



(12) 发明专利

(10) 授权公告号 CN 102316030 B

(45) 授权公告日 2014. 04. 09

(21) 申请号 201110257124. 4

CN 101702665 A, 2010. 05. 05, 全文.

(22) 申请日 2011. 09. 01

US 20090274156 A1, 2009. 11. 05, 全文.

(73) 专利权人 杭州华三通信技术有限公司

审查员 谭美玲

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 周万

(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 12/715(2013. 01)

(56) 对比文件

CN 1870588 A, 2006. 11. 29, 说明书第 1 页第 4、第 2 页第 16 行 - 第 4 页第 22 行、第 6 页第 6 段.

US 20080095155 A1, 2008. 04. 24, 全文.

CN 101272403 A, 2008. 09. 24, 全文.

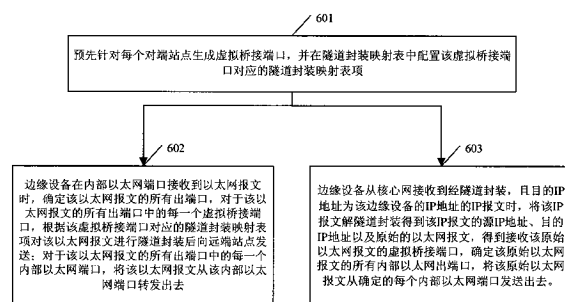
权利要求书5页 说明书13页 附图4页

(54) 发明名称

一种实现数据中心二层互联的方法和装置

(57) 摘要

本发明提供了一种实现数据中心二层互联的方法和装置,该方法中,边缘设备针对每个远端站点生成一个虚拟桥接端口,并配置相应的隧道封装映射表项;在发送以太网报文到远端站点时,根据该远端站点对应的隧道封装映射表项,将以以太网报文进行隧道封装后发送到核心网;在接收到核心网的 IP 报文时,将 IP 报文解封装得到原始以太网报文在站点内转发.应用本发明,不需要核心网支持组播。



1. 一种实现数据中心二层互联的方法,其特征在于,预先针对每个远端站点生成一个虚拟桥接端口,并在隧道封装映射表中配置该虚拟桥接端口对应的隧道封装映射表项;所述隧道封装映射表项包括:虚拟桥接端口、远端站点的边缘设备的 IP 地址、本地站点的边缘设备的 IP 地址,该方法包括:

边缘设备在内部以太网端口接收到以太网报文时,如果该以太网报文为单播报文,则根据该以太网报文的 MAC 地址在 MAC 地址映射表中查找对应的出端口,若未查找到,则确定该以太网报文的出端口为所有其它内部以太网端口;如果该以太网报文为广播报文,则确定该以太网报文的出端口为所有虚拟桥接端口和所有其它内部以太网端口;如果该以太网报文为组播报文,则根据该以太网报文的 MAC 查找该以太网报文对应的组播组得到所有出端口,若未查找到,则确定该以太网报文的出端口为所有其它内部以太网接口;对于该以太网报文的所有出端口中的每一个虚拟桥接端口,根据该虚拟桥接端口对应的隧道封装映射表项对该以太网报文进行隧道封装后向远端站点发送;对于该以太网报文的所有出端口中的每一个内部以太网端口,将该以太网报文从该内部以太网端口转发出去;

边缘设备从核心网接收到经隧道封装,且目的 IP 地址为该边缘设备的 IP 地址的 IP 报文时,将该 IP 报文解隧道封装得到该 IP 报文的源 IP 地址、目的 IP 地址以及原始的以太网报文,得到接收该原始以太网报文的虚拟桥接端口,确定该原始以太网报文的所有内部以太网出端口,将该原始以太网报文从确定的每个内部以太网出端口发送出去。

2. 如权利要求 1 所述的实现数据中心二层互联的方法,其特征在于,在边缘设备上设置有水平分割属性;所述边缘设备将 IP 报文解封得到原始以太网报文,根据所述水平分割属性,确定该原始以太网报文的出端口仅限于内部以太网出端口。

3. 如权利要求 2 所述的实现数据中心二层互联的方法,其特征在于,所述水平分割属性是通过端口 Source ID 实现的,其中每个内部以太网端口设置有不同的源标识 Source ID 值,所有虚拟桥接端口设置相同的 Source ID 值;且虚拟桥接端口的 Source ID 值与每个内部以太网端口的 Source ID 值均不相同;

所述确定该原始以太网报文的所有内部以太网出端口的方法为:将该原始以太网报文的所有出端口中 Source ID 值与所述虚拟桥接端口的 Source ID 值不同出端口确定为该原始以太网报文的内部以太网出端口。

4. 如权利要求 2 所述的实现数据中心二层互联的方法,其特征在于,所述边缘设备在内部以太网端口接收到以太网报文之后,确定该以太网报文的所有出端口之前,进一步包括:进行 MAC 地址学习,并将学习到的新的 MAC 地址信息通告给每个远端站点的边缘设备。

5. 如权利要求 4 所述的实现数据中心二层互联的方法,其特征在于,所述将学习到的新的 MAC 地址信息通告给每个远端站点的边缘设备的方法为:将学习到的新的 MAC 地址信息携带在 IGP 协议报文中,根据该远端站点对应的隧道封装映射表项,将该 IGP 协议报文进行隧道封装后发送到核心网;

所述得到接收该原始以太网报文的虚拟桥接端口之后,进一步包括:判断该原始以太网报文的协议类型,如果是 IGP 协议报文,则将该 IGP 协议报文中携带的远端站点学习到的 MAC 地址信息以及接收该原始以太网报文的虚拟桥接端口分别作为 MAC 地址映射表项的 MAC 地址信息、出端口写入 MAC 地址映射表。

6. 如权利要求 5 所述的实现数据中心二层互联的方法,其特征在于,所述得到接收该

原始以太网报文的虚拟桥接端口之后,确定该原始以太网报文的所有内部以太网出端口之前,进一步包括:学习该原始以太网报文的源 MAC 地址信息到接收该原始以太网报文的虚拟桥接端口。

7. 如权利要求 6 所述的实现数据中心二层互联的方法,其特征在于,确定将 IP 报文解隧道封装后得到的在虚拟桥接端口接收到的原始以太网报文的所有内部以太网出端口的方法为:

如果该原始以太网报文为单播报文,则根据该原始以太网报文的的目的 MAC 地址在 MAC 地址映射表中查找对应的出端口,若未查找到,则确定该原始以太网报文的出端口为所有内部以太网端口和所有其它虚拟桥接端口;

如果该原始以太网报文是广播报文,则确定该原始以太网报文的出端口为所有内部以太网端口和所有其它虚拟桥接端口;

如果该原始以太网报文为组播报文,则根据该原始以太网报文的的目的 MAC 查找该原始以太网报文对应的组播组得到所有出端口,若未查找到,则确定该原始以太网报文的出端口为所有内部以太网端口和所有其它虚拟桥接端口。

8. 如权利要求 6 所述的实现数据中心二层互联的方法,其特征在于,针对每个远端站点生成一个虚拟桥接端口时,进一步在该虚拟桥接端口上配置 VLAN;

边缘设备在内部以太网端口接收到以太网报文时,如果该以太网报文为单播报文,则根据该以太网报文的的目的 MAC 地址和该以太网报文所属的 VLAN 在 MAC 地址映射表中查找对应的出端口,若未查找到,则确定该以太网报文的出端口为该以太网报文所属 VLAN 中的所有其它内部以太网端口;如果该以太网报文为广播报文,则确定该以太网报文的出端口为该以太网报文所属 VLAN 中的所有虚拟桥接端口和所有其它内部以太网端口;如果该以太网报文为组播报文,则根据该以太网报文的的目的 MAC 和该以太网报文所属的 VLAN 查找该以太网报文对应的组播组得到所有出端口,若未查找到对应的组播组,则确定该以太网报文的出端口为该以太网报文所属 VLAN 中的所有其它内部以太网接口。

9. 如权利要求 8 所述的实现数据中心二层互联的方法,其特征在于,确定将 IP 报文解隧道封装后得到的在虚拟桥接端口接收到的原始以太网报文的所有出端口的方法为:

如果该原始以太网报文为单播报文,则根据该原始以太网报文的的目的 MAC 地址该原始以太网报文所属 VLAN 在 MAC 地址映射表中查找对应的出端口,若未查找到,则确定该原始以太网报文的出端口为该原始以太网报文所属 VLAN 中的所有内部以太网端口和所有其它虚拟桥接端口;

如果该原始以太网报文是广播报文,则确定该原始以太网报文的出端口为该原始以太网报文所属 VLAN 中的所有内部以太网端口和所有其它虚拟桥接端口;

如果该原始以太网报文为组播报文,则根据该原始以太网报文的的目的 MAC 和该以太网报文所属的 VLAN 查找该原始以太网报文对应的组播组得到所有出端口,若未查找到,则确定该原始以太网报文的出端口为该以太网报文所属 VLAN 中的所有内部以太网端口和所有其它虚拟桥接端口。

10. 如权利要求 7 或 9 任一权项所述的实现数据中心二层互联的方法,其特征在于,在边缘设备上运行组播侦听协议;

边缘设备确定在内部以太网接口接收到的以太网报文为组播报文后,进一步包括:判

断该以太网报文是否是组播查询报文或组播加入报文,如果是组播查询报文,则将接收该组播查询报文的内部以太网端口设置为该组播查询报文对应的组播组的路由器接口,并确定该组播查询报文的出端口为所有其它内部以太网端口;如果是组播加入报文,则将接收该组播加入报文的内部以太网端口加入到该组播加入报文对应的组播组的出接口列表,并确定该组播加入报文的出端口为该组播加入报文对应的组播组的路由器接口;

边缘设备将 IP 报文解隧道封装得到原始以太网报文,并确定该原始以太网报文为组播报文后,进一步包括:判断该原始以太网报文是否是组播查询报文或组播加入报文,如果是组播查询报文,则将该组播查询报文对应的虚拟桥接端口设置为该组播查询报文对应的组播组的路由器接口,并确定该组播查询报文的出端口为所有内部以太网端口;如果是组播加入报文,则将该组播加入报文对应的虚拟桥接端口加入到该组播加入报文对应的组播组的出接口列表,并确定该组播加入报文的出端口为该组播加入报文对应的组播组的路由器接口。

11. 一种边缘设备,应用于数据中心二层互联,该边缘设备包括:配置单元、第一报文收发单元、控制单元、封装/解封装单元、第二报文收发单元;

所述配置单元,用于预先针对每个远端站点生成一个虚拟桥接端口,并在隧道封装映射表中配置该虚拟桥接端口对应的隧道封装映射表项;所述隧道封装映射表项包括:虚拟桥接端口、远端站点的边缘设备的 IP 地址、本地站点的边缘设备的 IP 地址;

所述第一报文收发单元,用于在内部以太网端口接收以太网报文,并将该以太网报文发送到控制单元;用于在控制单元接收到第一报文收发单元发送的以太网报文,并确定该以太网报文的所有出端口后,对于该以太网报文的所有出端口中的每一个内部以太网端口,将该以太网报文从该内部以太网端口转发出去;用于在控制单元接收到封装/解封装单元发送的原始以太网报文,并确定该原始以太网报文的所有内部以太网出端口后,将该原始以太网报文从确定的每个内部以太网出端口发送出去;

所述控制单元,用于接收第一报文收发单元发来的以太网报文,如果该以太网报文为单播报文,则根据该以太网报文的 MAC 地址在 MAC 地址映射表中查找对应的出端口,若未查找到,则确定该以太网报文的出端口为所有其它内部以太网端口;如果该以太网报文为广播报文,则确定该以太网报文的出端口为所有虚拟桥接端口和所有其它内部以太网端口;如果该以太网报文为组播报文,则根据该以太网报文的 MAC 查找该以太网报文对应的组播组得到所有出端口,若未查找到对应的组播组,则确定该以太网报文的出端口为所有其它内部以太网接口;用于接收封装/解封装单元发来的原始以太网报文,得到接收该原始以太网报文的虚拟桥接端口,确定该原始以太网报文的所有内部以太网出端口;

所述封装/解封装单元,用于在控制单元接收到第一报文收发单元发送的以太网报文,并确定该以太网报文的所有出端口后,对于该以太网报文的所有出端口中的每一个虚拟桥接端口,根据该虚拟桥接端口对应的隧道封装映射表项,对该以太网报文进行隧道封装,并将经隧道封装后的以太网报文发送到第二报文收发单元;用于接收第二报文收发单元发来的经隧道封装,且目的 IP 地址为该边缘设备的 IP 地址的 IP 报文,将所述 IP 报文解隧道封装得到该 IP 报文的源 IP 地址、目的 IP 地址以及原始的以太网报文,将该原始以太网报文发送到控制单元;

所述第二报文收发单元,用于接收封装/解封装单元发来的经隧道封装的以太网报

文,并将该经隧道封装的以太网报文向远端站点发送;用于从核心网接收所述 IP 报文,将所述 IP 报文发送到封装/解封装单元。

12. 如权利要求 11 所述的边缘设备,其特征在于,所述配置单元,用于设置该边缘设备的水平分割属性;

所述控制单元在接收到封装/解封装单元发来的原始以太网报文之后,根据所述水平分割属性,确定该原始以太网报文的出端口仅限于所有内部以太网出端口。

13. 如权利要求 12 所述的边缘设备,其特征在于,所述配置单元设置实现该边缘设备的水平分割属性是通过端口 Source ID 实现的,其中,每个内部以太网端口设置有不同的源标识 Source ID 值,所有虚拟桥接端口设置相同的 Source ID 值;且虚拟桥接端口的 Source ID 值与每个内部以太网端口的 Source ID 值均不相同;

所述控制单元在确定该原始以太网报文的所有内部以太网出端口时,用于:将该原始以太网报文的所有出端口中 Source ID 值与所述虚拟桥接端口的 Source ID 值不同出端口确定为该原始以太网报文的内部以太网出端口。

14. 如权利要求 12 所述的边缘设备,其特征在于,所述的配置单元,用于存储 MAC 地址映射表;

所述控制单元接收到第一报文收发单元发来的以太网报文之后,确定该以太网报文的所有出端口之前,进一步用于:进行 MAC 地址学习,并将学习到的新的 MAC 地址信息携带在 IGP 协议报文中发送到封装/解封装单元;

所述封装/解封装单元,用于接收控制单元发来的 IGP 协议报文,针对每个远端站点,根据该远端站点对应的隧道封装映射表项,对该 IGP 协议报文进行隧道封装后发送到第二报文收发单元;

所述第二报文收发单元,用于接收到封装/解封装单元发来的经隧道封装的 IGP 协议报文后,将该经隧道封装的 IGP 协议报文发送到核心网;

所述控制单元,在得到接收该原始以太网报文的虚拟桥接端口之后,进一步用于:判断该原始以太网报文的协议类型,如果是 IGP 协议报文,则将该 IGP 协议报文中携带的远端站点学习到的 MAC 地址信息以及接收该原始以太网报文的虚拟桥接端口分别作为 MAC 地址映射表项的 MAC 地址信息、出端口写入 MAC 地址映射表。

15. 如权利要求 14 所述的边缘设备,其特征在于,所述控制单元在得到接收该原始以太网报文的虚拟桥接端口之后,确定该原始以太网报文的所有出端口之前,进一步用于:学习该原始以太网报文的源 MAC 地址信息到接收该原始以太网报文的虚拟桥接端口。

16. 如权利要求 15 所述的边缘设备,其特征在于,所述控制单元在确定封装/解封装单元发来的原始以太网报文的所有出端口时,用于:

如果该原始以太网报文为单播报文,则根据该原始以太网报文的源 MAC 地址在 MAC 地址映射表中查找对应的出端口,若未查找到,则确定该原始以太网报文的出端口为所有内部以太网端口和所有其它虚拟桥接端口;

如果该原始以太网报文是广播报文,则确定该原始以太网报文的出端口为所有内部以太网端口和所有其它虚拟桥接端口;

如果该原始以太网报文为组播报文时,则根据该原始以太网报文的源 MAC 查找该原始以太网报文对应的组播组得到所有出端口,若未查找到,则确定该原始以太网报文的出

端口为所有内部以太网端口和所有其它虚拟桥接端口。

17. 如权利要求 15 所述的边缘设备,其特征在于,所述配置单元,用于针对每个远端站点生成一个虚拟桥接端口时,进一步在该虚拟桥接端口上配置 VLAN;

所述控制单元在接收到第一报文收发单元发来的以太网报文时,如果该以太网报文为单播报文,则根据该以太网报文的目的地 MAC 地址和该以太网报文所属的 VLAN 在 MAC 地址映射表中查找对应的出端口,若未查找到,则确定该以太网报文的出端口为该以太网报文所属 VLAN 中的所有其它内部以太网端口;如果该以太网报文为广播报文,则确定该以太网报文的出端口为该以太网报文所属 VLAN 中的所有虚拟桥接端口和所有其它内部以太网端口;如果该以太网报文为组播报文,则根据该以太网报文的目的地 MAC 和该以太网报文所属的 VLAN 查找该以太网报文对应的组播组得到所有出端口,若未查找到对应的组播组,则确定该以太网报文的出端口为该以太网报文所属 VLAN 中的所有其它内部以太网接口。

18. 如权利要求 17 所述的边缘设备,其特征在于,所述控制单元在确定封装/解封装单元发来的原始以太网报文的所有出端口时,用于:

如果该原始以太网报文为单播报文,则根据该原始以太网报文的目的地 MAC 地址该原始以太网报文所属 VLAN 在 MAC 地址映射表中查找对应的出端口,若未查找到,则确定该原始以太网报文的出端口为该原始以太网报文所属 VLAN 中的所有内部以太网端口和所有其它虚拟桥接端口;

如果该原始以太网报文是广播报文时,则确定该原始以太网报文的出端口为该原始以太网报文所属 VLAN 中的所有内部以太网端口和所有其它虚拟桥接端口;

如果该原始以太网报文为组播报文时,则根据该原始以太网报文的目的地 MAC 和该原始以太网报文所属的 VLAN 查找该原始以太网报文对应的组播组得到所有出端口,若未查找到,则确定该原始以太网报文的出端口为该原始以太网报文所属 VLAN 中的所有内部以太网端口和所有其它虚拟桥接端口。

19. 如权利要求 16 或 18 任一权项所述的边缘设备,其特征在于,在边缘设备上运行组播侦听协议;

所述控制单元确定第一报文收发单元发来的以太网报文为组播报文后,进一步用于:判断该以太网报文是否是组播查询报文或组播加入报文,如果是组播查询报文,则将接收该组播查询报文的内部以太网端口设置为该组播查询报文对应的组播组的路由器接口,并确定该组播查询报文的出端口为所有其它内部以太网端口;如果是组播加入报文,则将接收该组播加入报文的内部以太网端口加入到该组播加入报文对应的组播组的出接口列表,并确定该组播加入报文的出端口为该组播加入报文对应的组播组的路由器接口;

所述控制单元在确定封装/解封装单元发来的原始以太网报文为组播报文后,进一步用于:判断该原始以太网报文是否是组播查询报文或组播加入报文,如果是组播查询报文,则将该组播查询报文对应的虚拟桥接端口设置为该组播查询报文对应的组播组的路由器接口,并确定该组播查询报文的出端口为所有内部以太网端口;如果是组播加入报文,则将该组播加入报文对应的虚拟桥接端口加入到该组播加入报文对应的组播组的出接口列表,并确定该组播加入报文的出端口为该组播加入报文对应的组播组的路由器接口。

一种实现数据中心二层互联的方法和装置

技术领域

[0001] 本发明涉及网络通信技术领域,特别涉及一种实现数据中心二层互联的方法和装置。

背景技术

[0002] 数据中心,是各种业务的提供中心,是数据处理、数据存储和数据交换的中心。为了提供更好的服务,提供商通常在异地部署多个数据中心,实现负载分担和高可靠性,并实现虚拟机在数据中心之间的自由迁移。由于虚拟机迁移过程对用户透明,不能改变 IP 地址,所以必须在分布于异地的数据中心之间实现二层网络互联。

[0003] 数据中心互联,需要满足以下基本要求:各站点互相独立,站点内的拓扑和故障互不影响;对站点间传输数据时使用的技术与站点位置、提供商的网络无特殊要求,具有传输无关性;使用多归宿提供冗余接入,并在站点间避免流量环路,保证高可靠性;站点内拓扑结构实现灵活;运营维护简单,可快速新增和减少站点。

[0004] 目前,在公网只支持 IP 技术时,只有思科的覆盖传输虚拟化 (OTV) 技术能够实现数据中心互联。在各数据中心的边缘设备的连接到核心网的物理端口上配置 IP 地址,并关联一个或多个 OTV 虚拟接口,每个 OTV 虚拟接口可以作为二层转发表的出接口。下面结合附图对 OTV 技术进行说明。

[0005] 参见图 1,图 1 是现有技术应用 OTV 技术实现数据中心二层互联的组网示意图。其中网络 10 为核心网,其上运行 pUMRP 协议,网络 12、16 分别是数据中心站点,其上运行 cUMRP 协议,网络 14 是构建在网络 10 上的虚拟网络,其上运行 oUMRP 协议。在网络 12 中,包含边缘设备 X11,交换机 S11,路由器 R11、R12,主机 H3 通过交换机 S11 接入,主机 H12 通过路由器 R11 接入。在网络 16 中,包含边缘设备 X12,交换机 S12、S13,路由器 R14、R15,主机 H1 通过交换机 S13 接入。

[0006] 参见图 2,图 2 是图 1 中边缘设备 X11 的数据平面内部实现图。其中,E1 和 E2 分别是连接主机 H2 和连接主机 H3 的内部接口,E3 是连接核心网的外部接口,OTV 虚拟接口 O1:IPA 关联到 E3。在 X11 中存储 MAC 地址映射表(如图 2 中所示的 MAC 地址映射表),其中的 H2/E1、H3/E2 表项是通过 MAC 地址学习得到的,H1/O1:IPA 表项则是从远端 OTV 设备通告的 IGP 协议报文中得到的(如图 2 中的线条 21 所示,表示将远端站点通告的 IGP 协议报文中携带的 MAC 地址写入 MAC 地址映射表),在图 1 中,远端 OTV 设备即为边缘设备 X12。例如,当 X11 从 E1 接收 H2 发来的报文时,进行 MAC 地址学习即可得到 H2/E1 表项并写入 MAC 地址映射表(如图 2 中的线条 22 所示,表示对在内部以太网端口接收到的报文进行 MAC 地址学习),同时还将学习到的 MAC 地址映射表项通过 IGP 协议报文通告给远端 OTV 设备;同样的道理,X11 也会从远端 OTV 设备通告的 IGP 协议报文中学习得到 H1/O1:IPA 表项。根据 MAC 地址映射表,X11 可以将来自于 H2 或 H3 且目的地址为 H1 的报文进行 OTV 封装,并从 O1:IPA 关联的物理接口 E3 发送出去;将来自于 H1 且目的地址为 H2 或 H3 的报文解封装后从相应的 E2 或 E3 接口发送出去。

[0007] 在图 1 所示的组网中,如果需要支持广播,则需要预先在 X11 和 X12 上配置由核心网分配的任意源组播 (ASM/Bidir) 类型的 IP 组播组。当 X11 在内部接口接收到广播报文时,将广播报文进行 OTV 封装后发送到核心网的组播树中,由核心网组播转发到远端 OTV 设备 X12, X12 将接收到的广播报文解封装后,在其所在的数据中心网络内广播。如果需要支持组播,则需要网络 12 和 16 上运行 IGMP/MLD 侦听,在 X11 和 X12 上记录组播信息到组播转发表,在核心网中构建组播组的特定源组播 (SSM) 树,并根据该组播树组播转发数据中心之间的组播报文。

[0008] 可见,如果在应用 OTV 技术实现数据中心二层互联的组网中,核心网必须支持组播,否则,将无法实现广播转发和组播转发需求,进而也无法实现数据中心二层互联。

发明内容

[0009] 有鉴于此,本发明的目的在于提供一种实现数据中心二层互联的方法,该方法不需要核心网支持组播。

[0010] 为了达到上述目的,本发明提供了一种实现数据中心二层互联的方法,预先针对每个远端站点生成一个虚拟桥接端口,并在隧道封装映射表中配置该虚拟桥接端口对应的隧道封装映射表项;所述隧道封装映射表项包括:虚拟桥接端口、远端站点的边缘设备的 IP 地址、本地站点的边缘设备的 IP 地址,该方法包括:

[0011] 边缘设备在内部以太网端口接收到以太网报文时,确定该以太网报文的所有出端口,对于该以太网报文的所有出端口中的每一个虚拟桥接端口,根据该虚拟桥接端口对应的隧道封装映射表项对该以太网报文进行隧道封装后向远端站点发送;对于该以太网报文的所有出端口中的每一个内部以太网端口,将该以太网报文从该内部以太网端口转发出去;

[0012] 边缘设备从核心网接收到经隧道封装,且目的 IP 地址为该边缘设备的 IP 地址的 IP 报文时,将该 IP 报文解隧道封装得到该 IP 报文的源 IP 地址、目的 IP 地址以及原始的以太网报文,得到接收该原始以太网报文的虚拟桥接端口,确定该原始以太网报文的所有内部以太网出端口,将该原始以太网报文从确定的每个内部以太网出端口发送出去。

[0013] 本发明还提供了一种实现数据中心二层互联的装置,该装置为数据中心的边缘设备,该装置包括:配置单元、第一报文收发单元、封装/解封单元、控制单元、第二报文收发单元;

[0014] 所述配置单元,用于预先针对每个远端站点生成一个虚拟桥接端口,并在隧道封装映射表中配置该虚拟桥接端口对应的隧道封装映射表项;所述隧道封装映射表项包括:虚拟桥接端口、远端站点的边缘设备的 IP 地址、本地站点的边缘设备的 IP 地址;

[0015] 所述第一报文收发单元,用于在内部以太网端口接收以太网报文,并将该以太网报文发送到控制单元;用于在控制单元接收到第一报文收发单元发送的以太网报文,并确定该以太网报文的所有出端口后,对于该以太网报文的所有出端口中的每一个内部以太网端口,将该以太网报文从该内部以太网端口转发出去;用于在控制单元接收到封装/解封单元发送的原始以太网报文,并确定该原始以太网报文的所有内部以太网出端口后,将该原始以太网报文从确定的每个内部以太网出端口发送出去;

[0016] 所述控制单元,用于接收第一报文收发单元发来的以太网报文,确定该以太网报

文的所有出端口；用于接收封装 / 解封装单元发来的原始以太网报文，得到接收该原始以太网报文的虚拟桥接端口，确定该原始以太网报文的所有内部以太网出端口；

[0017] 所述封装 / 解封装单元，用于在控制单元接收到第一报文收发单元发送的以太网报文，并确定该以太网报文的所有出端口后，对于该以太网报文的所有出端口中的每一个虚拟桥接端口，根据该虚拟桥接端口对应的隧道封装映射表项，对该以太网报文进行隧道封装，并将经隧道封装后的以太网报文发送到第二报文收发单元；用于接收第二报文收发单元发来的经隧道封装，且目的 IP 地址为该边缘设备的 IP 地址的 IP 报文，将所述 IP 报文解隧道封装得到该 IP 报文的源 IP 地址、目的 IP 地址以及原始的以太网报文，将该原始以太网报文发送到控制单元；

[0018] 所述第二报文收发单元，用于接收封装 / 解封装单元发来的经隧道封装的以太网报文，并将该经隧道封装的以太网报文向远端站点发送；用于从核心网接收所述 IP 报文，将所述 IP 报文发送到封装 / 解封装单元。

[0019] 由上面的技术方案可知，本发明通过在本地边缘设备上针对每个远端站点生成一个虚拟桥接端口，通过该虚拟桥接端口建立与该远端站点的隧道；在发送以太网报文到远端站点时，根据该远端站点对应的隧道封装映射表项将该以太网报文进行隧道封装后向该远端站点发送；在接收到来自核心网的，经过隧道封装的 IP 报文时，将该 IP 报文解隧道封装得到原始的以太网报文，将该原始以太网报文从内部以太网端口转发。本发明中，站点之间通过隧道进行报文传输，不需要核心网支持组播。

附图说明

[0020] 图 1 是现有技术应用 OTV 技术实现数据中心二层互联的组网示意图；

[0021] 图 2 是图 1 中边缘设备 X11 的数据平面内部实现图；

[0022] 图 3 是数据中心站点二层互联的组网示意图；

[0023] 图 4 是本发明实施例边缘设备功能结构示意图；

[0024] 图 5 是本发明数据中心站点二层互联的组网示意图；

[0025] 图 6 是本发明实施例实现数据中心二层互联的方法流程图；

[0026] 图 7 是本发明实施例边缘设备的结构示意图。

具体实施方式

[0027] 为了使本发明的目的、技术方案及优点更加清楚明白，下面结合附图并举实施例，对本发明的技术方案进行详细说明。

[0028] 数据中心站点的二层互联，主要是通过各数据中心站点的边缘设备实现。参见图 3，图 3 是数据中心站点二层互联的组网示意图，只画出了各站点的边缘设备，站点内部拓扑不是本发明的重点，因此未在图中画出。图 3 中，站点 A、B、C、D、E 分别通过各自的边缘设备 1、2、3、4、5 接入核心网。这里，需要说明的是，本申请文件中所述的站点，也即是数据中心站点。

[0029] 本发明的一个实施方案中，本地站点的边缘设备针对每个远端站点生成虚拟桥接端口。以图 3 为例，站点 A 的边缘设备 1 分别针对远端站点 B、C、D、E 生成四个虚拟桥接端口；站点 B 的边缘设备 2 分别针对远端站点 A、C、D、E 生成四个虚拟桥接端口，依次类推。任

意两个站点之间通过针对端站点生成的虚拟桥接端口建立点到点的隧道,在隧道进行报文传输,实现数据中心站点间的二层互联。

[0030] 参见图 4,图 4 是本发明实施例边缘设备功能结构示意图,将边缘设备的功能分为 3 个部分,分别为 P(Provider)、O(Over)、C(Customer)。其中,P 主要用作虚拟的核心网设备,提供边缘设备连接到核心网的通道;O,主要用作虚拟的边缘设备,提供站点间互联的通道;C 主要用作虚拟的私网设备,提供内部设备联结到边缘设备的通道。每部分功能相当于一台逻辑设备,由管理平面 MP、控制平面 CP 和数据平面 DP 组成。

[0031] 这里,管理平面不是本发明的重点,不再详述。下面分别对控制平面和数据平面进行详细说明:

[0032] 边缘设备的控制平面分为 P-CP,O-CP,C-CP 三个部分。其中,P-CP:主要用于控制与公网交互协议报文,例如,控制与公网交互三层路由信息。O-CP,主要用于控制与指定的远端站点建立点对点隧道,进行 MAC 地址信息交互,感知远端站点的隧道是否可达。C-CP,主要用于控制站点间二层协议交互信息,例如 IGMP Snooping;用于将单播 MAC 地址信息发布给 O-CP,再由 O-CP 传播给其它站点。

[0033] 边缘设备的数据平面分为 P-DP,O-DP,C-DP 三个部分。其中,C-DP 主要用于对从站点内部接收到的报文进行二层转发处理;用于对从 O-DP 进入的报文进行二层转发处理;用于触发 MAC 地址学习,并将学习到的 MAC 地址信息送入 O-DP。O-DP 主要用于对报文进行隧道封装和解隧道封装;用于针对从 C-DP 进入的 MAC 地址信息,触发与远端站点的 MAC 地址信息交互。P-DP 主要用于与公网之间的三层协议交互。

[0034] 参见图 5,图 5 是本发明数据中心站点二层互联的组网示意图,包括站点 A、B、C,其中,站点 A、B、C 接入核心网的 IP 地址分别为 ipA、ipB、ipC。图中还给出了站点 C 中边缘设备 3 的内部结构图,边缘设备 3 主要包含处理器、连接站点内主机 H3 的内部以太网端口 E1、连接站点内主机 H4 的内部以太网端口 E2、针对远端站点 A 生成的虚拟桥接端口 vEth1、针对远端站点 B 生成的虚拟桥接端口 vEth2。其中,处理器,主要用于对从内部以太网端口接收到的以太网报文和从核心网的接收到的报文进行处理。按照处理器对报文的处理功能,又分为两部分功能模块:隧道封装/解封装功能模块、二层交换处理功能模块。

[0035] 为了实现在站点之间通过隧道进行报文传输,边缘设备需要维护一个隧道封装映射表(如图 5 中所示的隧道封装映射表),在隧道封装映射表中为针对每个远端站点生成的虚拟桥接端口配置对应的隧道封装映射表项。其中,每个隧道封装映射表项包含针对远端站点生成的虚拟桥接端口、远端站点的边缘设备的 IP 地址,本地站点的边缘设备的 IP 地址。边缘设备在向远端站点发送报文时,先根据该远端站点对应的隧道封装映射表项,对报文进行隧道封装,再根据路由表将封装后的报文发送到核心网;边缘设备接收到来自核心网的经过隧道封装,且目的 IP 地址为本机 IP 地址的 IP 报文时,先将报文解隧道封装得到该 IP 报文的源 IP 地址、目的 IP 地址以及原始的以太网报文,得到接收该原始以太网报文的虚拟桥接端口;最后确定该原始以太网报文的内部以太网出端口,并从确定的每个内部以太网出端口转发。这里,根据 IP 报文的源 IP 地址和目的 IP 地址查找隧道封装映射表,确定接收该原始以太网报文的虚拟桥接端口。

[0036] 此外,为了实现报文的二层转发,边缘设备还需要维护一个 MAC 地址映射表(如图 5 中所示的 MAC 地址映射表),该 MAC 地址映射表与现有技术的 MAC 地址映射表相同,可以

通过对接收到的以太网报文进行 MAC 地址学习得到。不同之处在于：本发明中的 MAC 地址映射表中，还包含虚拟桥接端口对应的 MAC 地址映射表项，也即包含以虚拟桥接端口为出端口的 MAC 地址映射表项。获得虚拟桥接端口对应的 MAC 地址映射表项的方法有两种。一种是按照现有技术的方法，对在虚拟桥接端口接收到的以太网报文进行 MAC 地址学习。另外一种，边缘设备将学习到的站点内部的新的 MAC 地址信息通过隧道通过给每个远端站点；边缘设备在接收到远端站点通告的新的 MAC 地址信息时，将远端站点通告的新的 MAC 地址信息写入 MAC 地址映射表。

[0037] 上面是对本发明的一个实施方案的简单介绍，下面对该实施方案的具体实现过程进行详细说明。

[0038] 参加图 6，图 6 是本发明实施例实现数据中心二层互联的方法流程图，包括以下步骤：

[0039] 步骤 601、预先针对每个远端站点生成虚拟桥接端口，并在隧道封装映射表中配置该虚拟桥接端口对应的隧道封装映射表项。

[0040] 这里，虚拟桥接端口是针对每个远端站点生成的虚拟以太网端口，没有物理实体，但具有与物理以太网端口相同的功能。任意两个站点之间通过各自边缘设备上生成的针对对端站点的虚拟桥接端口建立点对点的隧道，实现数据中心站点之间的二层互联。

[0041] 虚拟桥接端口是本地站点的边缘设备针对每个远端站点生成的，本地站点的公网 IP 地址、远端站点的公网 IP 地址、以及本地站点针对该远端站点生成的虚拟桥接端口之间具有相互对应关系。实际上，由于各站点是通过边缘设备接入核心网的，边缘设备接入核心网的 IP 地址就是该边缘设备所在站点的公网 IP 地址，也即该边缘设备的 IP 地址。本发明实施例中，在隧道封装映射表中存储该对应关系，每个对应关系称为一个隧道封装映射表项，包括虚拟桥接端口、远端站点的边缘设备的 IP 地址、本地站点的边缘设备的 IP 地址。隧道封装映射表项用于边缘设备与远端站点之间进行通信时报文的隧道封装和解隧道封装。

[0042] 步骤 602、边缘设备在内部以太网端口接收到以太网报文时，确定该以太网报文的所有出端口，对于该以太网报文的所有出端口中的每一个虚拟桥接端口，根据该虚拟桥接端口对应的隧道封装映射表项对该以太网报文进行隧道封装后向远端站点发送；对于该以太网报文的所有出端口中的每一个内部以太网端口，将该以太网报文从该内部以太网端口发送出去。

[0043] 本步骤中，当边缘设备在内部以太网端口接收到以太网报文时，需要根据该以太网报文的的目的 MAC 确定对应的出端口，并根据对应的出端口进行报文发送。当出端口为虚拟桥接端口时，说明该报文的的目的地址为该虚拟桥接端口对应的远端站点，因此需要根据该虚拟桥接端口对应的隧道封装映射表项对该以太网报文进行隧道封装后发送出去；当出端口是内部以太网端口时，则只需按照现有技术的方法将报文从相应的内部以太网端口直接发送出去即可。

[0044] 这里，所述根据该虚拟桥接端口对应的隧道封装映射表项对该以太网报文进行隧道封装的方法为：将该虚拟桥接端口对应的隧道封装映射表项中的本地站点的边缘设备的 IP 地址、远端站点的边缘设备的 IP 地址分别作为源 IP 地址、目的 IP 地址对该以太网报文进行封装。

[0045] 步骤 603、边缘设备从核心网接收到经隧道封装，且目的 IP 地址为该边缘设备的

IP 地址的 IP 报文时,将该 IP 报文解隧道封装得到该 IP 报文的源 IP 地址、目的 IP 地址以及原始的以太网报文,根据该 IP 报文的源 IP 地址和目的 IP 地址查找隧道封装映射表,得到接收该原始以太网报文的虚拟桥接端口,确定该原始以太网报文的所有内部以太网出端口,将该原始以太网报文从确定的每个内部以太网端口发送出去。

[0046] 本步骤中,边缘设备从核心网接收到经过隧道封装,且目的 IP 地址为该边缘设备的 IP 地址的 IP 报文时,说明该 IP 报文是远端站点通过与本地站点之间建立的点对点的隧道发来的报文,需要将 IP 报文解隧道封装得到原始以太网报文,再根据该原始以太网报文的 MAC 确定报文的出端口。

[0047] 这里,所述根据该 IP 报文的源 IP 地址和目的 IP 地址查找隧道封装映射表,确定对应的虚拟桥接端口的方法为:将该 IP 报文的源 IP 地址、目的 IP 地址分别与隧道封装映射表中每个隧道封装映射表项中的远端站点的边缘设备的 IP 地址、本地站点的边缘设备的 IP 地址进行比较,如果相同,则确定对应的虚拟桥接端口为该隧道封装映射表项中的虚拟桥接端口。

[0048] 图 6 所示的本发明实施例中,边缘设备从核心网接收到经过隧道封装,且目的 IP 地址为该边缘设备的 IP 地址的 IP 报文时,其中封装的原始以太网报文是针对本地站点发送的,不允许再重新发回到核心网中,可以在边缘设备上设置水平分割属性;边缘设备将 IP 报文解封得到原始以太网报文,根据所述水平分割属性,确定该原始以太网报文的出端口仅限于内部以太网出端口。通过设置的水平分割属性,使得该原始以太网报文不会被发送回核心网。

[0049] 上述的水平分割属性,可以通过端口的源标识 (Source ID) 实现。现有技术中,每个端口都分配有一个源标识 (Source ID),在该端口接收到以太网报文时,为该以太网报文设置相同的 Source ID 值;在转发该以太网报文时,先比较该以太网报文的出端口的 Source ID 值是否与该以太网报文的 Source ID 值相同,如果相同,则不允许从该出端口。这样可以避免将该以太网报文从接收该以太网报文的端口发送,达到环路避免的目的。

[0050] 本实施例中,为了实现将来自其它站点的原始以太网报文仅在本站点内部传播的目的,通过端口的源标识 (Source ID) 实现的水平分割属性具体可以为:为每个内部以太网端口设置不同的 Source ID 值;将所有虚拟桥接端口的 Source ID 值设置为相同的值;且虚拟桥接端口的 Source ID 值与每个内部以太网端口的 Source ID 值均不相同;

[0051] 那么,所述边缘设备确定该原始以太网报文的所有内部以太网出端口的方法为:将该原始以太网报文的所有出端口中与所述虚拟桥接端口的 Source ID 值不同的出端口确定为该原始以太网报文的内部以太网出端口。

[0052] 这样,由于所有虚拟桥接端口的 Source ID 值都是相同的,从虚拟桥接端口接收到的原始以太网报文,不会再从其它的虚拟桥接端口发送出去。无论是单播报文、广播报文或组播报文,均只能在本地站点内传播。

[0053] 图 6 所示的本发明实施例中,边缘设备从内部以太网端口接收到报文时,可以进行 MAC 地址学习,如果学习到的是新的 MAC 地址信息,则还可以将学习到的新的 MAC 地址信息通告给每个远端站点。因此,步骤 602 中,边缘设备在内部以太网端口接收到以太网报文之后,确定该以太网报文的所有出端口之前,进一步包括:进行 MAC 地址学习,并将该学习到的新的 MAC 地址信息通告给每个远端站点。这里,如果该以太网报文属于某个 VLAN,则所

述 MAC 地址信息除了包括 MAC 地址外,还包括该以太网报文所属的 VLAN,此与现有技术相同。

[0054] 这里,判断学习到的源 MAC 地址是否是新的 MAC 地址的方法为:在 MAC 地址映射表查找该源 MAC 地址对应的 MAC 地址映射表项,如果查找到对应的 MAC 地址映射表项,则该源 MAC 地址不是新的 MAC 地址,否则,该源 MAC 地址是新的 MAC 地址。

[0055] 边缘设备可以通过与每个远端站点间建立的点对点的隧道,将学习到的新的 MAC 地址通告给每个远端站点。因此,所述将该学习到的新的 MAC 地址通告给每个远端站点的方法为:将该学习到的新的 MAC 地址信息携带在 IGP 协议报文中,根据该远端站点对应的隧道封装映射表项,对该 IGP 协议报文进行隧道封装后发送到核心网;

[0056] 另一方面,边缘设备也会接收到远端站点通过隧道通告的新的 MAC 地址信息。因此,所述得到接收该原始以太网报文的虚拟桥接端口之后,进一步包括:判断该原始以太网报文的协议类型,如果是 IGP 协议报文,则将该 IGP 协议报文中携带的 MAC 地址信息以及接收该原始以太网报文的虚拟桥接端口分别作为 MAC 地址映射表项的 MAC 地址信息、出端口写入 MAC 地址映射表。

[0057] 另外,边缘设备从核心网接收到的经隧道封装,且目的 IP 地址为本机 IP 地址的 IP 报文后,也可以按照与在内部以太网端口上接收到报文时相同的 MAC 地址学习方法,学习该 IP 报文中封装的原始以太网报文的源 MAC 地址。因此,在步骤 603 中,所述得到接收该原始以太网报文的虚拟桥接端口之后,确定该原始以太网报文的所有内部以太网出端口之前,进一步包括:学习该原始以太网报文的源 MAC 地址到接收该原始以太网报文的虚拟桥接端口。

[0058] 现有技术中,当交换机在物理以太网端口接收到报文时,如果接收到的报文为单播报文,则可以查找 MAC 地址映射表确定报文的出端口;如果接收到的报文是广播报文,则将该报文在路由设备的所有其它物理端口上泛洪,也即交换机上所有其它物理端口均是该报文的出端口;如果接收到的报文是组播报文,则根据该报文所属组播组的出接口列表进行报文转发。本发明中,虚拟桥接端口和实际的物理以太网端口具有相同的功能,也同样可以参与单播报文、广播报文、组播报文的转发,例如,当边缘设备接收到广播报文,需要在所有端口泛洪时,所述所有端口中应当包含虚拟桥接端口。

[0059] 因此,步骤 602 中所述确定该在内部以太网端口接收到的以太网报文的所有出端口的的方法为:

[0060] 如果该以太网报文为单播报文,则根据该以太网报文的的目的 MAC 地址在 MAC 地址映射表中查找对应的出端口,若未查找到,则确定该以太网报文的出端口为所有其它内部以太网端口;这里,当在 MAC 地址映射表中查找不到该以太网报文对应的出端口时,为了减少流量,只将该以太网报文在站点内部泛洪,不泛洪到虚拟桥接端口,因此,该以太网报文的出端口为除接收该以太网报文的内部以太网端口外的所有其它内部以太网端口。

[0061] 如果该以太网报文为广播报文,则确定该报文的出端口为所有虚拟桥接端口和所有其它内部以太网端口;这里,当在内部以太网端口接收到的以太网报文为广播报文时,按照现有技术的方法,需要将该以太网报文在所有端口上泛洪,因此,该以太网报文的出端口包括所有虚拟桥接端口、以及除接收该以太网报文的内部以太网端口外的其它内部以太网端口。

[0062] 如果该以太网报文为组播报文,则根据该以太网报文的目的地 MAC 查找该以太网报文对应的组播组得到所有出端口,若未查找到,则确定该以太网报文的出端口为所有其它内部以太网接口。这里,当在内部以太网端口接收到的以太网报文为组播报文时,按照现有技术的方法,需要将该以太网报文在其对应的组播组的各成员端口复制发送,因此,该以太网报文的出端口为该以太网报文对应的组播组的成员端口,也即该以太网报文对应的组播组的出接口列表中的所有出端口。

[0063] 步骤 603 中,确定将 IP 报文解隧道封装得到的在虚拟桥接端口接收到的原始以太网报文的所有出端口的方法为:

[0064] 如果该原始以太网报文为单播报文,则根据该原始以太网报文的目的地 MAC 地址在 MAC 地址映射表中查找对应的出端口,若未查找到,则确定该原始以太网报文的出端口为所有内部以太网端口和所有其它虚拟桥接端口;这里,当在 MAC 地址映射表中未查找到该原始以太网报文的 MAC 地址映射表项时,和现有技术相同,应当将该原始以太网报文在所有端口泛洪,因此,该原始以太网报文的出端口为所有内部以太网端口和除接收该以太网报文的虚拟桥接端口外的所有其它虚拟桥接端口。另外,由于该原始以太网报文是远端站点针对本地站点发送的,即使该原始以太网报文对应的所有出端口中包含虚拟桥接端口,也不允许重新发送回核心网。因此,在对该原始以太网报文进行转发时,根据在该边缘设备上设置的水平分割属性,只将该原始以太网报文从所有出端口中的内部以太网出端口发送出去。

[0065] 如果该原始以太网报文是广播报文,则确定该原始以太网报文的出端口为所有内部以太网端口和所有其它虚拟桥接端口;这里,当该原始以太网报文为广播报文时,按照现有技术的方法,应当将该原始以太网报文在所有端口泛洪,因此,该原始以太网报文的出端口为所有内部以太网端口和除接收该以太网报文的虚拟桥接端口外的所有其它虚拟桥接端口。另外,由于该原始以太网报文是远端站点针对本地站点发送的,即使出端口中包含虚拟桥接端口,也不允许重新发送回核心网。因此,在对该原始以太网报文进行转发时,根据在该边缘设备上设置的水平分割属性,只将该以太网报文从所有出端口中的内部以太网出端口发送出去。

[0066] 如果该原始以太网报文为组播报文,则根据该原始以太网报文的目的地 MAC 查找该原始以太网报文对应的组播组得到所有出端口,若未查找到,则确定该原始以太网报文的出端口为所有内部以太网端口和所有其它虚拟桥接端口。这里,当该原始以太网报文为组播报文时,按照现有技术的方法,需要将该以太网报文在其对应的组播组的各成员端口复制发送,因此,该以太网报文的出端口为该以太网报文对应的组播组的当前已经获取到的成员端口,存储在组播组的出接口列表中。另外,由于该原始以太网报文是远端站点针对本地站点发送的,即使出端口中包含虚拟桥接端口,也不允许重新发送回核心网。因此,在对该原始以太网报文进行转发时,根据在该边缘设备上设置的水平分割属性,只将该以太网报文从所有出端口中的内部以太网出端口发送出去。

[0067] 前面所述内容是对基于边缘设备不支持 VLAN 的情况下确定报文的出端口的方法进行说明,然而,目前的多数设备都支持 VLAN,因此,下面对边缘设备支持 VLAN 的情况下的确定报文的出端口的方法进行说明。

[0068] 图 6 所示本发明实施例中,在针对每个远端站点生成一个虚拟桥接端口时,还可

以进一步在该虚拟桥接端口上配置 VLAN。在边缘设备的虚拟桥接端口上配置 VLAN,在数据中心之间进行报文传输时,只将该报文在其所属的 VLAN 中进行转发。在虚拟桥接端口配置 VLAN 的情况下,可以控制报文的传播范围,在一定程度上缩小广播域。

[0069] 步骤 602 中,所述确定该在内部以太网端口接收到的以太网报文的所有出端口的方法为:

[0070] 如果该以太网报文为单播报文,则根据该以太网报文的目的地 MAC 地址和该以太网报文所属的 VLAN 在 MAC 地址映射表中查找对应的出端口,若未查找到,则确定该以太网报文的出端口为该以太网报文所属 VLAN 中的所有其它内部以太网端口;

[0071] 如果该以太网报文为广播报文,则确定该以太网报文的出端口为该以太网报文所属 VLAN 中的所有虚拟桥接端口和所有其它内部以太网端口;

[0072] 如果该以太网报文为组播报文,则根据该以太网报文的目的地 MAC 和该以太网报文所属的 VLAN 查找该以太网报文对应的组播组得到所有出端口,若未查找到对应的组播组,则确定该以太网报文的出端口为所有其它内部以太网接口。

[0073] 步骤 603 中,确定将 IP 报文解隧道封装后得到的在虚拟桥接端口接收到的原始以太网报文的所有出端口的方法为:

[0074] 如果该原始以太网报文为单播报文,则根据该原始以太网报文的目的地 MAC 地址和该原始以太网报文所属 VLAN 在 MAC 地址映射表中查找对应的出端口,若未查找到,则确定该原始以太网报文的出端口为该原始以太网报文所属 VLAN 中的所有内部以太网端口和所有其它虚拟桥接端口;

[0075] 如果该原始以太网报文是广播报文,则确定该原始以太网报文的出端口为该原始以太网报文所属 VLAN 中的所有内部以太网端口和所有其它虚拟桥接端口;

[0076] 如果该原始以太网报文为组播报文,则根据该原始以太网报文的目的地 MAC 地址和该原始以太网报文所属 VLAN 查找该原始以太网报文对应的组播组得到所有出端口,若未查找到,则确定该原始以太网报文的出端口为所有内部以太网端口和所有其它虚拟桥接端口。

[0077] 在实际应用中,组播组路由器会定期发送组播查询报文,例如,IGMP 查询报文,当交换机接收到该组播查询报文时,将接收该组播查询报文的端口设置为该组播查询报文对应的组播组的路由器端口,并将该组播查询报文从交换机的所有端口转发出去。当主机需要加入一个组播组时,会发送组播加入报文,例如 IGMP 加入报文,运行组播侦听协议的交换机接收到该组播加入报文时,根据该组播加入报文的目的地 MAC 确定对应的组播组,并将接收该组播加入报文的端口作为该组播组的成员端口,加入到该组播组的出接口列表中,然后将该组播加入报文从该组播组的路由器接口转发出去。当交换机接收到组播数据报文时,则先根据该组播数据报文的目的地 MAC 确定对应的组播组,再将组播数据报文在该对应的组播组的出接口列表中的所有出接口上复制发送。

[0078] 图 6 所示本发明实施例中,可以在边缘设备上运行组播侦听协议,用以获取每个组播组的成员端口信息。这里,所述组播侦听协议可以是 IGMP 协议、PIM 协议、或 MLD 协议。

[0079] 这样,边缘设备确定在内部以太网接口接收到的以太网报文为组播报文后,进一步包括:判断该以太网报文是否是组播查询报文或组播加入报文,如果是组播查询报文,则将接收该组播查询报文的内部以太网端口设置为该组播查询报文对应的组播组的路由器

接口,并确定该组播查询报文的出端口为所有其它内部以太网端口;如果是组播加入报文,则将接收该组播加入报文的内部以太网端口加入到该组播加入报文对应的组播组的出接口列表,并确定该组播加入报文的出端口为该组播加入报文对应的组播组的路由器接口;

[0080] 边缘设备将从核心网接收到 IP 报文解隧道封装得到原始以太网报文,并确定该原始以太网报文为组播报文后,进一步包括:判断该原始以太网报文是否是组播查询报文或组播加入报文,如果是组播查询报文,则将该组播查询报文对应的虚拟桥接端口设置为该组播查询报文对应的组播组的路由器接口,并确定该组播查询报文的出端口为所有内部以太网端口;如果是组播加入报文,则将该组播加入报文对应的虚拟桥接端口加入到该组播加入报文对应的组播组的出接口列表,并确定该组播加入报文的出端口为该组播加入报文对应的组播组的路由器接口。

[0081] 以上是对本发明实施例的实现方法进行说明,本发明还提供了一种边缘设备,能够实现数据中心二层互联,且不需要核心网支持组播。

[0082] 参见图 7,图 7 是本发明实施例边缘设备的结构示意图。该边缘设备包括:配置单元 701、第一报文收发单元 702、控制单元 703、封装/解封装单元 704、第二报文收发单元 705;其中,

[0083] 配置单元 701,用于预先针对每个远端站点生成一个虚拟桥接端口,并在隧道封装映射表中配置该虚拟桥接端口对应的隧道封装映射表项;所述隧道封装映射表项包括:虚拟桥接端口、远端站点的边缘设备的 IP 地址、本地站点的边缘设备的 IP 地址;

[0084] 第一报文收发单元 702,用于在内部以太网端口接收以太网报文,并将该以太网报文发送到控制单元 703;用于在控制单元 703 接收到第一报文收发单元 702 发送的以太网报文,并确定该以太网报文的所有出端口后,对于该以太网报文的所有出端口中的每一个内部以太网端口,将该以太网报文从该内部以太网端口转发出去;用于在控制单元 703 接收到封装/解封装单元 704 发送的原始以太网报文,并确定该原始以太网报文的所有内部以太网出端口后,将该原始以太网报文从确定的每个内部以太网出端口发送出去;

[0085] 控制单元 703,用于接收第一报文收发单元 702 发来的以太网报文,确定该以太网报文的所有出端口;用于接收封装/解封装单元 704 发来的原始以太网报文,得到接收该原始以太网报文的虚拟桥接端口,确定该原始以太网报文的所有内部以太网出端口;

[0086] 封装/解封装单元 704,用于在控制单元 703 接收到第一报文收发单元 702 发送的以太网报文,并确定该以太网报文的所有出端口后,对于该以太网报文的所有出端口中的每一个虚拟桥接端口,根据该虚拟桥接端口对应的隧道封装映射表项,对该以太网报文进行隧道封装,并将经隧道封装后的以太网报文发送到第二报文收发单元 705;用于接收第二报文收发单元 705 发来的经隧道封装,且目的 IP 地址为该边缘设备的 IP 地址的 IP 报文,将所述 IP 报文解隧道封装得到该 IP 报文的源 IP 地址、目的 IP 地址以及原始的以太网报文,将该原始以太网报文发送到控制单元 703;

[0087] 第二报文收发单元 705,用于接收封装/解封装单元 704 发来的经隧道封装的以太网报文,并将该经隧道封装的以太网报文向远端站点发送;用于从核心网接收所述 IP 报文,将所述 IP 报文发送到封装/解封装单元 704。

[0088] 较佳地,所述配置单元 701,用于设置该边缘设备的水平分割属性;

[0089] 所述控制单元 703 在接收到封装/解封装单元 704 发来的原始以太网报文之后,

- 根据所述水平分割属性,确定该原始以太网报文的出端口仅限于所有内部以太网出端口。
- [0090] 所述配置单元 701 设置实现该边缘设备的水平分割属性是通过端口 Source ID 实现的,其中,每个内部以太网端口设置有不同的源标识 Source ID 值,所有虚拟桥接端口设置相同的 Source ID 值;且虚拟桥接端口的 Source ID 值与每个内部以太网端口的 Source ID 值均不相同;
- [0091] 所述控制单元 703 在确定该原始以太网报文的所有内部以太网出端口时,用于:将该原始以太网报文的所有出端口中 Source ID 值与所述虚拟桥接端口的 Source ID 值不同出端口确定为该原始以太网报文的内部以太网出端口。
- [0092] 所述的配置单元 701,用于存储 MAC 地址映射表;
- [0093] 所述控制单元 703 接收到第一报文收发单元 702 发来的以太网报文之后,确定该以太网报文的所有出端口之前,进一步用于:进行 MAC 地址学习,并将学习到的新的 MAC 地址信息携带在 IGP 协议报文中发送到封装/解封装单元 704;
- [0094] 所述封装/解封装单元 704,用于接收控制单元 703 发来的 IGP 协议报文,针对每个远端站点,根据该远端站点对应的隧道封装映射表项,对该 IGP 协议报文进行隧道封装后发送到第二报文收发单元 705;
- [0095] 所述第二报文收发单元 705,用于接收到封装/解封装单元 704 发来的经隧道封装的 IGP 协议报文后,将该经隧道封装的 IGP 协议报文发送到核心网;
- [0096] 所述控制单元 703,在得到接收该原始以太网报文的虚拟桥接端口之后,进一步用于:判断该原始以太网报文的协议类型,如果是 IGP 协议报文,则将该 IGP 协议报文中携带的远端站点学习到的 MAC 地址信息以及接收该原始以太网报文的虚拟桥接端口分别作为 MAC 地址映射表项的 MAC 地址信息、出端口写入 MAC 地址映射表。
- [0097] 所述控制单元 703 在得到接收该原始以太网报文的虚拟桥接端口之后,确定该原始以太网报文的所有出端口之前,进一步用于:学习该原始以太网报文的源 MAC 地址信息到接收该原始以太网报文的虚拟桥接端口。
- [0098] 所述控制单元 703 在确定第一报文收发单元 702 发来的以太网报文的所有出端口时,用于:
- [0099] 如果该以太网报文为单播报文,则根据该以太网报文的的目的 MAC 地址在 MAC 地址映射表中查找对应的出端口,若未查找到,则确定该以太网报文的出端口为所有其它内部以太网端口;
- [0100] 如果该以太网报文为广播报文,则确定该以太网报文的出端口为所有虚拟桥接端口和所有其它内部以太网端口;
- [0101] 如果该以太网报文为组播报文,则根据该以太网报文的的目的 MAC 查找该以太网报文对应的组播组得到所有出端口,若未查找到对应的组播组,则确定该以太网报文的出端口为所有其它内部以太网接口。
- [0102] 所述控制单元 703 在确定封装/解封装单元 704 发来的原始以太网报文的所有出端口时,用于:
- [0103] 如果该原始以太网报文为单播报文,则根据该原始以太网报文的的目的 MAC 地址在 MAC 地址映射表中查找对应的出端口,若未查找到,则确定该原始以太网报文的出端口为所有内部以太网端口和所有其它虚拟桥接端口;

[0104] 如果该原始以太网报文是广播报文,则确定该原始以太网报文的出端口为所有内部以太网端口和所有其它虚拟桥接端口;

[0105] 如果该原始以太网报文为组播报文时,则根据该原始以太网报文的目的地 MAC 查找该原始以太网报文对应的组播组得到所有出端口,若未查找到,则确定该原始以太网报文的出端口为所有内部以太网端口和所有其它虚拟桥接端口。

[0106] 所述配置单元 701,用于针对每个远端站点生成一个虚拟桥接端口时,进一步在该虚拟桥接端口上配置 VLAN;

[0107] 所述控制单元 703 在确定第一报文收发单元 702 发来的以太网报文的出端口时,用于:

[0108] 如果该以太网报文为单播报文,则根据该以太网报文的目的地 MAC 地址和该以太网报文所属的 VLAN 在 MAC 地址映射表中查找对应的出端口,若未查找到,则确定该以太网报文的出端口为该以太网报文所属 VLAN 中的所有其它内部以太网端口;

[0109] 如果该以太网报文为广播报文,则确定该以太网报文的出端口为该以太网报文所属 VLAN 中的所有虚拟桥接端口和所有其它内部以太网端口;

[0110] 如果该以太网报文为组播报文,则根据该以太网报文的目的地 MAC 和该以太网报文所属的 VLAN 查找该以太网报文对应的组播组得到所有出端口,若未查找到对应的组播组,则确定该以太网报文的出端口为该以太网报文所属 VLAN 中的所有其它内部以太网接口。

[0111] 所述控制单元 703 在确定封装/解封装单元 704 发来的原始以太网报文的所有出端口时,用于:

[0112] 如果该原始以太网报文为单播报文,则根据该原始以太网报文的目的地 MAC 地址该原始以太网报文所属 VLAN 在 MAC 地址映射表中查找对应的出端口,若未查找到,则确定该原始以太网报文的出端口为该原始以太网报文所属 VLAN 中的所有内部以太网端口和所有其它虚拟桥接端口;

[0113] 如果该原始以太网报文是广播报文时,则确定该原始以太网报文的出端口为该原始以太网报文所属 VLAN 中的所有内部以太网端口和所有其它虚拟桥接端口;

[0114] 如果该原始以太网报文为组播报文时,则根据该原始以太网报文的目的地 MAC 和该原始以太网报文所属的 VLAN 查找该原始以太网报文对应的组播组得到所有出端口,若未查找到,则确定该原始以太网报文的出端口为该原始以太网报文所属 VLAN 中的所有内部以太网端口和所有其它虚拟桥接端口。

[0115] 在边缘设备上运行组播侦听协议;

[0116] 所述控制单元 703 确定第一报文收发单元 702 发来的以太网报文为组播报文后,进一步用于:判断该以太网报文是否是组播查询报文或组播加入报文,如果是组播查询报文,则将接收该组播查询报文的内部以太网端口设置为该组播查询报文对应的组播组的路由器接口,并确定该组播查询报文的出端口为所有其它内部以太网端口;如果是组播加入报文,则将接收该组播加入报文的内部以太网端口加入到该组播加入报文对应的组播组的出接口列表,并确定该组播加入报文的出端口为该组播加入报文对应的组播组的路由器接口;

[0117] 所述控制单元 703 在确定封装/解封装单元 704 发来的原始以太网报文为组播报文后,进一步用于:判断该原始以太网报文是否是组播查询报文或组播加入报文,如果是组

播查询报文,则将该组播查询报文对应的虚拟桥接端口设置为该组播查询报文对应的组播组的路由器接口,并确定该组播查询报文的出端口为所有内部以太网端口;如果是组播加入报文,则将该组播加入报文对应的虚拟桥接端口加入到该组播加入报文对应的组播组的出接口列表,并确定该组播加入报文的出端口为该组播加入报文对应的组播组的路由器接口。

[0118] 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。

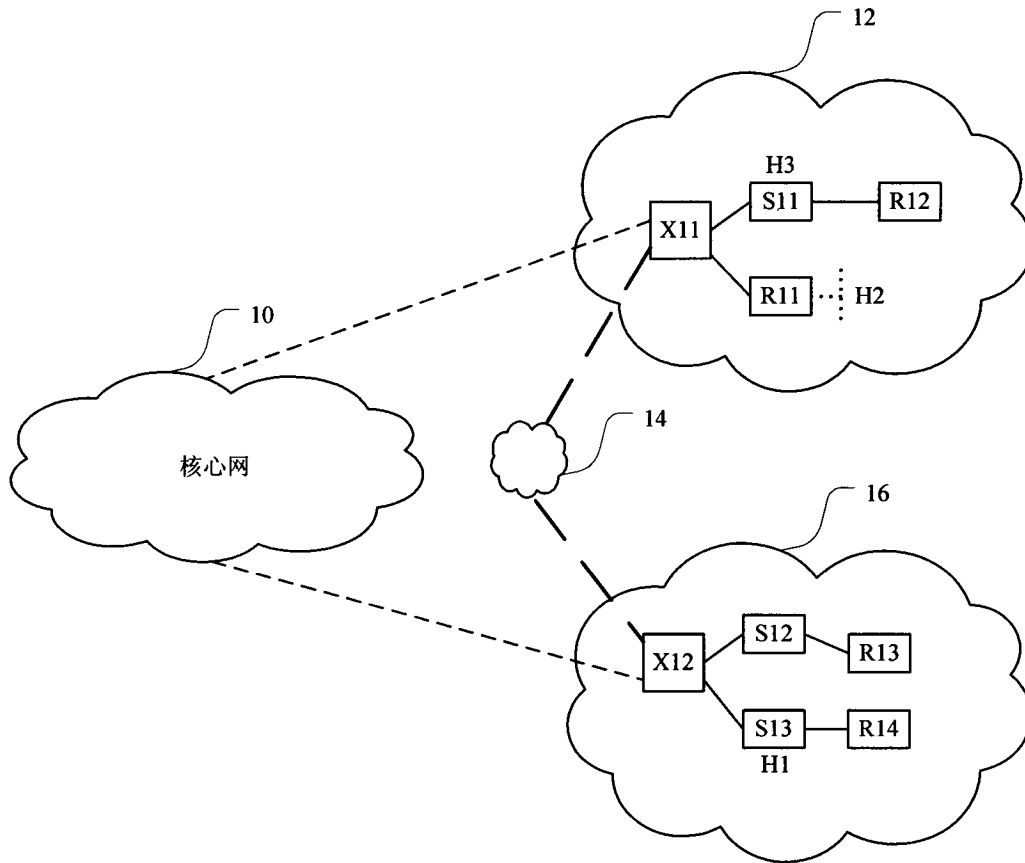


图 1

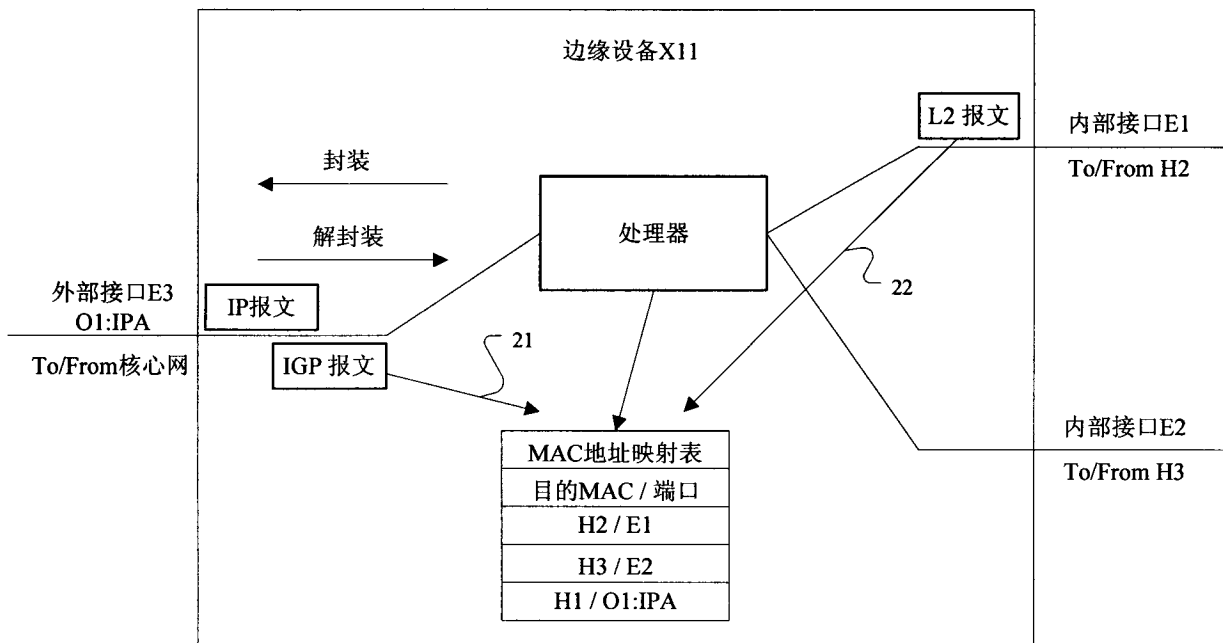


图 2

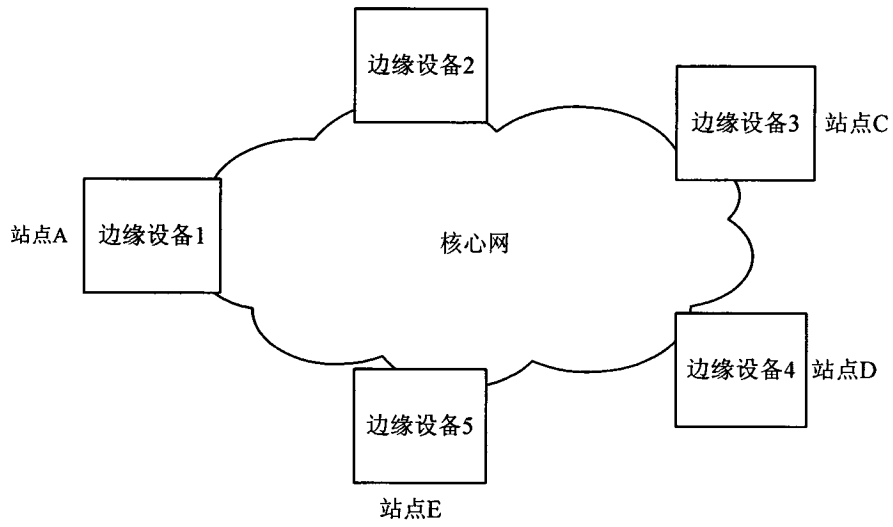


图 3

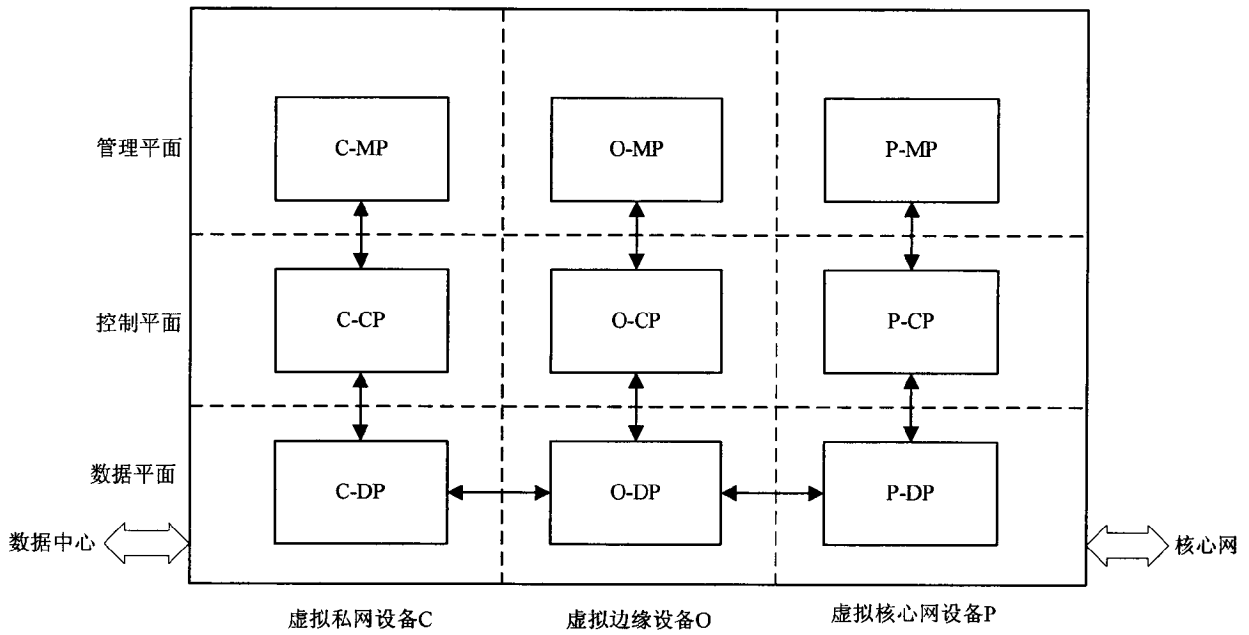


图 4

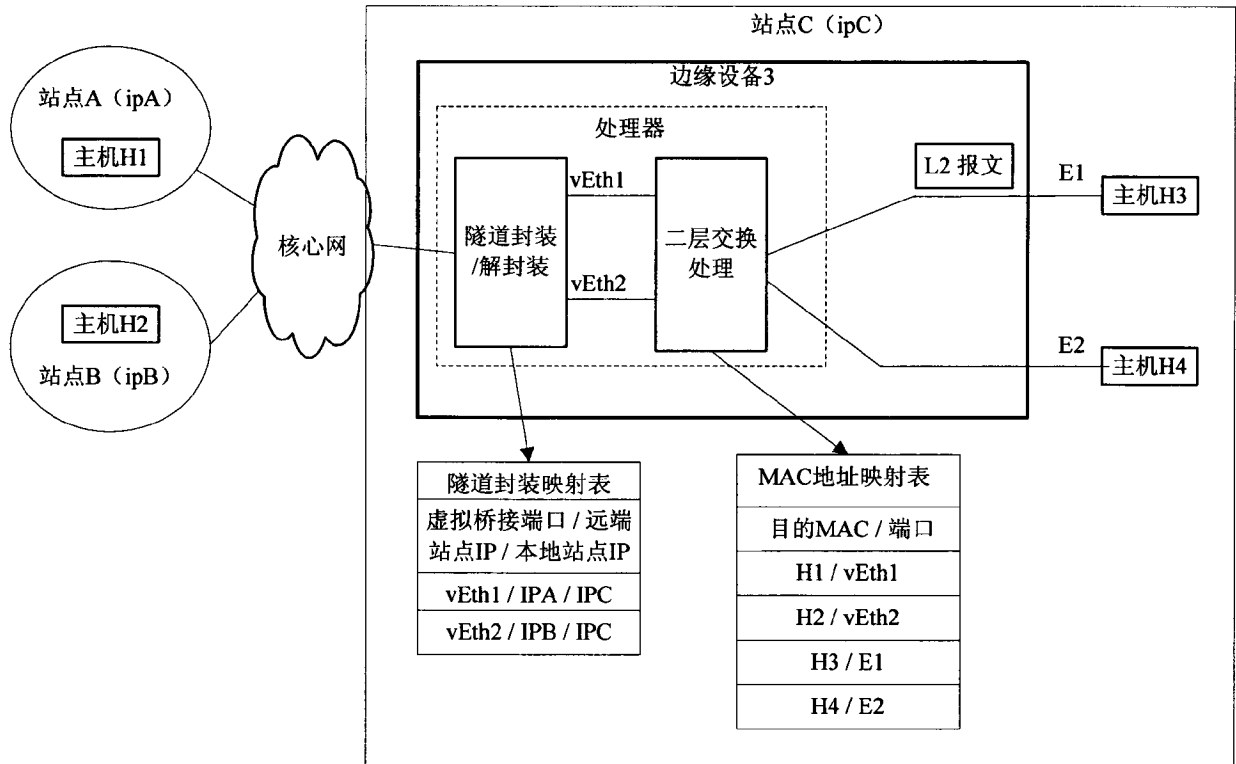


图 5

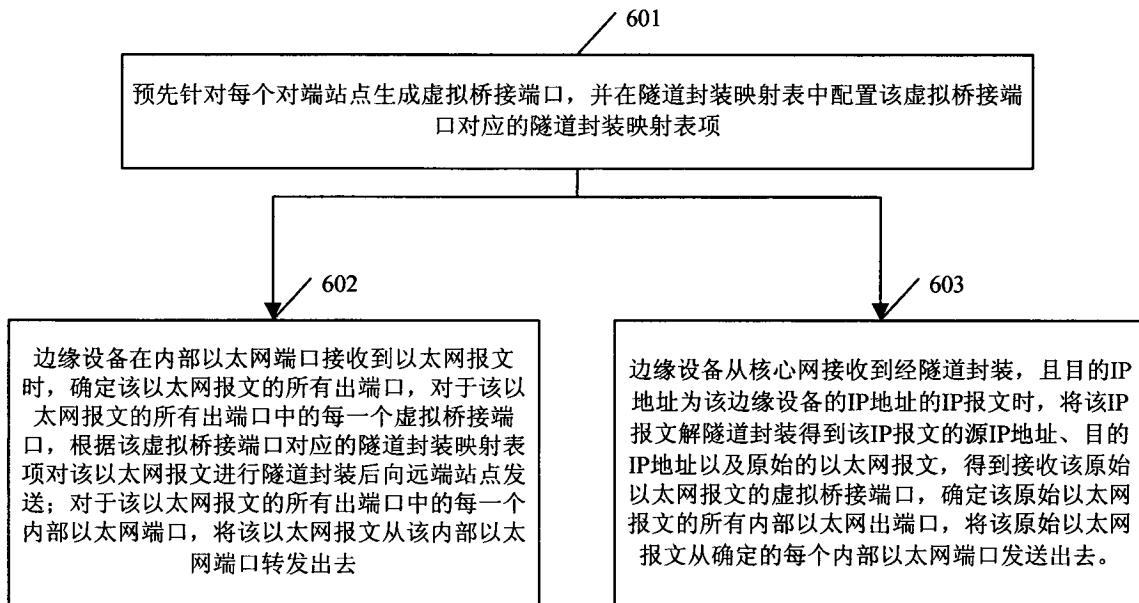


图 6

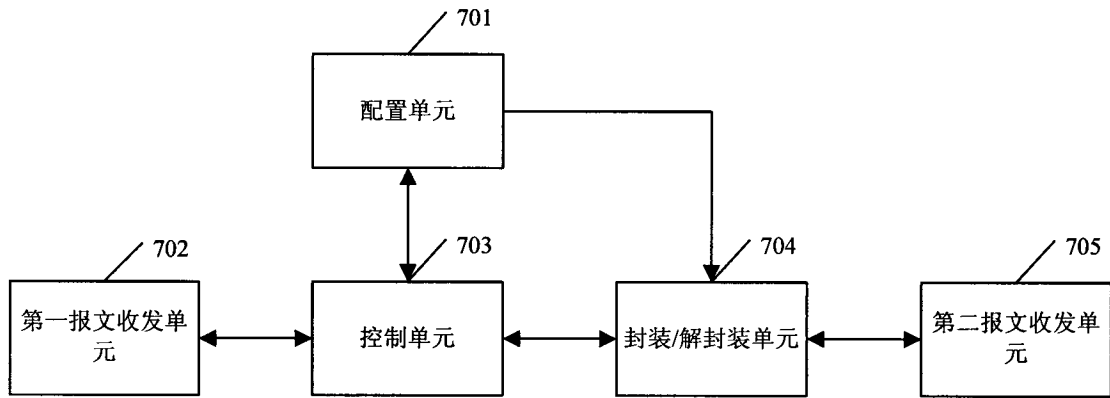


图 7