一种资源管理方法以及相关装置

本申请要求于 2022 年 5 月 17 日提交中国专利局、申请号为“202210536539.3”、申请名称为“一种资源管理方法以及相关装置”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本申请实施例涉及计算机领域，尤其涉及一种资源管理方法以及相关装置。

背景技术

随着云计算技术的发展，云服务的提供商将大量的计算资源通过网络连接起来构成一个巨大的资源池为用户设备提供云服务，用户设备通过资源管理服务器管理和调度资源池中的计算资源。

目前用户设备与云服务的提供商在云资源的交易过程中，用户设备向资源管理服务器发出云资源申请，云资源申请中包括计算资源的规格，例如处理器核心数和内存大小等。云服务的提供商基于当前资源池中的云资源的使用情况为用户设备提供云资源申请相应的价格信息，用户设备根据云服务提供商提供的价格信息选择订购云服务。

目前云服务的提供商仅根据用户设备在资源申请时刻的资源池中的可用资源状态信息为用户设备提供价格信息，而由于资源池中的可用资源状态信息会随着不同时刻使用的用户设备发生变化，因此，目前交易方式云服务的提供商提供的价格信息不准确，进一步导致资源池中的计算资源浪费。

发明内容

本申请实施例提供了一种资源管理方法以及相关装置，用于提升云资源的利用率。

本申请实施例第一方面提供了一种资源管理方法，该方法可以由管理服务器执行，也可以由管理服务器的部件，例如管理服务器的处理器、芯片或芯片系统等执行，还可以由能实现全部或部分管理服务器功能的逻辑模块或软件实现。以管理服务器执行为例，本申请实施例第一方面提供的资源管理方法包括：管理服务器获取云服务器集群的可用资源状态信息，可用资源状态信息包括以下一项或多项：云服务器集群在未来不同时段的可用云资源容量和未来不同时段的可用云资源规格，其中可用云资源容量包括未来时段可用的虚拟机、物理主机或容器的数量，可用云资源规格包括处理器核数、内存大小或显卡型号。管理服务器根据可用资源状态信息生成云服务价格信息，云服务价格信息包括未来不同时段的一个或者多个规格的云资源的价格信息。管理服务器向用户设备发送云服务价格信息。

本申请实施例中管理服务器能够获取云服务器集群在未来时段的可用云资源状态信息，并根据可用云资源状态信息生成相应的云服务价格，由于该云服务价格是根据未来不同时段的可用云资源类型和可用云资源容量计算得到的价格，因此，该云服务价格能准确反映出可用云资源的充足或紧缺情况，更利于云资源的销售，从而提升云服务器中云资源

的利用效率。

一种可能的实施方式中，管理服务器获取云服务器集群的可用资源状态信息之前，管理服务器接收用户设备发送的计划信息，计划信息用于向管理服务器提供用户设备对云资源的需求计划，计划信息包括以下一项或多项信息：云资源的需求时段或云资源的需求规格。管理服务器根据可用资源状态信息生成云服务价格信息的过程中，管理服务器根据计划信息和可用资源状态信息生成云服务价格信息，该云服务价格信息为计划信息所述对应的云服务价格信息。

本申请实施例中管理服务器接收用户设备提交的计划信息，并基于用户设备提交的计划信息和可用资源状态生成计划信息对应的云服务价格，由于用户设备预先提交了未来时段对云资源的需求容量和需求规格，因此，管理服务器能够基于计划信息提供更准确的云服务价格，进一步提升了云资源的利用效率。

一种可能的实施方式中，可用资源状态信息中的可用云资源规格与计划信息中的需求规格一致，管理服务器根据用户设备提交的计划信息中的需求规格确定相应的规格的云资源。

本申请实施例提供管理服务器提供的可用云资源规格与计划信息中的需求规格一致，从而可用云资源规格满足计划信息的需求规格，提升了方式的可实现性。

一种可能的实施方式中，管理服务器根据计划信息和可用资源状态信息生成云服务价格信息的过程中，当可用资源状态信息满足计划信息，且可用资源状态信息中可用云资源容量大于或等于目标阈值时，管理服务器生成云服务折扣信息。

本申请实施例中管理服务器对对于预先提交计划信息的用户设备可以提供云服务的折扣价格，从而提升了云资源的利用率。

一种可能的实施方式中，管理服务器向用户设备发送云服务价格信息之后，管理服务器接收用户设备发送的订购信息，订购信息用于确认计划信息对应的云服务价格信息。

本申请实施例中管理服务器向用户设备发送云服务价格信息之后，接收用户设备发送计划信息对应的订购请求，从而提升了方案可实现性。

一种可能的实施方式中，管理服务器向用户设备发送云服务价格信息前，管理服务器接收用户设备发送的查询请求。管理服务器根据查询请求向用户设备发送可用资源状态信息和云服务价格信息，管理服务器将未来时段的可用云资源规格以及可用云资源容量以及对应的价格都发送给用户，由用户设备选择需要订购的时段、云资源规格以及云资源容量。

本申请实施例中管理服务器可以基于用户设备发送的查询请求将未来时段的可用资源状态信息和相应的云服务价格信息发送至用户设备，由用户设备自己选择要订购的云资源需求时段、云资源规格以及云资源容量，从而提升了云资源的利用率。

一种可能的实施方式中，管理服务器根据查询请求向用户设备发送可用资源状态信息和云服务价格信息之后，管理服务器接收用户设备发送的订购信息，订购信息包括用户设备选择的至少一个规格的云资源和至少一个规格的云资源的价格。

本申请实施例中管理服务器向用户设备发送可用资源状态信息和云服务价格信息之后，接收用户设备选择的云资源需求时段、云资源规格以及云资源容量，从而提升了方案

可实现性。

一种可能的实施方式中，管理服务器获取云服务器集群的可用资源状态信息的过程中，管理服务器获取云服务器集群的调度信息，调度信息包括以下一项或多项信息：已提交的云资源计划、基于历史云资源变化预测的云资源计划、需预留的空闲云资源计划、总的云资源计划、单设备的最大云资源计划和使用中云资源的延续计划。管理服务器器根据调度信息生成云服务器集群的可用资源状态信息。具体的，管理服务器将总的云资源计划减去已提交的云资源计划、基于历史云资源变化预测的云资源计划、需预留的空闲云资源计划、使用中云资源的延续计划即可以得到不同时段的可用云资源容量。

本申请实施例中管理服务器可以定期获取云服务器集群的调度信息，并基于调度信息计算得到可用资源状态信息，从而提升了可用资源状态信息的准确性。

本申请实施例第二方面提供了一种资源管理装置，该资源管理装置包括收发单元和处理单元。其中，收发单元用于获取云服务器集群的可用资源状态信息，可用资源状态信息包括以下一项或多项：云服务器集群在未来不同时段的可用云资源容量和未来不同时段的可用云资源规格。处理单元用于根据可用资源状态信息生成云服务价格信息，云服务价格信息包括多个规格的云资源的价格信息。收发单元还用于向用户设备发送云服务价格信息。

一种可能的实施方式中，收发单元还用于接收用户设备发送的计划信息，计划信息用于向管理服务器提供用户设备对云资源的需求计划，计划信息包括以下一项或多项信息：云资源的需求时段或云资源的需求规格。处理单元具体用于根据计划信息和可用资源状态信息生成云服务价格信息。

一种可能的实施方式中，可用资源状态信息中的可用云资源规格与计划信息中的需求规格一致。

一种可能的实施方式中，处理单元具体用于当可用资源状态信息满足计划信息，且可用资源状态信息中可用云资源容量大于或等于目标阈值时，生成云服务折扣信息。

一种可能的实施方式中，收发单元还用于接收用户设备发送的订购信息，订购信息用于确认计划信息对应的云服务价格信息。

一种可能的实施方式中，收发单元还用于接收用户设备发送的查询请求。处理单元还用于根据查询请求向用户设备发送可用资源状态信息和云服务价格信息。

一种可能的实施方式中，收发单元还用于接收用户设备发送的订购信息，订购信息包括用户设备选择的至少一个规格的云资源和至少一个规格的云资源的价格。

一种可能的实施方式中，收发单元还用于获取云服务器集群的调度信息，调度信息包括以下一项或多项信息：已提交的云资源计划、基于历史云资源变化预测的云资源计划、需预留的空闲云资源计划、总的云资源计划、单设备的最大云资源计划和使用中云资源的延续计划。处理单元还用于根据调度信息生成云服务器集群的可用资源状态信息。

本申请实施例第三方面提供了一种管理服务器，包括处理器，处理器与存储器耦合，处理器用于存储指令，当指令被处理器执行时，以使得管理服务器执行上述第一方面或第一方面任意一种可能的实施方式所述的方法。

本申请实施例第四方面提供了一种计算机可读存储介质，其上存储有指令，指令被执行时，以使得计算机执行上述第一方面或第一方面任意一种可能的实施方式所述的方法。

本申请实施例第五方面提供了一种计算机程序产品，计算机程序产品中包括指令，指令被执行时，以使得计算机实现上述第一方面或第一方面任意一种可能的实施方式所述的方法。

可以理解，上述提供的任一种资源管理装置、管理服务器、计算机可读介质或计算机程序产品等所能达到的有益效果可参考对应的方法中的有益效果，此处不再赘述。

附图说明

图 1 为本申请实施例提供的一种云资源调度系统的系统架构示意图；

图 2 为本申请实施例提供的一种资源管理方法的流程示意图；

图 3 为本申请实施例提供的一种云资源调度计划示意图；

图 4a 为本申请实施例提供的另一种资源管理方法的流程示意图；

图 4b 为本申请实施例提供的一种管理服务器 UI 界面示意图；

图 5 为本申请实施例提供的另一种资源管理方法的流程示意图；

图 6 为本申请实施例提供的一种资源管理装置的结构示意图；

图 7 为本申请实施例提供的一种管理服务器的结构示意图。

具体实施方式

本申请实施例提供了一种资源管理方法以及相关装置，用于提升云服务器的资源利用效率。

本申请的说明书和权利要求书及上述附图中的术语“第一”、“第二”、“第三”、“第四”等（如果存在）是用于区别类似的对象，而不必用于描述特定的顺序或先后次序。应该理解这样使用的数据在适当情况下可以互换，以便这里描述的实施例能够以除了在这里图示或描述的内容以外的顺序实施。此外，术语“包括”和“具有”以及他们的任何变形，意图在于覆盖不排他的包含，例如，包含了一系列步骤或单元的过程、方法、系统、产品或设备不必限于清楚地列出的那些步骤或单元，而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它步骤或单元。

在本申请实施例中，“示例性的”或者“例如”等词用于表示作例子、例证或说明。本申请实施例中被描述为“示例性的”或者“例如”的任何实施例或设计方案不应被解释为比其它实施例或设计方案更优选或更具优势。确切而言，使用“示例性的”或者“例如”等词旨在以具体方式呈现相关概念。

下面结合附图介绍本申请实施例提供的资源管理方法以及相关装置。

请参阅图 1，图 1 为本申请实施例提供的一种云资源调度系统的系统架构示意图。如图 1 所示，该云资源调度系统 100 包括管理服务器 101、云服务器 102 和用户设备 103，其中，用户设备 103 通过管理服务器 101 与云服务器 102 相连接。下面具体介绍云资源调度系统 100 中各部分的功能。

管理服务器 101 用于对云服务器 102 提供的云资源进行管理和调度，包括收集云服务器 102 中所有物理主机的资源状态，对用户设备 103 发出新的云资源申请进行调度，以及为用户设备 103 提供云服务的价格信息。

管理服务器 101 中包括计划资源交易模块 1011 和资源调度模块 1012，其中计算资源交易模块 1011 用于测算云服务器 102 的可用资源状态信息是否满足用户设备 103 的云资源申请需求，以及根据可用资源状态信息生成云服务的价格信息。资源调度模块 1012 用于获取云服务器 102 的可用资源状态信息，以及对云服务器 102 提供的云资源进行调度。

在一些可能的实现方式中，管理服务器 101 还能够对外提供可进行操作的 UI 界面，以及可以被调用的应用程序编程接口 API。管理服务器 101 还能够提供对外提供云服务器 102 中的云资源的交易服务。管理服务器 101 例如 OpenStack 和 VMware 等。

云服务器 102 用于为用户设备 103 提供用云资源。云服务器 103 中包括多个物理主机，其中每个物理主机上部署多个虚拟机。云服务器 102 提供的云资源包括物理主机、虚拟机和容器。云服务器 102 根据管理服务器 101 的调度计划为用户设备 103 提供可用的云资源。可以理解的是，多个云服务器 102 也组成云服务集群，由服务器集群为用户设备 103 提供云服务。

用户设备 103 用于向管理服务器 101 请求云资源。具体的，用户设备 103 可以基于管理服务器 101 提供的 UI 界面或者 web UI 界面向管理服务器 101 发送云资源的请求，也可以通过用户设备 103 的应用软件向管理服务器 101 发送资源请求，具体不做限定。

下面以图 1 所示的系统架构为例，介绍本申请实施例提供的资源管理方法。

请参阅图 2，图 2 为本申请实施例提供的一种资源管理方法的流程示意图。如图 2 所示，本申请实施例提供的资源管理方法包括但不限于步骤：

201.管理服务器获取云服务器集群的可用资源状态信息，可用资源状态信息包括以下一项或多项信息：云服务器集群在未来不同时段的可用云资源容量和可用云资源规格。

管理服务器获取云服务器集群的可用资源状态信息，可用资源状态信息包括以下一项或多项信息：云服务器在未来不同时段的可用云资源容量和可用云资源规格。本申请实施例中的云资源的类型包括虚拟机、物理主机和容器，其中每种类型的云资源下包括不同的规格，例如，虚拟机的规格包括处理器核数、内存大小和显卡型号。可用云资源容量包括云服务器集群可以提供给用户设备使用的虚拟机、物理主机和容器数量。

具体的，管理服务器在获取云服务器集群的可用资源状态信息的过程中，管理服务器的计划交易模块向资源调度模块查询调度信息，管理服务器的计划交易模块基于调度信息生成云服务器集群的可用资源状态信息。其中云服务器集群的调度信息包括以下一项或多项信息：已提交的云资源计划、基于历史云资源变化预测的云资源计划、需预留的空闲云资源计划、总的云资源计划、单设备的最大云资源计划和使用中云资源的延续计划。

其中，已经提交资源计划申请包括未来时段内用户设备对云资源的需求容量和需求规格，上述用户设备包括当前已提交计划信息但还未完成订购的用户设备和其他已经提交完成计划信息并完成订购的用户设备。

基于历史云资源变化预测的云资源计划是指管理服务器根据历史云资源使用情况预测

得到的未来时段云资源使用变化情况的计划。

需预留的空闲云资源计划是指管理服务为了应对突发的云资源使用需求而计划预留的云资源，需预留的空闲资源包括当前时间到计划申请开始时间中的可能出现的新用户设备的资源申请，或者计划申请开始时间之后已有用户设备的新增资源申请。

总的云资源申请计划是指云服务器集群能够提供的可供计划的最大云资源容量，包括不同资源类型下每种资源类型能够提供的最大资源容量。

单设备的最大云资源计划是指管理服务器为单个用户设备所能分配的最大云资源容量设置的云资源计划。使用中云资源的延续计划是指管理服务器为用户设备延续正在使用的云资源预留的云资源计划。

请参阅图 3，图 3 为本申请实施例提供的一种云服务器集群的云资源容量随着时间变化的示意图。在图 3 所示的示例中，从 0 至 t6 时间段云资源容量的计划分配情况可以看出，在 0 至 t6 的时间段中每个时刻云服务器集群的总的云资源容量为 S7，在 0 至 t6 的时间段中每个时刻云服务器集群需预留的空闲云资源容量为 S6-S5，在 0 至 t6 的时间段中每个时刻使用中云资源的延续的云资源容量为 S0。在 t1 至 t4 的时间段中每个时刻用户设备已提交的云资源计划 1 的云资源容量为 S5-S3，在 t0 至 t2 的时间段中每个时刻基于历史云资源变化预测的云资源计划 1 中的云资源容量为 S1-S0，在 t3 至 t5 的时间段中每个时刻基于历史云资源变化预测的云资源计划 2 中的云资源容量为 S1-S0。

从图 3 所示的示例中可以看出，在 0 至 t6 的时间段中每个时刻单设备的最大云资源计划中的云资源容量为 S3-S2，该单设备的最大云资源计划中的云资源容量是指管理服务器为单个用户设备所能分配的最大云资源容量，图中用户设备已提交的资源计划 1、或者用户设备已提交的资源计划 2 以及用户设备已提交的资源计划 3 都不能超出该单设备的最大资源容量。

在图 3 所示的示例中，管理服务器进一步可以根据不同时段的云资源计划计算得到不同时段云服务器集群的可用云资源容量。例如，图 3 中总的云资源申请计划中减去需预留的空闲云资源计划、已提交的云资源计划 1、已提交的云资源计划 2、已提交的云资源计划 3、基于历史预测的云资源计划 1、基于历史预测的云资源计划 2 和使用中云资源的延续即可以得到可用云资源计划。该可用云资源计划可以指示未来不同时段下可用云资源容量，从图 3 所示的示例中可以看出，在 0 至 t0 时段以及在 t5 至 t6 时段，可用云资源容量达到最大值。

需要说明的是，在图 3 所示的示例中，需预留的空闲云资源计划、已提交的云资源计划、基于历史云资源变化预测的云资源计划以及使用中云资源的延续计划中云资源容量在时段内都未发生变化。在申请中的其他示例中，这些计划中的云资源容量在时段内也可以变化，例如，这些计划的云资源容量在时段内随时间变化增加或者减少。

此外，图 3 所示的示例仅为某一种云资源类型下的未来时段的云资源计划。对于不同的云资源类型，每一种云资源类型都有其对应的云资源计划，管理服务器可以基于未来时段的云资源计划计算不同云资源类型下的可用云资源计划。

本申请实施例中管理服务器可以定期获取云服务器集群的可用资源状态信息，也可以

基于用户设备的云资源请求触发管理服务器获取云服务器集群的可用资源状态信息，具体不做限定。

202.管理服务器根据可用资源状态信息生成云服务价格信息。

管理服务器获取可用资源状态信息之后，根据可用资源状态信息生成云服务价格信息。具体的，管理服务器可以根据可用资源那状态信息中不同时段下的可用云资源容量来生成不同时段的云服务的价格信息。当某个时段内可用云资源容量充足时，管理服务器可以生成该时段内云服务的折扣信息。

本申请实施例中管理服务器根据可用资源状态信息生成云服务价格信息的过程中，管理服务器可以根据用户设备提交的计划信息和可用资源状态信息生成计划信息对应的云服务价格信息，也可以将可用资源状态信息和相应的云服务价格信息发送至用户设备供用户设备进行选择，具体不做限定。后续图 4a 和图 5 所示的实施例会分别针对上述两种方式进行介绍。

可以理解的是，管理服务器对不同的用户设备也可以生成不同的云服务折扣信息。例如，在用户设备提交计划信息的场景下，管理服务器可以根据用户设备的云资源需求总量生成不同的折扣信息。再例如，管理服务器可以针对不同的用户等级生成不同的折扣信息。

203.管理服务器向用户设备发送云服务价格信息。

管理服务器向用户设备发送云服务价格信息。用户设备接收管理服务器发送的云服务价格信息之后，向管理服务器发送订购信息，该订购信息用于确认云服务价格信息。管理服务器根据订购信息为用户设备预留订购信息对应的云资源。

本申请实施例中管理服务器能够获取云服务器集群在未来时段的可用云资源状态信息，并根据可用云资源状态信息生成相应的云服务价格，由于该云服务价格是根据未来不同时段的可用云资源类型和可用云资源容量计算得到的价格，因此，该云服务价格能准确反映出可用云资源的充足或紧缺的情况，管理服务器按照该云服务价格对与服务器的云资源进行销售，从而提升云服务器中云资源的利用率。

下面分别结合图 4a 和图 5 分别对用户设备提交计划信息和管理服务器提供可用资源状态信息两种方式进行介绍。

请参阅图 4a，图 4a 为本申请实施例提供的另一种资源管理方法的流程示意图。图 4a 所示的资源管理方法包括以下步骤：

401.用户设备向管理服务器提交计划信息。

用户设备向管理服务器提交计划信息，该计划信息用于向管理服务器提供用户设备对云资源的需求计划，计划信息包括以下一项或多项信息：云资源的需求时段、云资源需求容量和云资源的需求规格。云资源的类型代号可以代表云资源的类型，例如，云资源的类型代号 g5 代表 16 核处理器、32g 内存、RTX5000 显卡的虚拟机。

本申请实施例中用户设备可以通过多种方式提交计划信息，例如，管理服务器可以通过管理服务器的 UI 界面提交计划信息，也可以通过应用程序接口 API 调用管理服务器中的计划交易模块提交计划信息，具体不做限定。

在步骤 401 的一个示例中，用户设备在 2022 年 3 月 30 日向管理服务器的计划交易模块的计划信息中包括：用户设备申请 100 台虚拟机，虚拟机的类型为 g5、需求开始时间为 2022 年 4 月 1 日 0：00，结束时间为 2022 年 4 月 15 日 24：00。

请参阅图 4b，图 4b 为本申请实施例提供的一种管理服务器的 UI 界面示意图。在图 4b 所示的示例中，用户设备可以通过管理服务器的 UI 界面提交用户设备的计划信息，用户设备提交的计划信息包括不同计划时间段的云资源的需求规格和云资源的需求容量。例如，用户设备在 2022 年 4 月 1 日至 2022 年 4 月 30 日的 6:00—12:00，云资源的需求规格为 G4 型虚拟机，云资源的需求容量为 100 台。

402. 管理服务器的计划交易模块从资源调度模块获取云服务器集群的调度信息。

管理服务器的计划交易模块从资源调度模块获取云服务器集群的调度信息，调度信息包括以下一项或多项信息：已提交的云资源计划、基于历史云资源变化预测的云资源计划、需预留的空闲云资源计划、总的云资源计划、单设备的最大云资源计划和使用中云资源的延续计划。

本申请实施例中，管理服务器的计划交易模块可以定期从资源调度模块查询云服务器集群的调度信息，也可以基于用户设备提交计划信息触发查询调度信息，或者资源调度模块也可以定期向计划交易模块上报调度信息，具体不做限定。

403. 管理服务器的计划交易模块根据云服务器集群的调度信息生成可用资源状态信息。

管理服务器的计划交易模块根据云服务器集群的调度信息生成可用资源状态信息。具体的，管理服务器的计划交易模块根据调度信息生成的可用资源状态信息包括以下一项或多项：未来不同可用时段内的可用云资源容量或可用云资源规格。

例如，在步骤 403 的一个示例中，管理服务器的计划交易模块根据调度信息生成的可用资源状态信息为 2022 年 4 月 1 日至 2022 年 4 月 15 日 0：00 至 6：00，虚拟机类型为 g4 的可用云资源容量为 5000 台虚拟机，2022 年 4 月 1 日至 2022 年 4 月 15 日 6：00 至 18：00，虚拟机类型为 g4 的可用云资源容量为 3000 台虚拟机。

404. 管理服务器的计划交易模块根据计划信息和可用资源状态信息生成云服务价格信息。

管理服务器的计划交易模块根据计划信息和可用资源状态信息生成云服务价格信息，该云服务价格信息包括计划信息所对应的云服务价格。进一步，当可用资源状态信息满足计划信息，且可用资源状态信息中可用云资源容量大于或等于第一阈值，则管理服务器生成云服务折扣信息。另外可用资源状态信息中可用云资源规格需要与计划信息中云资源的需求规格一致。

本申请实施例中由于用户设备向管理服务器提交了计划信息，因此管理服务器能够更好的计算可用资源状态信息，管理服务器可以向提交计划信息的用户设备设计更多的云服务价格折扣。

405. 管理服务器的计划交易模块向用户设备发送云服务价格信息。

管理服务器的计划交易模块向用户设备发送云服务价格信息，该云服务价格信息中包

括计划信息对应的云服务折扣信息。具体的，管理服务器可以通过 UI 界面返回云服务价格信息，可以通过应用程序接口 API 发送云服务价格信息，具体不做限定。

406.用户设备向管理服务器发送订购信息。

用户设备确认管理服务器的计划交易模块发送的云服务价格信息之后，向计划交易模块发送订购信息，该订购信息用于确认云服务价格信息。

407.管理服务器的计划交易模块向资源调度模块发送的订购信息对应的云资源计划。

管理服务器的计划交易模块向资源调度模块发送的订购信息对应的云资源计划，该云资源计划指示资源调度模块对云服务器集群执行云资源的分配任务。

408.管理服务器的资源调度模块保存云资源计划，预留所需的云资源。

管理服务器的资源调度模块保存云资源计划，预留云资源计划所需的云资源。

409.管理服务器的资源调度模块向用户设备发送通知消息。

管理服务器的资源调度模块向用户设备发送通知消息，该通知消息用于向用户设备通知云资源已经准备完成。

本申请实施例中管理服务器向用户设备发送通知消息的方式可以是消息服务、短信服务、邮件服务或基于超文本传输协议 HTTP 的请求调用等方式。通知消息的内容可以包括申请的虚拟机 ID、容器 ID 或者裸金属物理 ID，具体不做限定。

410.当到达计划时间时，管理服务器的资源调度模块通知云服务器集群分配对应的云资源。

当到达云资源计划的开始时间时，管理服务器的资源调度模块通知云服务器集群分配对应的云资源。例如，管理服务器通知云服务器集群启动虚拟机或容器。

本申请实施例中管理服务器接收用户设备提交的计划信息，并基于用户设备提交的计划信息和可用资源状态生成计划信息对应的云服务价格，由于用户设备预先提交了未来时段对云资源的需求容量和需求规格，因此，管理服务器能够基于计划信息提供更准确的云服务价格，提升了云资源的利效率。

请参阅图 5，图 5 为本申请实施例提供的另一种资源管理方法的流程示意图。图 5 所示的资源管理方法包括以下步骤：

501.管理服务器的计划交易模块从资源调度模块获取云服务器集群的调度信息。

502.管理服务器的计划交易模块根据云服务器集群的调度信息生成可用资源状态信息。

本申请实施实施例步骤 501 至步骤 502 中管理服务器所执行的操作与上述图 4a 所示的实施例中的步骤 402 至步骤 403 中管理服务器所执行的操作类似，具体不再赘述。

503.管理服务器的计划交易模块根据可用资源状态信息生成云服务价格信息。

管理服务器的计划交易模块根据可用资源状态信息生成云服务价格信息。具体的，管理服务器的计划交易模块可以根据未来不同时段的可用云资源规格和对应的可用云资源容量生成不同时段的云服务价格信息。

例如，在步骤 503 的一个示例中，管理服务器生成不同时段可用云资源容量以及对应

的云服务价格，例如，00:00 至 6:00 的 g5 型可用虚拟机数量为 2000 台，云服务价格为 10 元每台，06:00 至 18:00 的 g5 型可用虚拟机数量为 1000 台，云服务价格为 20 元每台。

504. 用户设备查询可用资源状态信息和对应的云服务价格信息。

用户设备查询可用资源状态信息和对应的云服务价格，用户设备可以通过 UI 界面或者 API 获取待销售的可用资源状态信息和对应的云服务价格。

本申请实施例中管理服务器可以将不同时段的云服务价格信息全部发送至用户设备，由用户设备选择云资源的需求时段、云资源需求容量和云资源的需求规格。可以理解的是，用户设备在不同时刻查询到的云服务价格信息会产生浮动。

505. 用户设备向管理服务器发送订购信息。

用户设备基于查询到的可用资源状态信息和对应的云服务价格确定订购信息，订购信息用于确认用户设备选择购买的可用资源状态信息以及对应的云服务价格信息。订购信息中包括用户设备确认订购的云资源的需求时段、云资源需求容量和云资源的需求规格。

506. 管理服务器的计划交易模块向资源调度模块提交订购信息对应的云资源。

507. 管理服务器的资源调度模块保存云资源计划，预留所需的云资源。

508. 管理服务器的资源调度模块向用户设备发送通知消息。

509. 当到达计划时间时，管理服务器的资源调度模块通知云服务器集群分配对应的云资源。

本申请实施实施例步骤 506 至步骤 509 中管理服务器所执行的操作与上述图 4a 所示的实施例中的步骤 407 至步骤 410 中管理服务器所执行的操作类似，具体不再赘述。

本申请实施例中管理服务器可以基于用户设备发送的查询请求将未来时段的可用资源状态信息和相应的云服务价格信息发送至用户设备，由用户设备自己选择要订购的云资源需求时段、云资源规格以及云资源容量，从而提升了云资源的利用率。

以上介绍本申请实施例中提供的资源管理方法，下面结合附图介绍本申请实施例提供的相关装置。

请参阅图 6，图 6 为本申请实施例提供的一种资源管理装置的结构示意图。该装置用于实现上述各实施例中管理服务器所执行的各个步骤，如图 6 所示，该资源管理装置 600 包括收发单元 601 和处理单元 602。

其中，收发单元 601 用于获取云服务器集群的可用资源状态信息，可用资源状态信息包括以下一项或多项：云服务器集群在未来不同时段的可用云资源容量和未来不同时段的可用云资源规格。处理单元 602 用于根据可用资源状态信息生成云服务价格信息，云服务价格信息包括多个规格的云资源的价格信息。收发单元 601 还用于向用户设备发送云服务价格信息。

一种可能的实施方式中，收发单元 601 还用于接收用户设备发送的计划信息，计划信息用于向管理服务器提供用户设备对云资源的需求计划，计划信息包括以下一项或多项信息：云资源的需求时段或云资源的需求规格。处理单元 602 具体用于根据计划信息和可用资源状态信息生成云服务价格信息。

一种可能的实施方式中，可用资源状态信息中的可用云资源规格与计划信息中的需求规格一致。

一种可能的实施方式中，处理单元 602 具体用于当可用资源状态信息满足计划信息，且可用资源状态信息中可用云资源容量大于或等于目标阈值时，生成云服务折扣信息。

一种可能的实施方式中，收发单元 601 还用于接收用户设备发送的订购信息，订购信息用于确认计划信息对应的云服务价格信息。

一种可能的实施方式中，收发单元 601 还用于接收用户设备发送的查询请求。处理单元 602 还用于根据查询请求向用户设备发送可用资源状态信息和云服务价格信息。

一种可能的实施方式中，收发单元 601 还用于接收用户设备发送的订购信息，订购信息包括用户设备选择的至少一个规格的云资源和至少一个规格的云资源的价格。

一种可能的实施方式中，收发单元 601 还用于获取云服务器集群的调度信息，调度信息包括以下一项或多项信息：已提交的云资源计划、基于历史云资源变化预测的云资源计划、需预留的空闲云资源计划、总的云资源计划、单设备的最大云资源计划和使用中云资源的延续计划。处理单元 602 还用于根据调度信息生成云服务器集群的可用资源状态信息。

应理解以上装置中单元的划分仅仅是一种逻辑功能的划分，实际实现时可以全部或部分集成到一个物理实体上，也可以物理上分开。且装置中的单元可以全部以软件通过处理元件调用的形式实现；也可以全部以硬件的形式实现；还可以部分单元以软件通过处理元件调用的形式实现，部分单元以硬件的形式实现。例如，各个单元可以为单独设立的处理元件，也可以集成在装置的某一个芯片中实现，此外，也可以以程序的形式存储于存储器中，由装置的某一个处理元件调用并执行该单元的功能。此外这些单元全部或部分可以集成在一起，也可以独立实现。这里所述的处理元件又可以成为处理器，可以是一种具有信号的处理能力的集成电路。在实现过程中，上述方法的各步骤或以上各个单元可以通过处理器元件中的硬件的集成逻辑电路实现或者以软件通过处理元件调用的形式实现。

值得说明的是，对于上述方法实施例，为了简单描述，故将其都表述为一系列的动作组合，但是本领域技术人员应该知悉，本申请本申请并不受所描述的动作顺序的限制，其次，本领域技术人员也应该知悉，说明书中所描述的实施例均属于优选实施例，所涉及的动作并不一定是本申请本申请所必须的。

本领域的技术人员根据以上描述的内容，能够想到的其他合理的步骤组合，也属于本申请本申请的保护范围内。其次，本领域技术人员也应该熟悉，说明书中所描述的实施例均属于优选实施例，所涉及的动作并不一定是本申请本申请所必须的。

请参阅图 7，图 7 为本申请实施例提供的一种管理服务器的结构示意图。如图 7 所示，该管理服务器 700 包括：处理器 710、存储器 720 和接口 730，处理器 710、存储器 720 与接口 730 通过总线（图中未标注）耦合。存储器 720 存储有指令，当存储器 720 中的执行指令被执行时，管理服务器 700 执行上述方法实施例中管理服务器所执行的方法。

管理服务器 700 可以是被配置成实施以上方法的一个或多个集成电路，例如：一个或多个特定集成电路（application specific integrated circuit，ASIC），或，一个或多个微处理器(digital singnal processor,DSP),或,一个或者多个现场可编程门阵列(field

programmable gate array，FPGA），或这些集成电路形式中至少两种的组合。再如，当装置中的单元可以通过处理元件调度程序的形式实现时，该处理元件可以是通用处理器，例如中央处理器（central processing unit，CPU）或其它可以调用程序的处理器。再如，这些单元可以集成在一起，以片上系统（system-on-a-chip，SOC）的形式实现。

处理器 710 可以是中央处理单元（central processing unit，CPU），还可以是其它通用处理器、数字信号处理器（digital signal processor，DSP）、专用集成电路（application specific integrated circuit，ASIC）、现场可编程门阵列（field programmable gate array，FPGA）或者其它可编程逻辑器件、晶体管逻辑器件，硬件部件或者其任意组合。通用处理器可以是微处理器，也可以是任何常规的处理器。

存储器 720 可以是易失性存储器或非易失性存储器，或可包括易失性和非易失性存储器两者。其中，非易失性存储器可以是只读存储器（read-only memory，ROM）、可编程只读存储器（programmable ROM，PROM）、可擦除可编程只读存储器（erasable PROM，EPROM）、电可擦除可编程只读存储器（electrically EPROM，EEPROM）或闪存。易失性存储器可以是随机存取存储器（random access memory，RAM），其用作外部高速缓存。通过示例性但不是限制性说明，许多形式的 RAM 可用，例如静态随机存取存储器（static RAM，SRAM）、动态随机存取存储器（DRAM）、同步动态随机存取存储器（synchronous DRAM，SDRAM）、双倍数据速率同步动态随机存取存储器（double data date SDRAM，DDR SDRAM）、增强型同步动态随机存取存储器（enhanced SDRAM，ESDRAM）、同步连接动态随机存取存储器（synchlink DRAM，SLDRAM）和直接内存总线随机存取存储器（direct rambus RAM，DR RAM）。

总线除包括数据总线之外，还可以包括电源总线、控制总线和状态信号总线等。总线可以是快捷外围部件互连标准（peripheral component interconnect express，PCIe）总线，或扩展工业标准结构（extended industry standard architecture，EISA）总线、统一总线（unified bus，Ubus 或 UB）、计算机快速链接（compute express link，CXL）、缓存一致互联协议（cache coherent interconnect for accelerators，CCIX）等。总线可以分为地址总线、数据总线、控制总线等。

可以理解的是，本申请实施例中图 6 所示的资源管理装置 600 和图 7 所示的管理服务器 700 可以是上述图 1 所示的系统架构中的管理服务器 101。

在本申请的另一个实施例中，还提供一种计算机可读存储介质，计算机可读存储介质中存储有计算机执行指令，当设备的处理器执行该计算机执行指令时，设备执行上述方法实施例中管理服务器所执行的方法。

在本申请的另一个实施例中，还提供一种计算机程序产品，该计算机程序产品包括计算机执行指令，该计算机执行指令存储在计算机可读存储介质中。当设备的处理器执行该计算机执行指令时，设备执行上述方法实施例中管理服务器所执行的方法。

所属领域的技术人员可以清楚地了解到，为描述的方便和简洁，上述描述的系统，装置和单元的具体工作过程，可以参考前述方法实施例中的对应过程，在此不再赘述。

在本申请所提供的几个实施例中，应该理解到，所揭露的系统，装置和方法，可以通过其它的方式实现。例如，以上所描述的装置实施例仅仅是示意性的，例如，所述单元的

划分，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式，例如多个单元或组件可以结合或者可以集成到另一个系统，或一些特征可以忽略，或不执行。另一点，所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口，装置或单元的间接耦合或通信连接，可以是电性，机械或其它的形式。

所述作为分离部件说明的单元可以是或者也可以不是物理上分开的，作为单元显示的部件可以是或者也可以不是物理单元，即可以位于一个地方，或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

另外，在本申请各个实施例中的各功能单元可以集成在一个处理单元中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现，也可以采用软件功能单元的形式实现。

所述集成的单元如果以软件功能单元的形式实现并作为独立的产品销售或使用时，可以存储在一个计算机可读取存储介质中。基于这样的理解，本申请的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的全部或部分可以以软件产品的形式体现出来，该计算机软件产品存储在一个存储介质中，包括若干指令用以使得一台计算机设备（可以是个人计算机，服务器，或者网络设备等）执行本申请各个实施例所述方法的全部或部分步骤。而前述的存储介质包括：U 盘、移动硬盘、只读存储器（ROM，read-only memory）、随机存取存储器（RAM，random access memory）、磁碟或者光盘等各种可以存储程序代码的介质。

权 利 要 求

1. 一种资源管理方法，其特征在于，包括：

获取云服务器集群的可用资源状态信息，所述可用资源状态信息包括以下一项或多项：所述云服务器集群在未来不同时段的可用云资源容量和未来不同时段的可用云资源规格；

根据所述可用资源状态信息生成云服务价格信息，所述云服务价格信息包括多个规格的云资源的价格信息；

向用户设备发送所述云服务价格信息。

2. 根据权利要求 1 所述的方法，其特征在于，所述获取云服务器集群的可用资源状态信息之前，所述方法还包括：

接收用户设备发送的计划信息，所述计划信息用于向管理服务器提供所述用户设备对云资源的需求计划，所述计划信息包括以下一项或多项信息：云资源的需求时段或云资源的需求规格；

所述根据所述可用资源状态信息生成云服务价格信息包括：

根据所述计划信息和所述可用资源状态信息生成云服务价格信息。

3. 根据权利要求 2 所述的方法，其特征在于，所述可用资源状态信息中的可用云资源规格与所述计划信息中的需求规格一致。

4. 根据权利要求 2 或 3 所述的方法，其特征在于，所述根据所述计划信息和所述可用资源状态信息生成云服务价格信息包括：

当所述可用资源状态信息满足所述计划信息，且所述可用资源状态信息中可用云资源容量大于或等于目标阈值时，生成云服务折扣信息。

5. 根据权利要求 1 至 4 中任一项所述的方法，其特征在于，所述向用户设备发送所述云服务价格信息之后，所述方法还包括：

接收所述用户设备发送的订购信息，所述订购信息用于确认所述计划信息对应的云服务价格信息。

6. 根据权利要求 1 所述的方法，其特征在于，所述向用户设备发送所述云服务价格信息前，所述方法还包括：

接收所述用户设备发送的查询请求；

根据所述查询请求向所述用户设备发送所述可用资源状态信息和所述云服务价格信息。

7. 根据权利要求 6 所述的方法，其特征在于，所述根据所述查询请求向所述用户设备发送所述可用资源状态信息和所述云服务价格信息之后，所述方法还包括：

接收所述用户设备发送的订购信息，所述订购信息包括所述用户设备选择的至少一个规格的云资源和所述至少一个规格的云资源的价格。

8. 根据权利要求 1 至 7 中任一项所述的方法，其特征在于，所述获取云服务器集群的可用资源状态信息包括：

获取所述云服务器集群的调度信息，所述调度信息包括以下一项或多项信息：已提交的云资源计划、基于历史云资源变化预测的云资源计划、需预留的空闲云资源计划、总的

云资源计划、单设备的最大云资源计划和使用中云资源的延续计划；

根据所述调度信息生成所述云服务器的可用资源状态信息。

9. 一种资源管理装置，其特征在于，包括：

收发单元，用于获取云服务器集群的可用资源状态信息，所述可用资源状态信息包括以下一项或多项：所述云服务器集群在未来不同时段的可用云资源容量和未来不同时段的可用云资源规格；

处理单元，用于根据所述可用资源状态信息生成云服务价格信息，所述云服务价格信息包括多个规格的云资源的价格信息；

所述收发单元还用于向用户设备发送所述云服务价格信息。

10. 根据权利要求 9 所述的装置，其特征在于，所述收发单元还用于接收用户设备发送的计划信息，所述计划信息用于向管理服务器提供所述用户设备对云资源的需求计划，所述计划信息包括以下一项或多项信息：云资源的需求时段或云资源的需求规格；

所述处理单元具体用于根据所述计划信息和所述可用资源状态信息生成云服务价格信息。

11. 根据权利要求 10 所述的装置，其特征在于，所述可用资源状态信息中的可用云资源规格与所述计划信息中的需求规格一致。

12. 根据权利要求 10 或 11 所述的装置，其特征在于，所述处理单元具体用于：

当所述可用资源状态信息满足所述计划信息，且所述可用资源状态信息中可用云资源容量大于或等于目标阈值时，生成云服务折扣信息。

13. 根据权利要求 9 至 12 中任一项所述的装置，其特征在于，所述收发单元还用于：

接收所述用户设备发送的订购信息，所述订购信息用于确认所述计划信息对应的云服务价格信息。

14. 根据权利要求 9 所述的装置，其特征在于，所述收发单元还用于接收所述用户设备发送的查询请求；

所述处理单元还用于根据所述查询请求向所述用户设备发送所述可用资源状态信息和所述云服务价格信息。

15. 根据权利要求 14 所述的装置，其特征在于，所述收发单元还用于：

接收所述用户设备发送的订购信息，所述订购信息包括所述用户设备选择的至少一个规格的云资源和所述至少一个规格的云资源的价格。

16. 根据权利要求 9 至 15 中任一项所述的装置，其特征在于，所述收发单元还用于：

获取所述云服务器集群的调度信息，所述调度信息包括以下一项或多项信息：已提交的云资源计划、基于历史云资源变化预测的云资源计划、需预留的空闲云资源计划、总的云资源计划、单设备的最大云资源计划和使用中云资源的延续计划；

所述处理单元还用于根据所述调度信息生成所述云服务器集群的可用资源状态信息。

17. 一种管理服务器，其特征在于，包括处理器，所述处理器与存储器耦合，所述处理器用于存储指令，当所述指令被所述处理器执行时，以使得所述管理服务器执行权利要求 1 至 8 中任一项所述的方法。

18. 一种计算机可读存储介质，其上存储有指令，其特征在于，所述指令被执行时，以使得计算机执行权利要求 1 至 8 中任一项所述的方法。

19. 一种计算机程序产品，所述计算机程序产品中包括指令，其特征在于，所述指令被执行时，以使得计算机实现权利要求 1 至 8 中任一项所述的方法。