

(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
2018年3月8日 (08.03.2018)



(10) 国际公布号
WO 2018/040488 A1

- (51) 国际专利分类号:
G06F 17/30 (2006.01)
- (21) 国际申请号: PCT/CN2017/071568
- (22) 国际申请日: 2017年1月18日 (18.01.2017)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201610797295.9 2016年8月31日 (31.08.2016) CN
- (71) 申请人: 华为技术有限公司 (HUAWEI TECHNOLOGIES CO., LTD.) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。
- (72) 发明人: 王振华 (WANG, Zhenhua); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。
- (74) 代理人: 北京中博世达专利商标代理有限公司 (BEIJING ZBSD PATENT&TRADEMARK AGENT LTD.); 中国北京市海淀区交大东路31号11号楼8层, Beijing 100044 (CN)。
- (81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT,

(54) Title: METHOD AND DEVICE FOR PROCESSING JOIN QUERY

(54) 发明名称: 一种处理连接查询的方法及装置

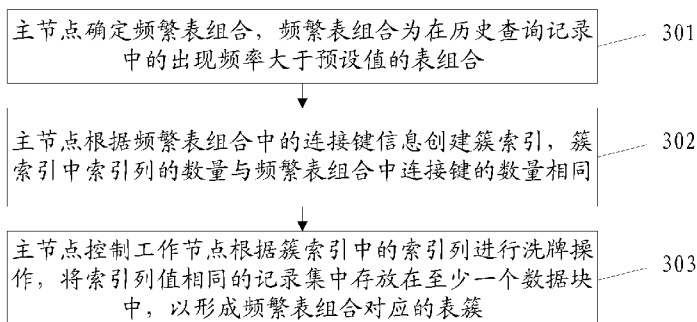


图 3

- 301 A master node determines a frequent table combination, the frequent table combination being a table combination appearing in a query history at a frequency greater than a preset value
- 302 The master node creates a cluster index according to join key information in the frequent table combination, the number of index columns in the cluster index being the same as the number of join keys in the frequent table combination
- 303 The master node controls an operation node to perform, according to the index columns in the cluster index, a shuffling operation, and stores records having the same index column value in at least one data block to form a table cluster corresponding to the frequent table combination

(57) Abstract: The invention relates to the technical field of communications. Provided are a method and device for processing a join query capable of solving a problem in which a join query has low efficiency. The method comprises: determining a frequent table combination, the frequent table combination being a table combination appearing in a query history at a frequency greater than a preset value, the table combination comprising a join key and tables joined by means of the join key; creating a cluster index according to join key information in the frequent table combination; and performing a shuffling operation according to an index column in the cluster index, and storing records having the same index column value in at least one data block to form a table cluster corresponding to the



WO 2018/040488 A1

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))。

frequent table combination. The method is suitable for performing a join query on a table.

(57) 摘要: 一种处理连接查询的方法及装置, 涉及通信技术领域, 能够解决连接查询的效率低的问题。方法包括: 确定频繁表组合, 频繁表组合为在历史查询记录中的出现频率大于预设值的表组合, 表组合包括连接键以及通过连接键进行连接的表, 然后根据频繁表组合中的连接键信息创建簇索引, 再根据簇索引中的索引列进行洗牌操作, 将索引列值相同的记录存放在至少一个数据块中, 以形成频繁表组合对应的表簇。该方法适用于对表格进行连接查询。

一种处理连接查询的方法及装置

本申请要求于 2016 年 08 月 31 日提交中国专利局、申请号为 201610797295.9、发明名称为“一种处理连接查询的方法及装置”的中国专利申请优先权，其全部内容通过引用结合在本申请中。

技术领域

本发明涉及通信技术领域，尤其涉及一种处理连接查询的方法及装置。

背景技术

网络技术的飞速发展使得数据量急剧增长，为了对大规模的数据进行高效的处理，可以采用基于 MapReduce（映射归约）的分布式计算框架进行大数据的查询分析任务，然而由于在基于 MapReduce 的分布式计算框架中执行查询分析任务时，需要针对每个任务编写复杂的程序，对于 OLAP（On-Line Analytical Processing，联机分析处理）这种复杂的查询，实现过程更为复杂，易用性较低。相比之下，SQL（Structured Query Language，结构化查询语言）的易用性较高，所以通常将 SQL 应用于基于 MapReduce 的分布式计算框架中以进行大数据的查询分析。

SQL 查询中的 Join（连接）查询可以将数据库中的两张表通过连接属性连接起来，所以可以采用分布式 Hash Join（散列连接）技术在分布式环境下确定两个表之间对某一系列的等值连接，如图 1 所示，首先需确定待处理表格的数据的存储节点，在确定的节点上启动 map 任务，对每个数据块执行一次本地 hash（哈希）过程，分别将每个 block（数据块）中的数据根据 key 值的哈希值进行分桶，例如，图 1 中的每个数据块中的数据分别被划分至三个分桶，同一分桶中的数据的 key 值的 hash 值相同，之后进行 shuffle（洗牌）过程，将相同分桶的记录传输至同一 reduce 节点上，最终在 reduce 阶段对两张表中的 key 值相同的记录执行连接操作。

然而，在实现上述方法的过程中，当数据块中数据的 key 值较为分散时，在 map 阶段对数据进行分桶的过程需要进行大量的计算，所需的时间较长，且由于每个数据块中都存在多个分桶的数据，在 shuffle 过程中会产生大量的网络连接开销和数据传输开销，结合图 1，每个数据块中的数据都分别需要传输至三个不同 reduce 节点，传输过程需要消耗一定的时间，最终导致连接查询的效率较低。

发明内容

本发明实施例提供一种处理连接查询的方法及装置，能够提高连接查询的效率。

第一方面，本发明实施例提供一种处理连接查询的方法，包括：主节点确定频繁表组合，频繁表组合为在历史查询记录中的出现频率大于预设值的表组合，表组合包括连接键以及通过连接键进行连接的表，然后根据频繁表组合中的连接键信息创建簇索引，簇索引中索引列的数量与频繁表组合中连接键的数量相同，簇索引用于指示频繁表组合中索引列值相同的记录的存储位置，之后主节点控制工作节点根据簇索引中的索引列进行洗牌操作，将索引列值相同的记录集中存放在至少一个数

据块中，以形成频繁表组合对应的表簇，其中索引列值相同的记录为通过连接键进行连接的表中的记录。可见将连接键相同的记录尽可能的存放在一个数据块中，使得数据块中数据的连接键值较为集中，大大减少了执行连接查询时，在map阶段对数据分桶时的计算量，且由于一个数据块中存储的基本都是连接键相同的数据，此时，一个数据块中就不会出现太多分桶的数据，可以减少reduce阶段将数据传输至reduce节点所需的时间，提高了连接查询的效率。

在一种可能的设计中，在完成洗牌操作，形成频繁表组合对应的表簇之后，可以进行连接查询的操作，首先接收查询请求，查询请求中包含待查询表组合，然后查找待查询表组合对应的待查询表簇，分别将待查询表簇对应的节点所包含的每个数据块中索引列值的哈希值相同的记录划分至一个分桶中，将待查询表簇中所有数据块中相同分桶中的记录传输至同一数据节点中。由于索引列值相同的数据已经被存储在一起，所以执行map任务时无需进行本地hash的过程也能够完成分桶，减少了对数据分桶时的计算量，节省了CPU开销，且表簇中只存储了包含相同索引列值的记录，对于索引列值不同的记录不需要进行处理，减少了需要处理的数据量以及启动的任务数量和磁盘I/O，之后在洗牌过程中，由于大部分数据块都只对应一个分桶，相当于一个数据块对应一个索引列值，且reduce任务优先安排在map任务的本地执行，reduce节点所需的map任务的输出数据大部分都不需要网络传输，所以大大减小了将索引列值相同的数据传输至reduce节点的过程产生的网络传输开销，缩短了传输时间，提高了连接查询的效率。

在一种可能的设计中，确定频繁表组合时，需从历史查询记录中提取表组合，生成表组合集合，然后从表组合集合中筛选出出现频率大于预设值的表组合，再从出现频率大于预设值的表组合中删除冗余表组合，确定剩余的表组合为频繁表组合。通过筛选出现频率大于预设值的表组合，可以使得最后存储的表组合为用户常用的表组合，可以提高查询效率，且将冗余表组合删除，可以节省存储空间，还能够减少后续处理的数据量，提高处理效率。

在一种可能的设计中，从出现频率大于预设值的表组合中删除冗余表组合的具体方法为：当存在包含的表相同，但是连接键不同的至少两个表组合时，保留连接键最多的表组合；当存在两个表组合所包含的连接键相同，且一个表组合中的表组成的集合是另一个表组合中表组成的集合的子集时，删除包含表较少的表组合；当存在相同表包含于至少两个表组合中时，只将所述相同表保留于所述至少两个表组合中出现频率最高的一个表组合中。

在一种可能的设计中，根据簇索引中的索引列进行洗牌操作，将索引列值相同的记录集中存放在至少一个数据块中的实现方式为：当包含同一个索引列值的记录的总大小达到一个数据块的存储空间大小的第一预设比例，且未超过一个数据块的存储空间大小时，将包含同一个索引列值的记录存储在一个数据块中；当包含同一个索引列值的记录的总大小超过一个数据块的存储空间大小时，将包含同一个索引列值的记录存储在多个数据块中；当包含同一个索引列值的记录的总大小小于一个数据块的存储空间大小的第二预设比例时，将多个包含同一个索引列值的记录存储在一个数据块中。可见，根据包含相同索引列值的记录的总大小和一个数据块存储

空间大小的关系，来尽可能地将包含相同索引列值的记录集中存储在一个或多个数据块中，相比较于分散存储在多个数据块中，可以减少连接查询过程中分桶所需的时间，所以可以减少连接查询过程所需的时间，提高连接查询的效率。

在一种可能的设计中，表组合的表示形式为：

$TG = (tab_1, \dots, tab_i, \dots, tab_N)$ $key = (key_1, \dots, key_j, \dots, key_M)$ ，其中 tab_i 为表组合中的第 i 张表， key_j 为第 j 个连接键， N 为表组合中表的个数， M 为表组合中连接键的个数。

另一方面，本发明实施例提供了一种处理连接查询的方法，包括：主节点接收查询请求，查询请求中包含待查询表组合，然后在分布式文件系统DFS中查找与待查询表组合对应的待查询表簇，主节点控制工作节点分别将待查询表簇对应的节点所包含的每个数据块中索引列值的哈希值相同的记录划分至一个分桶中，再将待查询表簇中所有数据块中相同分桶中的记录传输至同一数据节点中。其中，DFS中存储了频繁表组合对应的表簇，频繁表组合为在历史查询记录中的出现频率大于预设值的表组合，表组合包括连接键以及通过连接键进行连接的表，频繁表组合中连接键的数量与簇索引中索引列的数量相同，簇索引用于指示频繁表组合中索引列值相同的记录的存储位置，其中，频繁表组合中索引列值相同的记录集中存放在至少一个数据块中，形成频繁表组合对应的表簇。由于索引列值相同的数据已经被存储在一起，所以执行map任务时无需进行本地hash的过程也能够完成分桶，减少了对数据分桶时的计算量，节省了CPU开销，且表簇中只存储了包含相同索引列值的记录，对于索引列值不同的记录不需要进行处理，减少了需要处理的数据量以及启动的任务数量和磁盘I/O，之后在洗牌过程中，由于大部分数据块都只对应一个分桶，相当于一个数据块对应一个索引列值，且 reduce任务优先安排在map任务的本地执行，reduce节点所需的map任务的输出数据大部分都不需要网络传输，所以大大减小了将索引列值相同的数据传输至reduce节点的过程产生的网络传输开销，缩短了传输时间，提高了连接查询的效率。

另一方面，本发明实施例提供了一种处理连接查询的装置，该装置可以实现上述方法实施例中主节点的功能，所述功能可以通过硬件实现，也可以通过硬件执行相应的软件实现。所述硬件或软件包括一个或多个上述功能相应的模块。

在一种可能的设计中，该装置的结构中包括处理器和收发器，该处理器被配置为支持主节点执行上述方法中相应的功能。该收发器用于支持主节点与其他网元之间的通信。主节点还可以包括存储器，该存储器用于与处理器耦合，其保存该装置必要的程序指令和数据。

又一方面，本发明实施例提供了一种大数据分析系统，该系统包括上述方面所述的主节点、用于存储表簇的元数据存储单元 metastore、用于发送连接查询请求的客户端、用于承载上述方法中的数据块的分布式文件系统 DFS，以及用于读取 DFS 中数据，并对数据进行计算的工作节点。

再一方面，本发明实施例提供了一种计算机存储介质，用于储存为上述、主节点所用的计算机软件指令，其包含用于执行上述方面所设计的程序。

相比于现有技术，本发明实施例提供的技术方案中，可以将连接键相同的记录尽可能的存放在一个数据块中，使得数据块中数据的连接键值较为集中，大大减少

了执行连接查询时，在map阶段对数据分桶时的计算量，且由于一个数据块中存储的基本都是连接键相同的数据，此时，一个数据块中就不会出现太多分桶的数据，可以减少reduce阶段将数据传输至reduce节点所需的时间，提高了连接查询的效率。

附图说明

为了更清楚地说明本发明实施例中的技术方案，下面将对实施例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发明的一些实施例，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它的附图。

图1为背景技术例提供的一种处理连接查询的方法的示例性示意图；

图2为本发明实施例提供的一种集群式的大数据分析系统的结构示意图；

图3为本发明实施例提供的一种处理连接查询的方法的流程图；

图4为本发明实施例提供的另一种处理连接查询的方法的流程图；

图5为本发明实施例提供的一种处理连接查询的方法的示例性示意图；

图6为本发明实施例提供的另一种处理连接查询的方法的流程图；

图7为本发明实施例提供的另一种处理连接查询的方法的示例性示意图；

图8为本发明实施例提供的一种处理连接查询的装置的逻辑结构示意图；

图9为本发明实施例提供的另一种处理连接查询的装置的逻辑结构示意图；

图10为本发明实施例提供的另一种处理连接查询的装置的逻辑结构示意图；

图11为本发明实施例提供的另一种处理连接查询的装置的逻辑结构示意图。

具体实施方式

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它实施例，都属于本发明保护的范围。

为了提高连接查询的效率，本发明实施例提供一种处理连接查询的方法，该方法具体可以应用于集群式的大数据分析系统，如图2所示，该系统包括Client(客户端)、Metastore(元数据存储单元)、一个主节点(master)、多个工作节点(workers)以及DFS(Distribute File System, 分布式文件系统)。

其中，master中包含表簇管理模块和SQL引擎，表簇管理模块中还包括工作负载分析子模块以及簇索引维护子模块。DFS中包含多个节点(node)，每个节点中包含至少一个数据块(block)。

结合图2，工作负载(workload)是由多个历史SQL查询组成的一个查询集合，可以通过记录客户端在一段预设时间周期内提交的历史SQL查询来得到。主节点可以通过对工作负载中的历史SQL查询记录进行分析来构建表簇和簇索引结构，将建立好的表簇存储在Metastore中，之后由工作节点对表簇中的数据进行重组，并将重组后的数据存储在DFS中相应节点中。当主节点接收到客户端发送的SQL查询请求时，通过SQL引擎执行连接查询的过程，在连接查询的过程中，SQL引擎可以调用工作节点读取DFS中存储的数据，且工作节点还能够完成一些计算工作，具体的处理连接查询的方法可参考下述实施例。

结合图2所示的大数据分析系统，为了提高连接查询的效率，本发明实施例提供一种处理连接查询的方法，如图3所示，该方法包括：

301、主节点确定频繁表组合。

其中，频繁表组合为在历史查询记录中的出现频率大于预设值的表组合，表组合包括连接键以及通过连接键进行连接的表。

连接键为表格之间的连接属性，即表格之间公共列的属性。表组合的表示形式为：

$$TG = (tab_1, \dots, tab_i, \dots, tab_N) \quad key = (key_1, \dots, key_j, \dots, key_M)$$

其中， tab_i 为表组合中的第*i*张表， key_j 为第*j*个连接键，*N*为表组合中表的个数，*M*为表组合中连接键的个数。

需要说明的是，本步骤中的历史查询记录为主节点接收到的终端上报的工作负载，工作负载是由多个历史SQL查询记录组成的一个查询集合。主节点可以根据预设的时间段周期性地从工作负载中确定频繁表组合。

302、主节点根据频繁表组合中的连接键信息创建簇索引，簇索引中索引列的数量与频繁表组合中连接键的数量相同。

其中，一个连接键对应一个索引列，当连接键的个数*M*等于1时，创建单列索引，当连接键的个数*M*大于或等于2时，创建复合索引。

需要说明的是，簇索引用于指示频繁表组合中索引列值相同的记录的存储位置，例如，频繁表组合中包含表A和表B，表A和表B具有一个公共列，簇索引用于指示表A和表B中的索引列值相同的记录的存储位置。

303、主节点控制工作节点根据簇索引中的索引列进行洗牌操作，将索引列值相同的记录集中存放在至少一个数据块中，以形成频繁表组合对应的表簇。

其中，索引列值相同的记录为通过连接键进行连接的表中的记录，一个表簇由一组共享数据块的表组成，表组合中的表的公共列为表簇的簇索引，将表组合中的所有表中包含相同索引列值的记录存放在至少一个数据块中，使得包含相同索引列值的记录的存放位置较为集中。

例如，如果一个数据块的存储空间足以存储表组合中的所有表包含相同索引列值的记录，则将这些索引列值相同的记录集中存放在一个数据块中。

需要说明的是，当需要进行洗牌操作时，主节点会向工作节点发送控制指令，控制工作节点进行洗牌操作，并将索引列值相同的记录集中存放在DFS中的至少一个数据块中。

还需说明的是，步骤301至步骤302具体由主节点中的表簇管理模块来执行。

本发明实施例提供的处理连接查询的方法，确定频繁表组合，然后根据频繁表组合中的连接键信息创建簇索引，再根据簇索引中的索引列进行shuffle操作，将索引列值相同的记录集中存放在至少一个数据块中，从而形成频繁表组合对应的表簇，通过这种方法，可以将连接键相同的记录尽可能的存放在一个数据块中，使得数据块中数据的连接键值较为集中，大大减少了执行连接查询时，在map阶段对数据分桶时的计算量，且由于一个数据块中存储的基本都是连接键相同的数据，此时，一个数据块中就不会出现太多分桶的数据，可以减少reduce阶段将数据传输至reduce节点

所需的时间，提高了连接查询的效率。

结合图2，需要说明的是，本发明实施例中的频繁表组合是主节点根据客户端上报的工作负载确定的，基于此，在本发明实施例提供的另一种实现方式中，对确定频繁表组合的方法进行了详细说明，如图4所示，上述步骤301、主节点确定频繁表组合，具体可以实现为步骤3011至步骤3013。

3011、主节点从历史查询记录中提取表组合，生成表组合集合。

其中，主节点会对工作负载中的每一条查询记录进行形式转换，从SQL查询语句中提取出使用连接键进行连接操作的表，以形成一个表组合。如果一条查询记录中有多条查询语句，则根据每条查询语句生成一个表组合，将生成的所有表组合组成表组合集合。

例如，其中一条查询记录为：

```
web_sales ws JOIN warehouse w ON ws.ws_warehouse_sk=w.w_warehouse_sk,  
catalog_sales cs JOIN ship_mode sm ON cs.cs_mode_sk=sm.sm_ship_mode_sk
```

可见，这条查询语句中包含两个连接查询，对第一个连接查询进行形式转换，得到的表组合为：

TQ1: (web_sales, warehouse) key= warehouse_sk。

其中，web_sales和warehouse是两个表格，这两个表格之间的连接键为warehouse_sk。

对第二个连接查询进行形式转换，得到的表组合为：

TQ2: (catalog_sales, ship_mode) key= ship_mode_sk。

其中，catalog_sales和ship_mode是两个表格，这两个表格之间的连接键为ship_mode_sk。

3012、主节点从表组合集合中筛选出出现频率大于预设值的表组合。

在生成表组合集合之后，可以利用频繁项集挖掘算法计算在表组合集合中出现频率大于预设值的表组合，将在表组合集合中出现频率大于预设值的表组合组成频繁表组合集合。

3013、主节点从出现频率大于预设值的表组合中删除冗余表组合，确定剩余的表组合为频繁表组合。

需要说明的是，上述步骤3012生成的频繁表组合集合中可能会存在冗余表组合，可以采用以下三种规则中的任一种或多种来过滤冗余组合。

规则一、

当存在包含的表相同，但是连接键不同的至少两个表组合时，保留连接键最多的表组合；

例如，存在表组合TG1: (catalog_sales, catalog_returns) key= item_sk以及表组合TG2: (catalog_sales, catalog_returns) key= item_sk and order_number。

可以看出，TG1和TG2中包含的表都是catalog_sales和catalog_returns，但是TG1只有一个连接键item_sk，而TG2有两个连接键item_sk 和order_number。显然，TG1和TG2包含的表相同，但是TG2的连接键比TG1的连接键多，所以保留表组合TG2，将表组合TG1删除。

规则二、

当存在两个表组合所包含的连接键相同，且一个表组合中的表组成的集合是另一个表组合中的表组成的集合的子集时，删除包含表较少的表组合。

例如，存在表组合TG1: (catalog_sales, catalog_returns) key= item_sk以及表组合TG2: (catalog_sales, catalog_returns, item) key= item_sk。

可以看出，TG1和TG2包含的连接键相同，都是item_sk，TG1中的表组成的集合{ catalog_sales, catalog_returns }是TG2中的表组成的集合{ catalog_sales, catalog_returns , item }的子集，所以删除表组合TG1，保留表组合TG2。

规则三、

当存在相同表包含于至少两个表组合中时，只将相同表保留于至少两个表组合中出现频率最高的一个表组合中。

例如，存在表组合TG1: (catalog_sales, ship_mode) key= item_sk、表组合TG2: (catalog_sales, web_sales, item) key= item_sk and order_number，以及表组合TG3: (catalog_sales, catalog_returns, item) key=item_sk。

可以看出，TG1、TG2和TG3中都包含catalog_sales，假设通过频繁表项挖掘算法确定TG3为出现频率最高的表组合，则将catalog_sales从TG1和TG2中删除，只保留在TG3中，删除之后，TG1变为(ship_mode)key= item_sk，TG2变为(web_sales, item) key= item_sk and order_number。

还需说明的是，如果将相同表从某个表组合删除之后，如果该表组合只剩下一个成员表，则无法构成表组合，此时直接将该表组合删除，例如TG1中的catalog_sales被删除后，只剩下一个表ship_mode，此时可以直接将表组合TG1删除。

如图5所示，图5为本发明实施例提供的分布式表簇的创建流程的示例性示意图，在生成表组合后，还需要创建簇索引，进而根据簇索引生成表簇并存储在数据块中，以下对表组合的存储方法进行说明，上述步骤303、主节点控制工作节点根据簇索引中的索引列进行shuffle操作，将通过连接键进行连接的表中索引列值相同的记录存放在至少一个数据块中，具体可以实现为步骤3031至步骤3033。

3031、当包含同一个索引列值的记录的总大小达到一个数据块的存储空间大小的第一预设比例，且未超过一个数据块的存储空间大小时，将包含同一个索引列值的记录存储在一个数据块中。

其中，第一预设比例可以为80%。

结合图5，以频繁表组合TG1: (catalog_sales, catalog_returns) key= item_sk and order_number为例，连接键item_sk和order_number分别对应一个索引列，连接键item_sk对应的索引列的索引列值包括A和B，连接键order_number对应的索引列的索引列值包括1、2、3、4。

包含第一个索引列值 (A, 1) 的记录的大小达到了block1总大小的80%，且未超出block1的总存储空间大小，所以可以只将包含索引列值1的记录存储在block1中。

3032、当包含同一个索引列值的记录的总大小超过一个数据块的存储空间大小时，将包含同一个索引列值的记录存储在多个数据块中。

例如，包含第四个索引列值 (B, 4) 的记录的大小已经超出了一个block存储

空间大小,则将包含索引列值(B,4)的记录存储在block3和block4这两个数据块中。

3033、当包含同一个索引列值的记录的总大小小于一个数据块的存储空间大小的第二预设比例时,将多个包含同一个索引列值的记录存储在一个数据块中。

例如,包含第二个索引列值(A,3)的记录和包含第三个索引列值(B,2)的记录的总大小未超出block2的总存储空间大小,则将包含第二个索引列值(A,3)的记录和包含第三个索引列值(B,2)的记录均存储在block2中。

本发明实施例提供的处理连接查询的方法,通过筛选出现频率大于预设值的表组合,可以使得最后存储的表组合为用户常用的表组合,可以提高查询效率,且将冗余表组合删除,可以节省存储空间,最后,根据包含相同索引列值的记录的总大小和一个数据块存储空间大小的关系,来尽可能地将包含相同索引列值的记录集中存储在一个或多个数据块中,相比较于分散存储在多个数据块中,由于包含相同索引列值的记录集中存储在一个或多个数据块中,可以认为在数据存储于数据块中时已经完成了分桶,减少连接查询过程中分桶所需的时间,所以可以减少连接查询过程所需的时间,提高连接查询的效率。

结合图3所示的方法流程,在建立好表簇之后,即可进行后续的连接查询操作,基于本发明实施例建立的表簇,在本发明实施例提供的一种实现方式中,还提供了基于上述实施例中建立的表簇处理连接查询的过程,如图6所示,该方法包括:

601、主节点接收查询请求,查询请求中包含待查询表组合。

需要说明的是,在建立好表簇之后,当主节点中的SQL引擎接收到SQL查询请求时,需通过查询编译器来确定查询请求所请求进行连接计算的待查询表组合。

602、主节点查找待查询表组合对应的待查询表簇。

在确定待查询表组合之后,主节点中的SQL引擎可以通过表簇管理模块确定待查询表组合对应的待查询表簇。

603、主节点控制工作节点分别将待查询表簇所包含的每个数据块中索引列值的哈希值相同的记录划分至一个分桶中。

需要说明的是,当主节点中的SQL引擎接收到SQL查询请求时,主节点可向工作节点发送控制指令,控制工作节点执行后续的mapreduce过程。

在确定待查询表簇之后,SQL引擎可通过工作节点读取分布式文件系统中的关于待查询表簇的数据,对待查询表簇进行连接查询的过程如图7所示,主节点控制工作节点分别在两个节点上启动map程序,而由于索引列值(key值)相同的数据已经被集中存储在一起,所以map程序中无需进行本地hash的过程,即可完成分桶操作,例如,图8中,block1中只存储了索引列值为1的记录,所以block1中存储的所有记录就是一个分桶,block2中也只存储了索引列值为1的记录,所以block2中存储的所有记录就是一个分桶,block3中分别存储了索引列值为2的记录和部分索引列值为3的记录,且索引列值为2的记录和索引列值为3的记录是独立存储的,所以无需进行本地hash也能将block3中存储的数据分为两个分桶,block4中存储的都是索引列值为3的记录,所以block3中存储所有记录为一个分桶。

604、主节点控制工作节点将待查询表簇中所有数据块中相同分桶中的记录传输至同一数据节点中。

本发明实施例提供的处理连接查询的方法，在建立好表簇之后再接收到查询请求时，确定查询请求中的待查询表组合，然后查找待查询表组合对应的待查询表簇，在待查询表簇对应的节点上执行map任务，由于索引列值相同的数据已经被存储在一起，所以执行map任务时无需进行本地hash的过程也能够完成分桶，减少了对数据分桶时的计算量，节省了CPU开销，且表簇中只存储了包含相同索引列值的记录，对于索引列值不同的记录不需要进行处理，减少了需要处理的数据量以及启动的任务数量和磁盘I/O，之后在洗牌过程中，由于大部分数据块都只对应一个分桶，相当于一个数据块对应一个索引列值，且 reduce任务优先安排在map任务的本地执行，reduce节点所需的map任务的输出数据大部分都不需要网络传输，所以大大减小了将索引列值相同的数据传输至reduce节点的过程产生的网络传输开销，缩短了传输时间，提高了连接查询的效率。

上述主要从各个网元之间交互的角度对本发明实施例提供的方案进行了介绍。可以理解的是，各个网元，例如连接查询的装置等为了实现上述功能，其包含了执行各个功能相应的硬件结构和/或软件模块。本领域技术人员应该很容易意识到，结合本文中所公开的实施例描述的各示例的单元及算法步骤，本发明能够以硬件或硬件和计算机软件的结合形式来实现。某个功能究竟以硬件还是计算机软件驱动硬件的方式来执行，取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能，但是这种实现不应认为超出本发明的范围。

本发明实施例可以根据上述方法示例对图2所示的主节点等进行功能模块的划分，例如，可以对应各个功能划分各个功能模块，也可以将两个或两个以上的功能集成在一个处理模块中。上述集成的模块既可以采用硬件的形式实现，也可以采用软件功能模块的形式实现。需要说明的是，本发明实施例中对模块的划分是示意性的，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。

在采用对应各个功能划分各个功能模块的情况下，图8示出了上述实施例中所涉及的主节点的一种可能的结构示意图，图8具体为图2所示的主节点中的表簇管理模块的结构示意图，主节点包括：确定单元801，创建单元802，洗牌单元803。确定单元801用于支持主节点执行图3中的步骤301，图4中的步骤3011至3013；创建单元802用于支持主节点执行图3中的步骤302；洗牌单元803用于支持主节点控制工作节点执行图3中的步骤303，支持主节点执行图4中的步骤3031至3033。

在采用对应各个功能划分各个功能模块的情况下，图9还示出了上述实施例中所涉及的主节点的另一可能的结构示意图，图9具体为图2所示的主节点中的SQL引擎的结构示意图，主节点中包括：接收单元901，查找单元902，分桶单元903，传输单元904。其中，接收单元901用于主节点执行图6中的步骤601；查找单元902用于支持主节点执行图6中的步骤602；分桶单元903用于支持主节点控制工作节点执行图6中的步骤603，传输单元904用于支持主节点控制工作节点执行图6中的步骤604。

其中，上述方法实施例涉及的所有相关内容均可以援引到对应功能模块的功能描述，在此不再赘述。

在采用集成的单元的情况下，图 10 示出了上述实施例中所涉及的主节点的一种可能的结构示意图。主节点包括：处理模块 1002 和通信模块 1003。处理模块 1002 用于对主节点的动作进行控制管理，例如，处理模块 1002 用于支持图 3 中的步骤 301 至 303，图 4 中的步骤 3011 至 3033，图 6 中的步骤 602 至 603；通信模块 1003 用于支持主节点与其他网络实体的通信，例如，通信模块 1003 用于支持图 6 中的步骤 601 和 604，可以实现与图 2 或图 5 中示出的功能模块或网络实体之间的通信。该装置还包括存储模块 1001，用于存储主节点的程序代码和数据。

其中，处理模块 1002 可以是处理器或控制器，例如可以是中央处理器（Central Processing Unit, CPU），通用处理器，数字信号处理器（Digital Signal Processor, DSP），专用集成电路（Application-Specific Integrated Circuit, ASIC），现场可编程门阵列（Field Programmable Gate Array, FPGA）或者其他可编程逻辑器件、晶体管逻辑器件、硬件部件或者其任意组合。其可以实现或执行结合本发明公开内容所描述的各种示例性的逻辑方框，模块和电路。所述处理器也可以是实现计算功能的组合，例如包含一个或多个微处理器组合，DSP 和微处理器的组合等等。通信模块 1003 可以是收发器、收发电路或通信接口等。存储模块 1001 可以是存储器。

图 8 至图 10 为主节点实现为软件时的结构示意图，而主节点还可以以硬件的方式存在，即当处理模块 1002 为处理器，通信模块 1003 为收发器，存储模块 1001 为存储器时，本发明实施例所涉及的主节点可以为图 11 所示的主节点。

参阅图 11 所示，当主节点实现为硬件时，包括：处理器 1102、收发器 1103、存储器 1101 以及总线 1104。其中，收发器 1103、处理器 1102 以及存储器 1101 通过总线 1104 相互连接；总线 1104 可以是外设部件互连标准（Peripheral Component Interconnect PCI）总线或扩展工业标准结构（Extended Industry Standard Architecture, EISA）总线等。所述总线可以分为地址总线、数据总线、控制总线等。为便于表示，图 11 中仅用一条粗线表示，但并不表示仅有一根总线或一种类型的总线。

需要说明的是，本发明实施例中的主节点和工作节点可以分别为一个单机设备，也可以以软件的方式存在，例如在一个计算机集群系统中，主节点的功能和工作节点的功能可以由该系统中不同的虚拟机实现。

结合本发明公开内容所描述的方法或者算法的步骤可以硬件的方式来实现，也可以由处理器执行软件指令的方式来实现。软件指令可以由相应的软件模块组成，软件模块可以被存放于随机存取存储器（Random Access Memory, RAM）、闪存、只读存储器（Read Only Memory, ROM）、可擦除可编程只读存储器（Erasable Programmable ROM, EPROM）、电可擦可编程只读存储器（Electrically EPROM, EEPROM）、寄存器、硬盘、移动硬盘、只读光盘（CD-ROM）或者本领域熟知的任何其它形式的存储介质中。一种示例性的存储介质耦合至处理器，从而使处理器能够从该存储介质读取信息，且可向该存储介质写入信息。当然，存储介质也可以是处理器的组成部分。处理器和存储介质可以位于 ASIC 中。另外，该 ASIC 可以位于核心网接口设备中。当然，处理器和存储介质也可以作为分立组件存在于核心网接口设备中。

本领域技术人员应该可以意识到，在上述一个或多个示例中，本发明所描述的

功能可以用硬件、软件、固件或它们的任意组合来实现。当使用软件实现时，可以将这些功能存储在计算机可读介质中或者作为计算机可读介质上的一个或多个指令或代码进行传输。计算机可读介质包括计算机存储介质和通信介质，其中通信介质包括便于从一个地方向另一个地方传送计算机程序的任何介质。存储介质可以是通用或专用计算机能够存取的任何可用介质。

本说明书中的各个实施例均采用递进的方式描述，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，对于设备实施例而言，由于其基本相似于方法实施例，所以描述得比较简单，相关之处参见方法实施例的部分说明即可。

以上所述，仅为本发明的具体实施方式，但本发明的保护范围并不局限于此，任何熟悉本技术领域的技术人员在本发明揭露的技术范围内，可轻易想到的变化或替换，都应涵盖在本发明的保护范围之内。因此，本发明的保护范围应该以权利要求的保护范围为准。

权 利 要 求 书

1、一种处理连接查询的方法，其特征在于，包括：

确定频繁表组合，所述频繁表组合为在历史查询记录中的出现频率大于预设值的表组合，所述表组合包括连接键以及通过连接键进行连接的表；

根据所述频繁表组合中的连接键信息创建簇索引，所述簇索引中索引列的数量与所述频繁表组合中连接键的数量相同，所述簇索引用于指示所述频繁表组合中索引列值相同的记录的存储位置；

根据所述簇索引中的索引列进行洗牌操作，将索引列值相同的记录集中存放在至少一个数据块中，以形成所述频繁表组合对应的表簇，所述索引列值相同的记录为通过所述连接键进行连接的表中的记录。

2、根据权利要求1所述的处理连接查询的方法，其特征在于，在所述根据所述簇索引中的索引列进行洗牌操作，将索引列值相同的记录集中存放在至少一个数据块中之后，所述方法还包括：

接收查询请求，所述查询请求中包含待查询表组合；

查找所述待查询表组合对应的待查询表簇；

分别将所述待查询表簇对应的节点所包含的每个数据块中索引列值的哈希值相同的记录划分至一个分桶中；

将所述待查询表簇中所有数据块中相同分桶中的记录传输至同一数据节点中。

3、根据权利要求2所述的处理连接查询的方法，其特征在于，所述确定频繁表组合，包括：

从所述历史查询记录中提取表组合，生成表组合集合；

从所述表组合集合中筛选出出现频率大于预设值的表组合；

从出现频率大于预设值的表组合中删除冗余表组合，确定剩余的表组合为频繁表组合。

4、根据权利要求3所述的处理连接查询的方法，其特征在于，所述从出现频率大于预设值的表组合中删除冗余表组合，包括：

当存在包含的表相同，但是连接键不同的至少两个表组合时，保留连接键最多的表组合；

当存在两个表组合所包含的连接键相同，且一个表组合中的表组成的集合是另一个表组合中表组成的集合的子集时，删除包含表较少的表组合；

当存在相同表包含于至少两个表组合中时，只将所述相同表保留于所述至少两个表组合中出现频率最高的一个表组合中。

5、根据权利要求1至4任一项所述的处理连接查询的方法，其特征在于，所述根据所述簇索引中的索引列进行洗牌操作，将索引列值相同的记录集中存放在至少一个数据块中，包括：

当包含同一个索引列值的记录的总大小达到一个数据块的存储空间大小的第一预设比例，且未超过一个数据块的存储空间大小时，将所述包含同一个索引列值的记录存储在一个数据块中；

当包含同一个索引列值的记录的总大小超过一个数据块的存储空间大小时，将所

述包含同一个索引列值的记录存储在多个数据块中；

当包含同一个索引列值的记录的总大小小于一个数据块的存储空间大小的第二预设比例时，将多个所述包含同一个索引列值的记录存储在一个数据块中。

6、根据权利要求1至5任一项所述的处理连接查询的方法，其特征在于，表组合的表示形式为：

$$TG = (tab_1, \dots, tab_i, \dots, tab_N) \text{ key} = (key_1, \dots, key_j, \dots, key_M)$$

其中， tab_i 为表组合中的第 i 张表， key_j 为第 j 个连接键， N 为表组合中表的个数， M 为表组合中连接键的个数。

7、一种处理连接查询的方法，其特征在于，包括：

接收查询请求，所述查询请求中包含待查询表组合；

在分布式文件系统 DFS 中查找与所述待查询表组合对应的待查询表簇，所述 DFS 中存储了频繁表组合对应的表簇，所述频繁表组合为在历史查询记录中的出现频率大于预设值的表组合，所述表组合包括连接键以及通过连接键进行连接的表，所述频繁表组合中连接键的数量与簇索引中索引列的数量相同，所述簇索引用于指示所述频繁表组合中索引列值相同的记录的存储位置，其中，所述频繁表组合中索引列值相同的记录集中存放在至少一个数据块中，形成所述频繁表组合对应的表簇；

分别将所述待查询表簇对应的节点所包含的每个数据块中索引列值的哈希值相同的记录划分至一个分桶中；

将所述待查询表簇中所有数据块中相同分桶中的记录传输至同一数据节点中。

8、一种处理连接查询的装置，其特征在于，包括：

确定单元，用于确定频繁表组合，所述频繁表组合为在历史查询记录中的出现频率大于预设值的表组合，所述表组合包括连接键以及通过连接键进行连接的表；

创建单元，用于根据所述确定单元确定的所述频繁表组合中的连接键信息创建簇索引，所述簇索引中索引列的数量与所述频繁表组合中连接键的数量相同；

洗牌单元，用于根据所述创建单元创建的所述簇索引中的索引列进行洗牌操作，将索引列值相同的记录集中存放在至少一个数据块中，以形成所述频繁表组合对应的表簇，所述索引列值相同的记录为通过所述连接键进行连接的表中的记录。

9、根据权利要求8所述的处理连接查询的装置，其特征在于，所述装置还包括

接收单元，用于接收查询请求，所述查询请求中包含待查询表组合；

查找单元，用于查找所述待查询表组合对应的待查询表簇；

分桶单元，用于分别将所述待查询表簇对应的节点所包含的每个数据块中索引列值的哈希值相同的记录划分至一个分桶中；

传输单元，用于根据所述分桶单元的分桶结果，将所述待查询表簇中所有数据块中相同分桶中的记录传输至同一数据节点中。

10、根据权利要求9所述的处理连接查询的装置，其特征在于，

所述确定单元，还用于从所述历史查询记录中提取表组合，生成表组合集合；从所述表组合集合中筛选出出现频率大于预设值的表组合；从出现频率大于预设值的表组合中删除冗余表组合，确定剩余的表组合为频繁表组合。

11、根据权利要求10所述的处理连接查询的装置，其特征在于，

所述确定单元，还用于当存在包含的表相同，但是连接键不同的至少两个表组合时，保留连接键最多的表组合；当存在两个表组合所包含的连接键相同，且一个表组合中的表组成的集合是另一个表组合中表组成的集合的子集时，删除包含表较少的表组合；当存在相同表包含于至少两个表组合中时，只将所述相同表保留于所述至少两个表组合中出现频率最高的一个表组合中。

12、根据权利要求8至11任一项所述的处理连接查询的装置，其特征在于，

所述洗牌单元，还用于当包含同一个索引列值的记录的总大小达到一个数据块的存储空间大小的第一预设比例，且未超过一个数据块的存储空间大小时，将所述包含同一个索引列值的记录存储在一个数据块中；当包含同一个索引列值的记录的总大小超过一个数据块的存储空间大小时，将所述包含同一个索引列值的记录存储在多个数据块中；当包含同一个索引列值的记录的总大小小于一个数据块的存储空间大小的第二预设比例时，将多个所述包含同一个索引列值的记录存储在一个数据块中。

13、根据权利要求8至12任一项所述的处理连接查询的装置，其特征在于，表组合的表示形式为：

$$TG = (tab_1, \dots, tab_i, \dots, tab_N) \text{ key} = (key_1, \dots, key_j, \dots, key_M)$$

其中， tab_i 为表组合中的第 i 张表， key_j 为第 j 个连接键， N 为表组合中表的个数， M 为表组合中连接键的个数。

14、一种连处理接查询的装置，其特征在于，包括：

接收单元，用于接收查询请求，所述查询请求中包含待查询表组合；

查找单元，用于在分布式文件系统 DFS 中查找与所述接收单元接收的所述待查询表组合对应的待查询表簇，所述 DFS 中存储了频繁表组合对应的表簇，所述频繁表组合为在历史查询记录中的出现频率大于预设值的表组合，所述表组合包括连接键以及通过连接键进行连接的表，所述频繁表组合中连接键的数量与簇索引中索引列的数量相同，所述簇索引用于指示所述频繁表组合中索引列值相同的记录的存储位置，其中，所述频繁表组合中索引列值相同的记录集中存放在至少一个数据块中，形成所述频繁表组合对应的表簇；

分桶单元，用于分别将所述待查询表簇对应的节点所包含的每个数据块中索引列值的哈希值相同的记录划分至一个分桶中；

传输单元，用于根据所述分桶单元的分桶结果将所述待查询表簇中所有数据块中相同分桶中的记录传输至同一数据节点中。

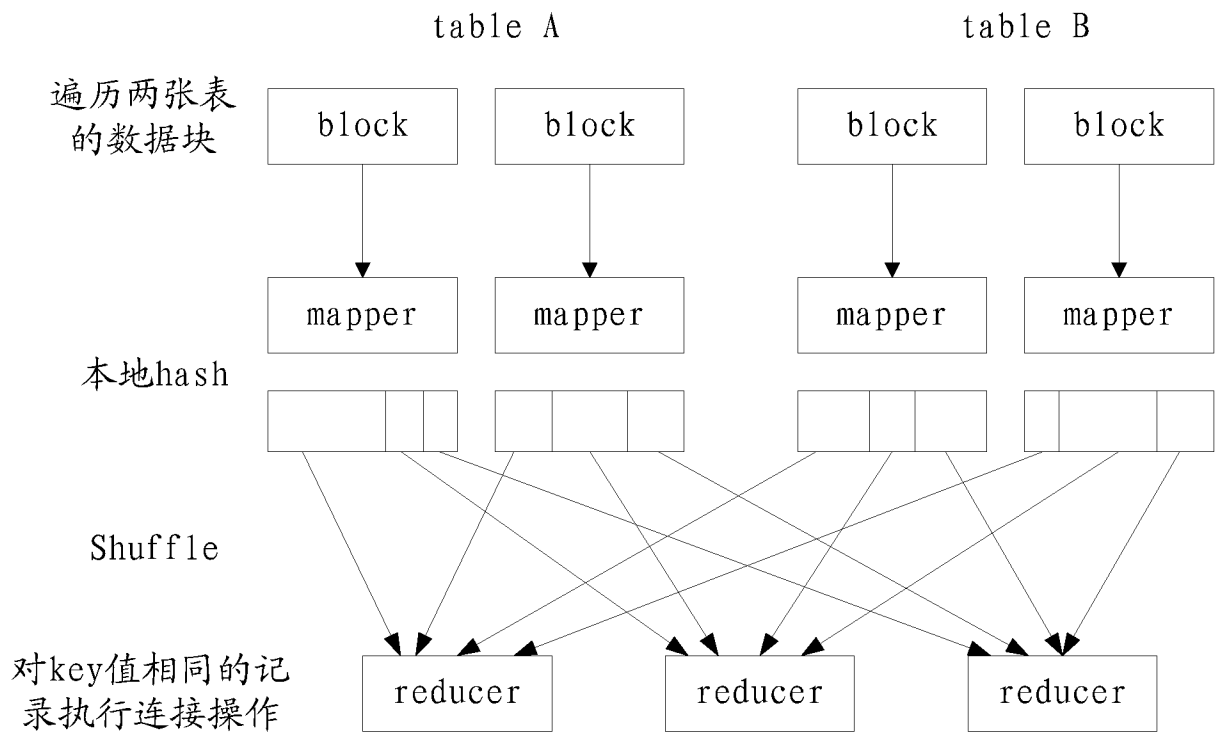


图 1

2/7

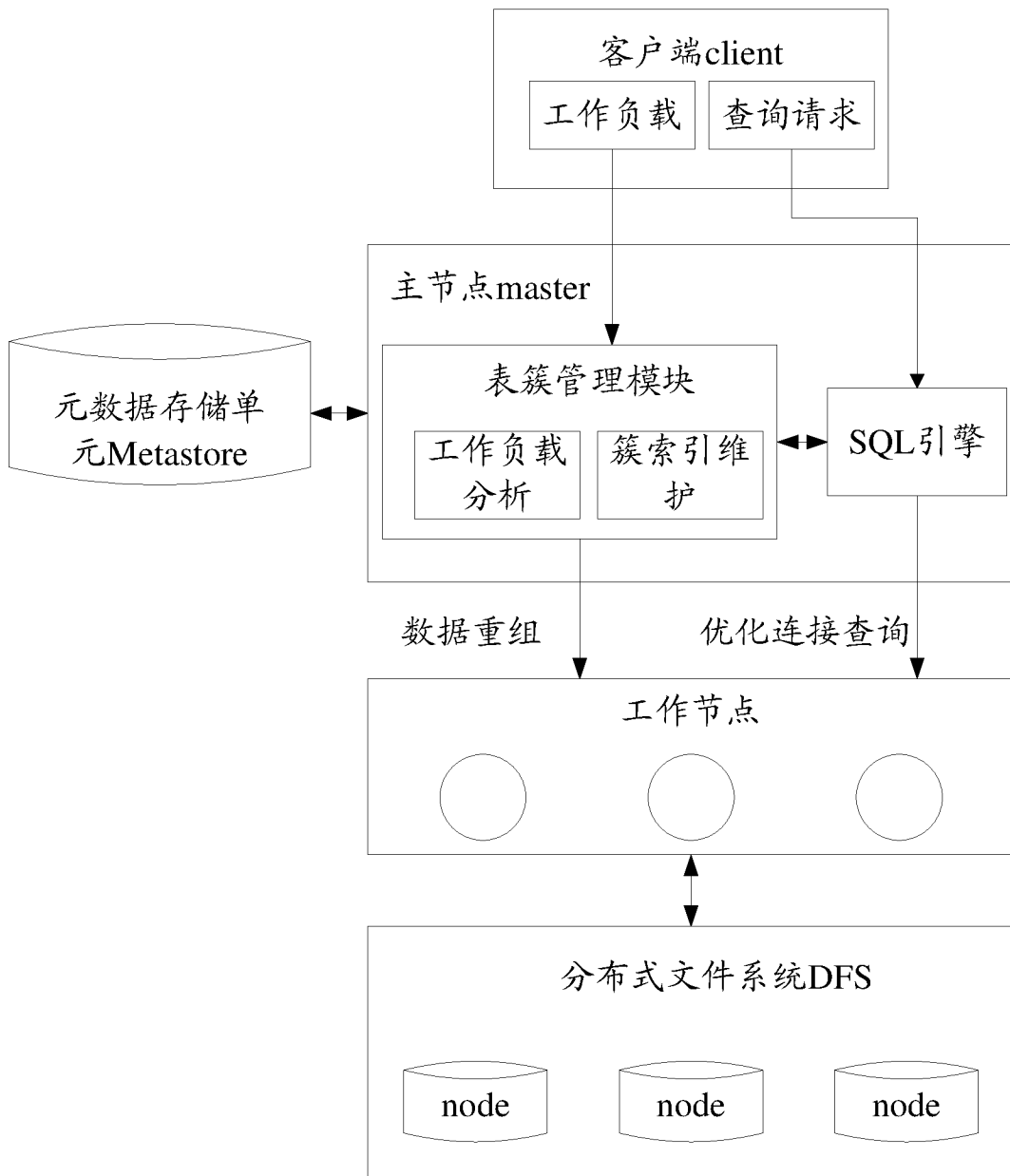


图 2

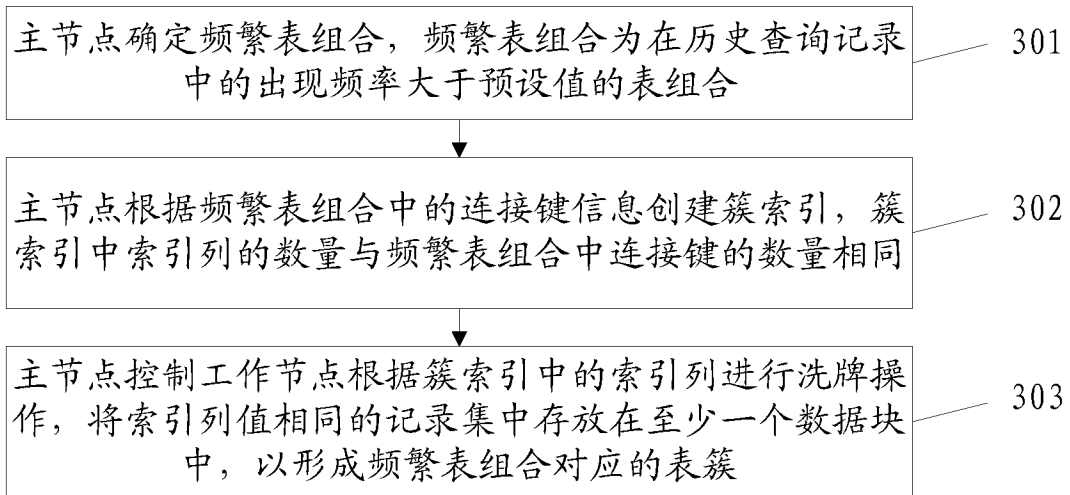


图 3

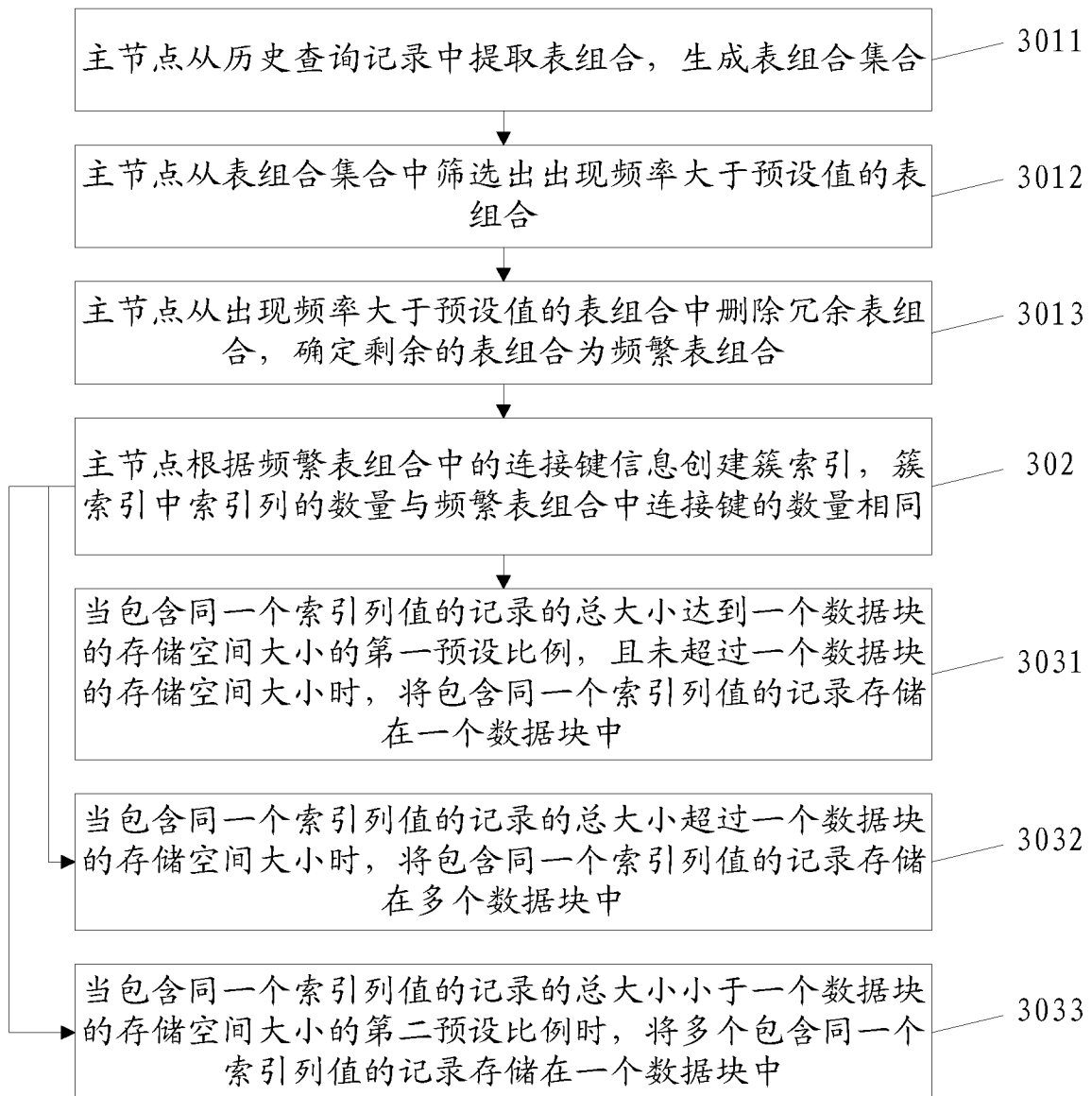


图 4

4/7

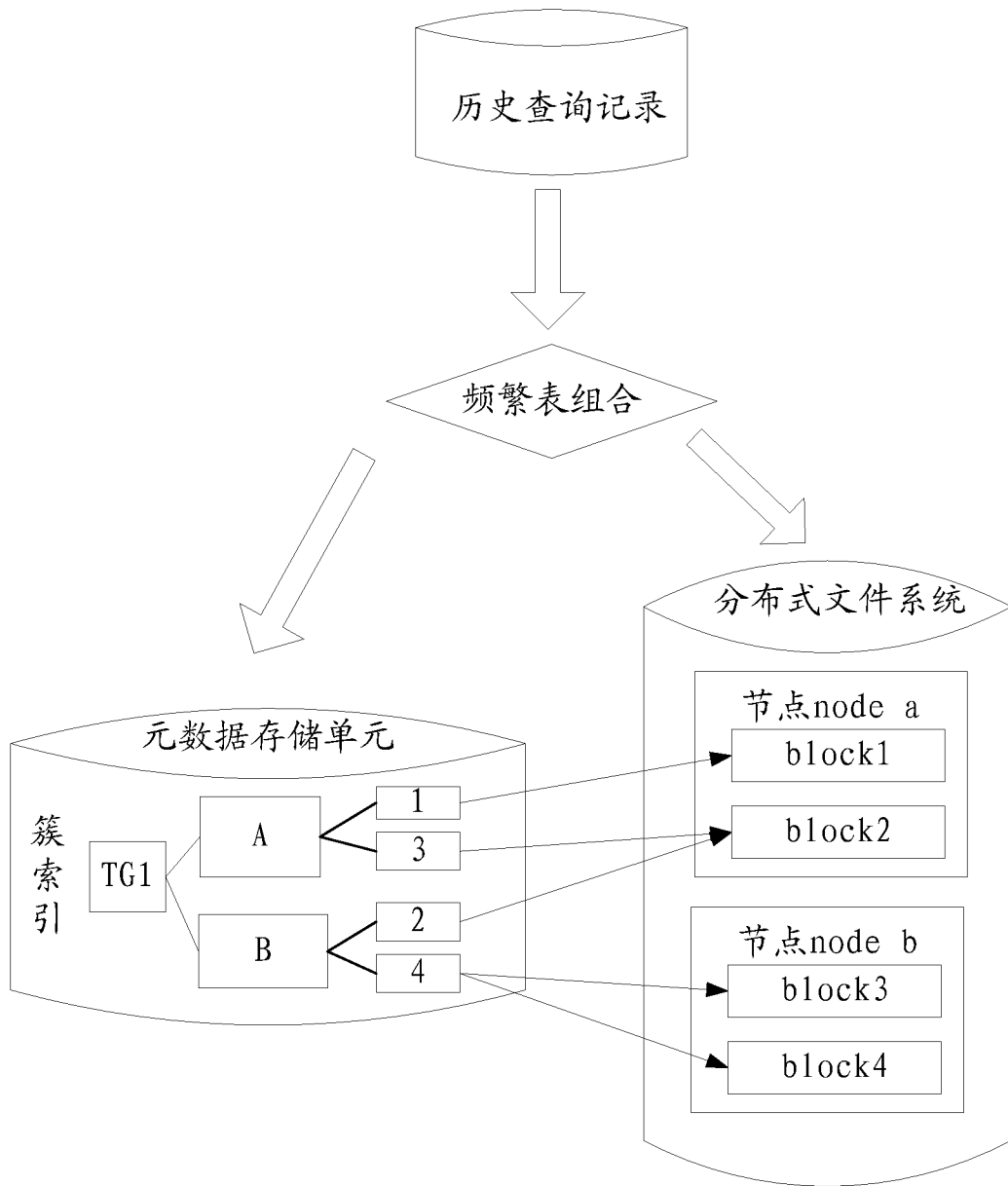


图 5

5/7

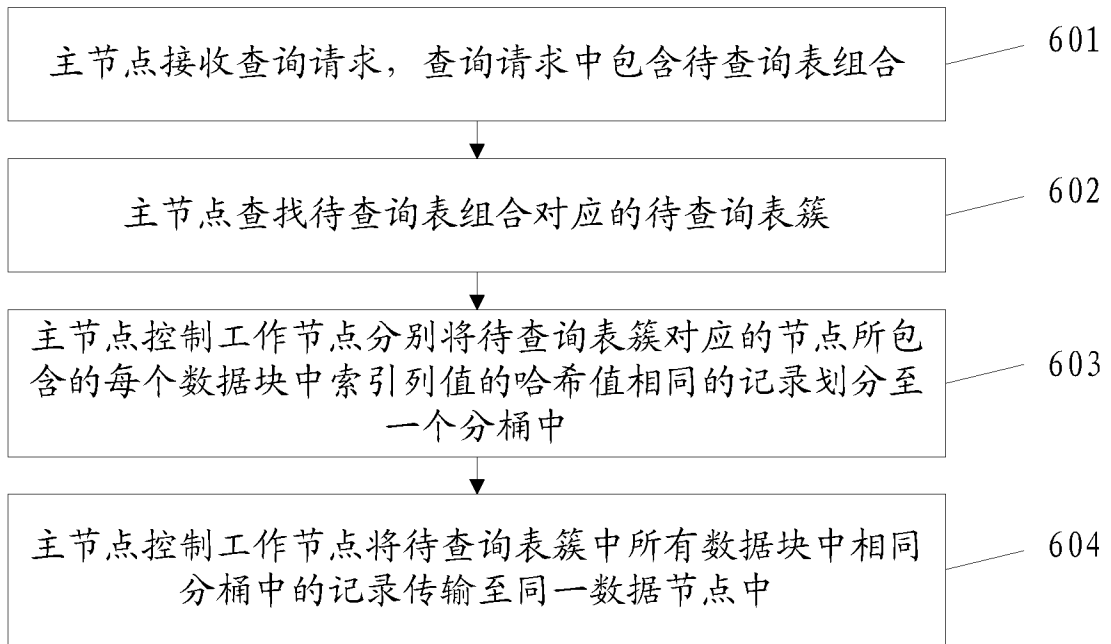


图 6

表簇

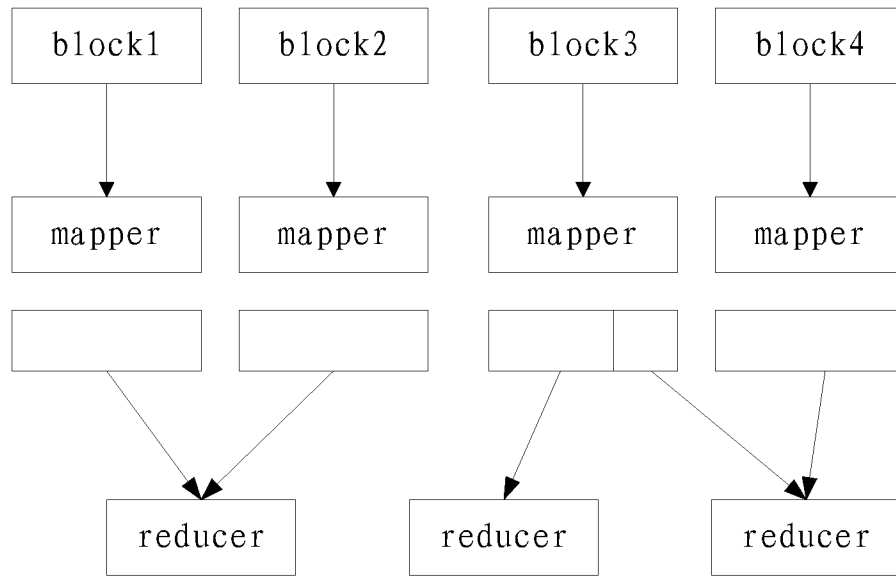


图 7

6/7

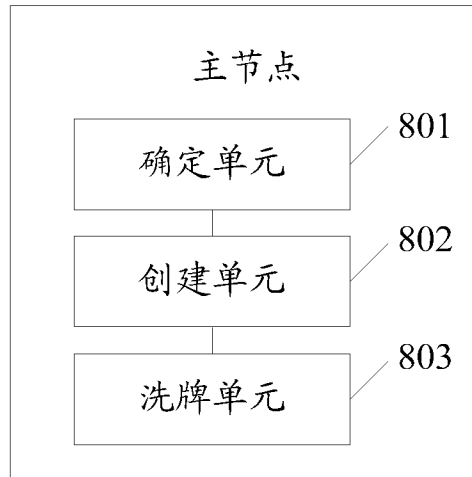


图 8

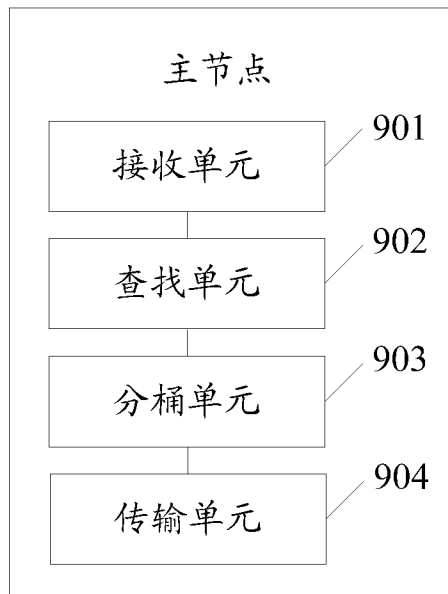


图 9

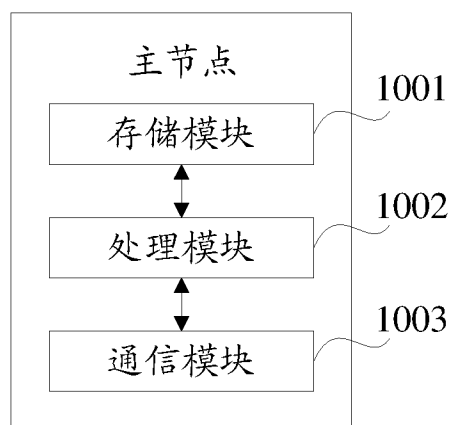


图 10

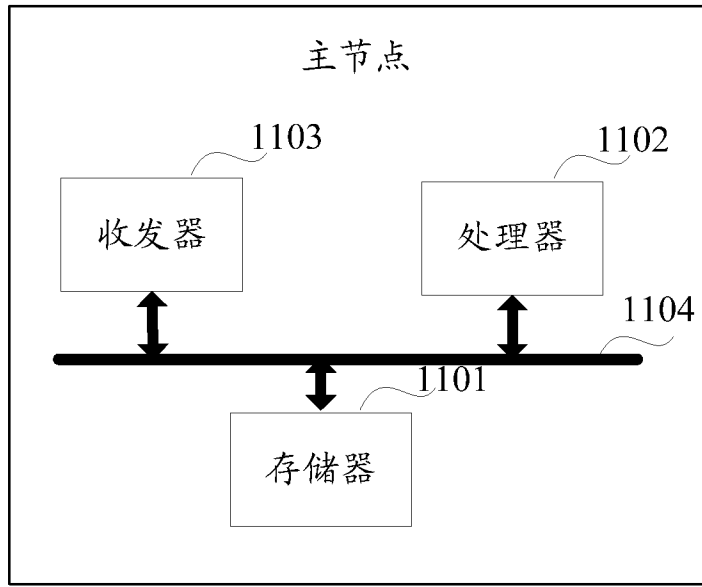


图 11

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2017/071568

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; H04L

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, CNKI, EPODOC, WPI, IEEE: 中兴, 华为, 分布, 数据, 连接, 查询, 历史, 表, 散列, 洗牌, 索引, 分桶, distributed, join, joint, history, table, map, reduce, DFS, shuffle, index, cluster

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 105357311 A (CHINA SOUTHERN POWER GRID CO., LTD.) 24 February 2016 (24.02.2016), description, paragraphs [0039]-[0045]	1-14
A	CN 105095515 A (BEIJING KINGSOFT SECURITY SOFTWARE CO., LTD.) 25 November 2015 (25.11.2015), entire document	1-14
A	CN 103488657 A (HUAWEI TECHNOLOGIES CO., LTD.) 01 January 2014 (01.01.2014), entire document	1-14
A	US 2009055370 A1 (BUSINESS. COM) 26 February 2009 (26.02.2009), entire document	1-14
A	CN 102426609 A (XIAMEN MEIYA PICO INFORMATION CO., LTD.) 25 April 2012 (25.04.2012), entire document	1-14

Further documents are listed in the continuation of Box C.

See patent family annex.

<p>* Special categories of cited documents:</p> <p>“A” document defining the general state of the art which is not considered to be of particular relevance</p> <p>“E” earlier application or patent but published on or after the international filing date</p> <p>“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>“O” document referring to an oral disclosure, use, exhibition or other means</p> <p>“P” document published prior to the international filing date but later than the priority date claimed</p>	<p>“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</p> <p>“&” document member of the same patent family</p>
---	---

<p>Date of the actual completion of the international search</p> <p style="text-align: center;">17 April 2017</p>	<p>Date of mailing of the international search report</p> <p style="text-align: center;">23 May 2017</p>
<p>Name and mailing address of the ISA</p> <p>State Intellectual Property Office of the P. R. China</p> <p>No. 6, Xitucheng Road, Jimenqiao</p> <p>Haidian District, Beijing 100088, China</p> <p>Facsimile No. (86-10) 62019451</p>	<p>Authorized officer</p> <p style="text-align: center;">LI, Weihua</p> <p>Telephone No. (86-10) 82246738</p>

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2017/071568

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 105357311 A	24 February 2016	None	
CN 105095515 A	25 November 2015	None	
CN 103488657 A	01 January 2014	None	
US 2009055370 A1	26 February 2009	WO 2010042238 A1	15 April 2010
CN 102426609 A	25 April 2012	None	

<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; H04L</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>CNPAT, CNKI, EPODOC, WPI, IEEE: 中兴, 华为, 分布, 数据, 连接, 查询, 历史, 表, 散列, 洗牌, 索引, 分桶, distributed, join, joint, history, table, map, reduce, DFS, shuffle, index, cluster</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 105357311 A (中国南方电网有限责任公司) 2016年 2月 24日 (2016-02-24) 说明书第[0039]-[0045]段</td> <td>1-14</td> </tr> <tr> <td>A</td> <td>CN 105095515 A (北京金山安全软件有限公司) 2015年 11月 25日 (2015-11-25) 全文</td> <td>1-14</td> </tr> <tr> <td>A</td> <td>CN 103488657 A (华为技术有限公司) 2014年 1月 1日 (2014-01-01) 全文</td> <td>1-14</td> </tr> <tr> <td>A</td> <td>US 2009055370 A1 (BUSINESS.COM) 2009年 2月 26日 (2009-02-26) 全文</td> <td>1-14</td> </tr> <tr> <td>A</td> <td>CN 102426609 A (厦门市美亚柏科信息股份有限公司) 2012年 4月 25日 (2012-04-25) 全文</td> <td>1-14</td> </tr> </tbody> </table> <p><input type="checkbox"/> 其余文件在C栏的续页中列出。 <input checked="" type="checkbox"/> 见同族专利附件。</p> <p>* 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</p>			类型*	引用文件, 必要时, 指明相关段落	相关的权利要求	A	CN 105357311 A (中国南方电网有限责任公司) 2016年 2月 24日 (2016-02-24) 说明书第[0039]-[0045]段	1-14	A	CN 105095515 A (北京金山安全软件有限公司) 2015年 11月 25日 (2015-11-25) 全文	1-14	A	CN 103488657 A (华为技术有限公司) 2014年 1月 1日 (2014-01-01) 全文	1-14	A	US 2009055370 A1 (BUSINESS.COM) 2009年 2月 26日 (2009-02-26) 全文	1-14	A	CN 102426609 A (厦门市美亚柏科信息股份有限公司) 2012年 4月 25日 (2012-04-25) 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 105357311 A (中国南方电网有限责任公司) 2016年 2月 24日 (2016-02-24) 说明书第[0039]-[0045]段	1-14																		
A	CN 105095515 A (北京金山安全软件有限公司) 2015年 11月 25日 (2015-11-25) 全文	1-14																		
A	CN 103488657 A (华为技术有限公司) 2014年 1月 1日 (2014-01-01) 全文	1-14																		
A	US 2009055370 A1 (BUSINESS.COM) 2009年 2月 26日 (2009-02-26) 全文	1-14																		
A	CN 102426609 A (厦门市美亚柏科信息股份有限公司) 2012年 4月 25日 (2012-04-25) 全文	1-14																		
国际检索实际完成的日期	国际检索报告邮寄日期																			
2017年 4月 17日	2017年 5月 23日																			
ISA/CN的名称和邮寄地址	授权官员																			
中华人民共和国国家知识产权局 (ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088	李伟华																			
传真号 (86-10) 62019451	电话号码 (86-10) 82246738																			

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2017/071568

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	105357311	A	2016年 2月 24日	无			
CN	105095515	A	2015年 11月 25日	无			
CN	103488657	A	2014年 1月 1日	无			
US	2009055370	A1	2009年 2月 26日	WO	2010042238	A1	2010年 4月 15日
CN	102426609	A	2012年 4月 25日	无			