



(12) 发明专利

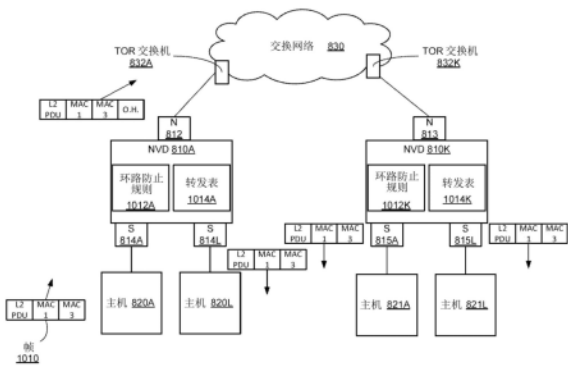
(10) 授权公告号 CN 115699698 B

(45) 授权公告日 2023. 12. 22

(21) 申请号 202180038294.5	(72) 发明人 J · S · 布拉尔 S · 贝克 L · T · 特雷西 S · 切尔维茨 B · E · 博克曼
(22) 申请日 2021.04.14	
(65) 同一申请的已公布的文献号 申请公布号 CN 115699698 A	(74) 专利代理机构 北京市汉坤律师事务所 11602 专利代理师 魏小微 吴丽丽
(43) 申请公布日 2023.02.03	
(30) 优先权数据 63/031,325 2020.05.28 US 17/192,681 2021.03.04 US	(51) Int.Cl. H04L 45/18 (2022.01) H04L 45/76 (2022.01)
(85) PCT国际申请进入国家阶段日 2022.11.25	(56) 对比文件 US 2013301553 A1,2013.11.14 US 6857027 B1,2005.02.15 US 2020044959 A1,2020.02.06 CN 1964360 A,2007.05.16 US 7606177 B1,2009.10.20
(86) PCT国际申请的申请数据 PCT/US2021/027281 2021.04.14	
(87) PCT国际申请的公布数据 W02021/242428 EN 2021.12.02	
(73) 专利权人 甲骨文国际公司 地址 美国加利福尼亚州	审查员 常交法 权利要求书4页 说明书43页 附图19页

(54) 发明名称
虚拟L2网络中的环路防止

(57) 摘要
描述了用于在虚拟L2网络中允许多路径同时防止环路的技术。在示例中,NIC支持虚拟L2网络。NIC被配置成经由NIC的第一端口接收包括源MAC地址和目的地MAC地址的L2帧。基于环路防止规则,NIC经由其除第一端口之外的端口传输L2帧。在附加示例中,NIC进一步被配置成经由NIC的第一端口向主机发送L2帧。L2帧可以是BPDU。在经由第一端口从主机接收到BPDU后,NIC确定BPDU被环回并禁用第一端口。



1. 一种方法,包括:

由网络虚拟化设备经由所述网络虚拟化设备的第一端口接收第一帧,所述第一帧包括作为所述第一帧的目的地的第一计算实例的第一媒体访问控制(MAC)地址、作为所述第一帧的源的第二计算实例的第二MAC地址、以及第2层(L2)协议数据单元(PDU),所述第一计算实例和所述第二计算实例是虚拟L2网络的成员,所述第一计算实例由第一主机托管,所述第一主机经由所述网络虚拟化设备的第二端口与所述网络虚拟化设备连接;

由所述网络虚拟化设备确定环路防止规则防止所述第一帧经由第一组端口的广播,所述第一组端口包括所述第一端口;

由所述网络虚拟化设备确定所述环路防止规则允许所述第一帧经由第二组端口的广播,所述第二组端口不包括所述第一组端口中的任何端口;

由所述网络虚拟化设备确定所述第一帧将基于所述第一MAC地址经由所述网络虚拟化设备的所述第二组端口进行广播;以及

由所述网络虚拟化设备经由所述网络虚拟化设备的所述第二组端口广播所述第一帧。

2. 如权利要求1所述的方法,其中,所述网络虚拟化设备与所述第一主机分离,并经由所述第二端口与所述第一主机的网络接口卡连接。

3. 如权利要求1所述的方法,进一步包括:

由所述网络虚拟化设备经由所述第二端口向所述第一计算实例传输桥接协议数据单元(BPDU);

响应于传输所述BPDU,由所述网络虚拟化设备经由所述第二端口从所述第一计算实例接收所述BPDU;以及

响应于接收所述BPDU,由所述网络虚拟化设备确定所述网络虚拟化设备与所述第一计算实例之间存在环路。

4. 如权利要求3所述的方法,进一步包括:

响应于确定所述网络虚拟化设备与所述第一计算实例之间存在所述环路,由所述网络虚拟化设备禁用所述第二端口;

由所述网络虚拟化设备接收第二帧,所述第二帧包括作为目的地地址的所述第一MAC地址;以及

由所述网络虚拟化设备防止经由所述第二端口传输所述第二帧。

5. 如权利要求1所述的方法,进一步包括:

由所述网络虚拟化设备经由所述第二端口向所述第一计算实例传输桥接协议数据单元(BPDU);

由所述网络虚拟化设备确定没有从所述第一计算实例接收回BPDU;

由所述网络虚拟化设备确定所述网络虚拟化设备与所述第一计算实例之间不存在环路;

由所述网络虚拟化设备接收第二帧,所述第二帧包括作为目的地地址的所述第一MAC地址;以及

由所述网络虚拟化设备基于确定不存在环路而经由所述第二端口向所述第一计算实例传输所述第二帧。

6. 如权利要求1所述的方法,进一步包括:

由所述网络虚拟化设备确定所述第一MAC地址不被包括在所述网络虚拟化设备的转发表中。

7. 如权利要求6所述的方法,进一步包括:

由所述网络虚拟化设备确定所述第二MAC地址不被包括在所述转发表中;

由所述网络虚拟化设备通过至少将所述第二MAC地址与所述第二端口之间的关联包括在所述转发表中来更新所述转发表;

由所述网络虚拟化设备经由所述第二端口接收第二帧,所述第二帧包括作为所述第二帧的目的地址的所述第二MAC地址;以及

由所述网络虚拟化设备基于所述转发表中的所述关联经由所述第一端口而不是经由所述网络虚拟化设备的其他端口传输所述第二帧。

8. 如权利要求1所述的方法,其中,所述方法进一步包括:

由所述网络虚拟化设备接收第二帧,所述第二帧包括作为目的地地址的所述第二MAC地址;以及

由所述网络虚拟化设备确定所述第一MAC地址不被包括在所述网络虚拟化设备的转发表中,其中,所述第二帧被广播到通过交换网络与所述网络虚拟化设备连接的另一网络虚拟化设备。

9. 如权利要求1所述的方法,其中,所述网络虚拟化设备通过交换网络与另一网络虚拟化设备连接,其中,所述第一帧是通过所述交换网络从所述另一网络虚拟化设备接收的,并且其中,所述方法进一步包括:

确定所述第一MAC地址不被包括在所述网络虚拟化设备的转发表中,其中,所述第一帧被广播到在与所述网络虚拟化设备连接的主机上托管的计算实例。

10. 一个或多个非暂态计算机可读存储介质,所述一个或多个非暂态计算机可读存储介质存储有指令,所述指令在由网络虚拟化设备的一个或多个处理器执行时使所述网络虚拟化设备执行包括以下各项的操作:

经由所述网络虚拟化设备的第一端口接收第一帧,所述第一帧包括作为所述第一帧的源的第一计算实例的第一媒体访问控制(MAC)地址、作为所述第一帧的目的地的第二计算实例的第二MAC地址、以及第2层(L2)协议数据单元(PDU),所述第一计算实例和所述第二计算实例是虚拟L2网络的成员,所述第一计算实例由第一主机托管,所述第一主机经由第二端口与所述网络虚拟化设备连接;

确定环路防止规则防止所述第一帧经由第一组端口的广播,所述第一组端口包括所述第一端口;

确定环路防止规则允许所述第一帧经由第二组端口的广播,所述第二组端口不包括所述第一组端口中的任何端口;

确定所述第一帧将基于所述第二MAC地址经由所述网络虚拟化设备的所述第二组端口进行广播;

经由所述网络虚拟化设备的所述第二组端口广播所述第一帧;

经由所述第二端口向所述第一计算实例传输包括桥接协议数据单元(BPDU)的第二帧;以及

基于从所述第一计算实例接收回所述BPDU来确定所述网络虚拟化设备与所述第一计

算实例之间存在环路。

11. 如权利要求10所述的一个或多个非暂态计算机可读存储介质,其中,所述操作进一步包括确定所述第一MAC地址不被包括在所述网络虚拟化设备的转发表中。

12. 如权利要求11所述的一个或多个非暂态计算机可读存储介质,其中,所述操作进一步包括:

确定所述第二MAC地址不被包括在所述转发表中;

通过至少将所述第二MAC地址与所述第一端口之间的关联包括在所述转发表中来更新所述转发表;

经由所述第二端口接收第三帧,所述第三帧包括作为所述第三帧的目的地址的所述第二MAC地址;以及

基于所述转发表中的所述关联经由所述网络虚拟化设备的所述第一端口而不是所述网络虚拟化设备的其他端口传输所述第三帧。

13. 如权利要求10所述的一个或多个非暂态计算机可读存储介质,其中,所述操作进一步包括:

确定所述第一MAC地址不被包括在所述网络虚拟化设备的转发表中,其中,所述第一帧被广播到通过交换网络连接至所述网络虚拟化设备的另一网络虚拟化设备。

14. 如权利要求10所述的一个或多个非暂态计算机可读存储介质,其中,所述操作进一步包括:

确定所述第一MAC地址不被包括在所述网络虚拟化设备的转发表中,其中,所述第一帧被广播到在与所述网络虚拟化设备连接的主机上托管的其他计算实例。

15. 如权利要求10所述的一个或多个非暂态计算机可读存储介质,其中,所述操作进一步包括:

基于所述环路禁用所述第二端口;

接收第三帧,所述第三帧包括作为目的地地址的所述第一MAC地址;以及

防止经由所述第二端口传输所述第三帧。

16. 一种网络虚拟化设备,包括:

多个端口,所述多个端口包括第一端口,以及与托管第一计算实例的第一主机连接的第二端口,所述第一计算实例和第二计算实例是虚拟第2层(L2)网络的成员;

一个或多个处理器;以及

一个或多个存储器,所述一个或多个存储器存储有指令,所述指令在由所述一个或多个处理器执行时将所述网络虚拟化设备配置成:

经由所述第一端口接收第一帧,所述第一帧包括作为所述第一帧的目的地的所述第一计算实例的第一媒体访问控制(MAC)地址、作为所述第一帧的源的所述第二计算实例的第二MAC地址、以及L2协议数据单元(PDU);

确定环路防止规则防止所述第一帧经由第一组端口的广播,所述第一组端口包括所述第一端口;

确定所述环路防止规则允许所述第一帧经由第二组端口的广播,所述第二组端口不包括所述第一组端口中的任何端口;

确定所述第一帧将基于所述第一MAC地址经由所述网络虚拟化设备的所述第二组端口

进行广播;以及

经由所述网络虚拟化设备的所述第二组端口广播所述第一帧。

17. 如权利要求16所述的网络虚拟化设备,其中,执行所述指令进一步将所述网络虚拟化设备配置成:

经由所述第二端口向所述第一计算实例传输桥接协议数据单元(BPDU);

响应于传输所述BPDU,经由所述第二端口从所述第一计算实例接收所述BPDU;以及

响应于接收所述BPDU,确定所述网络虚拟化设备与所述第一计算实例之间存在环路。

18. 如权利要求17所述的网络虚拟化设备,其中,执行所述指令进一步将所述网络虚拟化设备配置成:

响应于确定所述网络虚拟化设备与所述第一计算实例之间存在所述环路,禁用所述第二端口;

接收第二帧,所述第二帧包括作为目的地地址的所述第一MAC地址;以及

防止经由所述第二端口传输所述第二帧。

19. 如权利要求16所述的网络虚拟化设备,其中,执行所述指令进一步将所述网络虚拟化设备配置成:

确定所述第一MAC地址不被包括在所述网络虚拟化设备的转发表中;

确定所述第二MAC地址不被包括在所述转发表中;

通过至少将所述第二MAC地址与所述第二端口之间的关联包括在所述转发表中来更新所述转发表;

经由所述第二端口接收第二帧,所述第二帧包括作为所述第二帧的目的地地址的所述第二MAC地址;以及

基于所述转发表中的所述关联经由所述第一端口而不是经由所述网络虚拟化设备的其他端口传输所述第二帧。

20. 如权利要求16所述的网络虚拟化设备,其中,执行所述指令进一步将所述网络虚拟化设备配置成:

接收第二帧,所述第二帧包括作为目的地地址的所述第二MAC地址;以及

确定所述第一MAC地址不被包括在所述网络虚拟化设备的转发表中,其中,所述第二帧被广播到通过交换网络与所述网络虚拟化设备连接的另一网络虚拟化设备。

虚拟L2网络中的环路防止

[0001] 相关申请的交叉引用

[0002] 本国际专利申请要求于2021年3月4日提交的名称为“LOOP PREVENTION IN VIRTUAL L2 NETWORKS(虚拟L2网络中的环路防止)”的美国专利申请号17/192,681的优先权,所述美国专利申请要求于2020年5月28日提交的名称为“LOOP PREVENTION IN LAYER 2 VIRTUAL NETWORKS WITHOUT GLOBAL SPANNING TREE PROTOCOL(无全局生成树协议的第2层虚拟网络中的环路防止)”的美国临时申请号63/031,325的权益,所述美国临时申请出于所有目的通过引用以其整体并入本文。

背景技术

[0003] 如甲骨文云基础设施(OCI)等云基础设施可以提供一套云服务,这些云服务使得订阅这些服务的实体(例如,企业)能够在高度可用的云托管环境中构建和运行广泛的应用程序和服务。订阅实体被称为云服务提供商的客户。云基础设施可以在灵活的覆盖虚拟网络中提供高性能计算、存储和网络功能,所述灵活的覆盖虚拟网络在物理底层网络之上运行并且可从企业的本地部署网络安全地访问。如OCI等云基础设施通常允许客户以与管理其本地部署工作负载相同的方式来管理其基于云的工作负载。因此,组织可以在具有与其本地部署网络相同的控制力、隔离性、安全性和可预测性能的情况下获得云的所有益处。

[0004] 虚拟联网是云基础设施和云应用的基础,因为虚拟联网实现了访问、连接、保护和修改云资源的能力。虚拟联网实现了多个计算机、虚拟机(VM)、虚拟服务器或跨越不同物理位置的其他设备之间的通信。物理联网通过电缆或其他硬件来连接计算机系统,而虚拟联网使用软件管理通过因特网来连接不同物理位置的计算机和服务器。虚拟网络使用如网络交换机、路由器和适配器等传统网络部件的虚拟化版本,从而允许更高效的路由以及更简单的网络配置和重新配置。

发明内容

[0005] 本公开总体上涉及虚拟联网。更特别地,描述了用于在支持虚拟第2层(L2)网络中的多条路径的同时防止这种网络中的环路的技术。根据某些实施例,可以通过在网络接口卡(NIC)中实行某些规则和/或使用轻量级单端口STP来防止与虚拟L2网络相关联的环路,而不是使用全局生成树协议(STP)。本文中描述了各种发明实施例,包括方法、系统、存储有可由一个或多个处理器执行的程序、代码、指令的非暂态计算机可读存储介质等。

[0006] 根据某些实施例,在支持虚拟L2网络中的多条路径的同时防止环路的方法可以包括:由网络虚拟化设备(NVD)经由所述NVD的第一端口接收包括源媒体访问控制(MAC)地址和目的地MAC地址的L2帧;将所述源MAC地址与所述NVD的第一端口相关联;以及基于防止使用接收所述L2帧所经由的端口进行传输的规则,由所述NVD至少经由所述NVD的第二端口而不是所述NVD的第一端口传输所述L2帧。所述NVD可以提供虚拟NIC(VNIC)的实例。所述虚拟L2网络可以包括L2虚拟LAN(L2 VLAN)。所述第一端口可以连接至包括多个交换机的主机或交换网络,如Clos交换网络。

[0007] 根据某些实施例,在支持虚拟L2网络中的多条路径的同时防止环路的另一种方法可以涉及轻量级单端口STP,所述另一种方法包括:经由所述NVD的第一端口向执行计算实例的主机传输L2帧;经由所述NVD的第一端口从所述主机接收所述L2帧;由所述NVD确定所述L2帧被环回;以及由所述NVD禁用所述NVD的第一端口以停止使用所述第一端口传输和接收帧。

[0008] 根据某些实施例,非暂态计算机可读存储器可以存储可由一个或多个处理器执行的多个指令,所述多个指令包括当由所述一个或多个处理器执行时使所述一个或多个处理器执行以上方法中的任一种的指令。

[0009] 根据某些实施例,系统可以包括一个或多个处理器以及耦接到所述一个或多个处理器的存储器。所述存储器可以存储可由所述一个或多个处理器执行的多个指令,所述多个指令包括当由所述一个或多个处理器执行时使所述一个或多个处理器执行以上方法中的任一种的指令。

[0010] 根据某些实施例,NVD可以被配置成:经由所述NVD的第一端口接收包括源MAC地址和目的地MAC地址的L2帧;将所述源MAC地址与所述NVD的第一端口相关联;并且基于防止使用接收所述L2帧所经由的端口进行传输的规则,至少经由所述NVD的第二端口而不是经由所述NVD的第一端口传输所述L2帧。在一些实施例中,所述NVD可以进一步被配置成经由所述NVD的第一端口向计算实例的主机传输桥接协议数据单元(BPDU),通过所述NVD的第一端口从所述主机接收所述BPDU,确定所述BPDU被环回,并且禁用所述NVD的第一端口以停止使用所述第一端口传输和接收帧。

[0011] 已经采用的术语和表达被用作描述性而非限制性术语,并且在使用这种术语和表达时,不旨在排除所示出和所描述的特征的任何等同物或其部分。然而,应当认识到,在所要求保护的系统和方法的范围内可能有各种修改。因此,应当理解,尽管已经通过示例和可选特征具体地公开了本发明系统和方法,但是本领域技术人员应当认识到本文所公开的概念的修改和变化,并且这种修改和变化被认为在由所附权利要求限定的本发明系统和方法的范围内。

[0012] 本概述不旨在标识所要求保护的主题的关键特征或必要特征,也不旨在独立地用于确定所要求保护的主题的范围。主题应通过参考本公开的整个说明书的适当部分、任何或所有附图以及每个权利要求来理解。

[0013] 在参考以下说明书、权利要求和附图之后,前述内容以及其他特征和实施例将变得更加显而易见。

附图说明

[0014] 下文参考以下附图详细地描述了说明性示例。

[0015] 图1是示出了根据某些实施例的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。

[0016] 图2描绘了根据某些实施例的CSPI内的物理网络中的物理部件的简化架构图。

[0017] 图3示出了根据某些实施例的CSPI内的其中主机连接至多个网络虚拟化设备(NVD)的示例布置。

[0018] 图4描绘了根据某些实施例的主机与NVD之间的用于提供I/O虚拟化以支持多租户

的连通性。

[0019] 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。

[0020] 图6图示了根据某些实施例的包括虚拟L2网络的虚拟云网络 (VCN) 的示例。

[0021] 图7图示了根据某些实施例的包括L2和虚拟L3网络的VCN的示例。

[0022] 图8图示了根据某些实施例的支持VCN的虚拟L2网络的基础设施的示例。

[0023] 图9图示了根据某些实施例的网络虚拟化设备 (NVD) 之间的环路的示例。

[0024] 图10图示了根据某些实施例的防止NVD之间与虚拟L2网络相关联的环路的示例。

[0025] 图11图示了根据某些实施例的NVD与在主机上执行的计算实例之间的环路的示例。

[0026] 图12图示了根据某些实施例的防止NVD与在主机上执行的计算实例之间的环路的示例。

[0027] 图13描绘了用于防止与虚拟L2网络相关联的环路的流程的示例。

[0028] 图14描绘了用于防止NVD之间与虚拟L2网络相关联的环路的流程的示例。

[0029] 图15描绘了用于防止NVD与计算实例之间与虚拟L2网络相关联的环路的流程的示例。

[0030] 图16是图示了根据至少一个实施例的用于实施云基础设施即服务系统的一种模式的框图。

[0031] 图17是图示了根据至少一个实施例的用于实施云基础设施即服务系统的另一模式的框图。

[0032] 图18是图示了根据至少一个实施例的用于实施云基础设施即服务系统的另一模式的框图。

[0033] 图19是图示了根据至少一个实施例的用于实施云基础设施即服务系统的另一模式的框图。

[0034] 图20是图示了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

[0035] 本公开总体上涉及虚拟联网,并且更特别地涉及用于在支持虚拟第2层 (L2) 网络中的多条路径的同时防止环路的技术。根据某些实施例,可以通过实行涉及使用端口来传输帧的某些规则和/或实施每个端口的轻量级生成树协议 (STP) 来防止虚拟L2网络中的环路。本文中描述了各种实施例,包括方法、系统、存储有可由一个或多个处理器执行的程序、代码、指令的非暂态计算机可读存储介质等。

[0036] 如果引入环路,网络可能容易受到广播风暴的影响。网络可能还需要包括多条路径(本文称为“多路径”),以在链路故障的情况下提供冗余路径。L2交换机本身可能不允许多路径并且可能存在环路问题,因为L2控制协议本身可能没有环路防止机制。通常,STP可用于防止L2网络中的环路。然而,STP也防止多路径,而且通常实施起来很复杂。例如,对于STP,每个交换机或网络虚拟化设备 (NVD) 可能需要跨各种端口实施用于根桥选择或根桥优先级的逻辑以及用于根端口 (RP) 选择或RP优先级的逻辑,并且可能需要在环路检测后选择单个路径。

[0037] 根据一些实施例,交换机或NVD可以实施某些规则,以至少避免交换机和/或NVD之

间的环路,同时允许多路径。在一些实施例中,为了避免由软件错误(或者更一般地说,经由端口连接的主机上的计算实例的软件代码)引起的环路,可以实施轻量级STP,其中,轻量级STP可能只需要管理连接至主机的端口,并且可能不需要实施用于根桥选择或根桥优先级的逻辑以及用于RP选择或RP优先级的逻辑。

[0038] 示例虚拟联网架构

[0039] 术语云服务通常用于指云服务提供商(CSP)使用CSP提供的系统和基础设施(云基础设施)按需(例如,经由订阅模式)向用户或客户提供的服务。通常,构成CSP基础设施的服务器和系统与客户自己的本地部署服务器和系统分离。因此,客户可以使自身利用由CSP提供的云服务,而不必针对服务购买单独的硬件和软件资源。云服务旨在为订阅客户提供对应用程序和计算资源的轻松、可扩展的访问,而客户不必投资购买用于提供服务的基础设施。

[0040] 有几家提供各种类型的云服务的云服务提供商。云服务有各种不同的类型或模式,包括软件即服务(SaaS)、平台即服务(PaaS)、基础设施即服务(IaaS)等。

[0041] 客户可以订阅由CSP提供的一种或多种云服务。客户可以是如个体、组织、企业等任何实体。当客户订阅或注册由CSP提供的服务时,就会为该客户创建租约或账户。然后,客户可以经由该账户访问与订阅的该账户相关联的一个或多个云资源。

[0042] 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(称为云服务提供商基础设施或CSPI),客户可以使用所述基础设施来构建其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,在传统计算中,客户的资源和网络由客户提供的基础设施托管。

[0043] CSPI可以包括互连的高性能计算资源,所述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,所述物理网络也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布,所述一个或多个数据中心可以在地理上跨一个或者多个地理区域分布。虚拟化软件可以由这些物理资源执行,以提供虚拟化的分布式环境。虚拟化在物理网络上创建覆盖网络(也称为基于软件的网络、软件定义网络或虚拟网络)。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基础。虚拟或覆盖网络可以包括一个或多个虚拟云网络(VCN)。虚拟网络是使用软件虚拟化技术(例如,管理程序、由网络虚拟化设备(NVD)执行的功能(例如,智能NIC)、架顶式(TOR)交换机、实施由NVD执行的一个或多个功能的智能TOR以及其他机制)来实施的,以创建可以在物理网络之上运行的网络抽象层。虚拟网络可以采取多种形式,包括对等网络、IP网络等。虚拟网络通常是第3层IP网络或第2层VLAN。这种虚拟或覆盖联网方法通常被称为虚拟或覆盖第3层联网。为虚拟网络开发的协议的示例包括IP-in-IP(或通用路由封装(GRE))、虚拟可扩展LAN(VXLAN-IETF RFC 7348)、虚拟私有网络(VPN)(例如,MPLS第3层虚拟私有网络(RFC 4364))、VMware的NSX、GENEVE(通用网络虚拟化封装)等。

[0044] 对于IaaS,由CSP提供的基础设施(CSPI)可以被配置成通过公共网络(例如,因特网)提供虚拟化计算资源。在IaaS模型中,云计算服务提供商可以托管基础设施部件(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)等)。在一些情况下,IaaS提供商还可以提供各种服务来配合这些基础设施部件(例如,计费、监

测、日志记录、安全性、负载平衡和集群等)。因此,由于这些服务可能是策略驱动的,IaaS用户可以能够实施策略来驱动负载平衡,以维护应用程序可用性和性能。CSPI提供基础设施和一套互补的云服务,所述一套互补的云服务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程序和服务。CSPI在灵活的虚拟网络中提供高性能计算资源和功能以及存储容量,所述灵活的虚拟网络可从各种联网位置(如从客户的本地部署网络)安全地访问。当客户订阅或注册由CSP提供的IaaS服务时,为该客户创建的租赁是CSPI内的安全且隔离的分区,客户可以在所述分区中创建、组织和管理其云资源。

[0045] 客户可以使用由CSPI提供的计算资源、内存资源和联网资源来构建其自己的虚拟网络。可以在这些虚拟网络上部署一种或多种客户资源或工作负载(如计算实例)。例如,客户可以使用由CSPI提供的资源来构建一个或多个可定制的私有虚拟网络,称为虚拟云网络(VCN)。客户可以在客户VCN上部署如计算实例等一种或多种客户资源。计算实例可以采取虚拟机、裸机实例等形式。因此,CSPI提供基础设施和一套互补的云服务,所述一套互补的云服务使得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程序和服务。客户不管理或控制由CSPI提供的底层物理资源,但对操作系统、存储装置和部署的应用程序具有控制权;并且可能对选定联网部件(例如防火墙)具有有限的控制权。

[0046] 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的基于网络的用户接口。在一些实施方式中,控制台是由CSP提供的基于网络的应用程序。

[0047] CSPI可以支持单租户或多租户架构。在单租户架构中,软件(例如,应用程序、数据库)或硬件部件(例如,主机或服务器)服务于单个客户或租户。在多租户架构中,软件或硬件部件服务于多个客户或租户。因此,在多租户架构中,CSPI资源在多个客户或租户之间共享。在多租户情况下,CSPI内会采取预防措施并施行安全措施,以确保每个租户的数据都是隔离的并且保持对其他租户不可见。

[0048] 在物理网络中,网络端点(“端点”)是指连接至物理网络并与其所连接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至局域网(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、网桥、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都有可以用于与所述设备通信的固定网络地址。该固定网络地址可以是第2层地址(例如,MAC地址)、固定的第3层地址(例如,IP地址)等。在虚拟化环境或虚拟网络中,端点可以包括各种虚拟端点,如由物理网络的部件托管(例如,由物理主机托管)的虚拟机。虚拟网络中的这些端点由覆盖地址寻址,如覆盖第2层地址(例如,覆盖MAC地址)和覆盖第3层地址(例如,覆盖IP地址)。网络覆盖通过允许网络管理器使用软件管理(例如,经由实施虚拟网络的控制平面的软件)在与网络端点相关联的覆盖地址周围移动来实现灵活性。因此,与物理网络不同,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从一个端点移动到另一个端点。由于虚拟网络在物理网络之上构建,因此虚拟网络中的部件之间的通信涉及虚拟网络和底层物理网络两者。为了便于这种通信,CSPI的部件被配置成学习和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。然后使用这些映射来促进通信。客户流量被封装以便于在虚拟网络中路由。

[0049] 因此,物理地址(例如,物理IP地址)与物理网络中的部件相关联,而覆盖地址(例

如,覆盖IP地址)与虚拟网络中的实体相关联。物理IP地址和覆盖IP地址两者都是真实IP地址的类型。它们与虚拟IP地址不同,其中,虚拟IP地址映射到多个真实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。

[0050] 云基础设施或CSPI在物理上托管在全球范围的一个或多个地区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部件以及在物理网络部件之上构建的虚拟网络中的虚拟化部件(例如,虚拟网络、计算实例、虚拟机等)。在某些实施例中,CSPI是在领域、区域和可用性域中组织和托管的。区域通常是包含一个或多个数据中心的局部地理区域。区域通常是彼此独立的并且可能相隔很远,例如,跨越国家或甚至大陆。例如,第一个区域可能在澳大利亚,另一个在日本,还有一个在印度,等等。CSPI资源在各区域中进行划分,使得每个区域具有其自己的独立的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,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、档案存储);联网资源(例如,虚拟云网络(VCN)、负载均衡资源、到本地部署网络的连接)、数据库资源;边缘联网资源(例如,DNS);以及访问管理和监测资源等。每个区域通常有多条路径连接至领域中的其他区域。

[0051] 通常,应用程序被部署在其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。应用程序也可以出于各种原因被部署在不同的区域中,如冗余部署以减轻区域性事件(如大型天气系统或地震)的风险,从而满足法律管辖范围、税务领域和其他商业或社会标准等的不同要求。

[0052] 区域内的数据中心可以进一步被组织和细分为可用性域(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构成。在这种分布式环境中,CSPI资源要么是特定于区域的,如虚拟云网络(VCN),要么是特定于可用性域的,如计算实例。

[0053] 区域内的AD是彼此隔离的、具有容错性,并且被配置成使其不太可能同时发生故障。这是通过AD不共享如联网、物理电缆、电缆路径、电缆入口点等关键基础设施资源使得区域内一个AD的故障不太可能影响同一区域内其他AD的可用性来实现的。同一区域内的AD可以通过低延迟、高带宽的网络彼此连接,这使得能够为其他网络(例如,因特网、客户的内部部署网络等)提供高可用性连通性并在多个AD中构建复制系统以实现高可用性和灾难恢复。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,更多的区域和AD可能会增加附加的容量。可用性域之间的流量通常是加密的。

[0054] 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。领域是彼此隔离的,并且不共享任何数据。同一领域中的区域可以彼此通信,但不同领域中的区域无法通信。客户与CSP的租约或账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布。通常,当客户订阅IaaS服务时,会在领域内客户指定的区域(称为“主”区域)中为该客户创建租约或账户。客户可以跨领域内的一个或多个其他区域扩展客户的租约。客户无法访问不在客户的租约所在的领域中的区域。

[0055] IaaS提供商可以提供多个领域,每个领域为一组特定的客户或用户服务。例如,可以为商业客户提供商业领域。作为另一个示例,可以为特定国家内的客户提供该国家的领域。作为又一个示例,可以为政府提供政府领域,等等。例如,政府领域可以为特定政府服务,并且可以比商业领域具有更高的安全级别。例如,甲骨文云基础设施(OCI)目前为商业

区域提供一个领域,而为政府云区域提供两个领域(例如,经FedRAMP授权和IL5授权)。

[0056] 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内提供反亲和性的一组基础设施资源。故障域允许计算实例分散,使得这些实例不在单个AD内的同一物理硬件上。这称为反亲和力。故障域是指共享单点故障的一组硬件部件(计算机、交换机等)。计算池在逻辑上被划分为多个故障域。因此,影响一个故障域的硬件故障或计算硬件维护事件不会影响其他故障域中的实例。根据实施例,每个AD的故障域的数量可能有所不同。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。

[0057] 当客户订阅IaaS服务时,将为客户供应来自CSPI的资源并将其与客户的租约相关联。客户可以使用这些供应资源来构建私有网络并在这些网络上部署资源。由CSPI托管在云中的客户网络被称为虚拟云网络(VCN)。客户可以使用为其分配的CSPI资源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,如应用程序、负载均衡器、数据库等。部署在VCN上的计算实例可以通过如因特网等所公共网络与公共可访问端点(“公共端点”)通信,与同一VCN或其他VCN中的其他实例(例如,客户的其他VCN或不属于客户的VCN)通信,与客户的本地部署数据中心或网络通信,以及与服务端点和其他类型的端点通信。

[0058] CSP可以使用CSPI提供各种服务。在一些情况下,CSPI的客户可能自己充当服务提供商并使用CSPI资源提供服务。服务提供商可以公开用标识信息(例如,IP地址、DNS名称和端口)表征的服务端点。客户的资源(例如,计算实例)可以通过访问服务为特定服务公开的服务端点来使用该特定服务。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地址经由如因特网等公共通信网络公开访问的端点。可公开访问的网络端点有时也称为公共端点。

[0059] 在某些实施例中,服务提供商可以经由服务的端点(有时称为服务端点)来公开服务。然后,服务的客户可以使用该服务端点来访问服务。在某些实施方式中,为服务提供的服务端点可以由打算使用该服务的多个客户访问。在其他实施方式中,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。

[0060] 在某些实施例中,当创建VCN时,其与私有覆盖无类域间路由(CIDR)地址空间相关联,所述地址空间是指派给VCN的私有覆盖IP地址的范围(例如,10.0/16)。VCN包括相关子网、路由表和网关。VCN驻留在单个区域内,但可以跨越该区域的一个或多个或所有可用性域。网关是为VCN配置的虚拟接口,并且实现了往返VCN的流量与VCN外部的一个或多个端点通信。可以为VCN配置一个或多个不同类型的网关,以实现往返不同类型的端点的通信。

[0061] VCN可以被细分为一个或多个子网络,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子部分。VCN可以有一个或多个子网。VCN内的每个子网都与连续覆盖IP地址范围(例如,10.0.0.0/24和10.0.1.0/24)相关联,所述覆盖IP地址范围不与该VCN中的其他子网重叠并且表示VCN的地址空间内的地址空间子集。

[0062] 每个计算实例都与使得该计算实例能够加入VCN的子网的虚拟网络接口卡(VNIC)相关联。VNIC是物理网络接口卡(NIC)的逻辑表示。总的来说,VNIC是实体(例如,计算实例、服务)与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IP地址以及相关安全规则或策略。VNIC相当于交换机上的第2层端口。VNIC附接到计算实例和VCN内的子网。与

计算实例相关联的VNIC使得所述计算实例能够成为VCN的子网的一部分,并且使得所述计算实例能够与和所述计算实例处于同一子网上的端点、与VCN中不同子网中的端点或与VCN外部的端点通信(例如,发送和接收数据包)。因此,与计算实例相关联的VNIC确定了计算实例如何与VCN内外的端点连接。当创建计算实例并将其添加到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,所述子网包含与所述一组计算实例相对应的VNIC,每个VNIC附接到所述一组计算机实例内的计算实例。

[0063] 经由与计算实例相关联的VNIC为每个计算实例指派私有覆盖IP地址。当创建计算实例时,该私有覆盖IP地址被指派给与所述计算实例相关联的VNIC,并用于路由往返所述计算实例的流量。给定子网中的所有VNIC都使用相同的路由表、安全列表和DHCP选项。如上所述,VCN内的每个子网都与连续覆盖IP地址范围(例如,10.0.0.0/24和10.0.1.0/24)相关联,所述覆盖IP地址范围不与该VCN中的其他子网重叠并且表示VCN的地址空间内的地址空间子集。对于VCN的特定子网上的VNIC,指派给VNIC的私有覆盖IP地址是来自分配给所述子网的连续覆盖IP地址范围的地址。

[0064] 在某些实施例中,除私有覆盖IP地址之外,可以可选地为计算实例指派附加的覆盖IP地址,例如,在公共子网的情况下为其指派一个或多个公共IP地址。这些多个地址在同一VNIC上被指派,或者在与计算实例相关联的多个VNIC上被指派。然而,每个实例都有在实例启动期间创建并与指派给所述实例的覆盖私有IP地址相关联的主VNIC——该主VNIC不能移除。附加的VNIC(称为次级VNIC)可以被添加到与主VNIC处于同一可用性域中的现有实例中。所有VNIC都与实例处于同一可用性域中。次级VNIC可以与主VNIC处于同一VCN中的子网中,或者处于同一VCN或不同VCN中的不同子网中。

[0065] 如果计算实例处于公共子网中,则可以可选地为该实例指派公共IP地址。在创建子网时,可以将子网指定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VNIC不能有公共覆盖IP地址。公共子网意味着子网中的资源和相关VNIC可以有公共IP地址。客户可以指定子网存在于单个可用性域中,或者存在于区域或领域中的多个可用性域中。

[0066] 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)(称为VCN VR或仅VR)使得VCN的子网之间能够通信。对于VCN内的子网,VR表示该子网的逻辑网关,所述逻辑网关使得子网(即,该子网上的计算实例)能够与VCN内的其他子网上的端点以及VCN外部的其他端点通信。VCN VR是被配置成在VCN中的VNIC与和VCN相关联的虚拟网关(“网关”)之间路由流量的逻辑实体。下文参考图1进一步描述了网关。VCN VR是第3层/IP层概念。在一个实施例中,VCN有一个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,VCN的每个子网都有一个端口。以这种方式,VCN VR对于VCN VR所附接的VCN中的每个子网具有不同的IP地址。VR还连接至为VCN配置的各种网关。在某些实施例中,为子网的VCN VR的端口保留该子网的覆盖IP地址范围中的特定覆盖IP地址。例如,假设VCN有两个子网,这两个子网的相关地址范围分别为10.0/16和10.1/16。对于VCN内地址范围为10.0/16的第一子网,为该子网的VCN VR的端口保留该范围内的地址。在一些情况下,可以为VCN VR保留该范围内的第一个IP地址。例如,对于覆盖IP地址范围为10.0/16的子网,可以为该子网的VCN VR的端口保留IP地址10.0.0.1。对于同一VCN内地址范围为10.1/16的第二子网,VCN VR可以具有用于IP地址为10.1.0.1的该第二子网的端口。VCN VR

对于VCN中的每个子网有不同的IP地址。

[0067] 在一些其他实施例中,VCN内的每个子网可以具有其自己的相关VR,所述相关VR可由子网使用与VR相关联的保留或默认的IP地址寻址。例如,保留或默认的IP地址可以是与该子网相关联的IP地址范围中的第一个IP地址。子网中的VNIC可以使用该默认或保留的IP地址与和子网相关联的VR通信(例如,发送和接收数据包)。在这种实施例中,VR是该子网的入口/出口点。与VCN内的子网相关联的VR可以与和VCN内的其他子网相关联的其他VR通信。VR还可以与和VCN相关联的网关通信。子网的VR功能在为子网中的VNIC执行VNIC功能的一个或多个NVD上运行或由其执行。

[0068] 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特殊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。可以定制VCN的路由表以控制如何向VCN转发/路由数据包和从VCN转发/路由数据包。DHCP选项是指在实例启动时自动提供给这些实例的配置信息。

[0069] 为VCN配置的安全规则表示用于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)。客户可以选择给定的规则是有状态的还是无状态的。例如,客户可以通过设置源CIDR0.0.0.0/0和目的地TCP端口22的有状态入口规则以允许SSH流量从任何地方进入到一组实例。可以使用网络安全组或安全列表来实施安全规则。网络安全组由一组仅适用于该组中的资源的安全规则构成。另一方面,安全列表包括适用于使用安全列表的任何子网中的所有资源的规则。可以向VCN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。

[0070] 在某些实施例中,VCN的配置信息由VCN控制平面确定和存储。VCN的配置信息可以包括例如关于以下的信息: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分发服务向NVD发布由VCN控制平面或其部分存储的配置信息。所分发的信息可以用于更新由NVD存储和使用信息(例如,转发表、路由表等),以向VCN中的计算实例转发数据包和从所述计算实例转发数据包。

[0071] 在某些实施例中,VCN和子网的创建由VCN控制平面(CP)处理,并且计算实例的启动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控制平面来创建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成执行数据包转发和路由功能的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发服务。VCN控制平面的示例还在图16、图17、图18和图19中描绘(参见附图标记1616、1716、1816和1916)并在下文中描述。

[0072] 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。这些端点可以包括由CSPI托管的端点和CSPI外部的端点。

[0073] 用于使用CSPI来实施基于云的服务的各种不同的架构在图1、图2、图3、图4、图5、图16、图17、图18和图20中描绘并且在下文中描述。图1是示出了根据某些实施例的由CSPI托管的覆盖或客户VCN的分布式环境100的高级图。图1所描绘的分布式环境包括覆盖网络中的多个部件。图1所描绘的分布式环境100仅仅是示例,并不旨在不适当地限制所要求保

护的实施例的范围。可能有许多变化、替代和修改。例如,在一些实施方式中,图1所描绘的分布式环境可以具有比图1所示的那些系统或部件更多或更少的系统或部件,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。

[0074] 如图1所描绘的示例所示,分布式环境100包括CSPI 101,其提供客户可以订阅并用来构建其虚拟云网络 (VCN) 的服务和资源。在某些实施例中,CSPI 101向订阅客户提供IaaS服务。CSPI 101内的数据中心可以被组织成一个或多个区域。图1中示出了一个示例区域“美国地区”102。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程序、数据库、负载均衡器等。

[0075] 在图1所描绘的实施例中,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,而子网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,所述逻辑网关实现了VCN 104的子网之间以及与VCN外部的其他端点的通信。VCN VR105被配置成在VCN 104中的VNIC与和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IP地址为10.0.0.1的端口,并为子网2提供IP地址为10.1.0.1的端口。

[0076] 可以在每个子网上部署多个计算实例,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。计算实例经由与所述计算实例相关联的VNIC加入子网。例如,如图1所示,计算实例C1经由与所述计算实例相关联的VNIC成为子网1的一部分。同样地,计算实例C2经由与C2相关联的VNIC成为子网1的一部分。以类似的方式,可以作为虚拟机实例或裸机实例的多个计算实例可以是子网1的一部分。经由其相关VNIC,为每个计算实例指派私有覆盖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和MAC地址M1,而计算实例C2具有私有覆盖IP地址10.0.0.3和MAC地址M2。子网1中的每个计算实例(包括计算实例C1和C2) 都有使用IP地址10.0.0.1(其是子网1的VCN VR 105的端口的IP地址)到VCN VR 105的默认路由。

[0077] 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所示,计算实例D1和D2经由与相应计算实例相关联的VNIC成为子网2的一部分。在图1所描绘的实施例中,计算实例D1具有覆盖IP地址10.1.0.2和MAC地址MM1,而计算实例D2具有私有覆盖IP地址10.1.0.3和MAC地址MM2。子网2中的每个计算实例(包括计算实例D1和D2) 都有使用IP地址10.1.0.1(其是子网2的VCN VR 105的端口的IP地址)到VCN VR 105的默认路由。

[0078] VCN A 104还可以包括一个或多个负载均衡器。例如,可以为子网提供负载均衡器,并且所述负载均衡器可以被配置成跨子网上的多个计算实例负载均衡流量。还可以提供用于跨VCN中的子网负载均衡流量的负载均衡器。

[0079] 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:与特定计算实例处于同一子网上的端点(例如,子网1中的两个计算实例之间的通信);在不同子网上但在同一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);同一区域中的不同VCN中的端点(例如,子网1中的计算实例与同一区域106或110中的VCN中

的端点之间的通信,子网1中的计算实例与同一区域中的服务网络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108中的VCN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中的端点、可经由如因特网等公共网络访问的公共端点114、以及其他端点。

[0080] 使用与源计算实例和目的地计算实例相关联的VNIC来促进同一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可能想要向子网1中的计算实例C2发送数据包。对于源自源计算实例并且其目的地是同一子网中的另一个计算实例的数据包,所述数据包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:根据数据包报头确定数据包的目的地信息,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),确定数据包的下一跳地址,根据需要执行任何数据包封装/解封装功能,并且然后将数据包转发/路由到下一跳地址以便于将数据包传送到其预期目的地。当目的地计算实例与源计算实例处于同一子网中时,与源计算实例相关联的VNIC被配置成识别与目的地计算实例相关联的VNIC并将数据包转发到该VNIC进行处理。然后,执行与目的地计算实例相关联的VNIC,并将数据包转发到目的地计算实例。

[0081] 对于要从子网中的计算实例传送到同一VCN中不同子网中的端点的数据包,由与源计算实例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中的子网1中的计算实例C1想要向子网2中的计算实例D1发送数据包,则所述数据包首先由与计算实例C1相关联的VNIC处理。与计算实例C1相关联的VNIC被配置成使用VCN VR的默认路由或端口10.0.0.1将数据包路由到VCN VR 105。VCN VR 105被配置成使用端口10.1.0.1将数据包路由到子网2。然后,由与D1相关联的VNIC接收并处理数据包,并且VNIC将数据包转发到计算实例D1。

[0082] 对于要从VCN 104中的计算实例传送到VCN 104外部的端点的数据包,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。一种或多种类型的网关可以与VCN 104相关联。网关是VCN与另一个端点之间的接口,其中,另一个端点在VCN外部。网关是第3层/IP层概念,并且使得VCN能够与VCN外部的端点通信。因此,网关促进了VCN与其他VCN或网络之间的业务流。可以为VCN配置各种不同类型的网关,以便于与不同类型的端点进行不同类型的通信。根据网关,通信可以通过公共网络(例如,因特网)或通过私有网络进行。各种通信协议可以用于这些通信。

[0083] 例如,计算实例C1可能想要与VCN 104外部的端点通信。数据包可以首先由与源计算实例C1相关联的VNIC处理。VNIC处理确定数据包的目的地在C1的子网1外部。与C1相关联的VNIC可以将数据包转发到VCN 104的VCN VR 105。然后,VCN VR 105对数据包进行处理,并且作为处理的一部分,基于数据包的目的地来确定与VCN 104相关联的特定网关作为数据包的下一跳地址。然后,VCN VR 105可以将数据包转发到特定的经识别网关。例如,如果目的地是客户的本地部署网络内的端点,则数据包可以由VCN VR 105转发到为VCN 104配置的动态路由网关(DRG)网关122。然后,可以将数据包从网关转发到下一跳地址,以便于将数据包传送到其最终预期目的地。

[0084] 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中描绘

并在下文中描述。与VCN相关联的网关的示例还在图16、图17、图18和图19中描绘(例如,由附图标记1634、1636、1638、1734、1736、1738、1834、1836、1838、1934、1936和1938引用的网关)并在下文中描述。如图1所描绘的实施例所示,动态路由网关(DRG)122可以被添加到客户VCN 104或者与其相关联,并为客户VCN 104与另一个端点之间的私有网络流量通信提供路径,其中,另一个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或者未由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的访问通常非常有限。对于既有客户本地部署网络116又有由CSPI 101部署或托管在云中的一个或多个VCN 104的客户,客户可能希望其本地部署网络116与其基于云的VCN 104能够彼此通信。这使得客户能够构建涵盖由CSPI 101托管的客户的VCN 104和其本地部署网络116的扩展混合环境。DRG 122实现了这种通信。为了实现这种通信,建立通信信道124,其中,信道的一个端点在客户本地部署网络116中,而另一个端点在CSPI 101中并连接至客户VCN 104。通信信道124可以在如因特网等公共通信网络或私有通信网络上。可以使用各种不同的通信协议,如公共通信网络(如因特网)上的IPsec VPN技术、使用私有网络而不是公共网络的甲骨文公司FastConnect技术等。形成通信信道124的一个端点的客户本地部署网络116中的设备或装置被称为客户本地部署装置(CPE),如图1所描绘的CPE 126。在CSPI 101一侧,端点可以是执行DRG 122的主机。

[0085] 在某些实施例中,远程对等连接(RPC)可以被添加到DRG,这允许客户将一个VCN与不同区域中的另一个VCN对等。使用这种RPC,客户VCN 104可以使用DRG 122与另一个区域中的VCN 108连接。DRG 122还可以用于与未由CSPI 101托管的其他远程云网络118(如微软Azure云、亚马逊AWS云等)通信。

[0086] 如图1所示,可以为客户VCN 104配置因特网网关(IGW)120,所述因特网网关使得VCN 104上的计算实例能够与可通过如因特网等公共网络访问的公共端点114通信。IGW 120是将VCN连接至如因特网等公共网络的网关。IGW 120使得如VCN 104等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)能够直接访问如因特网等公共网络114上的公共端点112。使用IGW 120,可以从VCN 104内的子网或从因特网启动连接。

[0087] 网络地址转换(NAT)网关128可以被配置用于客户的VCN 104,并使得客户的VCN中不具有专用公共覆盖IP地址的云资源能够访问因特网,而且这样做不会将那些资源暴露于直接进入的因特网连接(例如,L4-L7连接)。这使得VCN内的私有子网(如VCN 104中的私有子网1)能够对因特网上的公共端点进行私有访问。在NAT网关中,只能启动从私有子网到公共因特网的连接,而不能启动从因特网到私有子网的连接。

[0088] 在某些实施例中,服务网关(SGW)126可以被配置用于客户VCN 104,并为VCN 104与服务网络110中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。这种服务网络的示例是甲骨文公司的服务网络,其提供客户可以使用的各种服务。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,对象存储装置),而不需要公共IP地址或访问因特网。在某些实施例中,VCN可以仅具有一个SGW,并且连接只能从VCN内的子网启动而不能从服务网络110启动。如果一个VCN与另一个对等,则另一个VCN中的资源通常无法访问SGW。通过FastConnect或VPN Connect连接至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。

[0089] 在某些实施方式中,SGW 126使用服务无类域间路由 (CIDR) 标签的概念,所述服务无类域间路由标签是表示所关注服务或服务组的所有区域性公共IP地址范围的字符串。客户在配置SGW和相关路由规则以控制服务流量时使用服务CIDR标签。客户可以在配置安全规则时可选地使用所述标签,而无需在服务的公共IP地址将来发生变化的情况下对所述安全规则进行调整。

[0090] 本地对等网关 (LPG) 132是可以被添加到客户VCN 104的网关,并且使得VCN 104能够与同一区域中的另一个VCN对等。对等意味着VCN使用私有IP地址进行通信,而不需要流量穿过如因特网等公共网络,或者不需要通过客户的本地部署网络116来路由流量。在优选实施例中,VCN对于其所建立的每个对等具有单独的LPG。本地对等或VCN对等是一种用于在不同的应用程序或基础设施管理功能之间建立网络连通性的常见做法。

[0091] 服务提供商(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,服务可以被公开为公共端点,所述公共端点可由客户VCN中的计算实例经由如因特网等公共网络公开访问,或者可经由SGW 126以私有方式访问。根据特定的私有访问模型,可以将服务作为客户的VCN中的私有子网中的私有IP端点进行访问。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作为客户的私有网络中的实例公开。私有端点资源表示客户的VCN内的服务。每个PE在客户的VCN中由客户选择的子网中表现为VNIC (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式。由于端点被公开为VNIC,因此所有与VNIC相关联的特征(如路由规则、安全列表等)现在都可用于PE VNIC。

[0092] 服务提供商可以注册其服务以使得能够通过PE进行访问。提供商可以将策略与服务相关联,从而限制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,尤其是多租户服务。可能存在表示相同服务的多个这样的私有端点(在多个VCN中)。

[0093] 然后,私有子网中的计算实例可以使用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向客户VCN中PE的私有IP地址发送流量来访问服务。私有接入网关 (PAGW) 130是可以被附接到充当往返客户子网私有端点的所有流量的入口/出口点的服务提供商VCN(例如,服务网络110中的VCN)的网关资源。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,而是附接到客户希望与之交互的服务。指定用于私有端点的流量经由PAGW 130路由到服务。这些被称为客户到服务私有连接(C2S连接)。

[0094] 通过允许流量流经FastConnect/IPsec链路和客户VCN中的私有端点,PE概念还可以用于将对服务的私有访问扩展到客户的本地部署网络和数据中心。通过允许流量在LPG 132与客户的VCN中的PE之间流动,对服务的私有访问还可以扩展到客户的对等VCN。

[0095] 客户可以在子网级别上控制VCN中的路由,因此客户可以指定客户的VCN(如VCN 104)中的哪些子网使用每个网关。使用VCN的路由表来决定是否允许流量通过特定网关流出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。同一客户VCN 104内的私有子网的路由表可以通过SGW 126发送指定用于CSP服务的流量。所有剩余的流量可以经由NAT网关128发送。路由表仅控制从VCN离开的流量。

[0096] 使用与VCN相关联的安全列表来控制通过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。安全列表可以用于控制允许进出VCN的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,SSH为22,Windows RDP为3389)等。在某些实施方式中,实例的操作系统可以实施其自己的与安全列表规则一致的防火墙规则。规则可以是有状态的(例如,跟踪连接并且在没有用于响应流量的显式安全列表规则的情况下自动允许响应)或无状态的。

[0097] 来自客户VCN的访问(即,由部署在VCN 104上的资源或计算实例进行的访问)可以被分类为公共访问、私有访问或专用访问。公共访问是指一种使用公共IP地址或NAT来访问公共端点的访问模型。私有访问使得VCN 104中具有私有IP地址的客户工作负载(例如,私有子网中的资源)能够在不穿过如因特网等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与位于客户的私有网络外部的服务公共端点之间建立虚拟链路来提供私有访问模型。

[0098] 另外地,CSPI可以使用如FastConnect公共对等的技术来提供专用公共访问,其中,客户本地部署实例可以使用FastConnect来访问客户VCN中的一项或多项服务,而无需穿过如因特网等公共网络。CSPI还可以使用FastConnect私有对等来提供专用私有访问,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是使用公共因特网将客户的本地部署网络连接至CSPI及其服务的网络连通性替代方案。与基于因特网的连接相比,FastConnect提供了一种简单、灵活和经济的方式来创建具有更高的带宽选项和更可靠、更一致的联网体验的专用和私有连接。

[0099] 图1和以上附带说明描述了示例虚拟网络中的各种虚拟化部件。如上所述,虚拟网络建立在底层物理或基底网络上。图2描绘了根据某些实施例的CSPI 200内的物理网络中为虚拟网络提供底层的物理部件的简化架构图。如图所示,CSPI 200提供了包括由云服务提供商(CSP)提供的部件和资源(例如,计算资源、存储器资源和联网资源)的分布式环境。这些部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,为客户供应CSPI 200的资源子集(例如,计算资源、存储器资源和联网资源)。然后,客户可以使用由CSPI200提供的物理计算资源、存储器资源和联网资源来构建其自己的基于云(即,CSPI托管)的可定制私有虚拟网络。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上部署如计算实例等一种或多种客户资源。计算实例可以采取虚拟机、裸机实例等形式。CSPI 200提供基础设施和一套互补的云服务,所述一套互补的云服务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程序和服务。

[0100] 在图2所描绘的示例实施例中,CSPI 200的物理部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所描绘的各种计算实例可以由图2所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所描绘的VNIC和VCN VR可以由图2所描绘的NVD执行。图1所描绘的网关可以由主机和/或图2所描绘的NVD执行。

[0101] 主机或服务器可以执行在主机上创建并启用虚拟化环境的管理程序(也称为虚拟机监视器或VMM)。虚拟化或虚拟化环境促进了基于云的计算。主机上的管理程序可以在该主机上创建、执行和管理一个或多个计算实例。主机上的管理程序使得主机的物理计算资源(例如,计算资源、存储器资源和联网资源)能够在由主机执行的各种计算实例之间共享。

[0102] 例如,如图2所描绘的,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实施。通常,管理程序是位于主机的操作系统(OS)之上的进程或软件层,而所述操作系统又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,如处理器/核心等处理资源、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,处理资源、存储器资源和联网资源)能够在由主机202执行的计算实例(例如,虚拟机)之间共享。虚拟机可以有其自己的操作系统(称为访客操作系统),所述操作系统可以与主机的OS相同或不同。由主机执行的虚拟机的操作系统可以与由同一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并行执行,同时共享主机的相同计算资源。图2所描绘的主机可以具有相同或不同类型的管理程序。

[0103] 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。

[0104] 在某些情况下,整个主机可以被供应给单个客户,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)属于同一客户。在其他情况下,主机可以在多个客户(即,多个租户)之间共享。在这种多租户场景中,主机可以托管属于不同客户的虚拟机计算实例。这些计算实例可以是不同客户的不同VCN的成员。在某些实施例中,裸机计算实例由不具有管理程序的裸机服务器托管。当供应裸机计算实例时,单个客户或租户保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。

[0105] 如前所述,作为VCN的一部分的每个计算实例都与VNIC相关联,这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促进了往返计算实例的数据包或帧的通信。当创建计算实例时,VNIC与所述计算实例相关联。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由连接至主机的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由连接至主机202的NVD 210执行。作为另一个示例,由主机206托管的裸机实例272与由连接至主机206的NVD 212执行的VNIC 280相关联。作为又一个示例,VNIC 284与由主机208执行的计算实例274相关联,并且VNIC 284由连接至主机208的NVD 212执行。

[0106] 对于由主机托管的计算实例,连接至该主机的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所描绘的实施例中,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。

[0107] 主机可以包括使得主机能够连接至其他设备的一个或多个网络接口卡 (NIC)。主机上的NIC可以提供使得主机能够通信地连接至另一个设备的一个或多个端口 (或接口)。例如,主机可以使用主机和NVD上提供的一个或多个端口 (或接口)连接至NVD。主机也可以连接至其他设备,如另一个主机。

[0108] 例如,在图2中,主机202使用链路220连接至NVD 210,所述链路在由主机202的NIC 232提供的端口234与NVD 210的端口236之间延伸。主机206使用链路224连接至NVD 212,所述链路在由主机206的NIC 244提供的端口246与NVD 212的端口248之间延伸。主机208使用链路226连接至NVD 212,所述链路在由主机208的NIC250提供的端口252与NVD 212的端口254之间延伸。

[0109] NVD进而经由通信链路连接至架顶式 (TOR) 交换机,所述交换机连接至物理网络218 (也称为交换结构)。在某些实施例中,主机与NVD之间以及NVD与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连接至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。连接至TOR的主机和NVD的集合有时被称为机架。

[0110] 物理网络218提供了使得TOR交换机能够彼此通信的通信结构。物理网络218可以是多层网络。在某些实施方式中,物理网络218是交换机的多层Clos网络,其中TOR交换机214和216表示多层和多节点物理交换网络218的叶级节点。可能有不同的Clos网络配置,包括但不限于2层网络、3层网络、4层网络、5层网络以及通常的“n”层网络。Clos网络的示例在图5中描绘并在下文中描述。

[0111] 主机与NVD之间可能有各种不同的连接配置,如一对一配置、多对一配置、一对多配置等。在一对一配置的实施方式中,每个主机都连接至其自己的单独NVD。例如,在图2中,主机202经由主机202的NIC 232连接至NVD 210。在多对一配置中,多个主机连接至一个NVD。例如,在图2中,主机206和208分别经由NIC 244和250连接至同一NVD 212。

[0112] 在一对多配置中,一个主机连接至多个NVD。图3示出了CSPI 300内的主机连接至多个NVD的示例。如图3所示,主机302包括网络接口卡 (NIC) 304,所述网络接口卡包括多个端口306和308。主机302经由端口306和链路320连接至第一NVD 310,并且经由端口308和链路322连接至第二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可以是以太网链路。NVD 310进而连接至第一TOR交换机314,而NVD 312连接至第二TOR交换机316。NVD 310和312与TOR交换机314和316之间的链路可以是以太网链路。TOR交换机314和316表示多层物理网络318中的第0层交换设备。

[0113] 图3所描绘的布置提供了往返物理交换网络318到主机302的两条单独的物理网络路径:第一条路径穿过TOR交换机314到NVD 310到主机302,而第二条路径穿过TOR交换机316到NVD 312到主机302。单独的路径提供了主机302的增强可用性 (称为高可用性)。如果其中一条路径 (例如,其中一条路径中的链路出现故障) 或设备 (例如,特定NVD不起作用) 存在问题,则另一条路径可以用于往/返主机302的通信。

[0114] 在图3所描绘的配置中,主机使用由主机的NIC提供的两个不同端口连接至两个不同的NVD。在其他实施例中,主机可以包括允许主机至多个NVD的连通性的多个NIC。

[0115] 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部件。NVD可以是任何具有一个或多个处理单元(例如,CPU、网络处理单元(NPU)、FPGA、数据包处理流水线等)、包括缓存的存储器和端口的设备。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。

[0116] NVD可以以各种不同的方式实施。例如,在某些实施例中,NVD被实施为称为智能NIC或智慧NIC的具有板载嵌入式处理器的接口卡。智能NIC是独立于主机上的NIC的设备。在图2中,NVD 210和212可以被实施为分别连接至主机202以及主机206和208的智能NIC。

[0117] 然而,智能NIC仅是NVD实施方式的一个示例。可能有各种其他实施方式。例如,在一些其他实施方式中,NVD或由NVD执行的一个或多个功能可以被结合到一个或多个主机、一个或多个TOR交换机以及CSPI 200的其他部件中或者由这些部件来执行。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。作为另一个示例,NVD可以是TOR交换机的一部分,或者TOR交换机可以被配置成执行由NVD执行的功能,所述NVD使得TOR交换机能够执行用于公共云的各种复杂的数据包转换。执行NVD的功能的TOR有时被称为智能TOR。在向客户提供虚拟机(VM)实例而不是裸机(BM)实例的又其他实施方式中,由NVD执行的功能可以在主机的管理程序内部实施。在一些其他实施方式中,NVD的一些功能可以被卸载到在主机机群上运行的集中式服务。

[0118] 在某些实施例中,如当实施为如图2所示的智能NIC时,NVD可以包括多个物理端口,所述多个物理端口使其能够连接至一个或多个主机和一个或多个TOR交换机。NVD上的端口可以分为面向主机的端口(也称为“南端口”)或面向网络或TOR的端口(也称为“北端口”)。NVD的面向主机的端口是用于将NVD连接至主机的端口。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连接至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以及NVD 212上的端口258。如图2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连接至TOR交换机214。同样地,NVD 212使用从NVD 212的端口258延伸到TOR交换机216的链路230连接至TOR交换机216。

[0119] NVD经由面向主机的端口从主机接收数据包和帧(例如,由主机托管的计算实例生成的数据包和帧),并且在执行必要的数据包处理之后,可以经由NVD的面向网络的端口将数据包和帧转发到TOR交换机。NVD可以经由NVD的面向网络的端口从TOR交换机接收数据包和帧,并且在执行必要的数据包处理之后,可以经由NVD的面向主机的端口将数据包和帧转发到主机。

[0120] 在某些实施例中,在NVD与TOR交换机之间可以有多个端口和相关链路。这些端口和链路可以被聚合以形成多个端口或链路的链路聚合器组(称为LAG)。链路聚合使得两个端点之间(例如,NVD与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下运行。LAG有助于增加两个端点之间连接的带宽和可靠性。如果LAG中的物理链路之一发生故障,则流量将被动态且透明地重新指派给LAG中的其他物理链路之一。聚合的物理链路传递的带宽比每个单独链路传递的带宽高。与LAG相关联的多个端口被视为单个逻辑端口。流量可以在LAG的多个物理链路上实现负载平衡。可以在两个端点之间配置一个或多个LAG。这两个端点可以介于NVD与TOR交换机之间、介于主机与NVD之间等。

[0121] NVD实施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：数据包封装和解封装功能；用于创建VCN网络的功能；用于实施网络策略的功能，如VCN安全列表(防火墙)功能；有助于在VCN中路由和转发往返计算实例的数据包的功能；等等。在某些实施例中，在接收到数据包后，NVD被配置成执行数据包处理流水线，以处理数据包并确定要如何转发或路由数据包。作为该数据包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，如执行与VCN中的cis相关联的VNIC、执行与VCN相关联的虚拟路由器(VR)、封装和解封装数据包以促进虚拟网络中的转发或路由、执行某些网关(例如，本地对等网关)、实施安全列表、网络安全组、网络地址转换(NAT)功能(例如，逐个主机地将公共IP转换为私有IP)、节流功能以及其他功能。

[0122] 在某些实施例中，NVD中的数据包处理数据路径可以包括多个数据包流水线，每个数据包流水线由一系列数据包转换级构成。在某些实施方式中，在接收到数据包后，对数据包进行解析并将其分类到单个流水线。然后，以线性方式一级接一级地处理数据包，直到数据包被丢弃或通过NVD的接口发送出去。这些级提供了基本的功能性数据包处理构造块(例如，验证报头、实施节流、插入新的第2层报头、实施L4防火墙、VCN封装/解封装等)，使得可以通过组合现有的级来构建新的流水线，并且可以通过创建新的级并将其插入到现有流水线中来添加新的功能。

[0123] 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功能和数据平面功能。VCN控制平面的示例还在图16、图17、图18和图19中描绘(参见附图标记1616、1716、1816和1916)并在下文中描述。VCN数据平面的示例在图16、图17、图18和图19中描绘(参见附图标记1618、1718、1818和1918)并在下文中描述。控制平面功能包括用于配置网络的功能(例如，建立路由和路由表、配置VNIC等)，所述网络控制要如何转发数据。在某些实施例中，提供了VCN控制平面，所述VCN控制平面集中计算所有覆盖-基底映射并将其发布到NVD和虚拟网络边缘设备，如各种网关，如DRG、SGW、IGW等。防火墙规则也可以使用相同的机制发布。在某些实施例中，NVD仅获得与该NVD相关的映射。数据平面功能包括基于使用控制平面设置的配置来实际路由/转发数据包的功能。VCN数据平面是通过在客户的网络数据包穿过基底网络之前对其进行封装来实施的。封装/解封装功能在NVD上实施。在某些实施例中，NVD被配置成拦截进出主机的所有网络数据包并执行网络虚拟化功能。

[0124] 如上文所指示的，NVD执行各种虚拟化功能，包括VNIC和VCN VR。NVD可以执行与由连接至VNIC的一个或多个主机托管的计算实例相关联的VNIC。例如，如图2所描绘的，NVD 210执行与由连接至NVD 210的主机202托管的计算实例268相关联的VNIC 276的功能。作为另一个示例，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(其属于不同客户)的计算实例，并且连接至主机的NVD可以执行与计算实例相对应的VNIC(即，执行VNIC相关功能)。

[0125] 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所描绘的实施例中，NVD 210执行与计算实例268所属的VCN相对应的VCN VR 277。NVD 212执行与由主机206和208托管的计算实例所属的一个或多个VCN相对应的一个或多个VCN VR 283。在某些实施例中，与该VCN相对应的VCN VR由连接至主机的所有NVD执行，所述主机托管属于该VCN的至少一个计算实例。如果主机托管属于不同VCN的计算实例，则连接至该主机的NVD可以

执行与那些不同VCN相对应的VCN VR。

[0126] 除VNIC和VCN VR之外,NVD还可以执行各种软件(例如,守护进程)并且包括便于NVD执行各种网络虚拟化功能的一个或多个硬件部件。为了简洁起见,这些不同部件被组合在一起作为“数据包处理部件”,如图2所示。例如,NVD 210包括数据包处理部件286,而NVD 212包括数据包处理部件288。例如,用于NVD的数据包处理部件可以包括数据包处理器,所述数据包处理器被配置成与NVD的端口和硬件接口交互,以监测由NVD接收和使用其传送的所有数据包并存储网络信息。网络信息可以例如包括识别由NVD处理的不同网络流的网络流信息和每流信息(例如,每流统计数据)。在某些实施例中,网络流信息可以在每个VNIC的基础上存储。数据包处理器可以执行逐个数据包操作,并且实施有状态NAT和L4防火墙(FW)。作为另一个示例,数据包处理部件可以包括复制代理,所述复制代理被配置成将由NVD存储的信息复制到一个或多个不同的复制目标存储器。作为又一个示例,数据包处理部件可以包括日志记录代理,所述日志记录代理被配置成执行NVD的日志记录功能。数据包处理部件还可以包括用于监测NVD的性能和健康并且还可能监测连接至NVD的其他部件的状态和健康的软件。

[0127] 图1示出了示例虚拟或覆盖网络的部件,包括VCN、VCN内的子网、部署在子网上的计算实例、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所描绘的覆盖部件可以由图2所描绘的物理部件中的一个或多个来执行或托管。例如,VCN中的计算实例可以由图2所描绘的一个或多个主机来执行或托管。对于由主机托管的计算实例,与该计算实例相关联的VNIC通常由连接至该主机的NVD执行(即,VNIC功能由连接至该主机的NVD提供)。VCN的VCN VR功能由所有连接至主机的NVD执行,所述主机托管或执行作为该VCN的一部分的计算实例。与VCN相关联的网关可以由一种或多种不同类型的NVD执行。例如,某些网关可以由智能NIC执行,而其他网关可以由一个或多个主机或NVD的其他实施方式执行。

[0128] 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,端点可以与源计算实例处于同一子网内,与源计算实例处于不同子网但在同一的VCN内)或者与源计算实例的VCN外部的端点通信。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这些通信。

[0129] 对于VCN中的同一子网上的两个计算实例之间的通信,使用与源计算实例和目的地计算实例相关联的VNIC来促进通信。源计算实例和目的地计算实例可以由同一主机托管,也可以由不同的主机托管。源自源计算实例的数据包可以从托管源计算实例的主机转发到连接至该主机的NVD。在NVD上,使用数据包处理流水线来处理数据包,所述数据包处理流水线可以包括执行与源计算实例相关联的VNIC。由于数据包的目的地端点在同一子网内,因此执行与源计算实例相关联的VNIC会导致数据包被转发到执行与目的地计算实例相关联的VNIC的NVD,所述NVD然后对数据包进行处理并将其转发到目的地计算实例。与源计算实例和目的地计算实例相关联的VNIC可以在同一NVD上执行(例如,当源计算实例和目的地计算实例由同一主机托管时),也可以在不同的NVD上执行(例如,当源计算实例和目的地计算实例由连接至不同NVD的不同主机托管时)。VNIC可以使用由NVD存储的路由/转发表来确定数据包的下一跳地址。

[0130] 对于要从子网中的计算实例传送到同一VCN中的不同子网中的端点的数据包,源自源计算实例的数据包从托管源计算实例的主机传送到连接至该主机的NVD。在NVD上,使

用数据包处理流水线来处理数据包,所述数据包处理流水线可以包括执行一个或多个VNIC以及与VCN相关联的VR。例如,作为数据包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)。由VNIC执行的功能可以包括查看数据包上的VLAN标签。由于数据包的目的地在子网外部,因此VCN VR功能接下来由NVD调用和执行。然后,VCN VR将数据包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对数据包进行处理并将数据包转发到目的地计算实例。与源计算实例和目的地计算实例相关联的VNIC可以在同一NVD上执行(例如,当源计算实例和目的地计算实例由同一主机托管时),也可以在不同的NVD上执行(例如,当源计算实例和目的地计算实例由连接至不同NVD的不同主机托管时)。

[0131] 如果数据包的目的地在源计算实例的VCN外部,则将源自源计算实例的数据包从托管源计算实例的主机传送到连接至该主机的NVD。NVD执行与源计算实例相关联的VNIC。由于数据包的目的地端点在VCN外部,因此数据包由该VCN的VCN VR处理。NVD调用VCN VR功能,这可能导致数据包被转发到执行与VCN相关联的适当网关的NVD。例如,如果目的地是客户的本地部署网络内的端点,则数据包可以由VCN VR转发到执行为VCN配置的DRG网关的NVD。VCN VR可以与执行与源计算实例相关联的VNIC的NVD在同一NVD上执行,或者由不同的NVD执行。网关可以由NVD执行,所述NVD可以是智能NIC、主机或其他NVD实施方式。然后,数据包由网关处理并且被转发到有助于将数据包传送到其预期目的地端点的下一跳地址。例如,在图2所描绘的实施例中,源自计算实例268的数据包可以通过链路220(使用NIC 232)从主机202传送到NVD 210。在NVD 210上,调用VNIC 276,因为其与源计算实例268相关联的VNIC。VNIC 276被配置成检查数据包中的封装信息,并确定用于转发数据包的下一跳地址以便于将数据包转发到其预期目的地端点,然后将数据包转发到所确定的下一跳地址。

[0132] 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同一VCN或其他VCN(其可以是客户的VCN,也可以是不属于客户的VCN)中的实例。由CSPI 200托管的端点之间的通信可以通过物理网络218来执行。计算实例还可以与不由CSPI 200托管或者在CSPI 200外部的端点通信。这些端点的示例包括客户的本地部署网络或数据中心内的端点或可通过如因特网等公共网络访问的公共端点。可以使用各种通信协议通过公共网络(例如,因特网)(图2中未示出)或私有网络(图2中未示出)来执行与CSPI 200外部的端点的通信。

[0133] 图2所描绘的CSPI 200的架构仅是示例,并不旨在进行限制。在替代实施例中,可能有变化、替代和修改。例如,在一些实施方式中,CSPI 200可以具有比图2所示的那些系统或部件更多或更少的系统或部件,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所描绘的系统、子系统和其他部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)、使用硬件或其组合实施。软件可以存储在非暂态存储介质上(例如,存储器设备上)。

[0134] 图4描绘了根据某些实施例的主机与NVD之间的用于提供I/O虚拟化以支持多租户的连通性。如图4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,即属于客户/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连接至NVD 412的物理NIC 410。每个计算实例被附接到由NVD 412执行的VNIC。在图4的实施例中,VM1 406附接到VNIC-VM1 420,而VM2 408附接到VNIC-VM2 422。

[0135] 如图4所示,NIC 410包括两个逻辑NIC,即逻辑NIC A 416和逻辑NIC B 418。每个虚拟机都附接到其自己的逻辑NIC并且被配置成用其自己的逻辑NIC运行。例如,VM1 406附接到逻辑NIC A 416,而VM2 408附接到逻辑NIC B 418。即使主机402仅包括由多个租户共享的一个物理NIC 410,但由于逻辑NIC,每个租户的虚拟机也认为其具有自己的主机和NIC。

[0136] 在某些实施例中,为每个逻辑NIC指派其自己的VLAN ID。因此,为租户#1的逻辑NIC A 416指派特定的VLAN ID,而为租户#2的逻辑NIC B 418指派不同的VLAN ID。当从VM1 406传送数据包时,由管理程序将指派给租户#1的标签附接到所述数据包,然后将所述数据包通过链路414从主机402传送到NVD 412。以类似的方式,当从VM2 408传送数据包时,由管理程序将指派给租户#2的标签附接到所述数据包,然后将所述数据包通过链路414从主机402传送到NVD 412。因此,从主机402传送到NVD 412的数据包424具有标识特定租户和相关VM的相关标签426。在NVD上,对于从主机402接收的数据包424,使用与所述数据包相关联的标签426来确定所述数据包是要由VNIC-VM1 420处理还是由VNIC-VM2 422处理。然后,数据包由对应的VNIC处理。图4所描绘的配置使得每个租户的计算实例能够相信其拥有自己的主机和NIC。图4所描述的设置提供了用于支持多租户的I/O虚拟化。

[0137] 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所描绘的实施例被构造为Clos网络。Clos网络是一种特定类型的网络拓扑,其被设计成提供连接冗余,同时保持高二分带宽和最大资源利用率。Clos网络是一种非阻塞、多级或多层交换网络,其中,级或层的数量可以为两个、三个、四个、五个等。图5所描绘的实施例是包括层1、2和3的3层网络。TOR交换机504表示Clos网络中的第0层交换机。一个或多个NVD连接至TOR交换机。第0层交换机也称为物理网络的边缘设备。第0层交换机连接至第1层交换机(也称为叶交换机)。在图5所描绘的实施例中,一组“n”个第0层TOR交换机连接至一组“n”个第1层交换机,并且一起形成pod。pod中的每个第0层交换机与所述pod中的所有第1层交换机互连,但是pod之间没有交换机连通性。在某些实施方式中,两个pod称为块。每个块由一组“n”个第2层交换机(有时称为主干交换机)提供服务或与其连接。物理网络拓扑中可以有几个块。第2层交换机进而连接至“n”个第3层交换机(有时称为超级主干交换机)。物理网络500上的数据包通信通常使用一个或多个第3层通信协议来执行。通常,除了TOR层之外,物理网络的所有层都是n向冗余的,从而允许高可用性。可以为pod和块指定策略,以控制物理网络中交换机之间的可见性,从而实现物理网络的扩展。

[0138] Clos网络的特征是,从一个第0层交换机到另一个第0层交换机(或从连接至第0层交换机的NVD到另一个连接至第0层交换机的NVD)要达到的最大跳数是固定的。例如,在3层Clos网络中,数据包从一个NVD到达另一个NVD最多需要七跳,其中,源和目标NVD连接至Clos网络的叶层。同样地,在4层Clos网络中,数据包从一个NVD到达另一个NVD最多需要九跳,其中,源和目标NVD连接至Clos网络的叶层。因此,Clos网络架构在整个网络中保持一致的延迟,这对于数据中心内部和之间的通信非常重要。Clos拓扑水平扩展并且具有经济效益。通过在不同层处添加更多的交换机(例如,更多的叶交换机和主干交换机)以及通过增加相邻层处的交换机之间的链路数量,可以容易地增加网络的带宽/吞吐量。

[0139] 在某些实施例中,为CSPI内的每种资源指派被称为云标识符(CID)的唯一标识符。该标识符作为资源信息的一部分被包括在内,并且可以用于例如经由控制台或API来管理

资源。CID的示例语法如下：

[0140] ocidl.<RESOURCETYPE>.<REALM>.[REGION][.FUTUREUSE].<UNIQUEID>

[0141] 其中，

[0142] ocidl:指示CID版本的文字字符串；

[0143] resource type:资源类型(例如,实例、卷、VCN、子网、用户、组等)；

[0144] realm:资源所在的领域。商业领域的示例值为“c1”,政府云领域的示例值为“c2”,或者联邦政府云领域的示例值为“c3”,等等。每个领域可以有其自己的域名；

[0145] region:资源所在的区域。如果区域不适用于资源,则该部分可能为空白；

[0146] future use:保留以供将来使用；

[0147] unique ID:ID的唯一部分。格式可能会根据资源或服务的类型而有所不同。

[0148] 环路防止的示例

[0149] 图6图示了根据某些实施例的为客户(如企业)配置和部署的VCN 600的示例。VCN 600包括虚拟L2网络610,并且可以在甲骨文云基础设施(OCI)中实施。

[0150] 通常,VCN 600可以是包括子网、路由表和网关的传统网络的软件定义的虚拟版本,各种计算实例620A、620B至620K可以在其上运行(本文统称为“计算实例620”或单个地称为“计算实例620”)。例如,VCN 600是客户在云基础设施中设置的虚拟私有网络。云基础设施可能驻留在某个特定区域内,并且其包括所述区域的所有可用性域。子网是VCN 600的子部分。云基础设施中定义的每个子网可以位于单个可用性域中,也可以跨越区域中的所有可用性域。在可以启动计算实例620之前,可能需要建立至少一个云基础设施。云基础设施可以配置有可选的因特网网关以处理公共流量,以及可选的IP安全(IPSec)、虚拟私有网络(VPN)连接或OCI FastConnect以安全地扩展客户的本地部署网络。VCN 600可以被私有地连接至另一个VCN,使得流量不穿过因特网。两个VCN的无类域间路由(CIDR)可能不重叠。

[0151] 在图6的图示中,虚拟L2网络610支持任何上述虚拟联网并且被实施为云基础设施上的覆盖层,其中,所述覆盖层实施L2协议。计算实例620可以属于虚拟L2网络610,并且使用L2协议经由所述网络进行连接。特别地,虚拟L2网络610可以包括虚拟交换机、VNIC和/或用于从计算实例620接收帧和向所述计算实例路由帧的其他虚拟计算资源。覆盖层在云基础设施的底层计算资源上实施这种虚拟化。

[0152] 例如,计算实例620A被托管在经由第一端口连接至NVD的第一主机上。NVD托管计算实例620A的第一VNIC,其中,该第一VNIC是虚拟L2网络610的部件。为了向计算实例620K传输帧,根据L2协议,计算实例620A在帧中包括作为源地址的计算实例620A的MAC地址和作为目的地地址的计算实例620K的MAC地址。这些MAC地址也可以被称为覆盖MAC地址或逻辑MAC地址。虚拟L2网络610经由第一VNIC接收帧,并将其发送到与计算实例620K相对应的第二VNIC。实际上,第一主机经由第一端口将帧转发到NVD,并且NVD将帧路由到计算实例620K的第二主机。路由包括用相关信息(例如,主机的MAC地址,也可以称为物理MAC地址或基底MAC地址)封装帧。

[0153] 图7图示了根据某些实施例的包括为客户建立的虚拟L2和第3层(L3)网络的VCN 700的示例。在示例中,VCN 700包括虚拟L3网络710。虚拟L3网络710可以是实施OSI模型的L3层的协议的虚拟网络,并且在其中可以建立多个子网712A至712N(本文统称为“子网712”或者单个地称为“子网712”)。另外,VCN 700包括实施OSI模型的L2层的协议的虚拟L2网络

720,如图6的虚拟L2网络610。交换虚拟接口(SVI)路由器730连接虚拟L3网络710和虚拟L2网络720。虚拟L2网络720可以具有指定的SVI IP地址。子网之间的流量可以通过SVI路由器730。

[0154] 在虚拟L3网络中,每个子网712可以具有不与VCN 700中的其他子网重叠的连续的IP版本4(IPv4)或版本6(IPv6)地址范围(例如,10.0.0.0/24和10.0.1.0/24)。子网充当配置单元。给定子网中的所有计算实例使用相同的路由表、安全列表和动态主机配置协议(DHCP)选项。子网712还可以包括VNIC。每个计算实例620可以附接到VNIC,所述VNIC可以驻留在子网中并为计算实例620启用网络连接。子网可以是公共的,也可以是私有的。私有子网可以具有无公共因特网协议(IP)地址的VNIC。公共子网可以具有可以有公共IP地址的VNIC。子网可以存在于单个可用性域中,也可以存在于特定区域的多个可用性域中。

[0155] 在虚拟L2网络中,VNIC也可以附接到计算实例,以使用L2协议实现计算实例之间的通信。无论是在L2还是L3中,VNIC都可以确定计算实例620如何与端点连接,并且可以被托管在NVD(也称为智能服务器适配器(ISA))上。L2 VNIC使用L2协议实现计算实例之间的通信。而L3 VNIC使用L3协议实现通信。

[0156] 每个计算实例620可以具有在计算实例启动期间创建并且不能被移除的主VNIC。次级VNIC(与主VNIC在同一可用性域中)可以被添加到现有的计算实例620中并且可以根据需要被移除。

[0157] L2 VNIC可以托管在实施交换机功能的NVD上。NVD可以基于来自源的传入帧逐渐了解计算实例的源MAC地址,维护L2转发表,并基于MAC地址和转发表传输帧。传输帧包括传输实际帧(例如,转发帧)或帧的一个或多个副本(例如,广播帧)。例如,当从某个端口接收到具有源MAC地址的帧时,NVD可能会了解到源MAC地址的帧可以通过所述端口发送。NVD还可以缓存和老化MAC地址。NVD可以保存SVI路由器730的静态(非老化)MAC地址。当接收到具有未知目的地MAC地址或广播目的地MAC地址的帧时,NVD可以广播(泛洪)这些帧。

[0158] 通常,OSI模型包含七层。第一层(L1)是与通过物理介质传输数据位相关联的物理层。L2是数据链路层,其指定物理层上连接节点之间的帧传输。L3是网络层,其描述多节点网络的寻址、路由和流量控制。

[0159] L2是广播MAC级网络,而L3是IP网络上的分段路由。L2可以包括两个子层,其中,MAC层批准设备访问和传输媒体,并且逻辑链路层(LLC)负责管理通信链路和处理帧流量,如识别网络层上的协议以及检查错误和帧同步。L3用IP地址运行,而L2用MAC地址运行。MAC地址是每个设备中存在的网络适配器的唯一标识符。IP地址是比MAC地址更高的抽象层,因此可能会更慢。IP地址通常由DHCP服务器“租用”或“指派”并且可以更改,而MAC地址是网络适配器的固定地址,并且在不更改硬件适配器的情况下无法在设备上更改。帧是L2网络上的协议数据单元。帧具有定义的结构,并且可以用于错误检测、控制平面活动等。网络可以使用一些帧来控制数据链路本身。

[0160] 在L2处,单播是指将帧从一个节点发送到另一个节点,多播是指将流量从一个节点发送到多个节点,并且广播是指将帧传输到网络中的所有节点。广播域是网络的逻辑分区,其中,该网络的所有节点都可以通过广播到达L2。LAN的区段可以使用网桥在帧级连接。桥接在LAN上创建单独的广播域,从而创建VLAN,VLAN是将相关设备组合成单独区段的独立逻辑网络。VLAN上的设备分组与设备在LAN中的物理位置无关。在VLAN上,源和目的地处于

同一VLAN中的帧只能在本地VLAN内转发。

[0161] L2网络设备是一种多端口设备,其使用硬件地址和MAC地址在数据链路层处理和转发数据。帧基于目的地MAC地址被发送到特定的交换机端口。L2交换机可以自动学习MAC地址,从而构建可以用于选择性地转发帧的表。例如,如果交换机在端口1上接收到来自MAC地址X的帧,则交换机可以了解到指定用于MAC地址X的帧可以从端口1转发出去,而不必尝试每个可用的端口。因为L2信息很容易取得,所以帧可以(通常以网络的线速)被非常快速地转发(交换)。同一L2区段中的设备不需要路由就可以到达本地对等设备。因此,L2交换对网络性能或带宽几乎没有影响。另外地,L2交换机便宜且易于部署。但是,L2交换机在转发帧时可能不会应用任何智能。例如,L2交换机本身可能不允许多路径并且可能存在环路问题,因为L2控制协议本身可能没有环路防止机制。其可能无法基于IP地址路由帧,也不能对由特定应用程序发送的帧进行优先级排序。在L3网络中,基于目的IP地址将数据包发送到特定的下一跳IP地址。L3虚拟网络(例如,L3 VCN)可以更好地扩展,可以本身支持多路径,并且不会遭受环路。但许多企业工作负载仍然使用L2虚拟网络(L2 VLAN),至少是因为L2虚拟网络可以为数百个或甚至数千个终端站提供简单、廉价、高性能的连通性。

[0162] 因为L2网络实施L2协议而且L2协议本身并不能防止环路,所以L2网络在引入环路时很容易受到广播风暴的影响。相比之下,L3网络可能不易受环路问题的影响。进一步地,L2和L3网络两者可能都需要包括多路径,以便在出现链路故障时提供冗余路径。可以在L2网络中实施STP来减轻环路问题。然而,这样做可能会妨碍L2网络中的多路径。

[0163] 特别地,STP可以调节交换机之间的互连并且可以禁用某些链路,使得所得L2网络是定义避免环路的单个路径的树。STP是为网络构建无环路逻辑拓扑的网络协议。之所以需要STP是因为LAN中的交换机通常使用冗余链路互连,以提高在一个连接出现故障时的恢复能力。然而,这种连接配置产生了导致广播辐射和MAC表不稳定的交换环路。如果使用冗余链路来连接交换机,则需要避免交换环路。因此,通常在交换机上实施STP以监测网络拓扑。STP的基本功能是防止桥接环路以及由此产生的广播辐射。生成树还允许网络设计包括备用链路,从而在活动链路发生故障时提供容错能力。STP创建了表征连接的L2交换机或网桥的网络内节点的关系的生成树。STP禁用不属于生成树的链路,从而在任何两个网络节点之间(如在树根与任何叶子之间)留下单个活动路径。所有其他路径强制进入备用(或阻塞)状态。基于STP的L2环路防止可以使用各种STP风格或扩展来实现,如快速生成树协议(RSTP)、多生成树协议(MSTP)和VLAN生成树协议(VSTP)。

[0164] 因此,对于L2网络,使用STP可能无法支持优化后的实施方式,因为这会导致多路径防止。如虚拟L2网络720等虚拟L2网络被实施为云基础设施上的L2网络的覆盖层。因此,虚拟L2网络也容易受到环路问题的影响。也可以在虚拟L2网络中实施STP,但是同样会导致多路径能力的损失。如结合接下来的附图进一步描述的,本公开的实施例支持环路防止和多路径。

[0165] 图8图示了根据某些实施例的支持VCN的虚拟L2网络的基础设施的示例。图8所描绘的部件通常由CSP提供,并且是其用于向订阅客户提供云服务(例如,IaaS服务)的CSPI的一部分。在示例中,部件与图2的架构图中描述的部件相同或类似。为了简洁起见,本文不再重复两个附图的部件之间的相似之处。

[0166] 如图8所描绘的,主机820A连接至NVD 810A,后者又连接至物理交换网络830的TOR

交换机832A。进一步地,主机820A至820L(本文统称为“主机820”或单个地称为“主机820”)经由以太网链路连接至NVD 810A,并且NVD 810A经由另一个以太网链路连接至TOR交换机832A。同样如图8所描绘的,主机821A至821L(本文统称为“主机821”或单个地称为“主机821”)经由以太网链路连接至NVD 810K,并且NVD 810K经由以太网链路连接至TOR交换机832K。

[0167] 每个主机820或821可以为一个或多个客户执行一个或多个计算实例。客户的计算实例可以是由客户配置并由CSPI托管在云中的一个或多个虚拟云网络(VCN)的一部分。NVD 810中的每一个还可以为每个计算实例托管一个或多个VNIC。客户的计算实例经由VNIC连接至虚拟L2网络,如图6的虚拟L2网络610或图7的L2虚拟网络720。该虚拟L2网络可以被实施为包括NVD 810、主机820和821以及交换网络830的基础设施上的覆盖层。

[0168] 在图8的图示中,NVD 810A包括将NVD 810A连接至TOR交换机832A的北端口812。NVD 810A还包括一组南端口814A至814L,每个端口连接至主机820A至820L之一。进一步地,NVD 810A托管多个VNIC 816A至816M,每个VNIC都附接到在主机820上执行的计算实例。

[0169] 同样在图8的图示中,主机820A经由南端口814A连接至NVD 810A。主机820A可以为客户执行计算实例822A。在NVD 810A上托管的VNIC 816A对应于计算实例822A。计算实例822A可以属于客户的VCN内的虚拟L2网络。类似地,主机820L经由南端口814L连接至NVD 810A并且为客户托管至少一个计算实例822M,其中,VNIC 816M对应于该计算实例822M。尽管图8图示了每个主机820连接至单个NVD(例如,NVD 810A),但是一些或所有主机820可以连接至多个NVD。

[0170] NVD 810K可以类似于NVD 810A。特别地,NVD 810K包括将NVD 810K连接至交换网络830的TOR交换机832K的北端口813。NVD 810K还包括一组南端口815A至815L,每个端口连接至主机821A至821L之一。进一步地,NVD 810K托管多个VNIC 817A至817M,每个VNIC都附接到在主机821上执行的计算实例。主机821A经由南端口815A连接至NVD 810K并为客户托管计算实例823A。VNIC 817A主机对应于计算实例823A。类似地,主机821L经由南端口815L连接至NVD 810K并且为客户托管至少一个计算实例823M,其中,VNIC 817M对应于该计算实例823M。尽管图8针对NVD 810A和810K中的每一个图示了相同数量的端口、相同数量的主机、相同数量的VNIC和相同数量的计算实例,但是这些数量中的每一个都可以不同。

[0171] 在虚拟级和硬件级处理从第一计算实例传输到第二计算实例的帧。特别地,第一计算实例822A生成包括有效载荷、作为源地址的第一计算实例的MAC地址和作为目的地地址的第二计算实例的MAC地址的帧。这些MAC地址位于虚拟网络中,因此是覆盖MAC地址。NVD 810A接收所述帧,并基于物理或虚拟端口(例如,VLAN)将其映射到正确的VNIC(例如,VNIC 816A)。NVD 810A然后执行必要的覆盖功能,包括数据包过滤、日志记录、目的地地址查找和覆盖封装。例如,第一NVD 810A使用来自帧的覆盖目的地MAC地址从NVD 810A的转发表中确定第二主机的物理MAC地址和可以发送帧的端口。如果在转发表中找到该信息,则NVD 810A在帧中用其MAC地址替换主机820A的物理MAC地址作为源,并用第二主机的物理MAC地址替换其MAC地址作为目的地,并且经由端口传输帧。如果在转发表中没有找到该信息,则NVD 810A确定帧将被泛洪,并且相应地,在帧中用其MAC地址替换第一主机820A的物理MAC地址作为源,并用广播MAC地址替换其MAC地址作为目的地,并且广播帧。传输帧可以包括传输实际帧或传输其多个副本(相当于广播帧)。

[0172] 相反,在经由北端口接收到帧后,NVD 810A执行必要的数据包转换,包括帧过滤、日志记录、源和/或目的地MAC地址查找、解封装和封装。这些转换的结果是NVD 810A确定数据包属于哪个VNIC以及数据包是否合法。如果帧通过了转发表中的查找过程,则NVD 810A确定相关端口(或VNIC和VLAN)经由所述端口将帧发送到主机。此处,与计算实例相对应的VNIC也可以接收和处理帧,从而改变帧中包括的覆盖地址。如果在转发表中没有发现匹配,则NVD 810A确定帧将被泛洪到面向主机的端口。

[0173] 图9图示了根据某些实施例的NVD之间的环路900的示例。环路900可以在图8所描述的基础设施的NVD 810A与NVD 810K之间以指数方式发生和增长。特别地,帧由NVD 810A接收并且具有不包括在NVD 810A的转发表中的目的地MAC地址。在这种情况下,接收NVD 810A可以泛洪该帧(例如,将帧发送到为该VCN或该客户的该VLAN服务的所有NVD)。如果NVD 810K也没有该目的地MAC地址的转发表条目,则其也必须泛洪该帧。由于泛洪,帧被发送回原始发送NVD 810A,从而在这两个NVD 810A与810K之间创建环路。再次重新发送帧时,广播可能会再次发生,从而增加环路。在短时间段内(例如,几毫秒),所述增加可能会显著泛洪网络并消耗网络带宽。

[0174] 在图9的示例中,帧910由在主机820A上执行的第一计算实例(例如,计算实例822A)生成,并被发送到在主机821A上执行的第二计算实例(例如,计算实例823A)。帧是包括L2报头和有效载荷的L2帧。有效载荷包括L2协议数据单元(L2 PDU)。报头包括作为源地址的第一计算实例的第一MAC地址(在图9中图示为“MAC 1”)和作为目的地地址的第二计算实例的第二MAC地址(在图9中图示为“MAC 3”)。然而,本公开的实施例不限于此,并且所述帧或其他帧可以在两个计算实例之间发送,这两个计算实例在同一主机上执行、在连接至同一NVD的不同主机上执行和/或在连接至不同NVD的不同主机上执行。在所有这些情况下,帧传输可以类似地导致环路900。

[0175] 主机820A经由南端口814A向NVD 810A发送帧910。NVD 810A查找转发表,并且在第二MAC地址与该表中的条目之间未识别到匹配。因此,NVD 810A确定泛洪包含未知目的地MAC地址的帧910。泛洪可以涉及经由NVD 810A的所有端口传输帧910。

[0176] 因此,主机820L接收包括第二MAC地址的帧910(例如,该帧910的副本)。因为该第二MAC地址不与由主机820L托管的计算实例的MAC地址相对应,所以帧副本930被丢弃。

[0177] NVD 810K接收帧910(例如,该帧的副本,所述副本已经封装有适当的信息,以便通过例如包括覆盖报头(在图中图示为“O.H.”)的方式通过交换网络830进行路由;覆盖报头可以包括VCN报头、源地址和目的地IP地址以及源地址和目的地MAC地址)。此处,尽管接收到的帧包括第二MAC地址,并且该第二MAC地址对应于在主机821A上执行的第二计算实例,但是NVD 810K尚未学习并且在其转发表中没有第二MAC地址的条目。因此,NVD 810K也确定接收到的帧将被泛洪,并经由其端口广播接收到的帧。因此,主机821中的每一个接收帧910(例如,该帧910的副本)。接收到的帧被丢弃,除非其目的地MAC地址与主机821上的计算实例的目的地MAC地址相匹配。

[0178] 另外,由NVD 810K进行的泛洪包括经由北端口813通过交换网络830将帧910(或其封装有适当信息以通过交换网络830进行路由的副本)传输回NVD 810A。由NVD 810A将帧910重新传输到NVD 810K以及由NVD 810AK将帧910传输回NVD 810A产生了环路900。所述环路可以在短时间内重复多次,从而导致大量网络带宽使用。

[0179] 图10图示了根据某些实施例的防止NVD之间与虚拟L2网络相关联的环路的示例。为了清楚起见,参考图8所描述的基础设施的NVD 810A和NVD 810K。然而,环路防止可以应用于支持虚拟L2网络的网络资源之间的任何其他环路,如交换机和。不同于图9,每个NVD实施减少环路发生的环路防止规则。环路防止规则是当应用时防止网络内产生一个或多个环路的规则。环路防止规则可以指定一个或多个参数和适用于支持该环路防止的一个或多个动作。(多个)参数和(多个)动作可以被编码为例如一组if-then语句。在示例中,环路防止规则规定,如果网络资源经由所述网络资源的端口接收到帧,则网络资源不能经由同一端口传输帧。特别地,当网络资源广播帧时,例如,因为帧具有在网络资源的转发表中不匹配的目的地MAC地址,所以广播不能经由接收帧的端口发生。

[0180] 在图10的图示中,NVD 810存储环路防止规则1012A。帧1010由在主机820A上执行的第一计算实例(例如,计算实例822A)生成,并被发送到在主机821A上执行的第二计算实例(例如,计算实例823A)。帧是包括报头和有效载荷的L2帧。有效载荷包括L2 PDU。报头包括作为源地址的第一计算实例的第一MAC地址(在图10中图示为“MAC 1”)和作为目的地地址的第二计算实例的第二MAC地址(在图10中图示为“MAC 3”)。然而,本公开的实施例不限于此,并且所述帧或其他帧可以在两个计算实例之间发送,这两个计算实例在同一主机上执行、在连接至同一NVD的不同主机上执行和/或在连接至不同NVD的不同主机上执行。

[0181] 主机820A经由南端口814A向NVD 810A发送帧1010。NVD 810A查找其转发表1014A,并且在第二MAC地址与该表中的条目之间未识别到匹配。此时,NVD 810A了解到第一MAC地址与南端口814A相关联,并更新其转发表1014A以存储该关联。进一步地,NVD 810A经由其南端口814(南端口814A除外)和北端口812广播帧1010。

[0182] 主机820L接收包括第二MAC地址的帧1010。因为该第二MAC地址不与由主机820L托管的计算实例的MAC地址相对应,所以帧副本被丢弃。

[0183] 进一步地,NVD 810K经由北端口813接收帧1010(或具有适当封装的帧1010的副本)。NVD 810K存储环路防止规则1012K,所述规则规定经由端口接收的帧不能经由同一端口传输回来。此处,尽管接收到的帧包括第二MAC地址,并且该第二MAC地址对应于在主机821A上执行的第二计算实例,但是NVD 810K尚未学习并且在其转发表1014K中没有第二MAC地址的条目。此时,NVD 810K了解到第一MAC地址与北端口813相关联,并更新其转发表1014K以存储该关联。进一步地,由于未知的目的地MAC地址,NVD 810K确定接收到的帧将被泛洪,并且因此,基于环路防止规则1012K经由其除了北端口813之外的端口广播该帧。因此,主机821接收帧1010(或其副本)。因此,主机821中的每一个接收帧1010(例如,该帧1010的副本)。接收到的帧被丢弃,除非其目的地MAC地址与主机821上的计算实例的目的地MAC地址相匹配。由于环路防止规则1012K,NVD 810K不会经由北端口813将帧1010传输回NVD 1010A。因此,防止了NVD 810A与NVD 810K之间的环路。

[0184] 上述环路不仅会发生在“未知的单播目的地MAC地址”上,还会发生在“广播MAC地址”上。由于“广播MAC地址”就其定义而言是在所有端口上发送的,因此其可能导致与上述相同的环路。上述环路防止机制也可以应用于这种“广播MAC地址”。

[0185] 图11图示了根据某些实施例的NVD与在主机上执行的计算实例之间的环路1100的示例。为了清楚起见,参考图8所描述的基础设施的NVD 810A、主机820A和计算实例822A。然而,在NVD与其他计算实例之间和/或计算实例与其他NVD之间可能存在环路。在示例中,NVD

810A发送寻址到计算实例822A的MAC地址的帧1110。帧经由南端口814A被发送到主机820A。在接收到帧1110后,并且由于计算实例822A的软件错误或软件代码,计算实例822A复制帧1110并将其发送回主机820A(在图11中示出为帧1120)。反过来,主机820A经由南端口814A将帧1120发送到接收所述帧的NVD 810A。因此,存在环路1100,其中,NVD 810A接收回其经由南端口814A发送的帧。

[0186] 图12图示了根据某些实施例的防止NVD与在主机上执行的计算实例之间的环路的示例。此处,同样为了解释清楚起见,参考图8所描述的基础设施的NVD 810A、主机820A和计算实例822A。然而,环路防止同样适用于任何其他NVD。在示例中,NVD 810A为NVD 810A的每个南端口实施轻量级STP 1200。在图12的图示中,轻量级STP 1200管理南端口814A。进一步地,可以为每个剩余的南端口实施类似的轻量级STP。

[0187] 对南端口814A的管理包括确定是否存在环路(如存在环路1100),并且如果存在环路,则禁用1202南端口814A。禁用1202南端口814A可以包括存储指示环路的标志并停用南端口814A。标志可以存储在与计算实例822A的MAC地址相关联的条目中,如存储在NVD 810A的转发表1014A中。可以使用多种技术来实施停用。在一种技术中,NVD 810A经由软件控制断开南端口814A。在附加或替代技术中,NVD 810A停止经由南端口814A进行帧传输,并忽略经由南端口814A接收的帧。所述停用可以应用于经由南端口814A发送到主机820A或从所述主机发送的所有帧,而不管主机820A托管的特定计算实例。可替代地,停用可以仅应用于发送到计算实例822A或从所述计算实例发送的帧。

[0188] 为了确定是否存在环路,轻量级STP 1200可以指定帧1210需要经由南端口814A进行传输。在计算实例822A启动后、在计算实例822A更新后和/或在去往和/或来自NVD 810A的帧拥塞达到拥塞水平后,可以周期性地触发传输。帧1210可以是BDPU。如果存在环路,则计算实例822A复制BDPU,并且主机820A将所得帧1220发送到NVD 810A。轻量级STP 1200检测到接收到的帧1220是BDPU,并且因此确定环路的存在并禁用1202南端口814A。如果不存在环路,则经由南端口814A接收的帧不是BDPU,并且轻量级STP 1200可以区分这样的帧与BDPU。因此,南端口814保持启用。

[0189] 图13至图15图示了用于在允许多路径的同时防止环路的流程示例。流程的操作可以由NVD来执行。用于执行流程操作的一些或所有指令可以被实施为硬件电路和/或存储为NVD的非暂态计算机可读介质上的计算机可读指令。如所实施的,指令表示包括可由NVD的处理器执行的电路或代码的模块。这些指令的使用将NVD配置成执行本文所述的特定操作。与(多个)相关处理器结合的每个电路或代码表示用于执行(多个)相应操作的装置。虽然以特定顺序图示了操作,但是应当理解,特定顺序不是必需的,并且一个或多个操作可以被省略、跳过、并行执行和/或重新排序。

[0190] 图13描绘了用于防止与虚拟L2网络(如图6的虚拟L2网络610或图7的虚拟L2网络720)相关联的环路的流程的示例。NVD可以属于提供虚拟L2网络的云基础设施。流程可以开始于操作1302,其中,NVD存储环路防止规则。环路防止规则可以指示由NVD经由NVD的端口接收的帧可以由NVD经由所述端口传输。在操作1304处,NVD可以维护转发表。例如,转发表可以包括条目,每个条目将端口与计算实例的MAC地址和/或计算实例的主机的MAC地址相关联。转发表可用于转发由NVD接收的并且包括目的地MAC地址的帧。特别地,如果目的地MAC地址与转发表中的条目匹配,则NVD会基于条目中的关联而转发帧。否则,NVD会经由除

接收帧所经由的端口之外的所有端口按照环路防止规则来传输帧。在操作1306处,NVD维护每个南端口的轻量级生成树。轻量级STP可以依靠传输BPDU和接收回BPDU来检测环路并禁用南端口。在操作1308处,NVD基于环路防止规则、转发表和轻量级生成树来路由帧。以这种方式,当经由NVD的一个或多个北端口接收的帧的目的地MAC地址与转发表中的条目不匹配时,这些帧不会经由这些端口被环回。另外地,检测到由于计算实例的软件错误或软件代码而产生的环路并对其进行管理。

[0191] 图14描绘了用于防止NVD之间与虚拟L2网络相关联的环路的流程的示例。此处,第一NVD经由交换网络与第二NVD连接,其中,帧可以经由其相关北端口在两个NVD之间交换。两个NVD中的每一个都可以存储环路防止规则。所述流程结合第一NVD的操作进行描述,并且可以类似地应用于第二NVD的操作。

[0192] 如所示的,所述流程可以开始于操作1402,其中,第一NVD可以经由第一NVD的第一端口(例如,当从第二NVD接收帧时为北端口)接收帧。帧可以源自第二计算实例,并且可以包括L2 PDU、作为目的地地址(例如,第一帧的目的地)的第一计算实例的第一MAC地址、以及作为源地址(例如,作为第一帧的源)的第二计算实例的第二MAC地址。第一计算实例和第二计算实例是虚拟L2网络的成员。第一计算实例由经由第二端口与网络接口卡连接的第一主机托管。在操作1404处,第一NVD可以查找其转发表。查找可以使用MAC地址来确定来自第一帧的MAC地址与转发表中的条目之间是否存在匹配。在操作1406处,第一NVD确定第二MAC地址是否与转发表中的条目相匹配。如果匹配,则所述流程移动到操作1408。否则,所述流程移动到操作1420,其中,第一NVD通过包括将第二MAC地址与第一端口相关联的条目来更新其转发表,使得如果后续帧被接收并且包括作为目的地地址的第二MAC地址,则该帧可以经由第一端口被转发。在操作1408处,第一NVD确定第一MAC地址是否与转发表中的条目相匹配。如果不匹配,则所述流程移动到操作1410。否则,所述流程移动到操作1430,其中,第一NVD将第一帧转发到目的地(例如,通过经由第二端口传输第一帧来转发到第一计算实例)。在操作1410处,第一NVD确定第一帧将基于第二MAC地址经由网络接口卡的所有端口进行传输。换句话说,由于不存在匹配的目的地,第一NVD确定需要通过其端口广播第一帧来泛洪第一帧。在操作1412处,第一NVD确定其环路防止规则防止使用接收帧所经由的端口传输帧。因此,该环路防止规则防止第一NVD经由第一端口发送第一帧。在操作1414处,NVD基于环路防止规则经由网络接口卡的除第一端口之外的所有端口传输第一帧。换句话说,通过应用环路防止规则,第一帧经由第一NVD的除第一端口之外的所有端口进行广播。

[0193] 图15描绘了用于防止NVD与计算实例之间与虚拟L2网络相关联的环路的流程的示例。所述流程同样适用于防止NVD与托管在一个或多个主机上的其他计算实例之间的环路,所述一个或多个主机经由NVD的南端口与NVD连接。所述示例流程可以开始于操作1502,其中,NVD经由NVD的第一端口(例如,南端口)传输第一帧。该端口与计算实例的主机相连。第一帧可以是BPDU,并且可以基于在NVD上实施的轻量级STP来传输。在操作1504处,NVD经由第一端口接收第二帧。在操作1506处,NVD确定第二帧是否是第一帧的副本。例如,如果第二帧也是其已经发送的BPDU(例如,在操作1504处接收回BPDU,并且BPDU具有其自己的MAC地址作为BPDU内部的网桥ID),则NVD确定存在副本并且操作1506之后是操作1520。否则,确定没有副本,并且操作1506之后是操作1510。在操作1510处,NVD不禁用第一端口,因为不存在环路。相比之下,在操作1520处,NVD确定存在环路。在操作1522处,NVD确定要执行指示一个

或多个动作的环路防止规则以管理环路。例如,该规则由NVD存储,并规定NVD的发生环路所经由的端口将被禁用。在操作1524处,NVD基于环路防止规则来执行动作(或所有动作)以防止环路。例如,NVD禁用第一端口。

[0194] 示例基础设施即服务架构

[0195] 如上所述,基础设施即服务(IaaS)是一种特定类型的云计算。IaaS可以被配置成通过公共网络(例如,因特网)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部件(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)等)。在一些情况下,IaaS提供商还可以提供各种服务来配合这些基础设施部件(例如,计费、监测、日志记录、安全性、负载平衡和集群等)。因此,由于这些服务可能是策略驱动的,IaaS用户可以能够实施策略来驱动负载平衡,以维护应用程序可用性和性能。

[0196] 在某些情况下,IaaS客户可以通过如因特网等广域网(WAN)来访问资源和服务,并且可以使用云提供商的服务来安装应用程序栈的其余元素。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),部署如数据库等中间件,为工作负载和备份创建存储桶,并且甚至在该VM上安装企业软件。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、解决应用程序问题、监测性能、管理灾难恢复等。

[0197] 在大多数情况下,云计算模型将需要云提供商的参与。云提供商可以是(但不一定是)专门提供(例如,供应、租赁、销售)IaaS的第三方服务提供商。实体也可能选择部署私有云,从而成为其自己的基础设施服务提供商。

[0198] 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等上的过程。其还可以包括准备服务器的过程(例如,安装库、守护进程等)。这通常由管理程序层(例如,服务器、存储装置、网络硬件和虚拟化)之下的云提供商管理。因此,客户可以负责处理(OS)、中间件和/或应用部署(例如,在(例如,可以按需启动的)自助服务虚拟机上)等。

[0199] 在一些示例中,IaaS供应可以指获取计算机或虚拟主机以供使用,并且甚至在其上安装所需的库或服务。在大多数情况下,部署不包括资源供应,并且可能需要先执行供应。

[0200] 在一些情况下,IaaS供应面临两种不同的挑战。首先是在任何事物开始运行之前供应初始的一组基础设施的初始挑战。其次是在所有事物已经供应时发展现有基础设施(例如,添加新服务、更改服务、移除服务等)的挑战。在一些情况下,这两种挑战可以通过使得基础设施的配置能够以声明方式定义来解决。换句话说,基础设施(例如,需要什么部件以及它们如何交互)可以由一个或多个配置文件来定义。因此,基础设施的整体拓扑(例如,哪些资源取决于哪些资源,以及它们如何一起工作)可以以声明方式描述。在一些情况下,一旦定义了拓扑,就可以生成创建和/或管理配置文件中描述的不同部件的工作流。

[0201] 在一些示例中,基础设施可以具有许多互连的元素。例如,可以有一个或多个虚拟私有云(VPC)(例如,可配置和/或共享的计算资源的潜在按需池),也称为核心网络。在一些示例中,还可以供应一个或多个安全组规则来定义将如何设置网络的安全性以及一个或多个虚拟机(VM)。也可以供应其他基础设施元素,如负载平衡器、数据库等。随着期望和/或添加越来越多的基础设施元素,基础设施可以逐渐发展。

[0202] 在一些情况下,可以采用连续部署技术来实现跨各种虚拟计算环境部署基础设施

代码。另外地,所描述的技术可以实现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被部署到一个或多个(但通常是许多)不同的生产环境的代码(例如,跨各种不同的地理位置,有时跨越整个世界)。然而,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,供应可以手动完成,供应工具可以用于供应资源,和/或一旦供应基础设施,部署工具就可以用于部署代码。

[0203] 图16是图示了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可以通信地耦接到安全主机租赁设备1604,所述安全主机租赁设备可以包括虚拟云网络(VCN)1606和安全主机子网1608。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,所述客户端计算设备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,GoogleGlass®头戴式显示器)、如Microsoft Windows Mobile®等运行软件和/或各种移动操作系统(如iOS、Windows电话、安卓(Android)、黑莓(BlackBerry)8、Palm OS等),并且是因特网、电子邮件、短消息服务(SMS)、Blackberry®或支持其他通信协议的。可替代地,客户端计算设备可以是通用个人计算机,通过示例的方式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各种可商购获得的UNIX®或类UNIX操作系统中的任一种的工作站计算机,包括但不限于各种GNU/Linux操作系统,例如,Google Chrome OS。可替代地或另外地,客户端计算设备可以是能够通过可以访问VCN 1606和/或因特网的网络进行通信的任何其他电子设备,如瘦客户端计算机、支持因特网的游戏系统(例如,具有或不具有Kinect®姿势输入设备的微软(Microsoft)Xbox游戏机)和/或个人消息传递设备。

[0204] VCN 1606可以包括本地对等网关(LPG)1610,所述本地对等网关可以经由包含在SSH VCN 1612中的LPG 1610通信地耦接到安全外壳(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通信地耦接到控制平面VCN 1616。而且,SSH VCN 1612可以经由LPG 1610通信地耦接到数据平面VCN 1618。控制平面VCN 1616和数据平面VCN 1618可以包含在服务租赁设备1619中,所述租赁设备可以为IaaS提供商所有和/或由其运营。

[0205] 控制平面VCN 1616可以包括充当外围网络的控制平面非军事区(DMZ)层1620(例如,公司内联网与外部网络之间的公司网络的部分)。基于DMZ的服务器可能具有有限的职责,并且有助于控制安全漏洞。另外地,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(多个)app子网1626的控制平面app层1624、可以包括(多个)数据库(DB)子网1630(例如,(多个)前端DB子网和/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(多个)LB子网1622可以通信地耦接到包含在控制平面app层1624中的(多个)app子网1626和可以包含在控制平面VCN 1616中的因特网网关1634,并且(多个)app子网1626可以通信地耦接到包含在控制平面数据层1628中的(多个)DB子网1630、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。

[0206] 控制平面VCN 1616可以包括数据平面镜像app层1640,所述数据平面镜像app层可以包括(多个)app子网1626。包含在数据平面镜像app层1640中的(多个)app子网1626可以

包括可以执行计算实例1644的虚拟网络接口控制器 (VNIC) 1642。计算实例1644可以将数据平面镜像app层1640的(多个)app子网1626通信地耦接到可以包含在数据平面app层1646中的(多个)app子网1626。

[0207] 数据平面VCN 1618可以包括数据平面app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(多个)LB子网1622,所述LB子网可以通信地耦接到数据平面app层1646的(多个)app子网1626和数据平面VCN 1618的因特网网关1634。(多个)app子网1626可以通信地耦接到数据平面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(多个)DB子网1630,所述DB子网可以通信地耦接到数据平面app层1646的(多个)app子网1626。

[0208] 控制平面VCN 1616和数据平面VCN 1618的因特网网关1634可以通信地耦接到元数据管理服务1652,所述元数据管理服务可以通信地耦接到公共因特网1654。公共因特网1654可以通信地耦接到控制平面VCN 1616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可以通信地耦接到云服务1656。

[0209] 在一些示例中,控制平面VCN 1616或数据平面VCN 1618的服务网关1636可以在不通过公共因特网1654的情况下对云服务1656进行应用编程接口 (API) 调用。从服务网关1636到云服务1656的API调用可以是单向的:服务网关1636可以对云服务1656进行API调用,并且云服务1656可以向服务网关1636发送所请求的数据。但是,云服务1656可能不会向服务网关1636发起API调用。

[0210] 在一些示例中,安全主机租赁设备1604可以直接连接至服务租赁设备1619,所述服务租赁设备可能以其他方式隔离。安全主机子网1608可以通过LPG 1610与SSH子网1614通信,所述LPG可以在以其他方式隔离的系统上实现双向通信。将安全主机子网1608连接至SSH子网1614可以使安全主机子网1608能够访问服务租赁设备1619内的其他实体。

[0211] 控制平面VCN 1616可以允许服务租赁设备1619的用户建立或以其他方式供应期望的资源。控制平面VCN 1616中供应的期望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app层1640可以经由可以包含在数据平面镜像app层1640和数据平面app层1646中的VNIC 1642与数据平面VCN 1618的数据平面app层1646通信。

[0212] 在一些示例中,系统的用户或客户可以通过公共因特网1654进行请求操作,例如创建、读取、更新或删除 (CRUD) 操作,所述公共因特网可以将请求传送到元数据管理服务1652。元数据管理服务1652可以通过因特网网关1634将该请求传送给控制平面VCN 1616。所述请求可以由包含在控制平面DMZ层1620中的(多个)LB子网1622接收。(多个)LB子网1622可以确定请求是有效的,并且响应于该确定,(多个)LB子网1622可以将请求传输到包含在控制平面app层1624中的(多个)app子网1626。如果请求被验证并且需要对公共因特网1654的调用,则对公共因特网1654的调用可以被传输到可以对公共因特网1654进行调用的NAT网关1638。可能期望请求存储的存储器可以存储在(多个)DB子网1630中。

[0213] 在一些示例中,数据平面镜像app层1640可以促进控制平面VCN 1616与数据平面VCN 1618之间的直接通信。例如,可能期望对包含在数据平面VCN 1618中的资源应用对配置的改变、更新或其他合适的修改。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,并且可以由此执行对这些资源的配置的改变、更新或其

他合适的修改。

[0214] 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设备1619中。在这种情况下,系统的用户或客户可能无法拥有或操作控制平面VCN 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设备1619中。该实施例可以能够隔离可能阻止用户或客户与其他用户或其他客户的资源进行交互的网络。而且,该实施例可以允许系统的用户或客户私有地存储数据库,而不需要依赖可能不具有期望的安全级别的公共因特网1654进行存储。

[0215] 在其他实施例中,包含在控制平面VCN 1616中的(多个)LB子网1622可以被配置成从服务网关1636接收信号。在该实施例中,控制平面VCN 1616和数据平面VCN 1618可以被配置成由IaaS提供商的客户调用,而不调用公共因特网1654。IaaS提供商的客户可能需要该实施例,因为客户使用的(多个)数据库可以由IaaS提供商控制,并且可以存储在服务租赁设备1619上,所述服务租赁设备可以与公共因特网1654隔离。

[0216] 图17是图示了根据至少一个实施例的IaaS架构的另一个示例模式的框图1700。服务运营商1702(例如,图16的服务运营商1602)可以通信地耦接到安全主机租赁设备1704(例如,图16的安全主机租赁设备1604),所述安全主机租赁设备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),所述本地对等网关可以经由包含在SSH VCN 1712中的LPG 1610通信地耦接到安全外壳(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通信地耦接到控制平面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设备1719(例如,图16的服务租赁设备1619)中,并且数据平面VCN 1718(例如,图16的数据平面VCN 1618)可以包含在用户租赁设备1721中,所述用户租赁设备可以为系统的用户或客户所有或由其操作。

[0217] 控制平面VCN 1716可以包括:可以包括(多个)LB子网1722(例如图16的(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(多个)app子网1726(例如图16的(多个)app子网1626)的控制平面app层1724(例如图16的控制平面app层1624)、可以包括(多个)数据库(DB)子网1730(例如类似于图16的(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(多个)LB子网1722可以通信地耦接到包含在控制平面app层1724中的(多个)app子网1726和可以包含在控制平面VCN 1716中的因特网网关1734(例如图16的因特网网关1634),并且(多个)app子网1726可以通信地耦接到包含在控制平面数据层1728中的(多个)DB子网1730和服务网关1736(例如图16的服务网关)以及网络地址转换(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。

[0218] 控制平面VCN 1716可以包括数据平面镜像app层1740(例如图16的数据平面镜像app层1640),所述数据平面镜像app层可以包括(多个)app子网1726。包含在数据平面镜像app层1740中的(多个)app子网1726可以包括虚拟网络接口控制器(VNIC)1742(例如1642的VNIC),所述虚拟网络接口控制器可以执行计算实例1744(例如类似于图16的计算实例

1644)。计算实例1744可以经由包含在数据平面镜像app层1740中的VNIC 1742和包含在数据平面app层1746中的VNIC 1742来促进数据平面镜像app层1740的(多个)app子网1726与可以包含在数据平面app层1746(例如图16的数据平面app层1646)中的(多个)app子网1726之间的通信。

[0219] 包含在控制平面VCN 1716中的因特网网关1734可以通信地耦接到元数据管理服务1752(例如图16的元数据管理服务1652),所述元数据管理服务可以通信地耦接到公共因特网1754(例如图16的公共因特网1654)。公共因特网1754可以通信地耦接到包含在控制平面VCN 1716中的NAT网关1738。包含在控制平面VCN 1716中的服务网关1736可以通信地耦接到云服务1756(例如图16的云服务1656)。

[0220] 在一些示例中,数据平面VCN 1718可以包含在客户租赁设备1721中。在这种情况下,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设备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设备1719中的控制平面VCN 1716与包含在客户租赁设备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设备1719中的控制平面VCN 1716中供应的资源在包含在客户租赁设备1721中的数据平面VCN 1718中部署或以其他方式使用。

[0221] 在其他示例中,IaaS提供商的客户可能具有位于客户租赁设备1721中的数据库。在该示例中,控制平面VCN 1716可以包括数据平面镜像app层1740,所述数据平面镜像app层可以包括(多个)app子网1726。数据平面镜像app层1740可以驻留在数据平面VCN 1718中,但是数据平面镜像app层1740可以不位于数据平面VCN 1718中。也就是说,数据平面镜像app层1740可以访问客户租赁设备1721,但是数据平面镜像app层1740可以不存在于数据平面VCN 1718中或者为IaaS提供商的客户所有或由其操作。数据平面镜像app层1740可以被配置成对数据平面VCN 1718进行调用,但是可以不被配置成对包含在控制平面VCN 1716中的任何实体进行调用。客户可能期望部署或以其他方式使用控制平面VCN 1716中供应的数据平面VCN 1718中的资源,并且数据平面镜像app层1740可以促进客户的资源的期望部署或其他使用。

[0222] 在一些实施例中,IaaS提供商的客户可以对数据平面VCN 1718应用过滤程序。在该实施例中,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访问公共因特网1754。IaaS提供商可能无法应用过滤程序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。由客户对包含在客户租赁设备1721中的数据平面VCN 1718应用过滤程序可以有助于将数据平面VCN 1718与其他客户和公共因特网1754隔离。

[0223] 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因特网1754上、控制平面VCN 1716上或数据平面VCN 1718上的服务。云服务1756与控制平面VCN 1716或数据平面VCN 1718之间的连接可能不是实时的或连续的。云服务1756可以存在于为IaaS提供商所有或由其运营的不同网络上。云服务1756可以被配置成接收来自服务网关1736的调用,并且可以被配置成不接收来自公共因特网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可以与可能不与控制平面VCN 1716处于同一区域中的云服务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的调用是由位于区域1中的控制平面VCN 1716中包含的服务网关1736进行的,则所述调用可以被传输到区域1中的部署16。在该示例中,控制平面VCN 1716或区域1中的部署16可以不与区域2中的部署16通信地耦接或通信。

[0224] 图18是图示了根据至少一个实施例的IaaS架构的另一个示例模式的框图1800。服务运营商1802(例如,图16的服务运营商1602)可以通信地耦接到安全主机租赁设备1804(例如,图16的安全主机租赁设备1604),所述安全主机租赁设备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),所述LPG可以经由包含在SSH VCN 1812中的LPG 1810通信地耦接到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通信地耦接到控制平面VCN 1816(例如图16的控制平面VCN 1616)并且经由包含在数据平面VCN 1818中的LPG 1810通信地耦接到数据平面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设备1819中(例如图16的服务租赁设备1619)。

[0225] 控制平面VCN 1816可以包括:可以包括(多个)负载均衡器(LB)子网1822(例如图16的(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(多个)app子网1826(例如类似于图16的(多个)app子网1626)的控制平面app层1824(例如图16的控制平面app层1624)、可以包括(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(多个)LB子网1822可以通信地耦接到包含在控制平面app层1824中的(多个)app子网1826和可以包含在控制平面VCN 1816中的因特网网关1834(例如图16的因特网网关1634),并且(多个)app子网1826可以通信地耦接到包含在控制平面数据层1828中的(多个)DB子网1830和服务网关1836(例如图16的服务网关)以及网络地址转换(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。

[0226] 数据平面VCN 1818可以包括数据平面app层1846(例如图16的数据平面app层1646)、数据平面DMZ层1848(例如图16的数据平面DMZ层1648)和数据平面数据层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(多个)LB子网1822,所述LB子网可以通信地耦接到数据平面app层1846的(多个)可信app子网1860和(多个)不可信app子网1862以及包含在数据平面VCN 1818中的因特网网关1834。(多个)可信app子网1860可以通信地耦接到包含在数据平面VCN 1818中的服务网关1836、包含在数据平面VCN 1818中的NAT网关1838以及包含在数据平面数据层1850中的(多个)DB子网1830。(多个)不可信app子网1862可以通信地耦接到包含在数据平面VCN 1818中的服务网关1836和包含在数据平面数据层1850中的(多个)DB子网1830。数据平面数据层1850可以包括(多个)DB子网1830,所述DB子网可以通信地耦接到包含在数据平面VCN 1818中的服务网关1836。

[0227] (多个)不可信app子网1862可以包括一个或多个主VNIC 1864(1) - (N),所述一个或多个主VNIC可以通信地耦接到租户虚拟机(VM) 1866(1) - (N)。每个租户VM 1866(1) - (N)可以通信地耦接到可以包含在相应容器出口VCN 1868(1) - (N)中的相应app子网1867(1) - (N),所述相应容器出口可以包含在相应客户租赁设备1870(1) - (N)中。相应次级VNIC 1872

(1) - (N) 可以促进包含在数据平面VCN 1818中的(多个)不可信app子网1862与包含在容器出口VCN 1868(1) - (N) 中的app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, 所述NAT网关可以通信地耦接到公共因特网1854(例如图16的公共因特网1654)。

[0228] 包含在控制平面VCN 1816中并且包含在数据平面VCN 1818中的因特网网关1834可以通信地耦接到元数据管理服务1852(例如图16的元数据管理系统1652), 所述元数据管理服务可以通信地耦接到公共因特网1854。公共因特网1854可以通信地耦接到包含在控制平面VCN 1816中并且包含在数据平面VCN 1818中的NAT网关1838。包含在控制平面VCN 1816中并且包含在数据平面VCN 1818中的服务网关1836可以通信地耦接到云服务1856。

[0229] 在一些示例中, 数据平面VCN 1818可以与客户租赁设备1870集成在一起。在一些情况下, 如在执行代码时可能需要支持的情况下, 这种集成对于IaaS提供商的客户来说可能是有用的或期望的。客户可以提供要运行的代码, 所述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期望的影响。响应于此, IaaS提供商可以确定是否运行由客户提供给IaaS提供商的代码。

[0230] 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, 并请求将功能附加到数据平面层app 1846。运行所述功能的代码可以在VM 1866(1) - (N) 中执行, 并且所述代码可以不被配置成在数据平面VCN 1818上的任何其他地方运行。每个VM 1866(1) - (N) 可以连接至一个客户租赁设备1870。包含在VM 1866(1) - (N) 中的相应容器1871(1) - (N) 可以被配置成运行代码。在这种情况下, 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(多个)不可信app子网1862中包含的VM 1866(1) - (N) 中), 这可以有助于防止不正确的或在其他方面不期望的代码破坏IaaS提供商的网络或破坏不同客户的网络。容器1871(1) - (N) 可以通信地耦接到客户租赁设备1870, 并且可以被配置成传输或接收来自客户租赁设备1870的数据。容器1871(1) - (N) 可以不被配置成从数据平面VCN 1818中的任何其他实体传输或接收数据。在完成运行代码后, IaaS提供商可以终止或以其他方式处置容器1871(1) - (N)。

[0231] 在一些实施例中, (多个)可信app子网1860可以运行为IaaS提供商所有或由其运营的代码。在该实施例中, (多个)可信app子网1860可以通信地耦接到(多个)DB子网1830, 并且被配置成在(多个)DB子网1830中执行CRUD操作。(多个)不可信app子网1862可以通信地耦接到(多个)DB子网1830, 但是在该实施例中, (多个)不可信app子网可以被配置成在(多个)DB子网1830中执行读取操作。可以包含在每个客户的VM 1866(1) - (N) 中并且可以运行来自客户的代码的容器1871(1) - (N) 可以不与(多个)DB子网1830通信地耦接。

[0232] 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接。在该实施例中, 控制平面VCN 1816与数据平面VCN 1818之间可能不存在直接通信。然而, 通信可以通过至少一种方法间接发生。IaaS提供商可以建立LPG 1810, 这可以促进控制平面VCN 1816与数据平面VCN 1818之间的通信。在另一个示例中, 控制平面VCN 1816或数据平面VCN 1818可以经由服务网关1836调用云服务1856。例如, 从控制平面VCN 1816调用云服务1856可以包括请求可以与数据平面VCN 1818通信的服务。

[0233] 图19是图示了根据至少一个实施例的IaaS架构的另一个示例模式的框图1900。服务运营商1902(例如, 图16的服务运营商1602)可以通信地耦接到安全主机租赁设备1904

(例如,图16的安全主机租赁设备1604),所述安全主机租赁设备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),所述LPG可以经由包含在SSH VCN 1912中的LPG 1910通信地耦接到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通信地耦接到控制平面VCN 1916(例如图16的控制平面VCN 1616)并且经由包含在数据平面VCN 1918中的LPG 1910通信地耦接到数据平面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设备1919中(例如图16的服务租赁设备1619)。

[0234] 控制平面VCN 1916可以包括:可以包括(多个)LB子网1922(例如图16的(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(多个)app子网1926(例如图16的(多个)app子网1626)的控制平面app层1924(例如图16的控制平面app层1624)、可以包括(多个)DB子网1930(例如图18的(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(多个)LB子网1922可以通信地耦接到包含在控制平面app层1924中的(多个)app子网1926和可以包含在控制平面VCN 1916中的因特网网关1934(例如图16的因特网网关1634),并且(多个)app子网1926可以通信地耦接到包含在控制平面数据层1928中的(多个)DB子网1930和服务网关1936(例如图16的服务网关)以及网络地址转换(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。

[0235] 数据平面VCN 1918可以包括数据平面app层1946(例如图16的数据平面app层1646)、数据平面DMZ层1948(例如图16的数据平面DMZ层1648)和数据平面数据层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(多个)LB子网1922,所述LB子网可以通信地耦接到数据平面app层1946的(多个)可信app子网1960(例如图18的(多个)可信app子网1860)和(多个)不可信app子网1962(例如图18的(多个)不可信app子网1862)以及包含在数据平面VCN 1918中的因特网网关1934。(多个)可信app子网1960可以通信地耦接到包含在数据平面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(多个)DB子网1930。(多个)不可信app子网1962可以通信地耦接到包含在数据平面VCN 1918中的服务网关1936和包含在数据平面数据层1950中的(多个)DB子网1930。数据平面数据层1950可以包括(多个)DB子网1930,所述DB子网可以通信地耦接到包含在数据平面VCN 1918中的服务网关1936。

[0236] (多个)不可信app子网1962可以包括主VNIC 1964(1)-(N),所述主VNIC可以通信地耦接到驻留在(多个)不可信app子网1962内的租户虚拟机(VM) 1966(1)-(N)。每个租户VM 1966(1)-(N)可以在相应容器1967(1)-(N)中运行代码,并且通信地耦接到可以包含在数据平面app层1946中的app子网1926,所述数据平面app层可以包含在容器出口VCN 1968中。相应次级VNIC 1972(1)-(N)可以促进包含在数据平面VCN 1918中的(多个)不可信app子网1962与包含在容器出口VCN 1968中的app子网之间的通信。容器出口VCN可以包括NAT网关1938,所述NAT网关可以通信地耦接到公共因特网1954(例如图16的公共因特网1654)。

[0237] 包含在控制平面VCN 1916中并且包含在数据平面VCN 1918中的因特网网关1934可以通信地耦接到元数据管理服务1952(例如图16的元数据管理系统1652),所述元数据管

理服务可以通信地耦接到公共因特网1954。公共因特网1954可以通信地耦接到包含在控制平面VCN 1916中并且包含在数据平面VCN 1918中的NAT网关1938。包含在控制平面VCN 1916并且包含在数据平面VCN 1918中的服务网关1936可以通信地耦接到云服务1956。

[0238] 在一些示例中,图19的框图1900的架构所图示的模式可以被认为是图18的框图1800的架构所图示的模式的例外,并且如果IaaS提供商不能直接与客户(例如,断开连接的区域)通信,则所述模式可能是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)可以被客户实时访问。容器1967(1)-(N)可以被配置成对包含在数据平面app层1946的(多个)app子网1926中的相应次级VNIC 1972(1)-(N)进行调用,所述数据平面app层可以包含在容器出口VCN 1968中。次级VNIC 1972(1)-(N)可以将调用传输到NAT网关1938,所述NAT网关可以将调用传输到公共因特网1954。在该示例中,可以被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)还可以与来自其他客户的资源隔离。

[0239] 在其他示例中,客户可以使用容器1967(1)-(N)来调用云服务1956。在该示例中,客户可以运行容器1967(1)-(N)中请求来自云服务1956的服务的代码。容器1967(1)-(N)可以将该请求传输到次级VNIC 1972(1)-(N),所述次级VNIC可以将所述请求传输到NAT网关,所述NAT网关可以将所述请求传输到公共因特网1954。公共因特网1954可以经由因特网网关1934将所述请求传输到包含在控制平面VCN 1916中的(多个)LB子网1922。响应于确定请求是有效的,(多个)LB子网可以将请求传输到(多个)app子网1926,所述app子网可以经由服务网关1936将请求传输到云服务1956。

[0240] 应了解,图中所描绘的IaaS架构1600、1700、1800、1900可以具有除了所描绘的部件之外的其他部件。进一步地,图中所示的实施例只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部件、可以组合两个或更多个部件或可以具有不同的部件配置或布置。

[0241] 在某些实施例中,本文所述的IaaS系统可以包括以自助、基于订阅、可弹性扩展、可靠、高度可用和安全的方式传递给客户的一套应用程序、中间件和数据库服务产品。这种IaaS系统的示例是本受让人所提供的甲骨文云基础设施(OCI)。

[0242] 图20图示了各种实施例可以在其中实施的示例计算机系统2000。系统2000可以用于实施上述计算机系统任何一种。如图所示,计算机系统2000包括经由总线子系统2002与多个外围子系统通信的处理单元2004。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。

[0243] 总线子系统2002提供用于使计算机系统2000的各个部件和子系统按预期彼此通信的机构。虽然总线子系统2002被示意性地示出为单个总线,但是总线子系统的替代性实施例可以利用多个总线。总线子系统2002可以是几种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、外围总线以及使用各种总线架构中的任何一种的本地总线。例如,这种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)局部总线以及外围部件互连(PCI)总线(所述外围部件互连总线可以被实施为根据IEEE P1386.1标准制造的夹层(Mezzanine)总线)。

[0244] 可以被实施为一个或多个集成电路(例如,传统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。处理单元2004中可以包括一个或多个处理器。这些处理器可以包括单核处理器或多核处理器。在某些实施例中,处理单元2004可以被实施为一个或多个独立的处理单元2032和/或2034,每个处理单元中包括单核处理器或多核处理器。在其他实施例中,处理单元2004还可以被实施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。

[0245] 在各种实施例中,处理单元2004可以响应于程序代码而执行各种程序,并且可以维护多个同时执行的程序或过程。在任何给定时间,要执行的程序代码中的一些或全部可以驻留在(多个)处理器2004和/或存储子系统2018中。通过适当的编程,(多个)处理器2004可以提供上述各种功能。计算机系统2000可以另外包括处理加速单元2006,所述处理加速单元可以包括数字信号处理器(DSP)、专用处理器等。

[0246] 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、如鼠标或轨迹球等指向设备、合并到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、具有话音命令识别系统的音频输入设备、麦克风以及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/或姿势识别设备,如Microsoft Kinect®运动传感器,其使得用户能够通过使用姿势和口头命令的自然用户接口来控制如Microsoft Xbox®360游戏控制器等输入设备并与其交互。用户接口输入设备还可以包括眼部姿势识别设备,如检测来自用户的眼部活动(例如,在拍照和/或进行菜单选择时‘眨眼’)并将眼部姿势变换为到输入设备(例如,Google Glass®)的输入的Google Glass®眨眼检测器。另外地,用户接口输入设备可以包括使得用户能够通过话音命令与话音识别系统(例如,Siri®导航器)交互的话音识别感测设备。

[0247] 用户接口输入设备还可以包括但不限于三维(3D)鼠标、操纵杆或指向杆、游戏手柄和图形板、以及音频/视觉设备(如扬声器、数码相机、数码摄像机、便携式媒体播放器、网络摄像机、图像扫描仪、指纹扫描仪、条形码读取器3D扫描仪、3D打印机、激光测距仪、以及眼睛注视跟踪设备)。另外地,用户接口输入设备可以包括例如医学成像输入设备,如计算机断层扫描、磁共振成像、正电子发射断层扫描、医学超声检查设备。用户接口输入设备还可以包括例如音频输入设备,如MIDI键盘、数码乐器等。

[0248] 用户接口输出设备可以包括显示子系统、指示灯或如音频输出设备等非视觉显示器。显示子系统可以是阴极射线管(CRT)、平板设备(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,使用术语“输出设备”旨在包括用于从计算机系统2000向用户或其他计算机输出信息的所有可能类型的设备和机构。例如,用户接口输出设备可以包括但不限于在视觉上传达文本、图形和音频/视频信息的各种显示设备,如监视器、打印机、扬声器、头戴式耳机、汽车导航系统、绘图仪、话音输出设备和调制解调器。

[0249] 计算机系统2000可以包括存储子系统2018,所述存储子系统包括被示出为当前定位于系统存储器2010内的软件元件。系统存储器2010可以存储可在处理单元2004上加载和执行的程序指令以及在这些程序执行期间生成的数据。

[0250] 取决于计算机系统2000的配置和类型,系统存储器2010可以是易失性的(如随机存取存储器(RAM))和/或非易失性的(如只读存储器(ROM)、闪速存储器等)。RAM通常包含可立即访问和/或当前由处理单元2004操作和执行的程序和/或程序模块。在一些实施方式

中,系统存储器2010可以包括多种不同类型的存储器,如静态随机存取存储器(SRAM)或动态随机存取存储器(DRAM)。在一些实施方式中,包含在启动期间帮助计算机系统2000内的元件之间传送信息的基本例程的基本输入/输出系统(BIOS)通常可以存储在ROM中。通过示例而非限制的方式,系统存储器2010还图示了应用程序2012(所述应用程序可以包括客户端应用程序、Web浏览器、中间层应用程序、关系数据库管理系统(RDBMS)等)、程序数据2014和操作系统2016。通过示例的方式,操作系统2016可以包括各种版本的Microsoft Windows®、Apple Macintosh®和/或Linux操作系统、各种可商购获得的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/或如iOS、Windows®电话、Android® OS、BlackBerry® 200S和Palm® OS操作系统等移动操作系统。

[0251] 存储子系统2018还可以提供用于存储提供一些实施例的功能的基本编程和数据构造的有形计算机可读存储介质。当由处理器执行时提供上述功能的软件(程序、代码模块、指令)可以存储在存储子系统2018中。这些软件模块或指令可以由处理单元2004执行。存储子系统2018还可以根据本公开提供用于存储所使用的数据的储存库。

[0252] 存储子系统2000还可以包括可以进一步连接至计算机可读存储介质2022的计算机可读存储介质读取器2020。与系统存储器2010一起并且可选地结合的计算机可读存储介质2022可以全面地表示远程、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、传输和取得计算机可读信息的存储介质。

[0253] 包含代码或代码的部分的计算机可读存储介质2022还可以包括本领域已知或使用的任何适当的介质,所述介质包括存储介质和通信介质,如但不限于以用于存储和/或传输信息的任何方法或技术实施的易失性和非易失性、可移动和不可移动介质。所述介质可以包括有形计算机可读存储介质,如RAM、ROM、电可擦除可编程ROM(EEPROM)、闪速存储器或其他存储器技术、CD-ROM、数字通用盘(DVD)或其他光存储装置、磁带盒、磁带、磁盘存储装置或其他磁存储设备,或其他有形计算机可读介质。所述介质还可以包括非有形计算机可读介质,如数据信号、数据传输或可以用于传输期望的信息并且可以由计算系统2000访问的任何其他介质。

[0254] 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器、从可移动非易失性磁介质读取或向其写入的磁盘驱动器,以及从可移动非易失性光盘(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器。计算机可读存储介质2022可以包括但不限于Zip®驱动器、闪速存储器卡、通用串行总线(USB)闪速存储器驱动器、安全数字(SD)卡、DVD盘、数字录像带等。计算机可读存储介质2022还可以包括如基于闪速存储器的固态驱动器(SSD)、企业级闪速存储器驱动器、固态ROM等基于非易失性存储器的SSD、基于如固态RAM、动态RAM、静态RAM等易失性存储器的SSD、基于DRAM的SSD、磁阻RAM(MRAM) SSD以及使用DRAM和基于闪速存储器的SSD的组合的混合SSD。磁盘驱动器和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。

[0255] 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用作用于从其他系统接收数据并且从计算机系统2000向其他系统传输数据的接口。例如,通信子系统2024可以使得计算机系统2000能够经由因特网连接至一个或多个设备。在一些实施例

中,通信子系统2024可以包括用于访问无线声音和/或数据网络的射频(RF)收发器部件(例如,使用蜂窝电话技术、如3G、4G或EDGE(全球演进增强型数据速率)等先进的数据网络技术、WiFi(IEEE 802.11家庭标准或其他移动通信技术或其任何组合)、全球定位系统(GPS)接收器部件和/或其他部件。在一些实施例中,除了无线接口之外或替代无线接口,通信子系统2024可以提供有线网络连通性(例如,以太网)。

[0256] 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的一个或多个用户接收结构化和/或非结构化数据馈送2026、事件流2028、事件更新2030等形式的输入通信。

[0257] 举例来说,通信子系统2024可以被配置成从社交网络和/或其他通信服务的用户实时地接收数据馈送2026(如Twitter®馈送、Facebook®更新、web馈送(如丰富站点摘要(RSS)馈送))和/或来自一个或多个第三方信息资源的实时更新。

[0258] 另外地,通信子系统2024还可以被配置成接收连续数据流形式的数据,所述连续数据流可以包括(可以没有显式结束的本质连续的或无界的)实时事件的事件流2028和/或事件更新2030。生成连续数据的应用程序的示例可以包括例如传感器数据应用程序、财务收报机、网络性能测量工具(例如,网络监测和流量管理应用程序)、点击流分析工具、汽车交通监测等。

[0259] 通信子系统2024还可以被配置成将结构化和/或非结构化数据馈送2026、事件流2028、事件更新2030等输出到可以与耦接到计算机系统2000的一个或多个流数据源计算机通信的一个或多个数据库。

[0260] 计算机系统2000可以是各种类型之一,包括手持便携式设备(例如,iPhone®蜂窝电话、iPad®计算平板电脑、PDA)、可穿戴设备(例如,GoogleGlass®头戴式显示器)、PC、工作站、主机、自助服务终端、服务器机架或任何其他数据处理系统。

[0261] 在前述描述中,出于解释的目的,阐述了具体细节以提供对本公开的示例的透彻理解。然而,将显而易见的是,可以在没有这些具体细节的情况下实践各种示例。随后的描述仅提供了示例并且不旨在限制本公开的范围、适用性或配置。而是,随后对示例的描述将为本领域技术人员提供用于实施示例的使能描述。应当理解,在不脱离所附权利要求中阐述的本公开的精神和范围的情况下,可以对元件的功能和布置进行各种改变。附图和描述不旨在是限制性的。电路、系统、网络、过程和其他部件可以以框图形式示出为部件,以免在不必要的细节上模糊示例。在其他实例中,可以示出公知的电路、过程、算法、结构和技术,而没有不必要的细节,以避免模糊示例。本文所公开的传授内容还可以应用于如移动应用程序、非移动应用程序、桌面应用程序、web应用程序、企业应用程序等各种类型的应用程序。进一步地,本公开的传授内容不限于特定的操作环境(例如,操作系统、设备、平台等),相反,可以应用于多个不同的操作环境。

[0262] 此外,应注意,可以将单独的示例描述为被描绘成流程图、流图、数据流图、结构图或框图的过程。虽然流程图可以将操作描述为顺序过程,但是许多操作可以并行地或同时地执行。另外,可以重新布置操作的顺序。当过程的操作完成时,所述过程终止,但是所述过程可以具有图中没有包括的另外的步骤。过程可以对应于方法、函数、进程、子例程、子程序等。当过程对应于函数时,过程的终止可以对应于所述函数返回到调用函数或主函数。

[0263] 词语“示例”和“示例性”在本文中用于意指“用作示例、实例或说明”。在本文中被

描述为“示例性”或“示例”的任何实施例或设计不必被解释为比其他实施例或设计更优选或有利。

[0264] 术语“机器可读存储介质”或“计算机可读存储介质”包括但不限于便携式或非便携式存储设备、光学存储设备和能够存储、包含或携带(多个)指令和/或数据的各种其他介质。机器可读存储介质或计算机可读存储介质可以包括非暂态介质,数据可以存储在所述非暂态介质中并且不包括无线地或通过有线连接传播的载波和/或暂态电子信号。非暂态介质的示例可以包括但不限于磁盘或磁带、如光盘(CD)或数字通用盘(DVD)等光学存储介质、闪速存储器或存储器或存储器设备。计算机程序产品可以包括代码和/或机器可执行指令,所述代码和/或机器可执行指令可以表示进程、函数、子程序、程序、例程、子例程、模块、软件包、类或者指令、数据结构或程序语句的任何组合。代码段可以通过传递和/或接收信息、数据、自变量、参数或存储器内容来耦接到另一个代码段或硬件电路。信息、自变量、参数、数据等可以通过任何适合的手段(包括存储器共享、消息传递、令牌传递、网络传输等)传递、转发或传输。

[0265] 此外,可以由硬件、软件、固件、中间件、微代码、硬件描述语言或其任何组合来实施示例。当在软件、固件、中间件或微代码中实施时,用于执行必要任务的程序代码或代码段(例如,计算机程序产品)可以存储在机器可读介质中。(多个)处理器可以执行必要任务。可以以各种配置提供一些图中描绘的系统。在一些示例中,可以将系统配置成分布式系统,其中,系统的一个或多个部件跨一个或多个网络分布在云计算系统中。在将部件描述为“被配置成”执行某些操作的情况下,这种配置可以通过例如将电子电路或其他硬件设计成执行该操作、通过对电子电路(例如,微处理器或其他适合的电子电路)进行编程或控制以执行操作或其任何组合来完成。

[0266] 尽管已经描述了本公开的具体实施例,但是各种修改、更改、替代构造和等同物也涵盖在本公开的范围。本公开的实施例不局限于在某些具体数据处理环境中的操作,而是自由地在多个数据处理环境中操作。另外地,尽管已经使用特定的一系列处理和步骤描述了本公开的实施例,但是对于本领域技术人员来说显而易见的是,本公开的范围不限于所描述的一系列处理和步骤。可以单独地或联合地使用以上所述的实施例的各种特征和方面。

[0267] 进一步地,虽然已经使用硬件和软件的特定组合描述了本公开的实施例,但是应当认识到,硬件和软件的其他组合也在本公开的范围。本公开的实施例可以仅以硬件或仅以软件或使用其组合实施。本文所描述的各种过程可以在相同处理器或不同处理器上以任何组合实施。因此,在将部件或模块描述为被配置成执行某些操作的情况下,这种配置可以例如通过将电子电路设计成执行操作、通过对可编程电子电路(如微处理器)进行编程以执行操作或其任何组合来完成。过程可以使用包括但不限于用于过程间通信的传统技术的各种技术进行通信,并且不同的过程对可以使用不同的技术,或者相同的过程对可以在不同的时间使用不同的技术。

[0268] 因此,应当从说明性而非限制性意义上看待本说明书和附图。然而,将明显的是,在不脱离权利要求中阐述的更广泛的精神和范围的情况下,可以对其作出添加、减少、删除以及其他修改和改变。因此,虽然已经描述了具体的公开实施例,但是这些实施例不旨在是限制性的。各种修改和等同物均在以下权利要求的范围内。

[0269] 除非本文中另有指示或明显与上下文相矛盾,否则在描述所公开的实施例的上下文中(特别是在以下权利要求的上下文中)使用的术语“一个/种(a/an)”和“所述(the)”以及类似的指代词应被解释为涵盖单数和复数两者。除非另外指出,否则术语“包括(comprising)”、“具有(having)”、“包含(including)”和“包含(containing)”应被解释为开放式术语(即,意指“包含但不限于”)。术语“连接”应被解释为部分或全部地包含在内、附接到或接合在一起,即使在存在中间物的情况下也是如此。除非本文中另有指示,否则对本文中的值的范围的叙述仅旨在用作单独地提及落入所述范围内的每个单独值的速记方法,并且每个单独值并入本说明书中,如同所述单独值在本文中单独地叙述一样。除非本文中另有指示或明显与上下文相矛盾,否则本文所描述的所有方法均可以以任何合适的顺序执行。除非另外声明,否则本文提供的任何和所有实例或示范性语言(例如,“如”)的使用仅旨在更好地说明本公开的实施例,而不对本公开的范围构成限制。本说明书中的任何语言都不应被解释为将任何未要求保护的要素指示为是实践本公开所必不可少的。

[0270] 除非另有明确说明,否则如短语“X、Y或Z中的至少一个”等析取语言旨在如通常用于呈现项目、术语等的上下文中进行理解,其可以是X、Y或Z或其任何组合(例如,X、Y和/或Z)。因此,此类析取语言通常不旨在并且不应暗示某些实施例要求X中的至少一个、Y中的至少一个或Z中的至少一个各自都存在。

[0271] 本文中描述了本公开的优选实施例,包括已知用于实施本公开的最佳模式。在阅读前文描述之后,那些优选实施例的变化对于本领域普通技术人员可以变得显而易见。这些普通技术人员应当适当地采用这些变化,并且可以以不同于本文具体描述的方式来实践本公开。因此,在适用法律允许的情况下,本公开包含所附权利要求中叙述的主题的所有修改和等同物。此外,本公开涵盖本公开的所有可能变体中的上述元件的任何组合,除非本文中另外指明。

[0272] 本文引用的所有参考文献(包含出版物、专利申请和专利)均通过引用的方式并入本文,其程度如同每篇参考文献被单独并且具体地指出通过引用的方式并入并且在本文中被整体阐述。

[0273] 在前述说明书中,参考本公开的具体实施例描述了本公开的各方面,但是本领域技术人员将认识到,本公开并不限于此。可以单独地或联合地使用上文所描述的公开的各种特征和方面。进一步地,在不脱离说明书的更广泛的精神和范围的情况下,可以在除了本文所描述的那些环境和应用程序外的任何数量的环境和应用环境中利用实施例。因此,说明书和附图应被视为说明性的而非限制性的。

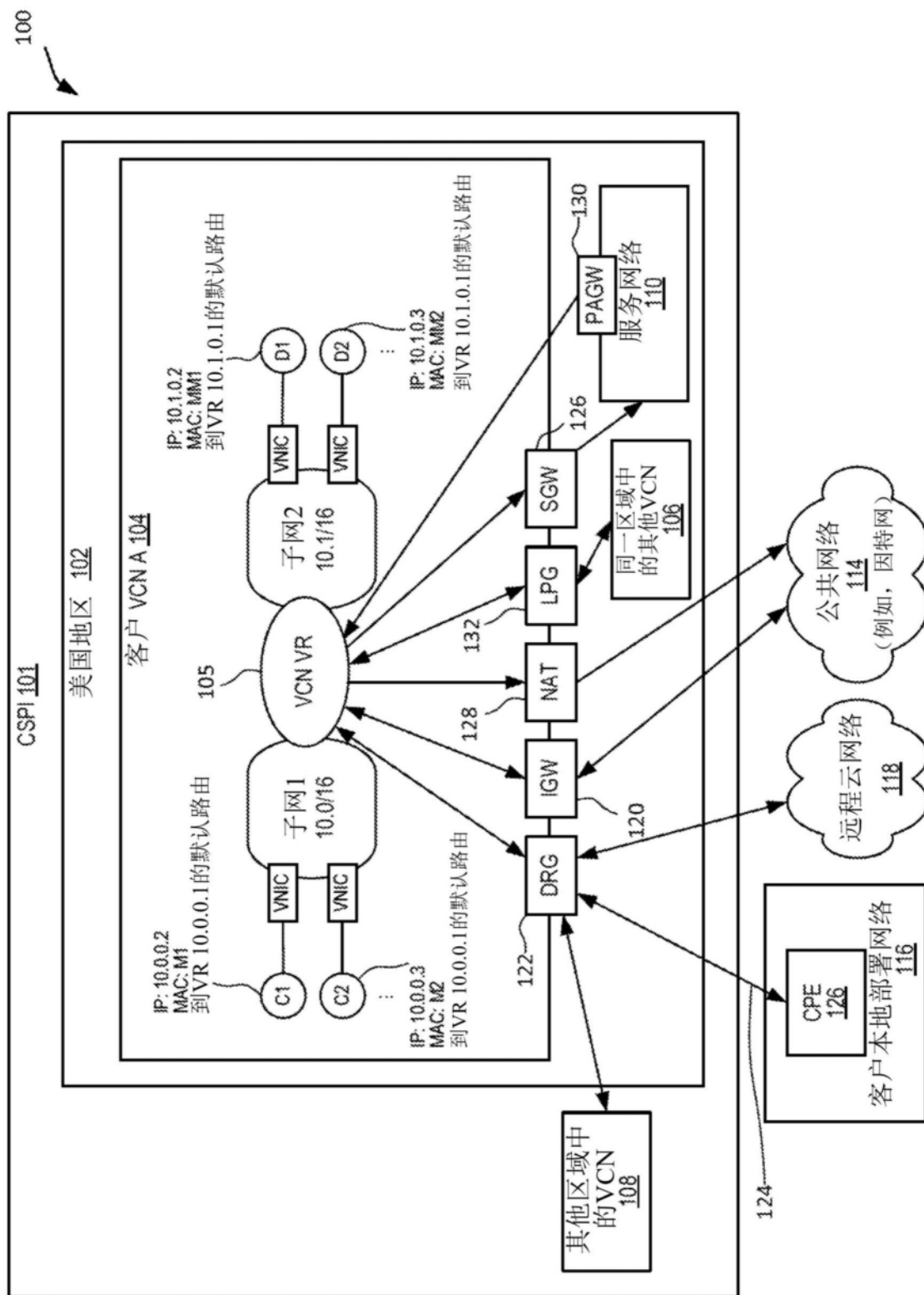


图1

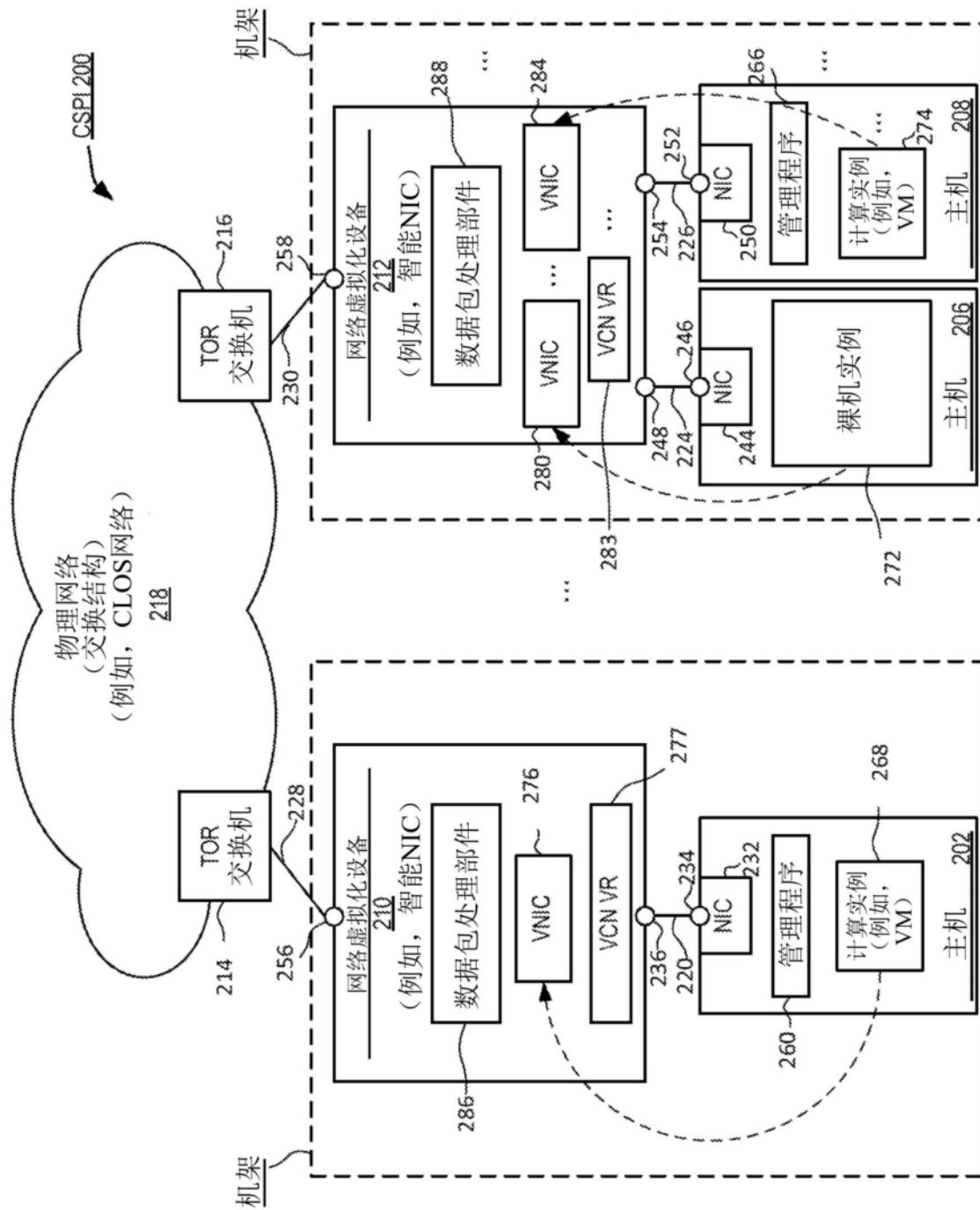


图2

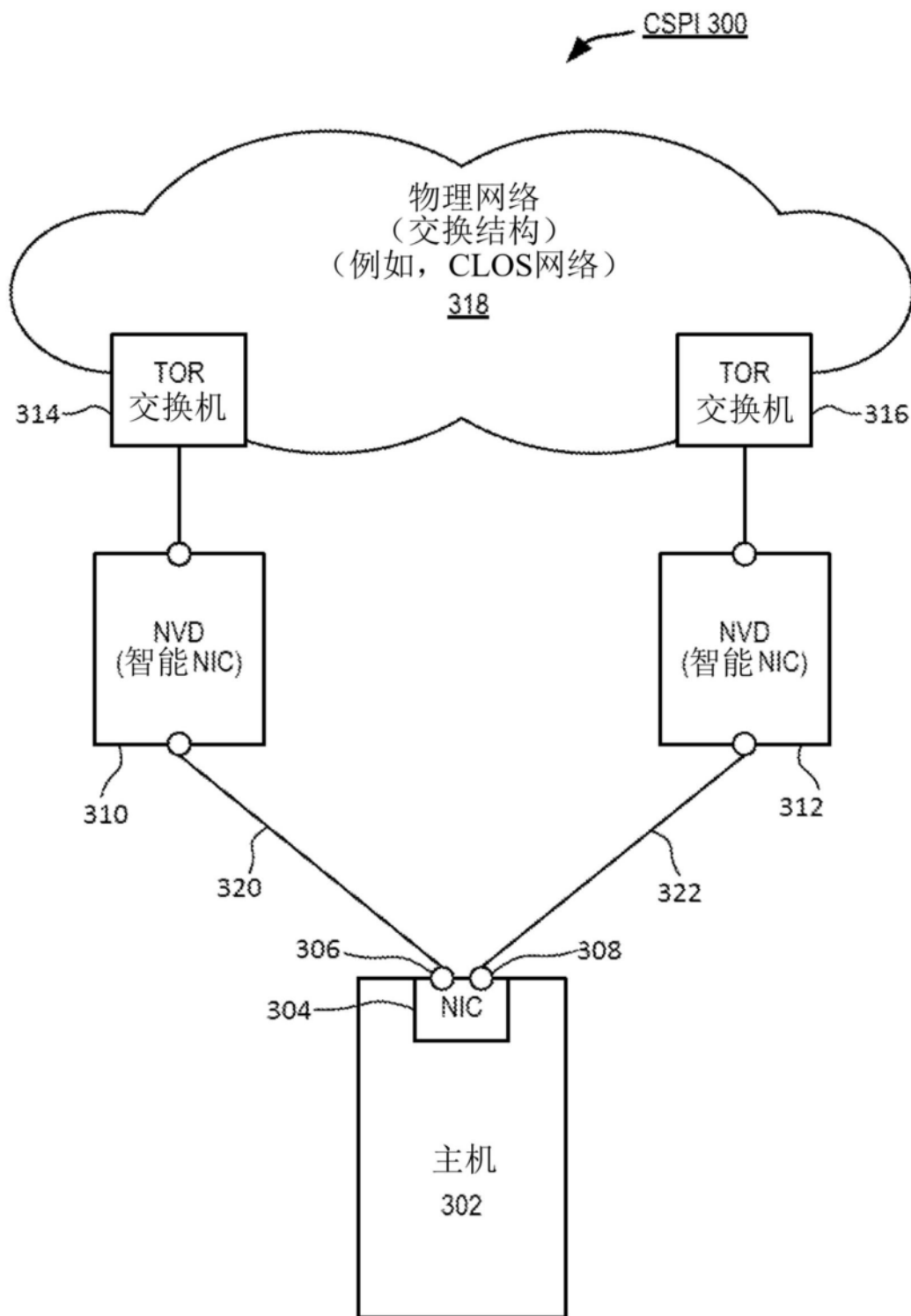


图3

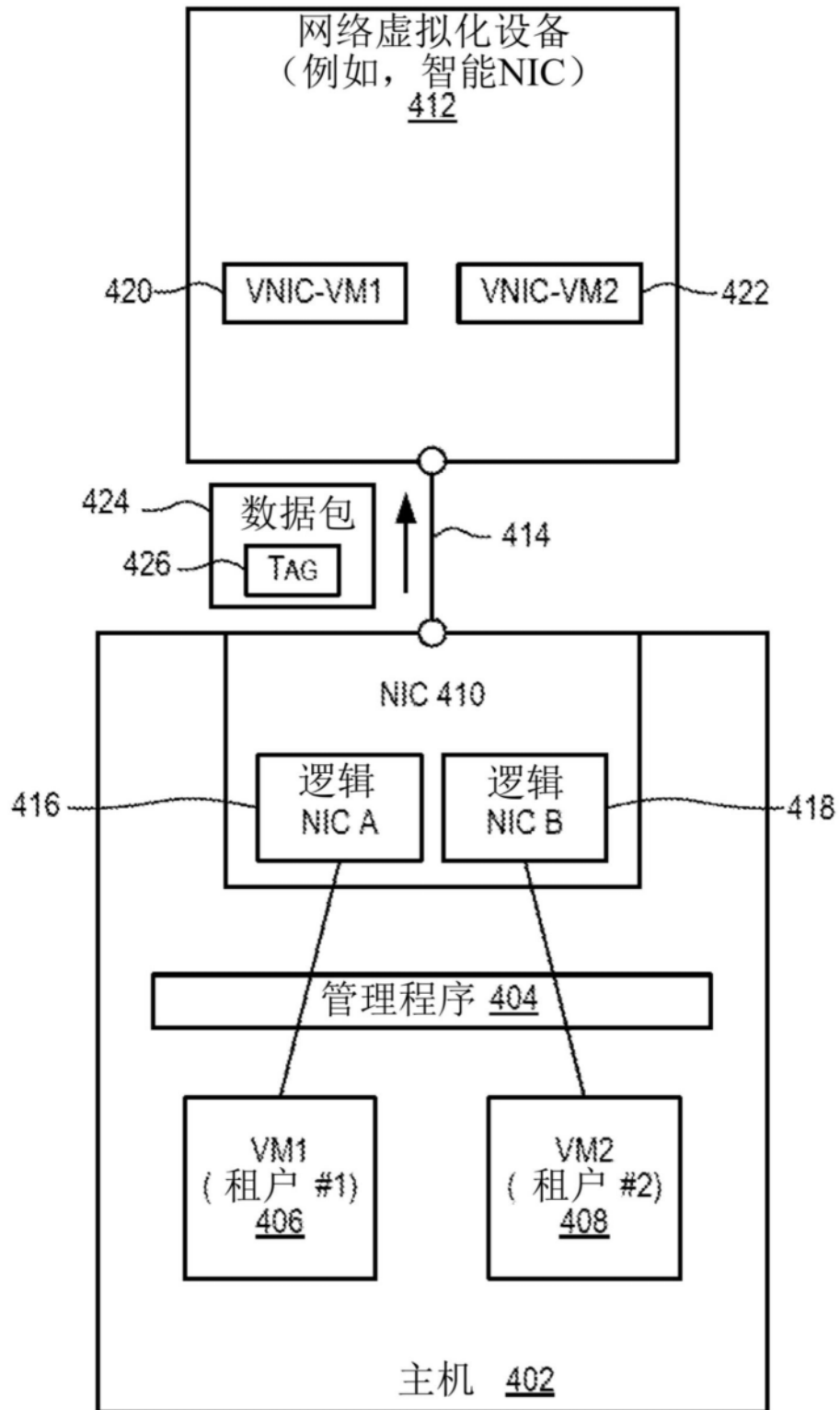


图4

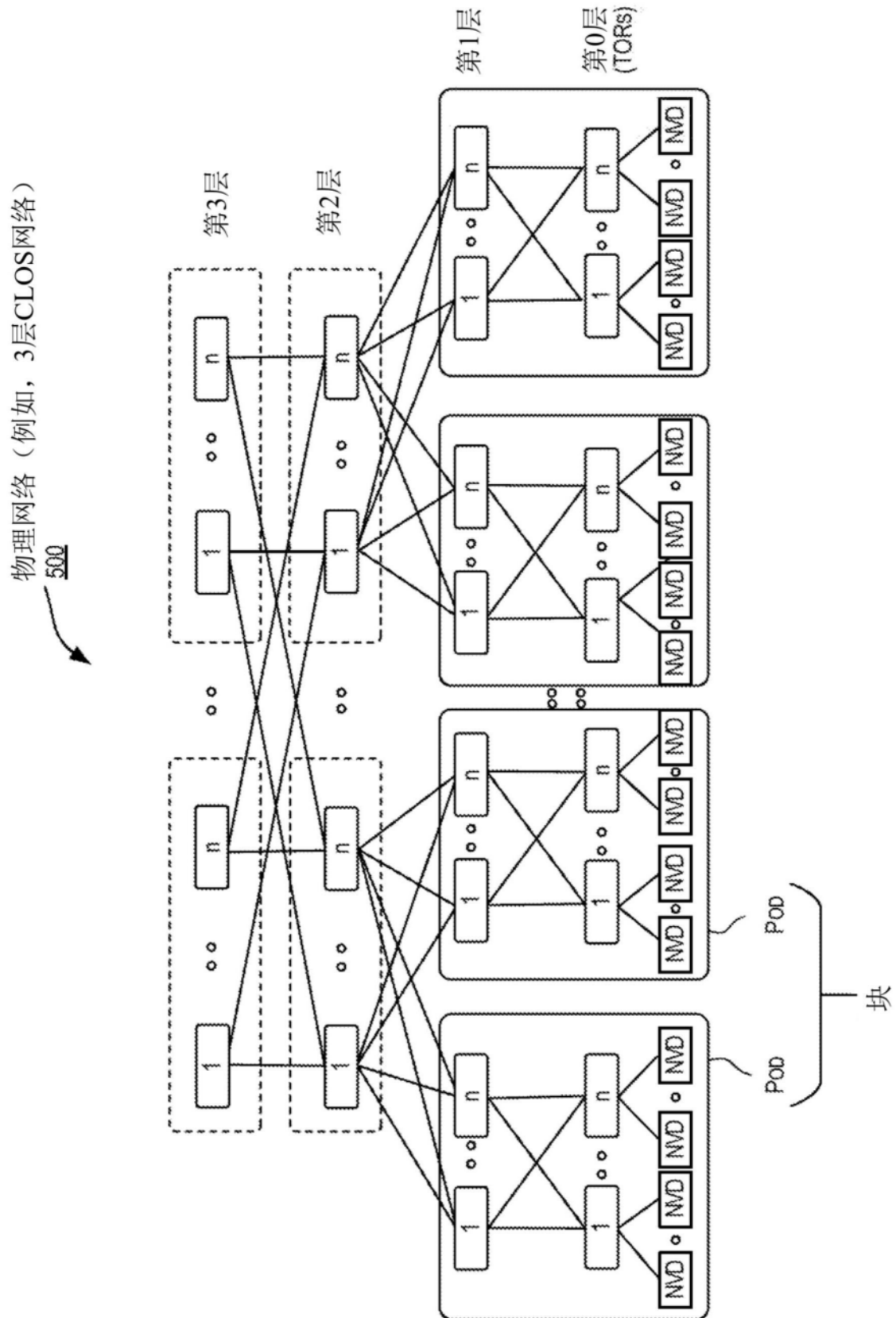


图5

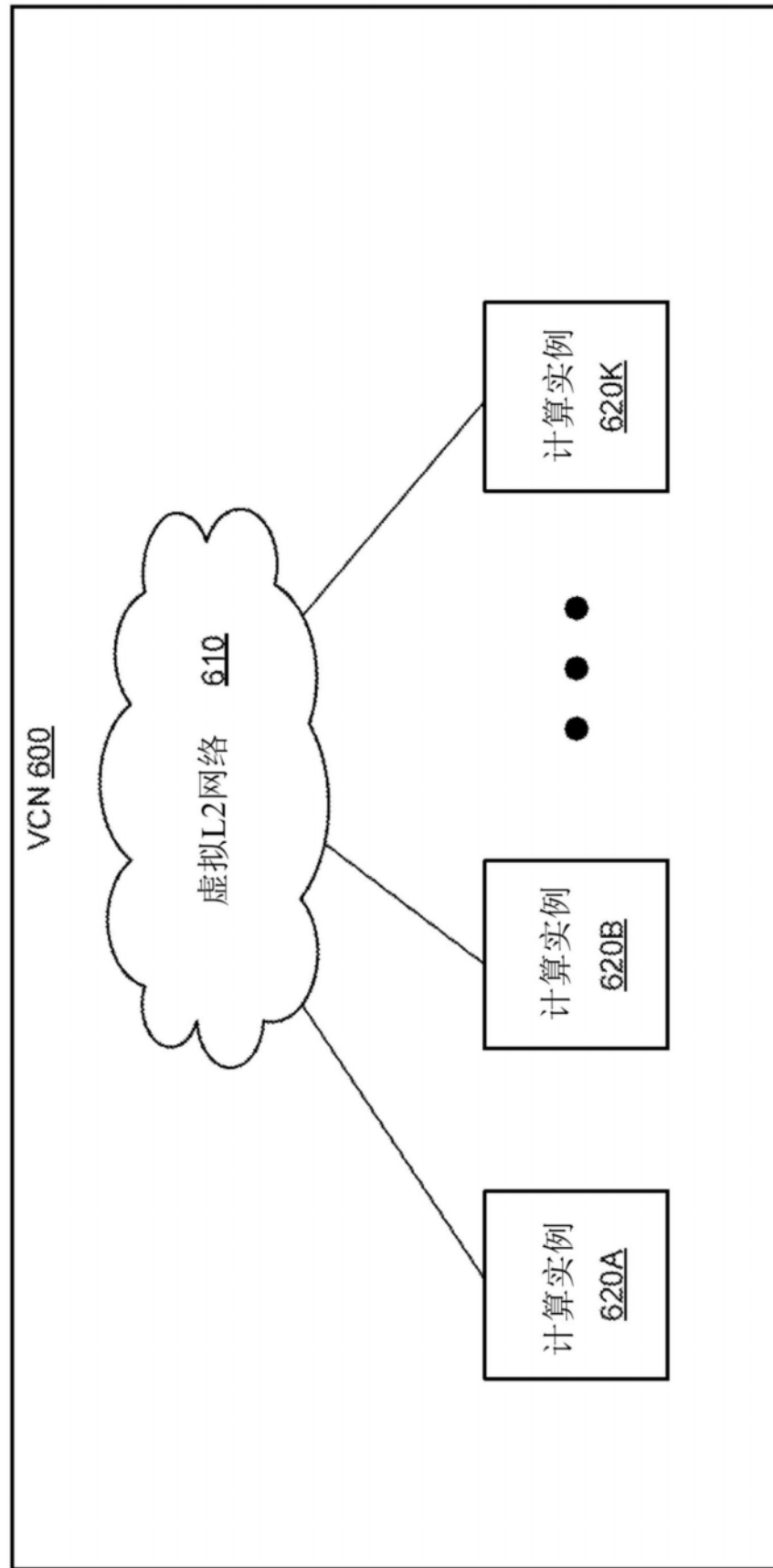


图6

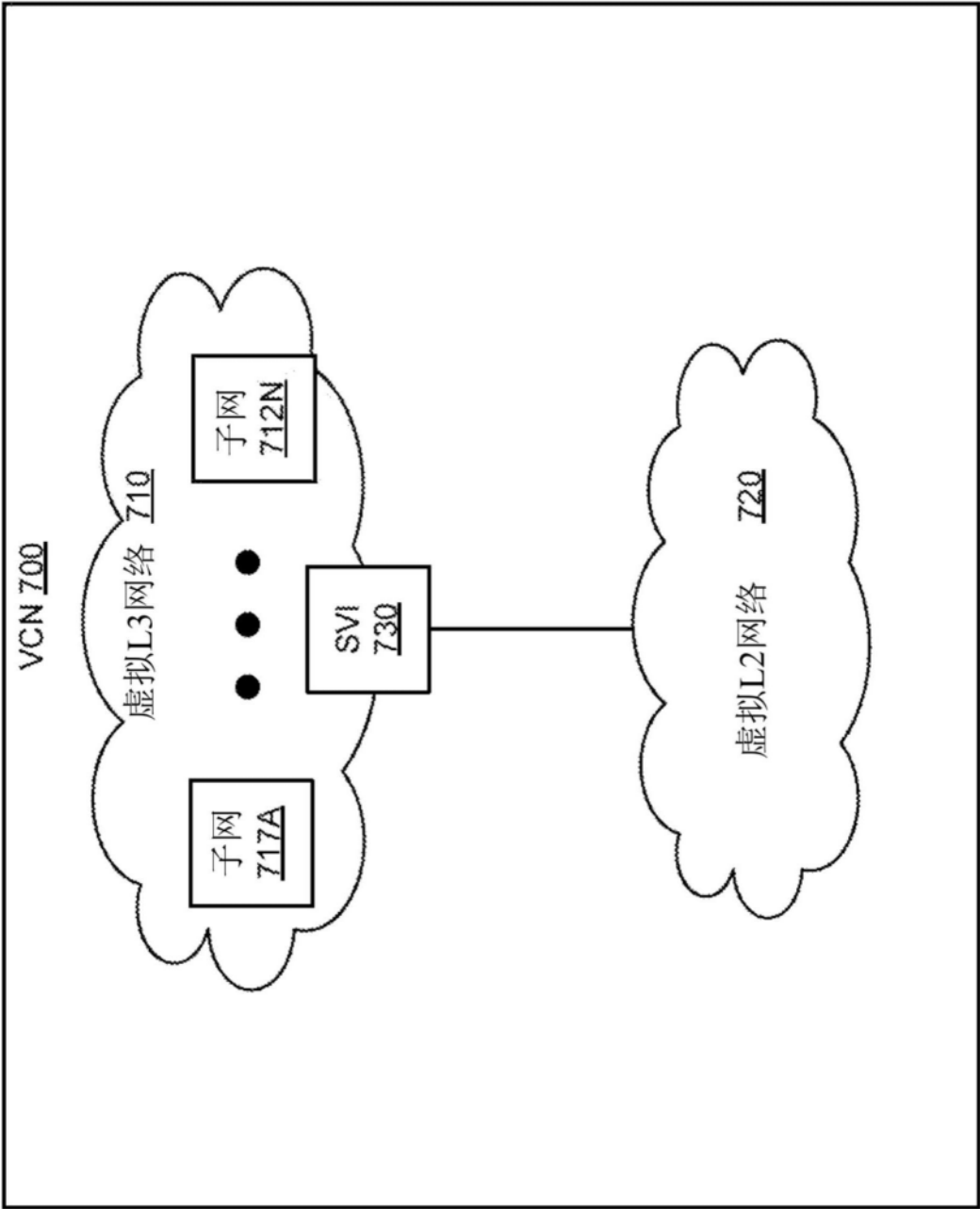


图7

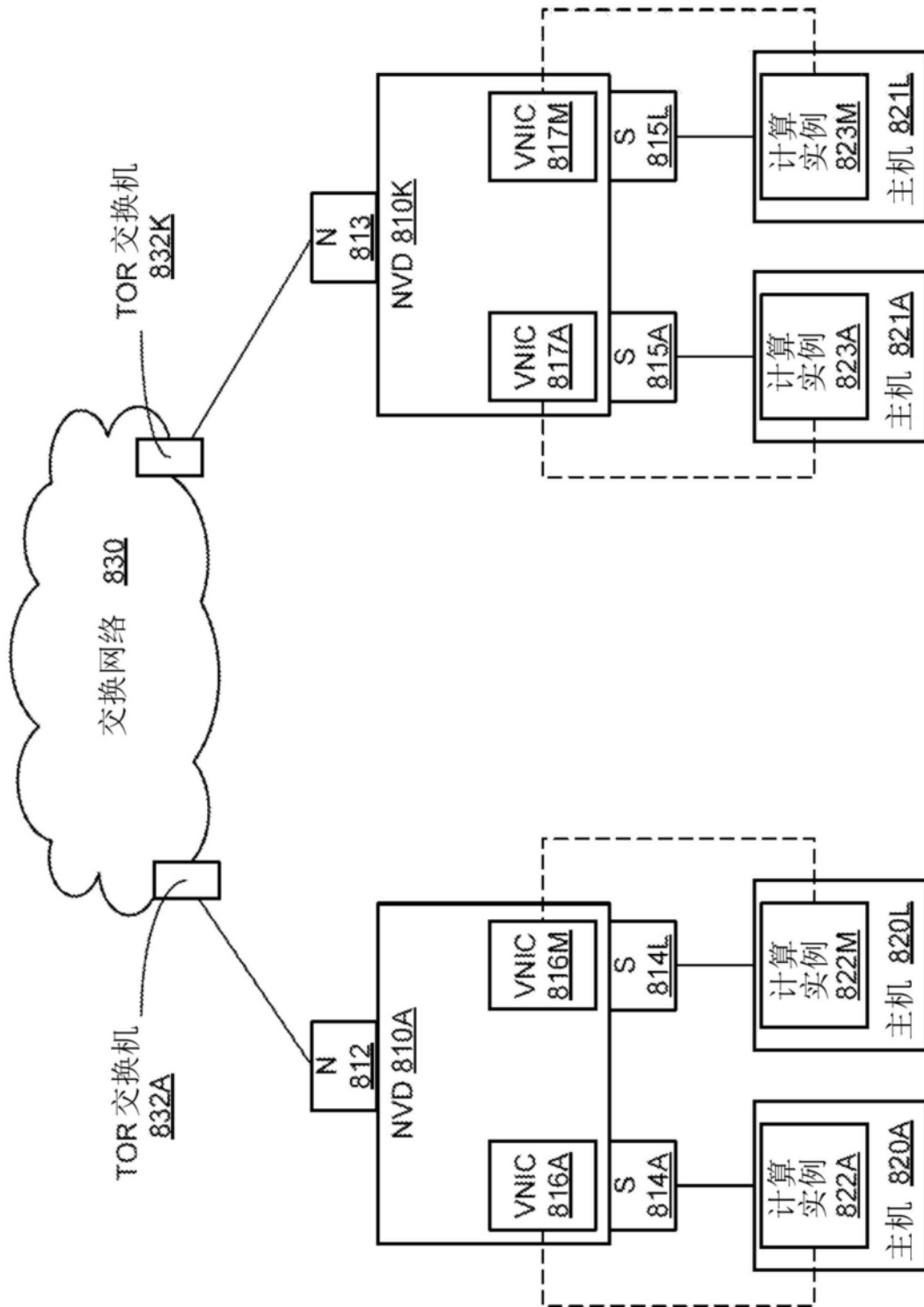


图8

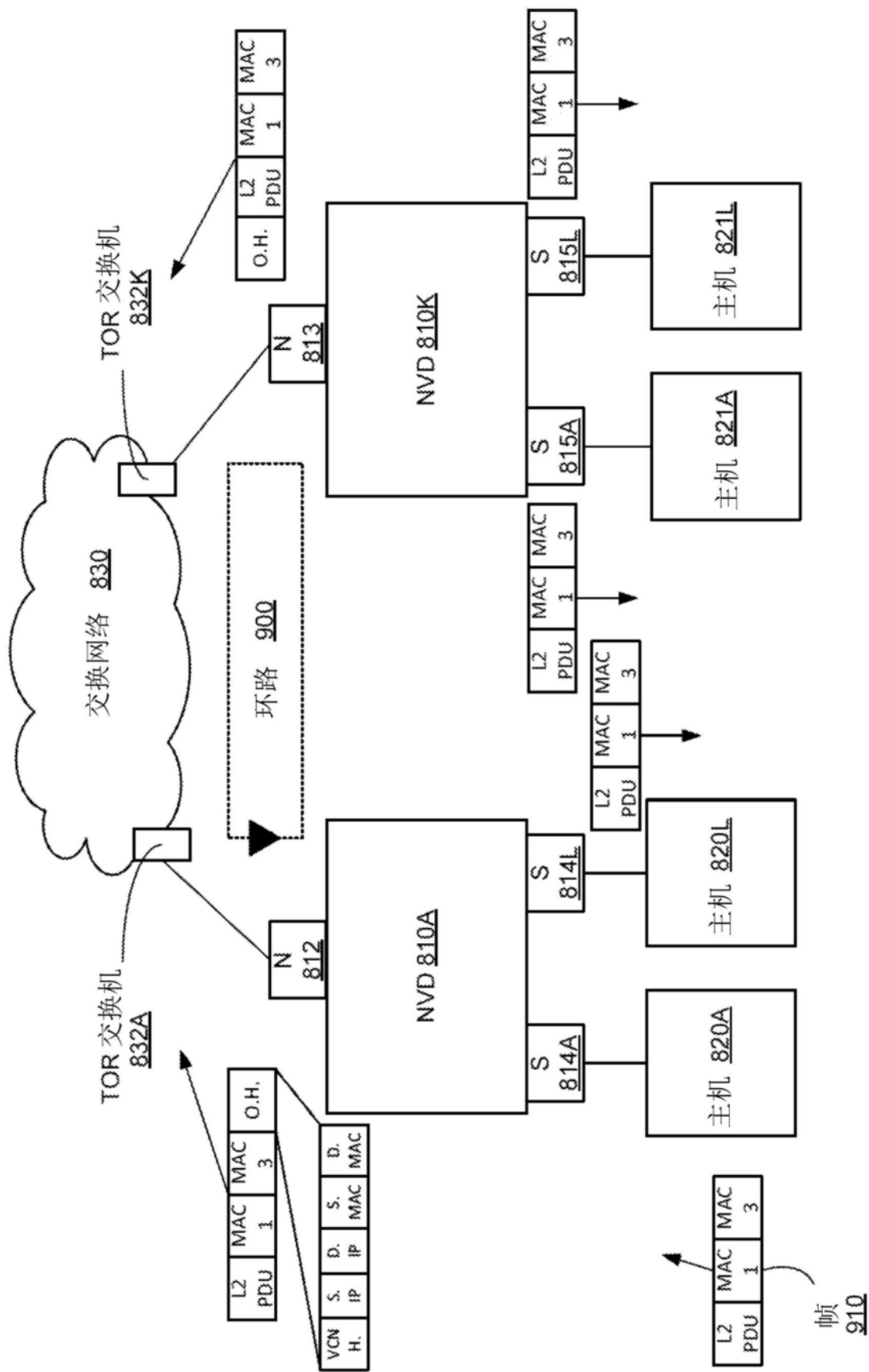


图9

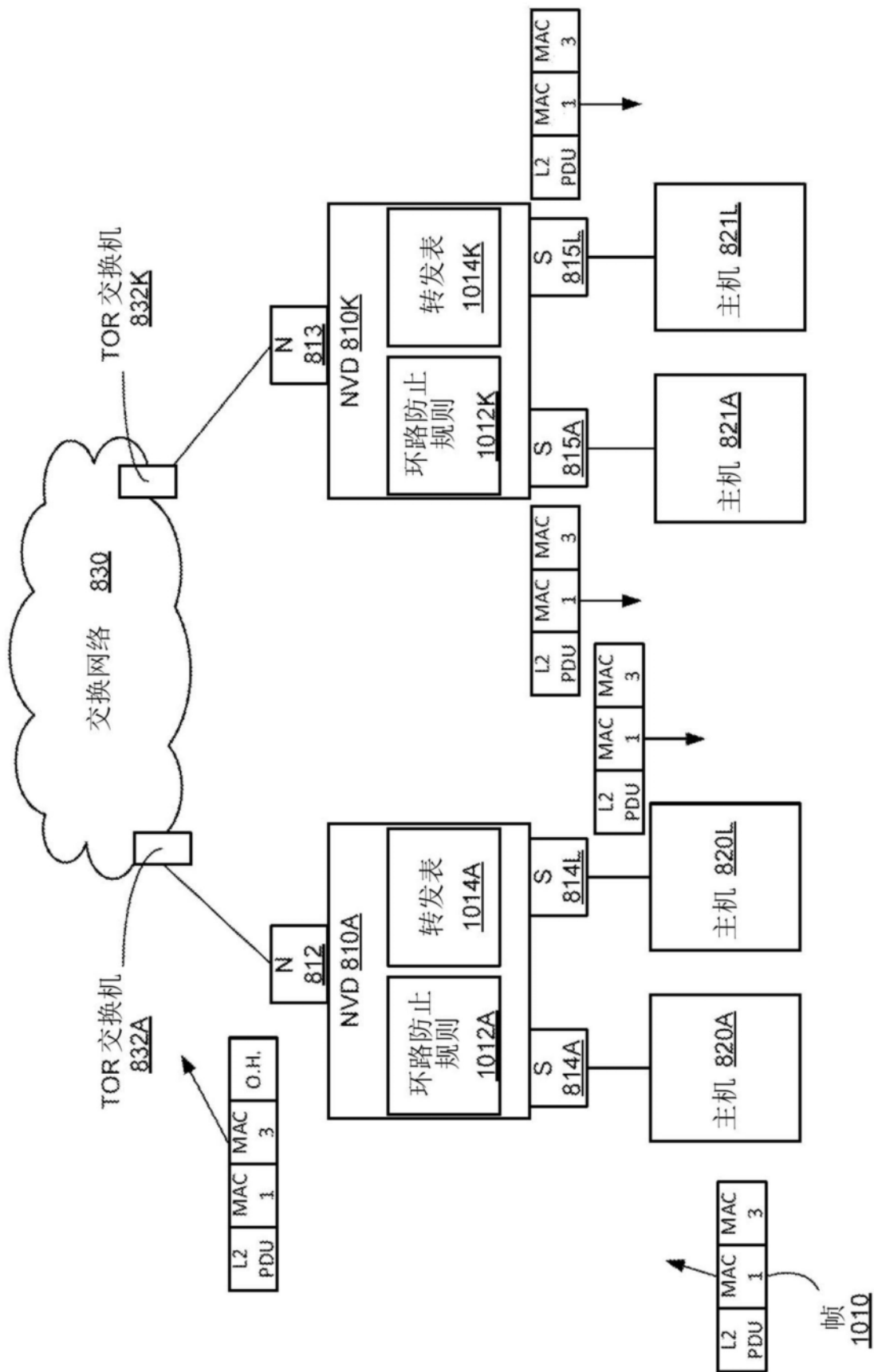


图10

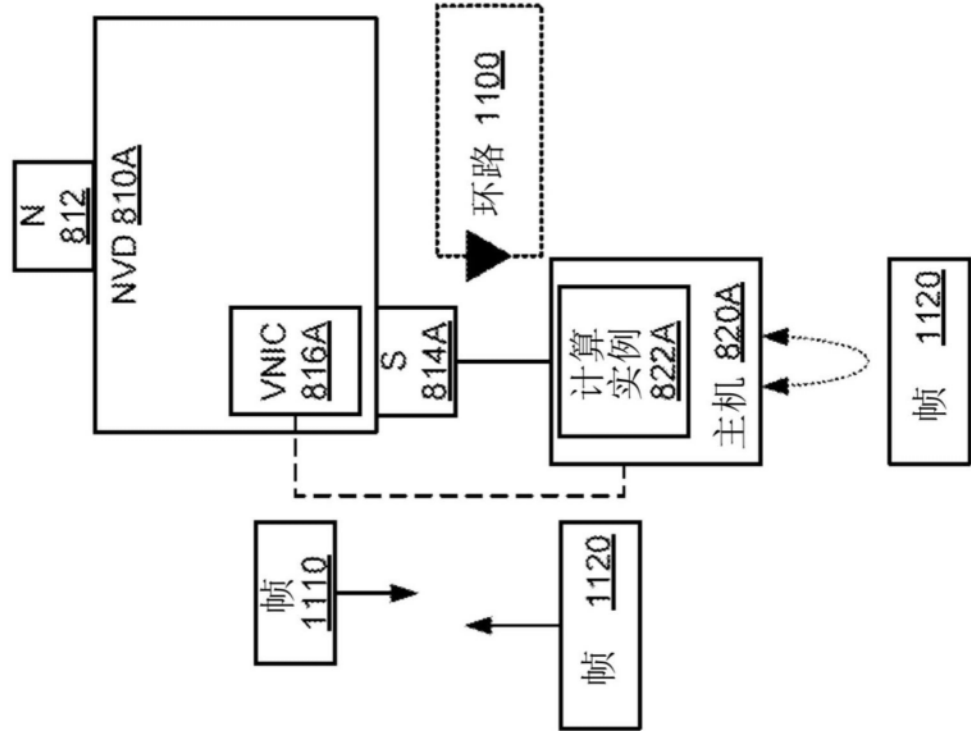


图11

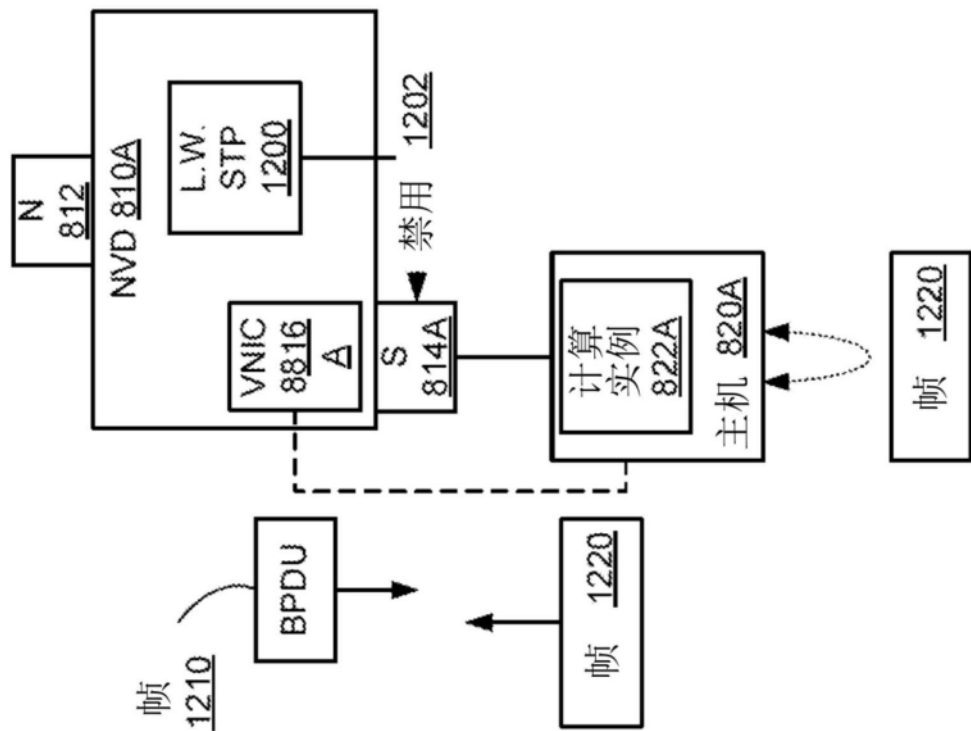


图12

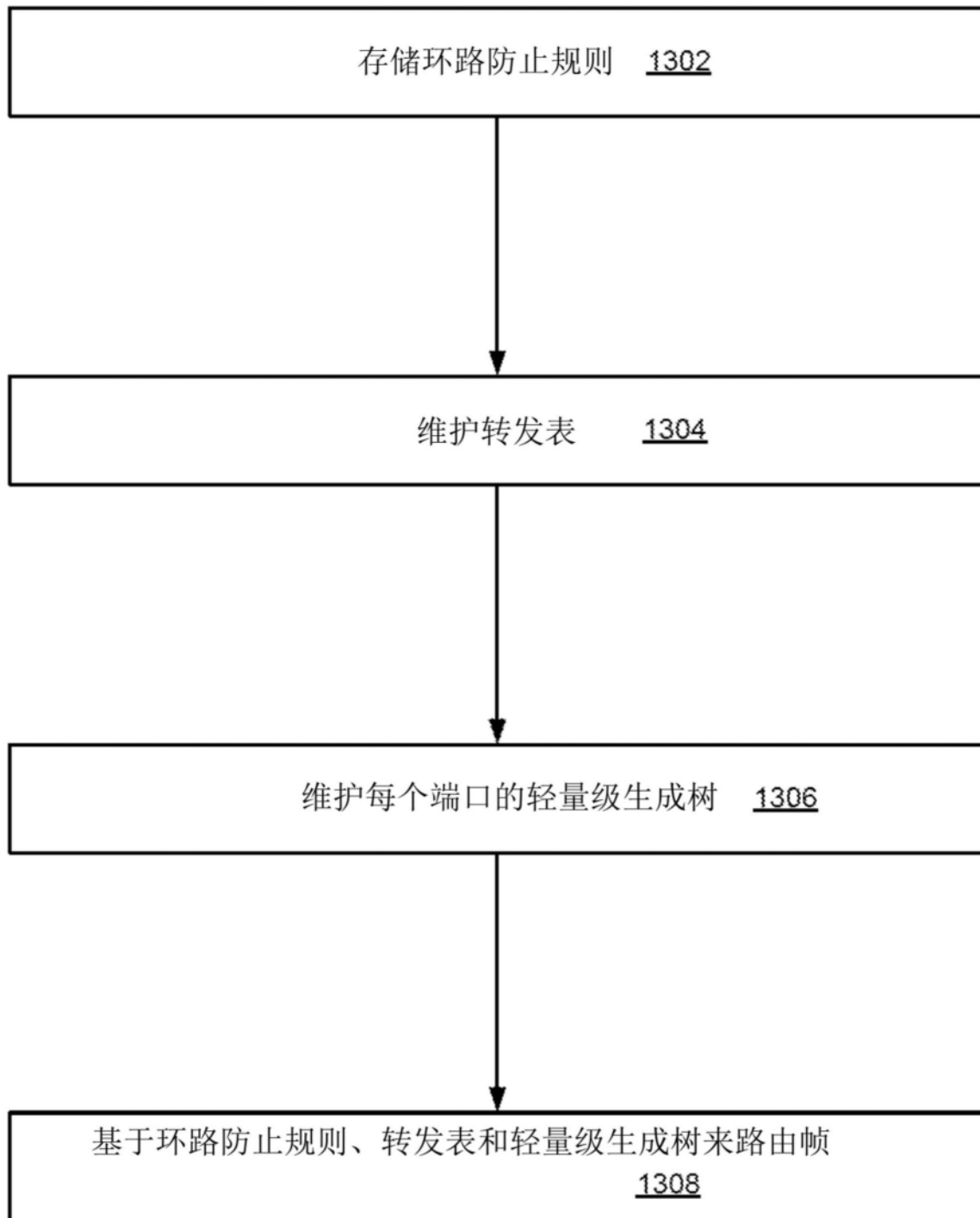


图13

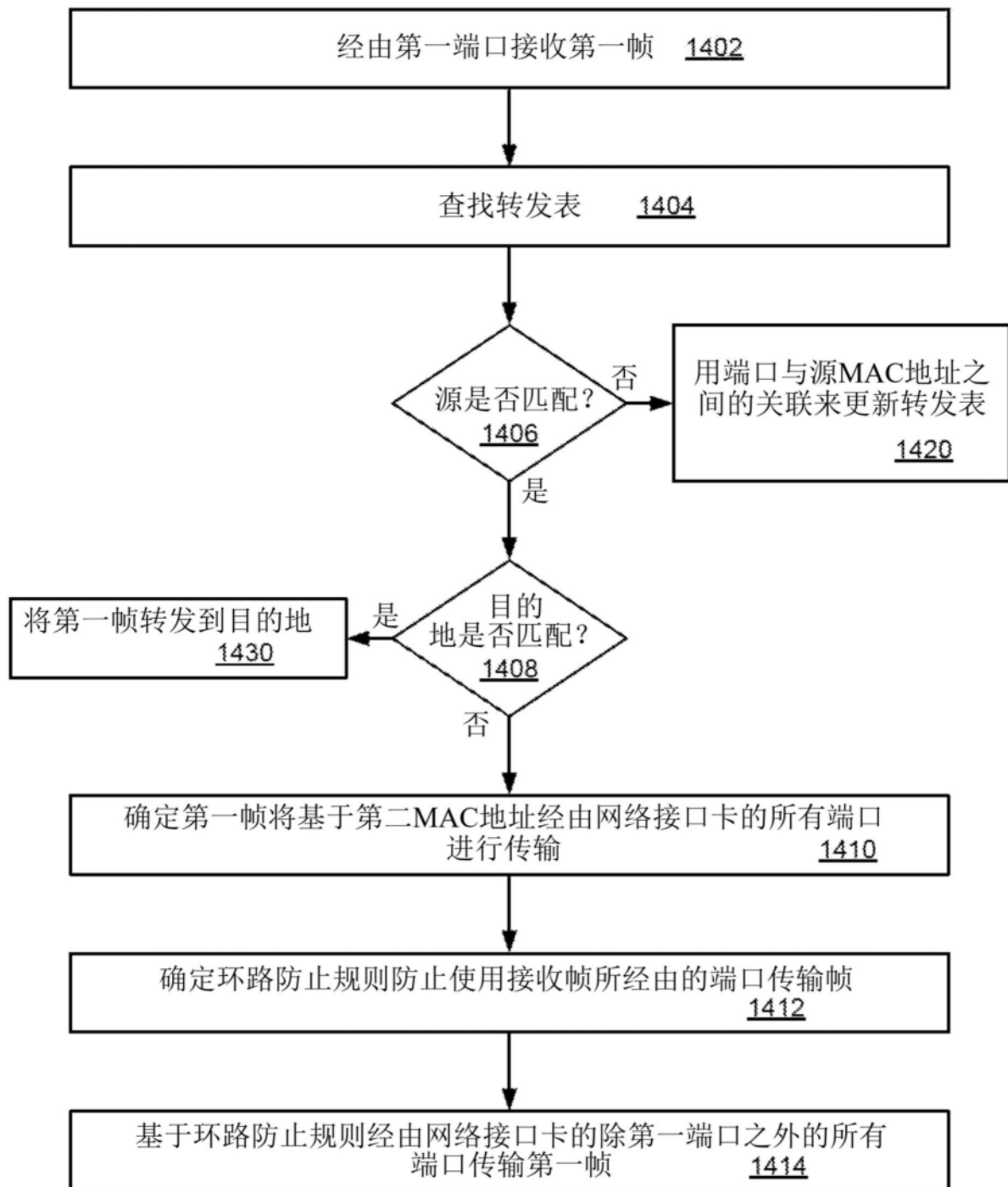


图14

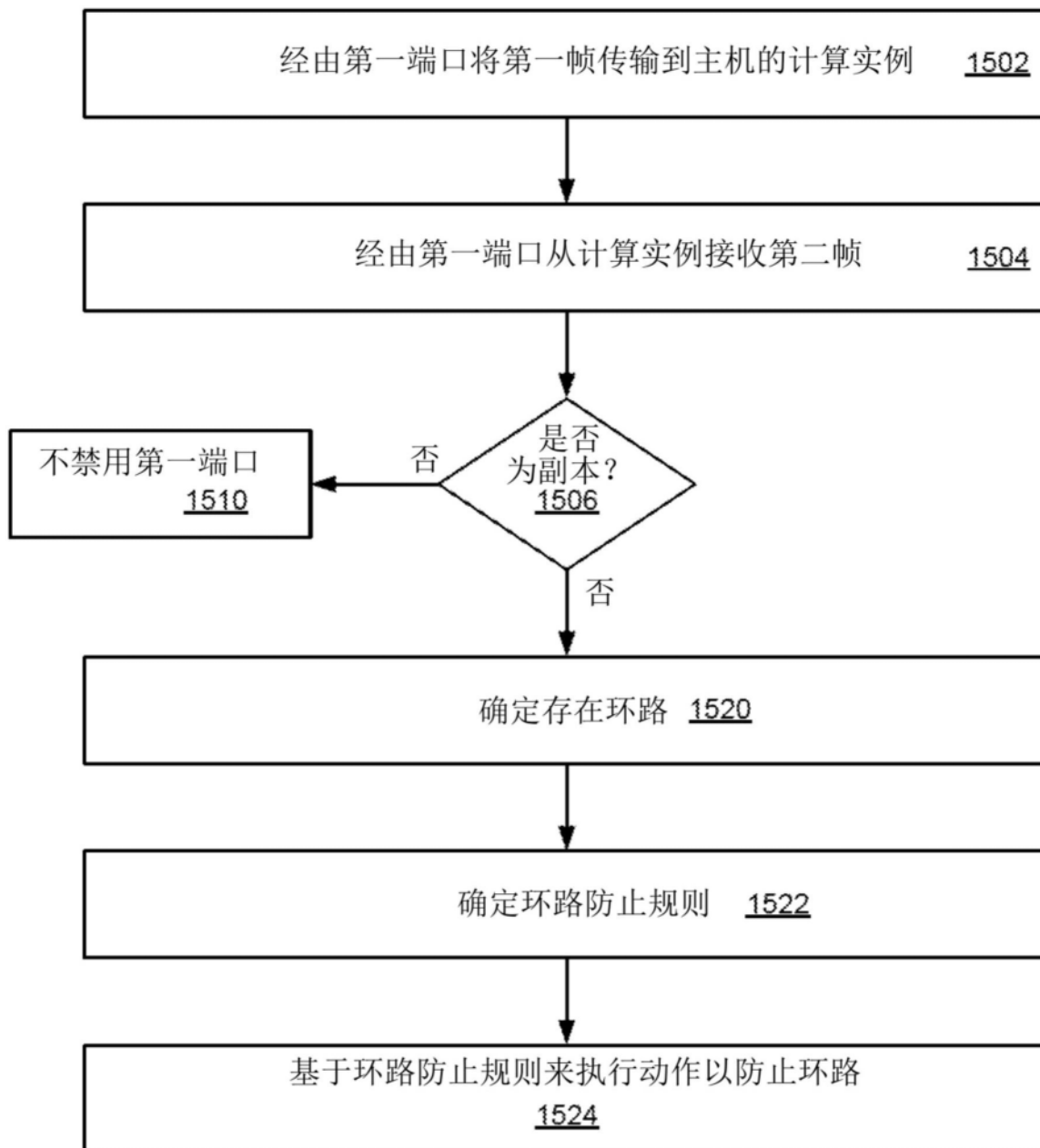


图15

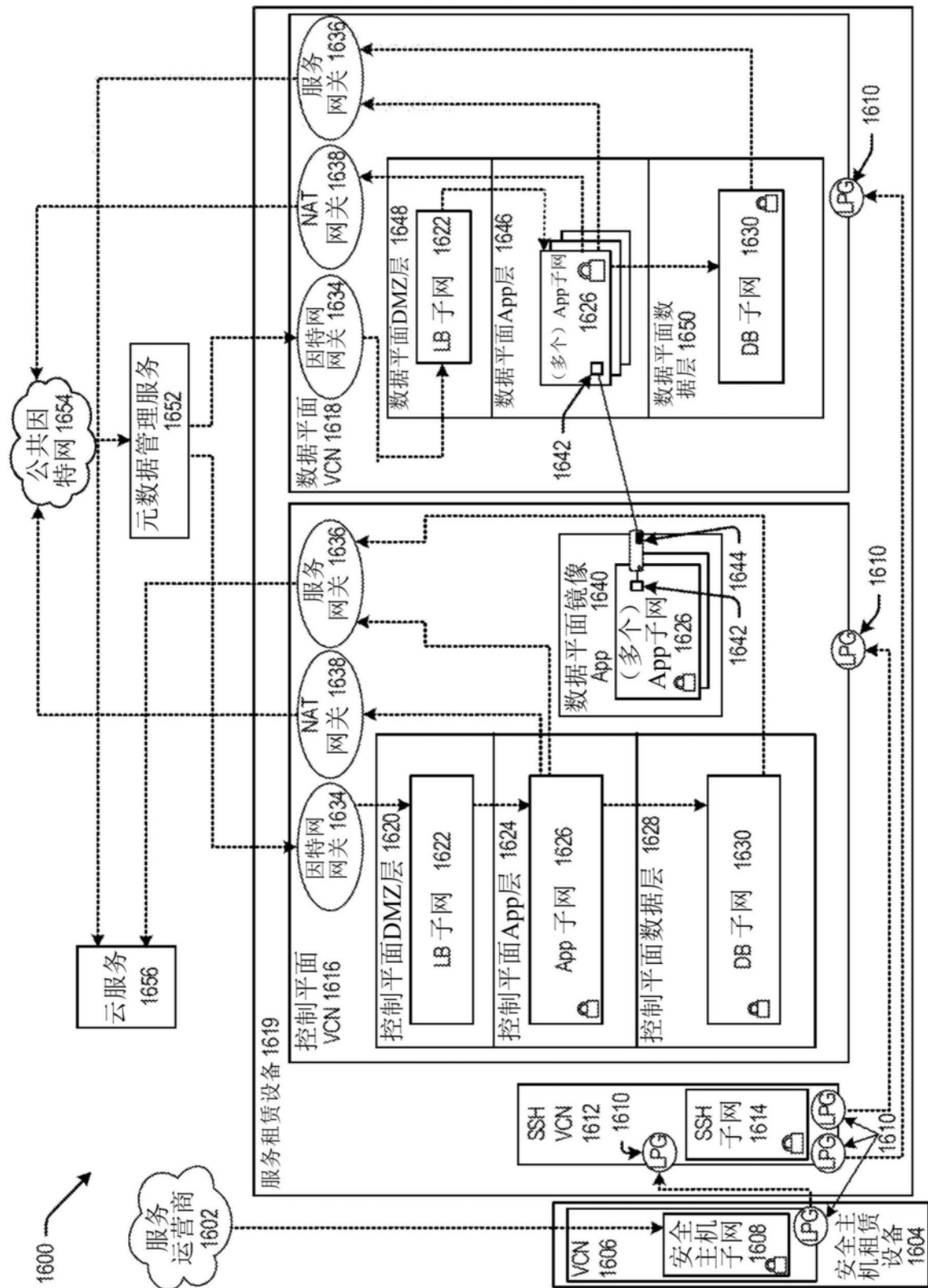


图16

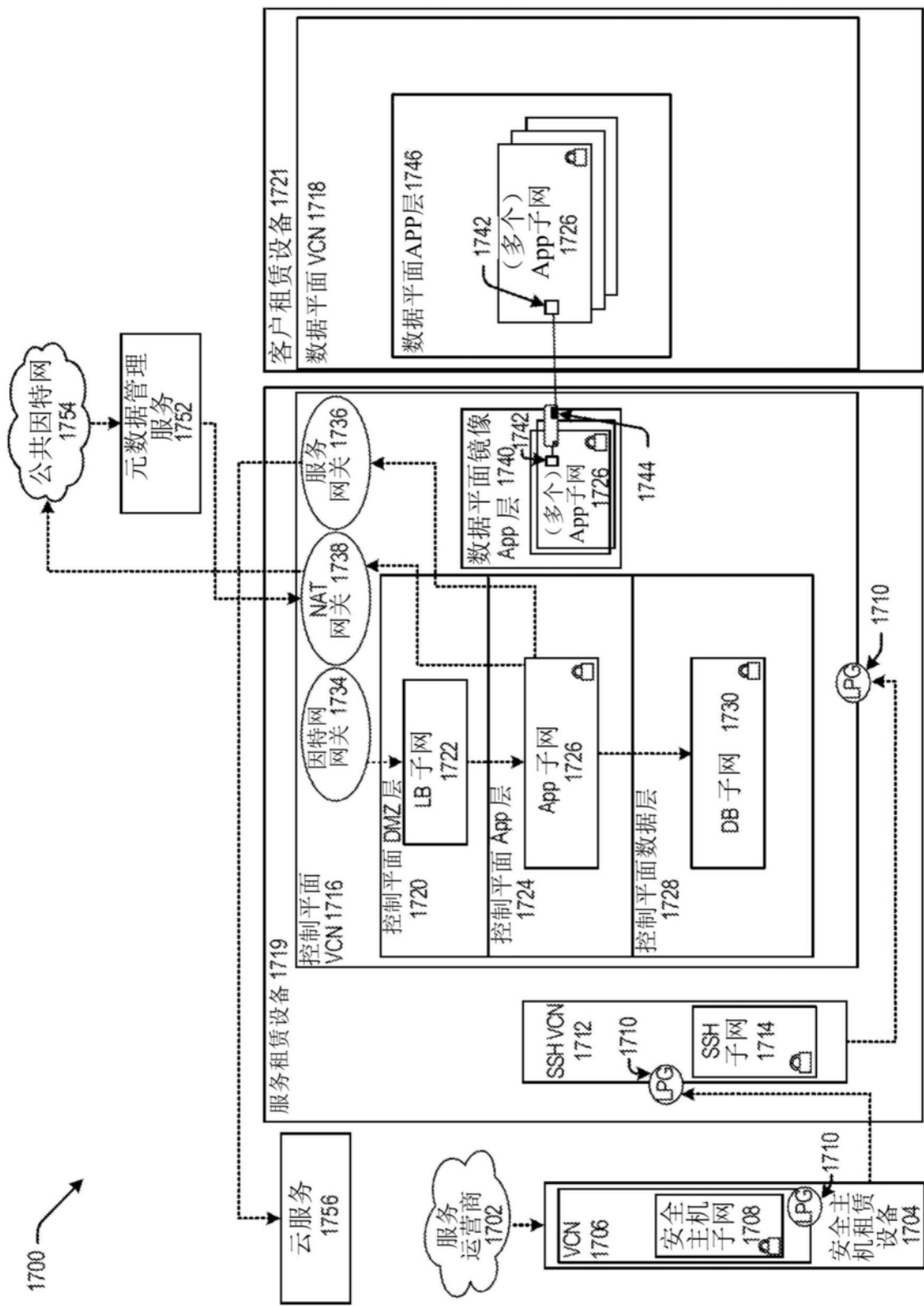


图17

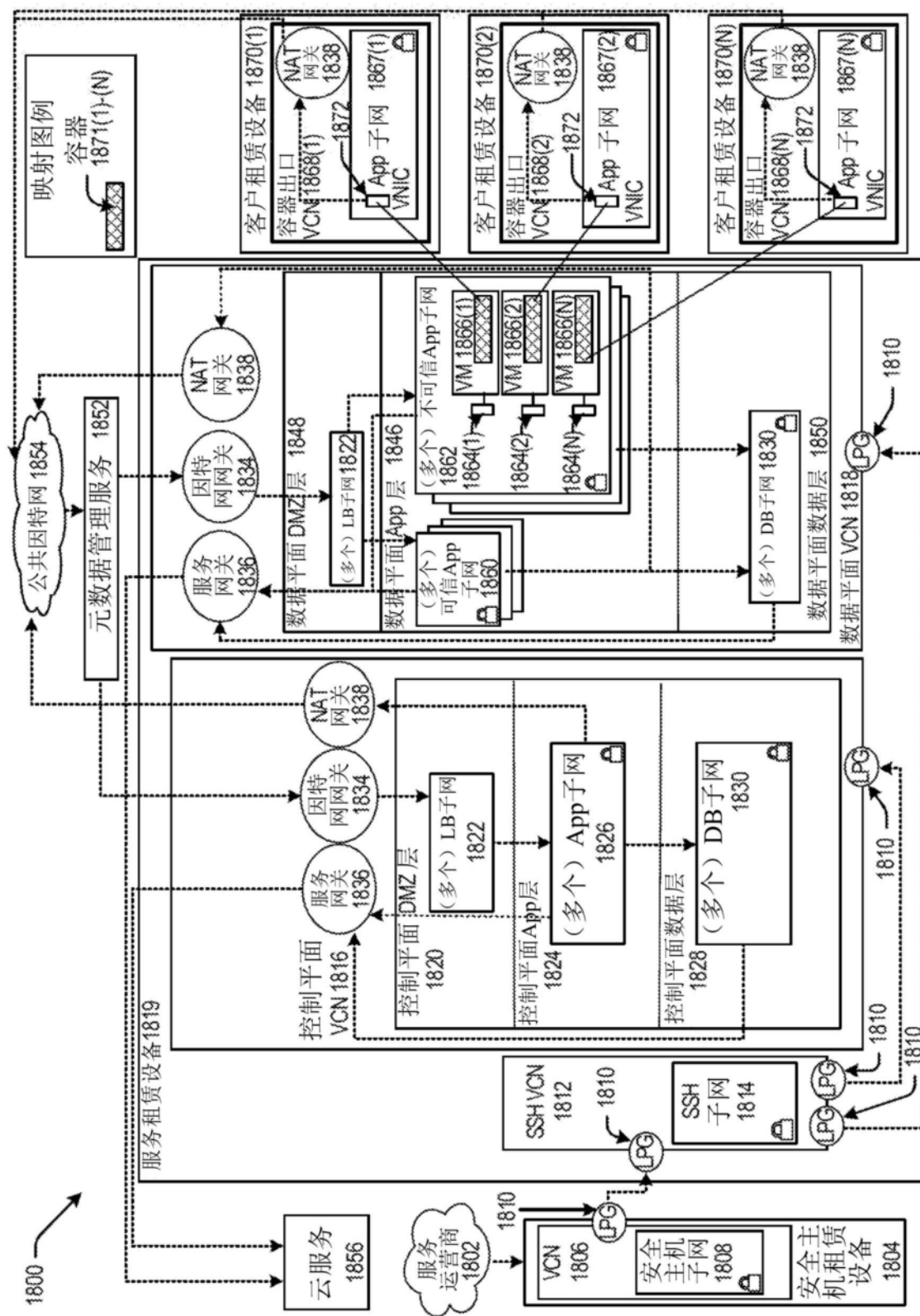


图18

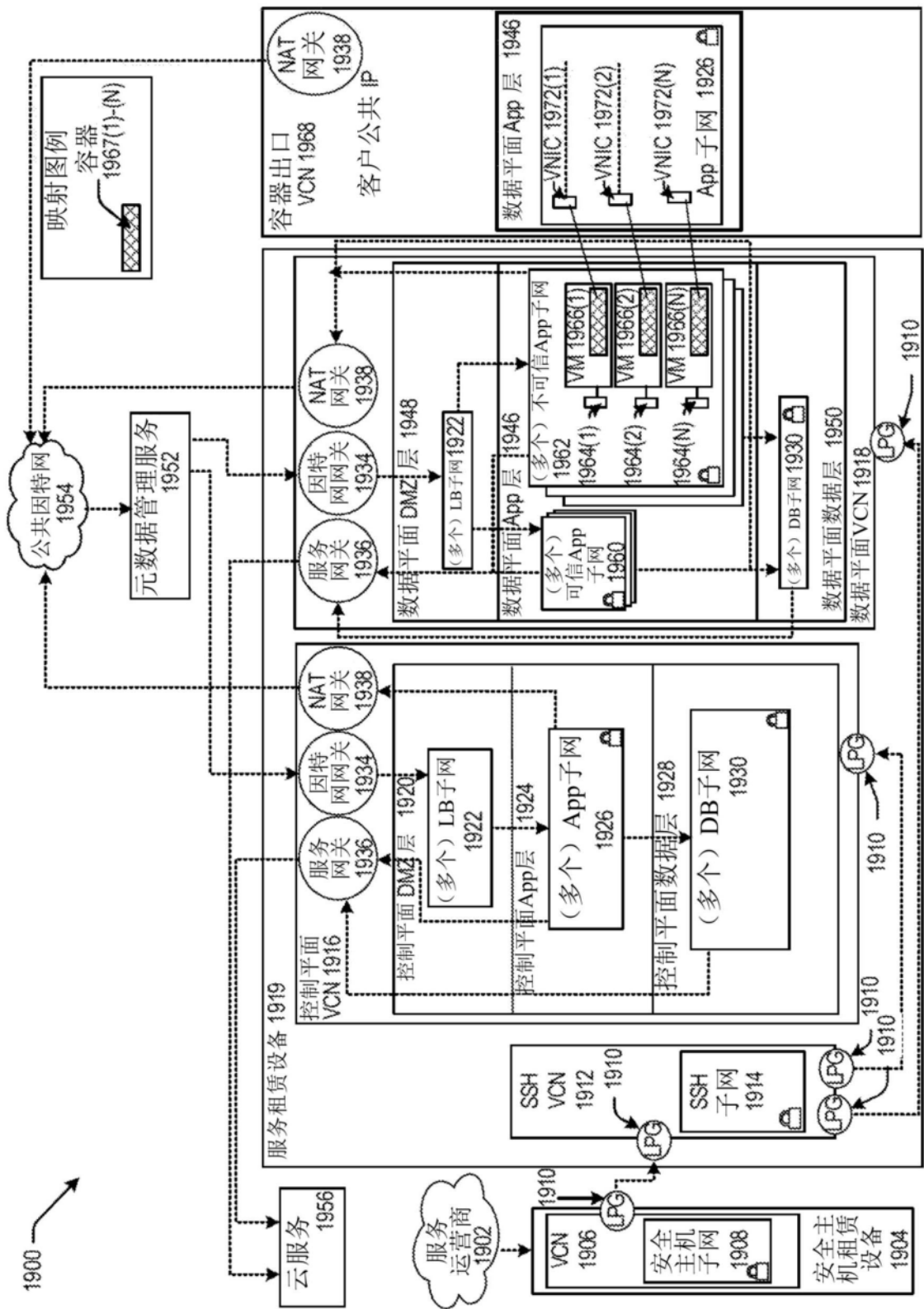


图19

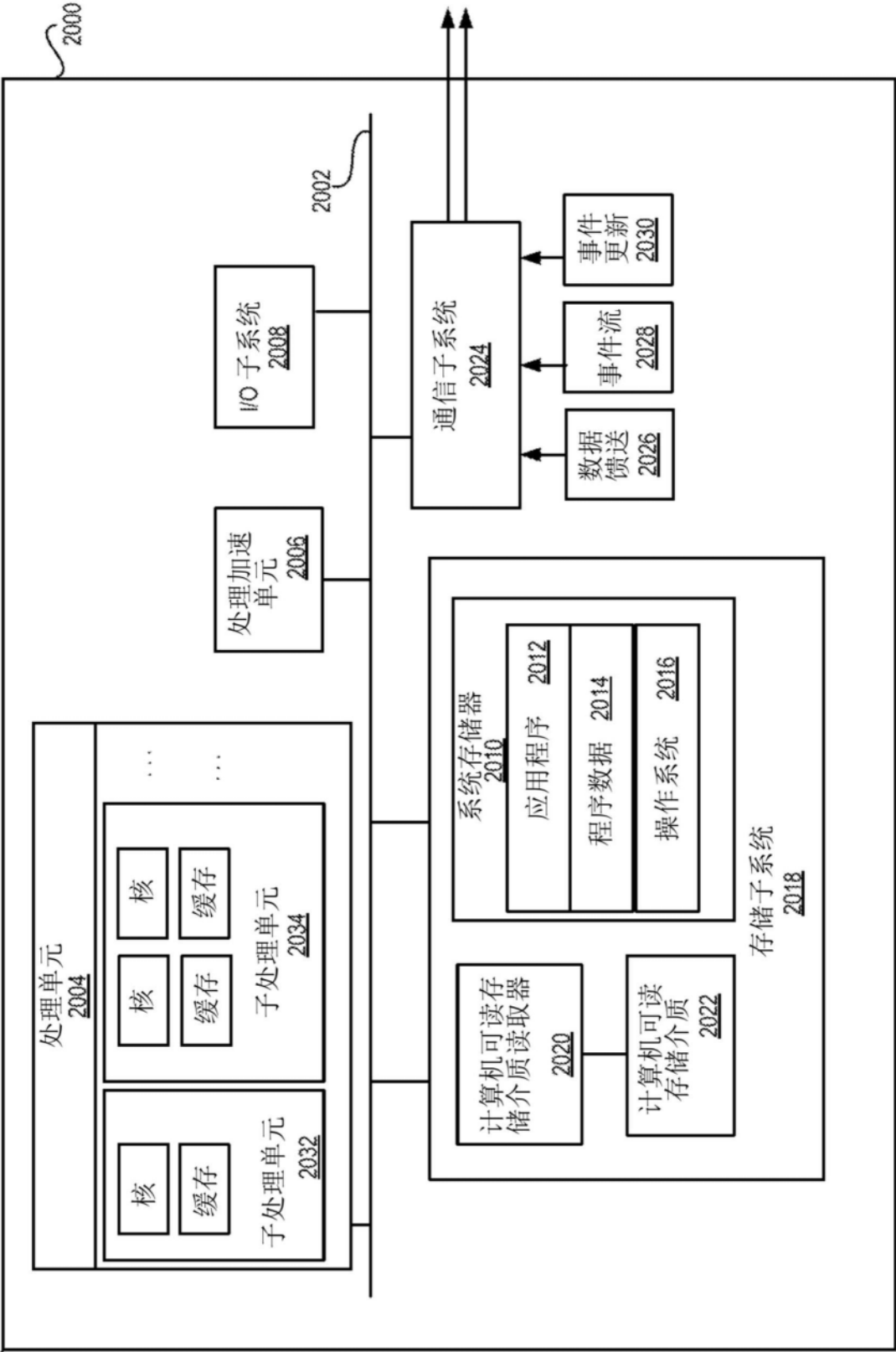


图20