

(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
2015 年 12 月 3 日 (03.12.2015)



(10) 国际公布号
WO 2015/180394 A1

- (51) 国际专利分类号:
G06F 17/30 (2006.01)
- (21) 国际申请号: PCT/CN2014/089458
- (22) 国际申请日: 2014 年 10 月 24 日 (24.10.2014)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201410231326.5 2014 年 5 月 28 日 (28.05.2014) 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: 谢勇 (XIE, Yong); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。李育国 (LI, Yuguo); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。钟延辉 (ZHONG, Yanhui); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。付旭东 (FU, Xudong); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (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, IR, IS, JP, KE, KG, KN, KP, KR, 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, 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,

[见续页]

(54) Title: METHOD AND DEVICE FOR RECOVERING CHECKPOINT IN COPY ON WRITE BASED FILE SYSTEM

(54) 发明名称: 一种基于写时复制的文件系统中回收检查点的方法和装置

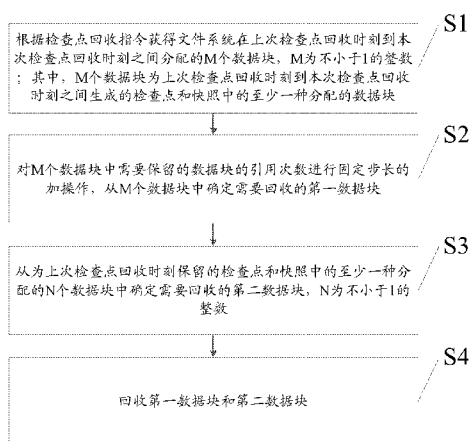


图 2 Fig. 2

S1 According to a checkpoint recovery instruction, obtaining M data blocks which are allocated between the last checkpoint recovery moment and this checkpoint recovery moment by a file system, where M is an integer which is not smaller than 1, wherein, the M data blocks are data blocks allocated for at least one of a checkpoint and a snapshot which are generated between the last checkpoint recovery moment and this checkpoint recovery moment.

S2 Conducting a fixed-step add operation on the number of times that a data block among the M data blocks needing to be reserved is quoted, and determining a first data block needing to be recovered from the M data blocks.

S3 Determining a second data block needing to be recovered from N data blocks allocated for at least one of the checkpoint and the snapshot which are reserved at the last checkpoint recovery moment, where N is an integer which is not smaller than 1.

S4 Recovering the first data block and the second data block.

(57) Abstract: Disclosed are a method and device for recovering a checkpoint in a copy on write based file system. The method for recovering a checkpoint comprises: according to a checkpoint recovery instruction, obtaining M data blocks which are allocated between the last checkpoint recovery moment and this checkpoint recovery moment by a file system, where M is an integer which is not smaller than 1, wherein, the M data blocks are data blocks allocated for at least one of a checkpoint and a snapshot which are generated between the last checkpoint recovery moment and this checkpoint recovery moment; conducting a fixed-step addition operation on the number of times that a data block among the M data blocks needing to be reserved is quoted, and determining a first data block needing to be recovered from the M data blocks; determining a second data block needing to be recovered from N data blocks allocated for at least one of a checkpoint and a snapshot which are reserved at the last checkpoint recovery moment, where N is an integer which is not smaller than 1; and recovering the first data block and the second data block.

(57) 摘要: 本发明实施例公开了一种基于写时复制的文件系统中回收检查点的方法和装置, 所述回收检查点的方法包括: 根据检查点回收指令获得所述文件系统中上次检查点回收时刻到本次检查点回收时刻之间分配的 M 个数据块, M 为不小于 1 的整数; 其中, 所述 M 个数据块为上次检查点回收时刻到本次检查点回收时刻之间生成的检查点和快照中的至少一种分配的数据块; 对所述 M 个数据块中需要保留的数据块的引用次数进行固定步长的加操作, 从所述 M 个数据块中确定需要回收的第一数据块; 从为所述上次检查点回收时刻保留的检查点和快照中的至少一种分配的 N 个数据块中确定需要回收的第二数据块, N 为不小于 1 的

整数; 回收所述第一数据块和所述第二数据块。



WO 2015/180394 A1



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))。

一种基于写时复制的文件系统中回收检查点的方法和装置

技术领域

本发明涉及数据处理领域，尤其涉及一种基于写时复制的文件系统中回收检查点的方法和装置。

背景技术

写时复制（简称：COW；英文：Copy On Write）是指当对文件系统中的数据进行改动的时候，并不真地去改动原数据，而是把要改动的数据复制到在磁盘的空白区。可以看出，由于在写的过程中，原数据没有被破坏，所以能保障数据的一致性，而且不存在二次写惩罚，同时避免了由二次写惩罚带来的小写问题，所以 COW 文件系统的应用领域越来越广。

在 COW 文件系统中，每次生成检查点（checkpoint，简称为 ckp）或者快照（snapshot）的时候，都会从检查点或者快照的根部向下遍历进行加引用计数，若数据块的引用计数在被加 1 后则会继续向下遍历，若数据块的引用计数在被加 1 后大于 1 就不再向下遍历。在 COW 文件系统中，包括引用树的最上层为超级块（简称：sb；英文：super block），被超级块引用的可以是检查点、快照、根区（ROOT）、索引节点（inode）等等。

请参考图 1A，图 1A 是现有技术中 COW 文件系统中具有检查点 1（ckp1）的引用示意图，如图 1A 所示，该检查点 1 中包括编号为 A—H 的 8 个数据块，对检查点 1 执行加引用计数遍历可知其引用计数为：

数据块编号	引用计数
A	1
B	1
C	1
C	1
D	1
E	1
F	1
G	1
H	1

请参考图 1B, 图 1B 是现有技术中在检查点 1 的基础上生成检查点 2 的示意图, 如图 1B 所示, COW 文件系统修改了索引节点 2 右边的子块 (即将 H 修改为 L) 而生成了检查点 2, 从检查点 2 根部执行加引用计数遍历可知其引用计数为:

数据块编号	引用计数
I	1
J	1
K	1
L	1

C	2
G	2

在删除检查点或者快照的时候, 需要从检查点或者快照的根部向下遍历进行减引用计数, 若数据块的引用计数被减 1 后为 0 则继续向下遍历, 若数据块的引用计数被减 1 后大于 0 就不再向下遍历, 其中, 引用计数为 0 的数据块即为需要回收的数据块。

以删除如图 1B 所示的检查点 1 为例, 请参考图 1C, 图 1C 是现有技术中删除检查点 1 的示意图, 如图 1C 所示, 对检查点 1 执行减引用计数遍历可知其引用计数为:

数据块编号	引用计数
A	0
B	0
C	1
D	0
E	1
F	1
G	1
H	0

在对检查点 1 执行减引用计数遍历后, 可知在检查点 1 删除后, 编号为 A、B、D、H 对应的数据块引用计数为 0, 需要回收, 从而才能够减少这些数据块

所占用的空间。

可以看出，在 COW 文件系统中，对生成的每一个检查点和每一个快照都需要进行加引用计数遍历，对删除的每一个检查点和快照都需要进行减引用计数遍历，所以遍历的规模较大，遍历时所计算的数据量也较大。

发明内容

本发明实施例通过提供一种基于写时复制的文件系统中回收检查点的方法和装置，用以解决现有技术中存在的在 COW 文件系统中遍历的规模较大，遍历时所计算的数据量也较大的技术问题。

本发明实施例第一方面提供了一种基于写时复制的文件系统中回收检查点的方法，包括：

根据检查点回收指令获得所述文件系统中上次检查点回收时刻到本次检查点回收时刻之间分配的 M 个数据块， M 为不小于 1 的整数；其中，所述 M 个数据块为上次检查点回收时刻到本次检查点回收时刻之间生成的检查点和快照中的至少一种分配的数据块；

对所述 M 个数据块中需要保留的数据块的引用次数进行固定步长的加操作，从所述 M 个数据块中确定需要回收的第一数据块；

从为所述上次检查点回收时刻保留的检查点和快照中的至少一种分配的 N 个数据块中确定需要回收的第二数据块， N 为不小于 1 的整数；

回收所述第一数据块和所述第二数据块。

结合第一方面，在第一方面的第一种可能的实现方式中，所述对所述 M

个数据块中需要保留的数据块的引用次数进行固定步长的加操作，从所述 M 个数据块中确定需要回收的第一数据块，具体包括：

当上次检查点回收时刻到本次检查点回收时刻之间只生成检查点时，则对为当前最新生成的检查点分配的 K 个数据块的引用次数进行所述固定步长的加操作，获得所述 K 个数据块的第一引用计数；所述 M 个数据块中除所述 K 个数据块外的数据块中确定所述第一数据块；或

当上次检查点回收时刻到本次检查点回收时刻之间既生成检查点又生成快照时，则对为上次检查点回收时刻到本次检查点回收时刻之间生成的快照和当前最新生成的检查点分配的 K 个数据块的引用次数进行所述固定步长的加操作，获得所述 K 个数据块的第一引用计数；从所述 M 个数据块中除所述 K 个数据块外的数据块中确定所述第一数据块。

结合第一方面的第一种可能的实现方式，在第一方面的第二种可能的实现方式中，所述从为所述上次检查点回收时刻保留的检查点和快照中的至少一种分配的 N 个数据块中确定需要回收的第二数据块，具体包括：

对所述 N 个数据块中的为所述上次检查点回收时刻保留的检查点分配的 L 个数据块的引用次数进行所述固定步长的减操作，确定所述 L 个数据块的第二引用计数；其中，所述获得所述 K 个数据块的第一引用计数之前，所述 N 个数据块的引用计数为第一参考引用计数；

根据所述第一引用计数、所述第二引用计数和所述第一参考引用计数，从所述 N 个数据块中确定所述第二数据块。

结合第一方面的第二种可能的实现方式，在第一方面的第三种可能的实现方式中，所述方法还包括：

根据所述 K 个数据块的第一引用计数, 所述 N 个数据块的第一参考引用计数和所述 L 个数据块当前的引用计数, 确定所述本次检查点回收需要保留的数据块的引用计数作为当前第一参考引用计数。

结合第一方面的第二种可能的实现方式, 在第一方面的第四种可能的实现方式中, 在所述文件系统包括上次检查点回收截止根区、本次检查点回收截止根区和实时根区时, 其中, 所述上次检查点回收截止根区为对所述 N 个数据块进行索引的根区, 所述实时根区为对所述文件系统在所述 K 个数据块进行索引的根区, 所述本次检查点回收截止根区为在所述文件系统获得所述检查点回收指令时复制所述实时根区内索引关系的根区。

结合第一方面的第四种可能的实现方式, 在第一方面的第五种可能的实现方式中, 在所述回收所述第一数据块和所述第二数据块之后, 所述方法还包括:

删除所述上次检查点回收截止根区内的数据;

将所述本次检查点回收截止根区内的数据复制到所述上次检查点回收截止根区中。

本发明实施例第二方面提供一种基于写时复制的文件系统中回收检查点的装置, 包括:

获得单元, 用于根据检查点回收指令获得文件系统在上次检查点回收时刻到本次检查点回收时刻之间分配的 M 个数据块, M 为不小于 1 的整数; 其中, 所述 M 个数据块为上次检查点回收时刻到本次检查点回收时刻之间生成的检查点和快照中的至少一种分配的数据块;

第一确定单元, 用于对所述 M 个数据块中需要保留的数据块的引用次数进行固定步长的加操作, 从所述 M 个数据块中确定需要回收的第一数据块;

第二确定单元,用于从为所述上次检查点回收时刻保留的检查点和快照中的至少一种分配的 N 个数据块中确定需要回收的第二数据块, N 为不小于 1 的整数;

回收单元,用于回收所述第一数据块和所述第二数据块。

结合第二方面,在第二方面的第一种可能的实现方式中,所述第一确定单元具体用于当上次检查点回收时刻到本次检查点回收时刻之间只生成检查点时,则对为当前最新生成的检查点分配的 K 个数据块的引用次数进行所述固定步长的加操作,获得所述 K 个数据块的第一引用计数,从所述 M 个数据块中除所述 K 个数据块外的数据块中确定所述第一数据块;或当上次检查点回收时刻到本次检查点回收时刻之间既生成检查点又生成快照时,则对为上次检查点回收时刻到本次检查点回收时刻之间生成的快照和当前最新生成的检查点分配的 K 个数据块的引用次数进行所述固定步长的加操作,获得所述 K 个数据块的第一引用计数,所述 M 个数据块中除所述 K 个数据块外的数据块中确定所述第一数据块。

结合第二方面的第一种可能的实现方式,在第二方面的第二种可能的实现方式中,所述第二确定单元具体用于对所述 N 个数据块中的为所述上次检查点回收时刻保留的检查点分配的 L 个数据块的引用次数进行所述固定步长的减操作,确定所述 L 个数据块当前的引用计数;其中,所述获得所述 K 个数据块的第一引用计数之前,所述 N 个数据块的引用计数为第一参考引用计数,并根据所述第一引用计数、所述第二引用计数和所述第一参考引用计数,从所述 N 个数据块中确定所述第二数据块。

结合第二方面的第二种可能的实现方式,在第二方面的第三种可能的实现方式中,所述装置还包括第三确定单元,所述第三确定单元用于根据所述第一

引用计数、所述第二引用计数和所述第一参考引用计数，确定所述本次检查点回收需要保留的数据块的引用计数作为当前第一参考引用计数。

结合第二方面的第二种可能的实现方式，在第二方面的第四种可能的实现方式中，在所述文件系统包括上次检查点回收截止根区、本次检查点回收截止根区和实时根区时，其中，所述上次检查点回收截止根区为对所述 N 个数据块进行索引的根区，所述实时根区为对所述文件系统在所述 K 个数据块进行索引的根区，所述本次检查点回收截止根区为在所述文件系统获得所述检查点回收指令时复制所述实时根区内索引关系的根区。

结合第二方面的第四种可能的实现方式，在第二方面的第五种可能的实现方式中，所述回收单元具体还用于在回收所述第一数据块和所述第二数据块之后，删除所述上次检查点回收截止根区内的数据，并将所述本次检查点回收截止根区内的数据复制到所述上次检查点回收截止根区中。

本发明实施例第三方面还提供一种设备，该设备包括第二方面提供的任一种装置。

本发明实施例中提供的一个或多个技术方案，至少具有如下技术效果或优点：

由于采用了根据检查点回收指令获得所述文件系统在上次检查点回收时刻到本次检查点回收时刻之间分配的 M 个数据块，M 为不小于 1 的整数，其中，所述 M 个数据块为上次检查点回收时刻到本次检查点回收时刻之间生成的检查点和快照中的至少一种分配的数据块，并对所述 M 个数据块中需要保留的数据块的引用次数进行固定步长的加操作，从所述 M 个数据块中确定需要回收的第一数据块，并从为所述上次检查点回收时刻保留的检查点和快照中的至少一种分配的 N 个数据块中确定需要回收的第二数据块，N 为不小于 1

的整数,并回收所述第一数据块和所述第二数据块的技术方案,不需要对 COW 文件系统中上次检查点回收时刻到本次检查点回收时刻之间生成的数据块中需要回收的数据块进行遍历操作,而只需要对需要保留的数据块进行遍历操作,所以实现了减少对 COW 文件系统中上次检查点回收时刻到本次检查点回收时刻之间需要回收的数据块进行遍历操作时对应的数据量,缩减 COW 文件系统回收空间时遍历的规模的技术效果。

附图说明

图 1A 为现有技术中 COW 文件系统中具有检查点 1 (ckp1) 的引用示意图;

图 1B 为现有技术中在检查点 1 的基础上生成检查点 2 的示意图;

图 1C 为现有技术中删除检查点 1 的示意图;

图 2 为本发明实施例提供的回收检查点的方法的流程图;

图 3 为本发明实施例提供的 COW 文件系统中上次检查点回收时刻保留的数据块的示意图;

图 4 为本发明实施例提供的 COW 文件系统中上次检查点回收时刻保留的数据块发生第一次修改的示意图;

图 5 为本发明实施例提供的 COW 文件系统中上次检查点回收时刻保留的数据块发生第二次修改的示意图;

图 6 为本发明实施例提供的 COW 文件系统中上次检查点回收时刻保留的数据块发生第三次修改的示意图;

图 7 为本发明实施例提供的 COW 文件系统包括上次检查点回收截止根区、本次检查点回收截止根区和实时根区的示意图;

图 8 为本发明实施例提供的 COW 文件系统本次检查点回收时刻的示意图;

图 9 为本发明实施例提供的 COW 文件系统在接收到检查点回收指令后的示意图;

图 10 为本发明实施例提供的将本次检查点回收截止根区内的内容复制到上次检查点回收截止根区内的示意图;

图 11 为本发明实施例提供的检查点装置的结构图。

具体实施方式

本发明实施例通过提供一种基于写时复制的文件系统中回收检查点的方法和装置,用以解决现有技术中存在的在 COW 文件系统中遍历的规模较大,遍历时所计算的数据量也较大的技术问题。

本发明实施例中的技术方案为解决上述技术问题,总体思路如下:

本发明实施例提供一种基于写时复制的文件系统中回收检查点的方法,该方法包括:根据检查点回收指令获得文件系统在上次检查点回收时刻到本次检查点回收时刻之间分配的 M 个数据块, M 为不小于 1 的整数;其中, M 个数据块为上次检查点回收时刻到本次检查点回收时刻之间生成的检查点和快照中的至少一种分配的数据块;对 M 个数据块中需要保留的数据块的引用次数进行固定步长的加操作,从 M 个数据块中确定需要回收的第一数据块;从为

上次检查点回收时刻保留的检查点和快照中的至少一种分配的 N 个数据块中确定需要回收的第二数据块, N 为不小于 1 的整数; 回收第一数据块和第二数据块。

通过上述部分可以看出, 由于采用了根据检查点回收指令获得文件系统在上次检查点回收时刻到本次检查点回收时刻之间分配的 M 个数据块, M 为不小于 1 的整数, 其中, M 个数据块为上次检查点回收时刻到本次检查点回收时刻之间生成的检查点和快照中的至少一种分配的数据块, 并对 M 个数据块中需要保留的数据块的引用次数进行固定步长的加操作, 从 M 个数据块中确定需要回收的第一数据块, 并从而为上次检查点回收时刻保留的检查点和快照中的至少一种分配的 N 个数据块中确定需要回收的第二数据块, N 为不小于 1 的整数, 并回收第一数据块和第二数据块的技术方案, 不需要对 COW 文件系统中上次检查点回收时刻到本次检查点回收时刻之间生成的数据块中需要回收的数据块进行遍历操作, 而只需要对需要保留的数据块进行遍历操作, 所以实现了减少对 COW 文件系统中上次检查点回收时刻到本次检查点回收时刻之间需要回收的数据块进行遍历操作时对应的数据量, 缩减 COW 文件系统回收空间时遍历的规模的技术效果。

为了更好地理解上述技术方案, 下面将结合说明书附图以及具体的实施方式对上述技术方案进行详细的说明。

请参考图 2, 图 2 是本发明实施例提供的回收检查点的方法的流程图, 如图 2 所示, 该方法包括:

S1: 根据检查点回收指令获得文件系统在上次检查点回收时刻到本次检查点回收时刻之间分配的 M 个数据块, M 为不小于 1 的整数; 其中, M 个数据块为上次检查点回收时刻到本次检查点回收时刻之间生成的检查点和快照中

的至少一种分配的数据块;

S2: 对 M 个数据块中需要保留的数据块的引用次数进行固定步长的加操作, 从 M 个数据块中确定需要回收的第一数据块;

S3: 从为上次检查点回收时刻保留的检查点和快照中的至少一种分配的 N 个数据块中确定需要回收的第二数据块, N 为不小于 1 的整数;

S4: 回收第一数据块和第二数据块。

在具体实施过程中, 实现本发明实施例提供的回收检查点的方法有多种实现方式, 在接下来的部分中, 将详细介绍其中的两种。

实施例一

请参考图 3, 图 3 是本发明实施例提供的 COW 文件系统上次检查点回收时刻保留的数据块的示意图, 如图 3 所示, COW 文件系统上次检查点回收时刻保留了快照 1 (snp1) 和检查点 1 (ckp1), 其中快照 1 包括编号 A、B、C、D 的四个数据块, 或称为快照 1 引用的编号为 A、B、C、D 数据块, 检查点 1 包括编号 E、F、C、G 的四个数据块, 或称为检查点 1 引用的编号为 E、F、C、G 的四个数据块。对快照 1 和检查点 1 引用的数据块分别执行固定步长的加操作, 这里执行加引用计数遍历操作, 则可以得到上次回收时刻的第一参考引用计数, 本发明实施例中, 数据块被快照 1 或检查点 1 引用一次, 则引用计数加 1, 也就是说固定步长为 1。在本实施例中, 所有的引用计数都以表的形式呈现, 第一参考引用计数为:

数据块编号	引用计数
A	1

B	1
C	2
D	1
E	1
F	1
G	1

表 1

请继续参考图 4，图 4 是本发明实施例提供的 COW 文件系统中上次检查点回收时刻保留的数据块发生第一次修改的示意图，如图 4 所示，数据块 C 被修改为 J，则此时会生成一个检查点 2 (ckp2)，对检查点 2 做一个快照，可以生成如图 4 所示的快照 2 (snp2)。

请继续参考图 5，图 5 是本发明实施例提供的 COW 文件系统中上次检查点回收时刻保留的数据块发生第二次修改的示意图，如图 5 所示，数据块 J 被修改为 M，则此时会生成一个检查点 3 (ckp3)。

请继续参考图 6，图 6 是本发明实施例提供的 COW 文件系统中上次检查点回收时刻保留的数据块发生第三次修改的示意图，如图 6 所示，编号为 G 的数据块被修改为编号为 P 的数据块，则此时会生成一个检查点 4 (ckp4)。

设定此时 COW 文件系统接收到一检查点回收指令，在具体实施过程中，该检查点回收指令可以由用户手动触发而生成，也可以是由 COW 文件系统的回收策略触发而生成，在此不做限制。

当然，需要说明的是，在本实施例中，从 COW 文件系统上次检查点回收时刻之后，到 COW 文件系统接收到检查点回收指令时刻之间，数据块都不进行遍历操作和引用计数更新操作。

在步骤 S1 中，根据检查点回收指令获得文件系统在上次检查点回收时刻到本次检查点回收时刻之间分配的 M 个数据块，M 为不小于 1 的整数，其中，M 个数据块为上次检查点回收时刻到本次检查点回收时刻之间生成的检查点和快照中的至少一种分配的数据块，具体来讲，可以通过 COW 文件系统的数据块分配模块获得 COW 文件系统从上次检查点回收时刻到本次检查点回收时刻之间分配的 M 个数据块，COW 文件系统的数据块分配模块会在上次检查点回收时刻与本次检查点回收时刻之间生成检查点和快照中的至少一种时，为检查点和快照中的至少一种对应的数据块分配对应的编号，由图 3 至图 6 可知，一个检查点或快照会对应多个数据块，所以 M 为不小于 1 的整数。

需要说明的是，本实施例中的“分配”是指基于写时复制的文件系统在生成检查点或快照时，该检查点或快照引用现有检查点或快照对应的数据块时，该数据块也属于为新生成的检查点或快照分配的数据块。

在本实施例中，请同时参考图 3、图 4、图 5 和图 6，可以从 COW 文件系统的数据块分配模块获得从上次检查点回收时刻到本次检查点回收时刻之间已经分配的数据块的编号为“G—P”这一信息，由于数据块编号为 COW 文件系统的数据块分配模块连续分配的，数据块分配模块可以只记录 COW 文件系统在上次检查点回收时刻和本次检查点回收时刻之间分配的第一个数据块编号和最后一个数据块编号即可，占用的空间较小。

在步骤 S2 中，对 M 个数据块中需要保留的数据块的引用次数进行固定步长的加操作，从 M 个数据块中确定需要回收的第一数据块，具体来讲，可以

是当上次检查点回收时刻到本次检查点回收时刻之间只生成检查点时,则对为当前最新生成的检查点分配的 K 个数据块的引用次数进行固定步长的加操作,这里步长为 1,获得 K 个数据块的第一引用计数;从 M 个数据块中除 K 个数据块外的数据块中确定第一数据块;或当上次检查点回收时刻到本次检查点回收时刻之间既生成检查点又生成快照时,则对为上次检查点回收时刻到本次检查点回收时刻之间生成的快照和当前最新生成的检查点分配的 K 个数据块的引用次数进行固定步长的加操作,这里固定步长为 1,获得 K 个数据块的第一引用计数;从 M 个数据块中除 K 个数据块外的数据块中确定第一数据块。

由于 COW 文件系统中一般会保留所有的快照,同时只会保留最后生成的一个检查点,因此,在本实施例中,同时参考图 3、图 4、图 5 和图 6,快照 2 和检查点 4 需要保留,因此需要对快照 2 和检查点 4 对应的 K 个数据块的引用次数进行固定步长的加操作,获得 K 个数据块的第一引用计数。

具体来讲,即对快照 2 和检查点 4 对应的 K 个数据块执行加引用计数遍历操作,从而能够获得快照编号和检查点编号对应的数据块的第一引用计数,本发明实施例中,为快照 2 和检查点 4 分配的数据块编号为 G、H、I、J、M、N、O 和 P,对快照 2 和检查点 4 对应的数据块执行加引用计数遍历操作,即能够获得快照 2 和检查点 4 引用的数据块的第一引用计数:

数据块编号	引用计数
G	1
H	1
I	1
J	1

M	1
N	1
O	1
P	1

表 2

通过步骤 S1 获得的 COW 文件系统在上次检查点回收时刻到本次检查点回收时刻之间已经分配的数据块的编号为 G—P，所以结合第一引用计数，可以确定 COW 文件系统在上次检查点回收时刻到本次检查点回收时刻之间已经分配的数据块编号中可以回收的数据块的编号为 K、L（因为其没有任一需要被保留的检查点 4 和快照 2 引用），也即第一数据块的编号为 K、L。

当然了，如上述部分所介绍的，在本实施例中，上次检查点回收时刻到本次检查点回收时刻之间既生成检查点又生成快照，所以按上述部分所介绍的步骤，为快照 2 和检查点 4 分配的数据块编号为 G、H、I、J、M、N、O 和 P，即可以从 M 个数据块中除 K 个数据块之外的数据块中确定需要回收的第一数据块，在本发明实施例中，因为需要回收检查点 3，从编号为 K 和 L 的数据块中确定需要回收的数据块的编号为 K 和 L。在其他实施例中，上次检查点回收时刻到本次检查点回收时刻之间只生成检查点时，则只对当前最新生成的检查点分配的 K 个数据块的引用次数进行固定步长的加操作，固定步长为 1，从而能够获得该 K 个数据块的第一引用计数，从而可以从 M 个数据块中除 K 个数据块之外的数据块中确定第一数据块，在此就不再赘述了。

在通过步骤 S2 从 M 个数据块中确定需要回收的第一数据块之后，本发明实施例进入步骤 S3，即：从为上次检查点回收时刻保留的检查点和快照中的

至少一种分配的 N 个数据块中确定需要回收的第二数据块，N 为不小于 1 的整数。

在步骤 S3 中，具体来讲，可以包括：对 N 个数据块中的为上次检查点回收时刻保留的检查点分配的 L 个数据块的引用次数进行固定步长的减操作，这里固定步长为 1，确定 L 个数据块当前的引用计数；其中，获得快照 2 和检查点 4 引用的数据块的第一引用计数之前，N 个数据块的引用计数为第一参考引用计数；根据第一引用计数、L 个数据块当前的引用计数和第一参考引用计数，从 N 个数据块中确定第二数据块。

在本实施例中，具体来讲，如前述部分所介绍的，表 1 即是 COW 文件系统上次检查点回收时刻的第一参考引用计数。在获得第一参考引用计数之后，即可以对 N 个数据块中的为上次检查点回收时刻保留的检查点分配的 L 个数据块的引用次数进行固定步长的减操作，这里固定步长为 1，也即对该检查点对应的 L 个数据块执行减引用计数遍历操作，例如，请参考图 6，上次检查点回收时刻保留的数据块中，对检查点 1 对应的数据块进行固定步长的减操作，也即执行减引用计数遍历操作，可以得到 N 个数据块中为检查点 1 分配的 4 个数据块的第二引用计数，如下表所示：

数据块编号	引用计数
C	-1
E	-1
F	-1
G	-1

表 3

在获得第二引用计数后，即能够根据第一参考引用计数（表 1）、第一引用计数（表 2）和第二引用计数（表 3），从 N 个数据块中确定出第二数据块，具体来讲，可以是合并第一参考引用计数、第一引用计数和第二计数，在本实施例中，“合并第一参考引用计数、第一引用计数和第二计数”是指将第一参考引用计数、第一引用计数和第二计数中相同数据块的引用计数进行合并，不相同的数据块的引用计数保留，其结果如下表所示：

数据块编号	引用计数
A	1
B	1
C	1
D	1
E	0
F	0
G	1
H	1
I	1
J	1
M	1
N	1
O	1

P	1
---	---

表 4

从表 4 可以确定，上次检查点回收时刻保留的数据块中编号为 E、F 的数据块是需要回收的数据块，第二数据块也即 E、F。

通过上述部分可以看出，现有技术中 COW 文件系统的遍历规模相比，本发明实施例提供的回收检查点的方法不需要对 COW 文件系统中上次检查点回收时刻到本次检查点回收时刻之间生成的数据块中需要回收的数据块进行遍历操作，而只需要对需要保留的数据块进行遍历操作，所以实现了减少了对 COW 文件系统中上次检查点回收时刻到本次检查点回收时刻之间需要回收的数据块进行遍历操作时对应的数据量，缩减 COW 文件系统回收空间时遍历的规模的技术效果。

在确定第二数据块之后，本发明实施例提供的回收检查点的方法还包括：根据 K 个数据块的第一引用计数，N 个数据块的第一参考引用计数和 L 个数据块当前的引用计数，确定本次检查点回收需要保留的数据块的引用计数作为当前第一参考引用计数。也就是说，在 COW 文件系统确定出需要回收的第一数据块和第二数据块之后，即能够根据第一引用计数、第二引用计数和第一参考引用计数，确定出本次检查点回收需要保留的数据块的引用计数作为当前第一参考引用计数（也即表 4 所示的内容），以供 COW 文件系统下一次检查点回收的时候使用，在此就不再赘述了。

在通过步骤 S2 确定需要回收的第一数据块编号、以及通过步骤 S3 确定需要回收的第二数据块编号之后，本发明实施例提供的回收检查点的方法进入步骤 S4，即：回收第一数据块和第二数据块。

在步骤 S4 中，即是回收编号为 E、F 的数据块，以及回收编号为 K、L 的数据块，在此就不再赘述了。

上述本发明实施例中的技术方案，至少具有如下的技术效果或优点：

由于采用了根据检查点回收指令获得文件系统在上次检查点回收时刻到本次检查点回收时刻之间分配的 M 个数据块，M 为不小于 1 的整数，其中，M 个数据块为上次检查点回收时刻到本次检查点回收时刻之间生成的检查点和快照中的至少一种分配的数据块，并对 M 个数据块中需要保留的数据块的引用次数进行固定步长的加操作，从 M 个数据块中确定需要回收的第一数据块，并从为上次检查点回收时刻保留的检查点和快照中的至少一种分配的 N 个数据块中确定需要回收的第二数据块，N 为不小于 1 的整数，并回收第一数据块和第二数据块的技术方案，不需要对 COW 文件系统中上次检查点回收时刻到本次检查点回收时刻之间生成的数据块中需要回收的数据块进行遍历操作，而只需要对需要保留的数据块进行遍历操作，所以实现了减少对 COW 文件系统中上次检查点回收时刻到本次检查点回收时刻之间需要回收的数据块进行遍历操作时对应的数据量，缩减 COW 文件系统回收空间时遍历的规模的技术效果。

实施例二

请参考图 7，图 7 是本发明实施例提供的 COW 文件系统包括上次检查点回收截止根区、本次检查点回收截止根区和实时根区的示意图，如图 7 所示，其中，上次检查点回收截止根区为对 N 个数据块进行索引的根区，实时根区为对文件系统在 K 个数据块进行索引的根区，本次检查点回收截止根区为在文件系统获得检查点回收指令时复制实时根区内索引关系的根区。当然，需要

说明的是，在本实施例中，COW 文件系统的快照或者检查点的生成和删除都由对 COW 文件系统的树的插入和删除来实现，挂在 COW 文件系统的根区下面，以保证根据 COW 文件系统的树来回收需要回收的数据块对应的空间时不会遗漏。

如图 7 所示，上次检查点回收截止根区内为 COW 文件系统上次检查点回收时刻保留的数据块的索引，COW 文件系统上次检查点回收时刻，只有检查点，因此，为保留的检查点引用的数据块的编号为包括 A、B、C、D、E、F、G，本实施例以实时根区内为 COW 文件系统在上次检查点回收时刻之后对新生成的检查点分配的数据块的索引为例，上次检查点回收时刻之后为新生成的检查点分配的数据块的编号为 B、C、D、F、G 和 H，编号为 H 的数据块直接或间接引用数据块的编号为 B、C、D、E、F、G。

请继续参考图 8，图 8 是本发明实施例提供的 COW 文件系统本次检查点回收时刻的示意图，如图 8 所示，COW 文件系统在本次检查点回收时刻与上次检查点回收时刻相比，经过了一些事务，又生成了另一检查点，为该新生成的检查点分配的数据块的编号为 B、C、D、F、G 和 I，编号为 I 的数据块的索引在实时根区。

请继续参考图 9，图 9 为本发明实施例提供的 COW 文件系统在接收到检查点回收指令后的示意图，如图 9 所示，COW 文件系统在接收到检查点回收指令后，即会将图 8 所示的实时根区内的数据复制到本次检查点回收截止根区内，实时根区内的数据是指如图 8 所示对编号为 I 的数据块及其直接引用或间接引用的数据块的索引，即对编号为 B、C、D、F、G 和 I 的数据块的索引。当然，如图 9 所示，COW 文件系统在接收到检查点回收指令之后，会继续根据 COW 文件系统的运行情况，在实时根区索引在本次检查点回收时刻之后新增长的数据块，如图 9 所示的编号为 J 的数据块，在此就不再赘述了。

如图 9 所示, 在将实时根区内的数据复制到本次检查点回收截止根区之后, 即可以对本次检查点回收截止根区内的数据块执行加引用计数遍历操作 (需要说明的是, 此时不需要考虑实时根区的影响, 也即不计编号为 J 的数据块的引用关系), 获得各个数据块的引用计数如下:

数据块编号	引用计数
I	1
B	2
C	2
D	1
E	1
F	1
G	1

表 5

在步骤 S1 中, 获得 COW 文件系统在上次检查点回收时刻到本次检查点回收时刻之间已经分配的 M 个数据块, 具体来讲, 与实施例一中的过程类似, 同样是可以 COW 文件系统的数据库分配模块来获得该 M 个数据块, 在实施例一中已经介绍了详细的过程, 在此为了说明书的简洁, 就不再赘述了。

在本实施例中, 通过 COW 文件系统的数据库分配模块可知, COW 文件系统在上次检查点回收时刻与本次检查点回收时刻之间已经分配的 M 个数据块编号分别为“B—I”。

在通过步骤 S1 获得 COW 文件系统在上次检查点回收时刻到本次检查点回收时刻之间已经分配的数据块之后,本发明实施例提供的回收检查点的方法即进入步骤 S2,即:对 M 个数据块中需要保留的数据块的引用次数进行固定步长的加操作,从 M 个数据块中确定需要回收的第一数据块。如图 9 所示,由于只生成了检查点,因此,只需要保留最新的检查点即可,在截止根区最新的检查点引用的数据块的编号为 B、C、D、F、G 和 I,因此 M 个数据块中编号为 B、C、D、F、G 和 I 的数据块需要保留,编号为 H 的数据块为可回收数据块。

在本实施例中,由于 COW 文件系统在上次检查点回收时刻到本次检查点回收时刻之间已经分配的数据块编号为“B—I”,并结合表 5,可知编号 H 的数据块为需要回收的第一数据块,在对本次检查点回收截止根区索引的数据块执行加引用计数遍历操作时,由于编号为 H 的数据块未被引用,所以其不需要被保留,也不需要对其执行加引用计数遍历操作,在回收时也不需要对其执行减引用技术遍历操作,而在现有技术中,对编号为 H 的数据块会分别执行加引用技术遍历操作与减引用技术遍历操作。

在通过步骤 S2 确定需要回收的第一数据块编号后,本发明实施例提供的回收检查点的方法进入步骤 S3,即:从为上次检查点回收时刻保留的检查点和快照中的至少一种分配的 N 个数据块中确定需要回收的第二数据块,N 为不小于 1 的整数。

具体来讲,在本实施例中,由于只有检查点,因此,可以是对上次检查点回收截止根区索引的检查点引用的数据块执行减引用计数遍历操作,其结果如下:

数据块编号	引用计数
-------	------

A	0
B	1
C	1
D	1
E	1
F	1
G	1

表 6

根据表 6 可以看出,编号为 A 的数据块是不再被引用的数据块,编号为 A 的数据块为需要回收的数据块,也即 A 为第二数据块编号。

在通过步骤 S2 确定需要回收的第一数据块,以及通过步骤 S3 确定需要回收的第二数据块后,本发明实施例提供的回收检查点的方法进入步骤 S4,即:回收第一数据块和第二数据块编号。

与实施例一类似的,在本实施例中,回收第一数据库编号对应的第一空间即是回收编号为 H 的数据块对应的空间,回收第二数据块编号对应的第二空间即是回收编号为 A 的数据块对应的空间,为了说明书的简洁,在此就不再赘述了。

在具体实施过程中,在回收第一数据块编号对应的第一空间和第二数据块编号对应的第二空间之后,方法还包括:删除上次检查点回收截止根区内的数据;将本次检查点回收截止根区内的数据复制到上次收截止根区中。

具体来讲, 请继续参考图 9, 如图 9 所示, 在 COW 文件系统的回收检查点完成之后, 即清空上次检查点回收截止根区, 如删除上次检查点回收截止根区内的索引等等, 并将本次检查点回收截止根区内的数据如数据块的索引信息等等复制到上次检查点回收截止根区内, 以供 COW 文件系统下一次空间回收时使用, 请参考图 10, 图 10 是本发明实施例提供的将本次检查点回收截止根区内的数据复制到上次检查点回收截止根区内的示意图, 在此就不再赘述了。

在具体实施过程中, 如果是已经被删除的检查点或快照, 本次检查点回收截止根区中就已经不再有其索引 (如新的检查点生成后, 旧的检查点就不会再被索引), 从上次检查点回收截止根区向下执行减引用计数遍历操作时, 发现其引用为 0, 就会继续向下减, 另外如果是当前保留的检查点或快照 (如一个快照生成后未被删除, 实时根区内就会一直存在该快照的索引), 本次检查点回收截止根区有其索引, 从上次检查点回收截止根区向下执行减引用计数遍历操作时, 发现其引用不为 0, 就不再向下减, 因此, 在 COW 文件系统接收到检查点回收指令后, 不需要知道哪些检查点、快照需要回收, 根据实施根区内的索引情况即能够完成空间回收。

上述本发明实施例中的技术方案, 至少具有如下的技术效果或优点:

由于采用了根据检查点回收指令获得文件系统在上次检查点回收时刻到本次检查点回收时刻之间分配的 M 个数据块, M 为不小于 1 的整数, 其中, M 个数据块为上次检查点回收时刻到本次检查点回收时刻之间生成的检查点和快照中的至少一种分配的数据块, 并对 M 个数据块中需要保留的数据块的引用次数进行固定步长的加操作, 从 M 个数据块中确定需要回收的第一数据块, 并从为上次检查点回收时刻保留的检查点和快照中的至少一种分配的 N 个数据块中确定需要回收的第二数据块, N 为不小于 1 的整数, 并回收第一数据块和第二数据块的技术方案, 不需要对 COW 文件系统中上次检查点回收时

刻到本次检查点回收时刻之间生成的数据块中需要回收的数据块进行遍历操作，而只需要对需要保留的数据块进行遍历操作，所以实现了减少对 COW 文件系统中上次检查点回收时刻到本次检查点回收时刻之间需要回收的数据块进行遍历操作时对应的数据量，缩减 COW 文件系统回收空间时遍历的规模的技术效果。

基于同一发明构思，本发明实施例还提供一种基于写时复制的文件系统中回收检查点的装置，请参考图 11，图 10 是本发明实施例提供的装置的模块图，如图 11 所示，该装置包括：

获得单元 101，用于根据检查点回收指令获得文件系统在上次检查点回收时刻到本次检查点回收时刻之间分配的 M 个数据块，M 为不小于 1 的整数；其中，M 个数据块为上次检查点回收时刻到本次检查点回收时刻之间生成的检查点和快照中的至少一种分配的数据块；

第一确定单元 102，用于对 M 个数据块中需要保留的数据块的引用次数进行固定步长的加操作，从 M 个数据块中确定需要回收的第一数据块；

第二确定单元 103，用于从为上次检查点回收时刻保留的检查点和快照中的至少一种分配的 N 个数据块中确定需要回收的第二数据块，N 为不小于 1 的整数；

回收单元 104，用于回收第一数据块和第二数据块。

在具体实施过程中，第一确定单元 102 具体用于当上次检查点回收时刻到本次检查点回收时刻之间只生成检查点时，则对为当前最新生成的检查点分配的 K 个数据块的引用次数进行固定步长的加操作，获得 K 个数据块的第一引

用计数, 从 M 个数据块中除 K 个数据块外的数据块中确定第一数据块; 或当上次检查点回收时刻到本次检查点回收时刻之间既生成检查点又生成快照时, 则对为上次检查点回收时刻到本次检查点回收时刻之间生成的快照和当前最新生成的检查点分配的 K 个数据块的引用次数进行固定步长的加操作, 获得 K 个数据块的第一引用计数, 从 M 个数据块中除 K 个数据块外的数据块中确定第一数据块。

在具体实施过程中, 第二确定单元 103 具体用于对 N 个数据块中的为上次检查点回收时刻保留的检查点分配的 L 个数据块的引用次数进行固定步长的减操作, 确定 L 个数据块当前的引用计数; 其中, 获得 K 个数据块的第一引用计数之前, N 个数据块的引用计数为第一参考引用计数, 并根据第一引用计数、第二引用计数和第一参考引用计数, 从 N 个数据块中确定第二数据块。

在具体实施过程中, 装置还包括第三确定单元 105, 第三确定单元 105 用于根据第一引用计数、第二引用计数和第一参考引用计数, 确定本次检查点回收需要保留的数据块的引用计数作为当前第一参考引用计数。

在具体实施过程中, 在文件系统包括上次检查点回收截止根区、本次检查点回收截止根区和实时根区时, 其中, 上次检查点回收截止根区为对 N 个数据块进行索引的根区, 实时根区为对文件系统在 K 个数据块进行索引的根区, 本次检查点回收截止根区为在文件系统获得检查点回收指令时复制实时根区内索引关系的根区。

在具体实施过程中, 回收单元 104 具体还用于删除上次检查点回收截止根区内的数据; 将本次检查点回收截止根区内的数据复制到上次检查点回收截止根区中。

需要说明的是, 本实施例中的装置与前述实施例中的方法是基于同一发明

构思下的两个方面，在前面已经对方法的实施过程作了详细的描述，所以本领域技术人员可根据前述描述清楚地了解本实施例中的装置的结构及实施过程，为了说明书的简洁，在此就不再赘述了。

本发明实施例中用于表示数据块个数的符号与表示数据块的编号的符号相同，但各自具有独立的含义，数据块个数的符号用于表示数据块的数量，而表示数据块的编号的符号仅仅用于区分不同的数据块，数据块的编号还可以以其他形式来表示，例如数字、字符串等。另外，本发明实施例中提到的对数据块进行固定步长加操作、固定步长减操作，步长可以为 1，对数据块进行固定步长加操作的含义与对数据块进行加引用计数操作含义相同，对数据块进行固定步长减操作的含义与对数据块进行减引用计数操作含义相同。具体地，对需要保留的快照和/或检查点引用的一个数据块进行加引用计数操作，即当该数据块被快照或检查点引用一次，引用计数加 1，则该数据块是为该快照或检查点分配的数据块。

本领域内的技术人员应明白，本发明的实施例可提供为方法、系统、或计算机程序产品。因此，本发明可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且，本发明可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质（包括但不限于磁盘存储器、CD-ROM、光学存储器等）上实施的计算机程序产品的形式。

本发明是参照根据本发明实施例的方法、设备（系统）、和计算机程序产品的流程图和/或方框图来描述的。应理解可由计算机程序指令实现流程图和/或方框图中的每一流程和/或方框、以及流程图和/或方框图中的流程和/或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器，使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一

个流程或多个流程和 / 或方框图一个方框或多个方框中指定的功能的装置。

这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中,使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品,该指令装置实现在流程图一个流程或多个流程和 / 或方框图一个方框或多个方框中指定的功能。

这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上,使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理,从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和 / 或方框图一个方框或多个方框中指定的功能的步骤。

显然,本领域的技术人员可以对本发明进行各种改动和变型而不脱离本发明的精神和范围。这样,倘若本发明的这些修改和变型属于本发明权利要求及其等同技术的范围之内,则本发明也意图包含这些改动和变型在内。

1、一种基于写时复制的文件系统中回收检查点的方法，其特征在于，包括：

根据检查点回收指令获得所述文件系统在上次检查点回收时刻到本次检查点回收时刻之间分配的 M 个数据块， M 为不小于 1 的整数；其中，所述 M 个数据块为上次检查点回收时刻到本次检查点回收时刻之间生成的检查点和快照中的至少一种分配的数据块；

对所述 M 个数据块中需要保留的数据块的引用次数进行固定步长的加操作，从所述 M 个数据块中确定需要回收的第一数据块；

从为所述上次检查点回收时刻保留的检查点和快照中的至少一种分配的 N 个数据块中确定需要回收的第二数据块， N 为不小于 1 的整数；

回收所述第一数据块和所述第二数据块。

2、如权利要求 1 所述的方法，其特征在于，所述对所述 M 个数据块中需要保留的数据块的引用次数进行固定步长的加操作，从所述 M 个数据块中确定需要回收的第一数据块，具体包括：

当上次检查点回收时刻到本次检查点回收时刻之间只生成检查点时，则对为当前最新生成的检查点分配的 K 个数据块的引用次数进行所述固定步长的加操作，获得所述 K 个数据块的第一引用计数；从所述 M 个数据块中除所述 K 个数据块外的数据块中确定所述第一数据块；或

当上次检查点回收时刻到本次检查点回收时刻之间既生成检查点又生成快照时，则对为上次检查点回收时刻到本次检查点回收时刻之间生成的快照和当前最新生成的检查点分配的 K 个数据块的引用次数进行所述固定步长的加

操作，获得所述 K 个数据块的第一引用计数；从所述 M 个数据块中除所述 K 个数据块外的数据块中确定所述第一数据块。

3、如权利要求 2 所述的方法，其特征在于，所述从为所述上次检查点回收时刻保留的检查点和快照中的至少一种分配的 N 个数据块中确定需要回收的第二数据块，具体包括：

对所述 N 个数据块中的为所述上次检查点回收时刻保留的检查点分配的 L 个数据块的引用次数进行所述固定步长的减操作，确定所述 L 个数据块的第二引用计数；其中，所述获得所述 K 个数据块的第一引用计数之前，所述 N 个数据块的引用计数为第一参考引用计数；

根据所述第一引用计数、所述第二引用计数和所述第一参考引用计数，从所述 N 个数据块中确定所述第二数据块。

4、如权利要求 3 所述的方法，其特征在于，所述方法还包括：

根据所述第一引用计数、所述第二引用计数和所述第一参考引用计数，确定所述本次检查点回收需要保留的数据块的引用计数作为当前第一参考引用计数。

5、如权利要求 3 所述的方法，其特征在于，在所述文件系统包括上次检查点回收截止根区、本次检查点回收截止根区和实时根区时，其中，所述上次检查点回收截止根区为对所述 N 个数据块进行索引的根区，所述实时根区为对所述 K 个数据块进行索引的根区，所述本次检查点回收截止根区为在所述文件系统获得所述检查点回收指令时复制所述实时根区内索引关系的根区。

6、如权利要求 5 所述的方法，其特征在于，在所述回收所述第一数据块和所述第二数据块之后，所述方法还包括：

删除所述上次检查点回收截止根区内的数据;

将所述本次检查点回收截止根区内的数据复制到所述上次检查点回收截止根区中。

7、一种基于写时复制的文件系统中回收检查点的装置,其特征在于,包括:

获得单元,用于根据检查点回收指令获得文件系统在上次检查点回收时刻到本次检查点回收时刻之间分配的 M 个数据块, M 为不小于 1 的整数;其中,所述 M 个数据块为上次检查点回收时刻到本次检查点回收时刻之间生成的检查点和快照中的至少一种分配的数据块;

第一确定单元,用于对所述 M 个数据块中需要保留的数据块的引用次数进行固定步长的加操作,从所述 M 个数据块中确定需要回收的第一数据块;

第二确定单元,用于从为所述上次检查点回收时刻保留的检查点和快照中的至少一种分配的 N 个数据块中确定需要回收的第二数据块, N 为不小于 1 的整数;

回收单元,用于回收所述第一数据块和所述第二数据块。

8、如权利要求 7 所述的装置,其特征在于,所述第一确定单元具体用于当上次检查点回收时刻到本次检查点回收时刻之间只生成检查点时,则对为当前最新生成的检查点分配的 K 个数据块的引用次数进行所述固定步长的加操作,获得所述 K 个数据块的第一引用计数,从所述 M 个数据块中除所述 K 个数据块外的数据块中确定所述第一数据块;或当上次检查点回收时刻到本次检查点回收时刻之间既生成检查点又生成快照时,则对为上次检查点回收时刻到本次检查点回收时刻之间生成的快照和当前最新生成的检查点分配的 K 个数

据块的引用次数进行所述固定步长的加操作，获得所述 K 个数据块的第一引用计数，所述 M 个数据块中除所述 K 个数据块外的数据块中确定所述第一数据块。

9、如权利要求 8 所述的装置，其特征在于，所述第二确定单元具体用于对所述 N 个数据块中的为所述上次检查点回收时刻保留的检查点分配的 L 个数据块的引用次数进行所述固定步长的减操作，确定所述 L 个数据块当前的引用计数；其中，所述获得所述 K 个数据块的第一引用计数之前，所述 N 个数据块的引用计数为第一参考引用计数，并根据所述第一引用计数、所述第二引用计数和所述第一参考引用计数，从所述 N 个数据块中确定所述第二数据块。

10、如权利要求 9 所述的装置，其特征在于，所述装置还包括第三确定单元，所述第三确定单元用于根据所述第一引用计数、所述第二引用计数和所述第一参考引用计数，确定所述本次检查点回收需要保留的数据块的引用计数作为当前第一参考引用计数。

11、如权利要求 9 所述的装置，其特征在于，在所述文件系统包括上次检查点回收截止根区、本次检查点回收截止根区和实时根区时，其中，所述上次检查点回收截止根区为对所述 N 个数据块进行索引的根区，所述实时根区为对所述文件系统在所述 K 个数据块进行索引的根区，所述本次检查点回收截止根区为在所述文件系统获得所述检查点回收指令时复制所述实时根区内索引关系的根区。

12、如权利要求 11 所述的装置，其特征在于，所述回收单元具体还用于在回收所述第一数据块和所述第二数据块之后，删除所述上次检查点回收截止根区内的数据，并将所述本次检查点回收截止根区内的数据复制到所述上次检查点回收截止根区中。

1/10

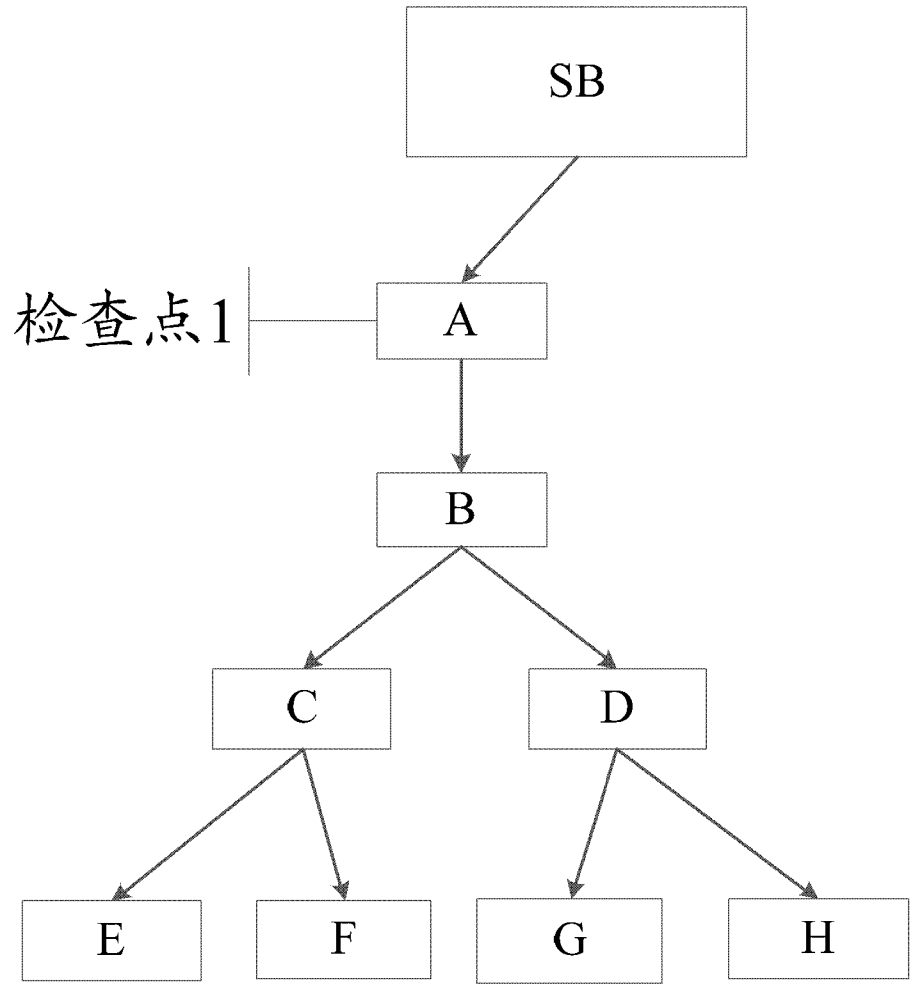


图 1A

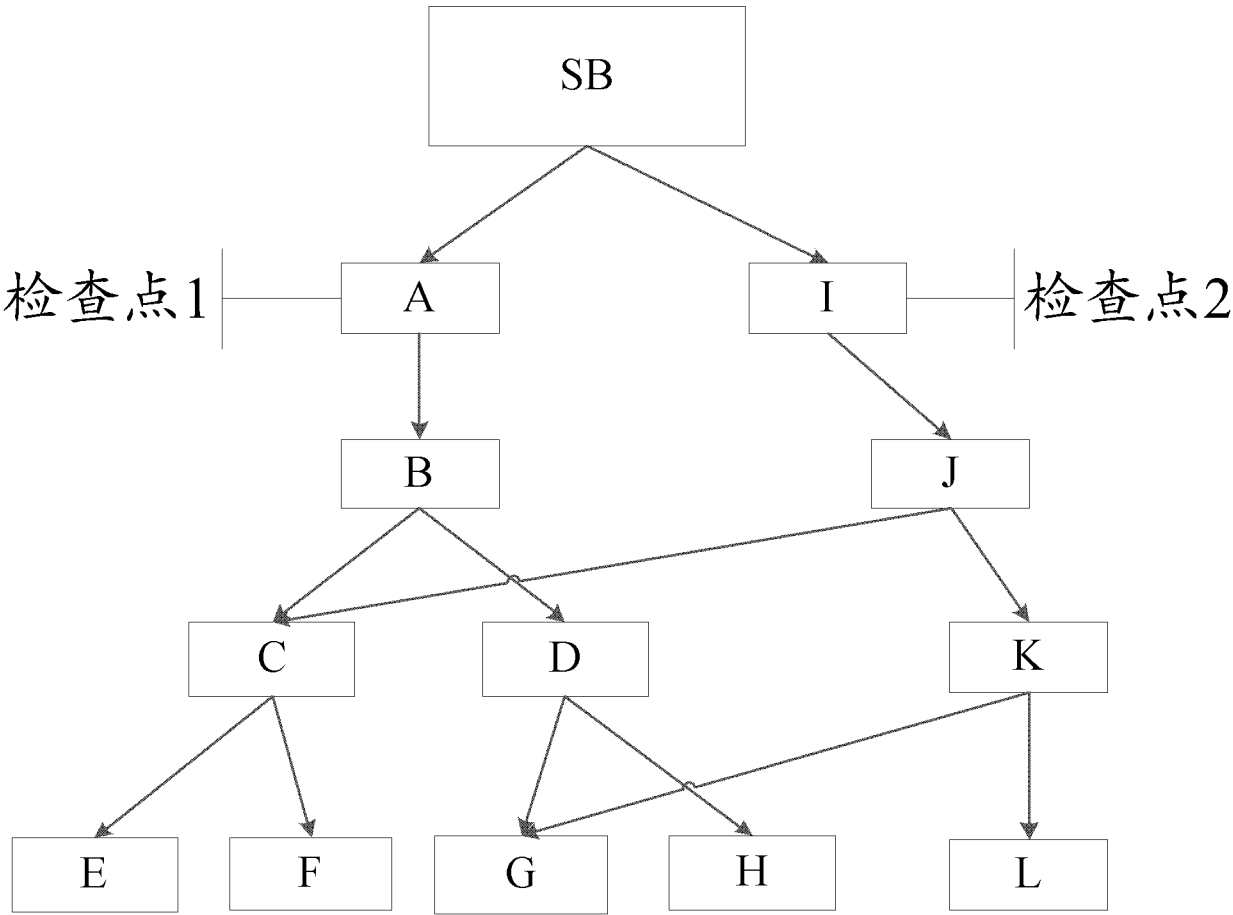


图 1B

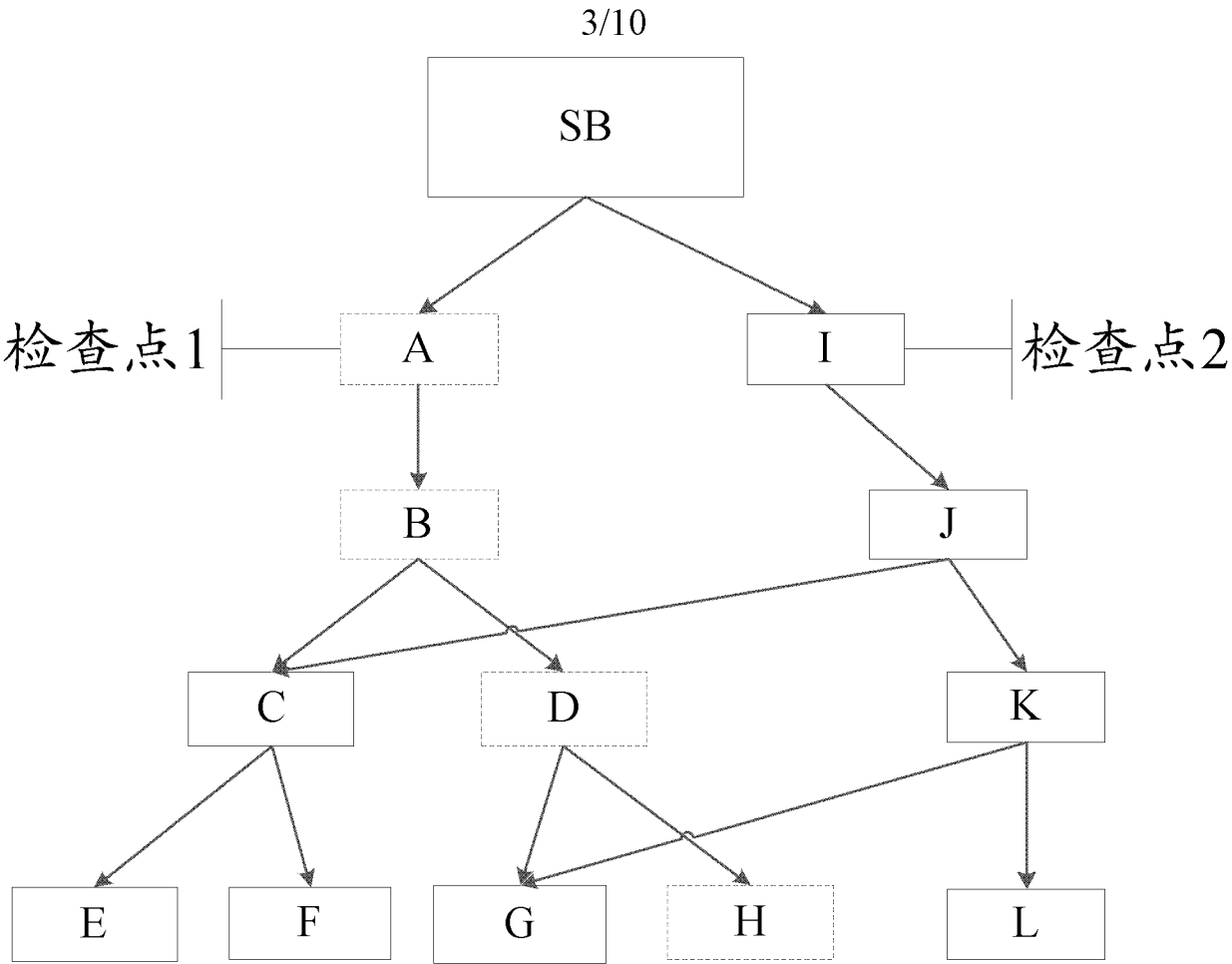


图 1C

4/10

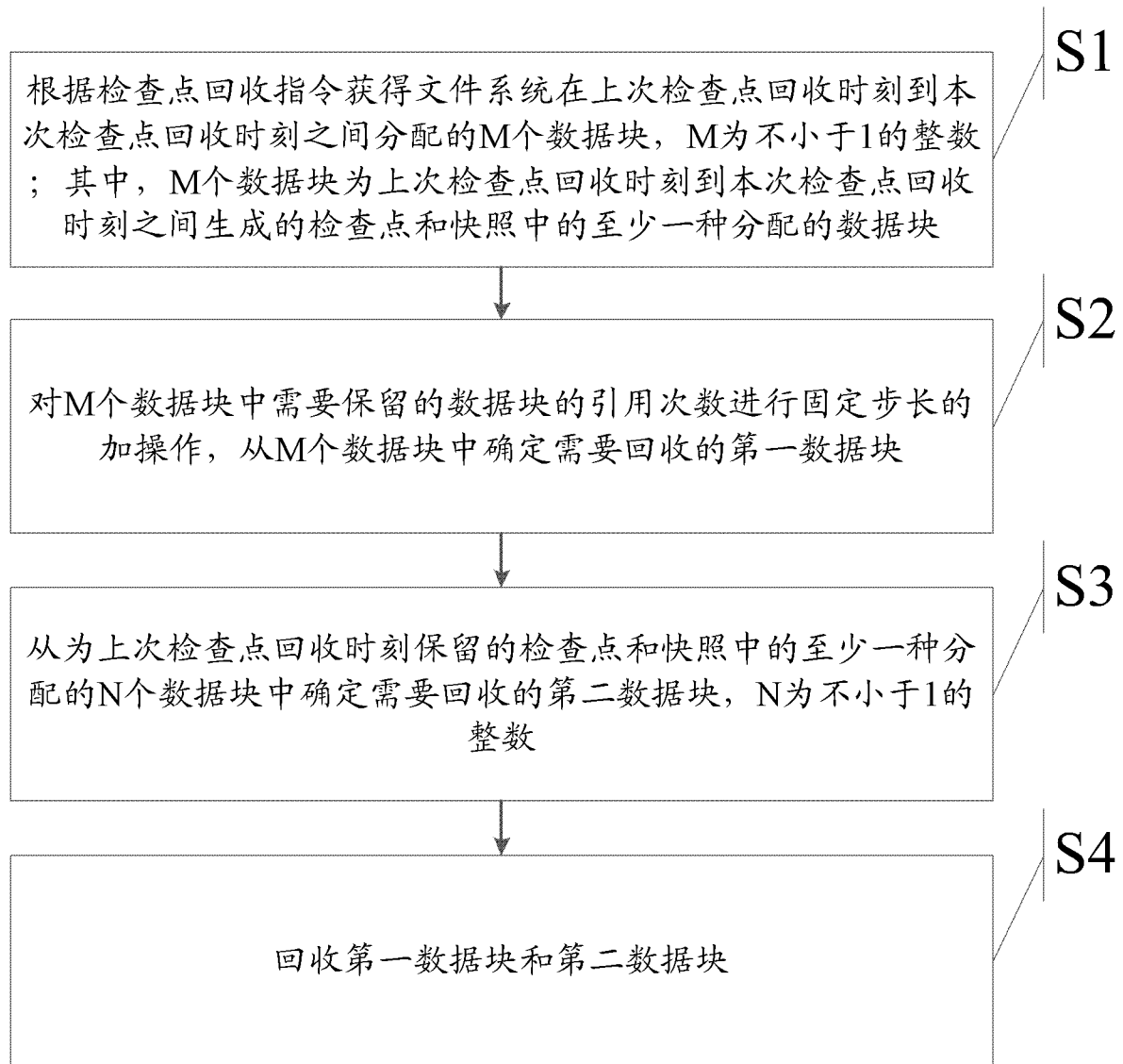


图 2

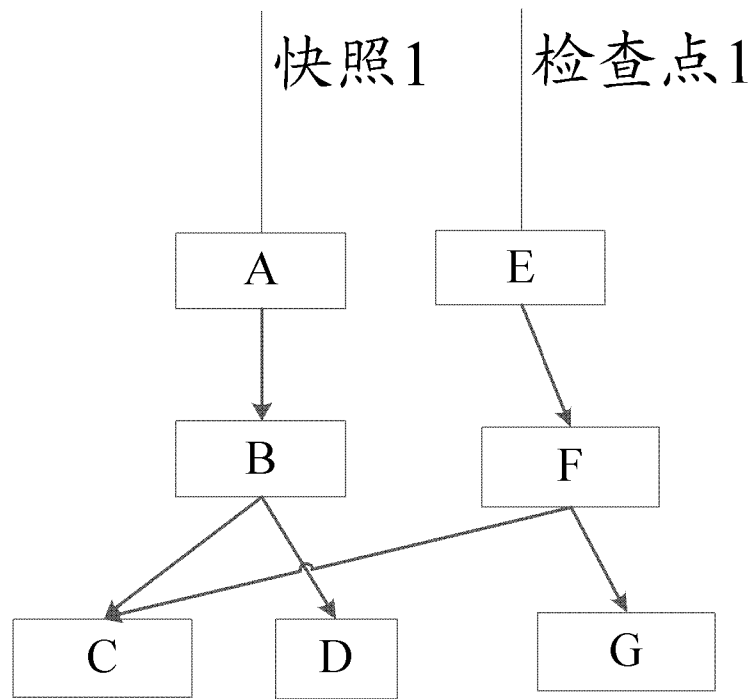


图 3

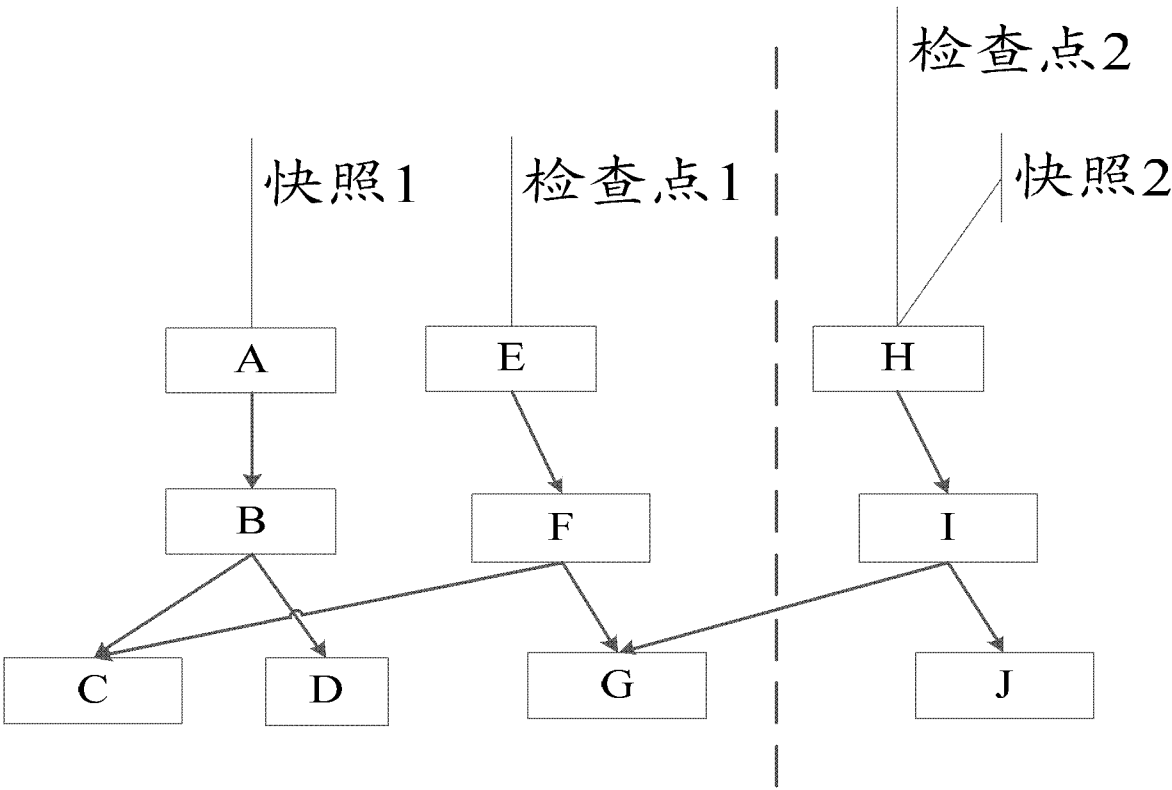


图 4

6/10

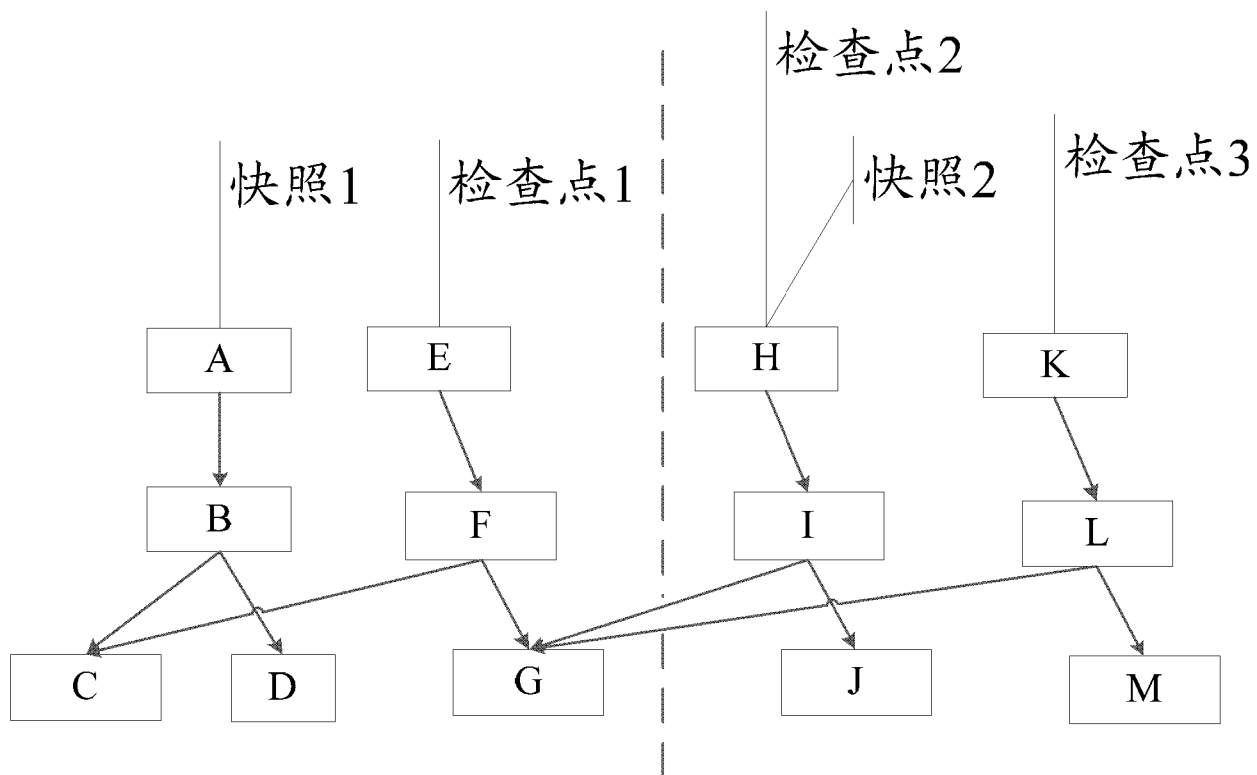


图 5

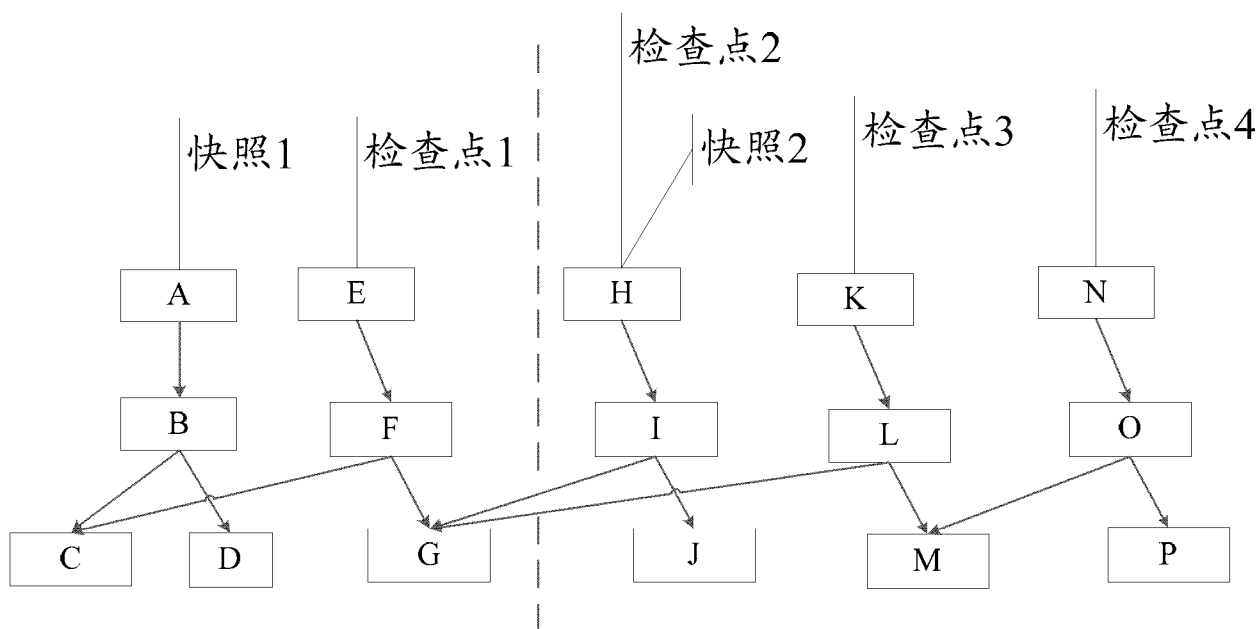


图 6

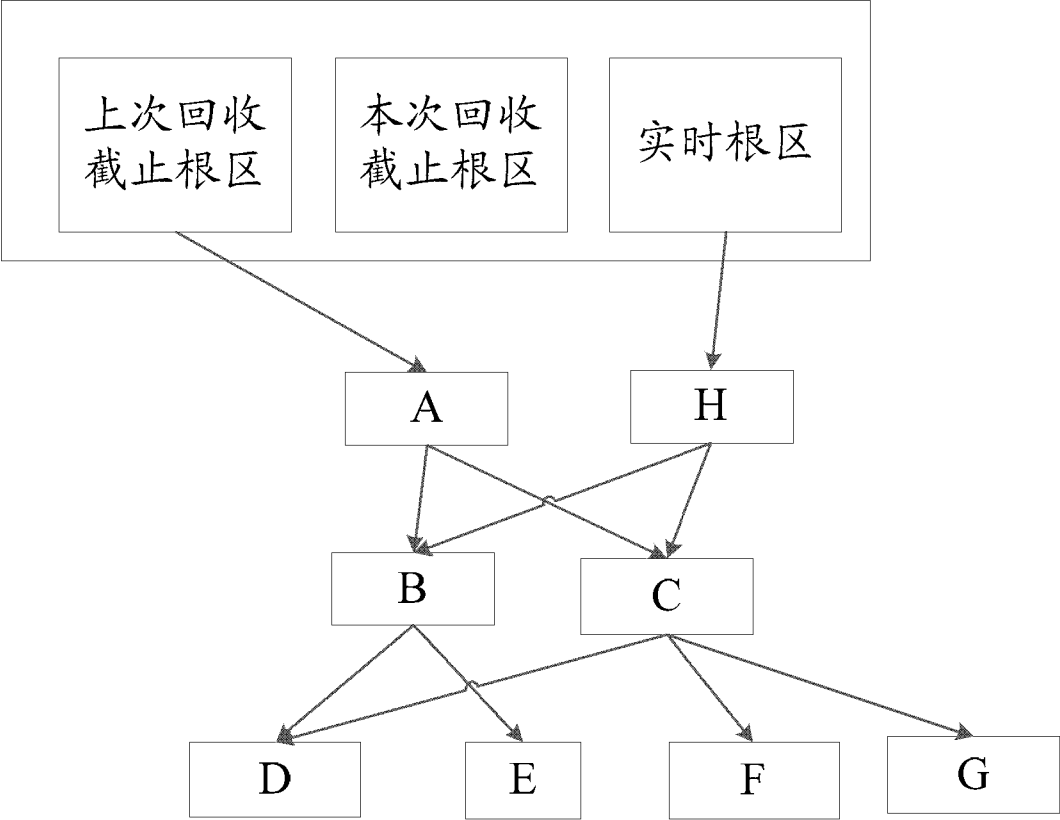


图 7

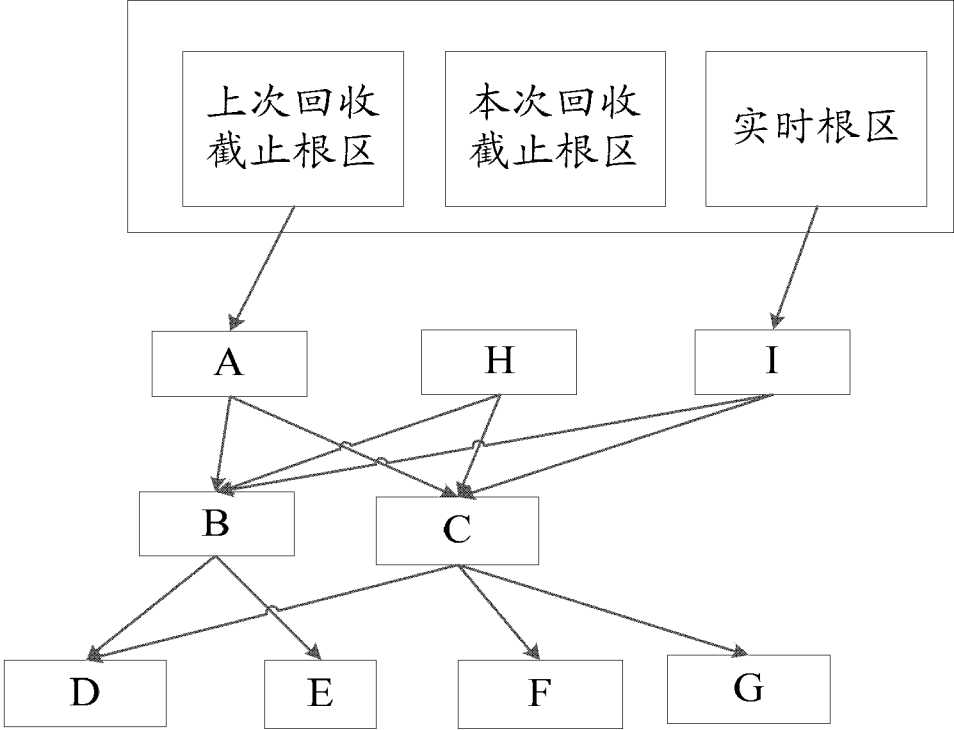


图 8

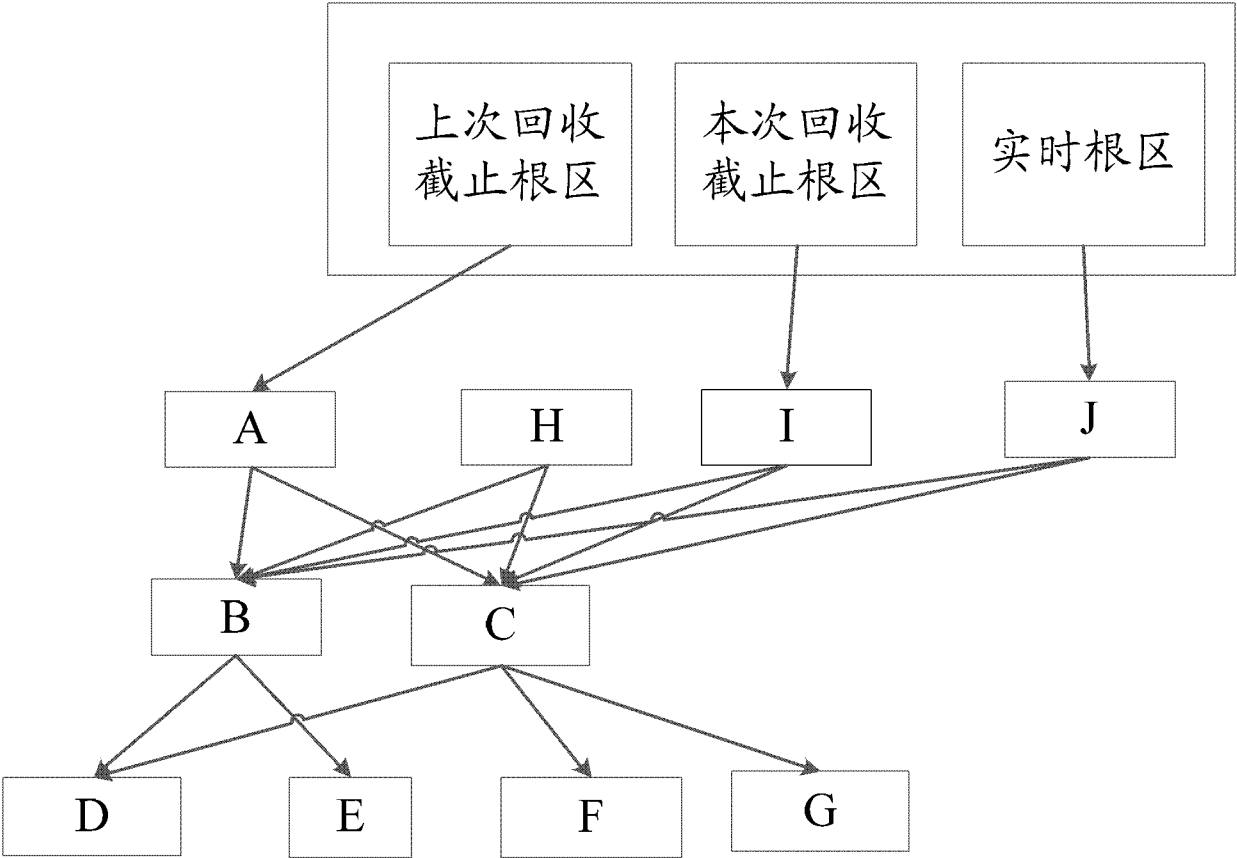


图 9

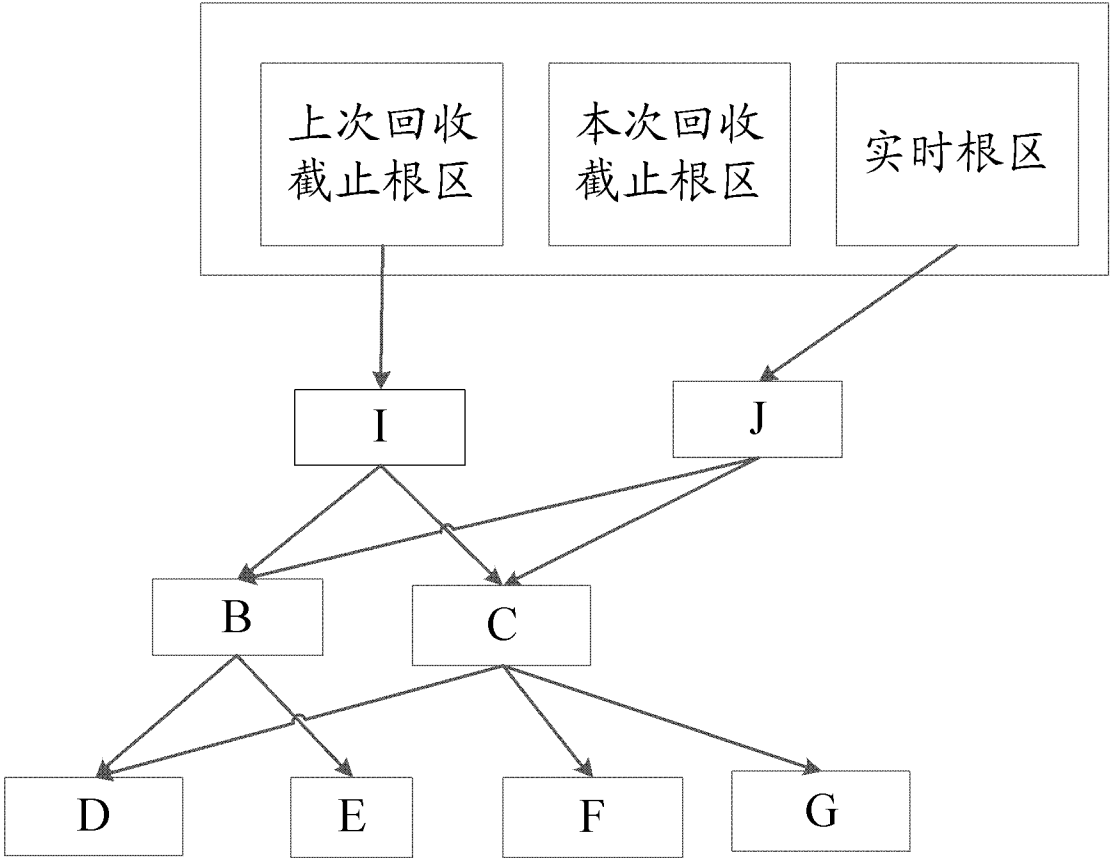


图 10

10/10

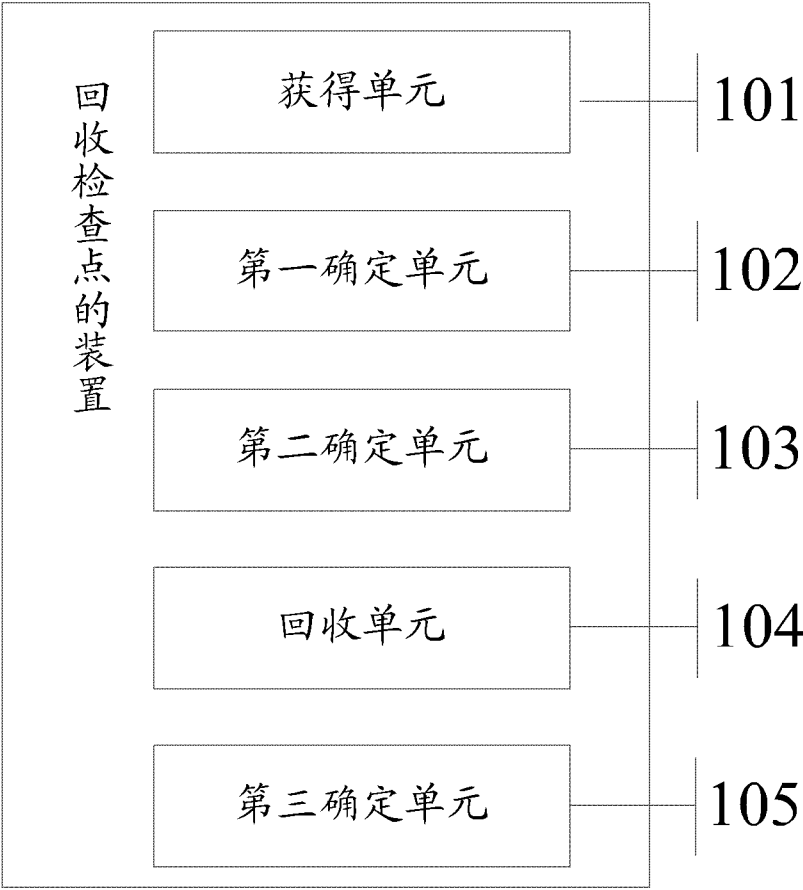


图 11

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2014/089458

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

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

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

CNABS; DWPI; CNKI: XIE, Yong; number of times, ergodic; COW, copy on write, recover, checkpoint, snapshot, check, snap, count+, ergodic

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 103984609 A (HUAWEI TECHNOLOGIES CO., LTD.), 13 August 2014 (13.08.2014), claims 1-12	1-12
A	CN 103678715 A (WUXI CLOUD COMPUTING CENTER CO., LTD.), 26 March 2014 (26.03.2014), the whole document	1-12
A	CN 1532707 A (INSTITUTE OF COMPUTING TECHNOLOGY, CHINESE ACADEMY OF SCIENCES), 29 September 2004 (29.09.2004), the whole document	1-12
A	CN 102968381 A (INSPUR ELECTRONIC INFORMATION INDUSTRY CO., LTD.), 13 March 2013 (13.03.2013), the whole document	1-12
A	US 2003159007 A1 (INTERNATIONAL BUSINESS MACHINES CORPORATION), 21 August 2003 (21.08.2003), the whole document	1-12
A	US 2011161381 A1 (WANG, Wenguang et al.), 30 June 2011 (30.06.2011), the whole document	1-12

☐ 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 04 February 2015 (04.02.2015)	Date of mailing of the international search report 16 February 2015 (16.02.2015)
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 WANG, Xiaoyuan Telephone No.: (86-10) 010-62089294

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2014/089458

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 103984609 A	13 August 2014	None	
CN 103678715 A	26 March 2014	None	
CN 1532707 A	29 September 2004	CN 1258715 C	07 June 2006
CN 102968381 A	13 March 2013	None	
US 2003159007 A1	21 August 2003	TW 591394 B	11 June 2004
		US 6748504 B2	08 June 2004
		TW 200303468 A	01 September 2003
US 2011161381 A1	30 June 2011	US 8849876 B2	30 September 2014
		US 2015006495 A1	01 January 2015

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) G06F 包含在检索领域中的除最低限度文献以外的检索文献 在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用)) CNABS;DWPI;CNKI:谢勇, 写时拷贝, 写时考贝, 写时复制, 回收, 检查, 快照, 计数, 次数, 遍历; COW, copy on write, recover, checkpoint, snapshot, check, snap, count+, ergodic		
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 103984609 A (华为技术有限公司) 2014年 8月 13日 (2014 - 08 - 13) 权利要求1-12	1-12
A	CN 103678715 A (无锡城市云计算中心有限公司) 2014年 3月 26日 (2014 - 03 - 26) 全文	1-12
A	CN 1532707 A (中国科学院计算技术研究所) 2004年 9月 29日 (2004 - 09 - 29) 全文	1-12
A	CN 102968381 A (浪潮电子信息产业股份有限公司) 2013年 3月 13日 (2013 - 03 - 13) 全文	1-12
A	US 2003159007 A1 (国际商业机器公司) 2003年 8月 21日 (2003 - 08 - 21) 全文	1-12
A	US 2011161381 A1 (WANG WENGUANG等) 2011年 6月 30日 (2011 - 06 - 30) 全文	1-12
☐ 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 2015年 2月 04日	国际检索报告邮寄日期 2015年 2月 16日	
ISA/CN的名称和邮寄地址 中华人民共和国国家知识产权局(ISA/CN) 北京市海淀区蓟门桥西土城路6号 100088 中国 传真号 (86-10)62019451	授权官员 王晓渊 电话号码 (86-10)010-62089294	

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2014/089458

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	103984609	A	2014年 8月 13日	无			
CN	103678715	A	2014年 3月 26日	无			
CN	1532707	A	2004年 9月 29日	CN	1258715	C	2006年 6月 07日
CN	102968381	A	2013年 3月 13日	无			
US	2003159007	A1	2003年 8月 21日	TW	591394	B	2004年 6月 11日
				US	6748504	B2	2004年 6月 08日
				TW	200303468	A	2003年 9月 01日
US	2011161381	A1	2011年 6月 30日	US	8849876	B2	2014年 9月 30日
				US	2015006495	A1	2015年 1月 01日