

# (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



(10) 国际公布号

WO 2020/207014 A1

(43) 国际公布日

2020 年 10 月 15 日 (15.10.2020)

(51) 国际专利分类号:

G06F 11/36 (2006.01)

(21) 国际申请号:

PCT/CN2019/118587

(22) 国际申请日:

2019 年 11 月 14 日 (14.11.2019)

(25) 申请语言:

中文

(26) 公布语言:

中文

(30) 优先权:

201910279738.9 2019 年 4 月 9 日 (09.04.2019) CN

(71) 申请人: 平安科技 (深圳) 有限公司 (PING AN

TECHNOLOGY (SHENZHEN) CO., LTD.) [CN/CN];

中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

(72) 发明人: 陈珍妮 (CHEN, Zhenni); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

(74) 代理人: 深圳市隆天联鼎知识产权代理有限公司 (SHENZHEN LUNGTIN LIANDING INTELLECTUAL PROPERTY AGENT LTD.); 中国广东省深圳市福田区南园路上田大厦 4A 刘抗美, Guangdong 518000 (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, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX,

(54) Title: AUTOMATED REGRESSION TESTING METHOD AND APPARATUS FOR BIG DATA, TEST CENTER SERVER AND STORAGE MEDIUM

(54) 发明名称: 大数据自动回归测试方法、装置、测试中心服务器和存储介质

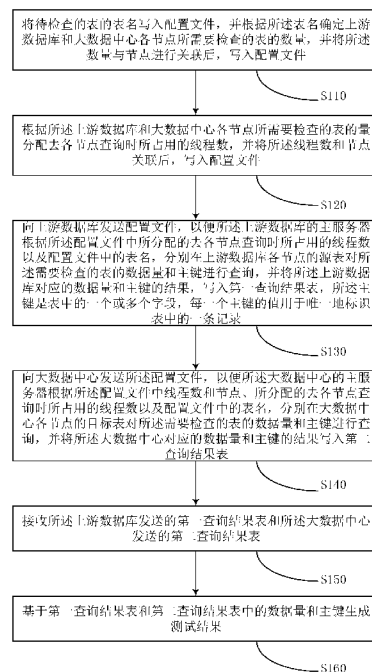


图 2

S110 WRITE THE NAMES OF TABLES TO BE CHECKED INTO A CONFIGURATION FILE, DETERMINE THE QUANTITY OF TABLES TO BE CHECKED AT ALL NODES OF THE UPSTREAM DATABASE AND THE BIG DATA CENTER, ASSOCIATE THE QUANTITY WITH THE NODES, AND WRITE INTO THE CONFIGURATION FILE  
S120 ACCORDING TO THE QUANTITY OF TABLES TO BE CHECKED AT ALL NODES OF THE UPSTREAM DATABASE AND THE BIG DATA CENTER, ASSIGN OCCUPIED THREAD COUNTS FOR QUERYING AT ALL NODES, ASSOCIATE THE THREAD COUNTS WITH THE NODES, AND WRITE INTO THE CONFIGURATION FILE  
S130 SEND THE CONFIGURATION FILE TO THE UPSTREAM DATABASE SO THAT THE MAIN SERVER OF THE UPSTREAM DATABASE RESPECTIVELY QUERIES THE DATA SIZE AND PRIMARY KEYS OF THE TABLES TO BE CHECKED IN SOURCE TABLES AT ALL NODES OF THE UPSTREAM DATABASE ACCORDING TO THE OCCUPIED THREAD COUNTS THAT ARE ASSIGNED FOR QUERYING AT ALL NODES AND THE TABLE NAMES IN THE CONFIGURATION FILE, AND WRITE THE RESULTS OF DATA SIZE AND PRIMARY KEYS CORRESPONDING TO THE UPSTREAM DATABASE INTO A FIRST QUERY RESULT TABLE, THE PRIMARY KEY REFERRING TO ONE OR MORE FIELDS IN A TABLE AND THE VALUE OF EACH PRIMARY KEY IS USED FOR UNIQUELY IDENTIFYING A RECORD IN THE TABLE  
S140 SEND THE CONFIGURATION FILE TO THE BIG DATA CENTER SO THAT THE MAIN SERVER OF THE BIG DATA CENTER RESPECTIVELY QUERIES THE DATA SIZE AND PRIMARY KEYS OF THE TABLES TO BE CHECKED IN TARGET TABLES AT ALL NODES OF THE BIG DATA CENTER ACCORDING TO THE THREAD COUNTS AND NODES IN THE CONFIGURATION FILE, THE OCCUPIED THREAD COUNTS THAT ARE ASSIGNED FOR QUERYING AT ALL NODES AND THE TABLE NAMES IN THE CONFIGURATION FILE, AND WRITE THE RESULTS OF DATA SIZE AND PRIMARY KEYS CORRESPONDING TO THE BIG DATA CENTER INTO A SECOND QUERY RESULT TABLE  
S150 RECEIVE THE FIRST QUERY RESULT TABLE SENT FROM THE UPSTREAM DATABASE AND THE SECOND QUERY RESULT TABLE SENT FROM THE BIG DATA CENTER  
S160 GENERATE A TEST RESULT ON THE BASIS OF THE FIRST QUERY RESULT TABLE AND THE SECOND QUERY RESULT TABLE

(57) Abstract: Disclosed are an automated regression testing method and apparatus for a big data storage center, a computer device and a storage medium, belonging to the technical field of regression testing. The main process of the automated regression testing method is as follows: first, the names of tables to be checked and the quantity of tables to be checked are associated with assigned nodes and written into a configuration file, and the assigned thread counts are associated with the nodes and written into the configuration file. The configuration file is then sent to an upstream database and the big data center, so that the upstream database and the big data center

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))。

---

extract data according to the settings in the configuration file, generate query result tables, and send the query result tables to a test center server. The test center server then compares the query result tables sent from the upstream database and the big data center to find differences between the two tables, and writes these differences into a check result table for feedback. The method improves testing efficiency, achieves full coverage of testing scenarios, and realizes real-time monitoring of data.

(57) 摘要: 一种大数据存储中心自动化回归测试方法、装置、计算机设备和存储介质, 属于回归测试技术领域, 所述大数据存储中心自动化回归测试方法主要流程有: 先将所需要检查的表的表名、所需要检查的表的数量与分配的节点关联写入配置文件, 并将分配的线程数和节点关联写入配置文件。然后将所述配置文件发送给上游数据库和大数据中心, 使上游数据库和大数据中心按照配置文件中的设置抽取数据, 生成查询结果表, 发送至测试中心服务器。然后测试中心服务器对比上游数据库和大数据中心分别发来的查询结果表, 找出两个表的不同, 并将这些不同写入检查结果表反馈。通过这种方法, 提高了测试效率, 实现了测试场景全覆盖, 且能实时监控数据。

# 发明名称：大数据自动回归测试方法、装置、测试中心服务器和存储介质

## 技术领域

[0001] 本申请要求2019年4月9日递交、申请名称为“大数据自动回归测试方法、装置、计算机设备和存储介质”的中国专利申请201910279738.9的优先权，在此通过引用将其全部内容合并于此。

[0002] 本申请涉及回归测试技术领域，尤其涉及一种大数据存储中心自动化回归测试方法、装置、计算机设备和存储介质。

## 背景技术

[0003] 由于大数据中心是数据存储中心，需要存储所有上游系统的业务数据，数据表存储的数量非常大。

[0004] 而现有技术，在进行大数据中心测试时，一般都是由相关测试人员通过手动的形式进行回归测试，而大数据存储中心涉及的数据量往往很大，这导致了测试人员在测试过程中工作量巨大，测试时间过长，效率较为低下。这也会导致回归测试只能抽检部分数据无法完全覆盖的所有数据，造成系统风险，有一定的风险隐患。

## 发明概述

## 技术问题

## 问题的解决方案

## 技术解决方案

[0005] 为了解决上述技术问题，本申请的一个目的在于提供一种大数据存储中心自动化回归测试方法及装置。

[0006] 其中，本申请所采用的技术方案为：

[0007] 一方面，一种大数据存储中心自动化回归测试方法，包括：将待检查的表的表名写入配置文件，并根据所述表名确定上游数据库和大数据中心各节点所需要

检查的表的数量，并将所述数量与节点进行关联后，写入配置文件；根据所述上游数据库和大数据中心各节点所需要检查的表的量分配去各节点查询时所占用的线程数，并将所述线程数和节点关联后，写入配置文件；向上游数据库发送配置文件，以便所述上游数据库的主服务器根据所述配置文件中所分配的去各节点查询时所占用的线程数以及配置文件中的表名，分别在上游数据库各节点的源表对所述需要检查的表的数据量和主键进行查询，并将所述上游数据库对应的数据量和主键的结果，写入第一查询结果表，所述主键是表中的一个或多个字段，每一个主键的值用于唯一地标识表中的一条记录；向大数据中心发送所述配置文件，以便所述大数据中心的主服务器根据所述配置文件中线程数和节点、所分配的去各节点查询时所占用的线程数以及配置文件中的表名，分别在大数据中心各节点的目标表对所述需要检查的表的数据量和主键进行查询，并将所述大数据中心对应的数据量和主键的结果写入第二查询结果表；接收所述上游数据库发送的第一查询结果表和所述大数据中心发送的第二查询结果表；基于第一查询结果表和第二查询结果表中的数据量和主键生成测试结果。

[0008] 另一方面，一种大数据存储中心自动化回归测试装置，包括：配置文件写入单元，用于将所需要检查的表的表名写入配置文件，并根据所述表名确定上游数据库和大数据中心各节点所需要检查的表的量，并将所述数量与节点关联写入配置文件；节点线程分配单元，用于根据所述上游数据库和大数据中心各节点所需要检查的表的数量分配去各节点查询时所占用的线程数，并将所述线程数和节点关联写入配置文件；上游数据查询单元，用于向上游数据库发送配置文件，以便所述上游数据库的主服务器根据所述配置文件中线程数和节点，按照所分配的线程数，根据配置文件中的表名，分别在上游数据库各节点的源表对所述需要检查的表的数据量和主键进行查询，并将所述数据量和主键的结果，写入第一查询结果表，所述主键是表中的一个或多个字段，它的值用于唯一地标识表中的某一条记录；下游数据查询单元，用于向大数据中心发送所述配置文件，以便所述大数据中心的主服务器根据所述配置文件中线程数和节点，以及，按照所分配的线程数，根据配置文件中的表名，分别在大数据中心各节点的目标表对所述需要检查的表的数据量和主键进行查询，并将所述数据量和主

键的结果写入第二查询结果表；查询结果接收单元，用于接收所述上游数据库发送的第一查询结果表和所述大数据中心发送的第二查询结果表；测试结果生成单元，用于基于第一查询结果表和第二查询结果表中的数据量和主键中不同的数量和主键，生成测试结果。

[0009] 另一方面，一种测试中心服务器，包括处理器及存储器，所述存储器上存储有计算机可读指令，所述计算机可读指令被所述处理器执行时实现如上所述的大数据存储中心自动化回归测试方法。

[0010] 另一方面，一种计算机可读存储介质，其上存储有计算机程序，所述计算机程序被处理器执行时实现如上所述的大数据存储中心自动化回归测试方法。

[0011] 本申请的实施例提供的技术方案可以包括以下有益效果：

[0012] 上述大数据存储中心自动化回归测试方法、装置、测试中心服务器和存储介质，通过将所需要检查的表的表名写入配置文件，然后根据所述表名确定上游数据库和大数据中心各节点所需要检查的表的数量，并将所述数量与节点关联写入配置文件。然后根据所述上游数据库和大数据中心各节点所需要检查的表的数量分配去各节点查询时所占用的线程数，并将所述线程数和节点关联写入配置文件，这样配置文件中就包含有要检查的表的表名、数量以及数量与节点的关系、线程数与节点的关系。然后向上游数据库发送配置文件，这样，所述上游数据库的主服务器就可以根据所述配置文件中线程数和节点，按照所分配的线程数，根据配置文件中的表名，分别在上游数据库各节点的源表对所述需要检查的表的数据量和主键进行查询，并将所述数据量和主键的结果，写入第一查询结果表，所述主键是表中的一个或多个字段，它的值用于唯一地标识表中的某一条记录，通过主键能够很快的查询需要检查的表，提高了测试的效率的同时还能覆盖所有要检查的内容。同时，也要向大数据中心这一下游数据库发送所述配置文件，同理所述大数据中心的主服务器根据所述配置文件中线程数和节点，以及，按照所分配的线程数，根据配置文件中的表名，分别在大数据中心各节点的目标表对所述需要检查的表的数据量和主键进行查询，并将所述数据量和主键的结果写入第二查询结果表。获取所述上游数据库发送的第一查询结果表和所述大数据中心发送的第二查询结果表后，就可以根据第一查询结

果表和第二查询结果表中的数据量和主键中不同的数量和主键，找出数据的差异，生成测试结果，这样就自动完成了回归测试，既避免了人工测试中的低效，也覆盖了所有的测试内容。

[0013] 应当理解的是，以上的一般描述和后文的细节描述仅是示例性和解释性的，并不能限制本申请。

发明的有益效果

对附图的简要说明

附图说明

[0014] 此处的附图被并入说明书中并构成本说明书的一部分，示出了符合本申请的实施例，并于说明书一起用于解释本申请的原理。

[0015] 图1是一实施例中提供的大数据存储中心自动化回归测试方法的实施环境图。

[0016] 图2是根据一示例性实施例示出的一种大数据存储中心自动化回归测试方法的流程图。

[0017] 图3是根据图2对应实施例示出的大数据存储中心自动化回归测试方法中步骤S160的一种具体实现流程图。

[0018] 图4是根据图2对应实施例示出的另一种大数据存储中心自动化回归测试方法的流程图。

[0019] 图5是根据图2对应实施例示出的又一种大数据存储中心自动化回归测试方法的流程图。

[0020] 图6是根据一示例性实施例示出的一种大数据存储中心自动化回归测试装置的框图。

[0021] 图7示意性示出一种用于实现上述大数据存储中心自动化回归测试方法的测试中心服务器示例框图。

[0022] 图8示意性示出一种用于实现上述大数据存储中心自动化回归测试方法的计算机可读存储介质。

[0023] 通过上述附图，已示出本申请明确的实施例，后文中将有更详细的描述，这些附图和文字描述并不是为了通过任何方式限制本申请构思的范围，而是通过参考特定实施例为本领域技术人员说明本申请的概念。

## 发明实施例

### 本发明的实施方式

[0024] 这里将详细地对示例性实施例执行说明，其示例表示在附图中。下面的描述涉及附图时，除非另有表示，不同附图中的相同数字表示相同或相似的要素。以下示例性实施例中所描述的实施方式并不代表与本发明相一致的所有实施方式。相反，它们仅是与如所附权利要求书中所详述的、本发明的一些方面相一致的装置和方法的例子。

[0025] 图1为一个实施例中提供的大数据存储中心自动化回归测试方法的实施环境图，如图1所示，在所述实施环境中，包括测试中心服务器100、上游数据库200、大数据中心300。上游数据库200是存储有上游数据的服务器，大数据中心300是存储有下游数据的大数据服务器，包含有多个节点，测试中心服务器100用于从上游数据库200和大数据中心300抽取数据并执行测试。具体过程为，先将所需要检查的表的表名、所需要检查的表的数量与分配的节点关联写入配置文件，并将分配的线程数和节点关联写入配置文件。然后将所述配置文件发送给上游数据库200和大数据中心300，使上游数据库200和大数据中心300按照配置文件中的设置抽取数据，生成查询结果表，发送至测试中心服务器100。然后测试中心服务器100对比上游数据库200和大数据中心300分别发来的查询结果表，找出两个表的不同，并将这些不同写入检查结果表反馈。需要说明的是测试中心服务器100可为智能手机、平板电脑、笔记本电脑、台式计算机等，但并不局限于此。上游数据库200、大数据中心300可以是多个不同服务器组成的服务器集群，但并不局限于此。中心服务器100、上游数据库200、大数据中心300之间可以通过蓝牙、USB(Universal Serial Bus，通用串行总线)或者其他通讯连接方式进行连接，本发明在此不做限制。

[0026] 如图2所示，在一个实施例中，提出了一种大数据存储中心自动化回归测试方法，所述大数据存储中心自动化回归测试方法可以应用于上述的测试中心服务器100中，具体可以包括以下步骤：步骤S110，将待检查的表的表名写入配置文件，并根据所述表名确定上游数据库和大数据中心各节点所需要检查的表的数量，并将所述数量与节点进行关联后，写入配置文件。本发明要进行的测试为

回归测试，本申请主要的目的在于提高回归测试的效率，使测试覆盖大多数场景。故在步骤S110中，本申请先配置一个配置文件，将待检查的表的表名信息写入所述配置文件，一个表的表名信息是其唯一的标识信息，需要检查的表的表名信息可以使上游数据库和大数据中心根据表名信息精确定位到所需要检查的表。然后由于一个表名唯一对应一张表，故再根据所述表名确定上游数据库和大数据中心各节点所需要检查的表的数量（即通过统计表名的数量来确定表的数量），并将所述数量与节点关联写入配置文件。以便上游数据库和大数据中心各节点根据各自所需要检查的表的数量，统计各自所需要检查的表的量。所述所需要检查的表的量可以是需要检查的表的数量，也可以是需要检查的表所占用存储空间的大小，还可以是需要检查的表的数量和需要检查的表所占用存储空间的大小的加权和。而之所以将上游数据库和大数据中心各节点所需要检查的表的数量与节点关联写入配置文件而不是所需要检查的表的量，是因为如果要统计所需要检查的表的量就有可能需要向上游数据库和大数据中心各节点查询，这个工作量过于庞杂，而所需要检查的表的数量可以直接根据表名统计，每一个表的表名对应一张表，其工作量大大减少，且所需要检查的表的量为数量时，可以直接将配置文件中的数量作为需要检查的表的量。步骤S120，根据所述上游数据库和大数据中心各节点所需要检查的表的量分配去各节点查询时所占用的线程数，并将所述线程数和节点关联后，写入配置文件。将待检查的表的表名以及上游数据库和大数据中心各节点所需要检查的表的数量均写入配置文件后，就可以根据所述上游数据库和大数据中心各节点所需要检查的表的量分配去各节点查询时所占用的线程数，对于所需要检查的表的量较多的节点，分配较多的线程数，对于所需要检查的表的量较少的节点，分配较少的线程数。这样分配所占用的线程数，在按照分配的线程数查询的时候，可以使查询过程更稳定，不会出现由于某一节点的大文件将所有线程占用而导致所述节点工作负载过大而查询失败的情况，也不会出现由于某一节点的大文件将所有线程占用而导致其它节点断连过久而查询失败的情况，也不会出现因为某一节点卡死而导致查询失败的情况。使得查询过程的工作负载分配得更为均衡更为合理，又好又快地完成查询过程。步骤S130，向上游数据库发送配置文件



，以便所述上游数据库的主服务器根据所述配置文件中所分配的去各节点查询时所占用的线程数以及配置文件中的表名，分别在上游数据库各节点的源表对所述需要检查的表的数据量和主键进行查询，并将所述上游数据库对应的数据量和主键的结果，写入第一查询结果表，所述主键是表中的一个或多个字段，每一个主键的值用于唯一地标识表中的一条记录。步骤S140，向大数据中心发送所述配置文件，以便所述大数据中心的主服务器根据所述配置文件中线程数和节点、所分配的去各节点查询时所占用的线程数以及配置文件中的表名，分别在大数据中心各节点的目标表对所述需要检查的表的数据量和主键进行查询，并将所述大数据中心对应的数据量和主键的结果写入第二查询结果表。将去各节点查询时所占用的线程数和节点关联写入配置文件后，所述配置文件中包含的信息已满足进行查询任务的需求，就可以发送给上游数据库和大数据中心，以便上游数据库和大数据中心根据配置文件中包含的信息进行查询任务。在进行查询任务时主要对所述需要检查的表的数据量和主键进行查询。主键，即主关键字，是表中的一个或多个字段，它的值用于唯一地标识表中的一条记录。主关键字是一种唯一关键字，一个表的主键可以由多个关键字共同组成，并且主关键字的列不能包含空值。查询所述需要检查的表的主键相比查询表的内容，能够在表示表的内容的同时，可以加快数据库的查询速度，大幅度缩减查询结果的大小。最后所述数据量和主键的结果分别写入一张查询结果表，即得到第一查询结果表和第二查询结果表。步骤S150，接收所述上游数据库发送的第一查询结果表和所述大数据中心发送的第二查询结果表。得到第一查询结果表和第二查询结果表后，上游数据库200和大数据中心300就可以将所述第一查询结果表和第二查询结果表发送给测试中心服务器100，由测试中心服务器100进行下一步操作。在发送第一查询结果表和第二查询结果表时，可以按照第一查询结果表和第二查询结果表所占用存储空间的大小分配所占用的线程数。也可以将第一查询结果表和第二查询结果表分成若干部分，按照各部分所占用存储空间的大小来分配所占用的线程数。步骤S160，基于第一查询结果表和第二查询结果表中的数据量和主键生成测试结果。最后对比第一查询结果表和第二查询结果表的数据量和主键，得出两个表中不同的数据量和主键。若第一查询

结果表和第二查询结果表的数据量不同，则存在有漏传的情况；若第一查询结果表和第二查询结果表存在有不同的主键，则存在有误传或者漏传的情况，这些情况都需要写入测试结果中。以便根据测试结果进行一些后续的处理，例如重新传输或者进行系统维护。

[0027] 本申请要进行的测试为回归测试，具体来说，就是检测上游数据库向大数据中心传输的数据是否有漏传或者误传。故本申请事先配置一个配置文件，配置文件中包括需要检查的表的表名信息，以便上游数据库和大数据中心根据表名信息查询所需要检查的表的数据量和主键。之所以检查数据量是为了检查有没有漏传，而检查主关键字是为了检查有没有误传。由于本申请涉及的是大数据集群间的测试，数据量会比较大，故需要合理安排测试中各大数据集群节点的传输的速度，故本申请中的配置文件中还包括了事先统计出来的各节点的任务量，以及按照任务量分配的线程数，以保证在数据测试过程中不会断线。可选的，图3是根据图2对应实施例示出的大数据存储中心自动化回归测试方法中步骤S160的细节描述，所述大数据存储中心自动化回归测试方法中，步骤S160可以包括以下步骤：步骤S161，将所述第一查询结果表中的数据量和主键和第二查询结果表中对应的数据量和主键输入机器学习模型，获得每一对数据和主键的测试结果。在对比第一查询结果表和第二查询结果表的数据量和主键时可以通过机器学习模型来执行，这样会使得输出的对比结果有较高的准确率而且随着机器学习模型的深度学习，这个结果会越来越准确。步骤S162，将所述测试结果写入测试结果表。其中，所述机器学习模型的训练过程为：将正查询结果样本对输入机器学习模型，所述正查询结果样本对是两个数据量样本或者两个主键样本，其中一个数据量样本是另一个数据量样本的相同对象，或者其中一个主键样本是另一个主键样本的相同对象，若输出的判定结果不相同，则调整所述机器学习模型的系数；将负查询结果样本输入机器学习模型，所述负查询结果样本对是两个数据量样本或者两个主键样本，其中一个数据量样本是另一个数据量样本的不同对象，或者其中一个主键样本是另一个主键样本的不同对象，若输出的判定结果为相同，则调整所述机器学习模型的系数。本方案通过机器学习模型来对两个查询结果表中数据量和主键信息进行比对，然后将比对信息

写入检查结果表的方式来检测查询结果，找出失败的场景。由于机器学习模型经过训练，在对比测试过程中输出结果会越来越准，提高了测试的准确率。所述机器学习模型的学习的方式为：在外界输入样本的刺激下不断改变网络的连接权值。学习的本质是对各连接权重进行动态调整。由于每个样本对的两个样本是否相同是已知的，一旦机器学习模型根据各连接权重，对从各样本对施加运算，得到的字段名称和参数要求结果与已知结果不符，其就自动调整各连接权重，直到得到的字段名称和参数要求结果与已知结果相符。这样，就训练好了机器学习模型。然后将第一查询结果表中的的查询结果和第二查询结果表中对应的查询结果输入机器学习模型，就得到了该对查询结果的测试结果。

[0028] 图4示出了在一个实施例中，图2对应实施例中的步骤S160之后，所述大数据存储中心自动化回归测试方法还可以包括以下步骤：步骤S170，将所述测试结果发送给用户，以使用户根据所述测试结果对所述不相同的数据量和主键进行人工验证。当得到测试结果后，还需要把所述测试结果传递到相关测试维护人员（即用户）手中，以方便相关人员根据测试结果中记录的不相同的数据量和主键进行人工验证，找出漏传和误传的数据表，根据漏传和误传的数据表分析失败场景，然后就失败场景进行修复维护，这样就可以修正错误，使得上游数据库和大数据中心之间的数据保持一致，也完成了此次测试的最终目的。

[0029] 图5示出了在一个实施例中，图2对应实施例中的步骤S160之后，所述大数据存储中心自动化回归测试方法还可以包括以下步骤：步骤S180，将所述测试结果发送至上游数据库，以便所述上游数据库根据所述测试结果将所述不相同的数据量和主键重新发送至大数据中心。当得到测试结果后，也可以不用相关人员根据测试结果中记录的不相同的数据量和主键进行人工验证，而是系统自己修复，本方案是将所述测试结果发送至上游数据库，然后上游数据库根据所述测试结果将所述不相同的数据量和主键重新再一次发送至大数据中心，填补覆盖掉之前的错误数据，这样就完成了自我纠正，免去了人工验证维护的繁杂的操作量，提高了测试维护的效率。可选的，图2对应实施例示出的大数据存储中心自动化回归测试方法中所述上游数据库和大数据中心各节点所需要检查的表的量是所需要检查的表的数量。可选的，图2对应实施例示出的大数据存储中心自

动化回归测试方法中所述上游数据库和大数据中心各节点所需要检查的表的量是所需要检查的表所占用存储空间的大小。可选的，图2对应实施例示出的大数据存储中心自动化回归测试方法中所述上游数据库和大数据中心各节点所需要检查的表的量是所需要检查的表的数量和需要检查的表所占用存储空间的大小的加权和。表的数量和表所占用存储空间都能在一定程度上反映出上游数据库和大数据中心查询数据的任务量的大小，而将表的数量和表所占用存储空间这两项按照一定的权重比求加权和，能够覆盖更多的维度，更客观得表示上游数据库和大数据中心查询数据的任务量的大小。

[0030] 如图6所示，在一个实施例中，提供了一种大数据存储中心自动化回归测试装置，所述大数据存储中心自动化回归测试装置可以集成于上述的测试中心服务器100中，具体可以包括：配置文件写入单元110、节点线程分配单元120、上游数据查询单元130、下游数据查询单元140、查询结果接收单元150、测试结果生成单元160。配置文件写入单元110，用于将所需要检查的表的表名写入配置文件，并根据所述表名确定上游数据库和大数据中心各节点所需要检查的表的数量，并将所述数量与节点关联写入配置文件；节点线程分配单元120，用于根据所述上游数据库和大数据中心各节点所需要检查的表的量分配去各节点查询时所占用的线程数，并将所述线程数和节点关联写入配置文件；上游数据查询单元130，用于向上游数据库发送配置文件，以便所述上游数据库的主服务器根据所述配置文件中线程数和节点，按照所分配的线程数，根据配置文件中的表名，分别在上游数据库各节点的源表对所述需要检查的表的数据量和主键进行查询，并将所述数据量和主键的结果，写入第一查询结果表，所述主键是表中的一个或多个字段，它的值用于唯一地标识表中的某一条记录；下游数据查询单元140，用于向大数据中心发送所述配置文件，以便所述大数据中心的主服务器根据所述配置文件中线程数和节点，以及，按照所分配的线程数，根据配置文件中的表名，分别在大数据中心各节点的目标表对所述需要检查的表的数据量和主键进行查询，并将所述数据量和主键的结果写入第二查询结果表；查询结果接收单元150，用于接收所述上游数据库发送的第一查询结果表和所述大数据中心发送的第二查询结果表；测试结果生成单元160，用于基于第一查询结果表

和第二查询结果表中的数据量和主键中不同的数量和主键，生成测试结果。上述装置中各个模块的功能和作用的实现过程具体详见上述大数据存储中心自动化回归测试方法中对应步骤的实现过程，在此不再赘述。应当注意，尽管在上文详细描述中提及了用于动作执行的设备的若干模块或者单元，但是这种划分并非强制性的。实际上，根据本申请的实施方式，上文描述的两个或更多模块或者单元的特征和功能可以在一个模块或者单元中具体化。反之，上文描述的一个模块或者单元的特征和功能可以进一步划分为由多个模块或者单元来具体化。此外，尽管在附图中以特定顺序描述了本申请中方法的各个步骤，但是，这并非要求或者暗示必须按照该特定顺序来执行这些步骤，或是必须执行全部所示的步骤才能实现期望的结果。附加的或备选的，可以省略某些步骤，将多个步骤合并为一个步骤执行，以及/或者将一个步骤分解为多个步骤执行等。通过以上的实施方式的描述，本领域的技术人员易于理解，这里描述的示例实施方式可以通过软件实现，也可以通过软件结合必要的硬件的方式来实现。因此，根据本申请实施方式的技术方案可以以软件产品的形式体现出来，该软件产品可以存储在一个非易失性存储介质（可以是CD-ROM，U盘，移动硬盘等）中或网络上，包括若干指令以使得一台计算设备（可以是个人计算机、服务器、移动终端、或者网络设备等）执行根据本申请实施方式的方法。

[0031] 在本申请的示例性实施例中，还提供了一种能够实现上述方法的测试中心服务器。所属技术领域的技术人员能够理解，本申请的各个方面可以实现为系统、方法或程序产品。因此，本申请的各个方面可以具体实现为以下形式，即：完全的硬件实施方式、完全的软件实施方式（包括固件、微代码等），或硬件和软件方面结合的实施方式，这里可以统称为“电路”、“模块”或“系统”。下面参照图7来描述根据本申请的这种实施方式的测试中心服务器500。图7显示的测试中心服务器500仅仅是一个示例，不应对本申请实施例的功能和使用范围带来任何限制。如图7所示，测试中心服务器500以通用计算设备的形式表现。测试中心服务器500的组件可以包括但不限于：上述至少一个处理单元510、上述至少一个存储单元520、连接不同系统组件（包括存储单元520和处理单元510）的总线530。其中，所述存储单元存储有程序代码，所述程序代码可以被所述处理单元

510执行，使得所述处理单元510执行本说明书上述“示例性方法”部分中描述的根据本申请各种示例性实施方式的步骤。例如，所述处理单元510可以执行如图2中所示的步骤S110，将所需要检查的表的表名写入配置文件，并根据所述表名确定上游数据库和大数据中心各节点所需要检查的表的数量，并将所述数量与节点关联写入配置文件；步骤S120，根据所述上游数据库和大数据中心各节点所需要检查的表的数量分配去各节点查询时所占用的线程数，并将所述线程数和节点关联写入配置文件；步骤S130，向上游数据库发送配置文件，以便所述上游数据库的主服务器根据所述配置文件中线程数和节点，按照所分配的线程数，根据配置文件中的表名，分别在上游数据库各节点的源表对所述需要检查的表的数据量和主键进行查询，并将所述数据量和主键的结果，写入第一查询结果表，所述主键是表中的一个或多个字段，它的值用于唯一地标识表中的某一条记录；步骤S140，向大数据中心发送所述配置文件，以便所述大数据中心的主服务器根据所述配置文件中线程数和节点，以及，按照所分配的线程数，根据配置文件中的表名，分别在大数据中心各节点的目标表对所述需要检查的表的数据量和主键进行查询，并将所述数据量和主键的结果写入第二查询结果表；步骤S150，接收所述上游数据库发送的第一查询结果表和所述大数据中心发送的第二查询结果表；步骤S160，基于第一查询结果表和第二查询结果表中的数据量和主键中不同的数量和主键，生成测试结果。存储单元520可以包括易失性存储单元形式的可读介质，例如随机存取存储单元（RAM）5201和/或高速缓存存储单元5202，还可以进一步包括只读存储单元（ROM）5203。存储单元520还可以包括具有一组（至少一个）程序模块5205的程序/实用工具5204，这样的程序模块5205包括但不限于：操作系统、一个或者多个应用程序、其它程序模块以及程序数据，这些示例中的每一个或某种组合中可能包括网络环境的实现。总线530可以为表示几类总线结构中的一种或多种，包括存储单元总线或者存储单元控制器、外围总线、图形加速端口、处理单元或者使用多种总线结构中的任意总线结构的局域总线。测试中心服务器500也可以与一个或多个外部设备700（例如键盘、指向设备、蓝牙设备等）通信，还可与一个或者多个使得用户能与该测试中心服务器500交互的设备通信，和/或与使得该测试中心服务器50

0能与一个或多个其它计算设备进行通信的任何设备（例如路由器、调制解调器等等）通信。这种通信可以通过输入/输出（I/O）接口550进行。并且，测试中心服务器500还可以通过网络适配器560与一个或者多个网络（例如局域网（LAN），广域网（WAN）和/或公共网络，例如因特网）通信。如图所示，网络适配器560通过总线530与测试中心服务器500的其它模块通信。应当明白，尽管图中未示出，可以结合测试中心服务器500使用其它硬件和/或软件模块，包括但不限于：微代码、设备驱动器、冗余处理单元、外部磁盘驱动阵列、RAID系统、磁带驱动器以及数据备份存储系统等。通过以上的实施方式的描述，本领域的技术人员易于理解，这里描述的示例实施方式可以通过软件实现，也可以通过软件结合必要的硬件的方式来实现。因此，根据本申请实施方式的技术方案可以以软件产品的形式体现出来，该软件产品可以存储在一个非易失性存储介质（可以是CD-ROM，U盘，移动硬盘等）中或网络上，包括若干指令以使得一台计算设备（可以是个人计算机、服务器、终端装置、或者网络设备等）执行根据本申请实施方式的方法。

[0032] 在本申请的示例性实施例中，还提供了一种计算机可读存储介质，其上存储有能够实现本说明书上述方法的程序产品。在一些可能的实施方式中，本申请的各个方面还可以实现为一种程序产品的形式，其包括程序代码，当所述程序产品在终端设备上运行时，所述程序代码用于使所述终端设备执行本说明书上述“示例性方法”部分中描述的根据本申请各种示例性实施方式的步骤。参考图8所示，描述了根据本申请的实施方式的用于实现上述方法的计算机可读存储介质600，其可以采用便携式紧凑盘只读存储器(CD-ROM)并包括程序代码，并可以在终端设备，例如个人电脑上运行。然而，本申请的程序产品不限于此，在本文件中，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。上述计算机可读存储介质可以为非易失性可读存储介质，例如存储于CD-ROM，U盘，移动硬盘设备中的非易失性可读存储介质，可以包括若干指令以使得一台计算设备（可以是个人计算机、服务器、终端装置、或者网络设备等）执行根据本申请实施方式的方法。所述程序产品可以采用一个或多个可读介质的任意组合。可读介质可以是可

读信号介质或者可读存储介质。可读存储介质例如可以为但不限于电、磁、光、电磁、红外线、或半导体的系统、装置或器件，或者任意以上的组合。可读存储介质的更具体的例子（非穷举的列表）包括：具有一个或多个导线的电连接、便携式盘、硬盘、随机存取存储器（RAM）、只读存储器（ROM）、可擦式可编程只读存储器（EPROM或闪存）、光纤、便携式紧凑盘只读存储器（CD-ROM）、光存储器件、磁存储器件、或者上述的任意合适的组合。计算机可读信号介质可以包括在基带中或者作为载波一部分传播的数据信号，其中承载了可读程序代码。这种传播的数据信号可以采用多种形式，包括但不限于电磁信号、光信号或上述的任意合适的组合。可读信号介质还可以是可读存储介质以外的任何可读介质，该可读介质可以发送、传播或者传输用于由指令执行系统、装置或者器件使用或者与其结合使用的程序。可读介质上包含的程序代码可以用任何适当的介质传输，包括但不限于无线、有线、光缆、RF等等，或者上述的任意合适的组合。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申请操作的程序代码，所述程序设计语言包括面向对象的程序设计语言—诸如Java、C++等，还包括常规的过程式程序设计语言—诸如“C”语言或类似的设计语言。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、作为一个独立的软件包执行、部分在用户计算设备上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）。此外，上述附图仅是根据本申请示例性实施例的方法所包括的处理的示意性说明，而不是限制目的。易于理解，上述附图所示的处理并不表明或限制这些处理的时间顺序。另外，也易于理解，这些处理可以是例如在多个模块中同步或异步执行的。上述内容，仅为本申请的较佳示例性实施例，并非用于限制本申请的实施方案，本领域普通技术人员根据本申请的主要构思和精神，可以十分方便地进行相应的变通或修改，故本申请的保护范围应以权利要求书所要求的保护范围为准。



## 权利要求书

### [权利要求 1]

一种大数据存储中心自动化回归测试方法，包括：

将待检查的表的表名写入配置文件，并根据所述表名确定上游数据库和大数据中心各节点所需要检查的表的数量，并将所述数量与节点进行关联后，写入配置文件；

根据所述上游数据库和大数据中心各节点所需要检查的表的量分配去各节点查询时所占用的线程数，并将所述线程数和节点关联后，写入配置文件；

向上游数据库发送配置文件，以便所述上游数据库的主服务器根据所述配置文件中所分配的去各节点查询时所占用的线程数以及配置文件中的表名，分别在上游数据库各节点的源表对所述需要检查的表的数据量和主键进行查询，并将所述上游数据库对应的数据量和主键的结果，写入第一查询结果表，所述主键是表中的一个或多个字段，每一个主键的值用于唯一地标识表中的一条记录；

向大数据中心发送所述配置文件，以便所述大数据中心的主服务器根据所述配置文件中线程数和节点、所分配的去各节点查询时所占用的线程数以及配置文件中的表名，分别在大数据中心各节点的目标表对所述需要检查的表的数据量和主键进行查询，并将所述大数据中心对应的数据量和主键的结果写入第二查询结果表；

接收所述上游数据库发送的第一查询结果表和所述大数据中心发送的第二查询结果表；

基于第一查询结果表和第二查询结果表中的数据量和主键生成测试结果。

### [权利要求 2]

如权利要求1所述的方法，其中，所述基于第一查询结果表和第二查询结果表中的数据量和主键生成测试结果的步骤包括：

将所述第一查询结果表中的数据量和主键和第二查询结果表中对应的数据量和主键输入机器学习模型，获得每一对数据和主键的测试结果；

将所述测试结果写入测试结果表；

其中，所述机器学习模型的训练过程为：

将正查询结果样本对输入机器学习模型，所述正查询结果样本对是两个数据量样本或者两个主键样本，其中一个数据量样本是另一个数据量样本的相同对象，或者其中一个主键样本是另一个主键样本的相同对象，若输出的判定结果不相同，则调整所述机器学习模型的系数；  
将负查询结果样本输入机器学习模型，所述负查询结果样本对是两个数据量样本或者两个主键样本，其中一个数据量样本是另一个数据量样本的不同对象，或者其中一个主键样本是另一个主键样本的不同对象，若输出的判定结果为相同，则调整所述机器学习模型的系数。

[权利要求 3] 如权利要求1所述的方法，其中，所述测试结果表包括不相同的数据量和主键，在对比第一查询结果表和第二查询结果表中的数据量和主键，并生成测试结果之后，还包括：

将所述测试结果发送给用户，以使用户根据所述测试结果对所述不相同的数据量和主键进行人工验证。

[权利要求 4] 如权利要求1所述的方法，其中，所述测试结果表包括不相同的数据量和主键，在对比第一查询结果表和第二查询结果表中的数据量和主键，并生成测试结果之后，还包括：

将所述测试结果发送至上游数据库，以便所述上游数据库根据所述测试结果将所述不相同的数据量和主键重新发送至大数据中心。

[权利要求 5] 如权利要求1所述的方法，其中，所述上游数据库和大数据中心各节点所需要检查的表的量是所需要检查的表的数量。

[权利要求 6] 如权利要求1所述的方法，其中，所述上游数据库和大数据中心各节点所需要检查的表的量是所需要检查的表所占用存储空间的大小。

[权利要求 7] 如权利要求1所述的方法，其中，所述上游数据库和大数据中心各节点所需要检查的表的量是所需要检查的表的数量和需要检查的表所占用存储空间的大小的加权和。

[权利要求 8] 一种大数据存储中心自动化回归测试装置，包括：

配置文件写入单元，用于将待检查的表的表名写入配置文件，并根据所述表名确定上游数据库和大数据中心各节点所需要检查的表的数量，并将所述数量与节点进行关联后，写入配置文件；

节点线程分配单元，用于根据所述上游数据库和大数据中心各节点所需要检查的表的量分配各节点查询时所占用的线程数，并将所述线程数和节点关联后，写入配置文件；

上游数据查询单元，用于向上游数据库发送配置文件，以便所述上游数据库的主服务器根据所述配置文件中所分配的去各节点查询时所占用的线程数以及配置文件中的表名，分别在上游数据库各节点的源表对所述需要检查的表的数据量和主键进行查询，并将所述上游数据库对应的数据量和主键的结果，写入第一查询结果表，所述主键是表中的一个或多个字段，每一个主键的值用于唯一地标识表中的一条记录；

下游数据查询单元，用于向大数据中心发送所述配置文件，以便所述大数据中心的主服务器根据所述配置文件中线程数和节点、所分配的去各节点查询时所占用的线程数以及配置文件中的表名，分别在大数据中心各节点的目标表对所述需要检查的表的数据量和主键进行查询，并将所述大数据中心对应的数据量和主键的结果写入第二查询结果表；

查询结果接收单元，用于接收所述上游数据库发送的第一查询结果表和所述大数据中心发送的第二查询结果表；

测试结果生成单元，用于基于第一查询结果表和第二查询结果表中的数据量和主键生成测试结果。

[权利要求 9]

如权利要求8所述的装置，其中，所述测试结果生成单元被配置为：

将所述第一查询结果表中的数据量和主键和第二查询结果表中对应的数据量和主键输入机器学习模型，获得每一对数据和主键的测试结果；

将所述测试结果写入测试结果表；

其中，所述机器学习模型的训练过程为：

将正查询结果样本对输入机器学习模型，所述正查询结果样本对是两个数据量样本或者两个主键样本，其中一个数据量样本是另一个数据量样本的相同对象，或者其中一个主键样本是另一个主键样本的相同对象，若输出的判定结果不相同，则调整所述机器学习模型的系数；  
将负查询结果样本输入机器学习模型，所述负查询结果样本对是两个数据量样本或者两个主键样本，其中一个数据量样本是另一个数据量样本的不同对象，或者其中一个主键样本是另一个主键样本的不同对象，若输出的判定结果为相同，则调整所述机器学习模型的系数。

[权利要求 10] 如权利要求8所述的装置，其中，所述测试结果表包括不相同的数据量和主键，在对比第一查询结果表和第二查询结果表中的数据量和主键，并生成测试结果之后，所述装置还被配置为：  
将所述测试结果发送给用户，以使用户根据所述测试结果对所述不相同的数据量和主键进行人工验证。

[权利要求 11] 如权利要求8所述的装置，其中，所述测试结果表包括不相同的数据量和主键，在对比第一查询结果表和第二查询结果表中的数据量和主键，并生成测试结果之后，所述装置还被配置为：  
将所述测试结果发送至上游数据库，以便所述上游数据库根据所述测试结果将所述不相同的数据量和主键重新发送至大数据中心。

[权利要求 12] 如权利要求8所述的装置，其中，所述上游数据库和大数据中心各节点所需要检查的表的量是所需要检查的表的数量。

[权利要求 13] 如权利要求8所述的装置，其中，所述上游数据库和大数据中心各节点所需要检查的表的量是所需要检查的表所占用存储空间的大小。

[权利要求 14] 如权利要求8所述的装置，其中，所述上游数据库和大数据中心各节点所需要检查的表的量是所需要检查的表的数量和需要检查的表所占用存储空间的大小的加权和。

[权利要求 15] 一种测试中心服务器，包括处理器及存储器，所述存储器上存储有计算机可读指令，所述计算机可读指令被所述处理器执行时，所述处理

器用于执行如下处理：

将待检查的表的表名写入配置文件，并根据所述表名确定上游数据库和大数据中心各节点所需要检查的表的数量，并将所述数量与节点进行关联后，写入配置文件；

根据所述上游数据库和大数据中心各节点所需要检查的表的量分配去各节点查询时所占用的线程数，并将所述线程数和节点关联后，写入配置文件；

向上游数据库发送配置文件，以便所述上游数据库的主服务器根据所述配置文件中所分配的去各节点查询时所占用的线程数以及配置文件中的表名，分别在上游数据库各节点的源表对所述需要检查的表的数据量和主键进行查询，并将所述上游数据库对应的数据量和主键的结果，写入第一查询结果表，所述主键是表中的一个或多个字段，每一个主键的值用于唯一地标识表中的一条记录；

向大数据中心发送所述配置文件，以便所述大数据中心的主服务器根据所述配置文件中线程数和节点、所分配的去各节点查询时所占用的线程数以及配置文件中的表名，分别在大数据中心各节点的目标表对所述需要检查的表的数据量和主键进行查询，并将所述大数据中心对应的数据量和主键的结果写入第二查询结果表；

接收所述上游数据库发送的第一查询结果表和所述大数据中心发送的第二查询结果表；

基于第一查询结果表和第二查询结果表中的数据量和主键生成测试结果。

[权利要求 16]

如权利要求15所述的测试中心服务器，其中，所述基于第一查询结果表和第二查询结果表中的数据量和主键生成测试结果的步骤包括：

将所述第一查询结果表中的数据量和主键和第二查询结果表中对应的数据量和主键输入机器学习模型，获得每一对数据和主键的测试结果；

将所述测试结果写入测试结果表；

其中，所述机器学习模型的训练过程为：

将正查询结果样本对输入机器学习模型，所述正查询结果样本对是两个数据量样本或者两个主键样本，其中一个数据量样本是另一个数据量样本的相同对象，或者其中一个主键样本是另一个主键样本的相同对象，若输出的判定结果不相同，则调整所述机器学习模型的系数；  
将负查询结果样本输入机器学习模型，所述负查询结果样本对是两个数据量样本或者两个主键样本，其中一个数据量样本是另一个数据量样本的不同对象，或者其中一个主键样本是另一个主键样本的不同对象，若输出的判定结果为相同，则调整所述机器学习模型的系数。

[权利要求 17] 如权利要求15所述的测试中心服务器，其中，所述测试结果表包括不相同的数据量和主键，所述计算机可读指令被所述处理器执行时，所述处理器还用于执行以下处理：在对比第一查询结果表和第二查询结果表中的数据量和主键，并生成测试结果之后，  
将所述测试结果发送给用户，以使用户根据所述测试结果对所述不相同的数据量和主键进行人工验证。

[权利要求 18] 如权利要求15所述的测试中心服务器，其中，所述测试结果表包括不相同的数据量和主键，所述计算机可读指令被所述处理器执行时，所述处理器还用于执行以下处理：在对比第一查询结果表和第二查询结果表中的数据量和主键，并生成测试结果之后，  
将所述测试结果发送至上游数据库，以便所述上游数据库根据所述测试结果将所述不相同的数据量和主键重新发送至大数据中心。

[权利要求 19] 如权利要求15所述的测试中心服务器，其中，所述上游数据库和大数据中心各节点所需要检查的表的量是所需要检查的表的数量。

[权利要求 20] 如权利要求15所述的测试中心服务器，其中，所述上游数据库和大数据中心各节点所需要检查的表的量是所需要检查的表所占用存储空间的大小。

[权利要求 21] 如权利要求15所述的测试中心服务器，其中，所述上游数据库和大数据中心各节点所需要检查的表的量是所需要检查的表的数量和需要检

查的表所占用存储空间的大小的加权和。

[权利要求 22] 一种计算机可读存储介质，其上存储有计算机程序，其中，所述计算机程序被处理器执行时，所述处理器用于执行如权利要求1至7所述的方法。

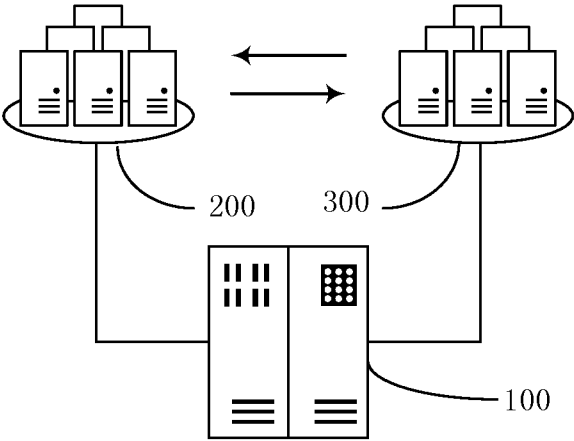


图 1



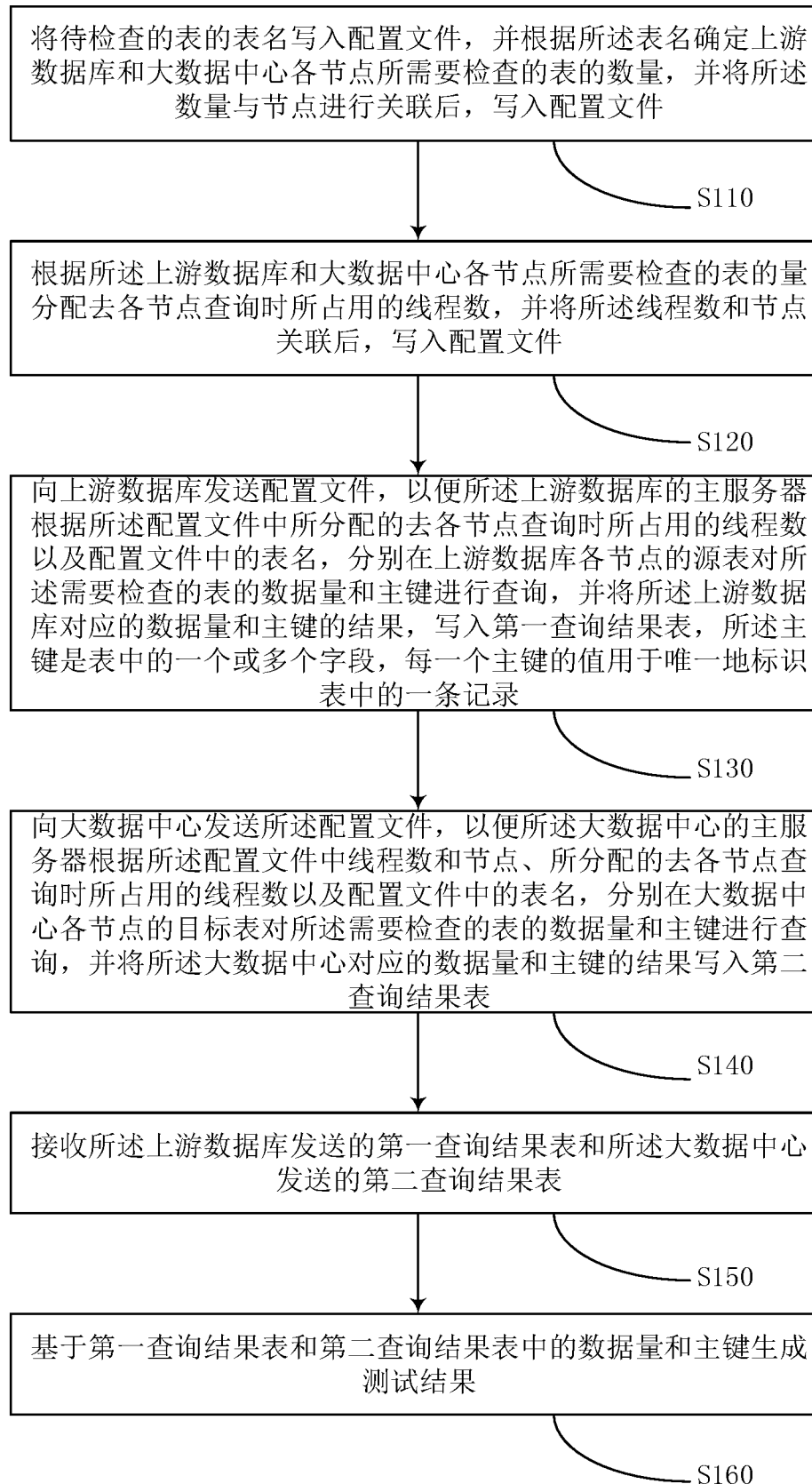


图 2

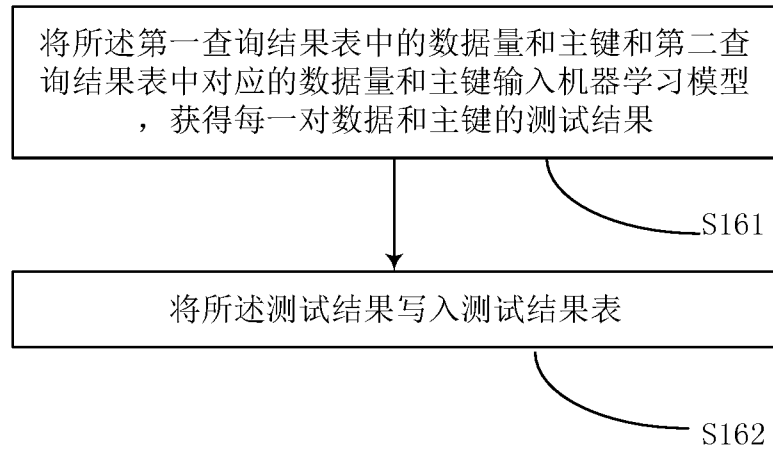


图 3

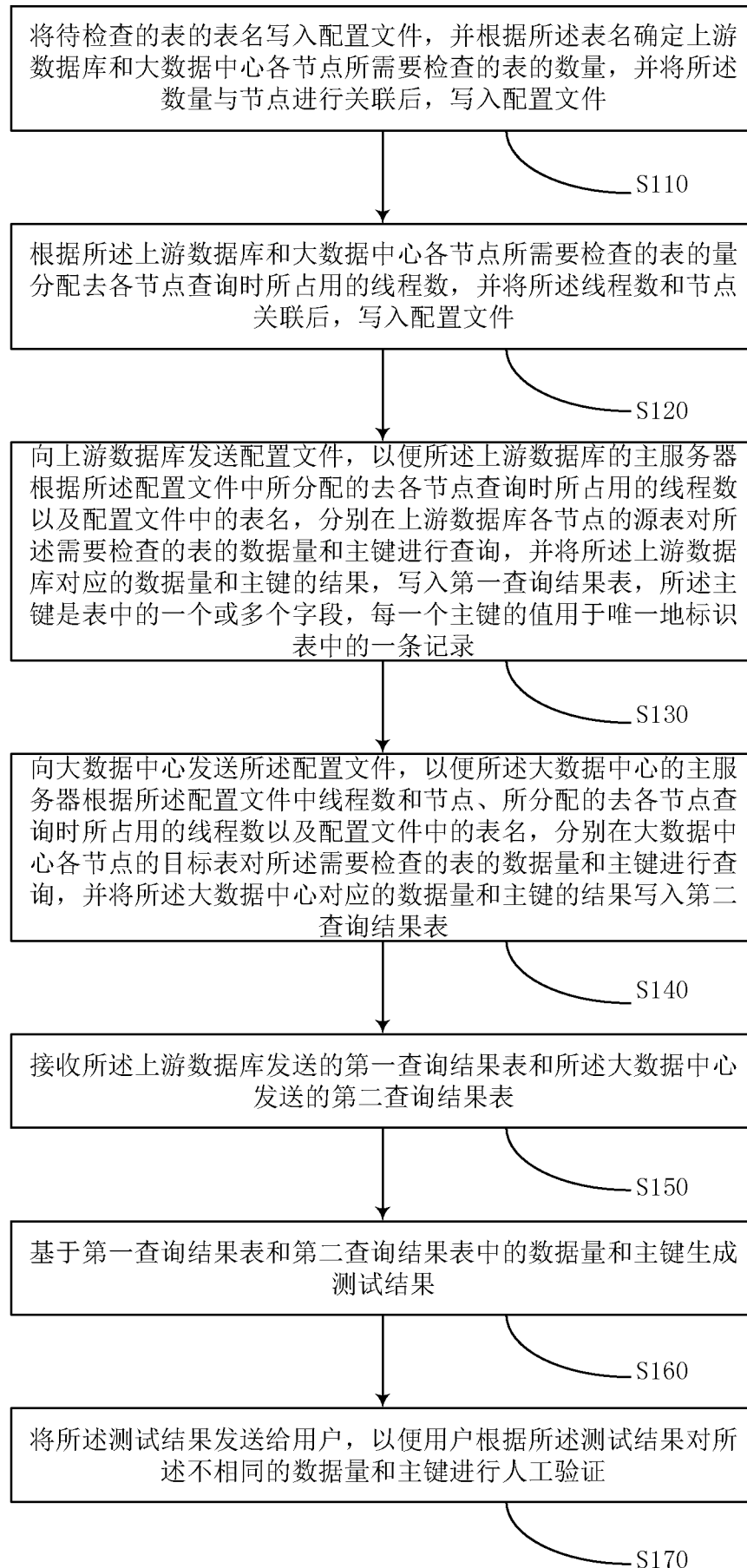


图 4

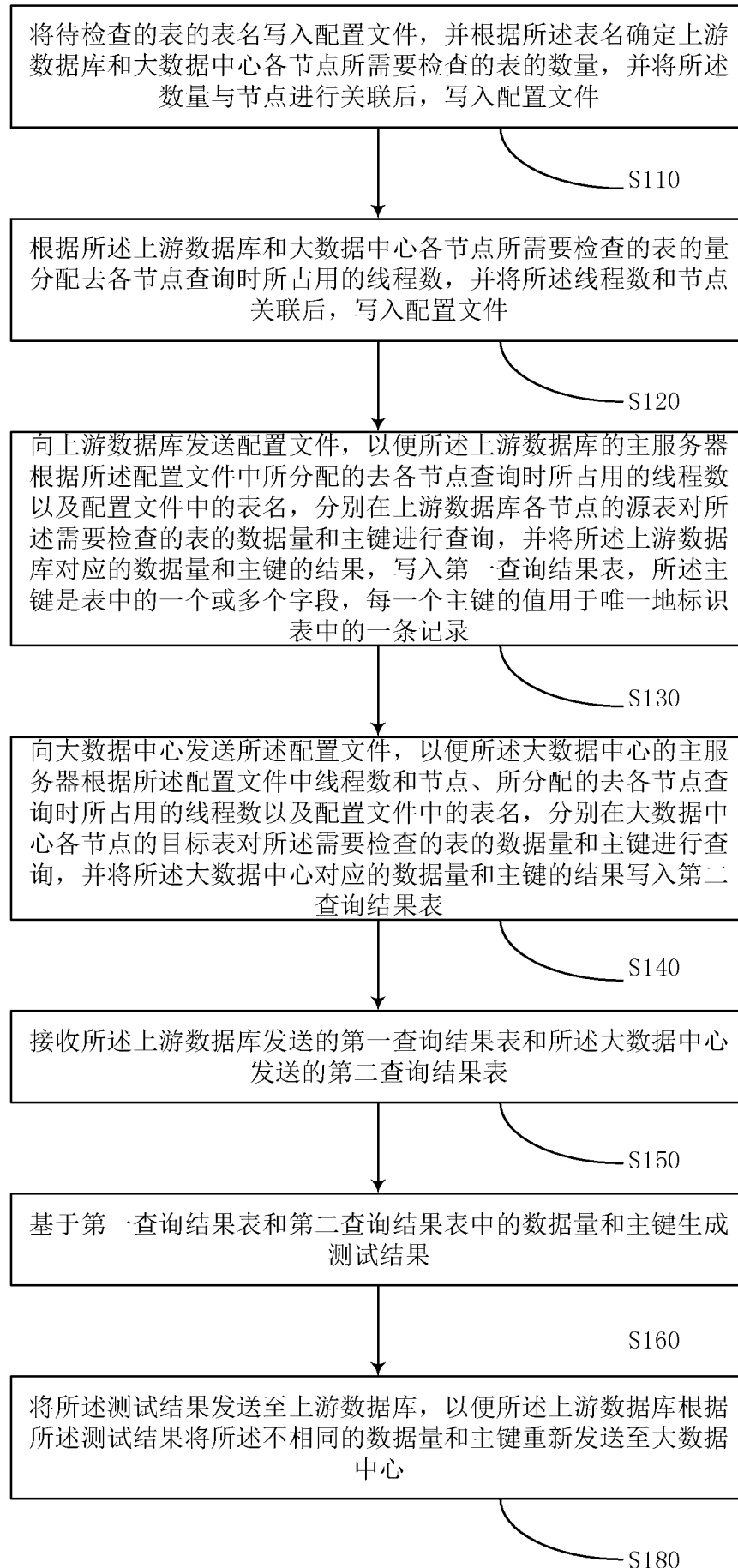


图 5

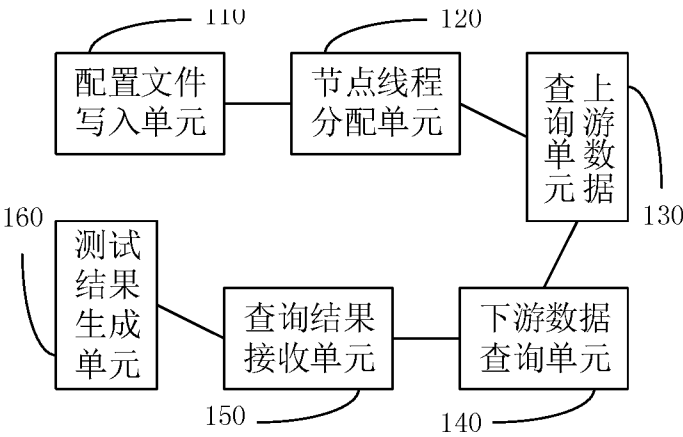


图 6

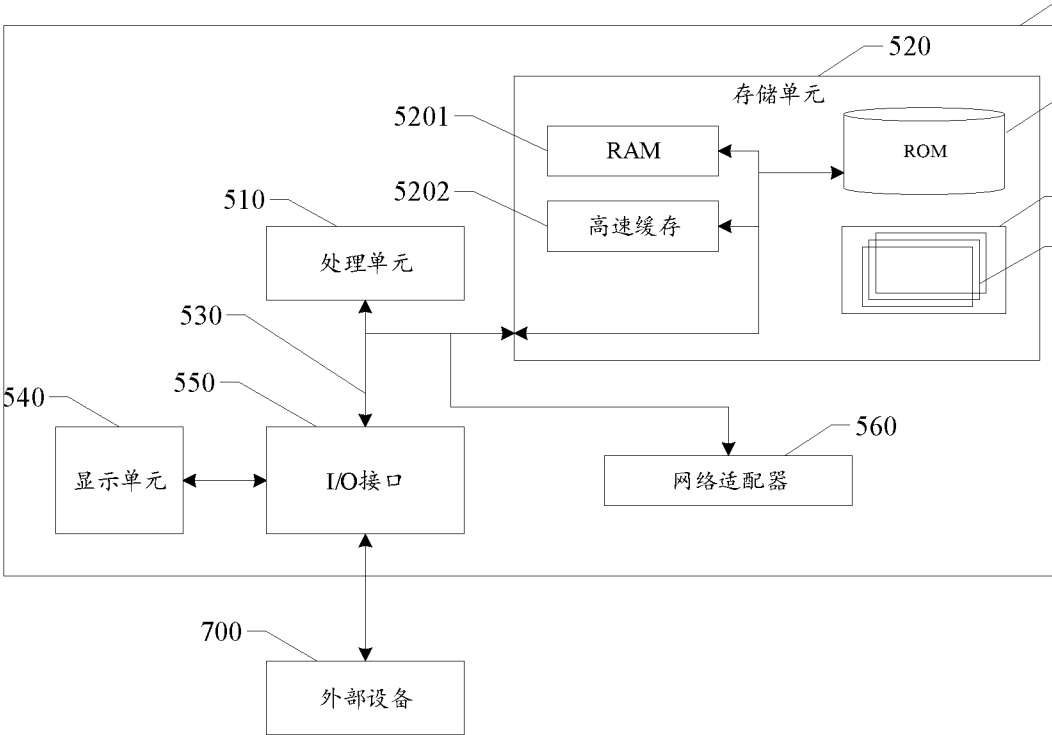


图 7

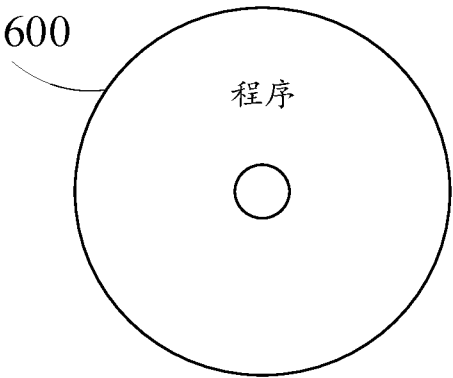


图 8

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/118587

**A. CLASSIFICATION OF SUBJECT MATTER**

G06F 11/36(2006.01)i

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

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

G06F

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

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

CNPAT, WPI, EPODOC, CNKI, IEEE, GOOGLE: 数据库, 表, 测试, 比较, 对比, 检查, 配置, 不同, 结果, database, table, test, compare, check, configuration, different, result

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

| Category* | Citation of document, with indication, where appropriate, of the relevant passages                                              | Relevant to claim No. |
|-----------|---------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| PX        | CN 110109824 A (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) 09 August 2019 (2019-08-09)<br>description, paragraphs [0045]-[0067]   | 1-22                  |
| A         | CN 107679054 A (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) 09 February 2018 (2018-02-09)<br>description, paragraphs [0092]-[0108] | 1-22                  |
| A         | CN 105760291 A (ALIBABA GROUP HOLDING LIMITED) 13 July 2016 (2016-07-13)<br>entire document                                     | 1-22                  |
| A         | US 2014006366 A1 (INTERNATIONAL BUSINESS MACHINES CORPORATION) 02 January 2014 (2014-01-02)<br>entire document                  | 1-22                  |



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

“&amp;” document member of the same patent family

Date of the actual completion of the international search

20 January 2020

Date of mailing of the international search report

12 February 2020

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/  
CN)No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing  
100088

China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

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

International application No.

**PCT/CN2019/118587**

| Patent document<br>cited in search report |            |    | Publication date<br>(day/month/year) | Patent family member(s) |            |    | Publication date<br>(day/month/year) |
|-------------------------------------------|------------|----|--------------------------------------|-------------------------|------------|----|--------------------------------------|
| CN                                        | 110109824  | A  | 09 August 2019                       | None                    |            |    |                                      |
| CN                                        | 107679054  | A  | 09 February 2018                     | WO                      | 2018227880 | A1 | 20 December 2018                     |
| CN                                        | 105760291  | A  | 13 July 2016                         | None                    |            |    |                                      |
| US                                        | 2014006366 | A1 | 02 January 2014                      | US                      | 2011099515 | A1 | 28 April 2011                        |

## 国际检索报告

国际申请号

PCT/CN2019/118587

## A. 主题的分类

G06F 11/36 (2006.01) i

按照国际专利分类 (IPC) 或者同时按照国家分类和 IPC 两种分类

## B. 检索领域

检索的最低限度文献 (标明分类系统和分类号)

G06F

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库 (数据库的名称, 和使用的检索词 (如使用))

CNPAT, WPI, EPDOC, CNKI, IEEE, GOOGLE: 数据库, 表, 测试, 比较, 对比, 检查, 配置, 不同, 结果, database, table, test, compare, check, configuration, different, result

## 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 |
|------|---------------------------------------------------------------------------------------------------|---------|
| PX   | CN 110109824 A (平安科技深圳有限公司) 2019年 8月 9日 (2019 - 08 - 09)<br>说明书第[0045]-[0067]段                    | 1-22    |
| A    | CN 107679054 A (平安科技深圳有限公司) 2018年 2月 9日 (2018 - 02 - 09)<br>说明书第[0092]-[0108]段                    | 1-22    |
| A    | CN 105760291 A (阿里巴巴集团控股有限公司) 2016年 7月 13日 (2016 - 07 - 13)<br>全文                                 | 1-22    |
| A    | US 2014006366 A1 (INTERNATIONAL BUSINESS MACHINES CORPORATION) 2014年 1月 2日 (2014 - 01 - 02)<br>全文 | 1-22    |

☐ 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

“&amp;” 同族专利的文件

国际检索实际完成的日期

2020年 1月 20日

国际检索报告邮寄日期

2020年 2月 12日

ISA/CN的名称和邮寄地址

中国国家知识产权局 (ISA/CN)

中国北京市海淀区蓟门桥西土城路6号 100088

传真号 (86-10)62019451

授权官员

孙韬敏

电话号码 86-(10)-53961420



国际检索报告  
关于同族专利的信息

国际申请号

PCT/CN2019/118587

| 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          | 110109824  | A  | 2019年 8月 9日    | 无    |            |    |                |
| CN          | 107679054  | A  | 2018年 2月 9日    | W0   | 2018227880 | A1 | 2018年 12月 20日  |
| CN          | 105760291  | A  | 2016年 7月 13日   | 无    |            |    |                |
| US          | 2014006366 | A1 | 2014年 1月 2日    | US   | 2011099515 | A1 | 2011年 4月 28日   |