



(12) 发明专利申请

(10) 申请公布号 CN 105281947 A

(43) 申请公布日 2016. 01. 27

(21) 申请号 201510289270. 3

(51) Int. Cl.

(22) 申请日 2015. 05. 29

H04L 12/24(2006. 01)

H04L 12/741(2013. 01)

(30) 优先权数据

62/005, 177 2014. 05. 30 US

62/089, 028 2014. 12. 08 US

(71) 申请人 布洛凯通讯系统公司

地址 美国加利福尼亚州

(72) 发明人 穆罕默德·杜拉尼

赛义德·纳蒂夫·那瓦兹

埃斯瓦拉·陈塔拉帕蒂 张毅

辛图·帕扬库拉斯

(74) 专利代理机构 北京律盟知识产权代理有限

责任公司 11287

代理人 张世俊

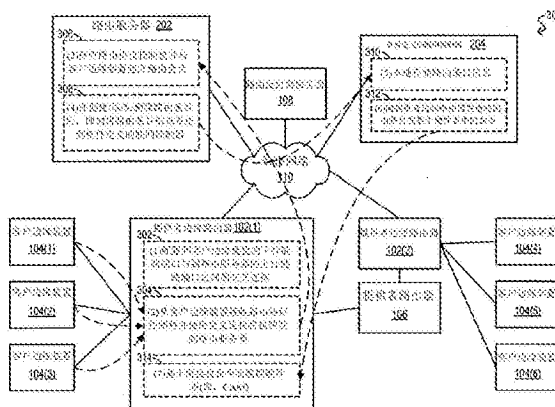
权利要求书3页 说明书7页 附图7页

(54) 发明名称

用于将遗留网络转换成具备 SDN 能力的网络的技术

(57) 摘要

本发明提供用于将遗留网络转换成具备软件定义的联网 SDN 的能力的网络的技术。在一个实施例中, 路由服务器可接收从网络装置始发的一或多个路由协议包, 其中所述一或多个路由协议包是经由在网络路由器上配置的交叉连接转发到所述路由服务器。所述路由服务器可进一步基于所述一或多个路由协议包而在所述路由服务器与所述网络装置之间建立路由协议会话, 且可将路由条目添加到本地路由表。在添加所述路由条目后, 所述路由服务器可即刻自动地调用应用程序编程接口 API 来将所述路由条目发射到软件定义的联网 SDN 控制器。



1. 一种方法,其包括:

由路由服务器接收从网络装置始发的一或多个路由协议包,所述一或多个路由协议包是通过网络路由器转发到所述路由服务器;

由所述路由服务器基于所述一或多个路由协议包在所述路由服务器与所述网络装置之间建立路由协议会话;

在建立所述路由协议会话之后,由所述路由服务器将路由条目添加到本地路由表;及

在添加所述路由条目后,即刻由所述路由服务器自动地调用应用程序编程接口 API 来将所述路由条目发射到软件定义的联网 SDN 控制器。

2. 根据权利要求 1 所述的方法,其中所述一或多个路由协议包是经由在所述网络路由器上配置的交叉连接转发到所述路由服务器,所述交叉连接在将所述网络路由器连接到所述网络装置的下行链路端口与将所述网络路由器连接到所述路由服务器的上行链路端口之间建立路径,使得在所述下行链路端口处接收到的所有路由协议业务自动地重新引导到所述上行链路端口。

3. 根据权利要求 2 所述的方法,其中所述交叉连接是使用与所述下行链路端口或所述上行链路端口相关联的接入控制列表 ACL 来实施。

4. 根据权利要求 1 所述的方法,其中所述一或多个路由协议包是通过所述网络路由器转发到所述路由服务器,而并不由所述网络路由器的控制平面进行本地处理。

5. 根据权利要求 1 所述的方法,其中使所述 API 在所述本地路由表中注册,使得在于所述本地路由表中进行路由条目插入、修改或删除后,即刻自动地调用所述 API。

6. 根据权利要求 1 所述的方法,其中将所述路由服务器实施为包括多个节点的群集,所述群集可操作以提供路由服务器冗余性。

7. 根据权利要求 6 所述的方法,其中所述多个节点包含现用节点及备用节点,且其中所述现用节点可操作以与所述网络装置建立所述路由协议会话。

8. 根据权利要求 7 所述的方法,其中由所述现用节点建立所述路由协议会话包括:

使用于所述路由协议会话的路由协议状态机从所述现用节点同步到所述备用节点;及致使所述备用节点将路由协议响应包发射到所述网络装置。

9. 根据权利要求 7 所述的方法,其中所述现用节点进一步可操作以在建立所述路由协议会话后即刻将对所述现用节点的路由数据库的更新同步到所述备用节点。

10. 根据权利要求 1 所述的方法,其中所述 API 为由所述 SDN 控制器公开的标准化代表性状态传送 REST API。

11. 根据权利要求 1 所述的方法,其中在接收到对所述 API 的所述调用后,所述 SDN 控制器即刻可操作以:

将所述路由条目添加到本地流条目数据库;及

向所述网络路由器发射用于将所述路由条目安装于所述网络路由器的硬件转发表中的命令。

12. 根据权利要求 11 所述的方法,其中所述 SDN 控制器进一步可操作以使用 fib 聚合算法在周期性基础上压缩所述本地流条目数据库。

13. 一种非暂时性计算机可读存储媒体,其上存储有可由路由服务器的处理器执行的程序代码,所述程序代码致使所述处理器:

接收从网络装置始发的一或多个路由协议包,所述一或多个路由协议包是通过网络路由器转发到所述路由服务器;

基于所述一或多个路由协议包在所述路由服务器与所述网络装置之间建立路由协议会话;

在建立所述路由协议会话之后,将路由条目添加到本地路由表;及

在添加所述路由条目后,即刻自动地调用应用程序编程接口 API 来将所述路由条目发射到软件定义的联网 SDN 控制器。

14. 根据权利要求 13 所述的非暂时性计算机可读存储媒体,其中所述一或多个路由协议包是经由在所述网络路由器上配置的交叉连接转发到所述路由服务器,所述交叉连接在将所述网络路由器连接到所述网络装置的下行链路端口与将所述网络路由器连接到所述路由服务器的上行链路端口之间建立路径,使得在所述下行链路端口处接收到的所有路由协议业务自动地重新引导到所述上行链路端口。

15. 根据权利要求 13 所述的非暂时性计算机可读存储媒体,其中所述一或多个路由协议包是通过所述网络路由器转发到所述路由服务器,而并不由所述网络路由器的控制平面进行本地处理。

16. 根据权利要求 13 所述的非暂时性计算机可读存储媒体,其中所述路由服务器被实施为包括多个节点的群集,所述群集可操作以提供路由服务器冗余性,

其中所述多个节点包含现用节点及备用节点,所述现用节点可操作以与所述网络装置建立所述路由协议会话,且

其中所述现用节点通过以下操作建立所述路由协议会话:

使用于所述路由协议会话的路由协议状态机从所述现用节点同步到所述备用节点;及致使所述备用节点将路由协议响应包发射到所述网络装置。

17. 一种计算机系统,其包括:

一或多个处理器;及

非暂时性计算机可读媒体,其上存储有程序代码,所述程序代码在由所述一或多个处理器执行时致使所述一或多个处理器:

接收从网络装置始发的一或多个路由协议包,所述一或多个路由协议包是通过网络路由器转发到所述计算机系统;

基于所述一或多个路由协议包在所述计算机系统与所述网络装置之间建立路由协议会话;

在建立所述路由协议会话之后,将路由条目添加到本地路由表;及

在添加所述路由条目后,即刻自动地调用应用程序编程接口 API 来将所述路由条目发射到软件定义的联网 SDN 控制器。

18. 根据权利要求 17 所述的计算机系统,其中所述一或多个路由协议包是经由在所述网络路由器上配置的交叉连接转发到所述计算机系统,所述交叉连接在将所述网络路由器连接到所述网络装置的下行链路端口与将所述网络路由器连接到所述计算机系统的上行链路端口之间建立路径,使得在所述下行链路端口处接收到的所有路由协议业务自动地重新引导到所述上行链路端口。

19. 根据权利要求 17 所述的计算机系统,其中所述一或多个路由协议包是通过所述网

络路由器转发到所述计算机系统,而并不由所述网络路由器的控制平面进行本地处理。

20. 根据权利要求 17 所述的计算机系统,其中所述计算机系统被实施为包括多个节点的群集,所述群集可操作以提供路由服务器冗余性,

其中所述多个节点包含现用节点及备用节点,所述现用节点可操作以与所述网络装置建立所述路由协议会话,且

其中所述现用节点通过以下操作建立所述路由协议会话:

使用于所述路由协议会话的路由协议状态机从所述现用节点同步到所述备用节点;及致使所述备用节点将路由协议响应包发射到所述网络装置。

## 用于将遗留网络转换成具备 SDN 能力的网络的技术

### [0001] 相关申请案交叉参考

[0002] 本申请案依据 35U. S. C. 119(e) 主张以下申请案的权益及优先权：2014 年 5 月 30 日提出申请的标题为“用以将遗留网络转换为具备 OpenFlow 能力的网络的方法及实施方案 (METHOD AND IMPLEMENTATION TO TRANSFORM LEGACY NETWORKS TO OPENFLOW ENABLED NETWORK)”的第 62/005, 177 号美国临时申请案, 及 2014 年 12 月 8 日提出申请的标题为“用于将遗留网络转换成具备 SDN 能力的网络的技术 (TECHNIQUES FOR TRANSFORMING LEGACY NETWORKS INTO SDN-ENABLED NETWORKS)”的第 62/089, 028 号美国临时申请案。这些临时申请案的全部内容出于所有目的而引用方式并入本文中。

### 技术领域

#### 背景技术

[0003] 软件定义的联网 (SDN) 及特定来说 OpenFlow (其为用于实施 SDN 的标准化通信协议) 具有用于将常规方法重新成像到层 3 联网的许多未经锁定的新工具。举例来说, SDN 使得远程控制器 (例如, 服务器计算机系统) 能够执行通常由 L3 网络中的专用网络装置 (例如, 路由器或交换机) 执行的控制平面功能。此类控制平面功能的实例包含路由协议会话建立、构建路由表等等。所述远程控制器可接着经由 OpenFlow (或某一其它类似协议) 将适当命令传达到专用网络装置以根据由远程控制器做出的路由决策来转发数据业务。对网络控制平面 (驻存于远程控制器上) 与网络转发平面 (驻存于专用网络装置上) 的此种分离可降低专用网络装置的复杂性 / 成本且可简化网络管理、规划及配置。

[0004] 遗憾地, 由于 SDN 及 OpenFlow 仍为相对新的技术, 因此客户将这些技术采纳于其生产环境中的速度一直很缓慢。因此, 将需要具有促进部署基于 SDN 的网络借此产生这些技术的可信度并促进对这些技术的采纳的技术。

#### 发明内容

[0005] 本发明提供用于将遗留网络转换成具备软件定义的联网 (SDN) 的能力的网络的技术。在一个实施例中, 路由服务器可接收从网络装置始发的一或多个路由协议包, 其中所述一或多个路由协议包是经由在网络路由器上配置的交叉连接转发到所述路由服务器。所述路由服务器可进一步基于所述一或多个路由协议包而在所述路由服务器与所述网络装置之间建立路由协议会话, 且可将路由条目添加到本地路由表。在添加所述路由条目后, 所述路由服务器可即刻自动地调用应用程序编程接口 (API) 来将所述路由条目发射到软件定义的联网 (SDN) 控制器。

[0006] 以下详细说明及附图提供对本发明的性质及优点的更佳理解。

#### 附图说明

[0007] 图 1 描绘根据一实施例的遗留 L3 网络。

- [0008] 图 2 描绘根据一实施例图 1 的网络的已被修改为支持 SDN 转换的版本。
- [0009] 图 3 描绘根据一实施例的遗留 /SDN 转换 / 变换 workflow。
- [0010] 图 4 描绘根据一实施例图 2 的网络的包含路由服务器群集的的版本。
- [0011] 图 5 描绘根据一实施例用于使路由协议状态在路由服务器群集的节点之间同步的 workflow。
- [0012] 图 6 描绘根据一实施例的网络路由器。
- [0013] 图 7 描绘根据一实施例的计算机系统。

## 具体实施方式

[0014] 在以下说明中,出于解释的目的,陈述了众多实例及细节以便提供对各种实施例的理解。然而,所属领域的技术人员将明白,某些实施例可在没有这些细节中的一些细节的情况下实践或可以对这些细节的修改或等效形式来实践。

### [0015] 1. 概述

[0016] 本发明的实施例提供用于将遗留层 3 网络(即,包括各自执行控制平面及数据平面两者的专用网络装置的网络)变换成具备 SDN 能力的网络(即,其中控制平面功能被分离并被合并到不同于专用网络装置的远程控制器(本文中称为“路由服务器”)中的网络)的技术。在一组实施例中,这些技术可包含通过管理员配置网络路由器以将从其它网络装置接收的路由协议包转发到路由服务器(而非在所述网络路由器上对所述路由协议包进行本地处理)。所述路由服务器可接着使用所接收路由协议包直接与其它网络装置建立路由协议会话(例如,OSPF、ISIS、BGP 等)。此步骤可包含填充路由数据库、计算针对各种目的地地址的最短/最佳路径,及构建具有针对下一目的地地址的下一跳点信息的路由表。

[0017] 在创建、修改或删除其路由表中的给定路由条目后,所述路由服务器可即刻自动地调用标准化应用程序编程接口(API)来将关于所述路由条目的信息传达到 SDN(例如,OpenFlow)控制器。在特定实施例中,标准化 API 可为由 SDN 控制器所理解的代表性状态传送(REST)API。所述 SDN 控制器可将所接收路由条目信息存储于其自身的数据库(本文中称为流条目数据库)中。所述 SDN 控制器可随后将命令发送到所述网络路由器,所述命令致使将所述路由条目安装/编程到路由器的硬件转发表(例如,CAM)中,借此使得所述路由器能够以线路速率根据所述路由条目来转发传入数据业务。

[0018] 借助上文所描述的方法,具有遗留 L3 网络的客户可较容易地将那些遗留网络迁移为支持具有单独控制及转发平面的 SDN 范型。此又可允许客户减少其资金/运营成本(因为不再需要所述成本来购买及部署具有复杂控制平面功能性的专用网络装置),且可确保其网络能够缩放以满足不断增加的带宽需求并支持新类型的网络服务。这特别有益于 IP 边缘网络(即,位于服务提供者维持的 IP/MPLS 网络与客户接入网络之间的边缘上的网络),因为 IP 边缘网络通常为处于增加规模及服务粒度的最大压力下的服务点。

[0019] 在某些实施例中,除促进到 SDN 范型的转变以外,本文中所描述的技术还可实现经配置以执行控制平面功能的路由服务器的高可用性(HA)。在这些实施例中,多个物理机器(即,节点)可协同工作以充当虚拟路由服务器群集(例如,使用虚拟路由器冗余性协议(VRRP)或其增强版本,例如 VRRP-e)。网络装置可与虚拟路由服务器群集的虚拟 IP 地址通信以便与所述群集中的现用节点建立路由协议会话。于是,当现用节点发生故障时,控制

平面处理可自动地被从现用节点失效恢复到备用节点,借此保存路由服务器的可能性。在特定实施例中,虚拟路由服务器群集可实施用于以下的新颖技术:(1)在会话建立期间使用于给定路由协议会话的路由协议状态机从现用节点同步到备用节点,及(2)允许备用节点(而非现用节点)将路由协议“发射”包(例如,响应消息)发出到始发网络装置。此技术可确保备用节点与现用节点恰当地同步且可避免在发生到备用节点的失效恢复的情况下在所述备用节点上重新构建路由协议状态机(及/或路由数据库)。

[0020] 在以下章节中进一步详细地描述本发明的这些及其它方面。

## [0021] 2. 网络环境

[0022] 图1描绘本发明的实施例可应用于的遗留L3网络100的实例。如所展示,网络100包含分别连接到客户边缘(CE)网络装置104(1)-(3)及104(4)-(6)的提供者边缘(PE)路由器102(1)及102(2)。网络100进一步包含内部提供者路由器106以及经由管理网络110连接到PE路由器102(1)及102(2)的路由反射器(RR)服务器108。如此项技术中所已知,RR服务器108可充当用于在网络100内传播路由协议信息的焦点且因此避免对PE路由器102(1)与102(2)(及网络100中的任何其它PE路由器)之间的全网状连接性的需要。

[0023] 在图1的实例中,每一CE网络装置104经配置以与其所连接PE路由器102创建路由协议会话。每一PE路由器102又经配置以执行在网络100的路由域内建立路线所需的控制平面功能(例如,建立/维持相邻关系、计算最佳路线、构建路由表等)以及物理转发网络业务。如背景技术章节中所述,关于在如PE路由器102(1)/102(2)等专用网络装置上执行控制平面及转发平面功能两者的一个问题是,此会限制网络的可缩放性及灵活性。这在如图1中所展示的提供者边缘网络中特别成问题,这常常是服务提供者在尝试增加网络规模及服务粒度时的“压力点”。

[0024] 为了解决这些及其它类似问题,图2描绘根据一实施例网络100的已被修改为促进将网络100转变或变换成具备SDN能力的网络的版本(即,网络200)。如所展示,网络200包含经由管理网络110与PE路由器102(1)及102(2)通信地耦合的路由服务器202及SDN控制器204。路由服务器202为可代表PE路由器102(1)及102(2)集中执行控制平面功能的基于软件或硬件的组件。在特定实施例中,路由服务器202可为在物理或虚拟机器上运行的博科通信系统公司(Brocade Communications Systems, Inc.)的Vyatta路由服务器软件的实例。SDN控制器204是可从路由服务器202接收(例如,经由适当“北向”协议)被引导到PE路由器102(1)及102(2)的命令且可将那些命令转发(例如,经由适当“南向”协议,例如OpenFlow)到路由器102(1)及102(2)以供在那些装置上执行的基于软件或硬件的组件。在特定实施例中,SDN控制器204可为OpenDaylight(ODL)控制器的实例。

[0025] 如下一章节中所描述,路由服务器202及SDN控制器204可结合每一PE路由器102来执行用于使将图1的遗留网络100转换或变换成具备SDN能力的网络自动化的工作流。换句话说,此转换/变换工作流可:(1)使得先前在PE路由器102(1)及102(2)上本地执行的L3控制平面功能(例如,建立路由协议会话、计算最佳路线、构建路由表等)能够自动集中在路由服务器202中;及(2)使得由路由服务器202确定的路由条目能够自动传播到(即,被编程到)PE路由器102(1)及102(2)的硬件转发表。对于此工作流,网络100的运营商可较快速且较容易地实现移动到基于SDN的网络范型的运营、成本及可缩放性益

处。

[0026] 应了解,图 1 及 2 为说明性的且并不打算限制本文中所论述的实施例。举例来说,虽然这些图描绘每一网络元件的某一数目(例如,两个 PE 路由器、六个 CE 装置等),但可支持这些元件的任何数目。此外,虽然这些图具体描绘提供者 /IP 边缘网络,但本发明的技术可应用于此项技术中已知的任何类型的遗留网络。另外,虽然路由服务器 202 及 SDN 控制器 204 被展示为两个单独实体,但在某些实施例中,归于这些组件的功能可由单个实体(例如,组合式路由服务器 / 控制器)执行。所属领域的普通技术人员将认识到其它变化形式、修改形式及替代方案。

### [0027] 3. 遗留 /SDN 变换 workflow

[0028] 图 3 描绘根据一实施例可在图 2 的网络 200 内执行的用于促进将所述网络转换 / 变换成具备 SDN 能力的网络的工作流 300。虽然工作流 300 描述具体来说与 PE 路由器 102(1) 有关的步骤,但应了解,可相对于 PE 路由器 102(2) (以及网络中的任何其它 PE 路由器) 执行类似工作流。

[0029] 以工作流 300 的步骤 (1) (参考编号 302) 开始, PE 路由器 102(1) 可经配置以在路由器的下行链路端口(即,将路由器连接到 CE 装置 104(1)-(3) 的端口)与路由器和  
管理网络 110 之间的上行链路端口(通向路由服务器 202) 之间实施“交叉连接”。可使用例如应用于下行链路或上行链路端口的一或多个接入控制列表 (ACL) 来实施的交叉连接适于自动地将  
从 CE 装置 104(1)-(3) 始发的路由协议(例如, BGP、OSPF、ISIS 等) 业务转发到路由服务器 202, 而并不在 PE 路由器 102(1) 的控制平面上对所述业务进行本地处理。

[0030] 在步骤 (2) (参考编号 304) 处, PE 路由器 102(1) 可从 CE 装置 104(1)-(3) 中的一或多者接收路由协议控制包且可经由交叉连接将所述包转发到路由服务器 202。作为响应, 路由服务器 202 可接收所述路由协议控制包且可与始发了所述包的 CE 装置建立 / 维持路由协议会话(步骤 (3), 参考编号 306)。例如, 此步骤可包含基于所接收路由协议包中所包含的信息而填充路由数据库、计算最佳路线并构建具有针对各种目的地 IP 地址的路由条目的一或多个路由表。

[0031] 在创建、修改或删除其路由表中的给定路由条目后, 路由服务器 202 可即刻将关于所述路由条目的信息传达到 SDN 控制器 204 (步骤 (4), 参考编号 308)。在某些实施例中, 路由服务器 202 可通过调用由 SDN 控制器 204 出于此目的而公布的 REST API 来执行此传达。在特定实施例中, 所述 API 可经配置以在路由服务器 202 的路由表中注册自身, 借此允许每当存在路由表修改事件(例如, 路由条目创建、更新、删除等) 时就通知(并自动地调用) 所述 API。

[0032] 在步骤 (5) (参考编号 310) 处, SDN 控制器 204 可接收路由条目信息并将其存储于本地流条目数据库中。在一个实施例中, 路由服务器 202 及 / 或 SDN 控制器 204 可使用 fib 聚合算法(例如在 <https://tools.ietf.org/html/draft-uzmi-smalta-01> 处所描述的 SMALTA 算法) 来压缩其相应数据库。此算法的使用可避免对路由服务器 202 及 / 或 SDN 控制器 204 上用于容纳大量路由条目的昂贵硬件的需要。

[0033] 最终, 在步骤 (6) 及 (7) (参考编号 312 及 314) 处, SDN 控制器 204 可将用于安装所创建 / 经修改路由条目的命令(例如, OpenFlow 命令) 发送到 PE 路由器 102(1), 这可致使路由器 102(1) 将所述路由条目编程到路由器的适当硬件转发表(例如, CAM) 中。这



将致使 PE 路由器 102(1) 随后根据新编程的路由条目来以硬件转发从 CE 装置 104(1)-(3) 接收的未来数据业务。应注意, 在一些实施例中, 这可需要使 PE 路由器 102(1) 支持 / 理解 OpenFlow (或由 SDN 控制器 204 在步骤 (6) 处用于传达命令的无论何种南向通信协议)。

#### [0034] 4. 路由服务器高可用性 (HA)

[0035] 关于图 2 中所展示的网络配置的一个潜在缺点是, 路由服务器 202 (其代表 PE 路由器 102(1) 及 102(2) 而集中执行控制平面功能) 是单个故障点; 如果路由服务器 202 失败, 那么整个网络将崩溃, 这是因为 CE 装置 104(1)-(6) 将不能够与路由服务器设置路由协议会话。为了避免此情景, 图 4 描绘具备 SDN 能力的网络 200 的替代实施方案 (展示为网络 400), 所述替代实施方案利用包括多个节点的路由服务器群集 402 而非单个路由服务器机器。在图 4 的特定实例中, 路由服务器群集 402 包含经由层 2 交换器 406 连接到管理网络 110 的两个节点 404(1) 及 404(2), 其中节点 404(2) 为所述群集中的现用节点且节点 404(1) 为所述群集中的备用节点。路由服务器群集 402 的各种节点可使用例如 VRRP 或 VRRP-e, 以在 CE 装置看来为单个服务器 (具有单个虚拟 IP 地址)。当在虚拟 IP 地址处接收到路由协议包时, 所述路由协议包可由现用节点 404(2) 处理。如果现用节点 404(2) 失效, 那么备用节点 404(1) 可从失效的现用节点接管处理职责, 借此确保路由服务器保持可接入及操作。

[0036] 在某些实施例中, 路由服务器群集 402 可执行用于以下操作的新颖工作流程: (1) 自动地使路由协议状态机及路由数据库在现用节点 404(2) 与备用节点 404(1) 之间同步; 及 (2) 通过备用节点 404(1) 将路由协议响应 (即, “发射”) 包发出到 CE 装置。这与其中路由协议发射包始终由现用节点发出的常规 VRRP 实施方案形成对比。对于此工作流程, 备用节点 404(1) 可始终恰当地与现用节点 404(2) 的状态同步, 这可在现用节点发生故障的情况中减少失效恢复时间。

[0037] 图 5 描绘根据一实施例用于在图 4 的网络 400 的背景中执行此 HA 同步的工作流程 500。在工作流程 500 中, 为清晰起见而省略了管理网络 110 及 L2 交换器 406, 但在各种实施例中, 其可被假定为存在且促进包在 PE 路由器 102(1)/102(2) 与路由服务器群集 402 之间的流动。

[0038] 在工作流程 500 的步骤 (1) (参考编号 502) 处, 可给 PE 路由器 102(1) 配置在 PE 路由器 102(1) 的连接到 CE 装置 104(1)-(3) 的下行链路端口与 PE 路由器 102(1) 的连接到路由服务器群集 402 (通过管理网络 110 及 L2 交换器 406) 的上行链路端口之间的交叉连接。此步骤类似于工作流程 300 的步骤 (1), 但涉及将路由协议业务 (经由交叉连接) 重新引导到路由服务器群集 402 的虚拟 IP 地址, 而非是特定路由服务器机器的物理 IP 地址。

[0039] 在步骤 (2) (参考编号 504) 处, PE 路由器 102(1) 可从给定 CE 装置 (例如, 装置 104(1)) 接收初始路由协议包且可使用交叉连接将所述包转发到路由服务器群集 402 的虚拟 IP 地址, 而并不在路由器 102(1) 上对所述包进行本地处理。这致使路由协议包将由群集的现用节点 404(2) 接收到。

[0040] 在步骤 (3) (参考编号 506) 处, 现用节点 404(2) 可处理从 CE 装置 104(1) 始发的路由协议包, 这引起对路由协议状态机的初始化以用于追踪会话建立过程。现用节点 404(2) 可接着经由直接通信信道 (有时称为 “心跳式连接”) 使此状态机与群集中的备用节点 404(1) 同步 (步骤 (4), 参考编号 508)。在一个实施例中, 此直接信道可为以太网连

接。这可致使备用节点 404(1) 接收并本地存储状态机 (步骤 (5), 参考编号 510)。

[0041] 一旦已根据步骤 (4) 及 (5) 使路由协议状态机同步, 备用节点 404(1) (而非现用节点 404(2)) 便可基于经同步状态机而经由 PE 路由器 102(1) 向 CE 装置 104(1) 发出响应 (即, “发射”包) (步骤 (6), 参考编号 512)。此使用备用节点 404(1) 来向 CE 装置 104(1) 发出发射包的步骤有利地确保状态机已在现用节点与备用节点之间恰当地同步。举例来说, 如果备用节点 404(1) 上的状态机未恰当地匹配现用节点 404(2) 上的状态机, 那么由备用节点 404(1) 发出的发射包将为不正确 / 不可靠的, 这将致使 CE 装置 104(1) 对会话进行复位。

[0042] 最终, 一旦已建立路由协议会话且现用节点 404(2) 已填充其路由数据库, 现用节点 404(2) 便可经由同一直信道在周期性基础上使路由数据库与备用节点 404(1) 同步 (步骤 (7) 及 (8), 参考编号 514 及 516)。这可确保从现用节点 404(2) 到备用节点 404(1) 的失效恢复时间是最小的, 因为不需要使备用节点 404(1) 在现用节点 404(2) 发生故障的情况中重新构建路由数据库。

#### [0043] 5. 网络路由器

[0044] 图 6 描绘根据一实施例的示范性网络路由器 600。网络路由器 600 可用于实施例如前述揭示内容中所描述的 PE 路由器 102(1) 及 102(2)。

[0045] 如所展示, 网络路由器 600 包含管理模块 602、组构模块 604 及若干个 I/O 模块 606(1)-606(N)。管理模块 602 表示网络路由器 600 的控制平面且因此包含用于管理 / 控制路由器的操作的一或多个管理 CPU 608。每一管理 CPU 608 均可在存储于相关联存储器 (未展示) 中的软件的控制下进行操作的通用处理器, 例如 PowerPC、Intel、基于 AMD 或 ARM 的处理器。

[0046] 组构模块 604 及 I/O 模块 606(1)-606(N) 共同地表示网络路由器 600 的数据或转发平面。组构模块 604 经配置以互连网络路由器 600 的各种其它模块。每一 I/O 模块 606 可包含由网络路由器 600 用于发送及接收数据包的一或多个输入 / 输出端口 610(1)-610(N)。每一 I/O 模块 606 还可包含包处理器 612。每一包处理器 612 为可做出关于如何处置传入或传出数据包的线速决策的硬件处理组件 (例如, FPGA 或 ASIC)。举例来说, 在各种实施例中, 每一包处理器 612 可包含 (或耦合到) 编程有由路由服务器 202 确定的路由条目的硬件转发表 (例如, CAM), 如前述实施例中所描述。

[0047] 应了解, 网络路由器 600 为说明性的且并不打算限制本发明的实施例。可存在具有比路由器 600 更多或更少组件的许多其它配置。

#### [0048] 5. 计算机系统

[0049] 图 7 描绘根据一实施例的示范性计算机系统 700。计算机系统 700 可用于实施例如前述揭示内容中所描述的路由服务器 202、路由服务器群集节点 404(1)-(2) 及 / 或 SDN 控制器 204。如图 7 中所展示, 计算机系统 700 可包含经由总线子系统 704 与若干个外围装置通信的一或多个处理器 702。这些外围装置可包含存储子系统 706 (包括存储器子系统 708 及文件存储子系统 710)、用户接口输入装置 712、用户接口输出装置 714 及网络接口子系统 716。

[0050] 总线子系统 704 可提供一种用于使计算机系统 700 的各种组件及子系统按照预期彼此通信的机制。虽然总线子系统 704 被示意性地展示为单个总线, 但总线子系统的替代

实施例可利用多个总线。

[0051] 网络接口子系统 716 可充当用于在计算机系统 700 与其它计算装置或网络之间传达数据的接口。网络接口子系统 716 的实施例可包含有线（例如，同轴、双绞线或光纤以太网）及 / 或无线（例如，Wi-Fi、蜂窝式、蓝牙等）接口。

[0052] 用户接口输入装置 712 可包含键盘、指向装置（例如，鼠标、轨迹球、触摸垫等）、扫描仪、条形码扫描仪、并入到显示器中的触摸屏、音频输入装置（例如，话音辨识系统、麦克风等）及其它类型的输入装置。一般来说，术语“输入装置”的使用打算包含用于将信息输入到计算机系统 700 中的所有可能类型的装置及机构。

[0053] 用户接口输出装置 714 可包含显示子系统、打印机、传真机或非视觉显示器（例如音频输出装置等）。所述显示子系统可为阴极射线管（CRT）、平板装置（例如液晶显示器（LCD））或投影装置。一般来说，术语“输出装置”的使用打算包含用于从计算机系统 700 输出信息的所有可能类型的装置及机构。

[0054] 存储子系统 706 可包含存储器子系统 708 及文件 / 磁盘存储子系统 710。子系统 708 及 710 表示可存储提供本文中所描述的各种实施例的功能性的程序代码及 / 或数据的非暂时性计算机可读存储媒体。

[0055] 存储器子系统 708 可包含若干个存储器，包含用于在程序执行期间存储指令及数据的主随机存取存储器（RAM）718 及其中存储固定指令的只读存储器（ROM）720。文件存储子系统 710 可为程序及数据文件提供持久性（即，非易失性）存储且可包含磁性或固态硬盘驱动器、光学驱动器连同相关联可拆卸媒体（例如，CD-ROM、DVD、蓝光（Blu-Ray）等）、可拆卸基于快闪存储器的驱动器或卡及 / 或此项技术中已知的其它类型的存储媒体。

[0056] 应了解，计算机系统 700 为说明性的且并不打算限制本发明的实施例。可存在具有比计算机系统 700 更多或更少组件的许多其它配置。

[0057] 以上说明图解说明本发明的各种实施例连同可如何实施本发明的方面的实例。以上实例及实施例不应被认为是仅有的实施例，且其呈现是为了图解说明由所附权利要求书所界定的本发明的灵活性及优点。举例来说，虽然已关于特定 workflow 及步骤描述了某些实施例，但所属领域的技术人员将明了，本发明的范围并不严格限于所描述的 workflow 及步骤。被描述为顺序的步骤可并行执行，可改变步骤的次序，且可修改、组合、添加或省略步骤。作为另一实例，虽然已使用硬件与软件的特定组合描述了某些实施例，但应认识到，可存在硬件与软件的其他组合，且被描述为以软件实施的特定操作也可以硬件来实施，且反之亦然。

[0058] 因此，应将本说明书及图式视为具有说明性意义而非限制性意义。所属领域的技术人员将明白且可采用其它布置、实施例、实施方案及等效形式，此并不背离所附权利要求书中所陈述的本发明精神及范围。

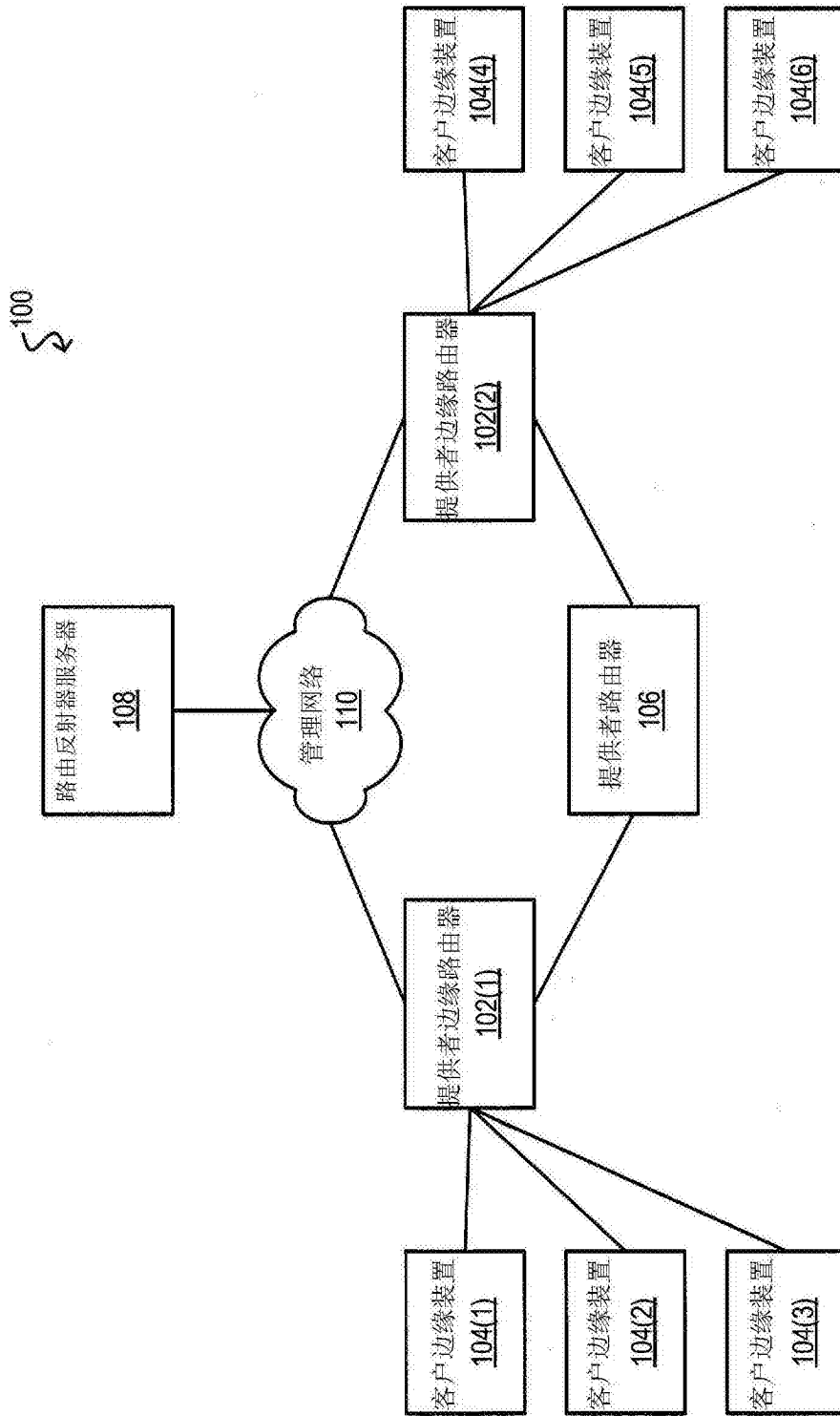


图1(现有技术)

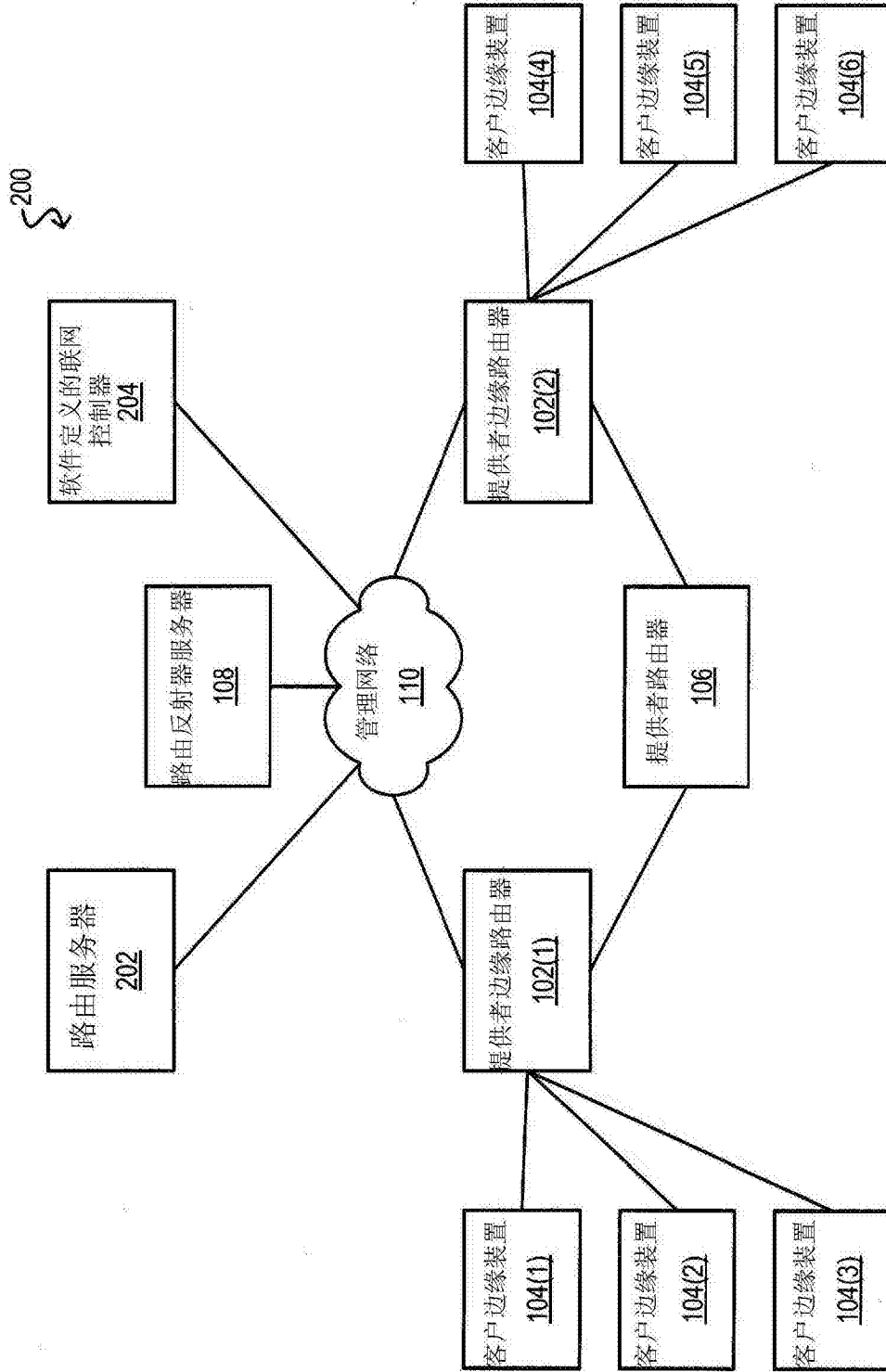


图 2

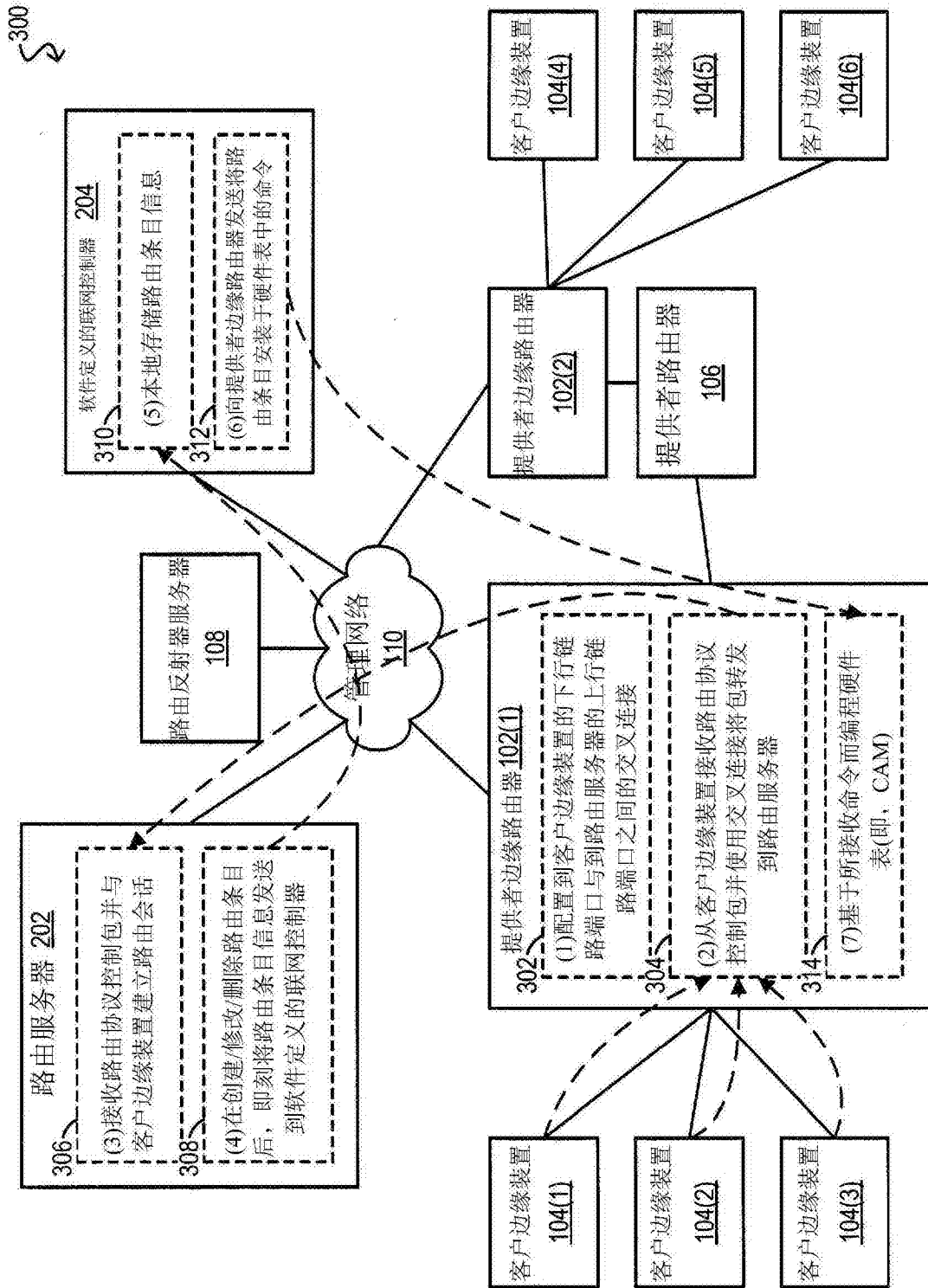


图 3

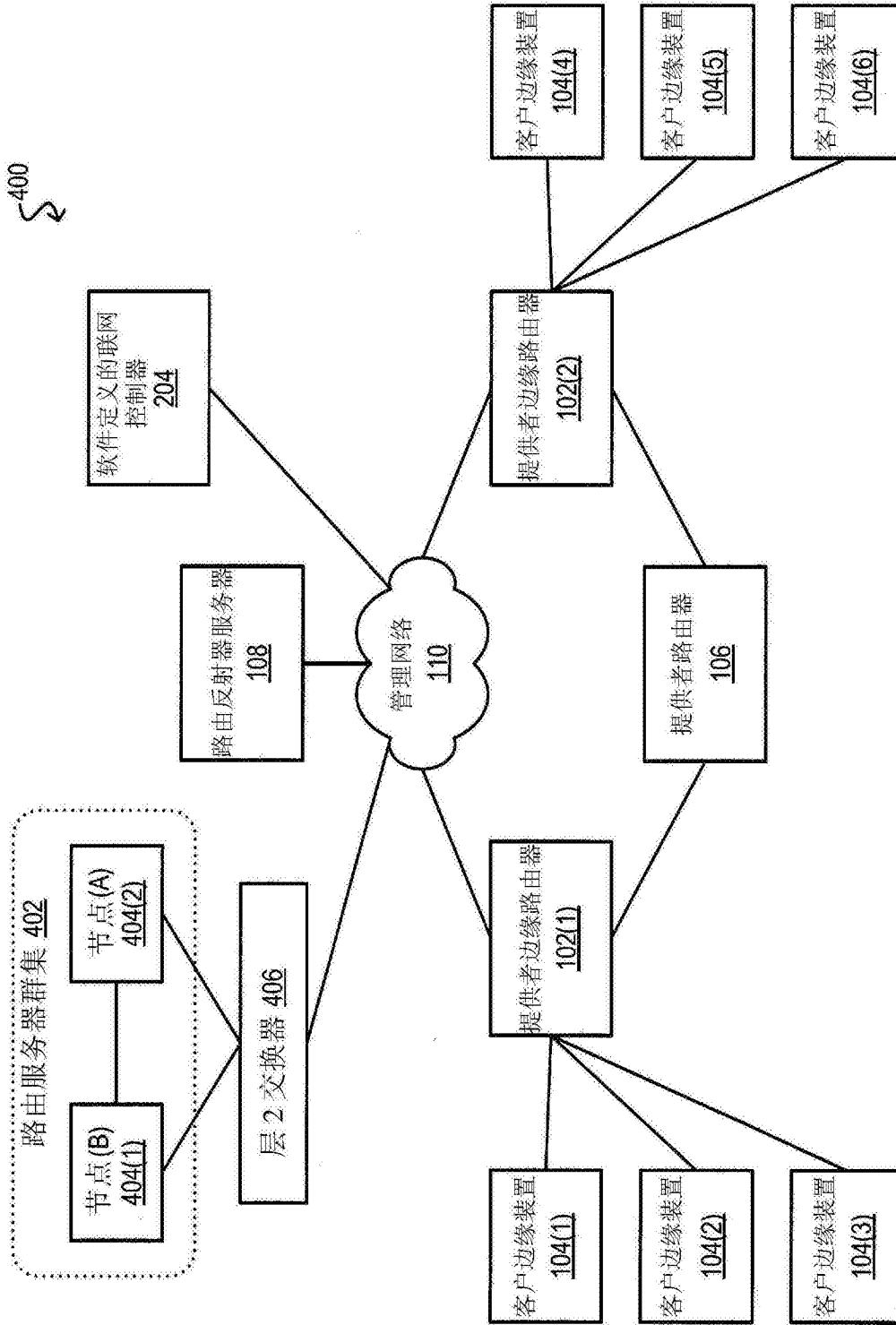


图 4

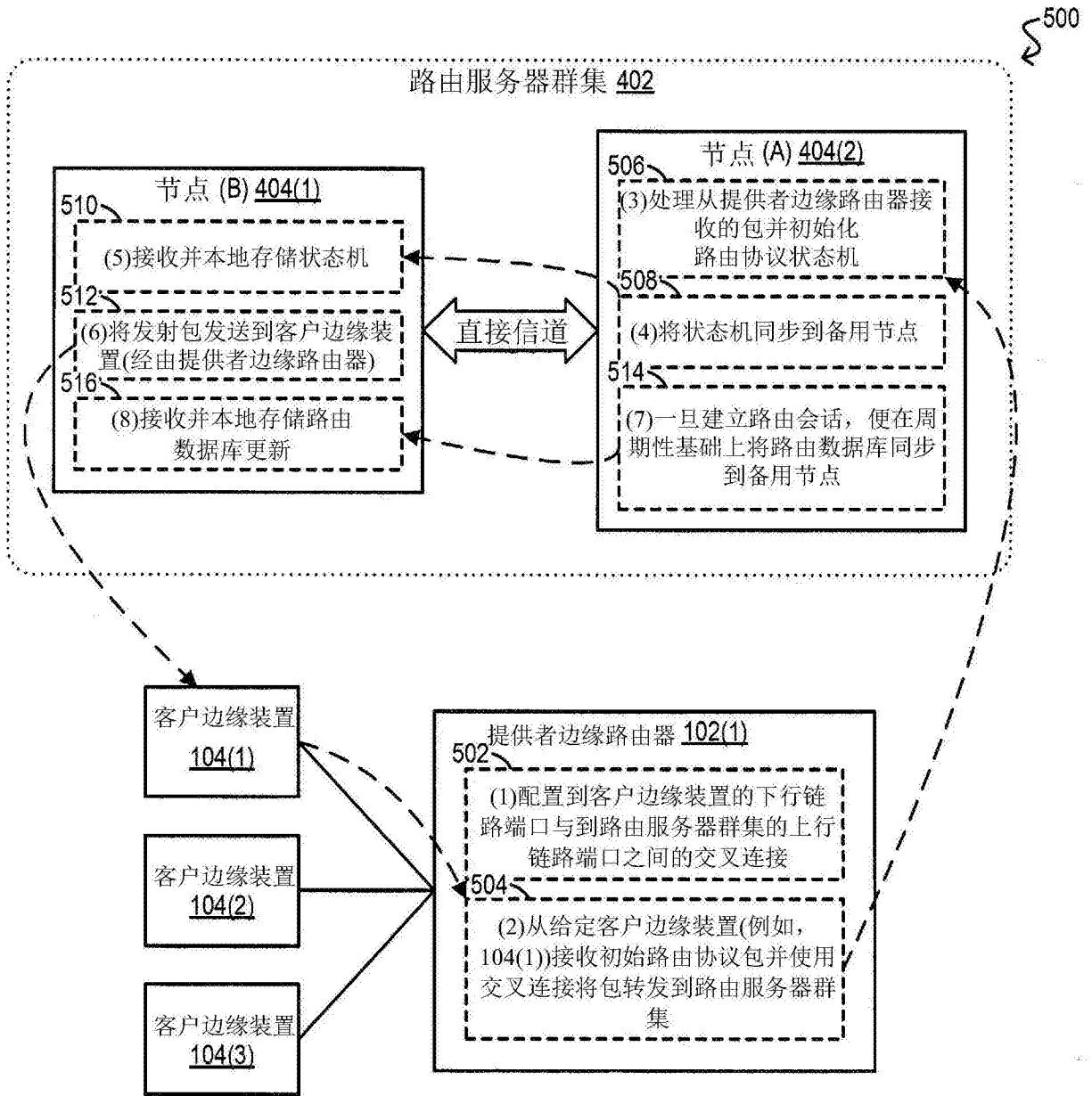


图 5



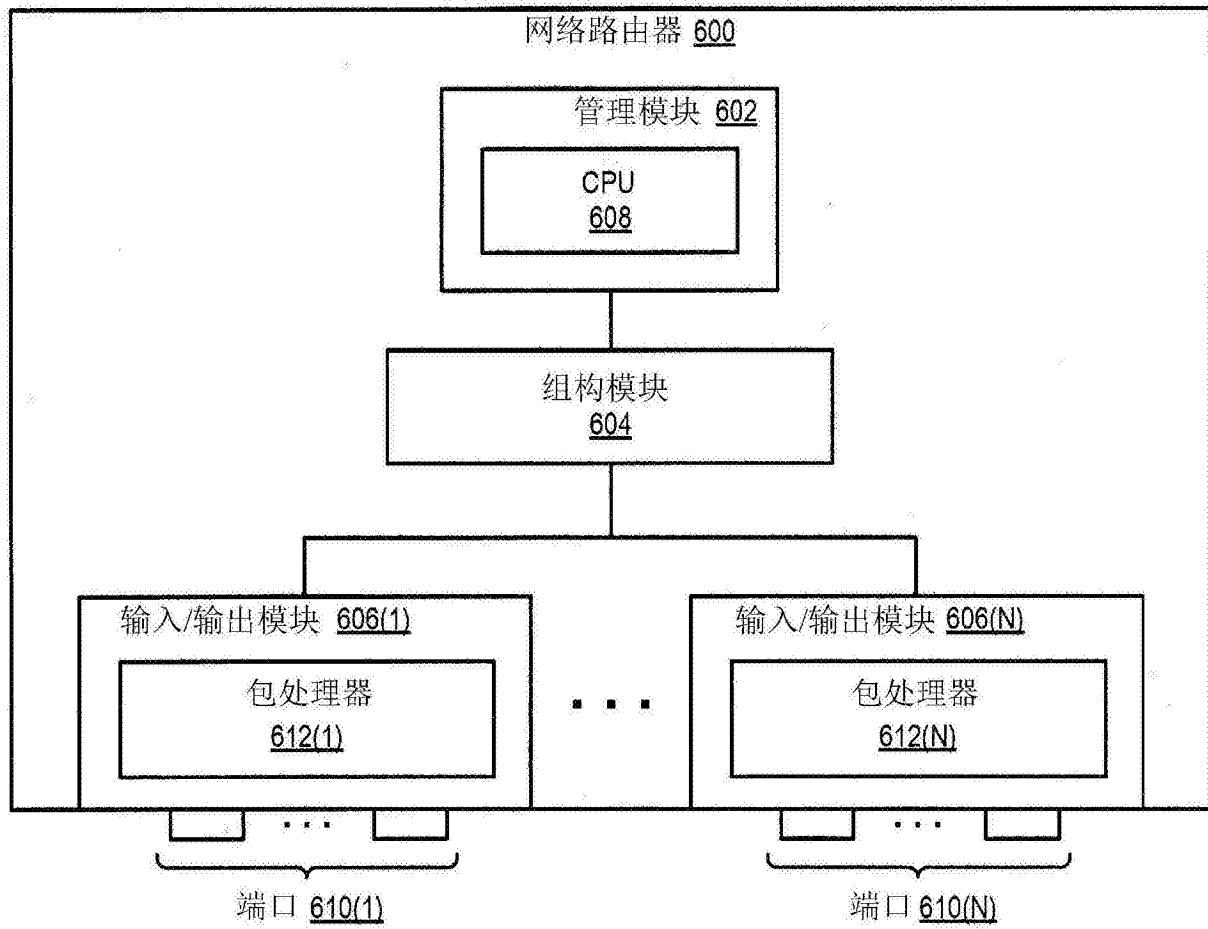


图 6

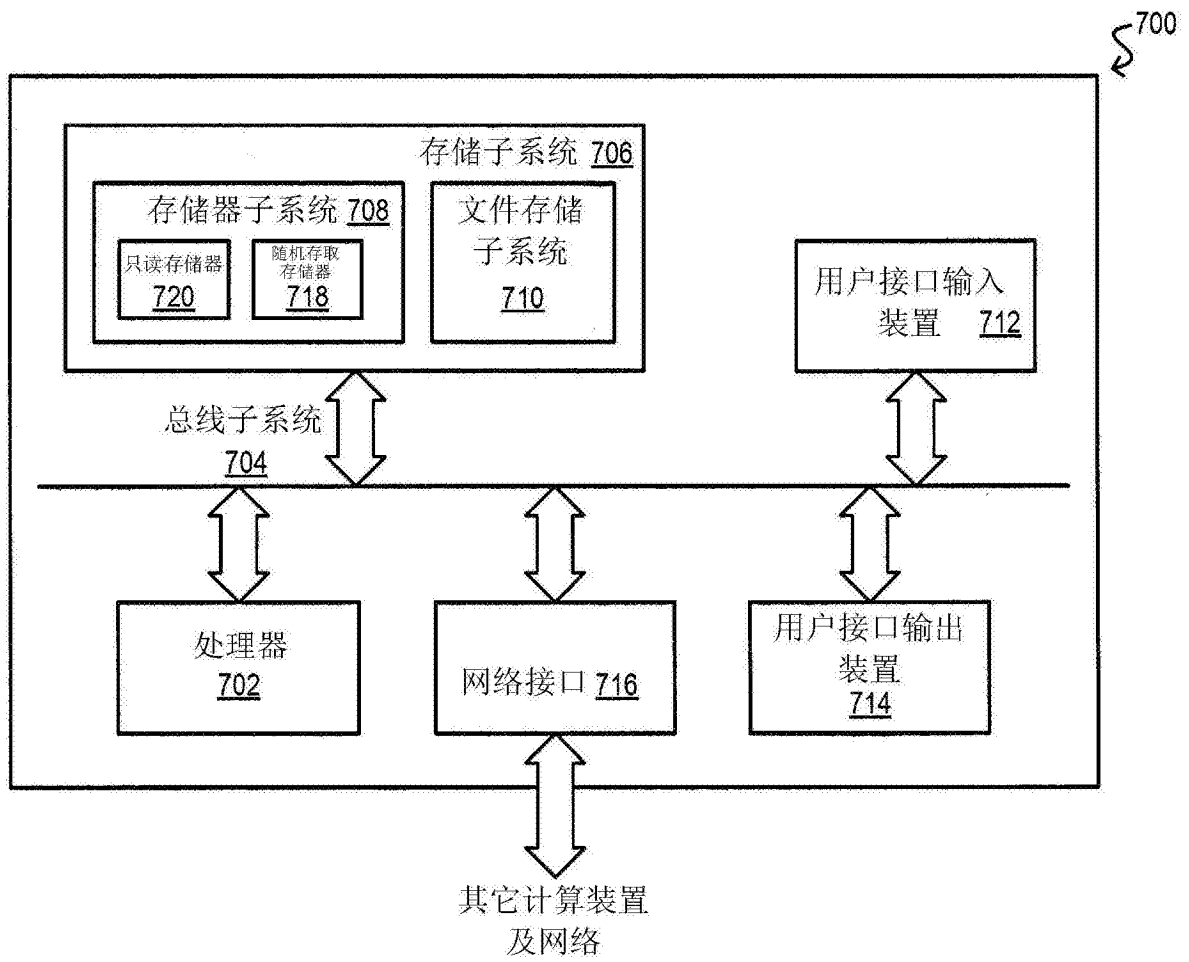


图 7