



(12)发明专利

(10)授权公告号 CN 106209636 B

(45)授权公告日 2019.08.02

(21)申请号 201510220886.5

(22)申请日 2015.05.04

(65)同一申请的已公布的文献号
申请公布号 CN 106209636 A

(43)申请公布日 2016.12.07

(73)专利权人 新华三技术有限公司
地址 310052 浙江省杭州市滨江区长河路
466号

(72)发明人 张洋

(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 杨春香 宋志强

(51)Int.Cl.

H04L 12/741(2013.01)

H04L 12/761(2013.01)

(56)对比文件

CN 104350714 A,2015.02.11,
CN 103259727 A,2013.08.21,
CN 103095546 A,2013.05.08,
WO 03073283 A1,2003.09.04,
US 2014192804 A1,2014.07.10,

审查员 张诗纬

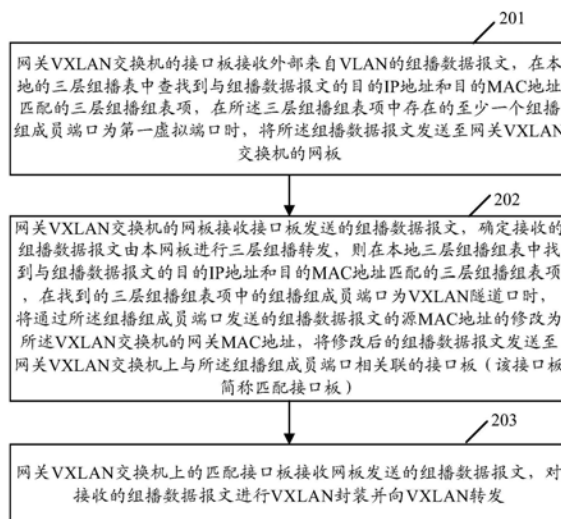
权利要求书3页 说明书10页 附图3页

(54)发明名称

从VLAN至VXLAN的组播数据报文转发方法和设备

(57)摘要

本申请提供了从VLAN至VXLAN的组播数据报文转发方法和设备。本发明中,网关VXLAN交换机的网板不再使用仅具有报文转发功能的网片,而是使用具有转发功能和表项功能的芯片比如与接口板相同的交换芯片。通过本发明中,网关VXLAN交换机实现VXLAN的三层网关功能,以实现从VLAN至VXLAN的组播数据报文转发,且整个过程只是在网关VXLAN交换机内部执行,没有浪费任何带宽资源,做到了组播数据报文线速转发的。



1. 一种从虚拟局域网VLAN至虚拟可扩展局域网VXLAN的组播数据报文转发方法,其特征在于,该方法包括:

VXLAN交换机的接口板接收外部来自VLAN的组播数据报文,在本地的三层组播表中找到与组播数据报文的目IP地址和目的MAC地址匹配的三层组播组表项,在所述三层组播组表项中存在的至少一个组播组成员端口为第一虚拟端口时,将所述组播数据报文发送至所述VXLAN交换机的网板;

所述VXLAN交换机的网板接收接口板发送的组播数据报文,确定接收的组播数据报文由本网板进行三层组播转发,则在本地三层组播组表中找到与组播数据报文的目IP地址和目的MAC地址匹配的三层组播组表项,在找到的三层组播组表项中的组播组成员端口为VXLAN隧道口时,将通过所述组播组成员端口发送的组播数据报文的源MAC地址的修改为所述VXLAN交换机的网关MAC地址,将修改后的组播数据报文发送至VXLAN交换机上与所述组播组成员端口相关联的接口板;

所述VXLAN交换机上与所述组播组成员端口相关联的接口板接收网板发送的组播数据报文,对接收的组播数据报文进行VXLAN封装并向VXLAN转发。

2. 根据权利要求1所述的方法,其特征在于,所述将组播数据报文发送至所述VXLAN交换机的网板包括:

对组播数据报文进行第一内部封装,第一内部封装包含:第一目的芯片标识mod和第一目的端口标识port,第一目的mod为第一虚拟芯片的标识,第一目的port为第一虚拟端口的标识;

在本接口板仅连接一个网板时,通过本接口板连接网板的互连口发送第一内部封装的组播数据报文;在本接口板连接多个网板时,通过与多个网板分别连接的互连口中的一个互连口发送第一内部封装的组播数据报文。

3. 根据权利要求2所述的方法,其特征在于,所述确定接收的组播数据报文由本网板进行三层组播转发包括:

从接收的组播数据报文的第一内部封装头中确定出第一目的mod和第一目的port;如果确定出的第一目的mod为第一虚拟芯片的标识,第一目的port为第一虚拟端口的标识,则对组播数据报文的第一内部封装进行解封装,如果解封装后的组播数据报文的目MAC地址为组播MAC地址、目的IP地址为组播IP地址,则确定接收的组播数据报文由本网板进行三层组播转发。

4. 根据权利要求1所述的方法,其特征在于,所述将修改后的组播数据报文发送至VXLAN交换机上与所述组播组成员端口相关联的接口板包括:

依据所述组播组成员端口确定目标出端口,选择所述目标出端口所在的接口板为与所述组播组成员端口相关联的接口板;

对修改后的组播数据报文进行第二内部封装,将第二内部封装的报文发送至所述目标出端口所在的接口板;第二内部封装包含:所述组播组成员端口对应的VXLAN隧道封装表项索引和虚拟转发实例VFI;

所述对接收的组播数据报文进行VXLAN封装并向VXLAN转发包括:

从组播数据报文的第二内部封装头中确定出VXLAN隧道封装表项索引和VFI并对组播数据报文的第二内部封装进行解封装,依据所述VXLAN隧道封装表项索引在本地找到对应

的VXLAN隧道封装表项,从预先设置的VFI和VNID之间的映射关系中找到与该确定出的VFI具有映射关系的VNID,利用所述VXLAN隧道封装表项和找到的VNID对组播数据报文进行VXLAN封装并通过该找到的VXLAN隧道封装表项中的出端口发送。

5. 根据权利要求4所述的方法,其特征在于,依据所述组播组成员端口确定目标出端口包括:

在本地找到与所述组播组成员端口相关联的VXLAN隧道对应的下一跳表项;

在所述下一跳表项中仅有一个下一跳时,确定所述下一跳表项中所述下一跳为目标下一跳,在所述下一跳表项中有两个以上下一跳时,从所述下一跳表项中的两个以上下一跳中选择一个下一跳确定为目标下一跳;

在所述下一跳表项中与所述目标下一跳对应的出端口为单一的物理端口时,确定所述单一的物理端口为目标出端口;在所述下一跳表项中与所述目标下一跳对应的出端口为至少两个物理端口聚合形成的聚合端口,确定所述聚合端口中一个物理端口为所述目标出端口。

6. 一种从虚拟局域网VLAN至虚拟可扩展局域网VXLAN的组播数据报文转发设备,其特征在于,该设备包括:

接口板处理单元,处于所述VXLAN交换机的接口板,用于接收外部来自VLAN的组播数据报文,在本地的三层组播表中查找到与组播数据报文的目IP地址和目的MAC地址匹配的三层组播组表项,在所述三层组播组表项中存在的至少一个组播组成员端口为第一虚拟端口时,将所述组播数据报文发送至所述VXLAN交换机的网板;以及,接收网板发送的组播数据报文,对接收的组播数据报文进行VXLAN封装并向VXLAN转发;

网板处理单元,处于所述VXLAN交换机的网板,用于接收接口板发送的组播数据报文,确定接收的组播数据报文由本网板进行三层组播转发,则在本地三层组播组表中找到与组播数据报文的目IP地址和目的MAC地址匹配的三层组播组表项,在找到的三层组播组表项中的组播组成员端口为VXLAN隧道口时,将通过所述组播组成员端口发送的组播数据报文的源MAC地址的修改为所述VXLAN交换机的网关MAC地址,将修改后的组播数据报文发送至VXLAN交换机上与所述组播组成员端口相关联的接口板。

7. 根据权利要求6所述的设备,其特征在于,所述接口板处理单元通过以下步骤将组播数据报文发送至所述VXLAN交换机的网板:

对组播数据报文进行第一内部封装,第一内部封装包含:第一目的芯片标识mod和第一目的端口标识port,第一目的mod为第一虚拟芯片的标识,第一目的port为第一虚拟端口的标识;

在本接口板仅连接一个网板时,通过本接口板连接网板的互连口发送第一内部封装的组播数据报文;在本接口板连接多个网板时,通过与多个网板分别连接的互连口中的一个互连口发送第一内部封装的组播数据报文。

8. 根据权利要求7所述的设备,其特征在于,所述网板处理单元通过以下步骤确定接收的组播数据报文由本网板进行三层组播转发:从接收的组播数据报文的第一内部封装头中确定出第一目的mod和第一目的port;如果确定出的第一目的mod为第一虚拟芯片的标识,第一目的port为第一虚拟端口的标识,则对组播数据报文的第一内部封装进行解封装,如果解封装后的组播数据报文的目MAC地址为组播MAC地址、目的IP地址为组播IP地址,则

确定接收的组播数据报文由本网板进行三层组播转发。

9. 根据权利要求6所述的设备,其特征在于,所述网板处理单元通过以下步骤将修改后的组播数据报文发送至VXLAN交换机上与所述组播组成员端口相关联的接口板:

依据所述组播组成员端口确定目标出端口,选择所述目标出端口所在的接口板为与所述组播组成员端口相关联的接口板;对修改后的组播数据报文进行第二内部封装,将第二内部封装的报文发送至所述目标出端口所在的接口板;第二内部封装包含:所述组播组成员端口对应的VXLAN隧道封装表项索引和虚拟转发实例VFI;

所述接口板处理单元通过以下步骤对接收的组播数据报文进行VXLAN封装并向VXLAN转发:

从组播数据报文的第二内部封装头中确定出VXLAN隧道封装表项索引和VFI并对组播数据报文的第二内部封装进行解封装,依据所述VXLAN隧道封装表项索引在本地找到对应的VXLAN隧道封装表项,从预先设置的VFI和VNID之间的映射关系中找到与该确定出的VFI具有映射关系的VNID,利用所述VXLAN隧道封装表项和找到的VNID对组播数据报文进行VXLAN封装并通过该找到的VXLAN隧道封装表项中的出端口发送。

10. 根据权利要求9所述的设备,其特征在于,所述网板处理单元通过以下步骤依据所述组播组成员端口确定目标出端口:

在本地找到与所述组播组成员端口相关联的VXLAN隧道对应的下一跳表项;

在所述下一跳表项中仅有一个下一跳时,确定所述下一跳表项中所述下一跳为目标下一跳,在所述下一跳表项中有两个以上下一跳时,从所述下一跳表项中的两个以上下一跳中选择一个下一跳确定为目标下一跳;

在所述下一跳表项中与所述目标下一跳对应的出端口为单一的物理端口时,确定所述单一的物理端口为目标出端口;在所述下一跳表项中与所述目标下一跳对应的出端口为至少两个物理端口聚合形成的聚合端口,确定所述聚合端口中一个物理端口为所述目标出端口。

从VLAN至VXLAN的组播数据报文转发方法和设备

技术领域

[0001] 本申请涉及网络通信技术,特别涉及从虚拟局域网(VLAN:Virtual Local Access Network)至虚拟可扩展局域网(VXLAN:Virtual eXtensible Local Access Network)的组播数据报文转发方法和设备。

背景技术

[0002] 目前,网关节点的硬件交换芯片即使支持三层组播转发,硬件交换芯片的转发逻辑也不允许三层组播转发,只能进行二层组播转发。

[0003] 比如,图1所示的组网,PC发送的组播数据报文在非网关节点进行VXLAN封装到达网关节点后,网关节点对组播数据报文进行VXLAN解封装,发现组播数据报文的目的地MAC地址是组播MAC地址,且组播数据报文的目的地IP地址是组播IP地址,按照正常的处理方式,网关节点会对VXLAN解封装后的组播数据报文进行三层组播复制转发,但限于网关节点上硬件交换芯片的转发逻辑,即使网关节点的硬件交换芯片支持三层组播转发,也不能实现三层组播转发,只能对VXLAN解封装后的组播数据报文进行二层转发。

[0004] 因此,在VXLAN交换机上实现VXLAN的三层网关功能以实现跨从VLAN至VXLAN的三层组播数据报文转发是当前亟待解决的技术问题。

发明内容

[0005] 本申请提供了从VLAN至VXLAN的组播数据报文转发方法和设备,用于在在作为网关的VXLAN交换机上实现VXLAN的三层网关功能,以实现从VLAN至VXLAN的组播数据报文转发。

[0006] 本申请提供的技术方案包括:

[0007] 一种从虚拟局域网VLAN至虚拟可扩展局域网VXLAN的组播数据报文转发方法,该方法包括:

[0008] VXLAN交换机的接口板接收外部来自VLAN的组播数据报文,在本地的三层组播表中查找到与组播数据报文的目的地IP地址和目的地MAC地址匹配的三层组播组表项,在所述三层组播组表项中存在的至少一个组播组成员端口为第一虚拟端口时,将通过所述组播组成员端口发送的组播数据报文发送至所述VXLAN交换机的网板;

[0009] 所述VXLAN交换机的网板接收接口板发送的组播数据报文,确定接收的组播数据报文由本网板进行三层组播转发,则在本地三层组播组表中找到与组播数据报文的目的地IP地址和目的地MAC地址匹配的三层组播组表项,在找到的三层组播组表项中的组播组成员端口为VXLAN隧道口时,将接收的组播数据报文的源MAC地址的修改为所述VXLAN交换机的网关MAC地址,将修改后的组播数据报文发送至VXLAN交换机上与所述组播组成员端口相关联的接口板;

[0010] 所述VXLAN交换机上与所述组播组成员端口相关联的接口板接收网板发送的组播数据报文,对接收的组播数据报文进行VXLAN封装并向VXLAN转发。

[0011] 一种从虚拟局域网VLAN至虚拟可扩展局域网VXLAN的组播数据报文转发设备,包括:

[0012] 接口板处理单元,处于所述VXLAN交换机的接口板,用于接收外部来自VLAN的组播数据报文,在本地的三层组播表中查找到与组播数据报文的目IP地址和目的MAC地址匹配的三层组播组表项,在所述三层组播组表项中存在的至少一个组播组成员端口为第一虚拟端口时,将所述组播数据报文发送至所述VXLAN交换机的网板;以及,接收网板发送的组播数据报文,对接收的组播数据报文进行VXLAN封装并向VXLAN转发;

[0013] 网板处理单元,处于所述VXLAN交换机的网板,用于接收接口板发送的组播数据报文,确定接收的组播数据报文由本网板进行三层组播转发,则在本地三层组播组表中找到与组播数据报文的目IP地址和目的MAC地址匹配的三层组播组表项,在找到的三层组播组表项中的组播组成员端口为VXLAN隧道口时,将接收的组播数据报文的源MAC地址的修改为所述VXLAN交换机的网关MAC地址,将修改后的组播数据报文发送至VXLAN交换机上与所述组播组成员端口相关联的接口板。

[0014] 由以上技术方案可以看出,本发明中,作为网关的VXLAN交换机通过接口板接收外部来自VLAN的组播数据报文,在本地的三层组播表中查找到与组播数据报文的目IP地址和目的MAC地址匹配的三层组播组表项,在所述三层组播组表项中存在的至少一个组播组成员端口为第一虚拟端口时,将所述组播数据报文发送至所述VXLAN交换机的网板,在网板本地的三层组播组表中找到与组播数据报文的目IP地址和目的MAC地址匹配的三层组播组表项,在找到的三层组播组表项中的组播组成员端口为VXLAN隧道口时,将通过所述组播组成员端口发送的组播数据报文的源MAC地址的修改为所述VXLAN交换机的网关MAC地址,将修改后的组播数据报文发送至VXLAN交换机上与所述组播组成员端口相关联的接口板以由接口板对组播数据报文进行VXLAN封装并向VXLAN转发,即通过作为网关的VXLAN交换机实现了VXLAN的三层网关功能,实现从VLAN至VXLAN的组播数据报文转发,且整个过程只是在作为网关的VXLAN交换机内部执行,没有浪费任何带宽资源,做到了组播数据报文线速转发的。

附图说明

[0015] 图1为现有跨网络转发的应用组网图;

[0016] 图2为本发明提供的方法流程图;

[0017] 图3为本发明提供的实施例应用图;

[0018] 图4为本发明提供的设备结构图;

[0019] 图5为本发明提供的图4所示设备的硬件结构图。

具体实施方式

[0020] 为了使本发明的目的、技术方案和优点更加清楚,下面结合附图和具体实施例对本发明进行详细描述。

[0021] 一般的框式交换机由主控板(main board)、接口板(I/O board)、网板(fabric board)组成。主控板为用于实现协议栈计算、转发表项下发及控制、设备管理的单板。接口板为用于实现数据报文转发(包括内部转发和外部转发)的单板。网板为用于实现数据报文

及控制报文在设备内部跨板、跨芯片转发的单板。其中，网板上的芯片只具有报文转发功能，用于在不同接口板之间中转报文。

[0022] VXLAN交换机的结构与现有框式交换机的结构一样，由主控板、接口板、网板组成。在通常的应用中，接口板的个数大于1。

[0023] 在VXLAN应用中，VXLAN交换机的网板与现有框式交换机的网板一样设置，网板上的芯片只具有报文转发功能，仅负责在不同接口板之间中转报文。

[0024] 本发明中，为了实现VXLAN的三层网关功能，可选择其中一个VXLAN交换机作为网关，对作为网关的VXLAN交换机（简称网关VXLAN交换机）进行改进，而对不作为网关的VXLAN交换机，可不作任何改进，仍按照现有的结构执行。

[0025] 其中，对网关VXLAN交换机进行的改进包括：

[0026] 网关VXLAN交换机的网板采用具有转发功能和表项功能的芯片，这里的表项功能至少包含接收并保存主控板下发的表项、表项查找等功能。优选地，作为一种较佳实施例，这里网板采用的芯片可以为网关VXLAN交换机的接口板所采用的交换芯片。

[0027] 当网关VXLAN交换机的网板采用具有转发功能和表项功能的芯片，这样，进入VXLAN交换机的组播数据报文可以通过VXLAN交换机的接口板和网板相互合作完成三层转发。下面通过图2对本发明提供的方法进行描述：

[0028] 参见图2，图2为本发明提供的方法流程图。该方法应用于网关VXLAN交换机。如上所述，网关VXLAN交换机的网板不再采用仅具有报文转发功能的芯片，而是采用具有转发功能和表项功能的芯片比如与接口板相同的交换芯片。基于此，如图2所示，该流程可包括以下步骤：

[0029] 步骤201，网关VXLAN交换机的接口板接收外部来自VLAN的组播数据报文，在本地的三层组播表中查找到与组播数据报文的目IP地址和目的MAC地址匹配的三层组播组表项，在所述三层组播组表项中存在的至少一个组播组成员端口为第一虚拟端口时，将所述组播数据报文发送至网关VXLAN交换机的网板。

[0030] 这里，接口板本地的三层组播组表项可以是主控板直接配置的，本发明并不具体限定。

[0031] 本发明中，三层组播组表项中的组播组成员端口可不止一个，但不管三层组播组表项中的组播组成员端口有多少个，只要有一个组播组成员端口为第一虚拟端口，就会将接收的组播数据报文发送至网关VXLAN交换机的网板。反之，在任何一个组播组成员端口都不为第一虚拟端口，比如为虚拟局域网络（VLAN）端口，则可执行其它操作。本专利重点是有组播组成员端口为第一虚拟端口的情况，至于组播组成员端口都不为第一虚拟端口的情况不是本专利的重点，这里不展开描述。

[0032] 至于本发明中的第一虚拟端口所表示的意义在下文进行了描述，这里暂不赘述。

[0033] 作为本发明的一个实施例，步骤201中，为了保证网关VXLAN交换机的接口板将组播数据报文顺利发送至网关VXLAN交换机的网板，一种较佳的实现方式是对发送至网板的组播数据报文进行用于转发至网板的内部封装，为便于描述，这里将该内部封装简称为第一内部封装。如此，步骤201中，接口板将组播数据报文发送至网关VXLAN交换机的网板可包括：

[0034] 步骤a1，对所述组播数据报文进行第一内部封装。

[0035] 这里,第一内部封装包含:第一目的芯片标识(mod)和第一目的端口标识(port),第一目的mod为第一虚拟芯片的标识,第一目的port为第一虚拟端口的标识;

[0036] 优选地,作为本发明一个实例,第一虚拟芯片可为预先配置的用于指示网板可继续执行查表进行组播数据报文转发的虚拟芯片,第一虚拟端口可为预先配置的用于指示由网板代理三层组播转发的虚拟端口,这里的代理三层组播转发具体为:继续执行查表进行组播数据报文转发。

[0037] 步骤a2,在本接口板仅连接一个网板时,通过本接口板连接网板的互连口发送第一内部封装的组播数据报文;在本接口板连接多个网板时,通过与多个网板分别连接的互连口中的一个互连口发送第一内部封装的组播数据报文。

[0038] 实际应用中,在本接口板连接多个网板时,本接口板与多个网板分别连接的互连口可聚合成一个聚合组。如此,在步骤a2中,通过与多个网板分别连接的互连口中的一个互连口发送第一内部封装的组播数据报文可包括:从所述聚合组中选择一个互连口,通过选择的一个互连口发送第一内部封装的组播数据报文。这里,从所述聚合组中选择一个互连口的方式有多种,比如随机选择一个互连口,或者按照一定的算法比如通过哈希算法选择一个互连口等,本发明并不具体限定。

[0039] 通过上述步骤a1和步骤a2,网关VXLAN交换机的网板最终接收的组播数据报文是带有第一内部封装的组播数据报文。

[0040] 步骤202,网关VXLAN交换机的网板接收接口板发送的组播数据报文,确定接收的组播数据报文由本网板进行三层组播转发,则在本地三层组播组表中找到与组播数据报文的目IP地址和目的MAC地址匹配的三层组播组表项,在找到的三层组播组表项中的组播组成员端口为VXLAN隧道口时,将通过所述组播组成员端口发送的组播数据报文的源MAC地址的修改为所述VXLAN交换机的网关MAC地址,将修改后的组播数据报文发送至网关VXLAN交换机上与所述组播组成员端口相关联的接口板(该接口板简称匹配接口板)。

[0041] 本发明中,网关本地的三层组播组表项不同于接口板本地的三层组播组表项,其中,网关本地的三层组播组表项可以是主控板按照类似现有三层组播组表的学习方式学习并下发给网板的,本发明并不具体限定。

[0042] 本步骤202中的VXLAN隧道口是一个虚拟的概念,作为本发明一个实施例,本步骤202中的VXLAN隧道口可通过VXLAN隧道封装索引表示,比如VXLAN隧道口为Tunnel1(Tunnel1为VXLAN隧道封装索引)。

[0043] 本发明中,网板找到的三层组播组表项中的组播组成员端口可不止一个,基于此,当网板找到的三层组播组表项中的组播组成员端口数量大于1时,本步骤202可针对每一组播组成员端口,其会先识别组播组成员端口是否为VXLAN隧道口,在该组播组成员端口为VXLAN隧道口时执行步骤202所述的以下操作:将通过组播组成员端口发送的组播数据报文的源MAC地址修改为所述VXLAN交换机的网关MAC地址,将修改后的组播数据报文发送至匹配接口板。反之,在该组播组成员端口不为VXLAN隧道口,比如为虚拟局域网络(VLAN)端口,则可执行其它操作。由于本专利重点是VLAN至VXLAN的组播数据报文转发,因此,组播组成员端口不为VXLAN隧道口的情况不是本专利的重点,这里不展开描述。

[0044] 本发明中,如上所述,网板找到的三层组播组表项中的组播组成员端口可不止一个,基于此,网板在找到的三层组播组表项中的组播组成员端口数量大于1时,由于网板只

接收到一份组播数据报文,基于此,网板可依据组播组成员端口的数量对组播数据报文进行复制,复制的组播数据报文的数量等于组播组成员端口的数量,之后,将复制的组播数据报文分配给各个组播组成员端口,以保证最终每一组播组成员端口都会分配到一份复制的组播数据报文,基于此,上述步骤202中所述的通过所述组播组成员端口发送的组播数据报文就为组播组成员端口分配的一份复制的组播数据报文。

[0045] 基于上面步骤201描述的第一内部封装,则较佳地,步骤202中,网板确定接收的组播数据报文由本网板进行三层组播转发可包括:

[0046] 从接收的组播数据报文的第一内部封装头中确定出第一目的mod和第一目的port;如果确定出的第一目的mod为上述的第一虚拟芯片的标识,第一目的port为上述的第一虚拟端口的标识,则对组播数据报文的第一内部封装进行解封装,如果解封装后的组播数据报文的目的地MAC地址为组播MAC地址、目的IP地址为组播IP地址,则确定接收的组播数据报文由本网板进行三层组播转发。

[0047] 作为本发明的一个实施例,则上述中,当确定目的mod不为上述的第一虚拟芯片的标识,和/或目的port不为上述的第一虚拟端口的标识,则可以不对组播数据报文的第一内部封装进行解封装,直接按照第一内部封装的目的port继续转发组播数据报文即可,这类类似现有网板转发组播数据报文的方式,不再展开描述。

[0048] 作为本发明的一个实施例,本步骤202中所述将修改后的组播数据报文发送至匹配接口板具体实现时包括:

[0049] 步骤b1,依据所述组播组成员端口确定目标出端口,选择所述目标出端口所在的接口板为匹配接口板。

[0050] 这里,依据所述组播组成员端口确定目标出端口可包括:

[0051] 步骤b11,在本地找到与所述组播组成员端口相关联的VXLAN隧道对应的下一跳表项;

[0052] 步骤b12,在所述下一跳表项中仅有一个下一跳时,确定所述下一跳表项中所述下一跳为目标下一跳,在所述下一跳表项中有两个以上下一跳时,从所述下一跳表项中的两个以上下一跳中选择一个下一跳确定为目标下一跳。这里,从所述下一跳表项中的两个以上下一跳中选择一个下一跳的方式可有多种方式,比如随机选择一个,或者按照设定的选择方式比如哈希算法等选择,本发明并不具体限定;

[0053] 步骤b13,在所述下一跳表项中与所述目标下一跳对应的出端口为单一的物理端口时,确定所述单一的物理端口为目标出端口;在所述下一跳表项中与所述目标下一跳对应的出端口为至少两个物理端口聚合形成的聚合端口,确定所述聚合端口中一个物理端口为所述目标出端口。

[0054] 这里,所述聚合端口中一个物理端口是从聚合端口中选择出的。其中,从聚合端口中选择一个物理端口的方式可有多种方式,比如随机选择一个,或者按照设定的选择方式比如哈希算法等选择,本发明并不具体限定。

[0055] 步骤b2,对修改后的组播数据报文进行第二内部封装,将第二内部封装的报文发送至匹配接口板;第二内部封装包含:组播组成员端口对应的VXLAN隧道封装表项索引和虚拟转发实例(VFI:Virtual Forwarding Instance)。

[0056] 优选地,组播组成员端口对应的VXLAN隧道封装表项索引一般存放在组播组成员

端口对应的VXLAN隧道封装表项中,组播组成员端口对应的VFI一般存放在组播组成员端口对应的VFI表项中。

[0057] 其中,VXLAN隧道封装表项、VFI表项可以是预先配置在网板上的,本发明并不具体限定。

[0058] 通过步骤b1和步骤b2的描述可以看出,最终,匹配接口板接收的来自网板的组播数据报文是执行了第二内部封装的报文。

[0059] 步骤203,网关VXLAN交换机上的匹配接口板接收网板发送的组播数据报文,对接收的组播数据报文进行VXLAN封装并向VXLAN转发。

[0060] 基于上面步骤202限定的网关VXLAN交换机上的网板如何将组播数据报文发送至网关VXLAN交换机上的匹配接口板,则优选地,本步骤203中,对接收的组播数据报文进行VXLAN封装并向VXLAN转发可包括:

[0061] 从组播数据报文的第二内部封装头中确定出VXLAN隧道封装表项索引和VFI并对组播数据报文的第二内部封装进行解封装,依据所述VXLAN隧道封装表项索引在本地找到对应的VXLAN隧道封装表项,从预先设置的VFI和VNID之间的映射关系中找到与该确定出的VFI具有映射关系的VNID,利用所述VXLAN隧道封装表项和找到的VNID对组播数据报文进行VXLAN封装并通过该找到的VXLAN隧道封装表项中的出端口发送。

[0062] 这里,利用所述VXLAN隧道封装表项和找到的VNID对组播数据报文进行VXLAN封装具体可为:利用所述VXLAN隧道封装表项中的VXLAN隧道封装信息对组播数据报文进行VXLAN外层头封装,利用找到的VNID对组播数据报文进行VXLAN字段(实质为VNID字段)的封装。这里VXLAN外层头至少包含外层以太(ETH)封装,其中,外层以太(ETH)封装至少包含外层源MAC地址、外层目的MAC地址、外层源IP地址、外层目的IP地址等。

[0063] 作为本发明的一个实施例,这里,通过该找到的VXLAN隧道封装表项中的出端口发送可为:识别该找到的VXLAN隧道封装表项中的出端口,如果该出端口为单一的物理端口,则直接通过该单一的物理端口发送,而如果该出端口为至少两个物理端口聚合形成的聚合端口,则需要从聚合端口中选择一个物理端口,通过该选择的物理端口发送。需要说明的是,这里从聚合端口中选择一个物理端口的方式不再随机,其必须最终保证选择的物理端口与上述网板在确定目标出端口时选择的物理端口一致。至于具体实现时如何保证接口板最终选择的物理端口与上述网板在确定目标出端口时选择的物理端口一致,本发明并不具体限定,其可有多种实现方式,比如网板按照哈希方式选择物理端口,接口板也同样按照哈希方式选择物理端口等。

[0064] 本发明中,作为一个实施例,接口板本地的VXLAN隧道封装表项可以是预先配置在接口板上的,本发明并不具体限定。

[0065] 至此,完成图2所示的流程。

[0066] 通过图2所示流程可以看出,本发明中,网关VXLAN交换机的网板不再使用仅具有报文转发功能的芯片,而是使用具有转发功能和表项功能的芯片比如与接口板相同的交换芯片,这样,当网关VXLAN交换机通过接口板接收到外部来自VLAN的组播数据报文,在本地的三层组播表中查找到与组播数据报文的目IP地址和目的MAC地址匹配的三层组播组表项,在所述三层组播组表项中存在的至少一个组播组成员端口为第一虚拟端口时,将所述组播数据报文发送至所述VXLAN交换机的网板,在本网板本地的三层组播组表中找到与组

播数据报文的目的IP地址和目的MAC地址匹配的三层组播组表项,在所述三层组播组表项中的组播组成员端口为VXLAN隧道口时,将通过所述组播组成员端口发送的组播数据报文的源MAC地址的修改为所述VXLAN交换机的网关MAC地址,将修改后的组播数据报文发送至VXLAN交换机上与所述组播组成员端口相关联的接口板,以由接口板对组播数据报文进行VXLAN封装并向VXLAN转发,即通过网关VXLAN交换机实现了VXLAN的三层网关功能,实现从VLAN至VXLAN的组播数据报文转发,且整个过程只是在网关VXLAN交换机内部执行,没有浪费任何带宽资源,做到了组播数据报文线速转发的。

[0067] 下面通过一个实施例对图2所示流程进行描述:

[0068] 以图3所示的组网为例,在图3中,物理机(PM:Physical Machine)1至PM5为VLAN中的设备,虚拟机(VM:Virtual Machine)1至VM5为VXLAN中的设备。网关VXLAN交换机连接在VLAN和VXLAN之间,其包含主控板、接口板和网板。主控板在本申请中不重点涉及,图3未示出。网板采用具有转发功能和表项功能的芯片,以与接口板中的交换芯片相同为例。图3仅示出VXLAN交换机的三个接口板为例。

[0069] 以PM1发送组播数据报文为例,则,

[0070] 网关VXLAN交换机的接口板1通过本地端口(图3以端口10为例)收到来自PM1的组播数据报文。为便于描述,这里将接收的组播数据报文称为报文0。

[0071] 接口板1发现报文0的目的MAC地址为组播MAC地址、目的IP地址为组播IP地址,则认为报文0为组播数据报文;

[0072] 接口板1在本地的三层组播表中查找到与报文0的目的IP地址和目的MAC地址匹配的三层组播组表项。

[0073] 假如接口板1发现三层组播组表项中的组播组成员端口有5个,其中有三个组播组成员端口为第一虚拟端口,剩下的两个组播组成员端口为VLAN端口,则,

[0074] 接口板1复制3份报文0,其中两份分配给两个为VLAN端口的组播组成员端口,剩下的一份分配给第一虚拟端口。为便于描述,这里将分配给第一虚拟端口的报文称为报文2。

[0075] 接口板1对报文2进行第一内部封装,第一内部封装的第一目的mod为第一虚拟芯片的标识(图3以chip01为例),第一内部封装的第一目的port为第一虚拟端口的标识(图3以Port1为例)。为便于描述,这里将第一内部封装后的报文2称为报文3。

[0076] 如图3所示,接口板1通过聚合了互连口11、互连口12、互连口13的聚合组与各网板连接,则接口板1从本接口板1连接各网板的聚合组中选择一个互连口,比如选择图3所示的互连口11将报文3发送至网板1。接口板1选择互连口的方式可预先设置,比如哈希算法,通过对组播组成员端口Port11的编号进行哈希运算,选择编号与哈希运算结果对应的互连口。

[0077] 网板1接收到报文3,从报文3的第一内部封装头中识别出第一目的mod和第一目的port分别为chip01和Port1,正好依次是第一虚拟芯片的标识和第一虚拟端口的标识,则对报文3的第一内部封装进行解封装(即恢复了上述的报文2)。

[0078] 网板1发现报文2的目的MAC地址为组播MAC地址、目的IP地址为组播IP地址,确定本网板1需要对报文2进行三层组播转发。

[0079] 网板1依据报文2的目的IP地址和目的MAC地址在本网板1本地的三层组播组表中查找到匹配的三层组播组表项。

[0080] 网板1发现三层组播组表项中存在三个组播组成员端口。假如这三个组播组成员端口分别记为Port11、Port12、Port13(这三个组播组成员端口在图3未示出)。

[0081] 网板1复制三份报文2,分别记为报文21、报文22、报文23,复制的这三份报文分别分配给三个组播组成员端口Port11、Port12、Port13。假如报文21分配给Port11、报文22分配给Port12、报文23分配给Port13。

[0082] 以Port11为例,Port12、Port13原理类似,网板1发现Port11为VXLAN隧道口,则将报文21的源MAC地址修改为网关VXLAN交换机的网关MAC地址;为便于描述,这里将源MAC地址修改后的报文21称为报文4。

[0083] 网板1在本地找到与VXLAN隧道口相关联的VXLAN隧道对应的下一跳表项,假如发现该找到的下一跳表项中仅有一个下一跳,且该下一跳的出端口为至少两个物理端口聚合形成的聚合端口,则选择聚合端口中一个物理端口为目标出端口,而该选择的物理端口连接的接口板为组播组成员端口Port11相关联的接口板,为便于描述,这里假设该接口板为接口板2。

[0084] 网板1在本地VXLAN隧道封装表项中找到组播组成员端口Port11对应的VXLAN隧道封装表项,以及在本地VFI表项中找到组播组成员端口Port11对应的VFI表项;

[0085] 网板1对报文4进行第二内部封装,为便于描述,这里将第二内部封装的报文4称为报文5。第二内部封装包含网板1找到的VXLAN隧道封装表项的索引、以及网板1找到的VFI表项中的VFI;

[0086] 网板1通过本网板1连接接口板2的互连口将报文5发送至接口板2。

[0087] 接口板2接收网板发送的报文5,从报文5的第二内部封装头中确定出VXLAN隧道封装表项索引和VFI并对报文5的第二内部封装进行解封装(即恢复了上述的报文4)。

[0088] 接口板2依据所述VXLAN隧道封装表项索引在本地找到对应的VXLAN隧道封装表项,从预先设置的VFI和VNID之间的映射关系中找到与该确定出的VFI具有映射关系的VNID(图3以VNID为VNID200为例),利用找到的VXLAN隧道封装表项和找到的VNID对报文4进行VXLAN封装。为便于描述,这里将VXLAN封装的报文4称为报文6。这里利用找到的VXLAN隧道封装表项和找到的VNID对报文4进行VXLAN封装具体为:利用找到的VXLAN隧道封装表项中的VXLAN隧道封装信息对报文4进行VXLAN外层头的封装,利用找到的VNID(图3以VNID为VNID200为例)对报文4进行VXLAN字段(实质为VNID字段)的封装。图3所示的报文4的VXLAN封装中以VXLAN外层头的外层ETH封装包含外层源MAC地址、外层目的MAC地址、外层源IP地址、外层目的IP地址为例。

[0089] 接口板2依据找到的VXLAN隧道封装表项中的出端口发送报文6。这里,具体实现时,接口板2依据找到的VXLAN隧道封装表项中的出端口发送报文6具体为:识别该找到的VXLAN隧道封装表项中的出端口,该识别的出端口实质是网板找到的一个下一跳的出端口,如上所述,网板找到的一个下一跳的出端口为至少两个物理端口聚合形成的聚合端口,基于此,这里识别出的VXLAN隧道封装表项中的出端口就为聚合端口,如此,接口板2按照网板选择目标出端口的方式从识别出的出端口中选择一个物理端口,最终保证发送报文6的端口实质与网板选择的目标出端口一致。

[0090] 假如Port11关联的组播组成员为VM11,则最终VXLAN的VM11会收到PM1发送的组播数据报文。类似的原理,假如Port12关联的组播组成员为VM12,Port13关联的组播组成员为

VM13,则采用类似于针对组播组成员端口Port11的处理方式处理Port12、Port13,VXLAN的VM12、VM13也会收到PM1发送的组播数据报文。即通过网关VXLAN交换机实现了从VLAN至VXLAN的组播数据报文转发。

[0091] 至此,完成图3所示的流程。

[0092] 以上对本发明提供的方法进行了描述,下面对本发明提供的设备进行描述:

[0093] 参见图4,图4为本发明提供的设备结构图。该设备应用于作为网关的VXLAN交换机,所述VXLAN交换机至少由接口板、网板组成,所述网板采用具有转发功能和表项功能的芯片,如图4所示,该设备包括:

[0094] 接口板处理单元,处于所述VXLAN交换机的接口板,用于接收外部来自VLAN的组播数据报文,在本地的三层组播表中查找到与组播数据报文的目IP地址和目的MAC地址匹配的三层组播组表项,在所述三层组播组表项中存在的至少一个组播组成员端口为第一虚拟端口时,将所述组播数据报文发送至所述VXLAN交换机的网板;以及,接收网板发送的组播数据报文,对接收的组播数据报文进行VXLAN封装并向VXLAN转发;

[0095] 网板处理单元,处于所述VXLAN交换机的网板,用于接收接口板发送的组播数据报文,确定接收的组播数据报文由本网板进行三层组播转发,则在本地三层组播组表中找到与组播数据报文的目IP地址和目的MAC地址匹配的三层组播组表项,在找到的三层组播组表项中的组播组成员端口为VXLAN隧道口时,将通过所述组播组成员端口发送的组播数据报文的源MAC地址的修改为所述VXLAN交换机的网关MAC地址,将修改后的组播数据报文发送至VXLAN交换机上与所述组播组成员端口相关联的接口板。

[0096] 优选地,所述接口板处理单元通过以下步骤将组播数据报文发送至所述VXLAN交换机的网板:

[0097] 对组播数据报文进行第一内部封装,第一内部封装包含:第一目的芯片标识mod和第一目的端口标识port,第一目的mod为第一虚拟芯片的标识,第一目的port为第一虚拟端口的标识;

[0098] 在本接口板仅连接一个网板时,通过本接口板连接网板的互连口发送第一内部封装的组播数据报文;在本接口板连接多个网板时,通过与多个网板分别连接的互连口中的一个互连口发送第一内部封装的组播数据报文。

[0099] 优选地,所述网板处理单元通过以下步骤确定接收的组播数据报文由本网板进行三层组播转发:从接收的组播数据报文的第一内部封装头中确定出第一目的mod和第一目的port;如果确定出的第一目的mod为第一虚拟芯片的标识,第一目的port为第一虚拟端口的标识,则对组播数据报文的第一内部封装进行解封装,如果解封装后的组播数据报文的目MAC地址为组播MAC地址、目IP地址为组播IP地址,则确定接收的组播数据报文由本网板进行三层组播转发。

[0100] 优选地,所述网板处理单元通过以下步骤将修改后的组播数据报文发送至VXLAN交换机上与所述组播组成员端口相关联的接口板:

[0101] 依据所述组播组成员端口确定目标出端口,选择所述目标出端口所在的接口板为与所述组播组成员端口相关联的接口板;对修改后的组播数据报文进行第二内部封装,将第二内部封装的报文发送至所述目标出端口所在的接口板;第二内部封装包含:所述组播组成员端口对应的VXLAN隧道封装表项索引和虚拟转发实例VFI;

[0102] 优选地,所述接口板处理单元通过以下步骤对接收的组播数据报文进行VXLAN封装并向VXLAN转发:

[0103] 从组播数据报文的第二内部封装头中确定出VXLAN隧道封装表项索引和VFI并对组播数据报文的第二内部封装进行解封装,依据所述VXLAN隧道封装表项索引在本地找到对应的VXLAN隧道封装表项,从预先设置的VFI和VNID之间的映射关系中找到与该确定出的VFI具有映射关系的VNID,利用所述VXLAN隧道封装表项和找到的VNID对组播数据报文进行VXLAN封装并通过该找到的VXLAN隧道封装表项中的出端口发送。

[0104] 优选地,所述网板处理单元通过以下步骤依据所述组播组成员端口确定目标出端口:

[0105] 在本地找到与所述组播组成员端口相关联的VXLAN隧道对应的下一跳表项;

[0106] 在所述下一跳表项中仅有一个下一跳时,确定所述下一跳表项中所述下一跳为目标下一跳,在所述下一跳表项中有两个以上下一跳时,从所述下一跳表项中的两个以上下一跳中选择一个下一跳确定为目标下一跳;

[0107] 在所述下一跳表项中与所述目标下一跳对应的出端口为单一的物理端口时,确定所述单一的物理端口为目标出端口;在所述下一跳表项中与所述目标下一跳对应的出端口为至少两个物理端口聚合形成的聚合端口,确定所述聚合端口中一个物理端口为所述目标出端口。

[0108] 至此,完成图4所示的设备结构图。

[0109] 作为本发明的一个实施例,本发明还提供了图4所示设备的硬件结构图。如图5所示,该设备硬件结构可包括:

[0110] 接口板和网板。

[0111] 其中,接口板上可包括:第一CPU和第一存储器。

[0112] 第一存储器,用于存放接口板处理单元;

[0113] 第一CPU,用于存放接口板处理单元运行的控制程序,以控制所述第一存储器中的接口板处理单元执行如上所述的操作。至于所述第一存储器中的接口板处理单元执行的操作,其如上述接口板处理单元执行的操作,这里不再赘述。

[0114] 网板上可包括:第二CPU和第二存储器。

[0115] 第二存储器,用于存放网板处理单元;

[0116] 第二CPU,用于存放网板处理单元运行的控制程序,以控制所述第二存储器中的网板处理单元执行操作。至于所述第二存储器中的网板处理单元执行的操作,其如上述网板处理单元执行的操作,这里不再赘述。

[0117] 至此,完成图5所示的设备硬件结构图。

[0118] 至此,完成本发明提供的设备描述。

[0119] 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。

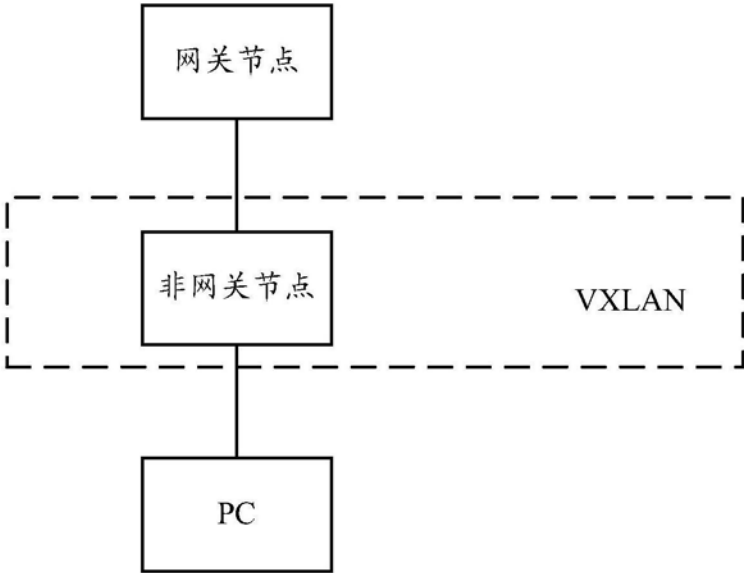


图1

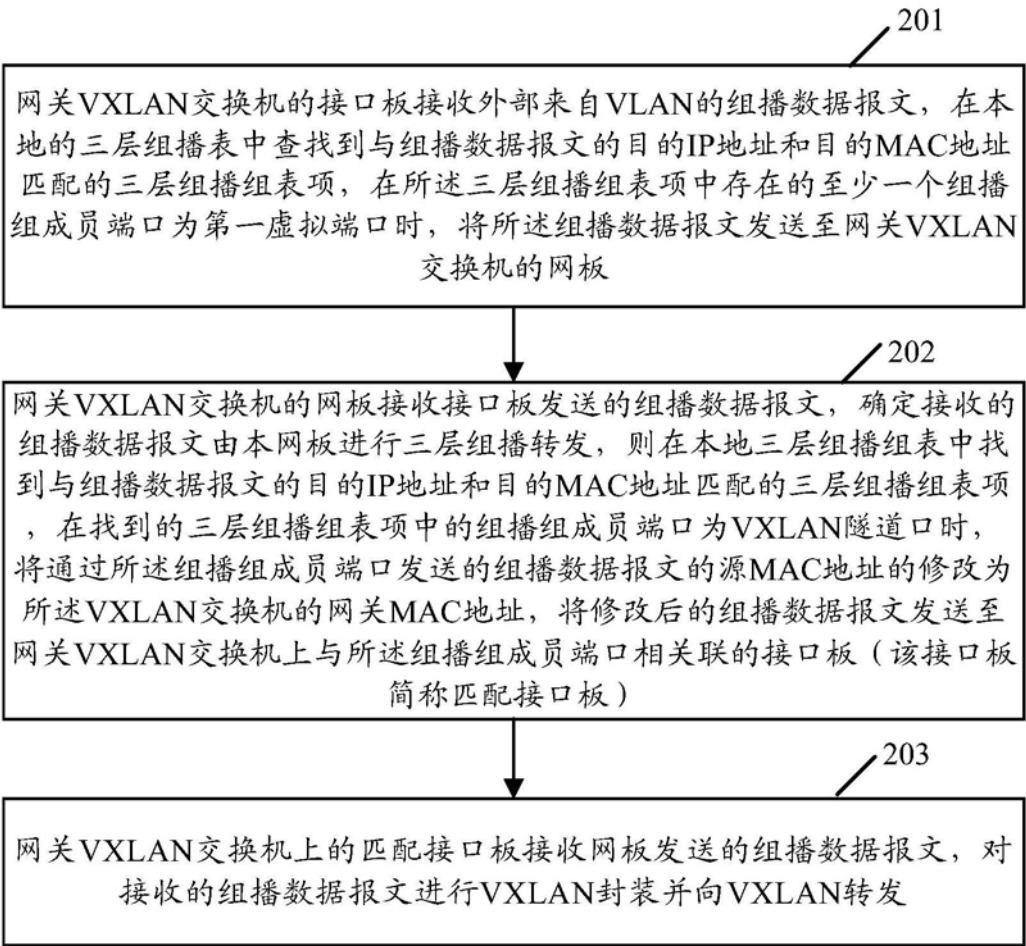


图2

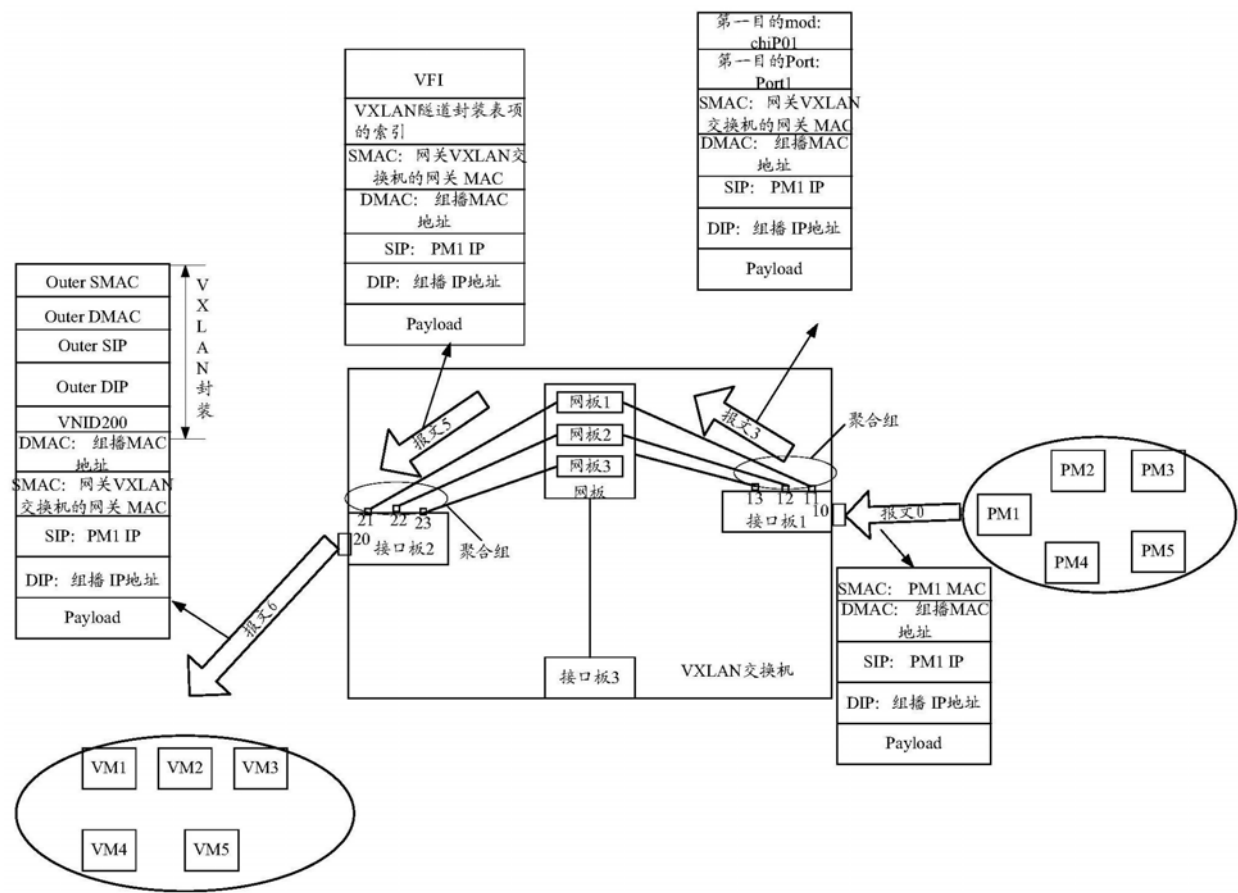


图3

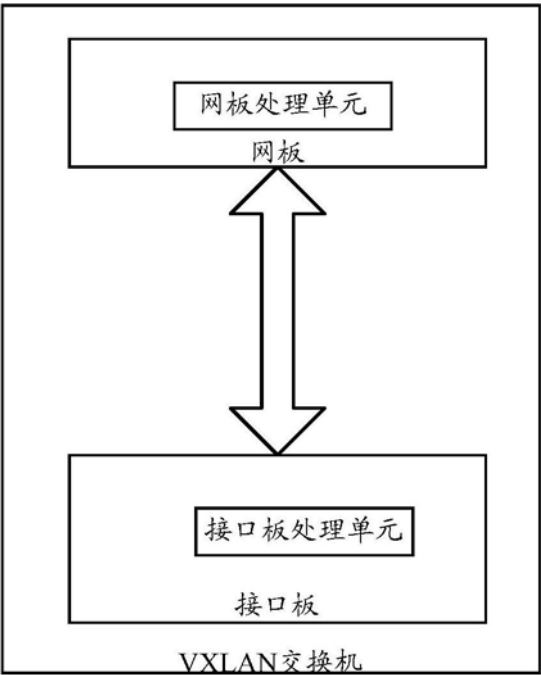


图4

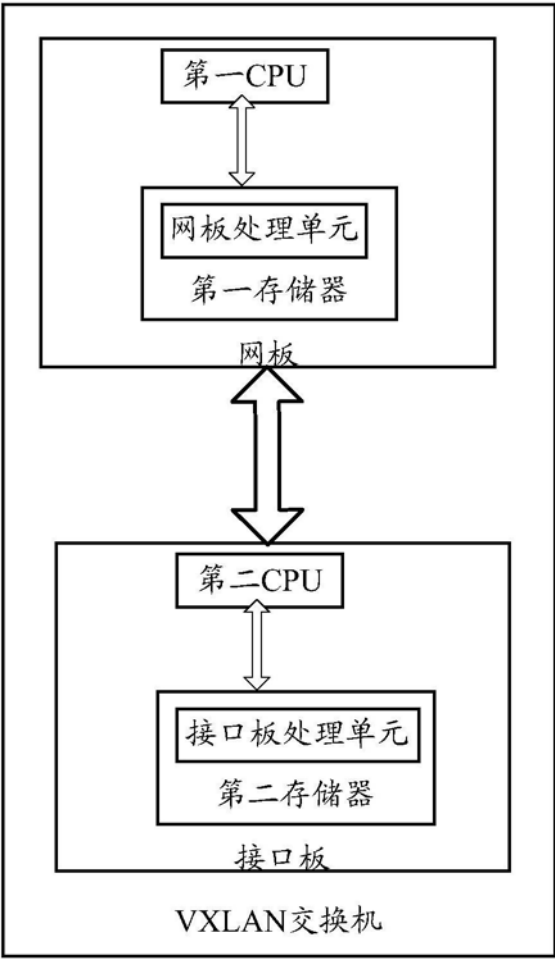


图5