

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2022 年 4 月 7 日 (07.04.2022)

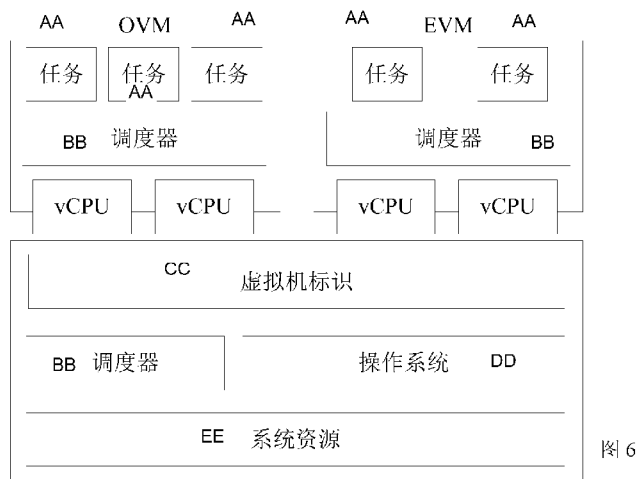


(10) 国际公布号
WO 2022/068697 A1

- (51) 国际专利分类号:
G06F 9/50 (2006.01)
- (21) 国际申请号: PCT/CN2021/120360
- (22) 国际申请日: 2021 年 9 月 24 日 (24.09.2021)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
202011053405.3 2020 年 9 月 29 日 (29.09.2020) CN
- (71) 申请人: 华为技术有限公司 (HUAWEI TECHNOLOGIES CO., LTD.) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。
- (72) 发明人: 张海亮 (ZHANG, Hailiang); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。
- (74) 代理人: 北京同立钧成知识产权代理有限公司 (LEADER PATENT & TRADEMARK FIRM); 中国北京市海淀区西直门北大街 32 号枫蓝国际 A 座 8F-6, Beijing 100082 (CN)。
- (81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,

(54) Title: TASK SCHEDULING METHOD AND APPARATUS

(54) 发明名称: 任务调度方法及装置



AA Task
BB Scheduler
CC Virtual machine identifier
DD Operating system
EE System resources

(57) Abstract: Provided are a task scheduling method and apparatus. When scheduling a task to be executed by a first target virtual machine, a scheduler first obtains at least one task with the highest priority in a queue of tasks to be executed by the first target virtual machine, so as to ensure that the task with the highest priority is preferentially executed; then, a first task with the minimum virtual running time is further determined from the at least one task with the highest priority, and finally the first target virtual machine is controlled to execute the first task, thereby ensuring the fair scheduling of a task from at least one task of the same priority. Therefore, by means of the present application, when scheduling the task on the first target virtual machine, the scheduler can ensure that a task with a high priority is preferentially scheduled, and can also achieve fairness when scheduling tasks of the same priority.

WO 2022/068697 A1

ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US,
UZ, VC, VN, WS, ZA, ZM, ZW。

- (84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

- 包括国际检索报告(条约第21条(3))。
-

(57) 摘要: 本申请提供一种任务调度方法及装置, 调度器在调度第一目标虚拟机待执行的任务时, 首先获取第一目标虚拟机待执行任务队列中优先级最高的至少一个任务, 以首先保证最高优先级的任务优先执行, 随后还在优先级最高的至少一个任务中进一步确定虚拟运行时间最小的第一任务, 并最终控制第一目标虚拟机执行第一任务, 保证了相同优先级的至少一个任务中对任务的公平调度, 因此, 本申请能够实现调度器能够在调度第一目标虚拟机上的任务时, 既能够保证高优先级的任务被优先调度、又能够满足相同优先级任务调度时的公平。

任务调度方法及装置

5 本申请要求于 2020 年 09 月 29 日提交中国专利局、申请号为 202011053405.3、申请名称为《任务调度方法及装置》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

本申请涉及计算机技术领域，尤其涉及一种任务调度方法及装置。

10

背景技术

在计算机领域，虚拟机技术允许一个服务器同时执行多个操作系统，每个操作系统可以作为一个虚拟机，在彼此独立的环境中执行而互不影响。因此，越来越多的服务商提供了虚拟机服务，用户可以使用服务商的虚拟机提供的计算服务。而对于服务商，由于不同用户所使用的虚拟机可以执行在同一个服务器上，极大地提高了服务商所部署的服务器使用率。同时，用户在使用服务商提供的虚拟机时，通常按照所需计算量的峰值申请虚拟机的资源，服务商也会将服务器的资源按照计算量的峰值划分给不同的虚拟机使用。但是在实际使用过程中，用户通常不会使用其所申请的全部的虚拟机资源，造成服务器向每个用户分配的虚拟机资源的利用率不高。

15 因此为了对服务器上有限的资源进行更加合理的使用，现有技术中服务商可以在同一个向用户分配的虚拟机资源中，执行不同用户的任务，由于不同用户的任务类型不同、优先级不同，因此可以在服务器中设置调度器，对虚拟机资源中需要执行的任务进行统一的管理与调度。①、在一种实现中，调度器通过完全公平调度策略对虚拟机中执行的任务进行管理，其中，调度器为虚拟机中执行的所有任务按照每个任务的优先权重值，分配不同的执行时间，保证所有任务都得到执行。但是，这种调度器只能够保证优先级较高的任务分配更多的执行时间，而无法保证优先调度优先级较高的任务，也无法保证高优先级任务抢占低优先级任务的执行。②、在另一种实现中，调度器通过实时调度器策略对虚拟机中执行的任务进行管理，其中，调度器按照所有任务的优先级，首先执行最高优先级的任务中最先进入队列的一个、或者轮流执行最高优先级的多个任务。但是，这种调度器在优先级较高的任务执行完之后才能够执行优先级较低的任务，因此无法保证相同优先级任务调度的公平性，同时无法解决调度引起的低优先级任务饥饿问题。

20 30 35 综上，在现有技术中，完全公平调度策略和实时调度器策略都各自存在着不足，当一个虚拟机上可以执行不同类型的任务时，无法同时满足调度优先调度高优先级的任务和保证低优先级任务也得到调度的公平。因此，如何设计一种调度器，使得调度器能够用于调度虚拟机上执行的不同优先级的任务时，还能够满足调度的公平，是本领域亟需解决的技术问题。

发明内容

本申请提供一种任务调度方法及装置，以使调度器能够在调度虚拟机的待执行的任务时，既能够保证高优先级的任务被优先调度、又能够满足相同优先级任务调度时的公平。

5 本申请第一方面提供一种任务调度方法，用于调度器在调度第一目标虚拟机待执行的任务时，可以满足对所调度的任务优先级的要求的同时，还能够保证任务的公平调度。

其中，调度器在调度第一目标虚拟机待执行的任务时，首先获取第一目标虚拟机待执行任务队列中优先级最高的至少一个任务，以首先保证最高优先级的任务优先执行，随后还在优先级最高的至少一个任务中进一步确定虚拟运行时间最小的第一任务，并最终控制第一目标虚拟机执行第一任务，保证了相同优先级的至少一个任务中对任务的公平调度，也就是说，本实施例能够实现了调度器能够在调度第一目标虚拟机上的任务时，既能够保证高优先级的任务被优先调度、又能够满足相同优先级任务调度时的公平。

15 在本申请第一方面一实施例中，作为执行主体的调度器可以同时多个虚拟机进行调度，并且调度器在调度多个虚拟机中的第一目标虚拟机执行的任务时，可以通过控制第一目标虚拟机的第一调度器实现。例如，调度器可以通过第一调度器中存储的待执行任务队列，确定第一目标虚拟机待执行的多个任务的优先级，又例如，调度器在控制第一目标虚拟机执行第一任务时，具体可以通过调整第一调度器中待执行任务队列的顺序，将第一任务调整到待执行任务队列的最前实现。

20 对于调度器确定第一目标虚拟机待执行的任务中，优先级最高的至少一个任务，本申请提供至少以下三种方式具体实现：

在本申请第一方面一实施例提供的第一种方式中，基于第一调度器中存储了第一目标虚拟机待执行任务队列，则调度器在确定优先级最高的至少一个任务时，可以从第一调度器中获取待执行任务队列中所有的任务后，再确定所有任务对应的优先级，并最终确定出优先级最高的至少一个任务。

25 在本申请第一方面一实施例提供的第二种方式中，第一调度器在对第一目标虚拟机的待执行任务进行调度时，除了存储所有待执行的任务，还对队列中所有任务的优先级进行记录，例如都记录在一个表格中。则调度器在确定第一目标虚拟机待执行任务的优先级时，可以直接从表格中获取所有任务对应的优先级，而不用获取每个任务本身，从而减少计算量、提高效率。

30 在本申请第一方面一实施例提供的第三种方式中，第一调度器在对第一目标虚拟机的待执行任务进行调度时，可以设置不同的多个队列来存储任务，并且每个队列所存储的队列的优先级不同。则调度器在确定第一目标虚拟机待执行任务中优先级最高的至少一个任务时，可以从第一调度器中存储的优先级最高的队列中，直接获取这至少一个任务，而不用处理其他队列上其他优先级的任务，同样能够减少计算量、提高效率。

35 本申请第二方面提供一种任务调度方法，用于调度器在调度第一目标虚拟机执行任务时，当第一目标虚拟机待执行任务队列中任务的优先级发生了调整，则调度器根

据优先级调整后的任务的优先级，对该任务的虚拟运行时间进行调整，来实现高优先级任务对第一目标虚拟机当前执行任务的抢占，以满足对第一目标虚拟机执行任务的动态调度。

5 其中，假设第一目标虚拟机正在执行第三任务，若此时第一目标虚拟机的待执行任务中，第二任务的优先级发生了调整，例如由第二优先级调整为第一优先级，则根据第一目标虚拟机待执行的与第一优先级相同的至少一个任务中的最小虚拟机运行时间，对第二任务的虚拟运行时间进行调整，例如由第一虚拟运行时间调整为第二虚拟运行时间，最终，在第二任务的优先级高于第三任务的优先级时，调度器随即控制第一目标虚拟机执行第二任务，从而实现了第二任务对第一目标虚拟机当前执行任务的
10 抢占执行，实现了对任务的动态调度。

在本申请第二方面一实施例中，在调度器对第二任务的虚拟运行时间进行调整时，调整的目标是将调整后的第二虚拟运行时间，调整为大于第一目标虚拟机所有待执行任务中，第一优先级对应的至少一个任务的虚拟运行时间中最小的虚拟运行时间。也就是说，对第二任务的虚拟运行时间进行调整，使得其虚拟运行时间尽可能的小，但
15 是不能小于同优先级中最小的虚拟运行时间。

在本申请第二方面一实施例中，调度器具体根据用户发送的指示信息，确定第二任务的优先级需要调整，使得用户通过调度器对第一目标虚拟机的待执行任务进行调整后，调度器即可调整任务的虚拟运行时间，实现用户对任务的动态调整，能够丰富
20 应用场景，以及提高用户体验。

在本申请第二方面一实施例中，调度器在对于第二任务的虚拟运行时间进行调整时，具体将该第二任务调整前的第一虚拟运行时间 T_a ，加上同优先级中最小的虚拟运行时间 $\min B$ 之后，再减去第二优先级对应的至少一个任务中虚拟运行时间的最小值记为 $\min A$ ，即通过公式第二虚拟运行时间 = $(T_a - \min A) + \min B$ ，来将第一虚拟运行时间
25 调整为第二虚拟运行时间，使得调整后的第二虚拟运行时间在第一优先级对应的任务中，略大于第一优先级的至少一个任务中虚拟运行时间的最小值。

本申请第三方面提供一种任务调度方法，应用于调度器在调度第一目标虚拟机中的任务时，当第一目标虚拟机的待运行任务队列中存在低优先级的任务一直得不到执行的情况，可以将低优先级的任务主动迁移到其他虚拟机（记为第二目标虚拟机）上
30 执行，从而防止由于第一目标虚拟机执行多种优先级的任务时，优先调度高优先级的任务进而导致的低优先级任务的饥饿问题。

其中，调度器在第一目标虚拟机执行第四任务时，同时获取第一目标虚拟机的第一调度器中对于第四任务的时间记录信息，并根据时间记录信息对优先级较低的第五任务是否出现饥饿问题进行判断，在时间记录信息满足预设条件之后，控制第一目标
35 虚拟机的第一调度器将第五任务发送给第二目标虚拟机的第二调度，使得第二目标虚拟机执行第五任务，从而实现了对第五任务在不同虚拟机之间的调度，使得第五任务不会由于优先级较低而一直停留在第一调度器的待执行队列中，防止饥饿问题的发生。

更为具体地，本申请提供的时间记录信息可以至少包括如下两种可能的实现方式：

在本申请第三方面一实施例中，时间记录信息一种可能的实现是：可用于记录第一目标虚拟机已经执行第四任务的执行时间，则预设条件可以是第四任务的执行时间

小于第一阈值。其中，能够防止由于第四任务的优先级较高，处于较低优先级的第五任务一直得不到执行，会导致第五任务的饥饿问题，因此，调度器可以确定第四任务执行时间大于第一阈值而满足预设条件后，调度由第二目标虚拟机执行第五任务。

5 在本申请第三方面一实施例中，时间记录信息的另一种可能的实现是：用于记录第五任务在第一调度器的待执行任务队列中的等待时间，则预设条件可以是第五任务的等待时间大于第二阈值。同样为了防止第五任务的饥饿问题，调度器可以确定第五任务的等待时间大于第二阈值而满足预设条件后，调度由第二目标虚拟机执行第五任务。

10 可选地，在调度器调度第二目标虚拟机执行第五任务时，可以参照一定的标准来选择第二目标虚拟机。

15 例如，在本申请第三方面一实施例中，可以按照如下顺序确定第二目标虚拟机：优先选择云服务器中空闲态的虚拟机、其次选择云服务器中虚拟机的调度队列上的任务优先级较低且负载较轻的虚拟机、再次选择云服务器中虚拟机的调度队列上的任务优先级较低但负载较重的虚拟机/服务器中虚拟机的调度队列上的任务优先级较高但负载较轻的虚拟机。

又例如，在本申请第三方面另一实施例中，第二目标虚拟机还可以是云服务器中专门设置的虚拟机，该虚拟机专用于在防止饥饿问题的场景中，执行调度器从其他虚拟机调度的任务，并且在没有被调度器调度时，该虚拟机不会执行其他任务。

20 本申请第四方面提供一种任务调度装置，可用于执行如本申请第一方面的任务调度方法，该装置包括：获取模块、优先级确定模块、虚拟运行时间确定模块和控制模块。其中，获取模块用于获取第一目标虚拟机待执行的多个任务的优先级；优先级确定模块用于确定多个任务中优先级最高的至少一个任务；虚拟运行时间确定模块用于从至少一个任务中，确定虚拟运行时间最小的第一任务；控制模块用于控制第一目标虚拟机执行第一任务。

25 在本申请第四方面一实施例中，获取模块具体用于，根据第一目标虚拟机中，第一调度器所存储的待执行任务队列，确定第一目标虚拟机待执行的多个任务的优先级。

在本申请第四方面一实施例中，获取模块具体用于，遍历第一调度器所存储的待执行任务队列上所有任务的优先级，从待执行任务的队列中确定优先级最高的至少一个任务。

30 在本申请第四方面一实施例中，获取模块具体用于，通过第一调度器中存储的优先级记录信息，确定多个任务中优先级最高的至少一个任务；其中，优先级记录信息用于记录第一调度器待执行的队列上所有任务的优先级。

35 在本申请第四方面一实施例中，获取模块具体用于，根据第一调度器中存储的最高优先级对应的队列中的任务，确定优先级最高的至少一个任务；其中，第一调度器待执行任务通过多个队列存储待执行的任务，每个队列对应不同的优先级。

本申请第五方面提供一种任务优先级调整装置，可用于执行如本申请第二方面提供的任务调度方法，该装置包括：获取模块、虚拟运行时间调整模块和控制模块。其中，获取模块用于当第一目标虚拟机待执行的第二任务的优先级由第二优先级调整为第一优先级，获取第一目标虚拟机待执行的与第一优先级相同的至少一个任务中，至

少一个任务对应的最小的虚拟运行时间；虚拟运行时间调整模块用于根据最小的虚拟运行时间，将第二任务对应的第一虚拟运行时间调整为第二虚拟运行时间；控制模块用于，当第一优先级高于第一目标虚拟机正在执行的第三任务的优先级，控制第一目标虚拟机执行第二任务。

5 在本申请第五方面一实施例中，第二虚拟运行时间大于第一目标虚拟机的第一调度器存储的待执行任务队列中，对应第一优先级的至少一个任务的虚拟运行时间中最小的虚拟运行时间。

在本申请第五方面一实施例中，所述任务调度装置还包括：接收模块，用于接收指示信息，并根据指示信息，确定第二任务的优先级由第二优先级调整为第一优先级。

10 在本申请第五方面一实施例中，虚拟时间调整模块具体用于，将第一虚拟运行时间加上最小的虚拟运行时间后，再减去待执行任务队列中对应第一优先级的至少一个任务的虚拟运行时间中最小的虚拟运行时间，得到第二虚拟运行时间。

本申请第六方面提供一种任务调度装置，可用于执行如本申请第三方面提供的任务调度方法，该装置包括：获取模块和控制模块。其中，获取模块用于获取第一目标虚拟机的第一调度器存储的时间记录信息；其中，时间记录信息用于记录第四任务执行的时间，第四任务对应第三优先级，或者，时间记录信息用于记录第五任务在第一调度器存储的待执行任务队列中的等待时间，第五任务对应的优先级低于第三优先级；控制模块用于当时间记录信息满足预设条件，控制第一目标虚拟机的第一调度器将第五任务发送到第二目标虚拟机的第二调度器，使第二目标虚拟机执行第五任务。

15 在本申请第六方面一实施例中，当时间记录信息用于记录第四任务执行的时间，预设条件包括：第四任务执行的时间大于第一阈值。

在本申请第六方面一实施例中，当时间记录信息记录第五任务在第一调度器的待执行任务队列中的等待时间，预设条件包括：等待时间大于第二阈值。

20 在本申请第六方面一实施例中，第二目标虚拟机为第一目标虚拟机所在主机部署的多个虚拟机中处于空闲态的虚拟机；或者，第二目标虚拟机为主机部署的多个虚拟机中负载小于第三阈值的虚拟机；或者，第二目标虚拟机为主机部署的多个虚拟机中用于执行第四优先级的任务的调度器，其中，第四优先级低于第三优先级。

在本申请第六方面一实施例中，第二目标虚拟机是专用于当时间记录信息满足预设条件时执行第五任务的调度器。

30 本申请第七方面提供一种电子设备，包括：处理器和通信接口。

所述通信接口用于实现所述通信装置与外设的连接通信。

所述处理器用于实现上述第一方面、第二方面或者第三方面任一项所述的方法。

作为一种可能的设计，上述通信装置还包括：存储器。

35 所述存储器用于存储计算机程序，所述处理器执行所述存储器中存储的计算机程序，以使得所述装置执行上述第一方面、第二方面或者第三方面任一项所述的方法。

本申请第八方面提供一种芯片，包括：处理器和通信接口；

所述通信接口用于实现与其他设备通信；

所述处理器用于读取指令以实现如上述第一方面、第二方面或者第三方面任一项所述的方法。

本申请第九方面提供一种计算机程序产品，所述计算机程序产品包括计算机程序代码，当所述计算机程序代码被计算机执行时，使得所述计算机执行如上述第一方面、第二方面或者第三方面任一项所述的方法。

5 附图说明

- 图 1 为本申请应用场景的示意图；
图 2 为一种使用虚拟机结构的云服务器的结构示意图；
图 3 为一种云服务器对虚拟机进行调度的模块示意图；
图 4 为完全公平调度器的工作原理示意图；
10 图 5 为实时调度器的工作原理示意图；
图 6 为本申请提供了一种任务调度方法所应用的云服务器的结构示意图；
图 7 为本申请提供了一种云服务器集群的部署示意图；
图 8 为本申请提供的任务调度方法一实施例的流程示意图；
图 9 为本申请提供的任务调度方法中调度流程的示意图；
15 图 10 为本申请提供的任务调度方法一实施例的流程示意图；
图 11 为本申请提供的任务调度方法中调度流程的示意图；
图 12 为本申请提供的任务调度方法一实施例的流程示意图；
图 13 为本申请提供的任务调度方法一实施例的流程示意图；
图 14 为本申请提供的任务调度方法中调度流程的示意图；
20 图 15 为本申请提供的任务调度装置一实施例的结构示意图；
图 16 为本申请提供的任务调度装置一实施例的结构示意图；
图 17 为本申请提供的任务调度装置一实施例的结构示意图；
图 18 为本申请提供的电子设备一实施例的结构示意图。

25 具体实施方式

图 1 为本申请应用场景的示意图，其中，本申请应用在数据中心的混合部署领域，例如，在图 1 所示的场景中，云计算服务的供应商可以在互联网 2 内设置多个云服务器 3，由云服务器 3 提供计算服务。当用户所使用的终端设备 1 需要一定的计算资源时，可以直接使用、或者向供应商申请、或者向供应商支付一定费用来获得由云服务器 3 提供的云计算服务。例如，用户使用的终端设备 1 在进行基因测序等计算量较大的计算任务时，仅仅依靠终端设备 1 的 CPU 提供的计算能力可能要计算几天时间效率较低，此时用户可以使用云服务器 3 的计算资源（例如 CPU），进行基因测序的计算任务，使得终端设备 1 能够在几分钟或者更短的时间内得到基因测序的计算结果，从而使用户获得了更高的计算效率，同时，使用设置在互联网 2 内的服务器 3 进行计算，35 还能够让用户在不需专门安装、升级其使用的终端设备 1 的情况下，就能够使用更多的计算资源，还减少了用户进行计算的经济成本。由于用户使用的计算资源是由供应商设置在网络侧的云服务器 3 来提供，故这种使用网络资源进行计算的场景又可被称为“云计算”。

而作为云服务器 3 的供应商，同样为了节约资源、减少成本，可以通过虚拟机技

术，在一个云服务器 3 上设置多台虚拟机，分别向不同的用户提供计算资源。其中，虚拟机技术是指通过软件模拟的具有完整硬件系统功能的、执行在各自独立环境中执行的完整计算机系统。例如，图 2 为一种使用虚拟机结构的云服务器的结构示意图，以如图 1 所示的场景中的云服务器 3 为例，供应商在云服务器 3 的系统资源（例如，CPU、内存、硬盘等）的基础上安装操作系统（所述操作系统又可被称为主机操作系统（host OS），例如：Linux 等），在操作系统中可以执行不同的虚拟机。当云服务器 3 向如图 1 所示的三个终端设备 11-13 提供计算服务，终端设备 11 请求使用的计算资源可以占云服务器 3 系统资源中 CPU 计算能力的 30%、终端设备 12 请求使用的计算资源可以占云服务器 3 系统资源中 CPU 计算能力的 50%、终端设备 13 请求使用的计算资源可以占云服务器 3 系统资源中 CPU 计算能力的 20%，供应商将云服务器 3 的系统资源按照三个终端设备所需要的计算量按照比例划分为三部分，在这三部分系统资源中各自执行一个虚拟机，记为虚拟机 1-3。最终，终端设备 11 可以使用云服务器 3 提供的虚拟机 1 进行计算、终端设备 12 可以使用云服务器 3 提供的虚拟机 2 进行计算、终端设备 13 可以使用云服务器 3 提供的虚拟机 3 进行计算，而在云服务器 3 中，每个用户可以像使用实体机一样对各自对应的虚拟机进行操作，每个虚拟机使用其所对应的部分系统资源执行，互不影响。

然而，在实际的使用过程中，由于用户在通过终端设备 1 使用云服务器 3 提供的虚拟机时，通常不会考虑对资源的使用率，是根据所需计算量的最大峰值来申请计算资源，造成了在如图 2 所示的云服务器 3 中，为虚拟机所分配的系统资源中的大部分并不会使用，而经常处于闲置状态，一些业界的研发数据表明，对系统资源中 CPU 的利用率在 20% 以下，而 CPU 的成本占据了云服务器 3 单板硬件成本的 40% 以上，因此云服务器 3 的提供商，为了提高对云服务器 3 的系统资源的利用率、节省在互联网 2 中设置的云服务器 3 的数量，可以将向用户已经分配的资源空闲时，用于向其他用户提供服务。例如，在如图 2 所示的云服务器中，在虚拟机 2 所申请的 50% 的系统资源空闲时，可以分配给虚拟机 4 使用，对于使用虚拟机 2 和虚拟机 4 的用户而言，各自都向提供商申请了云服务器 3 的 50% 的系统资源，而对于提供商一侧可以理解为使用相同的系统资源同时满足了虚拟机 2 和虚拟机 4 所申请的系统资源，从而通过资源共用的方式，提高对云服务器 3 的系统资源的利用效率。

同时，云服务器还可以对执行在各个虚拟机中任务进行调度，从而实现虚拟机资源更加合理的利用。示例性地，图 3 为一种云服务器对虚拟机进行调度的模块示意图，如图 3 示出了在云服务器中设置的调度器，其中，云服务器中可以包括主调度器和周期调度器，主调度器用于对任务进行直接的调度、周期调度器以固定的频率执行，用于对任务进行周期性的调度，二者结合可组成云服务器的核心调度器（core scheduler），又可被称为通用调度器（generic scheduler）。图 3 所示的场景中，以云服务器的核心调度器对云服务器中设置的任一虚拟机进行调度为例，其中，假设核心调度器正在调度的虚拟机中，待执行的任务有任务 1-任务 N。则核心调度器首先通过步骤①调用完全公平调度器（completely fair scheduler，简称：CFS）、RRS（round-robin）或者先入先出（first in, first out，简称：FIFO）等调度器类，每个调度器类可根据不同的调度规则对任务进行调度。则调度器类通过步骤②获取虚

虚拟机待执行的 N 个任务后,按照对应的调度规则确定调度虚拟机执行某个目标任务后,由核心调度器通过步骤③以上下文切换方式指示虚拟机,在步骤④中应执行的目标任务,从而实现对虚拟机执行的任务进行执行先后顺序的调度。

5 更为具体地,现有的调度器类至少可以划分为完全公平调度器(CFS)和实时调度器(RR、FIFO 或者 Deadline 等)两种不同的类型,下面分别进行说明。

图 4 为完全公平调度器的工作原理示意图,其中,当核心调度器调用 CFS 调度器类,对虚拟机待执行的任务 1、任务 2 和任务 3 进行调度时,CFS 调度器类首先将虚拟机的执行时间按照时间轴划分为不同的连续时间片,例如图 4 中的时间片 1、时间片 2... 随后,在每个时间片中,将三个任务按照每个任务的权重分配执行的时间,以保证在连续的时间片组成的时间段内,每个任务都会被执行,保证调度的公平性。例如,CFS 调度器类可以通过公式“时间片长度*任务权重/该任务所在调度队列中所有任务权重之和”计算每个任务在时间片内分配的时间。

15 从图 4 可以看出,完全公平调度器类在对虚拟机执行的多个任务进行调度时,以公平为原则,保证每个任务都可以被执行,但是只能够保证权重较高的任务比权重较低的任务获得更多的时间执行,在调度的不同任务属于不同的优先级(优先级越高对应权重越高、优先级越低对应权重越低)时,无法调度高优先级的任务优先执行、也无法保证高优先级任务能够抢占正在执行的低优先级任务的执行时间。

图 5 为实时调度器的工作原理示意图,其中,当核心调度器调用实时调度器类,对虚拟机待执行的任务进行调度时,可以将所有任务按照优先级顺序加入一个链表中,例如图 5 所示的上下方向排列的链表,向下方向的优先级更高,向上方向的优先级更低。假设此时虚拟机待执行的任务有高优先级的任务 A、中优先级的任务 B1 和 B2、低优先级的任务 C1 和 C2,则 FIFO 调度器类在调度上述任务时,将采用最先进入执行状态的任务首先获取虚拟机的 CPU 等计算资源,并一直占用该资源直到这个任务进入睡眠状态,例如,任务 A 进入队列后一直执行,任务 B1、B2、C1 和 C2 可能都不能获得执行时间,这种低优先级的任务不能获得执行时间的问题又被称为“饥饿问题”。而 RR 调度器类在调度上述任务时,可以通过轮流执行的方式,调度具有相同优先级的任务共享虚拟机的资源,例如,RR 调度器可以调度虚拟机轮流执行虚拟机的调度器中优先级的任务 B1 和 B2 的任务,来保证具有相同优先级任务的公平执行,但是同样可能会导致更低优先级的任务的饥饿问题。

30 从图 5 可以看出,实时调度器类以优先级为基础,优先调度优先级较高的任务执行,以保证任务优先级的要求,但是在调度高优先级任务的同时,会导致低优先级任务的饥饿问题,无法保障低优先级的公平执行需求。

35 综上,现有的调度器类中,完全公平调度器仅能够保证公平,而不能对不同优先级的任务进行区别调度;实时调度器仅对不同优先级的任务进行调度而不能公平调度,导致低优先级任务的饥饿问题。因此,在如图 2 所示的不同虚拟机共用系统资源的虚拟机混合部署的应用场景中,无法简单地使用现有的调度器类在既能够保证优先级高的任务被优先调度、又能够保证低优先级的任务得到公平的调度。因此,如何设计一种调度器,使得调度器能够用于调度混合部署的虚拟机上执行的不同优先级的任务时,还能够满足调度的公平,是本领域亟需解决的技术问题。

因此，本申请提供一种任务调度方法及装置，以使调度器能够在调度虚拟机的待执行的任务时，既能够保证高优先级的任务被优先调度由虚拟机执行、又能够满足虚拟机在执行相同优先级任务调度时的公平。下面以具体地实施例对本申请的技术方案进行详细说明。下面这几个具体的实施例可以相互结合，对于相同或相似的概念或过程可能在某些实施例不再赘述。

图 6 为本申请提供的一种任务调度方法所应用的云服务器的结构示意图，其中，本申请所提供的方法可以由图 6 所示的云服务器中的调度器来执行。具体地，如图 6 所示的云服务器中，可以部署不同类型的虚拟机，所述虚拟机的类型至少包括：按需虚拟机（on-demand virtual machine，简称：OVM）和弹性虚拟机（elastic virtual machine，简称：EVM）。其中，OVM 又可被称为非超卖虚拟机，OVM 的 vCPU 计算能力与物理机 CPU 计算能力基本保持一致，适合处理对计算和时延都敏感的任务；EVM 又可被称为超卖虚拟机或者共享型虚拟机，属于一种可以被抢占的计算实例，适合用于处理批处理和有容错机制的业务场景中的任务。则在如图 6 所示的云服务器中，在云服务器的主机的系统资源上可以安装操作系统，操作系统中执行的调度器可用于对每个虚拟机执行的任务进行调度。同时，为了对不同优先级的任务进行调度，云服务器的操作系统中还设置有虚拟机标识模块，用于对每个虚拟机待执行的任务的优先级进行标注。并且每个虚拟机中还可以各自设置调度器，用于对每个虚拟机待执行的任务进行调度。如图 6 所示的云服务器中以一个 OVM 虚拟机和一个 EVM 虚拟机为例，用以说明云服务器中部署有不同类型的虚拟机，可以理解的是，如图 6 所示的云服务器中还可以部署多个 OVM 虚拟机和多个 EVM 虚拟机，未在图中绘出。而对于每一个虚拟机，不论其类型是 OVM 还是 EVM，都相当于独立执行的计算机，此时虚拟机所使用的云服务器的系统资源例如 CPU，能够以独立的形式执行，在每个虚拟机中的部分 CPU 资源又可被称为 vCPU。

云服务器经过如图 6 所示的部署方式后，由于每个云服务器内可以同时部署 OVM 虚拟机和 EVM 虚拟机，则对于云服务器的供应商，就可以不用在集群中通过不同的云服务器部署不同类型的虚拟机，而是可以实现如图 7 所示的综合部署方式。图 7 为本申请提供的一种云服务器集群的部署示意图，其中，供应商通过集群统一调度平台，对不同区域的云服务器集群进行管理，而每个区域内包括的多个云服务器中，每个服务器上都可以部署不同类型的虚拟机，从而通过更加灵活的部署方式，实现了云服务器集群的部署，不用每个云服务器内仅部署一种类型的虚拟机。此外，每个云服务器上还可以设置动态资源代理，用于统计每个云服务器的资源使用情况，使得向集群统一调度器反馈每个云服务器的资源信息。

特别地，本申请提供的任务调度方法，可以由图 6 或图 7 中每个云服务器内的调度器来执行，虽然每个云服务器内部署有不同类型的虚拟机、每个虚拟机内有自己的调度器，但是由云服务器上设置的调度器对每个虚拟机执行的任务的调度时，可以实现对每个 OVM 虚拟机或者 EVM 虚拟机所执行的任务保证公性的前提下实现 OVM 的按需分配。其中，每个虚拟机内的调度器存储并调度该虚拟机的待执行任务队列，云服务器上的调度器可以通过对虚拟机内调度器的待执行任务队列进行调整的方式，实现对虚拟机待执行任务的调度。

下面以具体的实施例，详细介绍本申请提供的任务调度方法，下面这几个具体的实施例可以相互结合，并由同一个执行主体来同时实现其中的全部或部分，对于相同或相似的概念或过程可能在某些实施例不再赘述。本申请提供的任务调度方法可以由如图 6 或图 7 所示的云服务器中的调度器执行，或者，也可以由除图 6 或图 7 中其他

5 的云服务器中的其他装置执行，本申请仅使用图 6 所示的云服务器中的调度器作为执行主体作为示例，对任务调度方法的流程进行说明，而非对云服务器结构进行限定。

实施例一

本申请实施例一提供的任务调度方法，可用于调度器在调度虚拟机中的任务时，在满足任务优先级要求的同时，实现任务的公平调度，也就是在 CFS 调度器的基础上

10 实现了不同优先级任务的调度，可以理解为一种增强的 CFS 调度方法。

图 8 为本申请提供的任务调度方法一实施例的流程示意图，如图 8 所示的任务调度方法可以由如图 6 所示的云服务器中主机的操作系统内的调度器执行，应用在调度器对云服务器中第一目标虚拟机待执行的任务进行调度的场景，所述第一目标虚拟机可以是云服务器中部署的任一 OVM 虚拟机或者任一 EVM 虚拟机，并将所述第一目标虚拟机中的调度器记为第一调度器，第一调度器通过任务队列，存储并调度第一目标虚拟机待执行的

15 的任务，作为执行主体的调度器可以与第一调度器连接，获取第一调度器中存储的任务队列，以及调整第一调度器中的任务待执行任务，实现调整第一目标虚拟机待执行的任務。具体地，则本实施例提供的任务调度方法包括：

S101：调度器获取第一目标虚拟机待执行的所有任务的优先级。

具体地，调度器首先通过 S101 从第一目标虚拟机的第一调度器中，获取第一目标虚拟机当前待执行的所有任务，以及所有任务的优先级。其中，所述优先级可以是提前划分好的、也可以是由调度器根据任务的类型划分的。例如，对于基因测序、batch 业务等对计算时延不敏感的业务，对应较低的优先级；对于在线视频内容实时渲染等对计算时延敏感的业务，对应较高的优先级。在如图 6 所示的云服务器中，所有任务的

20 的优先级还可以由主机上设置的虚拟机标识模块进行统一管理。

下面结合图 9 的示例进行说明，图 9 为本申请提供的任务调度方法中调度流程的示意图，其中，假设第一调度器在调度的第一目标虚拟机时，第一调度器的任务队列中，存储有第一目标虚拟机当前待执行的 N 个任务分别为任务 1、任务 2……任务 N， $N > 1$ 。则在 S101 中，作为执行主体的云服务器上的调度器为了确定第一目标虚拟机的待执行任务，可以首先从第一目标虚拟机的第一调度器中获取这 N 个任务，以及每个任务对应的优先级。例如，在 N 个任务中，第一优先级的任务有任务 1、任务 2 和任务 3；第二优先级的任务有任务 4 和任务 5……。本申请实施例对第一目标虚拟机待执行任务的优先级的数量不做限定，并以第一优先级、第二优先级……的排列顺序中优先级由高至低为例进行说明。

25

S102：调度器确定所有任务中优先级最高的至少一个任务。

随后，调度器从 S101 中获取的所有 N 个任务中，确定优先级最高的至少一个任务。例如，在如图 9 所示的示例中，调度器确定优先级最高的第一优先级对应的任务有：任务 1、任务 2 和任务 3。

可选地，在 S102 第一种可能的实现方式中，调度器可以通过遍历第一调度器存储

的待执行任务队列上所有任务优先级的方式，从待执行任务队列上确定所有任务的优先级，并确定出所有任务中优先级最高的至少一个任务。可选地，所述队列可以是链表、数组、树类结构等，当所述待执行任务队列为树类结构中的红黑树时，调度器可以遍历整个红黑树，获取红黑树中优先级最高的至少一个任务。

5 或者，在 S102 第二种可能的实现方式中，调度器还可以在专门的存储空间中，存储优先级记录信息，通过优先级记录信息对加入第一调度器的待执行任务队列中的每个任务的优先级进行记录，例如，调度器可通过数组记录不同任务与优先级之间的对应关系，其形式可以为“任务 1-第一优先级、任务 2-第一优先级…”等。则在 S102 中调度器可以通过存储空间中所记录的优先级记录信息，确定第一调度器的待执行任
10 务队列中优先级最高的至少一个任务。

又或者，在 S102 第三种可能的实现方式中，第一调度器中还可以设置多个待执行任务队列，每个待执行任务队列对应一个优先级，所述队列同样可以是链表、数组、树类结构等，当所述待执行任务队列为树类结构中的红黑树时，调度器可以通过不同的红黑树表示不同优先级对应的待执行任务队列。例如，在如图 9 所示的示例中，可
15 以通过第一红黑树记录第一优先级对应的任务 1、任务 2 和任务 3，通过第二红黑树记录第二优先级对应的任务 4 和任务 5。则调度器在 S102 中可以直接从优先级最高的第一优先级对应的第一红黑树中，获取优先级最高的至少一个任务，从而 S102 中提高获取优先级最高的至少一个任务的效率。

S103：调度器确定至少一个任务中虚拟运行时间最小的任务，记为第一任务。

20 随后，在 S103 中，调度器进一步确定 S102 中所获取的至少一个任务中，虚拟运行时间最小的第一任务。其中，所述虚拟运行之间又可被称为“vruntime”，其中，每个任务的虚拟运行时间=任务的实际执行时间*1024/待执行的所有任务的权重之和，例如，对于第一调度器中待执行的任务 1、任务 2 和任务 3，权重分别为 1、2 和 3，则待执行所有任务的权重之和为 1+2+3=6。调度器通过每个任务的虚拟运行时间来衡
25 量哪个任务最值得被调度，其中，在 CFS 调度器中，所有待执行任务可以通过一棵虚拟运行时间为键值的红黑树表示，则虚拟运行时间越小的进程越靠近整个红黑树的最左端，第一调度器每次控制第一目标虚拟机执行位于红黑树最左端的任务，该进程的虚拟运行时间最小。因此，在本实施例中，当调度器通过 S102 确定出优先级最高的至少一个任务之后，可以在优先级最高的至少一个任务中确定虚拟运行时间最小的第一
30 任务。例如，在如图 9 所示的示例中，调度器确定出第一优先级的任务 1、任务 2 和任务 3 之后，从任务 1 的第一虚拟运行时间、任务 2 的第二虚拟运行时间和任务 3 的第三 vruntime 中确定任务 2 对应的第二虚拟运行时间最小，则将任务 2 作为第一任务。

S104：调度器控制第一调度器调度第一目标虚拟机执行第一任务。

35 最终，在 S104 中，调度器通过控制第一调度器的方式，使得第一调度器调整其任务队列，从而使得第一调度器调度第一目标虚拟机执行 S103 中所确定的第一任务。可选地，调度器可以通过上下文切换的方式，调整第一调度器的待执行任务队列中任务的顺序，实现控制第一目标虚拟机的第一调度器从待执行任务队列中首先获取第一任务并由第一目标虚拟机的处理器（vCPU）执行。

可以理解的是，如图 8-9 所示的实施例中，以第一调度器的待执行任务队列上有

多个优先级的任务为例，若第一调度器的待执行任务队列上只有一个优先级的任务，则在 S101 之后，调度器可以直接执行 S103，即，从所获取的所有任务中确定虚拟运行时间最小的第一任务，并通过 S104 控制第一目标虚拟机执行第一任务。

此外，本实施例中调度器在第一任务执行完之后，将重复执行如图 8 所示的任务调度方法 S101，并继续通过第一调度器调度第一目标虚拟机执行其待执行任务队列中优先级最高的至少一个任务中虚拟执行执行时间最小的任务，从而保证了当前第一目标虚拟机执行的是优先级最高的任务，并且对于相同优先级的任务能够得到公平的调度。因此在如图 9 所示的示例中，按照这个顺序，第一目标虚拟机将依次执行第一优先级、第二优先级……的任务，并且在调度每个优先级对应的相同任务时，通过虚拟运行时间的选择相当于采用了 CFS 调度器的完全公平的调度策略，实现了 CFS 调度策略。

综上，本实施例提供的任务调度方法中，调度器在调度第一目标虚拟机执行任务时，首先确定第一目标虚拟机待执行任务队列中优先级最高的至少一个任务，保证了优先级最高的任务优先执行；随后对于最高优先级的至少一个任务选择其中虚拟运行时间最小的第一任务执行，保证了相同优先级任务在调度时的公平。因此，本实施例提供的任务调度方法与图 4 所示的现有技术中完全公平调度器相比，能够调度高优先级的任务优先执行；与图 5 所示的实时调度器相比，还能够保证相同优先级的任务得到公平的调度，从而实现了调度器能够在调度第一目标虚拟机上的任务时，既能够保证高优先级的任务被优先调度、又能够满足相同优先级任务调度时的公平，能够应用与如图 6 或者图 7 所示的虚拟机混合部署调度的场景中。

实施例二

本申请实施例二提供的任务调度方法，可用于调度器在调度任务时，当第一目标虚拟机的第一调度器中，待执行任务队列中任务的优先级发生了调整，可以对第一调度器当前调度第一目标虚拟机执行的任务进行抢占，也就是实现调度器对第一目标虚拟机的任务进行动态的调度，以满足调度器对多种不同类型的任务的调度需求。

图 10 为本申请提供的任务调度方法一实施例的流程示意图，如图 10 所示的任务调度方法可以由如图 6 所示的云服务器中主机的操作系统内的调度器执行，应用在调度器对云服务器中第一目标虚拟机待执行的任务进行调度的场景，所述第一目标虚拟机可以是云服务器中部署的任一 OVM 虚拟机或者任一 EVM 虚拟机，并将所述第一目标虚拟机中的调度器记为第一调度器，第一调度器通过任务队列，存储并调度第一目标虚拟机待执行的任务，则该方法包括：

S201：当调度器确定第一目标虚拟机的待执行任务队列中的第二任务由第二优先级调整为第一优先级，获取第一目标虚拟机待执行的与第一优先级相同的至少一个任务中，最小的虚拟运行时间。

具体地，当第一调度器调度第一目标虚拟机执行任务（将第一虚拟机正在执行的任务记为第三任务）时，若此时第一目标虚拟机在第一调度器的待执行任务队列中，第二任务由第二优先级调整为第一优先级，或者在第一调度器的待执行任务队列中加入了第二任务，则调度器通过 S201 确定第一目标虚拟机待执行的第二任务的优先级发生了变化，并确定与第一优先级相同的至少一个任务中最小的虚拟运行时间。

S202: 调度器根据 S201 中确定的最小虚拟运行时间, 将第二任务对应的第一虚拟运行时间调整为第二虚拟运行时间。

下面结合图 11 的示例进行说明, 图 11 为本申请提供的任务调度方法中调度流程的示意图, 其中, 假设第一调度器调度第一目标虚拟机的当前待执行的 N 个任务分别为任务 1、任务 2……任务 N, $N > 1$ 。则在 S201 中, 作为执行主体的调度器首先获取其中优先级由第二优先级调整为第一优先级的任务, 记为第二任务, 或者还可以获取新加入的第 N+1 个任务记为第二任务。并进一步确定第二任务对应的第一优先级, 以及第二任务对应的第一虚拟运行时间。或者, 可选地, 所述第二任务还可以是从睡眠状态切换到唤醒状态后, 加入第一调度器待执行任务队列的任务, 则调度器可以直接从待执行任务队列中获取新加入的第二任务的第二虚拟运行时间。

随后, 调度器将 S201 中获取的第二任务的第一虚拟运行时间调整为第二虚拟运行时间, 其中, 对第二任务的虚拟运行时间进行调整的目的是使得第二任务能够得到执行时间, 故第二任务的虚拟运行时间的调整应与当前执行的第三任务或者与第二任务优先级相同的任务有关。

具体地, 本实施例的具体实现方式可以参照如图 12 所示的 S2021-S2023, 其中, 图 12 为本申请提供的任务调度方法一实施例的流程示意图, 如图 12 所示的方法在如图 11 所示方法的基础上, 调度器在获取第二任务后, 通过 S2021 确定第一调度器的待执行任务队列中是否存在其他与第二任务优先级相同的、同对应于第一优先级的任务。若第一调度器的待执行任务队列中没有其他第一优先级的任务, 则执行 S2022; 若第一调度器的待执行任务队列中有其他第一优先级的任务, 则执行 S2023。

在 S2022 中, 若待执行任务队列中, 所有的 N 个任务都不是第一优先级的任务, 则调度器参考第一调度器正在调度第一目标虚拟机执行的第三任务的虚拟运行时间, 对第二任务的第一虚拟运行时间进行调整。具体地, 调整原则可以使得第一虚拟运行时间调整后的第二虚拟运行时间略大于第三任务的虚拟运行时间, 例如, 当第一虚拟运行时间小于第三任务的虚拟运行时间, 则调度器将第一虚拟运行时间加上一个第一变量后, 得到第二虚拟运行时间, 使得第二虚拟运行时间大于第三任务的虚拟运行时间。所述变量可以根据实际情况实时进行调整。

在 S2023 中, 若第一调度器的待执行任务队列中存在其他第一优先级的任务, 示例性地, 在图 11 所示的场景中, 当确定第一调度器调度目标虚拟机所运行的任务 1、任务 2 和任务 3 与第二任务的优先级相同, 都对应第一优先级, 则调度器具体根据所确定的任务 1-任务 3 的虚拟运行时间中最小的虚拟运行时间对第二虚拟运行时间进行调整。具体地, 调整原则可以使得第一虚拟运行时间调整后的第二虚拟运行时间略大于第三任务的虚拟运行时间, 例如, 当第一虚拟运行时间小于第三任务的虚拟运行时间, 则调度器将第一虚拟运行时间加上一个第一变量后, 得到第二虚拟运行时间, 使得第二虚拟运行时间大于任务 1-任务 3 的虚拟运行时间中最小的虚拟运行时间。所述第一变量可以根据实际情况进行调整。

S203: 当第一优先级高于第一目标虚拟机正在执行的第三任务的优先级, 调度器控制第一目标虚拟机执行第二任务, 此时, 第二任务对应于第二虚拟运行时间。

最终, 当调度器通过 S202 对新加入待执行任务队列的第二任务的虚拟运行时间进

行调整后,可以控制第一调度器,使得第一调度器调度第一目标虚拟机执行第二任务。特别地,在本步骤中,还加入了调度器对第一调度器当前正在执行任务与第二任务优先级的比对,使得第二任务的第一优先级高于正在执行的第三任务的优先级时,实现第二任务对第三任务的抢占。

5 或者,可选地,在如图 10 所示的实施例中,S201 另一种可能的实现方式中,第二任务还可以是原本就存在第一调度器待执行任务队列中,则调度器在 S201 中获取第二任务后,进一步根据用户发出的指示信息,确定将第二任务对应的优先级由第二优先级调整为第一优先级。随后,在 S202 中,调度器具体根据第一调度器的待执行任务队列中对应第二优先级的至少一个任务的虚拟运行时间中最小的虚拟运行时间,对第一虚拟运行时间进行调整为第二虚拟运行时间。其中,假设第二任务在调整优先级之前,第二优先级对应的至少一个任务中,对应的第一虚拟运行时间为 T_a ,而第二优先级对应的至少一个任务中虚拟运行时间的最小值记为 $\min A$;第二任务在调整优先级之后,第一优先级对应的至少一个任务中虚拟运行时间的最小值记为 $\min B$,则调度器通过公式:第二虚拟运行时间 = $(T_a - \min A) + \min B$,来将第一虚拟运行时间调整为第二虚拟运行时间,并使得调整后的第二虚拟运行时间在第一优先级对应的任务中,略大于第一优先级的至少一个任务中虚拟运行时间的最小值。需要说明的是,上述公式仅为
10 15 示例性的说明,在实际调整过程中调度器并不限于使用这一种方式进行调整。

综上,本实施例提供的任务调度方法中,调度器在第一目标虚拟机执行第三任务的同时,若获取了对应第一优先级的第二任务,则根据第一目标虚拟机的待执行任务队列中,对应第一优先级的至少一个任务的虚拟运行时间中最小的虚拟运行时间或者根据第三任务的虚拟运行时间,对第二任务的第一虚拟运行时间进行调整为第二虚拟运行时间。最终,在第二任务的第一优先级高于第三任务的优先级时,控制第一目标虚拟机执行第二任务,从而实现了新加入第一目标虚拟机的待执行任务队列的第二任务的抢占执行。因此,本实施例提供的任务调度方法克服了如图 4 所示的完全公平调度器以及图 5 所示的实时调度器无法提供抢占的不足,使得调度器能够在第一目标虚拟机上运行多种优先级的任务时,能够实现高优先级的任务对低优先级任务的抢占运行,能够应用与如图 6 或者图 7 所示的虚拟机混合部署调度的场景中。
20 25

实施例三

本申请实施例三提供的任务调度方法,可用于调度器在调度第一目标虚拟机中的任务时,当第一目标虚拟机的第一调度器所存储的待运行任务队列中,存在低优先级的任务一直得不到执行的情况,可以将低优先级的任务主动迁移到其他虚拟机(记为第二目标虚拟机)上执行,从而防止由于第一虚拟机执行多种优先级的任务时,优先调度高优先级的任务进而导致的低优先级任务的饥饿问题。
30

图 13 为本申请提供的任务调度方法一实施例的流程示意图,如图 13 所示的任务调度方法可以由如图 6 所示的云服务器中主机的操作系统内的调度器执行,应用在调度器对云服务器中第一目标虚拟机待执行的任务进行调度的场景,所述第一目标虚拟机可以是云服务器中部署的任一 OVM 虚拟机或者任一 EVM 虚拟机,并将所述第一目标虚拟机中的调度器记为第一调度器,第一调度器通过任务队列,存储并调度第一目标虚拟机待执行的任务,则该方法包括:
35

S301: 调度器获取第一目标虚拟机执行第四任务时, 第一调度器的时间记录信息。

具体地, 当第一目标虚拟机执行任务 (将第一虚拟机正在执行的任务记为第四任务) 时, 调度器还同时通过时间记录信息对第一目标虚拟机的第一调度器的任务队列中的任务进行记录。所述 S301 可以由调度器在时钟中断时执行, 或者在调度器调度切换时执行, 又或者通过调度器设置定时器的方式执行。其中, 记第四任务对应第三优先级, 同时, 在第一目标虚拟机的第一调度器中, 存储的待运行任务队列中至少还包括第五任务, 第五任务对应的优先级低于所述第三优先级。

5 可选地, 所述时间记录信息可以存储在调度器内的存储空间中, 或者, 存储在第一目标虚拟机内的存储空间中。在一种可能的实现方式中, 所述时间记录信息用于记录第四任务的执行时间, 或者, 在另一种可能的实现方式中, 所述时间记录信息用于记录第五任务在待执行任务队列中等待的时间。

下面结合图 14 的示例进行说明, 图 14 为本申请提供的任务调度方法中调度流程的示意图, 其中, 假设第一调度器当前待执行的 N 个任务分别为任务 1、任务 2……任务 N, $N > 1$, 并且此时, 第一目标虚拟机执行待执行任务队列的 N 个任务中的第四任务。则对于调度器在 S301 中, 可以获取当前时刻的时间记录信息。

10 S302: 当时间记录信息满足预设条件, 则调度器控制第一目标虚拟机的第一调度器, 将第一调度器中存储的第五任务发送给第二目标虚拟机的第二调度器, 使得第二调度器控制第二目标虚拟机执行第五任务。

具体地, 调度器在通过 S301 获取时间记录信息之后, 根据时间记录信息确定调度第五任务到第二目标虚拟机执行, 例如, 当所述时间记录信息用于记录第一目标虚拟机已经执行第四任务的执行时间, 则预设条件可以是第四任务的执行时间小于第一阈值 (1s)。例如, 当第四任务已经执行了 1 秒, 由于第四任务的优先级较高, 处于较低优先级的第五任务一直得不到执行, 会导致第五任务的饥饿问题, 因此, 调度器可以确定第四任务执行时间大于第一阈值而满足预设条件后, 调度由第二目标虚拟机执行第五任务。

20 又例如, 当所述时间记录信息用于记录第五任务在第一目标虚拟机的待执行任务队列中的等待时间, 则预设条件可以是第五任务的等待时间大于第二阈值。同样为了防止第五任务的饥饿问题, 调度器可以确定第五任务的等待时间大于第二阈值而满足预设条件后, 调度由第二目标虚拟机执行第五任务。

30 进一步地, 所述第二目标虚拟机可以是云服务器中的任一个虚拟机, 则调度器在 S302 中控制第二目标虚拟机执行第五任务之前, 还可以在云服务器中确定第二目标虚拟机。例如, 所述第二目标虚拟机可以是云服务器中当前处于空闲态的虚拟机, 使得调度器控制第二目标虚拟机执行第五任务时, 不会影响第二目标虚拟机运行的其他任务; 或者, 第二目标虚拟机可以是云服务器中当前负载小于第三阈值的虚拟机; 使得调度器控制第二目标虚拟机执行第五任务时, 减少对第二目标虚拟机运行的其他任务的影响; 或者, 第二目标虚拟机还可以是云服务器部署的专用于执行第五任务对应的优先级的虚拟机, 也就是云服务器中将不同优先级的任务分配给不同的虚拟机执行, 则调度器在 S302 中调度第五任务时, 选择与第五任务的第四优先级对应的虚拟机作为所述第二目标虚拟机。

35

可选地，调度器在调度第五任务时，还可以按照如下顺序确定第二目标虚拟机：优先选择云服务器中空闲态的虚拟机、其次选择云服务器中虚拟机的调度队列上的任务优先级较低且负载较轻的虚拟机、再次选择云服务器中虚拟机的调度队列上的任务优先级较低但负载较重的虚拟机/服务器中虚拟机的调度队列上的任务优先级较高但负载较轻的虚拟机。

可选地，第二的目标虚拟机还可以是云服务器中专门设置的虚拟机，该虚拟机专用于在如图 14 所示的防止饥饿问题的场景中，执行调度器从其他虚拟机调度的任务，则在 S302 中，调度器可以直接控制第二目标虚拟机执行第五任务。

5 综上所述，本实施例提供的任务调度方法中，调度器在第一目标虚拟机执行第四任务时，还获取时间记录信息，用于对优先级较低的第五任务是否出现饥饿问题进行判断，并在时间记录信息满足预设条件之后，控制云服务器中第二目标虚拟机执行第五任务，从而实现了第五任务在不同目标虚拟机之间的调度，使得第五任务不会由于优先级较低而一直停留在第一目标虚拟机的待执行队列中，防止饥饿问题的发生。

10 在前述实施例中，对本申请实施例提供的任务调度方法进行了介绍，而为了实现上述本申请实施例提供的任务调度方法中的各功能，作为执行主体的调度器可以包括硬件结构和/或软件模块，以硬件结构、软件模块、或硬件结构加软件模块的形式来实现上述各功能。上述各功能中的某个功能以硬件结构、软件模块、还是硬件结构加软件模块的方式来执行，取决于技术方案的特定应用和设计约束条件。

20 例如，图 15 为本申请提供的任务调度装置一实施例的结构示意图，如图 15 所示的任务调度装置可用于执行如图 8 所示的任务调度方法，该装置包括：获取模块 1501、优先级确定模块 1502、虚拟运行时间确定模块 1503 和控制模块 1504。其中，获取模块 1501 用于获取第一目标虚拟机待执行的多个任务的优先级；优先级确定模块 1502 用于确定多个任务中优先级最高的至少一个任务；虚拟运行时间确定模块 1503 用于从至少一个任务中，确定虚拟运行时间最小的第一任务；控制模块 1504 用于控制第一目标虚拟机执行第一任务。

可选地，获取模块 1501 具体用于，根据第一目标虚拟机中，第一调度器所存储的待执行任务队列，确定第一目标虚拟机待执行的多个任务的优先级。

可选地，获取模块 1501 具体用于，遍历第一调度器所存储的待执行任务队列上所有任务的优先级，从待执行任务的队列中确定优先级最高的至少一个任务。

30 可选地，获取模块 1501 具体用于，通过第一调度器中存储的优先级记录信息，确定多个任务中优先级最高的至少一个任务；其中，优先级记录信息用于记录第一调度器待执行的队列上所有任务的优先级。

可选地，获取模块 1501 具体用于，根据第一调度器中存储的最高优先级对应的队列中的任务，确定优先级最高的至少一个任务；其中，第一调度器待执行任务通过多个队列存储待执行的任务，每个队列对应不同的优先级。

35 图 16 为本申请提供的任务调度装置一实施例的结构示意图，如图 16 所示的任务调度装置可用于执行如图 10 所示的任务调度方法，该装置包括：获取模块 1601、虚拟运行时间调整模块 1602 和控制模块 1603。其中，获取模块 1601 用于当第一目标虚拟机待执行的第二任务的优先级由第二优先级调整为第一优先级，获取第一目标虚拟

机待执行的与第一优先级相同的至少一个任务中，至少一个任务对应的最小的虚拟运行时间；虚拟运行时间调整模块 1602 用于根据最小的虚拟运行时间，将第二任务对应的第一虚拟运行时间调整为第二虚拟运行时间；控制模块 1603 用于，当第一优先级高于第一目标虚拟机正在执行的第三任务的优先级，控制第一目标虚拟机执行第二任务。

5 可选地，第二虚拟运行时间大于第一目标虚拟机的第一调度器存储的待执行任务队列中，对应第一优先级的至少一个任务的虚拟运行时间中最小的虚拟运行时间。

可选地，所述任务调度装置还包括：接收模块 1604，用于接收指示信息，并根据指示信息，确定第二任务的优先级由第二优先级调整为第一优先级。

10 可选地，虚拟时间调整模块 1602 具体用于，将第一虚拟运行时间加上最小的虚拟运行时间后，再减去待执行任务队列中对应第一优先级的至少一个任务的虚拟运行时间中最小的虚拟运行时间，得到第二虚拟运行时间。

图 17 为本申请提供的任务调度装置一实施例的结构示意图，如图 17 所示的任务调度装置可用于执行如图 13 所示的任务调度方法，该装置包括：获取模块 1701 和控制模块 1702。其中，获取模块 1701 用于获取第一目标虚拟机的第一调度器存储的时间记录信息；其中，时间记录信息用于记录第四任务执行的时间，第四任务对应第三优先级，或者，时间记录信息用于记录第五任务在第一调度器存储的待执行任务队列中的等待时间，第五任务对应的优先级低于第三优先级；控制模块 1702 用于当时间记录信息满足预设条件，控制第一目标虚拟机的第一调度器将第五任务发送到第二目标虚拟机的第二调度器，使第二目标虚拟机执行第五任务。

20 可选地，当时间记录信息用于记录第四任务执行的时间，预设条件包括：第四任务执行的时间大于第一阈值。

可选地，当时间记录信息记录第五任务在第一调度器的待执行任务队列中的等待时间，预设条件包括：等待时间大于第二阈值。

25 可选地，第二目标虚拟机为第一目标虚拟机所在主机部署的多个虚拟机中处于空闲态的虚拟机；或者，第二目标虚拟机为主机部署的多个虚拟机中负载小于第三阈值的虚拟机；或者，第二目标虚拟机为主机部署的多个虚拟机中用于执行第四优先级的任务的调度器，其中，第四优先级低于第三优先级。

可选地，第二目标虚拟机是专用于当时间记录信息满足预设条件时执行第五任务的调度器。

30 需要说明的是，应理解以上装置的各个模块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。且这些模块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；还可以部分模块通过处理元件调用软件的形式实现，部分模块通过硬件的形式实现。例如，确定模块可以为单独设立的处理元件，也可以集成在上述装置的某一个芯片中实现，此外，也可以以程序代码的形式存储于上述装置的存储器中，由上述装置的某一个处理元件调用并执行以上确定模块的功能。其它模块的实现与之类似。此外这些模块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可以是一种集成电路，具有信号的处理能力。在实现过程中，上述方法的各步骤或以上各个模块可以通过处理器元件中的硬件的集成逻辑电路或者软件形式的指令完成。

例如，以上这些模块可以是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 signal processor, DSP），或，一个或者多个现场可编程门阵列（field programmable gate array, FPGA）等。再如，当以上某个模块通过处理元件调度程序代码的形式实现时，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代码的处理器。再如，这些模块可以集成在一起，以片上系统（system-on-a-chip, SOC）的形式实现。

在上述实施例中，可以全部或部分地通过软件、硬件、固件或者其任意组合来实现。当使用软件实现时，可以全部或部分地以计算机程序产品的形式实现。所述计算机程序产品包括一个或多个计算机指令。在计算机上加载和执行所述计算机程序指令时，全部或部分地产生按照本申请实施例所述的流程或功能。所述计算机可以是通用计算机、专用计算机、计算机网络、或者其他可编程装置。所述计算机指令可以存储在计算机可读存储介质中，或者从一个计算机可读存储介质向另一个计算机可读存储介质传输，例如，所述计算机指令可以从一个网站站点、计算机、服务器或数据中心通过有线（例如同轴电缆、光纤、数字用户线（DSL））或无线（例如红外、无线、微波等）方式向另一个网站站点、计算机、服务器或数据中心进行传输。所述计算机可读存储介质可以是计算机能够存取的任何可用介质或者是包含一个或多个可用介质集成的服务器、数据中心等数据存储设备。所述可用介质可以是磁性介质，（例如，软盘、硬盘、磁带）、光介质（例如，DVD）、或者半导体介质（例如固态硬盘solid state disk (SSD)）等。

图 18 为本申请提供的电子设备一实施例的结构示意图，该电子设备可作为本申请前述实施例中任一所述的调度器，并执行调度器所执行的任务调度方法。如图 18 所示，该通信装置 1100 可以包括：处理器 111（例如 CPU）、收发器 113；其中，收发器 113 耦合至处理器 111，处理器 111 控制收发器 113 的收发动作。可选的，所述通信装置 1100 还包括存储器 112，存储器 112 中可以存储各种指令，以用于完成各种处理功能以及实现本申请实施例中调度器执行的方法步骤。

可选的，本申请实施例涉及的电子设备还可以包括：电源 114、系统总线 115 以及通信接口 116。收发器 113 可以集成在电子设备的收发信机中，也可以为电子设备上独立的收发天线。系统总线 115 用于实现元件之间的通信连接。上述通信接口 116 用于实现电子设备与其他外设之间进行连接通信。

在本申请实施例中，上述处理器 111 用于与存储器 112 耦合，读取并执行存储器 112 中的指令，以实现上述方法实施例中调度器执行的方法步骤。收发器 113 与处理器 111 耦合，由处理器 111 控制收发器 113 进行消息收发，其实现原理和技术效果类似，在此不再赘述。

该图 18 中提到的系统总线可以是外设部件互连标准（peripheral component interconnect, PCI）总线或扩展工业标准结构（extended industry standard architecture, EISA）总线等。所述系统总线可以分为地址总线、数据总线、控制总线等。为便于表示，图中仅用一条粗线表示，但并不表示仅有一根总线或一种类型的总线。通信接口用于实现数据库访问装置与其他设备（例如客户端、读写库和只读库）之间的通信。存储器可能包含 RAM，也可能还包括非易失性存储器（non-volatile memory），例如

至少一个磁盘存储器。

该图18中提到的处理器可以是通用处理器,包括中央处理器CPU、网络处理器(network processor, NP)等;还可以是数字信号处理器DSP、专用集成电路ASIC、现场可编程门阵列FPGA或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件。

5 可选的,本申请实施例还提供一种可读存储介质,所述存储介质中存储有指令,当其在计算机上运行时,使得计算机执行如本申请前述任一实施例中由调度器所执行的方法。

可选的,本申请实施例还提供一种运行指令的芯片,所述芯片用于执行如本申请前述任一实施例中由调度器所执行的方法。

10 本申请实施例还提供一种程序产品,所述程序产品包括计算机程序,所述计算机程序存储在存储介质中,至少一个处理器可以从所述存储介质读取所述计算机程序,所述至少一个处理器执行所述计算机程序时可实现如本申请前述任一实施例中由调度器所执行的方法。

15 在本申请实施例中,“至少一个”是指一个或者多个,“多个”是指两个或两个以上。“和/或”,描述关联对象的关联关系,表示可以存在三种关系,例如,A和/或B,可以表示:单独存在A,同时存在A和B,单独存在B的情况,其中A,B可以是单数或者复数。字符“/”一般表示前后关联对象是一种“或”的关系;在公式中,字符“/”,表示前后关联对象是一种“相除”的关系。“以下至少一项(个)”或其类似表达,是指的这些项中的任意组合,包括
20 单项(个)或复数项(个)的任意组合。例如,a,b,或c中的至少一项(个),可以表示:a,b,c,a-b,a-c,b-c,或a-b-c,其中,a,b,c可以是单个,也可以是多个。同时,在本申请实施例中涉及的各种数字编号仅为描述方便进行的区分,并不用来限制本申请实施例的范围。并且,在本发明的实施例中,上述各过程的序号的大小并不意味着执行顺序的先后,各过程的执行顺序应以其功能和内在逻辑确定,而不应对本申请实施例的实施过程构成任何限定。

25 最后应说明的是:以上各实施例仅用以说明本发明的技术方案,而非对其限制;尽管参照前述各实施例对本发明进行了详细的说明,本领域的普通技术人员应当理解:其依然可以对前述各实施例所记载的技术方案进行修改,或者对其中部分或者全部技术特征进行等同替换;而这些修改或者替换,并不使相应技术方案的本质脱离本发明各实施例技术方案的范围。

权 利 要 求 书

- 1、一种任务调度方法，其特征在于，包括：
获取第一目标虚拟机待执行的多个任务的优先级；
确定所述多个任务中优先级最高的至少一个任务；
5 从所述至少一个任务中，确定虚拟运行时间最小的第一任务；
控制所述第一目标虚拟机执行所述第一任务。
- 2、根据权利要求1所述的方法，其特征在于，所述获取第一目标虚拟机待执行的多个任务的优先级，包括：
根据所述第一目标虚拟机中，第一调度器所存储的待执行任务队列，确定所述第
10 一目标虚拟机待执行的多个任务的优先级。
- 3、根据权利要求2所述的方法，其特征在于，所述确定所述多个任务中优先级最高的至少一个任务，包括：
遍历所述第一调度器所存储的待执行任务队列上所有任务的优先级，从所述待执行任务的队列中确定所述优先级最高的至少一个任务。
- 15 4、根据权利要求2所述的方法，其特征在于，所述确定所述多个任务中优先级最高的至少一个任务，包括：
通过所述第一调度器中存储的优先级记录信息，确定所述多个任务中优先级最高的至少一个任务；其中，所述优先级记录信息用于记录所述第一调度器待执行的队列上所有任务的优先级。
- 20 5、根据权利要求2所述的方法，其特征在于，所述确定所述多个任务中优先级最高的至少一个任务，包括：
根据所述第一调度器中存储的最高优先级对应的队列中的任务，确定所述优先级最高的至少一个任务；其中，所述第一调度器待执行任务通过多个队列存储待执行的任务，每个队列对应不同的优先级。
- 25 6、一种任务优先级调整方法，其特征在于，包括：
当第一目标虚拟机待执行的第二任务的优先级由第二优先级调整为第一优先级，获取所述第一目标虚拟机待执行的与所述第一优先级相同的至少一个任务中，所述至少一个任务对应的最小的虚拟运行时间；
根据所述最小的虚拟运行时间，将所述第二任务对应的第一虚拟运行时间调整为
30 第二虚拟运行时间；
当所述第一优先级高于所述第一目标虚拟机正在执行的第三任务的优先级，控制所述第一目标虚拟机执行所述第二任务。
- 7、根据权利要求6所述的方法，其特征在于，
所述第二虚拟运行时间大于所述第一目标虚拟机的第一调度器存储的待执行任务
35 队列中，对应所述第一优先级的至少一个任务的虚拟运行时间中最小的虚拟运行时间。
- 8、根据权利要求6所述的方法，其特征在于，所述获取所述第一目标虚拟机待执行的与所述第一优先级相同的至少一个任务中，所述至少一个任务对应的最小的虚拟运行时间之前，还包括：
接收指示信息，并根据所述指示信息，确定所述第二任务的优先级由第二优先级

调整为第一优先级。

9、根据权利要求 6-8 任一项所述的方法，其特征在于，所述根据所述最小的虚拟运行时间，将所述第二任务对应的第一虚拟运行时间调整为第二虚拟运行时间，包括：

5 将所述第一虚拟运行时间加上所述最小的虚拟运行时间后，再减去所述待执行任务队列中对应所述第一优先级的至少一个任务的虚拟运行时间中最小的虚拟运行时间，得到所述第二虚拟运行时间。

10、一种任务调度方法，其特征在于，包括：

10 获取第一目标虚拟机的第一调度器存储的时间记录信息；其中，所述时间记录信息用于记录第四任务执行的时间，所述第四任务对应第三优先级，或者，所述时间记录信息用于记录第五任务在所述第一调度器存储的待执行任务队列中的等待时间，所述第五任务对应的优先级低于所述第三优先级；

当所述时间记录信息满足预设条件，控制所述第一目标虚拟机的第一调度器将所述第五任务发送到第二目标虚拟机的第二调度器，使所述第二目标虚拟机执行所述第五任务。

15 11、根据权利要求 10 所述的方法，其特征在于，

当所述时间记录信息用于记录所述第四任务执行的时间，所述预设条件包括：
所述第四任务执行的时间大于第一阈值。

12、根据权利要求 10 所述的方法，其特征在于，

20 当所述时间记录信息记录第五任务在所述第一调度器的待执行任务队列中的等待时间，所述预设条件包括：
所述等待时间大于第二阈值。

13、根据权利要求 10-12 任一项所述的方法，其特征在于，

25 所述第二目标虚拟机为所述第一目标虚拟机所在主机部署的多个虚拟机中处于空闲态的虚拟机；或者，所述第二目标虚拟机为所述主机部署的多个虚拟机中负载小于第三阈值的虚拟机；或者，所述第二目标虚拟机为所述主机部署的多个虚拟机中用于执行第四优先级的任务的调度器，其中，所述第四优先级低于所述第三优先级。

14、根据权利要求 10-12 任一项所述的方法，其特征在于，

所述第二目标虚拟机是专用于当所述时间记录信息满足预设条件时执行所述第五任务的调度器。

30 15、一种任务调度装置，其特征在于，包括：

获取模块，用于获取第一目标虚拟机待执行的多个任务的优先级；

优先级确定模块，用于确定所述多个任务中优先级最高的至少一个任务；

35 虚拟运行时间确定模块，用于从所述至少一个任务中，确定虚拟运行时间最小的第一任务；

控制模块，用于控制所述第一目标虚拟机执行所述第一任务。

16、根据权利要求 15 所述的装置，其特征在于，所述获取模块具体用于，

根据所述第一目标虚拟机中，第一调度器所存储的待执行任务队列，确定所述第一目标虚拟机待执行的多个任务的优先级。

17、根据权利要求 16 所述的装置，其特征在于，所述获取模块具体用于，

遍历所述第一调度器所存储的待执行任务队列上所有任务的优先级，从所述待执行任务的队列中确定所述优先级最高的至少一个任务。

18、根据权利要求 16 所述的装置，其特征在于，所述获取模块具体用于，

5 通过所述第一调度器中存储的优先级记录信息，确定所述多个任务中优先级最高的至少一个任务；其中，所述优先级记录信息用于记录所述第一调度器待执行的队列上所有任务的优先级。

19、根据权利要求 16 所述的装置，其特征在于，所述获取模块具体用于，

10 根据所述第一调度器中存储的最高优先级对应的队列中的任务，确定所述优先级最高的至少一个任务；其中，所述第一调度器待执行任务通过多个队列存储待执行的任务，每个队列对应不同的优先级。

20、一种任务优先级调整装置，其特征在于，包括：

获取模块，用于当第一目标虚拟机待执行的第二任务的优先级由第二优先级调整为第一优先级，获取所述第一目标虚拟机待执行的与所述第一优先级相同的至少一个任务中，所述至少一个任务对应的最小的虚拟运行时间；

15 虚拟运行时间调整模块，用于根据所述最小的虚拟运行时间，将所述第二任务对应的第一虚拟运行时间调整为第二虚拟运行时间；

控制模块，用于当所述第一优先级高于所述第一目标虚拟机正在执行的第三任务的优先级，控制所述第一目标虚拟机执行所述第二任务。

21、根据权利要求 20 所述的装置，其特征在于，

20 所述第二虚拟运行时间大于所述第一目标虚拟机的第一调度器存储的待执行任务队列中，对应所述第一优先级的至少一个任务的虚拟运行时间中最小的虚拟运行时间。

22、根据权利要求 20 所述的装置，其特征在于，还包括：

接收模块，用于接收指示信息，并根据所述指示信息，确定所述第二任务的优先级由第二优先级调整为第一优先级。

25 23、根据权利要求 20-22 任一项所述的装置，其特征在于，所述虚拟时间调整模块具体用于，

将所述第一虚拟运行时间加上所述最小的虚拟运行时间后，再减去所述待执行任务队列中对应所述第一优先级的至少一个任务的虚拟运行时间中最小的虚拟运行时间，得到所述第二虚拟运行时间。

30 24、一种任务调度装置，其特征在于，包括：

获取模块，用于获取第一目标虚拟机的第一调度器存储的时间记录信息；其中，所述时间记录信息用于记录第四任务执行的时间，所述第四任务对应第三优先级，或者，所述时间记录信息用于记录第五任务在所述第一调度器存储的待执行任务队列中的等待时间，所述第五任务对应的优先级低于所述第三优先级；

35 控制模块，用于当所述时间记录信息满足预设条件，控制所述第一目标虚拟机的第一调度器将所述第五任务发送到第二目标虚拟机的第二调度器，使所述第二目标虚拟机执行所述第五任务。

25、根据权利要求 24 所述的装置，其特征在于，

当所述时间记录信息用于记录所述第四任务执行的时间，所述预设条件包括：

所述第四任务执行的时间大于第一阈值。

26、根据权利要求 24 所述的装置，其特征在于，

当所述时间记录信息记录第五任务在所述第一调度器的待执行任务队列中的等待时间，所述预设条件包括：

5 所述等待时间大于第二阈值。

27、根据权利要求 24-26 任一项所述的装置，其特征在于，

10 所述第二目标虚拟机为所述第一目标虚拟机所在主机部署的多个虚拟机中处于空闲态的虚拟机；或者，所述第二目标虚拟机为所述主机部署的多个虚拟机中负载小于第三阈值的虚拟机；或者，所述第二目标虚拟机为所述主机部署的多个虚拟机中用于执行第四优先级的任务的调度器，其中，所述第四优先级低于所述第三优先级。

28、根据权利要求 24-26 任一项所述的装置，其特征在于，

所述第二目标虚拟机是专用于当所述时间记录信息满足预设条件时执行所述第五任务的调度器。

29、一种电子设备，其特征在于，包括：处理器和通信接口；

15 所述通信接口用于实现所述电子设备与其他装置进行通信；

所述处理器用于实现如权利要求 1-5 任一项、6-9 任一项或者 10-14 任一项所述的方法。

20 30、一种计算机可读存储介质，其特征在于，所述计算机可读存储介质存储有计算机程序，当所述计算机程序被运行时，用于实现如权利要求 1-5 任一项、6-9 任一项或者 10-14 任一项所述的方法。

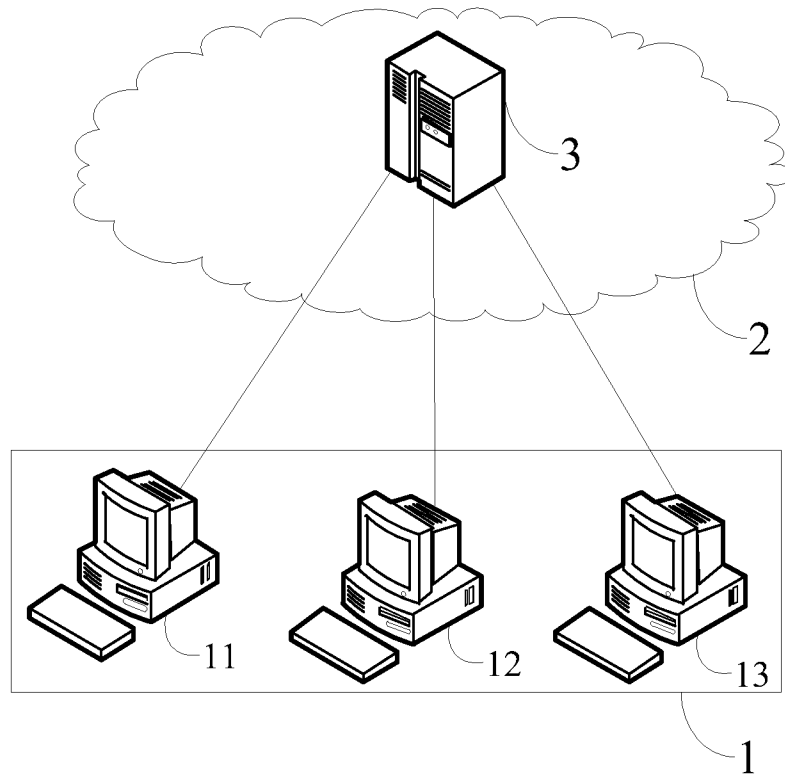


图 1

虚拟机1	虚拟机2 (虚拟机4)	虚拟机3
	操作系统	
	系统资源	

图 2

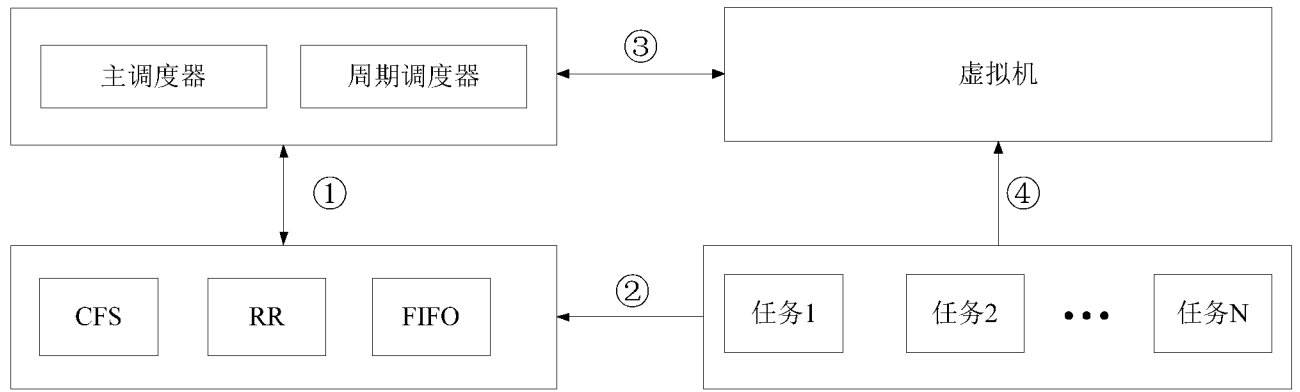


图 3

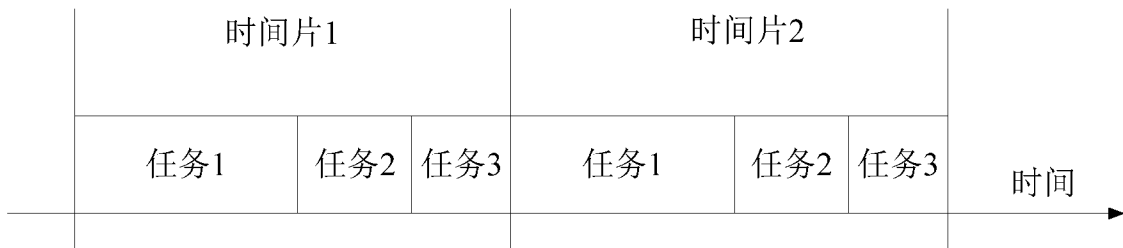


图 4

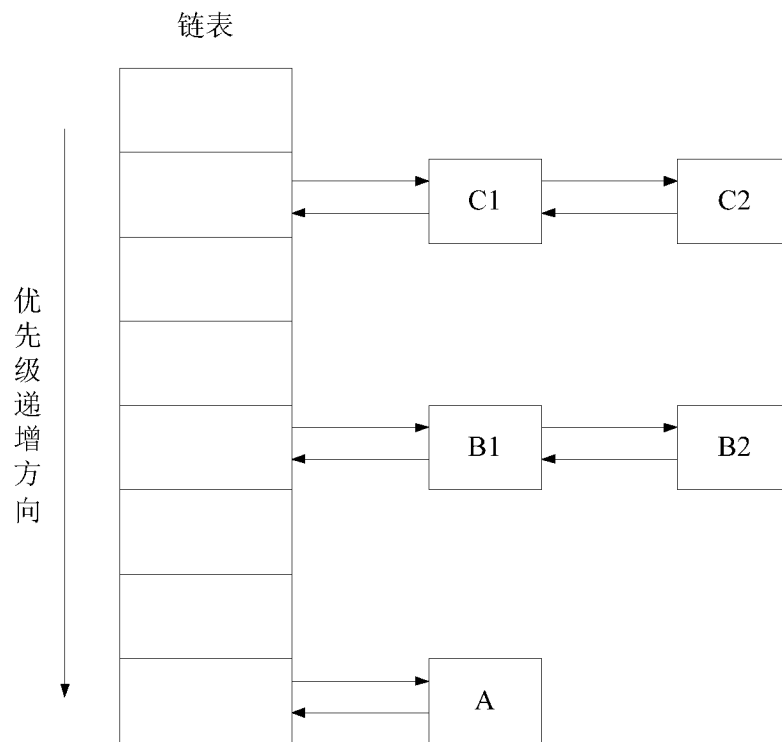


图 5

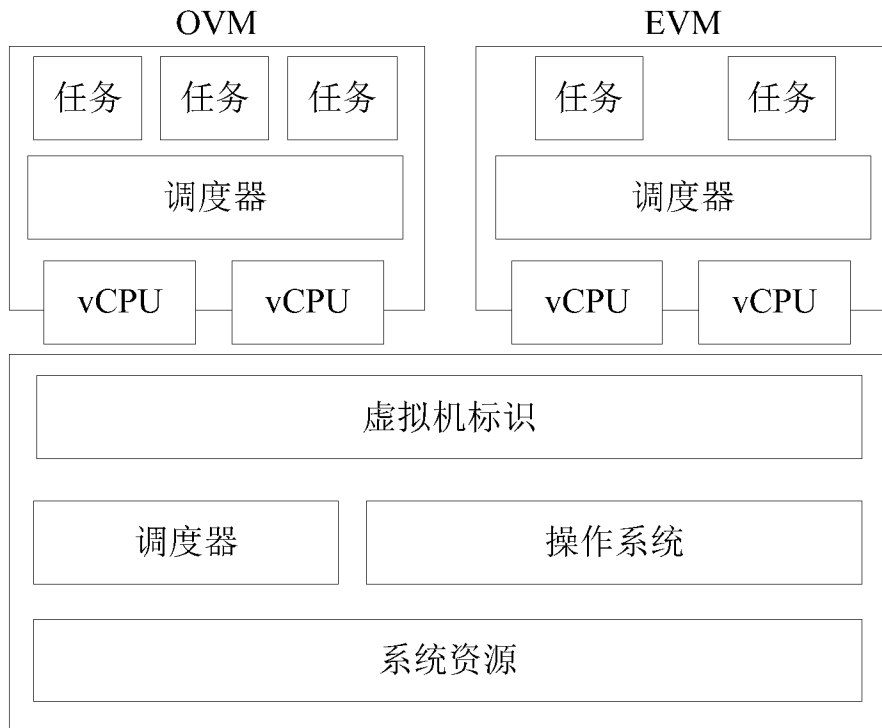


图 6

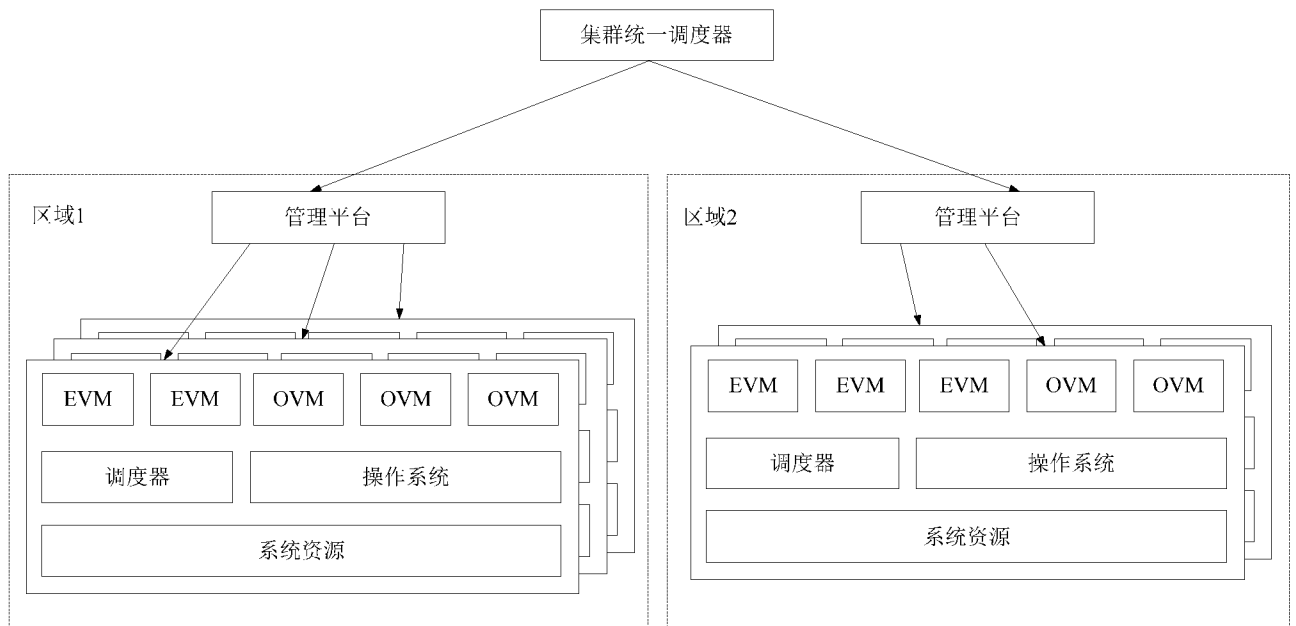


图 7

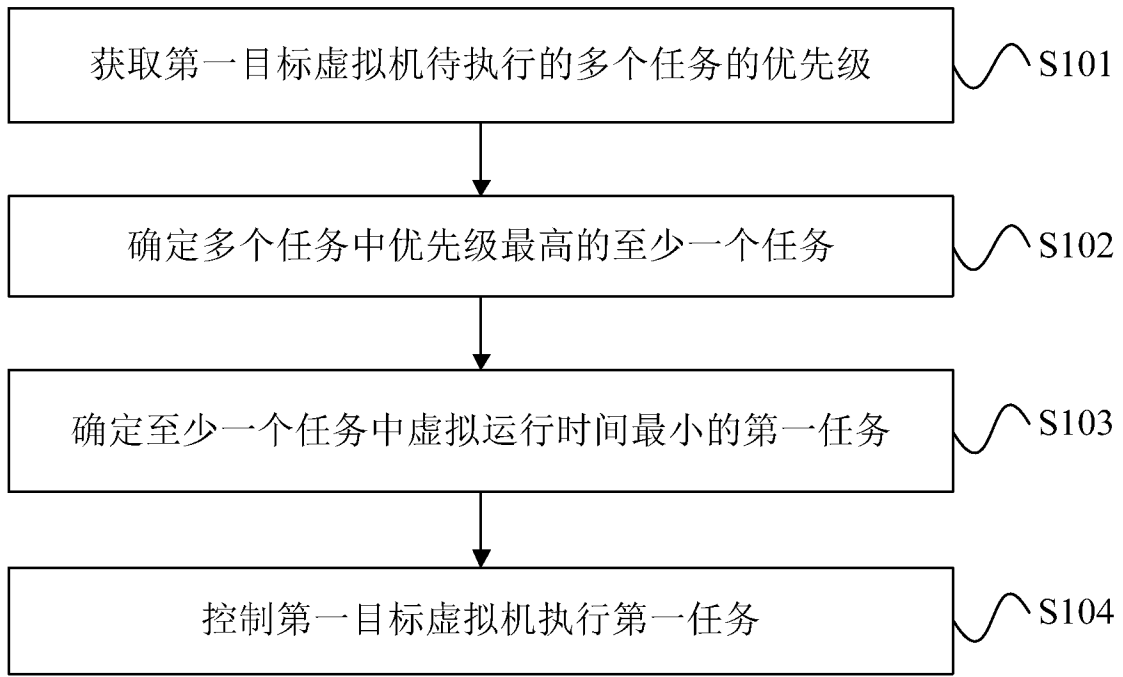


图 8

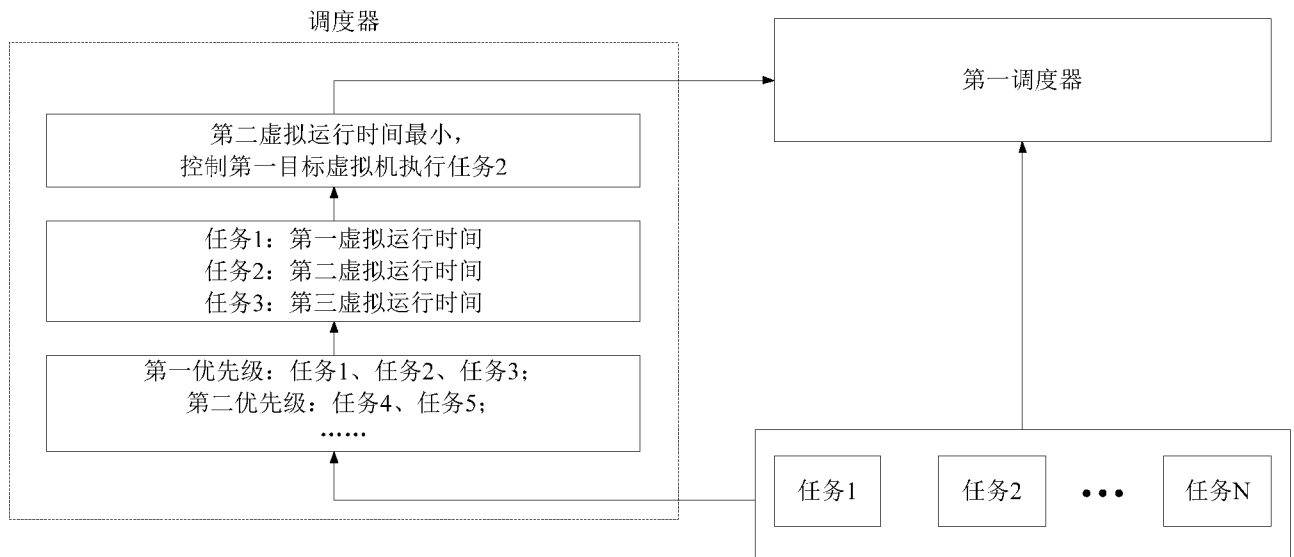


图 9

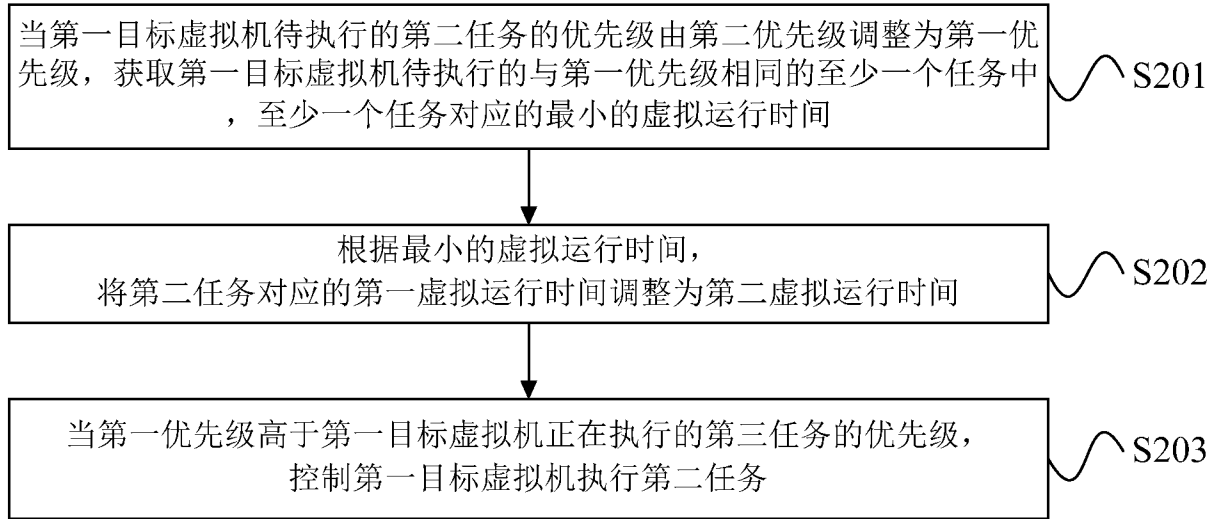


图 10

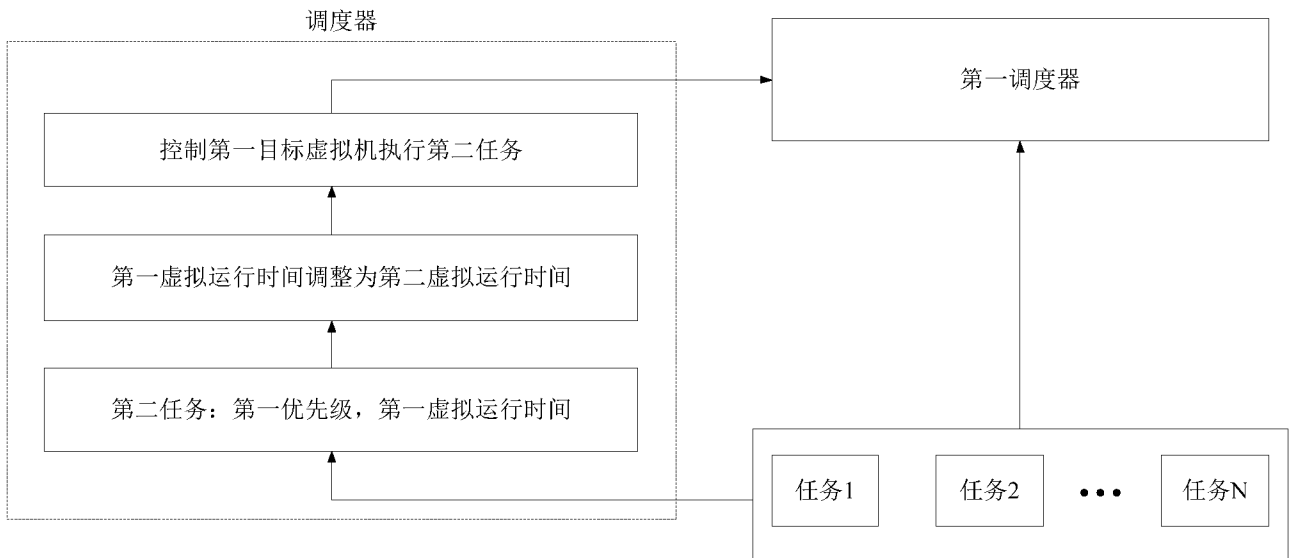
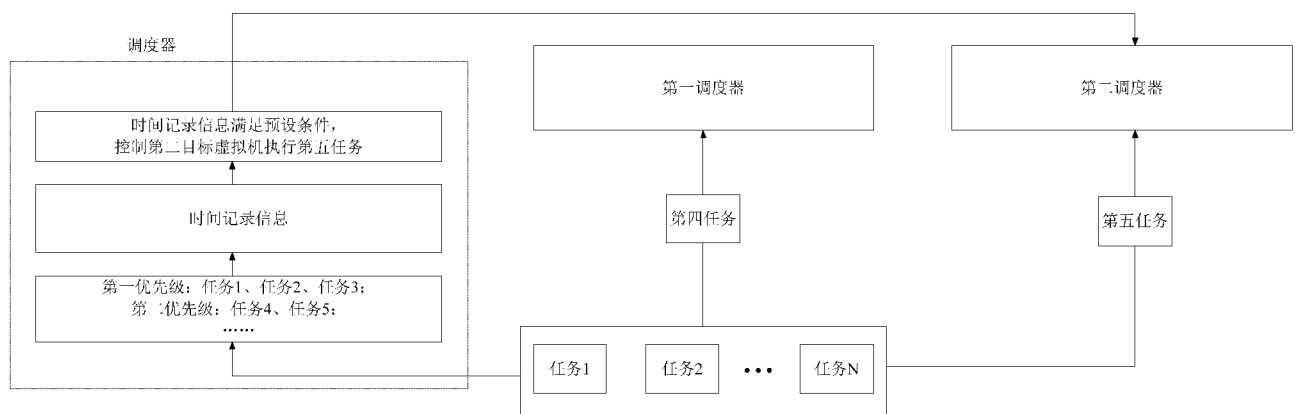
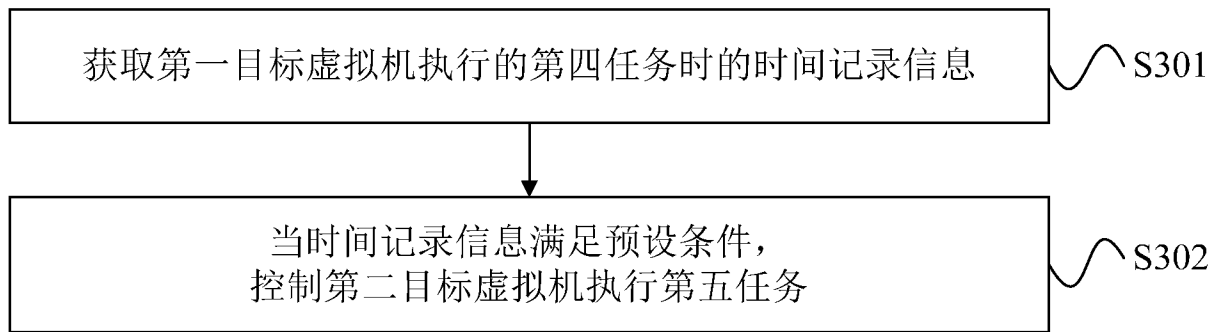
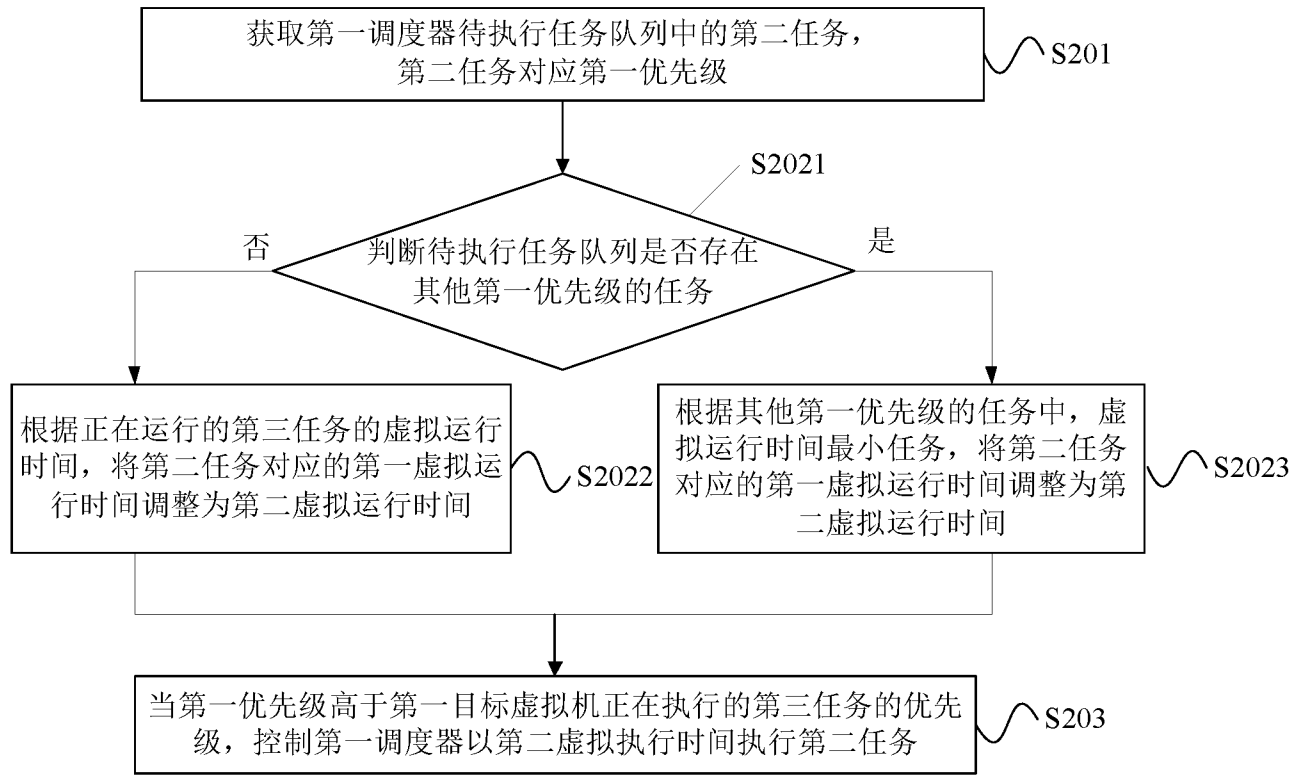


图 11



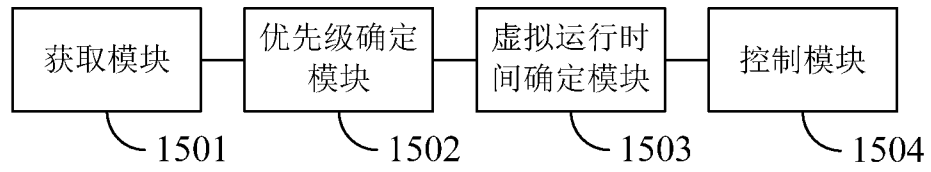


图 15

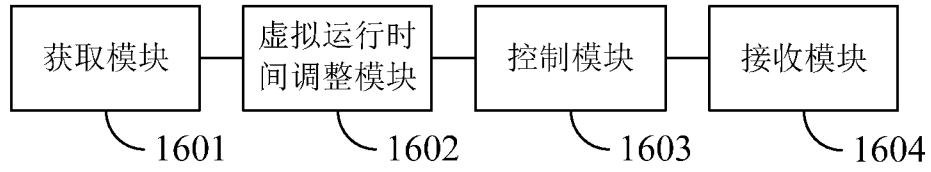


图 16

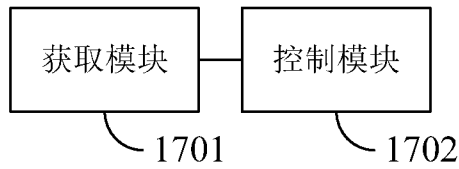


图 17

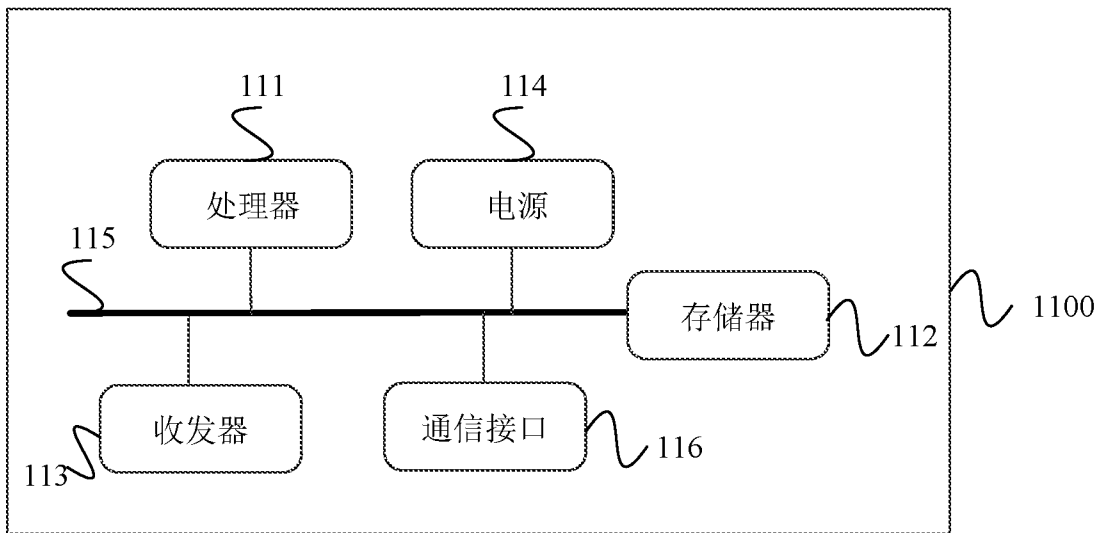


图 18

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2021/120360

A. CLASSIFICATION OF SUBJECT MATTER G06F 9/50(2006.01)i According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) G06F Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) CNABS; CNTXT; CNKI; DWPI; SIPOABS: 华为, 张海亮, 完全公平, 调度, 调整, 更改, 变更, 改变, 调节, 设置, 设定, 确定, 虚拟运行时间, 最小, 最少, 时间, 迁, 移, 换, 转, 挪, 任务, 作业, 优先级, 高, 低, 抢占, 空闲, 小, 低, 负载, 虚拟机, vruntime, CFS, Scheduler, adjust+, modulat+, chang+, modif+, least, time, task, mov+, PRI, high, low, load, vm		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 108595249 A (LENOVO (BEIJING) LIMITED) 28 September 2018 (2018-09-28) description, paragraphs 2-32	1-5, 15-19, 29, 30
Y	CN 108268310 A (DATANG MOBILE COMMUNICATIONS EQUIPMENT CO., LTD.) 10 July 2018 (2018-07-10) description, paragraphs 2-51, 59-128	1-9, 15-23, 29, 30
Y	CN 109684060 A (XI'AN AVIATION INSTITUTE OF COMPUTING TECHNOLOGY, AVIATION INDUSTRY CORPORATION OF CHINA) 26 April 2019 (2019-04-26) description, paragraphs 2-35	6-9, 20-23, 29, 30
Y	CN 103605567 A (HOHAI UNIVERSITY) 26 February 2014 (2014-02-26) description, paragraphs 2-17, 25-96	10-14, 24-30
Y	CN 111400087 A (GENEW TECHNOLOGIES CO., LTD.) 10 July 2020 (2020-07-10) description, paragraphs 2-28, 35-88	10-14, 24-30
A	US 7293136 B1 (EMC CORPORATION) 06 November 2007 (2007-11-06) entire document	1-30
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search 16 December 2021		Date of mailing of the international search report 28 December 2021
Name and mailing address of the ISA/CN China National Intellectual Property Administration (ISA/CN) No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088, China Facsimile No. (86-10)62019451		Authorized officer Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2021/120360

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	108595249	A	28 September 2018	None			
CN	108268310	A	10 July 2018	CN	108268310	B	28 August 2020
CN	109684060	A	26 April 2019	None			
CN	103605567	A	26 February 2014	CN	103605567	B	22 March 2017
CN	111400087	A	10 July 2020	None			
US	7293136	B1	06 November 2007	None			

A. 主题的分类 G06F 9/50 (2006.01) i 按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类		
B. 检索领域 检索的最低限度文献(标明分类系统和分类号) G06F 包含在检索领域中的除最低限度文献以外的检索文献 在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用)) CNABS;CNTXT;CNKI;DWPI;SIPOABS: 华为, 张海亮, 完全公平, 调度, 调整, 更改, 变更, 改变, 调节, 设置, 设定, 确定, 虚拟运行时间, 最小, 最少, 时间, 迁, 移, 换, 转, 挪, 任务, 作业, 优先级, 高, 低, 抢占, 空闲, 小, 低, 负载, 虚拟机, vruntime, CFS, Scheduler, adjust+, modulat+, chang+, modif+, least, time, task, mov+, PRI, high, low, load, vm		
C. 相关文件		
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求
Y	CN 108595249 A (联想北京有限公司) 2018年9月28日 (2018 - 09 - 28) 说明书第2-32段	1-5, 15-19, 29, 30
Y	CN 108268310 A (大唐移动通信设备有限公司) 2018年7月10日 (2018 - 07 - 10) 说明书第2-51, 59-128段	1-9, 15-23, 29, 30
Y	CN 109684060 A (中国航空工业集团公司西安航空计算技术研究所) 2019年4月26日 (2019 - 04 - 26) 说明书第2-35段	6-9, 20-23, 29, 30
Y	CN 103605567 A (河海大学) 2014年2月26日 (2014 - 02 - 26) 说明书第2-17, 25-96段	10-14, 24-30
Y	CN 111400087 A (深圳震有科技股份有限公司) 2020年7月10日 (2020 - 07 - 10) 说明书第2-28, 35-88段	10-14, 24-30
A	US 7293136 B1 (EMC CORP) 2007年11月6日 (2007 - 11 - 06) 全文	1-30
<input type="checkbox"/> 其余文件在C栏的续页中列出。		
<input checked="" type="checkbox"/> 见同族专利附件。		
* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件 “T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件		
国际检索实际完成的日期 2021年12月16日		国际检索报告邮寄日期 2021年12月28日
ISA/CN的名称和邮寄地址 中国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451		授权官员 孙蕾 电话号码 86-(010)-62412294

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2021/120360

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	108595249	A	2018年9月28日	无			
CN	108268310	A	2018年7月10日	CN	108268310	B	2020年8月28日
CN	109684060	A	2019年4月26日	无			
CN	103605567	A	2014年2月26日	CN	103605567	B	2017年3月22日
CN	111400087	A	2020年7月10日	无			
US	7293136	B1	2007年11月6日	无			