



(12) 发明专利申请

(10) 申请公布号 CN 105610614 A

(43) 申请公布日 2016. 05. 25

(21) 申请号 201510997234. 2

(22) 申请日 2015. 12. 28

(71) 申请人 中国银联股份有限公司

地址 200135 上海市浦东新区含笑路 36 号  
银联大厦

(72) 发明人 周雍恺 祖立军 刘国宝 李戈

(74) 专利代理机构 中国专利代理(香港)有限公司  
72001

代理人 俞华梁 付曼

(51) Int. Cl.

H04L 12/24(2006. 01)

H04L 12/26(2006. 01)

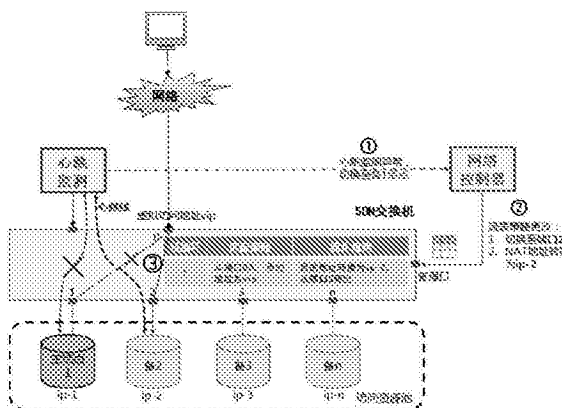
权利要求书1页 说明书5页 附图7页

(54) 发明名称

高可用访问系统以及高可用故障切换方法

(57) 摘要

本申请公开了一种高可用访问系统,所述系统包括访问资源池,所述访问资源池由多个待访问节点构成;心跳监测模块,用于通过定期发送心跳包来监测所述访问资源池中的节点状态;以及SDN模块,用于直接从外网接收访问请求,以及根据所述心跳监测模块提供的节点状态而对所述访问请求进行处理并直接转发到多个待访问节点中的一个适合节点,而没有经过其他代理主机。本申请还提供了一种高可用故障切换方法。



1. 一种高可用访问系统,所述系统包括:  
访问资源池,所述访问资源池由多个待访问节点构成;  
心跳监测模块,用于通过定期发送心跳包来监测所述访问资源池中的节点状态;以及  
SDN模块,用于直接从外网接收访问请求,以及根据所述心跳监测模块提供的节点状态而对所述访问请求进行处理并直接转发到多个待访问节点中的一个适合节点,而没有经过其他代理主机。
2. 如权利要求1所述的高可用访问系统,其中,所述SDN模块包括:  
SDN控制器,用于从所述心跳监测模块接收心跳监测异常信息,并通知SDN交换机进行流切换;以及  
所述SDN交换机,用于直接从外网接收访问请求,并根据所述SDN控制器的通知而对所述访问请求进行处理并直接转发到所述多个待访问节点中的一个适合节点。
3. 如权利要求2所述的高可用访问系统,其中,所述SDN交换机直接从外网接收访问请求并进行报文处理和转发,所述心跳监测模块和所述SDN控制器作为管理节点从旁路接入,并用于转发策略的制定与下发,而不直接影响业务流量的转发路径。
4. 如权利要求2或3所述的高可用访问系统,其中,所述多个待访问节点中的第一节点为设置为主服务节点,并且所述心跳监测模块通过经由所述SDN交换机定期发送心跳包来监测所述第一节点的状态。
5. 如权利要求4所述的高可用访问系统,其中,当所述心跳监测模块监测到所述第一节点宕机时,它配置成:  
选择所述多个待访问节点中的第二节点作为活动主机,并将该消息通知所述SDN控制器。
6. 如权利要求5所述的高可用访问系统,其中,所述SDN控制器进一步配置成:  
向所述SDN交换机下发流表更改的指令,通知网络流经由第二节点进行转发。
7. 一种高可用故障切换方法,所述方法包括:  
通过定期发送心跳包,心跳监测模块监测到访问资源池中作为活动节点的第一节点存在异常;  
心跳监测模块选取所述访问资源池中的第二节点作为新活动节点,并将消息通知SDN控制器;  
SDN控制器对流表进行更改,并向SDN交换机下发更改的流表;  
SDN交换机根据更改的流表而将从外网接收的访问请求转发到所述第二节点,而没有经过其他代理主机。
8. 如权利要求7所述的高可用故障切换方法,其中,所述SDN交换机根据所述控制器下发的流表,对数据层面流经的报文进行转发。

## 高可用访问系统以及高可用故障切换方法

### 技术领域

[0001] 本发明涉及高可用访问系统以及高可用故障切换方法。

### 背景技术

[0002] “高可用性”(High Availability)指一个系统经过专门的设计,达到降低停工时间、保持服务不间断的效果。服务系统的高可用通常通过多个冗余节点进行主备切换而实现,其系统的抽象模型如图1所示。

[0003] 如图1所示,服务端由多个访问节点构成一个资源池。高可用系统将外部的访问进行转换并导向资源池中某个可用的资源节点。高可用系统对外发布一个统一的“虚拟外部地址”以便于客户端进行访问。当客户端通过虚拟地址对服务系统进行访问时,高可用的前端代理将对该访问请求的目的地址进行转换,转换成内部的活动节点地址,并将访问流导向实际的活动节点。

[0004] 高可用系统会通过一定的机制对于活动的服务节点进行状态监测,当检测到活动节点宕机时,高可用系统将从访问资源池中选取备用节点,并将访问流切换至新的活动节点,从而达到不中断服务的效果。

[0005] 现有的高可用访问方案大多是基于主机实现的解决方案,例如Heartbeat、Corosync、Keepalived等。其通常的实现方式如图2所示。高可用核心功能通过前端的代理主机进行实现。当客户端的访问报文到达时,前端代理机将报文中的虚拟外部访问地址转换为内部活动主机的实际地址,并重新对访问报文进行封装,随后通过网络设备转发报文;当内部节点返回报文时,代理机再进行反向的地址转换与加解封装操作,并将报文传递至客户端。

[0006] 与此同时,前端代理机还需要定时对资源池中的节点进行心跳健康监测,当发现主节点宕机时,代理机需要选择其余的备用机作为新的活动节点,并更新转发表项,以便之后到达的报文能够及时进行切换。

[0007] 该类实现方式能够达到高可用访问的效果,不过存在如下的不足:

[0008] 首先,功能模块混杂,实现复杂度高。高可用代理主机同时实现了心跳健康监测与网络报文地址转换的功能。而后者属于纯网络层面的操作,应当通过专业的网络设备实现更为合适。

[0009] 其次,前端代理主机存在单点失效的风险。所有的外部访问报文在到达网络转发设备之前必须首先经过前端代理主机进行地址转换,一旦前端代理主机宕机,则整个系统都将停止工作。

[0010] 最后,在该类方案中,每个主数据流的进出都将通过代理主机进行网络报文地址的转换、加解封装以及反向代理等网络层面的操作。相比于直接应用专业的网络设备进行这类操作,其稳定性与吞吐性能有所欠缺。

### 发明内容

[0011] 为解决上述问题,根据本发明的一个方面,提供了一种高可用访问系统,所述系统包括:访问资源池,所述访问资源池由多个待访问节点构成;心跳监测模块,用于通过定期发送心跳包来监测所述访问资源池中的节点状态;以及SDN模块,用于直接从外网接收访问请求,以及根据所述心跳监测模块提供的节点状态而对所述访问请求进行处理并直接转发到多个待访问节点中的一个适合节点,而没有经过其他代理主机。

[0012] 在上述高可用访问系统中,所述SDN模块包括:SDN控制器,用于从所述心跳监测模块接收心跳监测异常信息,并通知SDN交换机进行流切换;以及所述SDN交换机,用于直接从外网接收访问请求,并根据所述SDN控制器的通知而对所述访问请求进行处理并直接转发到所述多个待访问节点中的一个适合节点。

[0013] 在上述高可用访问系统中,所述SDN交换机直接从外网接收访问请求并进行报文处理和转发,所述心跳监测模块和所述SDN控制器作为管理节点从旁路接入,并用于转发策略的制定与下发,而不直接影响业务流量的转发路径。

[0014] 在上述高可用访问系统中,所述多个待访问节点中的第一节点为设置为主服务节点,并且所述心跳监测模块通过经由所述SDN交换机定期发送心跳包来监测所述第一节点的状态。

[0015] 在上述高可用访问系统中,当所述心跳监测模块监测到所述第一节点宕机时,它配置成:选择所述多个待访问节点中的第二节点作为活动主机,并将该消息通知所述SDN控制器。

[0016] 在上述高可用访问系统中,所述SDN控制器进一步配置成:向所述SDN交换机下发流表更改的指令,通知网络流经由第二节点进行转发。

[0017] 根据本申请的另一个方面,提供了一种高可用故障切换方法,所述方法包括:通过定期发送心跳包,心跳监测模块监测到访问资源池中作为活动节点的第一节点存在异常;心跳监测模块选取所述访问资源池中的第二节点作为新活动节点,并将消息通知SDN控制器;SDN控制器对流表进行更改,并向SDN交换机下发更改的流表;SDN交换机根据更改的流表而将从外网接收的访问请求转发到所述第二节点,而没有经过其他代理主机。

[0018] 在上述高可用故障切换方法中,所述SDN交换机根据所述控制器下发的流表,对数据层面流经的报文进行转发。

[0019] 相比于现有的基于主机的高可用实现方案,本申请的技术方案至少具有如下优势:

[0020] 第一,本申请的技术方案充分利用SDN架构的优势,通过SDN提供的开放网络接口,将网络功能模块从主机中分离而出交由专业的网络设备进行处理,从而将心跳监测模块与网络模块进行最大程度的解耦合,简化系统设计,同时增强了可靠性。

[0021] 第二,所有的外部访问报文直接通过网络设备转发至最终服务节点,无需通过中间主机,由此克服了基于主机实现方案中的前端代理单点失效问题。

[0022] 第三,在性能方面,本申请的技术方案直接通过专业的网络设备进行报文地址的转换、加解封装以及反向代理、转发等操作,可以降低故障切换过程中的丢包概率并且提高吞吐性能,更好地达到高可用系统的连接保持、平滑切换效果。

## 附图说明

[0023] 在参照附图阅读了本发明的具体实施方式以后,本领域技术人员将会更清楚地了解本发明的各个方面。本领域技术人员应当理解的是:这些附图仅仅用于配合具体实施方式说明本发明的技术方案,而并非意在对本发明的保护范围构成限制。

[0024] 图1是现有的高可用访问系统的示意图;

[0025] 图2是现有的基于主机的高可用实现方案示意图;

[0026] 图3是SDN模型架构图;

[0027] 图4是图2中的前端代理机的模块分解图;

[0028] 图5是根据本申请的一个实施例、基于SDN架构高可用实现的模块分解图;

[0029] 图6是根据本申请的一个实施例、基于SDN架构的高可用系统的框架图;以及

[0030] 图7是根据本申请的一个实施例、基于SDN架构的高可用故障切换原理图。

### 具体实施方式

[0031] 下面介绍的是本发明的多个可能实施例中的一些,旨在提供对本发明的基本了解,并不旨在确认本发明的关键或决定性的要素或限定所要保护的范围。容易理解,根据本发明的技术方案,在不变更本发明的实质精神下,本领域的一般技术人员可以提出可相互替换的其它实现方式。因此,以下具体实施方式以及附图仅是对本发明的技术方案的示例性说明,而不应当视为本发明的全部或者视为对本发明技术方案的限定或限制。

[0032] 网络技术极大地推动了互联网的迅猛发展。不过,传统的网络技术一直较为封闭,核心技术与设备市场为少数几家厂商所垄断,并且上层的应用无法直接操作到网络层以下的设备。2009年,斯坦福团队正式提出了软件定义网络(Software Defined Networking, SDN)的概念,旨在打破网络层的封闭局面,通过对外提供标准的可编程网络接口,使得应用可以更好地与网络进行交互,从而提升系统的整体优化性能。

[0033] 典型的SDN模型如图3所示,其由SDN交换机与SDN控制器两部分组成。SDN控制器是整个网络的中心控制节点,负责维护整个网络的转发策略,并通过直连每个SDN交换机的管理端口下发流表指令。每个SDN交换机则存储着一张网络流表(Flow Table),以接收控制器所下发的流表策略,并通过匹配流表中的表项对数据层面流经的报文进行转发。

[0034] SDN控制器与交换机之间的通信协议被称为南向协议,最早提出的南向协议是OpenFlow,目前已经发展至1.5版本。在OpenFlow协议中,每个流表项(Flow Entry)由分类匹配域(Match Field)与操作规则(Instruction)等部分组成。分类匹配字域指示流入报文的匹配规则,一旦匹配成功,则进行Instruction部分所指示的操作。例如,在图3中的流表中,规则1表示所有从端口1进入的报文都从端口2进行转发;流表规则2则指示将所有源地址为0.0.0.0的报文进行丢弃。SDN交换机根据这些规则对所有的流经报文进行操作。

[0035] 继续参考图2,现有的前端代理系统可以划分为如下两个主功能模块,如图4所示:

[0036] 1)心跳监测模块:通过定期发送心跳包监测资源池中的主机状态;

[0037] 2)网络功能模块:对客户端的访问报文进行地址转换,并导流至实际工作节点。

[0038] 两个模块之间呈松耦合关系,仅当心跳监测模块发现主节点状态异常时,才会与网络功能模块进行交互,通知其进行流切换。

[0039] 与现有的技术方案不同,根据本申请的一个实施例,在基于SDN架构的高可用实现

的模块中仍保留心跳监测模块,同时将网络功能模块从前端代理主机中分离出,转而通过SDN架构结合专业的网络设备进行实现。新方案的系统模块框图如图5所示。

[0040] 如图5所示,在一个实施例中,SDN模块由网络控制器与SDN交换机两部分组成。对于外网的访问请求,数据报文将直接通过SDN交换机进行报文处理与转发,而不必经过其他代理主机。心跳监测模块与SDN控制器仅是作为管理节点从旁路接入,主要负责转发策略的制定与下发,并不直接影响到业务流量的转发路径。

[0041] 基于图5的系统模块分解图,将其中SDN模块进行实例化后,本方案的实际系统部署可由图6所示。在一个实施例中,SDN模块由网络控制器与SDN交换机两部分组成。在系统初始阶段,系统选取资源池中的节点1为主服务节点,网络控制器首先将如下两条流表规则写入SDN交换机:

[0042] 表1:系统初始阶段SDN交换机的流表

流编号	匹配策略	操作规则
[0043] 1	从端口0入, 目的地址为vip	目的地址转换为ip-1, 从端口1转出, 并且记录客户端地址
2	从端口1入, 目的地址在客户端地址列表中	源地址转换为vip, 从端口0转出

[0044] 流规则1的意义为将外网访问的报文转发至节点1进行处理,同时对报文的地址进行转换(转换为节点1的地址)。流规则2则是流规则1的反向规则。

[0045] 由图6可知,客户端的访问报文将被直接传送至SDN交换机,无需经过代理机等前置设备。根据流规则1,SDN交换机对于所有从端口0进入的报文进行匹配,如果目的地址为vip,则将报文的地址转换为节点1的地址ip-1,并从交换机的端口1将报文转出。

[0046] 对于从节点1返回的报文则通过流规则2进行反向的操作。

[0047] 与此同时,心跳监测模块通过定期发送心跳报文监测节点1的状态,该过程与上述的外部报文访问流程是完全独立的。

[0048] 图7示出了根据本申请的一个实施例、基于SDN架构的高可用故障切换原理图。

[0049] 结合图7,根据本申请的一个实施例,当心跳监测模块检测到节点1宕机时,系统将进行如下的处理流程:

[0050] (1)心跳监测模块检测到节点1宕机,选取备用节点2作为活动主机,并将该消息通知于网络控制器;

[0051] (2)SDN网络控制器向SDN交换器下发流表更改指令,更新后的流表如下表所示:

[0052] 表2:主备切换后SDN交换机的流表

流编号	匹配策略	操作规则
[0053] 1	从端口0入, 目的地址为vip	目的地址转换为ip-2, 从端口2转出, 并且记录客户端地址
2	从端口2入, 目的地址在客户端地址列表中	源地址转换为vip, 从端口0转出

[0054] (3)此后的网络流流经SDN交换器后都经由端口2进行转发,主备切换完成。

[0055] 由此,本申请的技术方案通过软件定义网络(SDN)的架构实现了连接访问的高可用,并且能够取得比原有基于主机的高可用方案更佳的系统性能以及更为简易的系统实现。

[0056] 综上所述,本申请提出了一种基于软件定义网络(SDN)架构的高可用实现方案。本申请的技术方案充分利用SDN架构的优势,通过SDN所提供的开放网络接口,将网络功能模块从主机中分离而出交由专业的网络设备进行处理,由此将高可用系统中的各模块进行最大程度的解耦,简化系统设计,同时增强了可靠性。

[0057] 上文中,参照附图描述了本发明的具体实施方式。但是,本领域中的普通技术人员能够理解,在不偏离本发明的精神和范围的情况下,还可以对本发明的具体实施方式作各种变更和替换。这些变更和替换都落在本发明权利要求书所限定的范围内。

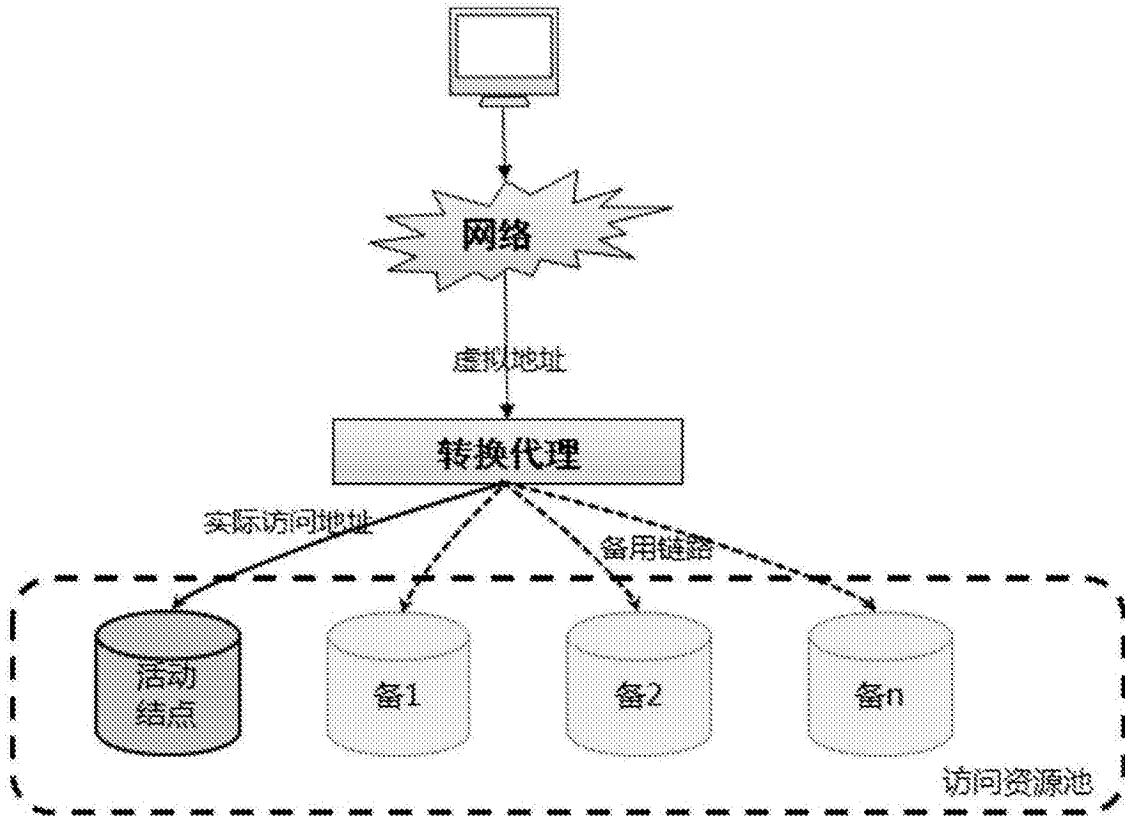


图1



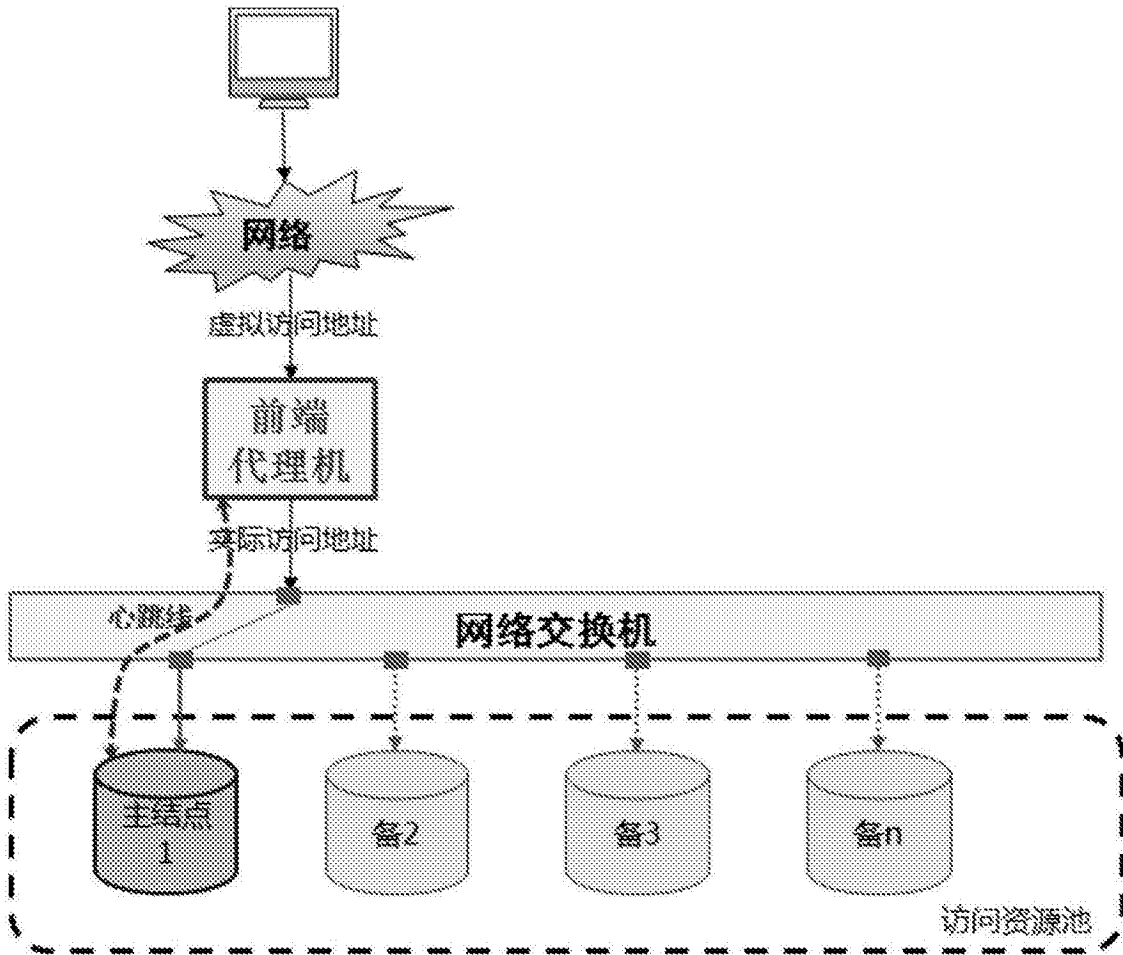


图2

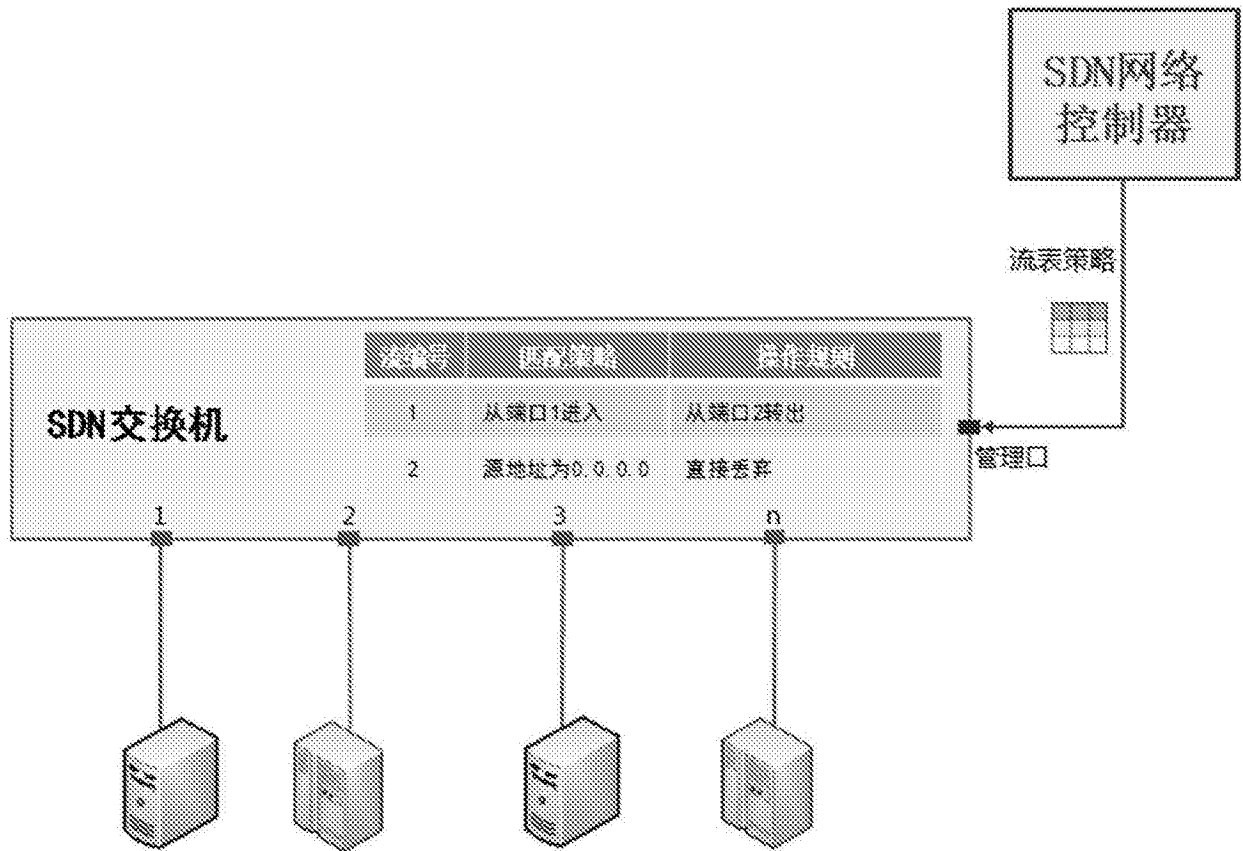


图3

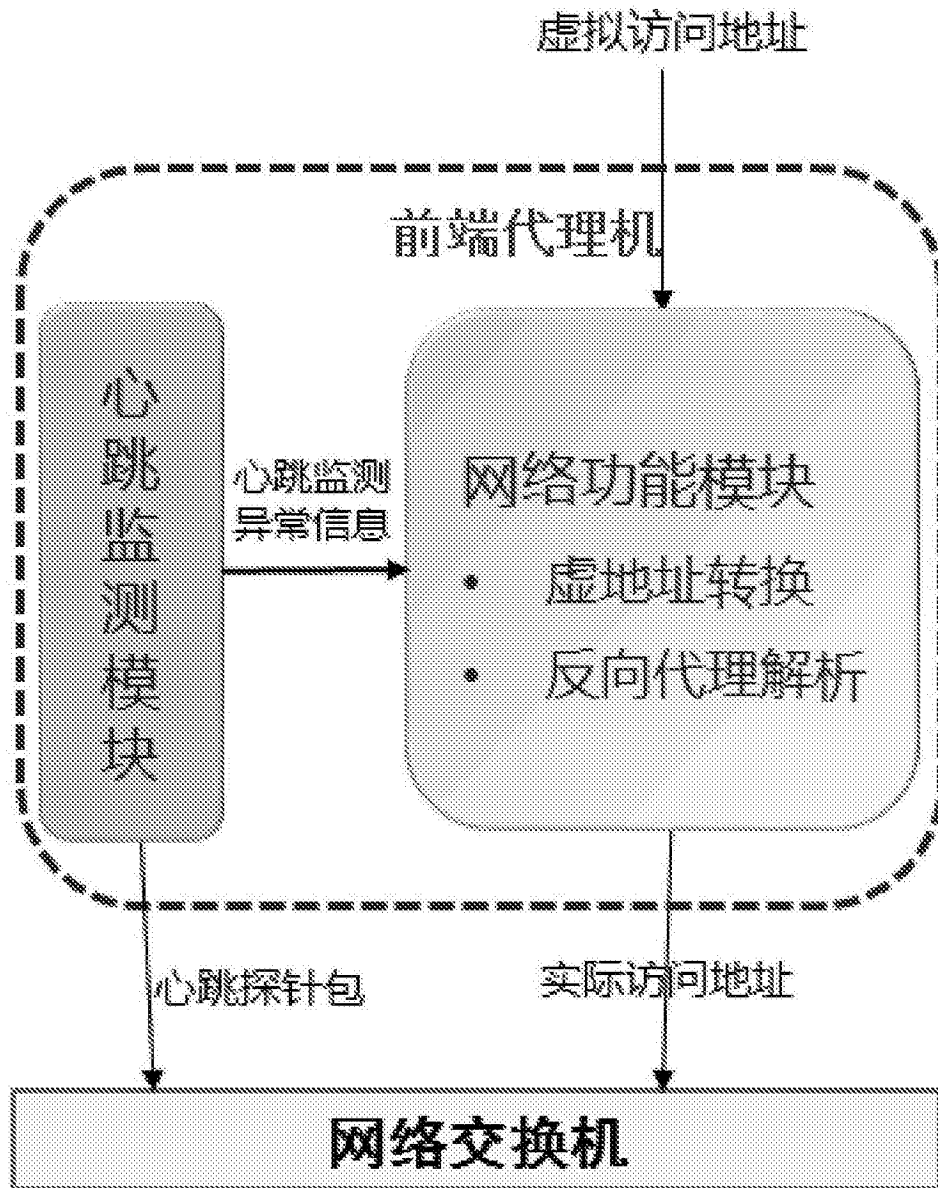


图4

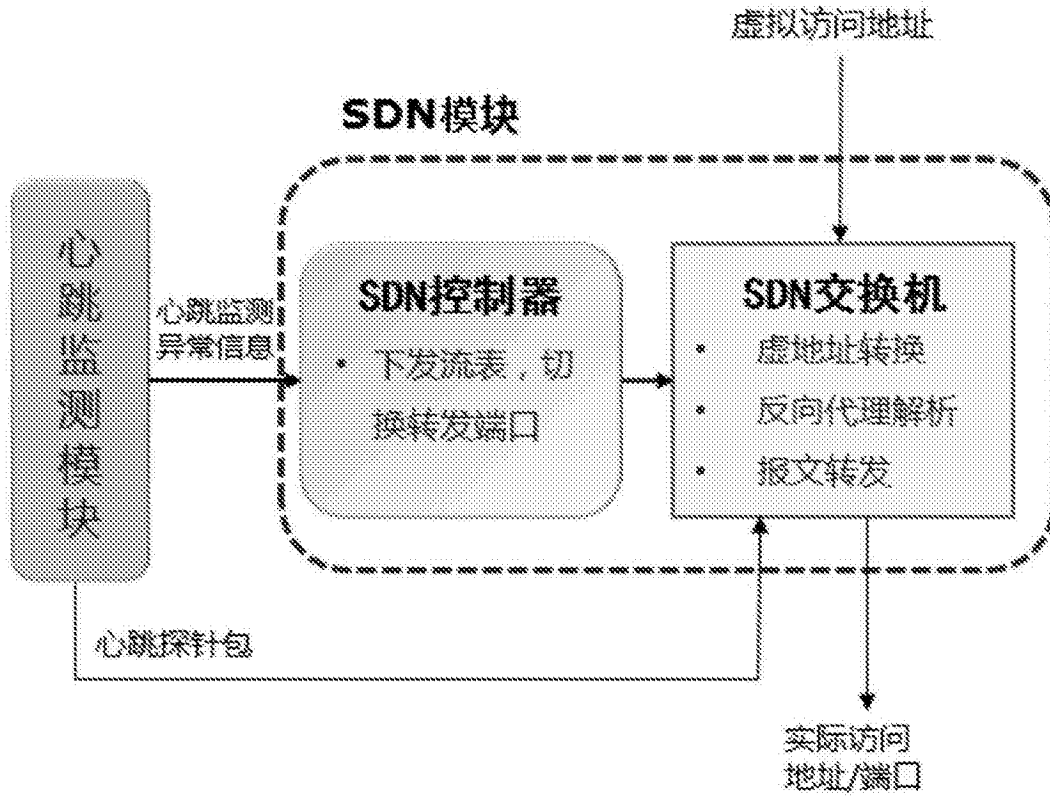


图5

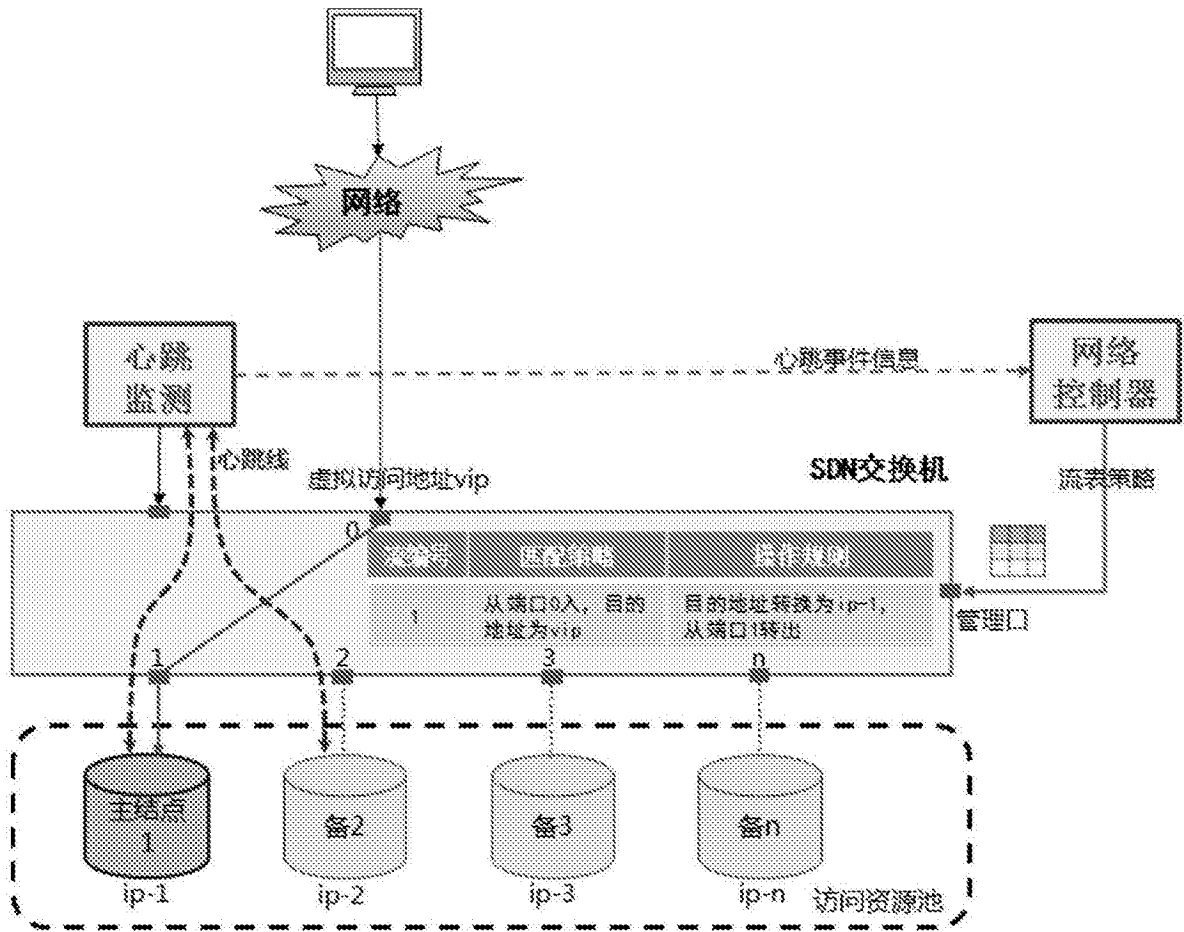


图6

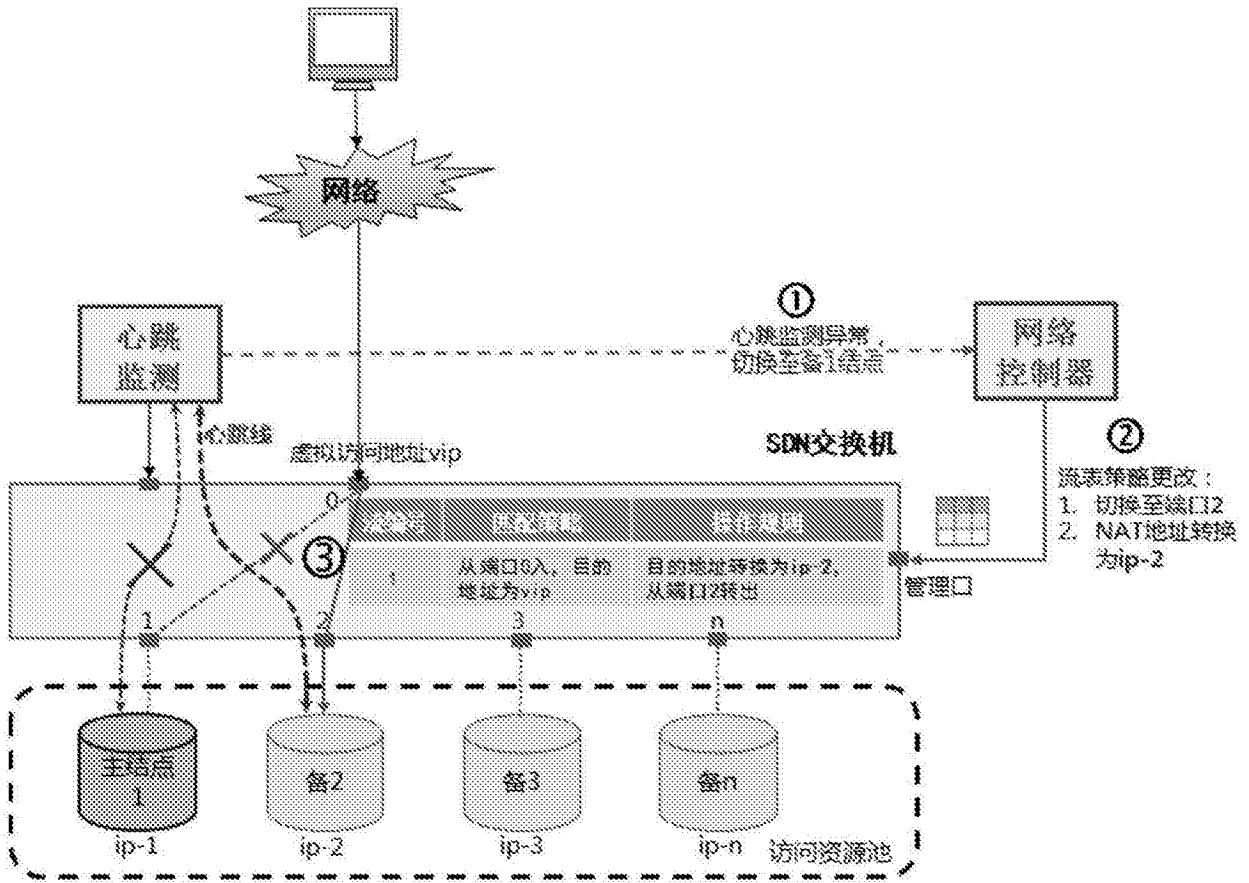


图7