



(12) 发明专利申请

(10) 申请公布号 CN 103905284 A

(43) 申请公布日 2014. 07. 02

(21) 申请号 201210583884. 9

(22) 申请日 2012. 12. 27

(71) 申请人 杭州华三通信技术有限公司

地址 310053 浙江省杭州市高新技术产业开发区之江科技工业园六和路 310 号华为杭州生产基地

(72) 发明人 杜一鸣

(74) 专利代理机构 北京鑫媛睿博知识产权代理有限公司 11297

代理人 龚家骅

(51) Int. Cl.

H04L 12/46 (2006. 01)

H04L 12/801 (2013. 01)

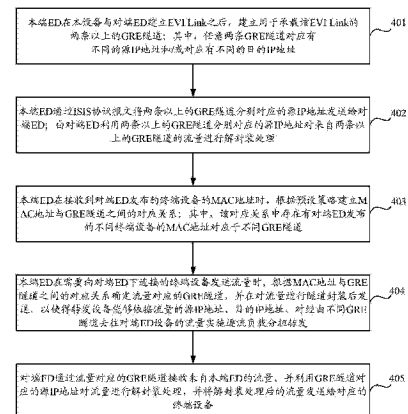
权利要求书3页 说明书9页 附图4页

(54) 发明名称

一种基于 EVI 网络的流量负载分担方法和设备

(57) 摘要

本发明公开了一种基于 EVI 网络的流量负载分担方法和设备,该方法包括:本端 ED 在本设备与对端 ED 建立 EVI Link 之后,建立用于承载所述 EVI Link 的两条以上的 GRE 隧道;所述本端 ED 在接收到所述对端 ED 发布的终端设备的 MAC 地址时,根据预设策略建立 MAC 地址与 GRE 隧道之间的对应关系;所述本端 ED 根据所述对应关系确定流量对应的 GRE 隧道,并通过该 GRE 隧道对流量进行发送。本发明中,可以使流量能够在核心网中进行负载分担,避免链路拥塞,并提高实际带宽利用率。



1. 一种基于以太网虚拟化互联 EVI 网络的流量负载分担方法,应用于包括本端边缘设备 ED 和对端 ED 的 EVI 网络中,其特征在于,包括:

所述本端 ED 在本设备与所述对端 ED 建立 EVI Link 之后,建立用于承载所述 EVI Link 的两条以上的通用路由封装 GRE 隧道,其中,任意两条 GRE 隧道对应有不同的源 IP 地址和/或对应有不同的目的 IP 地址;

所述本端 ED 在接收到所述对端 ED 发布的终端设备的介质访问控制 MAC 地址时,根据预设策略建立 MAC 地址与 GRE 隧道之间的对应关系,其中,所述对应关系中存在着有所述对端 ED 发布的不同终端设备的 MAC 地址对应于不同 GRE 隧道;

所述本端 ED 在需要向所述对端 ED 下连接的终端设备发送流量时,根据所述对应关系确定流量对应的 GRE 隧道,并在对流量进行隧道封装后发送,以使得转发设备能够依据流量的源 IP 地址、目的 IP 地址,对经由不同 GRE 隧道去往所述对端 ED 设备的流量实施逐流负载分担转发。

2. 如权利要求 1 所述的方法,其特征在于,所述本端 ED 建立用于承载所述 EVI Link 的两条以上的通用路由封装 GRE 隧道,之后还包括:

在所述任意两条 GRE 隧道对应有不同的源 IP 地址,且对应有相同的目的 IP 地址时,所述本端 ED 通过第一中间系统到中间系统 ISIS 协议报文将所述两条以上的 GRE 隧道分别对应的源 IP 地址发送给所述对端 ED;

由所述对端 ED 利用所述两条以上的 GRE 隧道分别对应的源 IP 地址对来自所述两条以上的 GRE 隧道的流量进行解封装处理。

3. 如权利要求 1 所述的方法,其特征在于,在所述任意两条 GRE 隧道对应有不同的源 IP 地址,且对应有相同的目的 IP 地址时,所述本端 ED 在接收到所述对端 ED 发布的终端设备的介质访问控制 MAC 地址时,根据预设策略建立 MAC 地址与 GRE 隧道之间的对应关系,具体包括:

所述本端 ED 接收来自所述对端 ED 的第二 ISIS 协议报文,所述第二 ISIS 协议报文中携带了所述对端 ED 下连接的多个终端设备的 MAC 地址;

所述本端 ED 通过静态配置方式建立所述对端 ED 下连接的多个终端设备的 MAC 地址与 GRE 隧道之间的对应关系,以使所述对应关系中存在着不同终端设备的 MAC 地址对应于不同 GRE 隧道;或者,

所述本端 ED 通过顺序配置方式建立所述对端 ED 下连接的多个终端设备的 MAC 地址与 GRE 隧道之间的对应关系,以使所述对应关系中存在着不同终端设备的 MAC 地址对应于不同 GRE 隧道;或者,

所述本端 ED 通过 Hash 配置方式建立所述对端 ED 下连接的多个终端设备的 MAC 地址与 GRE 隧道之间的对应关系,以使所述对应关系中存在着不同终端设备的 MAC 地址对应于不同 GRE 隧道。

4. 一种基于以太网虚拟化互联 EVI 网络的流量负载分担方法,应用于包括本端边缘设备 ED 和对端 ED 的 EVI 网络中,其特征在于,包括:

所述对端 ED 在本设备与所述本端 ED 建立 EVI Link 之后,接收来自所述本端 ED 的第一中间系统到中间系统 ISIS 协议报文,所述第一 ISIS 协议报文中携带了两条以上的通用路由封装 GRE 隧道分别对应的源 IP 地址;其中,所述两条以上的 GRE 隧道用于承载所述 EVI

Link,且任意两条 GRE 隧道对应有不同的源 IP 地址,并对应有相同目的 IP 地址或不同目的 IP 地址;

所述对端 ED 通过流量对应的 GRE 隧道接收来自所述本端 ED 的流量,通过 GRE 隧道对应的源 IP 地址对所述流量进行解封装处理,并将解封装处理后的流量发送给对应的终端设备;其中,所述流量对应的 GRE 隧道为本端 ED 根据所述对端 ED 发布的终端设备的介质访问控制 MAC 地址与 GRE 隧道之间的对应关系确定的,且所述对应关系中存在着有所述对端 ED 发布的不同终端设备的 MAC 地址对应于不同 GRE 隧道。

5. 如权利要求 4 所述的方法,其特征在于,在所述任意两条 GRE 隧道对应有不同的源 IP 地址,且对应有相同的目的 IP 地址时,所述方法还包括:

所述对端 ED 在学习到本设备下连接的多个终端设备的 MAC 地址后,通过第二 ISIS 协议报文将所述多个终端设备的 MAC 地址发送给所述本端 ED;由所述本端 ED 通过静态配置方式或顺序配置方式或 Hash 配置方式建立所述对端 ED 下连接的多个终端设备的 MAC 地址与 GRE 隧道之间的对应关系,以使所述对应关系中存在着不同终端设备的 MAC 地址对应于不同 GRE 隧道。

6. 一种本端边缘设备 ED,应用于包括所述本端 ED 和对端 ED 的以太网虚拟化互联 EVI 网络中,其特征在于,所述本端 ED 包括:

第一建立模块,用于在本设备与所述对端 ED 建立 EVI Link 之后,建立用于承载所述 EVI Link 的两条以上的通用路由封装 GRE 隧道,其中,任意两条 GRE 隧道对应有不同的源 IP 地址和 / 或对应有不同的目的 IP 地址;

第二建立模块,用于在接收到所述对端 ED 发布的终端设备的介质访问控制 MAC 地址时,根据预设策略建立 MAC 地址与 GRE 隧道之间的对应关系,其中,所述对应关系中存在着有所述对端 ED 发布的不同终端设备的 MAC 地址对应于不同 GRE 隧道;

发送模块,用于在需要向所述对端 ED 下连接的终端设备发送流量时,根据所述对应关系确定流量对应的 GRE 隧道,并在对流量进行隧道封装后发送,以使得转发设备能够依据流量的源 IP 地址、目的 IP 地址,对经由不同 GRE 隧道去往所述对端 ED 设备的流量实施逐流负载分担转发。

7. 如权利要求 6 所述的本端 ED,其特征在于,

所述发送模块,还用于在所述任意两条 GRE 隧道对应有不同的源 IP 地址,且对应有相同的目的 IP 地址时,通过第一中间系统到中间系统 ISIS 协议报文将所述两条以上的 GRE 隧道分别对应的源 IP 地址发送给所述对端 ED;

由所述对端 ED 利用所述两条以上的 GRE 隧道分别对应的源 IP 地址对来自所述两条以上的 GRE 隧道的流量进行解封装处理。

8. 如权利要求 6 所述的本端 ED,其特征在于,

所述第二建立模块,具体用于在所述任意两条 GRE 隧道对应有不同的源 IP 地址,且对应有相同的目的 IP 地址时,接收来自所述对端 ED 的第二 ISIS 协议报文,所述第二 ISIS 协议报文中携带了所述对端 ED 下连接的多个终端设备的 MAC 地址;并通过静态配置方式建立所述对端 ED 下连接的多个终端设备的 MAC 地址与 GRE 隧道之间的对应关系,以使所述对应关系中存在着不同终端设备的 MAC 地址对应于不同 GRE 隧道;或者,通过顺序配置方式建立所述对端 ED 下连接的多个终端设备的 MAC 地址与 GRE 隧道之间的对应关系,以使所述对应关

系中存在不同终端设备的 MAC 地址对应于不同 GRE 隧道 ;或者,通过 Hash 配置方式建立所述对端 ED 下连接的多个终端设备的 MAC 地址与 GRE 隧道之间的对应关系,以使所述对应关系中存在不同终端设备的 MAC 地址对应于不同 GRE 隧道。

9. 一种对端边缘设备 ED,应用于包括本端 ED 和所述对端 ED 的以太网虚拟化互联 EVI 网络中,其特征在于,所述对端 ED 包括 :

第一接收模块,用于在本设备与所述本端 ED 建立 EVI Link 之后,接收来自所述本端 ED 的第一中间系统到中间系统 ISIS 协议报文,所述第一 ISIS 协议报文中携带了两条以上的通用路由封装 GRE 隧道分别对应的源 IP 地址 ;其中,所述两条以上的 GRE 隧道用于承载所述 EVI Link,且任意两条 GRE 隧道对应有不同的源 IP 地址,并对应有相同目的 IP 地址或不同目的 IP 地址 ;

第二接收模块,用于通过流量对应的 GRE 隧道接收来自所述本端 ED 的流量,所述流量对应的 GRE 隧道为本端 ED 根据对端 ED 发布的终端设备的介质访问控制 MAC 地址与 GRE 隧道之间的对应关系确定的,且所述对应关系中存在有对端 ED 发布的不同终端设备的 MAC 地址对应于不同 GRE 隧道 ;

发送模块,用于通过 GRE 隧道对应的源 IP 地址对所述流量进行解封装处理,并将解封装处理后的流量发送给对应的终端设备。

10. 如权利要求 9 所述的对端 ED,其特征在于,

所述发送模块,还用于在任意两条 GRE 隧道对应有不同的源 IP 地址,且对应有相同的目的 IP 地址时,在学习到本设备下连接的多个终端设备的 MAC 地址后,通过第二 ISIS 协议报文将多个终端设备的 MAC 地址发送给本端 ED ;由所述本端 ED 通过静态配置方式或顺序配置方式或 Hash 配置方式建立对端 ED 下连接的多个终端设备的 MAC 地址与 GRE 隧道之间的对应关系,以使所述对应关系中存在不同终端设备的 MAC 地址对应于不同 GRE 隧道。

一种基于 EVI 网络的流量负载分担方法和设备

技术领域

[0001] 本发明涉及通信技术领域,特别涉及一种基于 EVI (Ethernet Virtualization Interconnection,以太网虚拟化互联)网络的流量负载分担方法和设备。

背景技术

[0002] 随着 Internet 的高速发展,网络地域跨度越来越大,为了更好的提供服务,用户通常会在异地部署数据中心,这些数据中心之间可以通过虚拟机进行自由迁移来实现负载分担和高可靠性;由于虚拟机迁移过程对用户透明,不能改变 IP 地址,因此要求异地数据中心之间实现二层网络互联,为此提出了 EVI 网络,该 EVI 网络中只是在站点网络的 ED (Edge Device,边缘设备)上维护路由和转发信息,无需改变站点网络内部和核心网络内部的路由和转发信息。

[0003] 如图 1 所示,部署 EVI 网络后的网络由核心网络、站点网络和 EVI 网络组成,该站点网络为通过一台或多台 ED 连接到核心网络的具有独立业务功能的二层网络;该 EVI 网络为站点网络的 ED 之间建立的虚拟网络,用于提供站点网络之间的二层互联;该核心网络为提供站点网络之间互联的网络。

[0004] EVI 网络由 EVI-Link(连接)接口和虚拟连接(即 EVI Link,ED 间的双向虚拟以太网通道,该通道由 GRE (Generic Routing Encapsulation,通用路由封装)隧道承载,一条 GRE 隧道上可承载多条 EVI Link)组成,用于承载站点网络间扩展 VLAN (Virtual Local Area Network,虚拟局域网)的二层流量,并通过 ENDP (EVI Neighbor Discovery Protocol,EVI 邻居发现协议)来自动发现站点网络;各站点网络可以通过 ISIS(Intermediate System to Intermediate System,中间系统到中间系统)协议学习异地站点网络的 MAC(Media Access Control,介质访问控制)地址,以实现不同站点网络之间的数据传输。

[0005] 在 EVI 网络中,ED 之间需要通告 MAC 地址信息,并使用这些 MAC 地址信息来指导报文的转发;且在通告 MAC 地址信息之前,ED 之间需要相互发现对方且形成邻居关系;为此,提出 ENDP 协议来实现邻居发现功能。

[0006] 具体的,(1)将 ENDS (EVI Neighbor Discovery Server,EVI 邻居发现服务器)、ENDC (EVI Neighbor Discovery Client,EVI 邻居发现客户端)部署在 ED 上,该 ENDS 用来维护同一 EVI 网络实例中所有客户端的信息(如 IP 地址等);(2)ENDC 向 ENDS 发送注册请求报文,其中携带 ENDC 的 IP 地址等信息,ENDS 向 ENDC 发送应答报文,其中携带所有 ENDC 信息;(3)ENDC 在收到应答报文后,与每个 ENDC 建立 EVI Link。

[0007] EVI 网络中所有站点网络的 ED 进行上述注册过程后,ED 可发现 EVI 网络的邻居,并建立 EVI Link;ED 间在建立 EVI Link 后,通过在该 EVI 网络上运行 ISIS 协议来通告单播 MAC 地址可达性信息,该 ISIS 协议主要包括 EVI ISIS 邻居协商、LSP(Link State Protocol Data Unit,链路状态协议数据单元)更新等过程,涉及到的报文都在邻居发现协议建立的 EVI Link 中传输。

[0008] 现有技术中,如图 2 所示,EVI 网络中的流量转发过程包括以下步骤:步骤 a、ED 在

收到流量之后,根据目的MAC地址查找本地MAC地址表,得到出接口为EVI Link对应的GRE隧道接口。步骤b、ED对流量进行封装处理;具体的,ED对流量进行GRE封装,添加外层IP头、链路层头以及校验和;外层IP头中的源IP地址为GRE隧道接口的IP地址,目的IP地址是异地站点ED的IP地址。步骤c、ED将封装后的流量从GRE隧道接口发送到核心网,最终到达异地站点ED。步骤d、异地站点ED对流量进行解封装处理。步骤e、异地站点ED根据解封装后的流量在本地进行目的MAC地址查找,得到出接口为本地接口,并且从本地接口发送该流量给终端设备。

[0009] 在上述实现方式中,两台ED间进行流量转发时,相同EVI实例内的流量通过一条EVI Link转发,所有流量被封装成相同源IP地址和目的IP地址的流量,无法实现负载分担,容易出现链路拥塞,且降低实际带宽利用率。

发明内容

[0010] 本发明提供一种基于EVI网络的流量负载分担方法和设备,为EVI网络的流量在核心网中实现负载分担提供支持,以避免链路拥塞,并提高实际带宽利用率。为了达到上述目的,则:

[0011] 本发明实施例提供一种基于以太网虚拟化互联EVI网络的流量负载分担方法,应用于包括本端边缘设备ED和对端ED的EVI网络中,包括:

[0012] 所述本端ED在本设备与所述对端ED建立EVI Link之后,建立用于承载所述EVI Link的两条以上的通用路由封装GRE隧道,其中,任意两条GRE隧道对应有不同的源IP地址和/或对应有不同的目的IP地址;

[0013] 所述本端ED在接收到所述对端ED发布的终端设备的介质访问控制MAC地址时,根据预设策略建立MAC地址与GRE隧道之间的对应关系,其中,所述对应关系中存在着有所述对端ED发布的不同终端设备的MAC地址对应于不同GRE隧道;

[0014] 所述本端ED在需要向所述对端ED下连接的终端设备发送流量时,根据所述对应关系确定流量对应的GRE隧道,并对流量进行隧道封装后发送,以使得转发设备能够依据流量的源IP地址、目的IP地址,对经由不同GRE隧道去往所述对端ED设备的流量实施逐流负载分担转发。

[0015] 所述本端ED建立用于承载所述EVI Link的两条以上的通用路由封装GRE隧道,之后还包括:

[0016] 在所述任意两条GRE隧道对应有不同的源IP地址,且对应有相同的目的IP地址时,所述本端ED通过第一中间系统到中间系统ISIS协议报文将所述两条以上的GRE隧道分别对应的源IP地址发送给所述对端ED;

[0017] 由所述对端ED利用所述两条以上的GRE隧道分别对应的源IP地址对来自所述两条以上的GRE隧道的流量进行解封装处理。

[0018] 在所述任意两条GRE隧道对应有不同的源IP地址,且对应有相同的目的IP地址时,所述本端ED在接收到所述对端ED发布的终端设备的介质访问控制MAC地址时,根据预设策略建立MAC地址与GRE隧道之间的对应关系,具体包括:

[0019] 所述本端ED接收来自所述对端ED的第二ISIS协议报文,所述第二ISIS协议报文中携带了所述对端ED下连接的多个终端设备的MAC地址;

[0020] 所述本端 ED 通过静态配置方式建立所述对端 ED 下连接的多个终端设备的 MAC 地址与 GRE 隧道之间的对应关系,以使所述对应关系中存在不同终端设备的 MAC 地址对应于不同 GRE 隧道;或者,

[0021] 所述本端 ED 通过顺序配置方式建立所述对端 ED 下连接的多个终端设备的 MAC 地址与 GRE 隧道之间的对应关系,以使所述对应关系中存在不同终端设备的 MAC 地址对应于不同 GRE 隧道;或者,

[0022] 所述本端 ED 通过 Hash 配置方式建立所述对端 ED 下连接的多个终端设备的 MAC 地址与 GRE 隧道之间的对应关系,以使所述对应关系中存在不同终端设备的 MAC 地址对应于不同 GRE 隧道。

[0023] 本发明实施例提供一种基于以太网虚拟化互联 EVI 网络的流量负载分担方法,应用于包括本端边缘设备 ED 和对端 ED 的 EVI 网络中,包括:

[0024] 所述对端 ED 在本设备与所述本端 ED 建立 EVI Link 之后,接收来自所述本端 ED 的第一中间系统到中间系统 ISIS 协议报文,所述第一 ISIS 协议报文中携带了两条以上的通用路由封装 GRE 隧道分别对应的源 IP 地址;其中,所述两条以上的 GRE 隧道用于承载所述 EVI Link,且任意两条 GRE 隧道对应有不同的源 IP 地址,并对应有相同目的 IP 地址或不同目的 IP 地址;

[0025] 所述对端 ED 通过流量对应的 GRE 隧道接收来自所述本端 ED 的流量,通过 GRE 隧道对应的源 IP 地址对所述流量进行解封装处理,并将解封装处理后的流量发送给对应的终端设备;其中,所述流量对应的 GRE 隧道为本端 ED 根据所述对端 ED 发布的终端设备的介质访问控制 MAC 地址与 GRE 隧道之间的对应关系确定的,且所述对应关系中存在有所述对端 ED 发布的不同终端设备的 MAC 地址对应于不同 GRE 隧道。

[0026] 在所述任意两条 GRE 隧道对应有不同的源 IP 地址,且对应有相同的目的 IP 地址时,所述方法还包括:

[0027] 所述对端 ED 在学习到本设备下连接的多个终端设备的 MAC 地址后,通过第二 ISIS 协议报文将所述多个终端设备的 MAC 地址发送给所述本端 ED;由所述本端 ED 通过静态配置方式或顺序配置方式或 Hash 配置方式建立所述对端 ED 下连接的多个终端设备的 MAC 地址与 GRE 隧道之间的对应关系,以使所述对应关系中存在不同终端设备的 MAC 地址对应于不同 GRE 隧道。

[0028] 本发明实施例提供一种本端边缘设备 ED,应用于包括所述本端 ED 和对端 ED 的以太网虚拟化互联 EVI 网络中,所述本端 ED 包括:

[0029] 第一建立模块,用于在本设备与所述对端 ED 建立 EVI Link 之后,建立用于承载所述 EVI Link 的两条以上的通用路由封装 GRE 隧道,其中,任意两条 GRE 隧道对应有不同的源 IP 地址和/或对应有不同的目的 IP 地址;

[0030] 第二建立模块,用于在接收到所述对端 ED 发布的终端设备的介质访问控制 MAC 地址时,根据预设策略建立 MAC 地址与 GRE 隧道之间的对应关系,其中,所述对应关系中存在有所述对端 ED 发布的不同终端设备的 MAC 地址对应于不同 GRE 隧道;

[0031] 发送模块,用于在需要向所述对端 ED 下连接的终端设备发送流量时,根据所述对应关系确定流量对应的 GRE 隧道,并在对流量进行隧道封装后发送,以使得转发设备能够依据流量的源 IP 地址、目的 IP 地址,对经由不同 GRE 隧道去往所述对端 ED 设备的流量实

施逐流负载分担转发。

[0032] 所述发送模块,还用于在所述任意两条 GRE 隧道对应有不同的源 IP 地址,且对应有相同的目的 IP 地址时,通过第一中间系统到中间系统 ISIS 协议报文将所述两条以上的 GRE 隧道分别对应的源 IP 地址发送给所述对端 ED;

[0033] 由所述对端 ED 利用所述两条以上的 GRE 隧道分别对应的源 IP 地址对来自所述两条以上的 GRE 隧道的流量进行解封装处理。

[0034] 所述第二建立模块,具体用于在所述任意两条 GRE 隧道对应有不同的源 IP 地址,且对应有相同的目的 IP 地址时,接收来自所述对端 ED 的第二 ISIS 协议报文,所述第二 ISIS 协议报文中携带了所述对端 ED 下连接的多个终端设备的 MAC 地址;并通过静态配置方式建立所述对端 ED 下连接的多个终端设备的 MAC 地址与 GRE 隧道之间的对应关系,以使所述对应关系中不存在不同终端设备的 MAC 地址对应于不同 GRE 隧道;或者,通过顺序配置方式建立所述对端 ED 下连接的多个终端设备的 MAC 地址与 GRE 隧道之间的对应关系,以使所述对应关系中不存在不同终端设备的 MAC 地址对应于不同 GRE 隧道;或者,通过 Hash 配置方式建立所述对端 ED 下连接的多个终端设备的 MAC 地址与 GRE 隧道之间的对应关系,以使所述对应关系中不存在不同终端设备的 MAC 地址对应于不同 GRE 隧道。

[0035] 本发明实施例提供一种对端边缘设备 ED,应用于包括本端 ED 和所述对端 ED 的以太网虚拟化互联 EVI 网络中,所述对端 ED 包括:

[0036] 第一接收模块,用于在本设备与所述本端 ED 建立 EVI Link 之后,接收来自所述本端 ED 的第一中间系统到中间系统 ISIS 协议报文,所述第一 ISIS 协议报文中携带了两条以上的通用路由封装 GRE 隧道分别对应的源 IP 地址;其中,所述两条以上的 GRE 隧道用于承载所述 EVI Link,且任意两条 GRE 隧道对应有不同的源 IP 地址,并对应有相同目的 IP 地址或不同目的 IP 地址;

[0037] 第二接收模块,用于通过流量对应的 GRE 隧道接收来自所述本端 ED 的流量,所述流量对应的 GRE 隧道为本端 ED 根据对端 ED 发布的终端设备的介质访问控制 MAC 地址与 GRE 隧道之间的对应关系确定的,且所述对应关系中不存在有对端 ED 发布的不同终端设备的 MAC 地址对应于不同 GRE 隧道;

[0038] 发送模块,用于通过 GRE 隧道对应的源 IP 地址对所述流量进行解封装处理,并将解封装处理后的流量发送给对应的终端设备。

[0039] 所述发送模块,还用于在任意两条 GRE 隧道对应有不同的源 IP 地址,且对应有相同的目的 IP 地址时,在学习到本设备下连接的多个终端设备的 MAC 地址后,通过第二 ISIS 协议报文将多个终端设备的 MAC 地址发送给本端 ED;由所述本端 ED 通过静态配置方式或顺序配置方式或 Hash 配置方式建立对端 ED 下连接的多个终端设备的 MAC 地址与 GRE 隧道之间的对应关系,以使所述对应关系中不存在不同终端设备的 MAC 地址对应于不同 GRE 隧道。

[0040] 与现有技术相比,本发明实施例至少具有以下优点:本发明实施例中,能够将相同 EVI Link 的流量封装成为目的 IP 地址相同,源 IP 地址不同的流量(即通过不同的 GRE 隧道的源 IP 地址对流量进行封装),从而为 EVI 网络的流量在核心网中实现负载分担提供技术支持,以使流量能够在核心网中进行负载分担,避免链路拥塞,并提高实际带宽利用率。

附图说明

- [0041] 图 1 是现有技术中部署 EVI 网络后的网络示意图；
- [0042] 图 2 是现有技术中 EVI 网络中的流量转发过程示意图；
- [0043] 图 3 是本发明实施例的应用场景示意图；
- [0044] 图 4 是本发明实施例提出的基于 EVI 网络的流量负载分担方法流程图；
- [0045] 图 5 是本发明实施例提出的本端 ED 的结构示意图；
- [0046] 图 6 是本发明实施例提出的对端 ED 的结构示意图。

具体实施方式

[0047] 下面结合附图对本发明实施例进行详细描述。

[0048] 在实际网络中，IP 核心网为了提供高可靠性和健壮性的服务，通常应用各种冗余的负载分担技术，比如链路聚合，等价路由等。在负载分担技术中，常用的两种负载分担方式有逐包分担和逐流分担。在逐包分担方式下，相同源 IP 地址和相同目的 IP 地址的 IP 报文能分担到冗余链路上；而在逐流分担方式下，相同源 IP 地址和相同目的 IP 地址的 IP 报文不能分担到冗余链路上。现有技术的 EVI 实例内的流量在通过一条 EVI Link 转发时，由于流量被封装成相同源 IP 地址和相同目的 IP 地址的报文，因此该流量在 IP 核心网中转发时，如果核心网中的设备仅支持逐流分担，则该流量无法在核心网中实现负载分担，容易导致链路拥塞，并降低网络实际带宽利用率。

[0049] 针对现有技术中存在的问题，本发明实施例提出一种基于 EVI 网络的流量负载分担方法，该方法应用于包括本端 ED 和对端 ED 的 EVI 网络中，本端 ED 与对端 ED 之间存在多条传输路径，且对端 ED 下连接有多个终端设备。

[0050] 以图 3 为本发明实施例的应用场景示意图，ED1 为本端 ED，ED2 为对端 ED，且 ED1 与 ED2 之间存在两条传输路径，分别为：ED1- 路由器 1- 路由器 2- 路由器 4- ED2，以及 ED1- 路由器 1- 路由器 3- 路由器 4- ED2；此外，ED2 下连接有终端设备 3（其 MAC 为 MAC3）和终端设备 4（其 MAC 为 MAC4）。

[0051] 基于上述应用场景，如图 4 所示，该方法包括以下步骤：

[0052] 步骤 401，本端 ED 在本设备与对端 ED 建立 EVI Link 之后，建立用于承载该 EVI Link 的两条以上的 GRE 隧道（即 2 条或者更多的 GRE 隧道）；其中，任意两条 GRE 隧道对应有不同的源 IP 地址和 / 或对应有不同的目的 IP 地址。

[0053] 本发明实施例中，针对 EVI Link 关联的 GRE 隧道，本端 ED 允许配置 GRE 隧道的源 IP 地址为多个 IP 地址（如 IP1 和 IP2），并配置 GRE 隧道的目的 IP 地址为相同 IP 地址；当本端 ED 与对端 ED 建立了 EVI Link 之后，本端 ED 利用多个源 IP 地址建立用于承载 EVI Link 的两条以上的 GRE 隧道；或者，

[0054] 针对 EVI Link 关联的 GRE 隧道，本端 ED 允许配置 GRE 隧道的目的 IP 地址为多个 IP 地址，并配置 GRE 隧道的源 IP 地址为相同 IP 地址；当本端 ED 与对端 ED 建立了 EVI Link 之后，本端 ED 利用多个目的 IP 地址建立用于承载 EVI Link 的两条以上的 GRE 隧道；或者，

[0055] 针对 EVI Link 关联的 GRE 隧道，本端 ED 允许配置 GRE 隧道的源 IP 地址为多个 IP 地址，并配置 GRE 隧道的目的 IP 地址为多个 IP 地址；当本端 ED 与对端 ED 建立了 EVI Link 之后，本端 ED 利用多个源 IP 地址和多个目的 IP 地址建立用于承载 EVI Link 的两条以上的 GRE 隧道。

[0056] 为了方便描述,本发明实施例中以任意两条 GRE 隧道对应有不同的源 IP 地址,且对应有相同的目的 IP 地址为例进行后续说明,对于其它情况的处理与此相类似,后续不再进行详加说明。

[0057] 例如,在图 3 所示的应用场景下,ED1 与 ED2 之间存在两条传输路径,基于此,ED1 可以为 EVI Link 建立两条 GRE 隧道,这两条 GRE 隧道分别为 Tunnel1 和 Tunnel2 ;其中,Tunnel1 对应传输路径 ED1- 路由器 1- 路由器 2- 路由器 4-ED2,且 Tunnel1 的源 IP 地址为 IP1, Tunnel1 的目的 IP 为 ED2 的 IP 地址 ;Tunnel2 对应传输路径 ED1- 路由器 1- 路由器 3- 路由器 4-ED2,且 Tunnel2 的源 IP 地址为 IP2, Tunnel2 的目的 IP 为 ED2 的 IP 地址。

[0058] 步骤 402,本端 ED 通过 ISIS 协议报文(为了区分方便,该 ISIS 协议报文称为第一 ISIS 协议报文)将两条以上的 GRE 隧道分别对应的源 IP 地址(如 IP1 和 IP2)发送给对端 ED ;由对端 ED 利用两条以上的 GRE 隧道分别对应的源 IP 地址对来自两条以上的 GRE 隧道的流量进行解封装处理。

[0059] 本发明实施例中,ED1 为 EVI Link 建立 Tunnel1 和 Tunnel2 之后,还需要通过第一 ISIS 协议报文将 Tunnel1 对应的源 IP 地址(如 IP1)和 Tunnel2 对应的源 IP 地址(如 IP2)通知给 ED2 ;ED2 在收到该第一 ISIS 协议报文后,分别创建 Tunnel1 和 Tunnel2 的解封装表项,并利用 Tunnel1 和 Tunnel2 的解封装表项分别对来自 Tunnel1 和 Tunnel2 的流量进行解封装处理。

[0060] 步骤 403,本端 ED 在接收到对端 ED 发布的终端设备的 MAC 地址时,根据预设策略建立 MAC 地址与 GRE 隧道之间的对应关系 ;其中,该对应关系中存在着对端 ED 发布的不同终端设备的 MAC 地址对应于不同 GRE 隧道。

[0061] 本发明实施例中,本端 ED 在接收到对端 ED 发布的终端设备的 MAC 地址时,根据预设策略建立 MAC 地址与 GRE 隧道之间的对应关系,具体包括 :

[0062] 步骤 1、对端 ED 在学习到本设备下连接的多个终端设备的 MAC 地址后,通过 ISIS 协议报文(为了区分方便,该 ISIS 协议报文称为第二 ISIS 协议报文)将本设备下连接的多个终端设备的 MAC 地址发送给本端 ED。

[0063] 例如,在图 3 所示的应用场景下,ED2 下连接有终端设备 3 (其 MAC 地址为 MAC3)和终端设备 4 (其 MAC 地址为 MAC4),因此,当 ED2 上学习到同一 EVI 实例内的两个 MAC 地址(MAC3 和 MAC4)后,ED2 会通过第二 ISIS 协议报文将该 MAC3 和 MAC4 发送给 ED1 ;其中,由于 ED2 与 ED1 之间存在两条传输路径,因此 ED2 可以基于系统指定情况或者负载分担策略选择其中的一条传输路径发送用于发布 MAC 地址的第二 ISIS 协议报文。

[0064] 步骤 2、本端 ED 接收来自对端 ED 的第二 ISIS 协议报文,该第二 ISIS 协议报文中携带了对端 ED 下连接的多个终端设备的 MAC 地址。

[0065] 步骤 3、本端 ED 根据预设策略建立 MAC 地址与 GRE 隧道之间的对应关系,以使该 MAC 地址与 GRE 隧道之间的对应关系中存在着对端 ED 发布的不同终端设备的 MAC 地址对应于不同 GRE 隧道。

[0066] 本发明实施例中,本端 ED 根据预设策略建立 MAC 地址与 GRE 隧道之间的对应关系,具体包括但不限于以下方式 :本端 ED 通过静态配置方式建立对端 ED 下连接的多个终端设备的 MAC 地址与 GRE 隧道之间的对应关系,以使对应关系中存在着不同终端设备的 MAC 地址对应于不同 GRE 隧道 ;或者,本端 ED 通过顺序配置方式建立对端 ED 下连接的多个终端设

备的 MAC 地址与 GRE 隧道之间的对应关系,以使对应关系中不存在不同终端设备的 MAC 地址对应于不同 GRE 隧道;或者,本端 ED 通过 Hash 配置方式建立对端 ED 下连接的多个终端设备的 MAC 地址与 GRE 隧道之间的对应关系,以使对应关系中不存在不同终端设备的 MAC 地址对应于不同 GRE 隧道。

[0067] 在图 3 所示的应用场景下,由于 ED2 下连接的终端设备的 MAC 地址为 MAC3 和 MAC4,且 GRE 隧道为 Tunnel1 和 Tunnel2;因此,基于静态配置方式,由用户手动指定 MAC3 与 Tunnel2 具有对应关系,MAC4 与 Tunnel1 具有对应关系,即 ED1 建立 MAC3 与 Tunnel2 的对应关系,并建立 MAC4 与 Tunnel1 的对应关系;基于顺序配置方式,ED1 根据 MAC 地址学习到的先后顺序建立 MAC 地址与 GRE 隧道之间的对应关系,如 ED1 先学习到 MAC3 时,建立 MAC3 与 Tunnel1 的对应关系,并建立 MAC4 与 Tunnel2 的对应关系;基于 Hash 配置方式,ED1 根据 MAC 地址计算 hash 值,并根据 hash 值建立 MAC 地址与 GRE 隧道之间的对应关系,如 ED1 基于 MAC 地址的 hash 值建立 MAC3 与 Tunnel1 的对应关系,并建立 MAC4 与 Tunnel2 的对应关系。

[0068] 步骤 4、本端 ED 将 MAC 地址与 GRE 隧道之间的对应关系保存在 EVI 转发表中,且该 EVI 转发表用于指导后续流量的转发。

[0069] 步骤 404,本端 ED 在需要向对端 ED 下连接的终端设备发送流量时,根据 MAC 地址与 GRE 隧道之间的对应关系(位于 EVI 转发表中)确定流量对应的 GRE 隧道(即流量的目的 MAC 地址对应的 GRE 隧道),并在对流量进行隧道封装后发送,以使得转发设备能够依据流量的源 IP 地址、目的 IP 地址,对经由不同 GRE 隧道去往对端 ED 设备的流量实施逐流负载分担转发。

[0070] 在图 3 所示的应用场景下,ED1 收到发送给终端设备 3 的流量时,通过流量中携带的目的 MAC 地址(MAC3)查询 MAC 与 Tunnel 的对应关系,确定 GRE 隧道为 Tunnel1;因此 ED1 通过 Tunnel1 的源 IP 地址(IP1)对流量进行封装 GRE 封装,即添加外层 IP 头、链路层头以及校验和;且外层 IP 头中的源 IP 地址为 Tunnel1 的源 IP 地址(IP1),目的 IP 地址是 ED2 的 IP 地址。

[0071] ED1 收到发送给终端设备 4 的流量时,通过流量中携带的目的 MAC 地址(MAC4)查询 MAC 与 Tunnel 的对应关系,确定 GRE 隧道为 Tunnel2;因此 ED1 通过 Tunnel2 的源 IP 地址(IP2)对流量进行封装 GRE 封装,即添加外层 IP 头、链路层头以及校验和;且外层 IP 头中的源 IP 地址为 Tunnel2 的源 IP 地址(IP2),目的 IP 地址是 ED2 的 IP 地址。

[0072] 进一步的,在图 3 所示的应用场景下,ED1 通过 Tunnel1 将对应终端设备 3 的流量(使用 Tunnel1 的源 IP 地址进行封装后的流量)发送给 ED2,即 ED1 通过传输路径 ED1-路由器 1-路由器 2-路由器 4-ED2 将流量发送给 ED2;ED1 通过 Tunnel2 将对应终端设备 4 的流量(使用 Tunnel2 的源 IP 地址进行封装后的流量)发送给 ED2,即 ED1 通过传输路径 ED1-路由器 1-路由器 3-路由器 4-ED2 将流量发送给 ED2;从而实现流量的负载分担。

[0073] 步骤 405,对端 ED 通过流量对应的 GRE 隧道接收来自本端 ED 的流量(即本端 ED 发送给对端 ED 下连接的终端设备的流量),并利用 GRE 隧道对应的源 IP 地址对流量进行解封装处理(例如,对端 ED 利用 Tunnel1 和 Tunnel2 的解封装表项分别对来自 Tunnel1 和 Tunnel2 的流量进行解封装处理),并将解封装处理后的流量发送给对应的终端设备。

[0074] 综上所述,本发明实施例中,能够将相同 EVI Link 的流量封装成为目的 IP 地址

相同,源 IP 地址不同的流量(即通过不同的 GRE 隧道的源 IP 地址对流量进行封装),从而为 EVI 网络的流量在核心网中实现负载分担提供技术支持,以使流量能够在核心网中进行负载分担,避免链路拥塞。

[0075] 基于与上述方法同样的发明构思,本发明实施例还提出了一种本端边缘设备 ED,应用于包括所述本端 ED 和对端 ED 的以太网虚拟化互联 EVI 网络中,如图 5 所示,所述本端 ED 包括:

[0076] 第一建立模块 11,用于在本设备与所述对端 ED 建立 EVI Link 之后,建立用于承载所述 EVI Link 的两条以上的通用路由封装 GRE 隧道,其中,任意两条 GRE 隧道对应有不同的源 IP 地址和 / 或对应有不同的目的 IP 地址;

[0077] 第二建立模块 12,用于在接收到所述对端 ED 发布的终端设备的介质访问控制 MAC 地址时,根据预设策略建立 MAC 地址与 GRE 隧道之间的对应关系,其中,所述对应关系中存有有所述对端 ED 发布的不同终端设备的 MAC 地址对应于不同 GRE 隧道;

[0078] 发送模块 13,用于在需要向所述对端 ED 下连接的终端设备发送流量时,根据所述对应关系确定流量对应的 GRE 隧道,并在对流量进行隧道封装后发送,以使得转发设备能够依据流量的源 IP 地址、目的 IP 地址,对经由不同 GRE 隧道去往所述对端 ED 设备的流量实施逐流负载分担转发。

[0079] 所述发送模块 13,还用于在所述任意两条 GRE 隧道对应有不同的源 IP 地址,且对应有相同的目的 IP 地址时,通过第一中间系统到中间系统 ISIS 协议报文将所述两条以上的 GRE 隧道分别对应的源 IP 地址发送给所述对端 ED;

[0080] 由所述对端 ED 利用所述两条以上的 GRE 隧道分别对应的源 IP 地址对来自所述两条以上的 GRE 隧道的流量进行解封装处理。

[0081] 所述第二建立模块 12,具体用于在所述任意两条 GRE 隧道对应有不同的源 IP 地址,且对应有相同的目的 IP 地址时,接收来自所述对端 ED 的第二 ISIS 协议报文,所述第二 ISIS 协议报文中携带了所述对端 ED 下连接的多个终端设备的 MAC 地址;并通过静态配置方式建立所述对端 ED 下连接的多个终端设备的 MAC 地址与 GRE 隧道之间的对应关系,以使所述对应关系中不存在不同终端设备的 MAC 地址对应于不同 GRE 隧道;或者,通过顺序配置方式建立所述对端 ED 下连接的多个终端设备的 MAC 地址与 GRE 隧道之间的对应关系,以使所述对应关系中不存在不同终端设备的 MAC 地址对应于不同 GRE 隧道;或者,通过 Hash 配置方式建立所述对端 ED 下连接的多个终端设备的 MAC 地址与 GRE 隧道之间的对应关系,以使所述对应关系中不存在不同终端设备的 MAC 地址对应于不同 GRE 隧道。

[0082] 其中,本发明装置的各个模块可以集成于一体,也可以分离部署。上述模块可以合并为一个模块,也可以进一步拆分成多个子模块。

[0083] 基于与上述方法同样的发明构思,本发明实施例还提出了一种对端边缘设备 ED,应用于包括本端 ED 和所述对端 ED 的以太网虚拟化互联 EVI 网络中,如图 6 所示,所述对端 ED 包括:

[0084] 第一接收模块 21,用于在本设备与所述本端 ED 建立 EVI Link 之后,接收来自所述本端 ED 的第一中间系统到中间系统 ISIS 协议报文,所述第一 ISIS 协议报文中携带了两条以上的通用路由封装 GRE 隧道分别对应的源 IP 地址;其中,所述两条以上的 GRE 隧道用于承载所述 EVI Link,且任意两条 GRE 隧道对应有不同的源 IP 地址,并对应有相同目的 IP

地址或不同目的 IP 地址；

[0085] 第二接收模块 22, 用于通过流量对应的 GRE 隧道接收来自本端 ED 的流量, 所述流量对应的 GRE 隧道为本端 ED 根据对端 ED 发布的终端设备的介质访问控制 MAC 地址与 GRE 隧道之间的对应关系确定的, 且所述对应关系中存在有对端 ED 发布的不同终端设备的 MAC 地址对应于不同 GRE 隧道；

[0086] 发送模块 23, 用于通过 GRE 隧道对应的源 IP 地址对所述流量进行解封装处理, 并将解封装处理后的流量发送给对应的终端设备。

[0087] 所述发送模块 23, 还用于在任意两条 GRE 隧道对应有不同的源 IP 地址, 且对应相同的 IP 地址时, 在学习到本设备下连接的多个终端设备的 MAC 地址后, 通过第二 ISIS 协议报文将多个终端设备的 MAC 地址发送给本端 ED ; 由本端 ED 通过静态配置方式或顺序配置方式或 Hash 配置方式建立对端 ED 下连接的多个终端设备的 MAC 地址与 GRE 隧道之间的对应关系, 以使所述对应关系中存在不同终端设备的 MAC 地址对应于不同 GRE 隧道。

[0088] 其中, 本发明装置的各个模块可以集成于一体, 也可以分离部署。上述模块可以合并为一个模块, 也可以进一步拆分成多个子模块。

[0089] 通过以上的实施方式的描述, 本领域的技术人员可以清楚地了解到本发明可以通过硬件实现, 也可以借助软件加必要的通用硬件平台的方式来实现。基于这样的理解, 本发明的技术方案可以以软件产品的形式体现出来, 该软件产品可以存储在一个非易失性存储介质(可以是 CD-ROM, U 盘, 移动硬盘等)中, 包括若干指令用以使得一台计算机设备(可以是个人计算机, 服务器, 或者网络设备等)执行本发明各个实施例所述的方法。

[0090] 本领域技术人员可以理解附图只是一个优选实施例的示意图, 附图中的模块或流程并不一定是实施本发明所必须的。

[0091] 本领域技术人员可以理解实施例中的装置中的模块可以按照实施例描述进行分布于实施例的装置中, 也可以进行相应变化位于不同于本实施例的一个或多个装置中。上述实施例的模块可以合并为一个模块, 也可以进一步拆分成多个子模块。

[0092] 上述本发明序号仅仅为了描述, 不代表实施例的优劣。

[0093] 以上公开的仅为本发明的几个具体实施例, 但是, 本发明并非局限于此, 任何本领域的技术人员能思之的变化都应落入本发明的保护范围。

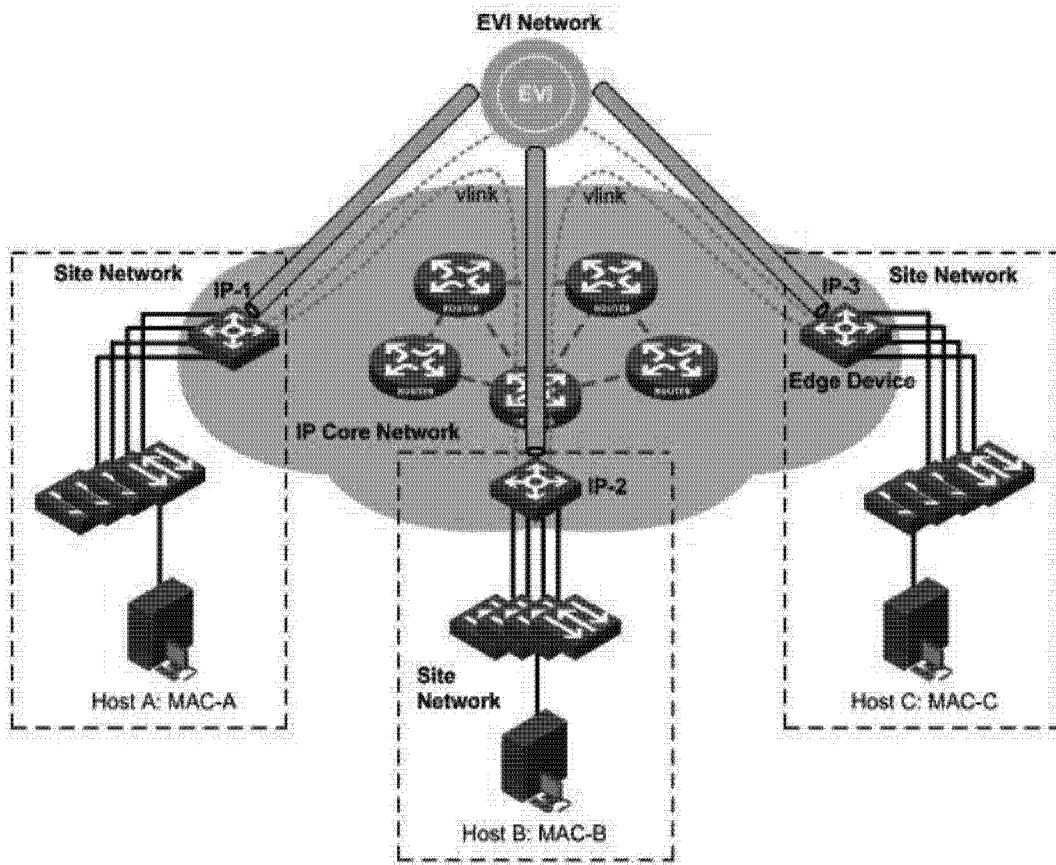


图 1

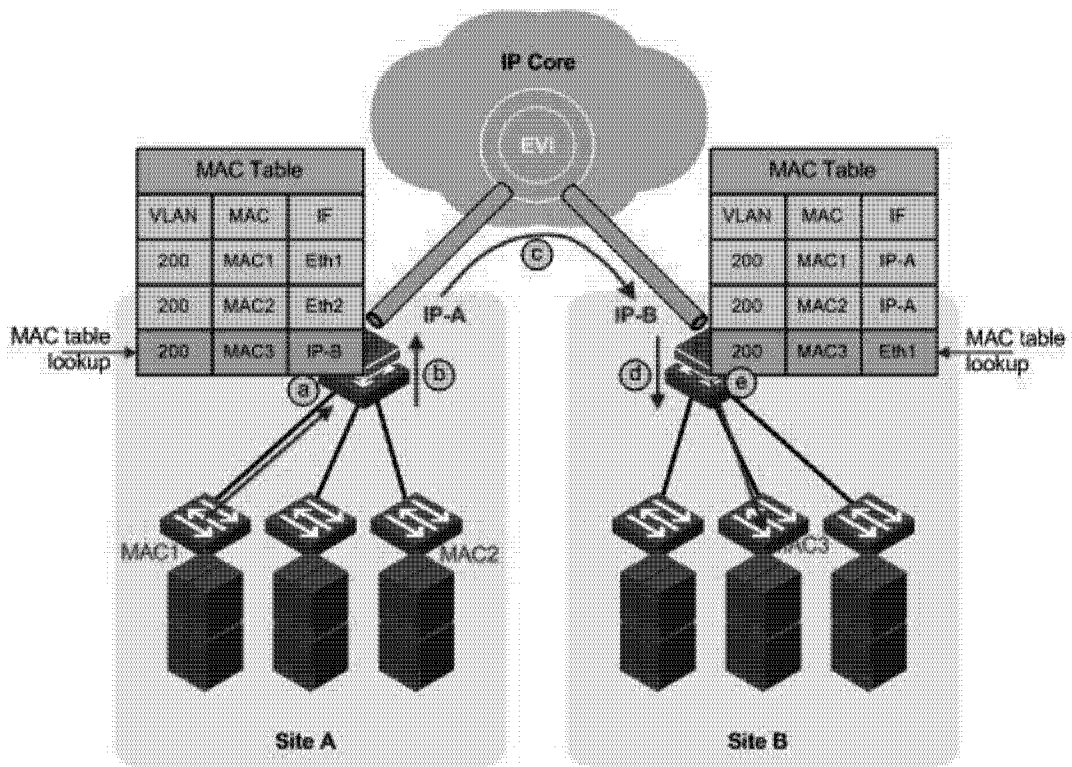


图 2

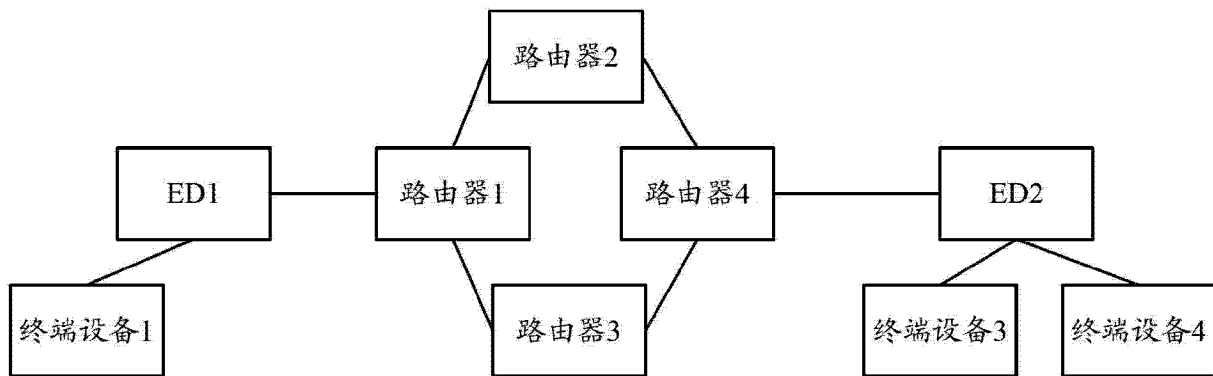


图 3

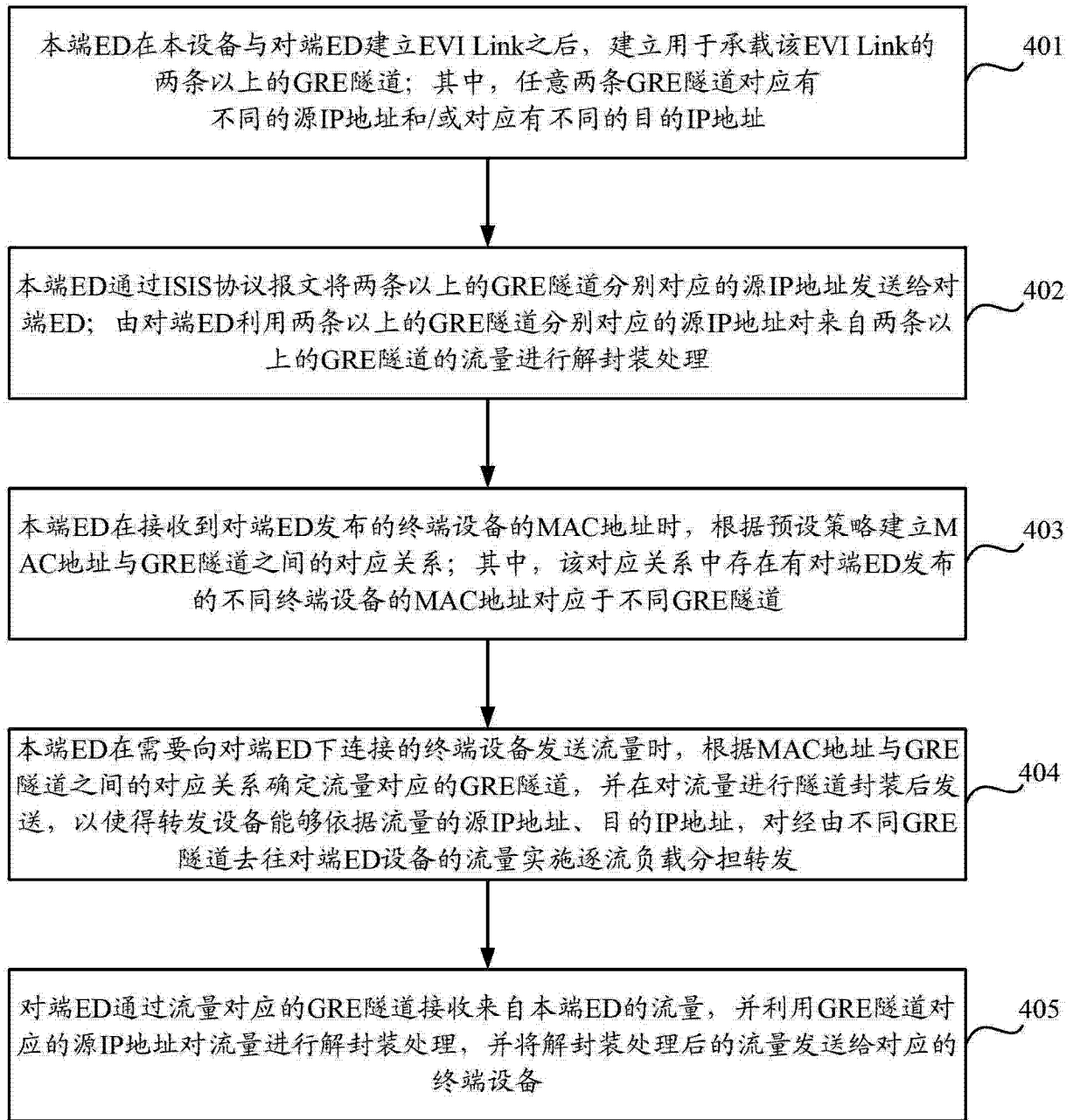


图 4

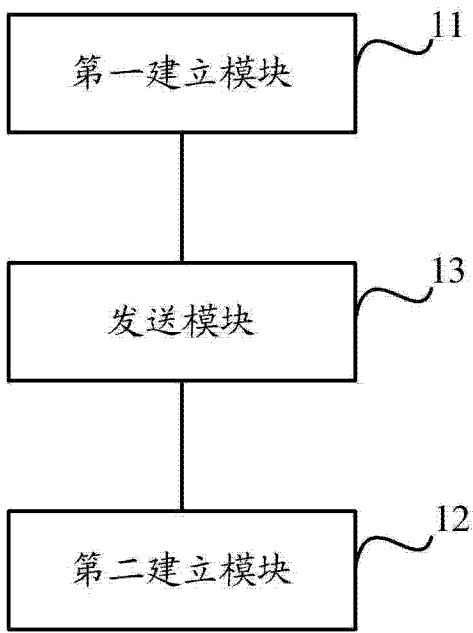


图 5

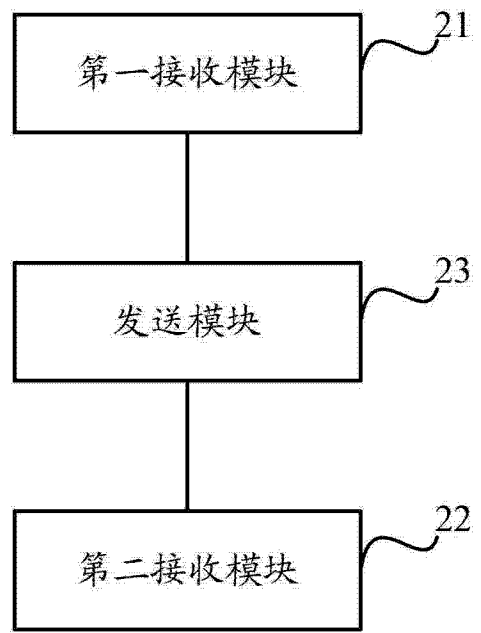


图 6