

(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
2019 年 11 月 21 日 (21.11.2019)



(10) 国际公布号
WO 2019/219005 A1

(51) 国际专利分类号:
G06F 16/00 (2019.01)

(21) 国际申请号: PCT/CN2019/086847

(22) 国际申请日: 2019 年 5 月 14 日 (14.05.2019)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201810470200.1 2018 年 5 月 16 日 (16.05.2018) CN

(71) 申请人: 杭州海康威视数字技术股份有限公司 (HANGZHOU HIKVISION DIGITAL TECHNOLOGY CO., LTD.) [CN/CN]; 中国浙江省杭州市滨江区阡陌路 555 号, Zhejiang 310051 (CN)。

(72) 发明人: 陈国栋(CHEN, Guodong); 中国浙江省杭州市滨江区阡陌路 555 号, Zhejiang 310051 (CN)。
管国辰(GUAN, Guochen); 中国浙江省杭州市滨江区阡陌路 555 号, Zhejiang 310051 (CN)。
邓长春(DENG, Changchun); 中国浙江省杭州市滨江区阡陌路 555 号, Zhejiang 310051 (CN)。

(74) 代理人: 北京柏杉松知识产权代理事务所(普通合伙)(PATENTSINO IP FIRM); 中国北京市西城区北三环中路 27 号 商 房 大 厦 413 室, Beijing 100029 (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, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK,

(54) Title: DATA PROCESSING SYSTEM AND METHOD

(54) 发明名称: 一种数据处理系统及方法

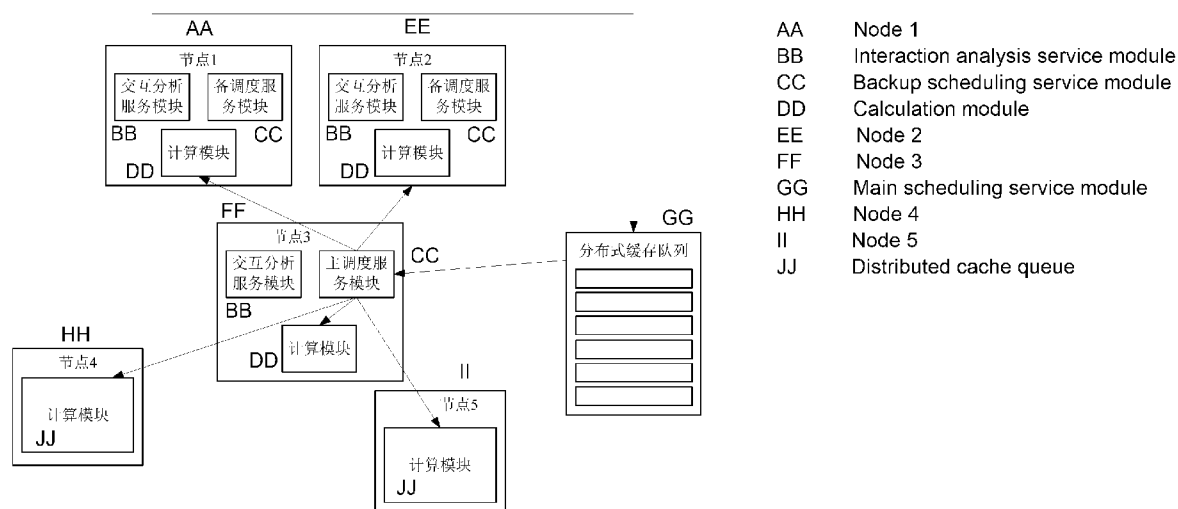


图 1

(57) Abstract: Provided by the embodiments of the present application are a data processing system and method, the data processing system comprising: an interaction analysis service module, a scheduling service module and a calculation module, wherein the interaction analysis service module receives a data request sent by each client, processes processing logic comprised in the received data requests, obtains a task set, and caches the obtained task set to a task set queue; a main scheduling service module obtains the task set from the task set queue and sends the obtained task set to the calculation module; the calculation module carries out calculations for the received task set, obtains a calculation result, and sends the obtained calculation result to the main scheduling service module; the main scheduling service module encapsulates the received calculation result, obtains a result set, and caches the obtained result set into a result set queue; and the interaction analysis service module obtains the result set from the result set queue, and sends the

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))。

obtained result set to a corresponding client. The concurrency of data interaction analysis processing may be improved by applying the embodiments of the present application.

(57) 摘要: 本申请实施例提供了一种数据处理系统及方法, 数据处理系统包括: 交互分析服务模块、调度服务模块和计算模块; 交互分析服务模块接收各个客户端发送的数据请求, 对接收的数据请求包括的处理逻辑进行处理, 得到任务集, 将得到的任务集缓存至任务集队列; 主调度服务模块从任务集队列中获取任务集, 将获取的任务集发送至计算模块; 计算模块对接收的任务集进行计算, 得到计算结果, 将得到的计算结果发送给主调度服务模块; 主调度服务模块对接收的计算结果进行封装, 得到结果集, 将得到的结果集缓存至结果集队列; 交互分析服务模块从结果集队列中获取结果集, 将获取的结果集发送给对应的客户端。应用本申请实施例, 能够提高数据交互分析处理的并发性。

一种数据处理系统及方法

本申请要求于2018年5月16日提交中国专利局、申请号为201810470200.1发明名称为“一种数据处理系统及方法”的中国专利申请的优先权，其全部内容
5 通过引用结合在本申请中。

技术领域

本申请涉及大数据处理技术领域，特别是涉及一种数据处理系统及方法。

背景技术

传统数据处理系统采用分布式架构，并且，传统数据处理系统屏蔽掉后
10 端复杂的数据处理实现和千差万别的接口，为用户提供统一标准 SQL
（Structured Query Language，结构化查询语言）接口。用户通过标准 SQL 接口来操作数据处理系统。这给用户带来了良好的用户体验。

在大数据处理领域，分布式架构诸如批处理系统 MapReduce、内存计算引擎 Spark、分布式数据库 Hbase、以及流式处理框架 Storm 等都有着自己的
15 编程模型。基于此，分布式架构的使用，对于用户来说，学习成本高，且业务无法平滑迁移。因而，近年来开始构建与传统数据处理系统类似的交互式数据处理系统。诸如 SQL On Hadoop 系统中的 Hive、Impala 和 Phoenix，以及 Spark 生态的 Spark SQL 等架构就是该类数据处理系统。在交互式数据处理系统中，数据以表的形式存在。交互式数据处理系统提供标准的 SQL 接口，
20 能够达到高并发、低延迟的线上数据分析效果。

然而，上述交互式数据处理系统在实际应用场景中效果不够理想，其中并发支持度是瓶颈之一。原因是当前交互式数据处理系统的SQL解析、分析、优化、计划生成及资源评估等阶段的交互分析服务都在单个节点上进行，这制约了高并发特性。因此需要一种新型数据处理系统解决该瓶颈，实现高并
25 发的交互分析处理需求。

发明内容

本申请实施例的目的在于提供一种数据处理系统及方法，以提高数据交互分析处理的并发性。具体技术方案如下：

-2-

为实现上述目的，本申请实施例提供了一种数据处理方法，应用于数据处理系统，所述数据处理系统包括：多个交互分析服务模块、多个调度服务模块和多个计算模块；所述方法包括：

5 所述多个交互分析服务模块接收各个客户端发送的数据请求；对接收的数据请求包括的处理逻辑进行处理，得到任务集；将得到的任务集缓存至任务集队列；

所述多个调度服务模块中的主调度服务模块从所述任务集队列中获取任务集；将获取的任务集发送至所述多个计算模块；

10 所述多个计算模块对接收的任务集进行计算，得到计算结果；将得到的计算结果发送给所述主调度服务模块；

所述主调度服务模块对接收的计算结果进行封装，得到结果集；将得到的结果集缓存至结果集队列；

所述多个交互分析服务模块从所述结果集队列中获取结果集；将获取的结果集发送给对应的客户端。

15 可选的，所述多个交互分析服务模块对接收的数据请求包括的处理逻辑进行处理，得到任务集，包括：

所述多个交互分析服务模块对接收的数据请求包括的处理逻辑进行 SQL 解析与优化、计划生成和资源预估处理，得到处理结果；对得到的处理结果进行序列化和封装，得到任务集。

20 可选的，所述多个交互分析服务模块将得到的任务集缓存至任务集队列，包括：

所述多个交互分析服务模块获取接收的数据请求所包括的处理逻辑的属性信息，将得到的任务集与获取的属性信息对应的缓存至任务集队列。

可选的，所述属性信息包括优先级；

25 所述主调度服务模块从所述任务集队列中获取任务集，包括：

所述主调度服务模块按照优先级从高到低的顺序，遍历并获取所述任务

—3—

集队列中的各个任务集。

可选的，所述主调度服务模块按照优先级从高到低的顺序，遍历并获取所述任务集队列中的各个任务集，包括：

5 所述主调度服务模块按照优先级从高到低的顺序，遍历所述任务集队列中的各个任务集，并判断任务集所需的计算资源是否超过所述多个计算模块中当前剩余的计算资源；若否，则从所述任务集队列中获取该任务集；若是，则将该任务集缓存至所述任务集队列中与该任务集的优先级相同的所有任务集之后。

可选的，所述主调度服务模块从所述任务集队列中获取任务集，包括：

10 所述主调度服务模块遍历所述任务集队列中的各个任务集，获取未被标记为已调度的任务集；

在所述主调度服务模块获取未被标记为已调度的任务集之后，还包括：

所述主调度服务模块将获取的任务集标记为已调度。

可选的，所述主调度服务模块从所述任务集队列中获取任务集，包括：

15 所述主调度服务模块遍历所述任务集队列中的各个任务集，获取未被标记为终止的任务集；

所述方法还包括：

20 所述多个交互分析服务模块若在将得到的任务集缓存至任务集队列的第一预设时长后，检测到所述结果集队列中不存在得到的任务集所对应的结果集，则将得到的任务集标记为终止。

可选的，所述数据处理系统还包括：删除模块；所述方法还包括：

所述删除模块检测所述结果集队列中是否存在缓存时长达到第二预设时长的第一结果集；若存在，删除所述第一结果集；或者，

25 所述删除模块检测所述结果集队列中是否存在最后一次被访问至当前时刻的时长达到第三预设时长的第二结果集；若存在，删除所述第二结果集。

可选的，所述多个交互分析服务模块对接收的数据请求包括的处理逻辑

-4-

进行处理，得到任务集，包括：

所述多个交互分析服务模块检测结果集队列中是否存在接收的数据请求所包括的处理逻辑对应的结果集；若不存在，则对接收的数据请求包括的处理逻辑进行数据处理，得到任务集。

5 可选的，所述方法还包括：

所述多个交互分析服务模块若确定所述结果集队列中存在接收的数据请求所包括的处理逻辑对应的结果集，则从所述结果集队列中获取接收的数据请求所包括的处理逻辑对应的结果集；将获取的结果集发送给对应的客户端。

可选的，所述交互分析服务模块和所述调度服务模块位于同一进程中。

10 为实现上述目的，本申请实施例还提供了一种数据处理系统，所述数据处理系统包括：多个交互分析服务模块、多个调度服务模块和多个计算模块；

所述多个交互分析服务模块，用于接收各个客户端发送的数据请求；对接收的数据请求包括的处理逻辑进行处理，得到任务集；将得到的任务集缓存至任务集队列；

15 所述多个调度服务模块中的主调度服务模块，用于从所述任务集队列中获取任务集；将获取的任务集发送至所述多个计算模块；

所述多个计算模块，用于对接收的任务集进行计算，得到计算结果；将得到的计算结果发送给所述主调度服务模块；

20 所述主调度服务模块，用于接收的计算结果进行封装，得到结果集；将得到的结果集缓存至结果集队列；

所述多个交互分析服务模块，用于从所述结果集队列中获取结果集；将获取的结果集发送给对应的客户端。

可选的，所述多个交互分析服务模块具体用于：

25 对接收的数据请求包括的处理逻辑进行 SQL 解析与优化、计划生成和资源预估处理，得到处理结果；对得到的处理结果进行序列化和封装，得到任务集。

—5—

可选的，所述多个交互分析服务模块具体用于：获取接收的数据请求所包括的处理逻辑的属性信息，将得到的任务集与获取的属性信息对应的缓存至任务集队列。

可选的，所述属性信息包括优先级；

- 5 所述主调度服务模块具体用于：按照优先级从高到低的顺序，遍历并获取所述任务集队列中的各个任务集。

可选的，所述主调度服务模块具体用于：

- 10 按照优先级从高到低的顺序，遍历所述任务集队列中的各个任务集，并判断任务集所需的计算资源是否超过所述多个计算模块中当前剩余的计
源；若否，则从所述任务集队列中获取该任务集；若是，则将该任务集缓存
至所述任务集队列中与该任务集的优先级相同的所有任务集之后。

可选的，所述主调度服务模块具体用于：遍历所述任务集队列中的各个任务集，获取未被标记为已调度的任务集；

所述主调度服务模块，还用于将获取的任务集标记为已调度。

- 15 可选的，所述主调度服务模块具体用于：遍历所述任务集队列中的各个任务集，获取未被标记为终止的任务集；

所述多个交互分析服务模块，还用于若在将得到的任务集缓存至任务集队列的第一预设时长后，检测到所述结果集队列中不存在得到的任务集所对应的结果集，则将得到的任务集标记为终止。

- 20 可选的，所述数据处理系统还包括：删除模块；

所述删除模块，用于检测所述结果集队列中是否存在缓存时长达到第二预设时长的第一结果集；若存在，删除所述第一结果集；或者，

- 25 所述删除模块，用于检测所述结果集队列中是否存在最后一次被访问至当前时刻的时长达到第三预设时长的第二结果集；若存在，删除所述第二结果集。

可选的，所述多个交互分析服务模块具体用于：

—6—

检测结果集队列中是否存在接收的数据请求所包括的处理逻辑对应的结果集；若不存在，则对接收的数据请求包括的处理逻辑进行数据处理，得到任务集。

5 可选的，所述多个交互分析服务模块，还用于若确定所述结果集队列中存在接收的数据请求所包括的处理逻辑对应的结果集，则从所述结果集队列中获取接收的数据请求所包括的处理逻辑对应的结果集；将获取的结果集发送给对应的客户端。

可选的，所述交互分析服务模块和所述调度服务模块位于同一进程中。

10 为实现上述目的，本申请实施例还提供了一种机器可读存储介质，所述机器可读存储介质内存储有计算机程序，所述计算机程序被处理器执行时实现上述任一数据处理方法步骤。

15 本申请实施例中，在数据处理系统中包括了任务集队列和结果集队列，交互分析服务模块将任务集缓存至任务集队列，调度服务模块从任务集队列中获取任务集。另外，调度服务模块将结果集缓存至结果集队列，交互分析服务模块从结果集队列中获取结果集，并将结果集发送给客户端。可见，通过任务集队列和结果集队列，实现了交互分析服务模块和调度服务模块的解耦，数据交互分析处理可以在多个节点上进行，提高了数据交互分析处理的并发性。当然，实施本申请的任一产品或方法必不一定需要同时达到以上所述的所有优点。

20 附图说明

为了更清楚地说明本申请实施例或相关技术中的技术方案，下面将对实施例或相关技术描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本申请的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获得其他的附图。

25 图 1 为本申请实施例提供的数据处理系统的一种结构示意图；

图 2 为本申请实施例提供的交互分析服务模块的请求处理流程的一种示意图；

图 3 为本申请实施例提供的调度服务模块的任务调度处理流程的一种示

意图；

图 4 为本申请实施例提供的数据处理方法的一种流程示意图。

具体实施方式

下面将结合本申请实施例中的附图，对本申请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申请一部分实施例，而不是全部的实施例。基于本申请中的实施例，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，都属于本申请保护的范围。

目前，数据处理系统提供的调度服务和交互分析服务之间为耦合关系。具体的，调度服务需要调度交互分析服务生成的任务集。基于此，SQL 解析、分析、优化、计划生成及资源评估等阶段的交互分析服务都在单个节点上，这制约了高并发特性。为解决上述问题，本申请实施例提供了一种数据处理系统。

该数据处理系统包括：多个交互分析服务模块、多个调度服务模块和多个计算模块。交互分析服务模块将任务集缓存至任务集队列，调度服务模块从任务集队列中获取任务集，并发送给计算模块进行计算，得到结果集。另外，调度服务模块将结果集缓存至结果集队列，交互分析服务模块从结果集队列中获取结果集，并将结果集发送给客户端。

可见，本申请实施例中，通过任务集队列和结果集队列，实现了交互分析服务模块和调度服务模块的解耦，多个交互分析服务模块可部署在多个节点上同时提供交互分析服务，即多个交互分析服务模块可同时提供数据交互分析处理，进而数据交互分析处理可以在多个节点上进行，提高了数据交互分析处理的并发性。

下面通过具体实施例，对本申请实施例提供的数据处理系统进行说明。

参考图 1，图 1 为本申请实施例提供的数据处理系统的一种结构示意图。该数据处理系统包括：多个交互分析服务模块 101-103、多个调度服务模块 201-203 和多个计算模块 301-305。这里，仅以 3 个交互分析服务模块、3 个调度服务模块和 5 个计算模块为例进行说明，本申请实施例并不限定交互分析服务模块、调度服务模块和计算模块的个数。

—8—

本申请实施例中，交互分析服务模块和调度服务模块可位于同一进程中，即交互分析服务模块和调度服务模块可运行于同一进程，为同一进程中的两个模块。交互分析服务模块之间关系对等，即多个交互分析服务模块可以同时接收并处理客户端发送的数据请求。调度服务模块是主备式结构，只有一个主调度服务模块提供服务，其它为备用调度服务模块，在主调度服务模块挂掉后，从备用调度服务模块中选举出一个主调度服务模块。这样可以避免

5 出现数据不同步的问题。

本申请实施例中，数据处理系统可依赖 zookeeper 等分布式协同组件技术，来实现主调度服务模块的选举和主备调度服务模块切换的功能，以及交互分

10 析服务模块的负载均衡功能。数据请求又可以称为数据处理请求。

交互分析服务模块为数据处理系统的数据请求处理及任务集产生层，调度服务模块为任务集的消费调度执行层。交互分析服务模块和调度服务模块之间通过分布式缓存队列进行解耦交互。

计算模块可单独位于一个进程中。多个计算模块隶属分布式计算框架，例如 Spark。计算模块主要完成调度服务模块分发的计算任务。交互分析服务

15 模块和调度服务模块与计算模块可以位于同一节点上，也可以位于不同的节点上。

本申请实施例中，数据处理系统中包括任务集队列和结果集队列。任务集队列和结果集队列均为分布式缓存队列，可通过分布式缓存技术实现，以

20 使用数据处理系统的全部内存资源。

多个交互分析服务模块 101-103 接收各个客户端发送的数据请求。

多个交互分析服务模块可同时接收并处理相同或不同的数据请求。例如，在某一时刻，交互分析服务模块 101 接收并处理数据请求 1，交互分析服务模块 102 接收并处理数据请求 2，交互分析服务模块 103 接收并处理数据请求 3。

25 多个交互分析服务模块 101-103 对接收的数据请求包括的处理逻辑进行处理，得到任务集。

数据请求中包括处理逻辑，处理逻辑为数据处理的逻辑关系。交互分析服务模块对数据请求进行处理，即为交互分析服务模块对数据请求包括的处

理逻辑进行处理。

可选的，交互分析服务模块对处理逻辑的处理包括：SQL 解析与优化、计划生成和资源预估等。其中，计划生成是根据 SQL 解析与优化的处理结果，生成逻辑计划和物理计划。资源预估为估算完成逻辑计划和物理计划所需的计算资源。

如图 2 所示的交互分析服务模块请求处理流程，包括如下步骤。

步骤 201，交互分析服务模块对接收的数据请求包括的处理逻辑进行 SQL 解析与优化处理。

步骤 202，交互分析服务模块对接收的数据请求包括的处理逻辑进行计划生成处理。

步骤 203，交互分析服务模块对接收的数据请求包括的处理逻辑进行资源预估处理。

步骤 204，交互分析服务模块得到处理结果，对得到的处理结果进行序列化和封装，得到任务集。

交互分析服务模块对数据请求包括的处理逻辑进行 SQL 解析与优化、计划生成以及资源预估处理后，得到处理结果为一个或多个任务。交互分析服务模块对一个或多个任务进行序列化和封装，得到任务集。

步骤 205，交互分析服务模块将任务集加入任务集队列。

多个交互分析服务模块 101-103 将得到的任务集缓存至任务集队列。

可选的，交互分析服务模块在将任务集缓存至任务集队列时，可以先获取数据请求所包括的处理逻辑的属性信息，将获取的属性信息与数据请求的任务集对应的缓存至任务集队列。

这里，属性信息可以为用户根据需要预先配置的。属性信息为主调度服务模块调度任务集的依据。例如，属性信息可包括优先级，任务集对应的优先级越高，主调度服务模块越优先调度该任务集。

以属性信息包括优先级为例，为了便于主调度服务模块调度任务集，任

务集队列可以按照优先级划分为多个子队列，每一优先级对应一个子队列。此时，交互分析服务模块将获取的属性信息与数据请求的任务集对应的缓存至任务集队列，可以为：将数据请求的任务集缓存至获取的优先级对应的子队列。例如，数据请求 1 的优先级为 1，数据请求 1 的任务集为任务集 1。交互分析服务模块 101 得到优先级 1 的任务集 1，将任务集 1 缓存至优先级 1 对应的子队列的尾部。

10 可选的，交互分析服务模块将任务集缓存至任务集队列后，定时检测结果集队列中是否存在该任务集所对应的结果集。若检测到结果集队列中存在该任务集所对应的结果集，则交互分析服务模块获取该任务集所对应的结果集，并将获取的结果集发送给客户端。

若将任务集缓存至任务集队列的第一预设时长后，交互分析服务模块检测结果集队列中还是不存在该任务集所对应的结果集，则将该任务集标记为终止，以告知主调度服务模块停止对该任务集的调度，节约计算资源。

15 多个调度服务模块 201-203 中的主调度服务模块从任务集队列中获取任务集，将获取的任务集发送至多个计算模块 301-305。

调度服务模块将任务集从任务集队列中获取并发送给计算模块的过程，即为一次任务集的调度。

20 可选的，任务集队列中缓存有优先级与任务集的对应关系。主调度服务模块按照优先级从高到低的顺序，遍历并获取任务集队列中的各个任务集。例如，任务集队列中包括任务集 1、任务集 2 和任务集 3，任务集 1 的优先级 1 > 任务集 2 的优先级 2 > 任务集 3 的优先级 3。主调度服务模块按照优先级从高到低的顺序，遍历并获取任务集队列中的各个任务集时，先获取任务集 1，再获取任务集 2，最后获取任务集 3。

25 可选的，为了便于主调度服务模块调度任务集，任务集队列可以按照优先级划分为多个子队列，每一优先级对应一个子队列。此时，主调度服务模块按照优先级从高到低的顺序，遍历各个子队列中的任务集，并获取子队列中的任务集。

可选的，为避免同一任务集重复调度，主调度服务模块在将一任务集发

送给计算模块后，将该任务集标记为已调度。此时，主调度服务模块遍历任务集队列中的各个任务集，获取未被标记为已调度的任务集，将获取的任务集发送给计算模块。之后，主调度服务模块再将该获取的任务集标记为已调度。

- 5 可选的，为节约计算资源，多个交互分析服务模块若在将得到的任务集缓存至任务集队列的第一预设时长后，检测到结果集队列中不存在得到的任务集所对应的结果集，则将得到的任务集标记为终止。主调度服务模块遍历任务集队列中的各个任务集，获取未被标记为终止的任务集，将获取的任务集发送给计算模块。之后，主调度服务模块在未获得获取的任务集对应的结果集的情况下，检测到获取的任务集被标记为终止，则停止对该获取的任务集的调度，计算模块停止对该获取的任务集计算。
- 10

本申请实施例中，主调度服务模块可根据实际需要调整从任务集队列中获取任务集的方式，任务集调度策略为可插拔式的，即上述 3 种主调度服务模块获取任务集的方式可以任意结合。例如，主调度服务模块可按照优先级从任务集队列中获取任务集，还可以依据任务集的标记获取任务集，例如，

15 获取未被标记为已调度的任务集，获取未被标记为终止的任务集，获取未被标记为已调度且未被标记为终止的任务集，或者按照优先级获取未被标记为已调度且未被标记为终止的任务集等。

可选的，为提高数据处理系统的数据处理效率，主调度服务模块获取任务集时，判断该任务集所需的计算资源是否超过多个计算模块中当前剩余的

20 计算资源。这里，任务集所需的计算资源可为交互分析服务模块在对数据请求包括的处理逻辑进行资源评估时获取到的。多个计算模块中当前剩余的计算资源为：多个计算模块中每个计算模块剩余计算资源的总和。

主调度服务模块若判定该任务集所需的计算资源未超过多个计算模块中

25 当前剩余的計算资源，则获取该任务集。之后，将获取的任务集发送至多个计算模块。主调度服务模块在将获取的任务集发送至多个计算模块时，可以将完整的一个任务集发送给一个计算模块进行计算，也可以将任务集拆分为多份，发送给多个计算模块进行计算。本申请实施例不进行限定。

主调度服务模块若判定该任务集所需的计算资源超过多个计算模块中当

前剩余的计算资源，则将该任务集缓存至任务集队列中与获取的任务集的优先级相同的所有任务集之后。

例如，优先级 $x1$ 的子队列中包括任务集 $x1 \rightarrow$ 任务集 $x2 \rightarrow$ 任务集 $x3$ 。主调度服务模块遍历到优先级 $x1$ 的任务集 $x1$ 后，若判定任务集 $x1$ 所需的计算资源超过多个计算模块中当前剩余的计算资源，则主调度服务模块将任务集 $x1$ 加入优先级 $x1$ 的子队列的尾部，即优先级 $x1$ 的子队列更新为：任务集 $x2 \rightarrow$ 任务集 $x3 \rightarrow$ 任务集 $x1$ 。若判定任务集 $x1$ 所需的计算资源未超过多个计算模块中当前剩余的计算资源，则主调度服务模块获取任务集 $x1$ 。

一个示例中，如图 3 所示的调度服务模块的任务调度处理流程，可包括如下步骤。

步骤 301，主调度服务模块按照优先级从高到低的顺序，遍历任务集队列中的各个任务集。

步骤 302，主调度服务模块每遍历到一个任务集，检测任务集是否被标记为终止。若未被标记为终止，执行步骤 303。若被标记为终止，则继续执行步骤 301，遍历任务集队列中的各个任务集，直至遍历结束。

这里，遍历结束可以理解为任务集队列均为被标记为终止和/或已调度的任务集。

步骤 303，主调度服务模块检测任务集是否被标记为已调度。若未被标记为已调度，执行步骤 304。若被标记为已调度，则继续执行步骤 301，遍历任务集队列中的各个任务集，直至遍历结束。

本申请实施例不限定步骤 302 和步骤 303 的执行顺序。若步骤 303 在步骤 302 之前执行，则步骤 303 中，主调度服务模块每遍历到一个任务集，检测到任务集未被标记为已调度，执行步骤 302。步骤 302 中主调度服务模块检测任务集未被标记为终止，执行步骤 304。

步骤 304，主调度服务模块检测执行任务集所需的计算资源是否超过当前剩余的计算资源。若未超过当前剩余的计算资源，则执行步骤 305。若超过当前剩余的计算资源，则执行步骤 306。

步骤 305, 主调度服务模块将任务集调度至计算模块。

步骤 306, 主调度服务模块将任务集加入任务集队列中该任务集的优先级对应的子队列的尾部。之后, 继续执行步骤 301, 遍历任务集队列中的各个任务集, 直至遍历结束。

- 5 多个计算模块 301-305 对接收的任务集进行计算, 得到计算结果; 将得到的计算结果发送给主调度服务模块。

其中, 多个计算模块 301-305 对接收的任务集进行计算, 即为多个计算模块 301-305 对接收的任务集包括的任务进行处理。

- 10 主调度服务模块对接收的计算结果进行封装, 得到结果集, 将得到的结果集缓存至结果集队列。

主调度服务模块对接收的计算结果进行序列化和封装, 得到结果集。主调度服务模块将得到的结果集缓存至结果集队列。

多个交互分析服务模块 101-103 从结果集队列中获取结果集; 将获取的结果集发送给对应的客户端。

- 15 本申请实施例中, 数据请求由哪个交互分析服务模块接收的, 数据请求的结果集就由哪个交互分析服务模块获取, 并发送给对应的客户端。这里, 对应的客户端即为发送获取的结果集对应的数据请求的客户端。

- 20 可选的, 上述数据处理系统还可包括删除模块。为了节约存储资源, 删除模块检测结果集队列中是否存在缓存时长达到第二预设时长的第一结果集。若存在, 删除模块删除第一结果集。

或者, 为了节约存储资源, 删除模块检测结果集队列中是否存在最后一次被访问至当前时刻的时长达到第三预设时长的第二结果集。若存在, 删除模块删除第二结果集。

- 25 可选的, 为了节约计算资源, 交互分析服务模块在接收到数据请求后, 检测结果集队列中是否存在该数据请求所包括的处理逻辑对应的结果集。若不存在, 则交互分析服务模块对该数据请求包括的处理逻辑进行处理, 得到任务集。若存在, 则交互分析服务模块直接从结果集队列中获取该数据请求

所包括的处理逻辑对应的结果集，并将获取的结果集发送给对应的客户端。

5 为了便于交互分析服务模块检测结果集队列中是否存在该数据请求所包括处理逻辑对应的结果集，交互分析服务模块在接收到数据请求后，可以生成能够唯一表示数据请求所包括的处理逻辑的标识。处理逻辑的标识可以通过 MD5 算法等计算生成。任务集队列中缓存有处理逻辑的标识与任务集的对应关系，结果集队列中缓存有处理逻辑的标识与结果集的对应关系。

10 交互分析服务模块检测结果集队列中是否存在接收的数据请求所包括的处理逻辑的标识（例如第一标识）。若存在，则交互分析服务模块确定结果集队列中存在该数据请求所包括的处理逻辑对应的结果集，从结果集队列中获取第一标识对应的结果集，并发送给对应的客户端。否则，交互分析服务模块确定结果集队列中不存在该数据请求所包括的处理逻辑对应的结果集，对该数据请求包括的处理逻辑进行处理，得到任务集。

15 应用本申请实施例，在数据处理系统中包括了任务集队列和结果集队列，交互分析服务模块将任务集缓存至任务集队列，调度服务模块从任务集队列中获取任务集。另外，调度服务模块将结果集缓存至结果集队列，交互分析服务模块从结果集队列中获取结果集，并将结果集发送给客户端。可见，通过任务集队列和结果集队列实现了交互分析服务模块和调度服务模块的解耦，多个交互分析服务模块可部署在多个节点上同时提供交互分析服务，即多个交互分析服务模块可同时提供数据交互分析处理，进而数据交互分析处理可以在多个节点上进行，提高了数据交互分析处理的并发性。

20 基于相同的发明构思，根据上述数据处理系统实施例，本申请实施例还提供了一种数据处理方法。参考图 4，图 4 为本申请实施例提供的数据处理方法的一种流程示意图。该方法应用于数据处理系统，数据处理系统包括：多个交互分析服务模块、多个调度服务模块和多个计算模块。

25 本申请实施例中，交互分析服务模块和调度服务模块可位于同一进程中，即交互分析服务模块和调度服务模块可运行于同一进程，为同一进程中的两个模块。交互分析服务模块之间关系对等，即多个交互分析服务模块可以同时接收并处理客户端发送的数据请求。调度服务模块是主备式结构，只有一个主调度服务模块提供服务，其它为备用调度服务模块，在主调度服务模块

挂掉后，从备用调度服务模块中选举出一个主调度服务模块。这样可以避免出现数据不同步的问题。

本申请实施例中，数据处理系统可依赖 zookeeper 等分布式协同组件技术，来实现主调度服务模块的选举和主备调度服务模块切换的功能，以及交互分析服务模块的负载均衡功能。

交互分析服务模块为数据处理系统的数据请求处理及任务集产生层，调度服务模块为任务集的消费调度执行层。交互分析服务模块和调度服务模块之间通过分布式缓存队列进行解耦交互。

计算模块可单独位于一个进程中。多个计算模块隶属分布式计算框架，例如 Spark。计算模块主要完成调度服务模块分发的计算任务。交互分析服务模块和调度服务模块与计算模块可以位于同一节点上，也可以位于不同的节点上。

本申请实施例中，数据处理系统中包括任务集队列和结果集队列。任务集队列和结果集队列均为分布式缓存队列，可通过分布式缓存技术实现，以使用数据处理系统的全部内存资源。

本申请实施例中，上述数据处理方法包括如下步骤。

步骤 401，多个交互分析服务模块接收各个客户端发送的数据请求；对接收的数据请求包括的处理逻辑进行处理，得到任务集；将得到的任务集缓存至任务集队列。

一个实施例中，多个交互分析服务模块对接收的数据包括的处理逻辑进行处理，得到任务集，可包括：对接收的数据请求包括的处理逻辑进行 SQL 解析与优化、计划生成和资源预估处理，得到处理结果；对得到的处理结果进行序列化和封装，得到任务集。可参考图 2 所示的交互分析服务模块的请求处理流程。

一个实施例中，多个交互分析服务模块将得到的任务集缓存至任务集队列，可包括：获取接收的数据请求所包括的处理逻辑的属性信息，将得到的任务集与获取的属性信息对应缓存至任务集队列。

步骤 402, 多个调度服务模块中的主调度服务模块从任务集队列中获取任务集; 将获取的任务集发送至多个计算模块。

一个实施例中, 任务集队列缓存了任务集与属性信息的对应关系, 属性信息包括优先级。这种情况下, 主调度服务模块从任务集队列中获取任务集,

5 可为: 按照优先级从高到低的顺序, 遍历并获取任务集队列中的各个任务集。

另一个实施例中, 主调度服务模块按照优先级从高到低的顺序, 遍历并获取任务集队列中的各个任务集, 可包括: 按照优先级从高到低的顺序, 遍历任务集队列中的各个任务集, 并判断任务集所需的计算资源是否超过所述多个计算模块中当前剩余的

10 计算资源; 若否, 则从任务集队列中获取该任务集; 若是, 则将该任务集缓存至任务集队列中与该任务集的优先级相同的所有任务集之后。

本申请实施例中, 主调度服务模块按照优先级从高到低的顺序, 遍历任务集队列中的各个任务集, 并判断任务集所需的计算资源是否超过所述多个计算模块中当前剩余的

15 计算资源, 包括: 主调度服务模块按照优先级从高到低的顺序, 遍历任务集队列中的各个任务集, 每遍历到一个任务集, 判断该任务集所需的计算资源是否超过所述多个计算模块中当前剩余的

一个实施例中, 主调度服务模块从任务集队列中获取任务集, 可包括: 遍历任务集队列中的各个任务集, 获取未被标记为已调度的任务集。

在主调度服务模块获取未被标记为已调度的任务集之后, 主调度服务模块将获取的任务集标记为已调度。

20 块将获取的任务集标记为已调度。

一个实施例中, 主调度服务模块从任务集队列中获取任务集, 可包括: 遍历任务集队列中的各个任务集, 获取未被标记为终止的任务集。

另外, 多个交互分析服务模块若在将得到的任务集缓存至任务集队列的第一预设时长后, 检测到结果集队列中不存在得到的任务集所对应的结果集,

25 则将得到的任务集标记为终止。

步骤 403, 多个计算模块对接收的任务集进行计算, 得到计算结果; 将得到的计算结果发送给主调度服务模块。

步骤 404, 主调度服务模块对接收的计算结果进行封装, 得到结果集; 将得到的结果集缓存至结果集队列。

主调度服务模块对接收的计算结果进行序列化和封装, 得到结果集。主调度服务模块将得到的结果集缓存至结果集队列。

- 5 步骤 405, 多个交互分析服务模块从结果集队列中获取结果集; 将获取的结果集发送给对应的客户端。

本申请实施例中, 数据请求由哪个交互分析服务模块接收的, 数据请求的结果集就由哪个交互分析服务模块获取, 并发送给对应的客户端。这里, 对应的客户端即为发送获取的结果集对应的数据请求的客户端。

- 10 一个实施例中, 数据处理系统还可包括删除模块。此时, 上述数据处理方法还可包括: 删除模块检测结果集队列中是否存在缓存时长达到第二预设时长的第一结果集; 若存在, 删除第一结果集。

- 15 另一个实施例中, 上述数据处理方法还可包括: 删除模块检测结果集队列中是否存在最后一次被访问至当前时刻的时长达到第三预设时长的第二结果集; 若存在, 删除第二结果集。

一个实施例中, 多个交互分析服务模块对接收的数据请求包括的处理逻辑进行处理, 得到任务集, 可包括: 检测结果集队列中是否存在接收的数据请求所包括的处理逻辑对应的结果集; 若不存在, 则对接收的数据请求包括的处理逻辑进行数据处理, 得到任务集。

- 20 若确定结果集队列中存在接收的数据请求所包括的处理逻辑对应的结果集, 则从结果集队列中获取接收的数据请求所包括的处理逻辑对应的结果集; 将获取的结果集发送给对应的客户端。

- 25 应用本申请实施例, 在数据处理系统中包括了任务集队列和结果集队列, 交互分析服务模块将任务集缓存至任务集队列, 调度服务模块从任务集队列中获取任务集。另外, 调度服务模块将结果集缓存至结果集队列, 交互分析服务模块从结果集队列中获取结果集, 并将结果集发送给客户端。可见, 通过任务集队列和结果集队列, 实现了交互分析服务模块和调度服务模块的解耦, 多个交互分析服务模块可部署在多个节点上同时提供交互分析服务, 即

多个交互分析服务模块可同时提供数据交互分析处理，进而数据交互分析处理可以在多个节点上进行，提高了数据交互分析处理的并发性。

基于相同的发明构思，根据上述数据处理系统实施例，本申请实施例还提供了一种机器可读存储介质，机器可读存储介质内存储有计算机程序，计算机程序被处理器执行时实现上述图 4 所示的数据处理方法。

需要说明的是，在本文中，诸如第一和第二等之类的关系术语仅仅用来将一个实体或者操作与另一个实体或操作区分开来，而不一定要求或者暗示这些实体或操作之间存在任何这种实际的关系或者顺序。而且，术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含，从而使得包括一系列要素的过程、方法、物品或者设备不仅包括那些要素，而且还包括没有明确列出的其他要素，或者是还包括为这种过程、方法、物品或者设备所固有的要素。在没有更多限制的情况下，由语句“包括一个……”限定的要素，并不排除在包括所述要素的过程、方法、物品或者设备中还存在另外的相同要素。

本说明书中的各个实施例均采用相关的方式描述，各个实施例之间相同相似的部分互相参见即可，每个实施例重点说明的都是与其他实施例的不同之处。尤其，对于数据处理方法和机器可读存储介质实施例而言，由于其基本相似于数据处理系统实施例，所以描述的比较简单，相关之处参见数据处理系统实施例的部分说明即可。

以上所述仅为本申请的较佳实施例而已，并非用于限定本申请的保护范围。凡在本申请的精神和原则之内所作的任何修改、等同替换、改进等，均包含在本申请的保护范围内。

权 利 要 求

1、一种数据处理方法，其特征在于，应用于数据处理系统，所述数据处理系统包括：多个交互分析服务模块、多个调度服务模块和多个计算模块；所述方法包括：

5 所述多个交互分析服务模块接收各个客户端发送的数据请求；对接收的数据请求包括的处理逻辑进行处理，得到任务集；将得到的任务集缓存至任务集队列；

所述多个调度服务模块中的主调度服务模块从所述任务集队列中获取任务集；将获取的任务集发送至所述多个计算模块；

10 所述多个计算模块对接收的任务集进行计算，得到计算结果；将得到的计算结果发送给所述主调度服务模块；

所述主调度服务模块对接收的计算结果进行封装，得到结果集；将得到的结果集缓存至结果集队列；

15 所述多个交互分析服务模块从所述结果集队列中获取结果集；将获取的结果集发送给对应的客户端。

2、根据权利要求 1 所述的方法，其特征在于，所述多个交互分析服务模块对接收的数据请求包括的处理逻辑进行处理，得到任务集，包括：

20 所述多个交互分析服务模块对接收的数据请求包括的处理逻辑进行结构化查询语言 SQL 解析与优化、计划生成和资源预估处理，得到处理结果；对得到的处理结果进行序列化和封装，得到任务集。

3、根据权利要求 1 所述的方法，其特征在于，所述多个交互分析服务模块将得到的任务集缓存至任务集队列，包括：

所述多个交互分析服务模块获取接收的数据请求所包括的处理逻辑的属性信息，将得到的任务集与获取的属性信息对应的缓存至任务集队列。

25 4、根据权利要求 3 所述的方法，其特征在于，所述属性信息包括优先级；
所述主调度服务模块从所述任务集队列中获取任务集，包括：

所述主调度服务模块按照优先级从高到低的顺序，遍历并获取所述任务集队列中的各个任务集。

5、根据权利要求4所述的方法，其特征在于，所述主调度服务模块按照优先级从高到低的顺序，遍历并获取所述任务集队列中的各个任务集，包括：

- 5 所述主调度服务模块按照优先级从高到低的顺序，遍历所述任务集队列中的各个任务集，并判断任务集所需的计算资源是否超过所述多个计算模块中当前剩余的计算资源；若否，则从所述任务集队列中获取该任务集；若是，则将该任务集缓存至所述任务集队列中与该任务集的优先级相同的所有任务集之后。

- 10 6、根据权利要求1所述的方法，其特征在于，所述主调度服务模块从所述任务集队列中获取任务集，包括：

所述主调度服务模块遍历所述任务集队列中的各个任务集，获取未被标记为已调度的任务集；

在所述主调度服务模块获取未被标记为已调度的任务集之后，还包括：

- 15 所述主调度服务模块将获取的任务集标记为已调度。

7、根据权利要求1所述的方法，其特征在于，所述主调度服务模块从所述任务集队列中获取任务集，包括：

所述主调度服务模块遍历所述任务集队列中的各个任务集，获取未被标记为终止的任务集；

- 20 所述方法还包括：

所述多个交互分析服务模块若在将得到的任务集缓存至任务集队列的第一预设时长后，检测到所述结果集队列中不存在得到的任务集所对应的结果集，则将得到的任务集标记为终止。

- 25 8、根据权利要求1所述的方法，其特征在于，所述数据处理系统还包括：删除模块；所述方法还包括：

所述删除模块检测所述结果集队列中是否存在缓存时长达到第二预设时

长的第一结果集；若存在，删除所述第一结果集；或者，

所述删除模块检测所述结果集队列中是否存在最后一次被访问至当前时刻的时长达到第三预设时长的第二结果集；若存在，删除所述第二结果集。

5 9、根据权利要求1所述的方法，其特征在于，所述多个交互分析服务模块对接收的数据请求包括的处理逻辑进行处理，得到任务集，包括：

所述多个交互分析服务模块检测结果集队列中是否存在接收的数据请求所包括的处理逻辑对应的结果集；若不存在，则对接收的数据请求包括的处理逻辑进行数据处理，得到任务集。

10、根据权利要求9所述的方法，其特征在于，所述方法还包括：

10 所述多个交互分析服务模块若确定所述结果集队列中存在接收的数据请求所包括的处理逻辑对应的结果集，则从所述结果集队列中获取接收的数据请求所包括的处理逻辑对应的结果集；将获取的结果集发送给对应的客户端。

11、根据权利要求1-10任一项所述的方法，其特征在于，所述交互分析服务模块和所述调度服务模块位于同一进程中。

15 12、一种数据处理系统，其特征在于，所述数据处理系统包括：多个交互分析服务模块、多个调度服务模块和多个计算模块；

所述多个交互分析服务模块，用于接收各个客户端发送的数据请求；对接收的数据请求包括的处理逻辑进行处理，得到任务集；将得到的任务集缓存至任务集队列；

20 所述多个调度服务模块中的主调度服务模块，用于从所述任务集队列中获取任务集；将获取的任务集发送至所述多个计算模块；

所述多个计算模块，用于对接收的任务集进行计算，得到计算结果；将得到的计算结果发送给所述主调度服务模块；

25 所述主调度服务模块，用于接收的计算结果进行封装，得到结果集；将得到的结果集缓存至结果集队列；

所述多个交互分析服务模块，用于从所述结果集队列中获取结果集；将

获取的结果集发送给对应的客户端。

13、根据权利要求 12 所述的系统，其特征在于，所述多个交互分析服务模块具体用于：

5 对接收的数据请求包括的处理逻辑进行结构化查询语言 SQL 解析与优化、计划生成和资源预估处理，得到处理结果；对得到的处理结果进行序列化和封装，得到任务集。

14、根据权利要求 12 所述的系统，其特征在于，所述多个交互分析服务模块具体用于：获取接收的数据请求所包括的处理逻辑的属性信息，将得到的任务集与获取的属性信息对应的缓存至任务集队列。

10 15、根据权利要求 14 所述的系统，其特征在于，所述属性信息包括优先级；

所述主调度服务模块具体用于：按照优先级从高到低的顺序，遍历并获取所述任务集队列中的各个任务集。

15 16、根据权利要求 15 所述的系统，其特征在于，所述主调度服务模块具体用于：

按照优先级从高到低的顺序，遍历所述任务集队列中的各个任务集，并判断任务集所需的计算资源是否超过所述多个计算模块中当前剩余的计算资源；若否，则从所述任务集队列中获取该任务集；若是，则将该任务集缓存至所述任务集队列中与该任务集的优先级相同的所有任务集之后。

20 17、根据权利要求 12 所述的系统，其特征在于，所述主调度服务模块具体用于：遍历所述任务集队列中的各个任务集，获取未被标记为已调度的任务集；

所述主调度服务模块，还用于将获取的任务集标记为已调度。

25 18、根据权利要求 12 所述的系统，其特征在于，所述主调度服务模块具体用于：遍历所述任务集队列中的各个任务集，获取未被标记为终止的任务集；

所述多个交互分析服务模块，还用于若在将得到的任务集缓存至任务集

—23—

队列的第一预设时长后，检测到所述结果集队列中不存在得到的任务集所对应的结果集，则将得到的任务集标记为终止。

19、根据权利要求 12 所述的系统，其特征在于，所述数据处理系统还包括：删除模块；

5 所述删除模块，用于检测所述结果集队列中是否存在缓存时长达到第二预设时长的第一结果集；若存在，删除所述第一结果集；或者，

所述删除模块，用于检测所述结果集队列中是否存在最后一次被访问至当前时刻的时长达到第三预设时长的第二结果集；若存在，删除所述第二结果集。

10 20、根据权利要求 12 所述的系统，其特征在于，所述多个交互分析服务模块具体用于：

检测结果集队列中是否存在接收的数据请求所包括的处理逻辑对应的结果集；若不存在，则对接收的数据请求包括的处理逻辑进行数据处理，得到任务集。

15 21、根据权利要求 20 所述的系统，其特征在于，

所述多个交互分析服务模块，还用于若确定所述结果集队列中存在接收的数据请求所包括的处理逻辑对应的结果集，则从所述结果集队列中获取接收的数据请求所包括的处理逻辑对应的结果集；将获取的结果集发送给对应的客户端。

20 22、根据权利要求 12-21 任一项所述的系统，其特征在于，所述交互分析服务模和所述调度服务模块位于同一进程中。

23、一种机器可读存储介质，其特征在于，所述机器可读存储介质内存储有计算机程序，所述计算机程序被处理器执行时实现权利要求 1-11 任一所述的方法步骤。

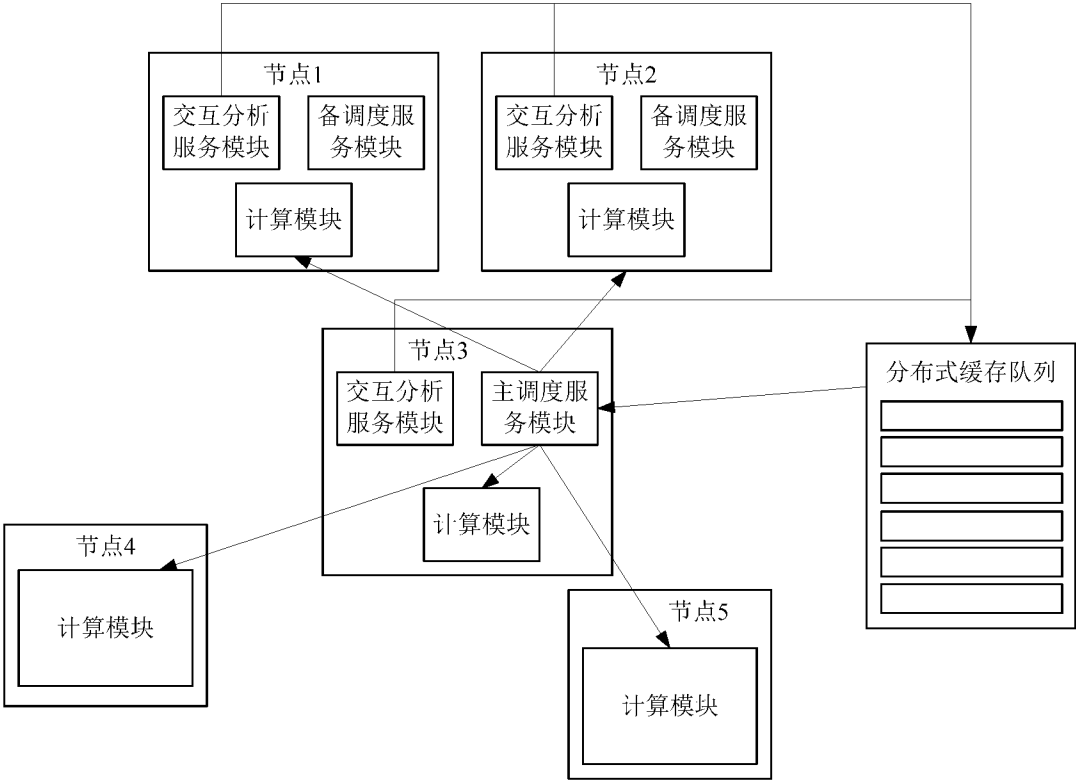


图 1

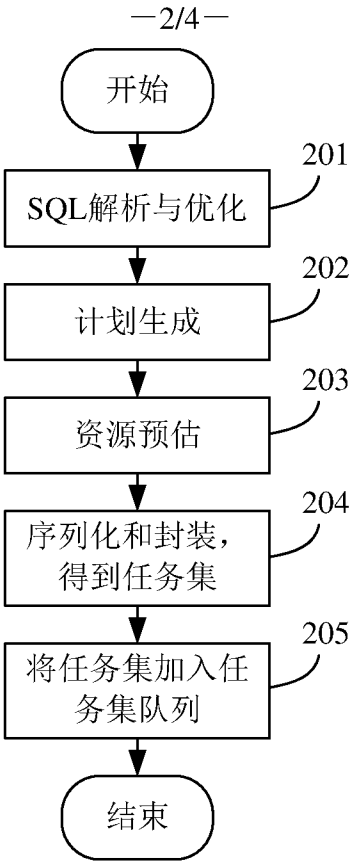


图 2

— 3/4 —

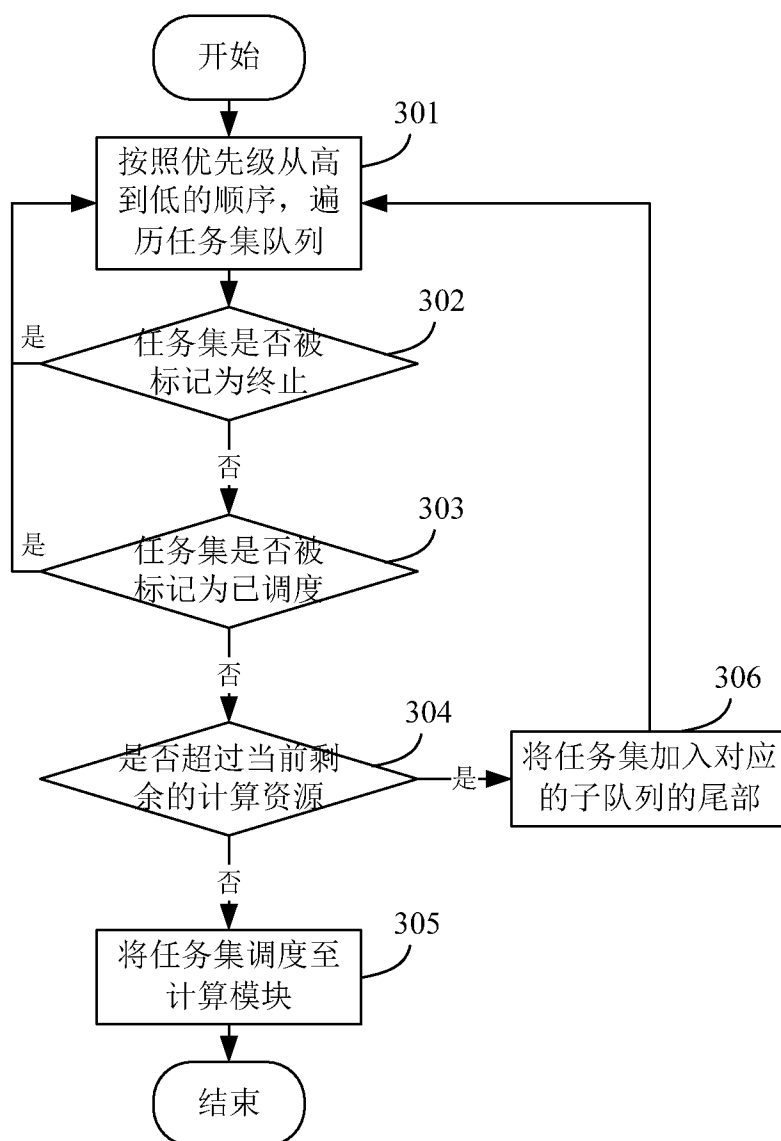


图 3

—4/4—

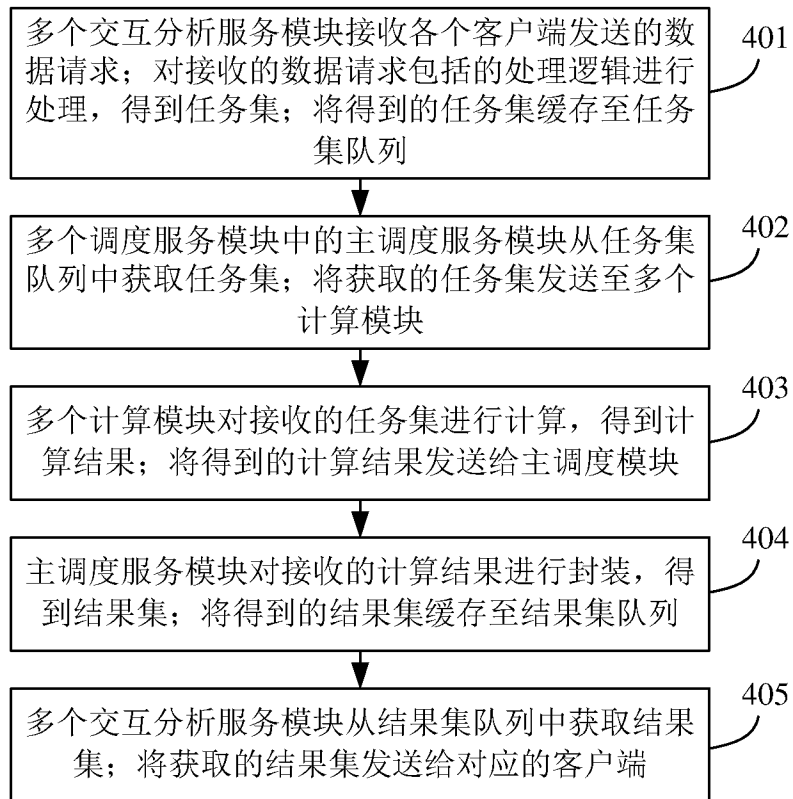


图 4

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/086847

A. CLASSIFICATION OF SUBJECT MATTER

G06F 16/00(2019.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNABS; CNTXT; DWPI; VEN; USTXT; EPTXT; WOTXT; CNKI: 数据处理, 调度, 任务集, 结果集, 缓存, 队列, 请求, 客户端, 优先级, data process, dispatch, schedule, task set, result set, buffer, cache, queue, request, client, priority

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 106095572 A (NETPOSA TECHNOLOGIES LTD.) 09 November 2016 (2016-11-09) description, paragraphs [0005]-[0051], and figure 1	1-23
A	CN 104657214 A (ECDATA INFORMATION TECHNOLOGY CO., LTD.) 27 May 2015 (2015-05-27) description, paragraphs [0006]-[0016]	1-23
A	CN 107665144 A (BEIJING JINGDONG SHANGKE INFORMATION TECHNOLOGY CO., LTD. ET AL.) 06 February 2018 (2018-02-06) entire document	1-23
A	CN 103077074 A (SHENZHEN INSTITUTES OF ADVANCED TECHNOLOGY) 01 May 2013 (2013-05-01) entire document	1-23
A	US 2007250681 A1 (IBM) 25 October 2007 (2007-10-25) entire document	1-23

☐ 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

07 August 2019

Date of mailing of the international search report

16 August 2019

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/CN2019/086847

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)		Publication date (day/month/year)
CN	106095572	A	09 November 2016	None		
CN	104657214	A	27 May 2015	None		
CN	107665144	A	06 February 2018	None		
CN	103077074	A	01 May 2013	CN	103077074	03 August 2016
US	2007250681	A1	25 October 2007	None		

国际检索报告

国际申请号

PCT/CN2019/086847

A. 主题的分类 G06F 16/00 (2019.01) i 按照国际专利分类 (IPC) 或者同时按照国家分类和 IPC 两种分类		
B. 检索领域 检索的最低限度文献 (标明分类系统和分类号) G06F 包含在检索领域中的除最低限度文献以外的检索文献 在国际检索时查阅的电子数据库 (数据库的名称, 和使用的检索词 (如使用)) CNABS; CNTXT; DWPI; VEN; USTXT; EPTXT; WOTXT; CNKI: 数据处理, 调度, 任务集, 结果集, 缓存, 队列, 请求, 客户端, 优先级, data process, dispatch, schedule, task set, result set, buffer, cache, queue, request, client, priority		
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 106095572 A (东方网力科技股份有限公司) 2016年 11月 9日 (2016 - 11 - 09) 说明书第5-51段, 附图1	1-23
A	CN 104657214 A (华存数据信息技术有限公司) 2015年 5月 27日 (2015 - 05 - 27) 说明书第6-16段	1-23
A	CN 107665144 A (北京京东尚科信息技术有限公司 等) 2018年 2月 6日 (2018 - 02 - 06) 全文	1-23
A	CN 103077074 A (深圳先进技术研究院) 2013年 5月 1日 (2013 - 05 - 01) 全文	1-23
A	US 2007250681 A1 (IBM) 2007年 10月 25日 (2007 - 10 - 25) 全文	1-23
☐ 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		国际检索报告邮寄日期
2019年 8月 7日		2019年 8月 16日
ISA/CN的名称和邮寄地址		受权官员
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		房琦 电话号码 86-(20)-28950714

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2019/086847

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	106095572	A	2016年 11月 9日	无	
CN	104657214	A	2015年 5月 27日	无	
CN	107665144	A	2018年 2月 6日	无	
CN	103077074	A	2013年 5月 1日	CN 103077074	B 2016年 8月 3日
US	2007250681	A1	2007年 10月 25日	无	