

(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
2021 年 2 月 4 日 (04.02.2021)



(10) 国际公布号
WO 2021/018183 A1

- (51) 国际专利分类号:
G06F 9/48 (2006.01)
- (21) 国际申请号: PCT/CN2020/105476
- (22) 国际申请日: 2020 年 7 月 29 日 (29.07.2020)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201910704756.7 2019年7月31日 (31.07.2019) 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: 曾艳 (ZENG, Yan); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 姜慧 (JIANG, Hui); 中国北京市海淀区中关村科学院南路6号中国科学院计算技术研究所, Beijing 100190 (CN)。 李继忠 (LI, Jizhong);

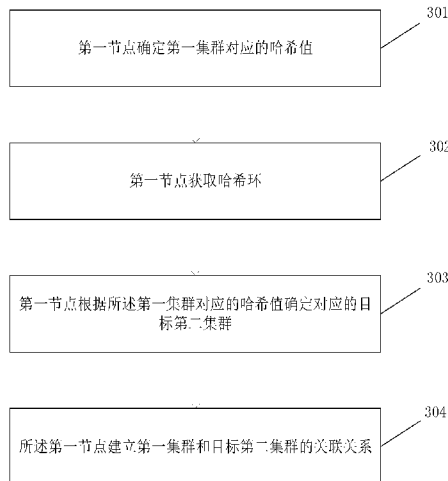
中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 刘敏 (LIU, Min); 中国北京市海淀区中关村科学院南路6号中国科学院计算技术研究所, Beijing 100190 (CN)。

(74) 代理人: 深圳市深佳知识产权代理事务所 (普通合伙) (SHENPAT INTELLECTUAL PROPERTY AGENCY); 中国广东省深圳市罗湖区南湖街道春风路庐山大厦B座18C2、18D、18E、18E2, Guangdong 518001 (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: RESOURCE ALLOCATION METHOD AND RESOURCE OFFLOADING METHOD

(54) 发明名称: 一种资源分配方法和资源借调方法



- 301 A first node determines a hash value corresponding to a first cluster
302 The first node acquires a hash ring
303 The first node determines a corresponding target second cluster according to the hash value corresponding to the first cluster
304 The first node establishes an association relationship between the first cluster and a target second cluster

图 3a

(57) Abstract: A resource allocation method, so as to prevent offloading conflict caused by multiple busy clusters concurrently offloading resources to the same idle cluster. The resource allocation method comprises: a first node determining a hash value corresponding to a first cluster; the first node acquiring a hash ring; the first node establishing an association relationship between the first cluster and a target second cluster (304), the target second cluster having an association relationship established with the first cluster permitting a resource to be applied for from the first cluster.

(57) 摘要: 一种资源分配方法, 不会出现由于多个忙集群向同一个空闲集群并发借调资源而导致的借调冲突。其中, 资源分配方法包括: 第一节点确定第一集群对应的哈希值; 第一节点获取哈希环; 第一节点建立第一集群和目标第二集群的关联关系 (304), 其中, 与第一集群建立了关联关系的目标第二集群被允许从第一集群申请资源。

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))。

一种资源分配方法和资源借调方法

本申请要求于2019年07月31日提交中国国家知识产权局、申请号为201910704756.7、发明名称为“一种资源分配方法和资源借调方法”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。

5

技术领域

本申请涉及分布式系统领域，尤其涉及一种资源分配方法和资源借调方法。

背景技术

随着人工智能（artificial intelligence, AI）和物联网（internet of things, IoT）的发展，许多新兴计算密集型和延迟敏感型终端应用，如虚拟现实（virtual reality, VR）、增强现实（augmented reality, AR）以及高清视频图片处理等正在迅速兴起，低延时、高带宽和强计算等应用需求日益增长。云计算技术受网络链路容量限制等因素影响，容易导致数据传输速率过慢以及响应延迟过高等问题，难以满足应用服务需求。利用靠近数据源的边缘设备进行计算以提高服务响应性能逐渐成为一种趋势。然而单个边缘设备算力小，仅使用单个边缘设备执行任务难以胜任大规模任务需求。因此，将多边缘设备的计算能力加以组合使用，通过多边缘设备资源共享互助，协同完成复杂任务的数据处理，从而可以有效保障服务质量。

设备间的资源互助场景主要分为两类，一类是层次化的异构设备的资源互助，如边缘设备和边缘服务器、中央云处理器之间不同运算能力设备的协同；另一类是同层设备间的资源协同，即当边缘设备资源不足时，可以向其他边缘设备借调资源。

现有技术中，当资源不足的忙集群需要向其他空闲集群借调资源时，需要忙集群中的主节点向随机选择的空闲集群发送资源查询请求，并在接收到返回后，再进行资源借调，在多个忙集群向同一个空闲集群并发借调资源时，会出现借调冲突的情况。

发明内容

本申请实施例提供了一种资源分配方法和资源借调方法，不会出现由于多个忙集群向同一个空闲集群并发借调资源而导致的借调冲突。

第一方面，本申请提供了一种资源分配方法，包括：

第一节点确定第一集群对应的哈希值；所述第一节点获取哈希环，所述哈希环包括至少一个哈希值区间，每个哈希值区间对应一个第二集群；所述第一节点根据所述第一集群对应的哈希值确定对应的目标第二集群，其中，所述第一集群对应的哈希值属于目标哈希区间，所述目标哈希区间为所述至少一个哈希值区间中的一个哈希值区间，所述目标哈希区间对应所述目标第二集群；所述第一节点建立第一集群和目标第二集群的关联关系，其中，与所述第一集群建立了关联关系的所述目标第二集群被允许从所述第一集群申请资源。且，由于资源借调的对应关系是基于将第一集群的哈希值在基于第二集群生成的哈希环上的对应关系确定的，该对应关系是唯一的，因此第一集群仅能被哈希环上对应的第二集群

-2-

借调资源，因此，不会出现由于多个集群向同一个集群并发借调资源而导致的借调冲突。

这样，第一节点可以通过第一集群对应的哈希值在哈希环（根据第二集群对应的哈希值生成的）上的分布，来确定第一集群对应的目标第二集群，建立目标第二集群和第一集群的关联关系，由于基于一致性哈希算法的哈希环上的第一集群对应的哈希值和第二集群对应的哈希节点的分布是确定的，每个集群只允许被固定的集群借调，因此，当目标第二集群需要借调第一集群的空闲资源时，不需要随机向某个集群发送资源查询请求，并在接收到返回后，再进行资源借调，可以直接向特定的第一集群借调资源，减少了一次信息往返的时延开销。

在第一方面的一种可选设计中，第一集群为包含多个空闲节点的空闲集群，第二集群为包含多个忙节点的忙集群，其中空闲节点和忙节点是根据节点的负载来区分的，比如负载高于特定阈值的节点为忙节点，负载不高于特定阈值的节点为空闲节点。

在第一方面的一种可选设计中，第一集群为包含多个胖节点的胖集群，第二集群为包含多个瘦节点的瘦集群，其中胖节点和瘦节点是根据节点具有的资源或能力来区分的，比如资源受限或能力偏弱的节点为瘦节点，资源丰富，处理能力强节点为胖节点。在第一方面的一种可选设计中，所述哈希环包括第一哈希节点和第二哈希节点，所述第一哈希节点对应所述目标第二集群，所述第一哈希节点为所述哈希环上所述第二哈希节点的相邻后继节点，所述目标哈希区间为所述第一哈希节点的哈希值与所述第二哈希节点的哈希值之间的哈希区间。

在第一方面的一种可选设计中，所述方法还包括：所述第一节点对包括多个计算节点的空闲集群进行划分，得到多个子集群，所述第一集群为所述多个子集群中的一个。

在第一方面的一种可选设计中，所述多个子集群包括第三子集群和第四子集群，所述第三子集群为所述多个子集群中资源空闲量最大的子集群，所述第四子集群为所述多个子集群中资源空闲量最小的集群，所述第三子集群的资源空闲量与所述第四子集群的资源空闲量的差值的绝对值小于第一阈值。这样，第一节点可以对空闲集群的多个计算节点进行划分，以使得划分后得到的多个子集群中的各个第一集群之间的资源空闲量的差异不会过大。

在第一方面的一种可选设计中，所述多个子集群中的每个子集群的资源空闲量与预设资源空闲量的差值的绝对值小于或等于第二阈值，所述预设资源空闲量与历史资源请求量相关。这样，第一节点基于历史资源请求在对空闲集群的多个计算节点进行划分，使得划分后的每个子集群的空闲资源量接近于历史资源请求量。

在第一方面的一种可选设计中，所述多个子集群的数量与所述第二集群的数量小于第三阈值。

在第一方面的一种可选设计中，所述方法还包括：所述第一节点接收所述目标第二集群发送的第一资源借调请求，所述第一资源借调请求包括任务数据和第一集群的标识；所述第一节点根据所述第一集群的标识向所述第一集群发送所述任务数据。

在第一方面的一种可选设计中，所述方法还包括：所述第一节点接收所述目标第二集群发送的第一资源借调请求，所述第一资源借调请求包括任务数据；所述第一节点向所述

第一集群发送所述任务数据。

在第一方面的一种可选设计中，所述第一节点向所述第一集群发送所述任务数据，包括：第一节点基于所述第一集群和目标第二集群的关联关系向所述第一集群发送所述任务数据。

5 在第一方面的一种可选设计中，所述哈希环包括 Y 个哈希节点，每个第二集群对应至少一个哈希节点，且所述 Y 个哈希节点的对应的哈希值和所述第一集群对应的哈希值为根据相同的哈希算法生成的，所述 Y 为正整数。这样，哈希环上的哈希节点的对应的哈希值和所述第一集群对应的哈希值基于相同的哈希算法生成，使得不会出现第一集群对应的哈希值与哈希环上的哈希节点重合的情况。

10 在第一方面的一种可选设计中，所述方法还包括：所述第一节点确定 X 个第二集群对应的 Y 个哈希值，其中，每个第二集群对应至少一个哈希值，所述 X 正整数，所述 Y 为大于或等于 X 的正整数；所述第一节点根据所述 Y 个哈希值生成所述哈希环，其中，所述哈希环包括 Y 个哈希节点，每个第二集群对应至少一个哈希节点，所述 Y 个哈希节点将所述哈希环划分为 Y 个哈希值区间，每个哈希节点对应一个哈希值区间。这样，第一节点可以自己生成哈希环。

15 在第一方面的一种可选设计中，所述方法还包括：所述第一节点接收所述第二集群发送的所述哈希环。

在第一方面的一种可选设计中，所述方法还包括：所述第一节点向所述目标第二集群发送所述第一集群的资源空闲量。

20 第二方面，本申请提供了一种资源借调方法，所述方法包括：第二节点确定与目标第二集群具有关联关系的第一集群，其中，所述第二节点属于所述目标第二集群，所述目标第二集群对应于哈希环中的目标哈希值区间，所述第一集群对应的哈希值属于所述目标哈希值区间；所述第二节点接收任务数据；所述第二节点向第一节点发送第一资源借调请求，所述第一借调请求用于指示所述第一节点将所述任务数据分配给所述第一集群处理。

25 这样，由于基于一致性哈希算法的哈希环上的第一集群对应的哈希值和第二集群对应的哈希节点的分布是确定的，每个空闲集群只允许被固定的空集群借调，因此，当目标第二集群需要借调第一集群的空闲资源时，不需要随机向选择的空闲集群发送资源查询请求，并在接收到返回后，再进行资源借调，可以直接向第一集群借调资源，减少了一次信息往返的时延开销。

30 在第二方面的一种可选设计中，所述方法还包括：所述第二节点基于向第一节点发送所述第一资源借调请求，确定所述第一集群处于资源不可借调状态。这样，第二节点向第一节点发送第一资源借调请求之后，相当于已经占用了第一集群的空闲资源，则可以确定第一集群处于资源不可借调状态。可选地，第二节点可以在映射表中标注第一集群的资源借调状态为资源不可借状态。相应的，第一节点也可以确定第一集群处于资源不可借调状态，即确定第一集群不能为其他忙集群提供可用资源，从而确保了资源锁定，实现了跨集群资源的无冲突调度。

35 在第二方面的一种可选设计中，与所述目标第二集群具有关联关系的集群有多个，所

—4—

述第二节点确定与目标第二集群具有关联关系的第一集群，包括：所述第二节点确定与
所述目标第二集群具有关联关系的多个集群中的一个集群为第一集群。

在第二方面的一种可选设计中，所述第二节点确定与所述目标第二集群具有关联关系
的多个集群中的一个集群为第一集群，包括：所述第二节点确定与所述目标第二集群具有
5 关联关系的多个集群的资源空闲量中，与所述任务数据的资源需求量差值最小的资源空闲
量对应的集群为所述第一集群。这样，所述第二节点确定与所述目标第二集群具有关联关
系的多个集群的资源空闲量中，与所述任务数据的资源需求量差值最小的资源空闲量对应
的集群为所述第一集群，提高了资源借调时的资源利用率。

在第二方面的一种可选设计中，其特征在于，所述方法还包括：所述第二节点接收第
10 三节点发送的所述任务数据的资源需求量；基于所述第一集群的资源空闲量大于或等于所
述资源需求量，所述第二节点向所述第三节点发送所述第一集群的标识和第一集群的资源
空闲量。

在第二方面的一种可选设计中，所述第二节点接收任务数据包括：所述第二节点接收
所述第三节点发送的第二资源借调请求，所述第二资源借调请求包括所述任务数据，所述
15 第二借调请求用于指示所述第二节点将所述任务数据分配给所述第一集群处理。

第三方面，本申请提供了一种资源借调方法，包括：第二节点接收第三节点发送的任
务数据的资源需求量；所述第二节点确定与目标第二集群具有关联关系的第一集群，其中，
所述第二节点属于所述目标第二集群，所述目标第二集群对应于哈希环中的目标哈希值区
间，所述第一集群对应的哈希值属于所述目标哈希值区间；基于所述第一集群的资源空闲
20 量大于或等于所述资源需求量，所述第二节点向所述第三节点发送第三资源借调请求，所
述第三资源借调请求用于指示所述第三节点被允许从所述第一集群申请资源。

这样，第一节点可以通过第一集群对应的哈希值在哈希环（根据第二集群对应的哈希
值生成的）上的分布，来确定第一集群对应的目标第二集群，建立目标第二集群和第一集
群的关联关系，由于基于一致性哈希算法的哈希环上的第一集群对应的哈希值和第二集
25 群对应的哈希节点的分布是确定的，每个空闲集群只允许被固定的空集群借调，因此，当第
三节点需要向第二节点借调资源时，第二节点不需要随机向选择的空闲集群发送资源查询
请求，并在接收到返回后，再进行资源借调，而是直接可以确定可以借调资源的第一集群，
减少了一次信息往返的时延开销。

在第三方面的一种可选设计中，与所述目标第二集群具有关联关系的集群有多个，所
30 述第二节点确定与目标第二集群具有关联关系的第一集群，包括：所述第二节点确定与所
述目标第二集群具有关联关系的多个集群中的一个集群为第一集群。

第四方面，本申请提供了一种资源借调方法，包括：第三节点接收任务数据；所述第
三节点向第二节点发送所述任务数据的资源需求量；所述第三节点接收所述第二节点发送
的第三资源借调请求，所述第三资源借调请求用于指示所述第三节点被允许从第一集群申
35 请资源，其中，所述第二节点属于目标第二集群，所述目标第二集群对应于哈希环中的目
标哈希值区间，所述第一集群对应的哈希值属于所述目标哈希值区间，且所述第一集群的
资源空闲量大于或等于所述资源需求量；所述第三节点向第一节点发送第一资源借调请求，

—5—

所述第一借调请求用于指示所述第一节点将所述任务数据分配给所述第一集群处理。

这样，由于基于一致性哈希算法的哈希环上的第一集群对应的哈希值和第二集群对应的哈希节点的分布是确定的，且相对均衡，因此，当第三节点需要向第二节点借调资源时，第二节点不需要随机向选择的空闲集群发送资源查询请求，并在接收到返回后，再进行资源借调，而是直接可以确定可以借调资源的第一集群，减少了一次信息往返的时延开销。

在第四方面的一种可选设计中，所述第三节点向第二节点发送所述任务数据的资源需求量，包括：所述第三节点向多个节点发送所述任务数据的资源需求量，所述第二节点为所述多个节点中的一个；所述第三节点接收所述第二节点发送的第三资源借调请求，包括：所述第三节点接收所述多个节点发送的第三资源借调请求，所述第三资源借调请求用于指示所述第三节点被允许从多个子集群申请资源，所述第一集群为所述多个子集群中的一个。

在第四方面的一种可选设计中，所述方法还包括：所述第三节点确定所述多个子集群的资源空闲量中，与所述任务数据的资源需求量的差值最小的资源空闲量对应的子集群为所述第一集群。

第五方面，本申请提供了一种第一管理节点，所述第一管理节点包括：

处理模块，用于确定第一集群对应的哈希值；获取哈希环，所述哈希环包括至少一个哈希值区间，每个哈希值区间对应一个第二集群；根据所述第一集群对应的哈希值确定对应的目标第二集群，其中，所述第一集群对应的哈希值属于目标哈希区间，所述目标哈希区间为所述至少一个哈希值区间中的一个哈希值区间，所述目标哈希区间对应所述目标第二集群；

收发模块，用于建立第一集群和目标第二集群的关联关系，其中，与所述第一集群建立了关联关系的所述目标第二集群被允许从所述第一集群申请资源。

在第五方面的一种可选设计中，所述哈希环包括第一哈希节点和第二哈希节点，所述第一哈希节点对应所述目标第二集群，所述第一哈希节点为所述哈希环上所述第二哈希节点的相邻后继节点，所述目标哈希区间为所述第一哈希节点的哈希值与所述第二哈希节点的哈希值之间的哈希区间。

在第五方面的一种可选设计中，所述处理模块，还用于对包括多个计算节点的空闲集群进行划分，得到多个子集群，所述第一集群为所述多个子集群中的一个。

在第五方面的一种可选设计中，所述多个子集群包括第三子集群和第四子集群，所述第三子集群为所述多个子集群中资源空闲量最大的子集群，所述第四子集群为所述多个子集群中资源空闲量最小的集群，所述第三子集群的资源空闲量与所述第四子集群的资源空闲量的差值的绝对值小于第一阈值。

在第五方面的一种可选设计中，所述多个子集群中的每个子集群的资源空闲量与预设资源空闲量的差值的绝对值小于或等于第二阈值，所述预设资源空闲量与历史资源请求量相关。

在第五方面的一种可选设计中，所述多个子集群的数量与所述第二集群的数量小于第三阈值。

在第五方面的一种可选设计中，所述收发模块还用于：

—6—

接收所述目标第二集群发送的第一资源借调请求，所述第一资源借调请求包括任务数据和第一集群的标识；

根据所述第一集群的标识向所述第一集群发送所述任务数据。

在第五方面的一种可选设计中，所述收发模块还用于：

5 接收所述目标第二集群发送的第一资源借调请求，所述第一资源借调请求包括任务数据，向所述第一集群发送所述任务数据。

在第五方面的一种可选设计中，所述收发模块具体用于：基于所述第一集群和目标第二集群的关联关系向所述第一集群发送所述任务数据。

10 在第五方面的一种可选设计中，所述哈希环包括 Y 个哈希节点，每个第二集群对应至少一个哈希节点，且所述 Y 个哈希节点的对应的哈希值和所述第一集群对应的哈希值为根据相同的哈希算法生成的，所述 Y 为正整数。

在第五方面的一种可选设计中，所述处理模块，还用于确定 X 个第二集群对应的 Y 个哈希值，其中，每个第二集群对应至少一个哈希值，所述 X 正整数，所述 Y 为大于或等于 X 的正整数；根据所述 Y 个哈希值生成所述哈希环，其中，所述哈希环包括 Y 个哈希节点，
15 每个第二集群对应至少一个哈希节点，所述 Y 个哈希节点将所述哈希环划分为 Y 个哈希值区间，每个哈希节点对应一个哈希值区间。

在第五方面的一种可选设计中，所述收发模块，还用于接收所述第二集群发送的所述哈希环。

20 在第五方面的一种可选设计中，所述收发模块，还用于向所述目标第二集群发送所述第一集群的资源空闲量。

第六方面，本申请提供了一种第二管理节点，其特征在于，所述第二管理节点包括：

处理模块，用于确定与目标第二集群具有关联关系的第一集群，其中，所述第二节点属于所述目标第二集群，所述目标第二集群对应于哈希环中的目标哈希值区间，所述第一集群对应的哈希值属于所述目标哈希值区间；

25 收发模块，用于接收任务数据；并向第一节点发送第一资源借调请求，所述第一借调请求用于指示所述第一节点将所述任务数据分配给所述第一集群处理。

在第六方面的一种可选设计中，所述处理模块，还用于基于收发模块向第一节点发送资源借调请求，确定所述第一集群处于资源不可借调状态。

30 在第六方面的一种可选设计中，与所述目标第二集群具有关联关系的集群有多个，所述处理模块，具体用于确定与所述目标第二集群具有关联关系的多个集群中的一个集群为第一集群。

在第六方面的一种可选设计中，所述处理模块，具体用于确定与所述目标第二集群具有关联关系的多个集群的资源空闲量中，与所述任务数据的资源需求量差值最小的资源空闲量对应的集群为所述第一集群。

35 在第六方面的一种可选设计中，所述收发模块，还用于接收第三节点发送的所述任务数据的资源需求量，并基于所述第一集群的资源空闲量大于或等于所述资源需求量，向所述第三节点发送所述第一集群的标识和第一集群的资源空闲量。

—7—

在第六方面的一种可选设计中，所述收发模块，具体用于接收所述第三节点发送的第二资源借调请求，所述第二资源借调请求包括所述任务数据，所述第二借调请求用于指示所述第二节点将所述任务数据分配给所述第一集群处理。

第七方面，本申请提供了一种第二管理节点，其特征在于，所述第二管理节点包括：

收发模块，用于接收第三节点发送的任务数据的资源需求量；

处理模块，用于确定与目标第二集群具有关联关系的第一集群，其中，所述第二节点属于所述目标第二集群，所述目标第二集群对应于哈希环中的目标哈希值区间，所述第一集群对应的哈希值属于所述目标哈希值区间；

所述收发模块，还用于基于所述第一集群的资源空闲量大于或等于所述资源需求量，向所述第三节点发送第三资源借调请求，所述第三资源借调请求用于指示所述第三节点被允许从所述第一集群申请资源。

在第七方面的一种可选设计中，与所述目标第二集群具有关联关系的集群有多个，所述处理模块，具体用于确定与所述目标第二集群具有关联关系的多个集群中的一个集群为第一集群。

第八方面，本申请提供了一种第三管理节点，其特征在于，所述第三管理节点包括：

收发模块，用于接收任务数据，向第二节点发送所述任务数据的资源需求量，接收所述第二节点发送的第三资源借调请求，所述第三资源借调请求用于指示所述第三节点被允许从第一集群申请资源，其中，所述第二节点属于目标第二集群，所述目标第二集群对应于哈希环中的目标哈希值区间，所述第一集群对应的哈希值属于所述目标哈希值区间，且所述资源空闲量大于或等于所述资源需求量，向第一节点发送第一资源借调请求，所述第一借调请求用于指示所述第一节点将所述任务数据分配给所述第一集群处理。

在第八方面的一种可选设计中，所述收发模块，具体用于向多个节点发送所述任务数据的资源需求量，所述第二节点为所述多个节点中的一个，接收所述多个节点发送的第三资源借调请求，所述第三资源借调请求用于指示所述第三节点被允许从多个子集群申请资源，所述第一集群为所述多个子集群中的一个。

在第八方面的一种可选设计中，所述第三管理节点还包括：

处理模块，用于确定所述多个子集群的资源空闲量中，与所述任务数据的资源需求量的差值最小的资源空闲量对应的子集群为所述第一集群。

第九方面，本申请提供了一种集群，所述集群包括多个计算节点，所述计算节点为任务数据提供所需的资源以执行所述任务数据，以及，所述集群还包括：上述任一方面所述的第一管理节点。

第十方面，本申请提供了一种集群，所述集群包括多个计算节点，所述计算节点为任务数据提供所需的资源以执行所述任务数据，以及，所述集群还包括：上述任一方面所述的第二管理节点。

第十一方面，本申请提供了一种集群，所述集群包括多个计算节点，所述计算节点为任务数据提供所需的资源以执行所述任务数据，以及，所述集群还包括：上述任一方面所述的第三管理节点。

—8—

第十二方面，本申请提供了一种分布式系统，其特征在于，所述分布式系统包括上述第九方面描述的集群以及上述第十方面描述的集群。

第十三方面，本申请提供了一种分布式系统，其特征在于，所述分布式系统包括上述第九方面描述的集群、上述第十方面描述的集群以及上述第十一方面描述的集群。

5 第十四方面，本申请提供了一种计算机存储介质，用于储存为上述第一管理节点所用的计算机软件指令，其包含用于执行上述方面所设计的程序。

第十五方面，本申请提供了一种计算机存储介质，用于储存为上述第二管理节点所用的计算机软件指令，其包含用于执行上述方面所设计的程序。

10 第十六方面，本申请提供了一种计算机存储介质，用于储存为上述第三管理节点所用的计算机软件指令，其包含用于执行上述方面所设计的程序。

本申请实施例中，第一节点确定第一集群对应的哈希值；第一节点获取哈希环，哈希环包括至少一个哈希值区间，每个哈希值区间对应一个第二集群；第一节点根据第一集群对应的哈希值确定对应的目标第二集群，其中，第一集群对应的哈希值属于目标哈希区间，目标哈希区间为至少一个哈希值区间中的一个哈希值区间，目标哈希区间对应目标第二集群；第一节点建立第一集群和目标第二集群的关联关系，其中，与所述第一集群建立了关联关系的所述目标第二集群被允许从所述第一集群申请资源。通过上述方式，第一节点可以通过第一集群对应的哈希值在哈希环（根据第二集群对应的哈希值生成的）上的分布，来确定第一集群对应的目标第二集群，建立目标第二集群和第一集群的关联关系，由于基于一致性哈希算法的哈希环上的第一集群对应的哈希值和第二集群对应的哈希节点的分布是确定的，每个空闲集群只允许被固定的空集群借调，因此，当目标第二集群需要借调第一集群的空闲资源时，不需要随机向选择的空闲集群发送资源查询请求，并在接收到返回后，再进行资源借调，可以直接向第一集群借调资源，减少了一次信息往返的时延开销。

15
20

附图说明

- 25 图 1 为一种一层架构的中心化分布式系统架构示意图；
图 2 为一种资源借调场景的架构示意图；
图 3a 为本申请实施例提供的一种资源分配方法的实施例的示意图；
图 3b 为一种集群闲忙状态分布示意图；
图 3c 为一种集群闲忙状态分布示意图；
30 图 3d 为本申请实施例中的一种装箱结果示意图；
图 4a 为本申请实施例提供的一种哈希环的结构示意图；
图 4b 为本申请实施例提供的一种哈希环的结构示意图；
图 4c 为本申请实施例提供的一种哈希环的结构示意图；
图 4d 为本申请实施例提供的一种哈希环的结构示意图；
35 图 5 为本申请实施例提供的一种资源借调方法的实施例的示意图；
图 6 为本申请实施例提供的一种资源借调方法的实施例的示意图；
图 7 为本申请实施例提供的一种第一管理节点的结构示意图；

图 8 为本申请实施例提供的一种第二管理节点的结构示意图；

图 9 为本申请实施例提供的一种第三管理节点的结构示意图；

图 10 为适用于本发明实施例的一种管理节点的系统实例；

图 11 为本申请实施例提供的一种集群的结构示意图；

5 图 12 为本申请实施例提供的一种集群的结构示意图；

图 13 为本申请实施例提供的一种集群的结构示意图；

图 14 为本申请实施例提供的一种分布式系统的结构示意图；

图 15 为本申请实施例提供的一种分布式系统的结构示意图。

10 具体实施方式

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，以便这里描述的本申请的实施例例如能够以除了在这里图示或描述的那些以外的顺序实施。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括，例如，包括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。

15 为使本发明的目的、技术方案和优点更加清楚，下面将结合附图对本发明实施方式作进一步地详细描述。

20 下面通过图 1 示例性的说明本发明实施例应用的的分布式系统的系统架构。

如图 1 所示的是一种一层架构的中心化分布式系统架构示意图。在该架构下，用户（user1、user2、……）可以将任务数据提交到节点集群的主节点（master）上，主节点通过调度器（scheduler）对任务数据进行调度分配，为任务数据分配所需要的资源，从而把任务分配到满足任务所需资源的从节点（slave1、slave2、……slaveN）上。

25 需要说明的是，本申请中的主节点（master）也可以被理解为管理节点，从节点（slave）可以被理解为计算节点。

在前述的多种架构中，调度器（scheduler）可以是集成在相关节点上的硬件设备，也可以通过节点的通用硬件通过软件实现。在本发明实施例中，不限制管理节点实施所述调度器功能的具体方式。

30 本发明实施例中所述的资源调度，是指将分布式中的计算节点的资源分配给任务数据。应当注意的是，在本发明的各类实施例中，计算节点指的是分布式系统中作为资源调度单位的一个资源集合，一般而言，计算节点是以一台服务器或者一个物理计算机为单位的。但是，在一些场景下，针对不同的资源类型，可以按照不同的划分单位来划分计算节点。例如，当以 CPU 作为被监测资源时，可以按照处理器或者处理核作为划分单位，分布式系统中每个处理器或者处理核作为一个计算节点，一个服务器中可能包含多个计算节点。又
35 例如，当仅以存储资源作为被监测资源时，可以按照数据分片作为计算节点的划分单位，则分布式数据库中一个数据分片作为一个计算节点，一个服务器中可能包含多个数据分片，

即包含多个计算节点。

随着人工智能 (artificial intelligence, AI) 和物联网 (internet of things, IoT) 的发展, 许多新兴计算密集型和延迟敏感型终端应用, 如虚拟现实 (virtual reality, VR)、增强现实 (augmented reality, AR) 以及高清视频图片处理等正在迅速兴起, 低延时、高带宽和强计算等应用需求日益增长。云计算技术受网络链路容量限制等因素影响, 容易导致数据传输速率过慢以及响应延迟过高等问题, 难以满足应用服务需求。利用靠近数据源的边缘设备进行计算以提高服务响应性能逐渐成为一种趋势。然而单个边缘设备算力小, 仅使用单个边缘设备执行任务难以胜任大规模任务需求。因此, 将多边缘设备的计算能力加以组合使用, 通过多边缘设备资源共享互助, 协同完成复杂任务的数据处理, 从而可以有效保障服务质量。

设备间的资源互助场景主要分为两类, 一类是层次化的异构设备的资源互助, 如边缘设备和边缘服务器、中央云处理器之间不同运算能力设备的协同; 另一类是同层设备间的资源协同, 即当边缘设备资源不足时, 可以向其他边缘设备借调资源。

参照图 2, 图 2 为一种资源借调场景的架构示意图, 如图 2 中示出的那样, 包括集群 201、集群 202、集群 203 和集群 204, 其中, 集群 201、集群 202、集群 203 为资源空闲量较多的空闲集群, 集群 204 为资源空闲量较少的忙集群, 集群 201 包括管理节点 2011 和多个计算节点 2012, 集群 202 包括管理节点 2021 和多个计算节点 2022, 集群 203 包括管理节点 2031 和多个计算节点 2032, 集群 204 包括管理节点 2041 和多个计算节点 2042。

在一种实施例中, 以集群 204 的空闲资源不足为例, 可以按照如下策略进行跨集群的资源借调:

管理节点 2041 对集群 201、集群 202、集群 203 按照资源空闲量的大小进行从大到小进行排序, 得到第一集群列表。

例如, 第一集群列表是 {集群 201, 集群 203, 集群 202}。

管理节点 2041 接收任务数据, 并确定多个计算节点 2042 可用的资源空闲量小于任务数据的资源需求量。

管理节点 2041 在第一集群列表中, 选择资源空闲量满足任务数据的资源需求的集群, 得到第二集群列表。

例如, 若集群 201、集群 202、集群 203 的资源空闲量都满足任务数据的资源需求, 则第二集群列表是 {集群 201, 集群 203, 集群 202}。

管理节点 2041 从第二集群列表中随机选择 2 个集群, 向选择的 2 个集群发送资源查询请求。

例如, 管理节点 2041 选择了集群 201 和集群 203, 此时, 管理节点 2041 可以向管理节点 2011 和管理节点 2021 发送资源查询请求。

管理节点 2041 选择第一个返回借调结果且满足任务数据的资源需求的集群, 进行跨集群的资源借调。

例如, 集群 201 为第一个返回借调结果且满足任务数据的资源需求的集群, 则管理节

点可以将任务数据发送给管理节点 2011, 使得管理节点 2011 调度多个计算节点处理任务数据。

然而, 在上述设计中, 由于管理节点 2041 在进行跨级群的资源借调时, 需要向随机选择的 k 个集群发送资源查询请求, 并在接收到返回借调结果后, 再进行资源借调, 因此, 资源借调前需要一次信息往返的时延开销, 增加了时延。

同时, 以集群为粒度进行资源借调, 对于集群数小、但集群中节点多的场景, 并发借调时, 借调冲突概率比较大, 即针对于同一个空闲集群, 可能存在多个忙集群进行资源借调。例如, 以上述实施例为例, 若集群 201 和集群 203 都正在被其他忙集群借调资源, 则集群 201 和集群 203 上会发生资源借调冲突, 导致没有一个集群返回借调结果。

为解决上述技术问题, 本申请提供了一种资源分配方法, 参照图 3a, 图 3a 为本申请实施例提供的一种资源分配方法的实施例的示意图, 如图 3a 中示出的那样, 本申请提供的资源分配方法包括:

301、第一节点确定第一集群对应的哈希值。

本申请实施例中, 第一节点可以是空闲集群的管理节点, 其中, 空闲集群包括第一节点和多个计算节点, 第一集群可以由多个计算节点中的全部节点或部分节点组成的节点集群, 第一集群也可以包括第一节点, 这里并不限定。

本申请实施例中, 第一节点在确定第一集群对应的哈希值之前, 可以先确定所在的集群的忙闲状态。

可选地, 在一种实施例中, 第一节点可以根据所在的集群下计算节点的主导资源的占用率评估节点的忙闲状态, 并根据所在的集群中忙节点个数和闲节点的个数评估集群的忙闲状态。

示例性的, 第一节点可以按照如下公式定义计算节点的忙闲状态:

$$P_i = \frac{Use(i)}{Total(i)};$$

其中, i 表示第 i 种资源, $Use(i)$ 表示正在使用的第 i 种资源的资源占用量, $Total(i)$ 表示第 i 种资源的总资源量;

$$state(\text{节点}) = \max\{P_i\} = \begin{cases} \geq a, & \text{节点为忙;} \\ < a, & \text{节点为闲;} \end{cases}$$

其中, a 为预设的阈值, $\max\{P_i\}$ 表示 i 个资源占用率中最大的。

本申请实施例中, 当计算节点中最大的资源占用率大于或等于阈值 a 时, 第一节点可以确定该计算节点为忙节点, 当计算节点中最大的资源占用率小于阈值 a 时, 第一节点可以确定该计算节点为空闲节点。

需要说明的是, 上述阈值 a 可以根据需求实时调整, 可根据不同场景进行配置, 本申请并不限定。

示例性的, 第一节点可以按照如下方式定义所在的集群的忙闲状态:

第一节点所在的集群中, 忙节点数量为 n, 闲节点数量为 m;

- 12 -

当 $m=1$ 时：即第一节点所在的集群内只有一个空闲节点时，第一节点可以确定所在的集群为忙集群；

当 $m \neq 1$ 时：计算 $p = \frac{m}{n}$ ， $state(\text{集群}) = p = \begin{cases} \geq b, \text{集群为闲} \\ < b, \text{集群为忙} \end{cases}$ ；

其中， b 为预设的阈值。

需要说明的是，上述阈值 b 可以根据需求实时调整，可根据不同场景进行配置，本申请并不限定。

参照图 3b 和图 3c，图 3b 和图 3c 是一种集群闲忙状态分布示意图，如图 3b 和图 3c 所示，在图 3b 中，集群内有 4 个忙节点，5 个空闲节点，即 $n=4$ ， $m=5$ ；

则 $p=1.25$ ，若确定阈值 b 为 0.3，由于 $1.25 > 0.3$ ，则图 3b 示出的集群为闲集群，可以提供空闲资源。

同理，图 3c 所示的集群内有 7 个忙节点，2 个空闲节点，即 $n=7$ ， $m=2$ ；

则 $p=0.286$ ，由于 $0.286 < 0.3$ ，则图 3c 示出的集群为忙集群。

本申请实施例中，第一节点在确定所在的集群的闲忙状态为闲时，可以确定第一集群对应的哈希值。

可选的，第一节点在所在的空闲集群包括多个计算节点，第一节点可以对空闲集群的多个计算节点进行划分，得到多个子集群。

进一步的，第一节点在所在的空闲集群包括多个空闲节点，第一节点可以对空闲集群的多个空闲节点进行划分，得到多个子集群。

在一种实施例中，多个子集群包括第三集群和第四集群，第三集群为多个子集群中资源空闲量最大的集群，第四集群为多个子集群中资源空闲量最小的集群，第三集群的资源空闲量与第四集群的资源空闲量的差值的绝对值小于第一阈值。

即，本申请实施例中，第一节点可以对空闲集群的多个计算节点进行划分，以使得划分后得到的多个子集群中的各个第一集群之间的资源空闲量的差异不会过大（最大的资源空闲量与最小的资源空闲量的差值的绝对值小于第一阈值）。

示例性的，第一节点可以在空闲集群的多个闲节点中随机选取一部分空闲节点或全部空闲节点作为预留资源，为空闲集群本地的忙节点提供资源；

本申请实施例中，第一节点可以确定空闲节点中除了作为预留资源之外的剩余空闲节点作为可借资源，并对其进行分组。

具体的，第一节点可以首先计算可借资源中的每个空闲节点的资源空闲率 $r_i = \alpha \times c_i + \beta \times m_i$ ，并根据资源空闲率将节点进行降序排序，其中， c_i 表示 CPU 资源空闲率， m_i 表示内存资源空闲率， α 和 β 表示 CPU 资源空闲率和内存资源空闲率的权重参数。

第一节点将可借资源分为 j 组，并将所有可借节点进行“s”型装箱操作，同时计算每个空闲资源组的总空闲率 $R_j = \sum r_i$ 以及 $\sigma_j = \max\{R_j\} - \min\{R_j\}$ ，迭代到 σ_j 大于 σ_{j-1} 时，

算法终止，将 $j-1$ 组作为分组结果，其中， n 表示可借节点个数，需要说明的是， j 为一个可配置参数，实际应用中可按照需求选择，这里并不限定。

例如，第一节点可以确定空闲节点中除了作为预留资源之外的剩余空闲节点包括：{节点 1，节点 2，节点 3，节点 4，节点 5，节点 6，节点 7，节点 8，节点 9，节点 10，节点 11，节点 12}，其中每个节点已经按照空闲率降序排序（即节点 1 空闲率最高，节点 12 空闲率最低），参照图 3d，图 3d 为本申请实施例中的一种装箱结果示意图，其中，当确定分
5 组数量 j 为 5 时，“S”型装箱结果示意图如图 3d 所示，在图 3d 中，划分后得到了 5 个子集群。

在另一种实施例中，所述多个子集群中的每个子集群的资源空闲量与预设资源空闲量的差值的绝对值小于或等于第二阈值，所述预设资源空闲量与历史资源请求量相关。

具体的，资源空闲量可以包括 CPU 资源空闲量和内存资源空闲量，预设资源空闲量可以包括预设 CPU 资源空闲量和预设内存资源空闲量，其中，预设资源空闲量与历史资源请求量相关，相应的，预设 CPU 资源空闲量与历史 CPU 资源请求量相关，预设内存资源空闲量与历史内存资源请求量相关。

本申请实施例中，所述多个子集群中的每个子集群的 CPU 资源空闲量与预设 CPU 资源空闲量的差值的绝对值小于或等于第三阈值，所述多个子集群中的每个子集群的内存资源空闲量与预设内存资源空闲量的差值的绝对值小于或等于第四阈值。

本申请实施例中预设 CPU 资源空闲量和预设内存资源空闲量可以根据历史借调请求中的资源请求量来确定，示例性的，第一节点可以记录所在的空闲集群最近 k 个周期内被请求的资源信息，其中 k 为可调参数。

可选的，第一节点可以首先根据历史资源请求表，分别对 k 个周期内的 CPU 资源请求量和内存资源请求量进行统计平均，得到单个周期内每个空闲资源组 CPU 历史平均请求量 $\overline{CPU}$ 和内存历史平均请求量 $\overline{MEM}$ ：

$$\overline{CPU} = \frac{\sum_{i=1}^k \frac{CPU_i}{s_i}}{k} ;$$

$$\overline{MEM} = \frac{\sum_{i=1}^k \frac{MEM_i}{s_i}}{k} ;$$

其中， CPU_i 表示第 i 个周期内的 CPU 资源请求总量， s_i 表示第 i 个周期内的空闲资源
25 组个数， MEM_i 表示第 i 个周期内的内存资源请求总量。

第一节点可以确定 CPU 历史平均请求量 $\overline{CPU}$ 为预设 CPU 资源空闲量，内存历史平均请求量 $\overline{MEM}$ 为预设内存资源量。

第一节点可以以 $\overline{CPU}$ 和 $\overline{MEM}$ 为指标，对可借资源进行分组，使得分组后的每个子集

群内的 CPU 资源空闲量和内存资源空闲量都接近于 $\overline{CPU}$ 和 $\overline{MEM}$ 。

第一节节点可以计算每个空闲节点的贡献率 $v_i = \alpha \times \frac{idle\_c_i}{CPU} + \beta \times \frac{idle\_m_i}{MEM}$ ，并根据贡献率将空闲节点进行降序排序，将可借资源分为 N 组，

$$s = \max \left\{ \left[\frac{sum\_cpu}{\overline{CPU}} \right], \left[\frac{sum\_mem}{\overline{MEM}} \right] \right\},$$

5 将所有可借节点进行“S”型装箱，得到多个子集群。

其中， $idle\_c_i$ 表示 CPU 空闲量和 $idle\_m_i$ 表示内存空闲量， α 和 β 表示 CPU 和内存的权重参数， sum_cpu 表示空闲集群内可借空闲 CPU 总量， sum_mem 表示空闲集群内可借空闲内存总量。

10 关于第一节节点如何将所有空闲节点进行“S”型装箱，得到多个子集群，可以参照上述实施例中的描述，这里不再赘述。

本申请实施例中，每隔预设的时间，集群内的管理节点可以确定各自所在的集群的闲忙状态，当管理节点确定所在的集为空闲集群时，可以按照上述方式对空闲集群中的全部或部分空闲节点进行划分，得到多个子集群。当管理节点确定所在的集为忙集群时，可以向其他管理节点广播自己的忙闲状态（忙状态）。

15 本申请实施例中，第一节节点可以确定第一集群对应的哈希值，具体的，第一节节点可以确定多个子集群中的每个子集群对应的哈希值，应当理解，第一集群中多个子集群中的一个。

20 具体的，第一集群可以唯一对应一个标识，例如可以用 IP 和 ID 来唯一标识第一集群，其中，IP 为第一节节点所在的空闲集群的标识，ID 为分组的标识，本申请实施例中，IP 可以为空闲集群所在的服务器的 IP 地址，IP 地址可以包含服务器的 IP 地址和端口号两项信息。举例来说，计算第一集群对应的哈希值可以是根据所在的服务器的 IP 地址、端口号信息以及用于标识分组的 ID，基于预设的哈希算法计算得到第一集群对应的哈希值。

25 第一节节点可以基于预设的哈希算法计算每个子集群对应的哈希值，具体的，第一节节点可以基于预设的哈希算法计算每个子集群的标识的哈希值，由于多个子集群的标识彼此不重复，则对应的哈希值也是不同的。

302、第一节节点获取哈希环，所述哈希环包括至少一个哈希值区间，每个哈希值区间对应一个第二集群。

30 本申请实施例中，哈希环可以理解为包括多个哈希节点的哈希值以及相邻两个哈希节点之间的哈希值区间的数据结构，该数据结构可以为列表，或者是数组的形式。本申请并不限定哈希值的具体形式，只要是可以反映至少一个第二集群的分布特征的数据结构都应落入本申请的保护范围。

在一种实施例中，第一节节点可以基于预设的哈希算法构建哈希环。

其中，在基于一致性哈希(hash)算法的分布式系统中，集群可以根据其分别对应的哈

希值而分布于一个数值空间逻辑环上，称为哈希环。该哈希环一般由管理节点生成并保存。在所述哈希环上的一段数据范围(如哈希环上1000~20000的hash值范围)称为一个哈希值区间。

具体的，当管理节点确定所在的集为忙集群时，可以向其他管理节点广播自己的忙闲状态（忙状态），相应的，第一节点可以确定分布式系统中包括的忙集群（至少一个第二集群）。

第一节点可以基于预设的哈希算法计算各个第二集群对应的哈希值，需要说明的是，上述预设的哈希算法可以为 jhash 哈希算法，也可以为其它类型的哈希算法，本申请实施例对此不作具体限定。第一节点可根据系统需要选择合适的哈希算法，只要计算第一集群的哈希值与计算第二集群的哈希值采用相同的哈希算法即可。

第一节点可以根据各个第二集群对应的哈希值生成哈希环，参照图4a，图4a为本申请实施例提供的一种哈希环的结构示意图，该哈希环的数据范围从0到 2^{32} ，哈希环中的每个大圆圈为根据第二集群的标识计算得到的哈希值，代表一个第二集群。第二集群根据各自对应的哈希值而分布在哈希环上的不同位置。如图4a所示，第二集群对应的哈希值将整个哈希环划分为与第二集群相同数量的哈希值区间，一个哈希值区间是哈希环上的一段哈希值的数据范围。

可选的，在一种实施例中，哈希环包括第一哈希节点和第二哈希节点，第一哈希节点对应目标第二集群，第一哈希节点为哈希环上第二哈希节点的相邻后继节点，目标哈希区间为第一哈希节点与第二哈希节点之间的哈希区间。

本申请实施例中，哈希环上的每个哈希值区间与一个第二集群对应。在一种可能的实现方式中，针对组成哈希环的任一哈希值区间，沿着第一方向从该哈希值区间开始查找，将查找到的第一个第二集群确定为该哈希值区间对应的第二集群。换句话说，若第二集群的哈希值为该哈希值区间在第一方向上的边界哈希值，则可认为该第二集群与该哈希值区间对应。其中，第一方向为哈希环的正方向，它可以是顺时针方向，也可以是逆时针方向，本申请实施例不作具体限定。示例性地，图4a中选取了顺时针方向为第一方向。

本申请实施例中，第一节点可以确定 X 个第二集群对应的 Y 个哈希值，其中，每个第二集群对应至少一个哈希值，X 正整数，Y 为大于或等于 X 的正整数。

在一种实施例中，Y 大于 X，即存在第二集群，该第二集群对应多个哈希值。

本申请实施例中，当第二集群的数量过少时，在哈希环上第二集群的分布会过于稀疏，参照图 4b，图 4b 为本申请实施例提供的一种哈希环的结构示意图，当第二集群的数量过少时，第二集群在哈希环上的分布会过于稀疏，因此，可以将第二集群对应哈希环上的多个节点，即引入虚拟节点，参照图 4c，图 4c 为本申请实施例提供的一种哈希环的结构示意图，如图 4c 中示出的那样，目标第二集群可以对应于哈希环上的两个哈希节点，通过引入虚拟节点，使得第二集群可以对应哈希环上的多个节点，进而第二集群在哈希环上的分布不会过于稀疏。关于如何通过引入虚拟节点来实现将第二集群对应哈希环上的多个节点可以参照现有方案中的实现，这里不再赘述。

本申请实施例中，第一节点根据 Y 个哈希值生成哈希环，其中，哈希环包括 Y 个哈希

节点，每个第二集群对应至少一个哈希节点，Y 个哈希节点将哈希环划分为 Y 个哈希值区间，每个哈希节点对应一个哈希值区间，即每个哈希节点对应一个第二集群，而一个第二集群可以对应哈希环上的多个哈希节点。

本申请实施例中，第一节点可以周期性的接收到其他集群广播的闲忙状态，并根据其他集群广播的闲忙状态周期性的更新哈希环。

可选的，在另一种实施例中，第一节点可以接收第二集群发送的哈希表。

本申请实施例中，第二集群中的管理节点（例如第二节点）可以接收到其他集群广播的闲忙状态，并根据其他集群广播的闲忙状态构建哈希环，关于第二集群如何根据其他集群广播的闲忙状态构建哈希环可以参照上述实施例中的描述，这里不再赘述。

303、第一节点根据第一集群对应的哈希值确定对应的目标第二集群，其中，第一集群对应的哈希值属于目标哈希区间，目标哈希区间为至少一个哈希值区间中的一个哈希值区间，目标哈希区间对应目标第二集群。

本申请实施例中，第一节点在确定第一集群对应的哈希值，以及获取到基于第二集群的哈希值生成的哈希环之后，可以根据第一集群对应的哈希值从哈希环中确定对应的目标第二集群。

本申请实施例中，由于哈希环上的每个哈希值区间对应一个第二集群，因此，第一节点可以确定第一集群对应的哈希值在哈希环上属于的哈希值区间为目标哈希区间，并确定该目标哈希区间对应的第二集群为目标第二集群。

参照图 4d，图 4d 为本申请实施例提供的一种哈希环的结构示意图，如图 4d 中示出的那样，哈希环包括：六个哈希节点{哈希节点 1，哈希节点 3，哈希节点 4，哈希节点 6，哈希节点 8，哈希节点 10}，其中，哈希节点 3 和哈希节点 8 对应第二集群 1，哈希节点 1 和哈希节点 4 对应第二集群 2，哈希节点 6 和哈希节点 10 对应第二集群 3。

哈希节点 1 与哈希节点 3 之间的哈希值区间对应于第二集群 1，哈希节点 3 与哈希节点 4 之间的哈希值区间对应于第二集群 2，哈希节点 4 与哈希节点 6 之间的哈希值区间对应于第二集群 3，哈希节点 6 与哈希节点 8 之间的哈希值区间对应于第二集群 1，哈希节点 8 与哈希节点 10 之间的哈希值区间对应于第二集群 3，哈希节点 10 与哈希节点 1 之间的哈希值区间对应于第二集群 2。

空闲集群 a 包括两个子集群{第一集群 a1，第一集群 a2}，空闲集群 b 包括两个子集群{第一集群 b1，第一集群 b2}。

各个第一集群{第一集群 a1，第一集群 a2，第一集群 b1，第一集群 b2}在哈希环上的分布如图 4 中示出的那样，其中，第一集群 a1 对应的哈希值属于哈希节点 8 与哈希节点 6 之间的哈希值区间，第一集群 a2 对应的哈希值属于哈希节点 4 与哈希节点 10 之间的哈希值区间，第一集群 b1 对应的哈希值属于哈希节点 1 与哈希节点 3 之间的哈希值区间，第一集群 b2 对应的哈希值属于哈希节点 4 与哈希节点 10 之间的哈希值区间。

结合上述各个第一集群在哈希环中的分布，空闲集群 a 中的第一节点可以确定第一集群 a1 对应于第二集群 6，空闲集群 a 中的第一节点可以确定第一集群 a2 对应于第二集群 4，空闲集群 b 中的第一节点可以确定第一集群 b1 对应于第二集群 3，空闲集群 b 中的第一节

点可以确定第一集群 b2 对应于第二集群 4。

304、所述第一节点建立第一集群和目标第二集群的关联关系，其中，与所述第一集群建立了关联关系的所述目标第二集群被允许从所述第一集群申请资源。

本申请实施例中，第一节点根据第一集群对应的哈希值确定对应的目标第二集群之后，可以向目标第二集群发送第一集群的标识，具体的，第一节点根据第一集群对应的哈希值确定对应的目标第二集群之后，可以向目标第二集群中的管理节点（第二节点）发送资源分配指示，该资源分配指示包括第一集群的标识，该资源分配指示可以指示目标第二节点被允许从所述第一集群申请资源。

可选的，在一种实施例中，第一节点还可以向目标第二集群发送第一集群的资源空闲量。

具体的，第一节点还可以向目标第二集群发送第一集群的 CPU 资源空闲量和内存资源空闲量。

一方面，第一节点可以通过第一集群对应的哈希值在哈希环（根据第二集群对应的哈希值生成的）上的分布，来确定第一集群对应的目标第二集群，建立目标第二集群和第一集群的关联关系，由于基于一致性哈希算法的哈希环上的第一集群对应的哈希值和第二集群对应的哈希节点的分布是确定的，每个空闲集群只允许被固定的空集群借调，因此，当目标第二集群需要借调第一集群的空闲资源时，不需要随机向选择的空闲集群发送资源查询请求，并在接收到返回后，再进行资源借调，可以直接向第一集群借调资源，减少了一次信息往返的时延开销。

另一方面，由于资源借调的对应关系是基于将第一集群的哈希值在基于忙集群生成的哈希环上的对应关系确定的，该对应关系是唯一的，因此空闲集群仅能被哈希环上对应的忙集群借调资源，因此，不会出现由于多个忙集群向同一个空闲集群并发借调资源而导致的借调冲突。

另一方面，第一节点通过对空闲集群的多个计算节点进行划分，得到多个子集群，之后可以以子集群为单位为其他忙集群借调资源，使得空闲集群中的空闲资源可以实现相对均衡的借调。

另一方面，当第二集群的数量过少时，第二集群在哈希环上的分布会过于稀疏，因此，可以将第二集群（例如目标第二集群）对应哈希环上的多个节点，即引入虚拟节点，使得不会出现由于第二集群在哈希环上分布过于稀疏，而导致一个第二集群会对应多个第一集群的情况，增加了资源分配的均衡性。

本申请实施例中，第一节点确定第一集群对应的哈希值；第一节点获取哈希环，哈希环包括至少一个哈希值区间，每个哈希值区间对应一个第二集群；第一节点根据第一集群对应的哈希值确定对应的目标第二集群，其中，第一集群对应的哈希值属于目标哈希区间，目标哈希区间为至少一个哈希值区间中的一个哈希值区间，目标哈希区间对应目标第二集群；第一节点建立第一集群和目标第二集群的关联关系，其中，与所述第一集群建立了关联关系的所述目标第二集群被允许从所述第一集群申请资源。通过上述方式，一方面，第一节点可以通过第一集群对应的哈希值在哈希环（根据第二集群对应的哈希值生成的）上

的分布，来确定第一集群对应的目标第二集群，建立目标第二集群和第一集群的关联关系，由于基于一致性哈希算法的哈希环上的第一集群对应的哈希值和第二集群对应的哈希节点的分布是确定的，每个空闲集群只允许被固定的空集群借调，因此，当目标第二集群需要借调第一集群的空闲资源时，不需要随机向选择的空闲集群发送资源查询请求，并在接收
5 到返回后，再进行资源借调，可以直接向第一集群借调资源，减少了一次信息往返的时延开销。且，由于资源借调的对应关系是基于将第一集群的哈希值在基于忙集群生成的哈希环上的对应关系确定的，该对应关系是唯一的，因此空闲集群仅能被哈希环上对应的忙集群借调资源，因此，不会出现由于多个忙集群向同一个空闲集群并发借调资源而导致的借调冲突。

10 需要说明的是，在另一种实施例，第一节点可以不基于第一集群对应的哈希值在哈希环上的分布确定对应的目标第二集群，而是通过随机分配的方式，或者顺序分配的方式确定对应的目标第二集群。

具体的，第一节点可以从分布式系统中的多个第二集群中随机选择一个第二集群作为目标第二集群。

15 具体的，可以对至少一个第二集群进行排序，并按照预设的顺序将多个子集群（包括第一集群）顺次挂载到相应的第二集群上（即建立各个子集群与第二集群的关联关系）。则，第一节点也可以通过上述方式确定对应的目标第二集群。

应当理解，实际应用中可以不限于上述实现方式，只要在资源借调关系分配时，分布式系统确定固定的资源分配关系（第一节点与某一个目标第二集群建立关联关系），都可以
20 实现，忙集群在进行资源借调时，不需要随机向选择的空闲集群发送资源查询请求，并在接收到返回后，再进行资源借调，可以直接向第一集群借调资源，减少了一次信息往返的时延开销。且，由于资源借调的对应关系是基于将第一集群的哈希值在基于忙集群生成的哈希环上的对应关系确定的，该对应关系是唯一的，因此空闲集群仅能被哈希环上对应的忙集群借调资源，因此，不会出现由于多个忙集群向同一个空闲集群并发借调资源而导致的
25 的借调冲突。

参照图 5，图 5 为本申请实施例提供的一种资源借调方法的实施例的示意图，如图 5 中示出的那样，本申请提供的资源借调方法包括：

501、第二节点确定与目标第二集群具有关联关系的第一集群，其中，所述第二节点属于所述目标第二集群，所述目标第二集群对应于哈希环中的目标哈希值区间，所述第一集群对应的哈希值属于所述目标哈希值区间。

本申请实施例中，第二节点属于目标第二集群，其中，目标第二集群为忙集群。具体的，第二节点可以接收到第一节点发送的资源分配指示，该资源分配指示包括第一集群的标识，该资源分配指示可以指示第二节点可以借调第一集群的空闲资源，关于第一节点如何向目标第二集群发送第一集群的标识可参照上述实施例，这里不再赘述。
35

本申请实施例中，第二节点可以接收到多个子集群的标识，即，与所述目标第二集群具有关联关系的集群有多个。

本实施例中，所述第二节点可以确定与所述目标第二集群具有关联关系的多个集群中的一个集群为第一集群。

可选的，在一种实施例中，第二节点还可以接收到多个子集群中每个子集群的资源空闲量。

502、第二节点接收任务数据。

本申请实施例中，第二节点可以接收到任务数据。

可选的，在一种实施例中，第二节点接收到任务数据之后，可以确定接收到的任务数据的资源需求量，并判断所在的忙集群（目标第二集群）中的空闲节点的资源空闲量，若所属的忙集群中的空闲节点的资源空闲量大于或等于任务数据的资源需求量，相当于第二节点可以确定本地的集群有足够的资源空闲量来处理该任务数据。若所属的忙集群中的空闲节点的资源空闲量小于任务数据的资源需求量，相当于第二节点可以确定本地的集群没有足够的资源空闲量来处理该任务数据。

具体的，若所属的忙集群中的空闲节点的 CPU 资源空闲量大于或等于任务数据的 CPU 资源需求量，且所属的忙集群中的空闲节点的内存资源空闲量大于或等于任务数据的内存资源需求量，相当于第二节点可以确定本地的集群有足够的资源空闲量来处理该任务数据。若所属的忙集群中的空闲节点的 CPU 资源空闲量小于任务数据的 CPU 资源需求量，且所属的忙集群中的空闲节点的内存资源空闲量小于任务数据的内存资源需求量，相当于第二节点可以确定本地的集群没有足够的资源空闲量来处理该任务数据。

503、第二节点向第一节点发送第一资源借调请求，所述第一借调请求用于指示所述第一节点将所述任务数据分配给所述第一集群处理。

若第二节点确定本地集群没有足够的资源空闲量来处理该任务数据，则需要向其他集群借调资源来执行该任务数据。

本申请实施例中，第二节点可以向第一节点发送第一资源借调请求，第一借调请求可以包括任务数据和第一集群的标识，相应的，第一节点接收到第一资源借调请求之后，可以将其中包括的任务数据分配给第一集群来处理。

在一种实施例中，第二节点可以接收到多个子集群的标识以及多个子集群的资源空闲量，此时第二节点需要从其中确定一个集群作为资源借调的对象，接下来论述第二节点如何从多个子集群中确定作为资源借调对象的第一集群。

本申请实施例中，第二节点可以维护一个映射表，该映射表中存储了多个子集群的标识，以及对应的资源空闲量，第二节点可以在映射表中查找多个子集群中资源空闲量大于或等于任务数据的资源需求量的集群集合，得到满足资源需求量的子集群集合 {slices 1, slices 2, …, slices n}，其中，slices n 表示第 n 个子集群的标识。

第二节点在得到子集群集合 {slices 1, slices 2, …, slices n} 之后，可以对子集群集合中的多个子集群进行优先级排序，其中，优先级排序的指标为子集群的资源空闲量与任务数据的资源需求量的接近程度，越接近则对应的子集群的优先级越高，第二节点可以选择优先级最高的子集群发起第一借调请求。

示例性的，第二节点可以维护如下的映射表：

第一集群标识	资源空闲量	
	CPU 资源空闲量	内存资源空闲量
slices1	0.5	0.45
slices2	0.45	0.45
slices3	0.63	0.6
slices4	0.4	0.46
slices5	0.55	0.65
slices6	0.65	0.75
slices7	0.74	0.7
slices8	0.68	0.55
slices9	0.52	0.65

第二节点接收到任务数据，并确定该任务数据对应的 CPU 资源需求量为 0.6，内存资源需求量为 0.65。

5 第二节点可以在映射表中查找多个第一集群中资源空闲量大于或等于任务数据的资源需求量的集群集合，得到满足资源需求量的第一集群集合 {slices6, slices7}。

第二节点可以根据如下公式来确定与任务数据的资源需求量最接近的第一集群：

$$\min_i \alpha \times (cpi_i^b - cpu_j^r) + \beta \times (mem_i^b - mem_j^r)$$

$$s.t. \quad \begin{aligned} &cpu_i^b \geq cpu_j^r \\ &mem_i^b \geq mem_j^r \end{aligned};$$

10 其中， α 表示 CPU 的权重参数， β 表示内存的权重参数， cpi_i^b 表示注册在第二集群 j 上的第一集群 i 的 CPU 资源空闲量， cpu_j^r 表示第二集群 j 中确定的任务数据需要的 CPU 资源空闲量， mem_i^b 表示注册在第二集群 j 上的第一集群 i 的空闲内存资源量， mem_j^r 表示第二集群 j 中确定的任务数据需要的空闲内存资源量。示例性的， α 为 0.6， β 为 0.4。

slices 6 与任务数据资源请求量的接近程度为：

$$(0.65-0.6) * 0.6 + (0.75-0.65) * 0.4 = 0.07$$

15 slices 7 与任务数据资源请求量的接近程度为：

$$(0.74-0.6) * 0.6 + (0.7-0.65) * 0.4 = 0.104$$

则第二节点可以确定 slices 6 对应的第一集群的资源空闲量最接近任务数据的资源需求，并向 slices 6 对应的第一集群发起第一借调请求。

需要说明的是，上述描述仅为一种示意，其中具体的参数选择仅为一种示意，本申请

并不限定。

通过上述方式，所述第二节点确定与所述目标第二集群具有关联关系的多个集群的资源空闲量中，与所述任务数据的资源需求量差值最小的资源空闲量对应的集群为所述第一集群，提高了资源借调时的资源利用率。

5 可选地，在一种实施例中，第二节点可以基于接收到第一节点发送的第一集群的标识，确定该第一集群处于资源可借调状态。

第二节点可以基于向第一节点发送资源借调请求，确定第一集群处于资源不可借调状态。

10 本申请实施例中，第二节点向第一节点发送资源借调请求之后，相当于已经占用了第一集群的空闲资源，则可以确定第一集群处于资源不可借调状态。可选地，第二节点可以在映射表中标注第一集群的资源借调状态为资源不可借状态。相应的，第一节点也可以确定第一集群处于资源不可借调状态，即确定第一集群不能为其他忙集群提供可用资源，从而确保了资源锁定，实现了跨集群资源的无冲突调度。

15 本申请实施例中，第一集群在处理完任务数据之后，可以将处理结果返回第一节点，第一节点将处理结果发送到第二节点。

20 本申请实施例中，第二节点确定与目标第二集群具有关联关系的第一集群，其中，所述第二节点属于所述目标第二集群，所述目标第二集群对应于哈希环中的目标哈希值区间，所述第一集群对应的哈希值属于所述目标哈希值区间；所述第二节点接收任务数据；所述第二节点向第一节点发送第一资源借调请求，所述第一借调请求用于指示所述第一节点将所述任务数据分配给所述第一集群处理。通过上述方式，一方面，当分布式系统中存在多个第一集群时，各个第一节点通过第一集群对应的哈希值在根据第二集群对应的哈希值生成的哈希环上的分布，来确定第一集群对应的目标第二集群，并指示目标第二集群可以借调第一集群的空闲资源，由于基于一致性哈希算法的哈希环上的第一集群对应的哈希值和第二集群对应的哈希节点的分布是确定的，且相对均衡，因此，当目标第二集群需要借调
25 第一集群的空闲资源时，不需要随机向选择的空闲集群发送资源查询请求，并在接收到返回后，再进行资源借调，而是直接确定需要借调资源的第一集群，并向第一节点发送第一资源借调请求，减少了一次信息往返的时延开销。且，由于资源借调的对应关系是基于将第一集群的哈希值在基于忙集群生成的哈希环上的对应关系确定的，该对应关系是唯一的，因此空闲集群仅能被哈希环上对应的忙集群借调资源，因此，不会出现由于多个忙集群向
30 同一个空闲集群并发借调资源而导致的借调冲突。

参照图 6，图 6 为本申请实施例提供的一种资源借调方法的实施例的示意图，如图 5 中示出的那样，本申请提供的资源借调方法包括：

35 601、第二节点确定与目标第二集群具有关联关系的第一集群，其中，所述第二节点属于所述目标第二集群，所述目标第二集群对应于哈希环中的目标哈希值区间，所述第一集群对应的哈希值属于所述目标哈希值区间。

步骤 601 的具体描述可参照图 5 对应的实施例中步骤 501 的具体描述，这里不再赘述。

602、第二节点接收第三节点发送的任务数据的资源需求量。

本申请实施例中，第三节点接收到任务数据之后，可以确定接收到的任务数据的资源需求量，并判断所属的忙集群中的空闲节点的资源空闲量，若所属的忙集群中的空闲节点的资源空闲量小于任务数据的资源需求量，即第三节点确定本地的集群没有足够的资源空闲量来处理该任务数据。

则第三节点可以遍历存储了多个子集群的标识，以及对应的资源空闲量的映射表，并在映射表中查找多个子集群中资源空闲量大于或等于任务数据的资源需求量的集群集合，若第三节点确定多个子集群中的每个子集群的资源空闲量都小于任务数据的资源需求量，或者确定多个子集群都为资源不可借调状态，则第三节点需要向其他忙集群借调资源。

具体的，第三节点可以从维护的哈希环中沿着第一方向确定第三节点对应的哈希节点的相邻后续节点，该相邻后续节点对应于第二集群。第三节点可以向第二集群的管理节点（第二节点）发送任务数据的资源需求量，具体的，第三节点可以向第二节点发送任务数据的资源需求量。

可选地，在一种实施例中，第三节点可以从维护的哈希环中沿着第一方向确定第三节点对应的哈希节点的多个后续节点，多个后续节点对应于多个第二集群。第三节点可以向多个第二集群的管理节点（第二节点）发送任务数据的资源需求量。

相应的，第二节点可以接收到第三节点发送的任务数据的资源需求量。

603、基于第一集群的资源空闲量大于或等于资源需求量，所述第二节点向第三节点发送第三资源借调请求，所述第三资源借调请求用于指示所述第三节点被允许从所述第一集群申请资源。

本申请实施例中，第二节点接收到资源需求量之后，可以查找自己维护的映射表，并确定资源空闲量大于或等于资源需求量的第一集群，并向第三节点发送第三资源借调请求，第三资源借调请求包括第一集群的标识，所述第三资源借调请求用于指示所述第三节点被允许从第一集群申请资源。

可选地，在一种实施例中，第二节点可以确定多个子集群的资源空闲量大于或等于资源需求量，则第二节点可以向第三节点发送确定的多个子集群的标识，其中第一集群为多个子集群中的一个。

可选地，在一种实施例中，第二节点可以确定多个子集群的资源空闲量大于或等于资源需求量，并从确定的多个子集群中随机选择一部分子集群，并向第三节点发送一部分子集群的标识。

可选地，在一种实施例中，第二节点可以接收到第三节点发送的第二资源借调请求，所述第二资源借调请求包括所述任务数据，所述第二借调请求用于指示所述第二节点将所述任务数据分配给所述第一集群处理。

本申请实施例中，第三节点接收到第二节点发送的第一集群的标识之后，相当于从第二节点获取到了可以进行资源借调的第一集群的标识。

可选地，在一种实施例中，第三节点可以接收到多个第二节点发送的多个子集群的标识，或者，第三节点可以接收到一个第二节点发送的多个子集群的标识之后，需要确定出

多个子集群中的一个集群作为资源借调对象，关于第三节点如何确定出多个第一集群中的一个第一集群作为资源借调对象，可以参照上述实施例中，第二节点如何确定出多个子集群中的第一集群作为资源借调对象的描述，这里不再赘述。

若第三节点确定多个子集群中的第一集群为资源借调对象，相应的，第二节点可以接收第三节点发送的第二资源借调请求，所述第二资源借调请求包括所述任务数据和第一集群的标识，所述第二借调请求用于指示所述第二节点将所述任务数据分配给所述第一集群处理。

本申请实施例中，第二节点可以向第一节点发送任务数据和第一集群的标识。

第一节点可以将所述任务数据分配给所述第一集群处理。

可选地，在另一种实施例中，所述第三节点在接收所述第二节点发送的第三资源借调请求之后，可以不向第二节点发送第二资源借调请求，来指示所述第二节点将所述任务数据分配给所述第一集群处理，而是直接向第一节点发送第一资源借调请求，所述第一借调请求用于指示所述第一节点将所述任务数据分配给所述第一集群处理，减少了一步信息的转发开销。

参照图 7，图 7 为本申请实施例提供的一种第一管理节点 700 的结构示意图，如图 7 中示出的那样，所述第一管理节点 700 包括：

处理模块 701，用于确定第一集群对应的哈希值；获取哈希环，所述哈希环包括至少一个哈希值区间，每个哈希值区间对应一个第二集群；根据所述第一集群对应的哈希值确定对应的目标第二集群，其中，所述第一集群对应的哈希值属于目标哈希区间，所述目标哈希区间为所述至少一个哈希值区间中的一个哈希值区间，所述目标哈希区间对应所述目标第二集群；

收发模块 702，用于建立第一集群和目标第二集群的关联关系，其中，与所述第一集群建立了关联关系的所述目标第二集群被允许从所述第一集群申请资源。

可选地，在一种实施例中，所述哈希环包括第一哈希节点和第二哈希节点，所述第一哈希节点对应所述目标第二集群，所述第一哈希节点为所述哈希环上所述第二哈希节点的相邻后继节点，所述目标哈希区间为所述第一哈希节点的哈希值与所述第二哈希节点的哈希值之间的哈希区间。

可选地，在一种实施例中，所述处理模块 701，还用于对包括多个计算节点的空闲集群进行划分，得到多个子集群，所述第一集群为所述多个子集群中的一个。

可选地，在一种实施例中，所述多个子集群包括第三子集群和第四子集群，所述第三子集群为所述多个子集群中资源空闲量最大的子集群，所述第四子集群为所述多个子集群中资源空闲量最小的集群，所述第三子集群的资源空闲量与所述第四子集群的资源空闲量的差值的绝对值小于第一阈值。

可选地，在一种实施例中，所述多个子集群中的每个子集群的资源空闲量与预设资源空闲量的差值的绝对值小于或等于第二阈值，所述预设资源空闲量与历史资源请求量相关。

可选地，在一种实施例中，所述收发模块 702 还用于：

接收所述目标第二集群发送的资源借调请求，所述资源借调请求包括任务数据和第一集群的标识；

根据所述第一集群的标识向所述第一集群发送所述任务数据。

可选地，在一种实施例中，所述哈希环包括 Y 个哈希节点，每个第二集群对应至少一个哈希节点，且所述 Y 个哈希节点的对应的哈希值和所述第一集群对应的哈希值为根据相同的哈希算法生成的，所述 Y 为正整数。

可选地，在一种实施例中，所述处理模块 701，还用于确定 X 个第二集群对应的 Y 个哈希值，其中，每个第二集群对应至少一个哈希值，所述 X 正整数，所述 Y 为大于或等于 X 的正整数；根据所述 Y 个哈希值生成所述哈希环，其中，所述哈希环包括 Y 个哈希节点，每个第二集群对应至少一个哈希节点，所述 Y 个哈希节点将所述哈希环划分为 Y 个哈希值区间，每个哈希节点对应一个哈希值区间。

可选地，在一种实施例中，所述收发模块 702，还用于接收所述第二集群发送的所述哈希环。

可选地，在一种实施例中，所述收发模块 702，还用于向所述目标第二集群发送所述第一集群的资源空闲量。

参照图 8，图 8 为本申请实施例提供的一种第二管理节点 800 的结构示意图，如图 8 中示出的那样，所述第二管理节点 800 包括：

处理模块 801，用于确定与目标第二集群具有关联关系的第一集群，其中，所述第二节点属于所述目标第二集群，所述目标第二集群对应于哈希环中的目标哈希值区间，所述第一集群对应的哈希值属于所述目标哈希值区间；

收发模块 802，用于接收任务数据；并向第一节点发送第一资源借调请求，所述第一借调请求用于指示所述第一节点将所述任务数据分配给所述第一集群处理。

可选地，在一种实施例中，所述处理模块 801，还用于基于收发模块向第一节点发送资源借调请求，确定所述第一集群处于资源不可借调状态。

可选地，在一种实施例中，与所述目标第二集群具有关联关系的集群有多个，所述处理模块 801，具体用于确定与所述目标第二集群具有关联关系的多个集群中的一个集群为第一集群。

可选地，在一种实施例中，所述处理模块 801，具体用于确定与所述目标第二集群具有关联关系的多个集群的资源空闲量中，与所述任务数据的资源需求量差值最小的资源空闲量对应的集群为所述第一集群。

可选地，在一种实施例中，所述收发模块 802，还用于接收第三节点发送的所述任务数据的资源需求量，并基于所述第一集群的资源空闲量大于或等于所述资源需求量，向所述第三节点发送所述第一集群的标识和第一集群的资源空闲量。

可选地，在一种实施例中，所述收发模块 802，具体用于接收所述第三节点发送的第二资源借调请求，所述第二资源借调请求包括所述任务数据，所述第二借调请求用于指示所述第二节点将所述任务数据分配给所述第一集群处理。

参照图 8，图 8 为本申请实施例提供的一种第二管理节点 800 的结构示意图，如图 8 中示出的那样，所述第二管理节点 800 包括：

收发模块 802，用于接收第三节点发送的任务数据的资源需求量；

处理模块 801，用于确定与目标第二集群具有关联关系的第一集群，其中，所述第二节点属于所述目标第二集群，所述目标第二集群对应于哈希环中的目标哈希值区间，所述第一集群对应的哈希值属于所述目标哈希值区间；

所述收发模块 802，还用于基于所述第一集群的资源空闲量大于或等于所述资源需求量，向所述第三节点发送第三资源借调请求，所述第三资源借调请求用于指示所述第三节点被允许从所述第一集群申请资源。

可选地，在一种实施例中，与所述目标第二集群具有关联关系的集群有多个，所述处理模块，具体用于确定与所述目标第二集群具有关联关系的多个集群中的一个集群为第一集群。

参照图 9，图 9 为本申请实施例提供的一种第三管理节点 900 的结构示意图，如图 9 中示出的那样，所述第三管理节点 900 包括：

收发模块 901，用于接收任务数据，向第二节点发送所述任务数据的资源需求量，接收所述第二节点发送的第三资源借调请求，所述第三资源借调请求用于指示所述第三节点被允许从第一集群申请资源，其中，所述第二节点属于目标第二集群，所述目标第二集群对应于哈希环中的目标哈希值区间，所述第一集群对应的哈希值属于所述目标哈希值区间，且所述第一集群的资源空闲量大于或等于所述资源需求量，向第一节点发送第一资源借调请求，所述第一借调请求用于指示所述第一节点将所述任务数据分配给所述第一集群处理。

可选地，在一种实施例中，所述收发模块 901，具体用于向多个节点发送所述任务数据的资源需求量，所述第二节点为所述多个节点中的一个，接收所述多个节点发送的第三资源借调请求，所述第三资源借调请求用于指示所述第三节点被允许从多个子集群申请资源，所述第一集群为所述多个子集群中的一个。

可选地，在一种实施例中，所述第三管理节点 900 还包括：

处理模块 902，用于确定所述多个子集群的资源空闲量中，与所述任务数据的资源需求量的差值最小的资源空闲量对应的子集群为所述第一集群。

图10说明了适用于本发明实施例的一种管理节点的系统实例，其中该管理节点可以为上述实施例中描述的第一管理节点、第二管理节点和第三管理节点中的任意一种。基于该实施例中的系统环境可以实现上一实施例中管理节点的各个逻辑模块的功能。该实施例只是一个适用于本发明的实例，并不试图建议对本发明所提供的管理节点的功能和结构构成任何限制。

本发明实施例以一种通用计算机系统环境作为示例来对管理节点进行说明。众所周知的，可适用于该管理节点还可以采用其他的硬件架构来实现类似的功能。包括并不限于，个人计算机，服务计算机，多处理器系统，基于微处理器的系统，可编程消费电器，网路

PC, 小型计算机, 大型计算机, 包括任何上述系统或设备的分布式计算环境, 等等。

参照图10, 实现本发明所举例的系统包括管理节点1000形式的通用计算设备。结合前述所述的系统场景及架构, 本实施例所述的管理节点可以为前述场景及架构中所说明的本发明实施例的执行主体。例如, 可以为主节点、管理节点或者去中心化架构中的任一节点。

5 管理节点1000的元件可以包括, 但并不限制于, 处理单元1020, 系统存储器1030, 和系统总线1010。系统总线将包括系统存储器的各种系统元件与处理单元1020相耦合。系统总线1010可以是几种类型总线结构中的任意一种总线, 这些总线可以包括存储器总线或存储器控制器, 外围总线, 和使用一种总线结构的局部总线。总线结构可以包括工业标准结构 (ISA) 总线, 微通道结构 (MCA) 总线, 扩展ISA (EISA) 总线, 视频电子标准协会 (VESA) 局域
10 总线, 以及外围器件互联 (PCI) 总线。

管理节点1000一般包括多种管理节点可读媒介。管理节点可读媒介可以是任何管理节点1000可有效访问的媒介, 并包括易失性或非易失性媒介, 以及可拆卸或非拆卸的媒介。例如, 但并不限制于, 管理节点可读媒介可以包括管理节点存储媒介和通讯媒介。管理管理节点可读指令, 数据结构, 程序模块或其他数据的信息的任何方法或技术来实现。管理
15 节点存储媒介包括, 但并不限制于, RAM, ROM, EEPROM, 闪存存储器或其他存储器技术, 或者硬盘存储、固态硬盘存储、光盘存储, 磁盘盒, 磁盘存储或其它存储设备, 或任何其它可以存储所要求信息和能够被管理节点1000访问的媒介。通讯媒介一般包括嵌入的计算机可读指令, 数据结构, 程序模块或在模块化数据信号 (例如, 载波或其他传输机制) 中的
20 其他数据, 并且还包括任何信息传递的媒介。术语“模块化数据信号”是指具有一个或多个信号特征组或采用对信号中的信息进行编码的方式来改变的信号。例如, 但并不限制, 通讯媒介包括诸如有线网络或直接有线连接的有线媒介, 和诸如声, RF红外和其它无线媒介的无线媒介。上述任何组合也应该包括在管理节点可读媒介的范围内。

系统存储器1030包括管理节点存储媒介, 它可以是易失性和非易失性存储器, 例如, 只读存储器 (ROM) 1031和随即存取存储器 (RAM) 1032。基本输入/输出系统1033 (BIOS) 一般存
25 储于ROM1031中, 包含着基本的例行程序, 它有助于在管理节点1010中各元件之间的信息传输。RAM 1032一般包含着数据和/或程序模块, 它可以被处理单元1020即时访问和/或立即操作。例如, 但并不限制于, 图10说明了操作系统1034, 应用程序1035, 其他程序模块1036和程序数据1037。

管理节点1000也可以包括其他可拆卸/非拆卸, 易失性/非易失性的管理节点存储媒介。
30 仅仅是一个实例, 图10说明了硬盘存储器1041, 它可以是非拆卸和非易失性的可读写磁媒介; 外部存储器1051, 它可以是可拆卸和非易失性的各类外部存储器, 例如光盘、磁盘、闪存或者移动硬盘等; 硬盘存储器1041一般是通过非拆卸存储接口 (例如, 接口1040) 与系统总线1010相连接, 外部存储器一般通过可拆卸存储接口 (例如, 接口1060) 与系统总线1010相连接。上述所讨论的以及图10所示的驱动器和它相关的管理节点存储媒介提供了管理节
35 点可读指令, 数据结构, 程序模块和管理节点1000的其它数据的存储。例如, 硬盘驱动器1041说明了用于存储操作系统1042, 应用程序1043, 其它应用程序1044以及程序数据1045。值得注意的是, 这些元件可以与操作系统1034, 应用程序1035, 其他程序模块1036, 以及

程序数据1037是相同的或者是不同的。

在本实施例中，前述实施例中的方法或者上一实施例中逻辑模块的功能可以通过存储在管理节点存储媒介中的代码或者可读指令，并由处理单元1020读取所述的代码或者可读指令从而执行所述方法。

5 用户可以通过各类输入设备1061管理节点1000输入命令和信息。各种输入设备经常都是通过用户输入接口1060与处理单元1020相连接，用户输入接口1060与系统总线相耦合，但也可以通过其他接口和总线结构相连接，例如，并行接口，或通用串行接口(USB)。显示设备1091也可以通过接口(例如，视频接口1090)与系统总线1010相连接。此外，诸如计算设备1000也可以包括各类外围输出设备1020，输出设备可以通过输出接口10100等来连接。

10 管理节点1000可以在使用逻辑连接着一个或多个计算设备，例如，远程计算机1071。远程计算节点包括管理节点，计算节点，服务器，路由器，网络PC，等同的设备或其它通用的网络结点，并且一般包括许多或所有与管理节点1000有关的上述所讨论的元件。结合前述图1所描述的架构中，远程计算节点可以是从节点、计算节点或者其他管理节点。在图10中所说明的逻辑连接包括局域网(LAN)和广域网(WAN)，也可以包括其它网络。通过逻辑连接，

15 管理节点可以与其他节点实现本发明中与其他主题之间的交互。例如，可以通过与用户的逻辑链接进行任务信息和数据的传输，从而获取用户的任务数据；通过和计算节点的逻辑链接进行资源数据的传输以及任务分配命令的传输，从而实现各个节点的资源信息的获取以及任务数据的分配。

20 参照图11，图11为本申请实施例提供的一种集群1100的结构示意图，如图11中示出的那样，所述集群1100包括多个计算节点1101，所述计算节点1101为任务数据提供所需的资源以执行所述任务数据，以及，所述集群1100还包括：图7对应的实施例中描述的第一管理节点700。

参照图12，图12为本申请实施例提供的一种集群的结构示意图，如图12中示出的那样，

25 所述集群1200包括多个计算节点1201，所述计算节点1201为任务数据提供所需的资源以执行所述任务数据，以及，所述集群1200还包括：图8对应的实施例中描述的第二管理节点800。

参照图13，图13为本申请实施例提供的一种集群的结构示意图，如图13中示出的那样，所述集群1300包括多个计算节点1301，所述计算节点1301为任务数据提供所需的资源以执行所述任务数据，以及，所述集群1300还包括：图9对应的实施例中描述的第三管理节点900。

30 参照图14，图14为本申请实施例提供的一种分布式系统1400的结构示意图，如图14中示出的那样，所述分布式系统1400包括图11对应的实施例中描述的集群1100以及图12对应的实施例中描述的集群1200。

参照图15，图15为本申请实施例提供的一种分布式系统的结构示意图，如图12中示出的那样，所述分布式系统1500包括图11对应的实施例中描述的集群1100、图12对应的实施

35 例中描述的集群1200以及图13对应的实施例中描述的集群1300。

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发明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（Digital Subscriber Line, 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储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。

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，可以存储在一个计算机可读存储介质中。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。

以上所述，以上实施例仅用以说明本申请的技术方案，而非对其限制；尽管参照前述实施例对本申请进行了详细的说明，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申请各实施例技术方案的精神和范围。

权 利 要 求

1、一种资源分配方法，其特征在于，所述方法包括：

第一节点确定第一集群对应的哈希值；

5 所述第一节点获取哈希环，所述哈希环包括至少一个哈希值区间，每个哈希值区间对应一个第二集群；

所述第一节点根据所述第一集群对应的哈希值确定对应的目标第二集群，其中，所述第一集群对应的哈希值属于目标哈希区间，所述目标哈希区间为所述至少一个哈希值区间中的一个哈希值区间，所述目标哈希区间对应所述目标第二集群；

10 所述第一节点建立第一集群和目标第二集群的关联关系，其中，与所述第一集群建立了关联关系的所述目标第二集群被允许从所述第一集群申请资源。

2、根据权利要求1所述的方法，其特征在于，所述哈希环包括第一哈希节点和第二哈希节点，所述第一哈希节点对应所述目标第二集群，所述第一哈希节点为所述哈希环上所述第二哈希节点的相邻后继节点，所述目标哈希区间为所述第一哈希节点的哈希值与所述第二哈希节点的哈希值之间的哈希区间。

15 3、根据权利要求1或2所述的方法，其特征在于，所述方法还包括：

所述第一节点对包括多个计算节点的空闲集群进行划分，得到多个子集群，所述第一集群为所述多个子集群中的一个。

4、根据权利要求3所述的方法，其特征在于，所述多个子集群包括第三子集群和第四子集群，所述第三子集群为所述多个子集群中资源空闲量最大的子集群，所述第四子集群为所述多个子集群中资源空闲量最小的集群，所述第三子集群的资源空闲量与所述第四子集群的资源空闲量的差值的绝对值小于第一阈值。

5、根据权利要求3所述的方法，其特征在于，所述多个子集群中的每个子集群的资源空闲量与预设资源空闲量的差值的绝对值小于或等于第二阈值，所述预设资源空闲量与历史资源请求量相关。

25 6、根据权利要求1至5任一所述的方法，其特征在于，所述方法还包括：

所述第一节点接收所述目标第二集群发送的第一资源借调请求，所述第一资源借调请求包括任务数据和第一集群的标识；

所述第一节点根据所述第一集群的标识向所述第一集群发送所述任务数据。

7、根据权利要求1至6任一所述的方法，其特征在于，所述哈希环包括Y个哈希节点，30 每个第二集群对应至少一个哈希节点，且所述Y个哈希节点的对应的哈希值和所述第一集群对应的哈希值为根据相同的哈希算法生成的，所述Y为正整数。

8、根据权利要求1至7任一所述的方法，其特征在于，所述第一节点获取哈希环，包括：

35 所述第一节点确定X个第二集群对应的Y个哈希值，其中，每个第二集群对应至少一个哈希值，所述X正整数，所述Y为大于或等于X的正整数；

所述第一节点根据所述Y个哈希值生成所述哈希环，其中，所述哈希环包括Y个哈希节点，每个第二集群对应至少一个哈希节点，所述Y个哈希节点将所述哈希环划分为Y个

哈希值区间，每个哈希节点对应一个哈希值区间。

9、根据权利要求 1 至 7 任一所述的方法，其特征在于，所述第一节点获取哈希环，包括：

所述第一节点接收所述第二集群发送的所述哈希环。

10、根据权利要求 1 至 9 任一所述的方法，其特征在于，所述方法还包括：

所述第一节点向所述目标第二集群发送所述第一集群的资源空闲量。

11、一种资源借调方法，其特征在于，所述方法包括：

第二节点确定与目标第二集群具有关联关系的第一集群，其中，所述第二节点属于所述目标第二集群，所述目标第二集群对应于哈希环中的目标哈希值区间，所述第一集群对应的哈希值属于所述目标哈希值区间；

所述第二节点接收任务数据；

所述第二节点向第一节点发送第一资源借调请求，所述第一资源借调请求用于指示所述第一节点将所述任务数据分配给所述第一集群处理。

12、根据权利要求 11 所述的方法，其特征在于，所述方法还包括：

所述第二节点基于向第一节点发送第一资源借调请求，确定所述第一集群处于资源不可借调状态。

13、根据权利要求 12 所述的方法，其特征在于，所述方法还包括：

所述第二节点接收第一节点发送的第一集群的资源空闲量。

14、根据权利要求 11 至 13 任一所述的方法，其特征在于，与所述目标第二集群具有关联关系的集群有多个，所述第二节点确定与目标第二集群具有关联关系的第一集群，包括：

所述第二节点确定与所述目标第二集群具有关联关系的多个集群中的一个集群为第一集群。

15、根据权利要求 14 所述的方法，其特征在于，所述第二节点确定与所述目标第二集群具有关联关系的多个集群中的一个集群为第一集群，包括：

所述第二节点确定与所述目标第二集群具有关联关系的多个集群的资源空闲量中，与所述任务数据的资源需求量差值最小的资源空闲量对应的集群为所述第一集群。

16、根据权利要求 11 至 15 任一所述的方法，其特征在于，所述方法还包括：

所述第二节点接收第三节点发送的所述任务数据的资源需求量；

基于所述第一集群的资源空闲量大于或等于所述资源需求量，所述第二节点向所述第三节点发送所述第一集群的标识和第一集群的资源空闲量。

17、根据权利要求 16 所述的方法，其特征在于，所述第二节点接收任务数据包括：

所述第二节点接收所述第三节点发送的第二资源借调请求，所述第二资源借调请求包括所述任务数据，所述第二借调请求用于指示所述第二节点将所述任务数据分配给所述第一集群处理。

18、一种资源借调方法，其特征在于，包括：

第二节点接收第三节点发送的任务数据的资源需求量；

- 31 -

所述第二节点确定与目标第二集群具有关联关系的第一集群，其中，所述第二节点属于所述目标第二集群，所述目标第二集群对应于哈希环中的目标哈希值区间，所述第一集群对应的哈希值属于所述目标哈希值区间；

基于所述第一集群的资源空闲量大于或等于所述资源需求量，所述第二节点向所述第三节点发送第三资源借调请求，所述第三资源借调请求用于指示所述第三节点被允许从所述第一集群申请资源。

19、根据权利要求 18 所述的方法，其特征在于，与所述目标第二集群具有关联关系的集群有多个，所述第二节点确定与目标第二集群具有关联关系的第一集群，包括：

所述第二节点确定与所述目标第二集群具有关联关系的多个集群中的一个集群为第一集群。

20、一种资源借调方法，其特征在于，包括：

第三节点接收任务数据；

所述第三节点向第二节点发送所述任务数据的资源需求量；

所述第三节点接收所述第二节点发送的第三资源借调请求，所述第三资源借调请求用于指示所述第三节点被允许从第一集群申请资源，其中，所述第二节点属于目标第二集群，所述目标第二集群对应于哈希环中的目标哈希值区间，所述第一集群对应的哈希值属于所述目标哈希值区间，且所述第一集群的资源空闲量大于或等于所述资源需求量；

所述第三节点向第一节点发送第一资源借调请求，所述第一借调请求用于指示所述第一节点将所述任务数据分配给所述第一集群处理。

21、根据权利要求 20 所述的方法，其特征在于，所述第三节点向第二节点发送所述任务数据的资源需求量，包括：

所述第三节点向多个节点发送所述任务数据的资源需求量，所述第二节点为所述多个节点中的一个；

所述第三节点接收所述第二节点发送的第三资源借调请求，包括：

所述第三节点接收所述多个节点发送的第三资源借调请求，所述第三资源借调请求用于指示所述第三节点被允许从多个子集群申请资源，所述第一集群为所述多个子集群中的一个。

22、根据权利要求 21 所述的方法，其特征在于，所述方法还包括：

所述第三节点确定所述多个子集群的资源空闲量中，与所述任务数据的资源需求量的差值最小的资源空闲量对应的子集群为所述第一集群。

23、一种第一管理节点，其特征在于，所述第一管理节点包括：

处理模块，用于确定第一集群对应的哈希值；获取哈希环，所述哈希环包括至少一个哈希值区间，每个哈希值区间对应一个第二集群；根据所述第一集群对应的哈希值确定对应的目标第二集群，其中，所述第一集群对应的哈希值属于目标哈希区间，所述目标哈希区间为所述至少一个哈希值区间中的一个哈希值区间，所述目标哈希区间对应所述目标第二集群；

收发模块，用于建立第一集群和目标第二集群的关联关系，其中，与所述第一集群建

立了关联关系的所述目标第二集群被允许从所述第一集群申请资源。

24、根据权利要求 23 所述的第一管理节点，其特征在于，所述哈希环包括第一哈希节点和第二哈希节点，所述第一哈希节点对应所述目标第二集群，所述第一哈希节点为所述哈希环上所述第二哈希节点的相邻后继节点，所述目标哈希区间为所述第一哈希节点的哈希值与所述第二哈希节点的哈希值之间的哈希区间。

25、根据权利要求 23 或 24 所述的第一管理节点，其特征在于，所述处理模块，还用于对包括多个计算节点的空闲集群进行划分，得到多个子集群，所述第一集群为所述多个子集群中的一个。

26、根据权利要求 25 所述的第一管理节点，其特征在于，所述多个子集群包括第三子集群和第四子集群，所述第三子集群为所述多个子集群中资源空闲量最大的子集群，所述第四子集群为所述多个子集群中资源空闲量最小的集群，所述第三子集群的资源空闲量与所述第四子集群的资源空闲量的差值的绝对值小于第一阈值。

27、根据权利要求 25 所述的第一管理节点，其特征在于，所述多个子集群中的每个子集群的资源空闲量与预设资源空闲量的差值的绝对值小于或等于第二阈值，所述预设资源空闲量与历史资源请求量相关。

28、根据权利要求 23 至 27 任一所述的第一管理节点，其特征在于，所述收发模块还用于：

接收所述目标第二集群发送的第一资源借调请求，所述第一资源借调请求包括任务数据和第一集群的标识；

根据所述第一集群的标识向所述第一集群发送所述任务数据。

29、根据权利要求 23 至 28 任一所述的第一管理节点，其特征在于，所述哈希环包括 Y 个哈希节点，每个第二集群对应至少一个哈希节点，且所述 Y 个哈希节点的对应的哈希值和所述第一集群对应的哈希值为根据相同的哈希算法生成的，所述 Y 为正整数。

30、根据权利要求 23 至 29 任一所述的第一管理节点，其特征在于，所述处理模块，还用于确定 X 个第二集群对应的 Y 个哈希值，其中，每个第二集群对应至少一个哈希值，所述 X 正整数，所述 Y 为大于或等于 X 的正整数；根据所述 Y 个哈希值生成所述哈希环，其中，所述哈希环包括 Y 个哈希节点，每个第二集群对应至少一个哈希节点，所述 Y 个哈希节点将所述哈希环划分为 Y 个哈希值区间，每个哈希节点对应一个哈希值区间。

31、根据权利要求 23 至 29 任一所述的第一管理节点，其特征在于，所述收发模块，还用于接收所述第二集群发送的所述哈希环。

32、根据权利要求 23 至 31 任一所述的第一管理节点，其特征在于，所述收发模块，还用于向所述目标第二集群发送所述第一集群的资源空闲量。

33、一种第二管理节点，其特征在于，所述第二管理节点包括：

处理模块，用于确定与目标第二集群具有关联关系的第一集群，其中，所述第二节点属于所述目标第二集群，所述目标第二集群对应于哈希环中的目标哈希值区间，所述第一集群对应的哈希值属于所述目标哈希值区间；

收发模块，用于接收任务数据，并向第一节点发送第一资源借调请求，所述第一借调

请求用于指示所述第一节点将所述任务数据分配给所述第一集群处理。

34、根据权利要求 33 所述的第二管理节点，其特征在于，所述处理模块，还用于基于收发模块向第一节点发送第一资源借调请求，确定所述第一集群处于资源不可借调状态。

35、根据权利要求 33 或 34 所述的第二管理节点，其特征在于，与所述目标第二集群具有关联关系的集群有多个，所述处理模块，具体用于确定与所述目标第二集群具有关联关系的多个集群中的一个集群为第一集群。

36、根据权利要求 35 所述的第二管理节点，其特征在于，所述处理模块，具体用于确定与所述目标第二集群具有关联关系的多个集群的资源空闲量中，与所述任务数据的资源需求量差值最小的资源空闲量对应的集群为所述第一集群。

37、根据权利要求 33 至 36 任一所述的第二管理节点，其特征在于，所述收发模块，还用于接收第三节点发送的所述任务数据的资源需求量，并基于所述第一集群的资源空闲量大于或等于所述资源需求量，向所述第三节点发送所述第一集群的标识和所述第一集群的资源空闲量。

38、根据权利要求 37 所述的第二管理节点，其特征在于，所述收发模块，具体用于接收所述第三节点发送的第二资源借调请求，所述第二资源借调请求包括所述任务数据，所述第二借调请求用于指示所述第二节点将所述任务数据分配给所述第一集群处理。

39、一种第二管理节点，其特征在于，所述第二管理节点包括：

收发模块，用于接收第三节点发送的任务数据的资源需求量；

处理模块，用于确定与目标第二集群具有关联关系的第一集群，其中，所述第二节点属于所述目标第二集群，所述目标第二集群对应于哈希环中的目标哈希值区间，所述第一集群对应的哈希值属于所述目标哈希值区间；

所述收发模块，还用于基于所述第一集群的资源空闲量大于或等于所述资源需求量，向所述第三节点发送第三资源借调请求，所述第三资源借调请求用于指示所述第三节点被允许从所述第一集群申请资源。

40、根据权利要求 39 所述的第二管理节点，其特征在于，与所述目标第二集群具有关联关系的集群有多个，所述处理模块，具体用于确定与所述目标第二集群具有关联关系的多个集群中的一个集群为第一集群。

41、一种第三管理节点，其特征在于，所述第三管理节点包括：

收发模块，用于接收任务数据，向第二节点发送所述任务数据的资源需求量，接收所述第二节点发送的第三资源借调请求，所述第三资源借调请求用于指示所述第三节点被允许从第一集群申请资源，其中，所述第二节点属于目标第二集群，所述目标第二集群对应于哈希环中的目标哈希值区间，所述第一集群对应的哈希值属于所述目标哈希值区间，且所述第一集群的资源空闲量大于或等于所述资源需求量，向第一节点发送第一资源借调请求，所述第一借调请求用于指示所述第一节点将所述任务数据分配给所述第一集群处理。

42、根据权利要求 41 所述的第三管理节点，其特征在于，所述收发模块，具体用于向多个节点发送所述任务数据的资源需求量，所述第二节点为所述多个节点中的一个，接收所述多个节点发送的第三资源借调请求，所述第三资源借调请求用于指示所述第三节点被

允许从多个子集群申请资源，所述第一集群为所述多个子集群中的一个。

43、根据权利要求 42 所述的第三管理节点，其特征在于，所述第三管理节点还包括：处理模块，用于确定所述多个子集群的资源空闲量中，与所述任务数据的资源需求量的差值最小的资源空闲量对应的子集群为所述第一集群。

5 44、一种集群，其特征在于，所述集群包括多个计算节点，所述计算节点为任务数据提供所需的资源以执行所述任务数据，以及，所述集群还包括：如权利要求23-32中任意一项权利要求所述的第一管理节点。

10 45、一种集群，其特征在于，所述集群包括多个计算节点，所述计算节点为任务数据提供所需的资源以执行所述任务数据，以及，所述集群还包括：如权利要求33-40中任意一项权利要求所述的第二管理节点。

46、一种集群，其特征在于，所述集群包括多个计算节点，所述计算节点为任务数据提供所需的资源以执行所述任务数据，以及，所述集群还包括：如权利要求41至43中任意一项权利要求所述的第三管理节点。

15 47、一种分布式系统，其特征在于，所述分布式系统包括如权利要求44所述的集群和权利要求45所述的集群。

48、一种分布式系统，其特征在于，所述分布式系统包括如权利要求 44 所述的集群、权利要求 45 所述的集群以及如权利要求 46 所述的集群。

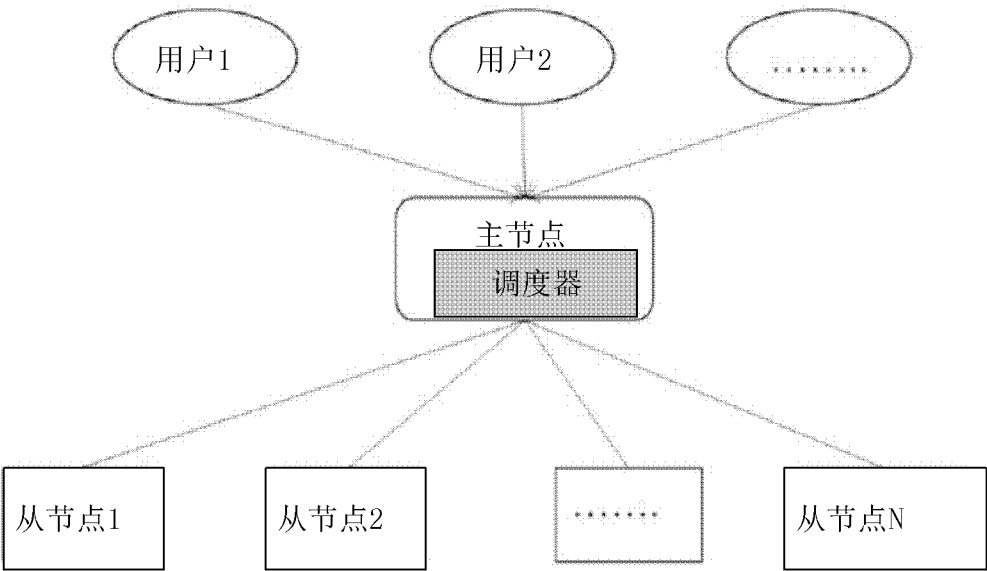


图 1

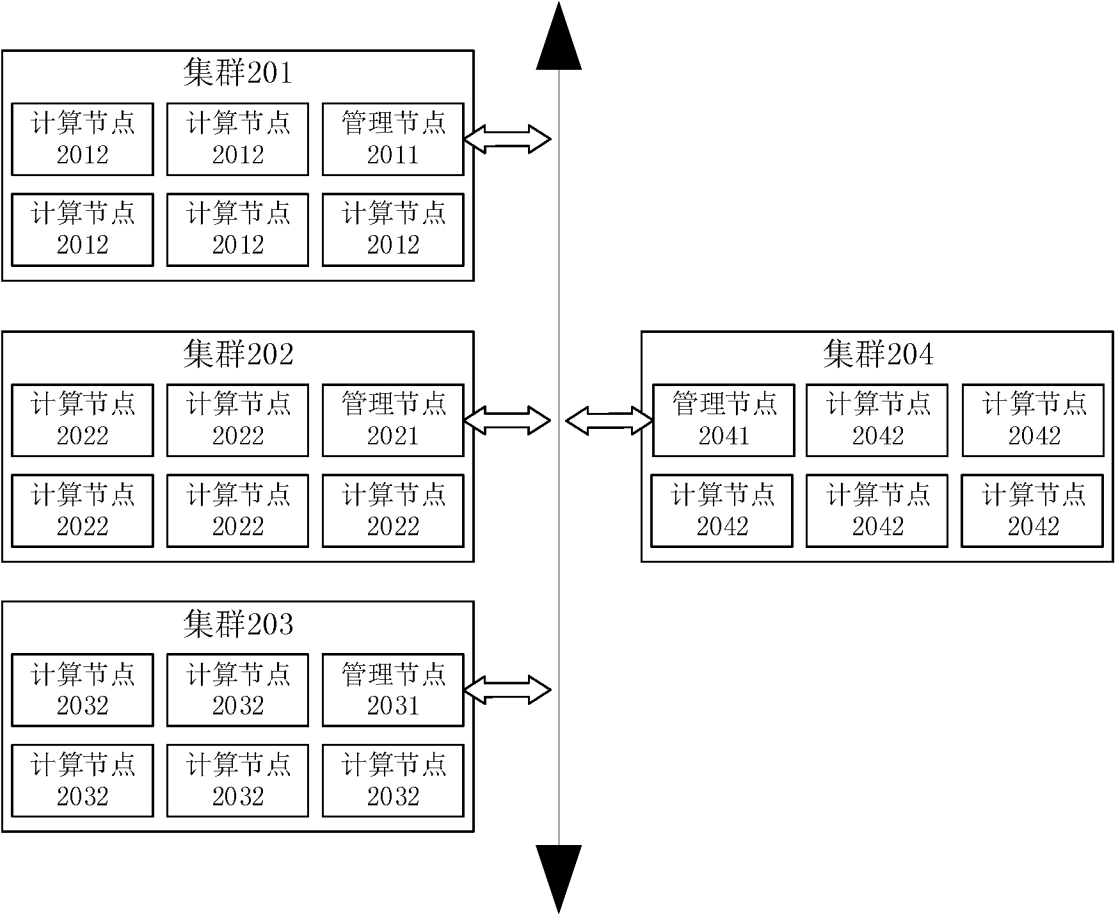


图 2

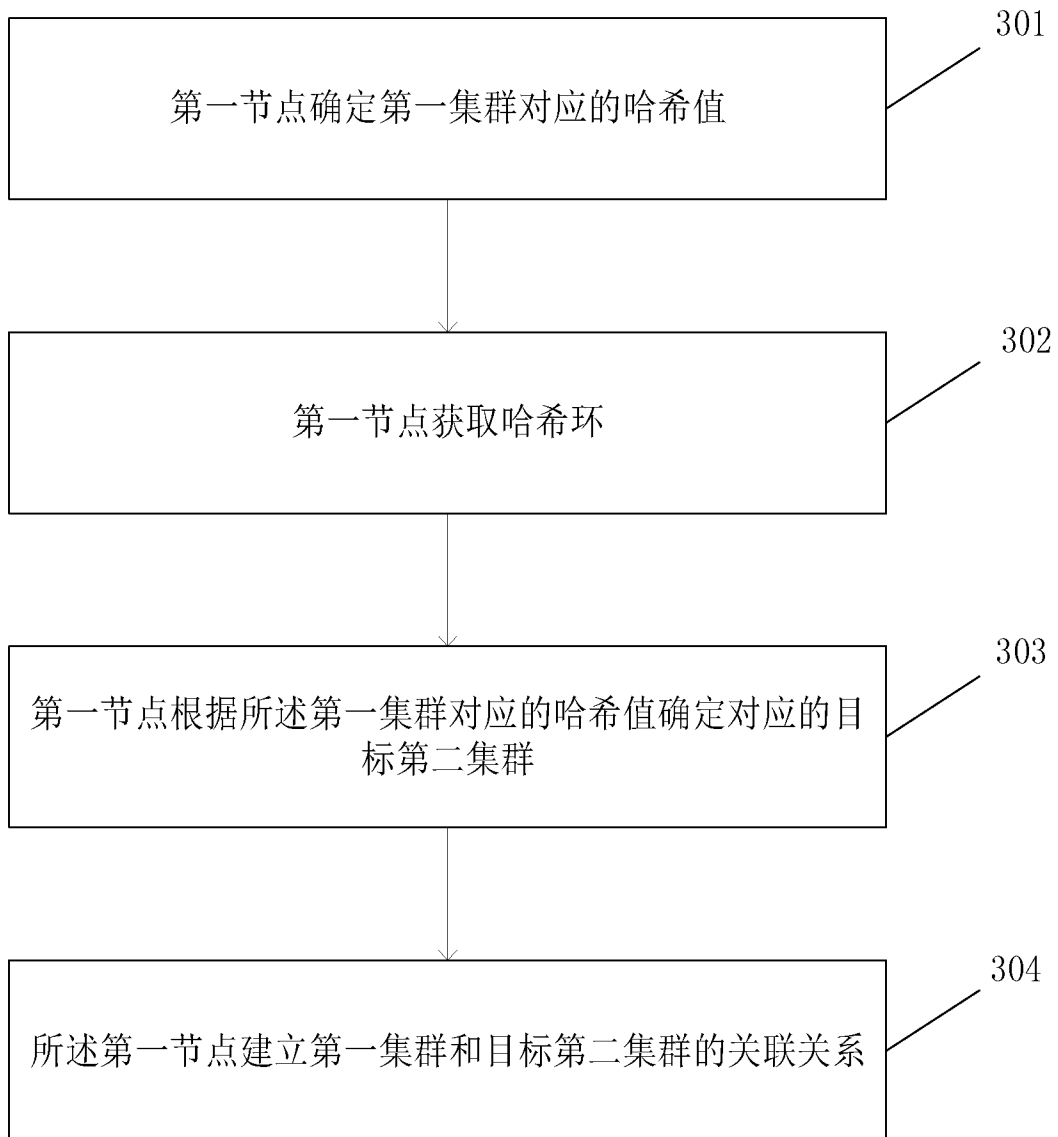


图 3a

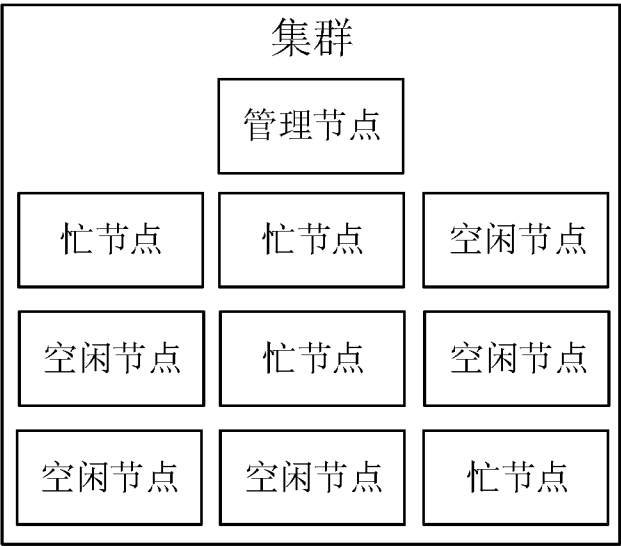


图 3b

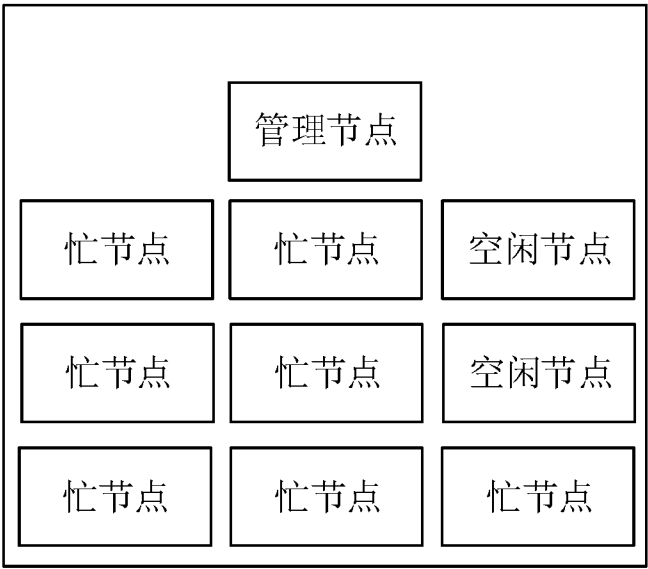


图 3c

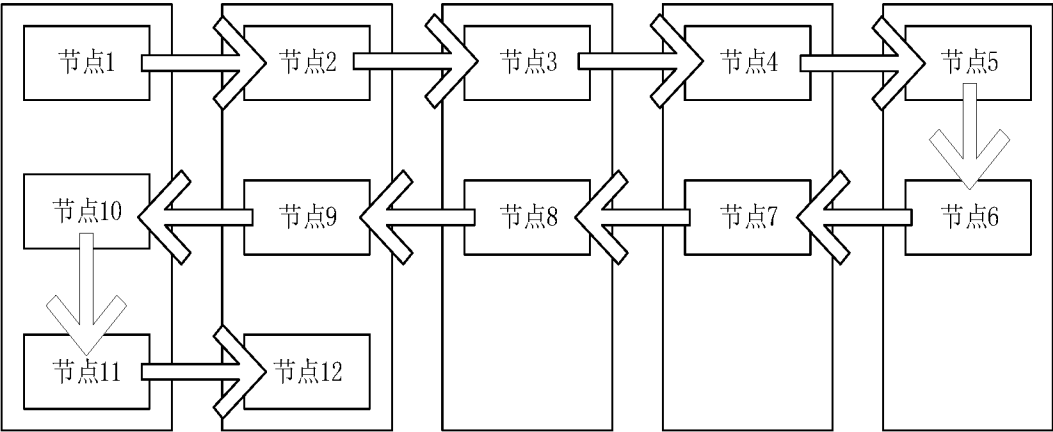


图 3d

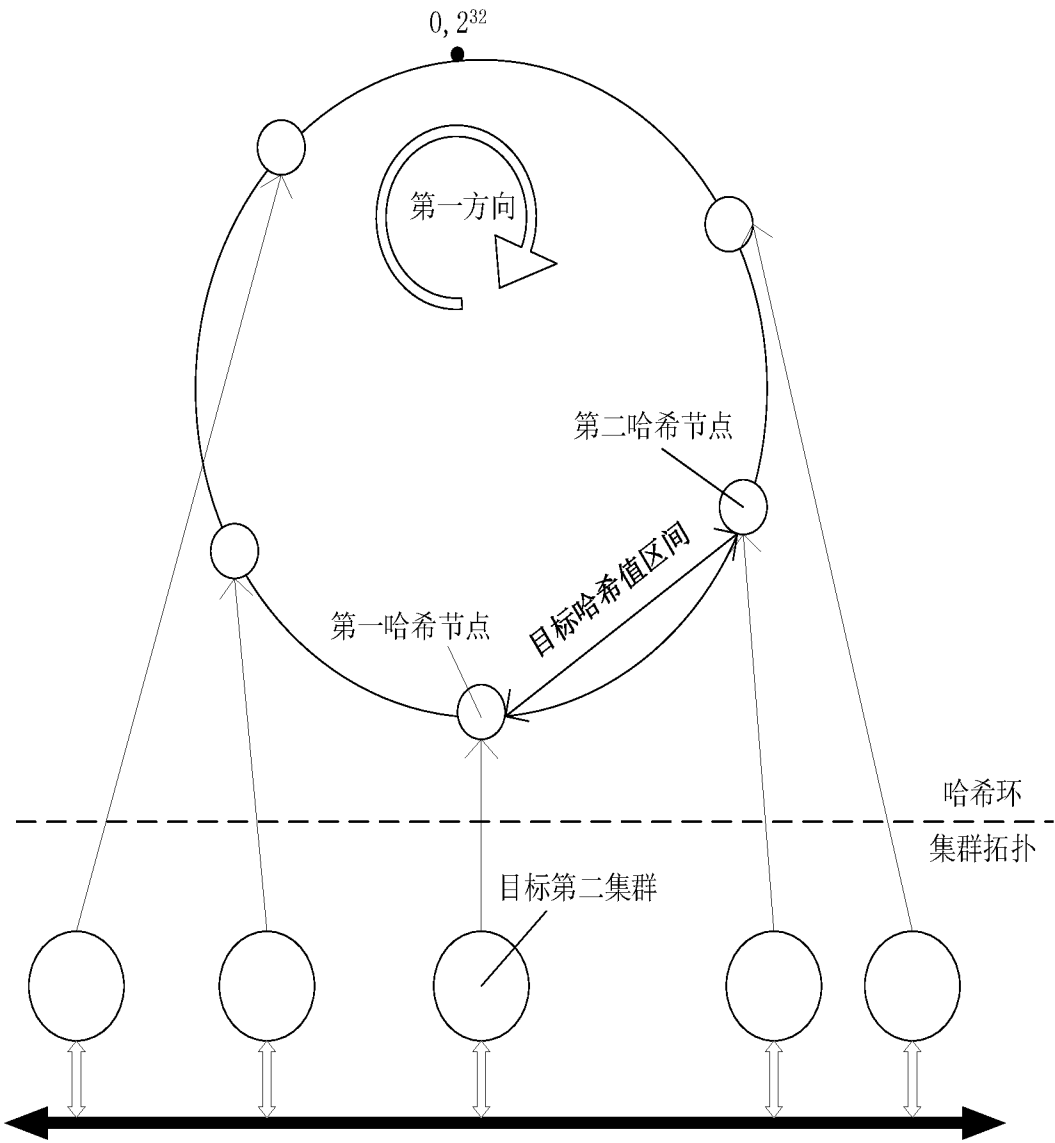


图 4a

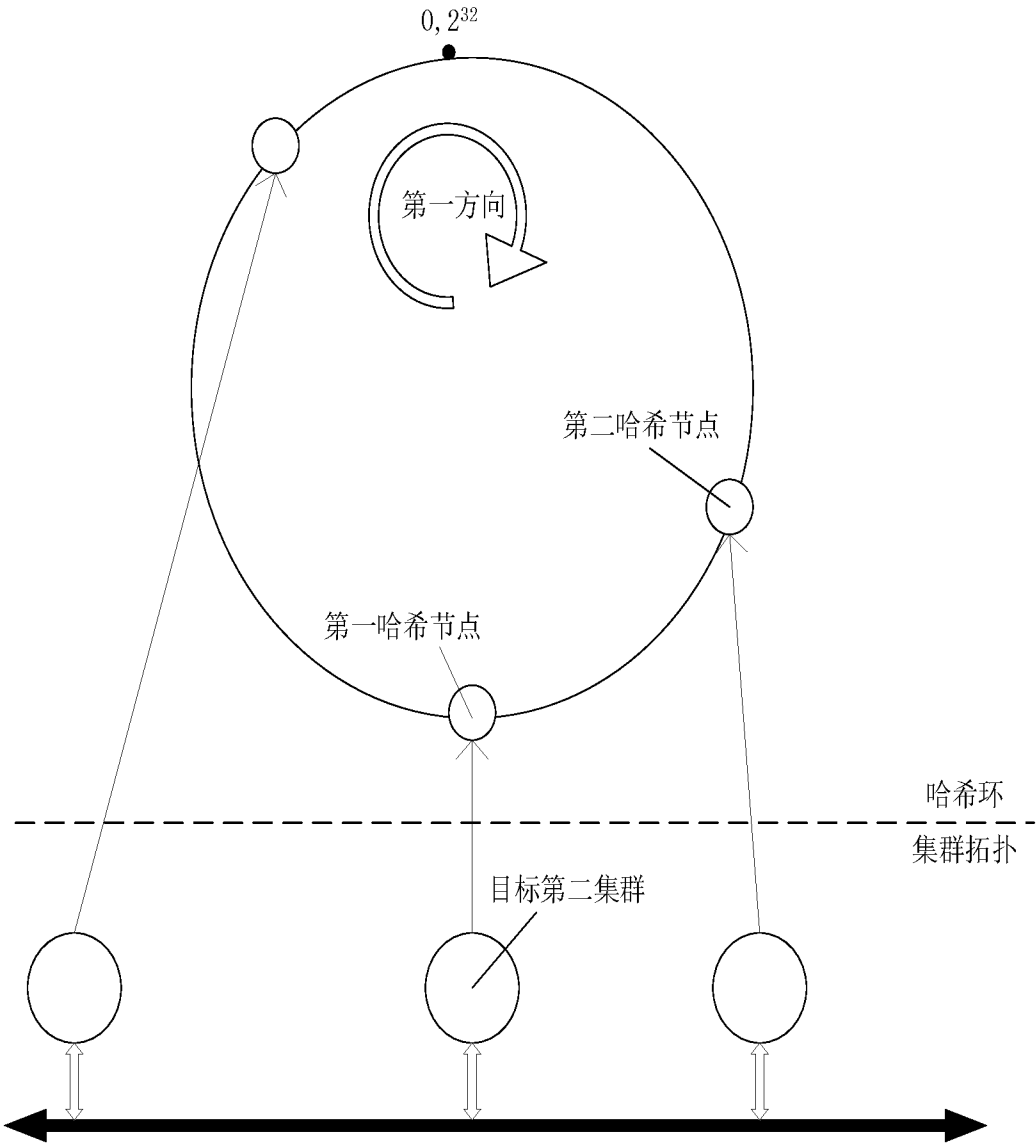


图 4b

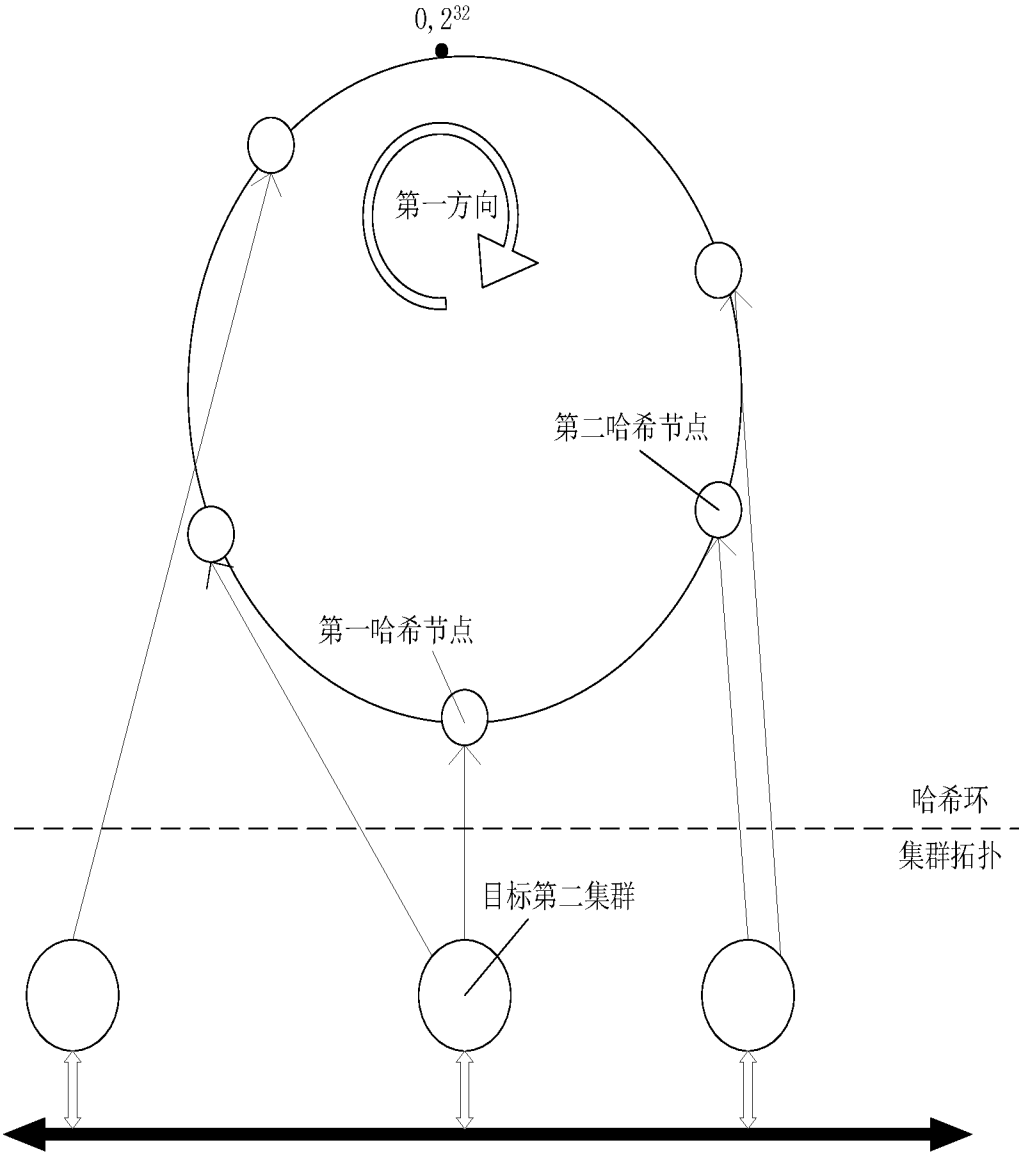


图 4c

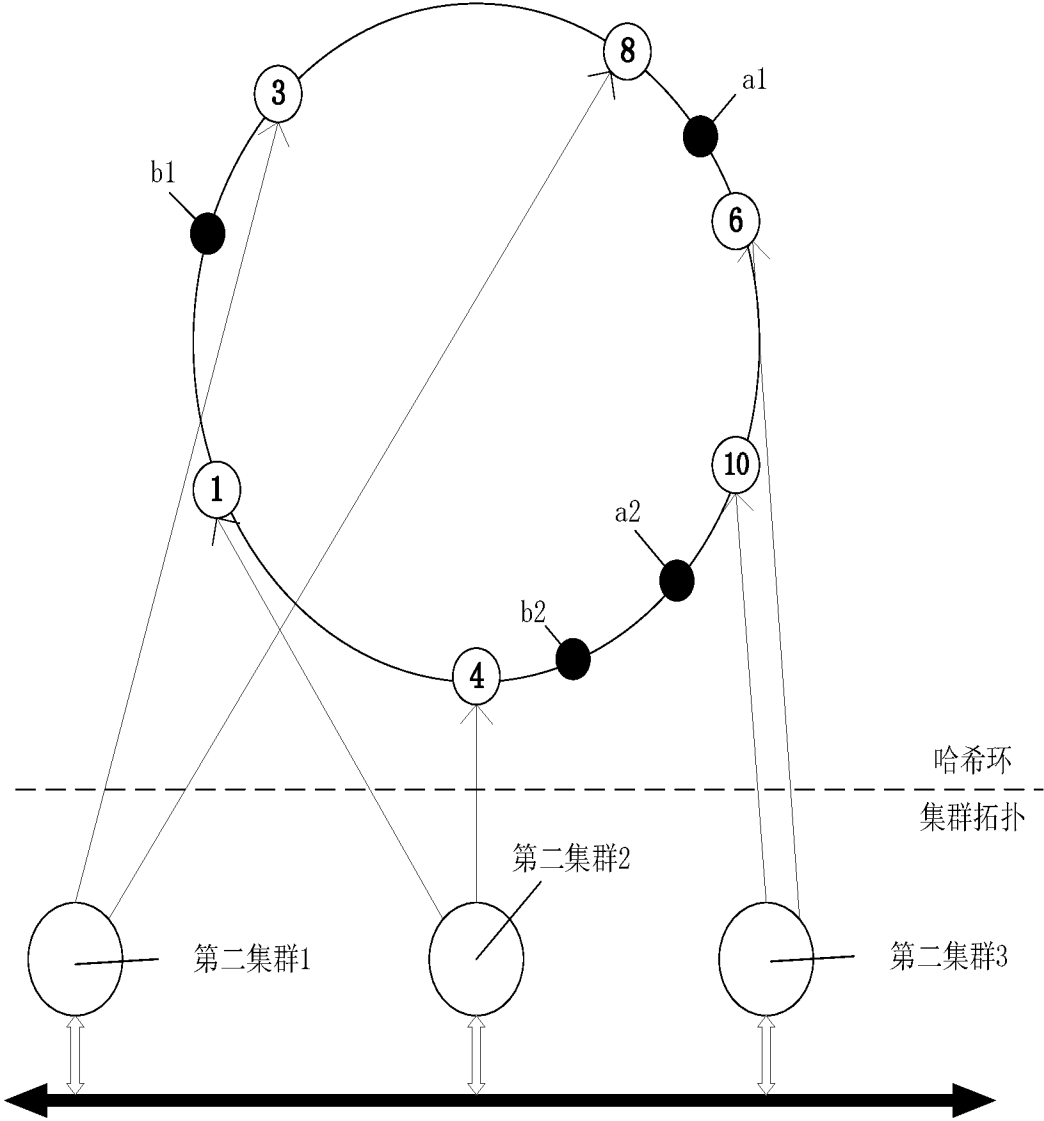


图 4d

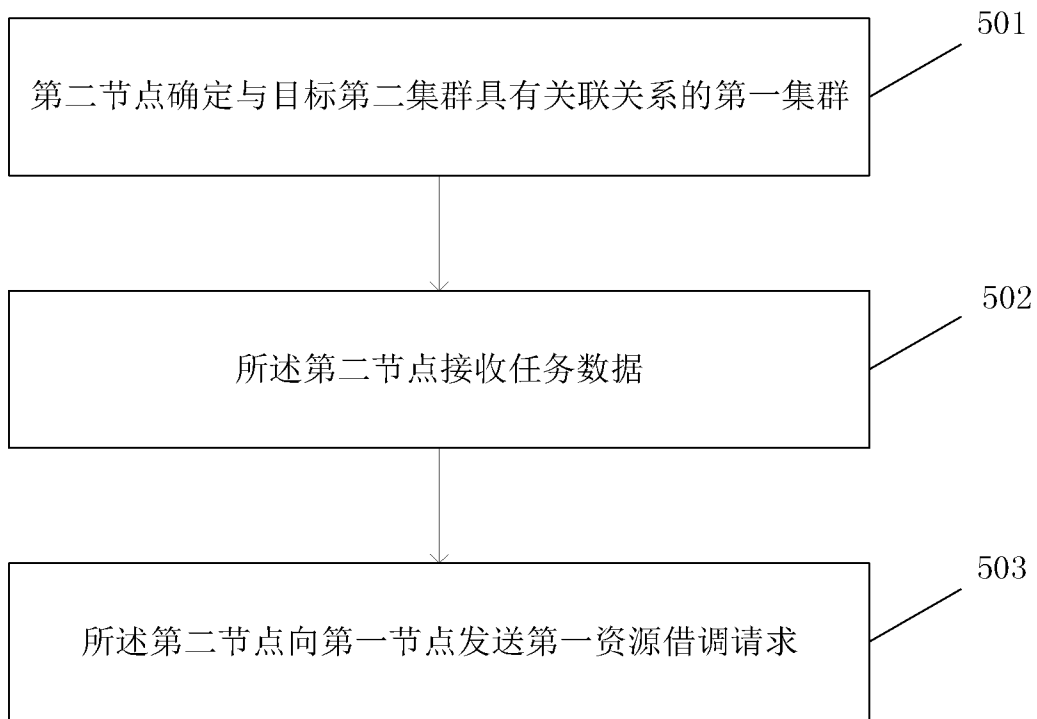


图 5

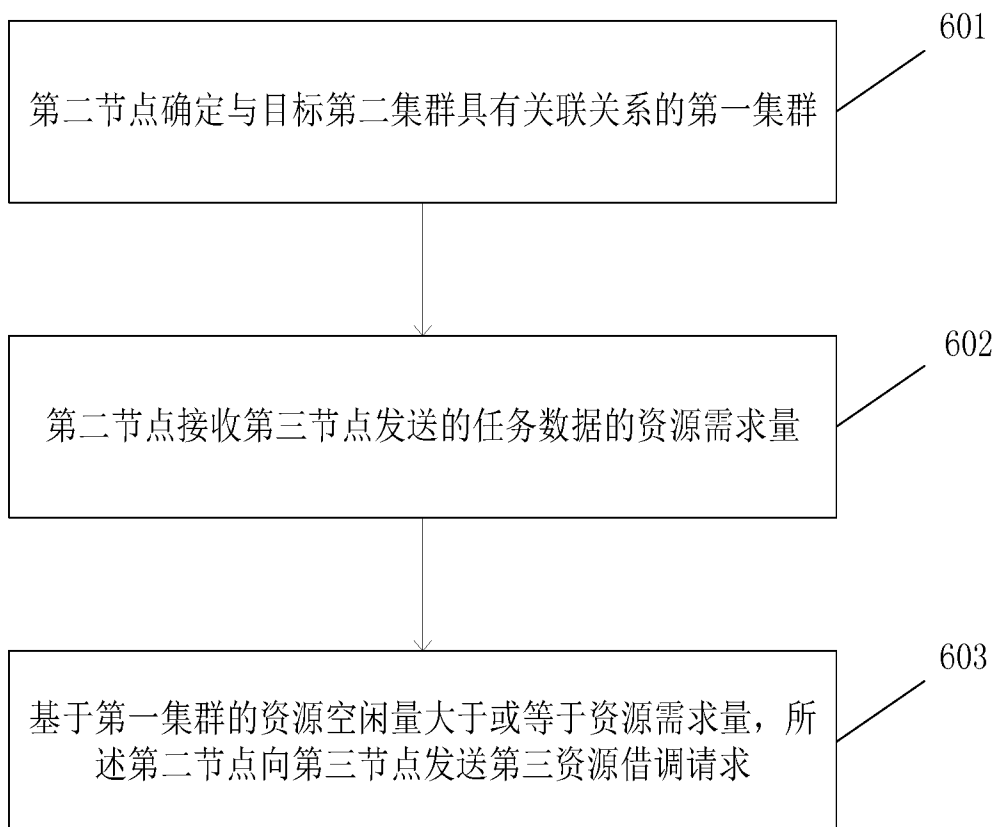


图 6

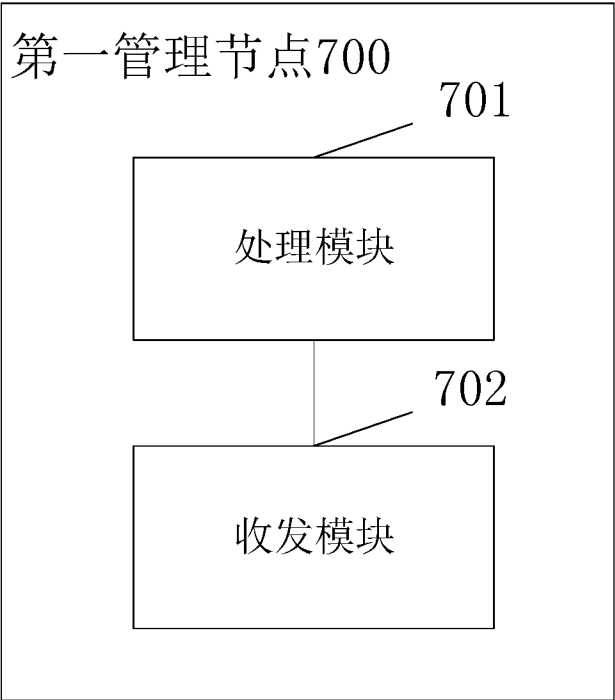


图 7

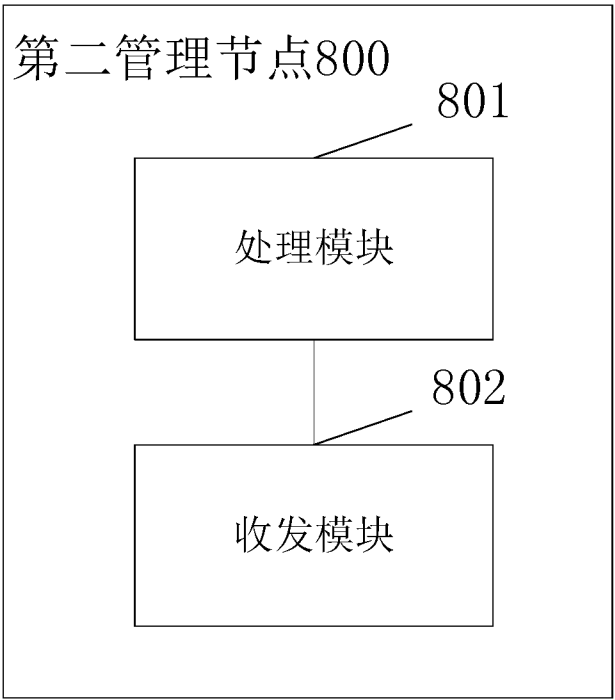


图 8

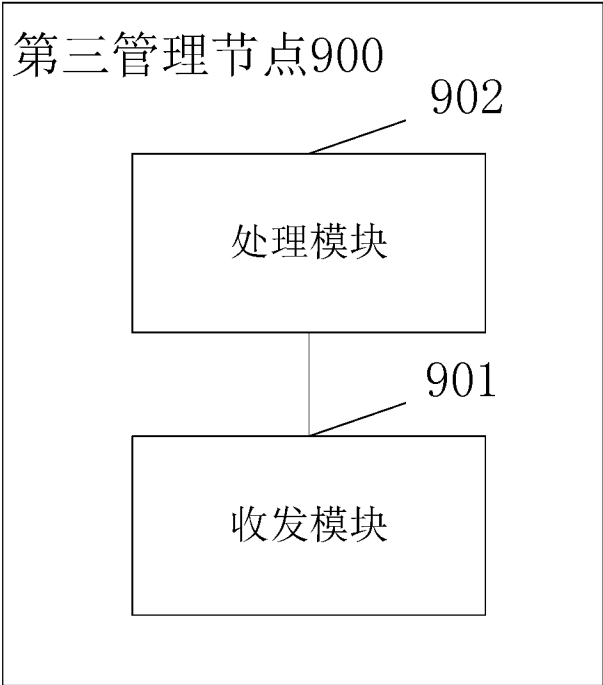


图 9

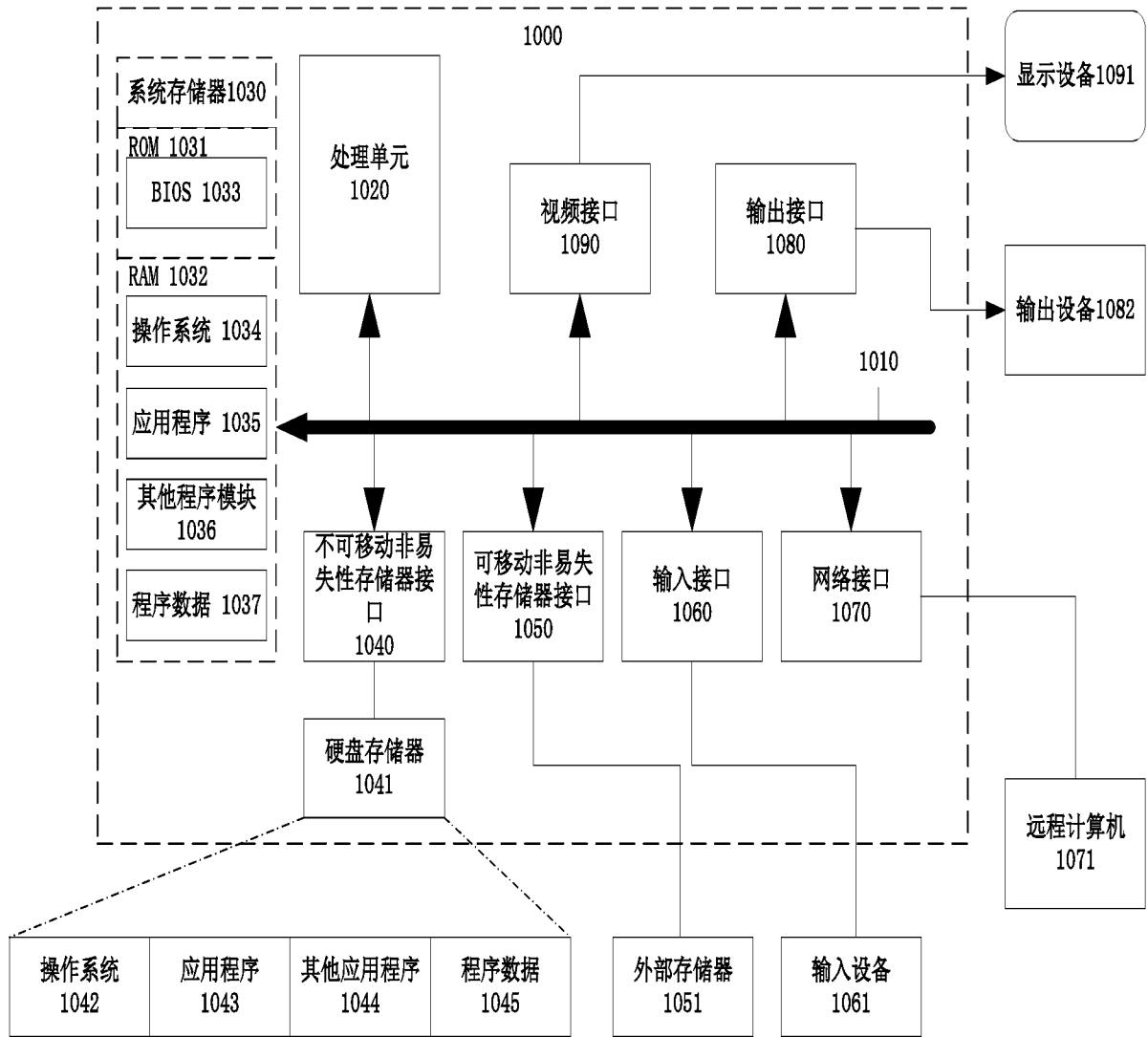


图 10



图 11



图 12



图 13

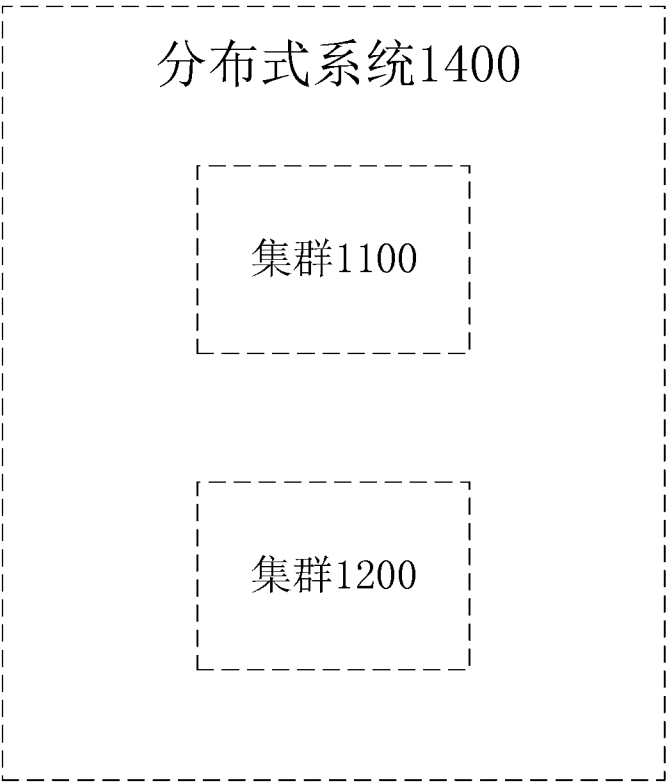


图 14

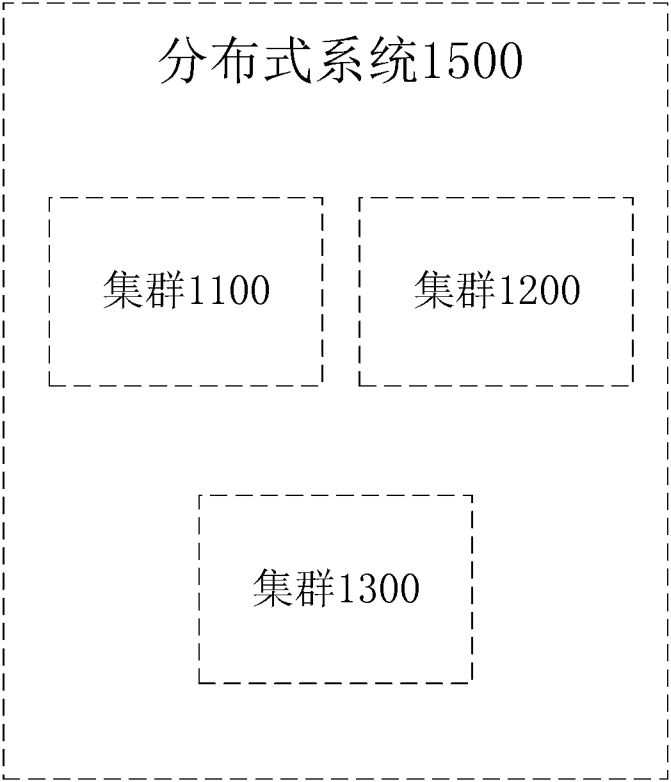


图 15

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/105476

A. CLASSIFICATION OF SUBJECT MATTER

G06F 9/48(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 H04L

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)

CNPAT, WPI, EPODOC, CNKI, BAIDU: 哈希, 资源, 分配, 调度, 借调, 调用, 共享, 分享, 一致性, 环, 集群, hash, consensus, resource, allocate+, share, ring, schedul+

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 108132830 A (BEIJING KINGSOFT CLOUD NETWORK TECHNOLOGY CO., LTD. et al.) 08 June 2018 (2018-06-08) description, paragraphs [0037]-[0084]	1-48
A	CN 108173937 A (BEIJING CHINA POWER INFORMATION TECHNOLOGY CO., LTD. et al.) 15 June 2018 (2018-06-15) entire document	1-48
A	CN 102130938 A (INSTITUTE OF SOFTWARE CHINESE ACADEMY OF SCIENCES) 20 July 2011 (2011-07-20) entire document	1-48
A	US 2017070567 A1 (ELECTRONICS & TELECOMMUNICATIONS RES. INST.) 09 March 2017 (2017-03-09) entire document	1-48



Further documents are listed in the continuation of Box C.



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

14 October 2020

Date of mailing of the international search report

28 October 2020

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/CN2020/105476

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	108132830	A	08 June 2018	None			
CN	108173937	A	15 June 2018	None			
CN	102130938	A	20 July 2011	CN	102130938	B	15 February 2012
US	2017070567	A1	09 March 2017	KR	20170029263	A	15 March 2017

国际检索报告

国际申请号

PCT/CN2020/105476

A. 主题的分类

G06F 9/48(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 H04L

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))

CNPAT, WPI, EPDOC, CNKI, BAIDU:哈希, 资源, 分配, 调度, 借调, 调用, 共享, 分享, 一致性, 环, 集群, hash, consensus, resource, allocate+, share, ring, schedul+

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
A	CN 108132830 A (北京金山云网络技术有限公司 等) 2018年 6月 8日 (2018 - 06 - 08) 说明书第[0037]-[0084]段	1-48
A	CN 108173937 A (北京中电普华信息技术有限公司 等) 2018年 6月 15日 (2018 - 06 - 15) 全文	1-48
A	CN 102130938 A (中国科学院软件研究所) 2011年 7月 20日 (2011 - 07 - 20) 全文	1-48
A	US 2017070567 A1 (ELECTRONICS & TELECOMMUNICATIONS RES. INST.) 2017年 3月 9日 (2017 - 03 - 09) 全文	1-48

☐ 其余文件在C栏的续页中列出。☒ 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

2020年 10月 14日

国际检索报告邮寄日期

2020年 10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

胡丽丽

电话号码 86-(10)-53961386

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2020/105476

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	108132830	A	2018年 6月 8日	无	
CN	108173937	A	2018年 6月 15日	无	
CN	102130938	A	2011年 7月 20日	CN 102130938	B 2012年 2月 15日
US	2017070567	A1	2017年 3月 9日	KR 20170029263	A 2017年 3月 15日