



(12) 发明专利申请

(10) 申请公布号 CN 106936609 A

(43) 申请公布日 2017. 07. 07

(21) 申请号 201511020591. X

(22) 申请日 2015. 12. 29

(71) 申请人 中兴通讯股份有限公司

地址 518057 广东省深圳市南山区高新技术产业园科技南路中兴通讯大厦法务部

(72) 发明人 许欣 李响

(74) 专利代理机构 北京安信方达知识产权代理有限公司 11262

代理人 解婷婷 龙洪

(51) Int. Cl.

H04L 12/24(2006. 01)

H04L 12/803(2013. 01)

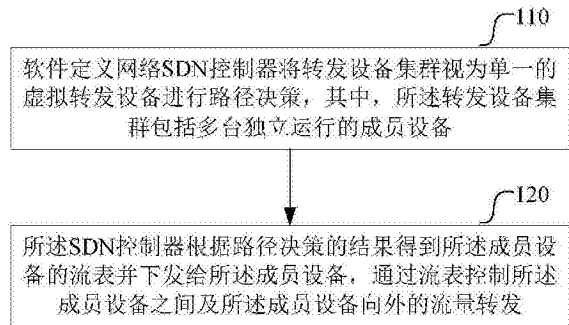
权利要求书6页 说明书10页 附图1页

(54) 发明名称

一种软件定义网络中控制转发设备集群的方法及控制器

(57) 摘要

一种软件定义网络中控制转发设备集群的方法和控制器, SDN 控制器将转发设备集群视为单一的虚拟转发设备进行路径决策, 其中, 所述转发设备集群包括多台独立运行的成员设备; 所述 SDN 控制器根据路径决策的结果得到所述成员设备的流表并下发给所述成员设备, 通过流表控制所述成员设备之间及所述成员设备向外的流量转发。本申请利用软件定义网络对网络设备进行集中控制的特点, 通过流表控制集群成员设备之间及成员设备向外的流量转发。成员设备之间无需运行私有协议, 可以实现异厂家互通, 且成员设备与控制器的链路维护无需特别处理, 比较简单。



1. 一种软件定义网络中控制转发设备集群的方法,包括:

软件定义网络 SDN 控制器将转发设备集群视为单一的虚拟转发设备进行路径决策,其中,所述转发设备集群包括多台独立运行的成员设备;

所述 SDN 控制器根据路径决策的结果得到所述成员设备的流表并下发给所述成员设备,通过流表控制所述成员设备之间及所述成员设备向外的流量转发。

2. 如权利要求 1 所述的方法,其特征在于:

所述路径决策包括:确定所述虚拟转发设备向外转发流量的作为出端口的虚拟端口,所述虚拟端口包括非捆绑端口和/或捆绑端口,其中:

每一非捆绑端口对应于一个成员设备上对外的一实体端口;

每一捆绑端口对应于一个或多个成员设备上对外的多个实体端口。

3. 如权利要求 2 所述的方法,其特征在于:

所述流量转发为单播流量转发时,所述 SDN 控制器通过流表控制所述成员设备之间及所述成员设备向外的流量转发,包括:

通过流表控制第一成员设备将流量从本设备直接向外转发,控制第二成员设备将流量从本设备向至少一个所述第一成员设备转发;其中,所述第一成员设备指具有与所述虚拟端口对应的实体端口的成员设备,所述第二成员设备指不具有与所述虚拟端口对应的实体端口的成员设备。

4. 如权利要求 3 所述的方法,其特征在于:

所述流量转发为单播流量转发时,所述 SDN 控制器根据路径决策的结果得到所述成员设备的流表,包括:

根据路径决策的结果生成所述虚拟转发设备的单播流表,所述单播流表包含作为出端口的所述虚拟端口;再针对每一所述成员设备,对所述单播流表进行修改,得到该成员设备的流表;

其中,针对每一所述第一成员设备,所述修改包括:将出端口替换为该第一成员设备上与所述虚拟端口对应的实体端口;针对每一所述第二成员设备,所述修改包括:将出端口替换为该第二成员设备上用于第一迂回链路的实体端口,所述第一迂回链路是该第二成员设备到所述第一成员设备的链路。

5. 如权利要求 3 所述的方法,其特征在于:

所述流量转发为单播流量转发时,所述 SDN 控制器根据路径决策的结果得到所述成员设备的流表并下发给所述成员设备之后,还包括:

在下发的流表中包括一个作为出端口的实体端口时,所述成员设备从该实体端口转发流量;

在下发的流表中包括多个作为出端口的实体端口时,所述成员设备从所述多个实体端口中选择一个实体端口转发流量。

6. 如权利要求 2 所述的方法,其特征在于:

所述流量转发为广播或组播流量转发时,所述 SDN 控制器通过流表控制所述成员设备之间及所述成员设备向外的流量转发,包括:

对每一个第一成员设备,通过流表控制该第一成员设备将流量从该第一成员设备直接向外转发,且对于来自所述转发设备集群外部的流量,还控制该第一成员设备将流量从该

第一成员设备向满足以下条件的其他第一成员设备转发：实体端口对应的虚拟端口中至少有一个不同于该第一成员设备上实体端口对应的虚拟端口；

对每一个第二成员设备，通过流表控制该第二成员设备将流量从该第二成员设备向所述转发设备集群中的第一成员设备转发，以将流量迂回到每一个所述虚拟端口再向外转发；

其中，所述第一成员设备指具有与所述虚拟端口对应的实体端口的成员设备，所述第二成员设备指不具有与所述虚拟端口对应的实体端口的成员设备。

7. 如权利要求 6 所述的方法，其特征在于：

所述流量转发为广播或组播流量转发时，所述 SDN 控制器根据路径决策的结果得到所述成员设备的流表，包括：

根据路径决策的结果生成所述虚拟转发设备的广播或组播流表，所述广播或组播流表包含作为出端口的所述虚拟端口；针对每一所述第一成员设备，对所述广播或组播流表进行第一修改和第二修改，分别得到该第一成员设备的第一流表和第二流表；

其中，所述第一修改包括：将表示报文来自集群内部的标志作为新增的匹配项，将出端口替换为该成员设备上与所述虚拟端口对应的实体端口；所述第二修改包括：将出端口替换为该成员设备上与所述虚拟端口对应的实体端口及用于第二迂回链路的实体端口，所述第二迂回链路是该第一成员设备到所述其他第一成员设备的链路。

8. 如权利要求 6 或 7 所述的方法，其特征在于：

所述流量转发为广播或组播流量转发时，所述 SDN 控制器根据路径决策的结果得到所述成员设备的流表，包括：

根据路径决策的结果生成所述虚拟转发设备的广播或组播流表，所述广播或组播流表包含作为出端口的所述虚拟端口；针对每一个所述第二成员设备，对所述广播或组播流表进行第三修改，得到该第二成员设备的流表；

其中所述第三修改包括：将出端口替换为该第二成员设备上用于第三迂回链路的实体端口，及增加命令或动作以在报文中打上表示报文来自集群内部的标志；所述第三迂回链路是该第二成员设备到所述第一成员设备的链路。

9. 如权利要求 8 所述的方法，其特征在于：

所述第三迂回链路包括除接收流量的链路之外的，该第二成员设备到满足以下条件一和条件二的第一成员设备的链路，以及按以下方式一选择的第一成员设备的链路：

条件一，该成员设备上实体端口对应的虚拟端口包括非捆绑端口；

条件二，该成员设备上实体端口对应的虚拟端口包括第一捆绑端口，所述第一捆绑端口指对应于一个成员设备上多个实体端口的捆绑端口；

方式一，从实体端口对应的虚拟端口均为第二捆绑端口且对应的所述第二捆绑端口相同的多个第一成员设备中，选择一个第一成员设备，其中，所述第二捆绑端口指对应于多个成员设备上多个实体端口的捆绑端口。

10. 如权利要求 5 或 9 所述的方法，其特征在于：

所述选择依据以下一种或多种策略进行：

负载均衡策略；

优先级策略；

随机选择策略。

11. 如权利要求 3-7、9 中任一所述的方法,其特征在于:

所述 SDN 控制器将流表下发给所述成员设备后,还包括:

所述 SDN 控制器确定所述转发设备集群中的一成员设备离线后,将该离线成员设备及其端口从所述转发设备集群存活的成员设备中删除;

对所述转发设备集群存活的每一成员设备,如果该成员设备的流表有出端口连接到该离线成员设备,将该成员设备的流表中连接到该离线成员设备的出端口删除,将更新后的该流表重新下发给该成员设备。

12. 如权利要求 3-7、9 中任一所述的方法,其特征在于:

所述 SDN 控制器将流表下发给所述成员设备后,还包括:

所述 SDN 控制器确定作为出端口的一虚拟端口的状态变为非活动后,将该虚拟端口对应的实体端口从相应流表的出端口中删除;

对删除后有出端口的每一流表,将更新后的该流表重新下发给对应的成员设备;

对删除后无出端口的每一流表,将该流表对应的成员设备上连接到所述第一成员设备的实体端口新增为该流表的出端口,将更新后的该流表重新下发给对应的成员设备。

13. 如权利要求 3-7、9 中任一所述的方法,其特征在于:

所述 SDN 控制器将流表下发给所述成员设备后,还包括:

所述 SDN 控制器确定一成员设备连接到一所述第一成员设备的实体端口或所述实体端口对应的链路失效后,将该成员设备的流表中作为出端口的所述实体端口更新为连接到另一所述第一成员设备的实体端口,将更新后的该流表重新下发给该成员设备。

14. 如权利要求 3-7、9 中任一所述的方法,其特征在于:

所述 SDN 控制器确定所述转发设备集群处于分裂状态后,记录处于非活动状态的虚拟端口和所述成员设备之间的互联端口,以及失效的所述成员设备之间的链路,并禁用所述成员设备上的联动端口;

所述 SDN 控制器对所述虚拟转发设备重新进行路径决策,根据路径决策的结果得到所述成员设备的流表并下发给所述成员设备。

15. 如权利要求 3-7、9 中任一所述的方法,其特征在于:

所述 SDN 控制器将流表下发给所述成员设备后,还包括:

所述 SDN 控制器确定有一成员设备加入所述转发设备集群后,根据所述路径决策的结果得到该成员设备的流表并下发给该成员设备,并在该成员设备具有与所述虚拟端口对应的实体端口时,按照所述转发设备集群新的拓扑对原有成员设备的流表进行更新并下发给相应的成员设备。

16. 一种软件定义网络中的控制器,其特征在于,包括集群控制装置,所述集群控制装置包括:

路径决策模块,用于将转发设备集群视为单一的虚拟转发设备进行路径决策,其中,所述转发设备集群包括多台独立运行的成员设备;

流表生成模块,用于根据路径决策的结果得到所述成员设备的流表,通过流表控制所述成员设备之间及所述成员设备向外的流量转发;

流表下发模块,用于将所述成员设备的流表下发给所述成员设备。

17. 如权利要求 16 所述的控制器,其特征在于:

所述路径决策模块进行路径决策,包括:确定所述虚拟转发设备向外转发流量的作为出端口的虚拟端口,所述虚拟端口包括非捆绑端口和/或捆绑端口,其中:每一非捆绑端口对应于一个成员设备上对外的一实体端口;每一捆绑端口对应于一个或多个成员设备上对外的多个实体端口。

18. 如权利要求 17 所述的控制器,其特征在于:

所述流表生成模块通过流表控制所述成员设备之间及所述成员设备向外的流量转发,包括:所述流量转发为单播流量转发时,通过流表控制第一成员设备将流量从本设备直接向外转发,控制第二成员设备将流量从本设备向至少一个所述第一成员设备转发;其中,所述第一成员设备指具有与所述虚拟端口对应的实体端口的成员设备,所述第二成员设备指不具有与所述虚拟端口对应的实体端口的成员设备。

19. 如权利要求 18 所述的控制器,其特征在于:

所述流表生成模块根据路径决策的结果得到所述成员设备的流表,包括:所述流量转发为单播流量转发时,根据路径决策的结果生成所述虚拟转发设备的单播流表,所述单播流表包含作为出端口的所述虚拟端口;再针对每一所述成员设备,对所述单播流表进行修改,得到该成员设备的流表;

其中,针对每一所述第一成员设备,所述修改包括:将出端口替换为该第一成员设备上与所述虚拟端口对应的实体端口;针对每一所述第二成员设备,所述修改包括:将出端口替换为该第二成员设备上用于第一迂回链路的实体端口,所述第一迂回链路是该第二成员设备到所述第一成员设备的链路。

20. 如权利要求 17 所述的控制器,其特征在于:

所述流表生成模块通过流表控制所述成员设备之间及所述成员设备向外的流量转发,包括:所述流量转发为广播或组播流量转发时,对每一个第一成员设备,通过流表控制该第一成员设备将流量从该第一成员设备直接向外转发,且对于来自所述转发设备集群外部的流量,还控制该第一成员设备将流量从该第一成员设备向满足以下条件的其他第一成员设备转发:实体端口对应的虚拟端口中至少有一个不同于该第一成员设备上实体端口对应的虚拟端口;及,对每一个第二成员设备,通过流表控制该第二成员设备将流量从该第二成员设备向所述转发设备集群中的第一成员设备转发,以将流量迂回到每一个所述虚拟端口再向外转发;其中,所述第一成员设备指具有与所述虚拟端口对应的实体端口的成员设备,所述第二成员设备指不具有与所述虚拟端口对应的实体端口的成员设备。

21. 如权利要求 20 所述的控制器,其特征在于:

所述流表生成模块根据路径决策的结果得到所述成员设备的流表,包括:所述流量转发为广播或组播流量转发时,根据路径决策的结果生成所述虚拟转发设备的广播或组播流表,所述广播或组播流表包含作为出端口的所述虚拟端口;针对每一所述第一成员设备,对所述广播或组播流表进行第一修改和第二修改,分别得到该第一成员设备的第一流表和第二流表;

其中,所述第一修改包括:将表示报文来自集群内部的标志作为新增的匹配项,将出端口替换为该成员设备上与所述虚拟端口对应的实体端口;所述第二修改包括:将出端口替换为该成员设备上与所述虚拟端口对应的实体端口及用于第二迂回链路的实体端口,所述

第二迂回链路是该第一成员设备到所述其他第一成员设备的链路。

22. 如权利要求 20 或 21 所述的控制器,其特征在于:

所述流表生成模块根据路径决策的结果得到所述成员设备的流表,包括:所述流量转发为广播或组播流量转发时,根据路径决策的结果生成所述虚拟转发设备的广播或组播流表,所述广播或组播流表包含作为出端口的所述虚拟端口;针对每一个所述第二成员设备,对所述广播或组播流表进行第三修改,得到该第二成员设备的流表;

其中所述第三修改包括:将出端口替换为该第二成员设备上用于第三迂回链路的实体端口,及增加命令或动作以在报文中打上表示报文来自集群内部的标志;所述第三迂回链路是该第二成员设备到所述第一成员设备的链路。

23. 如权利要求 22 所述的控制器,其特征在于:

所述第三迂回链路包括除接收流量的链路之外的,该第二成员设备到满足以下条件一和条件二的第一成员设备的链路,以及按以下方式一选择的第一成员设备的链路:

条件一,该成员设备上实体端口对应的虚拟端口包括非捆绑端口;

条件二,该成员设备上实体端口对应的虚拟端口包括第一捆绑端口,所述第一捆绑端口指对应于一个成员设备上多个实体端口的捆绑端口;

方式一,从实体端口对应的虚拟端口均为第二捆绑端口且对应的所述第二捆绑端口相同的多个第一成员设备中,选择一个第一成员设备,其中,所述第二捆绑端口指对应于多个成员设备上多个实体端口的捆绑端口;

其中所述选择依据以下一种或多种策略进行:负载均衡策略;优先级策略;及随机选择策略。

24. 如权利要求 17-21、23 中任一所述的控制器,其特征在于:

所述集群控制装置还包括以下模块中的一种或多种:

离线处理模块,用于在确定所述转发设备集群中的一成员设备离线后,将该离线成员设备及其端口从所述转发设备集群存活的成员设备中删除;及对所述转发设备集群存活的每一成员设备,如果该成员设备的流表有出端口连接到该离线成员设备,将该成员设备的流表中连接到该离线成员设备的出端口删除,将更新后的该流表重新下发给该成员设备;

端口失效处理模块,用于在确定作为出端口的一虚拟端口的状态变为非活动后,将该虚拟端口对应的实体端口从相应流表的出端口中删除;及,对删除后有出端口的每一流表,将更新后的该流表重新下发给对应的成员设备;对删除后无出端口的每一流表,将该流表对应的成员设备上连接到所述第一成员设备的实体端口新增为该流表的出端口,将更新后的该流表重新下发给对应的成员设备;

通信故障处理模块,用于在确定一成员设备连接到一所述第一成员设备的实体端口或所述实体端口对应的链路失效后,将该成员设备的流表中作为出端口的所述实体端口更新为连接到另一所述第一成员设备的实体端口,将更新后的该流表重新下发给该成员设备;

集群分裂处理模块,用于在确定所述转发设备集群处于分裂状态后,记录处于非活动状态的虚拟端口和所述成员设备之间的互联端口,以及失效的所述成员设备之间的链路,并禁用所述成员设备上的联动端口;及,对所述虚拟转发设备重新进行路径决策,根据路径决策的结果得到所述成员设备的流表并下发给所述成员设备;

成员加入处理模块,所述 SDN 控制器确定有一成员设备加入所述转发设备集群后,根

据所述路径决策的结果得到该成员设备的流表并下发给该成员设备,在该成员设备具有与所述虚拟端口对应的实体端口时,按照转发设备集群新的拓扑对原有成员设备的流表进行更新并下发给相应的成员设备。

## 一种软件定义网络中控制转发设备集群的方法及控制器

### 技术领域

[0001] 本发明涉及软件定义网络,更具体地,涉及一种软件定义网络中控制转发设备集群的方法及相应的控制器。

### 背景技术

[0002] 软件定义网络 (Software Defined Networking, SDN) 是一种新型网络创新架构,其核心思想是将网络的控制和转发平面分离,增加网络管理的灵活性和可扩展性。SDN 中的控制层由原有的交换机、路由器等网络设备中的控制功能抽取出来,采用独立的控制软件来实现。原有的交换机、路由器中的控制功能都从基础设施层剥离,由控制面统一实施转发管理。OpenFlow 是 SDN 网络中运行在控制器与转发设备间的一种协议,控制器通过 OpenFlow 协议向交换机写入流表和组表等转发表项,完成对整个网络的控制。

[0003] 类似于传统网络,在 SDN 网络中,为实现负载均衡和高可靠性,也需要在网络的特定位置部署多台转发设备组成集群,以主备或负载分担方式运行。为满足此要求,可以部署采用传统集群技术的设备,转发设备集群自行实现负载均衡转发或故障时的主备倒换,对 SDN 控制器呈现为单一的设备。

[0004] 但是,采用传统集群技术的转发设备集群,集群的成员设备之间需要运行私有协议,无法实现异厂家互通。呈现为整体的多个集群设备成员与 SDN 控制器维护单一链接的机制也较复杂。

### 发明内容

[0005] 有鉴于此,本发明提供了以下技术方案。

[0006] 一种软件定义网络中控制转发设备集群的方法,包括:

[0007] 软件定义网络 SDN 控制器将转发设备集群视为单一的虚拟转发设备进行路径决策,其中,所述转发设备集群包括多台独立运行的成员设备;

[0008] 所述 SDN 控制器根据路径决策的结果得到所述成员设备的流表并下发给所述成员设备,通过流表控制所述成员设备之间及所述成员设备向外的流量转发。

[0009] 一种软件定义网络中的控制器,其特征在于,包括集群控制装置,所述集群控制装置包括:

[0010] 路径决策模块,用于将转发设备集群视为单一的虚拟转发设备进行路径决策,其中,所述转发设备集群包括多台独立运行的成员设备;

[0011] 流表生成模块,用于根据路径决策的结果得到所述成员设备的流表,通过流表控制所述成员设备之间及所述成员设备向外的流量转发;

[0012] 流表下发模块,用于将所述成员设备的流表下发给所述成员设备。

[0013] 上述方案充分利用软件定义网络对网络设备进行集中控制的特点,将独立运行的多台转发设备作为集群,通过流表控制集群成员设备之间及成员设备向外的流量转发。成员设备之间无需运行私有协议,可以实现异厂家互通,且成员设备与控制器的链路维护

无需特别处理,比较简单。

## 附图说明

- [0014] 图 1 是本发明实施例一控制方法的流程图；  
[0015] 图 2 是本发明实施例一 SDN 控制器的模块图；  
[0016] 图 3 是本发明示例一转发设备集群的示意图。

## 具体实施方式

[0017] 为使本发明的目的、技术方案和优点更加清楚明白,下文中将结合附图对本发明的实施例进行详细说明。需要说明的是,在不冲突的情况下,本申请中的实施例及实施例中的特征可以相互任意组合。

### [0018] 实施例一

[0019] 本实施例提出一种软件定义网络中控制转发设备集群的方法,根据管理员的配置,SDN 控制器通过 OpenFlow 协议(或其他流控制协议)对独立运行的转发设备进行控制,组成转发设备集群,对外呈现为单一的虚拟转发设备。

[0020] 本实施例软件定义网络中控制转发设备集群的方法如图 1 所示,包括:

[0021] 步骤 110,SDN 控制器将转发设备集群视为单一的虚拟转发设备进行路径决策,其中,所述转发设备集群包括多台独立运行的成员设备;

[0022] 本步骤中,所述路径决策包括:确定所述虚拟转发设备向外转发流量(也即转发报文)的作为出端口的虚拟端口,所述虚拟端口可以分为两类:非捆绑端口和/或捆绑端口,其中:每一非捆绑端口对应于一个成员设备上对外的一实体端口;每一捆绑端口对应于一个或多个成员设备上对外的多个实体端口。

[0023] 步骤 120,所述 SDN 控制器根据路径决策的结果得到所述成员设备的流表并下发给所述成员设备,通过流表控制所述成员设备之间及所述成员设备向外的流量转发。

[0024] 文中,成员设备向外的流量转发是指向转发设备集群外的流量转发。

[0025] 所述流量转发为单播流量转发时,所述 SDN 控制器通过流表控制所述成员设备之间及所述成员设备向外的流量转发,包括:

[0026] 通过流表控制第一成员设备将流量从本设备直接向外转发,控制第二成员设备将流量从本设备向至少一个所述第一成员设备转发;其中,所述第一成员设备指具有与所述虚拟端口对应的实体端口的成员设备,所述第二成员设备指不具有与所述虚拟端口对应的实体端口的成员设备。

[0027] 为了实现上述单播流量转发的控制,所述 SDN 控制器根据路径决策的结果得到所述成员设备的流表,可以采用以下方式:根据路径决策的结果生成所述虚拟转发设备的单播流表,所述单播流表包含作为出端口的所述虚拟端口;再针对每一所述成员设备,对所述单播流表进行修改,得到该成员设备的流表;其中,针对每一所述第一成员设备,所述修改包括:将出端口替换为该第一成员设备上与所述虚拟端口对应的实体端口;针对每一所述第二成员设备,所述修改包括:将出端口替换为该第二成员设备上用于第一迂回链路的实体端口,所述第一迂回链路是该第二成员设备到所述第一成员设备的链路。

[0028] 在单播流量转发时,所述成员设备在下发的流表中包括一个作为出端口的实体端

口时,从该实体端口转发流量;而在下发的流表中包括多个作为出端口的实体端口时,可以从所述多个实体端口中选择一个实体端口转发流量。具体选择可以根据以下一种或多种(两种以上,任意组合)策略进行:负载均衡策略;优先级策略;及随机选择策略。控制器给成员设备下发的端口列表中,还可以为各端口设置权重,供成员设备选择出端口时使用。

[0029] 所述流量转发为广播或组播流量转发时,所述 SDN 控制器通过流表控制所述成员设备之间及所述成员设备向外的流量转发,包括:

[0030] 对每一个第一成员设备,通过流表控制该第一成员设备将流量从该第一成员设备直接向外转发,且对于来自所述转发设备集群外部的流量,还控制该第一成员设备将流量从该第一成员设备向满足以下条件的其他第一成员设备转发:实体端口对应的虚拟端口中至少有一个不同于该第一成员设备上实体端口对应的虚拟端口;

[0031] 对每一个第二成员设备,通过流表控制该第二成员设备将流量从该第二成员设备向所述转发设备集群中的第一成员设备转发,以将流量迂回到每一个所述虚拟端口再向外转发;

[0032] 其中,所述第一成员设备指具有与所述虚拟端口对应的实体端口的成员设备,所述第二成员设备指不具有与所述虚拟端口对应的实体端口的成员设备。

[0033] 为了实现上述广播或组播流量转发的控制,所述 SDN 控制器根据路径决策的结果得到所述成员设备的流表时,可以采用以下方式:

[0034] 根据路径决策的结果生成所述虚拟转发设备的广播或组播流表,所述广播或组播流表包含作为出端口的所述虚拟端口;

[0035] 针对每一所述第一成员设备,对所述广播或组播流表进行第一修改和第二修改,分别得到该第一成员设备的第一流表和第二流表;

[0036] 针对每一个所述第二成员设备,对所述广播或组播流表进行第三修改,得到该第二成员设备的流表;

[0037] 其中,所述第一修改包括:将表示报文来自集群内部的标志作为新增的匹配项,将出端口替换为该成员设备上与所述虚拟端口对应的实体端口;所述第二修改包括:将出端口替换为该成员设备上与所述虚拟端口对应的实体端口及用于第二迂回链路的实体端口,所述第二迂回链路是该第一成员设备到所述其他第一成员设备的链路;所述第三修改包括:将出端口替换为该第二成员设备上用于第三迂回链路的实体端口,及增加命令或动作以在报文中打上表示报文来自集群内部的标志;所述第三迂回链路是该第二成员设备到所述第一成员设备的链路。

[0038] 特别地,所述第三迂回链路可以包括除接收流量的链路之外的,该第二成员设备到满足以下条件一和条件二的第一成员设备的链路,以及按以下方式一选择的第一成员设备的链路:

[0039] 条件一,该成员设备上实体端口对应的虚拟端口包括非捆绑端口;

[0040] 条件二,该成员设备上实体端口对应的虚拟端口包括第一捆绑端口,所述第一捆绑端口指对应于一个成员设备上多个实体端口的捆绑端口;

[0041] 方式一,从实体端口对应的虚拟端口均为第二捆绑端口且对应的所述第二捆绑端口相同的多个第一成员设备中,选择一个第一成员设备,其中,所述第二捆绑端口指对应于多个成员设备上多个实体端口的捆绑端口,所述选择依据以下一种或多种策略进行:负载

均衡策略；优先级策略；及随机选择策略。

[0042] 本实施例针对转发设备集群可能出现的离线、故障、加入新设备等情况，还提出了相应的流表控制方法，如下：

[0043] 针对离线的情况：

[0044] 所述 SDN 控制器确定所述转发设备集群中的一成员设备离线后，将该离线成员设备及其端口从所述转发设备集群存活的成员设备中删除；对所述转发设备集群存活的每一成员设备，如果该成员设备的流表有出端口连接到该离线成员设备，将该成员设备的流表中连接到该离线成员设备的出端口删除，将更新后的该流表重新下发给该成员设备。

[0045] 针对端口失效的情况：

[0046] 所述 SDN 控制器确定作为出端口的一虚拟端口的状态变为非活动（即失效）后，将该虚拟端口对应的实体端口从相应流表的出端口中删除；对删除后有出端口的每一流表，将更新后的该流表重新下发给对应的成员设备；对删除后无出端口的每一流表，将该流表对应的成员设备上连接到所述第一成员设备的实体端口新增为该流表的出端口，将更新后的该流表重新下发给对应的成员设备。其中，对于一个非捆绑端口，其对应的一个实体端口变为非活动后，则确定该非捆绑端口变为非活动。对于一个捆绑端口，在其对应的多个实体端口均变为非活动后，再确定该捆绑端口变为非活动。

[0047] 针对成员设备间通信故障的情况：

[0048] 所述 SDN 控制器确定一成员设备连接到一所述第一成员设备的实体端口或所述实体端口对应的链路失效后，将该成员设备的流表中作为出端口的所述实体端口更新为连接到另一所述第一成员设备的实体端口，将更新后的该流表重新下发给该成员设备。

[0049] 针对集群分裂的情况：

[0050] 所述 SDN 控制器确定所述转发设备集群处于分裂状态（即集群变为互不相通的两个或更多部分）后，记录处于非活动状态的虚拟端口和所述成员设备之间的互联端口，以及失效的所述成员设备之间的链路，并禁用所述成员设备上的联动端口；及对所述虚拟转发设备重新进行路径决策，根据路径决策的结果得到所述成员设备的流表并下发给所述成员设备。

[0051] 针对加入新成员设备的情况：

[0052] 所述 SDN 控制器确定有一成员设备加入所述转发设备集群后，根据所述路径决策的结果得到该成员设备的流表并下发给该成员设备，并在该成员设备具有与所述虚拟端口对应的实体端口（如新加入的成员设备有实体端口加入现有的捆绑端口）时，按照所述转发设备集群新的拓扑对原有成员设备的流表进行更新并下发给相应的成员设备。

[0053] 本实施例还提供了一种软件定义网络中的控制器，包括集群控制装置，如图 2 所示，所述集群控制装置包括：

[0054] 路径决策模块 10，用于将转发设备集群视为单一的虚拟转发设备进行路径决策，其中，所述转发设备集群包括多台独立运行的成员设备；

[0055] 流表生成模块 20，用于根据路径决策的结果得到所述成员设备的流表，通过流表控制所述成员设备之间及所述成员设备向外的流量转发；

[0056] 流表下发模块 30，用于将所述成员设备的流表下发给所述成员设备。

[0057] 可选地，

[0058] 所述路径决策模块进行路径决策,包括:确定所述虚拟转发设备向外转发流量的作为出端口的虚拟端口,所述虚拟端口包括非捆绑端口和/或捆绑端口,其中:每一非捆绑端口对应于一个成员设备上对外的一实体端口;每一捆绑端口对应于一个或多个成员设备上对外的多个实体端口。

[0059] 可选地,

[0060] 所述流表生成模块通过流表控制所述成员设备之间及所述成员设备向外的流量转发,包括:所述流量转发为单播流量转发时,通过流表控制第一成员设备将流量从本设备直接向外转发,控制第二成员设备将流量从本设备向至少一个所述第一成员设备转发;其中,所述第一成员设备指具有与所述虚拟端口对应的实体端口的成员设备,所述第二成员设备指不具有与所述虚拟端口对应的实体端口的成员设备。

[0061] 可选地,

[0062] 所述流表生成模块根据路径决策的结果得到所述成员设备的流表,包括:所述流量转发为单播流量转发时,根据路径决策的结果生成所述虚拟转发设备的单播流表,所述单播流表包含作为出端口的所述虚拟端口;再针对每一所述成员设备,对所述单播流表进行修改,得到该成员设备的流表;

[0063] 其中,针对每一所述第一成员设备,所述修改包括:将出端口替换为该第一成员设备上与所述虚拟端口对应的实体端口;针对每一所述第二成员设备,所述修改包括:将出端口替换为该第二成员设备上用于第一迂回链路的实体端口,所述第一迂回链路是该第二成员设备到所述第一成员设备的链路。

[0064] 可选地,

[0065] 所述流表生成模块通过流表控制所述成员设备之间及所述成员设备向外的流量转发,包括:所述流量转发为广播或组播流量转发时,对每一个第一成员设备,通过流表控制该第一成员设备将流量从该第一成员设备直接向外转发,且对于来自所述转发设备集群外部的流量,还控制该第一成员设备将流量从该第一成员设备向满足以下条件的其他第一成员设备转发:实体端口对应的虚拟端口中至少有一个不同于该第一成员设备上实体端口对应的虚拟端口;及,对每一个第二成员设备,通过流表控制该第二成员设备将流量从该第二成员设备向所述转发设备集群中的第一成员设备转发,以将流量迂回到每一个所述虚拟端口再向外转发;其中,所述第一成员设备指具有与所述虚拟端口对应的实体端口的成员设备,所述第二成员设备指不具有与所述虚拟端口对应的实体端口的成员设备。

[0066] 可选地,

[0067] 所述流表生成模块根据路径决策的结果得到所述成员设备的流表,包括:所述流量转发为广播或组播流量转发时,根据路径决策的结果生成所述虚拟转发设备的广播或组播流表,所述广播或组播流表包含作为出端口的所述虚拟端口;针对每一所述第一成员设备,对所述广播或组播流表进行第一修改和第二修改,分别得到该第一成员设备的第一流表和第二流表;

[0068] 其中,所述第一修改包括:将表示报文来自集群内部的标志作为新增的匹配项,将出端口替换为该成员设备上与所述虚拟端口对应的实体端口;所述第二修改包括:将出端口替换为该成员设备上与所述虚拟端口对应的实体端口及用于第二迂回链路的实体端口,所述第二迂回链路是该第一成员设备到所述其他第一成员设备的链路。

[0069] 可选地，

[0070] 所述流表生成模块根据路径决策的结果得到所述成员设备的流表，包括：所述流量转发为广播或组播流量转发时，根据路径决策的结果生成所述虚拟转发设备的广播或组播流表，所述广播或组播流表包含作为出端口的所述虚拟端口；针对每一个所述第二成员设备，对所述广播或组播流表进行第三修改，得到该第二成员设备的流表；

[0071] 其中所述第三修改包括：将出端口替换为该第二成员设备上用于第三迂回链路的实体端口，及增加命令或动作以在报文中打上表示报文来自集群内部的标志；所述第三迂回链路是该第二成员设备到所述第一成员设备的链路。

[0072] 可选地，

[0073] 所述第三迂回链路包括除接收流量的链路之外的，该第二成员设备到满足以下条件一和条件二的第一成员设备的链路，以及按以下方式一选择的第一成员设备的链路：

[0074] 条件一，该成员设备上实体端口对应的虚拟端口包括非捆绑端口；

[0075] 条件二，该成员设备上实体端口对应的虚拟端口包括第一捆绑端口，所述第一捆绑端口指对应于一个成员设备上多个实体端口的捆绑端口；

[0076] 方式一，从实体端口对应的虚拟端口均为第二捆绑端口且对应的所述第二捆绑端口相同的多个第一成员设备中，选择一个第一成员设备，其中，所述第二捆绑端口指对应于多个成员设备上多个实体端口的捆绑端口；

[0077] 其中所述选择依据以下一种或多种策略进行：负载均衡策略；优先级策略；及随机选择策略。

[0078] 可选地，

[0079] 所述集群控制装置还包括以下模块中的一种或多种：

[0080] 离线处理模块，用于在确定所述转发设备集群中的一成员设备离线后，将该离线成员设备及其端口从所述转发设备集群存活的成员设备中删除；及对所述转发设备集群存活的每一成员设备，如果该成员设备的流表有出端口连接到该离线成员设备，将该成员设备的流表中连接到该离线成员设备的出端口删除，将更新后的该流表重新下发给该成员设备；

[0081] 端口失效处理模块，用于在确定作为出端口的一虚拟端口的状态变为非活动后，将该虚拟端口对应的实体端口从相应流表的出端口中删除；及，对删除后有出端口的每一流表，将更新后的该流表重新下发给对应的成员设备；对删除后无出端口的每一流表，将该流表对应的成员设备上连接到所述第一成员设备的实体端口新增为该流表的出端口，将更新后的该流表重新下发给对应的成员设备；

[0082] 通信故障处理模块，用于在确定一成员设备连接到一所述第一成员设备的实体端口或所述实体端口对应的链路失效后，将该成员设备的流表中作为出端口的所述实体端口更新为连接到另一所述第一成员设备的实体端口，将更新后的该流表重新下发给该成员设备；

[0083] 集群分裂处理模块，用于在确定所述转发设备集群处于分裂状态后，记录处于非活动状态的虚拟端口和所述成员设备之间的互联端口，以及失效的所述成员设备之间的链路，并禁用所述成员设备上的联动端口；及，对所述虚拟转发设备重新进行路径决策，根据路径决策的结果得到所述成员设备的流表并下发给所述成员设备；

[0084] 成员加入处理模块,所述 SDN 控制器确定有一成员设备加入所述转发设备集群后,根据所述路径决策的结果得到该成员设备的流表并下发给该成员设备,并在该成员设备具有与所述虚拟端口对应的实体端口时,按照所述转发设备集群新的拓扑对原有成员设备的流表进行更新并下发给相应的成员设备。

[0085] 下面再用几个应用中的示例进行说明。

[0086] 示例一

[0087] 本示例涉及集群的组建和配置。

[0088] SDN 控制器通过 OpenFlow 协议控制多台独立运行的转发设备,组成转发设备集群,对外呈现为一体化的逻辑设备。如图 3 所示,本示例中,集群由 2 台 OpenFlow 转发设备组成,标识为 DPID 1 和 DPID 2,2 台转发设备分别有编号为 1、2、3、4、5、6 的端口。

[0089] 转发设备集群的各个成员设备之间有互联的链路,即图 1 中 DPID 1 的端口 3 和 DPID 2 的端口 3 之间、DPID 1 的端口 4 和 DPID 2 的端口 4 之间的链路,对集群外不可见。成员设备之间的链路可由管理员配置,不限制是否是直连的,SDN 控制器通过链路探测或路径检测机制维护内部链路的状态。

[0090] 如图所示,转发设备集群呈现的虚拟转发设备对外暴露 6 个端口(图 1 中①、②、③、④、⑤、⑥),分为两类:

[0091] 非捆绑端口(①、②、③、④):一个非捆绑端口直接对应组成集群的一个成员设备上对外(即面向集群外)的一个实体端口,如 DPID1 上的端口 1、2,DPID2 上的端口 1、2。非捆绑端口通过非捆绑链路与邻居设备互联,端口状态直接对应实体端口的状态。

[0092] 捆绑端口(⑤、⑥):一个捆绑端口对应组成集群的一个或多个成员转发设备上对外的多个实体端口,如 DPID 1 上的端口 5 和 DPID 2 上的端口 5 捆绑组成端口⑤,DPID 1 上的端口 6 和 DPID 2 上的端口 6 捆绑组成端口⑥;当其中任一实体端口状态为活动时,捆绑端口状态即为活动,所有端口状态均为非活动时,捆绑端口状态为非活动。捆绑端口过捆绑链路与邻居互联。邻居设备相应的端口也需捆绑,可静态配置或通过聚合协议协商产生。

[0093] 示例二:

[0094] 本示例涉及对转发设备集群单播流量转发的控制。

[0095] 本示例 SDN 控制器通过 OpenFlow 协议下发转发表项,控制与转发设备集群相关单播流量转发,包括:

[0096] 步骤一,SDN 控制器将转发设备集群视为一个虚拟转发设备,根据网络拓扑情况进行路径决策,得到用于控制虚拟转发设备转发单播流量的单播流表;

[0097] 步骤二,SDN 控制器根据单播流表中作为出端口的虚拟端口对应的成员设备上的实体端口,对单播流表进行修改,得到各成员设备的流表并下发给各成员设备。

[0098] 具体地,对一成员设备,可通过流表实现以下流量转发的控制:

[0099] 如流量可经本设备直接转发出去,即本设备具有与所述虚拟端口对应的实体端口(对应于上文中的第一成员设备,本设备上的实体端口可以对应于非捆绑端口,也可以对应于捆绑端口,如果一实体端口是一捆绑端口对应的多个实体端口中的一个,即可认为该实体端口对应于该捆绑端口)时,通过流表控制该成员设备将流量从本设备直接转发出去。

[0100] 如流量需经集群其他成员设备迂回转发出去,即本设备不具有与所述虚拟端口对应的实体端口(对应于上文中的第二成员设备)时,通过流表控制该成员设备将流量先迂

回到其它成员设备再转发出去。

[0101] 对于从虚拟转发设备的单播流表如何得到各成员设备的流表,上文中已有详细描述,这里不再赘述。成员设备收到流表后,可以从流表中的出端口组中选择一个作为实际的转发端口。

[0102] 根据上述规则,按图 3 所示的集群拓扑进一步举例:

[0103] 例 1:虚拟出端口为①、⑤。

[0104] DPID 1 和 DPID 2 均可以直接转发流量,筛选以下端口作为向 DPID 1 和 DPID 2 下发的流表中的出端口:向 DPID 1 下发的流表的出端口为 1、5,向 DPID 2 下发的流表的出端口为 5。

[0105] 例 2:虚拟出端口为①、②。

[0106] DPID 1 可以直接转发流量,按例 1 中方法确定出端口,向 DPID 1 下发的转发表的出端口为 1、2。

[0107] DPID 2 需要将流量转发给 DPID 1 迂回,可以本设备前往有可用出端口的成员设备的迂回链路所对应端口作为出端口,向 DPID 2 下发的流表的出端口为 3、4。

[0108] 示例三

[0109] 本示例涉及对转发设备集群广播或组播流量转发的控制。

[0110] 本示例 SDN 控制器通过 OpenFlow 协议下发转发表项,控制与转发设备集群相关流量广播或组播转发,包括:

[0111] 步骤一,SDN 控制器将转发设备集群视为一个虚拟转发设备,根据网络拓扑情况进行路径决策,得到用于控制虚拟转发设备转发广播或组播流量的广播或组播流表;广播或组播时,需要从虚拟转发设备上的多个端口将报文转发出去。

[0112] 步骤二,SDN 控制器根据广播或组播流表中作为出端口的虚拟端口对应的成员设备上的实体端口,对广播或组播流表进行修改,得到各成员设备的流表并下发给各成员设备。

[0113] 具体地,对一成员设备,可通过流表实现以下流量转发的控制:

[0114] 如流量可经本设备直接转发出去(即本设备为第一成员设备)时,对于从集群成员设备收到的报文,通过流表控制该成员设备将流量从本设备直接转发出去;对于从集群以外设备收到的报文,除了通过流表控制该成员设备将流量从本设备直接转发出去外,还通过流表控制该成员设备从本设备向其他第一成员设备转发,此处的其他第一成员设备上实体端口对应的虚拟端口中至少有一个不同于本设备上实体端口对应的虚拟端口。也就是说,如果另一个第一成员设备具有与本设备完全相同的捆绑端口,则可以不再向该另一个第一成员设备转发。为了识别报文是从集群成员设备收到的,还是从集群以外设备收到的,可以在第二成员设备的流表中增加命令或动作,即在其转发的报文中打上表示报文来自集群内部的标志。同时给第一成员设备下发两个流表,在一个流表中增加该标记作为匹配项,如果匹配成功,即意味着是从成员设备收到的报文,没有匹配成功,再按另一个流表进行匹配处理。

[0115] 如流量需经集群其他成员设备迂回转发出去(即本设备为第二成员设备)时,通过流表控制该成员设备将流量从本设备向第一成员设备转发,以将流量迂回到每一个所述虚拟端口再向外转发。捆绑端口的成员端口分布在多个转发设备的,以及一对设备间存在

多条可用的内部链路的, 控制器需要统一选取将哪些端口加入出端口列表, 避免报文迂回后在多个成员端口输出。迂回时如何具体选择链路见上文的详细描述。并且, 对于需要迂回的流量, 在下发给第二成员设备的转发表中增加命令或动作以在转发的报文中打上表示报文来自集群内部的标志, 如通过添加 VLAN 标签、隧道标识等方式打上标志, 用于后续设备识别, 确定后续的转发动作。

[0116] 对于从虚拟转发设备的广播或组播流表如何得到各成员设备的流表, 上文中已有详细描述, 这里不再赘述。

[0117] 根据上述规则, 按图 3 中的所示的集群拓扑进一步举例:

[0118] 例: 虚拟出端口为①、⑤;

[0119] 向 DPID 1 下发 2 个流表(也可称为转发表项), 一个流表在原有匹配项的基础增加用于识别从 DPID 2 收到的报文的匹配项, 出端口为 1, 另一个流表的出端口为 1、5。

[0120] 在 DPID 2 上下发 1 个转发表项, 出端口为 5, 3。但对于从 DPID 1 发来的报文, 直接丢弃, 即不将接收报文的链路作为迂回链路。

[0121] 示例四

[0122] 本示例涉及集群成员设备发生故障离线时, SDN 控制器通过更新存活设备上的流表, 使流量恢复正常转发的方法, 包括:

[0123] 步骤一, SDN 控制器检测到转发设备集群中的某个成员设备已离线;

[0124] 步骤二, SDN 控制器维护转发设备集群的状态, 将离线设备及相关的端口从集群中剔除;

[0125] 步骤三, 集群其他存活成员设备收到的流量原来经该离线设备迂回转发的, 需要选取其他可用迂回路径。因此, SDN 控制器更新这些设备的流表, 即原有出端口连接所述离线设备的应予以删除, 仅保留前往其他可迂回设备的出端口。

[0126] 示例五

[0127] 本示例涉及集群成员设备上对外端口的状态变更为非活动时, SDN 控制器通过更新存活设备上的表项, 使流量恢复正常转发的方法, 包括:

[0128] 步骤一, 转发设备上报, 或者 SDN 控制器检测到集群成员上某一端口状态变更为非活动;

[0129] 步骤二, SDN 控制器将该端口从受影响的流表中删除, 即更新流表:

[0130] 对于该端口所在成员设备, 将该端口从该成员设备的流表的出端口中删除, 删除后无可用端口的成员设备, 流量需经其他设备迂回, 则增加本设备前往可迂回设备的端口作为出端口;

[0131] 对于其他成员设备, 原来迂回后经该端口转发的, 修改迂回路径, 将该端口从其流表的出端口中剔除, 保留前往其他可迂回设备的出端口。

[0132] 示例六

[0133] 本示例涉及集群内部链路发生中断时, SDN 控制器通过更新存活设备上的表项, 使流量恢复正常转发的方法, 包括:

[0134] 步骤一, 转发设备上报或者控制器检测到集群成员上连接其他成员的某端口变更为非活动, 或者控制器检测到集群内部路径失效;

[0135] 步骤二, 对于需要经过该端口或链路迂回转发的成员设备, 修改其流表中的出端

口,改从其他链路迂回。

[0136] 示例七

[0137] 本示例涉及集群内部链路故障和捆绑端口成员状态变更为非活动的叠加故障场景,SDN 控制器通过执行端口联动策略,禁用设备上部分端口,触发路径重计算并恢复正常转发,包括:

[0138] 步骤一,SDN 控制器判断集群已处于分裂状态;

[0139] 根据转发设备上报或者控制器检测到捆绑端口成员状态变更为非活动;或者捆绑端口成员已处于非活动状态,转发设备上报或者控制器检测到集群设备间互联端口状态变更为非活动等,SDN 控制器可以判断集群是否处于分裂状态,即变为互不相联的多个部分。

[0140] 步骤二,SDN 控制器根据联动策略禁用设备上的联动端口;联动端口如包括在某一端口失效后,也应关闭的端口。

[0141] 步骤三,SDN 控制器根据当前的集群拓扑重新计算路径,下发转发表,使流量经集群其他成员设备和捆绑端口其他成员端口转发。

[0142] 示例八:

[0143] 本示例涉及集群部分成员已处于正常工作状态,故障成员恢复或者新成员设备加入,SDN 控制器将设备加入集群,使流量可经该设备转发的方法,包括:

[0144] 步骤一,集群成员设备与控制器建立连接,完成初始化,SDN 控制器将设备加入集群,维护相关状态;

[0145] 步骤二,SDN 控制器为新加入设备生成下发的表项,并根据新的集群拓扑更新其他设备的流表;

[0146] 步骤三,下发流表,可以正常使用该新加入的设备。

[0147] 上述本发明实施例序号仅仅为了描述,不代表实施例的优劣。通过以上的实施方式描述,本领域的技术人员可以清楚地了解到上述实施例方法可借助软件加必需的通用硬件平台的方式来实现,当然也可以通过硬件,但很多情况下前者是更佳的实施方式。基于这样的理解,本发明实施例的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质(如 ROM/RAM、磁碟、光盘)中,包括若干指令用以使得一台终端设备(可以是手机,计算机,服务器,或者网络设备等)执行本发明各个实施例所述的控制器。

[0148] 以上仅为本发明的优选实施例,并非因此限制本发明的专利范围,凡是利用本发明说明书及附图内容所作的等效结构或等效流程变换,或直接或间接运用在其他相关的技术领域,均同理包括在本发明的专利保护范围内。

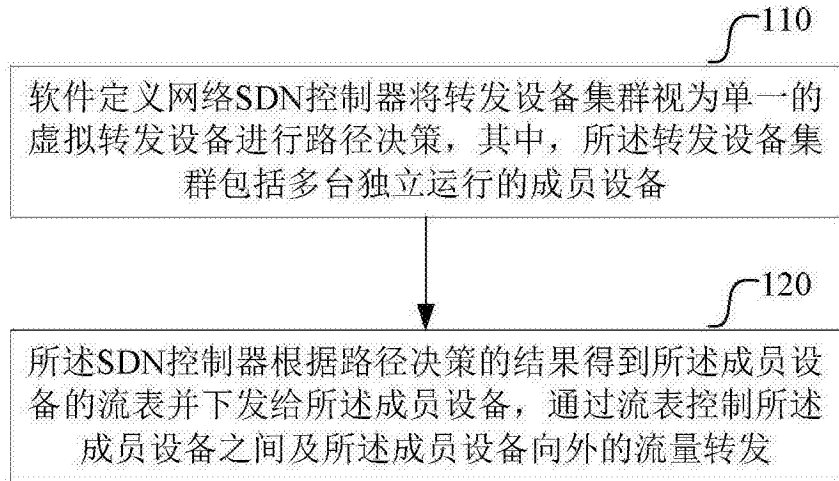


图 1

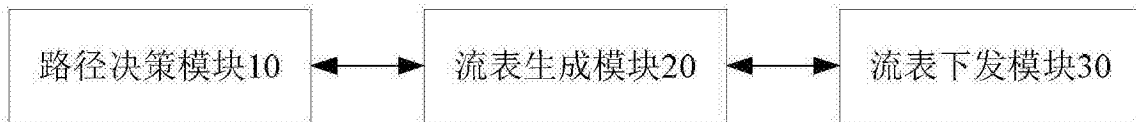


图 2

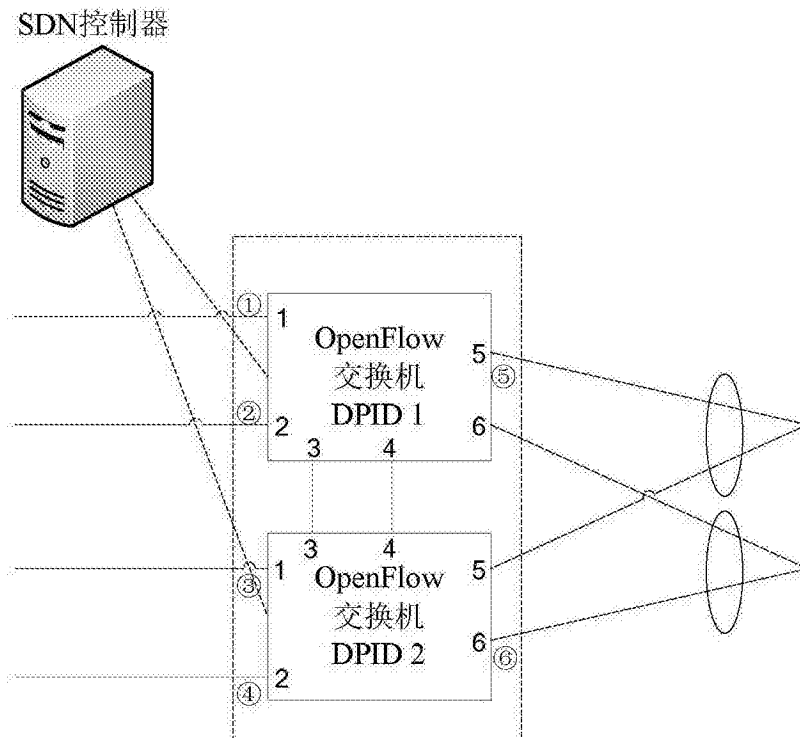


图 3