

(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
2014 年 4 月 24 日 (24.04.2014)

WIPO | PCT

(10) 国际公布号
WO 2014/059808 A1

- (51) 国际专利分类号:
G06F 17/30 (2006.01)
- (21) 国际申请号:
PCT/CN2013/079993
- (22) 国际申请日:
2013 年 7 月 24 日 (24.07.2013)
- (25) 申请语言:
中文
- (26) 公布语言:
中文
- (30) 优先权:
201210401421.6 2012 年 10 月 19 日 (19.10.2012) CN
- (71) 申请人 (对除美国外的所有指定国): 烽火通信科技股份有限公司 (FIBERHOME TELECOMMUNICATION TECHNOLOGIES CO., LTD) [CN/CN]; 中国湖北省武汉市东湖开发区关东科技园东信路 5 号, Hubei 430074 (CN)。
- (72) 发明人: 及
- (71) 申请人 (仅对美国): 邹昊 (ZOU, Hao) [CN/CN]; 中国湖北省武汉市东湖开发区关东科技园东信路 5 号, Hubei 430074 (CN)。

- (74) 代理人: 北京捷诚信通专利事务所 (普通合伙) (BEIJING PSCU PATENT OFFICE); 中国北京市西城区三里河一区 5-5, Beijing 100045 (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, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, 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, 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, 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)。

[见续页]

(54) Title: METHOD FOR MANAGING MASS HISTORICAL DATA OF NETWORK MANAGEMENT

(54) 发明名称: 一种网管海量历史数据管理方法

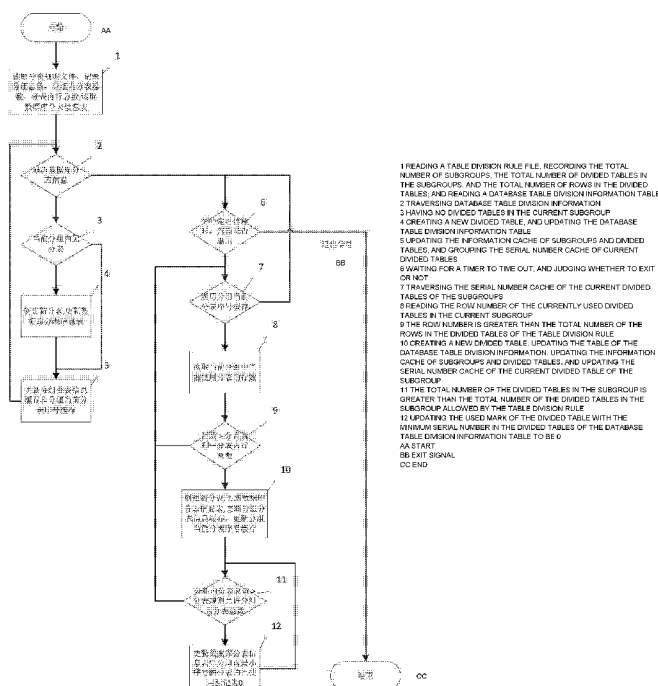


图 4 / Fig. 4

(57) Abstract: The present invention relates to management of historical data in a management system of network management, and particularly to a method for managing mass historical data of network management. A database is divided into tables via a second-level table division rule, the data stored in a single table in which mass historical data of network management is stored is segmented and then stored in a plurality of divided tables with the same structure so that the capacity of the single table is reduced, thereby improving the processing efficiency of the single table, and each divided table is managed as follows: A management of historical data in divided tables; B insertion management of historical data; C query management of historical data; and D transfer management of historical data. The method for managing mass historical data of network management described in the present invention can save mass historical data, and can efficiently perform data insertion, query and transfer without affecting the normal use of network management software, thereby improving the working efficiency of a management system of network management.

(57) 摘要:

[见续页]

WO 2014/059808 A1

根据细则 4.17 的声明:

— 关于申请人有权申请并被授予专利(细则 4.17(ii))

本国际公布:

— 包括国际检索报告(条约第 21 条(3))。

本发明涉及网管管理系统中历史数据的管理，具体说是一种网管海量历史数据管理方法，通过二级分表规则实现数据库分表，把保存有网管海量历史数据的单个表保存的数据，切分后保存到多个结构相同的分表中，减少单表的容量，从而提高单表的处理效率，并对各个分表进行以下管理：A 历史数据分表管理，B 历史数据插入管理，C 历史数据查询管理，D 历史数据转移管理。本发明所述的网管海量历史数据管理方法，能保存海量历史数据，并能高效的进行数据的插入、查询、转移而不影响网管软件的正常使用，提高了网管管理系统的工作效率。

一种网管海量历史数据管理方法

技术领域

本发明涉及网管管理系统中历史数据的管理，具体说是一种网管海量历史数据管理方法。

背景技术

目前网络技术日新月异，宽带用户规模不断扩大，电信网管系统管理的设备数量也在不断增长，电信网管系统日常管理的日志、设备的告警数据等数据都需要保存在数据库中，电信网管系统对这些数据的使用方式为记录数据、查询分析数据。经过日积月累，这些数据越来越庞大，如果不对这些数据进行有效的管理，无论是数据的插入或者是数据的查询，性能都会急剧下降，从而影响网管软件的正常使用。数据库使用硬盘空间来保存其中的数据，如果历史数据无限增长，最终将耗尽所有磁盘空间，所以需要定期对历史数据进行清理，将数据库中的一部分不常使用的数据转储到其它位置进行保存、归档。而在数据转移的过程中，网管则无法使用这些数据，如果转储的效率低下则影响网管系统的可用性，必须提供一种高效的数据转移方案。

发明内容

针对现有技术中存在的缺陷，本发明的目的在于提供一种网管海量历史数据管理方法，能保存海量历史数据，并能高效的进行数据的插入、查询、转移而不影响网管软件的正常使用。

为达到以上目的，本发明采取的技术方案是：

一种网管海量历史数据管理方法，其特征在于：通过二级分表规则实现数据库分表，把保存有网管海量历史数据的单个表保存的数据，切分后保存到多个结构相同的分表中，减少单表的容量，从而提

高单表的处理效率，并对各个分表进行以下管理：

- A 历史数据分表管理，
- B 历史数据插入管理，
- C 历史数据查询管理，
- D 历史数据转移管理。

在上述技术方案的基础上，二级分表规则用 xml 格式的文件记录，其中：

历史数据类型 data_type，表示数据表保存数据的类型，
分组总数 group_count，表示对这种数据表分组的总数，
分组内分表总数 split_count，表示每个分组内分表数最大值，
分表内行总数 row_count，表示分表内行总数，当一个分表的数据量超过这个值的时候，就需要创建一个新的分表。

在上述技术方案的基础上，所述历史数据分表管理的处理流程包括以下步骤：

A1、读取记载有数据表的二级分表规则的文件，获取历史数据类型 data_type、分组总数 group_count、分组内分表总数 split_count 和分表内行总数 row_count；

读取数据库分表信息表，获取所有分表的分组号 group_no、分表序号 table_no、当前正在使用标记 current_flag 和已使用标记 in_use_flag，将这些信息保存到一个分组号为键，分表信息列表为值的 map 对象中，所述分表信息列表包括分表序号 table_no、当前正在使用标记 current_flag 和已使用标记 in_use_flag，其中 map 的键的取值范围从 1 到分组总数 group_count，如果在数据库分表信息表中没有某个键对应的分表信息，那么在 map 中该键对应的分表信息列表就为空；

A2、遍历 A1 得到的 map 对象；

A3、根据遍历的结果，如果 map 对象中当前分组号对应的分表信息列表为空，则转至步骤 A4，否则，转至步骤 A5；

A4、在当前分组号的分组中创建一个可用的分表，分配分表序号

为 1, 设置分表当前正在使用标记为 1, 已使用标记为 1, 将该分表的信息插入数据库分表信息表后转至步骤 A5;

A5、将数据库分表信息表中记载的分表信息, 按分组号为键, 逐条插入到分组分表信息缓存中, 该缓存用于记录所有分组中已创建的分表的信息,

将当前正在使用标记为 1 的分表信息, 按分组号为键, 将分表序号插入分组当前分表序号缓存, 该缓存记录每个分组中 `current_flag` 为 1 的分表的分表序号, 转步骤 A6;

分组当前分表序号缓存内的分表, 需要检查其是否超过分表内行总数;

A6、使用定时器并等待, 如果定时器未超时则一直等待, 当定时器超时后, 如果接收到应用程序退出的信号, 则退出历史数据分表管理的处理流程; 否则, 转入步骤 A7;

A7、遍历分组当前分表序号缓存, 如果未遍历完则转步骤 A8, 若遍历完则转步骤 A6;

A8、读取当前分组中分表序号, 根据当前分组的组号和分表序号构造分表的名称, 获取分表数据行数;

A9、如果分表数据行数大于分表规则中分表内行总数 `row_count`, 则转步骤 A10、否则转步骤 A7;

A10、创建新分表, 新分表的分表序号为分表所在分组中原来 `current_flag` 为 1 的分表的分表序号值加上 1, 然后在数据库分表信息表中插入该新分表的信息, 并将该新分表的信息插入到分组分表信息缓存中, 然后将分组当前分表序号缓存中的当前分组对应的分表序号改为新分表的分表序号;

A11、判断当前分组中分表总个数是否大于分组内分表总数 `split_count`, 如果大于则转步骤 A12, 如果小于则转步骤 A7;

A12、更新数据库分表信息表中该分组内分表序号最小分表的已使用标记 `in_use_flag` 为 0, 标识该分表不会在数据插入和查询中使用, 并将被历史数据转移管理流程处理, 并从分组分表信息缓存中删

除该分表的记录，转步骤 A11。

在上述技术方案的基础上，所述历史数据插入管理的处理流程包括以下步骤：

B1、根据插入数据的内容确定数据业务分组的组号，根据分组当前分表序号缓存获取分表序号，根据分表和数据逻辑对应关系，重新排列插入数据内容，存入以分组号和分表序号为键，插入数据内存为值的 map 中；

B2、遍历 B1 中的 map，遍历完成则退出，否则转步骤 B3；

B3、根据当前分组号和分表序号，构造分表名称，插入数据到当前分表中，转步骤 B2。

在上述技术方案的基础上，所述历史数据查询管理的处理流程包括以下步骤：

C1、读取数据库分表信息表内已使用标记为 1 的分表信息；

C2、分析查询条件、获取所有相关分组分表的表名列表；

C3、根据查询条件查询所有 C2 中的分表。

在上述技术方案的基础上，所述历史数据转移管理的处理流程包括以下步骤：

D1、按分组号读取数据库分表信息表中所有已使用标记为 0 的分表信息记录，保存到以分组号为键，分表序号列表为值的 map 中；

D2、遍历 D1 中获取的 map，遍历完成后退出；

D3、根据当前分组号和分表序号，构造分表名称，使用批处理工具直接导出指定分表名称的分表中所有行记录；

D4、根据分组号和分表序号从数据库分表信息表中删除分表信息；

D5、根据分表名称从数据库中删除分表，转步骤 D2。

本发明所述的网管海量历史数据管理方法，通过二级分表规则实现数据库分表（数据库分表即：把单个表保存的数据，按照一定规则切分后保存到多个结构相同的表（称为分表）中，减少单表的容量，

从而提高单表的处理效率，在不同业务需求下可以访问不同的分表，减少锁竞争，提高运行效率)，将历史数据分散到多个分表中，且保证每个分表中记录数大致与限定的分表容量相差不大，不仅保证了历史数据查询的响应效率，还保证了数据的插入效率，提升了客户体验感受，提高了网管管理系统的工作效率。

本发明所述的网管海量历史数据管理方法，通过已使用标记将正在使用的和需要转移的数据表进行分离，网管系统和数据转移功能就不会相互影响，而且数据表整表处理的效率远远高于按行记录进行处理的效率，减少了对数据库系统的影响，提高了网管系统整体的效率。

附图说明

本发明有如下附图：

图 1 是根据分表规则划分后，分表展示；

图 2 是数据表分表规则文件的内容；

图 3 是数据库分表信息表包含的数据内容；

图 4 是分表管理处理流程图；

图 5 是数据插入处理流程图；

图 6 是数据查询处理流程图；

图 7 是数据转移处理流程图。

具体实施方式

以下结合附图对本发明作进一步详细说明。

本发明所述的网管海量历史数据管理方法，通过二级分表规则实现数据库分表，把保存有网管海量历史数据的单个表保存的数据，切分后保存到多个结构相同的分表中，减少单表的容量，从而提高单表的处理效率，并对各个分表进行以下管理：

A 历史数据分表管理，

B 历史数据插入管理，

- C 历史数据查询管理，
- D 历史数据转移管理。

在上述技术方案的基础上，图 2 示出了数据表分表规则文件的内容，此即所述的二级分表规则，二级分表规则用 xml 格式的文件记录，其中：

历史数据类型 `data_type`，表示数据表保存数据的类型，

分组总数 `group_count`，表示对这种数据表分组的总数，即表示一个数据表最多可以分成几个分组，每个分组对应唯一的一个分组号；

分组内分表总数 `split_count`，表示每个分组内分表数最大值，即表示一个分组内可以包括多少个分表，即分组内分表数最大值，每个分表对应唯一的一个分表序号；

分表内行总数 `row_count`，表示分表内行总数，即分表内行总数 `row_count` 表示一个分表内可以包括多少行数据，即分表中数据行最大值；当一个分表的数据量超过这个值的时候，就需要创建一个新的分表。

即：一个数据表最多会有 `group_count` 个分组，一个分组内有 `split_count` 个分表，一个分表最多纪录 `row_count` 行数据。

在上述技术方案的基础上，如图 4 所示，所述历史数据分表管理的处理流程包括以下步骤：

A1、读取记载有数据表的二级分表规则的文件，获取历史数据类型 `data_type`、分组总数 `group_count`、分组内分表总数 `split_count` 和分表内行总数 `row_count`；

读取数据库分表信息表，获取所有分表的分组号 `group_no`、分表序号 `table_no`、当前正在使用标记 `current_flag` 和已使用标记 `in_use_flag`，将这些信息保存到一个分组号为键，分表信息列表为值的 `map` 对象中，所述分表信息列表包括分表序号 `table_no`、当前

正在使用标记 `current_flag` 和已使用标记 `in_use_flag`, 其中 `map` 的键的取值范围从 1 到分组总数 `group_count`, 如果在数据库分表信息表中没有某个键对应的分表信息, 那么在 `map` 中该键对应的分表信息列表就为空;

数据库分表信息表的格式及内容可参见图 3 所示实施例; 图 3 示出了数据库分表信息表的内容, 并演示了一个可能出现的状态, 每一行记录一个分表的信息; 假设分表规则按照业务功能划分的组 `group_count = 3`, 每个组内最多允许 `split_count = 3` 个分表, 那么当创建了第 4 个分表后, 这个分表序号 `table_no` 为 4 的分表就成为当前使用状态 `current_flag` 为 1 的表, 且分表编号 `table_no` 为 1 的分表的已使用标记 `in_use_flag` 则被置为 0, 这样就可以保持所有可用的分表的总数不超过 `split_count` 的值;

A2、遍历 A1 得到的 `map` 对象;

A3、根据遍历的结果, 如果 `map` 对象中当前分组号对应的分表信息列表为空, 则转至步骤 A4, 否则, 转至步骤 A5;

A4、在当前分组号的分组中创建一个可用的分表, 分配分表序号为 1, 设置分表当前正在使用标记为 1, 已使用标记为 1, 将该分表的信息插入数据库分表信息表后转至步骤 A5;

A5、将数据库分表信息表中记载的分表信息, 按分组号为键, 逐条插入到分组分表信息缓存中, 该缓存用于记录所有分组中已创建的分表的信息,

将当前正在使用标记为 1 的分表信息, 按分组号为键, 将分表序号插入分组当前分表序号缓存, 该缓存记录每个分组中 `current_flag` 为 1 的分表的分表序号, 转步骤 A6;

分组当前分表序号缓存内的分表, 需要检查其是否超过分表内行总数;

A6、使用定时器并等待, 如果定时器未超时则一直等待, 当定时器超时后, 如果接收到应用程序退出的信号, 则退出历史数据分表管理的处理流程; 否则, 转入步骤 A7;

A7、遍历分组当前分表序号缓存，如果未遍历完则转步骤 A8，若遍历完则转步骤 A6；

A8、读取当前分组中分表序号，根据当前分组的组号和分表序号构造分表的名称，获取分表数据行数；

A9、如果分表数据行数大于分表规则中分表内行总数 row_count，则转步骤 A10、否则转步骤 A7；

A10、创建新分表，新分表的分表序号为分表所在分组中原来 current_flag 为 1 的分表的分表序号值加上 1，然后在数据库分表信息表中插入该新分表的信息，并将该新分表的信息插入到分组分表信息缓存中，然后将分组当前分表序号缓存中的当前分组对应的分表序号改为新分表的分表序号；

A11、判断当前分组中分表总个数是否大于分组内分表总数 split_count，如果大于则转步骤 A12，如果小于则转步骤 A7；

A12、更新数据库分表信息表中该分组内分表序号最小分表的已使用标记 in_use_flag 为 0，标识该分表不会在数据插入和查询中使用，并将被历史数据转移管理流程处理，并从分组分表信息缓存中删除该分表的记录，转步骤 A11。

在上述技术方案的基础上，如图 5 所示，所述历史数据插入管理的处理流程包括以下步骤：

B1、根据插入数据的内容确定数据业务分组的组号，根据分组当前分表序号缓存获取分表序号，根据分表和数据逻辑对应关系，重新排列插入数据内容，存入以分组号和分表序号为键，插入数据内存为值的 map 中；

B2、遍历 B1 中的 map，遍历完成则退出，否则转步骤 B3；

B3、根据当前分组号和分表序号，构造分表名称，插入数据到当前分表中，转步骤 B2。

在上述技术方案的基础上，如图 6 所示，所述历史数据查询管理

的处理流程包括以下步骤：

- C1、读取数据库分表信息表内已使用标记为 1 的分表信息；
- C2、分析查询条件、获取所有相关分组分表的表名列表；
- C3、根据查询条件查询所有 C2 中的分表。

在上述技术方案的基础上，如图 7 所示，所述历史数据转移管理的处理流程包括以下步骤：

- D1、按分组号读取数据库分表信息表中所有已使用标记为 0 的分表信息记录，保存到以分组号为键，分表序号列表为值的 map 中；
- D2、遍历 D1 中获取的 map，遍历完成后退出；
- D3、根据当前分组号和分表序号，构造分表名称，使用批处理工具直接导出指定分表名称的分表中所有行记录；
- D4、根据分组号和分表序号从数据库分表信息表中删除分表信息；
- D5、根据分表名称从数据库中删除分表，转步骤 D2。

图 1 示出了按照分表规则划分后可能出现的分表状态，所述的分表规则通常存储于静态文件中（例如：存储于数据表分表规则文件中），纵向表示按照业务功能划分的一个分组，如图划分了 M 个分组。每个分组内可能有多个分表，分表序号从 1 开始，向后排列。每个组内的分表数量不一定相同，而是根据实际数据的数量确定，使用分组号和分表序号对的方式可以唯一标识一个分表，例如 (2, 1) 标识分组 2 中序号为 1 的分表。

以下为一个具体实施例。

图 4 示出了网管海量历史数据分表管理方法：

A1、从静态文件读取数据表分表规则，从数据库分表信息表读取已使用标记 in_use_flag 为 1 的分表信息。

分表管理只管理已使用标记为 1 的分表，而已使用标记为 0 的分

表由数据转移功能使用。

A2、按分组号遍历从数据库分表信息表读取的分表信息，遍历完成转步骤 A6，否则转步骤 A3。

遍历进行分组分表信息缓存和分组当前分表序号缓存的初始化，完成后进行定时的分表管理步骤。

A3、若 A2 中分组号为 1 的分组内一个分表都没有，则转步骤 A4，若分组号为 1 的分组有至少一个分表，则转步骤 A5；

第一次使用网管时，历史数据是没有创建任何分表的，所以会进入步骤 A4。

A4、创建一个新的序号 `table_no` 为 1 分表，更新数据库分表信息表，置分表序号为 1 的当前正在使用标记 `current_flag` 为 1，每个分组内只有一个分表的当前正在使用标记 `current_flag` 为 1，标识在数据插入时访问的分表。

A5、更新分组分表信息缓存，增加 A2 中遍历的分表或者 A4 中创建的分表的分表信息；同时更新分组当前分表序号缓存，该缓存用来记录每个分组正在使用的分表序号。

A6、定时检查 A5 中获取的分组当前分表序号缓存，如果定时器超时，未接收到分表管理退出信号则转步骤 A7，如果接收到分表管理退出信号则退出分表管理流程，根据经验数据设置定时器超时间 `check_interval` 为 1 小时。

A7、遍历分组当前分表序号缓存，如果未遍历完则转步骤 A8，若遍历完则转步骤 A6；

检查所有当前正在使用标记为 1 的分表后，进入下一次检查的定时器等待周期，该周期的值为估算值，检查分表容量时，数据插入不可用，为了防止频繁检查，这个时间周期最好设定的长一些，而且需要保证两次检查点之间，分表的数据行超过分表规则允许的分表内行总数 `row_count` 不会太多。

A8、根据当前分组的组号和分表序号，获取当前正在使用标记 `current_flag` 为 1 的分表数据行数 `table_row_count`，假设分组号为

2, 分表序号为 3;

A9、如果 A8 中的 `table_row_count`, 大于分表规则中分表内行总数 `row_count`, 则转步骤 A10、否则转步骤 A7;

A10、创建分组号为 2, 分表序号为 4 的新分表, 设置分表序号为 4 的 `current_flag` 为 1, `in_use_flag` 为 1, 将分表序号为 4 的分表信息插入数据库分表信息表, 同时在分表信息表中修改分表序号为 3 的 `current_flag` 为 0, 更新分组分表信息缓存, 插入分表序号 4 的分表信息, 并更新分组当前分表序号缓存中分组 2 对应的当前分表序号为 4;

创建新的分表后, 需要更新数据库分表信息表, 最新增加的分表一定就是当前使用的分表, 同时也要更新分组分表信息缓存和分组当前分表序号缓存, 作为数据插入时获取分表的依据。

A11、根据分组分表信息缓存判断当前分组中分表总数是否大于分表规则中分组内分表数总数 `split_count`, 如果大于则转步骤 A12, 如果小于则转步骤 A7;

A12、更新数据库分表信息表中该分组号对应的分表序号最小的分表信息, 将该分表的已使用标记 `in_use_flag` 置为 0, 表示这个分表在查询时不会使用, 而且需要将数据进行转移, 然后从分组分表信息缓存中删除该分表的记录, 转步骤 A11。例如在分组 2 中, 若 `split_count` 为 3, 则将分表序号为 1 的分表的 `in_use_flag` 设置为 0, 然后在分组分表信息缓存中删除分组 2 内分表序号为 1 分表的信息。

图 5 示出了历史数据插入的处理流程, 包括以下步骤:

B1、根据插入数据的内容确定数据业务分组的组号, 根据分组当前分表信息缓存获取当前正在使用标记为 1 的分表序号, 根据数据对应的分表重新排列插入数据顺序, 插入以分组号和分表序号为键, 插入数据内容为值的 `map`;

B2、遍历 `map`, 遍历完成则退出, 未完成转步骤 B3。

B3、根据分组号为分表序号获取分表名, 将数据插入分表中。

由于分组分表的存在，一批数据可能写入到不同分组的分表中，为了防止不断切换数据写入时使用的分表，在插入前按照分表与数据对应的方式将数据重新排列，减少表切换使用的时间，这样既可以保证数据的顺序，也可以提高数据插入的效率。

图 6 示出了历史数据查询的处理流程，包括以下步骤：

C1、读取数据库中分表信息表内已使用标记 `in_use_flag` 为 1 的分表信息；

C2、解析查询条件，分析需要包含的分组，然后将相关分组中所有的分表设置为查询目标分表列表，例如用 (X, Y) 来标识一个唯一的分表，其中 X 标识分组号，Y 标识分组内的分表序号，当需要查询的分组号包括 1, 2, 3 时，则最终需要查询的分表列表为：
(1, 1), (2, 2), (2, 3), (2, 4), (3, 1), (3, 2), (3, 3)；

C3、遍历查询所有分表，若查询结果有优先级的要求，则 C2 中的列表需要按优先级排序，例如：
(1, 1), (2, 4), (2, 3), (2, 2), (3, 3), (3, 2), (3, 1)，然后依次串行查询；若查询结果无优先级的要求，则 C2 中的列表没有顺序，查询可以并行执行，合并所有的查询结果，返回。

图 7 示出了历史数据转移的处理流程，包括以下步骤：

D1、读取数据库分表信息表中所有已使用标记 `in_use_flag` 为 0 的分表信息，保存到以分组号为键，分表序号列表为值的 map 中，例如分组号为 2 的分组中，分表序号为 1, 2 的分表的已使用标记为 0；

D2、依次遍历 D1 中获取的 map，获取的分组号和分表序号列表，然后遍历分表序号列表，如果未完成遍历，则转步骤 D3，完成遍历后退出；

D3、如遍历的当前分表为分组 2 内分表序号为 1 的分表，构建的表名为 `T_HISDATA_2_1`，通过数据库工具或者 SQL 脚本直接将分表中所有记录导出到外部文件；

D4、从数据库分表信息表中删除分表号为 2，分表序号为 1 的信息；

D5、从数据库中删除分表 T_HISDATA_2_1，转步骤 D2。

本说明书中未作详细描述的内容属于本领域专业技术人员公知的现有技术。

权 利 要 求 书

1. 一种网管海量历史数据管理方法，其特征在于：通过二级分表规则实现数据库分表，把保存有网管海量历史数据的单个表保存的数据，切分后保存到多个结构相同的分表中，减少单表的容量，从而提高单表的处理效率，并对各个分表进行以下管理：

A 历史数据分表管理，B 历史数据插入管理，C 历史数据查询管理，D 历史数据转移管理。

2. 如权利要求 1 所述的网管海量历史数据管理方法，其特征在于：二级分表规则用 xml 格式的文件纪录，其中：

历史数据类型 data_type，表示数据表保存数据的类型，

分组总数 group_count，表示对这种数据表分组的总数，

分组内分表总数 split_count，表示每个分组内分表数最大值，

分表内行总数 row_count，表示分表内行总数，当一个分表的数据量超过这个值的时候，就需要创建一个新的分表。

3. 如权利要求 2 所述的网管海量历史数据管理方法，其特征在于，所述历史数据分表管理的处理流程包括以下步骤：

A1、读取记载有数据表的二级分表规则的文件，获取历史数据类型 data_type、分组总数 group_count、分组内分表总数 split_count 和分表内行总数 row_count；

读取数据库分表信息表，获取所有分表的分组号 group_no、分表序号 table_no、当前正在使用标记 current_flag 和已使用标记 in_use_flag，将这些信息保存到一个分组号为键，分表信息列表为值的 map 对象中，所述分表信息列表包括分表序号 table_no、当前正在使用标记 current_flag 和已使用标记 in_use_flag，其中 map 的键的取值范围从 1 到分组总数 group_count，如果在数据库分表信息表中没有某个键对应的分表信息，那么在 map 中该键对应的分表信息列表就为空；

A2、遍历 A1 得到的 map 对象；

A3、根据遍历的结果，如果 map 对象中当前分组号对应的分表信息列表为空，则转至步骤 A4，否则，转至步骤 A5；

A4、在当前分组号的分组中创建一个可用的分表，分配分表序号为 1，设置分表当前正在使用标记为 1，已使用标记为 1，将该分表的信息插入数据库分表信息表后转至步骤 A5；

A5、将数据库分表信息表中记载的分表信息，按分组号为键，逐条插入到分组分表信息缓存中，该缓存用于记录所有分组中已创建的分表的信息，

将当前正在使用标记为 1 的分表信息，按分组号为键，将分表序号插入分组当前分表序号缓存，该缓存记录每个分组中 current_flag 为 1 的分表的分表序号，转步骤 A6；

分组当前分表序号缓存内的分表，需要检查其是否超过分表内行总数；

A6、使用定时器并等待，如果定时器未超时则一直等待，当定时器超时后，如果接收到应用程序退出的信号，则退出历史数据分表管理的处理流程；否则，转入步骤 A7；

A7、遍历分组当前分表序号缓存，如果未遍历完则转步骤 A8，若遍历完则转步骤 A6；

A8、读取当前分组中分表序号，根据当前分组的组号和分表序号构造分表的名称，获取分表数据行数；

A9、如果分表数据行数大于分表规则中分表内行总数 row_count，则转步骤 A10、否则转步骤 A7；

A10、创建新分表，新分表的分表序号为分表所在分组中原来 current_flag 为 1 的分表的分表序号值加上 1，然后在数据库分表信息表中插入该新分表的信息，并将该新分表的信息插入到分组分表信息缓存中，然后将分组当前分表序号缓存中的当前分组对应的分表序号改为新分表的分表序号；

A11、判断当前分组中分表总个数是否大于分组内分表总数 split_count，如果大于则转步骤 A12，如果小于则转步骤 A7；

A12、更新数据库分表信息表中该分组内分表序号最小分表的已使用标记 `in_use_flag` 为 0, 标识该分表不会在数据插入和查询中使用, 并将被历史数据转移管理流程处理, 并从分组分表信息缓存中删除该分表的记录, 转步骤 A11。

4. 如权利要求 3 所述的网管海量历史数据管理方法, 其特征在于, 所述历史数据插入管理的处理流程包括以下步骤:

B1、根据插入数据的内容确定数据业务分组的组号, 根据分组当前分表序号缓存获取分表序号, 根据分表和数据逻辑对应关系, 重新排列插入数据内容, 存入以分组号和分表序号为键, 插入数据内存为值的 `map` 中;

B2、遍历 B1 中的 `map`, 遍历完成则退出, 否则转步骤 B3;

B3、根据当前分组号和分表序号, 构造分表名称, 插入数据到当前分表中, 转步骤 B2。

5. 如权利要求 3 所述的网管海量历史数据管理方法, 其特征在于, 所述历史数据查询管理的处理流程包括以下步骤:

C1、读取数据库分表信息表内已使用标记为 1 的分表信息;

C2、分析查询条件、获取所有相关分组分表的表名列表;

C3、根据查询条件查询所有 C2 中的分表。

6. 如权利要求 3 所述的网管海量历史数据管理方法, 其特征在于, 所述历史数据转移管理的处理流程包括以下步骤:

D1、按分组号读取数据库分表信息表中所有已使用标记为 0 的分表信息记录, 保存到以分组号为键, 分表序号列表为值的 `map` 中;

D2、遍历 D1 中获取的 `map`, 遍历完成后退出;

D3、根据当前分组号和分表序号, 构造分表名称, 使用批处理工具直接导出指定分表名称的分表中所有行记录;

D4、根据分组号和分表序号从数据库分表信息表中删除分表信息;

D5、根据分表名称从数据库中删除分表, 转步骤 D2。

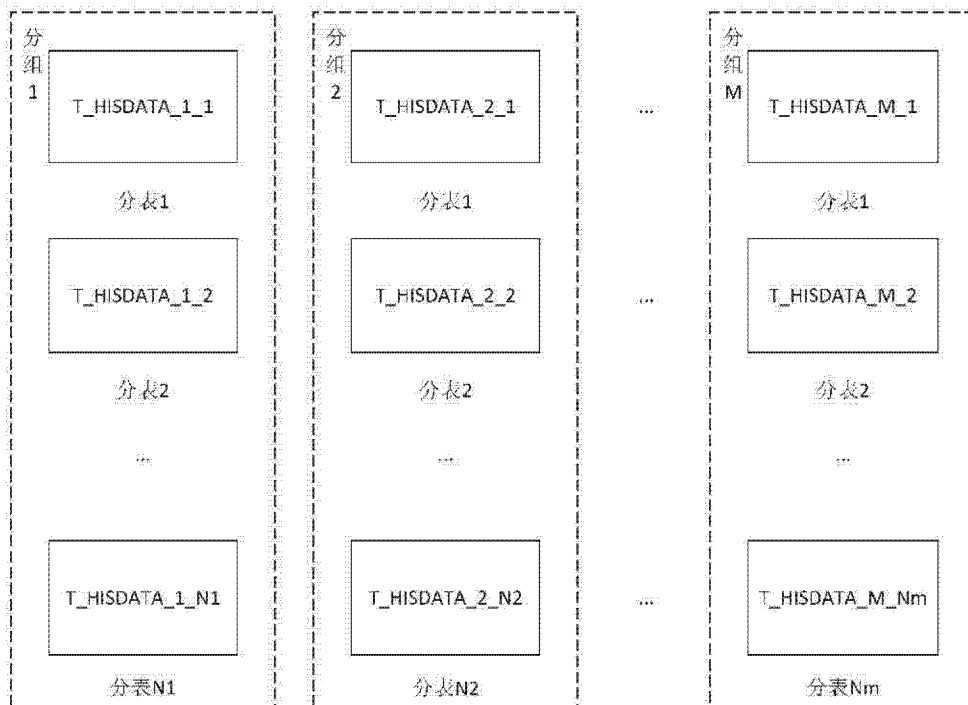


图 1

<split_rule>
 <data_type>历史数据类型</data_type>
 <group_count>分组总数</group_count>
 <split_count>分组内分表总数</split_count>
 <row_count>分表内行总数</row_count>
 </split_rule>

图 2

序号	分组号	分表序号	当前正在使用标记	已使用标记
seq_no	group_no	table_no	current_flag	in_use_flag
1	1	1	1	1
2	2	1	0	0
3	3	1	0	1
4	2	2	0	1
5	3	2	0	1
6	2	3	0	1
7	3	3	1	1
8	2	4	1	1

图 3

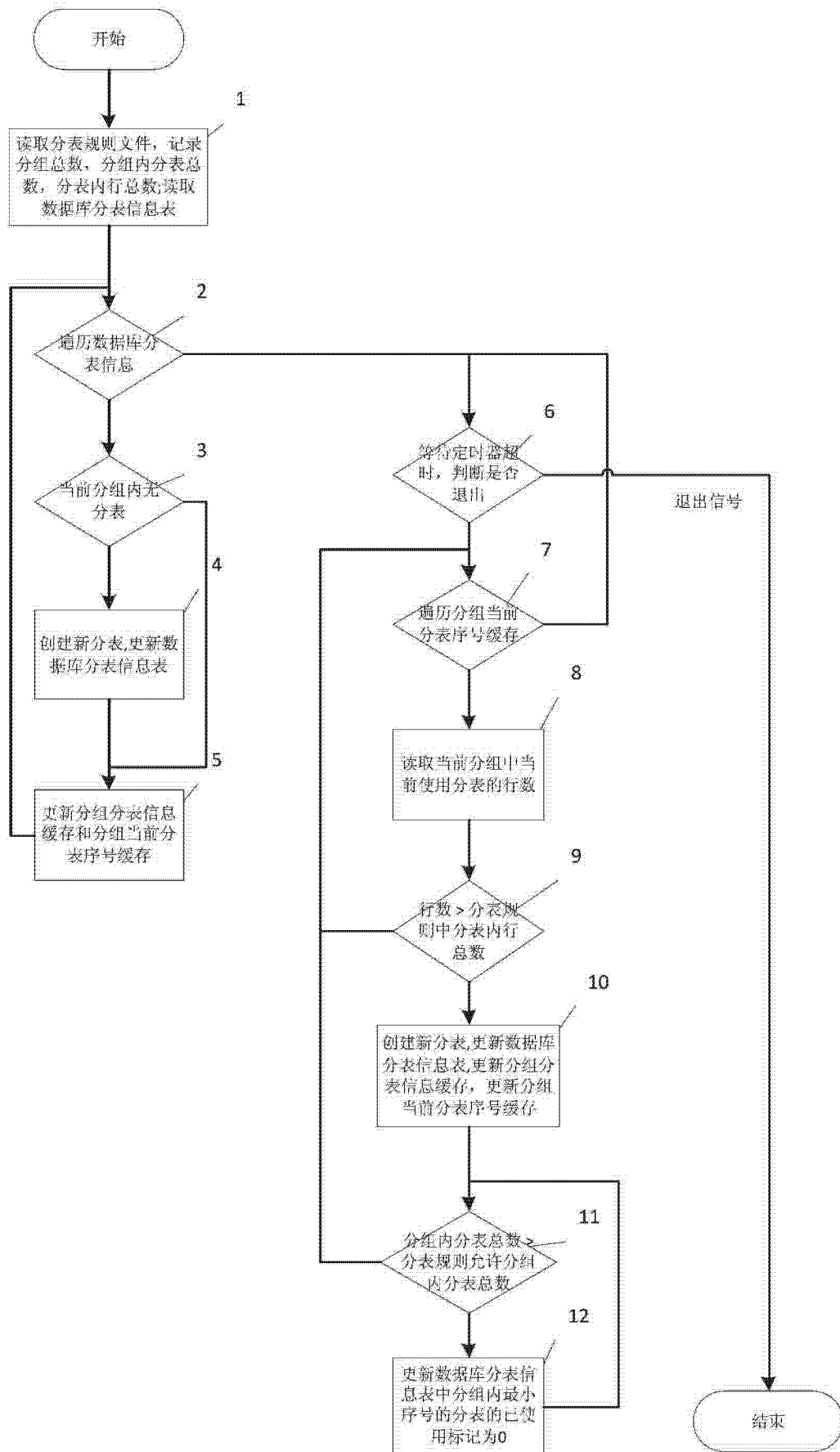


图 4

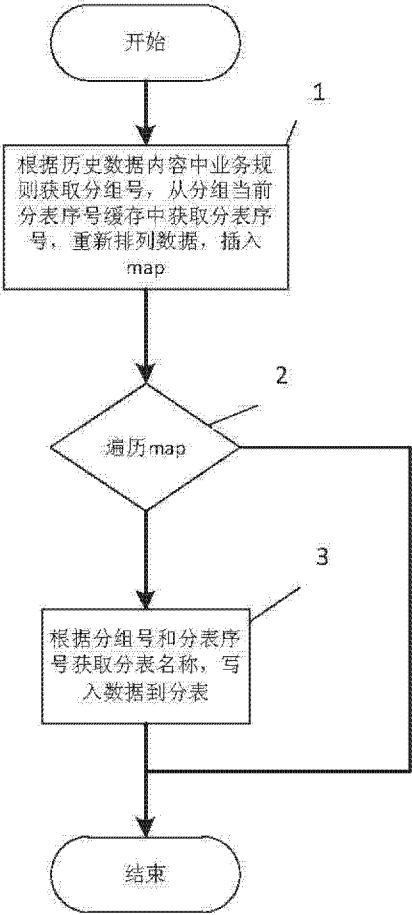


图 5

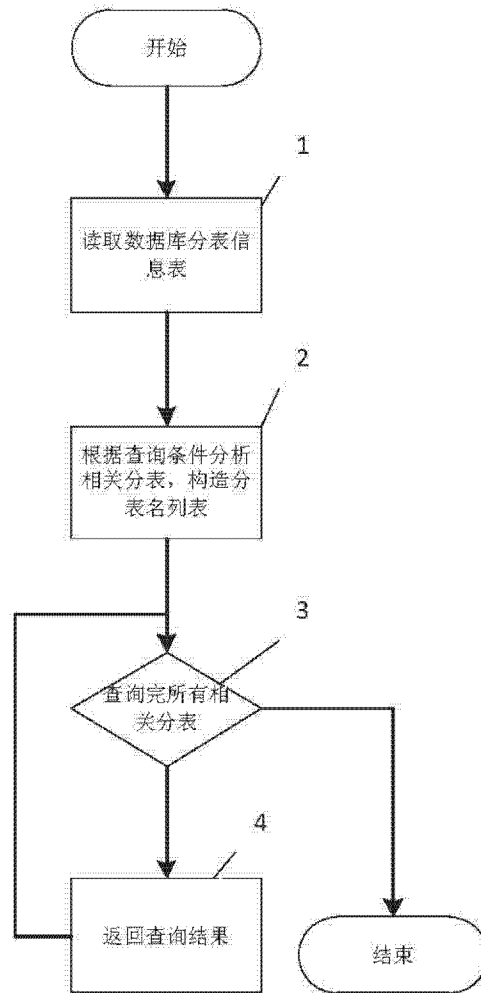


图 6

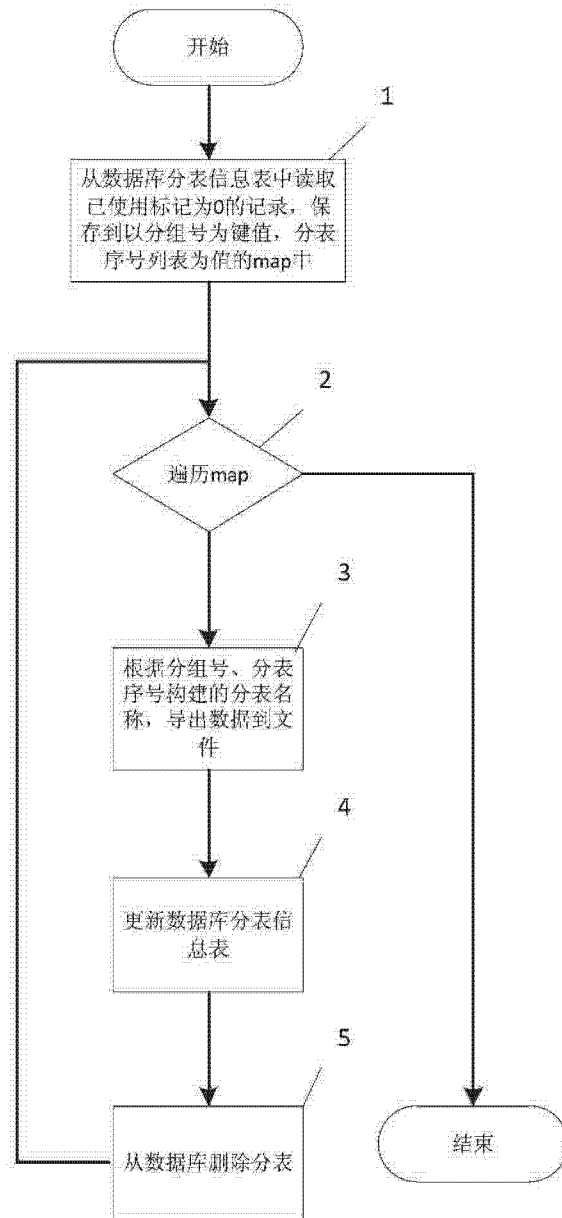


图 7

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2013/079993

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)

IPC: 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, CNKI, WPI, EPODOC, GOOGLE: webmaster, historical data, secondary table rule, submeter, subtable, data, segmentation, signal table, general table, management, historical, data, two w level, table w rule?, split, database, divid+, single w table, capacity

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 102867071 A (FIBERHOME TELECOMMUNICATION TECHNOLOGIES CO., LTD.), 09 January 2013 (09.01.2013), the whole document	1-6
X	CN 101566986 A (ALIBABA GROUP HOLDING LTD.), 28 October 2009 (28.10.2009), description, page 6, last line to page 7, line 16 and page 9, lines 14-23, and figure 2	1-6
A	CN 101149744 A (ZTE CORP.), 26 March 2008 (26.03.2008), the whole document	1-6
A	CN 102542054 A (XIAMEN YAXUN NETWORK CO., LTD.), 04 July 2012 (04.07.2012), the whole document	1-6
A	US 2003/0154206 A1 (SANYO ELECTRIC CO., LTD.), 14 August 2003 (14.08.2003), the whole document	1-6

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

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

Date of the actual completion of the international search 12 September 2013 (12.09.2013)	Date of mailing of the international search report 17 October 2013 (17.10.2013)
Name and mailing address of the ISA/CN: State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No.: (86-10) 62019451	Authorized officer YANG, Ying Telephone No.: (86-10) 62413310

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2013/079993

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 102867071 A	09.01.2013	None	
CN 101566986 A	28.10.2009	None	
CN 101149744 A	26.03.2008	None	
CN 102542054 A	04.07.2012	None	
US 2003/0154206 A1	14.08.2003	JP 2004005423 A	08.01.2004
		JP 2005050383 A	24.02.2005

国际检索报告

国际申请号

PCT/CN2013/079993

A. 主题的分类

G06F 17/30 (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)

IPC: 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,CNKI,WPI,EPODOC,GOOGLE: 网管, 历史数据, 二级分表规则, 数据库, 分表, 子表, 数据, 切分, 单表, 总表, 容量, management, historical, data, two w level, table w rule?, split, database, divid+, single w table, capacity

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	CN102867071A (烽火通信科技股份有限公司) 09.1 月 2013(09.01.2013) 全文	1-6
X	CN101566986A (阿里巴巴集团控股有限公司) 28.10 月 2009(28.10.2009) 说明书第 6 页最后 1 行至第 7 页第 16 行,第 9 页第 14-23 行及附图 2	1-2
A	CN101149744A (中兴通讯股份有限公司) 26.3 月 2008(26.03.2008) 全文	1-6
A	CN102542054A (厦门雅迅网络股份有限公司) 04.7 月 2012(04.07.2012) 全文	1-6
A	US2003/0154206A1 (SANYO ELECTRIC CO., LTD.) 14.8 月 2003(14.08.2003) 全文	1-6



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

12.9 月 2013(12.09.2013)

国际检索报告邮寄日期

17.10 月 2013 (17.10.2013)

ISA/CN 的名称和邮寄地址:

中华人民共和国国家知识产权局

中国北京市海淀区蓟门桥西土城路 6 号 100088

传真号: (86-10)62019451

受权官员

杨颖

电话号码: (86-10) **62413310**

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2013/079993

检索报告中引用的 专利文件	公布日期	同族专利	公布日期
CN102867071A	09.01.2013	无	
CN101566986A	28.10.2009	无	
CN101149744A	26.03.2008	无	
CN102542054A	04.07.2012	无	
US2003/0154206A1	14.08.2003	JP2004005423A	08.01.2004
		JP2005050383A	24.02.2005