



(12) 发明专利申请

(10) 申请公布号 CN 119278434 A

(43) 申请公布日 2025. 01. 07

(21) 申请号 202380042507.0

(22) 申请日 2023.05.10

(30) 优先权数据

63/353,462 2022.06.17 US

17/964,768 2022.10.12 US

(85) PCT国际申请进入国家阶段日

2024.11.25

(86) PCT国际申请的申请数据

PCT/US2023/021685 2023.05.10

(87) PCT国际申请的公布数据

W02023/244357 EN 2023.12.21

(71) 申请人 甲骨文国际公司

地址 美国

(72) 发明人 T·恩古延 D·克洛塞

S·帕蒂尔

(74) 专利代理机构 中国贸促会专利商标事务所

有限公司 11038

专利代理师 罗亚男

(51) Int.Cl.

G06F 9/455 (2006.01)

H04L 12/46 (2006.01)

H04L 41/0895 (2006.01)

H04L 41/40 (2006.01)

H04L 45/00 (2006.01)

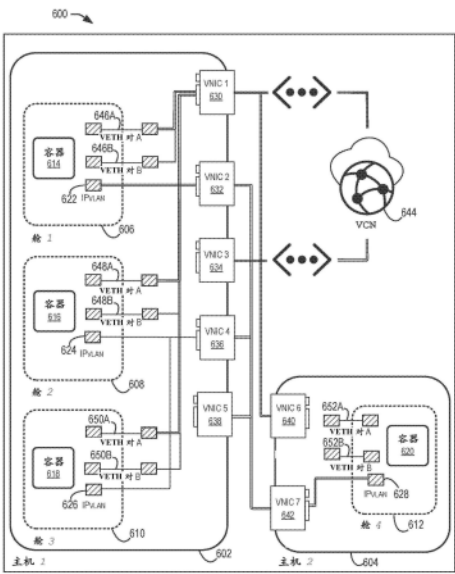
权利要求书3页 说明书41页 附图20页

(54) 发明名称

在容器环境内实现通信

(57) 摘要

描述了用于实现容器环境的技术,其中容器环境内的每个舱都设有唯一的IP地址和虚拟通信设备,诸如IPvlan设备。来自源舱的通信通过一个或多个虚拟化网络接口卡(VNIC)利用目的地舱的唯一IP地址被直接路由到容器环境内的目的地舱,而无需桥接和封装。这减少了要传送的数据的大小,并且也消除了在传输期间执行数据封装所必须的计算成本。



1. 一种用于在容器环境内通信的计算机实现的方法, 包括:

由计算机系统配置容器环境内的第一节点, 第一节点包括表示具有共享资源的一组一个或多个容器的第一舱;

由所述计算机系统配置从第一舱经由第一虚拟通信设备的通信;

由所述计算机系统在连接到所述容器环境内的第一节点的第一虚拟化网络接口卡(VNIC)处接收来自指配给第一节点内的第一舱的第一虚拟通信设备的数据, 所述数据被配置为被传送到目的地设备; 以及

由所述计算机系统经由第一虚拟化网络接口卡将所述数据路由到所述目的地设备, 所述数据至少包括标识第一舱的地址。

2. 根据权利要求1所述的计算机实现的方法, 其中:

所述目的地设备包括不同于第一舱的第二舱, 以及
第一舱和第二舱共享第一虚拟化网络接口卡。

3. 根据权利要求1或2所述的计算机实现的方法, 其中:

所述目的地设备包括不同于第一虚拟化网络接口卡的第二虚拟化网络接口卡, 以及
第二虚拟化网络接口卡与第一虚拟化网络接口卡两者都附接到所述容器环境的第一节点。

4. 根据权利要求1至3中任一项所述的计算机实现的方法, 其中:

所述目的地设备包括不同于第一虚拟化网络接口卡的第二虚拟化网络接口卡, 以及
第二虚拟化网络接口卡附接到不同于第一节点的所述容器环境的第二节点。

5. 根据权利要求1至4中任一项所述的计算机实现的方法, 其中所述目的地设备包括虚拟云网络(VCN)。

6. 根据权利要求1至5中任一项所述的计算机实现的方法, 其中第一虚拟通信设备包括互联网协议虚拟局域网(IPvlan)设备。

7. 根据权利要求1至6中任一项所述的计算机实现的方法, 包括:

由所述计算机系统在第一舱的第一虚拟通信设备处接收响应;

由所述计算机系统通过第一舱的一个或多个虚拟以太网(veth)对发送所述响应; 以及

由所述计算机系统将所述响应发送到位于第一舱内的容器。

8. 根据权利要求1至7中任一项所述的计算机实现的方法, 包括, 响应于由所述计算机系统确定所述目的地设备包括与第一舱共享第一虚拟化网络接口卡的第二舱:

由所述计算机系统利用发夹操作通过第一虚拟化网络接口卡路由所述数据; 以及

由所述计算机系统将所述数据从第一虚拟化网络接口卡路由到第二舱的第二虚拟通信设备。

9. 根据权利要求1至8中任一项所述的计算机实现的方法, 包括, 响应于由所述计算机系统确定所述目的地设备包括连接到与第一虚拟化网络接口卡分开的第二虚拟化网络接口卡的第一节点的第二舱:

由所述计算机系统将所述数据从第一虚拟化网络接口卡路由到第二虚拟化网络接口卡; 以及

由所述计算机系统将所述数据从第二虚拟化网络接口卡路由到第一节点内的第二舱的第二虚拟通信设备。

10. 根据权利要求1至9中任一项所述的计算机实现的方法, 包括, 响应于由所述计算机系统确定所述目的地设备包括与第一节点分开的第二节点的第二舱, 其中第二舱连接到与第一虚拟化网络接口卡分开的第二虚拟化网络接口卡:

由所述计算机系统将所述数据从第一虚拟化网络接口卡路由到第二虚拟化网络接口卡; 以及

由所述计算机系统将所述数据从第二虚拟化网络接口卡路由到第二节点内的第二舱的第二虚拟通信设备。

11. 一种系统, 包括:

一个或多个处理器, 被配置为:

配置容器环境内的第一节点, 第一节点包括表示具有共享资源的一组一个或多个容器的第一舱;

配置从第一舱经由第一虚拟通信设备的通信;

在连接到所述容器环境内的第一节点的第一虚拟化网络接口卡 (VNIC) 处接收来自指派给第一节点内的第一舱的第一虚拟通信设备的数据, 所述数据被配置为被传送到目的地设备; 以及

由第一虚拟化网络接口卡将所述数据路由到所述目的地设备, 所述数据至少包括标识第一舱的地址。

12. 根据权利要求11所述的系统, 其中:

所述目的地设备包括不同于第一舱的第二舱, 以及

第一舱和第二舱共享第一虚拟化网络接口卡。

13. 根据权利要求11或12所述的系统, 其中:

所述目的地设备包括不同于第一虚拟化网络接口卡的第二虚拟化网络接口卡, 以及

第二虚拟化网络接口卡与第一虚拟化网络接口卡两者都附接到所述容器环境的第一节点。

14. 根据权利要求11至13中任一项所述的系统, 其中:

所述目的地设备包括不同于第一虚拟化网络接口卡的第二虚拟化网络接口卡, 以及

第二虚拟化网络接口卡附接到不同于第一节点的所述容器环境的第二节点。

15. 根据权利要求11至14中任一项所述的系统, 其中所述目的地设备包括虚拟云网络 (VCN)。

16. 根据权利要求1至15中任一项所述的系统, 其中第一虚拟通信设备包括互联网协议虚拟局域网 (IPvlan) 设备。

17. 根据权利要求11至16中任一项所述的信息, 其中所述一个或多个处理器还被配置为:

在第一舱的第一虚拟通信设备处接收响应;

通过第一舱的一个或多个虚拟以太网 (veth) 对发送所述响应; 以及

将所述响应发送到位于第一舱内的容器。

18. 根据权利要求11至17中任一项所述的系统, 其中所述一个或多个处理器还被配置为, 响应于确定所述目的地设备包括与第一舱共享第一虚拟化网络接口卡的第二舱:

利用发夹操作通过第一虚拟化网络接口卡路由所述数据; 以及

将所述数据从第一虚拟化网络接口卡路由到第二舱的第二虚拟通信设备。

19. 根据权利要求11至18中任一项所述的系统,其中所述一个或多个处理器还被配置为,响应于确定所述目的地设备包括连接到与第一虚拟化网络接口卡分开的第二虚拟化网络接口卡的第一节点的第二舱:

将所述数据从第一虚拟化网络接口卡路由到第二虚拟化网络接口卡;以及

将所述数据从第二虚拟化网络接口卡路由到第一节点内的第二舱的第二虚拟通信设备。

20. 一种非暂时性计算机可读介质,存储一组指令,当所述一组指令由一个或多个处理器执行时,使得处理被执行,所述处理包括:

由计算机系统配置容器环境内的第一节点,第一节点包括表示具有共享资源的一组一个或多个容器的第一舱;

由所述计算机系统配置从第一舱经由第一虚拟通信设备的通信;

由所述计算机系统连接到所述容器环境内的第一节点的第一虚拟化网络接口卡(VNIC)处接收来自指配给第一节点内的第一舱的第一虚拟通信设备的数据,所述数据被配置为被传送到目的地设备;以及

由所述计算机系统经由第一虚拟化网络接口卡将所述数据路由到所述目的地设备,所述数据至少包括标识第一舱的地址。

21. 根据权利要求20所述的非暂时性计算机可读介质,其中所述一组指令还使得如权利要求2-10中任一项所述的方法的处理被执行。

22. 一种装置,包括:

用于配置容器环境内的第一节点的部件,第一节点包括表示具有共享资源的一组一个或多个容器的第一舱;

用于配置从第一舱经由第一虚拟通信设备的通信的部件;

用于在连接到所述容器环境内的第一节点的第一虚拟化网络接口卡(VNIC)处接收来自指配给第一节点内的第一舱的第一虚拟通信设备的数据的部件,所述数据被配置为被传送到目的地设备;以及

用于经由第一虚拟化网络接口卡将所述数据路由到所述目的地设备的部件,所述数据至少包括标识第一舱的地址。

23. 一种计算机程序产品,包括计算机指令,所述计算机指令当由处理器执行时,实现如权利要求1-10中任一项所述的方法的所述操作。

在容器环境内实现通信

[0001] 相关申请的交叉引用

[0002] 本申请要求于2022年6月17日提交的第63/353,462号美国临时专利申请和于2022年10月12日提交的第17/964,768号美国非临时专利申请的申请日的权益,这些申请中的每一个申请出于所有目的通过引用整体并入本文。

背景技术

[0003] 如今,包括诸如Kubernetes的编排系统的容器环境是用于实现和部署基于容器的应用的流行手段。然而,在传统系统内实现的应用和舱 (pod) 仅利用封装这样的通信的桥接设备彼此通信。此封装为这样的通信增加了计算和数据开销,并且还阻止了通信流日志中的特定的发送者和接收者信息的标识。

发明内容

[0004] 本公开的实施例涉及促进容器环境内舱之间的直接通信。公开了一种容器环境,其包括一个或多个节点,每个节点包含一个或多个舱。在容器环境内,虚拟通信设备包括在每个舱内,用于与其他舱通信,而没有对桥接器或封装的使用。一个或多个虚拟网络接口卡 (VNIC) 附接到包含舱的节点,并且这些VNIC通过接收和转发经其虚拟通信设备发送到每个舱和从每个舱发送的数据来促进舱通信。

[0005] 至少一个实施例针对一种计算机实现的方法。方法可以包括配置具有一个或多个操作系统级虚拟化实例的节点。方法还可以包括配置从一个或多个操作系统级虚拟化实例经由虚拟通信设备的通信。

[0006] 至少一个实施例针对一种计算机实现的方法。方法可以包括配置容器环境内的第一节点,该第一节点包括第一容器。方法还可以包括配置从第一容器经由第一虚拟通信设备的通信。方法还可以包括在连接到容器环境内的第一节点的第一虚拟化网络接口卡 (VNIC) 处接收来自指配给第一节点内的第一容器的第一虚拟通信设备的数据,该数据配置为传送到目的地设备。方法还可以包括由第一VNIC将数据路由到目的地设备,该数据至少包括标识第一容器的地址。

[0007] 另一实施例针对一种计算设备,该计算设备包括一个或多个处理器和指令,这些指令当由一个或多个处理器执行时,使得计算设备执行本文公开的(一个或多个)方法的任何适当组合。

[0008] 另一实施例针对一种非暂时性计算机可读介质,该非暂时性计算机可读介质存储计算机可执行指令,这些指令当由计算集群的一个或多个处理器执行时,使得计算集群执行本文公开的(一个或多个)方法的任何适当组合。

[0009] 又一实施例针对一种用于在容器环境内实现通信的装置,其包括:用于配置容器环境内第一节点的部件,第一节点包括表示具有共享资源的一组一个或多个容器的第一舱;用于配置从第一舱经由第一虚拟通信设备的通信的部件;用于在连接到容器环境内的第一节点的第一虚拟化网络接口卡 (VNIC) 处接收来自指配给第一节点内的第一舱的第一

虚拟通信设备的数据的部件,该数据配置为传送到目的地设备;以及用于经由第一虚拟化网络接口卡将数据路由到目的地设备的部件,该数据至少包括标识第一舱的地址。

[0010] 又另一实施例针对一种计算机程序产品,其包括计算机指令,这些指令当由处理器执行时,实现本文公开的(一个或多个)方法的任何适当组合。

附图说明

[0011] 为了容易地标识对任何特定元素或动作的讨论,附图标记中一个或多个最高数位指的是该元素被首次引入的图号。

[0012] 图1是根据某些实施例的示出由云服务提供商基础设施托管的虚拟或覆盖云网络的分布式环境的高级图。

[0013] 图2描绘了根据某些实施例的CSPI内的物理网络中的物理组件的简化架构图。

[0014] 图3示出了根据某些实施例的CSPI内的示例布置,其中主机连接到多个网络虚拟化设备(NVD)。

[0015] 图4描绘了根据某些实施例的用于提供I/O虚拟化以支持多租赁的主机和NVD之间的连接性。

[0016] 图5描绘了根据某些实施例的由CSPI提供的物理网络的简化框图。

[0017] 图6是根据至少一个实施例的用于促进直接的舱到舱通信的容器环境的框图。

[0018] 图7是用于图示根据至少一个实施例的在图1的示例性容器环境内共享单个VNIC的两个舱之间的示例性请求流的框图。

[0019] 图8是用于图示根据至少一个实施例的在图1的示例性容器环境内共享单个VNIC的两个舱之间的示例性响应流的框图。

[0020] 图9是用于图示根据至少一个实施例的图1的示例性容器环境的相同节点内经由不同VNIC通信的两个舱之间的示例性请求流的框图。

[0021] 图10是用于图示根据至少一个实施例的图1的示例性容器环境的相同节点内经由不同VNIC通信的两个舱之间的示例性响应流的框图。

[0022] 图11是用于图示根据至少一个实施例的图1的示例性容器环境的不同节点内经由不同VNIC通信的两个舱之间的示例性请求流的框图。

[0023] 图12是用于图示根据至少一个实施例的图1的示例性容器环境的不同节点内经由不同VNIC通信的两个舱之间的示例性响应流的框图。

[0024] 图13是用于图示根据至少一个实施例的图1的示例性容器环境内舱和虚拟云网络(VCN)之间的示例性请求流的框图。

[0025] 图14是用于图示根据至少一个实施例的图1的示例性容器环境内舱和虚拟云网络(VCN)之间的示例性响应流的框图。

[0026] 图15图示了根据至少一个实施例的用于在容器环境内实现通信的示例方法。

[0027] 图16是图示根据至少一个实施例的用于实现云基础设施即服务系统的一种模式的框图。

[0028] 图17是图示根据至少一个实施例的用于实现云基础设施即服务系统的另一模式的框图。

[0029] 图18是图示根据至少一个实施例的用于实现云基础设施即服务系统的另一模式

的框图

[0030] 图19是图示根据至少一个实施例的用于实现云基础设施即服务系统的另一模式的框图。

[0031] 图20是图示根据至少一个实施例的示例计算机系统的框图。

具体实施方式

[0032] 示例虚拟联网架构

[0033] 术语云服务通常用于指由云服务提供商 (CSP) 使得使用CSP提供的系统和基础设施 (云基础设施) 按需 (例如, 经由订阅模型) 对用户或客户可用的服务。通常, 组成CSP基础设施的服务器和系统与用户自己的本地部署服务器和系统是分开的。因此, 客户可以利用CSP提供的云服务自身, 而无需为服务购买分开的硬件和软件资源。云服务被设计为为订阅客户提供对应用和计算资源的轻松、可扩展的访问, 而无需客户投资购买用于提供服务的

基础设施。

[0034] 有供应各种类型的云服务的若干个云服务提供商。有各种不同的云服务类型或模型, 包括软件即服务 (SaaS)、平台即服务 (PaaS)、基础设施即服务 (IaaS) 等。

[0035] 客户可以订阅由CSP提供的一个或多个云服务。客户可以是任何实体, 诸如个人、组织、企业等。当客户订阅或注册CSP提供的服务时, 租赁或帐户针对该客户被创建。然后, 客户可以经由此帐户访问与帐户相关联的订阅的一个或多个云资源。

[0036] 如上所述, 基础设施即服务 (IaaS) 是一种特定类型的云计算服务。在IaaS模型中, CSP提供基础设施 (被称为云服务提供商基础设施或CSPI), 其可以被客户用来构建它们自己的可定制网络并部署客户资源。因此, 客户的资源和网络由CSP提供的基础设施托管在分布式环境中。这与传统计算不同, 传统计算中客户的资源和网络由客户提供的基础设施托管。

[0037] CSPI可以包括互连的高性能计算资源, 包括形成物理网络的各种主机、存储器资源和网络资源, 该物理网络也被称为基底网络或底层网络。CSPI中的资源可以跨一个或多个数据中心分散, 这些数据中心可以在地理上跨一个或多个地理区域分散。虚拟化软件可以由这些物理资源执行以提供虚拟化分布式环境。虚拟化在物理网络上创建覆盖网络 (也被称为基于软件的网络、软件定义的网络或虚拟网络)。CSPI物理网络为在物理网络之上创建一个或多个覆盖或虚拟网络提供了底层基础。物理网络 (或基底网络或底层网络) 包括物理网络设备, 诸如物理交换机、路由器、计算机和主机等。覆盖网络是在物理基底网络之上运行的逻辑 (或虚拟) 网络。给定的物理网络可以支持一个或多个覆盖网络。覆盖网络通常使用封装技术来区分属于不同覆盖网络的流量。虚拟或覆盖网络也被称为虚拟云网络 (VCN)。虚拟网络使用软件虚拟化技术 (例如, 管理程序、由网络虚拟化设备 (NVD) 实现的功能 (例如, smartNIC)、架顶式 (TOR) 交换机、实现由NVD执行的一个或多个功能的智能TOR和其他机制) 来实现以创建可以在物理网络之上运行的网络抽象层。虚拟网络可以采用许多形式, 包括对等网络、IP网络等。虚拟网络通常是层-3IP网络或层-2VLAN。此虚拟或覆盖联网的方法通常被称为虚拟或覆盖层-3联网。为虚拟网络开发的协议示例包括IP-in-IP (或通用路由封装 (GRE))、虚拟可扩展LAN (VXLAN—IETF RFC 7348)、虚拟专用网络 (VPN) (例如, MPLS层-3虚拟专用网络) (RFC 4364)、VMware的NSX、GENEVE (通用网络虚拟化封装) 等。

[0038] 对于IaaS,由CSP提供的基础设施(CSPI)可以被配置为通过公共网络(例如,互联网)提供虚拟化的计算资源。在IaaS模型中,云计算提供商可以托管基础设施组件(例如,服务器、存储设备、网络节点(例如,硬件)、部署软件、平台虚拟化(例如,管理程序层)或类似物)。在一些情况下,IaaS提供商还可以供应各种各样的服务来伴随那些基础设施组件(例如,计费、监控、日志记录、安全、负载均衡和集群等)。因此,由于这些服务可以是策略驱动的,IaaS用户可以能够实现策略来驱动负载均衡以维持应用可用性和性能。CSPI提供基础设施和一组互补的云服务,其使得客户能够在高度可用的托管分布式环境中构建和运行广泛的应用和服务。CSPI在灵活的虚拟网络中提供高性能计算资源和能力以及存储容量,该虚拟网络可从各种联网的位置(诸如从客户的本地部署网络)安全地访问。当客户订阅或注册CSP提供的IaaS服务时,针对该客户创建的租赁是CSPI内的安全且隔离的分区,客户可以在其中创建、组织和管理其云资源。

[0039] 客户可以使用CSPI提供的计算、存储器和联网资源来构建他们自己的虚拟网络。一个或多个客户资源或工作负载(诸如计算实例)可以部署在这些虚拟网络上。例如,客户可以使用CSPI提供的资源来构建被称为虚拟云网络(VCN)的一个或多个可定制且私有的(一个或多个)虚拟网络。客户可以在客户VCN上部署一个或多个客户资源,诸如计算实例。计算实例可以采用虚拟机、裸机实例等形式。因此,CSPI提供基础设施和一组互补的云服务,其使客户能够在高度可用的虚拟托管环境中构建和运行广泛的应用和服务。客户不管理或控制CSPI提供的底层物理资源,但具有对操作系统、存储和部署的应用的控制;以及对选择的联网组件(例如,防火墙)的可能的有限控制。

[0040] CSP可以提供使得客户和网络管理员能够使用CSPI资源来配置、访问和管理部署在云中的资源的控制台。在某些实施例中,控制台提供可以用于访问和管理CSPI的基于web的用户接口。在一些实现中,控制台是由CSP提供的基于web的应用。

[0041] CSPI可以支持单租赁或多租赁架构。在单租赁架构中,软件(例如,应用、数据库)或硬件组件(例如,主机或服务器)服务单个客户或租户。在多租赁架构中,软件或硬件组件服务多个客户或租户。因此,在多租赁架构中,CSPI资源在多个客户或租户之间共享。在多租赁情况下,在CSPI内采取预防措施并且实施保障措施,以确保每个租户的数据是隔离的并且对其他租户保持不可见。

[0042] 在物理网络中,网络端点(“端点”)是指连接到物理网络并且与其所连接到的网络来回通信的计算设备或系统。物理网络中的网络端点可以连接到局域网(LAN)、广域网(WAN)或其他类型的物理网络。物理网络中的传统端点的示例包括调制解调器、集线器、桥接器、交换机、路由器和其他联网设备、物理计算机(或主机)等。物理网络中的每个物理设备具有可以用于与设备通信的固定网络地址。此固定网络地址可以是层-2地址(例如,MAC地址)、固定层-3地址(例如,IP地址)等。在虚拟化环境中或在虚拟网络中,端点可以包括各种虚拟端点,诸如由物理网络的组件托管(例如,由物理主机托管)的虚拟机。虚拟网络中的这些端点通过诸如覆盖层-2地址(例如,覆盖MAC地址)和覆盖层-3地址(例如,覆盖IP地址)的覆盖地址来寻址。网络覆盖通过允许网络管理者使用软件管理(例如,经由实现虚拟网络的控制平面的软件)移动与网络端点相关联的覆盖地址来使能灵活性。因此,与在物理网络中不同,在虚拟网络中,覆盖地址(例如,覆盖IP地址)可以使用网络管理软件被从一个端点移动到另一个。由于虚拟网络构建在物理网络之上,因此虚拟网络中的组件之间的通信涉

及虚拟网络和底层物理网络两者。为了促进这样的通信,CSPI的组件被配置为学习并存储将虚拟网络中的覆盖地址映射到基底网络中的实际物理地址的映射,反之亦然。然后这些映射被用来促进通信。客户流量被封装以促进在虚拟网络中路由。

[0043] 因此,物理地址(例如,物理IP地址)与物理网络中的组件相关联,并且覆盖地址(例如,覆盖IP地址)与虚拟或覆盖网络中的实体相关联。物理IP地址是与基底或物理网络中的物理设备(例如,网络设备)相关联的IP地址。例如,每个NVD具有相关联的物理IP地址。覆盖IP地址是与覆盖网络中的实体(诸如与客户的虚拟云网络(VCN)中的计算实例)相关联的覆盖地址。两个不同的客户或租户(每个具有其自己的私有VCN)可能潜在地在彼此一无所知的情况下在其VCN中使用相同的覆盖IP地址。物理IP地址和覆盖IP地址两者都是真实IP地址的类型。这些与虚拟IP地址是分开的。虚拟IP地址通常是表示或映射到多个真实IP地址的单个IP地址。虚拟IP地址提供了虚拟IP地址与多个真实IP地址之间的一对多映射。例如,负载均衡器可以使用VIP来映射或表示多个服务器,每个服务器具有其自己的真实IP地址。

[0044] 云基础设施或CSPI物理地托管在世界各地一个或多个地区中的一个或多个数据中心中。CSPI可以包括物理或基底网络中的组件以及构建在物理网络组件之上的虚拟网络中的虚拟化组件(例如,虚拟网络、计算实例、虚拟机等)。在某些实施例中,CSPI在领域、区域和可用性域中组织和托管。区域通常是包含一个或多个数据中心的局部地理区。区域通常彼此独立,并且可以分开很远距离,例如跨越国家甚至大洲。例如,第一区域可以在澳大利亚,另一区域在日本,又另一区域在印度,等等。CSPI资源在区域之间划分,使得每个区域有其自己独立的CSPI资源子集。每个区域可以提供一组核心基础设施服务和资源,诸如,计算资源(例如,裸机服务器、虚拟机、容器和相关基础设施等);存储资源(例如,块卷存储、文件存储、对象存储、归档存储);联网资源(例如,虚拟云网络(VCN)、负载均衡资源、与本地部署网络的连接)、数据库资源;边缘联网资源(例如,DNS);以及访问管理和监控资源等。每个区域通常具有将其连接到领域内的其他区域的多个路径。

[0045] 一般地,应用部署在其被最频繁使用的区域中(即,部署在与该区域相关联的基础设施上),因为使用附近的资源比使用远处的资源更快。应用还可以出于各种原因部署在不同的区域,诸如冗余以减轻诸如大型天气系统或地震的区域范围事件的风险,以满足法律管辖区、税域和其他商业或社会标准等的变化的要求。

[0046] 区域内的数据中心可以被进一步组织并细分为可用性域(AD)。可用性域可以对应于位于区域内的一个或多个数据中心。区域可以由一个或多个可用性域组成。在这样的分布式环境中,CSPI资源是特定于区域的(诸如虚拟云网络(VCN)),或是特定于可用性域的,诸如计算实例。

[0047] 区域内的AD彼此隔离、容错,并且被配置为使得它们不太可能同时故障。这是通过AD不共享关键基础设施资源(诸如联网、物理线缆、线缆路径、线缆入口点等)来实现的,使得区域内的一个AD处的故障不太可能影响相同区域内其他AD的可用性。相同区域内的AD可以通过低延迟、高带宽网络彼此连接,这使得提供到其他网络(例如,互联网、客户本地部署网络等)高可用性连接性并在多个AD中针对高可用性构建复制系统和灾难恢复成为可能。云服务使用多个AD来确保高可用性并防止资源故障。随着IaaS提供商提供的基础设施的增长,更多区域和AD可以被添加附加的容量。可用性域之间的流量通常是加密的。

[0048] 在某些实施例中,区域被分组为领域。领域是区域的逻辑集合。领域彼此隔离并且不共享任何数据。相同领域中的区域可以彼此通信,但不同领域中的区域则不能。客户与CSP的租赁或帐户存在于单个领域中,并且可以跨属于该领域的一个或多个区域分散。通常,当客户订阅IaaS服务时,租赁或帐户在领域内的客户指定区域(被称为“归属”区域)中被针对该客户创建。客户可以跨领域内的一个或多个其他区域延伸客户的租赁。客户不能访问不在客户租赁存在的领域中的区域。

[0049] IaaS提供商可以提供多个领域,每个领域迎合一组特定的客户或用户。例如,可以为商业客户提供商业领域。作为另一示例,可以为特定国家内客户提供特定国家的领域。作为又另一示例,可以为政府提供政府领域等。例如,政府领域可以迎合特定政府,并且可以具有比商业领域提高的安全级别。例如,Oracle云基础设施(OCI)当前为商业区域供应一个领域并且为政府云区域供应两个领域(例如,FedRAMP授权和IL5授权)。

[0050] 在某些实施例中,AD可以被细分为一个或多个故障域。故障域是AD内的基础设施资源的分组,以提供反亲和性(anti-affinity)。故障域允许计算实例的分布,使得实例不在单个AD内的相同物理硬件上。这被称为反亲和性。故障域是指共享单点故障的一组硬件组件(计算机、交换机等)。计算池在逻辑上分为多个故障域。因此,影响一个故障域的硬件故障或计算硬件维护事件不影响其他故障域中的实例。取决于实施例,每个AD的故障域的数量可以变化。例如,在某些实施例中,每个AD包含三个故障域。故障域充当AD内的逻辑数据中心。

[0051] 当客户订阅IaaS服务时,来自CSPI的资源被供给到客户并且与客户的租赁相关联。客户可以使用这些供给的资源来构建私有网络并在这些网络上部署资源。由CSPI在云中托管的客户网络被称为虚拟云网络(VCN)。客户可以使用为客户分配的CSPI资源来建立一个或多个虚拟云网络(VCN)。VCN是虚拟或软件定义的私有网络。部署在客户的VCN中的客户资源可以包括计算实例(例如,虚拟机、裸机实例)和其他资源。这些计算实例可以表示各种客户工作负载,诸如应用、负载均衡器、数据库等。部署在VCN上的计算实例可以通过公共网络(诸如互联网)与公共可访问端点(“公共端点”)通信,可以与相同VCN或其他VCN(例如,客户的其他VCN或不属于客户的VCN)中的其他实例通信、可以与客户的本地部署数据中心或网络通信以及可以与服务端点和其他类型的端点通信。

[0052] CSP可以使用CSPI来提供各种服务。在一些实例中,CSPI的客户本身可以表现得像服务提供商并使用CSPI资源提供服务。服务提供商可以暴露服务端点,其特征在于标识信息(例如,IP地址、DNS名称和端口)。客户的资源(例如,计算实例)可以通过访问服务为特定服务暴露的服务端点来消费该特定服务。这些服务端点通常是可由用户使用与端点相关联的公共IP地址经由公共通信网络(诸如互联网)公开访问的端点。可公开访问的网络端点有时也被称为公共端点。

[0053] 在某些实施例中,服务提供商可以经由服务的端点(有时被称为服务端点)来暴露服务。然后,服务的客户可以使用此服务端点来访问服务。在某些实现中,为服务提供的服务端点可以由旨在消费该服务的多个客户访问。在其他实现中,可以为客户提供专用服务端点,使得只有该客户可以使用该专用服务端点来访问服务。

[0054] 在某些实施例中,当VCN被创建时,它与私有覆盖无类别域间路由(CIDR)地址空间相关联,该地址空间是指配给VCN的私有覆盖IP地址的范围(例如,10.0/16)。VCN包括相关

联的子网、路由表和网关。VCN驻留在单个区域内,但可以跨越区域的可用性域中的一个或多个或所有。网关是为VCN配置的虚拟接口并且使能进出VCN的流量到VCN外部的一个或多个端点的通信。一个或多个不同类型的网关可以被配置用于VCN以使能到不同类型的端点的或来自不同类型的端点的通信。

[0055] VCN可以被细分为一个或多个子网络,诸如一个或多个子网。因此,子网是可以在VCN内创建的配置单元或细分。VCN可以具有一个或多个子网。VCN内的每个子网都与连续范围的覆盖IP地址(例如10.0.0.0/24和10.0.1.0/24)相关联,这些地址不与该VCN中的其他子网以及表示VCN的地址空间内的地址空间子集的子网重叠。

[0056] 每个计算实例与虚拟网络接口卡(VNIC)相关联,其使得计算实例能够加入VCN的子网。VNIC是物理网络接口卡(NIC)的逻辑表示。一般地,VNIC是实体(例如,计算实例、服务)和虚拟网络之间的接口。VNIC存在于子网中,具有一个或多个相关联的IP地址并与安全规则或策略相关联。VNIC相当于交换机上的层-2端口。VNIC附接到计算实例以及VCN内的子网。与计算实例相关联的VNIC使得计算实例能够成为VCN子网的一部分,并使得计算实例能够与与和计算实例位于相同子网上的端点,与VCN中不同子网中的端点,或与VCN外部的端点通信(例如,发送和接收分组)。因此,与计算实例相关联的VNIC确定计算实例如何与VCN内部和外部的端点连接。当计算实例被创建并被添加到VCN内的子网时,计算实例的VNIC被创建并与该计算实例相关联。对于包括一组计算实例的子网,子网包含与该组计算实例相对应的VNIC,每个VNIC附接到该组计算机实例内的计算实例。

[0057] 每个计算实例经由与计算实例相关联的VNIC被指配私有覆盖IP地址。当计算实例被创建时,此私有覆盖IP地址被指配给与计算实例相关联的VNIC,并用于路由进出计算实例的流量。给定子网中的所有VNIC使用相同的路由表、安全列表和DHCP选项。如上所述,VCN内的每个子网与连续范围的覆盖IP地址(例如,10.0.0.0/24和10.0.1.0/24)相关联,这些地址不与该VCN中的其他子网以及表示VCN的地址空间内的地址空间子集的子网重叠。对于VCN的特定子网上的VNIC,指配给VNIC的私有覆盖IP地址是来自为该子网分配的连续范围的覆盖IP地址的地址。

[0058] 在某些实施例中,除了私有覆盖IP地址之外,计算实例还可以可选地被指配附加的覆盖IP地址,诸如,例如如果在公共子网中则是一个或多个公共IP地址。这些多个地址在相同VNIC上或在与计算实例相关联的多个VNIC上指配。然而,每个实例有在实例发起期间创建的主VNIC,并与指配给实例的覆盖私有IP地址相关联——此主VNIC不可以被移除。附加的VNIC(被称为次级VNIC)可被添加到与主VNIC相同的可用性域中的现有实例。所有VNIC在与实例相同的可用性域中。次级VNIC可以在与主VNIC相同的VCN中的子网中,或者可以在相同VCN或不同VCN中的不同子网中。

[0059] 如果计算实例在公共子网中,则计算实例可以被可选地指配公共IP地址。在子网被创建时,子网可以被指明为公共子网或私有子网。私有子网意味着子网中的资源(例如,计算实例)和相关联的VNIC不能具有公共覆盖IP地址。公共子网意味着子网中的资源和相关联的VNIC可以具有公共IP地址。客户可以指明子网以存在于单个可用性域中或跨区域或领域中的多个可用性域存在。

[0060] 如上所述,VCN可以被细分为一个或多个子网。在某些实施例中,为VCN配置的虚拟路由器(VR)(被称为VCN VR或仅VR)使能VCN的子网之间的通信。对于VCN内的子网,VR表示

该子网的逻辑网关,其使得该子网(即,该子网上的计算实例)能够与VCN内其他子网上的端点以及与VCN外部的其他端点通信。VCN VR是被配置为在VCN中的VNIC和与VCN相关联的虚拟网关(“网关”)之间路由流量的逻辑实体。网关在下面关于图1被进一步描述。VCN VR是层-3/IP层概念。在一个实施例中,有用于VCN的一个VCN VR,其中VCN VR潜在地具有由IP地址寻址的无限数量的端口,其中对于VCN的每个子网有一个端口。以此方式,VCN VR对于VCN VR所附接到的VCN中的每个子网具有不同的IP地址。VR还连接到被配置用于VCN的各种网关。在某些实施例中,来自子网的覆盖IP地址范围中的特定覆盖IP地址被保留用于该子网的VCN VR的端口。例如,考虑具有两个子网的VCN,该两个子网分别具有相关联的地址范围10.0/16和10.1/16。对于VCN内具有地址范围为10.0/16的第一子网,来自此范围的地址被保留用于该子网的VCN VR的端口。在一些实例中,来自范围内的第一IP地址可以被保留用于VCN VR。例如,对于具有覆盖IP地址范围10.0/16的子网,IP地址10.0.0.1可以针对该子网被保留用于VCN VR的端口。对于相同VCN内具有地址范围为10.1/16的第二子网,VCN VR可以具有针对该第二子网的端口,其具有IP地址10.1.0.1。VCN VR针对VCN中的子网中的每一个有不同的IP地址。

[0061] 在一些其他实施例中,VCN内的每个子网可以具有其自己相关联的VR,该VR可由子网使用与VR相关联的保留的或默认的IP地址来寻址。保留的或默认的IP地址可以例如是来自与该子网相关联的IP地址范围的第一IP地址。子网中的VNIC可以使用此默认的或保留的IP地址和与子网相关联的VR通信(例如,发送和接收分组)。在这样的实施例中,VR是该子网的入口/出口点。与VCN内的子网相关联的VR可以和与VCN内的其他子网相关联的其他VR通信。VR还可以与与VCN相关联的网关通信。子网的VR功能正在一个或多个NVD上运行或由一个或多个NVD执行,该一个或多个NVD执行子网中VNIC的VNIC的功能。

[0062] 可以为VCN配置路由表、安全规则和DHCP选项。路由表是VCN的虚拟路由表并且包括通过网关或专门配置的实例将流量从VCN内的子网路由到VCN外部的目的地的规则。VCN的路由表可被定制以控制分组如何被转发/路由到VCN或分组如何被从VCN转发/路由。DHCP选项是指当实例启动时自动提供给实例的配置信息。

[0063] 为VCN配置的安全规则表示VCN的覆盖防火墙规则。安全规则可以包括入口和出口规则,并且指定允许进出VCN内的实例的流量类型(例如,基于协议和端口)。客户可以选择给定规则是有状态的还是无状态的。例如,客户可以通过使用源CIDR 0.0.0.0/0和目的地TCP端口22建立有状态的入口规则来允许从任何地方到一组实例的传入SSH流量。安全规则可以使用网络安全组(network security group)或安全列表来实现。网络安全组由仅适用于该组中的资源的一组安全规则组成。另一方面,安全列表包括适用于使用该安全列表的任何子网中所有资源的规则。VCN可以被提供具有默认安全规则的默认安全列表。为VCN配置的DHCP选项提供在实例启动时自动提供给VCN中的实例的配置信息。

[0064] 在某些实施例中,VCN的配置信息由VCN控制平面确定并存储。VCN的配置信息可以包括例如关于以下信息:与VCN相关联的地址范围、VCN内的子网和相关信息、与VCN相关联的一个或多个VR、VCN中的计算实例和相关信息、VNIC、执行与VCN相关联的各种虚拟化网络功能(例如,VNIC、VR、网关)的NVD、VCN的状态信息以及其他VCN相关信息。在某些实施例中,VCN分布服务向NVD发布由VCN控制平面存储的配置信息或其部分。分布式信息可以用于更新由NVD存储和使用的信息(例如,转发表、路由表等),以将分组转发到VCN中的计算实

例以及从VCN中的计算实例转发分组。

[0065] 在某些实施例中,VCN和子网的创建由VCN控制平面(CP)处理,并且计算实例的发起由计算控制平面处理。计算控制平面负责为计算实例分配物理资源,然后调用VCN控制平面创建VNIC并将其附接到计算实例。VCN CP还将VCN数据映射发送到被配置为执行分组转发和路由功能的VCN数据平面。在某些实施例中,VCN CP提供负责向VCN数据平面提供更新的分布服务。VCN控制平面的示例也在图16、17、18和19(参见参考1616、1716、1816和1916)描绘并在下面描述。

[0066] 客户可以使用由CSPI托管的资源来创建一个或多个VCN。部署在客户VCN上的计算实例可以与不同的端点通信。这些端点可以包括CSPI托管的端点和CSPI外部的端点。

[0067] 用于使用CSPI来实现基于云的服务的各种不同架构在图1、2、3、4、5、16、17、18和20中描绘并在下面描述。图1是根据某些实施例的示出由CSPI托管的覆盖或客户VCN的分布式环境100的高级图。图1中描绘的分布式环境包括覆盖网络中的多个组件。图1中描绘的分布式环境100仅是示例并且不旨在不适当地限制所要求保护的实施例的范围。许多变化、替代和修改都是可能的。例如,在一些实现中,图1中描绘的分布式环境可以是具有比图1中所示的那些更多或更少的系统或组件,可以组合两个或更多个系统,或者可以具有不同的系统配置或布置。

[0068] 如图1中描绘的示例所示,分布式环境100包括CSPI 101,其提供客户可以订阅并用于构建他们的虚拟云网络(VCN)的服务和资源。在某些实施例中,CSPI 101向订阅客户供应IaaS服务。CSPI 101内的数据中心可以被组织成一个或多个区域。一个示例区域“区域US”102在图1中示出。客户已经为区域102配置了客户VCN 104。客户可以在VCN 104上部署各种计算实例,其中计算实例可以包括虚拟机或裸机实例。实例的示例包括应用、数据库、负载均衡器等。

[0069] 在图1中描绘的实施例中,客户VCN 104包括两个子网,即“子网-1”和“子网-2”,每个子网具有其自己的CIDR IP地址范围。在图1中,子网-1的覆盖IP地址范围是10.0/16并且子网-2的地址范围是10.1/16。VCN虚拟路由器105表示VCN的逻辑网关,其使能VCN 104的子网之间以及与VCN外部的其他端点的通信。VCN VR 105被配置为在VCN 104中的VNIC和与VCN 104相关联的网关之间路由流量。VCN VR 105为VCN 104的每个子网提供端口。例如,VR 105可以为子网-1提供具有IP地址10.0.0.1的端口并且为子网-2提供具有IP地址10.1.0.1的端口。

[0070] 多个计算实例可以部署在每个子网上,其中计算实例可以是虚拟机实例和/或裸机实例。子网中的计算实例可以由CSPI 101内的一个或多个主机托管。计算实例经由与计算实例相关联的VNIC加入子网。例如,如图1中所示,计算实例C1经由与计算实例相关联的VNIC来作为子网-1的一部分。同样,计算实例C2经由与C2相关联的VNIC来作为子网-1的一部分。以类似的方式,多个计算实例(其可以是虚拟机实例或裸机实例)可以是子网-1的一部分。经其关联的VNIC,每个计算实例被指配私有覆盖IP地址和MAC地址。例如,在图1中,计算实例C1具有覆盖IP地址10.0.0.2以及MAC地址M1,而计算实例C2具有私有覆盖IP地址10.0.0.3以及MAC地址M2。子网-1中的每个计算实例(包括计算实例C1和C2)具有到使用IP地址10.0.0.1的VCN VR 105的默认路由,该IP地址是VCN VR 105针对子网-1的端口的IP地址。

[0071] 子网-2可以具有部署在其上的多个计算实例,包括虚拟机实例和/或裸机实例。例如,如图1中所示,计算实例D1和D2经由与相应计算实例相关联的VNIC来作为子网-2的一部分。在图1中描绘的实施例中,计算实例D1具有覆盖IP地址10.1.0.2以及MAC地址MM1,而计算实例D2具有私有覆盖IP地址10.1.0.3以及MAC地址MM2。子网-2中的每个计算实例(包括计算实例D1和D2)具有到使用IP地址10.1.0.1的VCN VR 105的默认路由,该IP地址是VCN VR 105针对子网-2的端口的IP地址。

[0072] VCN A 104还可以包括一个或多个负载平衡器。例如,负载平衡器可以被提供用于子网并且可被配置为跨子网上的多个计算实例来负载平衡流量。负载平衡器还可以被提供来对跨VCN中的子网来负载平衡流量。

[0073] 部署在VCN 104上的特定计算实例可以与各种不同的端点通信。这些端点可以包括由CSPI 200托管的端点和CSPI 200外部的端点。由CSPI 101托管的端点可以包括:与特定计算实例在相同子网上的端点(例如,子网-1中的两个计算实例之间的通信);在不同子网上但在相同VCN内的端点(例如,子网-1中的计算实例与子网-2中的计算实例之间的通信);相同区域中不同VCN中的端点(例如,子网-1中的计算实例与相同区域中的VCN 106或110中的端点之间的通信、子网-1中的计算实例与相同区域中的服务网络110中的端点之间的通信);或者不同区域中的VCN中的端点(例如,子网-1中的计算实例与不同区域中的VCN108中的端点之间的通信)。由CSPI 101托管的子网中的计算实例还可以与不由CSPI 101托管(即,在CSPI 101外部)的端点通信。这些外部端点包括客户的本地部署网络116中的端点、其他远程云托管网络118内的端点、可经由诸如互联网的公共网络访问的公共端点114以及其他端点。

[0074] 使用与源计算实例和目的地计算实例相关联的VNIC来促进相同子网上的计算实例之间的通信。例如,子网-1中的计算实例C1可能想要将分组发送到子网-1中的计算实例C2。对于源自源计算实例处且其目的地是相同子网中的另一计算实例的分组,分组首先由与源计算实例相关联的VNIC处理。由与源计算实例相关联的VNIC执行的处理可以包括从分组报头确定分组的目的地信息、识别为与源计算实例相关联的VNIC配置的任何策略(例如,安全列表)、确定分组的下一跳,根据需要执行任何分组封装/解封装功能,然后将分组转发/路由到下一跳,目的是促进分组到其预期目的地的通信。当目的地计算实例与源计算实例在相同子网中时,与源计算实例相关联的VNIC被配置为识别与目的地计算实例相关联的VNIC,并将分组转发到该VNIC以供处理。然后与目的地计算实例相关联的VNIC被执行并将分组转发到目的地计算实例。

[0075] 对于要从子网中的计算实例传达到相同VCN中的不同子网中的端点的分组,由与源和目的地计算实例相关联的VNIC以及VCN VR来促进通信。例如,如果图1中子网-1中的计算实例C1想要向子网-2中的计算实例D1发送分组,则分组首先由与计算实例C1相关联的VNIC处理。与计算实例C1相关联的VNIC被配置为使用默认路由或VCN VR的端口10.0.0.1将分组路由到VCN VR 105。VCN VR 105被配置为使用端口10.1.0.1将分组路由到子网-2。然后,分组由与D1相关联的VNIC接收并处理,并且VNIC将分组转发到计算实例D1。

[0076] 对于要从VCN 104中的计算实例传达到VCN 104外部的端点的分组,由与源计算实例相关联的VNIC、VCN VR 105以及与VCN 104相关联的网关来促进通信。一个或多个类型的网关可以与VCN 104相关联。网关是VCN和另一端点之间的接口,其中另一端点在VCN外部。

网关是层-3/IP层概念并且使得VCN能够与VCN外部的端点通信。因此,网关促进VCN与其他VCN或网络之间的流量流。各种不同类型的网关可被配置用于VCN以促进与不同类型的端点的不同类型的通信。取决于网关,通信可以通过公共网络(例如,互联网)或通过私有网络。各种通信协议可以用于这些通信。

[0077] 例如,计算实例C1可能想要与VCN 104外部的端点通信。分组可以首先由与源计算实例C1相关联的VNIC处理。VNIC处理确定分组的目的地在C1的子网-1外部。与C1相关联的VNIC可以将分组转发到VCN 104的VCN VR 105。VCN VR 105然后处理分组并且作为处理的一部分,基于分组的目的地,确定与VCN 104相关联的特定网关作为分组的下一跳。VCN VR 105然后可以将分组转发到特定的识别的网关。例如,如果目的地是客户本地部署网络内的端点,则分组可以由VCN VR 105转发到为VCN 104配置的动态路由网关(DRG)网关122。然后分组可以从网关被转发到下一跳以促进分组的到其最终的预期目的地的通信。

[0078] 可以为VCN配置各种不同类型的网关。可以为VCN配置的网关的示例在图1中被描绘并在下面描述。与VCN相关联的网关的示例也在图16、17、18和19中被描绘(例如,由附图标记1634、1636、1638、1734、1736、1738、1834、1836、1838、1934、1936和1938所引用的网关)并在下面描述。如图1中描绘的实施例所示,动态路由网关(DRG) 122可以被添加到客户VCN 104或与客户VCN 104相关联,并且为客户VCN 104和另一端点之间的私有网络流量通信提供路径,其中另一端点可以是客户的本地部署网络116、CSPI 101的不同区域中的VCN 108、或不由CSPI 101托管的其他远程云网络118。客户本地部署网络116可以是使用客户的资源构建的客户网络或客户数据中心。对客户本地部署网络116的访问通常非常受限。对于具有客户本地部署网络116和由CSPI 101在云中部署或托管的一个或多个VCN 104两者的客户,客户可能希望他们的本地部署网络116和他们的基于云的VCN 104能够彼此通信。这使得客户能够构建包含由CSPI 101托管的客户的VCN 104和他们的本地部署网络116的延伸混合环境。DRG 122使能此通信。为了使能这样的通信,通信信道124被建立,其中信道的一个端点在客户本地部署网络116中,并且另一端点在CSPI 101中且连接到客户VCN 104。通信信道124可以通过诸如互联网的公共通信网络或私有通信网络。可以使用诸如通过公共通信网络(诸如互联网)的IPsec VPN技术、使用私有网络而不是公共网络的Oracle的FastConnect技术等等的各种不同的通信协议。形成通信信道124的一个端点的客户本地部署网络116中的设备或装备被称为客户驻地装备(CPE),诸如图1中描绘的CPE 126。CSPI 101侧,端点可以是执行DRG 122的主机。

[0079] 在某些实施例中,远程对等互联连接(RPC)可被添加到DRG,这允许客户将一个VCN与不同区域中的另一VCN对等。使用这样的RPC,客户VCN 104可以使用DRG 122来与另一区域中的VCN 108连接。DRG 122还可以用于与诸如微软Azure云、亚马逊AWS云等的不由CSPI 101托管的其他远程云网络118通信。

[0080] 如图1中所示,互联网网关(IGW) 120可以被配置用于客户VCN 104,其使得VCN 104上的计算实例能够与可通过诸如互联网的公共网络访问的公共端点114通信。IGW 120是将VCN连接到诸如互联网的公共网络的网关。IGW 120使得诸如VCN 104的VCN内的公共子网(其中公共子网中的资源具有公共覆盖IP地址)能够直接访问诸如互联网的公共网络114上的公共端点112。使用IGW 120,连接可以从VCN 104内的子网或从互联网开始。

[0081] 网络地址转换(NAT)网关128可以被配置用于客户的VCN 104并且使得客户的VCN

中不具有专用公共覆盖IP地址的云资源能够访问互联网,并且NAT网关128在不将那些资源暴露到直接传入互联网连接(例如,L4-L7连接)的情况下这样做。这使得VCN内的私有子网(诸如VCN 104中的私有子网-1)能够对互联网上的公共端点私有访问。在NAT网关中,只能从私有子网到公共互联网开始连接,而不能从互联网到私有子网开始连接。

[0082] 在某些实施例中,服务网关(SGW)126可以被配置用于客户VCN 104并且为VCN 104和服务网络110中的支持的服务端点之间的私有网络流量提供路径。在某些实施例中,服务网络110可以由CSP提供并且可以提供各种服务。这样的服务网络的示例是Oracle的服务网络,其提供可由客户使用的各种服务。例如,客户VCN 104的私有子网中的计算实例(例如,数据库系统)可以将数据备份到服务端点(例如,对象存储),而不需要公共IP地址或对互联网的访问。在某些实施例中,VCN可以仅具有一个SGW,并且连接只能从VCN内的子网开始而不能从服务网络110开始。如果VCN与另一VCN对等,则另一个VCN中的资源通常不可访问SGW。通过FastConnect或VPN Connect连接到VCN的本地部署网络中的资源也可以使用为该VCN配置的服务网关。

[0083] 在某些实现中,SGW 126使用服务无类别域间路由(CIDR)标签的概念,其是表示感兴趣的服务或服务组的所有区域性的公共IP地址范围的字符串。客户在配置SGW和相关路由规则时使用服务CIDR标签来控制到服务的流量。客户可以在配置安全规则时选择使用服务CIDR标签从而如果服务的公共IP地址在将来更改,无需调整他们。

[0084] 本地对等互联网网关(LPG)132是可以被添加到客户VCN 104并且使得VCN 104能够与相同区域中的另一VCN对等的网关。对等互联意味着VCN使用私有IP地址通信,而无需流量穿过诸如互联网的公共网络,或者无需通过客户的本地部署网络116路由流量。在优选实施例中,VCN对于其建立的每个对等互联具有分开的LPG。本地对等互联或VCN对等互联是用于在不同应用或基础设施管理功能之间建立网络连接性的常见实践。

[0085] 服务提供商(诸如服务网络110中的服务的提供商),可以使用不同的访问模型来提供对服务的访问。根据公共访问模型,服务可以被暴露为公共端点,这些端点可由客户VCN中的计算实例经由诸如互联网的公共网络公开访问和/或可以经由SGW 126私密地访问。根据特定的私有访问模型,使得服务可作为客户的VCN中的私有子网中的私有IP端点来访问。这被称为私有端点(PE)访问并且使得服务提供商能够将其服务暴露为客户的私有网络中的实例。私有端点资源表示客户的VCN内的服务。每个PE表现为由客户在客户的VCN中选择的子网中的VNIC(被称为PE-VNIC,具有一个或多个私有IP)。因此,PE提供了一种使用VNIC在私有客户VCN子网内呈现服务的方法。由于端点被暴露为VNIC,因此与VNIC相关的所有特征(诸如路由规则、安全列表等)现在可用于PE VNIC。

[0086] 服务提供商可以注册他们的服务以能通过PE的访问。提供商可以将策略与服务相关联,这限制了服务对客户租赁的可见性。提供商可以在单个虚拟IP地址(VIP)下注册多个服务,尤其是对于多租户服务。可以有表示相同的服务的多个这样的私有端点(在多个VCN中)。

[0087] 私有子网中的计算实例然后可以使用PE VNIC的私有IP地址或服务DNS名称来访问服务。客户VCN中的计算实例可以通过将流量发送到客户VCN中PE的私有IP地址来访问服务。私有访问网关(PAGW)130是可以附接到服务提供商VCN(例如,服务网络110中的VCN)的网关资源,其充当来自/去往客户子网私有端点的所有流量的入口/出口点。PAGW 130使得

提供商能够在不利用其内部IP地址资源的情况下扩展PE连接的数量。提供商只需为在单个VCN中注册的任何数量的服务配置一个PAGW。提供商可以将服务表示为一个或多个客户的多个VCN中的私有端点。从客户的角度来看,PE VNIC不是附接到客户的实例,而是看起来附接到客户希望与之交互的服务。经由PAGW 130将目的地为私有端点的流量路由至服务。这些被称为客户到服务私有连接(C2S连接)。

[0088] PE概念还可以用于通过允许流量流过FastConnect/IPsec链路和客户VCN中的私有端点来将服务的私有访问延伸到客户的本地部署网络和数据中心。通过允许流量在客户的VCN中的LPG 132和PE之间流动,服务的私有访问还可被延伸到客户的对等VCN。

[0089] 客户可以在子网级别控制VCN中的路由,因此客户可以指定客户的VCN(诸如VCN 104)中的哪些子网使用每个网关。VCN的路由表用于决定流量是否被允许通过特定网关离开VCN。例如,在特定实例中,客户VCN 104内的公共子网的路由表可以通过IGW 120发送非本地流量。相同客户VCN 104内的私有子网的路由表可以通过SGW 126发送目的地为CSP服务的流量。可以经由NAT网关128发送所有剩余流量。路由表仅控制出VCN的流量。

[0090] 与VCN相关联的安全列表用于控制经由入站连接经由网关进入VCN的流量。子网中的所有资源使用相同的路由表和安全列表。安全列表可用于控制允许进出VCN子网中的实例的特定类型的流量。安全列表规则可以包括入口(入站)和出口(出站)规则。例如,入口规则可以指定允许的源地址范围,而出口规则可以指定允许的目的地地址范围。安全规则可以指定特定协议(例如,TCP、ICMP)、特定端口(例如,用于SSH的22,用于Windows RDP的3389)等。在某些实现中,实例的操作系统可以强制执行与安全列表规则一致的其自己的防火墙规则。规则可以是有状态的(例如,在没有针对响应流量的明确的安全列表规则的情况下,连接被跟踪并且响应被自动允许)或无状态的。

[0091] 来自客户VCN的访问(即,通过部署在VCN 104上的资源或计算实例)可以被分类为公共访问、私有访问或专用访问。公共访问是指其中公共IP地址或NAT被用来访问公共端点的访问模型。私有访问使得具有私有IP地址的VCN 104中的客户工作负载(例如,私有子网中的资源)能够在不穿过诸如互联网的公共网络的情况下访问服务。在某些实施例中,CSPI 101使得具有私有IP地址的客户VCN工作负载能够使用服务网关来访问服务(的公共服务端点)。因此,服务网关通过在客户的VCN和驻留在客户的私有网络外部的服务的公共端点之间建立虚拟链路来供应私有访问模型。

[0092] 另外,CSPI可以使用诸如FastConnect公共对等互联的技术来供应专用公共访问,其中客户本地部署实例可以使用FastConnect连接来访问客户VCN中的一项或多项服务,而不穿过诸如互联网的公共网络。CSPI还可以使用FastConnect私有对等互联来供应专用私有访问,其中具有私有IP地址的客户本地部署实例可以使用FastConnect连接来访问客户的VCN工作负载。FastConnect是对使用公共互联网将客户的本地部署网络连接到CSPI及其服务的网络连接性替代方案。当与基于互联网的连接相比时,FastConnect提供了一种简单、弹性和经济的方式来创建具有更高的带宽选项和更可靠且一致的联网体验的专用和私有连接。

[0093] 图1和上面所附的描述描述了示例虚拟网络中的各种虚拟化组件。如上所述,虚拟网络构建在底层物理或基底网络之上。图2描绘了根据某些实施例的CSPI 200内的物理网络中的为虚拟网络提供底层的物理组件的简化架构图。如图所示,CSPI 200提供包括由云

服务提供商 (CSP) 提供的组件和资源 (例如, 计算、存储器和联网资源) 的分布式环境。这些组件和资源用于向订阅客户 (即, 已经订阅了由 CSP 提供的一项或多项服务的客户) 提供云服务 (例如, IaaS 服务)。基于客户订阅的服务, CSPI 200 的资源 (例如, 计算、存储器和联网资源) 子集被供给到客户。然后, 客户可以使用 CSPI 200 提供的物理计算、存储器和联网资源构建他们自己的基于云 (即, CSPI 托管的) 的可定制和私有虚拟网络。如前所述, 这些客户网络被称为虚拟云网络 (VCN)。客户可以在这些客户 VCN 上部署一个或多个客户资源, 诸如计算实例。计算实例可以是虚拟机、裸机实例等形式。CSPI 200 提供基础设施和一组互补的云服务, 其使得客户能够在高度可用的托管环境中构建和运行广泛的应用和服务。

[0094] 在图2中描绘的示例实施例中, CSPI 200 的物理组件包括一个或多个物理主机或物理服务器 (例如, 202、206、208)、网络虚拟化设备 (NVD) (例如, 210、212)、架顶式 (TOR) 交换机 (例如, 214、216) 和物理网络 (例如, 218) 以及物理网络 218 中的交换机。物理主机或服务器可以托管和执行加入 VCN 的一个或多个子网的各种计算实例。计算实例可以包括虚拟机实例和裸机实例。例如, 图1中描绘的各种计算实例可以由图2中描绘的物理主机托管。VCN 中的虚拟机计算实例可以由一个主机或由多个不同的主机执行。物理主机还可以托管虚拟主机、基于容器的主机或功能等。图1中描绘的 VNIC 和 VCN VR 可以由图2中描绘的 NVD 执行。图1中描绘的网关可以由图2中描绘的主机和/或 NVD 执行。

[0095] 主机或服务器可以执行管理程序 (也被称为虚拟机监视器或 VMM), 其在主机上创建并使能虚拟化环境。虚拟化或虚拟化环境促进基于云的计算。一个或多个计算实例可以由主机上的管理程序在该主机上创建、执行和管理。主机上的管理程序使得主机的物理计算资源 (例如, 计算、存储器和联网资源) 能够在由主机执行的各种计算实例之间共享。

[0096] 例如, 如图2中描绘的, 主机 202 和 208 分别执行管理程序 260 和 266。这些管理程序可以使用软件、固件或硬件或其组合来实现。通常, 管理程序是位于主机操作系统 (OS) 之上的进程或软件层, 该操作系统继而在主机的硬件处理器上执行。管理程序通过使得主机的物理计算资源 (例如, 诸如处理器/核的处理资源、存储器资源、联网资源) 能够在由主机执行的各种虚拟机计算实例之间共享来提供虚拟化环境。例如, 在图2中, 管理程序 260 可以位于主机 202 的 OS 之上并且使得主机 202 的计算资源 (例如, 处理、存储器和联网资源) 能够在由主机 202 执行的计算实例 (例如, 虚拟机) 之间共享。虚拟机可以具有其自己的操作系统 (称为访客操作系统), 其可以与主机的 OS 相同或不同。主机执行的虚拟机的操作系统可以与相同主机执行的另一个虚拟机的操作系统相同或不同。因此, 管理程序使得多个操作系统能够彼此同时执行, 同时共享主机的相同计算资源。图2中描绘的主机可以具有相同或不同类型的管理程序。

[0097] 计算实例可以是虚拟机实例或裸机实例。在图2中, 主机 202 上的计算实例 268 和主机 208 上的计算实例 274 是虚拟机实例的示例。主机 206 是提供给客户的裸机实例的示例。

[0098] 在某些实例中, 整个主机可以被供给到单个客户, 并且由该主机托管的所有一个或多个计算实例 (虚拟机或裸机实例) 属于该相同客户。在其他实例中, 主机可以在多个客户 (即, 多个租户) 之间共享。在这样的多租赁场景中, 主机可以托管属于不同客户的虚拟机计算实例。这些计算实例可能是不同客户的不同 VCN 的成员。在某些实施例中, 裸机计算实例由没有管理程序的裸机服务器托管。当裸金属计算实例被供给时, 单个客户或租户维持对托管裸机实例的主机的物理 CPU、存储器和网络接口的控制, 并且主机不与其他客户或租

户共享。

[0099] 如前所述,作为VCN的一部分的每个计算实例与VNIC相关联,该VNIC使得计算实例能够成为VCN的子网的成员。与计算实例相关联的VNIC促进进出计算实例的分组或帧的通信。当计算实例被创建时,VNIC与计算实例相关联。在某些实施例中,对于由主机执行的计算实例,与该计算实例相关联的VNIC由连接到主机的NVD执行。例如,在图2中,主机202执行与VNIC 276相关联的虚拟机计算实例268,并且VNIC 276由连接到主机202的NVD 210执行。作为另一示例,由主机206托管的裸机实例272与由连接到主机206的NVD 212执行VNIC 280相关联。作为又另一示例,VNIC 284与主机208执行的计算实例274相关联,并且VNIC 284由连接到主机208的NVD 212执行。

[0100] 对于由主机托管的计算实例,连接到该主机的NVD还执行与计算实例是其成员的VCN相对应的VCN VR。例如,在图2中描绘的实施例中,NVD 210执行与计算实例268是其成员的VCN相对应的VCN VR 277。NVD 212还可以执行与对应于由主机206和208托管的计算实例的VCN相对应的一个或多个VCN VR 283。

[0101] 主机可以包括一个或多个网络接口卡(NIC),其使得主机能够连接到其他设备。主机上的NIC可以提供使得主机能够通信地连接到另一设备的一个或多个端口(或接口)。例如,主机可以使用提供在主机和NVD上的一个或多个端口(或接口)连接到NVD。主机还可以连接到其他设备,诸如另一主机。

[0102] 例如,在图2中,主机202使用在由主机202的NIC 232提供的端口234与NVD 210的端口236之间延伸的链路220连接到NVD 210。主机206使用在由主机206的NIC 244提供的端口246和NVD 212的端口248之间延伸的链路224连接到NVD 212。主机208使用在由主机208的NIC 250提供的端口252和NVD 212的端口254之间延伸的链路226连接到NVD 212。

[0103] NVD继而经由通信链路连接至架顶式(TOR)交换机,该交换机连接至物理网络218(也称为交换机结构(switch fabric))。在某些实施例中,主机和NVD之间以及NVD和TOR交换机之间的链路是以太网链路。例如,在图2中,NVD 210和212分别使用链路228和230连接到TOR交换机214和216。在某些实施例中,链路220、224、226、228和230是以太网链路。连接到TOR的主机和NVD的集合有时被称为机架。

[0104] 物理网络218提供使得TOR交换机能够彼此通信的通信结构。物理网络218可以是多层级网络。在某些实现中,物理网络218是交换机的多层级Clos网络,其中TOR交换机214和216表示多层级和多节点物理交换网络218的叶级节点。不同的Clos网络配置是可能的,包括但不限于2层级网络、3层级网络、4层级网络、5层级网络以及一般的“n”层级网络。Clos网络的示例在图5中被描绘并在下面描述。

[0105] 主机和NVD之间诸如一对一配置、多对一配置、一对多配置等的各种不同的连接配置是可能的。在一对一配置实现中,每个主机连接到其自己分开的NVD。例如,在图2中,主机202经由主机202的NIC 232连接到NVD 210。在多对一配置中,多个主机连接到一个NVD。例如,在图2中,主机206和208分别经由NIC 244和250连接到同一NVD 212。

[0106] 在一对多配置中,一个主机连接到多个NVD。图3示出了CSPI 300内的示例,其中主机302连接到多个NVD。如图3中所示,主机302包括网络接口卡(NIC) 304,其包括多个端口306和308。主机302经由端口306和链路320连接到第一NVD 310,并经由端口308和链路322连接到第二NVD 312。端口306和308可以是以太网端口,并且主机302与NVD 310和312之间的链路

320和322可以是以以太网链路。NVD 310继而连接到第一TOR交换机314,并且NVD 312连接到第二TOR交换机316。NVD 310和312以及TOR交换机314和316之间的链路可以是以以太网链路。TOR交换机314和316表示多层级物理网络318中的层级0交换设备。

[0107] 图3中描绘的布置提供了进出物理交换机网络318到主机302的两条分开的物理网络路径:穿过TOR交换机314到NVD 310到主机302的第一路径,以及穿过TOR交换机316到NVD 312到主机302的第二路径。分开的路径提供了主机302的增强的可用性(称为高可用性)。如果在路径之一中有问题(例如,路径之一中的链路出故障)中或设备之一中有问题(例如,特定NVD不起作用),则另一个路径可以用于进/出主机302的通信。

[0108] 在图3中描绘的配置中,主机使用主机的NIC提供的两个不同端口连接到两个不同的NVD。在其他实施例中,主机可以包括使能主机到多个NVD的连接性的多个NIC。

[0109] 返回参考图2,NVD是执行一个或多个网络和/或存储虚拟化功能的物理设备或组件。NVD可以是具有一个或多个处理单元(例如,CPU、网络处理单元(NPU)、FPGA、分组处理管道等)、包括高速缓存的存储器和端口的任何设备。各种虚拟化功能可以由NVD的一个或多个处理单元执行的软件/固件来执行。

[0110] NVD可以以各种不同的形式来实现。例如,在某些实施例中,NVD被实现为称为smartNIC或智能NIC的接口卡,搭载有嵌入式处理器。smartNIC是与主机上的NIC分开的设备。在图2中,NVD 210和212可以被实现为分别连接到主机202以及主机206和208的smartNIC。

[0111] 然而,smartNIC只是NVD实现的一个示例。各种其他实现是可能的。例如,在一些其他实现中,NVD或由NVD执行的一项或多项功能可以并入到一个或多个主机、一个或多个TOR交换机以及CSPI 200的其他组件中或由其执行。例如,NVD可以体现在主机中,其中由NVD执行的功能由主机执行。作为另一示例,NVD可以是TOR交换机的一部分,或者TOR交换机可以被配置为执行由NVD执行的功能,该功能使得TOR交换机能够执行用于公共云的各种复杂的分组转换。执行NVD功能的TOR有时被称为智能TOR。在其中虚拟机(VM)实例而不是裸机(BM)实例被供应给客户的又其他实现中,由NVD执行的功能可以在主机的管理程序内部实现。在一些其他实现中,NVD的功能中的一些可以被卸载到在一系列主机上运行的集中式服务。

[0112] 在某些实施例中,诸如当实现为如图2所示的smartNIC时,NVD可以包括使得该NVD能够连接到一个或多个主机以及一个或多个TOR交换机的多个物理端口。NVD上的端口可以分类为面向主机的端口(也称为“南端口”)或面向网络或面向TOR的端口(也称为“北端口”)。NVD的面向主机的端口是用于将NVD连接到主机的端口。图2中面向主机的端口的示例包括NVD 210上的端口236以及NVD 212上的端口248和254。NVD的面向网络的端口是用于将NVD连接到TOR交换机的端口。图2中面向网络的端口的示例包括NVD 210上的端口256和NVD 212上的端口258。如图2中所示,NVD 210使用从NVD 210的端口256延伸到TOR交换机214的链路228连接到TOR交换机214。同样,NVD 212使用从NVD 212的端口258延伸到TOR交换机216的链路230连接到TOR交换机216。

[0113] NVD经由面向主机的端口从主机接收分组和帧(例如,由主机托管的计算实例生成的分组和帧),并且在执行必要的分组处理之后,可以经由NVD的面向网络的端口将分组和帧转发到TOR交换机。NVD可以经由NVD的面向网络的端口从TOR交换机接收分组和帧,并且在执行必要的分组处理之后,可以经由NVD的面向主机的端口将分组和帧转发到主机。

[0114] 在某些实施例中,NVD和TOR交换机之间可以有多个端口和相关联的链路。这些端口和链路可以聚合以形成多个端口或链路的链路聚合器组(称为LAG)。链路聚合允许两个端点之间(例如,NVD和TOR交换机之间)的多个物理链路被视为单个逻辑链路。给定LAG中的所有物理链路可以以相同的速度在全双工模式中操作。LAG有助于提高两个端点之间连接的带宽和可靠性。如果LAG中的物理链路之一出故障,则流量动态地且透明地重新指配给LAG中的其他物理链路之一。聚合的物理链路比每个单独的链路交付更高的带宽。与LAG相关联的多个端口被视为单个逻辑端口。流量可以被跨LAG的多个物理链路来负载平衡。在两个端点之间可以配置一个或多个LAG。两个端点可以位于NVD和TOR交换机之间、主机和NVD之间等。

[0115] NVD实现或执行网络虚拟化功能。这些功能由NVD执行的软件/固件来执行。网络虚拟化功能的示例包括但不限于:分组封装和解封装功能;用于创建VCN网络的功能;用于实现诸如VCN安全列表(防火墙)功能的网络策略的功能;促进路由和转发进出VCN中的计算实例的分组的功能;等等。在某些实施例中,在接收到分组后,NVD被配置为执行分组处理流水线,用于处理分组并确定分组被如何转发或路由。作为此分组处理流水线的一部分,NVD可以执行与覆盖网络相关联的一个或多个虚拟功能,诸如执行与VCN中的计算实例相关联的VNIC、执行与VCN相关联的虚拟路由器(VR)、分组的封装和解封装以促进在虚拟网络中转发或路由、某些网关(例如,本地对等互联网关)的执行、安全列表、网络安全组、网络地址转换(NAT)功能(例如,在逐个主机基础上的公共IP到私有IP的转换)、节流功能和其他功能的实现。

[0116] 在某些实施例中,NVD中的分组处理数据路径可以包括多个分組管道,每个分組管道由一系列分組变换阶段组成。在某些实现中,在接收到分組后,该分組被解析并分类到单个流水线。然后分組以线性方式处理,一个阶段又一个阶段,直到分組被丢弃或通过NVD的接口发送出去。这些阶段提供基本的功能分組处理构建块(例如,验证报头、施行节流、插入新的层2报头、施行L4防火墙、VCN封装/解封装等),使得新的流水线可以通过组合现有阶段来构造,并且新功能可以通过创建新阶段并将其插入现有的流水线来添加。

[0117] NVD可以执行与VCN的控制平面和数据平面相对应的控制平面和数据平面功能两者。VCN控制平面的示例也在图16、17、18和19(参见附图标记1616、1716、1816和1916)中被描绘并在下面描述。VCN数据平面的示例在图16、17、18和19(参见附图标记1618、1718、1818和1918)中被描绘并描述如下。控制平面功能包括用于配置控制数据被如何转发的网络的功能(例如,建立路由和路由表、配置VNIC等)。在某些实施例中,VCN控制平面被提供,其集中计算所有覆盖到基底映射并将它们发布到NVD以及到诸如各种网关(诸如DRG、SGW、IGW等)的虚拟网络边缘设备。防火墙规则也可以使用相同的机制来发布。在某些实施例中,NVD仅获取与该NVD相关的映射。数据平面功能包括用于基于使用控制平面建立的配置的针对分组的实际的路由/转发的功能。VCN数据平面是通过在客户的网络分組穿过底层网络之前对其封装来实现的。封装/解封装功能在NVD上实现。在某些实施例中,NVD被配置为拦截进出主机的所有网络分組并执行网络虚拟化功能。

[0118] 如上所述,NVD执行包括VNIC和VCN VR的各种虚拟化功能。NVD可以执行与由连接到VNIC的一台或多台主机托管的计算实例相关联的VNIC。例如,如图2中描绘的,NVD 210执行与由连接到NVD 210的主机202托管的计算实例268相关联的VNIC 276的功能。作为另一

示例,NVD 212执行与由主机206托管的裸机计算实例272相关联的VNIC 280,并执行与由主机208托管的计算实例274相关联的VNIC 284。主机可以托管属于不同VCN的计算实例,这些VCN属于不同的客户,并且连接到主机的NVD可以执行与计算实例相对应的VNIC (即,执行VNIC相关的功能)。

[0119] NVD还执行与计算实例的VCN相对应的VCN虚拟路由器。例如,在图2中所描绘的实施例中,NVD 210执行与计算实例268所属于的VCN相对应的VCN VR 277。NVD 212执行与主机206和208托管的计算实例所属于的一个或多个VCN相对应的一个或多个VCN VR 283。在某些实施例中,对应于该VCN的VCN VR由连接到托管属于该VCN的至少一个计算实例的主机的所有NVD执行。如果主机托管属于不同VCN的计算实例,则连接到该主机的NVD可以执行与那些不同VCN相对应的VCN VR。

[0120] 除了VNIC和VCN VR之外,NVD可以执行各种软件 (例如,守护进程) 并且包括促进由NVD执行的各种网络虚拟化功能的一个或多个硬件组件。为了简单起见,这些各个组件被分组在一起作为图2中所描绘的“分组处理组件”。例如,NVD 210包括分组处理组件286,并且NVD 212包括分组处理组件288。例如,NVD的分组处理组件可以包括分组处理器,其被配置为与NVD的端口和硬件接口交互以监控由NVD接收和使用NVD传达的所有分组并存储网络信息。网络信息可以例如包括标识由NVD处理的不同网络流的网络流信息和每条流信息 (例如,每条流统计)。在某些实施例中,网络流信息可以在每个VNIC的基础上来存储。分组处理器可以执行逐个分组的操作以及实现有状态NAT和L4防火墙 (FW)。作为另一示例,分组处理组件可以包括复制代理,其被配置为将由NVD存储的信息复制到一个或多个不同的复制目标存储。作为又另一示例,分组处理组件可以包括被配置为执行NVD的日志记录功能的日志记录代理。分组处理组件还可以包括用于监控NVD的性能和健康并且还可能监控连接到NVD的其他组件的状态和健康的软件。

[0121] 图1示出了示例虚拟或覆盖网络的组件,包括VCN、VCN内的子网、子网上部署的计算实例、与计算实例相关联的VNIC、VCN的VR以及为VCN配置的一组网关。图1中描绘的覆盖组件可以由图2中描绘的一个或多个物理组件来执行或托管。例如,VCN中的计算实例可以由图2中描绘的一个或多个主机执行或托管。对于由主机托管的计算实例,与该计算实例相关联的VNIC通常由连接到该主机的NVD执行 (即,VNIC功能由连接到该主机的NVD提供)。VCN的VCN VR功能由连接到托管或执行是该VCN的一部分的计算实例的主机的所有NVD执行。与VCN相关联的网关可以由一个或多个不同类型的NVD来执行。例如,某些网关可以由smartNIC执行,而其他可以由一个或多个主机或NVD的其他实现来执行。

[0122] 如上所述,客户VCN中的计算实例可以与各种不同的端点通信,其中端点可以在与源计算实例相同的子网内、在不同的子网中但在与源计算实例相同的VCN内,或与源计算实例的VCN外部的端点通信。这些通信使用与计算实例相关联的VNIC、VCN VR以及与VCN相关联的网关来促进。

[0123] 对于VCN中相同子网上的两个计算实例之间的通信,该通信使用与源和目的地计算实例相关联的VNIC来促进。源和目的地计算实例可以由相同主机或不同主机托管。源自源计算实例的分组可以从托管源计算实例的主机转发到连接到该主机的NVD。在NVD上,分组被使用分组处理流水线来处理,该分组处理流水线可以包括与源计算实例相关联的VNIC的执行。由于分组的目的地端点位于相同子网内,因此与源计算实例相关联的VNIC的执行

导致分组被转发到执行与目的地计算实例相关联的VNIC的NVD,该VNIC然后处理分组并将该分组转发到目的地计算实例。与源和目的地计算实例相关联的VNIC可以在相同NVD上执行(例如,当源和目的地计算实例两者由相同主机托管时)或不同NVD上执行(例如,当源和目的地计算实例由连接到不同NVD的不同主机托管)。VNIC可以使用NVD存储的路由/转发表来确定分组的下一跳。

[0124] 对于要从子网中的计算实例传达到相同VCN中的不同子网中的端点的分组,源自源计算实例的分组从托管源计算实例的主机传达到连接到该主机的NVD。在NVD上,使用分组处理流水线来处理分组,其可以包括一个或多个VNIC以及与VCN相关联的VR的执行。例如,作为分组处理流水线的一部分,NVD执行或调用对应于与源计算实例相关联的VNIC的功能(也称为执行VNIC)。VNIC执行的功能可以包括查看分组上的VLAN标记。由于分组的目的地在子网外部,因此VCN VR功能接下来由NVD调用并执行。然后,VCN VR将分组路由到执行与目的地计算实例相关联的VNIC的NVD。然后,与目的地计算实例相关联的VNIC处理分组并将分组转发到目的地计算实例。与源和目的地计算实例相关联的VNIC可以在相同NVD上执行(例如,当源和目的地计算实例两者由相同主机托管时)或不同NVD上执行(例如,当源和目的地计算实例由连接到不同NVD的不同主机托管)。

[0125] 如果分组的目的地在源计算实例的VCN外部,则源自源计算实例的分组从托管源计算实例的主机传达到连接到该主机的NVD。NVD执行与源计算实例相关联的VNIC。由于分组的目的地端点在VCN外部,因此分组然后由该VCN的VCN VR处理。NVD调用VCN VR功能,这可以导致分组被转发到执行与VCN相关联的适当网关的NVD。例如,如果目的地是客户本地部署网络内的端点,则分组可以由VCN VR转发到执行为VCN配置的DRG网关的NVD。VCN VR可以在与执行与源计算实例相关联的VNIC的NVD相同的NVD上执行或者由不同的NVD执行。网关可以由NVD执行,该NVD可以是smartNIC、主机或其他NVD实现。然后,分组由网关处理并转发到促进分组到其预期的目的地端点的通信的下一跳。例如,在图2中所描绘的实施例中,源自计算实例268的分组可以通过链路220(使用NIC 232)从主机202传达到NVD 210。在NVD 210上,VNIC 276被调用,因为它是与源计算实例268相关联的VNIC。VNIC 276被配置为检查分组中的封装信息,并确定用于转发分组的下一跳,目的是促进分组到其预期的目的地端点的通信,然后将分组转发到确定的下一跳。

[0126] 部署在VCN上的计算实例可以与各种不同的端点通信。这些端点可以包括由CSPI 200托管的端点和CSPI 200外部的端点。由CSPI 200托管的端点可以包括相同VCN或其他VCN中的实例,其可以是客户的VCN或不属于客户的VCN。由CSPI 200托管的端点之间的通信可以通过物理网络218执行。计算实例还可以与不由CSPI 200托管或在CSPI 200外部的端点通信。这些端点的示例包括客户本地部署网络内的端点或数据中心,或可通过诸如互联网的公共网络访问的公共端点。与CSPI 200外部的端点的通信可以使用各种通信协议通过公共网络(例如,互联网)(图2中未示出)或私有网络(图2中未示出)来执行。

[0127] 图2中描绘的CSPI 200的架构仅是示例并且不旨在限制。在替代实施例中,变化、替代和修改是可能的。例如,在一些实现中,CSPI 200可以具有比图2中所示的系统或组件更多或更少的系统或组件,可以组合两个或更多个系统,或者可以具有不同的系统配置或布置。图2中描绘的系统、子系统和其他组件可以由相应系统的一个或多个处理单元(例如,处理器、核)执行的软件(例如,代码、指令、程序)、使用硬件或其组合来实现。软件可以存储在

非暂时性存储介质上(例如,存储器设备上)。

[0128] 图4描绘了根据某些实施例的用于提供I/O虚拟化以支持多租赁的主机和NVD之间的连接性。如图4中所描绘的,主机402执行提供虚拟化环境的管理程序404。主机402执行两个虚拟机实例,属于客户/租户#1的VM1 406和属于客户/租户#2的VM2 408。主机402包括经由链路414连接到NVD 412的物理NIC 410。计算实例中的每一个附接到由NVD 412执行的VNIC。在图4中的实施例中,VM1 406附接到VNIC-VM1 420并且VM2 408附接到VNIC-VM2 422。

[0129] 如图4中所示,NIC 410包括两个逻辑NIC,逻辑NIC A 416和逻辑NIC B 418。每个虚拟机附接至其自己的逻辑NIC并且被配置为与其自己的逻辑NIC一起工作。例如,VM1 406附接到逻辑NIC A 416,并且VM2 408附接到逻辑NIC B 418。尽管主机402仅包括由多个租户共享的一个物理NIC 410,但由于逻辑NIC,每个租户的虚拟机相信它们有其自己的主机和NIC。

[0130] 在某些实施例中,每个逻辑NIC被指配其自己的VLAN ID。因此,特定的VLAN ID被指配给租户#1的逻辑NIC A 416,并且分开的VLAN ID被指配给租户#2的逻辑NIC B 418。当分组被从VM1 406传达时,指配给租户#1的标记由管理程序附接到分组,然后分组通过链路414被从主机402传达到NVD 412。以类似的方式,当分组被从VM2 408传达时,指配给租户#2的标记由管理程序附接到分组,然后分组通过链路414被从主机402传达到NVD 412。因此,从主机402传达到NVD 412的分组424具有标识特定租户和相关联VM的相关联标记426。在NVD上,对于从主机402接收到的分组424,与分组相关联的标记426被用于确定该分组是由VNIC-VM1 420还是由VNIC-VM2 422处理。然后,该分组由对应的VNIC处理。图4中所描绘的配置使得每个租户的计算实例能够相信它们拥有其自己的主机和NIC。图4中描述的设置提供了用于支持多租赁的I/O虚拟化。

[0131] 图5描绘了根据某些实施例的物理网络500的简化框图。图5中描绘的实施例被结构化为Clos网络。Clos网络是一种特定类型的网络拓扑,其被设计为提供连接冗余同时维持高对分带宽和最大资源利用率。Clos网络是一种非阻塞、多阶段或多层级交换网络,其中阶段或层级的数量可以是二、三、四、五等。图5中描绘的实施例是包括层级1、2和3的3层级网络。TOR交换机504表示Clos网络中的层级-0交换机。一个或多个NVD连接到TOR交换机。层级-0交换机也被称为物理网络的边缘设备。层级-0交换机连接到也被称为叶交换机层的层级-1交换机。在图5中描绘的实施例中,一组“n”个层级-0TOR交换机连接到一组“n”个层级-1交换机并一起形成舱(pod)。舱中的每个层级-0交换机都与舱中的所有层级-1交换机互连,但舱之间没有交换机的连接性。在某些实现中,两个舱被称为块(block)。每个块由一组“n”个层级-2交换机(有时称为骨干交换机)服务或连接到一组“n”个层级-2交换机。物理网络拓扑中可以有若干个块。层级-2交换机继而连接到“n”个层级-3交换机(有时称为超级骨干交换机)。物理网络500上的分组通信通常使用一个或多个层-3通信协议来执行。通常,除了TOR层之外,物理网络的所有层都是n路冗余,从而允许高可用性。策略可以被针对舱和块来指定以控制物理网络中交换机彼此的可见性,以便使能对物理网络的扩展。

[0132] Clos网络的特征是从一个层级-0交换机到达另一层级-0交换机(或者从连接到层级-0交换机的NVD到连接到层级-0交换机的另一NVD)的最大跳数是固定的。例如,在3层级Clos网络中,对于分组从一个NVD到达另一个NVD最多需要七跳,其中源NVD和目标NVD连接

到Clos网络的叶层级。同样,在4层Clos网络中,对于分组从一个NVD到达另一NVD最多需要九跳,其中源NVD和目标NVD连接到Clos网络的叶层级。因此,Clos网络架构贯穿网络维持一致的延迟,这对于数据中心内和数据中心之间的通信是重要的。Clos拓扑水平扩展并且具有成本效益。通过在各种层级处添加更多交换机(例如,更多叶和骨干交换机)以及通过增加相邻层级处交换机之间的链路数量,网络的带宽/吞吐容量可被容易地增加。

[0133] 在某些实施例中,CSPI内的每个资源被指配被称为云标识符(CID)的唯一标识符。此标识符作为资源的信息的一部分被包括并且可以用于管理资源,例如经由控制台或通过API。CID的示例语法为:

[0134] `ocid1.<RESOURCE TYPE>.<REALM>.[REGION][.FUTURE USE].<UNIQUE ID>`

[0135] 其中,

[0136] `ocid1`:指示CID版本的文字字符串;

[0137] `resource type`:资源的类型(例如,实例、卷、VCN、子网、用户、组等);

[0138] `realm`:资源所在的领域。示例值为“c1”用于商业领域,“c2”用于政府云领域,或者“c3”用于联邦政府云领域等。每个领域可以具有其自己的域名;

[0139] `region`:资源所在的区域。如果区域不适用于该资源,则此部分可以为空;

[0140] `future use`:保留供将来使用。

[0141] `unique ID`:ID的唯一部分。格式可以取决于资源或服务的类型而有所不同。

[0142] 在容器环境(例如,诸如Kubernetes等的容器编排系统)内,舱是表示具有共享资源的一组一个或多个容器的抽象,该资源包括命名空间、网络资源(诸如IP地址)和存储组件(诸如存储卷、文件系统卷等)。舱还包括详细说明容器如何运行的指令(这些指令包括可用的端口、容器镜像版本等)。舱的内容总是共址(co-located)并共同调度(co-scheduled)的。

[0143] 用于执行舱通信的当前方法需要桥接和封装。例如,当位于容器环境内不同节点(也称为主机)上的舱希望彼此通信时,这样的通信需要利用桥接设备来封装以便利用覆盖网络在主机之间路由这些通信。因此,由于封装,这些舱不能在流日志中被直接识别/区分。性能损失也被引发,因为封装的性能需要附加的处理资源来执行,并且需要封装数据添加到正在传送的分组。

[0144] 本公开提供了用于实现容器环境的技术,其中容器环境内的每个舱都被提供有唯一的IP地址和虚拟通信设备,诸如IPvlan设备。例如,IPvlan设备可以包括Linux联网设备。可以在网络命名空间中创建IPvlan设备,并且可以将另一设备指定为该IPvlan的父设备。每当网络请求被发送到IPvlan设备时,父设备都可以充当桥接器/路由器,并将流量路由到适当的位置(父设备知道所有连接的子IPvlan设备)。来自源舱的通信通过一个或多个虚拟化网络接口卡(VNIC)利用目的地舱的唯一IP地址被直接路由到容器环境内的目的地舱,而无需桥接和封装。每个VNIC可以包括虚拟机用来连接到虚拟云网络的接口设备。

[0145] 以此方式,VNIC可以利用指配给舱的IP地址,经由指配给舱的虚拟通信设备路由进出舱的所有流量。然后,此流量维持与舱相关联的IP地址,同时仍使能舱内的规则实现。此外,这样的流量可以由VNIC路由,而无需经历封装或通过桥接设备。

[0146] 这减少了正在传送的数据的大小(例如,通过消除添加到传送的网络分组中的封装期间的所需开销),也减少了实现容器环境的硬件执行的计算量(例如,通过消除在传输

期间执行数据封装所必须的计算成本)。这还减少了容器环境内的数据传输延迟(例如,通过移除数据传输期间执行封装的需要),这提高了实现容器环境内通信的硬件的性能。

[0147] 诸如Kubernetes的容器编排系统提供软件实现和管理。Kubernetes为云提供商提供了开放式(open-ended)接口以定义舱通信应该如何工作(这包括舱到舱、舱到主机、舱到互联网以及舱到自身)。此“开放式接口”被称为容器网络接口(CNI),并且可能包括用于管理集群上网络资源的Kubernetes规范。此规范使得Kubernetes能够与广泛的基于CNI的软件解决方案进行交互。CNI插件允许Kubernetes集群与特定的网络技术集成,并遵守CNI规范。此外,Kubernetes具有Kubernetes服务的概念,这是一种抽象的方式来指代一组舱(例如,可以定义一个服务,使得如果网络请求被发送到该服务,则该请求将被重新路由到多个预定的Pod)。

[0148] 本公开描述了用于以友好的方式定义舱通信如何在Kubernetes环境和虚拟云网络(VCN)内工作的技术。这包括以高效的方式在舱之间以及与VCN通信,并确保网络流量的路径可以被轻松跟踪。

[0149] 在一个示例性解决方案中,为每个pod提供IPvlan设备和点对点设备(ptp)。ptp设备专门用于与节点/主机对话。例如,ptp设备可以包括用于接通两个网络命名空间并允许从pod到主机的通信的Linux联网设备。否则,所有其他流量通过IPvlan设备。IPvlan可以在L3S模式下运行,这意味着主机IPtables被处理。Iptables可以包括计算实例上的过滤软件,该软件执行网络分组过滤和分组源/目的地修改。

[0150] 在一些实施例中,舱的流量流可以进行如下步骤:舱向另一舱发送网络请求。网络请求源自发送舱的IPvlan设备。如果需要,主机iptables将转换网络分组(例如,如果请求是针对Kubernetes服务而不是直接的舱)。在转换分组后,流量从IPvlan父设备(例如,指配给IPvlan设备的VLAN)出去,该设备路由到下一个路由器,并且路由器将决定(1)将流量路由回原始主机上的舱(又称利用“发夹(hairpin)”),或者(2)如果分组的目的地是另一目的地,则路由器将流量转发到该目的地。发送网络响应的舱将执行与上述步骤类似的流量流。

[0151] 图6是根据至少一个实施例的用于促进直接的舱到舱通信的容器环境600的框图。所示的示例性容器环境600可以包括容器编排系统(诸如Kubernetes)的数据平面,其中这样数据平面响应于从分开的控制平面(未示出)接收的指令来实现工作负载。如图所示,示例性容器环境600包括第一节点602和第二节点604。当然,在其他示例性实施例中可以包括更多或更少数量的节点。

[0152] 此外,第一节点602和第二节点604可以每个包括虚拟机或物理机,该虚拟机或物理机包含实现一个或多个舱所必须的一个或多个服务。每个节点(也称为主机)可以包括物理机或虚拟机,该物理机或虚拟机托管并运行Kubernetes集群中的舱。在一个示例中,节点上的组件包括Kubelet(每个节点上的代理,以确保容器在舱中运行)、容器运行时和kube-proxy(在每个节点上运行的网络代理)。如图所示,第一节点602实现第一舱606、第二舱608和第三舱610,而第二节点604实现第四舱612。每个舱606-612包括相应的容器614-620,其中每个容器包括要被运行的一个或多个应用以及相关库和依赖项。每个舱606-612还包括共享的存储/网络资源以及如何运行这样的容器614-620的规范。

[0153] 此外,每个舱606-612包括相应的虚拟通信设备622-628,其中每个虚拟通信设备

622-628被配置用于其相应的舱606-612。在一个实施例中,每个虚拟通信设备622-628可以包括互联网协议虚拟局域网(IPvlan)设备。例如,每个虚拟通信设备622-628可以促进其相应的舱606-612与容器环境600内的其他舱以及与虚拟化网络接口卡(VNIC) 638-642、诸如虚拟云网络644的外部网络等之间的通信。此外,每个舱606-612被指配唯一的互联网协议(IP)地址,该地址用于经由其相应的虚拟通信设备622-628和一个或多个VNIC 638-642将数据通信直接路由到该舱。

[0154] 此外,每个虚拟通信设备622-628被指配VNIC 638-642之一作为父设备。例如,第二VNIC 632被指配为第一舱606的虚拟通信设备622的父设备,第四VNIC 636被指配为第二pod 608的虚拟通信设备624以及第三舱610的虚拟通信设备626的父设备,并且第七VNIC 642被指配为第四舱612的虚拟通信设备628的父设备。每个VNIC 638-642存储VNIC 638-642被指配到的每个舱606-612的IP地址。每个舱606-612还包括相应的veth对646A-B-652A-B。每个veth对646A-B-652A-B包括网络连接,该网络连接将每个舱606-612内的容器网络连接到每个节点602-604的相应节点网络。例如,第一veth对646A-652A可以用于出口流量,并且第二veth对646B-652B可以用于入口流量。

[0155] 此外,第一VNIC 630、第二VNIC 632、第三VNIC 634、第四VNIC 636和第五VNIC 638连接到第一节点602。第六VNIC 640和第七VNIC 642连接到第二节点604。这些VNIC 630-642中的每一个都可以为其相应的节点602和604实现联网通信特征。在一个实施例中,第一VNIC 630可以包括第一节点602的主要父VNIC,而第二、第三、第四和第五VNIC 632-638可以包括第一节点602的次级父VNIC。同样,第六VNIC 640可以包括第二节点604的主要父VNIC,而第七VNIC 642可以包括第二节点604的次级父VNIC。第一节点602的第一/主要父VNIC 630与第二节点604的第六/主要父VNIC 640通信,而第一节点602的第二至第五/次级父VNIC 632-638与第二节点604的第七/次级父VNIC 642通信。

[0156] 如本文所述,图7-9图示了上述示例容器环境600内的各种示例性数据流。

[0157] 更具体地,图7是用于图示根据至少一个实施例的在图6的示例性容器环境600内共享单个VNIC的两个舱之间的示例性请求流700的框图。更具体地,第三舱610和第二舱608经由与第四VNIC 636的连接通信。当位于第三舱610内的容器618发送寻址到第二舱608的对数据的请求(例如,包括第二舱608的IP地址的请求)时,请求首先通过第三舱610的veth向650A-B被发送,使得一个或多个第三方接口可以对请求执行一个或多个评估操作。然后,从第三舱610的虚拟通信设备626向第四VNIC 636发送请求。

[0158] 当在第四VNIC 636处接收到请求时,利用另一发夹操作通过第四VNIC 636来路由请求,使得将与请求相关联的信息存储在一个或多个流日志中。此信息可以包括目的地IP地址(例如,第二舱608的IP地址)、源IP地址(例如,第三舱610的IP地址)等。在执行发夹操作后,第四VNIC 636利用包括在请求内的第二舱608的IP地址将请求路由到第二舱608的虚拟通信设备624。然后,请求由位于第二舱608内的容器616从第二舱608的虚拟通信设备624通过第二舱608的veth对648A-B发送,使得一个或多个第三方接口可以对请求执行一个或多个评估操作,并且请求在位于第二舱608内的容器616处被接收。

[0159] 此外,图8是用于图示根据至少一个实施例的在图6的示例性容器环境600内共享单个VNIC的两个舱之间的示例性响应流800的框图。例如,第三舱610和第二舱608经由与第四VNIC 636的连接通信。响应于在位于第二舱608内的容器616处从第三舱610的容器618接

收到对数据的请求,第二舱608的容器616生成对寻址到第三舱610的请求的响应(其包括第三舱610的IP地址)。此响应由容器616经由第二舱608的虚拟通信设备发送到第四VNIC 636。

[0160] 当在第四VNIC 636处接收到响应时,利用另一个发夹操作通过第四VNIC 636来路由响应,使得将与响应相关联的信息存储在一个或多个流日志中。此信息可以包括目的地IP地址(例如,第三舱610的IP地址)、源IP地址(例如,第二舱608的IP地址)等。在执行发夹操作后,第四VNIC 636利用包括在响应内的第三舱610的IP地址将响应路由到第三pod舱的虚拟通信设备626。然后,响应被从第三舱610的虚拟通信设备626通过第三舱610的veth对650A-B发送使得一个或多个第三方接口可以对响应执行一个或多个评估操作,并在位于第三舱610内的容器618处被接收。

[0161] 此外,图9是用于图示根据至少一个实施例的图6的示例性容器环境600的相同节点602内经由不同VNIC通信的两个舱之间的示例性请求流900的框图。更具体地,第一舱606经由与第一节点602的第二VNIC 632的连接通信,并且第二舱608经由与第一节点602的第四VNIC 636的连接通信。当位于第二舱608内的容器616发送寻址到第一舱606的对数据的请求(例如,包括第一舱606的IP地址的请求)时,请求首先通过第二舱608的veth对648A-B被发送,使得一个或多个第三方接口可以对请求执行一个或多个评估操作。然后,从第二舱608的虚拟通信设备624向第四VNIC 636发送请求。

[0162] 当在第四VNIC 636处接收到请求时,利用包括在请求内的第一舱606的IP地址将请求路由到第二VNIC 632。在在第二VNIC 632处接收到请求后,利用包括在请求内的第一舱606的IP地址,第二VNIC 632将请求路由到第一舱606的虚拟通信设备622。然后,请求被从第一舱606的虚拟通信设备622通过第一舱606的veth对646A-B发送,使得一个或多个第三方接口可以对请求执行一个或多个评估操作,并且请求在位于第一舱606内的容器614处被接收。

[0163] 此外,图10是用于图示根据至少一个实施例的图6的示例性容器环境600的相同节点602内经由不同VNIC通信的两个舱之间的示例性响应流1000的框图。例如,第一舱606经由与第一节点602的第二VNIC 632的连接通信,并且第二舱608经由与第一节点602的第四VNIC 636的连接通信。响应于在位于第一舱608内的容器614处从第二舱608的容器616接收到对数据的请求,第一舱606的容器614生成对寻址到第二舱608的请求的响应(其包括第二舱608的IP地址)。此响应由容器614经由第一舱608的虚拟通信设备622发送到第二VNIC 632。

[0164] 当在第二VNIC 632处接收到响应时,响应被路由到第四VNIC 636。在在第四VNIC 636处接收到响应后,利用包括在响应内的第二舱608的IP地址,第四VNIC 636将请求路由到第二舱608的虚拟通信设备624。然后,响应被从第二舱608的虚拟通信设备624通过第二舱608的veth对648A-B发送,使得一个或多个第三方接口可以对响应执行一个或多个评估操作,并且响应在位于第二舱608内的容器616处被接收。

[0165] 此外,图11是用于图示根据至少一个实施例的图6的示例性容器环境600的不同节点内经由不同VNIC通信的两个舱之间的示例性请求流1100的框图。更具体地,第三舱610经由与第一节点602的第四VNIC 636的连接通信,并且第四舱612经由与第二节点604的第七VNIC 642的连接通信。当位于第三舱610内的容器618发送寻址到第四舱612的对数据的请

求(例如,包括第四舱612的IP地址的请求)时,请求首先通过第三舱的veth对650A-B被发送,使得一个或多个第三方接口可以对请求执行一个或多个评估操作。然后,从第三舱610的虚拟通信设备626向第四VNIC 636发送请求。

[0166] 当在第四VNIC 636处接收到请求时,将请求路由到第七VNIC 642。在在第七VNIC 642处接收到请求后,利用包括在请求内的第四舱612的IP地址,第七VNIC 642将请求路由到第四舱612的虚拟通信设备628。然后,利用发夹操作,请求被从第四舱612的虚拟通信设备628通过第四舱612的veth对652A-B发送,使得一个或多个第三方接口可以对请求执行一个或多个评估操作,并且请求在位于第一舱612内的容器620处被接收。

[0167] 此外,图12是用于图示根据至少一个实施例的图6的示例性容器环境600的不同节点内经由不同VNIC通信的两个舱之间的示例性响应流1200的框图。例如,第三舱610经由与第一节点602的第四VNIC 636的连接通信,并且第四舱612经由与第二节点604的第七VNIC 642的连接通信。响应于在位于第四舱612内的容器620处从第三舱610的容器618接收到对数据的请求,第四舱612的容器620生成对寻址到第三舱610的请求的响应(其包括第三舱610的IP地址)。此响应由容器620经由第四舱612的虚拟通信设备628发送到第七VNIC 642。

[0168] 当在第七VNIC 642处接收到响应时,响应被路由到第四VNIC 636。在在第四VNIC 636处接收到响应后,利用包括在响应内的第三舱610的IP地址,第四VNIC 636将请求路由到第三舱610的虚拟通信设备626。然后,响应被从第三舱610的虚拟通信设备626通过第三舱610的veth对650A-B发送,使得一个或多个第三方接口可以对响应执行一个或多个评估操作,并且响应在位于第三舱610内的容器618处被接收。

[0169] 此外,图13是用于图示根据至少一个实施例的图6的示例性容器环境600内的舱和虚拟云网络(VCN)644之间的示例性请求流1300的框图。当位于第一舱606内的容器614发送寻址到VCN 644的对数据的请求(例如,包括VCN 644的IP地址的请求)时,利用发夹操作,请求首先通过第一舱606的veth对646A-B被发送,使得一个或多个第三方接口可以对请求执行一个或多个评估操作。然后,请求被从第一舱606的虚拟通信设备622发送到第二VNIC 632。当请求在第二VNIC 632处被接收到时,利用VCN 644的IP地址,请求被路由到VCN 644。然后,请求可以被引导到VCN 644内的适当目的地。

[0170] 此外,图14是用于图示根据至少一个实施例的图6的示例性容器环境600内的舱和虚拟云网络(VCN)644之间的示例性响应流1400的框图。响应于在VCN 644处从第一舱606的容器614接收到对数据的请求,VCN 644内的实体生成对寻址到第一舱606的请求的响应(其包括第一舱606的IP地址)。此响应经由VCN 644被发送到第二VNIC 632。

[0171] 当在第二VNIC 632处接收到响应时,利用包括在响应内的第一舱606的IP地址,响应被路由到第一舱606的虚拟通信设备622。然后,响应被从第一舱606的虚拟通信设备622通过第一舱606的veth对646A-B发送以使得一个或多个第三方接口可以对响应执行一个或多个评估操作,并在位于第一舱606内的容器614处被接收。

[0172] 图15图示了根据至少一个实施例的用于在容器环境内实现通信的示例方法1500。方法1500可以由图6的容器环境600的一个或多个组件执行。计算机可读存储介质包括计算机可读指令,该指令在通过计算设备的一个或多个处理器的执行时,使计算设备执行方法1500。方法1500可以以任何合适的次序执行。应当理解,方法1500可以包括比图15中所示的步骤更多数量或更少数量的步骤。

[0173] 方法1500可以在1502处起始,其中在容器环境内配置第一节点,第一节点包括第一舱。在一些实施例中,容器环境内的节点(包括第一节点)可以每个包括虚拟机或物理机,该虚拟机或物理机包含实现舱(诸如第一舱)所必须的一个或多个服务。在一些实施例中,第一舱可以包括操作系统(OS)级虚拟化。在一些实施例中,容器环境的第一节点可以包括容器环境内配置的主机。

[0174] 此外,在一些实施例中,容器环境包括用于管理一个或多个舱的系统(例如,Kubernetes容器编排系统等)。此外,在一些实施例中,第一舱可以包括一个或多个容器,其中每个容器包括一个或多个要运行的应用以及相关库和依赖项。此外,在一些实施例中,可以在第一节点内实现多个不同的舱(包括第一舱)。此外,在一些实施例中,第一节点内的每个舱可以具有其自己的命名空间。在一些实施例中,一个或多个舱可以共址于第一节点内。此外,在一些实施例中,第一节点内的每个pod可以被指配唯一的IP地址。

[0175] 在1504处,配置从第一舱经由第一虚拟通信设备的通信。在一些实施例中,第一虚拟通信设备可以包括互联网协议虚拟局域网(IPvlan)设备。在一些实施例中,可以为容器环境的每个节点内的每个舱配置虚拟通信设备。在一些实施例中,虚拟通信设备可以促进其指配的舱与容器环境内的其他舱以及与虚拟化网络接口卡(VNIC)、容器环境外部的网络(例如,虚拟云网络(VCN))等之间的通信。例如,第一舱的虚拟通信设备可以与VNIC通信以将对数据的请求(以及对对数据的请求的响应)发送到容器环境内的目的地舱,其中这样的请求和响应利用指配给那些目的地舱的唯一IP地址被引导到目的地舱。

[0176] 在1506处,来自指配给第一节点内的第一舱的第一虚拟通信设备的数据在连接到容器环境内的第一节点的第一虚拟化网络接口卡(VNIC)处被接收,其中数据被配置为传送到目的地设备。在一些实施例中,第一VNIC可以包括附接到第一节点的次级父设备,以实现第一节点和容器环境内的联网功能。在一些实施例中,数据可以包括对附加数据的请求(例如,来自第一节点内的第二舱、与第一节点分开的第二节点内的第二舱等)。

[0177] 此外,在一些实施例中,数据可以包括对第一舱所接收的请求的响应。在一些实施例中,数据可以包括一个或多个网络分组。在一些实施例中,数据可以按第一虚拟通信设备实现的一个或多个规则(例如,主机规则)处理。在一些实施例中,数据可以由第一舱内的本地主机实例化。

[0178] 在1508处,由第一VNIC将数据路由到目的地设备,其中数据至少包括标识第一舱的地址。在一些实施例中,目的地可以包括与请求所源自的第一舱不同的第二舱。例如,第一舱和第二舱可以共享第一节点内的第一VNIC。例如,第一舱和第二舱两者都可以由第一VNIC指配IP地址。

[0179] 此外,在一些实施例中,可以将数据直接从第一VNIC路由到第二舱。在一些实施例中,第一VNIC可以利用发夹操作将数据直接路由到第二舱。例如,发夹操作可以通过第一VNIC路由数据,使得传输信息存储在一个或多个日志(例如,流日志等)中。在另一个示例中,发夹可以使用利用模式标志选择的vepa模式来执行。在一些实施例中,第一VNIC可以利用桥接操作将数据直接路由到第二舱。

[0180] 此外,在一些实施例中,目的地可以包括与执行路由的第一VNIC不同的第二VNIC。例如,第二VNIC和第一VNIC两者可以附接到容器环境的第一节点。在另一示例中,第二VNIC可以附接到与第一节点不同的容器环境的第二节点。在又另一示例中,可以将数据直接从

第一VNIC路由到第二VNIC。

[0181] 此外,在一个示例中,第二VNIC可以接收数据并将数据路由到与第一舱不同的第二舱。例如,第二VNIC可被指配给其对应节点内的第二舱。这样,数据可以在容器环境内的VNIC之间被直接发送,而无需穿过桥接器,并且无需经历数据封装。

[0182] 此外,在一些实施例中,目的地可以包括虚拟云网络(VCN)。例如,可以将数据直接从第一VNIC路由到VCN。在一些实施例中,标识第一舱的地址可以包括指配给第一舱的第一互联网协议(IP)地址。例如,第一VNIC可以从IP地址的空闲池中将第一IP地址指配给第一舱的第一虚拟通信设备。以此方式,第一虚拟通信设备的第一IP地址可以在数据传输期间被记录在一个或多个流日志中。

[0183] 此外,在一些实施例中,数据还可以包括指配给接收者舱的第二互联网协议(IP)地址。例如,接收者舱可以包括数据的预期接收者。在另一示例中,接收者舱可以位于节点内,并且与节点相关联的VNIC可以将第二IP地址指配给被指配给该节点内的接收者舱的第二虚拟通信设备。

[0184] 此外,在一些实施例中,数据可以在由第一通信设备发送到第一VNIC之前由指配给第一舱的第二通信设备(例如,点对点设备)处理。例如,第二通信设备可以包括虚拟以太网设备(veth),诸如一个或多个veth对,其将第一网络(例如,容器网络等)连接到第二网络(例如,主机网络等)。在另一示例中,一个或多个veth对中的每一个可以包括存在于舱环境的舱命名空间中的第一部分,以及存在于舱环境的主机命名空间中的第二部分。

[0185] 此外,在一个示例中,第一舱可以利用发夹操作将数据引导通过第二通信设备。在另一示例中,第二通信设备可以使得第三方容器联网接口(诸如Calico)能够执行一个或多个评估操作。

[0186] 此外,在一些实施例中,响应于在第二舱处接收数据,数据可以由指配给第二舱的第三通信设备(例如,点对点设备)处理。例如,第三通信设备可以包括虚拟以太网设备(veth),诸如一个或多个veth对,其将第一网络(例如,容器网络等)连接到第二网络(例如,主机网络等)。在另一示例中,一个或多个veth对可以包括存在于容器环境的舱命名空间中的第一部分,以及存在于容器环境的主机命名空间中的第二部分。在又另一示例中,第二舱可以利用发夹操作将数据引导通过第三通信设备。在一个示例中,发夹操作可以包括将网络流量发送出网络设备的过程,其目的是网络流量将被重新路由回到相同网络设备中。以此方式,第三方容器联网接口(诸如Calico)可以执行一个或多个评估操作。

[0187] 此外,在一些实施例中,第一VNIC可以从目的地接收响应。例如,最初由第一VNIC接收/路由的数据可以包括对数据的请求。在另一示例中,第一VNIC可以从初始目的地接收对请求的响应。例如,响应可以包括指配给第一舱的第一互联网协议(IP)地址。在又另一示例中,第一VNIC可将响应路由到指配给第一节点内第一舱的第一虚拟通信设备。

[0188] 以此方式,VNIC可以利用标识第一舱的地址,经由第一虚拟通信设备(例如,IPvlan设备)路由进出第一舱的所有流量。这可以使得此流量能够维持与第一舱相关联的IP地址,同时仍使能通过第一虚拟通信设备在第一节点内的规则实现。此外,这样的流量可以由第一VNIC路由,而无需经历封装或通过桥接设备。这可以减少正在传送的数据的大小(例如,通过消除添加到传送的网络分组的封装期间的所需开销),并且也可以减少由实现容器环境的硬件执行的计算量(例如,通过消除在传输期间执行数据封装所必须的计算成

本)。这还可以减少容器环境内的数据传输的延迟(例如,通过移除数据传输期间执行封装的需要),这可以提高实现容器环境的硬件的性能。

[0189] 如上所述,基础设施即服务(IaaS)是一种特定类型的云计算。IaaS可以配置为通过公共网络(例如,互联网)提供虚拟化计算资源。在IaaS模型中,云计算提供商可以托管基础设施组件(例如,服务器、存储设备、网络节点(例如,硬件)、部署软件、平台虚拟化(例如,管理程序层)或类似物)。在一些情况下,IaaS提供商还可以供应各种各样的服务来伴随那些基础设施组件(例如,计费、监控、日志记录、负载平衡和集群等)。因此,由于这些服务可以是策略驱动的,IaaS用户可以能够实现策略来驱动负载平衡以维持应用可用性和性能。

[0190] 在一些情况下,IaaS客户可以通过广域网(WAN)(诸如互联网)访问资源和服务,并可以使用云提供商的服务来安装应用堆栈的剩余元素。例如,用户可以登录IaaS平台来创建虚拟机(VM)、在每个VM上安装操作系统(OS)、部署诸如数据库的中间件、为工作负载和备份创建存储桶,以及甚至将企业软件安装到该VM中。然后,客户可以使用提供商的服务来执行各种功能,包括平衡网络流量、故障排除应用问题、监控性能、管理灾难恢复等。

[0191] 在大多数情况下,云计算模型将需要云提供商的参与。但是,云提供商不一定是专门提供(例如,供应、出租、销售)IaaS的第三方服务。实体也可以选择部署私有云,变成其自己的基础设施服务提供者。

[0192] 在一些示例中,IaaS部署是将新应用或应用的新版本放到准备好的应用服务器或类似物上的过程。它还可以包括准备服务器的过程(例如,安装库、守护进程等)。这通常由云提供商在管理程序层(例如,服务器、存储、网络硬件和虚拟化)之下管理。因此,客户可以负责处理(OS)、中间件和/或应用部署(例如,在自助虚拟机上(例如,其可以按需启动(spun up))或类似物)。

[0193] 在一些示例中,IaaS供给可以指获取计算机或虚拟主机以供使用,以及甚至在其上安装所需的库或服务。在大多数情况下,部署不包括供给,并且供给可能需要先被执行。

[0194] 在一些情况下,对于IaaS供给有两个不同的挑战。首先,在运行任何事项之前,有供给初始基础设施集的初始挑战。其次,在一切都被供给后,有演进现有基础设施的挑战(例如,添加新服务、更改服务、移除服务等)。在一些情况下,这两个挑战可以通过使得基础设施的配置能够被声明式定义来解决。换句话说,基础设施(例如,需要什么组件以及它们如何交互)可以由一个或多个配置文件定义。因此,可以声明式描述基础设施的整体拓扑(例如,哪些资源依赖于哪些资源以及它们每个如何协同工作)。在某些情况下,在定义了拓扑后,可以生成创建和/或管理配置文件中描述的不同组件的工作流。

[0195] 在一些示例中,基础设施可以具有许多互连元素。例如,可以有一个或多个虚拟私有云(VPC)(例如,可配置和/或共享计算资源的潜在按需池),也被称为核心网络。在一些示例中,还可以有供给以定义网络的入站和/或出站流量将被如何建立的一个或多个入站/出站流量组规则以及一个或多个虚拟机(VM)。还可以供给其他基础设施元素,诸如负载平衡器、数据库,或类似物。随着越来越多的基础设施元素被期望和/或添加,基础设施可以逐步演进。

[0196] 在一些情况下,持续部署技术可以被采用来使能跨各个虚拟计算环境的基础设施代码的部署。此外,所描述的技术可以使能这些环境内的基础设施管理。在一些示例中,服务团队可以编写期望部署到一个或多个(但通常是多个)不同生产环境(例如,跨各种不同

的地理位置,有时甚至跨越整个世界)的代码。但是,在一些示例中,必须首先建立将在其上部署代码的基础设施。在一些情况下,供给可以被手动完成,供给工具可以被利用来供给资源,和/或当基础设施被供给后,部署工具可以被利用来部署代码。

[0197] 图16是图示根据至少一个实施例的IaaS架构的示例模式的框图1600。服务运营商1602可以通信地耦合到可以包括虚拟云网络(VCN)1606和安全主机子网1608的安全主机租赁1604。在一些示例中,服务运营商1602可以使用一个或多个客户端计算设备,其可以是便携式手持设备(例如,**iPhone®**、蜂窝电话、**iPad®**、计算平板电脑、个人数字助理(PDA))或可穿戴设备(例如,Google **Glass®**头戴式显示器),运行软件(诸如Microsoft Windows **Mobile®**)和/或各种各样的移动操作系统(诸如iOS、Windows Phone、Android、BlackBerry8、Palm OS等),并且启用互联网、电子邮件、短信服务(SMS)、**Blackberry®**或其他通信协议。替代地,客户端计算设备可以是通用个人计算机,举例来说,包括运行各种版本的Microsoft **Windows®**、Apple **Macintosh®**和/或Linux操作系统的个人计算机和/或膝上型计算机。客户端计算设备可以是运行任何各种各样商用**UNIX®**或类UNIX操作系统(包括但不限于各种GNU/Linux操作系统,诸如例如Google Chrome OS)的工作站计算机。替代地或此外,客户端计算设备可以是能够通过可以访问VCN 1606和/或互联网的网络通信的任何其他电子设备,诸如瘦客户端计算机、启用互联网的游戏系统(例如,具有或不具有**Kinect®**手势输入设备的Microsoft Xbox游戏控制台),和/或个人消息传递设备。

[0198] VCN 1606可以包括本地对等网关(LPG)1610,其可以经由包含在SSH VCN 1612中的LPG 1610通信地耦合到安全壳(SSH) VCN 1612。SSH VCN 1612可以包括SSH子网1614,并且SSH VCN 1612可以经由包含在控制平面VCN 1616中的LPG 1610通信地耦合到控制平面VCN 1616。而且,SSH VCN 1612可以经由LPG 1610通信地耦合到数据平面VCN 1618。控制平面VCN 1616和数据平面VCN 1618可以被包含在可以由IaaS提供商拥有和/或操作的服务租赁1619中。

[0199] 控制平面VCN 1616可以包括控制平面非军事区(DMZ)层1620,其充当外围网络(例如,公司内部网和外部网络之间的公司网络的部分)。基于DMZ的服务器可以具有有限的责任,并有助于保持违规被遏制。此外,DMZ层1620可以包括一个或多个负载平衡器(LB)子网1622、可以包括(一个或多个)应用程序子网1626的控制平面应用程序层1624、可以包括(一个或多个)数据库(DB)子网1630(例如,(一个或多个)前端DB子网和/或(一个或多个)后端DB子网)的控制平面数据层1628。包含在控制平面DMZ层1620中的(一个或多个)LB子网1622可以通信地耦合到包含在控制平面应用程序层1624中的(一个或多个)应用程序子网1626和可以包含在控制平面VCN 1616中的互联网网关1634,并且(一个或多个)应用程序子网1626可以通信地耦合到包含在控制平面数据层1628中的(一个或多个)DB子网1630和服务网关1636和网络地址转换(NAT)网关1638。控制平面VCN 1616可以包括服务网关1636和NAT网关1638。

[0200] 控制平面VCN 1616可以包括数据平面镜像应用程序层1640,其可以包括(一个或多个)应用程序子网1626。包含在数据平面镜像应用程序层1640中的(一个或多个)应用程序子网1626可以包括虚拟网络接口控制器(VNIC)1642,其可以执行计算实例1644。计算实

例1644可以通信地将数据平面镜像应用程序层1640的(一个或多个)应用程序子网1626耦合到可以包含在数据平面应用程序层1646中的(一个或多个)应用程序子网1626。

[0201] 数据平面VCN 1618可以包括数据平面应用程序层1646、数据平面DMZ层1648和数据平面数据层1650。数据平面DMZ层1648可以包括可通信地耦合到数据平面应用程序层1646的(一个或多个)应用程序子网1626和数据平面VCN 1618的互联网网关1634的(一个或多个)LB子网1622。(一个或多个)应用程序子网1626可以通信地耦合到数据平面VCN 1618的服务网关1636和数据平面VCN 1618的NAT网关1638。数据平面数据层1650还可以包括可以通信地耦合到数据平面应用程序层1646的(一个或多个)应用程序子网1626的(一个或多个)DB子网1630。

[0202] 控制平面VCN 1616的和数据平面VCN 1618的互联网网关1634可以通信地耦合到元数据管理服务1652,元数据管理服务1652可以通信地耦合到公共互联网1654。公共互联网1654可以通信地耦合到控制平面VCN 1616的和数据平面VCN 1618的NAT网关1638。控制平面VCN 1616和数据平面VCN 1618的服务网关1636可以通信地耦合到云服务1656。

[0203] 在一些示例中,控制平面VCN 1616的或数据平面VCN 1618的服务网关1636可以在不经过公共互联网1654的情况下对云服务1656做出应用编程接口(API)调用。来自服务网关1636的对云服务1656的API调用可以是单向的:服务网关1636可以对云服务1656做出API调用,并且云服务1656可以将请求的数据发送到服务网关1636。但是,云服务1656可以不发起对服务网关1636的API调用。

[0204] 在一些示例中,安全主机租赁1604可以直接连接到服务租赁1619,否则其会被隔离。安全主机子网1608可以通过LPG 1610与SSH子网1614通信,LPG 1610可以通过否则隔离的系统使能双向通信。将安全主机子网1608连接到SSH子网1614可以给予安全主机子网1608到服务租赁1619内的其他实体的访问权限。

[0205] 控制平面VCN 1616可以允许服务租赁1619的用户建立或以其他方式供给期望的资源。在控制平面VCN 1616中供给的期望的资源可以在数据平面VCN 1618中部署或以其他方式使用。在一些示例中,控制平面VCN 1616可以与数据平面VCN 1618隔离,并且控制平面VCN 1616的数据平面镜像应用程序层1640可以经由可以包含在数据平面镜像应用程序层1640和数据平面应用程序层1646中的VNIC 1642与数据平面VCN 1618的数据平面应用程序层1646通信。

[0206] 在一些示例中,系统的用户或客户可以通过公共互联网1654做出请求(例如创建、读取、更新或删除(CRUD)操作),该公共互联网1654可以将请求传达到元数据管理服务1652。元数据管理服务1652可以通过互联网网关1634将请求传达到控制平面VCN 1616。请求可以由包含在控制平面DMZ层1620中的(一个或多个)LB子网1622接收。(一个或多个)LB子网1622可以确定请求是有效的,并且响应于此确定,(一个或多个)LB子网1622可以将请求传送到包含在控制平面应用程序层1624中的(一个或多个)应用程序子网1626。如果请求是经验证的并且需要对公共互联网1654的调用,则对公共互联网1654的调用可以被传送到可以对公共互联网1654做出调用的NAT网关1638。请求可能期望被存储的元数据可以存储在(一个或多个)DB子网1630中。

[0207] 在一些示例中,数据平面镜像应用程序层1640可以促进控制平面VCN 1616和数据平面VCN 1618之间的直接通信。例如,可能期望将对配置的更改、更新或其他适当的修改应

用到包含在数据平面VCN 1618中的资源。经由VNIC 1642,控制平面VCN 1616可以直接与包含在数据平面VCN 1618中的资源通信,并由此可以执行对包含在数据平面VCN 1618中的资源的配置的更改、更新或其他适当的修改。

[0208] 在一些实施例中,控制平面VCN 1616和数据平面VCN 1618可以包含在服务租赁1619中。在此情况下,系统的用户或客户可以不拥有或操作控制平面1616或数据平面VCN 1618。相反,IaaS提供商可以拥有或操作控制平面VCN 1616和数据平面VCN 1618,这两者都可以包含在服务租赁1619中。此实施例可以使能可阻止用户或客户与其他用户的或其他客户的资源交互的网络隔离。此外,此实施例可以允许系统的用户或客户私密地存储数据库而不需要依靠可能不具有期望的威胁预防级别的公共互联网1654来存储。

[0209] 在其他实施例中,包含在控制平面VCN 1616中的一个或多个)LB子网1622可以被配置为从服务网关1636接收信号。在此实施例中,控制平面VCN 1616和数据平面VCN 1618可以被配置为由IaaS提供商的客户在不调用公共互联网1654的情况下调用。IaaS提供商的客户可能期望此实施例,因为客户使用的(一个或多个)数据库可以由IaaS提供商控制并且可以存储在可以与公共互联网1654隔离的服务租赁1619上。

[0210] 图17是图示根据至少一个实施例的IaaS架构的另一示例模式的框图1700。服务运营商1702(例如,图16的服务运营商1602)可以通信地耦合到安全主机租赁1704(例如,图16的安全主机租赁1604),其可以包括虚拟云网络(VCN)1706(例如,图16的VCN 1606)和安全主机子网1708(例如,图16的安全主机子网1608)。VCN 1706可以包括本地对等网关(LPG)1710(例如,图16的LPG 1610),其可以经由包含在SSH VCN 1712中的LPG 1610通信地耦合到安全壳(SSH)VCN 1712(例如,图16的(SSH)VCN 1612)。SSH VCN 1712可以包括SSH子网1714(例如,图16的SSH子网1614),并且SSH VCN 1712可以经由包含在控制平面VCN 1716中的LPG 1710通信地耦合到控制平面VCN 1716(例如,图16的控制平面VCN 1616)。控制平面VCN 1716可以包含在服务租赁1719(例如,图16的服务租赁1619)中,并且数据平面VCN 1718(例如,图16的数据平面VCN 1618)可以包含在可以由系统的用户或客户拥有或操作的客户租赁1721中。

[0211] 控制平面VCN 1716可以包括可以包括(一个或多个)LB子网1722(例如,图16的(一个或多个)LB子网1622)的控制平面DMZ层1720(例如,图16的控制平面DMZ层1620)、可以包括(一个或多个)应用程序子网1726(例如,图16的(一个或多个)应用程序子网1626)的控制平面应用程序层1724(例如,图16的控制平面应用程序层1624)、可以包括(一个或多个)数据库(DB)子网1730(例如,类似于图16的(一个或多个)DB子网1630)的控制平面数据层1728(例如,图16的控制平面数据层1628)。包含在控制平面DMZ层1720中的(一个或多个)LB子网1722可以通信地耦合到包含在控制平面应用程序层1724中的(一个或多个)应用程序子网1726和可以包含在控制平面VCN 1716中的互联网网关1734(例如,图16的互联网网关1634),并且(一个或多个)应用程序子网1726可以通信地耦合到包含在控制平面数据层1728中的(一个或多个)DB子网1730和服务网关1736(例如,图16的服务网关1636)和网络地址转换(NAT)网关1738(例如,图16的NAT网关1638)。控制平面VCN 1716可以包括服务网关1736和NAT网关1738。

[0212] 控制平面VCN 1716可以包括数据平面镜像应用程序层1740(例如,图16的数据平面镜像应用程序层1640),其可以包括(一个或多个)应用程序子网1726。包含在数据平面镜

像应用程序层1740中的(一个或多个)应用程序子网1726可以包括可以执行计算实例1744(例如,类似于图16的计算实例1644)的虚拟网络接口控制器(VNIC) 1742(例如,1642的VNIC)。计算实例1744可以促进数据平面镜像应用程序层1740的(一个或多个)应用程序子网1726和可以包含在数据平面应用程序层1746(例如,图16的数据平面应用程序层1646)中的(一个或多个)应用程序子网1726之间经由包含在数据平面镜像应用程序层1740中的VNIC 1742和包含在数据平面应用程序层1746中的VNIC 1742的通信。

[0213] 包含在控制平面VCN 1716中的互联网网关1734可以通信地耦合到元数据管理服务1752(例如,图16的元数据管理服务1652),该元数据管理服务1752可以通信地耦合到公共互联网1754(例如,图16的公共互联网1654)。公共互联网1754可以通信地耦合到包含在控制平面VCN 1716中的NAT网关1738。包含在控制平面VCN 1716中的服务网关1736可以通信地耦合到云服务1756(例如,图16的云服务1656)。

[0214] 在一些示例中,数据平面VCN 1718可以包含在客户租赁1721中。在此情况下,IaaS提供商可以为每个客户提供控制平面VCN 1716,并且IaaS提供商可以为每个客户建立包含在服务租赁1719中的唯一计算实例1744。每个计算实例1744可以允许包含在服务租赁1719中的控制平面VCN 1716和包含在客户租赁1721中的数据平面VCN 1718之间的通信。计算实例1744可以允许在包含在服务租赁1719中的控制平面VCN 1716中供给的资源被部署或以其他方式用在包含在客户租赁1721中的数据平面VCN 1718中。

[0215] 在其他示例中,IaaS提供商的客户可以拥有存在于客户租赁1721中的数据库。在此示例中,控制平面VCN 1716可以包括数据平面镜像应用程序层1740,其可以包括(一个或多个)应用程序子网1726。数据平面镜像应用程序层1740可以驻留在数据平面VCN 1718中,但是数据平面镜像应用程序层1740可以不存在于数据平面VCN 1718中。也就是说,数据平面镜像应用程序层1740可以具有对问客户租赁1721的访问权限,但是数据平面镜像应用程序层1740可以不存在于数据平面VCN 1718中或由IaaS提供商的客户拥有或操作。数据平面镜像应用程序层1740可以被配置为做出对数据平面VCN 1718的调用,但可以被配置为做出对包含在控制平面VCN 1716中的任何实体的调用。客户可能期望在数据平面VCN 1718中部署或以其他方式使用在控制平面VCN 1716中供给的资源,并且数据平面镜像应用程序层1740可以促进客户的期望的部署或对资源的其他使用。

[0216] 在一些实施例中,IaaS提供商的客户可以将过滤器应用到数据平面VCN 1718。在此实施例中,客户可以确定数据平面VCN 1718可以访问什么,并且客户可以限制从数据平面VCN 1718对公共互联网1754的访问。IaaS提供商可能无法应用过滤器或以其他方式控制数据平面VCN 1718对任何外部网络或数据库的访问。客户将过滤器和控制应用到包含在客户租赁1721中的数据平面VCN 1718上,可以帮助将数据平面VCN 1718与其他客户和与公共互联网1754隔离。

[0217] 在一些实施例中,云服务1756可以被服务网关1736调用以访问可能不存在于公共互联网1754上、控制平面VCN 1716上或数据平面VCN 1718上的服务。云服务1756和控制平面VCN 1716或数据平面VCN 1718之间的连接可以不是活动的或连续的。云服务1756可以存在于IaaS提供商拥有或操作的不同网络上。云服务1756可以被配置为接收来自服务网关1736的调用并且可以被配置为不接收来自公共互联网1754的调用。一些云服务1756可以与其他云服务1756隔离,并且控制平面VCN 1716可以与可以与控制平面VCN 1716不在相同区

域的云服务1756隔离。例如,控制平面VCN 1716可以位于“区域1”中,而云服务“部署16”可以位于区域1和“区域2”中。如果对部署16的调用由包含在位于区域1中的控制平面VCN 1716中的服务网关1736做出,则调用可以被传送到区域1中的部署16。在此示例中,区域1中的控制平面VCN 1716或部署16可以不通信地耦合到区域2中的部署16,或者以其他方式与区域2中的部署16通信。

[0218] 图18是图示根据至少一个实施例的IaaS架构的另一示例模式的框图1800。服务运营商1802(例如,图16的服务运营商1602)可以通信地耦合到安全主机租赁1804(例如,图16的安全主机租赁1604),该安全主机租赁1804可以包括虚拟云网络(VCN) 1806(例如,图16的VCN 1606)和安全主机子网1808(例如,图16的安全主机子网1608)。VCN 1806可以包括LPG 1810(例如,图16的LPG 1610),该LPG 1810可以经由包含在SSH VCN 1812中的LPG 1810通信地耦合到SSH VCN 1812(例如,图16的SSH VCN 1612)。SSH VCN 1812可以包括SSH子网1814(例如,图16的SSH子网1614),并且SSH VCN 1812可以经由包含在控制平面VCN 1816中的LPG 1810通信地耦合到控制平面VCN 1816(例如,图16的控制平面VCN 1616)和经由包含在数据平面VCN 1818中的LPG 1810通信地耦合到数据平面VCN 1818(例如,图16的数据平面1618)。控制平面VCN 1816和数据平面VCN 1818可以包含在服务租赁1819(例如,图16的服务租赁1619)中。

[0219] 控制平面VCN 1816可以包括可以包括(一个或多个)负载均衡器(LB)子网1822(例如,图16的(一个或多个)LB子网1622)的控制平面DMZ层1820(例如,图16的控制平面DMZ层1620),可以包括(一个或多个)应用程序子网1826(例如,类似于图16的(一个或多个)应用程序子网1626)的控制平面应用程序层1824(例如,图16的控制平面应用程序层1624),可以包括(一个或多个)DB子网1830的控制平面数据层1828(例如,图16的控制平面数据层1628)。包含在控制平面DMZ层1820中的(一个或多个)LB子网1822可以通信地耦合到包含在控制平面应用程序层1824中的(一个或多个)应用程序子网1826和到可以包含在控制平面VCN 1816中的互联网网关1834(例如,图16的互联网网关1634),并且(一个或多个)应用程序子网1826可以通信地耦合到包含在控制平面数据层1828中的(一个或多个)DB子网1830以及到服务网关1836(例如,图16的服务网关)和网络地址转换(NAT)网关1838(例如,图16的NAT网关1638)。控制平面VCN 1816可以包括服务网关1836和NAT网关1838。

[0220] 数据平面VCN 1818可以包括数据平面应用程序层1846(例如,图16的数据平面应用程序层1646)、数据平面DMZ层1848(例如,图16的数据平面DMZ层1648)、以及数据平面数据层1850(例如,图16的数据平面数据层1650)。数据平面DMZ层1848可以包括(一个或多个)LB子网1822,该(一个或多个)LB子网1822可以通信地耦合到数据平面应用程序层1846的(一个或多个)受信任应用程序子网1860和(一个或多个)非受信任应用程序子网1862以及包含在数据平面VCN 1818的互联网网关1834。(一个或多个)受信任应用程序子网1860可以通信地耦合到包含在数据平面VCN 1818中的服务网关1836、包含在数据平面VCN 1818中的NAT网关1838和包含在数据平面数据层1850中的(一个或多个)DB子网1830。(一个或多个)非受信任应用程序子网1862可以通信地耦合到包含在数据平面VCN 1818中的服务网关1836和包含在数据平面数据层1850中的(一个或多个)DB子网1830。数据平面数据层1850可以包括(一个或多个)DB子网1830,该(一个或多个)DB子网1830可以通信地耦合到包含在数据平面VCN 1818中的服务网关1836。

[0221] (一个或多个)非受信任应用程序子网1862可以包括一个或多个主要VNIC 1864(1)-(N),该主要VNIC 1864(1)-(N)可以通信地耦合到租户虚拟机(VM)1866(1)-(N)。每个租户VM 1866(1)-(N)可以通信地耦合到可以包含在相应容器出口VCN 1868(1)-(N)中的相应应用程序子网1867(1)-(N),该相应容器出口VCN 1868(1)-(N)可以包含在相应的客户租赁1870(1)-(N)中。相应的次级VNIC 1872(1)-(N)可以促进包含在数据平面VCN 1818中的(一个或多个)非受信任应用程序子网1862与包含在容器出口VCN 1868(1)-(N)中的应用程序子网之间的通信。每个容器出口VCN 1868(1)-(N)可以包括NAT网关1838,该NAT网关1838可以通信地耦合到公共互联网1854(例如,图16的公共互联网1654)。

[0222] 包含在控制平面VCN 1816中和包含在数据平面VCN 1818中的互联网网关1834可以通信地耦合到元数据管理服务1852(例如,图16的元数据管理系统1652),该元数据管理服务1852可以通信地耦合到公共互联网1854。公共互联网1854可以通信地耦合到包含在控制平面VCN 1816中和包含在数据平面VCN 1818中的NAT网关1838。包含在控制平面VCN 1816中和包含在数据平面VCN 1818中的服务网关1836可以通信地耦合到云服务1856。

[0223] 在一些实施例中,数据平面VCN 1818可以与客户租赁1870集成。在一些情况下,诸如在执行代码时可能期望支持的情况下,此集成对于IaaS提供商的客户可能是有用的或期望的。客户可以提供代码来运行,该代码可能是破坏性的,可能与其他客户资源通信,或者可能以其他方式导致不良影响。作为对此的响应,IaaS提供商可以确定是否运行客户给予IaaS提供商的代码。

[0224] 在一些示例中,IaaS提供商的客户可以向IaaS提供商授予临时网络访问权限并请求要附接到数据平面层应用程序1846的功能。运行该功能的代码可以在VM 1866(1)-(N)中执行,并且代码可以不被配置为在数据平面VCN 1818上的任何其他地方运行。每个VM 1866(1)-(N)可以连接到一个客户租赁1870。包含在VM 1866(1)-(N)中的相应容器1871(1)-(N)可以被配置为运行代码。在此情况下,可以存在双重隔离(例如,容器1871(1)-(N)运行代码,其中容器1871(1)-(N)可以至少包含在VM 1866(1)-(N)中,该VM 1866(1)-(N)包含在(一个或多个)非受信任应用程序子网1862中),这可以帮助防止不正确的或以其他方式不期望的代码损坏IaaS提供商的网络或损坏不同客户的网络。容器1871(1)-(N)可以通信地耦合到客户租赁1870并且可以被配置为从客户租赁1870发送或接收数据。容器1871(1)-(N)可以不被配置为发送或接收来自数据平面VCN 1818中的任何其他实体的数据。在运行代码完成后,IaaS提供商可以终止或以其他方式处置容器1871(1)-(N)。

[0225] 在一些实施例中,(一个或多个)受信任应用程序子网1860可以运行IaaS提供商拥有或操作的代码。在此实施例中,(一个或多个)受信任应用程序子网1860可以通信地耦合到(一个或多个)DB子网1830并且被配置为执行(一个或多个)DB子网1830中的CRUD操作。(一个或多个)非受信任应用程序子网1862可以通信地耦合到(一个或多个)DB子网1830,但在此实施例中,(一个或多个)非受信任应用程序子网可以被配置为执行(一个或多个)DB子网1830中的读取操作。可以包含在每个客户的VM 1866(1)-(N)中的并且可以运行来自客户的代码的容器1871(1)-(N)可以不与(一个或多个)DB子网1830通信地耦合。

[0226] 在其他实施例中,控制平面VCN 1816和数据平面VCN 1818可以不直接通信地耦合。在此实施例中,控制平面VCN 1816和数据平面VCN 1818之间可以没有直接通信。然而,通信可以通过至少一种方法间接发生。LPG 1810可以由IaaS提供商建立,其可以促进控制

平面VCN 1816和数据平面VCN 1818之间的通信。在另一示例中,控制平面VCN 1816或数据平面VCN 1818可以经由服务网关1836做出对云服务1856的调用。例如,从控制平面VCN 1816到云服务1856的调用可以包括对可以与数据平面VCN 1818通信的服务的请求。

[0227] 图19是图示根据至少一个实施例的IaaS架构的另一示例模式的框图1900。服务运营商1902(例如,图16的服务运营商1602)可以通信地耦合到安全主机租赁1904(例如,图16的安全主机租赁1604),该安全主机租赁1904可以包括虚拟云网络(VCN)1906(例如,图16的VCN 1606)和安全主机子网1908(例如,图16的安全主机子网1608)。VCN 1906可以包括LPG 1910(例如,图16的LPG 1610),该LPG 1910可以经由包含在SSH VCN 1912中的LPG 1910通信地耦合到SSH VCN 1912(例如,图16的SSH VCN 1612)。SSH VCN 1912可以包括SSH子网1914(例如,图16的SSH子网1614),并且SSH VCN 1912可以经由包含在控制平面VCN 1916中的LPG 1910通信地耦合到控制平面VCN 1916(例如,图16的控制平面VCN 1616)和经由包含在数据平面VCN 1918中的LPG 1910通信地耦合到数据平面VCN 1918(例如,图16的数据平面1618)。控制平面VCN 1916和数据平面VCN 1918可以包含在服务租赁1919(例如,图16的服务租赁1619)中。

[0228] 控制平面VCN 1916可以包括可以包括(一个或多个)LB子网1922(例如,图16的(一个或多个)LB子网1622))的控制平面DMZ层1920(例如,图16的控制平面DMZ层1620),可以包括(一个或多个)应用程序子网1926(例如,图16的(一个或多个)应用程序子网1626)的控制平面应用程序层1924(例如,图16的控制平面应用程序层1624)、可以包括(一个或多个)DB子网1930(例如,图18的(一个或多个)DB子网1830)的控制平面数据层1928(例如,图16的控制平面数据层1628)。包含在控制平面DMZ层1920中的(一个或多个)LB子网1922可以通信地耦合到包含在控制平面应用程序层1924中的(一个或多个)应用程序子网1926和可以包含在控制平面VCN 1916中的互联网网关1934(例如,图16的互联网网关1634),并且(一个或多个)应用程序子网1926可以通信地耦合到包含在控制平面数据层1928中的(一个或多个)DB子网1930以及服务网关1936(例如,图16的服务网关)和网络地址转换(NAT)网关1938(例如,图16的NAT网关1638)。控制平面VCN 1916可以包括服务网关1936和NAT网关1938。

[0229] 数据平面VCN 1918可以包括数据平面应用程序层1946(例如,图16的数据平面应用程序层1646)、数据平面DMZ层1948(例如,图16的数据平面DMZ层1648)、以及数据平面数据层1950(例如,图16的数据平面数据层1650)。数据平面DMZ层1948可以包括(一个或多个)LB子网1922,该(一个或多个)LB子网1922可以通信地耦合到数据平面应用程序层1946的(一个或多个)受信任应用程序子网1960(例如,图18的(一个或多个)受信任应用程序子网1860)和(一个或多个)非受信任应用程序子网1962(例如,图18的(一个或多个)非受信任应用程序子网1862)和包含在数据平面VCN 1918中的互联网网关1934。(一个或多个)受信任应用程序子网1960可以通信地耦合到包含在数据平面VCN 1918中的服务网关1936,包含在数据平面VCN 1918中的NAT网关1938以及包含在数据平面数据层1950中的(一个或多个)DB子网1930。(一个或多个)非受信任应用程序子网1962可以通信地耦合到包含在数据平面VCN 1918中的服务网关1936和包含在数据平面数据层1950中的(一个或多个)DB子网1930。数据平面数据层1950可以包括可以通信地耦合到包含在数据平面VCN 1918中的服务网关1936的(一个或多个)DB子网1930。

[0230] (一个或多个)非受信任应用程序子网1962可以包括主要VNIC 1964(1) - (N),该主

要VNIC 1964 (1) - (N) 可以通信地耦合到驻留在 (一个或多个) 非受信任应用程序子网1962内的租户虚拟机 (VM) 1966 (1) - (N) 。每个租户VM 1966 (1) - (N) 可以在相应的容器1967 (1) - (N) 中运行代码, 并且通信地耦合到可以包含在数据平面应用程序层1946中的应用程序子网1926, 该数据平面应用程序层1946可以包含在容器出口VCN 1968中。相应的次级VNIC 1972 (1) - (N) 可以促进包含在数据平面VCN 1918中的 (一个或多个) 非受信任应用程序子网1962与包含在容器出口VCN 1968中的应用程序子网之间的通信。容器出口VCN可以包括NAT网关1938, 该NAT网关1938可以通信地耦合到公共互联网1954 (例如, 图16的公共互联网1654) 。

[0231] 包含在控制平面VCN 1916中和包含在数据平面VCN 1918中的互联网网关1934可以通信地耦合到元数据管理服务1952 (例如, 图16的元数据管理系统1652), 该元数据管理服务1952可以通信地耦合到公共互联网1954。公共互联网1954可以通信地耦合到包含在控制平面VCN 1916中和包含在数据平面VCN 1918中的NAT网关1938。包含在控制平面VCN 1916中和包含在数据平面VCN 1918中的服务网关1936可以通信地耦合到云服务1956。

[0232] 在一些示例中, 由图19的框图1900的架构图示的模式可以被视为是图18的框图1800的架构所示的模式, 并且如果IaaS提供商不能直接与客户通信 (例如, 断开连接的区域), 则可能是IaaS提供商的客户所期望的。包含在每个客户的VM 1966 (1) - (N) 中的相应容器1967 (1) - (N) 可由客户实时访问。容器1967 (1) - (N) 可被配置为对包含在数据平面应用程序层1946的 (一个或多个) 应用程序子网1926中的相应次级VNIC 1972 (1) - (N) 做出调用, 该数据平面应用程序层1946可以包含在容器出口VCN 1968中。次级VNIC 1972 (1) - (N) 可以将调用传送到NAT网关1938, 该NAT网关1938可以将调用传送到公共互联网1954。在此示例中, 可以由客户实时访问的容器1967 (1) - (N) 可以与控制平面VCN 1916隔离, 并且可以与包含在数据平面VCN 1918中的其他实体隔离。容器1967 (1) - (N) 也可以与来自其他客户的资源隔离。

[0233] 在其他示例中, 客户可以使用容器1967 (1) - (N) 来调用云服务1956。在此示例中, 客户可以在容器1967 (1) - (N) 中运行请求来自云服务1956的服务的代码。容器1967 (1) - (N) 可以将此请求传送到次级VNIC 1972 (1) - (N), 该次级VNIC 1972 (1) - (N) 可以将请求传送到NAT网关, 该NAT网关可以将请求传送到公共互联网1954。公共互联网1954可以经由互联网网关1934将请求传送到包含在控制平面VCN 1916中的 (一个或多个) LB子网1922。响应于确定请求有效, (一个或多个) LB子网可以将请求传送到 (一个或多个) 应用程序子网1926, 该 (一个或多个) 应用程序子网1926可以经由服务网关1936将请求传送到云服务1956。

[0234] 应当理解, 图中描绘的IaaS架构1600、1700、1800、1900可以具有除描绘的那些之外的其他组件。此外, 图中所示的实施例仅是可以结合本公开的实施例的云基础设施系统的一些示例。在一些其他实施例中, IaaS系统可以具有比图中所示更多或更少的组件, 可以组合两个或更多组件, 或者可以具有组件的不同配置或布置。

[0235] 在某些实施例中, 本文描述的IaaS系统可以包括以自助服务、基于订阅、弹性可扩展、可靠、高度可用和安全的方式交付给客户的一套应用、中间件和数据库服务供应。这样的IaaS系统的示例是本申请人提供的Oracle云基础设施 (OCI) 。

[0236] 图20图示了其中可以实现各种实施例的示例计算机系统2000。系统2000可用于实现上面描述的任何计算机系统。如图所示, 计算机系统2000包括处理单元2004, 该处理单元

2004经由总线子系统2002与多个外围子系统通信。这些外围子系统可以包括处理加速单元2006、I/O子系统2008、存储子系统2018和通信子系统2024。存储子系统2018包括有形计算机可读存储介质2022和系统存储器2010。

[0237] 总线子系统2002提供了一种机制,用于让计算机系统2000的各个组件和子系统按预期彼此通信。尽管总线子系统2002被示意性地示为单个总线,但是总线子系统的替代实施例可以利用多个总线。总线子系统2002可以是若干类型的总线结构中的任何一种,包括存储器总线或存储器控制器、外围总线和使用各种各样的总线架构中的任何一种的局部总线。例如,这样的架构可以包括工业标准架构 (ISA) 总线、微通道架构 (MCA) 总线、增强型ISA (EISA) 总线、视频电子标准协会 (VESA) 局部总线和可以被实现为按照IEEE P1386.1标准制造的夹层总线的外围组件互连 (PCI) 总线。

[0238] 可以被实现为一个或多个集成电路 (例如,常规微处理器或微控制器) 的处理单元2004控制计算机系统2000的操作。一个或多个处理器可以被包括在处理单元2004中。这些处理器可以包括单核或多核处理器。在某些实施例中,处理单元2004可以被实现为一个或多个独立处理单元2032和/或2034,其中每个处理单元中包括单核或多核处理器。在其他实施例中,处理单元2004也可以被实现为通过将两个双核处理器集成到单个芯片中而形成的四核处理单元。

[0239] 在各种实施例中,处理单元2004可以响应于程序代码执行各种各样的程序并且可以维护多个并发执行的程序或进程。在任何给定时间,要被执行的一些或全部程序代码可以驻留在(一个或多个)处理器2004和/或存储子系统2018中。通过合适的编程,(一个或多个)处理器2004可以提供上面描述的各种功能。此外计算机系统2000可以包括处理加速单元2006,该处理加速单元2006可包括数字信号处理器 (DSP)、专用处理器和/或类似物。

[0240] I/O子系统2008可以包括用户接口输入设备和用户接口输出设备。用户接口输入设备可以包括键盘、诸如鼠标或轨迹球的指向设备、并入显示器中的触摸板或触摸屏、滚轮、点击轮、拨号盘、按钮、开关、小键盘、带有语音命令识别系统的音频输入设备、麦克风和其他类型的输入设备。用户接口输入设备可以包括,例如,运动感测和/或手势识别设备,诸如使得用户能够通过使用手势和口头命令的自然用户接口来控制输入设备 (诸如Microsoft **Xbox®** 360游戏控制器) 并与之交互的Microsoft **Kinect®** 运动传感器。用户接口输入设备还可以包括眼姿 (eye gesture) 识别设备,诸如Google **Glass®** 眨眼检测器,其检测来自用户的眼睛活动 (例如,在拍照和/或做出菜单选择时“眨眼”),并将眼姿转换为到输入设备 (例如,Google **Glass®**) 的输入。此外,用户接口输入设备可以包括使得用户能够通过语音命令与语音识别系统 (例如,**Siri®** 导航器) 交互的语音识别感测设备。

[0241] 用户接口输入设备还可以包括但不限于三维 (3D) 鼠标、操纵杆或指示杆、游戏手柄和图形输入板,以及音频/视频设备,诸如扬声器、数码相机、数码摄像机、便携式媒体播放器、网络摄像头、图像扫描仪、指纹扫描仪、条码阅读器3D扫描仪、3D打印机、激光测距仪和眼睛注视跟踪设备。此外,用户接口输入设备可以包括例如医学成像输入设备,诸如计算机断层扫描、磁共振成像、正电子发射断层扫描、医学超声设备。用户接口输入设备还可以包括例如音频输入设备,诸如MIDI键盘、数字乐器等。

[0242] 用户接口输出设备可以包括显示子系统、指示灯或诸如音频输出设备的非视觉显

示器等。显示子系统可以是阴极射线管 (CRT)、诸如使用液晶显示器 (LCD) 或等离子显示器的平板设备、投影设备、触摸屏等。通常,术语“输出设备”的使用旨在包括用于将信息从计算机系统2000输出到用户或其他计算机的所有可能类型的设备和机制。例如,用户界面输出设备可以包括但不限于以视觉方式输送文本、图形和音频/视频信息的各种各样的显示设备,诸如监视器、打印机、扬声器、耳机、汽车导航系统、绘图仪、语音输出设备和调制解调器。

[0243] 计算机系统2000可以包括存储子系统2018,该存储子系统2018可以提供用于存储提供本公开中描述的实施例的功能的软件和数据构造的有形非暂时性计算机可读存储介质。软件可以包括程序、代码模块、指令、脚本等,当其由处理单元2004的一个或多个核心或处理器执行时,提供上述功能。存储子系统2018还可以提供用于存储根据本公开使用的数据的存储库。

[0244] 如图20中的示例所示,存储子系统2018可以包括各种组件,包括系统存储器2010、计算机可读存储介质2022和计算机可读存储介质读取器2020。系统存储器2010可以存储可由处理单元2004加载和执行的程序指令。系统存储器2010还可以存储在指令执行期间使用的数据和/或在程序指令执行期间生成的数据。各种不同种类的程序可以被加载到系统存储器2010中,包括但不限于客户端应用、Web浏览器、中间层应用、关系型数据库管理系统 (RDBMS)、虚拟机、容器等。

[0245] 系统存储器2010还可以存储操作系统2016。操作系统2016的示例可以包括各种版本的Microsoft **Windows®**、Apple **Macintosh®**和/或Linux操作系统、各种各样的商用 **UNIX®**或类UNIX操作系统 (包括但不限于各种GNU/Linux操作系统、Google **Chrome®OS**等) 和/或移动操作系统,诸如iOS、**Windows®Phone**、**Android®OS**、**BlackBerry®OS**和**Palm®OS**操作系统。在计算机系统2000执行一个或多个虚拟机的某些实现中,虚拟机连同其客户操作系统 (GOS) 可以被加载到系统存储器2010中,并由处理单元2004的一个或多个处理器或核心执行。

[0246] 依赖于计算机系统2000的类型,系统存储器2010可以有不同的配置。例如,系统存储器2010可以是易失性存储器 (诸如随机存取存储器 (RAM)) 和/或非易失性存储器 (诸如只读存储器 (ROM)、闪存等)。可以提供不同类型的RAM配置,包括静态随机存取存储器 (SRAM)、动态随机存取存储器 (DRAM) 等。在一些实现中,系统存储器2010可以包括基本输入/输出系统 (BIOS),其包含有助于在计算机系统2000内的元素之间转移信息 (诸如在启动期间) 的基本例程。

[0247] 计算机可读存储介质2022可以表示远程、本地、固定和/或可移动存储设备以及用于临时和/或更永久地包含、存储计算机系统2000使用的计算机可读信息的存储介质,包括可由计算机系统2000的处理单元2004执行的指令。

[0248] 计算机可读存储介质2022可以包括本领域已知或使用的任何适当介质,包括存储介质和通信介质,诸如但不限于以任何方法或技术实现的用于存储和/或传输信息的易失性和非易失性、可移动和不可移动介质。这可以包括诸如RAM、ROM、电子可擦除可编程ROM (EEPROM)、闪存或其他存储器技术、CD-ROM、数字多功能磁盘 (DVD) 或其他光学存储、磁带盒、磁带、磁盘存储或其他磁存储设备的有形计算机可读存储介质或其他有形计算机可读

介质。

[0249] 举例来说,计算机可读存储介质2022可以包括从不可移动非易失性磁介质读取或向其写入的硬盘驱动、从可移动非易失性磁介质读取或向其写入的磁盘驱动、以及从可移动非易失性光盘(诸如CD ROM、DVD和**Blu-Ray®**光盘)或其他光学介质读取或向其写入的光盘驱动。计算机可读存储介质2022可以包括但不限于**Zip®**驱动、闪存卡、通用串行总线(USB)闪存驱动、安全数字(SD)卡、DVD盘、数字视频磁带等。计算机可读存储介质2022还可以包括基于非易失性存储器的固态驱动(SSD)(诸如基于闪存的SSD、企业闪存驱动、固态ROM等)、基于易失性存储器的SSD(诸如基于固态RAM、动态RAM、静态RAM、DRAM的SSD)、磁阻RAM(MRAM) SSD以及使用DRAM和基于闪存的SSD组合的混合SSD。磁盘驱动及其相关联的计算机可读介质可以为计算机系统2000提供计算机可读指令、数据结构、程序模块和其他数据的非易失性存储。

[0250] 可由处理单元2004的一个或多个处理器或核心执行的机器可读指令可以存储在非暂时性计算机可读存储介质上。非暂时性计算机可读存储介质可以包括包括易失性存储器存储设备和/或非易失性存储设备的物理上有形的存储器或存储设备。非暂时性计算机可读存储介质的示例包括磁存储介质(例如,磁盘或磁带)、光存储介质(例如DVD、CD)、各种类型的RAM、ROM或闪存、硬盘驱动、软盘驱动、可拆卸存储器驱动(例如,USB驱动)或其他类型的存储设备。

[0251] 通信子系统2024提供到其他计算机系统和网络的接口。通信子系统2024用作用于从计算机系统2000接收来自其他系统的数据以及向其他系统发送数据的接口。例如,通信子系统2024可以使得计算机系统2000能够经由互联网连接到一个或多个设备。在一些实施例中,通信子系统2024可以包括用于访问无线语音和/或数据网络(例如,使用蜂窝电话技术、先进数据网络技术,诸如3G、4G或EDGE(增强数据速率全球演进)、WiFi(IEEE 802.11系列标准,或其他移动通信技术,或其任何组合))的射频(RF)收发器组件、全球定位系统(GPS)接收器组件和/或其他组件。在一些实施例中,除了或替代无线接口,通信子系统2024可以提供有线网络连接性(例如,以太网)。

[0252] 在一些实施例中,通信子系统2024还可以代表可以使用计算机系统2000的一个或多个用户以结构化和/或非结构化的数据馈送2026、事件流2028、事件更新2030等形式接收输入通信。

[0253] 举例来说,通信子系统2024可以被配置为从社交网络和/或其他通信服务的用户实时接收数据馈送2026,诸如**Twitter®**馈送、**Facebook®**更新、诸如丰富站点摘要(RSS)馈送的web馈送和/或来自一个或多个第三方信息源的实时更新。

[0254] 此外,通信子系统2024还可以被配置为接收连续数据流形式的数据,该数据可以包括实时事件的事件流2028和/或事件更新2030,其本质上可以是连续的或无边界的,而没有明确的结束。生成连续数据的应用的示例可以包括例如传感器数据应用、金融股票、网络性能测量工具(例如,网络监控和流量管理应用)、点击流分析工具、汽车流量监控等。

[0255] 通信子系统2024还可以被配置为将结构化和/或非结构化数据馈送2026、事件流2028、事件更新2030等输出到一个或多个数据库,这些数据库可以与耦合到计算机系统2000的一个或多个流数据源计算机通信。

[0256] 计算机系统2000可以是各种类型中的一种,包括手持便携式设备(例如,**iPhone®**蜂窝电话、**iPad®**计算平板、PDA)、可穿戴设备(例如,Google**Glass®**头戴式显示器)、PC、工作站、大型机、信息亭、服务器机架或任何其他数据处理系统。

[0257] 由于计算机和网络不断更改的性质,图中所描绘的计算机系统2000的描述仅旨在作为特定示例。具有比图中所描绘的系统更多或更少的组件的许多其他配置是可能的。例如,定制硬件还可以被使用和/或特定元素可以以硬件、固件、软件(包括小应用程序)或组合来实现。此外,与其他计算设备(诸如网络输入/输出设备)的连接可以被采用。基于本文提供的公开和教导,本领域普通技术人员将理解实现各种实施例的其他方式和/或方法。

[0258] 尽管已经描述了特定实施例,但是各种修改、变更、替代构造和等同物也包含在本公开的范围。实施例不限于某些特定数据处理环境内的操作,而是可以自由地在多个数据处理环境内操作。此外,尽管实施例已经使用特定系列的事务和步骤被描述,但是对于本领域技术人员来说应显而易见的是,本公开的范围不限于所描述的系列的事务和步骤。上述实施例的各种特征和方面可以单独或联合使用。

[0259] 此外,虽然实施例已经使用硬件和软件的特定组合被描述,但是应该认识到硬件和软件的其他组合也在本公开的范围。实施例可以仅以硬件、或仅以软件、或使用其的组合来实现。本文描述的各种过程可以在相同的处理器或不同处理器上以任何组合实现。因此,在组件或模块被描述为被配置为执行某些操作的情况下,这样的配置可以例如通过设计电子电路来执行操作、通过对可编程电子电路(诸如微处理器)编程来执行操作、或其任何组合来实现。过程可以使用包括但不限于用于过程间通信的常规技术的各种各样的技术来通信,并且不同的过程对可以使用不同的技术,或者相同过程对可以在不同的时间使用不同的技术。

[0260] 因此,说明书和附图被认为是说明性的而不是限制性意义。然而,将显而易见,在不偏离如权利要求中阐述的更广泛的精神和范围的情况下,可以对其做出添加、删减、删除和其他修改和更改。因此,尽管已经描述了特定的公开实施例,但是这些并不旨在限制。各种修改和等同物都在所附权利要求的范围内。

[0261] 在描述所公开的实施例的上下文中(特别是在所附权利要求的上下文中)术语“(a)”和“一个(an)”和“该(the)”以及类似指代词的使用应被解释为涵盖单数和复数两者,除非本文另有说明或与上下文明显矛盾。除非另有说明,否则术语“包括(comprising)”、“具有(having)”、“包括(including)”和“包含(containing)”应被解释为开放式术语(即,意味着“包括,但不限于”)。术语“连接到”应被解释为部分或全部包含在内、附接到或结合在一起,即使有东西介入。除非本文另有说明,否则本文数值范围的列举仅旨在用作单独指代落入该范围内的每个分开值的简略表达方法,并且每个分开值都被并入说明书中,如同其在本文中被单独列举一样。除非本文另有说明或与上下文另有明显矛盾,否则本文描述的所有方法可以任何合适的次序执行。除非另有要求保护,否则本文提供的任何和所有示例或示例性语言(例如,“诸如”)的使用仅旨在更好地阐明实施例并且不对本公开的范围造成限制。说明书中的任何语言都不应被解释为指示任何未要求保护的元素对于本公开的实践是必不可少的。

[0262] 除非另有明确说明,诸如短语“X、Y或Z中的至少一个”的析取性语言旨在在上下文中被理解为通常用于表示项目、术语等,可以是X、Y或Z或其任何组合(例如X、Y和/或Z)。因

此,这样的析取性语言通常不旨在也不应该暗示某些实施例需要X中的至少一种、Y中的至少一种或Z中的至少一种每个存在。

[0263] 本文描述了本公开的优选实施例,包括已知的用于执行本公开的最佳模式。在阅读前述描述后,那些优选实施例的变化对于本领域的普通技术人员可以变得清晰。普通技术人员应该能够适当地采用这样的变化,并且本公开可以除如本文具体描述之外地被实践。因此,本公开包括如适用法律允许的所附权利要求中所列主题的所有修改和等同物。此外,除非本文另有说明,否则上述元素在其所有可能变化中的任何组合都包含在本公开中。

[0264] 本文引用的所有参考,包括出版物、专利申请和专利,通过引用并入本文,其程度如同每个参考被个别且具体地指示为通过引用整体被并入并在本文中阐述。

[0265] 在前述说明书中,参考本公开的特定实施例描述了本公开的方面,但是本领域技术人员将认识到本公开不限于此。上述公开的各种特征和方面可以单独或联合使用。此外,可以在超出本文描述的那些环境和应用的任何数量的环境和应用中利用实施例,而不脱离本说明书的更广泛的精神和范围。因此,说明书和附图被认为是说明性的而不是限制性的。

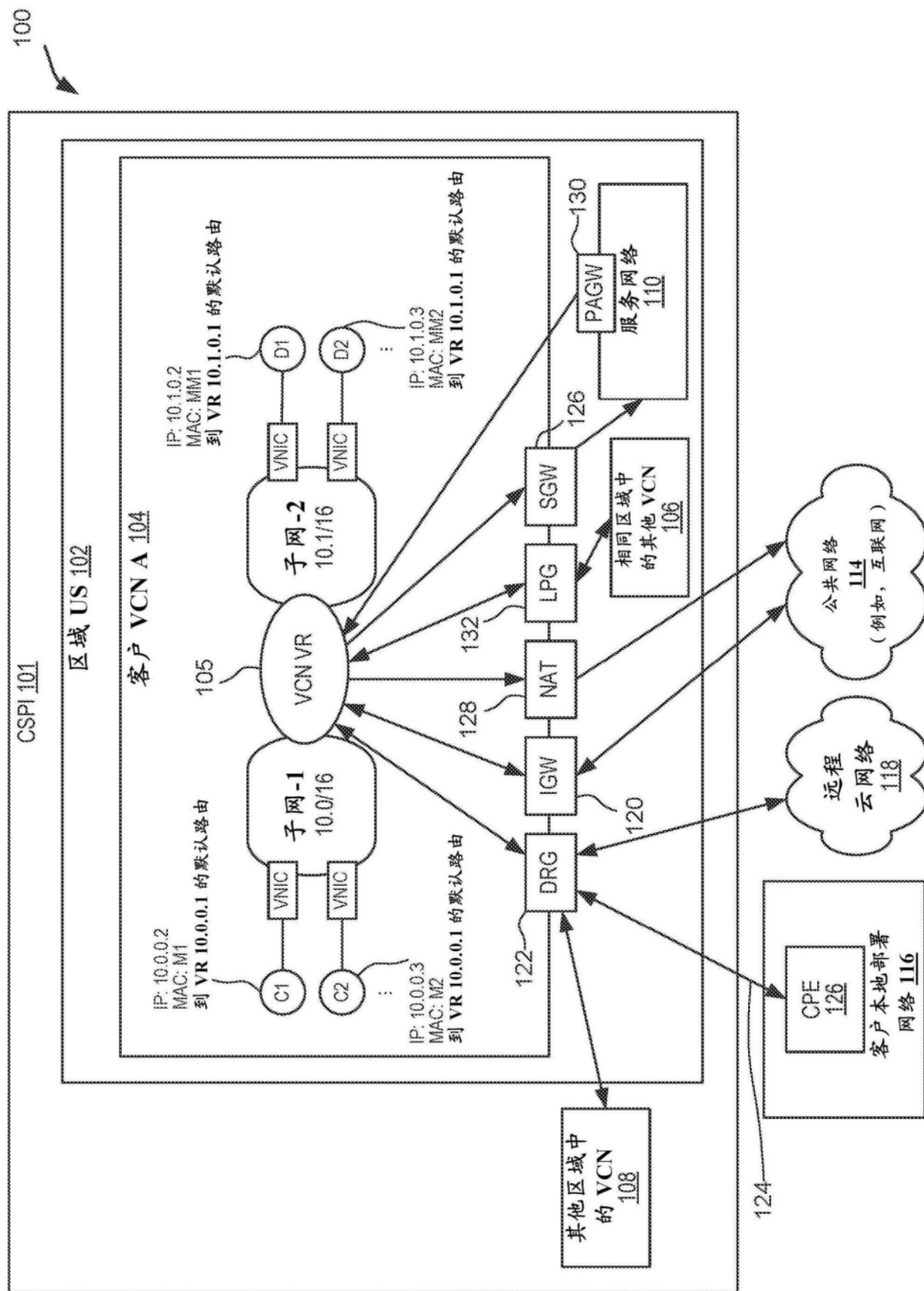


图1

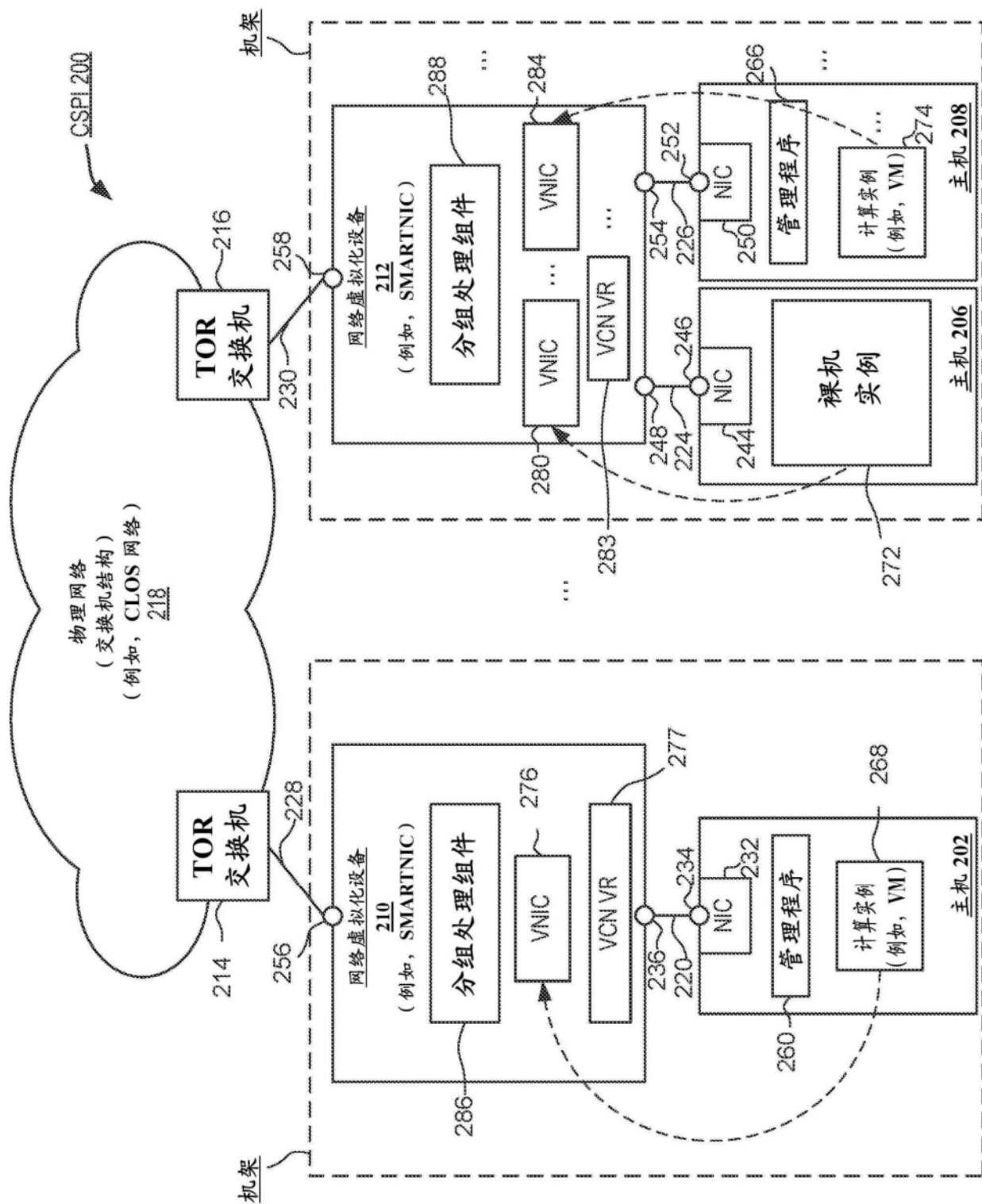


图2

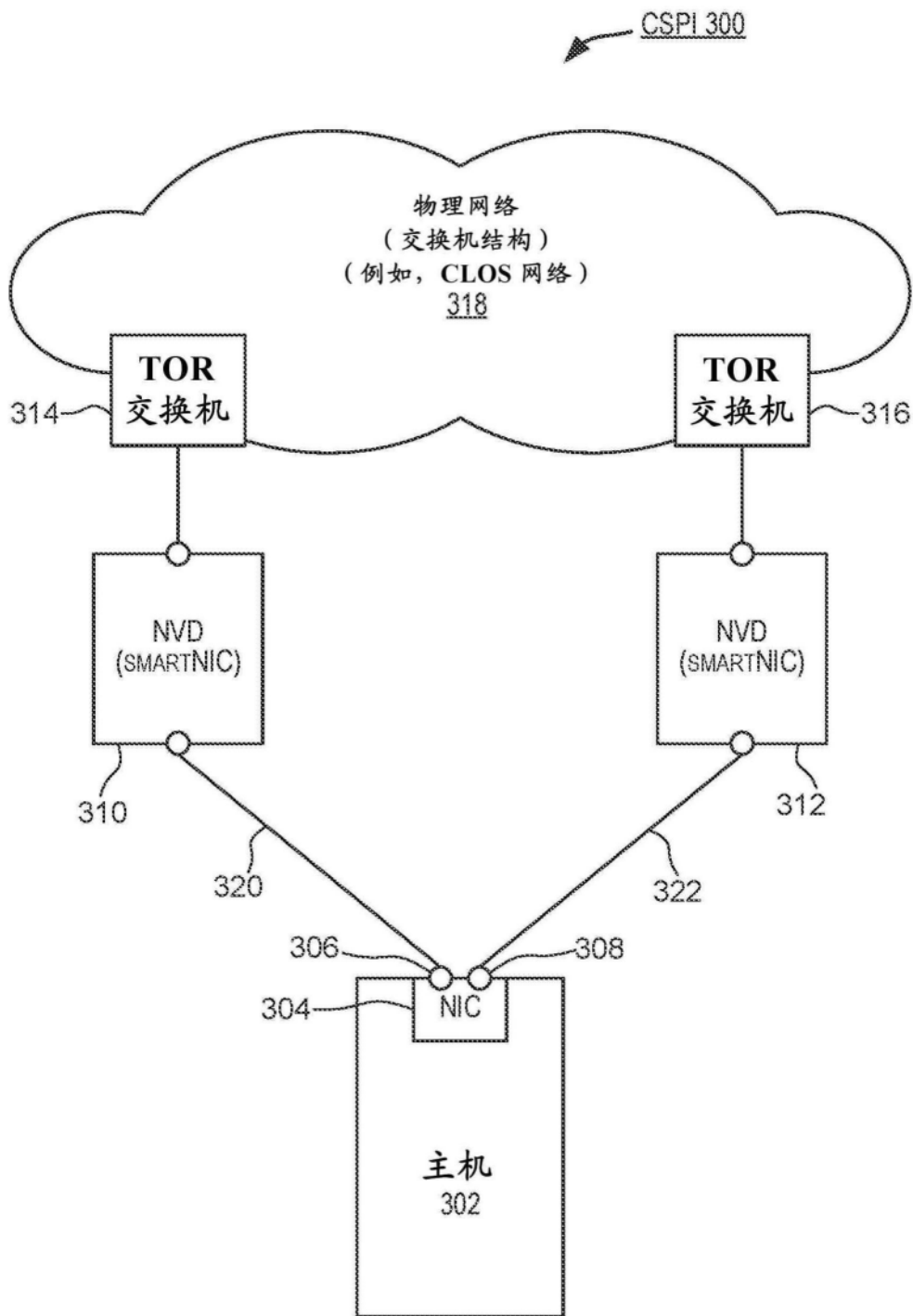


图3

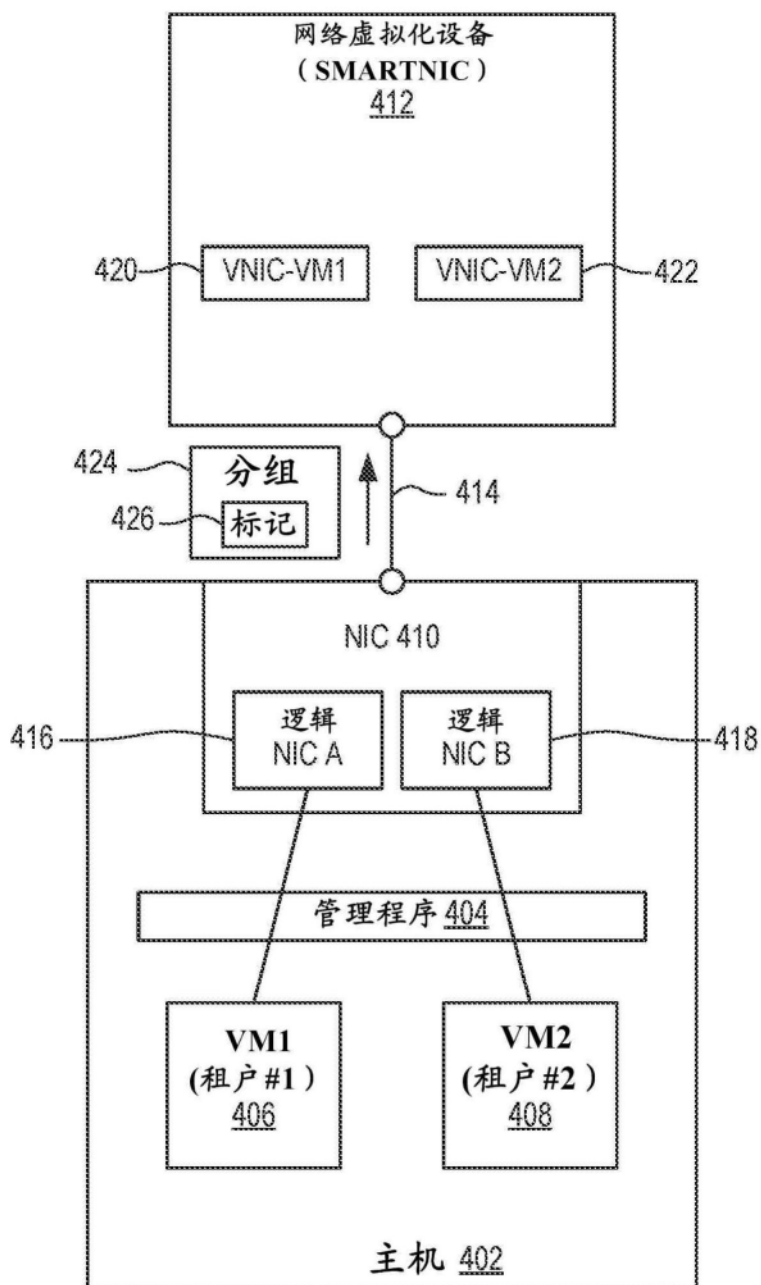


图4

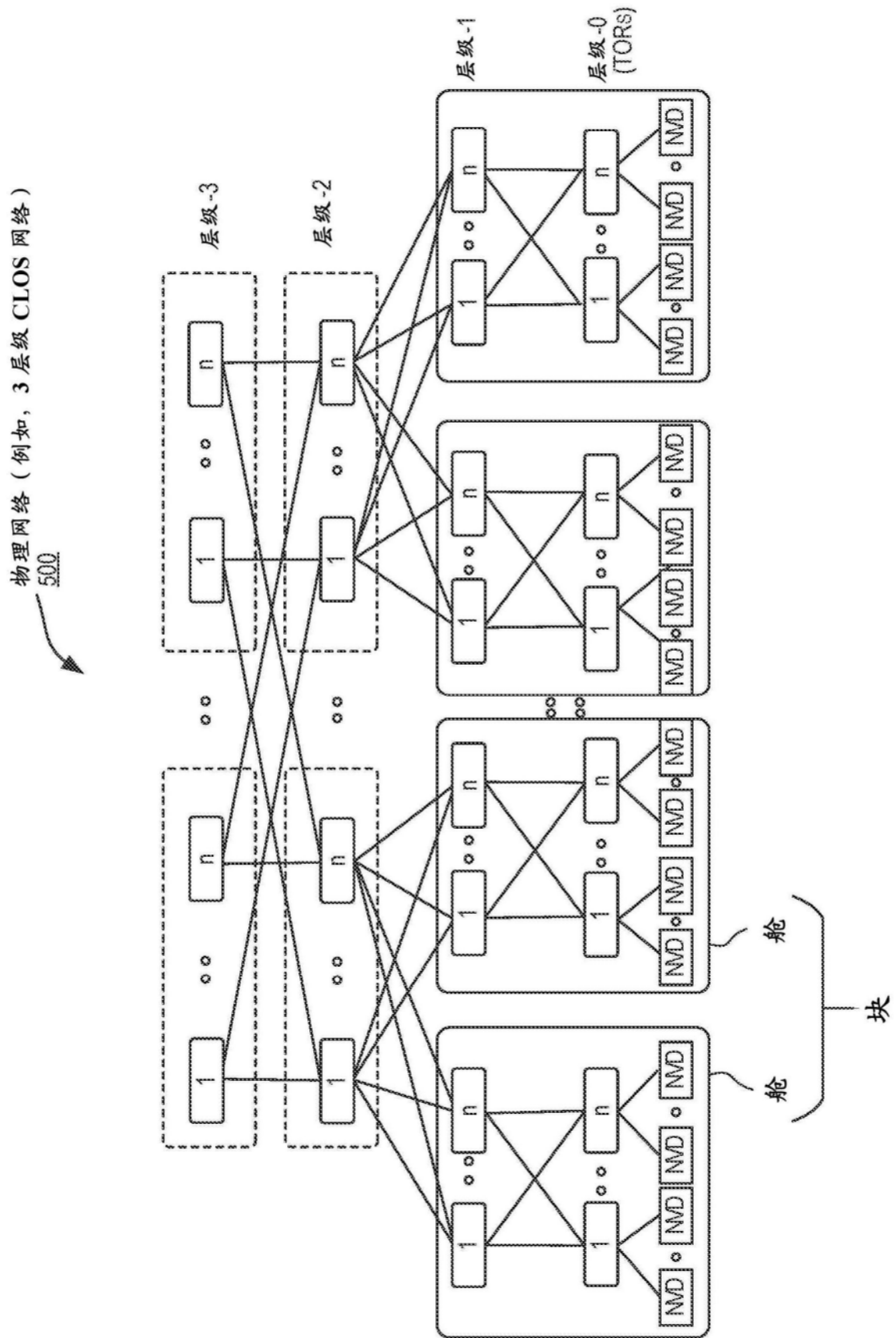


图5

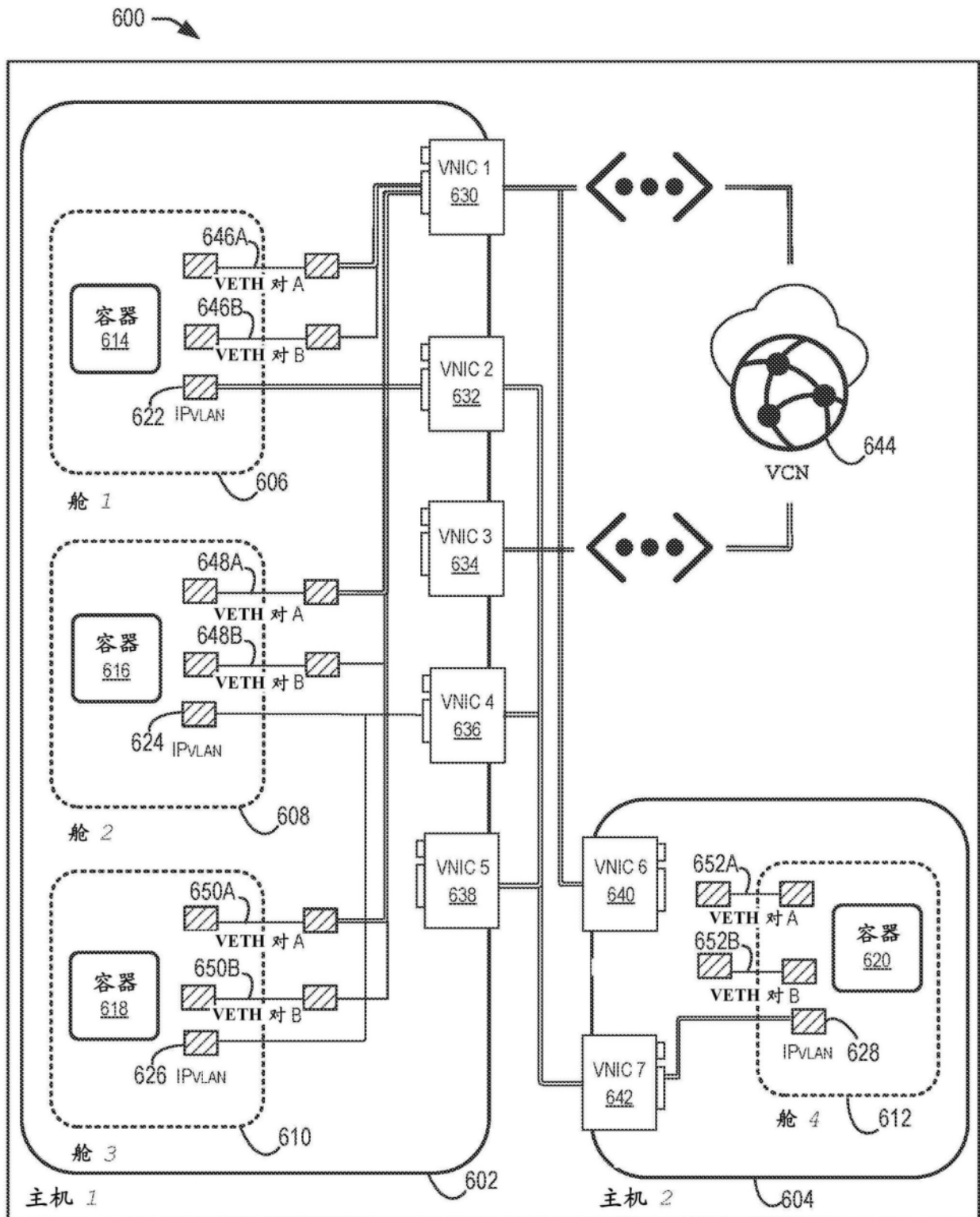


图6

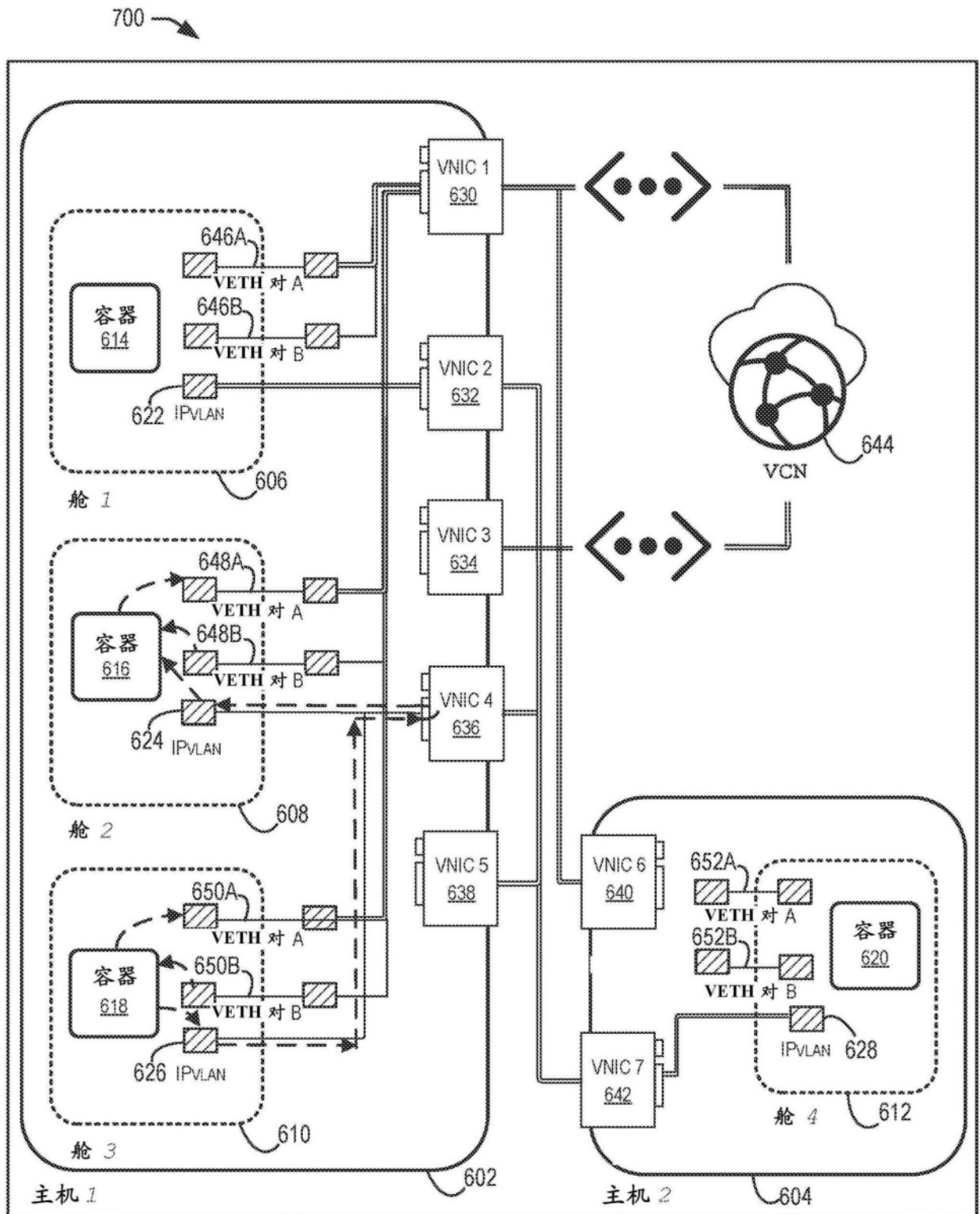


图7

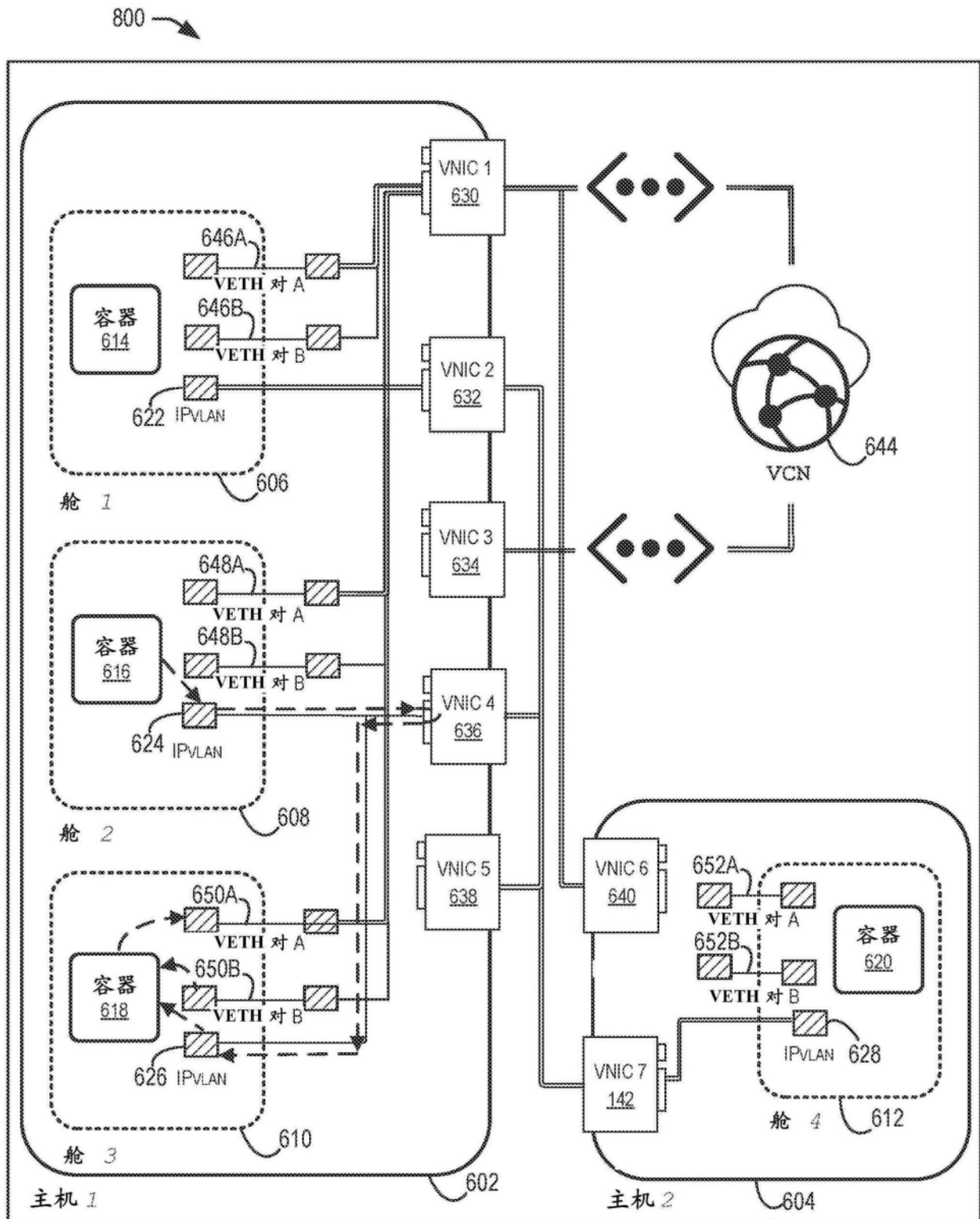


图8

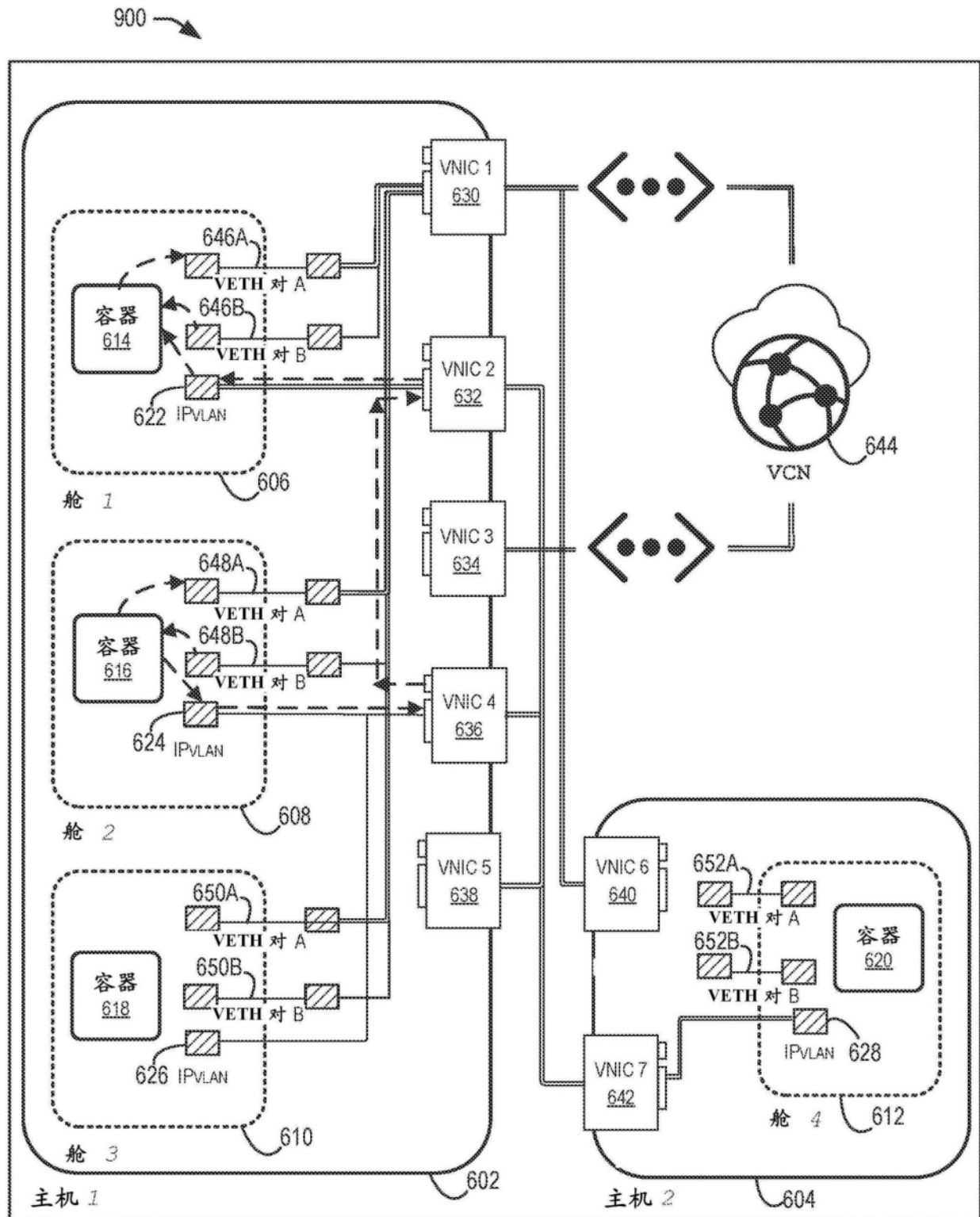


图9

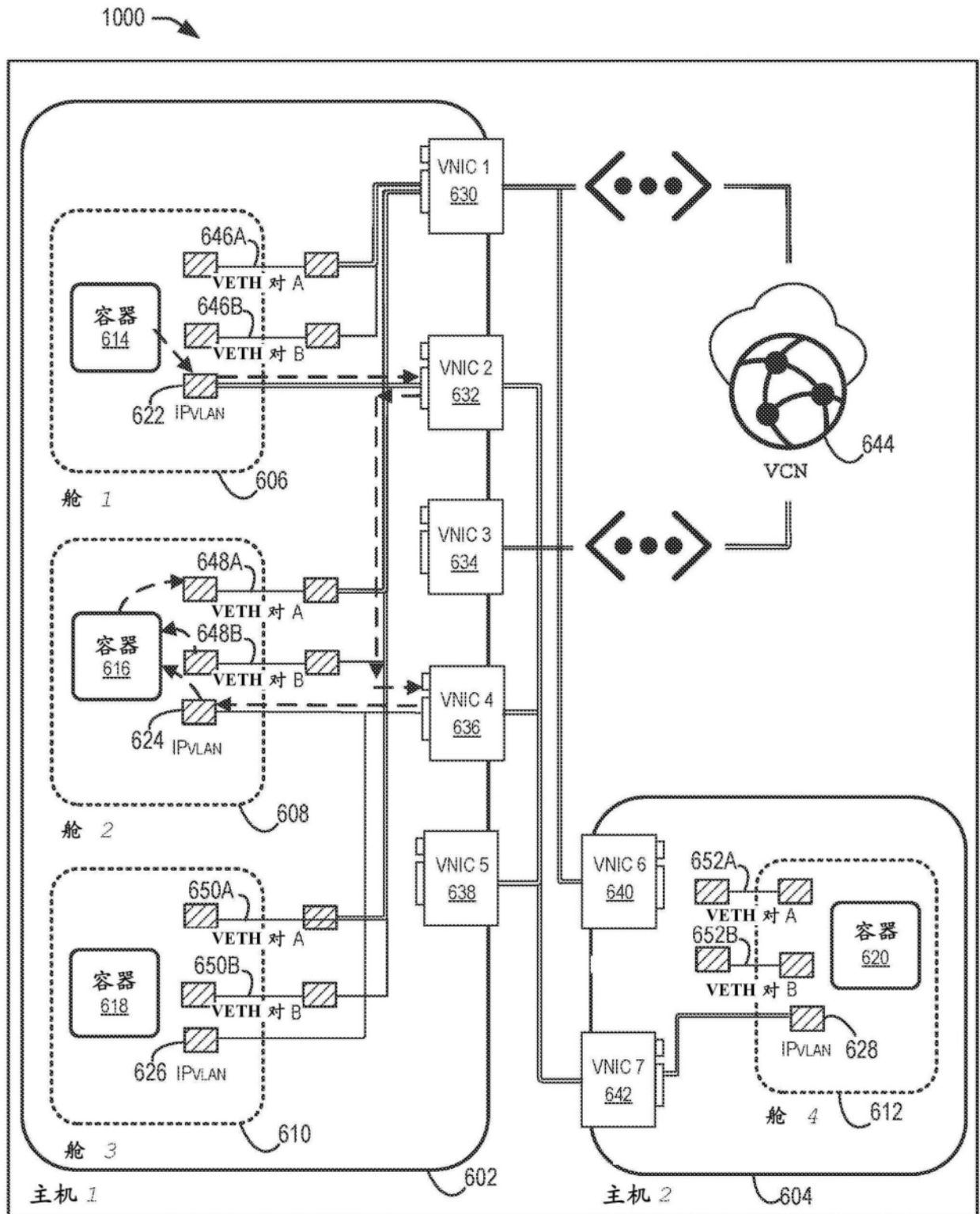


图10

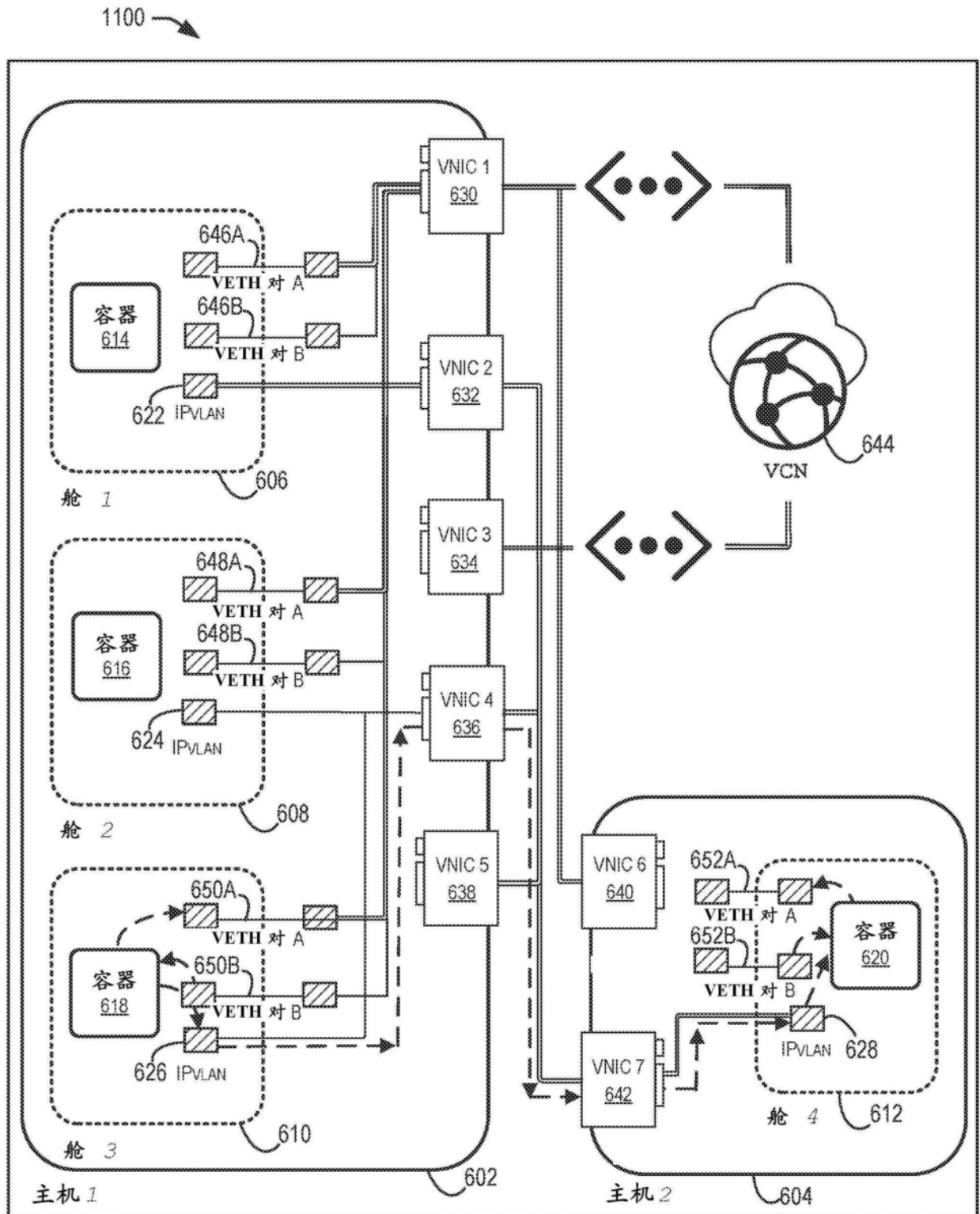


图11

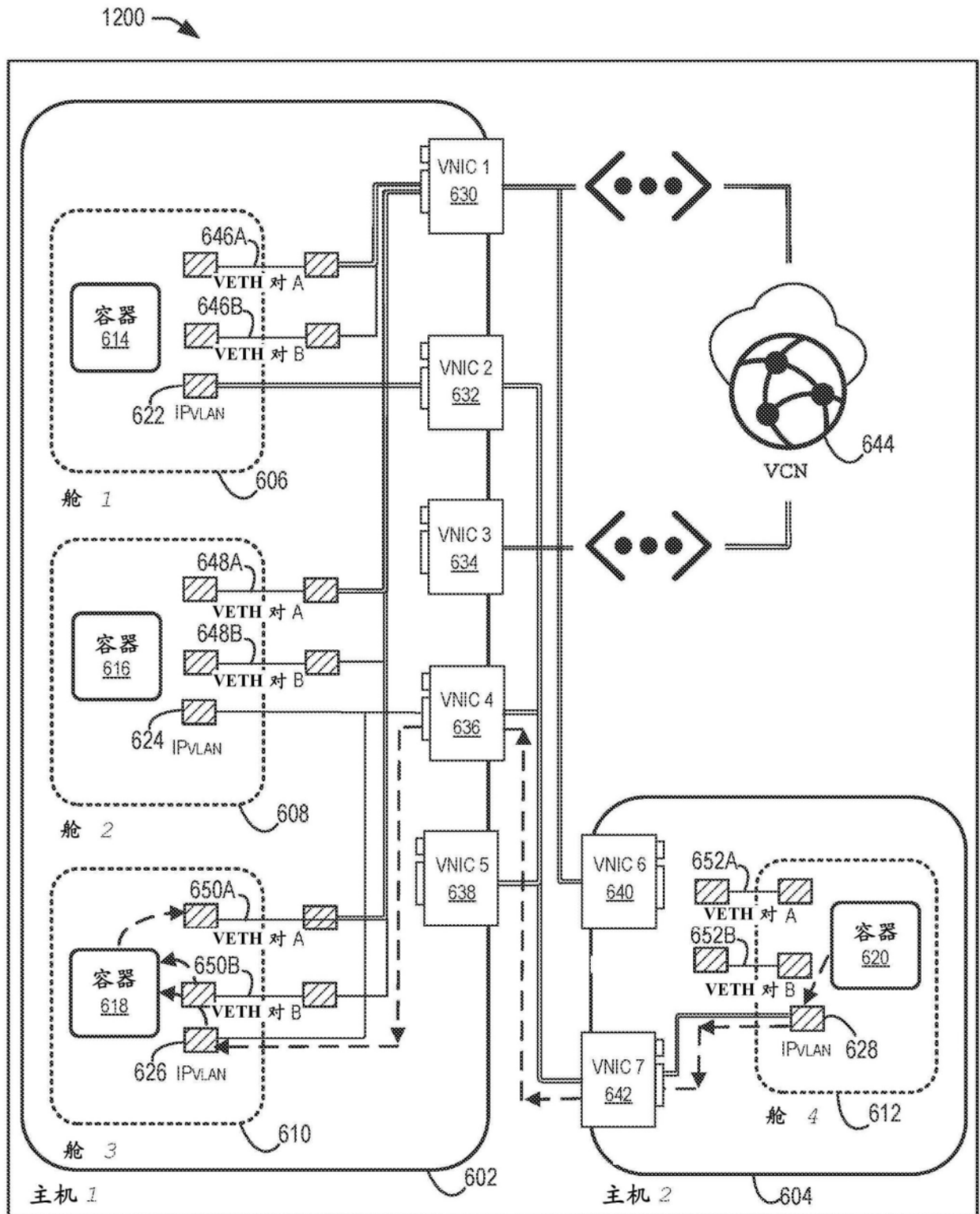


图12

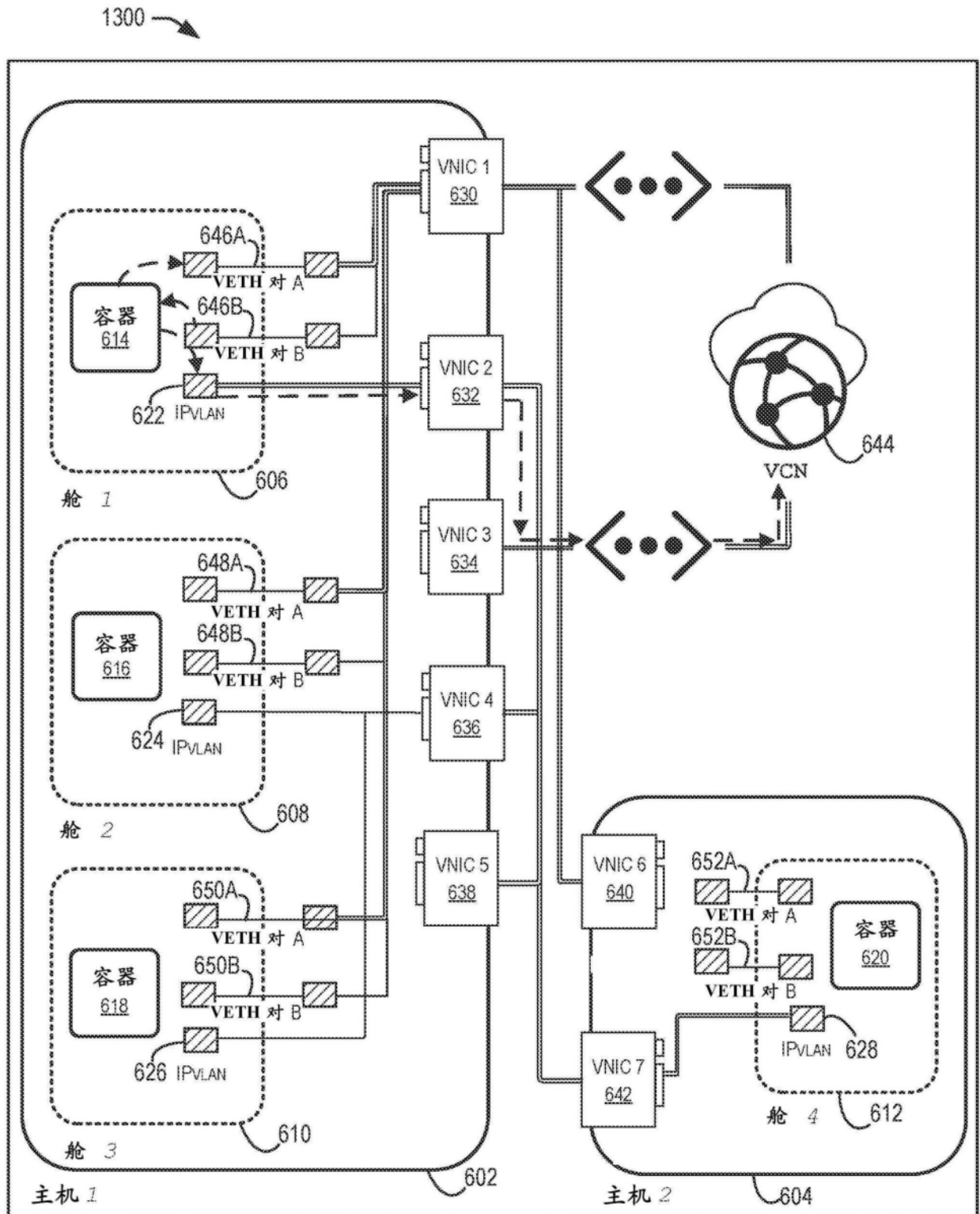


图13

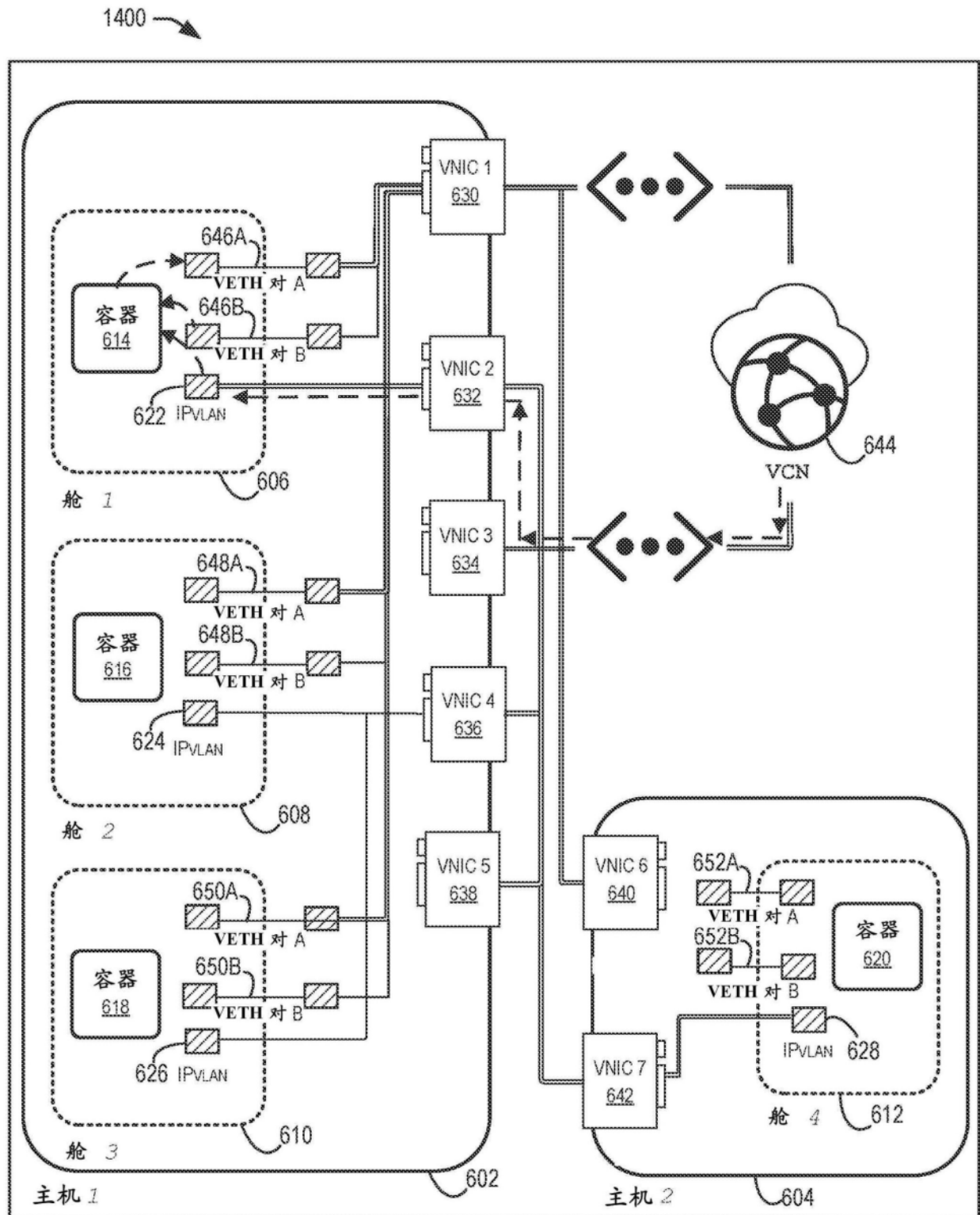


图14

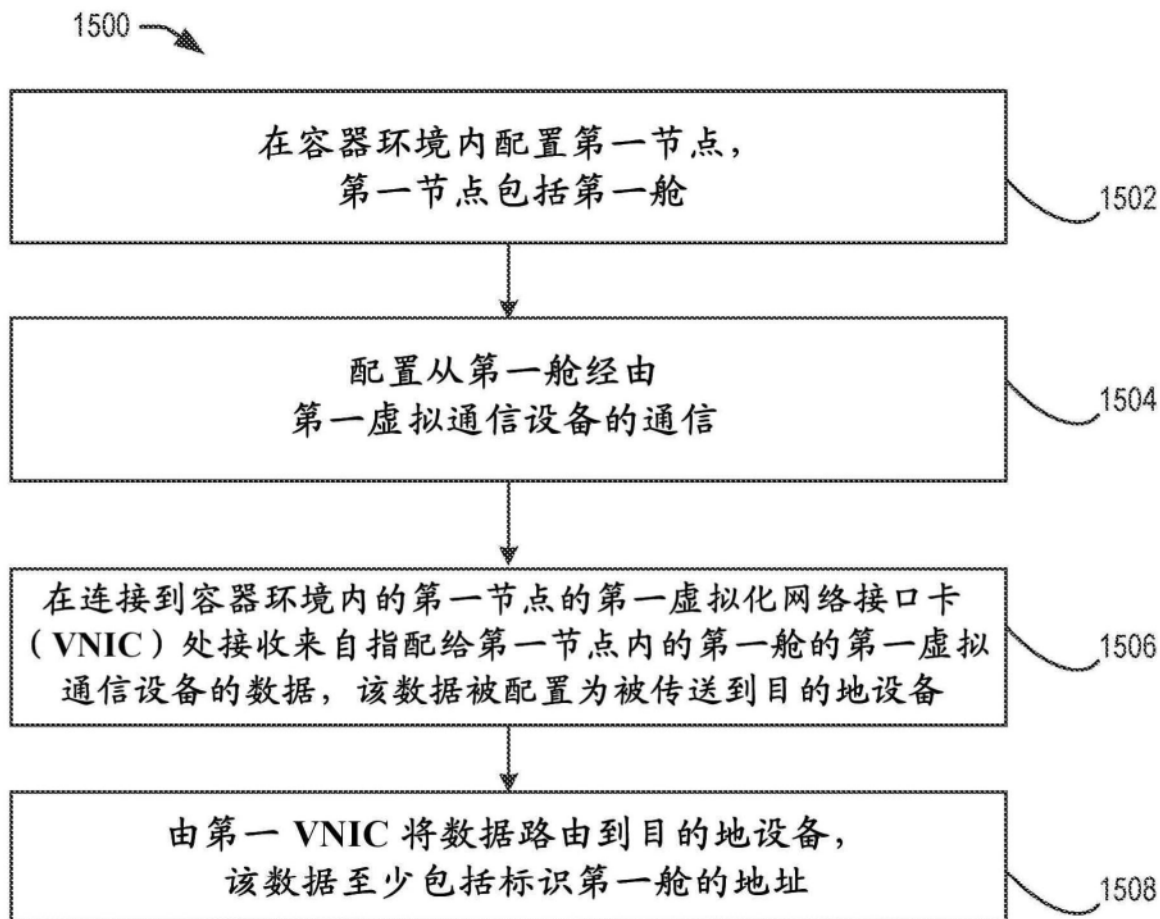


图15

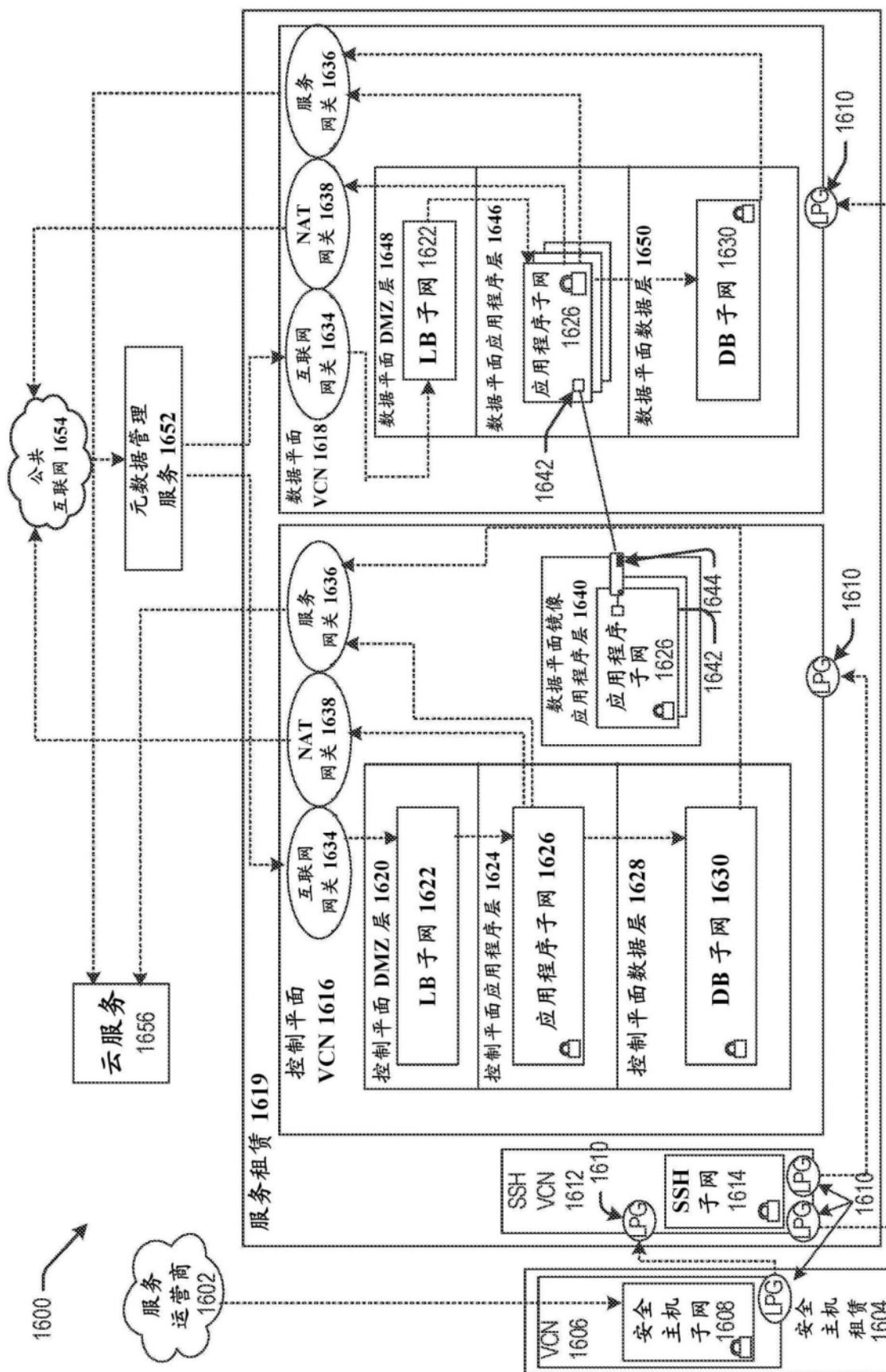


图16

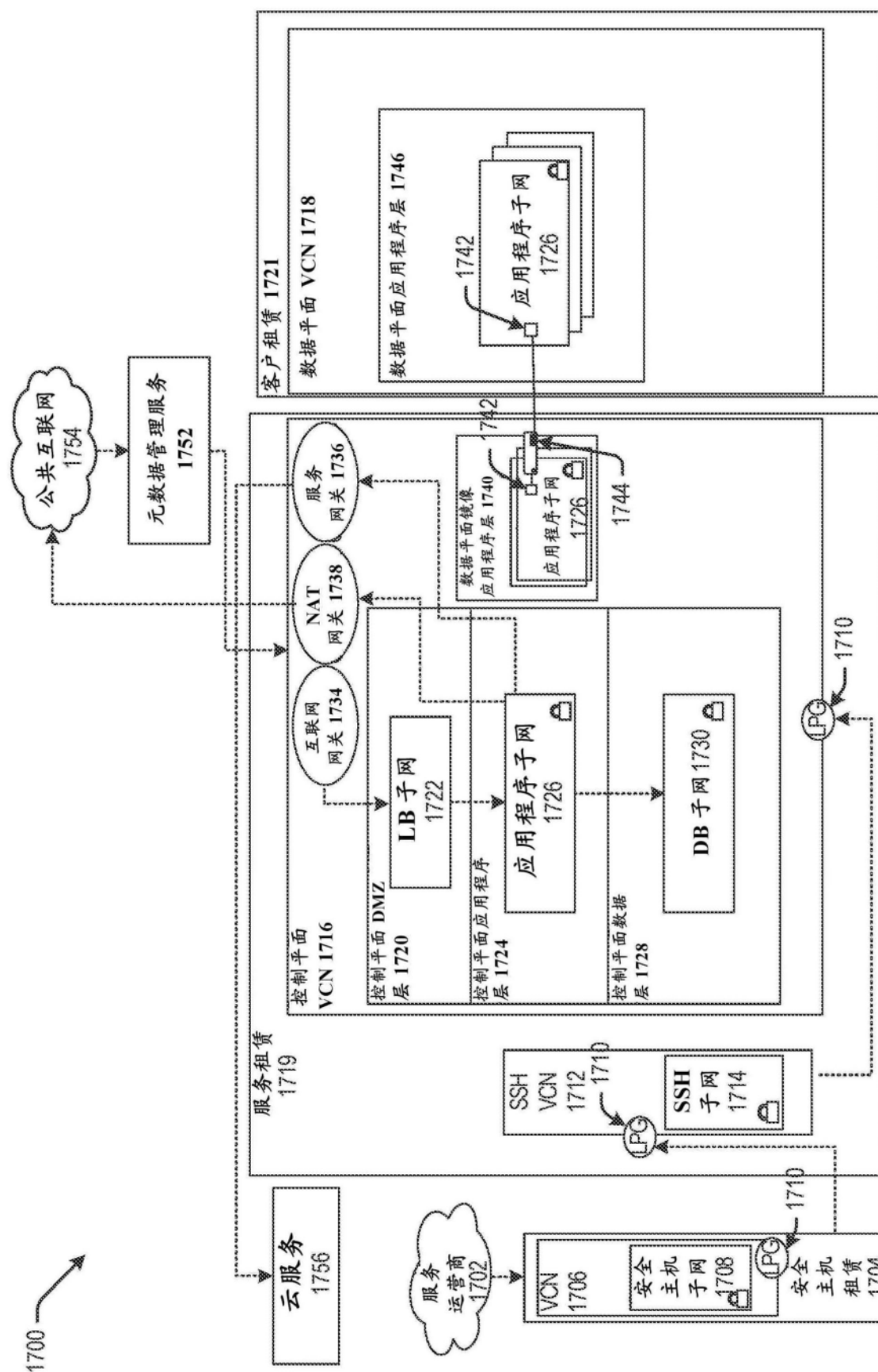


图17

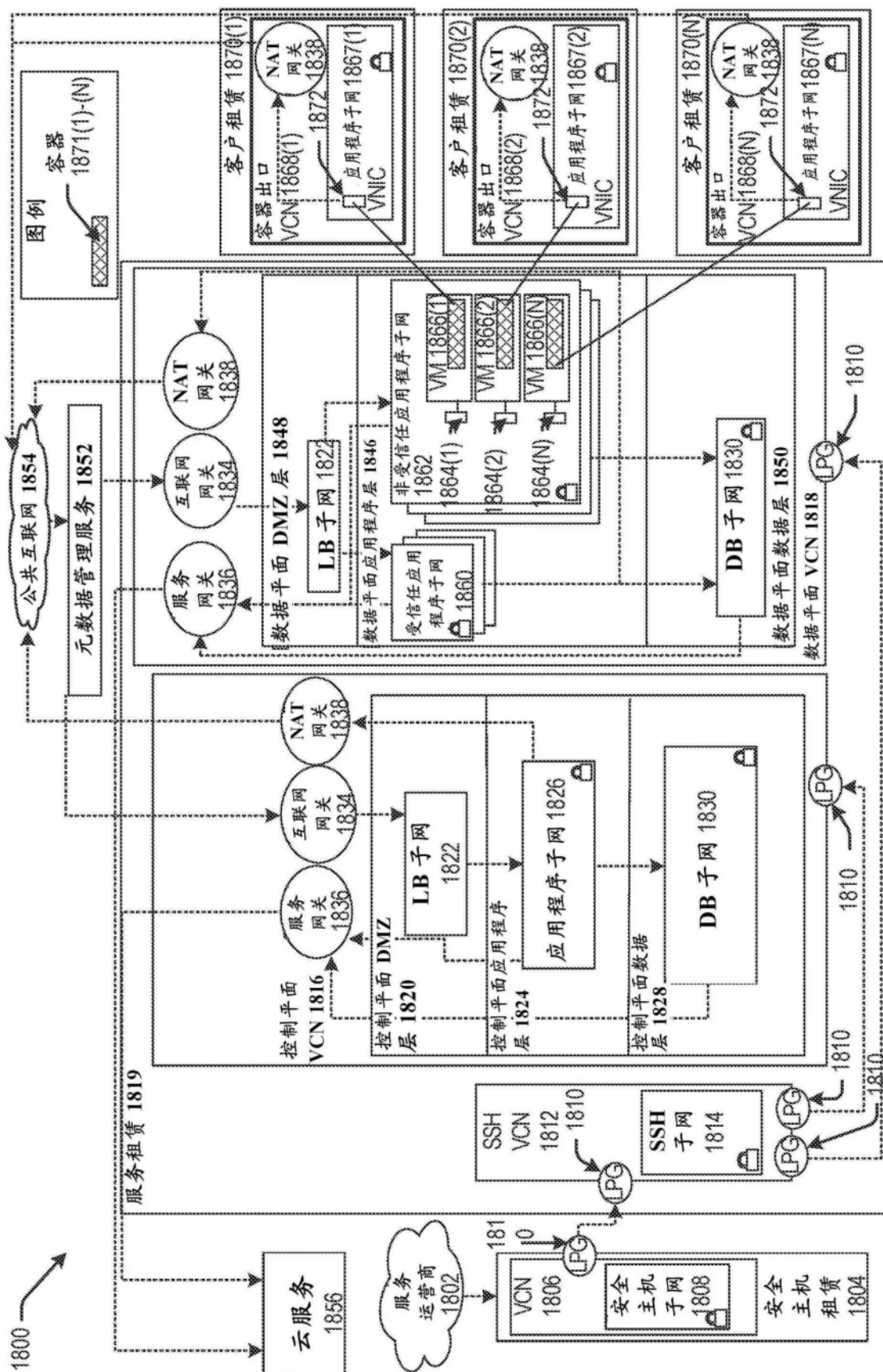


图18

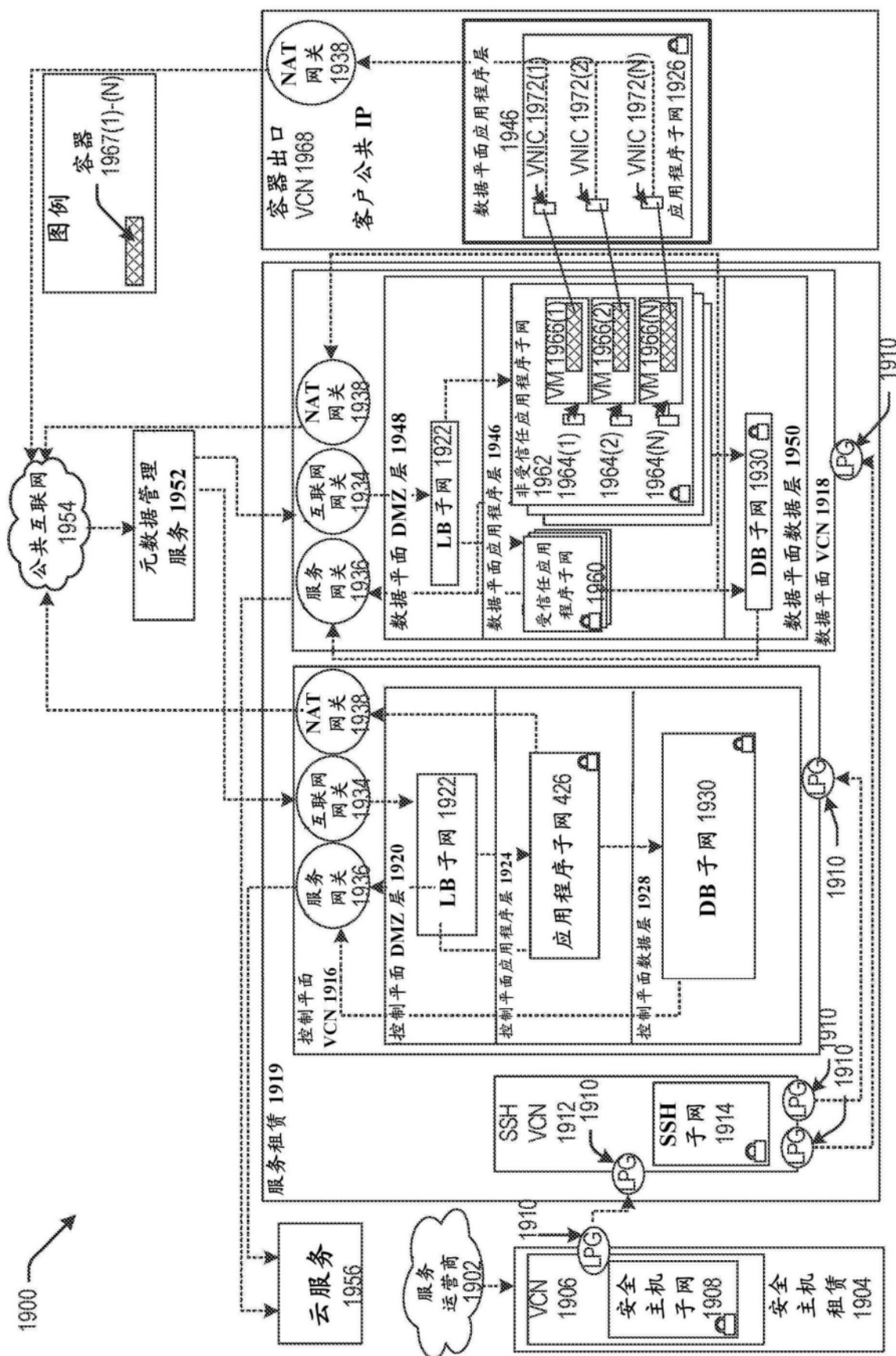


图19

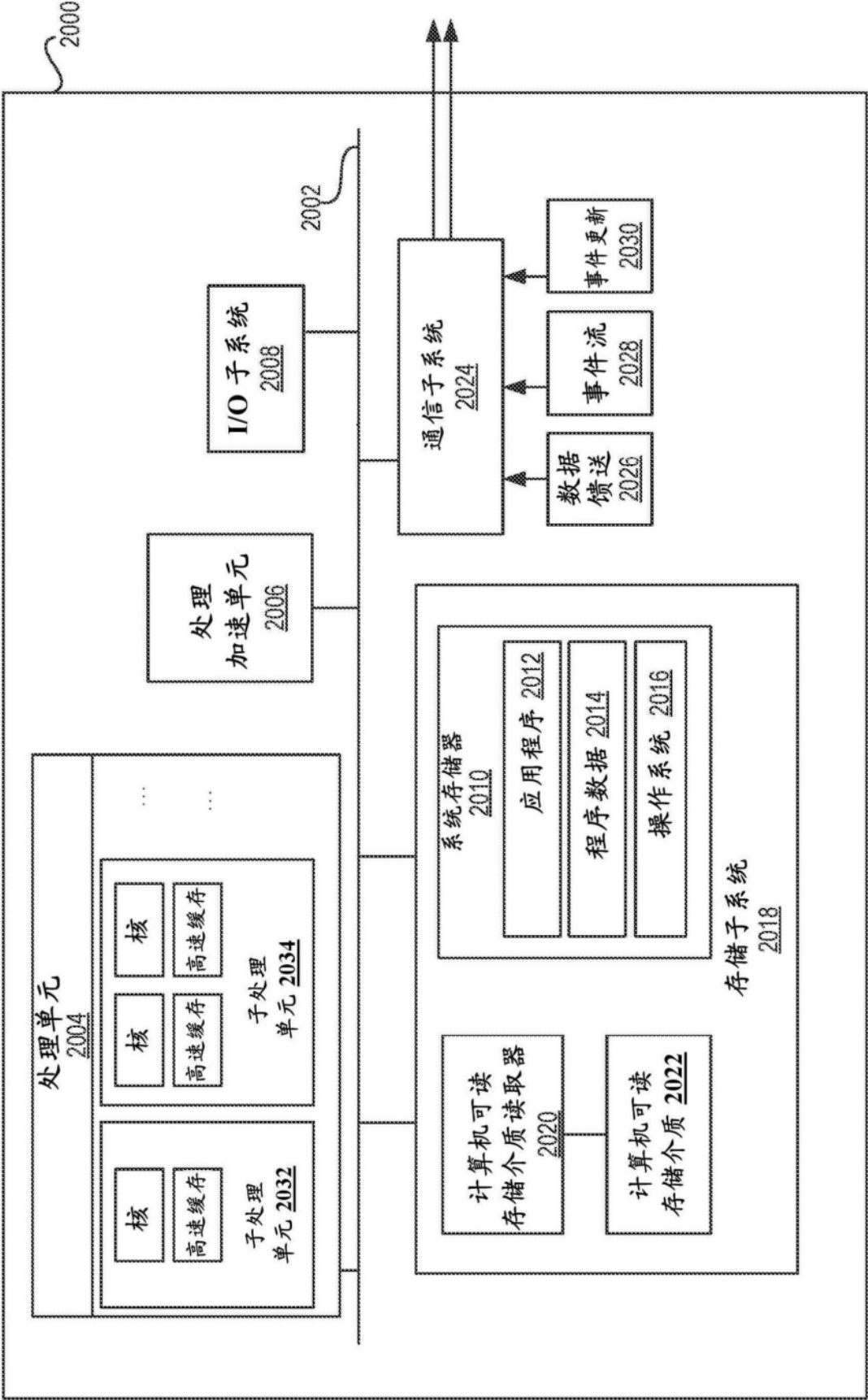


图20