基于抢占式调度的资源共享使用方法、系统及设备

本申请要求在 2018 年 6 月 25 日提交中国专利局、申请号为 201810659298.5 的中国专利申请的优先权，该申请的全部内容通过引用结合在本申请中。

技术领域

本申请涉及计算机技术领域，例如涉及一种基于抢占式调度的资源共享使用方法系统及设备。

背景技术

在资源共享的分布式系统中，多个租户共享使用资源；同时每个租户使用的资源又需要有一定的限制，以保证每个租户都能够有资源使用，不出现租户资源“饿死”的现象。分布式系统中的调度器是通过有效地调度多个租户的作业或任务，从而保证每个租户的作业或任务得到稳定而又快速地执行，同时分布式系统内的资源得到充分利用。

传统的分布式管理系统提供了多个调度策略，以保证多个租户的任务可以在分布式系统中均衡分配到物理节点，再由物理节点运行分配的任务。但是任务的处理方式存在资源没有充分利用的现象。

发明内容

本申请实施例提供一种基于抢占式调度的资源共享使用方法系统及设备，可以提高资源的利用率。

第一方面，本申请实施例提供了一种任务创建方法，包括：API 服务器获取任务的创建请求；以及当 API 服务器检测到所述任务所属租户的配额里包含与所述任务的优先级匹配的资源，且匹配资源满足所述任务的创建条件，根据所述创建请求创建所述任务。

第二方面，本申请实施例还提供了一种任务调度方法，包括：调度器从任务调度队列获取待调度的当前任务，并获取每个物理节点上大于等于所述当前任务优先级的任务，形成节点-任务映射表；所述调度器根据所述映射表以及预设筛选条件确定符合所述预设筛选条件的目标物理节点；以及所述调度器将所述当前任务与所述目标物理节点绑定，并将绑定的信息发送到 API 服务器。

第三方面，本申请实施例还提供了一种任务抢占方法，包括：当物理节点处理待运行的目标任务时，获取所述物理节点上正在运行的任务列表；所述物理节点检测其剩余资源是否满足所述目标任务运行所需的资源；若剩余资源不满足目标任务运行所需的资源，所述物理节点将所述任务列表中的优先级低于所述目标任务优先级的任务，按照优先级从低到高的顺序依次移入待移除队列，直至所述物理节点执行所述任务列表中的任务后所得到的剩余资源满足所述目标任务运行所需的资源，并采用所述目标任务抢占所述待移除队列中的任务；

以及所述物理节点调用执行环境，运行所述目标任务。

第四方面，本申请实施例还提供了一种基于抢占式调度的资源共享使用方法，包括：API 服务器获取任务的创建请求；当所述 API 服务器检测到所述任务所属租户的配额里包含与所述任务的优先级匹配的资源，且匹配资源满足所述任务的创建条件，根据所述创建请求创建所述任务；调度器获取所述 API 服务器创建的任务，并形成任务调度队列；所述调度器从所述任务调度队列获取待调度的当前任务，并获取每个物理节点上大于等于所述当前任务优先级的任务，形成节点-任务映射表；所述调度器根据所述映射表以及预设筛选条件确定符合所述预设筛选条件的目标物理节点；所述调度器将所述当前任务与所述目标物理节点绑定，并将绑定的信息发送到 API 服务器；物理节点监听所述 API 服务器中任务与物理节点的绑定信息，基于监听到的所述绑定信息获取对应的任务，并形成任务队列；当所述物理节点处理所述任务队列中待运行的目标任务时，获取所述物理节点上正在运行的任务列表；所述物理节点检测其剩余资源是否满足所述目标任务运行所需的资源；若剩余资源不满足目标任务运行所需的资源，所述物理节点将所述任务列表中的优先级低于所述目标任务优先级的任务，按照优先级从低到高的顺序依次移入待移除队列，直至所述物理节点执行所述任务列表中的任务后所得到的剩余资源满足所述目标任务运行所需的资源，并采用目标任务抢占所述待移除队列中的任务；以及所述物理节点调用执行环境，运行所述目标任务。

第五方面，本申请实施例还提供了一种 API 服务器，包括：请求获取模块，设置为获取任务的创建请求；任务创建模块，设置为当检测到所述任务所属租户的配额里包含与所述任务的优先级匹配的资源，且匹配资源满足所述任务的创建条件，根据所述创建请求创建所述任务。

第六方面，本申请实施例还提供了一种调度器，包括：映射表形成模块，设置为从任务调度队列中获取待调度的当前任务，并获取每个物理节点上大于等于所述当前任务优先级的任务，形成节点-任务映射表；筛选模块，设置为根据所述映射表以及预设筛选条件确定符合所述预设筛选条件的目标物理节点；绑定模块，设置为将所述当前任务与所述目标物理节点绑定，并将绑定的信息发送到 API 服务器。

第七方面，本申请实施例还提供了一种任务抢占装置，包括：任务列表获取模块，设置为当处理待运行的目标任务时，获取所述物理节点上正在运行的任务列表；检测模块，设置为检测物理节点上的剩余资源是否满足所述目标任务运行所需的资源；抢占模块，设置为当剩余资源不满足目标任务运行所需的资源时，将所述任务列表中的优先级低于所述目标任务优先级的任务，按照优先级从低到高的顺序依次移入待移除队列，直至所述物理节点执行所述任务列表中的任务所得到的剩余资源满足所述目标任务运行所需的资源，并采用所述目标任务抢占所述待移除队列中的任务；任务执行模块，设置为调用执行环境，运行所述目标任务。

第八方面，本申请实施例还提供了一种基于抢占式调度的资源共享使用系

统，包括本申请实施例提供的 API 服务器、调度器以及任务抢占装置。

第九方面，本申请实施例提供了一种设备，包括：一个或多个处理器；存储装置，用于存储一个或多个程序。

当所述一个或多个程序被所述一个或多个处理器执行，使得所述一个或多个处理器实现本申请实施例提供的任务创建方法，或者实现本申请实施例提供的任务调度方法，或者实现本申请实施例提供的任务抢占方法。

第十方面，本申请实施例提供了一种计算机可读存储介质，其上存储有计算机程序，该程序被处理器执行时实现本申请实施例提供的任务创建方法，或者实现本申请实施例提供的任务调度方法，或者实现本申请实施例提供的任务抢占方法。

本申请实施例提供的技术方案，通过对租户的资源配额进行优先级的设置，并通过将任务的优先级与所属租户下每个优先级的资源进行匹配，从而确定是否创建任务，可以让租户在资源紧张时优先使用资源，防止租户滥用高优先级的资源，导致低优先级的任务持续获取不了资源出现的“饿死”现象。调度器通过任务的优先级以及预设筛选条件对物理节点进行筛选，筛选出最合适的物理节点，将待调度的当前任务调度到最合适的物理节点，当资源紧张时，只进行逻辑上的资源抢占，并没有立即抢占资源，这种延后抢占的调度方法，可以在逻辑上为高优先级的任务腾出资源，在资源没有被充分利用时，继续运行被抢占的任务，可以提高资源的利用率。当物理节点处理待运行的目标任务时，若物理节点的剩余资源不满足目标任务运行的所需的条件，基于将低优先级的任务进行抢占，可以使物理节点优先处理重要任务，可以提高资源的利用率。

附图说明

图 1 是本申请实施例提供的一种任务创建方法流程图。

图 2 是本申请实施例提供的一种任务调度方法流程图。

图 3 是本申请实施例提供的一种任务调度方法流程图。

图 4 是本申请实施例提供的一种任务抢占方法流程图。

图 5 是本申请实施例提供的一种任务抢占方法流程图。

图 6 是本申请实施例提供的一种基于抢占式调度的资源共享使用方法流程图。

图 7 是本申请实施例提供的一种 API 服务器的结构框图。

图 8a 是本申请实施例提供的一种调度器的结构框图。

图 8b 是本申请实施例提供的一种调度系统的结构示意图。

图 9 是本申请实施例提供的一种任务抢占装置的结构框图。

图 10 是本申请实施例提供的一种基于抢占式调度的资源共享使用系统结构框图。

图 11 是本申请实施例提供的一种设备的结构示意图。

具体实施方式

下面结合附图和实施例对本申请作说明。可以理解的是，此处所描述的具体实施例仅仅用于解释本申请，而非对本申请的限定。另外还需要说明的是，为了便于描述，附图中仅示出了与本申请相关的部分而非全部内容。

图 1 是本申请实施例提供的一种任务创建方法流程图，所述方法可应用于应用编程接口（Application Programming Interface，API）服务器（API Server），由 API 服务器来执行，该 API 服务器可以是集群管理平台中的一个组件，并采用软件和/硬件的方式来实现，集群管理平台可以是对集群中的大量的资源进行管理的平台，集群管理平台包括但不限于 Kubernetes 和 Mesos，并可集成在多个计算机设备中。如图 1 所示，本申请实施例提供的任务创建方法包括 S110-S120。

在 S110 中，API 服务器获取任务的创建请求。

本申请实施例提供的方法可以应用在集群中，在集群中可以包含多个物理节点，每个物理节点上的资源可以是多个租户的共享资源，多个物理节点可以由集群管理平台进行管理，集群管理平台将任务分配给物理节点以使物理节点执行相应的任务。集群管理平台可以集成在多个计算机设备中，多个计算机设备可以由用户进行操作，用户可以登录集群管理平台，并递交任务的创建请求，集群管理平台中的 API 服务器获取用户提交的任务的创建请求，创建任务。由集群管理平台中的调度器进行任务调度，将任务合理分配给对应的物理节点。由物理节点执行该任务。

其中，API 服务器可以为集群管理平台中的一个组件，可以完成任务的创建；提供丰富的功能性插件，完善对集群的管理等。

其中，API 服务器可以获取用户提交的一个任务的创建请求，例如，可以是获取用户提交的创建一个应用的请求。其中，当用户提交任务的创建请求时，可以对任务的优先级进行设置。

在 S120 中，当 API 服务器检测到所述任务所属租户的配额里包含与所述任务的优先级匹配的资源，且匹配资源满足所述任务的创建条件，根据所述创建请求创建所述任务。

用户在管理粒度上被分到若干组内，每组称为一个租户，可以根据需要为每个租户预先设置配额，配额可以是一组资源，例如，该组资源可以包括处理器（Central Processing Unit，CPU）、内存、图形处理器（Graphics Processing Unit，GPU）、现场可编程门阵列（Field Programmable Gate Array，FPGA）、人工智能（Artificial Intelligence，AI）芯片、处理器的优先级以及内存的优先级、GPU 优先级、FPGA 优先级、AI 芯片优先级等，可以为每个租户预先设置资源配额。通过合理的设置配额，可以授权给租户使用合适优先级的资源，能够让租户在资源紧张的时候优先使用资源，还可以限制租户滥用高优先级的资源，导致低优先级的任务持续获取不了资源，出现“饿死”的现象。

当用户通过集群管理平台提交任务的创建请求时，每个任务携带标识信息，

API 服务器可以根据标识信息识别每个任务所属的租户，判断该租户的配额里是否包含与该任务的优先级匹配的资源，如果包含于该任务的优先级匹配的资源，继续判断该匹配资源是否满足该任务的创建条件，当该匹配资源满足该任务的创建条件时，则创建任务。

创建条件可以是 CPU 的数量和/或内存的占有率，还可以是其他条件。例如，当用户提交任务的创建请求时，可以通过集群管理平台设置该任务的优先级为高优先级，那么该任务所需资源的优先级也为高优先级，如：该任务所需 10 个高优先级的 CPU。若该任务所属的租户的配额中包含高优先级的 CPU，且高优先级的 CPU 的数量大于或等于 10，则创建该任务。

在本申请的一个实施方式中，可选的，本申请实施例提供的方法可以应用在 Kubernetes 环境下。所述服务器获取任务创建请求包括：API 服务器获取任务 pod 的创建请求。当 API 服务器检测到所述任务所属租户的配额里包含与所述任务的优先级匹配的资源，且匹配资源满足所述任务的创建条件，根据所述创建请求创建所述任务包括：当所述 API 服务器检测到所述任务 pod 所属的租户 namespace 的配额 quota 里包含与所述 pod 对应的优先级匹配的资源 quota 值，且该 quota 值满足所述任务 pod 的创建条件，则 API 服务器根据所述创建请求创建 pod。

其中，pod 是 kubernetes 中可以创建和部署的最小，也是最简的单位。一个 pod 代表着集群中运行的一个进程。pod 是 Kubernetes 中的组件，例如，可以创建一个应用，可以启动一个进程等。pod 中封装着应用的容器（有的情况下是好几个容器），该容器可以存储独立的网络 IP、管理容器如何运行的策略选项等。其中，pod 代表着部署的一个单位：kubernetes 中应用的一个实例，可能由一个或者多个容器组合在一起共享资源。其中，创建一个 pod 可以是创建一个应用等。

其中，Namespace 是对一组资源和对象的抽象集合，比如可以用来将 kubernetes 系统内部的对象划分为不同的项目组或用户组，namespace 常用来隔离不同的租户或者用户。在 kubernetes 中，quota 可以用来进行资源管理和资源限制，quota 值的大小可以代表资源的多少，例如，在一个租户下设置的资源是 20 个高优先级的 CPU，namespace 中的 quota 值可以是 20，即 quota 值可以代表资源的数量。

本申请实施例提供的一种任务创建方法，当获取到任务的创建请求时，通过检测任务所属租户的配额里是否包含与任务优先级匹配的资源，以及检测与任务优先级匹配的资源是否满足任务的创建条件，当两个条件均满足（所属租户的配额中既包含于该任务优先级匹配的资源，且该资源也满足该任务的创建条件）时创建该任务，本实施例通过对租户的资源配额进行优先级的设置，并通过将任务的优先级与所属租户下各优先级的资源进行匹配，从而确定是否创建任务，可以让租户在资源紧张时优先使用资源，防止租户滥用高优先级的资源，导致低优先级的任务持续获取不了资源出现的“饿死”现象。

图 2 是本申请实施例提供的一种任务调度方法流程图，所述方法可应用于调度器，调度器可以是集群管理平台的组件，并采用软件和/硬件的方式来实现，集群管理平台可以是对集群中的大量的硬件资源进行管理的平台，集群管理平台包括但不限于 Kubernetes 和 Mesos，并可集成在多个计算机设备中。本申请实施例提供的方法可以应用在该环境下：在集群中可以包含多个物理节点，多个物理节点上的资源可以是多个租户的共享资源，多个物理节点可以由集群管理平台进行管理，集群管理平台将任务分配给物理节点以使物理节点执行相应的任务。集群管理平台可以集成在多个计算机设备中，多个计算机设备可以由用户进行操作，用户可以登录集群管理平台，并递交任务的创建请求，集群管理平台中的 API 服务器获取用户提交的任务的创建请求，创建任务，由集群管理平台中的调度器进行任务调度，将任务合理分配给对应的物理节点，由物理节点执行该任务。其中，本申请实施例应用在调度器进行任务调度的阶段。

如图 2 所示，本申请实施例提供的技术方案包括 S210-S230。

在 S210 中，调度器从任务调度队列获取待调度的当前任务，并获取每个物理节点上大于等于所述当前任务的优先级的任务，形成节点-任务映射表。

其中，调度器可以是集群管理平台的组件，可以从 API 服务中监听 API 服务器中创建的任务，并从 API 服务器中读取任务，读取的任务形成任务调度队列。其中，调度器按照任务调度队列中任务的顺序，对任务进行调度。

物理节点可以是各种物理机，调度器可以从每个物理节点获取资源信息（包括全部资源和可用资源）和在每个物理节点上正在运行的任务队列。其中，该任务队列中的每个任务均有优先级。当调度器从任务调度队列获取待调度的当前任务时，获取每个物理节点上大于等于当前任务的优先级的任务，并形成节点-任务映射表。例如，当前任务的优先级为高优先级，物理节点 1 上大于或等于高优先级的任务有任务 1、任务 2 和任务 3，则调度器获取物理节点 1 上的任务 1、任务 2 和任务 3、并形成节点-任务映射表。其中，物理节点上大于等于当前任务的优先级的任务包括：物理节点上正在运行的且大于等于当前任务优先级的任务，以及物理节点上待运行的且大于等于当前任务优先级的任务。

在 S220 中，所述调度器根据所述映射表以及预设筛选条件确定最符合所述预设筛选条件的目标物理节点。

调度器从节点-任务映射表中根据预设筛选条件进行筛选，筛选出最符合预设筛选条件的目标物理节点。其中，预设筛选条件可以包括当前任务所需的资源与物理节点上的剩余资源之间的匹配条件、当前任务所需的端口与物理节点上的端口之间的匹配条件等。

在本申请一个实施方式中，可选的，所述调度器根据所述映射表以及预设筛选条件确定最符合所述预设筛选条件的目标物理节点包括：所述调度器从映射表筛选出符合第一阶段筛选条件的物理节点，形成节点组；根据映射表以及第二阶段优选条件，对所述节点组的物理节点进行评分，并筛选出分数最高的物理节点作为目标物理节点。

其中，第一阶段筛选条件和第二阶段优选条件并不相同。例如，第一阶段筛选条件可以是当前任务所需的端口与物理节点上的端口之间的匹配条件、物理节点是否有特殊标签等，第二阶段优选条件可以是当前任务所需的资源与物理节点上的剩余资源之间的匹配条件，并且第二阶段优选条件中可以包括一个条件，也可以包含多个条件。当第二阶段优选条件中包含多个条件时，也可以为每个条件设置权重，根据权重确定物理节点的评分。

对该实施方式进行举例说明，若第一阶段筛选条件为物理节点需要有 GPU 的标签，第二预设当前任务所需的资源与物理节点上的剩余资源之间的匹配条件。对照节点-任务映射表以及获取的物理节点的信息，调度器选取具有 GPU 的物理节点，形成节点组。调度器判断节点组中物理节点上的剩余资源是否满足当前任务所需的资源条件，将不满足条件的物理节点去除，并对满足当前任务所需的资源条件的物理节点进行打分，物理节点上的剩余资源越多，对该物理节点的打分越高，则分数最高的物理节点是目标物理节点。其中，筛选目标物理节点的方法包括但并不限于上述的方法。

通过上述两次筛选，筛选出分数最高的物理节点作为目标物理节点，即筛选出分数最高的物理节点作为最合适的物理节点，相对于一次筛选的情况，可以减少每次筛选时数据的处理量，提高任务调度的效率。

在 S230 中，所述调度器将所述当前任务与所述目标物理节点绑定，并将绑定的信息发送到 API 服务器。

调度器通过将当前任务与筛选出的最符合预设筛选条件的物理节点（即目标物理节点）进行绑定，将绑定的信息发送到 API 服务器，以使每个物理节点可以从 API 服务器读取各自执行的任务。

本申请实施例调度器基于任务的优先级以及预设筛选条件对多个物理节点进行筛选，筛选出每个任务对应的最合适的物理节点（即目标物理节点），将待调度的当前任务调度到最合适的物理节点。当资源紧张时，只进行逻辑上的资源抢占，并没有立即抢占资源，这种延后抢占的调度方法，可以在逻辑上为高优先级的任务让出资源，在资源没有被充分利用时，继续保留被抢占的任务，可以提高资源的利用率。

图 3 是本申请实施例提供的一种任务调度方法流程图，其中，该实施例提供的方法可以应用于 kubernetes 系统中，如图 3 所示，本实施例提供的技术方案包括 S310-S340。

在 S310 中，调度器从 pod 调度队列中获取待调度的当前 pod，并获取每个物理节点上大于等于所述当前 pod 的优先级的 pod，形成节点-pod 映射表。

在 S320 中，所述调度器从映射表筛选出符合第一阶段筛选条件的物理节点，形成节点组。

在 S330 中，根据映射表以及第二阶段优选条件，对所述节点组的物理节点进行评分，并筛选出分数最高的物理节点作为目标物理节点。

在 S340 中，所述调度器将所述当前 pod 与所述目标物理节点绑定，并将

绑定的信息发送到 API 服务器。

由此，调度器基于任务的优先级以及预设筛选条件对物理节点进行筛选，筛选出最合适的物理节点，将待调度的当前任务调度到最合适的物理节点。当资源紧张时，只进行逻辑上的资源抢占，并没有立即抢占资源，这种延后抢占的调度方法，可以在逻辑上为高优先级的任务让出资源，在资源没有被充分利用时，继续保留被抢占的任务，可以提高资源的利用率。

图 4 是本申请实施例提供的一种任务抢占方法流程图，所述方法可由任务抢占装置来执行，所述装置由软件和/或硬件来实现，所述装置可集成在计算机设备中。本申请实施例提供的任务抢占方法适用于物理节点处理任务的场景下。如图 4 所示，本申请实施例提供的技术方案包括：S410-S440。

在 S410 中，当物理节点处理待运行的目标任务时，获取所述物理节点上正在运行的任务列表。

其中，物理节点可以是计算机设备，例如，物理机等。物理节点可以通过监听 API 服务器中任务与物理节点的绑定信息，获取对应的任务，由获取的任务形成任务队列。物理节点根据任务队列中的任务的顺序依次进行处理，当前处理的任务称为待运行的目标任务。当物理节点处理待运行的目标任务时，物理节点获取该物理节点上正在运行的任务列表。其中，任务列表记载了物理节点上正在运行的任务的信息。物理节点上正在运行的任务可以有一个，也可以有多个。

在 S420 中，所述物理节点检测其剩余资源是否满足所述目标任务运行所需的资源。

若物理节点检测其剩余资源不满足目标任务运行所需的资源，执行 S430，若物理节点检测其剩余资源满足目标任务运行所需的资源，执行 S440。

在本步骤中，目标任务运行所需的资源可以包括 CPU、内存等。物理节点的剩余资源可以理解为物理节点上的可用资源。例如，若物理节点剩余 CPU 的数量为 10 个，内存为 1G。目标任务运行所需的 CPU 为 10，内存为 2G，则物理节点上的剩余资源不能满足目标任务运行所需的资源。

在 S430 中，所述物理节点将所述任务列表中的优先级低于所述目标任务的优先级的任务按照优先级从低到高的顺序依次移入待移除队列（该物理节点不执行被移入待移除队列中的任务），直至所述物理节点执行所述任务列表中的任务后所得到的剩余资源满足所述目标任务运行所需的资源，并采用所述目标任务抢占所述待移除队列中的任务。

在本步骤中，若物理节点检测其剩余资源可以满足目标任务运行所需的资源，直接调用执行环境，运行目标任务。若物理节点检测其剩余资源并不能满足目标任务运行所需的资源，将任务列表中的任务按照优先级由低到高的顺序进行排序，并将优先级低于目标任务的优先级的任务按照优先级从低到高的顺序移入到待移除队列，直至物理节点执行任务列表中的任务后所得到的剩余资源满足目标任务运行所需的资源，采用目标任务抢占待移除队列中的任务的资

源，即停止运行待移除队列中的任务。

在将任务列表中的任务移入待移除队列过程中，若物理节点上的剩余资源不满足目标任务的运行条件，则拒绝执行目标任务。

需要说明的是，当将任务列表中优先级低于目标任务优先级的任务移入待移除队列的过程中，并不停止执行待移除队列中的任务，当判断物理节点执行任务列表中的任务后所得到的剩余资源满足目标任务运行所需的资源时，才停止执行待移除队列中的任务。

对本步骤进行举例说明，若任务列表中共有 5 个任务（物理节点上正在运行的任务有 5 个），分别是 A，B，C，D 和 E，优先级分别为 1，2，3，4 和 5，其中，任务列表中的任务按照优先级由低到高的顺序排序是 A，B，C，D 和 E。待处理的目标任务的优先级为 4。若物理节点上的剩余资源并不能满足目标任务运行所需的资源条件，则将物理节点上正在运行的任务列表中的优先级低于目标任务优先级的任务（分别是 A，B 和 C），按照优先级从低到高的顺序依次移入待移除队列，即先将 A 移入待移除队列，判断物理节点运行任务列表中的 B，C，D 和 E 后所得到的剩余资源是否满足目标任务运行所需的资源条件。若物理节点运行任务 B，C，D 和 E 后所得到的剩余资源满足目标任务运行所需的资源条件，采用目标任务将正在运行的 A 抢占，即将 A 停止。

若将 A 移入待移除队列后，物理节点运行任务列表中的 B，C，D 和 E 所得到的剩余资源不满足目标任务运行所需的资源条件，则继续将 B 移入到待移除队列中，再次检测运行任务列表中的任务 C，D 和 E 后所得到的剩余资源是否满足目标任务运行所需的资源条件，重复上述的判断步骤，直至检测到运行任务列表中的任务后所得到的剩余资源满足目标任务运行所需的资源条件。若直至任务列表中的任务的优先级不低于目标任务的优先级时（即任务列表中的任务 A，B 和 C 均被移入到待移出队列中，只剩下任务 D 和 E 时），物理节点运行任务列表中的任务 D 和 E 后所得到的剩余资源不满足目标任务运行所需的资源条件，则拒绝运行目标任务。

在 S440 中，所述物理节点调用执行环境，运行所述目标任务。

本申请实施例当物理节点处理待运行的目标任务时，若物理节点的剩余资源不满足目标任务运行的所需的条件，将低优先级的任务进行抢占，可以使物理节点优先处理重要任务，可以提高资源的利用率。

在上述实施例的基础上，所述的任务抢占方法还包括：所述物理节点每间隔设定时间获取资源使用信息；所述物理节点若确定所述资源使用信息达到预设限制条件，将所述任务列表中的任务按照优先级由低到高的顺序移入到所述待移除队列，直至所述物理节点执行所述任务列表中的任务后所确定的资源使用信息没有达到预设限制条件，并停止所述待移除队列中的任务。

可选地，任务列表中的任务为当前物理节点上正在运行的任务。其中，每个物理节点间隔设定时间获取资源使用信息，判断资源的使用信息是否达到预设限制条件，来判断是否需要触发任务抢占。若资源的使用信息达到预设限制条件，则触发任务抢占，否资源使用信息未达到预设限制条件，不需要触发任

务抢占。任务抢占的过程是：将任务列表中的任务按照优先级进行排序，并将任务列表中任务按照优先级由低到高的顺序依次移入到待移除队列，直至所述物理节点执行任务列表中的任务后所确定的资源使用信息没有达到预设限制条件，并停止待移除队列中的任务

其中，预设限制条件可以是资源使用达到设置值，也可以是其他限制条件。例如，若物理节点的资源使用达到设定值，则触发任务抢占。

由此，物理节点通过根据资源使用信息触发任务抢占，可以提高资源利用率以及当资源紧张时，可以从低优先级任务中抢占资源，优先处理重要任务。

图 5 是本申请实施例提供的一种任务抢占方法流程图，本申请实施例提供的方法运行在 Kubernetes 系统中。如图 5 所示，本申请实施例提供的方法包括：S510-S540。

在 S510 中，当物理节点通过 kubelet 处理待运行的目标 pod 时，获取所述物理节点上正在运行的 pod 列表。

其中，kubelet 是 Kubernetes 系统的组件，可以监视 pod，挂载 pod 所需要的 volumes，下载 pod 的 secret，通过 docker/rkt 来运行 pod 中的容器，周期的执行 pod 中为容器定义的 liveness 探针，上报 pod 的状态给系统的其他组件，以及节点的状态。

在 S520 中，所述物理节点通过 kubelet 检测其剩余资源是否满足所述目标 pod 运行所需的资源。

若物理节点检测其剩余资源不满足目标 pod 运行所需的资源，执行 S530，若物理节点检测其剩余资源满足目标 pod 运行所需的资源，执行 S540。

在 S530 中，所述物理节点通过 kubelet 将所述 pod 列表中的优先级低于所述目标 pod 优先级的 pod，按照优先级由低到高的顺序依次移入待移除队列，直至所述物理节点执行所述 pod 列表中的 pod 后所得到的剩余资源满足所述目标 pod 运行所需的资源，并采用所述目标 pod 抢占所述待移除队列中的 pod。

在 S540 中，所述物理节点通过 kubelet 调用执行环境，运行所述目标 pod。

由此，当物理节点处理待运行的目标 pod 时，若物理节点的剩余资源不满足目标 pod 运行的所需的条件，将低优先级的 pod（优先级低于目标 pod 的优先级）进行抢占，可以使物理节点优先处理重要任务（在本实施例中，该重要任务指的是目标 pod），可以提高资源的利用率。

图 6 是本申请实施例提供的一种基于抢占式调度的资源共享使用方法流程图，所述方法由基于抢占式调度的资源共享使用系统来执行，所述系统可通过软件和/或硬件来实现。本申请实施例提供的方法可以应用在集群中，在集群中可以包含多个物理节点，每个物理节点上的资源可以是多个租户的共享资源，多个物理节点可以由集群管理平台进行管理，将任务分配给物理节点以使物理节点执行相应的任务。集群管理平台可以集成在多个计算机设备中，多个计算机设备可以由用户进行操作，用户可以登录集群管理平台，并递交任务的创建

请求，集群管理平台中的 API 服务器获取用户提交的任务的创建请求，创建任务，由集群管理平台中的调度器进行任务调度，将任务合理分配给对应的物理节点，由物理节点执行该任务。

如图 6 所示，本申请实施例提供的技术方案包括：S610-S692。

在 S610 中，API 服务器获取任务的创建请求。

在 S620 中，当所述 API 服务器检测到所述任务所属租户的配额里包含与所述任务的优先级匹配的资源，且匹配资源满足所述任务的创建条件，根据所述创建请求创建所述任务。

在 S630 中，调度器获取所述 API 服务器创建的任务，并形成任务调度队列。

API 服务器实时获取任务的创建请求，并为满足条件的创建请求创建任务。因此 API 服务器中有至少一个任务。调度器获取 API 服务器创建的所有任务，并将这些任务形成任务调度队列。

在 S640 中，所述调度器从所述任务调度队列获取待调度的当前任务，并获取每个物理节点上大于等于所述当前任务的优先级的任务，形成节点-任务映射表。

调度器根据任务调度队列中任务的排列顺序，依次对任务进行调度。待调度的当前任务指的是任务调度队列中调度器当前正在进行调度的任务。

在 S650 中，所述调度器根据所述映射表以及预设筛选条件确定最符合所述预设筛选条件的目标物理节点。

在 S660 中，所述调度器将所述当前任务与所述目标物理节点绑定，并将绑定的信息发送到 API 服务器。

调度器依次对任务调度队列中的任务进行调度，得到任务调度队列中每个任务与其对应的目标物理节点的绑定信息。

在 S670 中，物理节点通过监听所述 API 服务器中任务与物理节点的绑定信息，获取对应的任务，并形成任务队列。

物理节点监听 API 服务器中的绑定信息，通过绑定信息中的目标物理节点确定与该物理节点对应的任务。该任务可以是一个也可以是多个，将这些任务形成任务队列。

在 S680 中，当所述物理节点处理所述任务队列中待运行的目标任务时，获取所述物理节点上正在运行的任务列表。

物理节点根据任务队列中任务的排列顺序，依次对任务进行处理。待运行的目标任务指的是任务队列中物理节点当前正在处理的任务。

在 S690 中，所述物理节点检测其剩余资源是否满足所述目标任务运行所需的资源。

若物理节点检测其剩余资源不满足目标任务运行所需的资源，执行 S691，若物理节点检测其剩余资源满足目标任务运行所需的资源，执行 S692。

在 S691 中，所述物理节点将所述任务列表中的优先级低于所述目标任务的任务，按照优先级由低到高的顺序移入待移除队列，直至所述物理节点执行

所述任务列表中的任务所得到的剩余资源满足所述目标任务运行所需的资源，并采用目标任务抢占所述待移除队列中的任务。

在S692中，所述物理节点调用执行环境，运行所述目标任务。

相关技术的Kubernetes 1.3版本基于服务质量（Quality Of Service）的资源共享方案，用来进行管理共享资源，服务质量从高到低分别是Guarantee，Burstable以及Best Effort。其中，Best Effort的任务可以在集群资源没有充分使用时，进行调度并运行。当集群资源紧张时，Best Effort的任务优先被抢占。这种方案并没有考虑任务调度的环节，当集群调度满时，不能为高服务质量的任务腾出资源，并且无法限制租户内Best Effort任务的数目，以及无法区分Best Effort任务被抢占的顺序。

相关技术的Kubernetes 1.8版本引入的基于优先级的调度方案，任务可以设置优先级，当资源紧张的时，调度器会将低优先级的任务抢占，为高优先级的任务提供充分的资源。但该种方案中任务的抢占发生在调度器，即在集群逻辑调度满的时候发生，集群中会有资源没有充分利用的情况，资源的利用率并不高，并且无法精确限制每个优先级任务的数目。

相对于相关技术中的调度方案，本申请实施例提供的方法，对租户的资源配额进行优先级的设置，可以精确限制每个租户下多种优先级任务的数目，通过对任务优先级的设置，当任务被抢占时，可以区分任务被抢占的顺序。本申请实施例中基于优先级的任务抢占发生在物理节点，并没有发生在调度器，可以在逻辑上为高优先级的任务腾出资源，在资源没有充分利用时，继续运行被抢占的任务，可以提高资源的利用率。

本申请实施例提供的方法，通过对租户的资源配额进行优先级的设置，并通过将任务的优先级与所属租户下多个优先级的资源进行匹配，从而确定是否创建任务，可以让租户在资源紧张时优先使用资源，防止租户滥用高优先级的资源，导致低优先级的任务持续获取不了资源，出现“饿死”现象。调度器基于任务的优先级以及预设筛选条件对物理节点进行筛选，筛选出最合适的物理节点，将待调度的当前任务调度到最合适的物理节点，当资源紧张时，只进行逻辑上的资源抢占，并没有立即抢占资源，这种延后抢占的调度方法，可以在逻辑上为高优先级的任务让出资源，在资源没有被充分利用时，继续运行被抢占的任务，可以提高资源的利用率。当物理节点处理待运行的目标任务时，若物理节点的剩余资源不满足目标任务运行的所需的条件，基于将低优先级的任务进行抢占，可以使物理节点优先处理重要任务，提高资源的利用率。

图7是本申请实施例提供的一种API服务器结构框图，如图7所示，所述API服务器包括：请求获取模块710和任务创建模块720。

请求获取模块710设置为获取任务的创建请求。

任务创建模块720设置为当检测到所述任务所属租户的配额里包含与所述任务的优先级匹配的资源，且匹配资源满足所述任务的创建条件，根据所述创建请求创建所述任务。

可选地，所述装置应用于 Kubernetes 系统中，请求获取模块 710 设置为获取 pod 的创建请求。

任务创建模块 720 设置为当检测到所述 pod 所属的 namespace 的 quota 里包含与所述 pod 的优先级匹配的资源 quota 值，且 quota 值满足所述 pod 的创建条件，根据所述创建请求创建 pod。

上述任务创建装置可执行本申请任意实施例所提供的任务创建方法，具备执行任务创建方法相应的功能模块和有益效果。

图 8a 是本申请实施例提供的一种调度器的结构框图，如图 8a 所示，所述调度器包括：映射表形成模块 810、筛选模块 820 和绑定模块 830。

映射表形成模块 810 设置为从任务调度队列获取待调度的当前任务，并获取每个物理节点上大于等于所述当前任务的优先级的任务，形成节点-任务映射表。

筛选模块 820 设置为根据所述映射表以及预设筛选条件确定最符合所述预设筛选条件的目标物理节点。

绑定模块 830 设置为将所述当前任务与所述目标物理节点绑定，并将绑定的信息发送到 API 服务器。

可选地，筛选模块 820 设置为从所述映射表筛选出符合第一阶段筛选条件的物理节点，形成节点组；以及

根据所述映射表以及第二阶段优选条件，对所述节点组的物理节点进行评分，并筛选出分数最高的物理节点作为目标物理节点。

可选地，所述装置应用于 Kubernetes 系统中，映射表形成模块 810 设置为从 pod 调度队列中获取待调度的当前 pod，并获取每个物理节点上大于等于所述当前 pod 指定优先级的 pod，形成节点-pod 映射表。

相应的，绑定模块 830 设置为将所述当前 pod 与所述目标物理节点绑定，并将绑定的信息发送到 API 服务器。

其中，调度器结构还可以是其他的结构形式，以可以执行任务调度方法即可。例如，调度器中可包括调度系统，如图 8b 所示，调度系统可以包括四个部分：节点信息列表 840、筛选算法库 850、优选算法库 860 和未调度队列 870。

其中，节点信息列表 840 设置为记载当前可用的物理节点信息，包括物理节点上的资源信息（全部资源和可用资源），以及已经在物理节点上运行的任务队列。这部分信息是调度方法指定时关键的信息，需要实时同步，以保证调度系统对资源以及任务有全面的认知。

筛选算法库 850 设置为预先定义了多种筛选物理节点的算法，保证去除不满足任务执行条件的物理节点。

优选算法库 860 设置为预先定义了多种优选节点的算法以及算法的权重，优选算法计算出打分最高的物理节点会被选为调度节点，即目标物理节点。

调度队列 870 设置为未调度的任务形成的队列，是一个优先级队列以保证高优先级的任务先调度。

上述装置可执行本申请任意实施例所提供的任务调度方法，具备执行任务调度方法相应的功能模块和有益效果。

图 9 是本申请实施例提供的一种任务抢占装置的结构框图，如图 9 所示，所述任务抢占装置包括：任务列表获取模块 910、检测模块 920、抢占模块 930 和任务执行模块 940。

其中，任务列表获取模块 910 设置为当处理待运行的目标任务时，获取所述物理节点上正在运行的任务列表。

检测模块 920 设置为检测物理节点上的剩余资源是否满足所述目标任务运行所需的资源。

抢占模块 930 设置为当物理节点上的剩余资源不满足目标任务运行所需的资源时，所述物理节点将所述任务列表中的优先级低于所述目标任务的任务，按照优先级由低到高的顺序依次移入待移除队列（即按照优先级由低到高的顺序依次将任务移出任务列表），直至所述物理节点执行所述任务列表中的任务后所得到的剩余资源满足所述目标任务运行所需的资源，并采用所述目标任务抢占所述待移除队列中的任务。

任务执行模块 940 设置为所述物理节点调用执行环境，运行所述目标任务。

可选地，所述抢占模块还可以设置为每间隔设定时间获取资源使用信息。

若确定所述资源使用信息达到预设限制条件，将所述任务列表中的任务按照优先级由低到高的顺序依次移入到所述待移除队列，直至所述物理节点执行所述任务列表中的任务后所确定的资源使用信息没有达到预设限制条件，并停止所述待移除队列中的任务。

可选地，所述装置应用于 Kubernetes 系统中，所述目标任务为目标 pod，所述任务列表为 pod 列表，所述任务列表中的任务为所述物理节点上正在运行的 pod。

上述装置可执行本申请任意实施例所提供的任务抢占方法，具备执行任务抢占方法相应的功能模块和有益效果。

图 10 是本申请实施例提供的一种任务抢占系统的结构示意图，如图 10 所示，所述任务抢占系统包括上述实施例提供的 API 服务器 1010、上述实施例提供的调度器 1020 以及上述实施例提供的任务抢占装置 1030。

可选地，API 服务器 1010 和调度器 1020 分别是集群管理平台的组件，集群管理平台集成在用户所用的计算机设备上。任务抢占装置 1030 可集成在物理节点中的物理机上。

图 11 是本申请实施例提供的一种设备结构示意图，如图 11 所示，该设备包括：一个或多个处理器 1110，图 11 中以一个处理器 1110 为例；存储器 1120。

所述设备还可以包括：输入装置 1130 和输出装置 1140。

所述设备中的处理器 1110、存储器 1120、输入装置 1130 和输出装置 1140 可以通过总线或者其他方式连接，图 11 中以通过总线连接为例。

存储器 1120 作为一种非暂态计算机可读存储介质，可用于存储软件程序、计算机可执行程序以及模块，如本申请实施例中的一种任务创建方法对应的程序指令/模块（例如，附图 7 所示的请求获取模块 710 和任务创建模块 720）或者如本申请实施例中的一种任务调度方法对应的程序指令/模块（例如，附图 8 所示的映射表形成模块 810、筛选模块 820 和绑定模块 830），或者如本申请实施例中的一种任务抢占方法对应的程序指令/模块（例如，附图 9 所示的任务列表获取模块 910、检测模块 920、抢占模块 930 和任务执行模块 940）。处理器 1110 通过运行存储在存储器 1120 中的软件程序、指令以及模块，从而执行计算机设备的多种功能应用以及数据处理，即实现上述方法实施例的一种任务创建方法：API 服务器获取任务的创建请求；以及当 API 服务器检测到所述任务所属租户的配额里包含与所述任务的优先级匹配的资源，且匹配资源满足所述任务的创建条件，根据所述创建请求创建所述任务。

或者处理器 1110 通过运行存储在存储器 1120 中的软件程序、指令以及模块，实现上述方法实施例的一种任务调度方法：调度器从任务调度队列中获取待调度的当前任务，并获取每个物理节点上大于等于所述当前任务的优先级的任务，形成节点-任务映射表；所述调度器根据所述映射表以及预设筛选条件确定最符合所述预设筛选条件的目标物理节点；以及所述调度器将所述当前任务与所述目标物理节点绑定，并将绑定的信息发送到 API 服务器。

或者处理器 1110 通过运行存储在存储器 1120 中的软件程序、指令以及模块，实现上述方法实施例的一种任务抢占方法：当物理节点处理待运行的目标任务时，获取所述物理节点上正在运行的任务列表；所述物理节点检测其剩余资源是否满足所述目标任务运行所需的资源；若物理节点检测其剩余资源不满足目标任务运行所需的资源，所述物理节点将所述任务列表中的优先级低于所述目标任务优先级的任务，按照优先级从低到高的顺序依次移入待移除队列，直至所述物理节点执行所述任务列表中的任务后所得到的剩余资源满足所述目标任务运行所需的资源，并采用所述目标任务抢占所述待移除队列中的任务；以及所述物理节点调用执行环境，运行所述目标任务。

存储器 1120 可以包括存储程序区和存储数据区，其中，存储程序区可存储操作系统、至少一个功能所需要的应用程序；存储数据区可存储根据计算机设备的使用所创建的数据等。此外，存储器 1120 可以包括高速随机存取存储器，还可以包括非暂态性存储器，例如至少一个磁盘存储器件、闪存器件、或其他非暂态性固态存储器件。在一些实施例中，存储器 1120 可选包括相对于处理器 1110 远程设置的存储器，这些远程存储器可以通过网络连接至终端设备。上述网络的实例包括但不限于互联网、企业内部网、局域网、移动通信网及其组合。

输入装置 1130 可用于接收输入的数字或字符信息，以及产生与计算机设备的用户设置以及功能控制有关的键信号输入。输出装置 1140 可包括显示屏

等显示设备。

本申请实施例提供了一种计算机可读存储介质，其上存储有计算机程序，该程序被处理器执行时，实现如本申请实施例提供的一种任务创建方法：获取任务的创建请求；以及当检测到所述任务所属租户的配额里包含与所述任务的优先级匹配的资源，且匹配资源满足所述任务的创建条件，根据所述创建请求创建所述任务。

或者该程序被处理器执行时，实现上述方法实施例的一种任务调度方法，即：调度器从任务调度队列中获取待调度的当前任务，并获取每个物理节点上大于等于所述当前任务的优先级的任务，形成节点-任务映射表；所述调度器根据所述映射表以及预设筛选条件确定最符合所述预设筛选条件的目标物理节点；以及所述调度器将所述当前任务与所述目标物理节点绑定，并将绑定的信息发送到 API 服务器。

或者该程序被处理器执行时，实现上述方法实施例的一种任务抢占方法：物理节点处理待运行的目标任务时，获取所述物理节点上正在运行的任务列表；所述物理节点检测其剩余资源是否满足所述目标任务运行所需的资源；若物理节点检测其剩余资源不满足目标任务运行所需的资源，所述物理节点将所述任务列表中的优先级低于所述目标任务优先级的任务，按照优先级从低到高的顺序依次移入待移除队列，直至所述物理节点执行所述任务列表中的任务后所得到的剩余资源满足所述目标任务运行所需的资源，并采用所述目标任务抢占所述待移除队列中的任务；以及所述物理节点调用执行环境，运行所述目标任务。

可以采用一个或多个计算机可读的介质的任意组合。计算机可读介质可以是计算机可读信号介质或者计算机可读存储介质。计算机可读存储介质例如可以是——但不限于——电、磁、光、电磁、红外线、或半导体的系统、装置或器件，或者任意以上的组合。计算机可读存储介质的更具体的例子（非穷举的列表）包括：具有一个或多个导线的电连接、便携式计算机磁盘、硬盘、随机存取存储器（RAM）、只读存储器(ROM)、可擦式可编程只读存储器(EPROM 或闪存)、光纤、便携式紧凑磁盘只读存储器(CD-ROM)、光存储器件、磁存储器件、或者上述的任意合适的组合。在本文件中，计算机可读存储介质可以是任何包含或存储程序的有形介质，该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。

计算机可读的信号介质可以包括在基带中或者作为载波一部分传播的数据信号，其中承载了计算机可读的程序代码。这种传播的数据信号可以采用多种形式，包括——但不限于——电磁信号、光信号或上述的任意合适的组合。计算机可读的信号介质还可以是计算机可读存储介质以外的任何计算机可读介质，该计算机可读介质可以发送、传播或者传输用于由指令执行系统、装置或者器件使用或者与其结合使用的程序。

计算机可读介质上包含的程序代码可以用任何适当的介质传输，包括——但不限于——无线、电线、光缆、RF 等等，或者上述的任意合适的组合。

可以以一种或多种程序设计语言或其组合来编写用于执行本申请操作的计

算机程序代码，所述程序设计语言包括面向对象的程序设计语言—诸如 Java、Smalltalk、C++，还包括常规的过程式程序设计语言—诸如“C”语言或类似的程序设计语言。程序代码可以完全地在用户计算机上执行、部分地在用户计算机上执行、作为一个独立的软件包执行、部分在用户计算机上部分在远程计算机上执行、或者完全在远程计算机或服务器上执行。在涉及远程计算机的情形中，远程计算机可以通过任意种类的网络——包括局域网(LAN)或广域网(WAN)—连接到用户计算机，或者，可以连接到外部计算机（例如利用因特网服务提供商来通过因特网连接）。

权 利 要 求 书

1、一种任务创建方法，包括：

应用编码接口 API 服务器获取任务的创建请求；以及

当 API 服务器检测到所述任务所属租户的配额里包含与所述任务的优先级匹配的资源，且匹配资源满足所述任务的创建条件，根据所述创建请求创建所述任务。

2、根据权利要求 1 所述的方法，其中，

所述 API 服务器获取任务的创建请求包括：

API 服务器获取 pod 的创建请求；以及

当 API 服务器检测到所述任务所属租户的配额里包含与所述任务的优先级匹配的资源，且匹配资源满足所述任务的创建条件，根据所述创建请求创建所述任务，包括：

当所述 API 服务器检测到所述 pod 所属的 namespace 的 quota 里包含与所述 pod 对应的优先级匹配的资源 quota 值，且所述 quota 值满足所述 pod 的创建条件，根据所述创建请求创建 pod。

3、一种任务调度方法，包括：

调度器从任务调度队列中获取待调度的当前任务，并获取每个物理节点上大于等于所述当前任务优先级的任务，形成节点-任务映射表；

所述调度器根据所述映射表以及预设筛选条件确定符合所述预设筛选条件的目标物理节点；以及

所述调度器将所述当前任务与所述目标物理节点绑定，并将绑定的信息发送到 API 服务器。

4、根据权利要求 3 所述的方法，其中，所述调度器根据所述映射表以及预设筛选条件确定符合所述预设筛选条件的目标物理节点，包括：

所述调度器从所述映射表筛选出符合第一阶段筛选条件的物理节点，形成节点组；以及

根据所述映射表以及第二阶段优选条件，对所述节点组的物理节点进行评分，并筛选出分数最高的物理节点作为目标物理节点。

5、根据权利要求 3 或 4 所述的方法，其中，

所述调度器从任务调度队列中获取待调度的当前任务，并获取每个物理节点上大于等于所述当前任务指定优先级的任务，形成节点-任务映射表，包括：所述调度器从 pod 调度队列中获取待调度的当前 pod，并获取每个物理节点上大于等于所述当前 pod 优先级的 pod，形成节点-pod 映射表；以及

所述调度器将所述当前任务与所述目标物理节点绑定，并将绑定的信息发送到 API 服务器，包括：

所述调度器将所述当前 pod 与所述目标物理节点绑定，并将绑定的信息发送到 API 服务器。

6、一种任务抢占方法，包括：

当物理节点处理待运行的目标任务时，获取所述物理节点上正在运行的任务列表；

所述物理节点检测其剩余资源是否满足所述目标任务运行所需的资源;

若物理节点检测其剩余资源不满足目标任务运行所需的资源时，所述物理节点将所述任务列表中的优先级低于所述目标任务优先级的任务按照优先级从低到高的顺序依次移入待移除队列，直至所述物理节点执行所述任务列表中的任务后所得到的剩余资源满足所述目标任务运行所需的资源，并采用所述目标任务抢占所述待移除队列中的任务；以及

所述物理节点调用执行环境，运行所述目标任务。

7、根据权利要求 6 所述的方法，还包括:

所述物理节点每间隔设定时间获取资源使用信息;

所述物理节点若确定所述资源使用信息达到预设限制条件，将所述任务列表中的任务按照优先级由低到高的顺序依次移入到所述待移除队列，直至所述物理节点执行所述任务列表中的任务时所确定的资源使用信息没有达到预设限制条件，并停止所述待移除队列中的任务。

8、根据权利要求 6 或者 7 所述的方法，其中，所述目标任务为目标 pod，所述任务列表为 pod 列表，所述任务列表中的任务为所述物理节点上正在运行的 pod。

9、一种基于抢占式调度的资源共享使用方法，包括:

应用编码接口 API 服务器获取任务的创建请求;

当所述 API 服务器检测到所述任务所属租户的配额里包含与所述任务的优先级匹配的资源，且匹配资源满足所述任务的创建条件，根据所述创建请求创建所述任务;

调度器获取所述 API 服务器创建的任务，并形成任务调度队列;

所述调度器从所述任务调度队列获取待调度的当前任务，并获取每个物理节点上大于等于所述当前任务指定优先级的任务，形成节点-任务映射表;

所述调度器根据所述映射表以及预设筛选条件确定最符合所述预设筛选条件的目标物理节点;

所述调度器将所述当前任务与所述目标物理节点绑定，并将绑定的信息发送到 API 服务器;

物理节点监听所述 API 服务器中任务与物理节点的绑定信息，基于监听到的所述绑定信息获取对应的任务，并形成任务队列;

当所述物理节点处理所述任务队列中待运行的目标任务时，获取所述物理节点上正在运行的任务列表;

所述物理节点检测其执行所述任务列表中的任务后的剩余资源是否满足所述目标任务运行所需的资源;

若不满足目标任务运行所需的资源，所述物理节点将所述任务列表中的优先级低于所述目标任务的任务，按照优先级从低到高的顺序依次移入待移除队列，直至所述物理节点的剩余资源满足所述目标任务运行所需的资源，并采用目标任务抢占所述待移除队列中的任务；以及

所述物理节点调用执行环境，运行所述目标任务。

10、一种 API 服务器，包括:

请求获取模块，设置为获取任务的创建请求；以及

任务创建模块，设置为当检测到所述任务所属租户的配额里包含与所述任务的优先级匹配的资源，且匹配资源满足所述任务的创建条件，根据所述创建请求创建所述任务。

11、一种调度器，包括:

映射表形成模块，设置为从任务调度队列中获取待调度的当前任务，并获取每个物理节点上大于等于所述当前任务优先级的任务，形成节点-任务映射表；

筛选模块，设置为根据所述映射表以及预设筛选条件确定符合所述预设筛选条件的目标物理节点；以及

绑定模块，设置为将所述当前任务与所述目标物理节点绑定，并将绑定的信息发送到 API 服务器。

12、一种任务抢占装置，包括:

任务列表获取模块，设置为当处理待运行的目标任务时，获取物理节点上正在运行的任务列表；

检测模块，设置为检测物理节点上的剩余资源是否满足所述目标任务运行所需的资源；

抢占模块，设置为当检测模块检测到物理节点上的剩余资源不满足目标任务运行所需的资源时，将所述任务列表中的任务按照优先级由低到高的顺序依次移入待移除队列，直至所述物理节点执行所述任务列表中的任务后所得到的剩余资源满足所述目标任务运行所需的资源，且所述待移除队列中任务的优先级小于所述目标任务的优先级，并采用所述目标任务抢占所述待移除队列中的任务；

任务执行模块，设置为调用执行环境，运行所述目标任务。

13、一种基于抢占式调度的资源共享使用系统，包括如权利要求 10 所述的 API 服务器、权利要求 11 所述的调度器以及权利要求 12 所述的任务抢占装置。

14、一种设备，包括:

一个或多个处理器；

存储装置，用于存储一个或多个程序，

当所述一个或多个程序被所述一个或多个处理器执行，使得所述一个或多个处理器实现如权利要求 1 或 2 所述的任务创建方法，或者实现如权利要求 3-5 中任一所述的任务调度方法，或者实现如权利要求 6-8 中任一所述的任务抢占方法。

15、一种计算机可读存储介质，其上存储有计算机程序，该程序被处理器执行时实现如权利要求 1 或 2 所述的任务创建方法，或者实现如权利要求 3-5 中任一所述的任务调度方法，或者实现如权利要求 6-8 中任一所述的任务抢占方法。