



(12) 发明专利

(10) 授权公告号 CN 106708424 B

(45) 授权公告日 2021.05.04

(21) 申请号 201610857593.2

(22) 申请日 2016.09.27

(65) 同一申请的已公布的文献号  
申请公布号 CN 106708424 A

(43) 申请公布日 2017.05.24

(30) 优先权数据  
14/941,517 2015.11.13 US

(73) 专利权人 三星电子株式会社  
地址 韩国京畿道水原市

(72) 发明人 安德鲁·汤姆林 贾斯丁·琼斯

(74) 专利代理机构 北京铭硕知识产权代理有限公司 11286

代理人 刘灿强 陈晓博

(51) Int.Cl.

G06F 3/06 (2006.01)

(56) 对比文件

CN 103988184 A, 2014.08.13

CN 105051672 A, 2015.11.11

汪飞.可扩展操作系统中存储管理子系统的研究与实现.《中国优秀硕士学位论文全文数据库 信息科技辑》.2006,

审查员 王晨霞

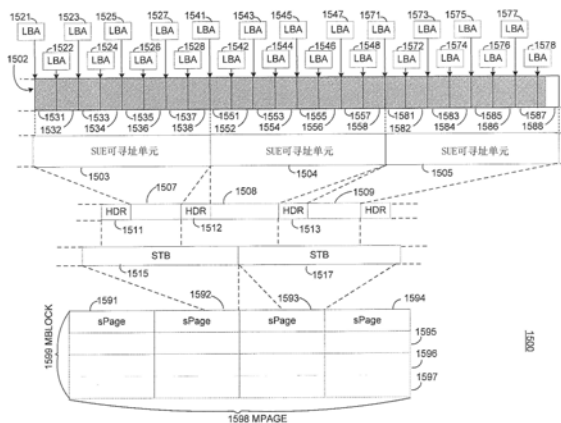
权利要求书3页 说明书22页 附图20页

(54) 发明名称

对用户数据执行选择性底层暴露映射的设备和方法

(57) 摘要

提供了一种用于将用户数据映射到选择性底层暴露(SUE)地址空间的设备,所述设备包括:存储器,存储计算机指令;处理器,执行指令以组合来自多个逻辑寻址的块的第一用户数据,以生成SUE页。SUE页与多个芯片中的每个上的各物理块的各物理页对应,对于所述多个芯片中的每个,存储单元的对应的物理块在存储装置中作为一个单元被共同管理。处理器还执行指令,以在存储装置的逻辑地址空间中存储使第一用户数据与SUE页关联的映射信息。



1. 一种用于将用户数据映射到选择性底层暴露地址空间的设备,所述设备包括:  
存储器,存储计算机指令;以及  
处理器,执行计算机指令以组合来自多个逻辑寻址的块的第一用户数据,以生成与多个芯片中的每个芯片上的各物理块的各物理页对应的选择性底层暴露页,并且在存储装置的逻辑地址空间中存储使第一用户数据与选择性底层暴露页关联的映射信息,其中,所述多个芯片中的每个包括在存储装置中作为一个单元共同管理的存储单元的对应的物理块,  
其中,选择性底层暴露地址空间包括底层物理地址空间的与管理活动关联的选择的维度。
2. 根据权利要求1所述的设备,其中,为了组合来自所述多个逻辑寻址的块的第一用户数据以生成选择性底层暴露页,处理器还执行计算机指令以:  
将逻辑寻址的块的组的第二用户数据集合在选择性底层暴露可寻址单元中;  
生成对应于选择性底层暴露可寻址单元的标头;  
将选择性底层暴露可寻址单元的至少一部分和标头放置在存储装置传输块中,  
其中,标头包括与用于促进回收过程或数据收集过程的第二用户数据关联的信息,所述多个逻辑寻址的块包括逻辑寻址的块的组,第一用户数据包括第二用户数据。
3. 根据权利要求2所述的设备,其中,为了组合来自所述多个逻辑寻址的块的第一用户数据以生成选择性底层暴露页,处理器还执行计算机指令以:  
对第二用户数据执行数据压缩算法;并  
将多个另外的选择性底层暴露可寻址单元和对应的标头放置在存储装置传输块中。
4. 根据权利要求3所述的设备,其中,为了组合来自所述多个逻辑寻址的块的第一用户数据以生成选择性底层暴露页,处理器还执行计算机指令以将整数数量的存储装置传输块结合在选择性底层暴露页中。
5. 一种用于将用户数据映射至选择性底层暴露地址空间中的方法,所述方法包括下述步骤:  
使用处理器组合来自多个逻辑寻址的块的第一用户数据,以生成与多个芯片中的每个芯片上的各物理块的各物理页对应的选择性底层暴露页,其中,所述多个芯片中的每个芯片包括在存储装置中作为一个单元共同管理的存储单元的对应的物理块;  
把使第一用户数据与选择性底层暴露页关联的映射信息存储在存储装置的逻辑地址空间中,  
其中,选择性底层暴露地址空间包括底层物理地址空间的与管理活动关联的选择的维度。
6. 根据权利要求5所述的方法,其中,组合来自所述多个逻辑寻址的块以生成选择性底层暴露页的步骤包括:  
将来自逻辑寻址的块的组的第二用户数据集合在选择性底层暴露可寻址单元中;  
生成对应于选择性底层暴露可寻址单元的标头;以及  
在存储装置传输块中放置选择性底层暴露可寻址单元的至少一部分和标头,  
其中,标头包括与用于促进回收过程或数据收集过程的第二用户数据关联的信息,所述多个逻辑寻址的块包括逻辑寻址的块的组,第一用户数据包括第二用户数据。
7. 根据权利要求6所述的方法,其中,组合来自所述多个逻辑寻址的块以生成选择性底

层暴露页的步骤包括：

对第二用户数据执行数据压缩算法。

8. 根据权利要求7所述的方法，其中，组合来自所述多个逻辑寻址的块的第一用户数据以生成选择性底层暴露页的步骤还包括：

在存储装置传输块中放置多个另外的选择性底层暴露可寻址单元和对应的标头。

9. 根据权利要求6所述的方法，其中，组合来自所述多个逻辑寻址的块的第一用户数据以生成选择性底层暴露页的步骤还包括：

包括源于存储装置传输块和连续的存储装置传输块中的单一逻辑寻址的块的第三用户数据。

10. 根据权利要求6所述的方法，其中，组合来自所述多个逻辑寻址的块的第一用户数据以生成选择性底层暴露页的步骤还包括：

包括源于选择性底层映射页和连续的选择性底层映射页中的单一逻辑寻址的块的第三用户数据。

11. 根据权利要求6所述的方法，其中，组合来自所述多个逻辑寻址的块的第一用户数据以生成选择性底层暴露页的步骤还包括：

在选择性底层暴露页中结合整数数量的存储装置传输块。

12. 根据权利要求5所述的方法，所述方法还包括从主机应用或虚拟机接收所述多个逻辑寻址的块，

其中，所述多个逻辑寻址的块中的每个包括与存储网络标准协议关联的大量数据。

13. 根据权利要求5所述的方法，所述方法还包括：

将选择性底层暴露页传送到存储装置，以存储在基于存储装置的用户区域的选择性底层暴露地址中；

发送与选择性底层暴露页关联的元数据至存储装置，以存储在存储装置的逻辑寻址的系统区域中。

14. 根据权利要求5所述的方法，所述方法还包括：将多个选择性底层暴露页与元页关联，其中，元页对应于所述多个芯片中的每个芯片上的各物理块的各物理页，所述多个芯片中的每个芯片包括在与存储系统关联的每个存储装置中作为一个单元共同管理的存储单元的对应的物理块，

其中，所述多个选择性底层暴露页构成元页。

15. 根据权利要求14所述的方法，所述方法还包括：将多个元页与元块关联，其中，元块对应于所述多个芯片中的每个芯片上的各物理块，所述多个芯片中的每个包括在与存储系统关联的每个存储装置中作为一个单元共同管理的存储单元的对应的物理块，

其中，所述多个元页中的每个对应于在所述多个芯片中的每个芯片上的各物理块的相关的物理页，所述多个元页构成元块。

16. 根据权利要求15所述的方法，其中，元块还与多个选择性底层暴露块关联，所述多个选择性底层暴露块中的每个对应于所述多个芯片中的每个芯片上的各物理块，其中，所述多个芯片中的每个包括在与存储系统关联的每个存储装置中作为一个单元共同管理的存储单元的对应的物理块。

17. 一种多模存储装置，所述多模存储装置包括：

存储单元的第一分区,组织在多个芯片上的物理块和物理页中,所述多个芯片包括作为由与选择性底层暴露页对应的用户数据编程的单元被共同管理的存储单元的对应的物理块;

存储单元的第二分区,由将选择性底层暴露页与多个逻辑寻址的块关联的第一映射来编程,

其中,每个物理块的每个物理页位于所述多个芯片的每个子集上,

其中,与选择性底层暴露页对应的选择性底层暴露地址空间包括底层物理地址空间的与管理活动关联的选择的维度。

18. 根据权利要求17所述的多模存储装置,其中,存储单元的第二分区还由将选择性底层暴露页与整数数量的存储装置传输块关联的第二映射编程,由将整数数量的存储装置传输块中的每个与多个选择性底层暴露可寻址单元以及对应的标头关联的第三映射编程,以及由将多个选择性底层暴露可寻址单元中的每个与逻辑寻址的块的组关联的第四映射编程,

其中,所述多个逻辑寻址的块包括逻辑寻址的块的组。

19. 根据权利要求18所述的多模存储装置,其中,标头包括与所述多个选择性底层暴露可寻址单元关联的信息,以促进回收过程或数据收集过程。

20. 根据权利要求18所述的多模存储装置,其中,整数数量的存储装置传输块和选择性底层暴露页包括压缩的数据。

## 对用户数据执行选择性底层暴露映射的设备和方法

### 技术领域

[0001] 本发明涉及信息存储。

### 背景技术

[0002] 许多电子技术(例如,数字计算机、计算器、音频装置、视频设备和/或电话系统等)在商业、科学、教育和娱乐的大多数领域内已经提高了生产率并且已经降低了成本。这些电子系统执行涉及信息存储系统的操作。信息存储操作进行的速度和容易度对信息存储系统的整体性能会有显著影响。然而,信息存储的传统尝试通常涉及速度与可管理的复杂度之间的反比关系。

[0003] 信息存储系统涉及可落入两种类别中的一种的操作。一种类别涉及与用户发起的活动关联的存储操作。另一种类别涉及由系统发起的管理和维护活动。这些操作进行的速度和容易度与用来存储信息的地址空间的类型关联。

[0004] 利用物理寻址空间的传统尝试理论上被认为以非常快的速度操作,但是在传统的物理寻址空间中对实际的管理和维护操作的尝试是非常复杂的并且没有被切实实施。传统的逻辑地址空间的管理和维护被认为比物理地址空间涉及更低的复杂度。然而,传统的逻辑地址空间无法以与物理地址空间同样快的速度操作。

[0005] 虽然传统的存储系统可以以可在先前已被认为是可接受的水平进行操作,但是这些存储系统越来越不足以满足用于改善应用和平台的需求以及长期以来切身的需要。实现提高的速度和可管理的复杂度两者以能够改进系统开发的传统尝试尚未成功。

### 发明内容

[0006] 根据本发明的一个实施例,一种用于将用户数据映射到选择性底层暴露(SUE)地址空间的设备包括:存储器,存储计算机指令;处理器,执行指令以组合来自多个逻辑寻址的块的第一用户数据以生成SUE页。SUE页与所述多个芯片中的每个上的各物理块的各物理页对应,对于所述多个芯片中的每个,存储单元的对应的物理块在存储装置中作为一个单元被共同管理。处理器还执行指令以在存储装置的逻辑地址空间中存储使第一用户数据与SUE页关联的映射信息。

[0007] 根据本发明的另一个实施例,用于将用户数据映射至SUE地址空间的计算机实施方法包括:组合来自多个逻辑寻址的块的第一用户数据,以生成SUE页。SUE页与多个芯片中的每个上的各物理块的各物理页对应,对于所述多个芯片中的每个,存储单元的对应的物理块在存储装置中作为一个单元被共同管理。所述方法还包括:把使第一用户数据与SUE页关联的映射信息存储在存储装置的逻辑地址空间中。

[0008] 根据本发明的另一实施例,一种多模存储装置包括:存储单元的第一分区,组织在多个芯片上的物理块和物理页中,所述多个芯片包括作为由与选择性底层暴露页对应的用户数据编程的单元被共同管理的存储单元的对应的物理块;所述装置还包括存储单元的第二分区,存储单元的第二分区由将SUE页与多个逻辑寻址的存储块关联的第一映射来编程。

[0009] 结合附图和下面的描述阐述本发明的一个或更多个实施例的细节。通过描述和附图以及通过权利要求,本发明的其他特征、目的和优点将清楚。

## 附图说明

[0010] 包括被并入本说明书并形成本说明书的一部分的附图用以对本发明的原理的示例性说明,并且不意图将本发明限制于其中所示的具体实施方式。除非另外明确指出,否则附图不按比例绘制。

[0011] 图1是示出根据一个实施例的具有SUE存储分区的示例性存储装置的框图。

[0012] 图2是示出根据一个实施例的示例性多模存储装置的框图。

[0013] 图3是示出根据一个实施例的另一个示例性多模存储装置的框图。

[0014] 图4是示出根据一个实施例的示例性多模固态驱动器 (MM-SSD) 的框图。

[0015] 图5是示出根据一个实施例的将地址空间信息转换为逻辑地址空间信息的示例性进程的框图。

[0016] 图6是示出根据一个实施例的示例性系统的框图。

[0017] 图7是示出根据一个实施例的系统的框图。

[0018] 图8是描述根据一个实施例的驱动多模底层暴露的方法的流程图。

[0019] 图9是示出示例性多模SSD装置的框图。

[0020] 图10是描绘根据本发明的实施例的在多模存储装置的用户区域中用于存储的示例性SUE块和对应的SUE页的框图。

[0021] 图11是描绘根据本发明的实施例的在多模存储装置的用户区域中用于存储的用户存储空间的示例性SUE块和对应的SUE页的框图。

[0022] 图12是描绘根据本发明的实施例的在多模存储装置的用户区域中用于存储的示例性SUE元页 (metapage) 和对应的SUE页的框图。

[0023] 图13是描绘根据本发明的实施例的在多模存储装置的用户区域中用于存储的示例性SUE元块 (metablock) 和对应的SUE元页的框图。

[0024] 图14是描绘根据本发明的实施例的在多模存储装置的用户区域中用于存储的另一示例性SUE元块和对应的SUE块的框图。

[0025] 图15是描绘根据本发明的实施例的可由多模存储系统实施以提供从逻辑地址到SUE地址的地址映射的示例性SUE映射方案的框图。

[0026] 图16是描绘可实施图15的SUE映射方案的示例性存储系统的示意图。

[0027] 图17是表示根据本发明的实施例的将逻辑地址空间映射到SUE地址空间的示例性方法的流程图。

[0028] 图18是示出根据本发明的实施例的采用SUE寻址方案以允许存储系统在存储装置中寻址逻辑和SUE存储空间的示例性多模存储管理系统的示意图。

[0029] 图19是示出根据本发明的实施例的采用SUE寻址方案以允许存储系统在存储装置中寻址逻辑和SUE存储空间的另一示例性多模存储管理系统的示意图。

[0030] 图20是示出根据本发明的实施例的采用SUE寻址方案以允许存储系统在存储装置中寻址逻辑和SUE存储空间的又一示例性多模存储管理系统的示意图。

[0031] 图21是示出根据本发明的实施例的可由多模存储管理系统实施的用户区域访问

管理器(UAAM)的示意图。

[0032] 图22是示出根据本发明的实施例的可由多模存储管理系统实施的用户区域映射引擎(UAME)的示意图。

[0033] 图23是示出根据本发明的实施例的可由多模存储管理系统实施的元块管理器(MBM)的示意图。

[0034] 图24是示出根据本发明的实施例的可由多模存储管理系统实施的存储装置控制管理器(SCM)的示意图。

[0035] 图25是示出根据本发明的实施例的可由多模存储管理系统实施的存储装置访问管理器(SAM)的示意图。

[0036] 图26是示出根据本发明的实施例的可由多模存储管理系统实施的全局状态管理器(GSM)的示意图。

### 具体实施方式

[0037] 提供所有上述特征和下面的描述以帮助本公开的理解。即,本公开不应该被解释为限制于这里阐述的“示例”实施例,并且可以以不同形式实施。下面的实施例仅是用于充分公开本发明的示例,并被描述以将本发明构思提供给本领域技术人员。因此,如果存在实施本公开的组件的许多方式,那么将被理解的是,可用任意一种方式或者它的任意一种相似事物来实施本公开。

[0038] 在本公开中,如果提到任何配置包括/包含特定组件或者任何过程包括/包含特定操作,那么这意味着可进一步包括其他组件或者其他操作。即,这里使用的术语仅意图描述具体的示例实施例,不意图限制本公开的范围。此外,描述的示例包括它们的补充实施例,其中,提供所述描述的示例以帮助对本公开的理解。

[0039] 这里使用的术语具有本领域技术人员通常理解的含义。通用术语将被解释为具有与本公开的上下文中一致的含义。另外,除非这里明确地如此定义,否则不应该以过于理想或正式的意义来解释这里使用的术语。

[0040] 描述了可采用多种不同类型的地址空间和地址空间活动的快速有效的多模存储方法。在一些实施例中,多模选择性底层暴露(multimode selective underlying exposure, SUE)存储装置能够使存储装置的一些底层方面选择性暴露而不暴露其他底层方面。多模存储和SUE方法可在限制可管理的范围的复杂度的同时改善性能。

[0041] 在一些示例性实施方式中,可选择性地暴露物理地址空间的底层方面。可实施整体存储分层方法,一个分层级别(hierarchical level)的底层方面可选择性地被暴露给另一分层级别。可通过地址空间配置和地址空间之间的映射发生选择性暴露。选择性地暴露的底层方面可更快速有效地实施在与具有被暴露的底层方面的分层级别不同的分层级别执行的各种活动。各种活动可包括存储管理操作。应理解的是,多模存储和SUE方法可包括各种配置和实施方式。

[0042] 在下文中,将参照附图在下面描述本公开的示例实施例。

[0043] I、多模存储装置

[0044] 图1是示出根据一个实施例的包括选择性底层暴露(SUE)存储分区101的示例性存储装置100的框图。SUE存储分区101可包括选择性底层暴露(SUE)接口102和底层存储区域

103。

[0045] 底层存储区域103可存储信息，SUE接口102可将底层存储区域103自身(例如涉及维度、典型几何、管理功能、写入操作、擦除操作等)的方面(例如，特性、特征、功能等)选择性地暴露给外部组件或存储系统分层级别(未示出)。该暴露可与存储在底层存储区域103中的信息(用户数据和元数据)的方面关联。SUE存储分区101可暴露底层方面的一部分(例如，特性、特征、功能等)。

[0046] 在暴露底层方面的一部分的一个示例性实施方式中，与未选择性地暴露底层方面的一部分的系统相比，可更有效地(例如，更快、更小的带宽以及更小的功耗)执行与暴露的方面关联的活动(例如，自由空间管理、用于自由空间使用的回收与调节、预留空间(over-provisioning)、修剪(trim)操作、功率循环等)。与暴露更多或全部底层方面的方法相比，该活动可以以更小的复杂性来执行。

[0047] 在一些实施例中，可基于速度对复杂度的比较或平衡来确定选择底层方面的哪部分被暴露。理解的是，SUE存储分区101可包括在具有单一分区的单模存储装置中或者SUE存储分区101可包括在具有多个分区的多模存储装置中。

[0048] 图2是示出根据一个实施例的示例性多模存储装置220的框图。存储装置220可包括第一分区230和第二分区240。

[0049] 理解的是，多模式和对应的分区可与各种事物关联或可基于各种事物。各种事物可包括底层存储区域的不同暴露、不同地址空间(例如，逻辑、虚拟或物理)、不同存储管理模式(例如，内部管理和外部管理)、不同底层存储信息(例如，元数据和用户数据)等。内部管理和外部管理可包括存储装置管理系统组件和操作(例如，闪存管理系统(FMS)、固态装置管理系统等)。分区和对应的组件也可有不同的类型。

[0050] 在多模存储装置220中的分区和对应的接口可与不同类型的地址空间(例如，逻辑地址空间和选择性底层暴露(SUE)地址空间)关联。在多模存储装置220中的不止一个分区和对应的接口也可与相同类型的地址空间关联，(例如，在多模存储装置220中的不止一个分区和对应的接口可与SUE地址空间关联)。

[0051] 第一分区230可包括第一类型的接口231和底层存储区域233。第二分区240可包括第二类型的接口241和底层存储区域243。在一些实施例中，第一分区230可为第一类型的地址空间分区(例如，逻辑地址空间分区)，第二分区240可为第二类型的地址空间分区(例如，SUE地址空间和虚拟地址空间分区)。理解的是，分区可以是SUE存储分区。

[0052] 图3是示出根据一个实施例的另一示例性多模存储装置350的框图。存储装置350可包括第一分区370和第二分区380。

[0053] 在一些实施例中，第一分区370可为第一类型的地址空间分区，第二分区380可为SUE地址空间分区。第一分区370可包括第一类型的接口371和底层存储区域373。第二分区380可包括SUE接口381和底层存储区域383。理解的是，诸如第一分区相关的活动372(例如，FMS)的一些活动可针对一个分区内部地(例如，在存储装置350中)执行，并可针对另一个分区外部地(未示出)执行。

[0054] 不同类型的信息可存储在不同的分区中。在一些实施例中，可存在两种类型的信息，例如，元数据和用户数据。用户数据可主要由用户应用生成，元数据可以是与用户数据关联的主要辅助信息(例如，在存储系统分层体系中的文件的位置、文件中的内容的大小、



访问时间、修改时间、用户ID等)。第一闪存管理系统可专注于管理元数据。元数据反过来可用于管理用户数据的存储。

[0055] 理解的是,存储系统可指示或实施与用户发起的活动关联的操作,此操作不同于与管理或维护活动关联的系统操作。例如,从用户角度,用户发起的读取或写入可指向具体的地址或位置,而从系统角度,系统操作可指向物理块和页。

[0056] 还理解的是,存储装置350可包括各种配置和实施方式。在一些实施例中,存储装置350可为固态装置。存储装置350可包括闪存组件(例如,NAND类型闪存组件、NOR类型闪存组件等)。

[0057] 图4是示出根据一个实施例的示例性多模固态驱动器(MM-SSD)的框图。多模固态驱动器(SSD)400可为多模存储装置的一个示例性实施方式。多模固态驱动器400可包括逻辑地址空间分区410、可包括闪存转换逻辑FTL413的逻辑接口411、底层物理地址空间412、SUE地址空间分区420、SUE接口421以及底层物理地址空间423。

[0058] 逻辑地址空间分区410可接收并存储逻辑寻址的系统数据(例如,元数据),SUE地址空间分区420可接收根据底层暴露地址空间寻址的用户数据(例如,应用数据)。用户数据可存储在可包括闪存组件(例如,不同类型的浮置栅极晶体管)的底层物理地址空间423中。闪存组件可以以各种配置和粒度(granularity)布置。例如,闪存组件可布置为多个芯片,多个芯片中的芯片470可包括块473和479及块内的页。

[0059] 在一些实施例中,SUE接口421可暴露底层物理地址空间423的一方面。可通过协调用户数据寻址与MM-SSD 400的底层操作来暴露底层物理地址空间423的选择性方面。所述协调可与底层物理地址空间423的管理操作的暴露关联。底层物理存储管理方面可包括被一起管理(例如,在单个操作中、作为单个管理单元、在块集中、在带(bank)中、以及响应于单个管理命令)的多个底层物理地址块(例如,471、472、473和474)的组。

[0060] 图5是示出根据一个实施例的将地址空间信息转换为逻辑地址空间信息的示例性进程的框图。SUE地址块503可包括与各种管理和维护操作关联的信息(例如,505、507和508)。物理地址空间502可包括多个芯片(511、512、513、514、521、522、523、524、531、532、533、534、541、542、543和544)。每个芯片可包括多个物理寻址块(例如,515和519),并且每个物理寻址块可包括多个物理地址页。

[0061] 物理地址空间502可基于物理块和物理页访问地址存储位置。SUE接口501可接收SUE地址块503信息,并可将接收的信息转换为或重新配置为与物理地址空间502兼容的配置。SUE地址块503信息可对应于涉及物理管理操作的信息。

[0062] 在一些实施例中,管理和维护操作可指向物理地址空间502中的物理块(例如,物理块515、519和539)。管理操作可指向物理地址空间或物理层面管理单元。物理层面管理单元可包括管理基本同时管理(例如,响应于管理操作或命令)的多个地址、页、块等。例如,擦除操作可指向来自每个芯片的物理块(类似于块515的以黑色示出的物理块)。

[0063] 由于SUE地址块503配置为与物理块匹配,因此用于每个对应的物理块的每段信息(例如,505、507和508)可包括在SUE地址块503中。在一些示例性实施方式中,SUE接口501可接收SUE地址块503信息,可将信息505、507和508识别为分别对应于物理块515、517和528,并可由此执行对应的管理和维护操作。在一些实施例中,可对多个物理块中的信息执行擦除管理操作,并可对页中的信息执行写入操作。

[0064] 两个地址空间的几何结构也可不同。在一些实施例中，逻辑地址空间可以是一维的（例如，逻辑块地址（LBA）偏移如何对齐），物理地址空间可以是多维的，其包括诸如纠错码（ECC）、物理页、物理块、物理芯片等（包括它们中的一些或子集）的各种方面。SUE地址空间可以是一维的或者有限数量或减少数量的维度。在SUE地址空间的一些示例性实施方式中，底层物理地址空间的维度可被概括为单维度或减少数量的维度。底层物理地址空间的与管理活动（例如，回收/垃圾收集、功率循环等）关联的选择的维度（例如，块和页）可被概括为SUE地址空间，但底层物理地址空间的其他方面或活动（例如，ECC）可不被概括为SUE地址空间。

[0065] 理解的是，底层方面的选择性暴露可包括在由整个系统中的其他组件（未示出）而非MM-SSD 400执行的将用户数据传送到MM-SSD 400之前的协调。在一些实施例中，MM-SSD 400可结合到在整个系统分层体系的不同级操作的管理组件。

[0066] 图6是示出根据一个实施例的系统600的框图。系统600可包括通信地结合至多模存储管理系统610的多个MM-SSD（例如，620、630、640和650）。

[0067] 理解的是，一些活动（例如，一些存储管理操作和闪存管理系统操作）可由多模存储管理系统610控制，其他活动（例如，其他存储管理操作和闪存管理系统操作）可由MM-SSD 620、630、640和650控制。在一些实施例中，MM-SSD 620、630、640和650可以分别包括可控制或指向针对MM-SSD 620、630、640和650的一些活动的控制器621、631、641和651，同时多模存储管理系统610可以包括可控制或指向针对MM-SSD 620、630、640和650的一些活动的控制器611。

[0068] 在一些示例性实施方式中，控制器621、631、641和651可分别控制或指向MM-SSD 620、630、640和650中的第一分区的活动，控制器611可控制或指向MM-SSD 620、630、640和650中的第二分区的活动。控制器611可经由选择性底层暴露接口来控制MM-SSD 620、630、640和650中的活动。

[0069] 在一些实施例中，系统600可包括多个卷（volume）（例如，671、672和673）。在一些示例性实施方式中，系统600可包括用户空间并且用户空间可映射到多个卷中，存储空间可以作为多个卷呈现给用户。理解的是，卷可具有不同的大小。还理解的是，不同大小的SUE可寻址单元可与多个卷关联。

[0070] 图7是示出根据一个实施例的系统700的框图。系统700可包括通信地结合到包括在设备710中的多模存储管理系统720的多模SSD（MM-SSD）750。设备710可为各种计算机/电子设备（装置）中的一种。理解的是，其他多模SSD可结合到多模存储管理系统720。系统700可管理元数据730和用户数据740的存储。

[0071] 多模存储管理系统720可包括控制器745。控制器745可包括闪存管理系统741（用于用户数据）和SUE映射器742。

[0072] 多模SSD 750可包括逻辑地址空间分区770和SUE地址空间分区780。逻辑地址空间分区770可以包括物理地址空间777和可包括闪存管理系统771（用于元数据）的控制器775。闪存管理系统771可包括逻辑接口772，逻辑接口772可包括FTL 773。物理地址空间777可包括NAND闪存。SUE地址空间分区780可以包括SUE接口782和可包括NAND闪存的物理地址空间787。

[0073] 可在逻辑地址块791中接收元数据730信息，在逻辑地址块792中转发元数据730信

息,并且将元数据730信息从多模管理系统720传送至逻辑地址空间分区770。理解的是,逻辑地址块791和792可相同(例如,可不改变逻辑地址块791并被简单地转发到逻辑地址空间分区770)。

[0074] 逻辑接口772可将与元数据关联的逻辑块地址(LBA)转换成与物理地址空间777关联的物理地址块793。FMS 771可指示与物理地址空间777关联的存储管理和维护操作。元数据可存储在物理地址空间777的NAND闪存中。

[0075] 逻辑地址块797中的用户数据可转发至FMS 741。由于物理地址空间787的底层特征和特性经由SUE接口782暴露,所以FMS 741可指示与物理地址空间787的底层特征和特性关联的闪存管理和维护操作。SUE映射器742可将逻辑地址块797映射到SUE地址块798。

[0076] SUE地址块798可转而由选择性底层暴露(SUE)接口782转换成物理地址块799(例如,与图5中的517和519类似),物理地址块799与包括在物理地址空间787中的NAND闪存组件关联。理解的是,逻辑地址块可具有与SUE地址块不同的大小,SUE地址块转而可具有与物理地址块不同的大小。

[0077] 相较于传统的方法,在上述的分层体系级别中执行的各种活动促进更有效的和更方便的管理。传统的方法在其处理影响多个层级的活动时的灵活性上受到限制。一些传统的方法必须对多个层级执行活动,导致对整体性能的指数级不利影响(例如,登陆日志(log-on-log)、在驱动级的FMS和在系统级的FMS)。

[0078] 例如,在raid(独立磁盘冗余阵列)存储系统中,存在对上存储分层体系级别(例如,raid系统管理级)和下层存储分层体系级别(例如,存储驱动级)均具有影响的需要一起管理的许多项(例如,数据存储和对应的奇偶校验存储)。对于每个级,信息的生命周期可不同(例如,用户可能想要重新写入信息但是raid系统可仍需要其以用于奇偶校验计算),导致驱动级为用户写入“新”数据,但系统级仍然为raid系统保留“旧”信息。在没有能力进行修剪的情况下,这导致写入放大因子(write amplification factor)。

[0079] 图8是描述根据一个实施例的驱动多模选择性底层暴露(MM-SUE)的方法的流程图。

[0080] 具有7%的预留空间的驱动器(例如,SSD)的工作效率是不具有驱动预留空间的直接重写系统(例如,HDD)的15倍,并且还是不具有系统预留空间的系统的15倍,总计为225(15×15)倍。允许FMS向上移动到上存储分层体系级别的多模存储装置有助于减少回落(back down)(例如,7%的预留空间工作在没有预留空间的系统的仅15倍工作效率范围内,28%的预留空间工作在没有预留空间的系统的仅3倍工作效率范围内),导致写入放大因子的减小。在一些示例性实施方式中,用于指示来自上存储分层体系级别的管理操作的选择的底层地址块和页可与底层物理级协调或匹配。这允许在用户级的信息的信息的生命周期变得与在系统级的信息的信息的生命周期不同,但是从管理立场看,两者的生命周期可被对齐为相同的,例如,在用户级的生命周期的持续时间和在系统级的生命周期的持续时间可与用户空间的使用和擦除对应。

[0081] 在操作810中,装置的第一部分可被配置或设计为用于存储第一类型的信息的第一区域。在一些实施例中,第一区域可为元数据区域并且第一类型的信息可为元数据。可改变纠错码(ECC)的大小。

[0082] 在操作820中,可在第一地址空间信息的基础上执行第一类型接口操作。在一些示

例性实施方式中,第一区域可为元数据区域并且第一类型的信息可为元数据。在一些实施例中,第一类型接口可为逻辑地址空间接口,并可基于逻辑寻址的信息来执行操作。逻辑接口操作可包括闪存转换逻辑 (FTL),闪存转换逻辑包括:接收元数据和逻辑地址;将在系统级配置可见的地址块转换为在物理级配置的地址块。

[0083] 在操作830中,装置的第二部分配置或设计为用于存储第二类型的信息的第二区域。在一些实施例中,第二区域可为用户数据区域,第二类型的信息可为用户数据。SUE地址空间可提取或去除与物理地址空间关联的复杂度,同时仍暴露与底层物理地址空间配置的关系或对应。在一些示例性实施方式中,物理空间维度可被提取为SUE地址页维度和SUE地址块维度。物理地址空间可由SUE地址提取。

[0084] 在操作840中,可基于第二地址空间信息执行第二类型接口操作,其中,第二类型接口可选择性地暴露底层方面。第二地址空间信息可为选择性底层暴露 (SUE) 地址空间信息,其中,SUE地址空间信息可对应于底层方面。底层方面可包括物理地址空间几何的典型几何或维度。SUE接口可暴露与底层系统管理操作关联的维度(例如,自由空间管理、自由空间使用的回收与调节)。在元数据区域中预留空间的百分比可不同于在用户数据区域中的预留空间的百分比。

[0085] 图9是示出根据一个实施例的示例性多模SSD 920的框图。在图9中,多模SSD 920与针对逻辑寻址的SSD 910和物理寻址的SSD 930的传统尝试进行比较。

[0086] 逻辑寻址的SSD 910可包括逻辑接口911、FTL 912和逻辑地址空间913。物理寻址的SSD 930可包括物理接口931和物理地址空间932。多模SSD 920可包括逻辑接口921、FTL 922、逻辑地址空间923、SUE接口924和物理地址空间925。

[0087] 多模SSD 920促进驱动器的底层方面的选择性暴露和便利性。有别于不是暴露过多就是过于复杂的传统方法,多模SSD 920允许在没有过度的复杂性的情况下的合适的量的暴露。

[0088] 然而实际上传统的SSD通常不是良好的线性地址空间,相反它们通常具有带有一束闪存芯片的控制器,闪存芯片具有配置为以由页组成的块操作的芯片,页具有将存储在晶体管的组或串中的数据。物理寻址的SSD 930试图暴露存储介质的全部底层物理地址空间方面,其允许被认为是非常快速的操作(例如,相比于逻辑寻址的SSD 910)但是产生非常复杂的方法。逻辑寻址的SSD 910具有被认为是具有隐藏存储介质的方面的全部或几乎全部底层细节的方案单线性平面映射空间(single linear flat mapping space),然而最终试图在具有隐藏的多个底层细节的物理区域中存储数据拖慢了系统(例如,与物理寻址的SSD 930相比)。

[0089] 多模SSD 920促进FMS操作的方便且灵活的配置和实施。多模SSD 920可主要在多模SSD 920的内部控制器(例如,图7的控制器775)中针对逻辑地址空间923执行FMS操作,同时可主要在多模SSD 920的外部的控制器(例如,图7的控制器745)中在系统级执行针对SUE地址空间925的FMS操作。

[0090] 这种分离或划分多模SSD 920的FMS操作的能力不同于不允许FMS操作的分离或划分的SSD 910和SSD 930的FMS操作方法。针对逻辑寻址的SSD 910的FMS操作主要在逻辑寻址的SSD 910的内部控制器中执行,而针对物理寻址的SSD 930的FMS操作主要在物理寻址的SSD 930的外部的控制器中在系统级执行。

[0091] 在一些实施例中,多模SSD 920可选择性地暴露一些底层地址空间特征,逻辑寻址的SSD 910和物理寻址的SSD 930不促进一些底层地址空间特征而非其他底层地址空间特征的选择性暴露。在一些实施例中,将底层方面暴露给外部FMS(例如,多模存储管理系统720)可涉及底层方面的选择的暴露的映射。

[0092] II、选择性底层暴露(SUE)映射

[0093] 本发明的另一实施例实施选择性底层暴露(SUE)映射方案,以针对存储系统中的用户数据生成从逻辑地址空间到SUE地址空间的映射。SUE映射方案可选择性地暴露底层物理存储介质的显著特征,以允许在多个存储装置的系统级而不是在单个存储装置级执行某些存储介质管理功能。

[0094] 例如,一些实施例可使存储设备中的多个NAND闪存非易失性存储装置的用户地址空间的方面的选择性暴露。SUE映射方案的SUE页和块可与在物理NAND闪存非易失性存储装置中的每个中作为单元共同地管理的对应的物理页及块对齐。物理NAND闪存非易失性存储装置中的单个芯片在SUE映射方案中不被区分,但是尽管如此,也间接地反映在SUE块大小中。

[0095] 存储装置的物理页和块与SUE映射方案的SUE页和块之间的关联性允许诸如擦除、编程、回收(垃圾收集)、自由空间管理等某些NAND闪存管理功能在存储系统中的全部NAND闪存非易失性存储装置的系统级被协调和实施。某些存储介质管理功能的系统级实施方式可提供关于存储资源配置的有利的效率。

[0096] 再次参照图3,诸如NAND闪存非易失性存储装置的多模存储装置(例如,350、400、620)可结合本公开中描述的SUE映射方案实施。例如,在一些实施中,多模存储装置可为基于NAND闪存的固态驱动器(SSD)。在一些实施例中,多模存储装置可遵照诸如标准磁盘驱动器形状因子或标准存储卡形状因子的标准化的物理形状因子。

[0097] 再次参照图5,如上面描述的,多模存储装置可包括具有许多NAND闪存存储单元的多个芯片511、512、513、514、521、522、523、524、531、532、533、534、541、542、543、544或存储芯片。每个芯片上的NAND闪存存储单元可被细分为存储单元的诸如物理块515、517、519、528和539的多个离散的物理块。

[0098] 可针对多模存储装置上的一个或多个离散组的芯片上的存储单元的块来执行自由空间的擦除和管理。例如,多模存储装置可包括一百二十八个芯片,并且可针对作为一个组或一个单元的来自一百二十八个芯片中的每个芯片的一个块进行擦除并管理自由空间。可选择地,多模存储装置可包括一百二十八个芯片,并且可针对作为一个组的来自芯片的子集(例如,三十二个芯片的组)的一个块进行擦除并管理自由空间。

[0099] 图10是描绘根据本发明的实施例的在多模存储装置的用户区域中用于存储的示例性SUE块和对应的SUE页的框图。

[0100] 现在参照图10,示出了形成SUE块1010的一系列物理块1012、1014、1016和1018。存储单元的每个物理块1012、1014、1016和1018还被细分为存储单元的多个离散的物理页(例如,1021、1022、1023和1024)。SUE页1030包括来自对应的SUE块1010中的每个物理块1012、1014、1016和1018的对应的物理页1032、1034、1036和1038。

[0101] 在本发明的一些实施例中,可生成多模存储装置(例如,350、400或620)中的存储单元的SUE配置。可针对存储装置上的针对编程和擦除而作为一个单元共同地管理的每组

芯片来组织SUE页和SUE块。SUE块可定义为包括物理块的存储单元,并且这些物理块可作为一个单元共同地被擦除和管理,其中,从多模存储装置的芯片的子集中所包括的每个芯片分别选择每个物理块。SUE页可定义为包括SUE块的离散的部分或片段,并且这些离散的部分或片段可被共同地编程。

[0102] 例如,在一些实施例中,多模存储装置可包括一百二十八个芯片,并可共同地擦除和管理每个芯片中所包括的各自物理块上的自由空间。对应的SUE块1010可定义为包括一百二十八个物理块的存储单元,其中,从多模存储装置的一百二十八个芯片中分别选择每个物理块。对应的SUE页1030可定义为包括与一百二十八个物理块分别对应的一百二十八个部分或片段。

[0103] 在另一实施例中,多模存储装置可包括一百二十八个芯片,并且例如可一次共同地擦除并管理在三十二个芯片的各自物理块上的自由空间。对应的SUE块1010可定义为包括三十二个物理块的存储单元,其中,从多模存储装置的三十二个芯片中分别选择每个物理块。在这种情况下,对应的SUE页1030可定义为包括与三十二个物理块分别地对应的三十二个部分或片段。

[0104] 在又一实施例中,多模存储装置可包括被划分为四个平面(plane)的一百二十八个芯片,并可管理各个平面的每个芯片中所包括的块上的自由空间。对应的SUE块1010可定义为包括存储装置中所包括的各个平面的一百二十八个块的存储单元。在这种情况下,对应的SUE页1030可定义为包括与各个平面的块对应的一百二十八个部分或片段。

[0105] 图11是描绘根据本发明的实施例的在多模存储装置的用户区域中的用于存储的用户存储空间示例性SUE块和对应的SUE页的框图。

[0106] 现在参照图11,用户存储空间的示例性SUE块(S-Block) 1110被表示为框图。在一些实施例中,SUE块1110可被认为与虚拟块(V-Block)相似。SUE块1110可为在单个存储装置级用于存储介质管理的操作的基本单元。SUE块可由多个SUE页(S-Page)组成。在一些示例性实施方式中,SUE页可被认为与虚拟页(V-Page)相似。例如,在图11中描绘的SUE块1110包括四个SUE页(S-Page) 1121、1122、1123和1124。

[0107] 如在图5中描绘的,分配至SUE块(S-Block)中的SUE页(S-Page)的物理存储单元可位于包括在单个多模存储装置(例如,350、400或620)中的多个芯片的对应的物理页以及物理块中。可选择的实施例可包括基于多模存储装置的物理擦除块大小和可编程物理页大小之间的关系而被划分为任意数量的页的块。

[0108] 图12是描绘根据本发明的实施例的在多模存储装置的用户区域中用于存储的示例性SUE元页和对应的SUE页的框图。

[0109] 现在参照图12,示例性元页(M-Page) 1210被表示为框图。元页1210可由在存储系统中的多个存储装置的多个SUE页(S-Page)组成。例如,图12中描绘的元页1210包括五个SUE页(S-Page) 1211、1212、1213、1214和1215。可选择的实施例可包括基于存储系统中单个多模存储装置的数量和在每个多模存储装置中作为一个单元共同地管理的芯片的数量而被划分为任意数量的SUE页的元页。

[0110] 分配给每个SUE页的物理存储单元可位于单个多模存储装置(例如,350、400或620)中。分配给形成元页1210的各个SUE页1211、1212、1213、1214和1215的存储单元可位于与存储系统(例如,存储设备)关联的多个存储装置(例如,620、630、640和650)中。

[0111] 因此,虽然SUE页1121、1122、1123和1124的大小或宽度可与在每个多模存储装置中作为一个单元共同地管理的芯片的数量对应,但是元页1210的大小或宽度可与包括在存储系统中的多模存储装置的数量对应。

[0112] 图13是描绘根据本发明的实施例的在多模存储装置的用户区域中用于存储的示例性SUE元块和对应的SUE元页的框图。

[0113] 参照图13,示例性元块(M-Block) 1310被表示为框图。元块(M-Block) 1310可由多个元页1311、1312、1313和1314组成。因为具有元页(M-Page) 1210,所以分配给元块(M-Block) 1310的物理存储单元可位于与存储系统关联的多个存储装置中。即,元块(M-Block) 1310可包括从包括在芯片的对应的子集中的每个芯片选择的各个块,所述芯片的对应的子集在包括在存储系统中的每个多模存储装置(例如,620、630、640和650)中作为一个单元共同地管理。因此,元块1310的大小可与在每个多模存储装置(例如,350、400或620)中共同地管理的芯片的数量和在存储系统中包括的多模存储装置(例如,620、630、640和650)的数量对应。

[0114] 图14是描绘根据本发明的实施例的在多模存储装置的用户区域中用于存储的另一示例性SUE元块和对应的SUE块的框图。

[0115] 参照图14,示例性元块(M-Block) 1410可被表示为由多个SUE块(S-Block) 1110组成的框图。元块1410可为从存储系统中所包括的多模存储装置(例如,620、630、640和650)中的每个中共同地管理的芯片的每个子集中选择的各个SUE块1411、1412、1413、1414和1415的集合。相似地,元页1210可为从元块1410中对应的SUE块1411、1412、1413、1414和1415中的每个选择的对应的SUE页(例如,1211、1212、1213、1214和1215)的集合。

[0116] 在本发明的一些实施例中,可以以元块级执行某些存储介质管理功能(例如,擦除、编程、回收(垃圾收集)、自由空间管理等)。也就是说,可以在存储系统级而不是在单个存储装置级协调这些存储介质管理功能。

[0117] 为了使期望的系统级存储管理成为可能,例如由应用和虚拟机(VR)操作系统寻址的专用于用户数据的逻辑地址空间可映射到SUE地址空间。因此,存储系统中所包括的多模存储装置的用户区域可通过底层暴露接口来寻址。系统级存储映射和管理导致更低的写放大因子,这使存储区域配置(storage area provisioning)减少,导致成本节约。

[0118] 图15是描绘根据本发明的实施例的可由多模存储系统实施以提供从逻辑地址到SUE地址的地址映射的示例性SUE映射方案的框图。

[0119] 参照图15,在本发明的实施例中,示出SUE映射方案1500,SUE映射方案1500可由诸如图6的多模存储管理系统610的存储系统实施,以提供从逻辑地址到SUE地址的地址映射。SUE映射方案1500可将逻辑地址空间与SUE地址空间关联。SUE地址空间可揭露底层物理存储介质的显著特征。SUE地址空间可用来对存储系统中所包括的多个存储装置的集合的物理存储空间进行寻址。

[0120] 可例如从主机应用和虚拟机操作系统接收用户数据1502作为输入。主机用户数据可被组织成对应于与本地主机文件系统、接口标准等关联的逻辑块大小(例如,512K字节的信息)的存储单元(例如,逻辑寻址的块或者逻辑块)。接收的用户数据的每个逻辑块可由逻辑块地址(LBA)寻址。例如,在一些实施例中,输入逻辑块寻址可与由美国国家标准协会(ANSI)公布的小型计算机系统接口(SCSI)标准关联。

[0121] 用户数据1502的逻辑寻址的块可组合成SUE可寻址单元或者混合映射系统(HMS)映射块(HMB)。在一些实施例中,整数数量的逻辑块可被分组以形成SUE可寻址单元。例如,在图15中,八个逻辑块被组合以形成每个SUE可寻址单元。在可选择的实施例中,任意全部或部分数量的逻辑块可被组合以形成SUE可寻址单元。

[0122] 在一些实施例中,SUE可寻址单元可具有用于存储系统的映射的最小粒度。在各种实施例中,SUE可寻址单元大小可包括4K字节、8K字节或任意其他合适的大小或大块的信息。

[0123] 在一些实施例中,存储系统可包括一组卷,每个卷可包括一组SUE可寻址单元,每个可寻址单元可包括一组逻辑单元。不同的卷可分别利用不同的SUE可寻址单元大小。理解的是,卷可具有许多特性。卷可对应于:应用、单个用户级文件系统、逻辑驱动器、命名空间(例如,与给出的命名空间关联的连续的逻辑地址的收集)、LUN(逻辑单元号)等。

[0124] 在描绘的示例实施方式中:由逻辑块地址1521、1522、1523、1524、1525、1526、1527和1528寻址的逻辑寻址的块1531、1532、1533、1534、1535、1536、1537和1538组合成SUE可寻址单元1503;由逻辑块地址1541、1542、1543、1544、1545、1546、1547和1548寻址的逻辑寻址的块1551、1552、1553、1554、1555、1556、1557和1558组合成SUE可寻址单元1504;由逻辑块地址1571、1572、1573、1574、1575、1576、1577和1578寻址的逻辑寻址的块1581、1582、1583、1584、1585、1586、1587和1588组合成SUE可寻址单元1505。逻辑块可跨越可寻址单元。每个可寻址单元可存在多个块。

[0125] 可以对SUE可寻址单元(例如,1503、1504和1505)中的用户数据1502选择性地执行数据压缩算法,以产生压缩的SUE可寻址单元(例如,1507、1508和1509)。可对应于每个压缩的SUE可寻址单元(例如,1507、1508和1509)来产生标题节(header section)(例如,1511、1512和1513)。标题节可包括例如在回收和数据恢复活动中使用的信息。

[0126] 压缩的SUE可寻址单元和标题节可位于存储装置传输块或者SSD传输块(STB)1515和1517中。在描绘的示例中,标题节1511、1512、1513和对应的压缩的SUE可寻址单元1507、1508和1509包括在STB 1515和1517中。在一些实施例中,压缩的SUE可寻址单元以及用户数据1502的附上的逻辑块可允许跨越两个或更多个存储装置传输块。

[0127] 整数数量的存储装置传输块可与SUE页1591、1592、1593和1594对齐,以传送到多模存储装置。在一些实施例中,压缩的SUE可寻址单元以及用户数据1502的附上的逻辑块可允许跨越两个或更多个SUE页1591、1592、1593和1594。

[0128] 在一些实施例中,可不在系统级针对用户数据1502执行诸如纠错码(ECC)的错误校验,而是相反的,错误校验必须由单个多模存储装置执行。

[0129] 与用户数据关联的元数据可存储在多模存储装置(例如,350、400或620)的逻辑寻址的系统区域中。例如,在一些实施例中,使用逻辑块寻址来寻址的多模存储装置中的存储单元的分区可存储可将SUE可寻址单元映射到SUE地址空间的映射表。例如,映射表可存储指针,每个指针可指向单个SUE可寻址单元。因此,可使用SUE可寻址单元的映射以及SUE可寻址单元和逻辑块的对应的偏移来确定在SUE地址空间中用户数据的逻辑块的对应的存储位置。

[0130] 在一些实施例中,信息可存储在卷(例如,671、672和673)中。可存在多个卷或命名空间,不同的卷或命名空间可与不同大小的SUE可寻址单元关联。还应理解的是,不同大小



的卷或命名空间可与相同大小的SUE可寻址单元关联。

[0131] 图16是描绘可实施图15的SUE映射方案的示例性存储系统的示意图。如图16中示出的,示例性存储系统1602可包括处理器1604、存储器1606、网络接口1608、输出/输入(I/O)装置1610、显示装置1612、总线1614和多个非易失性存储装置1616。存储系统1602的各个组件可通过本地数据链路1618结合,本地数据链路1618在各种实施例中包含例如地址总线、数据总线、串行总线、并行总线或它们的任意组合。

[0132] 处理器1604可包括适用于控制存储系统1602的任何通用或专用的数字处理器。存储器1606可包括适用于存储数据和用于由处理器1604访问的指令的任何数字存储装置。

[0133] 网络接口1608可包括适用于将存储系统1602通信地连接到通信网络(例如,局域网(LAN)或互联网协议(IP)网络)的任何网络接口。网络接口1608可实施存储网络标准,例如,互联网小型计算机系统接口(iSCSI)协议。

[0134] 输入/输出装置1610可包括用于将数字信息发送到存储系统1602或从存储系统1602接收数字信息的任何适合的装置。显示装置1612可包括用于显示文本或图形用户界面(GUI)的任何适合的装置。

[0135] 总线1614可包括例如外围组件互连快速(PCIe)总线或在存储系统中用于通信的任何其他适合的高速串行扩展总线。总线1614可利用标准NVM快速(NVMe)或非易失性存储主机控制器接口规范(NVMHCI)来访问存储系统1602中诸如非易失性存储装置1616的存储装置。非易失性存储装置1616可包括例如基于NAND闪存的固态驱动器(SSD)或任何其他适合的非易失性存储装置。

[0136] 在可选择的实施例中,通用计算装置可实现图15的SUE映射方案1500的功能。例如,通用计算装置可包括服务器、工作站或个人计算机等。

[0137] 在诸如非易失性存储装置1616的计算机可读介质上存储的诸如源码、目标码或可执行码的程序码可被加载到工作存储器或操作存储器(例如,存储器1606)中并由处理器1604执行,以执行图15的SUE映射方案1500的功能。在可选择的实施例中,可执行指令可存储在固件中,或者该功能可由专用硬件执行。

[0138] 图17是表示根据本发明的实施例的将逻辑地址空间映射到SUE地址空间的示例性方法的流程图。现在参照图17,示例性处理流程可例如由图16的存储系统1602执行,以实现在本公开中描述的用于将逻辑地址空间映射到SUE地址空间以对包括在存储系统中的多个存储装置的集合的物理存储空间寻址的SUE映射方案的实施例。

[0139] 处理开始于例如从主机应用或虚拟机操作系统接收用户数据的操作1702。接收的用户数据可在逻辑块中被组织并由逻辑块地址进行寻址。逻辑块可对应于与本地主机文件系统、数据库等关联的最小可寻址存储单元大小。

[0140] 在操作1704中,如上面描述的,逻辑块可组合成SUE可寻址单元。例如,整数数量的逻辑块可被分组,以形成每个SUE可寻址单元。在操作1706中,如上面解释的,可以对SUE可寻址单元中的用户数据可选择地执行数据压缩算法。(图17中以虚线示出的组件是可选择项。)

[0141] 在操作1708中,可产生标题节并将标题节添加到每个SUE可寻址单元。如上面描述的,标题节可产生为包括例如在回收和数据恢复活动中使用的信息。在操作1710中,如上面解释的,压缩的SUE可寻址单元和标题节可位于存储装置传输块中。

[0142] 如上面进一步解释的,在操作1712中,整数数量的存储装置传输块可与SUE页结合并对齐,在操作1714中,与SUE页对应的存储装置传输块可被传送给多模存储装置,以存储在用户区域中。在操作1716中,如上面描述的,关于SUE页中的用户数据的元数据可被发送到多模存储装置,以存储在系统区域中。

[0143] III、多模存储管理系统

[0144] 在图18中示出了本发明的另一实施例,其示出了采用SUE寻址方案以允许存储系统(例如,图16中的存储系统1602)在存储系统中寻址逻辑和SUE存储空间的示例性多模存储管理系统1802。多模存储管理系统1802可包括SUE存储管理器1804、逻辑存储管理器1806、回收管理器1808和存储阵列管理器1810。

[0145] SUE存储管理器1804可提供用户数据存储映射、读写功能。SUE存储管理器1804可使用SUE地址映射方案将用户数据映射到存储系统的用户区域。SUE存储管理器1804可通过存储系统的存储装置的SUE接口来访问存储在用户区域中的用户数据。

[0146] SUE映射方案可在存储系统与存储装置之间分配逻辑块地址到物理地址映射功能。也就是说,SUE映射方案可将从逻辑块地址到SUE地址的存储系统级映射或虚拟化与从SUE地址到物理地址的存储装置级映射或转换结合起来。

[0147] SUE映射方案可将存储装置的某些物理特征或典型几何暴露给存储系统,使得将跨越多个存储装置在存储系统级而非单个存储装置级执行的用户数据的某些非易失性存储管理功能成为可能。从单个存储装置级到从存储系统级的用户数据管理任务的这种再分布可导致系统效率,包括降低的写入放大因子,允许减小的资源配置和降低的成本。

[0148] 逻辑存储管理器1806可提供系统数据存储映射、读写功能。逻辑存储管理器1806可使用诸如逻辑块寻址(LBA)的逻辑地址映射方案将系统数据映射到存储装置的系统区域。逻辑存储管理器1806可通过存储装置的逻辑接口访问存储在系统区域中的系统数据。

[0149] 因此,在一些实施例中,关联的存储装置的存储空间或多个关联的存储装置中的每个存储装置可被细分或分区为包括逻辑寻址系统区域和SUE寻址用户区域的分开的存储区域或地址空间。存储装置可包括两个主机接口,逻辑主机接口可提供对逻辑寻址系统区域的访问,SUE主机接口可提供对SUE寻址用户区域的访问。可由单个存储装置控制器执行关于系统数据的非易失性存储管理功能。

[0150] 关于用户数据,回收管理器1808可在存储系统级提供包括自由空间管理与回收或垃圾收集的非易失性存储管理功能。因此,包括在存储系统中的单个存储装置可以不执行用于用户数据的局部回收(垃圾收集)。回收管理器1808可实施各种自由空间管理和回收方法。在一些实施例中,回收管理器1808也可执行在此公开中描述的新的自由空间管理和回收方法。

[0151] 关于用户数据,存储阵列管理器1810或单个磁盘冗余阵列(RAID)管理器可为包括在存储系统中的多个存储装置的阵列提供存储管理(包括数据恢复功能)。因此,包括在存储系统中的单个存储装置可不对用户数据执行芯片级RAID功能。存储阵列管理器1810可实施各种存储管理和数据恢复方法。在一些实施例中,存储阵列管理器1810也可执行在此公开中描述的新的存储管理和数据恢复方法。

[0152] 现在参照图19,示出了另一示例性多模存储管理系统1902,所述多模存储管理系统1902采用SUE寻址方案以允许存储系统在诸如图16的存储系统1602的存储系统中寻址逻

辑和SUE存储空间。多模存储管理系统1902可包括数据对齐单元(DAU)1904、SUE存储访问管理器1906、数据压缩管理器1908、卷映射引擎1910、缓冲管理器1912、元块管理器1914、回收管理器1916、存储阵列管理器1918和逻辑存储访问管理器1920。

[0153] 数据对齐单元(DAU)1904可从小型计算机系统接口(SCSI)目标接收逻辑寻址介质访问命令,例如,读取、写入和去映射命令。命令基于由整数索引指示存储块的线性寻址方案来使用逻辑块寻址(LBA),SCSI存储位置提取标准。在逻辑块寻址中,单精度-整数基地址用于识别数据的每个逻辑块的开始,每个线性基地址与单个逻辑块唯一地关联。因此,逻辑块寻址从操作系统、文件系统、硬盘驱动器和主机应用内隐藏或掩盖存储装置的具体细节或特征。

[0154] 在写入操作期间,数据对齐单元1904可将从SCSI目标接收的数据的逻辑块与SUE映射块结合。例如,在一些实施例中,整数数量的逻辑块可被分组以形成SUE映射块。数据压缩管理器1908可选择性地对SUE映射块中的用户数据执行数据压缩算法。

[0155] 在读取操作期间,数据对齐单元1904可从SCSI目标接收读取命令并将读取请求传递到SUE存储访问管理器1906。数据对齐单元1904可从SUE存储访问管理器1906接收请求的用户数据并将请求的用户数据传递到SCSI目标。

[0156] SUE存储访问管理器1906可提供用户数据读写功能。在写入操作期间,SUE存储访问管理器1906可针对每个SUE映射块生成标题节。标题节可包含例如在回收和数据恢复活动中使用的信息。SUE存储访问管理器1906可在存储装置传输块中放置压缩的SUE映射块和对应的标题节。在一些实施例中,压缩的SUE映射块和用户数据的附上的逻辑块允许跨越两个或更多个存储装置传输块。

[0157] SUE存储访问管理器1906还可将整数数量的存储装置传输块与SUE页对齐以传送至存储装置。SUE存储访问管理器1906可将与SUE页对应的存储装置传输块传送至写入缓冲器。

[0158] 在一些实施例中,压缩的SUE映射块和用户数据的附上的逻辑块允许跨越两个或更多个SUE页。每个SUE页可对应于存储系统的单个存储装置。SUE页可为SUE映射方案中的编程或写入操作的基本单元。

[0159] 在读取操作期间,SUE存储访问管理器1906可确定请求的用户数据的位置,并且请求将请求的用户数据从关联的存储装置读取到读取缓冲器。SUE存储访问管理器1906可将用户数据从读取缓冲器传送到数据对齐单元1904。

[0160] 作为SUE寻址方案的子功能或补充功能,数据压缩管理器1908可对用户数据执行压缩算法。由数据压缩管理器1908执行的数据压缩功能可帮助抵消导致写入放大的固有系统因子。

[0161] 卷映射引擎1910可协调SUE地址映射功能。卷映射引擎1910可保留记录用户数据的当前位置的用户区域映射表。用户区域映射表可包括将逻辑块地址关联到存储的用户数据的SUE地址的映射信息。用户区域映射表可存储在关联的存储装置的逻辑寻址的系统区域中。

[0162] 在写入操作期间,卷映射引擎1910可针对写入的用户数据基于从SUE存储访问管理器1906接收的新的或修订的SUE地址位置更新用户区域映射表。

[0163] 在读取操作期间,卷映射引擎1910可基于请求的逻辑块地址在用户区域映射表中

查找请求的用户数据的SUE地址位置并将SUE地址位置提供至SUE存储访问管理器1906。

[0164] 卷映射引擎1910可将用户数据组织为SUE页、SUE块、元页和元块。SUE块可映射单个存储装置上的多个物理块。在一些实施例中,映射至相同的SUE块的物理块可分别位于存储装置的不同芯片上。映射至同一SUE块的全部物理块可以在存储装置级作为一个单元被管理和擦除。因此,SUE块可对应于针对回收和自由空间管理而对各芯片共同管理的一组物理块。同样地,与SUE块对应的芯片上的各物理块的组可被管理为存储介质的单元。

[0165] 每个SUE块可包括多个SUE页,每个SUE页可与映射至SUE块的各物理块的物理页对齐。被定位成跨越存储系统中所包括的所有存储装置的所有SUE块的对应的SUE页可映射至元页。类似地,被定位成跨越存储系统中所包括的所有存储装置的对应的SUE块可映射至元块。

[0166] 可针对用户数据的元块执行在多模存储管理系统1902级的存储介质管理功能(例如,回收和自由空间管理)。因此,可针对在存储系统中所包括的每个存储装置中共同管理的对应物理块的组执行在多模存储管理系统1902级的存储介质管理功能。

[0167] 可针对用户数据的元页执行编程操作和读取操作。因此,可对在存储系统中包括的每个非易失性存储装置中共同管理的对应的物理页的组执行编程操作和读取操作。

[0168] 因此,以将物理存储区域的重要的组织或典型几何暴露至多模存储管理系统1902的方式,可以使存储系统中包括的存储装置虚拟化。在单一存储装置中包括的各芯片上共同管理的物理块的组可作为SUE块提供给多模存储管理系统1902,在被定位成跨越存储系统中所包括的全部存储装置的所有芯片上共同管理的物理块的对应的组可作为元块提供给多模存储管理系统1902。

[0169] 类似地,在单一存储装置中包括的各芯片上共同编程的物理页的组可作为SUE页提供给多模存储管理系统1902,在被定位成跨越存储系统中所包括的全部存储装置的所有芯片上共同编程的物理页的组可作为元页提供给多模存储管理系统1902。

[0170] 缓冲管理器1912可管理一批读取和写入缓冲器。在写入操作期间,缓冲管理器1912可在写入缓冲器中积累从SUE存储访问管理器1906接收的存储装置传输块,直到在用户数据作为SUE页通过存储阵列管理器1918被分别发送至单个存储装置之前已经积累了用户数据的几乎完整的元页。

[0171] 在读取操作期间,缓冲管理器1912可提供读取缓冲器以支持读取缓存功能。从存储阵列管理器1918接收作为存储装置传输块的用户数据的SUE页可保存在读取缓冲器中,直到被转发至SUE存储访问管理器1906。

[0172] 元块管理器1914可跟踪在存储装置的用户区域中定义的单个元块的当前状态,例如,被擦除、有效、关闭、回收或擦除。当前状态可存储在元块信息表中,元块信息表可存储在存储器中并在存储装置的系统区域中备份。元块管理器1914也可以保持当前在特定状态下的元块的对应列表,诸如被擦除列表、回收列表和擦除列表。元块管理器1914可选择用于提交至SUE存储访问管理器1906的特定的元块以进行回收活动。

[0173] 回收管理器1916可服务来自元块管理器1914的回收请求以恢复来自指定的元块的有效用户数据并将有效用户数据重定位到其他元块。回收管理器1916可请求擦除并回收与指定的元块对应的物理存储单元以在存储装置的用户区域中提供自由空间。

[0174] 存储阵列管理器1918可向SUE接口提供存储装置的用户区域,并且向逻辑接口提

供存储装置的系统区域。存储阵列管理器1918可提供数据保护功能,诸如RAID分条和奇偶校验。例如,在一些实施例中,存储装置传输块可用作RAID元素,RAID条带可包括被定位成跨越元页中所有SUE页的存储装置传输块。因此,如果存储系统中的单个存储装置故障,存储阵列管理器1918能够使用反向校验计算从故障的存储装置恢复数据。

[0175] 逻辑存储访问管理器1920可使用逻辑寻址方法提供系统数据读写功能。逻辑存储访问管理器1920可存储并检索关于用户数据的元数据,元数据包括用户区域映射表、元块信息表、卷表及存储系统文件、日志文件等。

[0176] 关于存储在用户区域中的用户数据,单个非易失性存储装置可负责某些存储介质管理功能,例如,读取重试、故障的物理块映射、纠错码(ECC)、高级增加阶跃脉冲编程(ISPP)等。关于存储在系统区域中的系统数据,单个非易失性存储装置可负责全部的存储介质管理功能,例如,回收、损耗均衡、读写缓存、读取重试、故障的物理块映射、纠错码(ECC)、高级增加阶跃脉冲编程(ISPP)等。

[0177] 现在参照图20,示出了另一示例性多模存储管理系统2002或混合映射系统(HMS),其采用SUE寻址方案以允许存储系统对存储装置中的逻辑和SUE存储空间进行寻址。

[0178] 多模存储管理系统2002可用作负责关于跨越存储系统中所包括的多个存储装置分布的用户区域的非易失性存储介质管理的全局闪存转换层(GFTL)。多模存储管理系统2002可执行非易失性存储介质访问功能、地址映射功能、回收和损耗均衡功能,其中,地址映射功能将主机应用逻辑地址空间元素映射为与物理非易失性存储位置对齐的SUE地址空间数据结构。

[0179] 多模存储管理系统2002可包括数据对齐单元(DAU)2004、用户区域访问管理器(UAAM)2006、用户区域映射引擎(UAME)2008、缓冲管理器(BM)2010、系统区域访问管理器(SAAM)2012、元块管理器(MBM)2014、元块信息管理器(MBI)2016、存储装置控制管理器(SCM)2018、存储装置访问管理器(SAM)2020、全局状态管理器(GSM)2022和全局错误管理器(GEM)2024。多模存储管理系统2002可通信地连接到系统状态管理器2026、系统记录和统计管理器2028、目标装置2030和多个非易失性存储(NVM)装置2032。

[0180] 全局错误管理器2024可管理在包括多模存储管理系统2002的存储系统中发生的全局错误。系统状态管理器2026可管理多模存储管理系统2002的状态,例如,操作条件。系统记录和统计管理器2028可基于包括多模存储管理系统2002的存储系统的操作提供系统记录/统计。非易失性存储装置2032可包括广泛使用的各种非易失性存储器。目标装置2030可以是作为读取/写入操作的目标的另一存储装置。

[0181] 数据对齐单元(DAU)2004可从目标装置2030接收逻辑寻址的介质访问命令,例如,读取、写入和去映射命令。数据对齐单元2004可接收逻辑块寻址(LBA)缓冲列表作为输入。在写入操作期间,数据对齐单元2004可将将从目标装置2030接收的数据的逻辑块组合为SUE映射块或混合映射块(HMB)。例如,在一些实施例中,整数数量的逻辑块可被分组以形成SUE映射块。

[0182] 数据对齐单元2004可合并来自目标装置2030的对齐的和未对齐的用户数据通信量,对未对齐的写入通信量执行读取/修改/写入操作以将数据与用于将逻辑地址映射到物理地址(SUE映射块)的单元对齐。数据对齐单元2004可将用户数据放置在SUE映射块对齐缓冲列表中。在各种实施例中,SUE映射块可包含固定数量的数据,诸如4KB、8KB、16KB等。

[0183] 在读取操作期间,数据对齐单元2004可从目标装置2030接收读取命令并将读取请求传递到用户区域访问管理器2006。数据对齐单元2004可从用户区域访问管理器2006接收请求的用户数据并将请求的用户数据传送到目标装置2030。

[0184] 图21是示出根据本发明的实施例的可由多模存储管理系统实施的用户区域访问管理器(UAAM)。参照图21,用户区域访问管理器(UAAM)2006可包括读取管理器(RM)2012、写入管理器(WM)2104、数据压缩管理器(DC)2106、数据解压缩管理器(DD)2108、回收管理器(RC)2110、自由空间账户管理器(FSA)2112、流控制管理器(FC)2114和服务质量管理器(QoS)2116。

[0185] 读取管理器(RM)2102可从数据对齐单元2004接收读取请求并服务读取请求。读取管理器2102可从用户区域映射引擎(UAME)2008请求相关的映射信息。读取管理器2102可将读取请求发送到存储装置访问管理器2020。在读取操作期间,读取管理器2102可从缓冲管理器2010请求读取缓冲器中的用户数据的释放。读取管理器2102可将关于读取请求的用户数据的解压缩的请求发送至用户解压缩管理器2108。

[0186] 写入管理器(WM)2104可从数据对齐单元2004接收写入请求。在写入操作期间,写入管理器2104可在SUE映射块流上生成元数据头,并生成关于用户数据的SUE寻址的用于数据区域映射引擎2008的映射信息。写入管理器2104可将压缩的请求发送至数据压缩管理器2106以调度数据压缩命令,并将写入请求发送至存储装置访问管理器2020。写入管理器2104可从缓冲管理器2010请求写入缓冲器中的用户数据的释放。当写入数据的当前元块已经变满时,写入管理器2104可请求元块管理器(MBM)2104以开启新的有效元块。

[0187] 数据压缩管理器(DC)2106可从写入管理器2104接收压缩请求并服务压缩请求。在一些实施例中,数据压缩管理器2106可对SUE映射块中的用户数据实施数据压缩算法。在一些其他实施例中,数据压缩管理器2106可将数据压缩任务调度至外部压缩单元(未示出)。

[0188] 数据解压缩管理器(DD)2108可从读取管理器2102接收解压缩请求并服务解压缩请求。在一些实施例中,数据解压缩管理器2108可对SUE映射块中的用户数据实施数据解压缩算法。在一些其他实施例中,数据解压缩管理器2108可将数据解压缩任务调度到外部解压缩单元(未示出)。

[0189] 回收管理器(RC)2110可从元块管理器2104接收回收请求并服务回收请求以恢复来自指定元块的有效数据以回收自由空间。回收管理器2110可从用户区域映射引擎2008请求相关的映射信息并将关于指定元块的读取请求发送至读取管理器2102。回收管理器2110可解析存储装置读取数据流中SUE映射块所附的元数据头并将关于指定的元块中保留的全部有效数据的写入请求发送至写入管理器2104。回收管理器2110也可服务来自存储装置控制管理器(SCM)2018的请求以回收部分元块数据。

[0190] 自由空间账户管理器(FSA)2112可在写入操作期间从写入管理器2104接收映射信息并生成关于存储在元块中的过期用户数据的自由空间信息。自由空间账户管理器2112可将自由空间信息发送至元块信息管理器2016以更新相应的元块信息表条目。

[0191] 流控制管理器(FC)2114可监测系统资源,诸如读取/写入缓冲器、压缩缓冲器、存储总线及其他队列深度等。如果系统级资源配置降至预设的限制之下,则流控制管理器2114可重置服务质量管理器2116中的节流等级。在一些实施例中,可使用系统管理员命令建立需要的等级的资源配置。流控制管理器2114可为系统管理员提供可用于接口级节流的

统计。

[0192] 服务质量管理器(QoS)可基于系统资源配置等级和延迟测量来定义服务质量策略。服务质量管理器2116可实施多个队列以服务不同的服务质量的策略池。关于基于延迟的策略,服务质量管理器2116可对队列条目实施时间戳。服务质量管理器2116可监测各种队列参数并选择请求以确保不违反策略。在流控制器2114的请求下,服务质量管理器2116可对基于配置的策略队列的通信量进行节流。

[0193] 图22是示出根据本发明的实施例的可由多模存储管理系统实施的用户区域映射引擎(UAME)的示意图。参照图22,用户区域映射引擎(UAME)2008可包括卷管理器(VM)2202、映射页读取管理器(MPRM)2204、映射页写入管理器(MPWM)2206和映射页缓存管理器(MPCM)2208。

[0194] 卷管理器(VM)2202可提供服务以创建、销毁和管理卷并处理多个配置策略。卷管理器2202可在卷表中维持相关信息,卷表可存储在存储器中并在系统区域中备份,并且卷管理器2202可为卷表中的条目提供访问服务。卷管理器2202可使用系统区域访问管理器2012以备份和还原卷表。

[0195] 当由映射页缓存管理器2208检测到映射页缺失时,映射页读取管理器(MPRM)2204可接收并服务针对缺失的映射页的由映射页缓存管理器2208提供的请求。映射页写入管理器(MPWM)2206可从映射页缓存管理器2208接收并服务针对映射页收回的请求。

[0196] 映射页缓存管理器(MPCM)2208可服务由读取管理器2102和回收管理器2110提供的映射条目信息请求,和由写入管理器2104提供的映射条目更新。当检测到映射页缺失时,映射页缓存管理器2208可从映射页读取管理器2204请求缺失的映射页。映射页缓存管理器2208可从映射页写入管理器2206请求映射页收回。

[0197] 缓冲管理器(BM)2010可管理一批读取缓冲器和写入缓冲器。在写入操作期间,缓冲管理器2010可分配并释放存储装置传输块以在写入缓冲器中积累从写入管理器2104接收的用户数据。当已经积累了用户数据的几乎完整的元页时缓冲管理器2010可从写入管理器2104接收对写入缓冲器中的用户数据的释放的请求,并将用户数据转发至存储装置访问管理器2020。

[0198] 在读取操作期间,缓冲管理器2010可分配并释放在读取缓冲器中的存储装置传输块以支持读取缓存功能。作为存储装置传输块从存储装置访问管理器2020接收的用户数据的SUE页可最初保存在读取缓冲器中。缓冲管理器2010可从读取管理器2102接收对读取缓冲器中的用户数据的释放的请求,并且缓冲管理器2010可将存储装置传输块转发至读取管理器2102。

[0199] 系统区域访问管理器(SAAM)2012可服务关于对存储系统中所包括的存储装置的系统区域中存储的系统数据的访问的请求。系统区域访问管理器2012可接收并服务来自卷管理器2202和元块信息管理器2016的请求以分别备份并还原卷表和元块信息表。系统区域访问管理器2012可接收并服务来自映射页写入管理器2206、映射页读取管理器2204和映射页缓存管理器2208的请求以访问用户区域映射表。

[0200] 图23是示出根据本发明的实施例的可由多模存储管理系统实施的元块管理器(MBM)的示意图。参照图23,元块管理器2014可包括回收元块拾取器(reclamation metablock picker,RCMBP)2302和元块状态管理器(MBSM)2304。

[0201] 回收元块拾取器 (RCMBP) 2302可监测关于用户区域元块的参数,诸如擦除计数、过期数据水平、停顿时间等。基于监测的参数,回收元块拾取器2302可选择用于回收或垃圾收集的元块。回收元块拾取器2302可实施各种损耗平衡策略。例如,回收元块拾取器2302可尝试将元块擦除计数保持在优选值范围内,并尝试在分开的元块中隔离相对动态(热)和相对静态(冷)的数据。

[0202] 元块状态管理器 (MBSM) 2304可跟踪用户区域元块的当前状态,例如,有效、关闭、擦除、被擦除、回收或垃圾收集。元块状态管理器2304可通过更新元块信息表将元块转换为各种状态。元块状态管理器2304也可保持在特定状态下的元块的各种列表,例如,被擦除的元块列表、回收元块列表和擦除元块列表。元块状态管理器2304可监测被擦除的元块列表以确定准备回收(垃圾收集)的单个元块。

[0203] 元块信息管理器 (MBI) 2016可保持元块信息表。元块信息管理器2016可针对其他模块提供对元块信息表中的条目的访问服务。元块信息管理器2016可向系统区域访问管理器2012发送请求以备份并恢复元块信息表。

[0204] 图24是示出根据本发明的实施例的可以由多模存储管理系统实施的存储装置控制管理器 (SCM) 的示意图。参照图24,存储装置控制管理器 (SCM) 2018或固态驱动器 (SSD) 控制管理器 (SCM) 可包括存储装置记录和统计管理器 (SLS) 2402、S块擦除引擎 (SBEE) 2404和存储装置错误管理器 (SEM) 2406。

[0205] 存储装置记录和统计管理器 (SLS) 2402可保持存储装置访问历史的日志。

[0206] S-Block擦除引擎 (SBEE) 2404可通过元块管理器2014从回收管理器2110接收擦除请求并管理擦除过程。S-Block擦除引擎2404可发送S-Block擦除请求至存储装置访问管理器2020。

[0207] 存储装置错误管理器 (SEM) 2406可发送请求至回收管理器2110以回收部分元块数据。

[0208] 图25是示出根据本发明的实施例的可由多模存储管理系统实施的存储装置访问管理器 (SAM) 的示意图。参照图25,存储装置访问管理器 (SAM) 2020可包括逻辑访问管理器 (SLA) 2502、RAID管理器 (RAID) 2504、读取查找引擎 (RLE) 2506和存储初始化管理器 (SI) 2508。

[0209] 读取查找引擎2506可提供/管理读取操作查找。存储初始化管理器2508可管理存储系统的初始化操作。

[0210] 逻辑访问管理器 (SLA) 2502,或SSD逻辑访问管理器可提供关于存储装置的系统区域中的系统数据的访问服务。逻辑访问管理器2502可使用各种逻辑块寻址对存储装置的系统区域中的系统数据进行寻址。逻辑访问管理器2502可利用标准NVM Express (NVMe) 或非易失性存储主机控制器接口规范 (NVMeHCI),并可命令访问在存储系统中包括的存储装置或固态驱动器 (SSD)。

[0211] 关于用户数据,RAID管理器 (RAID) 2504可为存储系统中所包括的多个存储装置的阵列提供存储管理(包括数据恢复功能)。因此,在存储系统中包括的单个存储装置可不对用户数据执行芯片级RAID功能。RAID管理器2504可实施各种存储管理和数据恢复方法。在一些实施例中,RAID管理器2504也可执行在此公开中描述的新的存储管理和数据恢复方法。



[0212] RAID管理器2504可向SUE接口提供存储装置的用户区域,并且向逻辑接口提供存储装置的系统区域。RAID管理器2504可提供数据保护功能,诸如RAID分条和奇偶校验。例如,在一些实施例中,存储装置传输块可用作RAID元素,RAID条带可包括被定位成跨越元页中的所有SUE页的存储装置传输块。因此,如果存储系统中的单个存储装置故障,RAID管理器2504能够使用反向奇偶计算从故障的存储装置中恢复数据。

[0213] 图26是示出根据本发明的实施例的可由多模存储管理系统实施的全局状态管理器(GSM)的示意图。参照图26,全局状态管理器(GSM)2022可包括电源故障管理器(PFM)2602以及错误和崩溃管理器(PFCM)2604。

[0214] 电源故障管理器2602可管理与多模存储管理系统2002的电源故障关联的问题。错误和崩溃管理器2604可管理在多模存储管理系统2002中发生的错误/崩溃问题。

[0215] 图18、图19和图20的多模存储管理系统1802、1902和2002的功能可由图16的存储系统1602实施。在可选的实施例中,多模存储管理系统1802、1902和2002的功能可由常规计算装置或由专用硬件实现。

[0216] 所提出的多模方法包括促进信息的高效和有效存储的各种特征和特性。所述特征和特性可改善与性能关联的许多不同方面。在一些实施例中,描述的分区方法的灵活性允许相对快的速度和可管理的复杂度的实现。相对大量的用户数据可存储在针对用户数据的能够非常快的存储和管理操作的SUE地址空间中。而相对少量的元数据可存储在允许存储系统利用用于减小复杂性的元数据的提取性质的逻辑寻址的区域中。

[0217] 另外,增加相对较小的元数据区域的预留空间的灵活性给出了更大百分比的预留空间影响,该影响有助于加快元数据存储操作并补偿原本将会由于复杂性的增加而发生的速度的减小。与其他存储系统相比,这允许更好的全局分配和更好的预留空间资源的影响。通过允许块的不同存储区域在两个分区之间重新指定或重新分配,所述灵活性也可促进改善的生命周期保护。存储在区域中的数据中的性质(例如,类型)可表明,一种类型的数据比另一种类型的数据在区域中写入/擦除得少(例如,与用户数据相比,大多数元数据没有改变很多),在一个分区中的物理块可重新指定给另一分区以使特定区域上的磨损均等。通过将电力循环责任上移到系统级,灵活性也允许电力循环改善。

[0218] 根据对计算机存储器内的数据位的操作的程序、逻辑块、处理和其他操作的符号表示来呈现详细描述的一些部分。这些描述和表示是由数据处理领域内的技术人员通常使用以将他们的工作的实质有效地传达给本领域其他技术人员的方式。程序、逻辑块或进程在这里通常被构思为导致期望的结果的前后照应的步骤或指令。所述步骤包括物理量的物理操作。通常地,但不是必需地,这些物理量采用可在计算机系统中进行存储、传送、组合、比较和另外的操作的电、磁、光或量子信号的形式。已经证明,有时为了方便起见,主要出于普遍使用的原因,将这些信号称为位、值、元素、符号、字符、项、数字等。

[0219] 然而,应该记住的是,所有这些和相似的术语与适当的物理量关联,并且仅仅是应用于这些物理量的方便的标签。除非特别地指出,否则从下面的讨论清楚可知,应理解的是,在整个本申请中,利用诸如“处理”、“计算(computing)”、“运算(calculating)”、“确定”、“显示”等的术语的讨论是指操作和转换被表示为物理(例如,电)量的数据的计算机系统或相似处理装置(例如,电、光或量子计算装置)的动作和过程。所述术语是指将计算机系统的组件(例如,寄存器、存储器、其他此类信息存储器、传输或显示装置)内将物理量操作

或转换为在其他组件内相似地表示为物理量的其他数据的处理装置的动作和过程。

[0220] 将仅从概念的角度理解在概念图的每个中示出的配置。为了帮助本公开的理解，已经夸大或缩小了在每个概念图中示出的每个组件的形式、结构和尺寸。实际实施的配置可具有与每个概念图中示出的物理形状不同的物理形状。每个概念图不意图限制组件的物理形状。

[0221] 提供在每个框图中示出的装置配置以帮助本公开的理解。每个块可根据功能由较小的块形成。可选择地，多个块可根据功能形成较大的块。即，本公开的精神或范围不限制于在框图中示出的配置。

[0222] 以上，已经基于一些示例实施例描述了本公开。然而，可以根据本领域的本质利用与包括本公开的主题的上述示例实施例不同的方式实现本公开的目的。因此，应该理解的是，上述示例实施例处于描述性和说明性的角度，而不是限制性的角度。即，包括本公开的主题并且可实现本公开的目标的精神或范围应该被包括在本公开的发明构思的精神或范围内。

[0223] 因此，不脱离本公开的范围或精神的修改或改变的技术构思被包括在权利要求的范围中。本公开的范围不限制于上面的示例实施例。

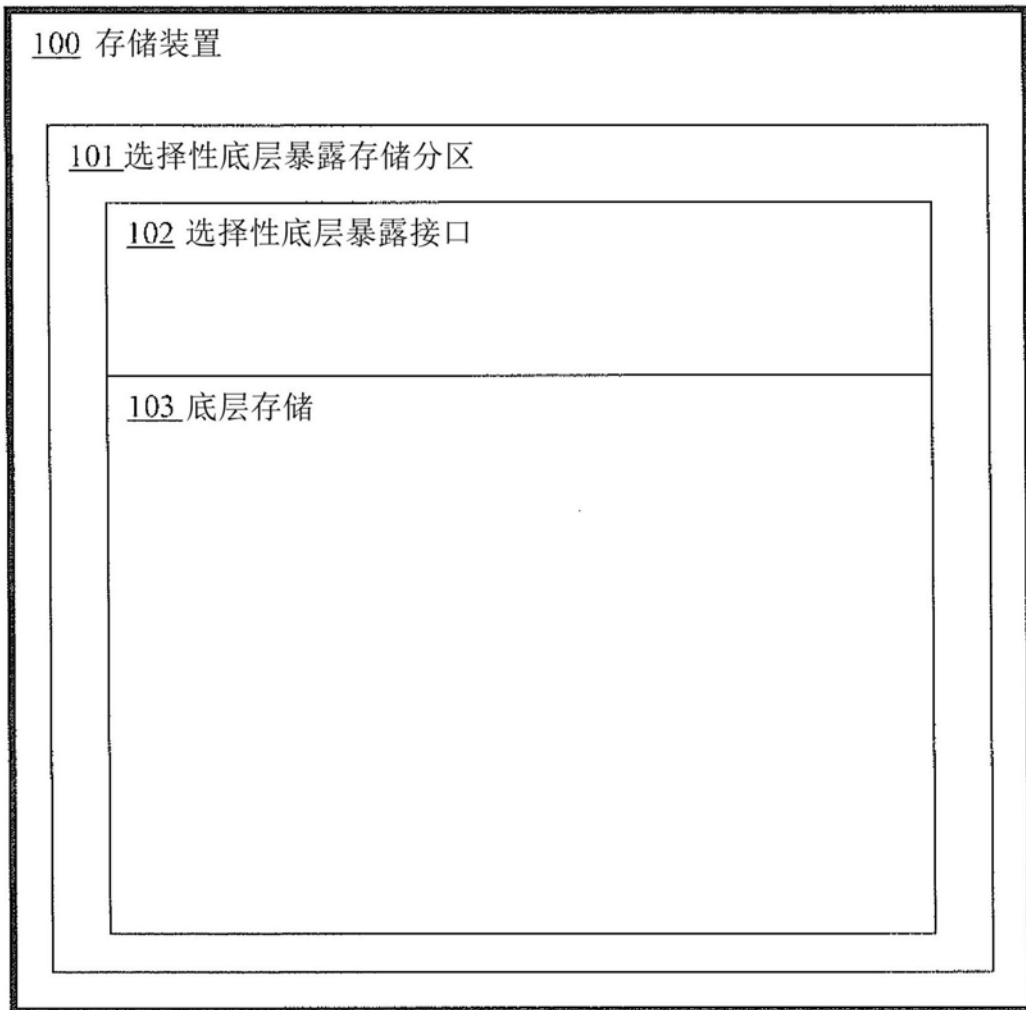


图1

220

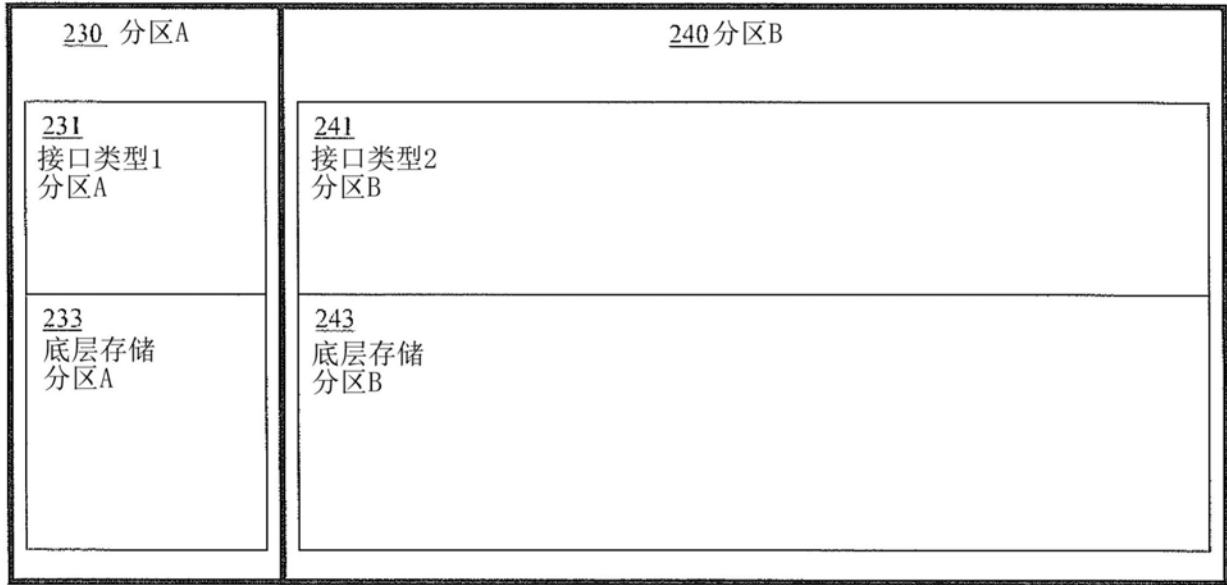


图2

350

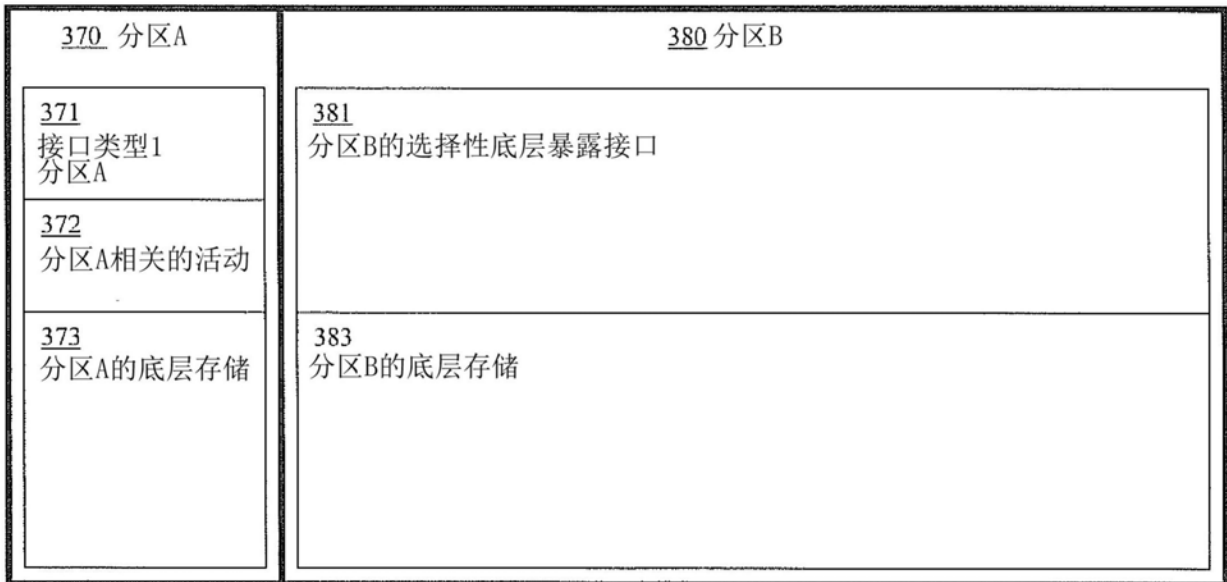


图3

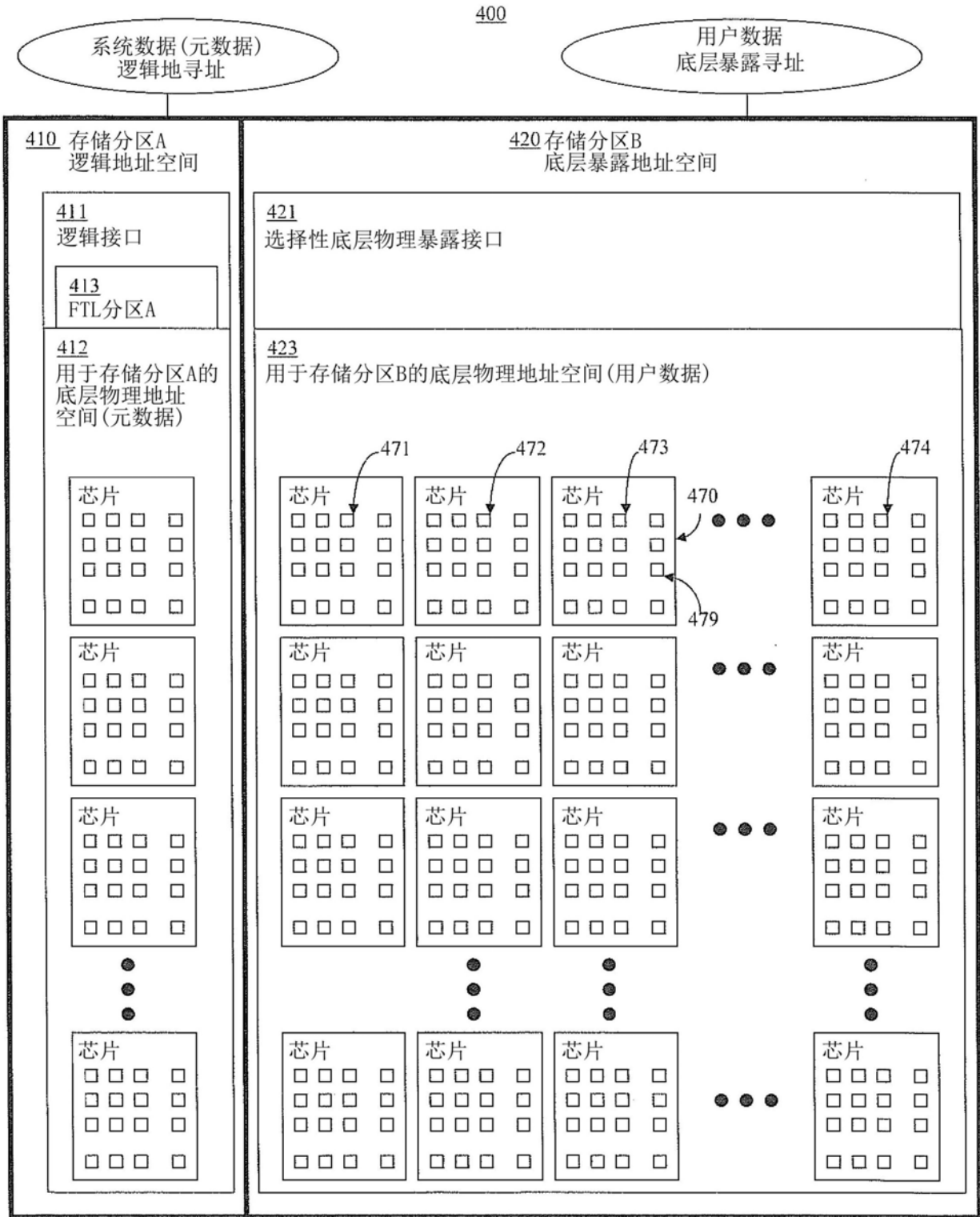


图4

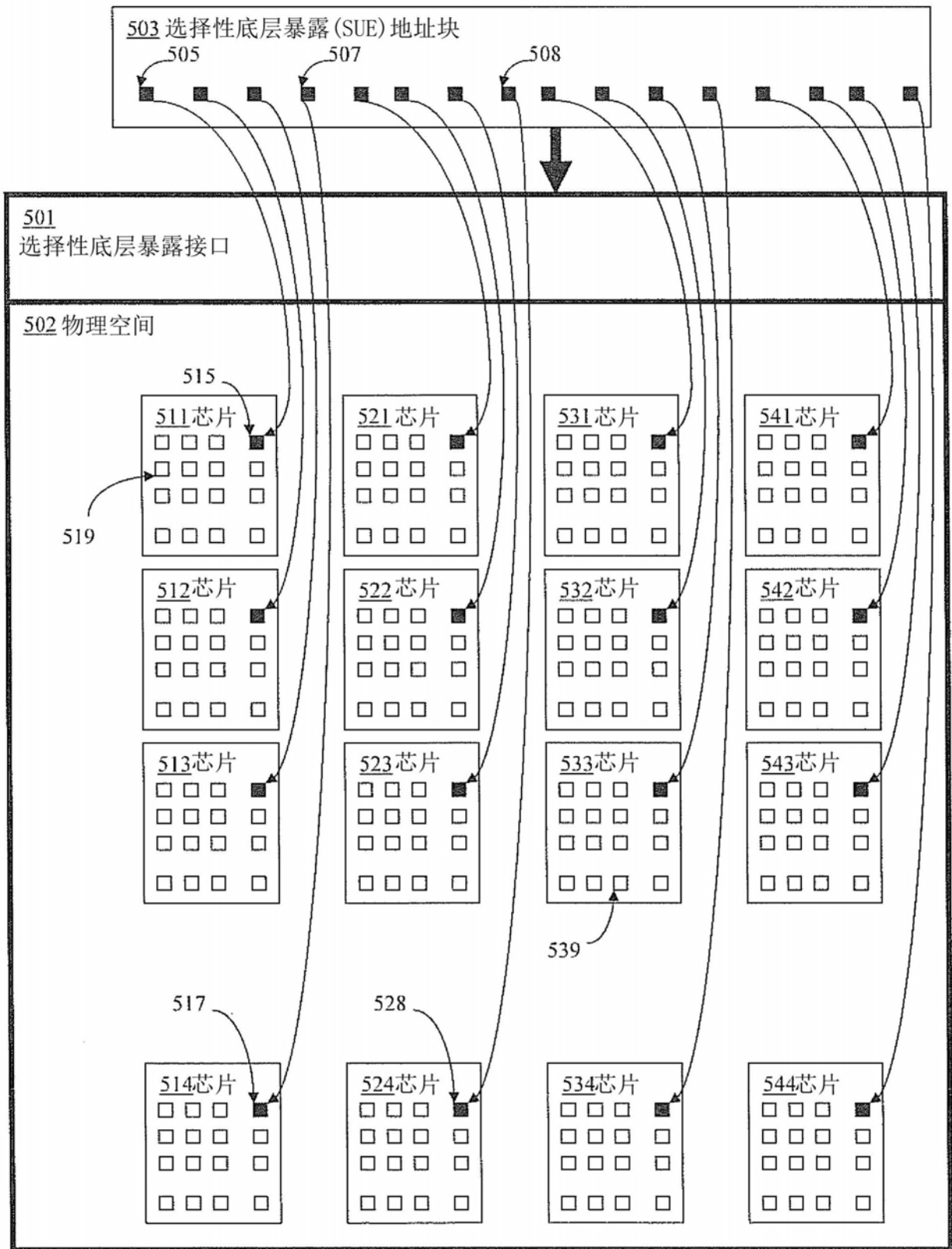


图5

600

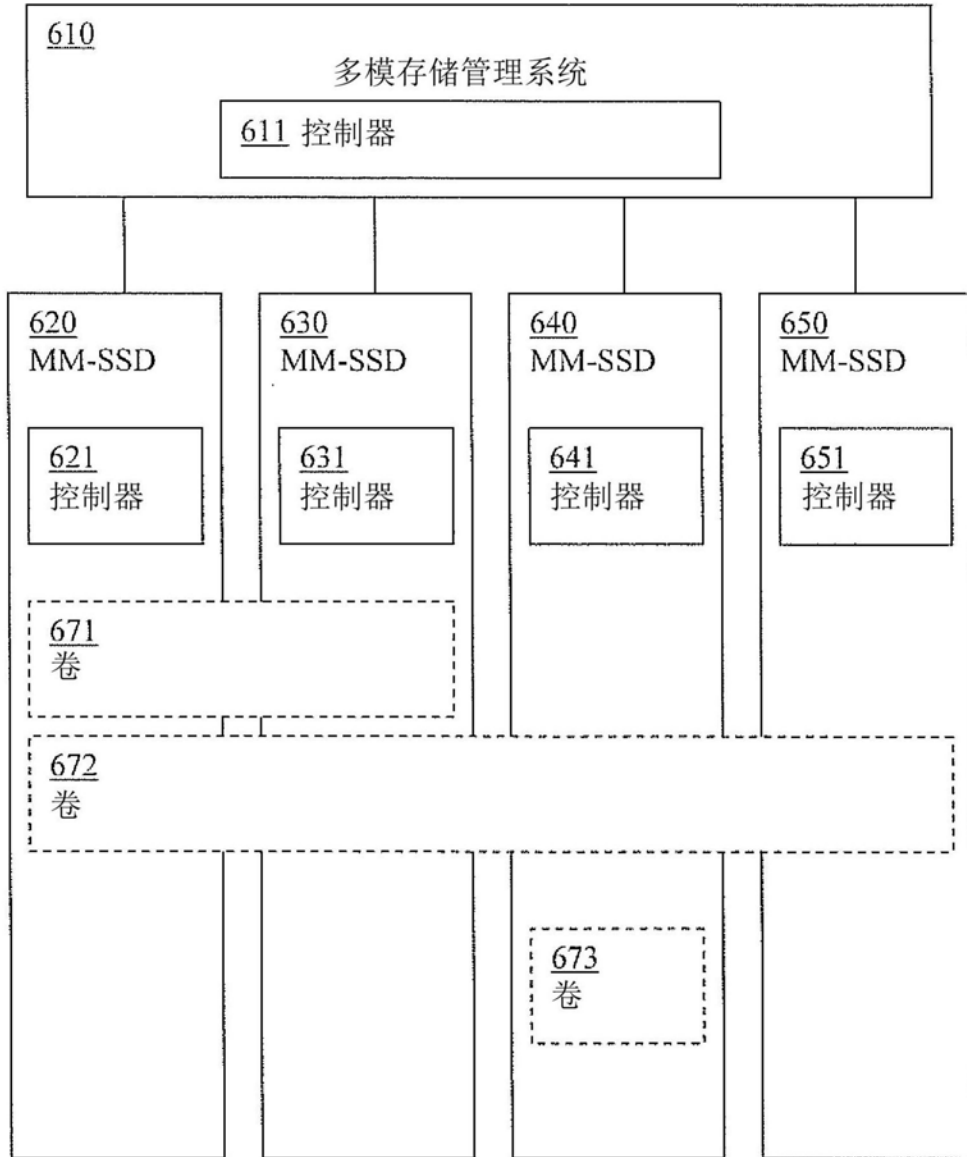


图6

700

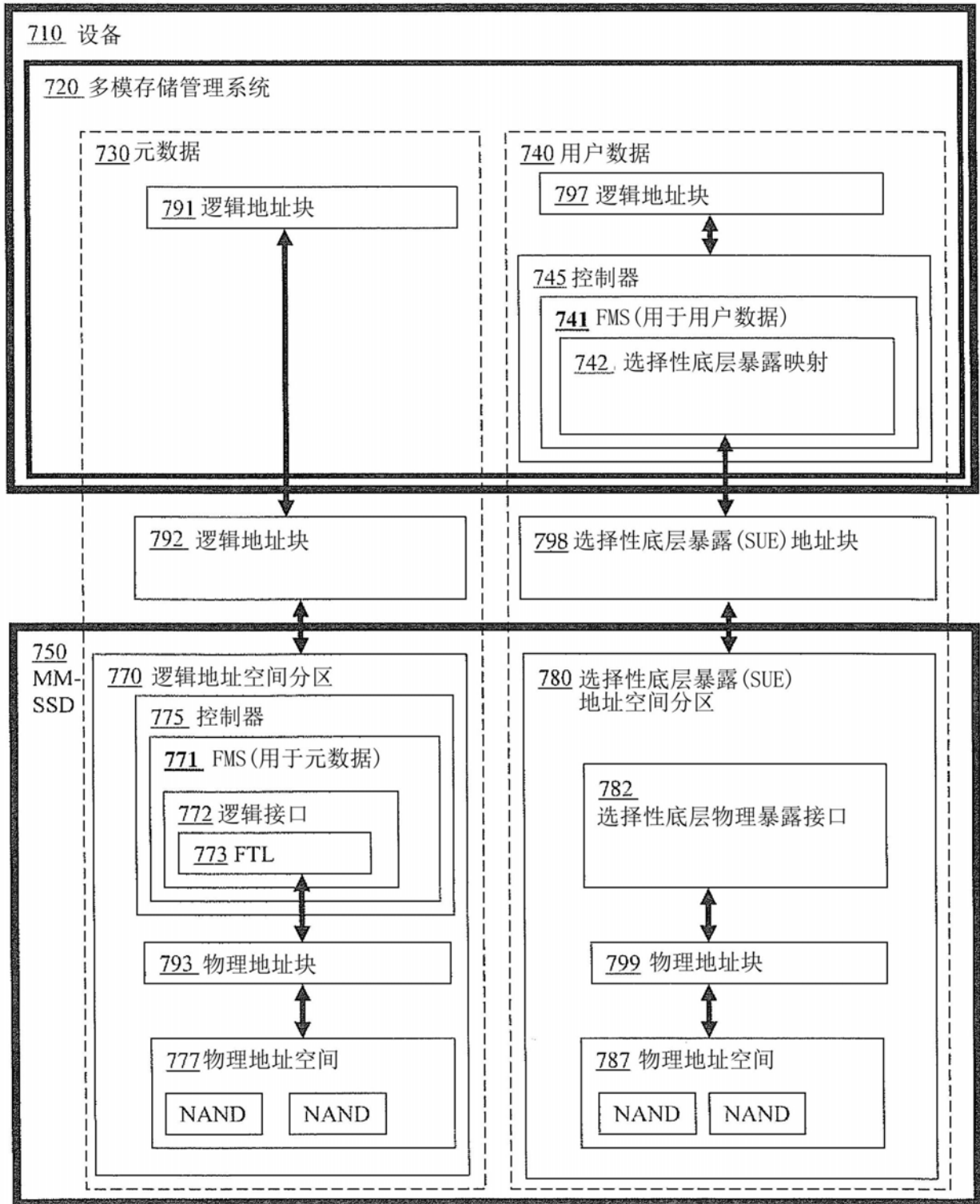


图7



800



图8

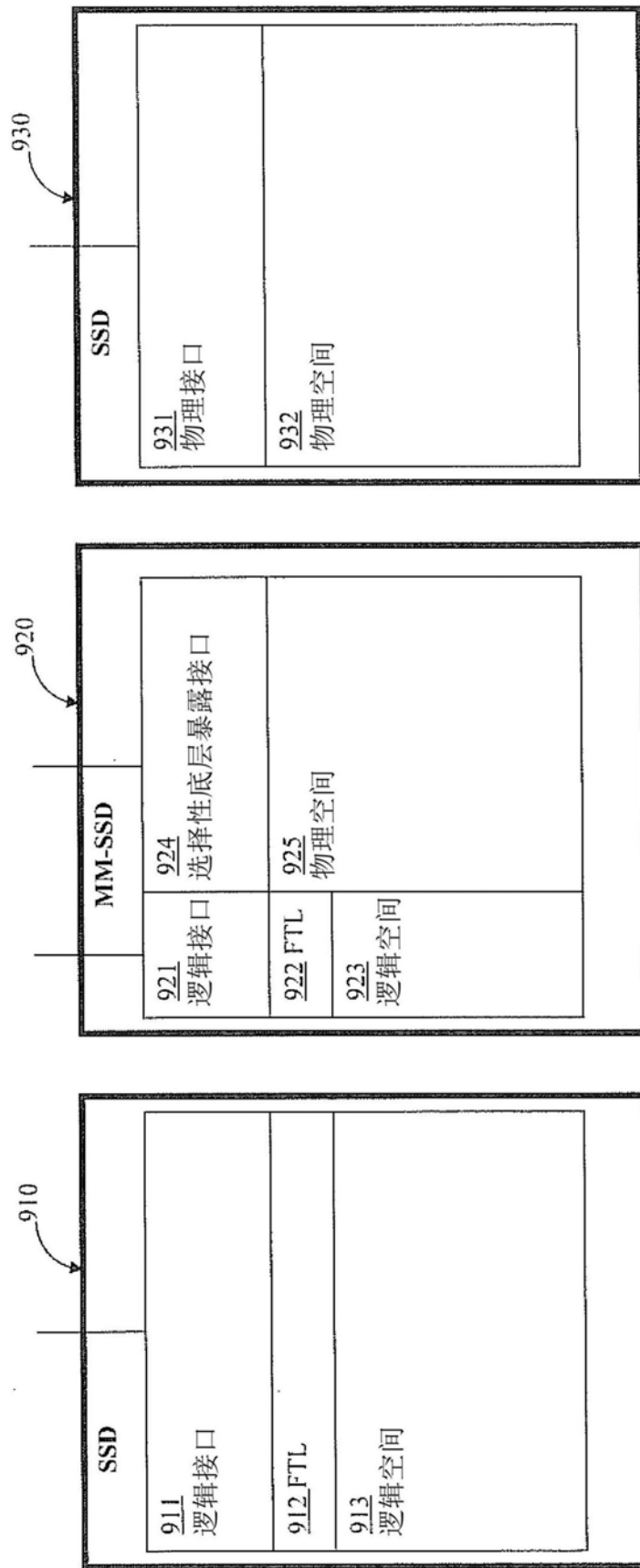


图9

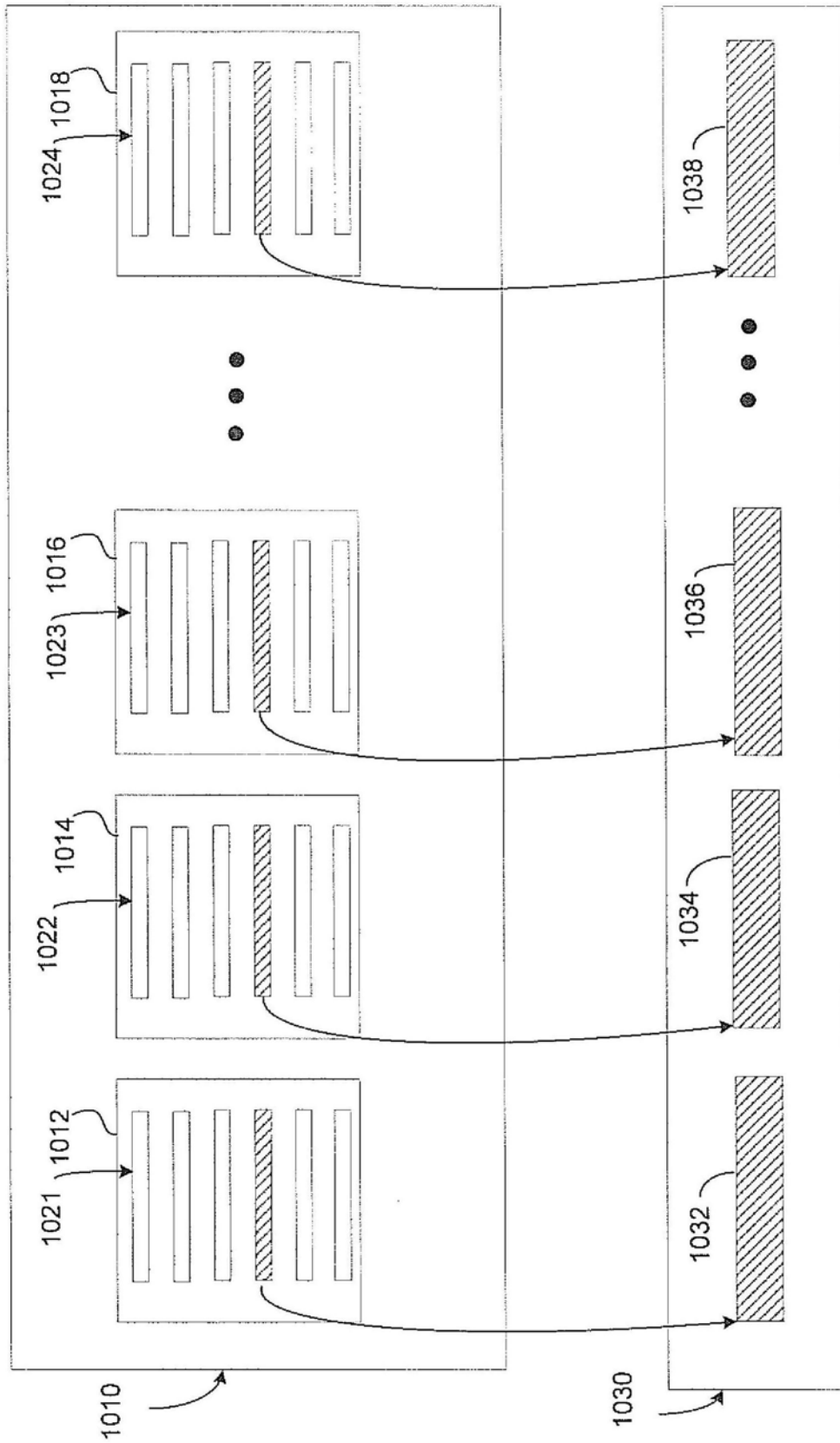


图10

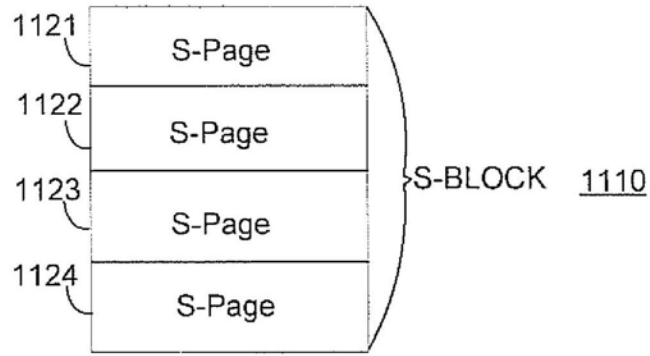


图11

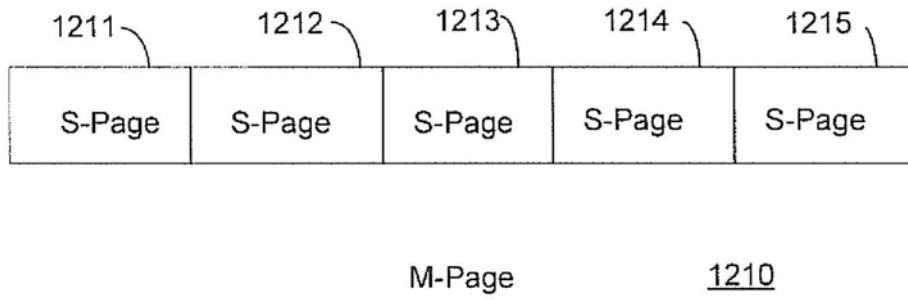


图12

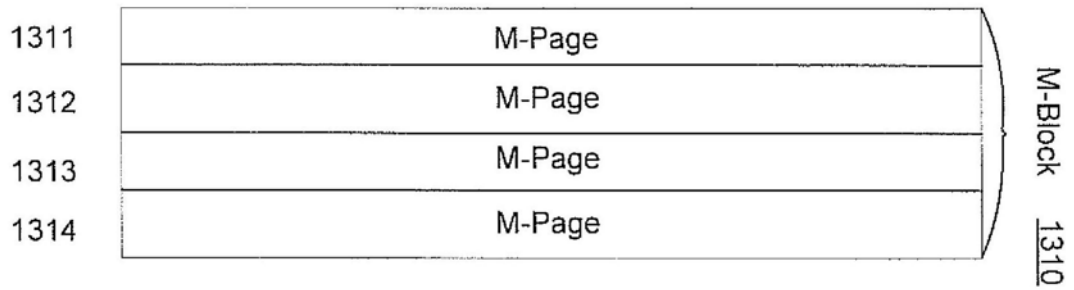


图13

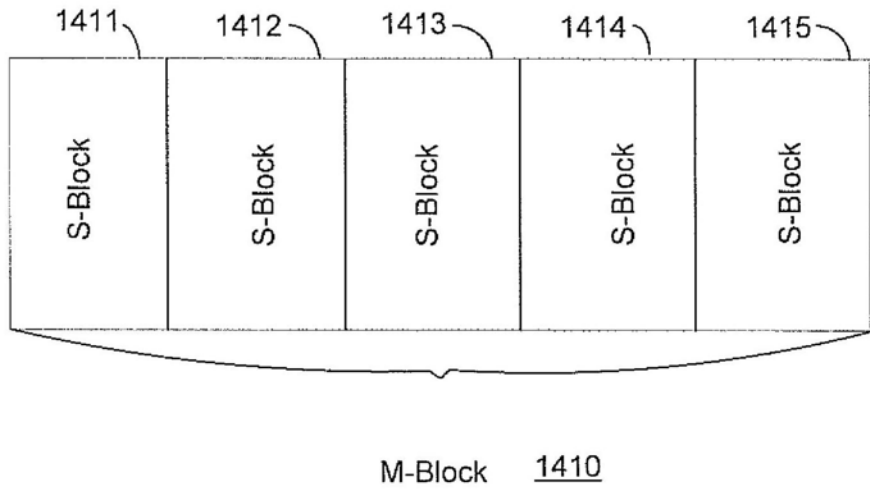


图14

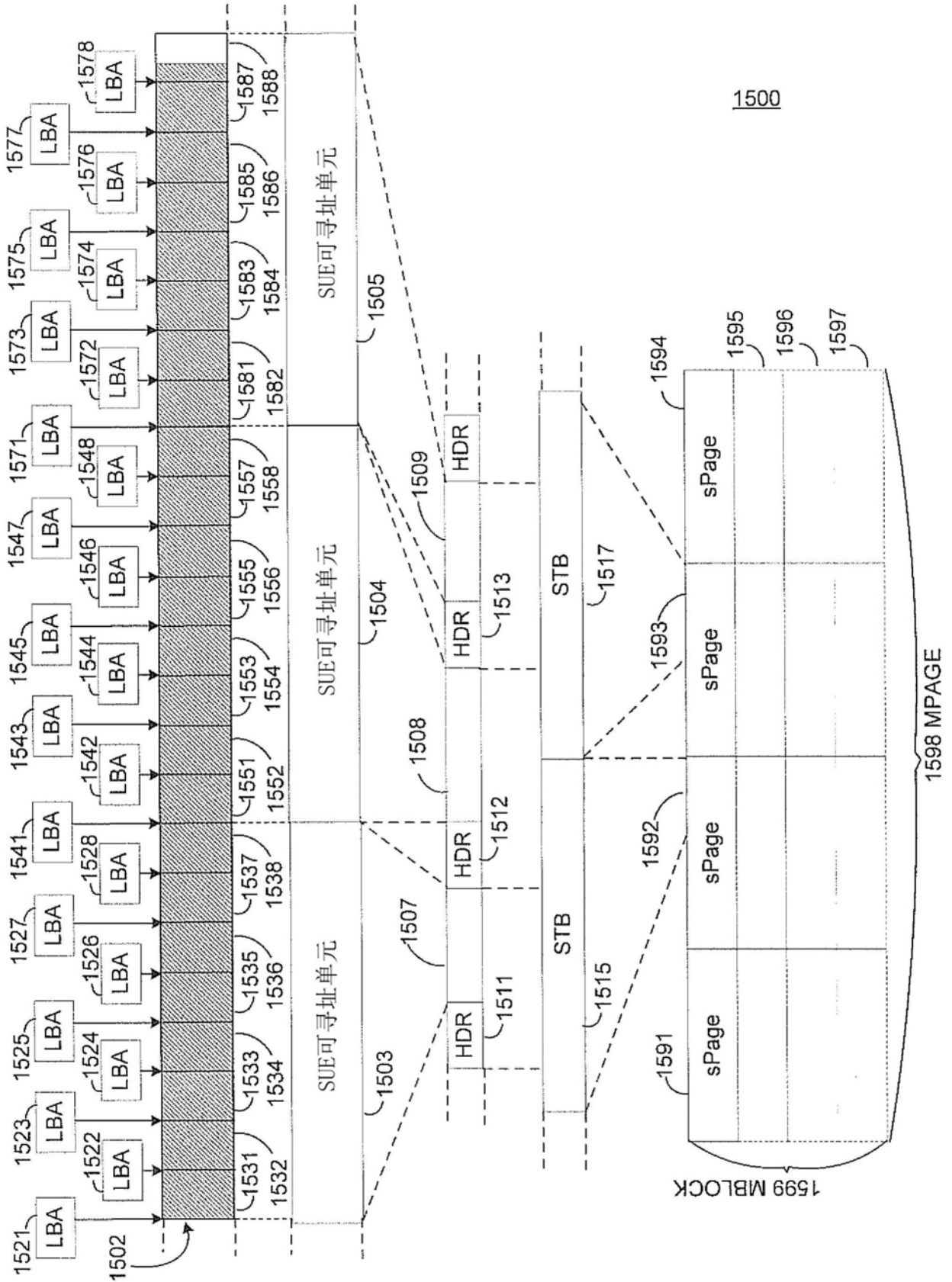


图15

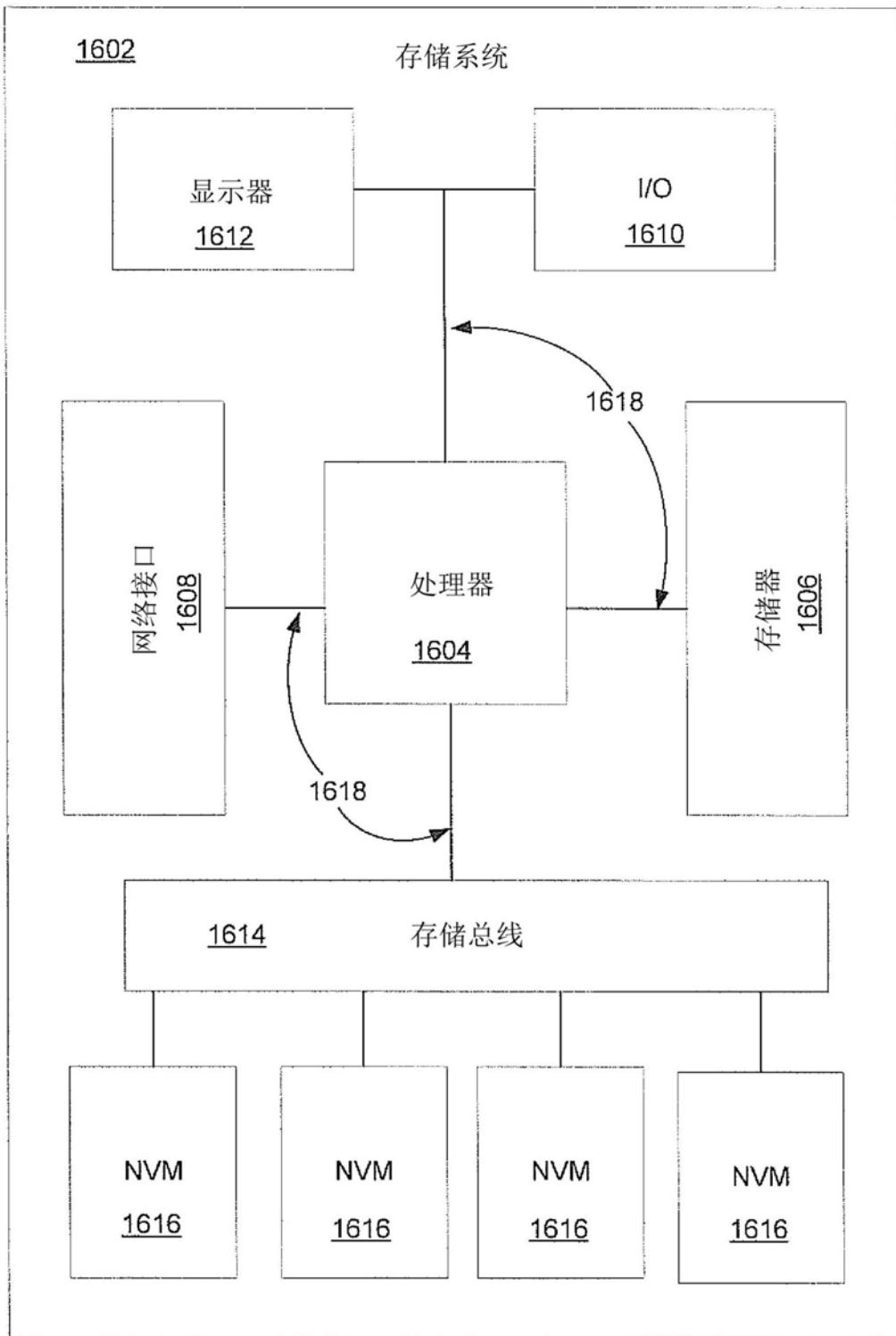


图16

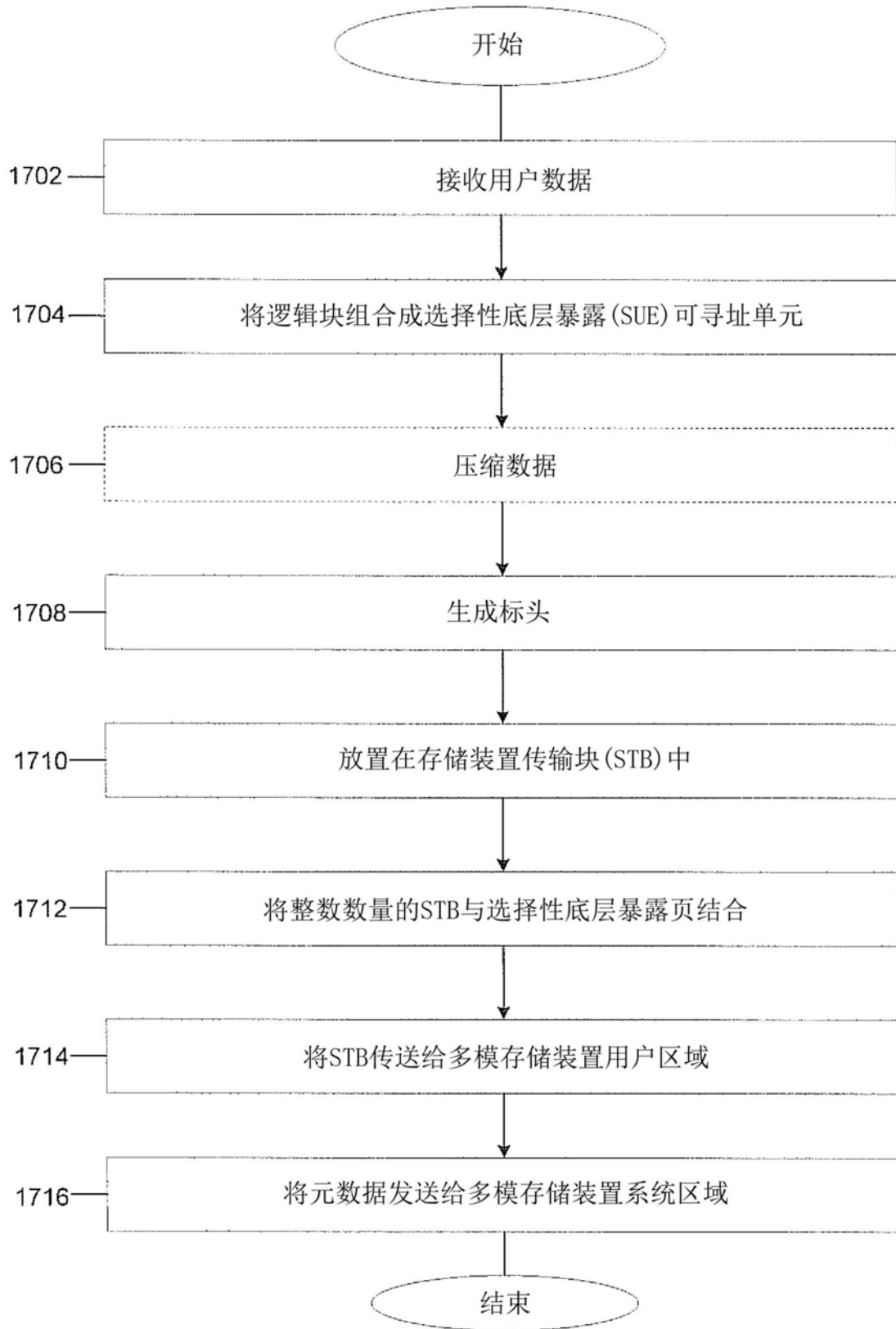


图17



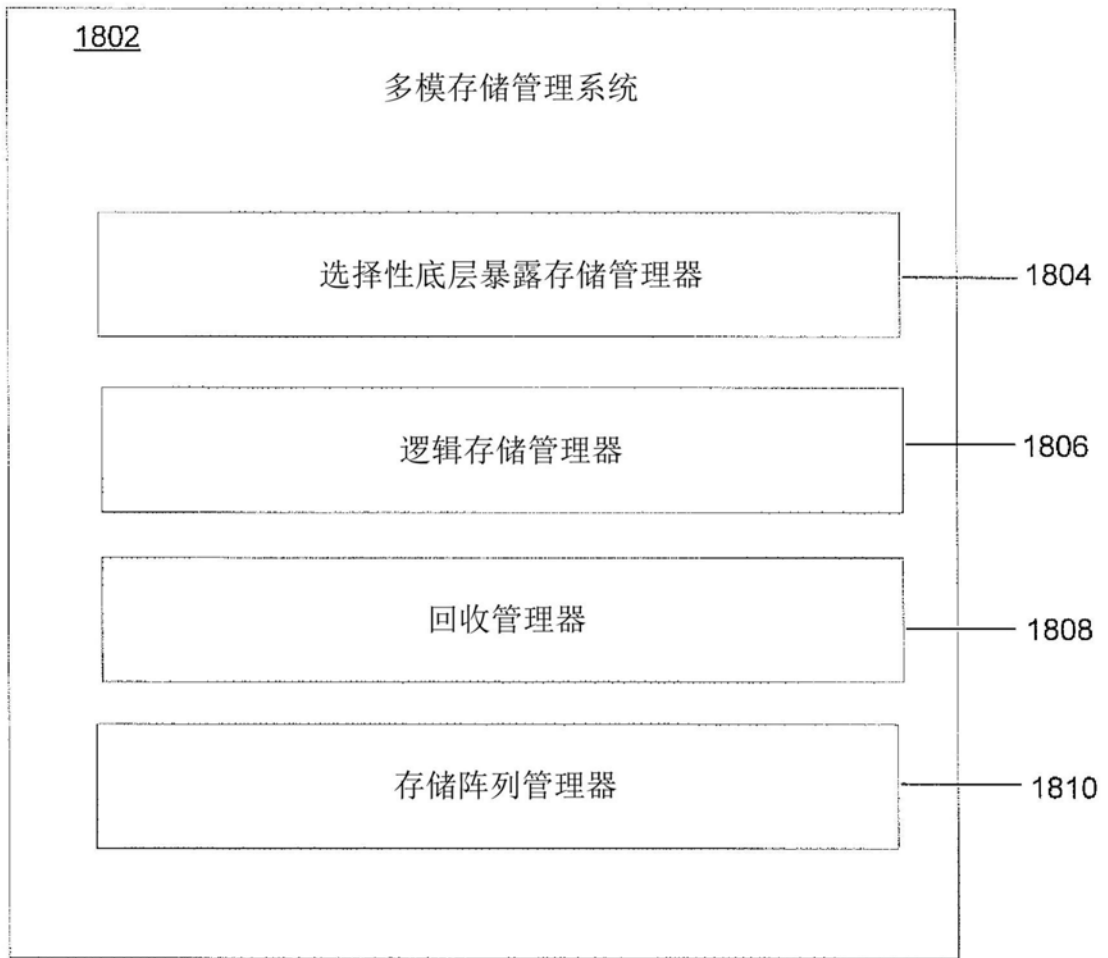


图18

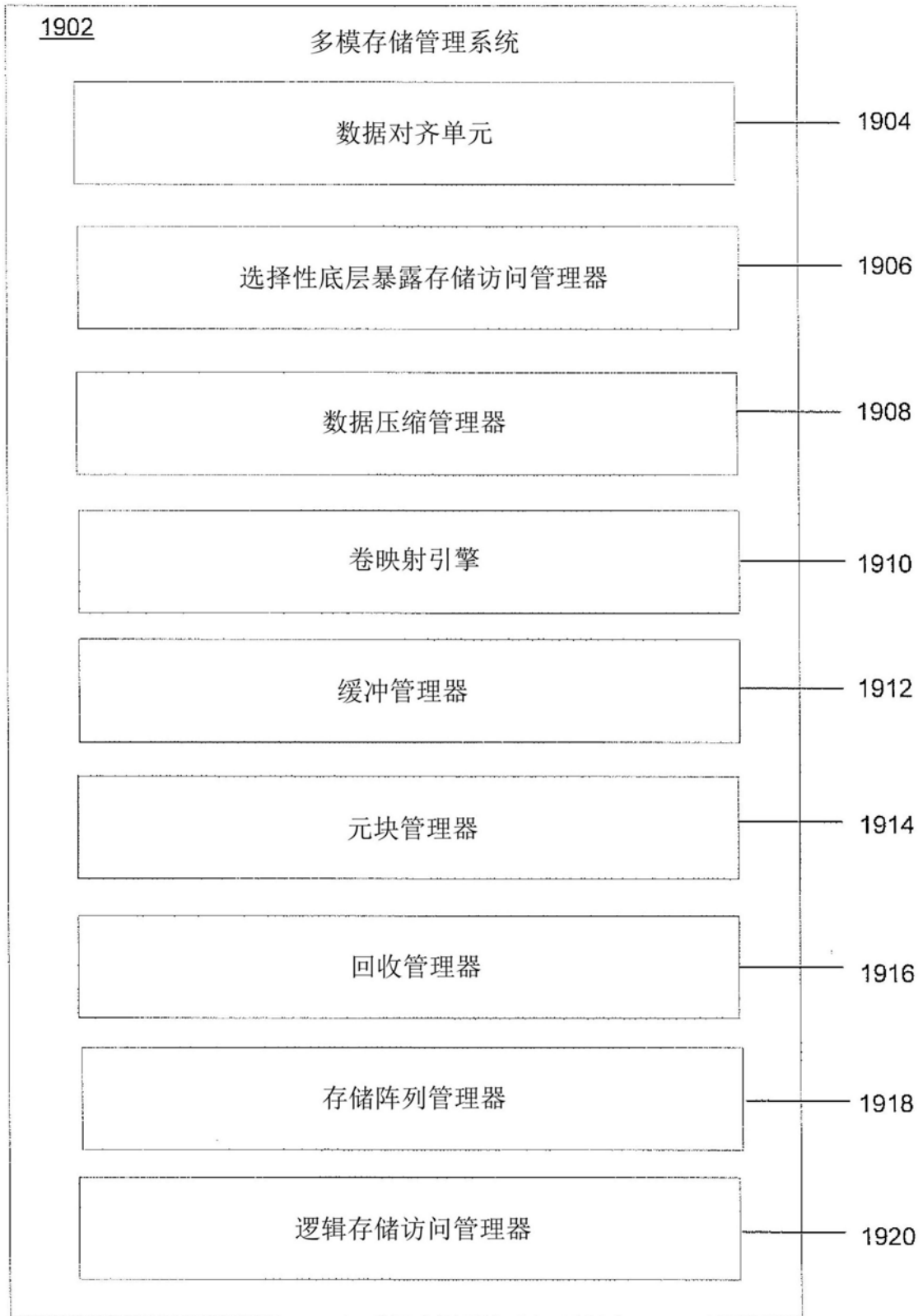


图19

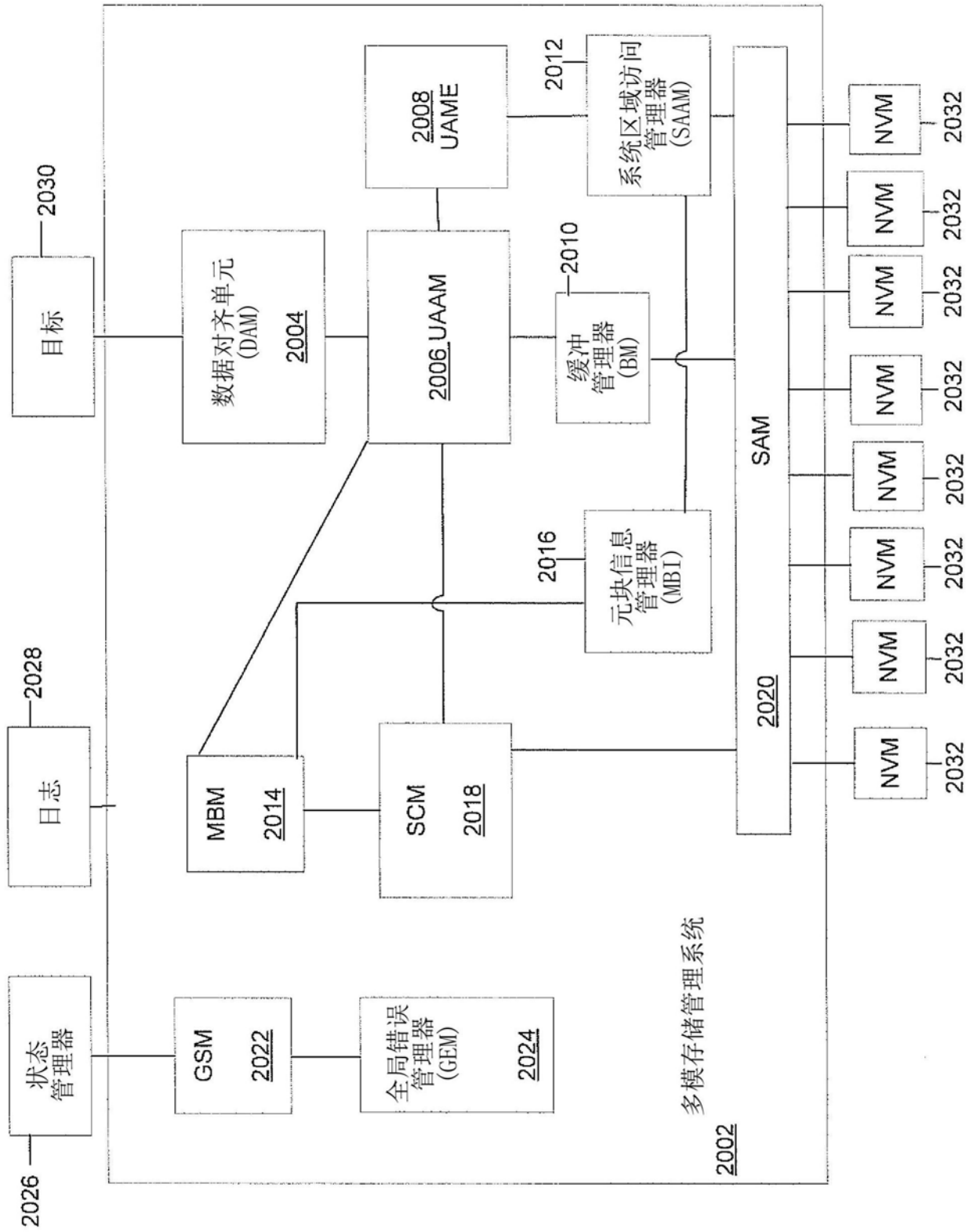


图20

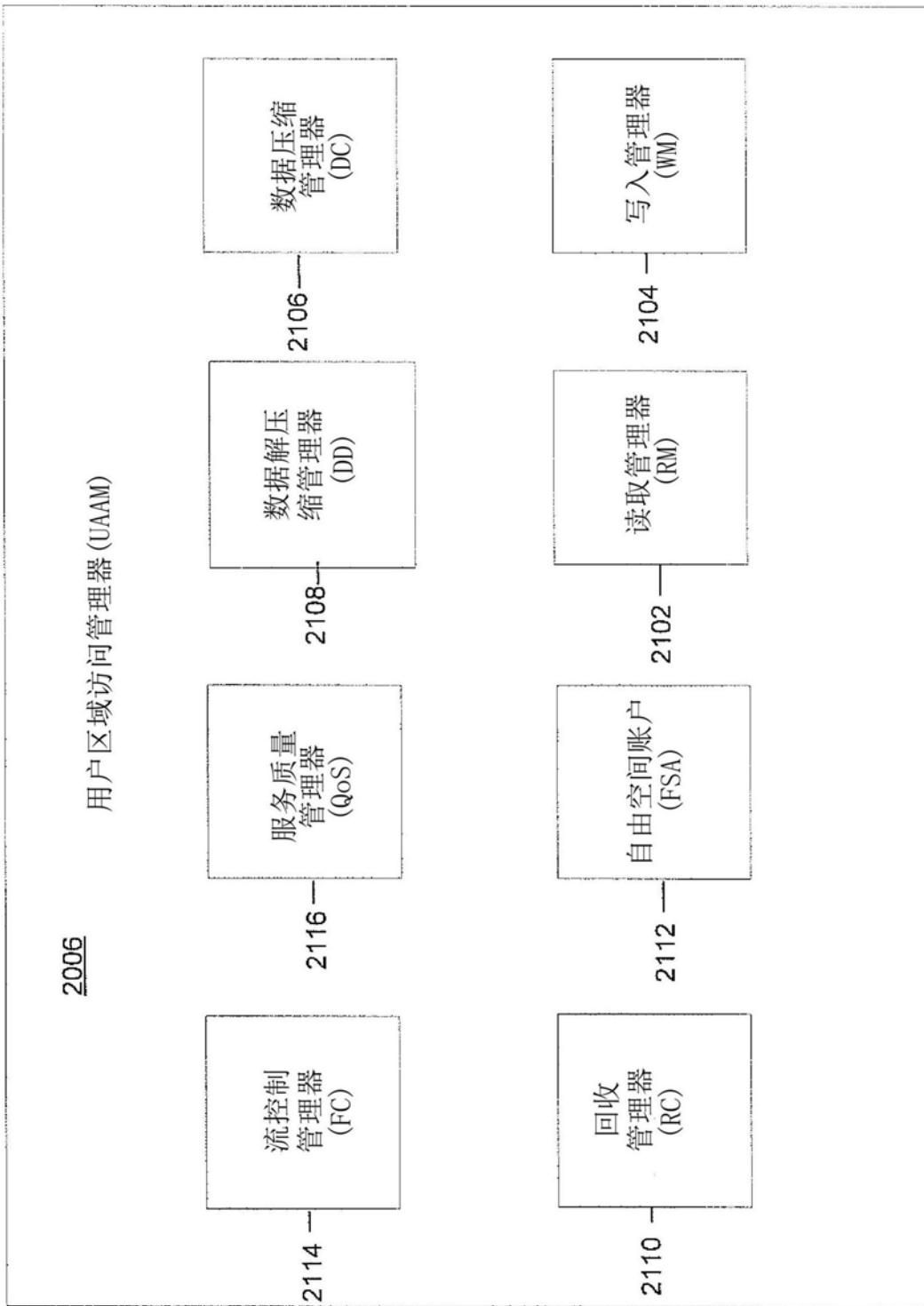


图21

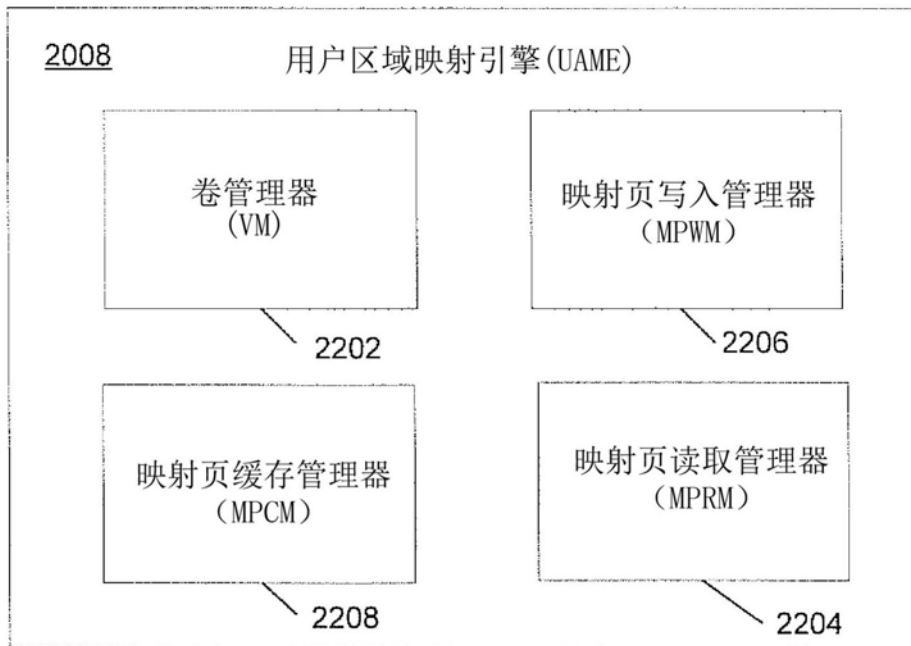


图22

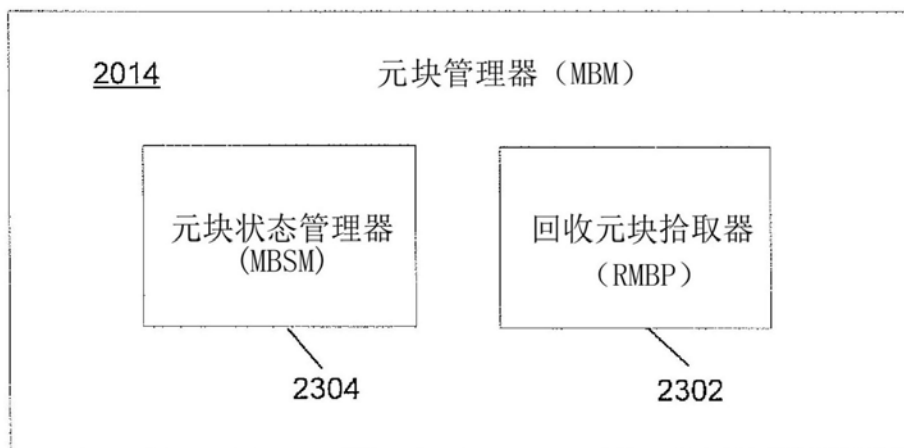


图23

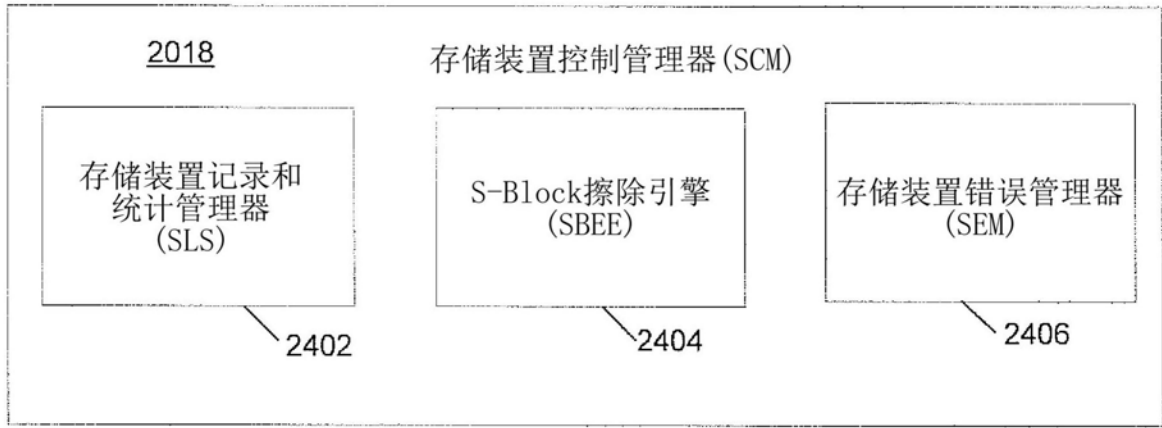


图24

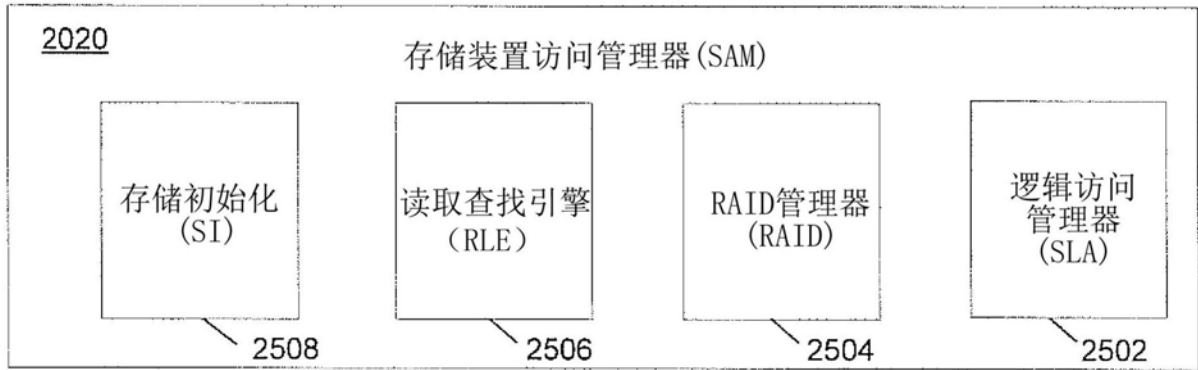


图25

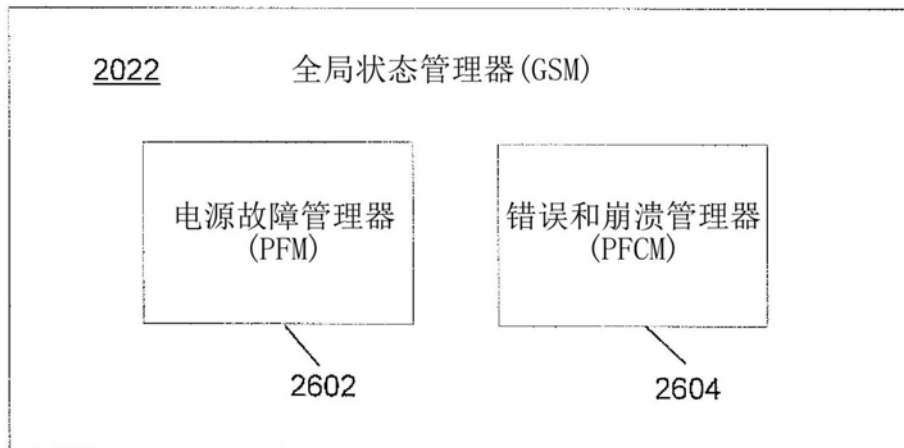


图26