



(12)发明专利

(10)授权公告号 CN 106209553 B

(45)授权公告日 2019.07.23

(21)申请号 201510219028.9

(22)申请日 2015.04.30

(65)同一申请的已公布的文献号
申请公布号 CN 106209553 A

(43)申请公布日 2016.12.07

(73)专利权人 华为技术有限公司
地址 518129 广东省深圳市龙岗区坂田华为总部办公楼

(72)发明人 高远 韩涛 蒋维廉

(51)Int.Cl.

H04L 12/46(2006.01)

H04L 12/723(2013.01)

H04L 12/741(2013.01)

H04L 12/755(2013.01)

(56)对比文件

CN 103580980 A,2014.02.12,

CN 104518940 A,2015.04.15,

CN 103607349 A,2014.02.26,

CN 102055647 A,2011.05.11,

US 2007058638 A1,2007.03.15,

US 2014086253 A1,2014.03.27,

US 2011032843 A1,2011.02.10,

审查员 胡冰舟

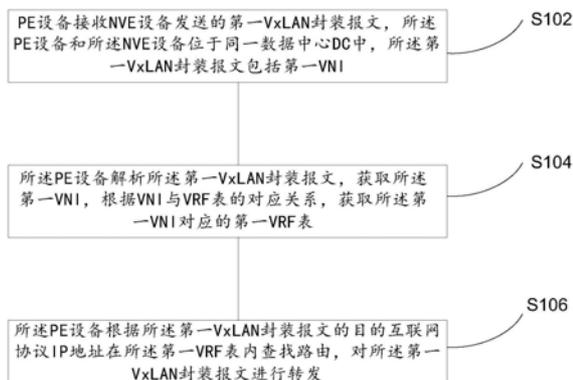
权利要求书2页 说明书10页 附图4页

(54)发明名称

报文处理方法、设备及系统

(57)摘要

一种报文处理方法、设备及系统,该方法包括,PE设备接收NVE设备发送的第一VxLAN封装报文,所述PE设备和所述NVE设备位于同一DC中,所述第一VxLAN封装报文包括第一VNI;所述PE设备解析所述第一VxLAN封装报文,获取所述第一VNI,根据VNI与VRF表的对应关系,获取所述第一VNI对应的第一VRF表;所述PE设备根据所述第一VxLAN封装报文的目的地IP地址在所述第一VRF内查找路由,对所述第一VxLAN封装报文进行转发。采用该方法,可以减少对接口资源的占用,简化配置。



1. 一种报文处理方法,其特征在于,所述方法包括:

运营商边缘PE设备接收网络虚拟边缘NVE设备发送的第一虚拟可扩展局域网VxLAN封装报文,所述PE设备和所述NVE设备位于同一数据中心DC中,所述第一VxLAN封装报文包括第一VxLAN网络标识符VNI;

所述PE设备解析所述第一VxLAN封装报文,获取所述第一VNI,根据VNI与虚拟专用网络VPN路由转发VRF表的对应关系,获取所述第一VNI对应的第一VRF表;

所述PE设备根据所述第一VxLAN封装报文的目的地互联网协议IP地址在所述第一VRF表内查找路由,对所述第一VxLAN封装报文进行转发。

2. 根据权利要求1所述的方法,其特征在于,所述PE设备部署边界网关协议BGP,所述PE设备接收所述NVE设备发送的BGP路由更新消息,所述BGP路由更新消息包括VxLAN隧道端点VTEP地址及所述VTEP地址对应的VNI。

3. 根据权利要求2所述的方法,其特征在于,还包括,所述PE设备根据所述VNI与VRF表的对应关系,将所述BGP路由更新消息中的所述VTEP地址学习到所述VRF表中。

4. 根据权利要求3所述的方法,其特征在于,还包括,所述PE设备通过BGP向远端PE设备发布所述VTEP地址。

5. 根据权利要求1-4中任一项所述的方法,其特征在于,还包括,所述PE设备在对所述第一VxLAN封装报文进行转发之前,对所述第一VxLAN封装报文进行多协议标签交换MPLS封装。

6. 一种运营商边缘PE设备,其特征在于,所述PE设备包括:

接收单元,用于接收网络虚拟边缘NVE设备发送的第一虚拟可扩展局域网VxLAN封装报文,所述PE设备和所述NVE设备位于同一数据中心DC中,所述第一VxLAN封装报文包括第一VxLAN网络标识符VNI;

处理单元,用于解析所述第一VxLAN封装报文,获取所述第一VNI,根据VNI与虚拟专用网络VPN路由转发VRF表的对应关系,获取所述第一VNI对应的第一VRF表;

转发单元,用于根据所述第一VxLAN封装报文的目的地互联网协议IP地址在所述第一VRF表内查找路由,对所述第一VxLAN封装报文进行转发。

7. 根据权利要求6所述的设备,其特征在于,所述PE设备部署有边界网关协议BGP,所述PE设备还包括路由更新消息接收单元,用于接收所述NVE设备发送的BGP路由更新消息,所述BGP路由更新消息包括VxLAN隧道端点VTEP地址及所述VTEP地址对应的VNI。

8. 根据权利要求7所述的设备,其特征在于,所述PE设备还包括学习单元,用于根据所述VNI与VRF表的对应关系,将所述BGP路由更新消息中的所述VTEP地址学习到所述VRF表中。

9. 根据权利要求8所述的设备,其特征在于,所述PE设备还包括发布单元,用于通过BGP向远端PE设备发布所述VTEP地址。

10. 根据权利要求6-9中任一项所述的设备,所述PE设备还包括MPLS封装单元,用于在对所述第一VxLAN封装报文进行转发之前,对所述第一VxLAN封装报文进行多协议标签交换MPLS封装。

11. 一种数据中心DC系统,其特征在于,包括权利要求6-10中任一项所述的PE设备和NVE设备,所述NVE设备包括:

BGP连接单元,用于建立所述NVE设备与运营商边缘PE设备之间的边界网关协议BGP连接,所述PE设备和所述NVE设备位于同一数据中心DC中;

对应单元,用于保存虚拟可扩展局域网VxLAN隧道端点VTEP地址与VxLAN网络标识符VNI的对应关系;

路由更新消息发送单元,用于通过所述BGP连接向所述PE设备发送BGP路由更新消息,所述BGP路由更新消息包括VTEP地址及所述VTEP地址对应的VNI。

报文处理方法、设备及系统

技术领域

[0001] 本发明涉及网络通信技术,尤其涉及一种报文处理方法、设备及系统。

背景技术

[0002] 虚拟可扩展局域网 (Virtual eXtensible Local Area Network, VxLAN) 是一种将二层报文用三层协议进行封装的技术, VxLAN 的主要技术原理就是引入一种 MAC-in-UDP 格式的封装报文, 即将以太帧封装在用户数据报协议 (User Datagram Protocol, UDP) 报文中传输, 可以使以太帧在三层网络中传送, 从而可以对二层网络在三层范围进行扩展。在数据中心 (DataCenter, DC) 与站点的虚拟网络互通业务场景中, DC 为多个企业提供虚拟机 (Virtual Machine, VM) 服务, 站点的主机通过 VxLAN 隧道与 DC 中的 VM 进行通信。在实现方面, DC 侧运营商边缘 (Provider Edge, PE) 设备需要为每个租户创建对应的子接口, 每个子接口绑定相应的虚拟专用网络 (Virtual Private Network, VPN) 路由转发 (Virtual Routing Forwarding, VRF) 表, 从而实现 DC 内网络和运营商网络之间互联。

[0003] 但是, 在实际应用中, 由于需要在 DC 内的 PE 设备上为每一个 VRF 分配一个子接口, 然后将分配的子接口和对应的 VRF 进行绑定, 因此存在如下问题: 新增加租户时, 需要在 DC 侧 PE 设备上分配新的子接口, 占用较多的所述 DC 侧 PE 设备接口资源, 而且相应的配置过程复杂。

发明内容

[0004] 有鉴于此, 本发明实施例提供了一种报文处理方法、设备及系统, 以实现在虚拟化网络的 DC 与站点互通的业务场景中, 简化 DC 中的 PE 设备的配置。

[0005] 本发明实施例提供的技术方案如下。

[0006] 第一方面, 提供了一种报文处理方法, 包括:

[0007] PE 设备接收网络虚拟边缘 (Network Virtualization Edge, NVE) 设备发送的第一 VxLAN 封装报文, 所述 PE 设备和所述 NVE 设备位于同一数据中心 DC 中, 所述第一 VxLAN 封装报文包括第一 VxLAN 网络标识符 (VxLAN Network Identifier, VNI);

[0008] 所述 PE 设备解析所述第一 VxLAN 封装报文, 获取所述第一 VNI, 根据 VNI 与 VRF 表的对应关系, 获取所述第一 VNI 对应的第一 VRF 表;

[0009] 所述 PE 设备根据所述第一 VxLAN 封装报文的目的地互联网协议 (Internet Protocol, IP) 地址在所述第一 VRF 表内查找路由, 对所述第一 VxLAN 封装报文进行转发。

[0010] 在第一方面的第一种可能的实现方式中, 所述方法还包括:

[0011] 所述 PE 设备部署边界网关协议 (Border Gateway Protocol, BGP), 所述 PE 设备接收所述 NVE 设备发送的 BGP 路由更新消息, 所述 BGP 路由更新消息包括 VxLAN 隧道端点 (VxLAN Tunnel End Point, VTEP) 地址及所述 VTEP 地址对应的 VNI。

[0012] 在第一方面的第二种可能的实现方式中, 所述方法还包括:

[0013] 所述 PE 设备根据所述 VNI 与 VRF 表的对应关系, 将所述 BGP 路由更新消息中的所述

VTEP地址学习到所述VRF表中。

[0014] 在第一方面的第三种可能的实现方式中,所述方法还包括:

[0015] 所述PE设备通过BGP向远端PE设备发布所述VTEP地址。

[0016] 结合上述第一方面,或第一方面的任意一种可能的实现方式,第一方面的第四种可能的实现方式中,所述PE设备在对所述第一VxLAN封装报文进行转发之前,对所述第一VxLAN封装报文进行多协议标签交换(Multiprotocol Label Switching,MPLS)封装。

[0017] 第二方面,提供了一种报文处理方法,包括:

[0018] NVE设备建立与PE设备之间的BGP连接,所述PE设备和所述NVE设备位于同一数据中心DC中;

[0019] 所述NVE设备保存VTEP地址与VNI的对应关系;

[0020] 所述NVE设备通过所述BGP连接向所述PE设备发送BGP路由更新消息,所述BGP路由更新消息包括VTEP地址及所述VTEP地址对应的VNI。

[0021] 第三方面,提供了一种PE设备,包括:

[0022] 接收单元,用于接收NVE设备发送的第一VxLAN封装报文,所述PE设备和所述NVE设备位于同一数据中心DC中,所述第一VxLAN封装报文包括第一VNI;

[0023] 处理单元,用于解析所述第一VxLAN封装报文,获取所述第一VNI,根据VNI与VRF表的对应关系,获取所述第一VNI对应的第一VRF表;

[0024] 转发单元,用于根据所述第一VxLAN封装报文的IP地址在所述第一VRF表内查找路由,对所述第一VxLAN封装报文进行转发。

[0025] 在第三方面的第一种可能的实现方式中,所述PE设备部署有边界网关协议BGP,所述PE设备还包括:

[0026] 路由更新消息接收单元,用于接收所述NVE设备发送的BGP路由更新消息,所述BGP路由更新消息包括VxLAN隧道端点VTEP地址及所述VTEP地址对应的VNI。

[0027] 在第三方面的第二种可能的实现方式中,所述PE设备还包括:

[0028] 学习单元,用于根据所述VNI与VRF表的对应关系,将所述BGP路由更新消息中的所述VTEP地址学习到所述VRF表中。

[0029] 在第三方面的第三种可能的实现方式中,所述PE设备还包括:

[0030] 发布单元,用于通过BGP向远端PE设备发布所述VTEP地址。

[0031] 结合上述第三方面,或第三方面的任意一种可能的实现方式,第三方面的第四种可能的实现方式中,所述PE设备还包括MPLS封装单元,用于在对所述第一VxLAN封装报文进行转发之前,对所述第一VxLAN封装报文进行MPLS封装。

[0032] 第四方面,提供了一种NVE设备,包括:

[0033] BGP连接单元,用于建立所述NVE设备与PE设备之间的BGP连接,所述PE设备和所述NVE设备位于同一数据中心DC中;

[0034] 对应单元,用于保存VTEP地址与VNI的对应关系;

[0035] 路由更新消息发送单元,用于通过所述BGP连接向所述PE设备发送BGP路由更新消息,所述BGP路由更新消息包括VTEP地址及所述VTEP地址对应的VNI。

[0036] 第五方面,提供了一种DC系统,包括:

[0037] 上述第三方面或第三方面的任意一种可能的实现方式所提供的PE设备和上述第

四方面或第四方面的任意一种可能的实现方式所提供的NVE设备。

[0038] 本发明实施方式的技术方案中,DC侧PE设备接收DC侧NVE设备发送的第一VxLAN封装报文后,解析所述第一VxLAN封装报文,获取第一VNI,根据VNI与VRF表的对应关系,获取所述第一VNI对应的第一VRF表,从而根据所述第一VxLAN封装报文的目的地IP地址在所述第一VRF表内查找路由,对第一VxLAN封装报文进行转发,所述PE设备和所述NVE设备位于同一DC中。相应的,NVE设备建立与PE设备之间的BGP连接,所述PE设备和所述NVE设备位于同一DC中,NVE设备保存VTEP地址与VNI的对应关系,通过BGP连接发送BGP路由更新消息,所述BGP路由更新消息包括携带有VNI信息的VTEP地址。本发明实施方式还提供了一种PE设备、NVE设备和DC系统,解决了由于新增租户时创建子接口,带来的占用接口资源多,相应的配置过程复杂的问题,简化了DC侧PE设备的配置;而且,通过扩展BGP协议,将VTEP地址自动更新和学习到相应的VRF表中,进一步简化了配置过程。

附图说明

[0039] 为了更清楚地说明本发明实施例或现有技术中的方案,下面将对实施例中所需要使用的附图作一简单地介绍,显而易见地,下面描述中的附图是本发明的一些实施例,对于本领域普通技术人员而言,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

[0040] 图1为本发明第一实施例的报文处理方法的流程图;

[0041] 图2为本发明第二实施例的报文处理方法的流程图;

[0042] 图3为本发明实施例方法的应用场景示意图;

[0043] 图4为本发明实施例的PE设备结构示意图;

[0044] 图5为本发明实施例的PE设备硬件结构示意图;

[0045] 图6为本发明实施例的NVE设备结构示意图;

[0046] 图7为本发明实施例的NVE设备硬件结构示意图;

[0047] 图8为本发明实施例的数据中心的结构示意图。

具体实施方式

[0048] 本发明实施例提供一种报文处理方法、设备及系统,以实现在虚拟化网络的DC与站点互通的业务场景中,简化DC中的PE设备的配置。另外,还可以通过扩展BGP协议,将VTEP地址自动更新和学习到相应的VRF表中,进一步简化配置过程。

[0049] 下面通过具体实施例,对本发明的技术方案进行举例说明。

[0050] 为使得本发明的发明目的、特征、优点能更加的明显和易懂,下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚地描述,显然下面所描述的实施例仅仅是本发明一部分实施例,而非全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动的前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0051] 本申请的说明书和权利要求书及附图中的术语“第一”、“第二”、“第三”和“第四”等是用于区别不同对象,而不是用于描述特定顺序。此外,术语“包括”和“具有”不是排他的。例如包括了一系列步骤或单元的过程、方法、系统、产品或设备没有限定于已列出的步

骤或单元,还可以包括没有列出的步骤或单元。

[0052] 图1为本发明第一实施例的报文处理方法的流程图,如图1所示,该方法可以包括:

[0053] S102,PE设备接收NVE设备发送的第一VxLAN封装报文,所述PE设备和所述NVE设备位于同一DC中,所述第一VxLAN封装报文包括第一VNI。

[0054] 举例说明,DC可以包括PE设备、NVE设备和VM设备,DC侧VM对业务报文封装后,发送封装后的业务报文给所述DC侧NVE设备。所述DC侧NVE设备接收所述DC侧VM发送的业务报文后,查找主机所属的NVE的IP地址,对业务报文进行VxLAN隧道封装,隧道目的端点VTEP是远端主机所归属的VxLAN网关,VxLAN网关可以是远端的NVE设备。DC侧PE设备接收DC侧NVE设备发送的第一VxLAN封装报文,所述第一VxLAN封装报文包括第一VNI。所述第一VNI可以具有24比特长度,可以支持多达16M的VxLAN段的网络隔离。例如在一个虚拟网络段中可以支持多达16M的用户,从而对用户进行隔离和标识不再受到限制,可满足海量用户。

[0055] S104,所述PE设备解析所述第一VxLAN封装报文,获取所述第一VNI,根据VNI与VRF表的对应关系,获取所述第一VNI对应的第一VRF表。

[0056] 举例说明,DC侧PE设备接收到第一VxLAN封装报文后,对第一VxLAN封装报文进行解析。所述解析可以理解为对报文的具体操作手段,例如“读取”手段,获取第一VxLAN封装报文中的第一VNI。在DC侧PE设备上保存有VNI与VRF表的对应关系。举例来讲,保存VNI与VRF表的对应关系的实现方式可以是:每个VPN实例可以对应一个VRF表,首先能够确定VNI与VPN实例的对应关系,然后通过VPN实例找到相应的VRF表,将VNI的标识信息添加到VRF表中,形成了VNI与VRF表的对应关系。非限定性的,一个VNI可以对应一个VRF表,也可以多个VNI对应一个VRF表。当多个VNI对应一个VRF表时,只需将多个VNI标识信息添加到同一个VRF表中即可。当获取所述第一VNI时,可以通过查找VRF表中的VNI标识信息,找到所述第一VNI对应的第一VRF表。

[0057] 举例说明,对于VNI与VRF表的对应关系的确定方式,可选的,例如可以在部署DC时或变更DC中的设备时将每个VNI与对应的VRF表进行对应,具体可以通过DC侧NVE设备将所有VNI发送到DC侧PE设备,从而使每个VNI与对应的VRF表进行对应。另外,DC侧NVE设备可以定期发送VNI,保持信息更新。

[0058] S106,所述PE设备根据所述第一VxLAN封装报文的IP地址在所述第一VRF表内查找路由,对所述第一VxLAN封装报文进行转发。

[0059] 举例说明,DC侧PE设备在根据所述第一VNI获取所述第一VRF表后,根据包括有所述第一VNI的所述第一VxLAN封装报文的IP地址在所述第一VRF表内查找相应的路由,对所述第一VxLAN封装报文进行转发,从而通过VxLAN隧道发送到远端的PE设备。所述远端的PE设备可以是站点侧的PE设备。可选的,所述远端的PE设备可以是运营商(provider,P)设备。

[0060] 本实施例提供的报文处理方法,DC侧PE设备通过解析第一VxLAN封装报文,获取第一VNI,根据VNI与VRF表的对应关系,获取第一VNI对应的第一VRF表,基于第一VRF表对所述第一VxLAN封装报文进行转发。这样,在新增租户时,不需要在DC侧PE设备上分配新的子接口,可以减少接口资源的占用,简化配置。

[0061] 可选的,所述PE设备部署边界网关协议BGP,所述PE设备接收所述NVE设备发送的BGP路由更新消息,所述BGP路由更新消息包括VxLAN隧道端点VTEP地址及所述VTEP地址对

应的VNI。

[0062] 举例说明,DC侧和站点侧在进行更新时,例如变更站点的企业主机,需要对DC侧PE设备的路由信息进行更新。可以在DC侧PE设备上部署BGP,DC侧PE设备接收DC侧NVE设备发送的BGP路由更新消息,BGP路由更新消息包括VTEP地址,DC侧NVE设备上保存有VTEP与VNI的对应关系,DC侧PE设备接收的VTEP地址可以包括相应的VNI信息。

[0063] 可选的,所述PE设备根据所述的VNI与VRF表的对应关系,将所述BGP路由更新消息中的所述VTEP地址学习到所述VRF表中。

[0064] 举例说明,DC侧PE设备上保存有VNI与VRF表的对应关系。DC侧PE设备在接收到携带有相应的VNI信息的VTEP地址后,根据所述VNI的关联,将所述BGP路由更新消息中的VTEP地址学习到VRF表中,从而保持路由信息的更新。另外,本发明对于更新的周期不进行限定,例如可以采用实时或定期更新。

[0065] 可选的,所述PE设备通过BGP向远端PE设备发布所述VTEP地址。

[0066] 举例说明,为了保证DC侧和站点侧的路由信息的一致,DC侧PE设备通过BGP向远端PE设备发布所述VTEP地址,其中远端PE设备可以包括站点侧PE设备。

[0067] 可选的,所述PE设备在对所述第一VxLAN封装报文进行转发之前,对所述第一VxLAN封装报文进行MPLS封装。

[0068] 举例说明,DC侧PE设备在对第一VxLAN封装报文进行转发之前,可以进行MPLS封装,MPLS封装后的报文格式是:外层以太网(Ethernet,ETH)头+基于流量工程的(Traffic Engineering,TE)隧道MPLS标签+VPN MPLS标签+外层VxLAN IP头(NVE目的、源IP地址)+UDP头+VxLAN头+内层ETH头+内层IP地址。

[0069] 在不采用本发明实施例方案的通常实现方式中,DC侧PE设备需要为站点侧的每一个租户创建对应的子接口,通过子接口绑定相应的VRF,从而通过构建子接口的方式将VxLAN封装报文发送到远端PE设备。但是,在近年的网络发展中,网络的集成度不断提升,用户量大幅增加,使用上述通常的实现方式,占用大量的DC侧PE设备的接口资源,并且需要进行大量配置工作,配置复杂,显然,这种常规的通过配备子接口的方式并不能满足如此灵活的应用需求。

[0070] 本发明实施例提供的技术方案中,通过解析第一VxLAN封装报文,获取第一VNI,根据VNI与VRF表的对应关系,获取第一VNI对应的第一VRF表,基于第一VRF表对所述第一VxLAN封装报文进行转发,从而在新增加租户时,不需要在PE设备上分配新的子接口,减少接口资源的占用,简化配置,进而有利于较好满足目前网络演进的需求。

[0071] 上述实施例描述了本发明报文处理方法中的DC侧PE设备执行的具体方法步骤,下面对本发明报文处理方法中的DC侧NVE设备执行的具体方法步骤进行详细说明。

[0072] 图2为本发明第二实施例的报文处理方法的流程图,如图2所示,该方法可以包括:

[0073] S202,NVE设备建立与PE设备之间的BGP连接,所述PE设备和所述NVE设备位于同一DC中。

[0074] 举例说明,DC侧NVE设备不仅向DC侧PE设备发送VxLAN封装报文,而且为了保障DC侧PE设备的路由信息的及时更新,例如增加租户、业务变更等,可以通过BGP扩展的方式进行路由信息更新。因此,DC侧NVE设备可以建立与DC侧PE设备之间的BGP连接。

[0075] S204,所述NVE设备保存VTEP地址与VNI的对应关系。

[0076] 举例说明,在DC侧NVE设备中,通过VTEP进行业务报文的VxLAN隧道封装,DC侧NVE设备可以保存VTEP地址与VNI的对应关系,这样,VTEP地址和VNI就形成了相应的对应关系。

[0077] S206,所述NVE设备通过所述BGP连接向所述PE设备发送BGP路由更新消息,所述BGP路由更新消息包括VTEP地址及所述VTEP地址对应的VNI。

[0078] 举例说明,DC侧NVE设备通过BGP连接向DC侧PE设备发送BGP路由更新消息。所述BGP路由更新消息包括VTEP地址,在VTEP地址中可以携带VNI信息。DC侧PE设备在接收到携带有相应的VNI信息的VTEP地址后,由于DC侧PE设备上保存有VNI与VRF表的对应关系,可以根据所述VNI的关联,将所述BGP路由更新消息中的VTEP地址学习到VRF表中,从而保持VRF表中的路由信息的更新,另外,本发明实施例对于VRF中的路由信息的更新周期不进行限定,可以实时或定期更新VRF表中的路由信息。

[0079] 举例说明,所述DC侧NVE设备接收DC侧VM发送的业务报文,对所述业务报文进行VxLAN封装,获得VxLAN封装报文;所述DC侧NVE设备向所述DC侧PE设备发送所述VxLAN封装报文。

[0080] 本实施例提供的报文处理方法,通过BGP扩展的方式,DC侧NVE设备将VTEP地址发布给DC侧PE设备,并且将发布的VTEP路由信息学习到相应的VRF表中,从而保证了DC侧PE设备中路由信息的及时更新,在面对例如新增加租户等业务需求时,提高业务部署的灵活性。

[0081] 上述实施例描述了本发明报文处理方法中的DC侧NVE设备执行的具体方法步骤,下面对本发明实施例方法的应用场景示意图进行详细说明。

[0082] 图3为本发明实施例方法的应用场景示意图,如图3所示,其中,报文从DC侧VM1经由DC侧NVE1及DC侧PE设备、MPLS网络到达主机1。图3示例性的说明了DC与企业站点之间的报文处理过程,具体执行过程是:

[0083] (1) DC可以为多个企业提供VM服务,每个企业在DC侧PE设备上具有对应的VRF表,DC侧VM1对业务报文进行封装,向DC侧NVE1设备发送业务报文。

[0084] (2) DC侧NVE1设备接收DC侧VM1发送的业务报文,查找企业主机1所属的NVE的IP地址,对业务报文进行VxLAN隧道封装,获取第一VxLAN封装报文,隧道目的端点VTEP是企业主机所归属的VxLAN网关,所述VxLAN网关可以是远端的NVE设备,DC侧NVE1设备通过建立的VxLAN隧道向企业主机所归属的VxLAN网关发送所述第一VxLAN封装报文。

[0085] 举例说明,对于查找企业主机1所属的NVE的IP地址的方法,示例性的可以是:DC侧NVE1设备中存储有企业主机1的MAC地址与主机1所属的NVE的IP地址的映射关系,通过这种映射关系可以查找到企业主机1所属的NVE的IP地址。

[0086] (3) DC侧PE设备接收DC侧NVE1设备发送的第一VxLAN封装报文,解析所述第一VxLAN封装报文,获取第一VNI,根据VNI与VRF表的对应关系,获取所述第一VNI对应的第一VRF表;DC侧PE设备根据所述第一VxLAN封装报文的的目的IP地址在所述第一VRF表内查找路由,对所述第一VxLAN封装报文进行转发。

[0087] 可选的,DC侧PE设备可以对第一VxLAN封装报文进行MPLS封装,MPLS封装后的报文格式是:外层ETH头+TE隧道MPLS标签+VPN MPLS标签+外层VxLAN IP头(NVE目的、源IP地址)+UDP头+VxLAN头+内层ETH头+内层IP地址。

[0088] (4) 站点侧PE设备接收MPLS封装后的报文,对报文进行MPLS解封,根据对应VRF表查找路径,将第一VxLAN封装报文发送到站点侧NVE设备。

[0089] (5) 站点侧NVE设备接收第一VxLAN封装报文,进行VxLAN解封,将业务报文发送到对应主机1。

[0090] 另外,DC侧NVE1可以通过扩展BGP的方式,实现路由信息的更新,具体过程是:

[0091] (1) DC侧NVE1设备建立与DC侧PE设备之间的BGP连接。

[0092] (2) 所述DC侧NVE1设备保存VTEP地址与VNI的对应关系。

[0093] (3) 所述DC侧NVE1设备通过所述BGP连接发送BGP路由更新消息,所述BGP路由更新消息包括VTEP地址及所述VTEP地址对应的VNI。

[0094] 其中,DC侧NVE1设备通过BGP连接发送BGP路由更新消息,BGP路由更新消息包括VTEP地址,在VTEP地址中可以携带VNI信息。DC侧PE设备在接收到携带有相应的VNI信息的VTEP地址后,由于DC侧PE设备上保存有VNI与VRF表的对应关系,可以根据所述VNI的关联,将所述BGP路由更新消息中的VTEP地址学习到VRF表中,从而保持路由信息的更新。另外,本发明对于更新的周期不进行限定,例如可以采用实时或定期更新。

[0095] 本实施例详细说明了DC与企业站点之间的报文处理方法对应的应用场景,其中每个步骤的执行方式请参见第一实施例和第二实施例中的相应描述,通过上述实施过程,解决了由于新增租户时创建子接口,从而带来的占用接口资源多,相应的配置过程复杂的问题,简化了DC侧PE设备的配置。而且,通过扩展BGP协议,将VTEP地址自动更新和学习到相应的VRF表中,进一步简化了配置过程。

[0096] 上述实施例描述了本发明实施例方法的应用场景示意图,下面对本发明实施例的PE设备结构示意图和硬件结构示意图进行详细说明。

[0097] 图4为本发明实施例的PE设备结构示意图,如图4所示,PE设备包括接收单元402、处理单元404和转发单元406:

[0098] 所述接收单元402,用于接收NVE设备发送的第一VxLAN封装报文,所述PE设备和所述NVE设备位于同一DC中,所述第一VxLAN封装报文包括第一VNI;

[0099] 所述处理单元404,用于解析所述第一VxLAN封装报文,获取所述第一VNI,根据VNI与VRF表的对应关系,获取所述第一VNI对应的第一VRF表;

[0100] 所述转发单元406,用于根据所述第一VxLAN封装报文的IP地址在所述第一VRF表内查找路由,对所述第一VxLAN封装报文进行转发。

[0101] 可选的,所述PE设备部署有BGP,所述PE设备还包括路由更新消息接收单元4001,用于部署边界网关协议BGP,接收所述NVE设备发送的BGP路由更新消息,所述BGP路由更新消息包括VxLAN隧道端点VTEP地址及所述VTEP地址对应的VNI。

[0102] 可选的,所述PE设备还包括学习单元4002,用于根据所述的VNI与VRF表的对应关系,将所述BGP路由更新消息中的所述VTEP地址学习到所述VRF表中。

[0103] 可选的,所述PE设备还包括发布单元4003,用于通过BGP向远端PE设备发布所述VTEP地址。

[0104] 可选的,所述PE设备还包括MPLS封装单元4004,用于在对所述第一VxLAN封装报文进行转发之前,对所述第一VxLAN封装报文进行MPLS封装。

[0105] 图4所示的PE设备可以执行上述实施例的方法中的相应步骤。DC侧PE设备通过解析第一VxLAN封装报文,获取第一VNI,根据VNI与VRF表的对应关系,获取第一VNI对应的第一VRF表,基于第一VRF表对VxLAN封装报文进行转发,从而减少接口资源的占用,简化配置。

[0106] 图5为本发明实施例的PE设备硬件结构图,图5所示的PE设备可以执行上述实施例的方法中DC侧PE设备执行的相应步骤。如图5所示,PE设备包括处理器501、存储器502、接口503和总线504,其中接口503可以通过无线或有线的方式实现,具体来讲可以是例如网卡(Network Interface Card,NIC)等元件,上述处理器501、存储器502、接口503通过总线504连接。

[0107] 所述存储器502用于存储程序代码,可选的,程序代码可以包括操作系统程序和应用程序。

[0108] 所述接口503,用于接收NVE设备发送的第一VxLAN封装报文,所述PE设备和所述NVE设备位于同一DC中,所述第一VxLAN封装报文包括第一VNI;

[0109] 所述处理器501,用于解析所述第一VxLAN封装报文,获取所述第一VNI,根据VNI与VRF表的对应关系,获取所述第一VNI对应的第一VRF表;

[0110] 所述接口503,还用于根据所述第一VxLAN封装报文的IP地址在所述第一VRF表内查找路由,对所述第一VxLAN封装报文进行转发。

[0111] 可选的,所述PE设备部署有BGP,所述接口503,还用于接收所述NVE设备发送的BGP路由更新消息,所述BGP路由更新消息包括VxLAN隧道端点VTEP地址及所述VTEP地址对应的VNI。

[0112] 可选的,所述处理器501,还用于根据所述的VNI与VRF表的对应关系,将所述BGP路由更新消息中的所述VTEP地址学习到所述VRF表中。

[0113] 可选的,所述接口503,还用于通过BGP向远端PE设备发布所述VTEP地址。

[0114] 可选的,所述处理器501,还用于在对所述第一VxLAN封装报文进行转发之前,对所述第一VxLAN封装报文进行MPLS封装。

[0115] 图5所示的PE设备可以用于执行前述实施例中方法中DC侧PE设备所执行的相应步骤。DC侧PE设备通过解析第一VxLAN封装报文,获取第一VNI,根据VNI与VRF表的对应关系,获取第一VNI对应的第一VRF表,基于第一VRF表对所述第一VxLAN封装报文进行转发,从而减少接口资源的占用,简化配置。

[0116] 上述实施例描述了本发明实施例的DC侧PE设备结构示意图和硬件结构示意图,下面对本发明实施例的DC侧NVE设备结构示意图和硬件结构示意图进行详细说明。

[0117] 图6为本发明实施例的NVE设备结构示意图,图6所示的NVE设备可以执行上述实施例的方法中DC侧NVE设备所执行的相应步骤。如图6所示,NVE设备包括BGP连接单元602、对应单元604和路由更新消息发送单元606:

[0118] 所述BGP连接单元602,用于建立所述DC侧NVE设备与DC侧PE设备之间的BGP连接,所述PE设备和所述NVE设备位于同一DC中;

[0119] 所述对应单元604,用于保存VTEP地址与VNI的对应关系;

[0120] 所述路由更新消息发送单元606,用于通过所述BGP连接向所述PE设备发送BGP路由更新消息,所述BGP路由更新消息包括VTEP地址及所述VTEP地址对应的VNI。

[0121] 可选的,NVE设备还包括接收单元6001、封装单元6002和发送单元6003。所述接收单元6001,用于接收DC侧VM发送的业务报文。所述封装单元6002,用于对所述业务报文进行VxLAN封装,获得VxLAN封装报文;所述发送单元6003,用于向所述DC侧PE设备发送所述VxLAN封装报文。

[0122] 图6所示的NVE设备可以用于执行前述实施例的方法中的DC侧NVE设备所执行的相应步骤。通过BGP扩展的方式,DC侧NVE设备将VTEP地址发布给DC侧PE设备,并且学习到相应的VRF表中,从而保证了路由信息的及时更新,在面对例如新增加租户等业务需求时,简化配置。

[0123] 图7为本发明实施例的NVE设备硬件结构图,图7所示的NVE设备可以执行上述实施例的方法中DC侧NVE设备所执行的相应步骤。如图7所示,NVE设备包括处理器701、存储器702、接口703和总线704,其中接口703可以通过无线或有线的方式实现,具体来讲可以是例如网卡等元件,上述处理器701、存储器702、接口703通过总线704连接。

[0124] 所述存储器702用于存储程序代码,可选的,程序代码可以包括操作系统程序和应用程序。

[0125] 所述接口703,用于建立所述NVE设备与PE设备之间的BGP连接,所述PE设备和所述NVE设备位于同一DC中;

[0126] 所述处理器701,用于保存VTEP地址与VNI的对应关系;

[0127] 所述接口703,用于通过所述BGP连接向所述PE设备发送BGP路由更新消息,所述BGP路由更新消息包括VTEP地址及所述VTEP地址对应的VNI。

[0128] 图7所示的NVE设备可以用于执行上述实施例的方法中DC侧NVE设备所执行的相应步骤。通过BGP扩展的方式,DC侧NVE设备将VTEP地址发布给DC侧PE设备,并且学习到相应的VRF表中,从而保证了路由信息的及时更新,在面对例如新增加租户等业务需求时,简化配置。

[0129] 上述实施例描述了本发明实施例的NVE设备结构示意图和硬件结构示意图,下面对本发明实施例的数据中心的结构进行详细说明。

[0130] 图8为本发明实施例的数据中心DC。如图8所示,数据中心DC包括PE设备和NVE设备。该数据中心DC可以执行第一实施例和第二实施例的方法,该PE设备和NVE设备可以具有图4-7及其对应的说明书内容所描述的结构。所述数据中心的实现原理和技术效果类似,此处,不再赘述。

[0131] 本发明实施例的数据中心、PE设备及NVE设备,均可执行上述实施例中的方法和步骤。

[0132] 举例说明,PE设备的接收单元402,用于接收NVE设备的发送单元6003发送的VxLAN封装报文,所述PE设备和所述NVE设备位于同一DC中,所述第一VxLAN封装报文包括第一VNI。所述PE设备的处理单元404,用于解析所述第一VxLAN封装报文,获取所述第一VNI,根据VNI与VRF表的对应关系,获取所述第一VNI对应的第一VRF表。所述PE设备的转发单元406,用于根据所述第一VxLAN封装报文的IP地址在所述第一VRF表内查找路由,对所述第一VxLAN封装报文进行转发。

[0133] 举例说明,NVE设备的BGP连接单元602,用于建立所述NVE设备与PE设备之间的BGP连接,所述PE设备和所述NVE设备位于同一DC中。所述NVE设备的对应单元604,用于保存VTEP地址与VNI的对应关系。所述NVE设备的路由更新消息发送单元606,用于通过所述BGP连接向所述PE设备的路由更新消息接收单元4001发送BGP路由更新消息,所述BGP路由更新消息包括VTEP地址及所述VTEP地址对应的VNI。

[0134] 本领域普通技术人员将会理解,本发明的各个方面、或各个方面的可能实现方式

可以被具体实施为系统、方法或者计算机程序产品。因此,本发明的各方面、或各个方面的可能实现方式可以采用完全硬件实施例、完全软件实施例(包括固件、驻留软件等等),或者组合软件和硬件方面的实施例的形式,在这里都统称为“电路”、“模块”或者“系统”。此外,本发明的各方面、或各个方面的可能实现方式可以采用计算机程序产品的形式,计算机程序产品是指存储在计算机可读介质中的计算机可读程序代码。

[0135] 计算机可读介质可以是计算机可读信号介质或者计算机可读存储介质。计算机可读存储介质包含但不限于电子、磁性、光学、电磁、红外或半导体系统、设备或者装置,或者前述的任意适当组合,如随机存取存储器(RAM)、只读存储器(ROM)、可擦除可编程只读存储器(EPROM或者快闪存储器)、光纤、便携式只读存储器(CD-ROM)。

[0136] 计算机中的处理器读取存储在计算机可读介质中的计算机可读程序代码,使得处理器能够执行在流程图中每个步骤、或各步骤的组合中规定的功能动作;生成实施在框图的每一块、或各块的组合中规定的功能动作的装置。

[0137] 计算机可读程序代码可以完全在用户的本地计算机上执行、部分在用户的本地计算机上执行、作为单独的软件包、部分在用户的本地计算机上并且部分在远程计算机上,或者完全在远程计算机或者服务器上执行。也应该注意,在某些替代实施方案中,在流程图中各步骤、或框图中各块所注明的功能可能不按图中注明的顺序发生。例如,依赖于所涉及的功能,接连示出的两个步骤、或两个块实际上可能被大致同时执行,或者这些块有时候可能被以相反顺序执行。

[0138] 显然,本领域的技术人员可以对本发明进行各种改动和变型而不脱离本发明的精神和范围。这样,倘若本发明的这些修改和变型属于本发明权利要求及其等同技术的范围之内,则本发明也意图包含这些改动和变型在内。

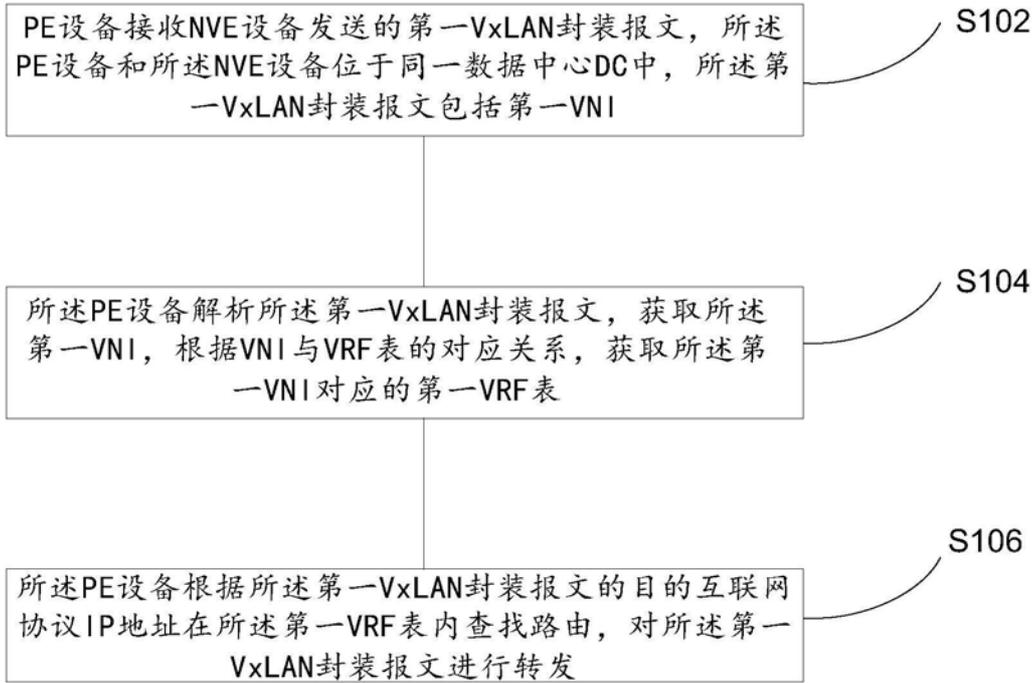


图1

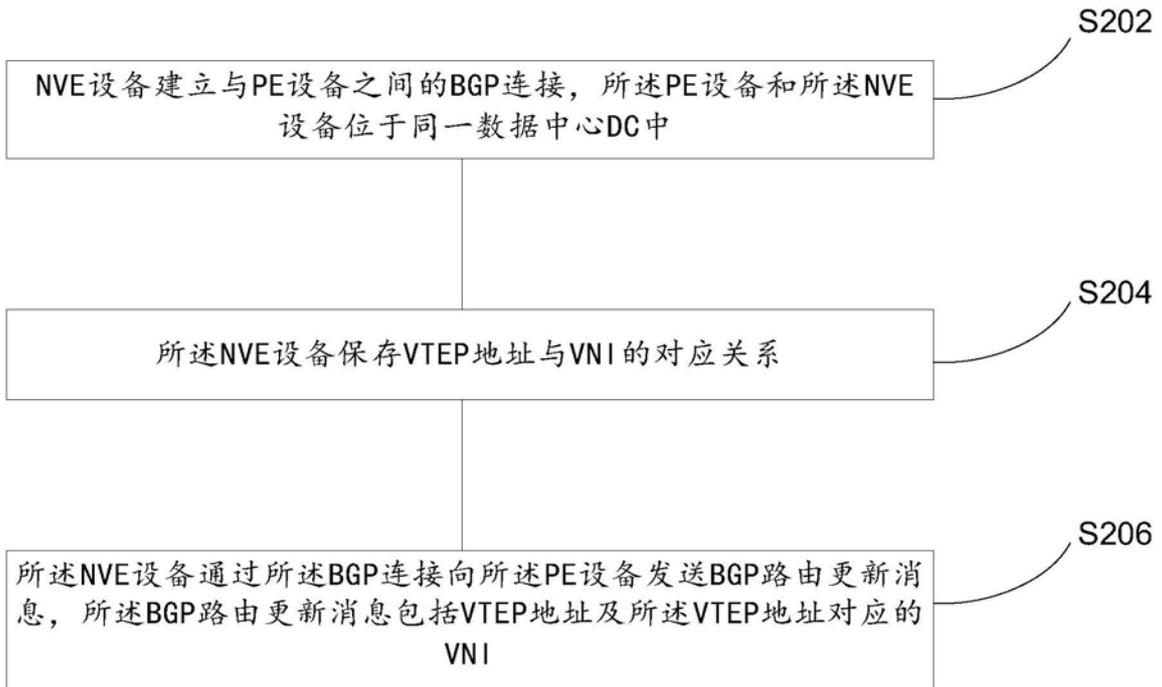


图2

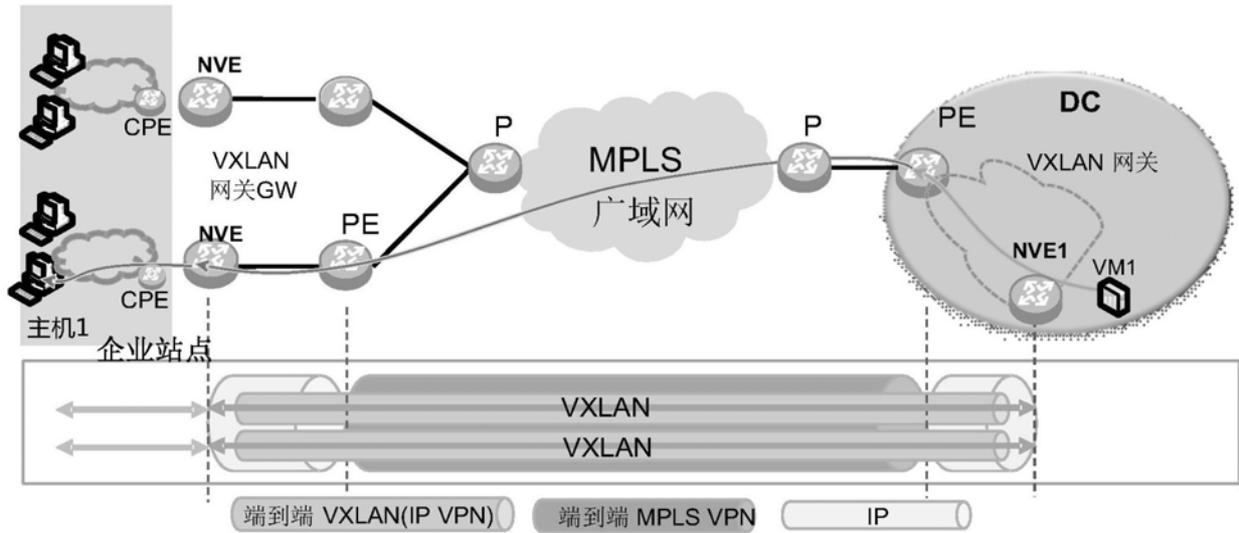


图3

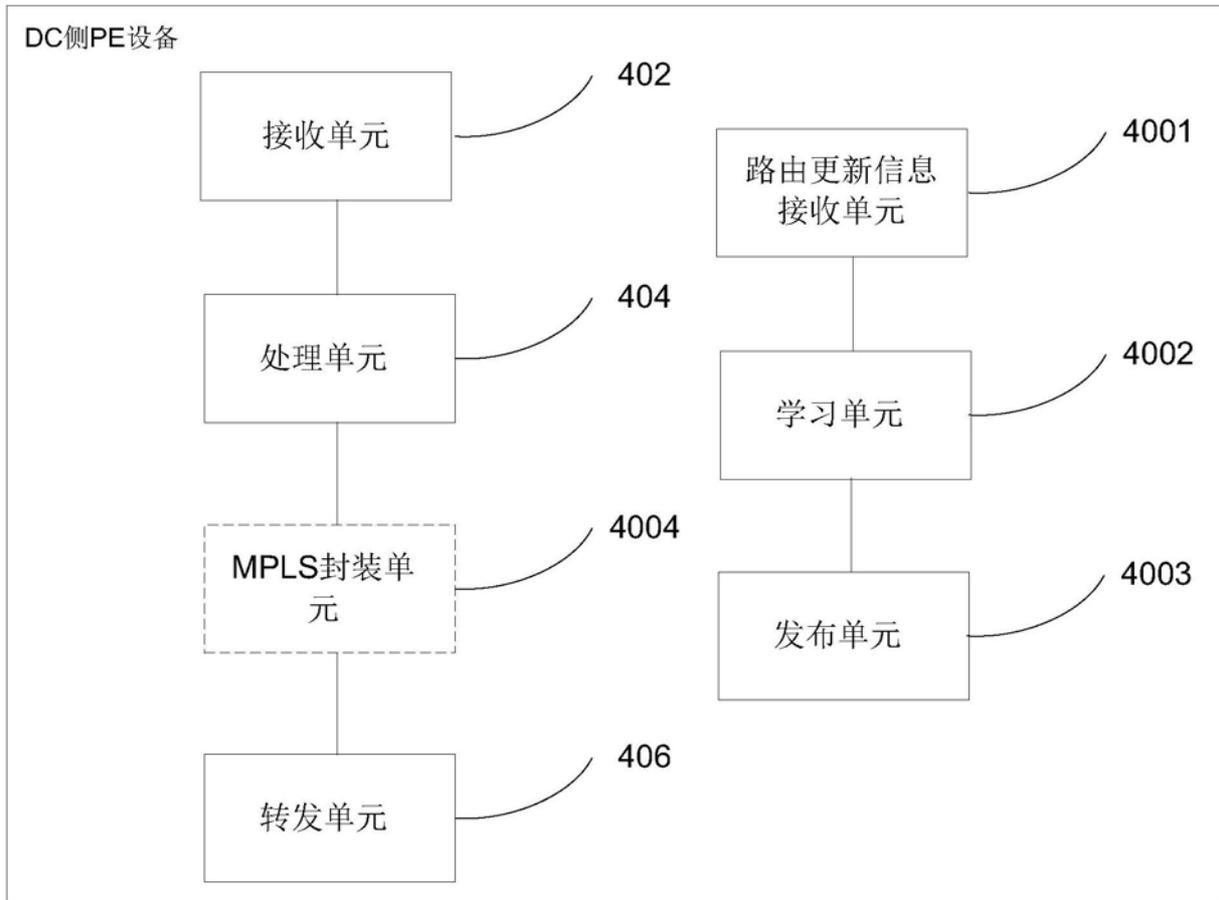


图4

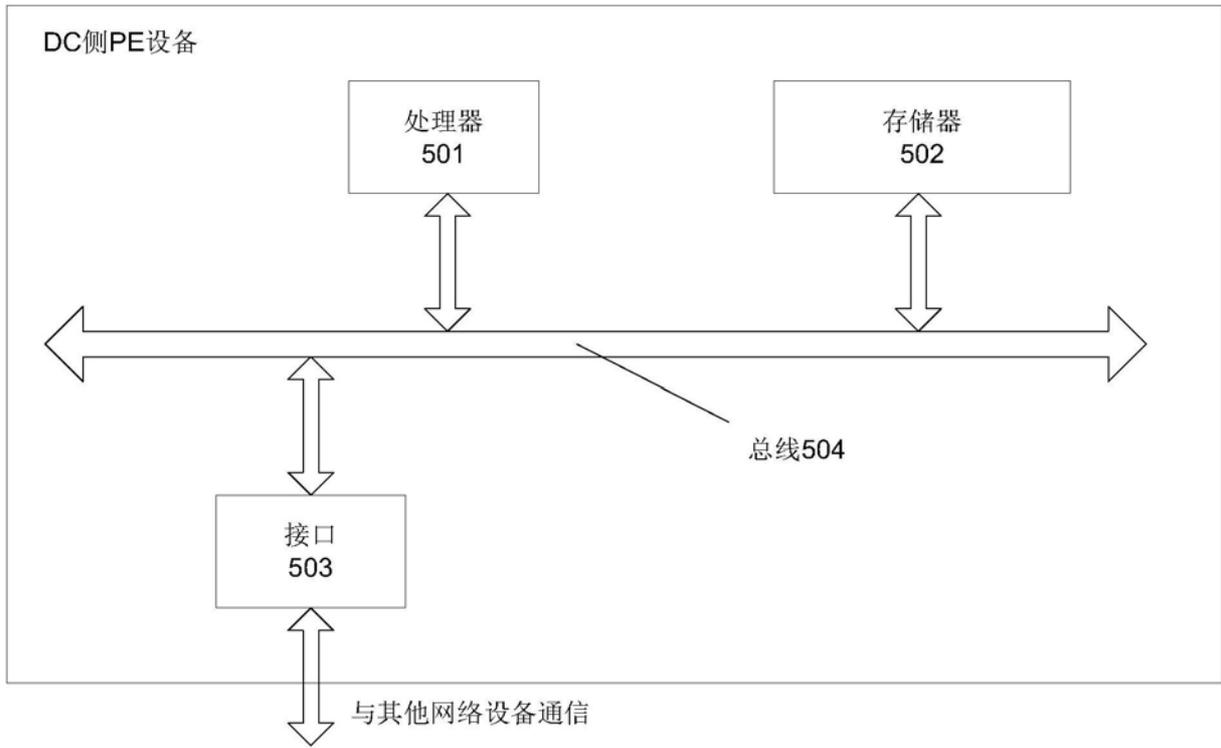


图5

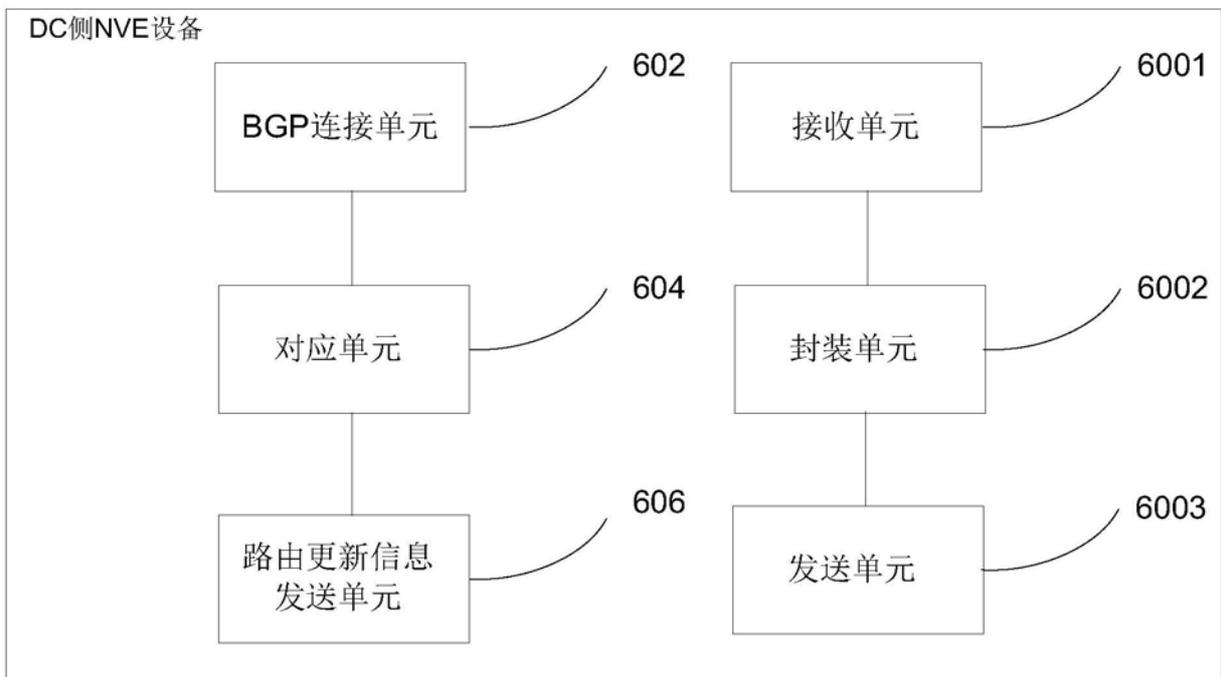


图6

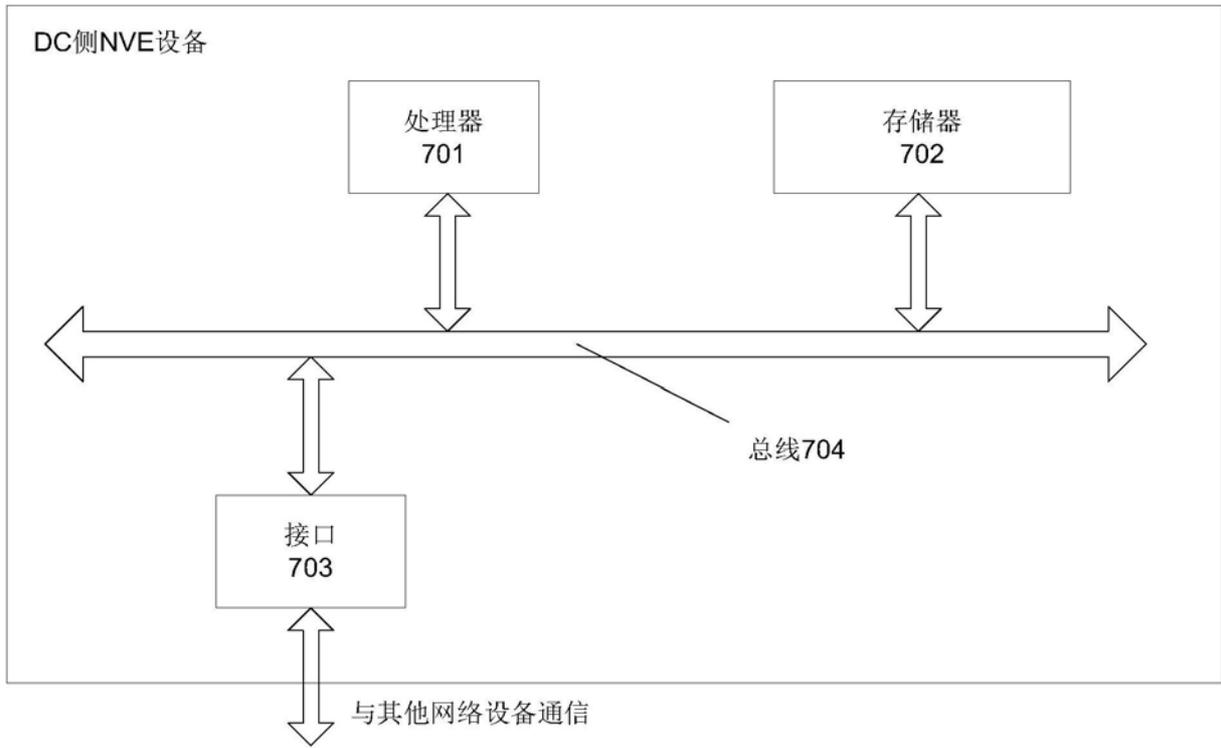


图7

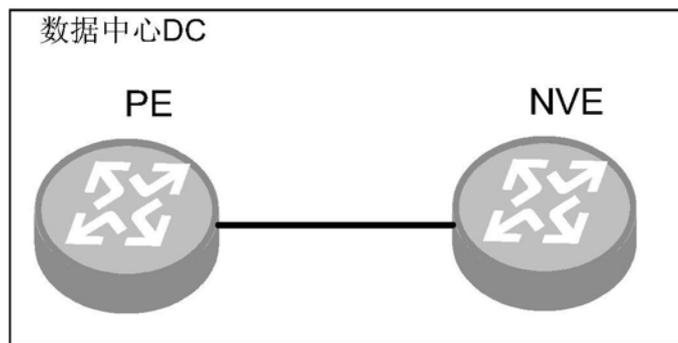


图8