(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 104160387 A
(43)申请公布日 2014.11.19

(21)申请号 201380005083.7
(22)申请日 2013.01.07
(30)优先权数据
13/346,375 2012.01.09 US
(85)PCT国际申请进入国家阶段日
2014.07.09
(86)PCT国际申请的申请数据
PCT/US2013/020447 2013.01.07
(87)PCT国际申请的公布数据
WO2013/106258 EN 2013.07.18
(71)申请人 微软公司
地址 美国华盛顿州
(72)发明人 B·G·卡尔德 J·王 V·贝德卡 S·桑卡兰 M·米克尼特二世 P·K·贡达 Y·张 S·安东尼 K·曼尼瓦纳 H·卡特里
(74)专利代理机构 上海专利商标事务所有限公司 31100
代理人 罗婷婷
(51)Int.Cl.
G06F 15/16(2006.01)
G06F 9/06(2006.01)

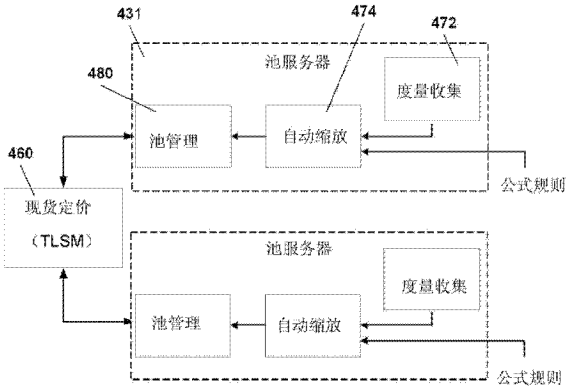
权利要求书2页 说明书19页 附图15页

(54)发明名称

对虚拟机池中的资源的定价

(57)摘要

本发明提供了用于经由现货定价过程来分派云计算环境中的资源的系统和方法。该现货定价过程允许基于与池相关联的投标以可抢占的基础向这些池分派虚拟机。可使用这些投标来确定针对可抢占虚拟机的分派的价格。随后,至少部分地基于与所确定的价格有关的所提交投标来将各可抢占虚拟机分派给池。



CN 104160387 A

1. 一种用于提供云计算环境中的资源的方法，包括：

接收针对可抢占虚拟机的分派的第一价格；

基于接收到的第一价格以及与虚拟机池相关联的第一投标将来自一个或多个虚拟机群集的多个可抢占虚拟机分派给所述虚拟机池；

在所分派的多个可抢占虚拟机上执行一个或多个任务；

接收对可抢占虚拟机的分派的第二价格；

基于接收到的第二价格和与所述虚拟机池相关联的第二投标将来自所述一个或多个虚拟机群集的至少一个可抢占虚拟机和来自附加虚拟机群集的至少一个可抢占虚拟机分派给所述虚拟机池；以及

在所分派的来自所述附加机器群集的至少一个可抢占虚拟机上执行一个或多个任务。

2. 如权利要求 1 所述的方法，其特征在于，来自所述附加虚拟机群集的虚拟机对应于在相对于对应于所述一个或多个虚拟机群集的物理机器而言分开的地理位置中的物理机。

3. 如权利要求 1 或 2 所述的方法，其特征在于，还包括基于来自所述第二虚拟机池的请求将来自所述一个或多个虚拟机群集的至少一个虚拟机分派给第二虚拟机池，所述请求包括对所述一个或多个虚拟机群集中的一虚拟机群集的亲和力，将来自附加虚拟机群集的至少一个可抢占虚拟机分派给所述虚拟机池响应于将来自所述一个或多个虚拟机群集的至少一个虚拟机分派给第二虚拟机池。

4. 如权利要求 1-3 中的任一项所述的方法，其特征在于，进一步包括：

聚集对应于多个虚拟机池的投标，每一投标包括所请求的多个可抢占虚拟机；

确定多个虚拟机可用于作为可抢占虚拟机来进行分派；

基于所聚集的投标来计算全局现货价格，所述全局现货价格被计算以使得所包括的具有大于所述全局现货价格的投标的所请求可抢占虚拟机的数目小于或等于所确定的虚拟机数目；以及

将计算出的全局现货价格作为针对可抢占虚拟机的分派的价格分发给所述多个虚拟机池。

5. 一个或多个存储计算机可使用指令的计算机存储介质，当所述指令由计算设备执行时执行一种用于提供云计算环境中的资源的方法，包括：

接收针对可抢占虚拟机的分派的价格；

基于接收到的价格和与第一虚拟机池相关联的第一投标将来自第一虚拟机群集的一个或多个可抢占虚拟机分派给所述第一虚拟机池，所述第一投标对应于包括对所述第一虚拟机群集的亲和力的对多个可抢占虚拟机的请求，其中所述对多个可抢占虚拟机的请求中的至少一个虚拟机在分派了所述第一虚拟机群集中的虚拟机之后未被履行；

基于接收到的价格和与第二虚拟机池相关联的第二投标将来自第二虚拟机群集的一个或多个可抢占虚拟机分派给所述第二虚拟机池，其中来自所述第二虚拟机群集的至少一个所分派的虚拟机是基于大于接收到的价格并小于与所述第一虚拟机池相关联的所述第一投标的投标而被分派给所述第二虚拟机池的；以及

在所述第二虚拟机池中的来自所述第二虚拟机群集的所分派的可抢占虚拟机上执行一个或多个任务。

6. 如权利要求 5 所述的计算机存储介质，其特征在于，至少一个未履行的虚拟机请求

在分派时间段内保持未被履行，并且来自所述第二虚拟机群集的至少一个虚拟机在所述分派时间段期间未被分派。

7. 如权利要求 5 或 6 所述的计算机存储介质，其特征在于，还包括：

聚集对应于多个虚拟机池的投标，每一投标包括所请求的多个可抢占虚拟机；

确定多个虚拟机可用于作为可抢占虚拟机进行分派；

基于所聚集的投标计算全局现货价格，所述全局现货价格被计算以使得大于所述全局价格的投标所包括的所请求的可抢占虚拟机的数目小于或等于所确定的虚拟机数目；以及

将计算出的全局现货价格作为针对可抢占虚拟机的分派的价格分发给所述多个虚拟机池。

8. 一种用于提供云计算环境中的资源的方法，包括：

接收针对可抢占虚拟机的分派的价格；

基于接收到的价格和与第一虚拟机池相关联的第一投标将来自一个或多个虚拟机群集的第一多个可抢占虚拟机分派给所述第一虚拟机池；

基于接收到的价格和与第二虚拟机池相关联的第二投标将来自所述一个或多个虚拟机群集的第二多个可抢占虚拟机分派给所述第二虚拟机池；

在所述第一多个可抢占虚拟机上以及所述第二多个可抢占虚拟机上执行一个或多个任务；

接收来自所述第一多个虚拟机池的增加可抢占虚拟机的数目的请求，该增加请求对应于与所述第一虚拟机池相关联的第三投标，所述第三投标大于与所述第二虚拟机池相关联的第二投标；

维护对所述第二多个虚拟机的分派直到分派时间段的结束；

从所述第二虚拟机池中移除对来自所述第二多个虚拟机的至少一个虚拟机的分派；以及

在后续分派时间段内将所移除的至少一个虚拟机分派给所述第一虚拟机池。

9. 如权利要求 8 所述的方法，其特征在于，对可抢占虚拟机的分派是在分派时间段的开始处执行的，针对可抢占虚拟机的分派的价格是在分派时间段的开始之前接收到的。

10. 如权利要求 9 所述的方法，其特征在于，可抢占虚拟机的数目的增加由所述第一虚拟机池在针对可抢占虚拟机的分派的价格被接收到之后并在所述后续分派时间段之前请求。

对虚拟机池中的资源的定价

[0001] 背景

[0002] 用于执行大规模计算作业的常规方法通常涉及用户购买计算机硬件来用作计算平台。这可导致各种低效，因为许多典型用户具有峰值水平的计算需求，该计算需求不同于对计算资源的常规需求。购买足够的硬件以满足峰值资源需求可导致对计算资源的低使用率。或者，使硬件与常规使用水平匹配可使得一些期望计算变得不切实际。最近，在处理速度和网络传输速度方面的改善已使得云计算环境变成本地计算平台的可行替换。

[0003] 概述

[0004] 在各实施例中，提供了用于经由现货定价过程来分派云计算环境中的资源的系统和方法。现货定价过程允许基于与各池相关联的投标以可抢占的基础向这些池分派虚拟机。投标可用于确定针对可抢占虚拟机的分派的价格。随后可至少部分地基于所提交的与所确定价格相关的投标来向池分派可抢占虚拟机。

[0005] 提供本概述是为了以简化的形式介绍将在以下具体实施方式中进一步描述的概念选择。该概述不旨在标识所要求保护的主题的关键特征或本质特征，也不旨在独立地用于帮助确定所要求保护的主题的范围。

[0006] 附图简述

[0007] 下面将参考附图详细描述本发明，其中：

[0008] 图 1 示意性地示出适于在实现云计算环境时使用的系统或组件的示例。

[0009] 图 2 示意性地示出适于在实现云计算环境时使用的系统或组件的示例。

[0010] 图 3 示意性地示出适于在实现云计算环境时使用的系统或组件的示例。

[0011] 图 4 示意性地示出适于在实现云计算环境时使用的系统或组件的示例。

[0012] 图 5 示意性地示出适于在实现云计算环境时使用的系统或组件的示例。

[0013] 图 6 示意性地示出适于在实现云计算环境时使用的系统或组件的示例。

[0014] 图 7-11 示意性地示出根据本发明的一实施例的管理云计算环境中的虚拟机的示例。

[0015] 图 12 示意性地示出适于执行本发明的各实施例的计算设备。

[0016] 图 13-15 示出根据本发明的过程流的示例。

[0017] 详细描述

[0018] 概览

[0019] 由于通过网络的数据传输速度的增加和其他网络特征的改善，越来越有可能在其中计算资源分布在大型网络上的环境中执行大规模计算任务。处于第一位置的用户可向计算服务提交作业或计算任务，并且使该任务在该用户不直接知晓的计算机群组上执行。用于执行该用户任务的计算资源可分布在多个位置上。位于一个或多个位置的第一计算资源群组可存储用于执行该用户的计算任务的数据和其他信息，而位于相同位置或可能位于一组不同的一个或多个位置的第二计算资源群组可被用于执行该计算任务。

[0020] 对各种分布式计算资源的访问允许用户执行作业任务，而无需关心这些计算资源位于何处。分布式资源还为用户提供了扩展 (scale out)(或收缩 (scale in)) 所使用的

资源量以便满足计算任务的目标（诸如用指定时间完成该计算任务）的机会。然而，为用户提供这种灵活性对分布式计算资源的运营者（或拥有者）提出了许多挑战。为了满足需求，分布式资源网络的运营者将优选地具有足够多的可用资源来满足在峰值需求时间的资源请求。

[0021] 在峰值需求时具有足够资源的云计算环境将可能至少在非峰值需求时段期间并且可能在所有时间都具有多余的虚拟机可用。多余的虚拟机可表示允许基于用户请求来扩展作业的虚拟机预留、对资源故障进行补偿的虚拟机预留、或仅峰值需求期间被使用但在非峰值时并非作为专用资源被需要的虚拟机。拍卖机制可用于允许用户为对这些多余的虚拟机的临时访问进行投标，而不是允许这些资源空闲。这向消费者提供了以较低成本对机器的访问，同时允许云计算运营商最大化满足峰值需求和 / 或系统冗余要求所需的资源的值。

[0022] 定义

[0023] “账户”是云计算环境内的全局唯一标识实体。在一实施例中，下面讨论的所有资源和任务均在一账户范围内。通常，用户将在使用云计算系统的资源之前首先创建账户。在创建账户之后，用户可以使用该账户向该系统提交工作项并且基于这些工作项来管理用于执行作业的资源。

[0024] “工作项”是要在云计算环境中运行的作业的静态表示。工作项可以指定作业的各个方面，包括作业二进制代码、指向要处理的数据的指针、以及可任选的启动用于执行该作业的任务的命令行。此外，工作项可以指定重现时间表、优先级和约束。例如，工作项可指定要在每天下午 5 点启动。

[0025] “作业”是工作项的运行实例。作业包含携手执行分布式计算的任务集合。这些任务可在云计算环境中的一个或多个虚拟机上运行。

[0026] “任务”是作业的基础执行单元。每个任务在一虚拟机上运行。用户可以为每个任务指定给命令行的附加输入以及到输入数据的指针。在任务的执行过程期间，该任务可以在其在执行该任务的虚拟机上的工作目录下创建文件的分层结构 (hierarchy)。

[0027] “作业管理器任务”(JM 任务）是作业中的特殊任务。作业管理器任务是可任选的，因此一些作业可在不使用 JM 任务的情况下执行。作业管理器任务可为作业内的所有任务提供单一控制指针，并且可被用作该作业的“主”任务。如果一作业具有 JM 任务，则该系统将该 JM 任务作为该作业的第一个任务启动。该 JM 任务随后可以向该作业提交更多任务，并且它可以监视这些任务的进程并且控制何时提交下一批次任务。以此方式，JM 任务可以协调对作业中的所有任务的调度并且管理各任务间的依赖关系。优选地，如果该作业管理器任务的节点或虚拟机发生故障，则可以自动在另一虚拟机上重启该 JM 任务，以便该 JM 任务总是为该相应作业而运行。此外，用户可以向该系统指定 :一旦该 JM 任务完成，该系统可以终止相应作业中的所有任务。

[0028] 虚拟机池和任务承租者

[0029] 虚拟机指的是处理能力的逻辑单元。虚拟机可与物理处理器具有一对一的对应关系，或者一虚拟机可对应于多个处理器，或者虚拟机可表示在一个或多个处理器上的处理时间的百分比。向池分派的虚拟机可在任何给定时间执行针对该池的一个或多个任务。

[0030] 在各实施例中，可潜在地基于工作项来执行作业的各虚拟机在使用之前与该工作

项的账户相关联。“池”是虚拟机的逻辑群组。工作项总是具有至少一个相关联的池来运行与该工作项相对应的（诸）作业。每个账户可创建该账户获得访问权的一个或多个池，以供在执行与该帐户相关联的工作项时使用。通常，账户对与该账户相关联的各池具有排他访问权。可在用户提交工作项时创建池，或者可以将工作项与现有的池相关联。池可被系统自动创建以执行作业。例如，在每天的特定时间运行的重现工作项可通过使池被自动创建以在开始时间执行该作业来处理。每天在完成该重现工作项之后可删除该池。可任选地，可关联一池以与单个工作项、单个作业、或与一账户相对应的工作项的另一子集一起使用。

[0031] 当用户提交一工作项时，可将该工作项与一个或多个虚拟机池相关联。可以按照任何方便的方式将各虚拟机组织在池内。例如，无论虚拟机的底层处理器的地理位置如何，都可以将所有虚拟机组织在单个池中。另一选项是基于地理位置来组织虚拟机，使得池的所有虚拟机均处于给定地理位置。又一选项是按照不同于地理位置的基础来组织虚拟机，诸如对其他变量（例如，存储资源、网络等待时间、用户位置 / 偏好、安全性要求）的接近度。又一选项是在创建工作项或作业时自动创建池，并且随后在该工作项或作业结束时拆卸该池。

[0032] 虚拟机池表示一种用于组织虚拟机的方法。虚拟机的另一组织单元是虚拟机群集。虚拟机群集表示由云环境中的过程（例如任务承租者过程）一起管理的虚拟机群组。虚拟机群集中的虚拟机可对应于以方便的方式分组在一起的物理机。例如，虚拟机群集可对应于位于同一地理区（诸如，位于美国或美国东北部）的物理机群组；位于同一通用位置（诸如，位于像西雅图或圣地亚哥郡等城市或大都市区域）的物理机群组；或位于同一特定位置（诸如位于形成计算或数据中心的一个或多个连接的或附近的建筑）的物理机群组。另一选项是基于具有有利数据传输速率的物理机群组用云环境中的特定存储部分形成虚拟机群集。又一选项是基于处于给定位置的物理机来形成多个虚拟机群集。虚拟机池可跨越多个虚拟机群集。用于管理虚拟机群集（诸如任务承租者）的过程可分派和取消分派虚拟机池中的各虚拟机。任务承租者（或用于管理虚拟机群集的其他过程）也可基于与向群集内的虚拟机分派的池相对应的作业队列来调度该虚拟机上的任务。当任务承租者需要附加的机器以便向虚拟机池分派足够的数目时，该任务承租者可以从通用云计算环境获得附加虚拟机。类似地，如果任务承租者具有多余的虚拟机，则该任务承租者向通用云计算环境返回这些多余的机器。

[0033] 专用、备用、和可抢占 (preemptible) 机器

[0034] 在将虚拟机分派给池时，可将该虚拟机作为两个类型之一来分派。可将该虚拟机作为专用虚拟机或可抢占虚拟机分派给该池。当虚拟机处于池中时，该虚拟机作为专用还是可抢占的状态也可以变化。“专用”虚拟机是被分派给池的供被分派给该池的工作项 / 作业专门使用的机器。任选地，专用虚拟机可被分配以供一个或多个相关联的工作项专门使用，而不是一般性地可用于被提交给池的任何作业。在虚拟机具有专用状态时，该虚拟机被预留以供与该池相关联的账户使用。不向专用机器提供来自其他账户的资源，并且专用机器不会代表其他账户执行作业。

[0035] “可抢占”虚拟机是当前正代表帐户执行池中的任务、但是不保证该虚拟机将持续对该池可用的虚拟机。当可抢占虚拟机变得对池可用时，将该可抢占机器添加到那个池中。

随后可提供该可抢占机器，并使用该可抢占机器来执行该池的作业。该可抢占机器可通过任何方便的方法而变得对该池可用，诸如通过（代表相应帐户）使该池在资源拍卖中赢得在该可抢占虚拟机上的处理时间。

[0036] 分派专用虚拟机和可抢占虚拟机的附加因素是对虚拟机的请求是否包括对特定虚拟机群集的亲和力。对一虚拟机群集的亲和力可基于各种理由。对虚拟机群集的亲和力的请求的一个示例由具有以下虚拟机的期望和需要引起：该虚拟机具有对用于将在虚拟机上执行的作业的数据存储的改善的访问（例如，高数据传输速度）。对于这种类型的存储亲和力，该亲和力请求可指定对一个或多个虚拟机群集中具有对数据的所需访问的虚拟机的分派。这可表示例如与具有到数据存储中心的所需物理数据连接的物理机相对应的虚拟机群组。另一类型的亲和力是作业亲和力。由虚拟机执行的一些类型的作业可涉及各虚拟机之间对相同或类似作业起作用的大量通信。在作业亲和力场合中，使对某作业起作用的所有虚拟机都位于单个虚拟机群集（或其他虚拟机组织单元）中以便促成在这些虚拟机之间的消息传递可能是有利的。从单个虚拟机群集选择虚拟机可对应于选择与同一地理位置中的物理机相对应的虚拟机。

[0037] 变得可用于作为可抢占虚拟机用于一账户的虚拟机通常将是云计算环境中具有另一目的的虚拟机。例如，可抢占虚拟机的一个来源是云计算环境所有者／运营者出于灾难恢复目的而提供的虚拟机。为了提供稳定操作，云计算环境可包括被预留的一组或多组虚拟机。这些预留虚拟机可用于取代由于处理器故障、网络故障、或导致云环境的一部分不再适于执行作业的任何其他种类的事件而失去的资源。当被分配给一池的一个或多个专用虚拟机由于一事件而被失去时，可使用预留的虚拟机来取代该失去的虚拟机。这改善了云计算环境中的资源的可用性。然而，由于期望故障事件是稀少的，因此预留灾难恢复机器通常将意味着大量虚拟机空闲并等待被使用。可将这些虚拟机的被指定用于处理故障事件的CPU循环周期作为可抢占虚拟机分派给各池以运行工作项／作业，而不是浪费这些虚拟机的CPU循环周期。如果故障发生并且该系统需要撤走可抢占资源以满足专用资源的要求，则在这种虚拟机上运行的可抢占作业在行得通时就将被停止（并且可能立即被停止），以便该可占用虚拟机可被用于其取代失去或故障资源的原始目的。

[0038] 可抢占机器的另一来源是能力多余的虚拟机。通常，任何网络的峰值负载将不同于平均负载。结果是，具有足够资源来处理峰值负载情形的计算环境在其他时间通常将具有多余的资源可用。这些多余资源提供了资源缓冲。当用户作出对附加专用虚拟机的请求时，多余的虚拟机可用于满足用户的请求。当云计算环境具有小于专用机器的峰值负载的负载时，一个或多个虚拟机将空闲。可在抢占基础上将这些虚拟机的被指定用于提供空闲能力的CPU循环周期分派给用户和池，而不是浪费这些虚拟机的CPU循环周期。随着对专用虚拟机的请求的负载的增加，在这些多余虚拟机上运行的可抢占作业一行得通就将被停止（并且可能被立即停止）。这允许可抢占虚拟机被用于其在需要时提供附加专用资源的原始目的。附加地或替代地，专用机器的负载方面的一些增加将由所排定的对专用机器的请求引起。如果虚拟机将由于在经排定的时间被用作专用机器而变得不可用，则被分配给该虚拟机的可抢占作业可在该经排定的时间之前被停止以便允许从该可抢占作业到专用资源的有序转变。

[0039] 在一些情形中，用户可能期望在某个将来时间具有对更大量的专用机器的访问

权。在该情形下，用户可将一个或多个虚拟机作为备用虚拟机来预留。“备用”的虚拟机预留是与一池或账户相关联的、供要分派给该池或账户的虚拟机在将来的某时使用的预留。提供虚拟机以供使用可能仅意味着在云计算环境中标识和 / 或预留足够的虚拟机资源，使得虚拟机资源在被请求时将可用于转换成专用虚拟机。任选地，提供备用机器还可包括向虚拟机提供数据、可执行代码或其组合。

[0040] 备用虚拟机预留不是对虚拟机的分配或分派。相反，备用虚拟机预留会预留在将来使空闲或可抢占虚拟机转换成向与该备用预留相关联的用户或池分派的专用虚拟机的权力。可抢占作业可以是与关联于备用预留的池或账户、另一不同的池、或另一不同的账户相关联的作业。当备用预留由池或账户作出时，不将来自虚拟机群集的虚拟机分派给该池或账户。相反，保持与该虚拟机群集相对应的备用预留的数目的计数，使得足够数目的空闲或可抢占虚拟机可用于满足与该虚拟机群集相对应的备用预留。

[0041] 虚拟机备用预留可出于各种理由而与池相关联。备用机器预留的一个用途针对具有仅在特定时间帧期间发生的高优先级计算作业的用户。例如，金融公司可能希望对一个或多个金融市场的日常活动（诸如股票交易或商品交易）执行分析。金融市场可能按明确的时间表开市和闭市，诸如在上午 9:30 开市并在下午 4:00 闭市。金融公司想要在金融市场开市的时间聚集数据以用于执行分析或模拟。分析的目标是在下一天市场开始之前为他们的员工提供信息。这种分析可能需要大量的虚拟机，但是虚拟机仅在例如从下午 6:00 直到第二天早晨 3:30 的时间之间被需要。在此时间期间，金融公司期望保证虚拟机的可用性。在当天的其他时间，该金融公司不需要这些机器。将虚拟机预留与金融公司的账户关联可以实现这一目标。作为支付预留价格的交换，向金融公司保证机器在所需时间期间的可用性。在所需时间窗口之外，这些虚拟机可用作该金融公司和 / 或其他用户的可抢占机器。

[0042] 备用预留可用于根据基于时间的标准或基于负载的标准将空闲或可抢占虚拟机转换成向对应于用户的池分派的专用机器。在某些情形下，备用预留可导致至少部分地基于预定的时间和 / 或日期将空闲或可抢占虚拟机转换成专用虚拟机。在这样的情形下，基于备用预留将可抢占虚拟机转换成专用虚拟机可在排定的可用性事件之前按有序方式停止。这被定义成具有基于时间的标准的备用预留。基于时间的标准不同于基于负载的标准，基于负载的标准被用于定义基于负载的阈值。基于负载的阈值对应于基于一个或多个云资源的使用和 / 或性能的阈值。优选地，基于负载的阈值不包括对基于时间的标准的使用。除了基于时间的标准和基于负载的标准以外，将对应于备用预留的一个或多个虚拟机转换成专用虚拟机的又一选项是基于来自用户或系统管理员的请求。

[0043] 备用留的另一用途是允许在扩展作业时的改善的性能。例如，零售商店可以使用云计算资源来在假日之前处理购物季期间的附加在线流量，诸如用于浏览零售商的网站并下订单的在线流量。基于过去的经验，零售商期望一定水平的在线活动，并且预留相应数量的专用虚拟机。然而，在在线活动大于预期的情况下，零售商还经由备用预留来预留附加的机器。零售商随后可以设置一个或多个阈值，该一个或多个阈值指示高于预期的活动水平。随着这些阈值的出现，可使用备用预留来将空闲或可抢占虚拟机转换成专用机器以允许零售商处理此附加在线流量，而无需使零售商的顾客经历慢响应时间。在此情形中，备用预留可在不可预测的时间被转换为专用机器，因为它可能不知道活动阈值何时将被满足。当活动阈值被满足时，空闲或可抢占虚拟机被转换成向与备用预留相关联的池分派的专用虚拟

机。如果可抢占任务在转换之前正在虚拟机上运行，则该可抢占任务在将该虚拟机转换成专用机器之前被停止。可任选地，活动阈值不包括基于时间的标准。

[0044] 基于现货定价来分派可抢占机器

[0045] 云计算环境中未作为专用机器与池相关联的任何虚拟机都潜在地可用于经由现货定价来进行分派。因此，这些可用于经由现货定价来进行分派的虚拟机可包括当前正在运行可抢占作业的虚拟机、供在灾难恢复时使用的虚拟机、或任何其他多余或空闲的虚拟机。可用于作为可抢占虚拟机来进行分派的多余或空闲虚拟机可包括满足虚拟机群集的备用预留计数所需的空闲虚拟机。

[0046] 为了经由现货定价获得虚拟机，针对与账户相关联的池的规范可包括所需的多个可抢占虚拟机的规范。通常，该规范进一步包括该账户的用户为了获得一个或多个可抢占虚拟机而愿意支付的投标或价格。针对池的规范不限于该投标数目。例如，池规范可包括滑动级别的投标，其中第一（较高的）投标被提供以获得两个可抢占虚拟机，第二（中间范围的）投标被提供以获得三个附加的可抢占虚拟机，以及第三（较低的）投标被提供以获得最后的两个可抢占虚拟机。取决于该现货价格，这样的投标模式可导致向用户分派零个、两个、五个、或七个可抢占虚拟机。

[0047] 经由现货定价来分派可抢占机器可周期性地发生，其中每一分派导致在分派时间段内分派可抢占机器。优选地，各分派时间段可以是连续的，使得一个分派时间段的结束对应于下一分派时间段的开始。通常，在每一分派时间段的开始处或附近重新计算现货价格。现货价格在分派时间段期间保持不变。

[0048] 池可在任何时间提交针对可抢占虚拟机的投标。然而，不保证高于该现货价格的投标将立即导致基于该投标对可抢占机器的分派。如果池已提交了高于现货价格的投标并且有足够的虚拟机可用，则将在不晚于下一分派时间段的开始处分派所请求的可抢占虚拟机。如果对可抢占虚拟机的投标是在某分派时间段期间提交的，则仅在有多余的虚拟机可用的情况下，可立即分派机器。尤其地，具有较低投标的池可能已经被分派了可抢占虚拟机。来自另一池的较高投标可在分派时间段的开始处而非在中间时间处取代该较低投标。被分派了可抢占虚拟机的账户可仅仅由于该虚拟机被需要用于非抢占目的（诸如转换成专用机器或用作灾难恢复机器）而在分派时间段的中间部分期间失去该虚拟机。

[0049] 各分派时间段的长度可被设置成任何方便的值。例如，分派时间段可为至少约 15 分钟、或至少约 30 分钟、或另一方便的间隔。可任选地，分派时间段可在需要的情况下在一天的整个过程中变化，或者该时间段可在工作日对周末上改变，或者可引入该时段中的任何其他变化。优选地，分派时间段可在预定时间（诸如每半个小时）开始。

[0050] 用于确定可抢占虚拟机的现货价格的一个选项是确定全局现货价格。为了确定全局现货价格，聚集来自云计算环境中的所有机器池的投标。随后将该现货价格与在分派时间段的开始处可用的虚拟机的总数进行比较。该现货价格可随后被设置成对于至少为大于该现货价格的所有投标分派可抢占机器而言所必需的全局价格。如果较大数目的投标处于分派机器的临界点，则仅部分满足处于现货价格的那些投标，可以以任何方便的方式处理处于市场出清价格的那些投标。例如，现货价格可被设置为处于次高投标，使得向处于或大于该现货价格的所有投标授权所请求数目的可抢占机器。或者，现货价格可被设置为等于市场出清价格，其中处于市场出清价格的各投标潜在地接收仅一部分所请求的机器。

[0051] 虽然现货价格被全局地设置，但对可抢占虚拟机的分派被局部地处理，诸如在任务承租者层级和 / 或在虚拟机池群组层级处被处理。例如，全局确定的现货价格可被分发给任务承租者。各任务承租者可随后将在各自任务承租者内的可用虚拟机分派给由该任务承租者服务的机器池。这些分派可通过以下开始 :履行来自该承租者内的池的最高投标，随后履行次高投标，并以此类推。该过程可继续直到没有更多高于该现货价格的投标可用，或直到该任务承租者内没有更多虚拟机可用于作为可抢占虚拟机来分派。

[0052] 在一些情形中，在全局投标被计算的时间和发生对可抢占虚拟机的分派的时间之间可用的可抢占虚拟机资源的数目可改变。如果发生这种情况，则虚拟机群集（诸如，由任务承租者管理的那些机器）可能不具有足够的虚拟机来为高于该现货价格的所有投标分派可抢占虚拟机。在这种情形下，任务承租者可任选地尝试添加更多虚拟机。如果在云计算环境内有不与另一任务承租者相关联的任何多余虚拟机可用，则这些多余的虚拟机可被添加并使用来以高于现货价格的投标履行对可抢占虚拟机的附加请求。然而，适于并入给定虚拟机群集的各附加虚拟机可能并不可用，诸如由于缺少对存储区域具有类似访问的附加虚拟机和 / 或缺少处于相同地理位置的附加虚拟机。

[0053] 也有可能任务承租者将具有比满足大于现货价格的所有投标所需的更多的可抢占虚拟机。再次，以投标的次序将可抢占虚拟机分派给虚拟机池。在满足了高于现货价格的所有投标后，任务承租者可能仍有附加的可抢占虚拟机剩下。这可指示该任务承租者应向通用云计算环境返回一些虚拟机以供重新分派给其他任务承租者。即使有附加可抢占虚拟机可用，低于现货价格的投标仍不会收到可抢占虚拟机。

[0054] 在将可抢占虚拟机分派给虚拟机池后，该可抢占虚拟机保持被分派给该池，要么直到下一次拍卖，要么直到该虚拟机被需要用于抢占当前使用的另一目的。抢占临时使用的使用示例包括将该虚拟机转换成专用机器的需要，或将该虚拟机用于灾难恢复的需要。当可抢占虚拟机被抢占时，任务承租者可按从最低投标到最高投标的次序来抢占适合的虚拟机。在标识可抢占虚拟机以供抢占时可考虑的另一因素是作业已在该可抢占虚拟机上运行的时间长度。与已运行达多个分派时间段的作业相比，刚开始的作业是抢占的更好选择。这种类型的因素可例如用作针对基于相同投标价值来分派的可抢占作业的附加考虑。在各个实施例中，如果向账户分派的虚拟机在分派时间段期间被抢占，则不针对该分派时间段向该账户收费。然而，如果可抢占虚拟机是在分派时间段期间自愿释放的，则针对该时间段被使用的部分向该账户收费。

[0055] 分布式网络环境中的计算资源的组织的示例

[0056] 云计算环境的用户通常希望使用云计算资源来执行作业。这些作业通常涉及基于存储在可经由云计算环境来访问的各位置中的数据执行作业。运营商提供云计算环境的一种方式是将该环境作为多个层来提供。图 1 示意性地示出适于执行云计算环境中的任务的系统的示例。图 1 中的系统包括任务运行时层 110、第三方任务运行时层 120、资源管理层 130 以及调度和执行层 140。

[0057] 在图 1 中示出的实施例中，任务运行时层 110 负责为来自用户 105 的任务设置执行环境以及安全上下文。任务运行时层 110 还可启动任务并监视任务的状态。任务运行时层 110 可采用在每个虚拟机上运行的系统代理的形式。该任务运行时层还可包括链接到用户的任务可执行代码的运行时库。具有作为任务运行时层 110 的一部分的运行时库可潜在

地向由该系统代理执行的任务提供更丰富的能力。运行时库的示例包括：用于允许任务间的快速通信的一个或多个高效通信库；用于从其他虚拟机和/或其他任务读取文件的高效远程文件访问库支持；用于允许任务进行检查点（例如，进入到二进制大对象中）和恢复的检查点库；日志记录库；以及用于提供跨虚拟机池内执行给定任务的各虚拟机使用的分布式文件系统的库。

[0058] 第三方任务运行时层120允许附加的运行时被构建在任务运行时层110之上并任务运行时层110之上运行。第三方任务运行时层120还可提供用于协调作业的各任务的运行的附加能力。示例可包括属于用于提供跨虚拟机池内执行给定任务的各虚拟机使用的分布式文件系统的库的MapReduce（映射减少）运行时。这允许用户按针对用户的作业或任务定制的方式组织云计算环境。在一些实施例中，作业管理器任务可促成允许用户使用第三方运行时层来运行和/或控制云计算资源。

[0059] 资源管理层130涉及管理云计算环境中可用的计算资源。一个选项是使资源管理层130以三个不同的层级管理这些资源。在第一层级，资源管理层130管理与作业（即，工作项的执行）相关联的虚拟机的分配和解除分配以及存储在每个虚拟机上的与任务相关联的文件。在第二层级，将与作业相关联的虚拟机分组成各机器池。池可包含与一个或多个作业和/或工作项相关联的虚拟机。取决于该实施例，单个池可跨越多个虚拟机群集，诸如一数据中心中的所有虚拟机群集、跨一地理区中的多个数据中心的多个虚拟机群集，或跨多个地理区中的各数据中心的多个虚拟机群集。单个池可包含大量虚拟机，诸如数百万的虚拟机。这些虚拟机可被包含在大量池中，诸如多达数十亿个池中。在第三层级，资源管理层管理给定池群组中可用于与各作业或工作项关联的虚拟机的量。这允许对基于系统的当前负载使用的计算资源的量进行动态调整。此外，未被当前池群组使用的虚拟机可被释放回云计算环境以供并入其他池群组。

[0060] 在图1中示出的实施例中，调度和执行层140管理用户正执行的工作项、作业和任务。调度和执行层140作出调度决策并且负责启动作业和任务以及在故障时重试。这种调度和执行层140可包括用于管理各个层的作业和/或任务的组件。

[0061] 以上描述的层可以在多个地理位置处包括处理器的云计算环境中。图2示意性地示出不同位置处的处理器可如何被集成在单个云计算架构中的示例。

[0062] 在图2中，一个或多个任务承租者215可被用来管理虚拟机池。任务承租者215可维持一组虚拟机。一个或多个用户的作业可在任务承租者215内作为一个或多个虚拟机池的一部分的各虚拟机上运行。一个或多个任务承租者215可在给定地理区中使用。任务承租者215的责任可包括维护该组虚拟机并基于该任务承租者内的资源利用来动态地增长或收缩该任务承租者。这允许任务承租者215增加任务承租者内的虚拟机的数量以便容纳增加的顾客需求。这还允许任务承租者215释放未使用的虚拟机，以便能将这些虚拟机分配给数据中心中的处理其他顾客的服务的其他托管服务。任务承租者215的另一责任可以是实现池分配/解除分配/管理逻辑的一部分。这允许任务承租者215参与确定如何将虚拟机分派给与顾客的任务相关联的池。任务承租者215还可负责各任务在该任务承租者内的各虚拟机上的调度和执行。

[0063] 在图2的实施例中，提供了控制多个任务承租者215的一个或多个任务位置服务225。此多个任务承租者可对应于给定地理区中的所有任务承租者、来自整个世界的各个任

务承租者、或任务承租者的任何其他方便的分组。在图 2 中，示出了服务于被标记为“美国北部”和“美国南部”的区域的任务位置服务 225。任务位置服务 225 的责任可包括管理该给定地理区的任务账户。任务位置服务 225 还可提供应用编程接口 (API) 以便允许用户与该云计算环境交互。这些 API 可包括与虚拟机池、池管理逻辑、以及跨给定地理区域中的任务承租者协调池管理逻辑相关联的处理 API。这些 API 还可包括用于处理用户所提交的任务、以及用于维护、调度以及终止与该用户任务相关联的工作项或作业的 API。这些 API 可进一步包括用于对地理区中的所有工作项、作业、任务和池进行统计数据收集、聚合和报告的 API。此外，这些 API 可包括用于允许基于虚拟机的现货市场短期地将可用虚拟机作为可抢占虚拟机向用户拍卖的 API。这些 API 还可包括用于计量使用并且提供计费支持的 API。

[0064] 这些任务位置服务 225 可被全局位置服务 235 链接在一起。全局位置服务 235 可负责账户创建和账户管理，包括结合任务位置服务承租者 225 来管理任务账户。这包括在存在主要数据中心灾难的情况下，负责灾难恢复并负责工作项和作业的可用性。这可包括由于数据中心因任何原因不可用而在不同的位置运行工作项或作业。这还可包括允许顾客将他们的工作项、作业和此从一个数据中心迁移到另一数据中心。通常，将仅存在一个活跃的全局位置服务 235。该活跃全局位置服务 235 与各个任务位置服务 225 以及各用于管理数据存储的服务组件（未示出）通信。全局位置服务可维护全局账户命名空间 237。

[0065] 作为图 2 中的系统的操作的示例，假想的顾客或用户 217 可经由全局位置服务 235 所提供的接口来创建任务账户。在此示例中，该假想的顾客被称为 Sally。创建任务账户的用户请求可任选地指定需要在其中创建账户的地理区域。在该示例中，Sally 请求于美国北部区域相关联的账户。作为响应，全局位置服务 235 联系与所请求的地理区域（例如，美国北部）相对应的任务位置服务 225 创建该账户。如果区域未被请求，则该任务账户可在用任何方便的方法（诸如基于与请求用户相关联的位置）选择的区域中创建。全局位置服务 235 还联系至少另一区域（诸如美国南部），使得该账户的灾难恢复副本被创建。可任选地，Sally 可请求将美国南部用作灾难恢复的故障转移区域，或者美国南部可被该系统用任何方便的方法自动分派。任务位置服务 225 为处于其地理区域中的所有账户维护所有信息。当成功地在针对美国北部和美国南部的任务位置服务 225 中创建了账户后，全局位置服务 235 为 Sally 的账户注册指向针对美国北部的任务位置服务 225 的虚拟 IP 地址的任务服务端点。例如，域名服务 (DNS) 记录可被创建以将诸如“sally.task.core.windows.net”等主机名映射到美国北部中的任务位置服务 225 的虚拟 IP 地址。这完成了对 Sally 的任务账户的创建。如果在将来的时间发生数据中心灾难，则全局位置服务 235 可将该 DNS 记录更新为指向美国南部。

[0066] 在创建该账户之后，顾客 Sally 可以访问该账户并且发送请求以访问用于针对主机名“sally.task.core.windows.net”与云计算环境进行交互的 API。例如，Sally 可访问 API 以发出创建新工作项或任务的请求。DNS 服务器随后可以解析该主机名，并且该请求将被路由到正确的任务位置服务承租者 225。在此示例中，该请求被路由到针对美国北部的任务位置服务承租者 225，该任务位置服务承租者处理该请求并创建所请求的工作项、作业或任务。

[0067] 图 3 示出了任务位置服务的潜在配置。在图 3 中示出的配置中，任务位置服务可包括一个或多个账户服务器 321。该账户服务器处理针对给定地理区域中的账户的账户管

理，包括创建、删除、或属性更新。账户前端 322 用作账户服务的前端节点。账户前端 322 在该图中所示出的账户虚拟 IP 地址 324 后。账户前端 322 处理来自全局位置服务的账户 API 请求，诸如用于创建账户或删除账户的 API 请求。

[0068] 图 3 中的配置还包括一个或多个池服务器 331。池服务器 331 处理针对给定地理区域中的虚拟机池的池管理和池事务。池服务器 331 处理池创建、删除和属性更新。池服务器 331 还跨多个任务承租者管理高等级虚拟机分配算法。虚拟机分配可将虚拟机与给定用户的存储的连接性纳入考虑。池服务器还可执行与虚拟机的分配有关的其他任务。

[0069] 图 3 的配置还包括一个或多个工作项或作业调度器 (WIJ)336。WIJ 调度器 336 处理对工作项和作业的创建、删除和更新。此外，如果用户请求了在工作项或作业开始或结束时对池的自动创建和 / 或破坏，则 WIJ 调度器 336 可发起对与这些工作项或作业相关联的池的创建和删除。WIJ 调度器 336 还使用用于缩放的类属划分机制。在一实施例中，在每个任务位置服务中存在多个 WIJ 调度器 618，并且每个 WIJ 调度器处理一定范围的工作项。

[0070] 池服务器 331 和 WIJ 调度器 336 经由任务位置服务前端 338 接收来自用户的请求。任务位置服务前端 338 还负责调用相应的组件以处理来自用户的请求。任务位置服务前端 338 在如该图中所示出的账户虚拟 IP 地址 334 后。

[0071] 图 3 中的配置还包括任务位置服务主机 342。在一实施例中，任务位置服务主机 342 具有两个主要的职责。首先，任务位置服务主机 325 用作用于为任务位置服务 225 中的相应服务器实现划分逻辑的主机系统。此外，任务位置服务主机 342 可负责在每一现货时段的开始处为该任务位置服务的整个地理区域计算可抢占虚拟机的新市场价。可以从各池服务器和任务承租者处收集当前投标和资源可用性信息，并据此计算该新市场价。或者，任务位置服务主机可向现货价格市场服务发送该投标和资源可用性信息。它还作出对池服务器的关于可抢占虚拟机跨地理区域中的所有任务承租者的高级分配指南。

[0072] 为了跟踪计算环境的活动和行为，任务位置服务主机 342 可与一个或多个统计数据聚集服务器 355 通信。统计数据聚集服务器负责收集和聚集任务、作业、工作项和池的详细统计数据。该系统中的其他组件发出任务和虚拟机的精细粒度统计数据。统计数据聚集服务器将来自任务层或虚拟机层统计数据的这些精细粒度统计数据聚集为工作项、帐户层、和 / 或池层统计数据。这些统计数据可经由 API 来展示以供使用。此外，统计数据聚集服务器可负责生成针对每个帐户的每小时计量记录以供计费使用。

[0073] 图 4 示意性地示出了可被包括为任务位置服务和 / 或任务位置服务主机的部分的附加模块。在图 4 中，现货定价模块 460 是可作为任务位置服务主机的部分的模块。现货定价模块是负责在每一现货时段的开始处确定市场价的全局模块。作为全局模块，现货定价模块 460 通常向多个池服务器 431 提供信息。现货定价模块 460 与作为任务位置服务的部分的池服务器维持“心跳”以同步经现货定价的可抢占虚拟机的当前市场价。

[0074] 度量收集模块 472 是可作为池服务器的部分的模块。度量收集模块 472 负责收集用于对池服务器所拥有的相应池进行自动缩放的度量。这些度量包括 CPU 的每池状态、网络、队列状态以及所有其他度量。该模块的输出馈送到自动缩放模块 474 中。自动缩放模块 474 也可以是池服务器的部分。自动缩放模块负责基于与每个池相关联的自动缩放公式来作出自动缩放判定。它取得这些度量以及用户所提供的公式 / 规则，并计算每个池的自动缩放动作。自动缩放动作可包括将池的专用虚拟机增加或减少特定量；将池的备用虚拟

机增加或减少特定量；以及将池的经现货定价的或可抢占虚拟机的目标数量增加或减少特定量，以及更新投标价格。自动缩放模块 474 的输出被馈送到池管理模块 480 中，该池管理模块 480 执行这些指令并以其他方式实现用于改变给定池的大小的机制。这些指令可以按与对更新池大小的用户请求相同的方式来处理。对于给定现货价格，池管理模块 480 根据当前市场价和未完结的投标来控制对池中的可抢占虚拟机的抢占和分配。

[0075] 图 5 示出任务承租者的一实施例的示例高级架构，该架构包括任务承租者的各组件及相应的责任的示例。如上所述，任务承租者可帮助管理虚拟机池。在图 5 中示出的实施例中，任务承租者包括一个或多个任务承租者前端 522。该任务承租者前端 522 位于任务承租者虚拟 IP 地址 524 之后，该任务承租者虚拟 IP 地址 524 内部用于任务承租者及其相应的任务地址服务之间的通信，该通信包括在任务位置服务和任务承租者之间传递请求。

[0076] 在图 5 所示的实施例中，任务承租者还包括任务调度器 536。任务调度器 536 可负责做出任务承租者内的局部任务调度判决。任务调度器 536 决定在它所控制的每个虚拟机上运行什么任务。例如，用户所提交的工作项或作业可具有一组队列，该组队列包含要被调度的任务的列表。任务调度器 536 从该组队列中取得任务，选择与该作业相关联的池中的一个或多个可用虚拟机，并且联系（诸）虚拟机来调度这些任务。任务调度器 536 还可基于与作业相关联的优先级值来做出调度判决。此外，任务调度器 536 跟踪任务承租者内的虚拟机。任务调度器 536 与池服务器一起工作以将虚拟机分配给池 / 从池中解除分配虚拟机。此外，任务调度器 536 与所有虚拟机保持心跳、经由心跳关于池成员与虚拟机同步，并且控制虚拟机的重启 / 重新映像。任务调度器 536 的又一功能可以是跟踪任务承租者的大小。基于任务承租者内的虚拟机的当前利用率，任务调度器可以增长或收缩该任务承租者，以便该任务承租者具有足够数量的虚拟机来运行与该任务承租者相关联的任务。类似地，如果该任务承租者中有太多虚拟机保持空闲，可以释放这些虚拟机以供该数据中心中的其他托管服务使用。

[0077] 图 5 还示出了与任务承租者相关联的多个虚拟机。在图 5 中示出的实施例中，每个虚拟机包括任务虚拟机代理 550(TVM)。在一实施例中，任务虚拟机代理 550 负责启动虚拟机上的任务，以及设置任务的目录结构和许可。在虚拟机上还配置了仅允许同一池内的虚拟机之间的流量（如果该池需要内部通信的话）的操作系统防火墙。如早先讨论的，任务调度器 536 经由任务虚拟机代理 550 与该虚拟机保持心跳。这允许任务调度器 536 监视虚拟机的健康以及为任务虚拟机代理同步池成员信息。

[0078] 现货定价流示例

[0079] 以下提供了如何在系统内以全局为基础实现现货定价的示例。在该示例中，以下三个组件或模块对全局现货定价有贡献：现货定价模块，诸如任务位置服务主机内的模块或任务系统外部的现货定价服务；池管理模块，诸如作为任务位置服务中的每一池服务器的部分的池管理模块；以及任务调度器，诸如潜在地作为每一任务承租者的部分的任务调度器。不同的组件具有各种责任。图 6 示意性地示出适于执行对可抢占虚拟机资源的全局现货定价的系统的示例。在图 6 所示的示例中，更新云计算环境中的全局现货价格包括至少三个过程。

[0080] 在图 6 中，现货定价模块 660 可负责在每一现货时段（诸如，分派时间段）的开始处计算全局市场价。现货定价模块 660 可提供跨所有池服务器 631 的现货可抢占虚拟机分

配的高级细目分类，但该现货定价模块并不涉及对每一个体投标的详细分配决策。在确定了市场价之后，现货定价模块 660 可负责更新价格历史表 670 和池服务器 631。在图 6 所示的示例中，价格历史表 670 对应于全局价格历史表。价格历史表 670 可在每一现货时段内跟踪市场价。一旦该价格被确定，现货定价模块 660 就可更新该表。现货定价模块 660 还可经由任务位置服务主机和池服务器之间有规律的心跳来向池服务器 631 发送市场价更新。现货定价模块 660 还可包括在每一任务承租者的不同池服务器之间的现货可抢占虚拟机分配的初始高级细目分类。

[0081] 优选地，现货定价模块 660 可首先更新价格历史表 670。现货定价模块 660 可随后在第二步骤经由心跳消息来更新池服务器 631。池服务器随后在第三过程更新各个任务承租者。优选地，价格更新消息可以用针对现货时段的相应时戳进行标记。由于现货定价模块 660 是全局模块，该现货定价模块可保证该时戳总是增加。价格历史表 670 可总是保持当前现货价格的真相。对当前现货价格没有把握的池服务器 631 可以经由价格历史表 670 来访问该当前现货价格。

[0082] 价格历史表 670 保持当前价格的真相。当设置新现货价格时，现货定价模块 660 将告知任何人该新价格，直到价格历史表 670 被更新。任务位置服务主机具有与每一池服务器的有规律的心跳。在每一心跳消息中可包括各种类型的信息。该心跳消息可包括当前现货时段的时戳。该时戳一直在增加，并且可用作用于确定哪个现货时段更新近的序列号。心跳消息还可包括针对当前现货时段的市场价。此外，心跳消息可包括直到下一现货时段将开始（其对应于该价格何时被再次更新）为止的持续时间。池服务器可使用该信息来判定如果它们没有及时收到任务位置服务的回音它们应在何时预期下一价格变化。

[0083] 如果现货定价模块（或任务位置服务主机）因任何原因停滞，则该系统的其余部分仍可正确工作，从而基本上通过另一时段来延续当前现货定价。价格表将不被更新，并且池服务器仍将使用目前的市场价，从而有效地延续了当前现货时段。现货价格时段优选地具有固定的 N 分钟范围。例如，如果 30 分钟的时段被使用，则这些时段可以是 1:00-1:30、1:30-2:00、2:00-2:30 并以此类推。当任务位置服务主机恢复时，它可针对当前时段开始新现货时段，如果该当前时段处于时段开始时间的 X 分钟内的话。如果它是过去的 X 分钟，则在直到下一间隔之前它将只是等待，以固定价格。然而，在该情形下，任务位置服务主机仍可向现货价格历史表添加未改变现货价格的新现货时段。例如，如果新现货价格在固定时窗（诸如，当前现货时段的预期开始的前 5 分钟）内可用，则现货价格可被更新。如果现货定价模块和 / 或任务位置服务主机晚了并错过了那个时窗，则该价格可在直到下一现货周期之前保持不变。

[0084] 每一池服务器 631 可包括池管理模块 680。池管理模块 632。在这个示例中，池管理模块 680 处理给定池内的预留和转换（在备用虚拟机和专用虚拟机之间）请求以及移除可抢占虚拟机的任何显式请求。此外，为了处理现货定价，池管理模块还可负责履行高于当前市场价的未完结投标，并基于不再有资格的投标而撤走可抢占虚拟机。池管理模块 680 可负责跟踪具有高于（或等于）当前市场价格的未完结或未履行投标的那组池。“未完结”意味着尚未收到其已请求的所有可抢占虚拟机的池。池管理模块可随后分配可抢占虚拟机以按降序次序来履行未完结的投标（即，先履行较高的投标）。此外，池管理模块可抢占来自池的现在具有低于当前市场价的投标的所有可抢占虚拟机。注意，池服务器 631 负责经

由现货定价为给定任务承租者615设置池中对分派可用的可抢占虚拟机的目标数目。池服务器631不跟踪任务承租者615中针对给定池分配的可抢占虚拟机的确切数目。为达到池服务器631所设的目标而添加/移除可抢占虚拟机均取决于任务承租者615。

[0085] 任务调度器636是任务承租者615中的模块。在该示例中，任务调度器636未活跃地跟踪现货价格。任务调度器636可维护跟踪给定承租者中针对每一池的目标可抢占虚拟机计数的“TenantPoolTable(承租者池表)”或其他类似的数据结构。当任务调度器636收到基于高于(或低于)现货价格的投标的针对可抢占虚拟机的池事务时，该任务调度器将更新这个表以记录给定池的目标可抢占虚拟机计数，并且从池服务器631的角度而言该事务被完成了。任务调度器负责使针对各池的可抢占虚拟机计数达到目标计数。在对专用虚拟机进行转换的情况下，如果不存在与任务承租者615相关联的足够空闲的虚拟机，则任务调度器636可抢占具有较低投标的一些可抢占虚拟机。这可在没有向池服务器631通知该抢占的情况下完成。

[0086] 分配和抢占可在新现货时段的开始处和在现货时段期间发生。在现货时段的开始处，任务位置服务主机向每一池服务器发送在跨每一任务承租者的池服务器之间的可抢占虚拟机分配的高级细目分类。池服务器可使用该信息来指导分配和抢占决策。池服务器跟踪所有这些未完结的投标及其它们的提交时间。对于在现货时段的开始之前或者可能在不同于现货时段的开始的截止时间之前提交的所有投标，池服务器可确保较高的投标在较低的投标之前被满足。结果，在先前时段基于较低投标分派的一些可抢占虚拟机可被抢占。池服务器还可使用由任务位置服务主机提供的全局信息来协调并最小化不必要的抢占。

[0087] 在任务位置服务主机计算市场价时，它还将计算跨不同池和不同任务承租者的可抢占虚拟机分配的高级细目分类。该信息被传递至所有这些池服务器以帮助其分配决策。该信息可包括对任务承租者内的每一池划分范围的每一投标价格和每一约束的详细可抢占虚拟机分配细目分类。例如，从任务位置服务主机的角度来说，具有相同投标价格和相同约束(例如，它们由于内部通信或存储亲和力约束而需要使用哪个(哪些)承租者)的所有投标将被分组在一起。任务位置服务主机为每一池划分范围提供详细分配。

[0088] 池服务器可基于由任务位置服务主机提供的分配信息来计算刷新分配(就好像所有潜在的可抢占虚拟机都空闲一样)。池服务器还可为每一池确定在每一任务承租者中的新的目标可抢占虚拟机计数。随后，可将该新分配与当前分配进行比较以计算需要被更新的那组池。池服务器随后为要求更新的池联系相关的任务承租者，以便为可抢占机器的数目设置新目标值。

[0089] 在池服务器分配可抢占虚拟机时，该池服务器可以在尝试向较低投标分配任何东西之前开始针对较高投标的分配事务。池服务器还可跟踪投标何时被提交。在相同价格的两个投标之间，较早的投标将占有优先权。注意，池服务器不需要在开始下一些事务之前等待之前的交易结束。相反，池服务器仅需要确保它在前进到下一组投标之前已开始了与相应任务承租者的分配事务。优选地，这些事务并行执行。

[0090] 在分派时间段期间，抢占可在需要将虚拟机作为专用机器来分派或者云计算环境因其他原因(诸如，灾难恢复)需要该机器时发生。专用虚拟机。如果在该系统中存在空闲虚拟机可用，则空闲机器可用于作为专用虚拟机来进行分派。如果附加的空闲虚拟机不可用，则任务承租者可抢占对应于较低投标的可抢占虚拟机。用于为供抢占的机器排定优

先级的另一选项是对抢占已运行作业达较短时间段的机器具有偏好。分配在更多可抢占虚拟机变得可用并且存在尚未被满足的未完结投标时发生。在那种情况下，可用的可抢占虚拟机可被分配，该分配伴随较高投标开始。

[0091] 优选地，一小组空闲虚拟机可保持被重新映像并准备好被使用，使得在需要将虚拟机用于专用使用时，可立即从该组中取得该专用虚拟机。任务承租者将在后台维护这些空闲虚拟机。当任务承租者中的空闲虚拟机的数目低于阈值量（诸如，该任务承租者中有1％的专用虚拟机）时，该任务承租者可开始抢占具有较低投标的虚拟机，直到空闲虚拟机计数达到第二阈值，该第二阈值可以与第一阈值相同或不同。任务承租者可在不涉及池服务器的情况下抢占这些可抢占虚拟机，从而允许该抢占快速发生。

[0092] 另一方面，如果任务调度器已经为它的由池服务器分派的所有池履行了它的所有目标可抢占虚拟机，并且仍存在大于第二阈值的额外空闲虚拟机时，该任务调度器可以经由池服务器有规律的心跳向这些池服务器报告这样的额外空闲虚拟机的计数。如果该计数大于第三阈值，则池服务器将开始向未完结的投标分配这些额外的虚拟机。

[0093] 以下提供了用于基于现货定价来分派可抢占虚拟机的过程流的高级示例。在现货时段的开始处，任务位置服务主机（诸如，任务位置服务主机内的全局现货定价模块）基于这些投标和资源可用性来为该现货时段计算新市场价。在决定了该价格后，任务位置服务主机如之前所描述的用新价格以及即将到来的现货时段的时戳来更新价格历史表。任务位置服务主机随后经由其有规律的心跳消息来向每一池服务器发送现货价格。此外，任务位置服务主机还为每一池服务器发送可抢占虚拟机分配的初始细目分类。这可帮助池服务器作出分配决策。当池服务器接收到来自任务位置服务主机的消息时，池服务器开始为它的所有池向未完结的投标分配可抢占虚拟机，并抢占低于新市场价的所有可抢占机器。尤其地，池服务器向任务承租者发送在给定池上设置新的可抢占虚拟机目标计数的命令。这以与设置专用虚拟机计数的池事务类似的方式完成，除了那个事务是在任务承租者一记录了该目标可抢占虚拟机计数就完成的以外。随后，任务调度器将尝试通过分配或移除池中的虚拟机来使虚拟计数达到该目标。在任务承租者侧，可抢占虚拟机分配以与针对专用虚拟机的相同的方式完成，除了可抢占虚拟机是从承租者中的全局空闲虚拟机集合中取得的，并且该任务调度器总是先向较高投标分配可抢占虚拟机以外。此外，在现货时段期间，由于资源短缺（例如，备用虚拟机向专用虚拟机的转换），任务调度器可能需要抢占可抢占虚拟机。在那种场景中，对应于具有较低投标的池的虚拟机先被抢占。

[0094] 在分派时间段期间，池服务器可发现一些可抢占虚拟机已变得对基于现货定价进行分派可用。例如，一些专用虚拟机可被转换成备用虚拟机，或者一些可抢占虚拟机可被用户释放。池服务器可通过为给定池设置新的（较高的）可抢占虚拟机目标计数来向未完结投标分配可用虚拟机，其中较高的投标取得优先权。在一些实施例中，在现货时段的结束之际（诸如在现货时段的最后 5 分钟内）不分配可抢占虚拟机，因为这些可抢占虚拟机在下一现货时段开始时可能很快就被抢占掉。

[0095] 对云计算环境中的虚拟机进行分派的示例

[0096] 提供了以下解说云计算环境中的专用虚拟机、备用虚拟机和可抢占虚拟机的操作和交互的假想示例。在这些示例中，将讨论较小数目的虚拟机，以便简化描述和附图。然而，本领域的技术人员将认识到本文中所描述的概念可扩大到任何所需数目的虚拟机。

[0097] 在以下假想示例中，将描述对各种专用虚拟机、备用虚拟机和可抢占虚拟机的分派。在相应的附图中，针对用户 Abel 的机器将被标记成 A，针对用户 Charlie 的机器将被标记成 C，针对用户 David 的机器将被标记成 D，并且针对用户 Frank 的机器将被标记成 F。一些机器将被标记成 L 以标识附加的大用户。除了指定虚拟机被分派给的用户以外，各虚拟机还可具有对专用机器的指定 (D) 或对可抢占机器的指定 (P)。该示例中各个用户所执行的作业可以是用于执行任何类型的计算（诸如，执行针对业务的数据挖掘和管理、执行科学计算、或处理零售顾客流量）的作业。

[0098] 图 7 示出了各虚拟机在两个任务承租者 710 和 711 中的初始状态的示例。任务承租者是代表性的，使得任何方便数目的任务承租者可以从现货定价模块 760 接收信息。类似地，在每一任务承租者内示出的虚拟机的数目是代表性的，使得更大或更新数目的虚拟机可被包括在任务承租者中。在任务承租者 710 或 711 内，具有相同开始指定字母的每一虚拟机对应于相同池内的机器。例如，具有“A(?)”格式的所有虚拟机都在与用户 Abel 的账户相关联的池中。

[0099] 在图 7 中，示出了任务承租者 710 和 711 在经由现货定价来将任何虚拟机作为可抢占机器进行分派之前的初始状态。在图 7 中，用户 Abel、Charles、David 和 Frank 各自使两个专用虚拟机被分派并运行作业。这些机器分别被示为机器 723、733、743 和 753。虚拟机 768 和 769 对应于未被分派给任何池的机器。此外，备用计数 793 被包括在任务承租者 710 的下面，并且备用计数 794 被包括在任务承租者 711 的下面。备用计数 793 和 794 表示当前与相应的任务承租者 710 和 711 相关联的备用虚拟机预留。在图 7 中，该备用预留对应于与任务承租者 710 和 711 中的每一者中的大用户相关联的 2 个预留。在这个示例中，备用预留基于系统的选择而与各任务承租者相关联。如果这些备用预留是亲和力请求的部分，则大用户可具有有合适亲和力的指定群集，并且这些备用预留可能已据此被关联。

[0100] 现货定价确定随后用于基于来自用户请求来向这些用户分派可抢占机器。现货定价模块收集从所有可用池中提供的投标，并确定每小时（或其他定价时段）0.6 分的现货价格。还可在定价时段的各片段内进行计费。现货价格基于对可抢占机器的各种投标，包括来自用户 Abel、Charles、David 和 Frank 的投标。用户 Abel 以 1.5 分的投标价格请求三个可抢占机器。用户 Charles 以 1.3 分的投标价格请求一个可抢占虚拟机，并以 0.6 分的投标价格请求第二个可抢占虚拟机。用户 David 以 0.5 分的投标价格请求三个可抢占机器。用户 Frank 以 1.0 分的投标价格请求一个机器，并以 0.8 分的投标价格请求另外三个机器。

[0101] 基于这些投标，向这些用户分派可抢占机器。图 8 示出了这些机器向用户的分派。在任务承租者 711 中，基于具有对可抢占机器的最高投标而将三个可用机器 826 分派给 Abel。这些机器对应于图 7 中的空闲机器 769 中的三个空闲机器。接着，任务承租者 710 和 711 尝试履行 Charles 的以 1.3 分的投标对一个可抢占虚拟机的请求。任务承租者 711 中的一个可用虚拟机 836 被作为可抢占虚拟机分派给 Charles。

[0102] 接着，基于投标价格按顺序解决 Frank 对虚拟机的投标。这些虚拟机是从任务承租者 710 处分派的，因为任务承租者 710 是具有剩余可用性的任务承租者。每时间段以 1.0 分对一个可抢占机器的投标由虚拟机 856 来履行。以 0.8 分对附加的三个机器的投标由虚拟机 857 来履行。在基于 Frank 的投标分派了虚拟机之后，下一分派时间段仍未履行来自 Charles 的处于或高于该价格的一个虚拟机请求。该请求通过将 Charles 的池扩展到任务

承租者 710 中并向 Charles 分派可抢占虚拟机 837 来履行。由于与 David 的请求相关联的投标低于针对可抢占虚拟机的分派的现货价格，因此 David 的以 0.5 分对三个机器的请求没有被履行。基于以上分派，具有最低的相应投标价格的可抢占机器位于任务承租者 710 中。如果大用户将备用预留转换成专用虚拟机，则一个选项将是为该大用户将来自任务承租者 710 和 711 中的每一者的两个可抢占虚拟机转换成专用虚拟机。这将导致任务承租者 711 中与较高投标相关联的可抢占作业被取代 (displaced)。为了以较低投标为代价而优先满足较高投标，这可导致第二取代 (displacement)，以使得在任务承租者 710 中以具有较低投标的作业（诸如针对 Charles 或 Frank 的作业）为代价而重新开始较高投标作业（诸如针对 Able 的作业）。另一选项是跨任务承租者 710 和 711 重新分配备用预留，使得备用预留与使得机器被基于最低投标分派的任务承租者相关联。这在图 8 中示出，其中备用计数 793 被调整为 0，而备用计数 794 现在为 4。注意，没有机器基于备用计数的改变而被重新分派。

[0103] 在图 8 中，使用了其中可抢占投标都不基于亲和力请求的情况下的示例。图 9 示出了其中来自 Charles 的可抢占投标包括对已经分派给 Charles 的其他专用虚拟机或可抢占虚拟机的亲和力请求的替换示例。基于该亲和力请求，来自 Charles 的可抢占作业请求对其中 Charles 具有两个专用虚拟机的任务承租者 711 具有亲和力。在图 9 中，相同数目的虚拟机被分派给用户 Abel、Charles 和 Frank 中的每一个。然而，在确定对虚拟机的分派时，任务承租者考虑来自 Charles 的亲和力请求。如果这些机器是在图 8 中描述的方法下分派的，则在任务承租者 711 中仅一个虚拟机将对 Charles 可用。由于该亲和力请求，Charles 不会使用来自任务承租者 710 的虚拟机，从而即使相应的投标价格处于或高于现货价格也留下来自 Charles 的请求不被履行。为了避免这种情况，在任务承租者 710 中的一个可抢占虚拟机 926 被分派给 Abel。可抢占虚拟机 936 随后对向 Charles 的分派可用。即使 Abel 的投标价格高于 Charles，来自 Charles 的亲和力请求也在履行了 Abel 的请求时才被考虑。这允许来自可抢占机器的利用和利润被增加。由于任务承租者 711 现在包括以最低投标价格分派的可抢占机器，任务承租者 711 的备用计数 793 为 1，而任务承租者 710 的备用计数为 3。

[0104] 继续图 9 中示出的示例，在稍后的时间，发生将大用户的 4 个备用预留转换成专用虚拟机的触发事件。由于大用户正在使用的虚拟机的活动或使用或其结合，该触发事件可基于时间、基于负载。在该示例中，该触发事件是基于活动或负载的，并在分派时间段的中间期间发生。在该相同分派时间段期间，Abel 还将所请求的可抢占机器的数目从三增加到四。该来自 Abel 的增加请求包括相同的投标价格。

[0105] 图 10 示出了以上改变的初始成果。为大用户对备用预留的转换导致为该大用户将虚拟机 1094 转换成专用机器。这些备用预留是通过抢占最低优先级可抢占作业而被转换成专用虚拟机的。在图 10 中示出的示例中，这对应于抢占具有最低关联投标的作业。在图 10 中示出的示例中，备用计数 793 和 794 反映出这些承租者包含基于最低可抢占投标来分派的虚拟机，但这并不是必须的。如上所述，备用预留可以出于各种原因而与所需任务承租者相关联，并且可抢占作业可以在对专用机器的转换之后在各任务承租者之间移动。在图 10 中，对应于最低可抢占投标的虚拟机是基于 0.6 分的投标向 Charles 分派的虚拟机。将该虚拟机转换成任务承租者 711 中针对大用户的专用虚拟机 1093。接着的三个最低投

标对应于任务承租者 710 中分派给 Frank 的可抢占虚拟机。这些虚拟机被转换成分派给大用户的专用虚拟机 1094。这留下一个可抢占虚拟机 856 被分派给 Frank。注意，虽然 Abel 具有更高的投标，但现货定价机制仅用于在时间段的开始处重新分派可抢占虚拟机。由于 Abel 的请求是在时间段的中间期间作出的，因此 Abel 的请求不会取代分派给 Frank 的可抢占虚拟机，即使 Abel 的请求包括更高的投标。此外，由于对针对大用户的备用预留的转换，任务承租者 710 和 711 两者的备用计数被减少成 0。

[0106] 图 11 示出了在下一分派时间段的开始处发生的附加改变。由于由大用户请求的额外资源，更少虚拟机可用于作为可抢占机器进行分派。这导致全局现货价格增加到每时间段 11 分。如图 11 所示，Abel 对附加机器的在先请求现在由虚拟机 1126 来履行。在任务承租者 710 中，全局现货价格的增加导致 Frank 的投标跌到现货价格以下，使得在该分派时间段期间未向 Frank 分派可抢占虚拟机。

[0107] 附加实施例

[0108] 在简要描述了本发明的各实施例的概览后，现在描述适于实现虚拟机的示例性操作环境。概括地参考附图，并首先具体参考图 12，示出了用于实现本发明的各实施例的示例性操作环境，并将其概括地指定为计算设备 1200。计算设备 1200 只是合适的计算环境的一个示例，并且不旨在对本发明的使用范围或功能提出任何限制。也不应该将计算设备 1200 解释为对所示出的任一组件或其组合有任何依赖性或要求。

[0109] 本发明的各实施例可以在由计算机或诸如个人数据助理或其它手持式设备之类的其它机器执行的计算机代码或机器可使用指令（包括诸如程序模块之类的计算机可执行指令）的一般上下文中描述。一般而言，包括例程、程序、对象、组件、数据结构等的程序模块指的是执行特定任务或实现特定抽象数据类型的代码。本发明可以在各种系统配置中实施，这些系统配置包括手持式设备、消费电子产品、通用计算机、更专用计算设备等等。本发明也可以在其中任务由通过通信网络链接的远程处理设备执行的分布式计算环境中实践。

[0110] 继续参考图 12，计算设备 1200 包括直接或间接耦合以下设备的总线 1210 :存储器 1212、一个或多个处理器 1214、一个或多个可任选的呈现组件 1216、输入 / 输出 (I/O) 端口 1218、可任选的 I/O 组件 1220、和说明性电源 1222。总线 1210 表示可以是一条或多条总线（诸如地址总线、数据总线、或其组合）。虽然为了清楚起见利用线条示出了图 12 的各框，但是实际上，各组件的轮廓并不是那样清楚，并且比喻性地来说，线条更精确地将是灰色的和模糊的。例如，可以将诸如显示设备等呈现组件认为是 I/O 组件。另外，许多处理器具有存储器。发明人关于此点认识到这是本领域的特性，并重申，图 12 的图示只是可以结合本发明的一个或多个实施例来使用的示例性计算设备的例示。诸如“工作站”、“服务器”、“膝上型计算机”、“手持式设备”等分类之间没有区别，它们全部都被认为是在图 12 的范围之内的并且被称为“计算设备”。

[0111] 计算设备 1200 通常包括各种计算机可读介质。计算机可读介质可以是可由计算设备 1200 访问的任何可用介质，而且包含易失性和非易失性介质、可移动和不可移动介质。作为示例而非限制，计算机可读介质可包括计算机存储介质和通信介质。计算机存储介质包括以存储诸如计算机可读的指令、数据结构、程序模块或其他数据之类的信息的任何方法或技术实现的易失性和非易失性、可移动和不可移动介质。计算机存储介质包括，但

不限于，随机存取存储器 (RAM)、只读存储器 (ROM)、电可擦除可编程只读存储器 (EEPROM)、闪存或其它存储器技术、CD-ROM、数字多功能盘 (DVD) 或其它全息存储器、磁带盒、磁带、磁盘存储或其它磁性存储设备或可用于编码所需信息且可以由计算设备 1200 访问的任何其它介质。在一个实施例中，计算机存储介质可以从有形计算机存储介质中选择。在另一实施例中，计算机存储介质可以从非临时性计算机存储介质中选择。

[0112] 通信介质通常以诸如载波或其他传输机制等已调制数据信号来体现计算机可读指令、数据结构、程序模块或其他数据，且包含任何信息传递介质。术语“已调制数据信号”指其一个或多个特征以这样的方式设置或改变以便在信号中对信息进行编码的信号。作为示例而非限制，通信介质包括诸如有线网络或直接线连接之类的有线介质，以及诸如声学、RF、红外及其他无线介质之类的无线介质。上述的任意组合也应包括在计算机可读介质的范围之内。

[0113] 存储器 1212 可包括易失性和 / 或非易失性存储器形式的计算机存储介质。存储器可以是可移动的，不可移动的，或两者的组合。示例性硬件设备包括固态存储器、硬盘驱动器、光盘驱动器等。计算设备 1200 包括从诸如存储器 1212 或 I/O 组件 1220 等各种实体读取数据的一个或多个处理器。（诸）呈现组件 1216 向用户或其他设备呈现数据指示。示例性呈现组件包括显示设备、扬声器、打印组件、振动组件等等。

[0114] I/O 端口 1218 可允许计算设备 1200 逻辑上耦合至包括 I/O 组件 1220 的其它设备，其中某些可以是内置的。说明性组件可包括话筒、操纵杆、游戏手柄、圆盘式卫星天线、扫描仪、打印机、无线设备等等。

[0115] 参考具体实施例描述了本发明的实施例，具体实施例在所有方面都旨在是说明性的而非限制性的。不偏离本发明范围的情况下，各替换实施例对于本发明所属领域的技术人员将变得显而易见。

[0116] 图 13 示出了根据本发明的方法的示例。在图 13 中，接收 1310 针对可抢占虚拟机的分派的第一价格。该接收价格可被使用，例如一个或多个虚拟机群集基于该接收价格以及与各个虚拟机池相关联的投标来分派可抢占虚拟机。随后将来自一个或多个虚拟机群集的多个可抢占虚拟机分派 1320 给虚拟机池。使用所分派的虚拟机来执行 1330 一个或多个任务。随后接收 1340 针对可抢占虚拟机的分派的第二价格。通常，这可对应于接收供在后续分派时间段使用的新价格。将来自一个或多个虚拟机群集的至少一个虚拟机和来自附加虚拟机群集的至少一个虚拟机分派 1350 给该虚拟机池。随后使用所分派的来自附加虚拟机群集的（诸）虚拟机来执行 1360 一个或多个任务。

[0117] 图 14 示出了根据本发明的方法的另一示例。在图 14 中，接收 1410 针对可抢占虚拟机的分派的价格。基于与第一虚拟机池相关联的第一投标将来自第一虚拟机群集的虚拟机分派 1420 给第一虚拟机池。该第一投标对应于对第一虚拟机群集具有亲和力的对虚拟机的请求。该请求中的至少一个虚拟机未被履行。基于与第二虚拟机池相关联的第二投标将来自第二虚拟机群集的虚拟机分派 1430 给第二虚拟机池。分派给第二虚拟机池的至少虚拟机是基于大于收到价格但小于对应于第一虚拟机池的投标的投标价值来分派的。随后使用所分派的可抢占虚拟机（诸如，分派给第二虚拟机池的可抢占虚拟机）来执行 1440 一个或多个任务。

[0118] 图 15 示出了根据本发明的方法的又一示例。在图 15 中，接收 1510 针对可抢占虚

拟机的分派的价格。基于相关联的第一投标将第一多个可抢占虚拟机分派 1520 给第一虚拟机池。基于相关联的第二投标将第二多个可抢占虚拟机分派 1530 给第二虚拟机池。使用所分派的虚拟机来执行 1540 一个或多个任务。随后接收到 1550 来自第一虚拟机池的增加可抢占虚拟机的数目的请求。对应于该请求的投标大于与第二虚拟机池相关联的投标。维护 1560 对第二多个虚拟机的分派直到分派时间段的结束。随后从第二虚拟机池中移除 1560 对来自第二多个虚拟机的至少一个虚拟机的分派。在后续分派时间段中将所移除的至少一个虚拟机分派 1570 给第一虚拟机池。

[0119] 在一实施例中，提供了一种用于提供云计算环境中的资源的方法。该方法包括：接收针对可抢占虚拟机的分派的第一价格；基于接收到的第一价格和与虚拟机池相关联的第一投标将来自一个或多个虚拟机群集的多个可抢占虚拟机分派给该虚拟机池；在所分派的多个可抢占虚拟机上执行一个或多个任务；接收针对可抢占虚拟机的分派的第二价格；基于接收到的第二价格以及与该虚拟机池相关联的第二投标将来自一个或多个虚拟机群集的至少一个可抢占虚拟机以及来自附加虚拟机群集的至少一个可抢占虚拟机分派给该虚拟机池；以及在所分派的来自附加虚拟机群集的至少一个可抢占虚拟机上执行一个或多个任务。

[0120] 在另一实施例中，提供了一种用于提供云计算环境中的资源的方法。该方法包括：接收针对可抢占虚拟机的分派的价格；基于接收到的价格和与第一虚拟机池相关联的第一投标将来自第一虚拟机群集的一个或多个可抢占虚拟机分派给该第一虚拟机池，该第一投标对应于包括对第一虚拟机群集的亲和力的对多个可抢占虚拟机的请求，其中在该对多个可抢占虚拟机的请求中的至少一个虚拟机在分派了第一虚拟机群集中的虚拟机后未被履行；基于接收到的价格和与第二虚拟机池相关联的第二投标将来自第二虚拟机群集的一个或多个可抢占虚拟机分派给该第二虚拟机池，其中来自该第二虚拟机群集的至少一个所分派虚拟机是基于大于收到的价格并小于与第一虚拟机池相关联的第一投标的投标而被分派给第二虚拟机池的；以及，在第二虚拟机池中的所分派的来自第二虚拟机群集的可抢占虚拟机上执行一个或多个任务。

[0121] 在又一实施例中，提供了一种用于提供云计算环境中的资源的方法。该方法包括：接收针对可抢占虚拟机的分派的价格；基于接收到的价格和与第一虚拟机池相关联的第一投标将来自一个或多个虚拟机群集的第一多个可抢占虚拟机分派给第一虚拟机池；基于接收到的价格和与第二虚拟机池相关联的第二投标将来自该一个或多个虚拟机群集的第二多个可抢占虚拟机分派给第二虚拟机池；在第一多个可抢占虚拟机上以及在第二多个可抢占虚拟机上执行一个或多个任务；接收来自第一虚拟机池的增加可抢占虚拟机的数目的请求，该增加请求对应于与第一虚拟机池相关联的第三投标，该第三投标大于与第二虚拟机池相关联的第二投标；维护对第二多个虚拟机的分派直到分派时间段的结束；从第二虚拟机池中移除对来自第二多个虚拟机的至少一个虚拟机的分派；以及，在后续分派时间段将所移除的至少一个虚拟机分派给第一虚拟机池。

[0122] 从前面的描述可以看出，本发明很好地适用于实现上文所阐述的所有目的和目标，并且具有对于该结构是显而易见且固有的其他优点。

[0123] 可理解的是，某些特征和子组合是有用的，并且可以加以利用而无需参考其它特征和子组合。这由权利要求所构想的，并在权利要求的范围内。

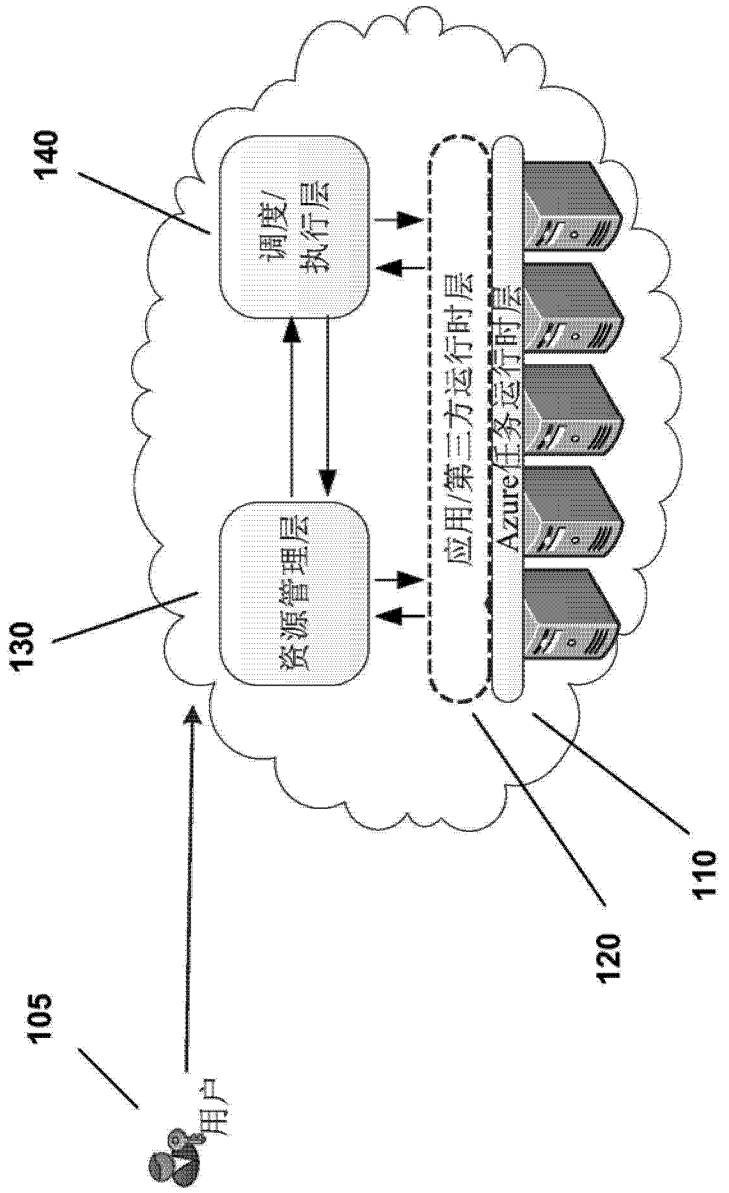


图 1

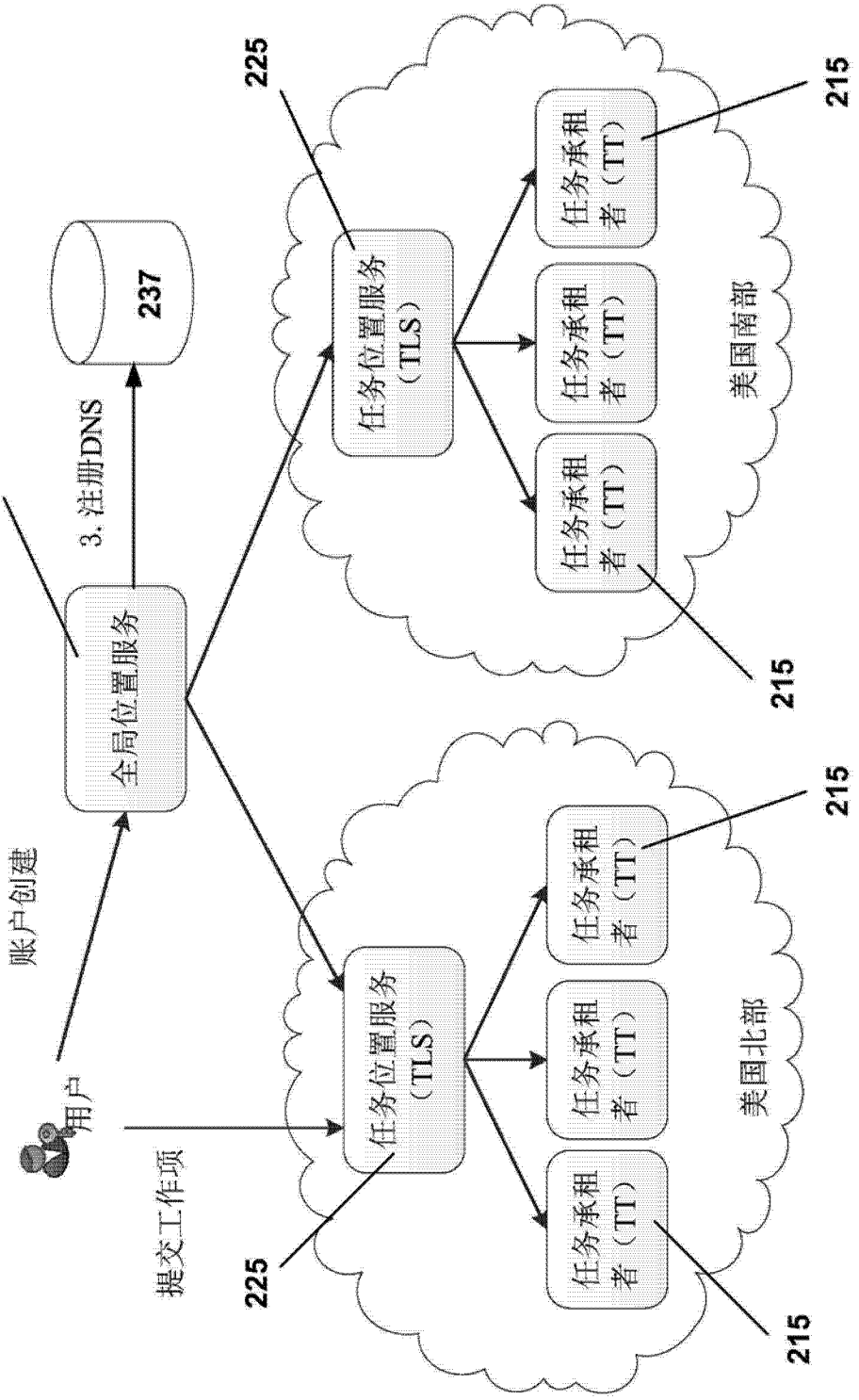

用户
账户创建
提交工作项
全局位置服务
235
3. 注册DNS
237
任务位置服务（TLS）
225
任务承租者（TT）
215
美国北部
美国南部

图 2

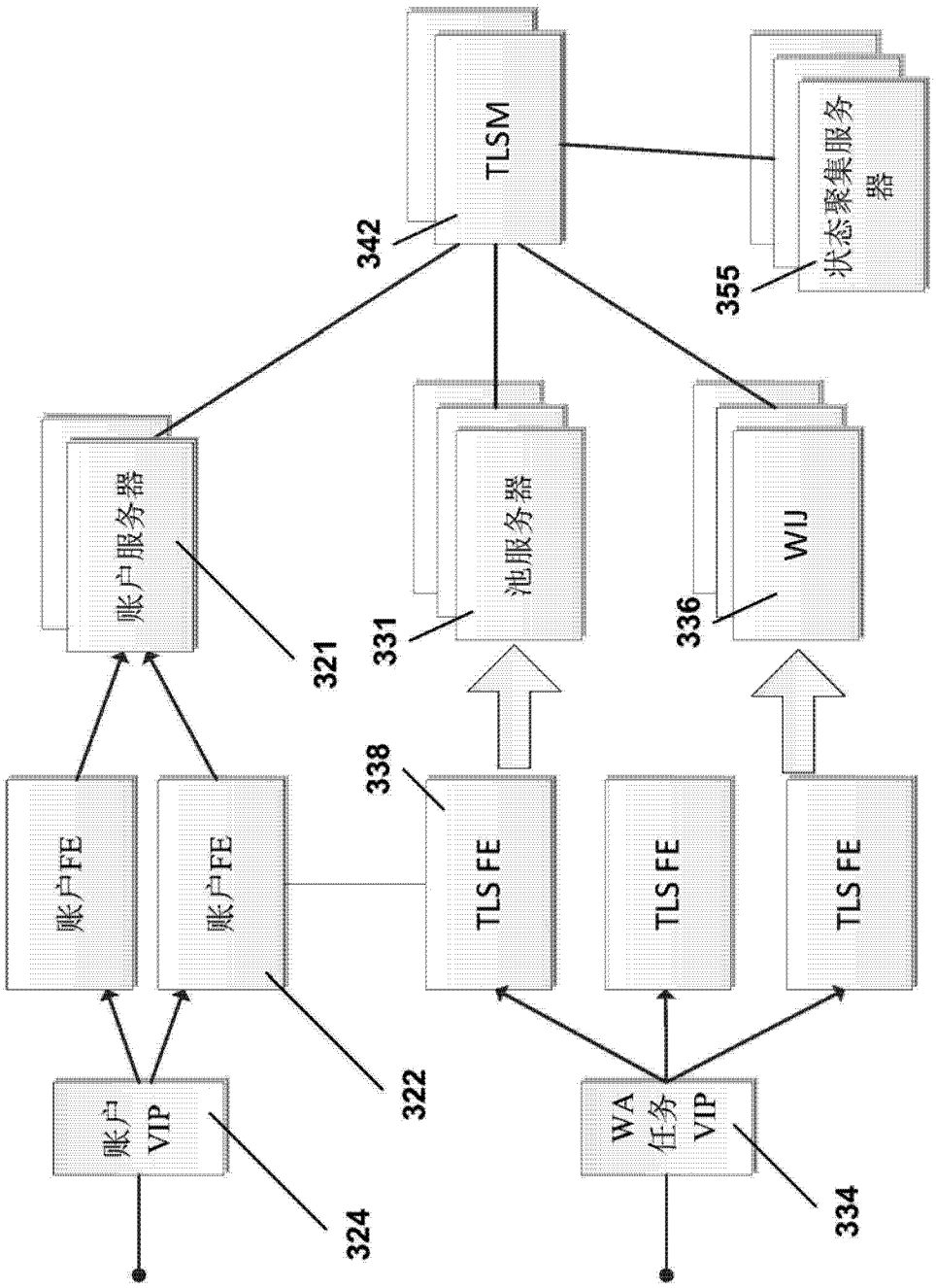

TLSM
342
状态聚集服务器
355
账户服务器
321
池服务器
331
WIJ
336
账户FE
账户FE
338
TLS FE
TLS FE
TLS FE
322
账户 VIP
324
WA 任务 VIP
334

图 3

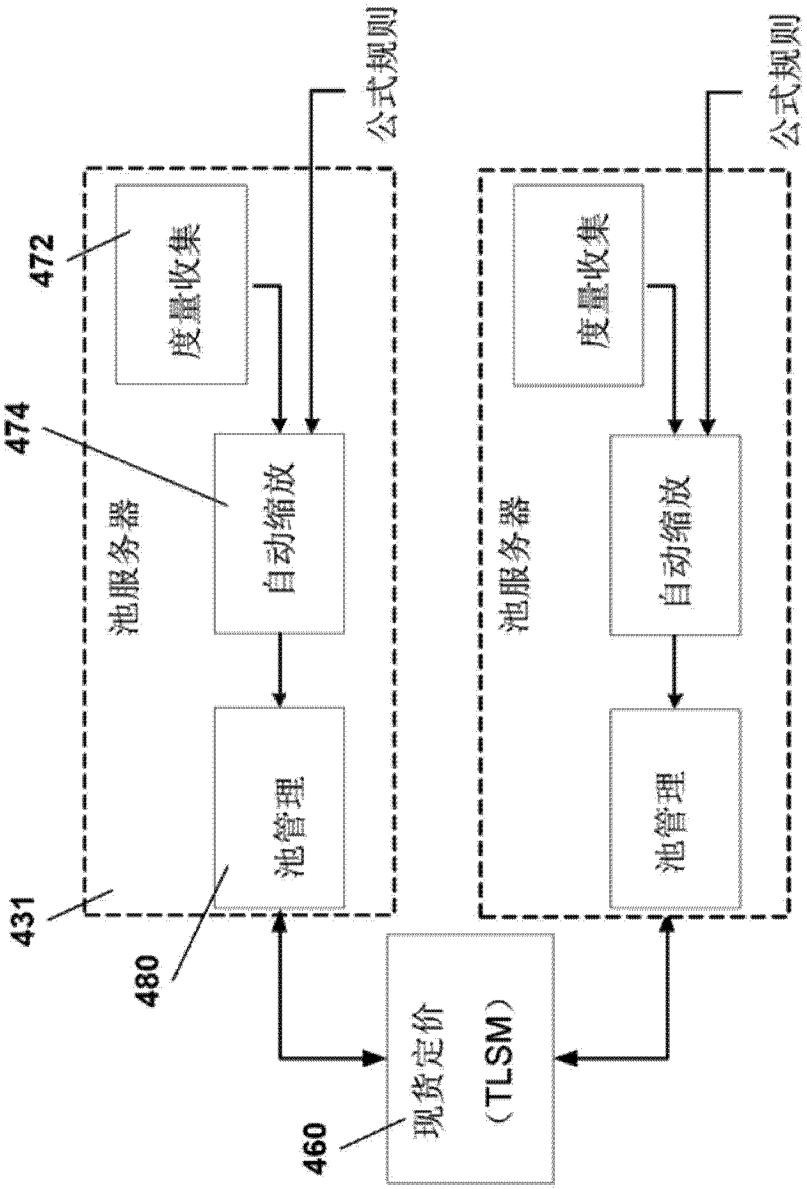

公式规则
472
度量收集
474
自动缩放
池服务器
池管理
431
480
460
现货定价
（TLSM）
公式规则
度量收集
自动缩放
池服务器
池管理

图 4

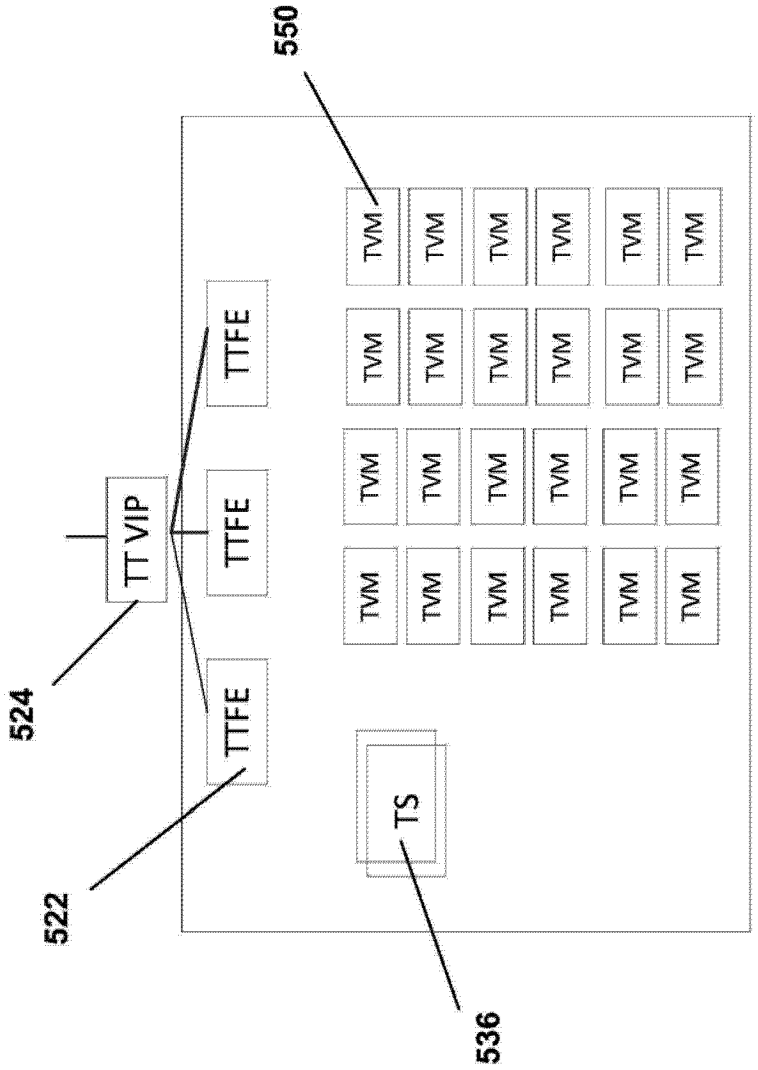


图 5

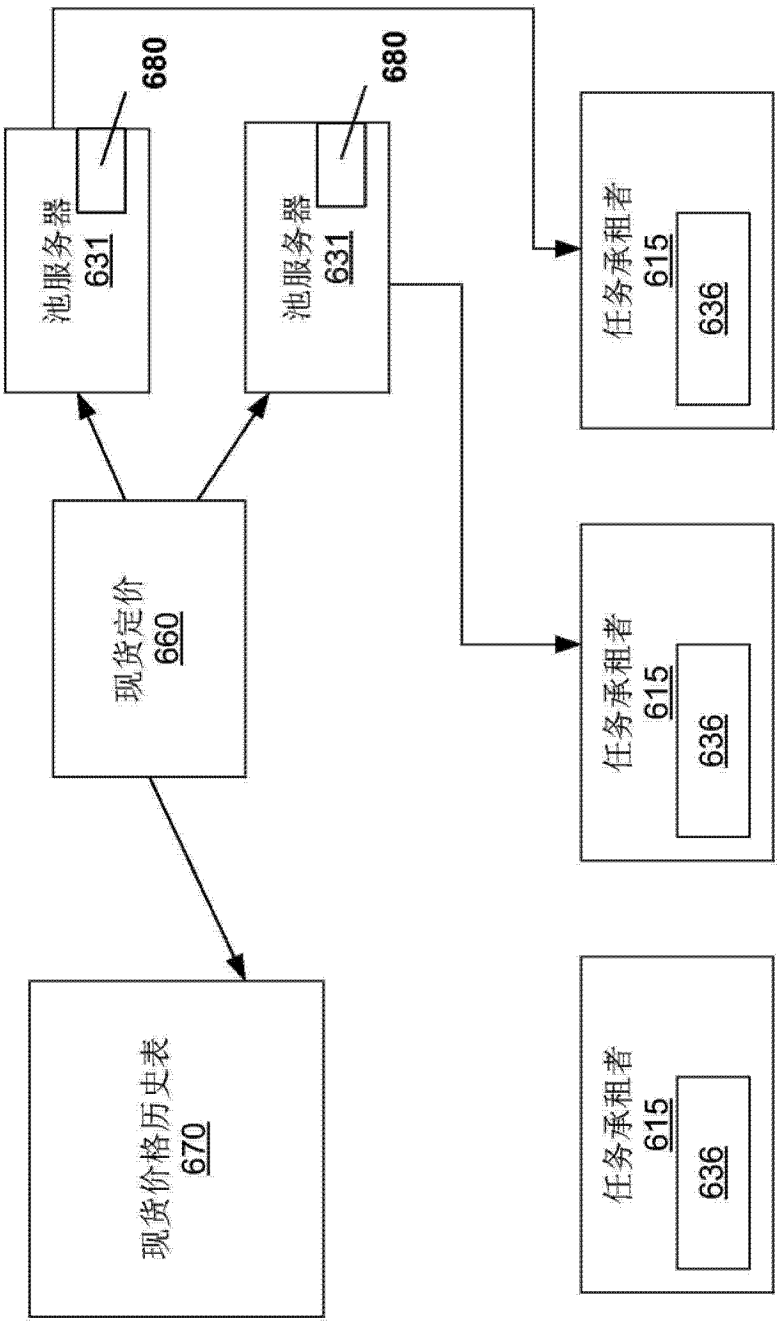

680
池服务器
631
680
池服务器
631
任务承租者
615
636
现货定价
660
任务承租者
615
636
现货价格历史表
670
任务承租者
615
636

图 6

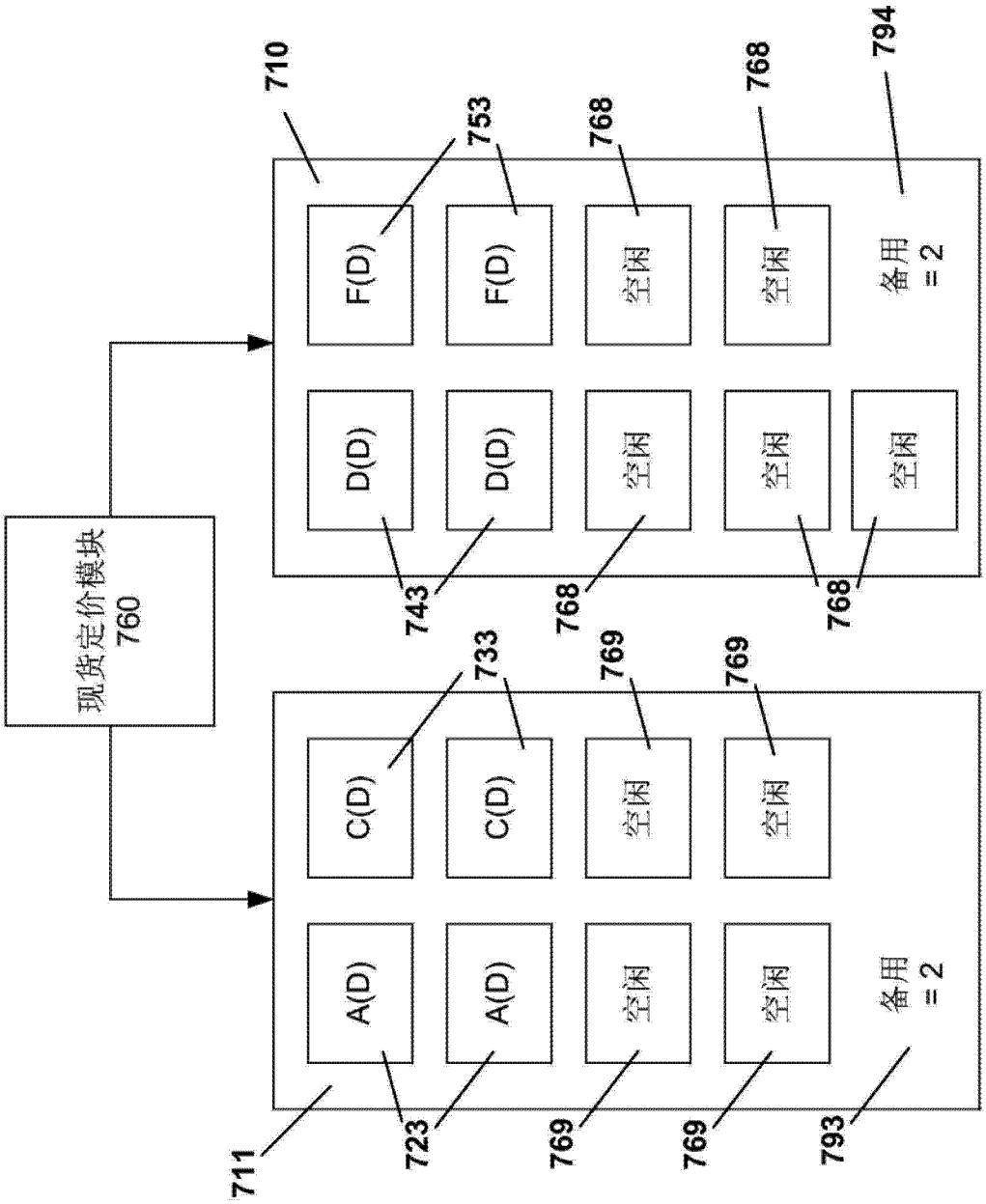

710
753
768
768
794
F(D)
F(D)
空闲
空闲
备用
= 2
D(D)
D(D)
空闲
空闲
空闲
743
768
768
现货定价模块
760
733
769
769
C(D)
C(D)
空闲
空闲
A(D)
A(D)
空闲
空闲
备用
= 2
711
723
769
769
793

图 7

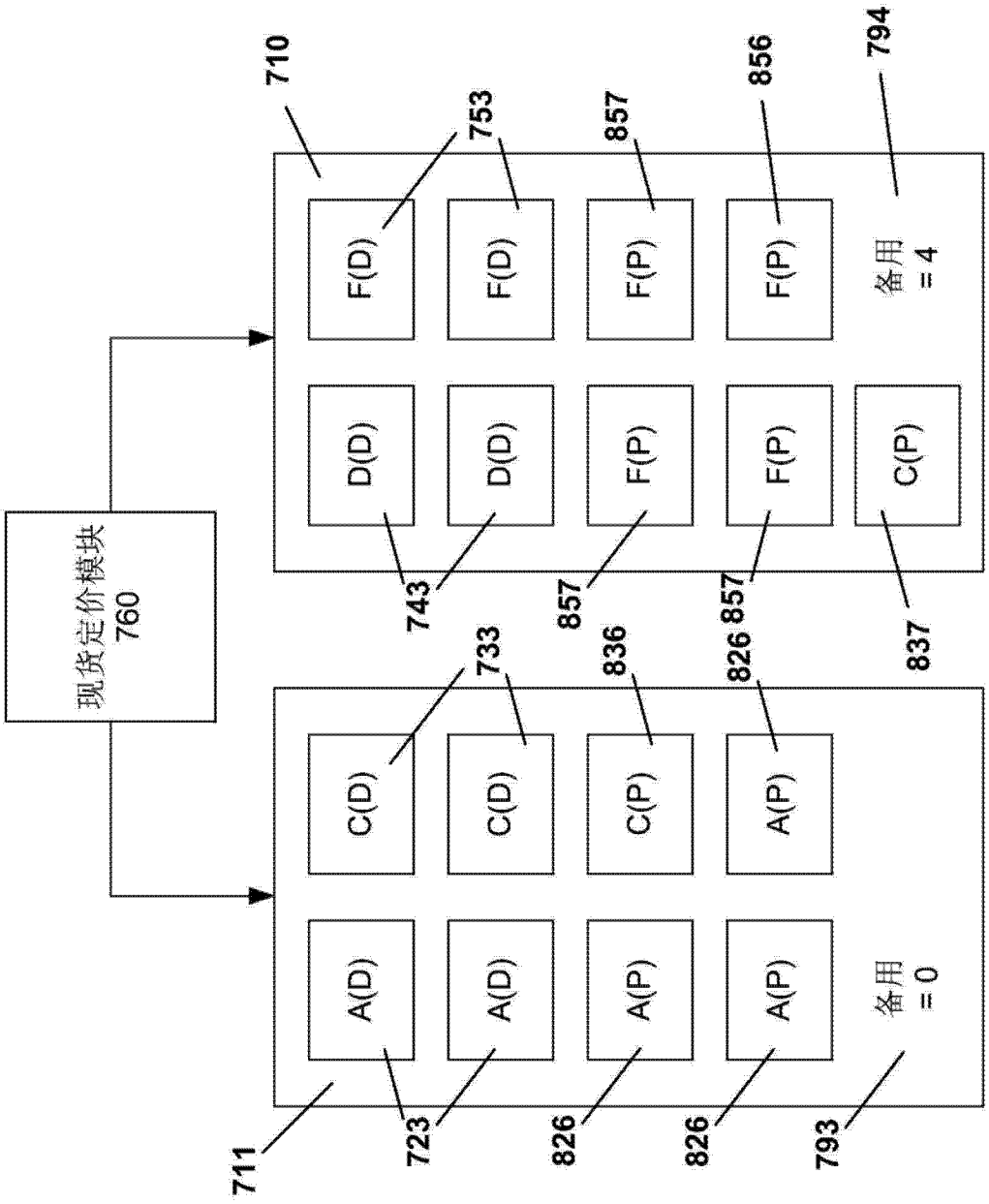

710
753
857
856
794
F(D)
F(D)
F(P)
F(P)
备用
= 4
D(D)
D(D)
F(P)
F(P)
C(P)
现货定价模块
760
743
857
857
826
837
733
836
C(D)
C(D)
C(P)
A(P)
A(D)
A(D)
A(P)
A(P)
备用
= 0
711
723
826
826
793

图 8

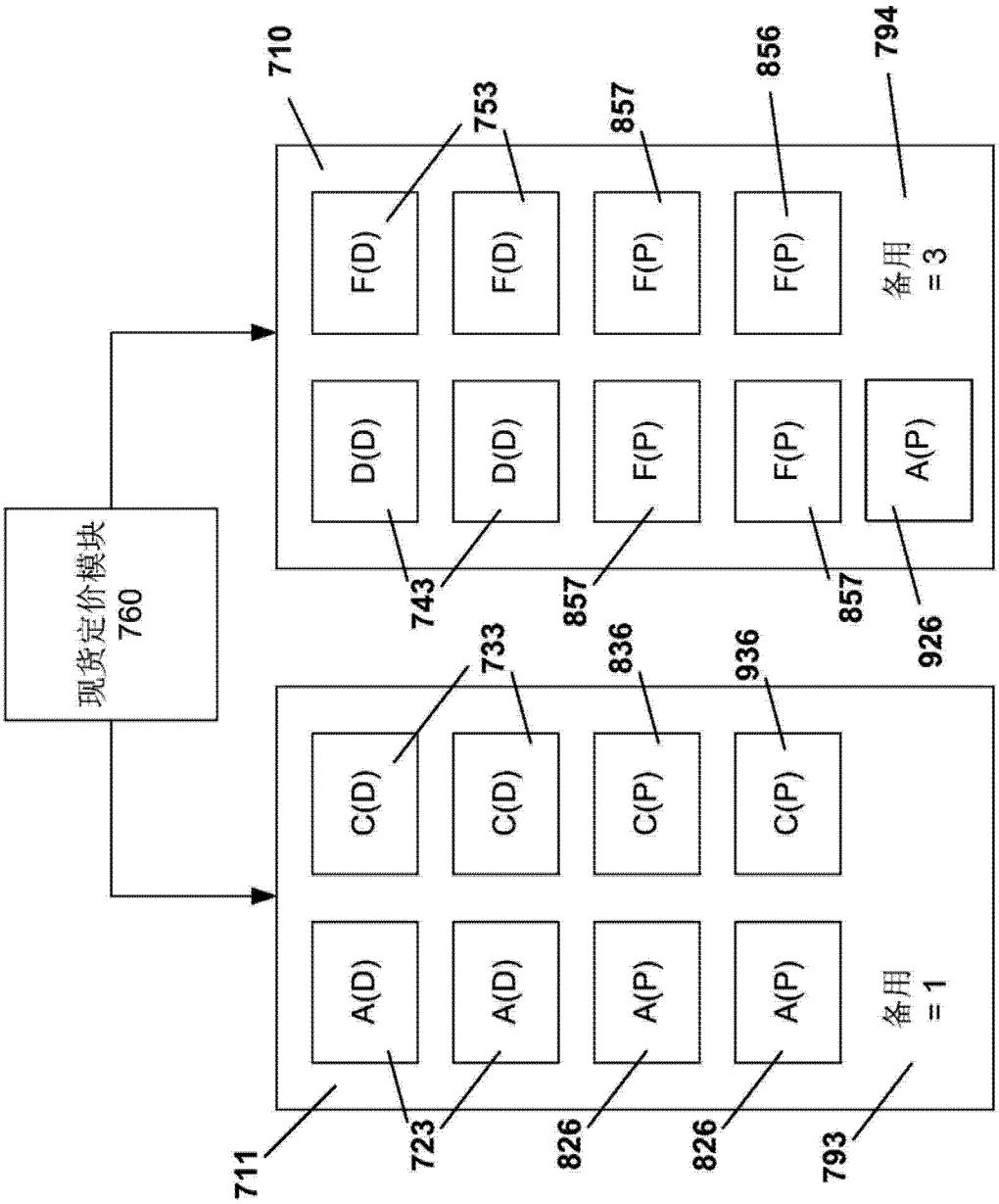

710
753
857
856
794
F(D)
F(D)
F(P)
F(P)
备用
= 3
D(D)
D(D)
F(P)
F(P)
A(P)
743
857
857
926
现货定价模块
760
733
836
936
C(D)
C(D)
C(P)
C(P)
A(D)
A(D)
A(P)
A(P)
备用
= 1
711
723
826
826
793

图 9

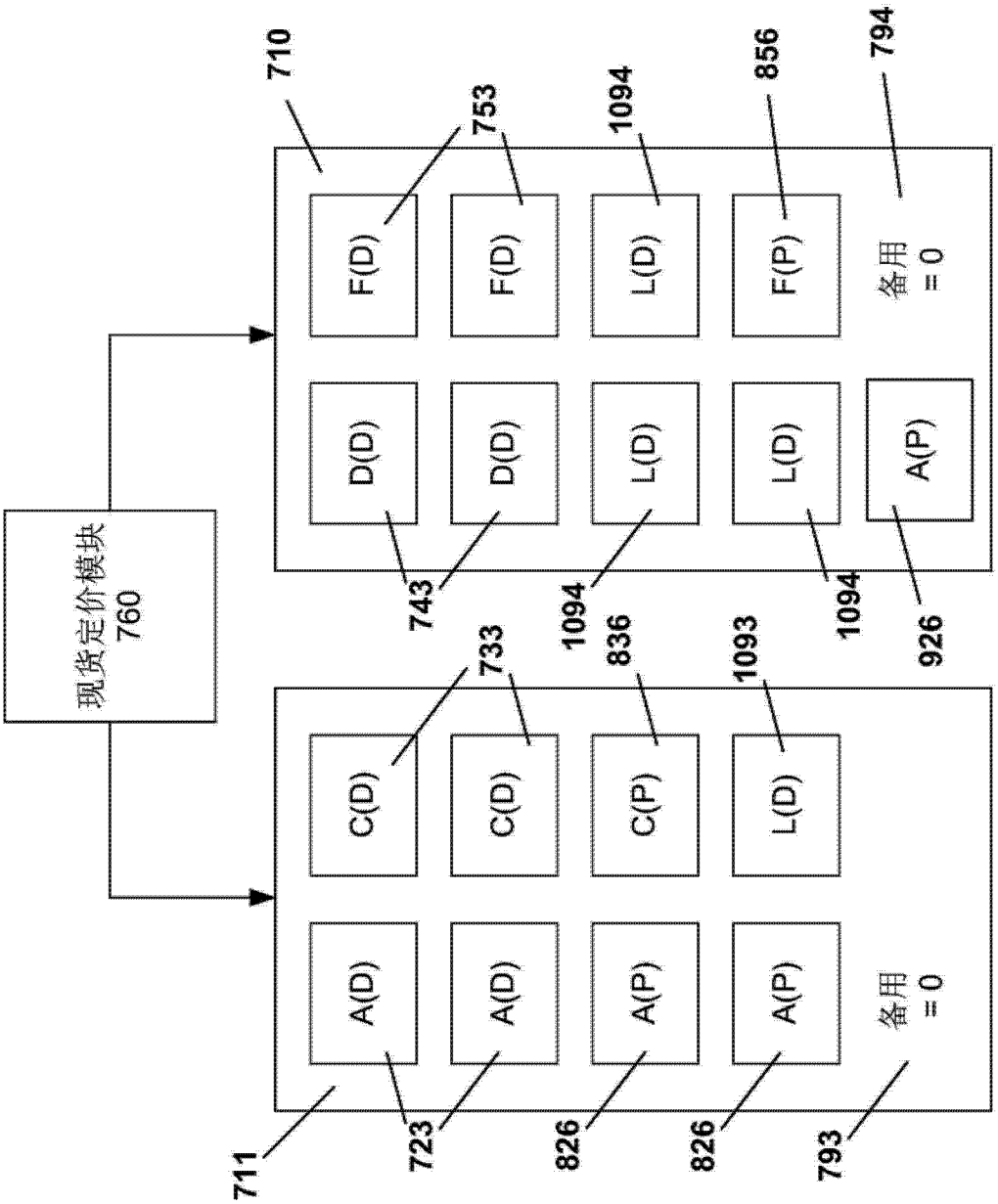

710
753
1094
856
794
F(D)
F(D)
L(D)
F(P)
备用
= 0
D(D)
D(D)
L(D)
L(D)
A(P)
743
1094
1094
926
现货定价模块
760
733
836
1093
C(D)
C(D)
C(P)
L(D)
A(D)
A(D)
A(P)
A(P)
备用
= 0
711
723
826
826
793

图 10

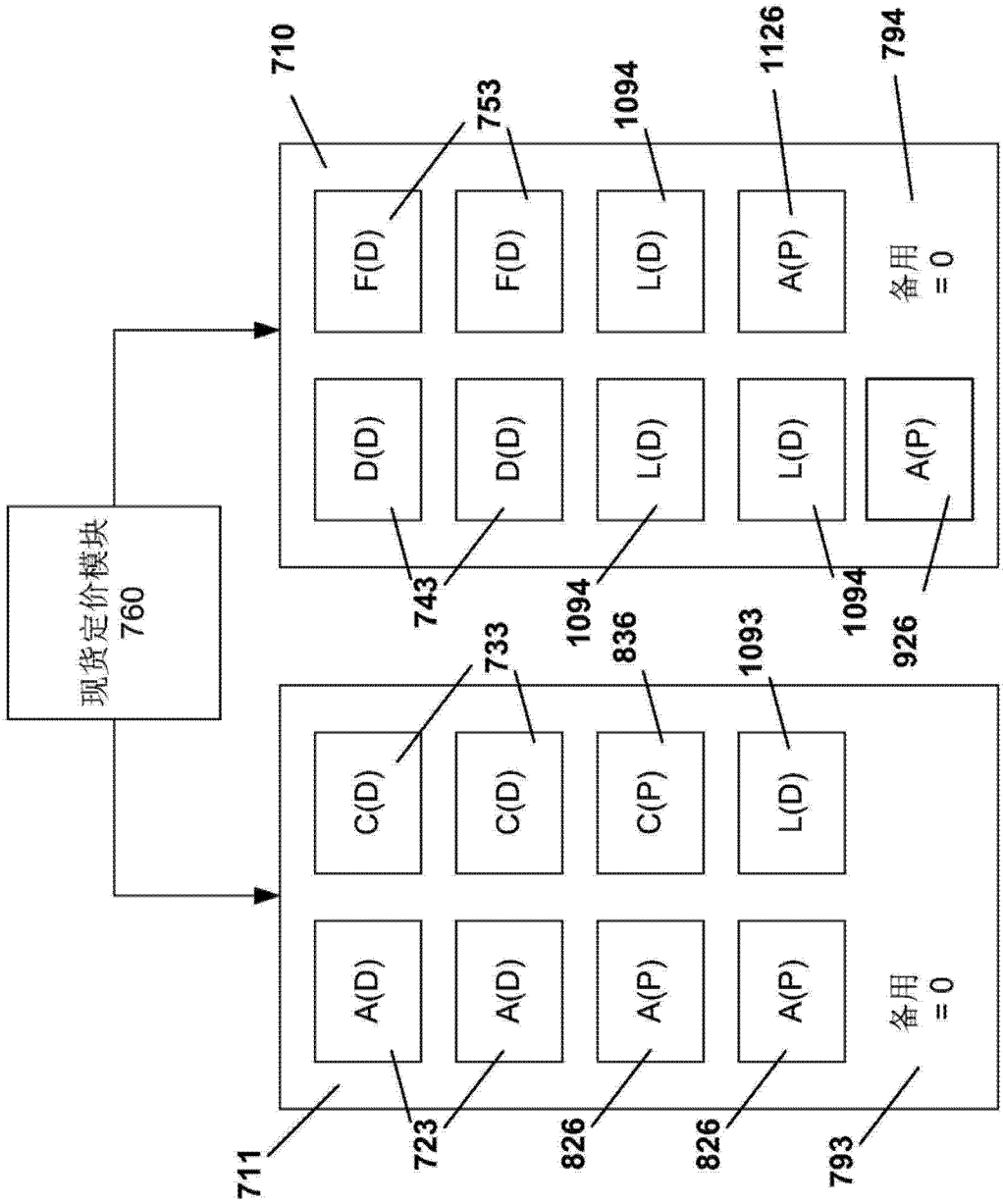

710
753
1094
1126
794
F(D)
F(D)
L(D)
A(P)
备用 = 0
D(D)
D(D)
L(D)
L(D)
A(P)
743
1094
1094
926
现货定价模块 760
733
836
1093
C(D)
C(D)
C(P)
L(D)
A(D)
A(D)
A(P)
A(P)
备用 = 0
711
723
826
826
793

图 11

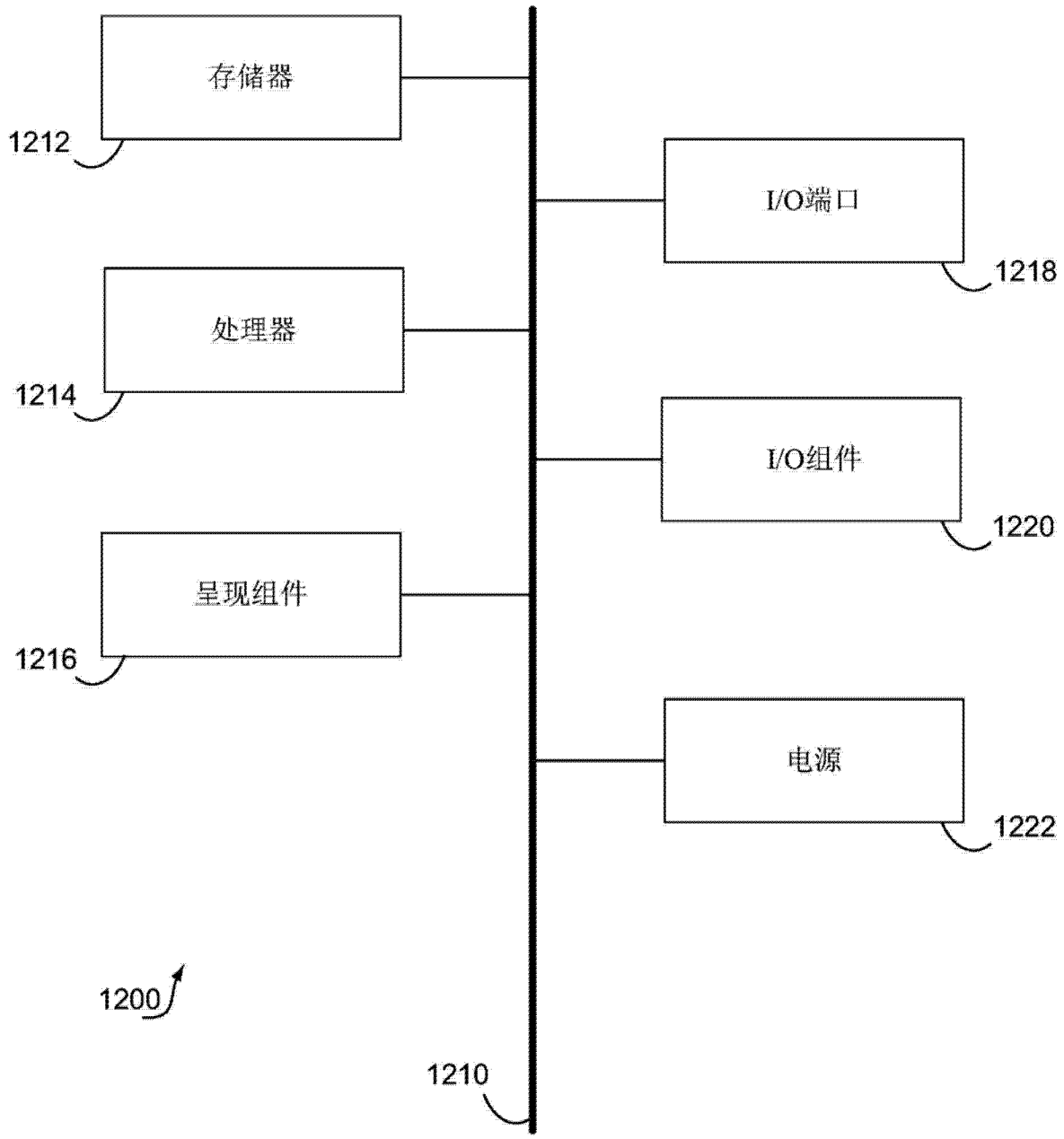

存储器
1212
处理器
1214
呈现组件
1216
I/O端口
1218
I/O组件
1220
电源
1222
1200
1210

图 12

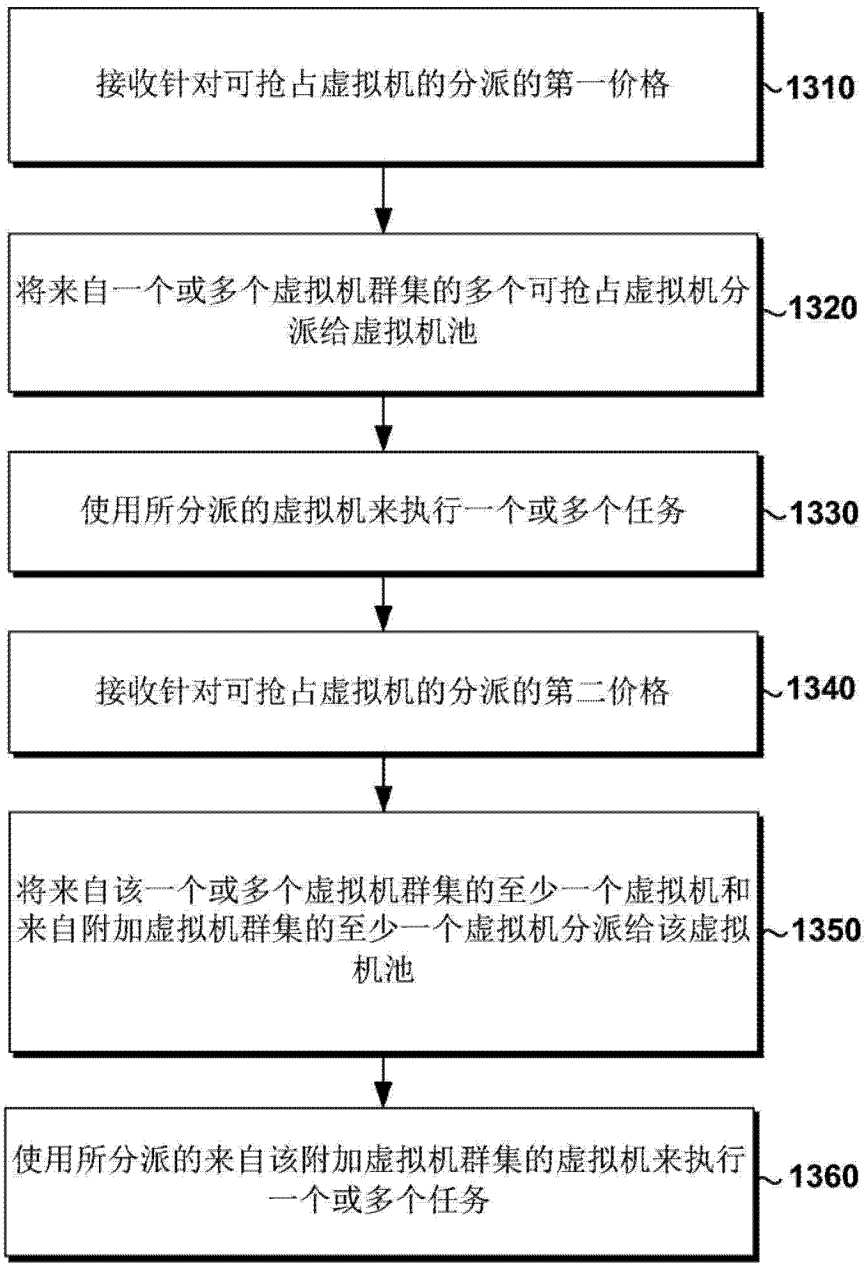

接收针对可抢占虚拟机的分派的第一价格
1310
将来自一个或多个虚拟机群集的多个可抢占虚拟机分派给虚拟机池
1320
使用所分派的虚拟机来执行一个或多个任务
1330
接收针对可抢占虚拟机的分派的第二价格
1340
将来自该一个或多个虚拟机群集的至少一个虚拟机和来自附加虚拟机群集的至少一个虚拟机分派给该虚拟机池
1350
使用所分派的来自该附加虚拟机群集的虚拟机来执行一个或多个任务
1360

图 13

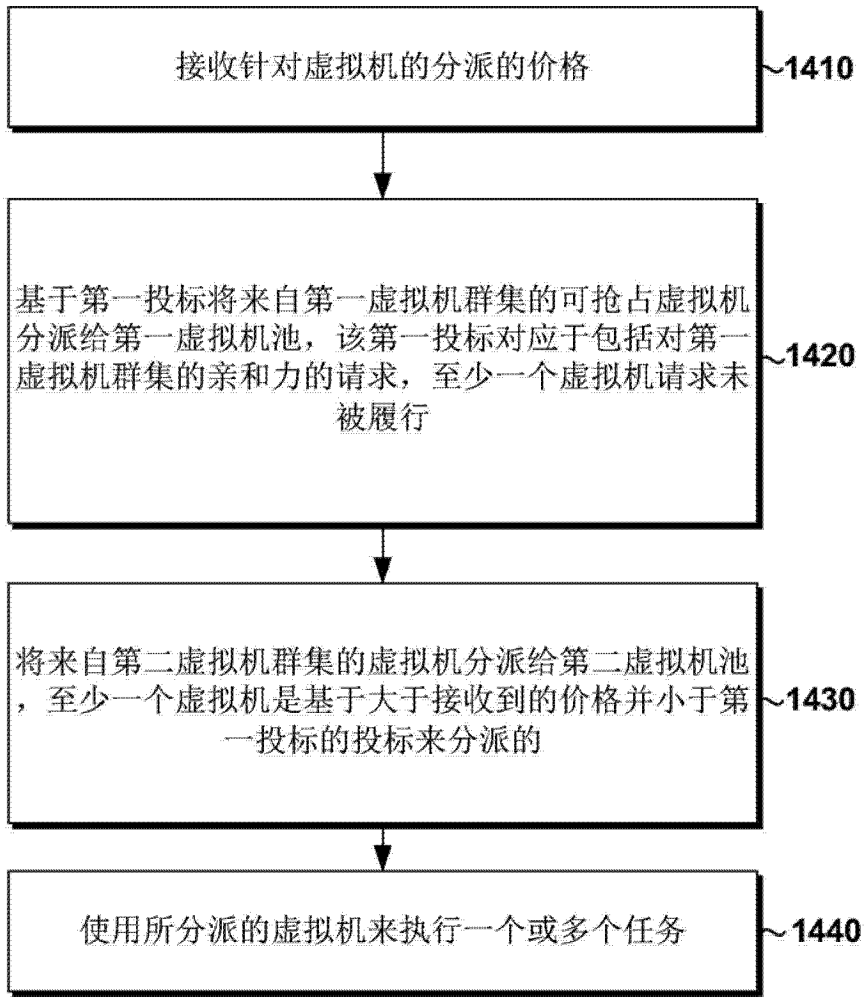

接收针对虚拟机的分派的价格
1410
基于第一投标将来自第一虚拟机群集的可抢占虚拟机分派给第一虚拟机池，该第一投标对应于包括对第一虚拟机群集的亲和力的请求，至少一个虚拟机请求未被履行
1420
将来自第二虚拟机群集的虚拟机分派给第二虚拟机池，至少一个虚拟机是基于大于接收到的价格并小于第一投标的投标来分派的
1430
使用所分派的虚拟机来执行一个或多个任务
1440

图 14

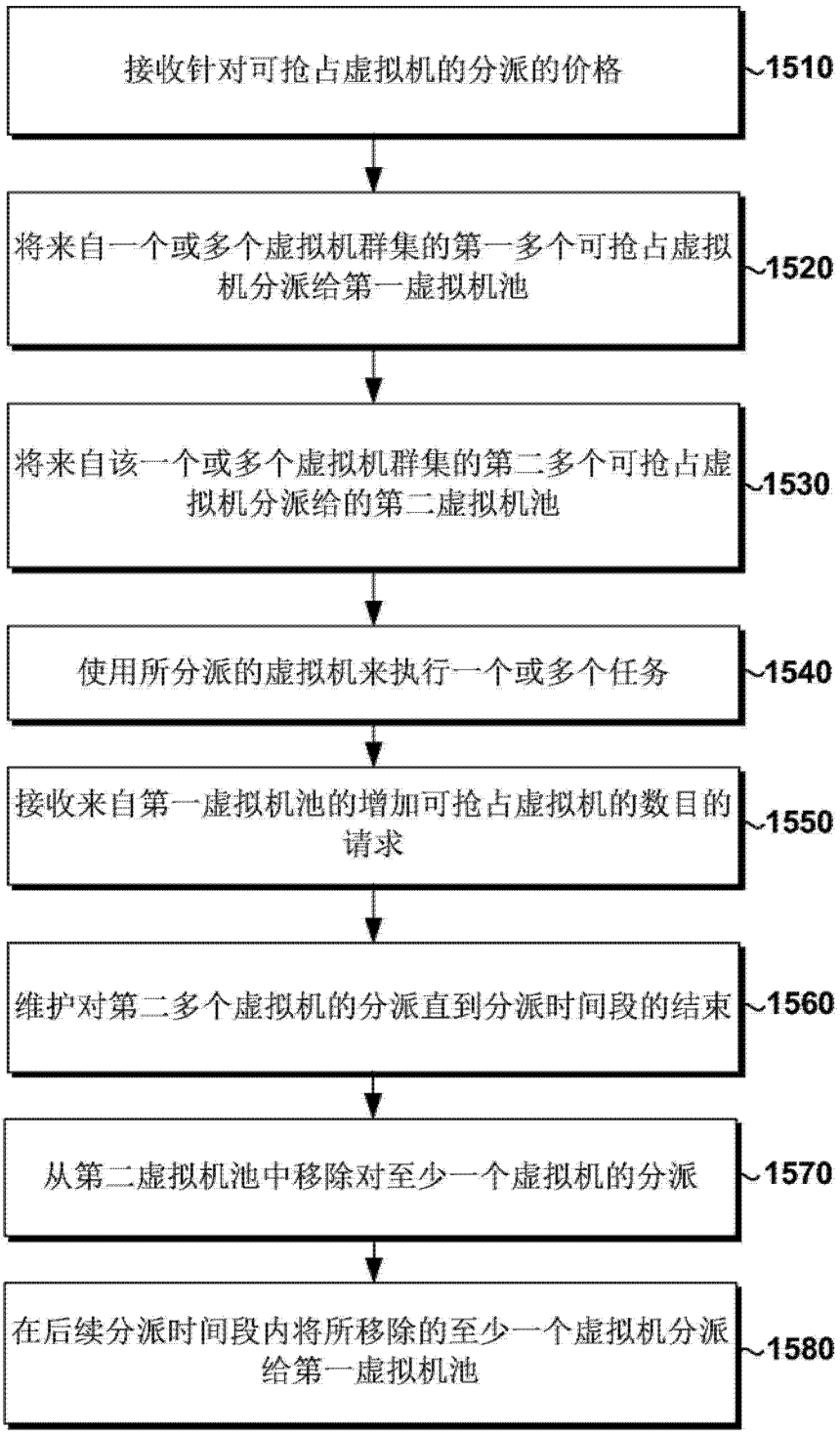

接收针对可抢占虚拟机的分派的价格
1510
将来自一个或多个虚拟机群集的第一多个可抢占虚拟机分派给第一虚拟机池
1520
将来自该一个或多个虚拟机群集的第二多个可抢占虚拟机分派给的第二虚拟机池
1530
使用所分派的虚拟机来执行一个或多个任务
1540
接收来自第一虚拟机池的增加可抢占虚拟机的数目的请求
1550
维护对第二多个虚拟机的分派直到分派时间段的结束
1560
从第二虚拟机池中移除对至少一个虚拟机的分派
1570
在后续分派时间段内将所移除的至少一个虚拟机分派给第一虚拟机池
1580

图 15