



## (12)发明专利

(10)授权公告号 CN 104067231 B

(45)授权公告日 2017.03.29

(21)申请号 201280060316.9

(74)专利代理机构 北京安信方达知识产权代理有限公司 11262

(22)申请日 2012.12.06

代理人 陆建萍 郑霞

(65)同一申请的已公布的文献号

申请公布号 CN 104067231 A

(51)Int.Cl.

G06F 9/455(2006.01)

(43)申请公布日 2014.09.24

(56)对比文件

(30)优先权数据

61/568,104 2011.12.07 US

EP 2355423 A1,2011.08.10,

(85)PCT国际申请进入国家阶段日

US 2010014526 A1,2010.01.21,

2014.06.06

US 2011/0090915 A1,2011.04.21,

(86)PCT国际申请的申请数据

PCT/US2012/068263 2012.12.06

Rob Sherwood等.“FlowVisor:A Network

(87)PCT国际申请的公布数据

W02013/086204 EN 2013.06.13

Virtualization Layer”.《<http://openflow.org/downloads/technicalreports/openflow-tr-2009-1-flowvisor.pdf>》.2009,摘要,第3页右侧分栏第3段,第4页右侧分栏第4段,图2.

(73)专利权人 西里克斯系统公司

审查员 曾璇

地址 美国佛罗里达州

(72)发明人 M·波塞尔

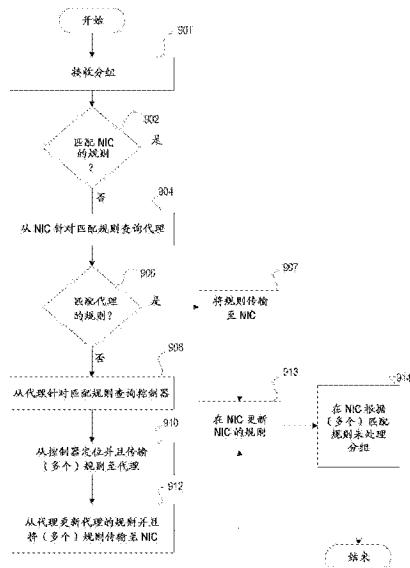
权利要求书2页 说明书22页 附图14页

(54)发明名称

使用虚拟交换代理控制网络接口

(57)摘要

本文中描述了用于允许虚拟化环境内的控制器控制对在硬件内被虚拟化的设备的访问的方法和系统。例如,NIC可以能够请求定义如何在NIC的虚拟交换机处对网络流量进行管理的规则。在一些布置中,NIC可以将针对匹配的规则的查询传输至代理,该代理可以确定其是否具有与该请求相匹配的规则。如果该代理没有匹配的规则,则该代理可以针对匹配的规则对控制器进行查询,该控制器可以将该规则传输至该代理。该代理可以更新其规则存储并且将匹配的规则传输至NIC。当接收到匹配的规则时,NIC可以更新其虚拟交换机中存储的规则并且可以依据该匹配的规则处理分组。



1. 一种用于使用虚拟交换机代理管理网络流量的方法，其中所述方法包括：

由代理使用OpenFlow协议与一个或多个控制器通信，其中所述代理包括定义如何在一个或多个虚拟交换机处对网络流量进行管理的一个或多个规则；

由所述代理与所述一个或多个虚拟交换机通信，其中所述一个或多个虚拟交换机遵从OpenFlow规范；

由所述代理从所述一个或多个虚拟交换机中的一个虚拟交换机接收分组；

由所述代理确定与所述分组相匹配的规则；以及

从所述代理将与所述分组相匹配的所述规则传输至所述虚拟交换机。

2. 根据权利要求1所述的方法，其中由所述代理确定与所述分组相匹配的所述规则包括：

将来自所述代理的规则库的规则与所述分组进行匹配，其中所述一个或多个规则存储在所述规则库中。

3. 根据权利要求1所述的方法，其中由所述代理确定与所述分组相匹配的所述规则包括：

由所述代理针对与所述分组相匹配的所述规则而查询所述一个或多个控制器中的控制器；以及

由所述代理从所述控制器接收与所述分组相匹配的所述规则。

4. 根据权利要求1所述的方法，其中由所述代理与所述一个或多个虚拟交换机的通信使用与所述OpenFlow协议不同的第二OpenFlow协议。

5. 根据权利要求1所述的方法，其中所述一个或多个虚拟交换机中的至少一个虚拟交换机被包括在支持单根输入/输出虚拟化(SR-IOV)的网络接口控制器(NIC)中。

6. 根据权利要求1所述的方法，其中所述代理在交换机域内操作。

7. 根据权利要求1所述的方法，其中所述代理在包括所述一个或多个虚拟交换机的主机内操作。

8. 根据权利要求7所述的方法，进一步包括：

由所述代理与一个或多个第二虚拟交换机通信，其中第二主机包括所述一个或多个第二虚拟交换机。

9. 根据权利要求1所述的方法，其中所述一个或多个虚拟交换机在第一网络部分内操作，并且所述一个或多个控制器在第二网络部分上操作，其中所述第一网络部分和所述第二网络部分互不兼容。

10. 一种用于使用虚拟交换机代理管理网络流量的装置，其中所述装置包括：

用于由代理使用OpenFlow协议与一个或多个控制器通信的模块，其中所述代理包括定义如何在一个或多个虚拟交换机处对网络流量进行管理的一个或多个规则；

用于由所述代理与所述一个或多个虚拟交换机通信的模块，其中所述一个或多个虚拟交换机遵从OpenFlow规范；

用于由所述代理从所述一个或多个虚拟交换机中的一个虚拟交换机接收分组的模块；

用于由所述代理确定与所述分组相匹配的规则的模块；以及

用于从所述代理将与所述分组相匹配的所述规则传输至所述虚拟交换机的模块。

11. 根据权利要求10所述的装置，其中，用于由所述代理确定与所述分组相匹配的规则

的所述模块包括：

用于将来自所述代理的规则库的规则与所述分组进行匹配的模块，其中所述一个或多个规则存储在所述规则库中。

12. 根据权利要求10所述的装置，其中，用于由所述代理确定与所述分组相匹配的规则的所述模块包括：

用于由所述代理针对与所述分组相匹配的所述规则而查询所述一个或多个控制器中的控制器的模块；

用于由所述代理从所述控制器接收与所述分组相匹配的所述规则的模块。

13. 根据权利要求10所述的装置，其中与所述一个或多个虚拟交换机的通信使用与所述OpenFlow协议不同的第二OpenFlow协议。

14. 根据权利要求10所述的装置，其中所述一个或多个虚拟交换机中的至少一个虚拟交换机被包括在支持单根输入/输出虚拟化(SR-IOV)的网络接口控制器(NIC)中。

15. 根据权利要求10所述的装置，其中所述代理在交换机域内操作。

16. 根据权利要求10所述的装置，其中所述一个或多个虚拟交换机在第一网络部分内操作，并且所述一个或多个控制器在第二网络部分上操作，其中所述第一网络部分和所述第二网络部分互不兼容。

17. 一种用于使用虚拟交换机代理管理网络流量的系统，包括：

代理；以及

被配置为遵从OpenFlow规范的一个或多个虚拟交换机；

其中所述代理包括：

用于存储定义如何在一个或多个虚拟交换机处对网络流量进行管理的一个或多个规则的规则库；

用于使用OpenFlow协议与一个或多个控制器通信的模块；

用于与所述一个或多个虚拟交换机通信的模块；

用于从所述一个或多个虚拟交换机中的一个虚拟交换机接收分组的模块；

用于确定与所述分组相匹配的规则的模块；以及

用于将与所述分组相匹配的所述规则传输至所述虚拟交换机的模块。

18. 根据权利要求17所述的系统，其中用于确定与所述分组相匹配的规则的所述模块包括：

用于将来自所述代理的规则库的规则与所述分组进行匹配的模块。

19. 根据权利要求17所述的系统，其中用于确定与所述分组相匹配的规则的所述模块包括：

用于针对与所述分组相匹配的所述规则而查询所述一个或多个控制器中的控制器的模块；以及

用于从所述控制器接收与所述分组相匹配的所述规则的模块。

20. 根据权利要求17所述的系统，其中所述一个或多个虚拟交换机在第一网络部分内操作，并且所述控制器在第二网络部分上操作，其中所述第一网络部分和所述第二网络部分互不兼容。

## 使用虚拟交换代理控制网络接口

[0001] 相关申请的交叉引用

[0002] 本申请是于2011年12月7日提交的题为“CONTROLLING HARDWARE VIRTUALIZED NETWORK INTERFACE CARDS USING VIRTUAL SWITCH PROXYING”的共同未决的美国临时申请序列号61/568,104的非临时申请并且要求其优先权,该申请通过引用全文而结合于此。

### 技术领域

[0003] 本公开内容总体上涉及促进在诸如虚拟化环境的计算环境内的网络通信。例如,一些方面允许在虚拟化环境内的控制器诸如通过使用与该控制器通信的代理以及包括虚拟交换机的网络接口卡而控制对在硬件内被虚拟化的设备的访问。

### 背景技术

[0004] 虚拟化环境通常能够利用单根输入/输出虚拟化(SR-IOV)来虚拟化具有单个端口的物理设备,以便该物理设备表现为具有多个端口。例如,SR-IOV能够被用来虚拟化网络接口卡(NIC),以使得具有单个端口的NIC能够表现为如同其具有多个端口。NIC的这种低级虚拟化提高了虚拟机能够访问并使用NIC的速度并且可以在对网络流量(network traffic)进行隔离和划分的能力之内提供控制虚拟机。虽然在硬件层内虚拟化NIC和其它单端口设备具有优势,但是存在着使得虚拟化的NIC在硬件层被控制而不是被虚拟化环境内的控制虚拟机所控制的缺陷。

### 发明内容

[0005] 以下给出本公开的简化概述,以便于提供对一些方面的基本理解。其并非意在标识出本公开内容的关键或重要要素或者对其范围加以界定。以下概述仅是以简化形式给出本公开的一些概念,作为以下所提供的更为详细的描述的前序,以下描述涉及需要在虚拟化环境内支持SR-IOV,除了其它优势之外,同时允许虚拟化环境内的控制器有从虚拟化环境之内控制对支持SR-IOV的NIC的访问的措施。

[0006] 本文中所描述的各个方面涉及允许虚拟化环境内的控制器控制对硬件内被虚拟化的设备的访问,该设备诸如支持单根I/O虚拟化(SR-IOV)的网络接口卡。在一些布置中,代理,诸如OpenFlow代理,可以使用OpenFlow协议与一个或多个控制器通信。该代理可以包括定义如何在一个或多个虚拟交换机处对网络流量进行管理的规则,该虚拟交换机诸如支持SR-IOV的NIC的虚拟交换机。该代理还与一个或多个虚拟交换机通信,其中每个虚拟交换机遵从OpenFlow规范。

[0007] 用于管理应当如何对网络流量进行管理的规则可以被分布在控制器、NIC和代理之间。例如,网络接口卡可以能够请求定义如何在网络接口卡的虚拟交换机处对网络流量进行管理的规则。在一些布置中,网络接口卡可以将针对匹配的规则的请求传输至代理,该代理可以确定其是否具有与该请求相匹配的规则。如果该代理没有匹配的规则,则该代理可以针对匹配的规则对控制器进行查询,该控制器可以将该规则传输至该代理。该代理可

以更新其规则存储并且将匹配的规则传输至网络接口卡。当接收到匹配的规则时,NIC可以更新其虚拟交换机中存储的规则并且可以依据该匹配的规则而处理分组。

[0008] 通过本文中所描述的各个实施例,以上和其它方面的细节将是显而易见的。

## 附图说明

[0009] 本公开内容在附图中以示例的方式进行图示而并非在其中被限制,附图中同样的附图标记指示相似的要素,并且其中:

[0010] 图1图示了可以在其中实施本公开内容的各个方面示例操作系统。

[0011] 图2图示了可以依据本文中所描述的一个或多个说明性方面而被使用的计算设备。

[0012] 图3图示了可以依据本文中所描述的一个或多个说明性方面而被使用的计算设备。

[0013] 图4是描绘依据本文中所描述的一个或多个说明性方面的虚拟化服务器的实施例的框图。

[0014] 图5A和5B是描绘实施网络策略引擎的虚拟化环境的实施例的框图。

[0015] 图6A-6C是描绘在OpenFlow控制器和OpenFlow客户端之间提供通信的实施例的框图。

[0016] 图7A-7C图示了其中虚拟化环境包括一个或多个网络接口卡的实施例的框图。

[0017] 图8图示了其中虚拟化环境包括SR-IOV NIC、OpenFlow控制器以及用于促进与OpenFlow控制器和SR-IOV NIC通信的OpenFlow代理的实施例的框图。

[0018] 图9图示了其中NIC根据定义如何在一个或多个虚拟交换机处对网络流量进行管理的一个或多个规则来处理所接收分组的示例方法。

## 具体实施方式

[0019] 在各个实施例的以下描述中对附图进行参考,该附图形成其一部分并且在附图中以图示的方式示出了可以对本文中所描述的各方面进行实践的各个实施例。所要理解的是,可以利用其它实施例并且可以做出其它结构和功能修改而并不背离本公开的范围和精神。

[0020] 如本领域技术人员在阅读以下公开时将会认识到的,本文中所描述的各个方面可以被体现为方法、数据处理系统或计算机程序产品。因此,那些方面可以采用完全硬件的实施例、完全软件的实施例或者结合软件和硬件方面的实施例的形式。此外,这样的方面可以采用由一个或多个计算机可读存储媒介所存储的计算机程序产品的形式,具有在存储媒介之中或之上所体现的计算机可读程序代码或可执行指令。可以利用任意适当的计算机可读存储媒介,包括硬盘、CD-ROM、光学存储设备、磁性存储设备和/或前述的任意组合。此外,表示如本文中所描述的数据或事件的各种信号可以以电磁波的形式在源和目的地之间进行传送,该电磁波通过诸如金属线、光纤之类的信号传导媒介和/或无线传输媒介(例如,空气和/或空间)而行进。

[0021] 图1图示了根据本公开内容的一个或多个说明性实施例的可以被使用的示例计算系统100中的通用计算设备101(例如,计算机服务器106a)的示例框图。根据一个或多个方

面,通用计算设备101可以是被配置用于为客户端访问设备提供虚拟机的单服务器或多服务器台式虚拟化系统(例如,云系统)中的服务器106a。通用计算设备101可以具有用于对服务器及其相关联组件的整体操作进行控制的处理器103,上述组件包括随机访问存储器(RAM)105、只读存储器(ROM)107、输入/输出(I/O)模块109和存储器115。

[0022] I/O模块109可以包括通用计算设备101的用户可以通过其提供输入的鼠标、键盘、触摸屏、扫描仪、光学读取器和/或触笔(或(多个)其它输入设备),并且还可以包括用于提供音频输出的扬声器以及用于提供文本、听觉和/或图形输出的视频显示设备中的一个或多个。软件可以被存储在存储器115和/或其它存储内,以向处理器103提供指令而使得通用计算设备101能够执行各种功能。例如,存储器115可以存储由通用计算设备101使用的软件,诸如操作系统119、应用程序119和相关联的数据库121。备选地,用于通用计算设备101的一些或全部计算机可执行指令可以被体现在硬件或固件(未示出)之中。

[0023] 通用计算设备101可以在支持到一个或多个远程计算机的连接的联网环境中进行操作,该远程计算机诸如终端140(也被称为客户端设备)。终端140可以是个人计算机或服务器,其包括以上关于通用计算设备101所描述的许多或全部元件。图1中所描绘的网络连接包括局域网(LAN)125和广域网(WAN)129,但是也可以包括其它网络。当在LAN联网环境中使用时,通用计算设备101可以通过网络接口或适配器123连接至LAN125。当在WAN联网环境中使用时,通用计算设备101可以包括用于通过WAN129、诸如计算机网络130(例如,互联网)而建立通信的调制解调器127或其它网络接口。将要认识到的是,所示出的网络连接是说明性的,并且可以使用在计算机之间建立通信链路的其它手段。

[0024] 通用计算设备101和/或终端140也可以是包括诸如电池、扬声器和天线(未示出)的各种其它组件的移动终端(例如,移动电话、智能电话、PDA、笔记本等)。

[0025] 本公开内容利用多种其它通用或专用计算系统环境或配置进行操作。可适于随本公开内容使用的公知计算系统、环境和/或配置的示例包括但并不局限于个人计算机、服务器计算机、手持或膝上型设备、多处理器系统、基于多处理器的系统、机顶盒、可编程消费者电器、网络PC、小型计算机、大型计算机、包括以上任意系统或设备的分布式计算环境,等等。

[0026] 如图1所示,一个或多个客户端设备140可以与一个或多个服务器106a-106n(在本文中被统称为“(多个)服务器106”)通信。在一个实施例中,计算环境100可以包括被安装在(多个)服务器106和(多个)客户端机器之间的器件。因此,该器件能够管理客户端/服务器连接,并且在一些情况下能够对多个后端服务器106之中的客户端连接进行负载平衡。

[0027] (多个)客户端机器140在一些实施例中能够被称为单客户端机器140或客户端机器140的单个群组,而(多个)服务器106可以被称为单个服务器106或服务器106的单个群组。在一个实施例中,单个客户端机器140与多于一个的服务器106通信,而在另一个实施例中,单个服务器106与多于一个的客户端机器140通信。在又另一个实施例中,单个客户端机器140与单个服务器106通信。

[0028] 在一些实施例中,客户端机器140能够通过以下任意一个术语来引用:(多个)客户端机器140;(多个)客户端;(多个)客户端计算机;(多个)客户端设备;(多个)客户端计算设备;本地机器;远程机器;(多个)客户端节点;(多个)端点;(多个)端点节点;或者第二机器。在一些实施例中,服务器106可以通过以下任意一个术语来引用:(多个)服务器;本地机器;

远程机器;(多个)服务器群;(多个)主机计算设备,或者(多个)第一机器。

[0029] 在一个实施例中,客户端机器140可以是虚拟机。虚拟机可以是任意虚拟机,而在一些实施例中,虚拟机可以是由Citrix Systems、IBM、VMware所以研发的监管程序或者任意其它监管程序所管理的任意虚拟机。在一些方面,虚拟机可以由监管程序进行管理,而在多个方面,虚拟机可以由在服务器106上执行的监管程序或者在客户端140上执行的监管程序进行管理。

[0030] 客户端机器140可以执行、操作或者以其它方式提供应用,该应用可以为以下的任意一个:软件;程序;可执行指令;虚拟机;监管程序;web浏览器;基于web的客户端;客户端-服务器应用;瘦客户端计算客户端;ActiveX控制;Java小应用程序(applet);如软IP电话的与通过互联网协议(VoIP)通信的语言相关的软件;用于流视频和/或音频的应用;用于促进实时数据通信的应用;HTTP客户端;FTP客户端;Oscar客户端;Telnet客户端;或者任意其它可执行指令集合。再其它的实施例包括客户端设备140,其显示服务器106或其它远程定位的机器上远程执行的应用所生成的应用输出。在这些实施例中,客户端设备140能够在应用窗口、浏览器或其它输出窗口中显示应用输出。在一个示例中,该应用是桌面,而在其它示例中,该应用是生成桌面的应用。桌面可以包括为本地和/或远程应用能够在其中进行整合的操作系统实例提供用户界面的图形壳。如本文中所使用的,应用是在操作系统实例(以及可选地还有桌面)已经被加载之后执行的程序。

[0031] 在一些实施例中,服务器106执行远程呈现客户端或者其它客户端或程序,其使用瘦客户端或远程显示协议来采集由服务器106上执行的应用所生成的显示输出并且将应用显示输出传输至远程客户端140。瘦客户端或远程显示协议可以是以下协议中的任意一个:由佛罗里达,劳德代尔堡的Citrix Systems公司制造的独立计算架构(ICA)协议;或者由华盛顿,雷德蒙德的微软公司制造的远程桌面协议(RDP)。

[0032] 计算环境可以包括多于一个的服务器106A-106N,以使得服务器106A-106N例如在云计算环境中在逻辑上被分组为服务器群106。服务器群106可以包括在地理上分散并且在逻辑上一起被分组在服务器群106中的服务器106,或者互相临近并且在逻辑上在服务器群106中分组在一起的服务器106。在一些实施例中,服务器群106内的地理上分散的服务器106A-106N使用WAN、MAN或LAN通信,其中不同地理区域能够被表征为:不同大洲;大洲的不同地区;不同国家;不同州;不同城市;不同校园;不同房间;或者之前的地理位置的任意组合。在一些实施例中,服务器群106可以作为单个实体进行管理,而在其它实施例中,服务器群106可以包括多个服务器群106。

[0033] 在一些实施例中,服务器群106可以包括执行基本上类似类型的操作系统平台(例如,由华盛顿,雷德蒙德的微软公司制造的WINDOWS NT、UNIX、LINUX或SNOW LEOPARD)的服务器106。在其它实施例中,服务器群106可以包括执行第一类型的操作系统平台的第一服务器106群组,以及执行第二类型的操作系统平台的第二服务器106群组。在其它实施例中,服务器群106可以包括执行不同类型的的操作系统平台的服务器106。

[0034] 在一些实施例中,服务器106可以是任意服务器类型。在其它实施例中,服务器106可以是以下的任意服务器类型;文件服务器;应用服务器;web服务器;代理服务器;器件;网络器件;网关;应用网关;网关服务器;虚拟化服务器;部署服务器;SSL VPN服务器;防火墙;web服务器;应用服务器或主应用服务器;执行活动目录的服务器106;或者执行提供防火墙

功能、应用功能或负载平衡功能的应用加速程序的服务器106。在一些实施例中，服务器106可以是RADIUS服务器，其包括远程认证拨入用户服务。在服务器106包括器件的实施例中，服务器106可以是由以下任一个制造商制造的器件：Citrix Application Networking集团；Silver Peak Systems公司；Riverbed Technology公司；F5Networks公司；或者Juniper Networks公司。一些实施例包括第一服务器A，其从客户端机器140接收请求，将这些请求转发至第二服务器106n，并且利用来自第二服务器106n的响应对客户端机器140所生成的请求作出响应。第一服务器106A能够获取客户端机器140可用的应用的列举以及与托管该应用的列举内所标示的应用的应用服务器106相关联的地址信息。第一服务器106A然后能够使用web接口向客户端请求给出响应，并且直接与客户端140通信以向客户端140提供以对所标示应用的访问。

[0035] 在一些实施例中，服务器106能够执行以下应用中的任意一个：使用瘦客户端协议向客户端传输应用显示数据的瘦客户端应用；远程显示呈现应用；Citrix Systems公司的CITRIX ACCESS SUITE的任意部分，如METAFRAME或CITRIX PRESENTATION SERVER；由微软公司制造的MICROSOFT WINDOWS终端服务；或者由Citrix Systems公司研发的ICA客户端。另一个实施例包括作为如下的应用服务器的服务器106，诸如：提供电子邮件服务的电子邮件服务器，诸如由微软公司制造的MICROSOFT EXCHANGE；web或互联网服务器；桌面共享服务器；协同服务器；或者任意其它类型的应用服务器。再其它的实施例包括执行以下类型的托管服务器应用中的任意一种的服务器106：由Citrix Online Division公司提供的GOTOMEETING；由加利福尼亚，圣克拉拉的WebEx公司提供的WEBEX；或者由微软公司提供的Microsoft Office LIVE MEETING。

[0036] 在一些实施例中，客户端机器140可以是寻求访问由服务器106提供的资源的客户端节点。在其它实施例中，服务器106可以为客户端140或客户端节点提供对所托管的资源的访问。在一些实施例中，服务器106用作主节点，以使得其与一个或多个客户端140或服务器106通信。在一些实施例中，该主节点能够识别与托管所请求应用的服务器106相关联的地址信息并将其提供给一个或多个客户端140或服务器106。在再其它的实施例中，主节点可以是服务器群106、客户端140、客户端节点140的集群、或者器件。

[0037] 一个或多个客户端140和/或一个或多个服务器106能够通过被安装在计算环境100内的机器和应用之间的网络130传输数据。网络130可以包括一个或多个子网络，并且可以被安装在包括于计算环境100内的客户端140、服务器106、计算机和器件的任意组合之间。在一些实施例中，网络130可以是：局域网(LAN)；城域网(MAN)；广域网(WAN)；由位于客户端机器140和服务器106之间的多个子网络所组成的主网络104；具有私有子网络的主公共网络130(例如，互联网)；具有公共子网络的主私有网络130；或者具有私有子网络的主私有网络130。再另外的实施例包括可以为以下任意网络类型的网络130：点对点网络；广播网络；电信网络；数据通信网络；计算机网络；ATM(异步传输模式)网络；SONET(同步光纤网络)网络；SDH(同步数字体系)网络；无线网络；有线网络；或者包括无线链路的网络，其中无线链路可以是红外信道或卫星带。网络130的网络拓扑可以在不同实施例内有所不同，可能的网络拓扑包括：总线网络拓扑；星形网络拓扑；环形网络拓扑；基于中继器的网络拓扑；或者层次化星形网络拓扑。另外的实施例可以包括使用协议在移动设备之中进行通信的移动电话网络的网络，其中该协议可以是以下的任意一个：AMPS；TDMA；CDMA；GSM；GPRS UMTS；或

者能够在移动设备之中传输数据的任意其它协议。

[0038] 图2中图示了计算设备200的实施例，其中图1中所图示的客户端机器140和服务器106可以被部署为本文中所图示和描述的计算设备200的任意实施例和/或在其上执行。计算设备200内包括与以下组件通信的系统总线250：中央处理单元221；主存储器222；储存存储器228；输入/输出(I/O)控制器223；显示设备224A-224N；安装设备216；以及网络接口218。在一个实施例中，储存存储器228包括：操作系统、软件例程和客户端代理220。在一些实施例中，I/O控制器223进一步被连接至键盘226和打印设备227。其它实施例可以包括被连接至多于一个的输入/输出设备230A-230N的I/O控制器223。

[0039] 图3图示了计算设备300的一个实施例，其中图1中所图示的客户端机器140和服务器106可以被部署为本文中所图示和描述的计算设备300的任意实施例和/或在其上执行。计算设备300内包括与以下组件通信的系统总线350：桥接器370和第一I/O设备330a。在另一个实施例中，桥接器370与主中央处理单元321通信，其中中央处理单元321可以进一步与第二I/O设备330b、主存储器322和高速缓存存储器340通信。中央处理单元321内包括I/O端口、存储器端口303和主处理器。

[0040] 计算机器300的实施例可以包括以下任一种组件配置为特征的中央处理单元321：对从主存储器单元322取得的指令作出响应并处理的逻辑电路；诸如由Intel公司制造的、由Motorola公司制造的、由加利福尼亚，圣克拉拉的Transmeta公司制造的微处理器单元；诸如由国际商业机器制造的RS/6000处理器；诸如由Advanced Micro Devices制造的处理器；或者逻辑电路的任意其它组合。中央处理单元322的再其它的实施例可以包括以下的任意组合：微处理器、微控制器、具有单个处理核的中央处理单元、具有两个处理核的中央处理单元、或者具有多于一个的处理核的中央处理单元。

[0041] 虽然图3图示了包括单个中央处理单元321的计算设备300，但是在一些实施例中，计算设备300可以包括一个或多个处理单元321。在这些实施例中，计算设备300可以存储并且执行固件或其它可执行指令，这些固件或其它可执行指令当被执行时指导一个或多个处理单元321同步执行指令或者同步执行单片数据上的指令。在其它实施例中，计算设备300可以存储并且执行固件或其它可执行指令，这些固件或其它可执行指令当被执行时指导一个或多个处理单元均执行指令群的一个分段。例如，每个处理单元321可以被指令以执行程序的一部分或者程序内的特定模块。

[0042] 在一些实施例中，处理单元321可以包括一个或多个处理核。例如，处理单元321可以具有两个核、四个核、八个核等。在一个实施例中，处理单元321可以包括一个或多个并行处理核。在一些实施例中，处理单元321的处理核可以访问作为全局地址空间的可用存储器，或者在其它实施例中，计算设备300内的存储器可以被划分并且被分配给处理单元321内的特定核。在一个实施例中，计算设备300中的一个或多个处理核或处理器均可以访问本地存储器。在又另一个实施例中，计算设备300内的存储器可以在一个或多个处理器或处理核之间进行共享，而其它存储器则能够由特定处理器或处理器子集进行访问。在计算设备300包括多于一个的处理单元的实施例中，多个处理单元可以被包括在单个集成电路(IC)中。在一些实施例中，这些多个处理器能够通过内部高速总线链接在一起，该内部高速总线在本文中可以被称为元件互连总线。

[0043] 在计算设备300包括一个或多个处理单元321或者包括具有一个或多个处理核的

处理单元321的实施例中,处理器能够同步执行多片数据上的单个指令(SIMD),或者在其它实施例中,能够同步执行多片数据块上的多个指令(MIMD)。在一些实施例中,计算设备100可以包括任意数量的SIMD或MIMD处理器。

[0044] 在一些实施例中,计算设备300可以包括图形处理器或图形处理单元(未示出)。图形处理单元可以包括软件和硬件的任意组合,并且能够进一步输入图形数据和图形指令,从所输入数据和指令渲染图形,并且输出经渲染的图形。在一些实施例中,图形处理单元可以被包括在处理单元321内。在其它实施例中,计算设备300可以包括一个或多个处理单元321,其中至少一个处理单元321专用于处理并渲染图形。

[0045] 计算机器300的一个实施例包括经由也被称为后方总线的二级总线来与高速缓存存储器340通信的中央处理单元321,而计算机器300的另一个实施例则包括经由系统总线350与高速缓存存储器通信的中央处理单元321。在一些实施例中,本地系统总线350也可以被中央处理单元用来与多于一种类型的I/O设备330a-330n通信。在一些实施例中,本地系统总线350可以是以下任一种类型的总线:VESA VL总线;ISA总线;EISA总线;MicroChannel架构(MCA)总线;PCI总线;PCI-X总线;PCI-Express总线;或者NuBus。计算设备300的其它实施例包括I/O设备330a-330n,它们包括与中央处理单元321通信的视频显示器224。计算机器300再其它的版本包括经由以下任一种连接而连接至I/O设备330a-330n的处理器321:HyperTransport、Rapid I/O或InfiniBand。计算机器300进一步的实施例包括使用本地互连总线与一个I/O设备330a通信并且使用直接连接与第二I/O设备330b通信的处理器321。

[0046] 在一些实施例中,计算设备300包括主存储器单元322和高速缓存存储器340。高速缓存存储器340可以是任意存储器类型,并且在一些实施例中,可以是以下任一种类型的存储器:SRAM;BSRAM或EDRAM。其它实施例包括高速缓存存储器340和主存储器单元322,它们可以是以下任一种类型的存储器:静态随机访问存储器(SRAM)、突发式SRAM或同步突发式SRAM(BSRAM);动态随机访问存储器(DRAM);快速页面模式DRAM(FPM DRAM);增强型DRAM(EDRAM);扩展数据输出RAM(EDO RAM);扩展数据输出DRAM(EDO DRAM);突发式扩展数据输出DRAM(BEDO DRAM);增强型DRAM(EDRAM);同步DRAM(SDRAM);JEDEC SRAM;PC100SDRAM;双倍数据速率SDRAM(DDR SDRAM);增强型SDRAM(ESDRAM);同步链接DRAM(SDRAM);直接Rambus DRAM(DRDARM);铁电RAM(FRAM);或者任意其它类型的存储器。进一步的实施例包括能够经由以下对主存储器322进行访问的中央处理单元321:系统总线350;存储器端口303;或者允许处理器321访问存储器322的任意其它连接、总线或端口。

[0047] 计算设备200/300的一个实施例为以下任一种安装设备216提供支持:CD-ROM驱动器、CD-R/RW驱动器、DVD-ROM驱动器、各种格式的带式驱动器、USB设备、可引导介质、可引导CD、用于诸如KNOPPIX®的GNU/Linux分布的可引导CD、硬盘或者适于安装应用或软件的任意其它设备。在一些实施例中,应用可以包括客户端代理220或者客户端代理220的任意部分。计算设备200/300可以进一步包括存储设备228,后者可以是一个或多个硬盘,或者独立磁盘的一个或多个冗余阵列;其中存储设备被配置为存储操作系统、软件、程序应用、或者客户端代理220的至少一部分。计算设备200、300的进一步的实施例包括用作存储设备228的安装设备216。

[0048] 计算设备200、300可以进一步包括网络接口218以通过各种连接对接至局域网(LAN)、广域网(WAN)或互联网,上述连接包括但并不局限于标准电话线、LAN或WAN链接(例

如,802.11、T1、T3、56kb、X.25、SNA、DECNET)、宽带连接(例如,ISDN、帧中继、ATM、兆比特以太网、Ethernet-over-SONET)、无线连接或者以上任意或全部的一些组合。还能够使用各种通信协议(例如,TCP/IP、IPX、SPX、NetBIOS、以太网、ARCNET、SONET、SDH、光纤分布数据接口(FDDI)、RS232、RS485、IEEE802.11、IEEE802.11a、IEEE802.11b、IEEE802.11g、CDMA、GSM、WiMax和直接异步连接)来建立连接。计算设备200、300的一个版本包括能够经由任意类型和/或形式的网关或隧道协议来与另外的计算设备200'、300'通信的网络接口218,上述网关或隧道协议诸如安全套接层(SSL)或传输层安全(TLS),或者由Citrix Systems公司制造的Citrix网关协议。网络接口218的版本可以包括以下任意一个:内建式网络适配器;网络接口卡;PCMCIA网卡;卡总线网络适配器;无线网络适配器;USB网络适配器;调制解调器;或者适于将计算设备200、300对接至能够通信并执行本文中所描述的方法和系统的网络的任意其它设备。

[0049] 计算设备200、300的实施例包括以下任意一个I/O设备230a-230n:键盘226;指向设备227;鼠标;轨迹板;光学笔;轨迹球;麦克风;画板;视频显示器;扬声器;喷墨打印机;激光打印机;以及热升华打印机;或者能够执行本文中所描述的方法和系统的任意其它输入/输出设备。I/O控制器223在一些实施例中可以连接至多个I/O设备230a-230n,以对一个或多个I/O设备进行控制。I/O设备230a-230n的一些实施例可以被配置为提供存储或安装介质216,而其它则可以提供通用串行总线(USB)接口以便接纳USB存储设备,诸如由Twintech Industry公司制造的设备的USB闪存驱动线路。再其它的实施例包括可以作为系统总线250和外部通信总线之间的桥接器的I/O设备230,外部通信总线诸如:USB总线;苹果桌面总线;RS-232串行连接;SCSI总线;FireWire总线;FireWire800总线;以太网总线;AppleTalk总线;兆比特以太网总线;异步传输模式总线;HIPPI总线;超级HIPPI总线;SerialPlus总线;SCI/LAMP总线;FibreChannel总线;或者串行连接的小型计算机系统接口总线。

[0050] 在一些实施例中,计算机器200、300可以连接至多个显示设备224a-224n,在再其它实施例中,计算设备100能够连接至单个显示设备224,而在再其它的实施例中,计算设备200、300连接至是相同类型或形式的显示器的显示设备224a-224n,或者连接至不同类型或形式的显示设备。显示设备224a-224n的实施例可以被以下所支持或使能:一个或多个I/O设备230a-230n;I/O控制器223;(多个)I/O设备230a-230n与I/O控制器223的组合;能够支持显示设备224a-224n的硬件和软件的任意组合;用于对接、通信、连接或以其它方式使用显示设备224a-224n的任意类型和/或形式的视频适配器、视频卡、驱动器和/或库。计算设备200、300在一些实施例中可以被配置为使用一个或多个显示设备224A-224N,这些配置包括:具有多个连接器对接至多个显示设备224A-224N;具有多个视频适配器,每个视频适配器连接至一个或多个显示设备224A-224N;具有被配置为支持多个显示器224A-224N的操作系统;使用被包括计算设备200内的电路和软件来连接到并且使用多个显示设备224A-224N;以及执行主计算设备200和多个二级计算设备上的软件以使得主计算设备200能够使用二级计算设备的显示器作为主计算设备200的显示设备224A-224N。计算设备200的再其它的实施例可以包括由多个二级计算设备提供并且经由网络连接至主计算设备200的多个显示设备。

[0051] 在一些实施例中,计算设备200能够执行任意操作系统,而在其它实施例中,计算设备200能够执行任意的以下操作系统:多个版本的MICROSOFT WINDOWS操作系统,诸如

WINDOWS3.x;WINDOWS95;WINDOWS98;WINDOWS2000;WINDOWS NT3.51;WINDOWS NT4.0;WINDOWS CE;WINDOWS XP;和WINDOWS VISTA;不同版本的Unix和Linux操作系统;由苹果计算机制造的任意版本的MAC OS;由国际商业机器所制造的OS/2;任意嵌入式操作系统;任意实时操作系统;任意开源操作系统;任意专有操作系统;任意用于移动计算设备的操作系统;或者任意其它操作系统。在再另一个实施例中,计算机器200可以执行多个操作系统。例如,计算机器200能够执行PARALLELS或另一种虚拟化平台,后者能够执行或管理第一操作系统的虚拟机,同时计算机器200执行不同于第一操作系统的第二操作系统。

[0052] 计算机器200可以在任意的以下计算设备中体现;计算工作站;台式计算机;膝上型或笔记本计算机;服务器;手持计算机;移动电话;便携式电信设备;媒体播放设备;游戏系统;移动计算设备;笔记本;由苹果计算机制造的IPOD设备族中的设备;由索尼公司制造的PLAYSTATION设备族中的任意一种;由Nintendo公司制造的Nintendo设备族中的任意一种;由微软公司制造的XBOX设备族中的任意一种;或者能够通信并且具有足够处理器能力和存储器容量来执行本文中所描述的方法和系统的任意其它类型和/或形式的计算、电信或媒介设备。在其它实施例中,计算机器100可以是移动设备,诸如以下任意一种移动设备:支持JAVA的蜂窝电话或个人数字助理(PDA),诸如i55sr、i58sr、i85s、i88s、i90c、i95cl或im1100,所有这些都由Motorola公司制造;由Kyocera制造的6035或7135;由三星电子有限公司制造的i300或i330;由Palm公司制造的TRE0180、270、600、650、680、700p、700w或750智能电话;具有与该设备相符的不同处理器、操作系统和输入设备的任意计算设备;或者能够执行本文中所描述的方法和系统的任意其它移动计算设备。在再其它的实施例中,计算设备200可以是以下任意一种移动计算设备:任意一个系列的Blackberry,或者由Research In Motion有限公司制造的其它手持设备;由苹果计算机制造的iPhone;Palm Pre;便携式PC;便携式PC电话;或者任意其它手持移动设备。

[0053] 在一些实施例中,计算设备200可以具有与该设备相符的不同处理器、操作系统和输入设备。例如,在一个实施例中,计算设备200是由Palm公司制造的TRE0180、270、600、650、680、700p、700w或750智能电话。在这些实施例中的一些中,TRE0智能电话在PalmOS操作系统的控制下进行操作并且包括触笔输入设备以及五向导航设备。

[0054] 在其它实施例中,计算设备200是移动设备,诸如支持JAVA的蜂窝电话或个人数字助理(PDA),诸如i55sr、i58sr、i85s、i88s、i90c、i95cl或im1100,所有这些都由伊利诺伊,绍姆堡的Motorola公司制造;由日本京都的Kyocera制造的6035或7135;由韩国首尔的三星电子有限公司制造的i300或i330。在一些实施例中,计算设备200是由芬兰的Nokia或瑞典Lund的Sony Ericsson移动通信公司制造的移动设备。

[0055] 在再其它的实施例中,计算设备200是Blackberry手持或智能电话,诸如由Research In Motion有限公司制造的设备,包括Blackberry7100系列、8700系列、8700系列、7700系列、7200系列,Blackberry7520或BlackBerry Pearl18100。在又其它的实施例中,计算设备200是智能电话、便携式PC、便携式PC电话或者支持微软Windows移动软件的其它手持移动设备。此外,计算设备200可以是任意工作站、台式计算机、膝上型或笔记本计算机、服务器、手持计算机、移动电话、任意其它计算机,或者能够通信并且具有足够处理器能力和存储器容量以执行本文中所描述的方法和系统的任意其它形式的计算或电信设备。

[0056] 在一些实施例中,计算设备200是数字音频播放器。在这些实施例之一中,计算设

备200是诸如由加利福尼亚,库比蒂诺的苹果计算机制造的苹果IPOD、IPOD Touch、IPOD NANO和IPOD SHUFFLE设备线的数字音频播放器。在这些实施例中的另一个中,数字音频播放器可以用作便携式媒体播放器和大型存储设备。在其它实施例中,计算设备200是如下的数字音频播放器,诸如由NJ的Ridgefield Park的美国三星电子制造的DigitalAudioPlayer Select MP3播放器,或者由IL的Schaumburg的Motorola公司制造的m500或m25数字音频播放器。在再其它的实施例中,计算设备200是便携式媒体播放器,诸如由Creative技术有限公司制造的Zen Vision W、Zen Vision系列、Zen便携式媒体中心设备或者数字MP3线的MP3播放器。在再其它的实施例中,计算设备200是便携式媒介播放器或数字音频播放器,其支持以下文件格式:MP3、WAV、M4A/AAC、WMA Protected AAC、AIFF、可听有声读物、苹果无损音频文件格式以及.mov、.m4v和.mp4/MPEG-4(H.264/MPEG-4AVC)视频文件格式。

[0057] 在一些实施例中,计算设备200包括诸如设备的组合,诸如结合数字音频播放器或便携式媒体播放器的移动电话。在这些实施例之一中,计算设备200是Motorola RAZR或Motorola ROKR线的组合数字音频播放器和移动电话。在这些实施例的另一个中,计算设备200是由加利福尼亚,库比蒂诺的苹果计算机制造的iPhone智能电话。虽然图1-3关于可以使用的硬件和/或软件的具体示例进行了描述,但是这样的示例并非以任何方式进行限制,而仅是作为可以用作技术进展的资源类型的说明。

[0058] 图1-3示出了说明性的桌面虚拟化系统的高级架构。如所示出的,桌面虚拟化系统可以是单个服务器或多个服务器的系统,或者是包括被配置为向一个或多个客户端访问设备140提供虚拟桌面和/或虚拟应用的至少一个虚拟化服务器106的云系统。如本文中所使用的,桌面是指可以在其中托管和/或执行一个或多个应用的图形环境或空间。桌面可以包括为可以在其中集成本地和/或远程应用的操作系统实例提供用户界面的图形壳。如本文中所使用的,应用是在操作系统的实例(以及可选地还有桌面)已经被加载之后执行的程序。每个操作系统实例可以是物理的(例如,每个设备一个操作系统)或虚拟的(例如,在单个设备上运行的许多OS实例)。每个应用可以在本地设备上执行,或者在远程定位的设备(例如,远程的)上执行。

[0059] 图4图示了被配置为虚拟化环境中的虚拟化服务器的计算机设备401的一个实施例,该虚拟化环境例如单服务器、多服务器或云计算环境。图1所示的虚拟化服务器401能够被部署为图1所示的服务器106或图2和3所示的计算设备200和300,和/或由它们的一个或多个实施例所实施。虚拟化服务器401中包括硬件层,其可以包括一个或多个物理磁盘404、一个或多个物理设备406、一个或多个物理处理器408和物理存储器416。在一些实施例中,固件412可以存储在物理存储器416中的存储器元件内并且能够由一个或多个物理处理器408所执行。虚拟化服务器401可以进一步包括可以存储在物理存储器416中的存储器元件中并且由一个或多个物理处理器408所执行的操作系统414。再进一步地,监管程序402可以存储在物理存储器416中的存储器元件中并且能够由一个或多个物理处理器408所执行。一个或多个虚拟机432A-C(统称为432)可以在一个或多个物理处理器408上执行。每个虚拟机432可以具有虚拟盘426A-C以及虚拟处理器428A-C。在一些实施例中,第一虚拟机432A可以在虚拟处理器428A上执行包括工具集(tool stack)424的控制程序420。在其它实施例中,一个或多个虚拟机432B-C能够在虚拟处理器428B-C上执行访客操作系统430A-B。

[0060] 进一步参考图4，并且更为详细地，虚拟化服务器401可以包括硬件层410，其具有与虚拟化服务器401通信的一个或多个硬件。在一些实施例中，硬件层410可以包括一个或多个物理磁盘404、一个或多个物理设备406、一个或多个物理处理器408以及一个或多个存储器416。物理组件404、406、408和416例如可以包括以上在图1-3中所描述的任意组件。例如，物理磁盘404可以永久性存储器存储、临时存储器存储、磁盘驱动器（例如，光盘、软盘、磁带）、硬盘、外部硬盘、闪存、网络连接的存储、存储域网络或者虚拟化服务器401所能够访问的任意其它存储库。物理设备406可以包括虚拟化服务器401中所包括的任意设备、和/或虚拟化服务器401中所包括的设备和与虚拟化服务器401通信的外部设备的任意组合。物理设备406例如可以是网络接口卡、视频卡、键盘、鼠标、输入设备、监视器、显示设备、扬声器、光驱、存储设备、通用串行总线连接、打印机、扫描仪、网络单元（例如，路由器、防火墙、网络地址转译器、负载平衡器、虚拟私人网络（VPN）网关、动态主机配置协议（DHCP）路由器等），或者连接到虚拟化服务器401或者与之进行通信的任意设备。硬件层410中的物理存储器416可以包括任意类型的存储器。物理存储器416可以存储数据，并且在一些实施例中可以存储一个或多个程序或者可执行指令的集合。图4图示了固件412存储在虚拟化服务器401的物理存储器416内的实施例。物理存储器416中所存储的程序或可执行指令能够由虚拟化服务器401的一个或多个处理器408所执行。

[0061] 虚拟化服务器401还可以包括监管程序（hypervisor）402。在一些实施例中，监管程序402可以是由虚拟化服务器401上的处理器408执行的程序，用以创建并管理任意数量的虚拟机432。监管程序402可以被称为虚拟机监视器或者平台虚拟化软件。在一些实施例中，监管程序402可以是对计算机器上执行的虚拟机进行监视的可执行指令和硬件的任意组合。监管程序402可以是2类（Type2）监管程序，或者是在虚拟化服务器401上执行的操作系统414内执行的监管程序。在一些实施例中，2类监管程序在操作系统414的环境内执行而虚拟机在高于监管程序的级别执行。在许多实施例中，2类监管程序在用户的操作系统的环境内执行，以使得2类监管程序与用户的操作系统进行交互。在其它实施例中，虚拟环境中的一个或多个虚拟化服务器401可以包括1类（Type1）监管程序（未示出）。1类监管程序可以通过直接访问硬件层410内的硬件和资源而在虚拟化服务器401上执行。也就是说，在2类监管程序402通过主机操作系统414访问系统资源的同时，1类监管程序可以直接访问所有系统资源而无需主机操作系统414。1类监管程序可以直接在虚拟化服务器401的408的一个或多个物理处理器上执行，并且可以包括存储在物理存储器416中的程序数据。

[0062] 在一些实施例中，监管程序402可以以模仿具有对系统资源的直接访问的操作系统430或控制程序420的任意方式，向在虚拟机432上执行的操作系统430或控制程序420提供虚拟资源。系统资源可以包括：物理设备406；物理磁盘；物理处理器；物理存储器416和虚拟化服务器401硬件层410中所包括的任意其它组件。在这些实施例中，监管程序402可以被用来仿真虚拟硬件、划分物理硬件、虚拟化物理硬件或者执行提供对计算环境的访问的虚拟机。在再其它的实施例中，监管程序402针对虚拟化服务器401上执行的虚拟机432控制处理器调度和存储器分区。监管程序可以包括由加利福尼亚，帕洛阿尔托的VMWare公司制造的监管程序；XEN监管程序，其研发由开源Xen.org团体所监督的开源产品；由微软提供的HyperV、VirtualServer或虚拟PC监管程序，等等。在一些实施例中，虚拟化服务器401执行监管程序402，监管程序402创建可以在其上执行访客操作系统的虚拟机平台。在这些实施

例中,虚拟化服务器401能够被称为主机服务器。这样的虚拟化服务器的示例是由佛罗里达州的劳德代尔堡的Citrix Systems公司提供的XEN SERVER。

[0063] 监管程序402可以创建访客操作系统430在其中执行的一个或多个虚拟机432B-C(统称为432)。在一些实施例中,监管程序402可以加载虚拟机镜像以创建虚拟机432。在其它实施例中,监管程序402可以执行虚拟机432内的访客操作系统430。在再其它的实施例中,虚拟机432可以执行访客操作系统430。

[0064] 除了创建虚拟机432之外,监管程序402可以控制至少一个虚拟机432的执行。在其它实施例中,监管程序402可以为至少一个虚拟机432给出虚拟化服务器401所提供的至少一个硬件资源(例如,在硬件层410内可用的任意硬件资源)的摘要。在其它实施例中,监管程序402可以对虚拟机432访问虚拟化服务器401内可用的物理处理器408的方式进行控制。控制对物理处理器408的访问可以包括确定虚拟机432是否应当具有对处理器408的访问,以及如何给虚拟机432物理处理器的能力。

[0065] 如图4的示例所示,虚拟化服务器401可以托管或执行一个或多个虚拟机432。虚拟机432是可执行指令的集合,其在被处理器408执行时模仿物理计算机的操作,以使得虚拟机432能够像物理计算设备那样执行程序和处理。虽然图4图示了虚拟化服务器401托管三个虚拟机432的实施例,但是在其它实施例中,虚拟化服务器401能够托管任意数量的虚拟机432。在一些实施例中,监管程序402为每个虚拟机432提供物理硬件、存储器、处理器和可供该虚拟机432使用的其它系统资源的唯一虚拟视图。在一些实施例中,该唯一虚拟视图可以基于以下的任一个:虚拟机许可;策略引擎对一个或多个虚拟机标识符的应用;访问虚拟机的用户;虚拟机上执行的应用;虚拟机所访问的网络;或者任意其它类似准则。例如,监管程序402可以创建一个或多个非安全虚拟机432以及一个或多个安全虚拟机432。可以防止非安全的虚拟机432访问安全虚拟机432被允许访问的资源、硬件、存储器位置和程序。在其它实施例中,监管程序402可以为每个虚拟机432提供物理硬件、存储器、处理器和可供该虚拟机432使用的其它系统资源的基本上类似的虚拟视图。

[0066] 每个虚拟机432可以包括虚拟磁盘426A-C(统称为426)和虚拟处理器428A-C(统称为428)。在一些实施例中,虚拟磁盘426是虚拟化服务器401的一个或多个物理磁盘404的虚拟化视图,或者是虚拟化服务器401的一个或多个物理磁盘404的一部分。物理磁盘404的虚拟化视图可以由监管程序402生成、提供并管理。在一些实施例中,监管程序402可以为每个虚拟机432提供物理磁盘404的唯一视图。因此,在这些实施例中,每个虚拟机432中所包括的虚拟磁盘426在与其它虚拟磁盘426相比时可以是唯一的。

[0067] 虚拟处理器428可以是虚拟化服务器401的一个或多个物理处理器408的虚拟化视图。在一些实施例中,物理处理器408的虚拟化视图能够由监管程序402生成、提供并管理。在一些实施例中,虚拟处理器428基本上具有完全与至少一个物理处理器408相同的所有特性。在其它实施例中,虚拟处理器408提供物理处理器408的修改视图,以使得虚拟处理器428的至少一些特性与相应的物理处理器408的特性有所不同。

[0068] 图5A中图示了执行虚拟化环境502并且具有可包括多个网络接口卡(NIC)516A-516N的硬件层510的计算设备501的实施例。虚拟化环境502可以包括诸如控制虚拟机540的控制器、一个或多个附加虚拟机532A-532B和监管程序502。虽然在图5A中示出了1类监管程序(无O/S),但是如以上所描述的,也可以使用2类监管程序。在一些情况下,每个附加虚拟

机532可以执行访客操作系统(访客OS)530A-530B,并且可以包括一个或多个虚拟接口(VIF)对象532A-532N。控制虚拟机540能够执行控制程序520和网络策略引擎522。在一些情况下,控制程序520可以包括工具集524以及一个或多个网络对象535A-535B,这些网络对象包括或实施网络策略520A-520B。

[0069] 进一步参考图5A,并且更为详细地,在一个实施例中,计算设备501可以是本文中所描述的任意计算设备。在一些实施例中,计算设备501可以是服务器或客户端机器。计算设备501可以被称为主机、计算机、第一计算机、第二计算机、第三计算机等。

[0070] 由于计算设备501可以是本文中所描述的任意计算设备,所以硬件层510也可以是本文中所描述的任意硬件层510并且可以包括本文中所描述的任意计算机硬件。在一个实施例中,硬件层510可以包括一个或多个NIC516A-516N。计算设备501可以包括单个NIC516、两个NIC516或者任意其它数量的NIC516。NIC516可以是任意NIC,诸如支持SR-IOV的NIC,并且硬件层510中所包括的每个NIC516可以是相同NIC或不同类型的NIC。

[0071] 在计算设备501上执行的虚拟化环境502可以是本文中所描述的任意虚拟化环境。虚拟化环境502可以包括任意监管程序配置,诸如图4中所图示的监管程序配置。在一个实施例中,虚拟化环境502中所包括的监管程序502可以是任意监管程序502,或者本文中所描述的任意监管程序502。

[0072] 图5A图示了包括虚拟化环境502的计算机器502,该虚拟化环境502包括诸如控制虚拟机540的控制器。该控制器可以是本文中所描述的任意控制器并且可以执行控制程序520,诸如本文中所描述的任意控制程序520。在一些实施例中,控制程序520可以包括工具集524,后者可以包括描述一个或多个网络配置的元数据。虽然图5A图示了包括控制虚拟机540和两个附加虚拟机532A-532B的虚拟化环境302,但是在一些实施例中,虚拟化环境502可以包括任意数量的虚拟机532。虚拟机532可以执行访客操作系统530,诸如本文中所描述的任意访客操作系统530,

[0073] 在一些实施例中,工具集524可以包括描述虚拟化环境502的虚拟机532中的一个或多个虚拟机532所包括的对象的元数据。例如,元数据可以包括与每个虚拟机532的当前配置相关的信息。配置信息可以包括:虚拟机532内的对象所使用的网络配置;虚拟机532内的接口对象的数量和类型;哪些对象当前在虚拟机532中运行;以及当前与虚拟机532对接的外部机器和网络。在一个实施例中,元数据可以包括网络对象535和网络策略520。在一些实施例中,工具集524能够被用来生成虚拟机532用来与其它虚拟机532通信或访问公共网络的网络实施。

[0074] 在一些实施例中,工具集524的元数据可以被存储在控制虚拟机540的存储库中。该元数据可以由虚拟机532以及在虚拟化环境502内执行的对象所生成。在其它实施例中,工具集524的元数据能够从位于主机机器501远程的存储库动态获取。例如,控制程序520能够从为虚拟化系统存储工具集元数据的远程计算机流传输或下载该源数据。在其它实施例中,工具集524的元数据能够使用从远程计算机所下载或流传输的元数据来更新或生成。例如,当控制虚拟机540引导或执行控制程序520时,控制虚拟机540或控制程序520能够为了工具集元数据更新而查询远程计算机。作为响应,远程计算机能够传输、流传输或以其它方式发送与虚拟化环境502相关的控制虚拟机540的元数据。

[0075] 控制程序520能够进一步包括一个或多个网络对象535A-535B。虽然图5A图示了包

括在控制程序520中的网络对象535A-535B(此后统称为网络对象535),但是在其它实施例中,网络对象535能够存储在任意的以下地方:控制虚拟机540;工具集524的元数据;或者位于远程虚拟机532或远程物理机器上的远程存储库。在一些实施例中,网络对象535可以是被配置为连接至一个或多个VIF对象532的对象。在再其它的实施例中,网络对象535可以执行一个或多个网络策略520A-520B。网络对象535能够定义特定网络策略。在另一个实施例中,网络对象535能够定义包括一个或多个网络策略520A-520B的网络配置。因此,当VIF对象532连接至网络对象535时,该VIF对象532能够根据网络对象535内所包括的网络策略或配置进行工作。

[0076] 图5A图示了包括两个网络对象535A-535B的虚拟化环境502。在其它实施例中,虚拟化环境502可以包括单个网络对象535或多个网络对象535。在一些实施例中,网络对象535可以与相同虚拟化环境502内的VIF对象532进行连接,例如包含网络对象535和VIF对象532的相同虚拟化环境502。在其它实施例中,网络对象535可以与和网络对象535位于相同物理机器501上的IVF对象532进行连接,和/或与位于不同于网络对象535所处的物理机器501的物理机器上的VIF对象532进行连接。例如,一个计算设备502上的网络对象535能够与从初始计算设备501远程定位的另一计算设备上的VIF对象532进行连接。

[0077] 在一些实施例中,网络策略520(例如,520A、520B)可以是与特定网络配置相关联的一个或多个策略。例如,网络策略520可以包括指南(或规则)的集合,虚拟机532内执行的对象能够通过该指南(或规则)的集合而与其它虚拟机532内的其它对象进行交互。在一些实施例中,网络策略能够被用来指定能够在一个或多个虚拟机532之间创建的网络的属性和/或定义能够如何对网络流量进行管理。这些属性可以包括:是否及如何对通过连接传输的数据进行加密;哪些NIC卡用来建立连接;是否将网络流量限制于特定物理计算机501;是否避免向特定虚拟机532传输网络流量;网络流量是否能够或应当通过物理NIC516传输至在远程物理计算机上执行的虚拟机532;是否在通过该网络传输之前对数据进行压缩;任意其它网络属性。对网络连接属性进行配置可以进一步包括指定诸如加密算法类型、NIC卡类型之类的子属性以及其它类似的子属性。

[0078] 网络策略520能够以一种或多种策略语言进行编写并且能够由管理员使用与控制程序520对接的配置应用进行配置。在一些实施例中,网络对象535能够实施指定网络配置的一个或多个属性的单个网络策略。在其它实施例中,网络对象535能够实施一个或多个网络策略,后者共同创建指定特定网络的一个或多个属性的网络配置。

[0079] 指定诸如要求网络流量不被传输至特定虚拟机或物理机器或者由特定虚拟机或物理机器所生成的那些属性能够通过将虚拟或物理对象与标签相关联来实现。当创建网络策略520时,策略语言能够参考与对象相关联的标签。例如,标识标签可以被存储在虚拟机532上或者以其它方式与之相关联。当创建指定网络流量不应当传输至带标签的虚拟机的网络策略520时,网络策略520能够使用该标签来识别虚拟机。因此,如果包括网络策略520的网络对象535被移动至不同的虚拟化环境502或计算机501,则网络对象535就无需被重新配置。相反,网络对象535将继续实施防止网络流量被路由或者以其它方式被传输至包含所指定标签的虚拟机的网络配置。因此,当实施网络策略520时,新的虚拟化环境502或计算机501中存储所指定标签或与之相关联的任意虚拟机532将得以被避免而使得其将不会接收到网络流量。

[0080] 网络策略的一个示例可以包括以下：

[0081] ALLOW\_CROSS\_SERVER=true

[0082] REQUIRE\_ENCRYPTION=true

[0083] PHYSICAL\_IFS\_ALLOWED={TAG(testdev\_network1),TAG(testdev\_network2)}

[0084] ISOLATE\_FROM={NAME(my\_insecure\_vm)}

[0085] 以上代码说明了网络策略的一个实施例，其指定了存在于器件上的一个或多个虚拟机能够存在于经由网络连接的单独的物理计算机上。在该实施例中，网络流量应当被加密(例如,REQUIRE\_ENCRYPTION=true)并且业务应当由具有以下标签或者与其相关联的网络接口卡516进行服务：testdev\_network1；和testdev\_network2。另外，该实施例说明了对于网络流量的限制，其中网络流量无法通过与具有标签“my\_insecure\_vm”的虚拟机532或器件相连接的任何网络。

[0086] 在一些实施例中，控制虚拟机540能够执行网络策略引擎522，后者能够被用来执行或实施一个或多个网络策略520。网络策略引擎522能够与一个或多个网络对象535以及网络对象535的一个或多个网络策略520进行对接。这些策略520可以被存储在诸如高速缓存存储器之类的存储库中，并且能够被存储在诸如数据库、列表、表格之类的数据单元或者任意其它存储单元中。在一些实施例中，策略引擎522能够在控制虚拟机540接收到API调用以改变一个或多个虚拟机532的状态时被控制虚拟机540所调用。在其它实施例中，策略引擎522能够由虚拟化环境502所调用或者直接由经历生命周期事件的虚拟机532所调用。在被调用时，网络策略引擎522能够执行一种或多种功能，包括识别对应于经历生命周期事件的虚拟机532中所包括的(多个)VIF对象的一个或多个网络对象535。当网络策略引擎522无法识别包括与虚拟机532的一个或多个VIF对象相匹配的网络配置的网络对象535时，由网络策略引擎522给出指示网络配置无法合成的错误。

[0087] 虽然图5A图示了在控制器上执行的网络策略引擎522，但是在一些实施例中，所示的网络策略引擎522的实例能够在任意虚拟机532上执行。在再其它的实施例中，网络策略引擎522能够在监管程序502上执行。

[0088] 在其它实施例中，每个虚拟机532能够包括一个或多个虚拟接口(VIF)对象532。VIF对象532能够被用来在两个或更多虚拟机532之间建立局部网络。在一些实施例中，VIF对象532能够与网络对象535进行对接而使得网络对象535执行指示连接VIF对象532的网络的属性应当如何的网络策略520。在其它实施例中，一个或多个VIF对象532能够与网络对象535进行注册，以使得网络对象535能够促进与该网络对象进行注册的VIF对象532之间的局部网络的建立。在一些实施例中，VIF对象532能够通过局部网络与其它VIF对象532通信。在其它实施例中，VIF对象532能够与远程物理计算机上的其它VIF对象532通信而使得在VIF对象532之间建立局部网络，然而，该局部网络需要与NIC516进行对接以建立与远程计算机上的那些VIF对象532的局部网络。网络对象535的网络策略或配置能够概括出在VIF对象532之间创建的网络的属性。

[0089] 在一些实施例中，每个虚拟机532能够在出现生命周期事件时生成通知。生命周期事件可以包括：开启或重新开启虚拟机；停止虚拟机；迁移虚拟机；或者其它类似的生命周期事件。该通知可以包括虚拟机532能够传输至控制虚拟机540的警告的生成。当出现生命周期事件时，能够由虚拟化环境502做出是否要求针对该虚拟机532的网络拓扑的决定。该

网络拓扑可以使用网络策略来创建，并且在一些实施例中可以包括在虚拟机532和其它虚拟机之间建立局部网络，或者在虚拟机532和物理网络接口卡516之间建立连接。控制程序520或虚拟化环境502能够依据一个或多个网络策略520来创建网络拓扑。

[0090] 图5B图示了包括一个或多个计算设备501A-501N的虚拟化系统590的一个实施例，以及包括工具集元数据555和网络接口卡516的控制计算设备503。每个计算设备501A-501N可以包括虚拟化环境502，后者包括控制虚拟机540以及一个或多个另外的虚拟机532、一个或多个NIC516，以及进一步包括工具集元数据550的控制虚拟机540。

[0091] 进一步参考图5B，并且更为详细地，在一个实施例中，虚拟化系统590可以是如下的环境，在该环境中执行它们自身的虚拟化环境502的多个计算设备501能够互相通信并且与为每个计算设备501提供对共享资源的访问的控制计算机503通信。在一些实施例中，计算设备501和控制计算机503能够经由诸如本文中所描述的任意网络的网络进行通信。在其它实施例中，计算设备501A-501N以及控制器可以是本文中所描述的任意计算设备。在一些实施例中，计算设备501A-501N能够经由网络和NIC516互相通信。

[0092] 在一些实施例中，控制计算机503能够包括诸如本文中所描述的任意NIC516的NIC516。进一步地，控制计算机503能够经由网络504以及控制计算机503的NIC516而与其它计算机501通信。

[0093] 在一些实施例中，控制计算机503能够进一步包括工具集元数据555，其可以是系统590中的每个计算设备501的元数据555的集合。在一些实施例中，工具集元数据555是每个计算设备501的工具集元数据550的组合。当控制虚拟机540引导(boot)时或者当虚拟环境502执行时，控制虚拟机540或在虚拟化环境502中执行的另一个应用能够从存储在控制计算机503上的元数据555的池中获得元数据550。所获得的元数据550能够存储在请求元数据550的虚拟机532上，并且能够被用来对存储在虚拟机532上的元数据进行更新。当控制虚拟机540停止执行时，元数据550能够被传输至元数据555的池。在一些实施例中，元数据555、550能够存储网络对象535。创建包括在远程物理计算机上执行的虚拟机532的局部网络可以要求使用存储在控制计算机503的工具集元数据555中的集中网络对象535。

[0094] 在一些情况下，能够在一台计算机501A上的一个虚拟机532A与从第一计算机501A远程定位的第二计算机501B上的第二虚拟机532B之间创建局部网络。在这种情况下，可以使用每个虚拟机532A、532B中的VIF对象来创建局部网络。VIF对象能够用网络对象535进行注册，以使得网络对象535能够依据网络对象535的网络策略或配置在VIF对象之间创建局部网络。在该示例中，网络对象535可以是从主元数据存储555所下载的元数据的一部分，或者网络对象535可以被计算设备501A、501B通过计算机501A、501B与控制计算机503之间的网络进行访问。

[0095] 可以对图5A和5B的示例配置进行修改。例如，如以下将要讨论的，可以将代理插入虚拟环境之中，例如通过插入在一个或多个控制器与一个或多个NIC之间。该代理可以被定位而使得其与控制器和NIC通信并且可以被配置为促进控制器与NIC之间的通信。

[0096] 现在将讨论使用NIC(包括SR-IOV NIC)、OpenFlow控制器的示例实施例以及能够与NIC和OpenFlow控制器通信的代理。

[0097] 图6A-6C是描绘在OpenFlow控制器和OpenFlow客户端之间提供通信的实施例的框图。特别地，图6A图示了其中一个或多个OpenFlow控制器可以与OpenFlow客户端通信的实

施例。如图6A所示,OpenFlow控制器601可以与一个或多个OpenFlow客户端602、604和606通信。OpenFlow客户端可以是各种类型的网络设备,诸如OpenFlow交换机或Open vSwitch。OpenFlow控制器601可以用作用于根据任意数量的网络策略而对业务进行路由和优先化的交换机。OpenFlow控制器601还可以用来管理多个虚拟客户端并且在那些虚拟客户端之间实施网络策略,上述虚拟客户端诸如OpenFlow客户端602、604和606。

[0098] OpenFlow是能够被包括在各种网络设备中的标准,上述网络设备包括交换机、路由器和无线访问点。例如,OpenFlow交换机可以将数据路径与控制路径分隔开来。数据路径可以仍然存在于交换机上,而一些路由决策则可以由例如OpenFlow控制器601的单独控制器来控制。OpenFlow交换机,诸如OpenFlow客户端602、604和606中的任意客户端,可以使用一个版本的OpenFlow协议与OpenFlow控制器601通信。

[0099] OpenFlow控制器601可以存储用于对某些类型的分组进行路由的各种规则。例如,控制器可以包括定义应当针对分组执行的各种动作的规则。在一些实施例中,这些规则可以被存储为具有定义分组标识符的字段以及要针对每个分组标识符采取的一个或多个动作的表格。动作可以包括诸如“发送至端口(send-out-port)”、“修改字段(modify-field)”、“丢弃(drop)”等的事情,并且各种动作类型可以在OpenFlow规范中有所定义。

[0100] 图6B图示了其中OpenFlow控制器与OpenFlow客户端通信的实施例的更为详细的描绘。在OpenFlow客户端和OpenFlow控制器之间存在两种主要操作模式。一种操作模式可以是在其中OpenFlow控制器向OpenFlow客户端至少发送规则的子集。第二操作模式可以是在其中OpenFlow客户端从OpenFlow控制器请求一个或多个附加规则。

[0101] 如图6B所示,OpenFlow控制器601包括定义网路业务应当如何进行管理的一个或多个规则612或策略,诸如通过定义在OpenFlow客户端接收到分组时所应当采取的动作。OpenFlow控制器612可以将这些规则612和策略广播至任意OpenFlow客户端,诸如OpenFlow客户端615。进而,OpenFlow客户端可以在用户空间组件614或内核组件620中存储这些规则。图6B图示了具有不同规则存储(规则618和622)的用户空间组件614和内核组件620。虽然规则618和规则622都可以被包括在OpenFlow客户端614上,但是OpenFlow客户端可以在用户空间组件或内核组件中仅包括一个规则。在存储了规则和策略时,OpenFlow客户端614可以使用这些规则和策略来管理网络流量。在一些实施例中,OpenFlow客户端614可以利用新的规则或者利用针对动作的属性的调整而对OpenFlow控制器610所存储的规则进行更新。

[0102] 当诸如示例的OpenFlow客户端614的OpenFlow客户端接收到分组时,客户端可以首先检查其本地规则以确定交换机是否已经知道如何处理分组(检查规则618和/或规则622的匹配)。OpenFlow客户端614可以在交换机的本地规则集合中所存储的规则之一与分组的分组标识符相匹配的情况下确定其能够应对该分组的处理。如果OpenFlow客户端614并没有在本地规则集合中找到匹配,则OpenFlow客户端614可以针对匹配规则而查询控制器,诸如通过向OpenFlow控制器610发送分组。OpenFlow控制器610可以搜索控制器的规则612以确定是否有条目匹配。如果条目在规则612中匹配,则该条目和/或针对该条目而设置的动作可以被发送至OpenFlow客户端614。该分组还可以从OpenFlow控制器610传输至具有规则或条目集合的OpenFlow客户端614。OpenFlow客户端614在接收到动作集合时可以利用与当前分组相匹配的新条目来更新其本地规则,并且可以执行该动作集合以对分组进行适

当处理。动作集合可以包括应当在接收到分组时执行的一个或多个动作。例如，分组可以被修改(减少存活时间字段或改变目的地地址)或者转发至端口。

[0103] 图6C图示了其中OpenFlow代理可以被用来促进与OpenFlow控制器和OpenFlow客户端的通信的实施例。OpenFlow控制器630可以类似于图6B中所描绘的OpenFlow控制器610，并且包括其自身的规则存储(例如，图6B的规则612)。OpenFlow客户端640、642和644中的每个客户端可以类似于图6B中所描绘的示例OpenFlow客户端614并且可以包括其自身的规则存储(例如，图6B的规则618、规则622)。在一些布置中，OpenFlow代理，诸如代理632，能够被用来通过引入冗余虚拟机而提供错误容忍度的度量，该冗余虚拟机能够被用来备份或补充规则和策略并且能够协助促进虚拟机执行各种处理，诸如虚拟机从一个域移动至另一个域。

[0104] 当OpenFlow客户端640、642或644接收到分组时，客户端可以针对与该分组相匹配的条目而搜索其本地规则。如果没有找到匹配，则OpenFlow客户端640、642或644可以将该分组发送至OpenFlow代理632以便查询定义如何处理该分组的规则。

[0105] 与控制器和客户端一样，OpenFlow代理632可以包括定义应当如何在一个或多个OpenFlow客户端处对网络流量进行管理的规则634。因此，当在OpenFlow代理632处接收到分组时，可以搜索规则634以确定条目是否与分组相匹配。

[0106] 在一些实施例中，规则备份638还可以为用于OpenFlow代理632的规则提供二级存储。另外，规则备份638可以是规则634的备份副本，或者规则638可以包括能够由OpenFlow代理在需要时进行检索的附加规则，诸如在OpenFlow代理632确定规则634并不包括该分组的匹配条目时。在这样的情况下，规则638可以被检索或搜索以确定针对该分组是否存在匹配条目。

[0107] 如果OpenFlow代理632找到了匹配规则，则OpenFlow代理可以将(多个)规则或动作集合传输至OpenFlow客户端640、642或644。否则，OpenFlow代理可以通过向OpenFlow控制器630发送分组而针对规则来查询OpenFlow控制器。当OpenFlow控制器630接收到该分组时，可以对控制器本地的规则进行搜索，并且如果在这些规则中找到了匹配的条目，则匹配的规则或事件集合可以被传输至OpenFlow代理632。

[0108] 当从OpenFlow控制器630接收到规则或事件集合时，OpenFlow代理632可以，诸如通过存储所接收的规则或者利用所接收的动作集合在规则634或638中创建新的条目，适当地更新规则634和规则638。OpenFlow代理632还可以将规则或动作集合传输至OpenFlow客户端640、642或644。

[0109] 在OpenFlow客户端640、642或644从OpenFlow代理632接收到规则或动作集合时，诸如可以通过将存储所接收的规则或者利用所接收的动作集合在客户端的规则中创建新的条目而对客户端的规则进行更新。OpenFlow客户端也可以执行该规则或动作集合以对分组进行处理。

[0110] 图7A-7C图示了其中虚拟化环境包括一个或多个网络接口卡的实施例的框图。图7A图示了其中网络接口卡712A和712B与受信任的域703通信的实施例。该虚拟环境还可以包括监管程序710，诸如Xen监管程序或者另一种I类或II类监管程序。在所描绘的实施例中，NIC712A和712B是不支持SR-IOV的。

[0111] 受信任的域703也可以被称为域零(Dom0)，并且可以执行虚拟化环境的管理和监

管功能,以便于执行各种虚拟化功能,诸如执行一个或多个访客域(例如,虚拟机701A和701B)。

[0112] 在所描绘的实施例中,虚拟机701A和701B可以经由Dom0访问网络。在该实施例中,并不要求安装访客驱动器以在虚拟机701A或701B内实现网络功能,并且虚拟化环境可以包括主机交换。例如,虚拟机的接口702A和702B可以沿网络路径与受信任的域703的接口704通信。在一些布置中,接口702A和702B都可以包括NetFront处理,而接口704则可以包括NetBack处理。Open vSwitch705A和705B可以执行从受信任的域703到相应的NIC的切换。如图所示,Open vSwitch705A映射到NIC712A,而Open vSwitch705B则映射到712B。

[0113] 图7B和7C都利用支持SR-IOV的NIC。简单来讲,SR-IOV是一种允许虚拟机与NIC(直接)通信而不必通过控制虚拟机来输送网络流量的技术,上述控制虚拟机诸如虚拟机管理器或Dom0(如图7A所描绘)。SR-IOV允许以硬件而不是软件的网络虚拟化,这具有各种优势和缺陷。例如,使用SR-IOV技术(而不是Dom0中介的联网)可以在每个虚拟机基础上提供带宽和/或延时方面的明显速度提升。对网络流量进行划分或分割的能力可能是SR-IOV的另一优势。例如,SR-IOV能够针对多个用户的划分/分割业务,而这在相同的软件交换机上可能无法进行(例如,参见图7A的Dom0中的单独交换机)。网络流量还可以被划分/分割为多个安全域(例如,用于DMZ中的业务的域,用于具体公司部分的业务的域,以及用于内部业务的域)。

[0114] 通过使用SR-IOV,虚拟机能够绕过虚拟化环境(例如,Citrix XenServer环境)所提供的并且由控制虚拟机所控制的虚拟交换机。

[0115] 然而,在虚拟化环境中使用时,SR-IOV产生了一些困难。例如,当访客在主机之间迁移时,虚拟机在之前主机上所具有的虚拟功能可能并没有被带到新的主机。另外,新的主机可能不支持SR-IOV(由于缺少适当的卡或驱动器)。此外,在诸如云部署的一些使用SR-IOV的实施例中,可能需要向虚拟化环境安装访客驱动器(这可能要求手工安装)。图7B和7C图示了使用SR-IOV但是却表现出以上的一种或多种缺陷等的实施例。图8图示了对使用SR-IOV的布置有所改进的示例实施例。

[0116] 图7B图示了具有支持SR-IOV的NIC720和受信任的域703的虚拟化环境的实施例。虚拟机701A和受信任的域703经由接口702A和704B通信。受信任的域703包括映射到NIC714的Open vSwitch706。与图7A所描绘的实施例不同,包括了支持SR-IOV的NIC720。受信任的域703可以经由物理功能(PF)722与支持SR-IOV的NIC720通信。虚拟机701B也可以与虚拟功能(VF)通信,虚拟功能(VF)诸如与虚拟机701B相映射的VF723B。VF723A可用于被映射到另一虚拟机。

[0117] 如所描绘的,SR-IOV NIC720可以由虚拟机使用PF722经由受信任的域703进行访问,或者直接使用VF723A-B进行访问。受信任的域703用来与NIC714通信的交换机与SR-IOV NIC720通信时可能无法使用(例如,可能无法经由类似于Open vSwitch706的交换机对SR-IOV NIC720进行访问)。针对SR-IOV NIC的直接通信可能需要附加的功能。因此,这样的实施例可能要求(例如,由NIC提供商)为操作系统创建特殊驱动器(例如,Windows驱动器)以对虚拟机的接口进行管理,以便虚拟机701经由受信任的域703与支持SR-IOV的NIC720通信或者直接使用VF723B与SR-IOV NIC720通信。例如,可能需要安装用来管理接口702B的驱动器。在一些布置中,这样的驱动器可以管理NetFront处理(其与Dom0的NetBack通信),包括

虚拟功能驱动器,以促进从虚拟机701B到SR-IOV NIC720的直接通信。

[0118] 图7C图示了包括受信任的域703和另一个域730的虚拟化环境的实施例。在一些布置中,域730可以是网络驱动器虚拟机,或者被连接至支持SR-IOV的NIC720所提供的VF集合(723A&723B)的DomNet(例如,通过使用域730中的VF的相对应映射,诸如VF733A和733B)。在所图示的实施例中,虚拟机701A和701B可以连接至域730,并且可以存在从每个虚拟机701A和701B到VF的直接映射。例如,虚拟机701A可以映射到VF733A和VF723A,并且虚拟机701B可以映射到VF733B和723B。然而,域730的使用可以导致性能相对图7B的实施例有所衰退。多种存储器映射技术可以允许对通过域730的通信进行路由的大量开销有所降低。

[0119] 图7C的实施例还可以扩展为具有其中存在多个支持SR-IOV的NIC的多个域或DomNet(数据的分隔和划分也可以被维持)。与图7B相同,该虚拟化环境也无法与主机交换进行整合而使得虚拟机从一个域迁移至另一个域。

[0120] 图8图示了其中虚拟化环境包括SR-IOV NIC820、OpenFlow控制器850以及用于促进与OpenFlow控制器850和SR-IOV NIC820通信的OpenFlow代理842的实施例。该实施例以及所描绘实施例的变化形式允许虚拟化环境维持支持SR-IOV的NIC的虚拟交换能力(并且因此保持性能增益),同时允许对访问那些NIC的虚拟机的管理进行控制以因此避免以上所描述的一些缺陷。虽然图8描绘了支持SR-IOV的NIC,但是能够使用各种其它的NIC。例如,除了支持SR-IOV的NIC820之外或者作为其替代,可以使用支持多根输入/输出虚拟化的NIC。包括虚拟交换机的其它NIC或者其中OpenFlow代理允许绕过主机的内部交换机的环境是图8所描绘的虚拟化环境的另外的变化形式。

[0121] OpenFlow代理842例如可以处于交换域840或域830(例如,DomNet)中。在这里,虚拟机801A可以与域830通信,后者在支持SR-IOV的NIC820上提供对VF823的访问。在一些情况下,NIC820可以具有如(与VF823和网络860通信的)虚拟交换机826所描绘的虚拟交换能力。在一些布置中,DomNet可以提供良好定义的应用编程接口(API)集合,以允许域被创建为插入式(drop-in)“器件”。在这样的情况下,可能存在可用来允许域的“堆叠”或“分层”以组合功能的能力。在一些配置中,域(例如,803、830和840)、代理842、NIC820和虚拟机(例如,826)可以被包括在同一主机中。

[0122] 对虚拟机所提供的访问的控制可以经由控制面(由虚拟交换机826、OpenFlow代理842和OpenFlow控制器850之间的虚线所示)来提供。如图8所描绘的,OpenFlow代理842包括在主机所提供的交换域840(“DomSwitch”)中。使用该控制面,虚拟交换机826的初始编程允许其成为交换架构的一部分,该主机参与在该部分之中。该主机可以参与到其中OpenFlow控制器850由第三方进行操作的架构之中,该第三方可以远离虚拟化环境的其它设备,包括交换域840、SR-IOV NIC820和虚拟机801A。此外,除了虚拟交换机826之外,OpenFlow代理可以与另一主机的虚拟交换机通信。

[0123] 图9图示了其中NIC根据一个或多个规则对所接收的分组进行处理的示例方法,上述规则定义了如何在一个或多个虚拟交换机处对网络流量进行管理。特别地,该示例方法适于在诸如图8所描绘的环境的虚拟化环境中使用。因此,图9的示例方法将使用图8中所描绘的虚拟化环境的设备进行图示。可以针对图8所描绘的虚拟化环境执行各种修改,例如使用支持MR-IOV的NIC或者取代支持SR-IOV的NIC820的、包括虚拟交换机的另一种NIC。

[0124] 在步骤901,可以在NIC处接收分组。在一些布置中,NIC,可以是图8中支持SR-IOV

的NIC820,可以经由域830以及从图8的虚拟机801A到达VF823。

[0125] 在步骤902,该NIC可以确定NIC所存储的规则是否与分组(或者分组所包括或基于该分组的分组标识符)相匹配。例如,该规则可以由支持SR-IOV的NIC的虚拟交换机826所存储并且可能与以上结合图6A-6C所讨论的规则相同或相似。如果没有在NIC规则中找到匹配,则该方法可以进行至步骤904。如果找到了匹配,则该方法可以进行至步骤914。在一些情况下,虚拟交换机826的规则可以是宽度减小的规则或异常规则的集合。

[0126] 在步骤904,NIC已经确定了无法在NIC本地的规则中找到匹配的规则,并且NIC可以针对用于处理该分组的匹配规则查询代理。例如,当支持SR-IOV的NIC820接收到虚拟交换机826针对其没有匹配的规则的分组时,虚拟交换机826能够通过向包括OpenFlow代理842的交换机840而对其进行查询。

[0127] OpenFlow代理842可以具有其自身的规则库。因此,在步骤906,OpenFlow代理842可以确定NIC所存储的规则是否与(从NIC所接收的)分组相匹配。如果找到了匹配的规则,则该方法可以进行至步骤907。如果没有找到匹配的规则,则该方法可以进行至步骤908。

[0128] 在步骤907,OpenFlow代理已经找到了匹配的规则,并且该代理可以将匹配的规则(或者在多个规则与分组相匹配的一些情况下是多个规则)传输至NIC。例如,OpenFlow代理可以使得交换机840将匹配的规则传输至支持SR-IOV的NIC820的虚拟交换机826。该传输可以包括一个或多个匹配的规则以及该分组。OpenFlow代理与NIC之间的通信可以被加密。

[0129] 在步骤908,OpenFlow代理并没有在其本地规则中找到匹配。因此,OpenFlow代理可以针对匹配的规则向OpenFlow控制器(例如,OpenFlow控制器850)传输查询。在一些情况下,OpenFlow代理所存储的规则可能不同于OpenFlow控制器所存储的规则。例如,本地规则的集合可以存储在OpenFlow代理处,并且由于OpenFlow代理在OpenFlow控制器之前从NIC接收查询,所以OpenFlow代理的规则可以占据优先(例如,可以在查询控制器之前搜索代理的规则)。在一些实施例中,OpenFlow代理与OpenFlow控制器之间的通信可以被加密。加密密钥可以与用于对OpenFlow代理和NIC之间的通信进行加密的密钥相同,然而也可以使用不同的加密密钥。

[0130] 此外,在一些情况下,与OpenFlow控制器的通信链路可能并不可靠或者有时可能断开。通过更可靠的通信链路提供为NIC能够OpenFlow代理,提供了更快且更为可靠的规则检索。此外,OpenFlow代理可以使得错误容忍度有所改善,这是因为只要OpenFlow控制器和OpenFlow代理之间存在链路以及在OpenFlow代理和NIC或虚拟交换机之间存在链路,规则就能够被分布。

[0131] 在一些实施例中,OpenFlow控制器和NIC可能互不兼容。例如,控制器所处的网络部分以及NIC所处的网络部分可能互不兼容。作为另一个示例,OpenFlow控制器可以使用与NIC所使用的不同的OpenFlow协议(例如,它们使用OpenFlow的不同版本或变化形式)。因此,OpenFlow代理可能需要对来自控制器/NIC的通信进行转换以使得其与NIC/控制器兼容。

[0132] 在步骤910,OpenFlow控制器可以定位来自其规则的批评规则,并且一旦找到了匹配的规则(或多个规则),OpenFlow控制器就可以将匹配的规则传输至OpenFlow代理。在一些布置中,该传输可以包括一个或多个匹配的规则以及该分组。

[0133] 在步骤912,OpenFlow代理已经从OpenFlow控制器接收到匹配的规则,并且

OpenFlow代理可以利用该匹配的规则来更新其规则并且将匹配的规则传输至NIC(或者NIC的虚拟交换机)。在一些情况下,去往NIC的传输可以包括一个或多个匹配的规则以及该分组。

[0134] 在一些实施例中,可以使用一个或多个OpenFlow控制器,并且可以由OpenFlow代理针对匹配的规则对每个OpenFlow控制器进行查询。例如,控制器和代理可以以层级进行布置,具有针对该层级中的每个控制器和代理所定义的不同职责。作为使用多个控制器的一个特定示例,第一OpenFlow控制器可以为所有设备提供缺省规则,而第二OpenFlow控制器可以提供诸如IP语音传输(VoIP)之类的规定于设备或协议的规则。在两个控制器都包括匹配的规则的情况下,OpenFlow代理可以从第一OpenFlow控制器接收匹配的规则并且从第二OpenFlow控制器接收匹配的规则。这些规则可以定义不同动作。OpenFlow代理可以能够确定哪个匹配的规则应当具有优先性,以便选择是来自第一OpenFlow控制器的匹配的规则还是来自第二OpenFlow控制器的匹配的规则应当被用来更新代理的规则并且被传输至NIC。在一些实施例中,OpenFlow代理可以基于各种准则来做出该决定,这些准则包括网络的业务水平、一天中的时间或控制器优先级。

[0135] 在步骤913,NIC或者NIC的虚拟交换机已经从OpenFlow代理接收到匹配的规则,并且NIC或虚拟交换机可以利用该匹配的规则来更新其规则。例如,支持SR-IOV的NIC820的虚拟交换机826从交换机842接收到一个或多个匹配的规则,并且可以利用所接收的一个或多个匹配的规则对虚拟交换机826的规则进行更新。

[0136] 在步骤914,NIC或虚拟交换机可能已经利用一个或多个匹配的规则来更新其规则。因此,NIC或虚拟交换机可以继续跟进对该匹配的规则以及该规则所指定的动作对分组进行处理。动作例如可以包括用于修改该分组的动作,诸如通过缩减存活时间的字段、改变目的地地址。动作还可以包括用于传输或丢弃分组的动作,诸如通过将该分组转发至端口。

[0137] 虽然图9将示例方法的步骤图示为以特定顺序进行,但是可以对步骤的顺序进行修改。例如,在一些实施例中,NIC或虚拟交换机可以在执行步骤913之前执行步骤914。

[0138] 已经在其说明性实施例的方面度本公开内容的各方面进行了描述。本领域技术人员通过回顾本公开内容将认识到处于所附权利要求的范围和精神之内的多种其它实施例、修改和变化。例如,本领域技术人员将会认识到,说明性附图中所图示的步骤可以与所引用顺序不同的顺序来执行,并且所图示的一个或多个步骤依据本公开内容的各方面而可以是可选的。本领域技术人员特别是可以鉴于前述教导来进行修改。例如,以上所提到的实施例的每个要素可以被单独利用,或者以与其它实施例的要素的组合或子组合而得以利用。还要认识到和理解的是,可以在不背离以下权利要求的精神和范围的情况下进行修改。

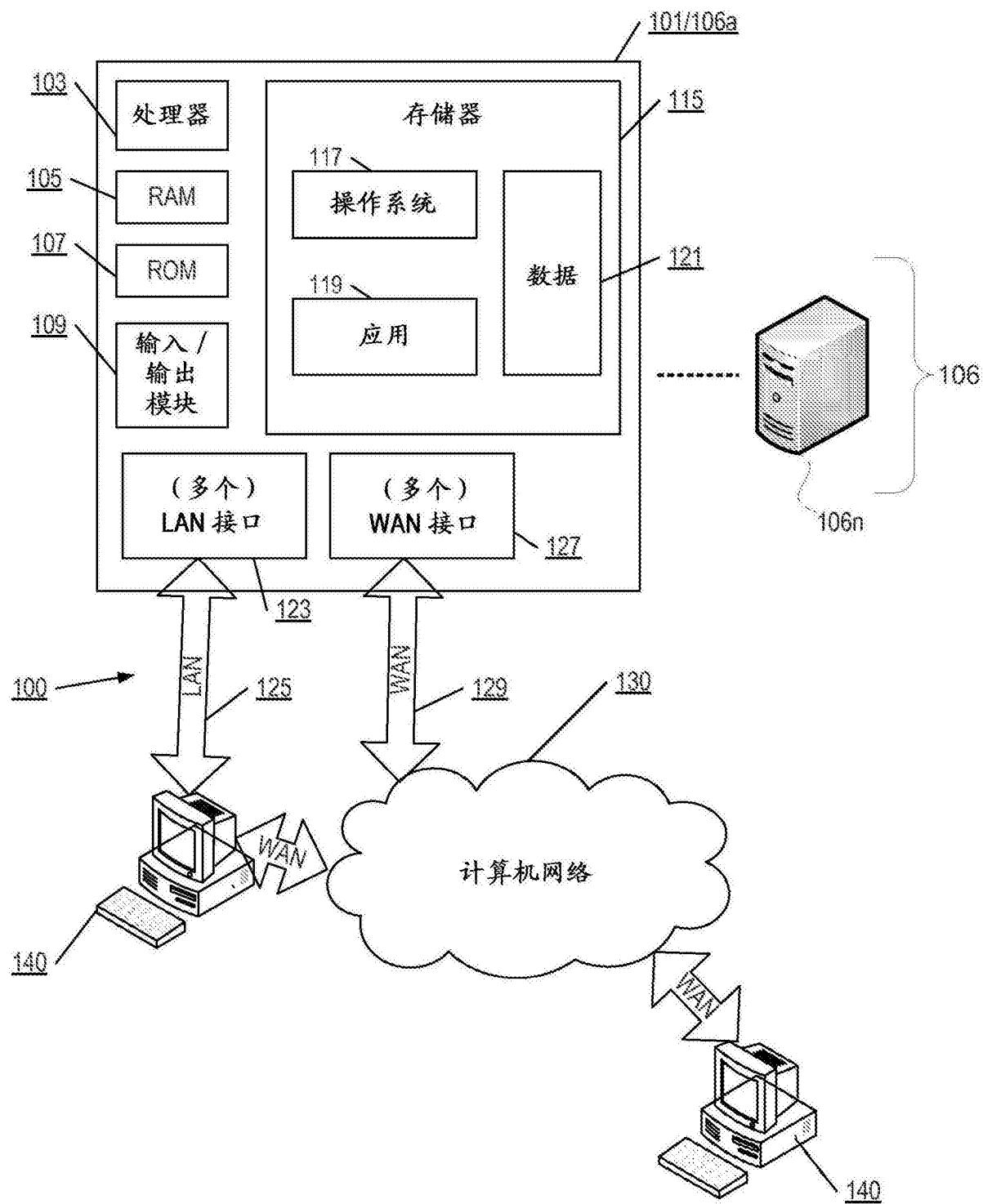


图1

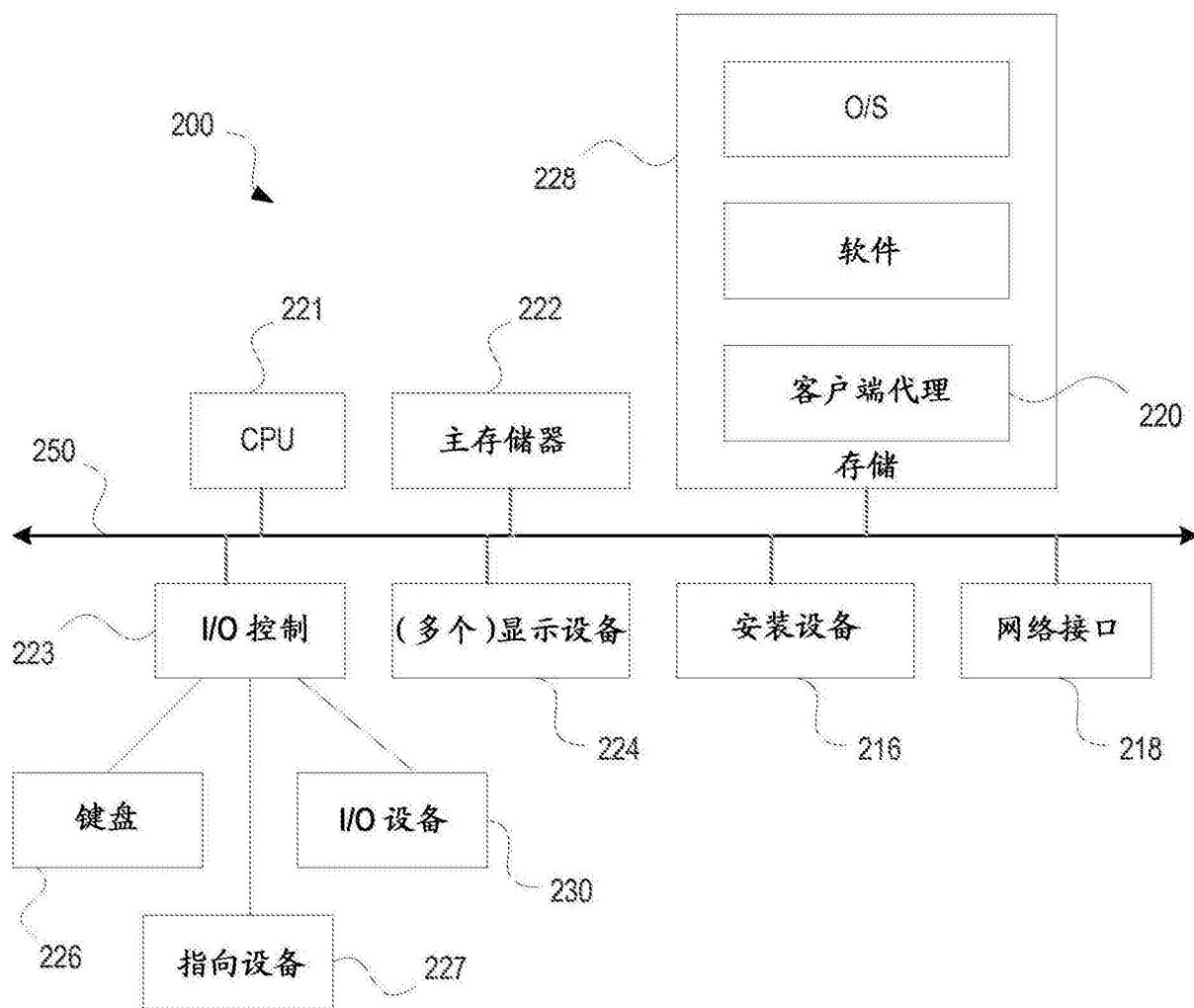


图2

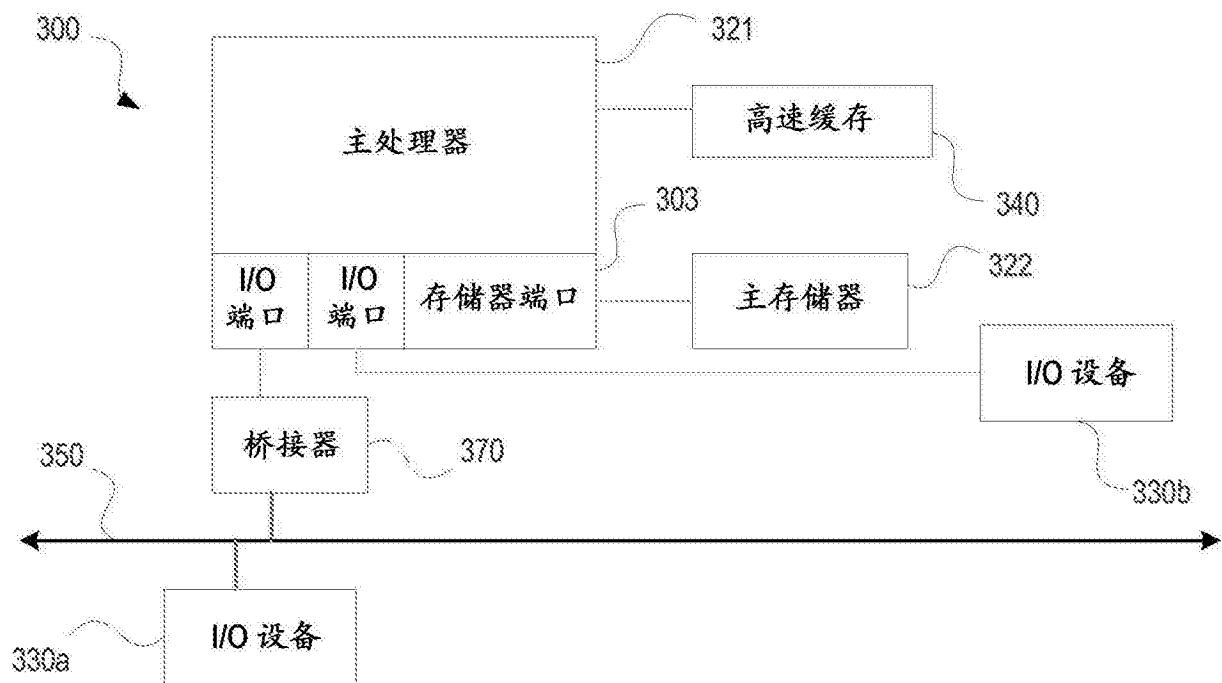


图3

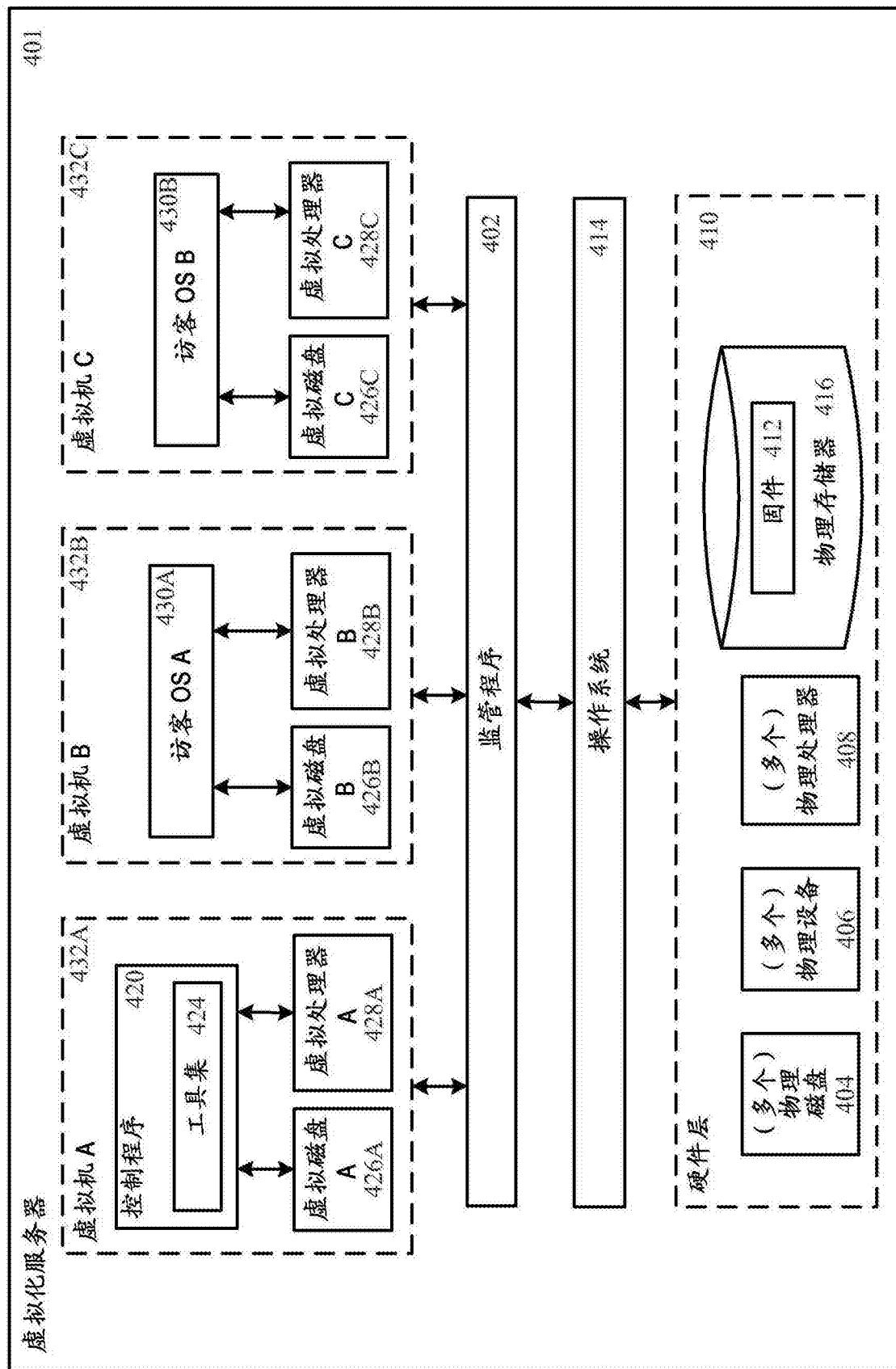


图4

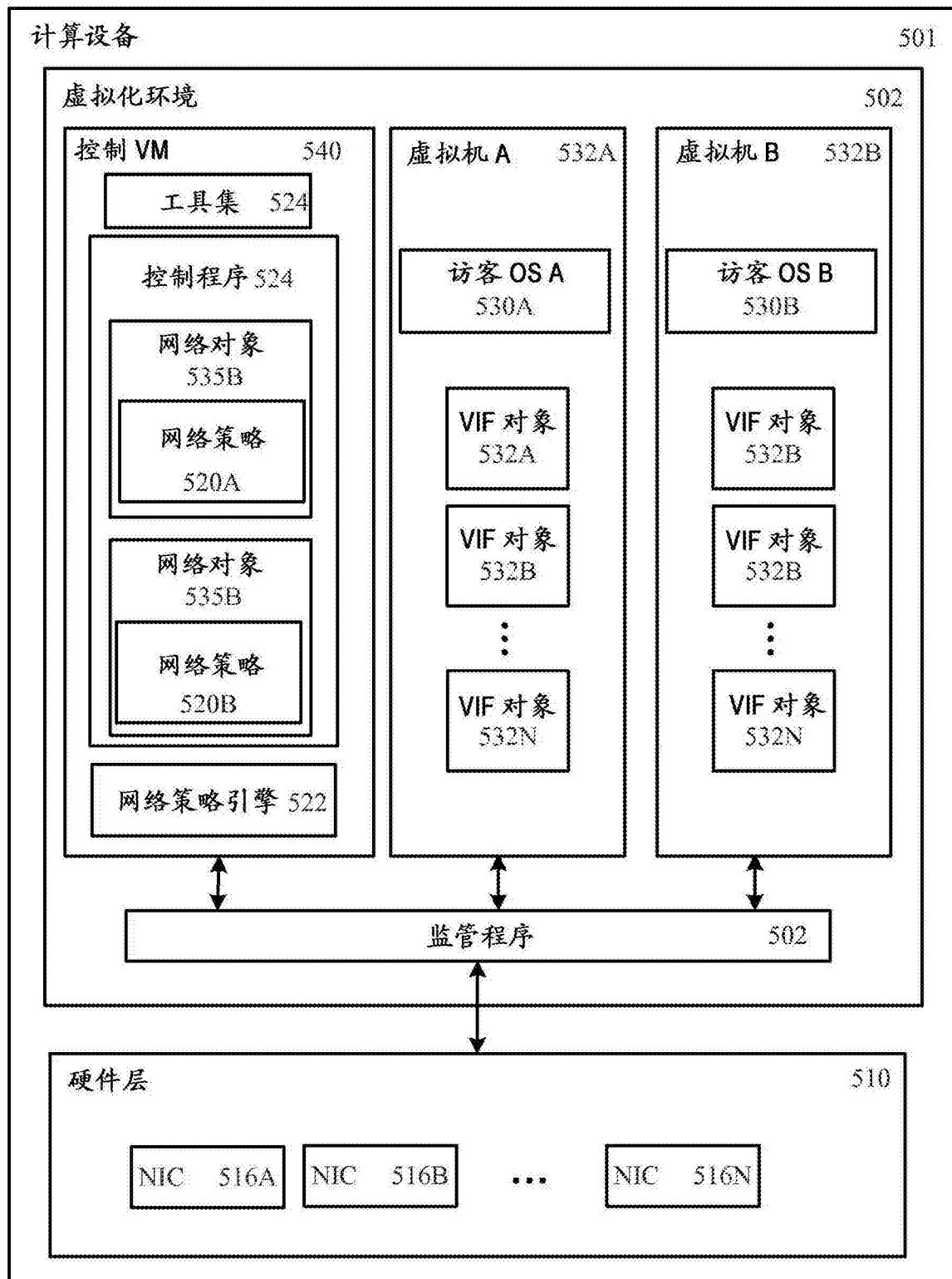


图5A

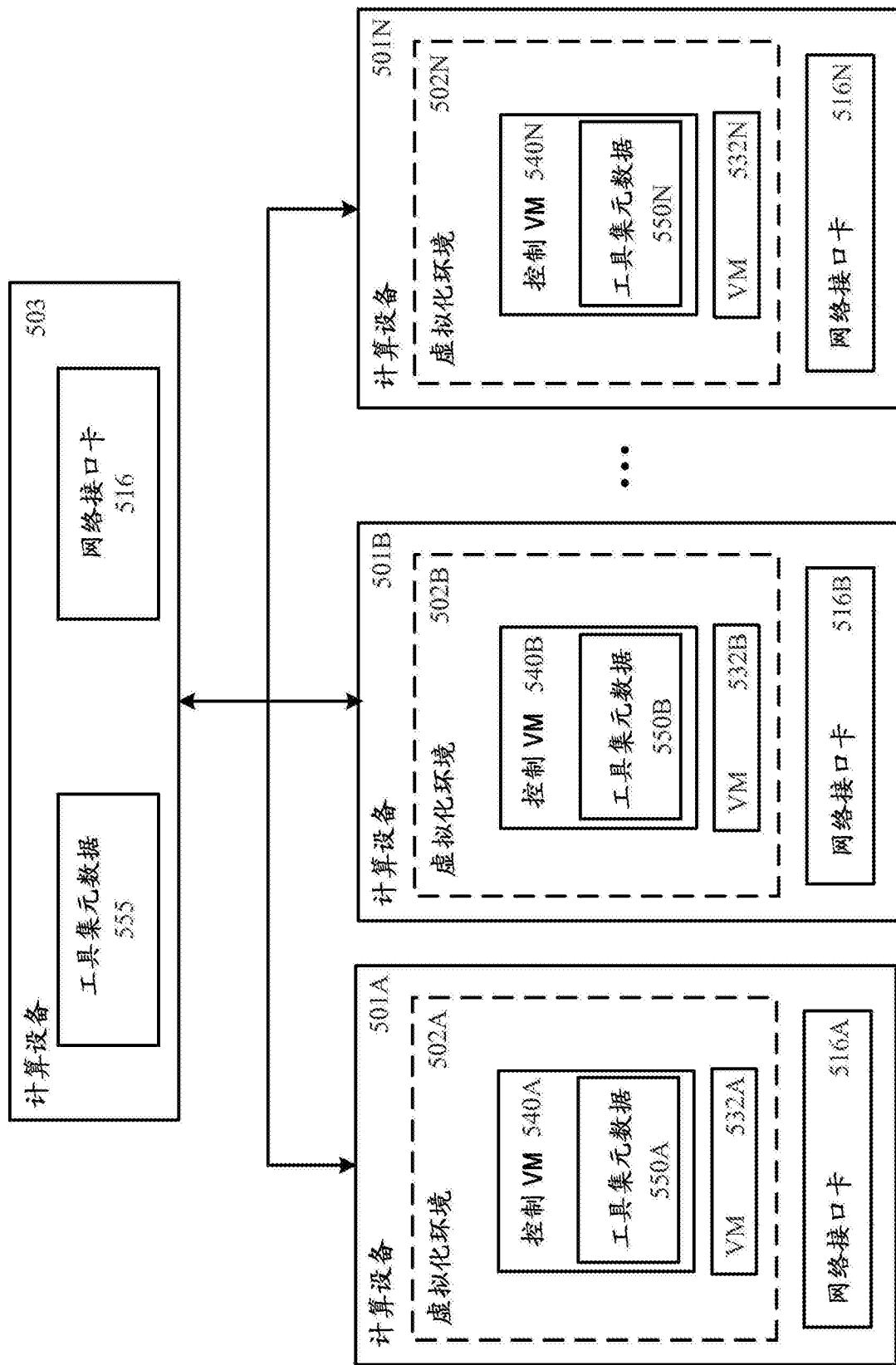


图5B

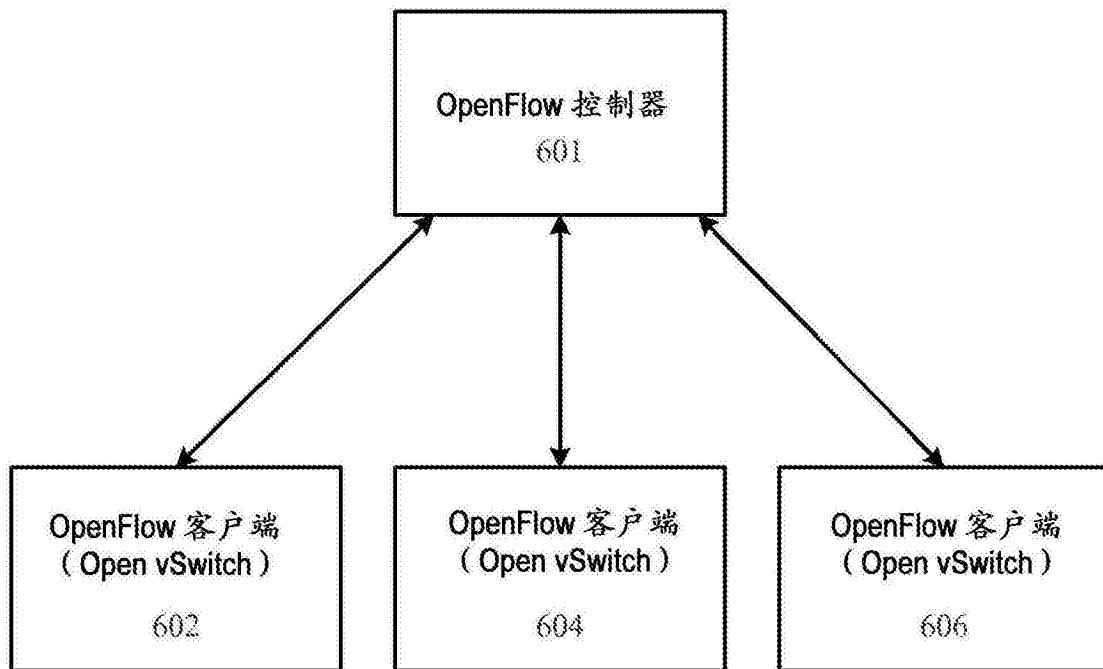


图6A

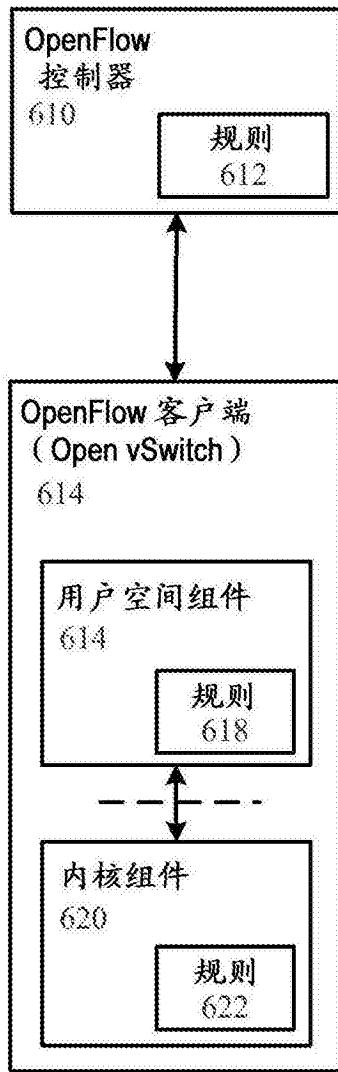


图6B

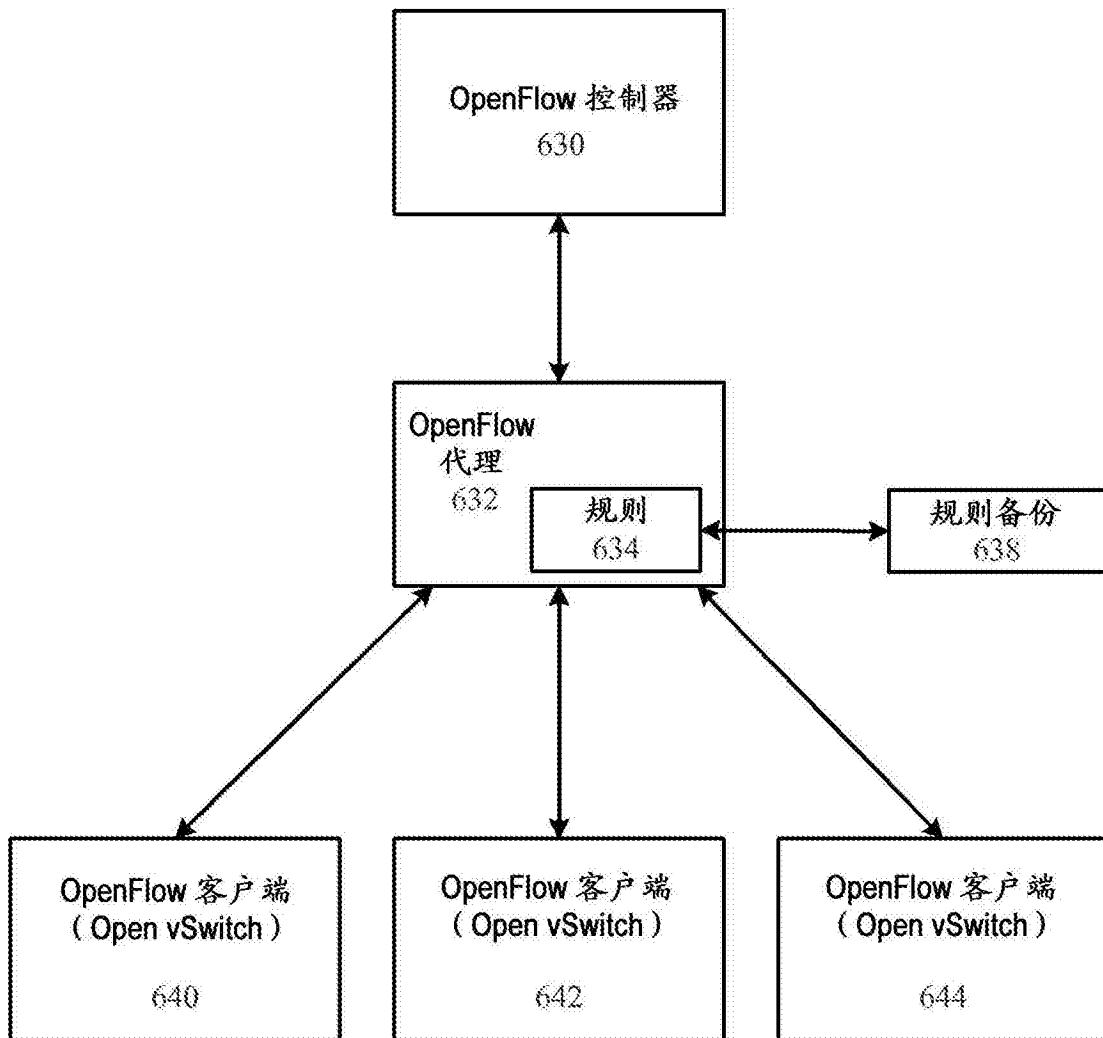


图6C

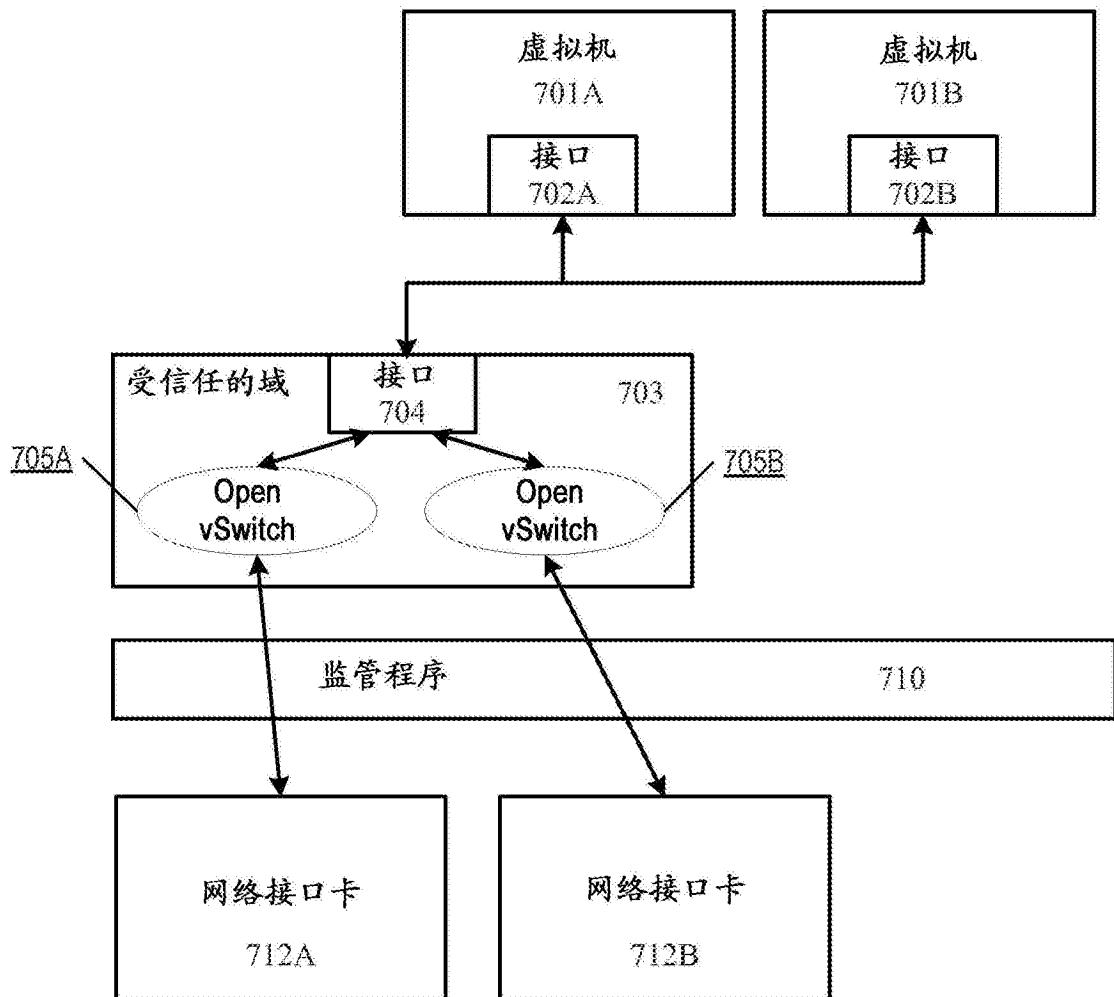


图7A

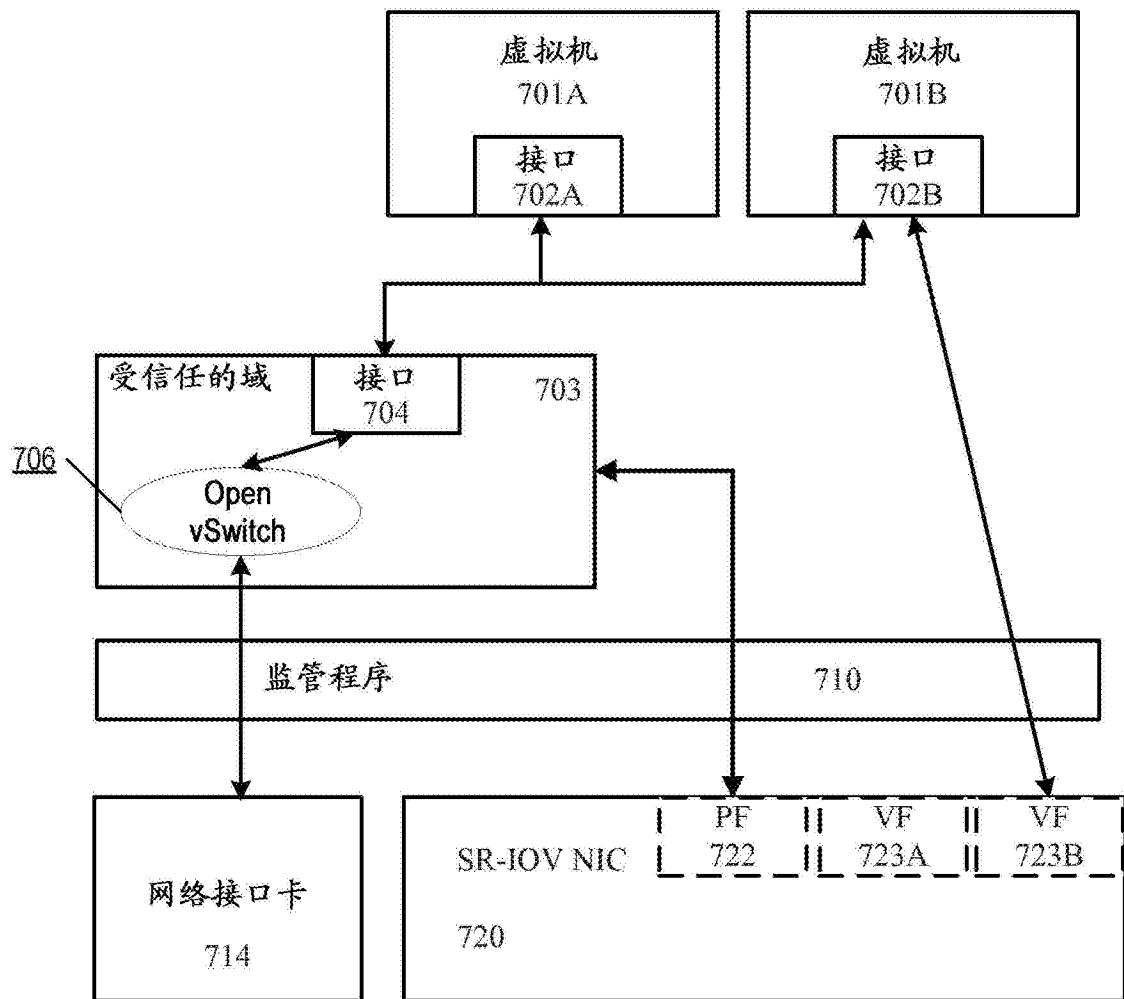


图7B

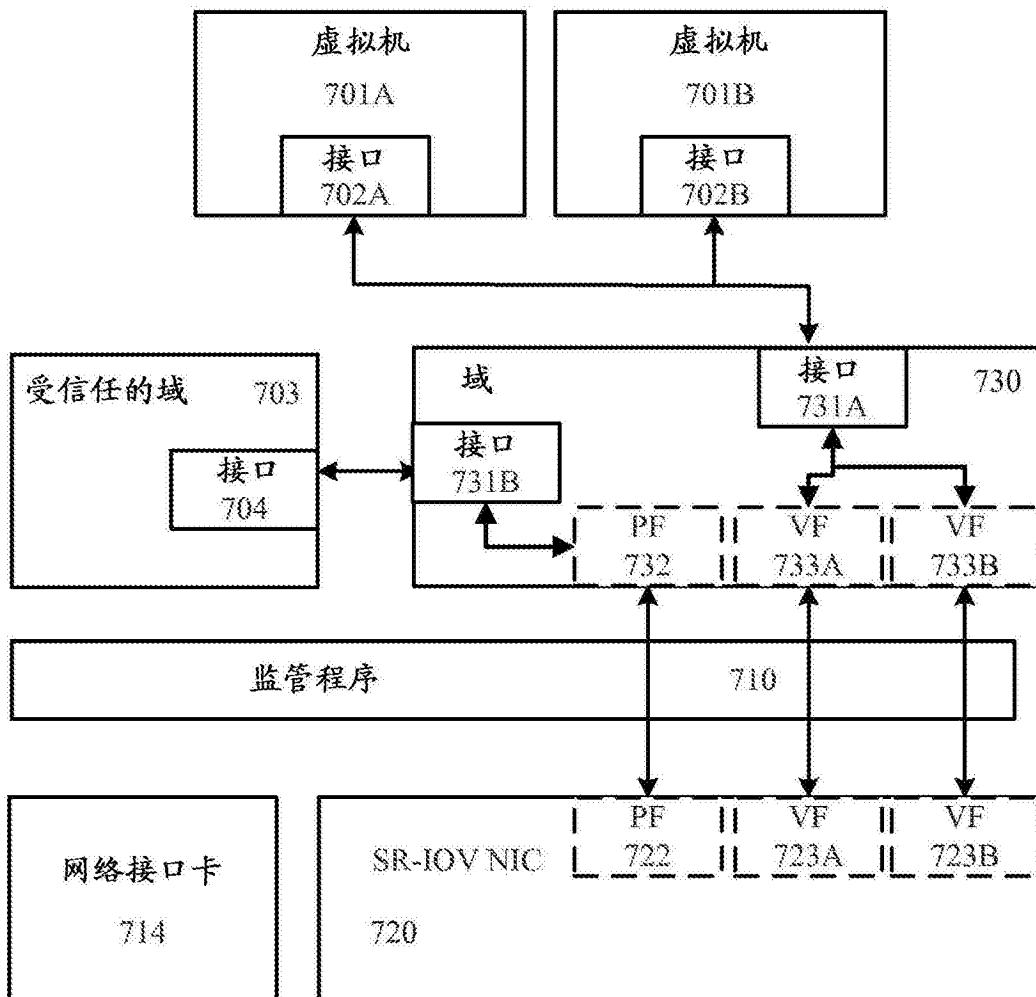


图7C

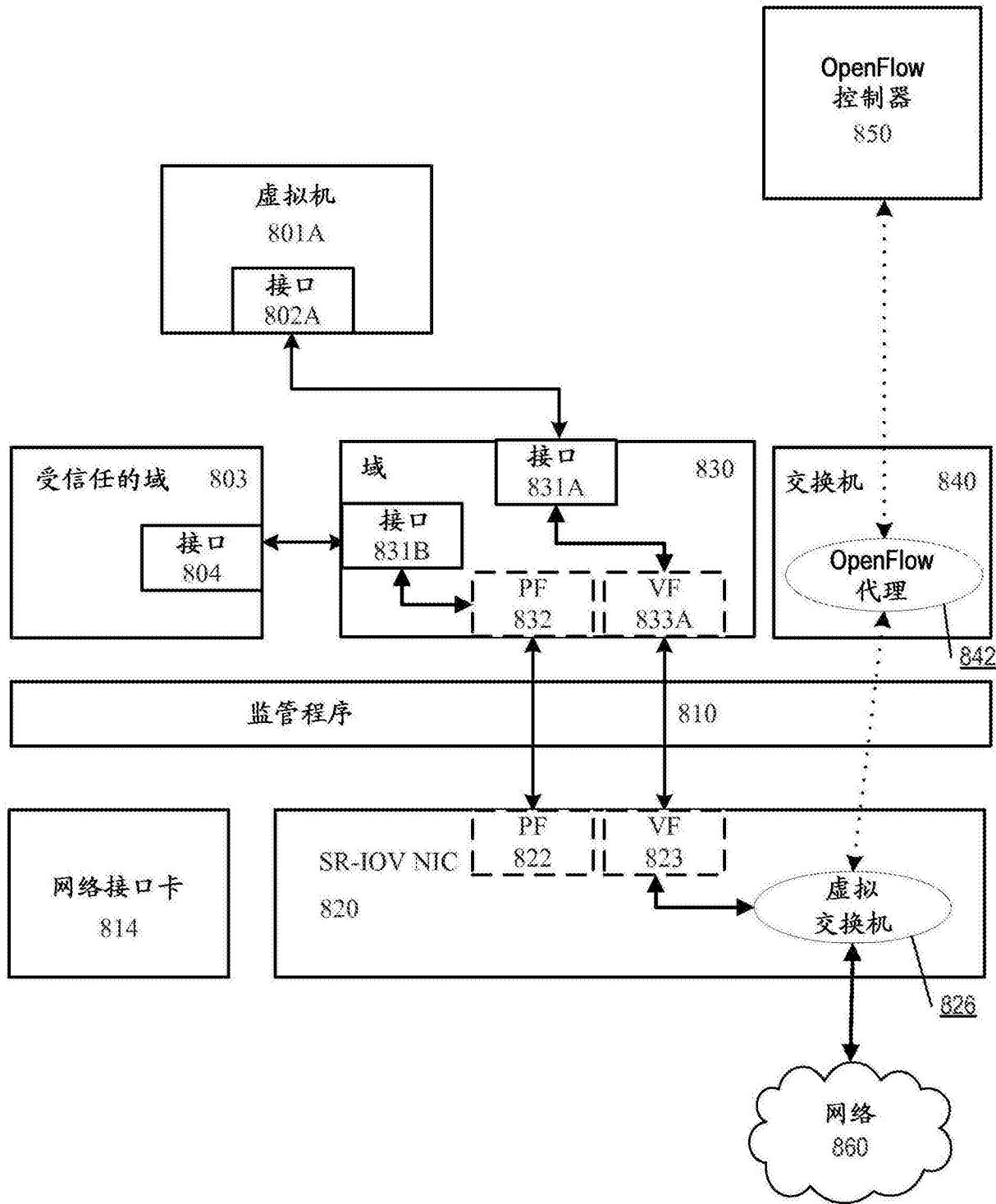


图8

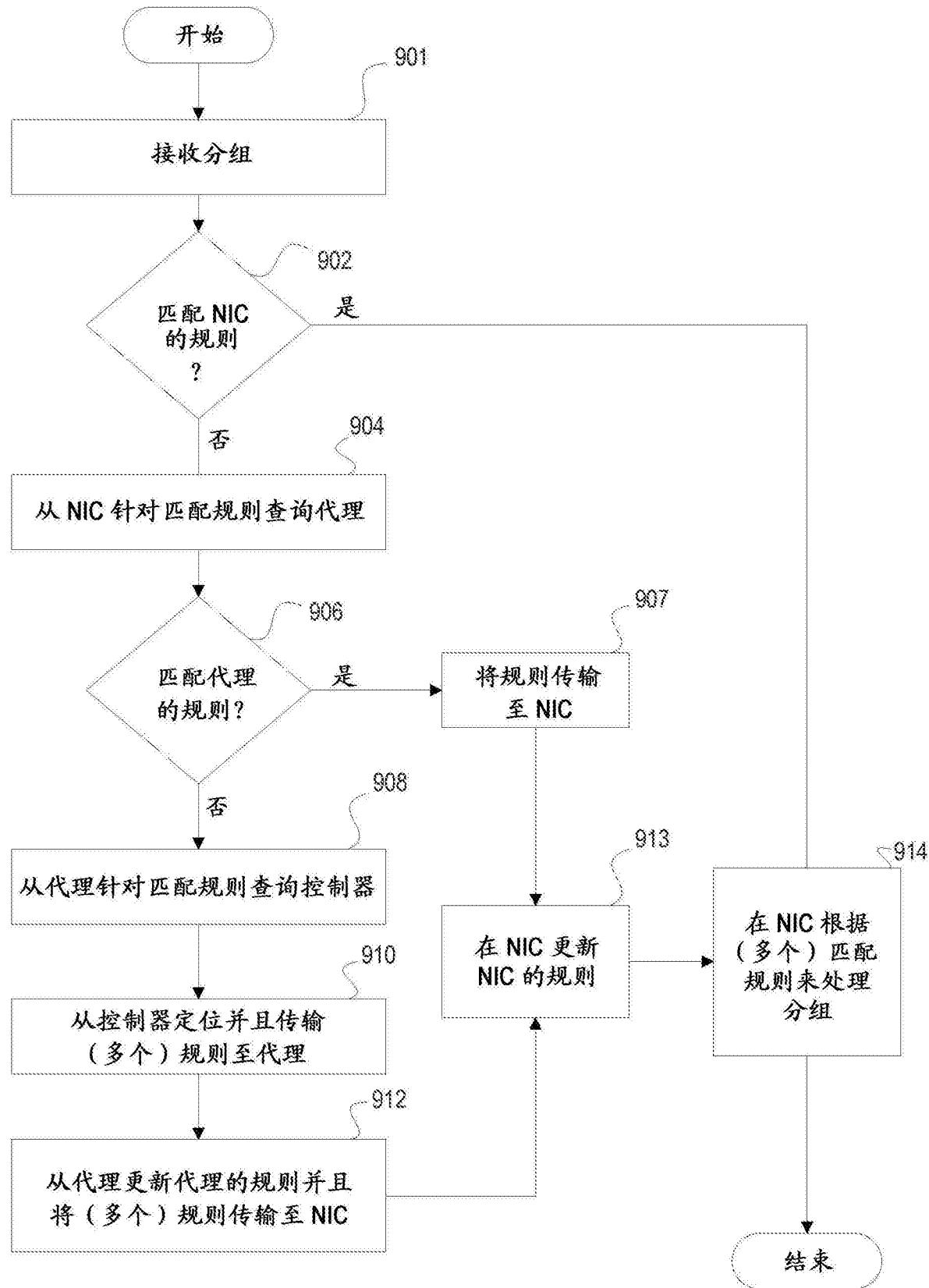


图9