



(12)发明专利申请

(10)申请公布号 CN 108845770 A

(43)申请公布日 2018. 11. 20

(21)申请号 201810651386.0

(22)申请日 2018.06.22

(71)申请人 深圳忆联信息系统有限公司

地址 518067 广东省深圳市南山区蛇口街
道蛇口后海大道东角头厂房D24/F-02

(72)发明人 付东松 吴娴 张健

(74)专利代理机构 深圳市精英专利事务所
44242

代理人 冯筠

(51) Int. Cl.

G06F 3/06(2006.01)

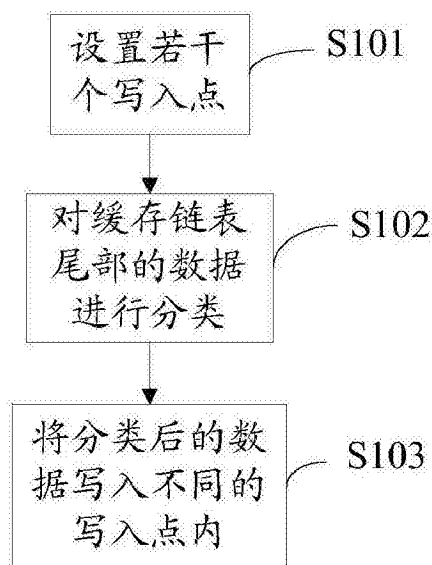
权利要求书1页 说明书7页 附图6页

(54)发明名称

降低SSD写放大的方法、装置及计算机设备

(57)摘要

本发明涉及降低SSD写放大的方法、装置及计算机设备,该方法包括设置若干个写入点;对缓存链表尾部的数据进行分类;将分类后的数据写入不同的写入点内。本发明通过设置若干个写入点,并对缓存中冷数据做更细地冷热程度划分,获取若干个类别的冷数据,将相似冷热程度的数据写到同一写入点所对应的闪存块中,优化缓存中冷数据写入到闪存的方式,降低了垃圾回收的触发,有效降低了写放大,延长了闪存的使用寿命。



1. 降低SSD写放大的方法,其特征在於,包括:
 - 设置若干个写入点;
 - 对缓存链表尾部的数据进行分类;
 - 将分类后的数据写入不同的写入点内。
2. 根据权利要求1所述的降低SSD写放大的方法,其特征在於,设置若干个写入点的步骤,包括以下具体步骤:
 - 选取若干个闪存块,形成写入点;
 - 设置若干个写入点的状态。
3. 根据权利要求1所述的降低SSD写放大的方法,其特征在於,对缓存链表尾部的数据进行分类的步骤,包括以下具体步骤:
 - 提取缓存链表尾部的数据;
 - 按照设定的分类规则将缓存链表尾部的数据进行分类,获取若干个类别的数据。
4. 根据权利要求3所述的降低SSD写放大的方法,其特征在於,所述按照设定的分类规则将缓存链表尾部的数据进行分类,获取若干个类别的数据的步骤中,所述分类规则包括按照设定若干个比例进行数据分类。
5. 降低SSD写放大的装置,其特征在於,包括有:
 - 写入点设置单元,用于设置若干个写入点;
 - 分类单元,用于对缓存链表尾部的数据进行分类;
 - 写入单元,用于将分类后的数据写入不同的写入点内。
6. 根据权利要求5所述的降低SSD写放大的装置,其特征在於,所述写入点设置单元包括有:
 - 选取模块,用于选取若干个闪存块,形成写入点;
 - 状态设置模块,用于设置若干个写入点的状态。
7. 根据权利要求6所述的降低SSD写放大的装置,其特征在於,所述分类单元包括有:
 - 数据提取模块,用于提取缓存链表尾部的数据;
 - 分类处理模块,用于按照设定的分类规则将缓存链表尾部的数据进行分类,获取若干个类别的数据。
8. 一种计算机设备,其特征在於,包括存储器、处理器及存储在所述存储器上并可在所述处理器上运行的计算机程序,所述处理器执行所述计算机程序时实现如权利要求1至4中任意一项所述的降低SSD写放大的方法。

降低SSD写放大的方法、装置及计算机设备

技术领域

[0001] 本发明涉及固态硬盘写操作处理方法,更具体地说是降低SSD写放大的方法、装置及计算机设备。

背景技术

[0002] 在固态硬盘(即SSD)中,主机发起的数据先缓存在DRAM中,DRAM的管理策略需要保证数据的一致性,即如果后续主机又下发了相同逻辑地址(即LBA)的数据到DRAM,最终写到存储介质(主流介质为Nand Flash)的数据必须为最后一次主机下发的数据;如果主机下发多次相同LBA的数据至DRAM后下发读该LBA的命令,那么SSD返回给主机的数据也必须是最后一次写入的数据。

[0003] 在数据中心、企业级应用场景下,主机会频繁下发一些活跃的LBA,即某些LBA会被频繁的复写,这些活跃的LBA和一些冷数据一起发往SSD,现有缓存技术中,用LRU算法将LBA信息和对应数据以链表方式存储,如图1和图2,当缓存已存满后,将最热的数据(这部分数据很有可能被主机复写)保留在缓存中,其他的LBA数据作为冷数据写入到闪存中,这些数据被复写的概率比较小,垃圾回收机制触发的几率也小,而剩下的最近最多使用(即位于链表头上)LBA将继续停留在缓存中,因为它们将来被复写的概率最大,一旦将来真发生命中,则在缓存中进行数据覆盖,在缓存方式上粗略地隔离了冷热数据,能降低垃圾回收的概率,即降低写方法。

[0004] 现有的缓存具体流程如下:采用链表方式存储LBA信息和对应数据,如图3所示,这样容易更改缓存中LBA的位置,譬如缓存容量为10个LBA,主机先发起了6个LBA,顺序为LBA8、LBA4、LBA3、LBA7、LBA0、LBA1,SSD固件处理每个LBA时会遍历链表中已缓存的LBA,如果发生命中则进行数据覆盖并且将对应LBA迁移至链表头部,否则直接将LBA挂在链表头部。这6个LBA都没有发生命中情况,最终在链表中位置如图3所示。LBA0和LBA7是主机端的活跃LBA,即主机经常会读写的LBA。接着主机发起6个LBA,顺序为LBA5、LBA2、LBA6、LBA0、LBA7、LBA9,LBA5、LBA2和LBA6没有命中缓存,而且是最新数据,直接将它们挂在链表头部,如图4所示;LBA0和LBA7命中缓存,则将它们从原来位置迁移到链表头部,如图5所示;LBA9没有命中缓存,将它挂在链表头部,如图6所示;由于缓存容量为10个LBA,此时,图6中展示的控制缓存已满,触发了刷缓存机制,而且链表头部缓存的很有可能是活跃LBA,所以需要将链表头部的一部分数据保留下来,假设保留百分之三十(该数值可以根据实际场景灵活设置)的数据,即3个LBA,那么链表尾部的7个LBA和对应数据需要写到闪存中,之后缓存状态如图7。可以看到图7中留在缓存中的LBA是最近最多写入的数据,即活跃LBA,它们被复写的概率也是最大的,而已经写入到闪存中的数据是最近最少被写入的,后续被复写的概率也是最低的,所以触发垃圾回收的几率也低。但是该技术将缓存链表尾部的数据统一当成冷数据写到闪存块中,不够精细,如图8所示,如果这些数据冷热程度差距比较大的话,比如主机写入的数据按冷热程度可以分为五档,档位越高数据越热,那么现有技术只能处理最热的N档数据,剩余5-N档数据写入到相同闪存块中,如图8,那么该闪存块也比较容易出现

部分无效页的情况,即垃圾回收机制也仍比较容易触发,依然存在写放大的问题。

[0005] 因此,有必要设计一种新的方法,优化缓存中冷数据写入到闪存的方式,以实现降低SSD写放大,延长闪存寿命。

发明内容

[0006] 本发明的目的在于克服现有技术的缺陷,提供降低SSD写放大的方法、装置及计算机设备。

[0007] 为实现上述目的,本发明采用以下技术方案:降低SSD写放大的方法,包括:

[0008] 设置若干个写入点;

[0009] 对缓存链表尾部的数据进行分类;

[0010] 将分类后的数据写入不同的写入点内。

[0011] 其进一步技术方案为:设置若干个写入点的步骤,包括以下具体步骤;

[0012] 选取若干个闪存块,形成写入点;

[0013] 设置若干个写入点的状态。

[0014] 其进一步技术方案为:对缓存链表尾部的数据进行分类的步骤,包括以下具体步骤:

[0015] 提取缓存链表尾部的数据;

[0016] 按照设定的分类规则将缓存链表尾部的数据进行分类,获取若干个类别的数据。

[0017] 其进一步技术方案为:所述按照设定的分类规则将缓存链表尾部的数据进行分类,获取若干个类别的数据的步骤中,所述分类规则包括按照设定若干个比例进行数据分类。

[0018] 本发明还提供了降低SSD写放大的装置,包括有:

[0019] 写入点设置单元,用于设置若干个写入点;

[0020] 分类单元,用于对缓存链表尾部的数据进行分类;

[0021] 写入单元,用于将分类后的数据写入不同的写入点内。

[0022] 其进一步技术方案为:所述写入点设置单元包括有:

[0023] 选取模块,用于选取若干个闪存块,形成写入点;

[0024] 状态设置模块,用于设置若干个写入点的状态。

[0025] 其进一步技术方案为:所述分类单元包括有:

[0026] 数据提取模块,用于提取缓存链表尾部的数据;

[0027] 分类处理模块,用于按照设定的分类规则将缓存链表尾部的数据进行分类,获取若干个类别的数据。

[0028] 本发明还提供了一种计算机设备,包括存储器、处理器及存储在所述存储器上并可在所述处理器上运行的计算机程序,所述处理器执行所述计算机程序时实现上述的降低SSD写放大的方法。

[0029] 本发明与现有技术相比的有益效果是:本发明的降低SSD写放大的方法,通过设置若干个写入点,并对缓存中冷数据做更细地冷热程度划分,获取若干个类别的冷数据,将相似冷热程度的数据写到同一写入点所对应的闪存块中,优化缓存中冷数据写入到闪存的方式,降低了垃圾回收的触发,有效降低了写放大,延长了闪存的使用寿命。

[0030] 下面结合附图和具体实施例对本发明作进一步描述。

附图说明

- [0031] 图1为现有技术的链表的示意性框图；
- [0032] 图2为现有技术的链表缓存中保留热数据的示意性框图；
- [0033] 图3为现有技术的链表缓存的示意性框图；
- [0034] 图4为现有技术的链表缓存中LBA5和LBA2以及LBA6未命中缓存的示意性框图；
- [0035] 图5为现有技术的链表缓存中LBA0和LBA7未命中缓存的示意性框图；
- [0036] 图6为现有技术的链表缓存中LBA9未命中缓存的示意性框图；
- [0037] 图7为现有技术的链表缓存中刷缓存的示意性框图；
- [0038] 图8为现有技术的链表缓存中冷数据写入同一闪存块的示意性框图；
- [0039] 图9为本发明具体实施例提供的降低SSD写放大的方法的示意流程图；
- [0040] 图10为本发明具体实施例提供的降低SSD写放大的方法中冷数据细分类写入闪存的示意性框图；
- [0041] 图11为本发明具体实施例提供的降低SSD写放大的方法的子步骤示意流程图；
- [0042] 图12为本发明具体实施例提供的降低SSD写放大的方法的子步骤示意流程图；
- [0043] 图13为本发明具体实施例提供的降低SSD写放大的装置的示意性框图；
- [0044] 图14为本发明具体实施例提供的写入点设置单元的示意性框图；
- [0045] 图15为本发明具体实施例提供的分类单元的示意性框图；
- [0046] 图16为本发明具体实施例提供的一种计算机设备的示意性框图。

具体实施方式

[0047] 为了更充分理解本发明的技术内容,下面结合具体实施例对本发明的技术方案进一步介绍和说明,但不局限于此。

[0048] 应当理解,当在本说明书和所附权利要求书中使用时,术语“包括”和“包含”指示所描述特征、整体、步骤、操作、元素和/或组件的存在,但并不排除一个或多个其它特征、整体、步骤、操作、元素、组件和/或其集合的存在或添加。

[0049] 还应当理解,在此本申请说明书中所使用的术语仅仅是出于描述特定实施例的目的而并不意在限制本申请。如在本申请说明书和所附权利要求书中所使用的那样,除非上下文清楚地指明其它情况,否则单数形式的“一”、“一个”及“该”意在包括复数形式。

[0050] 还应当进一步理解,在本申请说明书和所附权利要求书中使用的术语“和/或”是指相关联列出的项中的一个或多个的任何组合以及所有可能组合,并且包括这些组合。

[0051] 如图9~16所示的具体实施例,本实施例提供的降低SSD写放大的方法、装置及计算机设备,可以运用在固态硬盘对主机的数据进行写入的操作中,实现优化缓存中冷数据写入到闪存的方式,以实现降低SSD写放大,延长闪存寿命。

[0052] 请参阅图9,图9是本具体实施例提供的降低SSD写放大的方法的示意流程图;如图9所示,降低SSD写放大的方法,包括步骤S101~S103。

[0053] S101、设置若干个写入点。

[0054] 请参阅图10,举个例子,SSD固件选取闪存中3个闪存块作为3档数据的写入点,每

个写入点对应不同的闪存块,再按照数据的不同类别存储到不同的闪存块中,该闪存块上物理页在大概率上是同时失效的,不会触发垃圾回收,也不会产生写放大。

[0055] 在一实施例中,如图11所示,上述的步骤S101可包括有S1011~S1012:

[0056] S1011、选取若干个闪存块,形成写入点;

[0057] S1012、设置若干个写入点的状态。

[0058] 具体地,至于闪存块的个数,在实际使用中该值可根据主机实际业务特征灵活配置,也就是写入点的个数可以根据主机实际业务特征做灵活配置,并且一旦设定好写入点,则固件需要管理好每个写入点的状态,以确保每个写入点可写入数据。

[0059] S102、对缓存链表尾部的数据进行分类。

[0060] 在一实施例中,如图12所示,上述的步骤S102可包括S1021~S1022。

[0061] S1021、提取缓存链表尾部的数据;

[0062] S1022、按照设定的分类规则将缓存链表尾部的数据进行分类,获取若干个类别的数据。

[0063] 上述的分类规则包括按照设定若干个比例进行数据分类。

[0064] 具体地,举个例子:当主机写完LBA8、LBA4、LBA3、LBA7、LBA0、LBA1、LBA5、LBA2、LBA6、LBA0、LBA7、LBA9后,保留百分之三十的最热数据在缓存中,剩下的百分之七十(在本实施例中为7个LBA,实际使用时可能有成百上千个)分成3档,由于在链表尾部的为最近最少使用的数据,即最冷的数据,分类规则为:从缓存链表最尾部起头百分之三十数据为第1档数据,次百分之三十为第2档数据,再次百分之十为第3档数据,最后百分之三十的数据保留在缓存中,不写入闪存,如图10所示。

[0065] 当然,也可以按照其他百分比来划分冷门数据的档次或类别,将缓存链表尾部的冷数据进一步细化为X档,闪存上同样开辟出X个写入点,每个写入点对应不同的闪存块,SSD固件管理这些写入点,刷缓存时,将每一档数据写入到对应的写入点,这样就能将同一档位(相同或近似冷热程度)的数据写入到同一闪存块,那么该闪存块上物理页在大概率上是同时失效的,不会触发垃圾回收,也不会产生写放大。上述的X值可以依据实际情况而定。

[0066] S103、将分类后的数据写入不同的写入点内。

[0067] 当逻辑地址中缓存已满,则将第一档数据写入到第一写入点的闪存块中,将第二档数据写入到第二写入点的闪存块中,第三档数据写入到第三写入点的闪存块中,如果写入期间某写入点出现闪存块已满的情况,SSD固件负责重新分配一个闪存块给该写入点。

[0068] 上述的降低SSD写放大的方法,通过设置若干个写入点,并对缓存中冷数据做更细地冷热程度划分,获取若干个类别的冷数据,将相似冷热程度的数据写到同一写入点所对应的闪存块中,优化缓存中冷数据写入到闪存的方式,降低了垃圾回收的触发,有效降低了写放大,延长了闪存的使用寿命。

[0069] 请参阅图13,图13为本具体实施例提供的降低SSD写放大的装置的示意性框图;如图13所示,降低SSD写放大的装置,包括有:

[0070] 写入点设置单元1,用于设置若干个写入点。

[0071] 分类单元2,用于对缓存链表尾部的数据进行分类。

[0072] 写入单元3,用于将分类后的数据写入不同的写入点内。

[0073] 具体地,如图14所示,写入点设置单元1包括有:

[0074] 选取模块11,用于选取若干个闪存块,形成写入点;

[0075] 状态设置模块12,用于设置若干个写入点的状态。

[0076] 另外,如图15所示,分类单元2包括有:

[0077] 数据提取模块21,用于提取缓存链表尾部的数据;

[0078] 分类处理模块22,用于按照设定的分类规则将缓存链表尾部的数据进行分类,获取若干个类别的数据。

[0079] 上述的降低SSD写放大的装置,通过设置若干个写入点,并对缓存中冷数据做更细地冷热程度划分,获取若干个类别的冷数据,将相似冷热程度的数据写到同一写入点所对应的闪存块中,优化缓存中冷数据写入到闪存的方式,降低了垃圾回收的触发,有效降低了写放大,延长了闪存的使用寿命。

[0080] 所属领域的技术人员可以清楚地了解到,为了描述的方便和简洁,上述描述的降低SSD写放大的装置和单元的具体工作过程,可以参考前述方法实施例中的对应过程,在此不再赘述。

[0081] 上述降低SSD写放大的装置可以实现为一种计算机程序的形式,计算机程序可以在如图16所示的计算机设备上运行。

[0082] 请参阅图16,图16是本申请实施例提供的一种计算机设备的示意性框图。该计算机设备700设备可以是终端或服务器。

[0083] 参照图16,该计算机设备700包括通过系统总线710连接的处理器720、存储器和网络接口750,其中,存储器可以包括非易失性存储介质730和内存存储器740。

[0084] 该非易失性存储介质730可存储操作系统731和计算机程序732。该计算机程序732被执行时,可使得处理器720执行任意一种降低SSD写放大的方法。

[0085] 该处理器720用于提供计算和控制能力,支撑整个计算机设备700的运行。

[0086] 该内存存储器740为非易失性存储介质730中的计算机程序732的运行提供环境,该计算机程序732被处理器720执行时,可使得处理器720执行任意一种降低SSD写放大的方法。

[0087] 该网络接口750用于进行网络通信,如发送分配的任务等。本领域技术人员可以理解,图16中示出的结构,仅仅是与本申请方案相关的部分结构的框图,并不构成对本申请方案所应用于其上的计算机设备700的限定,具体的计算机设备700可以包括比图中所示更多或更少的部件,或者组合某些部件,或者具有不同的部件布置。其中,所述处理器720用于运行存储在存储器中的程序代码,以实现以下步骤:

[0088] 设置若干个写入点;

[0089] 对缓存链表尾部的数据进行分类;

[0090] 将分类后的数据写入不同的写入点内。

[0091] 在一实施例中,所述处理器720用于运行存储在存储器中的程序代码以实现所述设置若干个写入点的步骤,具体实现以下步骤:

[0092] 选取若干个闪存块,形成写入点;

[0093] 设置若干个写入点的状态。

[0094] 在一实施例中,所述处理器720用于运行存储在存储器中的程序代码以实现所述对缓存链表尾部的数据进行分类的步骤,具体实现以下步骤:

[0095] 提取缓存链表尾部的数据；

[0096] 按照设定的分类规则将缓存链表尾部的数据进行分类,获取若干个类别的数据。

[0097] 在一实施例中,所述处理器720用于运行存储在存储器中的程序代码以实现所述按照设定的分类规则将缓存链表尾部的数据进行分类,获取若干个类别的数据的步骤中,所述分类规则包括按照设定若干个比例进行数据分类。

[0098] 上述的一种计算机设备,通过设置若干个写入点,并对缓存中冷数据做更细地冷热程度划分,获取若干个类别的冷数据,将相似冷热程度的数据写到同一写入点所对应的闪存块中,优化缓存中冷数据写入到闪存的方式,降低了垃圾回收的触发,有效降低了写放大,延长了闪存的使用寿命。

[0099] 应当理解,在本申请实施例中,处理器720可以是中央处理单元(Central Processing Unit,CPU),该处理器720还可以是其他通用处理器、数字信号处理器(Digital Signal Processor,DSP)、专用集成电路(Application Specific Integrated Circuit,ASIC)、现成可编程门阵列(Field-Programmable Gate Array,FPGA)或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。其中,通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。

[0100] 本领域技术人员可以理解,图16中示出的计算机设备700结构并不构成对计算机设备700的限定,可以包括比图示更多或更少的部件,或者组合某些部件,或者不同的部件布置。

[0101] 本领域普通技术人员可以理解的是实现上述实施例方法中的全部或部分流程,是可以通过计算机程序来指令相关的硬件来完成,计算机程序可存储于一存储介质中,该存储介质为计算机可读存储介质。如本发明实施例中,该计算机程序可存储于计算机系统的存储介质中,并被该计算机系统至少一个处理器执行,以实现包括如上述各降低SSD写放大的方法的实施例的流程步骤。

[0102] 该计算机可读存储介质可以是磁碟、光盘、U盘、移动硬盘、只读存储器(ROM,Read-Only Memory)、磁碟或者光盘等各种可以存储程序代码的介质。

[0103] 本领域普通技术人员可以意识到,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,能够以电子硬件、计算机软件或者二者的结合来实现,为了清楚地说明硬件和软件的可互换性,在上述说明中已经按照功能一般性地描述了各示例的组成及步骤。这些功能究竟以硬件还是软件方式来执行,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申请的范围。

[0104] 在本申请所提供的几个实施例中,应该理解到,所揭露的降低SSD写放大的装置和方法,可以通过其它的方式实现。例如,以上所描述的降低SSD写放大的装置实施例仅仅是示意性的。例如,各个单元的划分,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。例如多个单元或组件可以结合或者可以集成到另一个系统,或一些特征可以忽略,或不执行。

[0105] 本申请实施例方法中的步骤可以根据实际需要进行顺序调整、合并和删减。本申请实施例装置中的单元可以根据实际需要进行合并、划分和删减。

[0106] 另外,在本申请各个实施例中的各功能单元可以集成在一个处理单元中,也可以

是各个单元单独物理存在,也可以是两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现,也可以采用软件功能单元的形式实现。

[0107] 该集成的单元如果以软件功能单元的形式实现并作为独立的产品销售或使用,可以存储在一个计算机可读取存储介质中。基于这样的理解,本申请的技术方案本质上或者说对现有技术做出贡献的部分,或者该技术方案的全部或部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质中,包括若干指令用以使得一台计算机设备(可以是个人计算机,终端,或者网络设备)执行本申请各个实施例所述方法的全部或部分步骤。

[0108] 上述仅以实施例来进一步说明本发明的技术内容,以便于读者更容易理解,但不代表本发明的实施方式仅限于此,任何依本发明所做的技术延伸或再创造,均受本发明的保护。本发明的保护范围以权利要求书为准。

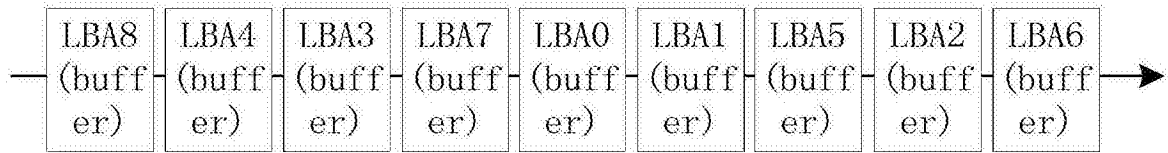


图1

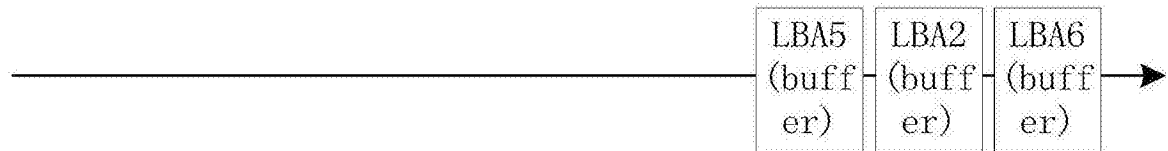


图2

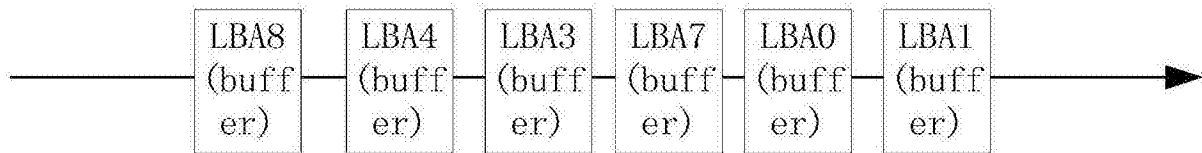


图3

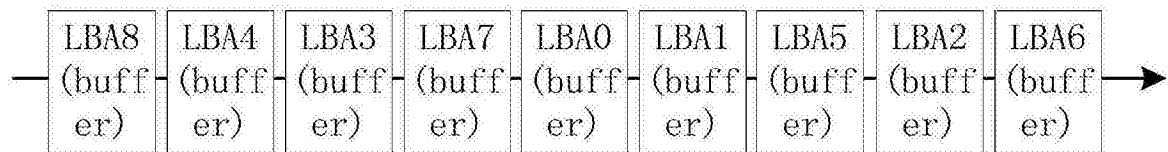


图4

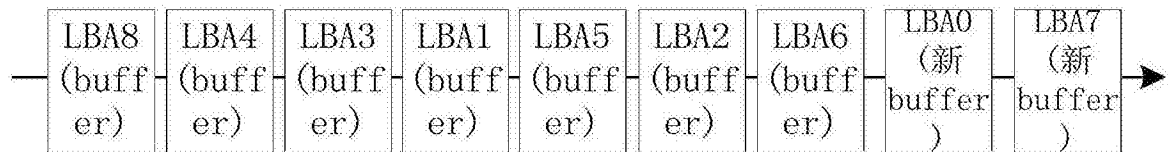


图5

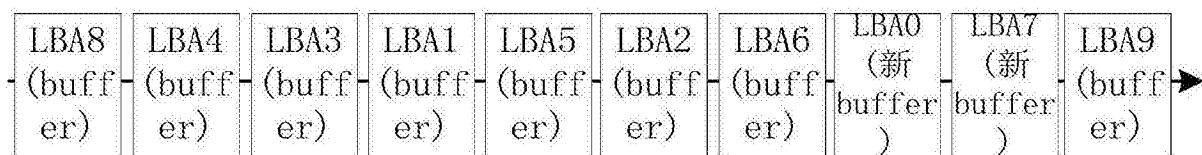


图6

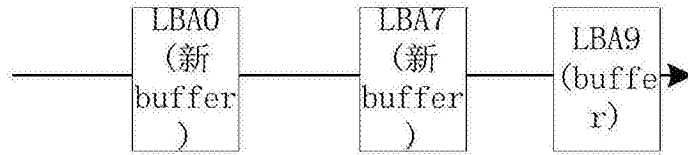


图7

闪存块

LBA8	一档
LBA4	
LBA3	
LBA1	二档
LBA5	
LBA2	
LBA6	三档

图8

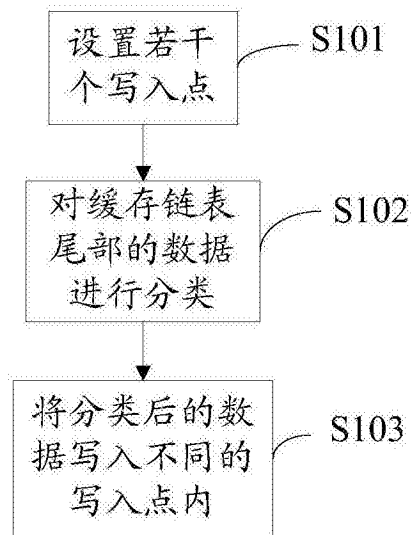


图9

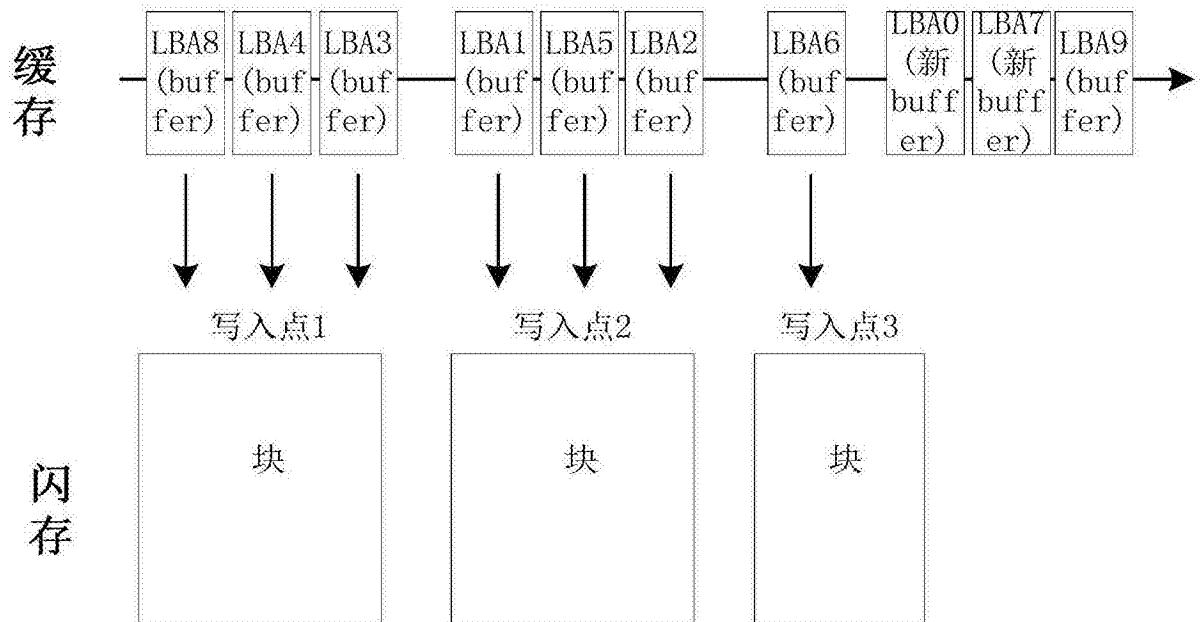


图10

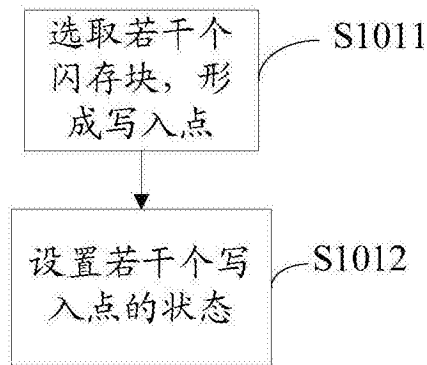


图11

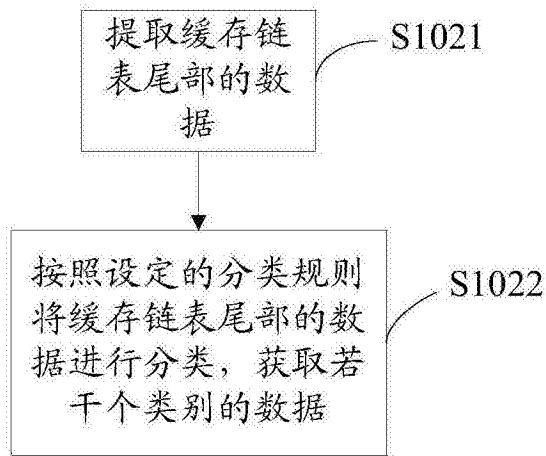


图12

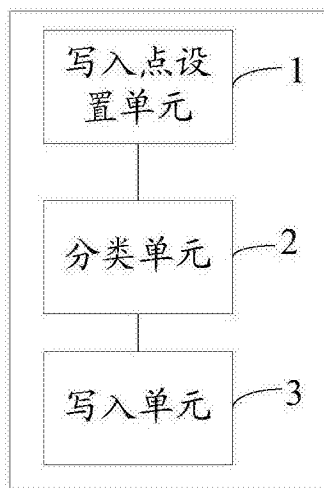


图13

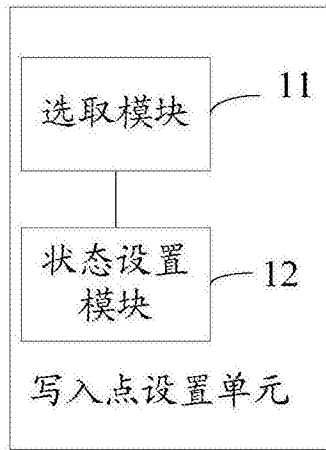


图14

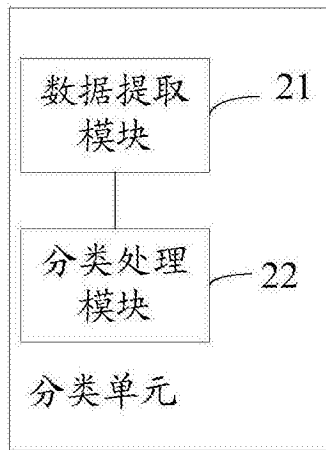


图15

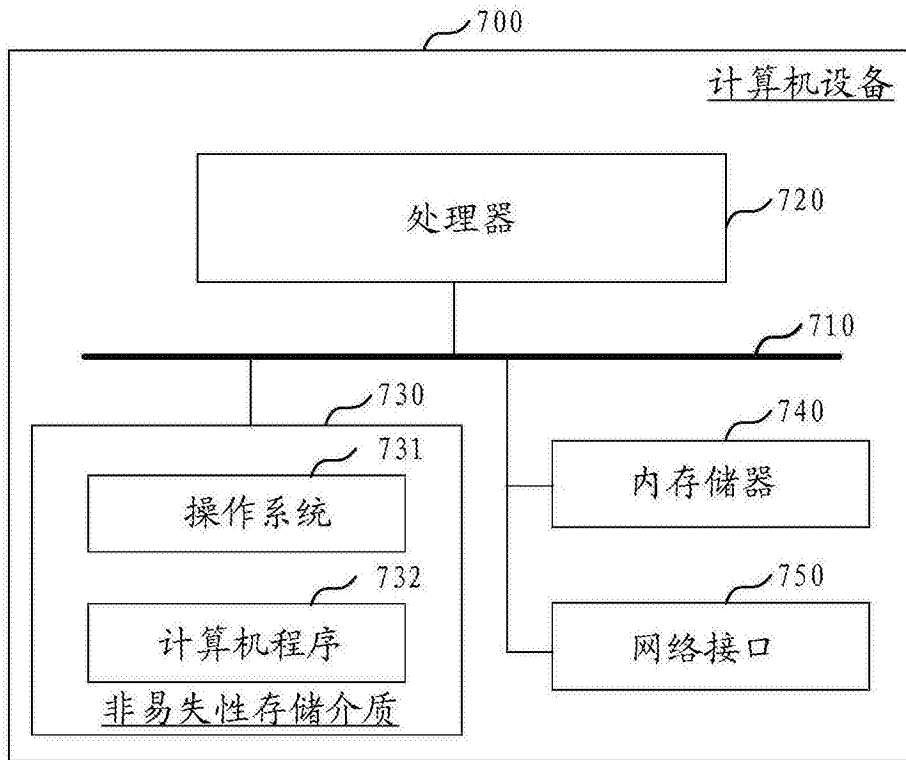


图16