



# (12) 发明专利申请

(10) 申请公布号 CN 114662468 A

(43) 申请公布日 2022. 06. 24

(21) 申请号 202210308942.0

(22) 申请日 2022.03.25

(71) 申请人 中国建设银行股份有限公司  
地址 100033 北京市西城区金融大街25号

(72) 发明人 李子锋

(74) 专利代理机构 中科专利商标代理有限责任  
公司 11021  
专利代理师 吴梦圆

(51) Int. Cl.  
G06F 40/205 (2020.01)  
G06Q 40/00 (2012.01)

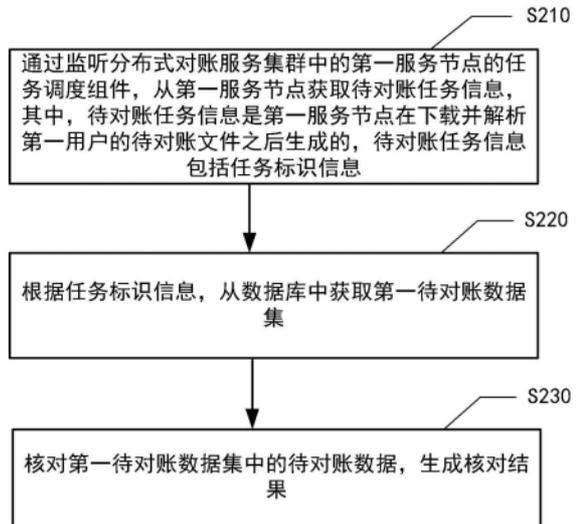
权利要求书2页 说明书12页 附图4页

## (54) 发明名称

对账数据处理方法、装置、设备及介质

## (57) 摘要

本公开提供了一种对账数据处理方法、装置、设备及介质,可以应用于大数据技术领域和金融科技领域。该对账数据处理方法包括:通过监听分布式对账服务集群中的第一服务节点的任务调度组件,从第一服务节点获取待对账任务信息,其中,待对账任务信息是第一服务节点在下载并解析第一用户的待对账文件之后生成的,待对账任务信息包括任务标识信息;根据任务标识信息,从数据库中获取第一待对账数据集;核对第一待对账数据集中的待对账数据,生成核对结果。



1. 一种对账数据处理方法,包括:

通过监听分布式对账服务集群中的第一服务节点的任务调度组件,从所述第一服务节点获取待对账任务信息,其中,所述待对账任务信息是所述第一服务节点在下载并解析第一用户的待对账文件之后生成的,所述待对账任务信息包括任务标识信息;

根据所述任务标识信息,从数据库中获得第一待对账数据集;

核对所述第一待对账数据集中的待对账数据,生成核对结果。

2. 根据权利要求1所述的方法,其中,所述根据所述任务标识信息,从数据库中获得第一待对账数据集,包括:

根据所述任务标识信息中的目标字段,确定目标索引信息;

根据所述目标索引信息,从数据库中获得所述第一待对账数据集。

3. 根据权利要求1所述的方法,所述核对结果包括差异结果和一致结果,还包括:

在所述核对结果为差异结果的情况下,记录所述差异结果对应的待对账数据;

在所述核对结果为一致结果的情况下,删除所述一致结果对应的待对账数据。

4. 根据权利要求3所述的方法,其中,所述在所述核对结果为一致结果的情况下,删除所述一致结果对应的待对账数据,包括:

在所述核对结果为一致结果的情况下,记录核对时刻信息;

在所述核对时刻信息与当前时刻信息的间隔超过预设时长的情况下,删除所述一致结果对应的待对账数据。

5. 根据权利要求1所述的方法,还包括:

通过下载并解析第二用户的待对账文件,得到所述第二待对账数据集;

根据所述第二待对账数据集和第二服务节点的计算线程量,创建第二待对账任务。

6. 根据权利要求5所述的方法,其中,第二待对账任务包括多个待对账子任务,所述根据所述第二待对账数据集和预设服务节点的计算线程量,创建第二待对账任务,包括:

根据所述第二待对账数据集和所述第二服务节点的计算线程量,确定所述待对账子任务的数量;

根据所述待对账子任务的数量,将所述第二待对账数据集拆分成多个待对账数据子集;

根据每一个所述待对账数据子集,创建所述待对账子任务。

7. 根据权利要求6所述的方法,所述根据所述待对账子任务的数量,将所述第二待对账数据集拆分成多个待对账数据子集,包括:

根据所述待对账子任务的数量,按照所述第二待对账数据集中的待对账数据的产生时刻信息,将所述第二待对账数据集拆分成多个所述待对账数据子集,其中,每一个所述待对账数据子集中的所述待对账数据的产生时刻信息相同。

8. 一种对账数据处理装置,包括:

第一获取模块,用于通过监听分布式对账服务集群中的第一服务节点的任务调度组件,从所述第一服务节点获取待对账任务信息,其中,所述待对账任务信息是所述第一服务节点在下载并解析第一用户的待对账文件之后生成的,所述待对账任务信息包括任务标识信息;

第二获取模块,用于根据所述任务标识信息,从数据库中获得第一待对账数据集;

核对模块,用于核对所述第一待对账数据集中的待对账数据,生成核对结果。

9.一种电子设备,包括:

一个或多个处理器;

存储装置,用于存储一个或多个程序,

其中,当所述一个或多个程序被所述一个或多个处理器执行时,使得所述一个或多个处理器执行根据权利要求1~7中任一项所述的方法。

10.一种计算机可读存储介质,其上存储有可执行指令,该指令被处理器执行时使处理器执行根据权利要求1~7中任一项所述的方法。

11.一种计算机程序产品,包括计算机程序,所述计算机程序被处理器执行时实现根据权利要求1~7中任一项所述的方法。

## 对账数据处理方法、装置、设备及介质

### 技术领域

[0001] 本公开涉及大数据技术领域,具体地涉及一种对账数据处理方法、装置、设备、介质和程序产品。

### 背景技术

[0002] 传统的对账过程中,是在下载并解析完待对账文件之后,服务器才开始执行对账任务。每个环节之间需要相互等待,串行化较严重,对账任务执行的时间段集中,使得在等待期间,服务器处于空闲或压力较小的状态,当执行对账任务时,服务器的瞬时压力较大,导致对账效率降低。

### 发明内容

[0003] 鉴于上述问题,本公开提供了一种对账数据处理方法、装置、设备、介质和程序产品。

[0004] 根据本公开的一个方面,提供了一种对账数据处理方法,包括:

[0005] 通过监听分布式对账服务集群中的第一服务节点的任务调度组件,从第一服务节点获取待对账任务信息,其中,待对账任务信息是第一服务节点在下载并解析第一用户的待对账文件之后生成的,待对账任务信息包括任务标识信息;

[0006] 根据任务标识信息,从数据库中获取第一待对账数据集;

[0007] 核对第一待对账数据集中的待对账数据,生成核对结果。

[0008] 根据本公开的实施例,根据任务标识信息,从数据库中获取第一待对账数据集,包括:

[0009] 根据任务标识信息中的目标字段,确定目标索引信息;

[0010] 根据目标索引信息,从数据库中,获取第一待对账数据集。

[0011] 根据本公开的实施例,核对结果包括差异结果和一致结果,上述对账数据处理方法还包括:

[0012] 在核对结果为差异结果的情况下,记录差异结果对应的待对账数据;

[0013] 在核对结果为一致结果的情况下,删除一致结果对应的待对账数据。

[0014] 根据本公开的实施例,在核对结果为一致结果的情况下,删除一致结果对应的待对账数据,包括:

[0015] 在核对结果为一致结果的情况下,记录核对时刻信息;

[0016] 在核对时刻信息与当前时刻信息的间隔超过预设时长的情况下,删除一致结果对应的待对账数据集。

[0017] 根据本公开的实施例,上述对账数据处理方法,还包括:

[0018] 通过下载并解析第二用户的待对账文件,得到第二待对账数据集;

[0019] 根据第二待对账数据集和第二服务节点的计算线程量,创建第二待对账任务。

[0020] 根据本公开的实施例,第二待对账任务包括多个待对账子任务,根据第二待对账

数据集和预设服务节点的计算线程量,创建第二待对账任务,包括:

[0021] 根据第二待对账数据集和第二服务节点的计算线程量,确定待对账子任务的数量;

[0022] 根据待对账子任务的数量,将第二待对账数据集拆分成多个待对账数据子集;

[0023] 根据每一个待对账数据子集,创建待对账子任务。

[0024] 根据本公开的实施例,根据待对账子任务的数量,将第二待对账数据集拆分成多个待对账数据子集,包括:

[0025] 根据待对账子任务的数量,按照第二待对账数据集中的待对账数据的产生时刻信息,将第二待对账数据集拆分成多个待对账数据子集,其中,每一个待对账数据子集中的待对账数据的产生时刻信息相同。

[0026] 本公开的另一个方面提供了一种对账数据处理装置,包括:第一获取模块、第二获取模块和核对模块。其中,第一获取模块,用于通过监听分布式对账服务集群中的第一服务节点的任务调度组件,从第一服务节点获取待对账任务信息,其中,待对账任务信息是第一服务节点在下载并解析第一用户的待对账文件之后生成的,待对账任务信息包括任务标识信息。第二获取模块,用于根据任务标识信息,从数据库中获取第一待对账数据集。核对模块,用于核对第一待对账数据集中的待对账数据,生成核对结果。

[0027] 根据本公开的实施例,第二获取模块包括第一确定单元和第一获取单元。其中,第一确定单元,用于根据任务标识信息中的目标字段,确定目标索引信息。第一获取单元,用于根据目标索引信息,从数据库中,获取第一待对账数据集。

[0028] 根据本公开的实施例,上述对账数据处理装置还包括记录模块和删除模块。其中,记录模块,用于在核对结果为差异结果的情况下,记录差异结果对应的待对账数据。删除模块,用于在核对结果为一致结果的情况下,删除一致结果对应的待对账数据。

[0029] 根据本公开的实施例,删除模块包括记录单元和删除单元。其中,记录单元,用于在核对结果为一致结果的情况下,记录核对时刻信息。删除单元,用于在核对时刻信息与当前时刻信息的间隔超过预设时长的情况下,删除一致结果对应的待对账数据集。

[0030] 根据本公开的实施例,上述对账数据处理装置还包括下载解析模块和创建模块。其中,下载解析模块,用于通过下载并解析第二用户的待对账文件,得到第二待对账数据集。创建模块,用于根据第二待对账数据集和第二服务节点的计算线程量,创建第二待对账任务。

[0031] 根据本公开的实施例,创建模块包括第二确定单元、拆分单元、创建单元。其中,第二确定单元,用于根据第二待对账数据集和第二服务节点的计算线程量,确定待对账子任务的数量。拆分单元,用于根据待对账子任务的数量,将第二待对账数据集拆分成多个待对账数据子集。创建单元,用于根据每一个待对账数据子集,创建待对账子任务。

[0032] 根据本公开的实施例,拆分单元包括拆分子单元。拆分子单元用于根据待对账子任务的数量,按照第二待对账数据集中的待对账数据的产生时刻信息,将第二待对账数据集拆分成多个待对账数据子集,其中,每一个待对账数据子集中的待对账数据的产生时刻信息相同。

[0033] 本公开的另一个方面提供了一种电子设备,包括:一个或多个处理器;存储器,用于存储一个或多个程序,其中,当所述一个或多个程序被所述一个或多个处理器执行时,使

得一个或多个处理器执行上述方法。

[0034] 本公开的另一个方面还提供了一种计算机可读存储介质,其上存储有可执行指令,该指令被处理器执行时使处理器执行上述对账数据处理方法。

[0035] 本公开的另一个方面还提供了一种计算机程序产品,包括计算机程序,该计算机程序被处理器执行时实现上述对账数据处理方法。

[0036] 根据本公开的实施例,通过监听分布式对账服务集群中的第一服务节点的任务调度组件,从第一服务节点获取待对账任务信息,根据任务标识信息,从数据库中获取第一待对账数据集;核对第一待对账数据集中的待对账数据,生成核对结果。由于待对账任务信息是第一服务节点在下载并解析第一用户的待对账文件之后生成的,第一用户的待对账文件下载解析完成之后,分布式对账服务集群中的其他服务节点就可以获取以用户为维度的待对账数据,因此,可以缩短对账环节之间的等待时间,降低服务节点的瞬时压力。

## 附图说明

[0037] 通过以下参照附图对本公开实施例的描述,本公开的上述内容以及其他目的、特征和优点将更为清楚,在附图中:

[0038] 图1示意性示出了根据本公开实施例的对账数据处理方法、装置、设备、介质和程序产品的应用场景图;

[0039] 图2示意性示出了根据本公开实施例的对账数据处理方法的流程图;

[0040] 图3示意性示出了根据本公开实施例的创建待对账任务的方法流程图;

[0041] 图4示意性示出了根据本公开实施例的对账数据处理方法的逻辑框图;

[0042] 图5示意性示出了根据本公开实施例的对账数据处理装置的结构框图;以及

[0043] 图6示意性示出了根据本公开实施例的适于实现对账数据处理方法的电子设备的方框图。

## 具体实施方式

[0044] 以下,将参照附图来描述本公开的实施例。但是应该理解,这些描述只是示例性的,而并非要限制本公开的范围。在下面的详细描述中,为便于解释,阐述了许多具体的细节以提供对本公开实施例的全面理解。然而,明显地,一个或多个实施例在没有这些具体细节的情况下也可以被实施。此外,在以下说明中,省略了对公知结构和技术的描述,以避免不必要地混淆本公开的概念。

[0045] 在此使用的术语仅仅是为了描述具体实施例,而并非意在限制本公开。在此使用的术语“包括”、“包含”等表明了所述特征、步骤、操作和/或部件的存在,但是并不排除存在或添加一个或多个其他特征、步骤、操作或部件。

[0046] 在此使用的所有术语(包括技术和科学术语)具有本领域技术人员通常所理解的含义,除非另外定义。应注意,这里使用的术语应解释为具有与本说明书的上下文相一致的含义,而不应以理想化或过于刻板的方式来解释。

[0047] 在使用类似于“A、B和C等中至少一个”这样的表述的情况下,一般来说应该按照本领域技术人员通常理解该表述的含义来予以解释(例如,“具有A、B和C中至少一个的系统”应包括但不限于单独具有A、单独具有B、单独具有C、具有A和B、具有A和C、具有B和C、和/或

具有A、B、C的系统等)。

[0048] 需要说明的是,本公开的对账数据处理方法和装置可用于金融领域和大数据技术领域,也可用于除金融领域之外的任意领域,本公开的对账数据处理方法和装置的应用领域不做限定。

[0049] 本公开的实施例提供了一种对账数据处理方法,通过监听分布式对账服务集群中的第一服务节点的任务调度组件,从第一服务节点获取待对账任务信息,根据任务标识信息,从数据库中获取第一待对账数据集;核对第一待对账数据集中的待对账数据,生成核对结果。由于待对账任务信息是第一服务节点在下载并解析第一用户的待对账文件之后生成的,第一用户的待对账文件下载解析完成之后,分布式对账服务集群中的其他服务节点就可以获取以用户为维度的待对账数据,因此,可以缩短对账环节之间的等待时间,降低服务节点的瞬时压力。

[0050] 图1示意性示出了根据本公开实施例的对账数据处理方法的应用场景图。

[0051] 如图1所示,根据该实施例的应用场景100可以包括服务节点101、102、103、网络104和数据库服务器105。网络104用以在服务节点101、102、103和数据库服务器105之间提供通信链路的介质。网络104可以包括各种连接类型,例如有线、无线通信链路或者光纤电缆等等。

[0052] 用户可以使用服务节点101、102、103通过网络104进行交互,以下载并解析待对账文件、发布待对账任务、处理待对账任务等等。也可以通过网络104使得服务节点101、102、103与数据库服务器105交互,以便于服务节点101、102、103从数据库服务器105查询待对账数据集。

[0053] 服务节点101、102、103可以包括一台或多台服务器。

[0054] 数据库服务器105可以是提供数据查询、数据存储服务的服务器,例如对服务节点101、102、103发出的数据查询请求服务(仅为示例)。数据库服务器可以对接收到的用户请求等数据进行分析等处理,并将处理结果(例如根据查询请求获取或生成的数据等)反馈给服务节点。

[0055] 需要说明的是,本公开实施例所提供的对账数据处理方法一般可以由服务节点101、102、103中的任意一个服务节点执行。相应地,本公开实施例所提供的对账数据处理装置一般可以设置于服务节点101、102、103中的任意一个服务节点。本公开实施例所提供的对账数据处理方法也可以由服务器集群执行。相应地,本公开实施例所提供的对账数据处理装置也可以设置于服务器集群中。

[0056] 应该理解,图1中的服务节点、网络 and 数据库服务器的数目仅仅是示意性的。根据实现需要,可以具有任意数目的服务节点、网络和数据库服务器。

[0057] 以下将基于图1描述的场景,通过图2~图4对公开实施例的对账数据处理方法进行详细描述。

[0058] 图2示意性示出了根据本公开实施例的对账数据处理方法的流程图。

[0059] 如图2所示,该实施例的对账数据处理方法包括操作S210~操作S230。

[0060] 在操作S210,通过监听分布式对账服务集群中的第一服务节点的任务调度组件,从第一服务节点获取待对账任务信息,其中,待对账任务信息是第一服务节点在下载并解析第一用户的待对账文件之后生成的,待对账任务信息包括任务标识信息。

[0061] 根据本公开的实施例,例如:待对账文件中可以包括多个用户的待对账文件,例如:可以包括用户A的待对账文件、用户B的待对账文件。分布式对账服务集群中可以包括多个服务节点,例如:服务节点A、服务节点B,第一服务节点可以为服务节点A。服务节点B可以通过监听服务节点A的任务调度组件,确定服务节点A下载并解析完用户A的待对账文件时,从服务节点A获取待对账任务信息。

[0062] 在操作S220,根据任务标识信息,从数据库中获取第一待对账数据集。

[0063] 根据本公开的实施例,根据本公开的实施例,任务标识信息可以包括批次信息、待对账数据的存储位置信息,其中,待对账数据的存储位置信息可以为存储所述待对账数据的数据库宽表的页码信息。例如:任务标识信息可以为1-表1-101,表示第一批宽表1中的第101页。可以从数据库中,获取第一批宽表1中第101页中的全部数据为第一待对账数据集。

[0064] 根据本公开的实施例,待对账数据的存储位置信息还可以包括待对账数据在数据库宽表中的行和/或列信息。例如:任务标识信息可以为表1-101-2-30,可以表示宽表1中的第101页的第2~30行的数据,或第2~30列的数据为第一代对账数据集。

[0065] 在操作S230,核对第一待对账数据集中的待对账数据,生成核对结果。

[0066] 根据本公开的实施例,第一待对账数据集中的待对账数据可以包括我方订单数据、对方订单数据,订单数据可以包括订单流水号、订单金额、订单产生的时刻信息等等。可以通过核对我方订单数据的订单流水号、订单金额、订单产生的时刻信息与对方订单数据的相关信息,生成核对结果。

[0067] 根据本公开的实施例,通过监听分布式对账服务集群中的第一服务节点的任务调度组件,从第一服务节点获取待对账任务信息,根据任务标识信息,从数据库中获取第一待对账数据集;核对第一待对账数据集中的待对账数据,生成核对结果。由于待对账任务信息是第一服务节点在下载并解析第一用户的待对账文件之后生成的,第一用户的待对账文件下载解析完成之后,分布式对账服务集群中的其他服务节点就可以获取以用户为维度的待对账数据,因此,可以缩短对账环节之间的等待时间,降低服务节点的瞬时压力。

[0068] 根据本公开的实施例,在分布式对账服务集群中,作为对账任务发布者的服务节点,可以通过定时任务组件,定时扫描商户的业务信息表,例如:每天的凌晨某一时刻扫描商户的业务信息表,在Java代码整合Mybait在数据库商户清算进度表中生成当日的商户清算进度控制表和对账批次任务表。商户清算进度控制表可以包括每个商户当日的对账单下载情况、对账单解析情况、清算情况、转账状态等等信息。对账批次任务表中包括当日对账任务的批次号。对账任务发布者的服务节点可以通过定时扫描商户清算进度控制表获取待对账任务信息。

[0069] 根据本公开的实施例,Java代码通过第三方组件HttpClient发起http请求下载待对账文件。当扫描商户清算进度控制表的下载任务的状态,当下载完成一个商户的待对账文件时,即可通过任务调度组件将任务发布。例如:对于下载并解析完成的待对账文件,作为对账任务发布者的服务节点可以通过扫描对账批次主任务表,日期为今天且执行状态为0且当前时间大于主任务中执行时间的最近一条主对账任务,得到主任务的批次号。

[0070] 根据本公开的实施例,上述定时任务组件、第三方组件HttpClient、以及任务调度组件均可以通过jar包的形式即成到对账清算微服务中。

[0071] 根据本公开的实施例,作为对账任务发布者的服务节点可以创建当前对账主任务的子任务。扫描商户对账清算进度表中当日对账单解析状态为“已完成”状态,对账状态为“未完成”待执行的商户。更新这些商户的订单表设置账子任务号(主任务的批次号+子任务序号),每次一万条(可配置),更新一次后将子任务编号以及主任务序号等信息保存到对账子任务表中,并通过任务调度组件将任务发布等待作为对账任务处理者的服务节点消费。子任务创建完成后将对账批次主任务表执行状态更新为“执行中”的状态。

[0072] 根据本公开的实施例,待对账文件的下载,也可以通过任务发布的形式由作为对账任务发布者的服务节点执行。例如:作为对账任务发布者的服务节点可以通过扫描商户清算进度控制表,为下载状态为“未下载”的商户创建待对账文件下载任务,通过任务调度组件将任务发布,并将商户清算进度控制表中的下载状态更改为“执行中”。待作为下载任务处理者的服务节点执行完下载任务之后,根据下载任务完成的情况,向作为对账任务发布者的服务节点反馈信息,作为对账任务发布者的服务节点根据反馈信息更改商户清算进度控制表的下载任务的状态,例如:将下载任务状态修改为“成功”或“失败”。

[0073] 根据本公开的实施例,待对账文件的解析,也可以通过任务发布的形式由作为对账任务发布者的服务节点执行。例如:作为对账任务发布者的服务节点可以通过扫描商户清算进度控制表,为下载状态为“成功”解析状态为“未完成”的商户创建待对账文件解析任务,通过任务调度组件将任务发布,并将商户清算进度控制表中的解析状态更改为“执行中”。待作为解析任务处理者的服务节点执行完解析任务之后,根据解析任务完成的情况,向作为对账任务发布者的服务节点反馈信息,作为对账任务发布者的服务节点根据反馈信息更改商户清算进度控制表的解析任务的状态,例如:将解析任务状态修改为“成功”或“失败”。

[0074] 根据本公开的实施例,分布式对账服务集群可以通过观察者模式解耦生产者服务与消费者服务,因此,在需要扩容时,新的服务节点只需要配置任务调度组件即可加入到该分布式对账服务集群中,实现快速扩容,以应对待对账数据量突发性增长的情况。

[0075] 根据本公开的实施例,根据任务标识信息,从数据库中获得第一待对账数据集,包括:

[0076] 根据任务标识信息中的目标字段,确定目标索引信息;

[0077] 根据目标索引信息,从数据库中,获取第一待对账数据集。

[0078] 根据本公开的实施例,数据库可以为SQL数据库,在数据库中建立树形结构,将待对账数据存储在不同节点上,根据树形结构中节点位置建立索引信息,索引信息可以为包括树形结构中的节点信息,例如:1-2可以表示第一级子树中的第二个节点。从数据库中第一级子树中的第二个节点获取第一待对账数据集。

[0079] 根据本公开实施例,数据库中也可以通过邻接链表的方式存储待对账数据,索引信息可以为邻接链表的路径信息,通过邻接链表的路径信息从数据库中查询路径信息中包括的每个存储节点的数据,进而获取第一待对账数据集。例如:1-3-5-8,可以将邻接链表中的第1、3、5、8个存储节点的数据作为第一待对账数据集中的待对账数据。

[0080] 根据本公开的实施例,通过索引信息从数据库中查询待对账数据,缩短遍历整个数据库的耗时,提高数据查询速度。

[0081] 根据本公开的实施例,核对结果包括差异结果和一致结果,还包括:

[0082] 在核对结果为差异结果的情况下,记录差异结果对应的待对账数据;

[0083] 在核对结果为一致结果的情况下,删除一致结果对应的待对账数据。

[0084] 根据本公开的实施例,由于导致差异结果的原因很多,还需要分析产生差异结果的具体原因,例如:解析对账文件时出现错误数据。因此,在核对结果为差异结果的情况下,可以记录差异结果对应的待对账数据。可以将差异结果对应的待对账数据与重新解析对账文件获得的待对账数据,若差异结果对应的待对账数据与重新解析对账文件获得的待对账数据一致,可以生成人工介入信号,申请人工介入进行重新对账。若不一致,可以根据重新解析对账文件获得的待对账数据重新对账。

[0085] 根据本公开的实施例,在核对结果为一致结果的情况下,删除一致结果对应的待对账数据。

[0086] 根据本公开的实施例,通过核对结果,确定记录或删除待对账数据,可以将已对平的数据及时删除,以优化数据库中的数据量;同时记录差异结果对应的待对账数据,以确定产生差异结果的原因,避免对账过程由于数据错误导致对账错误的情况出现。

[0087] 根据本公开的实施例,在核对结果为一致结果的情况下,删除一致结果对应的待对账数据,包括:

[0088] 在核对结果为一致结果的情况下,记录核对时刻信息;

[0089] 在核对时刻信息与当前时刻信息的间隔超过预设时长的情况下,删除一致结果对应的待对账数据。

[0090] 根据本公开的实施例,例如:核对结果为一致结果时的核对时刻信息为3月1日9:00,预设时长可以设置为3天,预设时长可以在配置文件中配置。如果当前时刻信息为3月2日12:00,那么,核对时刻信息与当前时刻信息的间隔为1天零3小时,未超过预设时长,不可以删除一致结果对应的待对账数据。如果当前时刻信息为3月5日9:00,那么,核对时刻信息与当前时刻信息的间隔为4天,超过预设时长,可以删除一致结果对应的待对账数据。

[0091] 根据本公开的实施例,通过设置预设时长,可以给删除一致结果对应的待对账数据一段缓冲时间,以防止一致结果对应的待对账数据发生变化导致对账结果有误。

[0092] 根据本公开的实施例,上述对账数据处理方法还包括:

[0093] 通过下载并解析第二用户的待对账文件,得到第二待对账数据集;

[0094] 根据第二待对账数据集和第二服务节点的计算线程量,创建第二待对账任务。

[0095] 根据本公开的实施例,处理待对账任务的服务节点,也可以创建待对账任务。例如:分布式对账服务集群中包括服务节点A和服务节点B。服务节点A按照本公开实施例提供的对账数据处理方法处理完待对账任务之后,可以通过下载并解析用户B的待对账文件,得到第二待对账数据集。

[0096] 根据本公开的实施例,可以根据第二待对账数据集的数据量和第二服务节点的计算线程量,创建第二待对账任务。例如:第二待对账数据集中包含100万条数据,服务节点B的计算线程量可以为单次读取量,例如:单次读取量为5000条,可以创建一条包含5000条数据的待对账数据的第二待对账任务,第二待对账任务中可以包括200条子任务,服务节点B一次仅从服务节点A获取一条子任务。

[0097] 根据本公开的实施例,分布式对账服务集群中任意一个服务节点既可以是待对账任务的处理器,也可以是待对账任务的分发者,可以提高服务节点的高可用性,任意一个服

务节点发生故障,都不影响其他服务节点的正常工作。

[0098] 图3示意性示出了根据本公开实施例的创建待对账任务的方法流程图。

[0099] 如图3所示,该实施例包括操作S310~S330。

[0100] 在操作S310,根据第二待对账数据集和第二服务节点的计算线程量,确定待对账子任务的数量。

[0101] 在操作S320,根据待对账子任务的数量,将第二待对账数据集拆分成多个待对账数据子集。

[0102] 在操作S330,根据每一个待对账数据子集,创建待对账子任务。

[0103] 根据本公开的实施例,例如:分布式对账服务集群中包括服务节点B、服务节点A,第二待对账数据集中包含100万条数据,服务节点B可以包括多台服务器,计算线程量可以为每一台服务器的单次读取量,例如:每一台服务器的单次读取量为5000条,服务节点B包括10台服务器。服务节点A可以创建20条包含5000条待对账数据的第二待对账子任务,服务节点B中的每一台服务器都可以从服务节点A获取一条待对账子任务。

[0104] 根据本公开的实施例,例如:分布式对账服务集群中还包括服务节点C,第二待对账数据集中包含100万条数据,服务节点B包括10台服务器,每一台服务器的单次读取量为5000条,服务节点C包括5台服务器,每一台服务器的单次读取量为10万条。服务节点A可以创建10条包含5000条待对账数据的第二待对账子任务,创建5条包含10万条待对账数据的第二待对账子任务。

[0105] 根据本公开的实施例,以创建20条包含5000条待对账数据的第二待对账子任务,可以将第二待对账数据集中的100万条数据,拆分成20个待对账数据子集,每一个待对账数据子集中包括5000条待对账数据。需要说明的是,在实际应用中,待对账数据的拆分之后,不是恰好形成整数个待对账数据子集,则可以按照每个待对账子集中的数据量小于或等于单次读取量进行拆分,每个待对账子集中的数据量不一定相同。

[0106] 根据本公开的实施例,通过服务节点的计算线程量,创建子任务,可以将数据量较大的任务根据每个服务节点的处理能力,拆分成数据量较小的任务,并行分发给多个分布式服务集群中的服务节点,以充分利用集群中服务节点的数据处理功能,环节单机计算对服务器造成的压力。

[0107] 根据本公开的实施例,根据待对账子任务的数量,将第二待对账数据集拆分成多个待对账数据子集,包括:

[0108] 根据待对账子任务的数量,按照第二待对账数据集中的待对账数据的产生时刻信息,将第二待对账数据集拆分成多个待对账数据子集,其中,每一个待对账数据子集中的待对账数据的产生时刻信息相同。

[0109] 根据本公开的实施例,例如:待对账子任务的数量为20,第二待对账数据集中共100万条待对账数据,则可以确定每个待对账数据子集中包含5000条待对账数据。通过查询待对账数据产生的时刻信息,可以将时刻信息相同的数据归集到一个待对账数据子集中,例如:第二待对账数据集中4980条待对账数据是2月1日10:00产生的,而第二待对账数据集中没有与上述待对账数据产生时刻信息相同的其他数据,可以将这4980条待对账数据作为一个待对账数据子集。

[0110] 根据本公开的实施例,将待对账数据产生时刻信息相同的数据作为一个待对账数

据子集,创建待对账子任务,以便于服务节点在处理该待对账任务时,可以减少从数据库查询待对账数据的次数。

[0111] 图4示意性示出了根据本公开实施例的对账数据处理方法的逻辑框图。

[0112] 如图4所示,该实施例包括操作S410~S480。

[0113] 在操作S410,通过监听分布式对账服务集群中的第一服务节点的任务调度组件,从第一服务节点获取待对账任务信息,其中,待对账任务信息是第一服务节点在下载并解析第一用户的待对账文件之后生成的,待对账任务信息包括任务标识信息。

[0114] 在操作S420,根据任务标识信息,从数据库中获取第一待对账数据集。

[0115] 在操作S430,核对第一待对账数据集中的待对账数据,生成核对结果。

[0116] 在操作S440,确定核对结果是否为差异结果,若是,执行操作S450,若不是,执行操作S460。

[0117] 在操作S450,记录差异结果对应的待对账数据。

[0118] 在操作S460,记录核对时刻信息。

[0119] 在操作S470,确定核对时刻信息与当前时刻信息的间隔是否超过预设时长,若是,则执行操作S480,若不是,则不作处理。

[0120] 在操作S480,删除一致结果对应的待对账数据。

[0121] 基于上述对账数据处理方法,本公开还提供了一种对账数据处理装置。以下将结合图5对该装置进行详细描述。

[0122] 图5示意性示出了根据本公开实施例的对账数据处理装置的结构框图。

[0123] 如图5所示,该实施例的对账数据处理装置500包括第一获取模块510、第二获取模块520和核对模块530。

[0124] 第一获取模块510,用于通过监听分布式对账服务集群中的第一服务节点的任务调度组件,从所述第一服务节点获取待对账任务信息,其中,所述待对账任务信息是所述第一服务节点在下载并解析第一用户的待对账文件之后生成的,所述待对账任务信息包括任务标识信息。在一实施例中,第一获取模块510,可以用于执行本公开实施例的操作S210,在此不做赘述。

[0125] 第二获取模块,用于根据所述任务标识信息,从数据库中获取第一待对账数据集。在一实施例中,第二获取模块520,可以用于执行本公开实施例的操作S220,在此不做赘述。

[0126] 核对模块,用于核对所述第一待对账数据集中的待对账数据,生成核对结果。在一实施例中,核对模块530,可以用于执行本公开实施例的操作S230,在此不做赘述。

[0127] 根据本公开的实施例,第二获取模块包括第一确定单元和第一获取单元。其中,第一确定单元,用于根据任务标识信息中的目标字段,确定目标索引信息。第一获取单元,用于根据目标索引信息,从数据库中,获取第一待对账数据集。

[0128] 根据本公开的实施例,上述对账数据处理装置还包括记录模块和删除模块。其中,记录模块,用于在核对结果为差异结果的情况下,记录差异结果对应的待对账数据。删除模块,用于在核对结果为一致结果的情况下,删除一致结果对应的待对账数据。

[0129] 根据本公开的实施例,删除模块包括记录单元和删除单元。其中,记录单元,用于在核对结果为一致结果的情况下,记录核对时刻信息。删除单元,用于在核对时刻信息与当前时刻信息的间隔超过预设时长的情况下,删除一致结果对应的待对账数据集。

[0130] 根据本公开的实施例,上述对账数据处理装置还包括下载解析模块和创建模块。其中,下载解析模块,用于通过下载并解析第二用户的待对账文件,得到第二待对账数据集。创建模块,用于根据第二待对账数据集和第二服务节点的计算线程量,创建第二待对账任务。

[0131] 根据本公开的实施例,创建模块包括第二确定单元、拆分单元、创建单元。其中,第二确定单元,用于根据第二待对账数据集和第二服务节点的计算线程量,确定待对账子任务的数量。拆分单元,用于根据待对账子任务的数量,将第二待对账数据集拆分成多个待对账数据子集。创建单元,用于根据每一个待对账数据子集,创建待对账子任务。

[0132] 根据本公开的实施例,拆分单元包括拆分子单元。拆分子单元用于根据待对账子任务的数量,按照第二待对账数据集中的待对账数据的产生时刻信息,将第二待对账数据集拆分成多个待对账数据子集,其中,每一个待对账数据子集中的待对账数据的产生时刻信息相同。

[0133] 根据本公开的实施例,第一获取模块510、第二获取模块520和核对模块530中的任意多个模块可以合并在一个模块中实现,或者其中的任意一个模块可以被拆分成多个模块。或者,这些模块中的一个或多个模块的至少部分功能可以与其他模块的至少部分功能相结合,并在一个模块中实现。根据本公开的实施例,第一获取模块510、第二获取模块520和核对模块530中的至少一个可以至少被部分地实现为硬件电路,例如现场可编程门阵列(FPGA)、可编程逻辑阵列(PLA)、片上系统、基板上的系统、封装上的系统、专用集成电路(ASIC),或可以通过对电路进行集成或封装的任何其他的合理方式等硬件或固件来实现,或以软件、硬件以及固件三种实现方式中任意一种或以其中任意几种的适当组合来实现。或者,第一获取模块510、第二获取模块520和核对模块530中的至少一个可以至少被部分地实现为计算机程序模块,当该计算机程序模块被运行时,可以执行相应的功能。

[0134] 图6示意性示出了根据本公开实施例的适于实现……方法的电子设备的方框图。

[0135] 如图6所示,根据本公开实施例的电子设备600包括处理器601,其可以根据存储在只读存储器(ROM) 602中的程序或者从存储部分608加载到随机访问存储器(RAM) 603中的程序而执行各种适当的动作和处理。处理器601例如可以包括通用微处理器(例如CPU)、指令集处理器和/或相关芯片组和/或专用微处理器(例如,专用集成电路(ASIC))等等。处理器601还可以包括用于缓存用途的板载存储器。处理器601可以包括用于执行根据本公开实施例的方法流程的不同动作的单一处理单元或者是多个处理单元。

[0136] 在RAM 603中,存储有电子设备600操作所需的各种程序和数据。处理器601、ROM 602以及RAM 603通过总线604彼此相连。处理器601通过执行ROM 602和/或RAM 603中的程序来执行根据本公开实施例的方法流程的各种操作。需要注意,所述程序也可以存储在除ROM 602和RAM 603以外的一个或多个存储器中。处理器601也可以通过执行存储在所述一个或多个存储器中的程序来执行根据本公开实施例的方法流程的各种操作。

[0137] 根据本公开的实施例,电子设备600还可以包括输入/输出(I/O)接口605,输入/输出(I/O)接口605也连接至总线604。电子设备600还可以包括连接至I/O接口605的以下部件中的一项或多项:包括键盘、鼠标等的输入部分606;包括诸如阴极射线管(CRT)、液晶显示器(LCD)等以及扬声器等的输出部分607;包括硬盘等的存储部分608;以及包括诸如LAN卡、调制解调器等的网络接口卡的通信部分609。通信部分609经由诸如因特网的网络执行通信

处理。驱动器610也根据需要连接至I/O接口605。可拆卸介质611,诸如磁盘、光盘、磁光盘、半导体存储器等等,根据需要安装在驱动器610上,以便于从其上读出的计算机程序根据需要被安装入存储部分608。

[0138] 本公开还提供了一种计算机可读存储介质,该计算机可读存储介质可以是上述实施例中描述的设备/装置/系统中所包含的;也可以是单独存在,而未装配入该设备/装置/系统中。上述计算机可读存储介质承载有一个或者多个程序,当上述一个或者多个程序被执行时,实现根据本公开实施例的方法。

[0139] 根据本公开的实施例,计算机可读存储介质可以是非易失性的计算机可读存储介质,例如可以包括但不限于:便携式计算机磁盘、硬盘、随机访问存储器(RAM)、只读存储器(ROM)、可擦式可编程只读存储器(EPROM或闪存)、便携式紧凑磁盘只读存储器(CD-ROM)、光存储器件、磁存储器件、或者上述的任意合适的组合。在本公开中,计算机可读存储介质可以是任何包含或存储程序的有形介质,该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。例如,根据本公开的实施例,计算机可读存储介质可以包括上文描述的ROM 602和/或RAM 603和/或ROM 602和RAM 603以外的一个或多个存储器。

[0140] 本公开的实施例还包括一种计算机程序产品,其包括计算机程序,该计算机程序包含用于执行流程图所示的方法的程序代码。当计算机程序产品在计算机系统中运行时,该程序代码用于使计算机系统实现本公开实施例所提供的对账数据处理方法。

[0141] 在该计算机程序被处理器601执行时执行本公开实施例的系统/装置中限定的上述功能。根据本公开的实施例,上文描述的系统、装置、模块、单元等可以通过计算机程序模块来实现。

[0142] 在一种实施例中,该计算机程序可以依托于光存储器件、磁存储器件等有形存储介质。在另一种实施例中,该计算机程序也可以在网络介质上以信号的形式进行传输、分发,并通过通信部分609被下载和安装,和/或从可拆卸介质611被安装。该计算机程序包含的程序代码可以用任何适当的网络介质传输,包括但不限于:无线、有线等等,或者上述的任意合适的组合。

[0143] 在这样的实施例中,该计算机程序可以通过通信部分609从网络上被下载和安装,和/或从可拆卸介质611被安装。在该计算机程序被处理器601执行时,执行本公开实施例的系统中限定的上述功能。根据本公开的实施例,上文描述的系统、设备、装置、模块、单元等可以通过计算机程序模块来实现。

[0144] 根据本公开的实施例,可以以一种或多种程序设计语言的任意组合来编写用于执行本公开实施例提供的计算机程序的程序代码,具体地,可以利用高级过程和/或面向对象的编程语言、和/或汇编/机器语言来实施这些计算程序。程序设计语言包括但不限于诸如Java,C++,python,“C”语言或类似的程序设计语言。程序代码可以完全地在用户计算设备上执行、部分地在用户设备上执行、部分在远程计算设备上执行、或者完全在远程计算设备或服务器上执行。在涉及远程计算设备的情形中,远程计算设备可以通过任意种类的网络,包括局域网(LAN)或广域网(WAN),连接到用户计算设备,或者,可以连接到外部计算设备(例如利用因特网服务提供商来通过因特网连接)。

[0145] 附图中的流程图和框图,图示了按照本公开各种实施例的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上,流程图或框图中的每个方框可以代

表一个模块、程序段、或代码的一部分,上述模块、程序段、或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意,在有些作为替换的实现中,方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如,两个接连地表示的方框实际上可以基本并行地执行,它们有时也可以按相反的顺序执行,这依所涉及的功能而定。也要注意的,框图或流程图中的每个方框、以及框图或流程图中的方框的组合,可以用执行规定的功能或操作的专用的基于硬件的系统来实现,或者可以用专用硬件与计算机指令的组合来实现。

[0146] 本领域技术人员可以理解,本公开的各个实施例和/或权利要求中记载的特征可以进行多种组合和/或结合,即使这样的组合或结合没有明确记载于本公开中。特别地,在不脱离本公开精神和教导的情况下,本公开的各个实施例和/或权利要求中记载的特征可以进行多种组合和/或结合。所有这些组合和/或结合均落入本公开的范围。

[0147] 以上对本公开的实施例进行了描述。但是,这些实施例仅仅是为了说明的目的,而并非为了限制本公开的范围。尽管在以上分别描述了各实施例,但是这并不意味着各个实施例中的措施不能有利地结合使用。本公开的范围由所附权利要求及其等同物限定。不脱离本公开的范围,本领域技术人员可以做出多种替代和修改,这些替代和修改都应落在本公开的范围之内。

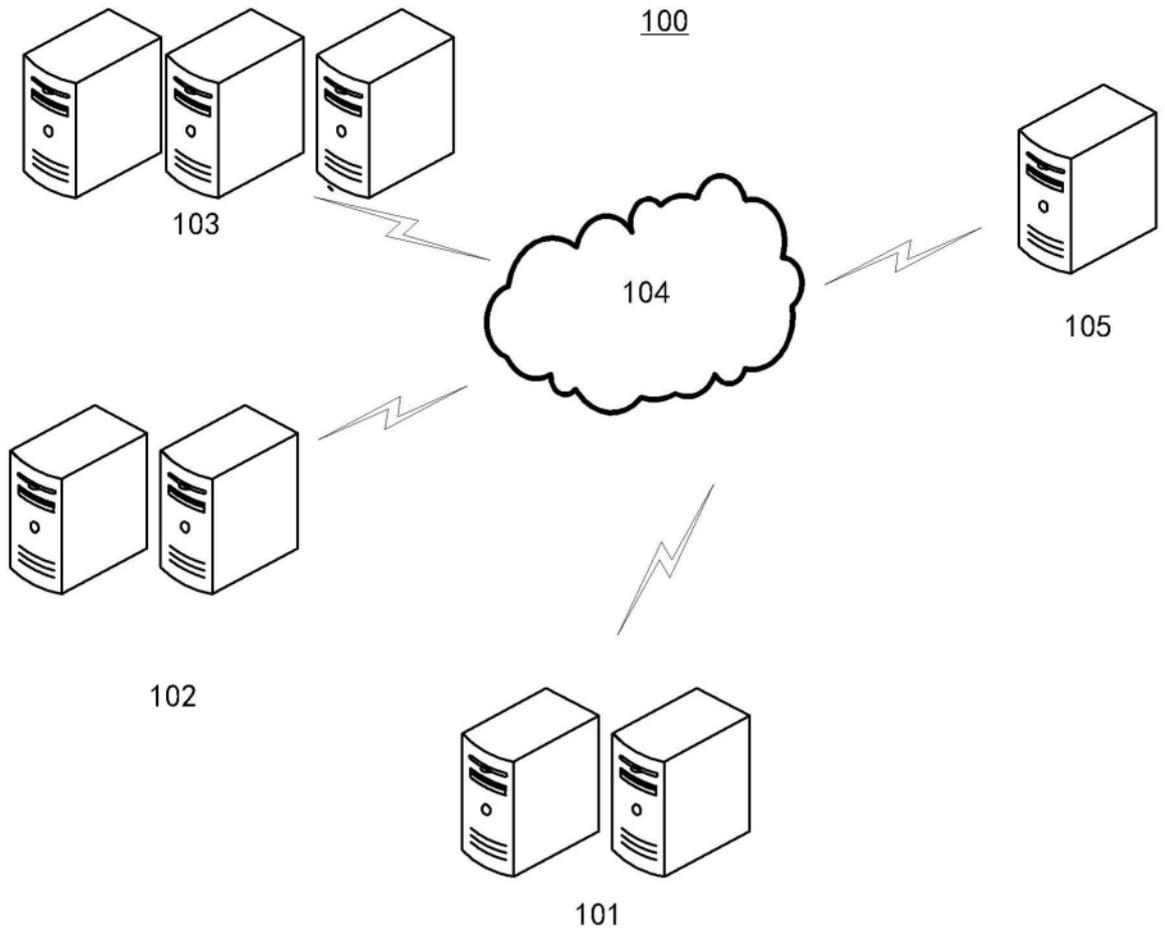


图1

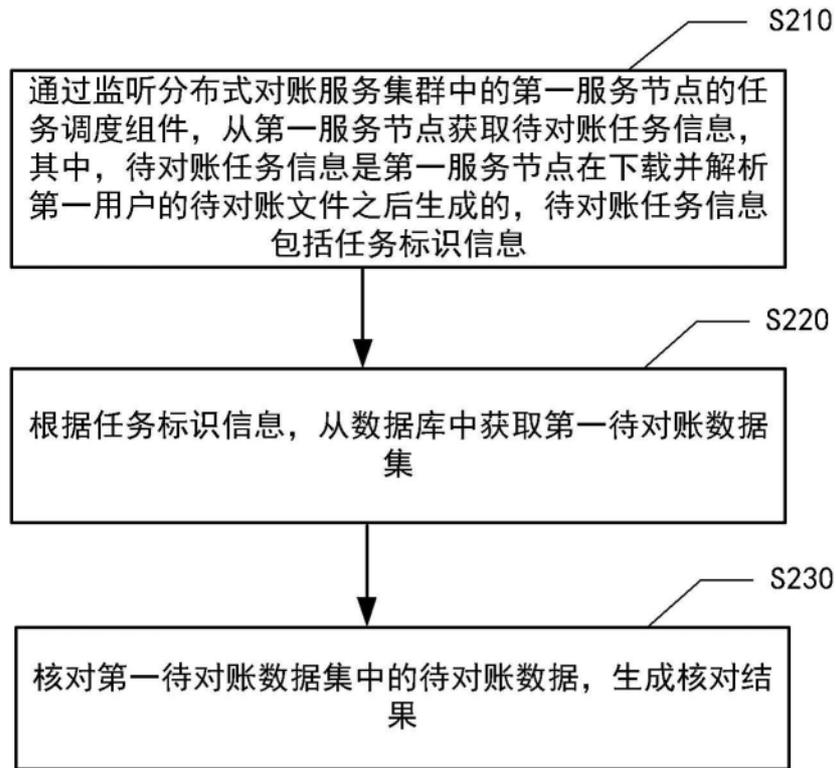


图2

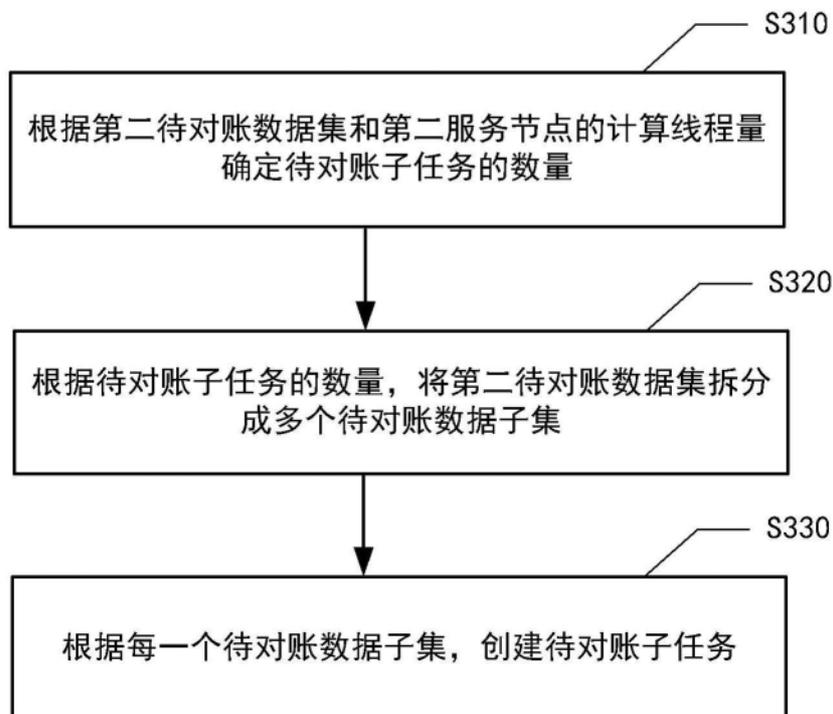


图3

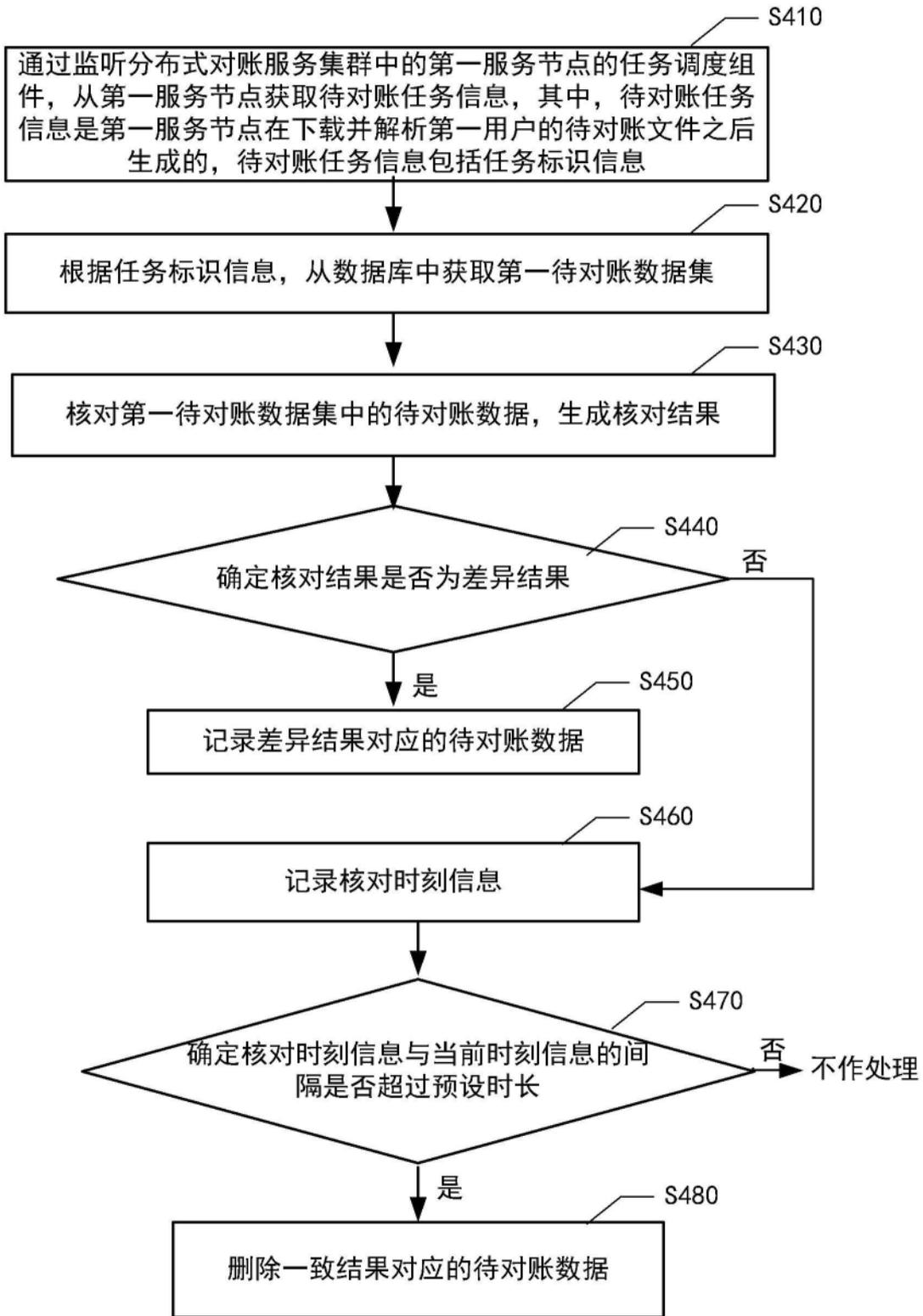


图4

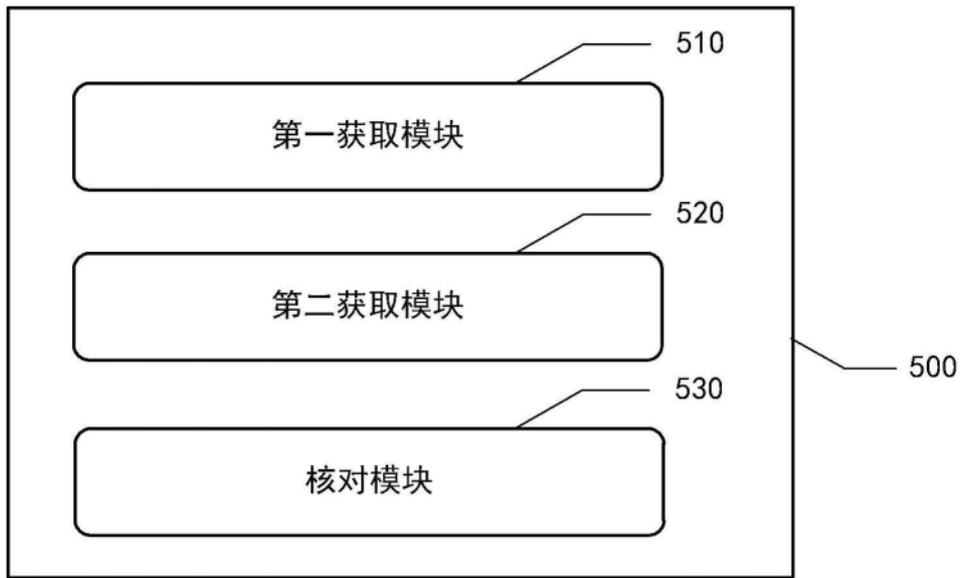


图5

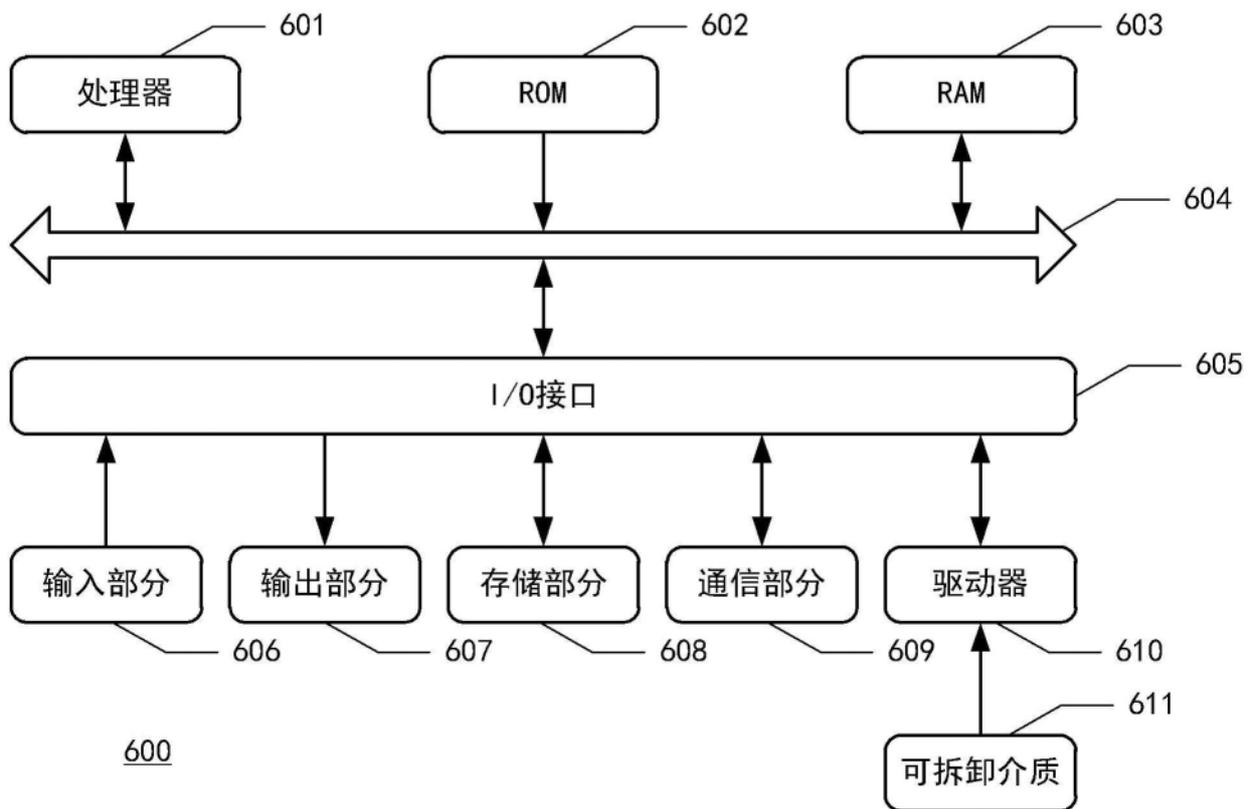


图6