

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2018 年 9 月 20 日 (20.09.2018)



(10) 国际公布号
WO 2018/166233 A1

(51) 国际专利分类号:
H04L 12/46 (2006.01)

(21) 国际申请号: PCT/CN2017/111608

(22) 国际申请日: 2017 年 11 月 17 日 (17.11.2017)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201710151473.5 2017年3月14日 (14.03.2017) CN

(71) 申请人: 华为技术有限公司 (HUAWEI TECHNOLOGIES CO., LTD.) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。

(72) 发明人: 王海波 (WANG, Haibo); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 陈大鹏 (CHEN, Dapeng); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 庄顺万 (ZHANG, Shunwan); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,

(54) Title: ROUTING PROCESSING METHOD, DEVICE AND SYSTEM

(54) 发明名称: 一种处理路由的方法、设备及系统

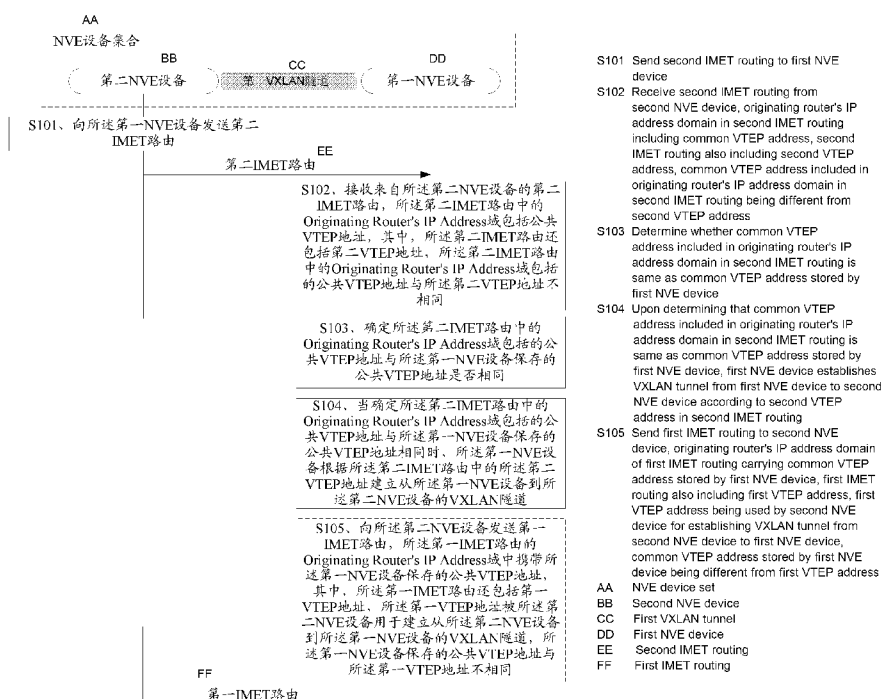


图2

(57) Abstract: A routing processing method, device and system. In the method, an NVE device set consisting of at least two NVE devices includes a common VTEP address, the common VTEP address being used for indicating a common VTEP and being deployed on each NVE device in the NVE device set. Each NVE device in the NVE device set also includes a VTEP address, the VTEP address being used for indicating a VTEP included in the corresponding NVE device. The NVE devices in the NVE device set establish VXLAN tunnels between the NVE devices in the NVE device set according to the VTEP addresses and the VTEP common address by means

SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG,
US, UZ, VC, VN, ZA, ZM, ZW。

- (84) 指定国 (除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

- 包括国际检索报告 (条约第21条(3))。

of IMET routing. The present invention thus helps to expand EVPN application scenarios. In addition, the present invention helps to increase reliability and reduce deployment complexity.

(57) 摘要: 一种处理路由的方法、设备及系统。所述方法包括, 由至少两个NVE设备组成的NVE设备集合包括公共VTEP地址, 所述公共VTEP地址用于标识公共VTEP, 并且被部署在所述NVE设备集合中的每个NVE设备上。而且, 所述NVE设备集合中的每个NVE设备还包括VTEP地址, 所述VTEP地址用于标识对应的NVE设备包括的VTEP。所述NVE设备集合中的NVE设备根据VTEP地址和公共VTEP地址, 通过IMET路由, 建立所述NVE设备集合中的NVE设备之间的VXLAN隧道。从而, 有助于扩大EVPN的应用场景。另外, 有助于提高可靠性和降低部署复杂程度。

一种处理路由的方法、设备及系统

本申请要求于2017年03月14日提交中国专利局、申请号为201710151473.5、发明名称为“一种处理路由的方法、设备及系统”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

5 技术领域

本申请涉及通信技术领域，尤其涉及一种处理路由的方法、设备及系统。更具体的，本申请涉及具有以太网虚拟私有网络（英文：Ethernet Virtual Private Network，缩写：EVPN）控制平面的虚拟扩展局域网（英文：Virtual Extensible Local Area Network，缩写：VXLAN）技术。

10 背景技术

VXLAN是一种将二层报文用三层协议进行封装的技术。VXLAN技术涉及MAC-in-UDP格式的报文。具体地，将基于媒体接入控制（英文：Media Access Control，缩写：MAC）协议的以太网帧封装在用户数据报协议（英文：User Datagram Protocol，缩写：UDP）报文中。进一步地，将UDP报文封装在因特网协议（英文：Internet Protocol，缩写：IP）报文中，可以在三层网络中传输IP报文。因此，实现了以太网帧在三层网络中传送。VXLAN技术使用VXLAN网络标识符（英文：VXLAN Network Identifier，缩写：VNI）标识VXLAN段（英文：VXLAN segment）。不同的VXLAN段分别对应不同的VNI。不同的VXLAN段之间是隔离的。同一个VNI内的两台虚拟机（英文：Virtual Machine，缩写：VM）进行通信时不需要经由VXLAN三层网关（英文：VXLAN L3 Gateway）。分别位于不同VNI中的两台VM通信需要经由VXLAN三层网关。VNI字段包含24比特。一个管理域最多可以包含 2^{16} 个VXLAN段。虚拟扩展局域网隧道端点（英文：VXLAN Tunnel End Point，缩写：VTEP）可以被集成在网络虚拟边缘（英文：Network Virtualization Edge，缩写：NVE）设备中，作为VXLAN中的边缘设备。NVE设备通过VXLAN隧道传输VXLAN的流量。VXLAN隧道是指两台NVE设备之间的点到点逻辑隧道。

EVPN是一种二层虚拟私有网络（英文：Virtual Private Network，缩写：VPN）技术。EVPN通过跨IP/多协议标签交换（英文：Multiprotocol Label Switching，缩写：MPLS）承载网将不同地域的客户站点（英文：customer site）连接起来，相当于这些客户站点位于同一个局域网（英文：Local Area Network，缩写：LAN）。

在应用场景中，VXLAN没有控制平面。VXLAN基于组播的洪泛方式，通过数据平面进行VTEP发现和远端主机地址学习，从而，导致数据中心网络中存在很多泛洪流量。基于边界网关协议（英文：Border Gateway Protocol，缩写：BGP）和MPLS的EVPN可以用于实现VXLAN的控制平面。在本文中，将这种具有EVPN控制平面的VXLAN场景命名为EVPN VXLAN。

在EVPN VXLAN应用场景中，两台NVE设备之间通过物理链路直接连接。例如，两台NVE设备之间通过PeerLink链路连接。但是，两台NVE设备之间通过物理链路直接连接的方式限制了EVPN的应用场景，且存在可靠性低和部署复杂的问题。

发明内容

有鉴于此，本申请实施例提供了一种处理路由的方法、设备及系统，应用在EVPN VXLAN应用场景中。EVPN的应用场景没有限制在NVE设备之间的链路需采用物理直

连链路的约束条件,有助于扩大EVPN的应用场景。另外,有助于提高可靠性和降低部署复杂程度。

本申请实施例提供的技术方案如下。

第一方面,提供了一种处理路由的方法。所述方法应用于VXLAN中,所述VXLAN
5 包括第一NVE设备和第二NVE设备。其中,所述第一NVE设备接收来自所述第二NVE设备的第二集成组播以太网标签(英文: Inclusive Multicast Ethernet Tag, 缩写: IMET)路由。所述第二IMET路由中的Originating Router's IP Address域包括公共VTEP地址,并且,所述第二IMET路由还包括第二VTEP地址。所述第二IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与
10 所述第二VTEP地址不相同。然后,所述第一NVE设备确定所述第二IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与
所述第一NVE设备保存的公共VTEP地址是否相同。当所述第一NVE设备确定所述第二IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与
所述第一NVE设备保存的公共VTEP地址相同时,所述第一NVE设备根据所述第二IMET路由中的所述第二VTEP地址建立从所述第一NVE设备到所述第二NVE设备的
15 VXLAN隧道。

基于实施例提供的方案,通过扩展IMET路由携带与公共VTEP地址不同的VTEP地址,以便NVE设备收到该扩展IMET路由后,根据VTEP地址建立该NVE设备到该NVE设备的BGP peer的VXLAN隧道。实现EVPN的应用场景没有限制在NVE设备之间的链路需采用物理直连链路的约束条件,有助于扩大EVPN的应用场景。另外,有助于提高
20 可靠性和降低部署复杂程度。

在第一方面的一种可能的实现方式中,所述方法还包括,所述第一NVE设备向所述第二NVE设备发送第一IMET路由。其中,所述第一IMET路由的Originating Router's IP Address域中携带所述第一NVE设备保存的公共VTEP地址,并且,所述第一IMET路由还包括第一VTEP地址。所述
25 第一VTEP地址被所述第二NVE设备用于建立从所述第二NVE设备到所述第一NVE设备的VXLAN隧道。所述第一NVE设备保存的公共VTEP地址与
所述第一VTEP地址不相同。

在第一方面的再一种可能的实现方式中,所述VXLAN还包括运营商边缘(英文: Provider Edge, 缩写: PE)设备。所述方法还包括,所述第一NVE设备向所述PE设备发送所述
30 第一IMET路由。所述第一NVE设备接收来自所述PE设备的IMET路由,所述来自所述PE设备的IMET路由中的Originating Router's IP Address域携带所述PE设备的VTEP地址。所述第一NVE设备根据所述来自所述PE设备的IMET路由中的所述PE设备的VTEP地址建立从所述第一NVE设备到所述PE设备的VXLAN隧道,所述PE设备的VTEP地址不等于所述第一NVE设备保存的公共VTEP地址。

在第一方面的又一种可能的实现方式中,所述VXLAN还包括第三NVE设备,所述
35 方法还包括,所述第一NVE设备接收来自所述第三NVE设备的第三IMET路由。其中,所述第三IMET路由中的Originating Router's IP Address域包括公共VTEP地址,并且,所述第三IMET路由还包括第三VTEP地址。所述第三IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与
所述第三VTEP地址不相同。所述第一NVE设备确定所述第三IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与
40 所述第一NVE设备保存的公共VTEP地址是否相同。当所述第一NVE设备确定所述第三

IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与所述第一NVE设备保存的公共VTEP地址相同时, 所述第一NVE设备根据所述第三IMET路由中的所述第三VTEP地址建立从所述第一NVE设备到所述第三NVE设备的VXLAN隧道。

可选的, 用户边缘(英文: Customer Edge, 缩写: CE)设备经由多个以太网链路多归属到所述第一NVE设备、所述第二NVE设备以及所述第三NVE设备, 所述多个以太网链路构成一个以太网段(英文: Ethernet segment, 缩写: ES)。所述方法还包括, 所述第一NVE设备根据来自所述第二NVE设备和所述第三NVE设备的ES路由, 将所述第一NVE设备确定为所述ES中的指定转发者(英文: Designated Forwarder, 缩写: DF)。

基于上述实现方式, 可以实现CE设备多归属到多于两台NVE设备的场景, 支持更多NVE设备的多活冗余模式, 有助于增加EVPN VXLAN的可靠性和增加网络部署的灵活性。

可选的, 所述方法还包括, 所述第一NVE设备经由所述第二NVE设备到所述第一NVE设备的VXLAN隧道接收所述第二NVE设备发送的广播、未知单播和组播BUM流量。所述第一NVE设备基于所述BUM流量是所述第一NVE设备经由所述第二NVE设备到所述第一NVE设备的VXLAN隧道接收的, 避免经由所述第一NVE设备到所述第三NVE设备的VXLAN隧道向所述第三NVE设备转发所述BUM流量。

基于上述实现方式, 利用NVE设备之间的VXLAN隧道不仅可以实现BUM流量的正常转发, 还可以避免BUM流量向其他NVE设备之间的VXLAN隧道的转发, 从而避免BUM流量的双发。

在第一方面的又一种可能的实现方式中, CE设备经由多个以太网链路双归属到所述第一NVE设备和所述第二NVE设备, 所述多个以太网链路构成一个ES。所述方法还包括, 所述第一NVE设备根据来自所述第二NVE设备的ES路由, 将所述第一NVE设备确定为所述ES中的DF。所述第一NVE设备向所述第二NVE设备发送Ethernet A-D per ES路由, 所述Ethernet A-D per ES路由携带所述第一NVE设备为所述CE设备与所述第一NVE设备之间的以太网链路分配的以太网段标识(英文: Ethernet Segment Identifier, 缩写: ESI)标签。所述第一NVE设备经由所述第二NVE设备到所述第一NVE设备的VXLAN隧道接收所述第二NVE设备发送的VXLAN报文, 所述VXLAN报文中的原始以太网载荷包括BUM流量, 所述VXLAN报文携带所述ESI标签, 其中, 所述BUM流量来自所述CE设备。所述第一NVE设备基于所述VXLAN报文携带所述ESI标签, 避免经由所述CE设备与所述第一NVE设备之间的ES向所述CE设备转发所述BUM流量。

可选的, 所述ESI标签被封装于所述VXLAN报文中的VXLAN头和所述原始以太网载荷之间, 所述VXLAN头携带指示信息, 所述指示信息用于指示所述VXLAN报文携带所述ESI标签。

基于上述实现方式, 通过扩展VXLAN报文的格式, 以便VXLAN报文能够携带ESI标签, 从而解决转发BUM流量过程中产生的环路问题。

在第一方面的又一种可能的实现方式中, CE设备经由多个以太网链路双归属到所述第一NVE设备和所述第二NVE设备, 所述多个以太网链路构成一个ES。所述方法还包括, 所述第一NVE设备接收来自所述第二NVE设备的MAC/IP Advertisement路由, 所述MAC/IP Advertisement路由包括MAC-VLAN ID。其中, 所述MAC-VLAN ID用于指示所述MAC/IP Advertisement路由携带的MAC地址所属的VLAN。所述MAC/IP

Advertisement路由携带的MAC地址是所述CE设备的MAC地址，或者所述MAC/IP Advertisement路由携带的MAC地址是所述CE设备所管辖的以太网虚拟私有网络站点EVPN site中的主机的MAC地址。所述第一NVE设备根据所述MAC/IP Advertisement路由携带的ESI和所述MAC-VLAN ID确定连接所述CE设备与所述第一NVE设备之间的以太网链路的所述第一NVE的本地接口。其中，所述MAC/IP Advertisement路由携带的ESI用于指示所述CE设备与所述第二NVE设备之间的所述ES。所述第一NVE设备的控制平面向所述第一NVE设备的转发平面发送第一MAC表项，所述第一MAC表项中的出接口为所述第一NVE设备的本地接口，其中，所述第一MAC表项包括所述MAC/IP Advertisement路由携带的MAC地址和所述MAC-VLAN ID。

可选的，所述第一NVE设备经由所述第二NVE设备到所述第一NVE设备的VXLAN隧道接收来自所述第二NVE设备的单播流量，所述单播流量是所述第二NVE设备基于所述第二NVE设备与所述CE设备之间的以太网链路故障发送的。所述第一NVE设备根据所述第一MAC表项经由所述第一NVE设备与所述CE设备之间的以太网链路向所述CE设备发送所述单播流量。

基于上述实现方式，可以通过MAC/IP Advertisement路由和NVE设备之间的VXLAN隧道实现转发路径的重新定向，在某个以太网链路出现链路故障时，依然可以确保单播流量的正常转发。

可选的，所述MAC/IP Advertisement路由被携带在MP_REACH_NLRI中，所述MP_REACH_NLRI中的下一跳域包括公共VTEP地址。所述MAC/IP Advertisement路由还包括所述第二VTEP地址。所述MP_REACH_NLRI中的下一跳域包括的公共VTEP地址与所述第二IMET路由中的Originating Router's IP Address域包括的公共VTEP地址相同。所述方法还包括，所述第一NVE设备的控制平面根据所述MP_REACH_NLRI中的下一跳域包括的公共VTEP地址和所述MAC/IP Advertisement路由包括的所述第二VTEP地址，向所述第一NVE设备的转发平面发送第二MAC表项。所述第二MAC表项中的出接口为连接所述第一NVE设备到所述第二NVE设备的VXLAN隧道的所述第一NVE设备的本地接口。其中，所述第二MAC表项包括所述MAC/IP Advertisement路由携带的MAC地址。

可选的，所述第一NVE设备接收从PE设备发往所述CE设备的单播流量。所述第一NVE设备根据所述第二MAC表项，经由所述第一NVE设备到所述第二NVE设备的VXLAN隧道发送所述单播流量，以便所述第二NVE设备根据所述第二NVE设备存储的MAC表项，将所述单播流量经由所述第二NVE设备与所述CE设备之间的以太网链路转发到所述CE设备。

基于上述实现方式，可以进一步实现单播流量的负载分担，有助于提高单播流量的传输带宽。

在第一方面的又一种可能的实现方式中，CE设备连接所述第二NVE设备。所述方法还包括，所述第一NVE设备接收来自所述第二NVE设备MAC/IP Advertisement路由，所述MAC/IP Advertisement路由被携带在MP_REACH_NLRI中，所述MP_REACH_NLRI中的下一跳域包括公共VTEP地址，所述MAC/IP Advertisement路由包括MAC-VLAN ID和所述第二VTEP地址，其中，所述MAC-VLAN ID用于指示所述MAC/IP Advertisement路由携带的MAC地址所属的VLAN，所述MAC/IP Advertisement

路由携带的MAC地址是所述CE设备的MAC地址,或者所述MAC/IP Advertisement路由携带的MAC地址是所述CE设备所管辖的EVPN site中的主机的MAC地址。所述第一NVE设备根据所述MAC/IP Advertisement路由携带的ESI和所述MAC-VLAN ID确定所述第一NVE设备中不存在连接所述CE设备与所述第一NVE设备之间的以太网链路的本地接口,其中,所述MAC/IP Advertisement路由携带的ESI用于指示所述CE设备与所述第二NVE设备之间的ES。所述第一NVE设备的控制平面根据所述MP_REACH_NLRI中的下一跳域包括的公共VTEP地址和所述MAC/IP Advertisement路由包括的所述第二VTEP地址,向所述第一NVE设备的转发平面发送MAC表项,所述MAC表项的出接口为连接所述第一NVE设备到所述第二NVE设备的VXLAN隧道的所述第一NVE设备的本地接口,其中,所述MAC表项包括所述MAC/IP Advertisement路由携带的MAC地址。

可选的,所述第一NVE设备接收从PE设备发往所述CE设备的单播流量。所述第一NVE设备根据所述MAC表项,经由所述第一NVE设备到所述第二NVE设备的VXLAN隧道发送所述单播流量,以便所述第二NVE设备根据所述第二NVE设备存储的MAC表项,将所述单播流量经由所述第二NVE设备与所述CE设备之间的以太网链路转发到所述CE设备。

基于上述实施方式,可以确保单归属连接到NVE设备的CE设备能够正常接收到单播流量。

在第一方面中,可选的,所述第一IMET路由中的VXLAN属性携带所述第一VTEP地址。

在第一方面中,可选的,所述第二IMET路由中的VXLAN属性携带所述第二VTEP地址。

在第一方面中,可选的,所述第一IMET路由包括VTEP VNI属性,所述VTEP VNI属性携带VTEP VNI,所述VTEP VNI与所述第一VTEP地址绑定,所述VTEP VNI用于指示所述第一NVE设备所属的BD。

在第一方面中,可选的,所述第二IMET路由包括VTEP VNI属性,所述VTEP VNI属性携带VTEP VNI,所述VTEP VNI与所述第二VTEP地址绑定,所述VTEP VNI用于指示所述第二NVE设备所属的BD。

第二方面,提供了第一NVE设备,所述第一NVE设备具有实现上述方法中第一NVE设备行为的功能。所述功能可以基于硬件实现,也可以基于硬件执行相应的软件实现。所述硬件或软件包括一个或多个与上述功能相对应的模块。

在一个可能的设计中,第一NVE设备的结构包括处理器和接口,所述处理器被配置为支持第一NVE设备执行上述方法中相应的功能。所述接口用于支持第一NVE设备与第一NVE设备的BGP peer之间的通信,向第一NVE设备的BGP peer发送上述方法中所涉及的信息或者指令,或者从第一NVE设备的BGP peer接收上述方法中所涉及的信息或者指令。所述第一NVE设备还可以包括存储器,所述存储器用于与处理器耦合,其保存第一NVE设备必要的程序指令和数据。

在另一个可能的设计中,所述第一NVE设备包括:处理器、发送器、随机存取存储器、只读存储器以及总线。其中,处理器通过总线分别耦接发送器、随机存取存储器以及只读存储器。其中,当需要运行第一NVE设备时,通过固化在只读存储器中的基本输入/输出系统或者嵌入式系统中的bootloader引导系统进行启动,引导第一NVE

设备进入正常运行状态。在第一NVE设备进入正常运行状态后，在随机存取存储器中运行应用程序和操作系统，使得该处理器执行第一方面或第一方面的任意可能的实现方式中的方法。

第三方面，提供一种第一NVE设备，所述第一NVE设备包括：主控板和接口板，进一步，还可以包括交换网板。所述第一NVE设备用于执行第一方面或第一方面的任意可能的实现方式中的方法。具体地，所述第一NVE设备包括用于执行第一方面或第一方面的任意可能的实现方式中的方法的模块。

第四方面，提供一种第一NVE设备，所述第一NVE设备包括控制器和第一NVE转发设备。所述第一NVE转发设备包括：接口板，进一步，还可以包括交换网板。所述第一NVE设备用于执行第三方面中的接口板的功能，进一步，还可以执行第三方面中交换网板的功能。所述控制器包括接收器、处理器、发送器、随机存取存储器、只读存储器以及总线。其中，处理器通过总线分别耦接接收器、发送器、随机存取存储器以及只读存储器。其中，当需要运行控制器时，通过固化在只读存储器中的基本输入/输出系统或者嵌入式系统中的bootloader引导系统进行启动，引导控制器进入正常运行状态。在控制器进入正常运行状态后，在随机存取存储器中运行应用程序和操作系统，使得该处理器执行第三方面中主控板的功能。

第五方面，提供一种处理路由的网络系统，所述网络系统包括至少两台NVE设备，所述至少两台NVE设备中的每台NVE设备为前述第二方面或第三方面或第四方面中的第一NVE设备。

第六方面，提供了一种计算机存储介质，用于储存为上述第一NVE设备所用的程序、代码或指令，当处理器或硬件设备执行这些程序、代码或指令时可以完成上述方面中第一NVE设备的功能或步骤。

通过上述方案，本申请实施例提供的用于处理路由的方法、设备及系统，在EVPN VXLAN应用场景中，由至少两个NVE设备组成的NVE设备集合包括公共VTEP地址，所述公共VTEP地址用于标识公共VTEP，并且所述公共VTEP被部署在所述NVE设备集合中的每个NVE设备上。而且，所述NVE设备集合中的每个NVE设备还包括VTEP地址，所述VTEP地址用于标识对应的NVE设备包括的VTEP。所述NVE设备集合中的NVE设备根据VTEP地址和公共VTEP地址，通过IMET路由，建立从所述NVE设备到所述NVE设备的BGP peer的VXLAN隧道。从而，在EVPN VXLAN应用场景中，EVPN的应用场景没有限制在NVE设备之间的链路需采用物理直连链路的约束条件，有助于扩大EVPN的应用场景。另外，有助于提高可靠性和降低部署复杂程度。

附图说明

图1为本申请实施例的一种EVPN VXLAN结构示意图；

图2为本申请实施例的一种处理路由的方法流程图；

图3为本申请实施例的另一种EVPN VXLAN结构示意图；

图4为本申请实施例的又一种EVPN VXLAN结构示意图；

图5为本申请实施例的VXLAN报文的格式示意图

图6为本申请实施例的VXLAN属性的格式示意图；

图7为本申请实施例的VTEP VNI属性的格式示意图；

图8为本申请实施例的MAC-VLAN属性的格式示意图；

图9为本申请实施例的又一种EVPN VXLAN结构示意图;

图10为本申请实施例的又一种EVPN VXLAN结构示意图;

图11为本申请实施例的第一NVE设备的结构示意图;

图12为本申请实施例的第一NVE设备的硬件结构示意图;

5 图13为本申请实施例的另第一NVE设备的硬件结构示意图;

图14为本申请实施例的又第一NVE设备的硬件结构示意图。

具体实施方式

本申请实施例提供了一种处理路由的方法、设备及系统, 应用在EVPN VXLAN应用场景中。EVPN的应用场景没有限制在NVE设备之间的链路需采用物理直连链路的约束条件, 有助于扩大EVPN的应用场景。另外, 有助于提高可靠性和降低部署复杂程度。

下面通过具体实施例, 分别进行详细的说明。

本申请所涉及的EVPN技术可以参见因特网工程任务组(英文: Internet Engineering Task Force, 缩写: IETF)请求注解(英文: Request For Comments, 缩写: RFC) 7432的说明, 所述RFC7432以全文引用的方式并入本申请中。

15 在本申请实施例中, “设备包括VTEP地址”的表述或“设备中的VTEP地址”的表述是指设备中保存有VTEP地址。“设备包括公共VTEP地址”的表述或“设备中的公共VTEP地址”的表述是指设备中保存有公共VTEP地址。其中, 所述设备可以是NVE设备、运营商边缘(英文: Provider Edge, 缩写: PE)设备或VTEP。

图1为本申请实施例的一种EVPN VXLAN结构示意图。EVPN VXLAN是指具有EVPN控制平面的VXLAN。该EVPN VXLAN包括NVE设备集合, 所述NVE设备集合至少包括两台NVE设备。如图1所示, 在NVE设备集合中, 包括第一NVE设备、第二NVE设备和第三NVE设备。应当理解, 本申请中, NVE设备集合中包括的NVE设备的数量不受限制, 可以为2台或者3台或者更多数量。所述NVE设备集合中的任意两台NVE设备是一对BGP对等体(英文: BGP peer)。在EVPN场景中, BGP peer也可以被称为EVPN peer。以图1为例, 所述第一NVE设备与所述第二NVE设备是一对BGP peer, 所述第一NVE设备与所述第三NVE设备是一对BGP peer, 所述第二NVE设备与所述第三NVE设备是一对BGP peer。其中, “一对BGP peer”可以被理解为: 一个设备是另一个设备的BGP peer。例如, 所述第一NVE设备与所述第二NVE设备是一对BGP peer可以被理解为指所述第一NVE设备是所述第二NVE设备的BGP peer, 或者被理解为所述第二NVE设备是所述第一NVE设备的BGP peer。所述BGP peer也可以被称为BGP邻居; 相应的, EVPN peer也可以被称为EVPN邻居。在本申请中, 为了说明方便, 后续实施例中统一使用BGP peer。所述BGP peer通过BGP中规定的OPEN消息建立, 并通过KEEPALIVE消息维持建立的BGP peer。所述OPEN消息和KEEPALIVE消息的实现可以参见IETF RFC2858和IETF RFC1771的相关说明。另外, 建立BGP peer的两端设备中可以部署路由反射器(英文: Route Reflector, 缩写: RR), 从而利用RR完成BGP peer的建立。NVE设备集合中包括的NVE设备也可以称为PE设备, 本申请在未加特殊说明的情况下, NVE设备等同于PE设备。本申请的描述中, 将NVE设备集合中包括的设备命名为NVE设备。实际上, NVE设备集合中包括的NVE设备也可以被称为PE设备。相应的, 所述NVE设备集合也可以被称为PE设备集合。

40 所述EVPN VXLAN还包括PE设备, 如图1所示, 所述PE设备位于所述NVE设备集

合之外。所述PE设备可以与所述NVE设备集合中的每台NVE设备组成BGP peer。以图1为例，所述PE设备分别是所述第一NVE设备、所述第二NVE设备和所述第三NVE设备的BGP peer。所述PE设备可以是一台物理设备，也可以是由多台物理设备组成的逻辑设备。而且，所述PE设备还可以连接其他网络设备，例如所述PE设备连接一台或多台用户边缘（英文：Customer Edge，缩写：CE）设备。

所述EVPN VXLAN还包括至少一台CE设备。图1示例性的示出了三台CE设备，分别为第一CE设备，第二CE设备和第三CE设备。其中，所述第一CE设备多归属到所述第一NVE设备、所述第二NVE设备和所述第三NVE设备，即，所述第一CE设备分别与所述第一NVE设备、所述第二NVE设备和所述第三NVE设备通信。所述第二CE设备双归属到所述第一NVE设备和所述第二NVE设备，即，所述第二CE设备分别与所述第一NVE设备和所述第二NVE设备通信。所述第三CE设备单归属到所述第二NVE设备，即，所述第三CE设备与所述第二NVE设备通信。另外，应当理解，多归属可以是双归属。本申请为了区别CE设备连接NVE设备的数量，因此使用两种表述方式并存，如上文所示。

在所述EVPN VXLAN中，CE设备与NVE设备之间通过以太网链路连接。而且，与同一台CE设备连接的所有以太网链路构成一个以太网段（英文：Ethernet segment，缩写：ES）。以图1为例，第一CE设备经由3个以太网链路多归属到所述第一NVE设备、所述第二NVE设备以及所述第三NVE设备，所述3个以太网链路构成一个ES，图1中使用ES1表示。以太网段标识（英文：Ethernet Segment Identifier，缩写：ESI）用于标识相对应的ES。以图1为例，所述第一CE设备连接所述第一NVE设备、所述第二NVE设备和所述第三NVE设备的3个以太网链路的ES1的ESI值为同一个值。所述第二CE设备连接所述第一NVE设备和所述第二NVE设备的2个以太网链路的ES2的ESI值为同一个值。其中，ES1的ESI值与ES2的ESI值不相等，且均为非零值；ES3的ESI值为零。ESI包括类型（英文：Type）域和ESI值域，其中Type域用于指示ESI的生成方式。常用的两种生成方式是Type0和Type1，其中Type0表示通过手工配置生成，Type1表示通过NVE和CE之间运行的链路聚合控制协议（英文：Link Aggregation Control Protocol，缩写：LACP）生成，所述ESI值域的取值范围为0至0xFF，其中“0x”表示16进制。ES和ESI的生成及设置可以参见RFC7432中的第5章的说明。

在图1所示的EVPN VXLAN中，NVE设备和PE设备可以是路由器或三层交换机。CE设备可以是路由器或交换机或主机。本申请实施例中的NVE设备、PE设备和CE设备是RFC7432中定义的相应设备。当所述CE设备是路由器或交换机时，可以连接一台或多台主机。其中，主机可以是物理设备或VM。图1所示的EVPN VXLAN可以应用于多种场景中。例如，所述EVPN VXLAN应用于移动承载网（英文：Mobile Bearer Network），典型的移动承载网是互联网协议化无线接入网（英文：Internet Protocol Radio Access Network，缩写：IP RAN）。在移动承载网中，所述第一CE设备、第二CE设备和第三CE设备可以是基站（英文：Base Transceiver Station，缩写：BTS），所述PE设备可以连接基站控制器（英文：Base Station Controller，缩写：BSC）或无线网络控制器（英文：Radio Network Controller，缩写：RNC），所述第一NVE设备、所述第二NVE设备和所述第三NVE设备可以是基站侧网关（英文：Cell Site Gateway，缩写：CSG），所述PE设备可以是基站控制器侧网关（英文：Radio Network Controller Site Gateway，

缩写：RSG）。又例如，所述EVPN VXLAN应用于固网（英文：Fixed Network）。在固网中，所述第一CE设备、第二CE设备和第三CE设备可以是用户侧的站点，所述第一NVE设备、所述第二NVE设备和所述第三NVE设备可以是数字用户线接入复接器（英文：Digital Subscriber Line Access Multiplexer，缩写：DSLAM），所述PE设备
5 可以是宽带接入服务器（英文：Broadband Access Server，缩写：BAS）。

在图1所示的EVPN VXLAN中，所述PE设备与所述NVE设备集合之间建立VXLAN隧道。对所述PE设备而言，所述NVE设备集合就如同单台NVE设备，即所述NVE设备集合中的所有NVE设备被虚拟化为单台NVE设备。在一种可能的实现场景中，从所述PE设备被转发到所述第一NVE设备、所述第二NVE设备和所述第三NVE设备的实际流量承载在从所述PE设备到所述第一NVE设备、所述第二NVE设备和所述第三NVE设备的多条物理链路上（如图1中，所述PE设备分别连接所述第一NVE设备、所述第二NVE设备和所述第三NVE设备的虚线所示）。在另一种可能的实现场景中，所述PE设备和NVE设备集合之间包括路由器。其中，所述路由器执行IP转发，实现流量分流的功能。所述PE设备通过物理链路与所述路由器连接。所述路由器经由多条物理链路连接所述第一NVE设备、所述
10 第二NVE设备和所述第三NVE设备。在上述两种可能的实现场景中，从所述PE设备到所述NVE设备集合中的多个NVE设备方向的物理链路被虚拟化为单条VXLAN隧道。因此，所述PE设备在向所述NVE设备集合发送报文时，会将所述NVE设备集合看作成单台NVE设备。为了达到上述目的，所述NVE设备集合包括一个公共VTEP。所述公共VTEP包括公共VTEP地址。所述公共VTEP地址用于标识所述公共VTEP。所述公共VTEP地址可以是IP地址。并且，所述公共VTEP被部署在所述NVE设备集合中的每台NVE设备上。如此这样，所述NVE设备集合中的每台NVE设备包括所述公共VTEP。所述NVE设备集合中的NVE设备包括相同的VTEP，所述NVE设备集合中的NVE设备包括相同的VTEP地址。因此，本申请中包括相同VTEP的NVE设备组成的集合被称为NVE设备集合。所述PE设备在收到所述NVE设备集合中的每台NVE设备发送的集成组播以太网标签
20 （英文：Inclusive Multicast Ethernet Tag，缩写：IMET）路由后，确定IMET路由的发送端只存在一个VTEP地址（公共VTEP地址），因此，对所述PE设备而言，所述NVE设备集合就如同一台NVE设备。所述PE设备将公共VTEP作为VXLAN隧道的端点建立所述VXLAN隧道。以图1为例，所述NVE设备集合的公共VTEP地址的值为9.9.9.9。包括所述公共VTEP地址的公共VTEP被分别部署在所述第一NVE设备、所述第二NVE设备和所述第三NVE设备上。所述PE设备可以基于接收到的来自NVE设备的IMET路由，建立从所述PE设备到所述NVE设备集合的VXLAN隧道。建立VXLAN隧道的具体实现方式请参见本申请后续实施例的说明。另外，从所述PE设备角度看，NVE设备集合是一台虚拟的NVE设备。技术术语“NVE设备集合”中的“集合”的使用不构成对本申请的限制。例如，可以使用“组”代替“集合”并表达相同的含义，即，“NVE设备组”。
30

在所述NVE设备集合中，NVE设备与NVE设备之间建立VXLAN隧道。即，所述NVE设备集合中的每对BGP peer之间建立VXLAN隧道。以图1为例，所述第一NVE设备与所述第二NVE设备之间建立第一VXLAN隧道。所述第二NVE设备与所述第三NVE设备之间建立第二VXLAN隧道。所述第一NVE设备与所述第三NVE设备之间建立第三VXLAN隧道。为了达到上述目的，所述NVE设备集合中的每台NVE设备还包括一个VTEP，
40 所述VTEP包括VTEP地址。而且，该VTEP地址在EVPN VXLAN中是唯一的。VTEP地址

用于标识对应的NVE设备中的VTEP。所述VTEP地址可以是IP地址。进一步，所述NVE设备集合的公共VTEP地址与所述NVE设备集合中的任意两个NVE设备之间的VXLAN隧道的端点（英文end point）地址均不相同。以图1为例，所述第一NVE设备包括前述提及的公共VTEP，所述公共VTEP包括的公共VTEP地址的值为9.9.9.9，所述第一NVE设备还包括第一VTEP，所述第一VTEP包括的第一VTEP地址的值为1.1.1.1。所述第二NVE设备包括前述提及的公共VTEP，所述公共VTEP地址的值为9.9.9.9，所述第二NVE设备还包括第二VTEP，所述第二VTEP包括的第二VTEP地址的值为2.2.2.2。所述第三NVE设备包括前述提及的公共VTEP，所述公共VTEP地址的值为9.9.9.9，所述第三NVE设备还包括第三VTEP，所述第三VTEP包括的第三VTEP地址的值为3.3.3.3。所述NVE设备集合中的NVE设备可以根据VTEP地址和公共VTEP地址，通过IMET路由，建立从所述NVE设备到所述NVE设备的BGP peer的VXLAN隧道。具体建立VXLAN隧道的实现方式请参见本申请后续实施例的说明。

通过上述实施方式，在EVPN VXLAN应用场景中，由至少两个NVE设备组成的NVE设备集合包括公共VTEP地址，所述公共VTEP地址用于标识公共VTEP，并且所述公共VTEP被部署在所述NVE设备集合中的每个NVE设备上。而且，所述NVE设备集合中的每个NVE设备还包括VTEP地址，所述VTEP地址用于标识对应的NVE设备包括的VTEP。所述NVE设备集合中的NVE设备根据VTEP地址和公共VTEP地址，通过IMET路由，建立从所述NVE设备到所述NVE设备的BGP peer的VXLAN隧道。从而，在EVPN VXLAN的应用场景中，NVE设备集合中的NVE设备通过VXLAN隧道实现互联，使得NVE设备集合中的NVE设备不受直连物理链路的限制。使用VXLAN隧道实现NVE设备的互联，即使NVE设备集合中的NVE设备之间存在交换节点，也不会影响EVPN VXLAN的实现。在部署NVE设备集合中的NVE设备时，不会受到部署地域和距离的影响，提高了部署的灵活性。进一步，NVE设备集合中的NVE设备之间的链路不限于直连物理链路，有助于提高部署的灵活性。

图2为本申请实施例的一种处理路由的方法流程图。图2所示的方法可以应用于图1所示的EVPN VXLAN中。即，图2所示的方法可以应用于图1所示的具有EVPN控制平面的VXLAN中。所述EVPN VXLAN包括包括第一NVE设备和第二NVE设备。所述第一NVE设备与所述第二NVE设备是一对BGP peer。其中，建立BGP peer的实现方式可以参见前述实施例的说明，此处不进行赘述。图2示出了NVE设备集合，所述第一NVE设备与所述第二NVE设备被包括在所述NVE设备集合中。应当理解，在实际的应用场景中，从PE设备角度看，NVE设备集合是单台虚拟NVE设备。所述第一NVE设备与所述第二NVE设备均包括公共VTEP。为了便于表述，在本申请实施例中使用NVE设备集合的表述。图2示出了2台NVE设备，应当理解，在实际的应用场景中，所述NVE设备集合中的NVE设备的数量可以多于2。当所述NVE设备集合中的NVE设备的数量可以多于2时，所述NVE设备集合中的任意两台NVE设备均可实施图2所示的处理路由的方法。图2所示的方法包括S101至S104，进一步还可以包括S105。

S101、所述第二NVE设备向所述第一NVE设备发送第二IMET路由。

S102、所述第一NVE设备接收来自所述第二NVE设备的第二IMET路由，所述第二IMET路由中的发端路由器因特网协议地址（英文：Originating Router's IP Address）域包括公共VTEP地址，其中，所述第二IMET路由还包括第二VTEP地址，所述第二IMET路由中

的Originating Router's IP Address域包括的公共VTEP地址与所述第二VTEP地址不相同。

根据前述实施例针对图1所示场景的描述,所述NVE设备集合包括公共VTEP,所述公共VTEP包括公共VTEP地址,所述VTEP地址可以是IP地址。该公共VTEP被部署在所述NVE设备集合中的每个NVE设备上。结合图2,该公共VTEP被部署在所述第一NVE设备和所述第二NVE设备上。可选的,所述公共VTEP地址可以通过静态配置的方式部署在所述NVE设备集合中的每个NVE设备上的所述公共VTEP中。如此这样,基于所述公共VTEP地址,从PE设备角度看,所述NVE设备集合可以被虚拟为单台NVE设备。

所述NVE设备集合中的NVE设备还包括VTEP,所述VTEP包括VTEP地址,所述VTEP地址可以是IP地址。结合图2,所述第一NVE设备还包括第一VTEP,所述第一VTEP包括第一VTEP地址;所述第二NVE设备还包括第二VTEP,所述第二VTEP包括第二VTEP地址。所述第一VTEP地址和所述第二VTEP地址在EVPN VXLAN中具有唯一性,并且,所述第一VTEP地址的值和所述第二VTEP地址的值与所述公共VTEP地址的值不相等。基于上述,所述NVE设备集合中的NVE设备既包括公共VTEP地址,也包括自身的VTEP地址。作为举例,公共VTEP地址和VTEP地址被保存在NVE设备的存储器中,以便NVE设备的处理器访问。

所述第二NVE设备向所述第一NVE设备发送第二IMET路由。本申请实施例使用的IMET路由利用了RFC7432中提及的IMET路由,并且本申请实施例对RFC7432中提及的IMET路由进行了扩展。IMET路由的格式和功能可参见RFC7432的说明。具体的,所述第二NVE设备向所述第一NVE设备发送第二IMET路由可以按照以下方式实现。所述第二NVE设备可以向所述第一NVE设备发送更新消息(英文: UPDATE message),所述UPDATE message携带有多协议可到达网络层可达信息(英文: Multiprotocol Reachable NLRI,缩写: MP_REACH_NLRI)属性(英文: attribute),NLRI的全称是Network Layer Reachability Information。所述MP_REACH_NLRI属性包括EVPN NLRI域和下一跳(英文: next hop)域。所述EVPN NLRI域用于携带所述第二IMET路由。其中,所述UPDATE message和所述MP_REACH_NLRI属性的定义参见RFC4760中的说明,所述EVPN NLRI域的定义参见RFC7432的第7章的说明。

所述第二IMET路由包括Originating Router's IP Address域。在本申请实施例中,下一跳域的值和所述第二IMET路由中的Originating Router's IP Address的值均被设置为公共VTEP地址9.9.9.9。在所述第二IMET路由向所述NVE设备集合中的其他NVE设备和NVE设备集合以外的PE设备(如图1所示)发布后,确保所述第二NVE设备能够接收到所述PE设备和所述其他NVE设备传送报文。

在本申请实施例中,还对RFC7432中提及的IMET路由进行扩展,以便IMET路由能够携带发送IMET路由的NVE设备的VTEP地址。例如,可以将NVE设备的管理IP地址(或者称设备IP地址)作为VTEP地址。如此这样设置,所述第二IMET路由既携带了公共VTEP地址9.9.9.9,也携带了第二VTEP地址1.1.1.1。

作为一种优选的方案,本申请实施例扩展了一种新的属性,用来携带VTEP地址。如图6所示,本申请实施例扩展了VXLAN属性,用来携带VTEP地址。其中,类型域和子类型域用于表示VXLAN属性所属类型,具体数值可以根据标准组织的要求进行设定。标志域可以用于指示所述VXLAN属性是否允许被读取,例如,标志域为0表示,所述VXLAN属性中的VTEP地址不允许被读取,标志域为1表示,所述VXLAN属性中的

VTEP地址允许被读取。应当理解，上述标志域的功能是一种示例，标志域的功能可以根据使用场景进行相应的变换。VTEP地址域用于携带VTEP地址的值。预留域为了实现进一步的扩展。VXLAN属性可以作为一种属性被携带在IMET路由中，以便实现本申请实施例的方案。以第二IMET路由为例，所述第二IMET路由携带有VXLAN属性，

5 所述VXLAN属性包括第二VTEP地址。

所述第二IMET路由还可以携带VNI，所述VNI用于指示所述第二NVE设备所属的BD，更为具体的，所述VNI用于指示所述第二NVE设备发送所述第二IMET路由的端口所属的BD。在本申请中，BD可以被称为广播域（英文：Broadcast Domain，缩写：BD）或者桥域（英文：Bridge Domain，缩写：BD）。所述VNI可以被携带在所述第二IMET

10 路由的运营商组播业务接口（英文：Provider Multicast Service Interface，缩写：PMSI）隧道属性中。所述PMSI隧道属性包括MPLS标签域，用于携带所述VNI。

在本申请实施例中，所述VNI可以被配置为两种形式，即，公共VTEP VNI和VTEP VNI。其中所述公共VTEP VNI用于与公共VTEP地址绑定，所述VTEP VNI用于与VTEP地址绑定。结合图1，PE设备与所述NVE设备集合之间建立的VXLAN隧道的端点地址

15 使用了公共VTEP地址，因此，PE设备与所述NVE设备集合之间建立的VXLAN隧道对应公共VTEP VNI。所述NVE设备集合中的NVE设备之间建立的VXLAN隧道的端点地址使用了VTEP地址，因此，所述NVE设备之间建立的VXLAN隧道对应VTEP VNI。在本申请实施例中，公共VTEP VNI和VTEP VNI可以设置为相同，也可以设置为不同。

如果将公共VTEP VNI和VTEP VNI设置为相同，所述PE设备与所述NVE设备集合之间建立的VXLAN隧道，和所述NVE设备集合中的NVE设备之间建立的VXLAN隧道

20 对应同一个BD。在所述公共VTEP VNI和所述VTEP VNI设置为相同的情况下，可以按照上述实现方式携带VNI信息。即，在所述PMSI隧道属性包括的MPLS标签域中携带VNI，该VNI既用于指示所述公共VTEP VNI，又用于指示所述VTEP VNI。在建立所述PE设备与所述NVE设备集合之间的VXLAN隧道和建立所述NVE设备集合中的NVE设备之间的VXLAN隧道过程中，均使用所述PMSI隧道属性包括的MPLS标签域中携带的

25 VNI。

如果将公共VTEP VNI和VTEP VNI设置为不同，所述PE设备与所述NVE设备集合之间建立的VXLAN隧道，和所述NVE设备集合中的NVE设备之间建立的VXLAN隧道

30 对应不同的BD。例如，所述PE设备与所述NVE设备集合之间建立的VXLAN隧道对应BD1，所述NVE设备集合中的NVE设备之间建立的VXLAN隧道对应BD2。上述所说的“将公共VTEP VNI和VTEP VNI设置为不同”是指：所述NVE设备集合中的NVE设备之间建立的VXLAN隧道都使用同一个VTEP VNI，并且该VTEP VNI与公共VTEP VNI不相同。在所述公共VTEP VNI和所述VTEP VNI设置为不同的情况下，可以按照下面实现方式携带VNI信息。所述公共VTEP VNI被携带在所述PMSI隧道属性包括的MPLS

35 标签域中。而且，本申请实施例还扩展了一种新的属性，用来携带所述VTEP VNI。如图7所示，本申请实施例扩展了VTEP VNI属性，用来携带VTEP VNI。其中，类型域和子类型域用于表示VTEP VNI属性所属类型，具体数值可以根据标准组织的要求进行设定。标志域可以用于指示所述VTEP VNI属性是否允许被读取，例如，标志域为0表示，所述VTEP VNI属性中的VTEP VNI不允许被读取，标志域为1表示，所述VTEP VNI属性中的VTEP VNI允许被读取。应当理解，上述标志域的功能是一种示例，标志域的功

40

能可以根据使用场景进行相应的变换。VTEP VNI域用于携带VTEP VNI的值。预留域为了实现进一步的扩展。VTEP VNI属性可以作为一种属性被携带在IMET路由中。所述PE设备在收到NVE设备集合中的NVE设备发布的IMET路由时，根据公共VTEP VNI建立VXLAN隧道，并且不解析VTEP VNI属性，从而忽略VTEP VNI。相应的，NVE设备集合中的NVE设备在接收到其他NVE设备发布的IMET路由后，优先按照VTEP VNI属性中的VTEP VNI建立VXLAN隧道。如果NVE设备集合中的NVE设备确定接收到的IMET路由中不存在VTEP VNI属性，则使用所述PMSI隧道属性包括的MPLS标签域中携带的公共VTEP VNI建立相应的VXLAN隧道。

进一步，所述NVE设备集合中的NVE设备之间建立的VXLAN隧道还可以使用不同的VTEP VNI。例如图1中，第一VXLAN隧道使用VTEP VNI1，对应BD21，第二VXLAN隧道使用VTEP VNI2，对应BD22，第三VXLAN隧道使用VTEP VNI3，对应BD23。如此这样，允许所述NVE设备配置多个VNI，以便所述NVE设备可以建立对应不同BD的VXLAN隧道。在本申请实施例中，为了使技术方案更容易理解，将公共VTEP VNI和VTEP VNI设置为相同，并且在后续说明中可以使用VNI表示所述公共VTEP VNI和所述VTEP VNI。应当理解，本申请实施例不限于将公共VTEP VNI和VTEP VNI设置为相同，可以将公共VTEP VNI和VTEP VNI设置为不同，进一步还可以设定不同的VTEP VNI，结合上述说明实现本申请的实施例。

所述第二IMET路由中的路由标识（英文：Route Distinguisher，缩写：RD）域用于指示所述第二NVE设备对应的EVPN实例（英文：EVPN instance，缩写：EVI）。所述第二IMET路由中的以太网标签（英文：Ethernet tag）域可以被设置为所述第二NVE设备连接的CE设备所包括的MAC地址对应的虚拟局域网标识（英文：Virtual Local Area Network Identifier，缩写：VLAN ID或VID）或者所述第二NVE设备对应的BD。所述第二IMET路由中的IP地址长度（英文：IP Address Length）域用于指示Originating Router's IP Address的长度。

所述第二IMET路由的PMSI隧道属性还包括隧道类型（英文：Tunnle Type）域，所述隧道类型域用于指示建立隧道的类型。在本申请实施例中，将所述隧道类型域的值设定为6，用于标识建立的隧道类型为头端复制（英文：Ingress Replication）。所述第二IMET路由还包括VXLAN隧道属性，所述VXLAN隧道属性的格式与图6或图7所示的属性格式类似，区别在于：将图6中的VTEP地址域或图7中的VTEP VNI域替换为VXLAN隧道类型域。所述VXLAN隧道类型域用于指示建立的隧道为VXLAN隧道。所述VXLAN隧道类型域的取值具体由标准组织确定，例如所述VXLAN隧道类型域的值为8。所述第一NVE设备利用所述第二IMET路由建立VXLAN的实现方式请参见本实施例的后续说明。所述第二IMET路由的PMSI隧道属性还包括隧道标识（英文：Tunnle Identifier）域，在本申请实施例中，所述隧道标识域可以填写所述公共VTEP地址。

S103、所述第一NVE设备确定所述第二IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与所述第一NVE设备保存的公共VTEP地址是否相同。

其中，所述第一NVE设备能够向所述第一NVE设备的BGP peer发送Originating Router's IP Address域中携带所述第一NVE设备保存的公共VTEP地址的IMET路由。所述第二IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与所述第一NVE设备保存的公共VTEP地址相同。

由于所述第二NVE设备向外发布所述第二IMET路由，且所述第一NVE设备与所述第二NVE设备是一对BGP peer，因此，所述第一NVE设备能够接收到所述第二IMET路由。

所述第一NVE设备接收到所述第二IMET路由后，解析所述第二IMET路由，获取所述第二IMET路由中的Originating Router's IP Address域包括的公共VTEP地址。所述第一NVE设备确定所述第二IMET路由中的Originating Router's IP Address域包括的所述公共VTEP地址与所述第一NVE设备保存的公共VTEP地址是否相同。根据本申请实施例，所述第一NVE设备和所述第二NVE设备均包括公共VTEP，因此，所述第二IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与所述第一NVE设备保存的公共VTEP地址相同。其中，当所述第一NVE设备在向外发布IMET路由时，将所述第一NVE设备保存的公共VTEP地址携带在该IMET路由的Originating Router's IP Address域中。也就是说，所述第一NVE设备保存的公共VTEP地址是所述第一NVE设备包括的公共VTEP中的公共VTEP地址。所述第一NVE设备确定所述第二IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与所述第一NVE设备保存的公共VTEP地址相同，则所述第一NVE设备可以确定所述第一NVE设备与所述第二NVE设备属于同一个NVE设备集合。

S104、当所述第一NVE设备确定所述第二IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与所述第一NVE设备保存的公共VTEP地址相同时，所述第一NVE设备根据所述第二IMET路由中的所述第二VTEP地址建立从所述第一NVE设备到所述第二NVE设备的VXLAN隧道。

基于上述判断的结果，所述第一NVE设备就可以根据所述第二VTEP地址建立从所述第一NVE设备到所述第二NVE设备的VXLAN隧道。具体的，所述第一NVE设备继续解析所述第二IMET路由，获取所述第二IMET路由中的所述第二VTEP地址，并且，以所述第二VTEP地址作为VXLAN隧道的目的端点地址建立从所述第一NVE设备到所述第二NVE设备的VXLAN隧道。

在一种可能的实现方式中，结合S101和S102的说明，所述第二VTEP地址被携带在VXLAN属性中。基于上述判断的结果，即所述第一NVE设备确定所述第二IMET路由中的Originating Router's IP Address域包括的所述公共VTEP地址与所述第一NVE设备保存的公共VTEP地址相同后，所述第一NVE设备检查所述第二IMET路由中是否携带所述VXLAN属性。所述第一NVE设备确定所述第二IMET路由中携带有所述VXLAN属性，获取所述VXLAN属性中的所述第二VTEP地址，并且，以所述第二VTEP地址作为VXLAN隧道的目的端点地址建立从所述第一NVE设备到所述第二NVE设备的VXLAN隧道。

举例说明，所述第一NVE设备根据所述第二IMET路由的PMSI隧道属性中的隧道类型域的值为6，所述第一NVE设备确定建立的隧道类型为头端复制。所述第一NVE设备解析所述第二IMET路由的VXLAN隧道属性，获取所述VXLAN隧道属性中的VXLAN隧道类型域的值，从而确定建立的隧道为VXLAN隧道。所述第一NVE设备在确定建立的隧道为VXLAN隧道后，生成VXLAN隧道表和VXLAN隧道头端复制表。其中，所述VXLAN隧道表包括隧道标识、源端点地址和目的端点地址。所述隧道标识用于标识VXLAN隧道。所述源端点地址为所述第一NVE设备包括的第一VTEP地址，例如1.1.1.1。所述目的端点地址为从所述第二IMET路由中获取的第二VTEP地址，例如2.2.2.2。其中，所述VXLAN隧道头端复制表包括VNI和BGP peer地址。所述VNI可以根据前述说明，在PMSI隧道属性包括的MPLS标签域中获取或者在VTEP VNI属性中获取。所述BGP peer地址为从所述第二

IMET路由中获取的所述第二VTEP地址，例如2.2.2.2。从而，所述所述第一NVE设备建立从所述第一NVE设备到所述第二NVE设备的VXLAN隧道。

如果所述第一NVE设备确定接收到的IMET路由中的Originating Router's IP Address域包括的VTEP地址与所述第一NVE设备保存的公共VTEP地址不相同，则所述第一NVE设备可以确定所述第一NVE设备与发送该IMET路由的设备不属于同一个NVE设备集合。所述第一NVE设备将根据接收到的IMET路由中的Originating Router's IP Address域包括的VTEP地址建立从所述第一NVE设备到发送该IMET路由的设备的VXLAN隧道，而不是根据VXLAN属性中的VTEP地址建立VXLAN隧道。类似的，可以参见后续实施例说明的PE设备与NVE设备集合之间建立VXLAN隧道的过程。

如果所述第一NVE设备确定所述第二IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与所述第一NVE设备保存的公共VTEP地址相同，但无法进一步找到所述第二VTEP地址或所述VXLAN属性，所述第一NVE设备将丢弃所述第二IMET路由，并且向所述第二NVE设备返回错误消息。进一步可选的，所述第一NVE设备还可以向EVPN VXLAN的网络管理设备反馈所述错误消息。这样设置有助于提高VXLAN隧道建立的可靠性。

可选的，S105、所述第一NVE设备向所述第二NVE设备发送第一IMET路由，所述第一IMET路由的Originating Router's IP Address域中携带所述第一NVE设备保存的公共VTEP地址，其中，所述第一IMET路由还包括第一VTEP地址，所述第一VTEP地址被所述第二NVE设备用于建立从所述第二NVE设备到所述第一NVE设备的VXLAN隧道，所述第一NVE设备保存的公共VTEP地址与所述第一VTEP地址不相同。

所述第一NVE设备向第二NVE设备发送第一IMET路由的实现过程，以及所述第一IMET路由的实现方式可以参见本实施例前述S101和S102的说明，此处不再进行赘述。

所述第二NVE设备在接收所述第一IMET路由后，可以根据所述第一IMET路由建立从所述第二NVE设备到所述第一NVE设备的VXLAN隧道。所述第二NVE设备根据所述第一IMET路由建立从所述第二NVE设备到所述第一NVE设备的VXLAN隧道的实现过程可以参见本实施例前述S103和S104的说明，此处不再进行赘述。

应当理解，上述实施例的说明不限定S105的执行过程与S101-S104的执行过程的先后顺序。例如，S105可以先于S101执行，或者S105和S101同时被触发，或者S105后于S101执行。

通过上述过程，所述第一NVE设备与所述第二NVE设备之间能够建立VXLAN隧道，所述VXLAN隧道的两端的端点地址为所述第一VTEP地址和所述第二VTEP地址。

通过上述实施方式，在EVPN VXLAN的应用场景中，由至少两个NVE设备组成的NVE设备集合包括公共VTEP地址，所述公共VTEP地址用于标识公共VTEP，并且被部署在所述NVE设备集合中的每个NVE设备上。而且，所述NVE设备集合中的每个NVE设备还包括VTEP地址，所述VTEP地址用于标识对应的NVE设备包括的VTEP。所述NVE设备集合中的NVE设备根据VTEP地址和公共VTEP地址，通过IMET路由，建立所述NVE设备集合中的NVE设备之间的VXLAN隧道。从而，在EVPN VXLAN的应用场景中，NVE设备集合中的NVE设备通过VXLAN隧道实现互联，使得NVE设备集合中的NVE设备不受直连物理链路的限制。使用VXLAN隧道实现互联，即使NVE设备集合中的NVE设备之间存在交换节点，也不会影响EVPN VXLAN的实现。而且，而且在部署

NVE设备集合中的NVE设备时，不会受到部署地域和距离的影响，降低了部署的复杂程度。进一步，NVE设备集合中的NVE设备的互联线路不再受直连物理链路的影响，有助于提高通信的可靠性。

根据前述实施例针对图1的说明，可选的，所述EVPN VXLAN还包括PE设备。其中，
5 所述PE设备与所述第一NVE设备是一对BGP peer，所述PE设备与所述第二NVE设备是一对BGP peer和所述PE设备与所述第三NVE设备是一对BGP peer。下面以图1为例，说明所述PE设备与所述NVE设备集合之间建立VXLAN隧道的过程。所述处理路由的方法还包括：

S201、所述第一NVE设备向所述PE设备发送所述第一IMET路由。

10 S202、所述第二NVE设备向所述PE设备发送所述第二IMET路由。

S203、所述第三NVE设备向所述PE设备发送所述第三IMET路由。

根据前述实施例针对图2的说明，所述第一NVE设备、所述第二NVE设备和所述第三NVE设备作为所述NVE设备集合中的NVE设备均向各自的BGP peer发布IMET路由。即，所述第一NVE设备向所述所述第一NVE设备的BGP peer发布所述第一IMET路由；
15 所述第二NVE设备向所述所述第二NVE设备的BGP peer发布所述第二IMET路由；所述第三NVE设备向所述所述第三NVE设备的BGP peer发布所述第三IMET路由。其中，所述第一IMET路由、所述第二IMET路由和所述第三IMET路由的实现方式参见前述实施例的说明，此处不进行赘述。

20 S204、所述PE设备接收所述第一IMET路由、所述第二IMET路由和所述第三IMET路由。

S205、所述PE设备根据所述第一IMET路由、所述第二IMET路由和所述第三IMET路由建立从所述PE设备到所述NVE设备集合的VXLAN隧道。

由于所述PE设备是所述NVE设备集合中每台NVE设备的BGP peer，因此，所述PE设备能够接收所述第一IMET路由、所述第二IMET路由和所述第三IMET路由。上述S204
25 和S205可以通过两种可能的方式实现。

在一种可能的实现方式中，所述PE设备等待收到所有BGP peer发送的IMET路由后，建立VXLAN隧道。所述PE设备在接收到上述3条IMET路由后，分别解析上述3条IMET路由，分别获取上述3条IMET路由中的Originating Router's IP Address域包括的公共VTEP地址。所述PE设备确定上述3条IMET路由中的Originating Router's IP Address域包
30 括的公共VTEP地址相同，可以确定发送上述3条IMET路由的NVE设备属于同一个NVE设备集合。所述PE设备在3条IMET路由中任意选择一条IMET路由，并且，所述PE设备以选择出的IMET路由中包括的公共VTEP地址作为VXLAN隧道的目的端点地址建立从所述PE设备到包括所述公共VTEP地址的公共VTEP的VXLAN隧道。

在另一种可能的实现方式中，所述PE设备接收到上述3条IMET路由中的任意一条
35 IMET路由后，建立VXLAN隧道。例如，所述PE设备接收到所述第一NVE设备发送的第一IMET路由，解析所述第一IMET路由，获取所述第一IMET路由中的Originating Router's IP Address域包括的公共VTEP地址。所述PE设备以第一IMET路由中包括的公共VTEP地址作为VXLAN隧道的目的端点地址建立从所述PE设备到包括所述公共VTEP地址的公共VTEP的VXLAN隧道。建立VXLAN隧道后，当所述PE设备收到所述
40 第二NVE设备发送的第二IMET路由或所述第三NVE设备发送的第三IMET路由，确定

所述第二IMET路由或所述第三IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与建立所述VXLAN隧道使用的公共VTEP地址相同,不再重复建立从所述PE设备到包括所述公共VTEP地址的公共VTEP的VXLAN隧道。

上述两种实现方式中,建立VXLAN隧道的具体实现方式可以参见前述实施例的说明,此处不进行赘述。上述两种实现方式中,上述3条IMET路由中的Originating Router's IP Address域包括的公共VTEP地址相同。所述PE设备在接收到具有相同公共VTEP地址的IMET路由后,所述PE设备认为所述PE设备的peer端只存在一个与所述公共VTEP地址对应的VTEP,因此,所述PE设备只针对该公共VTEP地址建立一次VXLAN隧道,而不是针对NVE设备集合中每个NVE设备保存的公共VTEP地址建立VXLAN隧道。如此这样,所述PE设备以公共VTEP地址作为VXLAN隧道的目的端点地址建立的从所述PE设备到包括所述公共VTEP地址的公共VTEP的VXLAN隧道可以被称为从所述PE设备到所述NVE设备集合的VXLAN隧道。

可选的,上述3条IMET路由中均包括VNI,对于建立VXLAN隧道来讲,该VNI是前述的公共VTEP VNI。所述PE设备在确定上述3条IMET路由中包括相同的公共VTEP地址后,进一步确定上述3条IMET路由中包括相同的公共VTEP VNI,从而建立从所述PE设备到所述NVE设备集合的所述VXLAN隧道。其中,对于所述公共VTEP VNI的解释可以参考前述实施例的相应说明,此处不进行赘述。

S206、所述PE设备分别向所述第一NVE设备、所述第二NVE设备和所述第三NVE设备发送IMET路由,下一跳域和所述IMET路由中的Originating Router's IP Address域均包括所述PE设备的VTEP地址,所述PE设备的VTEP地址用于标识所述PE设备包括的VTEP。

所述PE设备可以向所述PE设备的BGP peer发布所述IMET路由,所述IMET路由的格式和基本用法可参见RFC7432的说明。作为举例,在本申请实施例中,所述PE设备的VTEP中包括的VTEP地址的值设置为4.4.4.4。下面以所述第一NVE设备接收所述PE设备发送的IMET路由为例进行说明,应当理解,所述第二NVE设备和所述第三NVE设备可以执行同样的处理过程。

S207、所述第一NVE设备接收来自所述PE设备的所述IMET路由。

S208、所述第一NVE设备根据所述来自所述PE设备的VTEP地址建立从所述第一NVE设备到所述PE设备的VXLAN隧道,所述PE设备的VTEP地址不等于所述第一NVE设备保存的公共VTEP地址。

所述第一NVE设备接收到所述PE设备发送的IMET路由后,解析所述IMET路由,获取所述IMET路由中Originating Router's IP Address域携带的所述PE设备的VTEP地址,以所述VTEP地址作为VXLAN隧道的目的端点地址建立从所述第一NVE设备到所述PE设备的VXLAN隧道。可选的,所述IMET路由还包括公共VTEP VNI,所述所述第一NVE设备获取所述公共VTEP VNI,建立对应所述公共VTEP VNI的,从所述第一NVE设备到所述PE设备的VXLAN隧道。

可选的,所述第一NVE设备在获取所述PE设备的VTEP地址后,可以执行与前述S103和S104类似的确定过程,即,所述第一NVE设备确定所述PE设备的VTEP地址与所述第一NVE设备保存的公共VTEP地址不相同,所述第一NVE设备可以确定所述第一NVE设备与所述PE设备不属于同一个NVE设备集合。所述第一NVE设备将使用公共VTEP地址

作为源端点地址建立从所述第一NVE设备到所述PE设备的VXLAN隧道，而不是使用第一VTEP地址作为源端点地址建立从所述第一NVE设备到所述PE设备的VXLAN隧道。

根据S207和S208，所述第二NVE设备建立从所述第二NVE设备到所述PE设备的VXLAN隧道，所述第三NVE设备建立从所述第三NVE设备到所述PE设备的VXLAN隧道。结合前述，所述第一NVE设备、所述第二NVE设备和所述第三NVE设备向所述PE设备建立的3条VXLAN隧道是相同的，即，3条VXLAN隧道的源端点地址均是公共VTEP地址，目的端点地址均是所述PE设备的VTEP地址，且3条VXLAN隧道均对应公共VTEP VNI。相对应的，所述PE设备建立了1条从所述PE设备到所述NVE设备集合的VXLAN隧道，源端点地址均是所述PE设备的VTEP地址，目的端点地址均是公共VTEP地址。如此这样，所述PE设备会认为对端只存在一台NVE设备，实际上，该NVE设备是所述NVE设备集合。而且，所述第一NVE设备、所述第二NVE设备和所述第三NVE设备仍然可以通过各自与所述PE设备建立的VXLAN隧道发送流量。

基于前述实施例，本申请的实施例可以应用在多于2台NVE设备组成NVE设备集合的网络场景中。以图1为例，NVE设备集合包括3台NVE设备。根据图2所示的方法，第一NVE设备与第二NVE设备之间建立第一VXLAN隧道，第二NVE设备与第三NVE设备之间建立第二VXLAN隧道，第一NVE设备与第三NVE设备之间建立第三VXLAN隧道。如此这样，本申请的技术方案可以实现一台CE设备连接到三台或三台以上NVE设备的网络场景，即CE设备多归属到三台或三台以上NVE设备。以图1为例，第一CE设备多归属到所述第一NVE设备，所述第二NVE设备和所述第三NVE设备。上述实现方式是一种EVPN VXLAN的多活冗余模式（英文：All-Active Redundancy Mode），多台NVE设备通过ES连接至少一台CE设备，其中，多台NVE设备形成一个多活冗余模式的设备集合，均用于转发流量。相比较之下，已有的EVPN VXLAN的多活冗余模式中，均采用双活冗余模式，CE设备最大只能双归属到两台NVE设备；并且，两台NVE设备之间通过物理链路直接连接，例如，两台NVE设备之间通过PeerLink链路连接。因此，本申请的实施例进一步的有益效果是，支持更多NVE设备的多活冗余模式，有助于增加EVPN VXLAN的可靠性和增加网络部署的灵活性。

图3为本申请实施例的另一种EVPN VXLAN结构示意图。图3示出了利用前述实施例提供的技术方案建立NVE设备集合中的NVE设备之间的VXLAN隧道后，实现指定转发者（英文：Designated Forwarder，缩写：DF）选举和广播、未知单播和组播（英文：Broadcast, Unknown unicast, and Multicast, BUM）流量的转发。图3示例性的所示了三台CE设备，分别为第一CE设备，第二CE设备和第三CE设备。其中，所述第一CE设备多归属到所述第一NVE设备、所述第二NVE设备和所述第三NVE设备，即，所述第一CE设备分别与所述第一NVE设备、所述第二NVE设备和所述第三NVE设备通信。所述第二CE设备双归属到所述第一NVE设备和所述第二NVE设备，即，所述第二CE设备分别与所述第一NVE设备和所述第二NVE设备通信。所述第三CE设备单归属到所述第二NVE设备，即，所述第三CE设备与所述第二NVE设备通信。图3以BUM流量从PE设备转发到所述第一CE设备，所述第二CE设备和所述第三CE设备为例进行说明。

所述第一CE设备经由多个以太网链路多归属到所述第一NVE设备、所述第二NVE设备和所述第三NVE设备，因此，所述多个以太网链路构成一个ES，如图3中的ES1所示。所述第一CE设备连接所述第一NVE设备、所述第二NVE设备和所述第三NVE设备

的3个以太网链路的ES1的ESI值为同一个值。所述第一NVE设备，所述第二NVE设备和所述第三NVE设备分别向各自的BGP peer发送以太网段（英文：Ethernet Segment，缩写：ES）路由，所述ES路由用于DF选举。所述ES路由包括RD域、ESI域、IP地址长度域和Originating Router's IP Address域。上述各个字段域的定义类似前述IMET路由，此处不进行赘述。另外，ES路由的格式和基本用法也可以参见RFC7432第7.4章和第8.5章的说明。以图3中的第一CE设备为例，所述第一NVE设备接收到来自所述第二NVE设备和所述第三NVE设备的ES路由，所述第一NVE设备根据从所述第二NVE设备和所述第三NVE设备接收的ES路由，将所述第一NVE设备确定为所述多个以太网链路构成的ES中的DF。同样道理，所述第二NVE设备被确定为所述多个以太网链路构成的ES中的非指定转发者（英文：non-Designated Forwarder，缩写：NDF）；所述第三NVE设备被确定为所述多个以太网链路构成的ES中的NDF。如图3所示，所述第一CE设备与所述第一NVE设备之间的ES1为DF，所述第一CE设备与所述第二NVE设备之间的ES1为NDF，所述第一CE设备与所述第三NVE设备之间的ES1为NDF。所述第二CE设备经由2个以太网链路双归属到所述第一NVE设备和所述第二NVE设备，所述2个以太网链路构成一个ES，如图3中的ES2所示。所述第二CE设备连接所述第一NVE设备和所述第二NVE设备的2个以太网链路的ES2的ESI值为同一个值。同样道理，所述第二NVE设备被确定为所述2个以太网链路构成的ES中的DF；所述第一NVE设备被确定为所述2个以太网链路构成的ES中的NDF。所述第三CE设备单归属到所述第二NVE设备，因此不存在DF选举过程。

NVE设备向所述NVE设备的BGP peer发送ES路由可以按照以下方式实现。所述NVE设备可以向所述NVE设备的BGP peer发送UPDATE message，所述UPDATE message携带有MP_REACH_NLRI属性。所述MP_REACH_NLRI属性包括EVPN NLRI域和下一跳域。所述EVPN NLRI域用于携带ES路由。

PE设备基于VXLAN隧道向NVE设备集合发送BUM流量。具体的，所述PE设备根据哈希（英文：Hash）规则向所述NVE设备集合发送VXLAN报文，所述VXLAN报文携带所述BUM流量。可选的，所述VXLAN报文还携带有BUM标记，用于标识所述VXLAN报文携带的流量是BUM流量。在一种可能的实现方式中，所述BUM标记可以被携带在所述VXLAN报文的VXLAN头中。

以所述VXLAN报文被哈希到所述第二NVE设备为例进行说明。在本实施例和后续实施例中，为了区别PE设备与NVE设备集合之间传送的VXLAN报文与NVE设备集合中NVE设备之间传送的VXLAN报文，将PE设备与NVE设备集合之间传送的VXLAN报文命名为PE VXLAN报文。所述第二NVE设备接收到来自所述PE设备的PE VXLAN报文，解析所述PE VXLAN报文，根据PE VXLAN报文中携带的BUM标记，可以确定所述PE VXLAN报文携带的流量为BUM流量。所述第二NVE设备将所述BUM流量向与所述第二NVE设备具有连接关系的CE设备转发，并且所述第二NVE设备还将所述BUM流量向所述NVE设备集合中的其他NVE设备转发。

所述第二NVE设备将所述BUM流量向与所述第二NVE设备具有连接关系的CE设备转发：所述第二NVE设备确定所述第二NVE设备与所述第二CE设备之间的ES2处于DF状态，所述第二NVE设备将所述BUM流量传送到所述第二NVE设备与ES2连接的出口接口，经由ES2发送到所述第二CE设备。所述第二NVE设备确定所述第二NVE设备与

所述第一CE设备之间的ES1处于NDF状态，所述第二NVE设备不向所述第二NVE设备与ES1连接的出接口传送所述BUM流量，从而，所述第二NVE设备避免经由ES1向所述第一CE设备发送所述BUM流量。另外，所述第二NVE设备确定所述第二NVE设备与所述第三CE设备之间的ES3不存在DF或NDF状态，直接将所述BUM流量经由ES3发送到所述第三CE设备。

所述第二NVE设备将所述BUM流量向所述NVE设备集合中的其他NVE设备转发：所述第二NVE设备解析来自所述PE设备的PE VXLAN报文，获取所述BUM流量。所述第二NVE设备确定所述BUM流量来自所述PE设备，将所述BUM流量封装为VXLAN报文，经由第一VXLAN隧道向所述第一NVE设备发送，并且经由第二VXLAN隧道向所述第二NVE设备发送。

所述第一NVE设备经由所述第一VXLAN隧道接收到所述VXLAN报文后，能够确定所述VXLAN报文中的BUM流量是经由所述第一VXLAN隧道接收的，不再将所述BUM流量向与所述第一NVE设备连接的第三VXLAN隧道转发。如图3所示，所述所述第一NVE设备经由所述第一VXLAN隧道接收到所述VXLAN报文后，解析所述VXLAN报文，获取BUM流量，并且不再将所述BUM流量经由第三VXLAN隧道向所述第三NVE设备转发。也就是说，所述第一NVE设备根据BUM流量是经由所述第一VXLAN隧道接收的，避免经由所述第一NVE设备到所述第三NVE设备的VXLAN隧道向所述第三NVE设备转发所述BUM流量。基于上述的有益效果是，避免BUM流量双发。

所述第一NVE设备经由所述第一VXLAN隧道接收到所述VXLAN报文后，解析所述VXLAN报文，获取BUM流量，还将所述BUM流量发送到与所述第一NVE设备连接的CE设备。如图3所示，所述第一NVE设备确定所述第一NVE设备与所述第一CE设备之间的ES1处于DF状态，所述第一NVE设备将所述BUM流量传送到所述第一NVE设备与ES1连接的出接口，经由ES1发送到所述第一CE设备。所述第一NVE设备确定所述第一NVE设备与所述第二CE设备之间的ES2处于NDF状态，所述第一NVE设备不向所述第一NVE设备与ES2连接的出接口传送所述BUM流量，从而，所述第一NVE设备避免经由ES2向所述第二CE设备发送所述BUM流量。

所述第三NVE设备经由所述第二VXLAN隧道接收到所述VXLAN报文后，解析所述VXLAN报文，获取BUM流量。一方面，所述第三NVE设备避免经由第三VXLAN隧道向所述第一NVE设备转发所述BUM流量。另一方面，所述第三NVE设备确定所述第三NVE设备与所述第一CE设备之间的ES1处于NDF状态，所述第三NVE设备不向所述第三NVE设备与ES1连接的出接口传送所述BUM流量，从而，所述第三NVE设备避免经由ES1向所述第一CE设备发送所述BUM流量。

通过上述实施方式，利用NVE设备之间的VXLAN隧道不仅可以实现BUM流量的正常转发，还可以避免BUM流量向其他NVE设备之间的VXLAN隧道的转发，从而避免BUM流量的双发。

同样道理，从CE设备到PE设备和其他CE设备的BUM流量转发，是上述实施方式的逆向过程，因此可以参照上述方式完成BUM流量的转发。例如，BUM流量从所述第二CE设备被转发到所述PE设备、所述第一CE设备和所述第三CE设备。BUM经由所述第二NVE设备和所述第二CE设备之间的ES2到达所述第二NVE设备。所述第二NVE设备将BUM流量封装为PE VXLAN报文，经由所述第二NVE设备与所述PE设备之间的

VXLAN隧道转发到所述PE设备。所述第二NVE设备还将BUM流量封装为VXLAN报文，分别经由第一VXLAN隧道和第二VXLAN隧道转发到所述第一NVE设备和所述第三NVE设备。所述第二NVE设备还将BUM流量经由ES3直接转发给所述第三CE设备，但是，避免经由所述第二NVE设备与所述第一CE设备之间的ES1向所述第一CE设备转发所述BUM流量。所述第三NVE设备收到来自所述第二NVE设备的BUM流量后，避免经由第三VXLAN隧道向所述第三NVE设备转发所述BUM流量，并且避免经由VXLAN隧道向所述PE设备转发所述BUM流量。所述第三NVE设备将所述BUM流量经由ES1向所述第一CE设备转发，但是，避免经由ES2向所述第二CE设备转发所述BUM流量。所述第三NVE设备收到来自所述第二NVE设备的BUM流量后，避免经由第三VXLAN隧道向所述第一NVE设备转发所述BUM流量，并且，避免经由ES1向所述第一CE设备转发所述BUM流量，进一步，还避免经由VXLAN隧道向所述PE设备转发所述BUM流量。

在上述例举的BUM流量从所述第二CE设备被转发到所述PE设备、所述第一CE设备和所述第三CE设备过程中，所述第二CE设备发出的所述BUM经由所述第二NVE设备和所述第二CE设备之间的ES2到达所述第二NVE设备。但是，在实际的场景中，所述第二CE设备发出的所述BUM有可能经由所述第一NVE设备和所述第二CE设备之间的ES2到达所述第一NVE设备，因为，所述第二CE设备并不关心通信链路的DF或NDF状态。如图4所示，当所述第二CE设备将BUM经由ES2转发到所述第一NVE设备。所述第一NVE设备将所述BUM流量封装为VXLAN报文，经由第一VXLAN隧道转发到所述第二NVE设备。由于所述第二NVE设备与所述第二CE设备之间的ES2处于DF状态，因此，所述第二NVE设备会将所述BUM流量经由ES2向所述第二CE设备转发，如图4中的第二NVE设备到第二CE设备之间的虚线箭头所示。如此这样，所述第二CE设备、所述第一NVE设备和所述第二NVE设备形成了BUM流量环路。应当理解，为了更加清楚的呈现环路问题，图4中示出的BUM流量转发线路仅是出现环路的部分，图4没有示出BUM流量在EVPN VXLAN中的全部转发线路。

下面结合图4具体说明解决所述第二CE设备、所述第一NVE设备和所述第二NVE设备形成的BUM流量环路问题的方法。其中，所述第二CE设备经由2个以太网链路双归属到所述第一NVE设备和所述第二NVE设备，所述2个以太网链路构成一个ES，如图4中的ES2所示。所述第二CE设备连接所述第一NVE设备和所述第二NVE设备的2个以太网链路的ES2的ESI值为同一个值。所述第二NVE设备根据从所述第一NVE设备接收ES路由，将所述第二NVE设备确定为所述2个以太网链路构成的ES中的DF。

S301、所述第二NVE设备向所述第一NVE设备发送以太网自动发现每以太网段（英文：Ethernet A-D per ES）路由，所述Ethernet A-D per ES路由携带所述第二NVE设备为所述第二CE设备与所述第二NVE设备之间的以太网链路分配的ESI标签。

S302、所述第一NVE设备接收所述Ethernet A-D per ES路由。

由于所述第一NVE设备和所述第二NVE设备是一对BGP peer，因此，所述第二NVE设备能够向所述第一NVE设备发送Ethernet A-D per ES路由。其中，所述Ethernet A-D per ES路由属于以太网自动发现（英文：Ethernet Auto-discovery）路由。所述Ethernet Auto-discovery路由的格式和基本用法也可以参见RFC7432第7.1章和第8.2章的说明。所述Ethernet A-D per ES路由可以携带ESI标签扩展团体（英文：ESI Label Extended Community）属性，所述ESI Label Extended Community属性可以携带ESI标签。所述第

一NVE设备接收到所述Ethernet A-D per ES路由后，根据所述Ethernet A-D per ES路由中包括的ESI，可以确定所述第二CE设备不仅连接所述第一NVE设备，还连接所述第二NVE设备。而且，所述第一NVE设备根据所述Ethernet A-D per ES路由携带的所述ESI标签，可以确定所述第二NVE设备为所述第二CE设备与所述第二NVE设备之间的以太网链路分配的ESI标签的值。

NVE设备向所述NVE设备的BGP peer发送Ethernet A-D per ES路由可以按照以下方式实现。所述NVE设备可以向所述NVE设备的BGP peer发送UPDATE message，所述UPDATE message携带有MP_REACH_NLRI属性。所述MP_REACH_NLRI属性包括EVPN NLRI域和下一跳域。所述EVPN NLRI域用于携带Ethernet A-D per ES路由。

S303、所述第一NVE设备经由所述第一VXLAN隧道向所述第二NVE设备发送VXLAN报文，所述VXLAN报文中的原始以太网载荷包括BUM流量，所述VXLAN报文携带所述ESI标签，其中，所述BUM流量来自所述第二CE设备。

S304、所述第二NVE设备经由所述第一VXLAN隧道接收所述第一NVE设备发送的VXLAN报文。

如图4所示，当所述第二CE设备将BUM经由ES2转发到所述第一NVE设备。所述第一NVE设备接收到所述BUM流量后，确定所述第二CE设备与所述第一NVE设备之间的ESI的ESI，进而可以确定所述第二CE设备不仅连接所述第一NVE设备，还连接所述第二NVE设备。因此，所述第一NVE设备将所述BUM流量封装为VXLAN报文，其中，所述VXLAN报文中的原始以太网载荷包括所述BUM流量，所述VXLAN报文携带所述ESI标签。也就是说，所述第一NVE设备经由第一Bypass VXLAN隧道向所述第二NVE设备发送的VXLAN报文不仅包括BUM流量，还携带所述第二NVE设备为所述第二CE设备与所述第二NVE设备之间的以太网链路分配的ESI标签。所述原始以太网载荷的定义参见RFC7348的说明。

可选的，所述VXLAN报文携带所述ESI标签可以按照图5所示的格式实现。如图5所示，所述VXLAN报文包括VXLAN隧道部分和原始以太网载荷，所述VXLAN隧道部分包括外层Eth头（Eth头：Ethernet头）、外层IP头、外层UDP头和VXLAN头，所述原始以太网载荷中携带所述BUM流量。所述VXLAN报文还包括ESI标签，所述ESI标签被封装于所述VXLAN报文中的VXLAN头和所述原始以太网载荷之间。而且，所述VXLAN头携带指示信息，所述指示信息用于指示所述VXLAN报文携带所述ESI标签。

S305、所述第二NVE设备根据所述VXLAN报文携带所述ESI标签，避免经由所述第二CE设备与所述第二NVE设备之间的ES向所述CE设备转发所述BUM流量。

所述第二NVE设备解析所述VXLAN报文，获取所述VXLAN报文携带的所述ESI标签，根据所述ESI标签的值可以确定所述ESI标签是所述第二NVE设备为所述第二CE设备与所述第二NVE设备之间的以太网链路分配的ESI标签。所述第二NVE设备不再将所述BUM流量向所述第二NVE设备与ES2连接的出接口传送，从而，避免经由所述第二CE设备与所述第二NVE设备之间的ES向所述CE设备转发所述BUM流量。

上述方法是以图4所示出的环路为示例进行解释说明，应当理解，上述方法不仅局限解决所述第二CE设备、所述第一NVE设备和所述第二NVE设备形成的BUM流量环路问题。在EVPN VXLAN的应用场景中产生的环路问题均可以应用上述方法解决。

通过上述实施方式，通过扩展VXLAN报文的格式，以便VXLAN报文能够携带ESI标

签, 从而解决转发BUM流量过程中产生的环路问题。

本申请的实施例的EVPN VXLAN还可以用于MAC路由的处理, 从而根据图2所示的方法建立VXLAN隧道后, 所述EVPN VXLAN可以用于指导单播流量的转发。下面结合图1具体说明EVPN VXLAN用于MAC路由的处理的方法。其中, 所述第二CE设备经由2个以太网链路双归属到所述第一NVE设备和所述第二NVE设备, 所述2个以太网链路构成一个ES。所述第二CE设备连接所述第一NVE设备和所述第二NVE设备的2个以太网链路的ES2的ESI值为同一个值。

S401、所述第一NVE设备接收来自所述第二NVE设备的媒体接入控制/因特网协议通告(英文: MAC/IP Advertisement)路由, 所述MAC/IP Advertisement路由携带媒体接入控制-虚拟局域网标识(英文: Media Access Control-Virtual Local Area Network Identifier, 缩写: MAC-VLAN ID), 其中, 所述MAC-VLAN ID用于指示所述MAC/IP Advertisement路由携带的MAC地址所属的VLAN, 所述MAC/IP Advertisement路由携带的MAC地址是所述CE设备的MAC地址, 或者所述MAC/IP Advertisement路由携带的MAC地址是所述CE设备所管辖的以太网虚拟私有网络站点(英文: EVPN site)中的主机的MAC地址。

由于所述第一NVE设备和所述第二NVE设备是一对BGP peer, 因此, 所述第二NVE设备能够向所述第一NVE设备发送MAC/IP Advertisement路由。相应的, 所述第一NVE设备接收来自所述第二NVE设备的所述MAC/IP Advertisement路由。MAC/IP Advertisement路由的格式和基本用法也可以参见RFC7432第7.2章、第9章和第14章的说明。本申请实施例对RFC7432说明的MAC/IP Advertisement路由进行扩展, 使所述MAC/IP Advertisement路由能够携带MAC-VLAN ID。其中, 所述MAC-VLAN ID用于指示所述MAC/IP Advertisement路由携带的MAC地址所属的VLAN。所述MAC地址用于标识所述第二CE设备, 或者所述MAC地址用于标识所述第二CE设备连接的主机。其中, 主机是物理设备或VM。

可选的, 本申请实施例扩展了一种新的属性, 用来携带MAC-VLAN ID。如图8所示, 本申请实施例扩展了MAC-VLAN属性, 用来携带MAC-VLAN ID。其中, 类型域和子类型域用于表示MAC-VLAN属性所属类型, 具体数值可以根据标准组织的要求进行设定。标志域可以用于指示所述MAC-VLAN属性是否允许被读取, 例如, 标志域为0表示, 所述MAC-VLAN属性中的MAC-VLAN ID不允许被读取, 标志域为1表示, 所述MAC-VLAN属性中的MAC-VLAN ID允许被读取。应当理解, 上述标志域的功能是一种示例, 标志域的功能可以根据使用场景进行相应的变换。MAC-VLAN ID域用于携带MAC-VLAN ID的值。预留域为了实现进一步的扩展。MAC-VLAN属性可以作为一种属性被携带在MAC/IP Advertisement路由中, 以便实现本申请实施例的方案。

S402、所述第一NVE设备根据所述MAC/IP Advertisement路由携带的ESI和所述MAC-VLAN ID确定连接所述第二CE设备与所述第一NVE设备之间的以太网链路的所述第一NVE的本地接口, 其中, 所述MAC/IP Advertisement路由携带的ESI用于指示所述第二CE设备与所述第二NVE设备之间的ES。

所述第二NVE设备向所述第一NVE设备发送的所述MAC/IP Advertisement路由中包括ESI域, 所述ESI域用于携带指示所述第二CE设备与所述第二NVE设备之间的以太网链路的ESI。所述第二CE设备与所述第一NVE设备和所述第二NVE设备之间的ES2的ESI的值相等, 因此, 所述第一NVE设备能够根据所述MAC/IP Advertisement路由携带的ESI

和所述MAC-VLAN ID确定连接所述第二CE设备与所述第一NVE设备之间的以太网链路的所述第一NVE的本地接口。也就是说，所述第一NVE设备可以根据所述ESI和所述MAC-VLAN ID确定所述第一NVE设备上的本地接口，该本地接口用于与所述MAC/IP Advertisement路由携带的MAC地址所标识的设备通信。

- 5 S403、所述第一NVE设备的控制平面向所述第一NVE设备的转发平面发送第一MAC表项，所述第一MAC表项中的出接口为所述第一NVE设备的本地接口，其中，所述第一MAC表项包括所述MAC/IP Advertisement路由携带的MAC地址和所述MAC-VLAN ID。

10 所述第一NVE设备将包括有所述MAC地址和所述MAC-VLAN ID的所述第一MAC表项从所述第一NVE设备的控制平面下发到所述第一NVE设备的转发平面，以便所述第一NVE设备能够根据所述第一MAC表项指导单播流量的转发。其中，所述第一MAC表项中的出接口为所述第二CE设备与所述第一NVE设备之间的以太网链路的所述第一NVE的本地接口。

15 举例说明，如图1所示，PE设备向第二CE设备发送单播流量。假设所述PE设备将所述单播流量哈希到第二NVE设备。所述第二NVE设备接收来自所述PE设备的VXLAN报文，所述VXLAN报文携带所述单播流量。所述第二NVE设备获取所述VXLAN报文中的所述单播流量，基于所述第二NVE设备存储的MAC表项，将所述单播流量转发到所述第二CE设备。

20 如图9所示，如果所述第二NVE设备与所述第二CE设备之间的ES2发生故障，所述单播流量将无法经由所述第二NVE设备与所述第二CE设备之间的ES2到达所述第二CE设备。如图9中虚线箭头所示。所述第一NVE设备和所述第二NVE设备执行了上述S401-S403。因此所述第一NVE设备已经将包括有MAC地址和MAC-VLAN ID的第一MAC表项从所述第一NVE设备的控制平面下发到所述第一NVE设备的转发平面。所述第二NVE设备经由第一VXLAN隧道向所述第一NVE设备发送携带所述单播流量的VXLAN报文。所述第一NVE设备25 收到所述VXLAN报文后，获取所述单播流量，并根据所述第一MAC表项，将所述单播流量经由所述第一NVE设备与所述第二CE设备之间的ES2转发到所述第二CE设备。

30 通过上述实施方式，可以通过MAC/IP Advertisement路由和NVE设备之间的VXLAN隧道实现转发路径的重新定向，在某个以太网链路出现链路故障时，依然可以确保单播流量的正常转发。上述方法以CE设备双归属到两台NVE设备为例进行说明，应当理解，上述方法也能应用于CE设备多归属到多于两台NVE设备的场景中。

可选的，上述MAC路由的处理的方法还包括：

35 S404、所述MAC/IP Advertisement路由被携带在MP_REACH_NLRI中，所述MP_REACH_NLRI中的下一跳域包括公共VTEP地址，所述MAC/IP Advertisement路由还包括所述第二VTEP地址，所述MP_REACH_NLRI中的下一跳域包括的公共VTEP地址与所述第二IMET路由中的Originating Router's IP Address域包括的公共VTEP地址相同；所述方法还包括：所述第一NVE设备的控制平面根据所述MP_REACH_NLRI中的下一跳域包括的公共VTEP地址和所述MAC/IP Advertisement路由包括的所述第二VTEP地址，向所述第一NVE设备的转发平面发送第二MAC表项，所述第二MAC表项中的出接口为连接所述第一NVE设备到所述第二NVE设备的VXLAN隧道的所述第一NVE设备的本地接口，其中，所述40 第二MAC表项包括所述MAC/IP Advertisement路由携带的MAC地址。

所述第二NVE设备向所述第一NVE设备发送MAC/IP Advertisement路由可以按照以下方式实现。所述第二NVE设备可以向所述第一NVE设备发送UPDATE message, 所述UPDATE message携带有MP_REACH_NLRI属性。所述MP_REACH_NLRI属性包括EVPN NLRI域和下一跳域。所述EVPN NLRI域用于携带MAC/IP Advertisement路由。所述下一跳域包括公共VTEP地址, 而且所述MAC/IP Advertisement路由还包括所述第二VTEP地址。所述第二NVE设备根据所述公共VTEP地址可以确定所述第一NVE设备和所述第二NVE设备同属于一个NVE设备集合。所述第二NVE设备根据所述第二VTEP地址, 可以确定所述MAC/IP Advertisement路由是由所述第二NVE设备发布的。所述第一NVE设备向所述第一NVE设备的转发平面发送第二MAC表项, 所述第二MAC表项中的出接口为连接所述第一NVE设备到所述第二NVE设备的VXLAN隧道的所述第一NVE设备的本地接口, 其中, 所述第二MAC表项包括所述MAC/IP Advertisement路由携带的MAC地址。

举例说明, 如图1所示, PE设备向第二CE设备发送单播流量。假设所述PE设备将所述单播流量哈希到第一NVE设备。所述第一NVE设备接收来自所述PE设备的PE VXLAN报文, 所述PE VXLAN报文携带所述单播流量。所述第一NVE设备获取所述PE VXLAN报文中的所述单播流量, 基于所述第一NVE设备存储的MAC表项, 将所述单播流量经由所述第一NVE设备与所述第二CE设备之间的ES2转发到所述第二CE设备。而且, 所述第一NVE设备已经向所述第一NVE设备的转发平面发送第二MAC表项, 所述第二MAC表项中的出接口为连接所述第一NVE设备到所述第二NVE设备的VXLAN隧道的所述第一NVE设备的本地接口, 其中, 所述第二MAC表项包括所述MAC地址。因此, 所述第一NVE设备可以将所述单播流量经由第一VXLAN隧道转发到所述第二NVE设备。所述第二NVE设备基于所述第二NVE设备存储的MAC表项, 将所述单播流量经由所述第二NVE设备与所述第二CE设备之间的ES2转发到所述第二CE设备。

通过上述实施方式, 可以进一步实现单播流量的负载分担, 有助于提高单播流量的传输带宽。

上述MAC路由的处理适用于CE设备双归属连接到两台NVE设备的场景或CE设备多归属到多于两台NVE设备的场景。下面结合图1具体说明EVPN VXLAN用于CE设备单归到一台NVE设备场景的MAC路由的处理的方法。如图1所示, 第三CE设备单归属到第二NVE设备。

S501、所述第一NVE设备接收来自所述第二NVE设备MAC/IP Advertisement路由, 所述MAC/IP Advertisement路由被携带在MP_REACH_NLRI中, 所述MP_REACH_NLRI中的下一跳域包括公共VTEP地址, 所述MAC/IP Advertisement路由包括MAC-VLAN ID和所述第二VTEP地址, 其中, 所述MAC-VLAN ID用于指示所述MAC/IP Advertisement路由携带的MAC地址所属的VLAN, 所述MAC/IP Advertisement路由携带的MAC地址是所述CE设备的MAC地址, 或者所述MAC/IP Advertisement路由携带的MAC地址是所述CE设备所管辖的EVPN site中的主机的MAC地址。

所述S501中的执行过程可以参见前述实施例中的S401的相应解释, 此处不进行赘述。

S502、所述第一NVE设备根据所述MAC/IP Advertisement路由携带的ESI和所述MAC-VLAN ID确定不存在连接所述CE设备与所述第一NVE设备之间的以太网链路的本地接口, 其中, 所述MAC/IP Advertisement路由携带的ESI用于指示所述CE设备与所述第二NVE设备之间的ES。

结合前述实施例中的S402的相应解释，由于所述第三CE设备单归属连接所述第二NVE设备，因此所述第二NVE设备与所述第三CE设备之间的ES3的ESI的值为0或无效值。所述第一NVE设备接收到来自所述第二NVE设备MAC/IP Advertisement路由，获取所述MAC/IP Advertisement路由携带的ESI，发现所述ESI的值为0或无效值，则认为不存在ESI。因此，所述第一NVE设备根据所述MAC/IP Advertisement路由携带的ESI和所述MAC-VLAN ID确定不存在连接所述CE设备与所述第一NVE设备之间的以太网链路的本地接口。

S503、所述第一NVE设备的控制平面根据所述MP_REACH_NLRI中的下一跳域包括的公共VTEP地址和所述MAC/IP Advertisement路由包括的所述第二VTEP地址，向所述第一NVE设备的转发平面发送MAC表项。所述MAC表项的出接口为连接所述第一NVE设备到所述第二NVE设备的VXLAN隧道的所述第一NVE设备的本地接口。其中，所述MAC表项包括所述MAC/IP Advertisement路由携带的MAC地址。

所述第一NVE设备在确定不存在连接所述CE设备与所述第一NVE设备之间的ES的本地接口后，执行所述S503。所述S503中的执行过程可以参见前述实施例中的S404的相应解释，此处不进行赘述。

举例说明，如图10所示，PE设备向第三CE设备发送单播流量。假设所述PE设备将所述单播流量哈希到第一NVE设备。所述第一NVE设备接收来自所述PE设备的PE VXLAN报文，所述PE VXLAN报文携带所述单播流量。所述第一NVE设备获取所述PE VXLAN报文中的所述单播流量，并且，所述第一NVE设备在本地没有查找到从所述第一NVE设备到所述第三CE设备的MAC表项。所述第一NVE设备已经执行了上述S501-S503，因此，所述第一NVE设备可以将所述单播流量经由第一VXLAN隧道转发到所述第二NVE设备。所述第二NVE设备基于所述第二NVE设备存储的MAC表项，将所述单播流量经由所述第二NVE设备与所述第三CE设备之间的ES3转发到所述第三CE设备。

通过上述实施方式，可以确保单归属连接到NVE设备的CE设备能够正常接收到单播流量。

本申请实施例的EVPN VXLAN还可以应用在三层转发的场景中。具体的，如图1所示，分别在PE设备、第一NVE设备、第二NVE设备和第三NVE设备中部署虚拟路由转发（英文：Virtual Routing and Forwarding，缩写：VRF）实例。本实施例中，为了便于说明将PE设备、第一NVE设备、第二NVE设备和第三NVE设备中部署VRF实例设定为同一个VRF实例。应当理解，上述设备中的每个设备均可以设定多个VRF实例，例如不同的VRF实例对应不同的业务。所述第一NVE设备、所述第二NVE设备和所述第三NVE设备向各自的BGP peer发布因特网协议前缀（英文：IP Prefix）路由。所述IP Prefix路由包括RD域，ESI域，Ethernet tag ID域，IP Prefix长度域，IP Prefix域，网关因特网协议地址（英文：GW IP Address）域和MPLS Label域。上述各个域的定义类似本申请前述实施例中的定义，其中，IP Prefix域用于指示CE设备或CE设备连接的主机的IP地址；所述MPLS Label域用于指示三层VNI，所述三层VNI对应于VRF实例；所述GW IP Address域被设置为0，即无效值。

NVE设备向所述NVE设备的BGP peer发送IP Prefix路由可以按照以下方式实现。所述NVE设备可以向所述NVE设备的BGP peer发送UPDATE message，所述UPDATE message携带有MP_REACH_NLRI属性。所述MP_REACH_NLRI属性包括EVPN NLRI域和下一跳

域。所述EVPN NLRI域用于携带IP Prefix路由。

在本实施例中，所述第一NVE设备、所述第二NVE设备和所述第三NVE设备向外发布的携带IP Prefix路由的MP_REACH_NLRI属性中的下一跳域包括前述实施例的公共VTEP地址。作为对IP Prefix路由的扩展，所述IP Prefix路由还携带发送所述IP Prefix路由的NVE设备的VTEP地址。进一步可选的，所述IP Prefix路由可以携带前述实施例描述的VXLAN属性，所述VXLAN属性携带VTEP地址。通过IP Prefix路由的处理过程，可以实现IP路由的学习和将IP路由表项下发到转发平面，从而指导数据流量的三层转发。本实施例的IP Prefix路由的处理过程和三层流量转发过程类似前述实施例中的MAC路由处理的方法，此处不再进行赘述。

图11为本申请实施例的第一NVE设备1000的结构示意图。图11所示的第一NVE设备1000可以执行上述实施例的方法中第一NVE设备执行的相应步骤。如图11所示，所述第一NVE设备1000包括接收单元1002和处理单元1004。

所述接收单元1002，用于接收来自所述第二NVE设备的第二IMET路由，所述第二IMET路由中的Originating Router's IP Address域包括公共VTEP地址，其中，所述第二IMET路由还包括第二VTEP地址，所述第二IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与所述第二VTEP地址不相同；

所述处理单元1004，用于确定所述第二IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与所述第一NVE设备保存的公共VTEP地址是否相同。其中，所述第一NVE设备能够向所述第一NVE设备的BGP peer发送Originating Router's IP Address域中携带所述第一NVE设备保存的公共VTEP地址的IMET路由。所述第二IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与所述第一NVE设备保存的公共VTEP地址相同。

当所述处理单元1004确定所述第二IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与所述第一NVE设备保存的公共VTEP地址相同时，所述处理单元1004还用于根据所述第二IMET路由中的所述第二VTEP地址建立从所述第一NVE设备到所述第二NVE设备的VXLAN隧道。进一步可选的，所述第二IMET路由中的VXLAN属性携带所述第二VTEP地址。

可选的，所述第一NVE设备还包括发送单元1006。所述发送单元1006用于向所述第二NVE设备发送第一IMET路由。所述第一IMET路由的Originating Router's IP Address域中携带所述第一NVE设备保存的公共VTEP地址。其中，所述第一IMET路由还包括第一VTEP地址，所述第一VTEP地址被所述第二NVE设备用于建立从所述第二NVE设备到所述第一NVE设备的VXLAN隧道。所述第一NVE设备保存的公共VTEP地址与所述第一VTEP地址不相同。可选的，所述第一IMET路由中的VXLAN属性携带所述第一VTEP地址。

可选的，所述发送单元1006还用于向PE设备发送所述第一IMET路由。所述接收单元1002还用于接收来自所述PE设备的IMET路由。所述来自所述PE设备的IMET路由中的Originating Router's IP Address域携带所述PE设备的VTEP地址。所述处理单元1004还用于根据所述来自所述PE设备的IMET路由中的所述PE设备的VTEP地址建立从所述第一NVE设备到所述PE设备的VXLAN隧道。所述PE设备的VTEP地址不等于所述第一NVE设备保存的公共VTEP地址。

可选的，所述接收单元1002还用于接收来自所述第三NVE设备的第三IMET路由。所

述第三IMET路由中的Originating Router's IP Address域包括公共VTEP地址。其中，所述第三IMET路由还包括第三VTEP地址。所述第三IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与所述第三VTEP地址不相同。所述处理单元1004还用于确定所述第三IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与所述第一NVE设备保存的公共VTEP地址是否相同，其中，所述第三IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与所述第一NVE设备保存的公共VTEP地址相同；当所述处理单元1004确定所述第三IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与所述第一NVE设备保存的公共VTEP地址相同时，所述处理单元1004还用于根据所述第三IMET路由中的所述第三VTEP地址建立从所述第一NVE设备到所述第三NVE设备的VXLAN隧道。可选的，所述第三IMET路由中的VXLAN属性携带所述第三VTEP地址。

可选的，CE设备经由多个以太网链路多归属到所述第一NVE设备、所述第二NVE设备以及所述第三NVE设备，所述多个以太网链路构成一个ES；所述处理单元1004还用于根据来自所述第二NVE设备和所述第三NVE设备的ES路由，将所述第一NVE设备确定为所述ES中的DF。

可选的，所述接收单元1002还用于经由所述第二NVE设备到所述第一NVE设备的VXLAN隧道接收所述第二NVE设备发送的BUM流量；所述处理单元1004还用于基于所述BUM流量是所述第一NVE设备经由所述第二NVE设备到所述第一NVE设备的VXLAN隧道接收的，避免经由所述第一NVE设备到所述第三NVE设备的VXLAN隧道向所述第三NVE设备转发所述BUM流量。

可选的，CE设备经由多个以太网链路双归属到所述第一NVE设备和所述第二NVE设备，所述多个以太网链路构成一个ES；所述处理单元1004还用于根据来自所述第二NVE设备的ES路由，将所述第一NVE设备确定为所述ES中的DF；所述发送单元1006还用于向所述第二NVE设备发送Ethernet A-D per ES路由，所述Ethernet A-D per ES路由携带所述第一NVE设备为所述CE设备与所述第一NVE设备之间的以太网链路分配的ESI标签；所述接收单元1002还用于经由所述第二NVE设备到所述第一NVE设备的VXLAN隧道接收所述第二NVE设备发送的VXLAN报文，所述VXLAN报文中的原始以太网载荷包括BUM流量，所述VXLAN报文携带所述ESI标签，其中，所述BUM流量来自所述CE设备；所述处理单元1004还用于基于所述VXLAN报文携带所述ESI标签，避免经由所述CE设备与所述第一NVE设备之间的ES向所述CE设备转发所述BUM流量。进一步可选的，所述ESI标签被封装于所述VXLAN报文中的VXLAN头和所述原始以太网载荷之间，所述VXLAN头携带指示信息，所述指示信息用于指示所述VXLAN报文携带所述ESI标签。

可选的，CE设备经由多个以太网链路双归属到所述第一NVE设备和所述第二NVE设备，所述多个以太网链路构成一个ES。所述接收单元1002还用于接收来自所述第二NVE设备的MAC/IP Advertisement路由，所述MAC/IP Advertisement路由包括MAC-VLAN ID，其中，所述MAC-VLAN ID用于指示所述MAC/IP Advertisement路由携带的MAC地址所属的VLAN。所述MAC/IP Advertisement路由携带的MAC地址是所述CE设备的MAC地址，或者所述MAC/IP Advertisement路由携带的MAC地址是所述CE设备所管辖的以太网虚拟私有网络站点EVPN site中的主机的MAC地址。所述处理单元1004还用于根据所述MAC/IP Advertisement路由携带的ESI和所述MAC-VLAN ID确定连接所述CE设备与所述第一NVE设备之间的以太网链路的所述第一NVE的本地接口，其中，所述MAC/IP Advertisement路

由携带的ESI用于指示所述CE设备与所述第二NVE设备之间的所述ES; 所述处理单元1004还用于触发向所述第一NVE设备的转发平面发送第一MAC表项, 所述第一MAC表项中的出接口为所述第一NVE设备的本地接口, 其中, 所述第一MAC表项包括所述MAC/IP Advertisement路由携带的MAC地址和所述MAC-VLAN ID。

5 可选的, 所述MAC/IP Advertisement路由被携带在MP_REACH_NLRI中。所述MP_REACH_NLRI中的下一跳域包括公共VTEP地址。所述MAC/IP Advertisement路由还包括所述第二VTEP地址, 所述MP_REACH_NLRI中的下一跳域包括的公共VTEP地址与所述第二IMET路由中的Originating Router's IP Address域包括的公共VTEP地址相同。所述处理单元1004还用于根据所述MP_REACH_NLRI中的下一跳域包括的公共VTEP地址和所述MAC/IP Advertisement路由包括的所述第二VTEP地址, 触发向所述第一NVE设备的转发平面发送第二MAC表项, 所述第二MAC表项中的出接口为连接所述第一NVE设备到所述第二NVE设备的VXLAN隧道的所述第一NVE设备的本地接口, 其中, 所述第二MAC表项包括所述MAC/IP Advertisement路由携带的MAC地址。

15 可选的, CE设备连接所述第二NVE设备。所述接收单元1002还用于接收来自所述第二NVE设备MAC/IP Advertisement路由, 所述MAC/IP Advertisement路由被携带在MP_REACH_NLRI中, 所述MP_REACH_NLRI中的下一跳域包括公共VTEP地址, 所述MAC/IP Advertisement路由包括MAC-VLAN ID和所述第二VTEP地址, 其中, 所述MAC-VLAN ID用于指示所述MAC/IP Advertisement路由携带的MAC地址所属的VLAN, 所述MAC/IP Advertisement路由携带的MAC地址是所述CE设备的MAC地址, 或者所述MAC/IP Advertisement路由携带的MAC地址是所述CE设备所管辖的EVPN site中的主机的MAC地址; 所述处理单元1004还用于根据所述MAC/IP Advertisement路由携带的ESI和所述MAC-VLAN ID确定所述第一NVE设备中不存在连接所述CE设备与所述第一NVE设备之间的以太网链路的本地接口。其中, 所述MAC/IP Advertisement路由携带的ESI用于指示所述CE设备与所述第二NVE设备之间的ES。所述处理单元1004还用于根据所述MP_REACH_NLRI中的下一跳域包括的公共VTEP地址和所述MAC/IP Advertisement路由包括的所述第二VTEP地址, 向所述第一NVE设备的转发平面发送MAC表项。所述MAC表项的出接口为连接所述第一NVE设备到所述第二NVE设备的VXLAN隧道的所述第一NVE设备的本地接口。其中, 所述MAC表项包括所述MAC/IP Advertisement路由携带的MAC地址。

30 图11所示的第一NVE设备可以执行上述实施例的方法中第一NVE设备执行的相应步骤。确保EVPN的应用场景没有限制在NVE设备之间的链路需采用物理直连链路的约束条件, 有助于扩大EVPN的应用场景。另外, 有助于提高可靠性和降低部署复杂程度。应当理解, 所述图11中的结构同样适用于所述图1中的第二NVE设备和第三NVE设备。

图12为本申请实施例的第一NVE设备1100的硬件结构示意图。图12所示的第一NVE设备1100可以执行上述实施例的方法中第一NVE设备执行的相应步骤。

35 如图12所示, 所述第一NVE设备1100包括处理器1101、存储器1102、接口1103和总线1104。其中接口1103可以通过无线或有线的方式实现, 具体来讲可以是网卡。上述处理器1101、存储器1102和接口1103通过总线1104连接。

所述接口1103具体可以包括发送器和接收器, 用于第一NVE设备与上述实施例中的第二NVE设备和第三NVE设备之间收发信息; 或者用于第一NVE设备与PE设备之间收发信息;

或者用于第一NVE设备与所述第一NVE设备连接的CE设备之间收发信息。作为举例，所述接口1103用于支持图2中的过程S102和S105。所述处理器1101用于执行上述实施例中由第一NVE设备进行的处理。例如，所述处理器1101用于比较接收的IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与存储器1102中保存的公共VTEP地址是否相同；还用于基于接收的IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与存储器1102中保存的公共VTEP地址相同，根据所述接收的IMET路由中的VTEP地址建立VXLAN隧道；还用于处理接收到的ES路由、MAC/IP Advertisement路由或Ethernet A-D per ES路由；还用于处理BUM流量或单播流量；和/或用于本文所描述的技术的其他过程。作为举例，所述处理器1101用于支持图2中的过程S103和S104。存储器1102包括操作系统11021和应用程序11022，用于存储程序、代码或指令，当处理器或硬件设备执行这些程序、代码或指令时可以完成方法实施例中涉及第一NVE设备的处理过程。可选的，所述存储器1102可以包括只读存储器（英文：Read-only Memory，缩写：ROM）和随机存取存储器（英文：Random Access Memory，缩写：RAM）。其中，所述ROM包括基本输入/输出系统（英文：Basic Input/Output System，缩写：BIOS）或嵌入式系统；所述RAM包括应用程序和操作系统。当需要运行第一NVE设备1100时，通过固化在ROM中的BIOS或者嵌入式系统中的bootloader引导系统进行启动，引导第一NVE设备1100进入正常运行状态。在第一NVE设备1100进入正常运行状态后，运行在RAM中的应用程序和操作系统，从而，完成方法实施例中涉及第一NVE设备的处理过程。

可以理解的是，图12仅仅示出了第一NVE设备的简化设计。在实际应用中，第一NVE设备可以包含任意数量的接口，处理器或者存储器。另外，本实施例仅以第一NVE设备为例进行说明，应当理解，第二NVE设备、第三NVE设备，或者更多的NVE设备具有与所述第一NVE设备相同的功能，此处不再一一赘述。

图13为本申请实施例的另第一NVE设备1200的硬件结构示意图。图13所示的第一NVE设备1200可以执行上述实施例的方法中第一NVE设备执行的相应步骤。

如图13所述，第一NVE设备1200包括：主控板1210、接口板1230、交换网板1220和接口板1240。主控板1210用于完成系统管理、设备维护、协议处理等功能。交换网板1220用于完成各接口板（接口板也称为线卡或业务板）之间的数据交换。接口板1230和1240用于提供各种业务接口（例如，POS接口、GE接口、ATM接口等），并实现数据包的转发。主控板1210、接口板1230和1240，以及交换网板1220之间通过系统总线与系统背板相连实现互通。接口板1230上的中央处理器1231用于对接口板进行控制管理并与主控板上的中央处理器进行通信。

接口板1230上的物理接口卡1233接收来自所述第二NVE设备的第二IMET路由，并且，经由接口板1230上的中央处理器1231向主控板1210上的中央处理器1211上送所述第二IMET路由。

主控板1210上的中央处理器1211用于获取所述第二IMET路由。中央处理器1211还用于确定所述第二IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与所述第一NVE设备保存的公共VTEP地址是否相同。当中央处理器1211确定所述第二IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与所述第一NVE设备保存的公共VTEP地址相同时，所述中央处理器1211根据所述第二IMET路由中的所述第二VTEP地址建立从所述第一NVE设备到所述第二NVE设备的虚拟扩展局域网VXLAN隧道。其中，

所述第二IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与所述第二VTEP地址不相同。

主控板1210上的中央处理器1211还用于生成第一IMET路由。所述第一IMET路由的Originating Router's IP Address域中携带所述第一NVE设备保存的公共VTEP地址。其中，所述
5 所述第一IMET路由还包括第一VTEP地址。所述第一VTEP地址被所述第二NVE设备用于建立从所述第二NVE设备到所述第一NVE设备的VXLAN隧道。所述第一NVE设备保存的公共VTEP地址与所述第一VTEP地址不相同。主控板1210上的中央处理器1211将生成的第一IMET路由经由接口板1230上的中央处理器1231向物理接口卡1233发送。接口板1230上的物理接口卡1233向所述第一NVE设备的BGP peer发送所述第一IMET路由。

10 主控板1210上的中央处理器1211还用于从接口板1230上的物理接口卡1233获取来自所述第二NVE设备和所述第三NVE设备的ES路由。中央处理器1211还用于根据所述ES路由将所述第一NVE设备确定为ES中的DF或NDF。

主控板1210上的中央处理器1211还用于控制接口板1230完成BUM流量或单播流量的转发和处理。

15 主控板1210上的中央处理器1211还用于生成Ethernet A-D per ES路由，并且，通过接口板1230上的物理接口卡1233向所述第一NVE设备的BGP peer发送所述Ethernet A-D per ES路由。

主控板1210上的中央处理器1211还用于从接口板1230上的物理接口卡1233获取来自
20 所述第二NVE设备的MAC/IP Advertisement路由，并且，根据所述MAC/IP Advertisement路由携带的ESI和所述MAC-VLAN ID生成MAC表项。主控板1210上的中央处理器1211经由接口板1230上的中央处理器1231向接口板1230上的转发表项存储器1234传送所述MAC表项。

接口板1230上的转发表项存储器1234用于保存所述MAC表项。接口板1230上的中央
25 处理器1231用于控制网络存储器1232获取转发表项存储器1234中的MAC表项，并且，中央处理器1231用于控制网络存储器1232经由物理接口卡1233完成流量的接收和发送。

应理解，本发明实施例中接口板1240上的操作与所述接口板1230的操作一致，为了简洁，不再赘述。应理解，本实施例的第一NVE设备1200可对应于上述方法实施例所具有的功能和/或所实施的各种步骤。为了简洁，在此不再赘述。另外，本实施例仅以第一NVE
30 设备为例进行说明，应当理解，第二NVE设备、第三NVE设备，或者更多的NVE设备具有与所述第一NVE设备相同的功能，此处不再一一赘述。

此外，需要说明的是，主控板可能有一块或多块，有多块的时候可以包括主用主控板和备用主控板。接口板可能有一块或多块，第一NVE设备的数据处理能力越强，提供的接口板越多。接口板上的物理接口卡也可以有一块或多块。交换网板可能没有，也可能有一块或多块，有多块的时候可以共同实现负荷分担冗余备份。在集中式转发架构下，第一NVE
35 设备可以不需要交换网板，接口板承担整个系统的业务数据的处理功能。在分布式转发架构下，第一NVE设备可以有至少一块交换网板，通过交换网板实现多块接口板之间的数据交换，提供大容量的数据交换和处理能力。所以，分布式架构的第一NVE设备的数据接入和处理能力要大于集中式架构的设备。具体采用哪种架构，取决于具体的组网部署场景，此处不做任何限定。

40 图14为本申请实施例的又第一NVE设备1300的硬件结构示意图。图14所示的第一NVE

设备1300可以执行上述实施例的方法中第一NVE设备执行的相应步骤。

第一NVE设备1300的这种产品形态适用于基于控制与转发分离的网络架构（例如，软件定义网络（英文：Software Defined Network，缩写：SDN））。在SDN中，如图13所示的第一PE设备1200的主控板1210从设备中分离出来，形成新的独立的物理设备（即如图14所示的控制器1210A），剩下的形成另一独立的物理设备（即如图14所示的第一NVE转发设备1200A）。控制器1210A与第一NVE转发设备1200A通过控制通道协议实现交互。控制通道协议可以是开放流（英文：OpenFlow）协议、路径计算通信协议（英文：Path Computation Element Communication Protocol，缩写：PCEP）、BGP、路由系统接口（英文：Interface to the Routing System，缩写：I2RS）等。也就是说，与上述图13所对应的实施例相比，本实施中的第一NVE设备1300包括分离出去的控制器1210A和第一NVE转发设备1200A，即在该实施例中，该第一NVE设备1300也可以看成是一个系统。

控制器1210A可以是基于通用的物理服务器实现或者是专用的硬件结构实现，在一个设计示例中，所述控制器包括接收器、处理器、发送器、RAM、ROM以及总线（图中未示出）。其中，处理器通过总线分别耦接接收器、发送器、RAM以及ROM。其中，当需要运行控制器时，通过固化在ROM中的BIOS或者嵌入式系统中的bootloader引导系统进行启动，引导控制器进入正常运行状态。在控制器进入正常运行状态后，在RAM中运行应用程序和操作系统，使得该处理器执行上述图13中主控板1210的所有功能和步骤。

第一NVE转发设备1200A可以是基于专用的硬件结构实现，其功能和结构与上述图13中的接口板1230、接口板1240和交换网板1220的功能和结构保持一致，执行相应的功能和步骤。也可以是基于通用的物理服务器和网络功能虚拟化（英文：Network Function Virtualization，缩写：NFV）技术实现的虚拟第一NVE转发设备，所述虚拟第一NVE转发设备为虚拟路由器。在虚拟第一NVE转发设备的场景下，上述实体第一NVE转发设备实施例中提到的该第一NVE转发设备包括接口板、交换网板以及处理器在虚拟环境下可以认为是其所基于通用的物理服务器分配给该虚拟第一NVE转发设备所使用的接口资源、网络资源以及处理资源。采用通用物理服务器实施该第一转发NVE设备的功能或步骤，或者采用通用物理服务器并利用NFV技术实施该第一转发NVE设备的功能或步骤具体可以参考图12的实施例。

应理解，本实施例中第一NVE设备1300中的控制器1210A和该第一NVE转发设备1200A可以实现方法实施例中的第一NVE设备所实施的各种功能、步骤，为了简洁，在此不再赘述。另外，本实施例仅以第一NVE设备为例进行说明，应当理解，第二NVE设备、第三NVE设备，或者更多的NVE设备具有与所述第一NVE设备相同的功能，此处不再一一赘述。

另外，本申请实施例提供了一种计算机存储介质，用于储存为上述第一NVE设备所用的计算机软件指令，其包含用于执行上述方法实施例所设计的程序。

如图1所示，本申请实施例还包括一种处理路由的网络系统，所述网络系统包括至少两台网络虚拟边缘NVE设备，所述至少两台NVE设备中的每台NVE设备为前述图11或图12或图13或图14中的第一NVE设备。

结合本申请公开内容所描述的方法或者算法的步骤可以硬件的方式来实现，也可以是由处理器执行软件指令的方式来实现。软件指令可以由相应的软件模块组成，软件模块可以被存放于RAM存储器、闪存、ROM存储器、EPROM存储器、EEPROM存储器、寄存器、

硬盘、移动硬盘、CD-ROM或者本领域熟知的任何其它形式的存储介质中。一种示例性的存储介质耦合至处理器，从而使处理器能够从该存储介质读取信息，且可向该存储介质写入信息。当然，存储介质也可以是处理器的组成部分。处理器和存储介质可以位于ASIC中。另外，该ASIC可以位于用户设备中。当然，处理器和存储介质也可以作为分立组件存在于用户设备中。

本领域技术人员应该可以意识到，在上述一个或多个示例中，本申请所描述的功能可以用硬件、软件、固件或它们的任意组合来实现。当使用软件实现时，可以将这些功能存储在计算机可读介质中或者作为计算机可读介质上的一个或多个指令或代码进行传输。计算机可读介质包括计算机存储介质和通信介质，其中通信介质包括便于从一个地方向另一个地方传送计算机程序的任何介质。存储介质可以是通用或专用计算机能够存取的任何可用介质。

以上所述的具体实施方式，对本申请的目的、技术方案和有益效果进行了进一步详细说明。所应理解的是，以上所述仅为本申请的具体实施方式而已。

权 利 要 求

1、一种处理路由的方法,所述方法应用于虚拟扩展局域网VXLAN中,所述VXLAN包括第一网络虚拟边缘NVE设备和第二NVE设备,其特征在于,所述方法包括:

所述第一NVE设备接收来自所述第二NVE设备的第二集成组播以太网标签IMET路由,所述第二IMET路由中的发端路由器因特网协议地址Originating Router's IP Address域包括公共虚拟扩展局域网隧道端点VTEP地址,其中,所述第二IMET路由还包括第二VTEP地址,所述第二IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与所述第二VTEP地址不相同;

所述第一NVE设备确定所述第二IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与所述第一NVE设备保存的公共VTEP地址是否相同,其中,所述第一NVE设备能够向所述第一NVE设备的边界网关协议对等体BGP peer发送Originating Router's IP Address域中携带所述第一NVE设备保存的公共VTEP地址的IMET路由,所述第二IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与所述第一NVE设备保存的公共VTEP地址相同;

当所述第一NVE设备确定所述第二IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与所述第一NVE设备保存的公共VTEP地址相同时,所述第一NVE设备根据所述第二IMET路由中的所述第二VTEP地址建立从所述第一NVE设备到所述第二NVE设备的VXLAN隧道。

2、根据权利要求1所述的方法,其特征在于,所述方法还包括:

所述第一NVE设备向所述第二NVE设备发送第一IMET路由,所述第一IMET路由的Originating Router's IP Address域中携带所述第一NVE设备保存的公共VTEP地址,其中,所述第一IMET路由还包括第一VTEP地址,所述第一VTEP地址被所述第二NVE设备用于建立从所述第二NVE设备到所述第一NVE设备的VXLAN隧道,所述第一NVE设备保存的公共VTEP地址与所述第一VTEP地址不相同。

3、根据权利要求2所述的方法,所述VXLAN还包括运营商边缘PE设备,其特征在于,所述方法还包括:

所述第一NVE设备向所述PE设备发送所述第一IMET路由;

所述第一NVE设备接收来自所述PE设备的IMET路由,所述来自所述PE设备的IMET路由中的Originating Router's IP Address域携带所述PE设备的VTEP地址;

所述第一NVE设备根据所述来自所述PE设备的IMET路由中的所述PE设备的VTEP地址建立从所述第一NVE设备到所述PE设备的VXLAN隧道,所述PE设备的VTEP地址不等于所述第一NVE设备保存的公共VTEP地址。

4、根据权利要求2或3所述的方法,其特征在于,

所述第一IMET路由中的VXLAN属性携带所述第一VTEP地址;

所述第二IMET路由中的VXLAN属性携带所述第二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,所述VXLAN还包括第三

NVE设备, 所述方法还包括:

所述第一NVE设备接收来自所述第三NVE设备的第三IMET路由, 所述第三IMET路由中的Originating Router's IP Address域包括公共VTEP地址, 其中, 所述第三IMET路由还包括第三VTEP地址, 所述第三IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与所述第三VTEP地址不相同;

所述第一NVE设备确定所述第三IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与所述第一NVE设备保存的公共VTEP地址是否相同, 其中, 所述第三IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与所述第一NVE设备保存的公共VTEP地址相同;

当所述第一NVE设备确定所述第三IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与所述第一NVE设备保存的公共VTEP地址相同时, 所述第一NVE设备根据所述第三IMET路由中的所述第三VTEP地址建立从所述第一NVE设备到所述第三NVE设备的VXLAN隧道。

6、根据权利要求5所述的方法, 其特征在于, 用户边缘CE设备经由多个以太网链路多归属到所述第一NVE设备、所述第二NVE设备以及所述第三NVE设备, 所述多个以太网链路构成一个以太网段ES;

所述方法还包括:

所述第一NVE设备根据来自所述第二NVE设备和所述第三NVE设备的ES路由, 将所述第一NVE设备确定为所述ES中的指定转发者DF。

7、根据权利要求5或6所述的方法, 其特征在于, 所述方法还包括:

所述第一NVE设备经由所述第二NVE设备到所述第一NVE设备的VXLAN隧道接收所述第二NVE设备发送的广播、未知单播和组播BUM流量;

所述第一NVE设备基于所述BUM流量是所述第一NVE设备经由所述第二NVE设备到所述第一NVE设备的VXLAN隧道接收的, 避免经由所述第一NVE设备到所述第三NVE设备的VXLAN隧道向所述第三NVE设备转发所述BUM流量。

8、根据权利要求2-4中任一项所述的方法, 其特征在于, CE设备经由多个以太网链路双归属到所述第一NVE设备和所述第二NVE设备, 所述多个以太网链路构成一个ES;

所述方法还包括:

所述第一NVE设备根据来自所述第二NVE设备的ES路由, 将所述第一NVE设备确定为所述ES中的DF;

所述第一NVE设备向所述第二NVE设备发送以太网自动发现每以太网段Ethernet A-D per ES路由, 所述Ethernet A-D per ES路由携带所述第一NVE设备为所述CE设备与所述第一NVE设备之间的以太网链路分配的以太网段标识ESI标签;

所述第一NVE设备经由所述第二NVE设备到所述第一NVE设备的VXLAN隧道接收所述第二NVE设备发送的VXLAN报文, 所述VXLAN报文中的原始以太网载荷包括BUM流量, 所述VXLAN报文携带所述ESI标签, 其中, 所述BUM流量来自所述CE设

备;

所述第一NVE设备基于所述VXLAN报文携带所述ESI标签,避免经由所述CE设备与所述第一NVE设备之间的ES向所述CE设备转发所述BUM流量。

5 9、根据权利要求8所述的方法,其特征在于,所述ESI标签被封装于所述VXLAN报文中的VXLAN头和所述原始以太网载荷之间,所述VXLAN头携带指示信息,所述指示信息用于指示所述VXLAN报文携带所述ESI标签。

10 10、根据权利要求1-4中任一项所述的方法,其特征在于,CE设备经由多个以太网链路双归属到所述第一NVE设备和所述第二NVE设备,所述多个以太网链路构成一个ES;

所述方法还包括:

15 所述第一NVE设备接收来自所述第二NVE设备的媒体接入控制/因特网协议通告MAC/IP Advertisement路由,所述MAC/IP Advertisement路由包括媒体接入控制-虚拟局域网标识MAC-VLAN ID,其中,所述MAC-VLAN ID用于指示所述MAC/IP Advertisement路由携带的MAC地址所属的VLAN,所述MAC/IP Advertisement路由携带的MAC地址是所述CE设备的MAC地址,或者所述MAC/IP Advertisement路由携带的MAC地址是所述CE设备所管辖的以太网虚拟私有网络站点EVPN site中的主机的MAC地址;

20 所述第一NVE设备根据所述MAC/IP Advertisement路由携带的ESI和所述MAC-VLAN ID确定连接所述CE设备与所述第一NVE设备之间的以太网链路的所述第一NVE的本地接口,其中,所述MAC/IP Advertisement路由携带的ESI用于指示所述CE设备与所述第二NVE设备之间的所述ES;

25 所述第一NVE设备的控制平面向所述第一NVE设备的转发平面发送第一MAC表项,所述第一MAC表项中的出接口为所述第一NVE设备的本地接口,其中,所述第一MAC表项包括所述MAC/IP Advertisement路由携带的MAC地址和所述MAC-VLAN ID。

30 11、根据权利要求10所述的方法,其特征在于,所述MAC/IP Advertisement路由被携带在多协议可到达网络层可达信息MP_REACH_NLRI中,所述MP_REACH_NLRI中的下一跳域包括公共VTEP地址,所述MAC/IP Advertisement路由还包括所述第二VTEP地址,所述MP_REACH_NLRI中的下一跳域包括的公共VTEP地址与所述第二IMET路由中的Originating Router's IP Address域包括的公共VTEP地址相同;

所述方法还包括:

35 所述第一NVE设备的控制平面根据所述MP_REACH_NLRI中的下一跳域包括的公共VTEP地址和所述MAC/IP Advertisement路由包括的所述第二VTEP地址,向所述第一NVE设备的转发平面发送第二MAC表项,所述第二MAC表项中的出接口为连接所述第一NVE设备到所述第二NVE设备的VXLAN隧道的所述第一NVE设备的本地接口,其中,所述第二MAC表项包括所述MAC/IP Advertisement路由携带的MAC地址。

40 12、根据权利要求1-4中任一项所述的方法,其特征在于,CE设备连接所述第二

NVE设备;

所述方法还包括:

所述第一NVE设备接收来自所述第二NVE设备MAC/IP Advertisement路由, 所述MAC/IP Advertisement路由被携带在MP_REACH_NLRI中, 所述MP_REACH_NLRI中的下一跳域包括公共VTEP地址, 所述MAC/IP Advertisement路由包括MAC-VLAN ID和所述第二VTEP地址, 其中, 所述MAC-VLAN ID用于指示所述MAC/IP Advertisement路由携带的MAC地址所属的VLAN, 所述MAC/IP Advertisement路由携带的MAC地址是所述CE设备的MAC地址, 或者所述MAC/IP Advertisement路由携带的MAC地址是所述CE设备所管辖的EVPN site中的主机的MAC地址;

所述第一NVE设备根据所述MAC/IP Advertisement路由携带的ESI和所述MAC-VLAN ID确定所述第一NVE设备中不存在连接所述CE设备与所述第一NVE设备之间的以太网链路的本地接口, 其中, 所述MAC/IP Advertisement路由携带的ESI用于指示所述CE设备与所述第二NVE设备之间的ES;

所述第一NVE设备的控制平面根据所述MP_REACH_NLRI中的下一跳域包括的公共VTEP地址和所述MAC/IP Advertisement路由包括的所述第二VTEP地址, 向所述第一NVE设备的转发平面发送MAC表项, 所述MAC表项的出接口为连接所述第一NVE设备到所述第二NVE设备的VXLAN隧道的所述第一NVE设备的本地接口, 其中, 所述MAC表项包括所述MAC/IP Advertisement路由携带的MAC地址。

13、一种第一网络虚拟边缘NVE设备, 其特征在于, 所述第一NVE设备包括:

接收单元, 用于接收来自所述第二NVE设备的第二集成组播以太网标签IMET路由, 所述第二IMET路由中的发端路由器因特网协议地址Originating Router's IP Address域包括公共虚拟扩展局域网隧道端点VTEP地址, 其中, 所述第二IMET路由还包括第二VTEP地址, 所述第二IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与所述第二VTEP地址不相同;

处理单元, 用于确定所述第二IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与所述第一NVE设备保存的公共VTEP地址是否相同, 其中, 所述第一NVE设备能够向所述第一NVE设备的边界网关协议对等体BGP peer发送Originating Router's IP Address域中携带所述第一NVE设备保存的公共VTEP地址的IMET路由, 所述第二IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与所述第一NVE设备保存的公共VTEP地址相同;

当所述处理单元确定所述第二IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与所述第一NVE设备保存的公共VTEP地址相同时, 所述处理单元还用于根据所述第二IMET路由中的所述第二VTEP地址建立从所述第一NVE设备到所述第二NVE设备的虚拟扩展局域网VXLAN隧道。

14、根据权利要求13所述的第一NVE设备, 其特征在于, 所述第一NVE设备还包括:

发送单元, 用于向所述第二NVE设备发送第一IMET路由, 所述第一IMET路由的Originating Router's IP Address域中携带所述第一NVE设备保存的公共VTEP地址, 其中,

所述第一IMET路由还包括第一VTEP地址,所述第一VTEP地址被所述第二NVE设备用于建立从所述第二NVE设备到所述第一NVE设备的VXLAN隧道,所述第一NVE设备保存的公共VTEP地址与所述第一VTEP地址不相同。

- 5 15、根据权利要求14所述的第一NVE设备,其特征在于,
所述发送单元还用于向运营商边缘PE设备发送所述第一IMET路由;
所述接收单元还用于接收来自所述PE设备的IMET路由,所述来自所述PE设备的IMET路由中的Originating Router's IP Address域携带所述PE设备的VTEP地址;
所述处理单元还用于根据所述来自所述PE设备的IMET路由中的所述PE设备的
10 VTEP地址建立从所述第一NVE设备到所述PE设备的VXLAN隧道,所述PE设备的VTEP地址不等于所述第一NVE设备保存的公共VTEP地址。

- 16、根据权利要求14或15所述的第一NVE设备,其特征在于,
所述第一IMET路由中的VXLAN属性携带所述第一VTEP地址;
15 所述第二IMET路由中的VXLAN属性携带所述第二VTEP地址。

- 17、根据权利要求13-16中任一项所述的第一NVE设备,其特征在于,
所述接收单元还用于接收来自所述第三NVE设备的第三IMET路由,所述第三IMET路由中的Originating Router's IP Address域包括公共VTEP地址,其中,所述第三
20 IMET路由还包括第三VTEP地址,所述第三IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与所述第三VTEP地址不相同;

- 所述处理单元还用于确定所述第三IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与所述第一NVE设备保存的公共VTEP地址是否相同,其中,所述第三IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与所述
25 第一NVE设备保存的公共VTEP地址相同;

- 当所述处理单元确定所述第三IMET路由中的Originating Router's IP Address域包括的公共VTEP地址与所述第一NVE设备保存的公共VTEP地址相同时,所述处理单元还用于根据所述第三IMET路由中的所述第三VTEP地址建立从所述第一NVE设备到所述第三NVE设备的VXLAN隧道。
30

- 18、根据权利要求17所述的第一NVE设备,其特征在于,用户边缘CE设备经由多个以太网链路多归属到所述第一NVE设备、所述第二NVE设备以及所述第三NVE设备,所述多个以太网链路构成一个以太网段ES;

- 所述处理单元还用于根据来自所述第二NVE设备和所述第三NVE设备的ES路由,
35 将所述第一NVE设备确定为所述ES中的指定转发者DF。

- 19、根据权利要求17或18所述的第一NVE设备,其特征在于,
所述接收单元还用于经由所述第二NVE设备到所述第一NVE设备的VXLAN隧道接收所述第二NVE设备发送的广播、未知单播和组播BUM流量;
40 所述处理单元还用于基于所述BUM流量是所述第一NVE设备经由所述第二NVE

设备到所述第一NVE设备的VXLAN隧道接收的,避免经由所述第一NVE设备到所述第三NVE设备的VXLAN隧道向所述第三NVE设备转发所述BUM流量。

20、根据权利要求14-16中任一项所述的第一NVE设备方法,其特征在于,CE设备经由多个以太网链路双归属到所述第一NVE设备和所述第二NVE设备,所述多个以太网链路构成一个ES;

所述处理单元还用于根据来自所述第二NVE设备的ES路由,将所述第一NVE设备确定为所述ES中的DF;

所述发送单元还用于向所述第二NVE设备发送以太网自动发现每以太网段Ethernet A-D per ES路由,所述Ethernet A-D per ES路由携带所述第一NVE设备为所述CE设备与所述第一NVE设备之间的以太网链路分配的以太网段标识ESI标签;

所述接收单元还用于经由所述第二NVE设备到所述第一NVE设备的VXLAN隧道接收所述第二NVE设备发送的VXLAN报文,所述VXLAN报文中的原始以太网载荷包括BUM流量,所述VXLAN报文携带所述ESI标签,其中,所述BUM流量来自所述CE设备;

所述处理单元还用于基于所述VXLAN报文携带所述ESI标签,避免经由所述CE设备与所述第一NVE设备之间的ES向所述CE设备转发所述BUM流量。

21、根据权利要求20所述的第一NVE设备,其特征在于,所述ESI标签被封装于所述VXLAN报文中的VXLAN头和所述原始以太网载荷之间,所述VXLAN头携带指示信息,所述指示信息用于指示所述VXLAN报文携带所述ESI标签。

22、根据权利要求13-16中任一项所述的第一NVE设备,其特征在于,CE设备经由多个以太网链路双归属到所述第一NVE设备和所述第二NVE设备,所述多个以太网链路构成一个ES;

所述接收单元还用于接收来自所述第二NVE设备的媒体接入控制/因特网协议通告MAC/IP Advertisement路由,所述MAC/IP Advertisement路由包括媒体接入控制-虚拟局域网标识MAC-VLAN ID,其中,所述MAC-VLAN ID用于指示所述MAC/IP Advertisement路由携带的MAC地址所属的VLAN,所述MAC/IP Advertisement路由携带的MAC地址是所述CE设备的MAC地址,或者所述MAC/IP Advertisement路由携带的MAC地址是所述CE设备所管辖的以太网虚拟私有网络站点EVPN site中的主机的MAC地址;

所述处理单元还用于根据所述MAC/IP Advertisement路由携带的ESI和所述MAC-VLAN ID确定连接所述CE设备与所述第一NVE设备之间的以太网链路的所述第一NVE的本地接口,其中,所述MAC/IP Advertisement路由携带的ESI用于指示所述CE设备与所述第二NVE设备之间的所述ES;

所述处理单元还用于触发向所述第一NVE设备的转发平面发送第一MAC表项,所述第一MAC表项中的出接口为所述第一NVE设备的本地接口,其中,所述第一MAC表项包括所述MAC/IP Advertisement路由携带的MAC地址和所述MAC-VLAN ID。

23、根据权利要求22所述的第一NVE设备，其特征在于，所述MAC/IP Advertisement路由被携带在多协议可到达网络层可达信息MP_REACH_NLRI中，所述MP_REACH_NLRI中的下一跳域包括公共VTEP地址，所述MAC/IP Advertisement路由还包括所述第二VTEP地址，所述MP_REACH_NLRI中的下一跳域包括的公共VTEP地址与
5 所述第二IMET路由中的Originating Router's IP Address域包括的公共VTEP地址相同；

所述处理单元还用于根据所述MP_REACH_NLRI中的下一跳域包括的公共VTEP地址和所述MAC/IP Advertisement路由包括的所述第二VTEP地址，触发向所述第一NVE设备的转发平面发送第二MAC表项，所述第二MAC表项中的出接口为连接所述第
10 一NVE设备到所述第二NVE设备的VXLAN隧道的所述第一NVE设备的本地接口，其中，所述第二MAC表项包括所述MAC/IP Advertisement路由携带的MAC地址。

24、根据权利要求13-16中任一项所述的第一NVE设备，其特征在于，CE设备连接所述第二NVE设备；

15 所述接收单元还用于接收来自所述第二NVE设备MAC/IP Advertisement路由，所述MAC/IP Advertisement路由被携带在MP_REACH_NLRI中，所述MP_REACH_NLRI中的下一跳域包括公共VTEP地址，所述MAC/IP Advertisement路由包括MAC-VLAN ID和所述第二VTEP地址，其中，所述MAC-VLAN ID用于指示所述MAC/IP Advertisement路由携带的MAC地址所属的VLAN，所述MAC/IP Advertisement路由携带
20 的MAC地址是所述CE设备的MAC地址，或者所述MAC/IP Advertisement路由携带的MAC地址是所述CE设备所管辖的EVPN site中的主机的MAC地址；

所述处理单元还用于根据所述MAC/IP Advertisement路由携带的ESI和所述MAC-VLAN ID确定所述第一NVE设备中不存在连接所述CE设备与所述第一NVE设备之间的以太网链路的本地接口，其中，所述MAC/IP Advertisement路由携带的ESI用于
25 指示所述CE设备与所述第二NVE设备之间的ES；

所述处理单元还用于根据所述MP_REACH_NLRI中的下一跳域包括的公共VTEP地址和所述MAC/IP Advertisement路由包括的所述第二VTEP地址，触发向所述第一NVE设备的转发平面发送MAC表项，所述MAC表项的出接口为连接所述第一NVE设备到所述第二NVE设备的VXLAN隧道的所述第一NVE设备的本地接口，其中，所述MAC
30 表项包括所述MAC/IP Advertisement路由携带的MAC地址。

25、一种处理路由的网络系统，所述网络系统包括至少两台网络虚拟边缘NVE设备，所述至少两台NVE设备中的每台NVE设备为权利要求13至24中任一所述的第一NVE设备。

35

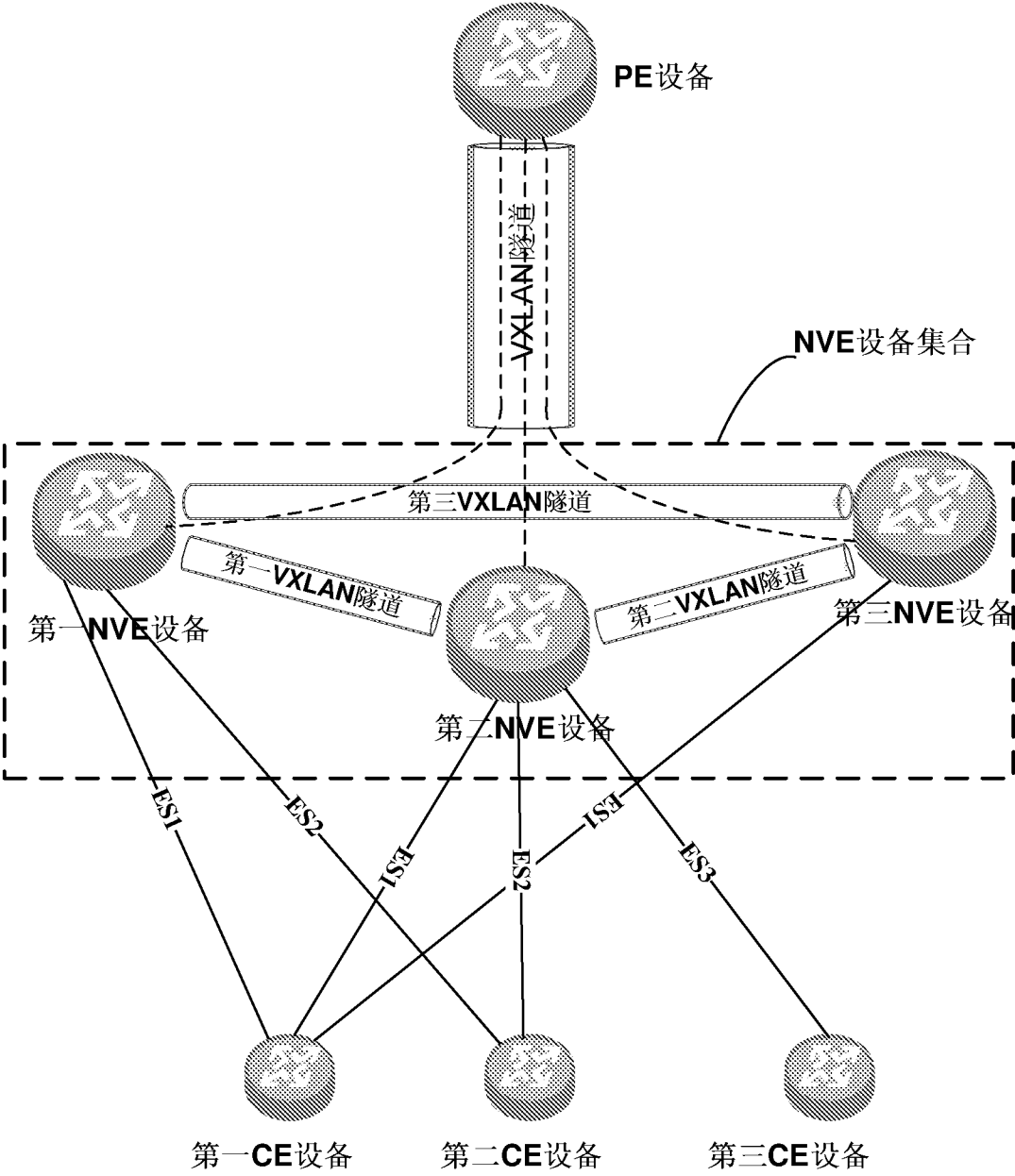


图 1

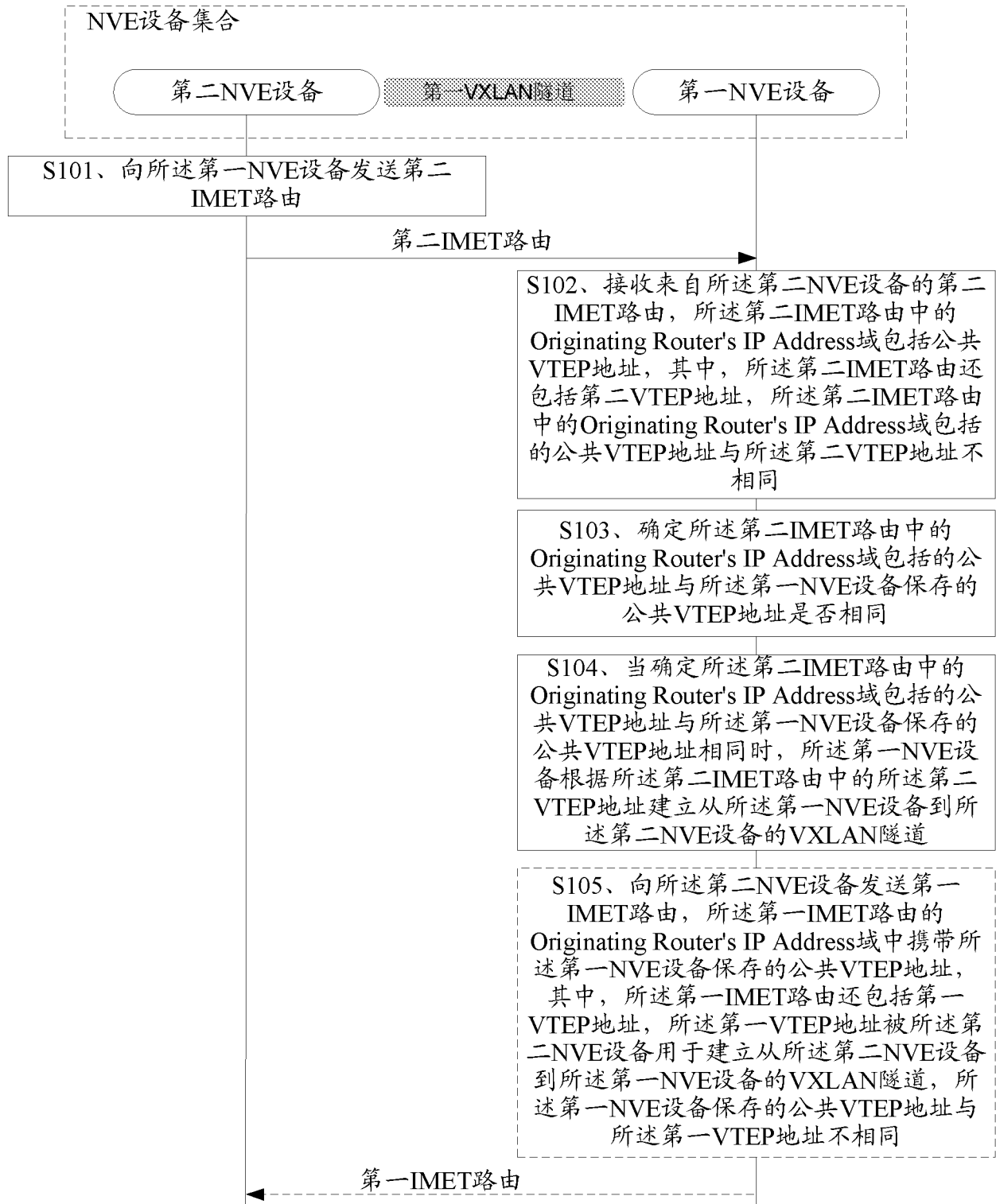


图 2

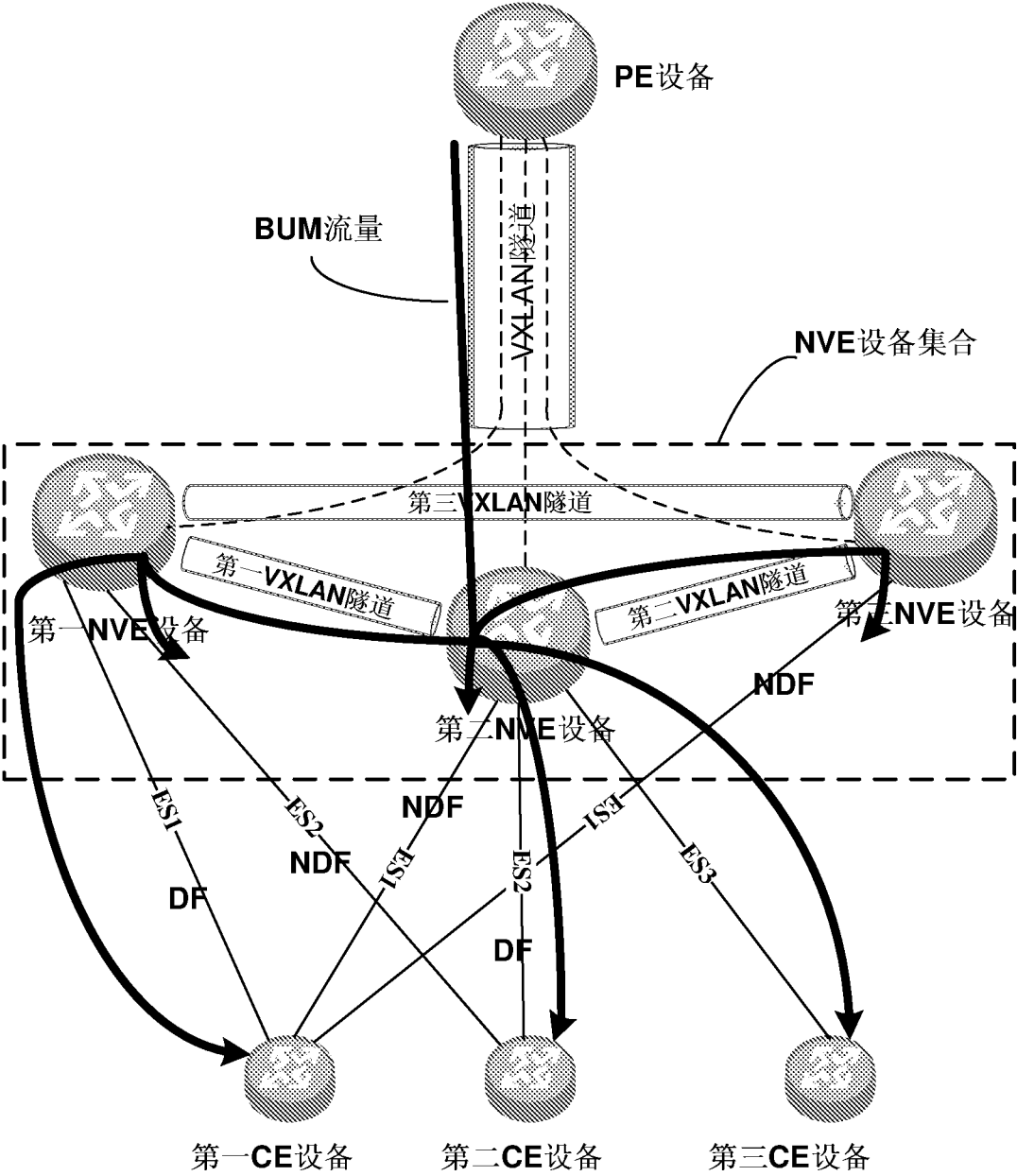


图 3

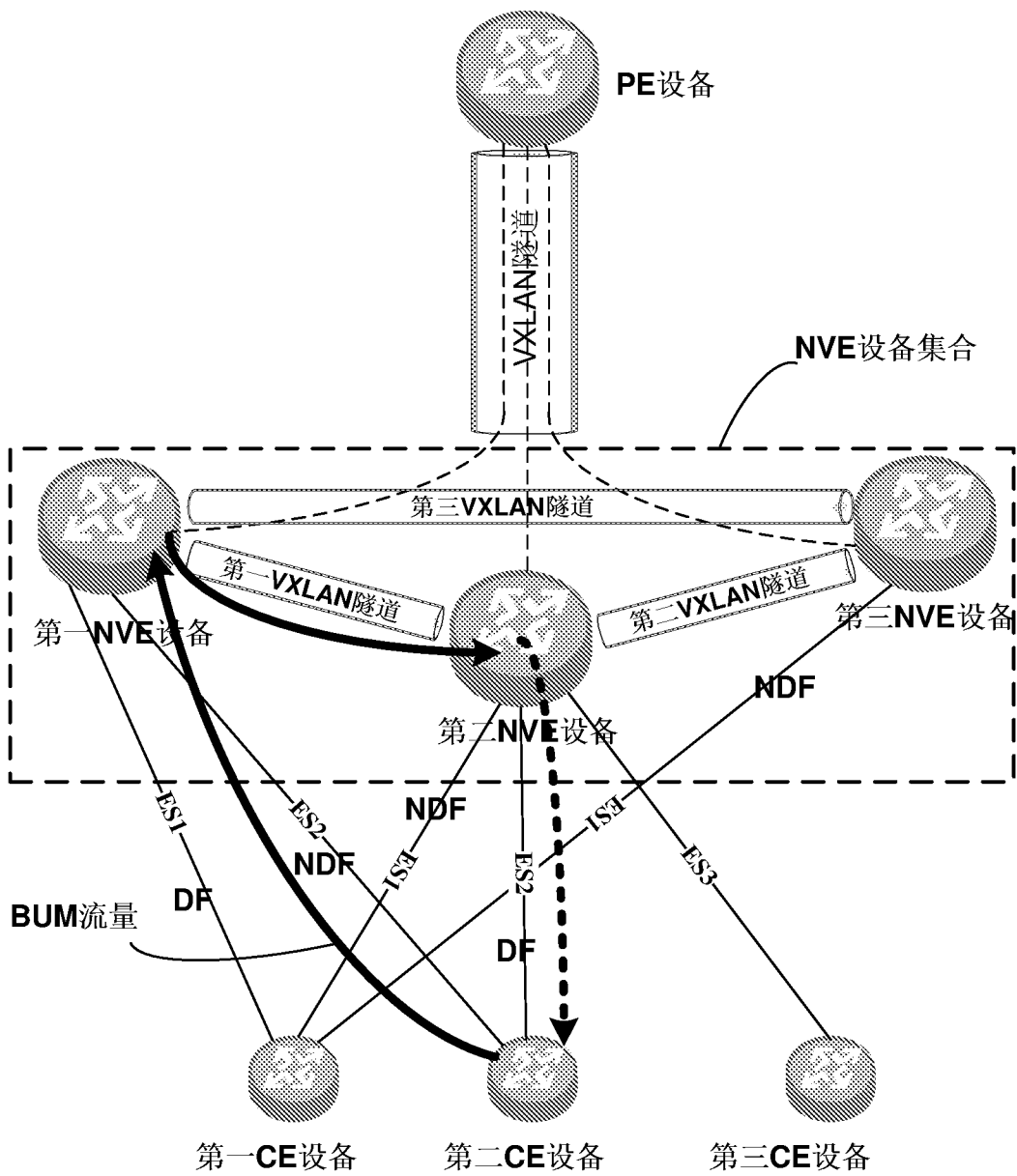


图 4

外层Eth头	外层IP头	外层UDP头	VXLAN头	ESI标签	原始以太网载荷
--------	-------	--------	--------	-------	---------

图 5

0	1	2	3
0 1 2 3 4 5 6 7	8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1		
类型 (Type)	子类型 (Sub Type)	标志 (Flag)	预留 (Reserved)
VTEP地址 (VTEP address)			

图 6

0								1								2								3							
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1
类型 (Type)								子类型 (Sub Type)								标志 (Flag)								预留 (Reserved)							
VTEP VNI																															

图7

0								1								2								3							
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1
类型 (Type)								子类型 (Sub Type)								标志 (Flag)								预留 (Reserved)							
MAC-VLAN标识 (MAC-VLAN ID)																															

图8

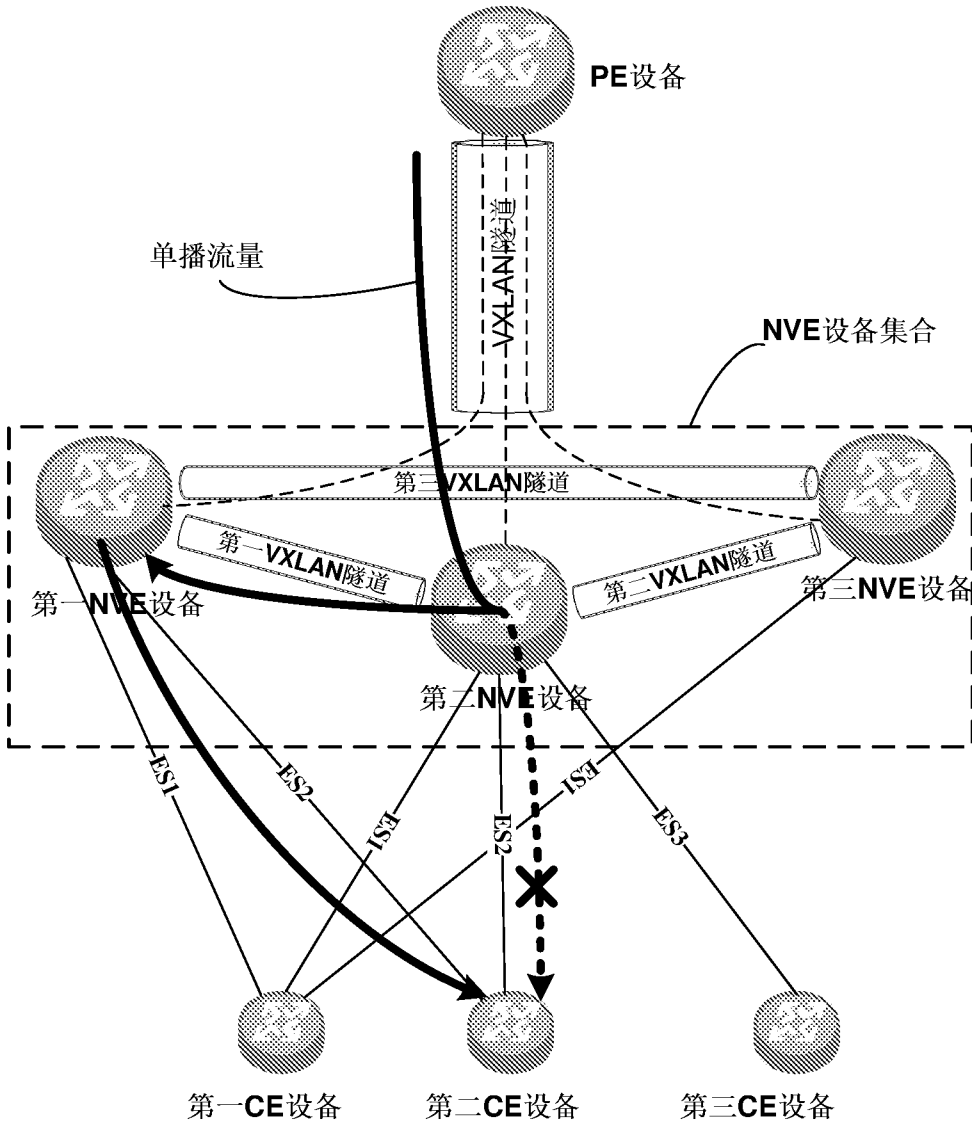


图9

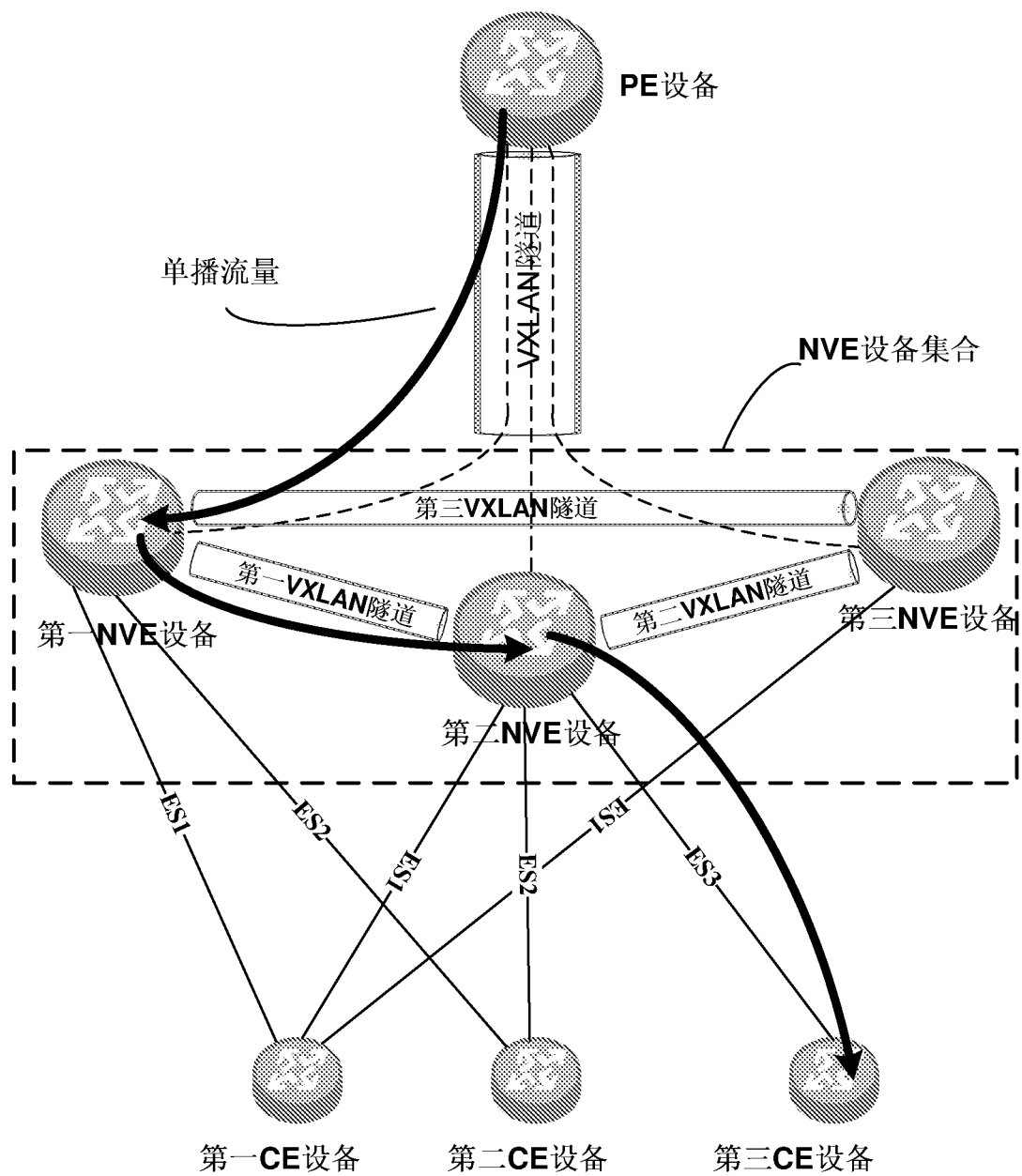


图10

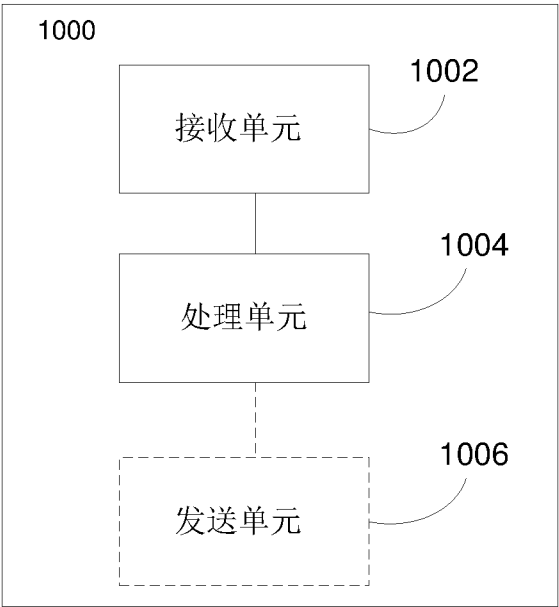


图11

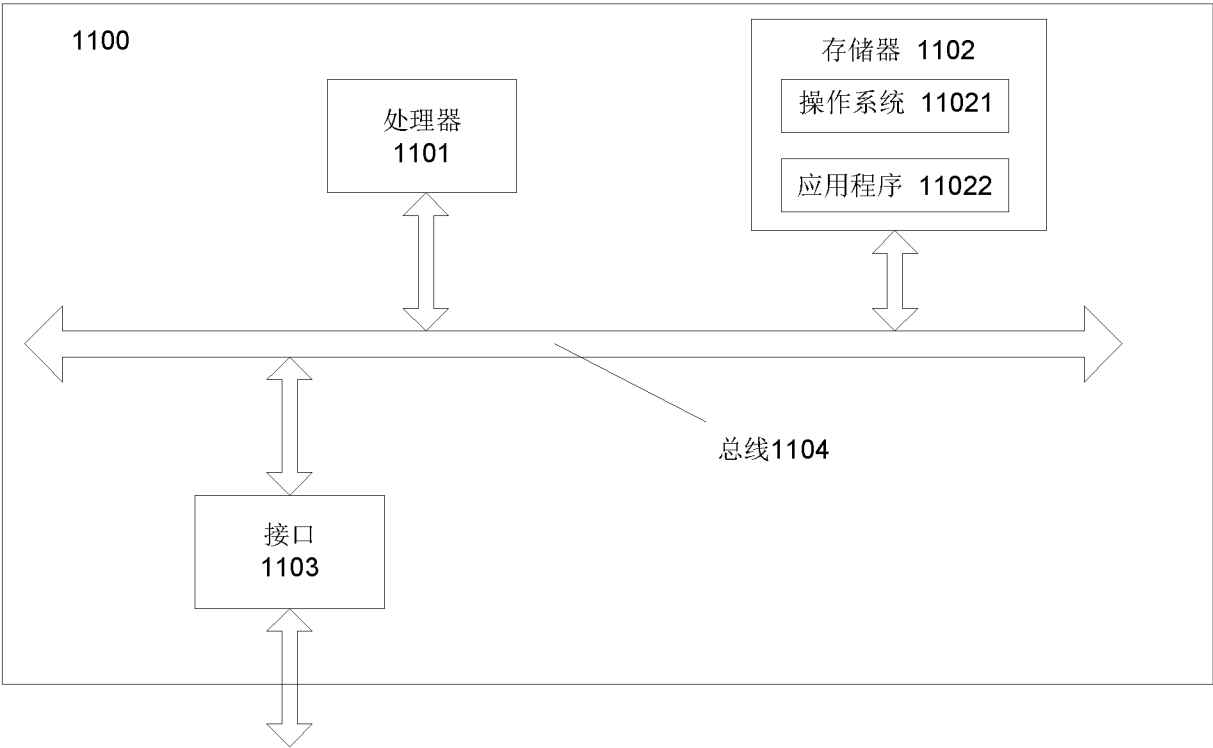


图 12

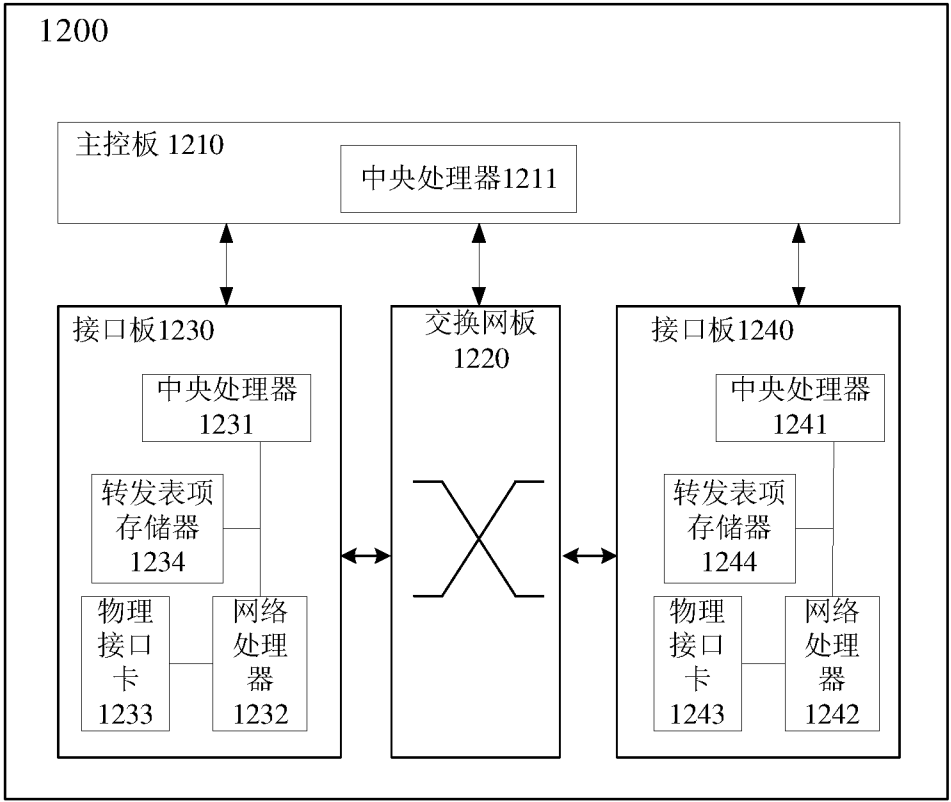


图 13

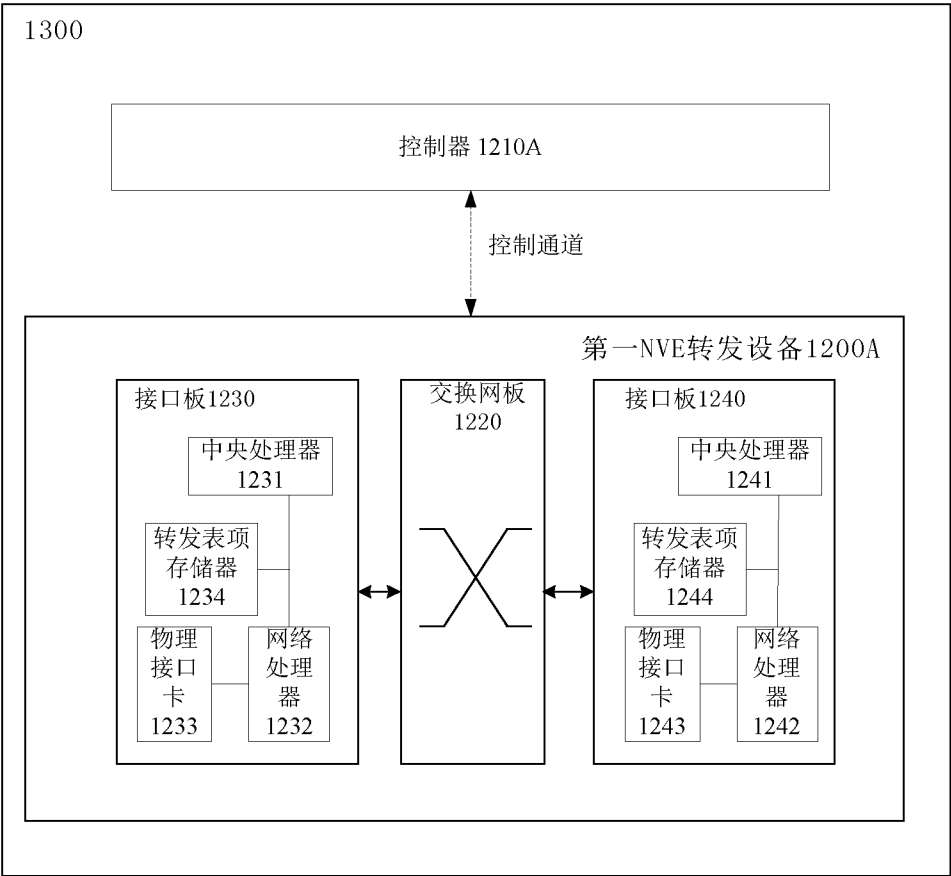


图 14

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2017/111608

A. CLASSIFICATION OF SUBJECT MATTER

H04L 12/46 (2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

H04L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNPAT, EPODOC, WPI, CNKI: 虚拟扩展局域网, 隧道端点, 隧道, 网络虚拟边缘, 建立, 路由, 组播以太网标签, VXLAN, VTEP, TEP, tunnel, NVE, set up, establish+, build+, route, routing, IMET

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 104702476 A (HUAWEI TECHNOLOGIES CO., LTD.), 10 June 2015 (10.06.2015), entire document	1-25
A	CN 105991432 A (JUNIPER NETWORKS INC.), 05 October 2016 (05.10.2016), entire document	1-25
A	CN 105591868 A (H3C TECHNOLOGIES CO., LIMITED), 18 May 2016 (18.05.2016), entire document	1-25
A	US 2015003458 A1 (FUTUREWEI TECHNOLOGIES INC.), 01 January 2015 (01.01.2015), entire document	1-25
A	US 2017019331 A1 (FUTUREWEI TECHNOLOGIES INC.), 19 January 2017 (19.01.2017), entire document	1-25

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:	“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
“A” document defining the general state of the art which is not considered to be of particular relevance	“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
“E” earlier application or patent but published on or after the international filing date	“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	“&” document member of the same patent family
“O” document referring to an oral disclosure, use, exhibition or other means	
“P” document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 18 January 2018	Date of mailing of the international search report 01 February 2018
Name and mailing address of the ISA State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No. (86-10) 62019451	Authorized officer ZHANG, Yujie Telephone No. (86-10) 53961590

International application No.
PCT/CN2017/111608

Form PCT/ISA/210 (patent family annex) (July 2009)

国际检索报告

国际申请号

PCT/CN2017/111608

A. 主题的分类

H04L 12/46(2006.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

H04L

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

CNPAT, EPODOC, WPI, CNKI:虚拟扩展局域网, 隧道端点, 隧道, 网络虚拟边缘, 建立, 路由, 组播以太网标签, VLAN, VTEP, TEP, tunnel, NVE, set up, establish+, build+, route, routing, IMET

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
A	CN 104702476 A (华为技术有限公司) 2015年 6月 10日 (2015 - 06 - 10) 全文	1-25
A	CN 105991432 A (丛林网络公司) 2016年 10月 5日 (2016 - 10 - 05) 全文	1-25
A	CN 105591868 A (杭州华三通信技术有限公司) 2016年 5月 18日 (2016 - 05 - 18) 全文	1-25
A	US 2015003458 A1 (FUTUREWEI TECHNOLOGIES INC.) 2015年 1月 1日 (2015 - 01 - 01) 全文	1-25
A	US 2017019331 A1 (FUTUREWEI TECHNOLOGIES INC.) 2017年 1月 19日 (2017 - 01 - 19) 全文	1-25

☐ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

2018年 1月 18日

国际检索报告邮寄日期

2018年 2月 1日

ISA/CN的名称和邮寄地址

中华人民共和国国家知识产权局(ISA/CN)
中国北京市海淀区蓟门桥西土城路6号 100088

受权官员

张玉洁

传真号 (86-10)62019451

电话号码 (86-10)53961590

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2017/111608

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	104702476	A	2015年 6月 10日	无			
CN	105991432	A	2016年 10月 5日	EP	3070877	A1	2016年 9月 21日
				US	2016277210	A1	2016年 9月 22日
CN	105591868	A	2016年 5月 18日	无			
US	2015003458	A1	2015年 1月 1日	无			
US	2017019331	A1	2017年 1月 19日	无			