



(12)发明专利

(10)授权公告号 CN 105468534 B

(45)授权公告日 2019.01.18

(21)申请号 201510631086.2

(51)Int.Cl.

G06F 12/02(2006.01)

(22)申请日 2015.09.29

(56)对比文件

(65)同一申请的已公布的文献号
申请公布号 CN 105468534 A

US 8259506 B1, 2012.09.04,
JP 特开2011-100519 A, 2011.05.19,
US 8189379 B2, 2012.05.29,
US 2013/0094286 A1, 2013.04.18,
JP 特开2013-176784 A, 2013.09.09,
CN 103902234 A, 2014.07.02,
US 2011/0173484 A1, 2011.07.14,
US 2014/0006688 A1, 2014.01.02,
CN 102150140 A, 2011.08.10,
US 6016275 A, 2000.01.18,
US 2013/0176784 A1, 2013.07.11,

(43)申请公布日 2016.04.06

(30)优先权数据

14/501,800 2014.09.30 US

(73)专利权人 EMC公司

地址 美国马萨诸塞州

(72)发明人 H·塔巴雷茨 R·阿加瓦尔
M·F·巴里恩托斯 J·P·费雷拉
J·S·邦威克 M·W·夏皮罗

(74)专利代理机构 北京英赛嘉华知识产权代理
有限责任公司 11204

代理人 王达佐 王艳春

权利要求书3页 说明书9页 附图9页

(54)发明名称

使用读取阈值表提高闪存利用率的方法和
系统

(57)摘要

本发明公开了一种使用读取阈值表提高闪存利用率的方法和系统，其中，所述方法包括从客户端接收包括用于数据的逻辑地址的客户端读请求，使用逻辑地址来确定物理地址，其中，所述物理地址包括用于持久性存储器中的物理页的页号，确定用于数据的保持时间，确定与物理页相关联的编程/擦除(P/E)循环值，使用P/E循环值、保持时间以及页号来获得至少一个读取阈值，向包括物理页的存储模块发布包括所述至少一个读取阈值的控制模块读请求，以及使用所述至少一个读取阈值从物理页获得数据。



1.一种用于从持久性存储装置读取数据的方法,所述方法包括:

从客户端接收用于数据的客户端读请求,其中,所述客户端读请求包括逻辑地址;

确定对应于所述逻辑地址的物理地址,其中,所述物理地址包括用于所述持久性存储装置中的物理页的页号;

使用选自由所述物理地址和所述逻辑地址组成的组中的一个地址来确定用于所述数据的保持时间;

确定与所述物理页相关联的编程/擦除循环值;

识别与包括所述编程/擦除循环值、所述保持时间以及所述页号的元组相对应的多个读取阈值,其中,所述多个读取阈值包括至少一个默认读取阈值和至少一个非默认读取阈值;

由控制模块向存储模块发布包括所述多个读取阈值的控制模块读请求,其中,所述存储模块包括所述物理页、存储模块控制器以及存储器;

由所述存储模块接收所述控制模块读请求;

由所述存储模块控制器向所述存储器发布读命令;以及

使用所述多个读取阈值从所述物理页获得所述数据。

2.根据权利要求1所述的方法,还包括:

从所述客户端接收用于第二数据的第二客户端读请求,其中,所述第二客户端读请求包括第二逻辑地址;

确定对应于所述第二逻辑地址的第二物理地址,其中,所述第二物理地址包括用于所述持久性存储装置中的第二物理页的第二页号;

使用选自由所述第二物理地址和所述第二逻辑地址组成的组中的一个来确定用于存储在所述第二物理页上的所述第二数据的第二保持时间;

确定与所述第二物理页相关联的第二编程/擦除循环值;

识别与包括所述第二编程/擦除循环值、所述第二保持时间和所述第二页号的第二元组相对应的至少一个读取阈值;

向所述存储模块发布包括所述至少一个读取阈值的第二控制模块读请求,其中,所述存储模块包括所述第二物理页,其中,所述第二物理页不同于所述物理页;以及

使用所述至少一个读取阈值从所述第二物理页获得所述第二数据。

3.根据权利要求1所述的方法,其中,所述物理页位于所述存储模块中的固态模块上,其中,所述固态模块包括多层单元。

4.根据权利要求3的所述方法,其中,所述至少一个默认读取阈值被应用于选自由A阈值、B阈值以及C阈值组成的一组阈值中的至少一个,其中,所述至少一个非默认读取阈值被应用于选自所述一组阈值中的至少另一个。

5.根据权利要求1所述的方法,其中,确定所述保持时间包括:

确定所述数据被写入到所述物理页的第一时间;

确定与所述客户端读请求相关联的第二时间;以及

使用所述第一时间和所述第二时间来确定所述保持时间。

6.根据权利要求5所述的方法,其中,从存储器中的存储器内数据结构获得所述第一时间,其中,所述存储器位于控制模块中。

7. 根据权利要求5所述的方法,其中,从所述客户端读请求获得所述第二时间。
8. 根据权利要求1所述的方法,其中,识别所述多个读取阈值包括在存储器内数据结构中执行查找,其中,所述存储器内数据结构包括多个条目,其中,所述多个条目中的一个条目包括所述元组、所述至少一个默认读取阈值以及所述至少一个非默认读取阈值。
9. 根据权利要求1所述的方法,其中,所述多个读取阈值中的至少一个读取阈值包括电压值。
10. 根据权利要求1所述的方法,其中,所述多个读取阈值中的至少一个读取阈值包括位移值,其中,所述位移值对应于默认读取阈值的电压位移。
11. 根据权利要求1所述的方法,其中,所述编程/擦除循环值是编程/擦除循环值范围。
12. 一种数据管理系统,包括:
 - 存储模块,其包括存储模块控制器和持久性存储装置;以及
 - 控制模块,其被可操作地连接到所述存储模块和客户端,

其中,所述控制模块:

从客户端接收用于数据的客户端读请求,其中,所述客户端读请求包括逻辑地址;

确定对应于所述逻辑地址的物理地址,其中,所述物理地址包括用于所述持久性存储装置中的物理页的页号;

使用选自由所述物理地址和所述逻辑地址组成的组中的一个地址来确定用于存储在所述物理页上的所述数据的保持时间;

确定与所述物理页相关联的编程/擦除循环值;

识别与包括所述编程/擦除循环值、所述保持时间以及所述页号的元组相对应的多个读取阈值,其中,所述多个读取阈值包括至少一个默认读取阈值和至少一个非默认读取阈值;以及

向所述存储模块发布包括所述多个读取阈值的控制模块读请求,其中,所述存储模块包括所述物理页,

其中,所述存储模块:

接收所述控制模块读请求;以及

使用所述控制模块读请求中的所述多个读取阈值从所述物理页获得所述数据。
13. 根据权利要求12所述的系统,其中,所述持久性存储装置包括闪存,并且其中,所述物理页位于所述闪存中。
14. 根据权利要求13所述的系统,其中,所述闪存是选自由NOR闪存和NAND闪存组成的组中的一个。
15. 一种包括计算机可读程序代码的非临时计算机可读介质,所述计算机可读程序代码在被计算机处理器执行时使得计算机处理器能够:
 - 从客户端接收用于数据的客户端读请求,其中,所述客户端读请求包括逻辑地址;
 - 确定对应于所述逻辑地址的物理地址,其中,所述物理地址包括用于持久性存储装置中的物理页的页号;

使用选自由所述物理地址和所述逻辑地址组成的组中的一个地址来确定用于所述数据的保持时间;

确定与所述物理页相关联的编程/擦除循环值;

识别与包括所述编程/擦除循环值、所述保持时间以及所述页号的元组相对应的多个读取阈值，其中，所述多个读取阈值包括至少一个默认读取阈值和至少一个非默认读取阈值；

向存储模块发布包括所述多个读取阈值的控制模块读请求，其中，所述存储模块包括所述物理页、存储模块控制器以及存储器；

由所述存储模块接收所述控制模块读请求；

由所述存储模块控制器向所述存储器发布读命令；以及

使用所述多个读取阈值从所述物理页获得所述数据。

使用读取阈值表提高闪存利用率的方法和系统

技术领域

[0001] 本文公开的实施方式总体上涉及用于提高闪存利用率的方法和系统。更具体地，本文公开的实施方式涉及用于从持久性存储器读取数据的方法以及与该方法相关联的一种系统和一种非临时计算机可读介质。

背景技术

[0002] 用于存储系统的一个重要性能度量是与检索存储在存储系统中的数据有关的延迟。存储系统的性能随着读取延迟的减小而改善。如果存储系统能够可靠地从存储介质检索无错误数据，则可以减小用于存储系统的读取延迟。当未检索到无错误数据时，存储系统可执行附加动作以便从检索数据去除错误。例如，存储系统可使用诸如纠错码(ECC)和/或RAID之类的纠错机制来从检索数据去除错误或者另外生成无错误数据。纠错机制的使用导致读取延迟的增加，其伴随有相应的性能下降。

发明内容

[0003] 一般地，在一方面，本发明涉及一种用于从持久性存储器读取数据的方法，所述方法包括从客户端接收用于数据的客户端读请求，其中，所述客户端读请求包括逻辑地址，确定对应于所述逻辑地址的物理地址，其中，所述物理地址包括用于持久性存储器中的物理页的页号，使用选自由所述物理地址和所述逻辑地址组成的组的一个来确定用于所述数据的保持时间，确定与所述物理页相关联的编程/擦除(P/E)循环值，使用所述P/E循环值、所述保持时间、所述页号来获得至少一个读取阈值，向存储模块发布包括所述至少一个读取阈值的控制模块读请求，其中，所述存储模块包括所述物理页，以及使用所述至少一个读取阈值从所述物理页获得所述数据。

[0004] 一般地，在一方面，本发明涉及一种系统，包括存储模块，所述存储模块包括存储模块控制器和持久性存储器；以及被可操作地连接到所述存储模块和客户端的控制模块，其中，所述控制模块：从客户端接收用于数据的客户端读请求，其中，所述客户端读请求包括逻辑地址，确定对应于所述逻辑地址的物理地址，其中，所述物理地址包括用于所述持久性存储器中的物理页的页号，使用选自由所述物理地址和所述逻辑地址组成的组中的一个来确定用于存储在所述物理页上的所述数据的保持时间，确定与所述物理页相关联的编程/擦除(P/E)循环值，使用所述P/E循环值、所述保持时间以及所述页号来获得至少一个读取阈值；以及向所述存储模块发布包括所述至少一个读取阈值的控制模块读请求，其中，所述存储模块包括所述物理页，其中，所述存储模块：接收所述控制模块读请求；以及使用所述控制模块读请求中的所述至少一个读取阈值从所述物理页获得所述数据。

[0005] 一般地，在一方面，本发明涉及一种包括计算机可读程序代码的非临时计算机可读介质，计算机可读程序代码在被计算机处理器执行时使得计算机处理器能够：从客户端接收用于数据的客户端读请求，其中，所述客户端读请求包括逻辑地址，确定对应于所述逻辑地址的物理地址，其中，所述物理地址包括用于持久性存储器中的物理页的页号，使用

选自由所述物理地址和所述逻辑地址组成的组中的一个来确定用于所述数据的保持时间，确定与所述物理页相关联的编程/擦除 (P/E) 循环值，使用所述P/E循环值、所述保持时间以及所述页号来获得至少一个读取阈值，向存储模块发布包括所述至少一个读取阈值的控制模块读请求，其中，所述存储模块包括所述物理页，以及使用所述至少一个读取阈值从所述物理页获得所述数据。

[0006] 根据以下描述和所附权利要求，本发明的其它方面将是显而易见的。

附图说明

- [0007] 图1A-1C示出了根据本发明的一个或多个实施方式的系统。
- [0008] 图2示出了根据本发明的一个或多个实施方式的存储设备。
- [0009] 图3示出了根据本发明的一个或多个实施方式的存储模块。
- [0010] 图4示出了根据本发明的一个或多个实施方式的各种部件之间的关系。
- [0011] 图5A-5B示出了根据本发明的一个或多个实施方式的用于从存储模块读取数据的方法。
- [0012] 图6A-6B示出了根据本发明的一个或多个实施方式的示例。

具体实施方式

[0013] 现在将参考附图来详细地描述本发明的具体实施方式。在本发明的实施方式的以下详细描述中，阐述了许多特定细节以便提供本发明的更透彻理解。然而，对于本领域的技术人员而言将显而易见的是可在没有这些特定细节的情况下实施本发明。在其它情况下，并未详细地描述众所周知的特征以避免不必要地使本描述复杂化。

[0014] 在图1-6B的以下描述中，在本发明的各种实施方式中，相对于附图描述的任何部件可等价于相对于任何其它图描述的一个或多个类似名称的部件。简洁起见，将不会相对于每个图的这些部件重复描述。因此，通过引用而结合每个图的部件的每个实施方式并假定为可选地存在于具有一个或多个类似名称的部件的每个图内。另外，根据本发明的各种实施方式，图的部件的任何描述将被解释为除相对于任何其它图中的相应类似名称部件所述的实施方式之外、与之相结合或作为其替代可实现的可选实施方式。

[0015] 一般地，本发明的实施方式涉及一种通过在闪速存储器的寿命内动态地修改读取阈值来增加固态存储器的利用率。更具体地，本发明的实施方式涉及使用P/E循环值、保持时间以及页号以便确定(一个或多个)适当的读取阈值以在读取先前已被存储在固态存储器中的数据时使用。基于每个读请求而动态地改变读取阈值允许从固态存储器将检索更多无错误数据。当从固态存储器检索到无错误数据时，不需要执行纠错机制。作为结果，系统的性能提升。

[0016] 以下描述对用于实现本发明的一个或多个实施方式的一个或多个系统和方法进行描述。

[0017] 图1A-1C示出了根据本发明的一个或多个实施方式的系统。参考图1A，该系统包括被可操作地连接到存储设备102的一个或多个客户端(客户端A 100A、客户端M 100M)。

[0018] 在本发明的应实施方式中，客户端100A、100M对应于包括向存储设备102发布读请求和/或向存储设备102发布写请求的功能的任何物理系统。虽然在图1A中未示出，但客户

端100A、100M中的每一个可包括客户端处理器(未示出)、客户端存储器(未示出)以及实现本发明的一个或多个实施方式所需的任何其它软件和/或硬件。

[0019] 在本发明的一个实施方式中,客户端100A—100M被配置成执行包括文件系统的操作系统(OS)。文件系统提供了用于从存储设备102进行文件的存储和检索的机制。更具体地,文件系统包括执行所需动作以向存储设备发布读请求和写请求的功能。文件系统还提供了编程界面以使得能够创建和删除文件、读和写文件、在文件内执行查找、创建和删除目录、管理目录内容等。另外,文件系统还提供了管理界面以创建和删除文件系统。在本发明的一个实施方式中,为了访问文件,操作系统(经由文件系统)通常提供文件操作界面以在每个文件内部打开、关闭、读取和写入数据,和/或操作相应元数据。

[0020] 继续图1A的讨论,在本发明的一个实施方式中,客户端100A、100M被配置使用以下协议中的一个或多个与存储设备102通信:外围部件互连(PCI)、快速PCI(PCIe)、扩展PCI(PCI-X)、快速非易失性存储器(NVMe)、快速PCI结构上的快速非易失性存储器(NVMe)、以太网结构上的快速非易失性存储器(NVMe)、以及无限带宽结构上的快速非易失性存储器(NVMe)。本领域的技术人员将认识到本发明不限于上述协议。

[0021] 在本发明的一个实施方式中,存储设备102是一种包括易失性和持久性存储器且被配置成从一个或多个客户端100A、100M读请求和/或写请求的服务的系统。下面在图2中描述存储设备102的各种实施方式。

[0022] 参考图1B,图1B示出了其中将客户端100A、100M连接到以网状配置(在图1B中表示为存储设备网格104)布置的多个存储设备104A、104B、104C、104D的系统。如图1B中所示,以完全连接网状配置示出了存储设备网格104——亦即,存储设备网格104中的每个存储设备104A、104B、104C、104D被直接地连接到存储设备网格104中的每个其它存储设备104A、104B、104C、104D。在本发明的一个实施方式中,可将客户端100A、100M中的每一个直接地连接到存储设备网格104中的一个或多个存储设备104A、104B、104C、104D。本领域的技术人员将认识到在不脱离本发明的情况下可使用其它网状配置(例如,部分连接网格)来实现存储设备网格。

[0023] 参考图1C,图1C示出了客户端100A、100M被连接到以扇出配置布置的多个存储设备104A、104B、104C、104D的系统。在此配置中,每个客户端100A、100M被连接到一个或多个存储设备104A、104B、104C、104D;然而,在单独存储设备104A、104B、104C、104D之间不存在通信。

[0024] 本领域的技术人员将认识到虽然图1A—1C示出了被连接到有限数目的客户端的存储设备,但在不脱离本发明的情况下可将该存储设备连接到任何数目的客户端。本领域的技术人员将认识到虽然图1A—1C示出了各种系统配置,但本发明不限于上述系统配置。此外,本领域的技术人员将认识到在不脱离本发明的情况下可使用任何其它物理连接将客户端(无论系统的配置如何)连接到(一个或多个)存储设备。

[0025] 图2示出了根据本发明的一个或多个实施方式的存储设备的实施方式。该存储设备包括控制模块200和存储模块组202。下面描述这些部件中的每一个。一般地,控制模块200被配置成管理来自一个或多个客户端的读请求和写请求的服务。特别地,控制模块被配置成经由IOM(下面讨论)从一个或多个客户端接收请求,处理请求(其可包括向存储模块发送请求),并在该请求已被服务之后向客户端提供响应。下面包括关于控制模块中的部件的

附加细节。此外，下面参考图5A-5B来描述关于服务读请求的控制模块的操作。

[0026] 继续图2的讨论，在本发明的一个或多个实施方式中，控制模块200包括输入/输出模块(IOM) 204、处理器208、存储器210以及可选地现场可编程门阵列(FPGA) 212。在本发明的一个实施方式中，IOM 204是客户端(例如，图1A—1C, 100A、100M与存储设备中的其它部件之间的物理接口。IOM支持以下协议中的一个或多个：PCI、PCIe、PCI-X、以太网(包括但不限于根据IEEE 802.3a-802.3bj定义的各种标准)、无限带宽以及融合以太网(RoCE)承载远程直接存储器访问(RDMA)。本领域的技术人员将认识到在不脱离本发明的情况下可使用除上文所列的那些之外的协议来实现IOM。

[0027] 继续图2，处理器208是具有被配置成执行指令的单核或多核的一组电子电路。在本发明的一个实施方式中，可使用复杂指令集(CISC)架构或精简指令集(RISC)架构来实现处理器208。在本发明的一个或多个实施方式中，处理器208包括根复合体(root complex)(由PCIe协议定义)。在本发明的一个实施方式中，如果控制模块200包括根复合体(可将其集成到处理器208中)，则存储器210经由根复合体而连接到处理器208。可替代地，存储器210使用另一点到点连接机制而直接地连接到处理器208。在本发明的一个实施方式中，存储器210对应于任何易失性存储器，包括但不限于动态随机存取存储器(DRAM)、同步DRAM、SDR SDRAM以及DDR SDRAM。

[0028] 在本发明的一个实施方式中，处理器208被配置成创建并更新存储器内数据结构(未示出)，其中，该存储器内数据结构被存储在存储器210中。在本发明的一个实施方式中，该存储器内数据结构包括在图4中描述的信息。

[0029] 在本发明的一个实施方式中，处理器被配置成将各种类型的处理卸载到FPGA 212。在本发明的一个实施方式中，FPGA 212包括计算用于正被写入到(一个或多个)存储模块的数据和/或从(一个或多个)存储模块正读取的数据的检查和的功能。此外，FPGA 212可包括出于使用RAID方案(例如，RAID 2—RAID 6)在(一个或多个)存储模块中存储数据的目的计算P和/或Q奇偶信息的功能和/或执行恢复使用RAID方案(例如，RAID 2—RAID 6)存储的已损坏数据所需的各种计算的功能。在本发明的一个实施方式中，存储模块组202包括每个被配置成存储数据的一个或多个存储模块214A、214N。下面在图3中描述存储模块的一个实施方式。

[0030] 图3示出了根据本发明的一个或多个实施方式的存储模块。存储模块300包括存储模块控制器302、存储器(未示出)以及一个或多个固态存储器模块304A、304N。下面描述这些部件中的每一个。

[0031] 在本发明的一个实施方式中，存储模块控制器300被配置成接收从一个或多个控制模块读取数据和/或向其写入数据的请求。此外，存储模块控制器300被配置成使用存储器(未示出)和/或固态存储器模块304A、304N来为读和写请求服务。

[0032] 在本发明的一个实施方式中，存储器(未示出)对应于任何易失性存储器，包括但不限于动态随机存取存储器(DRAM)、同步DRAM、SDR SDRAM以及DDR SDRAM。

[0033] 在本发明的一个实施方式中，固态存储器模块对应于使用固态存储器来存储持久性数据的任何数据存储器件。在本发明的一个实施方式中，固态存储器可包括但不限于NAND闪存和NOR闪存。此外，NAND闪存和NOR闪存可包括单层单元(SLC)、多层单元(MLC)或三层单元(TLC)。本领域的技术人员将认识到本发明的实施方式不限于存储类存储器。

[0034] 图4示出了根据本发明的一个或多个实施方式的各种部件之间的关系。更具体地，图4示出了存储在控制模块的存储器中的各种类型的信息。此外，控制模块包括更新存储在控制模块的存储器中的信息的功能。可将下面描述的信息存储在一个或多个存储器内数据结构中。此外，如果(一个或多个)数据结构类型保持信息之间的关系(如下所述)，可使用任何数据结构类型(例如，阵列、链接列表、散列表等)来组织(一个或多个)存储器内数据结构内的以下信息。

[0035] 存储器包括逻辑地址400到物理地址402的映射。在本发明的一个实施方式中，逻辑地址400是从客户端(例如，图1A中的100A、100M)的角度看数据看起来常驻在该处的地址。换言之，逻辑地址400对应于客户端上的文件系统在向存储设备发布读请求时所使用的地址。

[0036] 在本发明的一个实施方式中，逻辑地址是(或包括)通过把散列函数(例如，SHA-1、MD-5等)应用到n元组而生成的散列值，其中，n元组是<对象ID，偏移ID>。在本发明的一个实施方式中，对象ID定义文件，并且偏移ID定义相对于文件的起始地址的位置。在本发明的另一实施方式中，n元组是<对象ID，偏移ID，产生时间>，其中，产生时间对应于创建该文件(使用对象ID来识别)时的时间。可替代地，逻辑地址可包括逻辑对象ID和逻辑字节地址，或者逻辑对象ID和逻辑地址偏移。在本发明的另一实施方式中，逻辑地址包括对象ID和偏移ID。本领域的技术人员将认识到可将多个逻辑地址映射到单个物理地址，并且逻辑地址内容和/或格式不限于上述实施方式。

[0037] 在本发明的一个实施方式中，物理地址402对应于图3中的固态存储器模块304A、304N中的物理位置。在本发明的一个实施方式中，可将物理地址定义为以下n元组：<存储模块，通道，芯片使能，LUN，平面，块，页号，字节>。

[0038] 在本发明的一个实施方式中，每个物理地址402与编程/擦除(P/E)循环值404相关联。P/E循环值可表示：(i)已经在由物理地址定义的物理位置上执行的P/E循环的数目，或(ii)P/E循环范围(例如，5,000-9,999P/E循环)，其中，在已经由物理地址定义的物理位置上执行的P/E循环的数目在P/E循环范围内。在本发明的一个实施方式中，P/E循环是数据到擦除块(即，用于擦除操作的最小可寻址单元，通常为一组的多个页)中的一个或多个页的写入和该块的擦除，任一顺序均可。

[0039] 可基于每个页、基于每个块、基于每组块和/或以任何其它水平的量化度来存储P/E循环值。控制模块包括在数据被写入到固态存储模块(和/或从其擦除)时适当地更新P/E循环值402的功能。

[0040] 在本发明的一个实施方式中，所有数据(即，客户端上的文件系统已经请求被写入到固态存储模块的数据)406与产生时间408相关联。产生时间408可对应于：(i)数据被写入到固态存储模块中的物理位置(由于客户端写请求、由于由控制模块发起的垃圾收集操作等)的时间；(ii)客户端发布用以将数据写入到固态存储模块的写请求的时间；或者(iii)对应于(i)或(ii)中的写事件的无单位值(即，序号)。

[0041] 在本发明的一个实施方式中，存储器内数据结构包括<保持时间，页号，P/E循环值>到一个或多个读取阈值412的映射。上述映射还可包括影响读取阈值(例如，温度、工作负荷等)的任何(一个或多个)其它系统参数(即，除保持时间、页号、P/E循环值之外的一个或多个参数)。在本发明的一个实施方式中，保持时间对应于在数据到固态存储模块中的物

理位置的写入与正在从固态存储模块中的相同物理位置读取数据的时间之间所经历的时间。可用时间单位来表示保持时间,或者可将其表示为无单位值(例如,当产生时间被表示为无单位值时)。在本发明的一个实施方式中,可将<保持时间,页号,P/E循环值>中的P/E循环值表示为P/E循环或P/E循环范围。

[0042] 在本发明的一个实施方式中,(一个或多个)读取阈值412对应于电压或位移值,其中,该位移值对应于默认读取阈值的电压位移。可将每个读取阈值表示为电压或者对应于电压的无单位数。

[0043] 在本发明的一个实施方式中,默认读取阈值由固态存储器模块的制造商指定。此外,位移值的量化度可由位移值指定,其中,位移值对应于相应默认读取阈值的电压位移。

[0044] 在本发明的一个实施方式中,读取阈值(包括默认读取阈值)对应于用来读取存储在固态存储模块中的数据的电压值。更具体地,在本发明的一个实施方式中,通过将存储器单元中的电压与一个或多个读取阈值相比较来确定逻辑值(例如,用于作为SLC的存储器单元的1或0或者用于作为MLC的存储器单元的00、10、11、01)。然后可基于该比较的结果来确定存储在存储器单元中的逻辑值。例如,如果给定电压(V)在B阈值之上且在C阈值以下,则存储在存储器单元中的逻辑值是00(参见例如图6B)。在本发明的一个实施方式中,固态存储器模块中的每个页可包括4-8K之间的数据。因此,存储模块控制器通常从多个存储器单元获得逻辑值以便对服务读请求。必须从其获得逻辑值的存储器单元的指定数目基于正在请求的数据量(经由读请求)和存储器单元的类型(SLC、MLC等)而改变。

[0045] 在本发明的一个实施方式中,通过执行实验以确定当以下变量中的至少一个被修改时应如何修改读取阈值来确定(一个或多个)读取阈值412:保持时间、P/E循环值以及页号。优化(一个或多个)读取阈值412以便能够成功地从固态存储器模块读取数据。具体地,针对<保持时间,P/E循环值,页号>的每个组合,确定最佳读取阈值。用于给定<保持时间,P/E循环值,页号>的最佳读取阈值是导致针对数据的给定保持时间、数据被存储在其上面的物理位置的P/E循环值、以及数据在其上面存储在固态存储器模块中的页的页号,而从固态存储器模块检索的数据中的最低位出错率(BER)的读取阈值。

[0046] 通过基于保持时间、P/E循环值以及页号来修改(一个或多个)读取阈值,存储设备将在给定保持时间、P/E循环值以及页号下可改变存储在给定存储器单元中的电压的各种变量考虑在内。换言之,当逻辑值“01”将被存储在存储器单元中时,存储模块控制器将足够数目的电子存储在存储器单元中以便具有对应于“01”的电压。随时间推移,存储在存储器单元中的电压基于保持时间、P/E循环值以及页号而改变。通过理解电压如何基于上述变量随时间而改变,可在从存储器单元读取逻辑值时使用适当的读取阈值以便检索“01”。

[0047] 例如,当保持时间是4个月,P/E循环值是30,000,并且页号是3时,可使用(一个或多个)第一读取阈值来成功地读取数据,而当保持时间是5个月,P/E循环值是30,000,并且页号是3时,可使用(一个或多个)第二读取阈值来成功地读取数据。

[0048] 如果使用默认读取阈值(而不是非默认读取阈值),则存在可从读取存储器单元获得不正确逻辑值(例如,“11”而不是“01”)的较高可能性。这又导致需要ECC或诸如RAID重构(即,使用一个或多个奇偶值的检索数据内的错误修正)之类的其它纠错机制以便修正检索数据中的错误并最终向请求客户端提供无错误数据。纠错机制的使用增加了服务客户端读请求所需的时间并因此降低了存储设备的性能。

[0049] 在本发明的一个实施方式中,可针对每个<保持时间,P/E循环值,页号>组合来提供(一个或多个)读取阈值。用于给定<保持时间,P/E循环值,页号>的(一个或多个)指定读取阈值可对应于(一个或多个)默认读取阈值或(一个或多个)非默认读取阈值(即,除(一个或多个)默认读取阈值之外的读取阈值)。

[0050] 在本发明的另一实施方式中,存储器(图2中的210)仅针对与(一个或多个)非默认读取阈值相关联的每个<保持时间,P/E循环值,页号>组合存储(一个或多个)非默认读取阈值。在此情形中,当使用非默认读取阈值对比使用(一个或多个)默认读取阈值导致从固态存储器模块正在读取较高百分比的无错误数据时,非默认读取阈值与给定<保持时间,P/E循环值,页号>组合相关联。此外,在此情形中,当使用默认读取阈值对比使用(一个或多个)非默认读取阈值导致从固态存储器模块正在读取较高百分比的无错误数据时,针对任何<保持时间,P/E循环值,页号>组合不存储(一个或多个)默认读取阈值。

[0051] 转到流程图,虽然连续地提出并描述流程图中的各种步骤,但本领域的技术人员将认识到可按照不同的顺序执行某些或所有步骤,可将其组合或省略,并且可并行地执行某些或所有步骤。

[0052] 图5A示出了根据本发明的一个或多个实施方式的用于由存储设备处理客户端读请求的方法。

[0053] 在步骤500中,由控制模块从客户端接收客户端读请求,其中,客户端读请求包括逻辑地址。

[0054] 在步骤502中,从逻辑地址来确定物理地址(其包括页号)。如上文所讨论的,控制模块中的存储器包括逻辑地址到物理地址的映射(参见图4,400、402的讨论)。在本发明的一个实施方式中,通过使用逻辑地址到物理地址的映射以及在步骤500中从客户端请求获得的逻辑地址来执行查找(或查询)而确定物理地址。

[0055] 在步骤504中,针对存储在物理地址处的数据确定保持时间(t)。可使用数据的产生时间(参见图4,408)和客户端请求的时间(例如,客户端发布客户端请求的时间、通过存储设备接收客户端请求的时间等)来确定保持时间。从控制模块的存储器(例如,图2,210)获得数据的产生时间。可通过确定客户端请求的时间与产生时间之间的差来计算保持时间。

[0056] 在步骤506中,确定用于物理地址的P/E循环值。可通过使用物理地址作为密钥在存储器内数据结构(位于控制模块的存储器中)中执行查找来确定P/E循环值。步骤506的结果可以是与物理地址相关联的实际P/E循环值(例如,与对应于物理地址的物理位置位于其中的块相关联的P/E循环值)或者可以是P/E循环值范围(例如,5,000—9,999次P/E循环),其中,与物理地址相关联的实际P/E循环值位于P/E循环值范围内。

[0057] 在步骤508中,使用以下密钥<保持时间,P/E循环值,页号>从存储器内数据结构(参见图4,410、412)获得零或更多读取阈值。在本发明的一个实施方式中,当(一个或多个)默认读取阈值将被存储模块控制器用来从物理地址读取数据时,步骤508的结果可以是零读取阈值。如上文所讨论的,当使用(一个或多个)默认读取阈值对比使用(一个或多个)非默认读取阈值导致从固态存储器模块正在读取较高百分比的无错误数据(即,没有位错误的数据)时,使用(一个或多个)默认读取阈值。在本发明的一个实施方式中,可获得一个或多个非默认阈值(参见图4,412)。如上文所讨论的,当使用(一个或多个)非默认读取阈值对

比使用(一个或多个)默认读取阈值导致从固态存储器模块正在读取较高百分比的无错误数据(即,没有位错误的数据)时,使用(一个或多个)非默认读取阈值。

[0058] 在本发明的一个或多个实施方式中,是否使用非默认读取阈值的确定可基于P/E循环值(在步骤506中确定)或保持时间(在步骤504中确定)。例如,当P/E循环值在阈值P/E循环值以下时,使用(一个或多个)默认读取阈值,并且同样地,不执行步骤508。另外或可替代地,当保持时间在阈值保持时间以下时,使用(一个或多个)默认读取阈值,并且同样地,不执行步骤508。当P/E循环值(在步骤506中确定)在阈值P/E循环值以上和/或保持时间(在步骤504中确定)在阈值保持时间以上时,执行在步骤508中所述的查找。

[0059] 继续图5A中的讨论,在步骤510中,使用在步骤508中获得的(一个或多个)读取阈值和物理地址来生成控制模块读请求。如果在步骤508中未获得读取阈值,则控制模块请求可(i)不包括读取阈值或者(ii)可包括一个或多个默认读取阈值,其中,控制模块不响应于在步骤508中正获得读取阈值而获得默认读取阈值。控制模块读请求的格式可以是存储模块控制器所支持的任何格式。

[0060] 在本发明的一个实施方式中,如果存在与给定读请求相关联的多个读取阈值(例如,参见图6B,A阈值、B阈值、C阈值),则上述阈值中的一个或多个可对应于默认读取阈值,而其它阈值可对应于非默认阈值。例如,读取阈值可以是<默认A读取阈值,非默认B读取阈值,默认C读取阈值>。此外,在其中存在多个读取阈值的情形中,可选地,可仅在步骤508中确定非默认阈值。例如,步骤508的结果可以是<非默认B阈值,非默认C阈值>,其指示应连同用于阈值B和阈值C的非默认读取阈值一起使用默认A读取阈值。

[0061] 图5B示出了根据本发明的一个或多个实施方式的用于处理控制模块读请求的方法。更具体地,由存储模块控制器来执行图5B。

[0062] 在步骤520中,从控制模块接收控制模块读请求。在步骤522中,由存储控制器模块基于一个或多个读取阈值和控制模块读请求中的物理地址来生成读命令。在本发明的一个实施方式中,在步骤522中产生的任何给定读命令可指定一个或多个读取阈值。如果控制模块不包括任何读取阈值,则使用默认读取阈值来生成读命令。如果控制模块读请求包括位移值(上文所述)形式的读取阈值,则生成读命令可包括获得默认读取阈值并使用(一个或多个)位移值来修改一个或多个读取阈值。读命令可以采取固态存储器模块所支持的任何格式。

[0063] 在步骤524中,向固态存储器模块发布读命令。在步骤526中,由存储模块控制器响应于读命令而接收数据。在步骤528中,向控制模块提供检索数据。控制模块随后向客户端提供数据。在本发明的一个实施方式中,存储模块控制器可包括用以在不要求将数据临时地存储在控制模块上的存储器中的情况下直接地将检索数据传输到客户端的功能。

[0064] 图6A-6B示出了根据本发明的一个或多个实施方式的示例。以下示例并不意图限制本发明的范围。

[0065] 转到图6A,考虑其中客户端600向包括控制模块604和至少一个存储模块614的存储设备发布包括逻辑地址的客户端读请求602的情形。控制模块604接收客户端读请求并使用在控制模块上的处理器606和存储在存储器608中的一个或多个存储器内数据结构610而使用图5A中所述的方法来生成控制器读请求612。更具体地,控制模块使用一个或多个存储器内数据结构610确定对应于逻辑地址的物理地址(PA)。此外,控制模块使用一个或多个存

储器内数据结构610和物理地址来确定存储在物理地址处的数据的产生时间。控制模块随后使用该产生时间和接收到客户端读请求的时间以便确定存储在物理地址处的数据的保持时间。控制模块然后使用以下索引在(一个或多个)存储器内数据结构中执行查找：<保持时间，P/E循环值，页号>，以便获得至少一个读取阈值，其中，从存储器内数据结构中的一个获得P/E循环值并从物理地址提取页号。

[0066] 在本示例中，假设固态存储器模块620、622包括MLC，并且上述查找以用于阈值B和阈值C(参见图6B)的位移值的形式返回读取阈值。控制模块604随后生成控制器读请求，其包括物理地址、用于阈值B的位移值和用于阈值C的位移值。随后向存储模块614发布控制器读请求612，其包括包含与所述物理地址相对应的物理位置的固态存储器模块620、622。

[0067] 存储模块614随后接收控制器读请求612并对其服务。更具体地，存储模块控制器612生成并向包括对应于所述物理地址的物理位置的固态存储器模块发布读命令618。在本示例中，使用默认读取阈值A值、非默认读取阈值B值和/或非默认阈值C值来生成读命令。使用默认阈值B值和用于阈值B的位移值来确定非默认阈值B值。此外，使用默认阈值C值和用于阈值C的位移值来确定非默认阈值C值。

[0068] 存储模块控制器随后从固态存储器模块接收数据并向客户端600提供数据(响应于624)。可将该数据直接地从存储模块中的存储器(未示出)拷贝到客户端存储器(未示出)。

[0069] 图6B示出了用于多层单元的电压的示例性分布。更具体地，图6B示出了阈值A、阈值B和阈值C相对于电压表示的逻辑位值的两个不同映射的相对位置。在本示例中，用于阈值B和阈值C的读取阈值相对于<保持时间，P/E循环，页号>而改变，而读取阈值对于阈值A而言未改变。

[0070] 可使用由存储设备中的一个或多个处理器执行的指令来实现本发明的一个或多个实施方式。此外，此类指令可对应于存储在一个或多个非临时计算机可读介质上的计算机可读指令。

[0071] 虽然已相对于有限数目的实施方式描述了本发明，但受益于本公开的本领域的技术人员将认识到可以设计不脱离如在这里公开的本发明的范围的其它实施方式。因此，应仅由所附权利要求来限制本发明的范围。

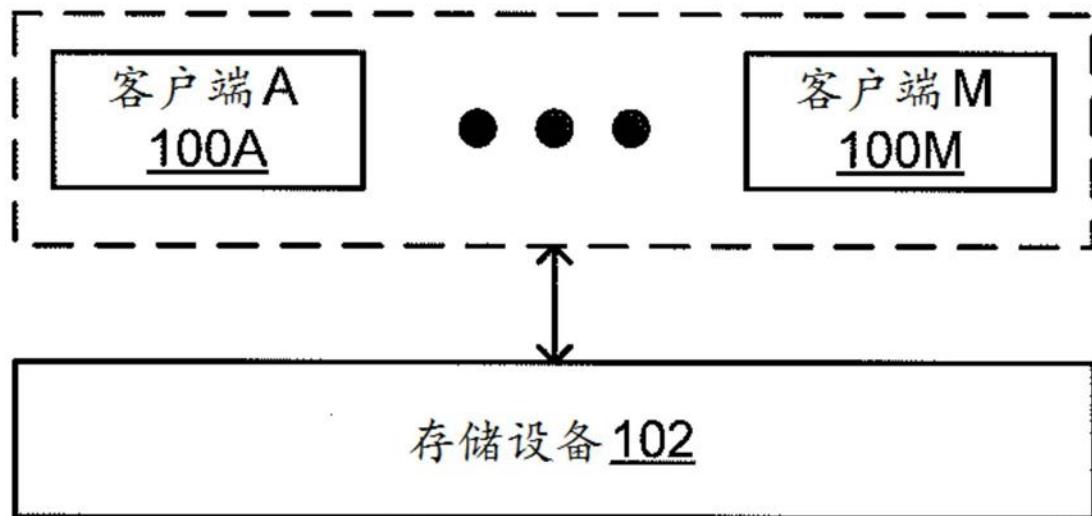


图1A

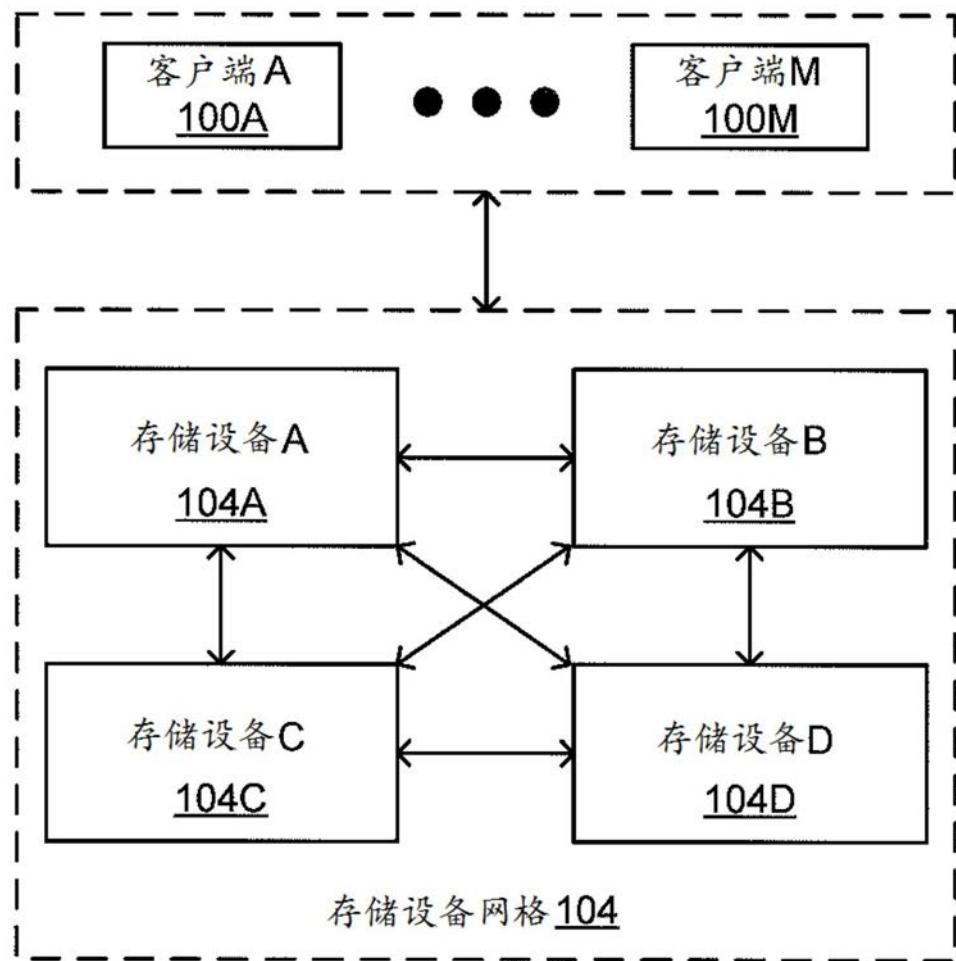


图1B

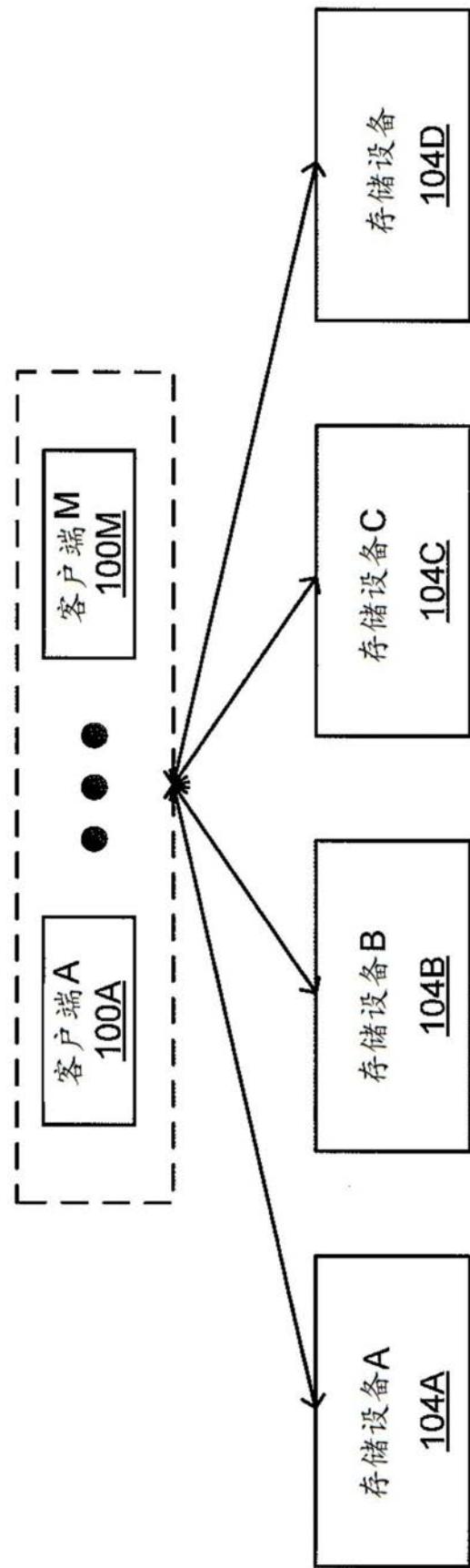


图1C

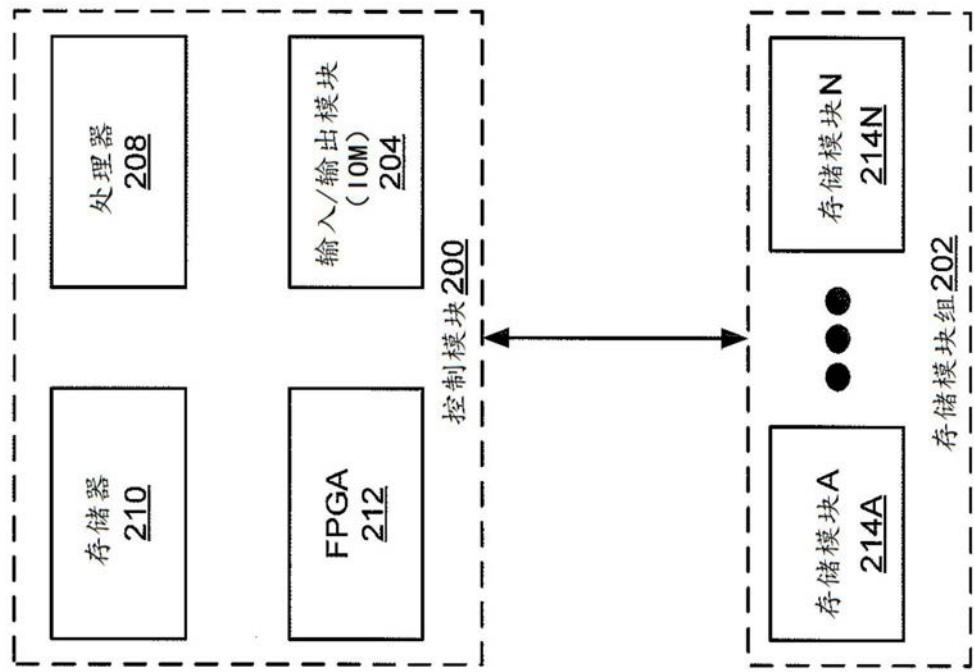


图2

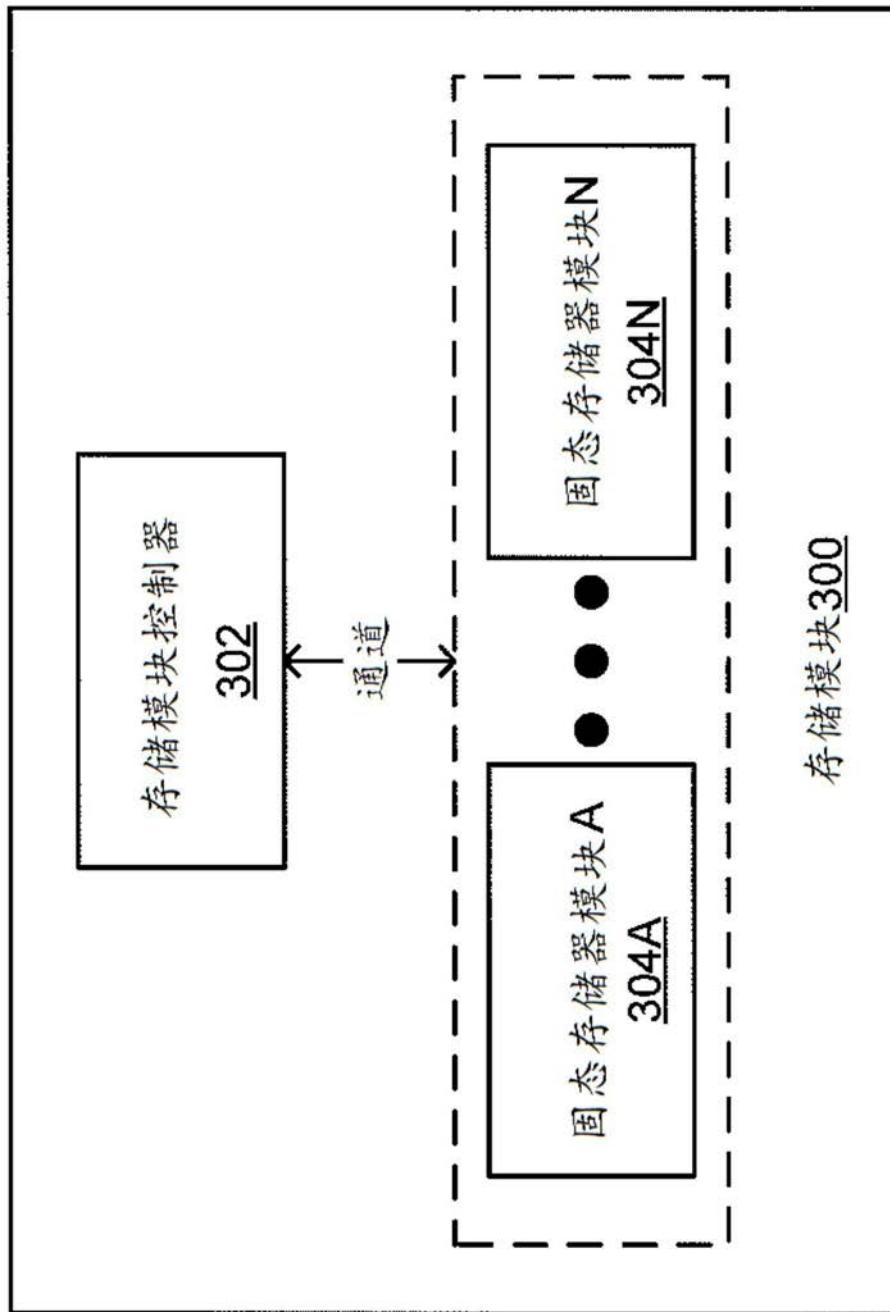


图3

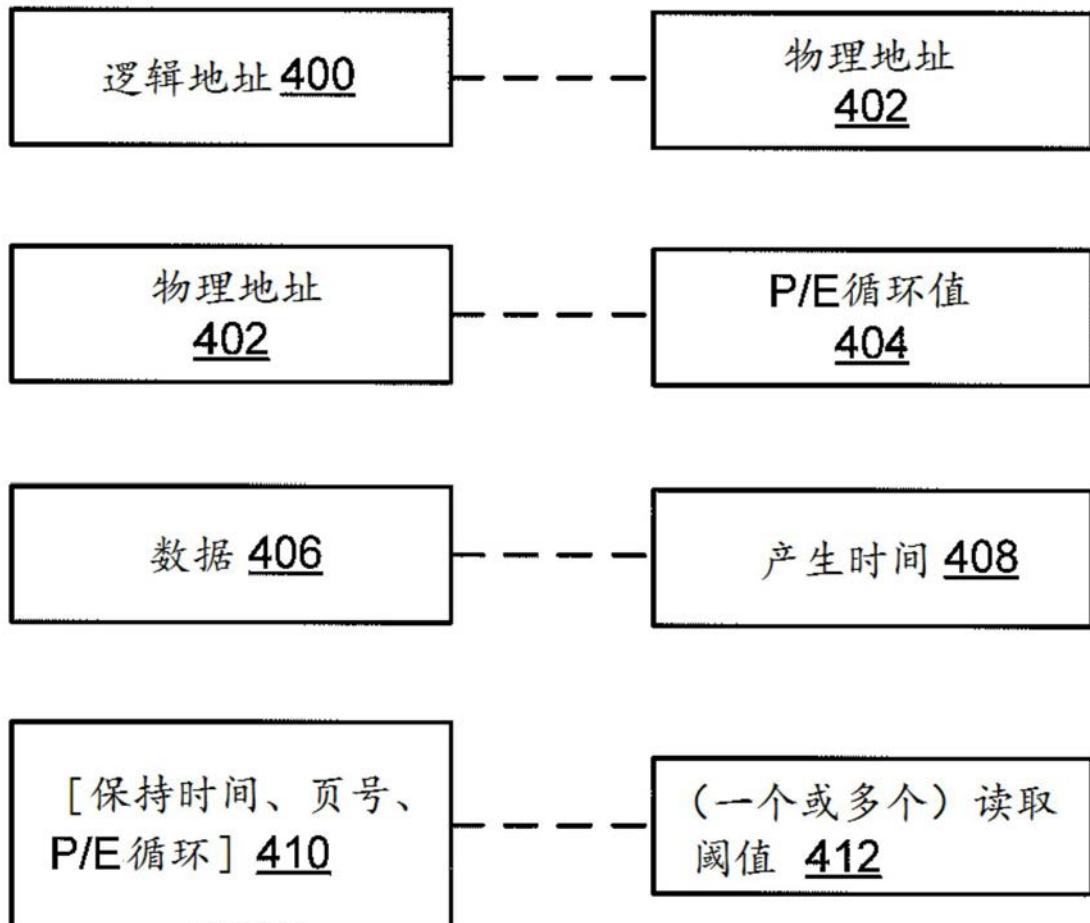


图4

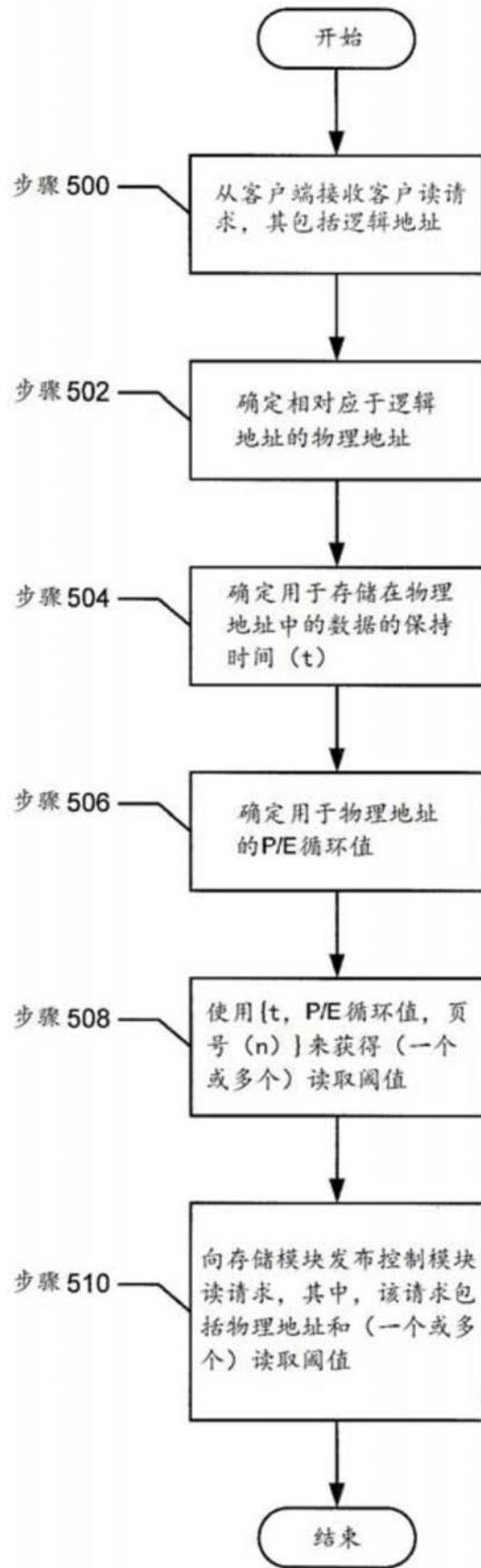


图5A

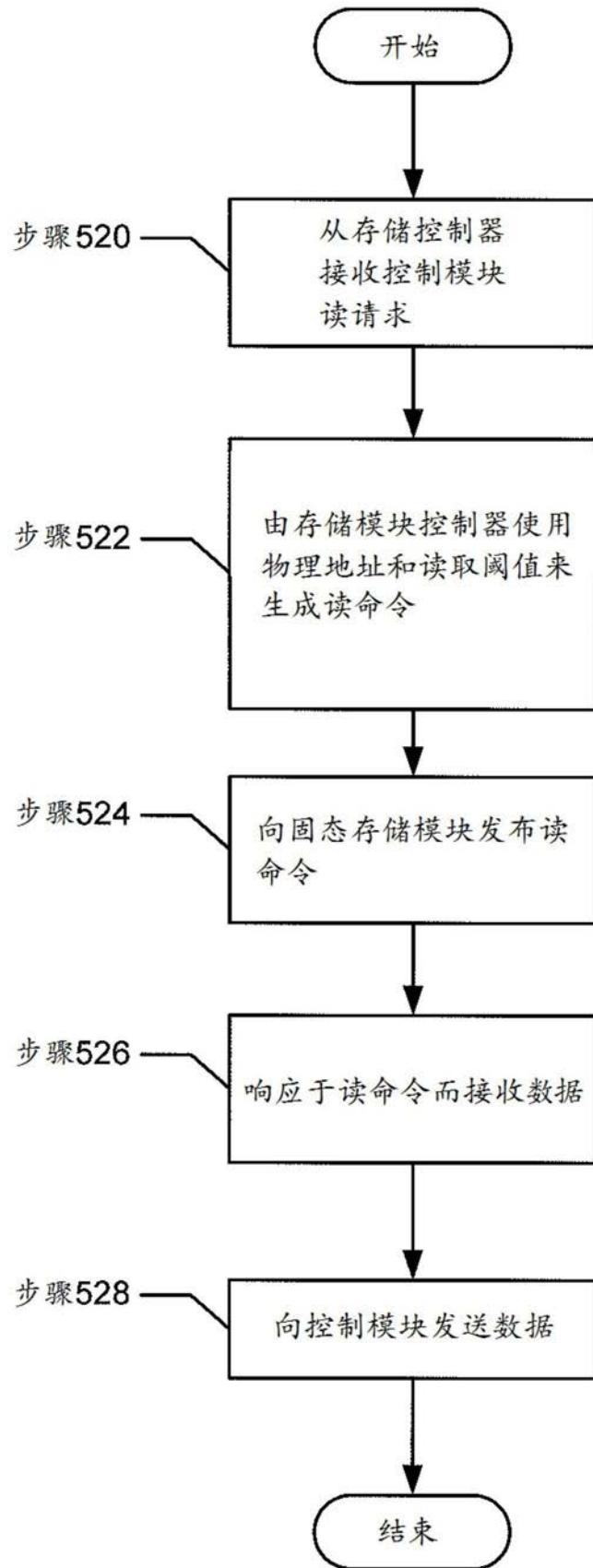


图5B

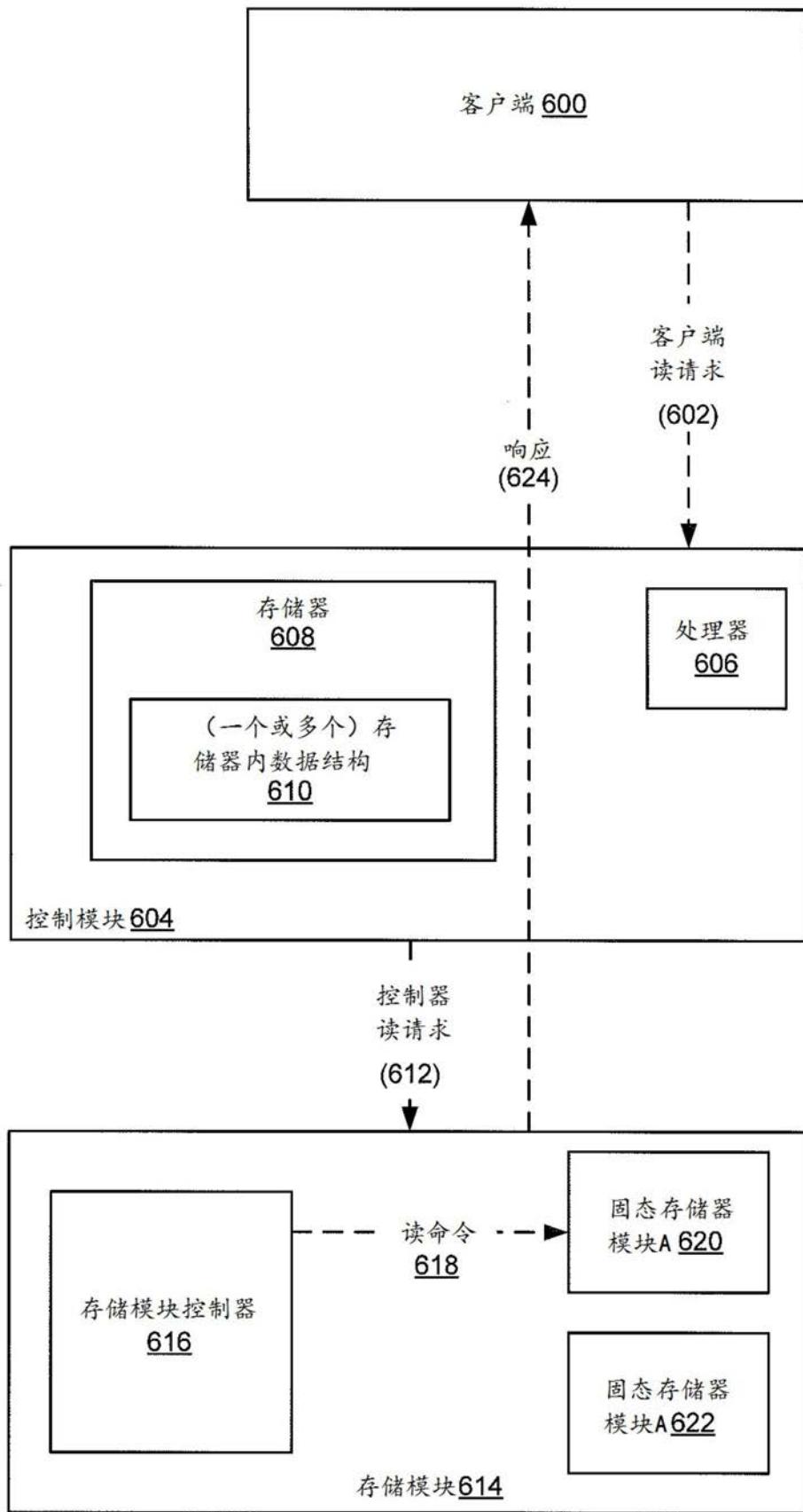


图6A

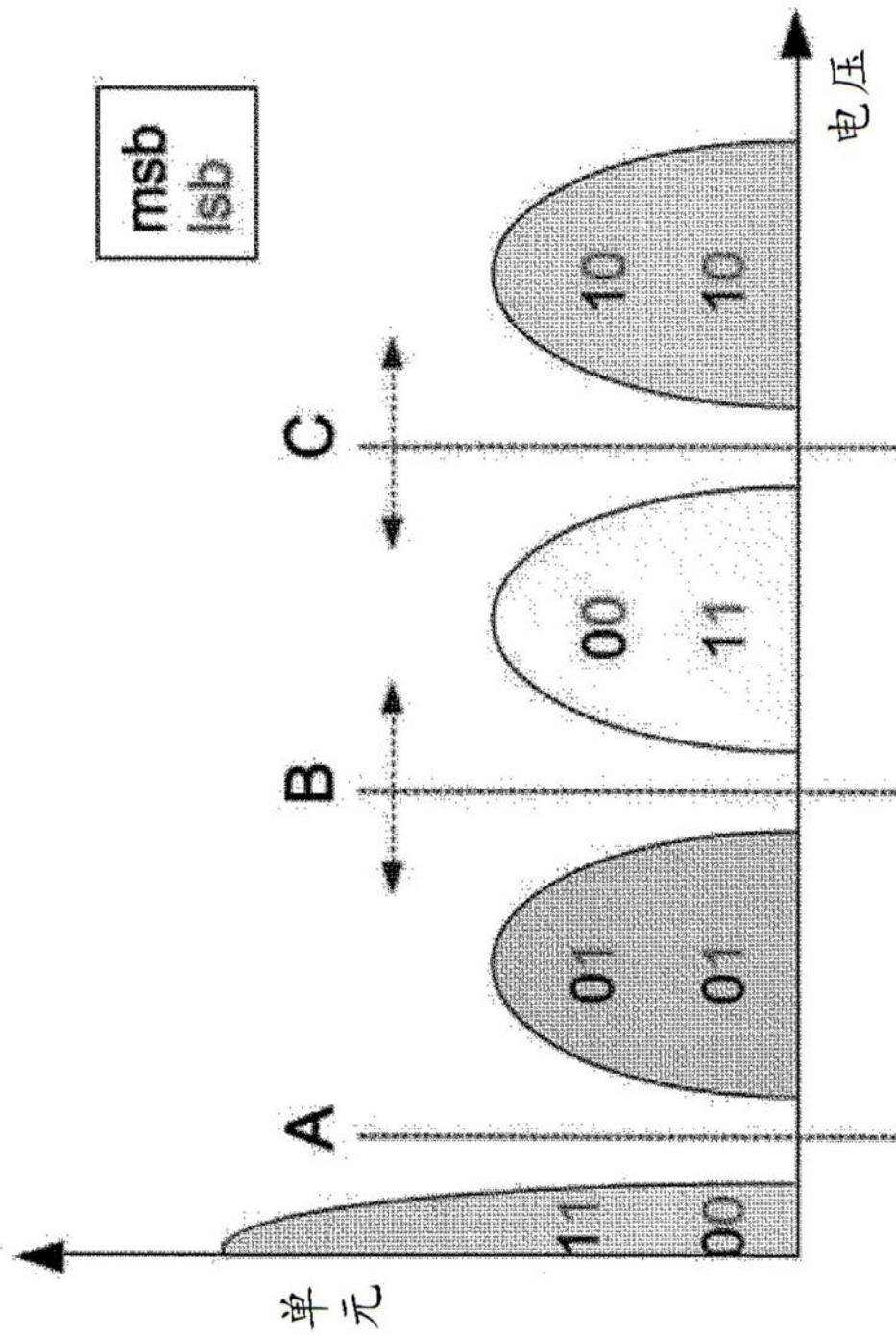


图6B