



(12) 发明专利

(10) 授权公告号 CN 1939004 B

(45) 授权公告日 2012. 07. 18

(21) 申请号 200580010032. 9

迈克尔·R·史密斯 艾力·戈尔杉

(22) 申请日 2005. 04. 05

(74) 专利代理机构 北京东方亿思知识产权代理
有限责任公司 11258

(30) 优先权数据

10/826, 888 2004. 04. 16 US

代理人 王怡

(85) PCT申请进入国家阶段日

2006. 09. 28

(51) Int. Cl.

H04L 12/56 (2006. 01)

(86) PCT申请的申请数据

PCT/US2005/011588 2005. 04. 05

(56) 对比文件

WO 03081451 A1, 2003. 10. 02,

US 5825772 A, 1998. 10. 20,

US 2002103921 A1, 2002. 08. 01,

(87) PCT申请的公布数据

W02005/107188 EN 2005. 11. 10

审查员 陈红英

(73) 专利权人 思科技术公司

地址 美国加利福尼亚州

(72) 发明人 斯塔拉姆·东图 珐塞·穆萨

士尔曼·博海德

权利要求书 3 页 说明书 17 页 附图 14 页

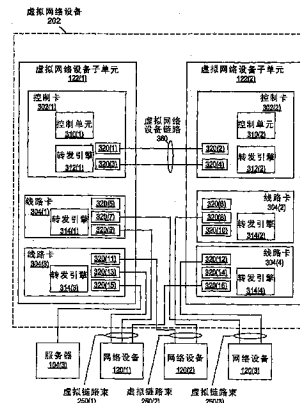
(54) 发明名称

虚拟网络设备中的分布式转发

(57) 摘要

虚拟网络设备子单元 (122(1)、122(2)) 包括到虚拟网络设备链路 (360) 的接口 (320(1)、320(2)、320(3)、320(4)) 和分布式转发模块 (312(1)、312(2))。接口 (320(1)、320(2)、320(3)、320(4)) 接收分组, 并且分布式转发模块 (312(1)、312(2)) 转发由接口 (320(1)、320(2)、320(3)、320(4)) 接收的分组。如果分组包括多播目的地地址, 则分布式转发模块 (312(1)、312(2)) 执行进入查找, 而如果分组包括单播目的地地址, 则分布式转发模块 (312(1)、312(2)) 执行外出查找。如果分组包括多播目的地地址, 分布式转发模块则为与多播目的地地址相关联的若干个传出 VLAN 中的每一个复制分组。如果经由未被耦合到虚拟网络设备链路 (306) 的接口 (320(1)、320(2)、320(3)、320(4)) 接收到额外的多播分组, 则分布式转发模块 (312(1)、312(2)) 经由虚拟网络设备链路 (360) 发送额外的多播分组的至多一个拷贝。

CN 1939004 B



1. 一种用于在虚拟网络设备中执行分布式转发的方法,所述虚拟网络设备包含经由虚拟网络设备链路互连的至少两个网络设备并被配置为作为单个虚拟网络设备工作,所述方法包括:

经由所述虚拟网络设备链路接收分组,所述分组包括单播目的地地址;以及
响应于接收所述分组在第一线路卡上的查找表中对所述分组执行外出查找,
其中,所述执行外出查找包括分配非初级条目,所述非初级条目指示第二线路卡。

2. 如权利要求 1 所述的方法,其中

与所述分组相关联的头部也经由所述虚拟网络设备链路被接收;并且
所述头部包括目的地标识符。

3. 如权利要求 2 所述的方法,还包括:

如果在所述外出查找期间找到与所述单播目的地地址相对应的所述非初级条目,则将
所述分组和所述头部发送到所述第二线路卡,其中

所述非初级条目指示所述第二线路卡包含在所述第二线路卡上的接口,所述分组应该
经由该接口被发送。

4. 如权利要求 2 所述的方法,还包括:

如果在所述外出查找期间找到与所述单播目的地地址相对应的初级条目,则:

从由所述初级条目所标识的接口发送所述分组,其中

所述初级条目指示所述第一线路卡上的接口,所述分组应该经由该接口被发送。

5. 如权利要求 4 所述的方法,还包括:

如果所述头部中包括的所述目的地标识符与所述初级条目中包括的目的地标识符不
匹配,则经由所述虚拟网络设备链路发送通知,其中

所述通知将所述单播目的地地址标识为与所述初级条目中包括的所述目的地标识符
相对应。

6. 如权利要求 5 所述的方法,其中

仅在所述接口未被包括在上行链路接口束中的情况下,所述分组才从所述接口被发
送。

7. 如权利要求 1 所述的方法,还包括:

接收第二分组,所述第二分组包括多播目的地地址;以及

经由虚拟网络设备链路将所述第二分组的至多一个拷贝发送到虚拟网络设备子单元,
所述虚拟网络设备子单元被包括在虚拟网络设备中。

8. 如权利要求 7 所述的方法,还包括:

经由所述虚拟网络设备链路接收第三分组,所述第三分组包括第二多播目的地地址;
以及

为与所述第二多播目的地地址相关联的多个传出 VLAN(虚拟局域网)中的每一个复制
所述第三分组。

9. 一种用于在虚拟网络设备中执行分布式转发的方法,所述虚拟网络设备包含经由虚拟网络设备链路互连的至少两个网络设备并被配置为作为单个虚拟网络设备工作,所述方法包括:

经由所述虚拟网络设备链路接收分组;

对所述分组执行进入查找和外出查找之一,其中

如果所述分组包括多播目的地地址,则对所述分组执行所述进入查找;并且

如果所述分组包括单播目的地地址,则在第一线路卡上的查找表中对所述分组执行所述外出查找,

其中,所述执行进入查找包括分配初级条目,所述初级条目指示在所述第一线路卡上的接口,并且

所述执行外出查找包括分配非初级条目,所述非初级条目指示第二线路卡。

10. 一种用于在虚拟网络设备中执行分布式转发的系统,所述虚拟网络设备包含经由虚拟网络设备链路互连的至少两个网络设备并被配置为作为单个虚拟网络设备工作,所述系统包括:

到所述虚拟网络设备链路的接口,其中

所述接口被配置为接收分组;以及

被耦合到所述接口的分布式转发模块,其中

所述分布式转发模块被配置为转发所述分组,

所述分布式转发模块被配置为在所述分组包括单播目的地地址的情况下对所述分组执行外出查找,并且

所述执行外出查找包括分配非初级条目,所述非初级条目指示第二线路卡。

11. 如权利要求 10 所述的系统,其中

所述分布式转发模块被配置为在所述分组包括多播目的地地址的情况下对所述分组执行进入查找,其中

所述执行进入查找包括分配初级条目,所述初级条目指示在所述第一线路卡上的接口。

12. 如权利要求 11 所述的系统,还包括:

被配置为接收第二分组的第二接口,其中

所述第二分组包括第二多播地址,并且

所述分布式转发模块被配置为经由所述虚拟网络设备链路发送所述第二分组的至多一个拷贝。

13. 如权利要求 10 所述的系统,其中

与所述分组相关联的头部也经由所述虚拟网络设备链路被接收,

所述头部包括目的地标识符,

所述分组包括所述单播目的地地址,并且

所述分布式转发模块被配置为在在所述外出查找期间找到与所述单播目的地地址相对应的所述非初级条目的情况下将所述分组和所述头部发送到所述第二线路卡,其中

所述非初级条目指示所述第二线路卡包含在所述第二线路卡上的接口,所述分组应该经由该接口被发送。

14. 如权利要求 13 所述的系统,其中

所述分布式转发模块被配置为在所述头部中包括的目的地标识符与所述初级条目中包括的目的地标识符不匹配的情况下经由所述虚拟网络设备链路发送通知,其中

所述初级条目指示所述第一线路卡上的接口,所述分组应该经由该接口被发送,并且所述通知将所述单播目的地地址标识为与所述初级条目中包括的所述目的地标识符相对

应。

15. 一种用于在虚拟网络设备中执行分布式转发的系统,所述虚拟网络设备包含经由虚拟网络设备链路互连的至少两个网络设备并被配置为作为单个虚拟网络设备工作,所述系统包括:

用于经由所述虚拟网络设备链路接收分组的装置,所述分组包括单播目的地地址;以及

用于在第一线路卡上的查找表中对所述分组执行外出查找的装置,

其中,所述执行外出查找包括分配指示非初级条目,所述非初级条目指示第二线路卡。

16. 如权利要求 15 所述的系统,其中

与所述分组相关联的头部也经由所述虚拟网络设备链路被接收;并且

所述头部包括通过对所述分组执行进入查找而获得的目的地标识符,其中

所述执行进入查找包括分配初级条目,所述初级条目指示在所述第一线路卡上的接口。

17. 如权利要求 16 所述的系统,还包括:

用于在在所述外出查找期间找到与所述单播目的地地址相对应的所述非初级条目的情况下将所述分组和所述头部发送到所述第二线路卡的装置,其中

所述非初级条目指示所述第二线路卡包含在所述第二线路卡上的接口,所述分组应该经由该接口被发送。

18. 如权利要求 16 所述的系统,还包括:

用于在在所述外出查找期间找到与所述单播目的地地址相对应的初级条目的情况下从由所述初级条目所标识的接口发送所述分组的装置,其中

所述初级条目指示所述第一线路卡上的接口,所述分组应该经由该接口被发送。

19. 如权利要求 18 所述的系统,还包括:

用于在所述头部中包括的所述目的地标识符与所述初级条目中包括的目的地标识符不匹配的情况下经由所述虚拟网络设备链路发送通知的装置,其中

所述通知将所述单播目的地地址标识为与所述初级条目中包括的所述目的地标识符相对应。

20. 如权利要求 18 所述的系统,其中

仅在所述接口未被包括在上行链路接口束中的情况下所述分组才从所述接口被发送。

21. 如权利要求 16 所述的系统,还包括:

用于接收第二分组的装置,所述第二分组包括多播目的地地址;以及

用于经由虚拟网络设备链路将所述第二分组的至多一个拷贝发送到虚拟网络设备子单元的装置,所述虚拟网络设备子单元被包括在虚拟网络设备中。

虚拟网络设备中的分布式转发

技术领域

[0001] 本发明涉及联网,更具体而言,涉及在虚拟网络设备中执行分布式转发。

背景技术

[0002] 为了提供更高的网络可靠性,常常在网络中包括冗余的交换机和链路。如果一个交换机或链路发生故障,已经在网络内就位的冗余交换机或链路可以迅速被启用以替换发生故障的交换机或链路。由于与替换或维修发生故障的组件相比,启用冗余交换机或链路作为替换通常更为迅速,因此拥有冗余的链路和 / 或交换装置能够提供更为可靠的网络。

[0003] 当在网络内包括冗余的组件时,常常希望能够在相应组件发生故障之前在正常网络工作期间使用冗余的组件。例如,如果在一对交换机之间实现了两条链路,则希望使用两条链路(而不是使一条链路空闲)来提供更大的带宽。但是,如果多条冗余链路同时活动,则对这些链路的管理的复杂程度可能是令人不快的(例如由于需要避免桥接环)。这种复杂性扩展到在正常操作期间使用多个冗余组件的其他情形。例如,如果多个冗余路由器被同时用于网络中,则由于需要对每个网络设备有不同的管理点,因而对网络的管理可能变得更加复杂。这些示例表明,希望能够降低由于在网络中使用多个冗余组件而导致的复杂性。

发明内容

[0004] 公开了用于在虚拟网络设备中执行分布式转发的方法和系统的各种实施例。虚拟网络设备包括若干个不同的虚拟网络设备子单元,它们共同作为单个逻辑网络设备工作。

[0005] 一种在虚拟网络设备内执行分布式转发的方法,包括经由虚拟网络设备链路接收分组。该分组包括单播目的地地址。响应于接收分组对分组执行外出查找(egress lookup)。如果在外出查找期间找到与单播目的地地址相对应的初级条目(primary entry),则从由初级条目所标识的接口发送分组。与分组相关联的头部也经由虚拟网络设备链路与分组一起被接收。头部包括目的地标识符。如果头部中的目的地标识符与初级条目中的目的地标识符不匹配,则经由虚拟网络设备链路发送差错通知。差错通知标识单播目的地地址与初级条目中的目的地标识符相对应。

[0006] 在虚拟网络设备中执行分布式转发的另一种方法包括经由虚拟网络设备链路接收分组并且对分组执行进入查找和外出查找之一。如果分组包括多播目的地地址,则对分组执行进入查找。如果分组包括单播目的地地址,则对分组执行外出查找。

[0007] 一种系统(例如虚拟网络设备子单元)包括到虚拟网络设备链路的接口和被耦合到接口的分布式转发模块。接口被配置为接收分组。分布式转发模块被配置为转发由接口接收的分组。在一个实施例中,分布式转发模块被配置为在分组包括多播目的地地址的情况下对分组执行进入查找,并且在分组包括单播目的地地址的情况下对分组执行外出查找。

[0008] 如果分组包括多播目的地地址,则分布式转发模块为与多播目的地地址相关联的

若干传出 VLAN 中的每一个复制分组。该系统还包括被配置为接收第二分组的第二接口（不是到虚拟网络设备链路的接口），该第二分组包括第二多播地址。分布式转发模块被配置为经由虚拟网络设备链路发送第二分组的至多一个拷贝。

[0009] 如果分组包括单播目的地地址，则分布式转发模块在外出查找期间找到与单播目的地地址相对应的非初级条目的情况下将分组和与分组相关联的头部发送到另一个线路卡。如果在外出查找期间找到与单播目的地地址相对应的初级条目并且该初级条目标识第二接口，则分布式转发模块从第二接口发送分组。分布式转发模块被配置为在头部中的目的地标识符与初级条目中的目的地标识符不匹配的情况下经由虚拟网络设备链路发送通知。通知标识单播目的地地址与初级条目中的目的地标识符相对应。

[0010] 以上只是一个概要，从而必然包括简化、概括和细节省略；因此，本领域的技术人员将会意识到，概要只是说明性的，而绝不想要是限制性的。这里公开的操作可以用多种方式来实现在，并且可以在不脱离本发明及其更宽的方面的情况下作出这种变化和修改。仅由权利要求书限定的本发明的其他方面将从以下阐述的非限制性详细描述中清楚显现出来。

附图说明

[0011] 通过参考下面的描述和附图可以获取对本发明的更彻底的理解，附图中类似的标号指示类似的特征。

[0012] 图 1 是根据本发明一个实施例的网络的框图。

[0013] 图 2A 和 2B 示出根据本发明一个实施例，同一网络层中的两个网络设备如何共同作为单个虚拟网络设备工作。

[0014] 图 3 示出根据本发明一个实施例，包括在虚拟网络设备中的每个虚拟网络设备子单元内的更多细节。

[0015] 图 4A-4F 示出根据本发明一个实施例，共同作为单个虚拟网络设备工作的两个虚拟网络设备子单元对具有多播目的地地址的分组执行分布式转发的方式。

[0016] 图 5A 是根据本发明一个实施例，由虚拟网络设备子单元用来将多播分组的至多一个拷贝转发到同一虚拟网络设备中的每个其他虚拟网络设备子单元的方法的流程图。

[0017] 图 5B 是由虚拟网络设备子单元用来转发接收自同一虚拟网络设备中的另一个虚拟网络设备子单元的多播分组的方法的一个实施例的流程图。

[0018] 图 6A-6K 示出根据本发明一个实施例，共同作为单个虚拟网络设备工作的两个虚拟网络设备子单元如何对具有单播目的地地址的分组执行分布式转发。

[0019] 图 7 是根据本发明一个实施例，虚拟网络设备子单元可用来转发接收自同一虚拟网络设备中的另一个虚拟网络设备子单元的单播分组的方法的流程图。

[0020] 图 8 是虚拟网络设备子单元可用来转发接收自同一虚拟网络设备中的另一个虚拟网络设备子单元的单播分组的方法的一个实施例的流程图。

[0021] 虽然本发明能够接受各种修改和替换形式，但在附图和详细描述中提供了本发明的特定实施例作为示例。应当理解，附图和详细描述并不是用来将本发明限制到所公开的特定形式的。相反，本发明将会覆盖处于由所附权利要求书所限定的本发明的精神和范围之内所有修改、等同物和替换。

具体实施方式

[0022] 图 1 是一个网络的框图。在图 1 中,若干个客户端 102(1)-102(n) 经由网络彼此通信,并且与若干个服务器 104(1)-104(n) 通信。客户端 102(1)-102(n) 可以包括访问互联网的服务的多种不同设备。例如,客户端 102(1) 可以是蜂窝电话,客户端 102(2) 可以是个人计算机,客户端 102(n) 可以是个人数字助理 (PDA)。服务器 104(1)-104(n) 提供各种服务,例如各种基于软件的服务和 / 或对共享的存储设备的访问。

[0023] 耦合客户端 102(1)-102(n) 和服务器 104(1)-104(n) 的网络是按照若干网络层来描述的。与客户端 102(1)-102(n) 最接近的层是访问层 110。访问层 110 包括若干个网络设备 120(1)-120(n)。在该示例中,访问层 110 是分组从客户端 102(1)-102(n) 进入网络的初级层。

[0024] 分发层 112 聚集经由访问层 110 接收到的流,并将这些聚集的流提供到核心层 114。在该示例中,分发层 112 包括网络设备 122(1)-122(n)。核心层 114 是网络的逻辑上集中的部分,各种聚集的流经过该层。核心层 114 包括网络设备 124(1)-124(n)。

[0025] 在该示例中,数据中心 116 包括两组网络设备:网络设备 126(1)-126(n) 和网络设备 128(1)-128(n)。网络设备 128(1)-128(n) 向各种服务器 104(1)-104(n) 提供对网络的访问。网络设备 126(1)-126(n) 聚集来自网络设备 128(1)-128(n) 的流,并且将聚集的流提供到核心层 114。

[0026] 注意,在某些实施例中,网络将不会包括图 1 所示的网络层(例如可以组合和 / 或消除某些层,并且附加于和 / 或取代图 1 所示的那些,还可以包括替换层)。此外,客户端和服务器可以按不同于图 1 所示的方式被耦合到网络(例如某些客户端和 / 或服务器可以耦合到核心和 / 或分发层中的个体网络设备)。此外,设备相对于彼此的物理位置可以不同于图 1 所示的逻辑位置。例如,相同网络层中的两个设备物理上可以位于不同的楼层上、不同的建筑中或不同的校园中。相反,不同网络层中的两个设备可以位于同一房间中。

[0027] 在某些实施例中,位于网络的外边缘的网络设备 120(1)-120(n) 和网络设备 128(1)-128(n) 的工作方式不同于位于网络的较内层的网络设备 122(1)-122(n)、124(1)-124(n) 和 126(1)-126(n)。例如,在一个实施例中,网络设备 120(1)-120(n) 是受网络的较内层(例如分发和核心层)中的网络设备控制或以其他方式从属于网络的较内层中的网络设备的附属网络设备。在这种实施例中,非附属网络设备提供 L2(第 2 层)和 L3(第 3 层)转发和路由选择,而附属网络设备只具有相对有限的转发和 / 或路由选择能力。在其他实施例中,附属网络设备不执行任何 L2 转发或 L3 路由选择。相反,附属网络设备只是简单地将所有分组转发到非附属网络设备,以进行 L2 转发和 L3 路由选择。在某些实施例中,被耦合到附属网络设备的非附属网络设备可以控制附属网络设备的工作。例如,网络设备 126(1)-126(n) 根据各种路由选择协议来配置网络设备 128(1)-128(n)。在某些实施例中,附属网络设备被视为附属所从属的网络设备的远程线路卡。还注意到,在替换实施例中,非附属网络设备可以取代附属网络设备被用于访问层和数据中心中。

[0028] 网络设备 120(1)-120(n)、122(1)-122(n)、124(1)-124(n)、126(1)-126(n) 和 128(1)-128(n) 可以包括各种路由器、交换机、网关和其他网络设备。在许多实施例中,要使网络工作,在每一层可能只需要一个网络设备。但是,如图 1 所示,在每一层可以包括多个网络设备,以提供冗余性。

[0029] 将会注意到,在这里描述的附图中的若干个实例中,使用了变量标识符“n”来更简单地指明一系列相关或类似元件的最后一个元件。这种变量标识符的反复使用并不一定是要暗示这种元件系列的大小之间的相关性,虽然这种相关性也可能是存在的。这种变量标识符的使用不要求每个元件系列与由相同的变量标识符定界的另一系列具有相同数目的元件(例如每个网络层中的网络设备的数目可能不同)。更确切地说,在每个使用实例中,由“n”(或任何其他这种标识符)标识的变量与同一变量标识符的其他实例相比可能拥有相同或不同的值。

[0030] 在不同网络层中的设备之间可以实现多个链路以提供额外的冗余性。例如,如图1所示,访问层110中的每个网络设备120(1)-120(n)可以被两条(或更多条)不同的链路耦合到分发层112。类似地,分发层112中的每个网络设备122(1)-122(n)可以被两条(或更多条)不同的链路耦合到核心层114。在一个实施例中,每条链路是一条以太网链路。

[0031] 在每个网络层内,多个冗余网络设备可以被配置为共同作为单个虚拟网络设备工作。例如,如图1所示,分发层112中的两个或更多个网络设备可以作为虚拟网络设备202工作。类似地,网络设备124(1)-124(n)中的两个或更多个可以作为单个虚拟网络设备204工作,而网络设备126(1)-126(n)中的两个或更多个可以作为单个虚拟网络设备206工作。关于两个分发层网络设备如何能够共同作为分发层虚拟网络设备202工作的更多细节在图2A、2B和3中示出。如下所述,虚拟网络设备可以被虚拟链路束耦合到其他虚拟网络设备、网络设备以及/或者客户端和/或服务器。一般来说,任何多端口设备(不论是物理设备,例如网络设备、客户端或服务器,还是虚拟网络设备)都可以被包括若干链路的虚拟链路束耦合到虚拟网络设备,所述链路中的某些在虚拟网络设备内的不同子单元上终止。

[0032] 图2A示出了一个网络的示例,在该网络中,在访问层110中有两个网络设备120(1)和120(2)。在分发层112中也有两个网络设备122(1)和122(2)。在该示例中,这两个网络设备122(1)和122(2)作为单个虚拟网络设备202工作。每个网络设备120(1)-120(2)被两条链路耦合到分发层112。在该示例中,这两条链路中的每一条被耦合到网络设备122(1)和122(2)中不同的一个。这提供了冗余性,从而即使在网络设备122(1)或122(2)之一发生故障或者给定访问层网络设备与给定分发层网络设备之间的链路之一发生故障的情况下也允许网络设备120(1)和120(2)继续与分发层112通信。

[0033] 将网络设备120(1)和120(2)中的每一个耦合到虚拟网络设备202的冗余链路可被当作单个逻辑链路来操作,在这里将该单个逻辑链路称为虚拟链路束。网络设备120(1)将把网络设备120(1)耦合到虚拟网络设备202的两条链路当作虚拟链路束250(1)来操作。在这种实施例中,网络设备120(1)中的每个被耦合到链路之一的接口被包括在接口束中,该接口束对应于虚拟链路束250(1)。类似地,网络设备120(2)将把网络设备120(2)耦合到虚拟网络设备202的两条链路当作虚拟链路束250(2)来操作。在某些实施例中,虚拟链路束250(1)和250(2)各自被当作EtherChannel(TM)或聚集链路(在IEEE 802.3中描述)来操作。在其他实施例中,每个虚拟链路束中的每条链路与该虚拟链路束中的每条其他链路相独立地被操作(而不是被当作单条逻辑链路的一部分操作)。

[0034] 如图2A所示,每个虚拟链路束250(1)和250(2)包括在分发层112中的不同网络设备处终止的链路。例如,虚拟链路束250(1)将网络设备120(1)耦合到网络设备122(1)和网络设备122(2)。这与传统的实现方式不同,在传统实现方式中,逻辑链路只允许存在于

单对网络设备之间。

[0035] 在某些实施例中,网络设备 120(1) 和 120(2) 意识到(例如,通过每个网络设备内维护的各种状态信息)每个虚拟链路束 250(1) 和 250(2) 包括在分发层 112 中的不同网络设备上终止的链路。在这种实施例中,网络设备 120(1) 和 120(2) 可以基于这种意识选择特定虚拟链路束内的一条链路以在其上发送分组。

[0036] 在其他实施例中,网络设备 122(1) 和 122(2) 进行操作以掩盖这种单逻辑链路实际上包括在不同网络设备处终止的链路这一事实。例如,如图 2A 所示,网络设备 122(1) 和 122(2) 作为单个虚拟网络设备 202 工作。图 2B 示出从访问层 110 中的网络设备 120(1) 的角度来看网络设备 120(1) 是如何被冗余链路对耦合到分发层 112 中的单个网络设备,即虚拟网络设备 202 的。网络设备 120(2) 对于虚拟网络设备 202 具有类似的观点。

[0037] 在网络设备 120(1) 和网络设备 120(2) 将其自身看成是被连接到单个网络设备的实施例中,例如图 2B 所示的实施例中,虚拟链路束的使用得以简化。例如,如果网络设备 120(1) 意识到虚拟链路束 250(1) 在两个不同的网络设备处终止,则网络设备 120(1) 可以基于生成树协议选择一条链路以在其上发送特定分组。与网络设备 120(1) 就简单地将虚拟网络设备 202 看作单个实体相比,使用生成树协议可能涉及更多开销,以及/或者对于哪些链路能被用来发送给定分组有更多的限制(例如,生成树协议可能阻塞除一条链路之外的所有链路,从而防止使用除一条未被阻塞的链路之外的所有链路)。例如,当将虚拟网络设备 202 看作单个实体时,网络设备 120(1) 可以简单地基于负载共享约束来选择一条链路以在其上发送分组。类似地,如果虚拟链路束 250(1) 内的链路发生故障,则网络设备 120(1) 无需改变生成树协议的应用方式。相反,网络设备 120(1) 可以就简单地继续使用虚拟链路束 250(1) 内的未发生故障的链路。

[0038] 虚拟网络设备 202 中包括的个体网络设备,例如网络设备 122(1) 和网络设备 122(2),在这里各自被称为“虚拟网络设备子单元”。在某些实施例中,虚拟网络设备子单元 122(1) 和 122(2) 各自是在单独的机箱中实现的(即,每个机箱容纳单个虚拟网络设备子单元)。例如,在图 2A 中,网络设备 122(1) 和 122(2) 可以各自具有其自己的机箱。即使虚拟网络设备子单元 122(1) 和 122(2) 共享机箱,也可以使每个虚拟网络设备子单元作为独立的网络设备工作,从而允许在虚拟网络设备中的一个(或多个)虚拟网络设备子单元发生故障的情况下其他虚拟网络设备子单元能够继续工作。例如,在一个实施例中,虚拟网络设备子单元 122(1) 和虚拟网络设备子单元 122(2) 共享相同的机箱,但是每个虚拟网络设备子单元具有独立的硬件、端口、上行链路接口和电源,并且每个虚拟网络设备子单元可以独立于另一个地从机箱中移除。如果虚拟网络设备子单元 122(1) 发生故障(例如由于电源故障或软件差错),虚拟网络设备子单元 122(2) 可以继续运行。在这种实施例中,可以移除虚拟网络设备子单元 122(1) 以便维修或替换,而不会干扰虚拟网络设备子单元 122(2) 的工作。

[0039] 在某些实施例中,将网络设备耦合到附属网络设备的虚拟链路束中的链路是用于将附属网络设备耦合到虚拟网络设备的专门链路,在这里被称为上行链路。每个上行链路可以输送网络设备之一内生成的分组和额外的信息。如果分组正在上行链路上从访问层附属网络设备被输送到分发层网络设备,则在上行链路上与分组一起被输送的额外信息可以包括标识附属网络设备的哪个端口接收到了该分组的信息。额外信息还可以包括指示发

送设备是否已对分组执行了任何转发或路由选择的信息。在某些实施例中,上行链路的使用允许了虚拟网络设备控制被耦合到该虚拟网络设备的附属网络设备。上行链路的使用还帮助虚拟网络设备有能力为从属附属网络设备执行路由选择和 / 或转发。网络设备或被耦合到上行链路的附属网络设备内的接口在这里被称为上行链路接口。

[0040] 图 3 示出了虚拟网络设备中包括的每个网络设备内的更多细节。在这里,虚拟网络设备 202 包括两个虚拟网络设备子单元 122(1) 和 122(2)。注意,在其他实施例中,虚拟网络设备 202 可以包括多于两个组件网络设备。在该示例中,虚拟网络设备 202 位于网络的分发层。但是,类似的虚拟网络设备也可以实现在其他网络层中(例如数据中心和 / 或核心层内)。

[0041] 虚拟网络设备 202 被耦合到若干个访问层网络设备 120(1)-120(3)。网络设备 120(2) 和 120(3) 各自被两条上行链路耦合到虚拟网络设备 202,这两条上行链路中的每一条去往虚拟网络设备子单元 122(1) 和 122(2) 中的一个。网络设备 120(2) 被虚拟链路束 250(2) 耦合到虚拟网络设备,网络设备 120(3) 被虚拟链路束 250(3) 耦合到虚拟网络设备 202。结果,即使这些上行链路之一和 / 或虚拟网络设备子单元 122(1) 和 122(2) 之一发生故障,网络设备 120(2) 和 120(3) 也能继续与分发层通信。网络设备 120(1) 被三条上行链路耦合到虚拟网络设备 202:两条去往虚拟网络设备子单元 122(1) 的上行链路和一条去往虚拟网络设备子单元 122(2) 的上行链路。这三条上行链路共同形成虚拟链路束 250(1)。即使三条上行链路中的两条和 / 或虚拟网络设备子单元 122(1) 和 122(2) 之一发生故障,网络设备 120(1) 也能继续与分发层通信。网络设备 120(1)-120(3) 各自将它们的多条去往虚拟网络设备 202 的上行链路当作单个逻辑上行链路来操作。此外,在某些实施例中,每个网络设备 120(1)-120(3) 可以像该网络设备被耦合到单个分发层设备(虚拟网络设备 202) 那样工作,而不是像该网络设备被耦合到了两个独立的分发层网络设备那样工作。

[0042] 分发层虚拟网络设备子单元 122(1) 也被单条链路耦合到服务器 104(3)。与访问层网络设备 120(1)-120(3) 不同,服务器 104(3) 并不将分发层网络设备单元 122(1) 和 122(2) 看作单个逻辑网络设备。在该示例中,如果网络设备 122(1) 或将服务器 104(3) 耦合到网络设备 122(1) 的链路发生故障,则服务器 104(3) 将无法经由分发层通信。注意,在替换实施例中,诸如服务器 104(3) 这样的但具有多个端口的服务器可以被虚拟链路束耦合到每个虚拟网络设备,并且这种服务器能够像虚拟网络设备子单元 122(1) 和 122(2) 是单个虚拟网络设备 202 那样与这些子单元交互。

[0043] 虚拟网络设备子单元 122(1) 包括若干个卡,其中包括控制卡 302(1) 以及线路卡 304(1) 和 304(3)。类似地,虚拟网络设备子单元 122(2) 包括控制卡 302(2) 以及线路卡 304(2) 和 304(4)。控制卡 302(1) 包括控制单元 310(1)、转发引擎 312(1) 以及接口 320(1) 和 320(3)。控制卡 302(2) 类似地包括控制单元 310(2)、转发引擎 312(2) 以及接口 320(2) 和 320(4)。

[0044] 在虚拟网络设备子单元 122(1) 中,线路卡 304(1) 包括转发引擎 314(1) 以及接口 320(5)、320(7) 和 320(9)。接口 320(7) 被耦合到网络设备 120(3)。接口 320(9) 也被耦合到网络设备 120(1)。接口 320(5) 在该示例中未被使用。线路卡 304(3) 包括转发引擎 314(3)、接口 320(11) 和 320(13),以及端口 320(15)。接口 320(11) 和 320(13) 分别被耦合到网络设备 120(2) 和 120(1)。接口 320(15) 被耦合到服务器 104(3)。在网络设

备 120(1)-120(3) 是受虚拟网络设备 202 控制的附属网络设备的实施例中,接口 320(7)、320(9)、320(11) 和 320(13) 被当作上行链路接口来操作,而未被耦合到附属网络设备的接口 320(15) 被当作正常端口来操作。

[0045] 在虚拟网络设备子单元 122(2) 中,线路卡 304(2) 包括转发引擎 314(2) 以及接口 320(6)、320(8) 和 320(10)。接口 320(8) 被耦合到附属网络设备 120(2),而接口 320(6) 和接口 320(10) 未被连接。线路卡 304(4) 包括转发引擎 314(4) 以及接口 320(12)、320(14) 和 320(16)。接口 320(12) 和接口 320(16) 分别被耦合到附属网络设备 120(3) 和 120(1)。接口 320(14) 未被使用。在网络设备 120(1)-120(3) 是受虚拟网络设备 202 控制的附属网络设备的实施例中,接口 320(8)、320(12) 和 320(16) 可以被当作上行链路接口来操作。

[0046] 注意,虽然图 2 中的接口被描述成既是进入接口又是外出接口,但是也可以使用充当仅限进入的接口或充当仅限外出的接口的接口。例如,图 2 所示的每个接口的功能可以利用一个仅限进入的接口和一个仅限外出的接口来实现。类似地,虚拟链路束 250(1)-250(3) 各自可以包括若干条只将分组从相应的网络设备 120(1)-120(3) 输送到虚拟网络设备 202 的链路,以及若干条只将分组从虚拟网络设备 202 输送到相应的网络设备 120(1)-120(3) 的链路。

[0047] 在图示实施例中,虚拟网络设备子单元 122(1) 中的控制卡 302(1) 经由虚拟网络设备链路 360 被耦合到虚拟网络设备子单元 122(2) 中的控制卡 302(2)。在该示例中,虚拟网络设备链路 360 包括两条链路(两条链路被用来提供更大的容错性和/或带宽;但是,在其他实施例中也可以使用一条链路)。在该示例中,这些链路是一类上行链路,其除了携带分组外还携带信息(例如与线路卡之间发送的那些类似的头部)。虚拟网络设备链路 360 中的上行链路被用于在虚拟网络设备子单元 122(1) 和 122(2) 之间交换控制虚拟网络设备 202 的工作的信息以及分组。通过经由这些上行链路进行通信,虚拟网络设备子单元 122(1) 和 122(2) 可以协调它们的行为,从而使得对于网络设备 120(1)-120(3),虚拟网络设备子单元 122(1) 和 122(2) 看起来是单个虚拟网络设备。

[0048] 从而,在虚拟网络设备子单元 122(1) 和 122(2) 之间提供互连允许了虚拟网络设备子单元 122(1) 和 122(2) 作为单个虚拟网络设备 202 工作。网络设备 120(1)-120(3) 以与网络设备 120(1)-120(3) 和单个物理设备通信的方式相同的方式与虚拟网络设备 202 通信。例如,如果网络设备 120(2) 正在处理寻址到服务器 104(3) 的分组,则网络设备 120(2) 可以选择网络设备束 250(2) 中的两条上行链路之一以在其上发送分组。这一选择可以基于负载共享标准。在这种情形下,由于虚拟网络设备 202 看起来是单个网络设备,因此,尽管只有虚拟网络设备子单元 122(1) 具有到服务器 104(3) 的直接连接,但网络设备 120(2) 选择去往虚拟网络设备子单元 122(2) 的上行链路的可能性与选择去往虚拟网络设备子单元 122(1) 的上行链路的可能性是相同的。如果分组被发送到虚拟网络设备子单元 122(2),则网络设备 122(2) 随后可以利用虚拟网络设备子单元 122(1) 和 122(2) 之间的虚拟网络设备链路 360 中包括的上行链路之一来将分组发送到虚拟网络设备子单元 122(1),而虚拟网络设备子单元 122(1) 进而又可以将分组提供到其目的地,即服务器 104(3)。

[0049] 在其他实施例中,网络设备 120(1)-120(3) 可能意识到它们的虚拟链路束 250(1) 和 250(2) 实际上在两个不同的网络设备上终止。网络设备 120(1)-120(3) 可以基于该信息控制分组传输。例如,在这种情形下,网络设备 120(2) 可以基于网络设备 120(2) 认识到

了逻辑链路内去往两个不同网络设备的不同的连接这一事实,通过选择被耦合到虚拟网络设备子单元 122(1) 的上行链路而不是被耦合到虚拟网络设备子单元 122(2) 的上行链路来处理寻址到服务器 104(3) 的分组。

[0050] 各自被虚拟链路束 250(1) 耦合到网络设备 120(1) 的接口 320(13)、320(9) 和 320(16) 形成接口束。如虚拟链路束 250(1) 所示,虚拟链路束可包括多于两条链路。此外,虚拟链路束可包括在相同虚拟网络设备子单元内的不同线路卡上终止的链路。类似地,接口 320(11) 和 320(8) 形成了另一个接口束,该接口束被虚拟链路束 250(2) 耦合到网络设备 120(2)。接口 320(7) 和 320(12) 形成第三接口束,该接口束被虚拟链路束 250(3) 耦合到网络设备 120(3)。在虚拟网络设备 202 内,相同接口束中的每个接口被分配以相同的逻辑标识符。例如,接口 320(13)、320(9) 和 320(16) 各自被分配以相同的逻辑标识符。在某些实施例中,经由这些接口之一接收到的分组可以被标记以该逻辑标识符或以其他方式与该逻辑标识符相关联,以指示这些分组是经由将虚拟网络设备 202 耦合到网络设备 120(1) 的虚拟链路束接收到的。注意,类似的接口束在每个网络设备 120(1)-120(3) 内实现,并且这种束中包括的接口也可以被每个网络设备(或者被虚拟网络设备 202,在虚拟网络设备 202 控制网络设备 120(1)-120(3) 的配置的实施例中)分配以相同的逻辑标识符。例如,网络设备 120(1) 可以向被耦合到虚拟链路束 250(1) 的每个接口分配相同的逻辑标识符。

[0051] 分组和特定逻辑标识符之间的关联可以被虚拟网络设备 202 内的转发引擎用来将分组路由和转发到以及路由和转发自网络设备 120(1)-120(3)。例如,当来自发送设备(例如被耦合到网络设备 120(1) 的客户端)的分组经由上行链路接口 320(13) 被接收到时,虚拟网络设备子单元 122(1) 可以通过使发送设备的 MAC(媒体访问控制)地址与上行链路接口 320(13) 的逻辑标识符相关联来获知该 MAC 地址位于上行链路接口 320(13) “之后”。虚拟网络设备子单元 122(1) 可以将该关联通知给虚拟网络设备子单元 122(1) 中的每个转发引擎以及虚拟网络设备子单元 122(2) 中的每个转发引擎。基于该关联,寻址到该 MAC 地址的分组将会从具有相关联的逻辑标识符的上行链路接口发送出来。由于在这种情况下,上行链路接口 320(9)(在虚拟网络设备子单元 122(1) 中)和 320(16)(在虚拟网络设备子单元 122(2))也具有与上行链路接口 320(13) 相同的逻辑标识符,因此寻址到该 MAC 地址的分组可以经由上行链路接口 320(9)、320(13) 和 320(16) 中的任何一个被转发。

[0052] 相同的逻辑标识符可以被虚拟网络设备子单元 122(1) 和 122(2) 中的每一个用来标识上行链路接口束,并且虚拟网络设备子单元协作以将相同的逻辑标识符分配给相同上行链路接口束内的每个上行链路接口。当经由特定逻辑标识符所标识的上行链路接口束转发分组时,每个虚拟网络设备子单元 122(1) 和 122(2) 生成散列值,以选择该上行链路接口束内的上行链路接口之一以在其上发送分组。每个虚拟网络设备子单元利用这些散列值来标识该虚拟网络内的本地上行链路接口。从而,每个虚拟网络设备子单元只会选择该虚拟网络设备子单元本地的上行链路接口。例如,如果虚拟网络设备子单元 122(1) 正在经由包括接口 320(9)、320(13) 和 320(16) 的上行链路接口束转发分组,则由虚拟网络设备子单元生成的散列值将会标识其接口 320(9) 或 320(13) 之一。

[0053] 在上述示例中,通过使每个散列值与上行链路接口束中的本地上行链路接口相关联,对虚拟交换机链路 360 的使用减少了。基本上,虚拟网络设备子单元 122(1) 更倾向于特定上行链路接口束内的其本地的上行链路接口,而不是相同上行链路接口束中处于虚拟

网络设备子单元 122(2) 上的远程上行链路接口。类似地,虚拟网络设备子单元 122(2) 更倾向于特定上行链路接口束内的其本地的上行链路接口,而不是包括在虚拟网络设备子单元 122(1) 中的上行链路接口。例如,如果虚拟网络设备子单元 122(2) 需要经由上行链路接口转发分组,则虚拟网络设备子单元 122(2) 将会经由上行链路接口 320(12) 发送该分组,而不是在虚拟网络设备链路 360 上转发该分组以经由上行链路接口 320(7) 发送。通过更倾向于本地接口,经由虚拟网络设备链路 360 发送的流量的量可以减小,这是因为每个虚拟网络设备子单元 122(1) 和 122(2) 将会从本地接口转发在本地接收到的分组(即经由除被耦合到虚拟网络设备链路 360 的接口之外的接口接收到的分组)。下文中描述的图 6A-6D 示出关于如何通过更倾向于用第一虚拟网络设备子单元内的本地接口处理特定分组从而来减小虚拟网络设备链路 360 上的流量的更详细示例。

[0054] 对于给定虚拟链路束,该虚拟链路束可在中心位置中被管理(例如针对诸如 L2 协议这样的控制协议)。例如,在一个实施例中,针对虚拟链路束 250(1) 的所有控制协议处理可以发生在虚拟网络设备子单元 122(1) 的控制单元 310(1) 中。该控制协议处理的结果随后可以被传输到虚拟网络设备子单元 122(2) 的控制单元 310(2) 和/或网络设备 120(1) 中的控制器。控制单元 310(2) 随后在控制如何处理从上行链路接口 320(16) (其处于被耦合到虚拟链路束 250(1) 的上行链路接口束中) 发送来的以及经由上行链路接口 320(16) 接收到的分组时可以使用(但不是修改)该信息。例如,控制单元 310(2) 可以利用该信息来设置或修改线路卡 304(2) 和/或 304(4) 上的查找表。这样,实际控制协议处理集中在控制单元 310(1) 中,而不是分布在虚拟网络设备 202 中的若干个控制单元上。

[0055] 控制协议处理的中心点在虚拟链路束之间可以发生变化。例如,针对虚拟链路束 250(1) 的控制协议处理由控制单元 310(1) 管理,而针对虚拟链路束 250(2) 的控制协议处理可以由控制单元 310(2) 管理。换言之,控制单元 310(2) 可以执行针对虚拟链路束 250(2) 的所有控制处理,并且由控制单元 310(2) 生成的信息随后可以被传输到控制单元 310(1) 以便在虚拟网络设备子单元 122(1) 内使用(但不是修改)。

[0056] 在在虚拟网络设备 202 内为每个虚拟链路束的控制协议处理实现一个管理中心点的实施例中,L2 协议在虚拟链路束上被运行,以及/或者接口束被用作经路由的 L3 接口(注意,在其他实施例中使用除 L2 和 L3 协议之外的其他协议)。如果虚拟网络设备 202 内的虚拟网络设备子单元独立于彼此地各自为其本地接口束执行控制协议处理,则这些能力将不可用。此外,在实现控制协议处理中心点的实施例中,用户可以通过访问虚拟网络设备子单元来修改虚拟链路束的控制协议行为。在上述示例中,在更新虚拟链路束 250(1) 的控制协议行为时,用户可以简单地就访问虚拟网络设备子单元 122(1) (而不是访问虚拟网络设备子单元 122(1) 和 122(2) 两者)。虚拟网络设备子单元 122(1) 随后可以将用户对控制协议作出的任何改变自动传播到网络设备 122(2)。此外,由于虚拟链路束的使用允许了若干个上行链路被当作单个逻辑上行链路来管理,因此与不使用虚拟链路束的情况相比,需要配置的上行链路接口更少。例如,如果每个虚拟链路束包括两条上行链路,则虚拟网络设备 202 内需要由用户配置的上行链路接口的数目减半。

[0057] 虚拟网络设备子单元 122(1) 和 122(2) 实现某些行为以充当虚拟网络设备 202,从网络设备 120(1)-120(3) 的角度来看,该虚拟网络设备 202 看起来是单个逻辑网络设备。例如,每当虚拟网络设备子单元 122(2) 接收到来自本地网络设备、客户端或服务端的分组

并且该分组的目的地逻辑标识符标识上行链路接口束时,虚拟网络设备子单元 122(2) 从所标识的上行链路接口束内的本地上行链路接口发送该分组。虚拟网络设备子单元 122(2) 也可以将该分组提供给虚拟网络设备子单元 122(1),但虚拟网络设备子单元 122(1) 不应该在虚拟链路束上输出该分组。这样,目的地设备只接收到来自虚拟网络设备 202 的分组的一个拷贝(而不是从每个虚拟网络设备子单元 122(1) 和 122(2) 接收到一个拷贝),虚拟网络设备 202 作为单个实体的外观得以维持。

[0058] 为了以这种方式工作,被耦合到虚拟链路束中的一条链路的每个外出上行链路接口被配置为过滤掉经由虚拟网络设备链路 360 接收到的流量。例如,虚拟网络设备子单元 122(1) 经由虚拟网络设备链路 360 接收分组。接收该分组的接口 320(1) 或 320(3) 可以更新与分组相关联的信息(例如在头部中)以指示该分组是经由虚拟网络设备链路 360 接收到的(在替换实施例中,虚拟网络设备子单元 122(2) 中的发送接口可以更新该信息)。当虚拟网络设备子单元 122(1) 在查找表中查找分组的目的地地址时,查找表 返回标识本地上行链路接口 320(9) 和 320(13) 的逻辑标识符。分组随后被转发到上行链路接口 320(13) (例如基于负载共享因素选择)。当上行链路接口 320(13) 接收到分组时,上行链路接口 320(13) 只会在分组不是经由虚拟交换机链路 360 接收到的情况下才输出分组,这是因为如果分组是经由虚拟交换机链路接收到的,则另一虚拟网络设备子单元 122(2) 已经经由虚拟链路束发送了分组。从而上行链路接口 320(13) 基于附加到分组的指示分组是否是经由虚拟网络设备链路 360 接收到的的信息来从经由上行链路接口 320(13) 发送的分组流中过滤掉该分组。

[0059] 在某些实施例中,MAC 通知帧被用来使虚拟网络设备子单元 122(1) 中的 L2 表的内容与虚拟网络设备子单元 122(2) 中的 L2 表的内容保持同步,反之亦然。每当在虚拟网络设备子单元内生成涉及位于虚拟链路束之后的端口或上行链路接口束中包括的上行链路接口的 MAC 通知时(这种通知可以由一个线路卡生成,以更新另一线路卡上的 L2 表),MAC 通知的拷贝经由虚拟网络设备链路 360 被发送。类似地,如果虚拟网络设备子单元确定分组应当被流播(flood),则虚拟网络设备子单元将会经由虚拟网络设备链路 360 发送该分组的拷贝,以确保虚拟网络设备子单元将会接收到由对等虚拟网络设备子单元中的转发引擎生成的任何 MAC 通知响应的拷贝。

[0060] 例如,假设由于虚拟网络设备子单元 122(1) 中包括的转发引擎不知道哪个端口或上行链路接口与分组的目的地地址相关联,因而虚拟网络设备子单元 122(1) 流播分组。作为分组流播的一部分,虚拟网络设备子单元 122(1) 经由虚拟交换机链路 360 将分组的拷贝发送到虚拟网络设备子单元 122(2)。如果虚拟网络设备子单元 122(2) 内的转发引擎已经知道目的地地址位于特定上行链路接口或端口之后(例如转发表已经包括将目的地地址与网络设备 120 之一的端口关联起来的条目),则该转发引擎生成标识该关联的 MAC 通知,该通知被分发到虚拟网络设备子单元 122(2) 内的任何其他转发引擎。由于分组最初是经由虚拟网络设备链路 360 接收到的,因此虚拟网络设备子单元 122(2) 也经由虚拟网络设备链路 360 发送回 MAC 通知的拷贝。该 MAC 通知随后在虚拟网络设备子单元 122(1) 中包括的转发引擎之间分发。在基于 MAC 通知被更新之后,虚拟网络设备子单元 122(1) 中的转发引擎现在就知道由目的地地址标识的设备的位置了。因此,随后接收到的寻址到该设备的分组将不会被流播。

[0061] 当虚拟链路束中连接到单个虚拟网络设备子单元的所有物理链路都发生故障时，虚拟链路束转变成被耦合到单个虚拟网络设备子单元的正常链路束。此时，每个虚拟网络设备子单元针对该网络设备束的行为被修改。例如，假设虚拟链路束 250(1) 中被耦合到虚拟网络设备子单元 122(2) 的所有上行链路都发生故障。此时，虚拟网络设备子单元 122(2) 不再拥有任何可以经由虚拟链路束 250(1) 发送分组的本地上行链路接口。因此，虚拟网络设备子单元 122(2) 将会在虚拟网络设备链路 360 上对所有需要经由虚拟链路束 250(1) 发送的流量进行重定向。此外，由于网络设备 122(2) 不再能够经由虚拟链路束 250(1) 发送分组，因此虚拟网络设备子单元 122(1) 将会停止阻止经由虚拟网络设备链路 360 接收到的流量被经由虚拟链路束 250(1) 发送。如果虚拟链路束 250(1) 中的上行链路中的至少一条被耦合到虚拟网络设备子单元 122(2) 的上行链路被恢复，则虚拟链路束 250(1) 将会转变回其正常工作模式，在这个模式中，虚拟网络设备子单元 122(2) 将会会经由虚拟链路束 250(1) 发送在本地接收到的分组，并且虚拟网络设备子单元 122(1) 将会过滤出经由虚拟网络设备链路 360 接收到的分组以使其不经由虚拟链路束 250(1) 发送。

[0062] 虚拟网络设备内的分布式转发

[0063] 图 4A-4F 示出共同作为单个虚拟网络设备工作的两个虚拟网络设备子单元 122(1) 和 122(2) 如何对具有多播目的地地址的分组执行分布式转发。在该示例中，具有多播目的地地址的分组的至多一个拷贝经由虚拟网络设备链路 (VNDL) 360 在虚拟网络设备子单元之间传送。

[0064] 在图 4A-4F 中，虚拟网络设备包括两个虚拟网络设备子单元 122(1) 和 122(2)。虚拟网络设备子单元 122(1) 包括两个转发引擎 (FE)，即 FE410(1) 和 FE 410(2)。虚拟网络设备子单元 122(1) 具有到蓝色虚拟局域网 (VLAN) (在本示例中用不同的颜色来指代不同的 VLAN) 中的设备的一个接口 420(1)。虚拟网络设备子单元 122(1) 还包括被耦合到虚拟链路束 (VLB) 450(1) 的一个接口和被耦合到虚拟链路束 450(2) 的另一个接口。VLB 450(1) 被耦合到蓝色 VLAN 中的设备，VLB 450(2) 被耦合到绿色 VLAN 中的设备。虚拟网络设备子单元 122(1) 还包括被耦合到虚拟网络设备链路 360 的接口。注意，在其他实施例中，虚拟网络设备包括多于两个虚拟网络设备子单元。

[0065] 虚拟网络设备子单元 122(2) 包括两个转发引擎，即 FE 410(3) 和 FE410(4)。虚拟网络设备子单元 122(2) 具有到红色虚拟局域网 (VLAN) 中的设备的一个接口 420(2)。虚拟网络设备子单元 122(2) 还包括被耦合到虚拟链路束 (VLB) 450(1) 的一个接口和被耦合到虚拟链路束 450(2) 的另一个接口。虚拟网络设备子单元 122(2) 还包括被耦合到虚拟网络设备链路 360 的接口。

[0066] 在某些实施例中，每个虚拟网络设备子单元 122(1) 和 122(2) 包括两个或更多个转发引擎。虚拟网络设备子单元中的转发引擎的数目可以与该虚拟网络设备子单元中包括的线路卡的数目成比例。例如，在一个实施例中，每个线路卡具有其自己的转发引擎。每个转发引擎 410(1)-410(4) 是被配置为对具有多播目的地地址的分组执行分布式转发的分布式转发模块的示例。

[0067] 在其中每个线路卡具有其自己的转发引擎的虚拟网络设备内，对于每个虚拟网络设备子单元，分组由最多两个转发引擎所处理。在每个虚拟网络设备子单元内，分组的给定拷贝被最多两个转发引擎所转发。这些转发引擎中之一是进入线路卡上的转发引擎，该进

入线路卡是包括分组经由其被接收进虚拟网络设备子单元中的接口的线路卡。另一个转发引擎位于外出线路卡上,该外出线路卡是包括分组的特定拷贝经由其从虚拟网络设备子单元输出的接口的线路卡。如果进入线路卡也是分组的给定拷贝的外出线路卡,则分组的该拷贝只被一个转发引擎处理。注意,对于给定分组,该分组的若干个拷贝可被输出。例如,如果分组被流播,则进入线路卡上的转发引擎使得该分组的拷贝被提供给每个其他线路卡。

[0068] 在某些实施例中,每个转发引擎 410(1)-410(4) 执行两个不同种类的查找。第一类查找在这里被称为进入查找。当对分组执行进入查找时,在查找表中可分配“初级”条目。初级条目对应于分组的源地址。如果特定线路卡上的查找表包括针对给定地址的初级条目,则初级条目标识在该特定线路卡上的应当经由其输出寻址到该特定地址的分组的输出接口。进入查找通常由虚拟网络设备内的第一转发引擎对分组执行,以处理该分组。当外出查找被执行时,初级条目不能被分配。外出查找通常由转发引擎而不是第一转发引擎执行,以处理给定分组。

[0069] 在图 4A 中,耦合到端口 420(1) 的设备向虚拟网络设备发送分组。该分组具有多播目的地地址。该分组被虚拟网络设备子单元 122(1) 经由端口 420(1) 所接收。该分组在 VLAN “蓝色”中传递(在该示例中,使用颜色来标识 VLAN)。换言之,蓝色 VLAN 对于该分组而言是传入 VLAN。

[0070] 如图 4B 所示,FE 410(1) 对经由端口 420(1) 接收的分组执行查找(例如,FE 410(1) 可以与端口 420(1) 包括在同一线路卡上,或者端口 420(1) 可以被包括在不包括其自身的转发引擎的线路卡上)。响应于经由端口 420(1) 接收的分组,FE 410(1) 执行对分组的进入查找。在该示例中,对于具有多播目的地地址的分组,进入查找包括 L2 查找和 L3 查找(注意,在其他实施例中,可以使用其他类型的查找)。进入查找识别出与分组被寻址到的多播组中的设备相关联的传出接口。

[0071] L2 查找返回一索引,该索引标识出蓝色 VLAN 中的所有接口。所标识的接口因此是处于耦合到虚拟链路束 (VLB) 450(1) 和端口 420(1) 的接口束中的本地接口。由于分组是经由端口 420(1) 接收的,因此分组不会经由该端口被输出。FE 410(1) 在蓝色 VLAN 中将分组的拷贝(标记为“B”)发送到 FE 410(2),FE 410(2) 执行对耦合到 VLB 450(1) 的本地接口的转发(例如 FE 410(2) 可以被包括在与耦合到 VLB 450(1) 的端口相同的线路卡上)。与发送到 FE 410(2) 的分组的拷贝相关联的信息(例如在附接到分组的头部中)指示已经执行了进入查找。

[0072] L3 查找(其标识除了传入 VLAN 之外的其他 VLAN 中的传出接口)识别出绿色 VLAN 中的耦合到 VLB 450(2) 的传出接口。由于在红色 VLAN 中存在传出接口,因此 FE 410(1) 复制该分组。通过复制该分组,FE 410(1) 针对每个外出 VLAN 创建至少一个分组拷贝。在该示例中,分组在绿色 VLAN 中被复制。FE 410(1) 将分组的复制拷贝发送到 FE 410(2)(例如,由于 FE 410(2) 被包括在与耦合到 VLB 450(2) 的接口相同的线路卡上)。

[0073] 如图 4C 所示,在绿色 VLAN 中的分组的复制拷贝(标记为“G”)被 FE 410(2) 所接收。响应于接收到蓝色 VLAN 中的分组的拷贝,FE 410(2) 对该分组拷贝执行外出查找。外出查找识别出传入 VLAN(蓝色 VLAN) 中的传出接口。在该示例中,外出查找识别出耦合到 VLB 450(1) 的接口。因此,FE 410(2) 致使分组拷贝经由蓝色 VLAN 中的 VLB 450(1) 输出。在外出查找期间,如果尚未对分组的源地址分配查找表条目,FE 410(2) 则还向该地

址分配非初级条目。非初级条目指示以该地址为目的地址的分组应该被转发到包括 FE 410(1) 的线路卡。

[0074] FE 410(1) 还标识蓝色 VLAN 是多播分组的传入 VLAN。因此, FE410(1) 经由虚拟网络设备链路 (VNDL) 360 将分组的拷贝发送到虚拟网络设备子单元 122(2)。FE 410(1) 还将信息与分组的该拷贝相关联 (例如通过将头部附加到分组), 并经由 VNDL 360 将相关联的信息发送到虚拟网络设备子单元 122(2)。相关联的信息标识最初接收到分组的传入 VLAN (蓝色 VLAN)。

[0075] 如图 4D 所示, FE 410(2) 也对在绿色 VLAN 中输送的分组的拷贝执行外出查找。外出查找将被耦合到 VLB 450(2) 的本地接口标识为分组的该拷贝的传出接口, 并且 FE 410(2) 使得分组的拷贝经由 VLB 450(2) 输出。

[0076] 此外, 虚拟网络设备子单元 122(2) 中的 FE 410(3) 对经由 VNDL 360 接收到的分组的拷贝执行进入查找。与分组相关联的信息将传入 VLAN 标识为蓝色 VLAN。L2 查找标识被耦合到 VLB 450(1) 的接口。L3 查找标识绿色 VLAN 中的端口 420(3) 和被耦合到 VLB 450(2) 的接口以及红色 VLAN 中的端口 420(2)。因此, FE 410(3) 使得分组在绿色 VLAN 和红色 VLAN 中被复制。

[0077] 在图 4E 中, 蓝色 VLAN 中的分组的拷贝不被虚拟网络设备子单元 122(2) 经由 VLB 450(1) 发送。例如, 被耦合到 VLB 450(1) 的接口响应于检测到分组已由虚拟网络设备子单元 122(1) 处理而从经由 VLB 450(1) 输出的分组流中过滤掉该分组。类似地, 绿色 VLAN 中的分组的拷贝不经由 VLB 450(2) 发送。因此, 经由每个虚拟链路束只输出分组的一个拷贝。

[0078] FE 410(3) 使得绿色 VLAN 中的分组的拷贝经由端口 420(3) 输出。由于该端口不是跨越虚拟网络设备子单元 122(1) 和 122(2) 两者的接口束的一部分, 因此分组不被从经由端口 420(3) 输出的分组流中过滤掉。FE 410(3) 还使得红色 VLAN 中的分组的拷贝 (被标记为“R”) 被发送到 FE 410(4) (例如因为 FE 410(4) 和端口 420(2) 被包括在同一线路卡上)。

[0079] 图 4F 示出 FE 410(4) 如何对在红色 VLAN 中发送的分组的拷贝执行外出查找。外出查找标识端口 420(2), 并且 FE 410(4) 使得分组的拷贝经由端口 420(2) 输出。由于端口 420(2) 不是与虚拟网络设备子单元 122(1) 共享的接口束的一部分, 因此分组不被从经由端口 420(2) 发送的输出流中过滤掉。

[0080] 注意, 在上述示例中, 每个虚拟网络设备子单元中的一个转发引擎对多播分组执行进入查找。在进入查找期间, 如果必要的话, 分组在每个传出 VLAN 中被复制。执行进入查找的转发引擎将分组的拷贝转发到每个包括传入和 / 或传出 VLAN 中的接口的线路卡。在图 4A-4F 的示例中, 如果线路卡包括多个 VLAN 中的接口 (例如, 线路卡包括 FE 410(2) 以及分别被耦合到 VLB 450(1) 和 450(2) 的接口), 则线路卡接收每个 VLAN 中的分组的一个拷贝。

[0081] 在替换实施例中, 分组的至多一个拷贝被发送到每个线路卡, 而不论该线路卡是否包括在多于一个传入和 / 或传出 VLAN 中的传出接口。在这种实施例中, 每个线路卡中的转发引擎对分组执行进入查找, 以在必要时在其他 VLAN 中复制分组。例如, 如果此类转发被用于图 4A 的实施例中, 则只有一个多播分组拷贝会被发送到 FE 410(2)。FE 410(2) 随

后将对该分组执行进入查找。作为进入查找的一部分,FE 410(2) 将会在绿色 VLAN 中复制分组。

[0082] 图 5A 是由虚拟网络设备子单元用来将多播分组的至多一个拷贝转发到同一虚拟网络设备中的另一个虚拟网络设备子单元的方法的流程图。在 501,多播分组被虚拟网络设备子单元接收。分组是经由未被耦合到虚拟网络设备链路的接口接收的(即,分组是从除同一虚拟网络设备中的另一虚拟网络设备子单元之外的设备接收的)。分组在特定 VLAN 中被输送。

[0083] 虚拟网络设备子单元对多播分组执行进入查找。作为该查找的一部分,虚拟网络设备子单元确定是否需要为其他传出 VLAN 复制分组,如 503 所示。如果由于在除传入 VLAN 之外的 VLAN 中有传出接口因而分组需要被复制,则为每个传出 VLAN 制作分组的至少一个拷贝,如 507 所示。

[0084] 不论在 507 处分组是否被复制,都只有一个分组拷贝经由虚拟网络设备链路被发送到另一个虚拟网络设备子单元,如 509 所示。与该分组相关联的信息(例如包括在附加到分组的头部中或包括在分组内的以太网标签内)也经由虚拟网络设备链路被发送,它标识了接收到分组的原始传入 VLAN。

[0085] 在 507,虚拟网络设备子单元(其在 501 处接收到分组)输出分组的一个或多个拷贝。分组的这些拷贝从分组的多播群组的传出和传入 VLAN 中的适当的端口和/或上行链路接口(包括那些被耦合到虚拟链路束的)输出。

[0086] 图 5B 是由虚拟网络设备子单元用来转发接收自包括在同一虚拟网络设备中的另一个虚拟网络设备子单元的多播分组的方法的一个实施例的流程图。在 551,虚拟网络设备子单元经由被耦合到虚拟网络设备链路的接口接收多播分组。换言之,虚拟网络设备子单元接收来自另一个虚拟网络设备子单元的多播分组。标识传入 VLAN 的信息也被接收。

[0087] 在 553,虚拟网络设备子单元(其在 551 处接收到多播分组)复制分组(如果必要的话)。例如,如果虚拟网络设备子单元包括在一个或多个传出 VLAN 中的接口,则虚拟网络设备子单元在这些 VLAN 中复制分组。在一个实施例中,复制是作为对多播分组的进入查找的一部分执行的。

[0088] 在 555,分组的拷贝从传入和传出 VLAN 中的接口输出。但是,没有分组拷贝从被耦合到接口束的接口输出。

[0089] 图 6A-6K 示出共同作为单个虚拟网络设备工作的两个虚拟网络设备子单元如何对具有单播目的地地址的分组执行分布式转发。注意,在其他实施例中,虚拟网络设备可包括多于两个虚拟网络设备子单元。在图 6A-6K 中,两个虚拟网络设备子单元 122(1)和 122(2) 被虚拟网络设备链路(VNDL)360 耦合。虚拟网络设备子单元 122(1) 包括端口 620(1)、转发引擎(FE)610(1)和 610(2)、被耦合到虚拟链路束(VLB)650(1)的接口,以及被耦合到 VNDL 360 的接口。虚拟网络设备子单元 122(2) 包括端口 620(2)、转发引擎(FE)610(3)和 610(4)、被耦合到虚拟链路束(VLB)650(1)的接口,以及被耦合到 VNDL 360 的接口。每个转发引擎 610(1)-610(4) 是被配置为对具有单播目的地地址的分组执行分布式转发的分布式转发模块的示例。

[0090] 在图 6A 中,虚拟网络设备子单元 122(1) 经由端口 610(1) 接收分组。分组是单播分组(即,分组具有单播目的地地址)。如图 6B 所示,FE610(1) 对单播分组执行进入查找。

在查找表中没有分组的目的地地址,因此转发引擎 610(1) 在分组的传入 VLAN 中流播分组。如果查找表尚未包括针对分组的源地址的条目,FE 610(1) 则向分组的源地址分配初级条目。被分配的初级条目指示该源地址与端口 620(1) 相关联。如果 FE 610(1) 后来转发将该地址作为其目的地地址的分组,则 FE 610(1) 将会响应于存储在初级条目中的信息使分组经由端口 620(1) 输出。

[0091] 在图 6C 中,FE 610(1) 使分组被转发到 FE 610(2) (例如,通过将分组转发到包括 FE 610(2) 的线路卡)。FE 610(1) 还发送指示已对分组执行了进入查找并且进入查找的结果是流播分组的信息(例如包括在附加到分组的头部中)。例如,在某些实施例中,FE 610(1) 将目的地索引附加到分组或以其他方式使目的地索引与分组相关联,该目的地索引在此情况下是流播索引(例如指示分组正被流播的目的地索引)。

[0092] 在图 6D 中,FE 610(2) 对分组执行外出查找。外出查找命中针对分组的目的地地址的非初级条目。非初级条目标识应当经由哪个传出接口(在虚拟网络设备内的)输出分组以到达其目的地。在本示例中,非初级条目标识分组应当经由 VNDL 360 输出以到达其被包括在虚拟网络设备子单元 122(2) 中的传出接口,即端口 620(2)。条目不是初级条目这一事实指示传出接口(在本示例中是 620(2))未被包括在与 FE 610(2) 相同的线路卡(或者 FE 610(2) 为其执行转发的线路卡)中。在外出查找期间,FE 610(2) 还向分组的源地址分配非初级条目(如果尚未分配任何与该源地址对应的条目)。这个非初级条目指示以该地址作为目的地地址的分组应当被转发到包括 FE 610(1) 的线路卡。

[0093] 由于 FE 610(2) 找到针对分组的目的地地址的匹配条目,因此 FE 610(2) 不继续流播分组。因此,分组不经由被耦合到 VLB 650(1) 的本地接口被发送。但是,由于匹配条目不是初级条目,因此 FE 610(2) 不更新与分组相关联的信息。相反,与分组相关联的信息保持由 FE 610(1) 生成的值,该值指示对分组的目的地地址的查找失败并且分组正被流播。FE 610(2) 经由 VNDL 360 将分组和相关信息转发到虚拟网络设备子单元 122(2),如图 6E 所示。

[0094] 在图 6F 中,FE 610(3) 响应于经由 VNDL 360 接收到分组对分组执行外出查找。从而,虽然是虚拟网络设备子单元 122(2) 对分组执行查找的第一个转发引擎,但 FE 610(3) 仍执行外出查找(而不是进入查找)。这与图 4A-4F 所示的多播转发场景不同,在多播转发场景中,对于经由虚拟网络设备链路接收到的具有多播虚拟地址的分组执行进入查找。

[0095] 由 FE 610(3) 执行的外出查找命中针对分组的目的地地址的非初级条目。该非初级条目例如可以标识包括端口 620(2) 的线路卡。由于匹配条目是非初级条目,因此 FE 610(3) 不修改与分组相关联的信息。相反,FE610(3) 使分组和相关信息(其仍具有由虚拟网络设备子单元 122(1) 中的 FE 610(1) 生成的值)被转发到 FE 610(4),如图 6G 所示。

[0096] 在外出查找期间,如果尚未分配任何与源地址相对应的条目,则 FE610(3) 还向分组的源地址分配非初级条目。在本示例中,非初级条目指示寻址到该地址的分组应当经由 VNDL 360 被转发到虚拟网络设备子单元 122(1)。但是,如果分组已经被虚拟网络设备子单元 122(1) 经由与虚拟网络设备子单元 122(2) 共享接口束接收,则分配给分组的源地址的非初级条目应当标识虚拟网络设备子单元 122(2) 中的接口。例如,如果分组经由 VLB 650(1) 被接收,然后经由被 VNDL 360 被转发到虚拟网络设备子单元 122(2),则 FE 410(3)

可向分组的源地址分配标识被耦合到 VLB 650(1) 的本地接口的非初级条目。

[0097] 在图 6H 中,FE 610(4) 对分组执行外出查找。外出查找命中查找表中的初级条目。初级条目标识端口 620(2)。例如,包括在初级条目中的目的地索引标识端口 620(2)。

[0098] FE 610(4) 将与分组相关联的信息与初级条目中的信息相比较。例如,FE 610(4) 将与分组相关联的目的地索引与包括在外出查找期间找到的初级条目中的目的地索引相比较。由于被 FE 610(1) 与分组关联起来的目的地索引是流播索引,并且初级条目中的目的地索引标识端口 620(2),因此两个目的地索引将不会匹配。

[0099] 如图 6I 所示,FE 610(4) 使分组经由端口 620(2) 输出。此外,由于与分组相关联的信息与在初级条目中找到的信息不匹配,因此 FE 610(4) 使差错通知被发送回对分组执行进入查找的转发引擎。在本示例中,差错通知是 MAC(媒体访问控制)通知(被标记为“MN”)。MAC 通知指示分组的目的地 MAC 地址。FE 610(4) 使 MAC 通知被发送到 FE 610(3)。

[0100] 如图 6J 所示,FE 610(3) 经由 VNDL 360 将 MAC 通知转发回虚拟网络设备子单元 122(1)。类似地,FE 610(2) 将 MAC 通知转发到 FE 610(1),如图 6K 所示。响应于接收到 MAC 通知,FE 610(1) 通过向包括在分组中的 MAC 地址分配非初级条目而“获知”包括在分组中的 MAC 地址。非初级条目包括标识 FE 610(1) 应当如何转发具有该目的地地址的分组的目的地索引或其他信息。在本示例中,FE 610(1) 在响应于 MAC 通知分配的非初级条目中包括标识包括 FE 610(2) 的线路卡的目的地索引,这是因为 FE 610(1) 是从该线路卡接收到 MAC 通知的。获知 MAC 地址的一个结果是 FE610(1) 随后将不会把接收到的分组流播到该目的地 MAC 地址。

[0101] 图 7 是虚拟网络设备子单元转发接收自同一虚拟网络设备中的另一个虚拟网络设备子单元的单播分组的方式的流程图。在 701,虚拟网络设备子单元经由虚拟网络设备链路接收单播分组(即单播分组是从另一个虚拟网络设备子单元接收到的)。与分组相关联的信息,例如目的地索引,也经由虚拟网络设备链路被接收。该关联信息是由对分组执行进入查找的转发引擎(处于同一虚拟网络设备内)生成的。

[0102] 在 703,对分组执行外出查找。不对分组执行进入查找,因为另一虚拟网络设备子单元中的转发引擎已经对分组执行了进入查找。执行外出查找可能包括将分组的目的地地址的全部或部分输入到查找表。

[0103] 如果在 705 处确定在外出查找期间找到非初级条目,则不修改与分组相关联的信息。根据查找中找到的信息转发分组和相关信息,如 707 处所示。例如,如果外出查找未找到匹配条目,则分组和相关信息被流播。

[0104] 如果在 705 处确定找到初级条目,则将与分组相关联的信息与初级条目中的信息相比较。例如,如 709 处所示,将包括在初级条目中的目的地索引与和该分组相关联的信息中的目的地索引相比较。如果初级条目中的目的地索引与和分组相关联的目的地索引不匹配,则用初级条目中找到的目的地索引替换与分组相关联的目的地索引,如 711 处所示。此外,将差错通知,例如 MAC 通知,发送回执行外出查找的转发引擎。在替换与分组相关联的目的地索引之后,如果必要的话,根据相关联的目的地索引转发分组,如 713 处所示。

[0105] 图 8 是示出虚拟网络设备子单元转发接收自同一虚拟网络设备中的另一个虚拟网络设备子单元的分组的方式的流程图。在 801,经由虚拟网络设备链路接收到分组。如果分组包括单播目的地地址,则对分组的目的地地址执行外出查找,如 803-805 处所示。如果

分组包括多播目的地地址,则对分组的目的地地址执行进入查找,如 803 和 805 处所示。

[0106] 注意,在某些实施例中,执行分布式转发所需的功能(例如通过实现执行图 5A、5B、7 和 8 所示的方法中的一种或多种的分布式转发模块)被实现在在虚拟网络设备子单元上执行的软件中。例如,每个虚拟网络设备子单元可包括计算机可读介质,可用于执行分布式转发的程序指令和/或数据被存储在其上。这种程序指令可执行以检测分组的接收,这些分组是由附属网络设备经由接口接收到的。计算机可读介质的示例性类型包括 CD(高密盘)、DVD(数字多功能盘)、硬盘、光盘、磁带设备、软盘和存储器(例如各种类型的 RAM(随机访问存储器)、ROM(只读存储器)、闪存、MEMS(微电子机械系统)存储器等等)。这种网络设备可以包括被配置为执行存储在计算机可读介质中的程序指令的一个或多个处理器(例如微处理器、PLD(可编程逻辑器件)或 ASIC(专用集成电路))。程序指令和/或数据也可以经由诸如因特网这样的网络或在承载介质上被传送到虚拟网络设备子单元。在某些实施例中,计算机可读介质是诸如网络和/或无线链路之类的承载介质,诸如电、电磁或数字信号之类的编码了数据和指令的信号在其上被输送。

[0107] 虽然已经针对本发明的特定实施例描述了本发明,但是可以向本领域的技术人员建议各种变化和修改。希望这种变化和修改落在所附权利要求书的范围之内。

[0108] 工业应用性

[0109] 本发明的实施例可用于联网领域。

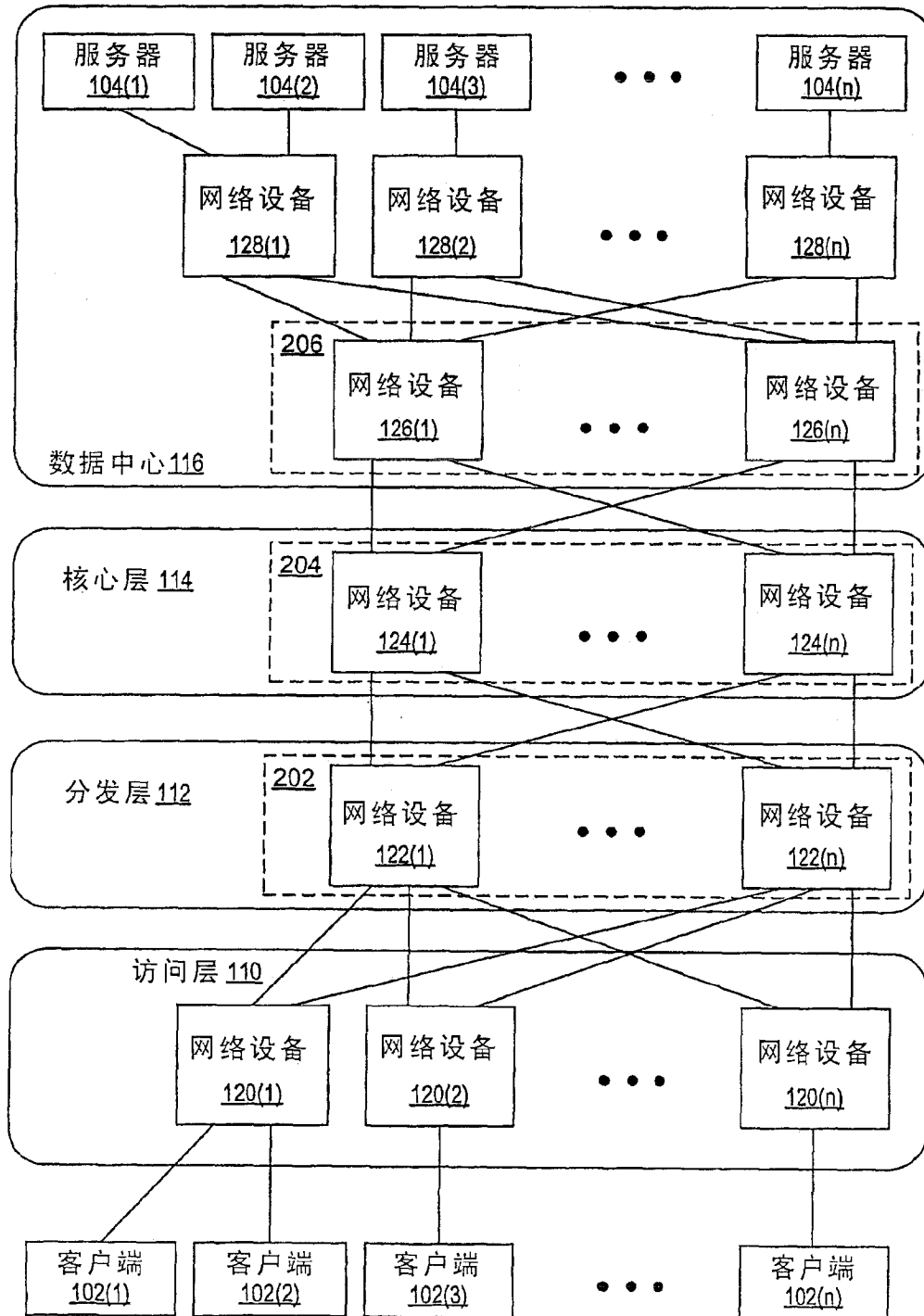


图 1

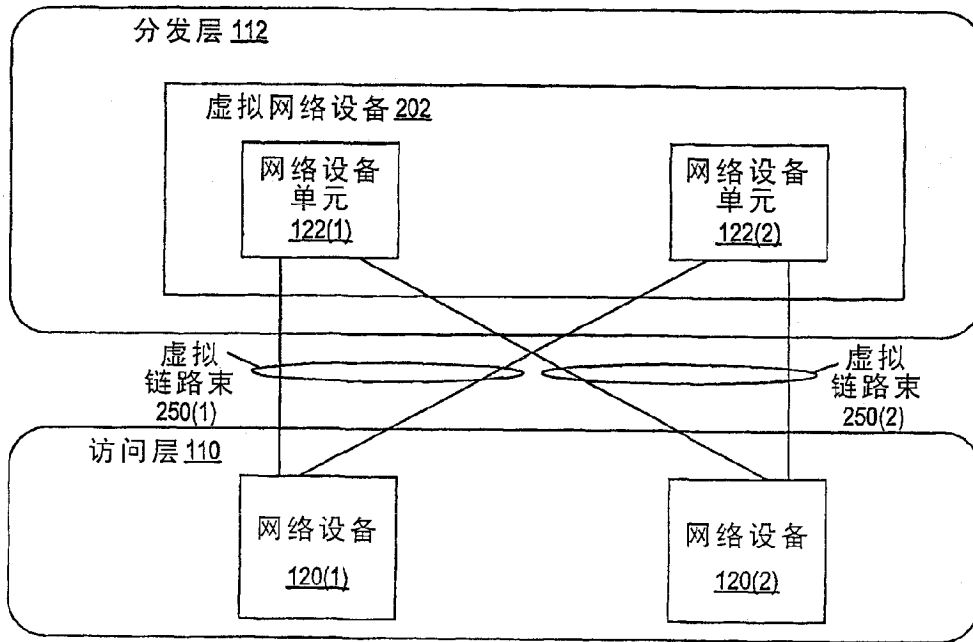


图 2A

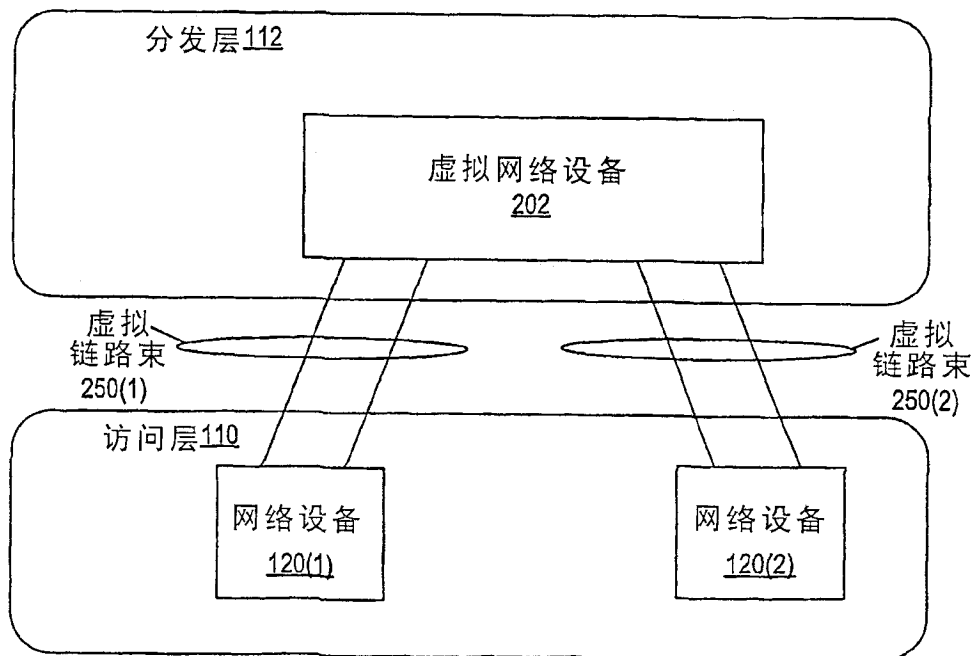


图 2B

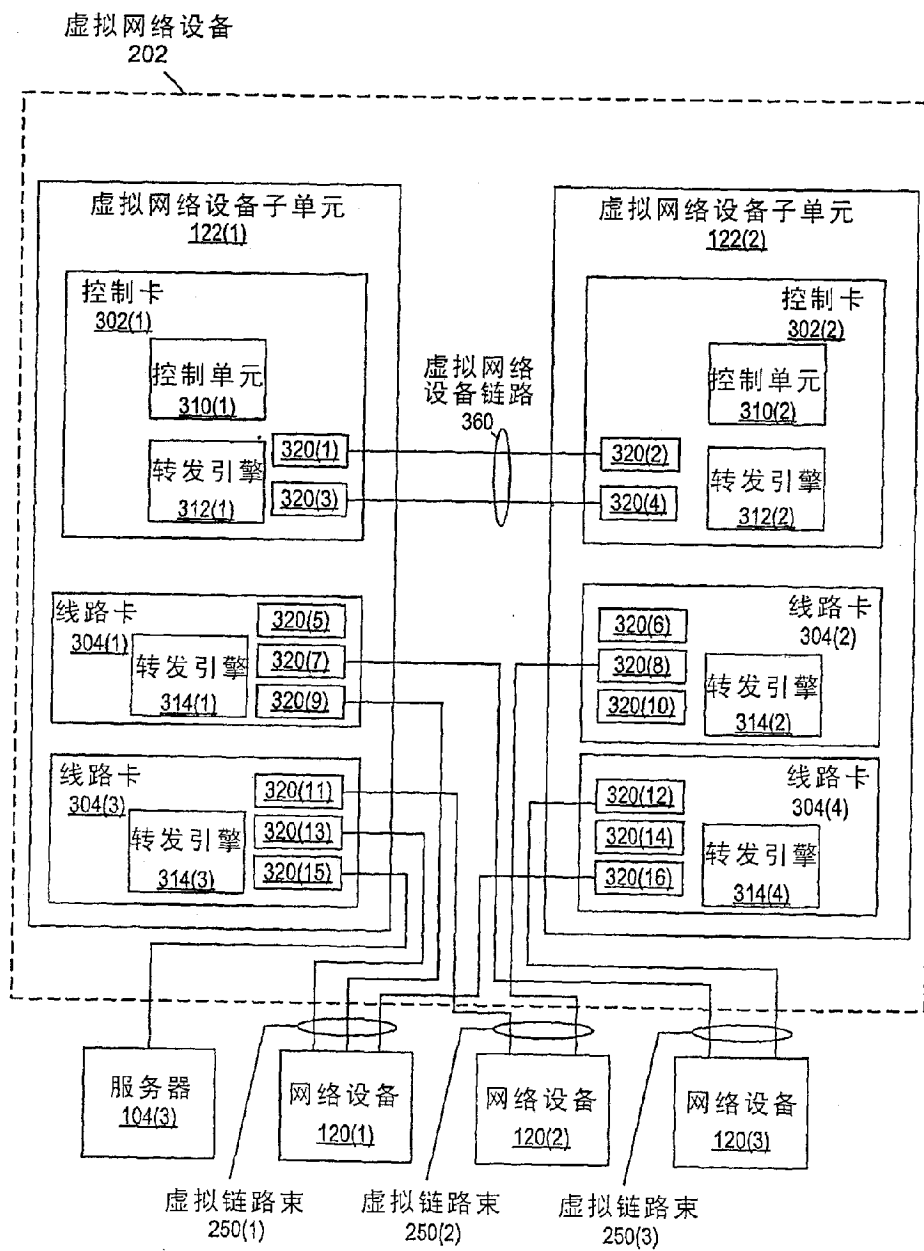


图 3

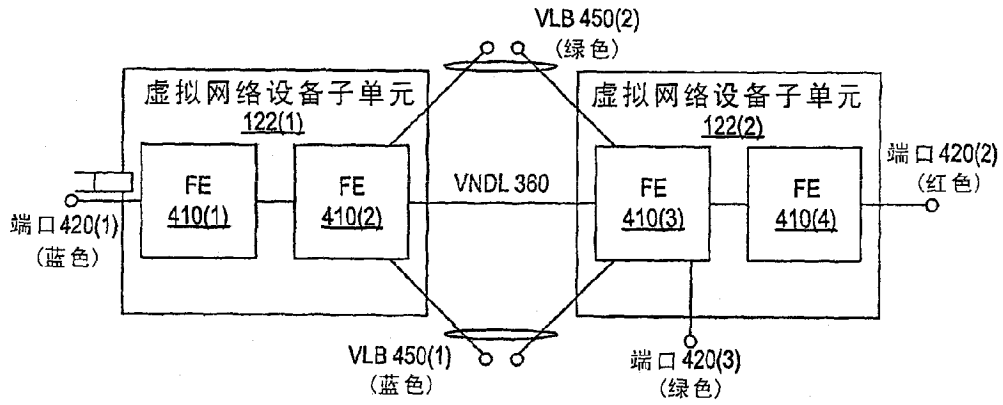


图 4A

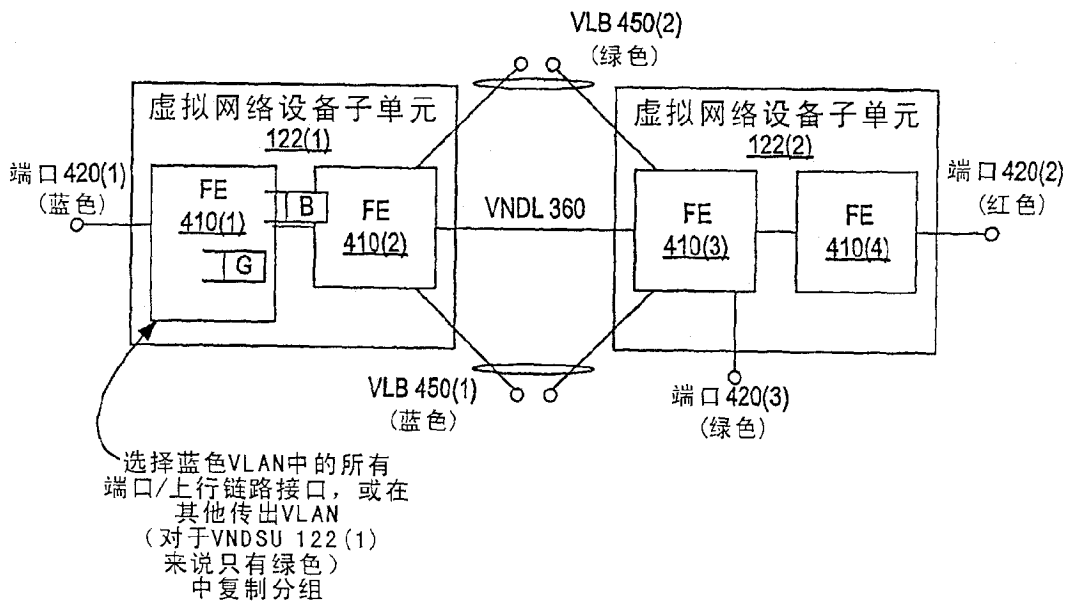


图 4B

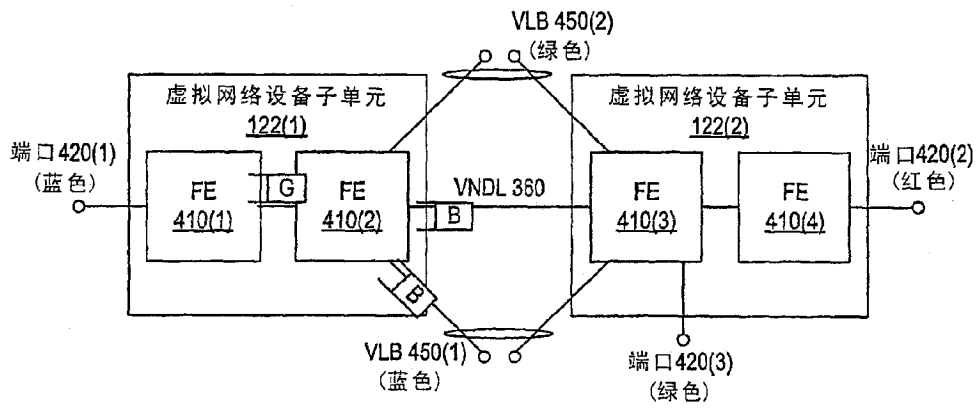


图 4C

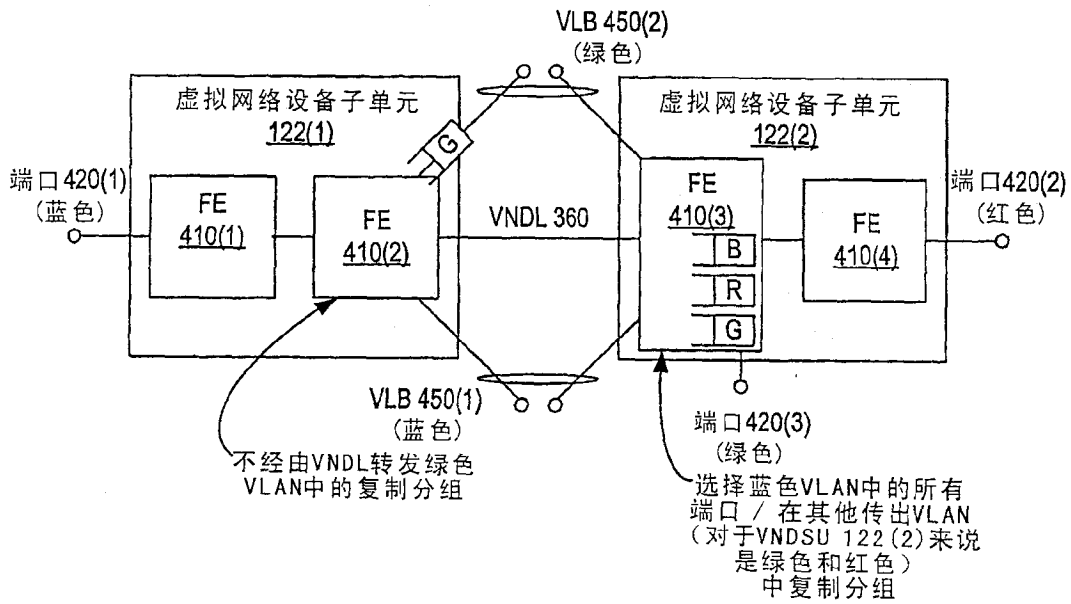


图 4D

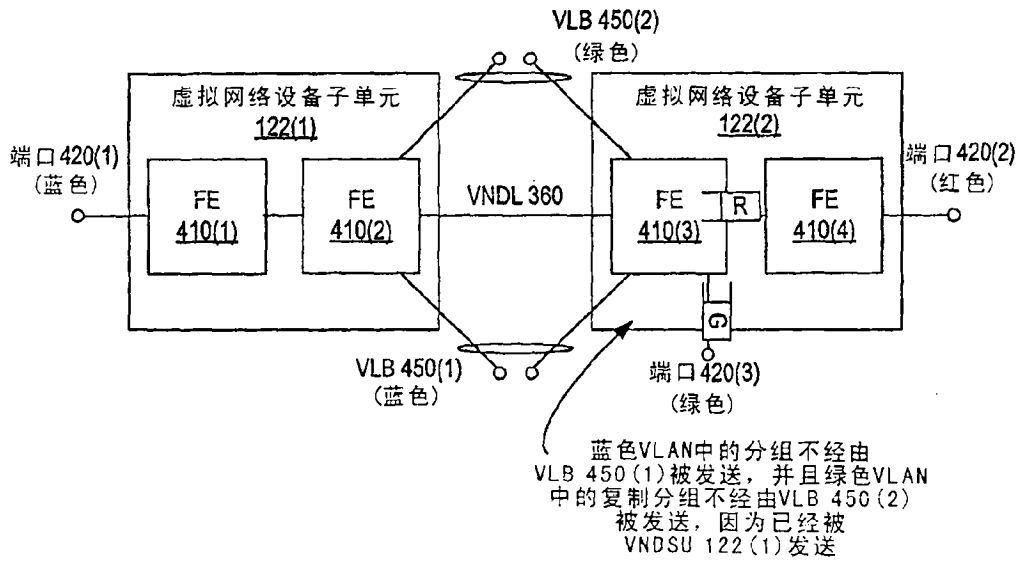


图 4E

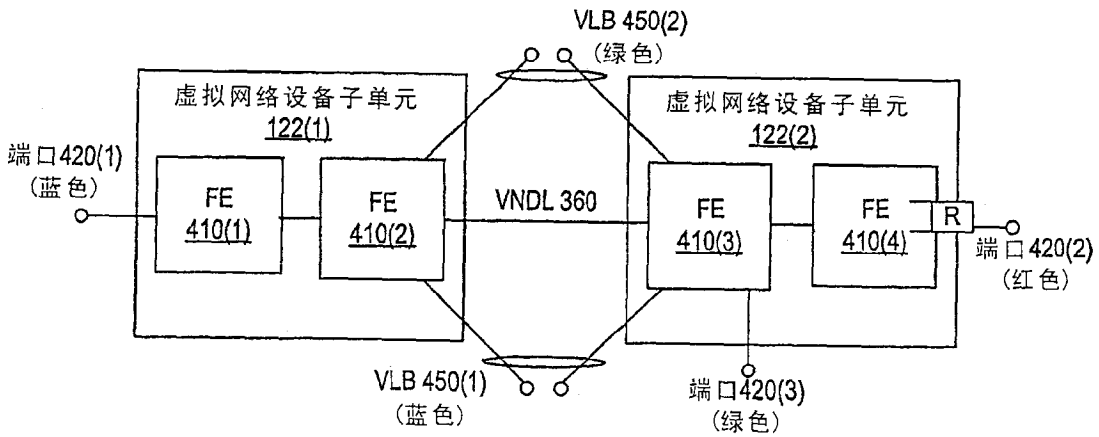


图 4F

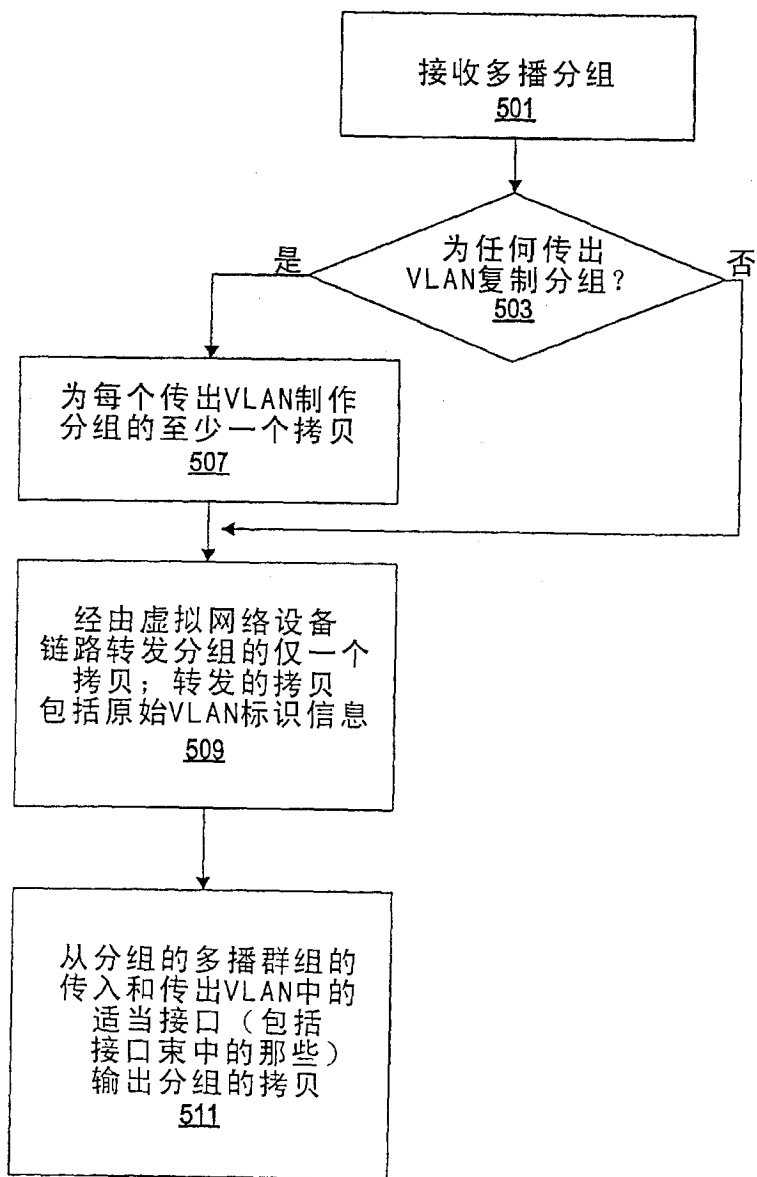


图 5A

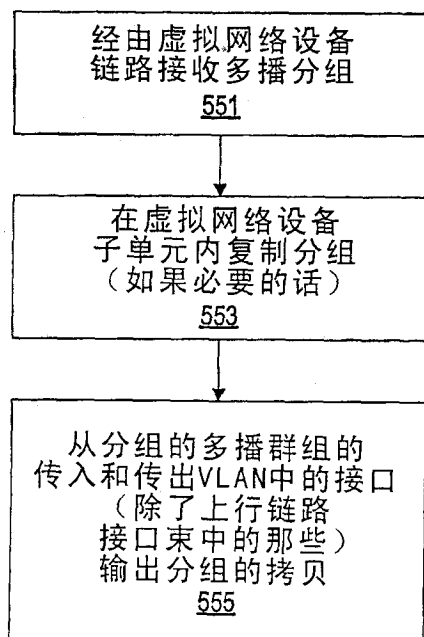


图 5B

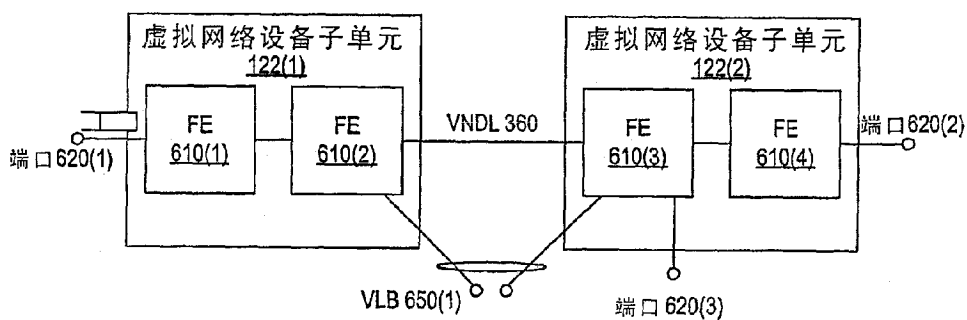


图 6A

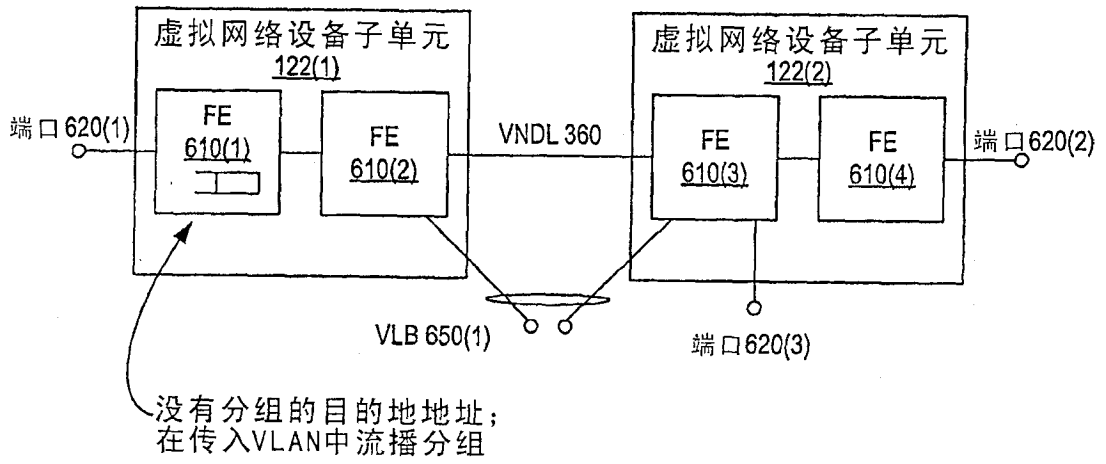


图 6B

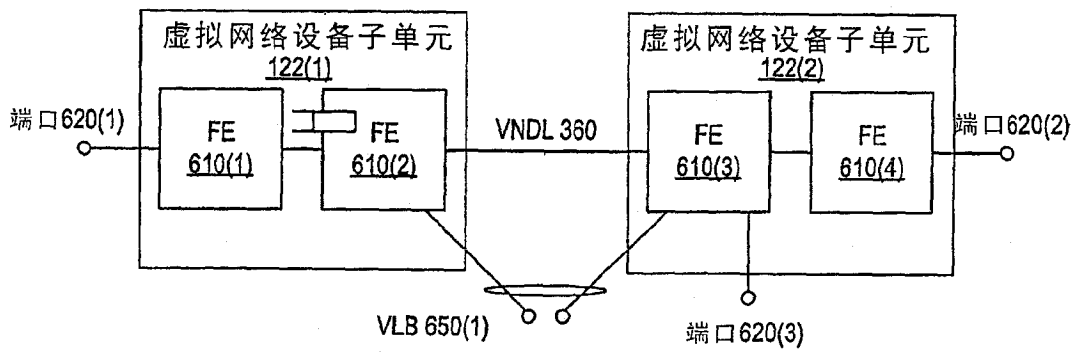


图 6C

FIG. 6C

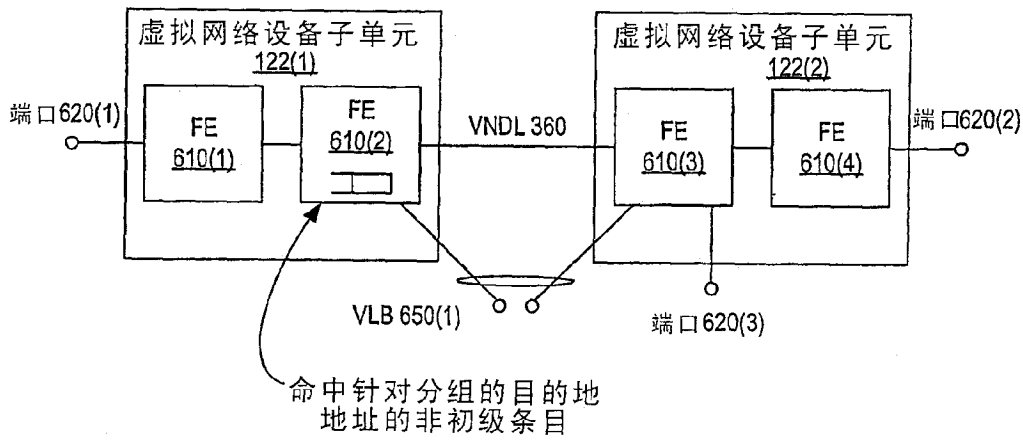


图 6D

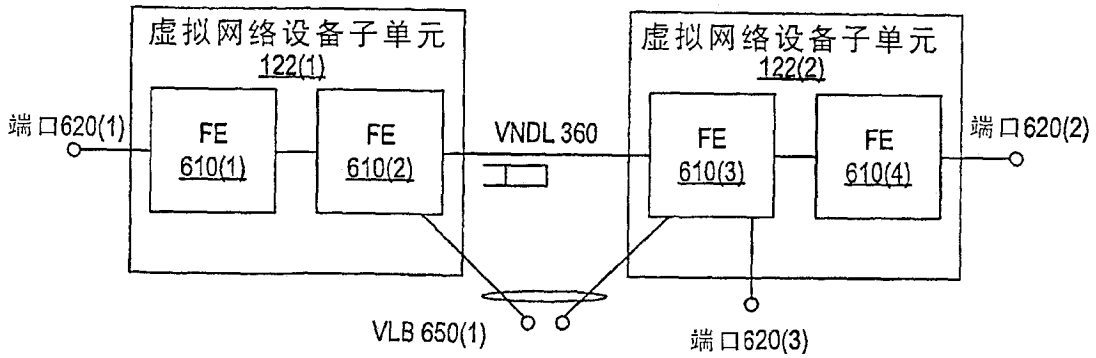


图 6E

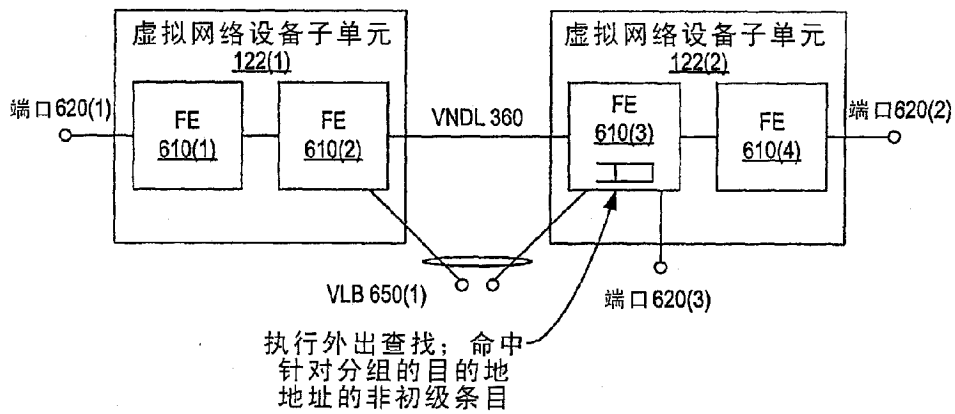


图 6F

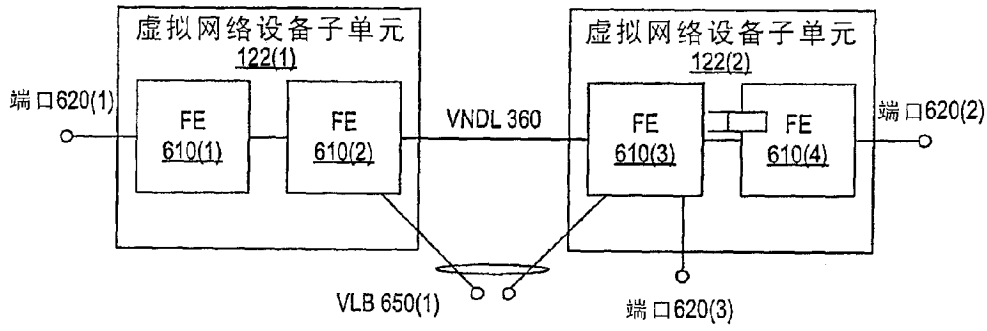


图 6G

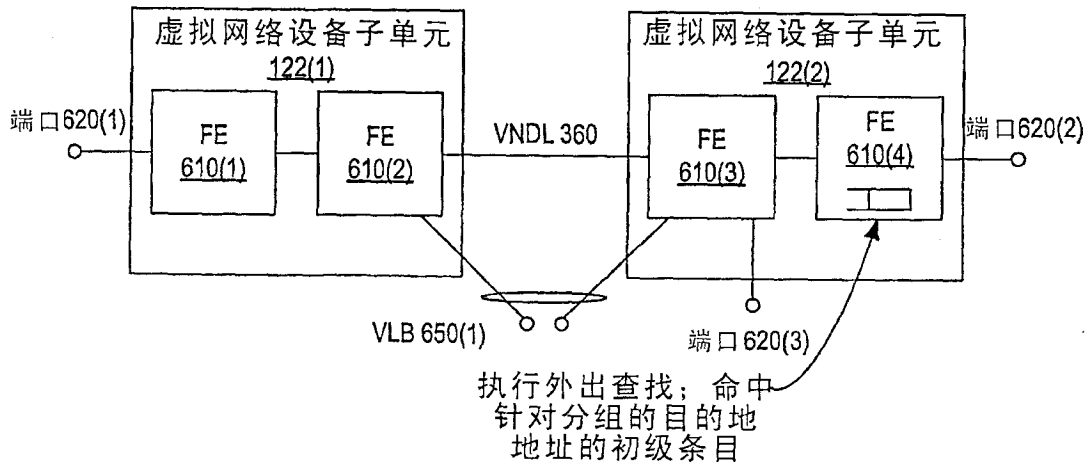


图 6H

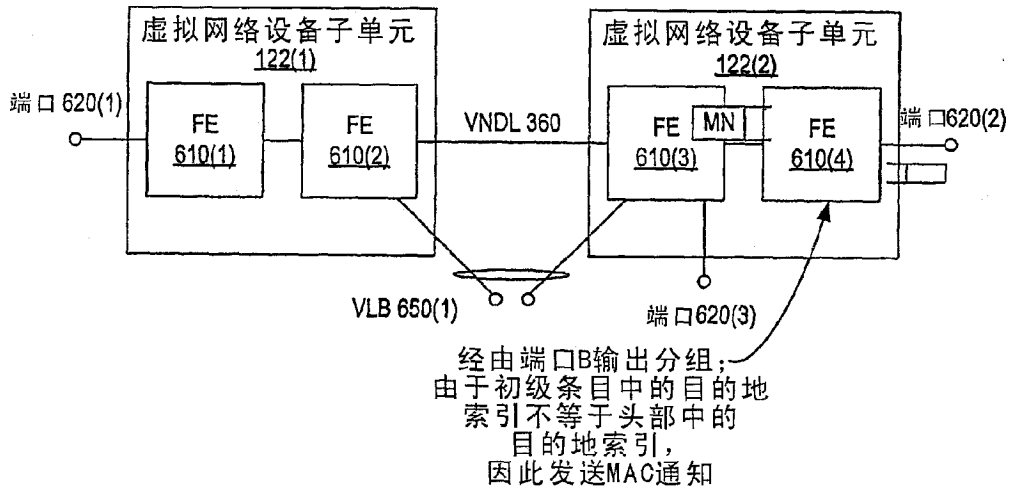
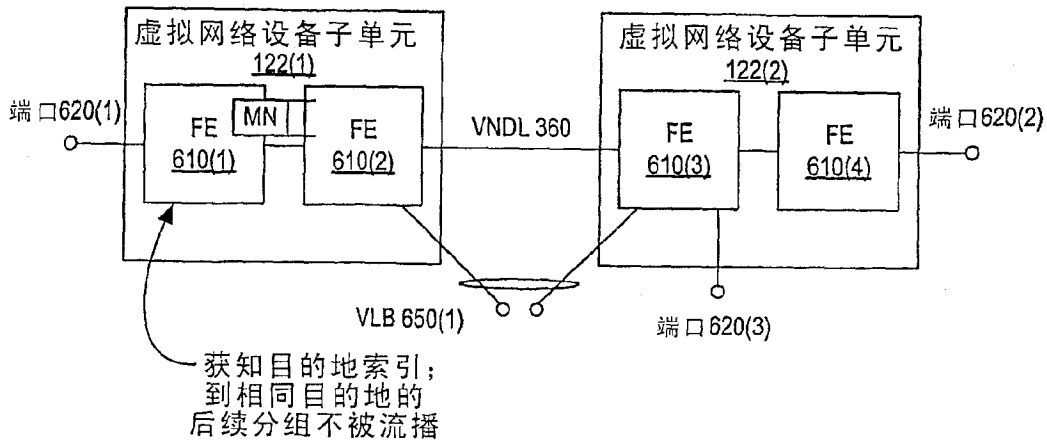
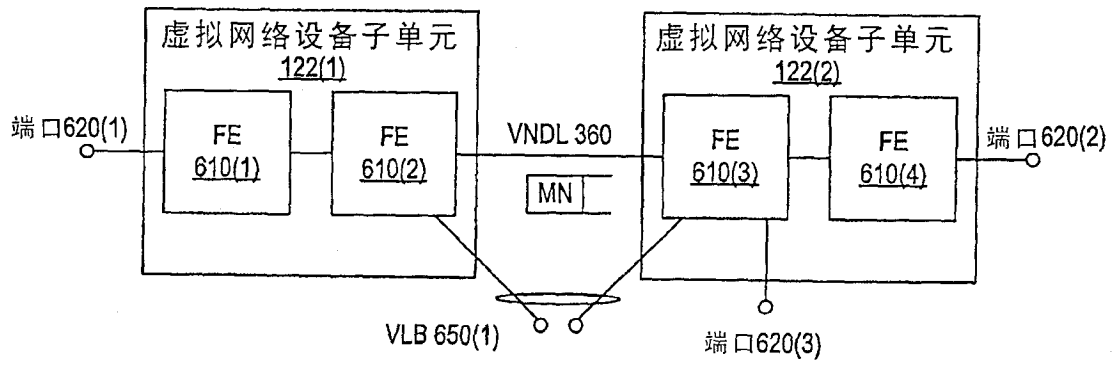


图 6I



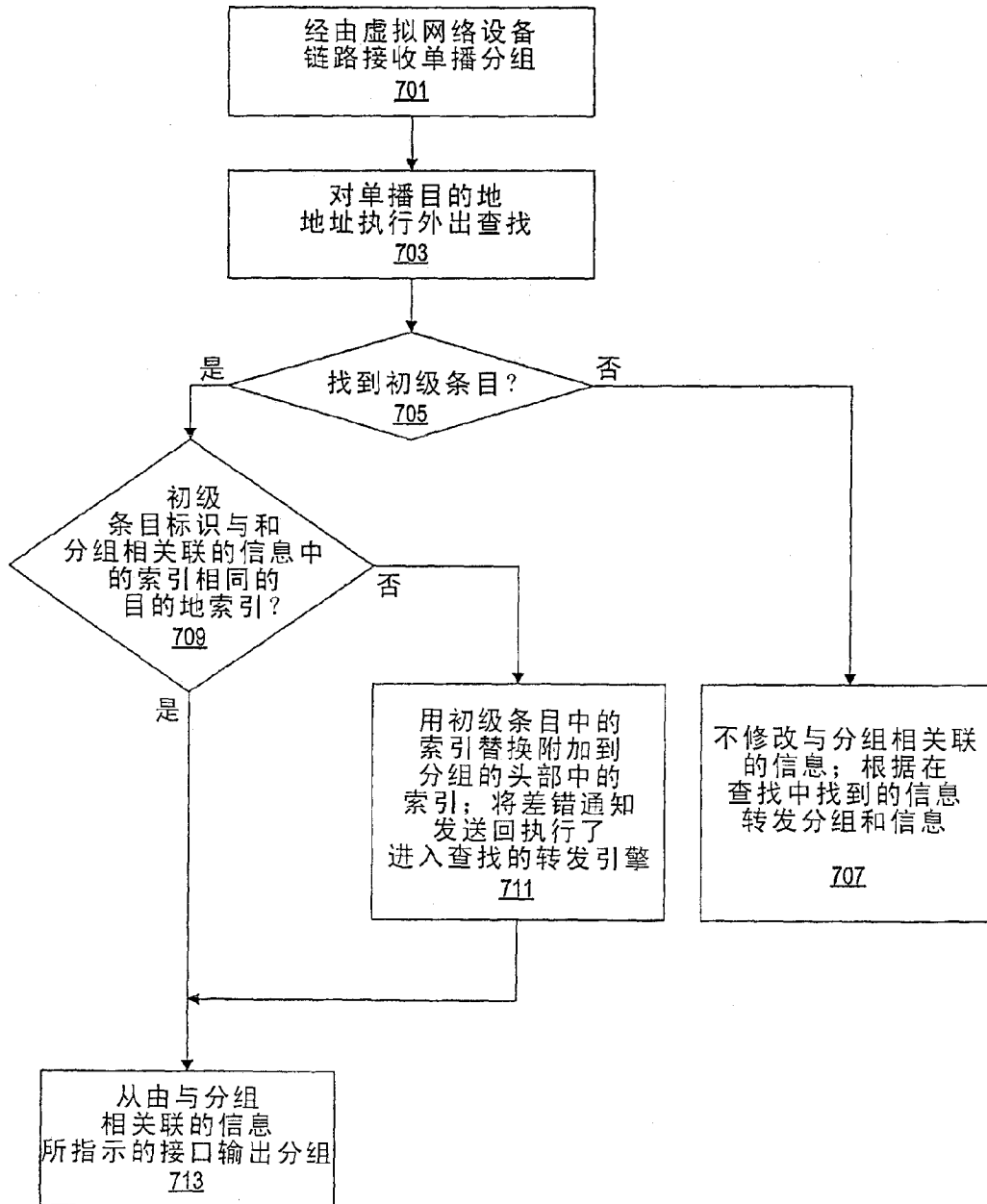


图 7

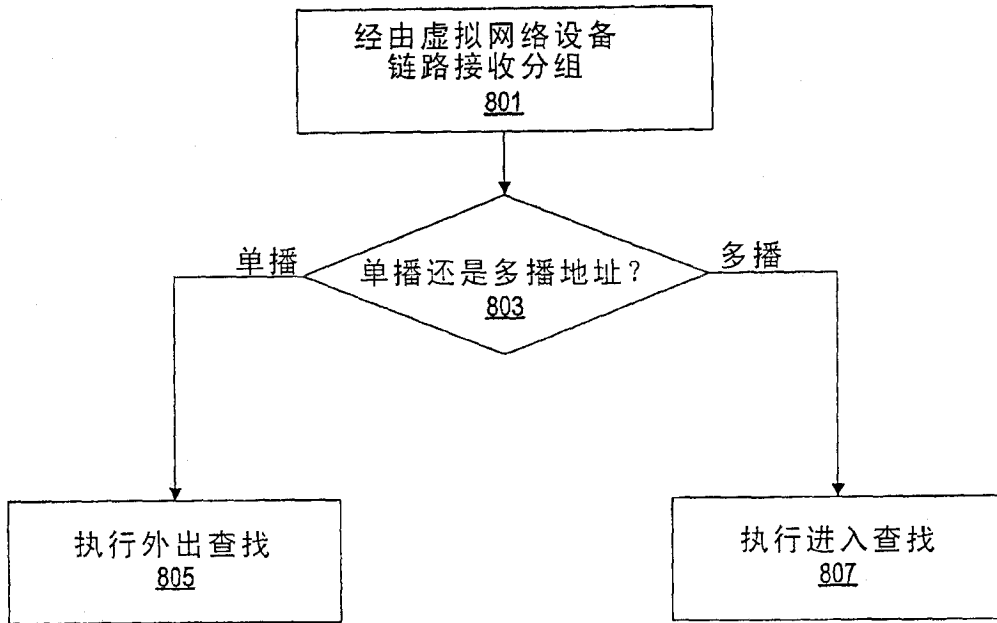


图 8