

# (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 6 月 15 日 (15.06.2017)



(10) 国际公布号  
**WO 2017/097124 A1**

- (51) 国际专利分类号:  
*G06F 17/30* (2006.01)
- (21) 国际申请号: PCT/CN2016/107409
- (22) 国际申请日: 2016 年 11 月 28 日 (28.11.2016)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:  
201510888403.9 2015 年 12 月 7 日 (07.12.2015) CN
- (71) 申请人: 阿里巴巴集团控股有限公司 (ALIBABA GROUP HOLDING LIMITED) [—/CN]; 英属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: 洪蛟 (HONG, Jiao); 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, 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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, 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))。

(54) Title: METHOD, APPARATUS AND SYSTEM FOR TRANSMITTING TASKS BASED ON SUB-LIBRARIES AND SUB-TABLES

(54) 发明名称: 基于分库分表的任务传输方法、装置及系统

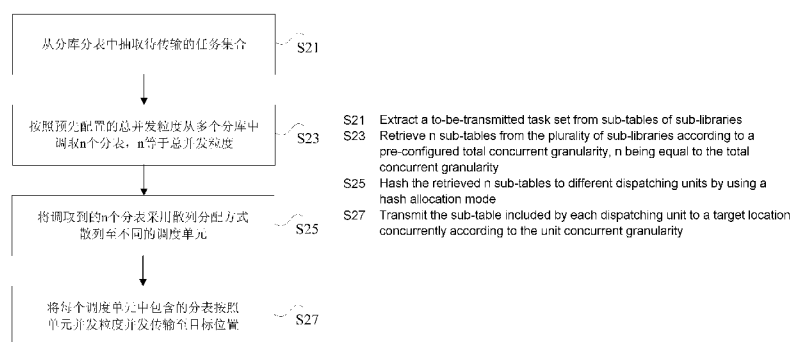


图 2

(57) Abstract: Disclosed are a method, apparatus and system for transmitting tasks based on sub-libraries and sub-tables. The method comprises: extracting a to-be-transmitted task set from sub-tables of sub-libraries, the task set comprising a plurality of sub-libraries, and sub-tables included by all the sub-libraries; retrieving n sub-tables from the plurality of sub-libraries according to a pre-configured total concurrent granularity, n being equal to the total concurrent granularity; hashing the retrieved n sub-tables to different dispatching units by using a hash allocation mode, pre-configured unit concurrent granularities of all the dispatching units being the same as each other; and transmitting the sub-table included by each dispatching unit to a target location concurrently according to the unit concurrent granularity. The present invention resolves the technical problem of low concurrent transmission efficiency of tasks of sub-tables of sub-libraries due to the fact of too large pressure of concurrent data reading, that is performed in DB of the sub-tables of the sub-libraries, of an extraction terminal during concurrent transmission of tasks based on the sub-libraries and the sub-tables in the prior art.

(57) 摘要:

[见续页]

WO 2017/097124 A1



---

本发明公开了一种基于分库分表的任务传输方法、装置及系统。其中，该方法包括：从分库分表中抽取待传输的任务集合，其中，任务集合包括：多个分库，以及每个分库所包含的分表；按照预先配置的总并发粒度从多个分库中调取  $n$  个分表， $n$  等于总并发粒度；将调取到的  $n$  个分表采用散列分配方式散列至不同的调度单元，其中，预先配置每个调度单元的单元并发粒度相同；将每个调度单元中包含的分表按照单元并发粒度并发传输至目标位置。本发明解决了现有技术中在基于分库分表的任务进行并发传输时，抽取端从分库分表 DB 中并发读取数据压力过大，导致并发传输分库分表的任务效率低的技术问题。

## 基于分库分表的任务传输方法、装置及系统

本申请要求 2015 年 12 月 07 日递交的申请号为 201510888403.9、发明名称为“基于分库分表的任务传输方法、装置及系统”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

本发明涉及数据库技术领域，具体而言，涉及一种基于分库分表的任务传输方法、装置及系统。

### 10 背景技术

随着互联网应用的广泛普及，对于一个大型的互联网应用，每天有几十亿的页面访问量，大量数据存储在数据库中。在大数据平台处理数据之前，需要把数据导入大数据平台的存储系统，一般采用 ETL（数据提取、转换和加载，Extract-Transform-Load 的缩写）技术。

15 在 ETL 技术中，数据同步作为数据仓库的出口和入口，扮演着非常重要的角色，特别是离线数据同步，往往一个任务就需要同步上百 GB 甚至 TB 量级的数据，这对数据同步工具的稳定性要求非常高，同时因为抽取端数据库(DB, Database)支持并发读取，对抽取端和写入端数据库 DB 的压力也很大。正是由于大量数据的同步，使得人们可以更方便的关注到原本忽略的数据，使得数据同步过程中产生了更多的长尾任务，导致抽  
20 取端 DB 的压力和长尾渐渐成为数据同步的瓶颈。

现有技术中的解决方案都是针对单库的（即一个任务只抽取一个库），即在底层同步工具上层做一层服务，在服务中进行调度控制，避免同时对一个库抽取的并发任务数过多。

25 随着数据量的增加，单库已经无法满足大数据量的需求，需要将单个数据库进行拆分成多库多表存储数据。针对分库分表的单个任务而言，由于分库分表任务会对多个分库进行抽取，分库的抽取策略直接决定着任务抽取速度，使得上述针对单库的解决方案不再适用。

针对现有技术中在基于分库分表的任务进行并发传输时，抽取端从分库分表 DB 中并发读取数据压力过大，导致并发传输分库分表的任务效率低的技术问题，目前尚未提出有效的解决方案。  
30

## 发明内容

本发明实施例提供了一种基于分库分表的任务传输方法、装置及系统，以至少解决现有技术中在基于分库分表的任务进行并发传输时，抽取端从分库分表 DB 中并发读取数据压力过大，导致并发传输分库分表的任务效率低的技术问题。

5 根据本发明实施例的一个方面，提供了一种基于分库分表的任务传输方法，包括：  
从分库分表中抽取待传输的任务集合，其中，任务集合包括：多个分库，以及每个分库所包含的分表；按照预先配置的总并发粒度从多个分库中调取  $n$  个分表， $n$  等于总并发粒度；将调取到的  $n$  个分表采用散列分配方式散列至不同的调度单元，其中，预先配置每个调度单元的单元并发粒度相同；将每个调度单元中包含的分表按照单元并发粒度并  
10 发传输至目标位置。

根据本发明实施例的另一方面，还提供了一种基于分库分表的任务传输装置，包括：  
抽取模块，用于从分库分表中抽取待传输的任务集合，其中，任务集合包括：多个分库，以及每个分库所包含的分表；调取模块，用于按照预先配置的总并发粒度从多个分库中共调取  $n$  个分表， $n$  等于总并发粒度；处理模块，用于将调取到的  $n$  个分表采用散列分  
15 配方式散列至不同的调度单元，其中，预先配置每个调度单元的单元并发粒度相同；并发模块，用于将每个调度单元中包含的分表按照单元并发粒度并发传输至目标位置。

根据本发明实施例的另一方面，还提供了一种基于分库分表的任务传输系统包括：  
源数据终端，用于存储分库分表；调度终端，与源数据终端通信，用于从分库分表中抽取待传输的任务集合，其中，任务集合包括：多个分库，以及每个分库所包含的分表，  
20 并按照预先配置的总并发粒度从多个分库中共调取  $n$  个分表， $n$  等于总并发粒度，在将调取到的  $n$  个分表采用散列分配方式散列至不同的调度单元之后，将每个调度单元中包含的分表按照单元并发粒度并发传输，其中，预先配置每个调度单元的单元并发粒度相同；目标终端，与调度终端通信，用于接收调度终端并发传输来的任务集合。

在本发明实施例中，如果希望减小分库分表任务的并行传输，可以在抽取待传输的  
25 任务集合之后，按照预先配置的总并发粒度从多个分库中调取  $n$  个分表，通过将调取到的  $n$  个分表采用散列分配方式散列至不同的调度单元，将每个调度单元中包含的分表按照单元并发粒度并发传输至目标位置，本方案可以根据总并发粒度和单元并发粒度对待传输的任务集合进行拆分，以均衡并行传输分库分表的任务。

容易注意到，由于在抽取待传输的任务集合之后，需要在按照预先配置的总并发粒  
30 度从多个分库中调取  $n$  个分表，并通过将调取到的  $n$  个分表采用散列分配方式散列至不

同的调度单元，就可以根据散列后的分表并行发送任务集合，达到均衡并发传输分库分表的任务，因此，通过本申请实施例所提供的方案，可以实现基于分库分表的任务传输，这样不仅实现了分库分表的任务传输，而且，在传输分库分表的任务时，在满足预先配置的总并发粒度的情况下，均衡传输分库分表的任务，因此，可以均衡分库分表中的多个任务，降低并发读取数据的压力，提高并发传输效率。

由此，本申请提供的上述实施例的方案解决了现有技术中在基于分库分表的任务进行并发传输时，抽取端从分库分表 DB 中并发读取数据压力过大，导致并发传输分库分表的任务效率低的技术问题。

## 10 附图说明

此处所说明的附图用来提供对本发明的进一步理解，构成本申请的一部分，本发明的示意性实施例及其说明用于解释本发明，并不构成对本发明的不当限定。在附图中：

图 1 是根据本申请实施例的一种基于分库分表的任务传输方法的计算机终端的硬件结构框图；

15 图 2 是根据本申请实施例的一种基于分库分表的任务传输方法的流程图；

图 3 是根据本申请实施例的分库分表对应调度管理单元的示意图；

图 4 是根据本申请实施例的分库分表散列至调度管理单元的示意图；

图 5 是根据本申请实施例的调度管理单元获取最优解的示意图；

20 图 6 是根据本申请实施例的一种可选的基于分库分表的任务传输方法的交互流程图；

图 7 是根据本申请实施例的一种基于分库分表的任务传输装置的示意图；

图 8 是根据本申请实施例的一种可选的基于分库分表的任务传输装置的示意图；

图 9 是根据本申请实施例的一种可选的基于分库分表的任务传输装置的示意图；

图 10 是根据本申请实施例的一种可选的基于分库分表的任务传输装置的示意图；

25 图 11 是根据本申请实施例的一种可选的基于分库分表的任务传输装置的示意图；

图 12 是根据本申请实施例的一种可选的基于分库分表的任务传输装置的示意图；

图 13 是根据本申请实施例的一种基于分库分表的任务传输系统的示意图；以及

图 14 是根据本申请实施例的一种计算机终端的结构框图。

## 30 具体实施方式

为了使本技术领域的人员更好地理解本发明方案，下面将结合本发明实施例中的附图，对本发明实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例仅仅是本发明一部分的实施例，而不是全部的实施例。基于本发明中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都应当属于本发明保护的

范围。

需要说明的是，本发明的说明书和权利要求书及上述附图中的术语“第一”、“第二”等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，以便这里描述的本发明的实施例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含，例如，包含了一系列步骤或单元的过程、方法、系统、产品或设备不必限于清楚地列出的那些步骤或单元，而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它步骤或单元。

首先，在对本申请实施例进行描述的过程中出现的部分名词或术语适用于如下解释：

ETL：Extract-Transform-Load 简称，用来描述将数据从来源端经过抽取（extract）、转换（transform）、加载（load）至目的端的过程。ETL 是构建数据仓库的重要一环，用户从数据源抽取出所需的数据，经过数据清洗，最终按照预先定义好的数据仓库模块，将数据假造到数据仓库中取。

分库分表：把原本存储于一个库的数据分块存储到多个库上，把原本存储于一个表的数据分块存储到多个表上。分库分表有垂直切分和水平切分两种，其中，垂直切分是指将表按照功能模块、关系密切程度划分出来，部署到不同的库上，例如，可以建立商品数据库存储商品定义表，建立用户数据库存储用户数据表等。水平切分是指当一个表中的数据量过大时，可以把该表的数据按照某种规则进行划分，然后存储到多个结构相同的表和不同的库上。

DB：（Database），数据库，是按照数据结构来组织、存储和管理的数据仓库。

## 实施例 1

根据本发明实施例，还提供了一种基于分库分表的任务传输方法实施例，需要说明的是，在附图的流程图示出的步骤可以在诸如一组计算机可执行指令的计算机系统中执行，并且，虽然在流程图中示出了逻辑顺序，但是在某些情况下，可以以不同于此处的顺序执行所示出或描述的步骤。

本申请实施例一所提供的方法实施例可以在移动终端、计算机终端或者类似的运算装置中执行。以运行在计算机终端上为例，图 1 是根据本申请实施例的一种基于分库分表的任务传输方法的计算机终端的硬件结构框图。如图 1 所示，计算机终端 10 可以包括

一个或多个(图中仅示出一个)处理器 102(处理器 102 可以包括但不限于微处理器 MCU 或可编程逻辑器件 FPGA 等的处理装置)、用于存储数据的存储器 104、以及用于通信功能的传输模块 106。本领域普通技术人员可以理解,图 1 所示的结构仅为示意,其并不对上述电子装置的结构造成限定。例如,计算机终端 10 还可包括比图 1 中所示更多或者更少的组件,或者具有与图 1 所示不同的配置。

存储器 104 可用于存储应用程序的软件程序以及模块,如本发明实施例中的基于分库分表的任务传输方法对应的程序指令/模块,处理器 102 通过运行存储在存储器 104 内的软件程序以及模块,从而执行各种功能应用以及数据处理,即实现上述的应用程序的漏洞检测方法。存储器 104 可包括高速随机存储器,还可包括非易失性存储器,如一个或者多个磁性存储装置、闪存、或者其他非易失性固态存储器。在一些实例中,存储器 104 可进一步包括相对于处理器 102 远程设置的存储器,这些远程存储器可以通过网络连接至计算机终端 10。上述网络的实例包括但不限于互联网、企业内部网、局域网、移动通信网及其组合。

传输装置 106 用于经由一个网络接收或者发送数据。上述的网络具体实例可包括计算机终端 10 的通信供应商提供的无线网络。在一个实例中,传输装置 106 包括一个网络适配器(Network Interface Controller, NIC),其可通过基站与其他网络设备相连从而可与互联网进行通讯。在一个实例中,传输装置 106 可以为射频(Radio Frequency, RF)模块,其用于通过无线方式与互联网进行通讯。

在上述运行环境下,本申请提供了如图 2 所示的基于分库分表的任务传输方法。图 2 是根据本申请实施例的一种基于分库分表的任务传输方法的流程图,如图 2 所示的方法可以包括如下步骤:

步骤 S21,从分库分表中抽取待传输的任务集合,其中,任务集合包括:多个分库,以及每个分库所包含的分表。

可选地,一个任务集合可以包括多个分库,每个分库可以包含多个分表,且每个分表可以记录有多条数据信息,该数据信息可以是注册用户数据信息、访问网页数据信息、购买商品数据信息等等。从分库分表中抽取待传输的任务集合之后,可以根据抽取顺序进行编号。

在一种可选的方案中,在从分库分表中抽取待传输的任务集合之前,可以获取配置文件,该配置文件中记录待传输的多个分库的名称以及每个分库中包含的分表的名称。在读取配置文件之后,按照读取到的多个分库的名称以及多个分表的名称,从源数据库

的分库分表中抽取待传输的多个分库以及每个分库包含的分表。

结合图 3 所示，在一个可选的实施例中，任务集合可以包含 3 个分库，分别是分库 A、分库 B 和分库 C，分库 A 可以包含 4 个分表，分别是分表 T1、分表 T2、分表 T3 和分表 T4，分库 B 可以包含 2 个分表，分别是分表 T1 和分表 T2，分库 C 可以包含 4 个分表，分别是分表 T1、分表 T2、分表 T3 和分表 T4。

步骤 S23，按照预先配置的总并发粒度从多个分库中调取  $n$  个分表， $n$  等于总并发粒度。

可选地，总并发粒度（totalChannel）表示同时并发的任务个数，可以按照预先配置的总并发粒度，依次从任务集合中顺序抽取满足总并发粒度数的  $n$  个分表。上述预先配置的总并发粒度可以按照用户的实际需求进行配置，也可以根据数据库的并发能力进行配置。

仍以上述抽取的任务集合为例，在一种可选的方案中，预先配置的总并发粒度可以为 9，按照总并发粒度依次从分库 A 中提取分表 T1、分表 T2、分表 T3 和分表 T4，从分库 B 中提取分表 T1 和分表 T2，从分库 C 中提取分表 T1、分表 T2 和分表 T3。提取的分表个数与总并发粒度相同。

步骤 S25，将调取到的  $n$  个分表采用散列分配方式散列至不同的调度单元，其中，预先配置每个调度单元的单元并发粒度相同。

可选地，在按照预先配置的总并发粒度从多个分库中调取  $n$  个分表之后，按照散列分配方式，将  $n$  个分表依次散列至不同的调度单元中。上述预先配置的每个调度单元的单元并发粒度（tgChannel）用于表示每个调度单元同时并发的任务个数，单元并发粒度小于等于总并发粒度。

如图 4 所示，在一种可选的方案中，上述调度单元可以是调度管理模块（TG，taskGroup 的简写），分库分表任务可以包括分库 A、分库 B 和分库 C，每个分库都可以包括  $n$  个分表，每一个分表都是一个待传输的任务，将  $3n$  个分表散列到  $n$  个调度管理单元中，每个调度管理单元可以包括三个分表。

仍以上述抽取的任务集合为例，在一种可选的方案中，预先配置每个调度单元的单元并发粒度可以为 3，将提取的 9 个分表散列至不同的调度单元，调度单元 1 可以包括：AT3、BT2 和 CT3，调度单元 2 可以包括：AT1、AT4 和 CT1，调度单元 3 可以包括：AT2、BT1 和 CT2。

需要说明的是，上述散列方式将相同分库中的分表散列在不同的调度单元中，尽量

满足不同分库中的多个分表均衡并发，减小抽取段 DB 的并发压力。

步骤 S27，将每个调度单元中包含的分表按照单元并发粒度并发传输至目标位置。

可选地，上述目标位置可以是目标数据库，用于存储提取的数据信息。将调取到的 n 个分表采用散列分配方式散列至不同的调度单元之后，每个调度单元将散列后的 n 个分表按照单元并发粒度并发传输至目标数据库。

本申请上述实施例一公开的方案中，如果希望减小分库分表任务的并行传输，可以在抽取待传输的任务集合之后，按照预先配置的总并发粒度从多个分库中调取 n 个分表，通过将调取到的 n 个分表采用散列分配方式散列至不同的调度单元，将每个调度单元中包含的分表按照单元并发粒度并发传输至目标位置，本方案可以根据总并发粒度和单元并发粒度对待传输的任务集合进行拆分，以均衡并行传输分库分表的任务。

容易注意到，由于在抽取待传输的任务集合之后，需要在按照预先配置的总并发粒度从多个分库中调取 n 个分表，并通过将调取到的 n 个分表采用散列分配方式散列至不同的调度单元，就可以根据散列后的分表并行发送任务集合，达到均衡并发传输分库分表的任务，因此，通过本申请实施例所提供的方案，可以实现基于分库分表的任务传输，这样不仅实现了分库分表的任务传输，而且，在传输分库分表的任务时，在满足预先配置的总并发粒度的情况下，均衡传输分库分表的任务，因此，可以均衡分库分表中的多个任务，降低并发读取数据的压力，提高并发传输效率。

由此，本申请提供的上述实施例一的方案解决了现有技术中在基于分库分表的任务进行并发传输时，抽取端从分库分表 DB 中并发读取数据压力过大，导致并发传输分库分表的任务效率低的技术问题。

在本申请上述实施例中，在步骤 S25，将调取到的 n 个分表采用散列分配方式散列至不同的调度单元之前，方法还包括如下步骤 S251 和步骤 S255：

步骤 S251，根据总并发粒度和单元并发粒度，确定调度单元的数量，并对每个调度单元分配对应的编号。

可选地，在按照预先配置的总并发粒度从多个分库中调取 n 个分表之后，可以根据预先配置的总并发粒度和单元并发粒度，根据公式  $tgCount = totalChannel / tgChannel$  得到调度单元的数量，然后对每个调度单元分配对应的编号。

结合图 3 所示，在一种可选的方案中，总并发粒度 T 可以为 9，单元并发粒度 t 可以为 3，则调度单元数量  $tgCount = totalChannel / tgChannel = 3$ ，3 个调度单元的编号分别为 TG0、TG1 和 TG2。

步骤 S253, 通过如下公式计算得到任意一个分库中的每个分表  $T_i$  的散列分配值, 其中, 散列分配值用于表征分表  $T_i$  所散列至对应的调度单元的编号  $T_{pos}$ :

$T_{pos} = (TCount + offset) \% tgCount$ ,  $tgCount = totalChannel / tgChannel$ , 其中,  $TCount$  为任意一个分库中每个分表  $T_i$  的编号,  $offset$  为每个分库向对应的调度单元分配的偏移量, 初始值为 0,  $totalChannel$  为总并发粒度,  $tgChannel$  为单元并发粒度。

可选地, 任意一个分库向对应的调度单元分配的偏移量可以为该分库对应前一个分库最后一个分表  $T_n$  的  $lastOffset$ , 例如 A 分库最后一个分表  $T_{4de}$  的  $lastOffset$  为 1, 则 B 分库的偏移量  $BOffset = lastOffset = 1$ , 在计算得到每个调度单元的数量之后, 可以根据公式  $T_{pos} = (TCount + offset) \% tgCount$ , 对每个分库中每个分表对应的调度单元分配的编号  $TGi$ 。

结合图 3 所示, 在一种可选的实施例中, 分库 A、分库 B 和分库 C 的编号分别为 0, 1, 和 2, 分库 A 的偏移量  $Aoffset = 0$ , 分表  $AT1$  对应的调度单元的编号为  $TGi = (1 + Aoffset) \% 3 = 1$ , 依次计算出分库 A 中其他分表对应的调度单元的编号,  $lastOffset = 4 \% 3 = 1$ , 分库 B 的偏移量  $BOffset = lastOffset = 1$ , 则分表  $BT1$  对应的调度单元的编号为  $TGi = (1 + Boffset) \% 3 = 2$ , 依次可以得到 3 个分库中每个分库对应的调度单元的编号。

步骤 S255,  $n$  个分表分别按照计算得到的散列分配值散列至对应的调度单元。

可选地, 在计算得到  $n$  个分表对应的调度单元分配的编号  $TGi$  之后, 将  $n$  个分表散列至对应的调度单元。

结合图 3 所示, 在一种可选的实施例中, 基于步骤 S253 的计算方法, 可以将分库 A 中的分表  $T1$ 、分表  $T2$ 、分表  $T3$  和分表  $T4$ , 分库 B 中的分表  $T1$  和分表  $T2$ , 分库 C 中的分表  $T1$ 、分表  $T2$ 、分表  $T3$  和分表  $T4$  散列到对应的调度单元中, 散列结果可以是分表  $AT3$ 、分表  $BT2$  和分表  $CT3$  对应散列至调度单元  $TG0$ , 分表  $AT1$ 、分表  $AT4$ 、分表  $CT1$  和分表  $CT4$  对应散列至调度单元  $TG1$ , 分表  $AT2$ 、分表  $BT1$  和分表  $CT2$  对应散列至调度单元  $TG2$ 。

由上可知, 在本申请上述实施例中, 按照预先配置的总并发粒度从多个分库中调取  $n$  个分表之后, 可以根据预先配置的总并发粒度和单元并发粒度, 确定调度单元的数量, 并对每个调度单元分配对应的编号, 并通过如下公式计算得到每个分库中的每个分表所对应的调度单元的编号, 实现将分库分表的任务进行切分的目的, 将分库分表的任务均匀的分散到不同的调度单元中, 避免现有技术中分库分表的任务顺序传输造成的资源浪费。

在本申请上述实施例中，在执行步骤 S27，将每个调度单元中包含的分表按照单元并发粒度并发传输至目标位置的过程中，步骤 S22701，实时获取多个分库中除 n 个分表之外的 m 个分表，步骤 S2703，按照总并发粒度将 m 个分表采用散列分配方式散列至不同的调度单元，步骤 S2705，如果调度单元的线程被占满的情况下，将分配至对应的调度单元的分表放置在对应的调度单元的等待队列中，在调度单元的至少一个线程被释放之后，将等待队列中的分表调度至对应的调度单元中。

可选地，将按照预先配置的总并发粒度从所述多个分库中调取 n 个分表并发传输至目标位置时，将多个分库中未传输的多个分表散列至不同的调度单元，如果调度单元的线程被占满，则将散列后的分表放置在等待队列中，直到调度单元中任意一个线程被释放之后，才能被调度单元并发传输至目标位置，如果调度单元的线程未被占满，则散列后的分表直接被调度单元并发传输至目标位置。

仍旧结合图 3 所示，在一种可选的方案中，预先配置的总并发粒度可以为 9，按照总并发粒度提取的分表为分表 AT1、分表 AT2、分表 AT3、分表 AT4、分表 BT1、分表 BT2、分表 CT1、分表 CT2 和分表 CT3。此时，三个分库中剩余的未被传输的分表为分表 CT4，经过散列分配，该分表 CT4 应该被分配至调度单元 TG1，由于调度单元 TG1 线程被占满，因此先将该分表 CT4 进入对调度单元 TG1 的等待队列，如图 3 中虚线所示。等待分表 AT1、AT4 和 CT1 中任意一个分表被传输至目标位置，线程释放之后，等待队列中的分表 CT4 被调度至调度单元 TG1 中，并发传输至目标位置。

由上可知，在本申请上述实施例中，在 n 个分表并发传输至目标位置的过程中，获取多个分库中其余的 m 个分表，在散列分配之后分配给对应的调度单元，如果调度单元线程被占满，则进入等待队列等待调度单元的至少一个线程被释放，实现切分分配分表的目的，在线程全部被占用的情况下进入等待队列，在至少一个线程释放后进入调度单元，保证并发传输分库分表的任务时资源最大化利用。

需要说明的是，如果相同分库不同分表的表的记录数基本相同，抽取时间相差不大，那么通过上述实施例就可以达到目的，使分库分表任务的 DB 抽取压力均衡，但是这种任务只存在于理想情况下，大多数任务相同分库不同分表的记录数不会保持一致，抽取时间也无法预测，这时候需要调度模块去做一个控制调度，选出最优解。

在本申请上述实施例中，在步骤 S27，将每个调度单元中包含的分表按照单元并发粒度并发传输至目标位置之后，上述方法还包括步骤 S271 至步骤 S275：

步骤 S271，在任意一个调度单元中存在空闲的线程的情况下，获取每个分库的当前

并发数，其中，当前并发数用于表征分库中已经被调度至对应的调度单元中的分表数量。

可选地，上述每个分库的当前并发数可以保存在线程并发管理器（TC-MGR，Thread-Concurrency-Manager 的简写）中，在启动线程时，将调度模块中线程对应的分库的并发数加 1，如果线程执行完毕，则将线程并发管理器中对应的分库的并发数减 1，这两个过程标记为同步（synchronized），线程之间互斥访问，为方便将上述两个成为启动通道（holdChannel）和释放通道（releaseChannel）。采用贪心算法，在需要启动线程时，获取调度模块中的每个分库的当前并发数。

如图 5 所示，在一种可选的方案中，线程并发管理器中记录了分库 A、分库 B 和分库 C 的当前并发数 TAC、TAB 和 TAC，分库 A 的当前并发数 TAC=X，分库 B 的当前并发数 TBC=Y，分库 C 的当前并发数 TCC=Z，调度管理单元（即上述的调度单元）包括 3 个通道（即上述的线程），通道 1、通道 2 和通道 3。调度单元 TG0 在 3 个线程中有至少一个线程空闲的情况下，例如线程 1 空闲，获取调度模块中每个分库的当前并发数。

步骤 S273，根据每个分库的当前并发数，从任意一个调度单元的等待队列中调度对应数量的分表。

可选地，当线程被释放之后，获取每个分库的当前并发数，并根据每个分库的当前并发数调度单元的等待队列中相应数量的分表，例如，如果调度单元释放了两个线程，则根据每个分库的当前并发数，调度该调度单元的等待队列中两个分表。

在一种可选的方案中，如果线程并发管理器中记录的当下分库 A 的当前并发数可以为 3，分库 B 的当前并发数可以为 1，分库 C 的当前并发数可以为 2，此时如果需要启动的线程个数为 2，则可以从调度单元的等待队列中调度两个分库 B 的分表，如果等待队列中只有一个分库 B 的分表，则进一步调度等待队列中一个分库 C 的分表。如果调度单元 TG1 的等待队列中保存了分库 B 的一个分表 T4 和其他分库的分表，此时，可以根据每个分库的并发数的比例，以及线程被释放，优先选则调取分表 T4 进入调度单元 TG1，如果 TG1 的线程被全部释放，则可以调取等待队列中除 T4 以外的其他分表。其他任意一个调度单元的调取规则同理。

步骤 S275，将从对应的等待队列中调度对应数量的分表并发传输至目标位置。

可选地，在从任意一个调度单元的等待队列中调度对应数量的分表之后，将该对应数量的分表并发传输至目标位置。

由上可知，在本申请上述实施例中，在任意一个调度单元中存在空闲的线程的情况

下，获取每个分库的当前并发数，并根据每个分库的当前并发数，从任意一个调度单元的等待队列中调度对应数量的分表，并发传输至目标位置，采用贪心算法，每当调度单元有通道线程饥饿时就先做启动通道，从而实现从局部最优到全局最优，进一步实现均衡并发传输分库分表的任务。

- 5        在本申请上述实施例中，步骤 S273，根据每个分库的当前并发数，从任意一个调度单元的等待队列中调度对应数量的分表，包括如下步骤 S2731 至步骤 S2735：

      步骤 S2731，按照每个分库的当前并发数进行排序，确定属于不同分库的分表的调度优先级，其中，分库的当前并发数越低，当前并发数越低的分库中的分表的调度优先级越高。

- 10        可选地，获取到每个分库的当前并发数之后，将每个分库按照当前并发数进行升序排序，当前并发数最低的分库中的分表优先级最高，当前并发数最高的分库中的分配的优先级最低，确定调度单元的等待队列中每个分库分表的调度优先级。

      步骤 S2733，根据任意一个调度单元中存在的空闲线程数量，确定调度数量。

- 可选地，在空闲线程数量小于等于调度单元的等待队列中的分表个数时，确定调度数量  
15        数量为空闲线程数量，在空闲线程数量大于调度单元的等待队列中的分表个数时，确定调度数量为调度单元的等待队列中的分表个数。

      步骤 S2735，在确定第一分库中的分表为调度优先级最高的分表之后，按照调度数量从等待队列中调度属于第一分库的分表。

- 可选地，在确定属于不同分库的分表的调度优先级和调度数量之后，从调度单元  
20        等待队列中调度符合调度数量和调度优先级的分表。

- 在一种可选的方案中，分库 A 的当前并发数可以为 3，分库 B 的当前并发数可以为 1，分库 C 的当前并发数可以为 2，通过分析三个分库的并发度的比例可以确定分库 B 的当前并发数最低，表示在之前的并发过程中，分库 B 中分表的分发或同步效率最低，因此分库 B 的分表的调度优先级最高，系统需要尽快处理分库 B 中的分表，而分库 A 的当前并发数最高，因此分库 A 的分表的调度优先级最低。调度单元的等待队列中有分表  
25        AT5，分表 AT6，分表 BT3 和分表 CT5，在有一个线程空闲的情况下，按照调度优先级的分析结果，此时调度单元从等待队列中调度分表 BT3 进行同步。

- 由上可知，在本申请上述实施例中，获取每个分库的当前并发数之后，根据排序结果确定调度单元中每个分表的调度优先级，并根据空闲线程数从调度单元的等待队列中  
30        调度对应调度优先级和对应数量的分表，从而实现在分表记录数均不相同的情况下，根

据每个分库的当前并发数选取最优解，使分库分表任务的抽取压力均衡。

需要说明的是，为了获取每个分库的当前并发数之后，根据排序结果确定调度单元中每个分表的调度优先级，线程并发管理其器中还需要保存每个调度管理单元的分库使用情况。

- 5        在本申请上述实施例中，在步骤 S2735，按照调度数量从等待队列中调度对应数量的分表之前，上述方法还包括步骤 S27351 至步骤 S27353：

步骤 S27351，读取等待队列中属于第一分库的分表的数量。

步骤 S27353，判断属于第一分库的分表的数量是否大于等于调度数量。

- 10        可选地，如果判断属于第一分库的分表的数量大于等于调度数量，则进入步骤 S2735，按照调度数量从等待队列中调度对应数量的分表；如果判断属于第一分库的分表的数量小于调度数量，则按照调度数量从等待队列中调度属于第一分库和属于其他分库的分表，其中，其他分库为当前并发数大于第一分库的分库。

- 15        在一种可选的方案中，分库 A 的当前并发数可以为 3，分库 B 的当前并发数可以为 1，分库 C 的当前并发数可以为 2，通过分析三个分库的并发度的比例可以确定分库 B 的当前并发数最低，因此分库 B 的分表的调度优先级最高，分库 A 的当前并发数最高，因此分库 A 的分表的调度优先级最低，分库 C 的当前并发数位于两者之间，因此分库 C 的分表的调度优先级也位于两者之间。调度单元 TG0 的等待队列中有分表 AT5，分表 AT6，分表 BT3 和分表 CT5，按照上述一种可选方案所确定的调度优先级的调度规则，可以读取等待队列中属于分库 B 的分表数量为 1，属于分库 C 的分表数量为 1，属于分库 A 的分表数量为 2，在调度单元 TG0 有两个线程空闲的情况下，调度单元 TG0 从等待队列中调度分表 BT3 和 CT5。
- 20

- 25        在另一种可选的方案中，分库 A 的当前并发数可以为 3，分库 B 的当前并发数可以为 1，分库 C 的当前并发数可以为 2，通过分析三个分库的并发度的比例可以确定分库 B 的当前并发数最低，因此分库 B 的分表的调度优先级最高，分库 A 的当前并发数最高，因此分库 A 的分表的调度优先级最低，分库 C 的当前并发数位于两者之间，因此分库 C 的分表的调度优先级也位于两者之间。调度单元的等待队列中有分表 AT5，分表 AT6，分表 BT3 和分表 CT5。按照上述另一种可选方案所确定的调度优先级的调度规则，可以读取等待队列中属于分库 B 的分表数量为 1，属于分库 C 的分表数量为 1，属于分库 A 的分表数量为 2，在调度单元 TG0 有三个线程空闲的情况下，调度单元 TG0 从等待队列中调度分表 BT3、分表 CT5 和分表 AT5。
- 30

在又一种可选的方案中，分库 A 的当前并发数可以为 3，分库 B 的当前并发数可以为 2，分库 C 的当前并发数可以为 2，通过分析三个分库的并发度的比例可以确定分库 B 和分库 C 的当前并发数最低，因此分库 B 和分库 C 的分表的调度优先级最高，分库 A 的当前并发数最高，因此分库 A 的分表的调度优先级最低。调度单元的等待队列中有分表 AT5，分表 AT6，分表 BT3，分表 BT4，分表 BT5，分表 CT5 和分表 CT6。按照上述又一种可选方案所确定的调度优先级的调度规则，可以读取等待队列中属于分库 B 的分表数量为 3，属于分库 C 的分表数量为 2，属于分库 A 的分表数量为 2，在调度单元 TG0 有三个线程空闲的情况下，调度单元 TG0 从等待队列中调度分表 BT3、分表 BT4 和分表 CT5。

由上可知，在本申请上述实施例中，在判断读取到的属于第一分库的分表的数量大于等于调度数量的情况下，按照调度数量从等待队列中调度对应数量的分表；在判断读取到的属于第一分库的分表的数量小于调度数量的情况下，在判断读取到的属于第一分库的分表的数量大于等于调度数量的情况下，按照调度数量从等待队列中调度属于第一分库和属于其他分库的分表，从而实现动态调度均衡，降低分库分表任务对抽取段 DB 的压力，减少长尾任务。

根据本申请上述实施例，对散列至不同的所述调度单元的分表标记分库标识，所述分库标识用于表征所述分表原始对应的分库。

下面结合图 4、图 5 和图 6 详细介绍本申请的一种优选实施例。

如图 4，图 5 和图 6 所示，以任务集合包含多个分库和多个分表为应用场景，提供了一种可选的基于分库分表的任务传输方法，该方法可以包括如下步骤 S61 至步骤 S67：步骤 S61，获取分库分表的任务集合。

可选地，为了获取分库分表的任务集合，调度终端 133 从配置文件中读取分库分表的名称，并从源数据终端 131 的分库分表中提取待传输的任务集合，任务集合中包括多个分库以及分库对应的分表。

在一种可选的方案中，配置文件中包含分库 A、分库 B 和分库 C，分库 A 可以包含 4 个分表，分别是分表 T1、分表 T2、分表 T3 和分表 T4，分库 B 可以包含 2 个分表，分别是分表 T1 和分表 T2，分库 C 可以包含 4 个分表，分别是分表 T1、分表 T2、分表 T3 和分表 T4。

步骤 S63，将分库分表的任务集合散列至多个调度管理单元中。

可选地，将分库分表的任务切分成 n 个单个分表粒度的任务（即上述的 n 个分表），

并按照用户配置的总并发粒度和单个调度管理单元（即上述的调度单元）的并发粒度（即上述的预先配置的总并发粒度和单元并发粒度），将任务按照分库均匀的分配到调度管理单元中，并对每个任务属于哪个分库打标。

在一种可选的方案中，调度终端 133 按照用户配置的总并发粒度从多个分库中调取  
5 n 个分表，按照用户配置的总并发粒度和单个调度管理单元的并发粒度确定调度管理单元的个数，并对每个调度管理单元进行编号；计算 n 个单个分表粒度的任务中每个任务对应的调度管理单元的编号，并将 n 个单个分表粒度的任务散列分配到对应调度管理单元中。

在该实施例中，上述步骤 S63 的实现方式与本申请上述实施例中的步骤 S23 和步骤  
10 S25 的实现方式一致，在此不再赘述。

步骤 S65，多个调度单元从调度模块中获取最优解。

可选地，最优解等于当前并发数最小的 N 个分库，且这 N 个分库在当前 TG 中还存在未消费的任务（结果按照并发数从小到大排列）（即上述的多个分库中除 n 个分表之外的 m 个分表），N 为保持通道的请求参数，一般为调度管理单元中当前饥饿的通道数。

在一种可选的方案中，调度模块中保存了分库分表的任务集合中每个分库的当前并发数和每个调度单元对每个分库的使用情况，调度终端 133 从调度模块中获取每个分库的当前并发数，并将每个分库的当前的并发数进行排序确定每个分库的调度优先级，读取调度单元的空闲线程数量和当前并发数最小的分库中未传输的分表的数量，在空闲线程数量大于分表的数量时，调取当前并发数最小的分库的分表和其他分库的分表；在空  
15  
20 闲线程数量小于等于分表的数量时，调取当前并发数最小的分库的分表。

步骤 S67，并发传输任务集合至目标位置。

可选地，按照用户配置的总并发粒度和单个调度管理单元的并发粒度，并发传输任务集合，在调度管理单元线程被占满的情况下，将未传输的任务放置在调度管理单元的等待队列中，直到调度管理单元存在空闲线程，则并行发送最优解。

在该实施例中，上述步骤 S63 的实现方式与本申请上述实施例中的 S27 的实现方式  
25 一致，在此不再赘述。

需要说明的是，对于前述的各方法实施例，为了简单描述，故将其都表述为一系列的动作组合，但是本领域技术人员应该知悉，本发明并不受所描述的动作顺序的限制，因为依据本发明，某些步骤可以采用其他顺序或者同时进行。其次，本领域技术人员也应该知悉，说明书中所描述的实施例均属于优选实施例，所涉及的动作和模块并不一定  
30

是本发明所必须的。

通过以上的实施方式的描述，本领域的技术人员可以清楚地了解到根据上述实施例的基于分库分表的任务传输方法可借助软件加必需的通用硬件平台的方式来实现，当然也可以通过硬件，但很多情况下前者是更佳的实施方式。基于这样的理解，本发明的技术方案本质上或者说对现有技术做出贡献的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（如 ROM/RAM、磁碟、光盘）中，包括若干指令用以使得一台终端设备（可以是手机，计算机，服务器，或者网络设备等）执行本发明各个实施例所述的方法。

## 实施例 2

根据本发明实施例，还提供了一种用于实施上述基于分库分表的任务传输方法的基于分库分表的任务传输装置，如图 7 所示，该装置包括：抽取模块 71，调取模块 73，处理模块 75 和并发模块 77。

其中，抽取模块 71 用于从分库分表中抽取待传输的任务集合，其中，任务集合包括：多个分库，以及每个分库所包含的分表。调取模块 73 用于按照预先配置的总并发粒度从多个分库中共调取  $n$  个分表， $n$  等于总并发粒度。处理模块 75 用于将调取到的  $n$  个分表采用散列分配方式散列至不同的调度单元，其中，预先配置每个调度单元的单元并发粒度相同。并发模块 77 用于将每个调度单元中包含的分表按照单元并发粒度并发传输至目标位置。

此处需要说明的是，上述抽取模块 71，调取模块 73，处理模块 75 和并发模块 77 对应于实施例一中的步骤 S21 至步骤 S27，四个模块与对应的步骤所实现的实例和应用场景相同，但不限于上述实施例一所公开的内容。需要说明的是，上述模块作为装置的一部分可以运行在实施例一提供的计算机终端 10 中。

本申请上述实施例二公开的方案中，如果希望减小分库分表任务的并行传输，可以在抽取待传输的任务集合之后，按照预先配置的总并发粒度从多个分库中调取  $n$  个分表，通过将调取到的  $n$  个分表采用散列分配方式散列至不同的调度单元，将每个调度单元中包含的分表按照单元并发粒度并发传输至目标位置，本方案可以根据总并发粒度和单元并发粒度对待传输的任务集合进行拆分，以均衡并行传输分库分表的任务。

容易注意到，由于在抽取待传输的任务集合之后，需要在按照预先配置的总并发粒度从多个分库中调取  $n$  个分表，并通过将调取到的  $n$  个分表采用散列分配方式散列至不同的调度单元，就可以根据散列后的分表并行发送任务集合，达到均衡并发传输分库分

表的任务，因此，通过本申请实施例所提供的方案，可以实现基于分库分表的任务传输，这样不仅实现了分库分表的任务传输，而且，在传输分库分表的任务时，在满足预先配置的总并发粒度的情况下，均衡传输分库分表的任务，因此，可以均衡分库分表中的多个任务，降低并发读取数据的压力，提高并发传输效率。

5 由此，本申请提供的上述实施例二的方案解决了现有技术中在基于分库分表的任务进行并发传输时，抽取端从分库分表 DB 中并发读取数据压力过大，导致并发传输分库分表的任务效率低的技术问题。

在本申请上述实施例中，如图 8 所示，上述装置还包括：第一确定模块 81 和计算模块 83，其中，处理模块 75 包括：第一散列分配模块 85。

10 其中，第一确定模块 81 用于根据总并发粒度和单元并发粒度，确定调度单元的数量，并对每个调度单元分配对应的编号。计算模块 83 用于通过如下公式计算得到任意一个分库中的每个分表  $T_i$  的散列分配值，其中，散列分配值用于表征分表  $T_i$  所散列至对应的调度单元的编号  $T_{pos}$ ： $T_{pos} = (TCount + offset) \% tgCount$ ， $tgCount = totalChannel / tgChannel$ ，其中， $TCount$  为任意一个分库中每个分表  $T_i$  的编号， $offset$  为每个分库向对应的调度单元分配的偏移量，初始值为 0， $totalChannel$  为总并发粒度， $tgChannel$  为单元并发粒度。  
15 第一散列分配模块 85 用于  $n$  个分表分别按照计算得到的散列分配值散列至对应的调度单元。

此处需要说明的是，上述第一确定模块 81，计算模块 83 和第一散列分配模块 85 对应于实施例一中的步骤 S251 至步骤 S255，该模块与对应的步骤所实现的实例和应用场景相同，但不限于上述实施例一所公开的内容。需要说明的是，上述模块作为装置的一部分可以运行在实施例一提供的计算机终端 10 中。  
20

在本申请上述实施例中，如图 9 所示，上述装置还包括：第一获取模块 91，第二散列分配模块 93 和子处理模块 95。

其中，第一获取模块 91 用于实时获取多个分库中除  $n$  个分表之外的  $m$  个分表。第二散列分配模块 93 用于按照总并发粒度将  $m$  个分表采用散列分配方式散列至不同的调度单元。子处理模块 95 用于如果调度单元的线程被占满的情况下，将分配至对应的调度单元的分表放置在对应的调度单元的等待队列中，在调度单元的至少一个线程被释放之后，将等待队列中的分表调度至对应的调度单元中。  
25

此处需要说明的是，上述第一获取模块 91，第二散列分配模块 93 和子处理模块 95 对应于实施例一中的步骤 S2701 至步骤 S2705，该模块与对应的步骤所实现的实例和应  
30

用场景相同，但不限于上述实施例一所公开的内容。需要说明的是，上述模块作为装置的一部分可以运行在实施例一提供的计算机终端 10 中。

在本申请上述实施例中，如图 10 所示，上述装置还包括：第二获取模块 101，调度模块 103 和传输模块 105。

5 其中，第二获取模块 101 用于在任意一个调度单元中存在空闲的线程的情况下，获取每个分库的当前并发数，其中，当前并发数用于表征分库中已经被调度至对应的调度单元中的分表数量。调度模块 103 用于根据每个分库的当前并发数，从任意一个调度单元的等待队列中调度对应数量的分表。传输模块 105 用于将从对应的等待队列中调度对应数量的分表并发传输至目标位置。

10 此处需要说明的是，上述第二获取模块 101，调度模块 103 和传输模块 105 对应于实施例一中的步骤 S271 至步骤 S275，该模块与对应的步骤所实现的实例和应用场景相同，但不限于上述实施例一所公开的内容。需要说明的是，上述模块作为装置的一部分可以运行在实施例一提供的计算机终端 10 中。

在本申请上述实施例中，如图 11 所示，调度模块 103 包括：排序模块 111，第二确定模块 113 和子调度模块 115。

15 其中，排序模块 111 用于按照每个分库的当前并发数进行排序，确定属于不同分库的分表的调度优先级，其中，分库的当前并发数越低，当前并发数越低的分库中的分表的调度优先级越高。第二确定模块 113 用于根据任意一个调度单元中存在的空闲线程数量，确定调度数量。子调度模块 115 用于在确定第一分库中的分表为调度优先级最高的分表之后，按照调度数量从等待队列中调度属于第一分库的分表。

20 此处需要说明的是，上述排序模块 111，第二确定模块 113 和子调度模块 115 对应于实施例一中的步骤 S2731 至步骤 S2735，该模块与对应的步骤所实现的实例和应用场景相同，但不限于上述实施例一所公开的内容。需要说明的是，上述模块作为装置的一部分可以运行在实施例一提供的计算机终端 10 中。

25 在本申请上述实施例中，如图 12 所示，上述装置还包括：读取模块 121，判断模块 123，第一执行模块 125 和第二执行模块 127。

30 其中，读取模块 121 用于读取等待队列中属于第一分库的分表的数量。判断模块 123 用于判断属于第一分库的分表的数量是否大于等于调度数量。第一执行模块 125 用于如果大于等于调度数量，则执行子调度模块的功能。第二执行模块 127 用于如果小于调度数量，则按照调度数量从等待队列中调度属于第一分库和属于其他分库的分表，其中，

其他分库为当前并发数大于第一分库的分库。

此处需要说明的是，上述读取模块 121，判断模块 123，第一执行模块 125 和第二执行模块 127 对应于实施例一中的步骤 S27351 至步骤 S27353，该模块与对应的步骤所实现的实例和应用场景相同，但不限于上述实施例一所公开的内容。需要说明的是，上述模块作为装置的一部分可以运行在实施例一提供的计算机终端 10 中。

### 实施例 3

根据本申请实施例，还提供了一种基于分库分表的任务传输系统，如图 13 所示，该系统可以包括：源数据终端 131，调度终端 133 和目标终端 135。

其中，源数据终端 131 用于存储分库分表。

调度终端 133，与源数据终端 131 通信，用于从分库分表中抽取待传输的任务集合，其中，任务集合包括：多个分库，以及每个分库所包含的分表，并按照预先配置的总并发粒度从多个分库中共调取  $n$  个分表， $n$  等于总并发粒度，在将调取到的  $n$  个分表采用散列分配方式散列至不同的调度单元之后，将每个调度单元中包含的分表按照单元并发粒度并发传输，其中，预先配置每个调度单元的单元并发粒度相同。

可选地，一个任务集合可以包括多个分库，每个分库可以包含多个分表，且每个分表可以记录有多条数据信息，该数据信息可以是注册用户数据信息、访问网页数据信息、购买商品数据信息等等。从分库分表中抽取待传输的任务集合之后，可以根据抽取顺序进行编号。

在一种可选的方案中，在从分库分表中抽取待传输的任务集合之前，可以获取配置文件，该配置文件中记录带传输的多个分库的名称以及每个分库中包含的分表的名称。在读取配置文件之后，按照读取到的多个分库的名称以及多个分表的名称，从源数据库的分库分表中抽取带传输的多个分库以及每个分库包含的分表。

结合图 3 所示，在一个可选的实施例中，任务集合可以包含 3 个分库，分别是分库 A、分库 B 和分库 C，分库 A 可以包含 4 个分表，分别是分表 T1、分表 T2、分表 T3 和分表 T4，分库 B 可以包含 2 个分表，分别是分表 T1 和分表 T2，分库 C 可以包含 4 个分表，分别是分表 T1、分表 T2、分表 T3 和分表 T4。

可选地，总开发粒度（totalChannel）表示同时并发的任务个数，可以按照预先配置的总并发粒度，依次从任务集合中顺序抽取满足总开发粒度数的  $n$  个分表。上述预先配置的总并发粒度可以按照用户的实际需求进行配置，也可以根据数据库的并发能力进行配置。

仍以上述抽取的任务集合为例，在一种可选的方案中，预先配置的总开发粒度可以为 9，按照总开发粒度依次从分库 A 中提取分表 T1、分表 T2、分表 T3 和分表 T4，从分库 B 中提取分表 T1 和分表 T2，从分库 C 中提取分表 T1、分表 T2 和分表 T3。提取的分表个数与总开发粒度相同。

5 可选地，在按照预先配置的总开发粒度从多个分库中调取  $n$  个分表之后，按照散列分配方式，将  $n$  个分表依次散列至不同的调度单元中。上述预先配置的每个调度单元的单元并发粒度（tgChannel）用于表示每个调度单元同时并发的任务个数，单元并发粒度小于等于总并发粒度。

如图 4 所示，在一种可选的方案中，上述调度单元可以是调度管理模块（TG，  
10 taskGroup 的简写），分库分表任务可以包括分库 A、分库 B 和分库 C，每个分库都可以包括  $n$  个分表，每一个分表都是一个待传输的任务，将  $3n$  个分表散列到  $n$  个调度管理单元中，每个调度管理单元可以包括三个分表。

仍以上述抽取的任务集合为例，在一种可选的方案中，预先配置每个调度单元的单元并发粒度可以为 3，将提取的 9 个分表散列至不同的调度单元，调度单元 1 可以包括：  
15 AT3、BT2 和 CT3，调度单元 2 可以包括：AT1、AT4 和 CT1，调度单元 3 可以包括：AT2、BT1 和 CT2。

需要说明的是，上述散列方式将相同分库中的分表散列在不同的调度单元中，尽量满足不同分库中的多个分表均衡并发，减小抽取段 DB 的并发压力。

目标终端 135，与调度终端 133 通信，用于接收调度终端并发传输来的任务集合。

20 可选地，上述目标终端可以是目标数据库，用于存储提取的数据信息。将调取到的  $n$  个分表采用散列分配方式散列至不同的调度单元之后，每个调度单元将散列后的  $n$  个分表按照单元并发粒度并发传输至目标数据库。

本申请上述实施例三公开的方案中，如果希望减小分库分表任务的并行传输，可以在抽取待传输的任务集合之后，按照预先配置的总并发粒度从多个分库中调取  $n$  个分表，  
25 通过将调取到的  $n$  个分表采用散列分配方式散列至不同的调度单元，将每个调度单元中包含的分表按照单元并发粒度并发传输至目标位置，本方案可以根据总并发粒度和单元并发粒度对待传输的任务集合进行拆分，以均衡并行传输分库分表的传输。

容易注意到，由于在抽取待传输的任务集合之后，需要在按照预先配置的总并发粒度从多个分库中调取  $n$  个分表，并通过将调取到的  $n$  个分表采用散列分配方式散列至不同的调度单元，就可以根据散列后的分表并行发送任务集合，达到均衡并发传输分库分  
30

表的任务，因此，通过本申请实施例所提供的方案，可以实现基于分库分表的任务传输，这样不仅实现了分库分表的任务传输，而且，在传输分库分表的任务时，在满足预先配置的总并发粒度的情况下，均衡传输分库分表的任务，因此，可以均衡分库分表中的多个任务，降低并发读取数据的压力，提高并发传输效率。

5 由此，本申请提供的上述实施例三的方案解决了现有技术中在基于分库分表的任务进行并发传输时，抽取端从分库分表 DB 中并发读取数据压力过大，导致并发传输分库分表的任务效率低的技术问题。

在本申请提供的一种可选实施例中，上述调度终端 133 还用于根据总并发粒度和单元并发粒度，确定调度单元的数量，并对每个调度单元分配对应的编号，通过如下公式  
10 计算得到任意一个分库中的每个分表  $T_i$  的散列分配值，其中，散列分配值用于表征分表  $T_i$  所散列至对应的调度单元的编号  $T_{pos}$ ： $T_{pos} = (TCount + offset) \% tgCount$ ， $tgCount = totalChannel / tgChannel$ ，其中， $TCount$  为任意一个分库中每个分表  $T_i$  的编号， $offset$  为每个分库向对应的调度单元分配的偏移量，初始值为 0， $totalChannel$  为总并发粒度， $tgChannel$  为单元并发粒度，并将  $n$  个分表分别按照计算得到的散列分配值散列至对  
15 应的调度单元。

可选地，在按照预先配置的总并发粒度从多个分库中调取  $n$  个分表之后，可以根据预先配置的总并发粒度和单元并发粒度，根据公式  $tgCount = totalChannel / tgChannel$  得到调度单元的数量，然后对每个调度单元分配对应的编号。

结合图 3 所示，在一种可选的方案中，总并发粒度  $T$  可以为 9，单元并发粒度  $t$  可以  
20 为 3，则调度单元数量  $tgCount = totalChannel / tgChannel = 3$ ，3 个调度单元的编号分别为  $TG_0$ 、 $TG_1$  和  $TG_2$ 。

可选地，任意一个分库向对应的调度单元分配的偏移量可以为该分库对应前一个分库最后一个分表  $T_n$  的  $lastOffset$ ，例如 A 分库最后一个分表  $T_{4de}$  的  $lastOffset$  为 1，则 B 分库的偏移量  $BOffset = lastOffset = 1$ ，在计算得到每个调度单元的数量之后，可以根据公  
25 式  $T_{pos} = (TCount + offset) \% tgCount$ ，对每个分库中每个分表对应的调度单元分配的编号  $TG_i$ 。

结合图 3 所示，在一种可选的实施例中，分库 A、分库 B 和分库 C 的编号分别为 0，1 和 2，分库 A 的偏移量  $AOffset = 0$ ，分表  $AT_1$  对应的调度单元的编号为  $TG_i = (1 + AOffset) \% 3 = 1$ ，依次计算出分库 A 中其他分表对应的调度单元的编号， $lastOffset = 4 \% 3 = 1$ ，分库  
30 B 的偏移量  $BOffset = lastOffset = 1$ ，则分表  $BT_1$  对应的调度单元的编号为  $TG_i = (1 + BOffset)$

%3=2, 依次可以得到 3 个分库中每个分库对应的调度单元的编号。

可选地, 在计算得到  $n$  个分表对应的调度单元分配的编号  $TG_i$  之后, 将  $n$  个分表散列至对应的调度单元。

结合图 3 所示, 在一种可选的实施例中, 基于步骤 S253 的计算方法, 可以将分库 A 中的分表 T1、分表 T2、分表 T3 和分表 T4, 分库 B 中的分表 T1 和分表 T2, 分库 C 中的分表 T1、分表 T2、分表 T3 和分表 T4 散列到对应的调度单元中, 散列结果可以是分表 AT3、分表 BT2 和分表 CT3 对应散列至调度单元 TG0, 分表 AT1、分表 AT4、分表 CT1 和分表 CT4 对应散列至调度单元 TG1, 分表 AT2、分表 BT1 和分表 CT2 对应散列至调度单元 TG2。

由上可知, 在本申请上述实施例中, 按照预先配置的总并发粒度从多个分库中调取  $n$  个分表之后, 可以根据预先配置的总并发粒度和单元并发粒度, 确定调度单元的数量, 并对每个调度单元分配对应的编号, 并通过如下公式计算得到每个分库中的每个分表所对应的调度单元的编号, 实现将分库分表的任务进行切分的目的, 将分库分表的任务均匀分散到不同的调度单元中, 避免现有技术中分库分表的任务顺序传输造成的资源浪费。

在本申请提供的一种可选实施例中, 上述调度终端 133 还用于将每个调度单元中包含的分表按照单元并发粒度并发传输至目标位置的过程中, 实时获取多个分库中除  $n$  个分表之外的  $m$  个分表, 并按照总并发粒度将  $m$  个分表采用散列分配方式散列至不同的调度单元, 如果调度单元的线程被占满的情况下, 将分配至对应的调度单元的分表放置在对应的调度单元的等待队列中, 在调度单元的至少一个线程被释放之后, 将等待队列中的分表调度至对应的调度单元中。

可选地, 将按照预先配置的总并发粒度从所述多个分库中调取  $n$  个分表并发传输至目标位置时, 将多个分库中未传输的多个分表散列至不同的调度单元, 如果调度单元的线程被占满, 则将散列后的分表放置在等待队列中, 直到调度单元中任意一个线程被释放之后, 才能被调度单元并发传输至目标位置, 如果调度单元的线程未被占满, 则散列后的分表直接被调度单元并发传输至目标位置。

仍旧结合图 3 所示, 在一种可选的方案中, 预先配置的总开发粒度可以为 9, 按照总开发粒度提取的分表为分表 AT1、分表 AT2、分表 AT3、分表 AT4、分表 BT1、分表 BT2、分表 CT1、分表 CT2 和分表 CT3。此时, 三个分库中剩余的未被传输的分表为分表 CT4, 经过散列分配, 该分表 CT4 应该被分配至调度单元 TG1, 由于调度单元 TG1

线程被占满,因此先将该分表 CT4 进入对调度单元 TG1 的等待队列,如图 3 中虚线所示。等待分表 AT1、AT4 和 CT1 中任意一个分表被传输至目标位置,线程释放之后,等待队列中的分表 CT4 被调度至调度单元 TG1 中,并发传输至目标位置。

由上可知,在本申请上述实施例中,在  $n$  个分表并发传输至目标位置的过程中,获取多个分库中其余的  $m$  个分表,在散列分配之后分配给对应的调度单元,如果调度单元线程被占满,则进入等待队列等待调度单元的至少一个线程被释放,实现切分分配分表的目的,在线程全部被占用的情况下进入等待队列,在至少一个线程释放后进入调度单元,保证并发传输分库分表的任务时资源最大化利用。

需要说明的是,如果相同分库不同分表的表的记录数基本相同,抽取时间相差不大,那么通过上述实施例就可以达到目的,使分库分表任务的 DB 抽取压力均衡,但是这种任务只存在于理想情况下,大多数任务相同分库不同分表的记录数不会保持一致,抽取时间也无法预测,这时候需要调度模块去做一个控制调度,选出最优解。

在本申请提供的一种可选实施例中,上述调度终端 133 还用于将每个调度单元中包含的分表按照单元并发粒度并发传输至目标位置之后,在任意一个调度单元中存在空闲的线程的情况下,获取每个分库的当前并发数,其中,当前并发数用于表征分库中已经被调度至对应的调度单元中的分表数量,并根据每个分库的当前并发数,从任意一个调度单元的等待队列中调度对应数量的分表,将从对应的等待队列中调度对应数量的分表并发传输至目标位置。

可选地,上述每个分库的当前并发数可以保存在线程并发管理器(TC-MGR, Thread-Concurrency-Manager 的简写)中,在启动线程时,将调度模块中线程对应的分库的并发数加 1,如果线程执行完毕,则将线程并发管理器中对应的分库的并发数减 1,这两个过程标记为同步(synchronized),线程之间互斥访问,为方便将上述两个成为启动通道(holdChannel)和释放通道(releaseChannel)。采用贪心算法,在需要启动线程时,获取调度模块中的每个分库的当前并发数。

如图 5 所示,在一种可选的方案中,线程并发管理器中记录了分库 A、分库 B 和分库 C 的当前并发数 TAC、TAB 和 TAC,分库 A 的当前并发数 TAC=X,分库 B 的当前并发数 TBC=Y,分库 C 的当前并发数 TCC=Z,调度管理单元(即上述的调度单元)包括 3 个通道(即上述的线程),通道 1、通道 2 和通道 3。调度单元 TG0 在 3 个线程中有至少一个线程空闲的情况下,例如线程 1 空闲,获取调度模块中每个分库的当前并发数。

可选地，当线程被释放之后，获取每个分库的当前并发数，并根据每个分库的当前并发数调度单元的等待队列中相应数量的分表，例如，如果调度单元释放了两个线程，则根据每个分库的当前并发数，调度该调度单元的等待队列中两个分表。

在一种可选的方案中，如果线程并发管理器中记录的当下分库 A 的当前并发数可以为 3，分库 B 的当前并发数可以为 1，分库 C 的当前并发数可以为 2，此时如果需要启动的线程个数为 2，则可以从调度单元的等待队列中调度两个分库 B 的分表，如果等待队列中只有一个分库 B 的分表，则进一步调度等待队列中一个分库 C 的分表。如果调度单元 TG1 的等待队列中保存了分库 B 的一个分表 T4 和其他分库的分表，此时，可以根据每个分库的并发数的比例，以及线程被释放，优先选则调取分表 T4 进入调度单元 TG1，  
10 如果 TG1 的线程被全部释放，则可以调取等待队列中除 T4 以外的其他分表。其他任意一个调度单元的调取规则同理。

可选地，在从任意一个调度单元的等待队列中调度对应数量的分表之后，将该对应数量的分表并发传输至目标位置。

由上可知，在本申请上述实施例中，在任意一个调度单元中存在空闲的线程的情况下，  
15 下，获取每个分库的当前并发数，并根据每个分库的当前并发数，从任意一个调度单元的等待队列中调度对应数量的分表，并发传输至目标位置，采用贪心算法，每当调度单元有通道线程饥饿时就先做启动通道，从而实现从局部最优到全局最优，进一步实现均衡并发传输分库分表的任务。

在本申请提供的一种可选实施例中，上述调度终端 133 还用于按照每个分库的当前  
20 并发数进行排序，确定属于不同分库的分表的调度优先级，其中，分库的当前并发数越低，当前并发数越低的分库中的分表的调度优先级越高，并根据任意一个调度单元中存在的空闲线程数量，确定调度数量，在确定第一分库中的分表为调度优先级最高的分表之后，按照调度数量从等待队列中调度属于第一分库的分表。

可选地，获取到每个分库的当前并发数之后，将每个分库按照当前并发数进行升序  
25 排序，当前并发数最低的分库中的分表优先级最高，当前并发数最高的分库中的分配的优先级最低，确定调度单元的等待队列中每个分库分表的调度优先级。

可选地，在空闲线程数量小于等于调度单元的等待队列中的分表个数时，确定调度数量为空闲线程数量，在空闲线程数量大于调度单元的等待队列中的分表个数时，确定调度数量为调度单元的等待队列中的分表个数。

30 可选地，在确定属于不同分库的分表的调度优先级和调度数量之后，从调度单元的

等待队列中调度符合调度数量和调度优先级的分表。

在一种可选的方案中，分库 A 的当前并发数可以为 3，分库 B 的当前并发数可以为 1，分库 C 的当前并发数可以为 2，通过分析三个分库的并发度的比例可以确定分库 B 的当前并发数最低，表示在之前的并发过程中，分库 B 中分表的分发或同步效率最低，因此分库 B 的分表的调度优先级最高，系统需要尽快处理分库 B 中的分表，而分库 A 的当前并发数最高，因此分库 A 的分表的调度优先级最低。调度单元的等待队列中有分表 AT5，分表 AT6，分表 BT3 和分表 CT5，在有一个线程空闲的情况下，按照调度优先级的分析结果，此时调度单元从等待队列中调度分表 BT3 进行同步。

由上可知，在本申请上述实施例中，获取每个分库的当前并发数之后，根据排序结果确定调度单元中每个分表的调度优先级，并根据空闲线程数从调度单元的等待队列中调度对应调度优先级和对应数量的分表，从而实现在分表记录数均不相同的情况下，根据每个分库的当前并发数选取最优解，使分库分表任务的抽取压力均衡。

需要说明的是，为了获取每个分库的当前并发数之后，根据排序结果确定调度单元中每个分表的调度优先级，线程并发管理其器中还需要保存每个调度管理单元的分库使用情况。

在本申请提供的一种可选实施例中，上述调度终端 133 还用于按照调度数量从等待队列中调度对应数量的分表之前，读取等待队列中属于第一分库的分表的数量，并判断属于第一分库的分表的数量是否大于等于调度数量，如果判断属于第一分库的分表的数量大于等于调度数量，则按照调度数量从等待队列中调度对应数量的分表；如果判断属于第一分库的分表的数量小于调度数量，则按照调度数量从等待队列中调度属于第一分库和属于其他分库的分表，其中，其他分库为当前并发数大于第一分库的分库。

在一种可选的方案中，分库 A 的当前并发数可以为 3，分库 B 的当前并发数可以为 1，分库 C 的当前并发数可以为 2，通过分析三个分库的并发度的比例可以确定分库 B 的当前并发数最低，因此分库 B 的分表的调度优先级最高，分库 A 的当前并发数最高，因此分库 A 的分表的调度优先级最低，分库 C 的当前并发数位于两者之间，因此分库 C 的分表的调度优先级也位于两者之间。调度单元 TG0 的等待队列中有分表 AT5，分表 AT6，分表 BT3 和分表 CT5，按照上述一种可选方案所确定的调度优先级的调度规则，可以读取等待队列中属于分库 B 的分表数量为 1，属于分库 C 的分表数量为 1，属于分库 A 的分表数量为 2，在调度单元 TG0 有两个线程空闲的情况下，调度单元 TG0 从等待队列中调度分表 BT3 和 CT5。

在另一种可选的方案中，分库 A 的当前并发数可以为 3，分库 B 的当前并发数可以为 1，分库 C 的当前并发数可以为 2，通过分析三个分库的并发度的比例可以确定分库 B 的当前并发数最低，因此分库 B 的分表的调度优先级最高，分库 A 的当前并发数最高，因此分库 A 的分表的调度优先级最低，分库 C 的当前并发数位于两者之间，因此分库 C 的分表的调度优先级也位于两者之间。调度单元的等待队列中有分表 AT5，分表 AT6，分表 BT3 和分表 CT5。按照上述另一种可选方案所确定的调度优先级的调度规则，可以读取等待队列中属于分库 B 的分表数量为 1，属于分库 C 的分表数量为 1，属于分库 A 的分表数量为 2，在调度单元 TG0 有三个线程空闲的情况下，调度单元 TG0 从等待队列中调度分表 BT3、分表 CT5 和分表 AT5。

在又一种可选的方案中，分库 A 的当前并发数可以为 3，分库 B 的当前并发数可以为 2，分库 C 的当前并发数可以为 2，通过分析三个分库的并发度的比例可以确定分库 B 和分库 C 的当前并发数最低，因此分库 B 和分库 C 的分表的调度优先级最高，分库 A 的当前并发数最高，因此分库 A 的分表的调度优先级最低。调度单元的等待队列中有分表 AT5，分表 AT6，分表 BT3，分表 BT4，分表 BT5，分表 CT5 和分表 CT6。按照上述又一种可选方案所确定的调度优先级的调度规则，可以读取等待队列中属于分库 B 的分表数量为 3，属于分库 C 的分表数量为 2，属于分库 A 的分表数量为 2，在调度单元 TG0 有三个线程空闲的情况下，调度单元 TG0 从等待队列中调度分表 BT3、分表 BT4 和分表 CT5。

由上可知，在本申请上述实施例中，在判断读取到的属于第一分库的分表的数量大于等于调度数量的情况下，按照调度数量从等待队列中调度对应数量的分表；在判断读取到的属于第一分库的分表的数量小于调度数量的情况下，在判断读取到的属于第一分库的分表的数量大于等于调度数量的情况下，按照调度数量从等待队列中调度属于第一分库和属于其他分库的分表，从而实现动态调度均衡，降低分库分表任务对抽取段 DB 的压力，减少长尾任务。

在本申请提供的一种可选实施例中，上述调度终端 133 还用于对散列至不同的所述调度单元的分表标记分库标识，所述分库标识用于表征所述分表原始对应的分库。

#### 实施例 4

本发明的实施例可以提供一种计算机终端，该计算机终端可以是计算机终端群中的任意一个计算机终端设备。可选地，在本实施例中，上述计算机终端也可以替换为移动终端等终端设备。

可选地，在本实施例中，上述计算机终端可以位于计算机网络的多个网络设备中的至少一个网络设备。

在本实施例中，上述计算机终端可以执行应用程序的漏洞检测方法中以下步骤的程序代码：从分库分表中抽取待传输的任务集合，其中，任务集合包括：多个分库，以及  
5 每个分库所包含的分表；按照预先配置的总并发粒度从多个分库中调取  $n$  个分表， $n$  等于总并发粒度；将调取到的  $n$  个分表采用散列分配方式散列至不同的调度单元，其中，预先配置每个调度单元的单元并发粒度相同；将每个调度单元中包含的分表按照单元并发粒度并发传输至目标位置。

可选地，图 14 是根据本发明实施例的一种计算机终端的结构框图。如图 14 所示，  
10 该计算机终端 A 可以包括：一个或多个（图中仅示出一个）处理器、存储器、以及传输装置。

其中，存储器可用于存储软件程序以及模块，如本发明实施例中的安全漏洞检测方法和装置对应的程序指令/模块，处理器通过运行存储在存储器内的软件程序以及模块，从而执行各种功能应用以及数据处理，即实现上述的系统漏洞攻击的检测方法。存储器  
15 可包括高速随机存储器，还可以包括非易失性存储器，如一个或者多个磁性存储装置、闪存、或者其他非易失性固态存储器。在一些实例中，存储器可进一步包括相对于处理器远程设置的存储器，这些远程存储器可以通过网络连接至终端 A。上述网络的实例包括但不限于互联网、企业内部网、局域网、移动通信网及其组合。

处理器可以通过传输装置调用存储器存储的信息及应用程序，以执行下述步骤：从  
20 分库分表中抽取待传输的任务集合，其中，任务集合包括：多个分库，以及每个分库所包含的分表；按照预先配置的总并发粒度从多个分库中调取  $n$  个分表， $n$  等于总并发粒度；将调取到的  $n$  个分表采用散列分配方式散列至不同的调度单元，其中，预先配置每个调度单元的单元并发粒度相同；将每个调度单元中包含的分表按照单元并发粒度并发传输至目标位置。

25 可选的，上述处理器还可以执行如下步骤的程序代码：根据总并发粒度和单元并发粒度，确定调度单元的数量，并对每个调度单元分配对应的编号；通过如下公式计算得到任意一个分库中的每个分表  $T_i$  的散列分配值，其中，散列分配值用于表征分表  $T_i$  所散列至对应的调度单元的编号  $T_{pos}$ ： $T_{pos} = (TCount + offset) \% tgCount$ ， $tgCount = totalChannel / tgChannel$ ，其中， $TCount$  为任意一个分库中每个分表  $T_i$  的编号，  
30  $offset$  为每个分库向对应的调度单元分配的偏移量，初始值为 0， $totalChannel$  为总并发粒

度， $tgChannel$  为单元并发粒度；其中， $n$  个分表分别按照计算得到的散列分配值散列至对应的调度单元。

可选的，上述处理器还可以执行如下步骤的程序代码：实时获取多个分库中除  $n$  个分表之外的  $m$  个分表，并按照总并发粒度将  $m$  个分表采用散列分配方式散列至不同的调度单元，其中，如果调度单元的线程被占满的情况下，将分配至对应的调度单元的分表放置在对应的调度单元的等待队列中，在调度单元的至少一个线程被释放之后，将等待队列中的分表调度至对应的调度单元中。

可选的，上述处理器还可以执行如下步骤的程序代码：在任意一个调度单元中存在空闲的线程的情况下，获取每个分库的当前并发数，其中，当前并发数用于表征分库中已经被调度至对应的调度单元中的分表数量；根据每个分库的当前并发数，从任意一个调度单元的等待队列中调度对应数量的分表；将从对应的等待队列中调度对应数量的分表并发传输至目标位置。

可选的，上述处理器还可以执行如下步骤的程序代码：按照每个分库的当前并发数进行排序，确定属于不同分库的分表的调度优先级，其中，分库的当前并发数越低，当前并发数越低的分库中的分表的调度优先级越高；根据任意一个调度单元中存在的空闲线程数量，确定调度数量；在确定第一分库中的分表为调度优先级最高的分表之后，按照调度数量从等待队列中调度属于第一分库的分表。

可选的，上述处理器还可以执行如下步骤的程序代码：读取等待队列中属于第一分库的分表的数量；判断属于第一分库的分表的数量是否大于等于调度数量；其中，如果大于等于调度数量，则进入按照调度数量从等待队列中调度对应数量的分表的步骤；如果小于调度数量，则按照调度数量从等待队列中调度属于第一分库和属于其他分库的分表，其中，其他分库为当前并发数大于第一分库的分库。

采用本发明实施例，如果希望减小分库分表任务的并行传输，可以在抽取待传输的任务集合之后，按照预先配置的总并发粒度从多个分库中调取  $n$  个分表，通过将调取到的  $n$  个分表采用散列分配方式散列至不同的调度单元，将每个调度单元中包含的分表按照单元并发粒度并发传输至目标位置，本方案可以根据总并发粒度和单元并发粒度对待传输的任务集合进行拆分，以均衡并行传输分库分表的任务。

容易注意到，由于在抽取待传输的任务集合之后，需要在按照预先配置的总并发粒度从多个分库中调取  $n$  个分表，并通过将调取到的  $n$  个分表采用散列分配方式散列至不同的调度单元，就可以根据散列后的分表并行发送任务集合，达到均衡并发传输分库分

表的任务，因此，通过本申请实施例所提供的方案，可以实现基于分库分表的任务传输，这样不仅实现了分库分表的任务传输，而且，在传输分库分表的任务时，在满足预先配置的总并发粒度的情况下，均衡传输分库分表的任务，因此，可以均衡分库分表中的多个任务，降低并发读取数据的压力，提高并发传输效率。

5 由此，本申请提供的上述实施例的方案解决了现有技术中在基于分库分表的任务进行并发传输时，抽取端从分库分表 DB 中并发读取数据压力过大，导致并发传输分库分表的任务效率低的技术问题。

本领域普通技术人员可以理解，图 14 所示的结构仅为示意，计算机终端也可以是智能手机（如Android 手机、iOS 手机等）、平板电脑、掌上电脑以及移动互联网设备  
10 （MobileInternetDevices, MID）、PAD 等终端设备。图 14 其并不对上述电子装置的结构造成限定。例如，计算机终端 10 还可包括比图 14 中所示更多或者更少的组件（如网络接口、显示装置等），或者具有与图 14 所示不同的配置。

本领域普通技术人员可以理解上述实施例的各种方法中的全部或部分步骤是可以通  
过程序来指令终端设备相关的硬件来完成，该程序可以存储于一计算机可读存储介质中，  
15 存储介质可以包括：闪存盘、只读存储器（Read-Only Memory, ROM）、随机存取器（Random Access Memory, RAM）、磁盘或光盘等。

## 实施例 5

本发明的实施例还提供了一种存储介质。可选地，在本实施例中，上述存储介质可以用于保存上述实施例一所提供的基于分库分表的任务传输方法所执行的程序代码。

20 可选地，在本实施例中，上述存储介质可以位于计算机网络中计算机终端群中的任意一个计算机终端中，或者位于移动终端群中的任意一个移动终端中。

可选地，在本实施例中，存储介质被设置为存储用于执行以下步骤的程序代码：从分库分表中抽取待传输的任务集合，其中，任务集合包括：多个分库，以及每个分库所包含的分表；按照预先配置的总并发粒度从多个分库中调取 n 个分表，n 等于总并发粒  
25 度；将调取到的 n 个分表采用散列分配方式散列至不同的调度单元，其中，预先配置每个调度单元的单元并发粒度相同；将每个调度单元中包含的分表按照单元并发粒度并发传输至目标位置。

可选地，在本实施例中，存储介质被设置为存储用于执行以下步骤的程序代码：根据总并发粒度和单元并发粒度，确定调度单元的数量，并对每个调度单元分配对应的编  
30 号；通过如下公式计算得到任意一个分库中的每个分表 Ti 的散列分配值，其中，散列分

配值用于表征分表  $T_i$  所散列至对应的调度单元的编号  $T_{pos}$  :

$T_{pos} = (TCount + offset) \% tgCount$ ,  $tgCount = totalChannel / tgChannel$ , 其中,  $TCount$  为任意一个分库中每个分表  $T_i$  的编号,  $offset$  为每个分库向对应的调度单元分配的偏移量, 初始值为 0,  $totalChannel$  为总并发粒度,  $tgChannel$  为单元并发粒度; 其中,  $n$  个分表分别按照计算得到的散列分配值散列至对应的调度单元。

可选地, 在本实施例中, 存储介质被设置为存储用于执行以下步骤的程序代码: 在将每个调度单元中包含的分表按照单元并发粒度并发传输至目标位置的过程中, 实时获取多个分库中除  $n$  个分表之外的  $m$  个分表, 并按照总并发粒度将  $m$  个分表采用散列分配方式散列至不同的调度单元, 其中, 如果调度单元的线程被占满的情况下, 将分配至对应的调度单元的分表放置在对应的调度单元的等待队列中, 在调度单元的至少一个线程被释放之后, 将等待队列中的分表调度至对应的调度单元中。

可选地, 在本实施例中, 存储介质被设置为存储用于执行以下步骤的程序代码: 在任意一个调度单元中存在空闲的线程的情况下, 获取每个分库的当前并发数, 其中, 当前并发数用于表征分库中已经被调度至对应的调度单元中的分表数量; 根据每个分库的当前并发数, 从任意一个调度单元的等待队列中调度对应数量的分表; 将从对应的等待队列中调度对应数量的分表并发传输至目标位置。

可选地, 在本实施例中, 存储介质被设置为存储用于执行以下步骤的程序代码: 按照每个分库的当前并发数进行排序, 确定属于不同分库的分表的调度优先级, 其中, 分库的当前并发数越低, 当前并发数越低的分库中的分表的调度优先级越高; 根据任意一个调度单元中存在的空闲线程数量, 确定调度数量; 在确定第一分库中的分表为调度优先级最高的分表之后, 按照调度数量从等待队列中调度属于第一分库的分表。

可选地, 在本实施例中, 存储介质被设置为存储用于执行以下步骤的程序代码: 读取等待队列中属于第一分库的分表的数量; 判断属于第一分库的分表的数量是否大于等于调度数量; 其中, 如果大于等于调度数量, 则进入按照调度数量从等待队列中调度对应数量的分表的步骤; 如果小于调度数量, 则按照调度数量从等待队列中调度属于第一分库和属于其他分库的分表, 其中, 其他分库为当前并发数大于第一分库的分库。

可选地, 在本实施例中, 存储介质被设置为存储: 对散列至不同的调度单元的分表标记分库标识, 分库标识用于表征分表原始对应的分库。

上述本发明实施例序号仅仅为了描述, 不代表实施例的优劣。

在本发明的上述实施例中, 对各个实施例的描述都各有侧重, 某个实施例中

述的部分，可以参见其他实施例的相关描述。

在本申请所提供的几个实施例中，应该理解到，所揭露的技术内容，可通过其它的方式实现。其中，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，单元或模块的间接耦合或通信连接，可以是电性或其它的形式。

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，在本发明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取存储介质中。基于这样的理解，本发明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（可为个人计算机、服务器或者网络设备等）执行本发明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、只读存储器（ROM, Read-Only Memory）、随机存取存储器（RAM, Random Access Memory）、移动硬盘、磁碟或者光盘等各种可以存储程序代码的介质。

以上所述仅是本发明的优选实施方式，应当指出，对于本技术领域的普通技术人员来说，在不脱离本发明原理的前提下，还可以做出若干改进和润饰，这些改进和润饰也应视为本发明的保护范围。

## 权利要求书

1. 一种基于分库分表的任务传输方法，其特征在于，包括：

从分库分表中抽取待传输的任务集合，其中，所述任务集合包括：多个分库，以及每个分库所包含的分表；

5 按照预先配置的总并发粒度从所述多个分库中调取  $n$  个分表， $n$  等于所述总并发粒度；

将调取到的所述  $n$  个分表采用散列分配方式散列至不同的调度单元，其中，预先配置每个调度单元的单元并发粒度相同；

将所述每个调度单元中包含的分表按照所述单元并发粒度并发传输至目标位置。

10 2. 根据权利要求 1 所述的方法，其特征在于，在将调取到的所述  $n$  个分表采用散列分配方式散列至不同的调度单元之前，所述方法还包括：

根据所述总并发粒度和所述单元并发粒度，确定所述调度单元的数量，并对所述每个调度单元分配对应的编号；

15 通过如下公式计算得到任意一个分库中的每个分表  $T_i$  的散列分配值，其中，所述散列分配值用于表征所述分表  $T_i$  所散列至对应的调度单元的编号  $T_{pos}$ ：

$T_{pos} = (T_{Count} + offset) \% tgCount$ ， $tgCount = totalChannel / tgChannel$ ，其中， $T_{Count}$  为所述任意一个分库中每个分表  $T_i$  的编号， $offset$  为每个分库向对应的调度单元分配的偏移量，初始值为 0， $totalChannel$  为所述总并发粒度， $tgChannel$  为所述单元并发粒度；

其中，所述  $n$  个分表分别按照计算得到的散列分配值散列至对应的调度单元。

20 3. 根据权利要求 2 所述的方法，其特征在于，在将所述每个调度单元中包含的分表按照所述单元并发粒度并发传输至目标位置的过程中，实时获取所述多个分库中除所述  $n$  个分表之外的  $m$  个分表，并按照所述总并发粒度将所述  $m$  个分表采用所述散列分配方式散列至所述不同的调度单元，其中，如果所述调度单元的线程被占满的情况下，将分配至对应的调度单元的分表放置在所述对应的调度单元的等待队列中，在所述调度  
25 单元的至少一个线程被释放之后，将所述等待队列中的分表调度至所述对应的调度单元中。

4. 根据权利要求 1 至 3 中任意一项所述的方法，其特征在于，在将所述每个调度单元中包含的分表按照所述单元并发粒度并发传输至目标位置之后，所述方法还包括：

30 在任意一个调度单元中存在空闲的线程的情况下，获取每个分库的当前并发数，其中，所述当前并发数用于表征分库中已经被调度至对应的调度单元中的分表数量；

根据所述每个分库的当前并发数，从所述任意一个调度单元的等待队列中调度对应数量的分表；

将从对应的等待队列中调度对应数量的分表并发传输至所述目标位置。

5. 根据权利要求 4 所述的方法，其特征在于，根据所述每个分库的当前并发数，  
5 从所述任意一个调度单元的等待队列中调度对应数量的分表，包括：

按照所述每个分库的当前并发数进行排序，确定属于不同分库的分表的调度优先级，其中，分库的当前并发数越低，所述当前并发数越低的分库中的分表的调度优先级越高；

根据所述任意一个调度单元中存在的空闲线程数量，确定调度数量；

- 10 在确定第一分库中的分表为所述调度优先级最高的分表之后，按照所述调度数量从所述等待队列中调度属于所述第一分库的分表。

6. 根据权利要求 5 所述的方法，其特征在于，在按照所述调度数量从所述等待队列中调度对应数量的分表之前，所述方法还包括：

读取所述等待队列中属于所述第一分库的分表的数量；

- 15 判断所述属于所述第一分库的分表的数量是否大于等于所述调度数量；

其中，如果大于等于所述调度数量，则进入按照所述调度数量从所述等待队列中调度对应数量的分表的步骤；

如果小于所述调度数量，则按照所述调度数量从所述等待队列中调度属于所述第一分库和属于其他分库的分表，其中，所述其他分库为当前并发数大于所述第一分库的  
20 库。

7. 根据权利要求 1 所述的方法，其特征在于，对散列至不同的所述调度单元的分表标记分库标识，所述分库标识用于表征所述分表原始对应的分库。

8. 一种基于分库分表的任务传输装置，其特征在于，包括：

抽取模块，用于从分库分表中抽取待传输的任务集合，其中，所述任务集合包括：

- 25 多个分库，以及每个分库所包含的分表；

调取模块，用于按照预先配置的总并发粒度从所述多个分库中共调取  $n$  个分表， $n$  等于所述总并发粒度；

处理模块，用于将调取到的所述  $n$  个分表采用散列分配方式散列至不同的调度单元，其中，预先配置每个调度单元的单元并发粒度相同；

- 30 并发模块，用于将所述每个调度单元中包含的分表按照所述单元并发粒度并发传输

至目标位置。

9. 根据权利要求 8 所述的装置, 其特征在于, 所述装置还包括:

第一确定模块, 用于根据所述总并发粒度和所述单元并发粒度, 确定所述调度单元的数量, 并对所述每个调度单元分配对应的编号;

5 计算模块, 用于通过如下公式计算得到任意一个分库中的每个分表  $T_i$  的散列分配值, 其中, 所述散列分配值用于表征所述分表  $T_i$  所散列至对应的调度单元的编号  $T_{pos}$ :

$T_{pos} = (TCount + offset) \% tgCount, tgCount = totalChannel / tgChannel$  其中,  $TCount$  为所述任意一个分库中每个分表  $T_i$  的编号,  $offset$  为每个分库向对应的调度单元分配的偏移量, 初始值为 0,  $totalChannel$  为所述总并发粒度,  $tgChannel$  为所述单元并发粒度;

10 其中, 所述处理模块包括: 第一散列分配模块, 用于所述  $n$  个分表分别按照计算得到的散列分配值散列至对应的调度单元。

10. 根据权利要求 9 所述的装置, 其特征在于, 所述装置还包括,

第一获取模块, 用于实时获取所述多个分库中除所述  $n$  个分表之外的  $m$  个分表;

15 第二散列分配模块, 用于按照所述总并发粒度将所述  $m$  个分表采用所述散列分配方式散列至所述不同的调度单元;

子处理模块, 用于如果所述调度单元的线程被占满的情况下, 将分配至对应的调度单元的分表放置在所述对应的调度单元的等待队列中, 在所述调度单元的至少一个线程被释放之后, 将所述等待队列中的分表调度至所述对应的调度单元中。

11. 根据权利要求 8 至 10 中任意一项所述的装置, 其特征在于, 所述装置还包括:

20 第二获取模块, 用于在任意一个调度单元中存在空闲的线程的情况下, 获取每个分库的当前并发数, 其中, 所述当前并发数用于表征分库中已经被调度至对应的调度单元中的分表数量;

调度模块, 用于根据所述每个分库的当前并发数, 从所述任意一个调度单元的等待队列中调度对应数量的分表;

25 传输模块, 用于将从对应的等待队列中调度对应数量的分表并发传输至所述目标位置。

12. 根据权利要求 11 所述的装置, 其特征在于, 所述调度模块包括:

排序模块, 用于按照所述每个分库的当前并发数进行排序, 确定属于不同分库的分表的调度优先级, 其中, 分库的当前并发数越低, 所述当前并发数越低的分库中的分表的调度优先级越高;

30

第二确定模块，用于根据所述任意一个调度单元中存在的空闲线程数量，确定调度数量；

子调度模块，用于在确定第一分库中的分表为所述调度优先级最高的分表之后，按照所述调度数量从所述等待队列中调度属于所述第一分库的分表。

5 13. 根据权利要求 12 所述的装置，其特征在于，所述装置还包括：

读取模块，用于读取所述等待队列中属于所述第一分库的分表的数量；

判断模块，用于判断所述属于所述第一分库的分表的数量是否大于等于所述调度数量；

第一执行模块，用于如果大于等于所述调度数量，则执行所述子调度模块的功能；

10 第二执行模块，用于如果小于所述调度数量，则按照所述调度数量从所述等待队列中调度属于所述第一分库和属于其他分库的分表，其中，所述其他分库为当前并发数大于所述第一分库的分库。

14. 一种基于分库分表的任务传输系统，其特征在于，包括：

源数据终端，用于存储分库分表；

15 调度终端，与所述源数据终端通信，用于从所述分库分表中抽取待传输的任务集合，其中，所述任务集合包括：多个分库，以及每个分库所包含的分表，并按照预先配置的总并发粒度从所述多个分库中共调取  $n$  个分表， $n$  等于所述总并发粒度，在将调取到的所述  $n$  个分表采用散列分配方式散列至不同的调度单元之后，将每个调度单元中包含的分表按照单元并发粒度并发传输，其中，预先配置所述每个调度单元的所述单元并发粒  
20 度相同；

目标终端，与所述调度终端通信，用于接收所述调度终端并发传输来的所述任务集合。

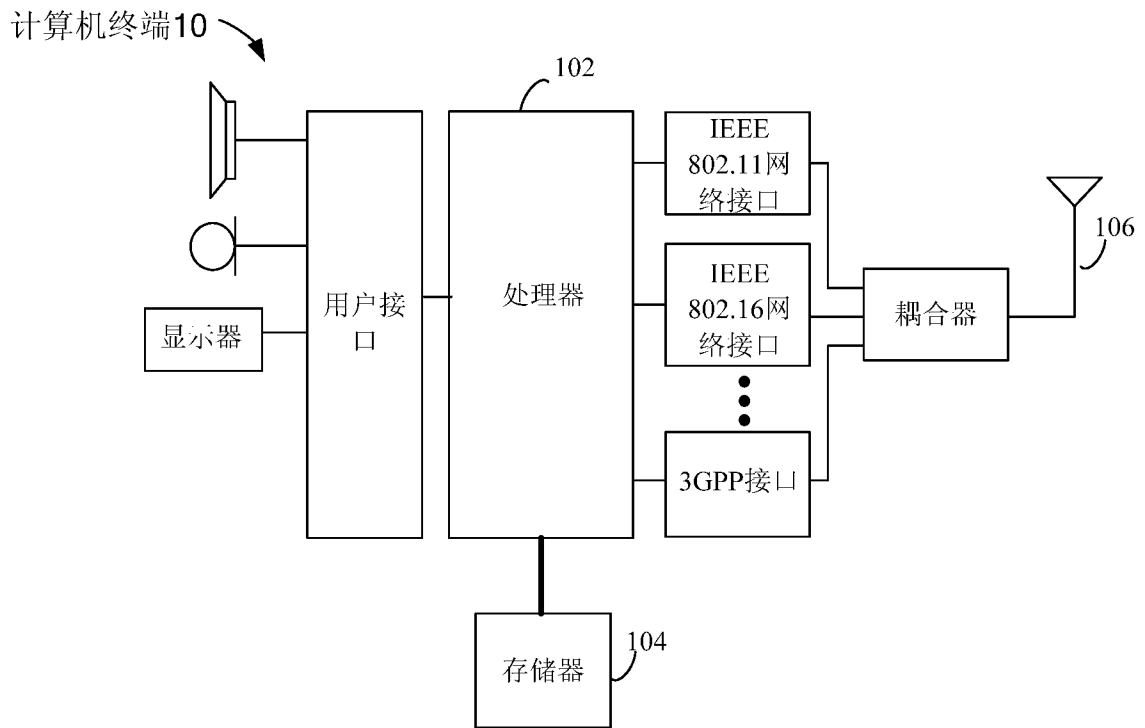


图 1

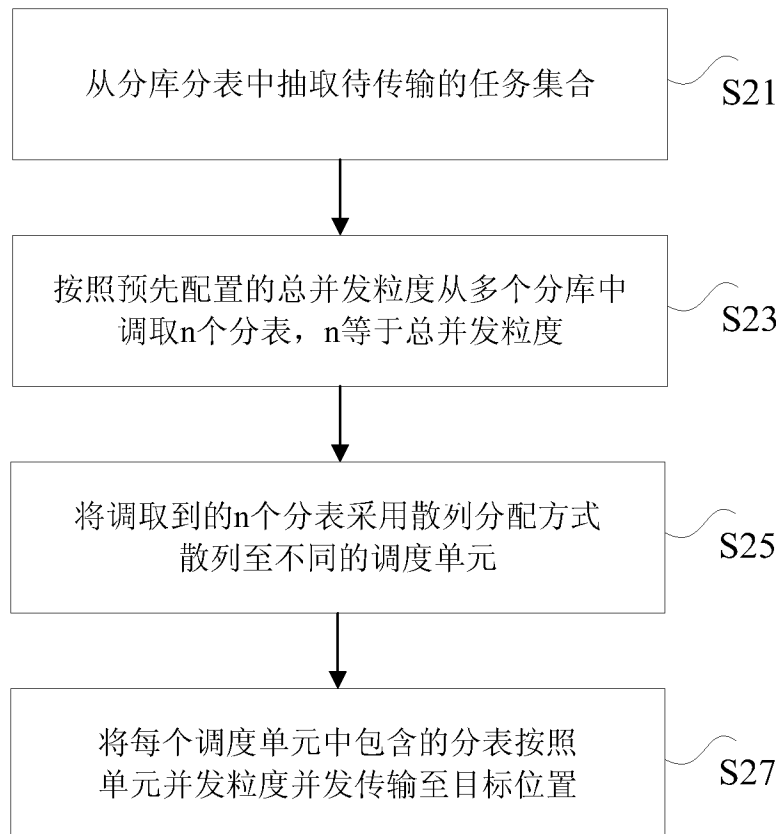


图 2

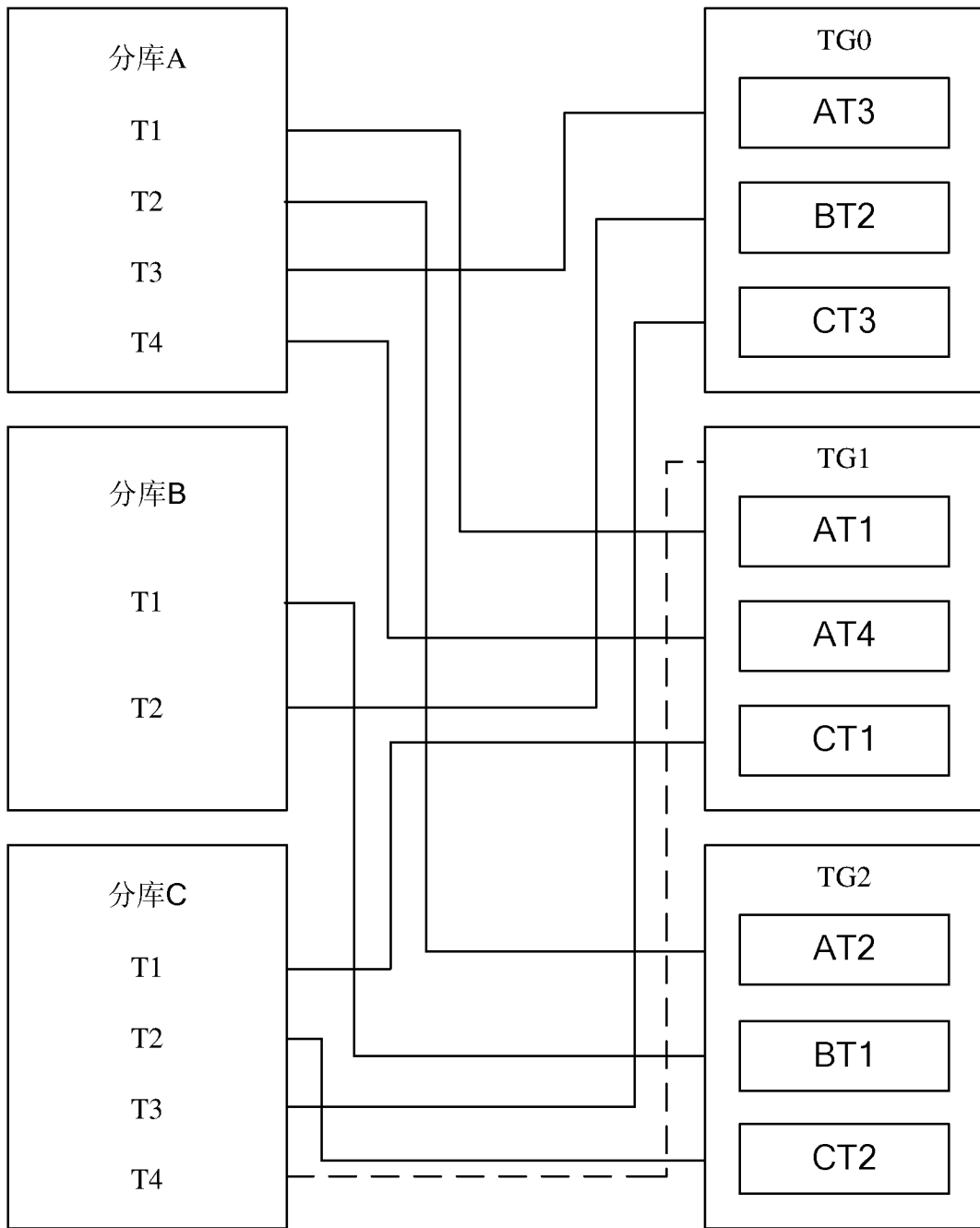


图 3

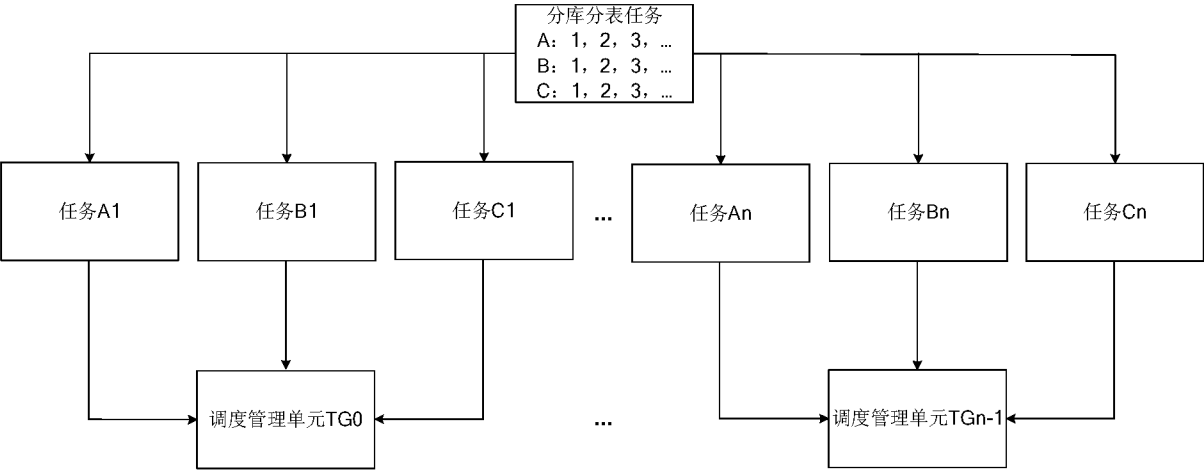


图 4

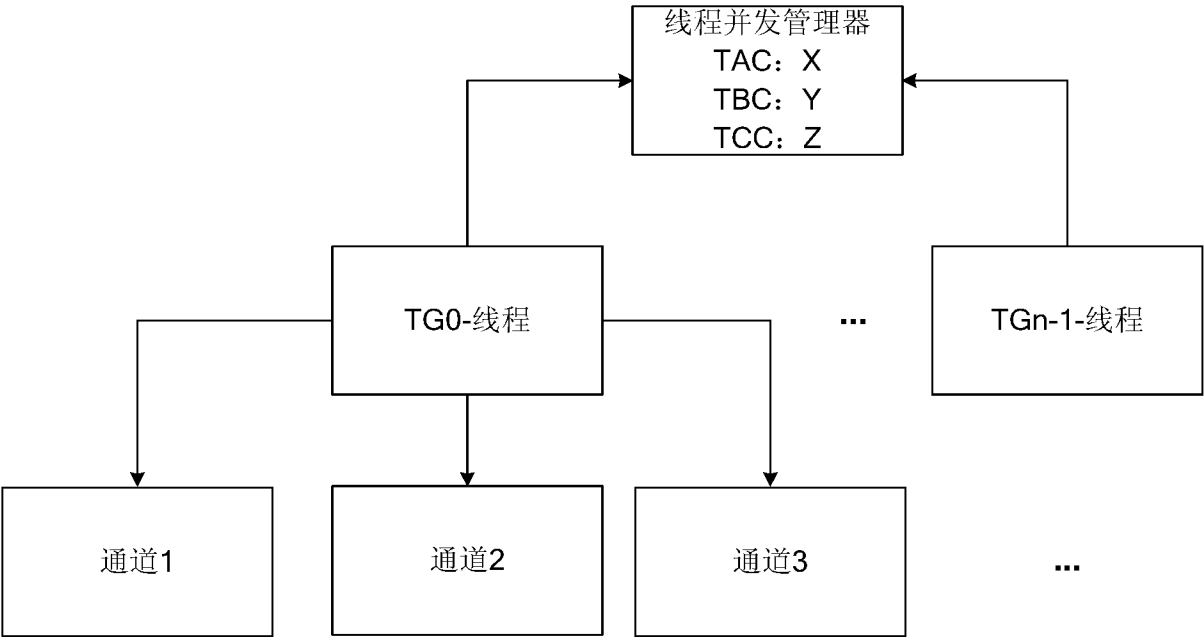


图 5

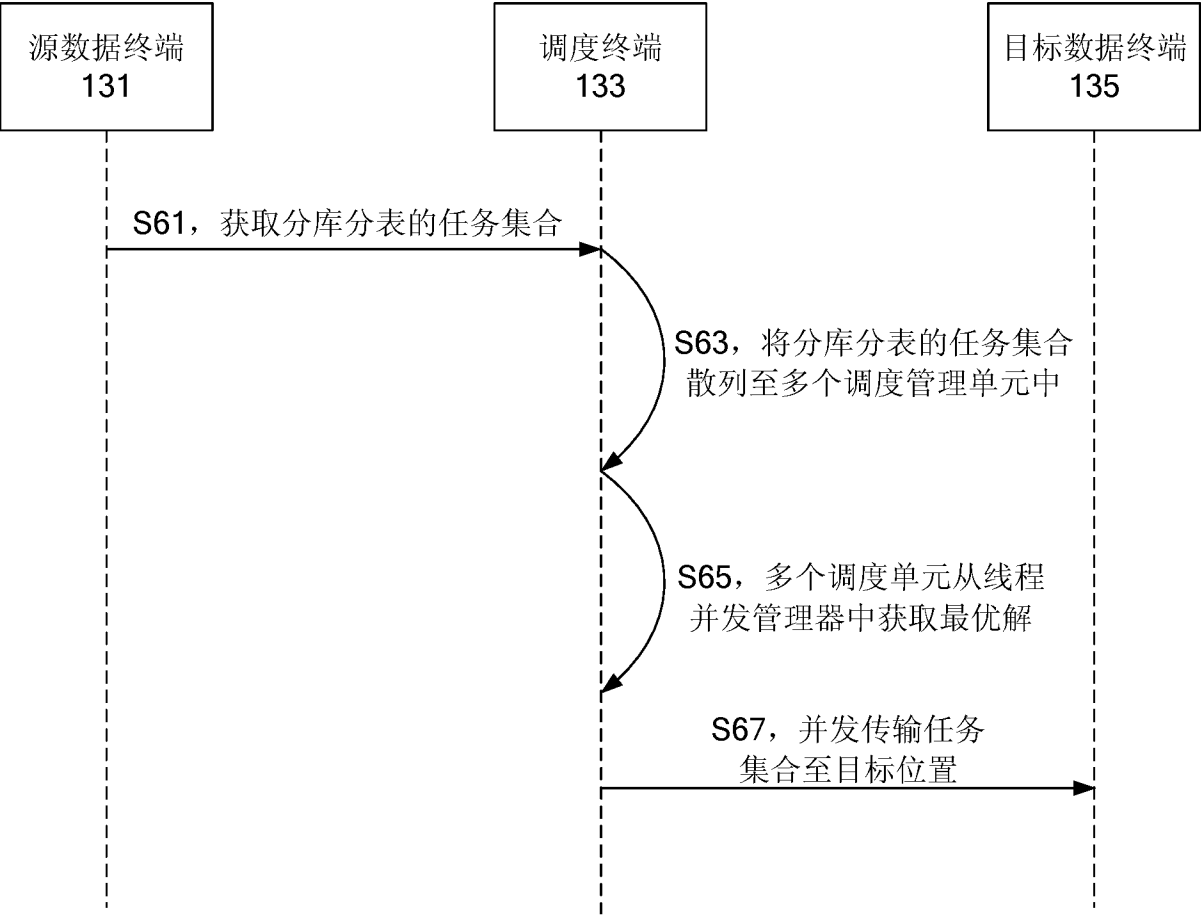


图 6



图 7

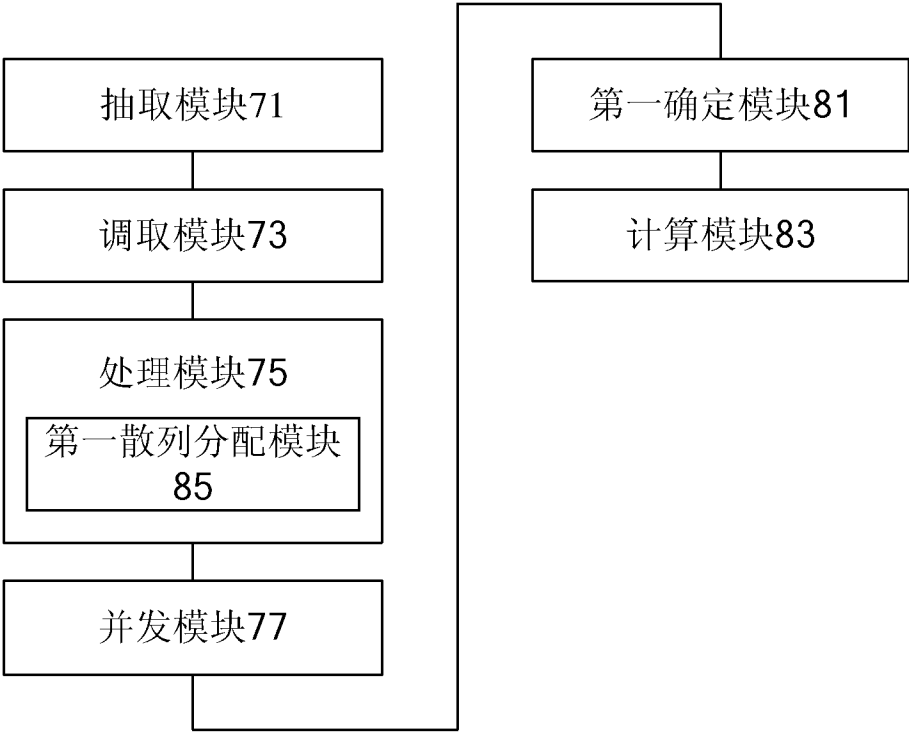


图 8



图 9

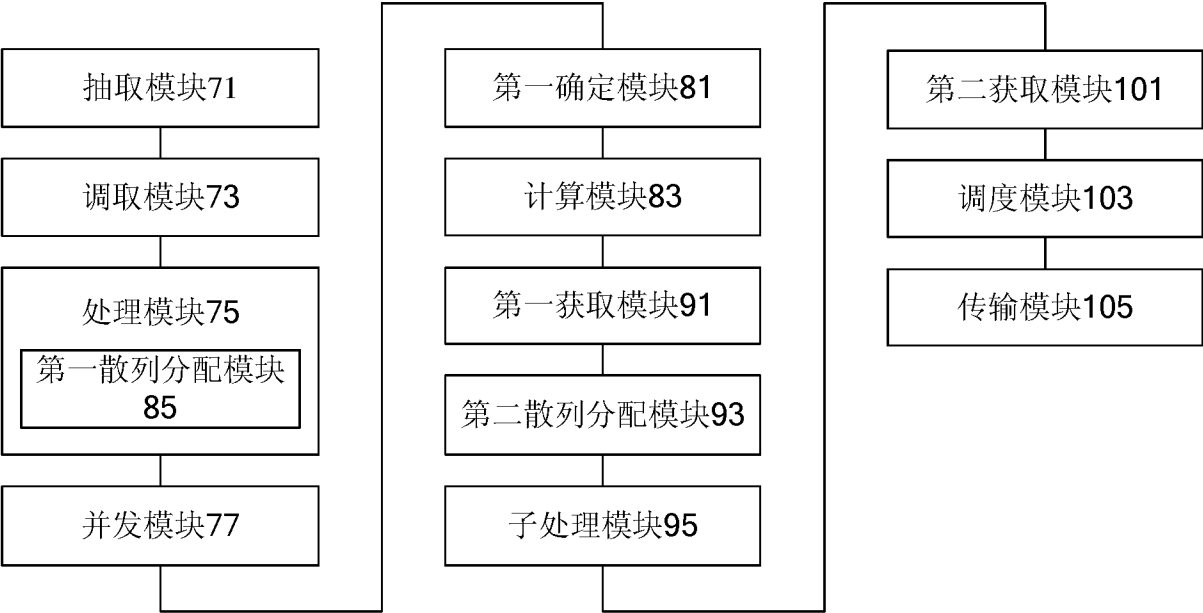


图 10

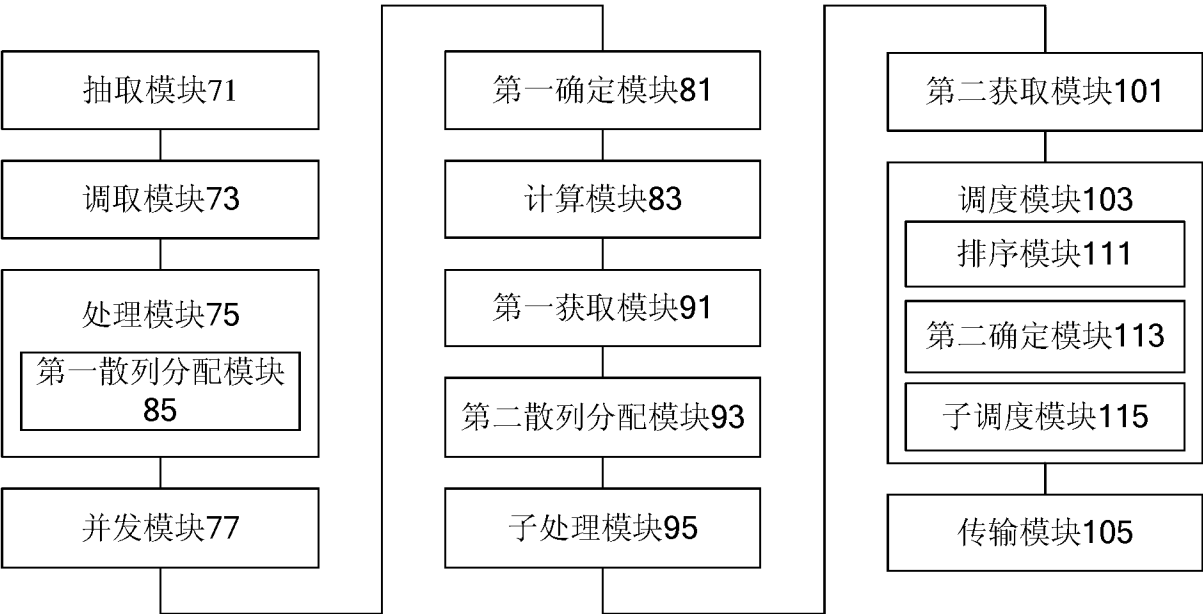


图 11

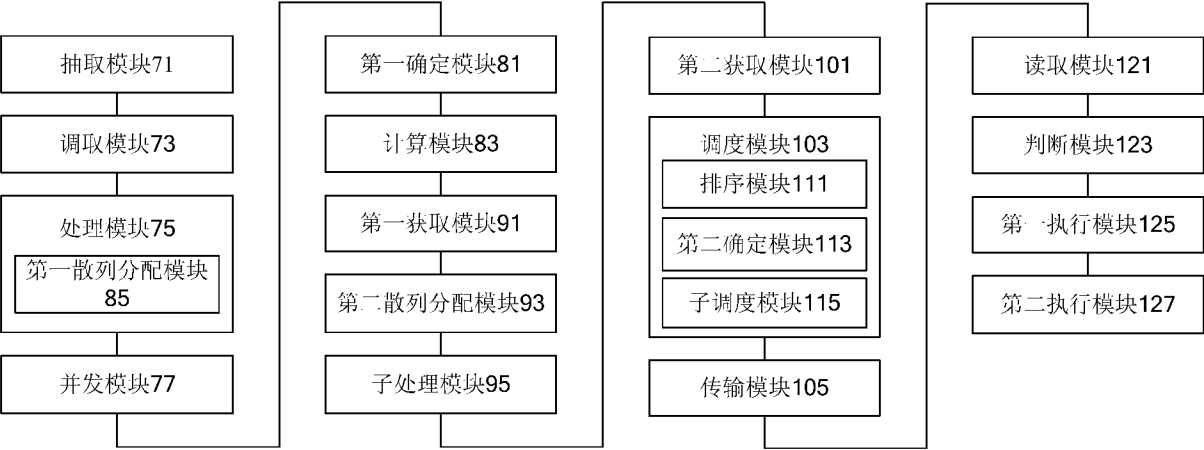


图 12



图 13

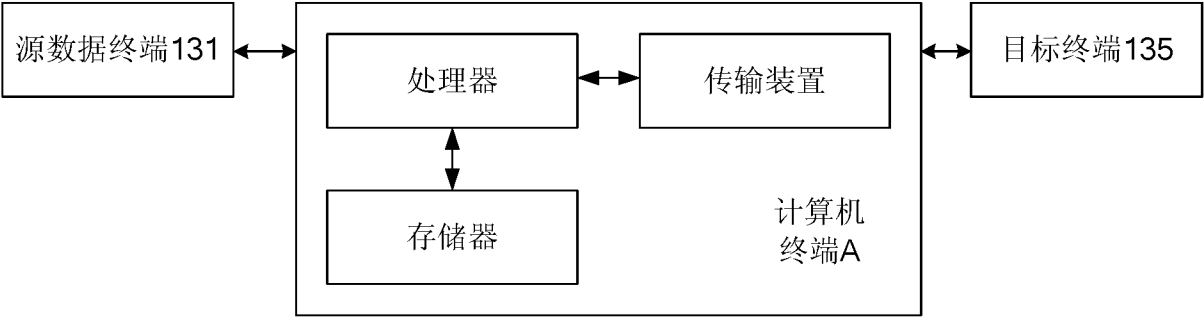


图 14

# INTERNATIONAL SEARCH REPORT

International application No.

**PCT/CN2016/107409****A. CLASSIFICATION OF SUBJECT MATTER**

G06F 17/30 (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 17/-

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)

DWPI, CNABS, SIPOABS, CPRSABS, CNKI: sub-library, sub-list, transmission, transport, sharding, database, concurrence, data, query, transform

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

| Category* | Citation of document, with indication, where appropriate, of the relevant passages             | Relevant to claim No. |
|-----------|------------------------------------------------------------------------------------------------|-----------------------|
| A         | CN 103942209 A (ALIBABA GROUP HOLDING LTD.), 23 July 2014 (23.07.2014), the whole document     | 1-14                  |
| A         | CN 104317749 A (XIAOMI TECHNOLOGY CO., LTD.), 28 January 2015 (28.01.2015), the whole document | 1-14                  |
| A         | CN 103678408 A (ALIBABA GROUP HOLDING LTD.), 26 March 2014 (26.03.2014), the whole document    | 1-14                  |

☐ Further documents are listed in the continuation of Box C.☒ See patent family annex.

|                                                                                                                                                                         |                                                                                                                                                                                                                                                  |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| * Special categories of cited documents:                                                                                                                                | “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                                              |
| “A” document defining the general state of the art which is not considered to be of particular relevance                                                                | “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                                                                     |
| “E” earlier application or patent but published on or after the international filing date                                                                               | “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 |
| “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) | “&” document member of the same patent family                                                                                                                                                                                                    |
| “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                                                                  |                                                                                                                                                                                                                                                  |

|                                                                                                                                                                                                                    |                                                                                         |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------|
| Date of the actual completion of the international search<br>28 February 2017 (28.02.2017)                                                                                                                         | Date of mailing of the international search report<br><b>06 March 2017 (06.03.2017)</b> |
| Name and mailing address of the ISA/CN:<br>State Intellectual Property Office of the P. R. China<br>No. 6, Xitucheng Road, Jimenqiao<br>Haidian District, Beijing 100088, China<br>Facsimile No.: (86-10) 62019451 | Authorized officer<br><b>TIAN, Jing</b><br>Telephone No.: (86-10) <b>62089417</b>       |

## INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/CN2016/107409

| Patent Documents referred in the Report | Publication Date | Patent Family | Publication Date  |
|-----------------------------------------|------------------|---------------|-------------------|
| CN 103942209 A                          | 23 July 2014     | None          |                   |
| CN 104317749 A                          | 28 January 2015  | None          |                   |
| CN 103678408 A                          | 26 March 2014    | HK 1193484 A0 | 19 September 2014 |

| <p>A. 主题的分类</p> <p>G06F 17/30 (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 17/-</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>DWPI, CNABS, SIPOABS, CPRSABS, CNKI: 分库, 分表, 数据库, 并发, 数据, 查询, 传输, 传送, sharding, database, concurrence, data, query, transform</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>A</td> <td>CN 103942209 A (阿里巴巴集团控股有限公司) 2014年 7月 23日 (2014 - 07 - 23)<br/>全文</td> <td>1-14</td> </tr> <tr> <td>A</td> <td>CN 104317749 A (小米科技有限责任公司) 2015年 1月 28日 (2015 - 01 - 28)<br/>全文</td> <td>1-14</td> </tr> <tr> <td>A</td> <td>CN 103678408 A (阿里巴巴集团控股有限公司) 2014年 3月 26日 (2014 - 03 - 26)<br/>全文</td> <td>1-14</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 | A | CN 103942209 A (阿里巴巴集团控股有限公司) 2014年 7月 23日 (2014 - 07 - 23)<br>全文 | 1-14 | A | CN 104317749 A (小米科技有限责任公司) 2015年 1月 28日 (2015 - 01 - 28)<br>全文 | 1-14 | A | CN 103678408 A (阿里巴巴集团控股有限公司) 2014年 3月 26日 (2014 - 03 - 26)<br>全文 | 1-14 |
| 类 型*                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 引用文件, 必要时, 指明相关段落                                                                                                                                                                                          | 相关的权利要求                                            |                                                                                                                                                                                                                                       |                                                                                                                                                                                                            |         |   |                                                                   |      |   |                                                                 |      |   |                                                                   |      |
| A                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | CN 103942209 A (阿里巴巴集团控股有限公司) 2014年 7月 23日 (2014 - 07 - 23)<br>全文                                                                                                                                          | 1-14                                               |                                                                                                                                                                                                                                       |                                                                                                                                                                                                            |         |   |                                                                   |      |   |                                                                 |      |   |                                                                   |      |
| A                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | CN 104317749 A (小米科技有限责任公司) 2015年 1月 28日 (2015 - 01 - 28)<br>全文                                                                                                                                            | 1-14                                               |                                                                                                                                                                                                                                       |                                                                                                                                                                                                            |         |   |                                                                   |      |   |                                                                 |      |   |                                                                   |      |
| A                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | CN 103678408 A (阿里巴巴集团控股有限公司) 2014年 3月 26日 (2014 - 03 - 26)<br>全文                                                                                                                                          | 1-14                                               |                                                                                                                                                                                                                                       |                                                                                                                                                                                                            |         |   |                                                                   |      |   |                                                                 |      |   |                                                                   |      |
| <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> <table border="0"> <tr> <td> <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> </td> <td> <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> </td> </tr> </table> |                                                                                                                                                                                                            |                                                    | <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>* 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 2月 28日</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |                                                                                                                                                                                                            | <p>国际检索报告邮寄日期</p> <p>2017年 3月 6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)</p> <p>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>受权官员</p> <p>田竞</p> <p>电话号码 (86-10) 62089417</p> |                                                                                                                                                                                                                                       |                                                                                                                                                                                                            |         |   |                                                                   |      |   |                                                                 |      |   |                                                                   |      |

国际检索报告  
关于同族专利的信息

国际申请号

PCT/CN2016/107409

| 检索报告引用的专利文件 |           |   | 公布日<br>(年/月/日) | 同族专利 |         |    | 公布日<br>(年/月/日) |
|-------------|-----------|---|----------------|------|---------|----|----------------|
| CN          | 103942209 | A | 2014年 7月 23日   | 无    |         |    |                |
| CN          | 104317749 | A | 2015年 1月 28日   | 无    |         |    |                |
| CN          | 103678408 | A | 2014年 3月 26日   | HK   | 1193484 | A0 | 2014年 9月 19日   |