



(12) 发明专利申请

(10) 申请公布号 CN 103797463 A

(43) 申请公布日 2014. 05. 14

(21) 申请号 201280037075. 6

(74) 专利代理机构 北京市中咨律师事务所
11247

(22) 申请日 2012. 06. 18

代理人 杨晓光 于静

(30) 优先权数据

13/191, 599 2011. 07. 27 US

(51) Int. Cl.

G06F 9/50 (2006. 01)

(85) PCT国际申请进入国家阶段日

2014. 01. 26

(86) PCT国际申请的申请数据

PCT/US2012/042864 2012. 06. 18

(87) PCT国际申请的公布数据

W02013/015905 EN 2013. 01. 31

(71) 申请人 阿尔卡特朗讯公司

地址 法国巴黎

(72) 发明人 M·A·阿里彻丽

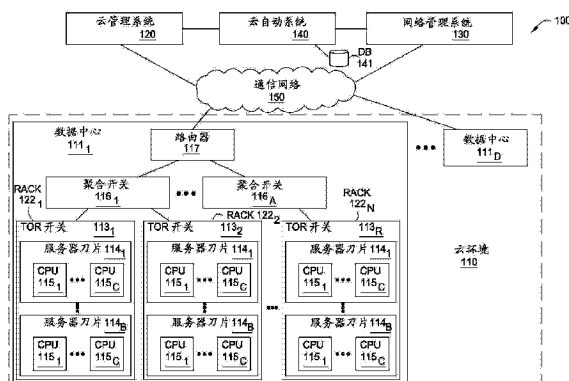
权利要求书2页 说明书17页 附图7页

(54) 发明名称

用于在云环境中指派虚拟资源的方法和设备

(57) 摘要

本发明公开了一种虚拟资源指派能力。该虚拟资源指派能力被配置为支持在云环境(110)内提供虚拟资源。在云环境中提供虚拟资源包括接收用来请求在云环境内供应虚拟资源的用户虚拟资源请求,确定指明在云环境内对虚拟资源的指派的虚拟资源指派信息,以及使用所述虚拟资源指派信息在云环境内供应虚拟资源。在云环境内对所请求的虚拟资源的指派包括将虚拟资源向将托管虚拟资源的云环境的数据中心(111)进行指派,以及特别是向将托管虚拟资源的云环境的数据中心内的物理资源进行指派。虚拟资源可包括虚拟处理器资源、虚拟存储器资源等。物理资源可包括处理器资源(115)、存储器资源等(例如云环境的数据中心的机架(112)的刀片式服务器(114)的物理资源)。



1. 一种设备,该设备包括:
处理器和存储器,所述处理器被配置为:
接收针对虚拟资源的请求,其中所述请求包括所请求的虚拟资源的数量和一个或多个虚拟资源指派限制;
接收与所述云环境相关联的管理信息;以及
使用所请求的虚拟资源的数量、所述一个或多个虚拟资源指派限制、以及所述管理信息来确定所请求的虚拟资源在所述云环境内的指派。
2. 根据权利要求 1 所述的设备,其中所述云环境包括多个数据中心,其中所述管理信息包括以下至少一者:
与所述云环境的数据中心之间的数据中心间通信相关联的网络管理信息;以及
包括针对所述数据中心中的至少一个数据中心与所述数据中心相关联的管理信息的云管理信息。
3. 根据权利要求 1 所述的设备,其中所述一个或多个虚拟资源指派限制包括以下至少一者:每个机架将被指派的虚拟资源的最大数量、每个机架将被指派的虚拟资源的最小数量、将用于每个数据中心的机架的最大数量、将用于每个数据中心的机架的最小数量、每个数据中心将被指派的虚拟资源的最大数量、每个数据中心将被指派的虚拟资源的最小数量、将被使用的数据中心的最大数量、以及将被使用的数据中心的的最小数量。
4. 根据权利要求 1 所述的设备,其中确定所请求的虚拟资源在所述云环境内的指派包括:
从在所述云环境内的可用数据中心的集合确定所述虚拟资源将被指派到的数据中心的集合;
确定将所述虚拟资源向所述虚拟资源将被指派到的数据中心的集合的指派;
针对将被指派所述虚拟资源的数据中心的集合的至少一个数据中心,选择所述数据中心的物理资源,以托管将被指派到所述数据中心的虚拟资源;以及
针对所述虚拟资源中的至少一个虚拟资源,确定将所述虚拟资源向所述虚拟资源被指派到的数据中心的物理资源的指派。
5. 根据权利要求 4 所述的设备,其中确定所述虚拟资源将被指派到的数据中心的集合包括:
将所述云环境表示为图,其中所述图包括表示所述可用的数据中心且具有与之相关联的多个权重的多个顶点,其中所述图包括表示所述可用的数据中心之间的多个通信路径且具有与之相关联的多个长度的多个边;
确定所述图的子图,该子图具有至少等于所请求的虚拟资源的数量的权重和以及最小直径;以及
从所述子图确定所述虚拟资源将被指派到的数据中心的集合。
6. 根据权利要求 4 所述的设备,其中确定将所述虚拟资源向所述虚拟资源将被指派到的数据中心的集合的指派包括:
将所述针对虚拟资源的请求表示为图,其中所述图包括多个顶点和多个边,所述顶点表示所述虚拟资源、且所述边表示由所述边所连接的虚拟资源之间的相关带宽;
将所述图分割成多个部分以由此形成分割图,其中所述分割图的每个部分表示在相同

的各个数据中心中将被调度的各个虚拟资源的集合,其中所述分割图的每个部分具有与之关联的尺寸,该尺寸以在所述各个数据中心中可用的虚拟资源的数量作为上限;以及

使用所述分割图来确定将所述虚拟资源向所述虚拟资源将被指派到的数据中心的集合的指派。

7. 根据权利要求 4 所述的设备,其中选择所述数据中心的物理资源以托管将被指派到所述数据中心的虚拟资源包括:

将所述数据中心表示为树型拓扑,其中所述树型拓扑包括表示所述数据中心的资源的多个节点,其中所述树型拓扑的节点中的每一个节点具有与之相关联的权重,该权重表示以所述节点为根的虚拟资源的数量,其中所述树型拓扑的节点中的每一个节点具有与之相关联的高度,该高度表示所述树型拓扑内的节点的高度,其中所述多个节点包括分别表示所述数据中心的多个机架、刀片式服务器或处理器的多个叶节点;

确定所述树型拓扑的子树,该子树的叶节点具有至少等于将被指派到所述数据中心的虚拟资源的所述数量的累积权重以及最小高度;以及

从所述子树确定将所述数据中心的物理资源以托管将被指派到所述数据中心的所述虚拟资源。

8. 根据权利要求 4 所述的设备,其中确定将所述虚拟资源向所述虚拟资源将被指派到的数据中心的物理资源的指派包括:

选择所述虚拟资源将被指派到的数据中心的机架;

选择所述虚拟资源将被指派到的数据中心的所选机架的刀片式服务器;以及

选择所述虚拟资源将被指派到的数据中心的所选机架的所选刀片式服务器的中央处理单元(CPU)。

9. 一种存储指令的计算机可读存储媒介,其中所述指令被计算机执行时使得所述计算机执行一种方法,该方法包括:

接收针对虚拟资源的请求,其中所述请求包括所请求的虚拟资源的数量和一个或多个虚拟资源指派限制;

接收与所述云环境相关联的管理信息;以及

使用所请求的虚拟资源的数量、所述一个或多个虚拟资源指派限制、以及所述管理信息来确定将所请求的虚拟资源在所述云环境内的指派。

10. 一种方法,该方法包括:

使用处理器和存储器以便:

接收针对虚拟资源的请求,其中所述请求包括所请求的虚拟资源的数量和一个或多个虚拟资源指派限制;

接收与所述云环境相关联的管理信息;以及

使用所请求的虚拟资源的数量、所述一个或多个虚拟资源指派限制、以及所述管理信息来确定将所请求的虚拟资源在所述云环境内的指派。

用于在云环境中指派虚拟资源的方法和设备

技术领域

[0001] 本发明总体涉及对虚拟资源的指派,尤其涉及但不限于在云环境中对虚拟资源的指派。

背景技术

[0002] 对经由云环境的云计算的使用变得越来越流行。云环境包括用来托管用于云计算用户的虚拟资源的物理资源。作为对来自云计算用户的虚拟资源请求的响应,所请求的虚拟资源被供应到云环境的物理资源上(例如位于一个或多个数据中心中的一个或多个服务器),以供用户在云计算过程中使用。

发明内容

[0003] 现有技术中的各种不足由用于云环境中的虚拟资源指派的实施方式解决。

[0004] 在一种实施方式中,一种设备包括处理器和存储器,其中所述处理器被配置为接收针对虚拟资源的请求,接收与所述云环境相关联的管理信息以及使用所请求的虚拟资源的数量、一个或多个虚拟资源指派限制和所述管理信息来确定对所述云环境内所请求的虚拟资源的指派,其中所述请求包括所述所请求的虚拟资源的数量和所述一个或多个虚拟资源指派限制。

[0005] 在一种实施方式中,一种计算机可读存储媒介存储了指令,所述指令在被计算机执行时使得所述计算机执行一种方法,该方法包括:接收针对虚拟资源的请求,接收与所述云环境相关联的管理信息以及使用所请求的虚拟资源的数量、一个或多个虚拟资源指派限制和所述管理信息来确定对所述云环境内所请求的虚拟资源的指派,其中所述请求包括所述所请求的虚拟资源的数量和所述一个或多个虚拟资源指派限制。

[0006] 在一种实施方式中,一种方法包括:使用处理器和存储器来接收针对虚拟资源的请求,接收与所述云环境相关联的管理信息以及使用所请求的虚拟资源的数量、一个或多个虚拟资源指派限制和所述管理信息来确定对所述云环境内所请求的虚拟资源的指派,其中所述请求包括所述所请求的虚拟资源的数量和所述一个或多个虚拟资源指派限制。

附图说明

[0007] 可通过结合附图考虑以下具体实施方式来容易的理解这里的教导,其中:

[0008] 图 1 描述被配置为支持向云环境的物理资源自动指派 VM 的系统的实施方式;

[0009] 图 2 描述用于响应于针对 VM 的用户请求在云环境的数据中心中供应 VM 的方法的实施方式;

[0010] 图 3 描述用于确定向数据中心内的物理资源指派 VM 以供在云环境的数据中心内提供 VM 使用的方法的实施方式;

[0011] 图 4A 和 4B 描述被配置为确定 VM 将被指派到的数据中心的集合的进程;

[0012] 图 5 描述被配置到 VM 将被指派到的机架(rack)、刀片(blade)和 CPU 的进程;

[0013] 图 6 描述了适用于在执行这里描述的功能中使用的计算机的高级框图。

[0014] 为了促进理解,尽可能地使用相同的参考符号来表示附图中共有的相同元素。

具体实施方式

[0015] 这里对虚拟资源指派能力进行一般性的说明和描述,但各种其他能力也可在这里呈现。

[0016] 在至少一些实施方式中,虚拟资源指派能力向云环境中的物理资源提供对虚拟资源的自动指派。虚拟资源可包括虚拟计算资源、虚拟存储资源等,以及其中的各种组合。物理资源可包括处理器资源(例如数据中心的机架的刀片式服务器的中央处理单元(CPU))、存储资源等。

[0017] 在至少一些实施方式中,虚拟资源指派能力响应于针对虚拟资源的用户请求向云环境中的物理资源提供对虚拟资源的自动指派。在至少一些实施方式中,接收针对虚拟资源的用户请求,确定将虚拟资源向云环境的物理资源进行的指派,并基于将所述虚拟资源向云环境的物理资源的指派在云环境中供应虚拟资源。针对虚拟资源的用户请求可包括虚拟资源请求信息(例如所请求的虚拟资源的数量、虚拟资源指派限制等),其可被用作针对确定将虚拟资源向云环境的物理资源进行的指派的输入。可使用虚拟资源指派信息(例如,对所请求的虚拟资源所位于的数据中心的选择、对所请求的虚拟资源所位于的数据中心内的设备的选择等等)来指定所确定的将虚拟资源向云环境的物理资源进行的指派。注意的是,各种其他类型的信息(例如网络拓扑和 / 或与数据中心之间的连接性相关联的状态信息、数据中心拓扑和 / 或状态信息等)可被用于确定将虚拟资源向云环境的物理资源进行的指派。注意的是,可以使用各种其他类型的信息(例如对托管虚拟资源的机架 / 刀片式服务器 / CPU 的选择、特定虚拟资源到托管虚拟资源的特定机架 / 刀片式服务器 / CPU 的映射等)来指定将虚拟资源向云环境的物理资源进行的指派。

[0018] 在至少一些实施方式中,虚拟资源指派能力以趋于提高云环境的性能的方式(例如提高对网络资源的利用)和趋于提高使用虚拟资源的用户应用的性能的方式(例如减小通信延迟、提高恢复能力等)向云环境中的物理资源提供对虚拟资源的自动指派。

[0019] 虽然这里主要在其中虚拟资源是虚拟机(VM)和物理资源是刀片式服务器的处理器的实施方式的语境中进行说明和描述,但是应该注意的是虚拟资源指派能力可被用于确定对各种其他类型的虚拟资源的指派和 / 或确定对各种其他类型物理资源的指派。

[0020] 图 1 描述被配置为支持向云环境的物理资源自动指派 VM 的系统的实施方式。

[0021] 系统 100 包括云环境 110、云管理系统(CMS)120、网络管理系统(NMS)130、云自动系统(CAS)140、和通信网络(CN)150。

[0022] 云环境 110 包括被配置为托管虚拟资源的物理资源。物理资源包括服务器和被配置为托管虚拟资源的其它类型的物理资源。虚拟资源可包括虚拟机(VM)、虚拟存储(VS)等,以及其各种组合。

[0023] 云环境 110 包括多个数据中心 111_1-111_D (统称为数据中心 111)。

[0024] 数据中心 111 包括数据中心资源(例如计算资源、通信资源等)。数据中心 111 被配置为经由通信网络 150 进行通信,从而能够彼此进行通信。数据中心 111 可以是地理上分布的。

[0025] 数据中心 111 每个包括被配置为托管虚拟资源的物理资源。每个数据中心 111 包括多个机架 112_1-112_r (统称为机架 112)。每个机架 112 包括多个刀片式服务器 114_1-114_b (统称为刀片式服务器 114)。每个刀片式服务器 114 包括多个 CPU 115_1-115_c (统称为 CPU115)。每个 CPU115 可具有一个或多个相关联的 CPU 核心(为了简明已略去)。CPU 被配置为托管 VM。数据中心 111 内的机架 112、刀片式服务器 114 和 CPU115 的典型排列和操作是本领域技术人员所理解的。注意的是,在任意的数据中心 111 内可使用任何适当类型、数量和 / 或排列的机架 112、刀片式服务器 114 和 CPU115。虽然这里主要是参照特定数量的数据中心 111、机架 112、刀片式服务器 114 和 CPU115 进行说明和描述的,但是应注意的是,在云环境 110 中可使用任何其它适当数量的数据中心 111、机架 112、刀片式服务器 114 和 / 或 CPU115 (例如云环境 110 可包括一个或多个数据中心 111,每个数据中心 111 可包括一个或多个机架 112,每个机架可包括一个或多个刀片式服务器 114,并且每个刀片式服务器 114 可包括一个或多个 CPU115)。

[0026] 每个数据中心 111 包括被配置为支持数据中心内通信和数据中心间通信的数据中心通信资源。

[0027] 数据中心 111 的数据中心通信资源是分层排列的(但是,注意也可使用其它类型的排列)。

[0028] 同一刀片式服务器 114 的 CPU115 能够直接进行通信(例如在不通过任何开关的情况下)。

[0029] 同一机架 112 的刀片式服务器 114 的 CPU 能够经由与机架 112 相关联的机架顶 (TOR)开关 113 进行通信(需要说明的是,TOR 开关 113_1-113_r 分别与机架 112_1-112_r 相关联)。

[0030] 不同机架 112 的刀片式服务器 114 的 CPU115 能够经由一个或多个开关 114(一般称作聚合开关)和 / 或路由器 117 进行通信。举例来讲,在示例性的数据中心 111_1 中,相邻的机架 112 对耦合到聚合 (aggregator)开关 116 (机架 112_1-112_2 的 TOR 开关 113_1-113_2 耦合到聚合开关 116_1 ,类似地,机架 112_r 耦合到聚合开关 116_a)。虽然为了简明进行了省略,但是应该理解的是聚合开关 116 (表示为第一层聚合开关)可被连接到数据中心中的第二层聚合开关(例如,其中第一层聚合开关 116 的集合包括两个或更多个相邻聚合开关 116,每个耦合到第二层聚合开关,其中第二层聚合开关的集合包括两个或更多个相邻的聚合开关 116,其中每个耦合到第三层聚合开关,以此类推)。如此,机架 112 和聚合开关 116 可被排列成分层的树结构。结果,数据中心 111 内的机架 112 中的刀片式服务器 114 的 CPU115 之间的时延取决于机架 112 在由机架 112 和聚合开关 116 构成的分层树结构内的位置(从而,为用户供应的 VM 之间的通信的时延取决于在其上供应所请求的 VM 的数据中心 111 的物理资源)。

[0031] 假设在本地通信的情况下排列数据中心 111 的数据中心通信资源。其中假设每个数据中心 111 内的通信的很大一部分将在相同机架 114 的 CPU115 之间发生。随着数据中心 111 的 CPU115 之间的距离增加,CPU115 之间的可用通信带宽减少。结果,在数据中心 111 内供应的所请求的 VM (一旦供应,其将经由数据中心内通信资源彼此通信)的可用通信带宽至少部分地取决于其上供应所请求的 VM 的数据中心 111 的 CPU115。类似地,数据中心 111 的总体效率还至少部分地取决于其上在数据中心 111 内供应所请求的 VM 的 CPU115。结果,可针对能被接受在云环境 110 中托管的用户请求的数量还分别至少部分地取决于其

上在数据中心 111 内供应所请求的 VM 的 CPU115。

[0032] 数据中心 111 的机架 112 经由路由器 117(并且按照说明可选地经由聚合开关 116 的一个或多个层)与数据中心 111 的外界进行通信。

[0033] 本领域技术人员将理解对数据中心通信资源的典型排列和操作。

[0034] 需要注意的是,为了支持任何数据中心 111 内的数据中心内和数据中心外通信,可使用任何适当类型、数量和 / 或排列的开关(例如 TOR 开关 113、聚合开关 116 等)和 / 或路由器(例如路由器 117)。

[0035] CMS120 被配置为针对云环境 110 的管理提供管理功能。本领域技术人员将会理解云环境的云管理系统所典型提供的管理功能。CMS120 维持与云环境 110 的管理相关联的云管理信息。CMS120 向 CAS140 提供云管理信息,以用于对响应于用户请求来确定将 VM 向云环境 110 的物理资源的指派。云管理信息可包括与数据中心 111 中的每一个相关联的数据中心资源信息(例如当前可用资源、期望在将来可用的资源、资源的当前分配、将来期望的资源分配、资源类型、资源排列、资源能力等)、与数据中心 111 中的每一个相关联的数据中心拓扑信息(例如与每个机架内的连接性相关联的、机架之间的连接性等)、与数据中心 111 中的每一个相关联的数据中心状态信息(例如资源状态信息、连接性状态信息等)等,以及其各种组合。

[0036] NMS130 被配置为提供针对 CN150 的管理的管理功能。本领域技术人员将理解由通信网络的网络管理功能提供的典型排列和操作。NMS130 维持与 CN150 的管理相关联的网络管理信息。NMS130 向 CAS140 提供网络管理信息,以用于响应于用户请求来确定将 VM 向云环境 110 的物理资源的指派。网络管理信息可包括网络拓扑信息、网络状态信息等,以及其各种组合。

[0037] CAS140 被配置为对响应于针对 VM 的用户请求来确定将所请求的 VM 向云环境 110 的物理资源的指派。

[0038] CAS140 被配置为接收请求在云环境 110 之内供应 VM 的用户 VM 请求。用户 VM 请求包括供 CAS140 在确定将所请求的 VM 向云环境 110 的物理资源的指派过程中使用的用户 VM 请求信息。用户 VM 请求信息包括多个所请求的 VM。用户 VM 请求信息还可包括针对所请求的 VM 之间的通信的通信要求。用户 VM 请求信息还可包括一个或多个 VM 指派限制(例如 VM 放置限制、VM 通信限制、VM 恢复限制、VM 类型限制等中的一个或多个)。VM 指派限制可包括将为每机架 112 指派的 VM 的最大数量、将为每机架 112 指派的 VM 的最小数量、每个数据中心 111 将使用的机架 112 的最大数量(例如为了降低机架间流量)、每个数据中心 111 将使用的机架 112 的最小数量(例如针对恢复)、将被指派给每个数据中心 111 的 VM 的最大数量(例如针对恢复)、将被指派给每个数据中心 111 的 VM 的最小数量(例如为了降低数据中心间流量)、将被使用的数据中心 111 的最大数量(例如为了降低数据中心间流量)、将被使用的数据中心 111 的最小数量(例如为了恢复)、基于将被指派的 VM 的类型的将被使用的刀片服务器 114 的类型等中的一个或多个,以及其各种组合。用户 VM 请求信息可包括适于在确定对所请求的 VM 在云环境 110 内的指派的过程中使用的任何其他信息。

[0039] CAS140 被配置为接收管理信息,以用于由 CAS140 在确定将所请求的 VM 向云环境 110 的物理资源的指派的过程中进行使用。例如,CAS140 可接收管理信息,该管理信息包括来自 CMS120 的云管理信息、来自 NMS130 的网络管理信息等。

[0040] CAS140 可使用与用户 VM 请求一同接收的用户 VM 请求信息、与系统 100 的部分相关联的管理信息等以及其各种组合来确定将所请求的 VM 向云环境 110 的物理资源的指派。

[0041] CAS140 可被配置为以用来降低资源使用的方式来确定将所请求的 VM 向云环境 110 的物理资源的指派(例如降低数据中心 111 内的机架间通信和降低数据中心之间的数据中心间通信)并向将使用 VM (一旦它们在云环境 110 中被供应)的用户应用提供提高的性能。数据中心 111 内的机架间通信的减少可通过降低被指派了 VM 的数据中心 111 的机架 112 的数量、将 VM 指派到在可能的情况下彼此相对较近的机架 112 等来提供。数据中心 111 之间的数据中心间通信的降低可通过降低被指派了 VM 的数据中心的数量、基于对数据中心 111 之间的通信能力的了解来向数据中心 111 指派 VM、基于对数据中心 111 之间的流量的了解向数据中心 111 指派 VM 以及被期待使用 VM (一旦它们在云环境 110 中被供应)的一个或多个客户端的一个或多个客户端位置等来提供。

[0042] CAS140 可被配置为可以基于用户 VM 请求的一个或多个其他特性(例如弹性、错误容忍度等)来确定将所请求的 VM 向云环境 110 的物理资源的指派。例如,CAS140 可被配置为以针对确保增长空间的方式(例如其中用户可能需要或想要在未来添加一个或多个附加 VM)指派 VM。例如,CAS140 可被配置为在考虑错误场景的同时指派 VM,从而对 VM 的结果指派使得能够从云环境 110 和 / 或 CN150 中的错误快速恢复。

[0043] 可参考图 2-6 来更好地理解 CAS140 确定将所请求的 VM 向云环境 110 的物理资源的指派的方式。

[0044] CAS140 可按照 VM 指派信息来指明对所请求的 VM 向云环境 110 的物理资源的指派。VM 指派信息包括所请求的 VM 将被放置的数据中心的标识、所请求的 VM 将被放置的数据中心内的设备的标识(例如根据每个 VM 将被放置的机架 112、刀片式服务器 114 和 CPU115 (以及可选地,核))等。VM 指派信息还可包括被配置为在供应所请求的 VM 之间的连接性的过程中使用的网络配置信息。

[0045] CAS140 可在相关联的数据库 141 中维持各种类型的信息。例如,与 CAS140 相关联的数据库 141 可存储由 CAS140 接收的信息(例如用户 VM 请求信息、管理信息等)、由 CAS140 确定的 VM 指派信息等,以及其各种组合。

[0046] CAS140 被配置为基于 VM 指派信息参与在云环境 110 中供应所请求的 VM 的过程。在一种实施方式中,CAS140 被配置为基于 VM 指派信息执行在云环境 110 内对所请求的 VM 的供应(例如经由基于 VM 指派信息生成相关联的 VM 供应请求以及将 VM 供应请求传递到云环境 110)。在一种实施方式中,CAS140 被配置为基于 VM 指派信息发起在云环境 110 中对所请求的 VM 的供应(例如通过将针对用户 VM 请求的 VM 指派信息提供给一个或多个管理系统(例如 CMS120、NMS130 等中的一个或多个)),以供一个或多个管理系统在基于 VM 指派信息在云环境 110 中供应所请求的 VM 的过程中使用。

[0047] CAS140 可被配置为执行支持虚拟资源指派能力的各种其它功能。

[0048] CN150 可以是任何类型的适合支持数据中心 111 之间的通信的通信网络。例如,可通过使用公共数据网络(例如因特网)中的一个或多个、一个或多个私有数据网络等以及其中的各种组合来提供 CN150。

[0049] 图 2 描述用于响应于针对 VM 的用户请求在云环境的数据中心中供应 VM 的方法的一种实施方式。虽然说明和描述是按先后顺序进行的,但是需要注意的是,方法 200 的步骤

可以同时执行和 / 或按照与所呈现的顺序不同的顺序执行。

[0050] 在步骤 210, 方法 200 开始。方法 200 可以响应于任何适当的状况或触发而开始。

[0051] 在步骤 220, 接收用户 VM 请求, 请求在云环境内供应 VM。用户 VM 请求包括用户 VM 请求信息。用户 VM 请求信息包括由用户请求的多个 VM。用户 VM 请求信息还可包括针对所请求的 VM 之间的通信的通信要求。用户 VM 请求信息还可包括一个或多个 VM 指派限制 (例如 VM 放置限制、VM 通信限制、VM 恢复限制等中的一个或多个)。用户 VM 请求信息可包括适于在确定所请求的 VM 在数据中心内的放置的过程中使用的任何其它信息。

[0052] 在步骤 230, 在数据中心内对所请求的 VM 的指派被确定。使用用户 VM 请求信息、网络管理信息、数据中心管理信息等以及其中的各种组合来确定在数据中心中对所请求的 VM 的指派。可按照 VM 指派信息来指明在数据中心内对所请求的 VM 的指派。VM 指派信息包括将要在其中放置所请求的 VM 的数据中心的标识、将要在其上放置所请求的 VM 的数据中心内的设备的标识 (例如按照每个 VM 将被放置的机架、刀片式服务器和 CPU (以及可选地, 核)) 等。VM 指派信息还可包括被配置为在所请求的 VM 之间供应连接性的过程中使用的网络配置信息。在一种实施方式中, 按照参照图 3 的描述和说明来确定在数据中心内对所请求的 VM 的指派。

[0053] 在步骤 240, 使用 VM 指派信息在云环境内供应所请求的 VM。可采用任何适当的模式来执行在云环境内对所请求的 VM 的供应 (例如使用发送到所请求的 VM 将被指派到的数据中心的供应消息)。

[0054] 在步骤 250, 方法 200 结束。

[0055] 图 3 描述用于确定在数据中心内对 VM 的指派以供在云环境的数据中心内供应 VM 的过程中使用的方法的一种实施方式。虽然说明和描述是按先后顺序进行的, 但是需要注意的是, 方法 300 的步骤可以同时执行和 / 或与按所呈现的顺序不同的顺序执行。

[0056] 在步骤 310, 方法 300 开始。

[0057] 在步骤 320, 为所请求的 VM 确定数据中心的集合。

[0058] 所述数据中心的集合包括所请求的 VM 将被指派到的一个或多个数据中心。所述数据中心的集合包括云环境的一个或多个可用数据中心。注意的是, 单个数据中心可能不具有足够的容量来托管所有所请求的 VM, 并且进一步讲, 即使单个数据中心中存在足够的容量来托管所有所请求的 VM, 用户也可能想要将所请求的 VM 分布到多个数据中心上, 以便提高恢复能力。

[0059] 可使用用户 VM 请求信息来确定数据中心的集合。可使用网络管理信息、数据中心管理信息等中的一个或多个以及其中的各种组合来确定数据中心的集合。可使用适合在确定数据中心的集合的过程中使用的任何其它信息来确定数据中心的集合。

[0060] 在一种实施方式中, 以趋于进行以下步骤的方式来确定数据中心的集合: (1) 最小化数据中心的集合中的数据中心的数量; 以及 (2) 降低与所请求的 VM 的放置相关联的数据中心间流量 (并且从而最小化所消耗的网络带宽)。

[0061] 在一种实施方式中, 通过选择可用数据中心中的满足以下条件的子集来确定数据中心的集合: (1) 最小化任何两个数据中心之间的最大距离 (例如跳的数量); (2) 受一个或多个限制 (例如所期望数量的 VM 的可用性、放在每个数据中心中的 VM 的最小和 / 或最大数量等)。需要注意的是, 对满足这些条件的数据中心的子集的选择可被看做子图选择问题

(这里将其表示为数据中心选择问题 MINDIAMETER (最小直径)),其描述如下。

[0062] 令可用数据中心的集合被表示为图(G)。图 $G=(V, E)$ 包括顶点(V)和边(E)。顶点(V)对应于数据中心。顶点(V)具有与之关联的顶点权重(w),其中顶点(V)的权重(w)对应于数据中心中的可用 VM 的数量。图是完全图,其中边(E)对应于连接数据中心的通信路径(例如经由因特网和 / 或任何其它相关联的分组网络)。边(E)具有与之关联的边权重或长度(l),其中两个数据中心之间的边(E)的长度(l)对应于数据中心之间的距离(例如跳的数量)。

[0063] 在一种实施方式中,可基于可在用户 VM 请求中指定的一个或多个 VM 指派限制对初始图 G 进行修改。如果用户 VM 请求包括关于可被放置在数据中心内的 VM 的最大数量的限制,则顶点(V)的权重不能大于该限制(即具有大于所述限制的权重的顶点被从图中移除)。如果用户 VM 请求包括关于可被放置在数据中心内的 VM 的最小数量的限制,则顶点(V)的权重不能小于该限制(即具有小于所述限制的权重的顶点被从图中移除)。

[0064] 令 s 表示用户 VM 请求中所请求的 VM 的数量。在一种实施方式中,数据中心选择问题 MINDIAMETER 对应于找到权重之和至少等于 s 并且具有最小的直径(即顶点之间的任何最小距离的最大长度)的 G 的子图。需要注意的是,由于原始图 G 是完全图,所以由所选择的顶点导致的子图也是完全的。如此,在一种实施方式中,数据中心选择问题 MINDIAMETER 对应于找到最长边的长度最小的子图。

[0065] 需要注意的是,数据中心选择问题 MINDIAMETER 是 NP-困难的,且对于任何的 $\epsilon > 0$ 都不能被近似到 $2^{-\epsilon}$ 内。还注意的是,数据中心选择问题 MINDIAMETER 可以从最大分团(clique)问题缩减,其中问题是找到最大尺寸的分团。缩减如下所述。给定具有最大分团问题的实例 $G=(V, E)$,则创建了具有最小直径子图问题的实例。新的完全图 $G'=(V', E', w, l)$ 被创建。G' 的顶点与 G 的顶点相同,并且具有权重 w 和长度 l 。如果 G 中的两个顶点 u 和 v 之间存在边,则 G' 的相应顶点之间的边的长度是 w ; 否则,顶点之间的边的长度是 l 。该图中的边满足三角不等式。当且仅当 G 中的相应子图是分团的情况下,G' 的子图具有直径 w 。这是因为,如果 G 具有尺寸为 k 的分团,则 G' 的相应顶点可被采用,以形成具有权重为 k 和直径为 w 的子图。类似地,如果 G' 具有权重 k 和直径 l 的子图,则 G 的相应顶点形成具有尺寸 k 的分团。当且仅当 G 中的相应顶点是分团时,G' 的子图具有直径 w 。从而,可通过找到存在尺寸为 s 和直径为 w 的子图的最大 $s \in \{n, n-1, \dots, 1\}$ 来找到对最大分团问题的解决方案。

[0066] 需要再次注意的是,除非 $P=NP$, 否则对于任何的 $\epsilon > 0$, 数据中心选择问题 MINDIAMETER 都不能被近似到因数 $2^{-\epsilon}$ 内。如果存在针对数据中心选择问题 MINDIAMETER 的 $2^{-\epsilon}$ 近似进程,则可按如下来解决最大分团问题。为了找到具有尺寸 k 的分团,使用以上的缩减来创建具有权重 k 的 MINDIAMETER 问题。如果存在尺寸为 k 的分团,则存在针对具有直径 w 的数据中心选择问题 MINDIAMETER 的子图。如果不存在这样的分团,则数据中心选择问题 MINDIAMETER 的直径至少为 l 。从而,如果针对数据中心选择问题 MINDIAMETER 的 $2^{-\epsilon}$ 进程返回直径小于 l 的子图,则在原始的图 G 中存在尺寸为 k 的分团。

[0067] 在一种实施方式中,为解决最小直径子图问题(即数据中心选择问题 MINDIAMETER)提供近似进程。在一种实施方式中,进程所输出的子图的直径至多为最优子

图的直径的两倍。对于近似进程,假定三角不等式对图中的边权重成立。需要注意的是,三角不等式在这一问题设置中时有效的,这是因为边的长度对应于对应数据中心之间的路径长度。还应注意的是,如果存在三个数据中心之间违反三角不等式,则仍可通过强制较长的边采用由较短的边形成的路径来维持对三角不等式的符合。

[0068] 在一种实施方式中,针对解决最小直径子图问题所提供的近似进程被实施为两个进程,这里表示为 FindMinStar(G, v, s) 进程(在图 4A 中说明和描述)和 MinDiameterGraph(G, s) 进程(在图 4B 中说明和描述),其描述如下。需要注意的是,虽然主要说明和描述为单独的进程,但是这些进程可被实施为更少或更多的进程,其中可为了提供所描述的功能执行这些进程。还需要注意的是,虽然参照进程的特定实施(例如特定变量、进程的步骤的特定排列等)进行说明和描述,但是可以使用适于提供所描述的功能的任何其它方式来实施这些进程。

[0069] 如此所述,图 4A 的 FindMinStar(G, v, s) 进程 410 和图 4B 的 MinDiameterGraph(G, s) 进程 420 被配置为确定 VM 将被指派到的数据中心的集合。

[0070] 如图 4A 中所述,FindMinStar(G, v, s) 进程 410 被配置为包括顶点 v 的权重至少为 s 的子图。通过以长度的增序向 v 增加节点来形成以顶点 v 为中心的星形拓扑,直到所述星形拓扑的权重至少为 s 为止。通过添加由星(star)中的节点引入的边来形成最终的子图。FindMinStar(G, v, s) 进程 410 还通过随着添加节点而维持直径来计算所得到的子图的直径(其中存在如下限制,当添加节点时,只有由该节点所引入的边的长度大于当前直径的情况下,子图的直径能够改变)。

[0071] 如图 4B 中所述,MinDiameterGraph(G, s) 进程 420 被配置为通过为每个顶点调用 FindMinStar(G, v, s) 进程 410 来找到权重至少为 s 的子图。MinDiameterGraph(G, s) 进程 420 随后选择具有最小直径的子图。

[0072] FindMinStar(G, v, s) 进程 410 找到权重至少为 s 的子图,其在入射到 v 上的任何边的长度是最小的。FindMinStar(G, v, s) 进程 410 首先向子图添加顶点 v 。FindMinStar(G, v, s) 进程 410 随后以边权重的增加顺序来添加与 v 相邻的顶点,直到子图中的顶点的权重之和为 s 为止。从而,FindMinStar(G, v, s) 进程 410 找到针对入射到 v 的边具有最小边权重的子图。

[0073] MinDiameterGraph(G, s) 进程 420 找到权重至少为 s 的子图,其直径至多为最优的两倍。MinDiameterGraph(G, s) 进程 420 为 G 的每个顶点调用 FindMinStar(G, v, s) 进程 410 并选择具有最小直径的一个。令 v' 作为 FindMinStar(G, v, s) 进程 410 所返回的图 G' 的直径最小的节点,并且令 l' 作为入射到子图 G' 中的 v' 上的最长边的长度。由于边长度遵循三角不等式,并且 G' 是完全图,所以 G' 中的任何边的长度至多为 $2l'$ 且从而 G' 的直径最多为 $2l'$ 。

[0074] 关于数据中心选择问题 MINDIAMETER,需要注意的是,针对解决数据中心选择问题 MINDIAMETER 的近似进程是 2- 近似进程。令 G_{opt} 作为数据中心选择问题 MINDIAMETER 的最优方案。令 l'' 作为 G_{opt} 的最长边。令 u'' 和 v'' 作为 l'' 的端点。由于 G_{opt} 的边满足三角不等式,所以 G_{opt} 中在 u'' 和 v'' 之间的最短距离为 l'' 。从而, G_{opt} 的直径至少为 l'' 。现在,在运行 MinDiameterGraph(G, s) 进程 420 的同时,考虑由 FindMinStar(G, v, s) 进程 410 所返回的图 G'' 。如上所述,入射到 G'' 上的 v'' 的边的长度至多为 l'' 。由于 G'' 的直径至多为

21”，所以由 $\text{MinDiameterGraph}(G, s)$ 进程 420 所返回的子图(即 G')的直径最多为 21”。从而，针对解决数据中心选择问题 MINDIAMETER 的近似进程是 2- 近似进程。

[0075] 现在返回图 3, 需要注意的是, 可以以其它方式(例如, 使用其它信息、以针对满足其他条件的方式、使用其它进程等, 以及其中的各种组合)执行针对所请求的 VM 的所确定的数据中心的集合。

[0076] 在步骤 330, 确定将所请求的 VM 向数据中心的集合中的数据中心的指派。

[0077] 将所请求的 VM 向数据中心的集合中的数据中心的指派针对数据中心的集合中的每个数据中心指明向该数据中心所指派的所请求的 VM 的数量。需要注意的是, 将所请求的 VM 向数据中心的集合中的数据中心的指派不必指明 VM 将被指派到的数据中心的设备(由于这可在步骤 340 处被确定)或 VM 中的哪一个将被指派到数据中心中的哪一个(由于这可在步骤 350 处被确定)。

[0078] 可使用用户 VM 请求信息、网络管理信息、数据中心管理信息等中的一个或多个以及其各种组合将所请求的 VM 指派到所述数据中心的集合中的数据中心的。在一种实施方式中, 以趋于最小化数据中心间流量的方式将所请求的 VM 指派到数据中心的集合中的数据中心的。在一种实施方式中, 根据用户 VM 请求中指明的一个或多个 VM 指派限制将所请求的 VM 指派到数据中心的集合中的数据中心的。在一种实施方式中, 例如其中用户 VM 请求信息包括与向不同数据中心指派不同类型的 VM 相关联的 VM 类型限制, 基于 VM 类型限制来将所请求的 VM 指派到数据中心的集合中的数据中心的。

[0079] 在一种实施方式中, 使用被配置用于向数据中心指派 VM 的 VM 指派进程将所请求的 VM 指派到数据中心的集合中的数据中心的。一般来讲, VM 指派进程的目标是将图分成多个等尺寸的部分, 从而图的这些部分之间存在很少连接(由此, 最小化支持所请求的 VM 所需的网络通信资源)。然而, 即使将图分成两个相等部分的最简单情况(典型的被称作图二分问题)也是 NP- 困难的。此外, 相关问题是 K- 切割问题, 其中目标是将图分成 K 个部分, 其中每部分的尺寸也是被输入所指明的。

[0080] 在 VM 指派进程中, 用户 VM 请求可被看作需要被分割的图。图的节点表示 VM, 且图的链接表示由链接所连接的 VM 对之间所需的相关带宽。对图的每个部分表示需要被指派到相同数据中心的 VM 的集合。每个部分的尺寸需要由数据中心的用于托管 VM 的可用资源的数量作为上限。与传统图分割问题不同, 该问题不可指明每个部分中的节点的具体数量; 反而, 其只指明每个部分中的最大节点数。这是由于, 在任一点, 在云环境中可用于托管 VM 的资源数量要多于由用户所请求的 VM 的数量。结果, VM 指派进程可通过在具有可用于托管 VM 的最大的资源的一个或多个数据中心之内指派 VM 来最优化(数据中心间的)通信。针对向数据中心指派 VM, 数据中心间通信成本取决于数据中心间的通信路由。从而, 为了最小化应用的通信成本, 对数据中心部分之间的通信距离进行考虑。

[0081] 关于 VM 指派进程, 可按如下一般化所述问题。给定具有指派的边权重 $w: E \rightarrow \mathbb{N}$ 的图 $G=(V, E)$, 以及具有将部分 i 和 j 之间的距离表示为 $d(i, j)$ 的分割容量 k_1, k_2, \dots 的集合。将 V 分成分离的集合 C_1, C_2, \dots, C_m , 以使得 $|C_i| \leq K_i$ 以及

$$\sum_{i=1}^k \sum_{j=i+1}^k \sum_{v_1 \in C_i, v_2 \in C_j} w(v_1, v_2) d(i, j)$$

[0082] 关于 VM 指派进程, 需要注意的是, 要么可以以递增的方式分离地处理单独的用户

请求,要么可将所有多重请求合并在一起以形成大的分割图,其中该分割图在不同用户请求的节点之间不包含任何边。

[0083] 关于 VM 指派进程,需要注意的是,图分割问题是 NP- 困难的,并且没有任何进程具有良好的(例如常量的)近似保证。

[0084] 在一种实施方式中,VM 指派进程被实施为贪婪启发式进程。

[0085] 在一种实施方式中,VM 指派进程一次选择一个数据中心。可以以任何适当的顺序来选择数据中心(例如执行每个选择以使得选择带有可用于托管 VM 的最大数量资源的数据中心、基于在用户 VM 请求中指明的一个或多个 VM 指派限制等,以及其各种组合)并向所选的数据中心指派尽可能多的 VM。VM 指派进程选择具有参数的最大值的 VM (例如进入 / 离开的带宽、邻居数量等)并向所选的数据中心指派所选的 VM。所选的 VM 被添加到集合 C,其是指派到所选数据中心的 VM 的集合。VM 指派进程随后考虑 C 的所有邻居并选择具有到往 / 离开 C 的最大流量的 VM,以及将所选的 VM 添加到集合 C。可重复该进程直到针对所选的数据中心满足条件为止(例如所选数据中心的所有可用资源已被耗尽,针对所选的数据中心满足按照数据中心限制的最大 VM 数量等)。VM 指派进程从而选择下一个数据中心并针对该下一个所选数据中心和尚未被指派到(V-C)数据中心的剩余 VM 的图重复所述进程。

[0086] 在一种实施方式中,通过交换特定节点来修改 VM 指派进程。VM 指派进程考虑在不同的部分中存在的节点对。VM 进程可确定交换所考虑的节点是否会改进所述方案(例如使用较少的数据中心间带宽)。VM 进程可确定将节点在部分之间移动(例如从其当前部分移动到具有可用容量的不同部分)是否会改进所述方案。VM 指派进程选择一个或多个所考虑的移动并进行所选的移动。需要注意的是,该进程可被重复(例如,直到达到了移动的阈值数量、直到所有移动已被考虑且 VM 指派中没有任何进一步的改进空间,等等)

[0087] 在步骤 340,针对数据中心的集合中的每个数据中心,选择将被用来托管将被指派到的数据中心的所请求的 VM 的数据中心的物理资源。

[0088] 对数据中心内的物理资源的选择可包括对将被用来托管指派到数据中心的 VM 的数据中心的机架的选择。对数据中心内的物理资源的选择可进一步包括对所选的机架上的刀片式服务器的选择。对数据中心内的物理资源的选择可进一步包括对所选的机架的所选刀片式服务器上的 CPU 的选择。对数据中心内的物理资源的选择可进一步包括对所选机架的所选刀片式服务器的所选 CPU 的核心的选择。如此,用于对数据中心内的物理资源的选择的进程在此可被更为一般地称为设备选择进程、机器选择进程和 / 或物理资源选择进程。需要注意的是,对物理资源的选择不必指明从将被指派到数据中心的 CM 到数据中心的所选物理资源的映射(这可在步骤 350 被确定)。

[0089] 可以以趋于最小化数据中心内的 VM 之间的机架间流量的方式执行对数据中心的物理资源的选择。这将趋于防止指派到数据中心内的 VM 之间的长通信路径,由此防止数据中心内的 VM 之间的通信中的较大时延并减少数据中心内对资源(例如通信资源)的使用。

[0090] 这可大大缩减较大数据中心(例如那些具有上百机架、上千机架甚至更多机架的数据中心)中的通信延迟,其中彼此位置相对不近的机架之间的任何通信可能必须遍历各个数量 / 级别的聚合开关。

[0091] 可以以趋于提供弹性的方式执行对数据中心的物理资源的选择。

[0092] 需要注意的是,对数据中心的物理资源的选择可改进数据中心利用以及请求了 VM

的一个或多个用户应用的性能。

[0093] 在一种实施方式中,使用机架选择进程(例如实施机器选择进程,其中选择了针对所请求的 VM 的一个或多个机架)执行对所选数据中心的一个或多个机架的选择。

[0094] 在一种实施方式中,机架选择进程被实施为用于数据中心选择的数据中心选择进程的修改版本。在该实施方式中,机架选择进程是 2- 近似进程,其被配置为最小化机架之间的最大距离。在这一实施方式中,图的顶点表示机架,且对于每个机架,机架的相关联的权重表示机架中用于支持指派到数据中心的 VM 的可用的资源数量。

[0095] 在一种实施方式中,其中数据中心的拓扑是分层网络,机架选择进程可确定针对机架选择问题的最优方案。所述拓扑可被认为是树型拓扑,其中树的根节点表示最高层的一个或多个开关(例如用于提供到往 / 来自数据中心的连接性的路由器),根节点的子节点表示开关(例如顶级聚合开关)的网络层,以此类推,直到表示机架的叶节点。需要注意的是,在这一树中,所有的叶节点在相同的级别。树被扩大为包括标签,标签用于针对每个机架指示对于支持指派到数据中心的 VM 可用的机架的资源数量。

[0096] 在一种实施方式中,使用刀片式服务器选择进程来执行对机架的刀片式服务器的选择(例如使用实施机器选择进程,其中选择针对所请求的 VM 的机架和刀片式服务器)。在一种实施方式中,刀片式服务器选择进程被实施为机架选择进程的修改版本,其中树的叶节点表示刀片式服务器级别,且叶节点的父节点表示机架。

[0097] 在一种实施方式中,使用 CPU 选择进程(例如使用对机器选择进程的实施,其中选择针对所请求的 VM 的机架、刀片式服务器和 CPU)来执行对机架的刀片式服务器的 CPU 的选择。在一种实施方式中,CPU 选择进程被实施为机架选择进程的修改版本,其中树的叶节点表示 CPU 级别,且刀片式服务器节点的父节点表示机架。

[0098] 如上所述,可针对数据中心执行机器选择进程,以用于选择数据中心内的指派到数据中心的 VM 被指派到的一个或多个机器。

[0099] 在一种实施方式中,针对 VM 被指派到的给定数据中心,机器选择进程被配置为最小化指派到数据中心的两个 VM 之间的最大通信距离。这可被译为,找到叶上的标签之和至少等于所需的 VM 的目标数量的具有最小高度的有根子树。如在数据中心选择进程中一样,机器选择进程可考虑到一个或多个 VM 指派限制(例如每个机架可被指派的 VM 的最大数量、每个机架可被指派的 VM 的最小数量等)。可通过根据 VM 指派限制改变表示机架的节点的权重来考虑这些类型的 VM 指派限制。

[0100] 在一种实施方式中,机器选择进程被实施为 FindMinHeightTree(T, r, s) 进程(如图 5 中描述和说明)。

[0101] 如图 5 中所述,FindMinHeightTree(T, r, s) 进程 500 被配置为针对给定的数据中心选择将被用于托管指派到该数据中心的 VM 的数据中心的物理资源。令 s 作为将被指派到数据中心内的 VM 的数量。令 T 作为数据中心资源的树表示(例如计算资源、网络资源等)。下列两个变量与 T 中的每个节点相关联:(1) 权重(w),表示以节点 v 处为根的可用于支持 VM 的数据中心资源的数量;以及(2) 高度(h),表示节点的高度。只初始化叶节点的权重变量。FindMinHeightTree(T, r, s) 进程 500 找到以 r 处为根的子树,其中该子树的叶节点具有至少为 s 的累积权重且具有最小高度。FindMinHeightTree(T, r, s) 进程 500 执行对树的顺次遍历,并维持每个节点的高度和权重以及该具有至少为 s 的权重的最小高度子树的

根。

[0102] 现在返回到图 3, 需要注意的是, 对数据中心的物理资源的选择可通过其它方式来执行(例如使用其它信息、以满足其它条件的方式、使用其它进程等, 以及其各种组合)。

[0103] 在步骤 350, 针对每个所请求的 VM, 确定将所请求的 VM 向所请求的 VM 所被指派的物理资源的指派。将 VM 向数据中心的物理资源的指派包括将 VM 向托管 VM 的机架、刀片式服务器和 CPU (以及可选的, CPU 上的核) 的指派。

[0104] 可使用用户 VM 请求信息、数据中心管理信息等中的一个或多个以及其各种组合将所请求的 VM 指派到所述数据中心的物理资源。在一种实施方式中, 以趋于最小化数据中心内的机架间流量的方式将 VM 指派到数据中心的物理资源。在一种实施方式中, 可根据用户 VM 请求中指明一个或多个 VM 指派限制执行将 VM 向将托管 VM 的数据中心的物理资源的指派。在一种实施方式中, 例如其中用户 VM 请求信息包括与向不同数据中心指派不同类型的 VM 相关联的 VM 类型限制, 可基于 VM 类型限制执行将 VM 向将托管 VM 的数据中心的物理资源的指派。

[0105] 在一种实施方式中, 使用被配置用于向数据中心的物理资源指派 VM 的 VM 指派进程执行将 VM 向将托管 VM 的数据中心的物理资源的指派。一般来讲, VM 指派进程的目标是将图分成多个等尺寸的部分, 从而图的这些部分之间存在很少连接(由此, 最小化支持将在数据中心中托管的所请求的 VM 所需的网络通信资源)。然而, 即使将图分成两个相等部分的最简单情况(典型的被称作图二分问题)也是 NP 困难的。此外, 相关问题是 K-切割问题, 其中目标是将图分成 K 个部分, 其中每部分的尺寸也是被输入所指明的。

[0106] 在 VM 指派进程中, 用户 VM 请求可被看作需要被分割的图。图的节点表示 VM, 且图的链接表示由链接所连接的 VM 对之间所需的相关带宽。对图的每个部分表示需要被指派到相同机架的 VM 的集合。每个部分的尺寸需要由机架中用于托管 VM 的可用资源的数量作为上限。与传统图分割问题不同, 该问题可以不指明每个部分中的节点的具体数量; 反而, 其只指明每个部分中的最大节点数。这是由于, 在任一点, 在云环境中针对托管 VM 的可用资源的数量要多于由用户所请求的 VM 的数量。结果, VM 指派进程可通过在针对托管 VM 的具有最大的可用资源的机架之内指派 VM 来最优化(机架间的)通信。为了在数据中心内指派 VM, 机架间的数据中心内通信成本取决于数据中心的开关分层中的机架的临近性。从而, 为了最小化应用的通信成本, 将机架部分之间的通信距离考虑在内。

[0107] 关于 VM 指派进程, 所述问题可概括如下。给定具有指派的边权重 $w: E \rightarrow N$ 的图 $G=(V, E)$, 以及具有将部分 i 和 j 之间的距离表示为 $d(i, j)$ 的部分容量 k_1, k_2, \dots 的集合。将 V 分成分离的集合 C_1, C_2, \dots, C_m , 以使得 $|C_i| \leq K_i$ 以及 $\sum_{i=1}^k \sum_{j=i+1}^k \sum_{v_1 \in C_i, v_2 \in C_j} w(v_1, v_2) d(i, j)$ 。

[0108] 关于 VM 指派进程, 需要注意的是, 要么可以以递增的方式分离地处理单独的用户请求, 要么可将所有多重请求合并在一起以形成大的分割图, 其中该分割图在不同用户请求的节点之间不包含任何边。

[0109] 关于 VM 指派进程, 需要注意的是, 图分割问题是 NP-困难的, 并且没有任何进程具有良好的(例如常量的)近似保证。

[0110] 在一种实施方式中, VM 指派进程被实施为贪婪启发进程。

[0111] 在一种实施方式中, VM 指派进程一次选择一个机架。可以以任何适当的顺序来选

择机架(例如执行每个选择以使得选择具有针对托管 VM 的可用的资源的最大数量的机架、基于在用户 VM 请求中指明的一个或多个 VM 指派限制等,以及其各种组合)并向所选的机架指派尽可能多的 VM。VM 指派进程选择具有参数(例如进入 / 离开的带宽、邻居数量等)的最大值的 VM,并向所选的机架指派所选的 VM。所选的 VM 被添加到集合 C,其是指派到所选的机架的 VM 的集合。VM 指派进程随后考虑 C 的所有邻居并选择具有到往 / 离开 C 的最大流量的 VM,并且将所选的 VM 添加到集合 C。可重复该进程直到所选机架的所有可用资源已被耗尽。VM 指派进程随后选择下一个机架(具有可用于托管 VM 的资源的最大数量)并针对该下一个所选机架和尚未被指派的剩余 VM 的图(即 V-C)重复所述进程。

[0112] 在一种实施方式中,通过交换特定节点来修改 VM 指派进程。VM 指派进程考虑在不同的部分中存在的节点对。VM 进程可确定交换所考虑的节点是否会改进所述方案(例如使用较少的机架间带宽)。VM 进程可确定将节点在部分之间移动(例如从其当前部分移动到具有可用容量的不同部分)是否会改进所述方案。VM 指派进程选择一个或多个所考虑的移动并进行所选的移动。注意的是,该进程可被重复(例如,直到达到了移动的阈值数量、直到所有移动已被考虑且 VM 指派中没有任何进一步的改进空间,等等)

[0113] 在步骤 360,方法 300 结束。

[0114] 虽然这里主要在其中虚拟资源是虚拟机(VM)和物理资源是刀片式服务器的处理器的实施方式的语境中进行说明和描述,但是应该注意的是虚拟资源指派能力可被用于确定对各种其他类型的虚拟资源的指派和 / 或确定对各种其他类型物理资源的指派。从而,这里使用的特定于对 VM 进行指派的各种术语可被更一般地理解为对虚拟资源的指派的提及。例如,这里对 VM 和 VM 指派的提及分别可被更一般地理解为虚拟资源和虚拟资源指派。例如,这里对在 VM 指派期间的 VM 指派限制(例如 VM 放置限制、VM 通信限制、VM 恢复限制、VM 类型限制等)的提及可被更一般地理解为在虚拟资源指派期间使用的虚拟资源指派限制(例如虚拟资源放置限制、虚拟资源通信限制、虚拟资源恢复限制、虚拟资源类型限制等)。需要注意的是,也可通过类似的方式被更一般地理解各种其它 VM- 特定术语。

[0115] 图 6 描述了适用于在执行这里描述的功能中使用的计算机的高级框图。

[0116] 如图 6 中所示,计算机 600 包括处理器元件 602 (例如中央处理单元(CPU)和 / 或其它适当的处理器)和存储器 604 (例如随机存取存储器(RAM)、只读存储器(ROM)等)。计算机 600 还可包括合作模块 / 进程 605 和 / 或各种输入 / 输出装置 606 (例如用户输入装置(比如键盘、小键盘、鼠标等)、用户输出装置(比如显示器、扩音器等)、输入端口、输出端口、接收机、发射机、和存储装置(例如磁带机、软驱、硬盘驱动、光驱等))。

[0117] 应该理解的是,这里说明和描述的功能可被实施为软件(例如经由在一个或多个处理器上的软件实现)和 / 或硬件(例如使用通用计算机、一个或多个应用特定集成电路(ASIC)、和 / 或其它硬件等同)。

[0118] 应该理解的是,这里说明和描述的功能可被实施在用于在通用计算机上执行的软件中(例如通过由一个或多个处理器来执行),从而实施特定目的计算机,和 / 或可被实施在硬件中(例如使用一个或多个应用特定集成电路(ASIC)和 / 或一个或多个其它硬件等同)。

[0119] 在一种实施方式中,合作进程 605 可被加载到存储器 604 中且被处理器 602 执行以实施按这里所述的功能。从而,合作进程 605 (包括相关联的数据结构)可被存储在计算机可读存储媒介(例如 RAM 存储器、磁或光驱或盘等)上。

[0120] 应该理解的是,图 6 中说明的计算机 600 提供适于实施这里所描述的功能性元素和 / 或功能性元素的部分的一般架构和功能。例如,计算机 600 提供适于实施刀片式服务器 114、CMS120、NMS130 和 CAS140 中的一个或多个的一般架构和功能。

[0121] 可以设想,这里作为软件方法进行讨论的一些步骤可在硬件中实现,例如作为与处理器合作执行各种方法步骤的电路。这里描述的功能 / 元素的部分可被实施为计算机程序产品,其中计算机指令(当被计算机处理器时)修改计算机的操作,以使得这里描述的方法和 / 或技术被调用或以其他方式提供。用于调用创造性方法的指令可被存储在固定的或可移除的媒介中,经由广播或其它信号承载媒介中的数据流进行传送,和 / 或存储在根据指令操作的计算装置内的存储器内。

[0122] 权利要求中说明了各种实施方式的方面。各种实施方式的这些和其它方面在下列编号条款中进行说明:

[0123] 1. 一种设备,包括:

[0124] 处理器和存储器,所述处理器被配置为:

[0125] 接收针对虚拟资源的请求,其中所述请求包括所请求的虚拟资源的数量和一个或多个虚拟资源指派限制;

[0126] 接收与所述云环境相关联的管理信息;以及

[0127] 使用所述所请求的虚拟资源的数量、所述一个或多个虚拟资源指派限制、以及所述管理信息确定将所请求的虚拟资源在所述云环境内的指派。

[0128] 2. 根据条款 1 所述的设备,其中所述虚拟资源包括虚拟机和虚拟存储资源中的至少一个。

[0129] 3. 根据条款 1 所述的设备,其中所述云环境包括多个数据中心,其中所述管理信息包括以下中的至少一个:

[0130] 网络管理信息,其与所述云环境的所述数据中心之间的数据中心间通信相关联;以及

[0131] 云管理信息,其包括针对所述数据中心中的至少一个与所述数据中心相关联的管理信息。

[0132] 4. 根据条款 1 所述的设备,其中所述一个或多个虚拟资源指派限制包括每个机架将被指派的虚拟资源的最大量、每个机架将被指派的虚拟资源的最小量、将用于每个数据中心的机架的最大数量、将用于每个数据中心的机架的最小数量、每个数据中心将被指派的虚拟资源的最大数量、每个数据中心将被指派的虚拟资源的最小数量、将被使用的数据中心的最大数量、以及将被使用的数据中心的的最小数量中的至少一个。

[0133] 5. 根据条款 1 所述的设备,其中确定将所请求的虚拟资源在所述云环境内的指派包括:

[0134] 从在所述云环境内的可用数据中心的集合确定将被指派所述虚拟资源的数据中心的集合;

[0135] 确定将所述虚拟资源向将被指派所述虚拟资源的所述数据中心的集合的指派;

[0136] 针对将被指派所述虚拟资源的所述数据中心的集合的至少一个数据中心,选择所述数据中心的物理资源,以托管将被指派到所述数据中心的所述虚拟资源;以及

[0137] 针对所述虚拟资源中的至少一个,确定将所述虚拟资源向将被指派所述虚拟资源

的数据中心的所述物理资源的指派。

[0138] 6. 根据条款 5 所述的设备,其中所述确定将被指派所述虚拟资源的所述数据中心的集合包括:

[0139] 将所述云环境表示为图,其中所述图包括表示所述可用的数据中心且具有与之相关联的多个权重的多个顶点,其中所述图包括表示所述可用的数据中心之间的多个通信路径且具有与之相关联的多个长度的多个边;

[0140] 确定所述图的子图,其中该子图具有至少等于所请求的虚拟资源的所述量的权重和以及最小直径;以及

[0141] 从所述子图确定将被指派所述虚拟资源的所述数据中心的集合。

[0142] 7. 根据条款 5 所述的设备,其中确定将所述虚拟资源向将被指派所述虚拟资源的所述数据中心的集合的指派包括:

[0143] 将所述针对虚拟资源的请求表示为图,其中所述图包括多个顶点和多个边,其中所述顶点表示所述虚拟资源且所述边表示由所述边所连接的所述虚拟资源之间的相关带宽;

[0144] 将所述图分割成多个部分以由此形成分割图,其中所述分割图的每个部分分别表示将分别在相同的数据中心中调度的所述虚拟资源的集合,其中所述分割图的每个部分具有与之关联的尺寸,该尺寸以分别在所述数据中心中可用的虚拟资源的数量作为上限;以及

[0145] 使用所述分割图来确定将所述虚拟资源向将被指派所述虚拟资源的所述数据中心的集合的指派。

[0146] 8. 根据条款 5 所述的设备,其中选择所述数据中心的物理资源以托管将被指派到所述数据中心的所述虚拟资源包括:

[0147] 将所述数据中心表示为树型拓扑,其中所述树型拓扑包括表示所述数据中心的资源的多个节点,其中所述树型拓扑的所述节点中的每一个具有与之相关联的权重,该权重表示以所述节点为根的虚拟资源的数量,其中所述树型拓扑的所述节点中的每一个具有与之相关联的高度,该高度表示所述树型拓扑内的所述节点的高度,其中所述多个节点包括分别表示所述数据中心的多个机架、刀片式服务器或处理器的多个叶节点;

[0148] 确定所述树型拓扑的子树,其叶节点具有至少等于将被指派到所述数据中心的虚拟资源的所述量的累积权重和最小高度;以及

[0149] 从所述子树确定将托管将被指派到所述数据中心的所述虚拟资源的所述数据中心的物理资源。

[0150] 9. 根据条款 5 所述的设备,其中确定将所述虚拟资源向将被指派所述虚拟资源的数据中心的所述物理资源的指派包括:

[0151] 选择将被指派所述虚拟资源的所述数据中心的机架;

[0152] 选择将被指派所述虚拟资源的所述数据中心的所选机架的刀片式服务器;以及

[0153] 选择将被指派所述虚拟资源的所述数据中心的所选机架的所选刀片式服务器的中央处理单元(CPU)。

[0154] 10. 根据条款 1 所述的设备,其中所述处理器还被配置为:

[0155] 基于所确定的在所述云环境内对所请求的虚拟资源的指派,发起在所述云环境内

对所述虚拟资源的供应。

[0156] 11. 一种存储指令的计算机可读存储媒介,其中所述指令被计算机执行时引起所述计算机执行方法,该方法包括:

[0157] 接收针对虚拟资源的请求,其中所述请求包括所请求的虚拟资源的数量和一个或多个虚拟资源指派限制;

[0158] 接收与所述云环境相关联的管理信息;以及

[0159] 使用所述所请求的虚拟资源的数量、所述一个或多个虚拟资源指派限制、以及所述管理信息确定将所请求的虚拟资源在所述云环境内的指派。

[0160] 12. 根据条款 11 所述的计算机可读存储媒介,其中所述虚拟资源包括虚拟机和虚拟存储资源中的至少一个。

[0161] 13. 根据条款 11 所述的计算机可读存储媒介,其中所述云环境包括多个数据中心,其中所述管理信息包括以下中的至少一个:

[0162] 网络管理信息,其与所述云环境的所述数据中心之间的数据中心间通信相关联;以及

[0163] 云管理信息,其包括针对所述数据中心中的至少一个与所述数据中心相关联的管理信息。

[0164] 14. 根据条款 11 所述的计算机可读存储媒介,其中所述一个或多个虚拟资源指派限制包括每个机架将被指派的虚拟资源的最大量、每个机架将被指派的虚拟资源的最小量、将用于每个数据中心的机架的最大数量、将用于每个数据中心的机架的最小数量、每个数据中心将被指派的虚拟资源的最大数量、每个数据中心将被指派的虚拟资源的最小数量、将被使用的数据中心的最大数量、以及将被使用的数据中心的最小数量中的至少一个。

[0165] 15. 根据条款 11 所述的计算机可读存储媒介,其中确定将所请求的虚拟资源在所述云环境内的指派包括:

[0166] 从在所述云环境内的可用数据中心的集合确定将被指派所述虚拟资源的数据中心的集合;

[0167] 确定将所述虚拟资源向将被指派所述虚拟资源的所述数据中心的集合的指派;

[0168] 针对将被指派所述虚拟资源的所述数据中心的集合的至少一个数据中心,选择所述数据中心的物理资源,以托管将被指派到所述数据中心的所述虚拟资源;以及

[0169] 针对所述虚拟资源中的至少一个,确定将所述虚拟资源向将被指派所述虚拟资源的数据中心的所述物理资源的指派。

[0170] 16. 根据条款 15 所述的计算机可读存储媒介,其中所述确定将被指派所述虚拟资源的所述数据中心的集合包括:

[0171] 将所述云环境表示为图,其中所述图包括表示所述可用的数据中心且具有与之相关联的多个权重的多个顶点,其中所述图包括表示所述可用的数据中心之间的多个通信路径且具有与之相关联的多个长度的多个边;

[0172] 确定所述图的子图,其中该子图具有至少等于所请求的虚拟资源的所述量的权重和以及最小直径;以及

[0173] 从所述子图确定将被指派所述虚拟资源的所述数据中心的集合。

[0174] 17. 根据条款 15 所述的计算机可读存储媒介,其中确定将所述虚拟资源向将被指

派所述虚拟资源的所述数据中心的集合的指派包括：

[0175] 将所述针对虚拟资源的请求表示为图，其中所述图包括多个顶点和多个边，其中所述顶点表示所述虚拟资源且所述边表示由所述边所连接的所述虚拟资源之间的相关带宽；

[0176] 将所述图分割成多个部分以由此形成分割图，其中所述分割图的每个部分分别表示将分别在相同的数据中心中调度的所述虚拟资源的集合，其中所述分割图的每个部分具有与之关联的尺寸，该尺寸以分别在所述数据中心中可用的虚拟资源的数量作为上限；以及

[0177] 使用所述分割图来确定将所述虚拟资源向将被指派所述虚拟资源的所述数据中心的集合的指派。

[0178] 18. 根据条款 15 所述的计算机可读存储媒介，其中选择所述数据中心的物理资源以托管将被指派到所述数据中心的所述虚拟资源包括：

[0179] 将所述数据中心表示为树型拓扑，其中所述树型拓扑包括表示所述数据中心的资源的多个节点，其中所述树型拓扑的所述节点中的每一个具有与之相关联的权重，该权重表示以所述节点为根的虚拟资源的数量，其中所述树型拓扑的所述节点中的每一个具有与之相关联的高度，该高度表示所述树型拓扑内的所述节点的高度，其中所述多个节点包括分别表示所述数据中心的多个机架、刀片式服务器或处理器的多个叶节点；

[0180] 确定所述树型拓扑的子树，其叶节点具有至少等于将被指派到所述数据中心的虚拟资源的所述量的累积权重和最小高度；以及

[0181] 从所述子树确定将托管将被指派到所述数据中心的所述虚拟资源的所述数据中心的物理资源。

[0182] 19. 根据条款 15 所述的计算机可读存储媒介，其中确定将所述虚拟资源向将被指派所述虚拟资源的数据中心的所述物理资源的指派包括：

[0183] 选择将被指派所述虚拟资源的所述数据中心的机架；

[0184] 选择将被指派所述虚拟资源的所述数据中心的所选机架的刀片式服务器；以及

[0185] 选择将被指派所述虚拟资源的所述数据中心的所选机架的所选刀片式服务器的中央处理单元(CPU)。

[0186] 20. 根据条款 11 所述的计算机可读存储媒介，其中所述处理器还被配置为：

[0187] 基于所确定的在所述云环境内对所请求的虚拟资源的指派，发起在所述云环境内对所请求的虚拟资源的供应。

[0188] 21. 一种方法，包括：

[0189] 针对以下步骤使用处理器和存储量：

[0190] 接收针对虚拟资源的请求，其中所述请求包括所请求的虚拟资源的数量和一个或多个虚拟资源指派限制；

[0191] 接收与所述云环境相关联的管理信息；以及

[0192] 使用所述所请求的虚拟资源的数量、所述一个或多个虚拟资源指派限制、以及所述管理信息确定将所请求的虚拟资源在所述云环境内的指派。

[0193] 虽然加入了本发明的教导的各种实施方式在这里被详细示出和描述，但是本领域技术人员能够轻易想出仍能加入这些教导的许多其它变形实施方式。

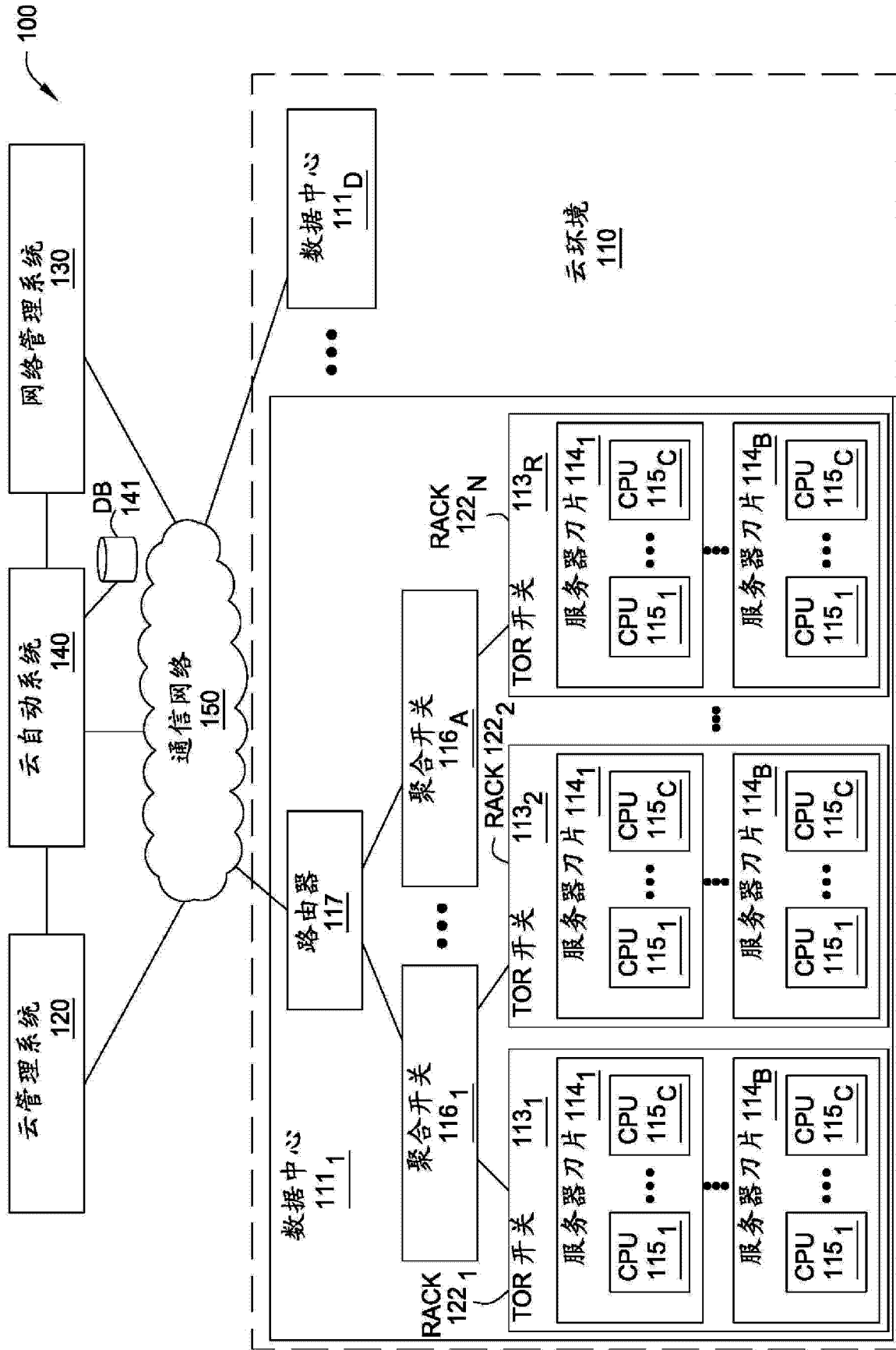


图 1

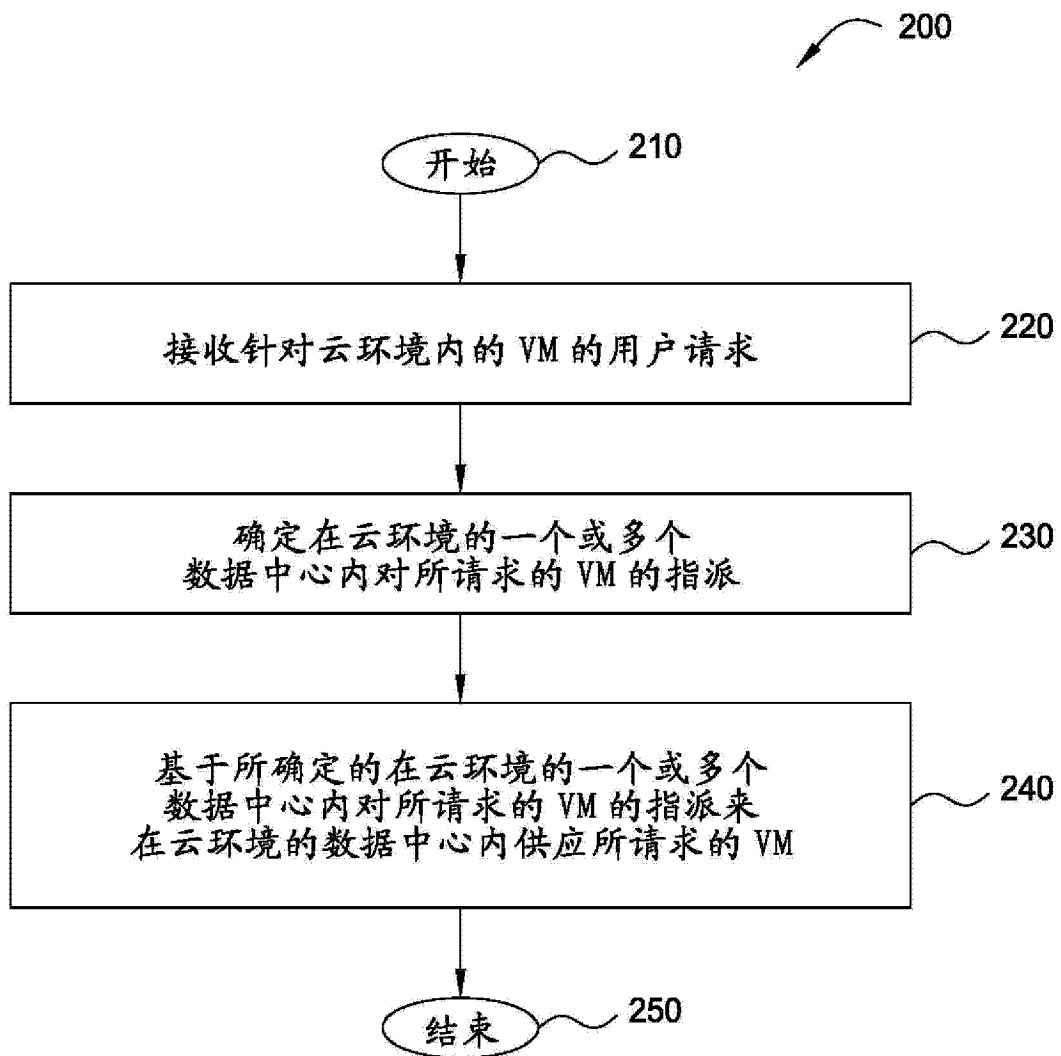


图 2

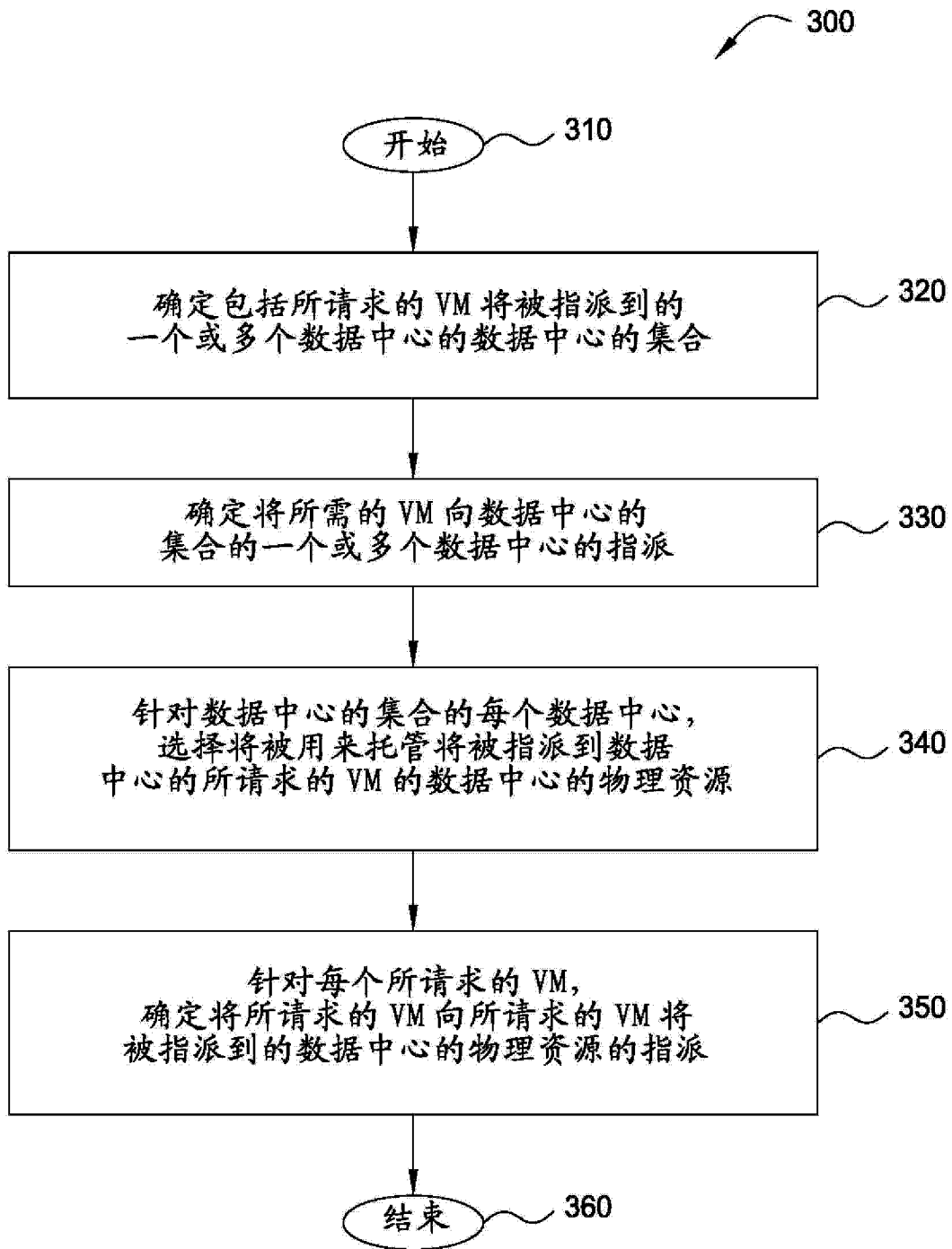


图 3

FindMinStar(G, v, s) 进程

输入: $G=(V,E,w,l)$: 具有顶点权重和边长度的完全图

v : 起始顶点

s : 子图的权重

输出: 由 v 和其最近邻居构成的权重至少为 s 的子图 $G'=(V',E')$ 。还返回直径。

$V' = \{v\}$

size = $w(v)$

Let u_1, u_2, \dots be the vertices of G sorted in increasing order of length to v . That is $l(v, u_i) \leq l(v, u_{i+1})$

$i=1$

diameter = 0

while (size < s)

do

$V' = V' \cup \{u_i\}$

size = size + $w(u_i)$

$i=i+1$

diameter = max(diameter, $\{l(v', u_i): \text{for all } v' \text{ in } V' - \{u_i\}\}$)

$E' = E' \cup \{(v', u_i): \text{for all } v' \text{ in } V' - \{u_i\}\}$

done

if ($i > n$)

Weight of G is less than s . Hence, no subgraph of size s exist in G .

Return NULL

Return $G'=(V',E')$ and diameter

410

图 4A

MinDiameterGraph(G, s) 进程

输入: $G=(V,E,w,l)$: 具有顶点权重、边长度的完全图
s: 子图的权重

输出: 权重至少为 s 的子图 $G'=(V',E')$ 。还返回直径。

```
mindiameter = ∞
For each vertex v in V
do
  G'' = FindMinStar(G, v, s)
  if (diameter(G'') < mindiameter)
  then
    mindiameter = diameter(G'')
    G' = G''
  endif
done
```

Return G' and mindiameter



420

图 4B

FindMinHeightTree(T, r, s) 进程

输入: T: 数据中心资源的树表示

r: 开始搜索的子树的根

s: 子树的所需权重

输出: 具有最小高度的权重至少为 s 的子树的根。

还计算当前子树的权重和高度。

```

if ( r is a leaf node)
then
    height(r) = 1
    if (weight(r) ≥ s)
        return r
    else
        return NULL
endif

height = 0
weight = 0
minheight = ∞
mintree = NULL
for each vertex v in children(r)
do
    n = FindMinHeightTree(T, v, s)
    if (n != NULL && minheight < height(n))
        minheight = height(n)
        mintree = n
    if (height(v) > height)
        height = height(v)
    weight = weight + weight(v)
done

height(r) = height + 1;
weight(r) = weight;

if ( mintree = NULL && weight ≥ s)
    return r
else
    return mintree

Return G' and mindiameter

```



500

图 5

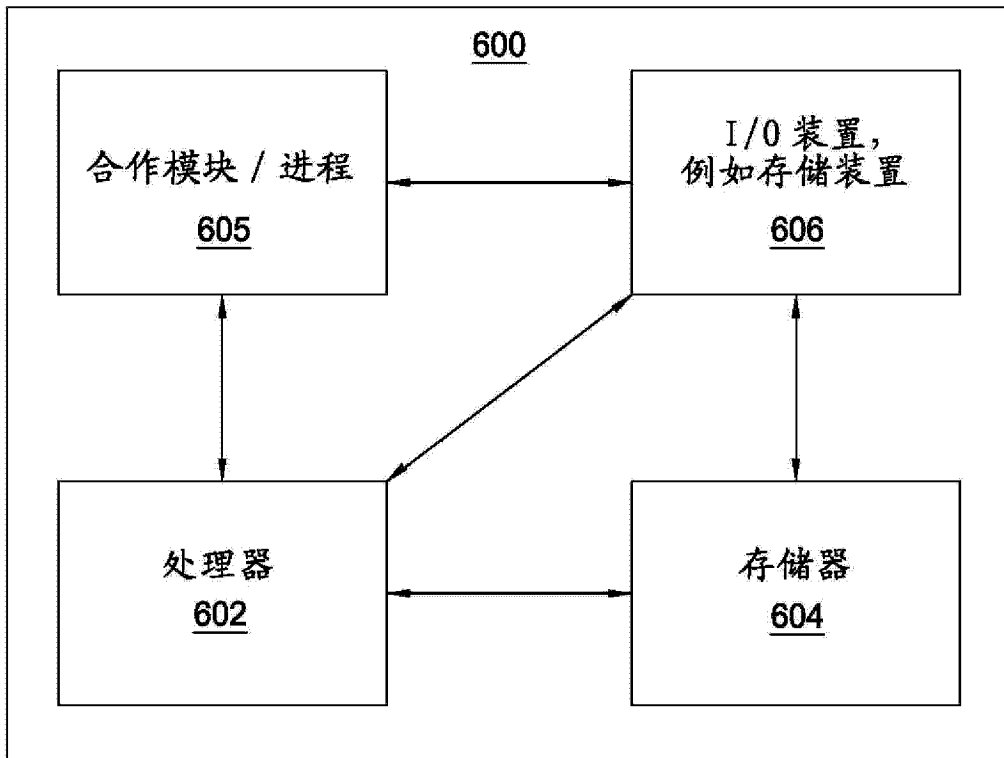


图 6