



(12)发明专利

(10)授权公告号 CN 104348727 B

(45)授权公告日 2018.05.15

(21)申请号 201310337052.3

审查员 于兰

(22)申请日 2013.08.05

(65)同一申请的已公布的文献号

申请公布号 CN 104348727 A

(43)申请公布日 2015.02.11

(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 李晶林

(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 郑红娟 宋志强

(51)Int.Cl.

H04L 12/741(2013.01)

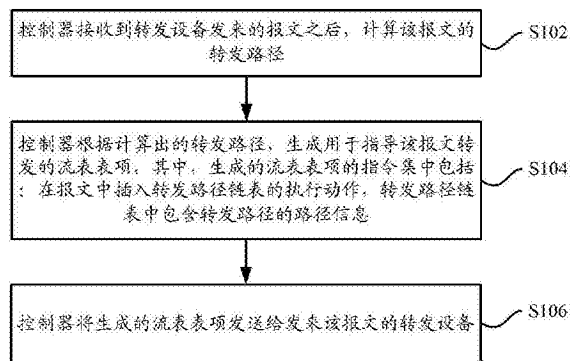
权利要求书3页 说明书11页 附图5页

(54)发明名称

OpenFlow网络中的流表表项处理方法及设备

(57)摘要

本申请公开了一种OpenFlow网络中的流表表项处理方法及设备,OpenFlow网络中包括:控制器和转发设备,该方法应用于控制器,该方法包括:接收到转发设备发来的报文之后,计算该报文的转发路径;根据计算出的转发路径,生成用于指导该报文转发的流表表项,其中,生成的流表表项的指令集中包括:在报文中插入转发路径链表的执行动作,转发路径链表中包含转发路径的路径信息;将流表表项发送给发来该报文的转发设备。本申请中,控制器只需向边缘转发设备下发流表表项,降低了控制器的流表表项下发负担,提高了业务部署效率。



1. 一种开放流OpenFlow网络中的流表表项处理方法,OpenFlow网络中包括:控制器和转发设备,所述方法应用于所述控制器,其特征在于,所述方法包括:

接收到转发设备发来的报文之后,计算该报文的转发路径;

根据计算出的转发路径,生成用于指导该报文转发的流表表项,其中,生成的流表表项的指令集中包括:在报文中插入转发路径链表的执行动作,所述转发路径链表中包含所述转发路径的路径信息;

将所述流表表项发送给发来该报文的转发设备,以使该转发设备在匹配该流表表项时根据指令集中的在报文中插入转发路径链表的执行动作在该报文中插入转发路径链表;

其中,所述转发路径的路径信息包括所述转发路径上的每一个转发设备的设备相关信息,所述设备相关信息包括该转发设备的设备标识以及该转发设备上用于转发该报文的转发接口的接口标识。

2. 根据权利要求1所述的方法,其特征在于,所述转发路径链表中还包含链表长度,所述链表长度的值置为所述转发路径上的转发设备的总数;并且,各个转发设备的设备相关信息按照在所述转发路径上的先后顺序排列。

3. 一种开放流OpenFlow网络中的流表表项处理方法,OpenFlow网络中包括:控制器和转发设备,所述方法应用于所述转发设备,其特征在于,所述方法包括:

接收到所述OpenFlow网络连接的外部网络发来的报文之后,根据该报文的报头中的信息在本地的流表中查找匹配表项;

若没有查找到匹配表项,则将该报文发送给控制器;

接收控制器发来的用于指导该报文转发的流表表项,将所述流表表项添加到所述本地的流表中,其中,所述流表表项的指令集中包括:在报文中插入转发路径链表的执行动作,所述转发路径链表中包含所述转发路径的路径信息;

根据该报文的报头中的信息在所述流表中查找到匹配表项,按照该匹配表项的指令集中的在报文中插入转发路径链表的执行动作,在该报文中插入转发路径链表,得到携带有转发路径链表的特定报文;

其中,所述转发路径的路径信息包括所述转发路径上的每一个转发设备的设备相关信息,所述设备相关信息包括该转发设备的设备标识以及该转发设备上用于转发该报文的转发接口的接口标识。

4. 根据权利要求3所述的方法,其特征在于,所述转发路径链表中还包含链表长度,所述链表长度的值置为所述转发路径上的转发设备的总数;并且,各个转发设备的设备相关信息按照在所述转发路径上的先后顺序排列。

5. 根据权利要求4所述的方法,其特征在于,在得到携带有转发路径链表的特定报文之后,还包括:

将特定报文的转发路径链表中的第一个设备相关信息弹出,并将该转发路径链表中的链表长度的值减1;

判断减1后的链表长度的值是否为0,若不为0,则将该特定报文通过弹出的设备相关信息中的接口标识所指示的接口转发出去。

6. 根据权利要求5所述的方法,其特征在于,还包括:

接收到邻居转发设备发来的特定报文之后,将该特定报文的转发路径链表中的第一个

设备相关信息弹出,并将该转发路径链表中的链表长度的值减1;

判断减1后的链表长度的值是否为0,若不为0,则将该特定报文通过弹出的设备相关信息中的接口标识所指示的接口转发出去。

7. 根据权利要求5或6所述的方法,其特征在于,在判断减1后的链表长度的值是否为0之后,还包括:

若为0,则将该特定报文中的转发路径链表弹出后,通过弹出的设备相关信息中的接口标识所指示的接口转发出去。

8. 一种开放流OpenFlow网络中的控制器,OpenFlow网络中包括:所述控制器和转发设备,其特征在于,所述控制器包括:

接收模块,用于接收转发设备发来的报文;

计算模块,用于在所述接收模块接收到转发设备发来的报文之后,计算该报文的转发路径;

表项生成模块,用于根据所述计算模块计算出的转发路径,生成用于指导该报文转发的流表表项,其中,生成的流表表项的指令集中包括:在报文中插入转发路径链表的执行动作,所述转发路径链表中包含所述转发路径的路径信息;

发送模块,用于将所述表项生成模块生成的流表表项发送给发来该报文的转发设备,以使该转发设备在匹配该流表表项时根据指令集中的在报文中插入转发路径链表的执行动作在该报文中插入转发路径链表;

其中,所述转发路径的路径信息包括所述转发路径上的每一个转发设备的设备相关信息,所述设备相关信息包括该转发设备的设备标识以及该转发设备上用于转发该报文的转发接口的接口标识。

9. 根据权利要求8所述的控制器,其特征在于,所述转发路径链表中还包含链表长度,所述链表长度的值置为所述转发路径上的转发设备的总数;并且,各个转发设备的设备相关信息按照在所述转发路径上的先后顺序排列。

10. 一种开放流OpenFlow网络中的转发设备,OpenFlow网络中包括:控制器和所述转发设备,其特征在于,所述转发设备包括:

接收模块,用于接收所述OpenFlow网络连接的外部网络发来的报文;还用于接收控制器发来的用于指导该报文转发的流表表项;

查找模块,用于在所述接收模块接收到外部网络发来的报文之后,根据该报文的报头中的信息在本设备的流表中查找匹配表项;

发送模块,用于若所述查找模块没有查找到匹配表项,则将该报文发送给控制器;

添加模块,用于将所述接收模块接收到的流表表项添加到所述本设备的流表中,其中,所述流表表项的指令集中包括:在报文中插入转发路径链表的执行动作,所述转发路径链表中包含所述转发路径的路径信息;

所述查找模块,还用于在所述添加模块将所述接收模块接收到的流表表项添加到所述流表中之后,根据该报文的报头中的信息在所述流表中查找到匹配表;

处理模块,用于按照所述查找模块查找到的匹配表项的指令集中的在报文中插入转发路径链表的执行动作,在该报文中插入转发路径链表,得到携带有转发路径链表的特定报文;

其中,所述转发路径的路径信息包括所述转发路径上的每一个转发设备的设备相关信息,所述设备相关信息包括该转发设备的设备标识以及该转发设备上用于转发该报文的转发接口的接口标识。

11. 根据权利要求10所述的转发设备,其特征在于,所述转发路径链表中还包含链表长度,所述链表长度的值置为所述转发路径上的转发设备的总数;并且,各个转发设备的设备相关信息按照在所述转发路径上的先后顺序排列。

12. 根据权利要求11所述的转发设备,其特征在于,还包括判断模块,其中,所述处理模块,还用于将特定报文的转发路径链表中的第一个设备相关信息弹出,并将该转发路径链表中的链表长度的值减1;

所述判断模块,用于判断经所述处理模块减1后的链表长度的值是否为0;

所述发送模块,还用于若所述判断模块的判断结果是不为0,则将该特定报文通过所述处理模块弹出的设备相关信息中的接口标识所指示的接口转发出去。

13. 根据权利要求12所述的转发设备,其特征在于,

所述接收模块,还用于接收邻居转发设备发来的特定报文;

所述处理模块,还用于在所述接收模块接收到邻居转发设备发来的特定报文之后,将该特定报文的转发路径链表中的第一个设备相关信息弹出,并将该转发路径链表中的链表长度的值减1。

14. 根据权利要求12或13所述的转发设备,其特征在于,

所述处理模块,还用于若所述判断模块的判断结果是为0,则将该特定报文中的转发路径链表弹出后,由所述发送模块通过所述处理模块弹出的设备相关信息中的接口标识所指示的接口转发出去。

OpenFlow网络中的流表表项处理方法及设备

技术领域

[0001] 本申请涉及网络通信技术领域,特别涉及一种OpenFlow网络中的流表表项处理方法及设备。

背景技术

[0002] 传统的IP(Internet Protocol,因特网协议)报文的路由转发方式是采用路由器或L3(Level3,3层)层交换机来完成的。以路由器为例,路由器的主要特点是单点行为,即,路由协议的交互和IP报文的转发动作,都是单点行为。路由器各自之间交互链路、端口状态、路由协议等信息,对于IP报文的转发判断和策略处理都是各自完成,相互之间可以不依赖于对方。转发过程只是根据报文头中的信息完成,如根据目的IP地址确定本机转发端口,根据IP TOS(Type Of Service,服务类型)标记确定转发优先级,等等。因此,传统路由器不适合构建端到端的业务转发路径。

[0003] SDN/OpenFlow(开放流)架构的提出,打破了路由器的传统转发方式与过程。SDN(Software Defined Network)即软件定义网络,从广义上说是指让网络专家用高级别的语言指明配置,而这些配置会告诉路由器和交换机如何优先处理数据流以及管理数据流。SDN分离了网络的转发平面和控制平面,把整个控制平面移到了单独的控制平面中。这个名为控制器的设备会为特定的工作负载计算最佳路径,为数据平面的数据转发进行规划。控制器可以是一个装置或虚拟机,也可以是一个物理服务器。

[0004] SDN的核心技术是OpenFlow,在网络数据转发平面和控制器之间使用OpenFlow协议,再由控制器来规划所有网络数据转发平面的转发动作。OpenFlow提出的最初出发点是为了解决校园内网络研究人员实验其创新网络架构、协议,考虑到实际的网络创新思想需要在实际网络上才能更好地验证,而研究人员又无法修改在网的网络设备的问题,而提出了OpenFlow的控制转发分离架构,将控制逻辑从网络设备盒子中引出来,研究者可以对其进行任意的编程从而实现新型的网络协议、拓扑架构而无需改动网络设备本身。

[0005] 如图1所示,OpenFlow网络的架构中有两种设备角色:控制器(Controller)和转发设备(Forwarder)。

[0006] 每一个转发设备维护一个或多个FlowTable(流表),并且只按照FlowTable对报文进行转发,Flowtable本身的生成、维护和下发完全由外置的控制器来执行。

[0007] 控制器上可以运行各种路由协议,根据需要将计算出来的流表表项下发给相应的转发设备。流表表项的下发可以是主动的,也可以是被动的,在主动模式中,控制器将自身收集的流表表项主动下发给转发设备,随后转发设备在接收到报文后可以直接根据流表对报文进行查表转发;在被动模式中,转发设备收到报文后,若在流表中没有查找到匹配表项,则将该报文转发给控制器,由控制器决策该报文的转发路径,并下发相应的流表表项。被动模式的好处是转发设备无需维护全部的流表表项,只有当实际的数据流产生时才向控制器获取流表表项并存储,当老化定时器超时可以删除相应的流表表项。

[0008] 流表中包括:Match Fields(匹配域)、计数器(Counter)和指令集(Instruction

Set)三个部分。Match Fields是报文匹配的输入关键字,Match Fields中包括:进入接口(In Port)、VLAN(Virtual Local Area Network,虚拟局域网)ID(标识)、Ethernet(以太网)源地址(SA)、Ethernet目标地址(DA)、Ethernet类型(Type)、IP源地址(SA)、IP目标地址(DA)、IP协议、TCP(Transmission Control Protocol,传输控制协议)/UDP(User Datagram Protocol,用户数据报协议)目标端口(Dst Port)、TCP/UDP源端口(Src Port)等。计数器是管理所需,用来统计流量的一些信息,例如活动表项、查找次数、发送包数等,可以针对每张表、每个流、每个端口、每个队列来维护。指令集中包括一系列用于决定报文如何转发的执行动作(Action),执行动作的类型包括:转发给某个端口、封装改写报文后转发、以及丢弃。

[0009] 在现有技术中,转发设备在接收到一个报文之后,根据该报文的报文头中的多个域的值在本地流表中查找匹配表项,如果查找到了匹配表项,则按照该匹配表项的指令集中的执行动作对该报文进行相应处理;否则,将该报文转发给控制器,控制器收到该报文后,计算该报文的转发路径以及该转发路径上每一个转发设备的转发端口等,生成流表表项,并将流表表项下发给该转发路径上的对应转发设备,其中,下发给该转发路径上的一个转发设备的流表表项的指令集中的执行动作是通过该转发设备上的某一个转发接口转发。这样,每增加一种新业务数据流,控制器就需要向相关转发路径上的所有转发设备下发对应的流表表项,增加了控制器的负担,降低了业务部署的效率。

发明内容

[0010] 本申请提供了一种OpenFlow网络中的流表表项处理方法及设备,以解决现有技术中存在的控制器的流表表项下发负担较大,业务部署效率较低的问题。

[0011] 本申请的技术方案如下:

[0012] 一方面,提供了一种OpenFlow网络中的流表表项处理方法,OpenFlow网络中包括:控制器和转发设备,该方法应用于控制器,该方法包括:

[0013] 接收到转发设备发来的报文之后,计算该报文的转发路径;

[0014] 根据计算出的转发路径,生成用于指导该报文转发的流表表项,其中,生成的流表表项的指令集中包括:在报文中插入转发路径链表的执行动作,转发路径链表中包含转发路径的路径信息;

[0015] 将流表表项发送给发来该报文的转发设备。

[0016] 另一方面,还提供了一种OpenFlow网络中的流表表项处理方法,OpenFlow网络中包括:控制器和转发设备,该方法应用于转发设备,该方法包括:

[0017] 接收到OpenFlow网络连接的外部网络发来的报文之后,根据该报文的报头中的信息在本地的流表中查找匹配表项;

[0018] 若没有查找到匹配表项,则将该报文发送给控制器;

[0019] 接收控制器发来的用于指导该报文转发的流表表项,将流表表项添加到该本地的流表中,其中,流表表项的指令集中包括:在报文中插入转发路径链表的执行动作,转发路径链表中包含转发路径的路径信息。

[0020] 又一方面,还提供了一种OpenFlow网络中的控制器,OpenFlow网络中包括:控制器和转发设备,该控制器包括:

[0021] 接收模块,用于接收转发设备发来的报文;

- [0022] 计算模块,用于在接收模块接收到转发设备发来的报文之后,计算该报文的转发路径;
- [0023] 表项生成模块,用于根据计算模块计算出的转发路径,生成用于指导该报文转发的流表表项,其中,生成的流表表项的指令集中包括:在报文中插入转发路径链表的执行动作,转发路径链表中包含转发路径的路径信息;
- [0024] 发送模块,用于将表项生成模块生成的流表表项发送给发来该报文的转发设备。
- [0025] 又一方面,还提供了一种OpenFlow网络中的转发设备,OpenFlow网络中包括:控制器和转发设备,该转发设备包括:
- [0026] 接收模块,用于接收OpenFlow网络连接的外部网络发来的报文;还用于接收控制器发来的用于指导该报文转发的流表表项;
- [0027] 查找模块,用于在接收模块接收到外部网络发来的报文之后,根据该报文的报头中的信息在本设备的流表中查找匹配表项;
- [0028] 发送模块,用于若查找模块没有查找到匹配表项,则将该报文发送给控制器;
- [0029] 添加模块,用于将接收模块接收到的流表表项添加到该本设备的流表中,其中,流表表项的指令集中包括:在报文中插入转发路径链表的执行动作,转发路径链表中包含转发路径的路径信息。
- [0030] 通过本申请的技术方案,控制器接收到某一个转发设备(为了描述方便,称为边缘转发设备)发来的报文之后,计算该报文的转发路径,根据计算出的转发路径生成用于指导该报文转发的流表表项,生成的流表表项的指令集中包括:在报文中插入转发路径链表的执行动作,转发路径链表中包含有该转发路径的路径信息,然后,控制器仅将该流表表项发送给边缘转发设备,而不需要再为该转发路径上的其他转发设备(称为中间转发设备)生成并下发流表表项。这样,边缘转发设备后续就可以按照该流表表项对该报文及其同类报文进行转发处理,即,在报文中插入包含有报文的转发路径的路径信息的转发路径链表并转发出去,这样,该转发路径上后续的中间转发设备就可以按照报文中的转发路径的路径信息对报文进行转发,从而,中间转发设备不再需要按照流表表项对报文进行转发处理,即不再需要控制器为其生成并下发相应地流表表项。由上,控制器只需向边缘转发设备下发流表表项,降低了控制器的流表表项下发负担,提高了业务部署效率。

附图说明

- [0031] 图1是OpenFlow网络的架构示意图;
- [0032] 图2是本申请实施例一的OpenFlow网络中的流表表项处理方法的流程图;
- [0033] 图3是本申请实施例一的转发路径链表的一种形式示意图;
- [0034] 图4是本申请实施例一的转发路径链表的另一种简化形式示意图;
- [0035] 图5是图3和图4中的转发链表指示的格式示意图;
- [0036] 图6是本申请实施例二的原始报文和插入了转发路径链表的特定报文的报文格式示意图;
- [0037] 图7是本申请实施例二的转发路径链表的形式示意图;
- [0038] 图8是弹出图7中的第一个设备相关信息后的转发路径链表的形式示意图;
- [0039] 图9是本申请实施例三的转发路径链表的形式示意图;

- [0040] 图10是本申请实施例四的OpenFlow网络中的控制器的结构示意图；
- [0041] 图11是本申请实施例四的OpenFlow网络中的转发设备的结构示意图；
- [0042] 图12是本申请实施例五的OpenFlow网络中的控制器的硬件结构示意图；
- [0043] 图13是本申请实施例五的OpenFlow网络中的转发设备的硬件结构示意图。

具体实施方式

[0044] 为了解决现有技术中存在的控制器的流表表项下发负担较大,业务部署效率较低的问题,本申请的以下实施例中提供了一种OpenFlow网络中的流表表项处理方法,以及一种控制器和转发设备。

[0045] 如图1所示,本申请以下实施例的OpenFlow网络中包括:控制器和转发设备。在本申请实施例中,控制器接收到某一个转发设备(为了描述方便,称为边缘转发设备)发来的报文之后,计算该报文的转发路径,根据计算出的转发路径生成用于指导该报文转发的流表表项,生成的流表表项的指令集中包括:在报文中插入转发路径链表的执行动作,转发路径链表中包含有该转发路径的路径信息,然后,控制器仅将该流表表项发送给边缘转发设备,而不需要再为该转发路径上的其他转发设备(称为中间转发设备)生成并下发流表表项。这样,边缘转发设备后续就可以按照该流表表项对该报文及其同类报文进行转发处理,即,在报文中插入包含有报文的转发路径的路径信息的转发路径链表并转发出去,这样,该转发路径上后续的中间转发设备就可以按照报文中的转发路径的路径信息对报文进行转发,从而,中间转发设备不再需要按照流表表项对报文进行转发处理,即不再需要控制器为其生成并下发相应地流表表项。由上,控制器只需向边缘转发设备下发流表表项,降低了控制器的流表表项下发负担,提高了业务部署效率。

[0046] 为简化描述,以下实施例中,针对一种业务数据流,将该业务数据流所经过的转发设备的角色做如下定义:

[0047] 1、边缘转发设备:接收到该业务数据流的第一个转发设备,也即,该业务数据流进入OpenFlow网络的第一个转发设备;

[0048] 2、中间转发设备:该业务数据流所经过的除边缘转发设备以外的其他转发设备。

[0049] 实施例一

[0050] 本申请实施例一的OpenFlow网络中的流表表项处理方法由控制器来执行。如图2所示,该方法包括以下步骤:

[0051] 步骤S102,控制器接收到转发设备发来的报文之后,计算该报文的转发路径;

[0052] 控制器收集所有转发设备的链路状态及拓扑信息,为需要转发的数据流计算路由(转发路径),同时对各个转发设备进行管理和监控。

[0053] 步骤S104,控制器根据计算出的转发路径,生成用于指导该报文转发的流表表项,其中,生成的流表表项的指令集中包括:在报文中插入转发路径链表的执行动作,转发路径链表中包含转发路径的路径信息;

[0054] 在实际实施过程中,转发路径链表中除了包含有转发路径的路径信息以外,还可以包含有链表长度,链表长度的值置为转发路径上的转发设备的总数;转发路径的路径信息包括:转发路径上的每一个转发设备的设备相关信息,并且,各个转发设备的设备相关信息按照在转发路径上的先后顺序排列,其中,设备相关信息可以用转发设备+转发接口来表

示,即,一个转发设备的设备相关信息为该转发设备的设备标识以及该转发设备上用于转发该报文的转发接口的接口标识,或者,设备相关信息也可以用采用一种简化形式,只用转发接口来表示,即,一个转发设备的设备相关信息为该转发设备上用于转发该报文的转发接口的接口标识。

[0055] 这样,当设备相关信息用转发设备+转发接口来表示时,转发路径链表的形式可以参见图3所示,图3中,所经第一个转发设备即为业务数据流的边缘转发设备,所经第二个转发设备至所经最后一个转发设备即为业务数据流的中间转发设备,转发链表指示中携带有链表长度;当设备相关信息只用转发接口来表示时,转发路径链表的形式可以参见图4所示,图4中,转发链表指示中携带有链表长度。在实际实施过程中,还可以在图3和图4中的转发链表指示字段中定义数据流的优先级等,此时,转发链表指示字段的格式可以参见图5。

[0056] 步骤S106,控制器将生成的流表表项发送给发来该报文的转发设备。

[0057] 上述的转发设备即为针对该报文所属的业务数据流的边缘转发设备。

[0058] 控制器只向一种业务数据流的边缘转发设备下发流表表项,对于中间转发设备,不再下发流表表项,也即中间转发设备中没有用于指导该业务数据流转发的流表表项。

[0059] 实施例二

[0060] 本申请实施例二的OpenFlow网络中的流表表项处理方法由转发设备来执行。当一个转发设备为一种业务数据流的边缘转发设备时,该方法包括以下步骤:

[0061] 步骤S202,转发设备接收到OpenFlow网络连接的外部网络发来的报文之后,根据该报文的报头中的信息在本地的流表中查找匹配表项,若没有查找到匹配表项,则执行步骤S204,若查找到了匹配表项,则执行步骤S210;

[0062] 在实际实施过程中,如果接收的报文中不携带转发路径链表,则说明是外部网络发来的报文。

[0063] 步骤S204,转发设备将该报文发送给控制器;然后,执行步骤S206;

[0064] 在步骤S202-S204中,转发设备接收到一个全新的业务数据流的报文,将首包(第一个报文)转发给控制器进行路径计算。在步骤S204中,转发设备将该报文发送给控制器之后,控制器执行的操作可以参见实施例一的步骤S102-S106,这里不再赘述。

[0065] 步骤S206,转发设备接收控制器发来的用于指导该报文转发的流表表项,将流表表项添加到该本地的流表中,其中,该流表表项的指令集中包括:在报文中插入转发路径链表的执行动作,转发路径链表中包含转发路径的路径信息;然后,执行步骤S208;

[0066] 转发路径链表的形式可以参见图3-5,这里不再赘述。

[0067] 步骤S208,根据该报文的报头中的信息在流表中查找到匹配表项;然后,执行步骤S210;

[0068] 步骤S210,按照该匹配表项的指令集中的在报文中插入转发路径链表的执行动作,在该报文中插入转发路径链表,得到携带有转发路径链表的特定报文;然后,执行步骤S212;

[0069] 在实际实施过程中,转发设备接收到的外部网络发来的报文(称为原始报文)的格式如图6所示,在该报文中插入如图3或图4所示的转发路径链表时,可以将转发路径链表插入到如图6所示的原始报文的链路层之后以及IP头之前的位置,插入了转发路径链表后得到的特定报文的格式如图6所示。

[0070] 步骤S212,将特定报文的转发路径链表中的第一个设备相关信息弹出,并将该转发路径链表中的链表长度的值减1;然后,执行步骤S214;

[0071] 由于在转发路径链表中,转发路径上的各个转发设备的相关信息都是按照在转发路径上的先后顺序排列的,因此,特定报文的转发路径链表中的第一个设备相关信息必定是本转发设备的设备相关信息。

[0072] 假设,在步骤S210中得到的特定报文中的转发路径链表的形式如图7所示,则经过步骤S212之后的特定报文中的转发路径链表的形式如图8所示。

[0073] 步骤S214,转发设备判断减1后的链表长度的值是否为0,若不为0,则执行步骤S216,若为0,则执行步骤S218;

[0074] 步骤S216,转发设备将经过步骤S210之后的特定报文通过弹出的设备相关信息中的接口标识所指示的接口转发出去;

[0075] 步骤S218,转发设备将经过步骤S302之后的特定报文中的转发路径链表弹出后,通过弹出的设备相关信息中的接口标识所指示的接口转发出去。

[0076] 在步骤S218中,若减1后的链表长度的值为0,则说明本转发设备是该报文所经的最后一个转发设备,则将特定报文中的转发路径链表整个弹出,得到原始报文,将原始报文转发出去。

[0077] 另外,当一个转发设备作为一种业务数据流的中间转发设备时,该转发设备还会接收到邻居转发设备(可以是该业务数据流的边缘转发设备,也可以是该业务数据流的中间转发设备)发来的特定报文,此时,该转发设备需要执行的操作包括以下步骤:

[0078] 步骤S302,转发设备接收到邻居转发设备发来的特定报文之后,将该特定报文的转发路径链表中的第一个设备相关信息弹出,并将该转发路径链表中的链表长度的值减1;

[0079] 步骤S304,转发设备判断减1后的链表长度的值是否为0,若不为0,则执行步骤S306,若为0,则执行步骤S308;

[0080] 步骤S306,转发设备将经过步骤S302之后的特定报文通过弹出的设备相关信息中的接口标识所指示的接口转发出去;

[0081] 步骤S308,转发设备将经过步骤S302之后的特定报文中的转发路径链表弹出后,通过弹出的设备相关信息中的接口标识所指示的接口转发出去。

[0082] 由此可见,只有业务数据流的边缘转发设备需要接收并保存控制器下发的包含有该业务数据流的转发路径的路径信息的流表表项,后续的中间转发设备不再需要接收并保存流表表项,进而不再有表项的插入、老化和刷新等事情,也不再需要在本地进行长流表匹配,从而极大地简化了中间转发设备的软硬件设计和开发设计难度,不再有表项规模限制和TCAM(Ternary Content Addressable Memory,三态内容寻址存储器)成本问题,提高了整个OpenFlow网络的业务容量。

[0083] 另外,中间转发设备不需要再根据报文头中的IP地址等信息对报文进行查表转发,因此,IP地址可以重叠部署,易于实现各种VPN(Virtual Private Network,虚拟专用网络)模型。

[0084] 在本实施例中,因为在原始报文中增加了新的字段:转发路径链表,导致报文长度增加,可能会超出链路层能够传输报文的最大长度,因此,边缘转发设备可能需要先对原始报文进行报文分片等动作,然后再将转发路径链表插入到各个分片中。

[0085] 实施例三

[0086] 以如图1所示的OpenFlow网络为例,对上述实施例一和实施例二中的方法进行详细说明。图1中的OpenFlow网络中包括:控制器和转发设备F1-F4。

[0087] 转发设备F1从外部网络接收到报文之后,根据报头中的信息在流表中查找匹配表项,若没有查找到匹配表项,则将该报文发送给控制器,由控制器决策该报文的转发路径,假设转发路径为F1→F2→F4,控制器根据该转发路径生成用于指导该报文转发的流表表项,并只将该流表表项发送给F1,该流表表项的形式例如可以参见表1所示。

[0088] 表1

[0089]

匹配域	计数器	指令集
目的 IP=XX, 源 IP=XX		在报文中插入如图9所示的转发路径链表

[0090] 转发设备F1收到如表1所示的流表表项之后,将该流表表项添加到流表中。再次根据该报文的报头中的信息在流表中查找匹配表项,此次就会查找到如表1中所示的流表表项,然后,按照该流表表项的指令集在该报文中插入如图9所示的转发路径链表,得到特定报文发送给转发模块;转发模块弹出该特定报文的如图9所示的转发路径链表中的第一个设备相关信息,即F1和PORT1,并将链表长度减1,判断此时的转发路径链表的长度是否等于0,判断结果是不等于0,则将该特定报文通过弹出的设备相关信息中的PORT1转发出去,从而发送给了转发设备F2。

[0091] 转发设备F2收到该特定报文之后,转发模块同样弹出该特定报文的转发路径链表中的第一个设备相关信息,即F2和PORT2,并将链表长度减1,判断此时的转发路径链表的长度是否等于0,判断结果是不等于0,则将该特定报文通过弹出的设备相关信息中的PORT2转发出去,从而发送给了转发设备F4。

[0092] 转发设备F4收到该特定报文之后,转发模块同样弹出该特定报文的转发路径链表中的第一个设备相关信息,即F4和PORT3,并将链表长度减1,判断此时的转发路径链表的长度是否等于0,判断结果是等于0,则将该特定报文中的转发路径链表整个弹出,得到原始报文,然后,通过弹出的设备相关信息中的PORT3转发出去,从而该原始报文能够到达目的设备。

[0093] 实施例四

[0094] 针对上述实施例一中的方法,本实施例四中提供了一种OpenFlow网络中的控制器,如图10所示,该控制器中包括以下模块:接收模块401、计算模块402、表项生成模块403和发送模块404,其中:

[0095] 接收模块401,用于接收转发设备发来的报文;

[0096] 计算模块402,用于在接收模块401接收到转发设备发来的报文之后,计算该报文的转发路径;

[0097] 表项生成模块403,用于根据计算模块402计算出的转发路径,生成用于指导该报文转发的流表表项,其中,生成的流表表项的指令集中包括:在报文中插入转发路径链表的执行动作,转发路径链表中包含转发路径的路径信息;

[0098] 发送模块404,用于将表项生成模块402生成的流表表项发送给发来该报文的转发设备。

[0099] 其中,转发路径链表中还包含链表长度,链表长度置为转发路径上的转发设备的总数;转发路径的路径信息包括:转发路径上的每一个转发设备的设备相关信息,并且,各个转发设备的设备相关信息按照在转发路径上的先后顺序排列,其中,设备相关信息包括:该转发设备的设备标识以及该转发设备上用于转发该报文的转发接口的接口标识,或该转发设备上用于转发该报文的转发接口的接口标识。

[0100] 在实际实施过程中,控制器中还可以包括:存储模块,用于存储表项生成模块403生成的流表表项,该流表表项中包含的转发路径链表,以及来自转发设备的报文。

[0101] 针对上述实施例二中的方法,本实施例四中还提供了一种OpenFlow网络中的转发设备,如图11所示,该转发设备中包括:接收模块501、查找模块502、发送模块503和添加模块504,其中:

[0102] 接收模块501,用于接收OpenFlow网络连接的外部网络发来的报文;还用于接收控制器发来的用于指导该报文转发的流表表项;

[0103] 查找模块502,用于在接收模块501接收到外部网络发来的报文之后,根据该报文的报头中的信息在本设备的流表中查找匹配表项;

[0104] 发送模块503,用于若查找模块502没有查找到匹配表项,则将该报文发送给控制器;

[0105] 添加模块504,用于将接收模块501接收到的流表表项添加到该本设备的流表中,其中,流表表项的指令集中包括:在报文中插入转发路径链表的执行动作,转发路径链表中包含转发路径的路径信息。

[0106] 其中,转发路径链表中还包含链表长度,链表长度的值置为转发路径上的转发设备的总数;转发路径的路径信息包括:转发路径上的每一个转发设备的设备相关信息,并且,各个转发设备的设备相关信息按照在转发路径上的先后顺序排列,其中,设备相关信息包括:该转发设备的设备标识以及该转发设备上用于转发该报文的转发接口的接口标识,或该转发设备上用于转发该报文的转发接口的接口标识。

[0107] 其中,该转发设备中还包括:处理模块和判断模块,其中,

[0108] 查找模块,还用于在添加模块将接收模块接收到的流表表项添加到流表中之后,根据该报文的报头中的信息在流表中查找到匹配表项;

[0109] 处理模块,用于按照查找模块查找到的匹配表项的指令集中的在报文中插入转发路径链表的执行动作,在该报文中插入转发路径链表,得到携带有转发路径链表的特定报文;还用于将特定报文的转发路径链表中的第一个设备相关信息弹出,并将该转发路径链表中的链表长度的值减1;

[0110] 判断模块,用于判断经处理模块减1后的链表长度的值是否为0;

[0111] 发送模块,还用于若判断模块的判断结果是不为0,则将该特定报文通过处理模块弹出的设备相关信息中的接口标识所指示的接口转发出去。

[0112] 另外,接收模块,还用于接收邻居转发设备发来的特定报文;处理模块,还用于在接收模块接收到邻居转发设备发来的特定报文之后,将该特定报文的转发路径链表中的第一个设备相关信息弹出,并将该转发路径链表中的链表长度的值减1。

[0113] 处理模块,还用于若判断模块的判断结果是为0,则将该特定报文中的转发路径链表弹出后,由发送模块通过处理模块弹出的设备相关信息中的接口标识所指示的接口转发出去。

[0114] 在实际实施过程中,转发设备中还可以包括:存储模块,用于存储本设备的流表,来自外部网络的报文,以及特定报文,其中,本设备的流表中记录有多条由控制器发来的流表表项。

[0115] 本申请实施例四中,控制器或转发设备中的模块或单元可以集成于一体,也可以分离部署;可以合并为一个模块或单元,也可以进一步拆分成多个子模块或子单元。本申请各实施例中的模块或单元可以以机械方式或电子方式实现。例如,一个硬件模块可以包括专门设计的永久性电路或逻辑器件(如专用处理器,如FPGA或ASIC)用于完成特定的操作。硬件模块也可以包括由软件临时配置的可编程逻辑器件或电路(如包括通用处理器或其它可编程处理器)用于执行特定操作。至于具体采用机械方式,或是采用专用的永久性电路,或是采用临时配置的电路(如由软件进行配置)来实现硬件模块,可以根据成本和时间上的考虑来决定。

[0116] 实施例五

[0117] 以上实施例四对本申请具体实施例中的控制器和转发设备进行了说明,本实施例五给出了本申请具体实施例中控制器和转发设备的硬件架构组成。

[0118] 控制器是可以软硬件结合的可编程设备,具体参见图12,图12是本申请具体实施例中控制器的硬件架构组成示意图,该控制器包括:非易失性存储器,CPU(中央处理器),转发芯片,内存和其他硬件;其中,

[0119] 内存,用于存储指令代码,当所述指令代码被执行时完成的操作主要为控制器中的接收模块、查找模块、发送模块和添加模块等模块完成的功能,这里不再详述。

[0120] CPU,用于与转发芯片通信,进行各种报文的收发,包括:接收来自转发设备的报文、向转发设备发送生成的流表表项;用于与内存通信,读取和执行内存中存储的所述指令代码,完成上述控制器中的接收模块、查找模块、发送模块和添加模块等模块完成的功能,以及对从转发芯片上送的报文的处理;用于与非易失性存储器通信,读/写非易失性存储器中的数据,包括:流表表项。

[0121] 其中,接收模块用于接收转发芯片上送的来自转发设备的报文,并写入存储模块进行保存。计算模块用于在接收模块接收到转发芯片上送的来自转发设备的报文之后,计算该报文的转发路径并输出给表项生成模块。表项生成模块用于根据计算模块输入的该报文的转发路径,生成用于指导该报文转发的流表表项,并写入存储模块中进行保存,其中,生成的流表表项的指令集中包括:在报文中插入转发路径链表的执行动作,转发路径链表中包含转发路径的路径信息。发送模块用于将表项生成模块生成的流表表项发送给转发芯片。

[0122] 转发芯片,用于通过该芯片上的端口连接到转发设备,负责上述的各种报文的收发处理,具体的,将CPU发来的流表表项发送到转发设备,并将接收到的来自转发设备的报文上送给CPU。

[0123] 非易失性存储器,用于存储各种数据,包括:来自转发设备的报文、CPU生成的流表表项以及该流表表项中包含的转发路径链表,完成控制器中的存储模块的功能,其中,存储

模块用于存储转发芯片上送的来自转发设备的报文、CPU生成的流表表项以及该流表表项中包含的转发路径链表。

[0124] 同样,转发设备也是可以软硬件结合的可编程设备,具体参见图13,图13是本申请具体实施例中转发设备的硬件架构组成示意图,该转发设备中包括:非易失性存储器,CPU,转发芯片,内存和其他硬件;其中,

[0125] 内存,用于存储指令代码,当所述指令代码被执行时完成的操作主要为转发设备中的接收模块、查找模块、发送模块、添加模块、处理模块和判断模块等模块完成的功能,这里不再详述。

[0126] CPU,用于与转发芯片通信,进行各种报文的收发,包括:接收来自OpenFlow网络连接的外部网络的报文、接收来自控制器的流表表项、向控制器发送报文;用于与内存通信,读取和执行内存中存储的所述指令代码,完成上述转发设备中的接收模块、查找模块、发送模块、添加模块、处理模块和判断模块等模块完成的功能,以及对从转发芯片上送的报文的处理;用于与非易失性存储器通信,读/写非易失性存储器中的数据,包括:流表中的流表表项。

[0127] 其中,接收模块用于接收转发芯片上送的来自外部网络的报文,并将该报文写入到存储模块中进行保存;还用于接收转发芯片上送的来自控制器的用于指导该报文转发的流表表项;还用于接收邻居转发设备发来的特定报文。

[0128] 查找模块用于在接收模块接收到来自外部网络的报文之后,根据该报文的报头中的信息在存储模块中的流表中查找匹配表项;还用于在添加模块将流表表项添加到存储模块中的流表中之后,根据该报文的报头中的信息在流表中查找到匹配表项。

[0129] 发送模块用于在查找模块没有查找到匹配表项时,将该报文发送给转发芯片,由转发芯片发送给控制器;还用于若判断模块的判断结果是不为0,则将该特定报文发送给转发芯片,由转发芯片通过处理模块弹出的设备相关信息中的接口标识所指示的接口转发出去;还用于将经处理模块弹出转发路径链表后的报文,发送给转发芯片,由转发芯片通过处理模块弹出的设备相关信息中的接口标识所指示的接口转发出去。

[0130] 添加模块用于将接收模块接收到的流表表项添加到存储模块中的流表中,其中,该流表表项的指令集中包括:在报文中插入转发路径链表的执行动作,转发路径链表中包含转发路径的路径信息。

[0131] 处理模块,用于按照查找模块查找到的匹配表项的指令集中的在报文中插入转发路径链表的执行动作,在该报文中插入转发路径链表,得到携带有转发路径链表的特定报文;还用于将特定报文的转发路径链表中的第一个设备相关信息弹出,并将该转发路径链表中的链表长度的值减1;还用于若判断模块的判断结果是为0,则将该特定报文中的转发路径链表弹出后,发送给转发模块;还用于在接收模块接收到邻居转发设备发来的特定报文之后,将该特定报文的转发路径链表中的第一个设备相关信息弹出,并将该转发路径链表中的链表长度的值减1。

[0132] 判断模块,用于判断经处理模块减1后的链表长度的值是否为0。

[0133] 转发芯片,用于通过该芯片上的端口连接到控制器和外部网络,负责上述的各种报文的收发处理,具体的,将CPU发来的来自外部网络的报文发送到控制器,将接收到的来自外部网络的报文上送给CPU,将接收到的来自控制器的流表表项上送给CPU,将接收到的

来自邻居转发设备的特定报文上送给CPU,将CPU发来的特定报文发送给邻居转发设备,将CPU发来的报文发送给外部网络。

[0134] 非易失性存储器,用于存储各种数据,包括:流表中的流表表项以及流表表项中包含的转发路径链表,特定报文,来自外部网络的报文,完成转发设备中的存储模块的功能。其中,存储模块用于存储本设备的流表,来自外部网络的报文,以及特定报文。

[0135] 需要说明的是,图12所示的控制器和转发设备只是一个具体的例子,也可以通过其他的与本实施例描述不同结构实现,例如,执行上述指令代码时所完成的操作,也可以由特定应用专用集成电路(ASIC)实现。另外,上述的CPU可以是一个或多个,如果是多个,则由多个CPU共同负责读取和执行所述指令代码。因此,本申请对控制器和转发设备的具体结构不作具体限定。

[0136] 综上,本申请以上实施例可以达到以下技术效果:

[0137] (1)控制器接收到某一个转发设备(为了描述方便,称为边缘转发设备)发来的报文之后,计算该报文的转发路径,根据计算出的转发路径生成用于指导该报文转发的流表表项,生成的流表表项的指令集中包括:在报文中插入转发路径链表的执行动作,转发路径链表中包含有该转发路径的路径信息,然后,控制器仅将该流表表项发送给边缘转发设备,而不需要再为该转发路径上的其他转发设备(称为中间转发设备)生成并下发流表表项。这样,边缘转发设备后续就可以按照该流表表项对该报文及其同类报文进行转发处理,即在报文中插入包含有报文的转发路径的路径信息的转发路径链表并转发出去,这样,该转发路径上后续的中间转发设备就可以按照报文中的转发路径的路径信息对报文进行转发,从而,中间转发设备不再需要按照流表表项对报文进行转发处理,即不再需要控制器为其生成并下发相应地流表表项。由上,控制器只需向边缘转发设备下发流表表项,降低了控制器的流表表项下发负担,提高了业务部署效率。

[0138] (2)只有业务数据流的边缘转发设备需要接收并保存控制器下发的包含有该业务数据流的转发路径的路径信息的流表表项,后续的中间转发设备不再需要接收并保存流表表项,进而不再有表项的插入、老化和刷新等事情,也不再需要在本地进行长流表匹配,从而极大地简化了中间转发设备的软硬件设计和开发设计难度,不再有表项规模限制和TCAM(Ternary Content Addressable Memory,三态内容寻址存储器)成本问题,提高了整个OpenFlow网络的业务容量。

[0139] (3)中间转发设备不需要再根据报文头中的IP地址等信息对报文进行查表转发,因此,IP地址可以重叠部署,易于实现各种VPN(Virtual Private Network,虚拟专用网络)模型。

[0140] 以上所述仅为本申请的较佳实施例而已,并不用以限制本申请,凡在本申请的精神和原则之内,所做的任何修改、等同替换、改进等,均应包含在本申请保护的范围之内。

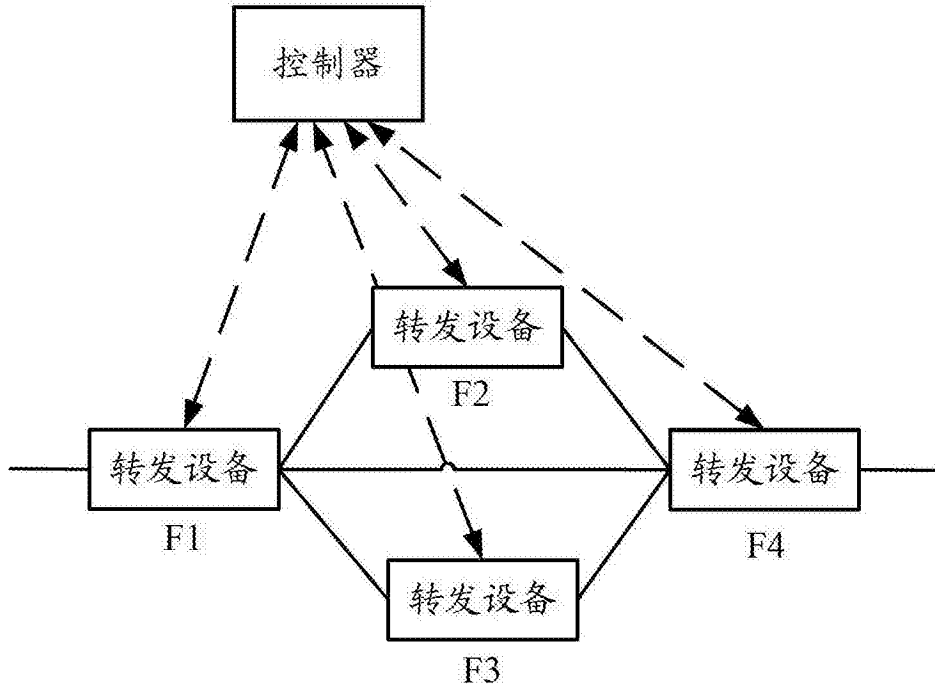


图1

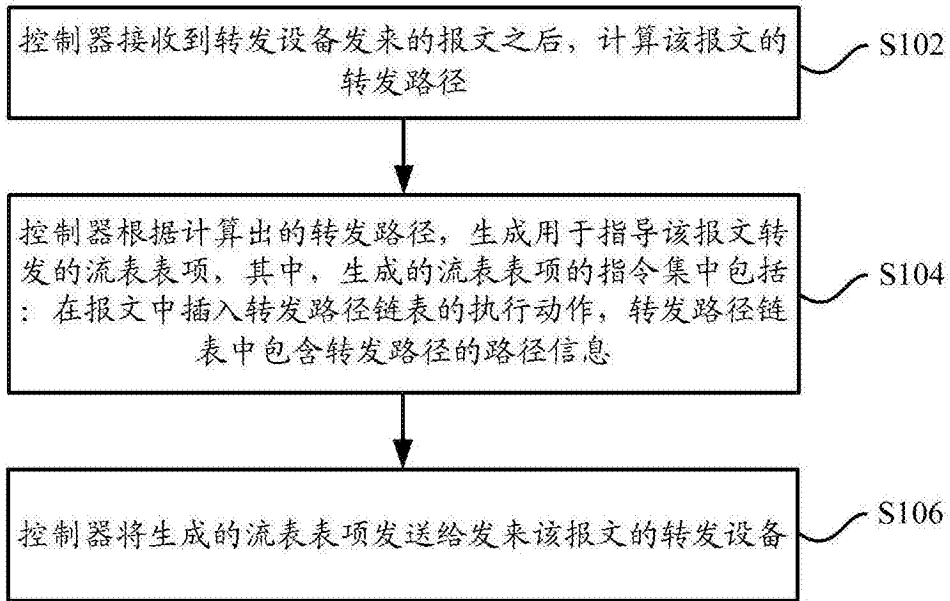


图2

转发链表指示	所经第一个转发设备的设备ID	所经第一个转发设备的转发接口的接口ID	所经第二个转发设备的设备ID	所经第二个转发设备的转发接口的接口ID	所经最后一个转发设备的设备ID	所经最后一个转发设备的转发接口的接口ID
--------	----------------	---------------------	----------------	---------------------	-------	-----------------	----------------------

图3

转发链表指示	所经第一个转发设备的转发接口的接口ID	所经第二个转发设备的转发接口的接口ID	所经最后一个转发设备的转发接口的接口ID
--------	---------------------	---------------------	--------	----------------------

图4

链表长度	数据流优先级 标记
------	--------------

图5



图6

链表长度=3	所经第一个转发设备的设备ID	所经第一个转发设备的转发接口的接口ID	所经第二个转发设备的设备ID	所经第二个转发设备的转发接口的接口ID	所经最后一个转发设备的设备ID	所经最后一个转发设备的转发接口的接口ID
--------	----------------	---------------------	----------------	---------------------	-----------------	----------------------

图7

链表长度=2	所经第二个转发设备的设备ID	所经第二个转发设备的转发接口的接口ID	所经最后一个转发设备的设备ID	所经最后一个转发设备的转发接口的接口ID
--------	----------------	---------------------	-----------------	----------------------

图8

链表长度=3	F1	PORT1	F2	PORT2	F4	PORT3
--------	----	-------	----	-------	----	-------

图9

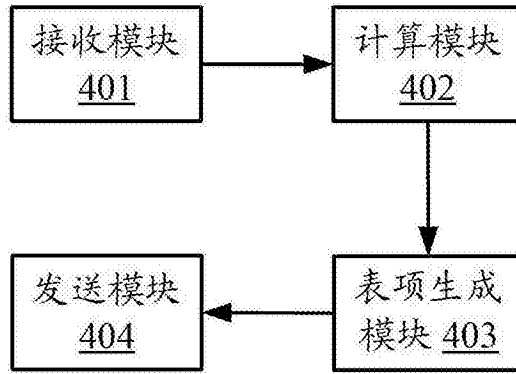


图10

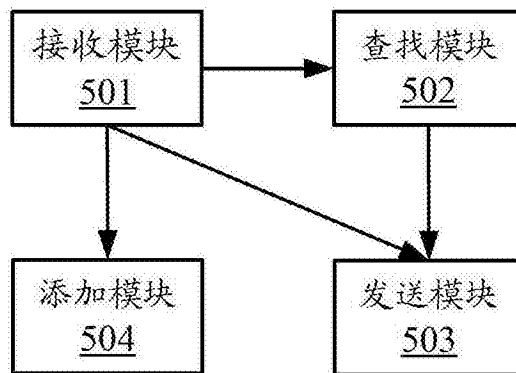


图11

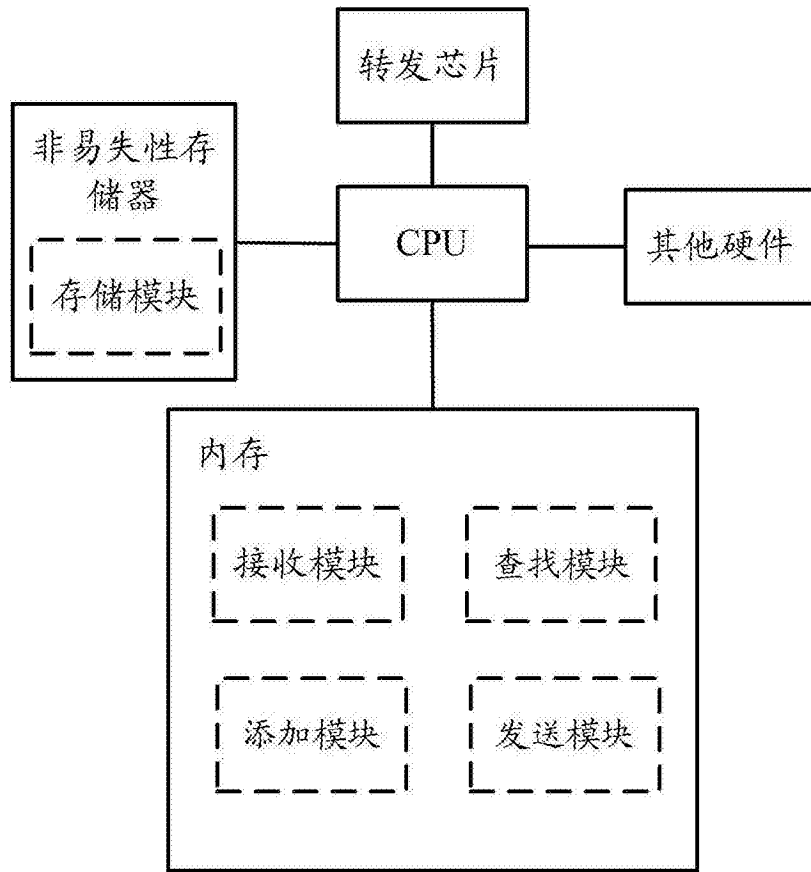


图12

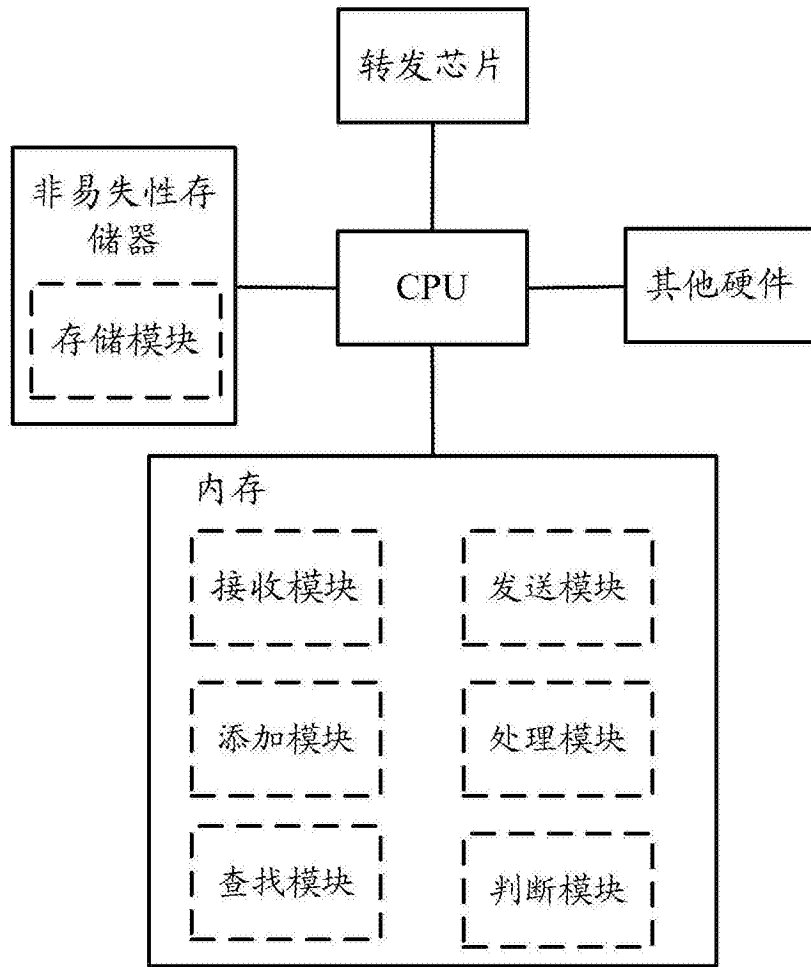


图13