

# (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 2018/166253 A1**

- (51) 国际专利分类号:  
*H04L 12/721* (2013.01)
- (21) 国际申请号: PCT/CN2017/115821
- (22) 国际申请日: 2017年12月13日 (13.12.2017)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:  
201710151472.0 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)。庄顺万 (ZHUANG, Shunwan); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。刘乔 (LIU, Qiao); 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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(54) Title: EVPN PACKET PROCESSING METHOD, DEVICE AND SYSTEM

(54) 发明名称: EVPN报文处理方法、设备及系统

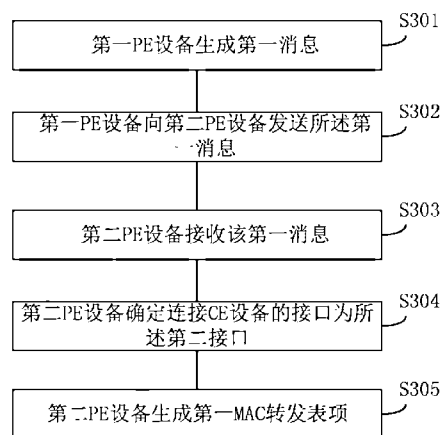


图 3

- S301 FIRST PE DEVICE GENERATES FIRST MESSAGE
- S302 FIRST PE DEVICE SENDS FIRST MESSAGE TO SECOND PE DEVICE
- S303 SECOND PE DEVICE RECEIVES FIRST MESSAGE
- S304 SECOND PE DEVICE DETERMINES THAT INTERFACE FOR CONNECTING CE DEVICE IS SECOND INTERFACE
- S305 SECOND PE DEVICE GENERATES FIRST MAC FORWARDING TABLE ITEM

(57) Abstract: An EVPN packet processing method, wherein a CE device accesses a first PE device and a second PE device in a multi-homing manner. The method comprises: the first PE device sends MAC routing and a VLAN identifier learned from the CE device to the second PE device, and the second PE device generates a MAC forwarding table item according to the MAC routing and the VLAN identifier, the MAC forwarding table item being used for directly forwarding, via the CE device, a packet for which a target MAC address is a MAC address of the CE device or a terminal device accessing the CE device. An outbound interface identifier



WO 2018/166253 A1

(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))。

---

included in the MAC forwarding table item is an identifier of an interface for connecting the CE device. The method enables the first PE device and the second PE device to form an effective load sharing for the reasonable utilisation of system bandwidth resources.

(57) 摘要: 一种EVPN中报文处理方法, CE设备多归接入第一PE设备和第二PE设备, 所述方法包括: 所述第一PE设备将从所述CE设备学习到的MAC路由和VLAN标识发送给第二PE设备, 第二PE设备根据所述MAC路由和VLAN标识生成MAC转发表项, 用于直接通过所述CE设备转发目的MAC地址为所述CE设备或接入所述CE设备的终端设备的MAC地址的报文。所述MAC转发表项中包括的出接口标识为连接所述CE设备的接口的标识。通过本申请的方法, 能够使得第一PE设备和第二PE设备形成有效的负载分担, 合理利用系统带宽资源。

## EVPN 报文处理方法、设备及系统

5 本申请要求于 2017 年 03 月 14 日提交中国专利局、申请号为 201710151472.0、申请名称为“EVPN 报文处理方法、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。

## 技术领域

10 本申请实施例涉及通信技术领域，特别涉及一种以太网虚拟私有网（英文：Ethernet Virtual Private Network, EVPN）报文处理的方法、设备及系统。

## 背景技术

以太虚网虚拟私有网络（英文：Ethernet Virtual Private Network, EVPN）是一种在多协议标签交换（英文：Multi-Protocol Label Switching, MPLS）网络上提供二  
15 层网络互联的虚拟私有网络（英文：Virtual Private Network, VPN）技术。目前，EVPN 作为承载二层业务的主流方案在各大运营商的网络设计中被使用。EVPN 技术使用边界网关协议（英文：Border Gateway Protocol, BGP）作为控制平面的协议，进行运营商边缘（英文：Provider Edge, PE）设备之间的媒体接入控制（英文：Media Access Control, MAC）地址学习，将 MAC 地址学习和发布过程从传统的数据平面转移到控制平面，从而大幅减少  
20 了流量洪泛方式的 MAC 地址扩散，以及可以支持用户边缘设备（英文：Customer Edge, CE）多归属接入 EVPN、便于管理 MAC 地址实现负载分担。EVPN 中，CE 设备可通过多种接入方式与 PE 设备相联，例如，通过虚拟局域网（英文：Virtual Local Area Network, VLAN）多归属接入 EVPN。EVPN 方案的重要优势之一是实现了 CE 设备的多归属接入。

在一种常见的场景中，如图 1 所示，网络 100 网络 100 包括服务商提供的骨干网和多个  
25 个 E-VPN 站点。所述骨干网包括第一 PE 设备 PE1、第二 PE 设备 PE2 和第三 PE 设备 PE3 以及多个 P（Provider）设备（图中未示出）。所述多个 E-VPN 站点包括 site1 和 site2。其中，site1、和 site2 属于同一个 E-VPN1。MAC 地址为 MAC1 的终端设备 A 接入 CE1，CE1 双归接入 PE1 和 PE2。MAC 地址为 MAC2 的终端设备 B 接入 CE2，CE2 接入 PE3。其中，PE1 从 CE1 学习到终端设备 A 的 MAC 地址。PE2 未从 CE1 学习到终端设备 A 的 MAC 地址。PE1  
30 通过 BGP 更新（Update）消息向 PE3 发送 EVPN 媒体接入控制/互联网协议通告路由（英文：MAC/IP Advertisement Route），即向 PE3 通告到达所述终端设备 A 的 MAC 路由。而 PE2 虽然没有学习到 MAC1，但是 PE2 可以向 PE3 发布 EVPN 以太网自动发现路由（英文：Ethernet A-D Route），PE3 就可以知道通过 PE1 和 PE2 均可以到达终端设备 A，即 PE3 根据别名（英文：Aliasing）方式形成了负载分担。需要说明的是，图 1 中以终端设备 A 通过 CE1 接入 PE 设备为例进行说明，在实际组网中，也可以是终端设备 A 直接作为 CE 设备接入 PE  
35 设备，此时，终端设备 A 的 MAC 地址即为 CE 设备的 MAC 地址。本文中以终端设备 A 通过 CE 设备接入 PE 设备为例进行说明。对于终端设备直接接入 PE 设备的方案，与终端设备通过 CE 设备接入 PE 设备的方案类似，此处不再赘述。

关于 MAC/IP 通告路由以及 Ethernet A-D 路由的具体细节，以及 PE3 根据别名方式形成负载分担的细节可以参照国际互联网工程任务组（Internet Engineering Task Force, IETF）发布的征求意见（英文：Request For Comments, RFC）7432 中的具体说明，此处不再赘述。但是，在现有的 EVPN 技术中，当来自终端设备 B 发送的目的 MAC 地址为 MAC1 的已知单播流量到达 PE3。PE3 对该已知单播流量进行负载分担处理，并发送到 PE2。由于 PE2 未从 CE1 学习到终端设备 A 的 MAC 地址，PE2 无法直接向 CE1 转发上述单播流量，导致 PE2 需要经由 PE1 将流量发送到 CE1。也就是说，PE1 和 PE2 没有对目的 MAC 地址为 MAC1 的已知单播流量进行有效的负载分担。不仅浪费了 PE2 到 CE1 的带宽资源，也消耗了 PE1 到 CE1 的带宽资源。上述方案中，浪费了系统资源，不能有效发挥 EVPN 的技术优势。

## 10 发明内容

本申请提供了一种报文处理的方法，用于解决现有技术中 PE 设备无法直接向多归属接入该 PE 设备的 CE 设备转发流量，从而导致多个 PE 设备之间无法形成有效的负载分担的技术问题。

第一方面，本申请提供了一种报文处理方法，该方法应用于 EVPN 中，其中，用户边缘 CE 设备经由第一链路连接到第一运营商边缘 PE 设备的第一接口，该 CE 设备经由第二链路连接到第二 PE 设备的第二接口。第一 PE 设备生成第一消息，所述第一消息携带到达所述 CE 设备的第一 MAC/IP Advertisement Route 以及虚拟局域网 VLAN 标识。该第一 MAC/IP Advertisement Route 包括 MAC 地址以及用于标识以太网段 ES 的以太网段标识 ESI，所述第一 MAC/IP Advertisement Route 包括的 MAC 地址为所述 CE 设备的 MAC 地址或所述 CE 设备所管辖的终端设备的 MAC 地址。其中，该以太网段 ES 包括所述第一链路和所述第二链路，所述 VLAN 标识用于指示所述终端设备所属的 VLAN。该第一 PE 设备向该第二 PE 设备发送所述第一消息，所述第一 MAC/IP Advertisement Route 以及所述 VLAN 标识被所述第二 PE 设备用于生成第一 MAC 转发表项。所述第一 MAC 转发表项包含所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址以及所述 VLAN 标识；所述第一 MAC 转发表项包含的出接口标识为所述第二接口的标识；所述第一 MAC 转发表项被所述第二 PE 设备用于向所述 CE 设备转发目的 MAC 地址为所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址的报文。

上述技术方案中，在上述的第二 PE 设备没有从 CE 设备学习到 MAC 路由，也无法获取终端设备所属的 VLAN 信息的情况下，第一 PE 设备向第二 PE 设备通告从 CE 学习到的 MAC 路由以及 VLAN 标识。所述第二 PE 设备能够根据接收到的 MAC 路由以及 VLAN 标识，生成所述第一 MAC 转发表项。当第二 PE 设备接收到目的 MAC 地址为所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址的数据流时，可以根据该第一 MAC 转发表项，直接通过第二链路向所述 CE 设备转发所述数据流。上述技术方案中，在 CE 设备多归属接入多个 PE 设备的场景中，多个 PE 设备能够有效的形成负载分担，实现带宽资源的合理利用。进一步的，在 EVPN 中，PE 设备可以通过控制平面从 CE 学习到 VLAN 信息。在一些可能的场景中，所述第二 PE 设备不能从 CE 设备直接学习到所述 VLAN 信息。本申请提供的上述技术方案，所述第二 PE 设备可以从所述第一 PE 设备学习到所述 VLAN 信息。进而，所述第二 PE 设备可以根据所述 VLAN 信息向 CE 设备转发流量。例如，在 CE 设备通过以太网标签 Tag 或者以太网 Tag 终结接入第二 PE 设备时，

由于所述第二 PE 未从 CE 设备直接学习到 VLAN 信息，导致第二 PE 设备无法获取所述 VLAN 信息，从而无法直接向 CE 设备转发报文。

在一个可选的设计中，所述第一消息为第一边界网关协议更新 BGP Update 消息，所述第一 BGP Update 消息包括用于携带所述 VLAN 标识的 VLAN 属性字段。可选的，该 VLAN 属性字段包括类型（英文：Type）字段和子类型（英文：Sub-Type）字段，该 Type 字段的取值用于标识复用 EVPN 扩展团体属性的类型，该 Sub-Type 字段的取值表示该扩展团体属性为 VLAN 扩展团体属性。该 VLAN 属性字段还包括 VLAN 值（Value）字段，用于携带所述 VLAN 标识。

通过在 BGP 更新消息中扩展 VLAN 属性，本申请可以有效利用现有的协议实现 VLAN 信息的通告。

在一个可选的设计中，在所述第一 PE 设备向所述第二 PE 设备发送所述第一消息之后，所述方法还包括：所述第一 PE 设备接收所述第二 PE 设备发送的第二消息。所述第二消息携带第二 MAC/IP Advertisement Route，下一跳网络地址和所述 VLAN 标识。所述第二 MAC/IP Advertisement Route 包括目的 MAC 地址以及所述以太网段标识 ESI；所述第二 MAC/IP Advertisement Route 中的目的 MAC 地址与所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址相同，所述第二消息携带的所述下一跳网络地址为所述第二 PE 设备的网络地址，例如，为所述第二 PE 设备的环回 loopback 地址。所述以太网段标识 ESI 被所述第一 PE 设备用于确定所述第一 PE 设备连接所述 CE 设备的接口为所述第一接口。所述确定的第一接口，所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址以及所述 VLAN 标识被所述第一 PE 设备用于生成第二 MAC 转发表项。所述第二 MAC 转发表项包括所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址，所述第二转发表项包括的出接口的标识为所述第一接口的标识，所述第二 MAC 转发表项被所述第一 PE 设备用于转发目的 MAC 地址为所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址的报文的表项。所述方法还包括：所述第一 PE 设备根据所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址和所述第二 PE 设备的网络地址，生成第三 MAC 转发表项。所述第三 MAC 转发表项包括所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址，所述第三 MAC 转发表项包括的下一跳网络地址为所述第二 PE 设备的网络地址。当所述第一链路发生故障时，所述第三 MAC 转发表项被所述第一 PE 设备用于转发目的 MAC 地址为所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址的报文的表项。

在上述方案中，所述第一 PE 设备生成所述第三转发表项和所述第四转发表项之后，根据所述第三 MAC 转发表项和所述第四 MAC 转发表项实现快速重路由 FRR。当第一 PE 设备收到已知单播流量（数据报文中携带的目的 MAC 地址为所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址）时，第一 PE 设备查询 MAC 转发表，当第一链路正常工作时，根据第三 MAC 转发表项的指示，直接通过 CE 设备转发报文。当第一链路存在故障时，第一 PE 设备收到上述的已知单播流量后，第一 PE 设备查询 MAC 转发表，根据备用转发表项的指示，即第四 MAC 转发表项的指示，向所述第二 PE 设备转发所述流量，通过所述第二 PE 设备向所述 CE 设备转发所述流量，从而提高了故障的收敛速度。

在一个可选的设计中，所述第二消息还携带有指示信息。所述第一 PE 设备接收所述第二 PE 设备发送的第二消息以后，所述方法还包括：所述第一 PE 设备根据所述指示信息的指示，避免向所述第二 PE 设备发送到达所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址的 MAC/IP Advertisement Route 和所述 VLAN 标识。由此可以有效的避免形成报文环路。

第二方面，本申请提供了一种报文处理方法，应用于 EVPN 中，用户边缘 CE 设备经由第一链路连接到第一运营商边缘 PE 设备的第一接口，所述 CE 设备经由第二链路连接到第二 PE 设备的第二接口。首先，该第二 PE 设备接收所述第一 PE 设备发送的第一消息。所述第一消息携带第一媒体接入控制/互联网协议通告路由 MAC/IP Advertisement Route 以及第一 VLAN 标识。所述第一 MAC/IP Advertisement Route 包括 MAC 地址以及用于标识以太网段 ES 的以太网段标识 ESI，所述第一 MAC/IP Advertisement Route 包括的 MAC 地址为所述 CE 设备的 MAC 地址或所述 CE 设备所管辖的终端设备的 MAC 地址。所述以太网段 ES 包括所述第一链路和所述第二链路，所述第一 VLAN 标识用于指示所述终端设备所属的 VLAN。然后，所述第二 PE 设备根据所述以太网段标识 ESI 确定所述第二 PE 设备连接所述 CE 设备的接口为所述第二接口。进一步地，所述第二 PE 设备根据所述第二接口、所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址以及所述第一 VLAN 标识，生成第一 MAC 转发表项，所述第一 MAC 转发表项包含所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址和所述第一 VLAN 标识，所述第一 MAC 转发表项包含的出接口标识为所述第二接口的标识，所述第一 MAC 转发表项被所述第二 PE 设备用于向所述 CE 设备转发携带有所述第一 VLAN 标识并且目的 MAC 地址为所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址的报文。

在上述方案中，在上述的第二 PE 设备没有从 CE 设备学习到 MAC 路由以及 VLAN 信息的情况下，第一 PE 设备向第二 PE 设备通告从 CE 学习到的 MAC 路由以及 VLAN 标识。所述第二 PE 设备能够根据接收到的 MAC 路由以及 VLAN 标识，生成所述第一 MAC 转发表项。当第二 PE 设备接收到目的 MAC 地址为所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址的数据流时，可以根据该第一 MAC 转发表项，直接通过第二链路向所述 CE 设备转发所述数据流。上述技术方案中，在 CE 设备多归多活接入多个 PE 设备的场景中，多个 PE 设备能够有效的形成负载分担，实现带宽资源的合理利用。进一步的，在 EVPN 中，PE 设备可以通过控制平面学习到所述 VLAN 信息。在一些可能的场景中，所述第二 PE 设备不能从 CE 设备直接学习所述 VLAN 信息。本申请提供的上述技术方案，所述第二 PE 设备可以从所述第一 PE 设备学习到所述 VLAN 信息。进而，所述第一 PE 设备可以根据所述 VLAN 信息向 CE 设备转发到达所述终端设备的流量。例如，在 CE 设备通过以太网标签 Tag 或者以太网 Tag 终结接入 PE 设备时，所述第一 PE 可以根据学习到的 VLAN 信息，向 CE 设备转发报文。

在一个可选的设计中，所述第一消息中还携带有下一跳网络地址，所述第一消息中的下一跳网络地址为所述第一 PE 设备的网络地址，例如，所述第一 PE 设备的环回 lookback 地址。所述方法还包括：所述第二 PE 设备根据所述第一消息，获取所述第一 PE 设备的网络地址；以及所述第二 PE 设备根据所述第一 MAC/IP Advertisement

Route 所包括的 MAC 地址以及所述第一 PE 设备的网络地址，生成第二 MAC 转发表项。其中，所述第二 MAC 转发表项包括所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址，所述第二 MAC 转发表项包含的下一跳网络地址为所述第一 PE 设备的网络地址。当所述第二链路发生故障时，所述第二 MAC 转发表项被所述第二 PE 用于转发携带有所述第一 VLAN 标识并且目的 MAC 地址为所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址的报文。

5 当第二 PE 设备收到已知单播流量（数据报文中携带的目的 MAC 地址为所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址）时，第二 PE 设备查询 MAC 转发表，在所述第二链路处于正常工作状态时，根据所述第一 MAC 转发表项的指示，经由所述第二链路，直接通过所述 CE 设备转发报文。当所述第二链路故障时，第二 PE 设备收到上述的已知单播流量时，根据第二 MAC 转发表项的指示，向所述第一 PE 设备转发所述流量，通过所述第一 PE 设备向所述 CE 设备转发所述流量，从而提高了故障的收敛速度。进一步的，在 EVPN 中，PE 设备可以通过控制平面从 CE 设备学习 VLAN 信息，在一些可能的场景中，所述第二 PE 设备不能从 CE 设备直接学习所述 VLAN 信息。本申请提供的上述技术方案，所述第二 PE 设备可以从所述第一 PE 设备学习到所述 VLAN 信息。进而，所述第二 PE 设备可以根据所述 VLAN 信息向 CE 设备转发流量。例如，在 CE 设备通过以太网标签 Tag 或者以太网 Tag 终结接入 PE 设备时，所述第二 PE 设备可以根据所述 VLAN 信息，向 CE 设备转发流量。

10 在一个可选的设计中，所述第一消息为第一边界网关协议更新 BGP Update 消息，所述第一 BGP Update 消息包括用于携带所述第一 VLAN 标识的 VLAN 属性字段。

通过在 BGP 更新消息中扩展 VLAN 属性，本申请可以有效利用现有的协议实现 VLAN 信息的通告。

25 在一个可选的设计中，所述第二 PE 设备根据所述确定的第二接口、所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址以及所述第一 VLAN 标识，生成第一 MAC 转发表项以后，所述第二 PE 设备接收携带有第二 VLAN 标识的第一报文，所述第一报文的目的是 MAC 地址为所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址。所述第二 PE 设备根据所述第一 MAC 转发表项将所述第二报文中携带的第二 VLAN 标识替换为所述第一 VLAN 标识，得到所述携带有所述第一 VLAN 标识的第二报文。所述第二 PE 设备向所述 CE 设备发送所述第二报文。

30 在一个可选的设计中，在所述第二 PE 设备接收所述第一 PE 设备发送的第一消息之后，所述方法还包括：所述第二 PE 设备生成第二消息，所述第二消息携带第二 MAC/IP Advertisement Route，下一跳网络地址和所述第一 VLAN 标识。所述第二 MAC/IP Advertisement Route 包括 MAC 地址和所述以太网段标识 ESI。所述第二 MAC/IP Advertisement Route 包括的 MAC 地址与所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址相同，所述第二消息携带的所述下一跳网络地址为所述第二 PE 设备的网络地址。所述第二 PE 设备向所述第一 PE 设备发送所述第二消息，所述第二消息被所述第一 PE 设备用于生成第三 MAC 转发表项和第四 MAC 转发表项。其中，所述第三 MAC 转发表项被所述第一 PE 设备用于转发目的 MAC 地址为所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址的报文，所述第三 MAC 转发表项包括所述第二 MAC/IP

Advertisement Route 所包括的 MAC 地址和所述第一 VLAN 标识, 所述第三 MAC 转发表项包含的出接口的标识为所述第一接口的标识。当所述第一链路发生故障时, 所述第四 MAC 转发表项被所述第一 PE 设备用于转发目的 MAC 地址为所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址的报文, 所述第四 MAC 转发表项包括所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址, 所述第四转发表项包含的下一跳网络地址为所述第二 PE 设备的网络地址。

通过上述方案, 第二 PE 设备在收到所述第一消息以后, 根据第一消息中携带的信息, 在控制平面生成本地的主用 MAC 路由表项以及用于快速重路由的备用 MAC 路由表项, 进一步的, 控制平面将上述主用 MAC 路由表项和备用 MAC 路由表项下发到转发平面, 生成用于实现快速重路由的第一 MAC 转发表项和第二 MAC 转发表项。并且, 第二 PE 设备在收到所述第一消息以后, 生成本地的所述第一 MAC 转发表项后, 还会将该本地的 MAC 路由回送给所述第一 PE 设备, 使得第一 PE 设备生成用于实现快速从路由的备用 MAC 路由表项。而第一 PE 设备从 CE 设备学习到的 MAC 路由可以作为本地 MAC 路由, 即主用 MAC 路由。当第一 PE 设备和 CE 设备连接的链路, 例如, 所述第一链路发生链路故障时, 所述第一 PE 设备的本地 MAC 路由被撤销。当所述第一链路故障恢复后, 第一 PE 设备可以根据所述第二 PE 设备通告的所述第二 MAC/IP Advertisement Route 路由和所述第一 VLAN 标识, 再次生成所述本地 MAC 路由, 用于指导向所述 CE 设备转发流量。从而使得, 在所述第一链路发送故障并再次恢复后, 能够快速实现路由重定向, 实现故障的快速收敛。

需要说明的是, 在本申请中所述的 PE 设备的本地 MAC 路由, 是指用于指导直接向所述 CE 设备转发单播流量的路由, 本地 MAC 路由中包括的目的 MAC 地址为所述 CE 设备的 MAC 地址或者接入所述 CE 设备的终端设备的 MAC 地址, 本地 MAC 路由中包括的出接口的标识为所述 PE 设备上的连接所述 CE 设备的接口标识。

在一个可选的设计中, 所述第二消息还携带有指示信息, 所述指示信息用于指示所述第一 PE 设备在收到所述第二消息后, 避免向所述第二 PE 设备发送到达所述第二 MAC/IP Advertisement Route 包括的 MAC 地址的 MAC/IP Advertisement Route 和所述第一 VLAN 标识, 从而避免形成报文环路。

第三方面, 本申请提供了一种第一运营商边缘 PE 设备, 第一 PE 设备用于执行第一方面以及第一方面中任一可能的设计中的方法。具体地, 第一 PE 设备包括用于实现第一方面以及第一方面中任一可能的设计中的方法的模块。

第四方面, 本申请提供了一种第二运营商边缘 PE 设备, 第二 PE 设备用于执行第二方面以及第二方面中任一可能的设计中的方法。包括用于实现第二方面以及第二方面中任一可能的设计中的方法的模块。

第五方面, 本申请提供了一种第一 PE 设备, 所述第一 PE 设备包括: 输入接口、输出接口、处理器和存储器。其中, 输入接口、输出接口、处理器以及所述存储器之间可以通过总线系统相连。该存储器用于存储程序, 所述处理器用于执行所述存储器中的程序, 从而执行第一方面或第一方面的任意可能的设计中的方法。

第六方面, 本申请提供了一种第二 PE 设备, 所述第二 PE 设备包括: 输入接口、输出接口、处理器和存储器。其中, 输入接口、输出接口、处理器以及所述存储器之

间可以通过总线系统相连。该存储器用于存储程序，所述处理器用于执行所述存储器中的程序，从而执行第二方面或第二方面的任意可能的设计中的方法。

5 第七方面，本申请实施例提供了一种第一 PE 设备，所述第一 PE 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 PE 设备用于执行第一方面或第一方面的任意可能的设计中的方法。

第八方面，本申请实施例提供了一种第一 PE 设备，所述第一 PE 设备包括：控制器和第一 PE 转发设备。所述第一 PE 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 PE 设备用于执行第一方面或第一方面的任意可能的设计中的方法。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。  
10 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七方面中主控板的功能。

15 第九方面，本申请实施例提供了一种第二 PE 设备，所述第二 PE 设备包括：主控板和接口板，进一步，还可以包括交换网板。所述第二 PE 设备用于执行第二方面或第二方面的任意可能的设计中的方法。

第十方面，本申请实施例提供了一种第二 PE 设备，所述第二 PE 设备包括：控制器和第二 PE 转发设备。所述第二 PE 转发设备包括：接口板，进一步，还可以包括交换网板。所述第二 PE 设备用于执行第二方面或第二方面的任意可能的设计中的方法。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。  
20 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七方面中主控板的功能。

第十一方面，本申请实施例提供了一种通信系统，该通信系统包括第三方面、第五方面，第七方面或第八方面所述的第一 PE 设备和第四方面、第六方面、第九方面或第十一方面所述的第二 PE 设备。

30 第八方面，本申请提实施例供了一种计算机可读存储介质或者计算机程序产品，用于存储计算机程序，该计算机程序用于执行第一方面、第二方面、第一方面任意可能的设计或第二方面任意可能的设计中的方法。

通过本申请实施例提供的方法，设备和系统，能够使得在 EVNP 中，CE 设备多归多活接入 PE 设备场景下，连接所述 CE 设备的各 PE 设备之间能够形成有效的负载分担，  
35 系统带宽资源得到合理利用。

## 附图说明

为了更清楚地说明本申请实施例中的技术方案，下面将对实施例描述中所需要使用的附图作简单地介绍。

图 1 为现有技术中提供的一种 EVPN 应用场景示意图；

图 2 为本申请实施例提供的一种报文处理方法的应用网络场景示意图；

图 3 为本申请实施例提供的另一种报文处理方法的流程示意图；

5 图 4a 为本申请实施例提供的一种 BGP Update 消息中 MP\_REACH\_NLRI 属性字段格式示意图；

图 4b 为本申请实施例提供的一种 EVPN NLRI 字段格式示意图；

图 4c 为本申请实施例提供的一种 EVPN MAC/IP Advertisement Route 字段格式示意图；

图 4d 为本申请实施例提供的一种 ESI 字段格式示意图；

10 图 5 为本申请实施例提供的一种 BGP Update 消息中 VLAN 属性字段格式示意图；

图 6 为本申请实施例提供的另一种报文处理方法的流程示意图；

图 7 为本申请实施例提供的另一种报文处理方法的流程示意图；

图 8 为本申请实施例提供的一种第一 PE 设备的结构示意图；

图 9 为本申请实施例提供的一种第二 PE 设备的结构示意图；

15 图 10 为本申请实施例提供的一种第一 PE 设备的硬件结构示意图；

图 11 为本申请实施例提供的一种第二 PE 设备的硬件结构示意图；

图 12 为本申请实施例提供的一种第一 PE 设备的结构示意图；

图 13 为本申请实施例提供的一种第一 PE 设备的结构示意图；

图 14 为本申请实施例提供的一种第二 PE 设备的结构示意图；

20 图 15 为本申请实施例提供的一种第二 PE 设备的结构示意图。

### 具体实施方式

下面将结合附图，对本申请实施例中的技术方案进行描述。本申请实施例描述的网络架构以及业务场景是为了更加清楚的说明本申请实施例的技术方案，并不构成对于本申请实施例提供的技术方案的限定，本领域普通技术人员可知，随着网络架构的演变和新业务场景的出现，本申请实施例提供的技术方案对于类似的技术问题，同样适用。

本申请描述的技术方案可以适用于基于 BGP MPLS 的 EVPN 中。EVPN 技术采用类似于 BGP/MPLS 互联网协议（英文：Internet Protocol, IP）VPN 的机制，通过扩展 BGP 协议，使用扩展后的可达性信息，使不同站点的二层网络间的 MAC 地址学习和发布过程从数据平面转移到控制平面。通过在控制平面学习 MAC 地址来实现 L2VPN 的功能。在控制平面学习 MAC 地址，可以解决数据平面进行 MAC 地址学习而导致的网络设备多归属实现困难、无法支持负载分担等问题。

CE 设备多归属接入 EVPN 包括通过以太网链路多归接入 EVPN。一个 CE 设备通过多条链路分别连接到多个网络侧设备的部署称为 CE 设备多归接入。

35 图 2 示出了典型的以太网链路多归多活接入 EVPN 的场景。如图 2 所示，EVPN 中包括四个 PE 设备，分别是 PE1-1，PE1-2，PE1-3 和 PE2。CE1 分别通过以太网链路（英文：Ethernet Link, EL）1、EL2 和 EL3 连接到 PE1-1、PE1-2 和 PE1-3。包含这三条以太网链路的一组以太网链路是一个以太网段（英文：Ethernet Segment, ES）。以太网段标识（英文：Ethernet Segment Identifier, ESI）是一个唯一的非零标识，

用于标识该以太网段 ES。

PE1-1 学习到 VPN1 的站点 (英文: site) 1 中的用户设备 (英文: User Equipment, UE) 1 的 MAC 地址, 例如 MAC A, PE1-1 通过 BGP Update 消息向 PE2 发布 MAC/IP 通告路由。PE1-2 没有学习到 UE1 的 MAC 地址。PE1-2 向 PE2 发布以太网自动发现路由 (英文: Ethernet Auto-discovery Route, Ethernet A-D route)。因此, PE2 根据别名 (英文: Aliasing) 方式知道, PE2 经由 PE1-1 可以到达 UE1, 和 PE2 经由 PE1-2 可以到达 UE1。由此, UE2 发往 UE1 的单播流量, 经由 PE2 传输时, PE2 可以对该单播流量进行负载分担处理。该单播流量经由 PE1-1 和 PE1-2, 被转发至 CE1。从而实现 VPN1 里的 UE1 和 UE2 的互通。

10 在以太网链路多归属部署场景中, EVPN 支持多种冗余模式。该多种冗余模式包括单活的冗余模式、多活的冗余模式和全活的冗余模式。所谓单活的冗余模式 (简称为单活模式) 是指以太网链路段中只有一条以太网链路的状态是活跃的, 其它的一条或者多条以太网链路的状态是非活跃的。活跃状态意味着该以太网链路可以用于承载、转发数据流。在主备保护的场景下, 通常用作主用以太网链路。相应地, 以太网链路的状态还可以为非活跃, 非活跃状态意味着该以太网链路不可以用于承载、转发数据流, 通常用作备用以太网链路。当主用以太网链路故障时, 会切换到备用以太网链路来承载、转发数据流。所以, 单活模式下的部署场景可以包括单活 (ES 中只有一条 EL)、单活单备 (ES 中有两条 EL, 一条的状态为活跃, 另一条的状态为非活跃) 和单活多备 (ES 中有至少三条 EL, 一条的状态为活跃, 另外至少两条的状态为非活跃)。结合图 20 2 进一步解释, 假如该 ES 中只有一条以太网链路 EL1 是活跃的, 用作主用 EL, 而其它 EL2 和 EL3 都为非活跃, 用作备份 EL, 则这种冗余模式为单活双备 (属于单活多备)。

在以太网链路多归属部署场景中, 所谓全活的冗余模式 (简称为全活模式) 是指以太网链路段中所有的以太网链路的状态都是活跃的, 即没有非活跃状态的以太网链路。所有这些活跃状态的以太网链路可以实现对数据流进行负载分担地转发, 从而提供更大带宽的传输能力。但是全活模式的场景不支持备份, 即没有作为备用的以太网链路, 当主用的一条或多条以太网链路出现故障时, 无法切换到备用以太网链路进行冗余保护。结合图 2 进一步解释, 假如该 ES 中的全部三条以太网链路 EL1、EL2 和 EL3 都为活跃, 则没有备份 EL, 则这种冗余模式为全活模式。

30 在以太网链路多归属部署场景中, 所谓多活的冗余模式 (简称为多活模式) 是指以太网链路段中的部分以太网链路的状态是活跃的, 另一部分以太网链路的状态是非活跃的。这些活跃状态的以太网链路 (用作主用以太网链路) 可以实现对数据流进行负载分担地转发, 从而提供更大带宽的传输能力。而另一部分非活跃状态的以太网链路用作备份, 当主用的一条或多条以太网链路出现故障时, 可以切换到这些备用以太网链路进行冗余保护。结合图 2 进一步解释, 假如该 ES 中的两条以太网链路 EL1 和 EL2 35 为活跃状态, EL3 为非活跃状态, 则 EL1 和 EL2 联合起来对数据流进行负载分担的转发, 而 EL3 为 EL1 或者 EL2 提供备份保护。

关于 MAC/IP 通告路由以及 Ethernet A-D 路由的具体细节, 以及 PE2 如何根据别名方式形成负载分担, 可以参考 RFC7432, 该文档与此相关部分的内容以引入的方式并入本文本中, 此处为了简洁, 不再赘述。

在图 2 所示的场景中，CE 设备发送给 PE 设备的流量，通过哈希（英文：Hash）算法选路，Hash 算法的具体实现依赖于 CE 设备。其中，CE 设备的实现方式并不能保证流量会流经 CE 设备与 PE 设备连接的每条链路。对于没有被选择的链路，与该链路对应的 PE 设备学习不到 CE 设备所接入的终端设备的 MAC 地址，也无法有效获取所述终端设备的 VLAN 信息。例如，CE 设备使用源 MAC 地址结合目的 MAC 地址作为哈希因子时，CE 设备发送给 PE 设备的流量可能会哈希到与 PE1-1 连接的链路上，从而导致 PE1-2 没有学习到 UE1 的 MAC 地址，导致 PE1-2 上没有到达 UE1 的本地 MAC 转发表项，当目的 MAC 地址为 MAC<sub>A</sub> 的单播流量到达 PE1-2 时，PE1-2 无法直接向 CE1 转发该单播流量。

需要说明的是，本领域技术人员应该理解，图 2 所示的场景仅是一种例举，不应构成对本申请的限制。图 2 示出终端设备通过 CE 设备接入所述 PE 设备，实际操作中，也可以是终端设备自身作为 CE 设备接入所述 PE 设备，此时，终端设备的 MAC 地址即为 CE 设备的 MAC 地址。图 2 所示的场景可以引用于多种场景中。例如，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 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）。又例如，所述 EVPN VXLAN 应用于固网（英文：Fixed Network）。在固网中，所述 CE 设备可以是用户侧的站点，所述 PE 设备可以是宽带接入服务器（英文：Broadband Access Server，缩写：BAS）。

还需要说明的是，本申请中，本申请实施例中的 CE 设备和 PE 设备可以是 RFC7432 中定义的相应设备。所述的 PE 设备可以是路由器或交换机。所述 CE 设备可以是路由器或交换机或终端设备。当所述 CE 设备是路由器时，可以连接一台或者多台终端设备。CE 设备通常一侧与 PE 设备相连，另一侧与 UE 相连，实现将用户设备中转、接入运营商网络。UE 又称之为终端设备（英文：Terminal Equipment，TE）或终端（英文：terminal），可以是具有无线通信功能的手持设备、车载设备、可穿戴设备、计算机设备、虚拟机或连接到无线调制解调器的其它处理设备。UE 还可以是用户设备或者移动台（英文：Mobile Station，MS）。

还需要说明的是，PE 和 PE 设备在本申请的各个实施例中是同一个意思。同理，CE 和 CE 设备是同一个意思。本申请中所述的数据流可以是已知 MAC 地址的单播数据流。

图 3 示出了本申请提供的一种报文处理方法 300，应用于 EVPN，用户边缘 CE 设备通过至少两条链路分别接入到至少两个 PE 设备的场景中。所述至少两条链路形成一个以太网段，所述至少两个 PE 设备包括第一 PE 设备和第二 PE 设备。需要说明的是，在本申请中，所述链路可以是以太网链路；用于标识所述以太网段的标识是以太网段标识 ESI。以太网段也可以称之为以太网链路段或者以太网链路集合。所述 CE 设备经由第一链路接入所述第一 PE 设备的第一接口。所述 CE 设备经由第二链路接入所述第二 PE 设备的第二接口。以太网段包括所述第一链路和所述第二链路。

图 3 所示的方法 300 可以应用于图 2 所示的场景中。具体地, 所述 CE 设备例如可以是图 2 所示的 CE1, 所述第一 PE 设备例如可以是图 2 所示的 PE1-1, 所述第二 PE 设备例如可以是图 2 所示第 PE1-2。该方法 300 包括 S301 至 S305。

S301. 第一 PE 设备生成第一消息。

5 具体地, 所述第一消息携带、第一媒体接入控制/互联网协议通告路由 MAC/IP Advertisement Route 以及第一 VLAN 标识。所述第一 MAC/IP Advertisement Route 包括 MAC 地址以及用于标识以太网段 ES 的链路以太网段标识 ESI, 所述第一 MAC/IP Advertisement Route 包括的 MAC 地址为所述 CE 设备的 MAC 地址或所述 CE 设备所管辖的终端设备的 MAC 地址。所述第一 VLAN 标识用于标识所述第一 MAC/IP  
10 Advertisement Route 包括的 MAC 地址所属的 VLAN。进一步的, 所述第一消息还包括下一跳网络地址, 该下一跳网络地址为该第一 PE 设备的网络地址, 例如, 为第一 PE 设备的环回 (英文: lookback) 地址。需要说明的是, 本申请所述的环回地址是在网络设备 (如路由器、交换机等) 的环回接口上配置的 IP 地址, 通常用作网络设备标识 (例如, 32 位掩码的 IPv4 地址: 10.10.1.1/32), 本领域技术人员可以理解。

15 在一个具体的实施方式中, 所述第一 PE 设备从所述第一接口接收所述 CE 设备通过所述第一链路发送的报文, 所述报文中携带所述 CE 设备的 MAC 地址或接入所述 CE 设备的终端设备的 MAC 地址以及所述第一 VLAN 标识, 所述第一 PE 设备从该报文中获取所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址和所述第一 VLAN 标识。所述第一 PE 设备根据所述第一接口确定所述以太网段标识 ESI。具体地, 所述第一 PE  
20 设备可以包括多个接口。所述多个接口可以是多个以太网接口。所述第一 PE 设备可以保存所述第一 PE 设备的每个接口的配置信息。所述第一接口的配置信息包括所述 ESI。也就是说, 所述第一接口与所述 ESI 具有对应关系。所述第一 PE 设备可以根据所述第一接口与所述 ESI 的对应关系确定所述 ESI。

S302. 该第一 PE 设备向第二 PE 设备发送所述第一消息。

25 具体地, 所述第一 PE 设备向所述第二 PE 设备发送所述第一消息, 所述第一 MAC/IP Advertisement Route 以及所述第一 VLAN 标识被所述第二 PE 设备用于生成第一 MAC 转发表项。所述第一 MAC 转发表项包含所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址以及所述第一 VLAN 标识, 所述第一 MAC 转发表项包含的出接口标识为所述  
30 第二接口的标识, 所述第一 MAC 转发表项被所述第二 PE 设备用于向所述 CE 设备转发目的 MAC 地址为所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址的报文。

需要说明的是, 在本申请中, 所述第一 PE 设备能够从所述 CE 设备学习到所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址, 所述第二 PE 设备不能从所述 CE 设备学习到所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址。

35 在一个具体的实施方式中, 所述第一消息为边界网关协议 (英文: Border Gateway Protocol, BGP) 更新 (英文: Update) 消息 (也可称为 BGP Update 报文)。在本申请中, 将该 BGP Update 消息称之为第一 BGP Update 消息。该第一 BGP Update 消息携带所述 MAC/IP Advertisement Route 和所述第一 VLAN 标识。为了方便, 将第一 BGP Update 消息中携带的 MAC/IP Advertisement Route 称之为第一 MAC/IP Advertisement Route。该 MAC/IP Advertisement Route 属于 BGP 协议定义的 EVPN 网络可达层信息 (英

文: Network Layer Reachability Information, NLRI) 中的一种路由类型, 用于指导单播流量转发。该 EVPN NLRI 承载在多协议网络层可达信息(英文: Multiprotocol Reachable NLRI, MP\_REACH\_NLRI) 属性中。MP\_REACH\_NLRI 属性是 BGP Update 消息里定义的一个属性, 具体格式如图 4a 所示, 该属性包括地址族标识(英文: Address Family Identifier, AFI) 字段和子地址族标识(英文: Subsequent Address Family Identifier, SAFI) 字段。该 AFI 字段的取值用于指示 L2VPN, 如为 25。该 SAFI 字段的取值用于指示 EVPN, 如为 70。该 MP\_REACH\_NLRI 属性还包括下一跳网络地址长度(英文: Length of Next Hop Network Address) 和下一跳网络地址(英文: Network Address of Next Hop) 字段。该下一跳网络地址字段用来携带所述的下一跳网络地址(如 loopback 地址)。该 MP\_REACH\_NLRI 属性还包括 NLRI 字段, 结合上述 AFI 和 SAFI 的取值指示 L2VPN 中的 EVPN, 该 NLRI 字段为 EVPN NLRI 字段。如图 4b 所示, 该 EVPN NLRI 字段包括例如 2 个字节的 5 路由类型(英文: Route Type) 字段、2 个字节的长度(英文: Length) 字段和变长的路由类型细节(英文: Route Type specific) 字段。

需要说明的是, 本申请对于 Route Type 字段以及 Length 字段的长度不作具体限定。其中, 该 Route Type 包括该 MAC/IP Advertisement Route, 例如, 取值为 2。该 Route Type specific 字段用于承载该 MAC/IP Advertisement Route 的细节。如图 4c 所示, 该 MAC/IP Advertisement Route 包括 8 个字节的 15 路由区分符(英文: Route Distinguisher, RD) 字段、10 个字节的以太网段标识(英文: Ethernet Segment Identifier, ESI) 字段、4 个字节的以太网标签标识(英文: Ethernet Tag ID) 字段、20 1 个字节的 MAC 地址长度字段、6 个字节的 MAC 地址字段、1 个字节的下一跳网络地址长度字段、0 字节或 4 字节或 16 个字节的下一跳网络地址字段、3 个字节的 MPLS 标签 1(英文: Label 1) 字段以及 0 字节或 3 字节 MPLS 标签 2 字段, MPLS 标签 2 用于指导 3 层流量转发。图 4c 中所示的 ESI 字段的格式如图 4d 所示, 包括类型(英文: Type, T) 字段和 ESI 值(英文: Value) 字段。其中 Type 字段用于指示 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 PE 和 CE 之间运行链路聚合控制协议(英文: Link Aggregation Control Protocol, 缩写: LACP), 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。其中, 关于 BGP update 消息和 MP\_REACH\_NLRI 属性的定义可以参见 RFC476 中的说明, 所述 EVPN NLRI 字段的定义参见 RFC7432 中的说明。

在一个示例中, 本申请通过对 BGP 协议进行扩展, 增加一个 VLAN 属性, 通过 VLAN 属性携带所述 VLAN 标识。所述 VLAN 属性的具体格式如图 5 所示。该 VLAN 属性包括类型(英文: Type) 字段和子类型(英文: Sub-Type) 字段。该 Type 字段的取值用于标识复用 EVPN 扩展团体属性的类型, 取值例如为 0X06。该 Sub-Type 字段的取值表示该扩展团体属性为 VLAN 扩展团体属性, 取值例如为 0xcc。上述 Type 字段和 Sub-Type 字段的取值仅为示例, 实际取值以因特网地址分配组织(英文: Internet Assigned Numbers Authority, IANA) 分配的数值为准。该 VLAN 属性还包括保留(英文: Reserved) 字段, 不使用时, 填充为 0。该 VLAN 属性还包括 VLAN 值(Value) 字段, 该 VLAN Value 字段的长度例如可以是 4 字节, 用于携带 VLAN 标识。在一个具体的实施方式中, 该

VLAN Value 字段包括业务 VLAN (S-VLAN) 字段以及客户 VLAN (C-VLAN) 字段, 该 S-VLAN 字段用于携带 S-VLAN 标识, 不能为 0。该 C-VLAN 字段用于携带 C-VLAN 标识, 当值为 0 时, 表示没有 C-VLAN, 非 0 值时用于表示 C-VLAN 标识。该 Type 字段和 Sub-Type 字段的长度例如可以分别是 1 个字节, 该 Reserved 字段的长度例如可以是 2 个字节, 该 S-VLAN 字段和 C-VLAN 字段的长度例如可以分别是 2 个字节, 本申请对此不作具体限定。

S303. 所述第二 PE 设备接收所述第一 PE 设备发送的所述第一消息。

S304. 所述第二 PE 设备确定连接所述 CE 设备的接口为所述第二接口。

在一个具体的实施方式中, 在 EVPN 中, 当多个 PE 设备连接同一个 CE 设备时, 用于连接该 CE 设备的接口配置相同的以太网段标识 ESI。CE 设备通过 E-TRUNK 多归连接到第一 PE 设备和第二 PE 设备, 该 CE 设备相当于连接了一个 PE 设备。第一 PE 设备和第二 PE 设备上具有多个接口, 该第一 PE 设备通过第一接口与该 CE 设备连接, 该第二 PE 设备通过第二接口与该 CE 设备连接。所述第一 PE 设备上保存了所述第一接口的配置信息。所述第二 PE 设备上保存了所述第二接口的配置信息。所述第一接口的配置信息包括 ESI。所述第二接口的配置信息包括 ESI。为该第一接口配置的 ESI 与为该第二接口配置的 ESI 是相同的。因此, 当所述第二 PE 设备接收所述第一消息以后, 获取所述第一消息中携带的所述第一 MAC/IP Advertisement Route。所述第二 PE 设备提取所述第一 MAC/IP Advertisement Route 中携带的所述以太网段标识 ESI, 以获取所述以太网段标识 ESI。该第二 PE 设备根据所述 ESI 确定连接所述 CE 设备的接口为所述第二接口。具体地, 所述第二 PE 设备可以包括多个接口。所述多个接口可以是多个以太网接口。所述第二 PE 设备可以保存所述第二 PE 设备的每个接口的配置信息。所述第二接口的配置信息包括所述 ESI。也就是说, 所述第二接口与所述 ESI 具有对应关系。所述第二 PE 设备可以以所述第一 MAC/IP Advertisement Route 中携带的所述 ESI 为查找关键字, 在所述第二 PE 设备保存的多个接口的配置信息中查找包含所述 ESI 的配置信息。当所述第二 PE 设备查找到包含所述 ESI 的配置信息时, 所述第二 PE 设备可以根据所述第二接口与所述 ESI 的对应关系确定所述第二 PE 设备连接所述 CE 设备的接口为所述第二接口。

S305. 所述第二 PE 设备生成第一 MAC 转发表项。

具体地, 所述第二 PE 设备根据所述第二接口, 所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址以及所述第一 VLAN 标识, 生成第一 MAC 转发表项。所述第一 MAC 转发表项被所述第二 PE 设备用于向所述 CE 设备转发目的 MAC 地址为所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址的报文。所述第一 MAC 转发表项包含所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址和所述第一 VLAN 标识, 所述第一 MAC 转发表项包含的出接口标识为所述第二接口的标识。作为示例, 当第二 PE 设备接收到去往所述终端设备的数据流 (已知单播数据流) 后, 基于该第一 MAC 转发表项, 将该数据流中的报文从该确定的第二接口向所述 CE 设备转发。

在一个具体的实施方式中, CE 设备通过以太 Tag, 以太 Tag 终结或者 802.1Q 嵌套 802.1Q (QinQ) 等方式接入所述第一 PE 设备和所述第二 PE 设备。第二 PE 设备接收到携带有第二 VLAN 标识的第一报文, 该第一报文的目的地 MAC 地址为所述第一 MAC/IP

Advertisement Route 所包括的 MAC 地址。所述第二 PE 设备通过查找该第一 MAC 转发表项，确定所述第二接口为转发该第一报文的出接口。所述第二 PE 设备以第一报文中的携带的 MAC 地址为关键字，在该第一 MAC 转发表项中查找与该 MAC 地址关联的所述第一 VLAN 标识，将该第一报文中的第二 VLAN 标识替换为该第一 VLAN 标识，并在完成所有的转发处理动作后，得到封装有所述第一 VLAN 标识的第二报文，将该第二报文通过所述第二接口转发给所述 CE 设备。需要说明的是，在本申请各实施例中，根据所述 CE 设备接入所述 PE 设备的不同接入方式，所述第一 VLAN 标识可以仅包括单个 VLAN 标识符（英文：Identifier, ID）也可以包括多个 VLAN ID。同理，所述第二 VLAN 标识可以仅包括单个 VLAN ID，也可以包括多个 VLAN ID。第一 VLAN 标识和第二 VLAN 标识的具体形式，本申请不作具体限定。

上述技术方案中，在上述的第二 PE 设备没有从 CE 设备学习到 MAC 路由，也无法获取有效的 VLAN 信息的情况下，第一 PE 设备向第二 PE 设备通告从所述 CE 设备学习到的 MAC 路由以及所述 VLAN 标识。所述第二 PE 设备能够根据接收到的 MAC 路由以及 VLAN 标识，生成所述第一 MAC 转发表项。当第二 PE 设备接收到目的 MAC 地址为所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址的数据流时，可以根据该第一 MAC 转发表项，直接通过第二链路向所述 CE 设备转发所述数据流。上述技术方案中，在 CE 设备多归多活接入多个 PE 设备的场景中，多个 PE 设备能够有效的形成负载分担，实现带宽资源的合理利用。进一步的，在 EVPN 中，PE 设备可以通过控制平面学习到所述 VLAN 标识。在一些可能的场景中，所述第二 PE 设备不能从 CE 设备直接学习所述 VLAN 标识。本申请提供的上述技术方案，所述第二 PE 设备可以从所述第一 PE 设备学习到所述 VLAN 标识。进而，所述第二 PE 设备可以根据所述 VLAN 标识向 CE 设备转发流量。例如，在 CE 设备通过以太网标签 Tag 或者以太网 Tag 终结接入第二 PE 设备时，根据所述 VLAN 标识，向 CE 设备转发流量。

在一个具体的实施方式中，在所述 S305 之后，所述方法 300 还可以包括 S306-308，如图 6 所示，所述方法 300 包括 S301-S308。

S301 至 S305 的相关说明参见前述实施例，此处不再赘述。

S306. 所述第二 PE 设备获取获取所述第一消息中携带的下一跳网络地址。

具体地，该下一跳网络地址可称之为第一下一跳网络地址，该第一下一跳网络地址为所述第一 PE 设备的网络地址，例如，为所述第一 PE 设备的 loopback 地址。

S307. 所述第二 PE 设备生成第二 MAC 转发表项。

具体地，所述第二 PE 设备根据所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址和所述第一 PE 设备的网络地址，生成所述第二 MAC 转发表项。所述第二 MAC 转发表项包括所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址，所述第二 MAC 转发表项包含的下一跳网络地址为所述第一 PE 设备的网络地址。

在一个具体地实施方式中，所述第二 PE 设备从第三接口 Intf1 接收所述第一 PE 设备发送的所述第一消息，所述第二 PE 设备将所述第一 PE 设备作为去往所述终端设备的下一跳节点。第二 PE 设备的控制平面（例如，控制板）生成 MAC 路由表项（如表 1 所示），所述 MAC 路由表项的目的 MAC 地址为所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址，下一跳网络地址为所述第一 PE 设备的环回地址。然后，所述第二

PE 设备根据所述 MAC 路由表项生成所述第二 MAC 转发表项（如表 2 所示），并将该第二 MAC 转发表项发送到第二 PE 设备的转发平面（例如，转发板）。所述第二 PE 设备根据所述 MAC 路由表项生成所述第二 MAC 转发表项时，所述第二 PE 设备可以根据所述 MAC 路由表项确定所述第二 MAC 转发表项中的目的 MAC 地址以及出接口。所述 MAC 转发表项的目的 MAC 地址为所述 MAC 路由表项中的目的地址（即所述 UE1 的 MAC 地址）。所述 MAC 转发表项的出接口为 Intf1。所述第二 PE 设备将所述 Intf1 确定为所述第二 MAC 转发表项中的出接口可以包括：首先，第二 PE 设备以该 MAC 路由表项中的第一 PE 设备的 loopback 地址为查找关键字，查找转发等价类（英文：Forwarding Equivalence Class, FEC）到下一跳标签转发项（英文：Next Hop Label Forwarding Entry, NHLFE）映射表（简称也叫 FTN 映射表或者 FTN 转发表），获得与第一 PE 设备的 loopback 地址相对应的出接口为第二 PE 设备到第一 PE 设备的隧道的隧道标识（英文：Tunnel Identifier, Tunnel ID）；然后，用该 Tunnel ID 查找隧道转发表，获得与该 Tunnel ID 对应的出接口为 Intf 1（即该第二 PE 设备到该第一 PE 设备的隧道在该第二 PE 设备上的接口）。所述第二 PE 设备将所述 Intf1 确定为所述第二 MAC 转发表项中的出接口。需要说明的是，该隧道可以是标签交换路径（Label Switched Path, LSP）隧道，也可以是资源预留协议-流量工程（Resource Reservation Protocol -Traffic Engineering, RSVP-TE）隧道等。该隧道用于承载已知单播数据流，为了简洁，本发明的实施例附图 2 中并未示出，本领域技术人员可以理解。

表 1: MAC 路由表

目的 MAC	下一跳网络地址列表
终端设备的 MAC 地址	第一 PE 设备的 loopback 地址

20

表 2: MAC 转发表

目的 MAC	出接口列表
终端设备的 MAC 地址	Intf 1

S308. 所述第二 PE 设备根据所述第一 MAC 转发表项和所述第二 MAC 转发表项实现快速重路由（英文：Fast Reroute, FRR）。所述第一 MAC 转发表项作为主用转发表项，所述第二 MAC 转发表项作为备用转发表项。

当第二 PE 设备收到已知单播流量（数据报文中携带的目的 MAC 地址为所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址）时，第二 PE 设备查询 MAC 转发表，在所述第二链路处于正常工作状态时，根据第一 MAC 转发表项的指示，经由所述第二链路，直接通过所述 CE 设备向终端设备转发报文；当所述第二链路故障时，第二 PE 设备收到上述的已知单播流量时，根据第二 MAC 转发表项的指示，向所述第一 PE 设备转发所述流量，通过所述第一 PE 设备向所述 CE 设备转发所述流量，从而提高了故障的收敛速度。

可选的，在所述 S305 之后，所述方法 300 还可以包括 S309-S312，下面结合图 7 对方法 300 进行说明。

S301 至 S305 的相关说明参见前述实施例，此处不再赘述。

S309. 所述第二 PE 设备生成第二消息。

35

在一个具体的实施方式中，在所述第二 PE 设备接收所述第一 PE 设备发送的第一消息之后，所述第二 PE 设备生成所述第二消息。具体地，所述第二 PE 设备根据所述第一消息生成所述第二消息。所述第二消息用于携带第二 MAC/IP Advertisement Route，下一跳网络地址和所述第一 VLAN 标识，所述第二 MAC/IP Advertisement Route 包括 MAC 地址和所述以太网段标识 ESI。所述第二 MAC/IP Advertisement Route 包括的 MAC 地址与所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址相同，具体地，所述第二 PE 设备获取所述第一消息中携带的所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址，并将其封装到所述第二消息携带的第二 MAC/IP Advertisement Route 中。所述第二消息携带的所述下一跳网络地址为所述第二 PE 设备的网络地址，例如，所述第二 PE 设备的 loopback 地址。

S310. 所述第二 PE 设备向所述第一 PE 设备发送所述第二消息。

在一个具体的实施方式中，所述第二消息为 BGP Update 消息，在本申请中，将该 BGP Update 消息称之为第二 BGP Update 消息。该第二 BGP Update 消息携带所述第二 MAC/IP Advertisement Route。所述第二 BGP Update 消息包括 VLAN 属性字段，用于携带所述第一 VLAN 标识。关于第二 BGP Update 消息的具体格式，第二 MAC/IP Advertisement Route 的具体格式以及携带所述第一 VLAN 标识的 VLAN 属性的具体格式的相关说明，参见 S202 中对于第一 BGP Update 消息的具体格式，第一 MAC/IP Advertisement Route 的具体格式以及携带所述第一 VLAN 标识的 VLAN 属性的具体描述，此处不再赘述。

在一个具体的实施方式中，所述第二消息中还携带指示信息，所述指示信息用于指示所述第一 PE 设备在收到所述第二消息后，避免向所述第二 PE 设备发送到达所述第二 MAC/IP Advertisement Route 包括的 MAC 地址的 MAC/IP Advertisement Route 和所述第一 VLAN 标识，进而避免形成报文环路。所述第二消息包括一个标记 (Flag) 字段，用于携带所述指示信息。可选的，在所述 VLAN 属性字段中设置一个标记 (Flag) 字段，用于携带该指示信息。该 Flag 字段的长度例如可以是一个比特，或者一个字节，本申请对此不作限定。可选的，还可以在第二消息中定义一个新的字段，例如 Flag 字段，用于携带该指示信息，不申请对此不作具体限定。

S311. 所述第一 PE 设备接收所述第二 PE 设备发送的第二消息。

S312. 所述第一 PE 设备生成第三 MAC 转发表项和第四 MAC 转发表项。

在一个具体的实施方式中，所述第二消息携带所述第一 VLAN 标识，所述第一 PE 设备接收所述第二消息以后，根据所述第二消息中携带的所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址，所述以太网段标识 ESI 和所述第一 VLAN 标识，生成第三 MAC 转发表项。所述第三 MAC 转发表项被所述第一 PE 设备用于转发目的 MAC 地址为所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址的报文，所述第三 MAC 转发表项包括的出接口的标识为所述第一接口的标识。所述第一 PE 设备生成第三 MAC 转发表项的方式，与前述 S305 中所述第二 PE 设备生成第一 MAC 转发表项的方式类似，具体的说明参见 S305，此处不再赘述。

需要说明的是，所述第一 PE 设备收到所述第二消息以后，可以在以下几种场景中生成所述第三 MAC 转发表项：

场景一：第一 PE 设备上没有到达所述终端设备的本地 MAC 路由，例如第一 PE 设备与所述 CE 设备之间的链路发生故障，导致第一 PE 设备从 CE 设备学习到的本地 MAC 路由被撤销。则所述第一 PE 设备可以根据所述第一 PE 设备携带的所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址以及所述以太网段标识 ESI，生成所述第三  
5 MAC 转发表项。

场景二：第一 PE 设备在收到所述第二消息时，第一 PE 设备上具有到达所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址的本地 MAC 路由，则第一 PE 设备保存所述第二消息中携带的路由信息。当第一 PE 设备与所述 CE 设备之间的链路发生故障，导致第一 PE 设备保存的本地 MAC 路由被撤销时，第一 PE 设备根据所述第二消息  
10 中携带的路由信息，即所述所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址以及所述以太网段标识 ESI，生成所述第三 MAC 转发表项。

上述两种场景仅是例举，本申请不限于此。

所述第一 PE 设备根据所述目的 MAC 地址和所述第二 PE 设备的网络地址，生成第四 MAC 转发表项，所述第四 MAC 转发表项为用于转发目的 MAC 地址为所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址的报文的备用转发表项，所述第四转发表项  
15 包含的下一跳网络地址为所述第二 PE 设备的网络地址。所述第一 PE 设备生成所述第四 MAC 转发表项的方式，与前述 S307 中所述第二 PE 设备生成所述第二 MAC 转发表项的方式类似，具体的说明参见 S307，此处不再赘述。

在一个具体的实施方式中，所述第一 PE 设备生成所述第四转发表项之后，根据所述  
20 所述第一 PE 设备到达所述 CE 设备的本地 MAC 转发表项（例如，第三 MAC 转发表项，或者在生成所述第三 MAC 转发表项之前已经存在的本地 MAC 转发表项）和所述第四 MAC 转发表项实现快速重路由 FRR。

由此，通过上述的方法，当第一 PE 设备收到已知单播流量（数据报文中携带的目的 MAC 地址为所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址）时，第一  
25 PE 设备查询 MAC 转发表，当第一链路正常工作时，根据本地 MAC 转发表项（也可称之为第五 MAC 转发表项，第五 MAC 转发表项例如可以是所述第三 MAC 转发表项，或者在生成所述第三 MAC 转发表项之前已经保存的本地 MAC 转发表项）的指示，直接通过 CE 设备转发报文。当第一链路存在故障时，第一 PE 设备收到上述的已知单播流量后，第一 PE 设备查询 MAC 转发表，根据备用转发表项的指示，即第四 MAC 转发表项的指示，  
30 向所述第二 PE 设备转发所述流量，通过所述第二 PE 设备向所述 CE 设备转发所述流量，从而提高了故障的收敛速度。

需要说明的是，在本申请所述的方法 300 中，可以同时包括 S306-S308 以及 S309-S312，S306-S308 与 S309-S312 的执行顺序不分先后。

在本申请实施例，当所述第一 PE 设备到所述 CE 设备的链路发生故障后，第一  
35 PE 设备到达该 CE 设备的 MAC 路由不可达，则该第一 PE 设备会向该第二 PE 设备发送 MAC 路由撤销消息，以撤销所述第一 MAC/IP Advertisement Route 路由。第二 PE 设备接收该第一 PE 设备发送的 MAC 路由撤销消息后，不会立即删除所述第一 MAC/IP Advertisement Route 路由，而是启动老化定时器，到达所述老化定时器设定的老化时间后，如果没有收到更新的到达所述 CE 设备的 MAC 路由，则删除所述第一 MAC/IP

Advertisement Route 路由。由此可以避免第二 PE 设备在收到所述第一 PE 设备发送的 MAC 路由撤销消息后，向远端的 PE 设备，例如图 2 所示的第三 PE 设备 PE2，发送 MAC 路由撤销消息，从而导致第一 PE 设备上本地 MAC 路由由于接口故障被撤销，远端的 MAC 路由被第二 PE 设备通告撤销后，第一 PE 设备上没有到达所述 CE 设备的 MAC 路由。此时，如果 PE2 发送的流量到达所述第一 PE 设备，第一 PE 设备查询不到 MAC 转发表而无法转发流量。

以上，结合图 2-图 7 详细说明书根据本申请实施例提供的报文处理方法。以下，结合图 8-图 15 详细说明根据本申请实施例提供的用于报文处理的运营商边缘 PE 设备和系统。

图 8 是根据本申请一实施例提供的第一 PE 设备 400 的示意图。该第一 PE 设备 400 可以是图 2 中的 PE1-1，可以用于执行图 3，图 6 或图 7 所示的方法。用户边缘 CE 设备经由第一链路连接到所述第一 PE 设备的第一接口，所述 CE 设备经由第二链路连接到第二 PE 设备的第二接口。如图 8 所示，该第一 PE 设备 400 包括：处理模块 401 和发送模块 402。

处理模块 401，用于生成第一消息，所述第一消息携带第一媒体接入控制/互联网协议通告路由 MAC/IP Advertisement Route 以及虚拟局域网 VLAN 标识，所述第一 MAC/IP Advertisement Route 包括 MAC 地址以及用于标识以太网段 ES 的以太网段标识 ESI，所述第一 MAC/IP Advertisement Route 包括的 MAC 地址为所述 CE 设备的 MAC 地址或所述 CE 设备所管辖的终端设备的 MAC 地址，其中，所述以太网段 ES 包括所述第一链路和所述第二链路，所述 VLAN 标识用于指示所述第一 MAC/IP Advertisement Route 包括的 MAC 地址所属的 VLAN。

发送模块 402，用于向所述第二 PE 设备发送所述第一消息。所述第一 MAC/IP Advertisement Route 以及所述 VLAN 标识被所述第二 PE 设备用于生成第一 MAC 转发表项，所述第一 MAC 转发表项包含所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址以及所述 VLAN 标识，所述第一 MAC 转发表项包含的出接口标识为所述第二接口的标识，所述第一 MAC 转发表项被所述第二 PE 设备用于向所述 CE 设备转发目的 MAC 地址为所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址的报文。

上述技术方案中，在上述的第二 PE 设备没有从 CE 设备学习到 MAC 路由，也无法获取有效的 VLAN 信息的情况下，第一 PE 设备向第二 PE 设备通告从 CE 设备学习到的 MAC 路由以及所述 VLAN 标识。所述第二 PE 设备能够根据接收到的 MAC 路由以及 VLAN 标识，生成所述第一 MAC 转发表项。当第二 PE 设备接收到目的 MAC 地址为所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址的数据流时，可以根据该第一 MAC 转发表项，直接通过第二链路向所述 CE 设备转发所述数据流。所述 CE 设备向所述终端设备转发所述数据流。上述技术方案中，在 CE 设备多归多活接入多个 PE 设备的场景中，多个 PE 设备能够有效的形成负载分担，实现带宽资源的合理利用。进一步的，在 EVPN 中，PE 设备可以通过控制平面学习到所述 VLAN 标识。在一些可能的场景中，所述第二 PE 设备不能从 CE 设备直接学习到所述 VLAN 标识。本申请提供的上述技术方案，所述第二 PE 设备可以从所述第一 PE 设备学习到所述 VLAN 标识。进而，所述第二 PE 设备可以根据所述 VLAN 标识向 CE 设备转发流量。例如，在 CE 设备通过以太网标签

Tag 或者以太网 Tag 终结接入第二 PE 设备时, 所述第二 PE 可以根据所述 VLAN 标识, 向所述 CE 设备转发流量。

5 在一个具体的实施方式中, 所述第一消息为第一边界网关协议更新 BGP Update 消息, 所述第一 BGP Update 消息包括用于携带所述 VLAN 标识的 VLAN 属性字段。关于所述第一 BGP Update 消息的具体格式以及所述 VLAN 属性字段的格式 (如采用哪些字段或扩展字段), 可以参考上述方法实施例中对对应部分的描述, 此处不再赘述。

通过在 BGP 更新消息中扩展 VLAN 属性, 本申请可以有效利用现有的协议实现 VLAN 信息的通告。

10 在一个具体的实施方式中, 所述第一 PE 设备 400 还包括接收模块 403。在所述发送模块向所述第二 PE 设备发送所述第一消息之后, 接收模块 403, 用于接收所述第二 PE 设备发送的第二消息。所述第二消息携带第二 MAC/IP Advertisement Route, 下一跳网络地址和所述 VLAN 标识。所述第二 MAC/IP Advertisement Route 包括目的 MAC 地址以及所述以太网段标识 ESI; 所述第二 MAC/IP Advertisement Route 中的目的 MAC 地址与所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址相同, 所述第  
15 二消息携带的所述下一跳网络地址为所述第二 PE 设备的网络地址。所述以太网段标识 ESI 被所述第一 PE 设备用于确定所述第一 PE 设备连接所述 CE 设备的接口为所述第一接口。所述确定的第一接口, 所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址以及所述 VLAN 标识被所述第一 PE 设备用于生成第二 MAC 转发表项, 所述第二 MAC 转发表项包括所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址, 所述第二转发表项包括的出接口的标识为所述第一接口的标识, 所述第二 MAC 转发表项被所述第一 PE 设备用于转发目的 MAC 地址为所述第二 MAC/IP Advertisement Route 所包括的  
20 MAC 地址的报文的表项。

所述处理模块 401, 还用于根据所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址和所述第二 PE 设备的网络地址, 生成第三 MAC 转发表项。所述第三 MAC 转发表项包括所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址, 所述第三 MAC 转发表项包括的下一跳网络地址为所述第二 PE 设备的网络地址, 当所述第一链路发生故障时, 所述第三 MAC 转发表项被所述第一 PE 设备用于转发目的 MAC 地址为所述第二  
25 MAC/IP Advertisement Route 所包括的 MAC 地址的报文的表项。

30 所述第一 PE 设备生成所述第三转发表项和所述第四转发表项之后, 根据所述第三 MAC 转发表项和所述第四 MAC 转发表项实现快速重路由 FRR。当第一 PE 设备收到已知单播流量 (数据报文中携带的目的 MAC 地址为所述第二 MAC/IP Advertisement Route 包括 MAC 地址) 时, 第一 PE 设备查询 MAC 转发表, 当第一链路正常工作时, 根据第三 MAC 转发表项的指示, 直接通过所述 CE 设备向终端设备转发报文。当第一链路存在故障时, 第一 PE 设备收到上述的已知单播流量后, 第一 PE 设备查询 MAC 转发表, 根据  
35 备用转发表项的指示, 即第四 MAC 转发表项的指示, 向所述第二 PE 设备转发所述流量, 通过所述第二 PE 设备向所述 CE 设备转发所述流量, 从而提高了故障的收敛速度。

在一个具体的实施方式中, 所述第二消息还携带有指示信息。所述处理模块 503, 还用于在所述接收模块接收所述第二 PE 设备发送的第二消息以后, 根据所述指示信息的指示, 避免向所述第二 PE 设备发送到达所述第二 MAC/IP Advertisement Route 包

括的 MAC 地址的 MAC/IP Advertisement Route 和所述 VLAN 标识。从而避免形成报文环路。

图 9 是根据本申请一实施例提供的第二 PE 设备 500 的示意图。该第二 PE 设备 500 可以是图 2 中的 PE1-2，可以用于执行图 3，图 6 或图 7 所示的方法。用户边缘 CE 设备经由第一链路连接到所述第一 PE 设备的第一接口，所述 CE 设备经由第二链路连接到第二 PE 设备的第二接口。如图 9 所示，该第一 PE 设备 500 包括：接收模块 501 和处理模块 502。

该接收模块 501，用于接收所述第一 PE 设备发送的第一消息。所述第一消息携带第一媒体接入控制/互联网协议通告路由 MAC/IP Advertisement Route 以及第一 VLAN 标识。所述第一 MAC/IP Advertisement Route 包括 MAC 地址以及用于标识以太网段 ES 的以太网段标识 ESI，所述第一 MAC/IP Advertisement Route 包括的 MAC 地址为所述 CE 设备的 MAC 地址或所述 CE 设备所管辖的终端设备的 MAC 地址。所述以太网段 ES 包括所述第一链路和所述第二链路。所述第一 VLAN 标识用于指示所述第一 MAC/IP Advertisement Route 包括的 MAC 地址所属的 VLAN。

该处理模块 502，用于根据所述接收模块 501 接收的所述第一消息中携带的所述以太网段标识 ESI 确定所述第二 PE 设备连接所述 CE 设备的接口为所述第二接口。

该处理模块 502，还用于根据所述确定的第二接口、所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址以及所述第一 VLAN 标识，生成第一 MAC 转发表项。所述第一 MAC 转发表项包含所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址和所述第一 VLAN 标识。所述第一 MAC 转发表项包含的出接口标识为所述第二接口的标识。所述第一 MAC 转发表项被所述第二 PE 设备用于向所述 CE 设备转发携带有所述第一 VLAN 标识并且目的 MAC 地址为所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址的报文。

上述技术方案中，在上述的第二 PE 设备没有从 CE 设备学习到 MAC 路由，也无法获取终端设备所属的 VLAN 信息的情况下，第一 PE 设备向第二 PE 设备通告到达所述终端设备的 MAC 路由以及所述终端设备所属的 VLAN 标识。所述第二 PE 设备能够根据接收到的 MAC 路由以及 VLAN 标识，生成所述第一 MAC 转发表项。当第二 PE 设备接收到目的 MAC 地址为所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址的数据流时，可以根据该第一 MAC 转发表项，直接通过第二链路向所述 CE 设备转发所述数据流。所述 CE 设备向所述终端设备转发所述数据流。上述技术方案中，在 CE 设备多归多活接入多个 PE 设备的场景中，多个 PE 设备能够有效的形成负载分担，实现带宽资源的合理利用。进一步的，在 EVPN 中，PE 设备可以通过控制平面学习到终端设备所属的 VLAN 的 VLAN 信息。在一些可能的场景中，所述第二 PE 设备不能从 CE 设备直接学习到终端设备的 VLAN 信息。本申请提供的上述技术方案，所述第二 PE 设备可以从所述第一 PE 设备学习到所述终端设备的 VLAN 信息。进而，所述第二 PE 设备可以根据所述终端设备的 VLAN 信息向 CE 设备转发到达所述终端设备的流量。例如，在 CE 设备通过以太网标签 Tag 或者以太网 Tag 终结接入第二 PE 设备时，所述第二 PE 可能不能从 CE 设备直接学习到终端设备的 VLAN 信息。

在一个具体的实施方式中，所述第一消息中还携带有下一跳网络地址，所述第一

消息中的下一跳网络地址为所述第一 PE 设备的网络地址，例如，所述第一 PE 设备的环回地址。所述处理模块 502，还用于根据所述第一消息，获取所述第一 PE 设备的网络地址。所述处理模块 502，进一步用于根据所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址以及所述第一 PE 设备的网络地址，生成第二 MAC 转发表项。所述第二 MAC 转发表项包括所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址，所述第二 MAC 转发表项包含的下一跳网络地址为所述第一 PE 设备的网络地址。其中，当所述第二链路发生故障时，所述第二 MAC 转发表项被所述第二 PE 用于转发携带有所述第一 VLAN 标识并且目的 MAC 地址为所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址的报文。

10 当第二 PE 设备收到去往所述终端设备的已知单播流量（数据报文中携带的目的 MAC 地址为终端设备的 MAC 地址）时，第二 PE 设备查询 MAC 转发表，在所述第二链路处于正常工作状态时，根据所述第一 MAC 转发表项的指示，经由所述第二链路，直接通过所述 CE 设备向终端设备转发报文。当所述第二链路故障时，第二 PE 设备收到去往终端设备的已知单播流量（数据报文中携带的目的 MAC 地址为终端设备的 MAC 地址）  
15 时，根据第二 MAC 转发表项的指示，向所述第一 PE 设备转发所述流量，通过所述第一 PE 设备向所述终端设备转发所述流量，从而提高了故障的收敛速度。

在一个具体的实施方式中，所述第一消息为第一边界网关协议更新 BGP Update 消息，所述第一 BGP Update 消息包括用于携带所述第一 VLAN 标识的 VLAN 属性字段。关于所述第一 BGP Update 消息的具体格式以及所述 VLAN 属性字段的格式（如采用哪些字段或扩展字段），可以参考上述方法实施例中对应部分的描述，此处不再赘述。  
20

通过在 BGP 更新消息中扩展 VLAN 属性，本申请可以有效利用现有的协议实现 VLAN 信息的通告。

在一个具体的实施方式中，所述第二 PE 设备还包括发送模块 503。在所述处理模块根据所述确定的第二接口、所述第一 MAC/IP Advertisement Route 所包括的 MAC  
25 地址以及所述第一 VLAN 标识，生成第一 MAC 转发表项以后，

所述接收模块 501，还用于接收携带有第二 VLAN 标识的第一报文，所述第一报文的  
目的 MAC 地址为所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址。

所述处理模块 502，还用于根据所述第一 MAC 转发表项将所述第二报文中携带的  
30 第二 VLAN 标识替换为所述第一 VLAN 标识，得到所述携带有所述第一 VLAN 标识的第二  
报文。

所述发送模块 503，用于向所述 CE 设备发送所述第二报文。

在一个具体的实施方式中，在所述接收模块接收所述第一 PE 设备发送的第一消息  
之后，

所述处理模块 502，还用于生成第二消息。具体地，所述处理模块 502 根据所述  
35 接收模块 501 接收的所述第一消息生成所述第二消息。所述第二消息携带第二 MAC/IP  
Advertisement Route，下一跳网络地址和所述第一 VLAN 标识，所述第二 MAC/IP  
Advertisement Route 包括 MAC 地址和所述以太网段标识 ESI。所述第二 MAC/IP  
Advertisement Route 包括的 MAC 地址与所述第一 MAC/IP Advertisement Route 所包  
括的 MAC 地址相同，具体地，所述处理模块 502 获取所述第一消息中携带的所述第一

MAC/IP Advertisement Route 所包括的 MAC 地址，并将其封装在所述第二消息携带的所述第二 MAC/IP Advertisement Route 中。所述第二消息携带的所述下一跳网络地址为所述第二 PE 设备的网络地址。

所述发送模块 503，还用于向所述第一 PE 设备发送所述第二消息，所述第二消息被所述第一 PE 设备用于生成第三 MAC 转发表项和第四 MAC 转发表项。其中，

所述第三 MAC 转发表项被所述第一 PE 设备用于转发目的 MAC 地址为所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址的报文，所述第三 MAC 转发表项包括所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址和所述第一 VLAN 标识，所述第三 MAC 转发表项包含的出接口的标识为所述第一接口的标识。当所述第一链路发生故障时，所述第四 MAC 转发表项被所述第一 PE 设备用于转发目的 MAC 地址为所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址的报文，所述第四 MAC 转发表项包括所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址，所述第四转发表项包含的下一跳网络地址为所述第二 PE 设备的网络地址。

通过上述方案，第二 PE 设备在收到所述第一消息以后，根据第一消息中携带的信息，在控制平面生成本地的主用 MAC 路由表项以及用于快速重路由的备用 MAC 路由表项，进一步的，控制平面将上述主用 MAC 路由表项和备用 MAC 路由表项下发到转发平面，生成用于实现快速重路由的第一 MAC 转发表项和第二 MAC 转发表项。并且，第二 PE 设备在收到所述第一消息以后，生成本地的所述第一 MAC 转发表项后，还会将该本地的 MAC 路由回送给所述第一 PE 设备，使得第一 PE 设备生成用于实现快速从路由的备用 MAC 路由表项。而第一 PE 设备从 CE 设备学习到的 MAC 路由可以作为本地 MAC 路由，即主用 MAC 路由。当第一 PE 设备和 CE 设备连接的链路，例如，所述第一链路发生链路故障时，所述第一 PE 设备的本地 MAC 路由被撤销。当所述第一链路故障恢复后，第一 PE 设备可以根据所述第二 PE 设备通告的所述第二 MAC/IP Advertisement Route 路由和所述第一 VLAN 标识，再次生成所述本地 MAC 路由，用于指导向所述 CE 设备转发流量。从而使得，在第一链路发送故障并再次恢复后，能够快速实现路由重定向，实现故障的快速收敛。

在一个具体的实施方式中，所述第二消息还携带有指示信息，所述指示信息用于指示所述第一 PE 设备在收到所述第二消息后，避免向所述第二 PE 设备发送到达所述第二 MAC/IP Advertisement Route 包括的 MAC 地址的 MAC/IP Advertisement Route 和所述第一 VLAN 标识，从而避免报文环路。

图 10 是根据本申请一实施例提供的第一 PE 设备 600 的示意图。该第一 PE 设备 600 可以是图 2 中的 PE1-1，可以用于执行图 3，图 6 或图 7 所示的方法。用户边缘 CE 设备经由第一链路连接到所述第一 PE 设备的第一接口，所述 CE 设备经由第二链路连接到第二 PE 设备的第二接口。如图 10 所示，所述第一 PE 设备 600 包括：输入接口 601、输出接口 602、处理器 603 和存储器 604。该输入接口 601、输出接口 602、处理器 603 和存储器 604 可以通过总线系统 605 相连。

所述存储器 604 用于存储包括程序。所述处理器 604，用于执行所述存储器 604 中的程序，以控制输入接口 601 接收信号、控制输出接口 602 发送信号以及实施图 3，或图 6 或图 7 所对应的实施方式中第一 PE 设备所实施的各步骤及功能，此处不再赘述。

上述输入接口 601、输出接口 602 以及处理器 603 的具体实施方式可以相应参考上述图 8 实施方式中的接收模块 403，发送模块 402 以及处理模块 401 中的具体说明，这里不再赘述。

5 图 11 是根据根据本申请一实施例提供的第二 PE 设备 700 的示意图。该第二 PE 设备 700 可以是图 2 中的 PE1-2，可以用于执行图 3，图 6 或图 7 所示的方法。用户边缘 CE 设备经由第一链路连接到所述第一 PE 设备的第一接口，所述 CE 设备经由第二链路连接到第二 PE 设备的第二接口。入图 11 所示，该第二 PE 设备 700 包括：输入接口 701、输出接口 702、处理器 703 和存储器 704。该输入接口 701、输出接口 702、处理器 703 和存储器 704 可以通过总线系统 705 相连。

10 所述存储器 704 用于存储包括程序。所述处理器 704，用于执行所述存储器 704 中的程序，以控制输入接口 701 接收信号、控制输出接口 702 发送信号以及实施图 3，或图 6 或图 7 所对应的实施方式中第一 PE 设备所实施的各步骤及功能，此处不再赘述。上述输入接口 701、输出接口 702、处理器 703 的具体实施方式可以相应参考上述图 9 实施方式中的接收模块 501，发送模块 503 以及处理模块 502 中的具体说明，这里不再赘述。

15 应理解，在本申请实施例中，该处理器 703 和处理器 803 可以是中央处理单元（Central Processing Unit，简称为“CPU”），还可以是其他通用处理器、数字信号处理器（DSP）、专用集成电路（ASIC）、现成可编程门阵列（FPGA）或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。

20 该存储器 704 和存储器 804 可以包括只读存储器和随机存取存储器，并分别向处理器 703 和处理器 803 提供指令和数据。存储器 704 或存储器 804 的一部分还可以包括非易失性随机存取存储器。例如，存储器 704 或存储器 804 还可以存储设备类型的信息。

25 该总线系统 705 和总线系统 805 除包括数据总线之外，还可以包括电源总线、控制总线和状态信号总线等。但是为了清楚说明起见，在图中将各种总线都标为总线系统。

30 在实现过程中，方法 300 的各步骤可以通过处理器 603 和处理器 703 中的硬件的集成逻辑电路或者软件形式的指令完成。结合本申请实施例所公开的定位方法的步骤可以直接体现为硬件处理器执行完成，或者用处理器中的硬件及软件模块组合执行完成。软件模块可以位于随机存储器，闪存、只读存储器，可编程只读存储器或者电可擦写可编程存储器、寄存器等本领域成熟的存储介质中。该存储介质分别位于存储器 604 和存储器 704 中，处理器 603 读取存储器 604 中的信息，处理器 703 读取存储器 704 中的信息，结合其硬件完成上述方法 300 的步骤。为避免重复，这里不再详细描述。

35 需要说明的是，一个具体的实施方式中，图 8 中的处理模块 401 可以用图 10 的处理器 603 实现，发送模块 402 可以由图 10 的输出接口 602 实现，接收模块 403 可以由图 10 的输入接口 601 实现。同理，图 9 中的处理模块 502 用图 11 的处理器 703 实现，发送模块 503 可以由图 11 的输出接口 702 实现，接收模块 501 可以用由图 11 的输入

接口 701 实现。

可以理解的是，图 8 至图 11 仅仅示出了第一 PE 设备和第二 PE 设备的简化设计。在实际应用中，第一 PE 设备和第二 PE 设备可以分别包含任意数量的接口，处理器和存储器等。

5 图 12 为本申请实施例提供的另一种第一 PE 设备 1200 的硬件结构示意图。图 12 所示的第一 PE 设备 1200 可以用于执行上述实施例的方法中第一 PE 设备执行的相应步骤。

如图 12 所示，所述第一 PE 设备包括：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。  
10 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，同步点（英文：Point of Synchronization, POS）接口、前兆以太网（英文：Gigabit Ethernet, GE）接口、异步传输模式（英文：Asynchronous Transfer Mode, 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。  
15

主控板 210 上的中央处理器 1211 生成第一消息，并经由接口板 1230 上的中央处理器向物理接口卡 1223 向第二 PE 设备发送所述第一消息。其中，所述第一消息携带第一 MAC/IP Advertisement Route 以及 VLAN 标识，所述第一 MAC/IP Advertisement  
20 Route 包括 MAC 地址以及用于标识以太网段 ES 的以太网段标识 ESI，所述第一 MAC/IP Advertisement Route 包括的 MAC 地址为所述 CE 设备的 MAC 地址或所述 CE 设备所管辖的终端设备的 MAC 地址。其中，所述以太网段 ES 包括所述第一链路和所述第二链路，所述 VLAN 标识用于指示所述第一 MAC/IP Advertisement Route 包括的 MAC 地址所属的 VLAN。所述第一 MAC/IP Advertisement Route 以及所述 VLAN 标识被所述第二 PE  
25 设备用于生成第一 MAC 转发表项，所述第一 MAC 转发表项包含所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址以及所述 VLAN 标识，所述第一 MAC 转发表项包含的出接口标识为所述第二接口的标识。所述第一 MAC 转发表项被所述第二 PE 设备用于向所述 CE 设备转发目的 MAC 地址为所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址的报文。  
30

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 PE 设备发送的第二消息。所述第二消息携带第二 MAC/IP Advertisement Route，下一跳网络地址和所述 VLAN 标识，所述第二 MAC/IP Advertisement Route 包括目的 MAC 地址以及所述以太网段标识 ESI。所述第二 MAC/IP Advertisement Route 中的目的 MAC 地址为所述第二 MAC/IP Advertisement Route 所  
35 包括的 MAC 地址。所述第二消息携带的所述下一跳网络地址为所述第二 PE 设备的网络地址；所述以太网段标识 ESI 被所述第一 PE 设备用于确定所述第一 PE 设备连接所述 CE 设备的接口为所述第一接口。所述确定的第一接口，所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址以及所述 VLAN 标识被所述第一 PE 设备用于生成第三 MAC 转发表项。所述第三 MAC 转发表项包括所述第二 MAC/IP Advertisement Route 所包括的

MAC 地址, 所述第三转发表项包括的出接口的标识为所述第一接口的标识。所述第三 MAC 转发表项被所述第一 PE 设备用于转发目的 MAC 地址为所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址的报文。

具体地, 主控板 1210 上的中央处理器 1211 根据所述 ESI 确定所述第一 PE 设备连接所述 CE 设备的接口为所述第一接口, 并根据所述确定的第一接口, 所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址以及所述 VLAN 标识生成第三 MAC 转发表项。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 发送所述第三 MAC 转发表项。

接口板 1230 上的转发表项存储器 1234 用于保存所述第三 MAC 转发表项。接口板 1230 上的中央处理器 1231 用于控制网络存储器 1232 获取转发表项存储器 1234 中的 MAC 转发表项。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 完成流量的接收和转发。

应理解, 本申请实施例中接口板 1240 上的操作与接口板 1230 的操作一致, 为了简洁, 不再赘述。应理解, 本实施例中第一 PE 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。

此外, 需要说明的是, 在本申请中, 第二 PE 设备可以具有和第一 PE 设备相同的结构。在 PE 设备中, 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 PE 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, PE 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, PE 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 PE 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。

图 13 为本申请实施例提供的另一种第一 PE 设备 1300 的硬件结构示意图。图 13 所示的第一 PE 设备 1300 可以用于执行上述实施例的方法中第一 PE 设备执行的相应步骤。

第一 PE 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 12 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 13 所示的控制器 1210A), 剩下的形成另一独立的物理设备(即如图 13 所示的第一 PE 转发设备 1200A)。控制器 1210A 与第一 PE 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 12 所对应的实施例相比, 本实施中的第一 PE 设备 1300 包括分离出去的控制 35 器 1210A 和第一 PE 转发设备 1200A, 即在该实施例中, 该第一 PE 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、随机读取存储器（英文：Random Access Memory, RAM）、只读存储器（英文：Read-only Memory, 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。

5 其中，当需要运行控制器时，通过固化在 ROM 中的基本输入/输出系统（英文：Basic Input/output System, BIOS）或者嵌入式系统中的引导器（英文：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 12 中主控板 1210 的所有功能和步骤。

10 第一 PE 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 12 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 PE 转发设备，所述虚拟第一 PE 转发设备为虚拟路由器。在虚拟第一 PE 转发设备的场景下，上述实体第一 PE

15 转发设备实施例中提到的该第一 PE 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 PE 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 PE 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 PE 设备的功能或步骤具体可以参考图 10 的实施例。

20 应理解，本实施例中第一 PE 设备 1300 中的控制器 1210A 和该第一 PE 转发设备 1200A 可以实现方法实施例中的第一 PE 设备所实施的各种功能、步骤，为了简洁，在此不再赘述。

图 14 提供了为本申请实施例提供的另一种第二 PE 设备 1400 的硬件结构示意图。图 14 所示的第二 PE 设备 1400 可以用于执行上述实施例的方法中第二 PE 设备执行的

25 相应步骤。

如图 14 所示，所述第二 PE 设备包括：主控板 1410、接口板 1430、交换网板 1420 和接口板 1440。主控板 1410 用于完成系统管理、设备维护、协议处理等功能。交换网板 1420 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 1430 和 1440 用于提供各种业务接口（例如，POS 接口、GE 接口、ATM 接口等），

30 并实现数据包的转发。主控板 1410、接口板 1430 和 1440，以及交换网板 1420 之间通过系统总线与系统背板相连实现互通。接口板 1430 上的中央处理器 1431 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 1430 上的物理接口卡 1433 接收第一 PE 设备发送的第一消息，其中，所述第一消息携带第一 MAC/IP Advertisement Route 以及第一 VLAN 标识，所述第一 MAC/IP

35 Advertisement Route 包括 MAC 地址以及用于标识以太网段 ES 的以太网段标识 ESI，所述第一 MAC/IP Advertisement Route 包括的 MAC 地址为所述 CE 设备的 MAC 地址或所述 CE 设备所管辖的终端设备的 MAC 地址。其中，所述以太网段 ES 包括所述第一链路和所述第二链路，所述第一 VLAN 标识用于指示所述第一 MAC/IP Advertisement Route 包括的 MAC 地址所属的 VLAN。并且，经由接口板 1430 上的中央处理器 1431 向

主控板 1410 上的中央处理器 1411 发送所述第一 MAC/IP Advertisement Route 和所述第一 VLAN 标识。

5 主控板 1410 上的中央处理器 1411 用于获取所述第一 MAC/IP Advertisement Route 和所述第一 VLAN 标识, 并且, 根据所述主控板 1410 上的中央处理器 1411 根据所述 ESI 确定所述第二 PE 设备连接所述 CE 设备的接口为所述第二接口。进一步的, 所述中央处理器 1411 根据所述确定的第二接口, 所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址以及所述第一 VLAN 标识生成第一 MAC 转发表项。所述第一 MAC 转发表项包含所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址和所述第一 VLAN 标识, 所述第一 MAC 转发表项包含的出接口标识为所述第二接口的标识, 所述第一 MAC 转发表项被所述第二 PE 设备用于向所述 CE 设备转发携带有所述第一 VLAN 标识并且目的 MAC 地址为所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址的报文。

10 主控板 1410 上的中央处理器 1411 经由接口板 1430 上的中央处理器 1431 向接口板 1430 上的转发表项存储器 1434 发送所述第一 MAC 转发表项。

15 接口板 1430 上的转发表项存储器 1434 用于保存所述第一 MAC 转发表项。接口板 1430 上的中央处理器 1431 用于控制网络存储器 1432 获取转发表项存储器 1434 中的 MAC 转发表项, 并且, 中央处理器 1431 用于控制网络存储器 1432 经由物理接口卡 1433 完成流量的接收和发送。

主控板 1410 上的中央处理器 1211 还用于控制接口板 1430 完成流量的转发和处理。

20 应理解, 本发明实施例中接口板 1440 上的操作与所述接口板 1430 的操作一致, 为了简洁, 不再赘述。应理解, 本实施例的第二 PE 设备 1400 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。

图 15 为本申请实施例提供的另一种第二 PE 设备 1500 的硬件结构示意图。图 15 所示的第二 PE 设备 1500 可以用于执行上述实施例的方法中第二 PE 设备执行的相应步骤。

25 第二 PE 设备 1500 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, SDN)。在 SDN 中, 如图 14 所示的第二 PE 设备 1400 的主控板 1410 从设备中分离出来, 形成新的独立的物理设备(即如图 15 所示的控制器 1410A), 剩下的形成另一独立的物理设备(即如图 14 所示的第二 PE 转发设备 1400A)。控制器 1410A 与第二 PE 转发设备 1400A 通过控制通道协议实现交互。控制通道协议可以是 OpenFlow 协议、PCEP、BGP、I2RS 等。也就是说, 与上述图 14 所对应的实施例相比, 本实施中的第二 PE 设备 1500 包括分离出去的控制器 1410A 和第二 PE 转发设备 1400A, 即在该实施例中, 30 该第二 PE 设备 1500 也可以看成是一个系统。

35 控制器 1410A 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 boot loader 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 14 中主控板 1410 的所有功能和步骤。

第二 PE 转发设备 1400A 可以是基于专用的硬件结构实现,其功能和结构与上述图 14 中的接口板 1430、接口板 1440 和交换网板 1420 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: Netwrk Function Virtulization, 缩写: NFV)技术实现的虚拟第二 PE 转发设备,所述虚拟第二 PE 转发设备为虚拟路由器。在虚拟第二 PE 转发设备的场景下,上述实体第二 PE 转发设备实施例中提到的该第一 PE 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二 PE 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二转发 PE 设备的功能或步骤,或者采用通用物理服务器并利用 NFV 技术实施该第二转发 PE 设备的功能或步骤具体可以参考图 11 的实施例。

应理解,本实施例中第二 PE 设备 1500 中的控制器 1410A 和该第二 PE 转发设备 1400A 可以实现方法实施例中的第二 PE 设备所实施的各种功能、步骤,为了简洁,在此不再赘述。

本申请还提供了一种通信系统,包括第一 PE 设备和第二 PE 设备,所述第一 PE 设备可以是图 8、图 10、图 12 或图 13 对应的实施例所提供的第一 PE 设备。所述第二 PE 设备可以是图 9、图 11、图 14 或图 15 对应的实施例所提供的第二 PE 设备。所述通信系统用于执行图 2-图 7 对应的实施例的方法 300。

应理解,在本申请的各种实施例中,各过程的序号的大小并不意味着执行顺序的先后,各过程的执行顺序应以其功能和内在逻辑确定,而不应对本申请实施例的实施过程构成任何限定。

本领域普通技术人员可以意识到,结合本文中所公开的实施例描述的各示例的模块及方法步骤,能够以电子硬件、或者计算机软件和电子硬件的结合来实现。这些功能究竟以硬件还是软件方式来执行,取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能。

所属领域的技术人员可以清楚地了解到,为描述的方便和简洁,上述描述的系统、装置和模块的具体工作过程,可以参考前述方法实施例中的对应过程,在此不再赘述。

在上述实施例中,可以全部或部分地通过软件、硬件、固件或者其任意组合来实现。当使用软件实现时,可以全部或部分地以计算机程序产品的形式实现。所述计算机程序产品包括一个或多个计算机指令。在计算机上加载和执行所述计算机程序指令时,全部或部分地产生按照本申请实施例所述的流程或功能。所述计算机可以是通用计算机、专用计算机、计算机网络、或者其他可编程装置。所述计算机指令可以存储在计算机可读存储介质中,或者从一个计算机可读存储介质向另一个计算机可读存储介质传输,例如,所述计算机指令可以从一个网站站点、计算机、服务器或数据中心通过有线(例如同轴电缆、光纤、数字用户线(DSL))或无线(例如红外、无线、微波等)方式向另一个网站站点、计算机、服务器或数据中心进行传输。所述计算机可读存储介质可以是计算机能够存取的任何可用介质或者是包含一个或多个可用介质集成的服务器、数据中心等数据存储设备。所述可用介质可以是磁性介质,(例如,软盘、硬盘、磁带)、光介质(例如,DVD)、或者半导体介质(例如固态硬盘 Solid State Disk (SSD))等。本说明书的各个部分均采用递进的方式进行描述,各个实施例之间相同相

似的部分互相参见即可，每个实施例重点介绍的都是与其他实施例不同之处。尤其，对于装置和系统实施例而言，由于其基本相似于方法实施例，所以描述的比较简单，相关之处参见方法实施例部分的说明即可。

5 以上所述，仅为本发明的具体实施方式，但本发明的保护范围并不局限于此，任何熟悉本技术领域的技术人员在本发明揭露的技术范围内，可轻易想到变化或替换，都应涵盖在本发明的保护范围之内。因此，本发明的保护范围应以所述权利要求的保护范围为准。

## 权 利 要 求 书

1. 一种报文处理方法, 应用于以太网虚拟私有网络 EVPN 中, 用户边缘 CE 设备经由第一链路连接到第一运营商边缘 PE 设备的第一接口, 所述 CE 设备经由第二链路连接到第二 PE 设备的第二接口, 其特征在于, 所述方法包括:

5 所述第一 PE 设备生成第一消息, 所述第一消息携带第一媒体接入控制/互联网协议通告路由 MAC/IP Advertisement Route 以及虚拟局域网 VLAN 标识, 所述第一 MAC/IP Advertisement Route 包括 MAC 地址以及用于标识以太网段 ES 的以太网段标识 ESI, 所述第一 MAC/IP Advertisement Route 包括的 MAC 地址为所述 CE 设备的 MAC 地址或所述 CE 设备所管辖的终端设备的 MAC 地址, 其中, 所述以太网段 ES 包括所述第一链  
10 路和所述第二链路, 所述 VLAN 标识用于指示所述第一 MAC/IP Advertisement Route 包括的 MAC 地址所属的 VLAN;

所述第一 PE 设备向所述第二 PE 设备发送所述第一消息, 所述第一 MAC/IP Advertisement Route 以及所述 VLAN 标识被所述第二 PE 设备用于生成第一 MAC 转发表项, 所述第一 MAC 转发表项包含所述第一 MAC/IP Advertisement Route 所包括的  
15 MAC 地址以及所述 VLAN 标识, 所述第一 MAC 转发表项包含的出接口标识为所述第二接口的标识, 所述第一 MAC 转发表项被所述第二 PE 设备用于向所述 CE 设备转发目的 MAC 地址为所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址的报文。

2. 根据权利要求 1 所述方法, 其特征在于, 所述第一消息为第一边界网关协议更新 BGP Update 消息, 所述第一 BGP Update 消息包括用于携带所述 VLAN 标识的 VLAN  
20 属性字段。

3. 根据权利要求 1 或 2 所述的方法, 其特征在于, 在所述第一 PE 设备向所述第二 PE 设备发送所述第一消息之后, 所述方法还包括:

所述第一 PE 设备接收所述第二 PE 设备发送的第二消息, 所述第二消息携带第二  
25 MAC/IP Advertisement Route, 下一跳网络地址和所述 VLAN 标识, 所述第二 MAC/IP Advertisement Route 包括 MAC 地址以及所述以太网段标识 ESI; 所述第二 MAC/IP Advertisement Route 中的 MAC 地址与所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址相同, 所述第二消息携带的所述下一跳网络地址为所述第二 PE 设备的网络地址; 所述以太网段标识 ESI 被所述第一 PE 设备用于确定所述第一 PE 设备连接所述 CE 设备的接口为所述第一接口; 所述确定的第一接口, 所述第二 MAC/IP Advertisement  
30 Route 所包括的 MAC 地址以及所述 VLAN 标识被所述第一 PE 设备用于生成第二 MAC 转发表项, 所述第二 MAC 转发表项包括所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址, 所述第二转发表项包括的出接口的标识为所述第一接口的标识, 所述第二 MAC 转发表项被所述第一 PE 设备用于转发目的 MAC 地址为所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址的报文;

35 所述第一 PE 设备根据所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址和所述第二 PE 设备的网络地址, 生成第三 MAC 转发表项, 所述第三 MAC 转发表项包括所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址, 所述第三 MAC 转发表项包括的下一跳网络地址为所述第二 PE 设备的网络地址, 当所述第一链路发生故障时, 所述第三 MAC 转发表项被所述第一 PE 设备用于转发目的 MAC 地址为所述第二 MAC/IP

Advertisement Route 所包括的 MAC 地址的报文。

4. 根据权利要求 3 所述的方法, 其特征在于, 所述第二消息还携带有指示信息;

所述第一 PE 设备接收所述第二 PE 设备发送的第二消息以后, 所述方法还包括:

5 所述第一 PE 设备根据所述指示信息的指示, 避免向所述第二 PE 设备发送到达所述第二 MAC/IP Advertisement Route 包括的 MAC 地址的 MAC/IP Advertisement Route 和所述 VLAN 标识。

5. 一种报文处理方法, 所述方法应用于以太网虚拟私有网络 EVPN 中, 用户边缘 CE 设备经由第一链路连接到第一运营商边缘 PE 设备的第一接口, 所述 CE 设备经由第二链路连接到第二 PE 设备的第二接口, 其特征在于, 所述方法包括:

10 所述第二 PE 设备接收所述第一 PE 设备发送的第一消息, 所述第一消息携带第一媒体接入控制/互联网协议通告路由 MAC/IP Advertisement Route 以及第一 VLAN 标识, 所述第一 MAC/IP Advertisement Route 包括 MAC 地址和用于标识以太网段 ES 的以太网段标识 ESI, 所述第一 MAC/IP Advertisement Route 包括的 MAC 地址为所述 CE 设备的 MAC 地址或所述 CE 设备所管辖的终端设备的 MAC 地址, 所述以太网段 ES 包括所述  
15 所述第一链路和所述第二链路, 所述第一 VLAN 标识用于指示所述第一 MAC/IP Advertisement Route 包括的 MAC 地址所属的 VLAN;

所述第二 PE 设备根据所述以太网段标识 ESI 确定所述第二 PE 设备连接所述 CE 设备的接口为所述第二接口;

20 所述第二 PE 设备根据所述确定的第二接口、所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址以及所述第一 VLAN 标识, 生成第一 MAC 转发表项, 所述第一 MAC 转发表项包含所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址和所述第一 VLAN 标识, 所述第一 MAC 转发表项包含的出接口标识为所述第二接口的标识, 所述第一 MAC 转发表项被所述第二 PE 设备用于向所述 CE 设备转发携带有所述第一 VLAN 标识并且目的 MAC 地址为所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址的  
25 报文。

6. 根据权利要求 5 所述的方法, 其特征在于, 所述第一消息中还携带有下一跳网络地址, 所述第一消息中的下一跳网络地址为所述第一 PE 设备的网络地址, 所述方法还包括:

所述第二 PE 设备根据所述第一消息, 获取所述第一 PE 设备的网络地址;

30 所述第二 PE 设备根据所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址以及所述第一 PE 设备的网络地址, 生成第二 MAC 转发表项, 所述第二 MAC 转发表项包括所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址, 所述第二 MAC 转发表项包含的下一跳网络地址为所述第一 PE 设备的网络地址;

35 其中, 当所述第二链路发生故障时, 所述第二 MAC 转发表项被所述第二 PE 用于转发携带有所述第一 VLAN 标识并且目的 MAC 地址为所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址的报文。

7. 根据权利要求 6 所述的方法, 其特征在于, 所述第一消息为第一边界网关协议更新 BGP Update 消息, 所述第一 BGP Update 消息包括用于携带所述第一 VLAN 标识的 VLAN 属性字段。

8. 根据权利要求 5-7 任一项所述的方法, 其特征在于, 所述第二 PE 设备根据所述确定的第二接口、所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址以及所述第一 VLAN 标识, 生成第一 MAC 转发表项以后, 所述方法还包括:

5 所述第二 PE 设备接收携带有第二 VLAN 标识的第一报文, 所述第一报文的目的地 MAC 地址为所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址;

所述第二 PE 设备根据所述第一 MAC 转发表项将所述第二报文中携带的第二 VLAN 标识替换为所述第一 VLAN 标识, 得到所述携带有所述第一 VLAN 标识的第二报文;

所述第二 PE 设备向所述 CE 设备发送所述第二报文。

10 9. 根据权利要求 5-8 中任一项所述的方法, 其特征在于, 在所述第二 PE 设备接收所述第一 PE 设备发送的第一消息之后, 所述方法还包括:

所述第二 PE 设备生成第二消息, 所述第二消息携带第二 MAC/IP Advertisement Route, 下一跳网络地址和所述第一 VLAN 标识, 所述第二 MAC/IP Advertisement Route 包括 MAC 地址和所述以太网段标识 ESI, 所述第二 MAC/IP Advertisement Route 包括的 MAC 地址与所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址相同, 所述第  
15 二消息携带的所述下一跳网络地址为所述第二 PE 设备的网络地址;

所述第二 PE 设备向所述第一 PE 设备发送所述第二消息, 所述第二消息被所述第一 PE 设备用于生成第三 MAC 转发表项和第四 MAC 转发表项; 其中,

20 所述第三 MAC 转发表项被所述第一 PE 设备用于转发目的 MAC 地址为所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址的报文, 所述第三 MAC 转发表项包括所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址和所述第一 VLAN 标识, 所述第三 MAC 转发表项包含的出接口的标识为所述第一接口的标识;

25 当所述第一链路发生故障时, 所述第四 MAC 转发表项被所述第一 PE 设备用于转发目的 MAC 地址为所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址的报文, 所述第四 MAC 转发表项包括所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址, 所述第四转发表项包含的下一跳网络地址为所述第二 PE 设备的网络地址。

30 10. 根据权利要求 9 所述的方法, 其特征在于, 所述第二消息还携带有指示信息, 所述指示信息用于指示所述第一 PE 设备在收到所述第二消息后, 避免向所述第二 PE 设备发送到达所述第二 MAC/IP Advertisement Route 包括的 MAC 地址的 MAC/IP Advertisement Route 和所述第一 VLAN 标识。

11. 一种第一运营商边缘 PE 设备, 应用于以太网虚拟私有网络 EVPN, 用户边缘 CE 设备经由第一链路连接到所述第一 PE 设备的第一接口, 所述 CE 设备经由第二链路连接到第二 PE 设备的第二接口, 其特征在于, 所述第一 PE 设备包括:

35 处理模块, 用于生成第一消息, 所述第一消息携带第一媒体接入控制/互联网协议通告路由 MAC/IP Advertisement Route 以及虚拟局域网 VLAN 标识, 所述第一 MAC/IP Advertisement Route 包括 MAC 地址以及用于标识以太网段 ES 的以太网段标识 ESI, 所述 MAC 地址为所述 CE 设备的 MAC 地址或所述 CE 设备所管辖的终端设备的 MAC 地址, 其中, 所述以太网段 ES 包括所述第一链路和所述第二链路, 所述 VLAN 标识用于指示所述第一 MAC/IP Advertisement Route 包括的 MAC 地址所属的 VLAN;

发送模块, 用于向所述第二 PE 设备发送所述第一消息, 所述第一 MAC/IP

Advertisement Route 以及所述 VLAN 标识被所述第二 PE 设备用于生成第一 MAC 转发表项, 所述第一 MAC 转发表项包含所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址以及所述 VLAN 标识, 所述第一 MAC 转发表项包含的出接口标识为所述第二接口的标识, 所述第一 MAC 转发表项被所述第二 PE 设备用于向所述 CE 设备转发目的 MAC 地址为所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址的报文。

12. 根据权利要求 11 所述的第一 PE 设备, 其特征在于, 所述第一消息为第一边界网关协议更新 BGP Update 消息, 所述第一 BGP Update 消息包括用于携带所述 VLAN 标识的 VLAN 属性字段。

13. 根据权利要求 11 或 12 所述的第一 PE 设备, 其特征在于, 所述第一 PE 设备还包括接收模块, 在所述发送模块向所述第二 PE 设备发送所述第一消息之后,

所述接收模块, 用于接收所述第二 PE 设备发送的第二消息, 所述第二消息携带第二 MAC/IP Advertisement Route, 下一跳网络地址和所述 VLAN 标识; 所述第二 MAC/IP Advertisement Route 包括 MAC 地址以及所述以太网段标识 ESI; 所述第二 MAC/IP Advertisement Route 中的 MAC 地址与所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址相同, 所述第二消息携带的所述下一跳网络地址为所述第二 PE 设备的网络地址; 所述以太网段标识 ESI 被所述第一 PE 设备用于确定所述第一 PE 设备连接所述 CE 设备的接口为所述第一接口; 所述确定的第一接口, 所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址以及所述 VLAN 标识被所述第一 PE 设备用于生成第二 MAC 转发表项, 所述第二 MAC 转发表项包括所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址, 所述第二 MAC 转发表项包含的出接口的标识为所述第一接口的标识, 所述第二 MAC 转发表项被所述第一 PE 设备用于转发目的 MAC 地址为所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址的报文的表项;

所述处理模块, 还用于根据所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址和所述第二 PE 设备的网络地址, 生成第三 MAC 转发表项, 所述第三 MAC 转发表项包括所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址, 所述第三 MAC 转发表项包括的下一跳网络地址为所述第二 PE 设备的网络地址, 当所述第一链路发生故障时, 所述第三 MAC 转发表项被所述第一 PE 设备用于转发目的 MAC 地址为所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址的报文的表项。

14. 根据权利要求 13 所述的第一 PE 设备, 其特征在于, 所述第二消息还携带有指示信息;

所述处理模块, 还用于在所述接收模块接收所述第二 PE 设备发送的第二消息以后, 根据所述指示信息的指示, 避免向所述第二 PE 设备发送到达所述第二 MAC/IP Advertisement Route 包括的 MAC 地址的 MAC/IP Advertisement Route 和所述 VLAN 标识。

15. 一种第二运营商边缘 PE 设备, 应用于以太网虚拟私有网络 EVPN 中, 用户边缘 CE 设备经由第一链路连接到第一运营商边缘 PE 设备的第一接口, 所述 CE 设备经由第二链路连接到第二 PE 设备的第二接口, 其特征在于, 所述第二 PE 设备包括:

接收模块, 用于接收所述第一 PE 设备发送的第一消息, 所述第一消息携带第一媒体接入控制/互联网协议通告路由 MAC/IP Advertisement Route 以及第一 VLAN 标识,

所述第一 MAC/IP Advertisement Route 包括 MAC 地址以及用于标识以太网段 ES 的以太网段标识 ESI, 所述第一 MAC/IP Advertisement Route 包括的 MAC 地址为所述 CE 设备的 MAC 地址或所述 CE 设备所管辖的终端设备的 MAC 地址, 所述以太网段 ES 包括所述第一链路和所述第二链路, 所述第一 VLAN 标识用于指示所述第一 MAC/IP Advertisement Route 包括的 MAC 地址所属的 VLAN;

处理模块, 用于根据所述以太网段标识 ESI 确定所述第二 PE 设备连接所述 CE 设备的接口为所述第二接口;

所述处理模块, 还用于根据所述确定的第二接口、所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址以及所述第一 VLAN 标识, 生成第一 MAC 转发表项, 所述第一 MAC 转发表项包含所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址和所述第一 VLAN 标识, 所述第一 MAC 转发表项包含的出接口标识为所述第二接口的标识, 所述第一 MAC 转发表项被所述第二 PE 设备用于向所述 CE 设备转发携带有所述第一 VLAN 标识并且目的 MAC 地址为所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址的报文。

16. 根据权利要求 15 所述的第二 PE 设备, 其特征在于, 所述第一消息中还携带有下一跳网络地址, 所述第一消息中的下一跳网络地址为所述第一 PE 设备的网络地址,

所述处理模块, 还用于根据所述第一消息, 获取所述第一 PE 设备的网络地址;

所述处理模块, 进一步用于根据所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址以及所述第一 PE 设备的网络地址, 生成第二 MAC 转发表项, 所述第二 MAC 转发表项包括所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址, 所述第二 MAC 转发表项包含的下一跳网络地址为所述第一 PE 设备的网络地址;

其中, 当所述第二链路发生故障时, 所述第二 MAC 转发表项被所述第二 PE 用于转发携带有所述第一 VLAN 标识并且目的 MAC 地址为所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址的报文。

17. 根据权利要求 16 所述第二 PE 设备, 其特征在于, 所述第一消息为第一边界网关协议更新 BGP Update 消息, 所述第一 BGP Update 消息包括用于携带所述第一 VLAN 标识的 VLAN 属性字段。

18. 根据权利要求 15-17 任一项所述的第二 PE 设备, 其特征在于, 所述第二 PE 设备还包括发送模块, 其中, 在所述处理模块根据所述确定的第二接口、所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址以及所述第一 VLAN 标识, 生成第一 MAC 转发表项以后,

所述接收模块, 还用于接收携带有第二 VLAN 标识的第一报文, 所述第一报文的目的 MAC 地址为所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址;

所述处理模块, 还用于根据所述第一 MAC 转发表项将所述第二报文中携带的第二 VLAN 标识替换为所述第一 VLAN 标识, 得到所述携带有所述第一 VLAN 标识的第二报文; 所述发送模块, 用于向所述 CE 设备发送所述第二报文。

19. 根据权利要求 15-18 任一项所述的第二 PE 设备, 其特征在于, 在所述接收模块接收所述第一 PE 设备发送的第一消息之后,

所述处理模块, 还用于生成第二消息, 所述第二消息携带第二 MAC/IP

Advertisement Route, 下一跳网络地址和所述第一 VLAN 标识, 所述第二 MAC/IP Advertisement Route 包括 MAC 地址和所述以太网段标识 ESI, 所述第二 MAC/IP Advertisement Route 包括的 MAC 地址与所述第一 MAC/IP Advertisement Route 所包括的 MAC 地址相同, 所述第二消息携带的所述下一跳网络地址为所述第二 PE 设备的网络地址;

5 所述发送模块, 还用于向所述第一 PE 设备发送所述第二消息, 所述第二消息被所述第一 PE 设备用于生成第三 MAC 转发表项和第四 MAC 转发表项; 其中,

10 所述第三 MAC 转发表项被所述第一 PE 设备用于转发目的 MAC 地址为所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址的报文, 所述第三 MAC 转发表项包括所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址和所述第一 VLAN 标识, 所述第三 MAC 转发表项包含的出接口的标识为所述第一接口的标识;

15 当所述第一链路发生故障时, 所述第四 MAC 转发表项被所述第一 PE 设备用于转发目的 MAC 地址为所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址的报文, 所述第四 MAC 转发表项包括所述第二 MAC/IP Advertisement Route 所包括的 MAC 地址, 所述第四转发表项包含的下一跳网络地址为所述第二 PE 设备的网络地址。

20. 根据权利要求 19 所述第二 PE 设备, 其特征在于, 所述第二消息还携带有指示信息, 所述指示信息用于指示所述第一 PE 设备在收到所述第二消息后, 避免向所述第二 PE 设备发送到达所述第二 MAC/IP Advertisement Route 包括的 MAC 地址的 MAC/IP Advertisement Route 和所述第一 VLAN 标识。

20 21. 一种通信系统, 其特征在于, 包括权利要求 11-14 任一项所述的第一运营商边缘 PE 设备和权利要求 15-20 任一项所述的第二运营商边缘 PE 设备。

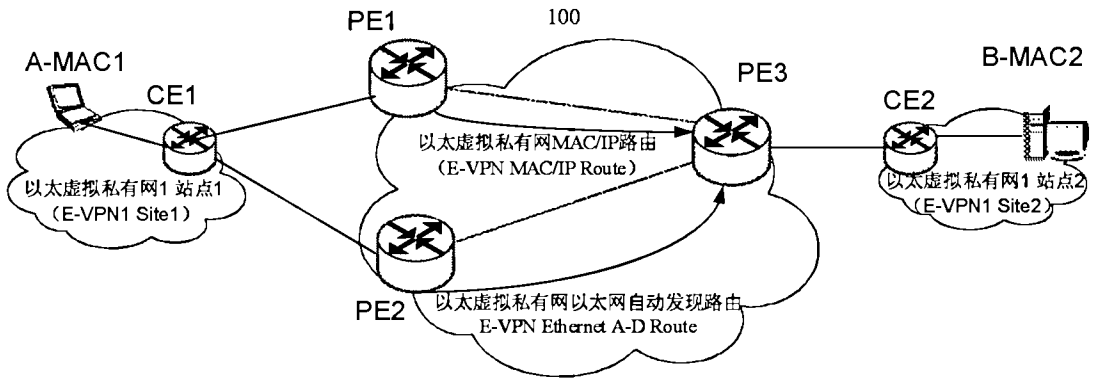


图 1

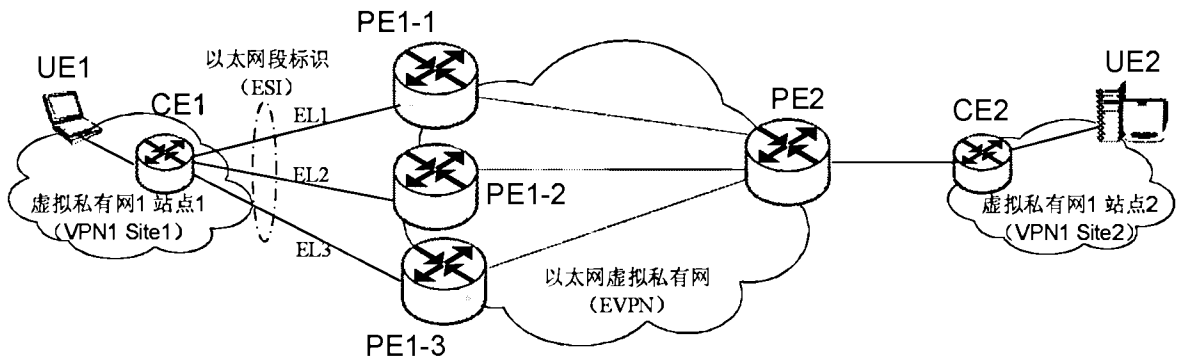


图 2

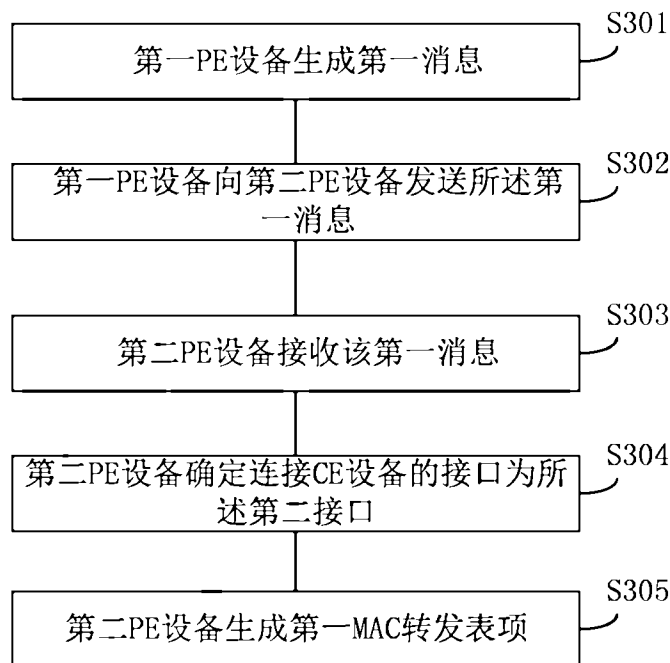


图 3

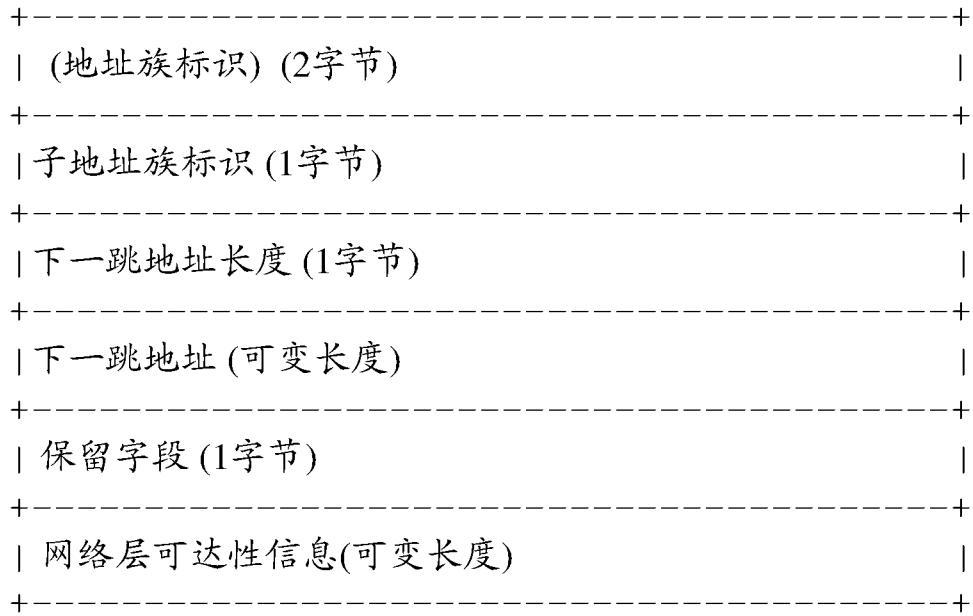


图 4a

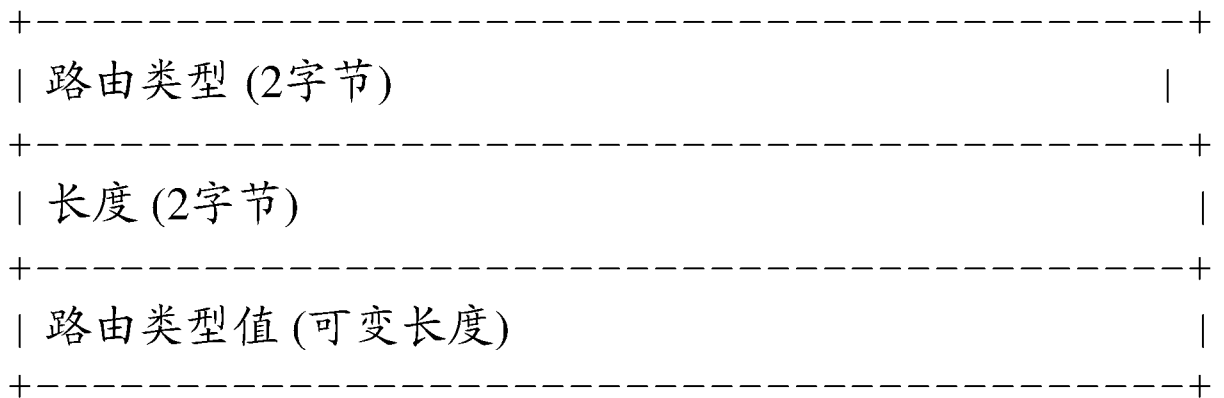


图 4b



图 4c

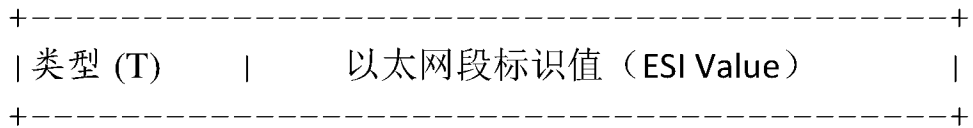


图 4d



图 5

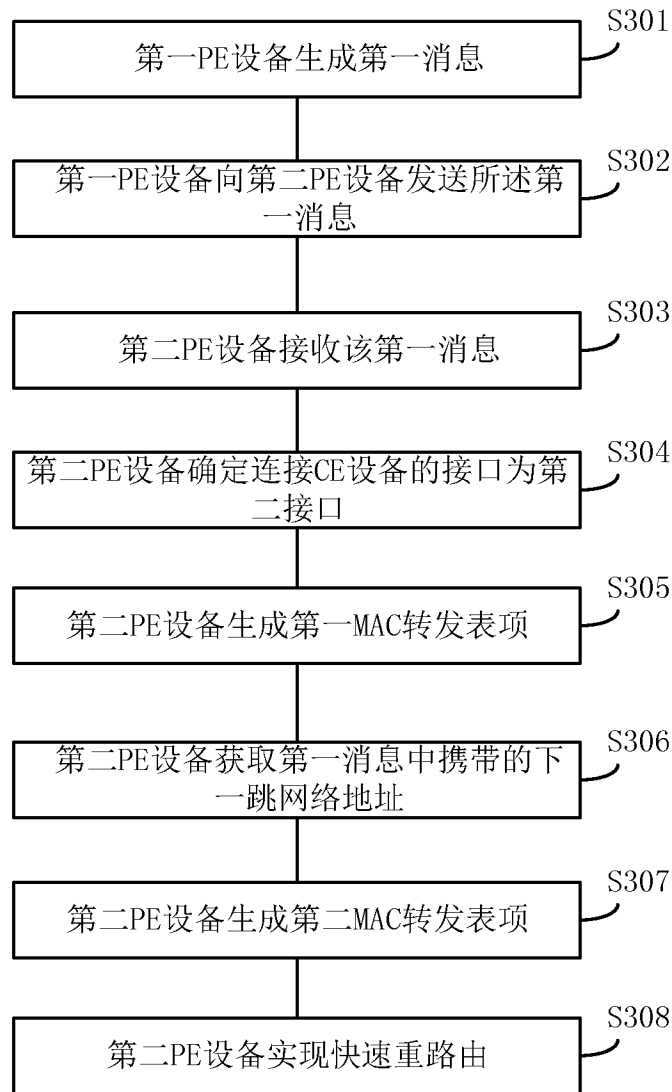


图 6

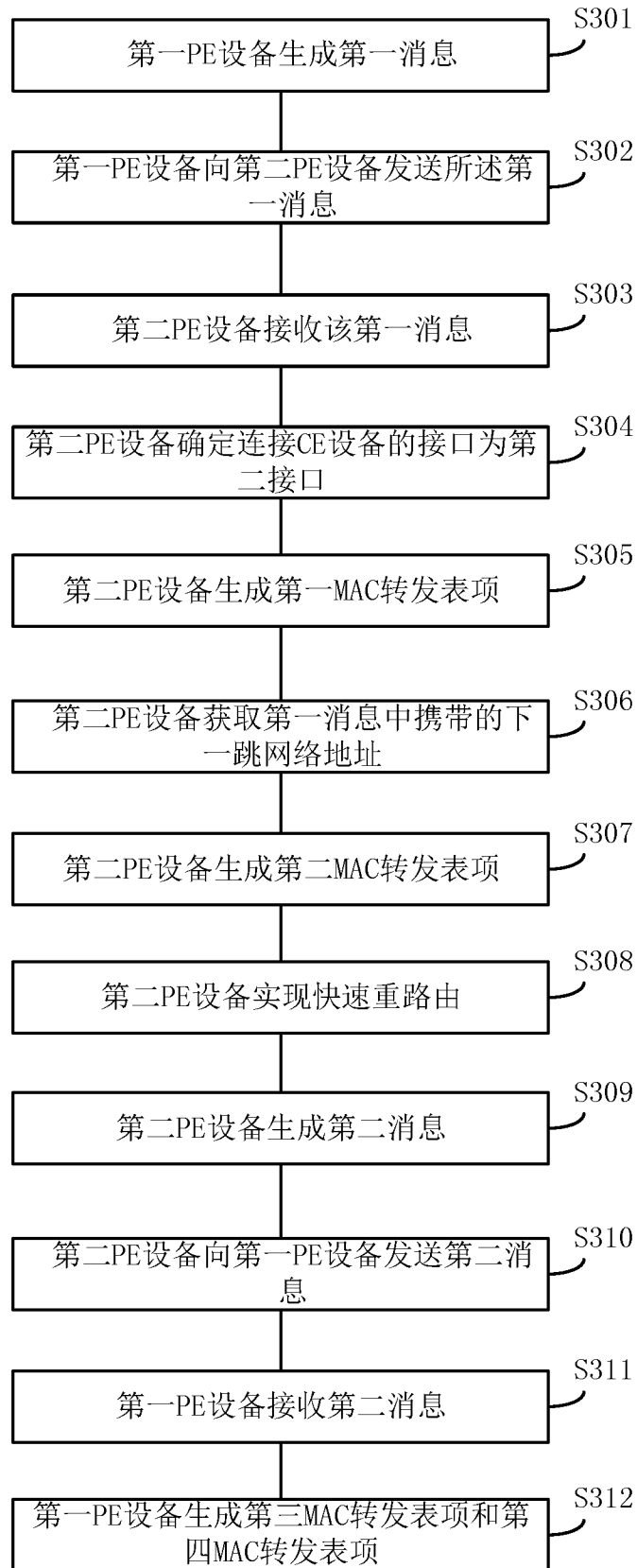


图 7

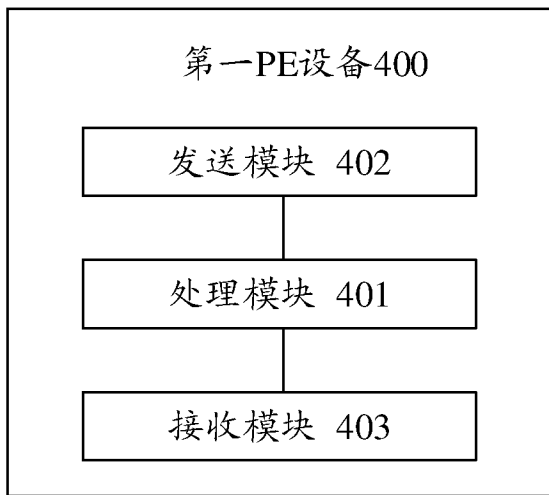


图 8

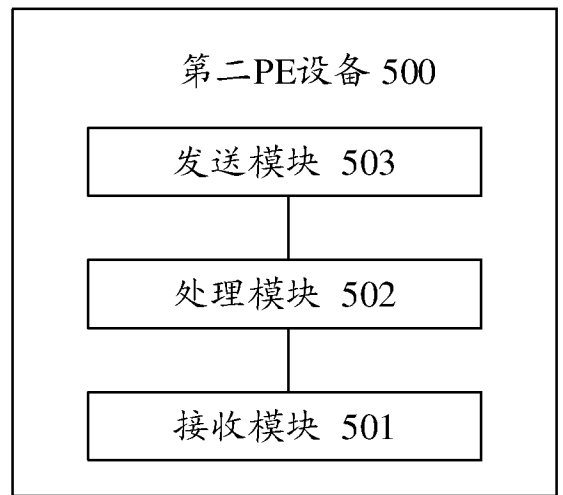


图 9

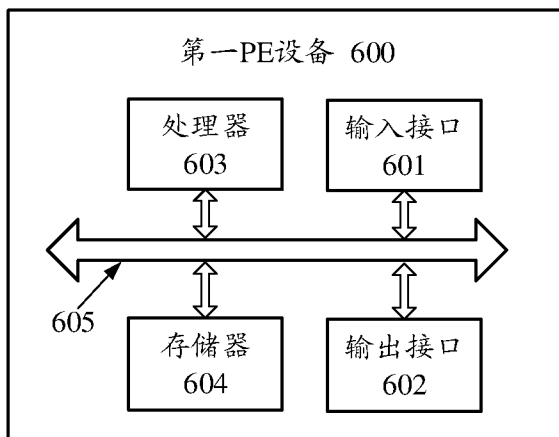


图 10

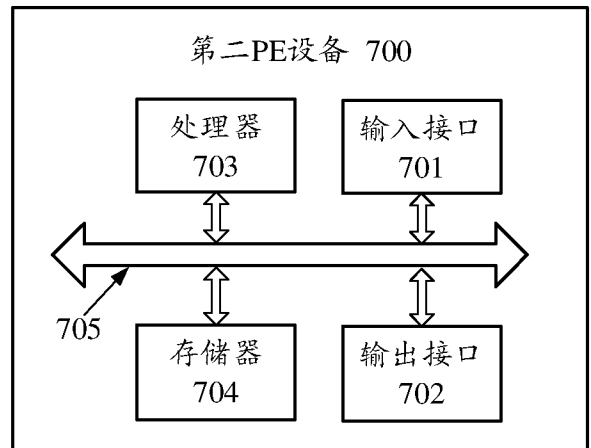


图 11

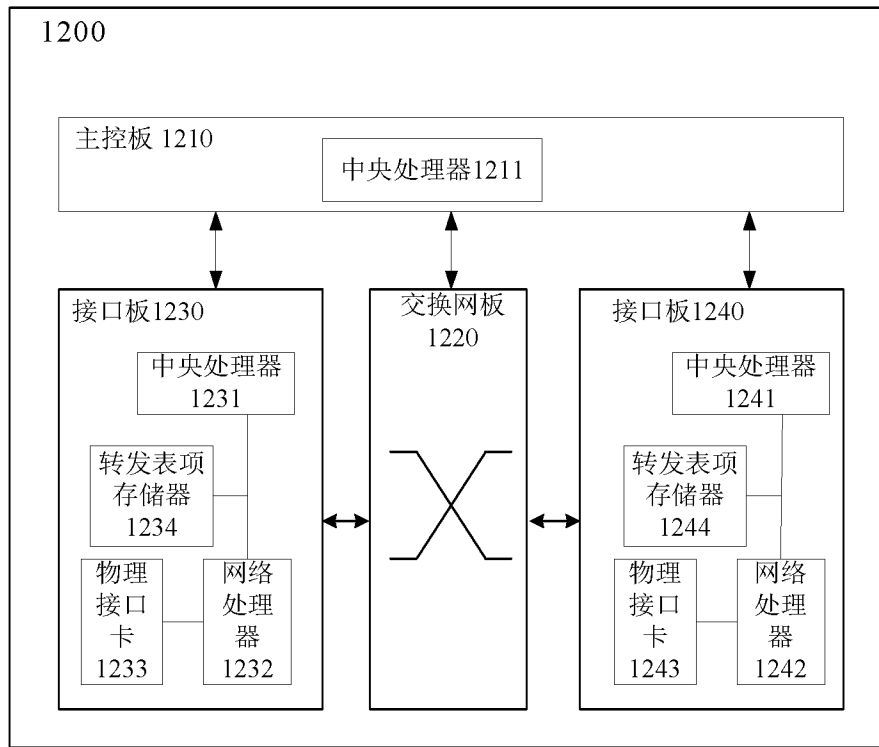


图 12

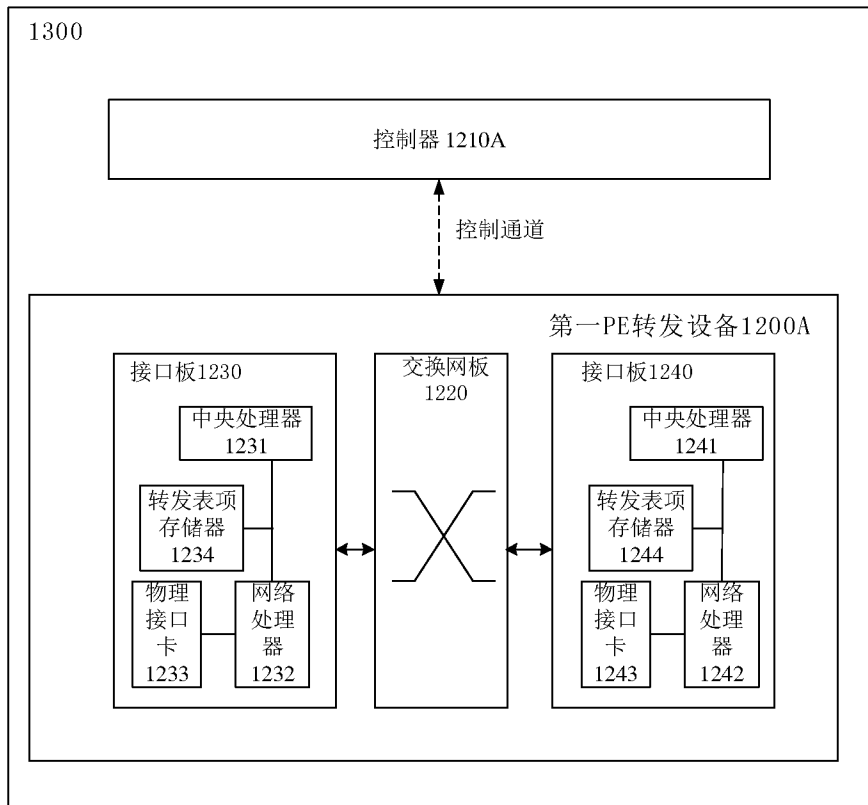


图 13

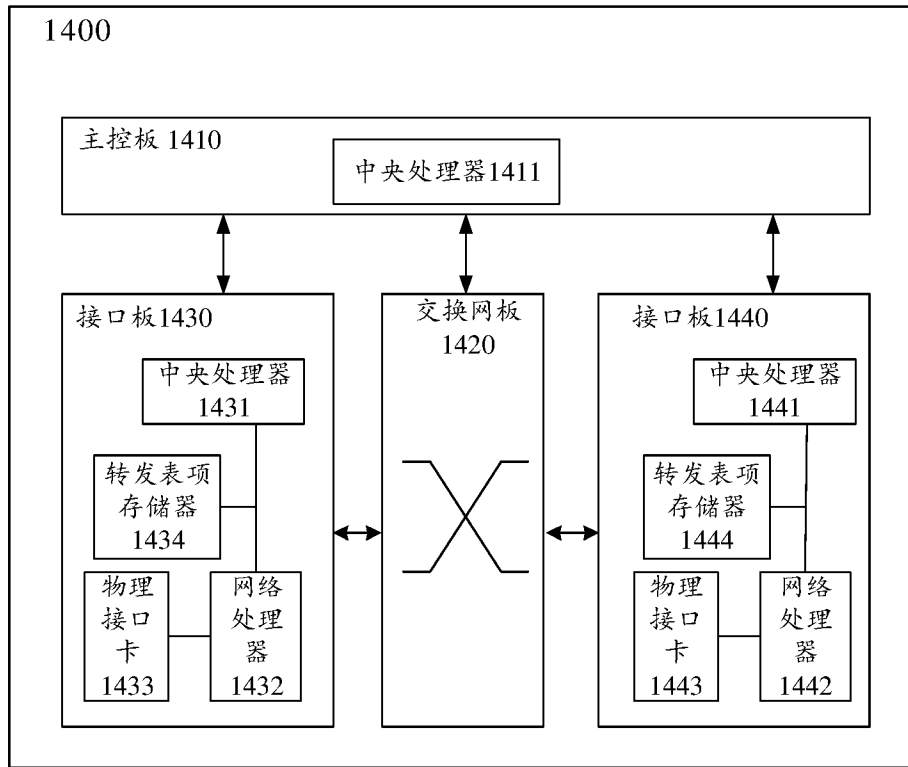


图 14

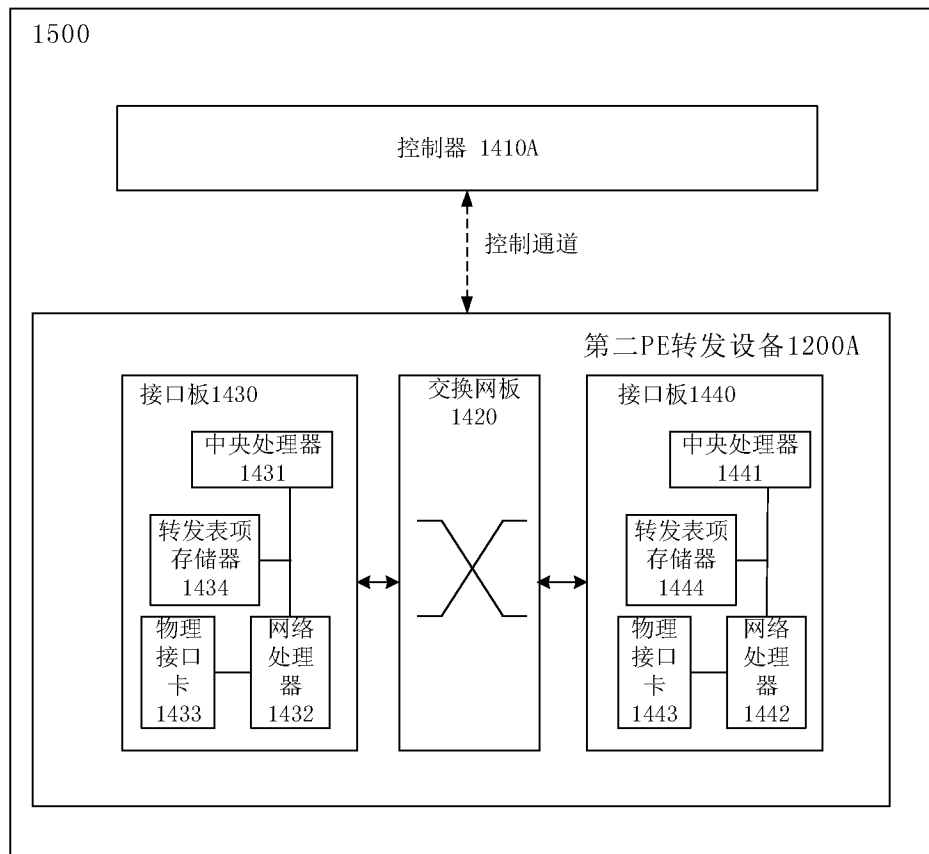


图 15

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/CN2017/115821

## A. CLASSIFICATION OF SUBJECT MATTER

H04L 12/721 (2013.01) 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)

WPI, EPODOC, CNPAT, CNKI: 以太网虚拟专用网, 以太网虚拟私有网, 用户边缘, 运营商边缘, 多归, 双归, 全活跃, 多活跃, 媒体接入控制/互联网协议通告路由, 负载, 分担, 均衡, 控制平面, 地址, 学习, 以太网段, 标识, 单播, 虚拟局域网, EVPN, EVN, PE, CE, MAC, advertisement route, all-active, active-active, multihome, multi-home, load, balanc+, aliasing, BGP, control, plane, learn+, ES, ESI, identifier, unicast, VLAN

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	kekekeke. “华为悦读汇技术发烧友: EVN-交换机-华为企业互动社区”, Support.huawei.com/huaweiconnect/enterprise/zh/thread-335619.html, 26 November 2015 (26.11.2015), chapters 1, 2.1, 2.2.2, 2.3, 2.4 and 2.4.2, and figures 2-11 and 2-18, non-official translation (kekekeke. “Huawei Forum-Technology Fan: EVN-Exchanger-Huawei Enterprise Support Community”)	1-21
A	DRAKE, J. et al., “BGP MPLS Based Ethernet VPN draft-ietf-12vpn-evpn-11”, Network Working Group Internet-Draft, 18 October 2014 (18.10.2014), entire document	1-21
A	CN 105743689 A (JUNIPER NETWORKS, INC.), 06 July 2016 (06.07.2016), entire document	1-21
A	CN 104365066 A (HUAWEI TECHNOLOGIES CO., LTD.), 18 February 2015 (18.02.2015), entire document	1-21
A	US 2013148657 A1 (CISCO TECHNOLOGY, INC.), 13 June 2013 (13.06.2013), entire document	1-21

Further documents are listed in the continuation of Box C.

See patent family annex.

<p>* Special categories of cited documents:</p> <p>“A” document defining the general state of the art which is not considered to be of particular relevance</p> <p>“E” earlier application or patent but published on or after the international filing date</p> <p>“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)</p> <p>“O” document referring to an oral disclosure, use, exhibition or other means</p> <p>“P” document published prior to the international filing date but later than the priority date claimed</p>	<p>“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</p> <p>“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</p> <p>“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</p> <p>“&amp;” document member of the same patent family</p>
---	---

Date of the actual completion of the international search  
06 February 2018

Date of mailing of the international search report  
26 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  
JIA, Yu  
Telephone No. (86-10) 61648539

**INTERNATIONAL SEARCH REPORT**  
Information on patent family members

International application No.  
PCT/CN2017/115821

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 105743689 A	06 July 2016	US 2016191374 A1	30 June 2016
		EP 3041178 A1	06 July 2016
CN 104365066 A	18 February 2015	WO 2014186978 A1	27 November 2014
		EP 2991284 A1	02 March 2016
US 2013148657 A1	13 June 2013	WO 2013086376 A1	13 June 2013
		EP 2789128 A1	15 October 2014

<p><b>A. 主题的分类</b></p> <p>H04L 12/721(2013.01)i</p> <p>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</p>																				
<p><b>B. 检索领域</b></p> <p>检索的最低限度文献(标明分类系统和分类号)</p> <p>H04L</p> <p>包含在检索领域中的除最低限度文献以外的检索文献</p> <p>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))</p> <p>WPI, EPODOC, CNPAT, CNKI: 以太网虚拟专用网, 以太网虚拟私有网, 用户边缘, 运营商边缘, 多归, 双归, 全活跃, 多活跃, 媒体接入控制/互联网协议通告路由, 负载, 分担, 均衡, 控制平面, 地址, 学习, 以太网段, 标识, 单播, 虚拟局域网, EVPN, EVN, PE, CE, MAC, advertisement route, all-active, active-active, multihome, multi-home, load, balanc+, aliasing, BGP, control, plane, learn+, ES, ESI, identifier, unicast, VLAN</p>																				
<p><b>C. 相关文件</b></p> <table border="1"> <thead> <tr> <th>类型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>A</td> <td>kekekeke. "华为悦读汇 技术发烧友: EVN-交换机-华为为企业互动社区" Support.huawei.com/huaweiconnect/enterprise/zh/thread-335619.html, 2015年 11月 26日 (2015 - 11 - 26), 第1、2.1、2.2.2、2.3、2.4、2.4.2节, 附图2-11, 2-18</td> <td>1-21</td> </tr> <tr> <td>A</td> <td>J.Drake 等. "BGP MPLS Based Ethernet VPN draft-ietf-l2vpn-evpn-11" Network Working Group INTERNET-DRAFT, 2014年 10月 18日 (2014 - 10 - 18), 全文</td> <td>1-21</td> </tr> <tr> <td>A</td> <td>CN 105743689 A (瞻博网络公司) 2016年 7月 6日 (2016 - 07 - 06) 全文</td> <td>1-21</td> </tr> <tr> <td>A</td> <td>CN 104365066 A (华为技术有限公司) 2015年 2月 18日 (2015 - 02 - 18) 全文</td> <td>1-21</td> </tr> <tr> <td>A</td> <td>US 2013148657 A1 (CISCO TECHNOLOGY, INC.) 2013年 6月 13日 (2013 - 06 - 13) 全文</td> <td>1-21</td> </tr> </tbody> </table>			类型*	引用文件, 必要时, 指明相关段落	相关的权利要求	A	kekekeke. "华为悦读汇 技术发烧友: EVN-交换机-华为为企业互动社区" Support.huawei.com/huaweiconnect/enterprise/zh/thread-335619.html, 2015年 11月 26日 (2015 - 11 - 26), 第1、2.1、2.2.2、2.3、2.4、2.4.2节, 附图2-11, 2-18	1-21	A	J.Drake 等. "BGP MPLS Based Ethernet VPN draft-ietf-l2vpn-evpn-11" Network Working Group INTERNET-DRAFT, 2014年 10月 18日 (2014 - 10 - 18), 全文	1-21	A	CN 105743689 A (瞻博网络公司) 2016年 7月 6日 (2016 - 07 - 06) 全文	1-21	A	CN 104365066 A (华为技术有限公司) 2015年 2月 18日 (2015 - 02 - 18) 全文	1-21	A	US 2013148657 A1 (CISCO TECHNOLOGY, INC.) 2013年 6月 13日 (2013 - 06 - 13) 全文	1-21
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求																		
A	kekekeke. "华为悦读汇 技术发烧友: EVN-交换机-华为为企业互动社区" Support.huawei.com/huaweiconnect/enterprise/zh/thread-335619.html, 2015年 11月 26日 (2015 - 11 - 26), 第1、2.1、2.2.2、2.3、2.4、2.4.2节, 附图2-11, 2-18	1-21																		
A	J.Drake 等. "BGP MPLS Based Ethernet VPN draft-ietf-l2vpn-evpn-11" Network Working Group INTERNET-DRAFT, 2014年 10月 18日 (2014 - 10 - 18), 全文	1-21																		
A	CN 105743689 A (瞻博网络公司) 2016年 7月 6日 (2016 - 07 - 06) 全文	1-21																		
A	CN 104365066 A (华为技术有限公司) 2015年 2月 18日 (2015 - 02 - 18) 全文	1-21																		
A	US 2013148657 A1 (CISCO TECHNOLOGY, INC.) 2013年 6月 13日 (2013 - 06 - 13) 全文	1-21																		
<p><input type="checkbox"/> 其余文件在C栏的续页中列出。 <input checked="" type="checkbox"/> 见同族专利附件。</p>																				
<p>* 引用文件的具体类型:</p> <p>"A" 认为不特别相关的表示了现有技术一般状态的文件</p> <p>"E" 在国际申请日的当天或之后公布的在先申请或专利</p> <p>"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)</p> <p>"O" 涉及口头公开、使用、展览或其他方式公开的文件</p> <p>"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</p> <p>"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</p> <p>"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</p> <p>"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</p> <p>"&amp;" 同族专利的文件</p>																				
<p>国际检索实际完成的日期</p> <p>2018年 2月 6日</p>		<p>国际检索报告邮寄日期</p> <p>2018年 2月 26日</p>																		
<p>ISA/CN的名称和邮寄地址</p> <p>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</p> <p>传真号 (86-10)62019451</p>		<p>受权官员</p> <p>加玉</p> <p>电话号码 (86-10)61648539</p>																		

国际检索报告  
关于同族专利的信息

国际申请号  
PCT/CN2017/115821

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	105743689	A	2016年 7月 6日	US	2016191374	A1	2016年 6月 30日
				EP	3041178	A1	2016年 7月 6日
CN	104365066	A	2015年 2月 18日	WO	2014186978	A1	2014年 11月 27日
				EP	2991284	A1	2016年 3月 2日
US	2013148657	A1	2013年 6月 13日	WO	2013086376	A1	2013年 6月 13日
				EP	2789128	A1	2014年 10月 15日

表 PCT/ISA/210 (同族专利附件) (2009年7月)