



(12) 发明专利

(10) 授权公告号 CN 111090595 B

(45) 授权公告日 2022.12.20

(21) 申请号 201911134392.X
 (22) 申请日 2019.11.19
 (65) 同一申请的已公布的文献号
 申请公布号 CN 111090595 A
 (43) 申请公布日 2020.05.01
 (73) 专利权人 中国航空工业集团公司西安航空
 计算技术研究所
 地址 710065 陕西省西安市雁塔区高新区
 锦业二路15号
 (72) 发明人 贺莹 张锐 王闯 刘婷婷 邓豹
 (74) 专利代理机构 北京清大紫荆知识产权代理
 有限公司 11718
 专利代理师 娄华
 (51) Int. Cl.
 G06F 12/02 (2006.01)
 (56) 对比文件
 US 2017017405 A1, 2017.01.19
 US 2012317345 A1, 2012.12.13
 CN 106339324 A, 2017.01.18
 WO 2007019198 A2, 2007.02.15
 CN 109739776 A, 2019.05.10
 US 2016188219 A1, 2016.06.30

CN 103218306 A, 2013.07.24
 CN 109710541 A, 2019.05.03
 CN 110347612 A, 2019.10.18
 CN 107301132 A, 2017.10.27
 US 2011029715 A1, 2011.02.03
 CN 106951187 A, 2017.07.14
 CN 101339808 A, 2009.01.07
 TW 644207 B, 2018.12.11
 CN 106528000 A, 2017.03.22
 CN 108628758 A, 2018.10.09
 覃仁谅等. 基于逻辑区间冷热分离的NAND闪存垃圾回收算法.《微电子学与计算机》.2019, (第05期),
 Ayumi Soga 等. NAND flash aware data management system for high-speed SSDs by garbage collection overhead suppression.《2014 IEEE 6th International Memory Workshop (IMW)》.2014, 第1-4页.
 潘沁等. 磨损均衡算法在NAND Flash管理中的改进.《微计算机信息》.2007, (第07期),
 拓晶等. 基于多阈值的闪存磨损均衡算法.《微电子学与计算机》.2018, (第01期), (续)

审查员 吴海旋

权利要求书1页 说明书3页 附图3页

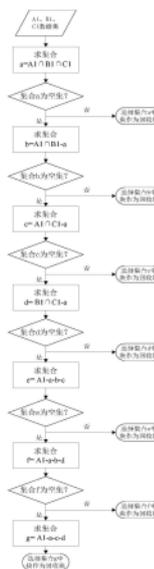
(54) 发明名称

一种NAND FLASH垃圾回收均衡优化方法

(57) 摘要

本发明提出一种NAND FLASH垃圾回收均衡优化方法, 综合考虑回收块有效数据量、写入频率、磨损度等垃圾回收相关因素, 并分别按照这些因素对NAND FLASH物理块进行聚类, 在聚类结果中选取低编程次数类集合、高写入频率类集合、低磨损度类集合作为初始筛选集, 求这三个初始筛选集的交集, 得到最优回收块集合。兼顾回收效率和磨损均衡, 均衡考虑垃圾回收过程相关因素, 采用集合求交思想逐步缩小可选回收块范围, 节省系统空间开销。

CN 111090595 B



[接上页]

(56) 对比文件

方才华 等. 全程优化的固态硬盘垃圾回收方法.《计算机应用》.2017,第37卷(第5期),第1257-1262页.

Yi Qin 等.DT-GC: Adaptive Garbage Collection with Dynamic Thresholds for SSDs.《2014 International Conference on Cloud Computing and Big Data》.2015,第182-187页.

1. 一种NAND FLASH垃圾回收均衡优化方法,其特征在于:综合考虑NAND FLASH物理块垃圾回收过程相关因素,并分别根据这些相关因素对NAND FLASH物理块采用聚类的方法进行分类,在聚类结果中选取相关目标集合作为初始筛选集,求解上述初始筛选集的交集,得到最优回收块集合;所述垃圾回收过程综合考虑回收块有效数据量X、写入频度Y和磨损度Z三个垃圾回收过程相关因素;垃圾回收过程包括以下步骤:

a) 建立多维模型 $F=f(X,Y,Z)$,其中X为有效数据量,Y为写入频度,Z为磨损度,f为关于X的正相关函数,即有效数据量越少引发的编程操作越少,f为Y的负相关函数,即选择写入频度越高的块引发的编程次数越少,f为Z的正相关函数,即选择磨损度低的块引发的磨损均衡度越高;

b) 优化目标为 $\max(F)$,即最大化回收效率与最大化磨损均衡度;

c) 采用聚类的方法分别根据有效数据量X、写入频度Y和磨损度Z对NAND FLASH物理块进行3个维度的聚类,根据有效数据量X将数据块分类为有效数据量少的块A1和有效数据量多的块A2,根据写入频度Y将数据块分类为高频写入块B1和低频写入块B2,根据磨损度Z将数据块分类为低磨损块C1和高磨损块C2,选取A1、B1和C1作为初始筛选集;

d) 求解A1、B1和C1三个初始筛选集的交集,得到最优回收块集合;求解A1、B1和C1三个初始筛选集的交集按如下步骤进行:

1) $a=A1 \cap B1 \cap C1$;

2) $b=A1 \cap B1 - a$;

3) $c=A1 \cap C1 - a$;

4) $d=B1 \cap C1 - a$;

5) $e=A1 - a - b - c$;

6) $f=A1 - a - b - d$;

7) $g=A1 - a - c - d$;

最终g即为最优回收块集合,以上步骤中“ \cap ”代表集合交集运算符,“-”代表集合差集运算符。

一种NAND FLASH垃圾回收均衡优化方法

技术领域

[0001] 本发明属于大容量存储技术领域,涉及一种NAND FLASH垃圾回收均衡优化方法。

背景技术

[0002] NAND FLASH凭着其非易失、读取速度快、存储容量大等诸多优点,近年来被广泛应用于大容量存储领域。由于NAND FLASH的固有属性,经过一段时间的编程后垃圾数据会越来越多,可供使用的空间越来越少,需要对垃圾数据占用的空间进行回收。垃圾回收过程应尽量减少回收操作,减少数据搬移的代价,减少系统开销,其中回收垃圾块的选择是最重要环节。目前存储器厂商通常使用的回收块选择方法包括静态策略和动态策略。如图1所示,静态策略是指根据回收时刻包含有效数据的多少,优先回收有效数据少的块,减少搬移数据量。动态策略是指根据数据更新的频率,优先回收更新频率低的有效数据所在的块,减少数据搬移频率。这两类方法或根据有效数据量,或根据数据更新频次选择需回收的垃圾块,没有兼顾考虑回收效率和磨损均衡。此外需根据记录的相关信息进行比较计算,比对算法复杂,系统开销大。

[0003] 专利“一种选择垃圾回收目标块的方法及固态硬盘”(201810209119.8)公开了一种选择垃圾回收目标块的方法及固态硬盘,其特征在于根据各个块block中无效页所占比例和擦写次数计算各个块的回收因数得分score,选择得分最少的block作为回收目标块。该方法中需记录每个块的无效页数量及擦写次数,占用较大的存储空间,同时需要计算最大擦写次数和最小擦写次数,系统开销大。

发明内容

[0004] 本发明的目的是:

[0005] 本发明是为了解决现有垃圾回收方法未兼顾回收效率和磨损均衡,且回收块选择算法复杂,系统开销大的问题。

[0006] 本发明的技术方案是:

[0007] 一种NAND FLASH垃圾回收均衡优化方法,综合考虑NAND FLASH物理块垃圾回收过程相关因素,并分别根据这些相关因素对NAND FLASH物理块采用聚类的方法进行分类,在聚类结果中选取相关目标集合作为初始筛选集,求解上述初始筛选集的交集,得到最优回收块集合。

[0008] 优选地,所述垃圾回收过程综合考虑回收块有效数据量X、写入频度Y和磨损度Z三个垃圾回收过程相关因素。

[0009] 优选地,垃圾回收优化过程包括以下步骤:

[0010] a) 建立多维模型 $F=f(X,Y,Z)$,其中X为有效数据量,Y为写入频度,Z为磨损度,f为关于X的正相关函数,即有效数据量越少引发的编程操作越少,f为Y的负相关函数,即选择写入频度越高的块引发的编程次数越少,f为Z的正相关函数,即选择磨损度低的块引发的磨损均衡度越高;

[0011] b) 优化目标为 $\max(F)$, 即最大化回收效率与最大化磨损均衡度;

[0012] c) 采用聚类的方法分别根据有效数据量 X 、写入频度 Y 和磨损度 Z 对NAND FLASH物理块进行3个维度的聚类, 根据有效数据量 X 将数据块分类为有效数据量少的块 $A1$ 和有效数据量多的块 $A2$, 根据写入频度 Y 将数据块分类为高频写入块 $B1$ 和低频写入块 $B2$, 根据磨损度 Z 将数据块分类为低磨损块 $C1$ 和高磨损块 $C2$, 选取 $A1$ 、 $B1$ 和 $C1$ 作为初始筛选集;

[0013] d) 求解 $A1$ 、 $B1$ 和 $C1$ 三个初始筛选集的交集, 得到最优回收块集合。

[0014] 优选地, 所述步骤d) 中, 求解 $A1$ 、 $B1$ 和 $C1$ 三个初始筛选集的交集按如下步骤进行:

[0015] 1) $a = A1 \cap B1 \cap C1$;

[0016] 2) $b = A1 \cap B1 - a$;

[0017] 3) $c = A1 \cap C1 - a$;

[0018] 4) $d = B1 \cap C1 - a$;

[0019] 5) $e = A1 - a - b - c$;

[0020] 6) $f = A1 - a - b - d$;

[0021] 7) $g = A1 - a - c - d$;

[0022] 最终 g 即为最优回收块集合, 以上步骤中“ \cap ”代表集合交集运算符, “ $-$ ”代表集合差集运算符。

[0023] 本发明具有的优点是: 本发明兼顾回收效率和磨损均衡, 均衡考虑NAND FLASH物理块的有效数据量、写入频度、磨损度等因素, 该方法将现有技术的复杂的数值问题转化为简单的离散聚类问题, 有效减少运算量, 并采用集合求交思想, 逐步缩小可选回收块范围, 节省系统空间开销。

附图说明

[0024] 附图1为垃圾回收静态和动态策略示意图。

[0025] 附图2为垃圾回收块均衡优化流程。

[0026] 附图3为初始筛选集交集求解示意图。

具体实施方式

[0027] 一种NAND FLASH垃圾回收均衡优化方法, 综合考虑NAND FLASH物理块垃圾回收过程相关因素, 并分别根据这些相关因素对NAND FLASH物理块采用聚类的方法进行分类, 在聚类结果中选取相关目标集合作为初始筛选集, 求解上述初始筛选集的交集, 得到最优回收块集合。

[0028] 本实施例以MICRON公司MT29F64G08AJABA型NAND FLASH的垃圾回收过程进行说明, 该芯片包含16384物理块, 每块包含128页。选择垃圾回收过程相关因素为有效数据量、写入频度和磨损度, 采用平均聚类分析方法对物理块进行分类, 回收块选择优化方法具体如下:

[0029] 1) 以物理块有效数据量的平均值为阈值进行分类, 将数据块分类为少量有效数据块集合 $A1$ 和大量有效数据量块集合 $A2$;

[0030] 2) 以写入次数平均值为阈值进行分类, 将数据块分类为高频写入块集合 $B1$ 和低频写入块集合 $B2$;

[0031] 3) 以擦除次数平均值为阈值进行分类,将数据块分类为低磨损块集合C1和高磨损块集合C2;

[0032] 4) 求集合 $a=A1 \cap B1 \cap C1$;

[0033] 5) 求集合 $b=A1 \cap B1 - a$;

[0034] 6) 求集合 $c=A1 \cap C1 - a$;

[0035] 7) 求集合 $d=B1 \cap C1 - a$;

[0036] 8) 求集合 $e=A1 - a - b - c$;

[0037] 9) 求集合 $f=A1 - a - b - d$;

[0038] 10) 求集合 $g=A1 - a - c - d$;

[0039] g即为最优回收块集合,以上步骤中“ \cap ”代表集合交集运算符,“-”代表集合差集运算符。

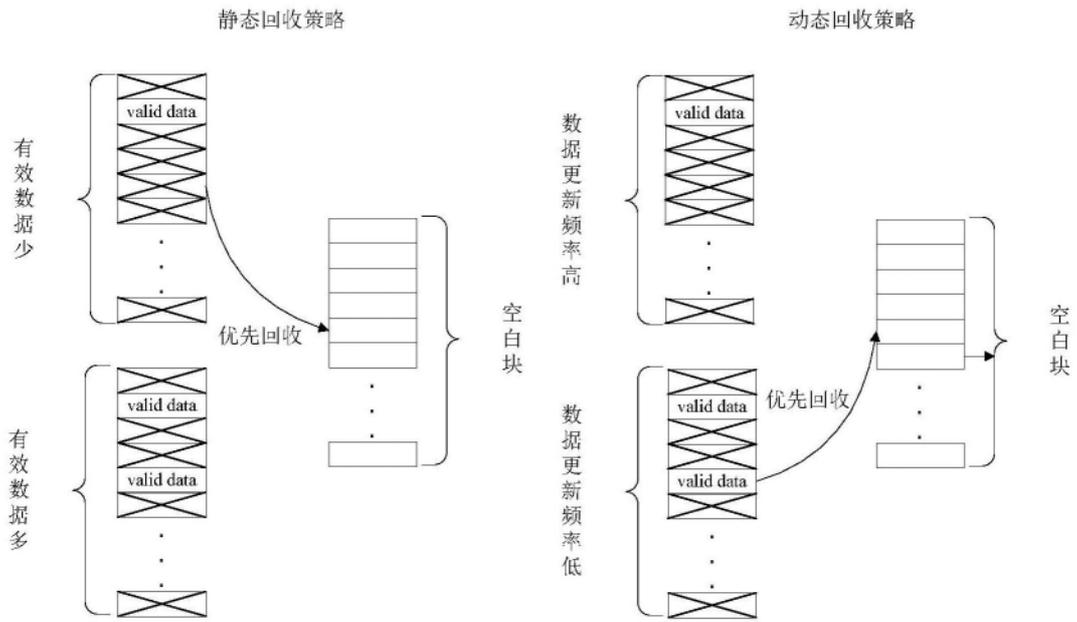


图1

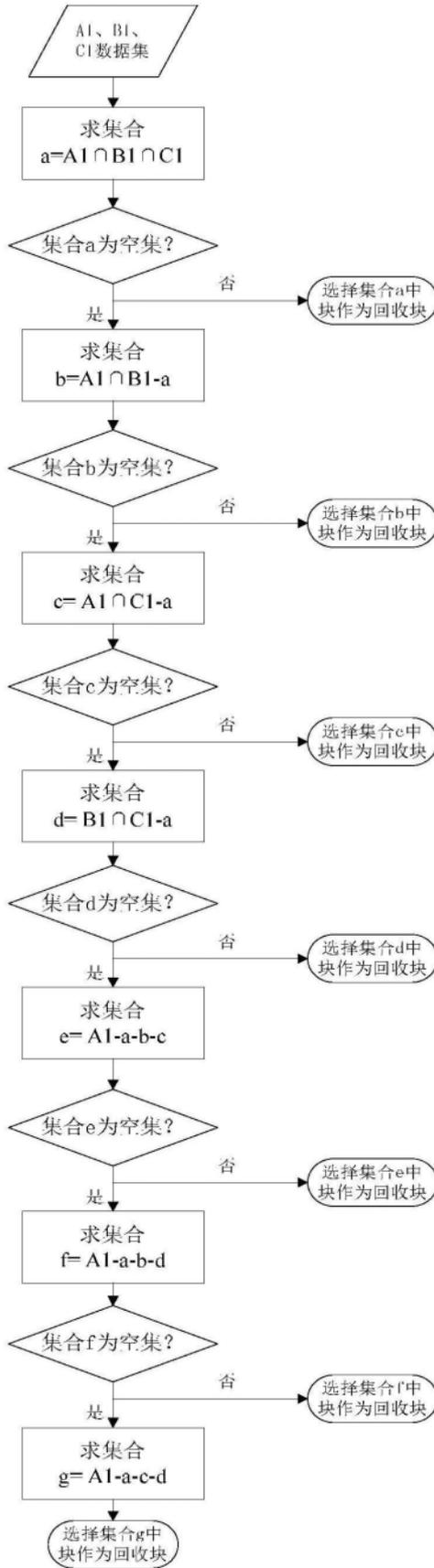


图2

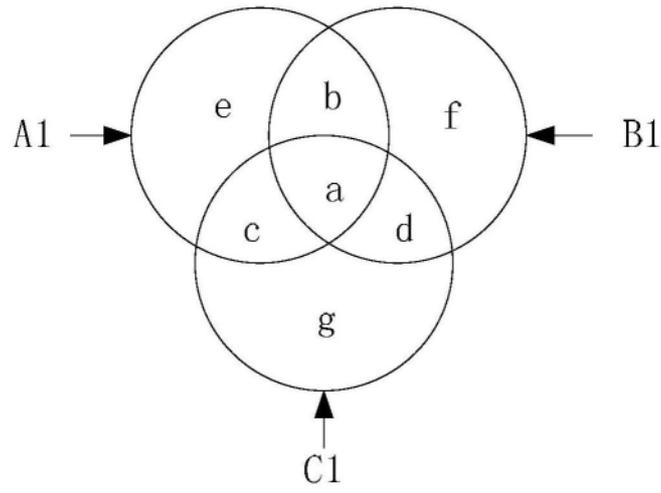


图3