



(12) 发明专利申请

(10) 申请公布号 CN 103795631 A

(43) 申请公布日 2014. 05. 14

(21) 申请号 201210422550. 3

H04L 12/46 (2006. 01)

(22) 申请日 2012. 10. 30

(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 王奕 王伟 马臻

(74) 专利代理机构 北京德琦知识产权代理有限公司 11018

代理人 谢安昆 宋志强

(51) Int. Cl.

H04L 12/741 (2013. 01)

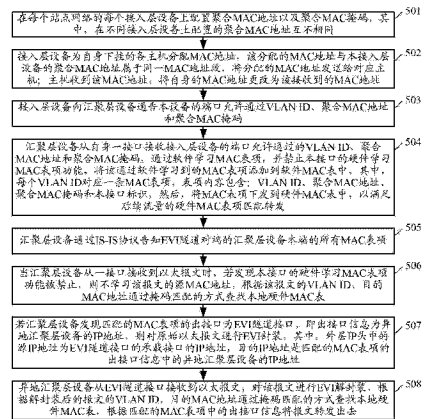
权利要求书2页 说明书10页 附图6页

(54) 发明名称

部署了以太网虚拟连接的网络中的流量转发方法及设备

(57) 摘要

本发明公开了部署了以太网虚拟连接的网络中的流量转发方法及设备,方法包括:在各站点的各接入层设备上分别配置聚合 MAC 地址及掩码,各接入层设备的聚合 MAC 地址互不相同;汇聚层设备从一接口接收接入层设备上报的端口允许通过的 VLAN ID、聚合 MAC 地址及掩码,通过软件学习 MAC 表项,将表项下发到硬件 MAC 表中,禁止接收接口的硬件学习 MAC 表项功能,将所有 MAC 表项发送给 EVI 隧道对端的汇聚层设备;汇聚层设备接收以太网报文,若发现接收接口的硬件学习 MAC 表项功能被禁止,不学习该报文的源 MAC 地址,根据报文的 VLAN ID 和目的 MAC 地址,采用掩码匹配方式在本地硬件 MAC 表中查找匹配的 MAC 表项,并转发该报文。本发明减少了汇聚层设备上的 MAC 表项。



1. 一种部署了以太网虚拟连接 EVI 的网络中的流量转发方法,其特征在于,该方法包括:

汇聚层设备从一接口接收接入层设备上报的该接入层设备的端口允许通过的虚拟局域网标识 VLAN ID、该接入层设备上配置的聚合 MAC 地址及掩码,通过软件学习 MAC 表项,将表项下发到硬件 MAC 表中,且禁止所述接口的硬件学习 MAC 表项功能,将通过软件学习到的所有 MAC 表项发送给 EVI 隧道对端的汇聚层设备;

所述汇聚层设备接收异地汇聚层设备发来的通过软件学习到的 MAC 表项,该 MAC 表项包括:该异地汇聚层设备下的接入层设备的端口允许通过的 VLAN ID、聚合 MAC 地址及掩码,将该 MAC 表项下发到硬件 MAC 表中;

所述汇聚层设备从所述接口接收以太报文,发现所述接口的硬件学习 MAC 表项功能被禁止,则不学习该报文的源 MAC 地址,根据报文的 VLANID 和目的 MAC 地址,采用掩码匹配方式在本地硬件 MAC 表中查找匹配的 MAC 表项,根据匹配的 MAC 表项转发该报文。

2. 根据权利要求 1 所述的方法,其特征在于,所述汇聚层设备根据匹配的 MAC 表项转发该报文包括:

若所述汇聚层设备发现匹配的 MAC 表项中的出接口信息为异地汇聚层设备的 IP 地址,则对该报文进行 EVI 封装,其中,外层 IP 头中的源 IP 地址为 EVI 隧道接口的承载接口的 IP 地址,目的 IP 地址是所述异地汇聚层设备的 IP 地址;

且,所述方法进一步包括:当所述汇聚层设备接收到异地汇聚层设备发来的以太报文时,对该报文进行 EVI 解封装,根据解封装后的报文的 VLANID、目的 MAC 地址通过掩码匹配的方式查找本地硬件 MAC 表,根据匹配的 MAC 表项中的出接口信息将报文转发出去。

3. 根据权利要求 1 所述的方法,其特征在于,所述汇聚层设备从一接口接收接入层设备上报的该接入层设备的端口允许通过的 VLAN ID、聚合 MAC 地址及掩码为:

所述汇聚层设备接收接入层设备发来的链路层发现协议 LLDP 协议报文,该报文的新增 TLV 字段中包含该接入层设备上报的端口允许通过的 VLAN ID、聚合 MAC 地址及掩码。

4. 一种部署了以太网虚拟连接 EVI 的网络中的流量转发方法,其特征在于,在各站点网络的各接入层设备上分别配置聚合 MAC 地址及聚合 MAC 掩码,其中,各接入层设备的聚合 MAC 地址互不相同;

接入层设备根据自身的聚合 MAC 地址,为自身下挂的各主机分配 MAC 地址,其中,各主机的 MAC 地址与接入层设备的聚合 MAC 地址属于同一 MAC 地址段。

5. 根据权利要求 4 所述的方法,其特征在于,所述接入层设备为自身下挂的各主机分配 MAC 地址进一步包括:

接入层设备通过数据中心桥交换 DCBX 协议报文中新增的长度类型值 TLV 字段将为各主机分配的 MAC 地址发送给对应主机。

6. 一种部署了 EVI 的网络中的汇聚层设备,其特征在于,包括:

聚合 MAC 表项学习模块:从一接口接收接入层设备上报的该接入层设备的端口允许通过的 VLAN ID、该接入层设备上配置的聚合 MAC 地址及掩码,通过软件学习 MAC 表项,将表项下发到硬件 MAC 表中,且禁止所述接口的硬件学习 MAC 表项功能,将通过软件学习到的所有 MAC 表项发送给 EVI 隧道对端的汇聚层设备;接收异地汇聚层设备发来的通过软件学习到的 MAC 表项,该 MAC 表项包括:该异地汇聚层设备下的接入层设备的端口允许通过的 VLAN

ID、聚合 MAC 地址及掩码,将该 MAC 表项下发到硬件 MAC 表中;

流量转发模块:从所述接口接收以太报文,发现所述接口的硬件学习 MAC 表项功能被禁止,则不学习该报文的源 MAC 地址,根据报文的 VLANID 和目的 MAC 地址,采用掩码匹配方式在本地硬件 MAC 表中查找匹配的 MAC 表项,根据匹配的 MAC 表项转发该报文。

7. 根据权利要求 6 所述的汇聚层设备,其特征在于,所述流量转发模块进一步包括:

当在本地硬件 MAC 表中查找匹配的 MAC 表项时,若发现匹配的 MAC 表项中的出接口信息为异地汇聚层设备的 IP 地址,则对该报文进行 EVI 封装,其中,外层 IP 头中的源 IP 地址为 EVI 隧道接口的承载接口的 IP 地址,目的 IP 地址是所述异地汇聚层设备的 IP 地址。

8. 根据权利要求 6 所述的汇聚层设备,其特征在于,所述流量转发模块进一步包括:当从 EVI 隧道接口接收到以太报文时,对该报文进行 EVI 解封装,根据解封装后的报文的 VLAN ID、目的 MAC 地址通过掩码匹配的方式查找本地硬件 MAC 表,根据匹配的 MAC 表项中的出接口信息将报文转发出去。

9. 一种部署了 EVI 的网络中的接入层设备,其特征在于,包括:聚合 MAC 上报模块,用于根据自身配置的聚合 MAC 地址,为自身下挂的各主机分配 MAC 地址,其中,各主机的 MAC 地址与本设备的聚合 MAC 地址属于同一 MAC 地址段,将本设备的端口允许通过 VLAN ID、聚合 MAC 地址及掩码上报给汇聚层设备。

## 部署了以太网虚拟连接的网络中的流量转发方法及设备

### 技术领域

[0001] 本发明涉及流量转发技术领域,具体涉及部署了以太网虚拟连接(EVI, Ethernet Virtual Interconnection)的网络中的流量转发方法及设备。

### 背景技术

[0002] EVI 是一种先进的“MAC in IP”技术,用于实现基于因特网协议(IP, Internet Protocol)核心网的二层虚拟专网(L2VPN, Layer2Virtual Private Network)技术。EVI 只是在站点的边缘设备上维护路由和转发信息,无需改变站点内部和核心网络的路由和转发信息。其中,边缘设备通常处于汇聚层交换机的位置。

[0003] 部署 EVI 后的网络由核心网络、站点网络、EVI 网络组成。其中:

[0004] 站点网络:通过一台或者多台边缘设备连接到核心网络的具有独立业务功能的二层网络,通常由单个组织管理控制,主要由主机和交换设备组成,边缘设备提供站点之间的二层交换功能。

[0005] EVI 网络:站点边缘设备之间建立的虚拟网络,提供站点网络之间的二层互联,通告连接到边缘设备的所有主机和设备的媒体访问控制(MAC, Media Access Control)地址,将多个站点互联形成更大的二层转发域。

[0006] 核心网络:主要由 IP 路由设备承载的提供站点网络之间互联的网络。

[0007] 图 1 为部署 EVI 后的网络模型示意图,如图 1 所示, EVI 网络实现虚拟局域网(VLAN, Virtual Local Area Network)在不同站点之间的扩展,主要由 EVI- 连接(Link)接口和虚拟连接组成,承载站点间扩展 VLAN 的二层流量,通过 EVI 邻居发现协议(ENDP, EVI Neighbor Discovery Protocol)来自动发现站点,简化网络的配置管理工作。通过 EVI 中间系统到中间系统(IS-IS, Intermediate System-to-Intermediate System)协议在站点之间通告主机和设备的 MAC 地址可达性信息。

[0008] 以下对 EVI 网络的运行机制进行详细说明。

[0009] 1、邻居发现

[0010] EVI 的重要特点是通过边缘设备之间通告 MAC 地址信息,用这些 MAC 地址信息来指导报文的转发。在通告 MAC 地址信息之前,边缘设备之间必须相互发现对方且形成邻居关系。

[0011] ENDP 主要用于实现邻居发现功能。在整个 EVI 邻居发现过程中有 EVI 邻居发现服务器(ENDS, EVI Neighbor Discovery Server)和 EVI 邻居发现客户端(ENDC, EVI Neighbor Discovery Client)两个实体。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, ENDS 通过接收 ENDC 的注册报文来学习 ENDC 的信息,同时通过注册应答报文向 ENDC 发布 EVI 网络实例中所有的成员 ENDC 信息, ENDC 收到应答报文后通知建立相应的 EVI Link。

[0012] 图 2 给出了 EVI 邻居自动发现过程,如图 2 所示,各个实体间的关系及 EVI 网络的建立过程如下:

[0013] (1) ENDS、ENDC 部署在站点边缘设备上。

[0014] (2) ENDC 向 ENDS 发起注册,发送注册请求报文,报文中携带本 ENDC 的 IP 地址等信息。

[0015] (3) ENDS 保存 ENDC 的信息,然后向 ENDC 发送应答报文,报文中携带所有 ENDC 的信息。

[0016] (4) ENDC 收到应答报文后,与其它每个 ENDC 建立 EVI 连接(Link)。

[0017] EVI 网络中所有成员站点的边缘设备进行相同的注册过程,最后每个成员均自动发现 EVI 网络实例内的所有成员信息。为了安全起见,可以配置 ENDP 认证功能来防止恶意的节点注册到 EVI 网络。为了保证 EVI 网络注册过程的高可靠性,可以配置多个 ENDS 进行相互备份。边缘设备发现 EVI 网络的邻居之后,就建立虚拟连接 EVI Link。

[0018] 2、MAC 地址学习

[0019] 经过邻居自动发现过程建立了边缘设备之间的 EVI 网络后,在该 EVI 网络上运行 EVI IS-IS 路由协议来通告单播 MAC 地址可达性信息,该路由协议与站点内部或核心网中的三层(L3)网络的路由协议之间相互独立,互不影响。在边缘设备上配置了隧道接口时会自动启用 EVI IS-IS 路由协议,无需额外的配置,但可以调节该路由协议的参数。

[0020] EVI 网络上的 EVI IS-IS 路由协议主要包括 EVI IS-IS 邻居协商、标签交换路径(LSP, Label Switched Path)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。EVI IS-IS 路由协议运行在点对点的虚拟连接之上,任意两个边缘设备之间为点对点的虚拟链路,边缘设备将 EVI IS-IS 问候(Hello)、LSP 报文进行单播 IP 地址封装,并进行复制后向每个远端边缘设备发送。

[0021] 3、单播流量

[0022] 一旦边缘设备之间建立了邻居关系且交互 MAC 地址可达性信息之后,就可以在站点之间转发流量。

[0023] (1) 站点内流量

[0024] 对于站点内流量,边缘设备进行 VLAN 内源 MAC 地址学习,根据目的 MAC 地址查找 MAC 表,从相应的本地接口转发。图 3 给出了站点内二层流量转发过程,如图 3 所示,主机 1 (MAC1) 发送以太报文到主机 2 (MAC2) 时,边缘设备从端口 Eth1 收到该以太报文后,学习 MAC1 到 Eth1 端口,查找 MAC 表得到 MAC2 的出端口为 Eth2,从而将以太报文从端口 Eth2 发送。

[0025] (2) 站点间流量

[0026] 对于站点之间的二层流量转发则与传统的二层转发大不相同。

[0027] 图 4 给出了站点间的二层流量转发过程,如图 4 所示,具体步骤如下:

[0028] 步骤 a:边缘设备收到以太报文之后,进行源 MAC 地址学习,根据目的 MAC 地址查找本地 MAC 表,但此时得到的出接口不是本地物理接口,而是 EVI 隧道接口,且邻居信息是异地站点通过 EVI IS-IS 路由协议发布该 MAC 路由信息的边缘设备的 IP 地址。

[0029] 步骤 b:边缘设备将原始的以太报文进行 EVI 封装,即对原始以太报文进行通用路由封装(GRE,Generic Routing Encapsulation)协议封装并添加外层 IP 头、链路层头以及校验和,其中,外层 IP 头中的源 IP 地址为 EVI 隧道接口的承载接口的 IP 地址,目的 IP 地址是通过 MAC 查找得到的异地边缘设备的 IP 地址。

[0030] 步骤 c :边缘设备将封装后的报文从 EVI 隧道接口的承载接口发送到核心网,最终到达异地站点的边缘设备。

[0031] 步骤 d :异地站点的边缘设备对报文进行解封装。

[0032] 步骤 e :异地站点的边缘设备根据解封装后的以太报文在本地进行目的 MAC 地址查找,得到的出接口为本地接口,从本地接口发送该以太报文,最终到达目的主机。

[0033] 现有方案中的缺点:

[0034] 现实应用中汇聚层交换机(也就是 EVI 网络中的边缘设备)下挂的服务器非常多,会使得汇聚层交换机上的 MAC 表项成倍增长,导致汇聚层交换机的 MAC 表规格不能满足大规模的服务器部署。以腾讯公司为例,一台汇聚层设备下挂超过 500 台服务器,每台服务器有多个 MAC 地址,这样,MAC 表项总数很容易达到汇聚层设备的 MAC 表项上限。

## 发明内容

[0035] 本发明提供部署了 EVI 的网络中的流量转发方法及设备,以减少部署了 EVI 的网络中的汇聚层设备上的 MAC 表项。

[0036] 本发明的技术方案是这样实现的:

[0037] 一种部署了以太网虚拟连接 EVI 的网络中的流量转发方法,该方法包括:

[0038] 汇聚层设备从一接口接收接入层设备上报的该接入层设备的端口允许通过的虚拟局域网标识 VLAN ID、该接入层设备上配置的聚合 MAC 地址及掩码,通过软件学习 MAC 表项,将表项下发到硬件 MAC 表中,且禁止所述接口的硬件学习 MAC 表项功能,将通过软件学习到的所有 MAC 表项发送给 EVI 隧道对端的汇聚层设备;

[0039] 所述汇聚层设备接收异地汇聚层设备发来的通过软件学习到的 MAC 表项,该 MAC 表项包括:该异地汇聚层设备下的接入层设备的端口允许通过的 VLAN ID、聚合 MAC 地址及掩码,将该 MAC 表项下发到硬件 MAC 表中;

[0040] 所述汇聚层设备从所述接口接收以太报文,发现所述接口的硬件学习 MAC 表项功能被禁止,则不学习该报文的源 MAC 地址,根据报文的 VLANID 和目的 MAC 地址,采用掩码匹配方式在本地硬件 MAC 表中查找匹配的 MAC 表项,根据匹配的 MAC 表项转发该报文。

[0041] 所述汇聚层设备根据匹配的 MAC 表项转发该报文包括:

[0042] 若所述汇聚层设备发现匹配的 MAC 表项中的出接口信息为异地汇聚层设备的 IP 地址,则对该报文进行 EVI 封装,其中,外层 IP 头中的源 IP 地址为 EVI 隧道接口的承载接口的 IP 地址,目的 IP 地址是所述异地汇聚层设备的 IP 地址;

[0043] 且,所述方法进一步包括:当所述汇聚层设备接收到异地汇聚层设备发来的以太报文时,对该报文进行 EVI 解封装,根据解封装后的报文的 VLANID、目的 MAC 地址通过掩码匹配的方式查找本地硬件 MAC 表,根据匹配的 MAC 表项中的出接口信息将报文转发出去。

[0044] 所述汇聚层设备从一接口接收接入层设备上报的该接入层设备的端口允许通过的 VLAN ID、聚合 MAC 地址及掩码为:

[0045] 所述汇聚层设备接收接入层设备发来的链路层发现协议 LLDP 协议报文,该报文的新增 TLV 字段中包含该接入层设备上报的端口允许通过的 VLAN ID、聚合 MAC 地址及掩码。

[0046] 一种部署了以太网虚拟连接 EVI 的网络中的流量转发方法,在各站点网络的各接

入层设备上分别配置聚合 MAC 地址及聚合 MAC 掩码,其中,各接入层设备的聚合 MAC 地址互不相同;

[0047] 接入层设备根据自身的聚合 MAC 地址,为自身下挂的各主机分配 MAC 地址,其中,各主机的 MAC 地址与接入层设备的聚合 MAC 地址属于同一 MAC 地址段。

[0048] 所述接入层设备为自身下挂的各主机分配 MAC 地址进一步包括:

[0049] 接入层设备通过数据中心桥交换 DCBX 协议报文中新增的长度类型值 TLV 字段将为各主机分配的 MAC 地址发送给对应主机。

[0050] 一种部署了 EVI 的网络中的汇聚层设备,包括:

[0051] 聚合 MAC 表项学习模块:从一接口接收接入层设备上报的该接入层设备的端口允许通过的 VLAN ID、该接入层设备上配置的聚合 MAC 地址及掩码,通过软件学习 MAC 表项,将表项下发到硬件 MAC 表中,且禁止所述接口的硬件学习 MAC 表项功能,将通过软件学习到的所有 MAC 表项发送给 EVI 隧道对端的汇聚层设备;接收异地汇聚层设备发来的通过软件学习到的 MAC 表项,该 MAC 表项包括:该异地汇聚层设备下的接入层设备的端口允许通过的 VLAN ID、聚合 MAC 地址及掩码,将该 MAC 表项下发到硬件 MAC 表中;

[0052] 流量转发模块:从所述接口接收以太报文,发现所述接口的硬件学习 MAC 表项功能被禁止,则不学习该报文的源 MAC 地址,根据报文的 VLANID 和目的 MAC 地址,采用掩码匹配方式在本地硬件 MAC 表中查找匹配的 MAC 表项,根据匹配的 MAC 表项转发该报文。

[0053] 所述流量转发模块进一步包括:

[0054] 当在本地硬件 MAC 表中查找匹配的 MAC 表项时,若发现匹配的 MAC 表项中的出接口信息为异地汇聚层设备的 IP 地址,则对该报文进行 EVI 封装,其中,外层 IP 头中的源 IP 地址为 EVI 隧道接口的承载接口的 IP 地址,目的 IP 地址是所述异地汇聚层设备的 IP 地址。

[0055] 所述流量转发模块进一步包括:当从 EVI 隧道接口接收到以太报文时,对该报文进行 EVI 解封装,根据解封装后的报文的 VLAN ID、目的 MAC 地址通过掩码匹配的方式查找本地硬件 MAC 表,根据匹配的 MAC 表项中的出接口信息将报文转发出去。

[0056] 一种部署了 EVI 的网络中的接入层设备,包括:聚合 MAC 上报模块,用于根据自身配置的聚合 MAC 地址,为自身下挂的各主机分配 MAC 地址,其中,各主机的 MAC 地址与本设备的聚合 MAC 地址属于同一 MAC 地址段,将本设备的端口允许通过 VLAN ID、聚合 MAC 地址及掩码上报给汇聚层设备。

[0057] 与现有技术相比,本发明通过在接入层设备上配置聚合 MAC 地址及掩码,汇聚层设备学习接入层设备的聚合 MAC 地址,减少了汇聚层设备的 MAC 表项规模,同时使得 EVI IS-IS 路由协议上传输的 MAC 表项数量大大减少,减少了系统压力。

## 附图说明

[0058] 图 1 为部署 EVI 后的网络模型示意图;

[0059] 图 2 为 EVI 邻居自动发现过程示意图;

[0060] 图 3 为部署了 EVI 的网络中站点内二层流量转发过程示意图;

[0061] 图 4 为部署了 EVI 的网络中站点间的二层流量转发过程示意图;

[0062] 图 5 为本发明实施例提供的部署了 EVI 的网络中的流量转发方法流程图;

[0063] 图 6 为本发明的一个应用实例图;

[0064] 图 7 为本发明实施例提供的部署了 EVI 的网络中的汇聚层设备的组成示意图。

### 具体实施方式

[0065] 下面结合附图及具体实施例对本发明再作进一步详细的说明。

[0066] 图 5 为本发明实施例提供的部署了 EVI 的网络中的流量转发方法流程图,如图 5 所示,其具体步骤如下:

[0067] 步骤 501:在每个站点网络的每个接入层设备上配置聚合 MAC 地址以及聚合 MAC 掩码,其中,在不同接入层设备上配置的聚合 MAC 地址互不相同。

[0068] 在接入层设备上配置了聚合 MAC 地址和聚合 MAC 掩码后,接入层设备下挂的各主机的 MAC 地址对上层的汇聚层设备而言,就是透明的了。

[0069] 步骤 502:接入层设备为自身下挂的各主机分配 MAC 地址,该分配的 MAC 地址与本接入层设备的聚合 MAC 地址属于同一 MAC 地址段,将分配的 MAC 地址发送给对应主机;主机收到该 MAC 地址,将自身的 MAC 地址更改为该接收到的 MAC 地址。

[0070] 这里,接入层设备可根据自身的聚合 MAC 地址,计算出该聚合 MAC 地址段内的 MAC 地址,将计算出的 MAC 地址顺序分配给下挂的各主机。

[0071] 其中,接入层设备可通过修改数据中心桥交换(DCBX, Data Center Bridging Exchange)协议下发为主机分配的 MAC 地址,具体实现可如下:

[0072] 在 DCBX 协议报文中增加如下类型长度值(TLV, Type-Length-Value)字段:

[0073]

TLV 名称	说明	Type	Length	Value
MAC	需要下发的 MAC 地址	100	6 字节	MAC 地址

[0074] 表 1 本发明实施例中在 DCBX 协议报文中增加的 TLV 字段

[0075] 步骤 503:接入层设备向汇聚层设备通告本设备的端口允许通过的 VLANID、聚合 MAC 地址和聚合 MAC 掩码。

[0076] 这里,接入层设备可在与汇聚层设备建立链路层发现协议(LLDP, Link Layer Discovery Protocol)邻居关系后,通过修改 LLDP 协议报文来向汇聚层设备通告本设备的端口允许通过的 VLAN ID、聚合 MAC 地址和聚合 MAC 掩码,具体实现可如下:

[0077] 在 LLDP 协议报文中增加如下 TLV 字段:

[0078]

TLV 名称	说明	Type	Length	Value
Port permit	端口允许	200	512 字节	位图方式填充,每位置 1 表示

[0079]



vlan ID	通过的 VLAN ID			vlan 允许通过, 0 表示不允许通过
Aggregate MAC	聚合 MAC 地址	201	6 字节	本设备的聚合 MAC 地址
Aggregate MAC Mask	聚合 MAC 掩码	202	6 字节	本设备使用的聚合 MAC 掩码, 用于判断 MAC 地址是否和本聚合 MAC 地址为同一 MAC 地址段

[0080] 表 2 本发明实施例中在 LLDP 协议报文中增加的 TLV 字段

[0081] 步骤 504 : 汇聚层设备从自身一接口接收接入层设备的端口允许通过的 VLAN ID、聚合 MAC 地址和聚合 MAC 掩码, 通过软件学习 MAC 表项, 并禁止本接口的硬件学习 MAC 表项功能, 将该通过软件学习到的 MAC 表项添加到软件 MAC 表中, 其中, 每个 VLAN ID 对应一条 MAC 表项, 表项内容包含 : VLAN ID、聚合 MAC 地址、聚合 MAC 掩码和本接口标识, 然后, 将 MAC 表项下发到硬件 MAC 表中, 以满足后续流量的硬件 MAC 表项匹配转发。

[0082] 本步骤中的“本接口”即, “汇聚层设备从自身一接口接收接入层设备的端口允许通过的 VLAN ID、聚合 MAC 地址和聚合 MAC 掩码”中的“接口”。

[0083] 当汇聚层设备的一个接口采用掩码方式时, 从该接口只会接收到聚合 MAC 地址段内的报文 ; 而对于未采用掩码方式的接口, 则只会接收到聚合 MAC 地址段之外的报文, 此时按照现有技术处理该报文即可, 即可以继续在硬件学习和对外发布。

[0084] 步骤 505 : 汇聚层设备通过 IS-IS 协议告知 EVI 隧道对端的汇聚层设备本端的所有 MAC 表项。

[0085] 举例说明, 设 : 汇聚层设备 SWA 通过 IS-IS 协议告知 EVI 隧道对端的汇聚层设备 SWB 本端的所有 MAC 表项, SWB 接收到各 MAC 表项后, 会将各 MAC 表项下发到硬件 MAC 表中, 其中, 各 MAC 表项的出接口信息为 SWA 的 IP 地址。

[0086] 步骤 506 : 当汇聚层设备从一接口接收到以太报文时, 若发现本接口的硬件学习 MAC 表项功能被禁止, 则不学习该报文的源 MAC 地址, 根据该报文的 VLAN ID、目的 MAC 地址通过掩码匹配的方式查找本地硬件 MAC 表。

[0087] 步骤 507 : 若汇聚层设备发现匹配的 MAC 表项的出接口为 EVI 隧道接口, 即出接口信息为异地汇聚层设备的 IP 地址, 则对原始以太报文进行 EVI 封装, 其中, 外层 IP 头中的源 IP 地址为 EVI 隧道接口的承载接口的 IP 地址, 目的 IP 地址是匹配的 MAC 表项的出接口信息中的异地汇聚层设备的 IP 地址。

[0088] 步骤 508 : 异地汇聚层设备从 EVI 隧道接口接收到以太报文, 对该报文进行 EVI 解封封装, 根据解封后的报文的 VLAN ID、目的 MAC 地址通过掩码匹配的方式查找本地硬件 MAC 表, 根据匹配的 MAC 表项中的出接口信息将报文转发出去。

[0089] 图6为本发明的一个应用实例图,以图6为例对本发明进行具体说明,具体实现如下:

[0090] 首先给出配置下发及上报过程,其中,步骤01~07针对SWA1,步骤08~14针对SWB1,步骤01~06与步骤08~12可并行进行,具体如下:

[0091] 步骤01:在部署好EVI网络后,在站点网络A的接入层设备SWA1上通过命令配置聚合MAC地址及聚合MAC掩码。

[0092] 这里,设配置的聚合MAC地址为0001-0001-0000,聚合MAC掩码为FFFF-FFFF-0000。

[0093] 步骤02:SWA1根据自身的聚合MAC地址及掩码计算要分配给下挂的各主机的MAC地址列表,每个主机对应一个端口,保存主机MAC地址与端口的对应关系。

[0094] 例如:对于主机A,设SWA1为主机A分配的MAC地址为0001-0001-0001,设主机A通过端口P1连接到SWA1,则SWA1保存P1->0001-0001-0001。

[0095] 步骤03:SWA1根据为各主机分配的MAC地址,依次组织DCBX协议报文并通过相应端口发送给各主机。

[0096] 例如:设主机A通过端口P1连接到SWA1,则主机A将从端口P1收到携带MAC TLV为0001-0001-0001的DCBX协议报文。

[0097] 步骤04:主机接收到DCBX协议报文,将报文的MAC TLV字段中携带的MAC地址设置为本网卡的MAC地址。

[0098] 需要说明的是,若SWA设备收到DCBX协议报文,不处理报文中的MAC TLV。

[0099] 步骤05:SWA1将端口允许通过的VLAN ID、聚合MAC地址及掩码通过LLDP协议报文发送给SWA。

[0100] 如图6所示,SWA1通过端口P2连接到SWA,且端口P2属于VLAN 100和200,则SWA1通过LLDP协议报文中的新增TLV字段将VLAN100和200、0001-0001-0000、FFFF-FFFF-0000通告给SWA。

[0101] 步骤06:SWA从接口Eth1接收到LLDP协议报文后,解析出SWA1的端口允许通过的VLAN ID、聚合MAC地址及掩码,禁止接口Eth1的硬件学习MAC表项功能,通过软件学习MAC表项,并将该MAC表项下发到硬件MAC表中,以便后续流量转发查询之用。

[0102] 如图6所示,本步骤中SWA学习到的软件MAC表项可如下:

[0103]

VLAN ID	MAC 地址	MAC 掩码	出接口信息
100	0001-0001-0000	FFFF-FFFF-0000	Eth1
200	0001-0001-0000	FFFF-FFFF-0000	Eth1

[0104] 表3SWA根据SWA1发来的LLDP协议报文通过软件学习到的MAC表项示例

[0105] 步骤07:SWA通过IS-IS协议告知EVI隧道对端的汇聚层设备SWB本端的所有MAC表项。

[0106] 步骤08:同步骤01,为SWB1配置聚合MAC地址为0002-0002-0000,聚合MAC掩码为FFFF-FFFF-0000。

[0107] 步骤 09~11 :同步骤 02~04,主机 B 将 0002-0002-0002 设置为本网卡的 MAC 地址。

[0108] 步骤 12~14 :同步骤 05~07,SWB 通过软件学习到如下 MAC 表项,将该表项下发到硬件 MAC 表中,并通过 IS-IS 协议通告给 SWA。

[0109] SWA、SWB 通告完毕后,各自的 MAC 表项如下 :

[0110] SWA :

[0111]

VLAN ID	MAC 地址	MAC 掩码	出接口信息
100	0001-0001-0000	FFFF-FFFF-0000	Eth1
200	0001-0001-0000	FFFF-FFFF-0000	Eth1
100	0002-0002-0000	FFFF-FFFF-0000	IP B
200	0002-0002-0000	FFFF-FFFF-0000	IP B

[0112] 表 4SWA 的硬件 MAC 表示例

[0113] SWB :

[0114]

VLAN ID	MAC 地址	MAC 掩码	出接口信息
100	0002-0002-0000	FFFF-FFFF-0000	Eth1
200	0002-0002-0000	FFFF-FFFF-0000	Eth1
100	0001-0001-0000	FFFF-FFFF-0000	IP A
200	0001-0001-0000	FFFF-FFFF-0000	IP A

[0115] 表 5SWB 的硬件 MAC 表示例

[0116] 可见,SWA、SWB 上的硬件 MAC 表项比采用 MAC 聚合前会大幅减少。

[0117] 然后给出流量转发过程,具体步骤如下 :

[0118] 步骤 01 :SWA 从接口 Eth1 接收到以太报文,发现接口 Eth1 的硬件学习 MAC 表项功能被禁止,则不学习该报文的源 MAC 地址,设报文中的 VLANID 为 VLAN 100,目的 MAC 地址为 0002-0002-0002,根据报文中的 VLANID、目的 MAC 地址通过掩码匹配的方式查找本地硬件 MAC 表,查找到的匹配表项为 :

[0119]

100	0002-0002-0000	FFFF-FFFF-0000	IP B
-----	----------------	----------------	------

[0120] 表 6SWA 查找到的硬件 MAC 表项示例

[0121] 可见,此时得到的出接口不是本地物理接口,而是 EVI 隧道接口,其表示为异地汇聚层设备的 IP 地址 :IP B。

[0122] 步骤 02 :SWA 对原始以太报文进行 EVI 封装,其中,外层 IP 头中的源 IP 地址为 EVI

隧道接口的承载接口的 IP 地址,目的 IP 地址是 IP B。

[0123] 步骤 03 :SWA 将封装后的报文从 EVI 隧道接口的承载接口发送到核心网,最终到达异地站点的边缘设备即 SWB。

[0124] 步骤 04 :SWB 对报文进行 EVI 解封装。

[0125] 步骤 05 :SWB 根据解封装后的以太报文,根据报文的 VLAN ID :VLAN100、目的 MAC 地址 :0002-0002-0002,通过掩码匹配的方式查找本地硬件 MAC 表,查找到的匹配表项为 :

[0126]

100	0002-0002-0000	FFFF-FFFF-0000	Eth1
-----	----------------	----------------	------

[0127] 表 7SWB 查找到的硬件 MAC 表项示例

[0128] 可见,出接口为本地接口 Eth1。

[0129] 步骤 06 :SWB 从接口 Eth1 发送该以太报文,该报文经过 SWB1 最终到达目的主机 :主机 B。

[0130] 图 7 为本发明实施例提供的部署了 EVI 的网络中的汇聚层设备的组成示意图,如图 7 所示,其主要包括 :聚合 MAC 表项学习模块 71、硬件 MAC 表项存储模块 72 和流量转发模块 73,其中 :

[0131] 聚合 MAC 表项学习模块 71 :从一接口接收接入层设备上报的该接入层设备的端口允许通过的 VLAN ID、该接入层设备上配置的聚合 MAC 地址及掩码,在软件 MAC 表中增加对应 MAC 表项,其中,每个 VLAN ID 对应一条 MAC 表项,表项内容包含 :VLAN ID、聚合 MAC 地址、聚合 MAC 掩码和接收接口标识,将 MAC 表项下发到硬件 MAC 表项存储模块 72,且禁止接收接口的硬件学习 MAC 表项功能,将学习到的所有 MAC 表项发送给 EVI 隧道对端的汇聚层设备 ;接收异地汇聚层设备发来的通过软件学习到的 MAC 表项,该 MAC 表项包括 :该异地汇聚层设备下的接入层设备的端口允许通过的 VLAN ID、聚合 MAC 地址及掩码,将该 MAC 表项下发到硬件 MAC 表中。

[0132] 硬件 MAC 表项存储模块 72 :存储 MAC 表项。

[0133] 流量转发模块 73 :从一接口接收以太报文,若发现接收接口的硬件学习 MAC 表项功能被禁止,则不学习该报文的源 MAC 地址,根据报文的 VLANID 和目的 MAC 地址,采用掩码匹配方式在硬件 MAC 表项存储模块 72 中查找匹配的 MAC 表项,根据匹配的 MAC 表项转发该报文。

[0134] 流量转发模块 73 进一步包括 :当在硬件 MAC 表项存储模块 72 中查找匹配的 MAC 表项时,若发现匹配的 MAC 表项中的出接口信息为异地汇聚层设备的 IP 地址,则对该报文进行 EVI 封装,其中,外层 IP 头中的源 IP 地址为 EVI 隧道接口的承载接口的 IP 地址,目的 IP 地址是匹配的 MAC 表项的出接口信息中的异地汇聚层设备的 IP 地址。

[0135] 流量转发模块 73 进一步包括 :当从 EVI 隧道接口接收到以太报文时,对该报文进行 EVI 解封装,根据解封装后的报文的 VLAN ID、目的 MAC 地址通过掩码匹配的方式查找本地硬件 MAC 表,根据匹配的 MAC 表项中的出接口信息将报文转发出去。

[0136] 以下给出本发明实施例提供的部署了 EVI 的网络中的接入层设备的组成,该设备包括 :聚合 MAC 上报模块,用于根据自身配置的聚合 MAC 地址,为自身下挂的各主机分配 MAC 地址,其中,各主机的 MAC 地址与本设备的聚合 MAC 地址属于同一 MAC 地址段,将本设备

的端口允许通过 VLAN ID、聚合 MAC 地址及掩码上报给汇聚层设备。

[0137] 以上所述仅为本发明的较佳实施例而已,并不用以限制本发明,凡在本发明的精神和原则之内,所做的任何修改、等同替换、改进等,均应包含在本发明保护的范围之内。

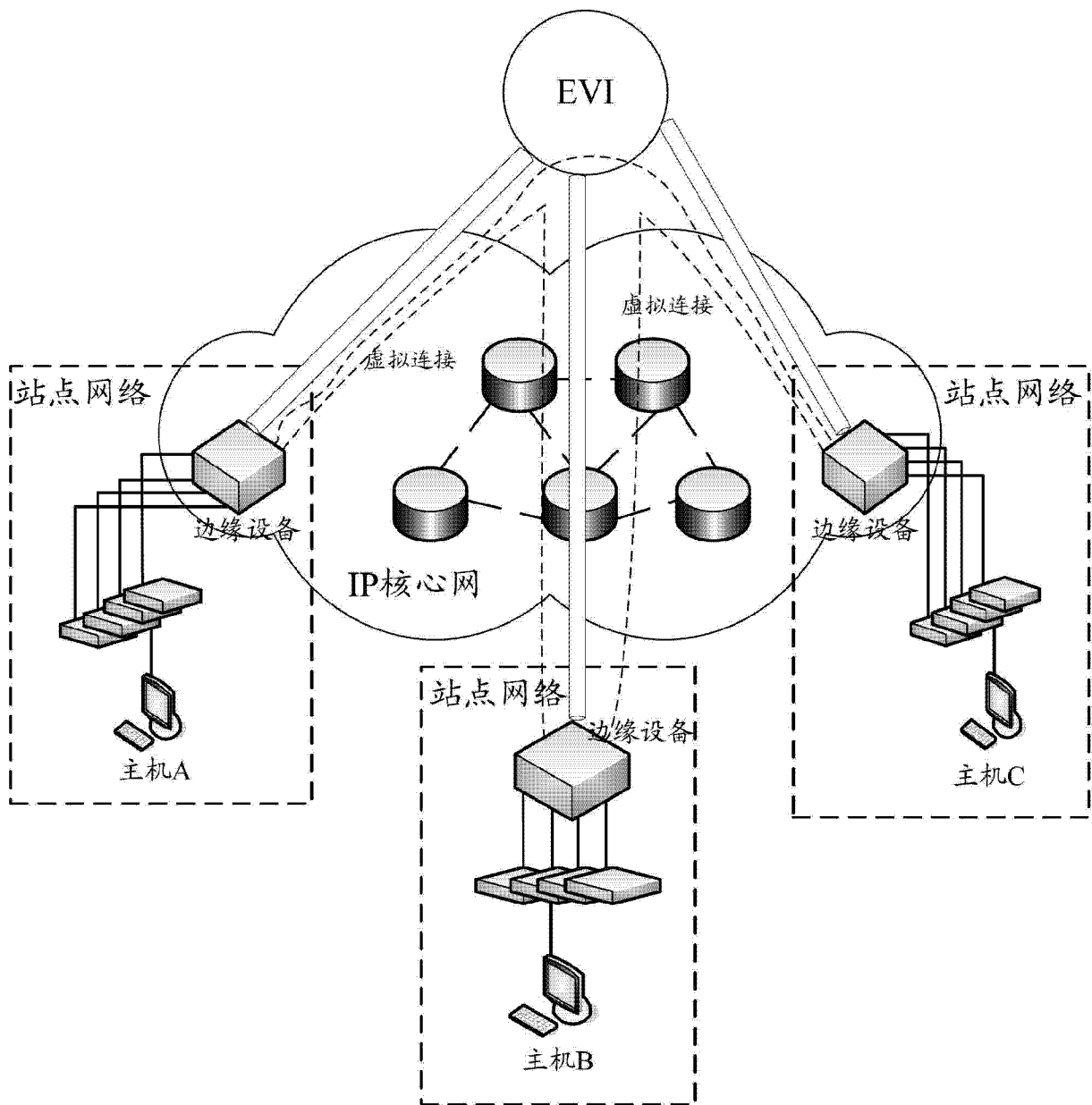


图 1

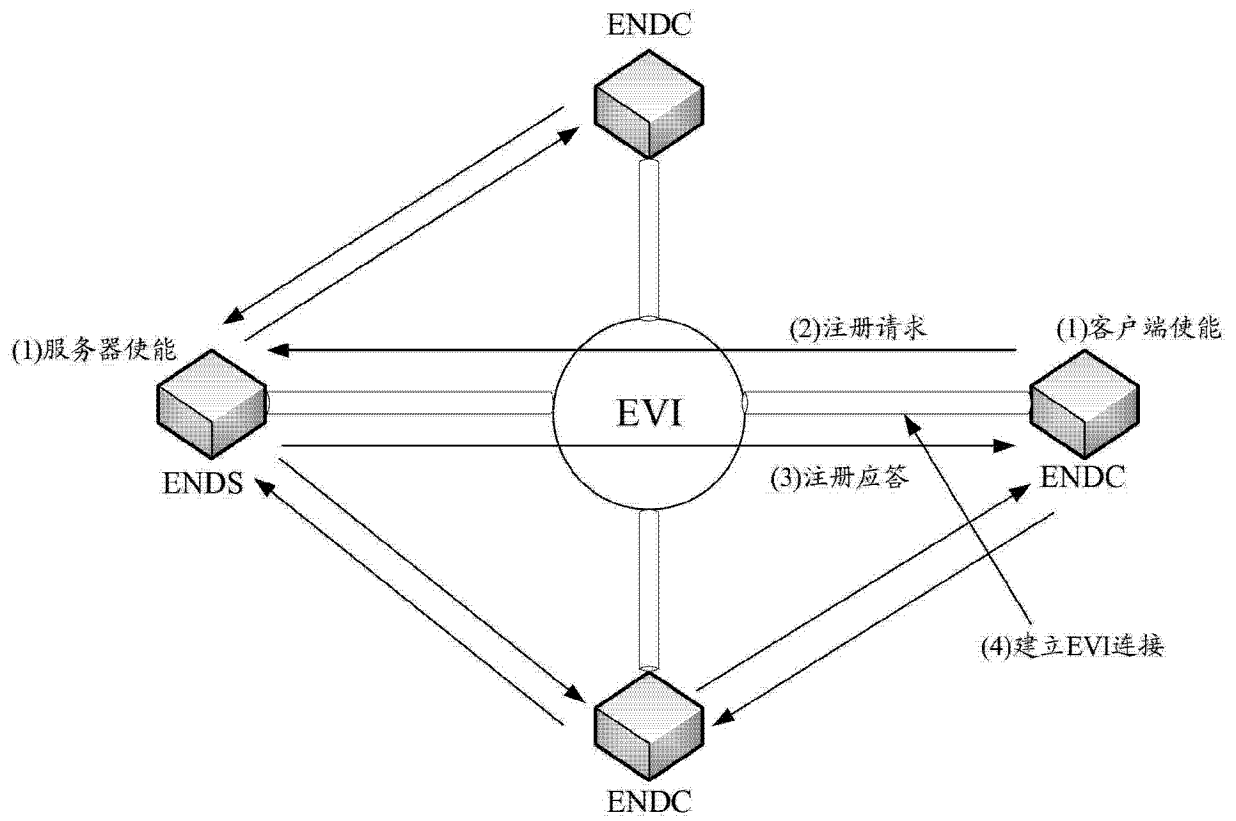


图 2

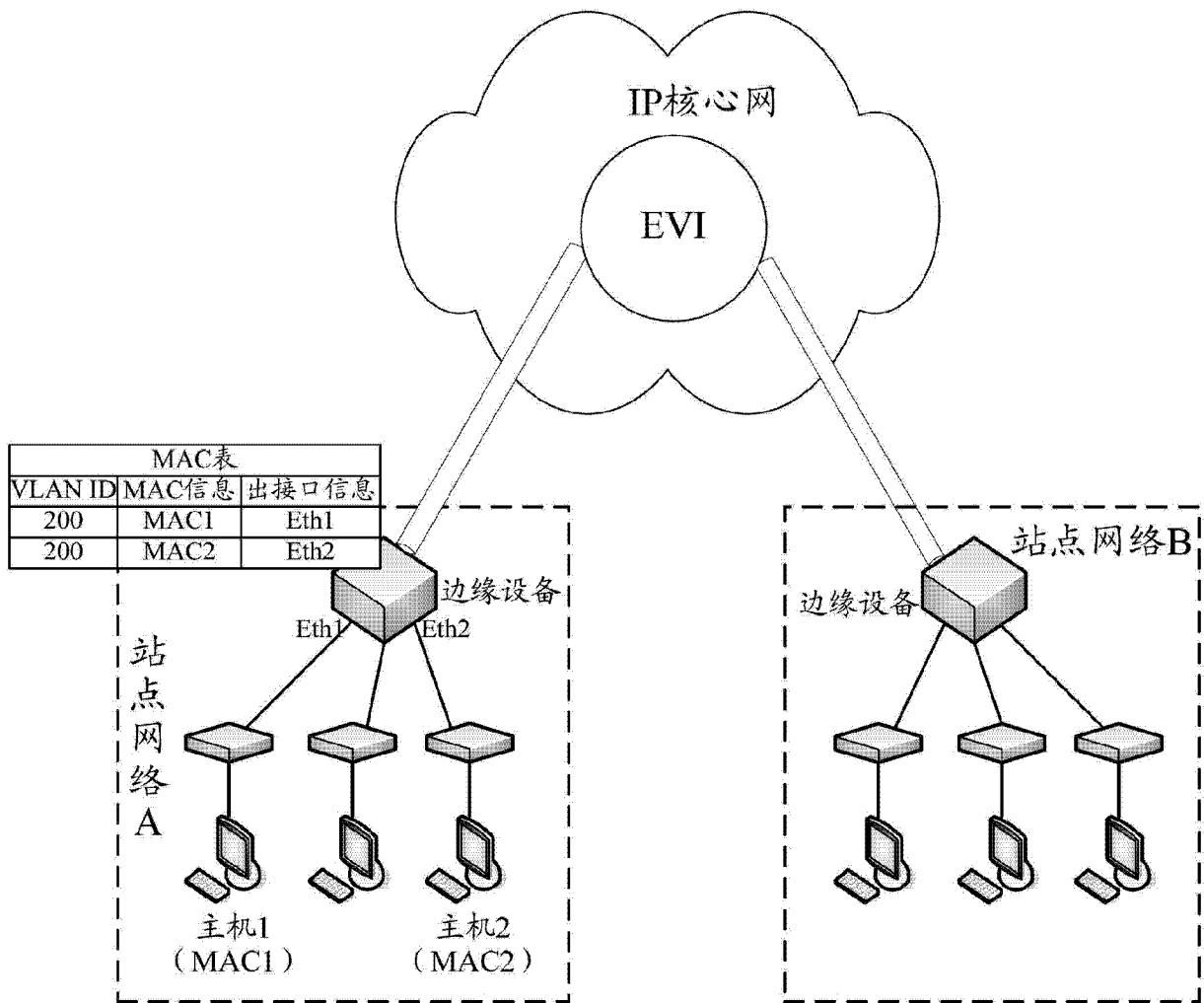


图 3



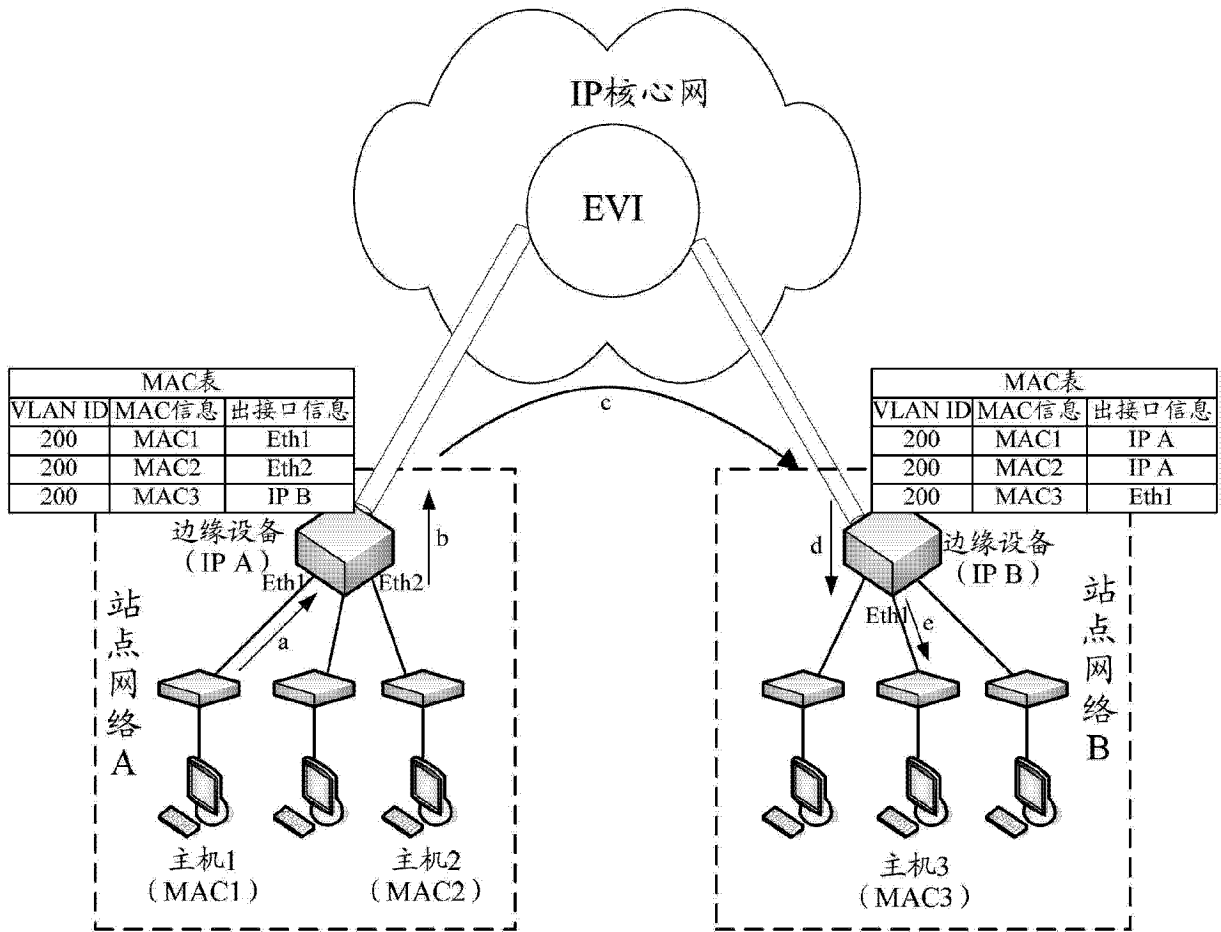


图 4

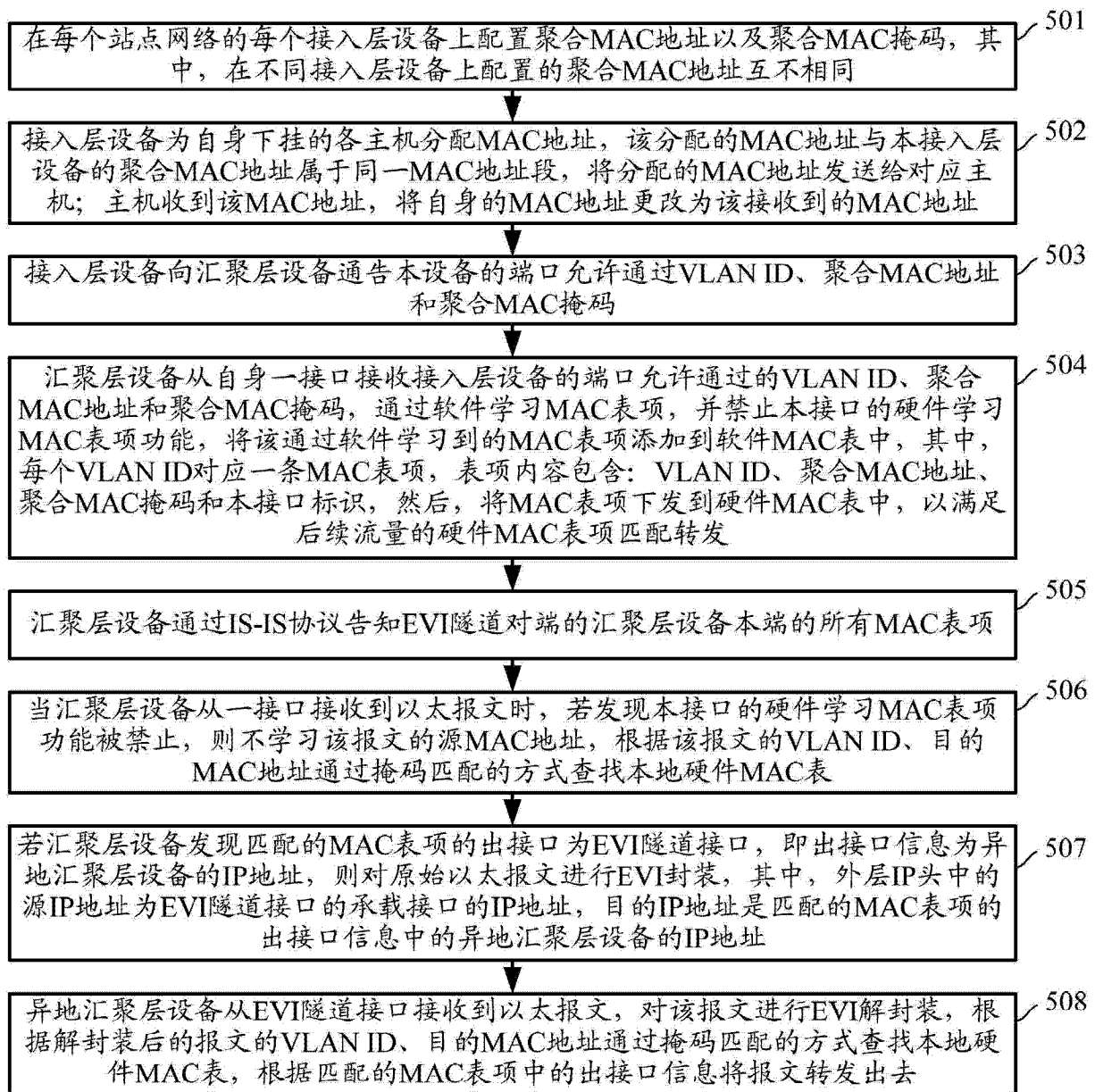


图 5

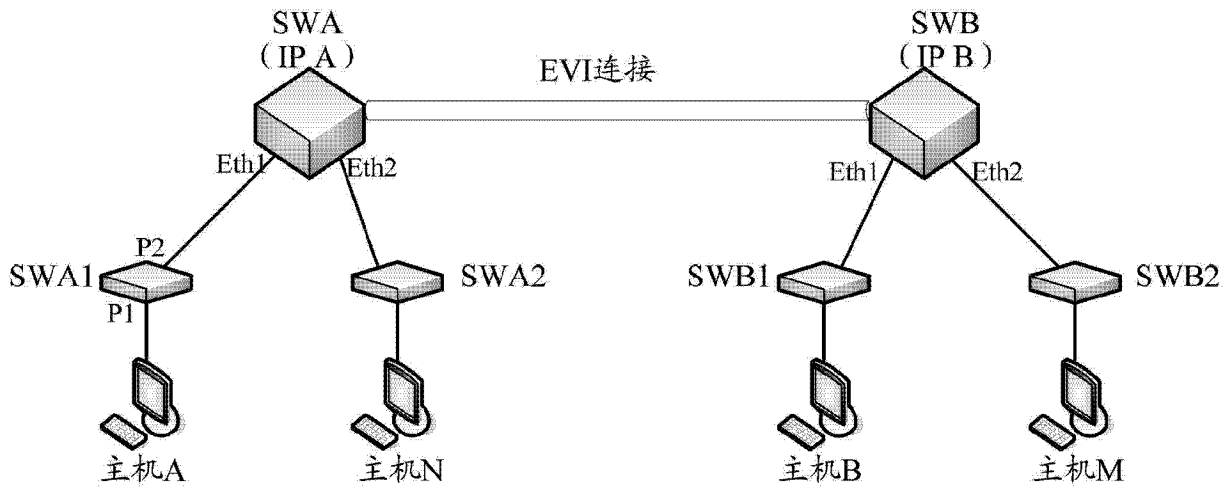


图 6

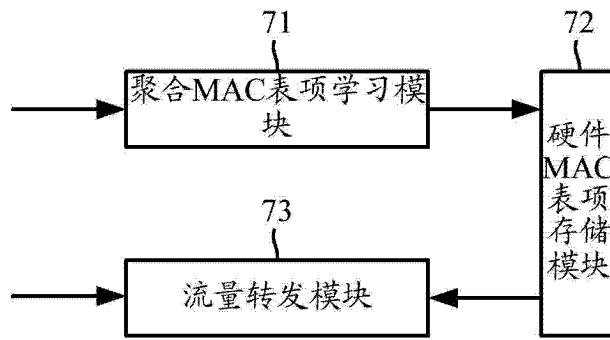


图 7