

(12) 发明专利

(10) 授权公告号 CN 1953087 B

(45) 授权公告日 2011.09.21

(21) 申请号 200610136110.6

(56) 对比文件

(22) 申请日 2006.10.11

US 6104561 A, 2000.08.15,

(30) 优先权数据

US 5163141 A, 1992.11.10, 说明书第2栏第  
34行到第53行.

11/255, 545 2005.10.21 US

审查员 刘晓华

(73) 专利权人 国际商业机器公司

地址 美国纽约阿芒克

(72) 发明人 尼尔斯·豪斯汀 马丁·鲁森

丹尼尔·J·威纳斯基

克雷格·A·克莱因

(74) 专利代理机构 北京市柳沈律师事务所

11105

代理人 黄小临 王志森

(51) Int. Cl.

G11B 20/10 (2006.01)

G06F 12/00 (2006.01)

G06F 12/16 (2006.01)

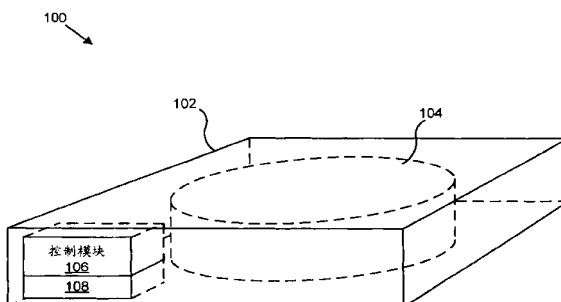
权利要求书 2 页 说明书 11 页 附图 11 页

(54) 发明名称

存储媒体中设置受保护分区保护状态的装  
置、系统及方法

(57) 摘要

一种用于设置存储媒体中受保护分区的保护  
状态的装置、系统以及方法。所述装置包括：控制  
模块，配置为从主机接收针对目标分区的保护状  
态转换命令；检查模块，配置为核查目标分区的  
保护状态；以及写模块，配置为根据多个所定义的  
可允许状态转换，执行保护状态转换。所述系统包  
括：主机设备；存储设备，耦合于主机，该存储设  
备包括具有存储媒体的封装和控制模块、以及  
所述装置。所述方法包括：从主机接收针对目标  
分区的保护状态转换命令；核查目标分区的保护  
状态；以及根据多个所定义的可允许状态转换执  
行保护状态转换。



1. 一种用于设置硬盘驱动器存储媒体中受保护分区的保护状态的装置, 所述装置包括 :

控制模块, 配置为从主机接收针对包括多个逻辑块地址 LBA 的目标分区的保护状态转换命令 ;

检查模块, 配置为核查所述目标分区的多个保护状态中的保护状态, 其中, 所述多个保护状态包括 :

未受保护状态, 配置为允许执行无限定数目的读 / 写操作 ;

在该目标分区不存在未写的逻辑块地址 LBA 的情况下, 一写多读 WORM 受保护状态, 配置为不能被重新设置为未受保护状态, 以及保护所述分区免遭任何试图的写、擦除以及格式化命令 ;

自动受保护状态, 配置为只允许一个写操作, 并随后自动转换至只读自动保护状态, 而该目标分区中尚未被写的其余的逻辑块地址保留在自动受保护状态 ; 以及

在该目标分区具有未写的逻辑块地址 LBA 的情况下, 只读自动受保护状态, 配置为不能被重新设置为未受保护状态, 以及一旦数据已写入该分区, 则防止写、擦除以及格式化命令 ; 以及

写模块, 配置为根据多个所定义的可允许状态转换, 把目标分区的保护状态写入该硬盘驱动器存储媒体的保留分区, 所述多个可允许转换包括 :

从未受保护状态向用于正常读 / 写操作的未受保护状态、自动受保护状态、以及一写多读 WORM 受保护状态之一的转换 ;

从自动受保护状态向未受保护状态或只读自动受保护状态之一的转换, 以及

当目标分区中所有逻辑块地址都被写时从只读自动受保护状态向一写多读 WORM 受保护状态的转换。

2. 根据权利要求 1 所述的装置, 还包括解析模块, 配置为从保护状态转换命令中抽取逻辑块地址。

3. 根据权利要求 1 所述的装置, 其中, 所述分区还包括至少一个固定大小的存储单元。

4. 根据权利要求 1 所述的装置, 其中, 所述分区还包括至少一个可变大小的存储单元。

5. 一种用于设置硬盘驱动器存储媒体中受保护分区的保护状态的系统, 所述系统包括 :

主机设备 ;

存储设备, 耦合于主机, 该存储设备包括具有存储媒体的封装和控制模块 ;

控制模块, 配置为从主机接收针对包括多个逻辑块地址 LBA 的目标分区的保护状态转换命令 ;

检查模块, 配置为核查所述目标分区的多个保护状态中的保护状态, 其中, 所述多个保护状态包括 :

未受保护状态, 配置为允许执行无限定数目的读 / 写操作 ;

在该目标分区不存在未写的逻辑块地址 LBA 的情况下, 一写多读 WORM 受保护状态, 配置为不能被重新设置为未受保护状态, 以及保护所述分区免遭任何试图的写、擦除以及格式化命令 ;

自动受保护状态, 配置为只允许一个写操作, 并随后自动转换至只读自动保护状态, 而

该目标分区中尚未被写的其余的逻辑块地址保留在自动受保护状态 ; 以及

在该目标分区具有未写的逻辑块地址 LBA 的情况下, 只读自动受保护状态, 配置为不能被重新设置为未受保护状态, 以及一旦数据已写入该分区, 则防止写、擦除以及格式化命令 ; 以及

写模块, 配置为根据多个所定义的可允许状态转换, 把目标分区的保护状态写入该硬盘驱动器存储媒体的保留分区, 所述多个可允许转换包括 :

从未受保护状态向用于正常读 / 写操作的未受保护状态、自动受保护状态、以及一写多读 WORM 受保护状态之一的转换 ;

从自动受保护状态向未受保护状态或只读自动受保护状态之一的转换, 以及

当目标分区中所有逻辑块地址都被写时从只读自动受保护状态向一写多读 WORM 受保护状态的转换。

6. 根据权利要求 5 所述的系统, 还包括解析模块, 配置为从保护状态转换命令抽取逻辑块地址。

7. 根据权利要求 5 所述的系统, 其中, 所述分区还包括至少一个固定大小的存储单元。

8. 根据权利要求 5 所述的系统, 其中, 所述分区还包括至少一个可变大小的存储单元。

9. 一种用于设置硬盘驱动器存储媒体中受保护分区的保护状态的方法, 所述方法包括 :

从主机接收针对包括多个逻辑块地址 LBA 的目标分区的保护状态转换命令 ;

核查目标分区的多个保护状态中的保护状态, 其中, 所述多个保护状态包括 :

未受保护状态, 配置为允许执行无限定数目的读 / 写操作 ;

在该目标分区不存在未写的逻辑块地址 LBA 的情况下, 一写多读 WORM 受保护状态, 配置为不能被重新设置为未受保护状态, 以及保护所述分区免遭任何试图的写、擦除以及格式化命令 ;

自动受保护状态, 配置为只允许一个写操作, 并随后自动转换至只读自动保护状态, 而该目标分区中尚未被写的其余的逻辑块地址保留在自动受保护状态 ; 以及

在该目标分区具有未写的逻辑块地址 LBA 的情况下, 只读自动受保护状态, 配置为不能被重新设置为未受保护状态, 以及一旦数据已写入该分区, 则防止写、擦除以及格式化命令 ; 以及

根据多个所定义的可允许状态转换来把目标分区的保护状态写入该硬盘驱动器存储媒体的保留分区, 所述多个可允许转换包括 :

从未受保护状态向用于正常读 / 写操作的未受保护状态、自动受保护状态、以及一写多读 WORM 受保护状态之一的转换 ;

从自动受保护状态向未受保护状态或只读自动受保护状态之一的转换, 以及

当目标分区中所有逻辑块地址都被写时从只读自动受保护状态向一写多读 WORM 受保护状态的转换。

10. 根据权利要求 9 所述的方法, 其中, 所述方法包括从读和写命令中抽取逻辑块地址和保护状态。

## 存储媒体中设置受保护分区保护状态的装置、系统及方法

### 技术领域

[0001] 本发明涉及存储媒体,更具体地讲,本发明涉及实现存储媒体中受保护的分区。

### 背景技术

[0002] 电子商务所创建的数据的剧增,使存储器成为各种规模的公司优先考虑的战略投资。鉴于存储器的优先性,开始主要关注:对以非可重写和非可擦除方式对数据进行归档的需求。这一需求部分地基于对满足例如由证券交易委员会与 HIPPA 所要求的金融数据的管理、以及法庭记录、客户记录、和其它长生命期信息的归档的法定要求的需求。传统上,光媒体已经用于以非可重写和非可擦除形式存储数据。也可把非可重写和非可擦除数据称为参照数据、固定内容数据、或一写多读 (WORM) 数据。

[0003] 由于存储子系统中盘价格的不断下降以及存储容量的不断增加,信息技术提供商正越来越多地把 WORM 数据转向基于盘的存储子系统。基于盘的存储子系统可以为存储区域网的一部分。存储网络工业协会 (SNIA) 把 SAN 定义为一种其主要目的是计算机系统和存储元件之间的数据传送的网络。SAN 可以包括:通信基础设施,其提供了物理连接;以及管理层,其对连接、存储元件以及计算机系统进行组织,从而可使数据传送安全与鲁棒性。SAN 还可以包括存储系统,该存储系统包括存储元件、存储设备、计算机系统、和 / 或器具、以及在网络上通信传送的所有控制软件。

[0004] 通常,存储区域网包括与存储器或盘控制器相连的多个存储设备,例如磁带驱动器或硬盘驱动器。总体上,盘控制器是配置为处理来自主机或客户机的读 / 写请求的服务器。主机可以运行诸如视窗 (Windows)、Linux、UNIX、AIX、zOS 等的各种操作系统。在大规模计算环境中,存储区域网是用于提供大量存储和可升级 (scalable) 服务器或存储控制器性能的理想方案。

[0005] 典型的情况是,在存储区域网环境中,主机从盘控制器请求数据。然后,盘控制器从包含所请求数据的、通常称为本地单元 (home location) 的具体存储设备提取数据。然后,盘控制器把数据发送至主机。如果主机修改了数据,则把数据发送回盘控制器,盘控制器把所修改的数据返回给本地单元。典型的情况是,主机等待来自盘控制器的指示读或写操作已经完成的响应。通常,本地单元呈盘驱动器上分区的形式。分区还包括至少一个固定大小的存储单元,或至少一个可变大小的存储单元。在一个实施例中,固定大小的存储单元为逻辑块地址 (LBA)。

[0006] 盘控制器还可以提供功能,诸如,提供异构服务器存取的能力、数据高速缓冲存储、诸如各种 RAID 实施以及集群的数据可用性特性、可升级性、设备可视化、复制服务、以及非可重写和非可擦除存储器 (WORM)。通常,由驻留在存储设备或盘驱动器之外、例如盘控制器中的微代码提供 WORM 功能,该功能防止了对存储在盘驱动器上的数据的修改、删除以及添加。通常,盘驱动器集成和封装于子系统中,并且不能被去除。也把提供 WORM 功能的存储子系统称为 WORM 存储子系统。

[0007] 尽管越来越多地把 WORM 存储子系统接受为对参照数据进行归档的法定手段,但

仍存在这样的安全性关注：盘驱动器上的数据固有地可重写和可擦除。防止归档数据被改变或删除的最好的防范措施由不允许修改或删除的控制器微代码、和封装在可封闭框架或壳体中的盘驱动器的物理保护提供。当把盘驱动器从存储子系统中去除时，盘驱动器上的数据不再安全，并且可以被盖写、删除或操纵。为了防止这一情况，可以按特定方式对数据进行格式化，以使对数据的改变更为困难，然而，仍有可能对该格式进行解码并且操纵数据。与媒体的物理性质提供数据修改的自然屏障的过去相比，会使归档数据的完整性冒更大的风险。而且，当前的 WORM 系统通常依赖于主计算机或盘控制器的专用（proprietary）接口，可能引发与现存信息技术基础设施相关的问题。

[0008] 从以上的论述可以看出，显然存在着对在硬盘驱动器中实现受保护分区的装置、系统以及方法的需求。有益的作法是，这样的装置、系统以及方法将利用标准化的接口，而不要求专用的硬件或接口。

## 发明内容

[0009] 针对这一技术的当前状态，具体地讲，针对这一技术中由当前可用的分区保护系统尚未完全解决的问题和需求，开发本发明。因此，开发本发明以提供一种用于设置克服以上所论述的这一技术的诸多或全部缺点的受保护分区的保护状态的装置、系统以及方法。

[0010] 设置受保护分区的装置配备有逻辑单元，该逻辑单元包含多个模块，所述模块配置为功能上执行保护分区所必须的步骤。所描述的实施例中的这些模块包括：控制模块，配置为从主机接收针对目标分区的保护状态转换命令；以及检查模块，配置为核查目标分区的保护状态。

[0011] 在另一实施例中，该装置包括写模块，该写模块被配置为根据多个所定义的可允许状态转换，来执行保护状态转换。所述多个保护状态还可以包括：自动受保护状态，将其配置为只允许一个写操作，并随后转换至只读自动受保护状态；以及只读自动受保护状态，配置为一旦数据已写入分区，则防止写、擦除以及格式命令。

[0012] 在一个实施例中，所述装置包括多个可允许保护状态转换。保护状态转换可以包括：从未受保护状态向用于正常读/写操作的未受保护状态、自动受保护状态、以及 WORM 受保护状态之一转换；从自动受保护状态向未受保护状态或只读自动受保护状态之一转换；以及从只读自动受保护状态向 WORM 受保护状态转换。

[0013] 在另一实施例中，所述装置包括解析（parse）模块，该解析模块配置为从保护状态转换命令抽取逻辑块地址。可以从由包括硬盘控制器、DVD-R、DVD-RW、DVD-RAM、HD-DVD、蓝光（Blu-Ray）、UDO、CD-R、CD-RW、磁光、相变、全息、匣式磁带（tape cartridge）驱动器、盒式磁带（tape cassette）驱动器、以及固态媒体的组中选择存储设备。分区可以包括：至少一个固定大小的存储单元或一个可变大小存储单元。

[0014] 本发明还提供一种实现存储媒体中受保护分区的系统。具体地讲，在一个实施例中，所述系统包括：以上所描述的装置、主机设备、以及耦合于主机的存储设备。该存储设备可以包括包含存储媒体的封装和控制模块。

[0015] 本发明还提供一种保护存储媒体中分区的方法。在所公开的实施例中，该方法实质包括针对所描述的装置与系统的操作而执行以上所提出的功能所必须的步骤。在一个实施例中，所述方法包括：从主机接收针对目标分区的保护状态转换命令；核查目标分区的

保护状态；以及根据多个所定义的可允许状态转换而执行保护状态转换。

[0016] 所述方法还包括从读和写命令中抽取逻辑块地址和保护状态。另外，该方法包括：允许从未受保护状态向用于正常读/写操作的未受保护状态、自动受保护状态、以及 WORM 受保护状态之一转换；允许从自动受保护状态向未受保护状态或只读自动受保护状态之一转换；以及允许从只读自动受保护状态向 WORM 受保护状态转换。

[0017] 贯穿这一说明书，对特性、优点或类似语言的引用，并不意味着通过本发明可实现的所有特性与优点应在或在本发明的任何单个实施例中。而是，应该把引用这些特性与优点的语言理解为意指：与某实施例相结合所描述的具体特性、优点或特征包含于本发明的至少一个实施例中。因此，在整个这一说明书中，对特性和优点以及类似语言的论述可以但不必涉及同一实施例。

[0018] 而且，可以以任何适当的方式，把所描述的本发明的特性、优点以及特征组合于一或多个实施例中。相关技术领域中的技术人员将会意识到，可在不采用具体实施例的一或多个具体特性或优点的情况下，实践本发明。在另一些例子中，在某些实施例中可以认识到的附加特性与优点可以不存在于本发明的所有实施例中。

[0019] 通过以下的描述和所附权利要求，本发明的这些特性与优点将会变得更加明显，也可以通过以下所阐述的本发明的实践获知本发明的这些特性与优点。

## 附图说明

[0020] 为了容易地理解本发明的优点，将参照附图中所说明的具体实施例，给出以上简要描述的本发明的更具体的说明。应该认识到，这些附图仅描述了本发明的一些典型实施例，因此，不应该把它们视为对本发明的范围的限制，将通过使用附图，更具体和详细地描述与解释本发明，其中：

- [0021] 图 1 是示意性的方框图，说明了根据本发明的存储设备的一个实施例；
- [0022] 图 2 是示意性的方框图，说明了根据现有技术的逻辑块寻址的一个实施例；
- [0023] 图 3 是示意性的方框图，说明了根据本发明的控制模块的一个实施例；
- [0024] 图 4 是示意性的方框图，说明了根据本发明的多个保护状态；
- [0025] 图 5 是示意性的方框图，说明了根据本发明的模式选择命令的一个实施例；
- [0026] 图 6 是示意性的方框图，说明了根据本发明的模式感测命令的一个实施例；
- [0027] 图 7 是示意性的方框图，说明了根据本发明的模式页的一个实施例；
- [0028] 图 8 是示意性的流程图，说明了根据本发明的一种用于实现可允许保护状态转换的方法的一个实施例；
- [0029] 图 9 是示意性的流程图，说明了一种用于处理写命令的方法的一个实施例；
- [0030] 图 10 是示意性的方框图，说明了根据本发明的写命令的一个实施例；
- [0031] 图 11 是示意性的方框图，说明了根据本发明的媒体分区模式页的一个实施例；以及
- [0032] 图 12 是示意性的方框图，说明了一种用于保护控制模块 106 免遭存储设备的有意损坏的方法的一个实施例。

## 具体实施方式

[0033] 已把本说明书中所描述的许多功能单元标记为模块,这样做是为了更特别地强调它们实现上的独立性。例如,可以把一个模块作为这样的一个硬件电路加以实现,该硬件电路包括:定制的(custom) VLSI 电路或门阵列;现货供应的(off-the-shelf) 半导体,例如逻辑芯片、晶体管、或其它离散部件。也可以以诸如场可编程门阵列的可编程硬件设备、可编程阵列逻辑、可编程逻辑设备等来实现模块。

[0034] 也可以由各种类型的处理器加以执行的软件来实现模块。例如,可执行代码的所标识的模块可以包括可以组织为对象、过程或函数的计算机指令的一或多个物理或逻辑块。尽管如此,所标识模块的可执行部分不需要物理地放置在一起,而可以包括存储在不同位置的异类的(disparate) 指令,当逻辑地结合在一起时,其包括模块,并且实现模块的所陈述的目的。

[0035] 实际上,可执行代码的模块可以为单个指令,也可以为多个指令,甚至可以在多个不同的代码段上、在不同的程序之间、以及跨越多个存储器设备加以分布。类似地,操作数据此处可以在模块内标识和说明,并且可以按任何适当形式体现,以及在任何适当类型的数据结构中加以组织。操作数据可以作为单个数据集合收集,或者可以在包括在不同存储设备上的不同位置上加以分布,以及可以仅作为系统或网络上的电子信号而存在,或至少部分地存在。

[0036] 贯穿这一说明书,对“一个实施例”、“实施例”或类似语言的引用,意指结合这一实施例所描述的具体特性、结构或特征包括于本发明的至少一个实施例中。因此,贯穿这一说明书,短语“在一个实施例中”、“在实施例中”或类似语言的出现,可以但不必全部涉及同一实施例。

[0037] 所提及的信号承载媒体可以呈能够生成信号、致使信号得以生成、或者导致机器可读指令的程序在数字处理装置上执行的任何形式。可以通过传输线、紧致盘、数字通用盘、磁带、Bernoulli(柏努利)驱动器、磁盘、穿孔卡、快闪存储器、集成电路、或其它数字处理装置存储器设备,实现信号承载媒体。

[0038] 而且,可以按任何适当的方式,把所描述的本发明的特性、结构或特征组合于一或多个实施例中。在以下的描述中,提供了诸多具体的细节,诸如程序设计、软件模块、用户选择、网络事务处理、数据库查询、数据库结构、硬件模块、硬件电路、硬件芯片等的例子,以提供对本发明各实施例的全面理解。然而,相关技术领域中的技术人员将会认识到,也可以在无一或多个所述具体细节,或通过其它方法、部件、设备等的情况下,实践本发明。在其它的例子中,没有详细示出或描述所熟知的结构、材料或操作,以避免淡化本发明的各个方面。

[0039] 图 1 是示意性的方框图,说明了根据本发明的存储设备 100 的一个实施例。存储设备 100 可以包括封装 102、存储媒体 104、控制模块 106、以及通信接口 108。存储媒体 104 可以包括但不限于磁、光以及固态存储媒体。磁存储媒体 104 的例子包括但不限于硬盘驱动器和磁带驱动器。通信接口 108 可以包括但不限于 HVD 或 LVD SCSI、光纤信道、ESCON、FICON 以及 RS-232。

[0040] 光媒体的例子包括 DVD-R(数字通用盘 - 一次写)、DVD-RW(可写 DVD)、DVD-RAM(DVD 随机存取存储器)、HD-DVD(高清晰度 DVD)、蓝光、UDO(超密光)、CD-R(紧致盘 - 一次写)、CD-RW(可重写 CD)、磁光(MO)、相变(PC)、全息等。固态存储媒体的例子包括:易失性存储器,例如 SDRAM;以及非易失性存储器,例如快闪存储器。

[0041] 控制模块 106 配置为通过通信接口 108 从主机接收输入和输出请求（以下称为“I/O 请求或命令”），并随后执行该 I/O 请求。可以把控制模块 106 可操作地与封装 102 中的存储媒体 104 相耦合。或者，可以把控制模块 106 配置为控制网络上的多个存储媒体 104。一个这样的实施包括存储区域网络（未示出）。

[0042] 控制模块 106 配置为执行对存储媒体 104 的最小存储单元进行存取的 I/O 请求、或读 / 写命令。在一个实施例中，最小存储单元为逻辑块地址 (LBA)。逻辑块寻址把存储媒体的物理位置 (location) 转换成可由主机加以使用的地址。例如，逻辑块寻址把硬盘驱动器的柱面、头 (head) 以及扇区规格翻译成可由操作系统或基本输入 / 输出系统 (BIOS) 所使用的地址。

[0043] 在另一实施例中，控制模块 106 可以经由通信链路连接于主机，并且可以类似地连接于存储媒体 104。如以上所描述的，存储媒体 104 可以经由备选的通信链路而包括盘驱动器、磁带驱动器和 / 或光驱动器。还把通信接口 108 配置为接收 I/O 请求，并且把 I/O 请求传送至控制模块。可以把控制模块 106 配置为执行从 ESCON 到 SCSI 或者从 FICON 到 SCSI 的协议转换。而且，还可以将其配置为把所附接的存储媒体 104 分组为逻辑驱动器或卷。控制模块 106 还可以提供先进的功能，例如，远程卷镜像、PPRC（对等远程拷贝）、XRC（扩展远程拷贝）、快照、快闪拷贝 (Flashcopy) 等。

[0044] 图 2 是示意性的方框图，说明了根据现有技术的逻辑块寻址 200 的一个实施例。逻辑块寻址 200 可用于诸如硬盘驱动器、光盘驱动器以及可去除存储驱动器的固定块体系结构 (FBA) 系统。在 FBA 外围中，通过执行存取 LBA 的读 / 写命令，来实现 I/O。LBA 引用固定量的数据。

[0045] 取决于存储媒体的容量，存储媒体 104 可以包括数百万个 LBA。图 2 说明了多个 LBA，每一个 LBA 表示为单个框 202。如这一技术领域中所熟知的，可以按任何次序写 LBA，不需要从第一 LBA202a 顺序地写至最后 LBA202b。

[0046] 也可以把最后 LBA202b 称为最大 LBA。根据媒体上 LBA 的数目和 LBA202 的固定大小，确定存储媒体 104 盘驱动器的容量。例如，假设典型的 LBA 的大小为 512 个字节，则具有 10,000,000 个 LBA 的硬盘驱动器占 51.2 亿个字节的容量。

[0047] 分区 204 是存储媒体 104 的孤立部分，可以用作为独立存储媒体 104。分区 204 包括至少一个 LBA202。例如，分区 204a 可以包括 3 个 LBA202，而分区 204b 包括一个 LBA202，以及分区 204c 包括 2 个 LBA202。或者，分区 204 可以包括数百万个 LBA202。存储媒体 104 可以包括多个分区，每一个分区 204 具有不同的保护状态。在另一实施例中，分区 204 从不交叠，这意味着任何两个分区不包含相同的 LBA202。

[0048] 图 3 是示意性的方框图，说明了根据本发明的控制模块 106 的一个实施例。在一个实施例中，控制模块 106 包括通信模块 302、写模块 304、解析模块 306、保护模块 308、以及检查模块 310。通信模块 302 配置为经由通信接口 108 与主机通信，以接收读 / 写命令，并随后把写命令传送至写模块 304，或者把数据作为读命令的结果返回至主机。

[0049] 在另一实施例中，解析模块 306 配置为从读 / 写命令抽取逻辑块地址和保护状态。检查模块 310 配置为核查存储媒体 104 的分区 204 和 LBA202 的保护状态。保护模块 308 配置为具有多个保护状态和保护状态之间的可允许转换。以下将参照图 4 更详细地论述保护模块 308。写模块 304 配置为把数据写至存储媒体 104，并且执行保护状态转换。

[0050] 图4是示意性的方框图,说明了根据本发明的多个保护状态。在一个实施例中,可以在保护模块308中实现保护状态。保护模块308可以包括未受保护状态402、一写多读(WORM)受保护状态404、自动受保护状态406、以及只读自动受保护状态(以下称为“只读状态”)408。未受保护状态402可以为针对整个存储媒体104的默认的出厂设置,并且允许针对任何LBA202进行无限定(indefinite)和重复次数的读/写操作。

[0051] 未受保护状态402保留针对存储媒体104上未明确地配备有WORM状态404或自动受保护状态406的那些区域的设置。把WORM状态404配置为保护与“WORM受保护”相关的任何分区204免于任何企图的写、擦除或格式操作。自动受保护状态406允许“自动受保护”分区204中的任何LBA202被仅写一次,而且在成功完成写操作之后,LBA202将自动转换至“只读自动受保护”状态408。尚未被写的分区中的其余的LBA202将保留在自动受保护状态406下。只读自动受保护状态408配置为一旦LBA202已被写,则防止任何写、擦除以及格式操作。

[0052] 箭头412代表说明分区204的保护状态的可允许状态转换的规则。例如,保护模块308配置为允许这样的分区204:其具有未受保护状态402,以允许重写数据,如箭头412a所指示的。可以通过发布以下将加以解释的适当的状态转换命令,把未受保护状态402转换412b至WORM受保护状态404。或者,也可以把未受保护状态作为状态转换命令的结果,而改变412c为自动受保护状态406。如图中所说明的,可以把WORM受保护状态404配置为不允许任何状态转换。

[0053] 可以把自动受保护状态406配置为仅允许向分区204发布一个成功的写命令。当已成功地处理了写命令时,将把自动受保护状态406的保护状态转换412d到只读自动受保护状态408。尽管存在未被写的LBA202,但把已被写的LBA202标记为只读自动受保护408,而未被写的LBA202保留在自动受保护状态406下。一旦已写了自动受保护的分区中的所有LBA202,则该分区中的所有LBA202变为412e WORM受保护404。

[0054] 在另一实施例中,也可以通过发送适当的状态转换命令412f,把具有自动受保护状态406的分区改变为未受保护状态402。例如,这使得能够纠正不希望的状态变化,例如,如果已错误地把分区的状态从未受保护设置为自动受保护。

[0055] 一旦分区处于WORM受保护状态404或只读自动受保护状态408下,则不能够盖写、删除或操纵数据。另外,一旦被赋值(assigned),则不能重新设置WORM保护状态404和只读自动受保护状态408。这意味着,控制模块106拒绝对这样的写、擦除或格式操作的任何I/O请求,所述写、擦除或格式操作对属于WORM保护状态404或只读自动受保护状态的任何LBA进行寻址。在另一实施例中,自动受保护406分区中的LBA202保持为自动受保护,直至其被写。如果自动受保护406分区中的至少一个LBA202已被写,则该分区改变为只读自动受保护408。如果自动受保护的分区406中的最后LBA202已被写,则只读自动受保护408分区将转换至WORM404受保护的分区。

[0056] 在一个实施例中,保护状态402、404、406、408可以在查询命令中表示为由b后缀表示的二进制值,或十进制值(括号中示出)。例如,可以把未受保护状态402表示为00b(十进制0),可以把WORM状态404表示为01b(十进制1),可以把自动受保护状态406表示为10b(十进制2),以及可以把只读自动受保护状态408表示为11b(十进制3)。

[0057] 在另一实施例中,保护模块308包括虚拟删除的受保护状态(未示出)。可以把

这一虚拟删除的状态配置为表示 WORM 或自动受保护的数据已被“虚拟删除”。出于历史目的,虚拟删除的数据将仍可由主机进行读访问,但出于任何原因,将把该数据标记为“虚拟删除”,例如,将该数据视为过期的或不再相关的。

[0058] 图 5 是示意性的方框图,说明了根据 SCSI-3 接口标准的模式选择命令 500 的一个实施例。所说明的模式选择命令 500 为 SCSI(小型计算机系统接口)命令,具有命令代码 15h 502。此处,针对状态转换命令,以举例的方式给出模式选择命令 500。例如,可以把模式选择命令 500 用于启动从未受保护状态 402 到 WORM 受保护状态 404 的转换 412b、或者从未受保护状态 402 到自动受保护状态 406 的转换 412c、或者从自动受保护状态 406 到未受保护状态 402 的转换 412f。模式选择命令 500 可以由主机系统(以下也称为启动器)加以启动,并且由控制模块 106 经由存储设备 100 的通信接口 108 加以接收。具有命令代码 15h 502 的模式选择命令 500 向主机提供了一种指定目标存储设备的参数列表长度 504 和逻辑单元 506 的手段。参数列表长度 504 指定按字节传送的模式页的大小。模式选择命令 500 后总是跟随从主机系统发送至存储设备 100 的模式页。

[0059] 图 6 是示意性的方框图,说明了根据本发明的模式感测命令 600 的一个实施例。模式感测命令 600 允许主机从例如来自存储设备 100 的模式页提取或查询模式页信息。模式感测命令 600 可以由主机系统(以下也称为启动器)启动,并且由控制模块 106 经由存储设备 100 的通信接口 108 接收。SCSI 模式感测命令的命令代码 602 为 1Ah。模式感测命令 600 允许主机指定页代码 604 和分配长度 606,该分配长度 606 然后指定最大返回数据大小。页代码 604 指定将提取的模式页的页代码、和目标存储设备的逻辑单元 608。分配长度指定了模式页的长度。作为模式感测命令 600 的结果,接收模式感测命令 600 的存储设备将返回适当的模式页。

[0060] 图 7 是示意性的方框图,说明了根据本发明的模式页 700 的一个实施例。图 6 的分配长度 606 指定作为模式感测命令 600 的结果、从存储媒体 104 传送到主机的字节的数目,即模式页 700 的大小。在一个实施例中,模式感测命令 600 将允许主机获取分区或个别 LBA 的保护状态。

[0061] 由存储设备作为对模式感测命令 600 的响应而发送的模式页 700 允许主机系统查询分区或 LBA 的状态。并非模式页 700 中的所有字段均由模式感测命令加以使用。例如,如果把模式感测命令 600 的字段 604 设置为指定模式页代码 0Dh 的 0Dh,则存储设备可以返回具有以下特征的模式页 700:如果存储设备能够支持 WORM 分区,则把字段 702 设置为 0Dh,把字段 704 设置为 17h,把字段 705 设置为 1;如果激活了至少一个受保护的分区,则把字段 706 设置为 1,字段 708 返回存在于存储设备中的受保护的分区的数目,字段 710 返回其中随后的字段 711、712 以及 713 施加至的分区的数目,字段 711 指定分区的开始地址(或 LBA),字段 712 概述了分区的结束地址,以及字段 713 返回分区状态的保护状态,该保护状态为未受保护的 00b、WORM 受保护的 01b、自动受保护的 10b 或只读自动受保护的 11b 之一。在这一序列中,存储设备将针对每一被配置的分区多次返回模式页 700。

[0062] 可以把模式选择和模式感测命令相结合地用于查询保护参数,诸如具体分区的开始和结束地址(LBA)以及保护状态。该序列开始于发送模式选择命令 500,该模式选择命令 500 在字段 502 中指定 15h、在字段 504 中指定 17h、以及在字段 506 中指定存储设备的逻辑单元地址。然后,发送带有字段 710 的模式页 700,该字段 710 设置为向其请求信息的分区

的数目，并且把模式页 700 中的所有其它字段设置为 00h,00h 表示这不是状态转换。把这两个命令从主机（启动器）发送至存储设备。

[0063] 在一个实施例中，然后，主机发送模式感测命令 600，其中字段 602 设置为 1Ah，把字段 604 设置为 0Dh，以及把字段 606 设置为 17h，该 17h 表示模式页的所期望的长度。把字段 608 设置为存储设备的地址。作为对这一命令序列的响应，存储设备返回模式页 700，在模式页 700 中，如果存储设备能够支持 WORM 分区，则把字段 702 设置为 0Dh，把字段 704 设置为 17h，把字段 705 设置为 1；如果激活了至少一个受保护的分区，则把字段 706 设置为 1，字段 708 返回存在于存储设备中的受保护的分区的数目，以及字段 710 返回向其请求信息的分区的数目。在一个实施例中，这一数目可以与先前发送的数目相同。字段 711 指定了分区的开始地址（或 LBA），字段 712 概述了分区的结束地址，以及字段 713 返回作为未受保护的 00b、WORM 受保护的 01b、自动受保护的 10b 或只读自动受保护的 11b 之一的该分区状态的保护状态。

[0064] 根据 SCSI 标准，在存储设备和主机系统之间双向地使用模式页。在本发明中，主机系统把模式页 700 与模式选择命令相关联地用于启动状态转换。主机系统还把这一模式页 700 与模式感测命令相关联地用于查询分区或 LBA 的状态。存储设备从模式页导出一或多个分区的保护状态。因此，把模式页存储在存储设备中。通常，存在多个模式页，每一个模式页具有唯一的页代码。示例性地把受保护的分区模式页的页代码设置为 0Dh。

[0065] 模式代码 702 指示将加以设置的受保护的分区。这是主机系统启动状态转换或查询状态转换所使用的唯一页代码。示例性地把受保护的分区模式页的页代码设置为 0Dh。字段 704 或模式页 700 定义了按照字节的这一模式页的长度，对于本例，将其设置为 25 个字节。字段 WORM 705 指定是否 WORM 保护机制由该产品支持，并且仅用于查询保护状态。字段 ACT 706 表示激活至少一个受保护的分区。这一字段也用于查询命令操作。如果把该字段设置为“0”，则不存在受保护的分区设置。否则，在这一盘驱动器上存在受保护的分区设置。

[0066] 分区 708 的字段最大数目表示可以使用开始和结束地址，例如 LBA 地址加以配置的分区的最大数目。这一字段用于查询命令操作。这是产品特定值，而且通常在存储设备的制造期间被赋值。这一值不能超过 FFFEh。通过该字段受保护的分区 710，可以指定现存的分区。把这一字段设置为值“0”，可以指定所有分区。值 FFFFh 表示将返回字段 711 中所指定的具体 LBA 的信息。

[0067] 开始边界 711 表示在该处这一分区将经历状态转换或查询操作开始的 LBA。结束边界 712 指定在哪个 LBA 处这一分区经历状态转换或查询操作结束。字段 711 和 712 将大于 0，而结束边界 712 必须等于或大于开始边界 711。保护状态 713 指定作为未受保护的 00b、WORM 受保护的 01b、自动受保护的 10b 或只读自动受保护的 11b 之一的这一分区状态的保护状态。

[0068] 如以上所概述的，可以根据图 4 中的可允许的转换，把保护状态配置为单一的 LBA 或一系列 LBA。与 00b、01b、10b 或 11b 的保护状态相关联的 LBA 的每一连续的序列是受保护的分区。因此，受保护的分区可以包括媒体上从 1 到 LBA 的最大数目的任何数目的 LBA。

[0069] 在一个实施例中，把控制模块 106 配置为接收现存的 SCSI 命令，例如模式选择 500 和模式感测 600，以创建具有 WORM 受保护状态 404 或自动受保护状态 406 的分区，或者以提

取可以是未受保护状态 402、WORM 受保护状态 404、自动受保护状态 406 或只读自动受保护状态 408 的分区状态。模式页 700 包括关于分区的信息 710、711、712、713、以及保护参数，例如将加以设置的 LBA 范围和保护状态。可以根据分区的创建与操纵，把保护模块 308 配置为利用可允许的状态 402、404、406、408。

[0070] 图 7 通过举例的方式，给出了作为实现本发明的一个实施例的模式页 700 代码 0Dh 702。或者，SCSI 命令技术领域中的技术人员将会意识到，也可以使用根据 SCSI 标准的保留而非使用的模式页的不同的页代码。在存储设备 100 和主机之间双向地使用模式页 700。主机可以结合模式选择 500 命令而使用模式页 700 以创建分区，即“选择模式”。主机结合模式感测 600 命令而使用这一模式页 700 以指定将针对其提取保护状态 402、404、406、408 的分区。控制模块 106 也可以使用模式页 700 来获取驻留在分区中的数据和相应的保护状态 402、404、406、408，并且当向主机请求时将其返回。

[0071] 下面论述的示意性的流程图是作为逻辑流程图总体性地加以描述的。因而，所描述的次序和所标识的步骤，是对所介绍的方法的一个实施例的说明性描述。可以构想功能、逻辑、或效果上等价于所说明的方法的一或多个步骤或所说明的方法的某些部分的其它步骤与方法。另外，所使用的格式与符号的提供，也旨在解释该方法的逻辑步骤，并且所使用的格式与符号应理解为不是对该方法的范围的限制。尽管在流程图中使用了各种类型的箭头和线型，但不应将它们视为对相应方法的范围的限制。实际上，可仅使用某些箭头或其它连接符来指示该方法的逻辑流。例如，箭头可以指示所描述的方法的所列步骤之间未指定持续时间的等待或监视周期。另外，具体方法出现的次序，可以、也可以不严格遵循所示相应步骤的次序。

[0072] 图 8 是示意性的流程图，说明了根据本发明的一种用于实现可允许保护状态转换的方法 800 的一个实施例。方法 800 开始 802，控制模块 106 的通信模块 302 经由接口 108 接收 804 模式选择命令 500 和模式页 700。然后，解析模块 306 从模式页 700 抽取 806 LBA 和新的保护状态。抽取 806 LBA 和保护状态可以包括生成 LBA 的列表和新保护状态。然后，检查模块 310 核查所获取的那些 LBA 的实际的保护状态。在一个实施例中，实际的状态为存储媒体 104 已经与每一所请求的 LBA 相关联的保护状态。然后，检查模块 310 针对从所述命令抽取的 LBA 的列表，获取 808 实际保护状态。

[0073] 然后，保护模块 308 把新保护状态与实际保护状态加以比较，并确定状态转换是否为可允许 810。在一个实施例中，预先确定和定义可允许的转换，如以上参照图 4 所描述的。如果对于从模式页抽取 804 的每一 LBA，允许 810 向新保护状态转换，则写模块 304 存储 812 该新保护状态，当通信模块 302 向主机返回成功响应时，方法 800 结束 814。然而，如果对于任何 LBA 都不允许转换 810，则保护模块 308 阻止这一转换，而且通信模块 302 向主机返回错误 813，例如 SCSI 检查条件状态。在一个实施例中，错误可以是“任务的异常结束”，或 abend(异常终止)、错误消息。然后，方法 800 结束 814。

[0074] 在一个实施例中，可以把存储媒体 104 配置为把模式选择命令 500 所给出的信息存储在内部存储器中，该内部存储器可以为检查模块 310 的一部分。例如，使用可负担的(affordable)1Mb 的存储器芯片，可以存储大约 40,000 个保护分区的信息。这一内部存储器允许针对例如处理写命令的随后操作对这一数据进行快速存取。在另一实施例中，可以周期性地把保护分区信息写至存储媒体 104 上的保留区域，通常在空闲期间进行此操作。

[0075] 图9是示意性的流程图,说明了一种用于处理写命令的方法900的一个实施例。方法900开始902,通信模块302经由通信接口108接收904写命令。然后,解析模块306从例如写命令1000的该写命令获取906所有命令地址(LBA)。命令地址包括将向其写数据的开始LBA,以及表示将被写的顺序LBA的数目的传送长度。然后,检查模块310获取908命令地址(LBA)的实际保护状态。可以把从写命令1000所获取906的LBA称为“CMD\_LBA”。然后,保护模块308把来自该命令的地址与存储媒体104上的相应的保护状态加以比较。如果保护模块308发现CMD\_LBA相当于910 WORM受保护的区域或只读(RO)自动受保护的区域,则保护模块308阻止写命令,而且通信模块302返回911错误消息,然后方法900结束912。在一个实施例中,错误消息可以包括以上所描述的abend错误消息。

[0076] 或者,如果检查模块310发现910所有CMD\_LBA赋予至可写分区,则写模块304把数据写914至存储媒体104。如果写未成功916,则控制模块106启动错误恢复过程918。在一个实施例中,可以通过执行写核查或读取刷新写入的数据,来确定写的成功。如果写成功916,则保护模块308判断920是否把刷新写入的数据写至自动受保护的分区406。如果确定为否,则方法900结束912。或者,如果确定为是,则把写模块304配置为确定922分区是否充满。如果属于该分区的所有LBA状态均处于只读自动受保护408,则该分区充满。从检查模块310中获取这一信息。

[0077] 如果该分区为充满的,则写模块304把该分区的保护状态改变924为WORM受保护的404。如果该分区为非充满的,则写模块304把刚被写的LBA的保护状态和分区本身改变926为只读自动保护408。然后,方法900结束912。

[0078] 图10是示意性的方框图,说明了根据本发明的具有逻辑单元号(LUN)1002的写命令1000的一个实施例。写命令1000还可以包括开始LBA1004、字节2~5、以及传送长度1006、字节7~8。在一个实施例中,从开始LBA1004获取CMD\_LBA,除以LBA大小的传送长度1006加1等于该写命令1000的结束LBA。在一个实施例中,可以把每个SCSI接口中固有的逻辑单元号1002用于从WORM数据中挑选(sort)可重写的数据。例如,可以跨越存储设备100的通信接口108的第一逻辑单元号(LUN-0)发送可重写数据I/O402。可以跨越同一存储设备100的同一通信接口108的LUN-2发送自动受保护406的数据I/O。可以跨越同一存储设备100的同一通信接口108的LUN-3发送WORM受保护404的数据I/O。可以把通过LUN-0发送的数据保存在与LUN-2和LUN-3的数据的分区不同的分区上。

[0079] 为了维持保护状态402、404、406、408的数据保护属性,禁止了某些命令。例如,在处于保护状态WORM受保护404、自动受保护406以及只读自动受保护408下的存储媒体104的分区上,将不支持格式单元命令和擦除命令。总体上,允许改变或删除存储在WORM受保护的或只读自动受保护的分区中的数据的所有命令将由存储控制器通过检查条件加以禁止和拒绝。

[0080] 图11是示意性的方框图,说明了根据本发明的媒体分区模式页1100的一个实施例。由页代码11h 1101指定模式页1100。可以容易地使控制模块106适合于控制单轴匣式磁带和双轴盒式磁带。匣式或盒式磁带可以具有可经由模式设置命令通过主机加以管理的多个分区。于是,对匣式或盒式磁带上的具体的分区,可以把保护状态设置为(a)未受保护、(b)WORM受保护、以及(c)自动受保护。

[0081] 可以使用模式页把有关保护状态的信息从主机传送至磁带驱动器。图11说明了

根据 INCITS T10/1434-D 的媒体分区模式页 1100。先前已经论述过的保留字段 1102 的两个比特 1104、00b ~ 11b, 可用于表示分区的保护状态。在一个实施例中, 特别是, 如果把磁带用作存储媒体, 则可以使用媒体格式识别字段 1106 表示保护状态。此处, 可以并入示出按字节和比特编码的具体分区的保护状态的方案。或者, 例如, 也可以开发把 SCSI 模式感测和模式选择命令转换成 ATA 或 IDE 命令的设备驱动器, 以把本发明延伸至非 SCSI 设备。

[0082] 对于独立 (stand-alone) 磁带驱动器, 磁带驱动器可以把用于分区和用于该分区中每一逻辑块的保护状态存储在每一可去除匣式或盒式磁带的每一个分区的初始化区域中。在一个实施例中, 可以把每一个分区的保护状态既存储在匣式或盒式磁带的初始化区域中, 也存储在匣式或封装存储器中。

[0083] 图 12 是示意性的方框图, 说明了一种用于保护控制模块 106 免遭存储设备 100 有意损坏的方法 1200 的一个实施例。存储设备 100 的有意损坏的例子是, 某人有意取出属于 WORM 受保护实体的盘, 并且修改包含其中的数据。在一个实施例中, 方法 1200 开始 1202, 控制模块 106 检测 1204 所破坏的存储媒体 104。

[0084] 当检测 1204 存储媒体 104 故障时, 控制模块 106 配置为启动 1206 适当的错误恢复, 并且公布错误消息。如果没有故障被检测 1204, 则方法 1200 继续检查盘故障。例如, 在 RAID 系统中, 适当的错误恢复的一个例子是, 重新建立 RAID。当已经替换 1208 了盘驱动器时, 控制模块 106 检查 1210 是否又插入了同一个盘, 如这可能是有意操纵, 即某人去除了盘、操纵它、并且重新插入它。在一个实施例中, 同一盘的检查 1210 包括把存储设备 100 的唯一系列号和制造商 ID 进行比较。如果盘与先前去除的盘不是同一盘, 则控制模块 106 把其视为新盘, 并且并入 1216 这一盘。

[0085] 如果盘是 1210 与以前的盘相同的盘, 则控制模块获取 1212 口令以越过 (override) 保护。必须由插入了该盘的用户输入口令。如果口令不正确 1214, 则控制模块 106 递增 1220 计数器。把该计数器配置为维持不正确口令尝试的次数。如果该计数器大于 1222 预先定义的最大值, 则控制模块 106 将隔离存储媒体 104, 并且返回 1224 错误状态。错误可以表示操纵数据的未经授权的尝试已检测到并且已清除的状态。如果该计数器不大于 1222 最大次数, 则控制模块 106 再次尝试从用户获取口令 1212。

[0086] 当验证 1214 了口令时, 控制模块 106 将并入 1216 所述盘。例如, 在 RAID 阵列的情况下, 重新建立将启动。如果盘与先前的盘相同, 则控制模块 106 将尝试执行对盘的诊断, 并且针对可能的缺陷对该媒体进行扫描。如果没有发现缺陷, 则把盘的数据视为有效, 因此, 并入过程可能不需要重新建立。然而, 如果盘不是 1210 同一个盘, 则控制模块 106 并入 1216 该盘, 并且将其视为新的空盘。这将要求重新建立过程。然后, 方法 1200 结束 1226。

[0087] 可以在不背离本发明的构思或实质特征的情况下, 按其它特定形式实施本发明。例如, 可以在光纤信道 (“FC”) 物理层上, 实现此处所描述的 SCSI 命令。或者, 也可以无限制地使用其它协议, 例如 FICON、以太网 (Ethernet)、千兆比特以太网、Infiniband (无限带宽)、TCP/IP、iSCSI、ATA、SATA 等。在所有方面, 所描述的实施例都仅应被视为说明性的, 而不是限制性的。因此, 本发明范围由所附权利要求, 而不是由以上的描述加以表示。落入权利要求的等同含义和范围的所有变化均将被包括在其范围内。

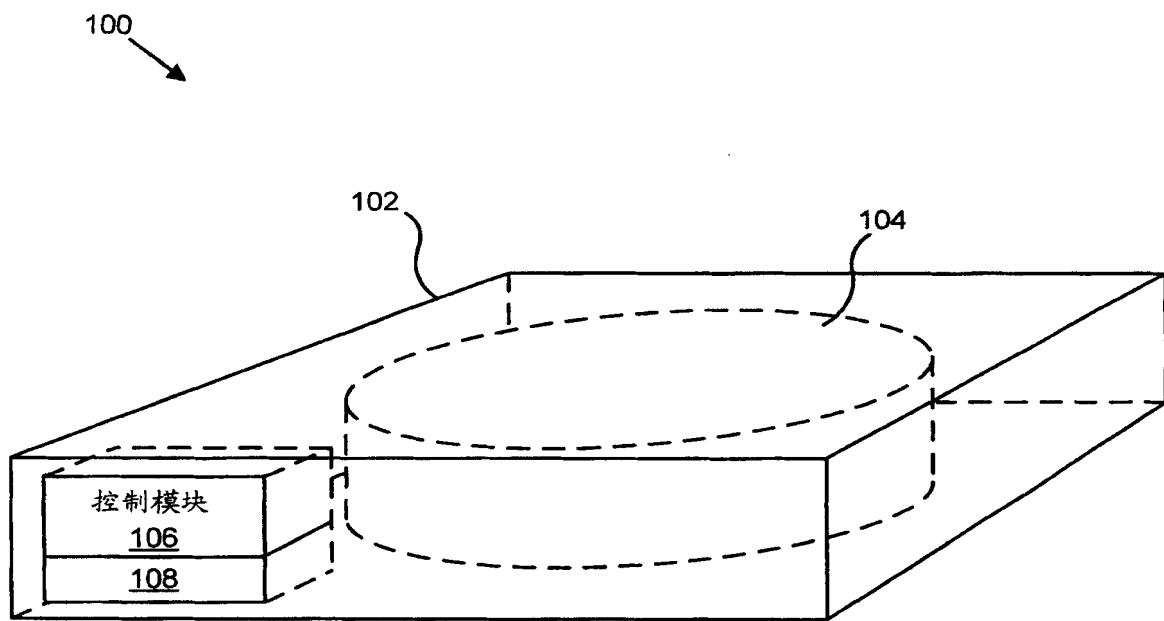


图 1

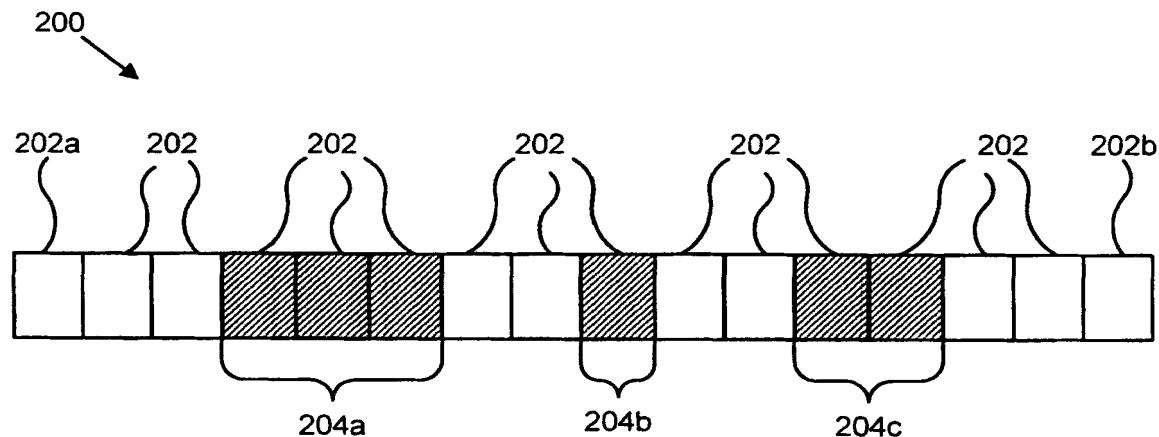


图 2

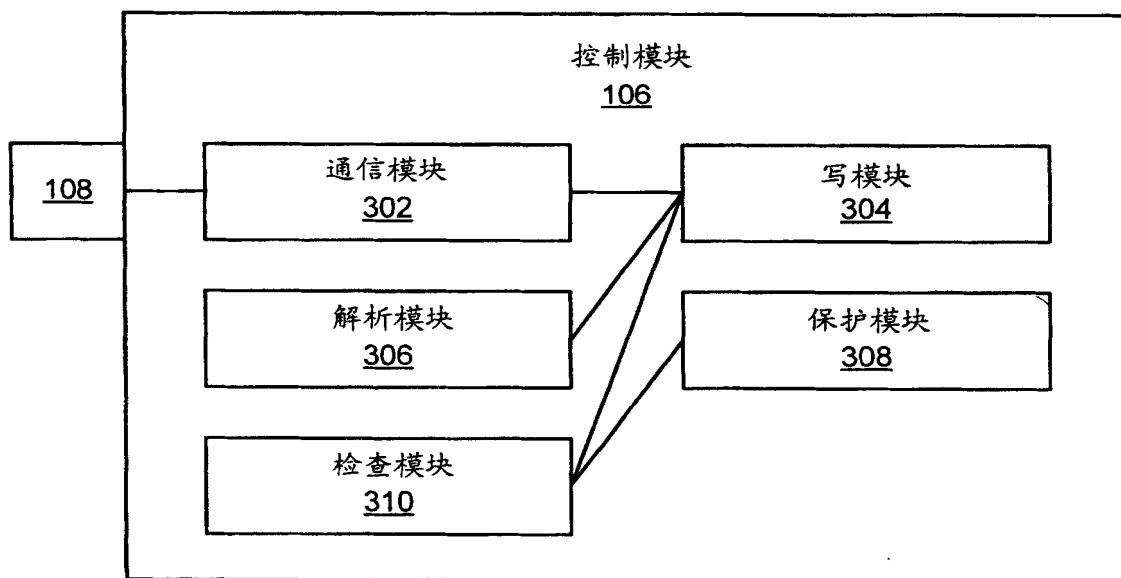


图 3

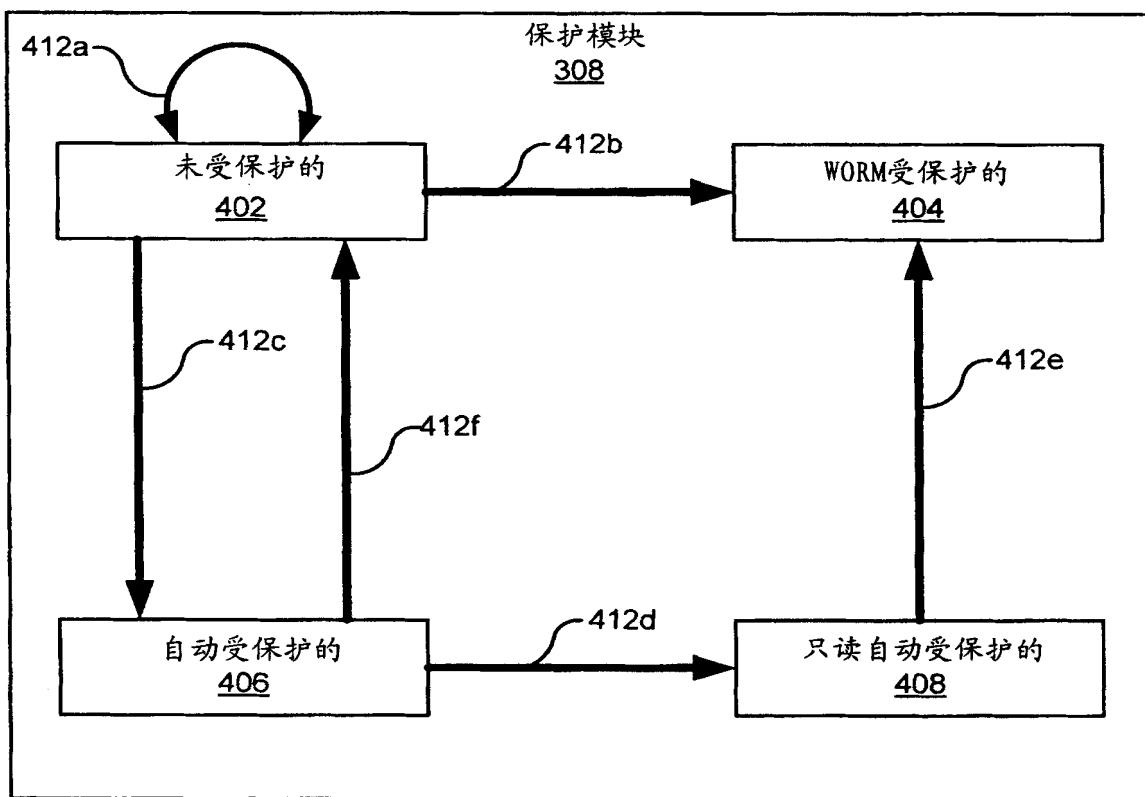


图 4

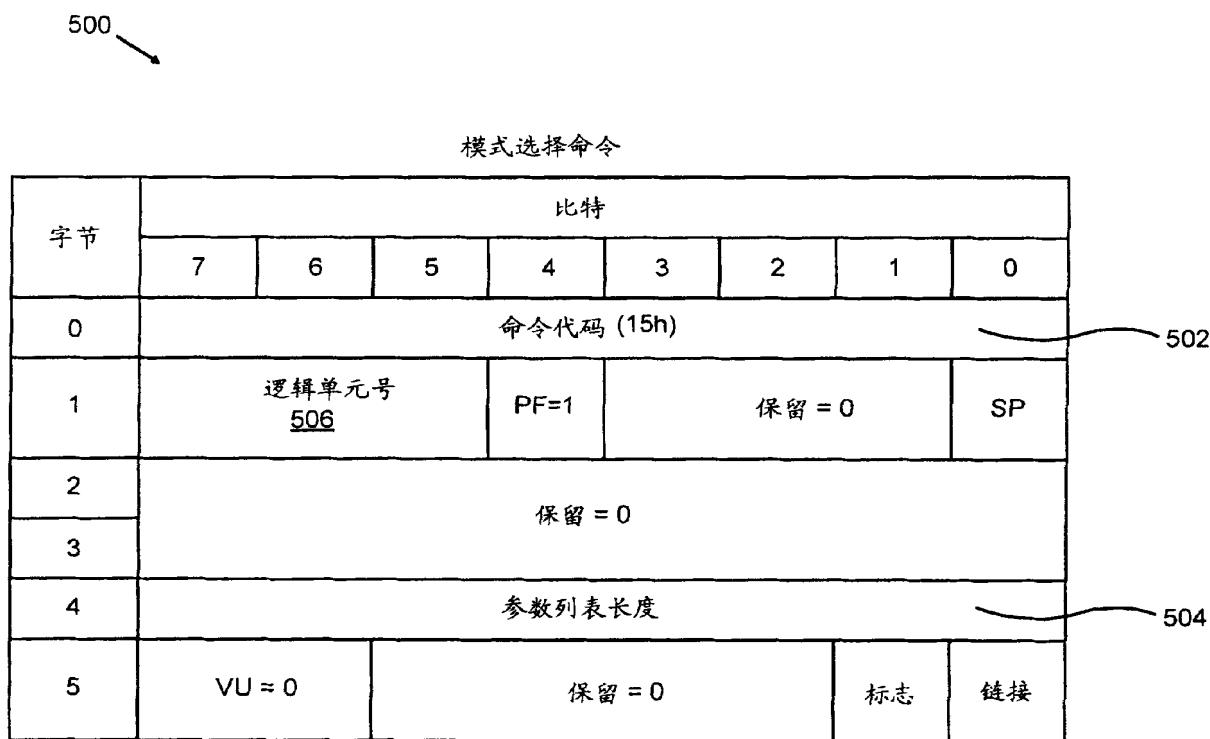


图 5

模式感测命令

字节	比特																
	7	6	5	4	3	2	1	0									
0	命令代码 (1Ah)																
1	逻辑单元号 <u>608</u>			RSVD	DBD	保留 = 0											
2	PCF		页代码														
3	子页代码																
4	分配长度																
5	VU = 0	保留 = 0			标志	链接											

图 6

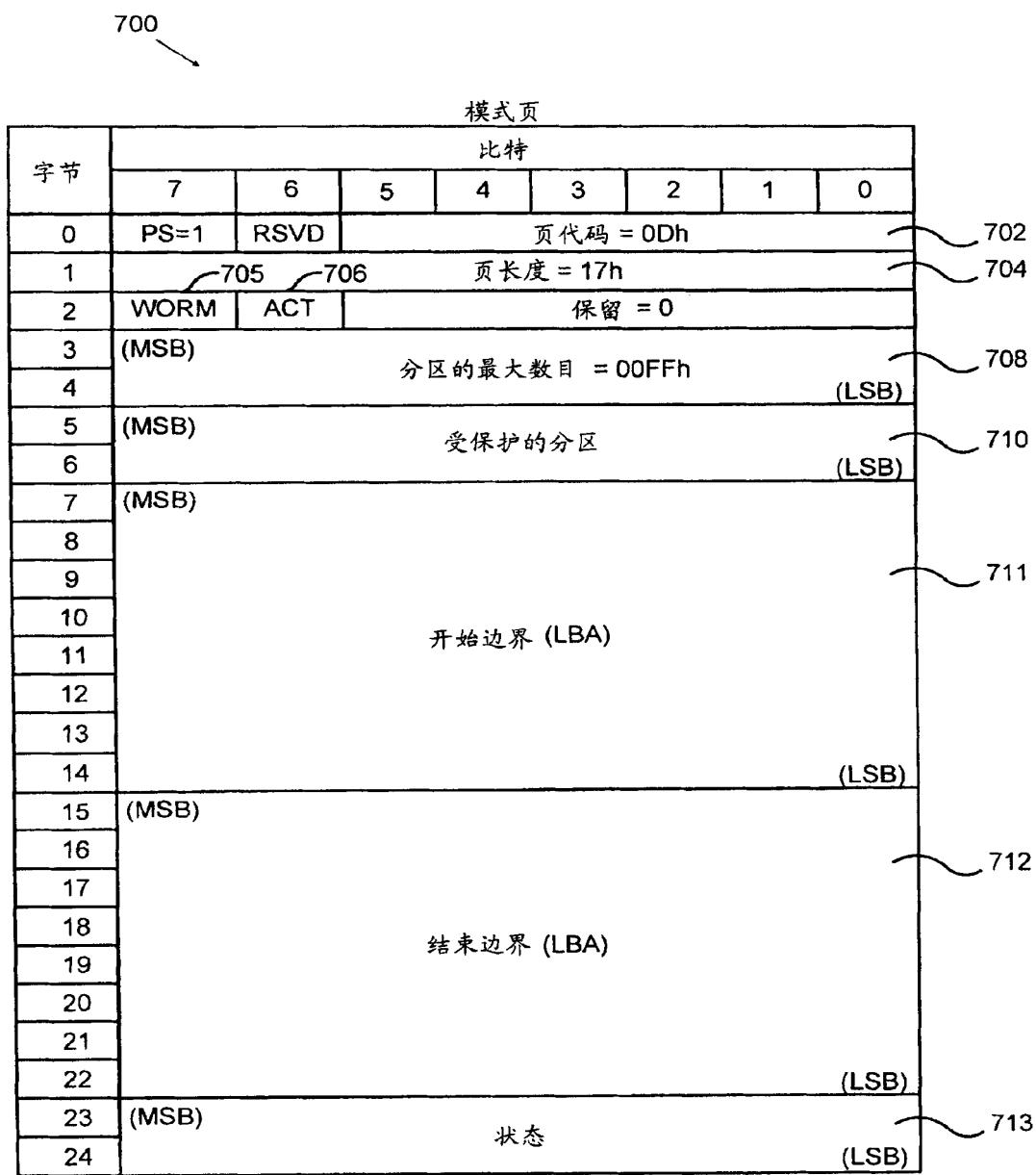


图 7

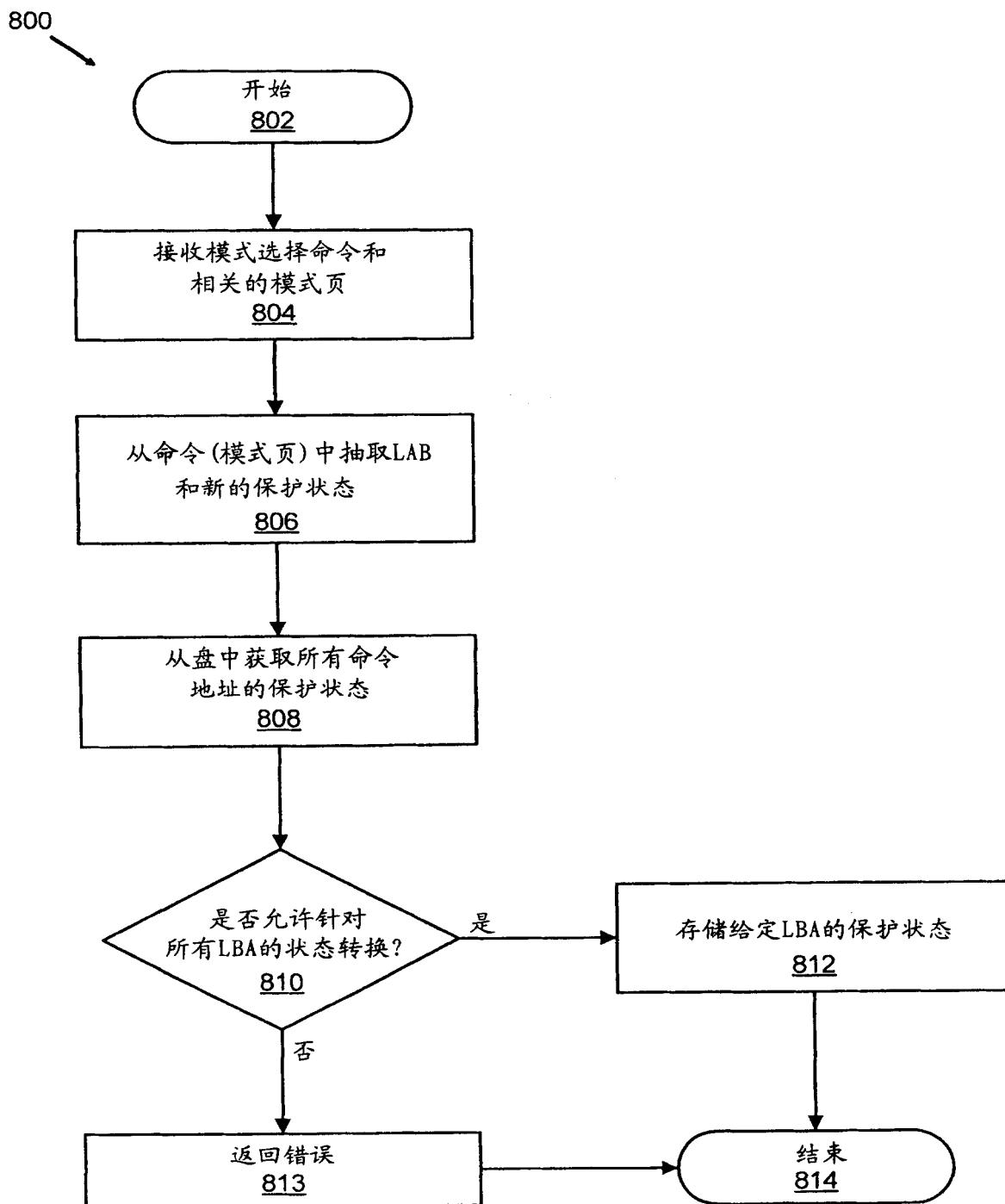


图 8

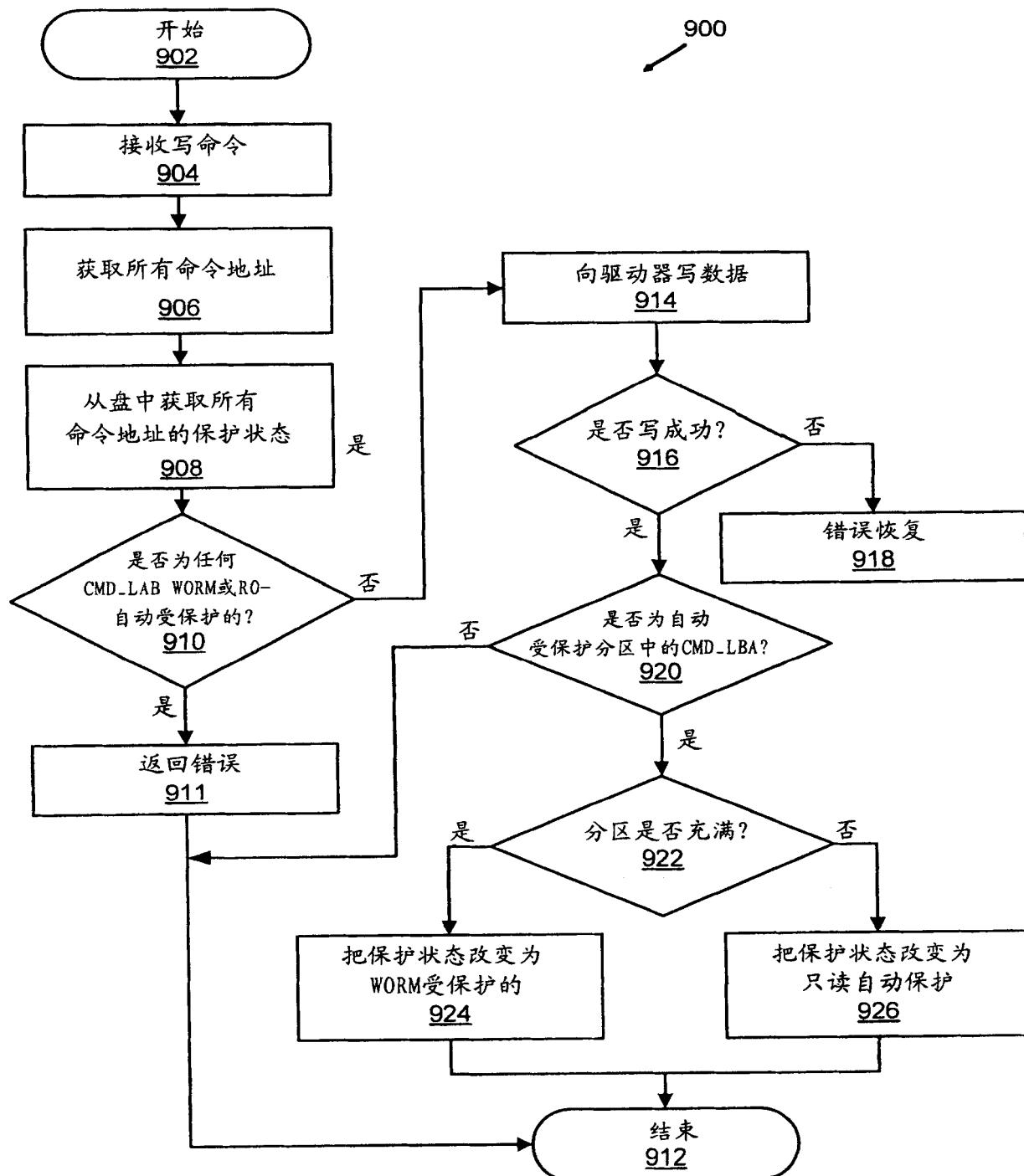


图 9

写命令

	比特7	比特6	比特5	比特4	比特3	比特2	比特1	比特0				
字节0	操作代码 (2Ah)											
字节1	逻辑单元号	DPO	FUA	保留的	RelAdr							
字节2	1002											
字节3	逻辑块地址 ↵1004											
字节4												
字节5												
字节6	保留											
字节7	传送长度 ↵1006											
字节8												
字节9	控制字节											

图 10

分区模式页

比特 字节	7	6	5	4	3	2	1	0
0	PS	Rsvd						页代码 (11h) ~1101
1								页长度
2								最大附加分区
3								所定义的附加分区
4	FDP	SDP	IDP	PSUM	POFM	CLEAR	ADDP	
5								媒体格式识别 ~1106
6	1104		保留					分区单元
7			1102	保留				
8	MSB							分区大小
9								LSB
n	MSB							分区大小
n-1								LSB

图 11

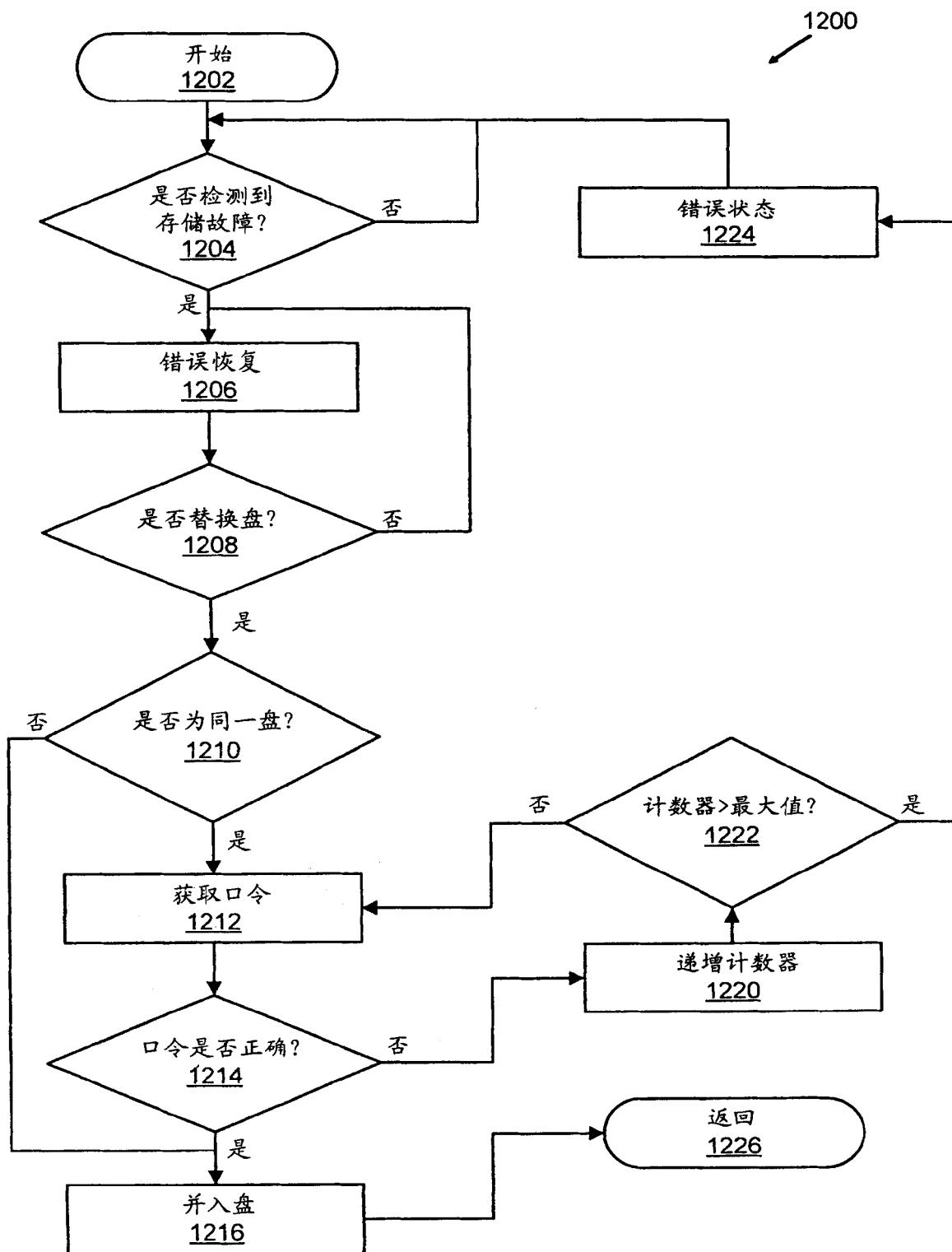


图 12