



(12) 发明专利申请

(10) 申请公布号 CN 103473179 A

(43) 申请公布日 2013. 12. 25

(21) 申请号 201310345300. 9

(22) 申请日 2013. 08. 09

(71) 申请人 记忆科技(深圳)有限公司

地址 518000 广东省深圳市南山区蛇口后海
大道东角头东南工贸大厦 5 楼

(72) 发明人 雷伟

(74) 专利代理机构 北京律诚同业知识产权代理
有限公司 11006

代理人 刘健 黄韧敏

(51) Int. Cl.

G06F 12/02 (2006. 01)

权利要求书1页 说明书4页 附图3页

(54) 发明名称

后台系统及其删除固态硬盘重复数据的方法

(57) 摘要

本发明适用于存储技术领域,提供了一种删除固态硬盘重复数据的方法,所述方法包括:通过错误检查和纠正算法计算固态硬盘数据的指纹,并保存为指纹库;查找所述固态硬盘的增量页,根据所述指纹库判断所述固态硬盘中是否有与所述增量页重复的数据;若所述固态硬盘中是有与所述增量页重复的数据,则更新所述固态硬盘的FTL映射表,并将所述增量页标记为无效。本发明还提供一种实现上述方法的后台系统。借此,本发明可以方便的删除固态硬盘的重复数据,提高固态硬盘的性能。

通过错误检查和纠正算法计算固态硬盘数据的指纹,并保存为指纹库 S401

查找固态硬盘的增量页,根据指纹库判断固态硬盘中是否有与增量页重复的数据 S402

若固态硬盘中是有与增量页重复的数据,则更新固态硬盘的FTL映射表,并将增量页标记为无效 S403

1. 一种删除固态硬盘重复数据的方法,其特征在于,所述方法包括:
通过错误检查和纠正算法计算固态硬盘数据的指纹,并保存为指纹库;
查找所述固态硬盘的增量页,根据所述指纹库判断所述固态硬盘中是否有与所述增量页重复的数据;
若所述固态硬盘中是有与所述增量页重复的数据,则更新所述固态硬盘的 FTL 映射表,并将所述增量页标记为无效。
2. 根据权利要求 1 所述的删除固态硬盘重复数据的方法,其特征在于,所述后台通过 LRU 算法处理所述指纹库中的指纹。
3. 根据权利要求 1 所述的删除固态硬盘重复数据的方法,其特征在于,所述根据所述指纹库判断所述固态硬盘中是否有与所述增量页相同的数据步骤包括:
判断所述指纹库中是否有与所述增量页的数据相同的指纹;
若所述指纹库中有与所述增量页的数据相同的指纹,则进一步判断所述指纹对应的数据与所述增量页数据是否相同,若相同则判定为重复数据。
4. 根据权利要求 1 所述的删除固态硬盘重复数据的方法,其特征在于,所述查找所述固态硬盘的增量页的步骤包括:
对所述 FTL 建立时间点快照,记录并获取两个时间点之间的修改的页。
5. 一种删除固态硬盘重复数据的后台系统,其特征在于,包括:
指纹库建立模块,用于通过错误检查和纠正算法计算固态硬盘数据的指纹,并保存为指纹库;
数据比较模块,用于查找所述固态硬盘的增量页,根据所述指纹库判断所述固态硬盘中是否有与所述增量页重复的数据;
数据处理模块,用于所述固态硬盘中是有与所述增量页重复的数据时,更新所述固态硬盘的 FTL 映射表,并将所述增量页标记为无效。
6. 根据权利要求 5 所述的删除固态硬盘重复数据的后台系统,其特征在于,所述后台系统还包括算法处理模块,用于通过 LRU 算法处理所述指纹库中的指纹。
7. 根据权利要求 5 所述的删除固态硬盘重复数据的后台系统,其特征在于,所述数据比较模块包括:
第一判断子模块,用于判断所述指纹库中是否有与所述增量页的数据相同的指纹;
第二判断子模块,用于所述指纹库中有与所述增量页的数据相同的指纹时,进一步判断所述指纹对应的数据与所述增量页数据是否相同,若相同则判定为重复数据。
8. 根据权利要求 5 所述的删除固态硬盘重复数据的后台系统,其特征在于,所述后台系统还包括:
增量页处理模块,用于对所述 FTL 建立时间点快照,记录并获取两个时间点之间的修改的页。

后台系统及其删除固态硬盘重复数据的方法

技术领域

[0001] 本发明涉及存储技术领域,尤其涉及一种后台系统及删除固态硬盘重复数据的方法。

背景技术

[0002] 目前,SSD(Solid State Disk,固态硬盘)产品在市场已经趋于成熟,其高速的性能,尤其是随机性能也广为人知,但仍有较多用户还是担心其寿命问题。影响 SSD 寿命的因素有多种,但其根本原因在于 NAND Flash 编程/擦除次数,即对 NAND Flash 编程/擦除次数越多,损耗越大,寿命越短。重复数据删除是一项减少 NAND Flash 磨损来延长 SSD 寿命的技术,目前大多数重复数据删除大都在 IO 路径上进行计算和匹配,这样会大大降低 IO 性能,即通过性能来换取空间。

[0003] 综上可知,现有的固态硬盘重复数据的处理方法,在实际使用上显然存在不便与缺陷,所以有必要加以改进。

发明内容

[0004] 针对上述的缺陷,本发明的目的在于提供一种后台系统及删除固态硬盘重复数据的方法,其可以方便的删除固态硬盘的重复数据,提高固态硬盘的性能。

[0005] 为了实现上述目的,本发明提供一种删除固态硬盘重复数据的方法,所述方法包括:

[0006] 通过错误检查和纠正算法计算固态硬盘数据的指纹,并保存为指纹库;

[0007] 查找所述固态硬盘的增量页,根据所述指纹库判断所述固态硬盘中是否有与所述增量页重复的数据;

[0008] 若所述固态硬盘中是有与所述增量页重复的数据,则更新所述固态硬盘的 FTL 映射表,并将所述增量页标记为无效。

[0009] 根据本发明的删除固态硬盘重复数据的方法,所述后台通过 LRU 算法处理所述指纹库中的指纹。

[0010] 根据本发明的删除固态硬盘重复数据的方法,所述根据所述指纹库判断所述固态硬盘中是否有与所述增量页相同的数据步骤包括:

[0011] 判断所述指纹库中是否有与所述增量页的数据相同的指纹;

[0012] 若所述指纹库中有与所述增量页的数据相同的指纹,则进一步判断所述指纹对应的数据与所述增量页数据是否相同,若相同则判定为重复数据。

[0013] 根据本发明的删除固态硬盘重复数据的方法,所述查找所述固态硬盘的增量页的步骤包括:

[0014] 对所述 FTL 建立时间点快照,记录并获取两个时间点之间的修改的页。

[0015] 本发明还提供一种删除固态硬盘重复数据的后台系统,包括:

[0016] 指纹库建立模块,用于通过错误检查和纠正算法计算固态硬盘数据的指纹,并保

存为指纹库；

[0017] 数据比较模块,用于查找所述固态硬盘的增量页,根据所述指纹库判断所述固态硬盘中是否有与所述增量页重复的数据；

[0018] 数据处理模块,用于所述固态硬盘中是有与所述增量页重复的数据时,更新所述固态硬盘的 FTL 映射表,并将所述增量页标记为无效。

[0019] 根据本发明的删除固态硬盘重复数据的后台系统,所述后台系统还包括算法处理模块,用于通过 LRU 算法处理所述指纹库中的指纹。

[0020] 根据本发明的删除固态硬盘重复数据的后台系统,所述数据比较模块包括：

[0021] 第一判断子模块,用于判断所述指纹库中是否有与所述增量页的数据相同的指纹；

[0022] 第二判断子模块,用于所述指纹库中有与所述增量页的数据相同的指纹时,进一步判断所述指纹对应的数据与所述增量页数据是否相同,若相同则判定为重复数据。

[0023] 根据本发明的删除固态硬盘重复数据的后台系统,所述后台系统还包括：

[0024] 增量页处理模块,用于对所述 FTL 建立时间点快照,记录并获取两个时间点之间的修改的页。

[0025] 本发明通过错误检查和纠正算法计算固态硬盘数据的指纹,并保存为指纹库,在固态硬盘有数据写入后,查找所述固态硬盘的增量页,根据所述指纹库判断所述固态硬盘中是否有与所述增量页重复的数据；若所述固态硬盘中是有与所述增量页重复的数据,则更新所述固态硬盘的 FTL 映射表,并将所述增量页标记为无效。借此,本发明可以方便的删除固态硬盘的重复数据,提高固态硬盘的性能。优选的是,本发明还可以通过后台系统以 LRU 算法处理所述指纹库中的指纹,减小指纹库的存储空间。

附图说明

[0026] 图 1 是本发明一实施例的后台系统的结构示意图；

[0027] 图 2 是本发明另一实施例的后台系统的结构示意图；

[0028] 图 3 是本发明一实施例的后台操作结构示意图；

[0029] 图 4 是本发明一实施例的删除固态硬盘重复数据的方法流程图；

[0030] 图 5 是本发明另一实施例的删除固态硬盘重复数据的方法流程图。

具体实施方式

[0031] 为了使本发明的目的、技术方案及优点更加清楚明白,以下结合附图及实施例,对本发明进行进一步详细说明。应当理解,此处所描述的具体实施例仅仅用以解释本发明,并不用于限定本发明。

[0032] 参见图 1,本发明提供了一种删除固态硬盘重复数据的后台系统,其与固态硬盘连接,该系统 100 包括指纹库建立模块 10、数据比较模块 20 以及数据处理模块 30,其中：

[0033] 指纹库建立模块 10,用于通过错误检查和纠正算法 (ECC) 计算固态硬盘数据的指纹,并保存为指纹库。本实施例中,固态硬盘以数据页为单位进行读写,即,指纹库中存储有各数据页对应的指纹。本发明通过 ECC 算法对数据页进行计算获取到 ECC 数据作为该数据页的指纹。

[0034] 数据比较模块 20,用于查找所述固态硬盘的增量页,根据所述指纹库判断所述固态硬盘中是否有与所述增量页重复的数据。且本发明中扫描增量页的方式可根据 FTL 设计的不同而不同,本发明通过对 FTL 建立时间点快照,以记录并获取两个时间点之间的修改的页。

[0035] 数据处理模块 30,用于所述固态硬盘中是有与所述增量页重复的数据时,更新所述固态硬盘的 FTL 映射表,并将所述增量页标记为无效。若固态硬盘中没有与所述增量页重复的数据,则将数据交由 FTL 层处理,将其写入到固态硬盘,包括分配物理页,写入逻辑关系等,同时将该数据的指纹写入指纹库进行储存,借此方便固态硬盘再次有数据写入后的判断。

[0036] 再参见图 2,本实施例中,后台系统还包括算法处理模块 40 和增量页处理模块 50,用于通过 LRU 算法处理所述指纹库中的指纹。指纹库保存在 SSD 的保留空间内,可通过对指纹进行替换的方式,来将最容易重复数据对应的指纹留在指纹库中,对不容易重复数据对应的指纹替换出指纹库,并且保证仅占用少量的 SSD 保留空间。

[0037] 所述增量页处理模块 50,用于对所述 FTL 建立时间点快照,记录并获取两个时间点之间的修改的页。

[0038] 进一步的,所述数据比较模块 20 包括:

[0039] 第一判断子模块 21,用于判断所述指纹库中是否有与所述增量页的数据相同的指纹,若有,则交由第二判断子模块 22 处理。

[0040] 第二判断子模块 22,用于所述指纹库中有与所述增量页的数据相同的指纹时,进一步判断所述指纹对应的数据与所述增量页数据是否相同,若相同则判定为重复数据。

[0041] 本发明的指纹采用 ECC 数据,节省计算时间。且本发明的实施例中,并非 ECC 数值相同就判定数据一致,其将 ECC 校验作为一级判断,相同后再比较数据,数据相同则可作为重复数据来处理。借此延长 SSD 的使用寿命,且能大大提高 SSD 的空间利用率及速率。

[0042] 结合图 3 所示的具体实施例中,后台系统对固态硬盘进行扫描,收集指纹库,获取到块 0 的页 0 与块 1 的页 0 的 ECC 值相同,比较两者数据,且判定匹配,则更新 FTL 映射表,并将重复的物理页(块 1 的页 0)标记为数据无效页。

[0043] 参见图 4,本发明提供了一种删除固态硬盘重复数据的方法,其可以通过如图 1 所示的后台系统 100 实现,该方法包括:

[0044] 步骤 S401,通过错误检查和纠正算法(ECC)计算固态硬盘数据的指纹,并保存为指纹库。本实施例中,固态硬盘以数据页为单位进行读写,即,指纹库中存储有各数据页对应的指纹。本发明通过 ECC 算法对数据页进行计算获取到 ECC 数据作为该数据页的指纹。

[0045] 步骤 S402,查找所述固态硬盘的增量页,根据所述指纹库判断所述固态硬盘中是否有与所述增量页重复的数据。且本发明中扫描增量页的方式可根据 FTL 设计的不同而不同,本发明通过对 FTL 建立时间点快照,以记录并获取两个时间点之间的修改的页。

[0046] 步骤 S403,所述固态硬盘中是有与所述增量页重复的数据时,更新所述固态硬盘的 FTL 映射表,并将所述增量页标记为无效。若固态硬盘中没有与所述增量页重复的数据,则将数据交由 FTL 层处理,将其写入到固态硬盘,包括分配物理页,写入逻辑关系等,同时将该数据的指纹写入指纹库进行储存,借此方便固态硬盘再次有数据写入后的判断。

[0047] 参见图 5,本发明提供了另一种删除固态硬盘重复数据的方法,其可以通过如图 2

所示的后台系统 100 实现,该方法包括:

[0048] 步骤 S501,通过错误检查和纠正算法计算固态硬盘数据的指纹,并保存为指纹库。

[0049] 步骤 S502,判断所述指纹库中是否有与所述增量页的数据相同的指纹,若是则执行步骤 S503。本发明中扫描增量页的方式可根据 FTL 设计的不同而不同,本发明通过对 FTL 建立时间点快照,以记录并获取两个时间点之间的修改的页。

[0050] 步骤 S503,判断所述指纹对应的数据与所述增量页数据是否相同。若是则执行步骤 S504。

[0051] 步骤 S504,判断数据为重复数据。

[0052] 步骤 S505,通过 LRU 算法处理所述指纹库中的指纹。

[0053] 综上所述,本发明通过错误检查和纠正算法计算固态硬盘数据的指纹,并保存为指纹库,在固态硬盘有数据写入后,查找所述固态硬盘的增量页,根据所述指纹库判断所述固态硬盘中是否有与所述增量页重复的数据;若所述固态硬盘中是有与所述增量页重复的数据,则更新所述固态硬盘的 FTL 映射表,并将所述增量页标记为无效。借此,本发明可以方便的删除固态硬盘的重复数据,提高固态硬盘的性能。优选的是,本发明还可以通过后台系统以 LRU 算法处理所述指纹库中的指纹,减小指纹库的存储空间。

[0054] 当然,本发明还可有其它多种实施例,在不背离本发明精神及其实质的情况下,熟悉本领域的技术人员当可根据本发明作出各种相应的改变和变形,但这些相应的改变和变形都应属于本发明所附的权利要求的保护范围。

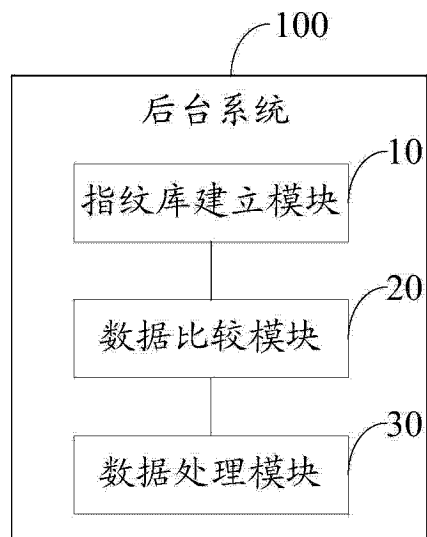


图 1

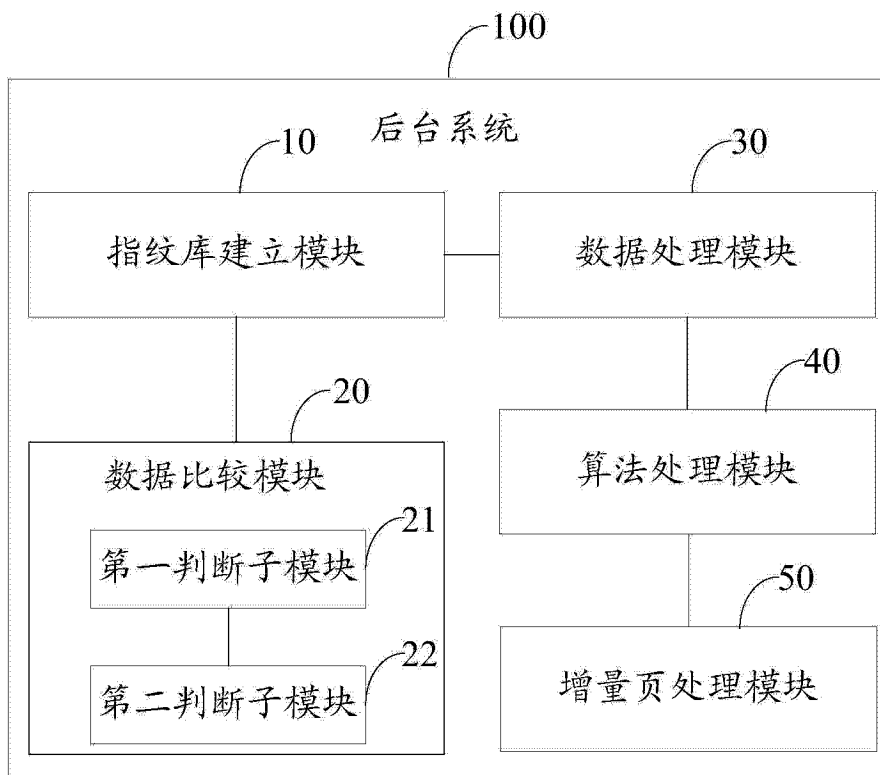


图 2

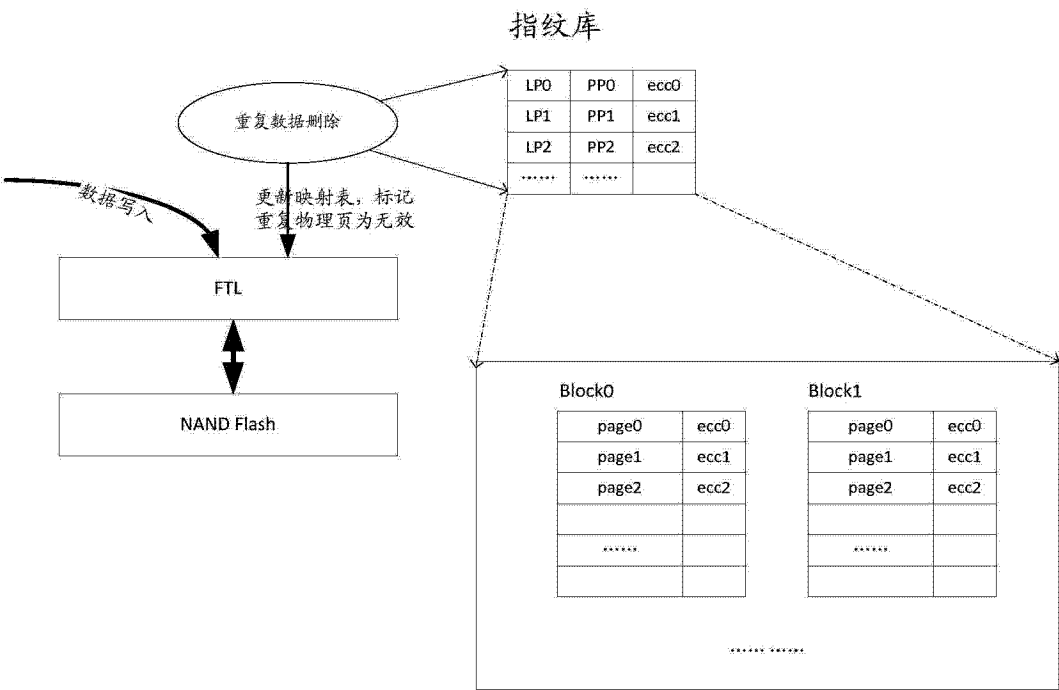


图 3

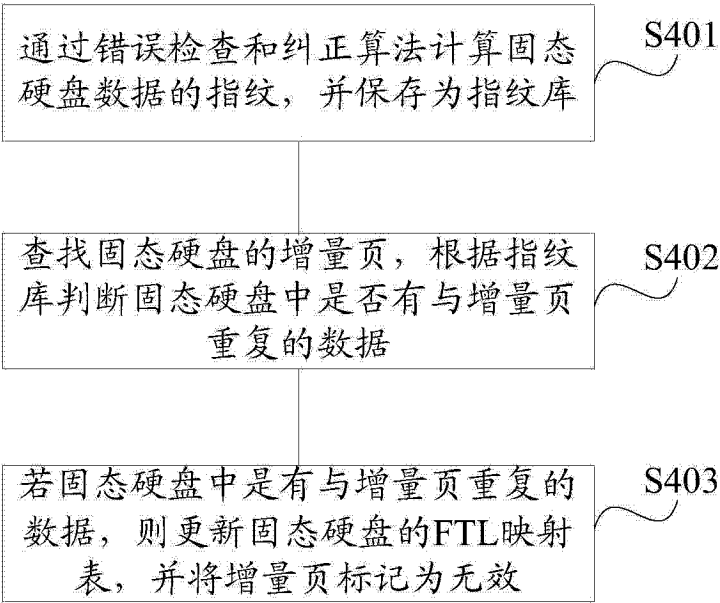


图 4

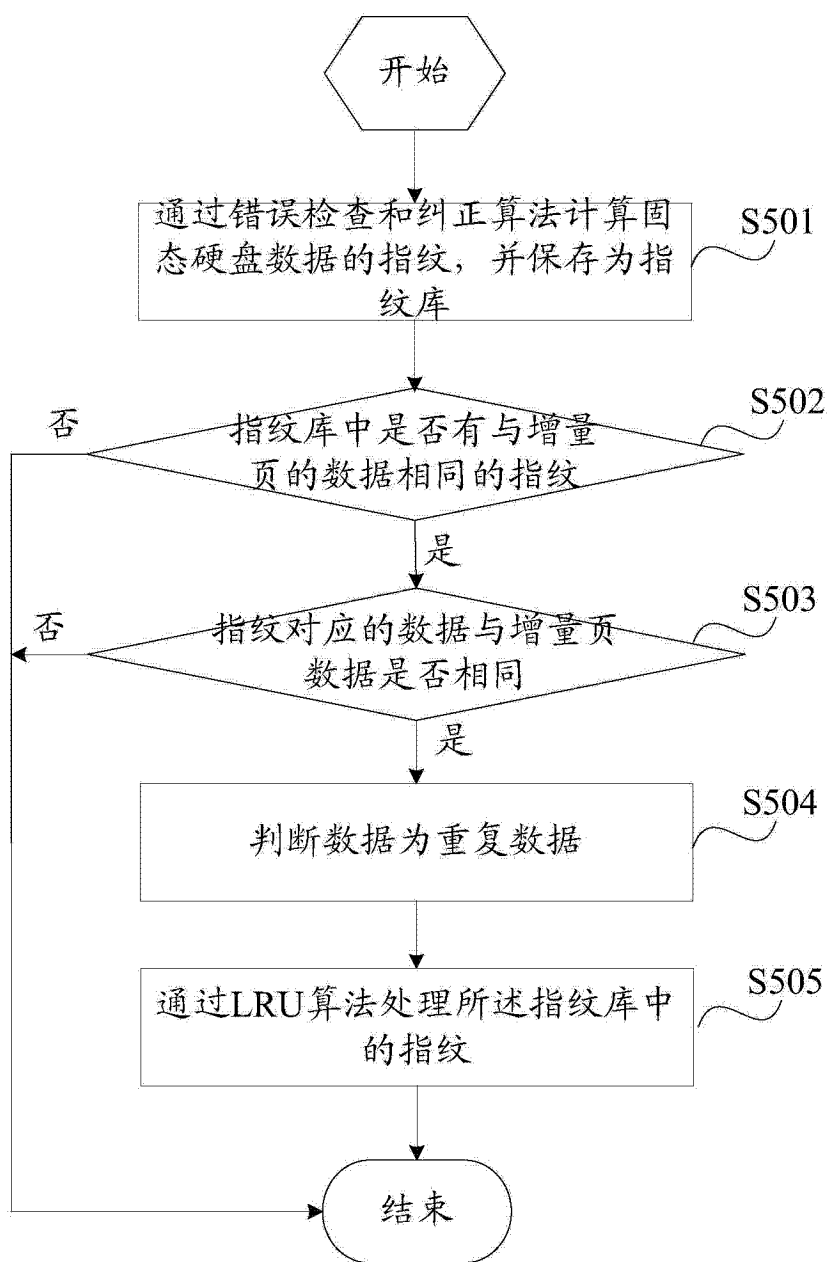


图 5