



(12)发明专利申请

(10)申请公布号 CN 107168884 A

(43)申请公布日 2017. 09. 15

(21)申请号 201610592346.4

(22)申请日 2016.07.25

(30)优先权数据

2016-044261 2016.03.08 JP

(71)申请人 东芝存储器株式会社

地址 日本东京

(72)发明人 菅野伸一

(74)专利代理机构 北京律盟知识产权代理有限

责任公司 11287

代理人 张世俊

(51)Int.Cl.

G06F 12/02(2006.01)

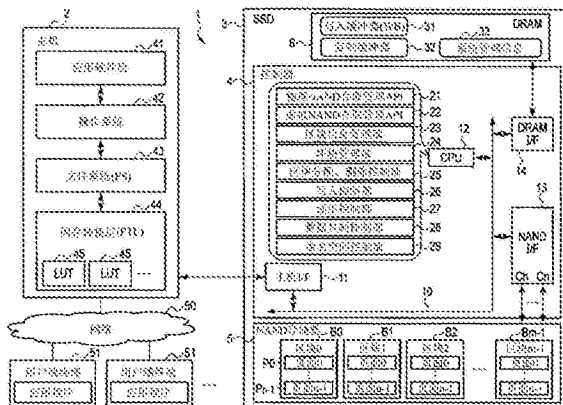
权利要求书3页 说明书47页 附图54页

(54)发明名称

存储系统、信息处理系统及非易失性存储器的控制方法

(57)摘要

本发明的实施方式实现能够取得非易失性存储器的控制相关的有用信息的存储系统、信息处理系统及非易失性存储器的控制方法。实施方式的存储系统执行将非易失性存储器的多个区块内的多个第1区块分配给第1命名空间的第1分配动作。所述控制器响应于来自主机的用来读出、写入或删除所述多个第1区块的一个的命令的接收,对所述多个第1区块的一个执行读出动作、写入动作、或删除动作,计数对所述多个第1区块执行的删除动作的次数,且响应于来自所述主机的请求取得与所述第1命名空间关联的删除次数的命令的接收,而将所述删除动作的次数的计数值通知给所述主机。



1. 一种存储系统,其特征在於具备:
非易失性存储器,包含多个区块;及
控制器,构成为电连接于所述非易失性存储器,
且执行将所述多个区块内的多个第1区块分配给第1命名空间的第1分配动作,
响应于来自主机的用来读出、写入或删除所述多个第1区块的一个的命令的接收,对所述多个第1区块的一个执行读出动作、写入动作、或删除动作,
计数对所述多个第1区块执行的删除动作的次数,
响应于来自所述主机的请求取得与所述第1命名空间关联的删除次数的命令的接收,
而将所述删除动作的次数的计数值通知给所述主机。

2. 根据权利要求1所述的存储系统,其特征在於:所述控制器构成为,
响应于来自所述主机的用来确保所述第1命名空间用的第1个数的区块的命令的接收,
将所述第1个数的区块确保用于所述第1命名空间,且能够分配给所述第1命名空间的所述多个第1区块的数的上限被限制在所述第1个数以下,
响应于来自所述主机的用来对所述第1命名空间用追加第2个数的区块的命令的接收,
将所述第2个数的区块作为应确保用于所述第1命名空间的区块进行追加,且能够分配给所述第1命名空间的所述多个第1区块的数的上限增加为所述第1个数与所述第2个数的和。

3. 根据权利要求1所述的存储系统,其特征在於:所述控制器执行第2分配动作,所述第2分配动作将所述多个区块内的各第2区块,作为能够写入与第2命名空间关联的数据的区块,分配给所述第2命名空间,

用来删除所述多个第1区块的一个的命令包含请求向所述第1命名空间的一个区块的分配及删除双方的命令,

所述第1分配动作包含:用来响应于请求所述区块的分配及删除双方的所述命令的接收,从所述第1命名空间及所述第2命名空间共有的共通空闲区块群中,分配一个区块用于所述第1命名空间的动作;将所述分配的区块删除的动作;及将指定所述分配且删除的区块的物理地址通知给所述主机的动作。

4. 根据权利要求3所述的存储系统,其特征在於:所述控制器构成为,
响应于来自所述主机的用来返还所述多个第1区块的一个的命令的接收,将所述多个第1区块的所述一个返还至所述共通空闲区块群,
响应于来自所述主机的用来返还所述多个第2区块的一个的命令的接收,将所述多个第2区块的所述一个返还至所述共通空闲区块群,且

所述第1分配动作包含选择所述共通空闲区块群之中具有最小删除次数的区块作为应分配给所述第1命名空间用的区块。

5. 根据权利要求4所述的存储系统,其特征在於:所述控制器构成为,
响应于来自所述主机的用来确保所述第1命名空间用的第1个数的区块的命令的接收,将所述第1个数的区块确保用于所述第1命名空间,能够分配给所述第1命名空间的所述多个第1区块的数的上限被限制为所述第1个数以下,
在执行了所述第1分配动作的情况下,将能够分配给所述第1命名空间用的区块的当前个数减去1,

在将所述多个第1区块的一个返还至所述共通空闲区块群的情况下,将能够分配给所

述第1命名空间用的区块的当前个数增加1。

6. 一种信息处理系统,其特征在于对包含含有多个区块的非易失性存储器的存储系统进行控制,且具备:

存储器;及

处理器,构成为执行所述存储器内储存的程序;

所述处理器构成为,

将用来对所述多个区块内分配给第1命名空间的多个第1区块的一个进行读出、写入或删除的命令发送至所述存储系统,且将请求取得与所述第1命名空间关联的删除次数的命令发送至所述存储系统,由此从所述存储系统取得通过计数对所述多个第1区块执行的删除动作的次数所得的计数值。

7. 根据权利要求6所述的信息处理系统,其特征在於:所述处理器构成为,

将用来确保所述第1命名空间用的区块的命令发送至所述存储系统,由此让所述存储系统确保第1个数的区块用于所述第1命名空间,能够分配给所述第1命名空间的所述多个第1区块的数的上限被限制为所述第1个数以下,

将用来追加应确保用于所述第1命名空间的区块的个数的命令,发送至所述存储系统,由此让所述存储系统将第2个数的区块作为应确保用于所述第1命名空间的区块进行追加,且能够分配给所述第1命名空间的所述多个第1区块的数的上限增加为所述第1个数与所述第2个数的和。

8. 一种非易失性存储器的控制方法,其特征在於其是包含多个区块的非易失性存储器的控制方法,且具备以下步骤:

执行将所述多个区块内的多个第1区块分配给第1命名空间的第1分配动作;

响应于来自主机的用来读出、写入或删除所述多个第1区块的一个的命令的接收,对所述多个第1区块的一个执行读出动作、写入动作、或删除动作;

计数对所述多个第1区块执行的删除动作的次数;及

响应于来自所述主机的请求取得与所述第1命名空间关联的删除次数的命令的接收,将所述删除动作的次数的计数值通知给所述主机。

9. 根据权利要求8所述的非易失性存储器的控制方法,其特征在於还包含以下步骤:

响应于来自所述主机的用来确保所述第1命名空间用的第1个数的区块的命令的接收,将所述第1个数的区块确保用于所述第1命名空间,且能够分配给所述第1命名空间的所述多个第1区块的数的上限被限制为所述第1个数以下;及

响应于来自所述主机的用来对所述第1命名空间用追加第2个数的区块的命令的接收,将所述第2个数的区块作为应确保用于所述第1命名空间的区块进行追加,且能够分配给所述第1命名空间的所述多个第1区块的数的上限增加为所述第1个数与所述第2个数的和。

10. 根据权利要求8所述的非易失性存储器的控制方法,其特征在於还具备以下步骤:执行第2分配动作,所述第2分配动作是将所述多个区块内的各第2区块,作为能写入与第2命名空间关联的数据的区块分配给所述第2命名空间;

用来将所述多个第1区块的一个删除的命令包含请求向所述第1命名空间的一个区块的分配及删除双方的命令,

所述第1分配动作包含:用来响应于请求所述区块的分配及删除双方的所述命令的接

收,从所述第1命名空间与所述第2命名空间共有的共通空闲区块群,分配一个区块用于所述第1命名空间的动作;将所述分配的区块删除的动作;及将指定所述分配且删除的区块的物理地址通知给所述主机的动作。

11.根据权利要求10所述的非易失性存储器的控制方法,其特征在于还具备以下步骤:

响应于来自所述主机的用来返还所述多个第1区块的一个的命令的接收,将所述多个第1区块的所述一个返还至所述共通空闲区块群;及

响应于来自所述主机的用来返还所述多个第2区块的一个的命令的接收,将所述多个第2区块的所述一个返还至所述共通空闲区块群;且

所述第1分配动作包含选择所述共通空闲区块群之中具有最小删除次数的区块,作为应分配给所述第1命名空间的区块。

12.根据权利要求11所述的非易失性存储器的控制方法,其特征在于还具备以下步骤:

响应于来自所述主机的用来确保所述第1命名空间用的第1个数的区块的命令的接收,将所述第1个数的区块确保用于所述第1命名空间,且能够分配给所述第1命名空间的所述多个第1区块的数的上限被限制为所述第1个数以下;

在执行了所述第1分配动作的情况下,将能够分配给所述第1命名空间用的区块的当前个数减去1;及

在将所述多个第1区块的一个返还至所述共通空闲区块群的情况下,将能够分配给所述第1命名空间用的区块的当前个数增加1。

存储系统、信息处理系统及非易失性存储器的控制方法

[0001] [相关申请]

[0002] 本申请享有以日本专利申请2016-44261号(申请日:2016年3月8日)为基础申请的优先权。本申请通过参照该基础申请而包含基础申请的全部内容。

技术领域

[0003] 本发明的实施方式涉及一种存储系统、信息处理系统及非易失性存储器的控制方法。

背景技术

[0004] 近年来,广泛利用有具备非易失性存储器的存储系统。作为此种存储系统之一,众所周知有基于NAND闪存技术的固态驱动器(SSD)。SSD因其低耗电、高性能等特征,而被用作各种计算机的主存储。

[0005] 最近,开始尝试利用主机来控制非易失性存储器。

[0006] 但是,为了能够让主机控制非易失性存储器,要求实现用来取得非易失性存储器的控制相关的有用信息的新功能。

发明内容

[0007] 本发明的实施方式提供一种能够取得非易失性存储器的控制相关的有用信息的存储系统、信息处理系统及非易失性存储器的控制方法。

[0008] 实施方式的存储系统具备包含多个区块的非易失性存储器、电连接于所述非易失性存储器的控制器。所述控制器执行将所述多个区块内的多个第1区块分配给第1命名空间的第1分配动作。所述控制器响应于来自主机的用来读出、写入或删除所述多个第1区块的一个的命令的接收,对所述多个第1区块的一个执行读出动作、写入动作、或删除动作,计数对所述多个第1区块执行的删除动作的次数,且响应于来自所述主机的请求取得与所述第1命名空间关联的删除次数的命令的接收,而将所述删除动作的次数的计数值通知给所述主机。

附图说明

[0009] 图1是表示包含实施方式的存储系统的信息处理系统的构成例的框图。

[0010] 图2是表示控制所述实施方式的存储系统的主机软件的框图。

[0011] 图3是表示所述实施方式的存储系统内的NAND接口与多个NAND存储器芯片的关系的框图。

[0012] 图4是表示由所述实施方式的存储系统管理的物理NAND存取管理应用程序接口(API)用的区块群、与虚拟NAND存取管理应用程序接口(API)用的区块群的图。

[0013] 图5是表示由所述实施方式的存储系统管理的虚拟区块的结构图。

[0014] 图6是表示在所述实施方式的存储系统内维持的管理各区块的删除次数的删除次

数管理表的图。

[0015] 图7是表示在所述实施方式的存储系统内维持的管理各虚拟区块的删除次数的另一删除次数管理表的图。

[0016] 图8是表示在所述实施方式的存储系统内维持的管理各命名空间的删除次数的又一删除次数管理表的图。

[0017] 图9是表示在所述实施方式的存储系统内维持的物理区块结构信息表的图。

[0018] 图10是表示在所述实施方式的存储系统内维持的虚拟区块结构信息表的图。

[0019] 图11是表示在所述实施方式的存储系统内维持的物理NAND存取管理API用的命名空间信息表的图。

[0020] 图12是表示在所述实施方式的存储系统内维持的虚拟NAND存取管理API用的命名空间信息表的图。

[0021] 图13是用来说明所述实施方式的存储系统内的虚拟区块与构成该虚拟区块的各区块的区块地址的关系的图。

[0022] 图14是用来说明由所述实施方式的存储系统执行的不良区块替换动作的图。

[0023] 图15是表示应用于所述实施方式的存储系统的物理NAND存取管理API用的坏块化命令的图。

[0024] 图16是表示应用于所述实施方式的存储系统的虚拟NAND存取管理API用的坏块化命令的图。

[0025] 图17是表示由所述实施方式的存储系统执行的坏块化处理的次序的流程图。

[0026] 图18是表示由所述实施方式的存储系统及主机执行的用于物理NAND存取及虚拟NAND存取的处理序列的图。

[0027] 图19是表示由所述实施方式的存储系统及主机执行的用于物理NAND存取的写入处理的处理序列的图。

[0028] 图20是表示由所述实施方式的存储系统及主机执行的用于虚拟NAND存取的写入处理的处理序列的图。

[0029] 图21是表示由所述实施方式的存储系统及主机执行的用于物理NAND存取的读出处理的处理序列的图。

[0030] 图22是表示由所述实施方式的存储系统及主机执行的用于虚拟NAND存取的读出处理的处理序列的图。

[0031] 图23是表示由所述实施方式的存储系统及主机执行的用于物理NAND存取的删除处理的处理序列的图。

[0032] 图24是表示由所述实施方式的存储系统及主机执行的用于虚拟NAND存取的删除处理的处理序列的图。

[0033] 图25是用来说明由所述实施方式的存储系统执行的命令优先级管理动作的图。

[0034] 图26是表示应用于所述实施方式的存储系统的区块分配、删除命令的图。

[0035] 图27是表示由所述实施方式的存储系统管理的使用中区块列表、空闲区块列表、使用中虚拟区块列表、及空闲虚拟区块列表的图。

[0036] 图28是表示由所述实施方式的存储系统及主机执行的区块分配、删除处理的处理序列的图。

- [0037] 图29是表示应用于所述实施方式的存储系统的物理NAND存取管理API用的写入命令的图。
- [0038] 图30是表示应用于所述实施方式的存储系统的虚拟NAND存取管理API用的写入命令的图。
- [0039] 图31是用来说明与向区块内的多个页面的数据写入顺序相关的制约的图。
- [0040] 图32是用来说明与从页面读出数据的时序相关的制约的图。
- [0041] 图33是对由所述实施方式的存储系统的NAND存储器执行的包含多个写入阶段的编程动作的例子进行说明的图。
- [0042] 图34是表示由所述实施方式的存储系统执行的数据写入处理的次序的流程图。
- [0043] 图35是表示响应于写入命令的命令完成回应的接收而由主机执行的处理的次序的流程图。
- [0044] 图36是表示响应于包含命名空间标识符(NSID)的写入命令的命令完成回应的接收而由主机执行的处理的次序的流程图。
- [0045] 图37是表示应用于所述实施方式的存储系统的物理NAND存取管理API用的读出命令的图。
- [0046] 图38是表示应用于所述实施方式的存储系统的虚拟NAND存取管理API用的读出命令的图。
- [0047] 图39是表示由所述实施方式的存储系统及主机执行的数据读出处理的处理序列的图。
- [0048] 图40是表示由所述实施方式的存储系统执行的数据读出处理的次序的流程图。
- [0049] 图41是对指定应复制的有效数据数作为结束条件时,由所述实施方式的存储系统执行的数据复制动作的例子进行说明的图。
- [0050] 图42是用来说明指定应检测的无效数据数作为结束条件时,由所述实施方式的存储系统执行的数据复制动作的例子图。
- [0051] 图43是用来说明指定多个复制源区块且指定应复制的有效数据数作为结束条件时,由所述实施方式的存储系统执行的数据复制动作的例子图。
- [0052] 图44是用来说明指定多个复制源区块且指定应检测的无效数据数作为结束条件时,由所述实施方式的存储系统执行的数据复制动作的例子图。
- [0053] 图45是用来说明数据大小比页面大小小且指定应复制的有效数据数作为结束条件时,由所述实施方式的存储系统执行的数据复制动作的例子图。
- [0054] 图46是用来说明数据大小比页面大小小且指定应检测的无效数据数作为结束条件时,由所述实施方式的存储系统执行的数据复制动作的例子图。
- [0055] 图47是表示应用于所述实施方式的存储系统的数据复制命令的输入参数的图。
- [0056] 图48是表示应用于所述实施方式的存储系统的数据复制命令的输出参数(返回值)的图。
- [0057] 图49是表示指定应复制的有效数据数作为结束条件时,由所述实施方式的存储系统执行的数据复制动作的次序的流程图。
- [0058] 图50是表示指定应检测的无效数据数作为结束条件时,由所述实施方式的存储系统执行的数据复制动作的次序的流程图。

- [0059] 图51是表示数据大小比页面大小小且指定应复制的有效数据数作为结束条件时,由所述实施方式的存储系统执行的数据复制动作的次序的流程图。
- [0060] 图52是表示数据大小比页面大小小且指定应检测的无效数据数作为结束条件时,由所述实施方式的存储系统执行的数据复制动作的次序的流程图。
- [0061] 图53是用来说明所述实施方式的存储系统的命名空间管理功能的图。
- [0062] 图54是表示所述实施方式的存储系统的命名空间管理架构的图。
- [0063] 图55是表示应用于所述实施方式的存储系统的命名空间分配命令的图。
- [0064] 图56是表示由所述实施方式的存储系统执行的命名空间分配处理的次序的流程图。
- [0065] 图57是表示应用于所述实施方式的存储系统的包含命名空间标识符(NSID)的区块分配、删除命令的图。
- [0066] 图58是表示响应于包含命名空间标识符(NSID)的区块分配、删除命令的接收而由所述实施方式的存储系统执行的区块分配、删除处理的次序的流程图。
- [0067] 图59是表示应用于所述实施方式的存储系统的命名空间用的删除命令的图。
- [0068] 图60是表示响应于包含命名空间标识符(NSID)的删除命令的接收而由所述实施方式的存储系统执行的删除处理的次序的流程图。
- [0069] 图61是表示应用于所述实施方式的存储系统的包含命名空间标识符(NSID)的区块返还命令的图。
- [0070] 图62是表示响应于包含命名空间标识符(NSID)的区块返还命令的接收而由所述实施方式的存储系统执行的区块返还处理的次序的流程图。
- [0071] 图63是表示应用于所述实施方式的存储系统的包含命名空间标识符(NSID)的删除次数获取命令的图。
- [0072] 图64是表示由所述实施方式的存储系统执行的删除次数通知处理的次序的流程图。
- [0073] 图65是说明主机的构成例的框图。
- [0074] 图66是表示包含所述实施方式的存储系统及主机的计算机的构成例的框图。

具体实施方式

- [0075] 以下,参照附图对实施方式进行说明。
- [0076] [系统构成]
- [0077] 首先,参照图1对包含一实施方式的存储系统的信息处理系统1的构成进行说明。
- [0078] 在信息处理系统1中,该存储系统能作为该信息处理系统1的主存储(外部存储装置)发挥作用。该存储系统构成为向非易失性存储器写入数据,并自非易失性存储器读出数据。该存储系统例如作为基于NAND闪存技术的固态驱动器(SSD)3而实现。SSD3是具备作为非易失性存储器的NAND闪存的存储装置。
- [0079] 信息处理系统1使用作为存储系统的SSD3来管理比如各种文件的数据。该信息处理系统1能作为构成为控制SSD3内的非易失性存储器的读出动作、写入动作、删除动作的计算机系统而发挥作用。
- [0080] 该信息处理系统1包含主机(主机装置)2、SSD3。主机2是构成为向SSD3保存数据的

信息处理装置。该信息处理装置的例子包含服务器计算机、个人计算机等。

[0081] SSD3可以内置于作为主机2发挥功能的信息处理装置,也可以经由缆线或网络连接于该信息处理装置。

[0082] 作为用来使主机2与SSD3相互连接的接口,能使用SCSI、Serial Attached SCSI (SAS)、ATA、Serial ATA(SATA)、PCI Express(PCIe)、NVM Express(NVMe)、Ethernet(注册商标)、Fibre channel等。

[0083] SSD3也可以具备控制器4、非易失性存储器(NAND存储器)5、及DRAM6。NAND存储器5并无限定,也可以包含多个NAND闪速存储器芯片。

[0084] NAND存储器5包含存储器单元阵列,该存储器单元阵列包含多个NAND区块(区块)B0~Bm-1。区块B0~Bm-1作为删除单位发挥功能。区块也被称为“物理区块”或“删除区块”。

[0085] 区块B0~Bm-1包含多个页面(物理页面)。也就是说,区块B0~Bm-1的分别包含页面P0~Pn-1。连接于同一字线的多个存储器单元是作为一个页面(物理页面)而编成。在NAND存储器5中,数据的读出及数据的写入是以页面单位执行。数据的删除是以包含多个页面的区块单位而执行。

[0086] 控制器4经由比如Toggle、ONFI的NAND接口13,电连接于作为非易失性存储器的NAND存储器5。控制器4也可以具有用来管理SSD3的物理资源、即NAND存储器5的物理资源管理功能。控制器4的物理资源管理功能能用来帮助主机2直接存取SSD3的物理资源。

[0087] 主机2为了能够直接控制及存取SSD3的物理资源,也可以在主机2中执行闪存转换层(FTL)43。主机2具有闪存转换层(FTL)43的系统构成,能够让主机2直接控制及存取SSD3的物理资源,且能够降低SSD3的处理负荷。

[0088] 闪存转换层(FTL)43也可以执行数据管理及NAND存储器5的区块管理。

[0089] 数据管理能包含(1)表示各逻辑区块地址(LBA)与NAND存储器5的各物理地址之间的对应关系的映射信息的管理、(2)用来隐藏以页面单位执行的读出/写入及以区块单位执行的删除动作的处理等。各LBA与各物理地址之间的映射的管理是使用作为逻辑物理地址变换表发挥功能的查找表(LUT)45而执行。

[0090] FTL44也可以支持管理多个命名空间的多命名空间功能。多命名空间功能为了能将一个存储装置(此处为SSD3)处理为宛如多个驱动器,而能管理与多个命名空间分别对应的多个逻辑地址空间(LBA空间)。对各命名空间分配有LBA范围(LBA0~LBA_n-1)。LBA范围的大小(即LBA的数)可以每个命名空间可变。各LBA范围从LBA0开始。

[0091] FTL44还可以通过使用创建的命名空间的数量及相同数量的查找表(LUT)45,按各命名空间管理各LBA与各物理地址之间的映射。

[0092] 在与某个特定的命名空间对应的查找表(LUT)45中,还可以管理关联于特定的命名空间的LAB范围内的各LBA与各物理地址之间的映射。LBA与物理地址之间的映射的管理是以特定的管理大小单位进行。特定的管理大小单位能够使用与系统设计相应的各种大小。特定的管理大小单位的例子并无限制,例如可为4K字节。

[0093] 与某个特定的LBA对应的物理地址表示储存该特定的LBA的数据的NAND存储器5内的位置(物理存储位置)。物理地址也可以通过区块地址与页面地址的组合来表现。区块地址是指定各个区块的地址,也称为“物理区块地址”或“区块编号”。页面地址是指定一个区

块内的各个页面的地址,也称为“物理页面地址”或“页面编号”。

[0094] 向页面的数据写入在1删除周期只能进行1次。换句话说,数据只能对删除状态的页面(可利用页面)写入。写入数据后的页面变成有效页面。

[0095] 另一方面,数据的最小删除单位是包含多个页面的区块。

[0096] 因此,FTL44将向相同LBA的写入(重写)映射至NAND存储器5上的不同页面。即,FTL44将数据(写入数据)无关于该数据的LBA而写入下一可利用页面。然后,FTL44更新查找表(LUT)45,将该LBA关联于与实际写入该数据的该页面对应的物理地址。将原本的页面(即关联该LBA的旧数据)无效化。

[0097] FTL44能够管理有效数据及无效数据。有效/无效数据也可以使用保持与各物理地址对应的有效/无效旗标的页面管理表进行管理。各有效/无效旗标可以特定的管理大小单位(例如4K字节)表示与各物理地址对应的数据有效还是无效。所谓数据有效是指此数据为最新数据。所谓数据无效是指此数据被更新(覆写)而不再利用。

[0098] 由FTL44执行的区块管理的例子中也可以包含损耗平均、及垃圾回收等。损耗平均是用来使各区块的删除次数平准化的动作。垃圾回收是用来作出NAND存储器5内的空闲空间的动作。该垃圾回收将混杂有效数据和无效数据的若干区块内的所有有效数据复制到另一区块(复制目标空闲区块)。并且,垃圾回收更新查找表(LUT)45,将复制的有效数据的各LBA映射至正确的物理地址。通过将有效数据复制到另一区块,仅剩无效数据的区块释放在空闲区块。由此,该区块在删除后能够再利用。

[0099] 主机2将比如写入命令、读出命令、删除命令之类的各种命令(存取请求)送出至SSD3。如上所述,在该信息处理系统1中,FTL44是在主机2上执行,因此所述各命令能够包含不指定LBA而是指定NAND存储器5上的位置的物理地址(区块地址、页面地址)。

[0100] 接着,对控制器4的构成进行说明。

[0101] 控制器4包含主机接口11、CPU12、NAND接口13、DRAM接口14等。所述CPU12、NAND接口13、DRAM接口14经由总线10而相互连接。

[0102] 主机接口11从主机2接收各种命令(写入命令、读出命令、删除命令等)。

[0103] CPU12是以控制主机接口11、NAND接口13、DRAM接口14的方式构成的处理器。CPU12执行管理NAND存储器5的物理资源管理处理、用来处理从主机2接收的各种命令的命令处理等。物理资源管理处理及命令处理也可以由CPU12执行的固件控制。

[0104] 该固件能够执行用来帮助主机2控制NAND存储器5的处理。

[0105] 该固件使CPU12作为物理NAND存取管理应用程序接口(API)21、虚拟NAND存取管理应用程序接口(API)22、区块信息管理部23、坏块管理部24、区块分配、删除控制部25、写入控制部26、读出控制部27、数据复制控制部28、命名空间控制部29发挥功能。

[0106] <物理NAND存取管理API与虚拟NAND存取管理API>

[0107] 所述物理NAND存取管理API21及虚拟NAND存取管理API22分别是与主机2和SSD3之间的通信相关的软件接口,主机2能够直接控制NAND存储器5内的区块。在物理NAND存取管理API21中,主机2对NAND存储器5内的区块的控制是以各个区块为单位、即物理区块为单位而执行。另一方面,在虚拟NAND存取管理API22中,主机2对NAND存储器5内的区块的控制基本上是以将多个区块(多个物理区块)汇集而得的区块组为单位执行。以下也将包含多个区块的区块组称为“虚拟区块”。

[0108] 物理NAND存取管理API21、虚拟NAND存取管理API22任一个之中,应存取的NAND存储器5上的位置能由来自主机2的命令所含的物理地址(区块地址、页面地址)来指定。

[0109] CPU12将NAND存储器5内的多个区块(物理区块)分类成多个第1区块、及多个第2区块。

[0110] 多个第1区块是物理NAND存取管理API21用的区块,这些区块是作为用来以单个区块(单一的物理区块)为单位存取NAND存储器5的区块而使用。

[0111] 多个第2区块是虚拟NAND存取管理API22用的区块,且是作为分别包含多个区块的多个区块组(多个虚拟区块)而编成的区块。这些第2区块是作为用来以将多个区块(多个物理区块)汇集而得的区块组为单位存取NAND存储器5的区块来使用。

[0112] 当主机2使用物理NAND存取管理API21存取SSD3的情况下,CPU12响应于来自主机2的包含指定多个第1区块内的一个区块(物理区块)的物理地址的第1读出、写入或删除命令的接收,而执行对该第1区块的读出动作、写入动作、或删除动作。

[0113] 当主机2使用虚拟NAND存取管理API22存取SSD3的情况下,CPU12响应于包含指定多个虚拟区块内的一个虚拟区块的物理地址的第2读出、写入或删除命令的接收,而执行对一个虚拟区块的读出动作、写入动作、或删除动作。

[0114] 在该虚拟NAND存取管理API22中,能不以单一的区块为单位而是以构成虚拟区块的多个区块的汇集为单位执行读出动作、写入动作、或删除动作。由此,虚拟NAND存取管理API22能够作为能高速地读出/写入/删除比如用户数据这样大小相对较大的数据的接口而利用。例如,当页面大小为16K字节,一个虚拟区块包含4个区块(4个物理区块)的情况下,能够实现最大64K字节的带宽。此外,例如当页面大小为16K字节一个虚拟区块包含8个区块的情况下,能够实现最大128K字节的带宽。

[0115] 另一方面,在物理NAND存取管理API21中,执行对单一区块的读出动作、写入动作、或删除动作。物理NAND存取管理API21能保证的最大带宽比虚拟NAND存取管理API22能保证的最大带宽窄。相反地,物理NAND存取管理API21能以比虚拟NAND存取管理API22小的粒度控制读出/写入/删除。由此,物理NAND存取管理API21作为例如将比如元数据这样的大小相对较小的数据配置在NAND存储器5上的所需位置的数据配置控制、及以较小的数据大小的粒度进行存取(读出/写入/删除)的接口而使用。

[0116] 主机2可以从SSD3预先取得指定物理NAND存取管理API21用的各区块的物理地址、及指定虚拟NAND存取管理API22用的各虚拟区块的物理地址,或者也可以通过向SSD3请求某一个区块或某一个虚拟区块的分配而从SSD3取得指定某一个区块的物理地址或指定某一个虚拟区块的物理地址。

[0117] 一个虚拟区块所含的多个区块的组合并无限定,例如可由能够并列(同时)存取的多个区块(多个物理区块)形成一个虚拟区块。

[0118] 例如,当NAND接口13具有多个信道,且各信道连接有1个以上的NAND闪存芯片的情况下,可以利用从分别连接于不同信道的NAND闪存芯片逐一选择的信道数量的区块来构成一个虚拟区块。由此,能够保证与NAND存储器5的结构上的最大带宽对应的存取速度。

[0119] 元数据也可为文件管理信息。文件管理信息也可以包含表示文件内的数据的存储位置的数据、表示该文件的创建日期的数据、表示该文件更新日期数据、或表示该文件最后读出日期的数据中至少一个。

[0120] 物理NAND存取管理API21用的读出命令、写入命令、删除命令也可以具有与虚拟NAND存取管理API22用的读出命令、写入命令、删除命令不同的操作代码。

[0121] <区块信息管理>

[0122] 区块信息管理部23能够管理与NAND存储器5内的各个区块及各个虚拟区块相关的信息,并将与各个区块及各个虚拟区块相关的信息提供给主机2。与各个区块及各个虚拟区块相关的信息也可以包含各个区块各自的删除次数、各个虚拟区块各自的删除次数。

[0123] 在数据中心等,连接于某个服务器计算机的SSD有更换成在另一服务器计算机等中过去被使用的另一SSD的情况。

[0124] 当SSD3从某个服务器计算机移设至另一服务器计算机的情况下,SSD3的区块信息管理部23能够将与NAND存储器5内的各个区块及各个虚拟区块的使用历史记录相关的信息(当前删除次数等)提供给该另一服务器计算机。由此,该另一服务器计算机能够在考虑了SSD3的过去的历史记录的状态下正确地理解各个区块或各个虚拟区块的实际的删除次数,因此例如能够基于从SSD3取得的删除次数的信息而精度良好地进行损耗平均用的处理等。

[0125] <坏块管理>

[0126] 坏块管理部24能够执行用来管理由初级缺陷列表(也称为“工厂出货缺陷列表”)指定的无法使用的不良区块(初级坏块)、及系统运行中由主机2指定的不良区块(成长坏块)的处理。

[0127] 例如,当由初级缺陷列表或主机2指定某个虚拟区块内的一个区块为无法使用的不良区块(坏块)的情况下,坏块管理部24也可以执行将虚拟区块内的坏块替换成虚拟NAND存取管理API22用的另一区块的处理。该另一区块可以从坏块以外的该虚拟区块内的其他所有区块及能并列存取的区块群中选择。

[0128] <区块分配、删除>

[0129] 区块分配、删除控制部25能够管理包含有效数据的各区块和不包含有效数据的各区块(空闲区块)。所谓包含有效数据的区块是指用户经由主机2利用SSD3而使用的区块。当若干用户终端51经由网络50而连接于主机2的情况下,这些用户终端51的用户可为经由主机2利用SSD3的用户。所谓不包含有效数据的区块是指哪个用户都无法使用的区块。

[0130] 而且,区块分配、删除控制部25还能够管理包含有效数据的各虚拟区块及不包含有效数据的各虚拟区块(空闲虚拟区块)。

[0131] 当从主机2请求一个区块的分配的情况下,区块分配、删除控制部25可以从空闲区块中将一个区块(物理区块)分配给主机2,然后将指定该分配的区块的物理地址(区块地址)通知给主机2。之后,主机2能使用该通知的物理地址对该分配的区块进行存取(读出/写入/删除)。

[0132] 当从主机2请求一个虚拟区块的分配的情况下,区块分配、删除控制部25可以从空闲虚拟区块中将一个虚拟区块分配给主机2,然后将通知该分配的虚拟区块的物理地址(虚拟区块地址)通知给主机2。之后,主机2能使用该通知的物理地址对该分配的虚拟区块进行存取(读出/写入/删除)。

[0133] 用来存取虚拟区块的物理地址也可以包含虚拟区块地址、虚拟区块内区块编号、及页面地址。虚拟区块地址指定分配给各虚拟区块的虚拟区块编号的任一个。虚拟区块内

区块编号指定存取对象区块是虚拟区块内的第几个区块。页面地址表示存取对象区块内的页面编号。

[0134] 该区块分配、删除控制部25能够在主机2自身不管理包含有效数据的区块及不含有效数据的区块的情况下,让主机2取得不含有效数据的区块。由此,能够降低主机2所致的NAND存储器5的管理成本。

[0135] 区块(或虚拟区块)的分配请求也可为只请求区块(或虚拟区块)的分配的命令。或者,也可以使用请求区块(或虚拟区块)的分配及区块(或虚拟区块)的删除两者的命令、即区块分配、删除命令、或虚拟区块分配、删除命令。

[0136] 区块分配、删除命令是将请求区块分配的命令的功能及请求区块删除的命令的功能组合而成的一个命令。同样地,虚拟区块分配、删除命令是将请求虚拟区块分配的命令的功能及请求虚拟区块删除的命令的功能组合而成的一个命令。

[0137] 响应于来自主机2的区块分配、删除命令的接收,区块分配、删除控制部25从空闲区块中将一个区块分配给主机2,然后自动删除该分配的区块,并将指定该分配及删除的区块的物理地址(区块地址)通知给主机2。

[0138] 有各空闲区块虽然不含有效数据但保持着某一用户过去写入的旧数据的情况。因此,通过自动删除分配区块的所述功能,能够预先防止用户数据的泄漏。此外,主机2能够不用将分配区块删除用的删除命令发送给SSD3,而立即开始向该分配区块的数据写入。

[0139] 响应于来自主机2的虚拟区块分配、删除命令的接收,区块分配、删除控制部25从空闲虚拟区块中将一个虚拟区块分配给主机2,然后自动删除该分配的虚拟区块,并将指定该分配及删除的虚拟区块的物理地址(虚拟区块地址)通知给主机2。

[0140] <写入控制>

[0141] 写入控制部26接收包含指定某个特定区块的区块地址及指定该特定区块内的多个页面内的某个特定页面的页面地址的写入命令,并将该写入命令所指定的数据写入特定区块内的特定页面(直接地址指定模式)。写入控制部26支持物理NAND存取管理API21及虚拟NAND存取管理API22的双方。在物理NAND存取管理API21中,特定区块是特定的物理区块,在虚拟NAND存取管理API22中,特定区块是特定的虚拟区块内的存取对象区块。

[0142] 写入控制部26具有将保持可读出数据的区块内的最新的页面通知给主机2的“可读出页面通知功能”。根据作为NAND存储器5使用的NAND存储器的种类,有即便向某个区块的最初的页面写入数据,在向该区块内的后续若干页面进一步写入数据之前,无法从该区块的最初的页面正常读出数据的情况。若向最初的页面后续的若干页面写入数据前该最初的页面已被主机2读出存取,有ECC无法订正的错误数据从最初的页面被读出,向主机2返回表示读出错误的状态的情况。最初的页面的数据即便在向后续若干页面写入数据后能正常读出,主机2也有可能将该读出错误误认为是物理存储器缺陷引起的。

[0143] 同样地,有写入该区块的第2个页面的数据在向该第2个页面后续的若干页面写入数据之前,无法正常读出的情况。各页面的数据的可读出时序因NAND存储器而不同。

[0144] 写入控制部26能够将保持可读出数据的区块内的最新的页面通知给主机2。例如,当向写入对象区块内的若干页面写入数据而写入对象区块内的最初的页面变得可读出时,写入控制部26可以将最初的页面的页面地址作为保持可读出数据的最新的页面通知给主机2。当向写入对象区块进一步写入而写入对象区块内的第2个页面变得可读出时,写入控

制部26可以将第2个页面的页面地址作为保持可读出数据的最新的页面通知给主机2。

[0145] 更详细来说,写入控制部26执行以下动作。

[0146] 写入控制部26从主机2接收包含指定NAND存储器5的多个区块内的第1区块的区块地址及指定该第1区块内的多个页面内的第1页面的页面地址的写入命令。写入控制部26将该写入命令所指定的数据写入至第1区块内的第1页面。并且,写入控制部26将表示向第1页面写入数据前由主机2写入数据的第1区块内的页面群之中,因向第1页面写入数据而变得可读出的最新的页面的页面地址,通知给主机2。

[0147] 主机2基于该通知能够识别出写入数据的区块内的哪个页面为止为可读出状态。由此,利用该“可读出页面通知功能”,能够帮助主机2直接存取NAND存储器5。

[0148] 写入控制部26还具有“不当写入顺序警告”功能。“不当写入顺序警告”功能是如下功能:当主机2未遵守区块内必须从最初的页面向最后的页面连续写入数据这样的写入顺序制约的情况下,向主机2返回不当写入顺序的警告。

[0149] 主机2能够指定应写入数据的物理地址(区块地址及页面地址)。这意味着主机2有以错误写入顺序执行写入存取的可能性。利用该“不当写入顺序警告”功能,能够支援主机2直接控制向NAND存储器5的写入。

[0150] 更详细来说,写入控制部26执行以下动作。

[0151] 写入控制部26从主机2接收包含指定NAND存储器5的多个区块内的第1区块的区块地址、及指定第1区块内的多个页面内的第1页面的页面地址的写入命令。写入控制部26基于该写入命令内的页面地址,判定该写入命令是否满足按照第1区块内的多个页面内的最初的页面到最后的页面的顺序写入数据这样的写入顺序相关的制约。当该写入命令满足写入顺序相关的制约的情况下,写入控制部26将由该写入命令指定的数据向第1区块内的第1页面写入数据。当该写入命令不满足写入顺序相关的制约的情况下,写入控制部26并非将该写入命令指定的数据向第1区块内的第1页面写入数据,而是将包含不当写入顺序的警告的命令完成回应通知给主机2。

[0152] 写入控制部26除了所述“直接地址指定模式”以外,还可以进一步支持“自动地址产生模式”。

[0153] 直接地址指定模式是主机2直接指定应写入数据的NAND存储器5内的区块及该区块内的页面双方的写入模式。在该直接地址指定模式中,主机2将包含区块地址及页面地址双方的写入命令发送给SSD3。

[0154] 另一方面,自动地址产生模式是主机2只指定应写入数据的NAND存储器5上的区块的写入模式。该指定区块可为物理区块,也可为虚拟区块内的存取对象区块。

[0155] 自动地址产生模式中利用的写入命令只包含区块地址,不包含页面地址。该指定区块内的写入对象页面的页面地址由SSD3自动发布。表示写入数据的页面的页面地址、即自动发布的页面地址从SSD3通知给主机2。

[0156] 所述“可读出页面通知功能”能在自动地址产生模式下利用。自动地址产生模式中,写入控制部26能够执行以下动作。

[0157] 写入控制部26从主机2接收包含区块地址但不包含页面地址的写入命令。写入控制部26按照从区块地址指定的第1区块内的多个页面内的最初的页面到最后的页面的顺序写入数据这样的写入顺序,自动发布指定第1区块内的多个页面内的下一可利用页面的页

面地址。写入控制部26将写入命令指定的数据写入第1区块内的下一可利用页面(自动发布的页面地址指定的页面)。写入控制部26将表示在向下一可利用页面写入数据前由主机2写入数据的第1区块内的页面群之中,因向下一可利用页面写入数据而变得可读出的最新的页面的页面地址,通知给主机2。

[0158] 写入控制部26还具有将写入数据的页面已达到当前写入对象区块的最后的页面的意旨通知给主机2的功能。根据该通知,主机2能够认识到必须分配新的区块。

[0159] 写入控制部26还具有将写入数据的当前写入对象区块内的页面数已达到特定页面数的意旨通知给主机2的功能。主机2有期望向各区块内的特定页面、例如最后的页面写入特定的管理信息(例如元数据)的情况。因此,通过将当前写入对象区块内写入数据的页面数已达到“特定页面数”的意旨通知给主机2,能够帮助主机2执行向该写入对象区块的最后的页面等写入特定的管理信息之类的作业。“特定页面数”能够由来自主机2的写入命令而指定。

[0160] <读出控制>

[0161] 当读出控制部27从主机2接收读出命令时,读出控制部27从该读出命令内的区块地址及页面地址指定的区块内的页面读出数据。该指定区块可为物理NAND存取管理API21用的物理区块,也可为虚拟NAND存取管理API22用的虚拟区块内的存取对象区块。

[0162] <数据复制控制>

[0163] 数据复制控制部28执行数据复制动作,用来支援主机2执行垃圾回收。数据复制动作是基于从主机2接收的数据复制命令,而在SSD3本地执行。即,用来从NAND存储器5的特定的复制源区块向特定的复制目标区块复制数据的数据传输动作,并不经由主机2,而是在SSD3内执行。

[0164] 因此,即便不进行将从NAND存储器5的特定的复制源区块读出的数据传输到主机2的存储器,且将该数据内的有效数据从主机2的存储器写入至NAND存储器5的特定的复制目标区块这样的处理,也能在SSD3内本地执行将垃圾回收用所必要的有效数据聚集于NAND存储器5内的特定区块的动作。

[0165] 数据复制命令能够指定复制源区块、复制源区块内的复制开始页面、复制目标区块、复制目标区块内的传输开始页面、复制结束条件(向复制目标区块应复制的有效数据数、或复制结束前应检测的无效数据数)。于此,有效数据数可为有效页面数,且无效数据数可为无效页面数。复制源区块可为物理NAND存取管理API21用的物理区块,也可为虚拟NAND存取管理API22用的虚拟区块内的一个区块。同样地,复制目标区块也是既可为物理NAND存取管理API21用的物理区块,也可为虚拟NAND存取管理API22用的虚拟区块内的一个区块。

[0166] 复制开始页面表示复制源区块内的复制对象的最初的页面。传输开始页面表示复制目标区块内的传输对象的最初的页面。通过指定所述复制开始页面及传输开始页面,能够执行将复制源区块内的任意的页面范围内的有效数据向复制目标区块内的任意的页面范围内复制(移动)这样极细微的复制动作。

[0167] 此外,在数据复制动作中,数据复制控制部28自动跳过无效数据的复制,而只将特定页面范围内的有效数据复制到复制目标区块。由此,主机2不用指定哪个数据能复制到何处,便能执行垃圾回收用所必须之有效数据的复制。

[0168] <命名空间控制>

[0169] 命名空间控制部29能够支持用来个别地处理多个命名空间的多命名空间功能。多命名空间功能能够管理分别分配有多个逻辑地址空间(LBA空间)的多个命名空间,以便能够将NAND存储器5逻辑分割成多个区域。各命名空间作为NAND存储器5内的一个区域发挥功能。与某个特定命名空间关联的数据被写入对该特定的命名空间分配的区块群。

[0170] 命名空间控制部29支持命名空间分配命令等。

[0171] 命名空间分配命令是对SSD3请求各个命名空间应确保的区块的个数。应确保的区块的个数可为物理NAND存取管理API21用的物理区块的个数,也可为虚拟NAND存取管理API22用的虚拟区块的个数。

[0172] 响应于来自主机2的命名空间分配命令的接收,命名空间控制部29能够对特定的命名空间确保(分配)由主机2指定的个数的区块。

[0173] 命名空间分配命令能够让主机2(主机软件)确保适于主机2内的工作负载(work load)的个数的区块用于各个命名空间。例如,关于较多使用随机写入存取的工作负载关联的命名空间,也可以确保相当于比该命名空间用的逻辑区块地址(LBA)的数对应的容量更多容量的个数的区块。例如,若与某个命名空间用的逻辑区块地址(LBA)的数(LBA范围)对应的容量为100G字节,对该命名空间用确保相当于150G字节的个数的区块,便能确保具有与LBA范围对应的容量(用户空间的容量)的50%的大小的过量供给区域。

[0174] 过量供给意味着向主机2分配从主机2侧不会被视作用户空间(用户可存取LBA空间)的存储容量。分配了从主机2侧不会被视作用户可存取LBA空间的存储容量的空间为过量供给区域。通过过量供给,能够将超过用户可存取LBA空间(用户区域的容量)的容量的区块群分配给特定的命名空间。

[0175] 关于处理更新频率高的数据的命名空间,数据覆写多次发生,因此有大量区块碎片化的情况。结果,垃圾回收的执行次数增加,由此写入扩增增加,使得各区块的删除次数也增加。删除次数的增加成为引发SSD3的耐久性及寿命劣化的主要原因。

[0176] 关于分配了较多过量供给区域的命名空间,能够延缓垃圾回收的开始时序。例如,假设对具有100G字节的用户空间(LBA范围)的特定的命名空间,确保相当于150G字节的个数的区块(或虚拟区块)的情况。这种情况下,即便当100G字节量的个数的区块写满数据,结果使得这些区块变成无可删除区块而不含可利用页面的状态时,也能代替这些区块而将对应于过量供给区域的区块分别用于数据写入。由此,能够延缓执行该特定的命名空间用的垃圾回收动作的时序。随着向过量供给区域的区块群写入数据,用户空间的区块群内的数据有可能因其更新而被无效化。所有数据被无效化的区块不进行垃圾回收便能再利用。由此,通过使过量供给区域的大小最佳化,能够抑制删除次数的增加。

[0177] 而且,命名空间控制部29能够按各个命名空间计数删除次数(也称为合计删除次数),并将计数的各命名空间的删除次数通知给主机2。

[0178] 主机2(主机2的管理者)能够将与特定的命名空间对应的合计删除次数,作为用来由该特定的命名空间(使用该命名空间的 用户)判定NAND存储器5消耗了多少的指标而利用。主机2能够将与命名空间分别对应的合计删除次数活用于所述多个命名空间的管理。

[0179] 例如,关于合计删除次数较多的命名空间,主机2可以向该命名空间追加应确保的区块的数。这种情况下,主机2可以通过向SSD3发送命名空间分配命令,而向SSD3请求特定数的区块的追加。通过向该命名空间追加特定数的区块而该命名空间的过量供给区域的

量增加,因此能够抑制与该命名空间对应的写入扩增,结果能够使SSD3的寿命最大化。

[0180] 在数据中心等,有进行向用户有偿租用各存储空间的租借服务的情况。这种情况下,NAND存储器5被逻辑分割成与多个命名空间分别对应的多个区域(存储空间)。某个用户对关联于某个命名空间的标识符的存储空间进行存取,另一用户对关联于另一命名空间的标识符的另一存储空间进行存取。数据中心提供者也可以将各个命名空间的合计删除次数、即各个命名空间的消耗度反映到各个存储空间的使用费(租借费)上。例如,对于使用合计删除次数非常多的命名空间的用户,除了缴纳由该命名空间的容量(相当于确保该命名空间所需的区块的数的容量)而决定的基本使用费以外,还要缴纳与消耗度相应的追加费。

[0181] 接着,对控制器4内的其他组件进行说明。

[0182] NAND接口13是在CPU12的控制之下以控制NAND存储器5的方式构成的NAND控制器。NAND接口13可以具有多个信道。各信道连接有若干NAND存储器芯片。控制器4能够对连接于NAND接口13的不同信道的多个NAND存储器芯片并列进行存取。

[0183] DRAM接口14是在CPU12的控制之下以控制DRAM6的方式构成的DRAM控制器。

[0184] DRAM6的存储区域的一部分可以作为临时储存应写入NAND存储器5的数据的写入缓冲器(WB)31而利用。此外,DRAM6的存储区域可以作为临时储存数据复制动作中从复制源区块读出的数据的复制缓冲器32而利用。此外,DRAM6的存储区域可以用于储存包含各种管理表的系统管理信息33。系统管理信息33可以在SSD3的电源接通时从NAND存储器5加载到DRAM6。当SSD3的电源须断开时,可以将更新后的系统管理信息33保存到NAND存储器5。系统管理信息33可以包含NAND存储器5的区块结构(或虚拟区块结构)、及比如删除次数之类的NAND存储器5的使用历史记录信息。

[0185] 接着,对主机2的构成进行说明。

[0186] 主机2是能够执行各种程序的信息处理装置。由主机2执行的程序包含应用软件层41、操作系统(OS)42、文件系统43、及所述FTL44。

[0187] 众所周知,操作系统(OS)42一般来说是构成为执行以下控制的软件:管理主机2全体,控制主机2内的硬件,让应用程序及各用户端终端51能够使用主机2的硬件及SSD3。

[0188] 文件系统43的作用是执行文件操作(创建、保存、更新、删除等)用的控制。例如,文件系统43可以使用ZFS(Zettabyte File System,zetta字节文件系统)、Btrfs、XFS、ext4、NTFS(New Technology File System,新技术文件系统)等。或者,文件系统43还可以使用文件目标系统、Key Value Store System。

[0189] 各种应用软件线程是在应用软件层41上运行。作为应用软件线程的示例,有用户端软件、数据库软件、虚拟机等。

[0190] 在应用软件层41需要向SSD3送出比如读出请求、写入请求、删除请求这样的请求时,应用软件层41会向OS42送出此请求。OS42将此请求经由文件系统43发送给FTL44。FTL44将此请求转码成命令(读出命令、写入命令、删除命令等)。这种情况下,FTL44进行逻辑物理地址变换,将请求所含的LBA变换成NAND存储器5的物理地址。可以执行与多个命名空间分别对应的多个FTL44。这种情况下,各逻辑地址(LBA)与各物理地址之间的映射的管理可以使用各命名空间不同的LUT45来执行。

[0191] FTL44将命令送出至SSD3。当接收来自SSD3的回应时,FTL44将此回应经由文件系统43而送出至OS42。OS42将此回应送出至应用软件层41。

[0192] [主机软件]

[0193] 图2表示在主机2上执行的软件(主机软件)的例子。

[0194] 有某个应用程序处理多个文件的情况。关于各个文件,通过连续写入来执行数据写入用的写入存取。若与这些文件分别对应的连续写入由一个FTL44合并,则这些连续写入各自的写入目标LBA混杂在一起。因此,合并的连续写入有可能作为随机写入被送至SSD3。随机写入的增加会变成SSD3的写入扩增增加的原因。

[0195] 如图2所示,在主机2中,与这些文件分别对应的若干FTL44能同时执行。

[0196] 在图2的例子中,假定执行4个FTL44A、FTL44B、FTL44C、FTL44D的情况。所述FTL44A、FTL44B、FTL44C、FTL44D能够相互独立第执行动作。例如,FTL44A管理对应于NSID#1的各LBA与各物理地址之间的映射,FTL44B管理对应于NSID#2的各LBA与各物理地址之间的映射,FTL44C管理对应于NSID#3的各LBA与各物理地址之间的映射,FTL44D管理对应于NSID#4的各LBA与各物理地址之间的映射。

[0197] 用来写入与NSID#1关联的数据(例如,文件“A”的数据)的若干写入请求经由文件系统43A而送至FTL44A。该FTL44A将对应于所述若干写入请求的若干写入命令送出至SSD3。SSD3能够将由所述多个写入命令指定的数据写入分配给NSID#1的区块。由此,能够将与某个LBA范围关联的文件“A”的数据连续写入分配给NSID#1的区块。

[0198] 同样地,用来写入与NSID#2关联的数据(例如,文件“B”的数据)的若干写入请求经由文件系统43B而送至FTL44B。该FTL44B将与所述若干写入请求对应的若干写入命令送出至SSD3。SSD3能够将由所述多个写入命令指定的数据写入至分配给NSID#2用的区块。

[0199] 因此,能够防止用来写入文件“A”的数据的连续写入与用来写入文件“B”的数据的连续写入合并,所以能够抑制SSD3的写入扩增的增加。

[0200] [信道]

[0201] 图3表示NAND接口13与多个NAND存储器芯片的关系。

[0202] 在图3中,例示了在NAND接口13具有的8个信道(Ch#1~Ch#8)上分别连接4个NAND存储器芯片的情况。在控制器4的控制之下,NAND接口13同时驱动连接于8个信道(Ch#1~Ch#8)的8个NAND存储器芯片,由此能够并列(同时)执行对8个区块的读出、写入、删除。

[0203] [物理区块与虚拟区块]

[0204] 图4表示由SSD3管理的物理NAND存取管理API21用的区块群与虚拟NAND存取管理API22用的区块群。

[0205] 于此,假定NAND接口13具有4个信道(Ch.A~Ch.D)的情况。

[0206] 在Ch.A~Ch.B上分别连接着一个以上的NAND存储器芯片。连接于各信道的一个以上的NAND存储器芯片包含多个区块、例如111个区块(区块地址0~110)。

[0207] SSD3的控制器4将NAND存储器5内的多个区块(物理区块)分类成虚拟NAND存取管理API22用的区块的组#X、及物理NAND存取管理API21用的区块的组#Y。区块的组#X可以每个信道包含例如101个区块(区块地址0~100)。区块的组#Y可以每个信道包含例如10个区块(区块地址101~110)。

[0208] 组#X内的区块是作为多个虚拟区块而编成。多个虚拟区块分别包含多个区块。多个虚拟区块可以分别包含能够并列存取的多个区块的组合。

[0209] 更详细来说,一个虚拟区块可包含能经由Ch.A存取的区块、能经由Ch.B存取的区

块、能经由Ch.C存取的区块、能经由Ch.D存取的区块。如图5所示,在NAND存储器芯片#0~#3连接于信道Ch.A~Ch.D的情况下,一个虚拟区块可以包含芯片#0内的一个区块、芯片#1内的一个区块、芯片#2内的一个区块、芯片#3内的一个区块。向该虚拟区块的数据的写入顺序为芯片#0内的区块内的页面P0、芯片#1内的区块内的页面P0、芯片#2内的区块内的页面P0、芯片#3内的区块内的页面P0、芯片#0内的区块内的页面P1、芯片#1内的区块内的页面P1、芯片#2内的区块内的页面P1、芯片#3内的区块内的页面P1…。

[0210] 在写入64K字节的写入数据的情况下,能够将构成该写入数据的各16K字节的4个数据部并列写入例如芯片#0内的区块内的页面P0、芯片#1内的区块内的页面P0、芯片#2内的区块内的页面P0、芯片#3内的区块内的页面P0。

[0211] 图4所示的组#Y内的区块分别作为物理NAND存取管理API21用的区块、即以单个区块存取的区块(物理区块)而使用。

[0212] 这样,NAND存储器5内的多个区块被分类成虚拟NAND存取管理API22用的区块的组#X及物理NAND存取管理API21用的区块的组#Y,因此能够防止同一个区块被物理NAND存取管理API21用的区块及虚拟NAND存取管理API22用的区块共用。即,物理NAND存取管理API21用的各区块均不属于任何虚拟区块。结果,能够防止物理NAND存取管理API21用的某个特定区块因对虚拟区块的存取而被误存取(读出、写入、删除),所以能够提高安全性。

[0213] [物理区块信息、虚拟区块信息、命名空间信息]

[0214] 图6表示SSD3内维持的删除次数管理表33A。

[0215] 删除次数管理表33A管理组#Y内的区块(物理区块)各自的删除次数。例如,当物理NAND存取管理API21用的特定区块(例如区块地址0)因来自主机2的命令(删除命令、区块分配、删除命令)而被删除时,SSD3内的控制器4将该特定区块(例如区块地址0)的删除次数递增1。

[0216] 图7表示SSD3内维持的另一删除次数管理表33B。

[0217] 删除次数管理表33B管理组#X内的虚拟区块各自的删除次数。例如,当虚拟NAND存取管理API22用的特定的虚拟区块(例如虚拟区块地址0)因来自主机2的命令(删除命令、虚拟区块分配、删除命令)而被删除时,SSD3内的控制器4将该特定的虚拟区块(例如虚拟区块地址0)的删除次数递增1。在虚拟NAND存取管理API22中,一个虚拟区块所含的多个区块是同时被删除。因此,删除次数是以虚拟区块为单位进行管理。

[0218] 图8表示SSD3内维持的又一删除次数管理表33C。

[0219] 删除次数管理表33C管理与各命名空间ID(NSID)对应的合计删除次数。某个NSID的合计删除次数是对该NSID的命名空间(区域)执行的删除动作的累积值,合计删除次数是每当对分配给该NSID的命名空间的区块中的任一区块执行删除动作时便增加1次。分配给该NSID的区块在物理NAND存取管理API21中为物理区块,在虚拟NAND存取管理API22中为虚拟区块。

[0220] 例如,当特定区块(例如区块地址0)因来自主机2的命令(删除命令、或区块分配、删除命令)而被删除时,SSD3内的控制器4特定分配有该特定区块(例如区块地址0)的NSID,并将特定的NSID的删除次数递增1。同样地,当特定的虚拟区块(例如虚拟区块地址0)因来自主机2的命令(删除命令、或虚拟区块分配、删除命令)而被删除时,SSD3内的控制器4特定分配有该特定的虚拟区块(例如虚拟区块地址0)的NSID,并将特定的NSID的删除次数递增

1。

[0221] 图9表示SSD3内维持的物理区块结构信息表33D。

[0222] 物理区块结构信息表33D是表示NAND存储器5内的各区块(物理区块)的结构的信息。物理区块结构信息表33D包含区块大小、页面大小、推定写入时间、删除时间等。区块大小表示一个区块的大小(容量)。页面大小表示一个页面的大小(容量)。推定写入时间表示将数据从页面缓冲器向存储器单元编程所需的时间(tProg)。

[0223] 图10表示SSD3内维持的虚拟区块结构信息表33E。

[0224] 虚拟区块结构信息表33E是表示各虚拟区块的结构的信息。虚拟区块结构信息表33E包含等价的区块大小、页面大小、推定写入时间、删除时间、一个虚拟区块所含的区块数等。等价的区块大小可以是一个虚拟区块所含的各区块的容量的合计。

[0225] 图11表示SSD3内维持的物理NAND存取管理API21用的命名空间信息表33F。

[0226] 命名空间信息表33F管理(1)存在的命名空间的总数、(2)各命名空间的区块分配数、(3)各命名空间的区块地址的列表、(4)各命名空间的合计删除次数。

[0227] 例如,关于NSID#1,对应于NSID#1的区块分配数可以表示确保NSID#1用的区块(物理区块)的总数。对应于NSID#1的区块地址的列表表示分配给NSID#1实际使用的各区块的区块地址。

[0228] 图12表示SSD3内维持的虚拟NAND存取管理API22用的命名空间信息表33G。

[0229] 命名空间信息表33G管理(1)存在的命名空间的总数、(2)各命名空间的虚拟区块分配数、(3)各命名空间的虚拟区块地址的列表、(4)各命名空间的合计删除次数。

[0230] 例如,关于NSID#1,对应于NSID#1的虚拟区块分配数可以表示确保NSID#1用的虚拟区块的总数。对应于NSID#1的虚拟区块地址的列表表示分配给NSID#1实际使用的各虚拟区块的虚拟区块地址。

[0231] [虚拟区块管理]

[0232] 图13表示SSD3内的多个虚拟区块与构成所述多个虚拟区块的各区块(物理区块)的区块地址的关系。

[0233] 属于各虚拟区块的多个区块(物理区块)各自的区块地址的组合是基于数学规则而利用各虚拟区块的虚拟区块地址唯一地决定。通过使用基于数学规则利用虚拟区块地址唯一地决定区块地址的组合的方法,不用对每个虚拟区块使用保持该虚拟区块所属的区块地址的专用管理表,只要利用各虚拟区块地址便能容易地特定属于该虚拟区块的区块地址的组合。

[0234] 作为数学规则能使用能够利用虚拟区块地址唯一地决定区块地址的组的任意规则。

[0235] 在图13中,例如应用如下数学规则的情况:对与Ch.A关联的区块地址0~100升序地分配虚拟区块地址VB0~VB100,对与Ch.B关联的区块地址0~100降序地分配虚拟区块地址VB0~VB100,对与Ch.C关联的区块地址0~100升序地分配虚拟区块地址VB0~VB100,对与Ch.D关联的区块地址0~100降序地分配虚拟区块地址VB0~VB100。

[0236] 这种情况下,例如属于具有虚拟区块地址VB0的虚拟区块的多个区块各自的区块地址的组合是作为Ch.A的区块地址0、Ch.B的区块地址100、Ch.C的区块地址0、及Ch.D的区块地址100进行决定。同样地,属于虚拟区块地址VB1的虚拟区块的多个区块各自的区块地

址的组合是作为Ch.A的区块地址1、Ch.B的区块地址99、Ch.C的区块地址1、及Ch.D的区块地址99进行决定。

[0237] 能够应用的数学规则的例子并不限于于此,例如也可以应用从各信道选择与虚拟区块地址的值相同值的区块地址的数学规则。这种情况下,例如,属于虚拟区块地址VB0的虚拟区块的多个区块各自的区块地址的组合是作为Ch.A的区块地址0、Ch.B的区块地址0、Ch.C的区块地址0、及Ch.D的区块地址0进行决定。同样地,属于虚拟区块地址VB1的虚拟区块的多个区块各自的区块地址的组合是作为Ch.A的区块地址1、Ch.B的区块地址1、Ch.C的区块地址1、及Ch.D的区块地址1进行决定。

[0238] 主机2只要识别能够指定各个虚拟区块的物理地址(虚拟区块地址)即可,无须识别各虚拟区块所含的区块各自的区块地址自身。

[0239] 例如,在虚拟区块地址VB0的虚拟区块中,VB0-0表示虚拟区块地址(VB0)与该虚拟区块内的最初的区块的区块编号(0)的组合。当由主机2指定VB0-0的情况下,SSD3能够将VB0-0变换成Ch.A的区块地址0,并存取该Ch.A的区块地址0。

[0240] [虚拟区块的坏块管理]

[0241] 图14表示由SSD3执行的不良区块(坏块)替换动作。

[0242] 例如,当指定虚拟区块(虚拟区块地址VB2)内的信道Ch.C的区块(VB2-2)作为无法使用的不良区块(坏块)的情况下,SSD3的控制器4能够基于数学规则利用VB2-2特定信道Ch.C的区块地址(=2)。并且,控制器4将该区块地址(=2)的区块登记到坏块列表。并且,控制器4将区块地址(=2)的区块替换成能够与虚拟区块(虚拟区块地址VB2)内的其他所有区块(于此,为Ch.A的区块地址2的区块、Ch.B的区块地址98的区块、Ch.D的区块地址98的区块)并列存取的另一区块。例如,能选择Ch.C的区块群内当前不用于物理NAND存取及虚拟NAND存取的区块、例如Ch.C的区块地址102的区块,作为该另一区块。

[0243] 控制器4只要对将虚拟区块地址VB2的信道Ch.C的区块地址2替换成Ch.C的另一区块地址102进行管理即可。

[0244] [坏块化命令]

[0245] 图15表示应用于SSD3的物理NAND存取管理API21用的坏块化命令。

[0246] 物理NAND存取管理API21用的坏块化命令向SSD3请求应将特定区块设为坏块。主机2可以基于读出错误的次数等来特定应坏块化的区块。坏块化命令包含以下输入参数。

[0247] (1)区块地址:该区块地址指定应坏块化的区块(物理区块)。

[0248] 坏块化命令包含以下输出参数。

[0249] (1)结束状态:向主机2返回表示坏块化命令的成功或错误的结束状态。

[0250] 图16表示应用于SSD3的虚拟NAND存取管理API22用的坏块化命令。

[0251] 虚拟NAND存取管理API22用的坏块化命令向SSD3请求应将特定的虚拟区块内区块设为坏块。主机2可以基于读出错误的次数等来决定应将特定的虚拟区块内的第几个区块坏块化。坏块化命令包含以下输入参数。

[0252] (1)虚拟区块地址与虚拟区块内区块编号

[0253] 虚拟区块地址与虚拟区块内区块编号指定应将哪个虚拟区块内的第几个区块坏块化。虚拟区块内区块编号可为指定信道编号的值。

[0254] 坏块化命令包含以下输出参数。

[0255] (1)结束状态

[0256] 向主机2返回表示坏块化命令的成功或错误的结束状态。

[0257] [坏块化处理的次序]

[0258] 图17的流程图表示由SSD3执行的虚拟NAND存取管理API22用的处理与与坏块化处理的次序。

[0259] SSD3的控制器4将NAND存储器5内的区块群分类成2个组(组#X、组#Y)(步骤S1),并将组#X的区块群作为多个虚拟区块而编成(步骤S2)。

[0260] 在步骤S2中,控制器4基于数学规则来决定应属于各虚拟区块的区块地址的组合。若根据所述初级缺陷列表指定某个虚拟区块内的区块作为坏块,则控制器4可以从组#Y选择与该坏块连接于相同信道的另一区块,并利用选择的区块替换该坏块。组#Y内的剩余区块则作为物理NAND存取管理API21用的区块而使用。

[0261] 并且,控制器4响应于来自主机2的包含指定某个虚拟区块的物理地址(虚拟区块地址、虚拟区块内区块编号、页面地址)的读出、写入或删除命令的接收,对该虚拟区块执行读出动作、写入动作、或删除动作(步骤S3)。在步骤S3中,控制器4基于数学规则利用物理地址(虚拟区块地址、虚拟区块内区块编号)而决定存取对象的虚拟区块内的存取对象区块的区块地址,并对该区块执行读出动作、写入动作、或删除动作。

[0262] 若控制器4接收来自主机2的坏块化命令,则控制器4判定坏块化命令是虚拟NAND存取管理API22用的坏块化命令或物理NAND存取管理API21用的坏块化命令的哪一个(步骤S4、S5)。当虚拟NAND存取管理API22用的坏块化命令与物理NAND存取管理API21用的坏块化命令具有不同操作代码的情况下,该判定基于操作代码而执行。当虚拟NAND存取管理API22用的坏块化命令与物理NAND存取管理API21用的坏块化命令具有相同操作代码的情况下,该判定也可以基于坏块化命令所含的地址的种类(区块地址/虚拟区块地址及虚拟区块内区块编号)而执行。

[0263] 若坏块化命令是物理NAND存取管理API21用的坏块化命令(步骤S5的是),则控制器4将坏块化命令指定的区块地址的区块登记到坏块列表,并将该区块地址的区块作为坏块进行管理(步骤S6)。不执行将该坏块替换成另一区块的处理。

[0264] 若坏块化命令是虚拟NAND存取管理API22用的坏块化命令(步骤S4的是),则控制器4基于在步骤S2中使用的数学规则,利用虚拟区块地址与虚拟区块内区块编号而特定应坏块化的区块的区块地址(步骤S7)。

[0265] 控制器4将特定出的区块地址的区块登记到坏块列表,并将该特定出的区块地址的区块作为坏块进行管理(步骤S8)。

[0266] 并且,控制器4从与该特定出的区块地址的区块(坏块)连接于相同信道的区块之中,选择当前不用于物理NAND存取及虚拟NAND存取的区块(步骤S9),并将该特定出的区块地址的区块(坏块)利用选择的区块替换(步骤S10)。

[0267] [用于物理NAND存取及虚拟NAND存取的处理序列]

[0268] 图18表示由SSD3及主机2执行的用于物理NAND存取及虚拟NAND存取的处理序列。

[0269] 当主机2希望使用物理NAND存取管理API21对SSD3进行存取时,主机2可以向SSD3请求一个区块的分配。该请求可以是所述区块分配、删除命令。SSD3的控制器4从组#Y内的区块中选择不含有有效数据的区块(当前不使用的区块),将该选择的区块分配给主机2,并向

主机2通知该分配的区块的物理地址(区块地址)。

[0270] 当主机2希望使用虚拟NAND存取管理API22对SSD3进行存取时,主机2可以向SSD3请求一个虚拟区块的分配。该请求可以是所述虚拟区块分配、删除命令。SSD3的控制器4从多个虚拟区块选择不含有效数据的虚拟区块(当前不使用的虚拟区块),将该选择的虚拟区块分配给主机2,并向主机2通知该分配的虚拟区块的物理地址(虚拟区块地址)。

[0271] 主机2将包含通知的区块地址的读出、写入或删除命令、即物理NAND存取管理API21用的读出、写入或删除命令发送至SSD3。响应于包含区块地址的读出、写入或删除命令、即物理NAND存取管理API21用的读出、写入或删除命令的接收,控制器4对该区块地址指定的特定且单一的区块执行读出动作、写入动作、或删除动作(步骤S11)。

[0272] 主机2将包含通知的虚拟区块地址的读出、写入或删除命令发送至SSD3。响应于包含虚拟区块地址的读出、写入或删除命令、即虚拟NAND存取管理API22用的读出、写入或删除命令的接收,控制器4对该虚拟区块地址指定的特定的虚拟区块所含的区块的汇集执行读出动作、写入动作、或删除动作(步骤S12)。

[0273] [写入处理]

[0274] 图19表示由SSD3与主机2执行的用来存取单一的物理区块的物理NAND存取的写入处理的处理序列。

[0275] 主机2将物理NAND存取管理API21用的写入命令发送至SSD3。该写入命令包含指定应写入数据的区块的区块地址。在所述直接地址指定模式的情况下,写入命令包含区块地址及页面地址双方。响应于物理NAND存取管理API21用的写入命令的接收,SSD3的控制器4将该写入命令指定的数据写入至写入命令内的区块地址指定的区块内的写入对象页面(步骤S13)。在直接地址指定模式的情况下,写入对象的页面由写入命令内的页面地址指定。在所述自动地址产生模式的情况下,写入对象的页面由控制器4自动产生的页面地址指定。

[0276] 并且,控制器4将写入命令的命令完成回应发送至主机2。

[0277] 图20表示由SSD3及主机2执行的用来对包含多个物理区块的虚拟区块进行存取的虚拟NAND存取的写入处理的处理序列。

[0278] 主机2将虚拟NAND存取管理API22用的写入命令发送至SSD3。该写入命令包含指定应写入数据的虚拟区块的虚拟区块地址。在所述直接地址指定模式的情况下,写入命令包含虚拟区块地址及页面地址双方。如上所述,该写入命令也可以包含虚拟区块地址、虚拟区块内区块编号、页面地址。

[0279] 响应于虚拟NAND存取管理API22用的写入命令的接收,SSD3的控制器4将该写入命令指定的数据写入至写入命令内的虚拟区块地址指定的虚拟区块内的写入对象页面(步骤S14)。在直接地址指定模式的情况下,写入对象的页面由写入命令内的页面地址指定。在所述自动地址产生模式的情况下,写入对象的页面由控制器4自动产生的页面地址指定。此外,在所述自动地址产生模式下,也可以自动产生虚拟区块内区块编号、及页面地址双方。

[0280] 并且,控制器4将写入命令的命令完成回应发送至主机2。

[0281] [读出处理]

[0282] 图21表示由SSD3及主机2执行的用来进行物理NAND存取的读出处理的处理序列。

[0283] 主机2将物理NAND存取管理API21用的读出命令发送至SSD3。该读出命令包含区块地址及页面地址。响应于物理NAND存取管理API21用的读出命令的接收,SSD3的控制器4从

读出命令内的区块地址及页面地址指定的区块内的读出对象页面读出数据(步骤S15)。并且,控制器4将读出的数据、及读出命令的命令完成回应发送至主机2。

[0284] 图22表示由SSD3及主机2执行的用来进行虚拟NAND存取的读出处理的处理序列。

[0285] 主机2将虚拟NAND存取管理API22用的读出命令发送至SSD3。该读出命令包含虚拟区块地址及页面地址。读出命令也可以包含虚拟区块地址、虚拟区块内区块编号、页面地址。

[0286] 响应于虚拟NAND存取管理API22用的读出命令的接收,SSD3的控制器4从读出命令内的虚拟区块地址、虚拟区块内区块编号、及页面地址指定的虚拟区块内的读出对象页面读出数据(步骤S16)。并且,控制器4将读出的数据、及读出命令的命令完成回应发送至主机2。

[0287] [删除处理]

[0288] 图23表示由SSD3及主机2执行的用来进行物理NAND存取的删除处理的处理序列。

[0289] 主机2将物理NAND存取管理API21用的删除命令发送至SSD3。该删除命令包含区块地址。响应于物理NAND存取管理API21用的删除命令的接收,SSD3的控制器4将删除命令内的区块地址指定的区块删除,并将该区块内的所有页面设为删除状态(步骤S17)。并且,控制器4将删除命令的命令完成回应发送至主机2。

[0290] 图24表示由SSD3及主机2执行的用来进行虚拟NAND存取的删除处理的处理序列。

[0291] 主机2将虚拟NAND存取管理API22用的删除命令发送至SSD3。该删除命令包含虚拟区块地址。响应于虚拟NAND存取管理API22用的删除命令的接收,SSD3的控制器4将删除命令内的虚拟区块地址指定的虚拟区块内的多个区块同时删除,并将所述多个区块内的所有页面设为删除状态(步骤S18)。并且,控制器4将删除命令的命令完成回应发送至主机2。

[0292] [优先级管理]

[0293] 图25表示控制器4的命令优先级的管理动作。

[0294] 从主机2接收的所有命令中,表示所述多个命令的执行的优先顺序(优先等级)的值(优先级)可以由主机2赋予。各命令可以具有表示其优先级的输入参数。

[0295] 优先顺序(优先等级)的种类数可以为2以上的任意数。优先等级的种类可以包含例如表示最高优先级的“高”、表示最低优先级的“低”、表示中间优先级的“中”。

[0296] 由控制器4执行的优先级管理中,优先级高的命令能够比优先级低的命令先执行。具有相同优先级的命令间的执行顺序可以通过先进先出方式来决定。关于像删除命令那样直到命令的处理完成为止需要较多时间的命令,可以中断此命令(例如删除命令)的执行,先执行优先顺序高的命令,当优先顺序高的命令完成后,再继续执行中断的命令(例如删除命令)的处理。

[0297] 为了进行优先级管理,可以对每个NAND芯片设置付优先级的队列61、62、63。为了能够识别执行哪个命令,也可以对所有命令及所有命令完成回应赋予队列ID。主机2可以对所有命令赋予队列ID。

[0298] 赋予了“高”优先级的各命令储存在付优先级的队列61(排队)。赋予了“中”优先级的各命令储存在付优先级的队列62。赋予了“低”优先级的各命令储存在付优先级的队列63。相比于从付优先级的队列62的命令取出及从付优先级的队列63的命令取出,优先执行从付优先级的队列61的命令的取出。相比于从付优先级的队列63的命令取出,优先执行从

付优先级的队列62的命令取出。

[0299] [区块分配、删除命令]

[0300] 参照图26,对物理NAND存取管理API21用的区块分配、删除命令、及虚拟NAND存取管理API22用的区块分配、删除命令进行说明。

[0301] 物理NAND存取管理API21用的区块分配、删除命令包含以下输入参数。

[0302] (1)区块类型=区块:区块类型表示应分配的区块的类型。物理NAND存取管理API21用的区块分配、删除命令的区块类型为区块(物理区块)。从空闲区块向主机2分配一个区块,且自动删除该区块。

[0303] (2)处理优先级:处理优先级表示该命令的优先级。

[0304] (3)NSID(可选):NSID表示应分配区块的命名空间的ID。

[0305] 物理NAND存取管理API21用的区块分配、删除命令包含以下输出参数。

[0306] (1)结束状态:向主机2返回表示区块分配、删除命令的成功或错误的结束状态。

[0307] (2)区块地址:向主机2返回分配的区块的区块地址。

[0308] (3)剩余区块数:只在指定NSID的情况下,向主机2返回确保该NSID用的剩余区块数。

[0309] 虚拟NAND存取管理API22用的区块分配、删除命令(也称为“虚拟区块分配、删除命令”)包含以下输入参数。

[0310] (1)区块类型=虚拟区块:虚拟NAND存取管理API22用的区块分配、删除命令的区块类型为虚拟区块。从空闲虚拟区块向主机2分配一个虚拟区块,且自动删除该虚拟区块内的多个区块。

[0311] (2)处理优先级:处理优先级表示该命令的优先级。

[0312] (3)NSID(可选):NSID表示应分配虚拟区块的命名空间的ID。

[0313] 虚拟NAND存取管理API22用的区块分配、删除命令包含以下输出参数。

[0314] (1)结束状态:向主机2返回表示区块分配、删除命令的成功或错误的结束状态。

[0315] (2)虚拟区块地址:向主机2返回分配的虚拟区块的虚拟区块地址。

[0316] (3)剩余区块数:只在指定NSID的情况下,向主机2返回确保该NSID用的剩余虚拟区块的数。

[0317] [物理区块管理与虚拟区块管理]

[0318] 图27表示由SSD3的控制器4管理的使用中区块列表71A、空闲区块列表71B、使用中虚拟区块列表72A、及空闲虚拟区块列表72B。

[0319] 使用中区块列表71A表示组#Y的区块群之中保持有效数据的区块(物理区块)的列表,即正被主机2使用的使用中区块的列表。空闲区块列表71B表示组#Y的区块群之中未保持有效数据的区块(物理区块)的列表,即未被主机2使用的空闲区块的列表。

[0320] 使用中虚拟区块列表72A表示组#X的虚拟区块群之中保持有效数据的虚拟区块的列表,即正被主机2使用的使用中虚拟区块的列表。空闲虚拟区块列表72B表示组#X的虚拟区块群之中未保持有效数据的虚拟区块的列表,即未被主机2使用的空闲虚拟区块的列表。

[0321] [区块分配、删除处理]

[0322] 图28表示由SSD3及主机2执行的区块分配、删除处理的处理序列。

[0323] 首先,说明物理NAND存取管理API21用的区块分配、删除处理。

[0324] 当前分配给主机2的区块(物理区块)写满来自主机2的数据的情况下,主机2可以向SSD3发送物理NAND存取管理API21用的区块分配、删除命令。

[0325] 当SSD3的控制器4从主机2接收该区块分配、删除命令时,控制器4从空闲区块列表71B选择一个区块(物理区块),并将选择的区块(物理区块)作为写入对象区块分配给主机2(步骤S21)。控制器4拥有来自空闲区块列表71B的区块(写入对象区块)的选择权限。因此,控制器4能够将可靠性高的区块作为写入对象区块分配给主机2。在步骤S21中,控制器4可以从空闲区块列表71B选择最小删除次数的区块,并将该最小删除次数的区块作为写入对象区块分配给主机2。

[0326] 控制器4删除该分配的区块,并更新该分配的区块的删除次数(步骤S22)。控制器4将该分配的区块的区块地址通知给主机2(步骤S23)。该区块地址也可以作为对于区块分配、删除命令的命令完成回应内的返回值而通知给主机2。

[0327] 接着,说明虚拟NAND存取管理API22用的区块分配、删除处理。

[0328] 在当前分配给主机2的虚拟区块写满来自主机2的数据的情况下,主机2可以向SSD3发送虚拟NAND存取管理API22用的区块分配、删除命令。

[0329] 当SSD3的控制器4从主机2接收该区块分配、删除命令时,控制器4从空闲虚拟区块列表72B选择一个虚拟区块,并将选择的虚拟区块作为写入对象虚拟区块而分配给主机2(步骤S21)。控制器4拥有来自空闲虚拟区块列表72B的虚拟区块(写入对象虚拟区块)的选择权限,因此控制器4能够将可靠性高的虚拟区块作为写入对象虚拟区块分配给主机2。在步骤S21中,控制器4可以从空闲虚拟区块列表72B选择最小删除次数的虚拟区块,并将该最小删除次数的虚拟区块作为写入对象虚拟区块分配给主机2。

[0330] 控制器4将该分配的虚拟区块所含的多个区块同时删除,并更新该分配的虚拟区块的删除次数(步骤S22)。控制器4将该分配的虚拟区块的虚拟区块地址通知给主机2(步骤S23)。该虚拟区块地址也可以作为对于区块分配、删除命令的命令完成回应内的返回值而通知给主机2。

[0331] [写入命令]

[0332] 图29表示应用于SSD3的物理NAND存取管理API21用的写入命令。

[0333] 写入命令包含以下输入参数。

[0334] (1)区块地址、或区块地址与页面地址:该输入参数值是指定应写入数据的NAND存储器5上的位置的物理地址。在只指定区块地址的情况下,写入对象页面由SSD3自动更新。

[0335] (2)NSID(可选):在未指定区块地址的情况下,指定NSID。在指定了NSID的情况下,区块地址与页面地址由SSD3自动发布。数据被写入最后分配给NSID的区块(当前写入对象区块)。

[0336] (3)处理优先级:处理优先级表示该写入命令的优先级。

[0337] (4)写入数据的起始地址:写入数据的起始地址表示储存着写入数据的输出缓冲器(主机的存储器)上的起始地址。

[0338] (5)写入页面数:写入页面数表示应写入数据的页面的数、即所述“特定页面数”。在写入数据的区块内的页面数达到写入页面数(“特定页面数”)的情况下,向主机2通知此意旨。

[0339] (6)NAND模式(可选):NAND模式包含SLC/MLC/TLC等。

[0340] 写入命令包含以下输出参数。

[0341] (1)结束状态:向主机2返回表示写入命令的成功或错误的结束状态。

[0342] (2)区块地址与页面地址:区块地址与页面地址表示写入数据的NAND存储器5上的位置。在写入命令只包含区块地址的情况下,或只包含NSID的情况下,主机2能够基于该返回值获知写入数据的NAND存储器5上的位置。

[0343] (3)写入页面数:该值表示写入数据的页面的数。

[0344] (4)不当写入顺序警告:在检测到不当写入顺序的情况下,向主机2返回警告或错误。

[0345] (5)可读出的最新页面地址:向主机2返回保持可读出数据的最新页面地址。主机2能够获知直到该区块内的哪个页面为止可读出。

[0346] 图30表示应用于SSD3的虚拟NAND存取管理API22用的写入命令。

[0347] 该写入命令包含以下输入参数。

[0348] (1)虚拟区块地址、或虚拟区块地址与页面地址:这些地址是指定应写入数据的NAND存储器5上的位置的物理地址。该物理地址可以用虚拟区块地址、虚拟区块内区块编号、页面地址来表现。在只指定了虚拟区块地址的情况下,虚拟区块内区块编号、页面地址可以由SSD3自动更新。

[0349] (2)NSID(可选):在未指定虚拟区块地址的情况下,指定NSID。在指定NSID的情况下,虚拟区块地址、虚拟区块内区块编号、页面地址可以由SSD3自动发布。数据被写入分配给NSID的虚拟区块。

[0350] (3)处理优先级:处理优先级表示该写入命令的优先级。

[0351] (4)写入数据的起始地址:写入数据的起始地址表示储存着写入数据的输出缓冲器(主机的存储器)上的起始地址。

[0352] (5)写入页面数:写入页面数表示应写入数据的页面的数。当写入数据的区块内的页面数达到写入页面数的情况下,向主机2通知此意旨。

[0353] (6)NAND模式(可选):NAND模式包含SLC/MLC/TLC等。

[0354] 写入命令包含以下输出参数。

[0355] (1)结束状态:向主机2返回表示写入命令的成功或错误的结束状态。

[0356] (2)虚拟区块地址与页面地址:该值是表示写入数据的NAND存储器5上的位置的物理地址。该物理地址可以用虚拟区块地址、虚拟区块内区块编号、页面地址来表现。在写入命令只包含虚拟区块地址或只包含NSID的情况下,主机2能够基于该返回值获知写入数据的NAND存储器5上的位置。

[0357] (3)写入页面数:该值表示写入数据的页面的数。

[0358] (4)不当写入顺序警告:在检测到不当写入顺序的情况下,向主机2返回警告或错误。

[0359] (5)可读出的最新页面地址:向主机2返回保持可读出数据的最新页面地址。主机2能够获知直到虚拟区块内存取对象区块内的哪个页面为止可读出。

[0360] [写入顺序制约]

[0361] 图31表示向区块内的多个页面的数据写入顺序相关的制约。

[0362] 于此,假定一个区块包含页面0~页面255的情况。关于数据读出,能够以任意顺序

读出区块内的任意的若干页面。另一方面,关于数据写入,在某一个区块内,数据必须按照页面0、页面1、页面2、页面3、…页面254、页面255的顺序连续地写入。因此,“不当写入顺序警告”功能能够帮助主机2使用直接地址指定模式、即帮助主机2控制应写入数据的区块地址及页面地址。

[0363] [来自页面的数据读出的时序相关的制约]

[0364] 有如下情况:即便NAND存储器向区块内的某个页面写入了数据,但写入该页面的数据无法在刚写入读出,写入该页面的数据在向该页面后续的1以上的页面的数据写入完成后才能读出。

[0365] 图32表示来自页面的数据读出的时序相关的制约。

[0366] 于此,例示向页面P0~P3写入数据后页面P0保持可读出数据的状态的情况。这样,有如下情况:即便向某个特定页面(例如页面P0)的写入完成,但在向后续若干页面(例如页面P1~P3)的写入完成之前,无法从该特定页面(例如页面P0)正确地读出数据。

[0367] 这种来自页面的数据读出的时序相关的制约是由于在NAND存储器内执行的编程动作而引起的。

[0368] 即,在NAND存储器中,单元不断微细化,因向单元的数据写入会引起相邻单元的阈值电平发生变动这样的编程干扰。因此,在NAND存储器中,有考虑编程干扰的影响,因向页面内的各单元的数据的写入,而执行对其前一个以上的页面内的各单元的阈值电平进行校正的编程动作的情况。若从校正完成前的页面读出数据,便会读出与原本的数据不同的错误数据。校正完成时序因使用的NAND存储器的类型而不同。

[0369] 控制器4能够向主机2通知校正完成时序。换句话说,控制器4能够向主机2通知能读出写入数据的区块内的最后的页面。

[0370] 更详细来说,控制器4执行以下处理。

[0371] 控制器4从主机2接收包含指定某个特定区块的区块地址及指定该特定区块内的写入对象页面的页面地址的写入命令。控制器4按照写入命令向特定区块内的写入对象页面写入数据。并且,控制器4向主机2通知表示因向该特定区块内的写入对象页面的数据写入而变成可读出状态的该特定区块内的最新页面(该特定区块内的页面群之中最后变得可读出的页面)的页面地址。

[0372] 主机2可以根据该通知更新表示可读出写入数据的页面的主机2上的可读出页面地址管理信息。在接收表示某个页面的数据可读出的通知后,主机2可以释放保持该页面的数据的主机2的存储器区域(写入缓冲器)。换句话说,主机2首先将应写入SSD3的NAND存储器5的数据临时储存在主机2的存储器内,然后用用来写入该应写入数据的写入命令发送至SSD3。并且,主机2在该数据变得从非易失性存储器5可读出之前,将该数据维持在主机2的存储器。

[0373] 此外,为了减小编程干扰的影响,也有NAND存储器按照图33所示的编程顺序执行编程动作的情况。在图33中,用来向各页面写入数据的编程动作包含多次写入阶段。向各页面的多次写入阶段内的至少最后的写入阶段是在向后续1以上的页面的1以上的写入阶段完成后执行。这时,有如下情况:即便通过将应写入某个特定页面(例如页面P0)的数据传输至NAND存储器而完成向该特定页面的写入,但在通过将应写入其他1以上的页面的数据传输至NAND存储器而完成向该1以上的页面的写入之前,无法从该特定页面准确地读出数据。

[0374] 在每个单元写入3比特的数据的数据的TLC写入中,可以按照以下的编程顺序执行编程动作。

[0375] (1)向页面0的第1写入阶段(向页面0的下位页面数据的写入)

[0376] (2)向页面1的第1写入阶段(向页面1的下位页面数据的写入)

[0377] (3)向页面0的第2写入阶段(向页面0的中位页面数据的写入)

[0378] (4)向页面2的第1写入阶段(向页面2的下位页面数据的写入)

[0379] (5)向页面1的第2写入阶段(向页面1的中位页面数据的写入)

[0380] (6)向页面0的第3写入阶段(向页面0的上位页面数据的写入)

[0381] (7)向页面3的第1写入阶段(向页面3的下位页面数据的写入)

[0382] (8)向页面2的第2写入阶段(向页面2的中位页面数据的写入)

[0383] (9)向页面1的第3写入阶段(向页面1的上位页面数据的写入)

[0384] 可使用的编程顺序并不限于该例,可以对每个NAND存储器使用不同的各种编程顺序。

[0385] [写入处理的次序]

[0386] 图34的流程图表示由SSD3执行的数据写入处理的次序。

[0387] SSD3的控制器4响应于区块分配、删除命令的接收而将写入对象区块分配给主机2,并且将写入对象页面设定为初始值(页面0)(步骤S31)。若控制器4从主机2接收写入命令(步骤S32的是),则控制器4判定接收的写入命令是否为包含页面地址的直接地址指定模式的写入命令(步骤S33)。

[0388] 若接收的写入命令是直接地址指定模式的写入命令(步骤S33的是),则控制器4判定该写入命令指定的页面地址是否与当前写入对象页面(于此为页面0)一致(步骤S34)。

[0389] 若写入命令指定的页面地址与当前写入对象页面不一致(步骤S34的否),则控制器4检测到不当写入顺序发生,向主机2通知包含不当写入顺序的警告的命令完成回应(步骤S35)。在步骤S35中,控制器4不向写入命令指定的页面地址写入数据,而是向主机2通知包含不当写入顺序的警告的命令完成回应。

[0390] 另一方面,若写入命令指定的页面地址与当前写入对象页面一致(步骤S34的是),则控制器4将写入命令指定的数据传输至NAND存储器5,将数据写入至写入命令的页面地址指定的写入对象区块内的页面(步骤S36)。控制器4基于按从最初的页面到最后的页面的顺序写入数据的写入顺序相关的制约,更新写入对象页面(于此,从页面0更新为页面1)(步骤S37)。

[0391] 若接收的写入命令并非直接地址指定模式的写入命令(步骤S33的否),则控制器4自动发布当前写入对象页面的页面地址(于此页面0),将写入命令指定的数据传输至NAND存储器5,并将数据写入至写入对象区块内的当前写入对象页面(步骤S38)。控制器4基于所述写入顺序相关的制约,更新写入对象页面(于此,从页面0更新为页面1)(步骤S39)。

[0392] 在步骤S37或S39之后,控制器4特定保持可读出数据的最新的页面地址(步骤S40)。

[0393] 并且,控制器4判定是否因此次写入而数据写入至写入对象区块内的最后的页面为止、及写入数据的页面数是否达到“写入页面数”(所述“特定页面数”)(步骤S41、S42)。

[0394] 并且,控制器4创建返回值,将包含返回值的命令完成回应发送至主机2。

[0395] 若数据写入至写入对象区块内的最后的页面为止(步骤S41的是),则控制器4向主机2发送包含结束状态、返回值的命令完成回应(步骤S43)。返回值包含以下值。

[0396] (1)表示写入数据的NAND存储器5上的位置的物理地址(该物理地址可为区块地址与页面地址、或者也可为虚拟区块地址与页面地址)

[0397] (2)可读出的最新页面地址

[0398] (3)表示区块内的所有页面的写入完成的状态

[0399] 若写入数据的页面数达到“写入页面数”(所述“特定页面数”)(步骤S42的是),则控制器4将包含结束状态与返回值的命令完成回应发送至主机2(步骤S44)。返回值包含以下值。

[0400] (1)表示写入数据的NAND存储器5上的位置的物理地址(该物理地址可为区块地址与页面地址,或者也可为虚拟区块地址与页面地址)

[0401] (2)可读出的最新页面地址

[0402] (3)表示特定页面数的写入完成的状态

[0403] 图35的流程图表示响应于写入命令的命令完成回应的接收而由主机2执行的处理的次序。

[0404] 响应于来自SSD3的写入命令的命令完成回应的接收,主机2判定该写入命令的命令处理是否成功(步骤S51)。

[0405] 若写入命令的命令处理成功(步骤S51的是),则主机2使用命令完成回应所含的区块地址与页面地址,更新地址变换表(即查找表LUT45),由此对与通过该写入命令而写入的数据对应的LBA,映射正确的物理地址(步骤S52)。并且,主机2基于可读出的最新页面地址更新所述可读出页面地址管理信息,并释放保持着向该可读出的最新页面地址写入的数据的主机2的存储器区域(写入缓冲器)(步骤S53)。

[0406] 即,在通过该写入命令写入的数据变得能够从SSD3读出之前,将通过该写入命令写入的数据维持在主机2的存储器区域(写入缓冲器)。在通过该写入命令写入的数据变得能够从SSD3读出之后,对于该数据的读出请求的存取目标从写入缓冲器切换到SSD3。

[0407] 并且,主机2判定特定页面数的写入是否完成、写入是否完成到区块的最后的页面为止(步骤S54、S56)。

[0408] 若特定页面数的写入完成(步骤S54的是),则主机2将请求比如元数据这样的管理信息的写入的写入命令发送至SSD3(步骤55)。由此,能够向各区块的最终页面等写入比如元数据这样的管理信息。

[0409] 若写入完成到区块的最后的页面为止(步骤S56的是),则主机2将区块分配、删除命令发送至SSD3(步骤57)。

[0410] 若写入命令的命令处理错误(步骤S51的否),则主机2判定是否因不当写入顺序所致的错误(步骤S58)。

[0411] 若为因不当写入顺序所致的错误(步骤S58的是),则主机2执行包含用来特定不当写入顺序的原因的处理的错误处理(步骤S59)。

[0412] 图36的流程图表示响应于包含NSID的写入命令的命令完成回应的接收而由主机2执行的处理的次序。

[0413] 响应于来自SSD3的写入命令的命令完成回应的接收,主机2判定写入命令的命令

处理是否成功(步骤S61)。

[0414] 若写入命令的命令处理成功(步骤S61的是),则主机2特定与通过该写入命令写入的数据关联的NSID(步骤S62)。命令完成回应也可以包含与写入命令内的NSID相同的NSID。

[0415] 主机2使用命令完成回应所含的区块地址与页面地址,更新与特定的NSID对应的地址变换表(即查找表LUT45),由此对与数据对应的LBA映射正确的物理地址(步骤S63)。并且,主机2通过将可读出的最新页面地址保存在所述可读出页面地址管理信息来更新可读出页面地址管理信息,并释放保持着向该可读出的最新页面地址写入的数据的主机2的存储器区域(写入缓冲器)(步骤S64)。

[0416] 并且,主机2判定特定页面数的写入是否完成、写入是否完成到区块的最后的页面为止(步骤S65、S67)。

[0417] 若特定页面数的写入完成(步骤S65的是),则主机2将请求比如元数据这样的管理信息的写入的写入命令发送至SSD3(步骤66)。

[0418] 若写入完成到区块的最后的页面为止(步骤S67的是),则主机2将区块分配、删除命令发送至SSD3(步骤68)。

[0419] 若写入命令的命令处理错误(步骤S61的否),则主机2判定是否为因不当写入顺序所致的错误(步骤S69)。

[0420] 若为因不当写入顺序所致的错误(步骤S69的是),则主机2执行包含特定与通过该写入命令应写入的数据关联的NSID的处理、用来特定不当写入顺序的原因的处理等的错误处理(步骤S70)。

[0421] [读出命令]

[0422] 图37表示应用于SSD3的物理NAND存取管理API21用的读出命令。

[0423] 读出命令包含以下输入参数。

[0424] (1)区块地址与页面地址:该地址是指定应读出数据的NAND存储器5上的位置的物理地址。

[0425] (2)处理优先级:处理优先级表示该读出命令的优先级。

[0426] (3)读出数据的传输目标地址:读出数据的传输目标地址表示应传输读出数据的输入缓冲器(主机的存储器)上的位置。

[0427] (4)读出页面数:读出页面数表示应读出的页面的数。

[0428] (5)允许等待时间:允许等待时间指定最小等待时间、正常等待时间、或长等待时间的任一个。

[0429] 读出命令包含以下输出参数。

[0430] (1)结束状态:向主机2返回表示读出命令的成功或错误的结束状态。

[0431] (2)区块地址与页面地址:区块地址与页面地址表示读出数据的NAND存储器5上的位置。

[0432] (3)页面数:该值表示读出数据的页面的数。

[0433] (4)起始地址:该地址表示读出数据的起始地址。

[0434] 图38表示应用于SSD3的虚拟NAND存取管理API22用的读出命令。

[0435] 该读出命令包含以下输入参数。

[0436] (1)虚拟区块地址与页面地址:该地址是指定应读出数据的NAND存储器5上的位置

的物理地址。

[0437] (2)处理优先级:处理优先级表示该读出命令的优先级。

[0438] (3)读出数据的传输目标地址:读出数据的传输目标地址表示应传输读出数据的输入缓冲器(主机的存储器)上的位置。

[0439] (4)读出页面数:读出页面数表示应读出的页面的数。

[0440] (5)允许等待时间:允许等待时间指定最小等待时间、正常等待时间、或长等待时间的任一个。

[0441] 读出命令包含以下输出参数。

[0442] (1)结束状态:向主机2返回表示读出命令的成功或错误的结束状态。

[0443] (2)虚拟区块地址与页面地址:虚拟区块地址与页面地址表示读出数据的NAND存储器5上的位置。

[0444] (3)页面数:该值表示读出数据的页面的数。

[0445] (4)起始地址:该地址表示读出数据的起始地址。

[0446] [读出处理的次序]

[0447] 参照图39及图40,表示由SSD3及主机2执行的数据读出处理的次序。

[0448] 如图39所示,主机2参照地址变换表(即查找表LUT45),将读出对象的数据的LBA变换成NAND存储器5的物理地址(步骤S71)。并且,主机2将包含该物理地址的读出命令发送至SSD3。

[0449] SSD3的控制器4从该物理地址指定的区块内的页面读出数据(步骤S72)。在该步骤S72中,控制器4执行图40所示的处理。

[0450] 即,如图40所示,控制器4执行从物理地址指定的物理位置(某个区块内的某个页面)读出数据的动作、及读出数据的错误订正动作(步骤S81)。并且,控制器4判定读出数据是否包含ECC无法订正的错误(步骤S82)。若读出仍未变成可读出状态的页面,则读出数据包含ECC无法订正的大量错误。

[0451] 在读出数据不含无法订正的错误的情况下(步骤S82的否),则控制器4将该读出数据传输至主机2,且将表示成功的命令完成回应传输至主机2(步骤S83)。

[0452] 另一方面,在读出数据包含无法订正的错误的情况下(步骤S82的是),控制器4将表示错误的命令完成回应传输至主机2(步骤S84)。

[0453] [数据复制]

[0454] 图41表示SSD3执行的数据复制动作的例子。

[0455] SSD3的控制器4并非将复制源区块内的所有数据复制到复制目标区块,而是跳过复制源区块内的指定页面范围内的无效数据,只将该页面范围内的有效数据复制到复制目标区块内的指定页面范围。该数据复制动作如上所述是以垃圾回收为目的而执行。

[0456] 在数据复制动作中,如上所述,控制器4自动跳过不含有有效数据的无效页面的复制。由此,主机2即便不以页面单位个别地指定应复制的页面,也能只将有效页面复制到复制目标区块。

[0457] 此外,主机2利用复制命令不仅能指定复制源区块及复制目标区块,还能指定复制源区块内的复制开始页面及复制目标区块内的传输开始页面。由此,能够执行将复制源区块内的特定页面群复制到复制目标区块内的特定页面群这样的极细微的复制动作。另外,

也可以指定多个复制源区块。

[0458] 而且,主机2也可以指定“到复制结束为止应复制的有效数据数”、或“到复制结束为止应检测的无效数据数”的任一个作为数据复制的结束条件。

[0459] 在指定“到复制结束为止应复制的有效数据数”作为数据复制动作的结束条件的情况下,持续进行数据复制动作直到向复制目标区块复制所需数的有效数据为止。当向复制目标区块复制了所需数的有效数据时,结束数据复制动作。例如,若指定一个区块的数据数作为“到复制结束为止应复制的有效数据数”,能够用从若干复制源区块复制的有效数据写满复制目标区块,且能够将若干复制源区块设为只包含无效数据的空闲区块。另外,并非必须在一次数据复制动作中增加一个空闲区块数,也可以在多次数据复制动作中增加一个空闲区块数。因此,“到复制结束为止应复制的有效数据数”可为任意数。

[0460] 在指定“到复制结束为止应检测的无效数据数”作为数据复制动作的结束条件的情况下,持续数据复制动作指导跳过无效数据的复制的次数变成所需次数为止。当跳过无效数据的复制的次数变成所需次数时,结束数据复制动作。通常,选择的若干复制源区块分别是混杂了有效数据和无效数据的区块。此外,这些选择的若干复制源区块所含的无效数据数的合计至少有一个区块的数据数以上。因此,例如,若指定一个区块的数据数作为“到复制结束为止应检测的无效数据数”,则到复制动作结束为止,能够将至少一个复制源区块设为只包含无效数据的空闲区块。如上所述,可以在多次数据复制动作中增加一个空闲区块数,因此“到复制结束为止应检测的无效数据数”也可为任意数。

[0461] 图41中,为了简化说明,假定利用来自主机2的数据复制命令而指定以下参数的情况。

[0462] (1)复制源区块=区块B0

[0463] (2)复制源开始页面=P31

[0464] (3)复制目标区块=区块B10

[0465] (4)传输目标开始页面=P11

[0466] (5)有效/无效位图=位图数据81

[0467] (6)到复制结束为止应复制的有效数据数=3

[0468] 位图数据81表示复制对象范围内的各页面的数据的有效/无效。控制器4首先判定复制源区块B0的复制开始页面P31的数据的有效/无效。在图41的情况下,页面P31的数据有效。因此,控制器4从页面P31读出数据,并将该读出的数据复制到复制目标区块B10的传输目标开始页面P11。复制到复制目标区块B10的有效数据数变成1。

[0469] 控制器4判定复制源区块B0的页面P32的数据的有效/无效。在图41的情况下,页面P32的数据无效。因此,控制器4跳过页面P32的数据的复制。复制到复制目标区块B10的有效数据数维持1。

[0470] 控制器4判定复制源区块B0的页面P33的数据的有效/无效。在图41的情况下,页面P33的数据有效。因此,控制器4从页面P33读出数据,并将该读出的数据复制到复制目标区块B10的页面P12。复制到复制目标区块B10的有效数据数变成2。

[0471] 控制器4判定复制源区块B0的页面P34的数据的有效/无效。在图41的情况下,页面P34的数据无效。因此,控制器4跳过页面P34的数据的复制。复制到复制目标区块B10的有效数据数维持2。

[0472] 控制器4判定复制源区块B0的页面P35的数据的有效/无效。在图41的情况下,页面P35的数据有效。因此,控制器4从页面P35读出数据,并将该读出的数据复制到复制目标区块B10的页面P13。复制到复制目标区块B10的有效数据数变成3。由于复制到复制目标区块B10的有效数据数达到结束条件(到复制结束为止应复制的有效数据数),因此结束复制动作。

[0473] 图42表示指定应检测的无效数据数作为结束条件的情况下的数据复制动作的例子。

[0474] 图42中,假定利用来自主机2的数据复制命令指定以下参数的情况。

[0475] (1)复制源区块=区块B0

[0476] (2)复制源开始页面=P31

[0477] (3)复制目标区块=区块B10

[0478] (4)传输目标开始页面=P11

[0479] (5)有效/无效位图=位图数据81

[0480] (6)到复制结束为止应检测的无效数据数=3

[0481] 控制器4首先判定复制源区块B0的复制开始页面P31的数据的有效/无效。在图42的情况下,页面P31的数据有效。因此,控制器4从页面P31读出数据,并将该读出的数据复制到复制目标区块B10的传输目标开始页面P11。

[0482] 控制器4判定复制源区块B0的页面P32的数据的有效/无效。在图42的情况下,页面P32的数据无效。因此,控制器4跳过页面P32的数据的复制。检测的无效数据数(即跳过复制的数据的数)变成1。

[0483] 控制器4判定复制源区块B0的页面P33的数据的有效/无效。在图42的情况下,页面P33的数据有效。因此,控制器4从页面P33读出数据,并将该读出的数据复制到复制目标区块B10的页面P12。检测的无效数据数维持1。

[0484] 控制器4判定复制源区块B0的页面P34的数据的有效/无效。在图42的情况下,页面P34的数据无效。因此,控制器4跳过页面P34的数据的复制。检测的无效数据数变成2。

[0485] 控制器4判定复制源区块B0的页面P35的数据的有效/无效。在图42的情况下,页面P35的数据有效。因此,控制器4从页面P35读出数据,并将该读出的数据复制到复制目标区块B10的页面P13。检测的无效数据数维持2。

[0486] 控制器4判定复制源区块B0的页面P36的数据的有效/无效。在图42的情况下,页面P36的数据有效。因此,控制器4从页面P36读出数据,并将该读出的数据复制到复制目标区块B10的页面P14。检测的无效数据数维持2。

[0487] 控制器4判定复制源区块B0的页面P37的数据的有效/无效。在图42的情况下,页面P37的数据无效。因此,控制器4跳过页面P37的数据的复制。检测的无效数据数变成3。由于检测的无效数据数达到结束条件(到复制结束为止应检测的无效数据数),因此结束复制动作。

[0488] 图43表示指定多个复制源区块且指定应复制的有效数据数作为结束条件的情况下的数据复制动作的例子。

[0489] 于此,为了简化图示,假定1区块所含的页面数为3、到复制结束为止应复制的有效数据数为3的情况。位图数据81A、81B是与区块B11、B20分别对应的有效/无效位图。

[0490] 区块B11、B20是由来自主机2的复制命令指定的复制源区块,区块B30是由复制命令指定的复制目标区块。

[0491] 控制器4首先判定复制源区块B11的复制开始页面P0的数据的有效/无效。在图43的情况下,页面P0的数据有效。因此,控制器4从页面P0读出数据,并将该读出的数据复制到复制目标区块B30的传输目标开始页面P0。复制到复制目标区块B30的有效数据数变成1。

[0492] 控制器4判定复制源区块B11的页面P1的数据的有效/无效。在图43的情况下,页面P1的数据有效。因此,控制器4从页面P1读出数据,并将该读出的数据复制到复制目标区块B30的页面P1。复制到复制目标区块B30的有效数据数变成2。

[0493] 控制器4判定复制源区块B11的页面P2的数据的有效/无效。在图43的情况下,页面P2的数据无效。因此,控制器4跳过页面P2的数据的复制。复制到复制目标区块B30的有效数据数维持2。

[0494] 控制器4判定下一复制源区块B20的页面P0的数据的有效/无效。在图43的情况下,复制源区块B20的页面P0的数据有效。因此,控制器4从复制源区块B20的页面P0读出数据,并将该读出的数据复制到复制目标区块B30的页面P2。复制到复制目标区块B30的有效数据数变成3。由于复制到复制目标区块B30的有效数据数达到结束条件(到复制结束为止应复制的有效数据数),因此结束复制动作。

[0495] 控制器4按复制的有效数据将表示有效数据ID及储存该有效数据的复制目标区块内的位置(复制目标位置)的数据复制信息,作为复制命令的命令完成回应通知给主机2。基于该数据复制信息,主机2更新地址变换表(LUT45),将各个复制的数据映射到正确的物理地址。区块B11的页面P0的数据、页面P1的数据无效化。由此,区块B11变成不含有效数据的空闲区块。同样地,区块B20的页面P0的数据也无效化。

[0496] 而且,控制器4将复制源区块B20的页面P1作为下一次的复制源开始页面进行决定,并且,将下一次的复制源开始页面的物理地址(复制源区块B20的区块地址、页面P1的页面地址)通知给主机2。

[0497] 而且,控制器4特定出表示因数据复制动作复制了有效数据的复制目标区块的页面群之中,保持可读出数据的最新的页面的页面地址(最新页面地址),并将特定出的最新页面地址通知给主机2。由此,主机2能够获知复制目标区块B30内的页面群中的可读出的最后的页面。

[0498] 图44表示指定多个复制源区块且指定应检测的无效数据数作为结束条件的情况下的数据复制动作的例子。

[0499] 于此,为了简化图示,假定1区块所含的页面数为3、到复制结束为止应检测的无效数据数为3的情况。位图数据81A、81B、81C是与区块B11、B20、B25分别对应的有效/无效位图。

[0500] 区块B11、B20、B25是由来自主机2的复制命令指定的复制源区块,区块B30、B31是由复制命令指定的复制目标区块。

[0501] 控制器4首先判定复制源区块B11的复制开始页面P0的数据的有效/无效。在图44的情况下,页面P0的数据有效。因此,控制器4从页面P0读出数据,并将该读出的数据复制到复制目标区块B30的传输目标开始页面P0。

[0502] 控制器4判定复制源区块B11的页面P1的数据的有效/无效。在图44的情况下,页面

P1的数据有效。因此,控制器4从页面P1读出数据,并将该读出的数据复制到复制目标区块B30的页面P1。

[0503] 控制器4判定复制源区块B11的页面P2的数据的有效/无效。在图44的情况下,页面P2的数据无效。因此,控制器4跳过页面P2的数据的复制。检测的无效数据数(即跳过复制的数据的数)变成1。

[0504] 控制器4判定下一复制源区块B20的页面P0的数据的有效/无效。在图44的情况下,复制源区块B20的页面P0的数据有效。因此,控制器4从复制源区块B20的页面P0读出数据,并将该读出的数据复制到复制目标区块B30的页面P2。由于检测的无效数据数仍未达到结束条件,因此继续进行复制动作。

[0505] 控制器4判定复制源区块B20的页面P1的数据的有效/无效。在图44的情况下,页面P1的数据无效。因此,控制器4跳过复制源区块B20的页面P1的数据的复制。检测的无效数据数变成2。

[0506] 控制器4判定复制源区块B20的页面P2的数据的有效/无效。在图44的情况下,复制源区块B20的页面P2的数据有效。因此,控制器4从复制源区块B20的页面P2读出数据,并将该读出的数据复制到复制目标区块B31的页面P0。

[0507] 控制器4判定下一复制源区块B25的页面P0的数据的有效/无效。在图44的情况下,复制源区块B25的页面P0的数据无效。因此,控制器4跳过复制源区块B25的页面P0的数据的复制。检测的无效数据数变成3。由于检测的无效数据数达到结束条件(到复制结束为止应检测的有效数据数),因此结束复制动作。

[0508] 控制器4按各复制数据将表示数据ID及复制目标位置的数据复制信息,作为复制命令的命令完成回应通知给主机2。基于该数据复制信息,主机2更新地址变换表(LUT45),并将各个复制的数据映射至正确的物理地址。区块B11的页面P0的数据、页面P1的数据无效化。由此,区块B11变成不含有有效数据的空闲区块。同样地,区块B20的页面P0的数据、页面P2的数据也无效化。由此,区块B11、B20变成不含有有效数据的空闲区块。

[0509] 而且,控制器4将复制源区块B25的页面P1作为下一次的复制源开始页面进行决定,并且,将下一次的复制源开始页面的物理地址(复制源区块B25的区块地址、页面P1的页面地址)通知给主机2。

[0510] 而且,控制器4特定出保持可读出数据的复制目标区块B30内的最新页面地址及保持可读出数据的复制目标区块B31内的最新页面地址,并将这些最新页面地址通知给主机2。由此,主机2能够获知各复制目标区块内的页面群中的可读出的最后的页面。

[0511] 图45表示数据大小比页面大小小且指定应复制的有效数据数作为结束条件的情况下的数据复制动作。

[0512] 于此,作为一例,假定页面大小为16K字节、数据大小为4K字节的情况。该数据大小对应于用来管理各LBA与各物理地址之间的映射的所述管理大小。在复制源区块B0的复制开始页面P31,储存有分别具有4K字节的数据大小的数据D1、数据D2、数据D3、数据D4。在各数据的复制源区块B0的下一页面P32,储存有分别具有4K字节的数据大小的数据D5、数据D6、数据D7、数据D8。位图数据81表示数据D1、数据D2、数据D3、数据D4、数据D5、数据D6、数据D7、数据D8、…各自的有效/无效。

[0513] 这样,在数据大小比页面大小小的情况下,复制源区块B0的各页面包含由位图数

据81分别表示其有效/无效的多个数据。

[0514] 控制器4(1)从包含一个以上的有效数据的复制源区块B0内的页面分别以页面单位读出数据,(2)从读出数据抽取有效数据,由此准备与1页面的页面大小对应的数的有效数据,(3)将准备的与1页面的页面大小对应的数的有效数据从复制目标区块B10的复制目标区域(从传输目标开始页面开始的区域)以页面单位写入,由此将有效数据复制到复制目标区域且跳过无效数据的复制。并且,控制器4在复制到复制目标区域的有效数据的数为到复制结束为止应复制的有效数据的数以上的情况下,或跳过复制的无效数据的数为到复制结束为止应检测的无效数据的数以上的情况下,结束数据复制动作。

[0515] 也就是说,关于包含有效数据的各个页面,控制器4从这些页面以页面单位依次读出数据。若准备了1页面的有效数据(于此为4个有效数据),则控制器4将该1页面的有效数据以页面单位写入至复制目标区块。由此,能够只将有效数据以这些有效数据与页面单位对准的状态效率良好地复制到复制目标区块。在1页面的有效数据齐备之前没有复制源区块的数据的情况下,控制器4向当前准备的有效数据填充虚拟数据,并将由此获得的1页面的数据写入至复制目标区块。

[0516] 更详细来说,执行以下复制动作。以下,为了简化图示,假定到复制结束为止应复制的有效数据数为2的情况。

[0517] 控制器4首先判定复制源区块B0的复制开始页面P31是否包含有效数据。在图45的情况下,页面P31包含有效数据D1、D3。因此,控制器4从页面P31读出1页面的数据(D1~D4)。读出数据可以临时储存在复制缓冲器32。

[0518] 由于读出的有效数据的数只有D1、D3这2个,与1页面的大小对应的4个有效数据并未齐备,因此控制器4继续执行从复制源区块B0以页面单位读出数据的处理。

[0519] 控制器4判定复制源区块B0的页面P32是否包含有效数据。在图45的情况下,页面P32包含有效数据D5、D6。因此,控制器4从页面P32读出1页面的数据(D5~D8)。读出数据可以临时储存在复制缓冲器32。

[0520] 控制器4从页面P31读出的1页面的数据中抽取有效数据D1、D3,从页面P32读出的1页面的数据中抽取有效数据D5、D6,产生1页面的有效数据(D1、D3、D5、D6)。并且,控制器4将1页面的有效数据(D1、D3、D5、D6)复制到复制目标区块B10的传输目标开始页面P11。由此,跳过无效数据D2、D3、D7、D8的复制,只将有效数据以页面单位复制到复制目标区块B10。这样,使有效数据与页面单位对准的处理优先执行,在有效数据与页面单位对准后执行结束条件的判定。

[0521] 复制到复制目标区块B10的有效数据数变成4。复制到复制目标区块B10的有效数据数4为结束条件指定的有效数据数(于此为2)以上,因此结束复制动作。

[0522] 利用以上的复制动作,即便在各个数据的大小比页面大小小的情况下,也能将有效数据以这些有效数据与页面单位对准的状态效率良好地复制到复制目标区块。

[0523] 控制器4按各复制数据将表示数据ID与复制目标位置的数据复制信息,作为复制命令的命令完成回应通知给主机2。复制目标位置也可以用区块地址、页面地址、及页面内偏移来表现。与某个4K字节的数据对应的页面内偏移是表示储存该4K字节的数据的页面内的偏移位置的页面内地址。基于该数据复制信息,主机2更新地址变换表(LUT45),将各个复制数据(于此为D1、D3、D5、D6)映射至正确的物理地址。区块B0的页面P31的数据D1、D3、页面

P32的数据D5、D6无效化。

[0524] 而且,控制器4将复制源区块B0的页面P33作为下一次的复制源开始页面进行决定,并且,将下一次的复制源开始页面的物理地址(复制源区块B0的区块地址、页面P33的页面地址)通知给主机2。

[0525] 而且,控制器4特定出保持可读出数据的复制目标区块B10内的最新页面地址,并将该最新页面地址通知给主机2。

[0526] 图46表示数据大小比页面大小小且指定应检测的无效数据数作为结束条件的情况下的数据复制动作。以下,假定到复制结束为止应检测的无效数据数为2的情况。

[0527] 控制器4首先判定复制源区块B0的复制开始页面P31是否包含有效数据。在图46的情况下,页面P31包含有效数据D1、D3。因此,控制器4从页面P31读出1页面的数据(D1~D4)。读出数据可以临时储存在复制缓冲器32。

[0528] 由于读出的1页面的数据(D1~D4)所含的无效数据数只有D2、D4这2个,因此控制器4继续执行从复制源区块B0以页面单位读出数据的处理。

[0529] 控制器4判定复制源区块B0的页面P32是否包含有效数据。在图46的情况下,页面P32包含有效数据D5、D6。因此,控制器4从页面P32读出1页面的数据(D5~D8)。读出数据可以临时储存在复制缓冲器32。

[0530] 控制器4从页面P31读出的1页面的数据中抽取有效数据D1、D3,从页面P32读出的1页面的数据中抽取有效数据D5、D6,产生1页面的有效数据(D1、D3、D5、D6)。并且,控制器4将1页面的有效数据(D1、D3、D5、D6)复制到复制目标区块B10的传输目标开始页面P11。由此,跳过无效数据D2、D4、D7、D8的复制,只将有效数据以页面单位复制到复制目标区块B10。

[0531] 检测的无效数据数、即跳过复制的无效数据数变成4。由于检测的无效数据数4为结束条件指定的无效数据数(于此为2)以上,因此结束复制动作。

[0532] [数据复制命令]

[0533] 图47表示应用于SSD3的数据复制命令的输入参数。数据复制命令包含以下输入参数。

[0534] (1)复制源区块地址(复制源区块地址列表):该参数值表示复制源区块的区块地址。复制源区块的区块地址也能由虚拟区块地址及虚拟区块内区块编号来指定。复制源区块地址列表包含多个复制源区块各自的区块地址。也就是说,主机2能够指定一个以上的复制源区块。

[0535] (2)复制源区块内的复制开始位置:该参数值表示应复制数据的复制源区块内的开始位置(复制源开始页面)。在指定了多个复制源区块的情况下,可以只有第1个复制源区块内的复制开始位置由该参数值指定。

[0536] (3)复制目标区块地址:复制目标区块地址表示复制目标区块的区块地址。复制目标区块的区块地址也能由虚拟区块地址及虚拟区块内区块编号来指定。

[0537] (4)复制目标区块内的开始位置:复制目标区块内的开始位置表示应传输数据的复制目标区块内的开始位置(传输目标开始页面)。

[0538] (5)有效/无效位图:有效/无效位图是表示复制源区块内的有效数据及无效数据的配置的信息(位图数据)。

[0539] (6)到复制结束为止应复制的有效数据数:该参数值指定应传输即应复制到复制

目标区块的数据数(例页面数)。该参数值作为复制动作的结束条件使用。数据复制动作一边跳过无效数据的复制一边只复制有效数据。

[0540] (7)到复制结束为止应检测的无效数据数:该参数值也能作为复制动作的结束条件使用。主机2在一个复制命令中只能指定(6)的结束条件或(7)的结束条件的任一个。

[0541] (8)数据的大小(数据大小)

[0542] 此外,数据复制命令可以包含处理优先级等其他输入参数值。

[0543] 图48表示数据复制命令的输出参数。数据复制命令包含以下输出参数。

[0544] (1)到复制结束为止检测的无效数据数

[0545] (2)复制到复制目标区块的数据的总数:该参数值表示实际复制到复制目标区块的数据的总数。在指定到复制结束为止应检测的无效数据数作为结束条件的情况下,实际复制到复制目标区块的数据的总数不固定。因此,该参数值在指定到复制结束为止应检测的无效数据数作为结束条件的情况下有用。

[0546] (3)复制到复制目标区块的数据的标识符与复制目标的记录位置的组:该参数值表示将哪个有效数据复制到复制目标区块内的何处。

[0547] (4)下一次应开始复制的数据的位置:该参数值表示下一次的复制中的复制源区块内的复制开始位置。

[0548] (5)可读出的最新页面地址:该参数值表示复制目标区块内保持可读出数据的最新页面地址。

[0549] [数据复制处理的次序]

[0550] 图49的流程图表示在指定应复制的有效数据数作为结束条件的情况下的数据复制动作的次序。

[0551] 控制器4首先将数据复制命令指定的复制源区块内的复制开始位置(复制源开始页面)设定为当前复制对象页面。并且,控制器4基于位图数据判定当前复制对象页面是否为有效数据(步骤S91)。

[0552] 若当前复制对象页面为无效数据(步骤S91的否),则控制器4跳过当前复制对象页面的复制动作(步骤S92),将当前复制对象页面变更为下一页面(步骤S93)。

[0553] 若当前复制对象页面为有效数据(步骤S91的是),则控制器4从当前复制对象页面读出该有效数据(步骤S94),并将该读出的有效数据写入至复制目标区块的传输目标开始页面(步骤S95)。控制器4更新复制的有效数据数(步骤S96),并判定复制的有效数据数是否达到应复制的有效数据数(步骤S97)。

[0554] 若复制的有效数据数未达到应复制的有效数据数(步骤S97的否),则控制器4将当前复制对象页面变更为下一页面(步骤S98),然后再次执行步骤S91~S97的处理。

[0555] 若复制的有效数据数达到应复制的有效数据数(步骤S97的是),则控制器4执行结束处理(步骤S99)。在步骤S99中,控制器4创建返回值的数据,并将包含返回值的数据的命令完成回应发送至主机2。

[0556] 图50的流程图表示指定应检测的无效数据数作为结束条件的情况下的数据复制动作的次序。

[0557] 控制器4首先将数据复制命令指定的复制源区块内的复制开始位置(复制源开始页面)设定为当前复制对象页面。并且,控制器4基于位图数据,判定当前复制对象页面是否

为有效数据(步骤S101)。

[0558] 若当前复制对象页面为有效数据(步骤S101的是),则控制器4从当前复制对象页面读出该有效数据(步骤S102),并将该读出的有效数据写入至复制目标区块的传输目标开始页面(步骤S103)。控制器4将当前复制对象页面变更为下一页面(步骤S104),并进入步骤S101的处理。

[0559] 若当前复制对象页面为无效数据(步骤S101的否),则控制器4跳过当前复制对象页面的复制(步骤S105),更新检测的无效数据数(步骤S106),并判定检测的无效数据数是否达到应检测的无效数据数(步骤S107)。

[0560] 若检测的无效数据数未达到应检测的无效数据数(步骤S107的否),则控制器4将当前复制对象页面变更为下一页面(步骤S108),进入步骤S101的处理。

[0561] 若检测的无效数据数达到应检测的无效数据数(步骤S107的是),则控制器4执行结束处理(步骤S109)。在步骤S109中,控制器4创建返回值的数据,并将包含返回值的数据的命令完成回应发送至主机2。

[0562] 图51的流程图表示数据大小比页面大小小且指定应复制的有效数据数作为结束条件的情况下的数据复制动作的次序。

[0563] 控制器4首先将数据复制命令指定的复制源区块内的复制开始位置(复制源开始页面)设定为当前复制对象页面。并且,控制器4基于位图数据判定当前复制对象页面是否包含至少一个有效数据(步骤S111)。

[0564] 若当前复制对象页面为只包含无效数据的页面(步骤S111的否),则控制器4跳过当前复制对象页面的复制动作(步骤S112),将当前复制对象页面变更为下一页面(步骤S113),并进入步骤S111的处理。

[0565] 若当前复制对象页面为包含至少一个有效数据的页面(步骤S111的是),则控制器4以页面单位读出当前复制对象页面内的数据,并将读出的数据储存在复制缓冲器32(步骤S114)。控制器4从读出的数据只抽取有效数据,由此跳过无效数据,准备与页面大小对准的有效数据的组(步骤S115)。控制器4判定是否能够准备与页面大小对准的有效数据的组(具有1页面的大小的有效数据)(步骤S116)。

[0566] 若准备的有效数据的大小比1页面的大小少(步骤S116的否),则控制器4将当前复制对象页面变更为下一页面(步骤S117),并在此执行步骤S111~S115的处理。

[0567] 若能够准备与页面大小对准的有效数据的组(具有1页面的大小的有效数据)(步骤S116的是),则控制器4将该具有1页面的大小的有效数据写入至复制目标区块的传输目标开始页面(步骤S118)。控制器4更新复制的有效数据数(步骤S119),并判定复制的有效数据数是否变成应复制的有效数据数以上(步骤S120)。

[0568] 若复制的有效数据数比应复制的有效数据数少(步骤S120的否),则控制器4将当前复制对象页面变更为下一页面(步骤S117),然后再次执行从步骤S111开始的处理。

[0569] 若复制的有效数据数为应复制的有效数据数以上(步骤S120的是),则控制器4执行结束处理(步骤S121)。在步骤S121中,控制器4创建返回值的数据,并将包含返回值的数据的命令完成回应发送至主机2。

[0570] 图52的流程图表示数据大小比页面大小小且指定应检测的无效数据数作为结束条件的情况下的数据复制动作的次序。

[0571] 在图52的处理中,代替图51的步骤S119、S120,而执行步骤S131、S132的处理。

[0572] 即,在将与页面大小对准的有效数据的组(具有1页面的大小的有效数据)写入至复制目标区块的传输目标开始页面后(步骤S118),控制器4更新检测的无效数据数(步骤S131),并判定检测的无效数据数是否为应检测的无效数据数以上(步骤S132)。

[0573] 若检测的无效数据数比应检测的无效数据数少(步骤S132的否),则控制器4将当前复制对象页面变更为下一页面(步骤S117),然后再次执行从步骤S111开始的处理。

[0574] 若检测的无效数据数为应检测的无效数据数以上(步骤S132的是),则控制器4执行结束处理(步骤S121)。

[0575] [命名空间管理]

[0576] 图53表示SSD3的命名空间管理功能。

[0577] 在SSD3中,能够确保(预约)对NSID#1的命名空间指定的个数的区块,同样地,能够确保(预约)对NSID#n的命名空间指定的个数的区块。连接于主机2的用户端终端51(用户A)能够使用NSID#1对SSD3进行存取(读出、写入、删除),且连接于主机2的另一用户端终端51(用户B)能够使用NSID#n对SSD3进行存取(读出、写入、删除)。

[0578] 以下,假定用户A处理更新频率高的数据,用户B处理更新频率低的数据的情况。这种情况下,在NSID#1的命名空间中有写入扩增增加的可能性。写入扩增(WA)是以如下方式定义。

[0579] $WA = \frac{\text{“写入至SSD的数据的总量”}}{\text{“因来自主机的写入命令而写入至SSD的数据的总量”}}$

[0580] “写入至SSD的数据的总量”相当于因来自主机的写入命令而写入的数据的总量与因垃圾回收(数据复制动作)等而内部写入至SSD的数据的总量的和。

[0581] 写入扩增(WA)的增加会引起SSD3内的各区块的删除次数的增加。也就是说,写入扩增(WA)越大则区块的删除次数越容易快速达到其删除次数的上限值。结果,引起SSD3的耐久性及其寿命的劣化。

[0582] 因此,起因于向NSID#1的命名空间的写入的SSD3的消耗度,比起因于向NSID#n的命名空间的写入的SSD3的消耗度大。

[0583] SSD3的命名空间管理功能能够按各个命名空间管理区块(或虚拟区块)的合计删除次数,并将与主机2指定的特定的命名空间对应的合计删除次数作为该特定的命名空间引起的SSD3的消耗度的指标,通知给主机2。NSID#1的合计删除次数是由合计删除次数计数器300-1计数,NSID#n的合计删除次数是由合计删除次数计数器300-n计数。某个NSID的合计删除次数通过计数对分配给该NSID的区块执行的删除动作的次数而获得。

[0584] 利用该通知,主机2能够评价各个命名空间对SSD3的消耗程度。结果,主机2能够基于该评价结果,进行比如对合计删除次数多的命名空间追加确保更多个数的区块的对策。

[0585] 例如,主机软件可以针对NSID#1的命名空间,向SSD3请求超过与NSID#1的LBA范围对应的容量(用户数据容量)的充分个数的区块的确保。响应于该请求,SSD3的控制器4将指定个数的区块确保用于NSID#1的命名空间。

[0586] 在与NSID#1的LBA范围对应的容量为100G字节的情况下,主机软件可以向SSD3请求相当于100G字节的个数的物理区块的追加,而将相当于合计200G字节的个数的物理区块确保用于NSID#1的命名空间。从200G字节减去用户数据容量后的剩余100GB的物理资源作

为NSID#1的命名空间的过量供给区域而发挥功能。

[0587] 在另一实施方式中,主机软件可以基于确保用于NSID#1的命名空间用的区块的数、与对应于该命名空间的合计删除次数,决定对正使用该命名空间的用户A应收取的存储使用费(租借费)。合计删除次数越多,租借费设定得越高。

[0588] 图54表示SSD3的命名空间管理架构。

[0589] 控制器4利用共通空闲区块池90管理NAND存储器4的空闲区块,将若干区块从共通空闲区块池90分配给NSID#1的命名空间。这些分配区块用来储存与NSID#1的命名空间关联的数据。即,控制器将若干区块作为用来储存与NSID#1的命名空间关联的数据的区块而分配给NSID#1的命名空间。响应于来自主机2的用来对这些区块的一个进行读出、写入或删除的命令的接收,控制器4对这些区块的一个执行读出动作、写入动作、或删除动作。控制器4计数对这些区块执行的删除动作的次数。控制器4响应于来自主机2的请求取得与NSID#1的命名空间关联的删除次数的命令的接收,将删除动作的次数的计数值(NSID#1的合计删除次数)通知给主机2。关于NSID#n的命名空间,控制器4也执行与对NSID#1的命名空间的处理相同的处理。

[0590] 以下,说明命名空间管理架构。

[0591] 在SSD3中,对每个命名空间设有独立的虚拟闪存池。虚拟闪存池81用来管理对NSID#1的命名空间用确保(预约)的物理资源量(确保的区块的总数)。同样地,虚拟闪存池82用来管理对NSID#n的命名空间用确保(预约)的物理资源量(确保的区块的总数)。这种情况下,无须考虑应确保(预约)哪个区块,利用各虚拟闪存池只管理应确保(预约)的区块的个数即可。应确保的区块的个数在物理NAND存取管理API21中为物理区块的个数,在虚拟NAND存取管理API22中为虚拟区块的个数。

[0592] 各空闲区块是利用由多个命名空间共有的共通空闲区块池90进行管理。从各命名空间的虚拟闪存池返还的区块利用共通空闲区块池90进行管理。

[0593] 损耗平均是在从共通空闲区块池90向各命名空间分配新的区块(例如写入对象区块、或写入对象虚拟区块)时执行。在控制器4从主机2接收包含特定的NSID的区块分配命令(例如所述区块分配、删除命令)时,控制器4从共通空闲区块池90选择一个空闲区块。选择的空闲区块在使用物理NAND存取管理API21的情况下为物理区块,在使用虚拟NAND存取管理API22的情况下为虚拟区块。控制器4将该选择的区块分配给与该特定的NSID对应的命名空间,且从确保该命名空间用的区块的总数减去1。在从共通空闲区块池90选择空闲区块时,控制器4也可以选择最小删除次数的区块(最小删除次数的物理区块或最小删除次数的虚拟区块)。由此,能够将从NSID#n的命名空间返还的删除次数少的区块分配给频繁覆写数据的NSID#1的命名空间,因此能够实现命名空间之间的损耗平均。

[0594] 控制器4将确保NSID#1用的区块的总数、分配给NSID#1用的区块地址的列表、及NSID#1的合计删除次数等,作为与NSID#1的命名空间对应的管理信息进行管理。NSID#1的合计删除次数是通过计数对分配给NSID#1的各区块执行的删除动作的次数而获得。

[0595] NSID#1的命名空间管理也可以如下方式执行。以下,例示物理NAND存取管理API21用的命名空间管理。

[0596] 在控制器4从主机2接收包含NSID#1的命名空间分配命令时,控制器4将由命名空间分配命令指定的个数的区块确保用于NSID#1。确保用于NSID#1的区块的总数是利用虚拟

闪存池81进行管理。能够分配给NSID#1的命名空间的区块数的上限被限制为确保用于NSID#1用的区块的总数以下。

[0597] 在控制器4从主机2接收包含NSID#1的区块分配、删除命令时,控制器4从共通空闲区块池90选择最小删除次数的区块,将该选择的空闲区块分配给NSID#1的命名空间用,删除该分配的区块,并将该分配且删除的区块的物理地址通知给主机2,且从利用虚拟闪存池81管理的区块的总数、即能够分配给NSID#1的剩余区块数减去1。能够分配给NSID#1的剩余区块数表示能够分配给NSID#1的命名空间用的区块的当前个数。该分配及删除的区块能够作为例如NSID#1用的写入对象区块91使用。

[0598] 在若利用虚拟闪存池81管理的区块的当前总数(剩余区块数)为零的情况下,即便从主机2接收包含NSID#1的区块分配、删除命令,控制器4也不会对NSID#1的命名空间分配新的区块。

[0599] 在控制器4从主机2接收包含NSID#1的写入命令时,控制器4将写入命令指定的数据写入至写入对象区块91。写入命令可以包含应写入数据的物理地址(区块地址与页面地址双方)(直接地址指定模式),也可以只包含应写入数据的区块地址(自动地址产生模式),或者也可以只包含NSID#1。

[0600] 在写入命令只包含NSID#1的情况下,与自动地址产生模式的情况同样地,由控制器4自动产生应写入数据的物理地址。这种情况下,由写入命令指定的数据按照当前写入对象区块91内的页面P0~P255的顺序依次被写入。控制器4将写入数据的物理地址(区块地址与页面地址双方)通知给主机2。

[0601] 在当前写入对象区块91写满数据时,可以将该写入对象区块91移动到活动区块池92。活动区块池92管理由NSID#1当前使用的区块(活动区块)的列表。在当前写入对象区块91写满数据时,主机2可以向SSD3发送包含NSID#1的区块分配、删除命令,请求新的写入对象区块的分配、删除。

[0602] 主机2能够将活动区块池92内的任意区块读出或删除。此外,主机2能够将区块返还命令发送至SSD3,向SSD3请求将活动区块池92内的区块返还至共通空闲区块池90。例如,返还删除的区块、只包含因数据更新而无效化的数据的区块、或者只包含因所述数据复制动作等而无效化的数据的区块等。控制器4响应于区块返还命令的接收,将该区块返还命令指定的区块移动至共通空闲区块池90,且将利用虚拟闪存池81管理的区块的总数(剩余区块数)增加1。

[0603] 控制器4也管理与NSID#n的命名空间对应的管理信息、确保的区块的总数、分配的区块地址的列表、及NSID#n的合计删除次数等。

[0604] NSID#n的命名空间管理是以如下方式执行。

[0605] 在控制器4从主机2接收包含NSID#n的命名空间分配命令时,控制器4将由命名空间分配命令指定的个数的区块确保用于NSID#n。确保用于NSID#n的区块的总数是利用虚拟闪存池82进行管理。

[0606] 在控制器4从主机2接收包含NSID#n的区块分配、删除命令时,控制器4从共通空闲区块池90选择最小删除次数的区块,将该选择的空闲区块分配给NSID#n的命名空间用,并删除该分配的区块,且从利用虚拟闪存池82管理的区块的总数、即能够分配给NSID#n的剩余区块数减去1。该分配及删除的区块能够作为例如NSID#n用的写入对象区块93使用。

[0607] 在若利用虚拟闪存池82管理的区块的当前总数(剩余区块数)为零的情况下,即便从主机2接收包含NSID#n的区块分配、删除命令,控制器4也不会对NSID#n的命名空间分配新的区块。

[0608] 在控制器4从主机2接收包含NSID#n的写入命令时,控制器4将写入命令指定的数据写入至写入对象区块93。写入命令可包含应写入数据的物理地址(区块地址与页面地址双方)(直接地址指定模式),也可只包含应写入数据的区块地址(自动地址产生模式),或者还可以只包含NSID#n。

[0609] 在写入命令只包含NSID#n的情况下,与自动地址产生模式的情况同样地,由控制器4自动产生应写入数据的物理地址。这种情况下,由写入命令指定的数据依次写入至当前写入对象区块93内的页面P0~P255。控制器4将写入数据的物理地址(区块地址与页面地址双方)通知给主机2。

[0610] 在当前写入对象区块93写满数据时,可以将该写入对象区块93移动至活动区块池94。活动区块池94管理由NSID#n当前使用的区块的列表。在当前写入对象区块93写满数据时,主机2可以将包含NSID#n的区块分配、删除命令发送至SSD3,请求新的写入对象区块的分配、删除。

[0611] 主机2能够将活动区块池94内的任意区块读出或删除。此外,主机2能够将区块返还命令发送至SSD3,向SSD3请求活动区块池94内的区块的返还。控制器4响应于区块返还命令的接收,将该区块返还命令指定的区块移动至共通空闲区块池90,并将利用虚拟闪存池82管理的区块的总数(剩余区块数)增加1。

[0612] 虚拟NAND存取管理API22用的命名空间管理也能够按照与物理NAND存取管理API21用的命名空间管理相同的次序执行。在虚拟NAND存取管理API22用的命名空间管理中,可以代替管理确保的区块的个数,而管理确保的虚拟区块的个数,代替管理分配的区块地址的列表,而管理分配的虚拟区块地址的列表。

[0613] 此外,在虚拟NAND存取管理API22用的命名空间管理中,可以将通过计数对分配给NSID#1的虚拟区块执行的删除动作的次数而获得的计数值作为NSID#1的合计删除次数进行管理,也可以将通过计数对分配给NSID#n的虚拟区块执行的删除动作的次数而获得的计数值作为NSID#n的合计删除次数进行管理。

[0614] [命名空间分配命令]

[0615] 图55表示命名空间分配命令。命名空间分配命令是向SSD3请求确保(或追加)由该命名空间分配命令指定的数的区块。

[0616] 命名空间分配命令包含以下输入参数。

[0617] (1)NSID:该输入参数值表示对象命名空间的标识符(ID)。

[0618] (2)物理资源量:该物理资源量表示应确保的区块的数。在物理NAND存取管理API21中,应确保的区块的数是以区块(物理区块)的粒度指定,在虚拟NAND存取管理API22中,应确保的区块的数是以虚拟区块(构成一个虚拟区块的多个区块)的粒度指定。

[0619] 此外,命名空间分配命令还可以包含表示处理优先级的输入参数。

[0620] 命名空间分配命令包含以下输出参数。

[0621] (1)物理资源量:该物理资源量表示确保的区块的数。在物理NAND存取管理API21中,确保的区块的数是以区块(物理区块)的粒度表示,在虚拟NAND存取管理API22中,确保

的区块的数是以虚拟区块(构成一个虚拟区块的多个区块)的粒度表示。

[0622] [命名空间分配处理的次序]

[0623] 图56的流程图表示SSD3执行的命名空间分配处理的次序。

[0624] SSD3的控制器4从主机2接收命名空间分配命令(步骤S141)。基于共通空闲区块池90内的剩余区块的个数,判定是否能够确保由命名空间分配命令内的输入参数(物理资源量)指定的个数的区块(步骤S142)。如上所述,在物理NAND存取管理API21中,应确保的区块的数是以区块(物理区块)的粒度指定,在虚拟NAND存取管理API22中,应确保的区块的数是以虚拟区块(构成一个虚拟区块的多个区块)的粒度指定。

[0625] 若剩余区块(或剩余虚拟区块)的个数为指定个数以上(步骤S142的是),则控制器4对由命名空间分配命令指定的NSID的命名空间确保指定个数的区块(或虚拟区块)(步骤S143),然后将包含表示确保的区块(或虚拟区块)的个数的输出参数的命令完成回应发送至主机2(步骤S144)。

[0626] 若剩余区块(或剩余虚拟区块)的个数比指定个数少(步骤S142的否),则控制器4向主机2通知错误(步骤S145)。得到错误报告的主机2可以变更应确保的区块(或虚拟区块)的个数。

[0627] [命名空间用的区块分配、删除命令]

[0628] 图57表示命名空间用的区块分配、删除命令。

[0629] 命名空间用的区块分配、删除命令包含以下输入参数。

[0630] (1)处理优先级:处理优先级表示该命令的优先级。

[0631] (2)NSID:NSID表示应分配区块(或虚拟区块)的命名空间的ID。

[0632] 命名空间用的区块分配、删除命令包含以下输出参数。

[0633] (1)结束状态:向主机2返回表示区块分配、删除命令的成功或错误的结束状态。

[0634] (2)区块地址:向主机2返回分配的区块的区块地址(或分配的虚拟区块的虚拟区块地址)。

[0635] (3)剩余区块数:向主机2返回确保用于该NSID用的剩余区块数(或剩余虚拟区块的数)。

[0636] [命名空间用的区块分配、删除处理的次序]

[0637] 图58的流程图表示由SSD3执行的区块分配、删除处理的次序。于此,例示说明物理NAND存取管理API21用的区块分配、删除处理。

[0638] SSD3的控制器4从主机2接收包含NSID的区块分配、删除命令(步骤S151)。控制器4判定是否有该NSID用的剩余区块(步骤S152)。

[0639] 若有该NSID用的剩余区块(步骤S152的是),则控制器4从共通空闲区块池90分配一个区块作为指定的NSID用的写入对象区块,并自动删除分配的区块(步骤S152)。控制器4从指定的NSID用的剩余区块数减去1(步骤S154)。控制器4更新与指定的NSID对应的合计删除次数(步骤S155)。在步骤S155中,控制器4将与该指定的NSID对应的合计删除次数增加1。并且,控制器4产生返回值(输出参数),并将包含返回值的命令完成回应发送至主机2(步骤S156)。

[0640] 另一方面,若无该NSID用的剩余区块(步骤S152的否),则控制器4将包含表示因确保该NSID用的剩余区块数为零所以无法分配新区块的错误状态的命令完成回应,发送至主

机2(步骤S157)。

[0641] 在虚拟NAND存取管理API22用的区块分配、删除处理中,控制器4判定是否有该NSID用的剩余虚拟区块(步骤S152)。

[0642] 若有该NSID用的剩余虚拟区块(步骤S152的是),则控制器4从共通空闲区块池90分配一个虚拟区块作为指定的NSID用的写入对象区块(写入对象虚拟区块),并自动删除分配的虚拟区块(步骤S152)。控制器4从指定的NSID用的剩余虚拟区块数减去1(步骤S154)。控制器4更新与指定的NSID对应的合计删除次数(步骤S155)。在步骤S155中,控制器4将与该指定的NSID对应的合计删除次数增加1。并且,控制器4产生返回值(输出参数),并将包含返回值的命令完成回应发送至主机2(步骤S156)。

[0643] 另一方面,若无该NSID用的剩余虚拟区块(步骤S152的否),则控制器4将包含表示因确保该NSID用的剩余虚拟区块数为零所以无法分配新区块的错误状态的命令完成回应,发送至主机2(步骤S157)。

[0644] [命名空间用的删除命令]

[0645] 图59表示用来删除分配给某个命名空间的特定区块的删除命令。

[0646] 该删除命令包含以下输入参数。

[0647] (1)区块地址:该输入参数值表示应删除的区块的区块地址。在虚拟NAND存取管理API22用的删除命令中,该输入参数值代替表示区块地址而表示应删除的虚拟区块的虚拟区块地址。

[0648] (2)处理优先级:该输入参数值表示该命令的优先级。

[0649] (3)NSID:该输入参数值表示与应删除的区块(或虚拟区块)对应的NSID。

[0650] 该删除命令包含以下输出参数。

[0651] (1)结束状态:向主机返回该删除命令的成功或错误。

[0652] (2)区块地址:该输出参数值表示删除的区块的区块地址。在虚拟NAND存取管理API22用的删除命令中,该输出参数值表示删除的虚拟区块的虚拟区块地址。

[0653] (3)合计删除次数:该输出参数值表示分配给该NSID的各区块的删除次数的合计。在虚拟NAND存取管理API22用的删除命令中,该输出参数值表示分配给该NSID的各虚拟区块的删除次数的合计。

[0654] [删除处理的次序]

[0655] 图60的流程图表示SSD3执行的删除处理的次序。于此,例示说明物理NAND存取管理API21用的删除处理。

[0656] SSD3的控制器4从主机2接收删除命令(步骤S171)。控制器4执行用来删除由该删除命令指定的区块的删除动作(步骤S172),将该删除的区块的删除次数增加1(步骤S173),并且将与分配该删除的区块的NSID对应的合计删除次数增加1(步骤S174)。在删除命令包含NSID的情况下,分配该删除的区块的NSID是由该删除命令内的NSID特定。另一方面,在删除命令不含NSID的情况下,分配该删除的区块的NSID可以基于该删除的区块的区块地址、及分配给各NSID的区块地址的列表来特定。

[0657] 并且,产生返回值(输出参数),将包含返回值的命令完成回应发送至主机2(步骤S175)。

[0658] 在虚拟NAND存取管理API22用的删除处理中,控制器4执行用来删除由该删除命令

指定的虚拟区块内的多个区块的删除动作(步骤S172),将该删除的虚拟区块的删除次数增加1(步骤S173),并且将与分配该删除的虚拟区块的NSID对应的合计删除次数增加1(步骤S174)。

[0659] [区块返还命令]

[0660] 图61表示应用于SSD3的区块返还命令。该区块返还命令用来将与特定的命名空间对应的活动区块池内的区块返还至共通空闲区块池90。

[0661] 该区块返还命令包含以下输入参数。

[0662] (1)区块地址:该输入参数值表示应返还的区块的区块地址。在虚拟NAND存取管理API22用的区块返还命令中,该输入参数值代替表示区块地址,而表示应返还的虚拟区块的虚拟区块地址。

[0663] (2)处理优先级:该输入参数值表示该命令的优先级。

[0664] (3)NSID:该输入参数值表示与应返还的区块(或虚拟区块)对应的NSID。

[0665] 该区块返还命令包含以下输出参数。

[0666] (1)结束状态:向主机返回该区块返还命令的成功或错误。

[0667] (2)剩余区块数:该输出参数值表示区块返还后的剩余区块数。在虚拟NAND存取管理API22用的区块返还命令中,该输出参数值表示虚拟区块返还后的剩余虚拟区块数。

[0668] [区块返还处理的次序]

[0669] 图62的流程图表示SSD3执行的区块返还处理的次序。于此,例示说明物理NAND存取管理API21用的区块返还处理。

[0670] SSD3的控制器4从主机2接收区块返还命令(步骤S181)。控制器4将由该区块返还命令指定的区块,从与由该区块返还命令指定的NSID对应的活动区块池移动至共通空闲区块池90(步骤S182)。由此,对该NSID的该区块的分配被解除,该区块作为不含有效数据的空闲区块而由共通空闲区块池90管理。

[0671] 控制器4将与指定的NSID对应的剩余区块数增加1(步骤S183)。并且,控制器4产生返回值(输出参数),并将包含返回值的命令完成回应发送至主机2(步骤S184)。

[0672] 在虚拟NAND存取管理API22用的区块返还处理中,控制器4将由该区块返还指定的虚拟区块移动至共通空闲区块池90(步骤S182),并将与指定的NSID对应的剩余虚拟区块数增加1(步骤S183),然后将包含返回值的命令完成回应发送至主机2(步骤S184)。

[0673] [删除次数获取命令]

[0674] 图63表示应用于SSD3的命名空间用的删除次数获取命令。该删除次数获取命令对SSD3请求应通知特定的命名空间的合计删除次数。

[0675] 删除次数获取命令包含以下输入参数。

[0676] (1)NSID:该输入参数值表示对象NSID。在该输入参数值为某个特殊值的情况下,也可以将所有NSID作为对象NSID进行决定。

[0677] (2)处理优先级:该输入参数值表示该命令的优先级。

[0678] 删除次数获取命令包含以下输出参数。

[0679] (1)结束状态:向主机返回该删除次数获取命令的成功或错误。

[0680] (2)合计删除次数:该输出参数值表示指定的NSID的合计删除次数。

[0681] [删除次数通知处理]

[0682] 图64的流程图表示SSD3执行的删除次数通知处理的次序。于此,例示说明物理NAND存取管理API21用的删除次数通知处理。

[0683] 在SSD3的控制器4从主机2接收删除次数获取命令时,控制器4判定指定由该删除次数获取命令特定的NSID、还是指定所有NSID(步骤S191)。

[0684] 若指定特定的NSID(步骤S191的是),则控制器4取得该特定的NSID的当前合计删除次数(计数值)(步骤S192),并将包含该特定的NSID的当前合计删除次数的命令完成回应发送至主机2(步骤S194)。

[0685] 若指定所有NSID(步骤S191的否),则控制器4取得与所有NSID分别对应的当前合计删除次数的列表(步骤S193),并将包含与所有NSID分别对应的当前合计删除次数的列表的命令完成回应发送至主机2(步骤S194)。

[0686] 在虚拟NAND存取管理API22用的删除次数通知处理中,计数对分配给指定的特定的NSID的各虚拟区块执行的删除动作的次数,并将该计数值作为指定的特定的NSID的合计删除次数通知给主机2。

[0687] 于此,对将由删除次数获取命令指定的NSID的合计删除次数通知给主机2的处理进行了说明,但也能代替NSID,而使用包含指定区块地址或虚拟区块地址的任一个的参数的删除次数获取命令。

[0688] 物理NAND存取管理API21用的删除命令包含区块地址(指定区块的物理地址)。控制器将从主机2接收的删除命令内的区块地址指定的区块删除,管理组#Y内的各区块的删除次数,在从主机2接收包含某个区块地址的删除次数获取命令的情况下,将包含由该删除次数获取命令内的区块地址指定的区块的删除次数通知给主机2。

[0689] 虚拟NAND存取管理API22用的删除命令包含虚拟区块地址(指定虚拟区块的物理地址)。控制器将从主机2接收的删除命令内的虚拟区块地址指定的虚拟区块内的多个区块删除,管理组#X内的各虚拟区块的删除次数,在从主机2接收包含某个虚拟区块地址的删除次数获取命令的情况下,将该删除次数获取命令内的虚拟区块地址指定的虚拟区块的删除次数通知给主机2。

[0690] [命名空间管理用的其他命令]

[0691] 控制器4还支持命名空间解体命令。命名空间解体命令向控制器4请求命名空间的删除(解体)。命名空间解体命令也可以包含表示当前分配给该特定的NSID的区块的数(应解体的区块数)的输入参数。该区块的数在物理NAND存取管理API21中是当前分配给该特定的NSID的物理区块的数,在虚拟NAND存取管理API22中是当前分配给该特定的NSID的虚拟区块的数。响应于来自主机2的包含特定的NSID的命名空间解体命令的接收,控制器4将与该特定的NSID对应的虚拟闪存池内的所有区块作为空闲区块移动至共通空闲区块90。并且,控制器4对于命名空间解体命令的命令完成的回应发送至主机2。该命令完成的回应也可以包含表示解体区块的数的返回值。

[0692] 主机2通过使用所述命名空间分配命令及命名空间解体命令,能够增减命名空间。

[0693] [主机构成]

[0694] 图65表示作为主机2发挥功能的信息处理装置的硬件构成例。

[0695] 该信息处理装置是作为服务器计算机、或个人计算机而实现。该信息处理装置包含处理器(CPU)101、主存储器102、BIOS—ROM103、网络控制器105、周边接口控制器106、控

制器107、及嵌入式控制器(EC)108等。

[0696] 处理器101是以控制该信息处理装置的各组件的動作的方式构成的CPU。该处理器101执行从多个SSD3的任1个加载到主存储器102的各种程序。主存储器102包含比如DRAM的随机存取存储器。处理器101执行的程序包含所述应用软件层41、OS42、文件系统43、及FTL44。处理器101执行的程序也可以还包含资源管理部45。

[0697] 该资源管理部45可以将删除次数获取命令发送至SSD3,从SSD3取得各命名空间的合计删除次数,基于取得的各命名空间的合计删除次数,判定起因于各个命名空间的SSD3的物理资源的消耗度。在假设起因于某个特定命名空间的SSD3的物理资源的消耗度大于阈值的情况下,该资源管理部45可以执行用来增加应确保用于该特定的命名空间的区块的个数的处理。这种情况下,该资源管理部45可以向SSD3发送请求应将指定个数的区块追加至该特定的命名空间用的命名空间分配命令。由此,该特定的命名空间的过量供给区域的大小增加,因此能够减少该特定的命名空间的写入扩增,结果,能够降低起因于该特定的命名空间的SSD3的物理资源的消耗度。

[0698] 此外,如上所述,数据中心提供者也可以基于确保用于该命名空间的区块的数(或虚拟区块的数)、及与该命名空间对应的合计删除次数,来决定与该命名空间对应的租借费。这种情况下,资源管理部45可以执行用来支援数据中心提供者决定租借费的服务。例如,某个命名空间相关的基本租借费用首先由与该命名空间对应的区域的容量(区块的个数)决定。并且,由合计删除次数的函数求出的追加费加上基本租借费用后的合计费可以作为该命名空间相关的租借费用进行决定。

[0699] 租借该命名空间的用户在请求追加确保用于该命名空间用的区块的个数的情况下,资源管理部45可以免除追加费收取,只基于追加区块的数与已确保的区块的数的合计来设定新的租借费用。并且,资源管理部45可以将请求应将用户指定的个数的区块追加至该特定的命名空间用的命名空间分配命令发送至SSD3。

[0700] 此外,资源管理部45能够与FTL44共同动作,让各应用程序使用物理NAND存取管理API/虚拟NAND存取管理API来控制SSD3的NAND存储器5。例如,资源管理部45响应于来自用户(某个应用程序、或某个用户端终端)的包含指定NAND存储器5内的物理NAND存取管理API21用的区块中的一个区块的物理地址的第1读出、写入或删除请求的接收,将所述读出命令、写入命令、或删除命令发送至SSD3,控制对该指定区块的读出动作、写入动作、或删除动作。此外,资源管理部45响应于来自用户(某个应用程序、或某个用户端终端)的包含指定NAND存储器5内的虚拟区块中的一个虚拟区块的物理地址(虚拟区块地址)的第2读出、写入或删除请求的接收,将所述读出命令、写入命令、或删除命令发送至SSD3,控制对该指定虚拟区块的读出动作、写入动作、或删除动作。

[0701] 而且,资源管理部45还可以执行以下控制:基于从SSD3通知的保持可读出数据的最新页面地址,自动变更应面向来自用户(某个应用程序、或某个用户端终端)的读出请求的存储区域(将写入数据保持一定期间的主机2内的存储器、或SSD3内的NAND存储器5)。这种情况下,资源管理部45可以在写入SSD3的某个特定数据变得可读出之前,响应于对该特定数据的读出请求对主机2内的存储器进行读出存取,当特定数据变得可读出之后,响应于与该特定数据的读出请求,对SSD3内的NAND存储器5进行读出存取。

[0702] 此外,资源管理部45能够与FTL44共同动作,管理多个命名空间。资源管理部45将

用来对SSD3内的NAND存储器5内的多个区块内分配给第1命名空间的多个第1区块的一个进行读出、写入或删除的命令,发送至SSD3。并且,资源管理部45通过将请求与所述第1命名空间关联的删除次数的取得的命令发送至SSD3,而从SSD3取得通过计数对多个第1区块执行的删除动作的次数所得的计数值。

[0703] 而且,资源管理部45通过将用来确保第1命名空间用的区块的命令发送至SSD3,让SSD3确保第1个数的区块用于第1命名空间,并将能够分配给第1命名空间的多个第1区块的数的上限限制为第1个数以下。此外,资源管理部45构成为通过将用来追加应确保用于第1命名空间的区块的个数的命令发送至SSD3,让SSD3追加第2个数的区块作为应确保用于第1命名空间的区块,并能够分配给第1命名空间的数的第1区块的数的上限增加为第1个数与第2个数的和。

[0704] 而且,资源管理部45通过将所述数据复制命令发送至SSD3,而控制垃圾回收(数据复制动作)。

[0705] 即,处理器101通过执行FTL44,来管理各逻辑区块地址与各NAND存储器5的物理地址之间的映射。处理器101进一步通过执行资源管理部45,来控制垃圾回收(数据复制动作)。

[0706] 这种情况下,处理器101将请求只复制有效数据的数据复制命令发送至SSD3。数据复制命令如上所述包含复制源区块、复制源区块内的复制开始页面、复制目标区块、复制目标区块内的传输开始页面、表示复制源区块内的各页面的数据为有效数据或无效数据的哪一个的位图数据、及指定到复制结束为止应复制的有效数据的数或到复制结束为止应检测的无效数据的数的任一个的结束条件。

[0707] 并且,处理器101从SSD3接收按复制到复制目标区块的各有效数据而表示有效数据的标识符将储存有效数据的复制目标区块内的位置的数据复制信息。处理器101基于数据复制信息来更新地址变换表。

[0708] 而且,处理器101基于从SSD3通知的下一应开始复制的数据的位置,控制下一轮的垃圾回收(数据复制动作)。

[0709] 此外,处理器101还执行作为非易失性存储器的BIOS-ROM103内储存的基本输入输出系统(BIOS)。BIOS是硬件控制用的系统程序。

[0710] 网络控制器105是比如有线LAN控制器、无线LAN控制器的通信装置。周边接口控制器106构成为执行与比如USB装置的周边装置的通信。

[0711] 控制器107构成为执行与分别连接于多个连接器107A的装置的通信。在本实施方式中,多个SSD3分别连接于多个连接器107A。控制器107为SAS expander、PCIe Switch、PCIe expander、闪存阵列控制器、或RAID控制器等。

[0712] EC108作为构成为执行信息处理装置的电力管理的系统控制器发挥功能。EC108根据用户对电源开关的操作而将信息处理装置通电及断电。EC108作为比如单片微控制器的处理电路而实现。EC108也可以内置控制键盘(KB)等输入装置的键盘控制器。

[0713] 图66表示包含多个SSD3及主机2的信息处理装置的构成例。

[0714] 该信息处理装置具备能够收容在支架上的薄箱形的壳体201。多个SSD3也可以配置在壳体201内。这种情况下,各SSD3可拆卸地插入至设置在壳体201的前表面201A的插槽中。

[0715] 系统板(主板)202配置在壳体201内。在系统板(主板)202上安装着包含CPU101、存储器102、网络控制器105、控制器107的各种电子零件。这些电子零件作为主机2发挥功能。

[0716] 如以上说明那样,根据本实施方式的命名空间管理功能,能够对每个命名空间计数删除动作的次数,根据来自主机2的请求,将各命名空间的删除次数通知给主机2。因此,能够对主机2提供关于各个命名空间引起的NAND存储器5的消耗程度的尺度,主机2能够执行用来抑制各个命名空间引起的NAND存储器5的消耗的对策。

[0717] 另外,在本实施方式中,例示了NAND存储器作为非易失性存储器。但是,本实施方式的功能例如也可以应用于例如MRAM(Magnetoresistive Random Access Memory,磁阻式随机存取存储器)、PRAM(Phase change Random Access Memory,相位变换随机存取存储器)、ReRAM(Resistive Random Access Memory,电阻式随机存取存储器)或FeRAM(Ferroelectric Random Access Memory,铁电随机存取存储器)等其他各种非易失性存储器。

[0718] 虽对本发明的若干实施方式进行了说明,但这些实施方式是作为示例而提出的,并不意图限定发明的范围。这些新颖的实施方式能以其他方式加以实施,且能够在不脱离发明主旨的范围内进行各种省略、替换、变更。这些实施方式或其变化包含在发明的范围或主旨内,并且包含在权利要求所记载的发明及其均等的范围内。

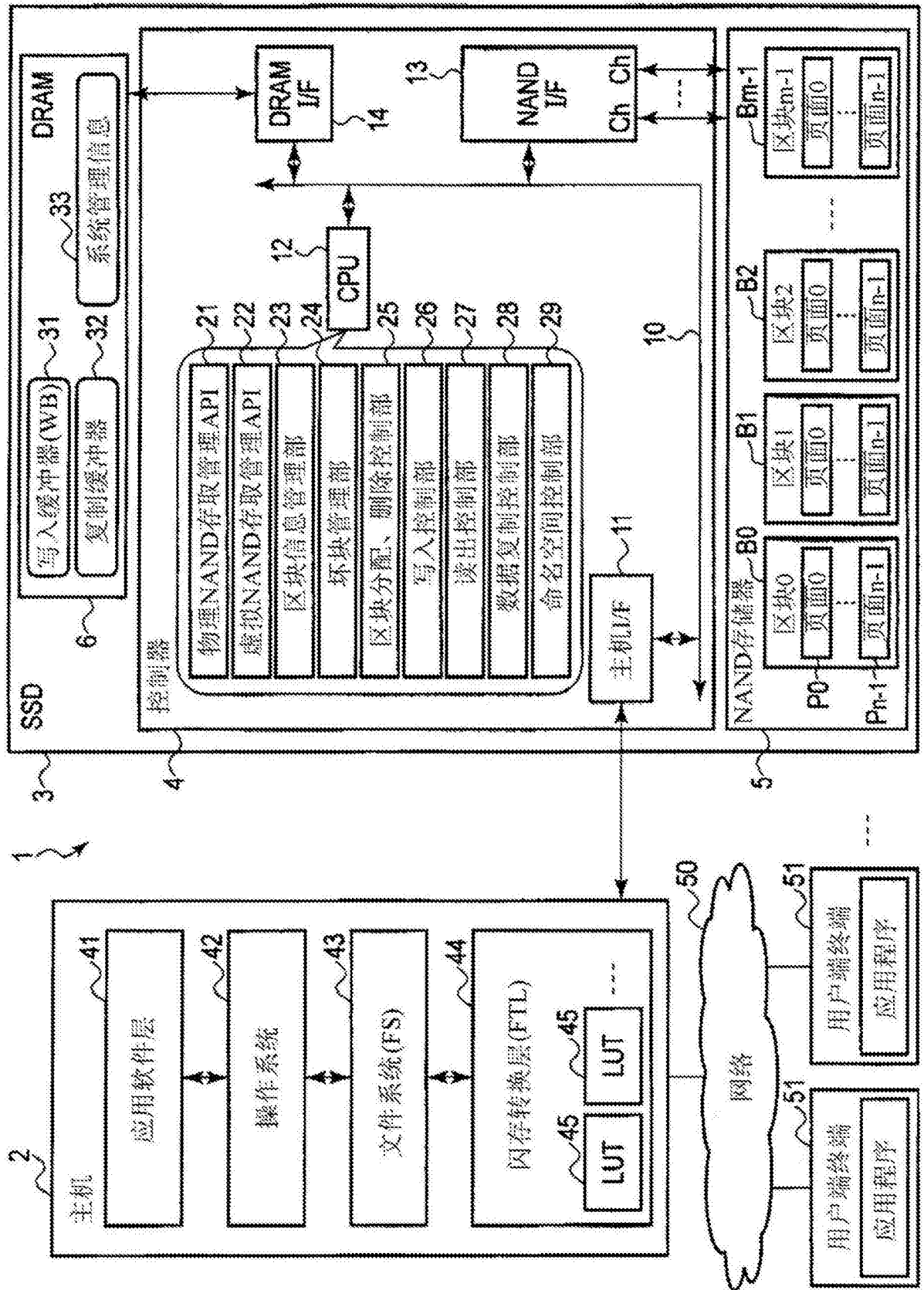


图1

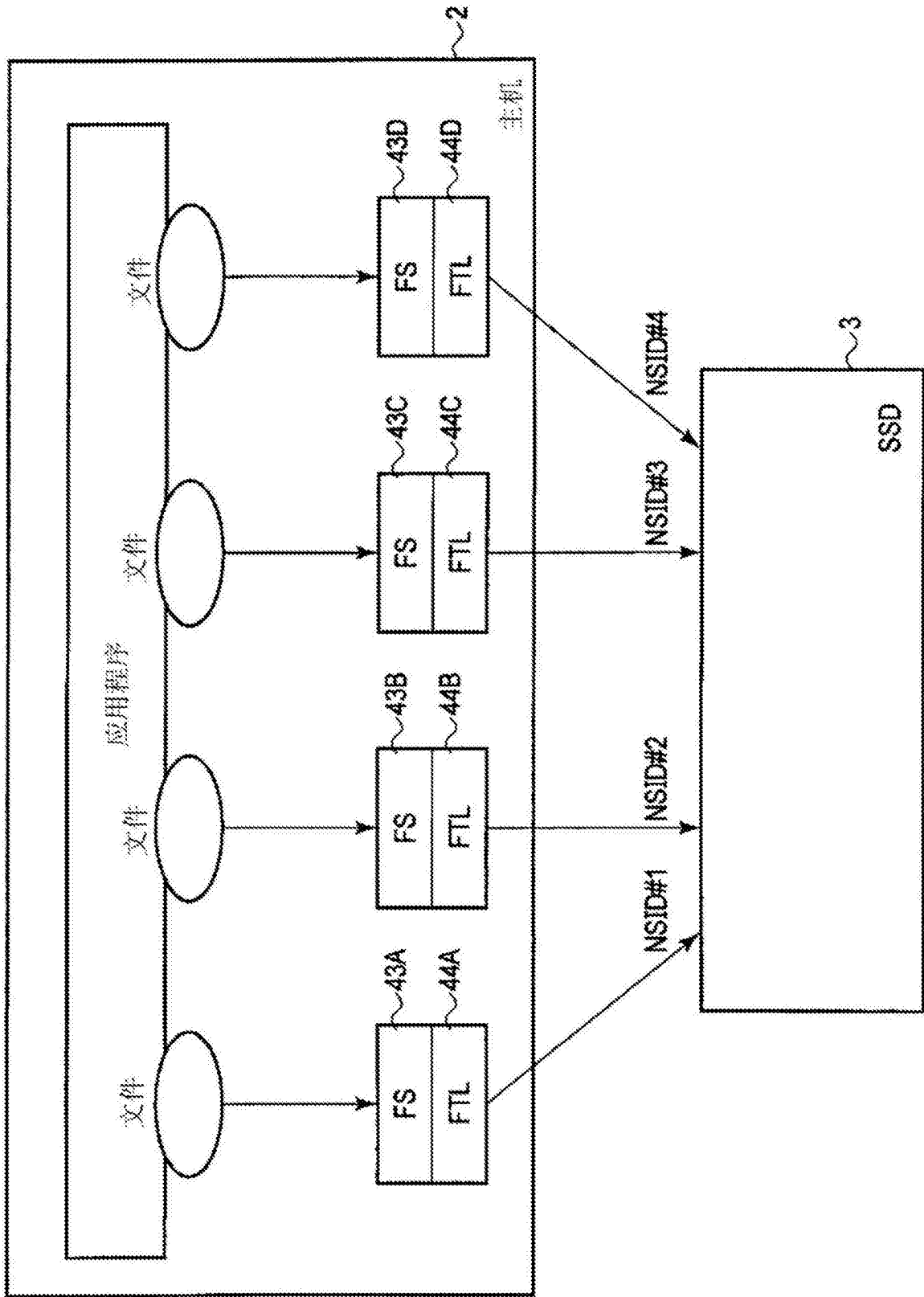


图2

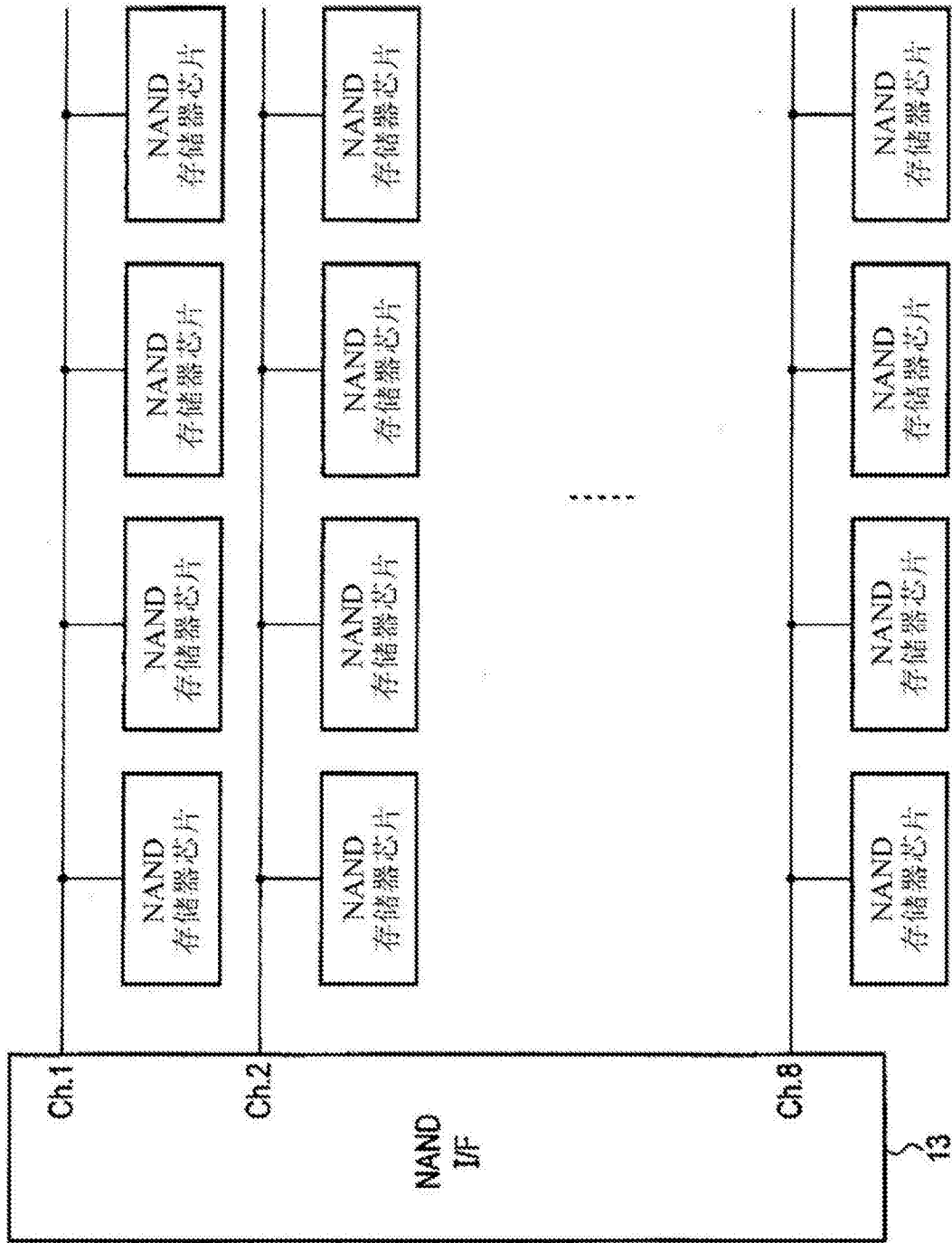


图3

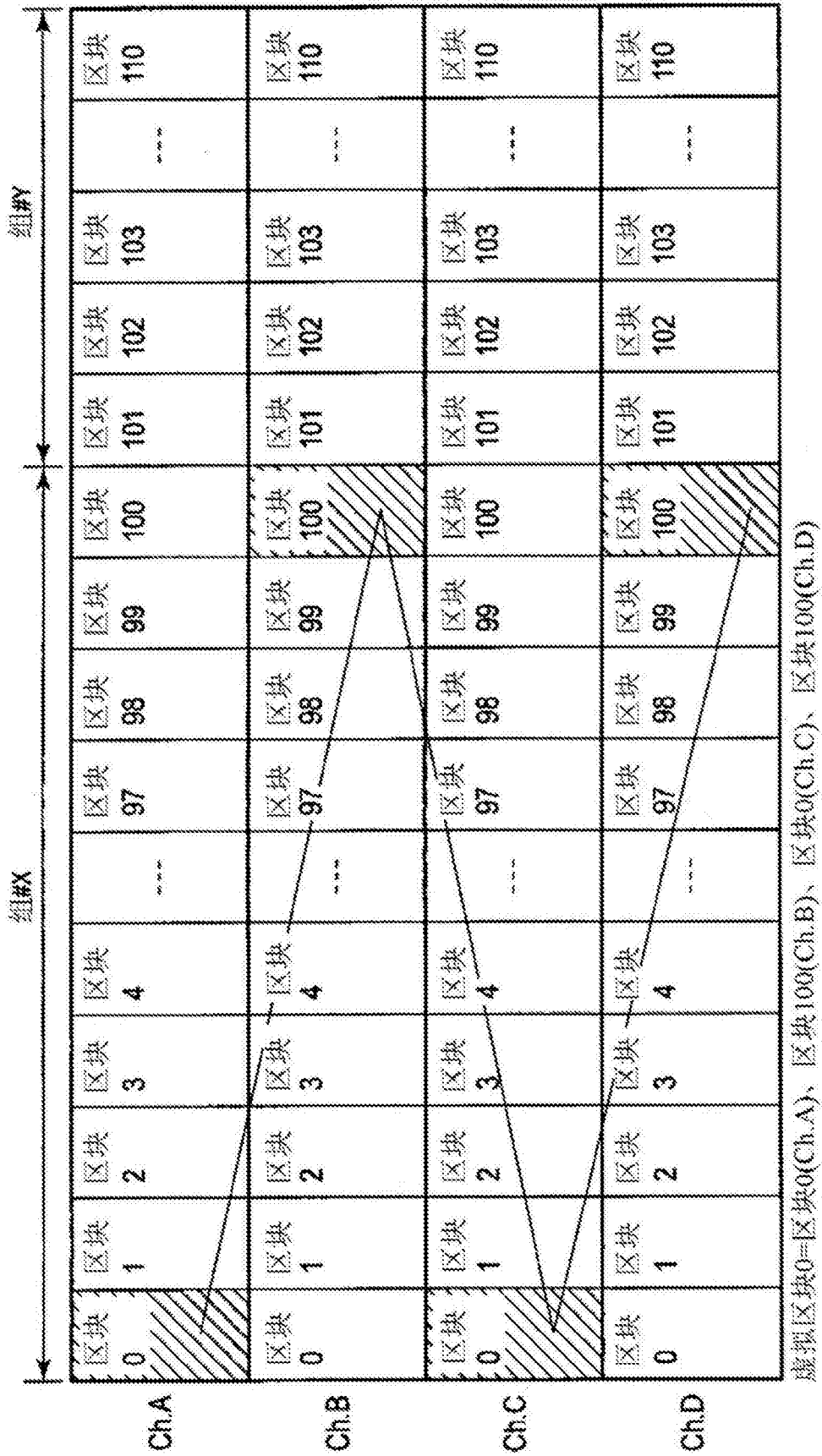


图4

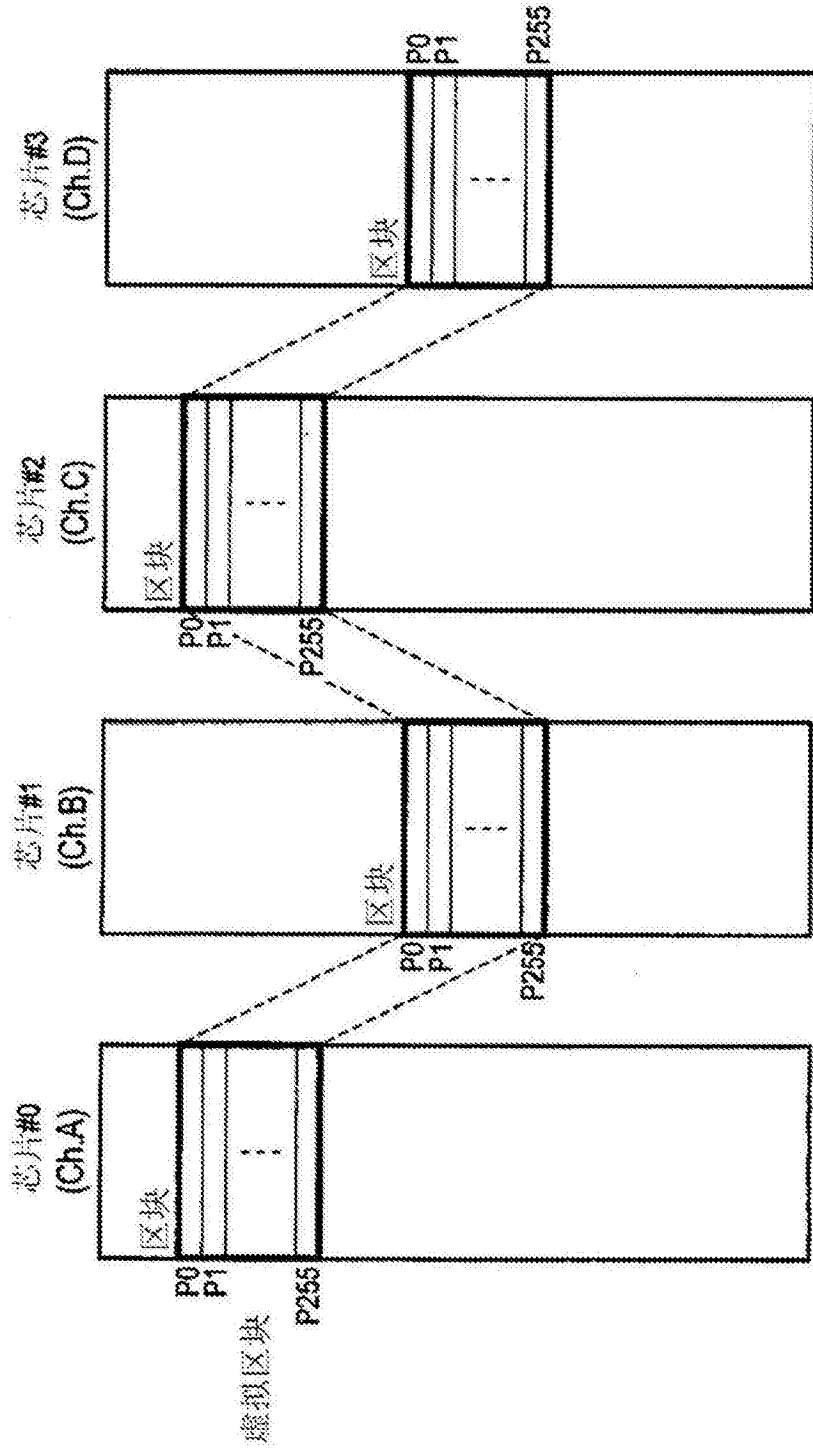


图5

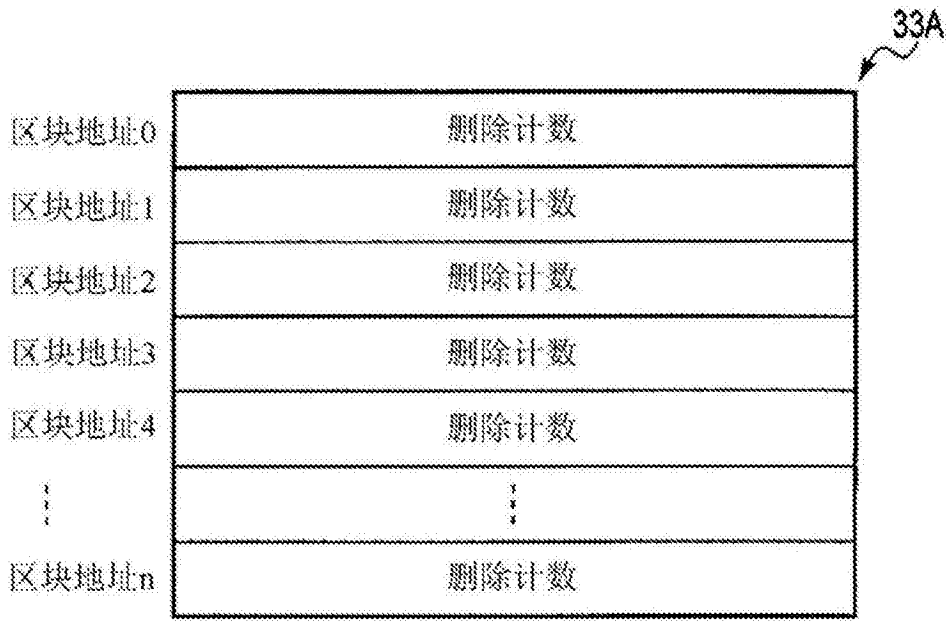


图6

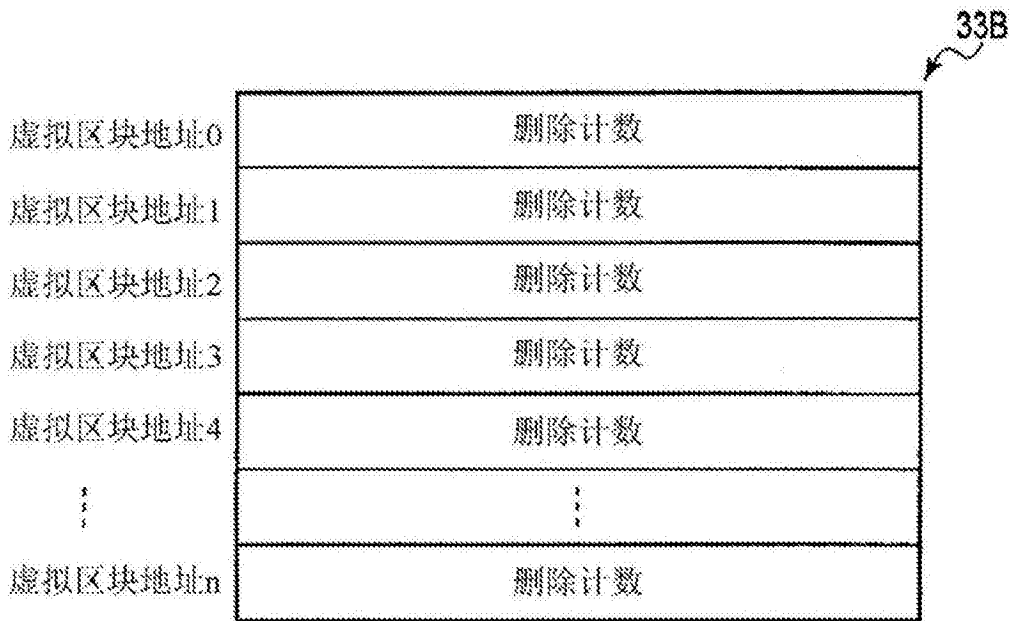


图7

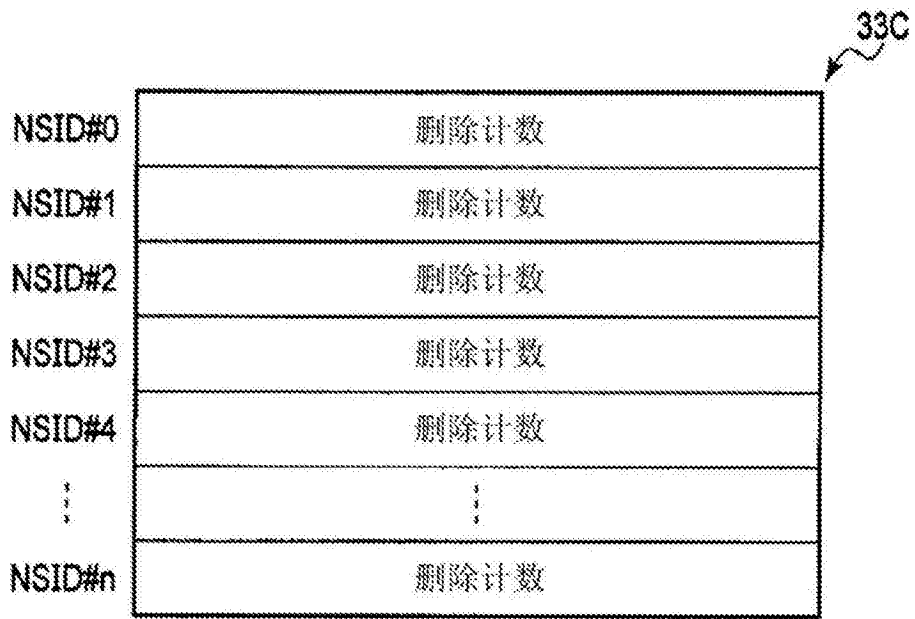


Diagram 33C shows a table with two columns. The left column lists NSID identifiers from NSID#0 to NSID#n, with vertical ellipsis between NSID#4 and NSID#n. The right column lists '删除计数' (Delete Count) for each corresponding NSID. A reference numeral 33C is located at the top right of the table.

NSID#0	删除计数
NSID#1	删除计数
NSID#2	删除计数
NSID#3	删除计数
NSID#4	删除计数
⋮	⋮
NSID#n	删除计数

图8



Diagram 33D shows a table with six rows. The rows contain the following text from top to bottom: '物理区块结构信息' (Physical Block Structure Information), '区块大小 (容量)' (Block Size (Capacity)), '页面大小 (容量)' (Page Size (Capacity)), '推定写入时间' (Estimated Write Time), '删除时间' (Delete Time), and '其他' (Other). A reference numeral 33D is located at the top right of the table.

物理区块结构信息
区块大小 (容量)
页面大小 (容量)
推定写入时间
删除时间
其他

图9

33E

虚拟区块结构信息
等价区块大小 (容量)
页面大小 (容量)
推定写入时间
删除时间
一个虚拟区块所含的区块数
其他

图10

33F

命名空间信息 (物理NAND存取管理API)
存在的命名空间总数
NSID#1的区块分配数 (用来确保NSID#1的区块总数)
分配给NSID#1的区块地址列表
NSID#1的合计删除次数
NSID#2的区块分配数 (用来确保NSID#2的区块总数)
分配给NSID#2的区块地址的列表
NSID#2的合计删除次数
其他

图11

33G

命名空间信息 (虚拟NAND存取管理API)
存在的命名空间总数
NSID#1的虚拟区块分配数 (用来确保NSID#1的虚拟区块总数)
分配给NSID#1的虚拟区块地址列表
NSID#1的合计删除次数
NSID#2的虚拟区块分配数 (用来确保NSID#2的虚拟区块总数)
分配给NSID#2的虚拟区块地址的列表
NSID#2的合计删除次数
其他

图12

虚拟区块 地址 VB	Ch.A 区块地址	Ch.B 区块地址	Ch.C 区块地址	Ch.D 区块地址	删除次数
VB0	区块 0 (VB0-0)	区块 100 (VB0-1)	区块 0 (VB0-2)	区块 100 (VB0-3)	100
VB1	区块 1 (VB1-0)	区块 99 (VB1-1)	区块 1 (VB1-2)	区块 99 (VB1-3)	102
VB2	区块 2 (VB2-0)	区块 98 (VB2-1)	区块 2 (VB2-2)	区块 98 (VB2-3)	103
VB3	区块 3 (VB3-0)	区块 97 (VB3-1)	区块 3 (VB3-2)	区块 97 (VB3-3)	101
VB4	区块 4 (VB4-0)	区块 96 (VB4-1)	区块 4 (VB4-2)	区块 96 (VB4-3)	95
∴	∴	∴	∴	∴	∴
VB100	区块 100 (VB100-0)	区块 0 (VB100-1)	区块 100 (VB100-2)	区块 0 (VB100-3)	100

图13

虚拟区块地址	Ch.A 区块地址	Ch.B 区块地址	Ch.C 区块地址	Ch.D 区块地址	删除次数
VB0	区块0 (VB0-0)	区块100 (VB0-1)	区块0 (VB0-2)	区块100 (VB0-3)	100
VB1	区块1 (VB1-0)	区块99 (VB1-1)	区块1 (VB1-2)	区块99 (VB1-3)	102
VB2	区块2 (VB2-0)	区块98 (VB2-1)	区块2→区块102 (VB2-2)	区块98 (VB2-3)	103
VB3	区块3 (VB3-0)	区块97 (VB3-1)	区块3 (VB3-2)	区块97 (VB3-3)	101
VB4	区块4 (VB4-0)	区块96 (VB4-1)	区块4 (VB4-2)	区块96 (VB4-3)	95
⋮	⋮	⋮	⋮	⋮	⋮
VB100	区块100 (VB100-0)	区块0 (VB100-1)	区块100 (VB100-2)	区块0 (VB100-3)	100

替换成与坏块相同信道的区块（例如区块102）

图14

坏块化命令（物理NAND存取管理API）

输入参数	说明
区块地址	应坏块化的区块的区块地址
输出参数	说明
结束状态	向主机返回是成功或错误

图15

坏块化命令（虚拟NAND存取管理API）

输入参数	说明
虚拟区块地址与虚拟区块内区块编号	指定哪个虚拟区块内的第几个区块是应坏块化的区块
输出参数	说明
结束状态	向主机返回是成功或错误

图16

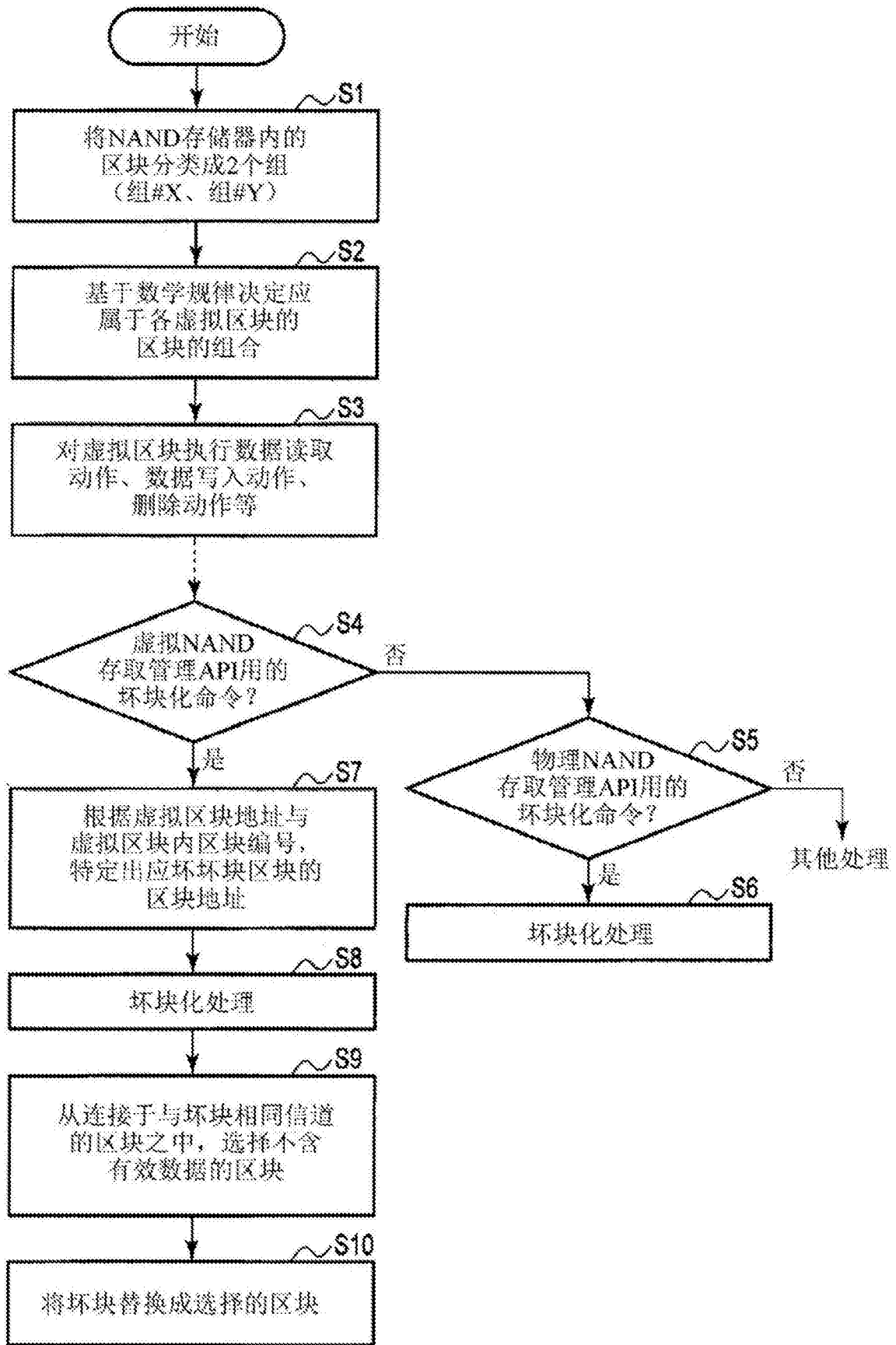


图17

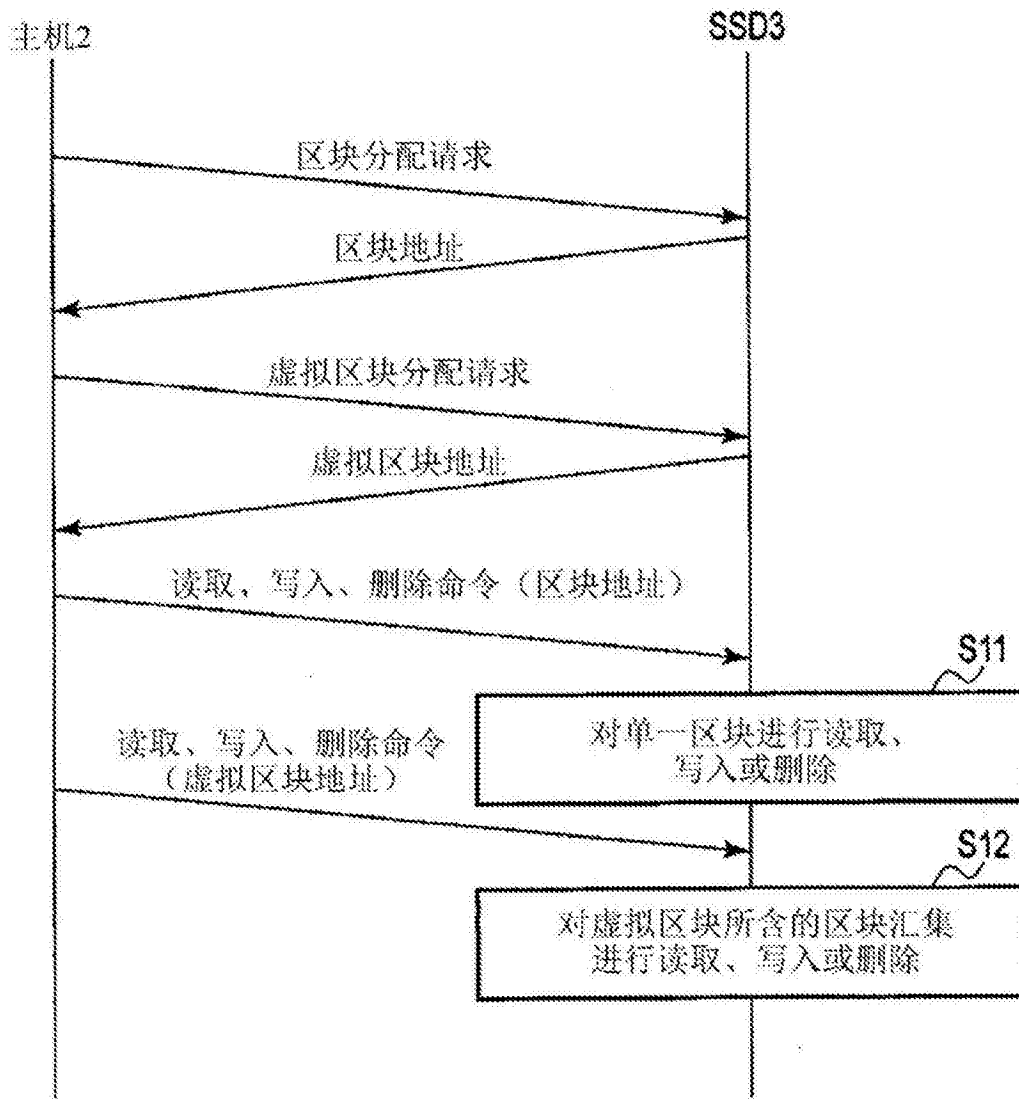


图18

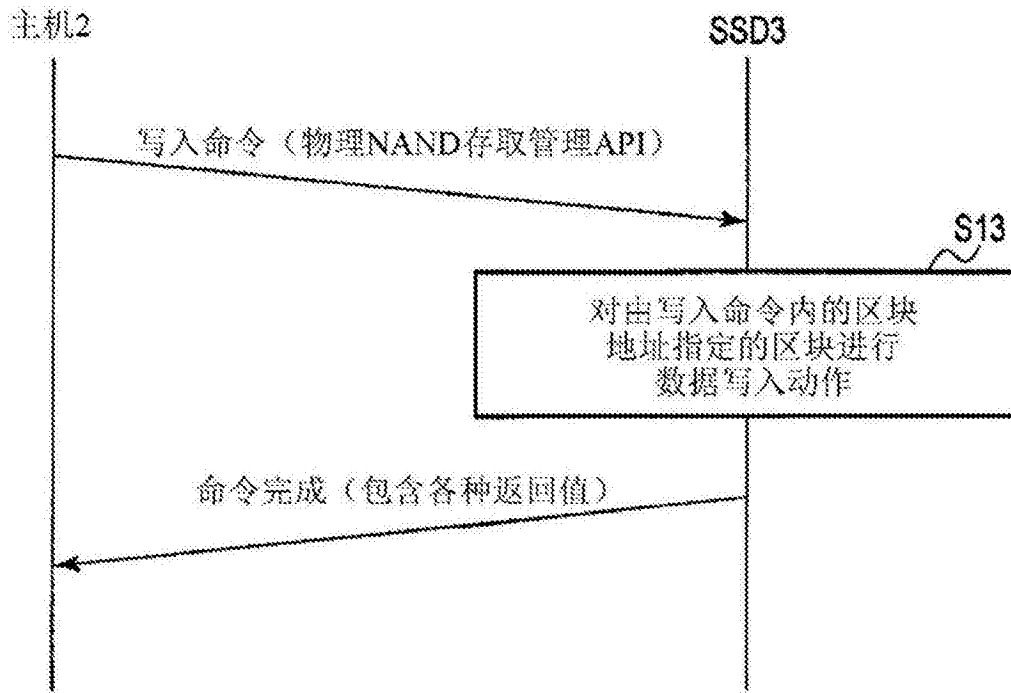


图19

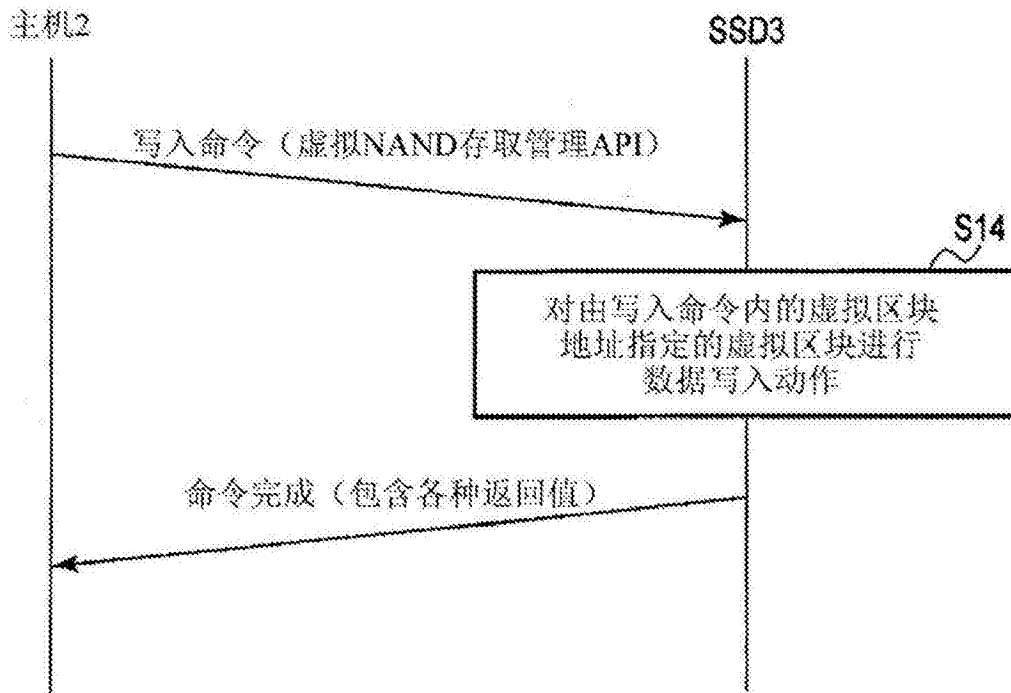


图20

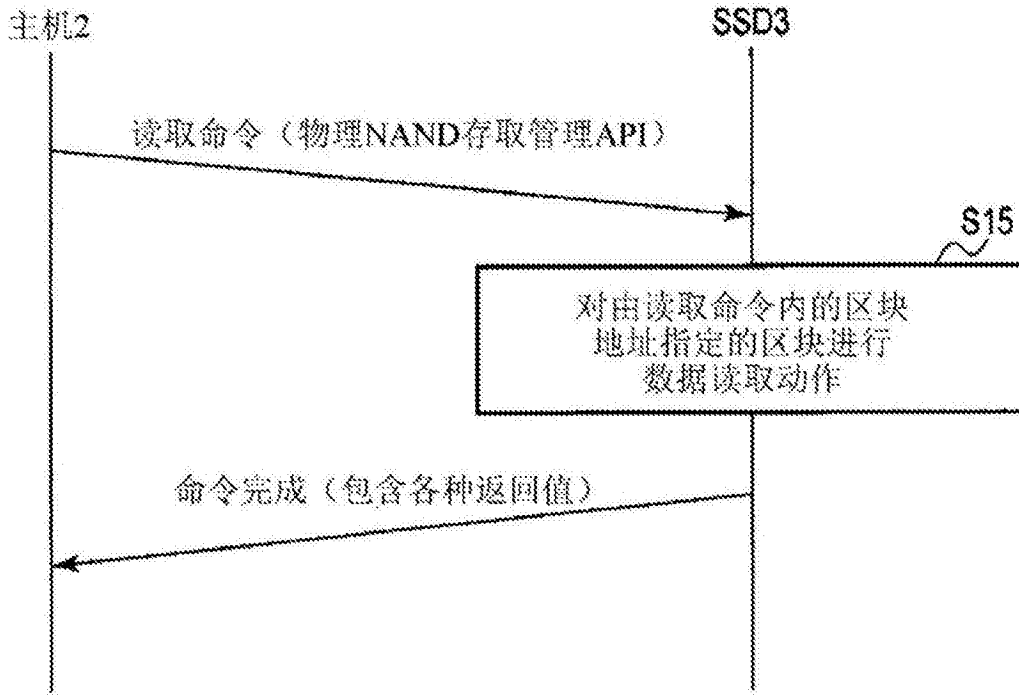


图21

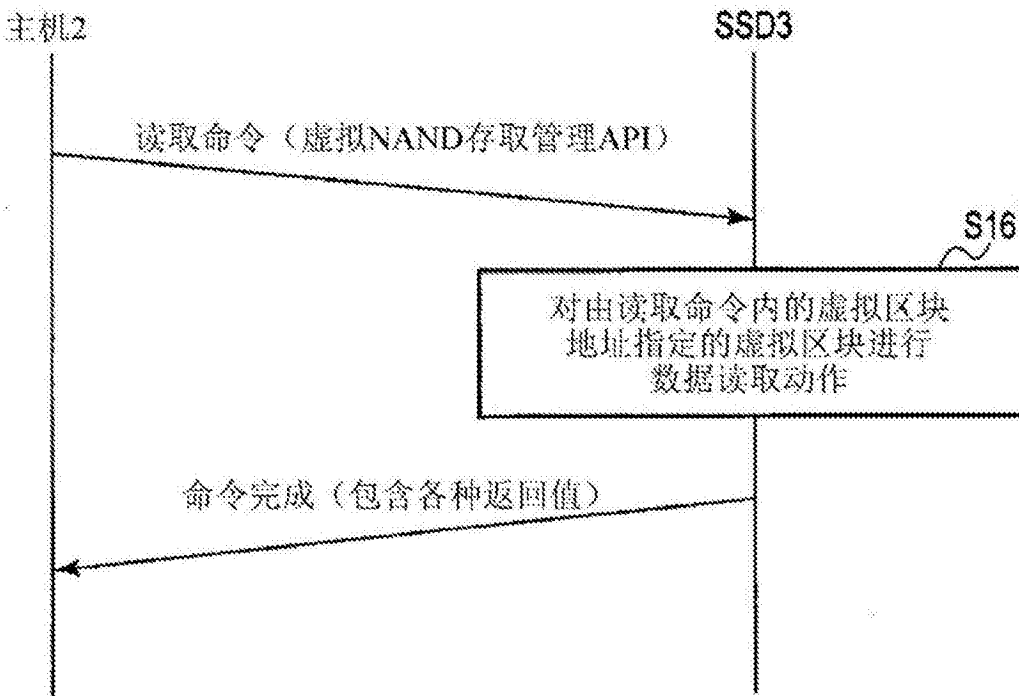


图22

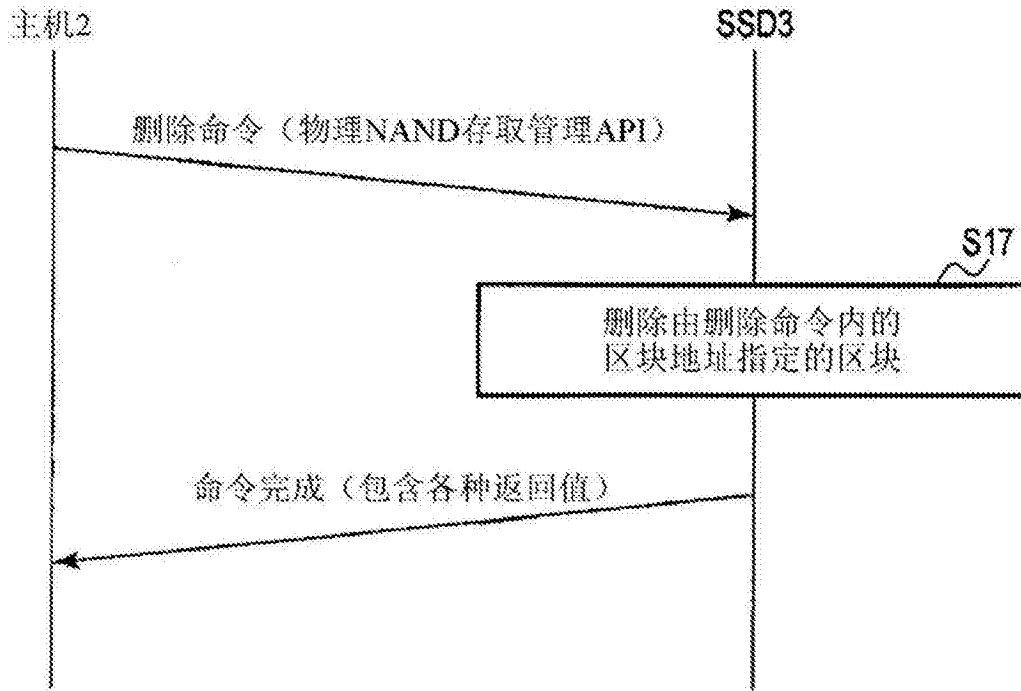


图23

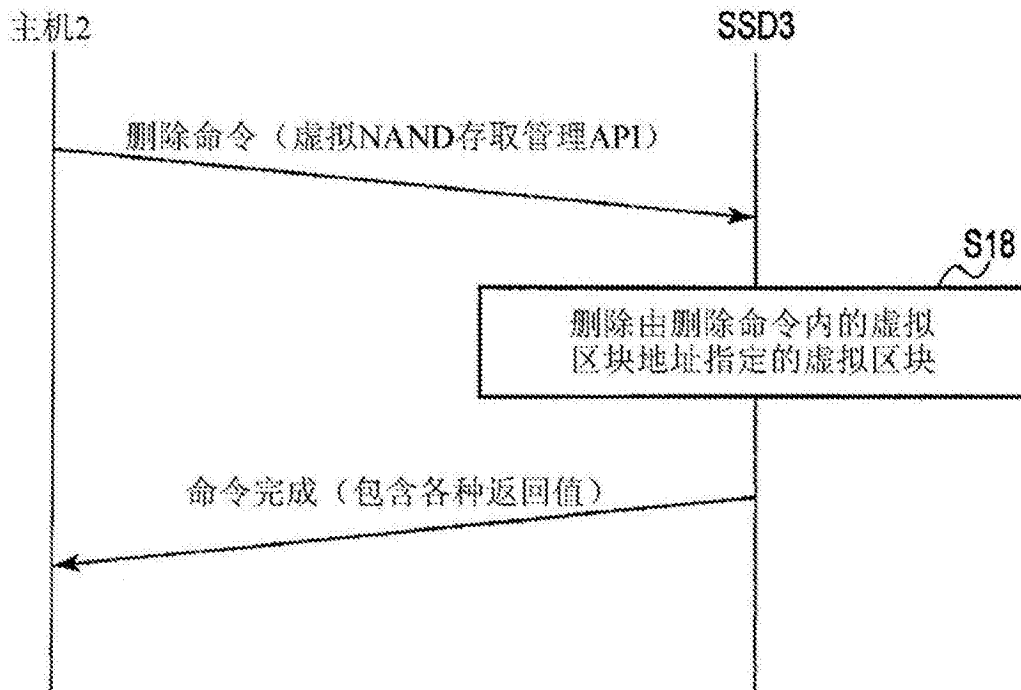


图24

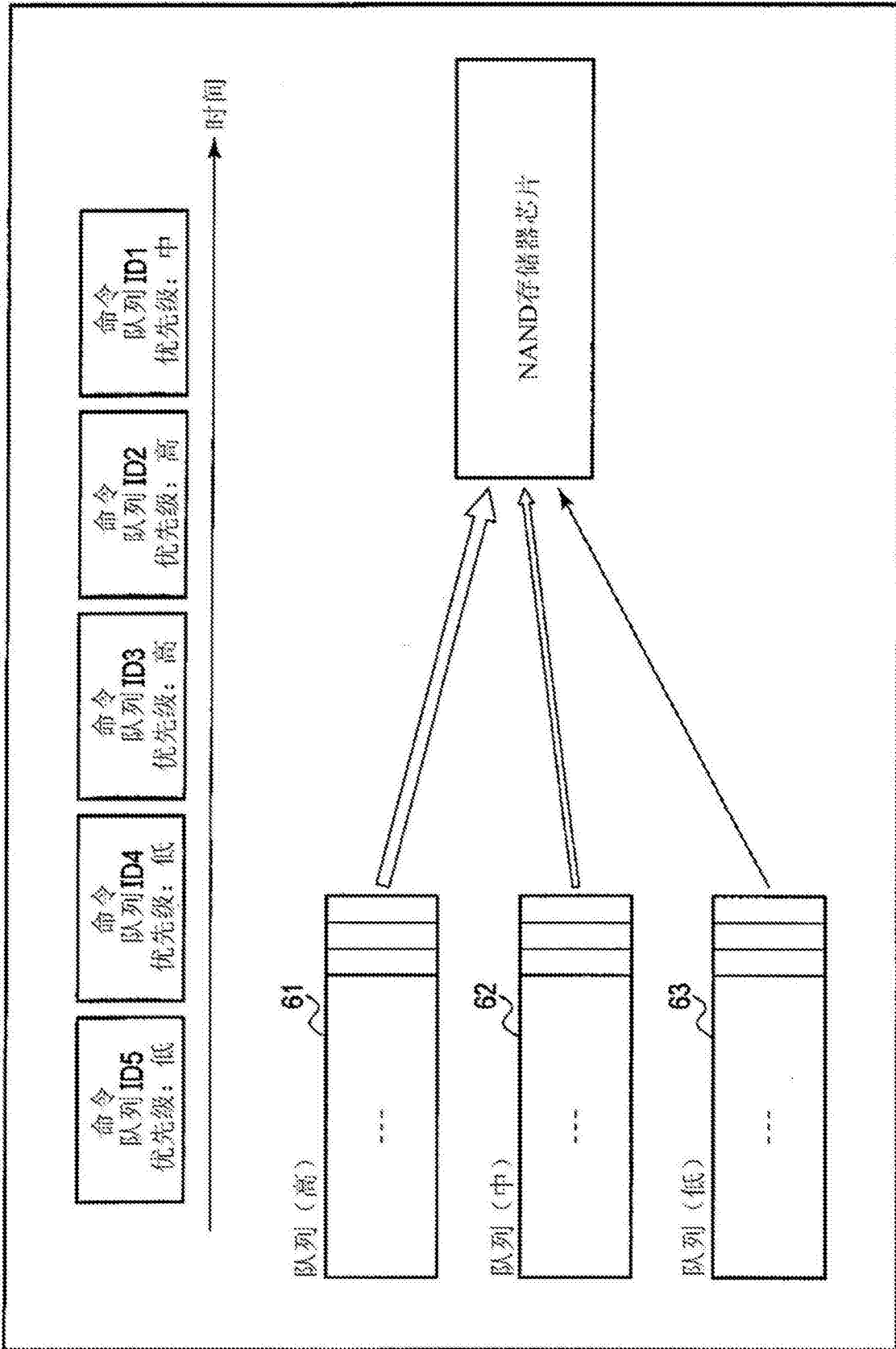


图25

区块分配、删除命令

输入参数	说明
区块类型 (区块或虚拟区块)	应分配区块类型 区块：从空闲区块分配 一个区块并删除 虚拟区块：从空闲虚拟区块 分配一个虚拟区块 (区块组)并删除
处理优先级	该命令的优先级
NSID (选择)	应分配区块/虚拟区块的NSID
输出参数	说明
结束状态	向主机返回是成功或错误
区块地址或 虚拟区块地址	向主机返回分配的区块的 区块地址或分配的虚拟区块 的虚拟区块地址
剩余区块数	只在指定了NSID的情况下， 向主机返回NSID用剩余区块数， 或NSID用剩余虚拟区块数

图26

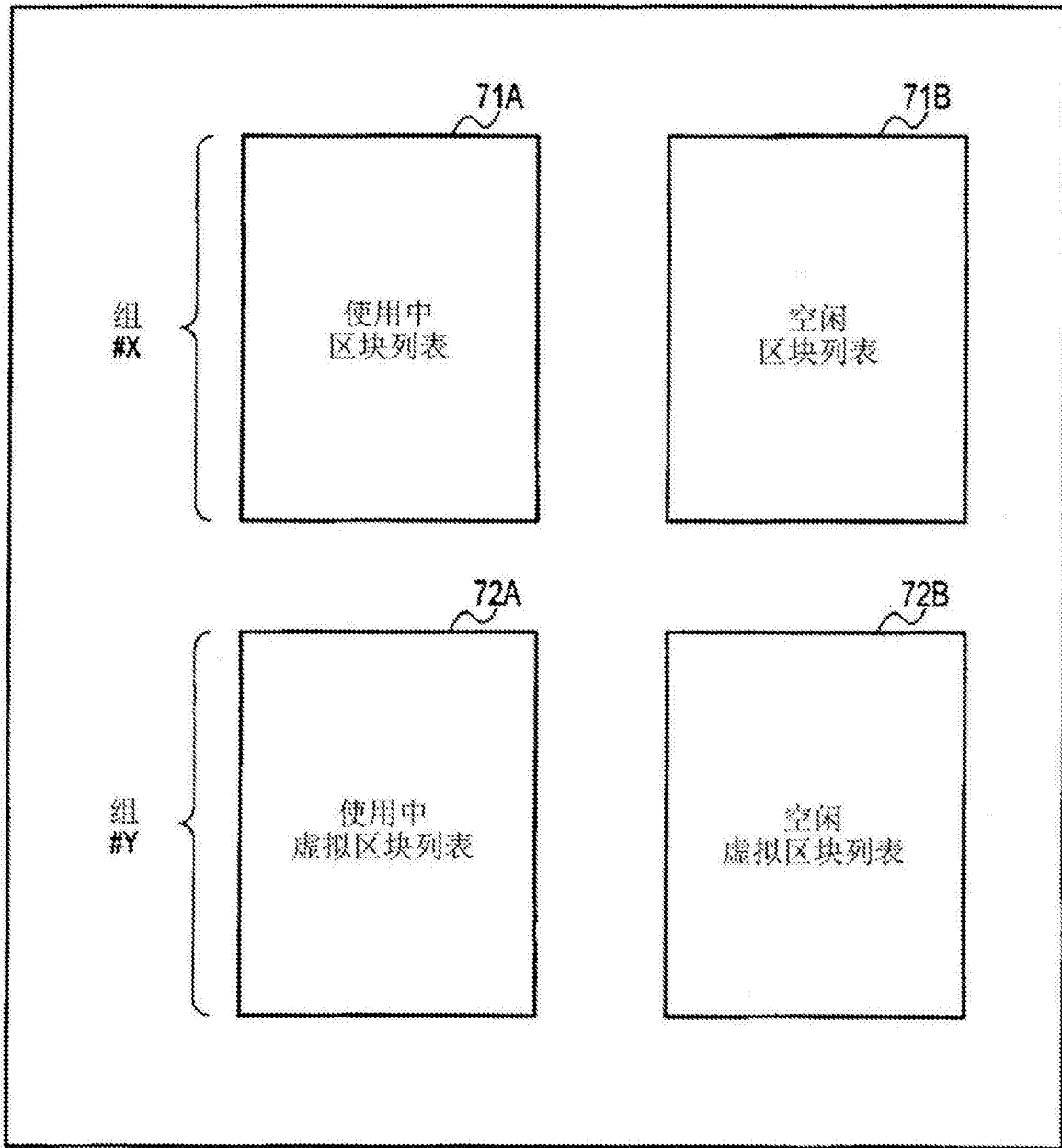


图27

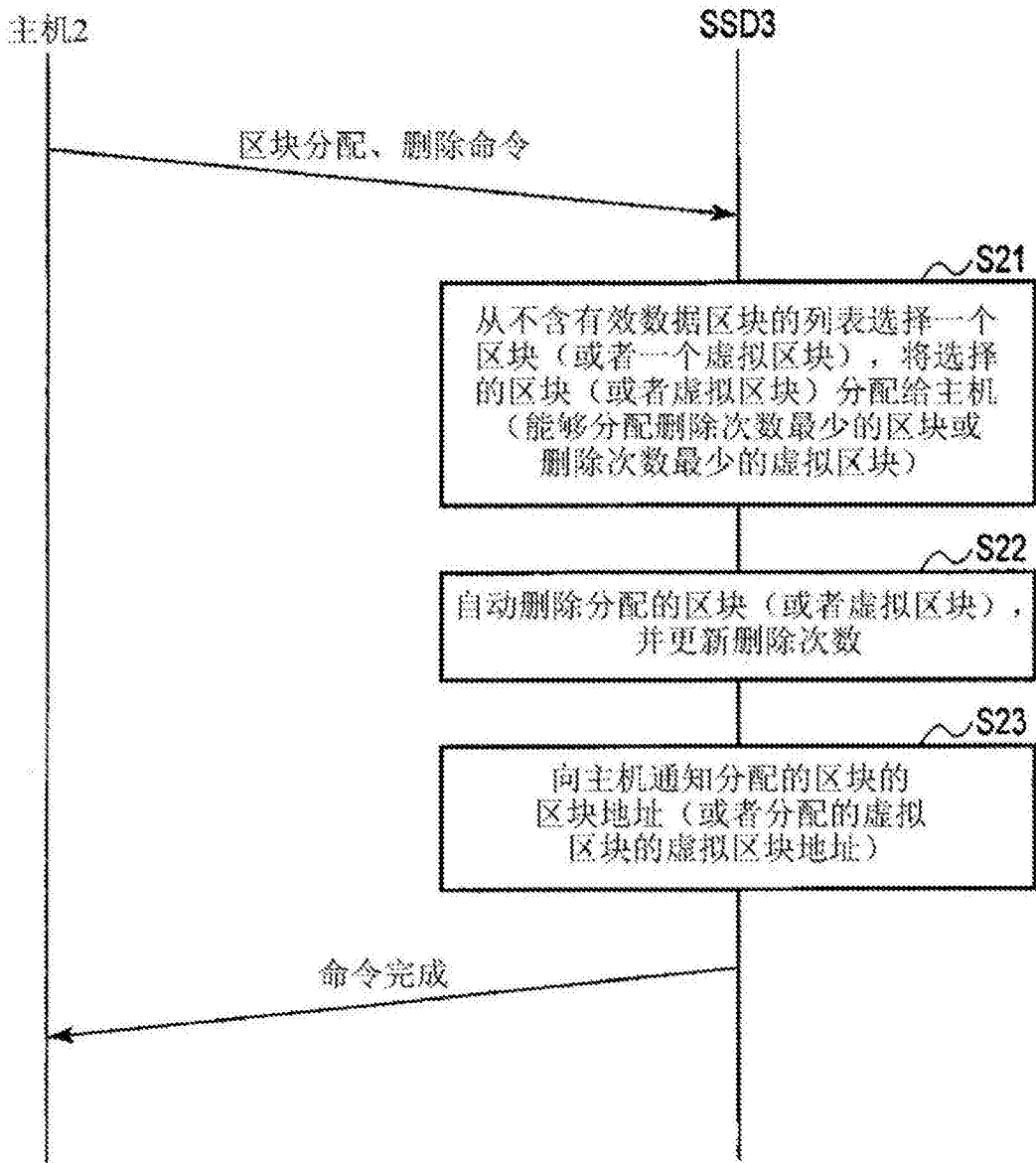


图28

写入命令（物理NAND存取管理API）

输入参数	说明
区块地址或 区块地址与页面地址	指定应写入数据的NAND存储器上的位置的物理地址（在只指定区块地址时，写入对象页面由SSD自动更新。）
NSID（可选）	当不指定区块时指定NSID。当指定NSID时，区块地址与页面地址由SSD自动发布。数据被写入分配了NSID的区块。
处理优先级	该写入命令的优先级
写入数据起始地址	储存写入数据的输出缓冲器（主机的存储器）上的起始地址
写入页面数	应写入数据的页面数
NAND模式（可选）	NAND模式（SLC/MLC/TLC等）
输出参数	说明
结束状态	向主机返回成功或错误
区块地址与页面地址	写入数据的NAND存储器上的位置
写入的页面数	写入数据的页面数
不当写入顺序警告	当检测出不当写入顺序时，向主机返回警告或错误
可读取的最新页面地址	保持可读出数据的最新页面地址

图29

写入命令（虚拟NAND存取管理API）

输入参数	说明
虚拟区块地址或 虚拟区块地址与页面地址	指定应写入数据的NAND存储器上的位置的物理地址（在只指定虚拟区块地址时，写入对象页面由SSD自动更新。）
NSID（可选）	当不指定虚拟区块时指定NSID。当指定NSID时，虚拟区块地址与页面地址由SSD自动发布。数据被写入分配了NSID的虚拟区块。
处理优先级	该写入命令的优先级
写入数据起始地址	储存写入数据的输出缓冲器（主机的存储器）上的位置
写入页面数	应写入数据的页面数
NAND模式（可选）	NAND模式（SLC/MLC/TLC等）
输出参数	说明
结束状态	向主机返回成功或错误
虚拟区块地址与页面地址	写入数据的NAND存储器上的位置
写入的页面数	写入数据的页面数
不当写入顺序警告	当检测出不当写入顺序时，向主机返回警告或错误
可读取的最新页面地址	保持可读出数据的最新页面地址

图30

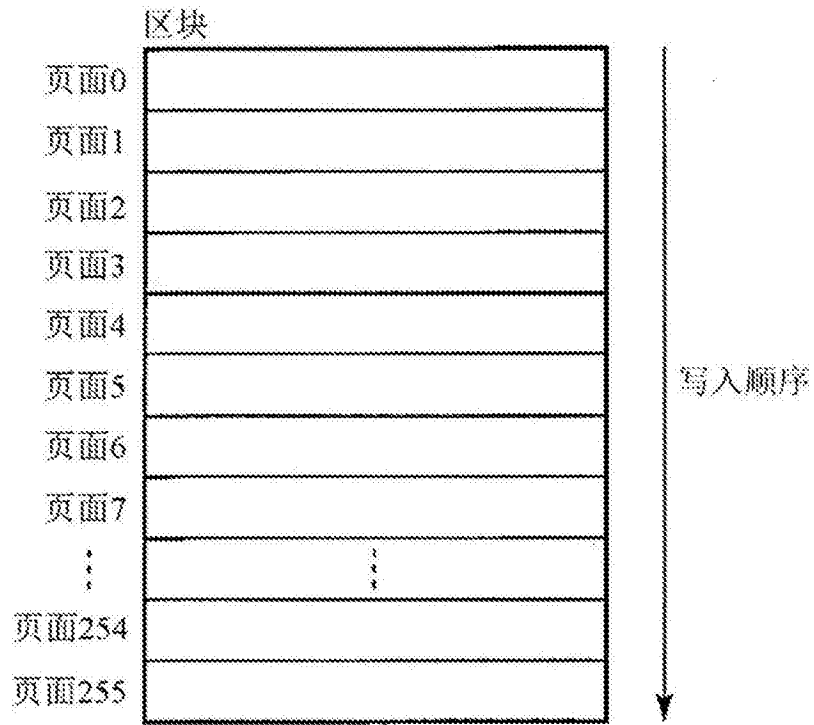


图31

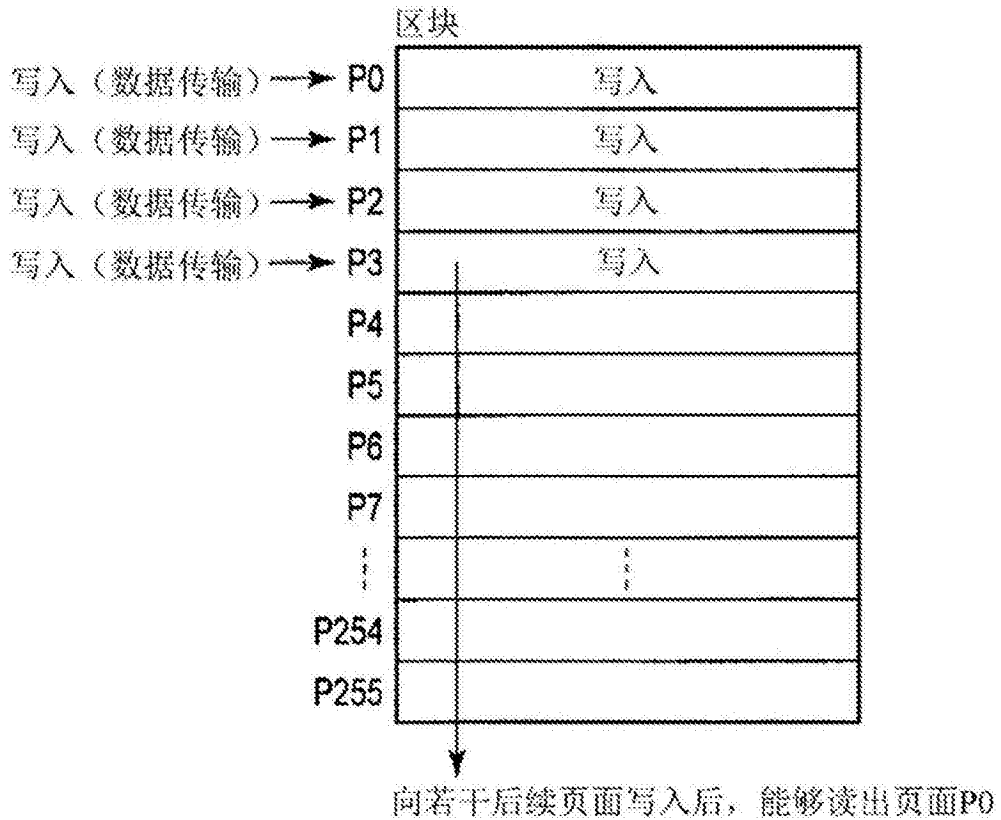


图32

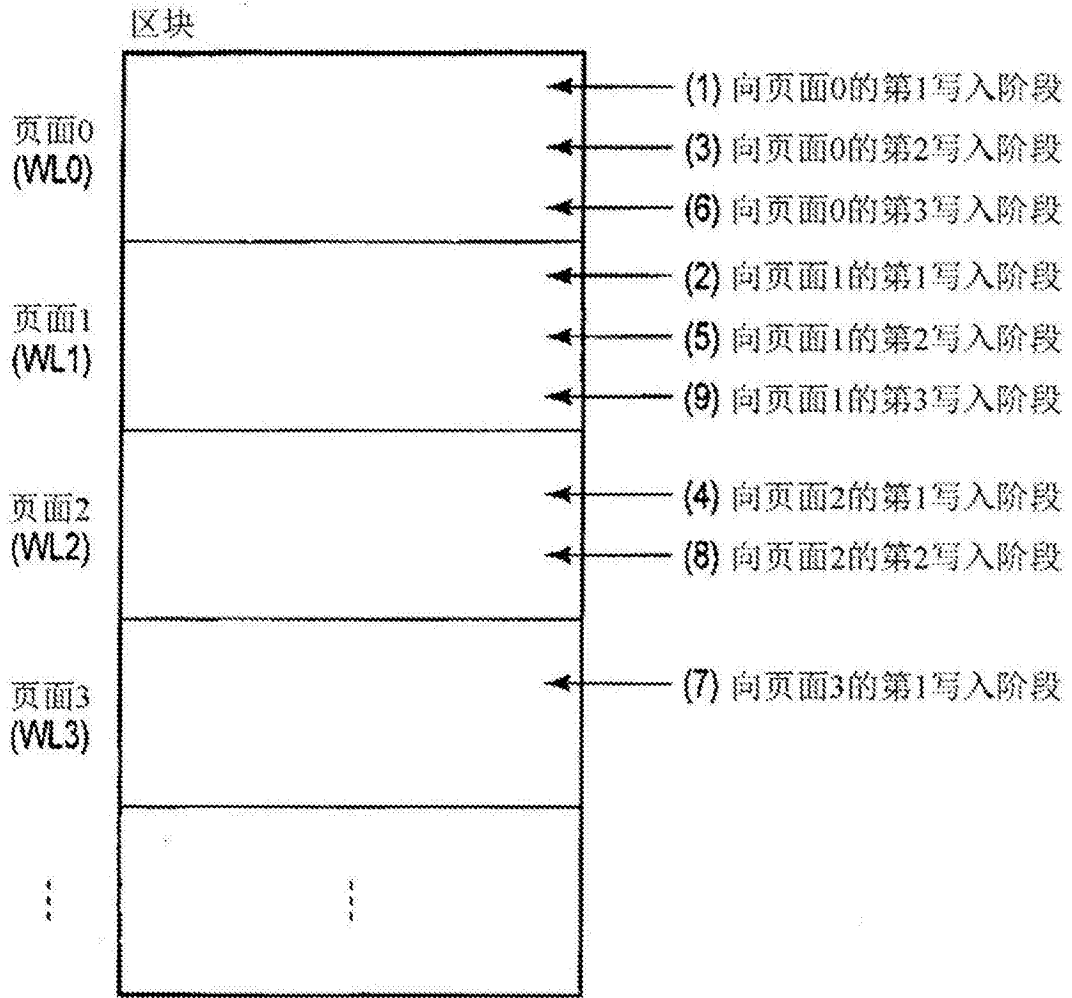


图33

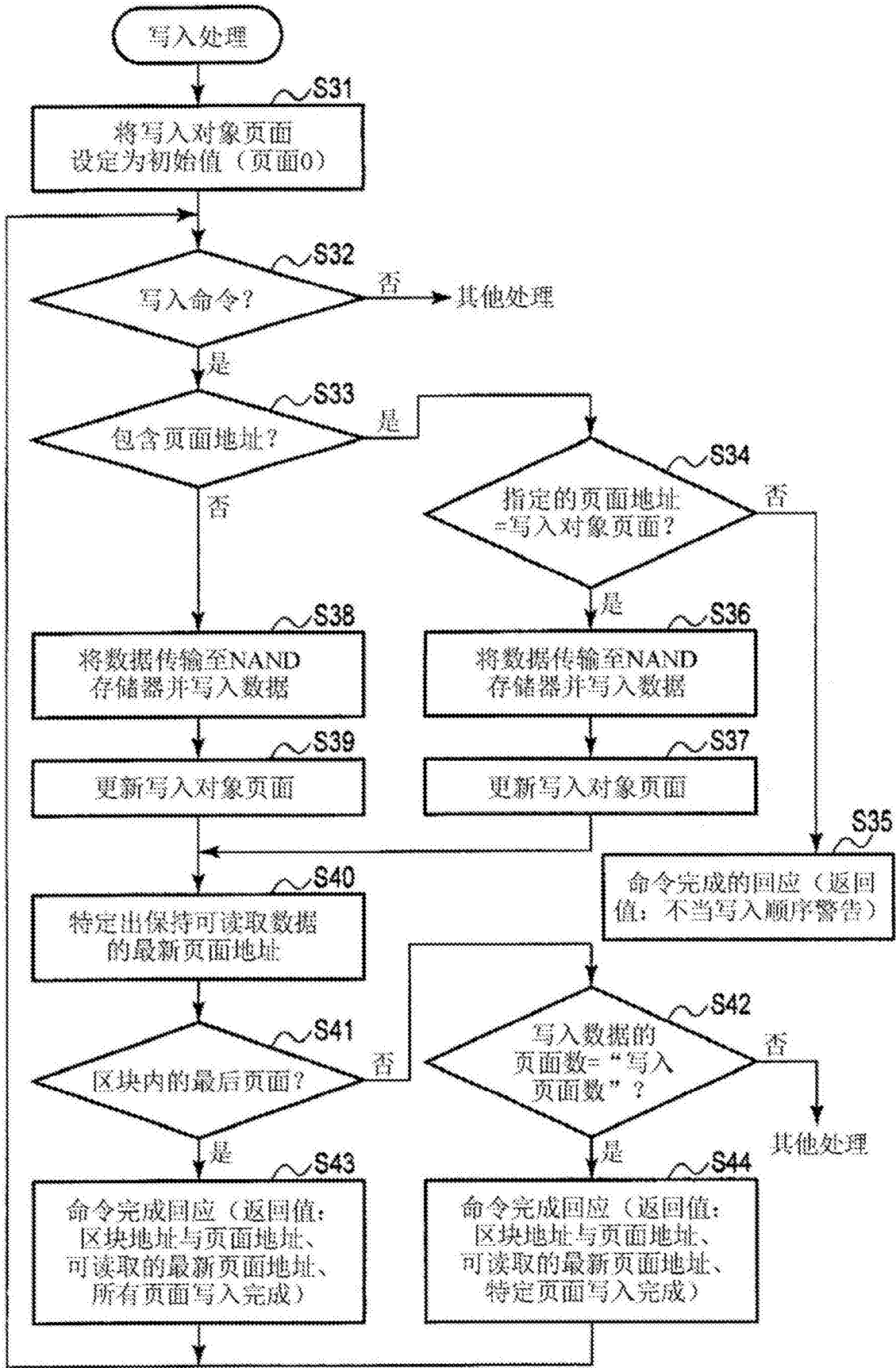


图34

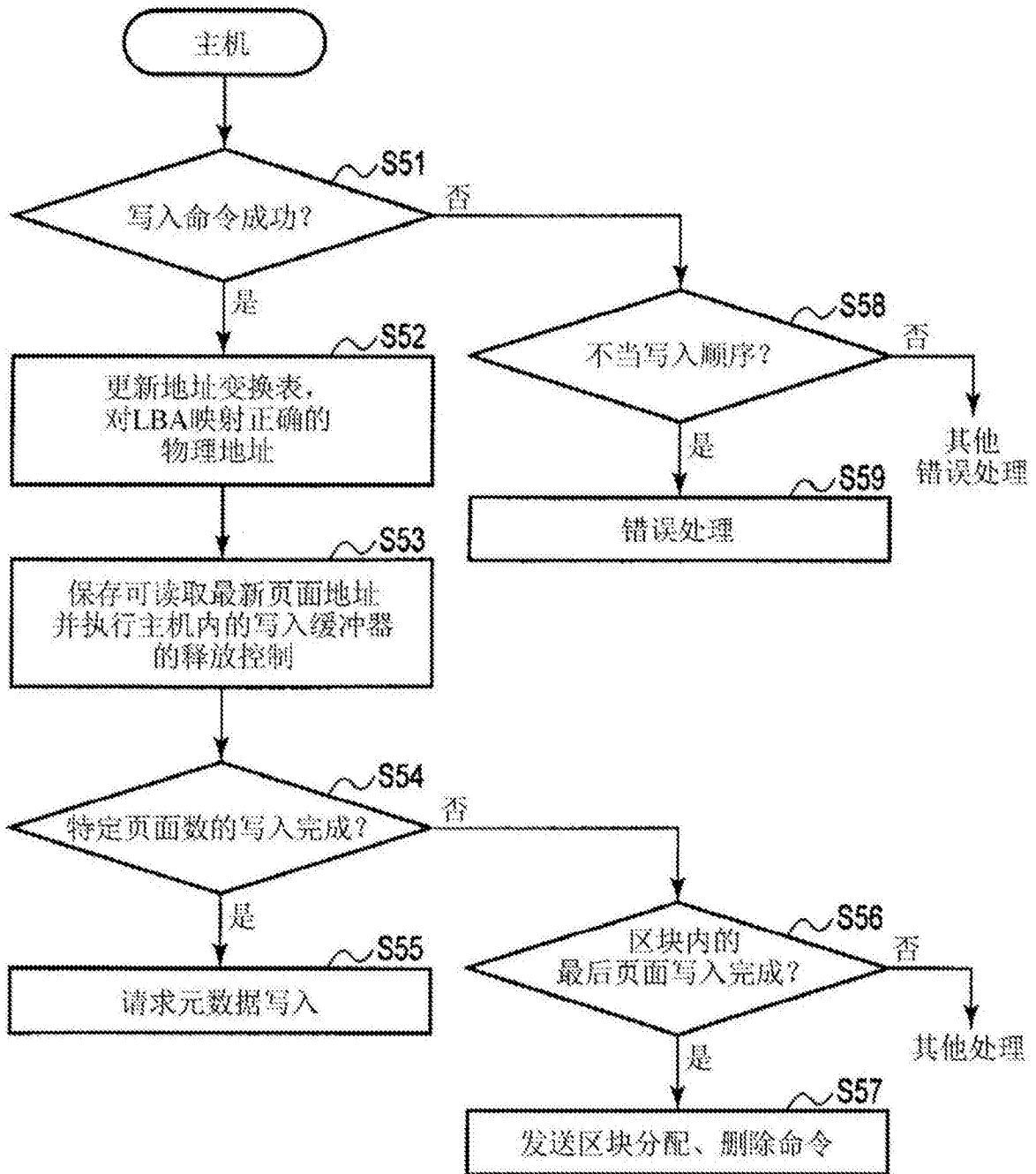


图35

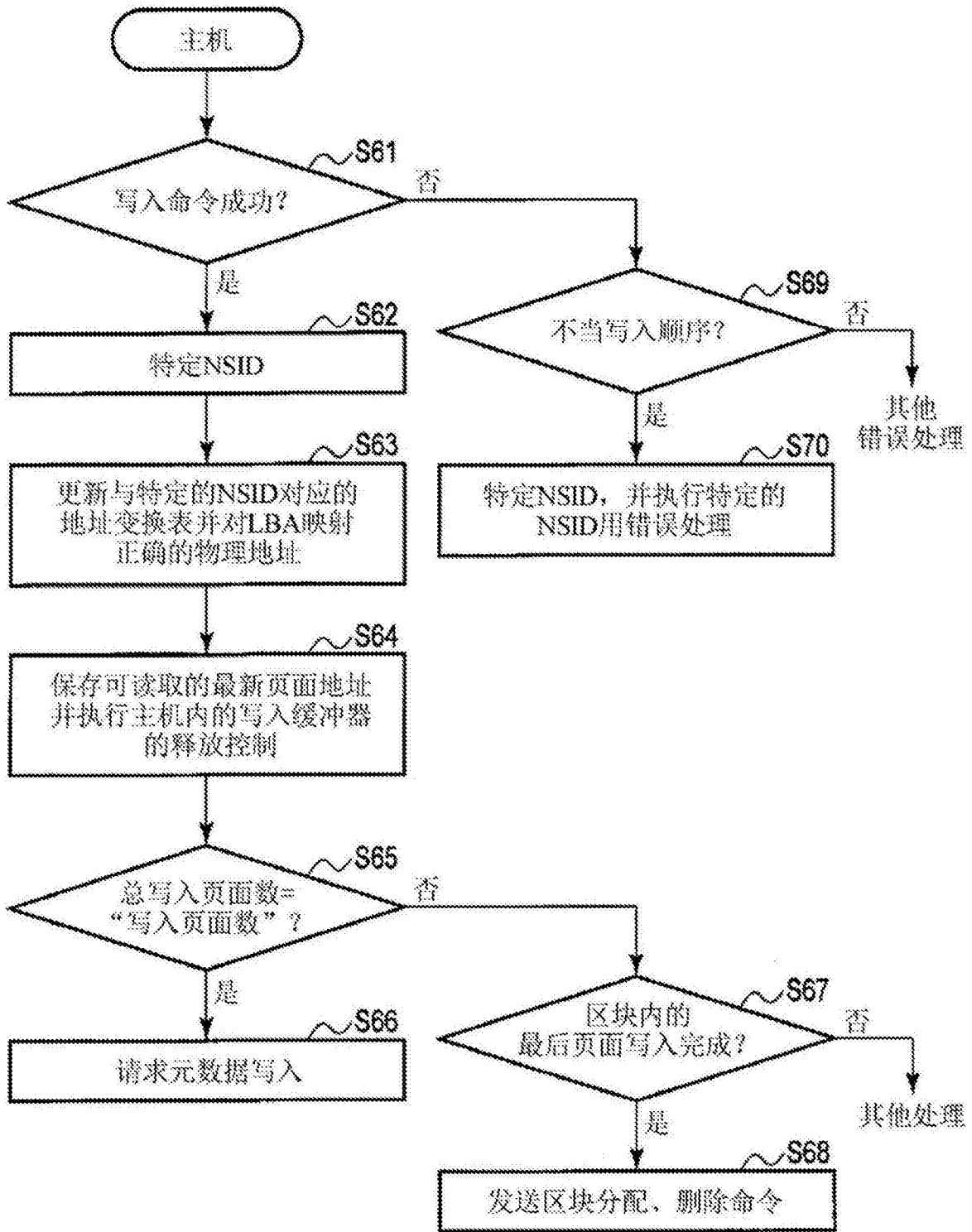


图36

读取命令（物理NAND存取管理API）

输入参数	说明
区块地址与页面地址	指定应读取数据的NAND存储器上的位置的物理地址
处理优先级	该读取命令的优先级
读出数据的传输目标地址	应传输读出数据的输入缓冲器（主机的存储器）上的位置
读出页面数	应读出的页面数
允许等待时间	最少等待时间/正常等待时间/长等待时间
输出参数	说明
结束状态	向主机返回成功或错误
区块地址与页面地址	读出数据的NAND存储器上的位置（区块地址与页面地址）
页面数	读出数据的页面数
起始地址	读出数据的起始地址

图37

读取命令（虚拟NAND存取管理API）

输入参数	说明
虚拟区块地址与页面地址	指定应读取数据的NAND存储器上的位置的物理地址
处理优先级	该读取命令的优先级
读出数据的传输目标地址	应传输读出数据的输入缓冲器（主机的存储器）上的位置
读出页面数	应读出的页面数
允许等待时间	最少等待时间/正常等待时间/长等待时间
输出参数	说明
结束状态	向主机返回成功或错误
虚拟区块地址与页面地址	读出数据的NAND存储器上的位置（虚拟区块地址与页面地址）
页面数	读出数据的页面数
起始地址	读出数据的起始地址

图38

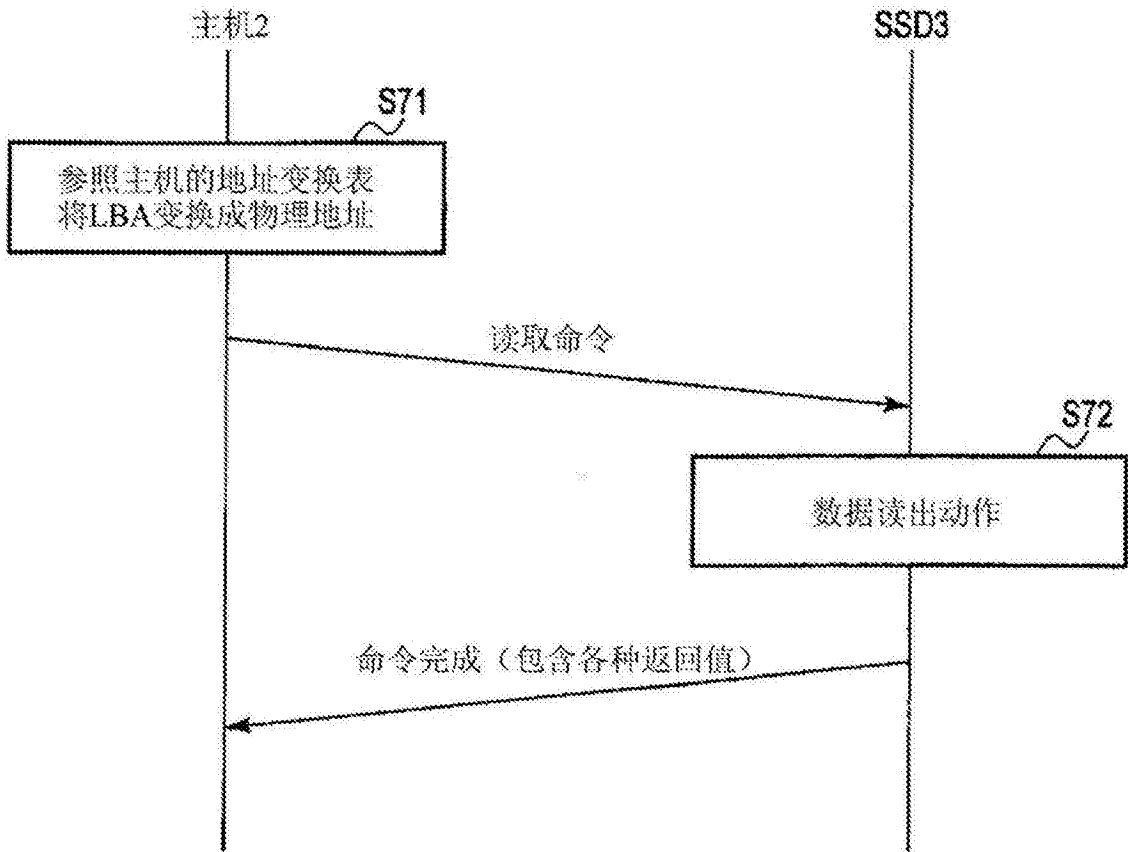


图39

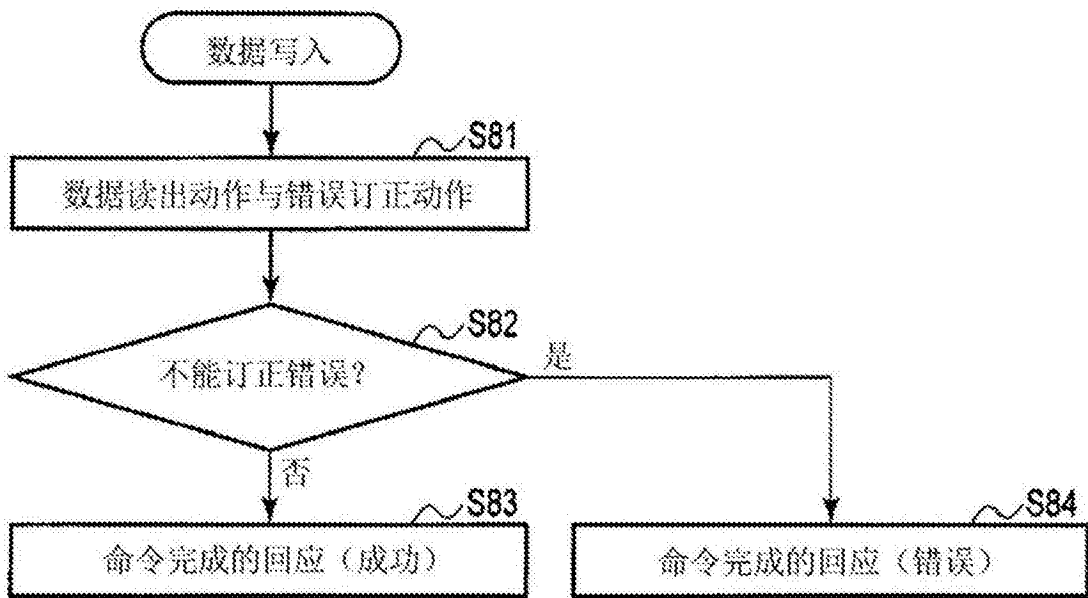


图40

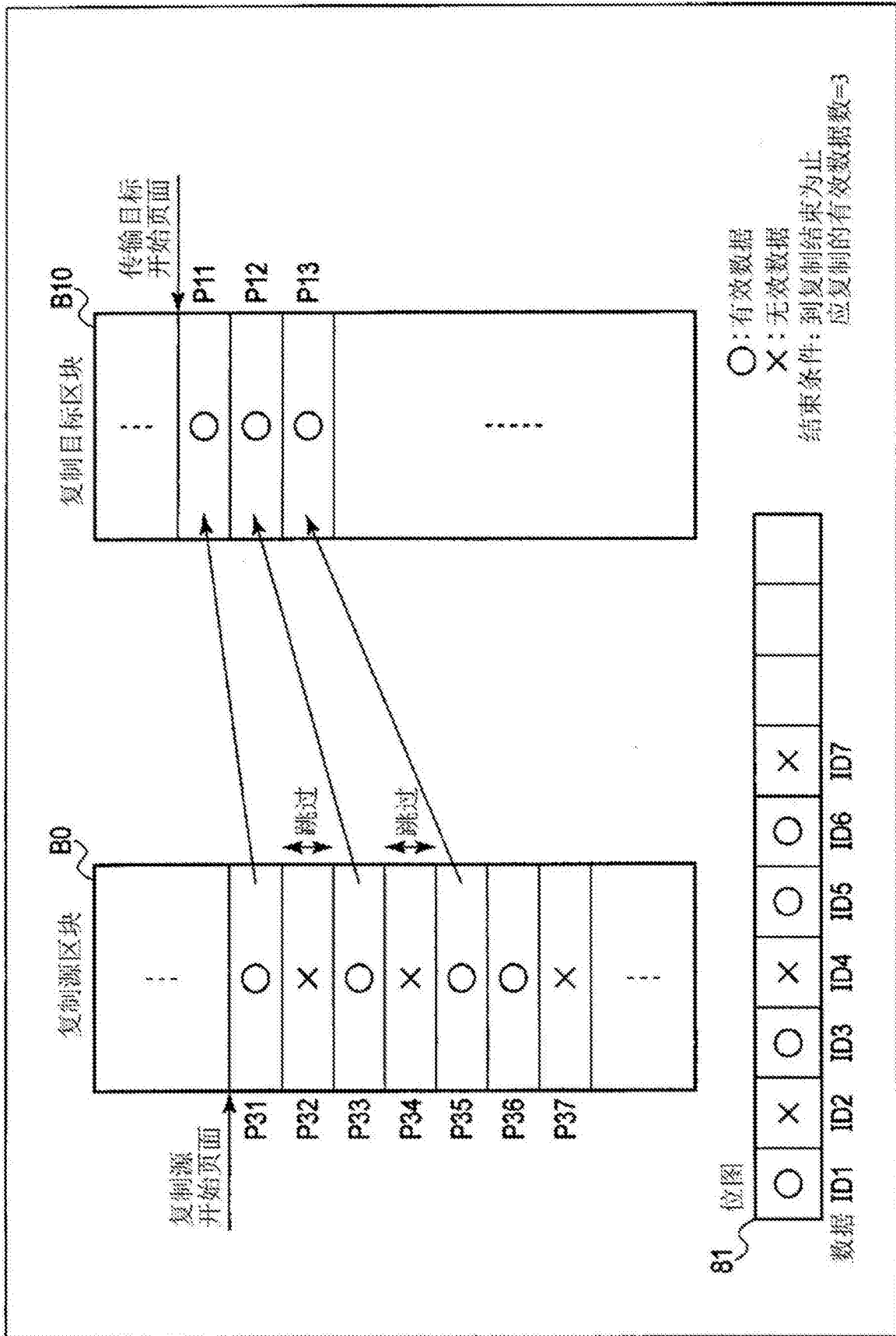


图41

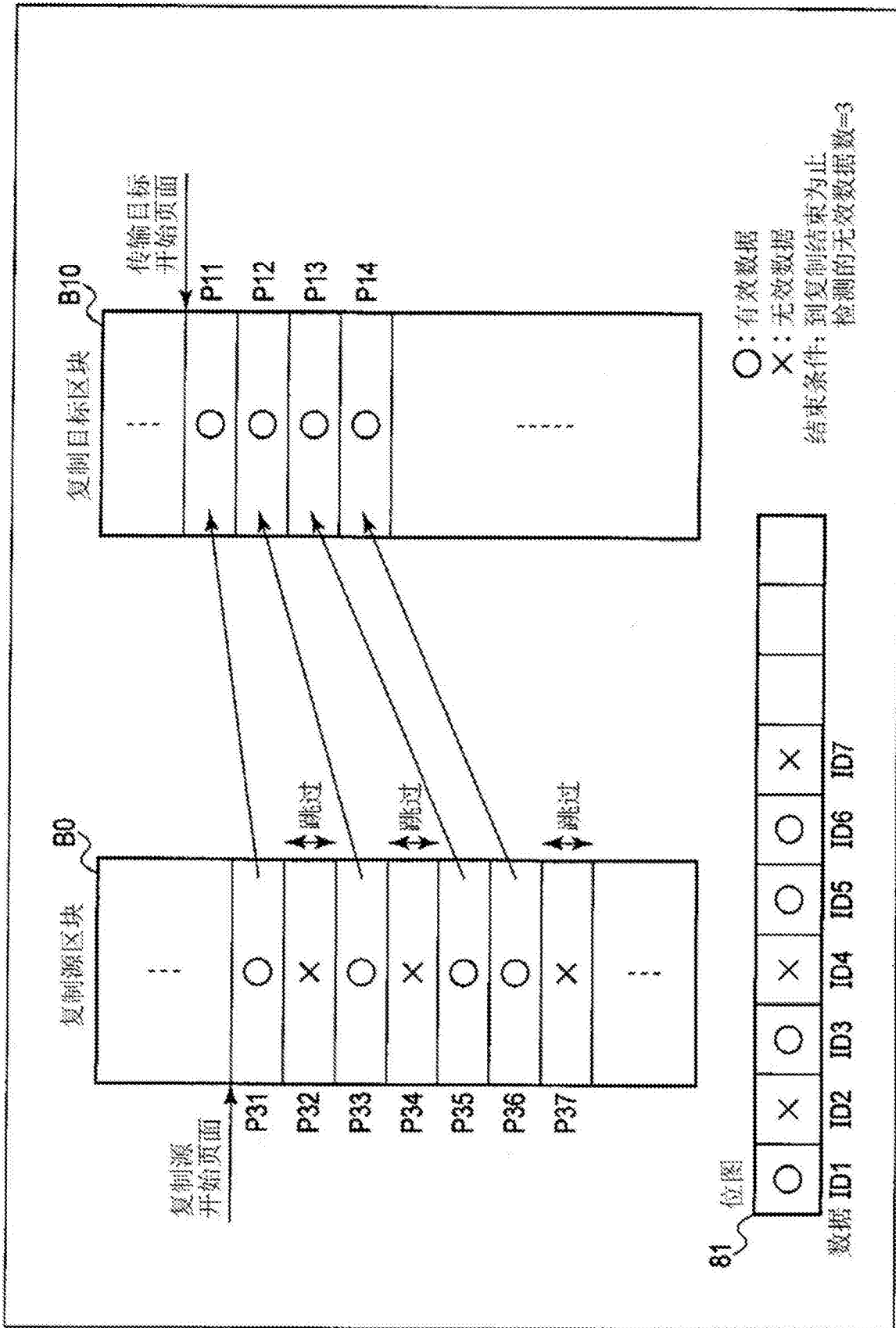


图42

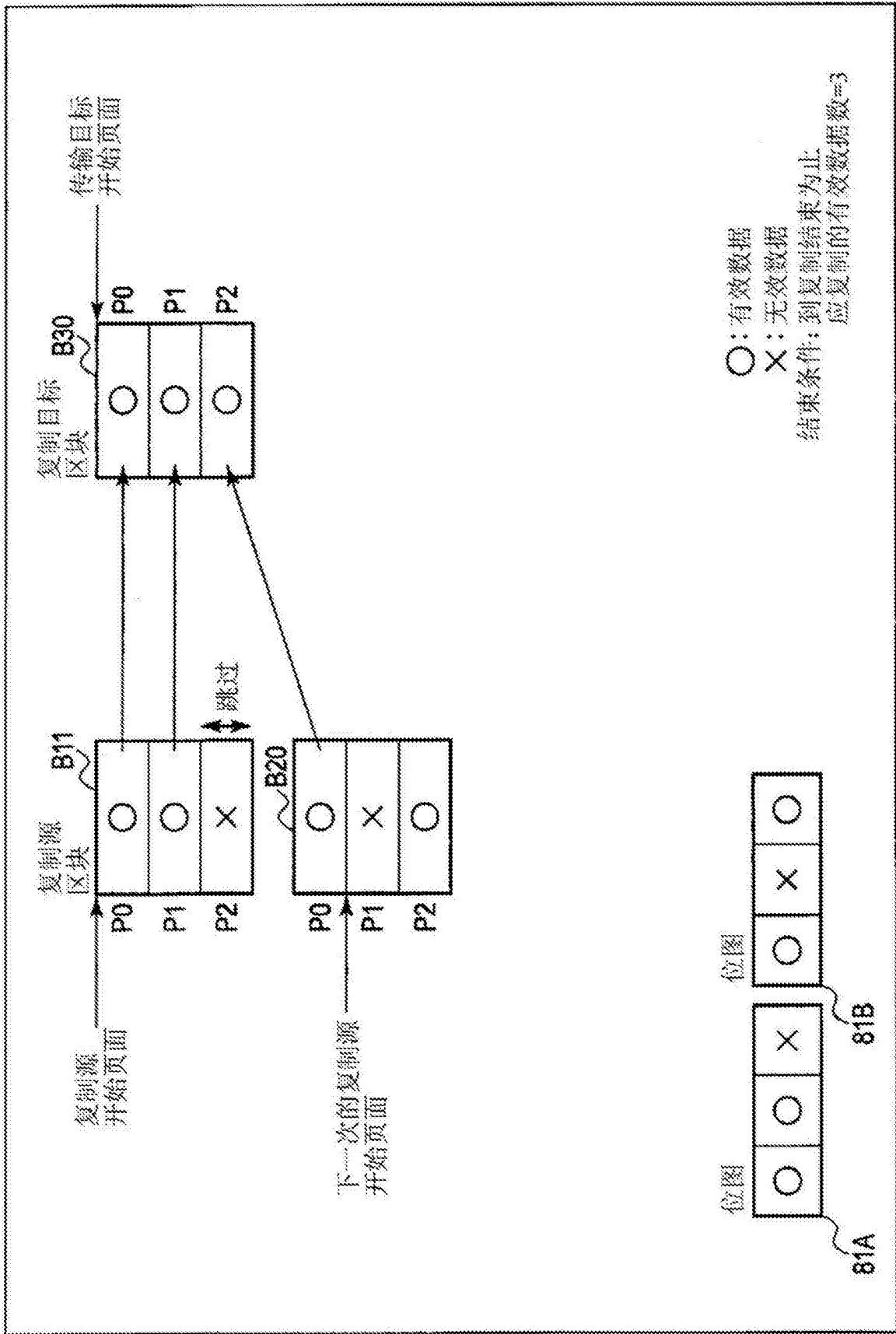


图43

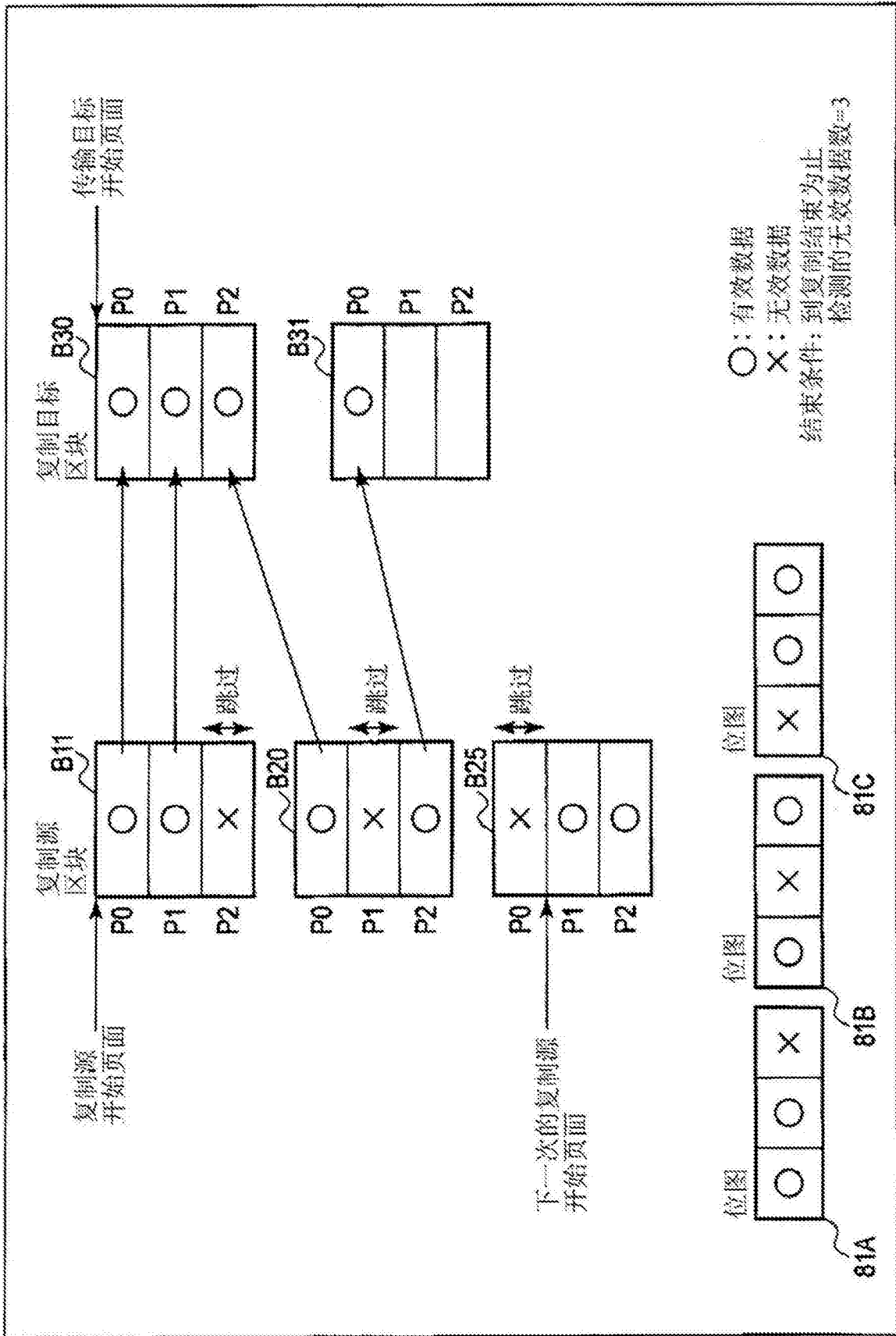


图44

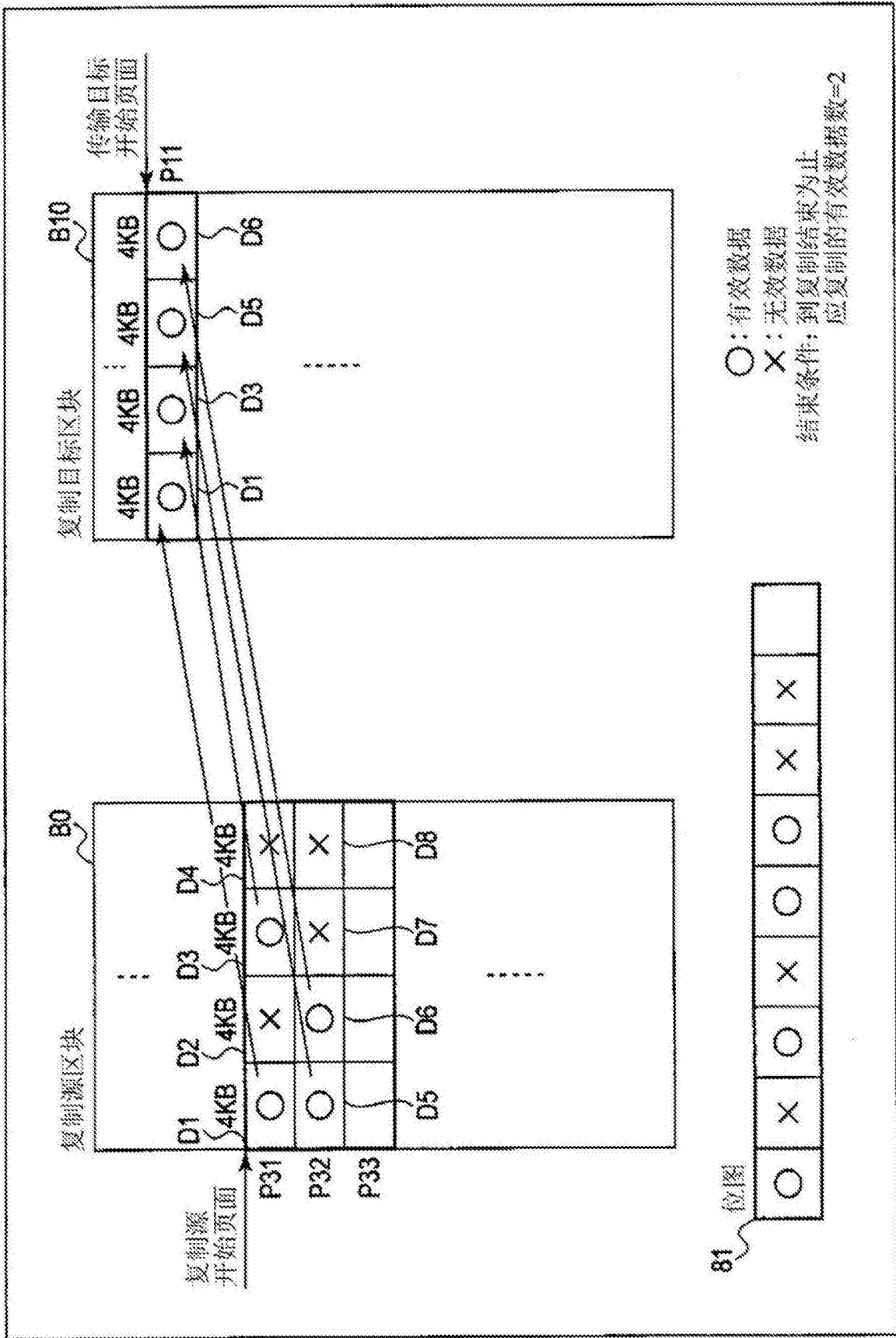


图45

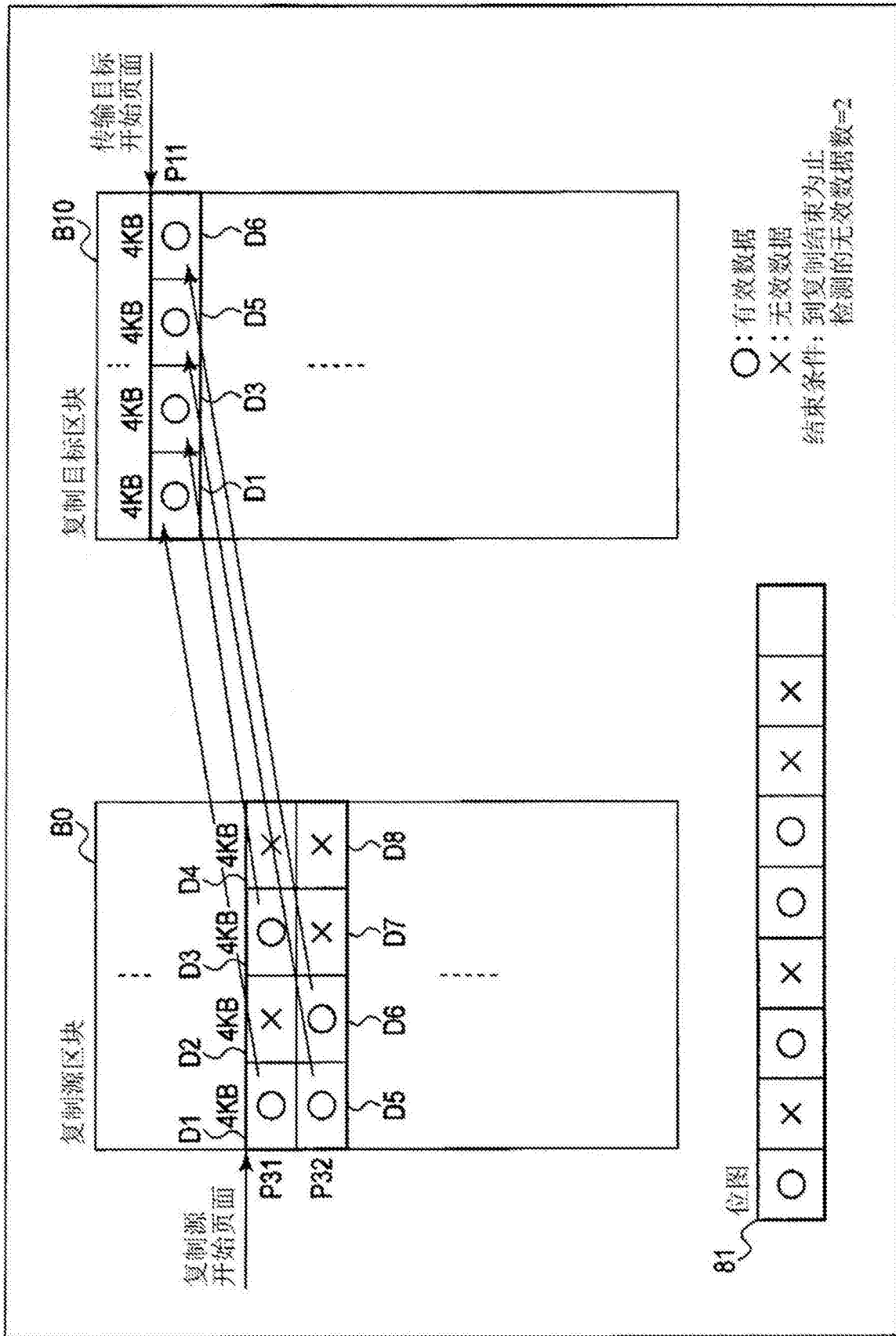


图46

数据复制命令

输入参数	说明
复制源区块地址 (复制源区块地址列表)	复制源区块的区块地址 (复制源区块的区块地址 也能利用虚拟区块地址与 虚拟区块内区块编号来指定。 复制源区块的区块地址列表 包含多个复制源区块各自的 区块地址)
复制源区块内的复制开始位置	应复制数据的复制源区块 内的开始位置(例如, 第1个复制源区块内的复制 开始位置)
复制目标区块地址	复制目标区块的区块地址 (复制目标区块的区块地址 也能利用虚拟区块地址与 虚拟区块内区块编号来指定)
复制目标区块内的开始位置	应传输数据的复制目标区块内的 开始位置
有效/无效位图	表示有效数据与无效数据的配置的信息
到复制结束为止应复制的 有效数据数	结束条件
到复制结束为止应检测的 无效数据数	结束条件
数据大小	

图47

数据复制命令

输出参数	说明
到复制结束为止检测的无效数据数	
复制到复制目标区块的数据总数	
复制到复制目标区块的数据的标识符与复制目标的记录位置的组	表示哪个数据复制到何处的信息
下一应开始复制的数据的位置	下一次复制中复制源区块内的复制开始位置
可读出的最新页面地址	复制目标区块内保持可读出数据的最新页面地址

图48

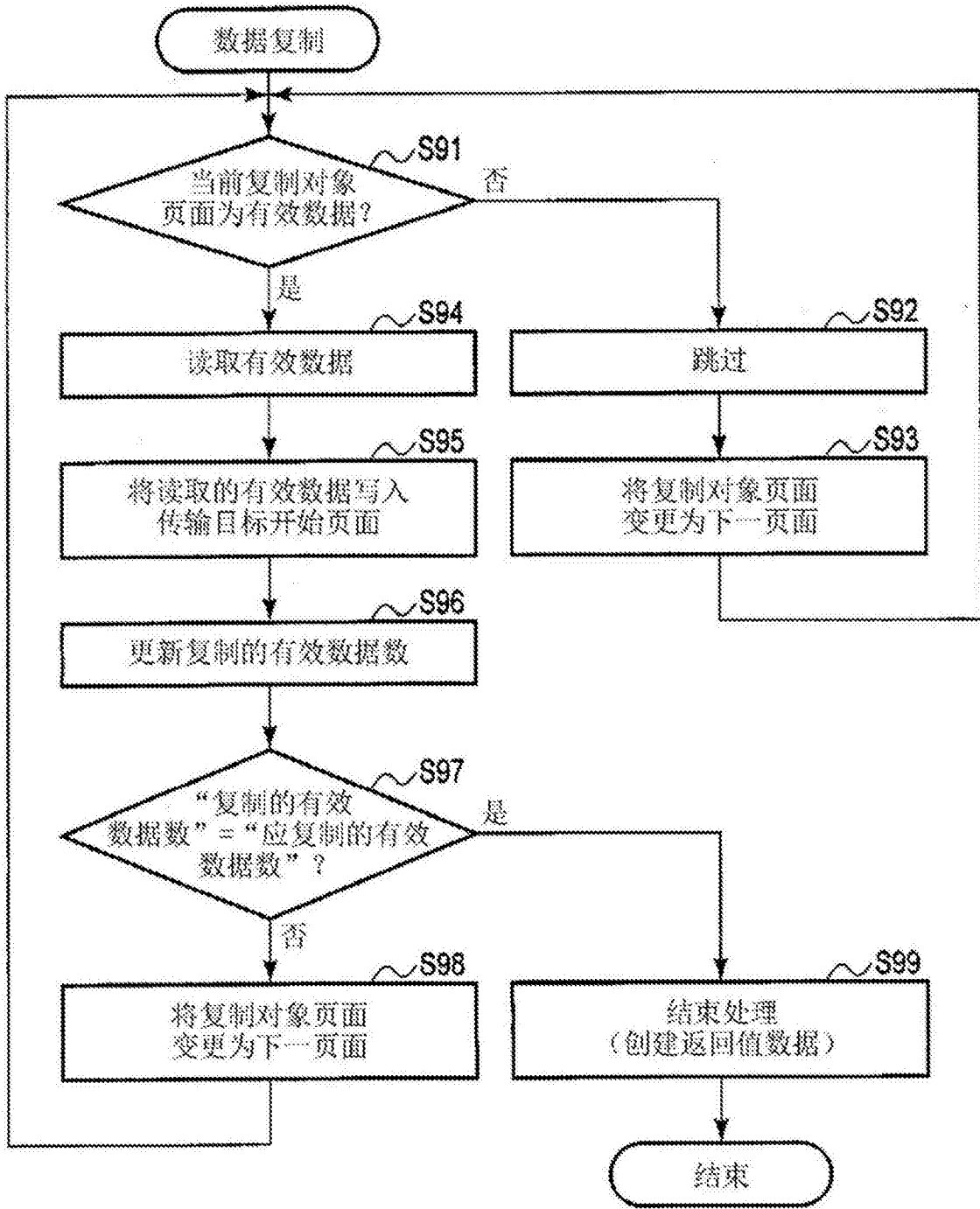


图49

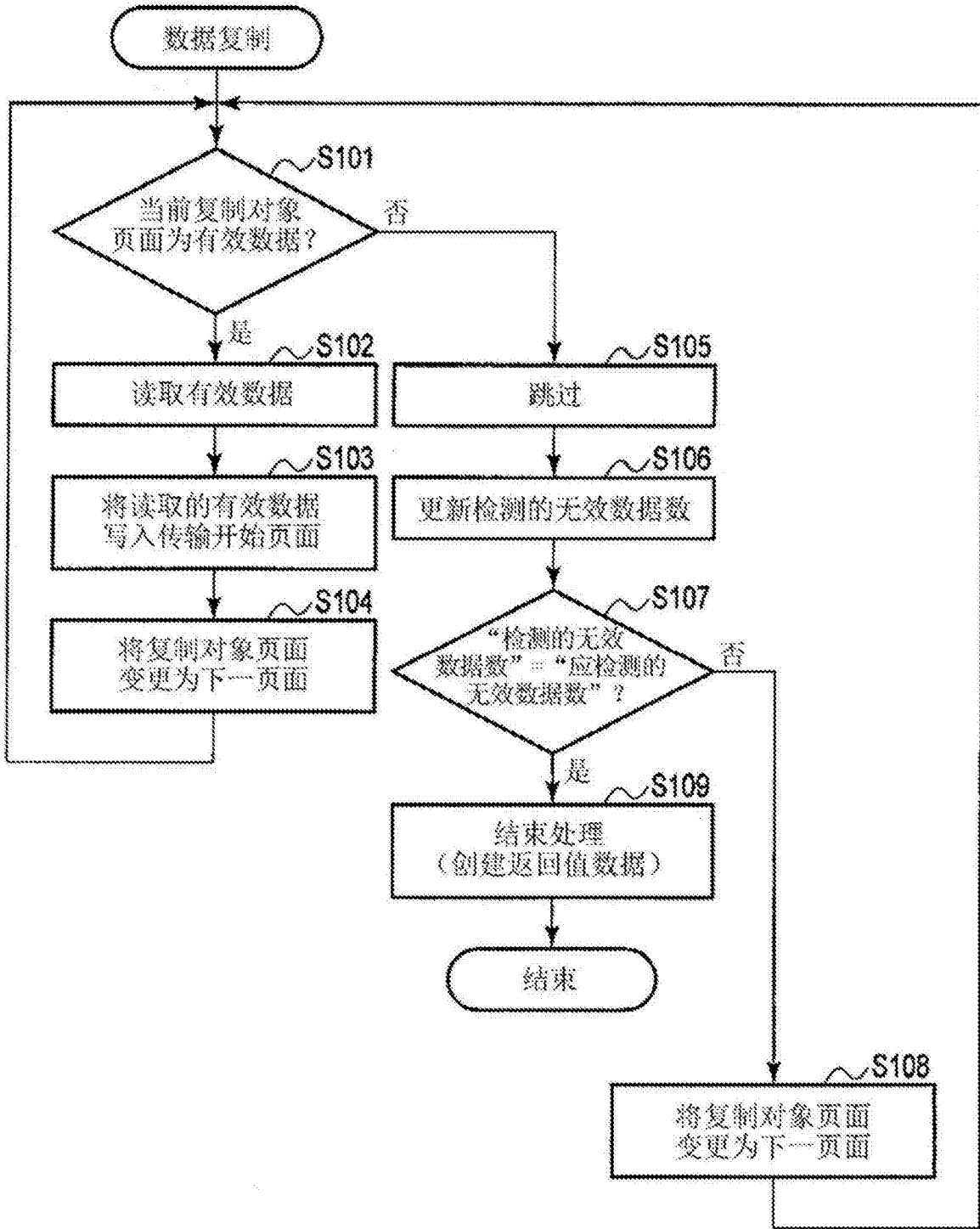


图50

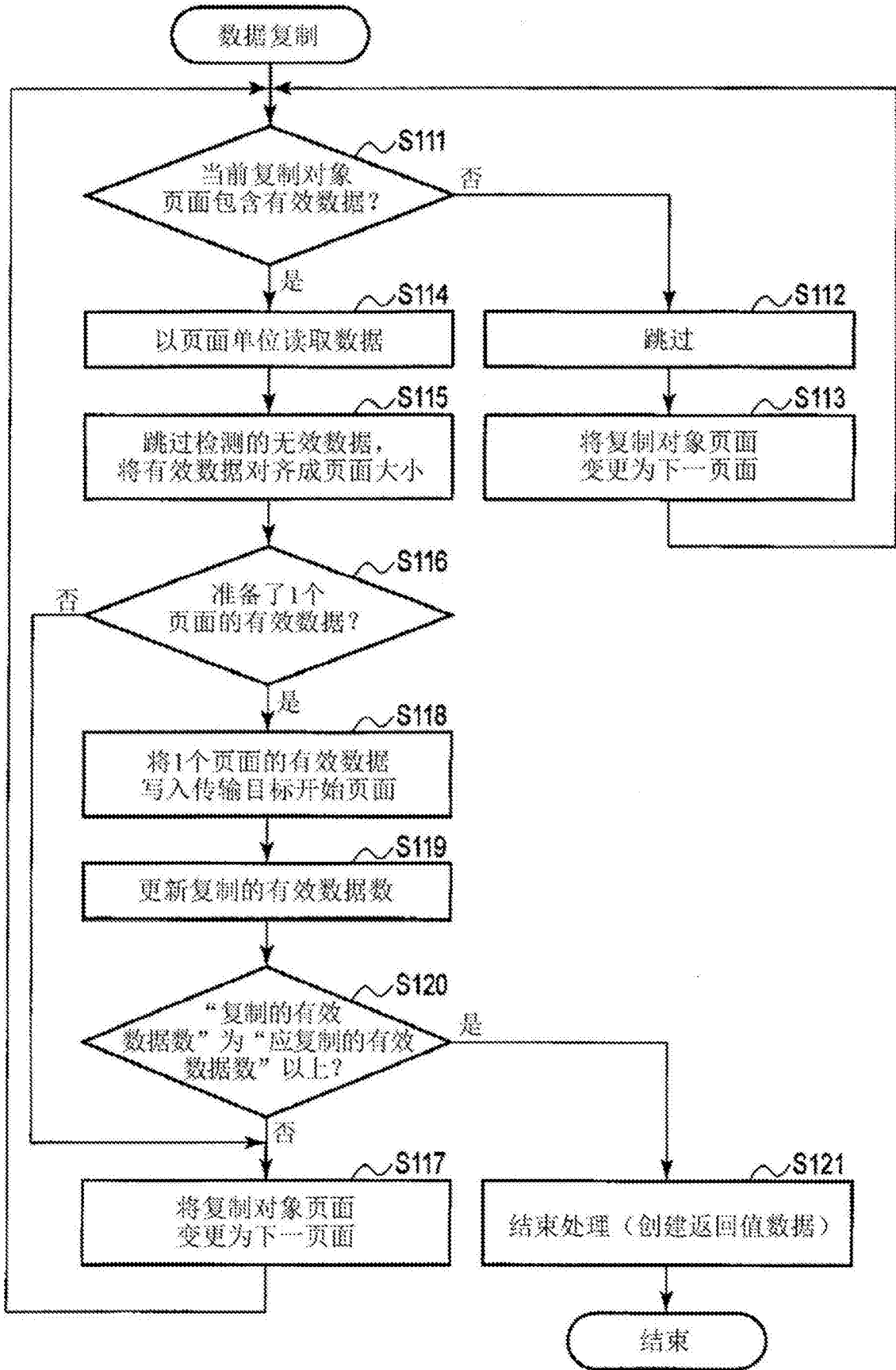


图51

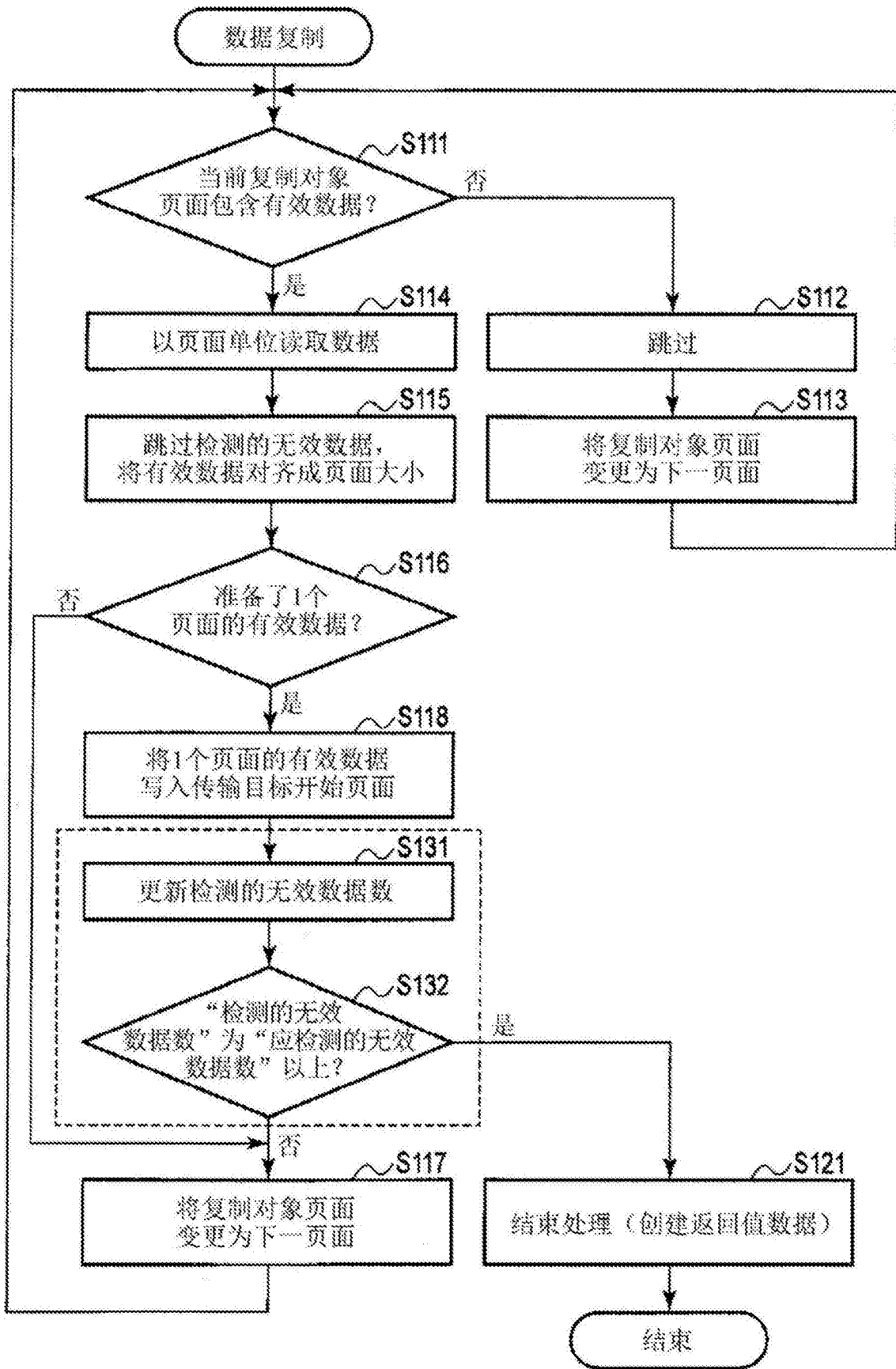


图52

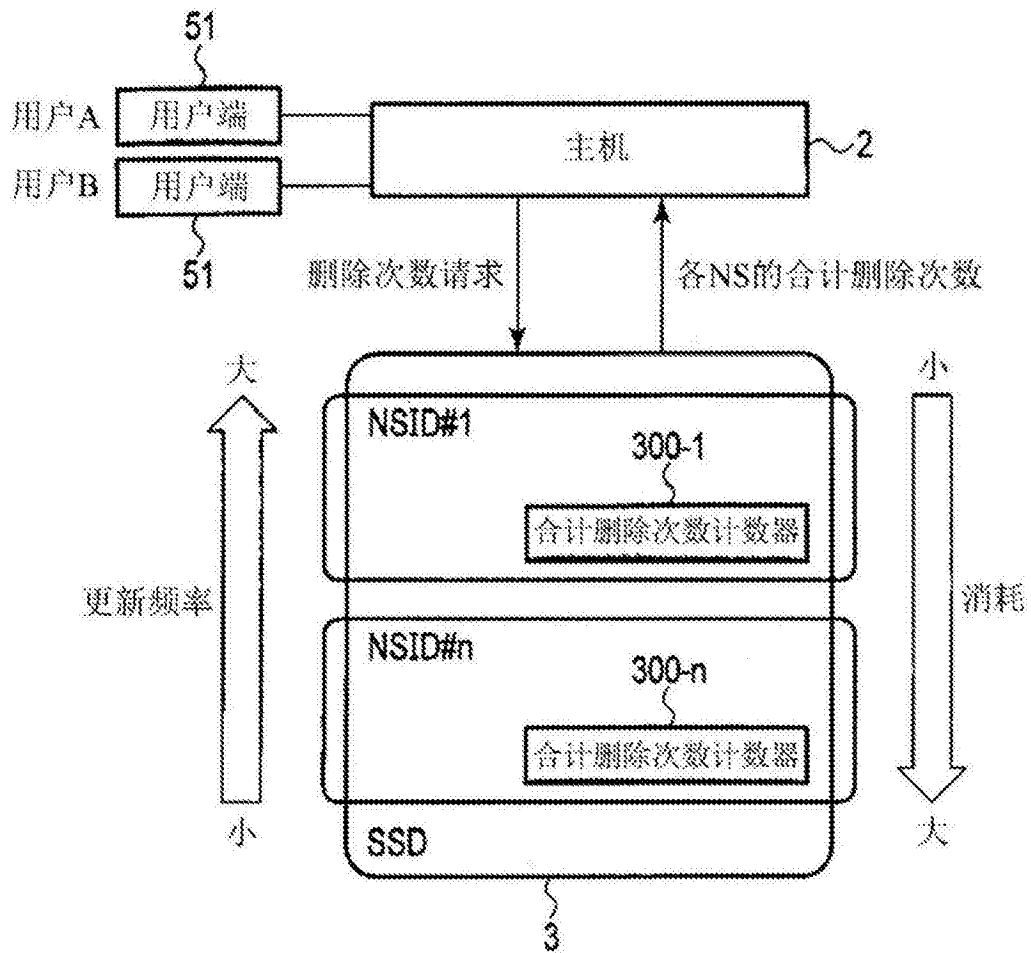


图53

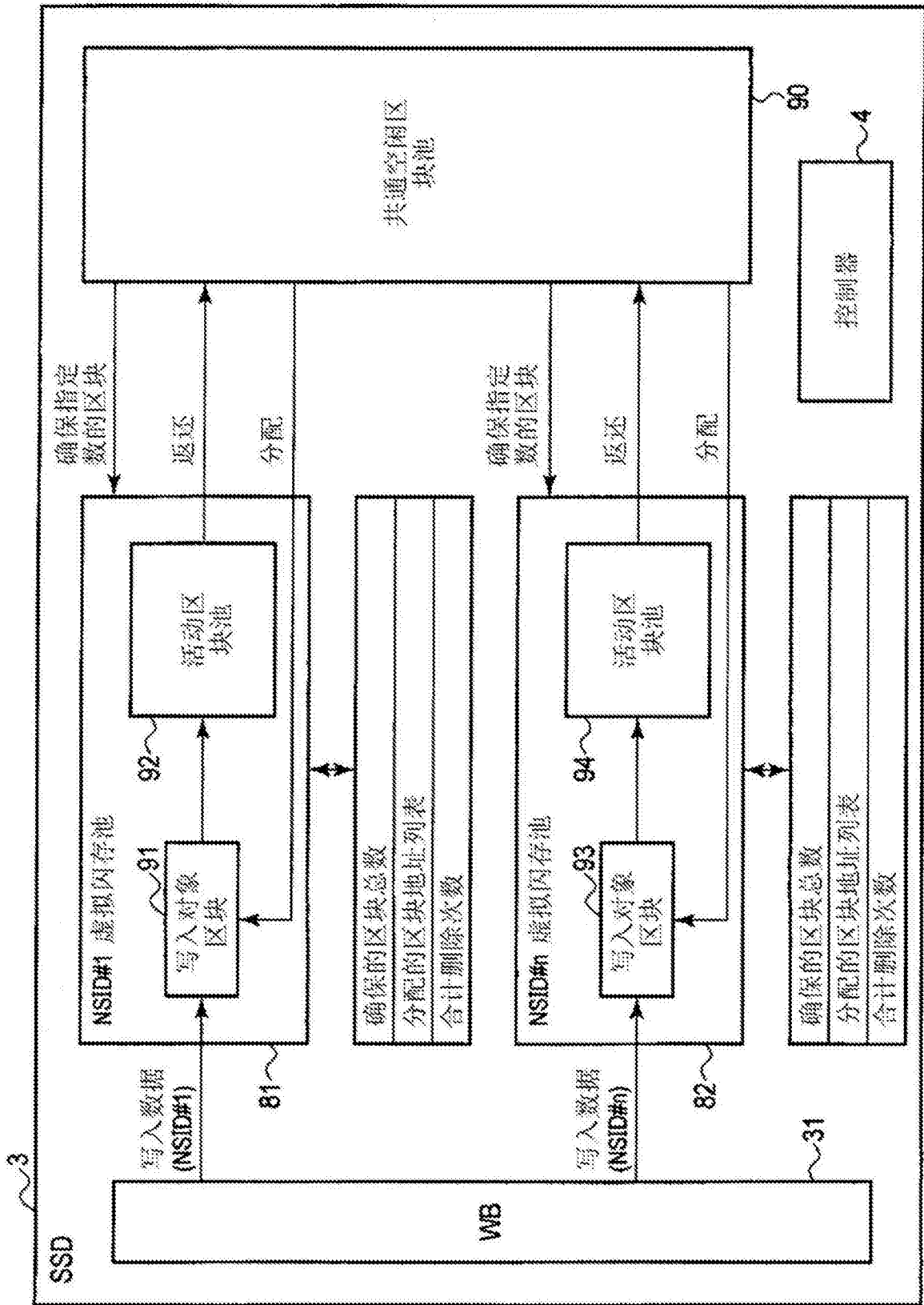


图54

命名空间分配命令

输入参数	说明
NSID	命名空间ID
物理资源量	应确保的区块数
输出参数	说明
物理资源量	确保的区块数

图55

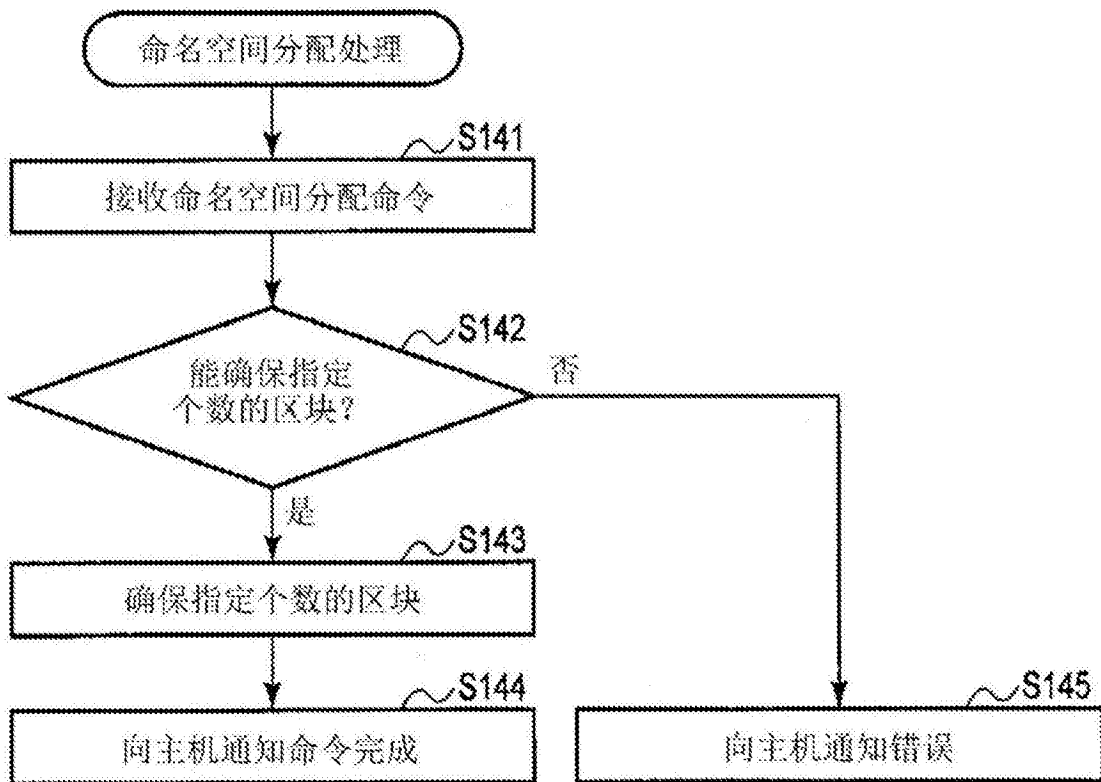


图56

区块分配、删除命令（命名空间用）

输入参数	说明
处理优先级	该命令的优先级
NSID	应分配区块的NSID
输出参数	说明
结束状态	向主机返回成功或错误
区块地址	向主机返回分配的区块地址
剩余区块数	向主机返回NSID用的剩余区块数

图57

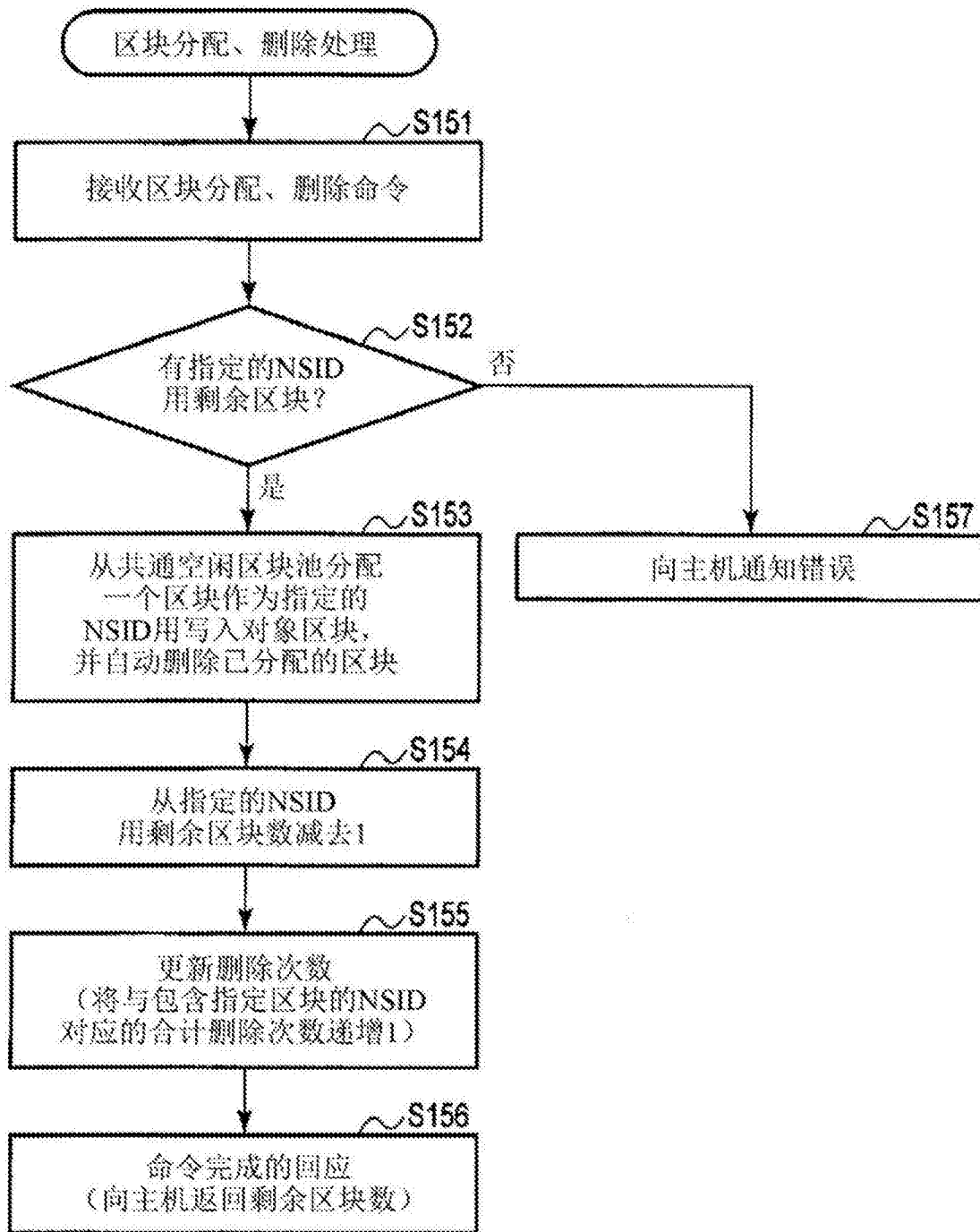


图58

删除命令（命名空间用）

输入参数	说明
区块地址	应删除的区块的区块地址
处理优先级	该命令的优先级
NSID	与应删除区块对应的NSID
输出参数	说明
结束状态	向主机返回成功或错误
区块地址	删除的区块的区块地址
合计删除次数	分配给该NSID的各区块的删除次数的合计

图59

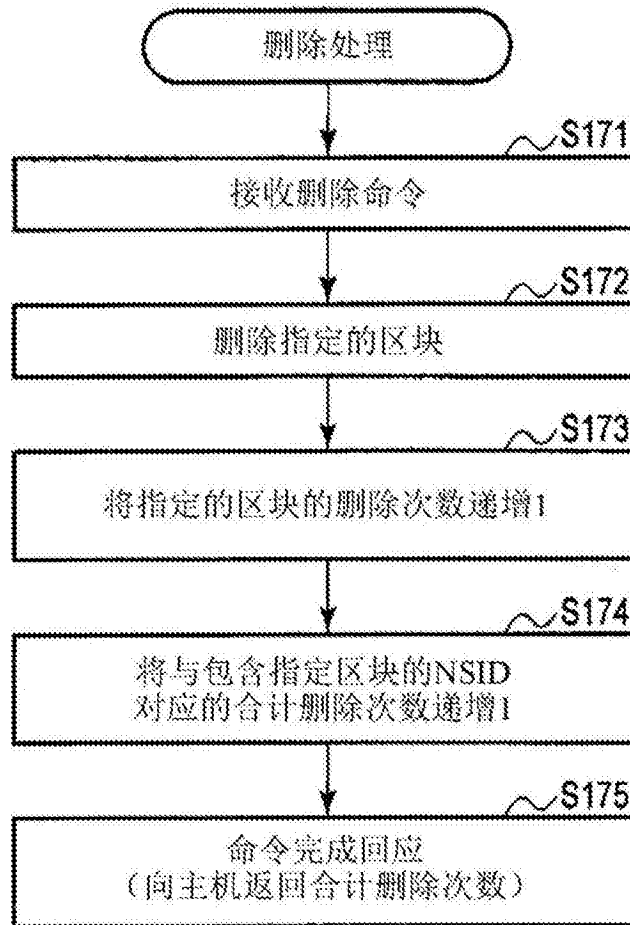


图60

区块返还命令 (命名空间用)

输入参数	说明
区块地址	应返还的区块的区块地址
处理优先级	该命令的优先级
NSID	与应返还的区块对应的NSID
输出参数	说明
结束状态	向主机返回成功或错误
剩余区块数	分配给NSID的剩余区块数

图61

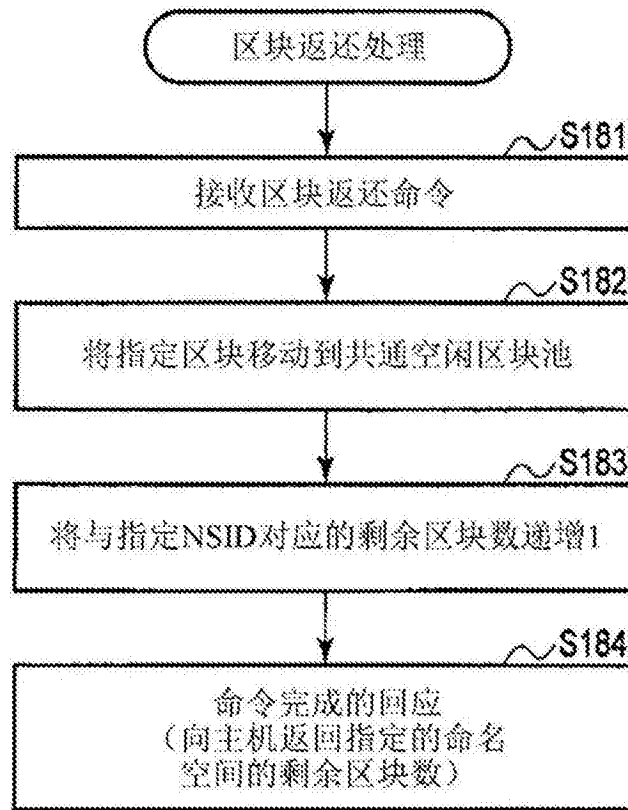


图62

删除次数获取命令 (命名空间用)

输入参数	说明
NSID	对象NSID
处理优先级	该命令的优先级
输出参数	说明
结束状态	向主机返回成功或错误
合计删除次数	指定的NSID的合计删除次数

图63

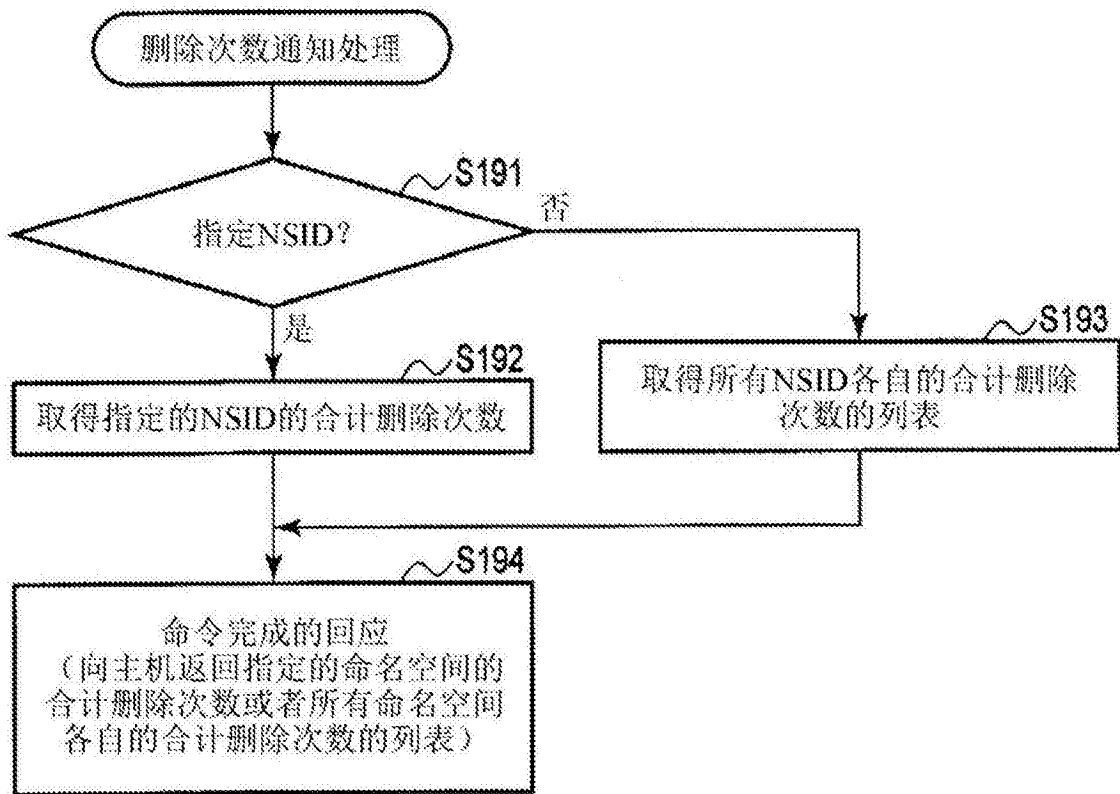


图64

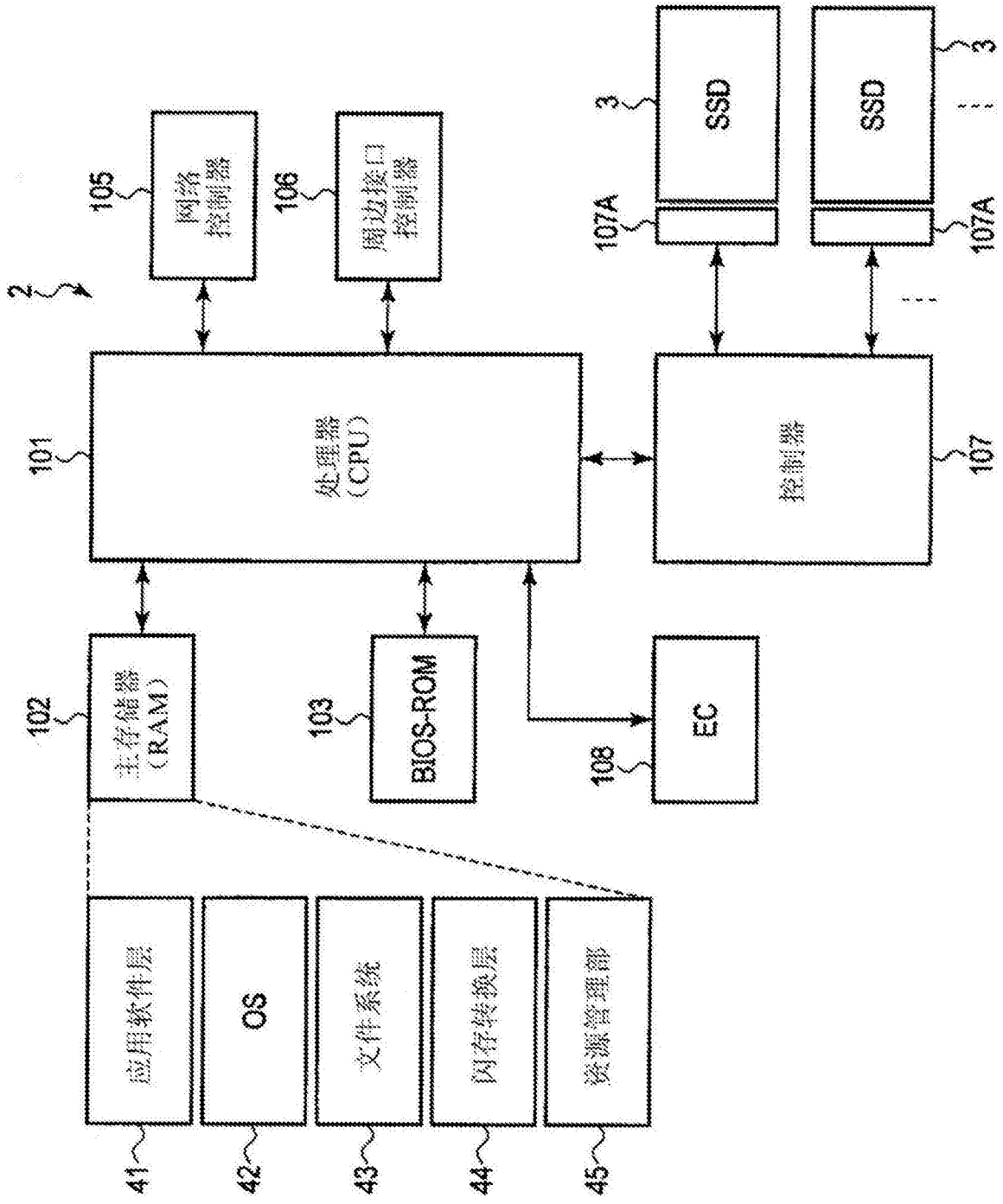


图65

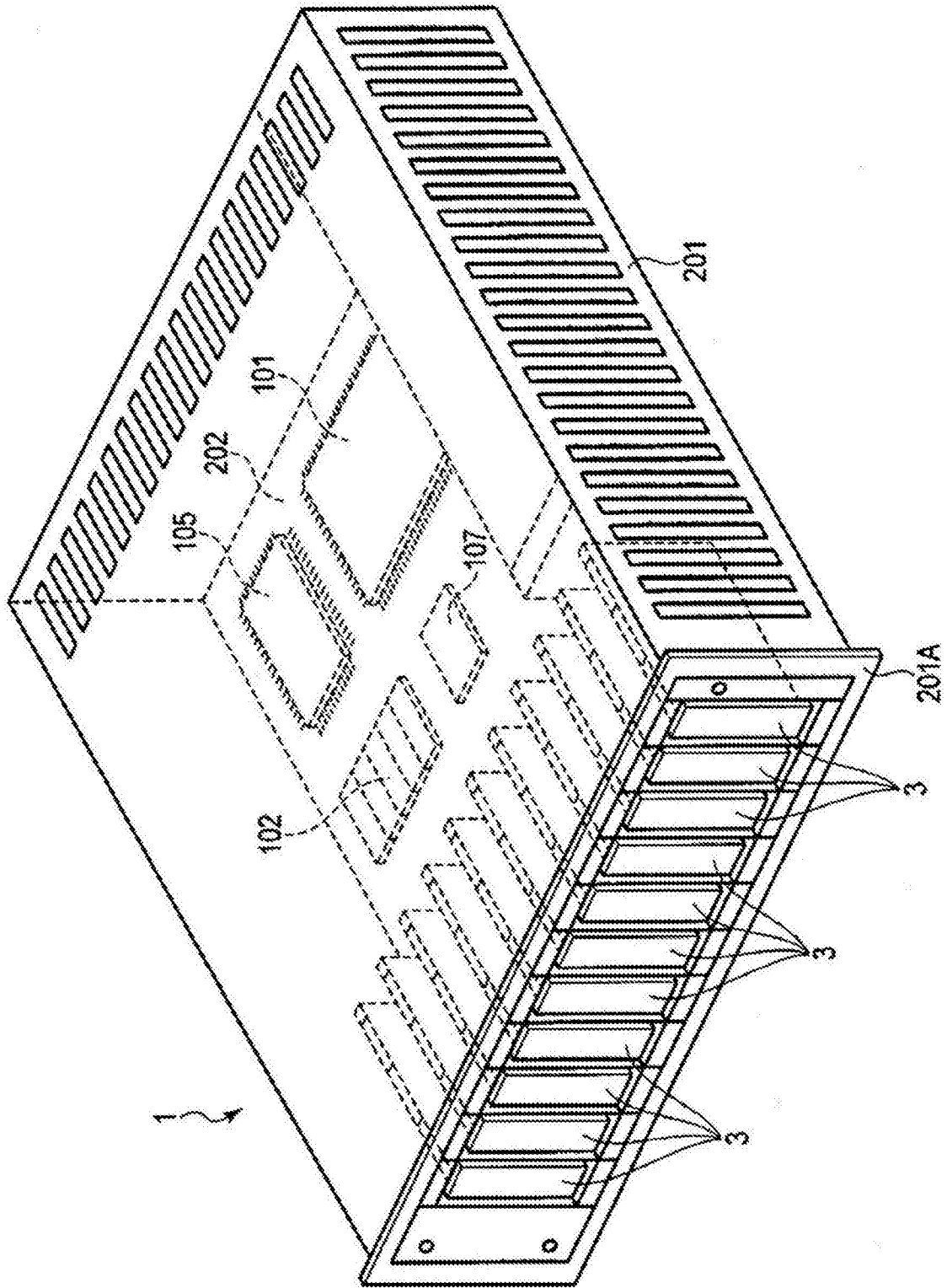


图66