

# (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  
2017年9月28日 (28.09.2017)



(10) 国际公布号  
WO 2017/162086 A1

- (51) 国际专利分类号:  
G06F 9/50 (2006.01)
- (21) 国际申请号: PCT/CN2017/076877
- (22) 国际申请日: 2017年3月16日 (16.03.2017)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:  
201610179807.5 2016年3月25日 (25.03.2016) CN
- (71) 申请人: 阿里巴巴集团控股有限公司 (ALIBABA GROUP HOLDING LIMITED) [—/CN]; 开曼群岛大开曼资本大厦一座四层 847 号邮箱, Grand Cayman (KY)。
- (72) 发明人: 何乐 (HE, Le); 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。黄俨 (HUANG, Yan); 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。史英杰 (SHI, Yingjie); 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。张杰 (ZHANG, Jie); 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。张辰 (ZHANG, Chen); 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。
- (74) 代理人: 北京三友知识产权代理有限公司 (BEIJING SANYOU INTELLECTUAL PROPERTY AGENCY LTD.); 中国北京市金融街 35 号国际企业大厦 A 座 16 层, Beijing 100033 (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, 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,

[见续页]

(54) Title: TASK SCHEDULING METHOD AND DEVICE

(54) 发明名称: 任务调度方法和装置

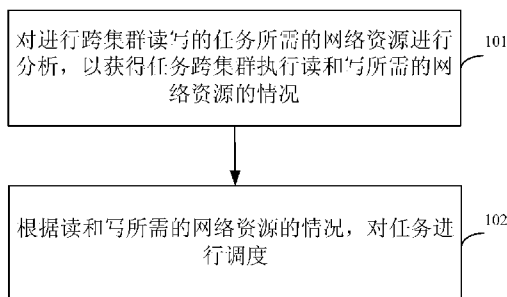


图 1

101 ANALYSE NETWORK RESOURCES REQUIRED BY TASK PERFORMING CROSS-CLUSTER READING AND WRITING, IN ORDER TO OBTAIN SITUATION REGARDING NETWORK RESOURCES REQUIRED BY TASK CROSS-CLUSTER READING AND WRITING  
102 SCHEDULE TASK ACCORDING TO SITUATION REGARDING NETWORK RESOURCES REQUIRED BY READING AND WRITING

(57) Abstract: Provided are a task scheduling method and device. Network resources required by a task performing cross-cluster reading and writing are analysed, in order to obtain the situation regarding network resources occupied by the reading and writing, and the task is scheduled according to the situation regarding network resources required by the reading and writing. The situation regarding network resources occupied by the reading and writing respectively reflects network resources which can be saved by scheduling the task to a cluster in which data accessed during the reading and writing is located. Therefore, accordingly determining the cluster to which the task is scheduled enables to the task to achieve less network resources occupation, solving inter-cluster excessive bandwidth occupation in the prior art.

(57) 摘要: 本发明提供了任务调度方法和装置, 通过对进行跨集群读写的任务所需的网络资源进行分析, 以获得读和写所占用网络资源的情况, 根据读和写所需的网络资源的情况, 对任务进行调度。由于读和写所占用的网络资源的情况, 分别体现了将任务调度至读和写时所访问数据所在集群能够节省的网络资源, 因而, 据此确定任务所调度至的集群能够使得任务实现较少的网络资源占用, 解决现有技术中集群间的带宽占用过高的情况。



WO 2017/162086 A1



PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

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)。

(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,

**本国际公布:**

— 包括国际检索报告(条约第 21 条(3))。

## 任务调度方法和装置

本申请要求 2016 年 03 月 25 日递交的申请号为 201610179807.5、发明名称为“任务调度方法和装置”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。

### 5 技术领域

本发明涉及计算机技术，尤其涉及一种任务调度方法和装置。

### 背景技术

为了提高系统的稳定性和网络中心的数据处理能力及服务能力，通常采用集群技术。集群技术的出现，能够使得服务器相互连接在一起，构成一个集群，多个集群相互连接构成一个分布式系统，该分布式系统内的各个集群运行一系列共同的应用程序。

在分布式系统中，可以将所运行的应用程序划分为多个任务，对于单个任务来说，可以将其按照运行的业务类型不同划分至不同的业务单元内，进而将同属于一个业务单元的任务运行于同一个集群上并将这些任务的任务数据也存储该集群上。

15 由于一个业务单元中的一个任务可能会需要读取另一业务单元中的另一任务的任务数据，也就是说在原集群上运行的任务需要依赖另一任务的任务数据。因此，当任务和其所依赖的另一任务的任务数据，即依赖数据，在不同集群上时，该任务会出现跨集群读写的情况，导致占用了大量的带宽。针对这一问题，现有技术中一旦发现某一任务存在跨集群读写的情况，便将该任务调度至其所读取的依赖数据所在的目标集群。但在  
20 实际运行过程中仍会出现集群间的带宽占用过高的情况。

### 发明内容

本发明提供一种任务调度方法和装置，用于解决现有技术中集群间的带宽占用过高的情况。

25 为达到上述目的，本发明的实施例采用如下技术方案：

第一方面，提供了一种任务调度方法，包括：

对进行跨集群读写的任务所需的网络资源进行分析，以获得所述任务跨集群执行读和写所需的网络资源的情况；

根据读和写所需的网络资源的情况，对所述任务进行调度。

30 第二方面，提供了一种任务调度装置，包括：

分析模块，用于对进行跨集群读写的任务所需的网络资源进行分析，以获得所述任务跨集群执行读和写所需的网络资源的情况；

调度模块，用于根据所述读和写所需的网络资源的情况，对所述任务进行调度。

本发明实施例提供的任务调度方法和装置，通过对进行跨集群读写的任务所需的网络资源进行分析，以获得任务跨集群执行读和写所占用网络资源的情况，根据读和写所需的网络资源的情况，对任务进行调度。由于读和写所占用的网络资源的情况，分别体现了将任务调度至读和写时所访问数据所在集群能够节省的网络资源，因而，据此确定任务所调度至的集群能够使得任务实现较少的网络资源占用，解决现有技术中集群间的带宽占用过高的情况。

10 上述说明仅是本发明技术方案的概述，为了能够更清楚了解本发明的技术手段，而可依照说明书的内容予以实施，并且为了让本发明的上述和其它目的、特征和优点能够更明显易懂，以下特举本发明的具体实施方式。

### 附图说明

15 通过阅读下文优选实施方式的详细描述，各种其他的优点和益处对于本领域普通技术人员将变得清楚明了。附图仅用于示出优选实施方式的目的，而并不认为是对本发明的限制。而且在整个附图中，用相同的参考符号表示相同的部件。在附图中：

图 1 为本发明实施例一提供的一种任务调度方法的流程示意图；

图 2 为本发明实施例二提供的一种任务调度方法的流程示意图；

20 图 3 为本发明实施例三提供的一种任务调度装置的结构示意图；

图 4 为本发明实施例四提供的一种任务调度装置的结构示意图。

### 具体实施方式

下面将参照附图更详细地描述本公开的示例性实施例。虽然附图中显示了本公开的示例性实施例，然而应当理解，可以以各种形式实现本公开而不应被这里阐述的实施例所限制。相反，提供这些实施例是为了能够更透彻地理解本公开，并且能够将本公开的范围完整的传达给本领域的技术人员。

下面结合附图对本发明实施例提供的任务调度方法和装置进行详细描述。

#### 实施例一

30 图 1 为本发明实施例一提供的一种任务调度方法的流程示意图，本实施例所提供的

方法，可以由分布式系统中的任务管理器执行，如图 1 所示，方法包括：

步骤 101、对进行跨集群读写的任务所需的网络资源进行分析，以获得任务跨集群执行读和写所需的网络资源的情况。

具体的，分布式系统一般会承载着很多的任务，其中有相当一部分是周期性运行的任务，这些任务可能会每小时、每天或者每周的定期运行，任务运行之前需要访问运行所需的数据，这里所说的访问即读取，任务定期访问的数据会是当时最新的数据，从而保证了运行所获得的结果数据的准确性。在读取数据之后，还需要将所读取到的数据输入到运行的任务中，因此，所读取的数据为任务的输入数据。若该任务访问数据的过程中涉及跨集群读的情况，则会占用一定带宽等网络资源。由于在这种跨集群访问数据的情况下，必然涉及将数据从另一集群读取到任务所运行的集群，此时任务的输入数据便能够体现出跨集群读所占用的网络资源的情况，例如：占用带宽的情况。

另外，任务在运行结束后，还需要将任务运行的结果数据返回至任务管理器所最初分配的默认集群，所返回的结果数据需要首先由任务进行输出，所输出的数据称为任务的输出数据，进而将结果数据写入默认集群。若该任务返回结果数据的过程中涉及跨集群写的情况，则同样会占用一定带宽等网络资源。由于跨集群返回结果数据必然涉及将结果数据从任务所运行的集群写入到另一集群，此时任务的输出数据便能够体现出跨集群写所占用的网络资源的情况，例如：占用带宽的情况。

通过执行跨集群读和写时，任务所需的网络资源的情况，一方面可以预测出若任务当前所在的原集群与所读取的依赖数据所在集群为不同集群时，任务所占用的网络资源情况，另一方面可以预测出若任务当前所在的原集群与所写入的结果数据所在集群为不同集群时，任务所占用的网络资源情况。

作为一种可能的实现方式，历史记录中记录有每一个任务运行在当前所分配的原集群上所输入和输出的数据量，可以根据历史记录，针对每一个任务，获得输入数据的数据量、输出数据的数据量，针对每一个任务计算输入输出比。其中，输入输出比等于输入数据的数据量与输出数据的数据量的比值。

步骤 102、根据读和写所需的网络资源的情况，对任务进行调度。

具体的，若读操作所需的网络资源多于写操作所需的网络资源，将任务调度至所读取的依赖数据所在的目标集群。

这是由于跨集群任务的数据存储和任务运行分别位于不同的集群上，输入数据和输出数据分别均存在三种情况：A. 存储在任务当前所分配的原集群上；B. 存储在任务所

待调度至的目标集群上；C. 存储在原集群和目标集群之外的集群上。因此，在现有技术的一旦发现某一任务存在跨集群读写的情况，便将该任务调度至其所读取的依赖数据所在的目标集群的情况下，输入数据或输出数据只要不在目标集群上就需要通过跨集群复制或直读直写的方式访问，跨集群任务所产生的跨集群数据量过大时会对网络资源造成压力。

在一种可能的情况下，若跨集群任务的输入数据存储于目标集群上并且数据量很大，那么调度该跨集群任务至目标集群可以省去将依赖数据从目标集群读取到原集群的网络资源占用，同时任务的输出数据的数据量很小，那么调度该跨集群任务至目标集群所导致增加的将结果数据从目标集群写入到默认集群的网络资源占用不会很大，从而使得调度引起的跨集群网络资源占用会出现显著降低。也就是说，输入数据的数据量与输出数据的数据量的比值，即输入输出比越大则跨集群网络资源占用降低的情况越显著。

作为一种可能的实现方式，可以通过判断任务输入输出比是否大于预设第一阈值预测调度至依赖数据所在的目标集群是否能够有效降低所占用的带宽等网络资源，若是，则将任务调度至所述任务的依赖数据所在的目标集群。其中，第一阈值大于1。

相反的，若读操作所需的网络资源不多于写操作所需的网络资源，可以将任务调度至所写入的结果数据所在的集群。

本实施例中，通过对进行跨集群读写的任务所需的网络资源进行分析，以获得读和写所占用网络资源的情况，根据读和写所需的网络资源的情况，对任务进行调度。由于读和写所占用的网络资源的情况，分别体现了将任务调度至读和写时所访问数据所在集群能够节省的网络资源，因而，据此确定任务所调度至的集群能够使得任务实现较少的网络资源占用，解决现有技术中集群间的带宽占用过高的情况。

## 实施例二

图2为本发明实施例二提供的一种任务调度方法的流程示意图，如图2所示，包括：步骤201、基于历史记录进行分析，从进行跨集群读写的任务中筛选出目标任务。

具体的，历史记录中记录有各个任务的任务信息和数据信息。其中，任务信息包括：输入数据的数据量和输出数据的数据量；数据信息包括：输入数据所在集群、输出数据所在集群、任务所属业务单元和任务当前所在集群，以及任务运行频率和运行开销。

基于历史记录，确定出输入数据或输出数据所在集群和任务当前所在集群为不同集群的跨集群任务。将任务当前所在集群作为原集群。

针对每一个跨集群任务，根据公式输入输出比=输入数据量/输出数据量，计算输入输出比。采用输入输出比、输出数据量、运行开销和运行频率、当前任务所在集群负载所构成的筛选条件筛选出读操作所需的网络资源多于写操作所需的网络资源的目标任务。

5 例如：筛选条件可以为输入输出比大于第一阈值，第一阈值为 100，且输出数据量小于 100GB，运行开销、运行频率、集群负载等分别小于预设集群配额。

这里所说的预设集群配额是根据目标集群能够提供的 CPU 和存储空间等集群资源所确定的。在筛选条件中增加运行开销、运行频率、集群负载等用于指示任务占用的集群资源的指标，主要是由于即使输入输出比较大也就是调度任务所节省的网络资源较多，  
10 但是当任务运行时所占用的集群资源较多，而所调度至的目标集群无法满足任务所需，从而大幅增加了目标集群的负载的情况下，将任务调度至目标集群后分布式系统的整体性能并没有得到明显提升，这种调度就是得不偿失的，因此，在当任务运行时所占用的集群资源较多时，不对任务进行调度。

步骤 202、对目标任务提取任务标识，生成记录任务标识的调度信息。

15 具体的，获得目标任务的任务标识，有两种方法，对于结构化查询语言(Structured Query Language, SQL)任务可以将查询语句中周期性的属性屏蔽掉，对任务进行哈希摘要处理，例如采用消息摘要算法第五版(Message Digest Algorithm MD5, 简称 MD5)算法进行哈希摘要处理，并将哈希摘要作为该任务的任务标识。另一种方法，针对非 SQL 任务可以直接将任务的固定编号作为任务的任务标识，这些固定编号可以来自如天网系  
20 统等外部系统。

在调度信息中，用任务标识对该目标任务进行标记。另外，调度信息可以用于记录目标任务所需调度至的目标集群。

由于基于历史记录进行分析，筛选出目标任务的过程运算量较大，因而，可以预先执行，从而基于执行结果，生成调度信息。这样，当接收到待调度的任务时，则无需对  
25 其进行分析，可以直接根据预先分析所获得的调度信息进行调度，节省了时间，提高了调度的时效性。在实际操作过程中，可以将基于历史记录进行分析，生成调度信息的过程称为训练过程，将后续根据调度信息进行调度的过程称为决策过程。

步骤 203、当接收到待调度的任务时，根据调度信息调度任务。

具体的，当接收到待调度的任务时，获取该任务的任务标识，从而对该任务进行识  
30 别。具体来说，可以在任务调度系统接收到待调度的任务后，判断任务的类型是否为 SQL

任务，如果是则提取哈希摘要作为任务标识，否则，提取固定编号作为任务标识。具体获取任务标识的过程参见步骤 202 中的相关描述，此处不再赘述。根据获取到的任务标识，在调度信息中进行匹配，匹配到了则调度到任务的目标集群上，否则，调度到任务所写入的结果数据所在的原集群。进一步，在调度之后，还可以为经过调度的任务分配  
5 计算资源。

因为当跨集群任务的输入数据存储于目标集群上并且数据量很大时，调度该跨集群任务至目标集群可以省去将依赖数据从目标集群读取到原集群的网络资源占用，同时任务的输出数据的数据量很小，那么调度该跨集群任务至目标集群所增加的将结果数据从目标集群到原集群写入的网络资源占用不会很大，因此，使得调度引起的跨集群网络资源  
10 占用会出现显著降低。

也就是说，输入数据的数据量与输出数据的数据量的比值，即输入输出比越大则跨集群网络资源占用降低的情况越显著，可以据此确定第一阈值的取值，若仅在跨集群网络资源占用降低的情况越显著时调度任务至目标集群，则可以将第一阈值确定的较大，否则，确定的较小，但第一阈值应大于 1。

15

### 实施例三

图 3 为本发明实施例三提供的一种任务调度装置的结构示意图，如图 3 所示，包括：分析模块 31 和调度模块 32。

分析模块 31，用于对进行跨集群读写的任务所需的网络资源进行分析，以获得所述  
20 任务读和写所需的网络资源的情况；

调度模块 32，用于根据所述读和写所需的网络资源的情况，对所述任务进行调度。

具体的，调度模块 32 具体用于若所述读操作所需的网络资源多于写操作所需的网络资源，将所述任务调度至所读取的依赖数据所在的目标集群。

本实施例中，通过对进行跨集群读写的任务所需的网络资源进行分析，以获得读和  
25 写所占用网络资源的情况，根据读和写所需的网络资源的情况，对任务进行调度。由于读和写所占用的网络资源的情况，分别体现了将任务调度至读和写时所访问数据所在集群能够节省的网络资源，因而，据此确定任务所调度至的集群能够使得任务实现较少的网络资源占用，解决现有技术中集群间的带宽占用过高的情况。

30

### 实施例四

图 4 为本发明实施例四提供的一种任务调度装置的结构示意图，在图 3 所提供的任务调度装置的基础上，分析模块 31，包括：获得单元 311 和计算单元 312。

获得单元 311，用于根据历史记录，针对每一个所述任务，获得输入数据的数据量、输出数据的数据量。

5 计算单元 312，用于针对每一个所述任务计算用于指示读和写所需的网络资源的比例的输入输出比。

其中，输入输出比等于输入数据的数据量与输出数据的数据量的比值。

进一步，调度模块 32，包括：判断单元 321、标识单元 322、生成单元 323 和调度单元 324。

10 判断单元 321，用于判断所述任务是否满足预设的筛选条件。

其中，筛选条件包括：所述输入输出比大于预设第一阈值；其中，第一阈值大于 1。筛选条件还包括：输出数据的数据量小于第二阈值；和/或，所占用的集群资源小于预设配额，其中所占用的集群资源包括运行开销、运行频率和集群负载中的至少一个。

标识单元 322，用于针对满足所述筛选条件的任务，获得任务标识。

15 生成单元 323，用于生成用于记录所述任务标识的调度信息。

调度单元 324，用于若满足所述筛选条件，则将所述任务调度至所述任务所读取的依赖数据所在的目标集群。

具体的，调度单元 324，具体用于当接收到待调度任务时，获取标识单元所获得所述待调度任务的任务标识；若所述待调度任务的任务标识与所述调度信息中的任务标识  
20 相匹配，则将所述待调度任务调度至所述待调度任务的依赖数据所在的目标集群。

进一步，标识单元 322，包括：判断子单元 3221、哈希子单元 3222 和编号子单元 3223。

判断子单元 3221，用于判断所述任务的类型是否为 SQL；

25 哈希子单元 3222，用于若所述任务的类型为 SQL，对所述任务进行哈希处理，获得哈希摘要，将所述哈希摘要作为所述任务标识；

编号子单元 3223，用于若所述任务的类型不为 SQL，将所述任务的编号作为所述任务标识。

需要说明的是，在前述各中所提及的网络资源可以为网络带宽和/或网络带宽时延积，本领域技术人员可以知晓，还可以采用其他用于衡量网络资源的指标，而不会影响  
30 各实施例的实现效果。

通过对进行跨集群读写的任务所需的网络资源进行分析，以获得读和写所占用网络资源的情况，根据读和写所需的网络资源的情况，对任务进行调度。由于读和写所占用的网络资源的情况，分别体现了将任务调度至读和写时所访问数据所在集群能够节省的网络资源，因而，据此确定任务所调度至的集群能够使得任务实现较少的网络资源占用，

5 解决现有技术中集群间的带宽占用过高的情况。

本领域普通技术人员可以理解：实现上述各方法实施例的全部或部分步骤可以通过程序指令相关的硬件来完成。前述的程序可以存储于一计算机可读取存储介质中。该程序在执行时，执行包括上述各方法实施例的步骤；而前述的存储介质包括：ROM、RAM、磁碟或者光盘等各种可以存储程序代码的介质。

10 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。

15

## 权利要求书

1. 任务调度方法，其特征在于，包括：

对进行跨集群读写的任务所需的网络资源进行分析，以获得所述任务跨集群执行读和写所需的网络资源的情况；

5 根据读和写所需的网络资源的情况，对所述任务进行调度。

2. 根据权利要求 1 所述的任務调度方法，其特征在于，所述根据读和写所需的网络资源的情况，对所述任务进行调度，包括：

若读操作所需的网络资源多于写操作所需的网络资源，将所述任务调度至所读取的依赖数据所在的目标集群。

10 3. 根据权利要求 1 所述的任務调度方法，其特征在于，所述对进行跨集群读写的任务所需的网络资源进行分析，包括：

根据历史记录，针对每一个所述任务，获得输入数据的数据量、输出数据的数据量；针对每一个所述任务计算用于指示读和写所需的网络资源的比例的输入输出比；其中，输入输出比等于输入数据的数据量与输出数据的数据量的比值。

15 4. 根据权利要求 3 所述的任務调度方法，其特征在于，所述根据读和写所需的网络资源的情况，对所述任务进行调度，包括：

判断所述任务是否满足预设的筛选条件；其中，所述筛选条件包括：所述输入输出比大于预设第一阈值；其中，第一阈值大于 1；

20 若满足所述筛选条件，则将所述任务调度至所述任务所读取的依赖数据所在的目标集群。

5. 根据权利要求 4 所述的任務调度方法，其特征在于，所述判断所述任务是否满足预设的筛选条件之后，还包括：

针对满足所述筛选条件的任务，获得任务标识；

生成用于记录所述任务标识的调度信息。

25 6. 根据权利要求 5 所述的任務调度方法，其特征在于，所述若满足所述筛选条件，则将所述任务调度至所述任务所读取的依赖数据所在的目标集群，包括：

当接收到待调度任务时，针对所述待调度任务，获得任务标识；

若所述待调度任务的任务标识与所述调度信息中的任务标识相匹配，则将所述待调度任务调度至所述待调度任务的依赖数据所在的目标集群。

30 7. 根据权利要求 5 或 6 所述的任務调度方法，其特征在于，所述获得任务标识，

包括:

判断所述任务的类型是否为 SQL;

若所述任务的类型为 SQL, 对所述任务进行哈希处理, 获得哈希摘要, 将所述哈希摘要作为所述任务标识;

5 否则, 将所述任务的编号作为所述任务标识。

8. 根据权利要求 4 所述的任务调度方法, 其特征在于, 所述筛选条件还包括: 输出数据的数据量小于第二阈值;

和/或, 所占用的集群资源小于预设配额, 其中所占用的集群资源包括运行开销、运行频率和集群负载中的至少一个。

10 9. 根据权利要求 1-6 任一项所述的任务调度方法, 其特征在于, 所述网络资源包括: 网络带宽和网络带宽时延积中的至少一个。

10. 一种任务调度装置, 其特征在于, 包括:

分析模块, 用于对进行跨集群读写的任务所需的网络资源进行分析, 以获得所述任务跨集群执行读和写所需的网络资源的情况;

15 调度模块, 用于根据所述读和写所需的网络资源的情况, 对所述任务进行调度。

11. 根据权利要求 10 所述的任务调度装置, 其特征在于,

所述调度模块, 具体用于若读操作所需的网络资源多于写操作所需的网络资源, 将所述任务调度至所读取的依赖数据所在的目标集群。

12. 根据权利要求 10 所述的任务调度装置, 其特征在于, 所述分析模块, 包括:

20 获得单元, 用于根据历史记录, 针对每一个所述任务, 获得输入数据的数据量、输出数据的数据量;

计算单元, 用于针对每一个所述任务计算用于指示读和写所需的网络资源的比例的输入输出比; 其中, 输入输出比等于输入数据的数据量与输出数据的数据量的比值。

13. 根据权利要求 12 所述的任务调度装置, 其特征在于, 所述调度模块, 包括:

25 判断单元, 用于判断所述任务是否满足预设的筛选条件; 其中, 所述筛选条件包括: 所述输入输出比大于预设第一阈值; 其中, 第一阈值大于 1;

调度单元, 用于若满足所述筛选条件, 则将所述任务调度至所述任务所读取的依赖数据所在的目标集群。

14. 根据权利要求 13 所述的任务调度装置, 其特征在于, 所述调度模块, 还包括:

30 标识单元, 用于针对满足所述筛选条件的任务, 获得任务标识;

生成单元，用于生成用于记录所述任务标识的调度信息。

15. 根据权利要求 14 所述的调度装置，其特征在于，  
所述标识单元，还用于当接收到待调度任务时，针对所述待调度任务，获得任务标识；

5 所述调度单元，具体用于当接收到待调度任务时，获取标识单元所获得所述待调度任务的任务标识；若所述待调度任务的任务标识与所述调度信息中的任务标识相匹配，则将所述待调度任务调度至所述待调度任务的依赖数据所在的目标集群。

16. 根据权利要求 14 或 15 所述的调度装置，其特征在于，所述标识单元，包括：

10 判断子单元，用于判断所述任务的类型是否为 SQL；

哈希子单元，用于若所述任务的类型为 SQL，对所述任务进行哈希处理，获得哈希摘要，将所述哈希摘要作为所述任务标识；

编号子单元，用于若所述任务的类型不为 SQL，将所述任务的编号作为所述任务标识。

15 17. 根据权利要求 13 所述的调度装置，其特征在于，所述筛选条件还包括：  
输出数据的数据量小于第二阈值；

和/或，所占用的集群资源小于预设配额，其中所占用的集群资源包括运行开销、运行频率和集群负载中的至少一个。

20 18. 根据权利要求 10-15 任一项所述的调度装置，其特征在于，所述网络资源  
包括：网络带宽和网络带宽时延积中的至少一个。

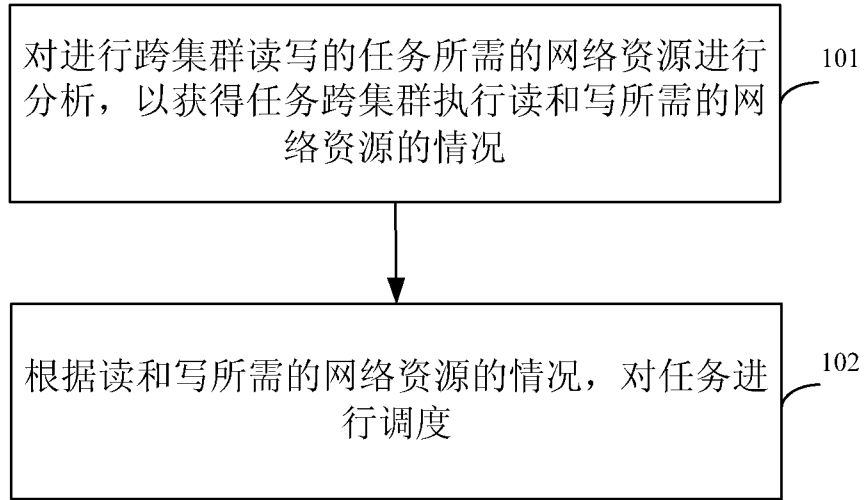


图 1

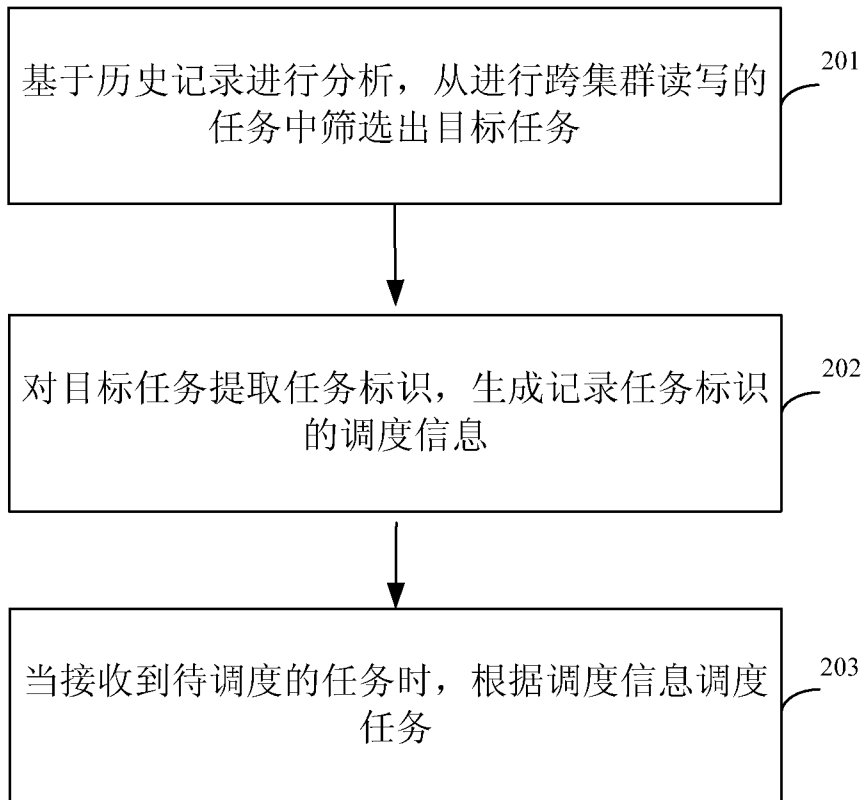


图 2

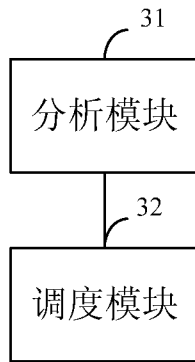


图 3

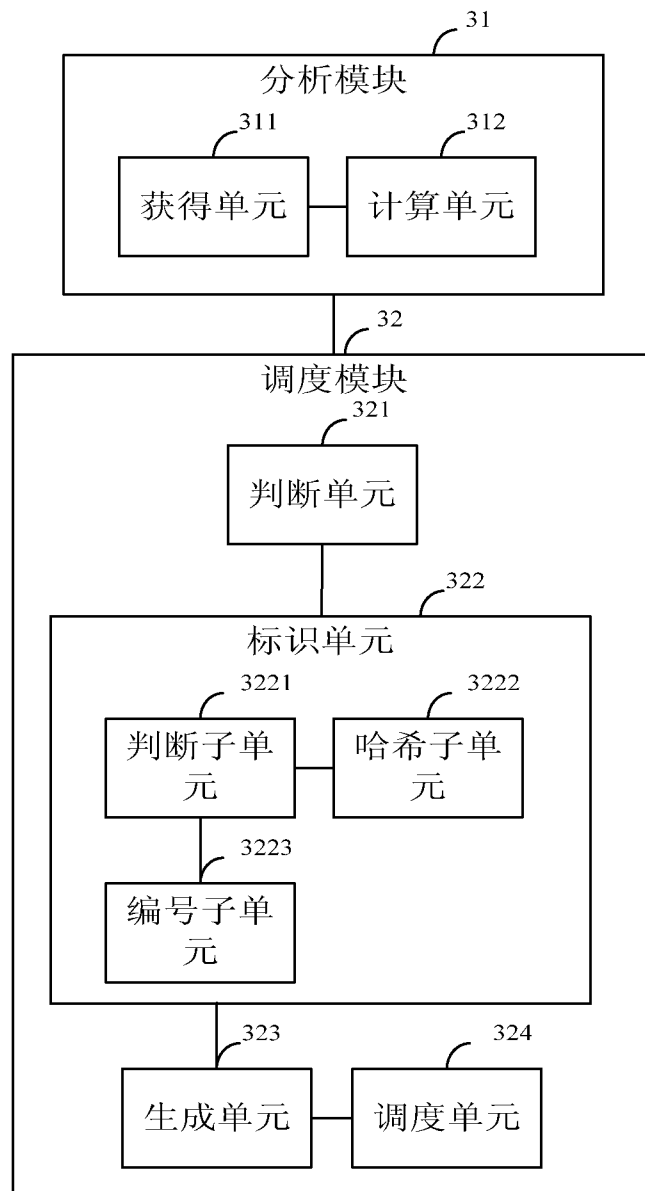


图 4

# INTERNATIONAL SEARCH REPORT

International application No.

**PCT/CN2017/076877**

## 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)

SIPOABS, CNABS, DWPI, CNKI: schedule, task, scheduler, distributing, cluster, read, write, input, output, data, resource, bandwidth

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 103019853 A (BEIJING IZP TECHNOLOGIES CO., LTD.), 03 April 2016 (03.04.2016), description, paragraphs [0050]-[0078]	1-18
A	CN 103377075 A (TENCENT TECHNOLOGY (SHENZHEN) CO., LTD.), 30 October 2013 (30.10.2013), the whole document	1-18
A	US 2009193427 A1 (INT BUSINESS MACHINES CORP.), 30 July 2009 (30.07.2009), the whole document	1-18

Further documents are listed in the continuation of Box C.

See patent family annex.

<p>* Special categories of cited documents:</p> <p>“A” document defining the general state of the art which is not considered to be of particular relevance</p> <p>“E” earlier application or patent but published on or after the international filing date</p> <p>“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)</p> <p>“O” document referring to an oral disclosure, use, exhibition or other means</p> <p>“P” document published prior to the international filing date but later than the priority date claimed</p>	<p>“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</p> <p>“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</p> <p>“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</p> <p>“&amp;” document member of the same patent family</p>
---	---

Date of the actual completion of the international search <p style="text-align: center;">16 June 2017 (16.06.2017)</p>	Date of mailing of the international search report <p style="text-align: center;"><b>22 June 2017 (22.06.2017)</b></p>
---	---

Name and mailing address of the ISA/CN: State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No.: (86-10) 62019451	Authorized officer <p style="text-align: center;"><b>ZHANG, Yan</b></p> Telephone No.: (86-10) <b>62089421</b>
--	---

**INTERNATIONAL SEARCH REPORT**  
Information on patent family members

International application No.

**PCT/CN2017/076877**

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 103019853 A	03 April 2016	None	
CN 103377075 A	30 October 2013	None	
US 2009193427 A1	30 July 2009	US 8281012 B2	02 October 2012

<p>A. 主题的分类</p> <p>G06F 9/50 (2006.01) i</p> <p>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</p>														
<p>B. 检索领域</p> <p>检索的最低限度文献 (标明分类系统和分类号)</p> <p>G06F</p> <p>包含在检索领域中的除最低限度文献以外的检索文献</p> <p>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))</p> <p>SIPOABS, CNABS, DWPI, CNKI: 任务, 调度, 分布, 集群, 群集, 读, 写, 输入, 输出, 数据, 资源, 带宽, task, scheduler, distributing, cluster, read, write, input, output, data, resource, bandwidth</p>														
<p>C. 相关文件</p> <table border="1"> <thead> <tr> <th>类型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>X</td> <td>CN 103019853 A (北京亿赞普网络技术有限公司) 2016年 4月 3日 (2016 - 04 - 03) 说明书第[0050]-[0078]段</td> <td>1-18</td> </tr> <tr> <td>A</td> <td>CN 103377075 A (腾讯科技深圳有限公司) 2013年 10月 30日 (2013 - 10 - 30) 全文</td> <td>1-18</td> </tr> <tr> <td>A</td> <td>US 2009193427 A1 (INT BUSINESS MACHINES CORP.) 2009年 7月 30日 (2009 - 07 - 30) 全文</td> <td>1-18</td> </tr> </tbody> </table>			类型*	引用文件, 必要时, 指明相关段落	相关的权利要求	X	CN 103019853 A (北京亿赞普网络技术有限公司) 2016年 4月 3日 (2016 - 04 - 03) 说明书第[0050]-[0078]段	1-18	A	CN 103377075 A (腾讯科技深圳有限公司) 2013年 10月 30日 (2013 - 10 - 30) 全文	1-18	A	US 2009193427 A1 (INT BUSINESS MACHINES CORP.) 2009年 7月 30日 (2009 - 07 - 30) 全文	1-18
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求												
X	CN 103019853 A (北京亿赞普网络技术有限公司) 2016年 4月 3日 (2016 - 04 - 03) 说明书第[0050]-[0078]段	1-18												
A	CN 103377075 A (腾讯科技深圳有限公司) 2013年 10月 30日 (2013 - 10 - 30) 全文	1-18												
A	US 2009193427 A1 (INT BUSINESS MACHINES CORP.) 2009年 7月 30日 (2009 - 07 - 30) 全文	1-18												
<p><input type="checkbox"/> 其余文件在C栏的续页中列出。</p> <p><input checked="" type="checkbox"/> 见同族专利附件。</p>														
<p>* 引用文件的具体类型:</p> <p>“A” 认为不特别相关的表示了现有技术一般状态的文件</p> <p>“E” 在国际申请日的当天或之后公布的在先申请或专利</p> <p>“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)</p> <p>“O” 涉及口头公开、使用、展览或其他方式公开的文件</p> <p>“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</p> <p>“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</p> <p>“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</p> <p>“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</p> <p>“&amp;” 同族专利的文件</p>														
<p>国际检索实际完成的日期</p> <p>2017年 6月 16日</p>		<p>国际检索报告邮寄日期</p> <p>2017年 6月 22日</p>												
<p>ISA/CN的名称和邮寄地址</p> <p>中华人民共和国国家知识产权局 (ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088</p> <p>传真号 (86-10)62019451</p>		<p>授权官员</p> <p>张妍</p> <p>电话号码 (86-10)62089421</p>												

国际检索报告  
关于同族专利的信息

国际申请号 PCT/CN2017/076877
----------------------------

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	103019853	A	2016年 4月 3日	无			
CN	103377075	A	2013年 10月 30日	无			
US	2009193427	A1	2009年 7月 30日	US	8281012	B2	2012年 10月 2日

表 PCT/ISA/210 (同族专利附件) (2009年7月)