



(12)发明专利

(10)授权公告号 CN 104412218 B

(45)授权公告日 2017.05.03

(21)申请号 201380032155.7

(22)申请日 2013.03.25

(65)同一申请的已公布的文献号

申请公布号 CN 104412218 A

(43)申请公布日 2015.03.11

(30)优先权数据

2012-137765 2012.06.19 JP

2013-020765 2013.02.05 JP

(85)PCT国际申请进入国家阶段日

2014.12.18

(86)PCT国际申请的申请数据

PCT/JP2013/059792 2013.03.25

(87)PCT国际申请的公布数据

W02013/190878 EN 2013.12.27

(73)专利权人 株式会社 东芝

地址 日本东京都

(72)发明人 桥本大辅

(74)专利代理机构 北京市中咨律师事务所
11247

代理人 杨晓光 于静

(51)Int.Cl.

G06F 3/06(2006.01)

(56)对比文件

CN 102246135 A, 2011.11.16,

US 7024526 B2, 2006.04.04,

US 2012/0144097 A1, 2012.06.07,

审查员 吴琼乐

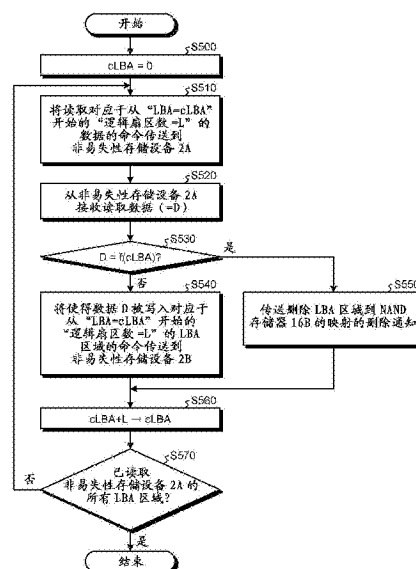
权利要求书3页 说明书29页 附图44页

(54)发明名称

存储控制程序的存储介质、控制信息处理设备的方法、信息处理系统,以及信息处理设备

(57)摘要

根据一个实施例,当从作为备份源的第一存储单元读取的数据与第一函数指示的数据不相对应时,将所述读取数据写入作为备份目的地的第二存储单元。当从所述第一存储单元读取的数据与所述第一函数指示的数据相对应时,不将所述读取数据写入所述第二存储单元,并将删除通知发送到所述第二存储单元。



1. 一种控制可与第一和第二存储单元相连的信息处理设备的方法，

所述第一存储单元包括至少可读的第一存储器，并且

所述第二存储单元包括可读和可写的第二非易失性存储器，以及控制单元，所述控制单元被配置为使得已接收的删除通知所指定的逻辑地址区域到所述第二非易失性存储器的一个或多个映射失效，并且在已接收的读取命令所指定的逻辑地址区域未映射到所述第二非易失性存储器时将第一函数指示的数据传送到所述信息处理设备，所述第一函数能够指示逻辑地址的特定数据，

所述方法包括：

从所述第一存储单元的第一逻辑地址区域读取数据；

判定从所述第一逻辑地址区域读取的数据是否与所述第一函数指示的数据相同；

在所述读取数据与所述第一函数指示的数据不同时，将所述读取数据写入所述第二存储单元的第二逻辑地址区域；以及

在所述读取数据与所述第一函数指示的数据相同时，将所述删除通知发送到所述第二存储单元以使所述第二逻辑地址区域到所述第二非易失性存储器的一个或多个映射失效，而不将所述读取数据发送到所述第二存储单元。

2. 根据权利要求1所述的控制信息处理设备的方法，其中所述第一函数指示与被发送到所述第二存储单元的逻辑地址无关的固定数据。

3. 根据权利要求1所述的控制信息处理设备的方法，其中所述第一函数针对被发送到所述第二存储单元的至少两个不同逻辑地址指示不同数据。

4. 根据权利要求1所述的控制信息处理设备的方法，其中所述第一存储单元是非易失性存储单元，并且所述第一和第二存储单元被包括在不同的非易失性存储设备中。

5. 根据权利要求1所述的控制信息处理设备的方法，

其中所述第一存储单元是非易失性存储单元，并且所述第一和第二存储单元被包括在同一非易失性存储设备中，并且

所述第一和第二逻辑地址区域的逻辑地址彼此不重叠。

6. 根据权利要求1所述的控制信息处理设备的方法，其中在执行所述判定之前，从所述第二存储单元接收所述第一函数的信息。

7. 根据权利要求1所述的控制信息处理设备的方法，其中所述第一函数指示的数据长度等于逻辑地址区域的大小。

8. 一种信息处理系统，包括：

第一存储单元，其包括至少可读的第一存储器；

第二存储单元，其包括可读和可写的第二非易失性存储器，以及控制单元，所述控制单元被配置为使得已接收的删除通知所指定的逻辑地址区域到所述第二非易失性存储器的一个或多个映射失效，并且在已接收的读取命令所指定的逻辑地址区域未映射到所述第二非易失性存储器时将第一函数指示的数据传送到信息处理设备，所述第一函数能够指示逻辑地址的特定数据；以及

所述信息处理设备，其可与所述第一和第二存储单元相连，

其中所述信息处理设备被配置为执行

从所述第一存储单元的第一逻辑地址区域读取数据；

判定从所述第一逻辑地址区域读取的数据是否与所述第一函数指示的数据相同；

在所述读取数据与所述第一函数指示的数据不同时，将所述读取数据写入所述第二存储单元的第二逻辑地址区域；以及

在所述读取数据与所述第一函数指示的数据相同时，将所述删除通知发送到所述第二存储单元以使所述第二逻辑地址区域到所述第二非易失性存储器的一个或多个映射失效，而不将所述读取数据发送到所述第二存储单元。

9. 根据权利要求8所述的信息处理系统，其中所述第一函数指示与被发送到所述第二存储单元的逻辑地址无关的固定数据。

10. 根据权利要求8所述的信息处理系统，其中所述第一函数针对被发送到所述第二存储单元的至少两个不同逻辑地址指示不同数据。

11. 根据权利要求8所述的信息处理系统，其中所述第一存储单元是非易失性存储单元，并且所述第一和第二存储单元被包括在不同的非易失性存储设备中。

12. 根据权利要求8所述的信息处理系统，

其中所述第一存储单元是非易失性存储单元，并且所述第一和第二存储单元被包括在同一非易失性存储设备中，并且

所述第一和第二逻辑地址区域的逻辑地址彼此不重叠。

13. 根据权利要求8所述的信息处理系统，其中所述信息处理设备还执行：

在判定从所述第一逻辑地址区域读取的数据是否与所述第一函数指示的数据相同之前，从所述第二存储单元接收所述第一函数的信息。

14. 根据权利要求8所述的信息处理系统，其中所述第一存储单元是易失性存储单元，并且在从程序接收将所述第一逻辑地址区域的数据存储在所述第二存储单元中的请求之后，执行从所述第一逻辑地址区域读取数据。

15. 根据权利要求8所述的信息处理系统，其中所述第一函数指示的数据长度等于逻辑地址区域的大小。

16. 一种可与第一和第二存储单元相连的信息处理设备，

所述第一存储单元包括至少可读的第一存储器，并且

所述第二存储单元包括可读和可写的第二非易失性存储器，以及控制单元，所述控制单元被配置为使得已接收的删除通知所指定的逻辑地址区域到所述第二非易失性存储器的一个或多个映射失效，并且在已接收的读取命令所指定的逻辑地址区域未映射到所述第二非易失性存储器时将第一函数指示的数据传送到所述信息处理设备，所述第一函数能够指示逻辑地址的特定数据，

所述信息处理设备被配置为执行：

从所述第一存储单元的第一逻辑地址区域读取数据；

判定从所述第一逻辑地址区域读取的数据是否与所述第一函数指示的数据相同；

在所述读取数据与所述第一函数指示的数据不同时，将所述读取数据写入所述第二存储单元的第二逻辑地址区域；以及

在所述读取数据与所述第一函数指示的数据相同时，将所述删除通知发送到所述第二存储单元以使所述第二逻辑地址区域到所述第二非易失性存储器的一个或多个映射失效，而不将所述读取数据发送到所述第二存储单元。

17. 根据权利要求16所述的信息处理设备, 其中所述第一函数指示与被发送到所述第二存储单元的逻辑地址无关的固定数据。

18. 根据权利要求16所述的信息处理设备, 其中所述第一函数针对被发送到所述第二存储单元的至少两个不同逻辑地址指示不同数据。

19. 根据权利要求16所述的信息处理设备, 其中所述第一存储单元是非易失性存储单元, 并且所述第一和第二存储单元被包括在不同的非易失性存储设备中。

20. 根据权利要求16所述的信息处理设备,

其中所述第一存储单元是非易失性存储单元, 并且所述第一和第二存储单元被包括在同一非易失性存储设备中, 并且

所述第一和第二逻辑地址区域的逻辑地址彼此不重叠。

21. 根据权利要求16所述的信息处理设备, 其中所述信息处理设备还执行:

在判定从所述第一逻辑地址区域读取的数据是否与所述第一函数指示的数据相同之前, 从所述第二存储单元接收所述第一函数的信息。

22. 根据权利要求16所述的信息处理设备, 其中所述第一存储单元是易失性存储单元, 并且在从程序接收将所述第一逻辑地址区域的数据存储在所述第二存储单元中的请求之后, 执行从所述第一逻辑地址区域读取数据。

23. 根据权利要求16所述的信息处理设备, 其中所述第一函数指示的数据长度等于逻辑地址区域的大小。

存储控制程序的存储介质、控制信息处理设备的方法、信息处理系统,以及信息处理设备

[0001] 相关申请交叉引用

[0002] 本申请基于并主张2012年6月19日提交的编号为2012-137765的日本专利申请,以及2013年2月5日提交的编号为2013-020765的日本专利申请的优先权益,所有这些申请的全部内容在此引入作为参考。

技术领域

[0003] 此处描述的实施例一般地涉及存储用于将第一存储器的数据备份到非易失性第二存储器的控制程序的存储介质、控制信息处理设备的方法、信息处理系统,以及信息处理设备。

背景技术

[0004] 在操作系统(OS)上,当非易失性存储设备(例如,硬盘或固态驱动器(SSD))的文件数据被删除时,数据实际并未从非易失性存储设备中删除,而是仅删除了与OS上的数据对应的管理信息。由于非易失性存储设备上的删除处理并非在每次删除OS上数据时执行,因此,操作性能被提高到不执行删除处理的程度。该系统对于访问缓慢的硬盘驱动器特别有效。

[0005] 另一方面,由于OS上被明确删除的数据被识别为在非易失性存储设备上仍然有效的数据(主机无效-设备有效数据),因此非易失性存储设备上不存在任何有效数据的可用区域少于被OS识别的可用区域。具体来说,用尽非易失性存储设备上的可用区域可能导致这样的SSD中出现大问题:其中在某些情况下,逻辑地址与非易失性存储设备的物理地址之间不再存在一一对应关系。

[0006] 可能出现以下问题:

[0007] 写入性能严重劣化的问题,因为通过以下方式确保新的可用块的过程需要一定的工作量:此方式是,SSD控制器检索被识别为具有有效数据的NAND闪存的物理块(活动块)中的无效数据区域,并且在SSD的可用区域被用尽(即,用尽被SSD控制器识别为没有有效数据的NAND闪存的物理块(可用块))的情况下,在将数据写入SSD时针对有效数据执行垃圾收集(NAND垃圾收集);以及

[0008] SSD可靠性严重劣化的问题,在SSD因为大量无效数据而被用尽可用区域的状态下,当执行写入时,频繁发生NAND垃圾收集,从而导致写入访问范围被限于特定区域或者额外的块被删除。

[0009] 上述问题可通过在删除主机上的数据时,执行从主机(信息处理设备)到SSD的删除通知来避免。数据删除通知的实例包括在INCITS ATA/ATAPT Command Set-2(ACS-2)中描述的数据集管理命令(一般称为修剪(trim)命令),以及在NVM Express Revision 1.1中描述的11h Dataset Management命令的Deallocate(AD)。当在该系统中删除主机上的数据时,主机通知SSD已删除数据所存在的逻辑地址空间,这样便可在以后将这

些区域视为可用区域(即使在SSD上)。这样,修剪命令实现将主机无效-设备有效数据更改为主机无效-设备无效数据,因而,SSD可重新确保可用区域。

[0010] 当执行删除通知时,SSD控制器通过访问数据管理区域,使得对应的区域失效。在失效过程中,通过重写SSD的管理信息,将对应的区域识别为无效。对应区域的数据可能不被实际删除,也可能被删除。

[0011] 专利文献1:JP 2010-157133 A

[0012] 专利文献2:JP 2010-161199 A

[0013] 专利文献3:JP 2011-29586 A

[0014] 专利文献4:JP 2011-128998 A

[0015] 专利文献5:US 2010/0217927

[0016] 专利文献6:JP 2009-230385 A

[0017] 非专利文献1:ATA/ATAPI Command Set-2 (ACS-2) d2015r6 2011年2月22日(<http://www.t13.org/>),98-99页,以及50页;Data Set Management命令(修剪命令),130-171页;ECh IDENTIFY DEVICE命令,342-365页;SCT Command Transport,217页;25h READ DMA EXT命令,218页;60h READ FPDMA QUEUED命令,322页;35h WRITE DMA EXT命令,324页;61h WRITE FPDMA QUEUED命令

[0018] 非专利文献2:Serial ATA国际组织:Serial ATA Revision 3.1 Gold Revision,2011年7月18日,<http://www.serialata.org/>

[0019] 非专利文献3:PCI Express Base Specification Revision 3.0,2010年11月10日,<http://www.pcisig.com/>

[0020] 非专利文献4:Serial ATA Technical Proposal:SATA31_TPR_C108标题:Device Sleep,<http://www.serialata.org/>

[0021] 非专利文献5:“Well-known Storage Network(公知的存储网络,Masaru Kitsuregawa编辑,第1版,Ohmsha,Ltd.,6-9页,67-93页,2011年5月20日)”

附图说明

[0022] 图1是示出根据第一实施例的信息处理系统的功能配置实例的框图;

[0023] 图2是示出SSD的内部配置实例的功能框图;

[0024] 图3是示出SSD的管理信息的示意图;

[0025] 图4是示出SSD的写入处理实例的流程图;

[0026] 图5是示出SSD的NAND存储器的垃圾收集处理实例的流程图;

[0027] 图6是示出当SSD接收删除通知时的处理实例的流程图;

[0028] 图7是示出SSD的读取处理实例的流程图;

[0029] 图8是示出当控制工具被存储在SSD中时,信息处理系统的功能配置实例的框图;

[0030] 图9是示出当控制工具被存储在另一外部存储设备中时,信息处理系统的功能配置实例的框图;

[0031] 图10是示出当从WEB安装控制工具时,信息处理系统的功能配置实例的框图;

[0032] 图11是示出当从光设备安装控制工具时,信息处理系统的功能配置实例的框图;

[0033] 图12是示出当从USB存储器安装控制工具时,信息处理系统的功能配置实例的框图;

图；

[0034] 图13是示出信息处理设备的分级功能配置实例的框图；

[0035] 图14是示出信息处理系统的外观配置的示意图；

[0036] 图15是示出信息处理系统的另一外观配置的示意图；

[0037] 图16是示出信息处理系统的外观配置的示意图；

[0038] 图17是示出信息处理系统的另一外观配置的示意图；

[0039] 图18是示出根据第一实施例的由控制程序执行的备份处理顺序的流程图；

[0040] 图19是示出根据第二实施例的由控制程序执行的备份处理顺序的流程图；

[0041] 图20是示出LBA与数据之间的关系的实例的示意图；

[0042] 图21是示出在图20中的数据的情况下，0-fill扇区的存在概率分布实例的示意图；

[0043] 图22是示出LBA与数据之间的关系的另一实例的示意图；

[0044] 图23是示出在图22中的数据的情况下，0-fill和1-fill扇区的存在概率分布实例的示意图；

[0045] 图24是示出LBA与数据之间的关系的另一实例的示意图；

[0046] 图25是示出在图24中的数据的情况下，0-fill、320a-fill和1-fill扇区的存在概率分布实例的示意图；

[0047] 图26是示出LBA与数据之间的关系的另一实例的示意图；

[0048] 图27是示出在图26中的数据的情况下的开放扇区的示意图；

[0049] 图28是示出判定函数 $f(LBA)$ 的方法的示意图；

[0050] 图29是示出根据第四实施例的由控制程序执行的备份处理顺序的流程图；

[0051] 图30是示出根据第四实施例的由控制程序执行的另一备份处理顺序的流程图；

[0052] 图31是概念性地示出其中网络存储驱动器或云存储被用作作为备份目的地的非易失性存储设备的实例的示意图；

[0053] 图32是示出根据第六实施例的信息处理系统的功能配置实例的框图；

[0054] 图33是示出信息处理设备的分级功能配置实例的框图；

[0055] 图34是示出根据第六实施例的由控制程序执行的备份处理顺序的流程图；

[0056] 图35是示出OS文件管理表的示意图；

[0057] 图36是示出信息处理设备的分级功能配置中的数据写入流的示意图；

[0058] 图37是示出信息处理设备的分级功能配置中的另一数据写入流的示意图；

[0059] 图38是示出根据第七实施例的使用网络的信息处理系统实例的框图；

[0060] 图39是示出根据第七实施例的使用网络的另一信息处理系统实例的框图；

[0061] 图40是示出根据第七实施例的使用网络的另一信息处理系统实例的框图；

[0062] 图41是示出根据第七实施例的使用网络的另一信息处理系统实例的框图；

[0063] 图42是示出根据第七实施例的由控制程序执行的备份处理顺序的流程图；

[0064] 图43是示出根据第八实施例的使用网络的另一信息处理系统实例的框图；

[0065] 图44是示出根据第八实施例的使用网络的另一信息处理系统实例的框图；

[0066] 图45是示出根据第九实施例的使用网络的信息处理系统实例的框图；

[0067] 图46是示出根据第九实施例的使用的元数据实例的示意图；

- [0068] 图47是示出根据第九实施例的由控制程序执行的备份处理顺序的流程图；
[0069] 图48是示出文件ID列表的数据结构的示意图；以及
[0070] 图49是示出根据第十实施例的系统配置的示意图。

具体实施方式

[0071] 根据一个实施例，第一存储单元包括至少可读的第一存储器。第二存储单元包括可读和可写的第二非易失性存储器，控制单元被配置为使得已接收的删除通知所指定的逻辑地址区域到所述第二非易失性存储器的一个或多个映射失效，并且在已接收的读取命令所指定的逻辑地址区域未映射到所述第二非易失性存储器时传送第一函数指示的数据。存储介质存储要被加载到可与第一和第二存储单元相连的信息处理设备的控制程序。所述控制程序导致信息处理设备执行：第一过程，其从所述第一存储单元的第一逻辑地址区域读取数据；第二过程，其判定从所述第一逻辑地址区域读取的数据是否与所述第一函数指示的数据相同；第三过程，其在所述读取数据与所述第一函数指示的数据不同时，将所述读取数据写入所述第二存储单元的第二逻辑地址区域；以及第四过程，其在所述读取数据与所述第一函数指示的数据相同时，将所述删除通知发送到所述第二存储单元以使所述第二逻辑地址区域到所述第二非易失性存储器的一个或多个映射失效，而不将所述读取数据发送到所述第二存储单元。

[0072] 在一个实施例中，在备份过程中使用修剪命令将主机有效-设备有效数据处理为主机有效-设备无效数据。因此，实现高速备份SSD的数据，增加作为备份目的地的SSD中的可用块数，以及减少作为备份目的地的SSD中的写入数据量。

[0073] 首先，将描述本说明书中使用的术语。

[0074] 主机无效-设备有效数据是指针对主机（信息处理设备）或OS无效，但是针对设备（非易失性存储设备）有效的数据。

[0075] 主机有效-设备有效数据是指不仅针对主机有效，而且还针对非易失性存储设备有效的数据。

[0076] 主机有效-设备无效数据是指针对主机有效，但是针对非易失性存储设备无效的数据。

[0077] 主机无效-设备无效数据是指不仅针对主机无效，而且还针对非易失性存储设备无效的数据。

[0078] 主机有效数据是指至少针对主机有效的数据。

[0079] 主机无效数据是指至少针对主机无效的数据。

[0080] 设备有效数据是指至少针对非易失性存储设备有效的数据。

[0081] 设备无效数据是指至少针对非易失性存储设备无效的数据。

[0082] 下面将参考附图详细地描述根据各实施例的存储控制程序的存储介质、控制信息处理设备的方法、信息处理系统，以及信息处理设备。本发明不限于这些实施例。

[0083] （第一实施例）

[0084] 图1是示出作为根据第一实施例的信息处理系统的实例的信息处理系统1的第一配置的示意图。信息处理系统1包括存储设备2A、信息处理设备3（主机3）、将非易失性存储设备2A连接到信息处理设备3的存储接口19A，以及上面可安装用于备份存储设备2A的非易

失性存储设备2B的存储接口19B。在该实施例中,存储设备2A被配置为非易失性存储设备2A。SSD(固态驱动器)被用作非易失性存储设备2A和2B。但是,例如,也可使用其它非易失性存储设备,例如硬盘驱动器、混合硬盘驱动器、SD卡、USB存储器、NAND型闪存、ReRAM(电阻式随机存取存储器)或磁带。非易失性存储设备2A和2B可以是不同种类的非易失性存储设备。作为可在非易失性存储设备2B中寻址的逻辑地址的LBA(逻辑块寻址)的最大值(LBA计数)优选地与非易失性存储设备2A的LBA计数相同,或者优选地大于非易失性存储设备2A的LBA计数。作为存储设备2A,可使用诸如DRAM(动态随机存取存储器)或SRAM(静态随机存取存储器)之类的易失性存储设备,而不是非易失性存储设备。此外,存储设备2A可以是仅执行读取操作和写入操作中的读取操作的存储设备。

[0085] 在该实施例中,SATA(串行高级技术附件)接口被用作接口19A和19B。也可使用其它接口,例如USB(通用串行总线)、SAS(串行连接SCSI)、Thunderbolt(注册商标)、PCI EXPRESS、LAN(局域网)、光纤通道和以太网。接口19A和19B可以是不同种类的接口。例如,接口19A和接口19B可以是分别是SATA接口和USB接口。

[0086] 作为备份目的地的非易失性存储设备2B可安装在物理上远离信息处理设备3的位置,也可作为网络驱动器或云存储,经由网络安装在信息处理设备3上。

[0087] 在该实施例中,例如,将描述SATA接口被用作接口19B的情况。

[0088] CPU(控制电路)4是信息处理设备3的中央处理单元。信息处理设备3中的各种计算和控制由CPU 4执行。CPU 4经由南桥7控制非易失性存储设备2A和2B或者诸如DVD-ROM之类的光驱10。CPU 4经由北桥5控制主存储器6。例如,DRAM可被用作主存储器6。

[0089] 用户通过诸如键盘14或鼠标15之类的输入设备控制信息处理设备3。来自键盘14或鼠标15的信号经由USB(通用串行总线)控制器13和南桥7被CPU 4处理。CPU 4经由北桥5和显示控制器8将图像数据、文本数据等传送到显示器(显示设备)9。用户可通过显示器9,识别来自信息处理设备3的图像数据、文本数据等。

[0090] CPU 4是处理器,其在安装之后用于控制信息处理系统1的进程并且 执行从非易失性存储设备2A加载到主存储器6的操作系统(OS) 100。当光驱10允许执行所加载的光盘中的读取处理和写入处理中的至少一个处理时,CPU 4执行此过程。CPU 4还执行存储在BIOS(基本输入/输出系统)-ROM 11中的系统BIOS。系统BIOS是用于信息处理系统1中的硬件控制的程序。在ROM 11中,可存储UEFI(统一可扩展固件接口)固件,并且该固件可由CPU 4而非系统BIOS执行。CPU 4还经由南桥7控制LAN(局域网)。

[0091] 北桥5是连接到CPU 4的局部总线的桥接设备。北桥5还包括存储控制器,该控制器执行主存储器6的访问和控制。此外,北桥5具有执行与显示控制器8的通信等的功能。

[0092] 主存储器6临时存储程序或数据,并且充当CPU 4的工作存储器。例如,DRAM(动态随机存取存储器)、MRAM(磁阻式随机存取存储器)、ReRAM(电阻式随机存取存储器)或SRAM(静态随机存取存储器)可被用作主存储器6。主存储器6包括存储OS 100的存储区域6A和存储控制工具(控制程序)200的存储区域6B。如一般所知的那样,OS 100是管理整个信息处理设备3的程序,例如通过以下方式执行上述管理:管理信息处理设备3的输入/输出设备,管理磁盘或存储器,以及执行控制以便软件可使用信息处理设备3的硬件。控制程序200是具有将数据从存储设备1A备份到非易失性存储设备2B的备份功能的程序。在该实施例中,OS 100和控制程序200彼此具有区别,但是控制程序200可以是OS 100的一部分。

[0093] 显示控制器8是图形处理单元,其控制信息处理系统1的显示器9。南桥7是连接到CPU 4的局部总线的桥接设备。南桥7控制非易失性存储设备2A,该设备是通过SATA接口19A存储各种类型软件和数据的存储设备。非易失性存储设备2A可以直接连接到CPU 4而不需要经过南桥7。

[0094] 在信息处理系统1中,对非易失性存储设备2A和2B的访问以逻辑扇区为单元执行,逻辑扇区是LBA的最小单元。例如,写入命令(写入请求)、读取命令(读取请求)、闪存命令等经由SATA接口19A和19B被输出到非易失性存储设备2A和2B。

[0095] 南桥7具有执行对BIOS-ROM 11、光驱10、LAN控制器12和USB控制器13的访问和控制的功能。键盘14和鼠标15连接到USB控制器13。

[0096] [SSD]

[0097] 图2是示出作为备份目的地的非易失性存储设备2B的框图。在该实施例中,将描述这样的实例:其中使用固态驱动器(SSD)作为非易失性存储设备2B。作为备份源的非易失性存储设备2A可与非易失性存储设备2B具有相同的配置。作为备份源的非易失性存储设备2A本身不接收修剪命令。因此,即使非易失性存储设备2A不满足此处指示的条件(例如,当非易失性存储设备2A不支持修剪命令时),也可应用本发明。在该实施例中,非易失性存储设备2B包括作为非易失性半导体存储器的NAND型闪存(下文简称为NAND存储器)16B;经由SATA接口19B将信号发送到信息处理设备3以及从信息处理设备3接收信号的接口控制器(IFC) 42B;作为半导体存储器并包括管理信息区域46B和高速缓冲存储器(CM) 44B的RAM(随机存取存储器) 40B,该管理信息区域46B存储包括逻辑-物理转换表(查找表(LUT))的管理信息,其中该转换表具有从LBA到NAND存储器16B的物理地址的映射,该高速缓冲存储器(CM) 44B充当IFC 42B与NAND存储器16B之间的中间缓冲区;SSD控制器(SSDC) 41B,其执行对NAND存储器16B的管理和控制;以及对IFC 42B的控制;以及将组成单元彼此连接的总线43B。

[0098] 作为RAM 40B,可使用DRAM(动态随机存取存储器)、SRAM(静态随机存取存储器)、FeRAM(铁电体随机存取存储器)、MRAM(磁阻式随机存取存储器)、PRAM(相变随机存取存储器)、ReRAM(电阻式随机存取存储器)等。RAM 40B可包括在SSDC 41B中。

[0099] NAND存储器16B存储信息处理设备3指定的用户数据,并将管理信息区域46B中的管理信息存储在管理信息区域45B中作为备份信息。NAND存储器16B包括存储单元阵列,其中多个存储单元以矩阵形式排列,并且各个存储单元可使用上页和下页存储多个位。NAND存储器16B被配置为包括多个存储芯片,并且每个存储芯片被配置为排列多个作为擦除操作单元的块。NAND存储器16B针对每一页写入数据和读取数据。块被配置为包括多个页。

[0100] RAM 40B包括高速缓冲存储器(CM) 44B,其充当信息处理设备2与NAND存储器16之间的数据传送高速缓存。RAM 40B充当存储管理信息的存储器和用于其它工作领域的存储器。例如,当非易失性存储设备2B通电时,SSDC 41B将管理信息从管理信息区域45B加载到管理信息区域46B。当关闭电源,或者当从信息处理设备3接收待机命令或清仓高速缓存命令时,管理信息区域46B中的管理信息被定期保存并且存储在NAND存储器16B的管理信息区域45B中。

[0101] 如图3所示,管理表46B包括可用块表、活动块表和活动页表。物理块ID和物理页地址包括在物理地址中。活动块表具有逻辑地址(LBA)到物理块ID(物理块地址)的映射。在活

动块表中注册的物理块(活动块)被配置为至少针对SSDC 41B存储有效数据。可用块表管理不从LBA映射的物理块。由于在可用块表中注册的物理块(可用块)不从LBA映射,因此物理块不存储有效数据。

[0102] 活动页表具有LBA到物理页地址的映射。在活动页表中注册的物理页(活动页)被配置为至少针对SSDC 41B存储有效数据。在活动页表中注册的物理页与属于该物理页的物理块ID(物理块地址)关联。在管理表中管理的LBA优选地为具有非易失性存储设备2B可管理的最小数据大小或具有大于最小数据大小的数据大小的地址。为了减小管理表46B的数据大小,通过对LBA执行右移计算获取的地址可被用作通过预定位存储在活动块表或活动页表中的逻辑地址。

[0103] 充当非易失性存储设备2B的控制器的SSDC 41B的功能通过执行存储在NAND存储器16B中的系统程序(固件)的处理器、各种硬件电路等实现。当SSDC 41B接收写入请求、高速缓存闪存请求、读取请求等各种命令时,SSDC 41B在信息处理设备3与NAND存储器16B之间传送数据,更新存储在区域64B和区域44B中的管理信息,写入NAND存储器16B的数据的ECC编码,从NAND存储器16B读取的数据的ECC解码等。

[0104] 当信息处理设备3向非易失性存储设备2B发出读取请求或写入请求时,信息处理设备3经由SATA接口19B将逻辑地址区域的信息发送到非易失性存储设备2B。逻辑地址区域的信息包括LBA和扇区大小。

[0105] IFC 42B具有从信息处理设备3接收读取请求、写入请求或其它请求和数据,并将所接收的请求和数据发送到SSDC 41B或在SSDC 41B的控制下将数据发送到RAM 40B的功能。

[0106] <写入处理>

[0107] 图4是示出作为备份目的地的非易失性存储设备2B的写入处理的流程图。非易失性存储设备2B从信息处理设备3接收写入命令(步骤S100)。写入命令包括写入命令、LBA、扇区计数和数据。

[0108] SSDC 41B从DRAM 40B读取可用块表(步骤S110)并且从可用块表选择物理块ID。当可用块未在可用块中注册时,SSDC 41B执行下面描述的NAND垃圾收集(NAND GC),创建可用块,将该可用块添加到可用块表并选择可用块的物理块ID(步骤S120和步骤S130)。

[0109] SSDC 41B在对应于选定的物理块ID的可用块中对数据(写入数据)进行编程(步骤S140)。接下来,SSDC 41B从可用块表删除经过编程的物理块ID(步骤S150)。SSDC 41B更新活动块表和活动页表(步骤S160)。也就是说,LBA和对应于经过编程的块的物理块ID被添加到活动块表,并且LBA、物理块ID和物理页地址被添加到活动页表。SSDC 41B可在步骤S140的编程处理之前擦除要写入的物理块。

[0110] <NAND垃圾收集>

[0111] 接下来,将描述非易失性存储设备2B的NAND垃圾收集处理(GC)。图5是示出非易失性存储设备2B的NAND垃圾收集处理的流程图。物理块内包括的所有页当中注册在活动页表中的活动页(有效页)从LBA映射。包括在物理块内的所有页并非始终为有效页。不是有效页的无效页不从LBA映射。物理块具有等于无效页数量的可用空间,可通过收集无效页和有效页创建可用块。

[0112] 首先,SSDC 41B设定行号=0,以及可用空间累计值s=0(步骤S200)。接下来,SSDC

41B从DRAM 40B读取活动块表的第一行,并选择物理块ID(步骤S210)。在步骤S210选定的物理块ID被称为当前物理块ID。接下来,SSDC 41B从DRAM 40B读取活动页表并选择与当前物理块ID对应的所有物理页地址(步骤S220)。

[0113] 接下来,SSDC 41B计算当前物理块当中无效页的大小并将无效页大小替换为变量 v (步骤S230)。接下来,SSDC 41B判定是否 $v>0$ (步骤S235)。当 $v>0$ 时,SSDC 41B将当前物理块ID添加到垃圾收集目标块列表(步骤S240)。另一方面,当SSDC 41B在步骤S235判定不存在无效页时(步骤S235的结果为否),SSDC 41B将行号 i 递增1以选择下一物理块ID作为处理目标(步骤S245)。

[0114] SSDC 41B将变量 v 添加到可用空间的累计值 s (步骤S250)。SSDC 41B判定可用空间的累计值 s 是否达到物理块大小,即,到目前为止获取的无效页大小的总量是否达到物理块大小(步骤S255)。当可用空间的累计值 s 未达到物理块大小时,SSDC 41B将行号 i 递增1(步骤S245),然后重复步骤S210之后的处理。

[0115] 相反地,当可用空间的累计值 s 达到物理块大小时,SSDC 41B从垃圾收集目标块列表的所有物理块读取所有活动页的数据,并将读取数据存储在DRAM 40B中(步骤S260)。SSDC 41B擦除垃圾收集目标块列表的所有物理块(步骤S265)。SSDC 41B更新活动块表、可用块和活动页表(步骤S270)。也就是说,已擦除的物理块被从活动块表删除并且被添加到可用块表,从中读取数据的活动页被从活动页表删除。

[0116] SSDC 41B将DRAM 40B中存储的所有活动页的数据写入可用块(步骤S280)。SSDC 41B更新活动块表、可用块表和活动页表(步骤S290)。也就是说,经过编程的物理块被从可用块表删除,并且被添加到活动块表,经过编程的物理页被添加到活动页表。

[0117] 这样,由于NAND GC对NAND存储器16B大量地执行读取处理、删除处理和编程处理,因此需要大量时间。因此,当非易失性存储设备2B中没有足够多的可用块时,SSDC 41B执行NAND垃圾收集。然后信息处理设备3的写入速度会降低。而且,由于NAND GC的写入处理和擦除处理,NAND存储器16B的可靠性会劣化。

[0118] 步骤S265的删除处理不能在图4中的步骤S130的NAND GC期间执行,并且可以在步骤S140的写入处理之前立即执行。

[0119] <删除通知处理>

[0120] 接下来,将描述非易失性存储设备2B的删除通知处理。用于删除通知的命令实例包括INCITS ATA/ATAPI Command Set-2 (ACS-2)中描述的数据集管理命令(通常称为修剪命令)。删除通知处理是能够在删除OS 100上的LBA区域的数据时通知非易失性存储设备LBA区域(LBA范围)的处理。在该实施例中,删除通知包括LBA区域的信息,其中包括LBA和扇区计数。然后,LBA区域可被视为非易失性存储设备2B上的可用区域。非易失性存储设备2B可通过删除通知处理创建可用块。修剪命令的功能不仅可通过Data Set Management命令实现,而且例如还可通过另一命令实现,如使用ACS2中描述的SCT命令的供应商特定命令。当SATA接口之外的接口被用作接口19B时,接口固有的用于删除通知处理的命令可被用作修剪命令。例如,可使用2012年10月11日发布的NVM Express Revision 1.1 (<http://www.nvmexpress.org/>)中描述的11h Dataset Management命令的Deallocate (AD)作为修剪命令。例如,当非易失性存储设备2B具有下面描述的Deterministic Zeroing TRIM功能时,修剪命令的功能不仅可通过上述命令实现,而且还可通过NVM Express

Revision 1.1中描述的08h Write Zeroes命令实现。而且,可使用2012年12月7日发布的SCSI Block Commands-3 (SBC-3) Revision 35 (<http://www.t10.org/>) 中描述的42h UNMAP命令作为修剪命令。在该实施例中,不仅在删除数据时使用修剪命令,而且还在备份和复制数据时使用修剪命令。

[0121] 图6是示出非易失性存储设备2B的删除通知处理的流程图。非易失性存储设备2B从信息处理设备3接收删除通知(步骤S300)。删除通知包括删除命令和LBA范围信息。LBA范围信息包括LBA和扇区计数(扇区大小)。SSDC 41B从DRAM 40B读取活动页表(步骤S310)。SSDC 41B从活动页表删除其所有区域从删除通知指定的LBA区域映射的所有页(步骤S320)。这样,删除通知目标的页中的数据从设备有效数据更改为设备无效数据。

[0122] SSDC 41B从DRAM 40B读取活动块表(步骤S330)。SSDC 41B从活动块表删除在活动页表中没有物理块ID的物理块,并将物理块添加到可用块表(步骤S340)。

[0123] 这样,SSDC 41B删除由删除通知指定的LBA区域到NAND存储器16B的物理地址的映射。SSDC 41B可通过删除通知处理以及NAND GC创建可用块。从删除通知处理指定的LBA区域映射的NAND区域上的数据从设备有效数据更改为设备无效数据。删除通知指定的LBA区域的数据大小可以是这样的大小:其等于、大于或小于非易失性存储设备2B在管理表中管理的最小数据大小。

[0124] 如上所述,非易失性存储设备可通过处理从信息处理设备3接收的删除通知来创建可用块。这样,删除通知指定的LBA区域的设备有效数据更改为设备无效数据。SSDC 41B仅在删除通知处理中更改管理信息。然后,与处理从信息处理设备3发送到SSD的写入命令相比(其中管理信息的更改和针对NAND存储器的用户数据写入由SSDC 41B处理),可以以更高的速度处理删除通知处理。

[0125] <读取处理>

[0126] 接下来,将描述作为备份目的地的非易失性存储设备2B的读取处理。图7是示出非易失性存储设备2B的读取处理的流程图。在完成作为备份源的非易失性存储设备2A的备份处理后,信息处理设备3可从作为备份目的地的非易失性存储设备2B读取备份数据。在这种情况下,非易失性存储设备2B从信息处理设备3接收读取命令。读取命令包括要读取的LBA区域的信息。例如,LBA区域的信息包括LBA和扇区计数。

[0127] SSDC 41B从DRAM 40B读取活动块表(步骤S410)。SSDC 41B选择从LBA区域映射的物理块ID(步骤S420)。当物理块ID位于活动块表中时(S420:是),SSDC 41B从具有这些块ID的块读取数据。也就是说,SSDC 41B读取活动页表(步骤S430),从活动页表搜索从LBA区域映射的物理页地址,并且从NAND存储器16B的物理地址读取数据(步骤S440)。SSDC 41B将读取数据发送到信息处理设备3(步骤S450)。

[0128] 当LBA区域未映射到活动块表中的有效物理块ID时(S420:否),SSDC 41B执行读取未写入区域的处理。也就是说,SSDC 41B不从NAND存储器16B读取数据并将通过函数f计算的数据传送到信息处理设备3(步骤S460)。换言之,SSDC 41B将函数f的输出数据传送到信息处理设备3。换言之,SSDC 41B将函数f指示的输出数据传送到信息处理设备3。数据的数据长度等于LBA区域的大小。具体而言,当假设 A_i 是地址, D_i 是数据时,函数f指示的数据满足:

[0129] $f(A_i) = D_i$ (其中i是等于或大于0的整数)。

[0130] 可任意配置函数 f 。例如,当LBA针对函数 f 被设定为独立变量,而函数 f 被定义为:

[0131] $f(LBA) = 0$ 时,

[0132] SSDC 41B响应于来自信息处理设备3的读取请求,发送填充“0”的数据,该数据具有LBA区域的数据长度。函数 f 可存储在SSDC 41B中,可存储在NAND存储器16B中,或者可存储在RAM 40B中。为了提高未写入区域的读取处理速度,优选地将函数 f 存储在SSDC 41B或RAM 40B中。在非易失性存储设备2B被断电期间,可将函数 f 存储在NAND存储器16B中,并且当非易失性存储设备2B被通电时,SSDC 41B可将函数 f 从NAND存储器16B读取到RAM 40B。

[0133] 在未写入区域的读取处理中,SSDC 41B不从NAND存储器16B读取数据,或者基本上从NAND存储器16B读取很少的数据。因此,与SSDC 41B读取活动页表和NAND存储器16B两者的写入区域读取处理相比,未写入区域的读取处理以更高的速度执行。

[0134] 这样,当信息处理设备3读取作为备份目的地的非易失性存储设备2B中的包括设备无效数据的LBA区域时,SSDC 41B可能不读取非易失性存储设备2B的NAND存储器16B。结果,与设备有效数据相比,SSDC 41B可以以更高的速度读取设备无效数据。当读取具有设备无效数据的LBA区域时,从非易失性存储设备2B传送到信息处理设备3的数据可以是任何数据。但是在该实施例中,对于每个LBA,要求所述数据至少是确定性数据。此外,当非易失性存储设备2B被激活或连接到信息处理设备3时,为了更清楚地将要备份的数据通知给信息处理设备3,非易失性存储设备2B优选地通知信息处理设备3该非易失性存储设备2B的设备无效数据的读取数据具有确定性。例如,如在非专利文献中的ACS-2中描述的那样,IDENTIFY DEVICE数据字169的位14被设定为1。这样,非易失性存储设备2B可提前通知信息处理设备3设备无效数据的值是确定性的。这样,在本发明中,至少针对每个LBA是确定性的数据使用函数 $f(LBA)$ 表示为如下:

[0135] $data = f(LBA)$

[0136] 具体而言,当 $f(LBA) = 0$ 时,IDENTIFY DEVICE数据字169的位5被设定为1,这样,非易失性存储设备2B可提前通知信息处理设备3设备无效数据的读取数据为固定值0(非易失性存储设备2B支持Deterministic Zeroing TRIM特性)。当 $f(LBA) = 0$ 时,具体而言,当逻辑扇区大小为512字节时,函数 f 输出长度为512字节的填充0的数据,例如“ $f(LBA) = 00000...0$ ”(0-fill)。

[0137] [存储控制程序的方法]

[0138] 接下来,将描述在信息处理设备3上执行的存储控制程序(控制工具)200的方法。例如,如图8所示,控制工具200在信息处理设备3断电期间被存储在作为备份源的非易失性存储设备2A的NAND存储器16A中。但是,当信息处理设备3断电或者当OS 100接收激活控制程序200的消息时,CPU 4将控制工具200从NAND存储器16加载到主存储器6上的区域6B。另一方面,当不同于非易失性存储设备2A和非易失性存储设备2B的非易失性存储设备20连接到信息处理设备3时,如图9所示,控制程序200可被存储在非易失性存储设备20的区域20B中。当信息处理设备3被激活或者程序被激活时,控制程序200可被从区域20B加载到区域6B。具体而言,当非易失性存储设备20被用作存储OS的系统驱动器,并且非易失性存储设备2A被用作存储用户数据(例如文档、静止图像数据或移动图像数据)的数据驱动器时,控制程序优选地被存储在非易失性存储设备20中,从驱动器2A和20的作用被清晰地区分(以便系统驱动器20被用作主要存储OS或应用程序的驱动器,数据驱动器2A被用作存储用户数据

的驱动器)这一事实的角度来看,非易失性存储设备20充当系统驱动器。备选地,当USB存储器被用作存储控制程序200的驱动器20时,USB存储器可在正常使用时断开与信息处理设备3的连接。仅当控制程序被使用时,例如非易失性存储设备2A备份时,USB存储器才可连接到信息处理设备3。控制程序可被存储在作为备份目的地的非易失性存储设备2B中。

[0139] 从省略用户设置控制程序的工作这一事实的角度来看,例如,如图8或9所示,信息处理系统1优选地在这样的状态(预安装状态)下发运:其中控制程序被存储在非易失性存储设备2A或非易失性存储设备20中,这样,信息处理系统1便可在商店出售或者交付到用户手上。另一方面,从用户可选择安装或不安装控制程序这一事实的角度来看,以及从最新控制程序被提供给用户的这一事实的角度来看,控制程序200优选地从WEB下载,或者从诸如DVD-ROM或USB存储器之类的外部存储介质安装并且被存储在非易失性存储设备2A或非易失性存储设备20中。

[0140] 图10是示出从WEB下载的实例的框图。控制程序200被存储在Web服务器21的存储介质22的区域22B中,因此,控制程序200例如经由诸如因特网、局域网或无线LAN之类的网络传送通过与LAN控制器12或USB控制器12相连的网络设备,并且被下载到(或安装到)NAND存储器16B。在图9的情况下,控制工具200被下载到或安装到非易失性存储设备20上的区域20B。

[0141] 图11是示出从诸如DVD-ROM或CD-ROM之类的光介质安装的实例的框图。控制工具200被存储在诸如DVD-ROM或CD-ROM之类的光介质23的区域23B中,因此,光介质23被放入光驱10,以便经由光驱10在非易失性存储设备2A的NAND存储器16A上安装控制工具200。图12是示出从USB存储器24安装的实例的框图。控制工具200被存储在USB存储器24的区域24B中,因此,USB存储器24连接到USB控制器13,以便经由USB控制器13在非易失性存储设备2A的NAND存储器16A上安装控制工具200。除了USB存储器24之外,也可使用诸如SD卡之类的外部存储器。从方便用户购买的角度来看,光介质23或USB存储器24优选地进行包装并在发运SSD 2时作为SSD 2的附件一起销售。另一方面,光介质23或USB存储器24可作为软件产品单独销售,或者作为杂志或书籍的附赠品提供。

[0142] 图13是示出信息处理系统1的软件层处的分级结构的示意图。诸如被加载到主存储器6的控制程序200之类的各种软件一般不直接与非易失性存储设备2A或非易失性存储设备2B通信,而是经由被加载到主存储器6的OS 100与非易失性存储设备2A或非易失性存储设备2B通信。当各种软件需要将诸如读取请求或写入请求之类的命令传送到非易失性存储设备2A时,各种软件将以文件为单元的访问请求传送到OS 100。OS 100参考OS 100中嵌入的文件管理表(OS文件系统)来指定与访问请求所针对的文件对应的非易失性存储设备2A或非易失性存储设备2B的逻辑地址(LBA),将逻辑地址转换成对应的接口固有命令(非专利文1的ACS-2中描述的60h READ FPDMA QUEUED命令或61h WRITE FPDMA QUEUED命令),然后将接口固有命令传送到非易失性存储设备2A或非易失性存储设备2B。

[0143] 当OS 100从非易失性存储设备2A或非易失性存储设备2B接收回复时,OS 100指定此回复所针对的软件,并且将回复返回给所指定的软件。此外,控制程序200可被配置为在不使用OS 100的文件管理表的情况下,直接经由OS 100访问非易失性存储设备2A或非易失性存储设备2B的LBA,或者可被配置为在不使用经由OS 100的路径的情况下,直接访问非易失性存储设备2A或非易失性存储设备2B的LBA。

[0144] [信息处理设备的物理配置]

[0145] 接下来,将描述信息处理系统1的配置实例。信息处理系统1例如可被实现为桌面计算机或笔记本型便携式计算机。图14是示出充当信息处理系统1的桌面计算机的示意图。

[0146] 桌面计算机包括计算机主体31、显示器9、键盘14和鼠标15。计算机主体31包括主板30,上面安装有主硬件单元;非易失性存储设备2A和电源设备32。非易失性存储设备2A经由SATA电缆物理连接到主体30,并且经由安装在主板30上的南桥7电连接到安装在同一主板30上的CPU4。电源设备30产生在桌面计算机中使用的各种电力并且经由电缆将电力提供给主板30、非易失性存储设备2A等。

[0147] 图15是示出充当信息处理系统1的便携式计算机的示意图。便携式计算机包括计算机主体34和显示单元35。例如,包括LCD(液晶显示屏)的显示设备9被嵌入显示单元35。安装之后的显示单元35可在打开位置与闭合位置之间旋转,在打开位置上,计算机主体34的上表面暴露出来,在闭合位置上,主体34的上表面合住。主体34包括薄的箱体式外壳。电源开关36、键盘14和触摸板33等被设置在上表面。主体34包括非易失性存储设备2A、主板和电源设备,这与桌面计算机相同。

[0148] 应用该实施例的信息处理系统1不仅可以是上述信息处理系统,而且还可以是诸如照相机或是摄像机之类的成像设备、平板计算机、智能电话、游戏设备、汽车导航系统、存储驱动器复印机等。

[0149] 在该实施例中,非易失性存储设备2B被用作作为备份源的非易失性存储设备2A的备份目的地。从信息处理系统1的轻型、省电和省空间的角度来看,在存储设备2A的可靠性不劣化的同时,非易失性存储设备2B 优选地不连接到信息处理设备3。在这种情况下,当非易失性存储设备2A的可靠性劣化时,优选地通过显示器9,从控制程序向用户显示消息“PLEASE CONNECT A BACKUP STORAGE DEVICE (请连接备份存储设备)”,以使用户可将备份非易失性存储设备2B连接到信息处理设备3。例如,当信息处理系统1为桌面计算机时,非易失性存储设备2B可经由SATA电缆或USB电缆连接到主板30,例如,如图14所示。非易失性存储设备2B可以安装在计算机主体31的内部。此外,当信息处理系统1为便携式计算机时,非易失性存储设备2B可经由SATA电缆或USB电缆连接到计算机主体34,例如,如图15所示。在这种情况下,从简化拔插的角度来看,接口19B优选地为eSATA接口。非易失性存储设备2B可以安装在计算机主体34的内部。在安装非易失性存储设备2B之后,控制程序200执行将用户数据从非易失性存储设备2A备份到非易失性存储设备2B。

[0150] 从信息处理系统1的省电、轻型和省空间的角度来看,当控制程序200执行的数据备份结束时,如图16和17所示,优选地将非易失性存储设备2A与信息处理系统1分离,将非易失性存储设备2B安装到非易失性存储设备2A的安装位置,并且用非易失性存储设备2B替换非易失性存储设备2A。

[0151] [备份处理]

[0152] 图18是示出控制程序200执行的备份处理的处理流程。控制程序200被CPU 4加载到控制程序区域6B,并且控制程序200的功能由CPU 4处理。控制程序200是具有从非易失性存储设备2A到非易失性存储设备2B的数据备份功能的程序。控制程序200的备份功能可以由用户通过键盘14或鼠标15手动启动,也可由CPU 4自动启动。在图18所示的流程中,备份目标LBA区域被分成长度为L的小区域,并且针对每个小区域执行备份处理的子例程。可使

用作为LBA指针的cLBA作为小区域的引用,并且通过将cLBA从0开始递增L来备份所有LBA区域的数据。小区域的长度L是等于或大于1的整数。

[0153] 当备份功能启动时,控制程序200将作为数据复制目标LBA区域的指针的cLBA设定为0(步骤S500)。控制程序200将读取LBA区域(从cLBA开始,长度为L)的数据的命令传送到非易失性存储设备2A(步骤S510)。例如,使用非专利文献1的ACS-2中描述的60h READ FPDMA QUEUED命令、25h READ DMA EXT命令等作为读取命令。接收读取命令的存储设备2A将LBA区域的读取数据(D)传送到信息处理设备3,并且信息处理设备3接收数据D(步骤S520)。控制程序将数据D与主存储器6中存储的函数f(cLBA)的输出数据相比较(步骤S530)。主存储器6中的函数f()与存储在非易失性存储设备2B中的函数f()相同。在该实施例中,当CPU 4加载控制工具200时,CPU 4将函数f()从非易失性存储设备2B加载到主存储器6。处理分支为数据D与函数f(cLBA)的输出数据不同(步骤S530的结果为否)的情况,以及数据D与函数f(cLBA)的输出数据相同(步骤S530的结果为是)的情况。

[0154] 当D与f(cLBA)的输出数据不同时(步骤S530的结果为否),控制程序200将使得数据D被写入LBA区域(从cLBA开始,长度为L)的命令传送到非易失性存储设备2B,控制程序200将数据D传送到非易失性存储设备2B,以及SSDC 41B将数据D写入NAND存储器16B(步骤S540)。例如,使用非专利文献1的ACS-2中描述的61h WRITE FPDMA QUEUED命令、35h WRITE DMA EXT命令等作为写入命令。

[0155] 当D与f(cLBA)的输出数据相同时(步骤S530的结果为是),控制程序200将取消映射LBA区域(从cLBA开始,长度为L)的删除通知传送到非易失性存储设备2B。非易失性存储设备2B的SSDC 41B更改管理信息46B以删除(使失效)LBA区域到NAND存储器45B的映射,以使得LBA区域的数据变为设备无效数据(步骤S550)。例如,使用非专利文献1的ACS-2中描述的06h DATA SET MANAGEMENT命令或SCT命令,NVM Express Revision 1.1中描述的11h Dataset Management命令的Deallocate(AD)等作为删除通知命令。

[0156] 当步骤S540或步骤S550完成时,控制程序200使cLBA增加L(步骤S560),然后针对所有小区域执行步骤S500到步骤S550的子例程。当cLBA超过非易失性存储设备2B的最大可访问LBA时,假设所有LBA区域的读取完成(步骤S570的结果为是),这样,备份处理便完成。

[0157] 当控制程序200完成备份处理时,控制程序200优选地强制非易失性存储设备2B转变为待机状态,例如通过将待机状态和睡眠状态转变命令(例如非专利文献1的ACS-2中描述的E0h Standby Immediate命令或E6h SLEEP命令)发出到非易失性存储设备2B实现此操作。结果,在步骤S540被写入非易失性存储设备2B的数据被可靠地存储在NAND型闪存中,并且在步骤S550中被删除通知处理更改的最新管理信息46B被可靠地存储在NAND型闪存中,以防止在非易失性存储设备2B意外停止工作时丢失数据。信息处理系统1的功耗在非易失性存储设备2B转变为待机状态之后得到节省。控制程序200可将清仓高速缓存命令(例如非专利文献1的ACS-2中描述的E7h FLUSH命令或EAh FLUSH CACHE EXT命令)发出到非易失性存储设备2B,从而将RAM 40B的管理信息46B和CM 44B的数据保存和存储在NAND型闪存16B中。

[0158] 当控制程序200完成备份处理时,控制程序200可强制存储设备2A转换为省电模式,以便节省信息处理系统1的功耗。例如,控制程序可将信息技术ATA/ATAPI Command Set-3(ACS-3)中描述的E0h Standby Immediate命令、E6h SLEEP命令等发出到存储设备

2A。例如,控制程序200可关闭非易失性存储设备2A的电源。例如,控制程序200可强制非易失性存储设备2A转变为Serial ATA Revision 3.1 Gold Revision中描述的Partial状态或Slumber状态。例如,控制程序200可通过激活“Serial ATA Technical Proposal: SATA31_TPR_C 108 Title:Devic Sleep”中描述的DEVSLP信号,强制非易失性存储设备2A转变为DevSleep状态。例如,控制程序200可强制非易失性存储设备2A转变为PCI Express Base Specification Revision 3.0中描述的D1状态、D2状态或D3状态。例如,控制程序200可强制非易失性存储设备2A转变为PCI Express Base Specification Revision 3.0中描述的L1状态、L2状态或L3状态。

[0159] 如上所述,与非易失性存储设备2B的写入处理相比,针对非易失性存储设备2B的删除通知处理可以以更高的速度执行。此外,对非易失性存储设备2B的写入处理一般使用可用块,但是对非易失性存储设备2B的删除通知处理很少使用可用块。因此,该实施例提高了非易失性存储设备2B的处理速度。此外,NAND存储器块一般通过对非易失性存储设备2B的写入处理擦除,这样,NAND存储器的可靠性劣化。但是,由于删除NAND存储器块的处理很少在对非易失性存储设备2B的删除通知处理中执行,因此,该实施例可减轻NAND存储器的可靠性劣化。

[0160] 在一个比较例中,由于步骤S540的写入处理针对备份目的地目标的每个LBA区域执行,因此备份速度很低,非易失性存储设备2B的处理速度降低,并且NAND存储器16B的可靠性劣化。另一方面,当使用根据该实施例的控制工具200执行备份时,不是针对与函数f(LBA)的输出数据相同的数据执行步骤S540,而是执行步骤S550。因此,与比较例的备份系统相比,可以以更高的速度执行备份,可抑制非易失性存储设备2B的处理速度降低,从而可抑制NAND存储器16B的可靠性劣化。

[0161] 在该实施例中,SSD被用作非易失性存储设备2B,但是也可使用其它非易失性存储设备(例如硬盘驱动器(HDD)和混合HDD)作为非易失性存储设备2B。例如,当HDD被用作非易失性存储设备2B时,到非易失性存储设备2B的写入量在步骤S550减少,这样,非易失性存储设备2B的可靠性劣化减少。这样,从信息处理设备3传送到非易失性存储设备2B的数据量减少,备份处理可以以更高的速度执行。

[0162] 在该实施例中,SSD被用作非易失性存储设备2A,但是也可使用其它存储设备(例如HDD、混合HDD、易失性RAM等)作为存储设备2A。

[0163] 在该实施例中,描述了备份非易失性存储设备2A的所有LBA的情况。但是,备份目标LBA区域可限于所有LBA区域中的一部分。

[0164] 在该实施例中,不同的存储设备被用作作为备份源的非易失性存储设备2A以及作为备份目的地的非易失性存储设备2B。但是作为备份源的非易失性存储设备2A和作为备份目的地的非易失性存储设备2B可以是同一SSD(2A=2B)。也就是说,例如,SSD被分成两个分区,分区A(起始LBA=0)和分区B(起始LBA=X),并且控制程序200可执行从分区A到分区B的备份处理。在这种情况下,例如,备份处理可在从cLBA开始的备份源LBA区域与从X+cLBA开始的备份目的地LBS区域之间执行。

[0165] (第二实施例)

[0166] 在第一实施例中,已经描述了对具有固定扇区长度L的每个逻辑分区(作为备份处理单元)执行备份处理的情况。备份处理可针对具有可变长度的每个分区执行。例如,在OS

100上管理的文件单元可被用作备份处理单元。图19是示出根据第二实施例的控制工具200的处理实例的流程图。

[0167] 控制程序200读取存储在作为备份源的非易失性存储设备2A中的NAND存储器16A内的OS文件管理表(元数据、OS文件系统),并将存储在非易失性存储设备2A中的文件的所有信息存储在主存储器6中作为备份目标列表(步骤S600)。备份目标列表可能不是非易失性存储设备2A的所有文件,而可以是这些文件的一部分。控制程序200搜索其中每个文件的数据位于备份目标列表中的LBA区域的信息(步骤S605)。控制程序200从作为备份源的非易失性存储设备2A读取LBA区域的数据(步骤S610)。例如,使用非专利文献1的ACS-2中描述的60h READ FPDMA QUEUED命令、25h READ DMA EXT命令等作为读取命令。

[0168] 在判定步骤S610中读取的数据之前,优选地将文件的读取数据以逻辑单元分割为数据D(步骤S620)。逻辑单元可以是逻辑扇区或多个逻辑扇区。控制程序200针对文件中的每个逻辑单元处理步骤S630-660。cLBA被用作选定逻辑单元的指针,最初等于LBA区域的起始LBA,当在步骤S660之后处理步骤S630时,该指针递增。控制程序200选择一个从cLBA开始的逻辑单元(步骤S630),并且判定数据D是否与函数f(cLBA)的输出数据相同(步骤S640)。当数据D与函数f(cLBA)的输出数据不同时,控制程序200使用写入命令(例如非专利文献1的ACS-2中描述的61h WRITE FPDMA QUEUED命令、35h WRITE DMA EXT命令等)将数据D写入作为备份目的地的非易失性存储设备2B(步骤S645)。当数据D与函数f(cLBA)的输出数据相同时,控制程序200不向作为备份目的地的非易失性存储设备2B发送写入命令,而是发送删除通知命令以删除逻辑单元到NAND存储器16B的映射(步骤S650)。例如,可使用非专利文献1的ACS-2中描述的06h DATA SET MANAGEMENT命令或SCT命令,NVM Express Revision 1.1中描述的11h Dataset Management命令的Deallocate(AD)等作为删除通知命令。控制程序200针对所有逻辑单元重复步骤S630到步骤S650的处理。

[0169] 控制程序200可更改非易失性存储设备2B中的OS文件管理表(元数据)以将选定文件的信息更新为备份目的地,这样OS 6A便可识别备份文件。控制程序200判定控制程序200是否已完成所有文件的处理(步骤S680)。当控制程序200未完成所有文件时,控制程序200选择下一文件,获取选定文件的LBA信息(步骤S690),然后将处理移到步骤S610。当控制程序200完成所有文件时,处理随之结束。在该实施例中,已经描述了备份非易失性存储设备2A的所有文件的情况。但是,备份目标文件可能限于所有文件当中的部分文件。

[0170] 这样,在第二实施例中,读取备份源的数据,并且以具有可变数据大小的文件单元执行写入。在该实施例中,可针对不存储文件数据的LBA区域跳过读取、写入和删除通知处理。因此,备份处理可以更高的速度执行。

[0171] (第三实施例)

[0172] 在第一和第二实施例中,在控制程序200中使用函数f(LBA)执行备份处理。作为备份目的地的非易失性存储设备2B的SSDC 41B通过计算非易失性存储设备2B中的函数f(LBA)恢复要读取的LBA区域的数据,然后在读取备份后的数据时将恢复后的数据发送到信息处理设备3。

[0173] 在第三实施例中,将描述在设计控制程序200时,在设计产品时,在工厂发货之前,或者在工厂发货之后,实现函数f(LBA)的方法。具体而言,当确定要在SSD中安装的OS或系统配置时,最适合系统的函数f(LBA)优选地设计控制程序200时,在设计产品时,或者在

工厂发货之前进行配置。例如,函数f (LBA) 可作为将LBA映射到数据的表数据被存储在SSDC 41B、NAND存储器16B、或存储设备2B的RAM 40B、或信息处理设备3的主存储器6中。函数f (LBA) 可被实现为软件数据或固件数据,通过在SSDC 41B、NAND存储器16B、或存储设备2B的RAM 40B、或信息处理设备3的主存储器6中实现为算法的函数f (LBA) 来实现这些软件数据或固件数据。例如,函数f (LBA) 可嵌入在存储设备2B的SSDC 41B上,作为使用函数f (LBA) 作为算法而实现的电子电路。

[0174] 备选地,例如,来自函数f (LBA) 的输出数据可以作为函数f (LBA) 的信息存储在非易失性存储设备3中作为修剪数据存储单元的存储单元内。修剪数据存储单元可以是NAND型闪存和RAM。在这种情况下,当SSDC 41B从信息处理设备3接收删除通知时,SSDC 41B更改管理信息46B以删除由删除通知指定的LBA区域到NAND存储器16B的映射,并且SSDC 41B更改管理信息46B以将LBA区域映射到修剪数据存储单元。当信息处理设备3将从LBA区域读取数据的读取命令发送到非易失性存储设备2B时,SSDC 41B读取管理信息46B,读取从LBA区域映射的修剪数据存储单元中的数据,并且将读取数据发送到信息处理设备3。

[0175] 当函数f (LBA) 被存储在存储设备2B的NAND存储器16B,以及信息处理设备3的主存储器6的控制程序200中时,控制程序200中的函数f (LBA) 需要与非易失性存储设备2B中的函数f (LBA) 相同。例如,控制程序200和非易失性存储设备2B可以分别使用同一函数f (LBA) 来设计。备选地,例如,函数f (LBA) 的最新信息可仅被存储在非易失性存储设备2B中。在这种情况下,控制程序200将函数f (LBA) 的信息从非易失性存储设备2B加载到主存储器6。当初始函数f (LBA) 或之前的函数f (LBA) 被存储在主存储器6中时,控制程序200使用主存储器6中已加载的函数f (LBA) 覆盖初始或之前的函数f (LBA)。相反地,SSDC 41B可将函数f (LBA) 的信息从控制程序200加载到NAND型闪存16B或RAM 40B,并且可使用已加载的函数f (LBA) 覆盖之前的函数f (LBA)。

[0176] 图20是示出逻辑地址和数据的实例的示意图。信息处理设备3使用逻辑地址 (LBA) 指定SSD的逻辑地址区域。在逻辑地址中,一个逻辑扇区是最小管理单元,并且一个逻辑扇区例如具有512字节。在图20中,逻辑地址 (0000h、0001h、0002h、...、ba02h、...) 被分配给每个扇区。作为数字后缀的“h”表示这些数字以十六进制表示。

[0177] 在图20的实例中,f (LBA) 被配置为假设 (模拟) 最终用户可备份的典型数据模型。当非易失性存储设备2B和控制程序200的设计者设计非易失性存储设备2B和控制程序200时,他或她可通过基于表示典型使用环境中典型用户数据的模型分析所有逻辑地址,然后比较逻辑地址的数据与0-fill扇区,来计算图21所示的0-fill (0填充) 扇区的存在概率分布 (直方图)。在比较时,0-fill扇区的确切比率可通过用扇区 (与0-fill扇区相同) 的总计数除以所有逻辑地址区域的总LBA容量进行计算。当配置“函数f (LBA) = 000...0h (固定值和0以512字节重复)”时,在备份期间,可通过取消映射逻辑扇区 (0-fill扇区) 的总扇区计数大小的LBA区域来最大程度上节省可用块。在这种情况下,例如,函数f (LBA) 的信息可作为512字节的数据000...0h、将所有LBA映射到512字节的数据000...0h的表数据,或软件数据或固件数据 (被实现为针对所有LBA产生512字节的数据000...0h的算法),被存储在SSDC 41B、NAND存储器16B、或存储设备2B的RAM 40B、或信息处理设备3的主存储器6中。

[0178] 图22是示出其中函数f (LBA) 不输出恒定的固定值的实例的示意图。在某些情况下,常用的典型数据模式可根据逻辑地址区域而有所不同,如这样的情况:其中给定的数据

模式经常在给定逻辑地址区域中使用,另一数据模式经常在其它给定的逻辑地址区域中使用。如图22所示,考虑这样一种情况:0-fill数据经常在逻辑地址区域的前半部分中使用,1-fill数据(所有位都是“1”的数据,以及当以十六进制表示时,通过Fh填充的数据)经常在逻辑地址区域的后半部分中使用。在图23中示出0-fill扇区的存在概率分布和1-fill扇区的存在概率分布。在这种情况下,函数 $f(LBA)$ 优选地在两个逻辑地址区域中输出不同的值,如在“函数 $f(LBA) = 000 \dots 0h$ (其中 $LBA \leq c000h$)和函数 $f(LBA) = fff \dots fh$ (其中 $LBA > c000h$)”中那样,而不是“函数 $f(LBA) = 000 \dots 0h$ (固定值)”和“函数 $f(LBA) = fff \dots fh$ (固定值)”。结果,备份后的可用块数得到节省。在这种情况下,例如,函数 $f(LBA)$ 作为表数据(将小于 $c000h$ 的LBA映射到512字节的数据 $000 \dots 0h$,并且将等于或大于 $c000h$ 的LBA映射到数据 $fff \dots fh$)被存储在SSDC 41B、NAND存储器16B、或存储设备2B的RAM 40B、或信息处理设备3的主存储器6中。函数 $f(LBA)$ 作为算法、软件数据或固件数据(产生512字节到小于 $c000h$ 的LBA的数据 $000 \dots 0h$,并且产生512字节到等于或大于 $c000h$ 的LBA的数据 $fff \dots fh$)被存储在SSDC 41B、NAND存储器16B、或存储设备2B的RAM 40B、或信息处理设备3的主存储器6中。

[0179] 图24是示出三个逻辑地址区域中常用的数据不同的示意图。图25示出0-fill扇区的存在概率分布,320a-fill扇区(一个扇区长度通过重复的320a数据填充的数据)的存在概率分布,以及1-fill扇区的存在概率分布。例如,在这种情况下,函数 $f(LBA)$ 优选地被配置为在三个逻辑地址区域中具有不同的值,如在“函数 $f(LBA) = 000 \dots 0h$ (其中 $LBA \leq 7fffh$)、函数 $f(LBA) = 320a320a \dots 320ah$ (其中 $8000h \leq LBA \leq c000h$),以及函数 $f(LBA) = fff \dots fh$ (其中 $LBA \geq c001h$)”中那样。结果,备份后的可用块数得到节省。在这种情况下,例如,函数 $f(LBA)$ 作为表数据(将“ $LBA \leq 7fffh$ ”范围中的LBA映射到512字节的数据 $000 \dots 0h$,将“ $8000h \leq LBA \leq c000h$ ”范围中的LBA映射到512字节的数据 $320a320a \dots 320ah$,并且将“ $LBA \geq c001h$ ”范围中的LBA映射到512字节的数据 $fff \dots fh$)被存储在SSDC 41B、NAND存储器16B、或存储设备2B的RAM 40B、或信息处理设备3的主存储器6中。函数 $f(LBA)$ 可作为算法、软件数据或固件数据(产生512字节到“ $LBA \leq 7fffh$ ”范围中的LBA的数据 $000 \dots 0h$,产生512字节到“ $8000h \leq LBA \leq c000h$ ”范围中的LBA的数据 $320a320a \dots 320ah$,并且产生512字节到“ $LBA \geq c001h$ ”范围中的LBA的数据 $fff \dots fh$)被存储在SSDC 41B、NAND存储器16B、或存储设备2B的RAM 40B、或信息处理设备3的主存储器6中。

[0180] 函数 $f(LBA)$ 可被配置为在通过针对每个大小L分割逻辑地址区域获取的每个管理单元中具有不同的值。大小L优选地等于或大于LBA的最小单元。函数 $f(LBA)$ 在管理单元中输出固定值,但是函数 $f(LBA)$ 可在不同的管理单元之间输出不同的数据。例如,函数 $f(LBA)$ 在第一管理单元区域中输出D0(固定值),函数 $f(LBA)$ 在第二管理单元区域中输出D1(固定值)等等。在这种情况下,函数 $f(LBA)$ 被表示为如下:

[0181] “ $f(LBA) = D0$ (其中 $0 \leq LBA \leq L-1$)、 $f(LBA) = D1$ (其中 $L \leq LBA \leq 2L-1$)、 $f(LBA) = D2$ (其中 $2L \leq LBA \leq 3L-1$)、 $\dots$ 、 Dk (其中 $kL \leq LBA \leq (k+1)L-1$)、 $\dots$ ”。

[0182] Dk (其中 $k = 0, 1, \dots$) 可以是大小为L的任意固定数据,例如8d892a8fd91423...987ah和0123456789abcdef0123...defh。在这种情况下,例如,函数 $f(LBA)$ 可作为表数据(针对每个 $k = 0, 1, \dots$,将“ $kL \leq LBA \leq (k+1)L-1$ ”范围中的LBA映射到512字节的数据 Dk)被存储在SSDC 41B、NAND存储器16B、或存储设备2B的RAM 40B、或信息处

理设备3的主存储器6中。备选地,例如,函数 $f(LBA)$ 可作为算法、软件数据或固件数据(针对每个 $k=0,1,\dots$,产生512字节到“ $kL \leq LBA \leq (k+1)L-1$ ”范围中的LBA的数据 D_k)被存储在SSDC 41B、NAND存储器16B、或存储设备2B的RAM 40B、或信息处理设备3的主存储器6中。

[0183] D_k 可针对每个 $k=0,1,\dots$ 进行任意配置。最优函数 $f(LBA)$ 优选地通过选择最优 D_k 来配置。例如,非易失性存储设备2B和控制程序200的设计者可基于各种 D_k 候选的评分选择最优函数 $f(LBA)$ 。评分通过以下应用于最终用户可执行备份的典型用户数据模型的方法来计算。

[0184] (1) 设计者读取模型的每个扇区的数据(或模型的多个扇区的数据,或者模型的每个管理单元的数据)。

[0185] (2) 设计者判定(1)中读取的数据是否与函数 $f(LBA)$ 的数据输出相同(当针对扇区的每个整数倍读取数据时,设计者判定数据中的所有扇区是否与函数 $f(LBA)$ 的数据输出相同)

[0186] (3) 当数据与函数 $f(LBA)$ 的数据输出相同时,设计者使函数 $f(LBA)$ 的评分递增1。当数据与函数 $f(LBA)$ 的数据输出不同时,设计者不递增函数 $f(LBA)$ 的评分。

[0187] (4) 设计者计算SSD的所有逻辑地址区域的评分,并且计算评分(或者代表性地,只有部分逻辑地址区域可进行分级)。

[0188] 然后,设计者可选择最优 D_k ,其中最高评分被设定为最优 D_k ,并且使用该 D_k 配置最优函数 $f(LBA)$ 。例如,当“ $L=128\text{ MiB}$ (其中 $\text{MiB}=1,048,576$ 字节)”时,其中SSD的所有逻辑地址区域从 $LBA=0$ 开始,均等地按照128 MiB长度分区,设计者读取每个分区并判定每个分区的数据是否与函数 $f(LBA)$ 的输出数据相同。当数据与函数 $f(LBA)$ 相同时,设计者递增 $f(LBA)$ 的评分。例如,设计者可计算 D_k 阵列(例如“($D_0, D_1, \dots$) = ($0000\dots h, 0000\dots h, h\dots$)、($ffff\dots h, ffff\dots h, \dots$)、($ffff\dots h, 0000\dots h, \dots$)、($0000\dots h, ffff\dots h, \dots$)”)的多个候选的评分,选择具有最高评分的最优候选作为最优($D_0, D_1, \dots$),并且使用最优($D_0, D_1, \dots$)配置最优函数 $f(LBA)$ 。作为最优函数 $f(LBA)$ 的信息的最优 D_k 阵列(其中 $k=0,1,\dots$)优选地被存储在SSDC 41B、RAM 40B或NAND存储器16B中。最优 D_k 阵列(其中 $k=0,1,\dots$)可能未被非易失性地存储在控制程序200中,当需要时,控制程序200可将 D_k 阵列从非易失性存储设备2B读取到主存储器6。例如,可使用非专利文献1的ACS-2中描述的ECh IDENTIFY DEVICE命令读取最优 D_k 阵列,并且SSDC 41B可输出IDENTIFY DEVICE数据,其中最优化 D_k 阵列被存储在指定为供应商特定字段的Word129字段。

[0189] 分区可通过按照可变长度分割逻辑地址区域而获取,并且可被用作管理单元。在这种情况下,函数 $f(LBA)$ 被表示为如下。

[0190] “ $f(LBA) = D_0$ (其中 $L_0=0 \leq LBA \leq L_1-1$)、 $f(LBA) = D_1$ (其中 $L_1 \leq LBA \leq L_2-1$)、 $f(LBA) = D_2$ (其中 $L_2 \leq LBA \leq L_3-1$)、 $\dots$ 、 D_k (其中 $L_k \leq LBA \leq L_{k+1}-1$)、 $\dots$ ”。

[0191] D_k (其中 $k=0,1,\dots$) 是大小为L的任何数据模式。满足“ $L_0 < L_1 < L_2 < \dots < L_k < \dots$ ”。在这种情况下,函数 $f(LBA)$ 作为表数据(针对每个 $k=0,1,\dots$ 将“ $L_k \leq LBA \leq L_{k+1}-1$ ”范围中的LBA映射到512字节的数据 D_k)被存储在SSDC 41B、NAND存储器16B、或存储设备2B的RAM 40B、或信息处理设备3的主存储器6中。备选地,例如函数 $f(LBA)$ 作为算法、软件数据或固件数据(针对每个 $k=0,1,\dots$ 产生512字节到“ $L_k \leq LBA \leq L_{k+1}-1$ ”范围中的LBA的数据 D_k)被存储在SSDC 41B、NAND存储器16B、或存储设备2B的RAM 40B、或信息处理设备3的主存储器6

中。

[0192] 图26示出因为没有太多逻辑扇区使用小固定数据填充,所以函数 $f(LBA)$ 无法充分优化的情况。在这种情况下,将固定数据输出到分区(通过按照固定长度或可变长度分割逻辑地址区域而获取)的函数可能不是最优函数。在这种情况下,数据和LBA具有关联关系,并且设计者优选地在较一般的函数 $f(LBA)$ 候选当中选择最优函数 $f(LBA)$ 。例如,设计者提前准备多个函数 $f(LBA)$ 候选,计算应用到典型用户数据模型(其中典型的最终用户执行备份)的每个候选的评分,然后选择具有最高评分的函数 $f(LBA)$ 作为最优函数 $f(LBA)$ 。在这种情况下,如图27所示,其中数据与LBA具有固定关联的逻辑扇区通过借助最优函数 $f(LBA)$ 的备份处理取消映射。

[0193] 例如,设计者可配置具有不同的 a 和 b 值的“函数 $f(LBA) = a \times LBA + b$ ”定义的候选。设计者基于一般用户数据模型(其中最终用户执行备份)计算每个参数集(a, b)的评分,从候选当中选择具有最高评分的最优(a, b),并且以最优(a, b)配置最优函数 $f(LBA)$ 。图28是示出选择最优函数 $f(LBA)$ 的方法的示意图。图28示出其参数 b 分别为 b_1, b_2 和 b_3 ,并且其参数 a 是诸如 a_0 之类的任意值的候选。在这种情况下,“ $(a, b) = (a_0, b_2)$ ”的评分最高。因此,设计者可选择“函数 $f(LBA) = a_0 \times LBA + b_2$ ”作为最优函数 $f(LBA)$ 。例如,函数 $f(LBA)$ 可作为其中实现函数 $f(LBA)$ 的算法、软件数据或固件数据被存储在SSDC 41B、NAND存储器16B、或存储设备2B的RAM 40B、或信息处理设备3的主存储器6中。

[0194] 当SSDC 41B支持ACS-2中描述的48位地址特性集时,可通过矩阵运算配置函数 $f(LBA)$,例如“函数 $f(LBA) = A \times LBA + b$ ”,其中位于LBA中的是48列向量阵列,例如“ $LBA = (d_0, d_1, d_2, \dots, d_{47})^T$ ”(其中 t 是转置标记), A 是 512×48 矩阵阵列, b 是4096列向量阵列(其中1扇区=512字节=4096位),以便函数 $f(LBA)$ 的输出数据以4096列向量格式计算。在这种情况下,例如, $f(LBA)$ 的输出向量阵列的第一元素表示输出数据的第一位,其第二元素表示输出数据的第二位,以此类推。更一般地说,当LBA包括 n 个位时,函数 $f(LBA)$ 可通过矩阵运算配置,例如“函数 $f(LBA) = A \times LBA + b$ ”,其中LBA是 n 列向量阵列,例如“ $LBA = (d_0, d_1, d_2, \dots, d_{n-1})^T$ ”, A 是 $n \times m$ 矩阵阵列, b 是 m 列向量阵列。在这种情况下,例如,函数 $f(LBA)$ 可作为包括矩阵 A 阵列和向量 b 阵列的阵列数据被存储在SSDC 41B、NAND存储器16B、或存储设备2B的RAM 40B、或信息处理设备3的主存储器6中。函数 $f(LBA)$ 可作为算法、软件数据或固件数据被存储在SSDC 41B、NAND存储器16B、或存储设备2B的RAM 40B、或信息处理设备3的主存储器6中。

[0195] 在该实施例中,描述了函数 $f(LBA)$ 的输出数据仅依赖于作为独立变量的LBA的情况,或函数 $f(LBA)$ 的输出数据为不依赖于LBA的固定数据的情况。但是,函数 f 可依赖于LBA以外的变量。例如,函数 $f(LBA, X)$ 可被配置为具有两个独立的变量:LBA以及由SSDC 41B指定的参数 X ,或者函数 $f(LBA, Y)$ 可被配置为具有两个独立的变量:LBA以及由信息处理设备3指定的参数 Y 。

[0196] (第四实施例)

[0197] 当函数 $f(LBA)$ 的输出为不依赖于LBA的固定数据(= Q)时,数据 Q 可被仅存储在NAND存储器16B、SSDC 41B或作为备份目的地的非易失性存储设备2B的RAM 40B中,并且不存储在控制程序200中。在第四实施例中,在控制程序200使用图18的步骤S530或图19的步骤S640比较读取数据与“ $f(LBA) = Q$ ”之前,控制程序200从非易失性存储设备2B检索固定数

据Q。例如,可使用非专利文献1的ACS-2中描述的Ech IDENTIFY DEVICE命令作为从非易失性存储设备2B检索固定数据Q的方法。例如,数据Q可以根据Word129字段(被指定为Ech IDENTIFY DEVICE命令的Identify Device输出数据中的供应商特定字段)的数据R计算。在这种情况下,Word129中的2字节长数据R和控制程序200产生以重复256次的数据R填充的512字节长的Q,如下所示:

[0198] $Q = RRRR \dots R$ (R连续重复256次)。作为函数f(LBA)的信息的数据R例如被存储在SSDC 41B、NAND存储器16B、或存储设备2B的RAM 40B中。一旦控制程序200从存储设备2B检索到数据R,控制程序便可将数据R存储在信息处理设备3的主存储器6中。

[0199] 图29是示出根据第四实施例的处理实例的流程图。控制程序200例如将检索函数f(LBA)的信息的ECH IDENTITY DEVICE命令发出到非易失性存储设备2B并从非易失性存储设备2B检索Word129字段中的数据R(步骤S700)。然后,控制程序200通过使用连续重复的数据R填充主存储器6中512字节长的区域产生数据Q(步骤S710)。

[0200] 控制程序200使用0替换cLBA作为数据复制目标LBA区域的指针(步骤S720)。控制程序200将读取LBA区域(从cLBA开始,长度为L)的数据的命令发送到非易失性存储设备2A(步骤S730)。接收读取命令的非易失性存储设备2A将LBA区域的数据D传送到信息处理设备3,然后信息处理设备3接收数据D(步骤S740)。控制程序200比较数据D与固定数据Q(步骤S750)。处理分支为不同的步骤,其中包括数据D与固定数据Q不同(步骤S750的结果为否)的情况和数据D与固定数据Q相同(步骤S750的结果为是)的情况。

[0201] 当D与Q不同时(步骤S750的结果为否),控制程序将使得数据D被写入LBA区域(从cLBA开始,长度为L)的命令传送到非易失性存储设备2B。控制程序200将数据D传送到非易失性存储设备2B,并且SSDC 41B将数据D写入NAND存储器16B(步骤S760)。

[0202] 当D与Q相同时(步骤S750的结果为是),控制程序200将删除LBA区域(从cLBA开始,长度为L)到NAND存储器16B的映射的删除通知发送到非易失性存储设备2B。SSDC 41B更改管理信息46B以删除映射并且数据D变为设备无效数据(步骤S770)。

[0203] 当步骤S760或步骤S770完成时,控制程序200使cLBA递增L(步骤S780),并且重复执行步骤S730到步骤S780的处理,直到非易失性存储设备2A的所有LBA区域的数据均被备份(步骤S790的结果为是)。

[0204] 图30是示出根据第四实施例的另一处理实例的流程图。在图30中的处理实例中,例如,针对每个cLBA将Identify Device命令发出到作为备份目的地的非易失性存储设备2B以产生数据D。

[0205] 控制程序200使用0替换作为数据复制目标LBA的指针的cLBA(步骤S800)。例如,控制程序200将检索函数f(LBA)的信息的ECH IDENTITY DEVICE命令发出到非易失性存储设备2B并且从非易失性存储设备2B检索Word129字段中的数据R(步骤S810)。然后,控制程序200通过使用连续重复的数据R填充主存储器6中512字节长的区域来产生数据Q(步骤S820)。

[0206] 控制程序200将读取LBA区域(从cLBA开始,长度为L)的数据的命令传送到非易失性存储设备2A(步骤S830)。接收读取命令的非易失性存储设备2A将LBA区域的数据D传送到信息处理设备3,然后信息处理设备3接收数据D(步骤S840)。控制程序200比较数据D与数据Q(步骤S850)。处理分支为不同的步骤,其中包括数据D与固定数据Q不同(步骤S850的结果

为否)的情况和数据D与固定数据Q相同(步骤S850的结果为是)的情况。

[0207] 当D与Q不同时(步骤S850的结果为否),控制程序200将使得数据D被写入LBA区域(从cLBA开始,长度为L)的命令传送到非易失性存储设备2B。控制程序200将数据D发送到非易失性存储设备2B,并且作为非易失性存储设备2B的存储控制器的SSDC 41B将数据D写入NAND存储器16B(步骤S860)。

[0208] 当D与Q相同时(步骤S850的结果为是),控制程序200将删除LBA区域(从cLBA开始,长度为L)到NAND存储器16B的映射的删除通知传送到非易失性存储设备2B。SSDC 41B更改管理信息46B以删除映射并且数据D变为设备无效数据(步骤S870)。

[0209] 当步骤S860或步骤S870完成时,控制程序200使cLBA递增L(步骤S880),并且重复执行步骤S810到步骤S880的处理,直到非易失性存储设备2A的所有LBA区域的数据均被备份(步骤S790的结果为是)。

[0210] 该实施例可应用于 $f(LBA)$ 不输出独立于LBA的固定数据,而是输出依赖于LBA的不同数据的情况。在这种情况下,可在步骤S810备选地发出具有用于获取函数 $f(LBA)$ 的信息的输入参数LBA的命令。例如,控制程序200可在步骤S810从非易失性存储设备2B接收 $R=f(LBA)$,并且控制程序200在步骤S820产生依赖于LBA的数据Q。例如,可使用ACS-2中描述的SCT命令作为这样的命令:其中控制程序200将B0h/D6h SMART WRITE LOG命令、3Fh WRITE LOG EXT命令或57h WRITE LOG DMA EXT命令传送到存储设备2B以便通过包括目标LBA的写入数据实现SCT命令传输;将B0h/D5h SMART READ LOG命令、2Fh READ LOG EXT命令或47h READ LOG DMA EXT命令传送到存储设备2B以实现SCT数据传输并从接收自存储设备2B的读取数据中接收数据“ $R=f(LBA)$ ”。

[0211] 这样,在该实施例中,信息处理系统1可将函数 $f(LBA)$ 的信息整合到作为备份目的地的非易失性存储设备2B中,此信息不需要在控制程序200中单独管理。因此,信息处理系统1的管理和开发得到简化。

[0212] (第五实施例)

[0213] 如图31所示,网络存储驱动器或云存储可被用作作为备份目的地的非易失性存储设备2B。例如,信息处理设备3和云存储2B经由诸如因特网或企业内联网之类的网络相互连接。控制软件200可根据该实施例,通过经由网络将删除通知命令和写入命令传送到云存储2B来执行备份处理。由于删除通知命令的数据传输量小于伴随大量数据传输的写入命令,因此可以以更高的速度执行备份处理,并且可抑制写入导致的云存储2的可靠性劣化。此外,当SSD被用作云存储2B的存储介质时,可抑制云存储2B的可用块数量的减少。

[0214] (第六实施例)

[0215] 在第一实施例和第二实施例中,描述了存储设备2A被用作备份源的情况。在第六实施例中,主存储器6充当数据复制源。也就是说,在第六实施例中,应用程序400将主存储器6中临时存储的应用程序数据500复制到非易失性存储设备2B。在应用程序或OS 100运行的同时,应用程序数据500被存储在主存储器6的应用程序数据区域6E中。在该实施例中,例如,第二实施例中描述的方法被用于将应用程序数据500存储在区域6E中。在该实施例中,存储设备2A不是必需的。

[0216] 例如,诸如DRAM(动态随机存取存储器)和SRAM(静态随机存取存储器)之类的易失性存储器或诸如MARM(磁阻式随机存取存储器)和ReRAM(电阻式随机存取存储器)之类的非

易失性存储设备可被用作主存储器6。

[0217] 图32是示出根据第六实施例的信息处理系统1的配置的示意图。在该实施例中,例如,控制工具200被嵌入OS 100。但是,OS 100和控制工具200可被存储在主存储器6内不同于区域6A的区域中。应用程序400是一般应用程序400,例如Microsoft Word (注册商标)或Microsoft PowerPoint (注册商标),并且被存储在主存储器6中的区域6D中。应用程序400是由信息处理设备2执行的程序。例如,当信息处理设备2接通电源时,应用程序400被从非易失性存储设备加载到主存储器6中的区域6D。信息处理设备3在CPU 4的控制下,与硬件资源协同执行应用程序400的处理。应用程序400使用主存储器6中的区域6E作为工作区域。例如,当用户通过显示器9和鼠标15按下应用程序400的GUI的“保存”按钮,或者调用应用程序400的自动保存功能时,如图33所示,应用程序400将命令发送到OS 100以存储所述存储设备2B中的区域6E的应用数据500,这样,OS 100将区域6E的数据备份到存储设备2B。

[0218] 图34是示出根据第六实施例的控制工具200的处理实例的流程图。例如,当信息处理系统1接通电源时,如图35所示,控制工具200将OS文件管理表(元数据、OS文件系统)300从作为备份目的地的非易失性存储设备2B的NAND存储器16B加载到主存储器6中的区域6C。当要在存储设备2B中存储区域6E的应用数据500的命令被从应用程序400传送到OS 100时(步骤S900),控制工具200读取存储在主存储器6中的区域6C中的OS文件管理表(元数据)(步骤S910)。这样,控制工具200(OS 100)搜索主存储器6的区域6E的地址(从中读取数据以及将数据的LBA区域存储在非易失性存储设备2B中)(步骤S920)。OS 100从主存储器6中的区域6E读取应用数据500(步骤S930)。

[0219] 为了比较读取的应用数据500,读取的应用数据500优选地被分成逻辑扇区单元(步骤S940)。控制工具200选择被分割的单元中的一个单元(步骤S950),并判定单元的数据D是否与函数 $f(cLBA)$ 相同(步骤S960)。当数据D与函数 $f(LBAc)$ 的输出数据不同时,控制工具例如使用非专利文献1的ACS-2中描述的61h WRITE FPDMA QUEUED命令或35h WRITE DMA EXT命令、NVM Express Revision 1.1中描述的01h Write命令等将数据D写入存储设备2B的LBA区域,如图36所示(步骤S980)。当数据D与函数 $f(LBA)$ 的输出数据相同时,控制工具不将写入命令传送到存储设备2B,而是将删除通知命令发送到存储设备2B以取消从NAND存储器16B到LBA区域的映射,如图37所示(步骤S970)。例如,可使用非专利文献1的ACS-2中描述的06h DATA SET MANAGEMENT命令或SCT命令,NVM Express Revision 1.1中描述的11h Dataset Management命令的Deallocate (AD)等作为删除通知命令。例如,当SSDC 41B具有下面将描述的Deterministic Zeroing TRIM特性时,不仅使用上述命令,而且还使用NVM Express Revision 1.1中描述的08h Write Zeroes命令作为删除通知命令。此外,可使用2012年12月7日发布的SCSI Block Commands-3 (SBC-3) Revision 35 (<http://www.t10.org/>)中描述的42h UNMAP命令作为删除通知命令。

[0220] 控制程序200重复对所有分割的单元的处理(步骤S950到步骤S990)。控制程序200将LBA(其中写入数据500)写入主存储器6的OS文件管理表(元数据)300。此外,主存储器6的OS文件管理表300被备份并且被存储在非易失性存储设备2B中的NAND存储器16B中,然后处理结束(步骤S995)。

[0221] 这样,在第六实施例中,可使用删除通知命令,以更高的速度复制主存储器6的数据并将其存储在存储设备2B中。

[0222] (第七实施例)

[0223] 在第七实施例中,将描述网络7002被用作存储设备2A和2B的接口的情况。在该实施例中,图38中的信息处理系统1被配置为包括与存储设备2A和2B连接的网络7002。

[0224] 网络7002是用于存储访问的存储网络。例如,可以使用光纤通道或以太网。具体而言,例如,SAN(存储区域网)或NAS(网络连接存储)可被用作存储网络。例如,FC-SAN(光线通道存储区域网)或IP-SAN(网际协议存储区域网)被用作SAN。例如,SCSI(小型计算机系统接口)可被用作高级协议层。在该实施例中,将描述IP-SAN被用作存储网络7002的实例。此外,在该实施例中,iSCSI(互联网小型计算机系统接口)被用作高级协议层。存储网络7002包括网络交换机7005和集线器(未示出)。

[0225] 网络7001是将一个或多个服务器以及一个或多个客户机相互连接的网络。例如,可使用LAN(局域网)或WAN(广域网)。例如,可使用以太网作为LAN。

[0226] 客户机7004是连接到网络7001并且执行所请求的处理的计算机。一般而言,客户机7004包括硬件资源,例如处理器、主存储器、通信接口和本地输入/输出设备,还包括软件资源,例如设备驱动程序、操作系统(OS)和应用程序(未示出)。这样,客户机7004在控制器的控制下与硬件资源协同执行多个程序的处理。例如,客户机7004经由数据库服务器7003执行对存储单元2A的I/O访问,并且通过在控制器的控制下执行业务应用程序来实现所需的业务系统。

[0227] 数据库服务器(DB服务器)7003是连接到网络7001和存储网络7002,并且其中运行数据库管理系统(DBMS)的服务器。当数据库服务器7003接收来自客户机7004的数据读取请求时,数据库服务器7003从存储单元2A或2B读取数据,然后将读取数据传送到客户机7004。当数据库服务器7003接收来自客户机7004的数据写入请求时,数据库服务器7003从客户机7004接收写入数据并将已接收的数据写入存储单元2A或2B。数据库服务器7003存储被配置为管理存储在存储单元2A或2B中的数据的管理信息(元数据)。

[0228] 存储单元2A可单独为DB服务器7003提供数据存储服务,或者可作为包括多个存储设备的虚拟存储设备为DB服务器7003提供数据存储服务。在存储设备2A中,一个或多个逻辑设备(LDEV)被配置为提供给DB服务器7003。根据RAID配置,逻辑设备可被配置为包括多个硬盘驱动器的虚拟设备(VDEV)。

[0229] 逻辑设备(LDEV)是被DB服务器7003识别的逻辑存储设备。逻辑单元(LU)被分配给逻辑设备。DB服务器7003识别在物理设备上被配置为逻辑单元的逻辑设备。逻辑设备号(LUN)被分配给每个逻辑单元。逻辑单元被分成逻辑扇区(逻辑块),LBA被分配给每个逻辑扇区。DB服务器7003可通过将包括根据逻辑设备号和LBA配置的逻辑地址的命令发送到存储设备来访问特定逻辑单元的特定逻辑扇区。在该使用iSCSI的实施例中,如上所述,DB服务器7003、存储单元2A和存储单元2B被配置为充当像被分配iSCSI名称的iSCSI节点那样的启动器和目标。因此,DB服务器7003、存储单元2A和存储单元2B经由IP地址和TCP端口号指定的网络入口发送iSCSI PDU。因此,DB服务器7003可通过指定iSCSI名称、IP地址和TCP端口号来识别网络7002上的存储单元2A和2B,并且可访问存储单元的逻辑单元中的逻辑扇区。此外,客户机7004可直接连接到存储网络7002而无需通过DB服务器7003。

[0230] 备份服务器7000连接到存储网络7002,因此控制工具200被存储在存储服务器7000的主存储器中。DB服务器7003可具有备份服务器7000的功能。如图39所示,控制工具

200的功能可能不被配置为作为备份服务器7000的一部分的控制工具200,而是可作为控制单元200嵌入网络交换机7005。如图40所示,控制工具200的功能可嵌入存储设备2A作为控制单元200。此外,控制工具200的功能可嵌入客户机7004作为控制单元200。

[0231] 当IT管理者或运营商将存储单元2A中存储的数据备份到存储单元2B以实现灾难恢复时,存储单元2A与2B之间的距离优选地等于或大于100公里。在这种情况下,如图41所示,存储单元2A和2B优选地经由通道扩展器7006a和7006b以及经由长距离网络7007(例如WAN(广域网))或专线相互连接。

[0232] 在该实施例中,描述了作为备份源的存储单元2A和作为备份目的地的存储单元2B是不同的存储设备的情况。但是,即使存储单元2A和2B被嵌入同一存储设备作为不同的存储区域,也可应用本发明。也就是说,本发明可应用于这样的情况:其中存储设备包括两个或更多被分配LU的LDEV,并且一个LU被用作存储单元2A,另一LU被用作存储单元2B,其中控制工具200(控制单元200)将存储单元2A的数据复制到存储单元2B。当控制工具200(控制单元200)将数据从作为源的一个LU复制到作为目的地的另一LU时,本发明可应用于复制源LU中的所有LBA区域的情况,以及复制源LU的部分LBA区域的情况。

[0233] 图42是示出根据该实施例的由控制工具200(控制单元200)执行的备份处理的示意图。当备份开始,备份服务器700中的控制程序200使用0替换作为数据复制目标LBA区域的指针的cLBA(步骤S1000)。控制程序200将读取LBA区域(从cLBA开始,长度为L)的数据的命令传送到存储单元2A(步骤S1010)。作为来自存储单元2A的读取命令,例如,可使用SCSI Block Commands-3(SBC-3),T10/BSR INCITS 514,Revision 35(<http://www.t10.org/>)中描述的28h READ(10)命令、A8h READ(12)命令、88h READ(16)命令、7Fh READ(32)命令等。接收读取命令的存储单元2A经由存储网络7002将LBA区域的数据(D)传送到备份服务器7000,然后备份服务器7000接收数据D(步骤S1020)。

[0234] 控制程序200比较数据D与函数f(cLBA)的输出数据(步骤S1030)。当数据D与函数f(cLBA)的输出数据不同时(步骤S1030的结果为否),控制程序200将使得数据D被写入LBA区域(从cLBA开始,长度为L)的命令传送到存储单元2B并且将数据D传送到存储单元2B(步骤S1040)。这样,SSDC 41B将数据D写入NAND存储器16B。作为被发送到存储单元2B的写入命令,例如,使用上述SBC-3中描述的2Ah WRITE(10)命令、AAh WRITE(12)命令、8Ah WRITE(16)命令、7Fh WRITE(32)命令等。

[0235] 当D与函数f(cLBA)的输出数据不同时(步骤S1030的结果为是),控制程序200将删除通知传送到存储单元2B以删除LBA区域到NAND存储器16B的映射(步骤S1050)。SSDC 41B更改管理信息46B以删除映射并且数据D变为设备无效数据。作为被发送到存储单元2B的删除通知命令,例如,可使用上述SBC-3中描述的42h UNMAP命令等。

[0236] 当步骤S1040或步骤S1050完成时,控制程序200使cLBA递增L(步骤S1060),然后针对所有小区域执行步骤S1010到步骤S1050的子例程。当cLBA超过存储设备2B的最大可访问LBA时,假设所有LBA区域的读取完成(步骤S1070的结果为是),这样,备份处理便完成。

[0237] 作为备份目标范围,控制程序200(控制单元200)可备份存储单元2A的整个LBA区域的数据,或者可备份存储单元2A的部分LBA区域的数据。当备份存储单元2A的部分LBA区域时,例如,可将存储单元2A的逻辑单元(LU)中的所有LBA备份到存储单元2B。在这种情况下,逻辑单元(LU)被分配给存储单元2B,并且数据被从存储单元2A的逻辑单元(LU)复制到

存储单元2B的逻辑单元(LU)。

[0238] 如图40所示,当控制单元200嵌入作为备份源的存储设备2A时,在该实施例中,控制单元200可处理无服务器备份,作为控制工具200(控制单元200)的备份处理。在这种情况下,例如可使用扩展复制命令。例如,可使用SCSI Primary Commands-3 (SPC-3), INCITS T10/1731-D, Revision 36e (<http://www.t10.org/>) 中描述的83h EXTENDED COPY命令作为扩展复制命令。当备份服务器7000、数据库服务器7003、网络交换机7005或客户机7004将包括备份源地址和备份目的地地址的扩展复制命令传送到存储设备2A时,控制单元200针对备份源和目的地区域执行备份处理,如图18所示的处理。也就是说,控制单元200重复从扩展复制命令指定的存储设备2A的源区域读取数据,然后判定读取数据D是否与f(cLBA)相同。当 $D \neq f(cLBA)$ 时,控制单元200经由存储网络7002将写入命令和数据D传送到存储设备2B。当 $D = f(cLBA)$ 时,控制单元200经由存储网络7002将删除通知传送到存储设备2B。这样,存储设备2B执行数据写入处理和删除通知处理。

[0239] (第八实施例)

[0240] 在第七实施例中,描述了经由存储网络7002备份数据的情况。如图43所示,与客户机7004A连接的存储单元2A的数据可被备份到与存储网络7002连接的存储单元2B。在备份处理时,控制单元200重复从存储设备2A的备份源区域读取数据,然后判定读取数据D是否与f(cLBA)的输出数据相同。然后,当 $D \neq f(cLBA)$ 时,控制单元200经由存储网络7002将写入命令和数据D发送到存储单元2B。当 $D = f(cLBA)$ 时,控制单元200经由存储网络7002将删除通知发送到存储单元2B。

[0241] 如图44所示,与存储网络7002连接的存储单元2A的数据可被备份到与客户机7004B连接的存储单元2B。在备份处理时,控制单元200重复从存储设备2A的备份源区域读取数据,然后判定读取数据D是否与f(cLBA)的输出数据相同。然后,当 $D \neq f(cLBA)$ 时,控制单元200将写入命令和数据D传送到存储设备2B。当 $D = f(cLBA)$ 时,控制单元200经由存储网络7002将删除通知传送到存储设备2B。

[0242] (第九实施例)

[0243] 在第九实施例中,将描述网络7002被用作存储单元2A和2B的接口的情况,如在第七实施例中那样,并且以文件为单元执行备份。图45中的信息处理系统1被配置为包括与存储网络7002连接的元数据服务器(MDS) 7008,以及被存储在作为备份源的存储单元2A和存储单元2C(在分布式管理方法中)中的元数据300。

[0244] 例如,元数据300可具有与第二实施例或第六实施例中描述的OS文件管理信息300相同的结构。例如,元数据300可以是图35所示的OS文件管理表300,或者可具有另一结构。可使用块算法或B+tree算法作为元数据300的管理算法。在该实施例中,将描述元数据300具有图46所示的结构的情况。元数据300包括文件标识ID、存储单元标识ID(用于指定存储文件的存储单元)、LBA(指示存储文件的LBA区域的起始LBA)、扇区计数(指示文件的数据长度)、文件所有者ID、文件最后更新时间,以及文件的其它属性信息。存储单元标识ID优选地是每个存储单元的唯一编号。例如,可使用全球通用名称(WWM)、IP地址或MAC地址。元数据300可仅存储在作为备份源的存储单元2A中。但是,从整个系统的可靠性和性能提升的角度来看,元数据300优选地借助分布式管理方法存储在存储单元2A和存储单元2C(其是存储单元2A和2B之外的与存储网络7002连接的存储单元)中。备选地,元数据300可以不存储在存

储网络7002上的存储单元中,而是可存储在直接连接到元数据服务器7008的本地输入/输出设备(未示出)上。

[0245] 元数据服务器(MDS)7008连接到存储网络7002并且通过管理存储在存储单元2A、2B和2C中的文件数据的元数据300来配置信息处理系统1的文件系统。元数据300包括文件ID到存储单元2A、2B和2C的映射。例如,SAN文件系统、集群文件系统、RAIN(廉价/独立节点的冗余阵列)文件系统、pNFS(并行网络文件系统)或Google文件系统(GFS)可被用作文件系统。元数据服务器7008包括硬件资源,例如处理器、主存储器、通信接口和本地输入/输出设备,还包括软件资源,例如设备驱动程序、操作系统(OS)和元数据管理程序(未示出)。这样,元数据服务器7008在控制器的控制下与硬件资源协同执行多个程序的处理。

[0246] 客户机7010是连接到存储网络7002并且执行所需处理的计算机。客户机7010包括硬件资源,例如处理器、主存储器、通信接口和本地输入/输出设备,还包括软件资源,例如设备驱动程序、操作系统(OS)和应用程序(未示出)。这样,客户机7010在控制器的控制下与硬件资源协同执行多个程序的处理。例如,客户机7010执行对存储单元2A、2B和2C的文件访问,并且通过在控制器的控制下执行业务应用程序来实现所需的业务系统处理。

[0247] 客户机7010从元数据服务器7008搜索存储单元中的文件地址和文件属性信息以访问文件数据。例如,当客户机7010读取文件数据时,客户机7010将文件信息读取请求和文件的文件标识ID传送到元数据服务器7008,元数据服务器7008搜索从元数据300的文件标识ID映射的存储单元标识ID、LBA、扇区计数和属性信息,元数据服务器7008将对应行的存储单元标识ID、LBA、扇区计数和属性信息等传送到客户机7010,客户机7010将数据读取命令、LBA和扇区计数传送到存储单元标识ID指定的存储单元,该指定的存储单元从LBA读取具有扇区计数长度的数据,然后该存储单元将数据传送到客户机7010。

[0248] 例如,当客户机7010写入文件数据时,客户机7010将写入区域分配请求、文件的文件标识ID和文件大小传送到元数据服务器7008,元数据服务器7008读取元数据300,元数据服务器7008通知客户机7010存储单元标识ID、将被写入文件的LBA、文件的属性信息等,客户机7010将数据写入命令、LBA和扇区计数传送到存储单元标识ID指定的存储单元,客户机7010将要写入的数据发送到存储单元,该存储单元将数据写入LBA区域(从LBA开始,长度为扇区计数)。

[0249] 图47是示出根据该实施例的由控制工具200(控制单元200)执行的备份处理的流程图。作为来自存储单元2A的读取命令,例如,可使用上述SBC-3中描述的28h READ(10)命令、A8h READ(12)命令、88h READ(16)命令、7Fh READ(32)命令等。作为存储单元2B的写入命令,例如,可使用上述SBC-3中描述的2Ah WRITE(10)命令、AAh WRITE(12)命令、8Ah WRITE(16)命令、7Fh WRITE(32)命令等。作为到存储单元2B的删除通知命令,可使用上述SBC-3中描述的42h UNMAP命令等。

[0250] 控制工具200将文件ID列表请求传送到元数据服务器7008(步骤S1100)。元数据服务器7008读取元数据300,创建在元数据300中注册的文件的文件ID列表,然后将文件ID列表传送到控制工具200(步骤S1110)。例如,文件ID列表可以具有图48所示的阵列数据结构。例如,文件ID列表可包括元数据300中注册的所有文件。例如,文件ID列表可包括元数据300中注册的部分文件,例如包括在存储单元2A中的所有文件、包括在特定文件夹中的所有文件,和包括在特定LU中的所有文件。在该实施例中,将描述备份包括在存储单元2A中的所有

文件的情况。控制工具200将作为备份目标文件列表的已接收文件ID列表存储在备份服务器7000的主存储器中(步骤S1120)。

[0251] 控制工具200将读取备份文件列表中每个文件的信息的请求和文件标识ID传送到元数据服务器7008。元数据服务器7008从元数据300搜索文件标识ID并将其中存储文件的LBA范围(从cLBA开始,长度为cL)的信息传送到控制工具200(步骤S1130)。

[0252] 控制工具200将写入区域分配请求、文件标识ID和文件的扇区计数cL传送到元数据服务器7008。元数据服务器7008基于元数据300检索存储单元标识ID和其中将写入文件的LBA(=tLBA),并将tLBA和作为备份目的地的存储单元2B的存储单元标识ID传送到控制工具200(S1140)。当控制工具200发送写入区域分配请求时,控制工具200优选地显式请求元数据服务器7008避免分配存储单元2A中要被写入的LBA区域。

[0253] 控制工具200将读取区域(从cLBA开始,长度为cL)的数据读取命令传送到存储单元2A,存储单元2A从LBA区域读取数据并将数据传送到控制工具200(S1150)。控制工具200将从存储单元2A接收的数据分成例如逻辑扇区单元(=D)(步骤S1160),并且控制工具200针对每个分割的单元的每个从LBA开始的tLBA判定每个数据D是否与f(tLBA)的输出数据相同(步骤S1170和步骤S1180)。

[0254] 当D与f(tLBA)的输出数据不同时,控制工具200将在LBA区域(从tLBA开始)中写入数据D的写入命令传送到存储单元2B,并且将数据D传送到存储单元2B(步骤S1190)。当D与f(tLBA)的输出数据相同时,控制工具200将删除LBA区域(从tLBA开始)到NAND存储器16B的映射的删除通知传送到存储单元2B(步骤S1200)。当多个写入命令指定的LBA范围是连续区域时,多个写入命令优选地进行统一并被传送到存储单元2B以提高步骤S1190的处理效率。备选地,当多个删除通知指定的LBA范围是连续区域时,多个删除通知优选地进行统一并被传送到存储单元2B以提高步骤S1200的处理效率。

[0255] 控制工具200请求元数据服务器7008更改元数据300以将备份文件的存储单元标识ID从存储单元2A的ID更新为存储单元2B的ID。cLBA与tLBA之间的关系可以是“cLBA=tLBA”,或者可以是“cLBA≠tLBA”。当cLBA≠tLBA时,控制工具200请求元数据服务器7008更改元数据300以将文件的LBA从备份文件的cLBA更新为tLBA(步骤S1170)。

[0256] (第十实施例)

[0257] 在使用根据第一至第九实施例的方法时,控制工具200可以更高的速度将存储单元2A中存储的用户数据复制到存储单元2B。购买存储单元2A的最终用户可通过以下方式抢救数据:当所购买的存储设备2A的可靠性劣化时,使用控制工具200将数据从存储单元2A复制到新购买的具有良好可靠性的存储设备2B。另一方面,最终用户自己可能不恢复数据,但是存储设备2A或2B的分销商,或者数据恢复服务提供商可执行将用户数据从存储设备2A备份到存储设备2B的抢救服务,此项服务可能收费,也可能免费。在该实施例中,数据抢救服务将被描述为商业服务。将在该实施例中描述的数据抢救服务可能使用,也可能不使用根据第一至第九实施例的方法。当组合根据第一至第九实施例的方法时,可以更高的速度抢救数据,并且写入作为备份目的地的存储设备2B的数据量可被减少,这样,存储设备2B的寿命可延长。

[0258] 当最终用户购买存储设备2A时,最终用户可通过接口2A将存储设备2A连接到信息处理设备3,例如个人计算机、服务器、游戏设备或数字照相机。在信息处理设备正在运行

时,信息处理设备3将数据写入存储设备2A并且从存储设备2A读取数据。

[0259] 存储设备2A可能有一天突然无法启动,无法正确地读取和写入数据,或者因为存储设备2A长期磨损而崩溃,此时错误会累积或者存储设备2A出现错误。当发生崩溃时,最终用户可能无法读取存储设备2A中存储的诸如商业文档、照片和电子邮件之类的用户数据。因此,在存储设备2A崩溃之前,优选地提前将用户数据复制和抢救到存储设备2B(其是具有良好的可靠性的存储设备)。

[0260] 数据抢救服务优选地尽快执行,以便用户数据不会在数据抢救期间意外丢失,并且在预测存储设备2A崩溃之后,存储设备2A的可靠性不会进一步劣化。

[0261] 在数据抢救服务中,即使在存储设备2A崩溃之后,也可从存储设备2A抢救数据。例如,当存储设备达到US 2009/0222617 A1中描述的只读模式或者写保护模式时,存储设备2A无法正常处理已接收的写入命令,如US 2012/0284453 A1中描述的那样。因此,诸如Windows(注册商标)之类的OS不能正常识别存储设备2A,从而使得最终用户很难抢救用户数据。即使存储设备崩溃,数据抢救服务提供商也可在用于抢救数据的信息处理设备3上使用例如实施例中描述的控制工具200,或者例如US 2012/0284453 A1中描述的控制程序来抢救用户数据。

[0262] 图49是示出从用户数据抢救服务中的存储设备2A故障预测时间(或崩溃时间)步骤到完成将用户数据抢救到存储设备2B的步骤的处理流的流程图。例如,用户数据抢救服务可由存储设备2A或2B的分销商作为免费服务或收费服务执行,或者用户数据抢救服务可由数据抢救服务提供商作为收费服务执行。

[0263] 存储设备2A的崩溃例如可通过收集存储设备2A的统计信息(例如S.M.A.R.T信息)提前预测或获知。可使用例如US 2012/0284453 A1、US 2010/0250450 A1或US 2011/0239045 A1中公开的技术,借助崩溃预测软件提前预测存储设备2A的崩溃。

[0264] 当崩溃预测软件例如通过显示器9通知最终用户崩溃预测或崩溃时,最终用户将存储设备2A与信息处理设备3分离,然后将存储设备2A送到数据抢救服务中心(发运给数据抢救服务提供商)。

[0265] 服务中心从用户处接收存储设备2A。服务中心准备存储设备2B(具有良好的可靠性的存储设备)作为存储设备2A的数据备份目的地(抢救目的地)。例如,服务中心可新制作存储设备2B作为存储设备2A的数据复制目的地,或者新购买存储设备2B。

[0266] 服务中心运营商将作为备份源的存储设备2A和作为备份目的地的存储设备2B连接到作为数据抢救服务设备并且用于抢救数据的信息处理设备3。运营商控制信息处理设备3以将用户数据D001从存储设备2A复制到存储设备2B。复制处理优选地作为第一至第九实施例中描述的复制处理(备份处理)执行。用于抢救数据的信息处理设备3优选地可包括所述实施例中描述的控制工具200,或者例如优选地可包括US 2012/0284453 A1中描述的控制程序。

[0267] 为了防止存储设备2A的可靠性进一步劣化,信息处理设备优选地避免将写入命令传送到存储设备2A。当存储设备2A的接口控制器的可靠性劣化时,出现这样的问题:即,信息处理设备3与存储设备2A通信所使用的接口19A中出现诸如CRC错误或R错误之类的通信错误。因此,用于数据抢救服务的信息处理设备3优选地限制接口19A的连接速度。例如,当接口19A或存储设备2A的接口控制器支持6.0 Gbps比特率(第三代Serial ATA的最大连接

速度)时,信息处理设备3优选地以较低的连接速度3.0 Gbps与存储设备2A通信,或者,更优选地以Serial ATA的最低连接速度1.5 Gbps与存储设备2A通信。当接口19A支持诸如PCI Express之类的多通道通信,并且存储设备2A的接口控制器可靠性劣化时,出现这样的问题:在与信息处理设备3通信期间出现诸如通道内时脉扭曲故障之类的通信错误。因此,用于数据抢救服务的信息处理设备3优选地限制连接通道的数量。例如,当接口19A或存储设备2A的接口控制器支持PCI Express 4通道连接时,信息处理设备3优选地以较小通道数量的2-通道连接与存储设备2A通信,或者,更优选地以最低通道数量的1-通道连接与存储设备2A通信。在备份处理结束之后,用于数据抢救的信息处理设备3读取存储设备2B中存储的用户数据D001,然后比较用户数据D001与读取自存储设备2A的数据,以检验用户数据D001是否被正确地写入存储设备2B(验证处理)。

[0268] 服务中心将作为备份目的地的存储设备2B发运给最终用户。备选地,服务中心可通过电子邮件将经过抢救的数据D001传送给用户,或者可以将经过抢救的数据D001上传到云存储以便通过电子的方式将经过抢救的数据D001传送给用户。

[0269] 用户获取经过抢救的数据D001。

[0270] (修正实例)

[0271] 在上述每个实施例中,描述了SSD被用作作为备份源的非易失性存储设备2A的实例。每个实施例可广泛地应用于SSD之外的非易失性存储设备。例如,每个实施例也可应用于使用磁介质(圆形磁盘片)的硬盘驱动器(HDD),以及同时使用NAND型闪存和磁介质(圆形磁盘片)的混合硬盘驱动器(混合HDD)。由于删除通知命令的传输目的地不是作为备份源的非易失性存储设备2A,而是作为备份目的地的非易失性存储设备2B,因此,本发明适用于这样的情况:其中作为备份源的非易失性存储设备2A不支持删除通知命令。

[0272] 尽管已经描述了特定实施例,但是这些实施例仅出于举例的目的给出,并非旨在限制本发明的范围。实际上,此处描述的新颖实施例可通过多种其它形式体现;而且,在不偏离本发明精神的情况下,可对此处描述的实施例做出各种形式上的省略、替换和更改。所附权利要求及其等同物旨在涵盖将落在本发明的范围和精神内的此类形式或修改。

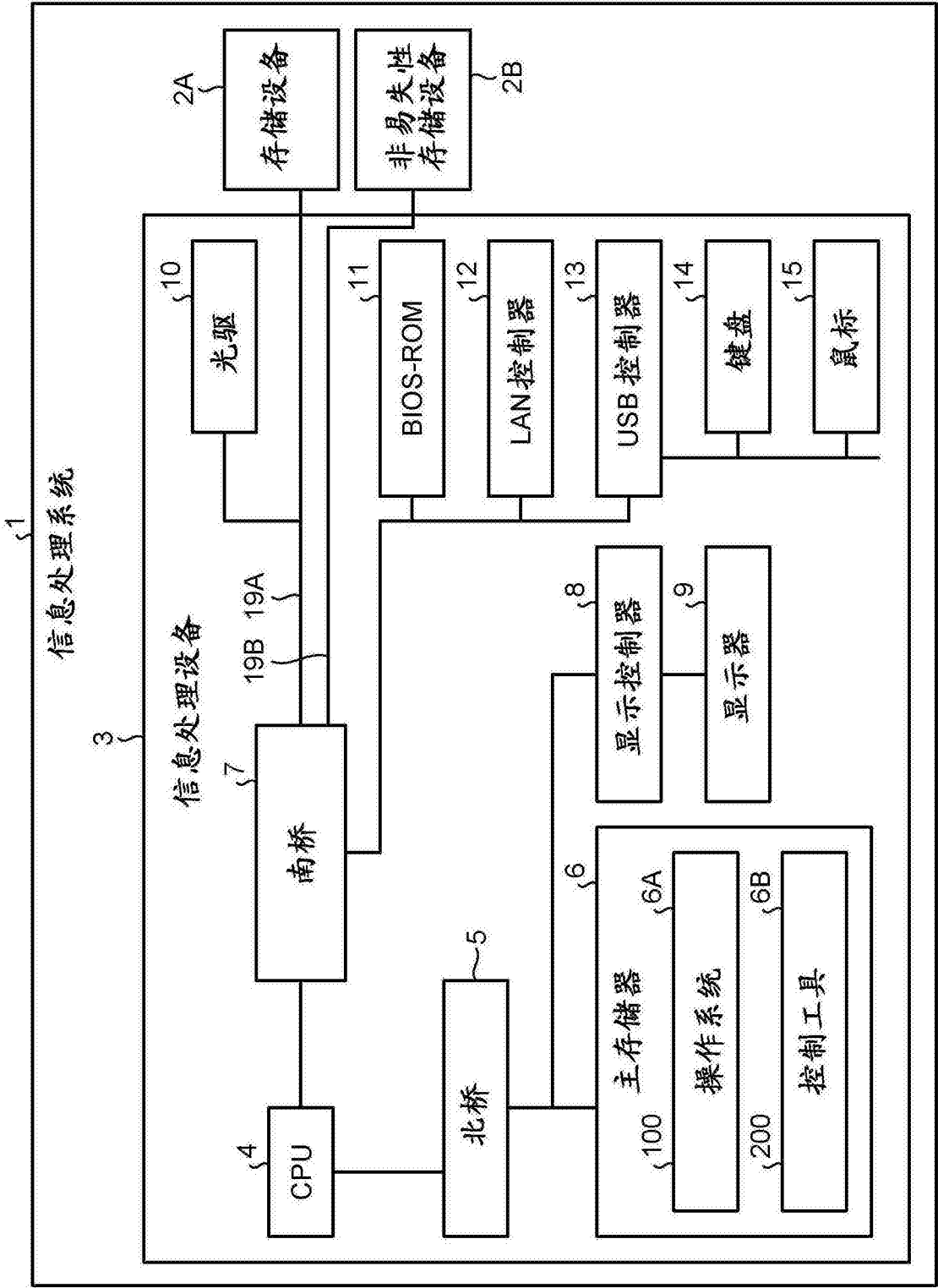


图1

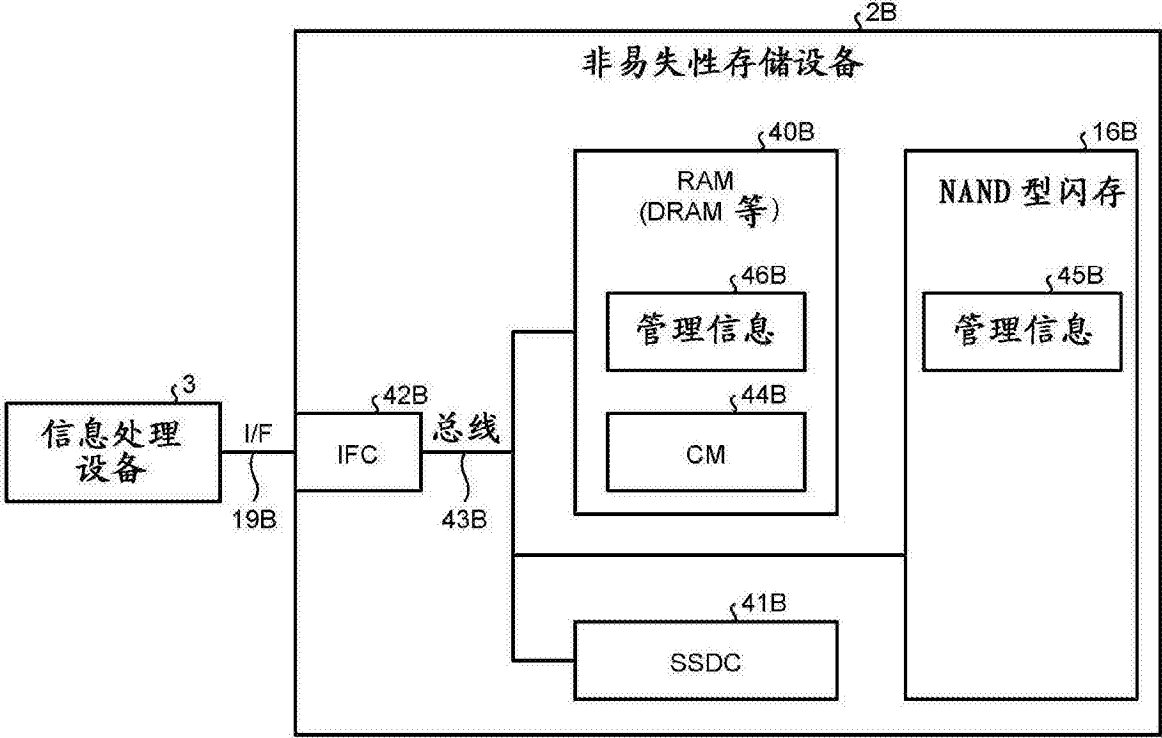


图2

546B

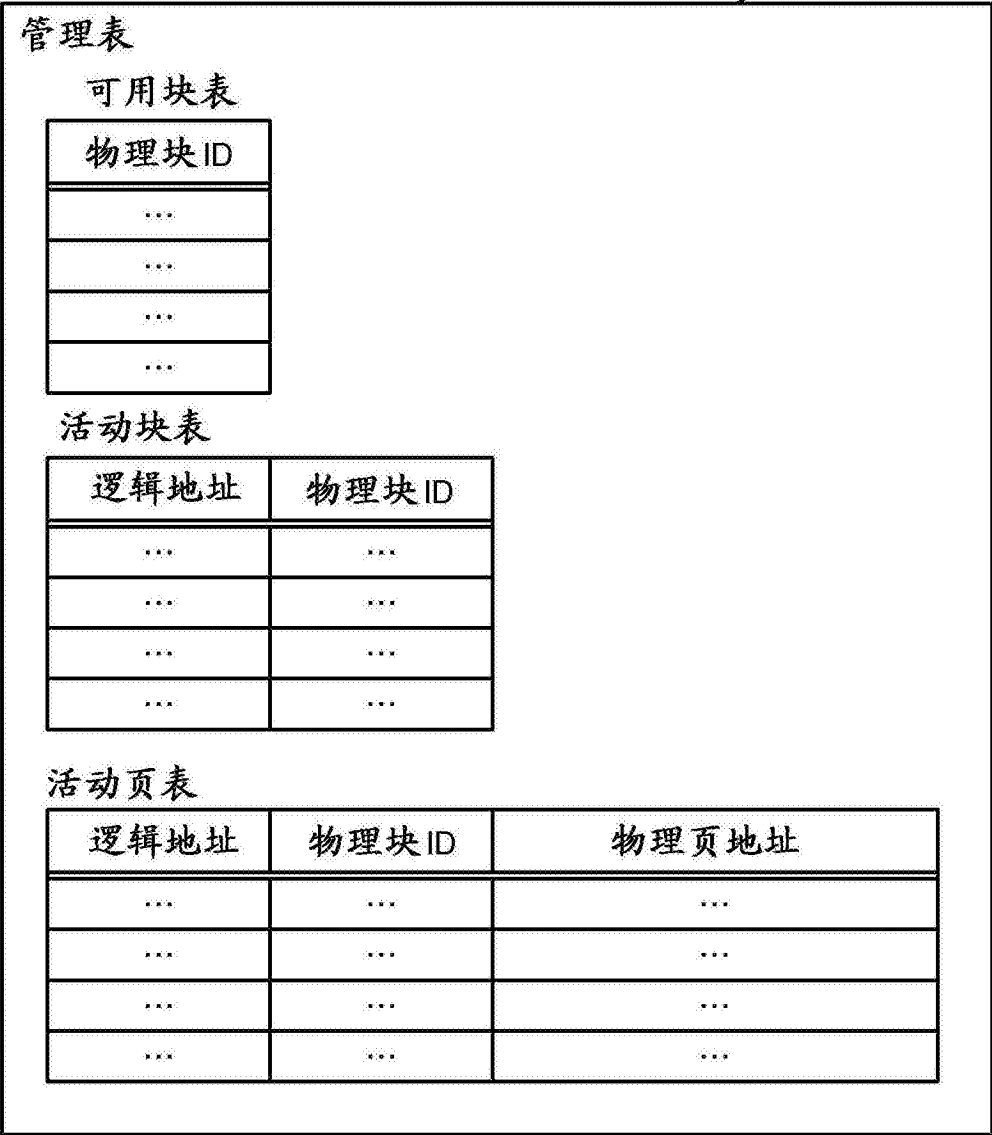


图3

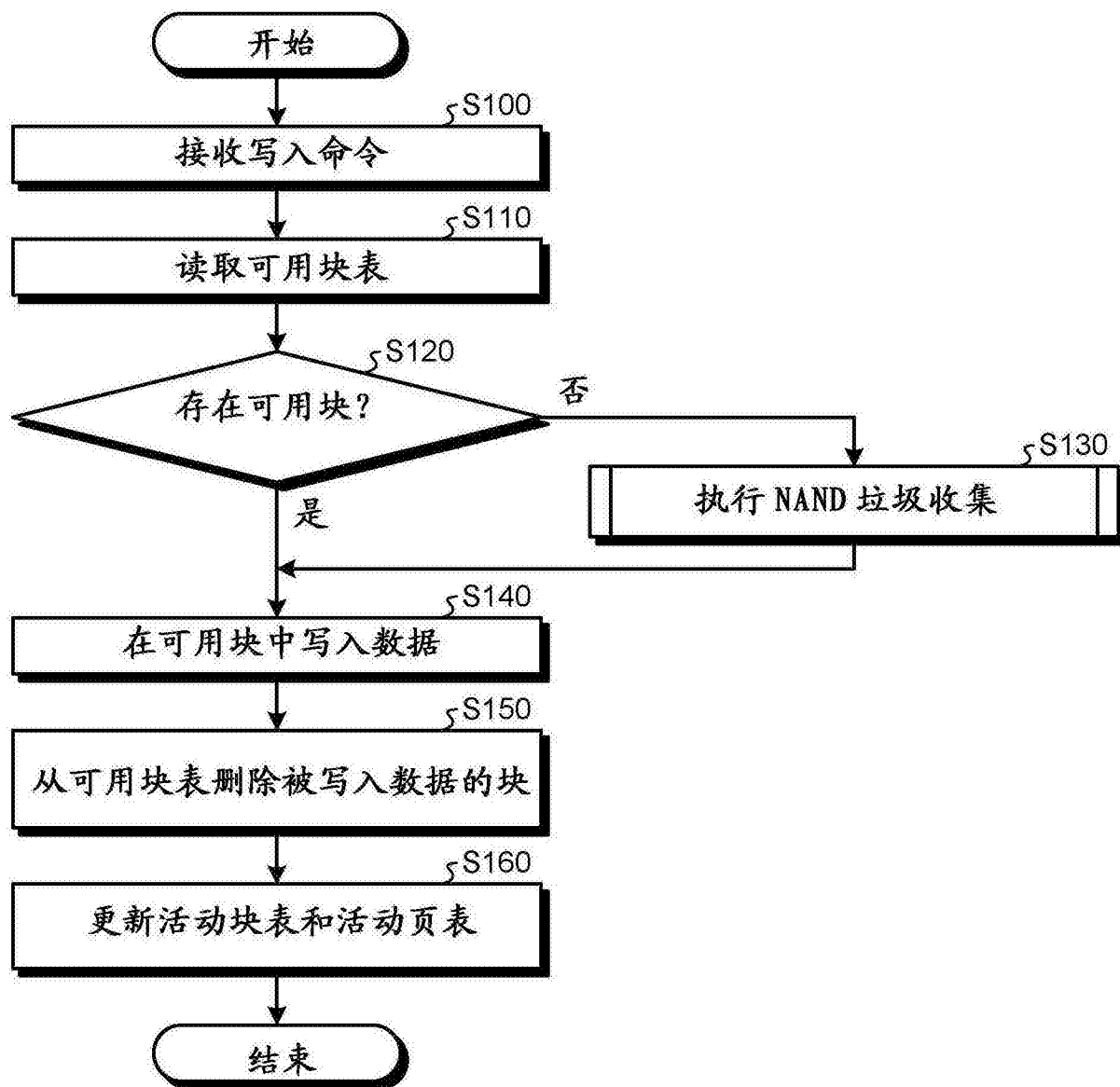


图4

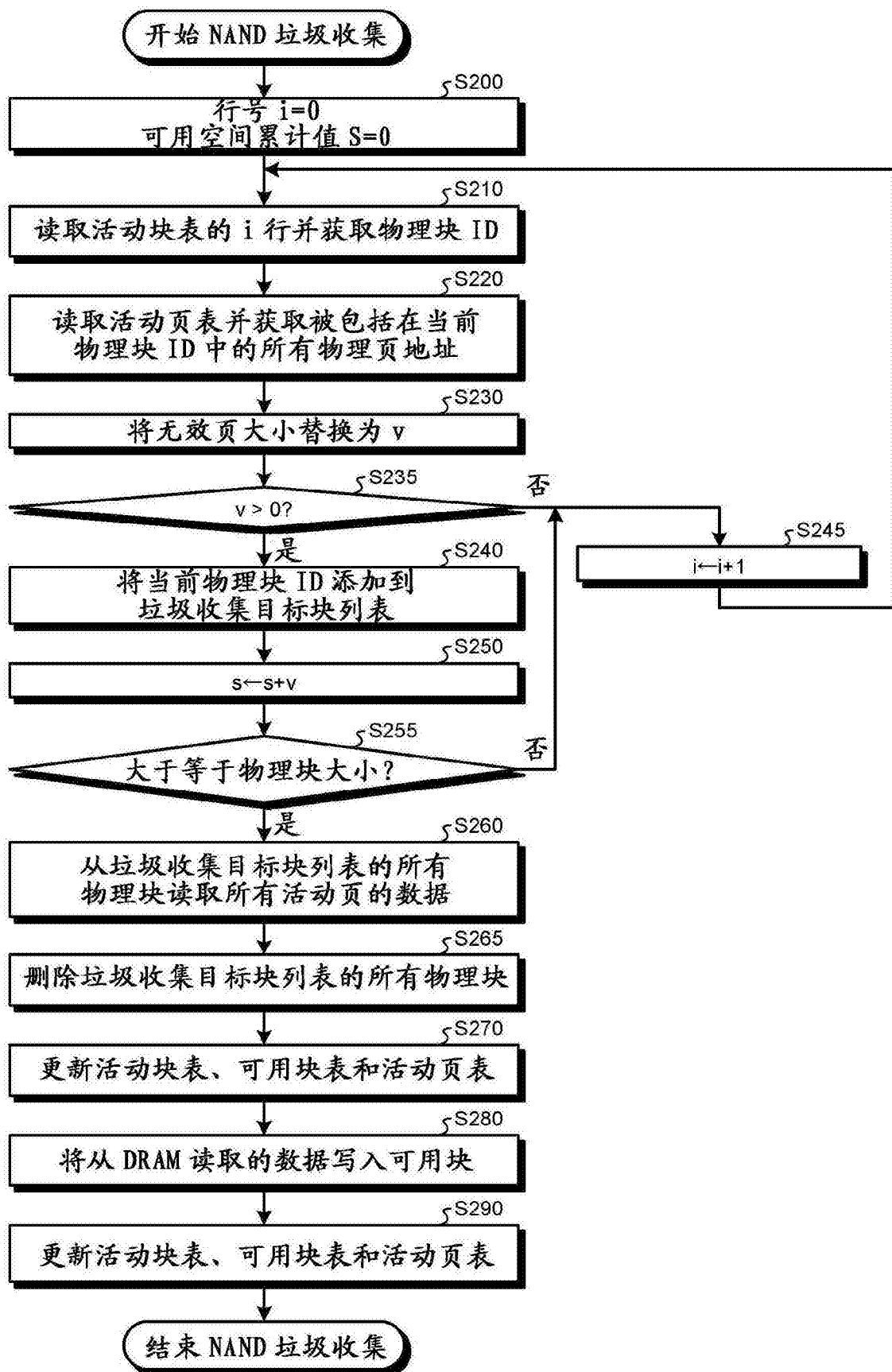


图5

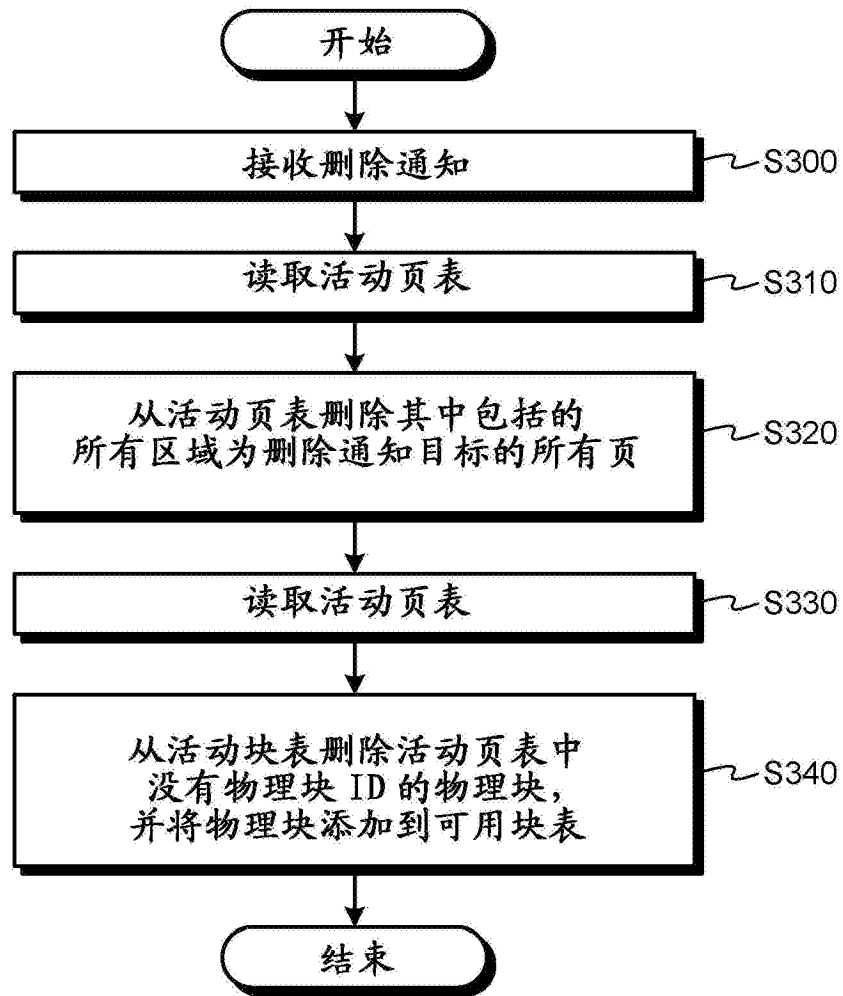


图6

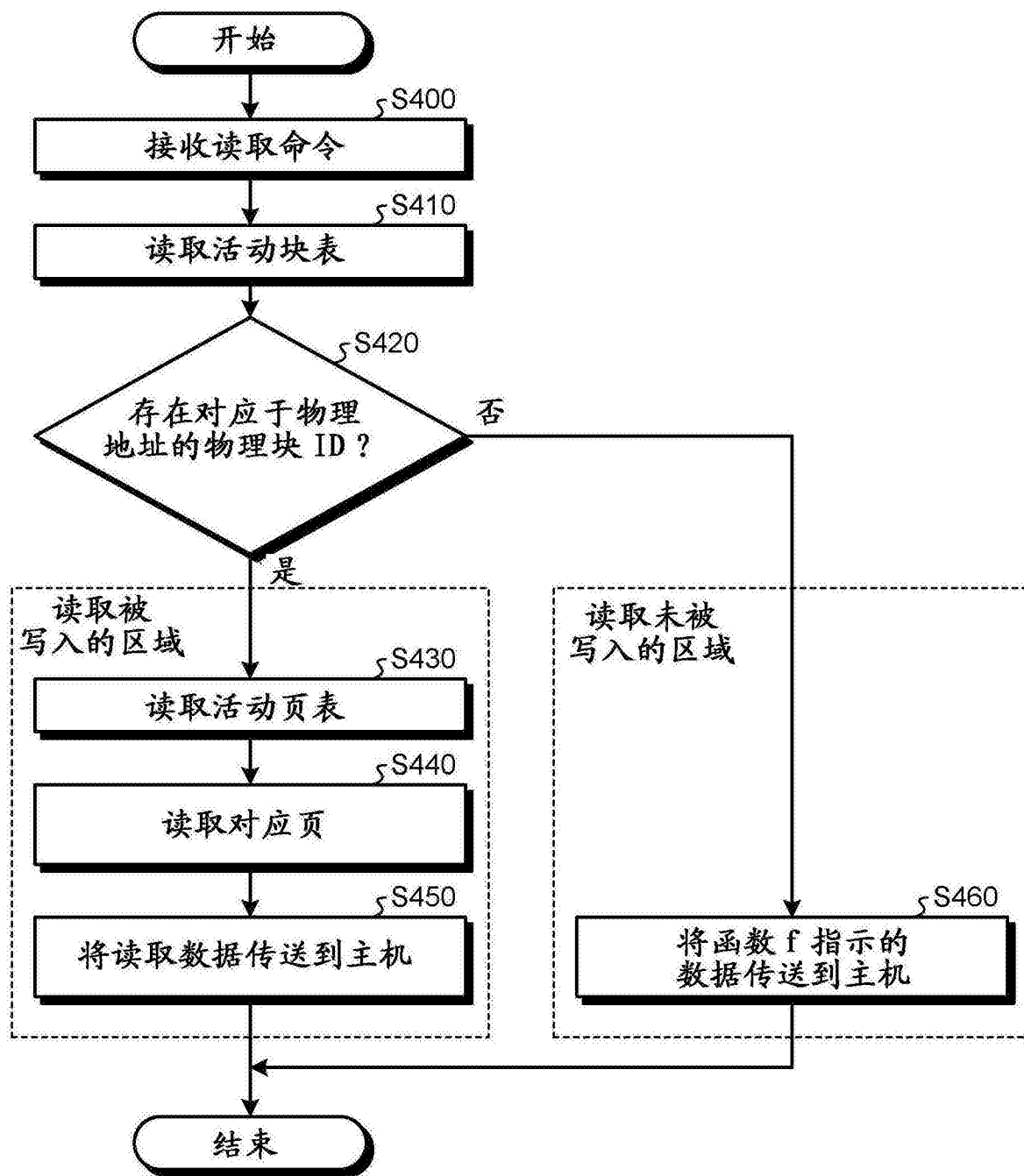


图7

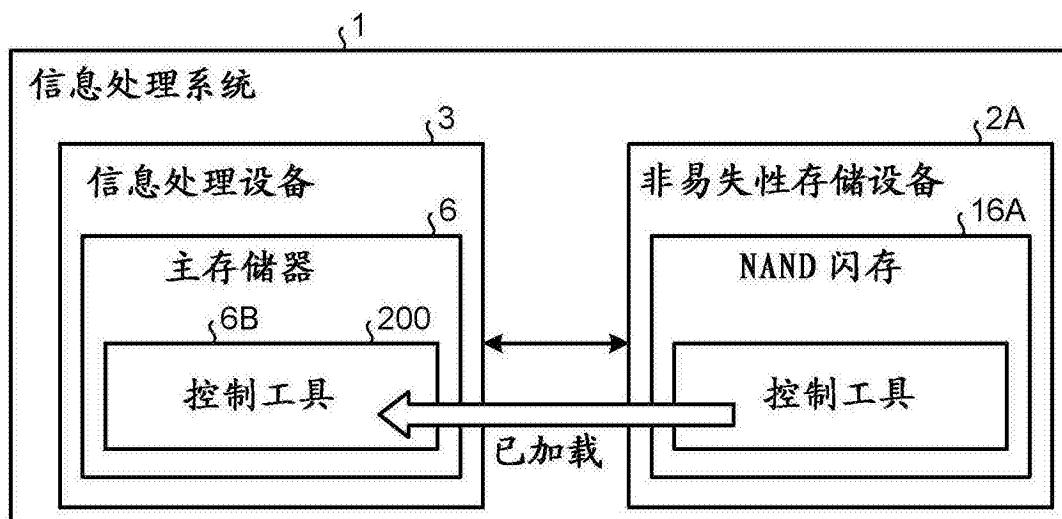


图8

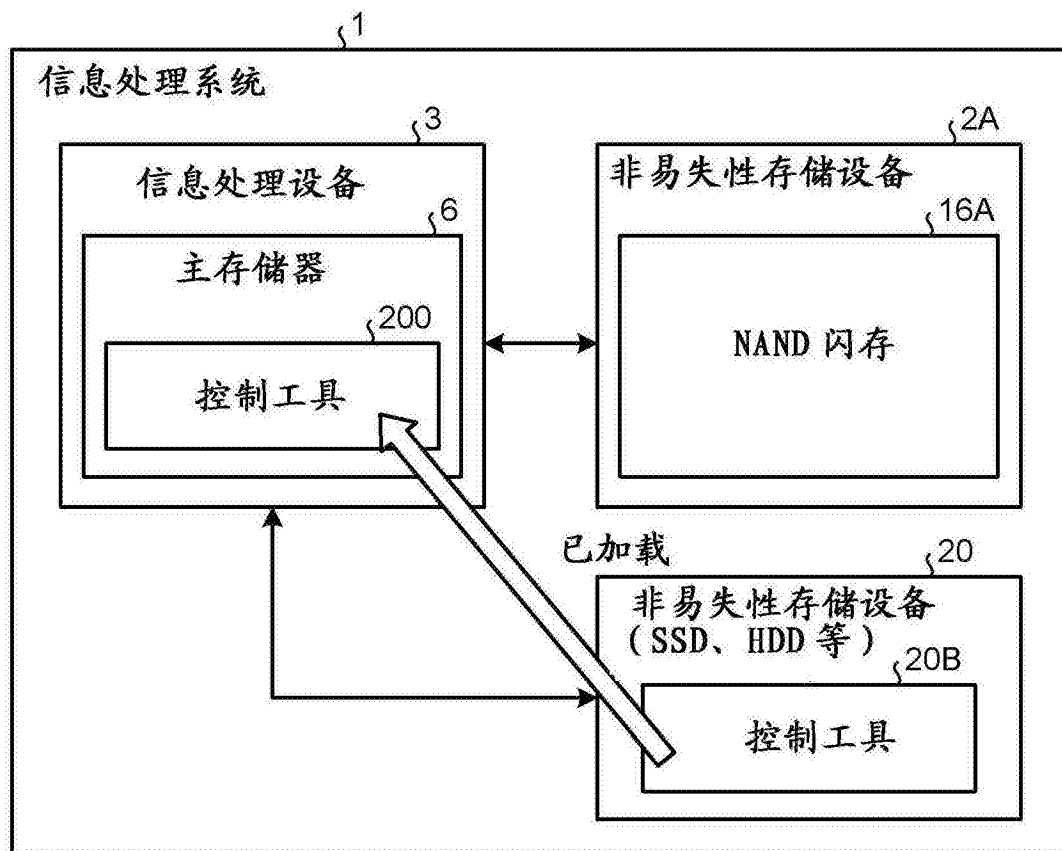


图9

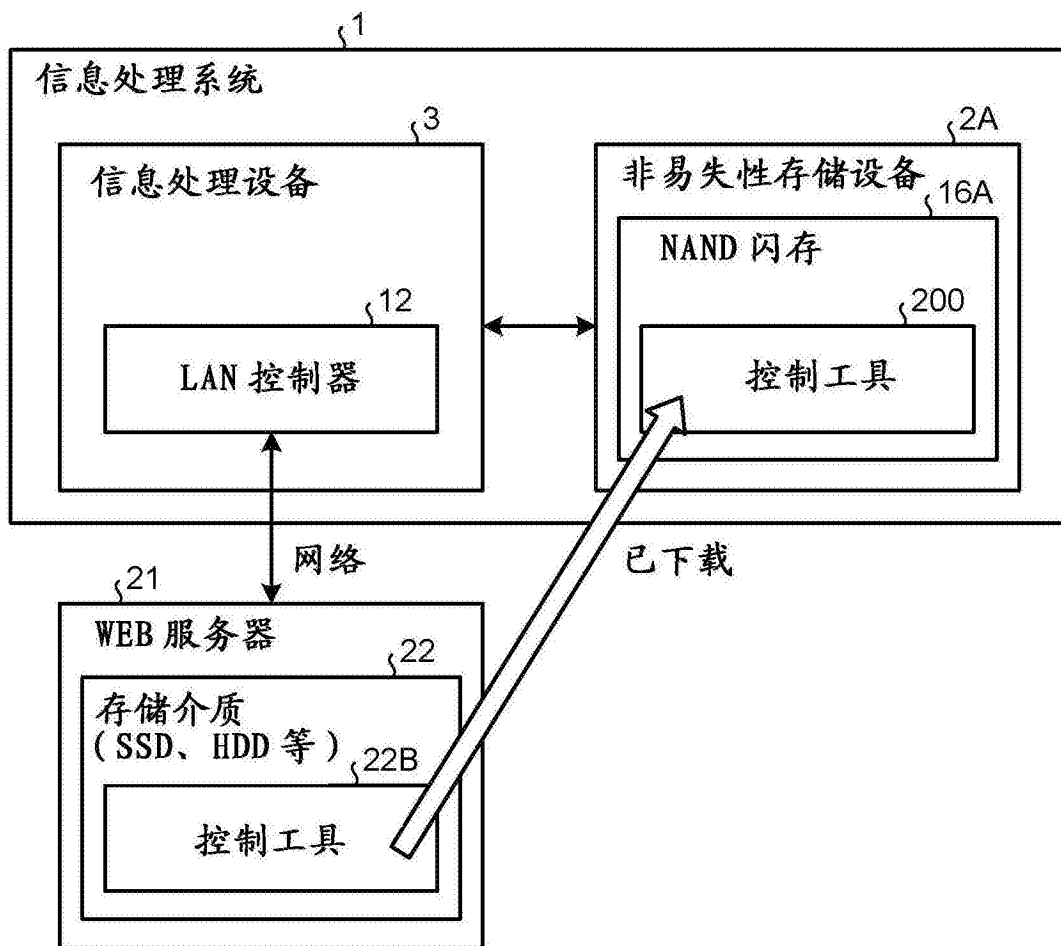


图10

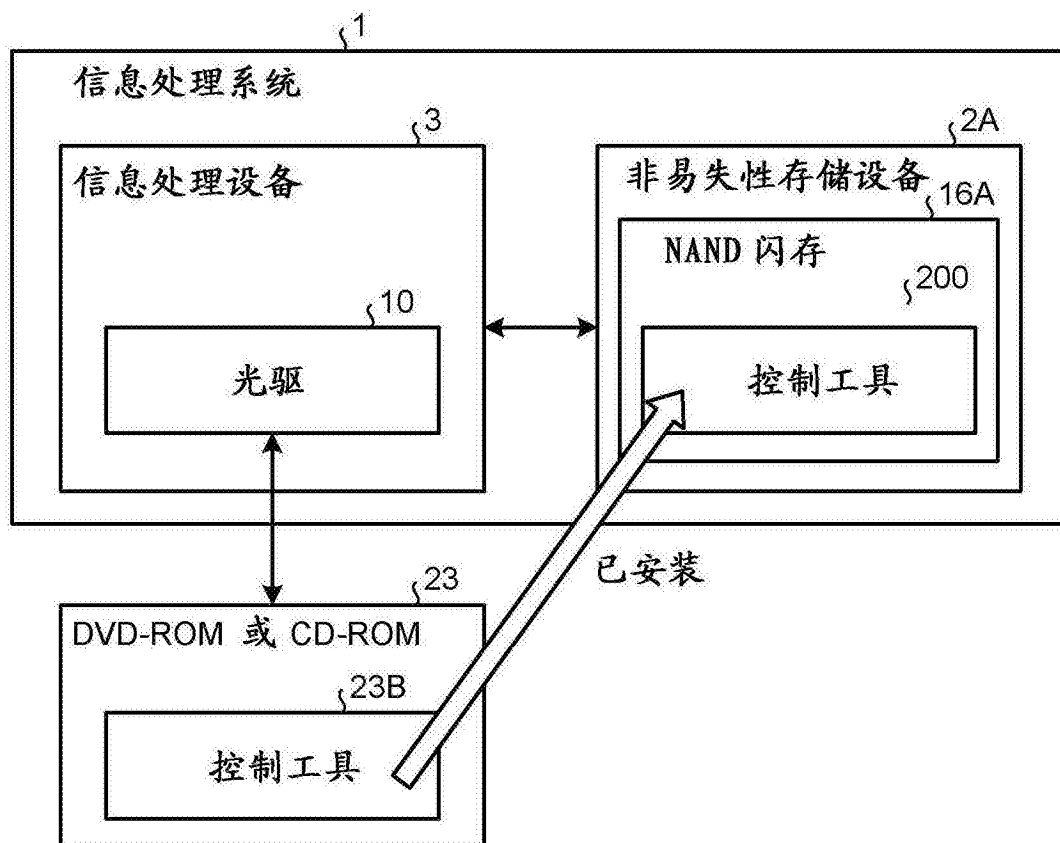


图11

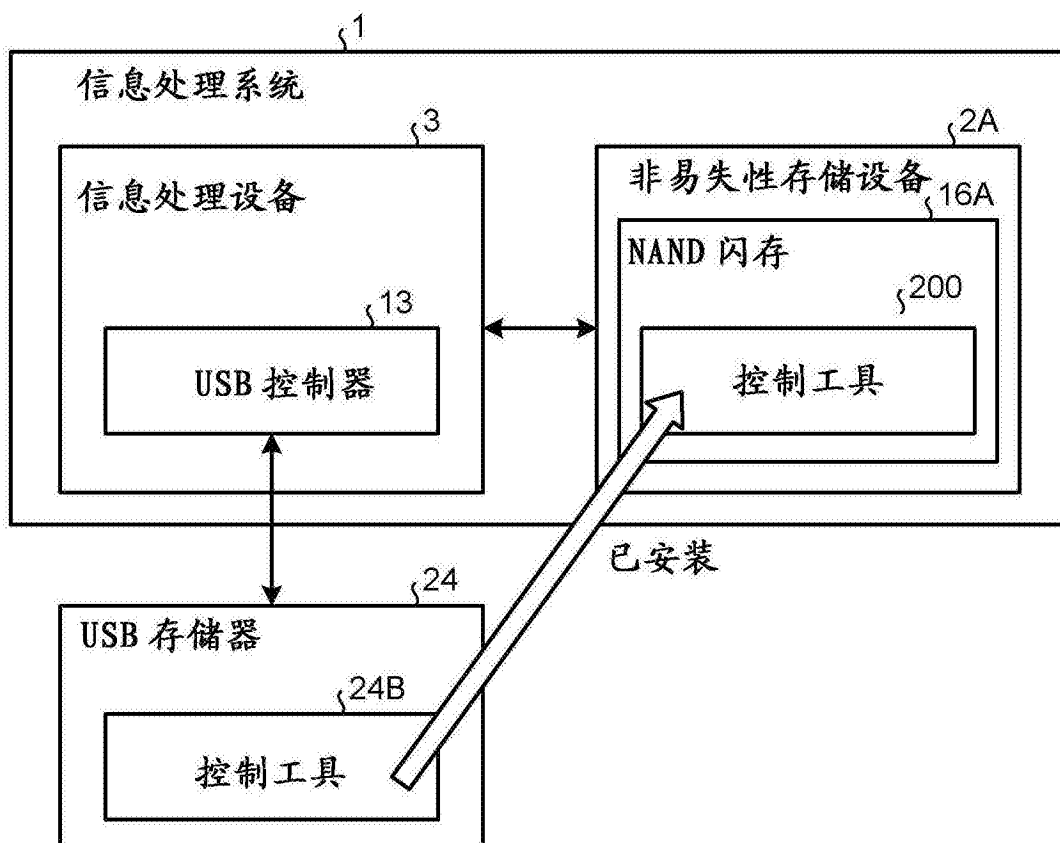


图12

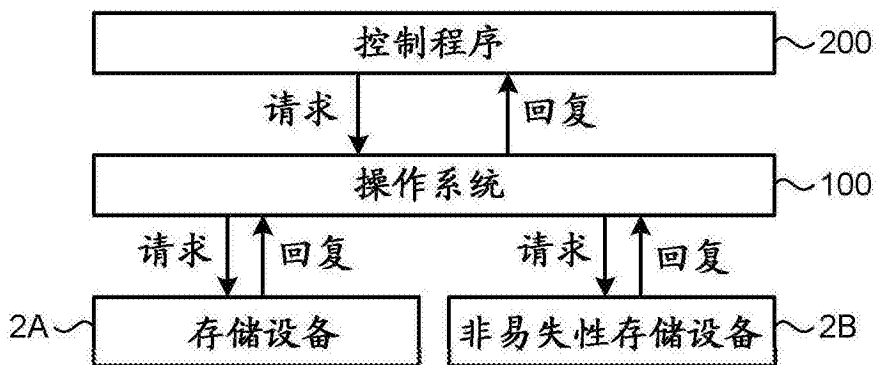


图13

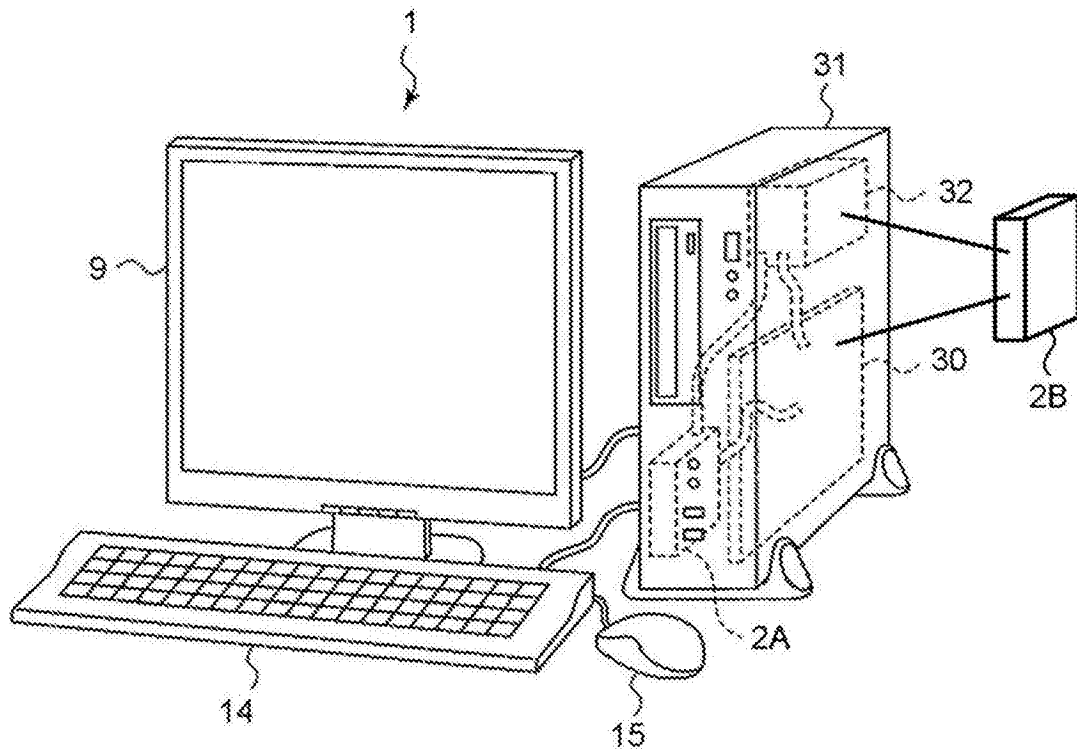


图14

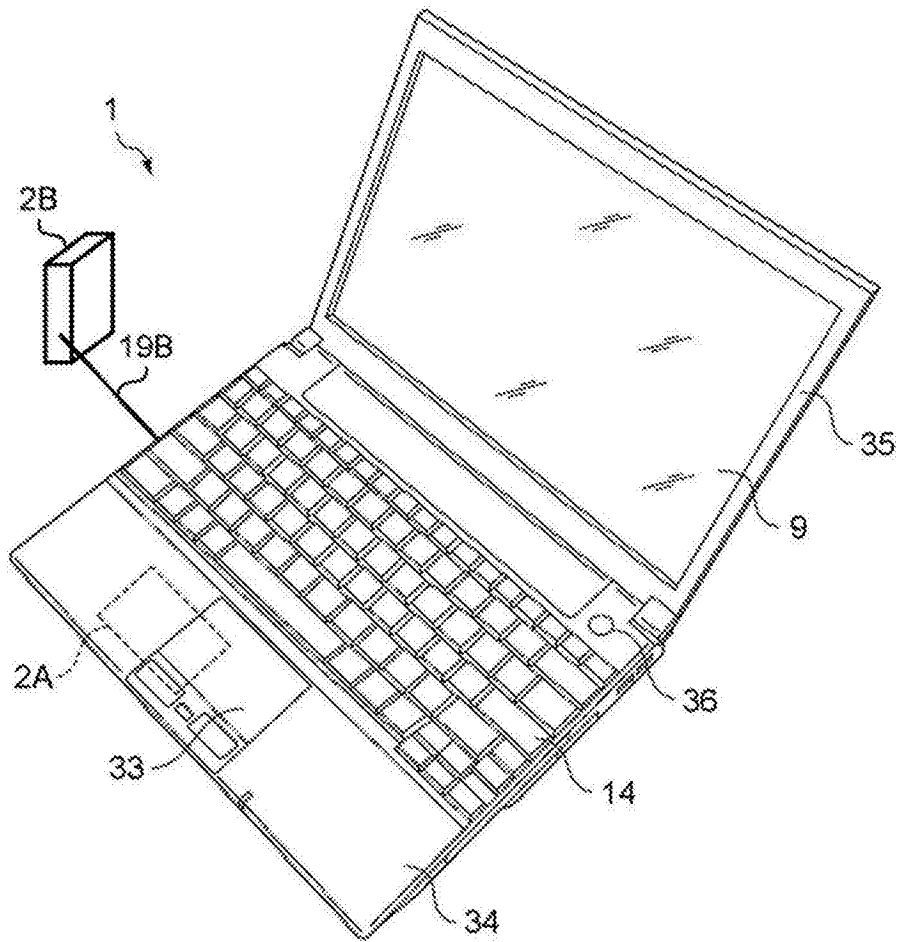


图15

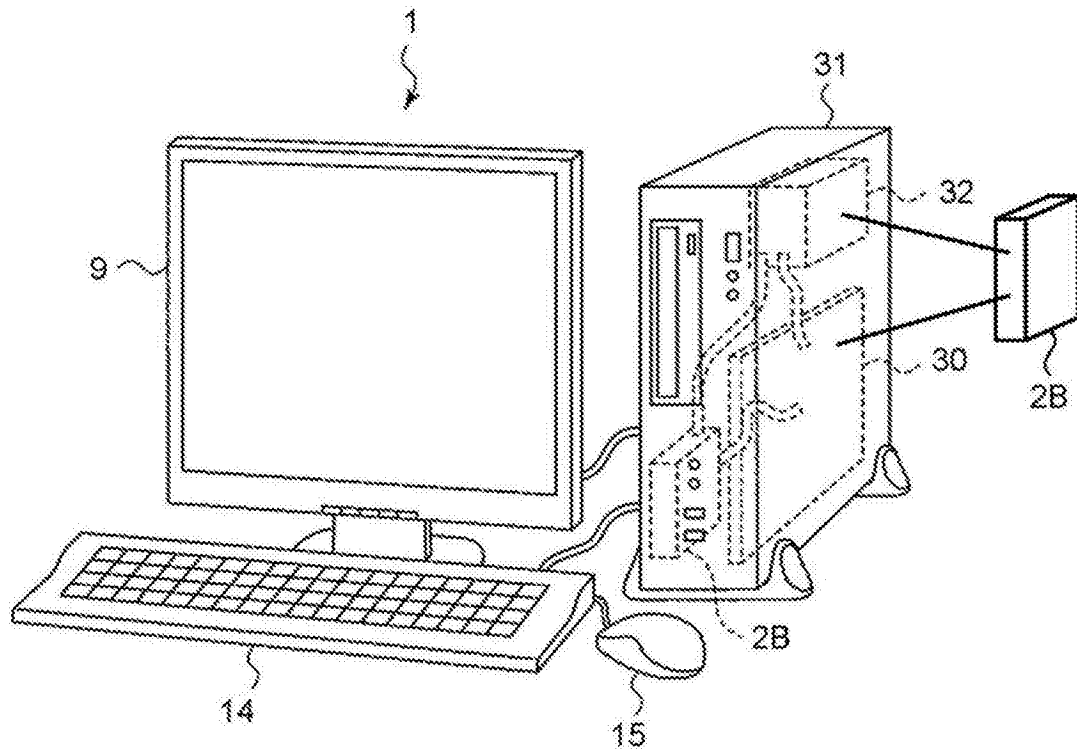


图16

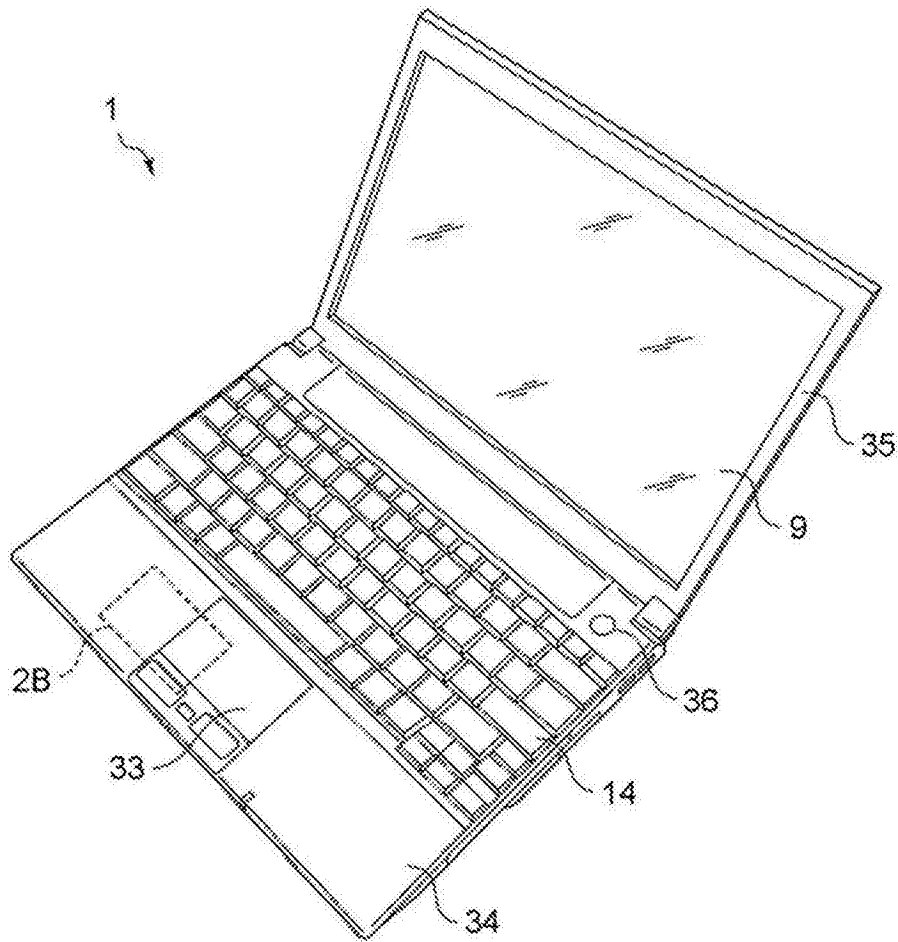


图17

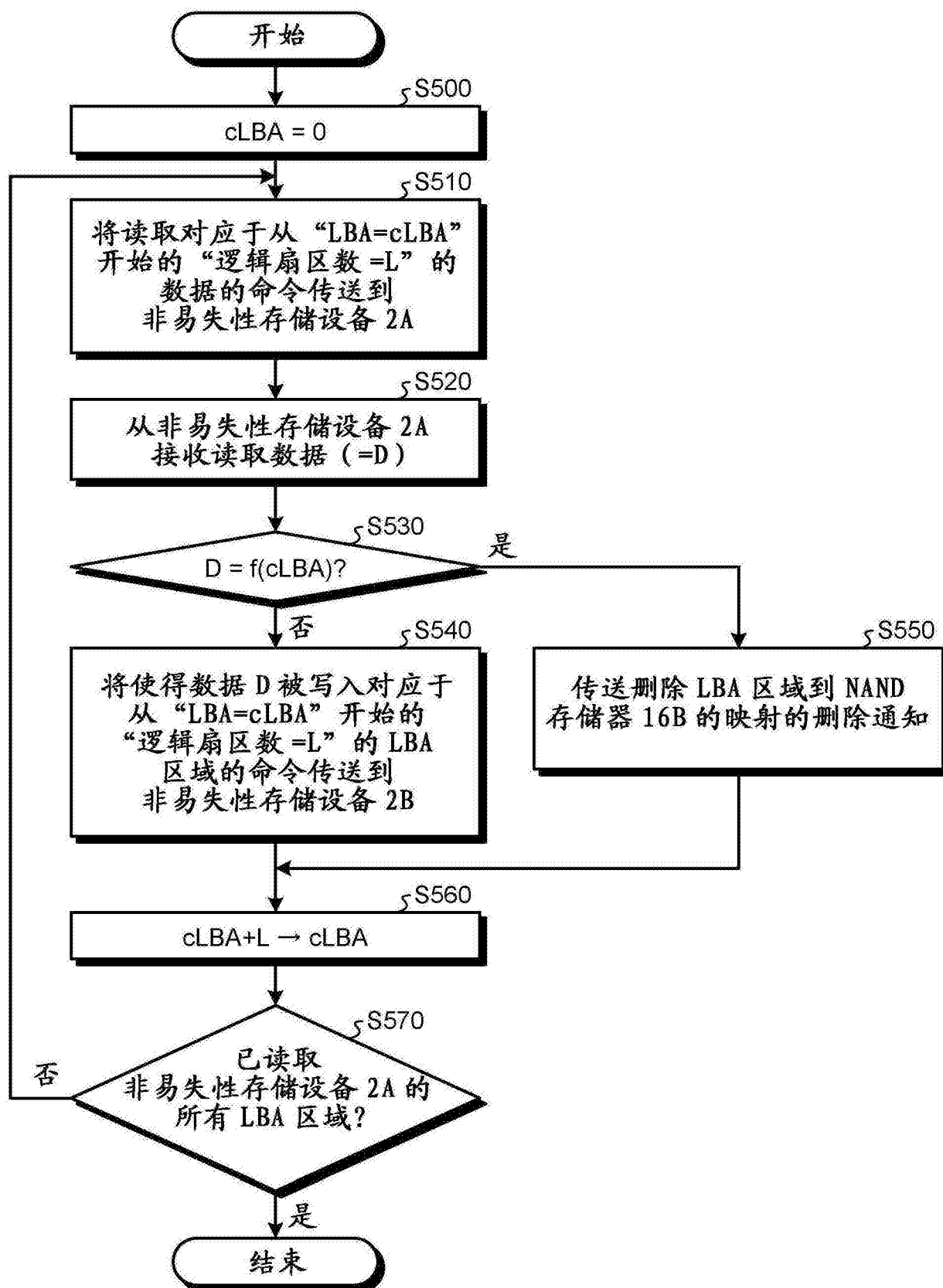


图18

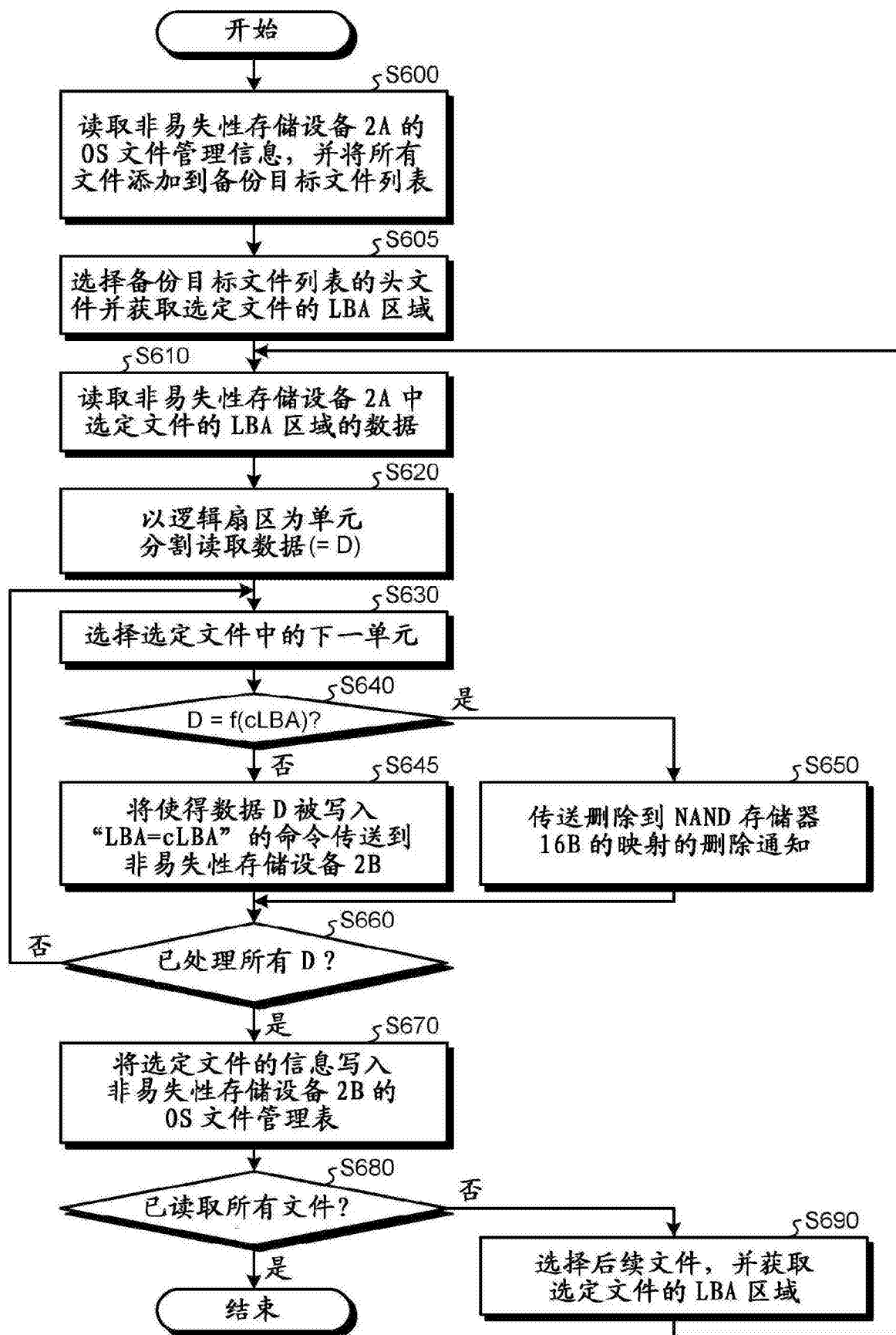


图19

LBA	数据
0000h	5ab543cd3763...
0001h	e564a641b763...
0002h	000000000000...
0003h	348957987908...
0004h	000000000000...
0005h	000000000000...
0006h	000000000000...
0007h	349058349058...
⋮	⋮
ba02h	954683f5484c...
ba03h	000000000000...
ba04h	000000000000...
ba05h	000000000000...
⋮	⋮
ffffh	000000000000...

1扇区 (= 512 字节)

全部为 0

图20

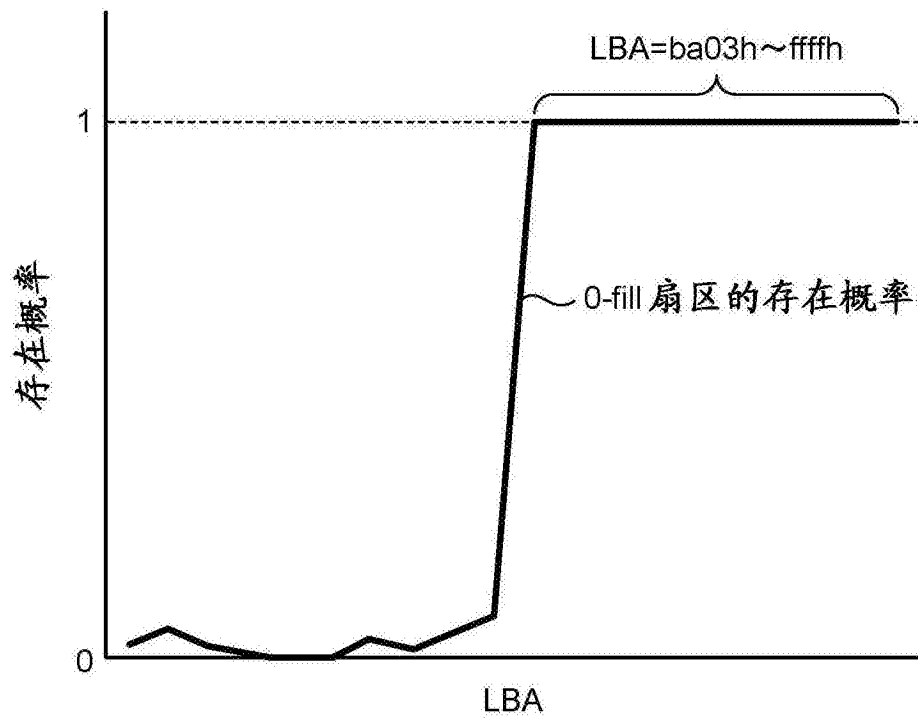


图21

LBA	数据	
0000h	5ab543cd3763...	0-fill 扇区的高存在概率很高
0001h	e564a641b763...	
0002h	000000000000...	
0003h	348957987908...	
0004h	000000000000...	
0005h	348957987908...	
0006h	000000000000...	
0007h	349058349058...	
⋮	⋮	1-fill 扇区的高存在概率很高
c000h	954683f5484c...	
c001h	ffffffffffffff...	
c002h	ffffffffffffff...	
c003h	349058349058...	
⋮	⋮	
ffffh	ffffffffffffff...	

图22

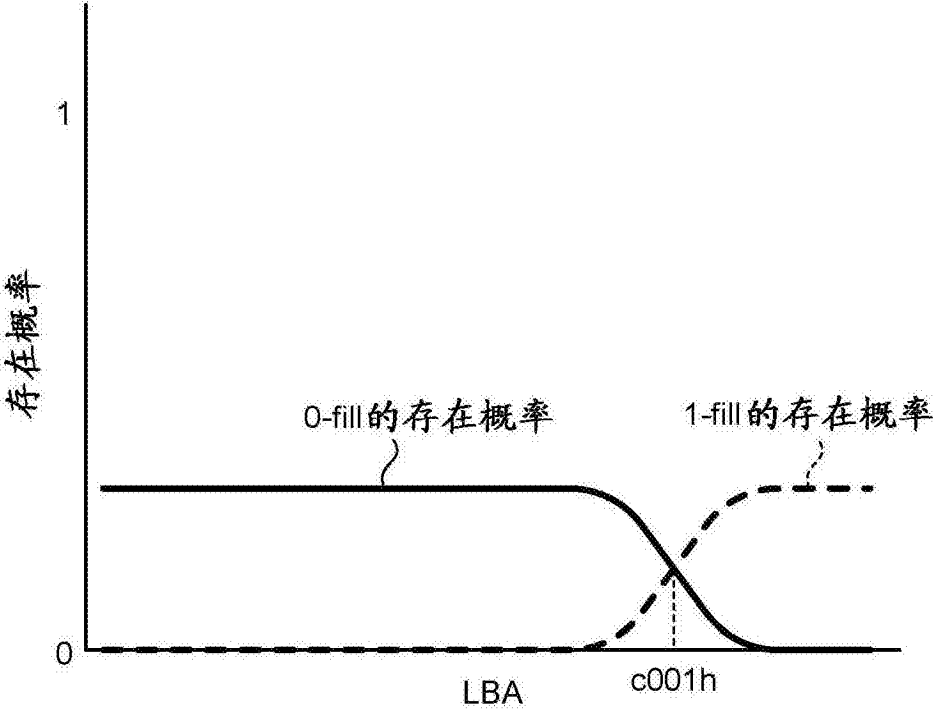


图23

LBA	数据	
0000h	5ab543cd3763...	0-fill 扇区的高存在概率很高
0001h	e564a641b763...	
0002h	000000000000...	
0003h	348957987908...	
0004h	000000000000...	
0005h	348957987908...	
0006h	000000000000...	
0007h	349058349058...	
⋮	⋮	320a-fill 扇区的高存在概率很高
8000h	954683f5484c...	
8001h	320a320a320a...	
8002h	320a320a320a...	
8003h	320a320a320a...	
⋮	⋮	1-fill 扇区的高存在概率很高
c000h	954683f5484c...	
c001h	ffffffffffffff...	
c002h	ffffffffffffff...	
c003h	349058349058...	
⋮	⋮	
ffffh	ffffffffffffff...	

图24

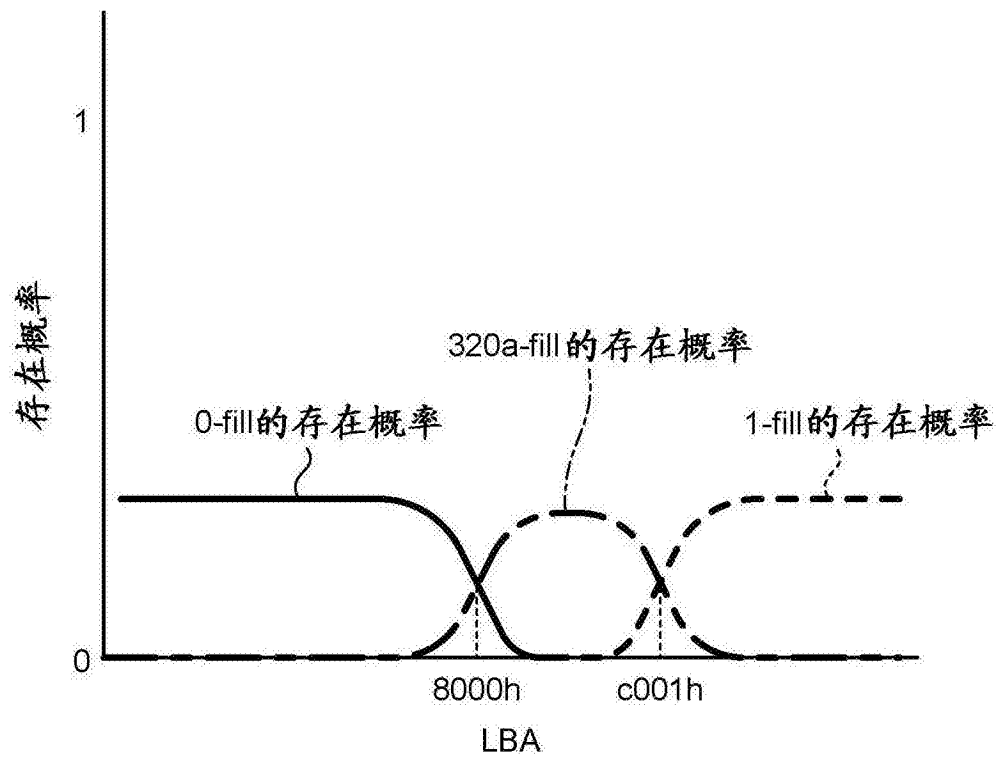


图25

LBA	数据	
0000h	5ab543cd3763...65183	
0001h	e564a641b763...ab552	
0002h	000000000000...00002	关联
0003h	000000000000...00003	关联
0004h	000000000000...00004	关联
0005h	000000000000...00005	关联
0006h	000000000000...00006	关联
0007h	349058349058...	
⋮	⋮	
ba02h	954683f5484c...86213	
ba03h	000000000000...0ba03	关联
ba04h	000000000000...0ba04	关联
ba05h	000000000000...0ba05	关联
⋮	⋮	
ffffh	000000000000...0ffff	关联

图26

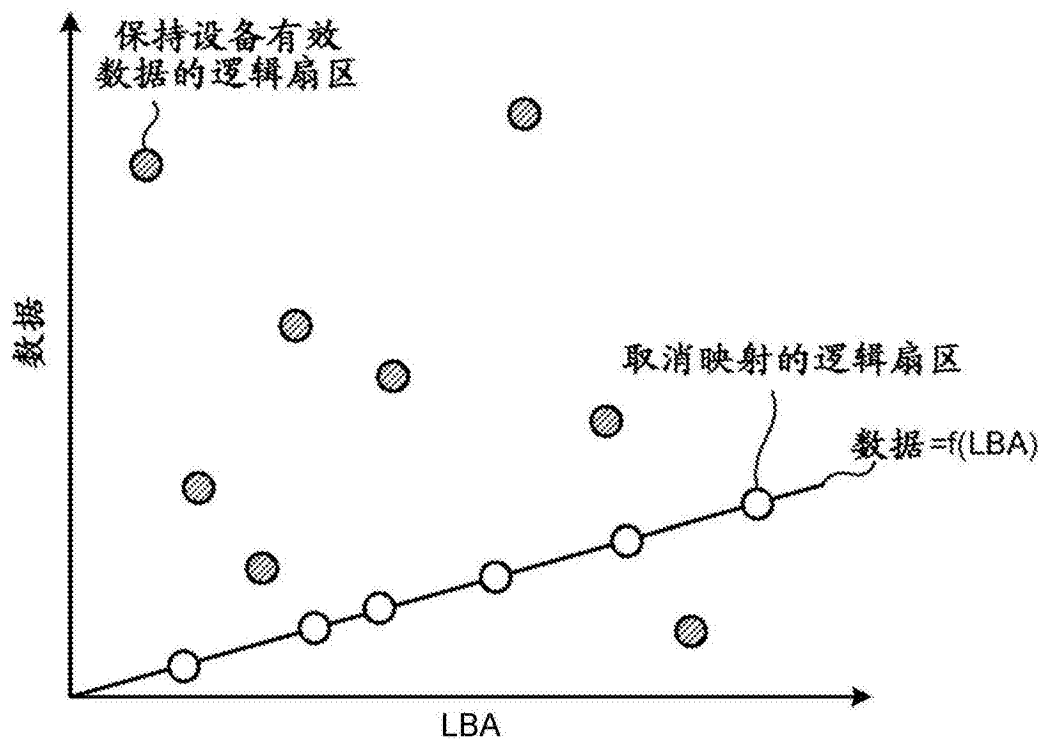


图27

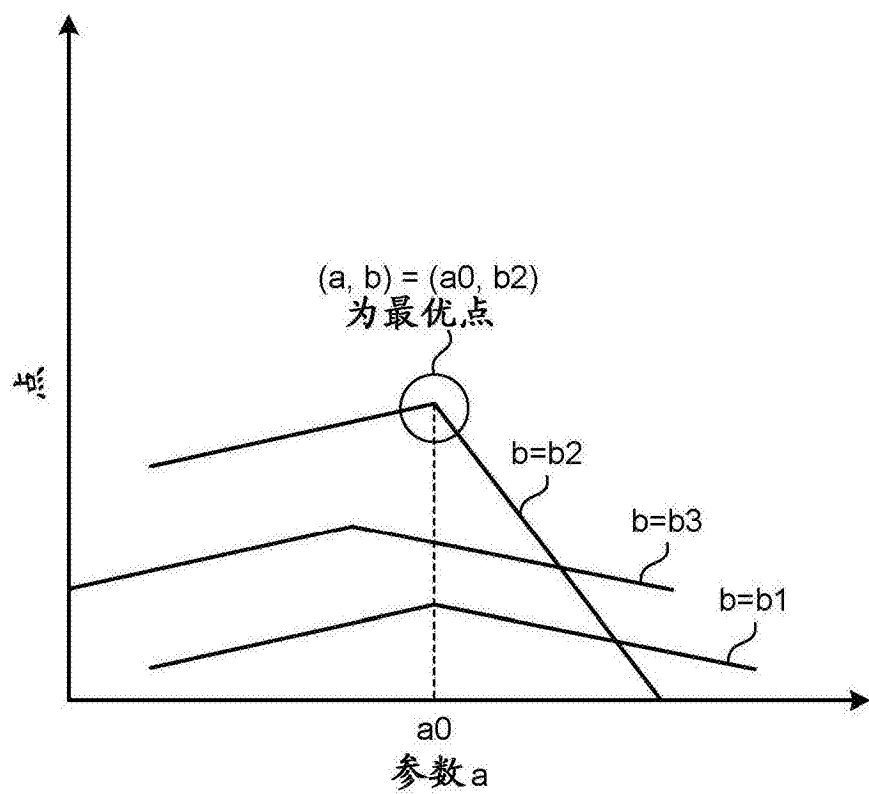


图28

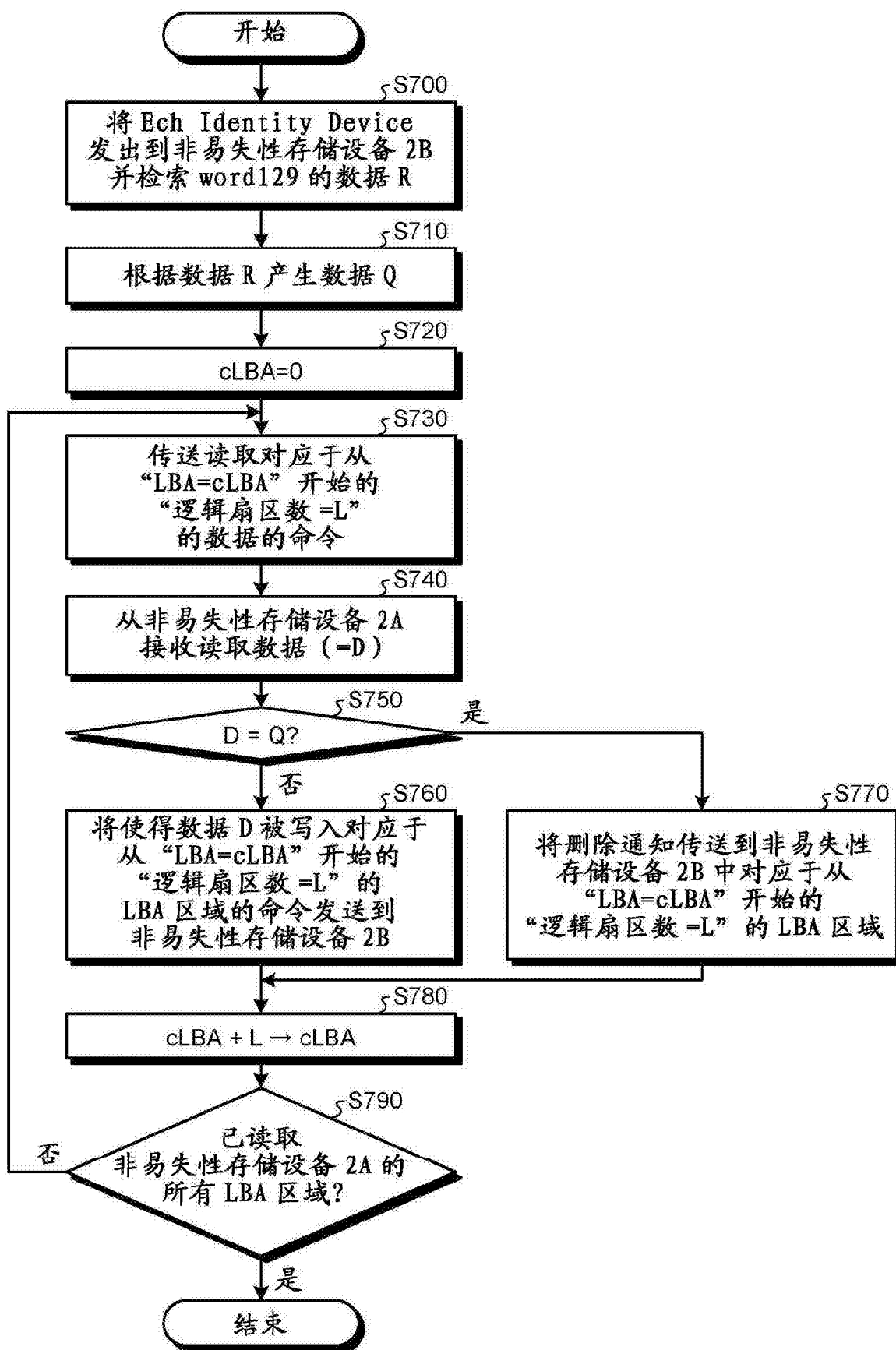


图29

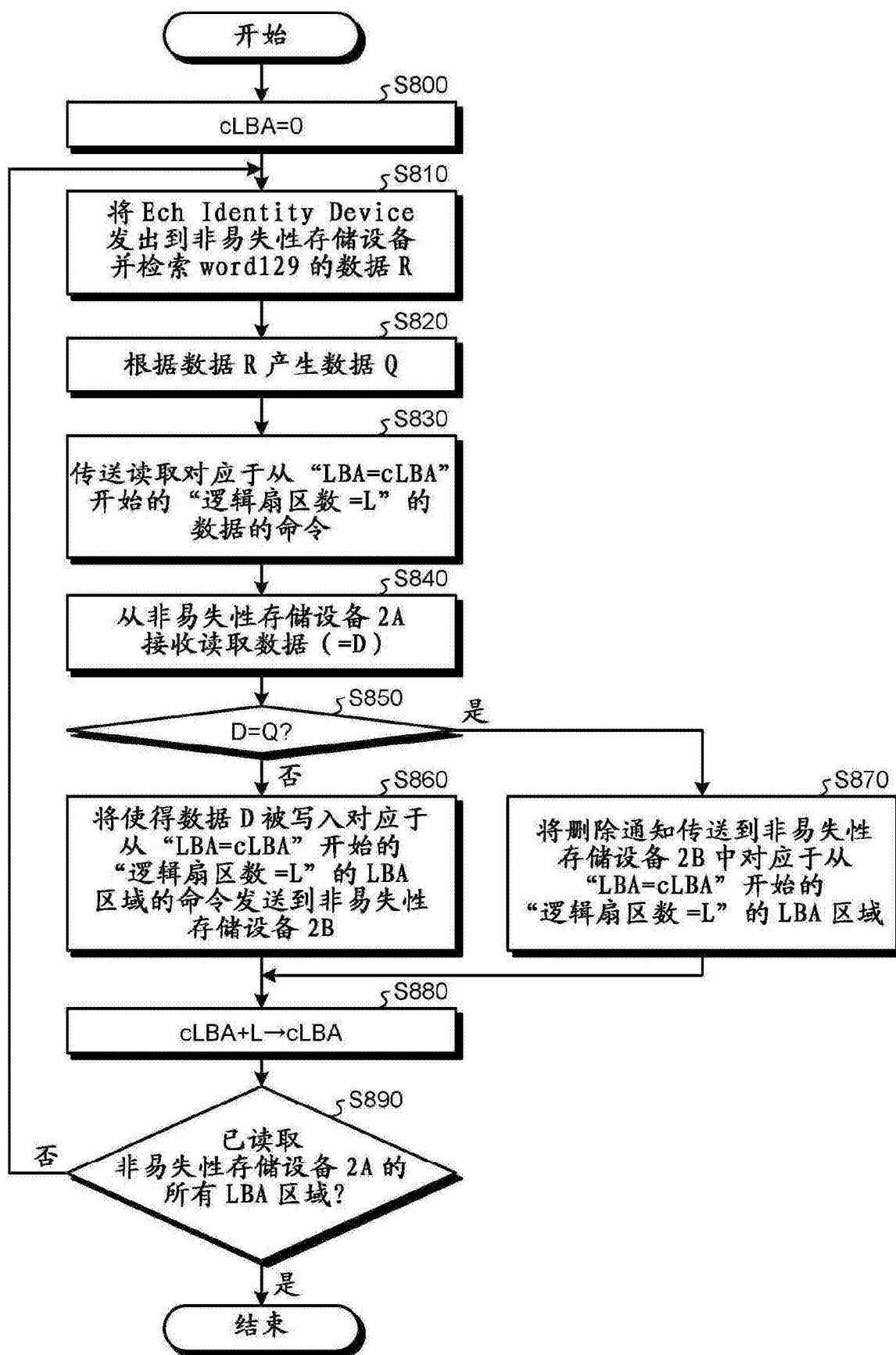


图30

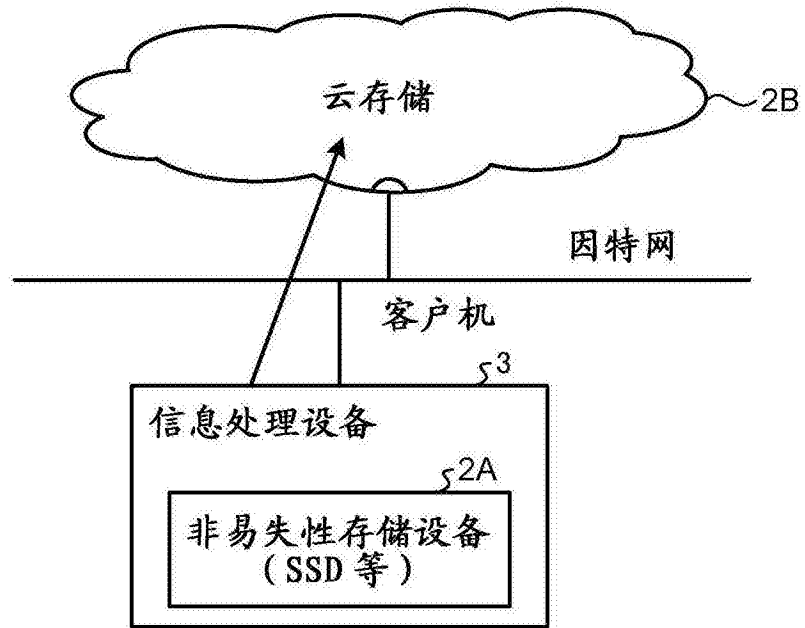


图31

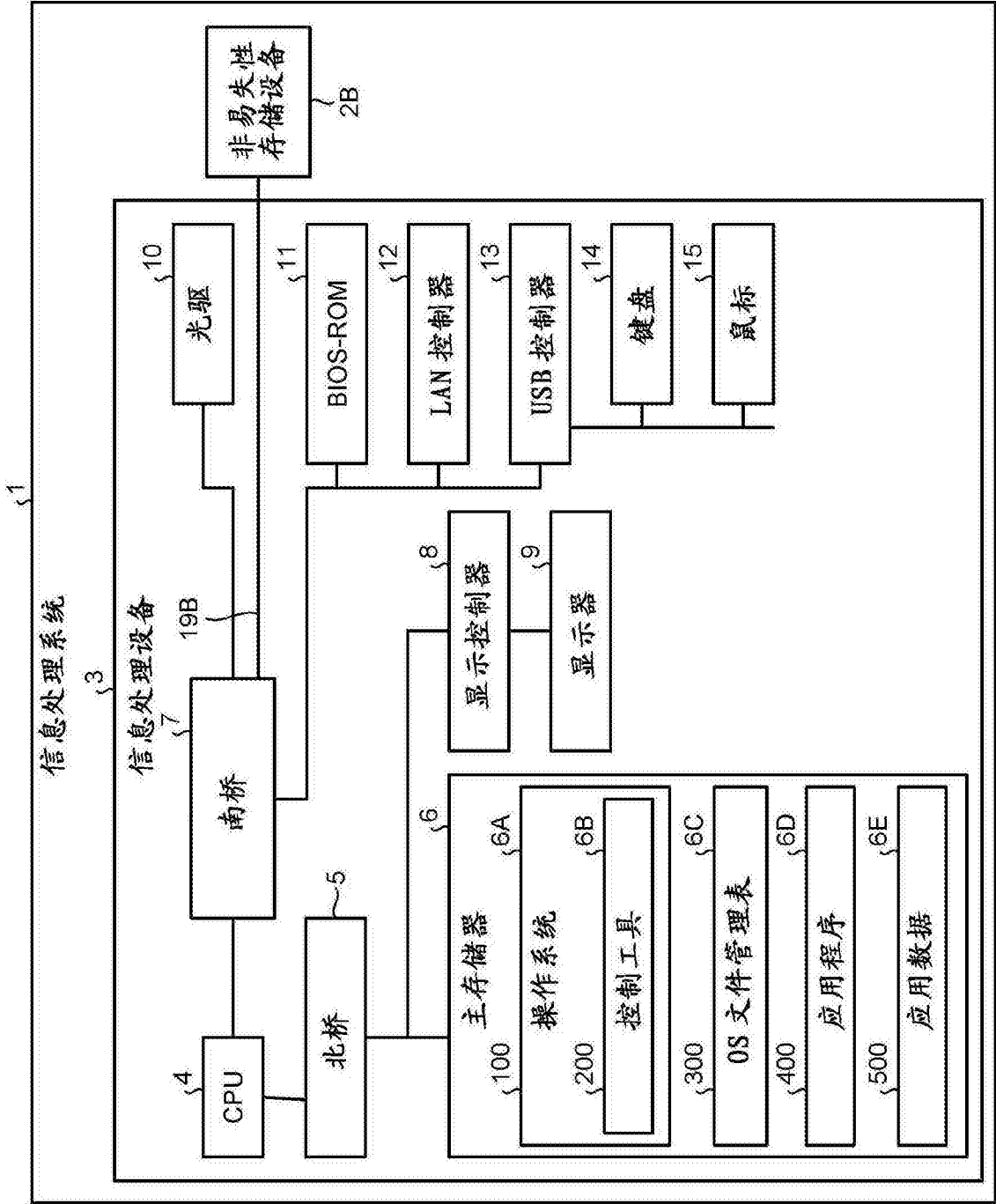


图32

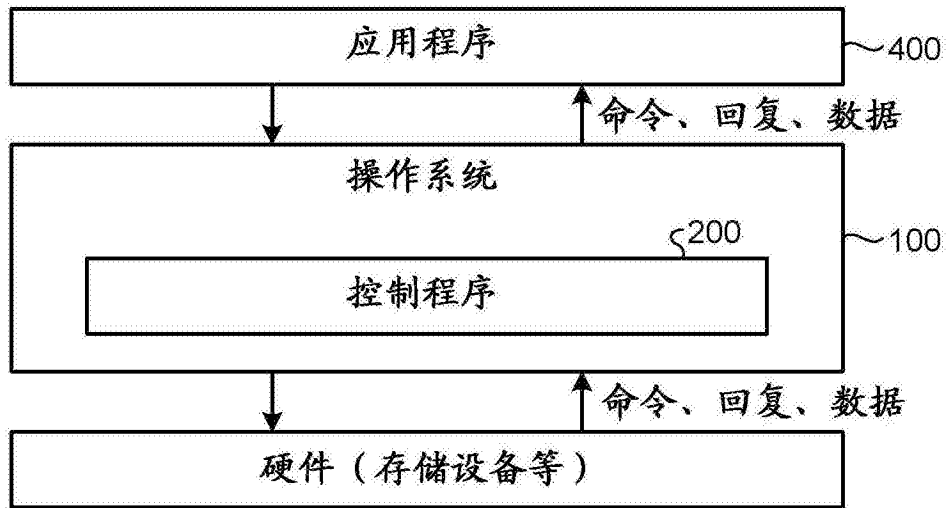


图33

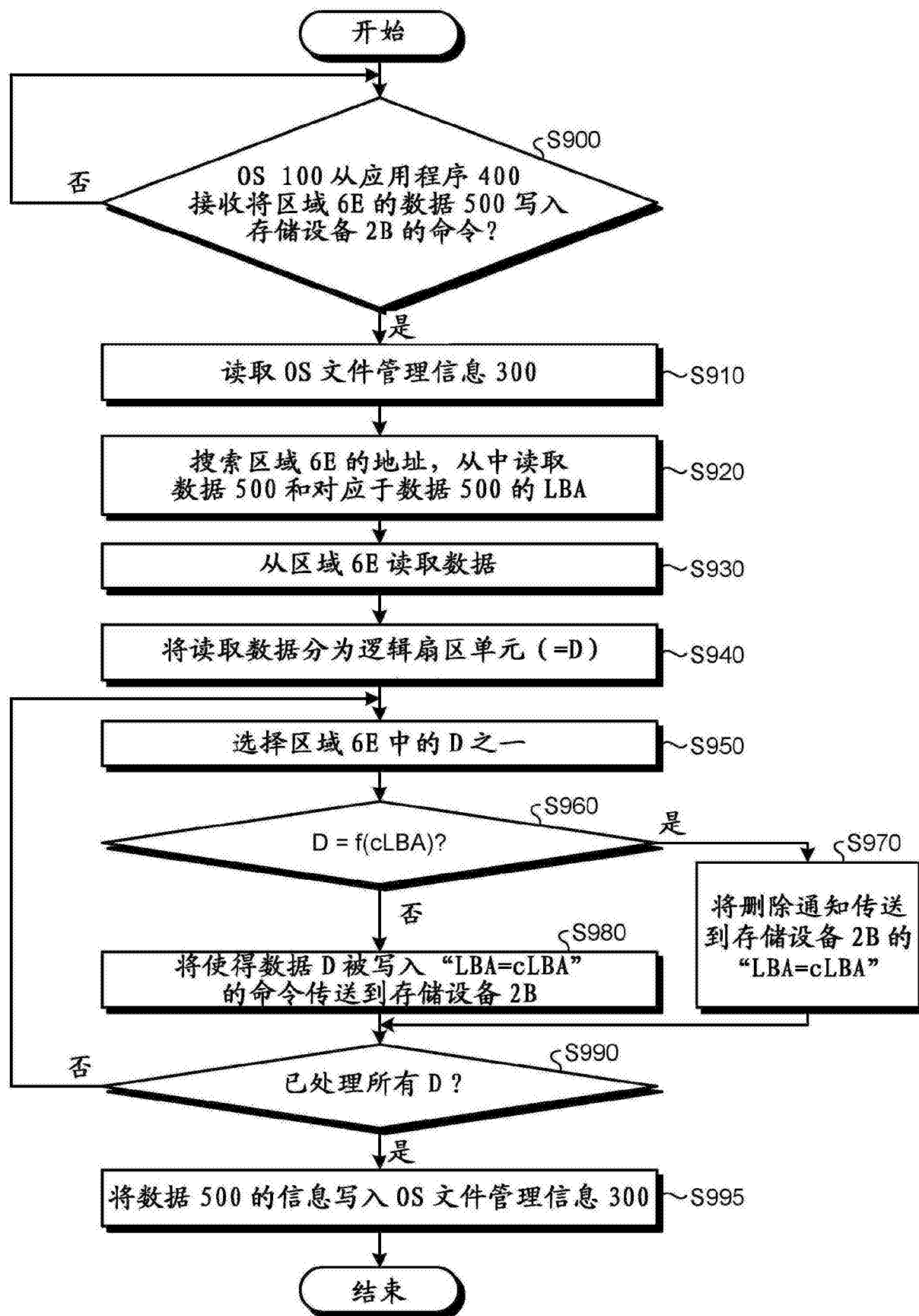


图34



图35

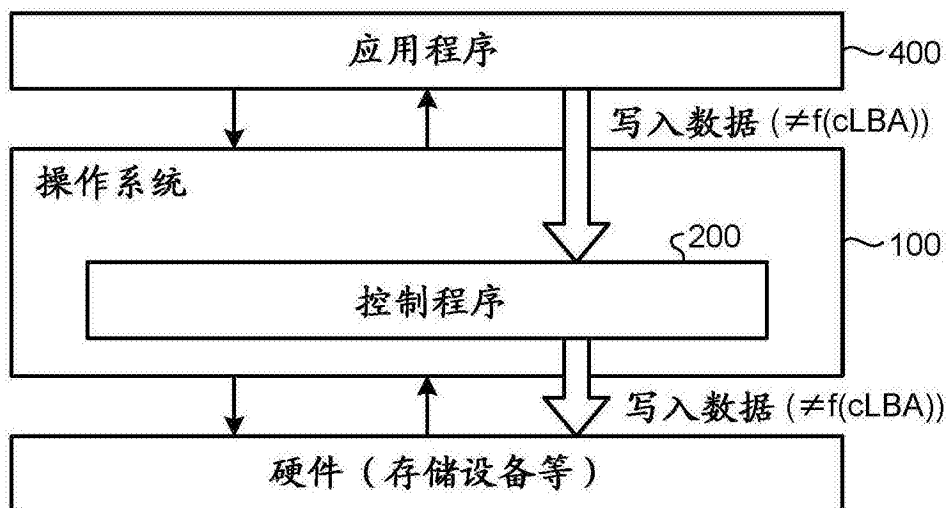


图36

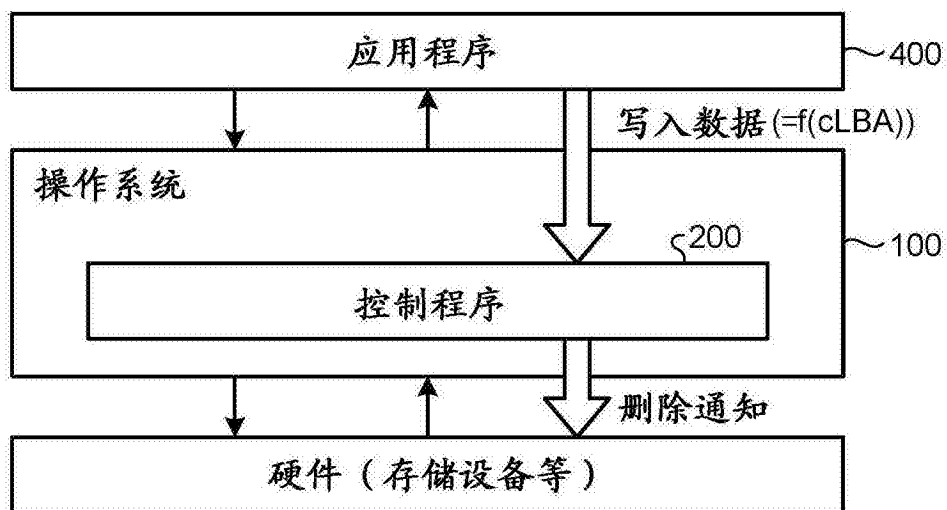


图37

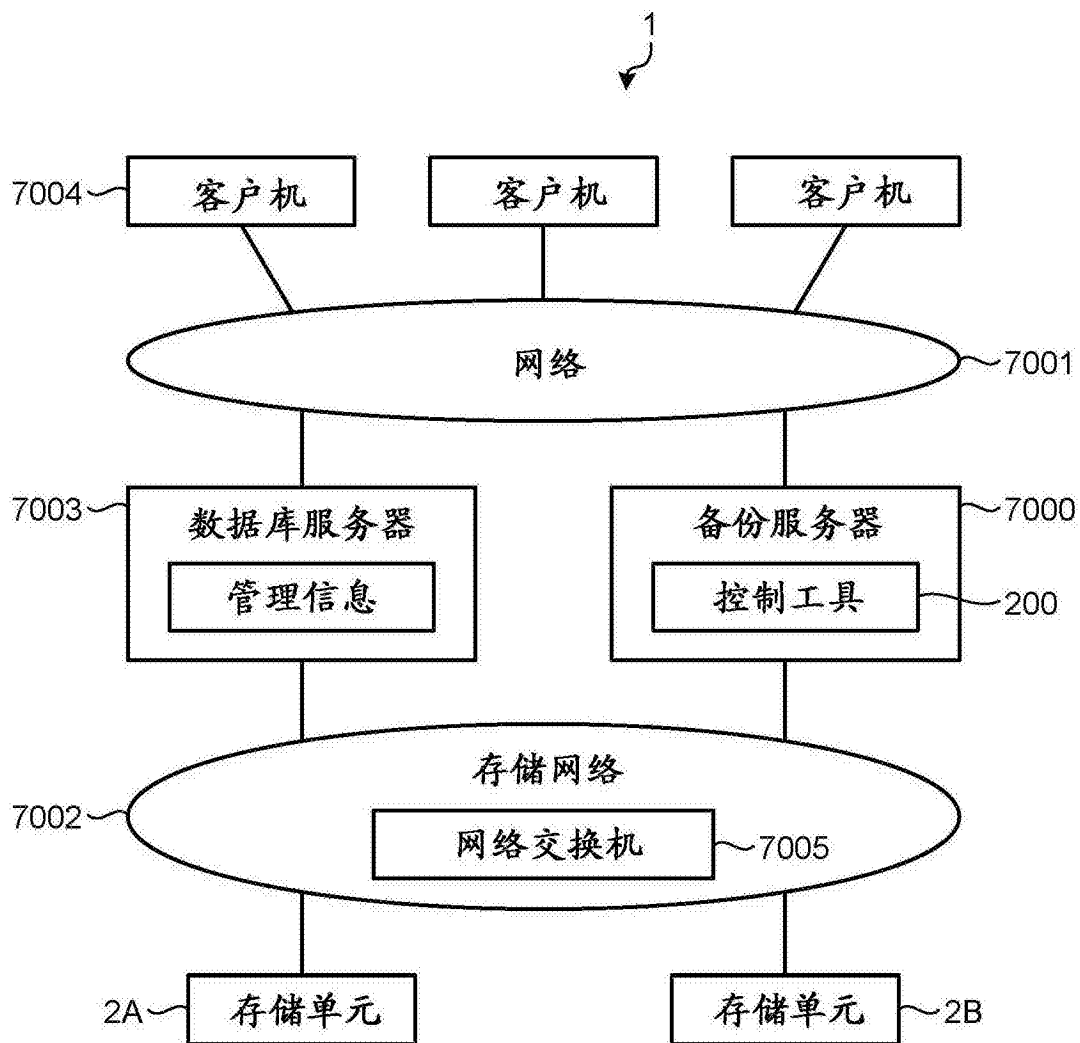


图38

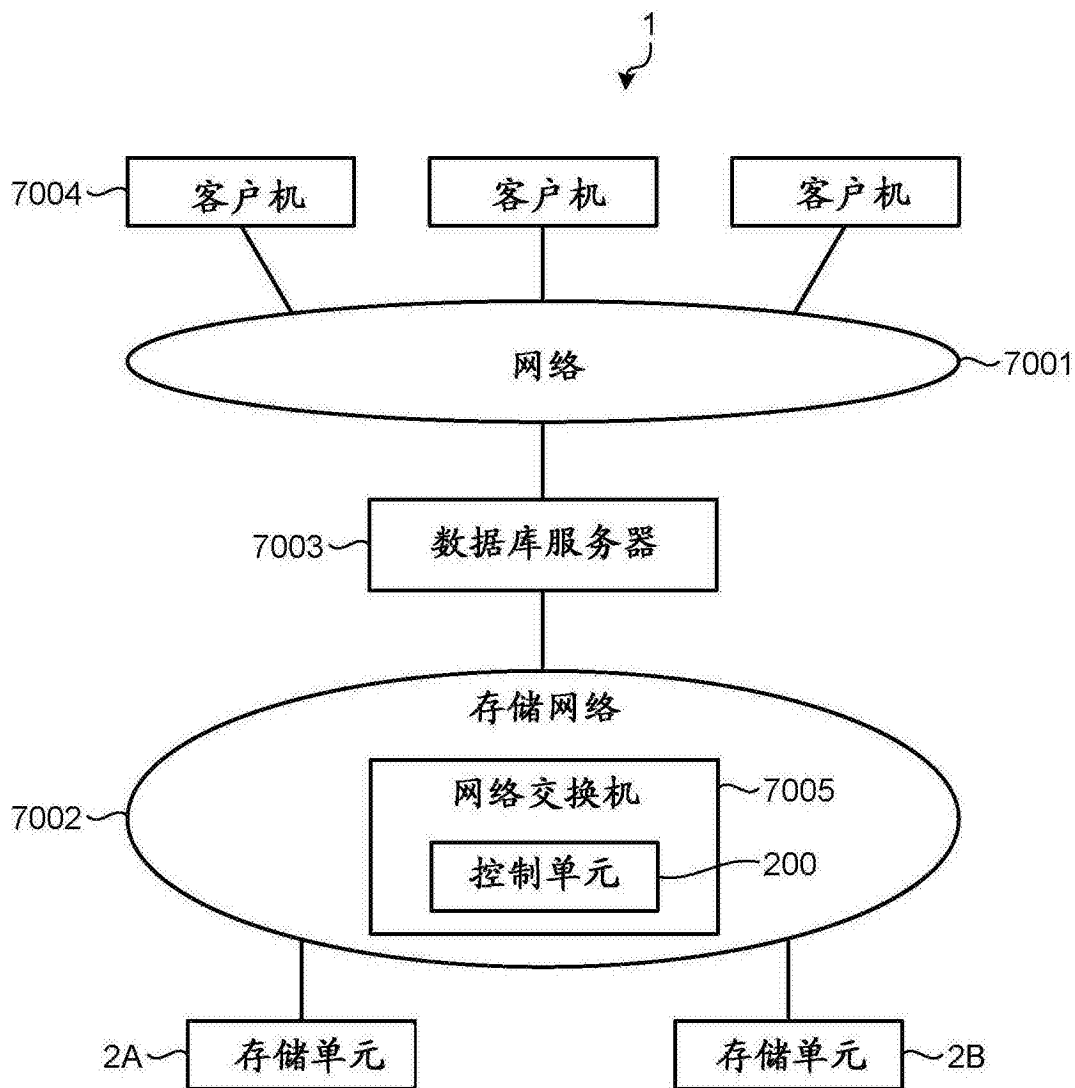


图39

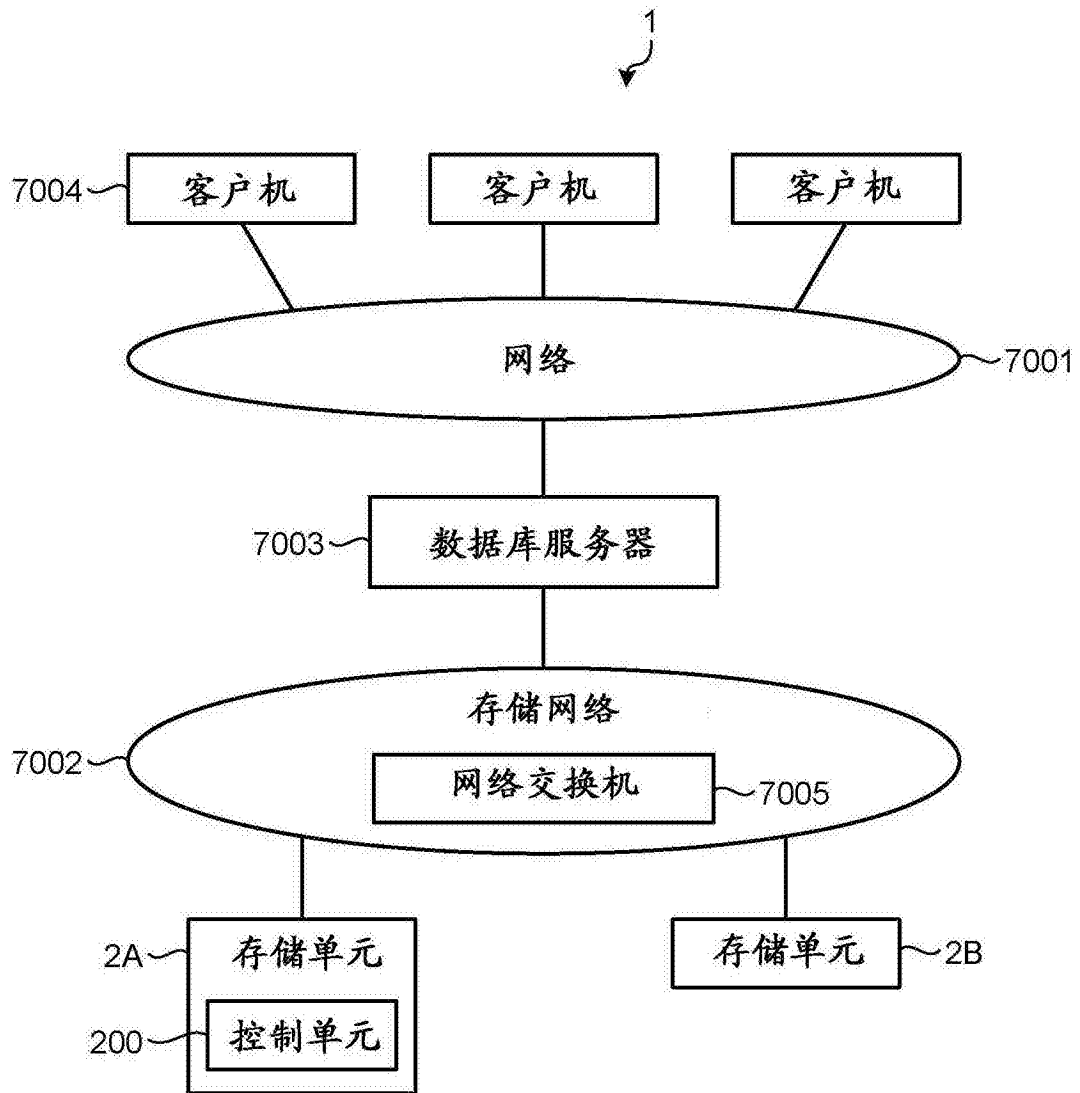


图40

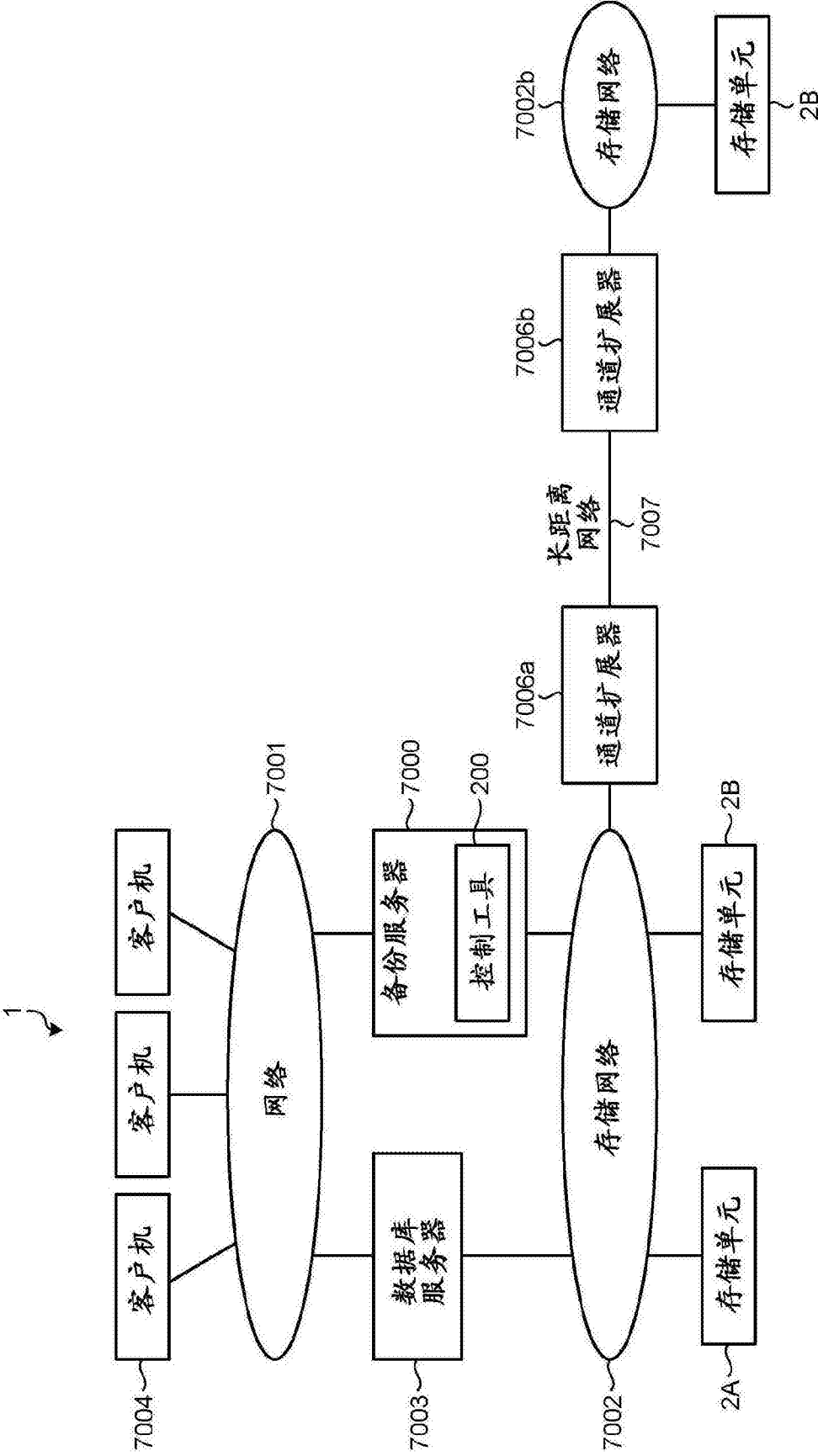


图41

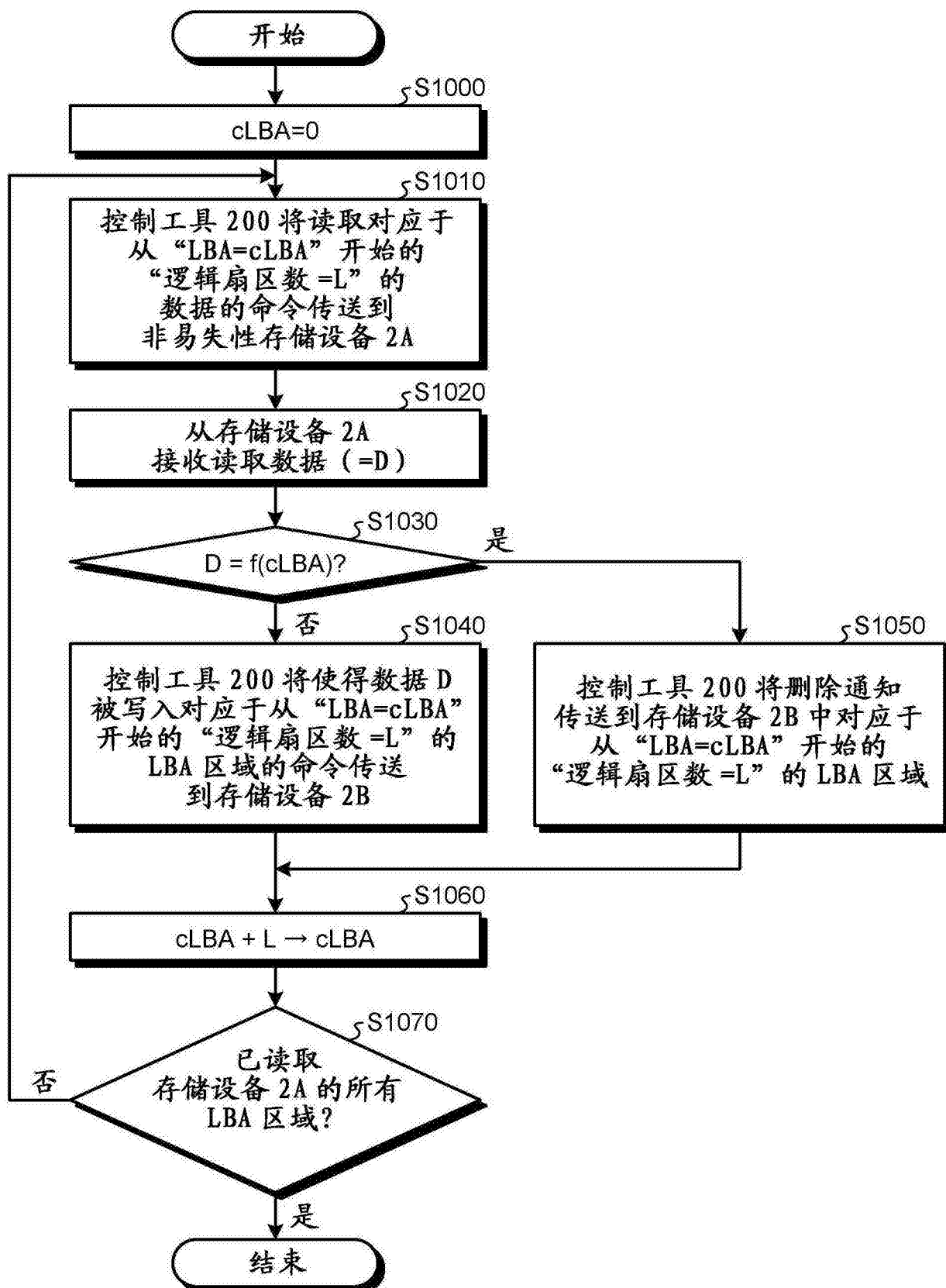


图42

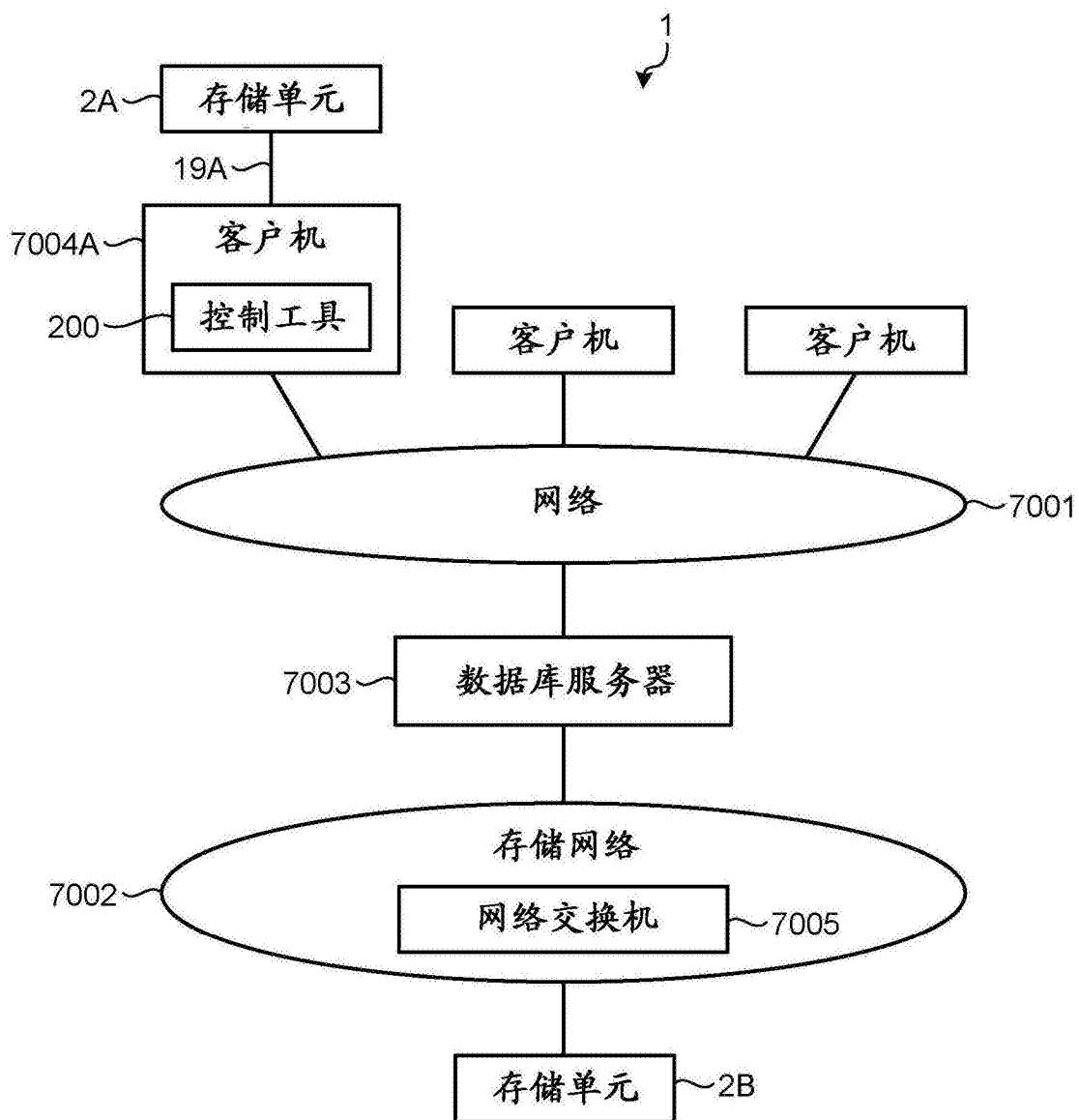


图43

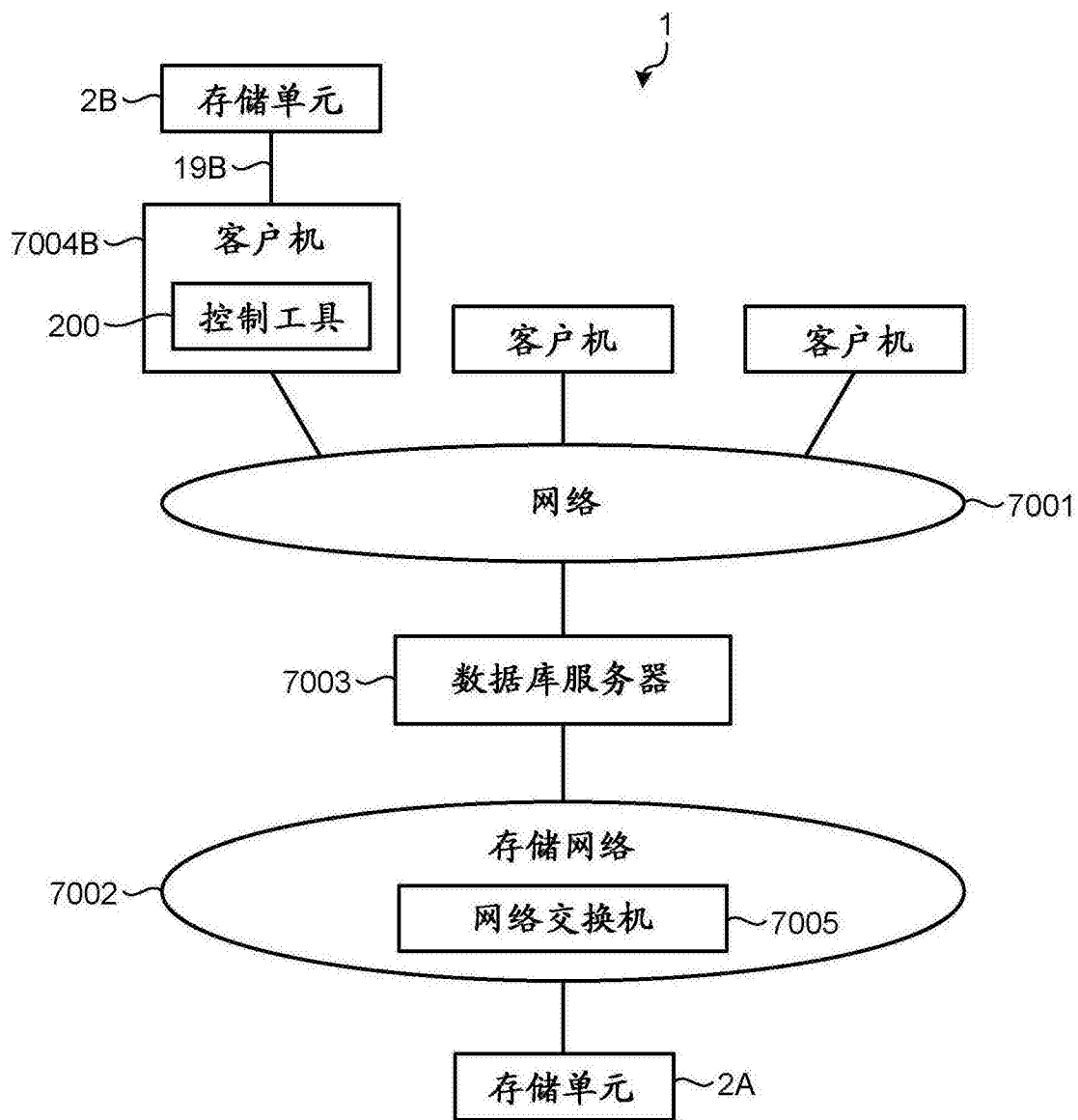


图44

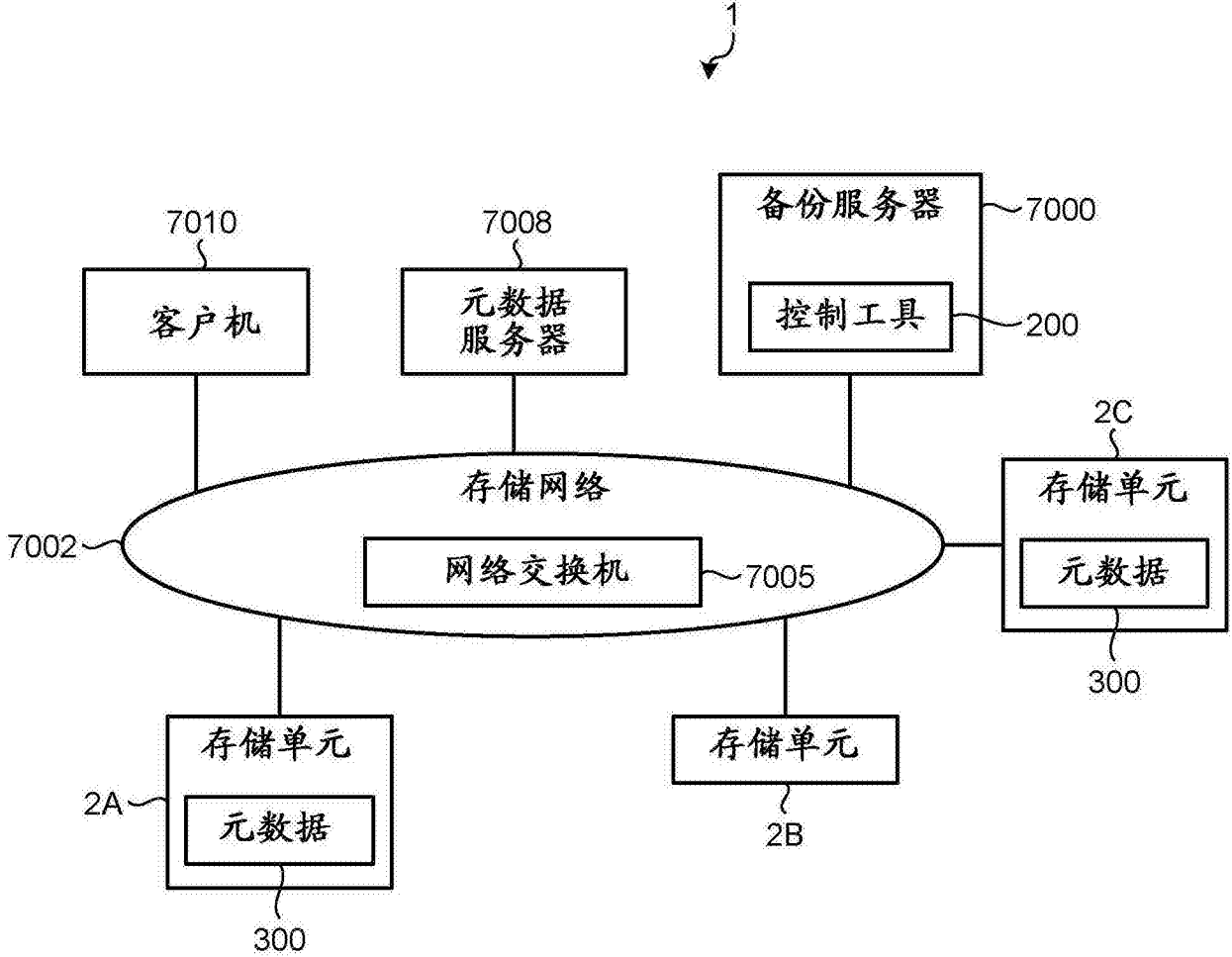


图45

300

元数据						
文件标识 ID	存储单元 标识 ID	LBA	扇区计数	所有者	更新 时间	其它属性 信息
***	***	***	***	***	***	***
***	***	***	***	***	***	***
***	***	***	***	***	***	***
***	***	***	***	***	***	***
***	***	***	***	***	***	***

图46

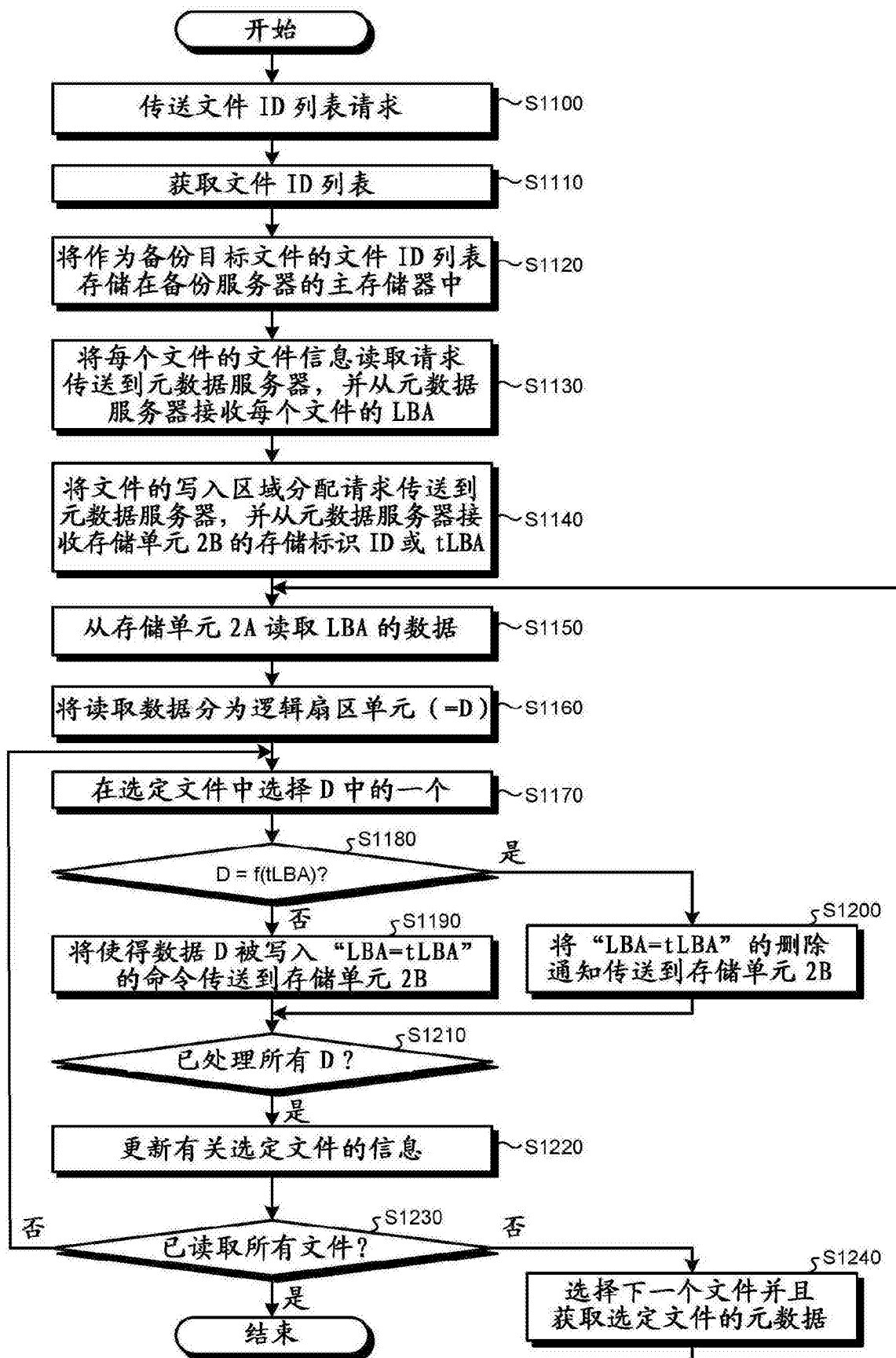


图47

文件 0 的 ID
文件 1 的 ID
文件 2 的 ID
文件 3 的 ID
文件 4 的 ID
文件 5 的 ID
文件 6 的 ID
...

图48

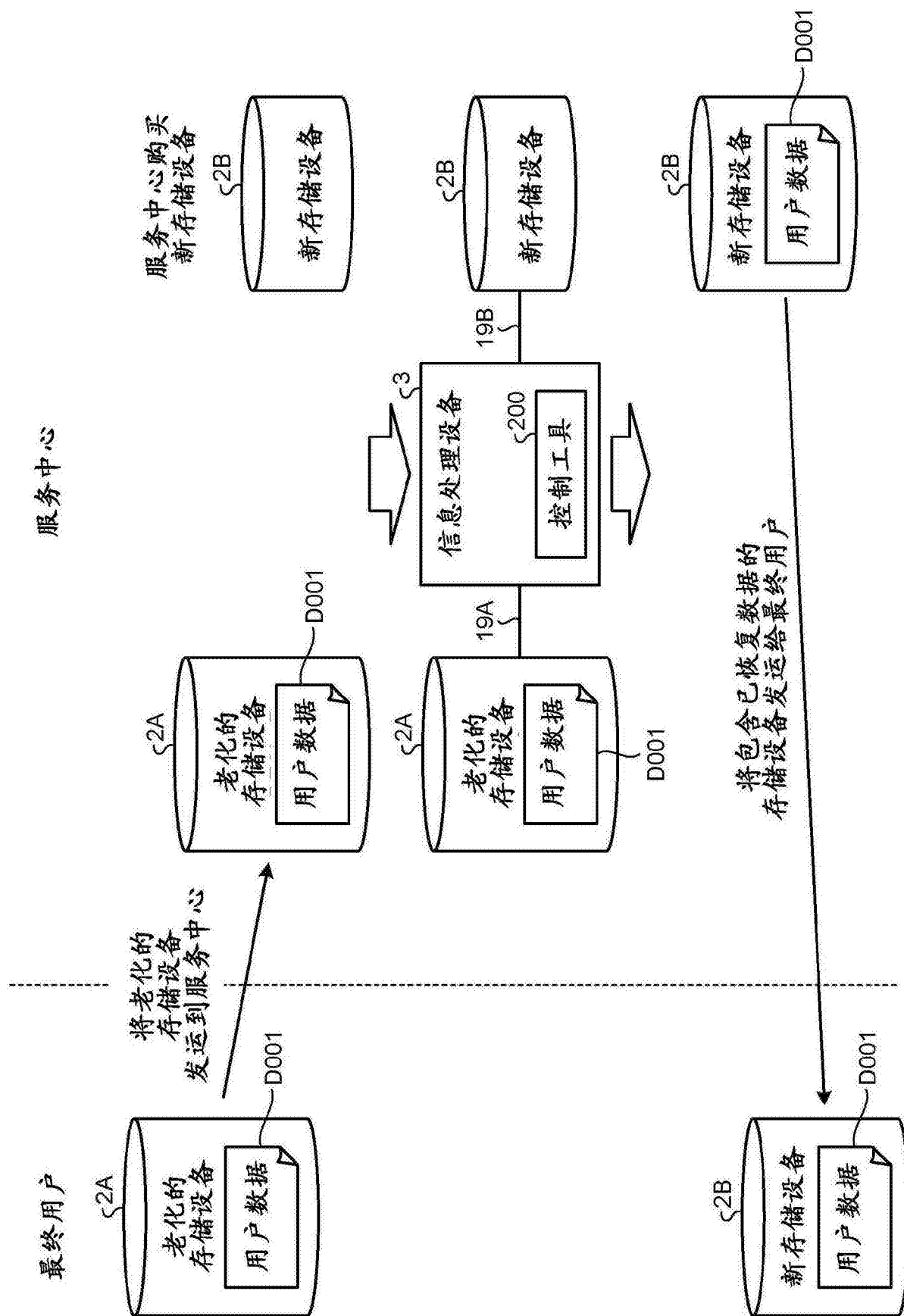


图49