-CE6851HI-1&CE6851HI-2"
[~Exit-Gateway-CE12808-1-40GE1/0/0] undo portswitch
[*Exit-Gateway-CE12808-1-40GE1/0/0] ip address 192.168.40.158 30
[*Exit-Gateway-CE12808-1-40GE1/0/0] commit
[~Exit-Gateway-CE12808-1-40GE1/0/0] quit
[~Exit-Gateway-CE12808-1] interface 40ge 1/0/1
[*Exit-Gateway-CE12808-1-40GE1/0/1] description "to-Leaf-CE6851HI-1&CE6851HI-2"
[*Exit-Gateway-CE12808-1-40GE1/0/1] undo portswitch
[*Exit-Gateway-CE12808-1-40GE1/0/1] ip address 192.168.40.170 30
[*Exit-Gateway-CE12808-1-40GE1/0/1] commit
[~Exit-Gateway-CE12808-1-40GE1/0/1] quit
[~Exit-Gateway-CE12808-1] interface 40ge 1/0/2
[~Exit-Gateway-CE12808-1-40GE1/0/2] description "to-Leaf-CE6851-3-40GE1/0/3"
[*Exit-Gateway-CE12808-1-40GE1/0/2] undo portswitch
[*Exit-Gateway-CE12808-1-40GE1/0/2] ip address 192.168.41.158 30
[*Exit-Gateway-CE12808-1-40GE1/0/2] commit
[~Exit-Gateway-CE12808-1-40GE1/0/2] quit
[~Exit-Gateway-CE12808-1] interface 40ge 1/0/3
[*Exit-Gateway-CE12808-1-40GE1/0/3] description "to-Leaf-CE6851-4-40GE1/0/4"
[*Exit-Gateway-CE12808-1-40GE1/0/3] undo portswitch
[*Exit-Gateway-CE12808-1-40GE1/0/3] ip address 192.168.41.170 30
[*Exit-Gateway-CE12808-1-40GE1/0/3] commit
[~Exit-Gateway-CE12808-1-40GE1/0/3] quit
[~Exit-Gateway-CE12808-1] interface 40ge 1/0/6
[~Exit-Gateway-CE12808-1-40GE1/0/6] undo portswitch
[*Exit-Gateway-CE12808-1-40GE1/0/6] ip address 192.168.46.158 30
[*Exit-Gateway-CE12808-1-40GE1/0/6] commit
[~Exit-Gateway-CE12808-1-40GE1/0/6] quit
[~Exit-Gateway-CE12808-1] interface 40ge 1/0/7
[~Exit-Gateway-CE12808-1-40GE1/0/7] undo portswitch
[*Exit-Gateway-CE12808-1-40GE1/0/7] ip address 192.168.46.170 30
[*Exit-Gateway-CE12808-1-40GE1/0/7] commit
[~Exit-Gateway-CE12808-1-40GE1/0/7] quit
```

步骤2 在Exit-Gateway-CE12808-2节点上配置接口IP地址。

[~Huawei] sysname Exit-Gateway-CE12808-2 [*Huawei] commit

```
[~Exit-Gateway-CE12808-2] interface 40ge 1/0/0
[~Exit-Gateway-CE12808-2-40GE1/0/0] undo portswitch
[*Exit-Gateway-CE12808-2-40GE1/0/0] ip address 192.168.40.162 30
[*Exit-Gateway-CE12808-2-40GE1/0/0] commit
[~Exit-Gateway-CE12808-2-40GE1/0/0] quit
[~Exit-Gateway-CE12808-2] interface 40ge 1/0/1
[~Exit-Gateway-CE12808-2-40GE1/0/1] undo portswitch
[*Exit-Gateway-CE12808-2-40GE1/0/1] ip address 192.168.40.166 30
[*Exit-Gateway-CE12808-2-40GE1/0/1] commit
[~Exit-Gateway-CE12808-2-40GE1/0/1] quit
[~Exit-Gateway-CE12808-2] interface 40ge 1/0/2
[~Exit-Gateway-CE12808-2-40GE1/0/2] description "to-Leaf-CE6851-4-40GE1/0/3"
[*Exit-Gateway-CE12808-2-40GE1/0/2] undo portswitch
[*Exit-Gateway-CE12808-2-40GE1/0/2] ip address 192.168.41.162 30
[*Exit-Gateway-CE12808-2-40GE1/0/2] commit
[~Exit-Gateway-CE12808-2-40GE1/0/2] quit
[~Exit-Gateway-CE12808-2] interface 40ge 1/0/3
[~Exit-Gateway-CE12808-2-40GE1/0/3] description "to-Leaf-CE6851-3-40GE1/0/4"
[*Exit-Gateway-CE12808-2-40GE1/0/3] undo portswitch
[*Exit-Gateway-CE12808-2-40GE1/0/3] ip address 192.168.41.166 30
[*Exit-Gateway-CE12808-2-40GE1/0/3] commit
[~Exit-Gateway-CE12808-2-40GE1/0/3] quit
[~Exit-Gateway-CE12808-2] interface 40ge 1/0/6
[~Exit-Gateway-CE12808-2-40GE1/0/6] undo portswitch
[*Exit-Gateway-CE12808-2-40GE1/0/6] ip address 192.168.46.166 30
[*Exit-Gateway-CE12808-2-40GE1/0/6] commit
[~Exit-Gateway-CE12808-2-40GE1/0/6] quit
[~Exit-Gateway-CE12808-2] interface 40ge 1/0/7
[~Exit-Gateway-CE12808-2-40GE1/0/7] undo portswitch
[*Exit-Gateway-CE12808-2-40GE1/0/7] ip address 192.168.46.162 30
[*Exit-Gateway-CE12808-2-40GE1/0/7] commit
[~Exit-Gateway-CE12808-2-40GE1/0/7] quit
```

----结束

3.2.2.2 配置 Leaf 上的路由

在本方案中,Spine节点和Gateway融合部署,因此Leaf上的BGP路由配置与三层架构组网所有不同。

步骤1 在堆叠组上配置BGP路由,对接Gateway。

```
[~Leaf-CE6851HI-1&CE6851HI-2] bgp 65021
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] router-id 11.11.11.11
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] timer keepalive 10 hold 30
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] group Gateway-CE12808 external
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] peer Gateway-CE12808 as-number 65000
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] peer 192.168.40.158 as-number 65000
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] peer 192.168.40.158 group Gateway-CE12808
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] peer 192.168.40.170 as-number 65000
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] peer 192.168.40.170 group Gateway-CE12808 [*Leaf-CE6851HI-1&CE6851HI-2-bgp] peer 192.168.40.166 as-number 65000
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] peer 192.168.40.166 group Gateway-CE12808
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] peer 192.168.40.162 as-number 65000
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] peer 192.168.40.162 group Gateway-CE12808
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] ipv4-family unicast
[*Leaf-CE6851HI-1&CE6851HI-2-bgp-af-ipv4] preference 20 200 10
[*Leaf-CE6851HI-1&CE6851HI-2-bgp-af-ipv4] network 11.11.11.11 255.255.255.255
[*Leaf-CE6851HI-1&CE6851HI-2-bgp-af-ipv4] maximum load-balancing 32
[*Leaf-CE6851HI-1&CE6851HI-2-bgp-af-ipv4] quit
[*Leaf-CE6851HI-1&CE6851HI-2-bgp] quit
[*Leaf-CE6851HI-1&CE6851HI-2] commit
```

步骤2 在Leaf-CE6851HI-3上配置BGP路由,对接Gateway。

```
[~Leaf-CE6851HI-3] bgp 65022
[*Leaf-CE6851HI-3-bgp] router-id 13.13.13.13
[*Leaf-CE6851HI-3-bgp] timer keepalive 10 hold 30
[*Leaf-CE6851HI-3-bgp] group Gateway-CE12808 external
```

```
[*Leaf-CE6851HI-3-bgp] peer Gateway-CE12808 as-number 65000
[*Leaf-CE6851HI-3-bgp] peer 192.168.41.158 as-number 65000
[*Leaf-CE6851HI-3-bgp] peer 192.168.41.158 group Gateway-CE12808
[*Leaf-CE6851HI-3-bgp] peer 192.168.41.166 as-number 65000
[*Leaf-CE6851HI-3-bgp] peer 192.168.41.166 group Gateway-CE12808
[*Leaf-CE6851HI-3-bgp] ipv4-family unicast
[*Leaf-CE6851HI-3-bgp-af-ipv4] preference 20 200 10
[*Leaf-CE6851HI-3-bgp-af-ipv4] network 11.11.11.12 255.255.255.255
[*Leaf-CE6851HI-3-bgp-af-ipv4] network 10.10.00.0 255.255.255.0
[*Leaf-CE6851HI-3-bgp-af-ipv4] network 13.13.13.13 255.255.255.255
[*Leaf-CE6851HI-3-bgp-af-ipv4] maximum load-balancing 32
[*Leaf-CE6851HI-3-bgp-af-ipv4] quit
[*Leaf-CE6851HI-3-bgp] quit
[*Leaf-CE6851HI-3-bgp] quit
```

步骤3 在Leaf-CE6851HI-4上配置BGP路由,对接Gateway。

```
[~Leaf-CE6851HI-4] bgp 65022
[*Leaf-CE6851HI-4-bgp] router-id 14.14.14.14
[*Leaf-CE6851HI-4-bgp] timer keepalive 10 hold 30
[*Leaf-CE6851HI-4-bgp] group Gateway-CE12808 external
[*Leaf-CE6851HI-4-bgp] peer Gateway-CE12808 as-number 65000
[*Leaf-CE6851HI-4-bgp] peer 192.168.41.170 as-number 65000
[*Leaf-CE6851HI-4-bgp] peer 192.168.41.170 group Gateway-CE12808
[*Leaf-CE6851HI-4-bgp] peer 192.168.41.162 as-number 65000
[*Leaf-CE6851HI-4-bgp] peer 192.168.41.162 group Gateway-CE12808
[*Leaf-CE6851HI-4-bgp] ipv4-family unicast
[*Leaf-CE6851HI-4-bgp-af-ipv4] preference 20 200 10
[*Leaf-CE6851HI-4-bgp-af-ipv4] network 11.11.11.12 255.255.255.255
[*Leaf-CE6851HI-4-bgp-af-ipv4] network 10.10.100.0 255.255.255.0
[*Leaf-CE6851HI-4-bgp-af-ipv4] network 14.14.14.14 255.255.255.255
[*Leaf-CE6851HI-4-bgp-af-ipv4] maximum load-balancing 32
[*Leaf-CE6851HI-4-bgp-af-ipv4] quit
[*Leaf-CE6851HI-4-bgp] quit
[*Leaf-CE6851HI-4] commit
```

步骤4 在SVF组上配置BGP路由,对接Gateway。

```
[~Leaf-CE6851HI-5&CE6851HI-6] bgp 65024
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] router-id 11.11.11.17
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] timer keepalive 10 hold 30
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] group Gateway-CE12808 external
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] peer Gateway-CE12808 as-number 65000
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] peer 192.168.46.158 as-number 65000
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] peer 192.168.46.158 group_Gateway-CE12808
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] peer 192.168.46.170 as-number 65000
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] peer 192.168.46.170 group Gateway-CE12808
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] peer 192.168.46.166 as-number 65000
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] peer 192.168.46.166 group Gateway-CE12808
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] peer 192.168.46.162 as-number 65000
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] peer 192.168.46.162 group Gateway-CE12808
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] ipv4-family unicast
[*Leaf-CE6851HI-5&CE6851HI-6-bgp-af-ipv4] preference 20 200 10
[*Leaf-CE6851HI-5&CE6851HI-6-bgp-af-ipv4] network 11.11.11.17 255.255.255.255
[*Leaf-CE6851HI-5&CE6851HI-6-bgp-af-ipv4] maximum load-balancing 32
[*Leaf-CE6851HI-5&CE6851HI-6-bgp-af-ipv4] quit
[*Leaf-CE6851HI-5&CE6851HI-6-bgp] quit
[*Leaf-CE6851HI-5&CE6851HI-6] commit
```

----结束

3.2.2.3 配置网关上的路由

在本方案中,Spine节点和Gateway融合部署,因此Gateway上的BGP路由配置与三层架构组网所有不同。

步骤1 在Exit-Gateway-CE12808-1上配置BGP路由。

```
[~Exit-Gateway-CE12808-1] bgp 65000
[*Exit-Gateway-CE12808-1-bgp] router-id 18.18.18
```

3 VXLAN 部署建议

```
[*Exit-Gateway-CE12808-1-bgp] timer keepalive 10 hold 30
[*Exit-Gateway-CE12808-1-bgp] group Router-1 external
                                                            //对接出口路由器
[*Exit-Gateway-CE12808-1-bgp] peer Router-1 as-number 65047
[*Exit-Gateway-CE12808-1-bgp] peer Router-1 ebgp-max-hop 10
[*Exit-Gateway-CE12808-1-bgp] peer Router-1 connect-interface loopback 2
[*Exit-Gateway-CE12808-1-bgp] peer 21.21.21.22 as-number 65047
[*Exit-Gateway-CE12808-1-bgp] peer 21.21.21.22 group Router-1
[*Exit-Gateway-CE12808-1-bgp] group Leaf-CE6851HI-1&CE6851HI-2 external //对接Leaf
[*Exit-Gateway-CE12808-1-bgp] peer Leaf-CE6851HI-1&CE6851HI-2 as-number 65021
[*Exit-Gateway-CE12808-1-bgp] peer 192.168.40.157 as-number 65021
[*Exit-Gateway-CE12808-1-bgp] peer 192.168.40.157 group Leaf-CE6851HI-1&CE6851HI-2
[*Exit-Gateway-CE12808-1-bgp] peer 192.168.40.169 as-number 65021
[*Exit-Gateway-CE12808-1-bgp] peer 192.168.40.169 group Leaf-CE6851HI-1&CE6851HI-2
[*Exit-Gateway-CE12808-1-bgp] group Leaf-CE6851HI-3&4 external
[*Exit-Gateway-CE12808-1-bgp] peer Leaf-CE6851HI-3&4 as-number 65022
[*Exit-Gateway-CE12808-1-bgp] peer 192.168.41.157 as-number 65022
[*Exit-Gateway-CE12808-1-bgp] peer 192.168.41.157 group Leaf-CE6851HI-3&4
[*Exit-Gateway-CE12808-1-bgp] peer 192.168.41.169 as-number 65022
[*Exit-Gateway-CE12808-1-bgp] peer 192.168.41.169 group Leaf-CE6851HI-3&4
[*Exit-Gateway-CE12808-1-bgp] group Leaf-CE6851HI-5&CE6851HI-6 external
[*Exit-Gateway-CE12808-1-bgp] peer Leaf-CE6851HI-5&CE6851HI-6 as-number 65024
[*Exit-Gateway-CE12808-1-bgp] peer 192.168.46.157 as-number 65024
[*Exit-Gateway-CE12808-1-bgp] peer 192.168.46.157 group Leaf-CE6851HI-5&CE6851HI-6
[*Exit-Gateway-CE12808-1-bgp] peer 192.168.46.169 as-number 65024
[*Exit-Gateway-CE12808-1-bgp] peer 192.168.46.169 group Leaf-CE6851HI-5&CE6851HI-6
[*Exit-Gateway-CE12808-1-bgp] ipv4-family unicast
[*Exit-Gateway-CE12808-1-bgp-af-ipv4] preference 20 200 10
[*Exit-Gateway-CE12808-1-bgp-af-ipv4] network 11.11.11.16 255.255.255.255
[*Exit-Gateway-CE12808-1-bgp-af-ipv4] network 18.18.18.18 255.255.255.255
[*Exit-Gateway-CE12808-1-bgp-af-ipv4] network 21.21.21.21 255.255.255.255
[*Exit-Gateway-CE12808-1-bgp-af-ipv4] network 192.168.44.156 255.255.255.252
[*Exit-Gateway-CE12808-1-bgp-af-ipv4] network 192.168.45.152 255.255.255.248
[*Exit-Gateway-CE12808-1-bgp-af-ipv4] maximum load-balancing 32
[*Exit-Gateway-CE12808-1-bgp-af-ipv4] quit
[*Exit-Gateway-CE12808-1-bgp] quit
[*Exit-Gateway-CE12808-1] commit
```

步骤2 在Exit-Gateway-CE12808-2上配置BGP路由。

```
[~Exit-Gateway-CE12808-2] bgp 65000
[*Exit-Gateway-CE12808-2-bgp] router-id 19.19.19.19
[*Exit-Gateway-CE12808-2-bgp] timer keepalive 10 hold 30
[*Exit-Gateway-CE12808-2-bgp] group Router-2 external //对接出口路由器
[*Exit-Gateway-CE12808-2-bgp] peer Router-2 as-number 65048
[*Exit-Gateway-CE12808-2-bgp] peer Router-2 ebgp-max-hop 10
[*Exit-Gateway-CE12808-2-bgp] peer Router-2 connect-interface loopback 2
[*Exit-Gateway-CE12808-2-bgp] peer 22.22.23 as-number 65048
[*Exit-Gateway-CE12808-2-bgp] peer 22.22.23 group Router-2
[*Exit-Gateway-CE12808-2-bgp] group Leaf-CE6851HI-1&CE6851HI-2 external
[*Exit-Gateway-CE12808-2-bgp] peer Leaf-CE6851HI-1&CE6851HI-2 as-number 65021
[*Exit-Gateway-CE12808-2-bgp] peer 192.168.40.165 as-number 65021
[*Exit-Gateway-CE12808-2-bgp] peer 192.168.40.165 group Leaf-CE6851HI-1&CE6851HI-2
[*Exit-Gateway-CE12808-2-bgp] peer 192.168.40.161 as-number 65021
[*Exit-Gateway-CE12808-2-bgp] peer 192.168.40.161 group Leaf-CE6851HI-1&CE6851HI-2
[*Exit-Gateway-CE12808-2-bgp] group Leaf-CE6851HI-3&4 external
[*Exit-Gateway-CE12808-2-bgp] peer Leaf-CE6851HI-3&4 as-number 65022
[*Exit-Gateway-CE12808-2-bgp] peer 192.168.41.165 as-number 65022
[*Exit-Gateway-CE12808-2-bgp] peer 192.168.41.165 group Leaf-CE6851HI-3&4
[*Exit-Gateway-CE12808-2-bgp] peer 192.168.41.161 as-number 65022
[*Exit-Gateway-CE12808-2-bqp] peer 192.168.41.161 group Leaf-CE6851HI-3&4
[*Exit-Gateway-CE12808-2-bgp] group Leaf-CE6851HI-5&CE6851HI-6 external
[*Exit-Gateway-CE12808-2-bgp] peer Leaf-CE6851HI-5&CE6851HI-6 as-number 65024
[*Exit-Gateway-CE12808-2-bgp] peer 192.168.46.165 as-number 65024
[*Exit-Gateway-CE12808-2-bgp] peer 192.168.46.165 group Leaf-CE6851HI-5&CE6851HI-6 [*Exit-Gateway-CE12808-2-bgp] peer 192.168.46.161 as-number 65024
[*Exit-Gateway-CE12808-2-bgp] peer 192.168.46.161 group Leaf-CE6851HI-5&CE6851HI-6
[*Exit-Gateway-CE12808-2-bgp] ipv4-family unicast
[*Exit-Gateway-CE12808-2-bgp-af-ipv4] preference 20 200 10
[*Exit-Gateway-CE12808-2-bgp-af-ipv4] network 11.11.11.16 255.255.255.255
```

```
[*Exit-Gateway-CE12808-2-bgp-af-ipv4] network 19.19.19.19 255.255.255.255
[*Exit-Gateway-CE12808-2-bgp-af-ipv4] network 22.22.22.22 255.255.255
[*Exit-Gateway-CE12808-2-bgp-af-ipv4] network 192.168.44.160 255.255.255.255
[*Exit-Gateway-CE12808-2-bgp-af-ipv4] network 192.168.45.152 255.255.255.258
[*Exit-Gateway-CE12808-2-bgp-af-ipv4] maximum load-balancing 32
[*Exit-Gateway-CE12808-2-bgp-af-ipv4] quit
[*Exit-Gateway-CE12808-2-bgp] quit
[*Exit-Gateway-CE12808-2] commit
```

----结束

3.2.2.4 配置 BGP EVPN

在本方案中,Spine节点和Gateway融合部署,当使用BGP EVPN作为VXLAN的控制平面时,不再需要出口网关与Spine之间建立对等体关系,而是出口网关和Leaf间建立对等体。

步骤1 在出口网关上使能EVPN作为VXLAN的控制平面功能,其他设备上的配置与之类似,这里不再赘述。

```
[~Exit-Gateway-CE12808-1] evpn-overlay enable
[~Exit-Gateway-CE12808-1] commit
```

步骤2 配置出口网关分别与Leaf节点建立BGP EVPN对等体关系。

在Exit-Gateway-CE12808-1上配置与Leaf建立IBGP EVPN对等体关系,其他设备与之 类似,不再赘述。

```
[~Exit-Gateway-CE12808-1] bgp 100 instance evpn1
[*Exit-Gateway-CE12808-1-bgp-instance-evpn1] router-id 18.18.18.18
[*Exit-Gateway-CE12808-1-bgp-instance-evpn1] peer 11.11.11 as-number 100
//与堆叠组建立BGP EVPN对等体关系
[*Exit-Gateway-CE12808-1-bgp-instance-evpn1] peer 11.11.11 connect-interface loopback 0
[*Exit-Gateway-CE12808-1-bgp-instance-evpn1] peer 13.13.13.13 as-number 100
//与Leaf-CE6851HI-3建立BGP EVPN对等体关系
[*Exit-Gateway-CE12808-1-bgp-instance-evpn1] peer 13.13.13 connect-interface loopback 0
[*Exit-Gateway-CE12808-1-bgp-instance-evpn1] peer 14.14.14.14 as-number 100
//与Leaf-CE6851HI-4建立BGP EVPN对等体关系
[*Exit-Gateway-CE12808-1-bgp-instance-evpn1] peer 14.14.14 connect-interface loopback 0
[*Exit-Gateway-CE12808-1-bgp-instance-evpn1] peer 11.11.11.17 as-number 100
//与SVF组建立BGP EVPN对等体关系
[*Exit-Gateway-CE12808-1-bgp-instance-evpn1] peer 11.11.11.17 connect-interface loopback 0
[*Exit-Gateway-CE12808-1-bgp-instance-evpn1] l2vpn-family evpn
[*Exit-Gateway-CE12808-1-bgp-instance-evpn1-af-evpn] peer 11.11.11.11 enable
[*Exit-Gateway-CE12808-1-bgp-instance-evpn1-af-evpn] peer 13.13.13.13 enable
[*Exit-Gateway-CE12808-1-bgp-instance-evpn1-af-evpn] peer 14.14.14.14 enable
[*Exit-Gateway-CE12808-1-bgp-instance-evpn1-af-evpn] peer 11.11.11.17 enable
[*Exit-Gateway-CE12808-1-bgp-instance-evpn1-af-evpn] quit
[*Exit-Gateway-CE12808-1-bgp-instance-evpn1] quit
[*Exit-Gateway-CE12808-1] commit
```

----结束