



(12)发明专利申请

(10)申请公布号 CN 110708229 A

(43)申请公布日 2020.01.17

(21)申请号 201810752158.2

(22)申请日 2018.07.10

(71)申请人 华为技术有限公司

地址 518129 广东省深圳市龙岗区坂田华为总部办公楼

(72)发明人 赵琨

(51) Int. Cl.

H04L 12/46(2006.01)

H04L 12/721(2013.01)

H04L 12/815(2013.01)

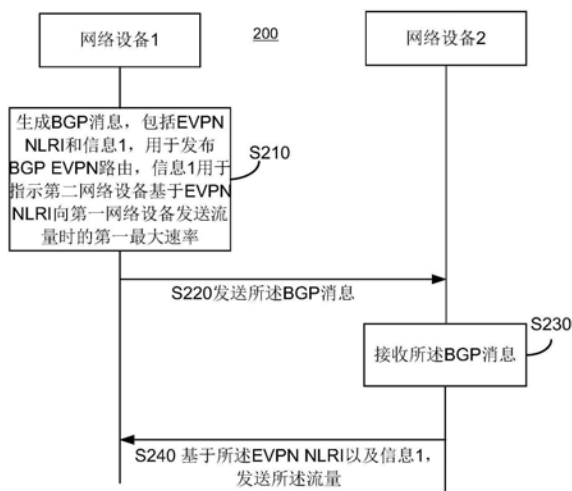
权利要求书3页 说明书16页 附图5页

(54)发明名称

一种收发消息的方法、装置和系统

(57)摘要

本申请请求保护一种收发信息的方法,该方法包括:第一网络设备向第二网络设备发送BGP消息,该BGP消息包括以太虚拟专用网网络层可达信息EVPN NLRI以及第一信息,所述第一信息指示所述第二网络设备基于所述EVPN NLRI向所述第一网络设备发送流量的第一最大速率。第二网络设备基于所述EVPN NLRI向第一网络设备发送流量时的第一类型的速率小于或者等于该第一最大速率。本申请的方法,能够节省公网带宽,减少网络设备开销。



1. 一种发送消息的方法,其特征在于,包括:

第一网络设备生成边界网关协议BGP消息,所述BGP消息包括以太虚拟专用网网络层可达信息EVPN NLRI以及第一信息,所述第一信息指示第二网络设备基于所述EVPN NLRI向所述第一网络设备发送流量的第一最大速率;

所述第一网络设备向所述第二网络设备发送所述BGP消息。

所述第一最大速率限制的速率是承诺信息速率CIR或峰值信息速率PIR。

2. 根据权利要求1所述的方法,其特征在于,所述BGP消息还包括第二信息,所述第二信息指示所述第二网络设备基于所述EVPN NLRI向所述第一网络设备发送所述流量的第二最大速率,其中,所述第一最大速率限制的速率的类型与所述第二最大速率限制的速率的类型不同。

所述第一最大速率限制的速率是承诺信息速率CIR,所述第二最大速率限制的速率是峰值信息速率PIR;或

所述第一最大速率限制的速率是峰值信息速率PIR,所述第二最大速率限制的速率是承诺信息速率CIR。

3. 根据权利要求1或2所述的方法,其特征在于,所述BGP消息还包括第三信息,所述第三信息指示所述第一最大速率限制的速率的类型。

4. 根据权利要求1-3任一项所述的方法,其特征在于,所述BGP消息还包括第四信息,所述第四信息指示所述流量的类型。

5. 根据权利要求1-4任一项所述的方法,其特征在于,所述第一信息携带于所述BGP消息的扩展团体属性中。

6. 一种接收消息的方法,其特征在于,所述方法包括:

第二网络设备接收来自第一网络设备的BGP消息,所述BGP消息包括以太虚拟专用网网络层可达信息EVPN NLRI以及第一信息,所述第一信息指示所述第二网络设备基于所述BGP EVPN路由向所述第一网络设备发送流量的第一最大速率;

所述第二网络设备基于所述EVPN NLRI以及所述第一信息,以第一类型的速率向所述第一网络设备发送所述流量,所述第一类型的速率小于或者等于所述第一最大速率。

7. 根据权利要求6所述的方法,其特征在于,所述BGP消息还包括第二信息,所述第二信息指示所述第二网络设备基于所述EVPN NLRI向所述第一网络设备发送所述流量的第二最大速率,其中,所述第一最大速率限制的速率的类型与所述第二最大速率限制的速率的类型不同;

所述第二网络设备基于所述EVPN NLRI以及所述第一信息,以所述第一类型的速率向所述第一网络设备发送所述流量包括:

所述第二网络设备基于所述EVPN NLRI,所述第一信息以及所述第二信息,以所述第一类型的速率和第二类型的速率向所述第一网络设备发送所述流量,所述第二类型的速率小于或者等于所述第二最大速率。

8. 根据权利要求6所述的方法,其特征在于,所述第一类型的速率是承诺信息速率CIR或峰值信息速率PIR。

9. 根据权利要求7所述的方法,其特征在于,

所述第一类型的速率是承诺信息速率CIR,所述第二类型的速率是峰值信息速率PIR;

或

所述第一类型的速率是峰值信息速率PIR,所述第二类型的速率是承诺信息速率CIR。

10. 根据权利要求6-9任一项所述的方法,其特征在于,所述BGP消息还包括第三信息,所述第三信息指示所述第一最大速率限制的速率的类型。

11. 根据权利要求6-10任一项所述的方法,其特征在于,所述BGP消息还包括第四信息,所述第四信息指示所述流量的类型。

12. 根据权利要求6-11任一项所述的方法,其特征在于,所述第一信息携带于所述BGP消息的扩展团体属性中。

13. 一种第一网络设备,其特征在于,包括:

存储器,该存储器包括计算机可读指令;

与该存储器相连的处理器,所述处理器用于执行所述计算机可读指令,从而执行以下操作:

生成边界网关协议BGP消息,所述BGP消息包括EVPN NLRI以及第一信息,所述第一信息指示第二网络设备基于所述EVPN NLRI向所述第一网络设备发送流量的第一最大速率;

向所述第二网络设备发送所述BGP消息。

所述第一最大速率限制的速率是承诺信息速率CIR或峰值信息速率PIR。

14. 根据权利要求13所述的第一网络设备,其特征在于,所述BGP消息还包括第二信息,所述第二信息指示所述第二网络设备基于所述EVPN NLRI向所述第一网络设备发送所述流量的第二最大速率,其中,所述第一最大速率限制的速率的类型与所述第二最大速率限制的速率的类型不同。

所述第一最大速率限制的速率是承诺信息速率CIR,所述第二最大速率限制的速率是峰值信息速率PIR;或

所述第一最大速率限制的速率是峰值信息速率PIR,所述第二最大速率限制的速率是承诺信息速率CIR。

15. 根据权利要求13或14所述的第一网络设备,其特征在于,所述BGP消息还包括第三信息,所述第三信息指示所述第一最大速率限制的速率的类型。

16. 根据权利要求13-15任一项所述的第一网络设备,其特征在于,所述BGP消息还包括第四信息,所述第四信息指示所述流量的类型。

17. 根据权利要求13-16任一项所述的第一网络设备,其特征在于,所述第一信息携带于所述BGP消息的扩展团体属性中。

18. 一种第二网络设备,其特征在于,包括:

存储器,该存储器包括计算机可读指令;

与该存储器相连的处理器,所述处理器用于执行所述计算机可读指令,从而执行以下操作:

接收来自第一网络设备的BGP消息,所述BGP消息包括以太虚拟专用网网络层可达信息EVPN NLRI以及第一信息,所述第一信息指示第二网络设备基于所述EVPN NLRI向所述第一网络设备发送流量的第一最大速率;以及

基于所述EVPN NLRI以及所述第一信息,以第一类型的速率向所述第一网络设备发送所述流量,所述第一类型的速率小于或者等于所述第一最大速率。

19. 根据权利要求18所述的第二网络设备,其特征在于,所述BGP消息还包括第二信息,所述第二信息指示所述第二网络设备基于所述EVPN NLRI向所述第一网络设备发送流量的第二最大速率,其中,所述第一最大速率限制的速率的类型与所述第二最大速率限制的速率的类型不同,所述处理器用于基于所述EVPN NLRI以及所述第一信息,以第一类型的速率向所述第一网络设备发送所述流量,包括:

所述处理器用于基于所述EVPN NLRI,所述第一信息以及所述第二信息,以所述第一类型的速率以及第二类型的速率向所述第一网络设备发送所述流量;其中,所述第二类型的速率小于或者等于所述第二最大速率。

20. 根据权利要求18所述的第二网络设备,其特征在于,所述第一类型的速率是承诺信息速率CIR或峰值信息速率PIR。

21. 根据权利要求19所述的第二网络设备,其特征在于,  
所述第一类型的速率是承诺信息速率CIR,所述第二类型的速率是峰值信息速率PIR;  
或

所述第一类型的速率是峰值信息速率PIR,所述第二类型的速率是承诺信息速率CIR。

22. 根据权利要求18-21任一项所述的第二网络设备,其特征在于,所述BGP消息还包括第三信息,所述第三信息指示所述第一最大速率限制的速率的类型。

23. 根据权利要求18-22任一项所述的第二网络设备,其特征在于,所述BGP消息还包括第四信息,所述第四信息指示所述流量的类型。

24. 根据权利要求18-23任一项所述的第二网络设备,其特征在于,所述第一信息携带于所述BGP消息的扩展团体属性中。

25. 一种计算机可读存储介质,所述计算机可读存储介质中存储有指令,当该指令在计算机上运行时,使得所述计算机执行权利要求1-5任一项所述的方法或者权利要求6-12任一项所述的方法。

26. 一种通信系统,其特征在于,所述通信系统包括权利要求13-17中任一项所述的第一网络设备以及权利要求18-24任一项所述的第二网络设备。

## 一种收发消息的方法、装置和系统

### 技术领域

[0001] 本申请涉及通信技术领域,特别涉及一种以太网虚拟私有网(英文:Ethernet Virtual Private Network,EVPN)中收发消息的方法、装置及系统。

### 背景技术

[0002] EVPN是一种在多协议标签交换(英文:Multi-Protocol Label Switching,MPLS)网络上提供二层网络互联的虚拟私有网络(英文:Virtual Private Network,VPN)的网络。目前,EVPN技术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。MAC地址学习和发布过程在控制平面实现,从而大幅减少了流量洪泛方式的MAC地址扩散,可以支持用户边缘设备(英文:Customer Edge,CE)多归属接入EVPN、便于实现负载分担。EVPN技术中,流量转发主要分为已知单播流量,以及广播,未知单播和组播(英文:broadcast,unknown unicast,and multicast,BUM)流量的转发。无论已知单播流量还是BUM流量,都要根据来自远端设备的路由对流量进行转发。在运营商网络中,链路带宽是有限的。如果不对流量进行限制,就可能出现拥塞,导致重要的流量被丢弃。因此需要对接口的流量进行限制。当前的解决方案是基于PE的本地策略对PE的流量进行限制,不能有效对EVPN的流量进行管理。例如,存在带宽浪费,设备处理压力大的问题。

### 发明内容

[0003] 本申请提供了一种发送和接收消息的方法,网络设备和系统,能够有效管理EVPN流量。有助于减少无效流量占用的公网带宽,减少网络设备的流量处理压力。

[0004] 第一方面,本申请提供了一种发送消息的方法,该方法包括:第一网络设备生成BGP消息,所述BGP消息包括EVPN网络层可达信息(英文:network layer reachability information,NLRI)以及第一信息,所述第一信息指示第二网络设备基于所述EVPN NLRI向所述第一网络设备发送流量的第一最大速率。所述第一网络设备向所述第二网络设备发送所述BGP消息。

[0005] 通过上述方法,第一网络设备在向其它网络设备通告BGP EVPN路由的同时,向其他网络设备动态发布第一网络设备对于接收流量的速率的限制。因此,本端设备能够控制远端设备发送的流量的速率,可以避免远端设备向本端设备发送超出本端设备允许的最大速率。当远端设备向本端设备发送的流量的速率超过本端设备发布的最大速率(即本端设备能够接收的流量的最大速率)时,本端设备可能没有能力对额外的流量进行处理,需要对额外的流量进行丢弃。本申请将上述额外的流量称为无效流量。因此,本申请提供的技术方案能够有效减少无效流量对公网带宽的占用。另外,远端设备生成并发送上述额外的流量会导致额外的开销,公网中设备传输上述额外的流量也会导致额外的开销。本申请提供的技术方案有助于降低远端设备以及公网设备的开销。无效流量可以指远端设备发送的,通

过公网传输到本端设备,但是由于超出本端速率限制而被丢弃的流量。

[0006] 第二方面,本申请提供了一种接收消息的方法,该方法包括:第二网络设备接收来自第一网络设备的BGP消息,所述BGP消息包括EVPN NLRI以及第一信息,所述第一信息指示所述第二网络设备基于所述EVPNNLRI向所述第一网络设备发送流量的第一最大速率;所述第二网络设备基于所述EVPN NLRI以及所述第一信息,以第一类型的速率向所述第一网络设备发送所述流量,所述第一类型的速率小于或者等于所述第一最大速率。

[0007] 通过上述方法,第二网络设备接收到第一网络设备发送的BGP消息后,根据BGP消息携带的信息生成BGP EVPN路由,并在基于该BGP EVPN路由以及第一信息的指示,以不大于第一最大速率的速率向第一网络设备发送流量。由此,第二网络设备向第一网络设备发送流量时,流量的速率符合第一网络设备对于接收的流量的速率限制。因此,减少了无效流量对公网带宽的占用。同时也降低了第二网络设备生成并发送无效流量导致的开销以及公网设备传输上述无效流量所导致的开销。

[0008] 在一种可能的设计中,所述BGP消息还包括第二信息,所述第二信息指示所述第二网络设备基于所述EVPN NLRI向所述第一网络设备发送所述流量的第二最大速率,其中,所述第一最大速率限制的速率的类型与所述第二最大速率限制的速率的类型不同。

[0009] 在一种可能的设计中,所述第二网络设备以所述第一类型的速率向所述第一网络设备发送所述流量包括:所述第二网络设备基于所述EVPN NLRI,所述第一类型的速率以及所述第二信息,以所述第一类型的速率以及第二类型的速率向所述第一网络设备发送所述流量,所述第二类型的速率小于或者等于所述第二最大速率。由此,在同一个BGP消息中,可以对不同速率类型的速率进行限制,例如,可以仅对第一类型的速率进行限制,也可以同时对第一类型的速率和第二类型的速率同时进行限制,增加了流量管理的灵活性。

[0010] 在一种可能的设计中,所述第一类型的速率是承诺信息速率(英文:committed information rate,CIR)或峰值信息速率(英文:peak information rate,PIR)。

[0011] 在一个可能的设计中,所述第二类型的速率是CIR或PIR。

[0012] 在一个可能的设计中,所述第一最大速率限制的速率是CIR或PIR。

[0013] 在一个可能的设计中,所述第二最大速率限制的速率是CIR或PIR。

[0014] 在一个可能的设计中,所述第一最大速率限制的速率的类型为所述第一类型,所述第二最大速率限制的速率的类型是所述第二类型。

[0015] 在一种可能的设计中,所述第一类型的速率是CIR,所述第一最大速率限制的速率是CIR,所述第二类型的速率是PIR,所述第一最大速率限制的速率是PIR。

[0016] 在一种可能的设计中,所述第一类型的速率是PIR,所述第一最大速率限制的速率是PIR,所述第二类型的速率是CIR,所述第二最大速率限制的速率是CIR。

[0017] 在一种可能的设计中,所述BGP消息还包括第三信息,所述第三信息指示所述第一最大速率限制的速率的类型。

[0018] 在一种可能的设计中,所述BGP消息还包括第四信息,所述第四信息指示所述流量的类型。由此,可以实现对不同类型的流量进行精确地流量管理。例如,可以分别对已知单播流量,广播流量,组播流量或者未知单播流量进行流量管理。

[0019] 在一种可能的设计中,所述第一信息,第二信息,第三信息和/或第四信息携带于所述BGP消息的扩展团体属性中。该扩展团体属性可以称之为流量抑制属性,是BGP协议新

定义的一个扩展团体属性,用于对EVPN流量进行管理。在一个BGP消息中,可以同时携带多个流量已知属性,实现对不同类型的速率分别进行限制,和/或对不同类型的流量的速率进行限制。

[0020] 第三方面,本申请的实施例提供一种第一网络设备。所述第一网络设备包括和与所述存储器相连的处理器。所述处理器用于执行所述存储器中的计算机可读指令从而执行以下操作:生成BGP消息,所述BGP消息包括EVPN NLRI以及第一信息,所述第一信息指示第二网络设备基于所述EVPN NLRI向所述第一网络设备发送流量的第一最大速率;向所述第二网络设备发送所述BGP消息。

[0021] 在一种可能的设计中,所述BGP消息还包括第二信息,所述第二信息指示所述第二网络设备基于所述EVPN NLRI向所述第一网络设备发送所述流量的第二最大速率,其中,所述第一最大速率限制的速率的类型与所述第二最大速率限制的速率的类型不同。

[0022] 由此,第一网络设备可以在一个BGP消息中,对不同速率类型的速率进行限制。例如,可以仅对第一类型的速率进行限制,也可以同时对第一类型的速率和第二类型的速率同时进行限制,增加了流量管理的灵活性。

[0023] 在一种可能的设计中,所述第一类型的速率是CIR或PIR。

[0024] 在一个可能的设计中,所述第二类型的速率是CIR或PIR。

[0025] 在一个可能的设计中,所述第一最大速率限制的速率是CIR或PIR。

[0026] 在一个可能的设计中,所述第二最大速率限制的速率是CIR或PIR。

[0027] 在一个可能的设计中,所述第一最大速率限制的速率的类型为所述第一类型,所述第二最大速率限制的速率的类型是所述第二类型。

[0028] 在一种可能的设计中,所述第一类型的速率是CIR,所述第一最大速率限制的速率是CIR,所述第二类型的速率是PIR,所述第一最大速率限制的速率是PIR。

[0029] 在一种可能的设计中,所述第一类型的速率是PIR,所述第一最大速率限制的速率是PIR,所述第二类型的速率是CIR,所述第二最大速率限制的速率是CIR。在一种可能的设计中,所述BGP消息还包括第三信息,所述第三信息指示所述第一最大速率限制的速率的类型。

[0030] 在一种可能的设计中,所述BGP消息还包括第四信息,所述第四信息指示所述流量的类型。由此,可以实现对不同类型的流量进行精确的流量管理。

[0031] 在一种可能的设计中,所述第一信息,第二信息,第三信息和/或第四信息携带于所述BGP消息的扩展团体属性中。该扩展团体属性可以称之为流量抑制属性,是BGP协议新定义的一个扩展团体属性,用于对EVPN流量进行管理。

[0032] 第四方面,本申请的实施例提供了一种第二网络设备。所述第二网络设备包括和与所述存储器相连的处理器。所述处理器用于执行所述存储器中的计算机可读指令从而执行以下操作:接收来自第一网络设备的BGP消息,所述BGP消息包括EVPN NLRI以及第一信息,所述第一信息指示所述第二网络设备基于所述EVPN NLRI向所述第一网络设备发送流量的第一最大速率;以及

[0033] 基于所述EVPN NLRI以及所述第一信息,以第一类型的速率向所述第一网络设备发送所述流量,所述第一类型的速率小于或者等于所述第一最大速率。

[0034] 在一种可能的设计中,所述BGP消息还包括第二信息,所述第二信息指示所述第二

网络设备基于所述EVPN NLRI向所述第一网络设备发送所述流量的第二最大速率,其中,所述第一最大速率限制的速率的类型与所述第二最大速率限制的速率的类型不同。基于所述EVPN NLRI以及所述第一信息,以第一类型的速率向所述第一网络设备发送所述流量包括:

[0035] 基于所述EVPN NLRI以及所述第二信息,以第二类型的速率向所述第一网络设备发送所述流量;

[0036] 其中,所述第二类型的速率小于或者等于所述第二最大速率。

[0037] 由此,在同一个BGP消息中,可以对不同速率类型的速率进行限制,例如,可以仅对第一类型的速率进行限制,也可以同时对第一类型的速率和第二类型的速率同时进行限制,增加了流量管理的灵活性。

[0038] 在一种可能的设计中,所述第一类型的速率是CIR或PIR。

[0039] 在一个可能的设计中,所述第二类型的速率是CIR或PIR。

[0040] 在一个可能的设计中,所述第一最大速率限制的速率是CIR或PIR。

[0041] 在一个可能的设计中,所述第二最大速率限制的速率是CIR或PIR。

[0042] 在一个可能的设计中,所述第一最大速率限制的速率的类型为所述第一类型,所述第二最大速率限制的速率的类型是所述第二类型。

[0043] 在一种可能的设计中,所述第一类型的速率是CIR,所述第一最大速率限制的速率是CIR,所述第二类型的速率是PIR,所述第一最大速率限制的速率是PIR。

[0044] 在一种可能的设计中,所述第一类型的速率是PIR,所述第一最大速率限制的速率是PIR,所述第二类型的速率是CIR,所述第二最大速率限制的速率是CIR。在一种可能的设计中,所述BGP消息还包括第三信息,所述第三信息指示所述第一最大速率限制的速率的类型。

[0045] 在一种可能的设计中,所述BGP消息还包括第四信息,所述第四信息指示所述流量的类型。由此,可以实现对不同类型的流量进行精确地流量管理。

[0046] 在一种可能的设计中,所述第一信息,第二信息,第三信息和/或第四信息携带于所述BGP消息的扩展团体属性中。该扩展团体属性可以称之为流量抑制属性,是BGP协议新定义的一个扩展团体属性,用于对EVPN流量进行管理。

[0047] 第五方面,本申请实施例提供了一种通信系统,包括第三方面或任一种可能的设计中提供的第一网络设备和第四方面或任意一种可能的设计中提供的第二网络设备。

[0048] 第六方面,本申请提实施例供了一种计算机可读存储介质,所述计算机可读存储介质中存储有指令,当其在计算机上运行时,使得计算机执行第一方面、第二方面、第一方面任意可能的实现方式或第二方面任意可能的实现方式中的方法。

[0049] 本申请实施例提供的方法,网络设备和系统,通过BGP消息向远端设备发布BGP EVPN路由时,携带与该BGP EVPN路由相关联的速率信息,使得网络设备基于该BGP EVPN路由向远端设备发送流量时,基于所述速率信息对EVPN流量进行管理,以不大于速率信息中所携带的最大速率向远端网络设备发送流量。因此,对于超出远端设备流量带宽的那部分流量不再发送到公网,减少了对公网带宽的占用,也减轻了设备处理流量的压力。

## 附图说明

[0050] 附图1为本申请实施例提供的EVPN网络架构示意图;



- [0051] 附图2为本申请实施例提供的一种收发消息的方法流程示意图；
- [0052] 附图3a为本申请实施例提供的MP\_REACH\_NLRI属性的格式示意图；
- [0053] 附图3b为本申请实施例提供的EVPN NLRI字段的格式示意图；
- [0054] 附图3c为本申请实施例提供的MAC/IP Advertisement Route格式示意图；
- [0055] 附图3d为本申请实施例提供的ESI字段的格式示意图；
- [0056] 附图4为本申请实施例提供的流量抑制属性的格式示意图；
- [0057] 附图5为本申请实施例提供的一个具体的流量抑制属性的功能示意图；
- [0058] 附图5a为本申请实施例提供的流量抑制属性用于MAC路由时的示意图；
- [0059] 附图6为本申请实施例提供的一种收发消息的方法流程示意图；
- [0060] 附图7为本申请实施例提供的第一网络设备的结构示意图；
- [0061] 附图8为本申请实施例提供的第二网络设备的结构示意图。

### 具体实施方式

[0062] 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。

[0063] 本申请中的“1”、“2”、“3”、“4”、“第一”、“第二”、“第三”以及“第四”等序数词用于对多个对象进行区分,不用于限定多个对象的顺序。

[0064] 本申请中提及的“A和/或B”,应该理解为包括以下情形:仅包括A,仅包括B,或者同时包括A和B。

[0065] 本申请中,方法步骤的执行主体称为“本端”,与执行主体通信的网元称为“远端”,这些都是本领域的技术人员可以理解其含义的。

[0066] 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。

[0067] 本申请中,“运营商网络”也可以被称之为“服务商提供的骨干网”,“服务提供商网络”或者“公网”,上述不同的术语可能经常交替使用。

[0068] 图1示出了本申请实施例应用的EVPN示意图。如图1所示,网络100包括服务商提供的骨干网和多个EVPN站点(英文:site)。网络100包括3个CE设备,3个PE设备,网络100还可以包括一个或者多个P设备(图中未示出)。多个P设备,例如P路由器,为骨干网中的骨干路由器,不与用户的CE设备直接相连。P设备具备基本的MPLS转发能力,维护到PE的路由,不需要了解任何VPN的路由信息。图1所示,3个CE设备分别是CE1,CE2和CE3。CE设备为用户网络边缘设备,有接口直接与PE设备相连。CE设备可以是路由器或交换机,也可以是一台主机。CE设备“感知”不到EVPN的存在,也无需支持多协议标签交换(Multiprotocol Label Switching,MPLS)。3个PE设备分别是PE1,PE2和PE3。PE设备为服务提供商网络的边缘设备,可以是一台路由器或三层交换机。PE设备与用户的CE设备直接相连,对EVPN的所有处理都发生在PE上。在移动承载网中,CE设备可以是基站(英文:Base Transceiver Station,缩写:BTS),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),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)。本端PE从CE学到MAC信息后,通过BGP与其它PE交换EVPN路由信息,本端PE维护与它直接相连的EVPN的路由信息以及远端PE发布过来的EVPN路由信息。所述多个EVPN站点包括site1-site3,site1,site2和site3属于同一个EVPN实例EVPN1。PE1与CE1直连;PE3与CE3直连;CE2分别通过以太网链路(英文:Ethernet Link,EL)1和EL2双归接入PE2和PE3。包含这两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与PE2是一对BGP对等体(英文:peer),PE1与PE3是一对BGP peer,PE2与PE3是一对BGP peer。在EVPN场景中,BGP peer也可以被称为EVPN 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,PE1与PE2是一对BGP peer可以被理解为指PE1是PE2的BGP peer,或者被理解为PE2是PE1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。

[0069] 应理解,图1中仅示例性的示出了3个PE设备,3个CE设备以及3个EVPN站点,该网络可以包括任意其它数量的PE设备、CE设备以及EVPN站点,可以包括任意数量的P设备,本申请实施例对此不做限定。

[0070] 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。

[0071] 下面结合图2对本申请实施例提供的一种发送和接收消息的方法200进行详细说明。应用方法200的网络架构至少包括网络设备1和网络设备2。网络设备1和网络设备2是一对BGP peer。网络设备1和网络设备1均可以是PE设备。举例来说,网络设备1可以是图1所示的PE2,网络设备2可以是图1所示的PE1。该网络架构可以是图1所示的网络架构。方法200包括以下操作。

[0072] S210、网络设备1生成BGP消息,所述BGP消息包括EVPN NLRI和信息1。

[0073] S220、网络设备1向网络设备2发送所述BGP消息。

[0074] 具体地,该BGP消息例如可以是BGP update消息。该EVPN NLRI为BGP协议定义的一个BGP NLRI。网络设备1向网络设备2发布BGP EVPN路由中包括该EVPN NLRI。BGP EVPN路由包括但不限于MAC/IP通告路由(英文:MAC/IP Advertisement Route),以太自动发现路由(英文:Ethernet Auto-Discovery (A-D) route),集成组播以太网标签路由(英文:Inclusive Multicast Ethernet Tag route,IMET route)以及以太网段路由(英文:Ethernet Segment route)。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3a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下一跳网络地址,该下一跳网络地址为该网络设备1的网络地址,例如,为网络设备1的环回(英文:loopback)地址。所述环回地址可以是在网络设备(例如路由器、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。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3b所示,该EVPN NLRI字段包括例如2个字节的路由类型(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。该Route Type用于指示不同的BGP EVPN路由类型。

[0075] 下面以MAC/IP Advertisement Route为例对EVPN NLRI字段进行介绍。本申请中,EVPN NLRI与EVPN NLRI字段是同一含义。MAC/IP Advertisement Route包括MAC Advertisement Route和IP Advertisement Route,分别用于指导二层报文转发和三层报文转发。

[0076] 对于MAC/IP Advertisement Route,Route type字段取值为2,Route Type specific字段用于承载该MAC/IP Advertisement Route的细节。如图3c所示,该MAC/IP Advertisement Route包括8个字节路由区分符(英文:Route Distinguisher,RD)字段、10个字节以太网段标识(英文:Ethernet Segment Identifier,ESI)字段、4个字节以太网标签标识(英文:Ethernet Tag ID)字段、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IP地址长度字段、0字节或4字节或16字节的IP地址字段、3个字节的MPLS标签1(英文:Label)字段以及0个字节或3个字节MPLS标签2字段,MPLS标签2用于指导3层流量转发,在MAC Advertisement Route中,MPLS标签2字段为0。图3c中所示的ESI字段的格式如图3d所示,包括类型(英文:Type,T)字段和ESI值(英文:Value,V)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。在该MAC Advertisement Route中包括的MAC地址为与网络设备1相连的CE设备的MAC地址或所述CE设备所管辖的用户设备的MAC地址。在本申请中,MAC Advertisement Route也被简称为MAC路由。

[0077] 其中,关于BGP update消息和MP\_REACH\_NLRI属性的定义可以参见RFC4760中的说明,所述EVPN NLRI字段的定义参见RFC7432中的说明。

[0078] 下面结合图1所示场景,举例介绍PE设备发布BGP EVPN路由的过程。

[0079] PE2学习到site2中的用户设备(英文:User Equipment,UE) 2的MAC地址,例如MAC1,PE2通过BGP Update消息向PE1发布MAC路由。PE3没有学习到site2中UE2的MAC地址。PE3向PE1发布以太网自动发现路由(英文:Ethernet Auto-discovery Route,Ethernet A-D route)。因此,PE1根据别名(英文:Aliasing)方式知道,PE1经由PE3可以到达该UE2。由此,UE1发往UE2的单播流量,经由PE1传输时,PE1可以对该单播流量进行负载分担处理。该单播流量经由PE2和PE3,被转发至CE2,从而实现EVPN1里的UE1和UE2的互通。

[0080] 除了上述的MAC路由以及Ethernet A-D路由,EVPN网络中,PE设备需要向同一个EVPN实例中的其它PE发送来自CE设备的BUM流量,在该场景下,接收了BUM流量的PE需要向其它的PE设备洪泛(英文:flood)未知单播流量。因此,在EVPN中,部署了PE设备后,每个PE设备需要向其它的PE设备通告IMET route,用来指导BUM报文的转发。接收该IMET路由的PE设备在IMET路由的指导下,向发送端PE转发BUM流量。

[0081] 在P2P EVPN业务中,例如PE1与PE2之间部署P2P EVPN,则PE2学习到site2中UE2的MAC地址后,通过BGP Update消息向PE1发布per-AD-EVI路由,per-AD-EVI路由属于Ethernet A-D路由的一种。在P2P场景下,所有的报文都是按照广播处理,per-AD-EVI路由用于在P2P场景中,指导报文转发。

[0082] 关于BGP EVPN路由,MAC通告路由,Ethernet A-D路由,IMET路由的具体细节,以及PE设备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。

[0083] 上述的信息1,也可以称之为速率信息1,用于指示当网络设备2基于网络设备1发送的BGP EVPN路由向网络设备1发送流量时的第一最大速率。上述第一最大速率所限制的速率的类型可以为CIR或PIR。举例来说,当第一最大速率限制的速率的类型为CIR时,网络设备2向网络设备1发送流量的CIR不能超过第一最大速率。

[0084] 作为一个具体的实施方式,本申请实施例提供了一种新的BGP扩展团体属性,可称之为流量抑制属性,用来携带信息1。本申请中的“流量抑制”指对流量的速率进行抑制。速率的类型可以是CIR或PIR。流量抑制属性的格式如图4所示,类型域和子类型域表示流量抑制属性所属类型,其中,类型域的取值为0x06,用于指示EVPN。子类型域表示该属性用于流量抑制,并且子类型域的取值不同,表示不同的功能,其取值根据不同的功能向标准组织进行申请。上述不同功能包括但不限于:子类型域的不同取值,代表流量抑制属性对不同类型的速率进行限制;或者子类型域的取值,表示该流量抑制属性通过其它字段携带的信息来表示对不同类型的速率进行限制,该其它字段例如可以是flags字段,也可以是保留字段。带宽(bandwidth)域用于承载速率信息,例如信息1带宽域的单位例如是kbps。如果流量抑制属性对于CIR进行限制,网络设备2向网络设备1发送流量时的CIR不能大于带宽域中所携带的值,类似的,如果流量抑制属性对于PIR进行限制,则网络设备2向网络设备1发送流量时的PIR不能大于带宽域中所携带的值。

[0085] 在一个具体的实施方式中,当需要对多个不同类型的速率进行限制时,可以根据需要在BGP update消息中携带多个流量抑制属性,即携带多个TLV。举例来说,在BGP update消息中,同时携带两个流量抑制属性。其中一个流量抑制属性用于对第一类型的速率(例如CIR)进行限制,另外一个流量抑制属性用于对第二类型的速率(例如PIR)进行限制。具体来说,所述BGP消息还可以包括信息2,信息2也可以称之为速率信息2,信息2指示网络设备2基于所述BGP EVPN路由向网络设备1发送流量的第二最大速率。信息2可以携带于流量抑制属性中。第二最大速率限制的速率的类型可以为CIR或PIR,CBS,PBS,或EBS中的一种。第二最大速率限制的速率的类型也可以是其它类型的速率。需要注意的是,第一最大速率限制的速率的类型与第二最大速率限制的速率的类型为不同类型。例如,第一最大速率限制的速率的类型为CIR时,第二最大速率限制的速率的类型为PIR。

[0086] 下面结合一个具体例子,来说明什么是第一类型的速率和第二类型的速率。第一

类型的速率例如可以是CIR,第二类型的速率例如是PIR。则第一类型的速率可以理解为CIR的值,速率类型是CIR,第二类型的速率可以理解为PIR的值,速率类型是PIR。本领域技术人员可以理解,CIR和PIR表示速率的类型的同时,也表示速率的值。当CIR的值是10兆比特/秒(Million bits per second,Mbps),则表示第一类型的速率为10Mbps,当PIR的值是20Mbps,则表示第二类型的速率为20Mbps。可以理解,对于同一个流量,可以同时遵循多个约束条件。多个约束条件可以包括第一最大速率和第二最大速率。例如,第一最大速率和第二最大速率分别是10Mbps和20Mbps。当流量1的CIR的值是10Mbps(例如等于第一最大速率)时,流量1的PIR的值可以是20Mbps(例如等于第二最大速率)。或者,当流量1的CIR的值是9Mbps(例如小于第一最大速率)时,流量1的PIR的值可以是19Mbps(例如小于第二最大速率)。

[0087] 在一个具体的实施方式中,子类型域携带信息3,该信息3也可以称之为速率类型信息,用于指示速率类型。例如,指示第一最大速率限制的速率的类型或第二最大速率限制的速率的类型。子类型域取值不同,可以代表对不同类型的速率进行限制。例如,当子类型域取值为0x07时,表示该流量抑制属性限制的速率类型是CIR,即该流量抑制属性对CIR的最大值进行限制,相应地,带宽域中携带的值就是CIR的最大值。再比如,当子类型域的取值为0x08时,可以表示该流量抑制属性限制的速率类型是PIR,即该流量抑制属性对PIR的值进行限制,相应地,带宽域中携带的值就是PIR的最大值。

[0088] 在另一个具体的实施方式中,为了对不同类型的速率进行流量抑制,可以通过定义一个新的子类型域,在该定义的子类型域下,使用flags域和/或保留字段中的不同比特位携带信息3,即通过flags域和/或保留字段中的不同比特位区分不同的速率类型。其中,每个比特位根据取值不同,可以标识两种不同的速率类型。举例来说,子类型域的取值为0x09,表示该流量抑制属性以flags域中比特位来区分不同的速率类型。例如,定义flags域中定义一个比特T。当Sub-type取值为0x07时,表示该子类型下的流量抑制属性采用flags域中的T来映射不同的速率类型,当T=0时,表示该流量抑制属性用于对CIR的值进行限制,当T=1时,表示该流量抑制属性对PIR的值进行限制。本领域技术人员可以理解的是,可以仅通过flags域的比特位或者仅通过保留字段中的比特位,以比特映射的方式来标识不同的速率类型,当然也可以通过组合flags域和保留字段中的比特位来限定不同的速率类型,这些都是在阅读本申请文件基础上,本领域技术人员能够得到的显而易见的方案。当需要对多个不同类型的速率同时进行限制时,可以BGP消息中同时携带多个流量抑制属性。具体的实现方式,本领域技术人员完全可以理解,此处不再赘述。

[0089] 在另一个具体的实施方式中,为了对不同类型的速率进行流量抑制,可以通过定义一个新的子类型域,在该子类型域下,通过使用flags域和/或保留字段中的不同比特位携带信息3,即通过flags域和/或保留字段中的不同比特位区分不同的速率类型。其中,每个比特位映射一种速率类型。举例来说,子类型域的取值为0x09,表示该流量抑制属性以flags域中比特位来区分不同的速率类型,其中flags域可以包括2个比特位,每个比特位对应一种速率类型,当每个比特位置1的时候,表示该流量抑制属性用于对该比特位映射的类型的速率进行速率抑制。

[0090] 在另一个具体的实施方式中,当BGP update消息发布MAC路由,或者在MP2MP场景中发布Ethernet A-D路由时,此时,MAC路由或者Ethernet A-D路由均用于指导单播报文转

发,因此,BGP update消息中携带的流量抑制属性可默认为对单播流量的速率进行限制,无需携带用于指示流量类型的信息。当BGP update消息用于发送IMET路由时,由于IMET路由用于指导BUM流量转发,因此,BGP update消息中携带的流量抑制属性可默认为对BUM流量进行限制,无需携带用于标识BUM流量类型的信息。此时,流量抑制属性不区分BUM流量的具体类型,对广播,组播和未知单播流量同时做出限制。即网络设备2向网络设备1转发BUM流量时,会根据接收到的IMET路由中携带的流量抑制属性中携带的信息1对所有的BUM流量进行流量控制。举例来说,当该流量抑制属性用于对PIR进行限制,并且信息1中携带的值是5000kbps时,则网络设备2向网络设备1发送广播,组播或者未知单播流量时,PIR的速率的最大值均不得大于5000kbps。上述方案实施简单,可以对不同类型的流量同时进行抑制,有助于节省报文开销。

[0091] 在另一个具体的实施方式中,在所述BGP消息中还携带信息4,该信息4也可以称之为流量类型信息,用于指示需要抑制的流量的类型。举例来说,在流量抑制属性的flags域中定义4个比特位,分别表示为B,M,U和K。该四个比特位分别以比特映射的方式,表示对不同的流量类型进行流量抑制。例如,当B=1时,表示对广播流量进行抑制。当M=1时,表示对组播流量的不同速率进行抑制,当U=1时,表示对未知单播流量的不同速率进行抑制,当K=1时,表示对单播流量的不同速率进行抑制。本领域技术人员可以理解,也可以采用比特位置0来标识对不同流量类型进行抑制。上述方案,可以精确的对不同类型的流量分别进行控制,提高了流量抑制的精确性。

[0092] 下面结合图5,举例说明一个具体的流量抑制属性的功能。如图5所示,类型域的取值0x06(对应二进制为00000110)表示EVPN,子类型域的取值0x07(对应二进制为00000111)表示流量抑制属性用于对CIR的值进行限制,flags域中的B=1,表示对广播流量进行抑制,带宽域的取值为10000kbps,表示网络设备1期望的接收广播流量的带宽值,也就是网络设备2向网络设备1发送广播流量时,CIR的值不能超过10000kbps。

[0093] S230、网络设备2接收网络设备1发送的所述BGP消息。

[0094] 所述BGP消息包括所述EVPN NLRI以及所述信息1,用于向第二网络设备发布BGP EVPN路由,所述第二网络设备是所述第一网络设备BGP对等体peer,所述第一信息指示所述第二网络设备基于所述EVPN NLRI向所述第一网络设备发送流量的所述第一最大速率。

[0095] S240、网络设备2基于所述EVPNNLRI以及所述信息1,以第一类型的速率向网络设备1发送所述流量,所述第一类型的速率小于或者等于所述第一最大速率。

[0096] 在一个具体的实施方式中,BGP消息中携带流量抑制属性,网络设备2获取BGP消息中携带的EVPN NLRI以及流量抑制属性,该流量抑制属性包括上述的信息1,还可以包括信息2和/或信息3。网络设备2的控制平面(例如控制板)根据所述EVPN NLRI以及流量抑制属性生成BGP EVPN路由表项,其中,根据EVPN NLRI所携带的不同类型的BGP EVPN路由,分别生成不同的BGP EVPN路由表项。例如,当所述BGP消息所发布的BGP EVPN路由是MAC路由时,网络设备2基于所述BGP消息生成MAC路由表项;当所述BGP消息所发布的BGP EVPN路由是Ethernet A-D路由时,网络设备2基于所述BGP消息生成Ethernet A-D路由表项;当所述BGP消息所发布的BGP EVPN路由是IMET路由时,网络设备2基于所述BGP消息生成IMET路由表项。在网络设备2中,不同类型的路由表项是单独维护,相互隔离的。下面以MAC路由为例,介绍本申请中生成MAC路由表以及MAC转发表的过程。对于IMET路由,以及Ethernet A-D路由,

网络设备生成相对应的路由表以及转发表的过程与MAC路由相类似,本申请中不再赘述。

[0097] 网络设备2收到网络设备1发送的BGP消息,并根据BGP消息所携带的MAC路由信息生成MAC路由表项。在一个具体的实现方式中,可以将流量抑制属性所携带的信息配置在MAC路由表项中(如表1所示)。所述MAC路由表项的目的MAC地址为所述MAC路由中所包括的MAC地址(例如MAC1),下一跳网络地址为所述网络设备1的环回地址,速率类型为流量抑制属性所限制的速率的类型(例如CIR),带宽值为每个流量抑制属性所携带的最大速率(例如10000kbps)。当BGP消息中携带了多个流量抑制属性信息时,MAC路由表项中可以保存多个不同速率类型以及相对应的带宽值。此处为了简便,仅以BGP消息中携带一个流量抑制属性为例进行介绍。网络设备2根据所述MAC路由表项生成MAC转发表项(如表2所示)。并将该MAC转发表项发送到网络设备2的转发平面(例如,转发板)。所述网络设备2根据所述MAC路由表项生成所述MAC转发表项时,可以根据所述MAC路由表项确定所述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MAC地址。所述MAC转发表项的出接口例如为Intf1。网络设备2将所述Intf1确定为所述MAC转发表项中的出接口可以包括:首先,网络设备2以该MAC路由表项中的网络设备1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),获得与网络设备1的loopback地址相对应的出接口为网络设备2到网络设备1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Intf1(即该网络设备2到该网络设备1的隧道在该网络设备2上的接口)。所述网络设备2将所述Intf1确定为所述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中并未示出,本领域技术人员可以理解。当BGP消息中携带多个流量抑制属性时,该多个不同的流量抑制属性所携带的速率类型以及对应的带宽值会分别保存在MAC路由表项和MAC转发表项中。例如,当BGP消息还携带用于对速率2进行抑制的流量抑制属性时,该流量抑制属性2中所携带的速率类型以及带宽值将分别保存在如下表1和表2所示的类型2和带宽2中。

[0098] 表1:MAC路由表项

[0099]

目的MAC	下一跳网络地址	类型1	带宽1	类型2	带宽2
MAC1	网络设备1的loopback地址	CIR	10000	...	...

[0100] 表2:MAC转发表项

[0101]

目的MAC	出接口	类型1	带宽1	类型2	带宽2
MAC1	Intf1	CIR	10000	...	...

[0102] 在上述的表1和表2中仅例举了MAC路由表项以及MAC转发表项中的部分信息,其它的信息为了简洁,并未示出,本领域技术人员可以理解。

[0103] 当网络设备2收到已知单播流量,例如,即报文中携带的目的MAC地址为上述MAC路由中所包括的MAC地址(MAC1),网络设备2查询MAC转发表,获知转发的出接口信息,并获知向MAC1转发流量时,远端设备所要求的CIR的速率不得超过10000kbps。此时,网络设备2确定本端流量抑制方案,例如,确定本端承诺接入速率(英文:committed access rate,CAR)限速中所规定的CIR的值,通过对两者进行比较,从中选取较小的值作为CIR的上限值,进行报文发送。由此,当网络设备2本端限速高于网络设备1能够接收的CIR的最大值时,网络设备2会根据获取的远端设备发送的流量需求来发送流量,有效避免发送超出带宽部分的无效流量,从而节省了对公网带宽资源的占用也有效减少了远端设备处理流量的压力。

[0104] 在本申请中,当网络设备1向网络设备2发布MAC路由时,可以基于MAC-VRF实例发布MAC路由,即从同一个MAC-VRF实例学到的所有MAC路由都会携带相同的流量抑制属性。或者,网络设备1向网络设备发布MAC路由时,可以基于不同的MAC地址,来分别进行流量抑制,由此,可以选择对发往某些MAC地址的流量进行流量抑制,或者对发往某些MAC地址的流量不做流量抑制。需要注意的是,如果流量抑制属性中带宽域中的值为0,网络设备2接收到发往相应MAC的流量时,做全部丢弃处理。

[0105] 下面结合具体的应用场景,来对流量抑制属性在不同类型的BGP EVPN路由中的应用进行举例介绍。应理解,下述的场景以及具体应用不应构成对本申请方案的限制。

[0106] 场景一、MP2MP场景中,流量抑制属性在MAC路由中的应用。

[0107] 在EVPN网络中,MAC路由用于指导单播报文转发,因此,流量抑制属性在MAC路由中用来抑制单播流量的带宽。流量抑制属性在MAC路由中的应用如图5a所示:

[0108] 结合图1,PE2向PE1发布MAC路由时,控制面信息建立过程如下:

[0109] PE2向PE1发送源MAC为MAC1的MAC路由,携带流量抑制属性。举例来说,流量抑制属性中flags域包括4个比特位,分别是B,M,U和K,分别表示限制了流量类型为广播流量,组播流量,未知单播流量和已知单播流量,在以下介绍的几个场景中,均以flags域中包括该4个比特位为例进行介绍,后续为了简化,不再赘述。在BGP消息发布MAC路由时,Sub-type=0x06,表示对CIR的速率进行限制,在以下介绍的几个场景中,均以Sub-type=0x06为例进行介绍,流量抑制属性中携带的带宽值均指代CIR的最大值,后续为了简化,不再赘述。K置1,其它比特位置0,可忽略检查,表示需要对发往该MAC的单播流量的CIR进行抑制,该流量抑制属性中所携带的最大允许带宽Bandwidth=10000kbps,即PE2能够接收的CIR的最大值不超过10000kbps,因此,PE1基于该MAC路由向PE2发送流量时的CIR要小于或者等于10000kbps。

[0110] PE1向PE2转发流量时,转发面信息建立过程简述如下:

[0111] PE1发送目的MAC为MAC1的单播报文时,查找MAC路由找到该MAC,其中单播报文的CIR的最大值为10000kbps,与本地接口带宽进行比较,如果该速率小于本端的接口带宽,则超出该值的流量丢弃,小于该值的流量正常转发。

[0112] 场景二、流量抑制属性在Per-AD-EVI路由的应用。

[0113] 在MP2MP业务中,发布Per-AD-EVI路由时所携带流量抑制属性用来对发往该EVPN实例的所有单播流量的总量做控制。其中,与场景一种MAC路由中应用相似,流量抑制属性flags域中K bit设置为1,其余位置0并忽略检查,远端EVPN发送的总的单播流超出bandwidth部分将被丢弃。



[0114] 在P2P EVPN业务中,per-AD-EVI路由用来指导报文的转发,因为所有报文都是按广播处理,因此,在P2P场景下,BGP消息发布per-AD-EVI路由时携带流量抑制属性,用于对广播报文进行流量抑制。举例来说:流量抑制属性中flags域中,B=1,其余位忽略检查,超出bandwidth限制的流量将被丢弃。如Bandwidth值为0,则表示所有发送给该MP2MP EVPN的单播流量或发送给该P2P EVPN的所有流量将被丢弃。

[0115] 场景三、流量抑制属性在IMET路由中的应用

[0116] 在EVPN网络中,IMET路由用来指导BUM报文的转发。在发布IMET路由时,流量抑制属性中flags域的B,U,M可以设置为1,K设置为0且忽略检查,表示对相应BUM流量做抑制。B=1时,表示需要对广播流量做抑制,超出bandwidth部分丢弃。M=1时,表示需要对组播流量做抑制,超出bandwidth部分丢弃。U=1时,表示需要对未知单播流量做抑制,超出bandwidth部分丢弃。B/M/U bit位可以单独设置,也可以一起存在,当有多个bit位一起设置为1时,表示对相应bit位代表的流量都根据bandwidth做抑制。bandwidth为0则表示不需要这种流量,远端要全部丢弃。当需要对不同BUM流量分别做抑制时,即使用不同的bandwidth,路由中需要带多个流量抑制属性,即携带多个TLV分别抑制不同类型的流量。对于多个接口接入PE的情况,不同的接口可能有不同的带宽抑制需求,此时IMET路由携带流量抑制属性中的bandwidth可以设置为接口中最大的带宽,远端发送过来BUM报文可以通过根据本地流量策略进行抑制。

[0117] 本申请中,将现有的静态部署流量抑制的方案更改为动态方案,即通过本端设备向远端设备通告EVPN路由时,将本端流量的需求动态发布给远端设备,简化了维护操作。具体来说,现有技术中,主要基于本端PE设备对本端PE向远端PE发送的流量或者接收到的来自远端PE的流量进行带宽控制,从而达到流量抑制的目的。本端PE设备仅能对本端发送或者接收的流量进行管理,无法控制远端PE设备流量的发送。远端PE发送的流量即使超出了本端PE的带宽限制,也依然会通过公网传输到本端PE,超出带宽限制的流量在到达本端PE以后,会被本端PE丢弃。被本端PE丢弃的流量在运营商网络传播过程中,占用了运营商网络的带宽。同时,流量传输和处理过程中,对端设备发送无效流量需要额外开销,运营商设备和本端PE设备的处理无效流量也需要开销。通过本申请提供的方案,只需要维护本端设备,就可以精确的控制远端设备发送的已知单播流量或者BUM流量的带宽。超出本端设备带宽限制的无效流量不会再通过公网传输到本端设备,减少了无效流量对公网带宽的占用,同时也减轻了本端设备,对端设备以及运营商设备处理大量无效流量的开销。

[0118] 图6为本申请实施例所提供的发送和接收消息的方法,应用方法600的网络架构至少包括第一网络设备和第二网络设备。第一网络设备和第二网络设备是一对BGP peer。第一网络设备和第二网络设备均可以是PE设备,举例来说,第一网络设备可以是图1所示的PE2,第二网络设备可以是图1所示的PE1。该网络架构可以是图1所示的网络架构。另外,图6所示的方法可以具体实现图2所示的方法。例如,图6中的第一网络设备和第二网络设备可以分别是图2中的网络设备1和网络设备2。图6所示的方法包括以下内容。

[0119] S610、第一网络设备生成边界网关协议BGP消息,所述BGP消息包括EVPN NLRI以及第一信息,,所述第一信息指示了当所述第二网络设备基于所述EVPN NLRI向所述第一网络设备发送流量时的第一最大速率。

[0120] 该BGP消息用于向第二网络设备发布BGP EVPN路由,所述第二网络设备是所述第

一网络设备BGP对等体peer。

[0121] S620、所述第一网络设备向所述第二网络设备发送所述BGP消息。

[0122] 关于所述第一网络设备如何向所述第二网络设备发布BGP EVPN路由,可参见方法200中的相关说明,此处不再赘述。

[0123] S630、第二网络设备接收所述BGP消息。

[0124] S640、所述第二网络设备基于所述EVPN NLRI以及所述第一信息,以第一类型的速率向所述第一网络设备发送所述流量,所述第一类型的速率小于或者等于所述第一最大速率。

[0125] 一个具体的实施方式中,所述第一最大速率限制的速率可以为CIR或PIR。所述的第一类型的速率与第一最大速率限制的速率相同。例如,第一最大速率限制的速率是CIR,第一类型的速率也是CIR;第一最大速率限制的速率是PIR,第一类型的速率也是PIR

[0126] 一个具体的实施方式中,所述BGP消息还可以包括第二信息,所述第二信息用于指示所述第二网络设备基于所述BGP EVPN NLRI路由向所述第一网络设备发送所述流量时的第二最大速率,其中,所述第一最大速率限制的速率的类型与所述第二最大速率限制的速率的类型不同,所述方法600还可以包括:

[0127] 所述第二网络设备基于所述EVPN NLRI以及所述第二信息,以第二类型的速率向所述第一网络设备发送所述流量,所述第二类型的速率小于或者等于所述第二最大速率。

[0128] 第二最大速率限制的速率的类型可以为CIR或PIR,所述第二类型的速率与所述第二最大速率限制的速率相同。例如,第二最大速率限制的速率是CIR,第二类型的速率也是CIR;第二最大速率限制的速率的是PIR,第二类型的速率也是PIR。

[0129] 一个具体的实施方式中,所述BGP消息还可以包括第三信息,所述第三信息用于指示所述第一最大速率限制的速率的类型。

[0130] 一个具体的实施方式中,所述BGP消息还可以包括第四信息,所述第四信息用于指示所述流量的类型。

[0131] 一个具体的实施方式中,所述第一信息,第二信息,第三信息和/或第四信息均携带于所述BGP消息的扩展团体属性中,该扩展团体属性例如可以定义为流量抑制属性,该流量抑制属性的相关说明参见方法200,此处不再赘述。

[0132] 以上,结合图2-图6说明了根据本申请实施例提供的发送和接收消息的方法。以下,结合图7和图8对与上述方法实施例对应的网络设备进行描述。

[0133] 图7是本申请实施例提供的一种网络设备700的示意图。该网络设备700可以应用于图1所示的网络架构中,例如可以是图1所示的网络架构中的PE2。用于执行方法200中网络设备1或者方法600中第一网络设备所执行的操作。如图7所示,网络设备700可以包括处理器710,与所述处理器710耦合连接的存储器720,收发器730。处理器710可以是中央处理器(英文:central processing unit,缩写:CPU),网络处理器(英文:network processor,缩写:NP)或者CPU和NP的组合。处理器还可以进一步包括硬件芯片。上述硬件芯片可以是专用集成电路(英文:application-specific integrated circuit,缩写:ASIC),可编程逻辑器件(英文:programmable logic device,缩写:PLD)或其组合。上述PLD可以是复杂可编程逻辑器件(英文:complex programmable logic device,缩写:CPLD),现场可编程逻辑门阵列(英文:field-programmable gate array,缩写:FPGA),通用阵列逻辑(英文:generic

array logic,缩写:GAL)或其任意组合。处理器710可以是指一个处理器,也可以包括多个处理器。存储器720可以包括易失性存储器(英文:volatile memory),例如随机存取存储器(英文:random-access memory,缩写:RAM);存储器也可以包括非易失性存储器(英文:non-volatile memory),例如只读存储器(英文:read-only memory,缩写:ROM),快闪存储器(英文:flash memory),硬盘(英文:hard disk drive,缩写:HDD)或固态硬盘(英文:solid-state drive,缩写:SSD);存储器还可以包括上述种类的存储器的组合。存储器720可以是指一个存储器,也可以包括多个存储器。在一个实施方式中,存储器720可以包括多个软件模块,例如发送模块721,处理模块722和接收模块723。处理器710执行各个软件模块后可以按照各个软件模块的指示进行相应的操作。在本实施例中,一个软件模块所执行的操作实际上是指处理器710根据所述软件模块的指示而执行的操作。处理模块722可以用于生成边界网关协议BGP消息,所述BGP消息包括EVPN NLRI以及第一信息,所述第一信息指示了所述第二网络设备基于所述EVPN NLRI向所述第一网络设备发送流量的第一最大速率;发送模块721可以用于向网络设备2发送所述BGP消息。此外,处理器710执行存储器720中的计算机可读指令后,可以按照所述计算机可读指令的指示,执行网络设备1或者第一网络设备可以执行的全部操作,例如,网络设备1在与图2-图5对应的实施例中执行的操作,或者第一网络设备在图6对应的实施例中执行的操作。

[0134] 图8所示为本申请提供的另一网络设备800的示意图。该网络设备800可以应用于图1所示的网络架构中,例如可以是图1所示的网络架构中的PE1。用于执行方法200中网络设备2或者方法600中第二网络设备所执行的操作。如图8所示,网络设备800可以包括处理器810,与所述处理器810耦合连接的存储器820,收发器830。处理器810可以是CPU,NP或者CPU和NP的组合。处理器还可以进一步包括硬件芯片。上述硬件芯片可以是ASIC,PLD或其组合。上述PLD可以是CPLD,FPGA,GAL或其任意组合。处理器810可以是指一个处理器,也可以包括多个处理器。存储器820可以包括易失性存储器(英文:volatile memory),例如RAM;存储器也可以包括非易失性存储器(英文:non-volatile memory),例如ROM,快闪存储器(英文:flash memory),HDD或SSD;存储器还可以包括上述种类的存储器的组合。存储器820可以是指一个存储器,也可以包括多个存储器。在一个实施方式中,存储器820可以包括多个软件模块,例如发送模块821,处理模块822和接收模块823。处理器810执行各个软件模块后可以按照各个软件模块的指示进行相应的操作。在本实施例中,一个软件模块所执行的操作实际上是指处理器810根据所述软件模块的指示而执行的操作。接收模块823用于接收来自第一网络设备的BGP消息,所述BGP消息包括以太虚拟专用网网络层可达信息EVPN NLRI以及第一信息,所述第一信息指示所述第二网络设备基于所述EVPN NLRI向所述第一网络设备发送流量的第一最大速率。处理模块822用于根据所述EVPN NLRI以及所述第一信息生成所述流量。发送模块821用于以第一类型的速率向所述第一网络设备发送处理模块822生成的所述流量。所述第一类型的速率小于或者等于所述第一最大速率。此外,处理器810执行存储器820中的计算机可读指令后,可以按照所述计算机可读指令的指示,执行网络设备2或者第二网络设备可以执行的全部操作。例如,网络设备2在与图2-图5对应的实施例中执行的操作,或者第二网络设备在图6对应的实施例中执行的操作。

[0135] 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,而不对本申请实施例的实施过程

构成任何限定。

[0136] 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操作,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。

[0137] 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。

[0138] 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。

[0139] 本说明书的各个部分均采用递进的方式进行描述,各个实施方式之间相同相似的部分互相参见即可,每个实施方式重点介绍的都是与其他实施方式不同之处。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。

[0140] 本申请说明书的上述描述可以使得本领域技术任何可以利用或实现本申请的内容,任何基于所公开内容的修改都应该被认为是本领域显而易见的,本申请所描述的基本原则可以应用到其它变形中而不偏离本申请的发明本质和范围。因此,本申请所公开的内容不仅仅局限于所描述的实施例和设计,还可以扩展到与本申请原则和所公开的新特征一致的最大范围。

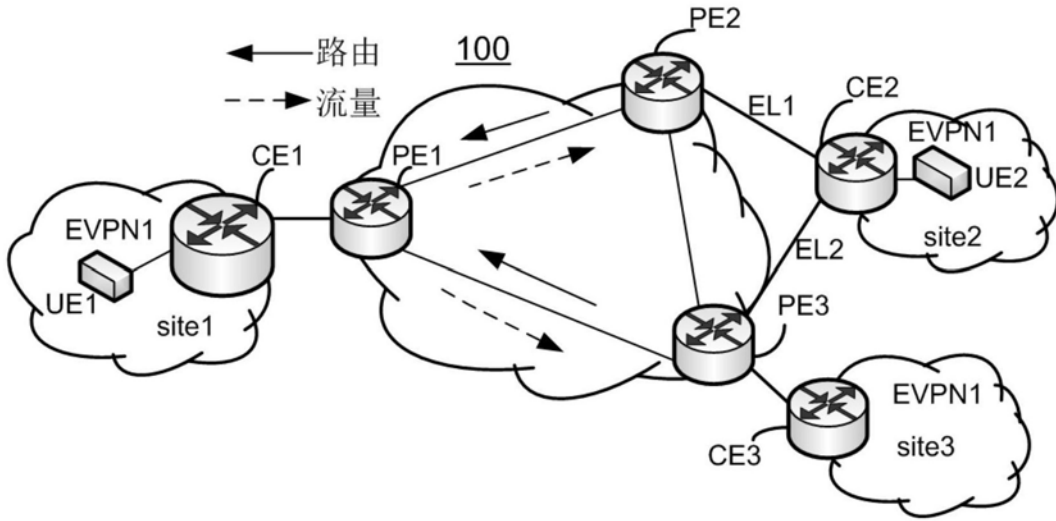


图1

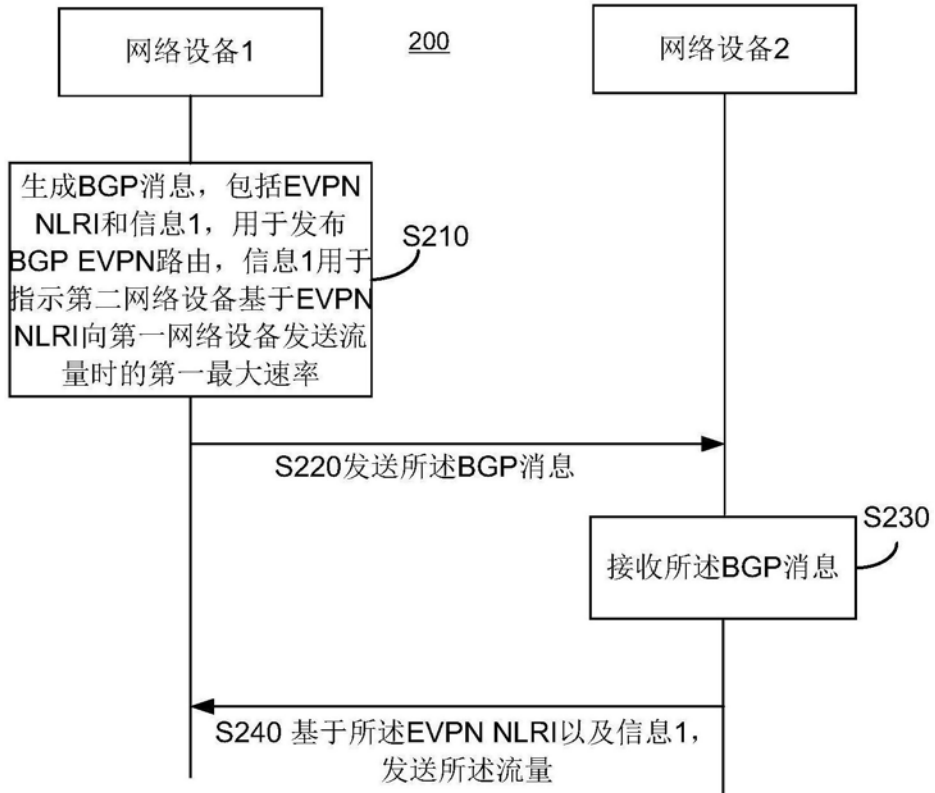


图2



图3a

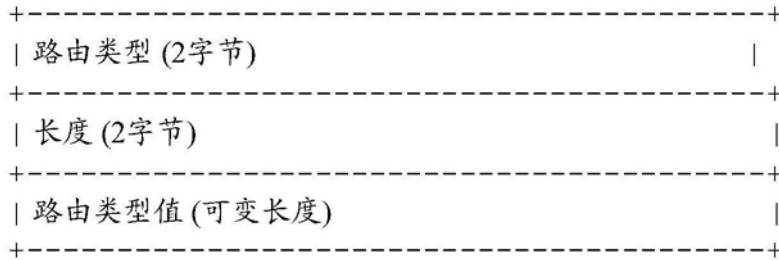


图3b



图3c

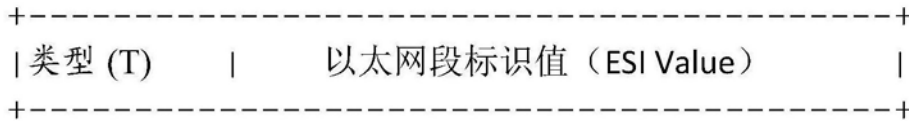


图3d



图4



图5

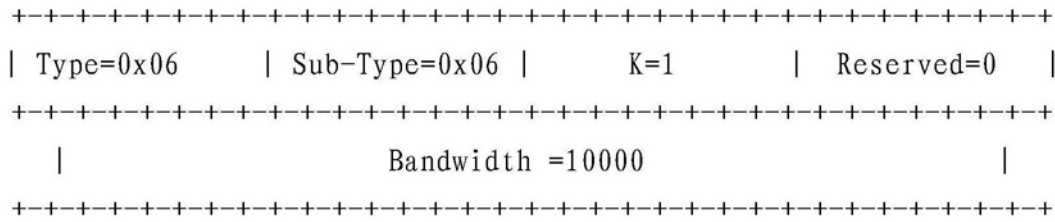


图5a

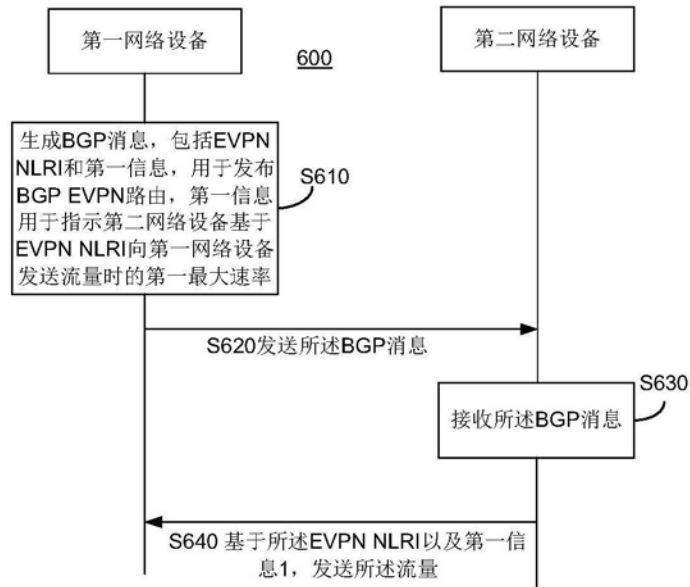


图6



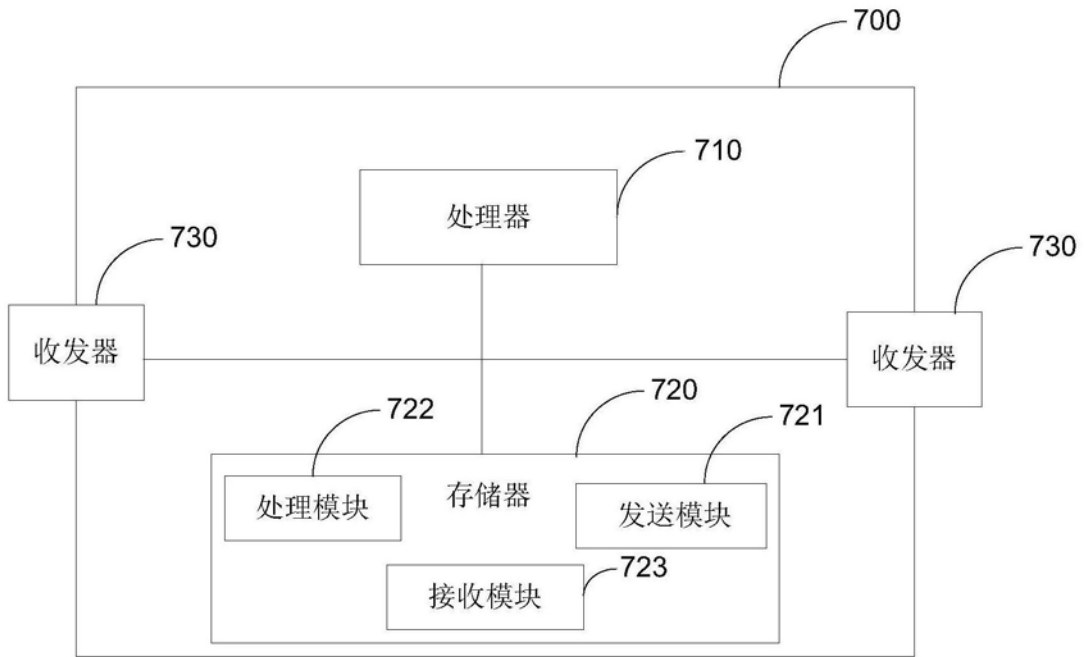


图7

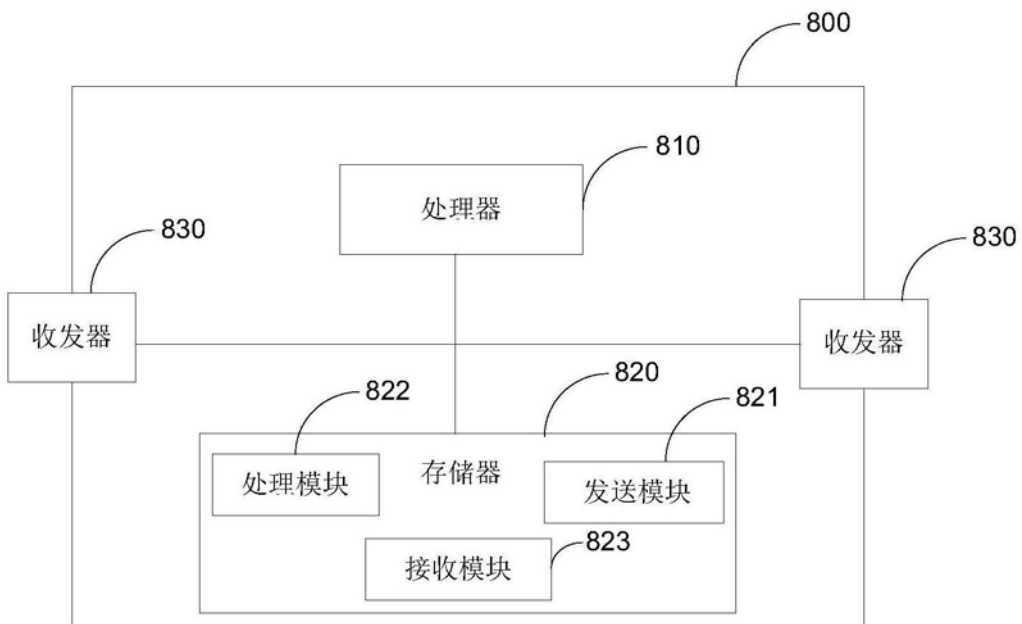


图8