



(12) 发明专利

(10) 授权公告号 CN 102197438 B

(45) 授权公告日 2014. 03. 05

(21) 申请号 200980142288. 3

(51) Int. Cl.

(22) 申请日 2009. 03. 31

G11C 29/00(2006. 01)

(30) 优先权数据

12/313, 205 2008. 11. 18 US

(56) 对比文件

US 5088081 A, 1992. 02. 11,

US 6427215 B2, 2002. 07. 30,

(85) PCT国际申请进入国家阶段日

2011. 04. 22

US 2007180239 A1, 2007. 08. 02,

CN 1679000 A, 2005. 10. 05,

(86) PCT国际申请的申请数据

PCT/US2009/001986 2009. 03. 31

审查员 张明霞

(87) PCT国际申请的公布数据

W02010/059173 EN 2010. 05. 27

(73) 专利权人 LSI 公司

地址 美国加利福尼亚州

(72) 发明人 布莱恩·麦克基恩

(74) 专利代理机构 北京纽乐康知识产权代理事

务所 11210

代理人 田磊

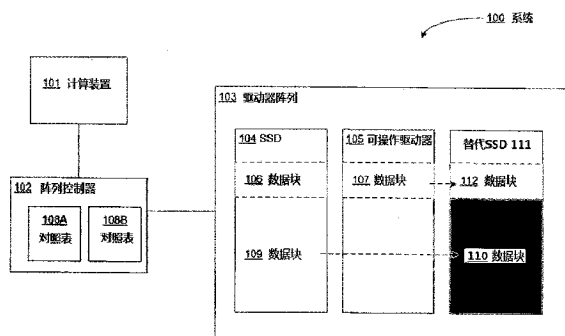
权利要求书2页 说明书4页 附图8页

(54) 发明名称

用于恢复固态驱动器数据的系统和方法

(57) 摘要

用于恢复固态驱动器 (SSD) 数据的方法可包括:检测包括一个或多个数据块的故障 SSD,接收将数据写到该故障 SSD 的该一个或多个数据块的请求,将该数据写到可操作驱动器的一个或多个数据块,和从该故障 SSD 和该可操作驱动器的该一个或多个数据块重建该故障 SSD。用于恢复固态驱动器 (SSD) 数据的系统可包括:用于检测包括一个或多个数据块的故障 SSD 的装置,用于接收将数据写到该故障 SSD 的该一个或多个数据块的请求的装置,用于将该数据写到可操作驱动器的一个或多个数据块的装置,和用于从该故障 SSD 和该可操作驱动器的该一个或多个数据块重建该故障 SSD 的装置。



1. 一种用于恢复固态驱动器(SSD)数据的方法,包括:
 - 检测包括一个或多个数据块的一故障 SSD;
 - 接收将数据写到该故障 SSD 的该一个或多个数据块的一请求;
 - 将该数据写到一可操作驱动器的一个或多个数据块;
 - 从该故障 SSD 和该可操作驱动器的该一个或多个数据块重建该故障 SSD,其中所述从该故障 SSD 和该可操作驱动器的该一个或多个数据块重建该故障 SSD 进一步包括:
 - 将未收到写请求的该故障 SSD 的一个或多个数据块拷贝到替代 SSD 中;
 - 将该可操作驱动器的该一个或多个数据拷贝到该替代 SSD 中。
2. 如权利要求 1 所述的方法,其中所述检测包括一个或多个数据块的一故障 SSD 进一步包括:
 - 检测指向该 SSD 的一个或多个数据块的一写故障。
3. 如权利要求 1 所述的方法,其中所述将该数据写到一可操作驱动器的一个或多个数据块进一步包括:
 - 根据一重定向写快照方法将该数据写到一可操作驱动器的一个或多个数据块。
4. 如权利要求 1 所述的方法,其中所述将该数据写到一可操作驱动器的一个或多个数据块进一步包括:
 - 更新一个或多个包括该故障 SSD 和一可操作 SSD 的数据块映射的查找表。
5. 如权利要求 1 所述的方法,其中所述将该可操作驱动器的该一个或多个数据拷贝到该替代 SSD 中进一步包括:
 - 将指向该可操作驱动器的该一个或多个数据块的读请求重定向到该替代 SSD 的一个或多个数据块。
6. 一种用于恢复固态驱动器(SSD)数据的系统,包括:
 - 用于检测包括一个或多个数据块的一故障 SSD 的装置;
 - 用于接收将数据写到该故障 SSD 的该一个或多个数据块的一请求的装置;
 - 用于将该数据写到一可操作驱动器的一个或多个数据块的装置;
 - 用于从该故障 SSD 和该可操作驱动器的该一个或多个数据块重建该故障 SSD 的装置,其中所述用于从该故障 SSD 和该可操作驱动器的该一个或多个数据块重建该故障 SSD 的装置进一步包括:
 - 用于将未收到写请求的故障 SSD 的一个或多个数据块拷贝到替代 SSD 中的装置;
 - 用于将该可操作驱动器的该一个或多个数据拷贝到该替代 SSD 中的装置。
7. 如权利要求 6 所述的系统,其中所述用于该检测包括一个或多个数据块的一故障 SSD 的装置进一步包括:
 - 用于检测指向该 SSD 的一个或多个数据块的写故障的装置。
8. 如权利要求 6 所述的系统,其中所述用于将该数据写到一可操作驱动器的一个或多个数据块的装置进一步包括:
 - 用于根据重定向写快照方法将该数据写到一可操作驱动器的一个或多个数据块的装置。
9. 如权利要求 6 所述的系统,其中所述用于将该数据写到一可操作驱动器的一个或多个

个数据块的装置进一步包括：

用于更新一个或多个包括该故障 SSD 和可操作 SSD 的数据块映射的查找表的装置。

10. 如权利要求 6 所述的系统,其中所述用于将该可操作驱动器的该一个或多个数据拷贝到该替代 SSD 中的装置进一步包括：

将指向该可操作驱动器的该一个或多个数据块的读请求重定向到该替代 SSD 的一个或多个数据块。

11. 一种用于恢复固态驱动器(SSD)数据的系统,包括：

用于检测包括一个或多个数据块的故障 SSD 的电路；

用于接收将数据写到该故障 SSD 的该一个或多个数据块的请求的电路；

用于将该数据写到可操作驱动器的一个或多个数据块的电路；

用于从该故障 SSD 和该可操作驱动器的该一个或多个数据块重建该故障 SSD 的电路,

其中所述用于从该故障 SSD 和该可操作驱动器的该一个或多个数据块重建该故障 SSD 的电路进一步包括：

用于将未收到写请求的故障 SSD 的一个或多个数据块拷贝到替代 SSD 中的电路；

用于将该可操作驱动器的该一个或多个数据拷贝到该替代 SSD 中的电路。

12. 如权利要求 11 所述的系统,其中所述用于检测包括一个或多个数据块的故障 SSD 的电路进一步包括：

用于检测指向该 SSD 的一个或多个数据块的写故障操作的电路。

13. 如权利要求 11 所述的系统,其中所述用于将该数据写到可操作驱动器的一个或多个数据块的电路进一步包括：

用于根据重定向写快照将数据写到可操作驱动器的一个或多个数据块的电路。

14. 如权利要求 11 所述的系统,其中所述用于将该数据写到可操作驱动器的一个或多个数据块的电路进一步包括：

用于更新一个或多个包括该故障 SSD 和可操作 SSD 的数据块映射的查找表的电路。

15. 如权利要求 11 所述的系统,其中所述用于将该可操作驱动器的该一个或多个数据拷贝到该替代 SSD 中的电路进一步包括：

将指向该可操作驱动器的该一个或多个数据块的读请求重定向到该替代 SSD 的一个或多个数据块。

用于恢复固态驱动器数据的系统和方法

背景技术

[0001] 当驱动器所基于的 NAND 技术具有有限数量或编程 / 擦除周期时, 固态驱动器 (SSDs) 会随着时间的推移而磨损。随着时间的推移, SSDs 会达到擦除周期无法将 NAND 闪存块重置为可写状态的点。这样的 SSD 会被认为发生故障。当前被设计用来处理传统的基于主轴的磁盘驱动器的磁盘阵列控制器可采用影响读和写二者的故障模式, 使该驱动器无用。

[0002] 但是, 一些 SSD 可采用只读故障模式, 其允许该驱动器成功执行读命令, 即使写命令导致故障。在当前磁盘阵列控制器对于写故障命令的通常响应是指示该驱动器发生故障时, 这样的控制器无法有效地利用 SSDs 的只读故障模式。

发明内容

[0003] 本发明涉及一种用于恢复固态驱动器 (SSD) 数据的系统和方法。

[0004] 一种用于恢复固态驱动器 (SSD) 数据的方法可包括: 检测包括一个或多个数据块的故障 SSD; 接收将数据写到该故障 SSD 的该一个或多个数据块的请求; 将该数据写到可操作驱动器的一个或多个数据块; 和从该故障 SSD 和该可操作驱动器的该一个或多个数据块重建该故障 SSD。

[0005] 一种用于恢复固态驱动器 (SSD) 数据的系统可包括: 用于检测包括一个或多个数据块的故障 SSD 的装置; 用于接收将数据写到该故障 SSD 的该一个或多个数据块的请求的装置; 用于将该数据写到可操作驱动器的一个或多个数据块的装置; 和用于从该故障 SSD 和该可操作驱动器的该一个或多个数据块重建该故障 SSD 的装置。

[0006] 应当可以理解, 不管是上述概括的描述还是后面详细的描述都仅仅是示例性和说明性的, 并不会是对所请求发明的必要限制。包括并组成说明书一部分的附图揭示了本发明的实施例, 并和说明书一起用来解释本发明的原理。

附图说明

[0007] 通过参考附图, 本发明的众多优点可以被本领域的技术人员更好地理解:

[0008] 图 1 示出了用于恢复固态驱动器 (SSD) 数据的高等级块的系统。

[0009] 图 2 示出了恢复固态驱动器 (SSD) 数据的过程。

[0010] 图 3 示出了恢复固态驱动器 (SSD) 数据的过程。

[0011] 图 4 示出了恢复固态驱动器 (SSD) 数据的过程。

[0012] 图 5 示出了恢复固态驱动器 (SSD) 数据的过程。

[0013] 图 6 示出了恢复固态驱动器 (SSD) 数据的过程。

[0014] 图 7A 示出了 SSD 配置。

[0015] 图 7B 示出了查找表配置。

[0016] 图 8A 示出了 SSD 配置。

[0017] 图 8B 示出了查找表配置。

具体实施方式

[0018] 在下面具体描述中,可以参考所附图,其组成了此处的一部分。在所述附图中,相同的符号通常识别为相同的部件,除非上下文另有规定。在具体实施方式、附图和权利要求中所述的说明性实施例不是限制性的。可采用其它实施例,以及进行不同的变化,而不脱离此处所公开主题的精神和范围。

[0019] 图 1 示出了在其中实施了一个或多个技术的示例系统。显示了包括计算装置 101、阵列控制器 102、和驱动器阵列 103 的数据存储系统 100。阵列控制器 102 可包括驱动器管理电路 / 软件,由此阵列控制器 102 能够处理访问驱动阵列 103 的各个驱动器的计算装置 101 的读 / 写请求。驱动器阵列 103 可包括一个或多个包括至少一个 SSD 104 的驱动器以及至少一个可操作驱动器 105。可操作驱动器 105 可以是 SSD 或硬盘驱动器 (HDD)。

[0020] 图 2 示出了展现与 SSD 恢复相关的示范操作的操作流程 200。在图 2 中和在后续包括操作流程的各个示例的图中,可参照图 1 所述示例,和 / 或参照其它示例和内容来提供讨论和说明。但是,应当理解,可在多个其它环境和内容中,和 / 或在图 1 的修改版本中来执行该操作流程。另外,尽管以所述顺序来展现该各个操作流程,应当理解,可按照非上述顺序的其它顺序来执行该各个操作,或可被同时执行。

[0021] 在开始操作后,该操作 210 示为检测包括一个或多个数据块的故障 SSD。例如,如图 1 所示,该阵列控制器 102 可检测出该驱动器阵列 103 不再能够处理对该驱动器阵列 103 的数据块 106 的写指令。

[0022] 操作 220 示为接收向该故障 SSD 的一个或多个数据块写数据的请求。例如,图 1 所示,该阵列控制器 102 可接收来自该计算装置 101 的将数据写给该 SSD 104 的数据块 106 的请求。

[0023] 操作 230 示为将数据写给可操作驱动器的一个或多个数据块。例如,如图 1 所示,该阵列控制器 102 可使得该驱动器阵列 103 将数据写给该可操作驱动器 105 的数据块 107。

[0024] 操作 240 示为从该故障 SSD 和该可操作驱动器的一个或多个数据块重建该故障 SSD。例如,如图 1 所示,该阵列控制器 102 可从该 SSD 104 和该可操作驱动器 105 的当前内容将 SSD 104 重建为替代 SSD 111。

[0025] 图 3 示出了图 2 的示范可操作流程 200 的可替换实施例。图 3 所示的示范实施例中检测操作 210 可包括至少一个附加操作。附加操作可包括操作 302。

[0026] 操作 302 示为检测指向该 SSD 的一个或多个数据块的故障的写操作。例如,如图 1 所示,该阵列控制器 102 可检测出该驱动器阵列 103 不再能够处理指向该 SSD104 的数据块 106 的写指令。

[0027] 图 4 示出了图 2 的示范可操作流程 200 的可替换实施例。图 4 所示的示范实施例中该可操作流程 200 可包括至少一个附加操作。附加操作可包括操作 402 和 / 或操作 404。

[0028] 操作 402 示为根据重定向写快照方法将该数据写给可操作驱动器的一个或多个数据块。例如,该阵列控制器 102 可使得该 SSD 104 将新写入数据放在与该同一数据的先前写入拷贝不同的位置。如图 7A 所示,多个数据段可被写入 SSD 104。该可操作驱动器 105 可以是已被擦除并准备被写入的 SSD。如图 7B 所示,查找表 108A 可保持该数据块到它们各自闪存块和页的映射。

[0029] 在参照上述相关操作 210 和 220 时,该阵列控制器 102 可接收该 SSD 104 的段 210 和 211 的写命令。如果该 SSD 104 的一个或多个数据块被检测出故障,数据存储系统 100 可采用重定向写 (redirect-on-write) 快照方法来保持该 SSD 104 中数据的完整性。

[0030] 如图 8A 所示,该阵列控制器 102 可接收段 210 和 211 的新数据并将其写入该可操作驱动器 105 的预先已擦除区域。该 SSD 104 可保留不变(图 8A 被修改以显示出 SSD 104 中该块 210 和 211 的数据是“旧”数据)。

[0031] 在该可操作驱动器 105 完成写操作后,该可操作驱动器 105 的状态如 8A 所示。段 210 和 211 的数据的旧拷贝可保留在该 SSD 104 中,直到整个 SSD 104 被要求收回并被擦除以准备再次使用。阵列控制器 102 中的查找表 108A 可被更新以反应段 210 和 211 的数据的新地址以及表现为如图 8B 所示。

[0032] 操作 404 示为更新包括该故障 SSD 和可操作 SSD 的数据块映射的一个或多个查找表。例如,如图 1 所示,阵列控制器 102 可包括一个或多个查找表(例如查找表 108A 和查找表 108B),该查找表保持存储器块组件和它们在该 SSD 104 和 / 或可操作驱动器 105 中各自闪存地址的映射表。

[0033] 当该 SSD 104 可为每一个接收到的写命令执行重定向写时,该 SSD 104 可保留用户数据的先前拷贝,该拷贝可被用来性能无损耗地重建当前数据或先前快照数据二者。通过增加能指出该 SSD 104 中数据的先前版本的附加查找表 108B,该数据存储系统 100 可提供对该装置中存储的多个时间点 (point-in-time) 拷贝的存取。

[0034] 例如,该阵列控制器 102 可接收存储先于段 201 和 211 写操作的数据的时间点拷贝的命令。如图 9A 所示,阵列控制器 102 可保留图 9B 中地址的查找表的拷贝,作为指向在查找表 108A 中该时间点拷贝的地址的指针。如图 9B 所示,当卷更新时,该数据的当前视图可被保持在第二查找表 108B 中。阵列控制器 102 可保留两个表的拷贝,只要其具有快照的空间以及该快照经过一些其它动作后未结束。

[0035] 图 5 示出了图 2 的示范可操作流程 200 的可替换实施例。图 5 所示的示范实施例中,该重建操作 240 可包括至少一个附加操作。附加操作可包括操作 502 和 / 或操作 504。

[0036] 操作 502 示为将未收到写请求的故障 SSD 的一个或多个数据块拷贝到替代 SSD 中。例如,如图 1 所示,该阵列控制器 102 可使得 SSD 104 的那些未通过写请求编址的部分(例如,未修改的数据块 109)拷贝到替代 SSD 111 的数据块 110 中。

[0037] 操作 504 示为将该可操作驱动器的一个或多个数据块拷贝到该替代 SSD 中。例如,如图 1 所示,该阵列控制器 102 可使得可操作驱动器 105 的那些响应于指向该 SSD 104 的写故障操作而被写入的部分(例如,数据块 107)拷贝到该替代 SSD 111 的数据块 112 中。

[0038] 图 6 示出了图 2 的示范可操作流程 200 的可替换实施例。图 6 所示的示范实施例中,该可操作流程 200 可包括至少一个附加操作。附加操作可包括操作 610。

[0039] 操作 610 示为将指向该可操作驱动器的一个或多个数据块的读请求重定向到该替代 SSD 的一个或多个数据块。例如,如图 1 所示,可以存在这种情况:该可操作驱动器 105 可包括 HDD。当 SSD 技术提供比 HDD 快的读性能时,在 HDD 可操作驱动器 105 的数据块拷贝期间,该阵列控制器 102 可保持指向已被写入该替代 SSD 111 中数据的长度的指针。当该阵列控制器 102 接收到已经被从该 HDD 可操作驱动器 105 拷贝到替代 SSD 111 的数据(例如,数据块 112)的读请求时,该阵列控制器 102 可将该读请求指向该替代 SSD 111。这样,

该替代 SSD 111 会能够服务被该 HDD 可操作驱动器 105 已服务过的 IO 流的一部分,其发生故障之前的 SSD 104 因此保持了增长的性能等级,与该 IO 被该 HDD 可操作驱动器 105 服务的情况相比较。

[0040] 应当相信,通过前述将理解本发明和许多其伴随的优点。同样应当相信,显而易见,其组成的形式、结构和布置可以进行不同的变化而不脱离本公开的范围和精神或不牺牲所有其物质优点。这里以前所述的形式只是示例说明,下列权利要求的目的是包含和包括此类变化。

[0041] 上述具体描述可包括通过使用方框图、流程和 / 或示例提出所述装置和 / 或过程的不同实施例。在这样方框图、流程和 / 或示例的范围内包含一个或多个功能和 / 或操作,应当理解,本领域技术人员可实施在这样方框图、流程和 / 或示例的内的每一个功能和 / 或操作,独立地和 / 或共同地,使用宽范围的硬件、软件、固件或实际上它们的任意组合。在实施例中,此处所述主题的几个部分可通过专用集成电路 (ASICs)、现场可编程门阵列 (FPGAs)、数字信号处理器 (DSPs) 或其它集成形式来进行实施。但是,本领域技术人员将认识到,此处公开的实施例的某些方面,整体地或部分地、可等效地以集成电路来实施,类似于在一个或多个计算机上运行的一个或多个计算机程序(例如,在一个或多个计算机系统上运行的一个或多个程序)、在一个或多个处理处理器上运行的一个或多个程序(例如,在一个或多个微处理器上运行的一个或多个程序)、固件或实际上它们的任意组合,以及上述设计电路和 / 或为软件和或固件写代码很好地处于根据本发明的本领域技术人员的技能内。

[0042] 另外,本领域技术人员将领会到,此处所述主题的结构会能够被分布为各种形式的程序产品,以及其被此处所述主题的示范性实施例所采用而不考虑用来实际执行所述分布的信号承载媒体的特定类型。信号承载媒体的示例包括,但可不限于:可记录类型的媒体,例如软盘、硬盘、光盘 (CD)、数字视频光盘 (DVD)、数字磁带、计算机存储器等;和传输类型的媒体,例如数字和 / 或模拟通信媒体(例如,光缆、波导、有线通信链路、无线通信链路(例如,发送器、接收器、发送逻辑、接收逻辑等)等)。

[0043] 本领域技术人员将认识到,所述技术的状态可包括被处理到在系统各个方面的硬件、软件和 / 或固件实施中保留较少区别处的位置;所述硬件、软件和 / 或固件的使用一般(但不总是,其中在特定语境中,在硬件和软件之间的选择会变得有意义)可以是一种展现成本效率平衡的设计选择。本领域技术人员将领会到,存在各种工具可被用来影响到这里所述的过程和 / 或系统和 / 或其它技术(例如,硬件、软件和 / 或固件),以及优选工具将随语境而变化,所述语境中所述过程和 / 或系统和 / 或其它技术会被展现。例如,如果实施者确定速度和精确度会是最重要的,实施者会选择大体上为硬件和 / 或固件的工具;或者,如果灵活性会是最重要的,实施者会选择大体上为软件的实施方案;或许,仍可选择地,实施者会选择硬件、软件和 / 或固件的一些组合。因此,存在几种可能的工具可被用来影响到这里所述的过程和 / 或系统和 / 或其它技术,没有一个会在本质上优于其它,其中任意将被使用工具的选择可以基于其中所述工具将被展现的语境以及实施者所关注的具体对象(例如,速度、灵活性或可预测性),其中的任意一个可发生变化。本领域技术人员将认识到,所述实现方案的视觉方面将一般地采用视觉主导的硬件、软件和或固件。

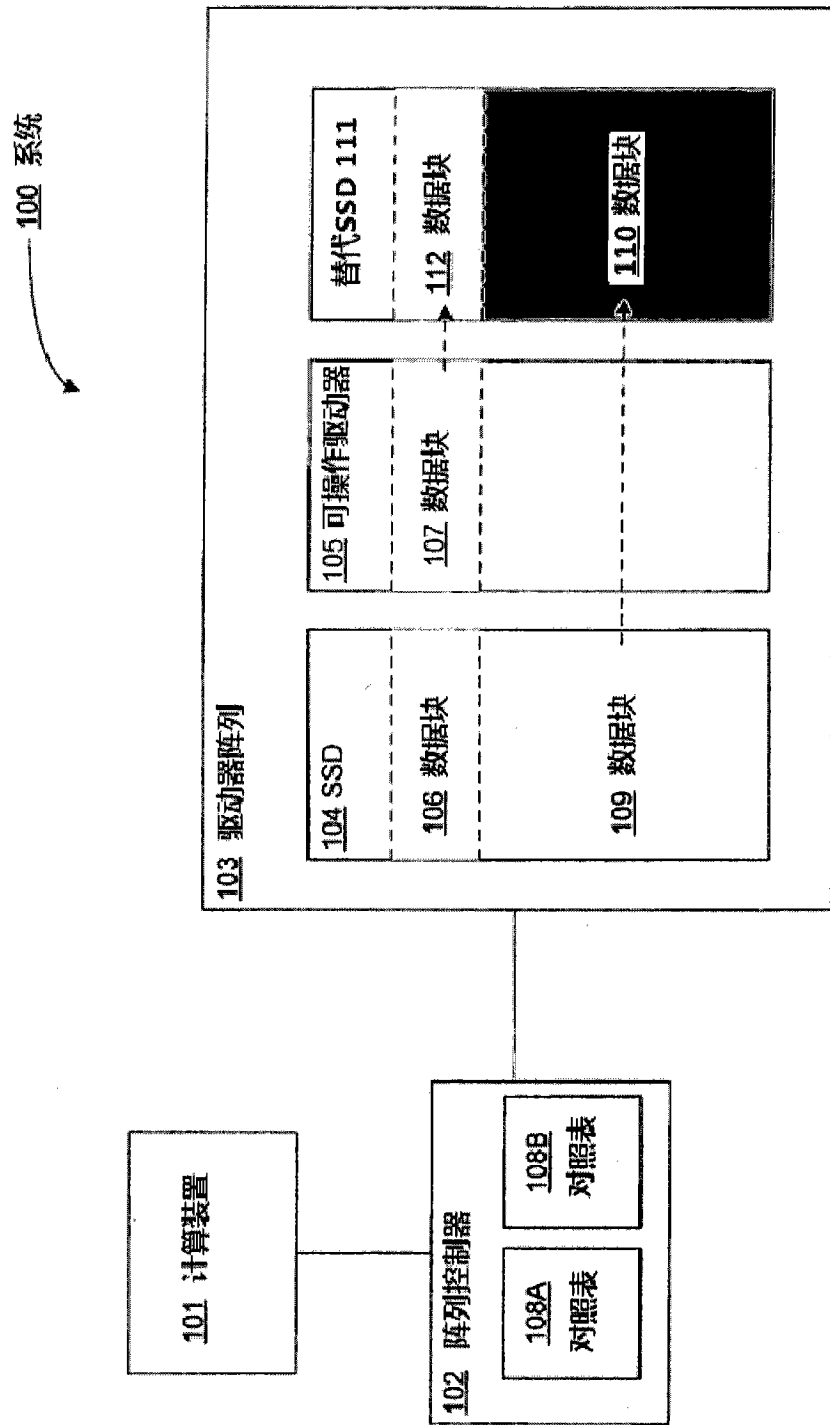


图 1

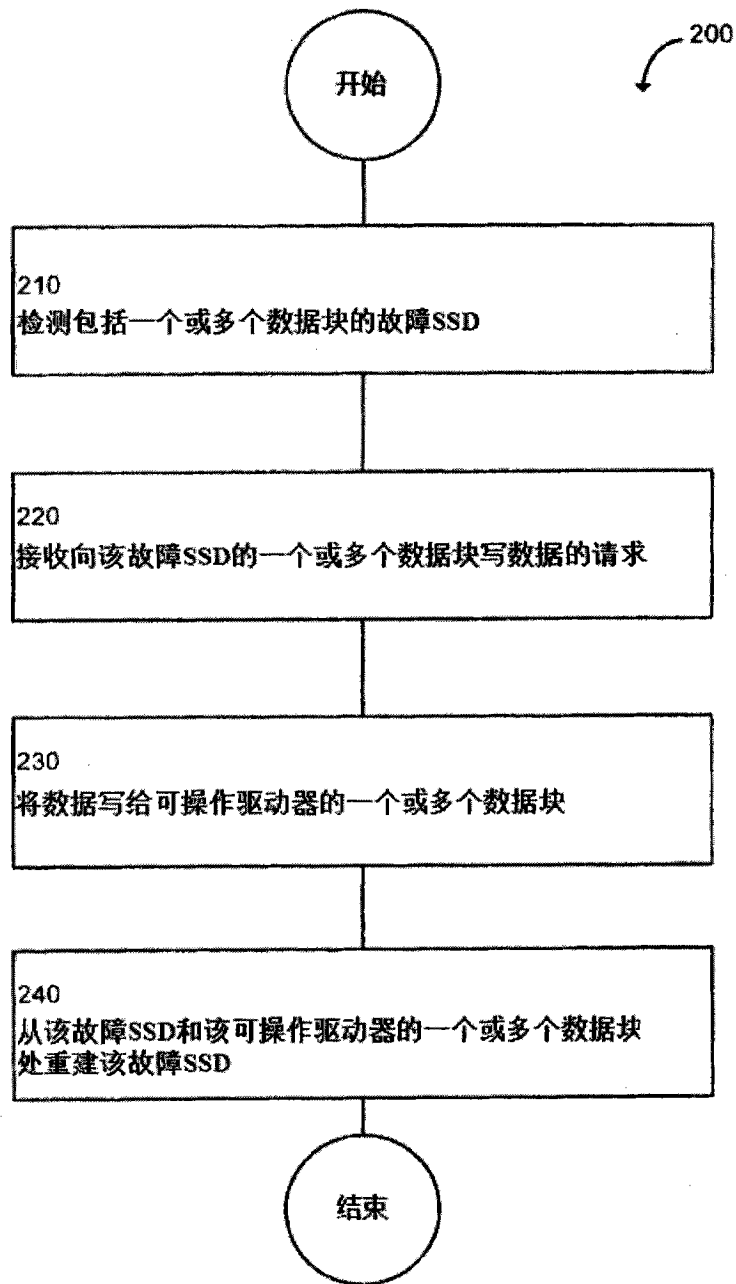


图 2

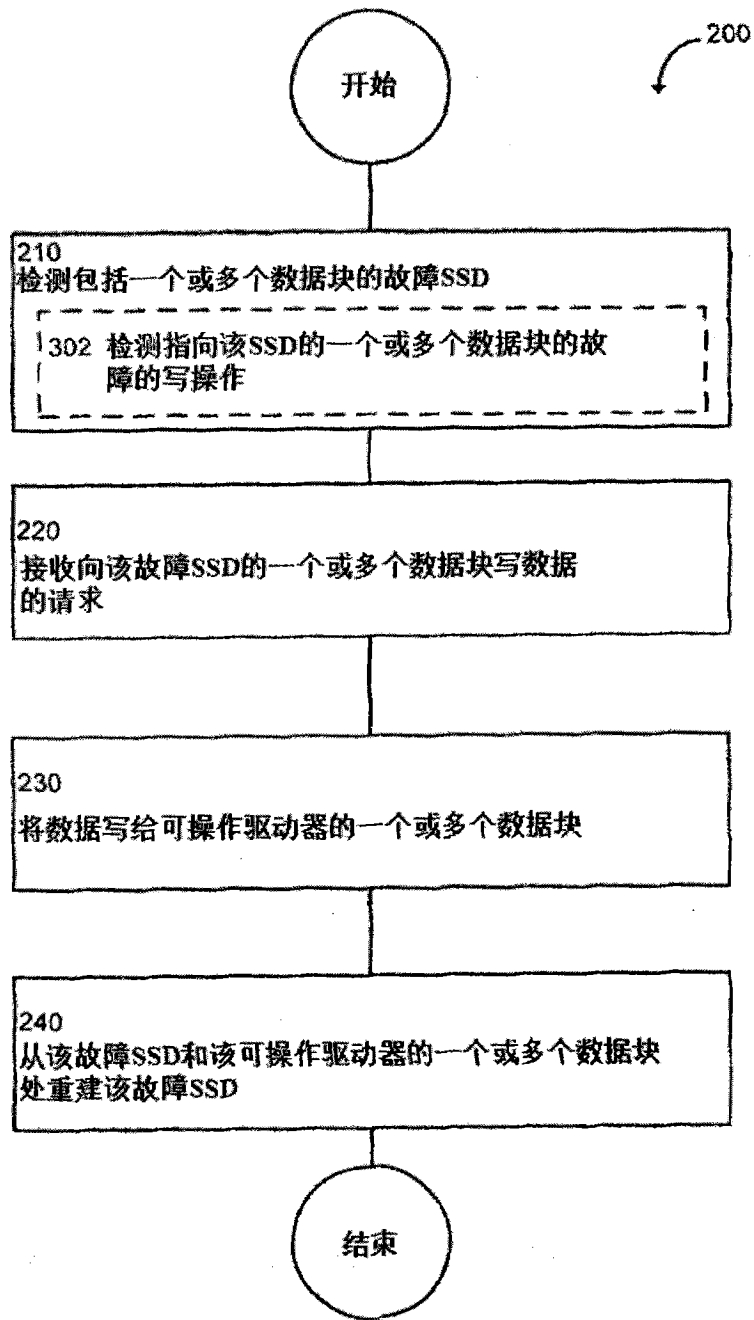


图 3

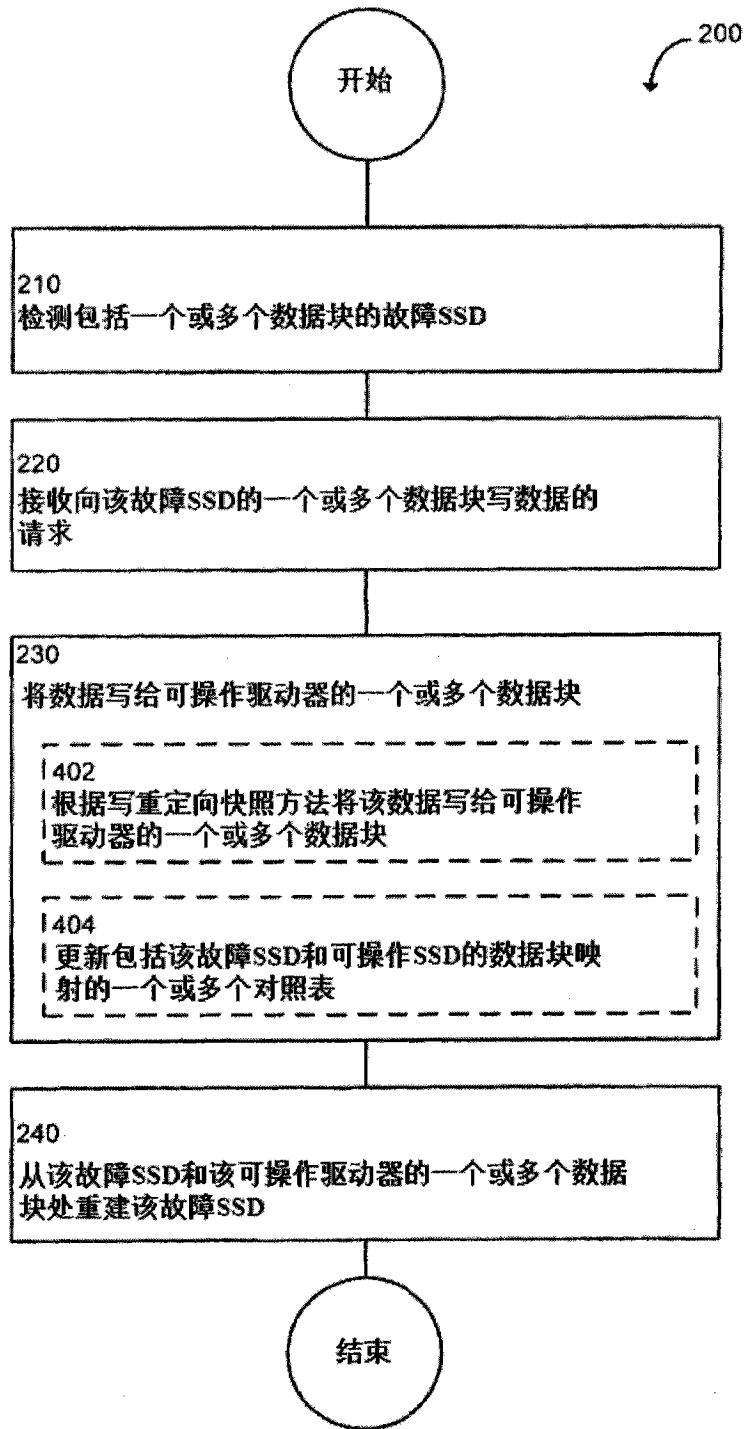


图 4

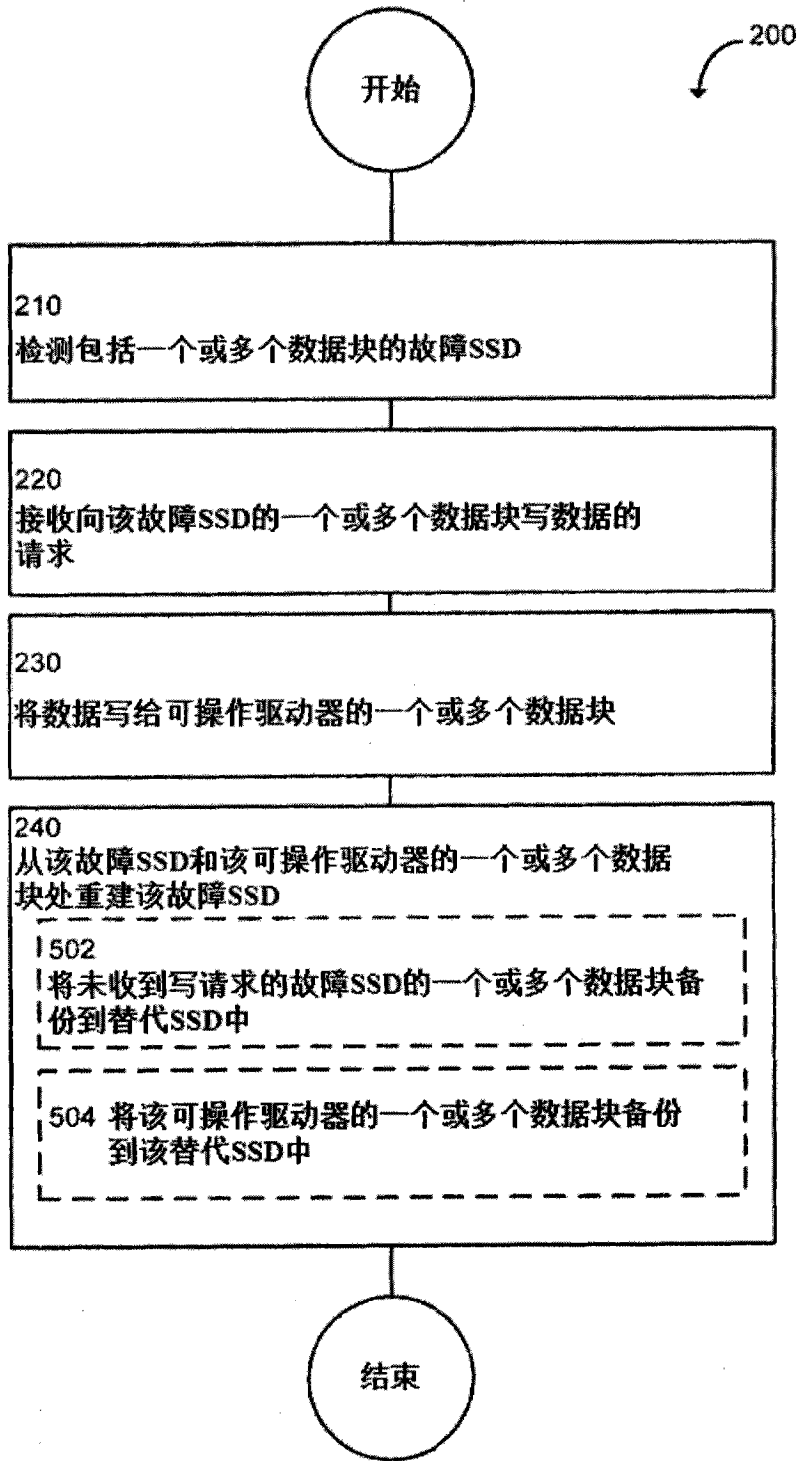


图 5

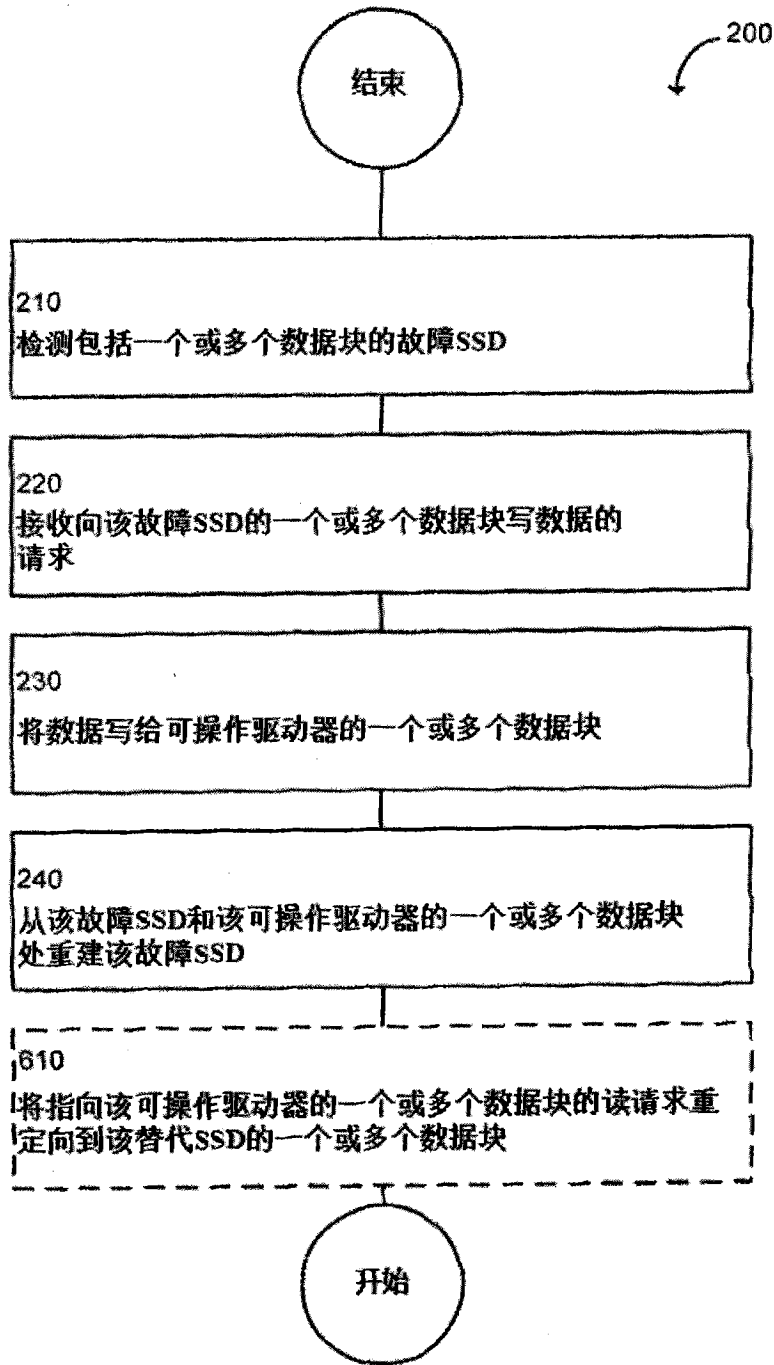


图 6

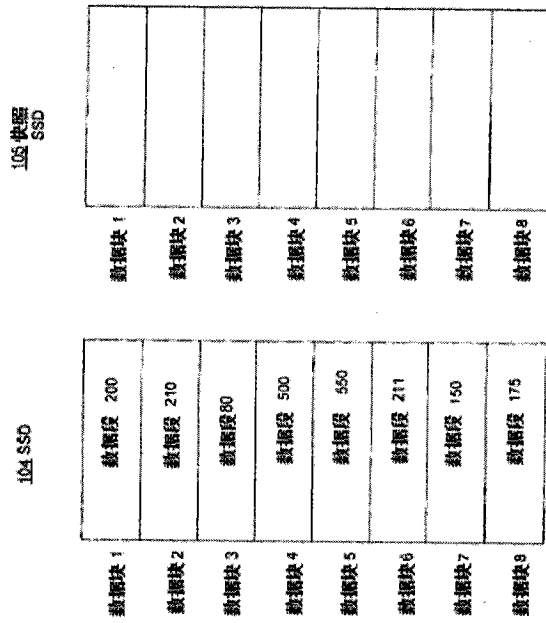


图 7A

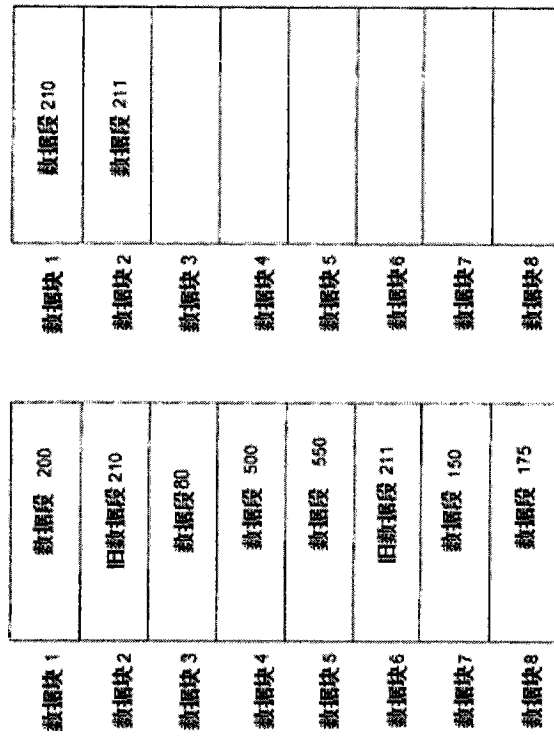


图 8A

数据段	SSD	数据块
80	104	3
...
150	104	7
...
175	104	8
...
200	104	1
...
210	104	2
211	104	6
...
500	104	4
...
550	104	5

图 7B

数据段	SDD	数据段
80	104	3
150	104	7
175	104	8
200	104	1
210	105	1
211	105	2
500	104	4
550	104	5

图 8B