



(12) 发明专利

(10) 授权公告号 CN 101498991 B

(45) 授权公告日 2011. 07. 27

(21) 申请号 200910003562. 0

(22) 申请日 2009. 01. 15

(30) 优先权数据

2008-017014 2008. 01. 29 JP

(73) 专利权人 株式会社日立制作所

地址 日本东京都

(72) 发明人 大平良德 儿玉升司 志贺贤太

江口贤哲

(74) 专利代理机构 北京银龙知识产权代理有限公司

11243

代理人 许静

(51) Int. Cl.

G06F 3/06 (2006. 01)

G06F 12/08 (2006. 01)

(56) 对比文件

CN 1453715 A, 2003. 11. 05, 全文.

JP 特开 2005-135116 A, 2005. 05. 26, 全文.

CN 1881167 A, 2006. 12. 20, 全文.

审查员 吴媛媛

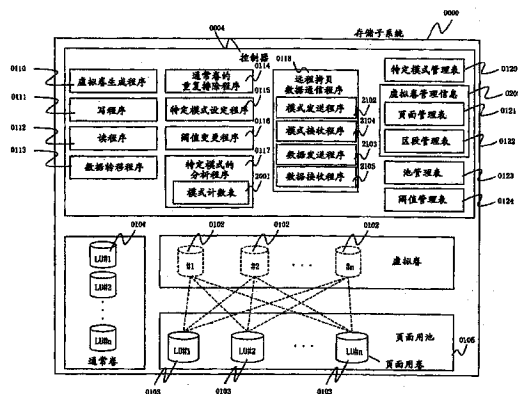
权利要求书 3 页 说明书 16 页 附图 27 页

(54) 发明名称

存储子系统

(57) 摘要

本发明提供一种可以在抑制管理信息量的同时高效率地使用存储容量的存储子系统。存储子系统具有虚拟卷、和具有分配给虚拟卷的物理区域的物理卷。将虚拟卷的地址空间分割为多个页面，将多个页面中的每个页面分类为包含至少第一状态和第二状态的多个状态中的某个状态来管理。而且，将分类为第二状态的页面进一步分割为多个区段来管理。在此，第一状态是指，在页面中接受写请求的结果为，向该页面分配物理区域，将写数据存储于物理区域中的状态。另外，第二状态是指，接受用于在页面内的区段中写入与预定的特定模式数据一致的写数据的写请求的结果为，将特定模式数据和区段对应起来管理的状态。



1. 一种存储子系统,其特征在于,
具有:
处理器;
存储器;
虚拟卷;以及
由多个存储介质构成的、具有分配给所述虚拟卷的页面用物理区域的页面用卷,
所述处理器,将所述虚拟卷的地址空间分割为多个页面,将该多个页面的各个页面分类为至少包含第一状态和第二状态的多个状态中的某一状态来进行管理,
所述处理器将被分类为所述第二状态的页面进一步分割为多个区段来进行管理,
所述第一状态是以下状态:接受用于在页面中写入写数据的写请求的结果为,从所述页面用卷向该页面分配页面用物理区域,在该页面用物理区域中存储了写数据,
所述第二状态是以下状态:接受用于在页面内的区段中写入与预定的特定模式数据一致的写数据的写请求的结果为,在所述存储器内将该特定模式数据与该区段对应起来进行管理。
2. 根据权利要求1所述的存储子系统,其特征在于,
在所述多个状态中还包含第三状态,
所述第三状态是以下状态:未从所述页面用卷向页面分配页面用物理区域,并且未将该页面分割为区段来进行管理。
3. 根据权利要求1所述的存储子系统,其特征在于,
所述处理器,当由接收到的写请求中包含的地址而指定的页面为所述第一状态时,在分配给该页面的页面用物理区域中存储写数据。
4. 根据权利要求3所述的存储子系统,其特征在于,
当处于所述第一状态的所述页面内的数据全部与特定模式数据一致时,所述处理器以所述区段为单位分割所述页面,针对各区段将该区段和所述特定模式数据对应起来在所述存储器内进行管理,
所述处理器,释放已分配给所述页面的页面用物理区域,将该页面作为所述第二状态的页面来进行管理。
5. 根据权利要求2所述的存储子系统,其特征在于,
所述处理器,当由接收到的写请求中包含的地址而指定的页面为所述第二状态或所述第三状态时,判定写数据与所述特定模式数据是否一致。
6. 根据权利要求5所述的存储子系统,其特征在于,
所述处理器,当由所述写请求中包含的所述地址而指定的页面为所述第二状态、所述写数据与所述特定模式数据一致时,将所述特定模式数据与由所述地址指定的区段对应起来在所述存储器内进行管理,并放弃所述写数据。
7. 根据权利要求5所述的存储子系统,其特征在于,
所述处理器,当由所述写请求中包含的所述地址而指定的页面为所述第二状态、所述写数据与所述特定模式数据不一致时,从所述页面用卷向由所述地址指定的页面分配页面用物理区域,根据在所述存储器内管理的所述特定模式数据与该页面内的区段之间的对应关系,在分配给所述页面的所述页面用物理区域内的与该区段对应的物理区域中存储该特

定模式数据,将该写数据存储在使用该地址确定的所述页面用物理区域内的物理区域中,所述处理器将所述页面作为所述第一状态的页面来进行管理。

8. 根据权利要求 5 所述的存储子系统,其特征在于,

所述处理器,当由所述写请求中包含的地址而指定的页面为所述第三状态、所述写数据与所述特定模式数据一致时,以所述区段为单位分割所述页面,将由所述地址指定的区段与所述特定模式数据对应起来记录在所述存储器内,并放弃所述写数据,

所述处理器将所述页面作为所述第二状态的页面来进行管理。

9. 根据权利要求 5 所述的存储子系统,其特征在于,

所述处理器,当由所述写请求中包含的地址而指定的页面为所述第三状态、所述写数据与所述特定模式数据不一致时,从所述页面用卷向所述页面分配页面用物理区域,在该页面用物理区域中存储该写数据,

所述处理器将所述页面作为所述第一状态的页面来进行管理。

10. 根据权利要求 4 所述的存储子系统,其特征在于,

所述处理器,对存储在处于所述第一状态的所述页面中的数据内的、与所述特定模式数据一致的数据的数据数量进行计数,

所述处理器,当按照所述写请求接收到的所述写数据与所述特定模式数据一致时,将该写数据存储在与所述页面分配的所述页面用物理区域中,并使所述数据数量增加 1,

所述处理器,当所述数据数量超过预定的第一阈值时,判定该页面内的全部数据是否与所述特定模式数据一致。

11. 根据权利要求 1 所述的存储子系统,其特征在于,

所述存储子系统还具有:由多个存储介质构成的、具有分配给所述虚拟卷的区段用物理区域的区段用卷,

所述处理器,当在所述第一状态的页面内存储的、与所述特定模式数据一致的数据的数量超过预定的第一阈值,并且在所述页面内存储的、与所述特定模式数据不一致的数据的数量在预定的第二阈值以下时,以所述区段为单位分割该页面,

针对所述页面内的各区段,当在该区段中存储了与所述特定模式数据一致的数据时,所述处理器将该区段与该特定模式数据对应起来在所述存储器内进行管理,当在该区段中存储了与所述特定模式数据不一致的数据时,所述处理器从所述区段用卷对该区段分配区段用物理区域,在该区段用物理区域中存储该数据,

所述处理器释放已分配给所述页面的所述页面用物理区域,将该页面作为所述第二状态的页面来进行管理。

12. 根据权利要求 11 所述的存储子系统,其特征在于,

所述处理器,在接收到针对所述第二状态的页面的写请求的情况下,当按照该写请求接收到的写数据是与所述特定模式数据不一致的数据、并且该页面中存储的与所述特定模式数据不一致的数据的数量在所述第二阈值以下时,从所述区段用卷向该页面新分配区段用物理区域,在该区段用物理区域中存储该写数据。

13. 根据权利要求 1 所述的存储子系统,其特征在于,

所述存储子系统还具有由多个存储介质构成的通常卷,

当将所述通常卷中存储的数据转移到所述虚拟卷时,

所述处理器从所述通常卷中顺次读取与所述区段相同大小的数据,将该读取的数据写入所述虚拟卷。

14. 根据权利要求 1 所述的存储子系统,其特征在于,
所述存储子系统,与具有由多个存储介质构成的物理卷的其它存储子系统相连,
所述存储子系统从所述其它存储子系统的所述物理卷的先头区段顺次读取数据,将该数据写入所述虚拟卷。

15. 根据权利要求 14 所述的存储子系统,其特征在于,
所述存储子系统,对从所述物理卷中读取的数据中有多少个相同数据进行计数,
所述存储子系统,当该计数数量的结果超过预先设定的阈值时,将该数据作为特定模式数据登录在所述存储器内。

16. 一种存储系统,具有发送源存储子系统、以及与所述发送源存储子系统可通信地连接的发送目的地存储子系统,该存储系统的特征在于,

所述发送源存储子系统具有拷贝源卷,
所述发送目的地存储子系统具有:虚拟卷;以及由多个记录介质构成的、具有分配给该虚拟卷的页面用物理区域的拷贝目的地物理卷,

所述发送目的地存储子系统,将所述虚拟卷的地址空间分割为多个页面,将该多个页面的每个页面分类为至少包含第一状态和第二状态的多个状态中的某个状态来进行管理,

所述发送目的地存储子系统,将处于所述第二状态的页面进一步分割为多个区段来进行管理,

所述第一状态是以下状态:接受用于在页面中写入数据的请求的结果为,从所述拷贝目的地物理卷向该页面分配页面用物理区域,将该数据分配给该页面用物理区域,

所述第二状态是以下状态:接受用于将与预定的特定模式数据一致的数据写入页面内的区段中的请求的结果为,将该区段与该特定模式数据对应起来进行管理的状态,

所述发送源存储子系统,从所述拷贝源卷读出作为远程拷贝对象的数据,当所述读出的数据与所述特定模式数据一致时,将识别该特定模式数据的 ID 发送给所述发送目的地存储子系统,

所述发送目的地存储子系统,当从所述发送源存储子系统接收到所述 ID 时,根据成为由该 ID 识别的所述特定模式数据的存储目的地的所述虚拟卷内的页面处于所述第一状态或所述第二状态的哪个状态,而执行不同的处理,

当成为由所述 ID 识别的所述特定模式数据的存储目的地的所述虚拟卷内的页面为所述第一状态时,所述发送目的地存储子系统在分配给该页面的页面用物理区域中存储由该 ID 识别的所述特定模式数据,

当成为由所述 ID 识别的所述特定模式数据的存储目的地的所述虚拟卷内的页面为所述第二状态时,所述发送目的地存储子系统将该 ID 与成为所述存储目的地的所述区段对应起来,在所述存储器内进行管理。

17. 根据权利要求 16 所述的存储系统,其特征在于,
所述拷贝源卷是所述发送源存储子系统具有的虚拟卷。

18. 根据权利要求 16 所述的存储系统,其特征在于,
所述拷贝源卷是由多个存储介质构成的物理卷。

存储子系统

技术领域

[0001] 本发明涉及具有多个磁盘装置的存储子系统 (Storage Subsystem), 尤其涉及用于有效地使用存储子系统具有的存储区域的技术。

背景技术

[0002] 作为在计算机系统中使用的存储装置系统, 存在具有多个磁盘装置的磁盘阵列系统 (以下也称为存储子系统)。在磁盘阵列系统中, 通过使多个磁盘装置并行地动作, 与单独地使用磁盘装置的存储装置系统相比, 实现了高速化。

[0003] 作为用于高效地利用磁盘阵列系统内的磁盘容量的技术, 存在虚拟卷技术。本技术是向计算机提供了虚拟存储区域, 仅对虚拟存储区域内存在写请求的区域分配物理存储区域的技术。

[0004] 另外, 作为用于高效地利用磁盘容量的其它技术, 还有重复排除技术。本技术是: 与接受写请求的数据相同的数据在接收该写请求之前已被存储在存储子系统中时, 仅存储表示与该已存储的数据是相同数据的意思, 接受写请求的数据本身不存储在存储子系统中。

[0005] 在重复排除技术中存在: 对于存储子系统中存储的全部数据进行重复排除控制的技术; 以及仅对预先登录的特定模式数据进行重复排除控制的技术。前者可以提高为了存储数据而需要的存储容量的削减效果 (以下称为容量削减效果), 但处理时间增长, 而且管理信息增多。相反, 后者可以减少管理信息, 但限制了容量削减效果。

[0006] 在专利文献 1 中记载了以下技术: 以某单位长的物理块分割由多个物理存储装置的集合构成的物理存储空间上的区域, 通过物理块来分配物理存储空间, 在由此构成逻辑存储装置的存储装置中具有逻辑存储管理表, 该逻辑存储管理表针对所述逻辑存储装置上的每个区域记录识别信息, 该识别信息表示是否分配了所述物理块、或是否符合根据规定的配置模式而生成数据, 当存在针对所述逻辑存储装置的区域的数据写入请求时, 当写入请求数据的配置模式与预先登录的配置模式一致时, 与所述逻辑存储管理表的存在写入请求的区域相对应地记录一致的配置模式的识别信息, 放弃所述写入请求数据。

[0007] 【专利文献 1】特开 2005-135116 号公报

发明内容

[0008] 在专利文献 1 中, 使用了虚拟卷技术和重复排除技术二者, 但虚拟卷的管理规模 (size) 与重复排除的管理规模相同。为了提高虚拟卷技术以及重复排除技术的容量削减效果而将两者的管理规模缩小时, 管理信息增多, 当想要将全部管理信息存储在高速缓冲存储器等可以高速访问的存储介质中时, 系统变得昂贵。相反, 当增大管理规模时, 管理信息减少, 但容量削减效果降低。即, 在管理信息量和容量削减效果之间存在折衷 (trade-off) 的关系。因此, 寻求用于抑制用于虚拟卷技术以及重复排除技术的管理信息量, 同时高效地

使用存储子系统的存储容量的技术。

[0009] 存储子系统具有：处理器、存储器、虚拟卷、以及由多个存储介质构成的具有分配给所述虚拟卷的页面用物理区域的页面用卷。并且，处理器将虚拟卷的地址空间分割为多个页面，将多个页面的各个页面分类为至少包含第一状态和第二状态的多个状态中的某一状态来进行管理。处理器将被分类为第二状态的页面进一步分割为多个区段来进行管理。在此，所谓第一状态是指，接受用于在页面中写入写数据的写请求的结果为，从页面用卷对该页面分配页面用物理区域，在页面用物理区域中存储了写数据的状态。所谓第二状态是指，接受用于在页面内的区段中写入与预定的特定模式数据一致的写数据的写请求的结果为，在存储器内将特定模式数据与区段相对应地进行管理的状态。

[0010] 可以在抑制用于卷虚拟化技术以及重复排除技术的管理信息量的同时，高效率地使用存储子系统的存储容量。

附图说明

[0011] 图 1 是表示实施例 1、实施例 2 以及实施例 5 中的存储系统的硬件结构的一例的说明图。

[0012] 图 2 是表示存储子系统的逻辑系统结构的一例的说明图。

[0013] 图 3 是表示虚拟卷的地址空间的一例的说明图。

[0014] 图 4 是表示虚拟卷管理信息的一例的说明图。

[0015] 图 5 是表示池管理表的一例的说明图。

[0016] 图 6 是表示特定模式管理表的一例的说明图。

[0017] 图 7 是表示阈值管理表的一例的说明图。

[0018] 图 8 是表示虚拟卷生成程序的处理流程的一例的说明图。

[0019] 图 9 是表示特定模式设定程序的处理流程的一例的说明图。

[0020] 图 10 是表示阈值变更程序的处理流程的一例的说明图。

[0021] 图 11 是表示读程序 (read programme) 的处理流程的一例的说明图。

[0022] 图 12 是表示写程序 (write programme) 的处理流程的一例的说明图。

[0023] 图 13 是表示 Write 处理 (1) 的处理流程的一例的说明图。

[0024] 图 14 是表示 Write 处理 (2) 的处理流程的一例的说明图。

[0025] 图 15 是表示 Write 处理 (3) 的处理流程的一例的说明图。

[0026] 图 16 是表示重复排除程序的处理流程的一例的说明图。

[0027] 图 17 是表示实施例 3 以及实施例 4 中的存储系统的硬件结构的一例的说明图。

[0028] 图 18(a) 是表示模式发送程序的处理流程的一例的说明图。

[0029] 图 18(b) 是表示模式接收程序的处理流程的一例的说明图。

[0030] 图 19 是表示向存储子系统发送的数据包的实例的说明图。

[0031] 图 20(a) 是表示数据发送程序的处理流程的一例的说明图。

[0032] 图 20(b) 是表示数据接收程序的处理流程的一例的说明图。

[0033] 图 21 是表示模式计数表的一例的说明图。

[0034] 图 22 是表示特定模式的分析程序的处理流程的一例的说明图。

[0035] 图 23 是表示数据转移程序的处理流程的一例的说明图。

- [0036] 图 24 是表示实施例 5 中的存储子系统的逻辑系统结构的一例的说明图。
- [0037] 图 25 是表示实施例 5 中的区段管理表的一例的说明图。
- [0038] 图 26 是表示实施例 5 中的阈值管理表的一例的说明图。
- [0039] 图 27 是表示实施例 5 中的 Write 处理 (1) 的处理流程的一例的说明图。
- [0040] 图 28 是表示实施例 5 中的 Write 处理 (2) 的处理流程的一例的说明图。
- [0041] 符号说明

[0042] 0000 : 存储子系统 ; 0110 : 虚拟卷生成程序 ; 0111 : 写程序 ; 0112 : 读程序 ; 0113 : 数据转移程序 ; 0114 : 通常卷的重复排除程序 ; 0115 : 特定模式设定程序 ; 0116 : 阈值表更新程序 ; 2001 : 特定模式的分析程序 ; 0118 : 远程拷贝数据通信程序 ; 0120 : 特定模式管理表 ; 0202 : 虚拟卷管理信息系统本部 ; 0123 : 池管理表 ; 0124 : 阈值管理表 ; 0121 : 页面管理表 ; 0122 : 区段管理表 ; 0102 : 虚拟卷 ; 0103 : 页面用卷 ; 0104 : 通常卷

具体实施方式

[0043] 下面说明本发明的实施方式。

[0044] (实施例 1)

[0045] 以下, 说明本发明的实施方式的一例。此外, 本发明不仅限于以下记载的实施方式。

[0046] 图 1 是表示应用本发明的存储系统的硬件结构的一例。存储系统具有存储子系统 0000 和主计算机 (主机) 0001, 存储子系统 0000 和主机 0001 通过主机接口 (主机 I/F) 0002 相连。存储子系统 0000 由主机接口 0002、管理接口 (管理 I/F) 0003、由存储器和 CPU 构成的控制器 0004、高速缓冲存储器 0005、磁盘接口 (磁盘 I/F) 0007 以及多个硬盘 (HDD) 0006 构成, 除 HDD 以外通过内部总线 0008 相连。通过在管理终端 0009 上运行的管理软件 0010 来进行存储子系统 0000 的构成管理。管理终端 0009 和存储子系统 0000 通过管理接口 0003 相连。

[0047] 此外, 虽未图示, 但管理终端 0009 以及主机 0001 是分别具有 CPU、存储器、用于与存储子系统 0000 进行通信的通信 I/F 等的一般计算机, 上述管理软件 0010 被存储在管理终端 0009 内的存储器内中, 通过管理终端 0009 的 CPU 而被执行。

[0048] 图 2 表示了存储子系统 0000 的逻辑系统结构的一例。在后面描述各种程序及各种表的细节。此外, 这些程序及表被存储在图 1 所示的控制器 0004 内的存储器中, 通过控制器 0004 的 CPU 被执行。其中, 0113、0114、0117、0118 的程序在实施例 1 中未使用, 因此在实施例 2 以后进行详细说明。

[0049] 0104 是通常卷, 存储子系统 0000 具有多个通常卷 0104。各通常卷 0104 是在由多个 HDD 0006 构成的 RAID 组等 HDD 组上构成的存储区域, 从主机 0001 以及管理终端 0009 识别为一个逻辑存储区域。

[0050] 0102 是虚拟卷, 存储子系统 0000 具有多个虚拟卷 0102。虚拟卷 0102 具有与通常卷 0104 同样的地址空间, 从主机 0001 与通常卷 0104 同样地识别为一个逻辑存储区域, 与通常卷 0104 同样地进行数据的读取或写入。虚拟卷 0102 与通常卷 0104 的不同点是 : 在数据写入前对通常卷 0102 分配了物理区域, 与之相对, 以写入了数据作为契机对虚拟卷 0104 分配物理区域。若使用虚拟卷 0104, 则当没有写入时不分配物理区域即可, 因此可以减少

安装在存储系统中的物理区域,可有效利用存储容量。此外,从页面用池 0105 确保分配给虚拟卷的物理区域。页面用池 0105 中包含多个页面用卷 0103。页面用卷 0103 与通常卷 0104 同样,是在由多个 HDD0006 构成的 RAID 组等 HDD 组上构成的存储区域。因此,对虚拟卷分配属于页面用池 0105 的页面用卷 0103 上的物理区域。

[0051] 图 3 表示了虚拟卷 0102 的地址空间的一例。虚拟卷 0102 的地址空间被分割成多个页面 (page) 来进行管理。为了容易进行页面的管理,通常使页面的大小为固定大小。在虚拟卷的地址空间 0201 上的页面中有 (a1) (a2) (a3) 这 3 种状态,根据各个状态进行不同的管理。而且,(a2) 的状态的页面,以区段单位被分割为多个来进行管理。这里,在图 3 中将页面内分割成 8 个区段,但只要页面内的区段在两个以上即可不限定个数。另外,为使区段的管理变得容易,所以通常使区段的大小为固定大小。

[0052] 详细说明 (a1) (a2) (a3) 的状态。(a1) 的状态是在页面内存在来自主机 0001 的写请求、结果向页面分配了物理区域的状态。

[0053] (a2) 的状态是,在页面内存在来自主机 0001 的写请求,但写数据与存储子系统 0000 管理的特定模式的数据一致,因此仅将表示是特定模式数据的意思作为管理信息,记录在后述的区段管理表 0122 中,未向页面分配物理区域的状态。即,该状态的页面通过前面所述的重复排除技术进行了存储容量的削减。另外,以区段单位管理处于 (a2) 状态的页面,是由于当如上所述进行重复排除控制时,在使管理单位较小的情况下,重复排除的存储容量的削减效果较高。即,关于实施重复排除控制的页面 ((a2) 状态的页面),通过以大小较小的区段作为管理单位,提高了存储系统中的存储容量的削减效果,关于未实施重复排除控制的页面 ((a1) 以及后述的 (a3) 状态的页面),通过不以区段单位、而以页面单位进行管理,防止了管理信息以及处理负荷的增大。

[0054] (a3) 的状态是,在页面内还没有来自主机 0001 的写请求,因此未分配物理区域的状态。该状态的页面通过前面所述的虚拟卷技术削减了存储容量。

[0055] 如上所述,根据是否从主机有写请求、以及当有写请求时写请求中包含的写数据是否是特定模式的数据,将虚拟卷 0102 的地址空间上的各页面分类为 (a1) (a2) 以及 (a3) 的状态来进行管理,由此可以在抑制管理信息量的同时提高容量削减效果。

[0056] 图 4 表示为了管理虚拟卷 0102 的地址空间而使用的、虚拟卷管理信息 0202 的一例。在生成虚拟卷 0102 时,针对每个虚拟卷 0102 生成虚拟卷管理信息 0202。各虚拟卷管理信息 0202 具有识别虚拟卷的虚拟卷号码 (虚拟卷 #) 0211 栏和一个页面管理表 0121。而且,当通过虚拟卷 #0211 识别的虚拟卷中存在状态为 (a2) 的页面时,具有状态为 (a2) 的页面的个数的区段管理表 0122。

[0057] 页面管理表 0121 是管理虚拟卷 0102 的地址空间上的各页面的状态的表。区段管理表 0122 是管理 (a2) 的状态的页面内的区段的表,并按照 (a2) 状态的页面的个数而被生成。

[0058] 首先,说明页面管理表 0121。在页面管理表 0121 的页面号码 (页面 #) 栏 0203 中记录了虚拟卷 0102 具有的页面的号码。在状态栏 0204 中记录了该页面的状态。在本实施例中,使对应于 (a1) 状态的值为 0、对应于 (a2) 状态的值为 1、对应于 (a3) 状态的值为 2。在区域号码 (区域 #) 栏 0205 中记录了表示向该页面分配了页面用池 0105 内的哪个物理区域的信息。因此,当状态 0204 为 0 时、即当页面的状态为 (a1) 时,将表示向该页面分配

的物理区域的识别号码记录在区域 #0205 中。而当状态 0204 为 1 或 2 时、即当页面状态为 (a2) 或 (a3) 时,由于未向该页面分配物理区域,所以在区域 #0205 中记录 null。此外,通过后述的池管理表 0123 来管理页面用池 0105 内的物理区域。在表号码 (表 #) 栏 0206 中,记录表示管理该页面的区段的区段管理表 0122 的识别号码。因此,当状态 0204 为 1 时、即页面的状态为 (a2) 时,与该页面对应的区段管理表 0122 的识别号码被记录在表 #0206 中,但在此外的情况下 (页面状态为 (a1) 或 (a3)),由于未以区段单位管理页面,因此区段管理表 0122 也不存在,在表 #0206 中记录 null。在特定模式计数数量栏 0207 中,记录表示针对该页面的写请求内的写数据与存储子系统管理的特定模式数据一致的次数的数。此外,关于状态为 (a1) 的页面,记录特定模式计数数量 0207,关于状态为 (a2) 或 (a3) 的页面,记录 null 值。当页面的状态为 (a2) 时,写入该页面的写数据应该全部与特定模式数据一致,当页面的状态为 (a3) 时,由于尚未从主机接收写请求,因此不需要对与特定模式一致的数据的数量进行计数。

[0059] 下面,说明区段管理表 0122。在区段管理表 0122 的区段号码栏 (区段 #) 0208 中,记录赋予一个页面内存在的各区段的识别号码。各区段可以取以下两种状态:从要写入特定模式数据的主机接受写请求,并进行了表示已写入特定模式数据的意思的记录的状态;以及未从主机接受写请求的状态。在标志栏 0209 中记录识别上述区段的状态的值。在本实施例中,使表示进行了已写入特定模式数据的意思的记录的状态为 0;使对于区段尚未取得写请求的状态为 1。在模式 ID0210 中记录表示从主机对于区段有写请求的数据的模式的识别号码。因此,仅在标志 0209 为 0 时,在模式 ID0210 栏中记录识别特定数据模式的 ID,当标志 0209 为 1 时,在模式 ID0210 中记录 null。此外,在后面描述为了识别从主机请求写入的写数据是否是特定模式数据而使用的、管理特定模式数据的表 0120。

[0060] 图 5 表示对构成页面用池 0105 的存储区域 (物理区域) 进行管理的池管理表 0123 的一例。在区域号码 (区域 #) 1231 栏中,记录构成页面用池 0105 的物理区域的识别号码。该区域 #,从前面所述的页面管理表 0121 的区域 #0205 被参照。在分配标志 1232 栏中,记录表示是否将该物理区域分配给了虚拟卷 0102 的标志。在本实施例中,当未分配区域时记录 0,当已分配时记录 1。在池 LU 号码 (池 LU#) 栏 1234 中,记录区域 #1231 表示的物理区域属于的页面用卷 0103 的识别号码,在开始地址栏 1235 中,记录用于确定该物理区域在页面用卷 0103 内的位置的开始地址。即,构成页面用池 0105 的各物理区域相当于,通过池 LU#1234 识别的页面用卷 0103 内的从开始地址 1235 起与页面大小相等大小的存储区域。

[0061] 图 6 是用于管理在判断向虚拟卷 0102 内存在写请求的写数据是否是特定模式数据时使用的特定模式数据的特定模式管理表 0120 的一例。在特定模式管理表 0120 中,将识别特定模式数据的 ID 记录在模式 ID1201 中,在模式数据 1202 的列中记录该特定模式的数据。所记录的特定模式的数据大小与区段大小相同。

[0062] 图 7 是管理在判断是否使页面的状态从 (a1) 向 (a2) 迁移时使用的阈值的阈值管理表 0124 的一例。此外,使用图 13 在后面描述使页面的状态从 (a1) 向 (a2) 迁移的处理。在阈值管理表 0124 中,识别虚拟卷 0102 的虚拟卷号码被记录在虚拟卷 #1241 栏中,针对每个虚拟卷记录了特定模式计数数量的阈值 γ 1242。用户可以在管理软件 0010 上对于各个虚拟卷 0102 指定阈值 γ 。并且,在前面所述的页面管理表 0121 中管理的特定模式计数数量 0207 达到阈值 r 时,调查在处于页面状态 (a1) 的页面中实际写入的数据,当断定该数据

全部是特定模式数据时,该数据的状态从(a1)向(a2)迁移。

[0063] 图8表示虚拟卷生成程序0110的处理流程0301的一例。首先,用户在管理软件0010上向存储子系统0000的虚拟卷生成程序0110指示生成虚拟卷0102。虚拟卷生成程序0110从管理软件0010接受虚拟卷生成命令(步骤0302),生成虚拟卷0102,而且生成虚拟卷管理信息0202(步骤0303)。即,虚拟卷生成程序0110向所生成的虚拟卷分配识别号码,记录在虚拟卷管理信息0202的虚拟卷#0211中。另外,对于此时生成的虚拟卷0102内的页面,尚未接收到来自主机的写请求,因此,在对于该虚拟卷生成的页面管理表0121的状态0204中,针对全部页面记录表示(a3)的2,在区域#0205、表#0206、特定模式计数数量0207中记录null值。而且不生成区段管理表0122。然后,虚拟卷生成程序0110,在阈值管理表0124的虚拟卷#1241中记录识别所生成的虚拟卷0102的号码,将用户在管理软件0010上任意指定的阈值记录在阈值 γ 1242中(步骤0304)。然后,虚拟卷生成程序0110对管理软件0010应答虚拟卷生成结束(步骤0305),结束处理。

[0064] 图9表示特定模式设定程序0115的处理流程0401的一例。用户在管理软件0010上指定作为特定模式而追加(或删除)的模式,对特定模式设定程序0115指示向特定模式管理表0120的特定模式数据的追加(或删除)。特定模式设定程序0115从管理软件0010接受特定模式数据的设定命令(步骤0402)。当设定命令指示追加新的数据作为特定模式数据时,特定模式设定程序0115向命令中所指定的数据分配ID,将该ID与所指定的数据一起记录在特定模式管理表0120中(步骤0403)。

[0065] 当设定命令指示从特定模式管理表0120中删除特定模式数据时,特定模式设定程序0115根据在设定命令中指定的ID来识别作为删除对象的特定模式数据,从特定模式管理表0120中删除该ID和所识别出的特定模式数据(步骤0403)。

[0066] 当特定模式管理表的更新结束时,特定模式设定程序0115向管理软件0010应答设定结束(步骤0404),结束处理。

[0067] 图10表示变更在判断是否使页面的状态从(a1)变迁为(a2)时所使用的阈值 γ 的阈值变更程序0116的处理流程0116的一例。用户在管理软件0010上指定识别想要变更阈值的虚拟卷的号码和变更后的阈值,管理软件0010向阈值变更程序0116指示变更阈值管理表0124上的阈值 γ 1242。阈值变更程序0116从管理软件0010接受阈值变更命令(步骤0502),将阈值管理表0124中指定的虚拟卷的阈值变更为指定的值(步骤0503)。当阈值的变更结束时,阈值变更程序0124向管理软件0010应答阈值变更结束(步骤0504)。

[0068] 图11是读程序0112的处理流程1001的一例。读程序0112当从主机0001接受针对虚拟卷0102的读请求时(步骤1001),参照所指定的虚拟卷的页面管理表0121,检查与存在读请求的地址对应的页面的状态0204(步骤1002)。此外,在读请求中包含:用于指定读对象虚拟卷的虚拟卷号码、该虚拟卷中的读对象区域的地址信息。另一方面,由于构成虚拟卷的各页面具有规定大小,因此可以根据读请求中包含的地址信息和各页面的大小确定读对象页面。

[0069] 当页面的状态为(a1)时,向该页面分配了物理区域,因此,读程序0112参照页面管理表0121的区域#0205,确定与该页面对应的物理区域的识别号码。而且,根据所确定的物理区域的识别号码,参照池管理表0123,读程序0012取得读对象页面用卷0103的号码1234、和该页面用卷0103内的读对象区域的开始地址1235(步骤1004)。然后,读程序

0112 从通过取得的页面用卷的号码和地址所指定的物理区域读取读对象数据,将读取的数据向主机 0001 发送(步骤 1005)。

[0070] 当页面的状态为 (a2) 时,该页面被进一步分割为区段来管理,因此,读程序 0112 参照关于所指定的虚拟卷 0102 的页面管理表 0121 的表 #0206,取得与该页面对应的区段管理表 0122(步骤 1006)。然后,读程序 0112 参照所取得的区段管理表 0122,确认读对象区段的标志 0209(步骤 1007)。此外,构成页面的区段也具有规定的大小,因此可以根据读请求中包含的地址信息和各区段的大小来确定读对象页面。

[0071] 当该标志为 0 时,向该区段分配了特定模式的数据,因此,读程序 0112 参照模式 ID0210,从特定模式管理表 0120 取得通过该 ID 识别的数据,将该数据向主机 0001 发送(步骤 1009)。另一方面,当标志 0209 为 1 时,未在该区段中写入数据,因此,向主机 0001 发送规定的格式数据(步骤 1010)。此外,在一般的存储子系统中,多数情况下使格式数据(format data)为 0,但不限于此。另外,格式数据也可以针对每个虚拟卷由用户指定。

[0072] 当页面的状态为 (a3) 时,在该页面中尚未写入数据。因此,读程序 0112 和步骤 1010 同样地向主机 0001 发送格式数据(步骤 1003)。

[0073] 图 12 是写程序 0111 的处理流程 0601 的一例。写程序 0111,当从主机 0001 接受针对虚拟卷 0102 的写请求时,参照关于所指定的虚拟卷的页面管理表 0121,检查有写请求的地址所对应的页面的状态 0204(步骤 1002)。此外,写对象页面的指定方法与前面所述的读处理时相同。

[0074] 写程序 0111,当页面的状态为 (a1) 时执行 Write 处理 (1)(步骤 0604),当页面的状态为 (a2) 时执行 Write 处理 (2)(步骤 0605),当页面的状态为 (a3) 时执行 Write 处理 (3)(步骤 0606)。使用图 13、图 14、图 15 说明 Write 处理 (1)(2)(3) 的细节。

[0075] 图 13 表示在写程序 0111 接受写请求时写对象页面的状态为 (a1) 时所执行的、前面所述的 Write 处理 (1)(步骤 0604) 的一例。写程序 0111 参照关于写对象虚拟卷的页面管理表 0121,确定分配给写对象页面的物理区域的区域 #0205。写程序 0111 从池管理表 0123 取得所确定的物理区域的池 LU#1234 和开始地址 1235(步骤 0701)。然后,在通过所取得的池 LU#1234 以及开始地址 1235 指定的物理区域中写入写数据(步骤 0702),将写完成报告发送至主机 0001(步骤 0703)。

[0076] 此后,写程序 0111 参照特定模式管理表 0120,判定有写请求的写数据是否与特定模式数据的某一个一致(步骤 0704)。当写数据与特定模式数据 1202 不一致时结束写处理。当写数据与特定模式数据 1202 一致时进入步骤 0705,参照页面管理表 0121,在接受了写请求的页面的特定模式计数数量 0207 上加 1。

[0077] 在步骤 0705 后进入步骤 0706,写程序 0111 参照阈值管理表 0124,确定与写对象虚拟卷对应的阈值 γ 1242,判定在步骤 0705 中加 1 后的特定模式计数数量 0207 是否超过了该阈值。当未超过该阈值时结束写处理。当特定模式计数数量超过了该阈值时,进入步骤 0707,写程序 0111 判定在该页面内写入的数据是否全部是与特定模式数据一致的数据。当在该页面中写入的数据即使有 1 个与特定模式数据不一致时,结束写处理。当在该页面中写入的写数据全部与特定模式数据一致时,进入步骤 0708,写程序 0111 针对写对象页面新生成区段表 0122。此时,关于该页面,由于全部数据是特定模式数据,因此在该区段表 0122 的标志 0209 中全部设定 0。然后,写程序 0111 参照特定模式管理表 0120,确定识别在该页

面的各区段中存储的特定模式数据的 ID,记录在模式 ID0210 中(步骤 0708)。此后,释放已分配给写对象页面的页面用池 0105 内的物理区域(步骤 0709)。通过步骤 0708、0709 使写对象页面的状态从 (a1) 迁移到 (a2),因此,写程序 0111 将页面管理表 0121 的与该页面对应的状态 0204 从 0 变更为 1(步骤 0710)。另外,在页面管理表 0121 的表 #0206 中,记录识别步骤 0708 中新生成的区段管理表 0122 的号码(步骤 0711)。

[0078] 通过以上 Write 处理 (1),即使是暂时分配了物理区域的页面,关于成为仅由特定模式数据构成的页面,释放已分配的物理区域,仅记录表示从主机写入了哪个特定模式数据的管理信息即可,可以提高容量削减效果。

[0079] 图 14 表示在写程序 0111 在写对象页面的状态为 (a2) 时执行的、前面所述的 Write 处理 (2)(步骤 0605) 的一例。写程序 0111 参照特定模式管理表 0120,判定模式数据 1202 中是否存在与接受了写请求的写数据一致的数据(步骤 0801)。

[0080] 在步骤 0801 中,当存在一致的特定模式数据时,进入步骤 0802,写程序 0111 根据特定模式管理表 0120 的模式 ID1201,确定由于识别与写数据一致的特定模式数据的 ID。然后,参照写对象页面的区段管理表 0122,在写对象区段的模式 ID0210 中记录所确定的模式数据的 ID(步骤 0802)。此外,可以与读对象区段通过同样的方法确定写对象区段。此后,写程序 0111 将写完成报告发送至主机 0001(步骤 0803),释放写数据(步骤 0804),结束写处理。

[0081] 在步骤 0801 中,当写数据与特定模式数据不一致时,进入步骤 0805,写程序 0111 从页面用池 0105 向写对象页面分配物理区域,因此参照池管理表 0123。然后,写程序 0111 确保表示未分配的物理区域的分配标志 1232 为 0 的区域(步骤 0805)。即,写程序 0111 将该未分配物理区域的标志 1232 变更为 1,同时在写对象虚拟卷的写对象页面的页面管理表 0121 中,记录所确保的物理区域的识别号码作为区域 #0205。

[0082] 写程序 0111 在步骤 0805 中确保的物理区域中存储原来分配给写对象页面的(即,从主机接受写请求,不进行向实际物理区域的数据的存储,而只进行写入了特定模式数据的意思的管理)特定模式数据(步骤 0806)。即,写程序 0111 参照关于写对象虚拟卷的页面管理表 0121,从接收写请求的页面所对应的表 #0206 中取得该页面的区段管理表 0122。然后,在该区段管理表 0122 中,关于记录了模式 ID0210 的各区段,写程序 0111 参照特定模式管理表 0120,取得通过该模式 ID 识别的特定模式数据,在与该区段对应的确保的物理区域上的存储区域中存储该特定模式数据。此外,写程序 0111 可以将此时在物理区域上写入的特定模式数据的数量作为该页面的特定模式计数数量 0207,记录在页面管理表 0121 中。

[0083] 而且,写程序 0111 在与写对象区段对应的、所确保的物理区域上的存储区域中写入写数据(步骤 0807)。

[0084] 通过步骤 0805 ~ 0807 使接受了写请求的页面的状态从 (a2) 迁移到 (a1),因此,写程序 0111 将与页面管理表 01021 的该页面对应的状态 0204 从 1 变更为 0,在表 #0206 中记录 null,同时放弃与该页面相关的区段管理表 0122(步骤 0808)。此后,写程序 0111 向主机 0001 发送写完成报告(步骤 0809),结束写处理。

[0085] 图 15 表示当写请求对象页面的状态为 (a3) 时,写程序 0111 执行的 Write 处理 (3) 的一例。写程序 0111 参照特定模式管理表 0120,判定在特定模式数据 1202 中是否存

在与从主机接收的写数据一致的数据（步骤 0901）。

[0086] 当步骤 0901 中存在一致的数据时进入步骤 0902, 写程序 0111 新生成写请求对象页面的区段表 0122 (步骤 0902)。然后, 关于写对象区段, 在标志 0209 中设定 0, 而且, 参照特定模式管理表 0120, 确定识别与写数据相同的特定模式数据的 ID, 将所确定的 ID 记录在模式 ID0210 中 (步骤 0903)。通过步骤 0902、0903, 写对象页面的状态从 (a3) 迁移到了 (a2), 因此, 写程序 0111 参照该页面的页面管理表 0121, 将该页面的状态 0204 从 2 变更为 1 (步骤 0904)。另外, 在页面管理表 0121 的表 #0206 中, 记录识别通过步骤 0902 新生成的区段管理表 0122 的号码 (步骤 0905)。最后, 写程序 0111 将写完成应答发送给主机 0001 (步骤 0906), 释放写数据 (步骤 0907), 结束写处理。

[0087] 当步骤 0901 中不存在一致的特定模式数据时进入步骤 0906, 写程序 0111 从页面用池 0103 向写对象页面分配物理区域, 参照池管理表 0123, 确保表示未分配的物理区域的分配标志 1232 为 0 的物理区域 (步骤 0906)。物理区域的确保方法, 与前面所述的 Write 处理 (2) 的步骤 0805 相同。然后, 根据写请求中包含的虚拟卷的识别号码和写对象区域的地址, 推断出对应的物理区域上的存储区域, 在推断出的存储区域中写入写数据 (步骤 0907)。通过步骤 0906、0907, 写对象页面的状态从 (a3) 迁移到了 (a1), 因此, 写程序 0111 参照该页面的页面管理表 0121, 将该页面的状态 0204 从 2 变更为 0 (步骤 0908)。最后, 写程序 0111 向主机 0001 发送写完成报告 (步骤 0909), 结束写处理。

[0088] (实施例 2)

[0089] 在本实施例中说明, 通过将通常卷 0104 中存储的数据转移到虚拟卷 0102 中, 由此排除通常卷 0104 中存储的特定模式数据的重复, 削减存储子系统的数据量的方法。

[0090] 图 16 表示图 2 所示的通常卷的重复排除程序 0114 的流程的一例。首先, 通常卷的重复排除程序 0114 选择具有与通常卷 0104 相同容量的虚拟卷 0102, 作为数据转移目的地的虚拟卷 (步骤 2401)。此外, 重复排除程序 0114 也可以指示虚拟卷生成程序 0110 生成数据转移目的地的虚拟卷 0102。然后, 通常卷的重复排除程序 0114 从通常卷中顺次读取与区段大小相同大小的数据 (步骤 2402), 使用实施例 1 中说明的写程序 0111 对虚拟卷 0102 写入所读取的数据 (步骤 2403)。

[0091] 写处理的方法与实施例 1 中说明的图 12 ~ 图 15 相同, 因此省略说明, 但数据的转移目的地虚拟卷 0102, 通过写程序 0111 的处理被分类为具有前面所述的状态 (a1)、(a2) 的某个状态的页面。关于状态 (a1) 的页面, 分配物理区域, 在该物理区域中写入数据, 但在状态 (a2) 的页面中, 使用重复排除技术, 不分配物理区域。结果, 可以削减存储系统的数据容量。

[0092] 如上所述, 通过本发明, 在将通常卷 0102 内的数据向虚拟卷 0102 转移时, 也可以进行容量削减效果高的重复排除。

[0093] (实施例 3)

[0094] 在本实施例中说明, 如图 17 所示经由网络 2101 ~ 2103 连接多台实施例 1 中说明的存储子系统 0000 时, 在该多个存储子系统 0000 间进行容量削减效果高的远程拷贝的方法。

[0095] 在远程拷贝中使用图 2 所示的远程拷贝通信程序 0118, 在成为发送源的存储子系统中使用模式发送程序 2102 和数据发送程序 2104, 在成为发送目的地的存储子系统中使

用模式接收程序 2103 和数据接收程序 2105, 在存储子系统 0000 间进行远程拷贝。在后面描述各程序 (2102 ~ 2105) 的细节。在本实施例中考虑将一个存储子系统的通常卷 0104 内的数据向其它存储子系统的虚拟卷 0102 拷贝的情况。

[0096] 作为远程拷贝的事先准备, 首先, 将发送源存储系统和发送目的地存储系统分别具有的特定模式管理表 0120 的数据同步。为此, 使用模式发送程序 2102 和模式接收程序 2104。

[0097] 图 18(a) 表示在发送源存储子系统中使用的模式发送程序的处理流程的一例。模式发送程序 2102, 将发送源存储子系统具有的特定模式表 0120 序列化 (步骤 2301)。关于序列化的方法, 考虑多个通用方式, 但在本说明书中不进行详细记载。此外, 所谓序列化, 是为了将软件内部处理的数据全部保存在文件中或者经由网络收发而进行变换。然后, 模式方法程序 2012 将序列化后的特定模式表 0120 的序列数据发送给发送目的地存储子系统的模式接收程序 2014 (步骤 2302)。然后, 接收来自发送目的地存储子系统的应答后结束 (步骤 2303)。

[0098] 图 18(b) 表示在发送目的地存储子系统中使用的模式接收程序的处理流程的一例。模式接收程序 2104 接收从模式发送程序 2102 发送的序列数据 (步骤 2304), 将该序列数据变换为特定模式表 0120 的数据 (步骤 2305)。然后, 用变换后的数据替换发送目的地存储子系统的特定模式管理表 0120 (步骤 2306)。完成替换后, 对于发送源存储子系统发送完成应答 (步骤 2307)。

[0099] 接着, 说明在多个存储子系统间实施远程拷贝时的数据的收发方法。图 19 表示从发送源存储子系统向发送目的地存储子系统发送的数据包的实例。数据包由存储确定发送目的地存储子系统的地址信息的头部 1801、和存储发送数据的数据部构成。数据部由数据识别标志 1802、虚拟卷信息 1803 和数据存储部 1804 构成。所谓数据识别标志 1802, 是表示使用数据包发送的数据是实际数据还是模式 ID 的标志。在本实施例中, 假定当该标志为 0 时, 在数据存储部 1804 中存储了实际数据。另一方面, 当该标志为 1 时, 在数据存储部 1804 中存储了模式 ID。虚拟卷信息 1803 是写入发送数据的、发送目的地存储子系统内的虚拟卷的号码和地址。

[0100] 图 20(a) 表示在发送源存储子系统中使用的数据发送程序 2103 的处理流程的一例。数据发送程序 2103 从作为远程拷贝的拷贝源的发送源存储子系统的通常卷 (以下称为拷贝源通常卷) 读出数据 (步骤 1701)。然后, 在步骤 1702 中, 数据发送程序 2103 参照发送源存储子系统具有的特定模式管理表 0120, 判定读出的数据是否与特定模式数据一致。与特定模式数据都不一致时, 进入步骤 1704, 数据发送程序 2103 将读出的该数据以原有形式作为数据包发送给发送目的地存储子系统。另一方面, 在步骤 1702 中, 当特定模式数据的某一个与步骤 1701 中读出的数据一致时, 数据发送程序 2103 根据特定模式管理表 0120 的模式 ID1201, 将该数据变换为模式 ID (步骤 1703), 使用前面所述的数据包, 将该模式 ID 发送给发送目的地存储子系统 (步骤 1710)。

[0101] 图 20(b) 表示在发送目的地存储子系统中使用的数据接收程序 2105 的处理流程的一例。数据接收程序 2105, 当从发送源存储子系统接收数据包时 (步骤 1705), 确认该数据包内的数据识别标志 1802 (步骤 1706)。当该数据识别标志 1802 为 0 时进入步骤 1708, 指示写程序 0111 在接收数据包的数据部信息 1803 表示的虚拟卷的存储区域中写入实际数

据。结果,通过与实施例 1 中使用图 12 ~ 图 15 说明的处理相同的处理,通过写程序 0111 将数据存储作为远程拷贝的发送目的地的拷贝目的地虚拟卷中。数据接收程序 2105,当写处理结束时,向发送源存储子系统发送完成应答,然后结束(步骤 1709)。在步骤 1706 中,当数据识别标志 1802 为 1 时,进入步骤 1708,数据接收程序 2105 参照发送目的地存储子系统具有的特定模式管理表 0120,将存储在数据存储部 1804 中的模式 ID 变换为实际数据(步骤 1707)。此后,数据接收程序 2105 执行与前面所述的相同的步骤 1708、步骤 1709 的处理,然后结束。

[0102] 通过以上方法,当从拷贝源通常卷读出的数据与特定模式数据一致时,从发送源存储子系统向发送目的地存储子系统不发送实际数据而发送模式 ID,因此可以降低远程拷贝所需要的数据的拷贝量。而且,在发送目的地存储子系统中,使用在实施例 1 中说明的写程序 0111 将拷贝数据写入发送目的地虚拟卷中,因此,关于特定模式数据,可以排除数据的重复。因此,可以削减拷贝目的地存储子系统的必要存储容量。

[0103] 在上述方法中,以成为远程拷贝源的发送源存储子系统内的拷贝源卷是通常卷 0104 为前提来进行说明。但是,拷贝源卷也可以是虚拟卷 0102。当拷贝源卷是虚拟卷 0102 时,在发送源存储子系统中使用页面管理表 0121,针对每个页面来管理该拷贝源虚拟卷 0102,因此,发送源存储子系统掌握了各页面的状态。从而,当页面状态为(a2)时,即使不执行图 20(a)的步骤 1701 以及 1702 的处理,也明了特定模式数据与该页面相对应。因此,也可以省略图 20(a)的步骤 1701、步骤 1702 以及步骤 1703。在这种情况下,首先,发送源存储子系统的发送程序 2103 参照发送源存储子系统具有的页面管理表 0121,判定拷贝源虚拟卷 0102 的各页面的状态。判定的结果,当该状态为(a1)时进入步骤 1702,当该状态为(a2)时省略步骤 1702、1703 而进入步骤 1710,当该状态为(a3)时进入步骤 1704。此后的步骤分别与前面所述的拷贝源为通常卷的情况下相同。

[0104] 而且,在上述方法中,当在发送目的地存储子系统中接收到的数据是模式 ID 时(即图 20(b)的步骤 1706 中标志的值为 1 时),暂时将模式 ID 变更为实际的实际数据(即特定模式数据)(图 20(b)的步骤 1707)后,使用写程序 0111 将该特定模式数据写入虚拟卷。当写程序 0111 将特定模式数据写入虚拟卷时,该特定模式数据再次被变换为模式 ID(图 14 步骤 0801 以及步骤 0802),因此,在特定模式数据与模式 ID 之间发生了无用的变换处理。为了省去该无用的处理,写程序 0111 可以接收模式 ID 来代替接收写数据。结果,数据接收程序 2105 在图 20(b)的步骤 1707 中不将模式 ID 变换为实际数据,而向写程序 0111 传递模式 ID 即可,接收到该模式 ID 的写程序 0111 解释为请求写入通过接收到的模式 ID 指定的规定的特定模式数据,执行写处理即可。

[0105] 在以上的方法中,通过进行远程拷贝,可以实现削减了网络的通信量的数据收发,而且,在多个存储子系统间可以进行容量削减效果高的重复排除。

[0106] 此外,也可以是以下形态。

[0107] 在具有发送源存储子系统、与发送源存储子系统可通信地连接的发送目的地存储子系统的存储系统中,发送源存储子系统以及发送目的地存储子系统具有逻辑存储区域。发送源存储子系统从卷中读出作为远程拷贝对象的数据。当读出的数据与特定模式数据一致时,将识别该特定模式数据的 ID 发送至发送目的地存储子系统。通过这种结构可以进行削减了网络的通信量的数据收发。

[0108] (实施例 4)

[0109] 在本实施例中说明,当与图 17 所示结构同样地经由网络连接了多台存储子系统 0000 时,在该多个存储子系统 0000 间进行容量削减效果高的数据转移。

[0110] 在本实施例中,使用转移源存储子系统具有的数据转移程序 0113,将转移源存储子系统内的全部或部分通常卷 0104 的数据,拷贝到转移目的地存储子系统的虚拟卷 0102。

[0111] 在本实施例中,首先,转移目的地存储子系统使用特定模式的分析程序 0117,分析在转移源存储子系统内的拷贝源通常卷 0104 内存储的数据,调查作为特定模式数据使用哪种模式数据的情况下容量削减效果较高。如后所述,执行特定模式的分析程序 0117 时,需要在数据转移前暂时从拷贝源通常卷读出数据,向拷贝目的地存储子系统进行发送,因此本程序 0117 的执行不是必须的。但是,通过事先对存储在拷贝源通常卷 0104 中的数据进行分析,如果转移目的地存储子系统在已经管理的特定模式数据以外,还可以发现容量削减效果高的特定模式数据,则可以高效地进行重复排除。

[0112] 在图 21 中表示特定模式的分析程序 0117 使用的模式计数表 2001 的一例。模式计数表 2001,用于在特定模式的分析程序 0117 对转移源存储子系统内的拷贝源通常卷 0104 的数据进行了分析时,预先登录该数据的模式和该模式的出现个数。在模式 ID 栏 2011 中记录赋予特定模式的分析程序 0117 从拷贝源通常卷 0104 读出的数据的 ID。与各个模式 ID 对应的数据模式被记录在模式数据栏 2012 中,对于从转移源存储子系统内的拷贝源通常卷 0104 中发现了多少个相同数据模式进行计数的结果,被记录在计数数量栏 2013 中。在模式 ID 2011 中,通过利用散列 (hash) 值等可以减小模式计数表 2001 的容量。

[0113] 图 22 表示特定模式的分析程序 0117 的处理流程的一例。首先,特定模式的分析程序 0117 生成模式计数表 2001 (步骤 2207)。然后,特定模式的分析程序 0117 从转移源存储子系统中针对各个拷贝源通常卷读出其构成信息 (步骤 2208)。所谓拷贝源通常卷的构成信息,是指卷的大小等信息。然后,对各拷贝源通常卷执行以下所示的处理。特定模式的分析程序 0117 将转移源卷的地址空间分割成区段,针对各区段顺次向转移源存储子系统发出读请求 (步骤 2211)。然后,特定模式的读请求的结果,将从转移源存储子系统读取的数据变换为数据 ID。具体而言,确认所读取的数据是否存在于模式计数表 2001 中 (步骤 2212),当数据存在于模式计数表 2001 上的模式数据 2012 中时,使该模式数据对应的计数数量 2013 增加 1 (步骤 2213)。另一方面,在步骤 2212 中,当所读取的数据不存在于模式计数表 2001 上时,向该数据分配新的模式 ID,将该数据和该模式 ID 一起新追加到模式计数表 2001 中,使计数数量为 1 (步骤 2214)。

[0114] 然后,判定特定模式的分析程序 0117 是否在步骤 2208 中针对该拷贝源通常卷内的全部区段读取了数据 (步骤 2215),当针对全部区段结束了读取时进入步骤 2216,判定是否针对成为数据转移对象的全部拷贝源通常卷,结束了步骤 2208 至步骤 2215 的处理 (步骤 2216),当结束时进入步骤 2217。另一方面,当步骤 2215 的判定的结果为,针对拷贝源通常卷内的全部区段读取未结束时,针对剩余的区段重复步骤 2211 ~ 2214,直到针对全部区段结束为止。另外,当步骤 2216 的判定的结果为,针对数据转移对象的全部拷贝源通常卷未结束读入时,针对剩余的拷贝源通常卷重复步骤 2208 ~ 2215,直到针对全部拷贝源通常卷结束读取为止。

[0115] 在步骤 2217 中,特定模式的分析程序 0117 使用模式计数表 2001,分析容量削减效果高的特定模式数据是哪个。在分析容量削减效果高的特定模式数据的方法中考虑了多种方法,但在本说明书中表示使用阈值 δ 的方法。阈值 δ 是在判断某模式的数据是否是特定模式的数据时所使用的阈值。即,通过执行特定模式的分析程序 0117,在某规定的卷中存储了多少个具有某规定模式的区段数据,作为计数数量 2013 被记录在模式计数表 2001 中。当该计数数量超过阈值 δ 时,将该数据认定为特定模式,追加到转移目的地存储子系统具有的特定模式管理表 0120 中。当将阈值 δ 设定得较大时,重复排除的特定模式数减少,因此难以获得容量削减效果。另一方面,当将阈值 δ 设定得较小时,容易获得容量削减效果,但识别的特定模式的个数增大,管理信息量增大。

[0116] 当特定模式的分析结束时,数据转移程序 0113 从转移源存储子系统内的拷贝源通常卷 0104 读取数据,使用写程序 0111 将该数据写入转移目的地存储子系统内的虚拟卷 0102。

[0117] 图 23 是数据转移程序 0113 的处理流程的一例。数据转移程序 0113,与特定模式的分析程序 0117 同样地读出关于转移源存储子系统内的数据转移对象的各拷贝源通常卷的构成信息(步骤 2201),对该拷贝源通常卷执行以下所示的处理。此外,当在先执行了特定模式的分析程序 0117 时,也可以省略步骤 2201 的处理,使用由特定模式的分析程序 0117 读出的构成信息。

[0118] 首先,数据转移程序 0113 在转移目的地存储子系统内生成与拷贝源通常卷相同大小的虚拟卷 0102(步骤 2203),把在此生成的虚拟卷作为数据的转移目的地卷。然后,为了从转移源通常卷的先头区段起顺次读出数据,向转移源存储子系统发出读请求(步骤 2205)。数据转移程序 0113 为了将读请求的结果、即从转移源存储子系统接收的数据写入转移目的地卷,而向写程序 0111 进行指示(步骤 2206)。结果,通过与在实施例 1 中使用图 12~图 15 说明的同样的处理,通过写程序 0111 将数据存储作为转移目的地的虚拟卷中。然后,进入步骤 2207,针对在步骤 2201 中读出构成信息的拷贝源通常卷内的全部区段,判定步骤 2205、2206 是否结束(步骤 2207)。当该拷贝源通常卷内还存在未执行步骤 2205 以及 2206 的处理的区段时,数据转移程序 0113 返回步骤 2205,重复步骤 2205、2206 的处理,直到剩余的区段的处理也结束为止。若针对该拷贝源通常卷内的全部区段结束步骤 2205 及 2206 的处理,则进入步骤 2208。在步骤 2208 中,针对成为数据转移对象的全部拷贝源通常卷,判定步骤 2201~2207 是否已结束,当未结束时返回步骤 2201,重复步骤 2201~2207,直到剩余的拷贝源通常卷的处理也结束为止。通过步骤 2208,当针对数据转移对象的全部拷贝源卷结束步骤 2201~2207 的处理时,结束数据转移程序 0113。

[0119] 通过以上方法,在发送目的地存储子系统中使用在实施例 1 中说明的写程序 0111 将数据写入转移目的地虚拟卷,因此,针对特定模式数据可以排除数据的重复,可以削减转移目的地存储子系统的必要存储容量。

[0120] 此外,在执行特定模式的分析程序 0117 后执行数据转移程序 0113 时,通过两个程序从拷贝源通常卷读出总计 2 次数据,从转移源存储子系统拷贝到转移目的地存储子系统,网络中的通信负荷增加。因此,为了防止这种情况,特定模式的分析程序 0117,可以通过图 22 的步骤 2211 从拷贝源存储子系统读出的数据暂时保存在拷贝目的地存储子系统的高速缓冲存储器 0005 或通常卷 0104 中。这样一来,数据转移程序 0113 如图 23 的步骤

2205 中描述的那样,代替从拷贝源存储子系统的拷贝源通常卷再次读出数据,而使用在拷贝目的地存储子系统的高速缓冲存储器 0005 或通常卷 0104 中存储的数据即可,可以减少从拷贝源存储子系统向拷贝目的地存储子系统的数据库发送量。

[0121] (实施例 5)

[0122] 在本实施例中,除了实施例 1 中说明的重复排除功能以外,说明通过使状态为 (a2) 的页面中也存在特定模式数据以外的数据(称为非特定模式数据),根据页面内的非特定模式数据的区段数使页面的状态细致地变迁,由此更高效地进行存储子系统的容量削减的方法。主要说明与实施例 1 的不同点。

[0123] 图 24 表示本实施例中的存储子系统 1111 的逻辑系统结构的一例。在图 24 所示的存储子系统 1111 中,与实施例 1 不同,为了在状态为 (a2) 的页面中存储非特定模式数据,存在分配存储区域时所使用的区段用池 0107。在区段用池 0107 中存在多个区段用卷 0106。区段用卷 0106 与通常卷 0104 或页面用卷 0103 相同,是在由多个 HDD0006 构成的 RAID 组等磁盘组上形成的存储区域。

[0124] 本实施例中的存储子系统 1111 还与管理页面用池 0105 的池管理表 0123 不同地,具有用于管理区段用池 0107 的区段用池管理表 0125。区段用池管理表的结构与在实施例 1 中使用图 5 说明的池管理表 0123 的结构相同,因此未图示。此外,在区段用池管理表 0125 中,池 LU# 是具有分配给该区段的物理区域的区段用卷 0106 的识别号码。另外,分配给区域的各物理区域的大小与区段的大小相等,因此,通过区域 # 识别的分配给区段的物理区域,是从通过池 LU# 识别的区段用卷 0106 的开始地址所表示的地址起区段大小的存储区域。

[0125] 在本实施方式的存储子系统 1111 中,区段管理表 0126 的结构与实施例 1 中的区段管理表 0122 不同。图 25 是表示本实施例中使用的区段管理表 0126 的一例的图。在实施例 1 中说明的图 4 的区段管理表 0122 中的标志栏 0209 中记录了:记录了表示已写入特定模式数据的意思的状态“0”;和对于区段尚未取得写请求的状态“1”的某个值。但在本实施例中,除了这些状态以外,也可以在区段管理表 0126 的标志栏 1262 中记录在区段中存储了非特定模式数据的状态“2”。而且,本实施方式中的区段管理表 0126 具有区域号码(区域 # 栏)1263。在标志栏 1262 中记录了“2”时使用区域 # 栏 1263,记录向该区段分配的区段用池 0107 内的物理区域的区域号码。

[0126] 在本实施方式中,阈值管理表 0127 也和实施例 1 中的阈值管理表 0124 不同。在本实施方式中使用的图 26 表示阈值管理表 0127 的一例。在本实施例中,当页面从 (a1) 向 (a2)、或者从 (a2) 向 (a1) 进行状态迁移时,使用与非特定模式数据数相关的阈值 α 、阈值 β 。在状态为 (a2) 的页面中,当非特定模式数据的数量增加,超过阈值 α 时,使页面的状态从 (a2) 向 (a1) 变迁。另外,在状态为 (a1) 的页面中,当非特定模式数据的数量减少,低于阈值 β 时,使页面的状态从 (a1) 向 (a2) 变迁。阈值 α 、阈值 β 可以使用阈值变更程序 0116 来变更。变更阈值的处理流程与实施例 1 相同,因此省略。此外,假定本实施例中的阈值 α 为 2 以上。

[0127] 然后,关于本实施例中的写程序 0128 的处理流程,表示与实施例 1 的不同点。

[0128] 首先,当写对象页面的状态为 (a1) 时所执行的 Write 处理 (1) 与实施例 1 不同。图 27 中表示本实施方式中的 Write 处理 (1) 的处理流程的一例。在本处理流程中,为了使页面的状态从 (a1) 向 (a2) 迁移而需要调查页面内的非特定模式数据数量的步骤,这一点

与实施例 1 中的 Write 处理 (1) 不同。

[0129] 从步骤 0701 到步骤 0706 进行与实施例 1 同样的处理,因此省略说明。在步骤 1301 中,写程序 0128 检查非特定模式数据的数量,确认是否小于阈值 β 。非特定模式数据的数量,可以通过对写对象页面中存储的全部数据进行检查来确认,也可以在页面管理表中新设置用于记录非特定模式计数数量的栏,每当写入非特定模式数据时使该栏中记录的计数值每次增加 1,由此来确认非特定模式数据的数量。

[0130] 当非特定模式数据的数量小于阈值 β 时,首先,写程序 0128 生成与该页面对应的区段管理表 0126 (步骤 1302)。然后,按照以下的步骤将该页面内的区段的信息记录在生成的区段管理表 0126 中。

[0131] 写程序 0128 参照特定模式管理表 0120 顺次读取页面内的区段数据,来判定区段内的数据是否与特定模式的某一种相一致 (步骤 1303)。当写入该区段的数据与特定模式数据的某一种一致时,在该区段所对应的标志栏 1262 中设定 0,将对应的 ID 记录在模式 ID 栏 1264 中 (步骤 1304)。在此,由于不是非特定模式数据,因此在区域 #1263 中记录 null。

[0132] 在步骤 1302 中读取的区段的数据与特定模式中任意一种都不一致时,写程序 0128 从区段用池 0107 中确保物理区域 (步骤 1306),在所确保的区域中写入该数据 (步骤 1307)。然后,在区段管理表 0126 的标志栏 1262 中记录 2,记录所确保的区段用池 0107 的区域 #1263。由于不是特定模式数据,因此在模式 ID1264 中记录 null (步骤 1308)。

[0133] 在步骤 1311 中,判定是否针对页面内的全部区段已结束从步骤 1303 至步骤 1308 的处理,当未结束时返回步骤 1303,针对剩余的区域重复上述处理。

[0134] 通过以上的步骤,页面的状态从 (a1) 迁移到 (a2),因此页面管理表 0121 的该页面的状态 0204 从 0 被变更为 1 (步骤 1309)。

[0135] 最后,释放被分配给该页面的页面用池 0105 内的物理区域 (步骤 1310)。

[0136] 通过以上的图 27 的流程,即使是暂时分配了物理区域的页面,若成为在该页面内特定模式数据多、非特定模式数据少的状态,则仅向非特定模式数据分配物理区域,因此可以实现容量削减效果高的管理。

[0137] 在本实施方式中,当写对象页面的状态为 (a2) 时,通过写程序 0128 执行的 Write 处理 (2) 也和实施例 1 不同。在图 28 中表示本实施方式中的 Write 处理 (2) 的处理流程的一例。

[0138] 写数据与特定模式数据的某一种一致时的处理流程 (步骤 0801 至步骤 0804) 与实施例 1 相同,因此省略说明。

[0139] 在步骤 0801 中,当写数据与特定模式中的任意一种都不一致时,写程序 0128 取得写对象页面用的区段管理表 0126,将写对象区段的标志栏 1262 设定为 2 (步骤 1404)。然后,进入步骤 1405,写程序 0128 参照区段管理表 0126 的标志栏 1262,对于在写对象页面中存在多少个标志为 2 的区段进行计数。而且,写程序 0128 参照阈值管理表 0127,取得与该页面对应的阈值 α 1272 的值,判定在先计数的标志为 2 的区段的数量 (即非特定模式数据的数量) 是否超过了阈值 α 的值 (步骤 1405)。在步骤 1405 中,非特定模式数据的数量小于阈值 α 时的处理成为步骤 1406 ~ 1408,但由于和图 27 的步骤 1306 ~ 1308 相同,因此省略说明。在步骤 1405 中,当非特定模式数据的数量超过阈值 α 的值时,写程序 0128 进入步骤 1410,从页面用池 0105 向写对象页面分配物理区域 (步骤 1410)。然后,参照与该页

面对应的区段管理表 0126, 针对标志 1262 为 0 的区段, 参照该区段所对应的模式 ID1264, 取得模式管理表 0121, 将对应的特定模式数据存储并确保区域的该区段中。针对标志 1262 为 1 的区段, 在所确保的区域的与该区段对应的区域中存储格式数据。当标志 1262 为 2 时, 参照区域 #1263, 从池管理表 0123 取得与该区段对应的池 LU#1234 和开始地址 1235, 将该地址中存储的数据拷贝到确保区域的该区段 (步骤 1411)。然后, 写程序 0128 释放与该页面对应的区段管理表 0126 (步骤 1412), 将与该页面对应的页面管理表 0121 的状态 0204 从 1 变更为 0, 然后结束 (步骤 1413)。

[0140] 通过以上的图 28 的流程, 即使在未分配物理区域的区域中接收到非特定模式数据的写请求的情况下, 当非特定模式数据数量少时也不向接收到写请求的页面全体分配物理, 而仅向非特定模式数据分配物理区域, 由此可以提高存储的容量削减效果。

[0141] 此外, 当状态为 (a3) 的页面中接收到写请求时的写程序 0128 的处理流程与实施例 1 中说明的 Write 处理 (3) 相同, 因此省略说明。

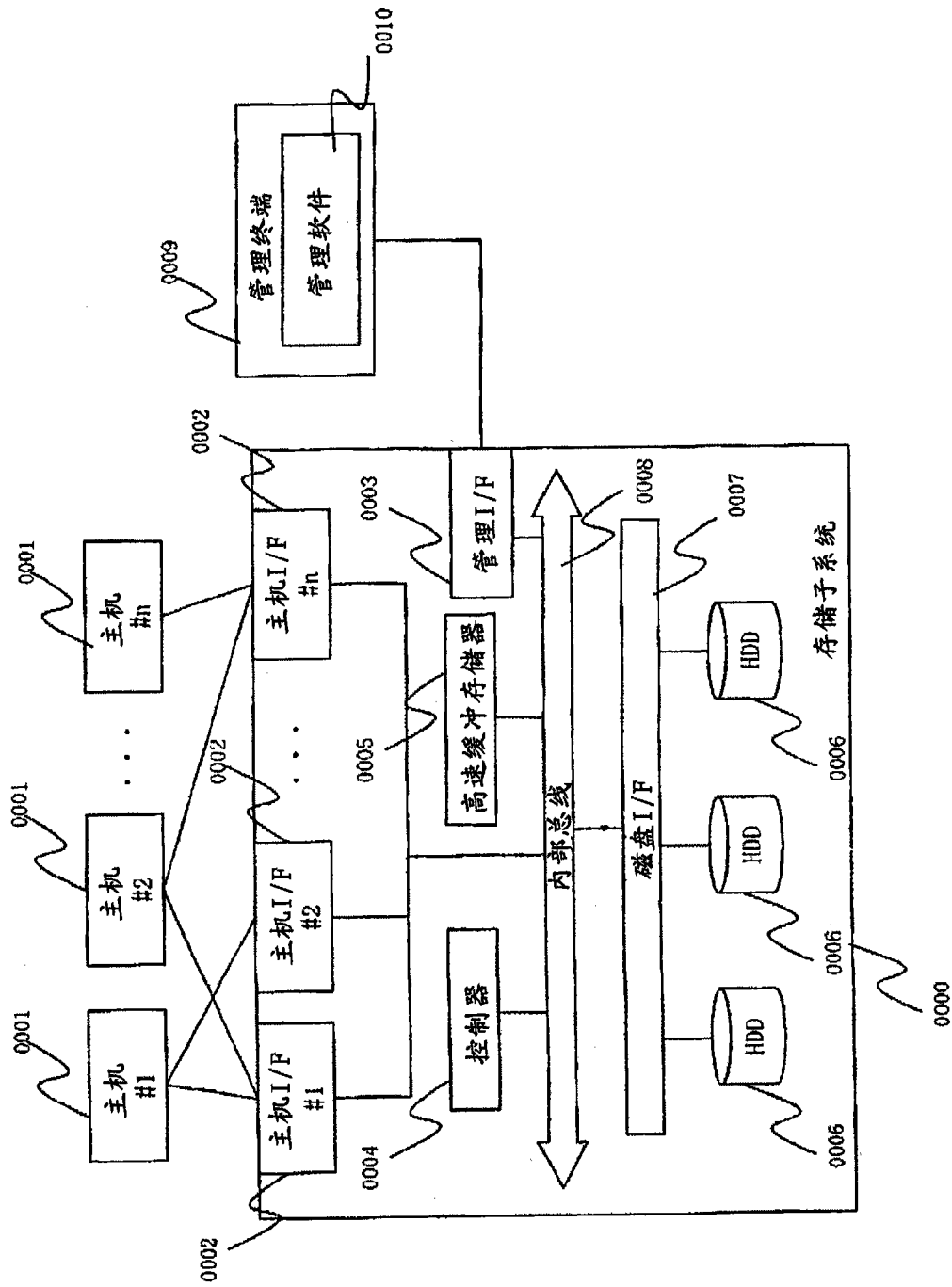


图 1

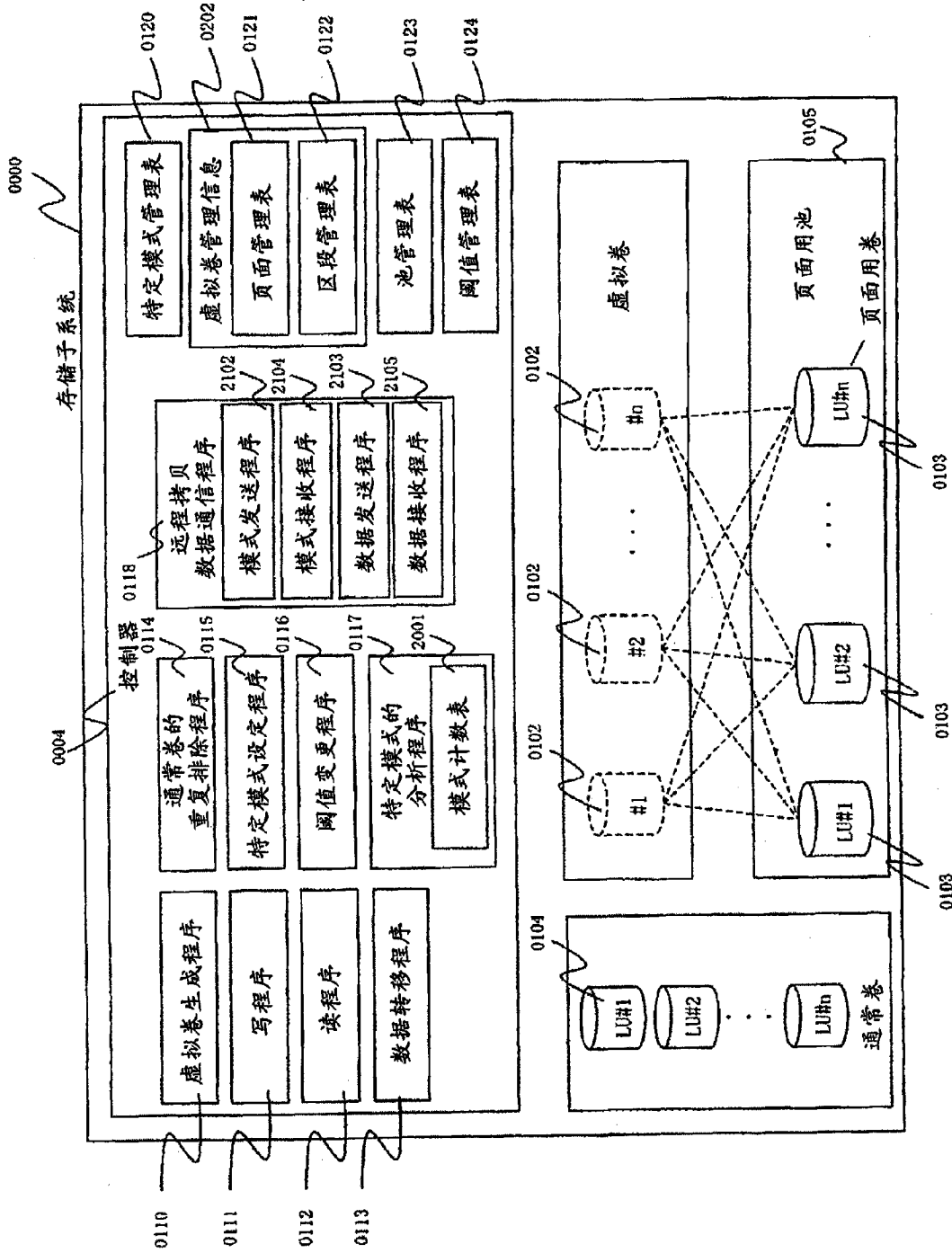


图 2

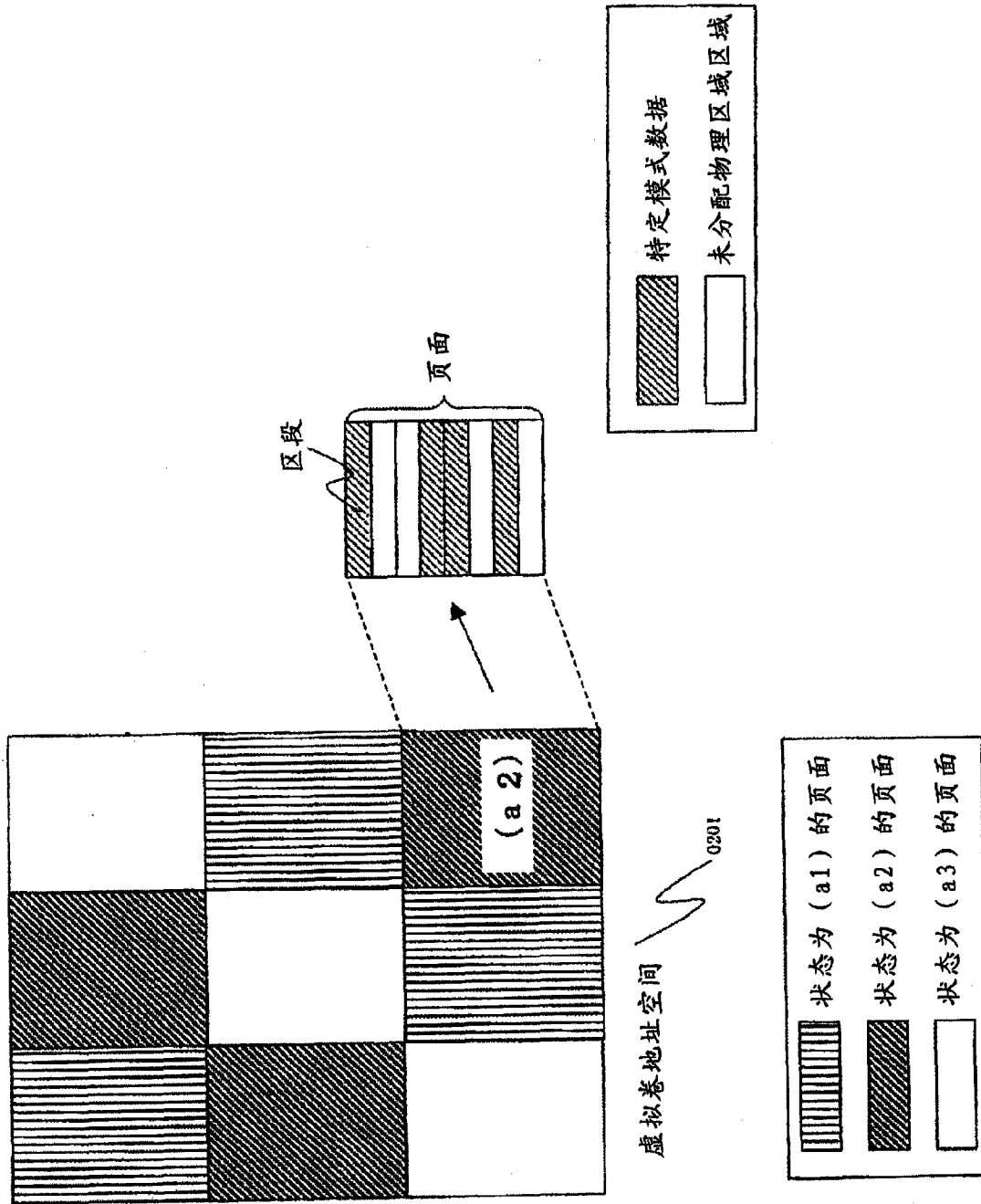


图 3

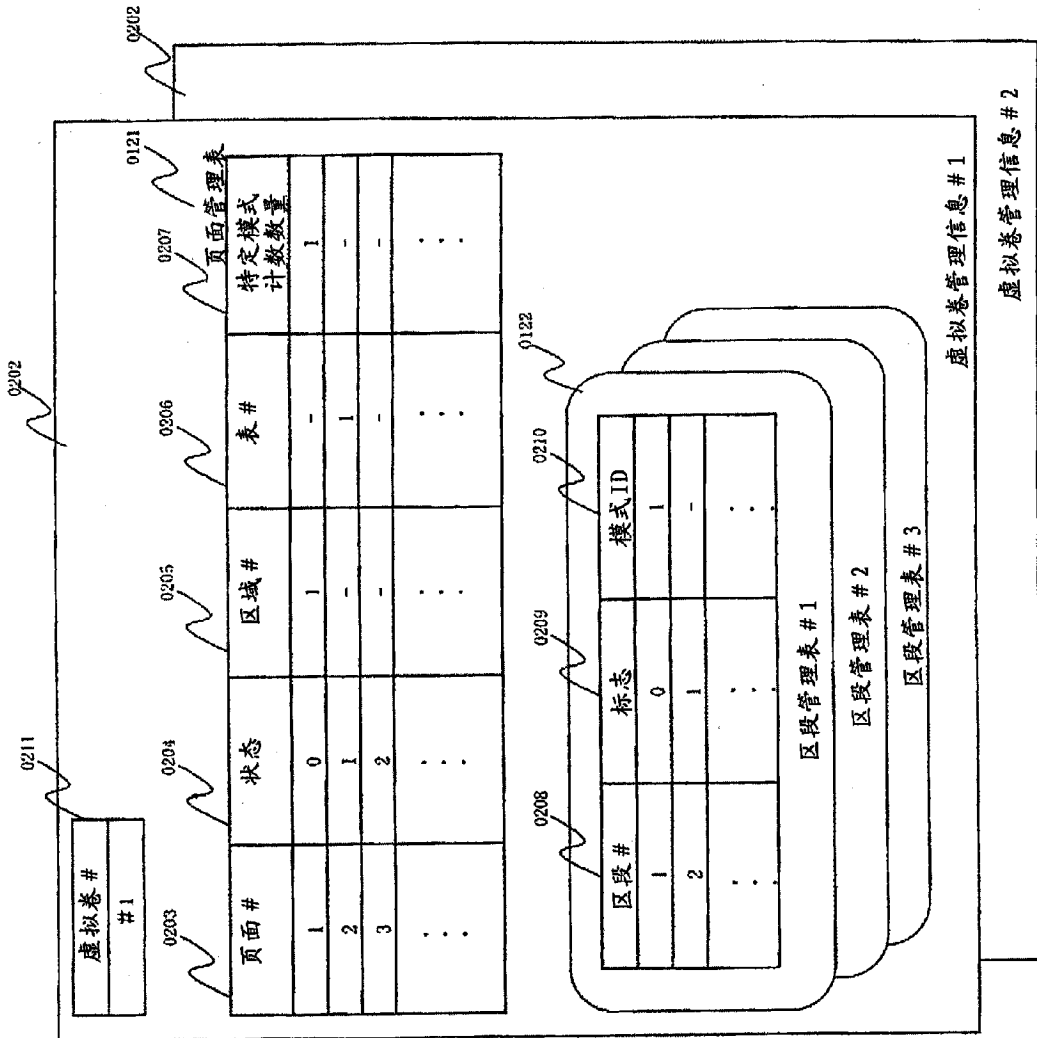


图 4

0123

1231

1232

1234

1235

池管理表

区域#	分配标志	池LU#	开始地址
1	0	1	AAA
2	1	1	BBB

图 5

特定模式管理表

模式ID	模式数据
1	00 00
2	abc xyz
.	.
.	.
.	.
n	ff ff

图 6

阈值管理表

虚拟卷 #	阈值 γ
1	5
2	6
.	.
.	.
.	.

图 7

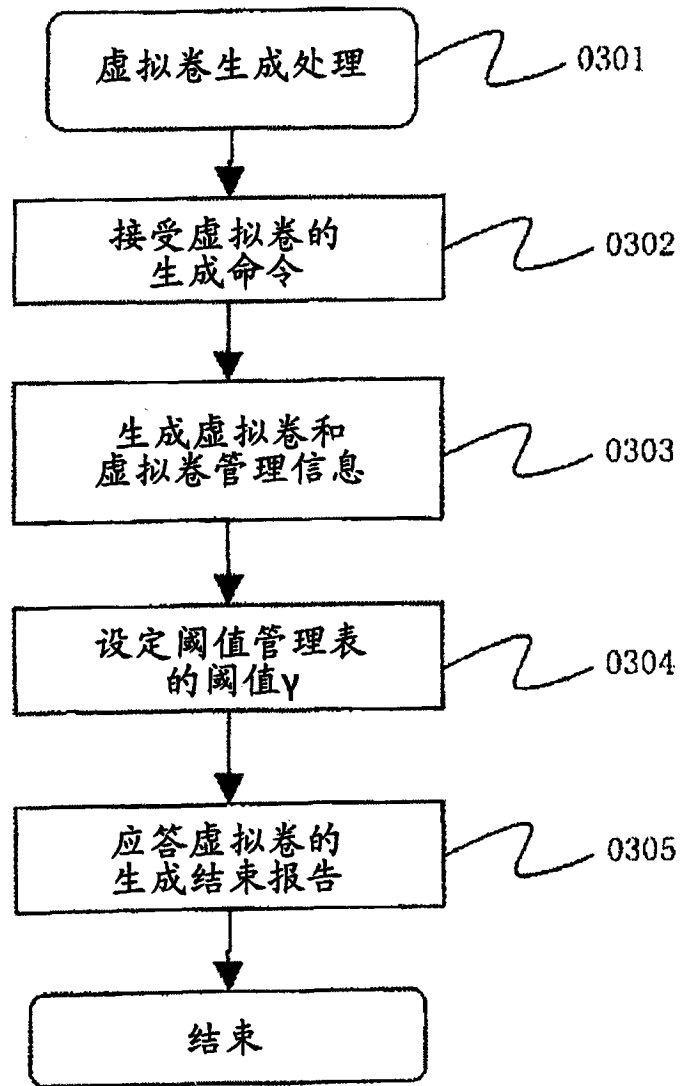


图 8

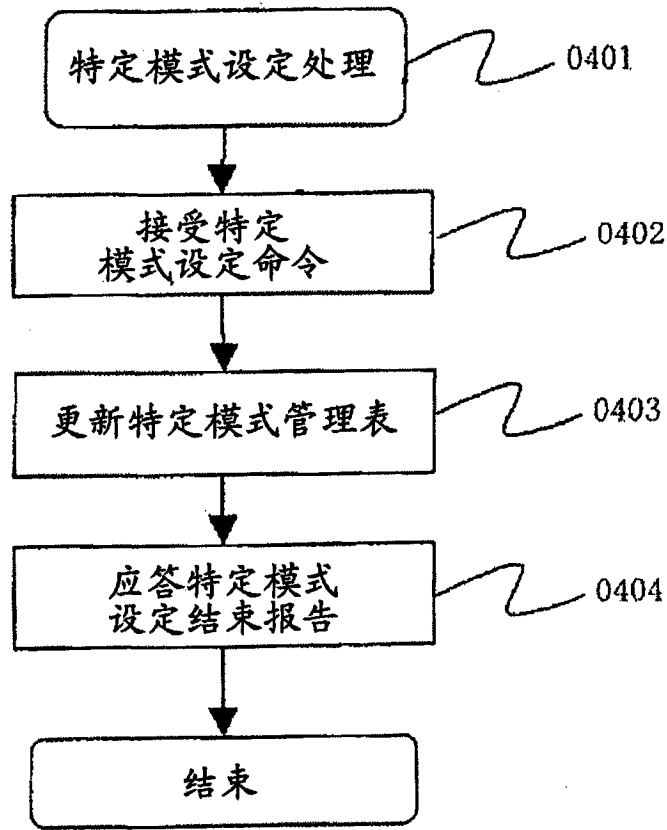


图 9

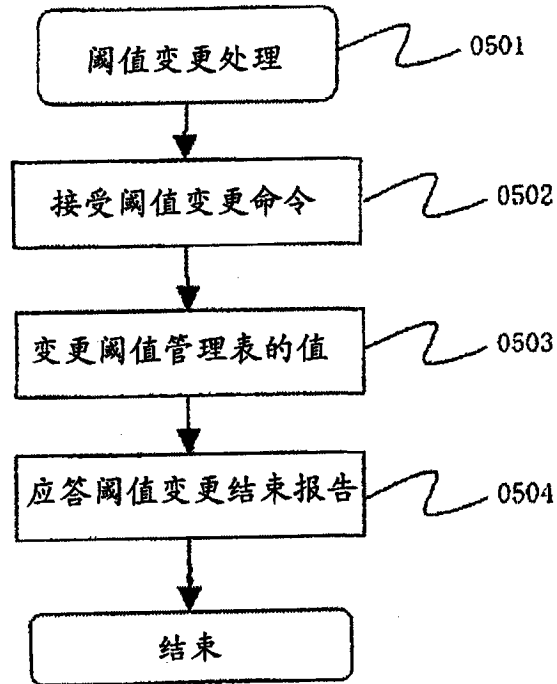


图 10

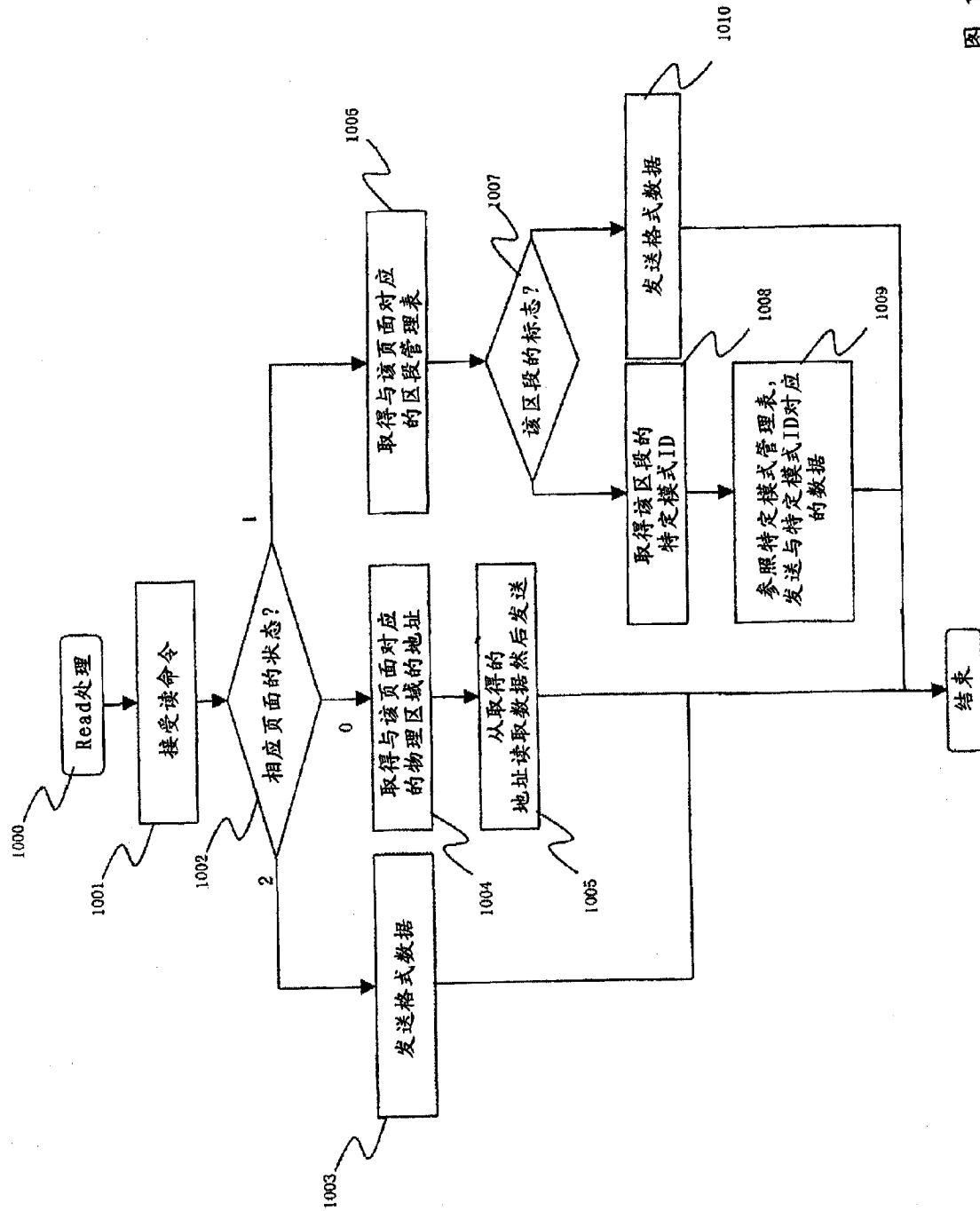


图 11

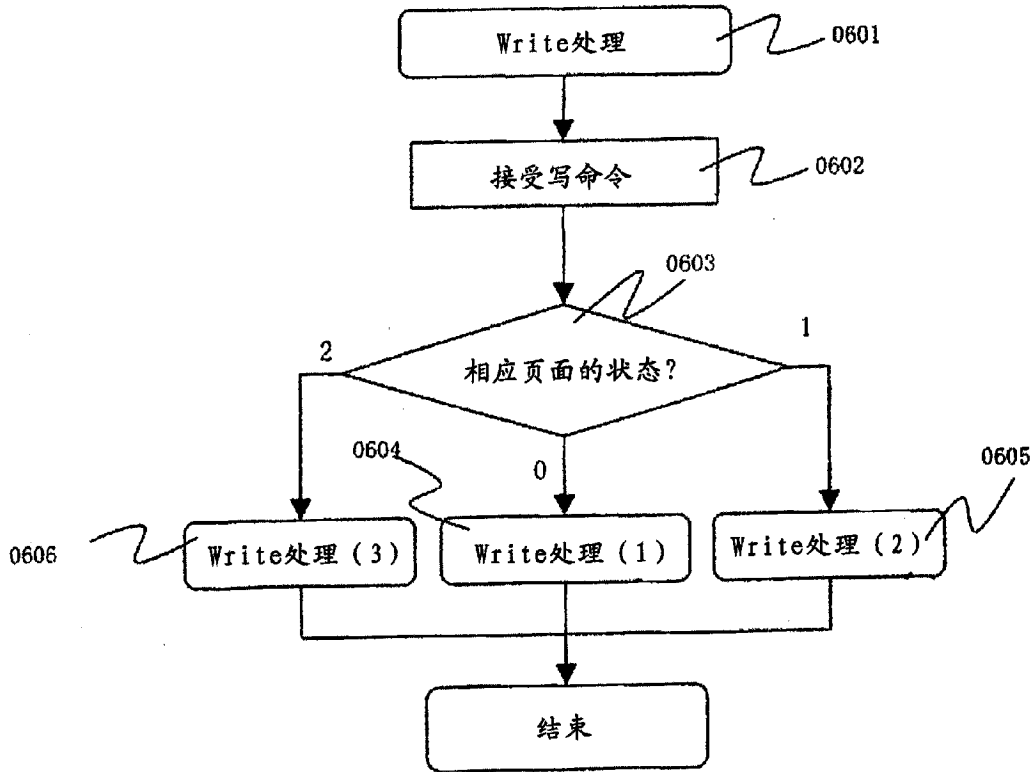


图 12

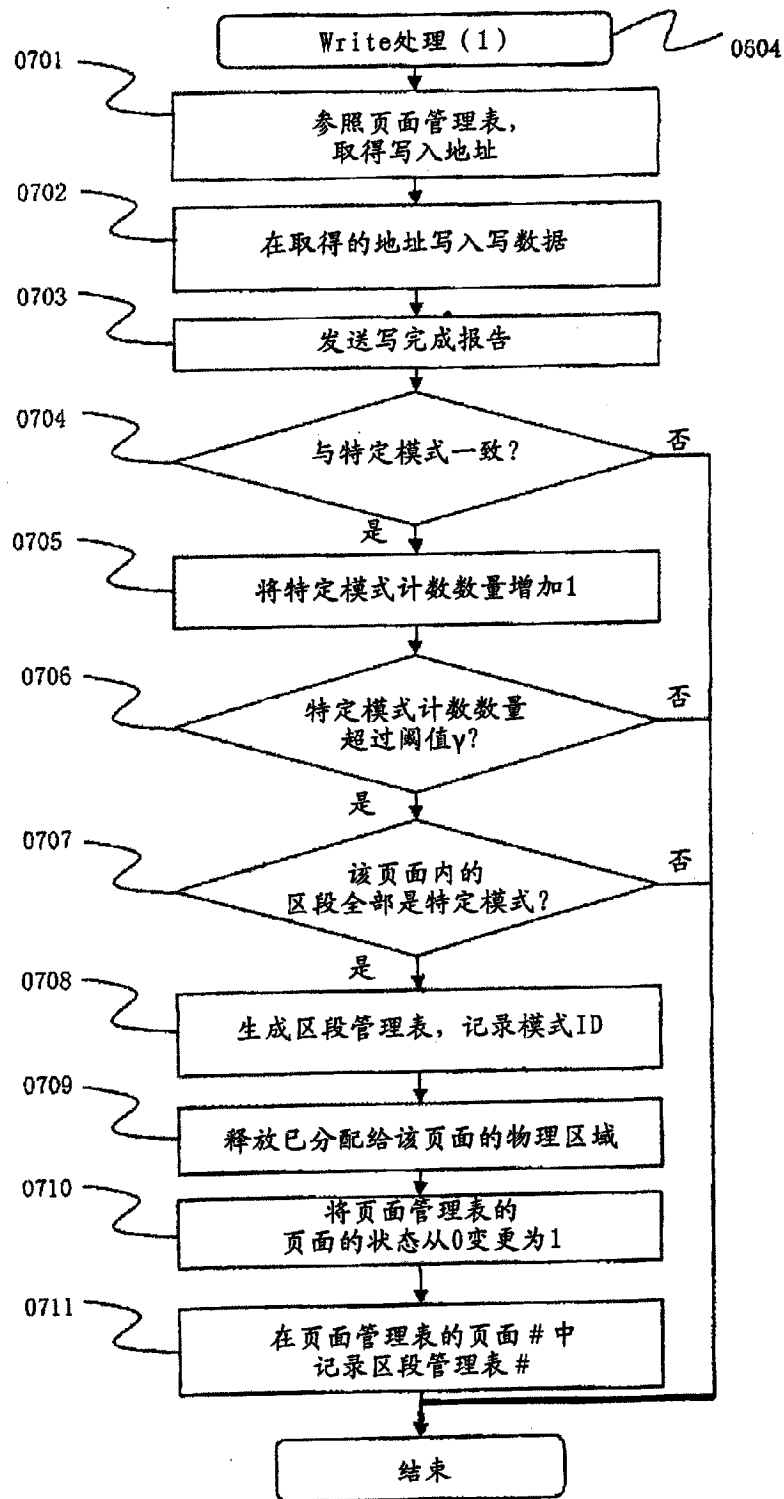


图 13

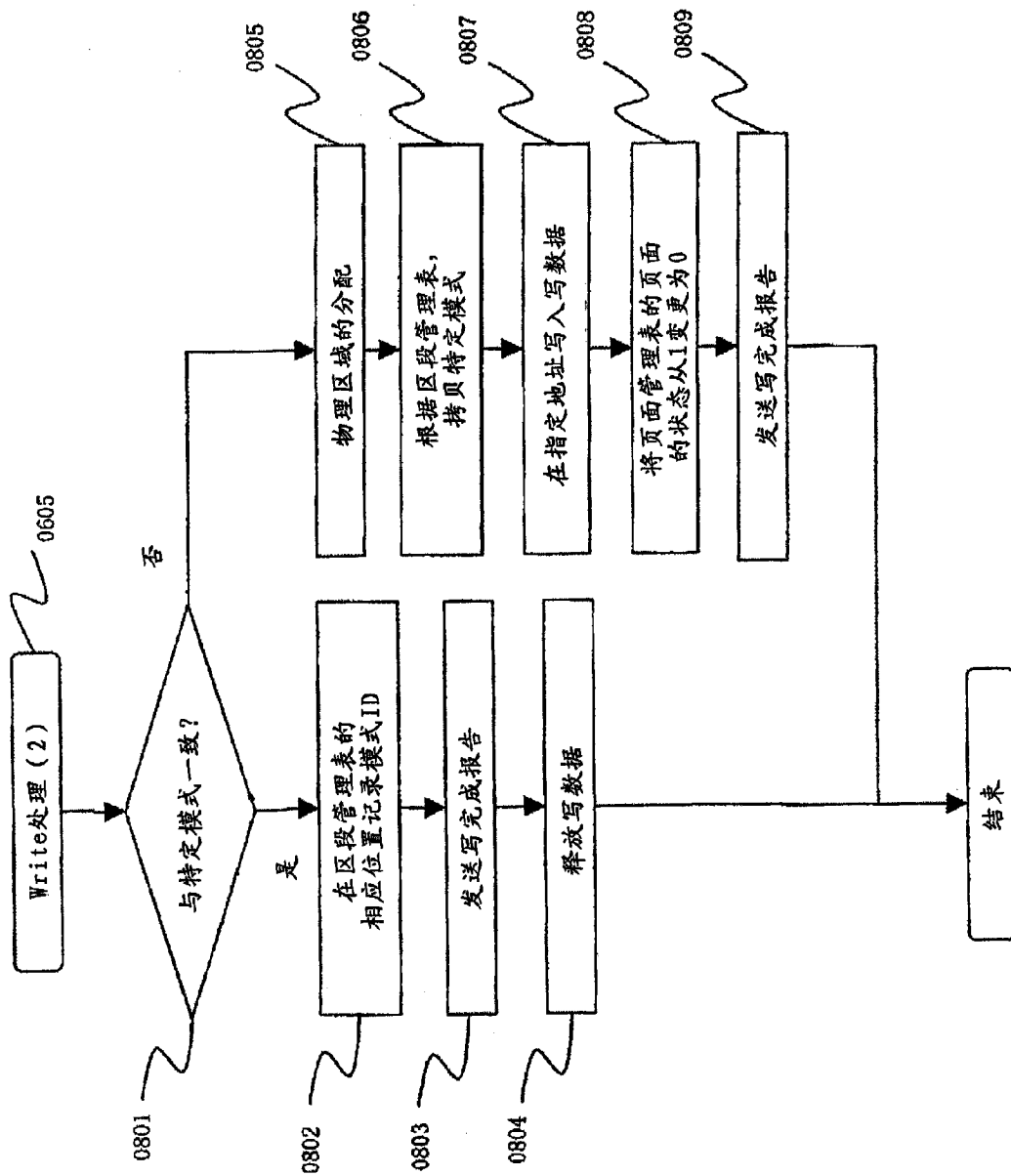


图 14

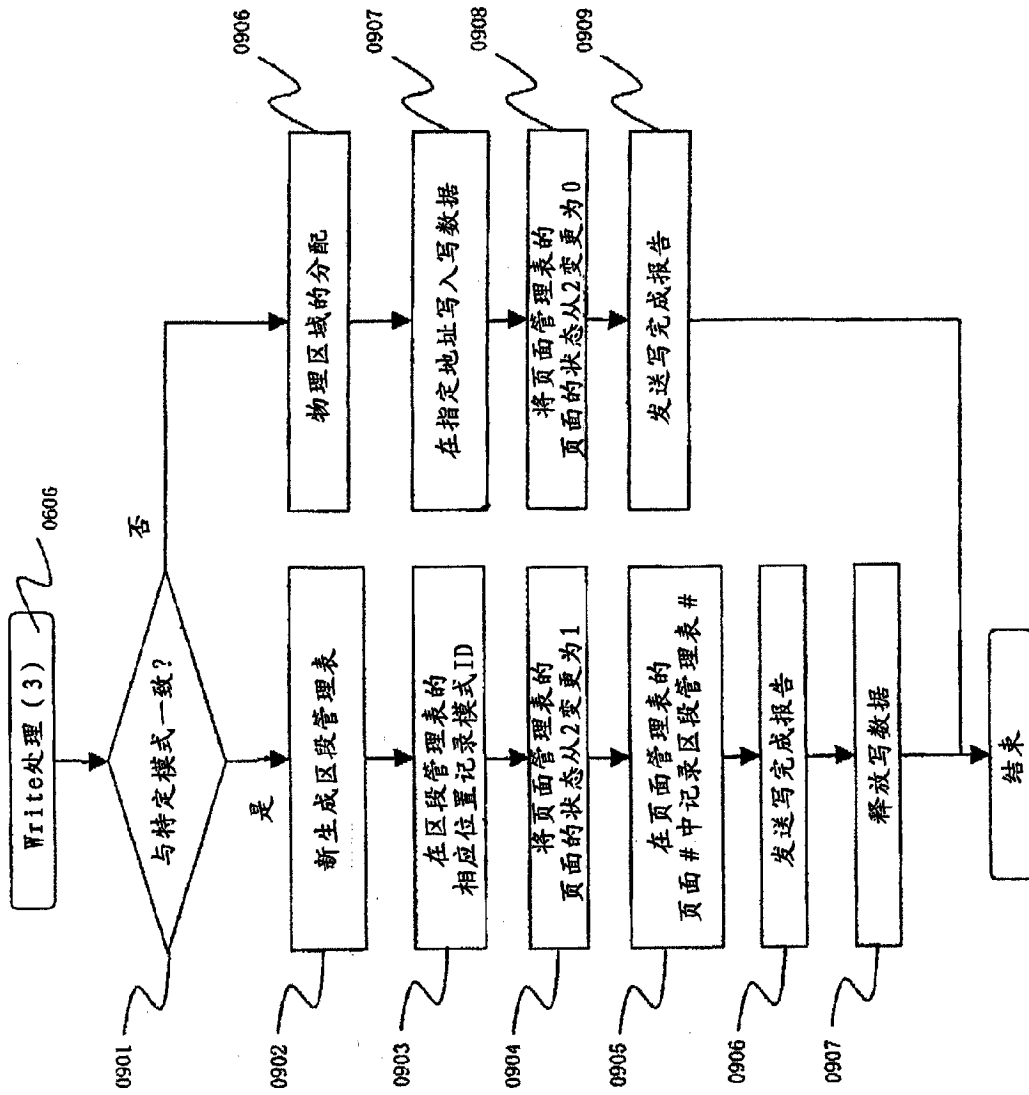


图 15

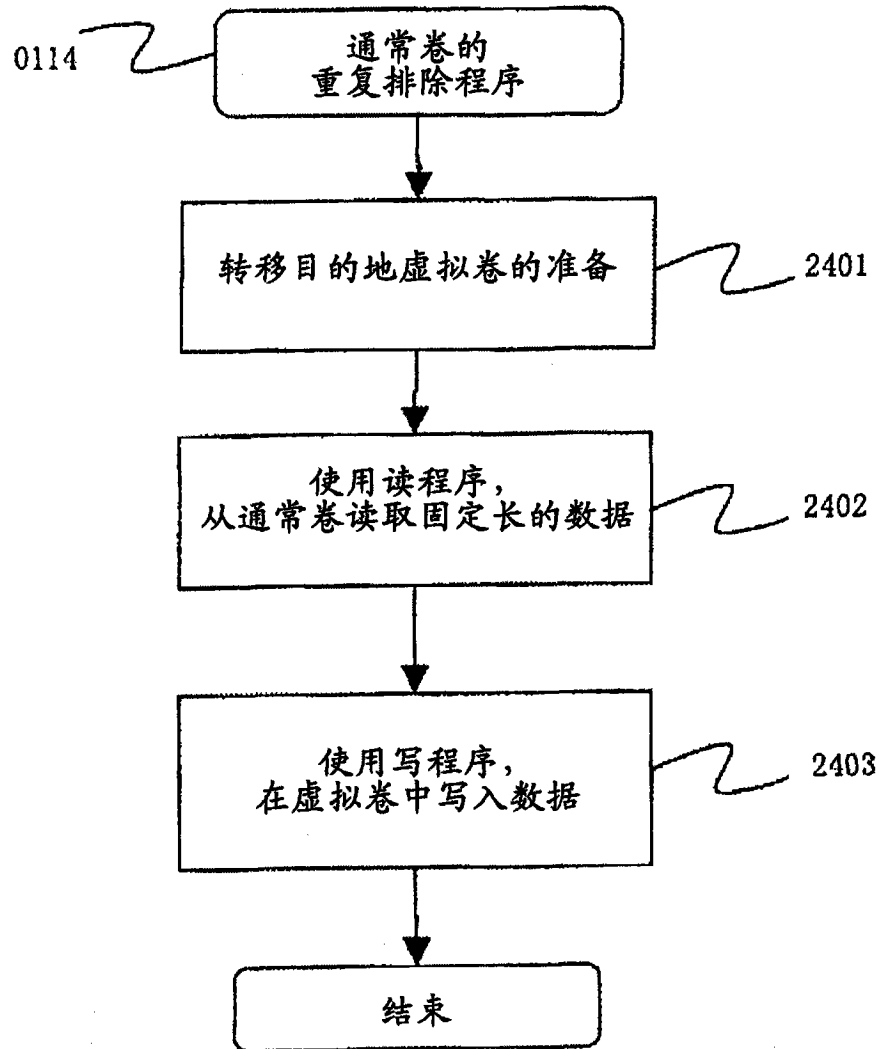


图 16

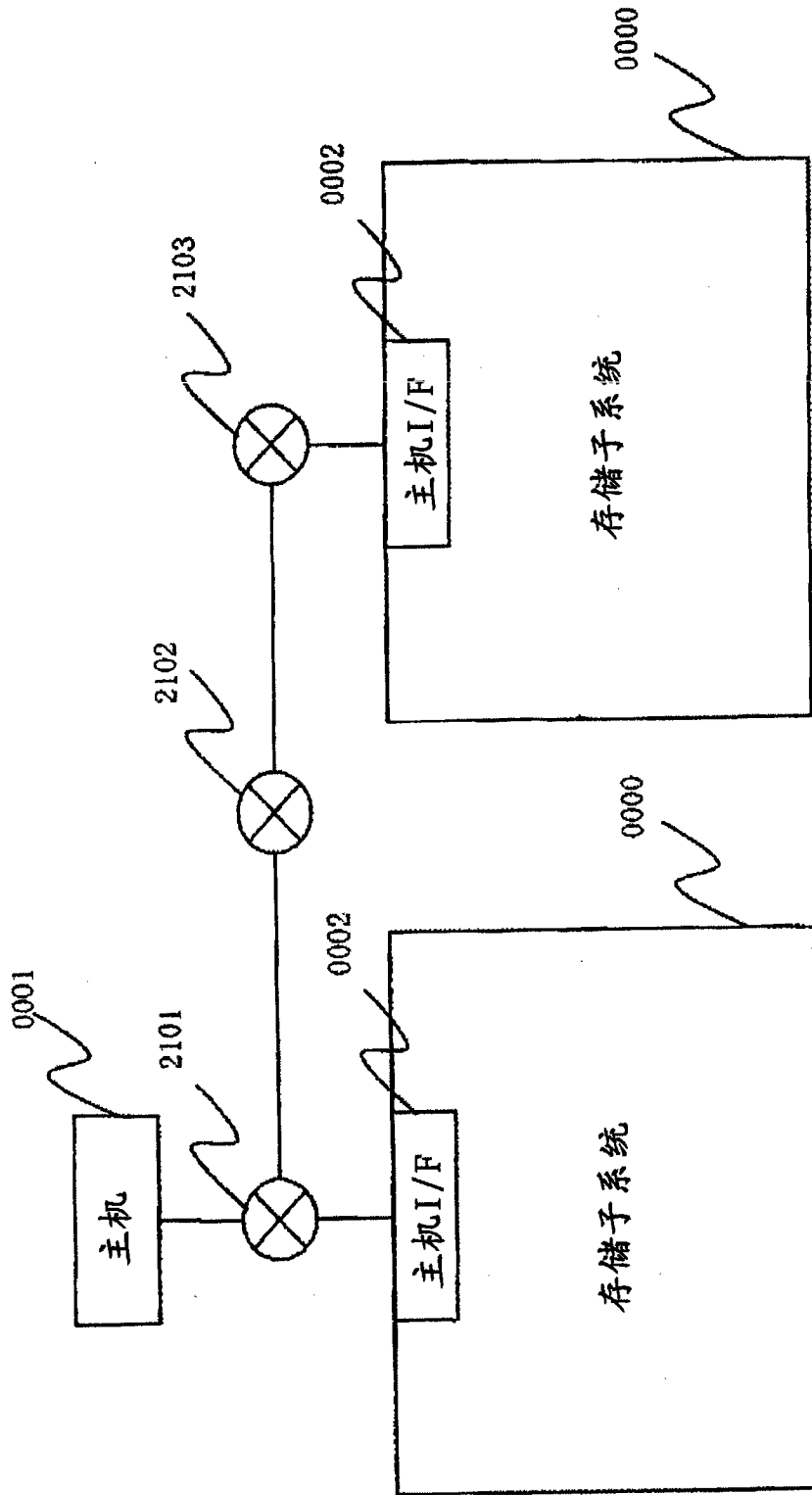


图 17

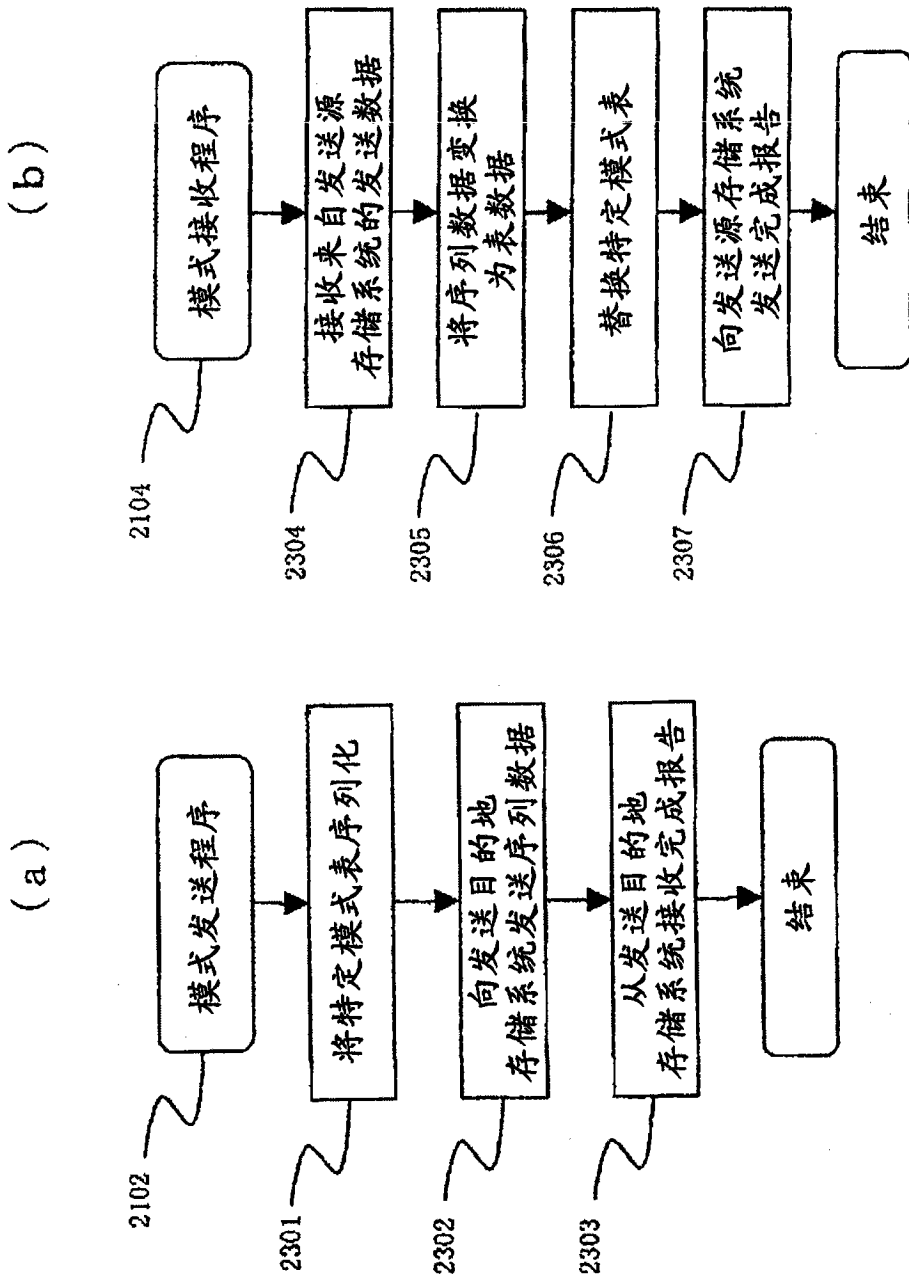


图 18

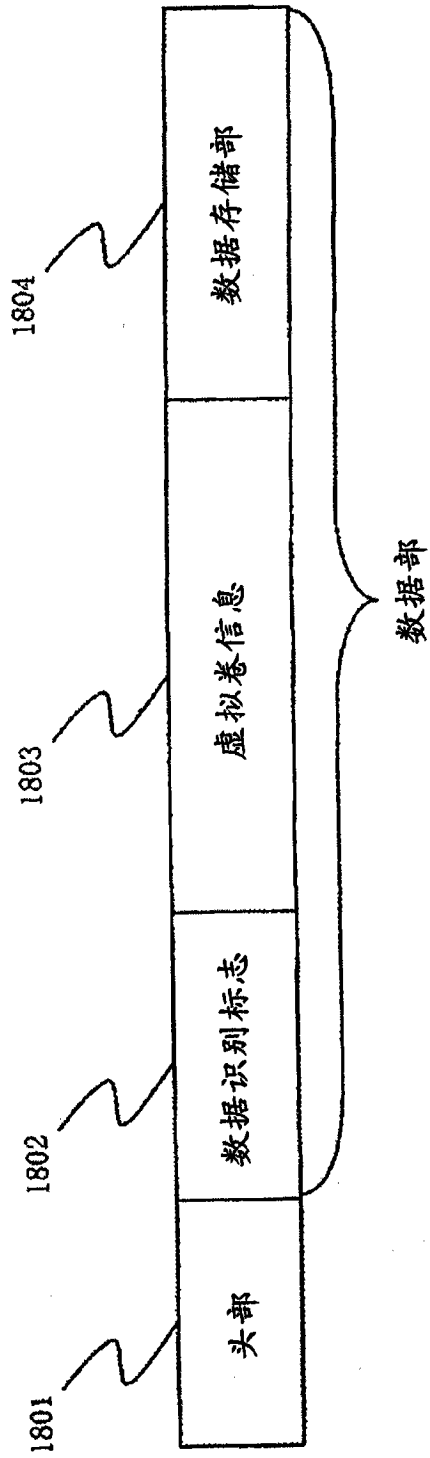


图 19

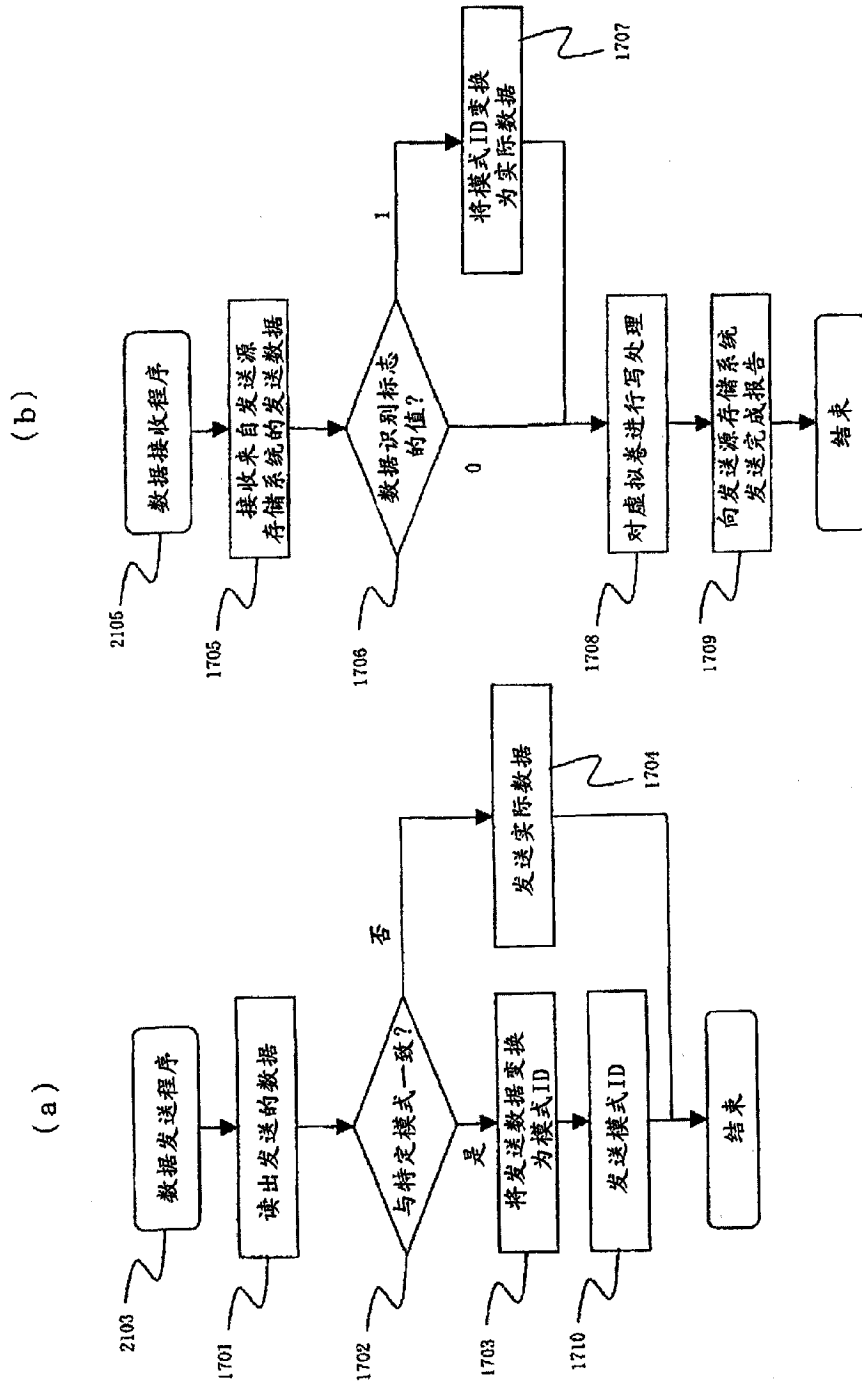


图 20

2001

2011 2012 2013

模式计数表

模式ID	模式数据	计数数量
1	00 00	2
2	abc xyz	3
⋮	⋮	⋮
n	ff ff	1

图 21

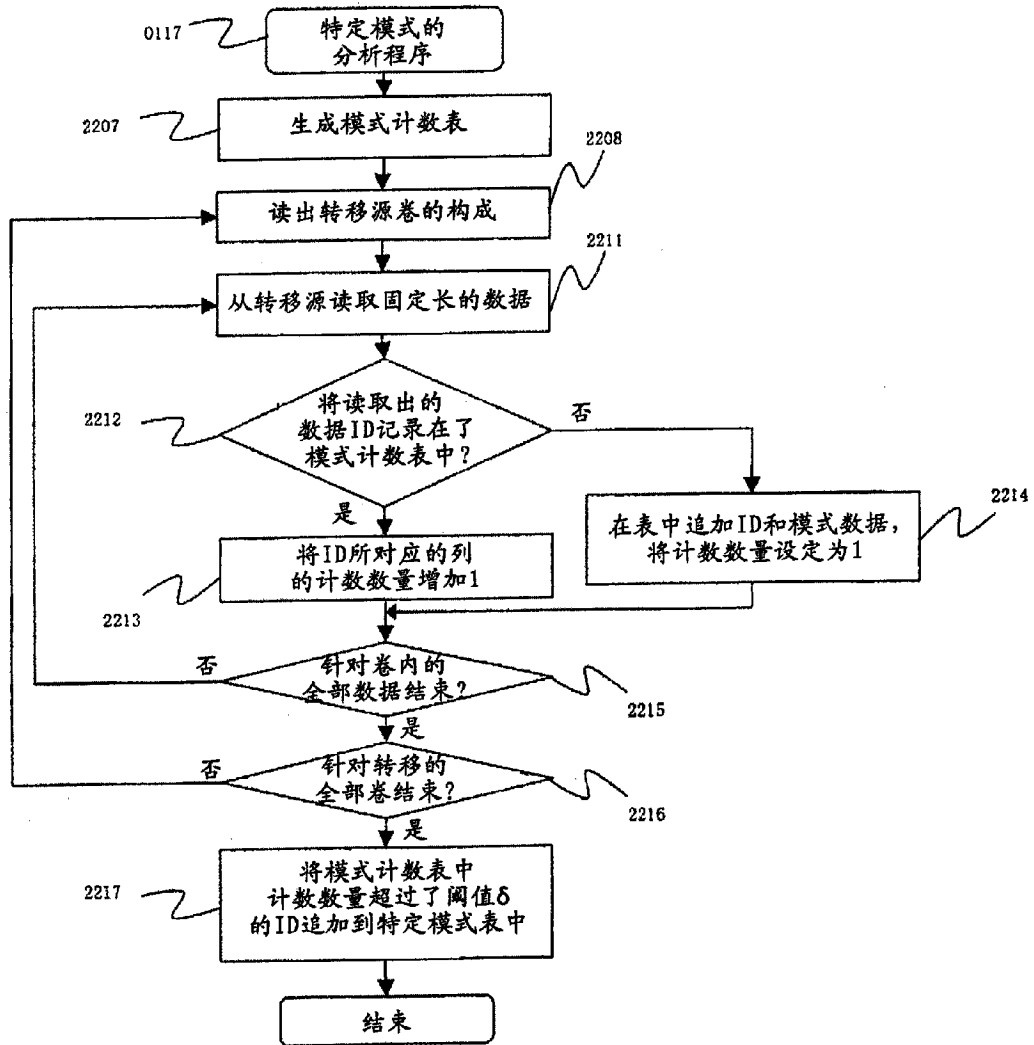


图 22

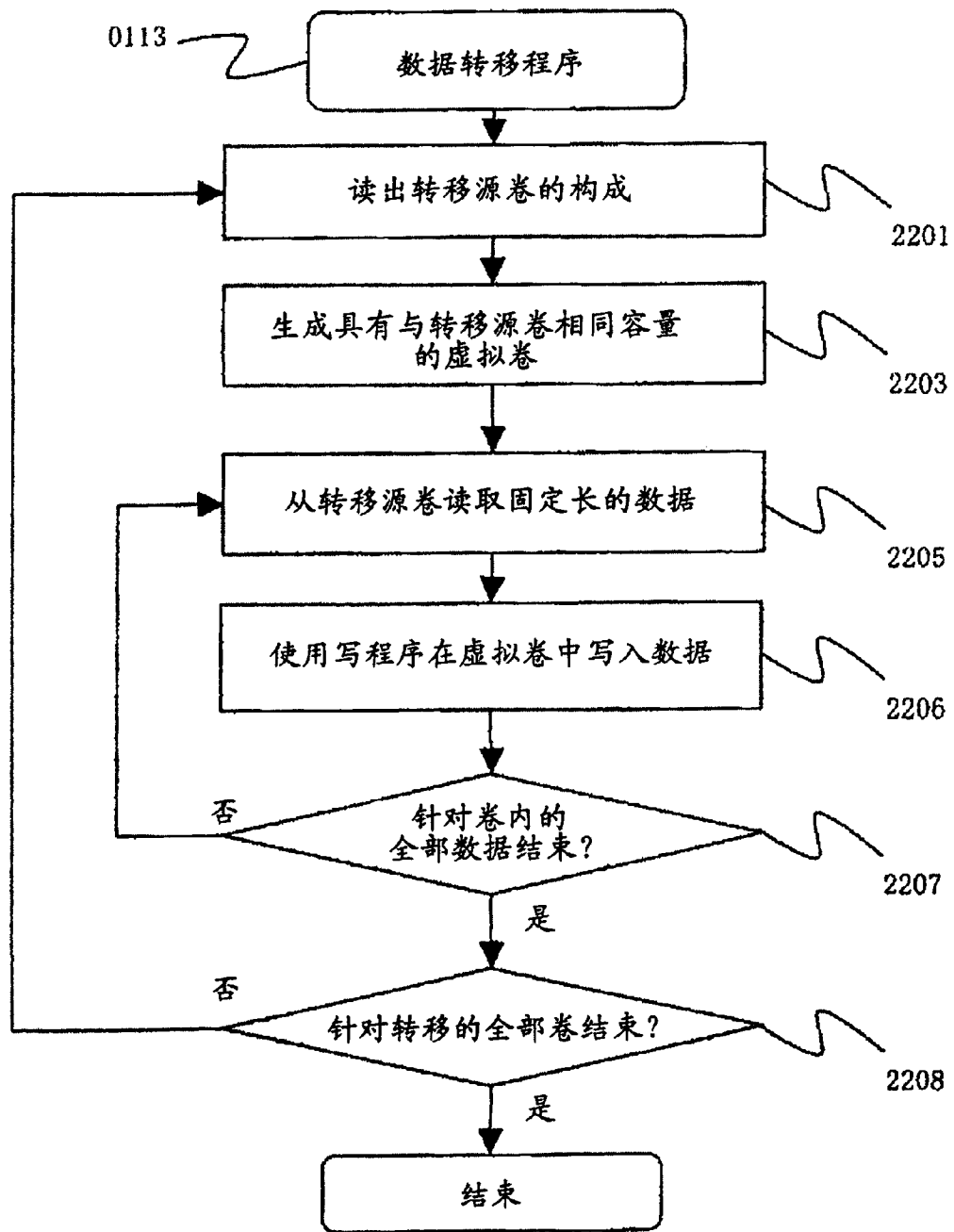


图 23

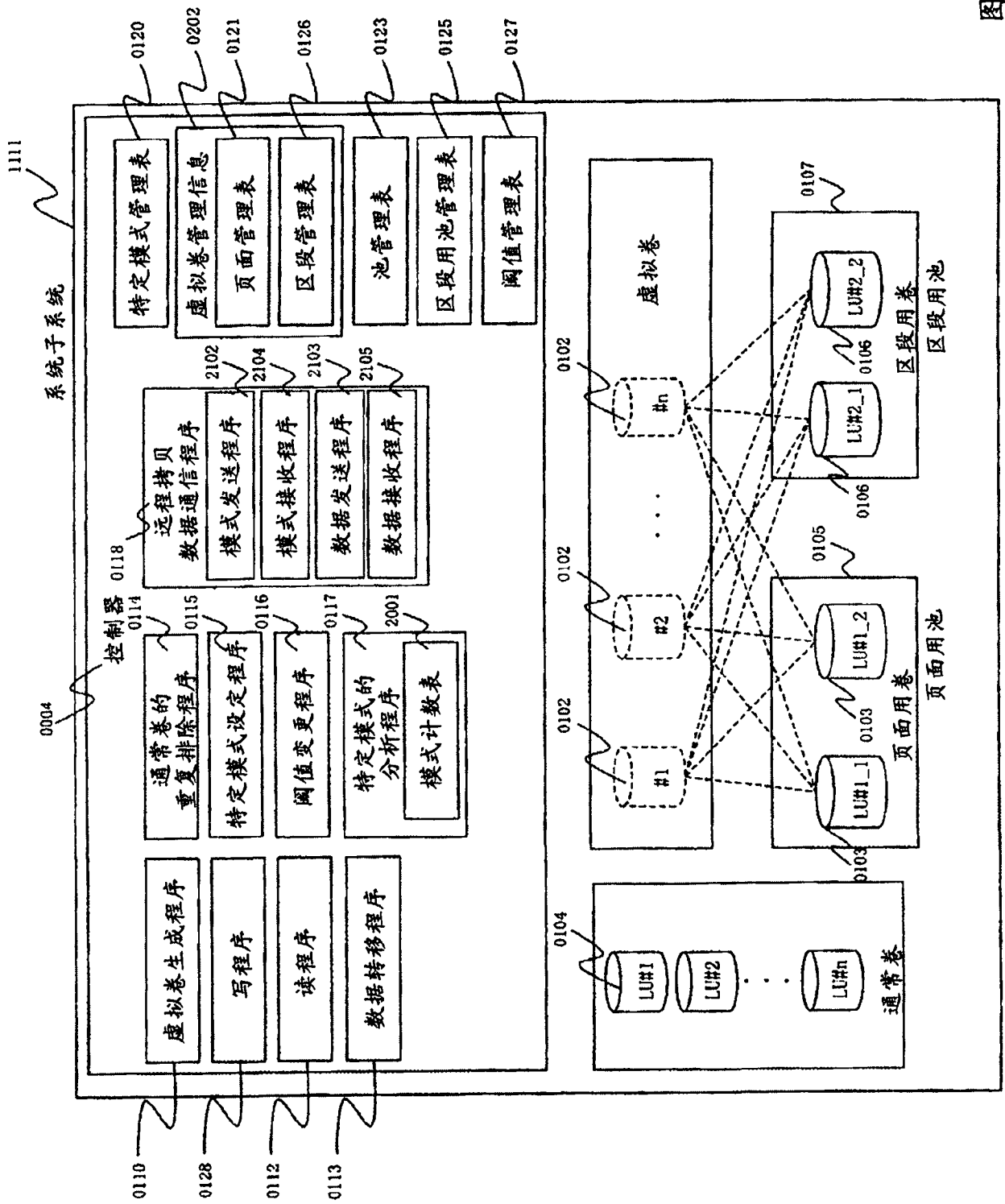


图 24

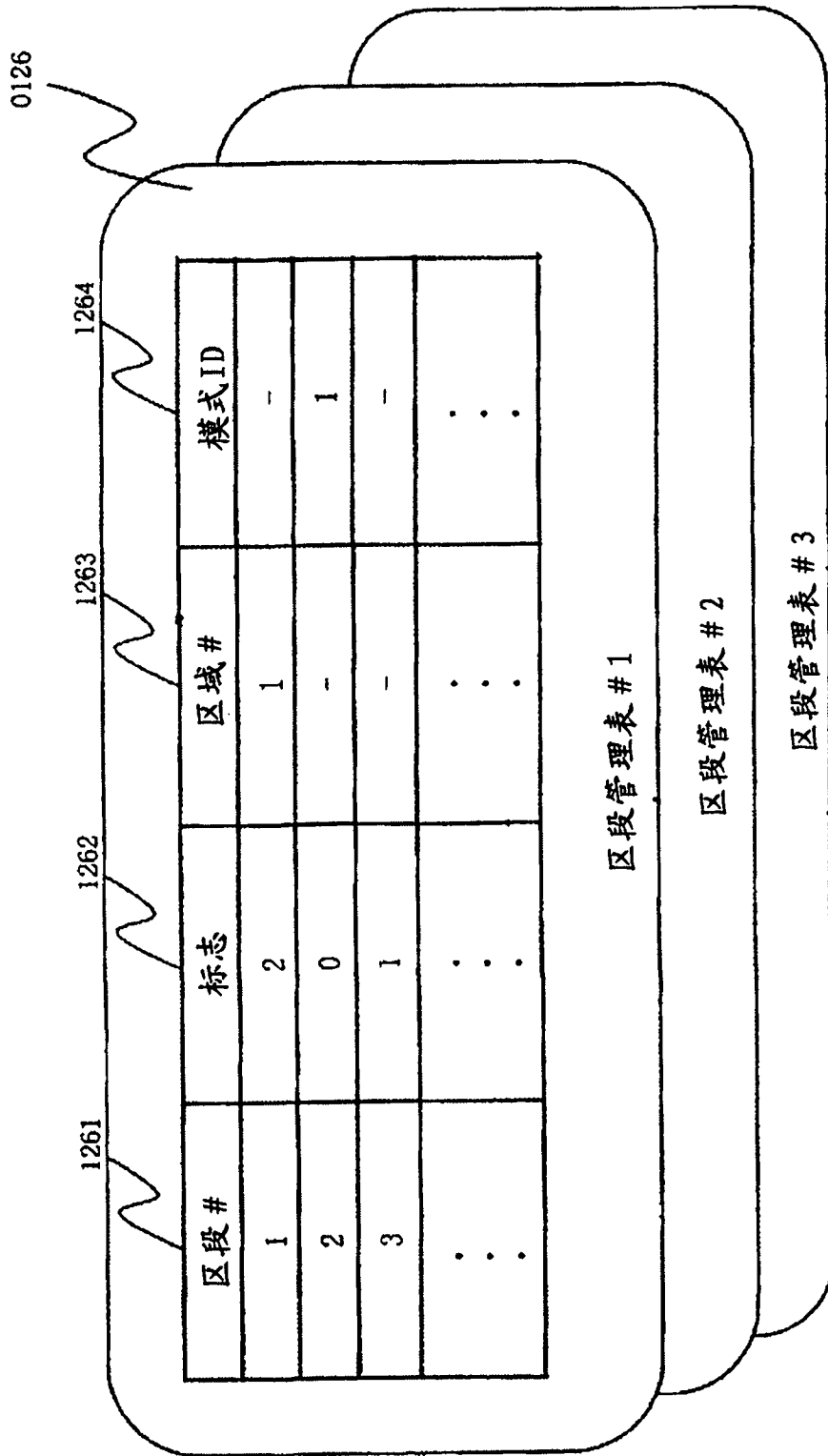


图 25

1271 1272 1273 1274 0127

阈值管理表

虚拟卷 #	阈值 α	阈值 β	阈值 γ
1	5	2	5
2	10	3	6
⋮	⋮	⋮	⋮

图 26

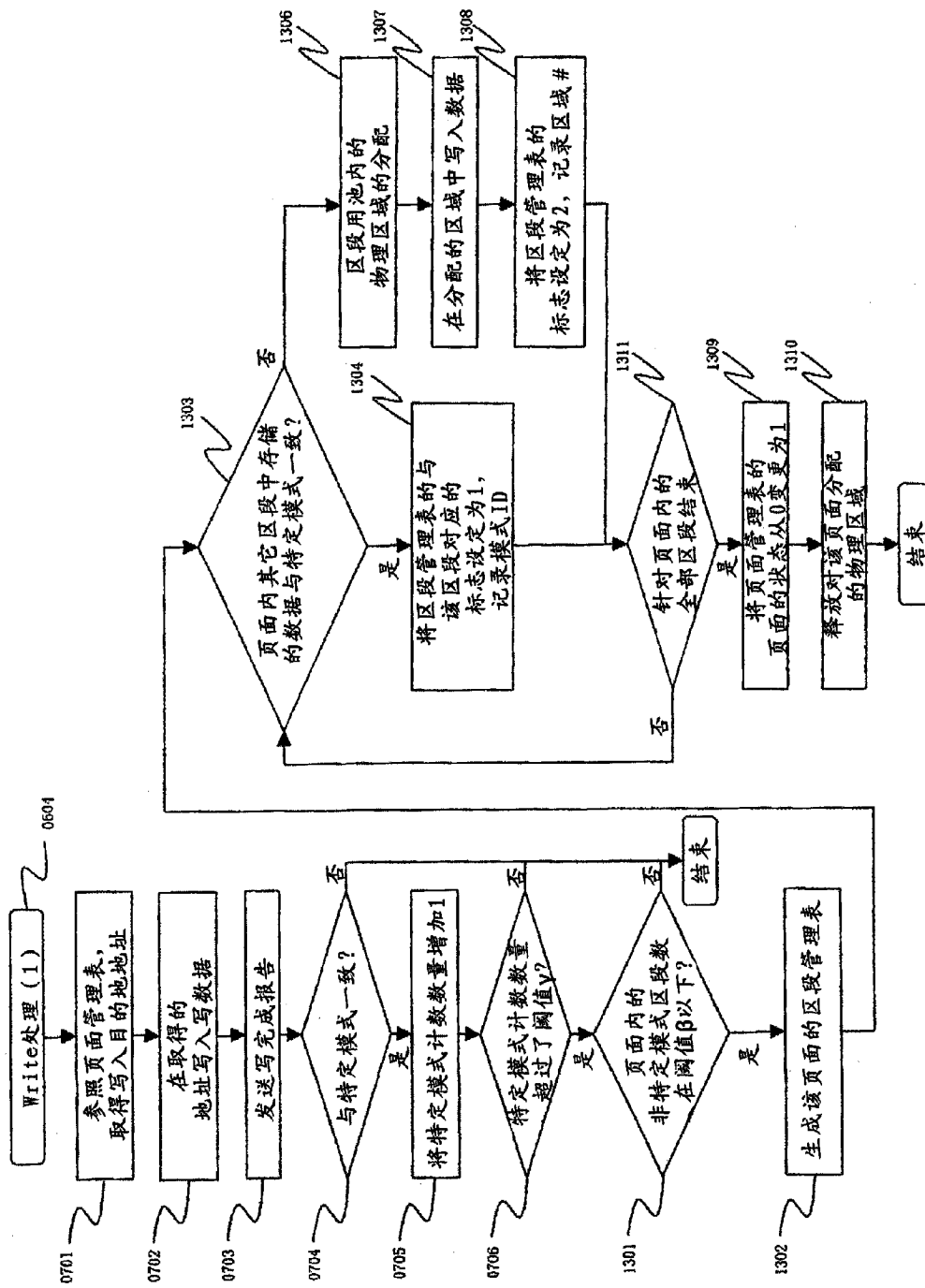


图 27

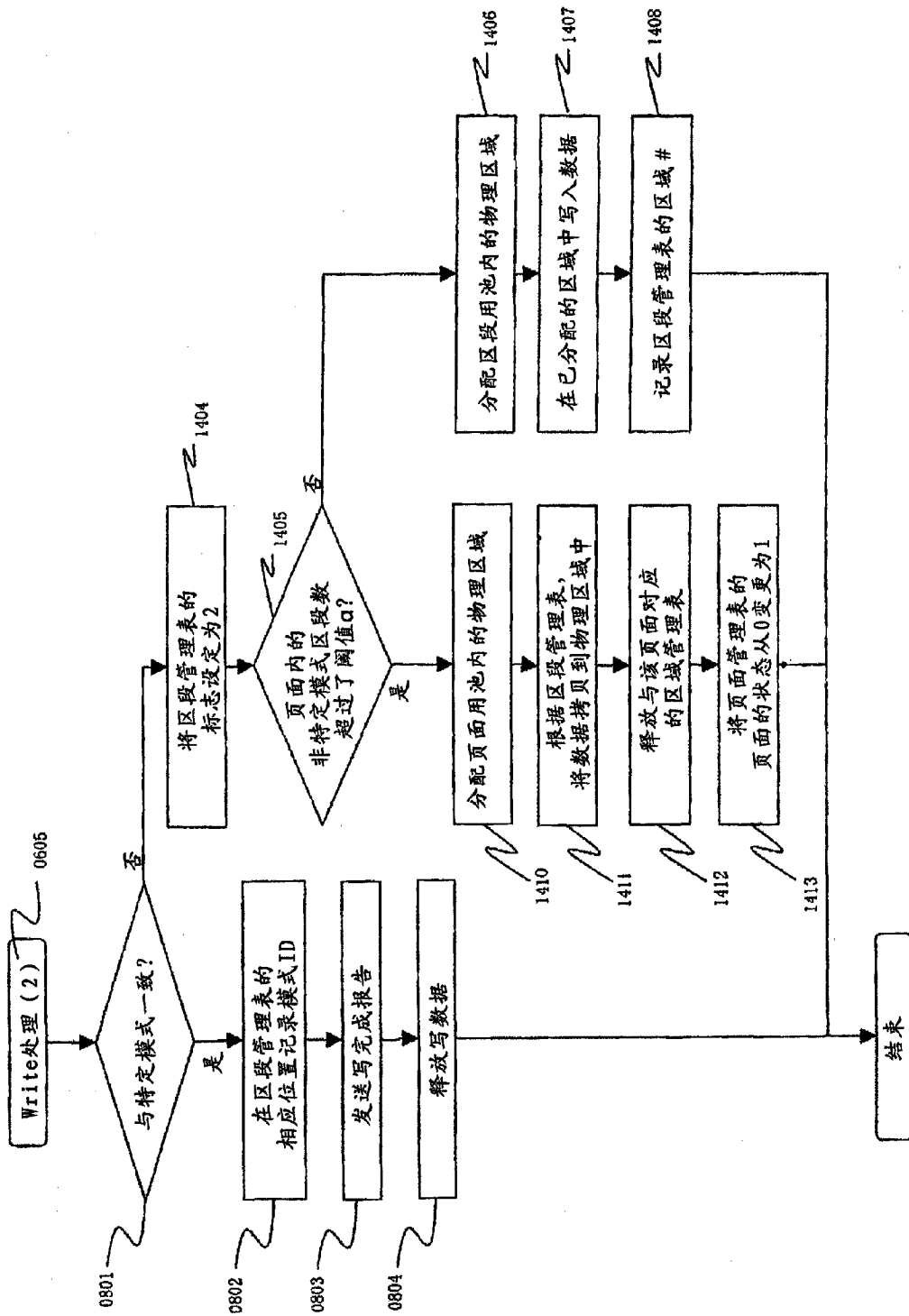


图 28