



(12) 发明专利

(10) 授权公告号 CN 108628552 B

(45) 授权公告日 2021.02.09

(21) 申请号 201810441749.8

(22) 申请日 2018.05.10

(65) 同一申请的已公布的文献号
申请公布号 CN 108628552 A

(43) 申请公布日 2018.10.09

(73) 专利权人 南京道熵信息技术有限公司
地址 211100 江苏省南京市江宁区秣周东路9号

(72) 发明人 胡晓宇

(74) 专利代理机构 南京苏高专利商标事务所
(普通合伙) 32204

代理人 孟红梅

(51) Int. Cl.
G06F 3/06 (2006.01)

(56) 对比文件

CN 106201340 A, 2016.12.07

CN 103092770 A, 2013.05.08

US 2012254513 A1, 2012.10.04

审查员 薛杰

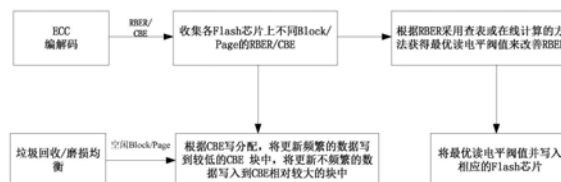
权利要求书2页 说明书8页 附图3页

(54) 发明名称

一种改善Flash磨损寿命的方法、控制装置及存储系统

(57) 摘要

本发明公开了一种改善Flash磨损寿命的方法、控制装置及存储系统,该方法包括:收集各个Flash芯片上不同Block/Page的原始误码率RBER和纠正误码数CBE;当RBER增加到一定程度时,计算得到相应Block/Page的最优读电平阈值并写入相应的Flash芯片;并根据CBE将空闲的Block/Page进行分类,将CBE较低的Block/Page分配给更新频率高的数据,将CBE较高的Block/Page分配给更新频率低的数据。本发明一方面通过调整读电平阈值,使得RBER降低,另一方面通过匹配数据的更新特质与Flash的Endurance磨损能够延缓块的损坏速度,提高Flash的磨损寿命。本发明可应用于固态硬盘,磁盘阵列,固态缓存,分布式存储系统,大数据应用等诸多领域,具有广泛的应用前景。



1. 一种改善Flash 磨损寿命的方法,其特征在于,包括:

收集各个Flash芯片上不同Block/Page的原始误码率RBER,当RBER增加到一定程度时,计算得到相应Block/Page的最优读电平阈值并写入相应的Flash芯片;所述最优读电平阈值从预先获取的RBER与读电平阈值关系表中查询得到,或者在线从多个可能的读电平阈值根据似然函数选取出最优读电平阈值;

收集各个Flash芯片上不同Block/Page的纠正误码数CBE,根据CBE将空闲的Block/Page进行分类,将CBE较低的Block/Page分配给更新频率高的数据,将CBE较高的Block/Page分配给更新频率低的数据;其中数据更新频率的高低信息根据用户逻辑地址LBA冷热计数器判断,或者,根据数据写请求的来源区分;

采用查表的方式获得最优读电平阈值,RBER与读电平阈值关系表根据如下方法获取:

对一组Flash 测试芯片反复进行读写操作,记录RBER的值;

当RBER值变化到下一个区间时,停止写操作,只进行读操作,反复尝试不同的读电平阈值,找到最佳值;

继续读写操作,使得芯片继续磨损,在更高的RBER区间校准最佳读电平阈值;

循环以上过程,直到获得所有的RBER区间所对应的读电平阈值;

采用在线计算的方式获得最优读电平阈值,在线计算最优读电平阈值的具体方法包括:

保存一段时间内每个Block的历史读出电平以及对应的存储数据信息;

在需要进行读电平阈值优化时,依据当前读电平阈值推测出多个可能的电平漂移值从而得到对个新的可能的读电平阈值;

使用得到多个新的可能的读电平阈值,计算上一时段内的ML函数,选择其中对应于最大ML函数的读电平阈值为最优的读电平阈值;

采用根据LBA冷热计数器获取数据冷热信息的方式实现数据写分配,具体包括:

根据CBE信息将垃圾回收/磨损均衡模块产生的可再分配的Block/Page分为若干等级;

接受用户数据写请求,查询写数据的LBA冷热计数器,获得数据冷热程度;

为较热的数据分配CBE较低的Flash Block/Page,为较冷的数据分配CBE较高的Flash Block/Page;

采用区分数据写请求来源的方式实现数据写分配,具体包括:

根据CBE信息将垃圾回收/磨损均衡模块产生的可再分配的Block/Page分为高级别和低级别两类;并将CBE低级别和高级别Block/Page分别按FIFO队列排列;

接受数据写请求,为用户写请求从CBE低级别队列中依次选取Block/Page并执行写操作,为垃圾回收/磨损均衡模块数据迁移写请求从CBE高级别队列中依次选取Block/Page并执行写操作。

2. 一种用于实现根据权利要求1所述的改善Flash 磨损寿命的方法的改善Flash 磨损寿命的控制装置,其特征在于,包括:

RBER/CBE收集模块,用于获取并存储各个Flash芯片上不同Block/Page的原始误码率RBER和纠正误码数CBE;

读电平阈值计算模块,用于在Flash芯片的Block/Page的RBER增加到一定程度时,计算得到相应Block/Page的最优读电平阈值;所述最优读电平阈值从预先获取的RBER与读电

阈值关系表中查询得到,或者在线从多个可能的读电平阈值根据似然函数选取出最优读电平阈值;

读电平阈值调节模块,用于接收计算得到的最优读电平阈值并写入相应的Flash芯片;

以及,数据写分配模块,用于根据CBE将空闲的Block/Page进行分类,将CBR较低的Block/Page分配给更新频率高的数据,将CBR较高的Block/Page分配给更新频率低的数据;其中数据更新频率的高低信息根据用户逻辑地址LBA冷热计数器判断,或者,根据数据写请求的来源区分。

3.一种存储系统,包括Flash芯片、ECC编解码模块、地址转换模块、垃圾回收模块、磨损均衡模块,其特征在于,还包括根据权利要求2所述的改善Flash 磨损寿命的控制装置。

4.一种存储系统,包括Flash控制器,所述控制器包括存储器、处理器及存储在存储器上并可在处理器上运行的程序,其特征在于,所述程序被加载至处理器时实现根据权利要求1所述的改善Flash 磨损寿命的方法。

一种改善Flash磨损寿命的方法、控制装置及存储系统

技术领域

[0001] 本发明属于存储技术领域,涉及一种提高NAND Flash磨损寿命的方法、控制装置及存储系统,尤其涉及的是利用原始误码率和纠错后误码率来自动适应Flash磨损带来的不良影响,通过调节读电平阈值(反馈策略)和匹配数据的更新特质与Flash的Endurance磨损(前馈策略),来实现延长Flash寿命的目的。

背景技术

[0002] NAND Flash(闪存)是一种非易失性随机访问存储介质,基于浮栅(Floating Gate)晶体管设计,通过浮栅来锁存电荷,电荷被储存在浮栅中,它们在无电源供应的情况下仍然可以保持。NAND Flash芯片在磁盘系统中已经被广泛的应用,比如固态硬盘(SSD),NVM-e固态硬盘,全闪存磁盘阵列等。常用的Flash芯片及技术有2D MLC NAND,2D TLC NAND,3D TLC NAND等。

[0003] NAND Flash芯片有多个Block(块)组成,Block是能够执行擦除操作的最小单元,通常由多个Page组成;Page是能够执行编程和读操作的最小单元,通常大小为4KB/8KB/16KB/32KB等。NAND Flash写入前必须擦除,Block擦除1次后再写入1次称为1次PE Cycle,Endurance(耐久性)用于衡量NAND Flash的擦写寿命的可靠性指标;Endurance指的是在一定的测试条件下NAND Flash能够反复擦写数据的能力,即对应NAND Flash的PE(Program/Erase)Cycle。

[0004] NAND Flash的原始误码率RBER通常较高,需要采用ECC(Error Correcting Code)纠错码进行纠错处理。常用的纠错码有BCH,RS,以及在Flash控制器使用最为广泛的LDPC码。其中,低密度奇偶校验码(LDPC)被认为是可以达到逼近香农极限(Shannon Limit)的帧差错率(Frame Error Rate,FER)的一类编码。对于这一类编码,要应用于NAND闪存,要求它必须不仅可以出色地解码而且适合在大规模集成电路(VLSI)上实现。准循环LDPC码(QC-LDPC)就可以满足这两点要求。QC-LDPC的H矩阵含有循环矩阵,不仅具有很好的纠错性能,循环矩阵的结构也利于VLSI实现。

[0005] RBER(Raw Bit Error Rate)指没有经过ECC纠错时出现一个Bit位发生错误的几率,RBER是衡量NAND品质的一项重要指标。随着PE次数的增加会变差,出错趋势呈指数分布,其中一个主要原因是擦写造成了浮栅氧化层的磨损,另外一个主要原因是随着PE增加,读数据的门限值会与起初设定的门限值(Thr1、Thr2、Thr3等等)发生漂移,从而导致RBER大幅上升,最终超过ECC纠错码的纠错能力。如图1显示,在t1时刻,Flash各个存储单位的状态和对应的读门限的位置,这个时刻对应于初始状态,各个状态在编程(写入)后的电压值与门限值之间保持了较好的距离,故具有较低的RBER。t2时刻表明经过一段时间的读写后,该存储单元经受了一定的磨损,每个编程状态变得更宽,彼此直接比较接近,容易导致误码。更严重的是,状态之间的门限值发生了漂移。如果不动态的调整这些阈值,很显然会导致较高的RBER,降低了Flash的使用寿命。这些门限阈值的漂移不仅随着Flash Blocks的PE值变化而变化,而且还和Blocks的品质,甚至在Flash芯片所处的相对位置,以及相邻块

的读写频率有关,很难通过预先设定的方式来适应整个动态变化,因此需要一个自适应反馈机制来应对。

[0006] Correctable Bit Error (CBE) 指ECC纠错码发现并纠正的错误bit数。当某个Page或Block上的错误Bit达到一定程度是,ECC将无法纠错,该Page或Block就不能被使用,被系统标注为块损坏。通常一个固态硬盘预留固定数目的冗余Block空间,当损坏的块超过冗余Block的数目时,该固态硬盘就不能正常工作。因此CBE是一个重要的经过ECC之后的Flash Blocks/Pages的品质参数,如何利用CBE来尽量延迟块损坏Block数目到达指定上限,是一种重要的延长固态硬盘寿命的技术。

[0007] NAND Flash芯片通常需要有一个或多个NAND控制器来完成读写控制。如图2,NAND控制器通常包含CPU, RAM/ROM, Firmware来执行控制读写过程和逻辑,以及与Flash芯片群相连的总线。从功能划分,NAND控制器里面的Firmware包含闪存逻辑/物理地址变换层(FTL),磨损均衡(Wear-leveling),和垃圾回收(Garbage Collection)等几个主要功能模块。

[0008] GC (Garbage Collection) 垃圾回收的产生主要由于Nand Flash的Block具有”Erase-Before-Write”特性,对于文件系统的Block更新实际上采用异地更新的做法,也就是说文件系统的Block第一次写在Nand物理Block 0的位置,下一次更新就有可能写在Nand物理Block 1的位置,原来写入的数据就变成脏数据。因此GC的主要目的是Nand Flash空间使用到一定程度之后进行脏数据空间回收的过程,同时要将有效数据搬移到新的位置。因此我们期望GC的候选Block能够全部是脏数据,这样回收起来只要Erase即可,再妥协一下就是包含极少量的有效数据,这样就减少了候选Block和有效数据的选择计算及有效数据搬移而带来的性能下降。

[0009] WL (Wear-Leveling) 磨损均衡的目的主要是防止Nand某些物理Block被频繁擦写而导致数据保持力很差,进而引发大量Bitflips,甚至是ECC Error或Bad Block的产生,数据因此出现错误而不可以再使用,这将是灾难性的。WL的出现正好解决这类问题,它能使所有的Block擦写处于同一水平,寿命均衡发展,用大家好才是真的好的理念保证数据安全性。WL有两种类型:动态WL和静态WL。简单来说动态WL每次挑选擦写次数最少的Block使用,而静态WL是把长期没有修改的Cold Data从擦写次数较少的Block中搬移出来放到擦写次数较多的Block中去,那么擦写次数较少的Block将被重新使用。一般动态WL发生在Write Request时候,而静态WL发生在空闲阶段定期检查触发条件执行,且具有全局Block的WL效果。

[0010] 目前的现状是这些功能模块都只利用了物理块(Flash Blocks)的PE信息,以及用户数据逻辑地址(LBA)的冷热信息(写频率高低),而忽略了每个物理Block/Page的CBE信息。

发明内容

[0011] 发明目的:针对现有技术的不足,本发明目的在于提供一种改善Flash磨损寿命的方法、控制装置及存储系统,利用原始误码率RBER和纠正误码数CBE来纠正Flash磨损带来的不良影响,通过调节读电平阈值和匹配数据的更新频率与Flash的Endurance磨损,来实现延长Flash寿命的目的。

[0012] 技术方案:为实现上述发明目的,本发明采用如下技术方案:

[0013] 一种改善Flash磨损寿命的方法,主要包括如下两个方面:

[0014] 收集各个Flash芯片上不同Block/Page的原始误码率RBER,当RBER增加到一定程度时,计算得到相应Block/Page的最优读电平阈值并写入相应的Flash芯片;所述最优读电平阈值从预先获取的RBER与读电平阈值关系表中查询得到,或者在线从多个可能的读电平阈值根据似然函数选取出最优读电平阈值;

[0015] 收集各个Flash芯片上不同Block/Page的纠正误码数CBE,根据CBE将空闲的Block/Page进行分类,将CBR较低的Block/Page分配给更新频率高的数据,将CBR较高的Block/Page分配给更新频率低的数据;其中数据更新频率的高低信息根据用户逻辑地址空间LBA冷热计数器判断,或者,根据数据写请求的来源区分。

[0016] 作为一种优选的实施方案,采用查表的方式获得最优读电平阈值,RBER与读电平阈值关系表根据如下方法获取:

[0017] 对一组Flash测试芯片反复进行读写操作,记录RBER的值;

[0018] 当RBER值变化到下一个区间时,停止写操作,只进行读操作,反复尝试不同的读电平阈值,找到最佳值;

[0019] 继续读写操作,使得芯片继续磨损,在更高的RBER区间校准最佳读电平阈值;

[0020] 循环以上过程,直到获得所有的RBER区间所对应的读电平阈值。

[0021] 作为另一种优选的实施方案,采用在线计算的方式获得最优读电平阈值,在线计算最优读电平阈值的具体方法包括:

[0022] 保存一段时间内每个Block的历史读出电平以及对应的存储数据信息;

[0023] 在需要进行读电平阈值优化时,依据当前读电平阈值推测出多个可能的电平漂移值从而得到对个新的可能的读电平阈值;

[0024] 使用得到多个新的可能的读电平阈值,计算上一时段内的ML函数,选择其中对应于最大ML函数的读电平阈值为最优的读电平阈值。

[0025] 作为一种优选的实施方案,采用根据LBA冷热计数器获取数据冷热信息的方式实现数据写分配,具体包括:

[0026] 根据CBE信息将垃圾回收/磨损均衡模块产生的可再分配的Block/Page分为若干等级;

[0027] 接受用户数据写请求,查询写数据的LBA冷热计数器,获得数据冷热程度;

[0028] 为较热的数据分配CBE较低的Flash Block/Page,为较冷的数据分配CBE较高的Flash Block/Page。

[0029] 作为另一种优选的实施方案,采用区分数据写请求来源的方式实现数据写分配,具体包括:

[0030] 根据CBE信息将垃圾回收/磨损均衡模块产生的可再分配的Block/Page分为高级别和低级别两类;并将CBE低级别和高级别Block/Page分别按FIFO队列排列;

[0031] 接受数据写请求,为用户写请求从CBE低级别队列中依次选取Block/Page并执行写操作,为垃圾回收/磨损均衡模块数据迁移写请求从CBE高级别队列中依次选取Block/Page并执行写操作;

[0032] 本发明另一方面公开的一种改善Flash磨损寿命的控制装置,包括:

[0033] RBER/CBE收集模块,用于获取并存储各个Flash芯片上不同Block/Page的原始误码率RBER和纠正误码数CBE;

[0034] 读电平阈值计算模块,用于在Flash芯片的Block/Page的RBER增加到一定程度时,计算得到相应Block/Page的最优读电平阈值;所述最优读电平阈值从预先获取的RBER与读电平阈值关系表中查询得到,或者在线从多个可能的读电平阈值根据似然函数选取出最优读电平阈值;

[0035] 读电平阈值调节模块,用于接收计算得到的最优读电平阈值并写入相应的Flash芯片;

[0036] 以及,数据写分配模块,用于根据CBE将空闲的Block/Page进行分类,将CBR较低的Block/Page分配给更新频率高的数据,将CBR较高的Block/Page分配给更新频率低的数据;其中数据更新频率的高低信息根据用户逻辑地址空间LBA冷热计数器判断,或者,根据数据写请求的来源区分。

[0037] 本发明另一方面公开的一种存储系统,包括Flash芯片、ECC编解码模块,地址转换模块、垃圾回收模块、磨损均衡模块,以及所述的改善Flash磨损寿命的控制装置。

[0038] 本发明另一方面公开的一种存储系统,包括Flash控制器,所述控制器包括存储器、处理器及存储在存储器上并可在处理器上运行的程序,所述程序被加载至处理器时实现根据权利要求所述的改善Flash磨损寿命的方法。

[0039] 有益效果:本发明通过跟踪Block/Page随数据写过程中的RBER变化,当Block/Page的RBER出现较大幅度的变化,采用策略调整这些块对应的读电平阈值,从而使得RBER降低。同时,根据Block/Page的CBE值,将Block/Page进行等级分类。相似的CBE值表明Block/Page具有相近的Endurance寿命。在写数据的过程中,根据Block/Page的Endurance等级分配数据块,将更新频繁的数据写到较低的CBE块中,而将更新不频繁的数据写入到CBE相对较大的块中,从而延缓块的损坏速度,提高Flash的磨损寿命。本发明可应用于固态硬盘,磁盘阵列,固态缓存,分布式存储系统,大数据应用等诸多领域,具有广泛的应用前景。

附图说明

[0040] 图1为读电平阈值在Flash芯片写的过程中发生漂移,导致RBER增大的原理示意图。

[0041] 图2为现有的Flash控制器硬件模块图。

[0042] 图3为本发明实施例的方法流程示意图。

[0043] 图4为本发明实施例中RBER-读电平阈值的查询表格示意图。

[0044] 图5为本发明实施例中RBER-读电平阈值的查询表格校准流程图。

[0045] 图6为本发明实施例中根据RBER变化在线计算最优读电平阈值的流程图。

[0046] 图7为本发明实施例中依据LBA写计算器实现CBE写分配的流程图。

[0047] 图8为本发明实施例中基于区分用户写操作与GC/WC数据迁移实现CBE写分配的流程图。

具体实施方式

[0048] 下面结合附图和具体实施例,对本发明做进一步说明。

[0049] 如图3所示,本发明实施例公开的一种基于原始误码率RBER和纠正误码数CBE改善Flash磨损寿命的方法,通过调节读电平阈值和匹配数据的更新频率与Flash的Endurance磨损,来实现延长Flash寿命的目的。该方法一方面是:收集各个Flash芯片上不同Page/Block的原始误码率,将RBER与历史数据对比,当增加到一定程度时,启动调整该Page/Block读电平阈值的算法,从而降低RBER。由于读电平阈值有反过来影响RBER,形成一种闭环反馈机制,故称通过RBER调整读电平阈值是一种反馈策略。另一方面是:收集每个Block/Page的CBE,并根据CBE将空闲的Block/Page进行分类,根据将要写入数据的冷热(数据更新频率)程度进行匹配,将CBR较低的Block/Page分配给热数据(更新频率高的数据),而将CBR较高的Block/Page分配给冷数据(更新频率低的数据)。由于热数据更新的频率比较高,因此驻留时间短,CBR较低的Block/Page就更容易被重新释放,磨损自然就相对增加,而CBR较高的Block/Page则受的磨损相对减少,起到了自然的磨损均衡的目的。尽管常用的Wear Level方法也具有磨损均衡的效果,但需要通过数据迁移实现,需要付出额外的性能和磨损的代价。而基于CBR的数据写分配策略可以在不引入额外代价的前提下实现磨损均衡,可有效延长Flash磨损寿命。由于基于CBE的数据写分配策略可以改善Flash磨损,但并不直接改变CBE,因此是一种前馈策略。

[0050] 本发明方法中,读电平阈值可以根据每个Page,每个Block,或多个Block组成的Zone进行不同颗粒度的计算,颗粒度越细,读电平阈值的精度就越高,但实现的复杂度就越高,在实际应用中可根据需要选择一个合适的折衷。本例中以Block为例。本发明提出两种具体的算法来实现计算适应RBER变化的读电平阈值。一种是利用表格查询的方式,即预先将Flash Block在生命周期内RBER变化的区间进行划分,N个区间,每个区间对应一组原先设定的读电平阈值,来适应RBER的变化。RBER-读电平阈值的查询表格如图4,表格的左栏表示的是RBER不同区间,右栏则显示其对应的最优读电平阈值。为了准确设定读电平阈值,需要事先做大量的校准工作,即对一定数额的Flash芯片反复进行读写操作,连续性地记录下RBER的值。当RBER值变化到另外一个区间时,停止写操作,只进行读操作,但反复尝试不同的读电平阈值,找到最佳值。之后,继续读写操作,使得芯片继续磨损,在更高的RBER区间校准最佳读电平阈值。循环以上过程,直到获得所有的RBER区间所对应的读电平阈值。

[0051] 图5详细给出了上述优化读电平阈值的校准流程。在步骤501中,选取一组具有典型性的Flash芯片,芯片具体数目取决于实验设备,时间/人力投入,成本等多个方面的考虑,通常芯片越多,得到的优化值越准确;在步骤502中,对上述测试芯片反复进行写操作,主动产生芯片磨损,然后统计RBER;在步骤503中,检查此刻RBER是否到达下一个需要重新校准的区间;如果还没有,则返回502继续执行写操作(磨损芯片);如果是,则进入504执行优化读电平阈值操作;在步骤504中,尝试逐步变化当前的读电平阈值,比如每次变化一个delta,然后执行大量读操作以统计RBER,最后在步骤505中选择一个对应于最低RBER的读电平阈值。在步骤506中,将优化后得到的读电平阈值保持在查询表格,然后返回502,继续执行写操作,使得RBER进入下一个需要重新优化的区间。该方法具有实现简单的特点,但不能适应环境温度的变化,以及PE对RBER和读电平阈值的交叉影响。

[0052] 本发明提出第二种算法来实现计算在线自适应RBER变化的读电平阈值。首先需要

保存一段时间内每个Block的历史读出电平以及对应的存储数据信息;然后根据块的RBER的变化决定该Block是否需要重新设置新的读电平阈值。当某块的RBER增大到某个设定值时,启动计算最优读电平阈值的流程,依据当前读电平阈值推测出多个可能的电平漂移值从而得到对个新的可能的读电平阈值,再基于最大似然函数选出最优的读电平阈值。具体在线计算过程如图6所示。在步骤601中,针对每个Flash Block保存在过去一段时间, t_1, t_2, t_3, \dots ,直到当前时刻,该Block的读出电平值,以及这些电平对应的真实存储信息(经过ECC纠错),比如101101…。这些信息将被用来计算似然函数。在步骤602中,检查当前时刻该Block对应的RBER是否偏大,如果偏大,则表明此时需要进行对该Block进行读电平阈值优化流程。步骤603中给出了一种具体的方法来寻找最优阈值。依据该块的当前读电平阈值,推测出多个可能的电平漂移值,按 $\text{delta}, 2*\text{delta}, \dots, \text{直到} N*\text{delta}$ 的方式得到N个新的可能的读电平阈值,即将当前电平偏移1到N个 delta ,得到N个可能的读电平阈值。在步骤604中,分别针对上述N个可能读电平阈值,计算在时间段 t_1, t_2, t_3, \dots 的似然函数。似然函数是依照某种分布函数和观察值去计算该观察值的出现概率,本例中分布函数是一种正太分布函数,其方差固定,均值取决于真实存储信息(经过ECC纠错)所对于的电平阈值(随读电平阈值调整而变化)。保存下来的读出电平值是观察样本值,用于计算似然函数。本步骤中我们在上一段时间内,使用N个可能的读电平阈值,按正太分布概率计算出N个maximum likelihood (ML) 函数。在步骤605中,设定最优读电平阈值为上述N个可能性中取得最大似然函数值。上述方法实现较为复杂,需要消耗更多的计算和存储资源,但可以自动适应温度和PE值的变化。

[0053] 本发明方法中,CBE写分配的工作流程可参考如下过程实施。CBE写分配模块从GC/WL模块接受可自由分配的Flash Block/Page,首先根据Block/Page的CBE进行分类,按CBE从低到高分为M个级别,CBE低级别表明该Block/Page的磨损程度低,相反,CBE高级别表明该Block/Page的磨损程度高。将用户逻辑地址空间LBA按预定块大小进行划分,并对每一个LBA对应的用户数据块分配一个冷热计数器,来记录在该块上的数据更新次数(写次数)。在单位时间内更新次数高的LBA就代表热数据,更新次数低的数据代表冷数据。所以根据LBA对应的更新次数划分数据冷热等级。在需要写入数据时,CBE写分配首先查询该数据LBA的数据更新次数,根据这个数字,判断该数据的冷热程度。根据冷热程度,分配相匹配的CBE等级对应的Flash Block/Page,将较热的数据写入CBE较低的Block/Page,而将较冷的数据写入CBE较高的Block/Page,从而达到自适应磨损均衡的目的。

[0054] 图7显示了上述利用LBA冷热计数器来获取数据冷热信息,并与CBE进行匹配的流程图。在步骤701中,垃圾回收(GC)/磨损均衡(WL)模块将产生可再分配的Flash Block/Page,在步骤702中,这些回收的Flash Block/Page将根据其CBE信息分为若干等级。CBE较低的等级表明磨损较轻微,可用于写入更新频繁的数据;更新频繁的数据在Flash/Page上驻留时间较短,所以这些Block/Page将更快地被再次回收,再次写入数据。相反,CBE较高的等级表明磨损相对较中,可用于写入更新频率低的数据,因这些数据距离下次被重写的的时间较长,故对CBE较高的Flash/Page起到的间接的保护作用。步骤703接受数据写请求,在步骤704中查询写数据的LBA冷热计数器,获得数据冷热程度。在步骤705中执行分配工作,为较热的数据分配CBE较低的Flash Block/Page,而为较冷的数据分配CBE较高的Flash Block/Page。步骤706完成写操作。注意步骤701和步骤703在逻辑上不存在先后次序,可以同时进行。

[0055] 采用LBA计数器方式虽然可以比较可靠的得到用户数据块的冷热程度,但实现需要消耗一定的存储和计算资源。本发明实施例提供了另一种更简单的实现方法是,将Block/Page按CBE分成两类,一类是CBE低级别,一类是CBE高级别。需要写入的数据按照其来源可分为两类,一类是来自用户写请求的数据,另一类是来自Flash控制器GC/WL模块发出的内部数据迁移的请求。内部数据迁移表明该涉及的数据在经历了较长时间任然没有被更新,可以被认定为冷数据,相反,来自用户写请求的数据可以被认定为热数据。因此,CBE写分配的工作流程就简化为:将来自用户写请求的数据分配到CBE低级别Block/Page,而将GC/WL迁移数据写到CBE高级别Block/Page。

[0056] 图8显示上述将CBE简单分为两类,高级别(代表磨损高)和低级别(代表磨损低),并分别与用户写入数据和GC/WL迁移数据进行匹配的流程。这种方法不需要依赖LBA计数器,因此实现更简单方便。在步骤801中,垃圾回收(GC)/磨损均衡(WL)模块将产生可再分配的Flash Block/Page;在步骤802中,上述Block/Page将根据其对应的CBE分成两类,高级别和低级别,分布代表磨损相对严重和模式相对较轻。步骤803将CBE低级别Flash Block/Page按FIFO队列排列,步骤804将CBE高级别按FIFO队列排列;步骤805接受数据写请求,在步骤806中区分数据写来源,将来自用户的数据写请求在步骤807中排队,将来自GC/WL的数据迁移写请求在808步骤中排序。在步骤809,为用户写请求从CBE低级别队列中依次选取Block/Page并执行写操作;在步骤810,为GC/WL数据迁移写请求从CBE高级别队列中依次选取Block/Page并执行写操作。注意步骤801和步骤805在逻辑上不存在先后次序,可以同时进行。

[0057] 本发明另一实施例公开的基于原始误码率RBER和纠正误码数CBE改善Flash磨损寿命的控制装置,在原有Flash控制器的基础上增加如下功能模块:

[0058] RBER/CBE收集模块,用于从ECC编解码器获取并存储各个Flash芯片上不同Block/Page的原始误码率RBER和纠正误码数CBE;读电平阈值计算模块,用于在Flash芯片的Block/Page的RBER增加到一定程度时,计算得到相应Block/Page的最优读电平阈值;所述最优读电平阈值从预先获取的RBER与读电平阈值关系表中查询得到,或者在线从多个可能的读电平阈值根据似然函数选取出最优读电平阈值;读电平阈值调节模块,用于接收计算得到的最优读电平阈值并写入相应的Flash芯片;以及,数据写分配模块,用于根据CBE将空闲的Block/Page进行分类,将CBR较低的Block/Page分配给更新频率高的数据,将CBR较高的Block/Page分配给更新频率低的数据;其中数据更新频率的高低信息根据用户逻辑地址LBA冷热计数器判断,或者,根据数据写请求的来源区分。

[0059] 在具体实施时,上述读电平阈值计算模块和数据写分配模块是在Flash控制器在传统的FTL、GC和WL三大软件功能模块基础上新增的两个软件功能模块,RBER/CBE收集模块和读电平阈值调节模块的功能可以固化到独立的芯片或硬件电路,通过RBER/CBE收集器和读电平阈值调节器两个硬件功能模块实现。RBER/CBE收集器负责从ECC编解码器收集并保存RBER和CBE信息,供CPU使用。CPU从ECC编解码器获取RBER信息,计算出新的读电平阈值,发送给读电平阈值调节器,读电平阈值调节器通过固定的芯片IO引脚将新设的读电平阈值写入到Flash芯片。

[0060] 本发明另一实施例提供了一种存储系统,包括Flash芯片、ECC编解码模块、地址转换模块、垃圾回收模块、磨损均衡模块,以及所述的改善Flash磨损寿命的控制装置。

[0061] 本领域技术人员可以理解,可以把上述实施例中的模块组合成一个模块或把它们分成多个子模块,可以以硬件实现,或者以在处理器上运行的软件模块实现,或者以它们的组合实现。

[0062] 基于相同的发明构思,本发明另一实施例提供了一种存储系统,包括Flash控制器,所述控制器包括存储器、处理器及存储在存储器上并可在处理器上运行的程序,所述程序被加载至处理器时实现所述的改善Flash磨损寿命的方法。

[0063] 综上,本发明提供了一种利用ECC纠错后得到的RBER和CBE信息,对Flash芯片的读电平阈值进行反馈控制,以及利用CBE来匹配冷热数据来有效延迟Flash磨损寿命的思路,具体实现该技术方案的方法和途径很多,以上所述仅是本发明的优选实施方式,应当指出,对于本技术领域的普通技术人员来说,在不脱离本发明原理的前提下,还可以做出若干改进和润饰,这些改进和润饰也应视为本发明的保护范围。本实施例中未明确的各组成部份均可用现有技术加以实现。

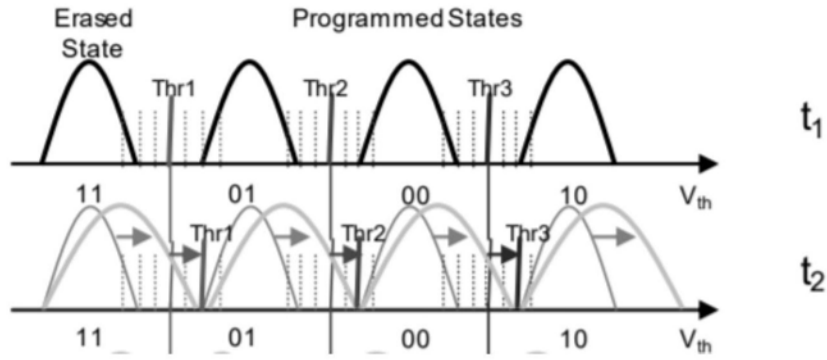


图1

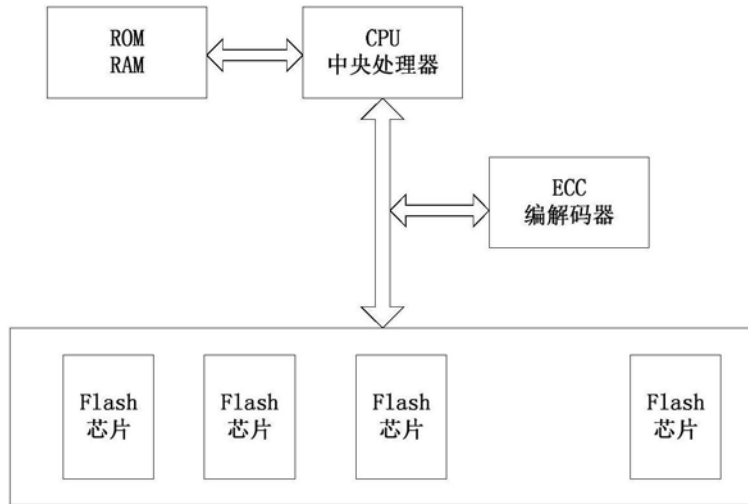


图2

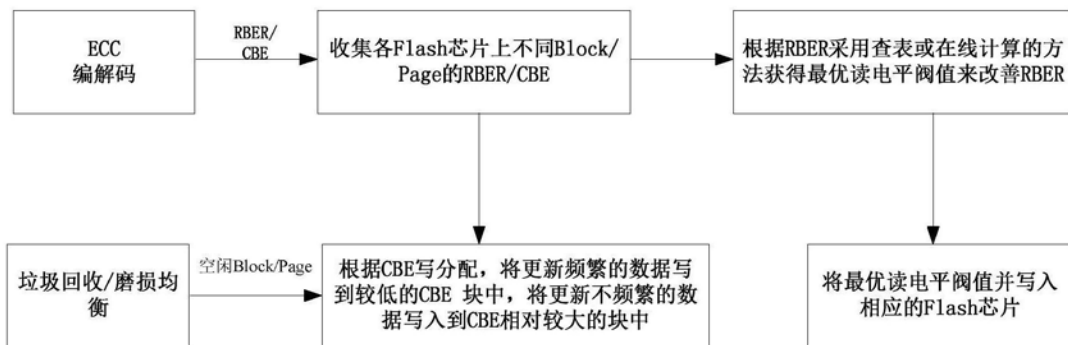


图3

RBER区间1	读电平预设阈值1
RBER区间2	读电平预设阈值2
RBER区间3	读电平预设阈值3
...	
RBER区间N	读电平预设阈值N

图4

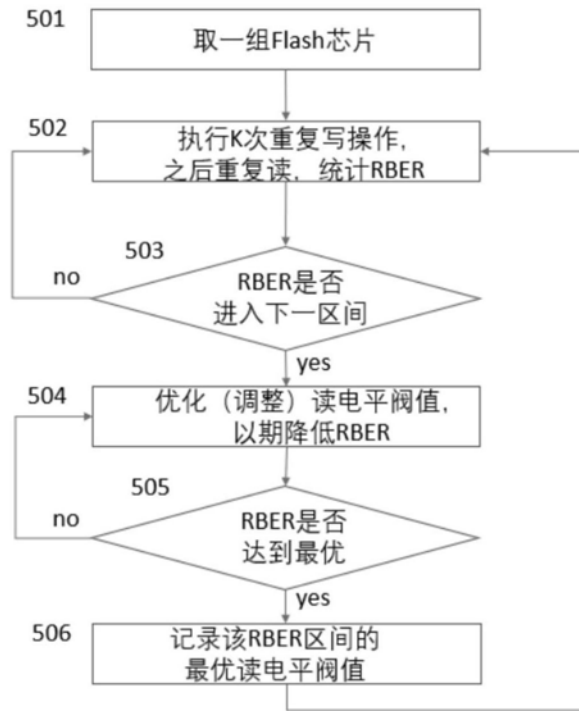


图5

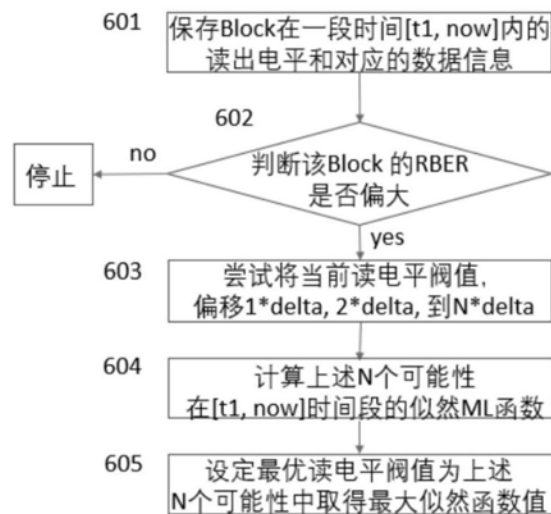


图6

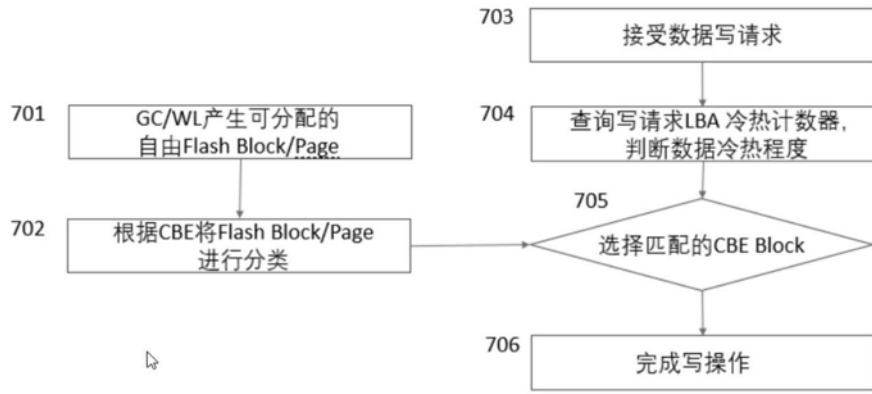


图7

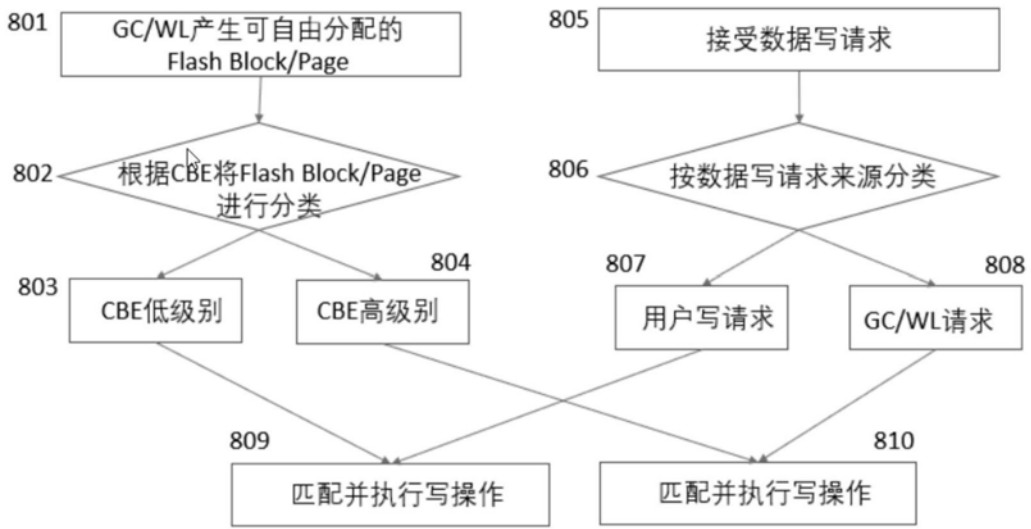


图8