



(12) 发明专利申请

(10) 申请公布号 CN 102483684 A

(43) 申请公布日 2012. 05. 30

(21) 申请号 200980161169. 2

(51) Int. Cl.

(22) 申请日 2009. 12. 24

G06F 3/06 (2006. 01)

(85) PCT申请进入国家阶段日

2012. 02. 28

(86) PCT申请的申请数据

PCT/JP2009/007207 2009. 12. 24

(87) PCT申请的公布数据

W02011/077489 JA 2011. 06. 30

(71) 申请人 株式会社日立制作所

地址 日本东京都

(72) 发明人 折笠大典 松田健 高田丰

江口贤哲 里山爱 水野阳一

(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 曾贤伟 范胜杰

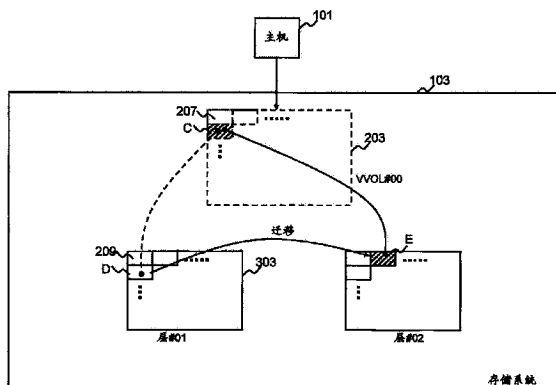
权利要求书 4 页 说明书 35 页 附图 34 页

(54) 发明名称

提供虚拟卷的存储系统

(57) 摘要

管理将由多个存储设备提供的多个存储区域依照存储区域的属性进行分类后的多个存储区域组。根据面向逻辑卷内多个地址中所含至少一个地址的写入请求,提供分配到所述多个存储区域中所含至少一个存储区域的多个所述逻辑卷。根据所述逻辑卷内所述至少一个地址的存取状况,将通过所述写入请求写到所述至少一个地址中的数据从所述多个存储区域组的一个中所含所述至少一个存储区域,向所述多个存储区域组所含其它存储区域组内的至少一个存储区域迁移。



1. 一种存储系统的控制方法,其特征在于,

管理将由多个存储设备提供的多个存储区域依照存储区域的属性进行分类后的多个存储区域组,

根据面向逻辑卷内多个地址中所含至少一个地址的写入请求,提供分配到所述多个存储区域中所含至少一个存储区域的多个所述逻辑卷,

根据所述逻辑卷内所述至少一个地址的存取状况,将通过所述写入请求写到所述至少一个地址中的数据从所述多个存储区域组的一个中所含所述至少一个存储区域,向所述多个存储区域组所含其它存储区域组内的至少一个存储区域迁移。

2. 根据权利要求 1 所述的存储系统的控制方法,其特征在于,

所述多个存储区域组中所含第一存储区域组和第二存储区域组在每个环中具有与各自的组相应的所述存取状况的程度作为属性,所述第一和所述第二存储区域组的所述存取状况的程度对于每个组不同,但是一部分重复,

在所述迁移处理中,在所述多个存储区域组的一个中所含所述至少一个存储区域包含在所述第一或者所述第二存储区域组中并且所述至少一个地址的所述存取状况位于所述存取状况的程度中所述一部分重复的范围内的情况下,不实施所述数据的迁移。

3. 一种存储系统,其与一台以上的主机连接,其特征在于,

该存储系统具备多种物理存储设备群和存储控制装置,

所述存储控制装置具有存储资源和控制器,所述控制器与所述多种物理存储设备群、所述存储资源以及所述主机连接,

物理存储设备群由一个以上的物理存储设备构成,

所述多种物理存储设备群成为多种实际页群的根基,

实际页群的种类基于成为该实际页群的根基的物理存储设备群的属性,

实际页群具有多个实际页,

所述控制器向所述一台以上的主机提供一个以上的虚拟卷,

所述虚拟卷是由多个虚拟页构成的虚拟逻辑卷,

所述虚拟页是虚拟的存储区域,

所述实际页是实体的存储区域,

所述存储资源存储分配信息、监视信息和实际页群管理信息,其中,所述分配信息是表示对哪个虚拟页分配了哪个实际页的信息,所述监视信息是表示各虚拟页的 I/O(Input/Output) 频度的信息,所述实际页群管理信息是包含表示对每种实际页群设定的 I/O 频度范围信息的信息,

所述控制器从所述主机接收写入命令,根据所述写入命令确定写入目的地虚拟页,更新所述分配信息,从而对所确定的虚拟页分配自由的实际页,在所分配的实际页中保存写入对象数据,

所述控制器根据针对虚拟页或者分配给该虚拟页的实际页的 I/O,更新所述监视信息中的、该虚拟页的 I/O 频度,

所述实际页是实体的存储区域,保存在实际页中的数据存储在作为该实际页的根基的物理存储设备,

所述控制器进行迁移处理,该迁移处理是包含以下处理 (A) 至 (C) 的处理:

(A) 确定根据所述监视信息所确定的 I/O 频度不符合预定条件的虚拟卷或者虚拟页，
(B) 将保存在分配给所确定的虚拟卷或者虚拟页的、第一种实际页群内第一实际页中的数据，迁移到第二种实际页群内的第二实际页，

(C) 更新所述分配信息，从而以所述第二实际页取代所述第一实际页向所述第一实际页分配目的地虚拟页、即第一虚拟页进行分配，

所述预定条件是指纳入所述第一种实际页群的 I/O 频度范围，

所述第二种实际页群是 I/O 频度范围纳入通过所述 (A) 所确定的虚拟卷或者虚拟页的 I/O 频度的实际页群，

所述控制器进行 I/O 频度范围的决定处理，在该决定处理中，所述控制器基于所述监视信息决定各种实际页群的 I/O 频度范围。

4. 根据权利要求 3 所述的存储系统，其特征在于，

所述控制器在所述决定处理中进行以下处理 (a) 以及 (b)：

(a) 使用所述监视信息生成表示 I/O 频度与虚拟页数的关系的信息、即关系信息，

(b) 基于所述关系信息决定各种实际页群的 I/O 频度范围。

5. 根据权利要求 3 所述的存储系统，其特征在于，

所述控制器在所述处理 (b) 中进行以下的处理 (b1) 至 (b3)：

(b1) 使用所述关系信息，计算从所述多种实际页群分配到实际页的一个以上虚拟卷内所有虚拟页的 I/O 频度的合计、即 I/O 频度合计，

(b2) 基于各种实际页群的性能值的比和所述 I/O 频度合计，针对实际页群的每个边界计算实际页群的边界的 I/O 频度、即边界值，

(b3) 基于各种实际页群的 I/O 频度和实际页群的每个边界的边界值，决定各种实际页群的 I/O 频度范围，

实际页群的所述性能值是基于作为该实际页群的根基的物理存储设备群的属性决定的、与性能相关的值，

所述实际页群的边界是指实际页群与其上位以及 / 或者下位实际页群的交界，

各种实际页群的 I/O 频度范围的最大值以及 / 或者最小值是基于该实际页群与其上位以及 / 或者下位实际页群的交界的边界值决定的值。

6. 根据权利要求 5 所述的存储系统，其特征在于，

实际页群的所述性能值是通过该实际页群中的、能够成为所述迁移处理中的迁移目的地的实际页的数量与该实际页群内每一个实际页的性能值的乘法运算计算出的值，

实际页群内每一个实际页的性能值是基于作为该实际页群的根基的物理存储设备群的属性决定的、与性能相关的值。

7. 根据权利要求 3 所述的存储系统，其特征在于，

所述 I/O 频度范围的最大值大于与该 I/O 频度范围对应的较大一方的边界值，以及 / 或者，该 I/O 频度范围的最小值小于与该 I/O 频度范围对应的较小一方的边界值。

8. 根据权利要求 3 所述的存储系统，其特征在于，

所述控制器在所述决定处理中进行以下处理：

(a) 使用所述监视信息，计算从所述多种实际页群分配到实际页的一个以上虚拟卷内所有虚拟页的 I/O 频度的合计、即 I/O 频度合计，

(b) 基于各种实际页群的性能值的比和所述 I/O 频度合计, 针对实际页群的每个边界计算实际页群的边界的 I/O 频度、即边界值,

(c) 基于各种实际页群的 I/O 频度和实际页群的每个边界的边界值, 决定各种实际页群的 I/O 频度范围,

实际页群的所述性能值是基于作为该实际页群的根基的物理存储设备群的属性决定的、与性能相关的值,

所述实际页群的边界是指实际页群与其上位以及 / 或者下位实际页群的交界,

各种实际页群的 I/O 频度范围的最大值以及 / 或者最小值是基于该实际页群与其上位以及 / 或者下位实际页群的交界的边界值决定的值。

9. 根据权利要求 3 所述的存储系统, 其特征在于,

所述控制器在所述决定处理中进行以下处理 (d) 以及 (e) :

(d) 假定分配给 I/O 频度纳入实际页群的 I/O 频度范围的所有虚拟页的实际页内的数据要素被迁移到了该实际页群中, 判定该实际页群的使用比例是否超出该实际页群的比例阈值,

(e) 在所述 (d) 中的判断的结果为否定性的情况下, 将所决定的 I/O 频度范围包含在所述实际页群管理信息中,

实际页群的所述使用比例是指分配给虚拟页的实际页的数量相对于构成该实际页群的实际页的数量的比例,

所述实际页群的比例阈值是指所述使用比例的阈值。

10. 根据权利要求 9 所述的存储系统, 其特征在于,

所述控制器在所述决定处理中还进行以下处理 (f) :

(f) 在所述 (d) 中的判断的结果为肯定性的情况下, 将 I/O 频度范围修正为实际页群的使用比例不超出该实际页群的比例阈值的大小。

11. 根据权利要求 10 所述的存储系统, 其特征在于,

所述控制器在所述决定处理中还进行以下处理 (g) :

(g) 将修正后的 I/O 频度范围包含在所述实际页群管理信息中。

12. 根据权利要求 3 所述的存储系统, 其特征在于,

根据所述监视信息所确定的、虚拟卷或者虚拟页的 I/O 频度是指该虚拟卷或者虚拟页的、包含在所述监视信息中的 I/O 频度。

13. 根据权利要求 3 所述的存储系统, 其特征在于,

根据所述监视信息所确定的、虚拟卷或者虚拟页的 I/O 频度是指所述监视信息中的、与该虚拟卷之外其它虚拟卷的 I/O 频度、或者该虚拟页以外其它虚拟页的 I/O 频度的相对值。

14. 根据权利要求 3 所述的存储系统, 其特征在于,

所述控制器接受使所述 (A) 中的 I/O 频度为绝对值还是相对值的指示,

在接受到使所述 (A) 中的 I/O 频度为绝对值的指定的情况下, 根据所述监视信息所确定的、虚拟卷或者虚拟页的 I/O 频度是指该虚拟卷或者虚拟页的、包含在所述监视信息中的 I/O 频度,

在接受到使所述 (A) 中的 I/O 频度为相对值的指定的情况下, 根据所述监视信息所确

定的、虚拟卷或者虚拟页的 I/O 频度是指所述监视信息中的、与该虚拟卷之外其它虚拟卷的 I/O 频度、或者该虚拟页以外其它虚拟页的 I/O 频度的相对值。

15. 根据权利要求 3 所述的存储系统,其特征在于,
所述 (A) 中的 I/O 频度是某时间的 I/O 数的平均。

提供虚拟卷的存储系统

技术领域

[0001] 本发明涉及应用 Thin Provisioning(自动精简配置,也被称作 Dynamic Provisioning(动态预配置))的存储控制。

背景技术

[0002] 数据迁移技术已为人所知。数据迁移通常以卷(volume)为单位进行。例如,根据专利文献 1,保存在第一逻辑卷中的所有数据从第一逻辑卷迁移到第二逻辑卷。

[0003] 另外,存储层次化技术也已为人所知。例如,根据专利文献 2,多个存储系统所具有的多个逻辑卷分别被分类至多个层(tier)中的某一个层,某一个层中所属的逻辑卷被迁移到其它的层。多个层中包括例如高可信度的层、低成本的层等。

[0004] 现有技术文献

[0005] 专利文献

[0006] 专利文献 1:日本特开 2006-302077 号公报

[0007] 专利文献 2:日本特开 2008-047156 号公报

发明内容

[0008] 发明要解决的课题

[0009] 但是,被称作 Thin Provisioning(也被称作 Dynamic Provisioning)的技术已为人所知。根据 Thin Provisioning,准备由多个虚拟存储区域(虚拟页)构成的虚拟逻辑卷(虚拟卷)和多个实体的存储区域(实际页)。在对虚拟页进行写入的情况下,对该虚拟页分配实际页,在所分配的实际页中写入作为写入对象的数据。

[0010] 设 Thin Provisioning 中应用了上述的数据迁移技术以及存储层次化技术。另外,设第一虚拟卷属于第一层,第二虚拟卷属于第二层。该情况下,能够从第一层中所属的第一虚拟卷向第二层中所属的第二虚拟卷迁移所有的数据。

[0011] 根据该技术,第一虚拟卷内的所有数据被从第一层迁移到第二层,但是并不一定优选所有的数据存在于第二层。

[0012] 于是,本发明的目的在于使依照 Thin Provisioning 的虚拟卷内的数据配置适当化。

[0013] 用于解决课题的手段

[0014] 管理将由多个存储设备提供的多个存储区域依照存储区域的属性进行分类后的多个存储区域组。根据面向逻辑卷内多个地址中所含至少一个地址的写入请求,提供分配到所述多个存储区域中所含至少一个存储区域的多个所述逻辑卷。根据所述逻辑卷内所述至少一个地址的存取状况,将通过所述写入请求写到所述至少一个地址中的数据从所述多个存储区域组的一个中所含所述至少一个存储区域,向所述多个存储区域组所含其它存储区域组内的至少一个存储区域迁移。

附图说明

- [0015] 图 1 表示本发明实施例 1 的概要。
- [0016] 图 2 表示存储系统 103 的构成。
- [0017] 图 3 表示存储系统 103 中各种存储区域的关系。
- [0018] 图 4 表示 CMPK119 所存储的信息以及计算机程序。
- [0019] 图 5 表示层定义表 401 的构成。
- [0020] 图 6 表示层管理表 402 的构成。
- [0021] 图 7 表示池管理表 403 的构成。
- [0022] 图 8 表示 VVOL 管理表 404 的构成。
- [0023] 图 9 表示 PVOL 管理表 405 的构成。
- [0024] 图 10 表示分配表 407 的构成。
- [0025] 图 11 表示实际页管理表 408 的构成。
- [0026] 图 12 表示监视表 409 的构成。
- [0027] 图 13 表示度数分布表 410 的构成。
- [0028] 图 14 表示迁移页表 411 的构成。
- [0029] 图 15 表示迁移定义表 412 的构成。
- [0030] 图 16 表示池生成处理的流程图。
- [0031] 图 17 表示池定义 GUI1700。
- [0032] 图 18 表示 VVOL 生成处理的流程。
- [0033] 图 19 表示 VVOL 生成 GUI1900。
- [0034] 图 20 表示本发明的实施例 5 所涉及的 VVOL 管理表 4404。
- [0035] 图 21 表示数据要素的汇总处理的概要。
- [0036] 图 22 表示信息显示处理的流程。
- [0037] 图 23 表示度数分布 GUI2300。
- [0038] 图 24 表示写入处理（向高速缓存的写入）的流程。
- [0039] 图 25 表示降级处理的流程。
- [0040] 图 26 表示读出处理的流程。
- [0041] 图 27 表示 I/O 频度的监视与重新配置的关系的概要。
- [0042] 图 28 表示重新配置处理的流程。
- [0043] 图 29 表示层范围决定处理的概要。
- [0044] 图 30 表示层范围决定处理的流程。
- [0045] 图 31 表示图 30 的 S4402 中层边界值计算处理的流程。
- [0046] 图 32 表示本发明的实施例 5 所涉及的池管理表 4403。
- [0047] 图 33 表示本发明的实施例 5 所涉及的比较方法 / 比较对象的设定处理的流程。
- [0048] 图 34 表示图 28 的 S2808 中迁移页表生成处理的流程。
- [0049] 图 35 表示页单位的迁移处理的概要。
- [0050] 图 36 表示页单位的迁移处理的流程。
- [0051] 图 37 表示层阈值的变更处理的流程。
- [0052] 图 38 表示阈值超出应对处理的流程。

具体实施方式

[0053] 以下,叙述基于本发明的后述多个实施例的观点。

[0054] <观点 1>

[0055] 一种存储系统的控制方法,其特征在于,管理将由多个存储设备提供的多个存储区域依照存储区域的属性进行分类后的多个存储区域组,根据面向逻辑卷内多个地址中所含至少一个地址的写入请求,提供分配到所述多个存储区域中所含至少一个存储区域的多个所述逻辑卷,根据所述逻辑卷内所述至少一个地址的存取状况,将通过所述写入请求写到所述至少一个地址中的数据从所述多个存储区域组的一个中所含所述至少一个存储区域,向所述多个存储区域组所含其它存储区域组内的至少一个存储区域迁移。

[0056] <观点 2>

[0057] 一种依照观点 1 的存储系统的控制方法,其特征在于,所述多个存储区域组中所含第一存储区域组和第二存储区域组在每个环中具有与各自的组相应的所述存取状况的程度作为属性,所述第一和所述第二存储区域组的所述存取状况的程度对于每个组不同,但是一部分重复,在所述迁移处理中,在所述多个存储区域组的一个中所含所述至少一个存储区域包含在所述第一或者所述第二存储区域组中并且所述至少一个地址的所述存取状况位于所述存取状况的程度中所述一部分重复的范围内,不实施所述数据的迁移。

[0058] <观点 3>

[0059] 一种与一台以上的主机连接的存储系统,其具备多种物理存储设备群和存储控制装置。所述存储控制装置具有存储资源和控制器,所述控制器与所述多种物理存储设备群、所述存储资源以及所述主机连接。物理存储设备群由一个以上的物理存储设备构成。所述多种物理存储设备群成为多种实际页群的根基。实际页群的种类基于成为该实际页群的根基的物理存储设备群的属性。实际页群具有多个实际页。所述控制器向所述一台以上的主机提供一个以上的虚拟卷。所述虚拟卷是由多个虚拟页构成的虚拟逻辑卷。所述虚拟页是虚拟的存储区域。所述实际页是实体的存储区域。所述存储资源存储分配信息、监视信息和实际页群管理信息,其中,所述分配信息是表示对哪个虚拟页分配了哪个实际页的信息,所述监视信息是表示各虚拟页的 I/O(Input/Output) 频度的信息,所述实际页群管理信息是包含表示对每种实际页群设定的 I/O 频度范围信息的信息(可以为表形式或者队列形式等,信息的形式任意)。所述控制器从所述主机接收写入命令,根据所述写入命令确定写入目的地虚拟页,更新所述分配信息,从而对所确定的虚拟页分配自由的实际页,在所分配的实际页中保存写入对象数据。所述控制器根据针对虚拟页或者分配给该虚拟页的实际页的 I/O,更新所述监视信息中的、该虚拟页的 I/O 频度。所述实际页是实体的存储区域,保存在实际页中的数据存储到作为该实际页的根基的物理存储设备。所述控制器进行迁移处理,该迁移处理是包含以下处理 (A) 至 (C) 的处理:(A) 确定根据所述监视信息所确定的 I/O 频度不符合预定条件的虚拟卷或者虚拟页,(B) 将保存在分配给所确定的虚拟卷或者虚拟页的、第一种实际页群内第一实际页中的数据,迁移到第二种实际页群内的第二实际页,(C) 更新所述分配信息,从而以所述第二实际页取代所述第一实际页向所述第一实际页分配目的地虚拟页、即第一虚拟页进行分配。所述预定条件是指纳入所述第一种实际页群的 I/O

频度范围。所述第二种实际页群是 I/O 频度范围纳入通过所述 (A) 所确定的虚拟卷或者虚拟页的 I/O 频度的实际页群。所述控制器进行 I/O 频度范围的决定处理,在该决定处理中,所述控制器基于所述监视信息决定各种实际页群的 I/O 频度范围。控制器例如可以具有至少一个处理器。存储资源例如可以为存储器。主机可以为计算机,也可以为其它的存储系统或者其它的存储控制装置。

[0060] 由此,能够将虚拟卷内的数据根据虚拟页的 I/O 频度以页单位重新配置。届时,能够将数据配置在与纳入有分配有保存该数据的实际页的虚拟页的 I/O 频度的 I/O 频度范围对应的种类的实际页群中。

[0061] < 观点 4>

[0062] 一种依照观点 3 的存储系统,所述控制器在所述决定处理中进行以下处理 (a) 以及 (b) :

[0063] (a) 使用所述监视信息生成表示 I/O 频度与虚拟页数的关系的信息、即关系信息,

[0064] (b) 基于所述关系信息决定各种实际页群的 I/O 频度范围。

[0065] < 观点 5>

[0066] 一种依照观点 4 的存储系统,所述控制器在所述处理 (b) 中进行以下的处理 (b1) 至 (b3) :

[0067] (b1) 使用所述关系信息,计算从所述多种实际页群分配到实际页的一个以上虚拟卷内所有虚拟页的 I/O 频度的合计、即 I/O 频度合计,

[0068] (b2) 基于各种实际页群的性能值的比和所述 I/O 频度合计,针对实际页群的每个边界计算实际页群的边界的 I/O 频度、即边界值,

[0069] (b3) 基于各种实际页群的 I/O 频度和实际页群的每个边界的边界值,决定各种实际页群的 I/O 频度范围。

[0070] 实际页群的所述性能值是基于作为该实际页群的根基的物理存储设备群的属性决定的、与性能相关的值。所述实际页群的边界是指实际页群与其上位以及 / 或者下位实际页群的交界。各种实际页群的 I/O 频度范围的最大值以及 / 或者最小值是基于该实际页群与其上位以及 / 或者下位实际页群的交界的边界值决定的值。

[0071] < 观点 6>

[0072] 一种依照观点 3 至观点 5 的任一观点的存储系统,所述 I/O 频度范围的最大值大于与该 I/O 频度范围对应的较大一方的边界值,以及 / 或者,该 I/O 频度范围的最小值小于与该 I/O 频度范围对应的较小一方的边界值。

[0073] < 观点 7>

[0074] 一种依照观点 3 或者观点 4 的存储系统,实际页群的所述性能值是通过该实际页群中的、能够成为所述迁移处理中的迁移目的地的实际页的数量与该实际页群内每一个实际页的性能值的乘法运算计算出的值。实际页群内每一个实际页的性能值是基于作为该实际页群的根基的物理存储设备群的属性决定的、与性能相关的值。

[0075] < 观点 8>

[0076] 一种依照观点 3 至观点 7 的任一观点的存储系统,所述控制器在所述决定处理中进行以下处理 :

[0077] (a) 使用所述监视信息,计算从所述多种实际页群分配到实际页的一个以上虚拟

卷内所有虚拟页的 I/O 频度的合计、即 I/O 频度合计，

[0078] (b) 基于各种实际页群的性能值的比和所述 I/O 频度合计，针对实际页群的每个边界计算实际页群的边界的 I/O 频度、即边界值，

[0079] (c) 基于各种实际页群的 I/O 频度和实际页群的每个边界的边界值，决定各种实际页群的 I/O 频度范围。

[0080] 实际页群的所述性能值是基于作为该实际页群的根基的物理存储设备群的属性决定的、与性能相关的值。所述实际页群的边界是指实际页群与其上位以及 / 或者下位实际页群的交界。各种实际页群的 I/O 频度范围的最大值以及 / 或者最小值是基于该实际页群与其上位以及 / 或者下位实际页群的交界的边界值决定的值。

[0081] < 观点 9 >

[0082] 一种依照观点 3 至观点 8 的任一观点的存储系统，所述控制器在所述决定处理中进行以下处理 (d) 以及 (e)：

[0083] (d) 假定分配给 I/O 频度纳入实际页群的 I/O 频度范围的所有虚拟页的实际页内的数据要素被迁移到了该实际页群中，判定该实际页群的使用比例是否超出该实际页群的比例阈值，

[0084] (e) 在所述 (d) 中的判断的结果为否定性的情况下，将所决定的 I/O 频度范围包含在所述实际页群管理信息中。

[0085] 实际页群的所述使用比例是指分配给虚拟页的实际页的数量相对于构成该实际页群的实际页的数量的比例。所述实际页群的比例阈值是指所述使用比例的阈值。

[0086] 由此，能够判断 I/O 频度范围是否合适。

[0087] < 观点 10 >

[0088] 一种依照观点 9 的存储系统，所述控制器在所述决定处理中还进行以下处理 (f)：

[0089] (f) 在所述 (d) 中的判断的结果为肯定性的情况下，将 I/O 频度范围修正为实际页群的使用比例不超出该实际页群的比例阈值的大小。

[0090] < 观点 11 >

[0091] 一种依照观点 10 的存储系统，所述控制器在所述决定处理中还进行以下处理 (g)：

[0092] (g) 将修正后的 I/O 频度范围包含在所述实际页群管理信息中。

[0093] < 观点 12 >

[0094] 一种依照观点 3 至观点 11 的任一观点的存储系统，根据所述监视信息所确定的、虚拟卷或者虚拟页的 I/O 频度是指该虚拟卷或者虚拟页的、包含在所述监视信息中的 I/O 频度。

[0095] < 观点 13 >

[0096] 一种依照观点 3 至观点 12 的任一观点的存储系统，根据所述监视信息所确定的、虚拟卷或者虚拟页的 I/O 频度是指所述监视信息中的、与该虚拟卷之外其它虚拟卷的 I/O 频度、或者该虚拟页以外其它虚拟页的 I/O 频度的相对值。

[0097] < 观点 14 >

[0098] 一种依照观点 3 至观点 12 的任一观点的存储系统，所述控制器接受使所述 (A) 中的 I/O 频度为绝对值还是相对值的指示。在接受到使所述 (A) 中的 I/O 频度为绝对值的指

定的情况下,根据所述监视信息所确定的、虚拟卷或者虚拟页的 I/O 频度是指该虚拟卷或者虚拟页的、包含在所述监视信息中的 I/O 频度。在接受到使所述 (A) 中的 I/O 频度为相对值的指定的情况下,根据所述监视信息所确定的、虚拟卷或者虚拟页的 I/O 频度是指所述监视信息中的、与该虚拟卷之外其它虚拟卷的 I/O 频度、或者该虚拟页以外其它虚拟页的 I/O 频度的相对值。

[0099] < 观点 15>

[0100] 一种依照观点 3 至观点 14 的任一观点的存储系统,所述 (A) 中的 I/O 频度是某时间的 I/O 数的平均。

[0101] < 观点 16>

[0102] 一种依照观点 3 至观点 14 的任一观点的存储系统,所述 (A) 中的 I/O 频度是与某时间带的多个时间分别对应的多个 I/O 数中最大的 I/O 数。

[0103] < 观点 17>

[0104] 一种依照观点 3 至观点 14 的任一观点的存储系统,所述控制器接受使所述 (A) 中的 I/O 频度为平均 I/O 数还是最大 I/O 数的指定。在接受到使所述 (A) 中的 I/O 频度为平均 I/O 数的指定的情况下,所述 (A) 中的 I/O 频度是某时间的 I/O 数的平均。在接受到使所述 (A) 中的 I/O 频度为最大 I/O 数的指定的情况下,所述 (A) 中的 I/O 频度是与某时间带的多个时间分别对应的多个 I/O 数中最大的 I/O 数。

[0105] < 观点 18>

[0106] 一种依照观点 3 至观点 17 的任一观点的存储系统,所述控制器从所述主机接收 I/O 命令,根据该 I/O 命令确定 I/O 目的地虚拟页。所述控制器在针对分配给所述 I/O 目的地虚拟页的实际页进行 I/O 的情况下更新所述监视信息中的、与所述 I/O 目的地虚拟页对应的 I/O 频度,而在不针对分配给所述 I/O 目的地虚拟页的实际页进行 I/O 的情况下不更新所述监视信息中的、与所述 I/O 目的地虚拟页对应的 I/O 频度。

[0107] 由此,能够仅在针对实际页进行了 I/O 的情况下更新分配有该实际页的虚拟页的 I/O 频度。

[0108] < 观点 19>

[0109] 一种依照观点 18 的存储系统,所述存储资源存储状态为激活的监视信息和状态为待机的监视信息作为所述监视信息。所述控制器更新所述激活的监视信息中的 I/O 频度。所述控制器在所述迁移处理中基于所述激活的监视信息确定所述第一虚拟页,并且将所述待机的监视信息的状态从待机变更为激活,然后,在依照确定虚拟页的 I/O 命令对实际页进行 I/O 的情况下,所述控制器更新状态被变更为激活的监视信息中的、该被确定的虚拟页的 I/O 频度。

[0110] 由此,即便进行上述迁移处理时也能够继续 I/O 频度的监视(更新)。

[0111] < 观点 20>

[0112] 一种依照观点 4 的存储系统,所述控制器输出所述关系信息,使得所述生成的关系信息得以显示。所述控制器在接受到开始迁移的指示时开始所述迁移处理。

[0113] < 观点 21>

[0114] 一种依照观点 3 至观点 20 中任一观点的存储系统,所述控制器在所述第一实际数据群的使用比例超出预定的比例阈值的情况下开始所述迁移处理。

[0115] 因此,在所述第一实际数据群的使用比例超出了预定比例阈值的情况下,通过所述迁移处理能够期待在第一实际数据群中自由的实际页有所增加。

[0116] <观点 22>

[0117] 一种依照观点 3 至观点 21 中任一观点的存储系统,所述第一种实际数据群是最上位的实际数据群。所述控制器必定从所述第一种实际数据群对根据所述写入命令所确定的写入目的地虚拟页分配实际页。

[0118] <观点 23>

[0119] 一种依照观点 3 至观点 22 中任一观点的存储系统,在分配给所述虚拟卷的所有实际页中预定比例以上的实际页存在于某种实际页群的情况下,所述控制器将分配给所述虚拟卷的所有实际页中的、所述某种实际页群以外的种类的实际页群内的实际页内的数据迁移到所述某种实际页群内的实际页。

[0120] 由此,能够根据虚拟卷的数据的配置状况,在考虑为适合该虚拟卷的种类的实际页群中汇总该虚拟卷内的数据。

[0121] <观点 24>

[0122] 一种依照观点 23 的存储系统,迁移到所述某种实际页群内的实际页后分配给所述虚拟页的实际页总是所述某种实际页群内的实际页。

[0123] 由此,在通过迁移处理汇总数据之后,从所述某种实际页群对所述第二虚拟卷分配实际页。

[0124] <观点 25>

[0125] 一种依照观点 3 至观点 24 的任一观点的存储系统,所述控制器在所述页单位迁移处理开始之后从所述主机接收确定所述第一虚拟页的写入命令,并且在没有开始从所述第一实际页向所述第二实际页的数据迁移的情况下将该迁移取消,向所述第二种实际页群内的实际页写入依照所述写入命令的数据,并且更新所述分配信息,从而以该写入目的地实际页取代所述第一实际页分配给所述第一虚拟页。

[0126] 由于数据在迁移之前写入迁移目的地实际页群,因此针对该数据的写入目的地虚拟页,可以不进行迁移。

[0127] <观点 26>

[0128] 一种依照观点 3 至观点 25 的任一观点的存储系统,所述控制器在所述页单位迁移处理开始之后从所述主机接收确定所述第一虚拟页的写入命令,并且在没有开始从所述第一实际页向所述第二实际页的数据迁移的情况下,向所述第一实际页写入依照所述写入命令的数据,然后,进行所述 (A) 以及 (B)。

[0129] <观点 27>

[0130] 一种依照观点 3 至观点 26 的任一观点的存储系统,所述控制器在所述页单位迁移处理开始之后从所述主机接收确定所述第一虚拟页的写入命令,并且在正从所述第一实际页向所述第二实际页迁移数据的情况下将该迁移取消,向所述第二种实际页群内的实际页写入依照所述写入命令的数据,并且更新所述分配信息,从而以该写入目的地实际页取代所述第一实际页分配给所述第一虚拟页。

[0131] <观点 28>

[0132] 一种依照观点 3 至观点 26 的任一观点的存储系统,所述控制器在所述页单位迁

移处理开始之后从所述主机接收确定所述第一虚拟页的写入命令,并且在正从所述第一实际页向所述第二实际页迁移数据的情况下,向所述第一实际页写入依照所述写入命令的数据,然后,重新开始从所述第一实际页向所述第二实际页群内的实际页的数据迁移。

[0133] < 观点 29>

[0134] 一种依照观点 3 至观点 26 的任一观点的存储系统,所述控制器在所述页单位迁移处理开始之后从所述主机接收确定所述第一虚拟页的写入命令,并且在正从所述第一实际页向所述第二实际页迁移数据的情况下,在该迁移结束后,依照所述分配信息向所述第二实际页写入依照所述写入命令的数据。

[0135] < 观点 30>

[0136] 一种依照观点 3 至观点 26 的任一观点的存储系统,所述控制器在所述页单位迁移处理开始之后从所述主机接收确定所述第一虚拟页的写入命令,并且在正从所述第一实际页向所述第二实际页迁移数据的情况下,向所述第一实际页和所述第二实际页双方写入依照所述写入命令的数据。

[0137] < 观点 31>

[0138] 一种依照观点 3 至观点 30 的任一观点的存储系统,所述控制器在所述页单位迁移处理开始之后从所述主机接收确定所述第一虚拟页的写入命令,并且在所述 (A) 以及 (B) 结束的情况下,依照所述分配信息,向所述第二实际页写入依照所述写入命令的数据。

[0139] < 观点 32>

[0140] 一种依照观点 3 至观点 31 的任一观点的存储系统,所述第一虚拟页是进行了 I/O 的最新时刻、即最终 I/O 时刻满足预定条件的虚拟页。

[0141] < 观点 33>

[0142] 一种依照观点 32 的存储系统,所述预定条件是指最终 I/O 时刻是当前时刻之前过去预定时间以上的时间。所述第二实际页群是比所述第一实际页群更下位的实际页群。

[0143] < 观点 34>

[0144] 一种依照观点 3 至观点 33 的任一观点的存储系统,所述第一虚拟页是发生依照顺序存取的 I/O 的虚拟页。所述第二实际页群是比所述第一实际页群更下位的实际页群。

[0145] < 观点 35>

[0146] 一种依照观点 3 至观点 34 的任一观点的存储系统,所述控制器仅在预定的时间带进行所述监视信息的更新。所述控制器在每经过预定时间带时,保存所述监视信息。所述第一虚拟页是根据与被指定的时间带对应的监视信息所确定的 I/O 频度不适合预定条件的虚拟页。

[0147] < 观点 36>

[0148] 一种与主机连接的存储系统的存储控制装置,其具有存储资源和控制器,所述控制器与所述存储资源以及所述主机连接。所述控制器向所述一台以上的主机提供一个以上的虚拟卷。所述虚拟卷是由多个虚拟页构成的虚拟逻辑卷。所述虚拟页是虚拟的存储区域。所述实际页是实体的存储区域。所述存储资源存储分配信息、监视信息和实际页群管理信息,其中,所述分配信息是表示对哪个虚拟页分配了哪个实际页的信息,所述监视信息是表

示各虚拟页的 I/O(Input/Output) 频度的信息,所述实际页群管理信息是包含表示对每种实际页群设定的 I/O 频度范围信息的信息。所述控制器从所述主机接收写入命令,根据所述写入命令确定写入目的地虚拟页,更新所述分配信息,从而对所确定的虚拟页从多种实际页群中的某一个分配自由的实际页,在所分配的实际页中保存写入对象数据。所述控制器根据针对虚拟页或者分配给该虚拟页的实际页的 I/O,更新所述监视信息中的、该虚拟页的 I/O 频度。所述实际页是实体的存储区域。实际页群具有多个实际页。所述控制器进行迁移处理,该迁移处理是包含以下处理 (A) 至 (C) 的处理:(A) 确定根据所述监视信息所确定的 I/O 频度不符合预定条件的虚拟卷或者虚拟页,(B) 将保存在分配给所确定的虚拟卷或者虚拟页的、第一种实际页群内第一实际页中的数据,迁移到第二种实际页群内的第二实际页,(C) 更新所述分配信息,从而以所述第二实际页取代所述第一实际页向所述第一实际页分配目的地虚拟页、即第一虚拟页进行分配。所述预定条件是指纳入所述第一种实际页群的 I/O 频度范围。所述第二种实际页群是 I/O 频度范围纳入通过所述 (A) 所确定的虚拟卷或者虚拟页的 I/O 频度的实际页群。所述控制器进行 I/O 频度范围的决定处理,在该决定处理中,所述控制器基于所述监视信息决定各种实际页群的 I/O 频度范围。

[0149] 上述的观点 1 以及 2 是以存储系统的控制方法为例的观点,观点 3 至 35 是以存储系统为例的观点,观点 36 是以存储控制装置为例的观点,它们之中的任何一种观点均能够成为存储系统、存储控制装置、存储控制方法、存储系统或者存储控制装置的控制方法、存储系统或者存储控制装置、计算机程序(例如用于存储控制的计算机程序或者在存储系统或者存储控制装置中执行的计算机程序)中的至少一个的观点。

[0150] 以下,参照附图说明本发明多个实施例中涉及的存储系统。需要说明的是,在以下的说明中,存在由名称和 ID 的组合取代参照符号来表示对象的情况。例如,在逻辑卷由参照符号 500 表示的情况下,表述为“逻辑卷 500”,在表示 ID 为 00 的逻辑卷的情况下,表述为“逻辑卷 #00”。另外,下文中将保存在一个实际页中的数据称为“数据要素”。另外,存在将保存在分配给虚拟页的实际页中的数据要素以“虚拟页内的数据要素”或者与其类似的表达来表示的情况。

[0151] 实施例 1

[0152] 图 1 表示本发明实施例 1 的概要。

[0153] 存储系统 103 上连接有主机装置(以下称作主机)101。主机 101 可以为一台或者多台。存储系统 103 向主机 101 提供虚拟卷(以下称作 VVOL(Virtual VOLUME))203,具有多种层(tier)303。

[0154] VVOL203 是依照 Thin Provisioning 的虚拟逻辑卷,即不以物理存储设备(以下称作 PDEV)为根基的逻辑卷。VVOL203 由多个虚拟页 207 构成。虚拟页 207 是虚拟的存储区域。设一个 VVOL#00 作为 VVOL203。下文中,将 VVOL#a 内的虚拟页 #b 表述为“虚拟页 #(a-b)”。Thin Provisioning 的 VVOL203 作为具有虚拟容量的逻辑卷被提供给主机 101,根据从主机 101 向具有虚拟页的地址的写入请求被分配了实际页。因此,除去为了满足虚拟容量而被分配了实际页的状态,分配给某一 VVOL203 的所有实际页的总容量小于虚拟容量。一个 VVOL203 被提供给一台以上的主机 101,在提供给多台主机 101 的情况下,由这些多台主机 101 共享。

[0155] 层 303 由多个实际页 209 构成。实际页 209 是实体的存储区域。例如,设层 #01

和 #02 这两个层作为层 303。下文中,将层 #c 内的实际页 #d 表述为“实际页 #(c-d)”。而且,层 303 例如可由一个以上的实际卷(以下称作 RVOL(Real VOLUME))构成。RVOL 是实体的逻辑卷,即以 PDEV 为根基的逻辑卷。一个池内的多个层 303 的每一个层在数据迁移前的状态下被设定为被一个或者多个 VVOL203 利用。

[0156] 主机 101 通常是计算机,也可由其它的存储系统来取代计算机。主机 101 向存储系统 103 发送例如 I/O(Input/Output) 命令。I/O 命令例如为写入命令或者读出命令,具有 I/O 目的地信息。I/O 目的地信息是表示 I/O 目的地的信息,例如包括 VVOL203 的 ID(例如 LUN(Logical Unit Number:逻辑单元号))和 I/O 目的地的地址(例如 LBA(Logical Block Address:逻辑块地址))。根据 I/O 目的地信息能够确定 I/O 目的地 VVOL203 以及虚拟页 207。

[0157] 设存储系统 103 从主机 101 接收写入命令,并基于该写入命令所具有的 I/O 目的地信息将虚拟页 #(00-C) 确定为写入目的地。如果所确定的虚拟页 #(00-C) 未被分配某一个实际页 209,则存储系统 103 对该虚拟页 207 分配某一个自由(未分配状态)的实际页 #(01-D),在所分配的实际页 #(01-D) 中写入依照写入命令的写入对象数据要素。

[0158] 本实施例中,数据迁移以页单位进行。具体地,例如如图 1 所示,存储系统 103 进行以下的处理:

[0159] * 将分配给虚拟页 #(00-C) 的实际页 #(01-D) 内的数据要素迁移到自由(未分配状态)的实际页 #(02-E)

[0160] * 将虚拟页 #(00-C) 的分配源从实际页 #(01-D) 变更为实际页 #(02-E)

[0161] * 将实际页 #(01-D) 的状态更新为自由(未分配状态)

[0162] 是否进行数据迁移以及将迁移目的地层配置为哪一个层是基于后述的层范围来决定的。层范围以自动方式来决定。

[0163] 以下,详细说明实施例 1。

[0164] 图 2 表示存储系统 103 的构成。

[0165] 存储系统 103 具有多个 PDEV105、与多个 PDEV105 连接的存储控制装置、以及电源单元 100。

[0166] 多个 PDEV105 包括多种 PDEV,例如 SSD(Solid State Drive:固态驱动器)以及 HDD(Hard Disk Drive:硬盘驱动器)。HDD 具有 HDD-FC(Fibre Channel:光纤通道)、HDD-SATA(Serial ATA:串行 ATA)、以及 HDD-SAS(Serial Attached SCSI:串行连接 SCSI)等。

[0167] 存储控制装置具有控制器和与控制器连接的多个 CMPK(高速缓存封装:Cache Memory Package)119。控制器具有多个 CHA(通道适配器:Channel Adapter)111、多个 DKA(盘适配器:Disk Adapter)113、多个 MPPK(微处理器封装:MicroProcessor Package)121、以及 SW(开关:Switch)117。从冗余化的观点出发,CHA111、DKA113、MPPK121、SW117 以及电源单元 100 分别存在多个,但是它们中至少一者的数量不限于附图所示的 2,可比 2 多或者少。CMPK119 的数量可比 2 多。

[0168] 电源单元 100 将基于来自商用电源电力的电力提供给各封装 119、111、113、121 以及 117。

[0169] CHA111 是与主机 101 连接的接口装置。CHA111 从主机 101 接收 I/O 命令(写入

命令或者读出命令),并将接收到的 I/O 命令转发给多个 MPPK121 中的某一个。

[0170] CHA111 具有协议芯片 112 和 LR(Local Router:本地路由器)114。协议芯片 112 进行用于与主机 101 的通信的协议转换。LR114 将接收到的 I/O 命令转发给与根据该 I/O 命令具有的 I/O 目的地信息确定的 I/O 目的地对应的 MPPK121。

[0171] DKA113 是与 HDD105 连接的接口装置。DKA113 从 HDD105 读出数据并转发给 CMPK119 的 DRAM(Dynamic Random Access Memory:动态随机存取存储器),或者向 HDD105 写入来自 CMPK119 的数据。

[0172] MPPK121 是具有一个或者多个 MP(微处理器:MicroProcessor)的装置。MP 处理来自 CHA111 的 I/O 命令。

[0173] SW117 上连接有多个封装、即 CHA111、DKA113、CMPK119 以及 MPPK121。SW117 控制 PK(封装:Package)间的连接。

[0174] CMPK119 具有易失性存储器以及 / 或者非易失性存储器。CMPK119 中具有暂时存储例如作为针对虚拟页 207(实际页 209)的 I/O 的对象的数据要素的存储区域(以下称作高速缓存或者将其略称为 CM)。另外,CMPK119 中具有存储各种信息以及计算机程序的存储区域。关于信息以及计算机程序,将在下文中进行说明。

[0175] 存储系统 103 上连接有管理装置 107。管理装置 107 例如为具有显示装置以及输入装置的计算机。管理者能够从管理装置 107 对存储系统 103 进行各种设定。

[0176] 图 3 表示存储系统 103 中各种存储区域的关系。

[0177] 如图 3 所示,从下位到上位,管理一个以上的 RG301、一个以上的池 201、一个以上的 VVOL203。

[0178] RG301 是 RAID 组(group)的略称。一个 RG301 由同一种类的 PDEV 构成。PDEV 的种类例如由性能以及单位成本中的至少一者定义。性能是指例如数据的 I/O 的速度、或者响应时间(从由主机接受到命令起到返回响应为止的时间长度)。单位成本是保存单位大小(size)的数据所需的价格(例如所谓的比特成本)。例如, RG#00 由多个 SSD 构成, RG#01 由多个 HDD-SAS 构成, RG#02 由多个 HDD-SATA 构成。而且,构成一个 RG 的多个 PDEV 的容量例如是相同的。

[0179] 池 201 是实际页 207 的集合。具体地,例如,池 201 由一个以上的 RVOL205 构成,各 RVOL205 由一个以上的实际页 207 构成。以下,对于构成池 201 的 RVOL,将池 VOL 加以省略地表示为“PVOL”。

[0180] 构成一个池 201 的实际页 207 进行了层次化。即,一个池 201 中存在多个层 303。层 303 例如在每种 RVOL 中存在。例如,根据图 3,由于 RVOL 的种类为 3 种,因此层 303 的数量也为 3。因此,一个层 303 中包括与该层对应种类的 RVOL205。具体地,例如,层 #00 是 SSD 这一种类,基于由 SSD 构成的 RG301 的 RVOL 属于层 #00,层 #01 是 HDD-SAS 这一种类,基于由 HDD-SAS 构成的 RG301 的 RVOL 属于层 #01,层 #03 是 HDD-SATA 这一种类,基于由 HDD-SATA 构成的 RG301 的 RVOL 属于层 #03。而且,层 303 不一定严密地针对每种 RVOL 进行设置,基于类似种类的 RG 的多个 RVOL205 可属于同一层 303。

[0181] VVOL203 与一个以上的池 201 中成为分配源的池 201 相关联。另外, VVOL203 与选择出的池 201 中可作为分配源使用的层(以下称作分配层)303 相关联。另外,一个以上的分配层 303 之一作为默认的层 303 与 VVOL203 相关联。根据图 3, VVOL#01 的分配层是层

#00 以及 #02。

[0182] 另外,根据上述的说明,对于多个 RVOL 而言,只要它们属于同一层,则存在根基为一个 RG 的可能性,但是如果它们分别属于不同的层,则成为根基的 RG 各不相同。在对与同一层 303 相关联的一个或者多个 VVOL203 的多个虚拟页依次分配多个实际页的情况下,优选跨越与该层 303 相关联的 RG 来依次选择实际页。具体地,在从该层 303 所含多个 RVOL205 选择多个实际页的情况下,考虑到与 RG 的关系,依次从不同的 RG 依次选择实际页。由此,能够提高存取性能。

[0183] 图 4 表示 CMPK119 所存储的信息以及计算机程序。需要说明的是,在本实施例中,采用表作为信息的一个例子,但是信息也可以是表以外的形式(例如队列形式)。

[0184] CMPK119 存储层定义表 401、层管理表 402、池管理表 403、VVOL 管理表 404、PVOL 管理表 405、分配表 407、实际页管理表 408、监视表 409、度数分布表 410、迁移页表 411 以及迁移定义表 412 作为信息。另外,CMPK119 存储控制 UI (User Interface :用户接口) 的程序即 UI 控制程序 413、控制 I/O 的程序即 I/O 控制程序 414、以及进行重新配置的程序即重新配置程序 415 作为计算机程序。而且,这些表以及计算机程序中的至少一个也可以存储在 CMPK119 以外的存储资源。例如,至少一个计算机程序以及 / 或者迁移页表 411 可以存储在 MPPK121 内的存储器中。

[0185] 下面说明各表。需要说明的是,在以下的说明中,对象使用 ID 来确定,但是也可以取代 ID 或者在其基础上,使用名称以及编号中的至少一个来确定。

[0186] 图 5 表示层定义表 401 的构成。

[0187] 层定义表 401 是表示层的定义的表。具体地,例如,表 401 表示 RVOL 种类与层 ID 的关系。根据表 401,作为与 RVOL 种类“SSD”对应的层,准备附加有“00”这一 ID 的层。

[0188] 图 6 表示层管理表 402 的构成。

[0189] 层管理表 402 是用于管理与层相关的信息的表。具体地,例如,表 402 针对每层具有池 ID、层 ID、层范围、PVOL ID、层阈值、使用比例、实际页 ID。以一个层(以下,在图 6 的说明中称作“对象层”)为例,这些信息要素具体如下。

[0190] * “池 ID”是具有对象层的池的 ID。

[0191] * “层 ID”是对象层的 ID。根据层 ID,成为属于对象层的 PVOL 的根基的 PDEV 的种类被唯一确定。

[0192] * “层范围”是应该存在于对象层中的实际页的负荷(例如 I/O 频度)的范围。

[0193] * “PVOL ID”是属于对象层的 PVOL 的 ID。

[0194] * “层阈值”是对象层的使用比例的阈值。

[0195] * “使用比例”是对象层的使用比例,即已分配的实际页的数量相对于构成对象层的实际页的数量的比例。

[0196] * “实际页 ID”是属于对象层的所有实际页的 ID。

[0197] 根据图 6 的表 402 的例子,层阈值是针对每层进行了设定,但是也可以取而代之,对一个池设定一个层阈值,该一个层对于该一个池所具有的所有层是共通的。

[0198] 图 7 表示池管理表 403 的构成。

[0199] 池管理表 403 是用于管理与池相关的信息的表。具体地,例如,表 403 针对每个池具有池 ID、层次化、重新配置、自动执行、重新配置执行周期、重新配置执行时刻、监视时间

带以及池状态。以一个池（以下，图 7 的说明中称作“对象池”）为例，这些信息要素具体如下。

[0200] * “池 ID”是对象池的 ID。

[0201] * “层次化”是是否对对象池内的实际页进行层次化。“ON”意味着进行层次化（即、对象池中设有多个层），“OFF”意味着不进行层次化（即、对象池中不设层）。

[0202] * “重新配置”是是否对对象池内的数据要素进行重新配置。“ON”意味着进行重新配置，“OFF”意味着不进行重新配置。“ON”的情况下，分配到对象池的 VVOL 或者其虚拟页的 I/O 频度受到监视，“OFF”的情况下，分配到对象池的 VVOL 或者其虚拟页的 I/O 频度不受监视。关于 I/O 频度的监视，重要的是，如果实际不对分配给 I/O 目的地 VVOL 或者虚拟页的实际页进行 I/O，则 I/O 频度不被更新，在对该分配的实际页进行 I/O 时，I/O 频度得到更新。这一点在下文中叙述写入处理以及读出处理时也将进行说明。

[0203] * “自动执行”是对象池内的数据要素的重新配置是自动开始还是手动开始。“ON”意味着重新配置以自动形式开始，“OFF”意味着重新配置以手动形式开始。

[0204] * “重新配置执行周期”是进行对象池内的数据要素的重新配置的周期。例如，“1 日”意味着每 1 日（24 小时）开始进行重新配置。

[0205] * “重新配置执行时刻”是开始对象池内的数据要素的重新配置的时刻。

[0206] * “监视时间带”是对分配给分配到对象池的 VVOL 的实际页的 I/O 频度进行监视的时间带。

[0207] * “池状态”是对象池的状态。作为“池状态”的值，例如有“监视中”、“重新配置中”、“非监视中”。“监视中”意味着正在监视分配到对象池的 VVOL 或者该 VVOL 内的虚拟页的 I/O 频度，并且没有进行数据要素的重新配置。“重新配置中”意味着正在进行数据要素的重新配置（可以是对象池内的重新配置，也可以是从对象池向其它池的数据要素的重新配置）。“非监视中”意味着既非正在监视 I/O 频度也非正在进行数据要素的重新配置。

[0208] 图 8 表示 VVOL 管理表 404 的构成。

[0209] VVOL 管理表 404 是用于管理与 VVOL 相关的信息的表。具体地，表 404 针对每个 VVOL 具有 VVOL ID、PVOL ID、分配层 ID、默认层 ID、容量、重新配置、自动执行、重新配置执行周期、重新配置执行时刻以及监视时间带。这些信息要素中，重新配置、自动执行、重新配置执行周期、重新配置执行时刻以及监视时间带与池管理表 403 中的信息要素相同，但是如果该表 404 中的信息要素（例如“重新配置”）的值与池管理表 403 中的相同信息要素（例如“重新配置”）的值不同，则可优先该表 404 中的值。即，在池管理表 403 中，针对一个池设定信息要素的值，其结果，能够使该值反映到针对分配到该池的所有 VVOL 的设定中，能够针对每个 VVOL 进行设定。如果未针对 VVOL 特别设定值，该针对分配给该 VVOL 的池设定的值作为针对该 VVOL 的值。以一个 VVOL（以下，图 8 的说明中称作“对象 VVOL”）为例，该表 404 中的上述信息要素具体如下。

[0210] * “VVOL ID”是对象 VVOL 的 ID。

[0211] * “PVOL ID”是分配给对象 VVOL 的池（以下称作分配池）的 ID。

[0212] * “分配层 ID”是分配给分配池所具有的多个层中的对象 VVOL 的层的 ID。从分配层向对象 VVOL 内的虚拟页分配实际页。

[0213] * “默认层 ID”是默认层的 ID。默认层是指成为最初分配给作为从主机 101 的写

入目的地的虚拟页的实际页的提供源的层。

[0214] * “容量”是对象 VVOL 的容量。

[0215] * “重新配置”是是否对对象 VVOL 内的数据要素进行重新配置。“ON”意味着进行重新配置，“OFF”意味着不进行重新配置。“ON”的情况下，对象 VVOL 或者其虚拟页的 I/O 频度受到监视，“OFF”的情况下，对象 VVOL 或者其虚拟页的 I/O 频度不受监视。

[0216] * “自动执行”是对象 VVOL 内的数据要素的重新配置是自动开始还是手动开始。“ON”意味着重新配置以自动形式开始，“OFF”意味着重新配置以手动形式开始。

[0217] * “重新配置执行周期”是进行对象 VVOL 内的数据要素的重新配置的周期。

[0218] * “重新配置执行时刻”是开始对象 VVOL 内的数据要素的重新配置的时刻。

[0219] * “监视时间带”是对对象 VVOL 或者虚拟页的 I/O 频度进行监视的时间带。

[0220] 图 9 表示 PVOL 管理表 405 的构成。

[0221] PVOL 管理表 405 是用于管理与 PVOL 相关的信息的表。具体地，例如，表 405 针对每个 PVOL 具有 RG ID、PDEV ID、种类、PVOL ID、以及容量。以一个 PVOL（以下，图 9 的说明中称作“对象 PVOL”）为例，这些信息要素具体如下。

[0222] * “RG ID”是作为对象 PVOL 的根基的 RG 的 ID。

[0223] * “PDEV ID”是构成作为对象 PVOL 的根基的 RG 的所有 PDEV 的 ID。

[0224] * “种类”是对象 PVOL 的种类。这与作为对象 PVOL 的根基的 PDEV 的种类相同。例如，RVOL 种类如果为“SSD”，则对象 PVOL 的种类也为“SSD”。

[0225] * “PVOL ID”是对象 PVOL 的 ID。

[0226] * “容量”是对象 PVOL 的容量。

[0227] 图 10 表示分配表 407 的构成。

[0228] 分配表 407 是表示对哪个虚拟页分配了哪个实际页的表。具体地，例如，表 407 针对每个虚拟页具有 VVOL ID、虚拟页 ID、池 ID、实际页 ID 以及层 ID。以一个虚拟页（以下，图 10 的说明中称作“对象虚拟页”）为例，这些信息要素具体如下。

[0229] * “VVOL ID”是具有对象虚拟页的 VVOL 的 ID。

[0230] * “虚拟页 ID”是对象虚拟页的 ID。

[0231] * “池 ID”是分配给具有对象虚拟页的 VVOL 的池的 ID。

[0232] * “实际页 ID”是分配给对象虚拟页的实际页的 ID。

[0233] * “层 ID”是具有分配给对象虚拟页的实际页的层的 ID。

[0234] 图 11 表示实际页管理表 408 的构成。

[0235] 实际页管理表 408 是用于管理各实际页的状态的表。具体地，例如，表 408 针对每个实际页具有池 ID、实际页 ID 以及状态。以一个实际页（以下，图 11 的说明中称作“对象实际页”）为例，这些信息要素具体如下。

[0236] * “池 ID”是具有对象实际页的池的 ID。

[0237] * “实际页 ID”是对象实际页的 ID。

[0238] * “状态”是对象实际页的状态。作为状态的值，例如有“使用中”和“自由”。“使用中”意味着对象实际页已经分配给某一个虚拟页。“自由”意味着如下状态：对象实际页未分配给任意一个虚拟页，因此能够对任意一个虚拟页进行分配。

[0239] 图 12 表示监视表 409 的构成。

[0240] 监视表 409 存在于每个 VVOL。表 409 是表示虚拟页的 I/O 频度的表。以一个 VVOL(以下,在图 12 的说明中称作“对象 VVOL”)为例,说明表 409。

[0241] 表 409 具体例如针对每一虚拟页具有 VVOL ID、虚拟页 ID、合计 I/O 数、平均 I/O 数、最大 I/O 数以及最终 I/O 时刻。以一个虚拟页(以下,图 12 的说明中称作“对象虚拟页”)为例,这些信息要素具体如下。

[0242] * “VVOL ID”是具有对象虚拟页的 VVOL(即对象 VVOL)的 ID。

[0243] * “虚拟页 ID”是对象虚拟页的 ID。

[0244] * “合计 I/O 数”是在监视时间带内对对象虚拟页进行的 I/O 的数量。

[0245] * “平均 I/O 数”是通过“合计 I/O 数”的值除以预定时间计算出的值。

[0246] * “最大 I/O 数”是构成监视时间带的多个时间带中多个 I/O 数(针对对象虚拟页的 I/O 数)中最大的 I/O 数。

[0247] * “最终 I/O 时刻”是对对象虚拟页进行了 I/O 的最新时刻。

[0248] 表 409 存在例如两个(也可多于两个)。两个表 409 交替使用。具体地,例如,在执行重新配置的情况下,使用关于该执行时刻点附近的监视时间带的表 409,生成后述的迁移页表 411,并且使用另一个表 409,进行 I/O 频度的监视。在以下的说明中,有时将正在使用的表 409 称作激活(active)的表 409,将未使用的表 409 称作待机(stand-by)的表 409。

[0249] 图 13 表示度数分布表 410 的构成。

[0250] 度数分布表 410 是表示 I/O 数与虚拟页数的相关的表。该表 410 基于监视表 409 生成。作为 I/O 数,可采用上述的表 409 中合计 I/O 数、平均 I/O 数以及最大 I/O 数中的至少一个。图 13 是采用平均 I/O 数的例子。即,图 13 所示的表 410 表示不同平均 I/O 数的虚拟页的数量。根据图 13 可知,平均 I/O 数为“0”的虚拟页的数量为“561”。该表 410(或者基于该表 410 生成的图)显示在管理装置 107 中。管理者能够基于表 401(或者基于其的图)通过手动指示存储系统 103 开始数据要素的重新配置。

[0251] 度数分布表 410 可随着监视表 409 的更新而更新,也可在从管理者指令显示度数分布时基于表 409 来生成。

[0252] 图 14 表示迁移页表 411 的构成。

[0253] 迁移页表 411 是进行页单位的数据迁移时基于监视表 409 生成的表。表 411 表示迁移源与迁移目的地的关系。具体地,例如,表 411 针对作为迁移对象的每个虚拟页具有虚拟页 ID、迁移源层 ID、迁移目的地层 ID 以及迁移状态。以一个虚拟页(以下,在图 14 的说明中称作“对象虚拟页”)为例,这些信息要素具体如下。

[0254] * “虚拟页 ID”是对象虚拟页的 ID。

[0255] * “迁移源层 ID”是具有分配给对象虚拟页的实际页(迁移源实际页)的层的 ID。

[0256] * “迁移目的地层 ID”是成为数据要素的迁移目的地的层的 ID。迁移目的地实际页是该层中某一个自由的实际页。

[0257] * “迁移状态”是关于对象虚拟页的迁移的状态。作为“迁移状态”的值,例如具有“已迁移”、“迁移中”以及“未迁移”。“已迁移”意味着从迁移源实际页到迁移目的地层中某一个实际页已经将数据要素迁移。“迁移中”意味着从迁移源实际页到迁移目的地层中某一个实际页正在迁移数据要素。“未迁移”意味着从迁移源实际页到迁移目的地层中某一个实际页尚未开始迁移数据要素。

[0258] 而且,迁移页表 411 可以取代虚拟页 ID 或者在其基础上具有与迁移源相关的其它种类的信息要素。例如,分配给虚拟页的实际页的 ID 可以包含在表 411 中。

[0259] 另外,迁移页表 411 可以取代迁移目的地层 ID 或者在其基础上具有与迁移目的地相关的其它种类的信息要素。例如,可基于表 402(图 6) 以及 408(图 11) 确定迁移目的地层中自由的实际页,将该确定出的实际页的 ID 包含在表 411 中。届时,确定出的实际页的状态可从“自由”更新为“使用中”(或者其它的状态(例如“保留”))。

[0260] 图 15 表示迁移定义表 412 的构成。

[0261] 迁移定义表 412 针对各 VVOL 表示与迁移有关的定义。具体地,表 412 针对每个 VVOL 具有 VVOL ID、层维持、分配层 ID 以及保存阈值。以一个 VVOL(以下,在图 15 的说明中称作“对象 VVOL”)为例,这些信息要素具体如下。

[0262] * “VVOL ID”是对象 VVOL 的 ID。

[0263] * “分配层 ID”是对象 VVOL 的分配层的 ID。

[0264] * “保存阈值”是分配层的保存比例的阈值。该阈值针对每个分配层设定。对于“保存比例”,如果针对一个分配层(以下,在该段落的说明中称作“对象分配层”)而言,则如下所述。即,保存比例是对象 VVOL 内所有数据要素中多大程度比例的数据要素保存在了对象分配层中。

[0265] 根据图 15 的表 412 的例子,保存阈值针对每个层设定,但是也可以取而代之,对于一个池设定一个保存阈值,该一个层对于该一个池所具有的所有层是共通的。

[0266] 以下说明本实施例中进行的处理。需要说明的是,在以下的说明中,计算机程序所进行的处理实际上是执行计算机程序的 MP(微处理器)在进行。MP 位于 MPPK121(参照图 2) 中。

[0267] <池的生成>

[0268] 图 16 表示池生成处理的流程。需要说明的是,该<池的生成>的说明中,将通过该池生成处理新生成的池称作“对象池”。

[0269] (S1601)

[0270] UI 控制程序 413 从管理者通过管理装置 107 接受池生成指示。然后,程序 413 进行例如下述准备处理。

[0271] * 程序 413 将对象池的新的记录追加到表 402(图 6)、403(图 7)、408(图 11)。

[0272] * 程序 413 从管理者通过管理装置 107 接受 PVOL 的指定。具体地,例如,程序 413 显示与 RVOL 相关的信息(例如 ID、种类等),接受作为 PVOL 的 RVOL 的指定。

[0273] 针对每个被指定的 RVOL,进行 S1602 ~ S1605。以一个 RVOL(以下,在图 16 的说明中称作“对象 RVOL”)为例,说明 S1602 ~ S1605。

[0274] (S1602)

[0275] 程序 413 确定被指定的 RVOL 的种类。具体地,例如,CMPK119 存储有针对每个 RVOL 具有 ID、容量、种类(例如 SSD、HDD-SAS) 等的 RVOL 管理表,程序 413 根据该 RVOL 管理表确定被指定的 RVOL 的种类。

[0276] (S1603)

[0277] 程序 413 根据层管理表 403 判断与被确定的 RVOL 种类对应的层是否已经存在于对象池中。在该判断的结果为否定性的情况下,进行 S1604,在该判断的结果为肯定性的情

况下,进行 S1605。

[0278] (S1604)

[0279] 程序 413 根据层定义表 401 确定与被确定的 RVOL 种类对应的层的 ID。然后,程序 413 针对对象池将被确定的层 ID 追加到表 402(图 6)中。

[0280] (S1605)

[0281] 程序 413 在与被确定的 RVOL 种类对应的层(以下,在该段落中称作“对象层”)中追加对象 RVOL。具体地,例如,程序 413 在表 402(图 6)中登记对象 RVOL 的 ID 作为与对象层的层 ID 对应的 PVOL ID。

[0282] (S1606)

[0283] 程序 413 针对所有被指定的 RVOL 判断是否进行了 S1602 以后的处理。如果在被指定的 RVOL 中存在没有进行 S1602 以后的处理的 RVOL,则针对该 RVOL 进行 S1602 以后的处理。

[0284] (S1607)

[0285] 程序 413 显示池定义 GUI(Graphical User Interface:图形用户接口)。然后,程序 413 通过该 GUI 接受各种设定。

[0286] 图 17 表示池定义 GUI1700。

[0287] GUI1700 中显示对象池的 ID。而且,程序 413 也能够响应管理者对预定工具(例如下拉菜单)的操作,显示对象池以外的池(即、既有的池)的 ID。

[0288] 另外,GUI1700 具有管理者为了针对正在显示的 ID 所表示的池输入应该在表 403(图 7)中设定的信息要素“层次化”、“重新配置”、“自动执行”、“执行周期”、“执行时刻”、“监视时间带”、“层阈值”的值而使用的输入工具。作为本实施例中所说的输入工具,能够采用无线电按钮、复选框、文本输入栏等任意的工具。而且,GUI1700 中,针对“层阈值”,能够使一个阈值在对象池所具有的全部层中共通,或者能够针对对象池所具有每个层设定阈值。

[0289] 而且,如上所述,由于 GUI1700 中也能够显示既有的池的 ID,因此使用 GUI1700 能够变更关于管理者所希望的既有的池的设定。届时,例如,如果管理者所希望的层阈值高于设定为该层阈值的层的使用比例,则程序 413 可将其作为不合适的层阈值在管理装置 105 中显示警告,并且不在表 402(图 6)中登记该层阈值。

[0290] 再次参照图 16。

[0291] (S1608)

[0292] 程序 413 更新表。具体地,例如,程序 413 进行下述处理。

[0293] * 程序 413 在表 403(图 7)中登记输入到图 17 的 GUI1700 的、信息要素“层次化”、“重新配置”、“自动执行”、“执行周期”、“执行时刻”、“监视时间带”、“层阈值”的值。

[0294] * 程序 413 将对象池划分为多个实际页,将该多个实际页的 ID 追加到表 402(图 6)、以及 408(图 11)中。

[0295] 实际页的容量在所有的池中可以是共通的。因此,构成对象池的实际页的数量可依存于对象池的容量(构成对象池的 PVOL 的总容量)。

[0296] 而且,对于每个 PVOL,实际页的容量可以不同。具体地,例如,构成基于 RAID5(4D+1P)的 RG 的 PVOL 的实际页的容量与构成基于 RAID6(4D+2P)的 RG 的 PVOL 的实

际页的容量可以不同。之所以这样,是因为保存在一个实际页中的奇偶校验位的数量不同。

[0297] <VVOL 的生成>

[0298] 图 18 表示 VVOL 生成处理的流程。需要说明的是,该 <VVOL 的生成> 的说明中,将通过该 VVOL 生成处理新生成的 VVOL 称作“对象 VVOL”。

[0299] (S1801)

[0300] UI 控制程序 413 显示 VVOL 生成 GUI。然后,程序 413 通过该 GUI 接受各种设定。

[0301] 图 19 表示 VVOL 生成 GUI1900。

[0302] GUI1900 中显示对象 VVOL 的 ID。而且,程序 413 也能够响应管理者对预定工具(例如下拉菜单)的操作,显示对象 VVOL 以外的 VVOL(既有的 VVOL) 的 ID。

[0303] GUI1900 具有管理者为了针对正在显示的 ID 所表示的 VVOL 输入应该在表 404(图 8) 中设定的信息要素“池 ID”、“容量”、“重新配置”、“自动执行”、“执行周期”、“执行时刻”、“监视时间带”的值而使用的输入工具。

[0304] 另外,GUI1900 根据表 402(图 6) 确定与输入的池 ID 对应的层 ID,并显示被确定的所有层 ID。GUI1900 具有管理者为了针对正在显示的 ID 所表示的 VVOL 输入“分配层”、“保存阈值”以及“默认层”的值而使用的输入工具。管理者能够使用那些输入工具输入将输入的池 ID 所表示的池(分配给 VVOL 的池)具有的多个层中的哪个作为分配层、将分配层中的哪个作为“默认层”、以及各分配层的保存阈值。

[0305] 再次参照图 18。

[0306] (S1802)

[0307] 程序 413 更新表。具体地,例如,程序 413 进行下述处理。

[0308] * 程序 413 在表 404(图 8) 中登记输入到图 19 的 GUI1900 的、信息要素“池 ID”、“容量”、“重新配置”、“自动执行”、“执行周期”、“执行时刻”、“监视时间带”、“分配层 ID”以及“默认层 ID”的值。

[0309] * 程序 413 基于对象 VVOL 的容量计算构成对象 VVOL 的虚拟页的数量。程序 413 在表 407(图 10) 中登记对象 VVOL 的 ID 和计算出的虚拟页数量个的虚拟页 ID。

[0310] * 针对对象 VVOL,如果“重新配置”为“ON”,则程序 413 生成表 409(图 12)。这是因为在监视时间带中进行 I/O 频度的监视。而且,表 409 中登记上述计算出的虚拟页数量个的虚拟页 ID。

[0311] * 程序 413 针对对象 VVOL 在表 412(图 15) 中登记“分配层 ID”以及“保存阈值”。

[0312] <信息的显示>

[0313] 图 22 表示信息显示处理的流程。

[0314] UI 控制程序 413 从管理装置 107 接受显示命令(S2201),在管理装置 107 中显示由显示命令指定的信息(S2202)。

[0315] 作为可显示的信息的种类,例如为表 401 ~ 412 中的至少一个(或者对它们进行加工而得到的信息(例如图、摘要等))。

[0316] 图 23 表示度数分布 GUI2300。

[0317] 度数分布 GUI2300 在图 22 的 S2202 中显示。具体地,例如,在图 22 的 S2201 中,管理装置 107 依照来自管理者的请求发送指定了管理者所希望的 VVOL ID 的度数分布显示命令。在图 22 的 S2202 中,程序 413 响应该命令,从表 410(图 13) 提取与由该命令指定的

VVOL ID 对应的度数分布,并如图 23 所示将提取出的度数分布显示在管理装置 107 中。需要说明的是,度数分布是以表形式显示的,但是也可取代其或者在其基础上以图形式等其它形式来显示。

[0318] 管理者看到显示在 GUI2300 中的度数分布来判断是否迁移所希望的 VVOL 内的数据要素。例如,管理者可以在认为存在大量平均 I/O 数高的虚拟页数的情况下指示存储系统 103 迁移所希望的 VVOL 内的数据要素(例如,可以按压 GUI2300 中的“迁移执行”按钮)。

[0319] 这样,度数分布作为用于管理者判断是否进行数据迁移(VVOL 内的数据要素的迁移)的判断材料来使用。换言之,如果数据执行自动执行,则可以不计算度数分布。

[0320] <写入处理(向高速缓存写入)>

[0321] 图 24 表示写入处理(向高速缓存写入)的流程。

[0322] (S2401)

[0323] I/O 控制程序 414 从主机 101 接收写入命令。

[0324] (S2402)

[0325] 程序 414 根据接收到的写入命令所具有的 I/O 目的地信息确定写入目的地的 VVOL 以及虚拟页。此处,确定出多个虚拟页的情况很多。之所以这样,是由于写入对象数据的大小通常大于一个虚拟页的容量。针对所确定的每个虚拟页,进行 S2403 以后的处理。以下,以一个虚拟页为例(在图 24 的说明中称作“对象虚拟页”),说明 S2403 以后的处理。另外,在图 24 的说明中,将针对对象虚拟页的写入对象数据要素称作“对象数据要素”。

[0326] (S2403)

[0327] 程序 414 判断对象虚拟页的 ID 是否存在于迁移页表 411 中。在该判断的结果是否定性的情况下,进行 S2404 ~ S2406,在该判断的结果为肯定性的情况下,进行 S2407 ~ S2408。

[0328] (S2404)

[0329] 程序 414 基于分配表 407(图 10)判断实际页是否分配给了对象虚拟页。如果该判断的结果为否定性的,则进行 S2405。

[0330] (S2405)

[0331] 程序 414 进行以下处理。

[0332] * 程序 414 基于 VVOL 管理表 404(图 8)确定与写入目的地 VVOL 对应的池以及该池内的默认层。

[0333] * 程序 414 基于层管理表 402(图 6)以及实际页管理表 408(图 11)确定被确定的默认层内的自由的实际页。

[0334] * 程序 414 对对象虚拟页分配所确定的实际页。即,程序 414 在分配表 407(图 10)中与对象虚拟页的 ID 对应的栏中登记所分配的实际页的 ID、具有该实际页的层的 ID、具有该层的池的 ID。

[0335] * 程序 414 在根据默认层不能确定自由的实际页的情况下,可以根据与写入目的地 VVOL 对应的分配层确定自由的实际页,并将所确定的实际页分配给对象虚拟页。

[0336] * 程序 414 更新表 402(图 6)中具有所分配的实际页的层(在该 S2403 的说明中称作“分配源层”)的“使用比例”的值。

[0337] * 程序 414 进行分配源层的“使用比例”的更新后的值是否超过了分配源层的“层

阈值”的值的判断（以下称作“判断 Z”）。分配源层的“层阈值”的值是登记在表 402（图 6）中的值、或者登记在表 412（图 15）中的值，但是可优先采用登记在表 412 中的值。

[0338] * 如果判断 Z 的结果是肯定性的，则程序 414 可以进行下述 (a) 以及 (b) 中的至少一个处理。

[0339] (a) 程序 414 可将错误 (error) 通知给管理装置 107（以及 / 或者作为写入命令发送源的主机 101）。管理者在针对分配源层知道了错误的情况下，能够在分配源层中追加 RVOL。

[0340] (b) 程序 414 调用重新配置程序 415。重新配置程序 415 进行图 38 所示的处理。图 38 所示的处理将在后面说明。

[0341] (S2406)

[0342] 程序 414 从高速缓存确保用于保存在对象虚拟页中的数据要素的 CM 区域。

[0343] (S2407)

[0344] 程序 414 根据迁移页表 411 中与对象虚拟页对应的“迁移状态”的值确保 CM 区域。具体如下。

[0345] << 情形 A：“迁移状态”为“未迁移” >>

[0346] 例如，进行以下（具体例 A1）以及（具体例 A2）中的某一个。

[0347] （具体例 A1）

[0348] * 程序 414 确保 CM 区域，将与对象虚拟页对应的迁移目的地层中的自由的实际页决定为写入目的地。更具体地，例如，如果迁移目的地层中某一个自由的实际页内的数据要素在高速缓存中，则程序 414 确保存在该数据要素的 CM 区域，并且将该自由的实际页决定为写入目的地。迁移目的地层中的自由的实际页如上所述能够基于表 402（图 6）以及 408（图 11）来确定。

[0349] * 程序 414 将与对象虚拟页对应的“迁移状态”更新为“已迁移”。其结果，针对对象虚拟页不进行数据要素的迁移。即，换言之，针对对象虚拟页，数据要素的迁移被取消。

[0350] * 程序 414 将所决定的写入目的地实际页取代原来的实际页分配给对象虚拟页。即，程序 414 在分配表 407 中与对象虚拟页的 ID 对应的栏中改写上述所决定的写入目的地实际页的 ID 以及具有该实际页的层的 ID（迁移目的地层的 ID）。

[0351] * 程序 414 将实际页管理表 408（图 11）中分配给对象虚拟页的原来的实际页的“状态”更新为“自由”。另外，程序 414 将该表 408 中新分配给对象虚拟页的实际页（所决定的写入目的地实际页）的“状态”更新为“使用中”。

[0352] （具体例 A2）

[0353] * 程序 414 确保 CM 区域，将分配给对象虚拟页的实际页（迁移源层内的实际页）决定为写入目的地。更具体地，例如，如果迁移目的地层中任何一个自由的实际页内的数据要素都不在高速缓存中，则程序 414 确保 CM 区域（例如存在分配给对象虚拟页的实际页内的数据要素的区域），并且将分配给对象虚拟页的实际页（迁移源层内的实际页）决定为写入目的地。

[0354] * 程序 414 可将与对象虚拟页对应的“迁移状态”更新为“中断”。由此，例如，针对对象虚拟页，不进行基于重新配置程序 415 的数据迁移。而且，该“迁移状态”的值例如可以通过程序 414 在后述的降级 (destage) 完成的情况下从“中断”恢复成“未迁移”。

[0355] <<情形 B：“迁移状态”为“迁移中”>>

[0356] 例如进行以下的（具体例 B1）～（具体例 B3）之一。

[0357] （具体例 B1）

[0358] * 程序 414 将与对象虚拟页对应的“迁移状态”更新为“中断”。然后，程序 414 进行上述的（具体例 A1）或者（具体例 A2）。在进行了（具体例 A1）的情况下，与对象虚拟页对应的“迁移状态”更新为“已迁移”。在进行了（具体例 A2）的情况下，与对象虚拟页对应的“迁移状态”更新为“未迁移”。

[0359] （具体例 B2）

[0360] * 程序 414 确保 CM 区域。在（具体例 B2）的情况下，在与对象虚拟页对应的“迁移状态”更新为“已迁移”之后进行降级。因此，写入目的地（降级目的地）为迁移目的地层内的实际页。之所以这样，是因为对象虚拟页被分配了迁移目的地层内的实际页。

[0361] （具体例 B3）

[0362] * 程序 414 确保 CM 区域。在（具体例 B3）的情况下，在与对象虚拟页对应的“迁移状态”为“迁移中”期间进行降级，数据要素被降级到分配给对象虚拟页的实际页和迁移目的地层内的实际页的双方（进行所谓的二重写入）。

[0363] <<情形 C：“迁移状态”为“已迁移”>>

[0364] 例如进行以下处理。

[0365] * 程序 414 确保 CM 区域。该情形中，写入目的地（降级目的地）是迁移目的地层内的实际页。之所以这样，是因为由于与对象虚拟页对应的“迁移状态”被更新为了“已迁移”，所以分配给对象虚拟页的实际页如后面说明的那样被变更为成为迁移目的地的实际页。

[0366] (S2408)

[0367] 在对于通过 S2402 所确定的所有虚拟页确保了 CM 区域的情况下，程序 414 向所确保的一个以上的 CM 区域写入作为写入对象的数据，并将写入完成向主机 101 进行响应。

[0368] < 写入处理（降级）>

[0369] 图 25 表示降级处理的流程。

[0370] I/O 控制程序 414 针对构成写入对象数据的各数据要素进行降级处理。

[0371] (S2501)

[0372] 程序 414 进行降级。即，程序 414 将存在于高速缓存中的多个数据要素中尚未写入实际页的数据要素（脏数据要素）写入实际页。降级目的地实际页是分配给成为该实际页的写入目的地的虚拟页的实际页（根据分配表 407 确定的实际页）。而且，此时，如果与该虚拟页对应的“迁移状态”为“中断”，则程序 414 可以将该“迁移状态”更新为“未迁移”。

[0373] (S2502)

[0374] 程序 414 基于表 403（图 7）或者 404（图 8）判断是否有更新监视表 409 的必要。可优先表 404（图 8）。具体地，例如，程序 414 进行以下的两个判断，即判断 P 以及判断 Q。

[0375] （判断 P）程序 414 判断与具有 S2501 中的降级目的地实际页的池（以下，在图 25 的说明中称作“对象池”）对应的“重新配置”是否为“ON”、或者与具有分配到该降级目的地实际页的虚拟页的 VVOL（以下，在图 25 的说明中称作“对象 VVOL”）对应的“重新配置”是否为“ON”。在该判断 P 的结果为否定性的情况下，不进行 S2503（即，没有更新监视表 409

的必要)。

[0376] (判断 Q) 在判断 P 的结果为肯定性的情况下,程序 414 判断当前的时刻是否在与对象池对应的“监视时间带”的值所表示的时间带内。在该判断 Q 的结果为否定性的情况下,不进行 S2503。另一方面,不仅判断 P,该判断 Q 的结果也是肯定性的情况下,进行 S2503(即,有更新监视表 409 的必要)。

[0377] (S2503)

[0378] 程序 414 更新监视表 409。具体如下。

[0379] * 如果表 412(图 15)中与对象 VVOL 对应的“页单位迁移”为“ON”,则程序 414 更新与对象 VVOL 对应的激活的监视表 409。具体地,例如,程序 414 更新激活的监视表 409 中的、与被降级的数据要素的写入目的地虚拟页对应的“合计 I/O 数”的值(例如增加 1),并且更新“最终 I/O 时刻”的值。另外,程序 414 针对例如该写入目的地虚拟页将降级时间点的时刻追加到上述的 I/O 履历表中。该情况下,I/O 履历表中可以记录写入目的地虚拟页的 ID。例如,可以基于该 I/O 履历表,而后由程序 414 更新激活的监视表 409 中“平均 I/O 数”以及“最大 I/O 数”。

[0380] < 读出处理 >

[0381] 图 26 表示读出处理的流程。

[0382] (S2600)

[0383] I/O 控制程序 414 从主机 101 接收读出命令。然后,程序 414 根据接收到的读出命令所具有的 I/O 目的地信息确定读出源 VVOL 以及虚拟页。

[0384] (S2601)

[0385] 程序 414 判断高速缓存上是否具有读出对象数据。该判断的结果为否定性的情况下,进行 S2603,该判断的结果为肯定性的情况下,进行 S2602。

[0386] (S2602)

[0387] 程序 414 将高速缓存上的读出对象数据发送给主机 101。该情况下,与读出源 VVOL 对应的监视表 409 不被更新。这是由于没有进行针对实际页的 I/O。

[0388] (S2603)

[0389] 程序 414 基于分配表 407 判断是否对通过 S2600 所确定的虚拟页分配了实际页。该判断的结果为否定性的情况下,进行 S2604,该判断的结果为肯定性的情况下,进行 S2605。

[0390] (S2604)

[0391] 程序 414 向主机 101 发送预定响应(例如表示零的数据)。

[0392] (S2605)

[0393] 程序 414 判断通过 S2600 所确定的一个以上的虚拟页的 ID 中的至少一个是否在迁移页表 411 中已登记、并且与该已登记的虚拟页 ID 对应的“迁移状态”是否为“迁移中”。该判断的结果为否定性的情况下,进行 S2607,该判断的结果为肯定性的情况下,进行 S2606。

[0394] (S2606)

[0395] 程序 414 从分别分配给通过 S2600 所确定的一个以上的虚拟页的一个以上的实际页(根据分配表 407 所确定的一个以上的实际页)读出数据要素。程序 414 将由那些数据

要素所构成的读出对象数据发送给主机 101。

[0396] (S2607)

[0397] 程序 414 针对“迁移状态”为“迁移中”的虚拟页等待至“迁移状态”更新为“已迁移”。然后,程序 414 从分别分配给通过 S2600 所确定的一个以上的虚拟页的一个以上的实际页(根据分配表 407 所确定的一个以上的实际页)读出数据要素。程序 414 将由那些数据要素所构成的读出对象数据发送给主机 101。

[0398] (S2608)

[0399] 程序 414 基于表 403(图 7) 或者 404(图 8) 判断是否有更新监视表 409 的必要。可优先表 404(图 8)。具体地,例如,程序 414 进行以下的两个判断,即判断 V 以及判断 W。

[0400] (判断 V) 程序 414 判断与分配给读出源 VVOL 的池(以下,在图 26 的说明中称作“读出源池”)对应的“重新配置”是否为“ON”、或者与读出源 VVOL 对应的“重新配置”是否为“ON”。在该判断 V 的结果为否定性的情况下,不进行 S2609(即,没有更新监视表 409 的必要)。

[0401] (判断 W) 在判断 V 的结果为肯定性的情况下,程序 414 判断当前的时刻是否在与对象池对应的“监视时间带”的值所表示的时间带内。在该判断 W 的结果为否定性的情况下,不进行 S2609。另一方面,不仅判断 V,该判断 W 的结果也是肯定性的情况下,进行 S2609(即,有更新监视表 409 的必要)。

[0402] (S2609)

[0403] 程序 414 更新监视表 409。具体如下。

[0404] * 如果表 412(图 15) 中与读出源 VVOL 对应的“页单位迁移”为“ON”,则程序 414 更新与读出源 VVOL 对应的激活的监视表 409。具体地,例如,程序 414 更新激活的监视表 409 中读出源的与写入目的地虚拟页对应的“合计 I/O 数”的值(例如增加 1),并且更新“最终 I/O 时刻”的值。另外,程序 414 针对例如该读出源虚拟页将读出时间点的时刻追加到上述的 I/O 履历表中。该情况下,I/O 履历表中还可以记录读出源虚拟页的 ID。例如,可以基于该 I/O 履历表,而后由程序 414 更新激活的监视表 409 中“平均 I/O 数”以及“最大 I/O 数”。

[0405] <I/O 频度的监视与重新配置处理的关系>

[0406] 图 27 表示 I/O 频度的监视与重新配置的关系的概要。

[0407] 在监视针对实际页的 I/O,并进行了针对实际页的 I/O 的情况下,分配到该实际页的虚拟页或者具有该虚拟页的 VVOL 的 I/O 频度被更新。即,监视表 409 被更新。换言之,如上所述,I/O 控制程序 414 虽然从主机 101 接受了读出命令,但是在没有从实际页读出数据要素的情况下,监视表 409 不被更新。

[0408] 随着监视表 409 的更新,度数分布表 410 被更新。而且,度数分布表 410 可以不预先准备,例如可以在由管理者指令了显示度数分布的情况下基于监视表 409 生成度数分布表 410,并基于该表 410 显示度数分布。而且,还可以取代度数分布,显示监视表 409。

[0409] 重新配置处理具有层判定处理和迁移处理。层判定处理中生成迁移页表 411。该表 411 可以基于监视表 409 所表示的 I/O 频度和层管理表 402(图 6) 内的层范围来生成。迁移处理中基于所生成的迁移页表 411 进行页单位的数据迁移。

[0410] <重新配置处理>

[0411] 重新配置处理可以在某层的使用比例超过该层的层阈值的情况下来进行,也可以定期性(例如每次到达“重新配置执行时刻”的值所表示的时刻时)进行。

[0412] 图 28 表示重新配置处理的流程。需要说明的是,图 28 是针对一个 VVOL(图 28 的说明中称作“对象 VVOL”)的流程。

[0413] (S2801)

[0414] 重新配置程序 415 判断是否针对对象 VVOL 执行数据迁移。具体进行如下处理。

[0415] * 程序 415 判断 VVOL 管理表 404(图 8)中与对象 VVOL 对应的“自动执行”是否为“ON”。在与对象 VVOL 对应的“自动执行”的值没有设定的情况下,程序 415 判断池管理表 403(图 7)中与分配给对象 VVOL 的池(以下,在图 28 的说明中称作“对象池”)对应的“自动执行”是否为“ON”。

[0416] * 在与对象 VVOL 或者对象池对应的“自动执行”为“ON”的情况下,程序 415 判断当前时刻是否到达了与对象 VVOL 或者对象池对应的“重新配置执行时刻”的值所表示的时刻。在该判断的结果为肯定性的情况下,S2801 的判断的结果为肯定性,在该判断的结果为否定性的情况下,S2801 的判断的结果为否定性。

[0417] * 在与对象 VVOL 或者对象池对应的“自动执行”为“OFF”的情况下,如果程序 415 从管理者通过管理装置 105 针对对象 VVOL 接受到了数据迁移的指示,则 S2801 的判断的结果为肯定性。

[0418] 如果与对象 VVOL 或者对象池对应的“自动执行”为“OFF”(S2802:否),则程序 415 进行 S2806。

[0419] 如果与对象 VVOL 或者对象池对应的“自动执行”为“ON”(S2802:是),则程序 415 进行 S2807。

[0420] (S2806)

[0421] 程序 415 判断是否由管理者设定了层范围。在该判断的结果为否定性的情况下,在进行了 S2807 之后进行 S2808,在该判断的结果为否定性的情况下,不进行 S2807,直接进行 S2808。

[0422] (S2807)

[0423] 程序 415 进行层范围决定处理。通过该处理来决定层范围。作为决定层范围的方法,可考虑到多种。例如,可预先准备多种层范围,也可由程序 415 针对每个层从该多种层范围中选择任意的层范围。

[0424] (S2808)

[0425] 程序 415 进行层判定处理。层判定处理中,如上所述,生成迁移页表 411。

[0426] (S2809)

[0427] 程序 415 依照所生成的迁移页表 411 进行页单位的迁移处理。在页单位的数据迁移中,在对象池内从某层内的实际页向其它层内的实际页迁移数据要素。针对页单位的迁移处理,将在后面详细说明。

[0428] << 页单位的迁移处理 >>

[0429] 图 34 表示图 28 的 S2808 中的迁移页表生成处理的流程。

[0430] (S3401)

[0431] 程序 415 选择对象 VVOL 的最初的虚拟页。

[0432] (S3402)

[0433] 程序 415 基于分配表 407 判断是否对被选择的虚拟页分配了实际页。如果该判断的结果是否定性的,则进行 S3405,如果该判断的结果是肯定性的,则进行 S3403。

[0434] (S3403)

[0435] 程序 415 基于监视表 409 判断是否应该变更分配给被选择的虚拟页的实际页内的数据要素所应该存在的层。具体如下。

[0436] * 程序 415 根据表 402(图 6) 确定具有分配给被选择的虚拟页的实际页的层(在 S3403 的说明中称作“对象层”)。

[0437] * 程序 415 基于被选择的虚拟页的 I/O 频度(例如与对象 VVOL 对应的激活的监视表 409 所表示的合计 I/O 数、平均 I/O 数、以及最大 I/O 数中的至少一个)和与对象层对应的“层范围”的值判断是否应该将分配给被选择的虚拟页的实际页内的数据要素从对象层迁移到其它的层。例如,如果被选择的虚拟页的 I/O 频度未纳入被选择的虚拟页的与对象层对应的层范围中,则判断为应该迁移数据要素。

[0438] 如果该 S3403 的判断的结果为否定性的,则进行 S3405,如果该判断的结果为肯定性的,则进行 S3404。

[0439] (S3404)

[0440] 程序 415 基于被选择的虚拟页的 I/O 频度(例如与对象 VVOL 对应的激活的监视表 409 所表示的合计 I/O 数、平均 I/O 数、以及最大 I/O 数中的至少一个)和与具有对象层的池内的其它分配层对应的“层范围”的值决定迁移目的地层。例如,将与纳入有被选择的虚拟页的 I/O 频度的层范围对应的分配层(对象 VVOL 的分配层)决定为迁移目的地层。然后,程序 415 在迁移目的地页表 411 中登记被选择的虚拟页的 ID、对象层的 ID、被决定的迁移目的地层的 ID、迁移状态“未迁移”。

[0441] (S3405)

[0442] 程序 415 判断被选择的虚拟页是否是对象 VVOL 的末端的虚拟页。如果该判断的结果为否定性的,则进行 S3406,如果该判断的结果为肯定性的,则该迁移页表生成处理结束。

[0443] (S3406)

[0444] 程序 415 从对象 VVOL 中选择上一次被选择的虚拟页的下一个虚拟页。然后,针对该被选择的虚拟页进行 S3402 以后的处理。

[0445] 通过该迁移页表生成处理,能够生成用于页单位的迁移处理的移动页表 411。

[0446] 图 35 表示页单位的迁移处理的概要。

[0447] 根据页单位的迁移处理,针对表 411 中登记的每个虚拟页,分配给该虚拟页的实际页内的数据要素被迁移到与该虚拟页对应的迁移目的地层内的自由的实际页。以迁移目的地实际页取代迁移源实际页对该虚拟页进行分配。如图 35 所示,通过该迁移处理,既有被迁移到更上位的层中的数据要素,也有被迁移到更下位的层中的数据要素。

[0448] 根据针对哪个层设定怎样的层范围,能够在适合虚拟页的 I/O 频度的层中配置该虚拟页内的数据要素。

[0449] 例如,假设对越上位的层设定越高的层范围,对越下位的层设定越低的层范围。具体地,假设例如层 #11 是高速且高信赖性的层(例如 RVOL 种类为“SSD”的层),并被设定为

高的层范围。另外,假设例如层 #13 是低速且低信赖性的层(例如 RVOL 种类为“HDD-SATA”的层),并被设定为低的层范围。另外,假设例如层 #12 是 VVOL#50 的默认层,并被设定为中程度的层范围。该情况下,针对构成 VVOL#50 的虚拟页的写入对象数据要素暂时被写入默认层 #12 内的实际页。然后,在例如 VVOL#50 内的某个虚拟页 #A 的 I/O 频度成为纳入在高的层范围内程度的高 I/O 频度的情况下,分配给该虚拟页 #A 的实际页 #X 内的数据要素被从层 #12 迁移到层 #11。另一方面,在例如 VVOL#50 内的某个虚拟页 #B 的 I/O 频度成为纳入在低的层范围内程度的低 I/O 频度的情况下,分配给该虚拟页 #B 的实际页 #Y 内的数据要素被从层 #12 迁移到层 #13。

[0450] 这样,通过对哪个层设定怎样的层范围、每个虚拟页的 I/O 频度的监视结果、页单位的数据迁移,能够将 VVOL 内的数据要素配置在适当的层(物理存储资源)中。

[0451] 图 36 表示页单位的迁移处理的流程。

[0452] (S3601)

[0453] 重新配置程序 415 将分配给对象 VVOL 的池的“池状态”更新为“重新配置中”。

[0454] (S3602)

[0455] 程序 415 选择所生成的迁移页表 411 的最初的记录。

[0456] (S3603)

[0457] 程序 415 将所选择的记录中的“迁移状态”的值更新为“迁移中”。

[0458] (S3604)

[0459] 程序 415 从所选择的记录所表示的迁移目的地层中选择自由的实际页,向所选择的实际页(迁移目的地实际页)迁移存储在迁移源层内的迁移源实际页中数据要素。迁移源实际页是分配给对象虚拟页的实际页。

[0460] (S3605)

[0461] 程序 415 更新表。具体如下。

[0462] * 程序 415 更新分配表 407,从而以迁移目的地实际页取代迁移源实际页分配给被选择的记录所表示的虚拟页(以下,在 S3605 的说明中称作“选择虚拟页”)。即,程序 415 使迁移目的地实际页的 ID、迁移目的地层的 ID 与选择虚拟页相对应。

[0463] * 程序 415 更新选择虚拟页的 I/O 频度(例如激活的表 409)。具体地,例如,选择虚拟页或者对象 VVOL 的合计 I/O 数增加 1。

[0464] * 程序 415 更新表 402(图 6)中的、迁移目的地层的“使用比例”的值。

[0465] (S3606)

[0466] 程序 415 将被选择的记录中的“迁移状态”的值更新为“已迁移”。

[0467] (S3607)

[0468] 程序 415 判断被选择的记录是否是虚拟页表 411 中末端的记录。如果该判断的结果是否定性的,则进行 S3608,如果该判断的结果是肯定性的,则该迁移处理结束。

[0469] (S3608)

[0470] 程序 415 从虚拟页表 411 选择上一次被选择的记录的下一个记录。然后,针对该被选择的记录,进行 S3603 以后的处理。

[0471] 以上是页单位的迁移处理的流程。而且,作为页单位的迁移处理的变形例,可以考虑到例如下述多个变形例中的至少一个。

[0472] * 在 S3602、S3603 之间,可以假设数据要素已经迁移到了迁移目的地层中,程序 415 判断迁移目的地层的使用比例是否超出了迁移目的地层的层阈值。在判断为超出的情况下,程序 415 可以跳过 S3603 ~ S3606 直接进行 S3607(即,程序 415 针对被选择的记录可以不进行数据要素的迁移)。或者,在判断为超出的情况下,程序 415 可以通过针对以迁移目的地层为分配层的 VVOL 进行重新配置处理,在迁移目的地层中增加自由的实际页,然后开始 S3603。

[0473] * 可以在通过 S3605 更新迁移目的地层的使用比例时由程序 415 判断迁移目的地层的使用比例是否超出了迁移目的地层的层阈值。在判断为超出的情况下,程序 415 可以在管理装置 107 中显示警告,以及 / 或者,通过针对以迁移目的地层为分配层的 VVOL 进行重新配置处理,在迁移目的地层中增加自由的实际页。

[0474] < 层阈值的变更 >

[0475] UI 控制程序 413 能够从管理者接受层阈值的变更。

[0476] 图 37 表示层阈值的变更处理的流程。

[0477] (S3701)

[0478] 程序 413 接受池 ID 以及层 ID、针对该层 ID 的变更后的层阈值。

[0479] (S3702)

[0480] 程序 413 判断与通过 S3701 接受到的池 ID 以及层 ID 对应的“使用比例”的值是否超出了变更后的层阈值。

[0481] (S3703)

[0482] 程序 413 在 S3702 的判断的结果是肯定性的情况下在管理装置 107 中显示错误。

[0483] < 使用比例的监视 >

[0484] 如上所述, I/O 控制程序 414 或者重新配置程序 415 在向实际页写入数据要素的情况下更新表 402(图 6) 中的、具有该实际页的层的“使用比例”的值。然后,程序 414 或者 415 可以判断“使用比例”的更新后的值是否超出了与该层对应的“层阈值”的值。如果该判断的结果是肯定性的,则程序 414 或者 415 可以进行图 38 所示的阈值超出应对处理。

[0485] 图 38 表示阈值超出应对处理的流程。

[0486] 程序 414 或者 415 判断具有使用比例超出了层阈值的层的池的“池状态”是否为“重新配置中”(S3801)。

[0487] 在 S3801 的判断的结果是肯定性的情况下,程序 414 或者 415 向管理装置 107 或者主机 101 发出警告(S3802)。该警告例如可以包括使用比例超出了层阈值的层的 ID、具有该层的池的 ID、该池的“池状态”的值(“重新配置中”)。

[0488] 在 S3801 的判断的结果是否定性的情况下,程序 414 或者 415 向管理装置 107 或者主机 101 发出警告(S3803),执行使该层的使用比例成为该层的层阈值以下的重新配置处理(图 28)(S3804)。而且,S3803 中的警告例如可以包括使用比例超出层阈值的层的 ID、具有该层的池的 ID、意味着针对该池进行重新配置处理的消息。

[0489] < 数据要素的汇总 >

[0490] 本实施例中,可以进行数据要素的汇总处理。

[0491] 图 21 表示数据要素的汇总处理的概要。

[0492] I/O 控制程序 414 或者重新配置程序 415 在针对 VVOL#39 定期性或者在将数据要

素写入实际页时（例如，页单位的迁移处理结束时）检查 VVOL#39 内的数据要素的配置。具体地，例如，程序 414 或者 415 针对 VVOL#39 的各分配层 #00、##01、#03 计算存在 VVOL#39 内所有数据要素中多大程度的比例的数据要素（即保存比例）。程序 414 或者 415 判断 VVOL#39 的分配层 #00、##01、#03 中是否存在保存比例超出保存阈值的分配层。

[0493] 在存在保存比例超出保存阈值的保存比例的分配层的情况下，程序 414 或者 415 将该分配层 #00 以外的分配层 #01 以及 #03 内的数据要素（VVOL#39 内的数据要素）迁移到分配层 #00（具体地，例如分配层 #00 的自由的实际页）。

[0494] 而且，在之后，例如程序 414 可以在针对 VVOL#39 内的虚拟页（尤其是未分配有实际页的虚拟页）发生写入的情况下，必定从分配层 #00 向该虚拟页分配实际页。

[0495] 而且，如果将分配层 #01 以及 #03 内的数据要素（VVOL#39 内的数据要素）迁移到分配层 #00 后导致分配层 #00 的使用比例超出分配层 #00 的层阈值，则程序 414 或者 415 可以进行下述的 (x) 或者 (y)。

[0496] (x) 程序 414 不将分配层 #01 以及 #03 内的数据要素迁移到分配层 #00。

[0497] (y) 程序 414 通过将对象 VVOL 以外的 VVOL 内的、分配层 #00 内的数据要素迁移到其它的层中，来在分配层 #00 中增加自由的实际页。

[0498] 另外，本实施例中，能够在图 28 的 S2807 中自动决定层范围。层范围例如通过由 MP 执行未图示的层范围决定程序来决定。

[0499] 以下，以一个池（在层范围的自动决定的说明中称作“对象池”）为例说明层范围决定处理。

[0500] 图 29 表示层范围决定处理的概要。

[0501] 设对象池具有层 #01、#02 以及 #03 这三个层。设层 #01 为最上位，层 #02 为第二上位，层 #03 为最下位的层。换言之，设层 #01 的性能潜力 (potential) 最高，层 #02 的性能潜力第二高，层 #03 的性能潜力最低。“性能潜力”是指重新配置处理中的数据要素配置权重。关于性能潜力，将在后面详细叙述。

[0502] 层范围决定程序通过汇总分配有对象池的所有 VVOL 的度数分布来取得针对对象池的度数分布 4300。度数分布的横轴为 I/O 频度（负荷），纵轴为虚拟页数。

[0503] 层范围决定程序从上位层到下位层依次（或者从下位层到上位层依次）决定层范围。以下，设以层 #01、#02、#03 的顺序决定层范围。以下，叙述用于决定针对各层的层范围所进行的处理（而且，图 29 中，层范围的上限为菱形的标记，层范围的下限为圆形标记）。

[0504] 《层 #01 的层范围 1 的决定》

[0505] * 层范围决定程序将层 #01 的层范围 1 的上限决定为无限大。即，最上位的层的层范围的上限设定为无限大。

[0506] * 层范围决定程序在度数分布 4300 中如箭头 4301 所示按照 I/O 频度从最高到较低累计虚拟页数。该累计值达到层 #01 的性能潜力的情况下，层范围决定程序将与即将到达层 #01 的性能潜力之前相加的虚拟页数对应的 I/O 频度设定为层 (0-1) 边界值（层 #00 与层 #01 的边界值）。

[0507] * 层范围决定程序将比层 (0-1) 边界值低某种程度容限 (margin) 的 I/O 频度设定为层 #01 的层范围 1 的下限。具体地，例如，该下限是低于层 (0-1) 边界值预定比例的 I/O 频度（例如低 10% 的 I/O 频度）。

[0508] *层范围决定程序假定 I/O 频度纳入层范围 1 的所有虚拟页内的数据要素被迁移到了层 #01,判断层 #01 的使用比例是否超出层 #01 的层阈值。在该判断的结果为肯定性的情况下,层范围决定程序将层范围 1 修正为层 #01 的使用比例不超出层 #01 的层阈值的大小。

[0509] 《层 #02 的层范围 2 的决定》

[0510] *层范围决定程序将比层 (0-1) 边界值高某种程度容限 (margin) 的 I/O 频度设定为层 #02 的层范围 1 的上限。具体地,例如,该上限是高于层 (0-1) 边界值预定比例的 I/O 频度 (例如高 10% 的 I/O 频度)。

[0511] *层范围决定程序在度数分布 4300 中从层 (0-1) 边界值到较低累计虚拟页数。该累计值达到层 #02 的性能潜力的情况下,层范围决定程序将与即将到达层 #02 的性能潜力之前相加的虚拟页数对应的 I/O 频度设定为层 (1-2) 边界值 (层 #01 与层 #02 的边界值)。

[0512] *层范围决定程序将比层 (1-2) 边界值低某种程度容限 (margin) 的 I/O 频度设定为层 #02 的层范围 2 的下限。

[0513] *层范围决定程序假定 I/O 频度纳入层范围 2 的所有虚拟页内的数据要素被迁移到了层 #02,判断层 #02 的使用比例是否超出层 #02 的层阈值。在该判断的结果为肯定性的情况下,层范围决定程序将层范围 2 修正为层 #02 的使用比例不超出层 #02 的层阈值的大小。

[0514] 《层 #03 的层范围 3 的决定》

[0515] *层范围决定程序将比层 (1-2) 边界值高某种程度容限 (margin) 的 I/O 频度设定为层 #03 的层范围 3 的上限。

[0516] *层范围决定程序将层 #02 的层范围 2 的下限设定为 0。即,最下位层的层范围的下限设定为 0。

[0517] *层范围决定程序假定 I/O 频度纳入层范围 3 的所有虚拟页内的数据要素被迁移到了层 #03,判断层 #03 的使用比例是否超出层 #03 的层阈值。在该判断的结果为肯定性的情况下,层范围决定程序将层范围 3 修正为层 #03 的使用比例不超出层 #03 的层阈值的大小。

[0518] 如上决定层范围 1 ~ 3。而且,之所以在层范围的上限以及 / 或者下限赋予层边界值的容限,是为了不随着 I/O 频度的变动频繁地发生数据要素的迁移。具体地,例如,如果没有容限、并且某虚拟页的 I/O 频度频繁超出或者低于层 (0-1) 边界值,则该虚拟页内的数据要素被频繁地迁移到层 #01 或者层 #02。但是,如果如上所述设置了容限,则能够减少进行该种迁移的频度。

[0519] 以下,进一步详细说明层范围决定处理。

[0520] 图 30 表示层范围决定处理的流程。

[0521] (S4401)

[0522] 层范围决定程序针对对象池的各层计算性能潜力。以一个层 (在 S4401 的说明中称作“对象层”) 为例,性能潜力例如通过下述 (式 1) 来计算。

[0523] 性能潜力 = 性能等级 × 实效实际页数 (实效容量) … (式 1)

[0524] 该 (式 1) 意味着即便性能等级低,只要执行实际页数多,性能潜力就高。

[0525] 而且,“实效实际页数”是指构成对象层的所有实际页中的、在重新配置处理中可

成为数据要素迁移目的地的实际页的总数。换言之,为了在池中写入重新配置处理中来自主机的写入对象数据要素,可以设置在重新配置处理中不会作为迁移目的地的实际页。

[0526] “性能等级”是指对象层的每单位容量(例如一个实际页)的性能潜力。性能等级例如基于通过下述(式2)计算的值K。

[0527] 值 $K = \text{RG 性能} / \text{RG 实效容量} \cdots$ (式2)

[0528] (式2)中的“RG性能”例如是指读出权重与写入权重的比为1:1时的RG性能,是成为对象层的根基的RG在每单位时间能够处理的最大I/O数。(式2)中的“RG实效容量”例如是指针对成为对象层的根基的RG的、保存数据要素的实际页的总数。而且,性能等级可以取代上述值K或者在其基础上参酌读出权重与写入权重的比。该比例例如可以基于来自主机的I/O的状况、高速缓存的状况(例如,脏数据要素与清洁数据要素(已保存在实际页中的数据要素)的分配)、RG的RAID等级中的至少一个来进行变更。另外,性能等级可以由管理者通过手动来设定。

[0529] (S4402)

[0530] 层范围决定程序进行层边界值计算处理。即,层范围决定程序基于针对对象池的各层计算出的性能潜力来计算层边界值。

[0531] (S4403)

[0532] 层范围决定程序基于计算出的各层边界值针对各层计算层范围。

[0533] (S4404)

[0534] 层范围决定程序针对各层判断层范围是否不合适。以一个层(以下,在S4403以及S4405的说明中称作“对象层”)为例具体而言,例如,层范围决定程序假定I/O频度纳入对象层的层范围的所有虚拟页内的数据要素被迁移到了对象层,判断对象层的使用比例是否超出对象层的层阈值。

[0535] (S4405)

[0536] 在S4404的判断的结果为肯定性的情况下,层范围决定程序或者在管理装置107中显示警告,或者将层范围修正为对象层的使用比例不超出对象层的层阈值的大小。

[0537] 图31表示图30的S4402中层边界值计算处理的流程。

[0538] (S4501)

[0539] 层范围决定程序通过汇总分配有对象池的所有VVOL的度数分布来取得对象池的度数分布。度数分布能够使用与分配有对象池的各VVOL对应的监视表409来取得。

[0540] (S4502)

[0541] 层范围决定程序计算对象池的I/O频度合计。具体地,层范围决定程序合计分配有对象池的所有VVOL的所有虚拟页的I/O频度。

[0542] (S4503)

[0543] 层范围决定程序基于通过S4502计算出的I/O频度合计计算针对对象池的各层的分配I/O频度。具体地,层范围决定程序将通过S4502计算出的I/O频度合计分配给对象池的多个层。具体地,例如,层范围决定程序基于通过图30的S4401计算出的、多个层的性能潜力的比,来对该多个层分配I/O频度合计。由此,例如在性能潜力的比为层#01:层#02:层#03=4:2:1、并且I/O频度合计为7000的情况下,层#01的分配I/O频度成为4000($7000 \times (4 / (4+2+1))$),层#02的分配I/O频度成为2000($7000 \times (2 / (4+2+1))$),层

#01 的分配 I/O 频度成为 $1000(7000 \times (1/(4+2+1)))$ 。

[0544] (S4504)

[0545] 层范围决定程序基于对象池的各层的分配 I/O 频度来决定层边界值。例如,以 S4503 的例子而言,层 (0-1) 边界值成为 2000,层 (1-2) 边界值成为 1000。

[0546] 根据以上的层范围决定处理,能够基于各层的性能潜力和对象池的度数分布针对每个层设定对应于对象池整体的 I/O 状况的合适的层范围。

[0547] 以上是针对实施例 1 的说明。

[0548] 而且,在实施例 1 中,例如针对 VVOL,默认层可以必定是 VVOL 的分配层中 I/O 性能(例如 I/O 速度或者响应时间)最大的层(例如分配层中最上位的层)。这是为了尽量增大 VVOL 的 I/O 性能。该情况下,写入 VVOL 的数据要素可以配置在尽量上位的分配层。但是,控制为分配层的使用容量不超出该分配层的层阈值。

[0549] 另外,例如,可以与各池、或者多个池共通地设置退避用的层。无论是依照来自主机的写入命令的写入,还是重新配置处理中的迁移处理,通常来讲,程序 414 以及 415 都不向退避用的层写入数据要素。但是,程序 414 或者 415 可以在例如数据要素写入目的地的层(以下,在该段落的说明中称作“对象层”)的使用比例超出该层的层阈值的情况下,将对象层内的任意实际页内的数据要素(或者分配给对象层的层范围内未纳入有 I/O 频度的虚拟页的、对象层中实际页内的数据要素)迁移到退避用的层,由此在对象层中增加自由的实际页。该情况下,程序 414 或者 415 对分配给对象层内的迁移源实际页的虚拟页分配退避用的层内的迁移目的地实际页来取代该迁移源实际页。然后,程序 414 或者 415 可以向设定为纳入有分配给退避用层内实际页的虚拟页的 I/O 频度的层范围的层迁移保存在该退避用层内的实际页中的数据要素。退避用层例如可以为由与最上位的层种类相同的 RVOL 构成的层。

[0550] 另外,可以不设定默认层。针对 VVOL 的来自主机 101 的写入对象数据要素可以写入尽量上位的层中。在上位的层的使用比例超出了层阈值的情况下,数据要素可以写入比该上位的层低一级的下位的层中。

[0551] 另外,监视表 409B 以及 409A 中,409A 可以不存在。该情况下,可以通过合计监视表 409B 所具有的每个虚拟页的 I/O 频度来计算 VVOL 的 I/O 频度。

[0552] 另外,页单位的迁移处理可以中断。例如,重新配置程序 415 能够在从管理者接受到中断指示的情况下,在针对迁移状态“迁移中”将处理进行到使迁移状态为“已迁移”之后,中断页单位的迁移处理。程序 415 可以在例如接受到中断指示后一定时间后或者接受到重新开始的指示时,基于迁移页表 411 重新开始迁移处理。或者,程序 415 可以在接受到中断指示的情况下废弃迁移页表 411,在下一次迁移处理时,基于激活的监视表 409 来生成迁移页表 411。

[0553] 另外,在本实施例中,Thin Provisioning 的 VVOL203 是根据从主机 101 向具有虚拟页的地址的写入请求分配到实际页的,但是本发明并不限于该情况,还可以优选列举以下的例子。这是因为即便是在这些例子的情况下也能够实现实际使用容量的削减。

[0554] 第一个例子如下。即, VVOL203 的多个虚拟页中一部分的虚拟页或者 VVOL203 的各虚拟页中包含的一部分区域在接收从主机 101 向具有虚拟页的地址的写入请求之前预先分配有实际区域(预分配),根据从主机 101 向具有虚拟页的地址的写入请求,在已经分

配完毕的区域不充足,还需要进一步分配的情况下,分配追加的实际区域。

[0555] 第二个例子如下。即, VVOL203 的多个虚拟页在接收从主机 101 向具有虚拟页的地址的写入请求之前预先暂时分配给保存有零数据的实际区域或者控制用的区域(这些区域与 SSD 等存储设备上的区域对应)(暂时分配),根据从主机 101 向具有虚拟页的地址的写入请求,分配目的地从暂时的分配区域变更为应该保存数据的实际页(正式分配)。该情况下,在实施正式分配之前接收到从主机 101 向具有虚拟页的地址的写入请求的情况下,向主机 101 响应保存在暂时分配目的地实际区域或者控制用区域中的零数据。

[0556] 另外,在本实施例中,为了一元管理多个池 201 而通过一个表来管理多个池 201,但是本发明并不限于此,也可以为了针对各个池 201 一元管理实际页 207,优选对每个池 201 分别使用管理用的表(可以是表形式以外形式的信息)。

[0557] 另外,层 303 的种类不限于按照本实施例中记载的 SSD、HSS-SAS、HDD-SATA 等各属性进行分类,还优选按照具有 FC 接口的 HDD、非 HDD 形式的存储设备、例如闪存等半导体存储介质构成的设备等各属性进行分类。该情况下,与层 303 相关联的虚拟页的属性或者各层 303 中所含各实际页的属性也不限于按照本实施例中记载的 SSD、HDD-SAS、HDD-SATA 等各属性进行分类,而是按照具有 FC 接口的 HDD、非 HDD 形式的存储设备、例如闪存等半导体存储介质构成的设备等各属性进行分类。

[0558] 并且,层 303 的种类不限于仅通过上述存储设备的类型这一属性来进行分类,还优选按照将上述存储设备的类型和 RAID1 到 RAID5 等的 RAID 类型组合后的属性、或者上述存储设备的类型与上述存储设备的性能(存取速度等)和存储容量与 RAID1 到 RAID5 等的 RAID 类型组合后的各属性进行分类。这些情况下,与各层 303 相关联的虚拟页的属性或者各层 303 中所含各实际页的属性也不限于仅通过上述存储设备的类型这一属性来进行分类,还优选按照将上述存储设备的类型和 RAID1 到 RAID5 等的 RAID 类型组合后的属性、或者上述存储设备的类型与上述存储设备的性能(存取速度等)和存储容量与 RAID1 到 RAID5 等的 RAID 类型组合后的各属性进行分类。

[0559] 实施例 2

[0560] 以下说明本发明的实施例 2。此时,主要说明与实施例 1 的区别点,关于与实施例 1 的共通点,则省略或者简略说明(这一点,后面的实施例 3 以后也同样)。

[0561] 本发明的实施例 2 中,取代 I/O 频度或者在其基础上,基于每个 VVOL 或者每个虚拟页的最终 I/O 时刻进行页单位的迁移处理。如实施例 1 中所说明的那样,在数据要素写入实际页的情况下,最终 I/O 时刻登记在监视表 409 中。重新配置程序取代虚拟页的 I/O 频度和具有该虚拟页内的数据要素的层的层范围的关系或者在其基础上,基于虚拟页的最终 I/O 时刻决定在迁移页表 411 中登记的虚拟页。例如,可将最终 I/O 时刻是当前时刻之前预定时间以上时刻的虚拟页登记在迁移页表 411 中。届时,该虚拟页内的数据要素的迁移目的地可以是比具有该虚拟页内的数据要素的层更下位的层。之所以这样,是由于考虑到最终 I/O 时刻久远的虚拟页内的数据要素今后被输入输出的可能性也低,优选配置在较下位的层中。

[0562] 实施例 3

[0563] 本发明的实施例 3 中,基于以顺序存取和随机存储之一的存取模式进行 I/O,来进行页的单位迁移处理。

[0564] 如果 I/O 对象数据是被顺序存取的数据,则在作为 I/O 目的地 RVOL 的根基的 PDEV 为 HDD 的情况下, HDD 的搜索时间很短。该情况下,即便是性能较低 HDD 中有 I/O 对象数据,也由于被顺序存取,而可以考虑为性能下降的程度小。另外,通过将顺序存取的数据迁移到廉价的性能低的 PDEV,由于被随机存取的其它数据,能够使更多的基于性能高的 PDEV 的实际页为自由的页。因此,能够期待提高存储系统整体的 I/O 性能。

[0565] 在实施例 3 中,例如在 I/O 控制程序 414 将数据要素写入实际页中的情况下,在例如分配表 407 中登记针对 I/O 目的地虚拟页的存取模式(是顺序存取还是随机存取)。分配表 407 中可以针对每个虚拟页写入每种存取模式的执行频度。

[0566] 例如在生成迁移页表 411 时,如果针对虚拟页的存取模式为顺序存取(或者如果顺序存取的执行频度高),则重新配置程序 415 可以选择较下位的层作为与该虚拟页对应的迁移目的地。另外,例如,即使虚拟页的 I/O 频度高于具有分配给该虚拟页的实际页的层的层范围的最大值,重新配置程序 415 可以不将该实际页内的数据要素作为迁移对象。

[0567] 实施例 4

[0568] 本发明的实施例 4 中,两个监视表 409 以预定的定时不可编辑地保存在 CMPK119 或者其它的存储资源中来取代交替使用。例如,表示星期一的监视结果的监视表 409 可作为星期一的表 409 保存,表示星期六的监视结果的监视表 409 可作为星期六的表 409 与星期一的表 409 相区分地保存。由此,增加了运用重新配置处理的运用方法的选择项。

[0569] 在实施例 4 中,例如重新配置程序 415 在重新配置即将开始前通过询问管理者等方法决定是否保存激活的监视表 409。在进行保存的情况下,在即将进行重新配置处理之前,程序 415 将监视表 409 保存在管理装置 107 或者其它的存储资源中。程序 415 能够从保存的多个监视表 409 中选择为了生成迁移页表 411 而参照的监视表 409,使用所选择的监视表 409 来生成迁移页表 411。

[0570] 实施例 5

[0571] 本发明的实施例 5 中,作为 I/O 频度与层范围的比较方法,能够选择设定绝对值和相对值比较之一。

[0572] “绝对值比较”是指,针对页单位的迁移处理而言,将虚拟页的 I/O 频度(例如,合计 I/O 数、平均 I/O 数、或者最大 I/O 数)直接与层范围进行比较,而不考虑具有该虚拟页的 VVOL 内其它虚拟页的 I/O 频度。根据绝对值比较,由于没有必要考虑其它虚拟页或者 VVOL 的 I/O 频度,因此能够缩短比较所需时间,从而能够期待缩短重新配置处理所需时间。

[0573] “相对值比较”是指,针对页单位的迁移处理而言,将虚拟页的 I/O 频度(例如,合计 I/O 数、平均 I/O 数、或者最大 I/O 数)基于具有该虚拟页的 VVOL 内其它虚拟页的 I/O 频度进行修正(正规化),将修正后的 I/O 频度与层范围进行比较。具体地,例如,即便某虚拟页或者 VVOL 的 I/O 频度低,只要其它虚拟页或者其它 VVOL 的 I/O 频度更低,则整体上修正为上述某虚拟页或者 VVOL 的 I/O 频度并不太低。或者,例如,即便某虚拟页或者 VVOL 的 I/O 频度高,只要其它虚拟页或者其它 VVOL 的 I/O 频度更高,则整体上修正为上述某虚拟页或者 VVOL 的 I/O 频度并不太高。根据相对值比较,能够期待对应于监视时间带中整体的 I/O 状况的最佳配置。

[0574] 另外,在本实施例中,作为比较对象的 I/O 频度,例如能够选择设定是平均 I/O 数还是最大 I/O 数。根据平均 I/O 数,能够期待对应于监视时间带中整体的 I/O 频度来配置

数据要素。根据最大 I/O 数,即便 I/O 频度暂时增大,也能够期待不使 I/O 性能下降的数据要素配置。

[0575] 使比较方法是绝对值比较还是相对值比较以及使比较对象是平均 I/O 数还是最大 I/O 数,例如能够由管理者通过上述的池定义 GUI 或者 VVOL 生成 GUI 来设定。针对池的比较方法以及比较对象例如图 20 所示登记在池管理表 4403 中,针对 VVOL 的比较方法以及比较对象例如图 20 所示登记在 VVOL 管理表 4404 中(图 32 以及图 20 中,省略了图 7 或者图 8 中所示的栏(column),但是在实施例 2 中,该栏是存在的)。而且,在针对 VVOL 与分配给其他的池的比较方法或者比较对象不同的情况下,可以与实施例 1 同样地优先针对 VVOL 的设定。

[0576] 图 33 表示比较方法 / 比较对象的设定处理的流程。

[0577] 比较方法 / 比较对象的设定处理例如可以在池生成处理以及 / 或者 VVOL 生成处理中进行。

[0578] UI 控制程序 413 在针对某池或者某 VVOL(图 33 的说明中称作“对象池或者对象 VVOL”)接受到相对值比较指定的情况下(S4201:否),针对对象池或者对象 VVOL 在表 4403 或者 4404 中设定比较方法为“相对值比较”(S4202)。

[0579] 程序 413 在针对对象池或者对象 VVOL 接受到绝对值比较指定的情况下(S4201:是),针对对象池或者对象 VVOL 在表 4403 或者 4404 中设定比较方法为“绝对值比较”(S4203)。

[0580] 程序 413 在针对对象池或者对象 VVOL 接受到最大 I/O 数指定的情况下(S4204:否),针对对象池或者对象 VVOL 在表 4403 或者 4404 中设定比较对象为“最大 I/O 数”(S4205)。

[0581] 程序 413 在针对对象池或者对象 VVOL 接受到平均 I/O 数指定的情况下(S4204:是),针对对象池或者对象 VVOL 在表 4403 或者 4404 中设定比较对象为“平均 I/O 数”(S4206)。

[0582] 而且,比较方法以及 / 或者比较对象可以自动设定或者变更,来取代由管理者手动设定。例如,程序 414 或者 415 基于监视表 409B 检查 VVOL 或者池的 I/O 的状况,根据所检查到的 I/O 的状况,可以将表 4403 以及 / 或者 4403 中已经设定的“比较方法”以及 / 或者“比较对象”的值进行变更。例如,在检测到一个 VVOL 中所有虚拟页的 I/O 频度(或者多个 VVOL 的 I/O 频度)的偏差未纳入在预定范围的情况下,“比较方法”从“相对值比较”变更为“绝对值比较”,反之,在检测到该偏差纳入在预定范围的情况下,可以将“比较方法”从“绝对值比较”变更为“相对值比较”。另外,例如,在检测到 I/O 频度暂时激增的情况下,“比较对象”从“平均 I/O 数”变更为“最大 I/O 数”,在检测到 I/O 频度的偏差纳入在预定范围的情况下,可以将“比较对象”从“最大 I/O 数”变更为“平均 I/O 数”。

[0583] 以上说明了本发明的若干个实施例,当然,本发明不限于这些实施例,在不脱离其主旨的范围内能够实现各种变更。

[0584] 例如,至少一个 PVOL 可以为虚拟的 PVOL,成为 PVOL 的根基的 PDEV 可依存在于其它的存储系统。该情况下,写入构成 PVOL 的实际页中的数据要素可以写入成为该 PVOL 的根基的、其它的存储系统内的 PDEV。

[0585] 另外,例如,无论是否对实际页进行了 I/O,监视表 409 可以在每次产生对虚拟页

的 I/O 时被更新。

[0586] 另外,例如,在页单位的迁移处理中,针对迁移目的地实际页的分配目的地虚拟页或者具有该虚拟页的 VVOL,可以在向迁移目的地实际页迁移数据时不更新 I/O 频度以及最终 I/O 时刻。迁移后,可以在根据针对该虚拟页或者 VVOL 的 I/O 对该虚拟页或者分配给 VVOL 的实际页进行 I/O 时,更新该虚拟页或者 VVOL 的 I/O 频度以及最终 I/O 时刻。

[0587] 另外,例如,RVOL 的种类可以取代构成作为 RVOL 的根基的 RG 的 PDEV 的种类或者在其基础上,基于该 RG 的 RAID 等级 (RAID1、RAID5 (3D+1P) 等)。因此,例如,即使分别构成第一 RG 和第二 RG 的 PDEV 的种类相同,只要第一 RG 和第二 RG 的 RAID 等级不同,基于第一 RG 的第一 RVOL 和基于第二 RG 的第二 RVOL 的种类就不同,因此,第一 RVOL 和第二 RVOL 可以属于不同的层。层的定义不限于 RVOL 种类,也可以从其它观点来定义。

[0588] 另外,例如一个实际页的容量可以是能够对一个虚拟页分配多个实际页这样的容量。该情况下,页单位的迁移例如可以是数据要素从分配给虚拟页的多个迁移源实际页分别迁移到多个迁移目的地实际页。

[0589] 另外,例如,例如一个实际页的容量可以是一个实际页能够分配给多个虚拟页这样的容量。该情况下,在页单位的迁移处理中,选择多个虚拟页,针对该多个虚拟页,从一个迁移源实际页向迁移目的地实际页迁移数据要素。

[0590] 另外,池 VOL 可以是与外部的存储系统所提供的 VOL 相对应的虚拟的逻辑卷。该情况下,写入池 VOL 的实际页的数据要素写入与该池 VOL 对应的、外部的存储系统的 VOL。

[0591] 另外,I/O 频度的单位可以是例如 IOPS (每一秒的 I/O 数)。

[0592] 另外,度数分布的 I/O 频度幅度可以是 1 (即,可针对每个 I/O 频度计数虚拟页数)。

[0593] 符号说明

[0594] 103... 存储系统

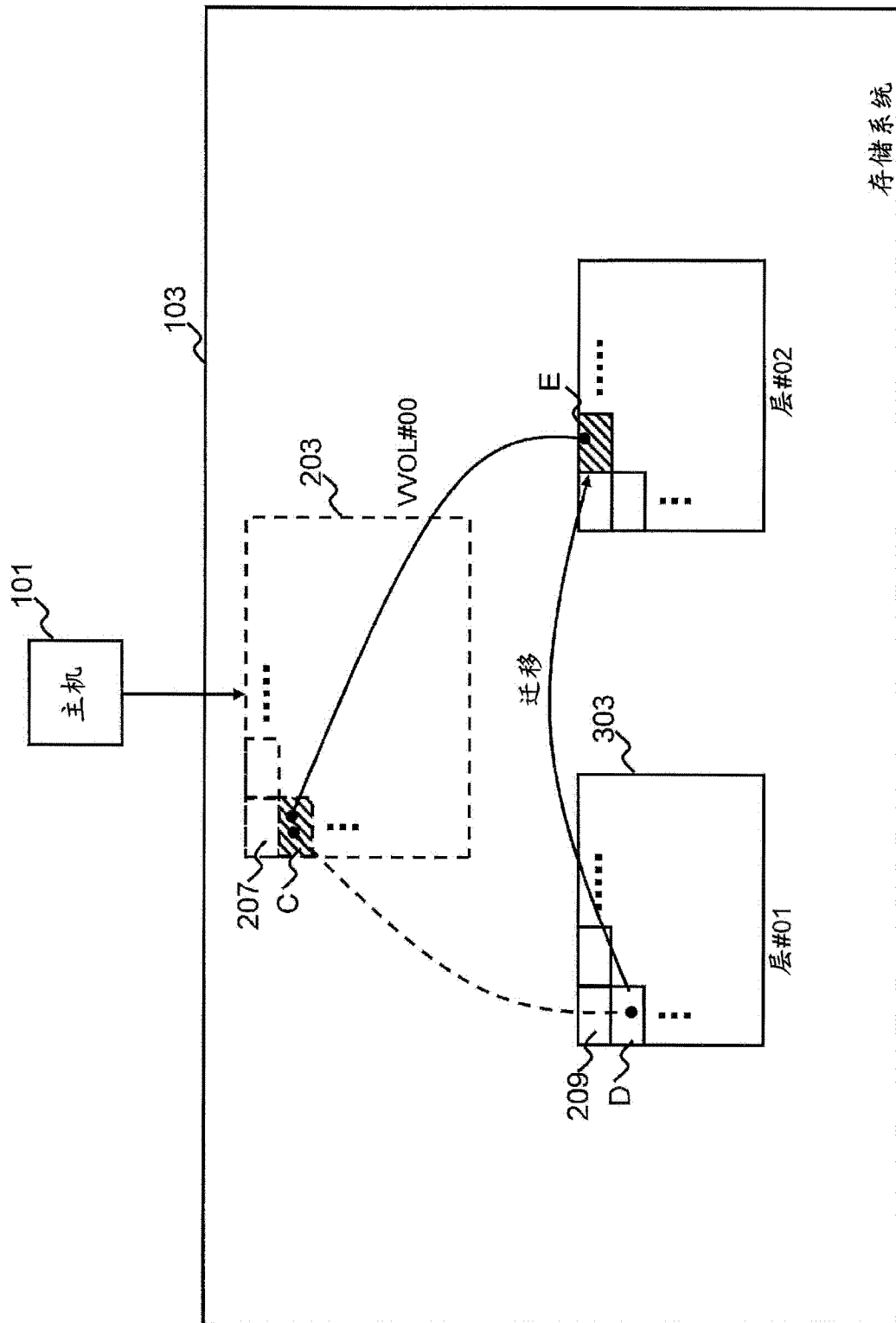


图 1

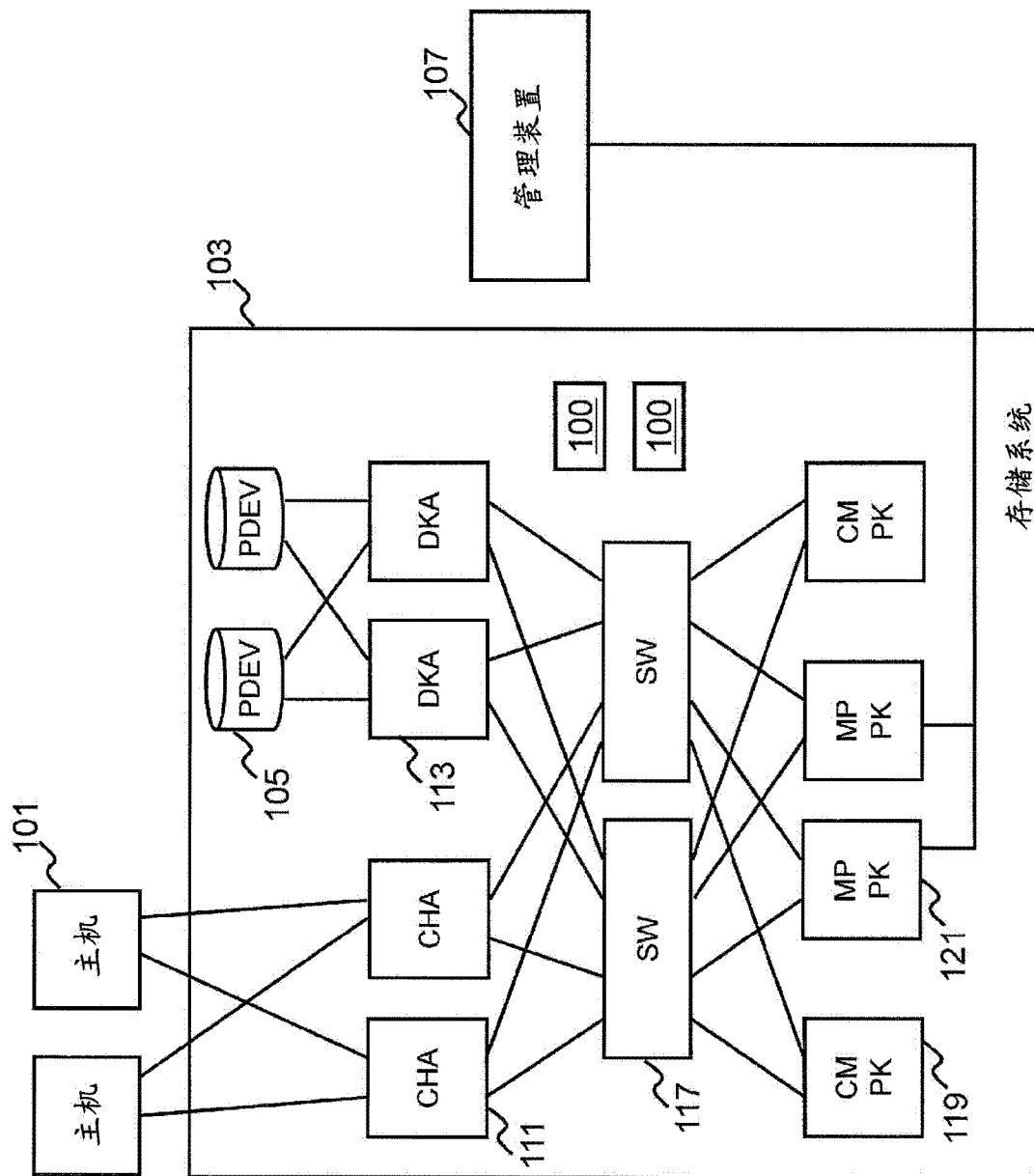


图 2

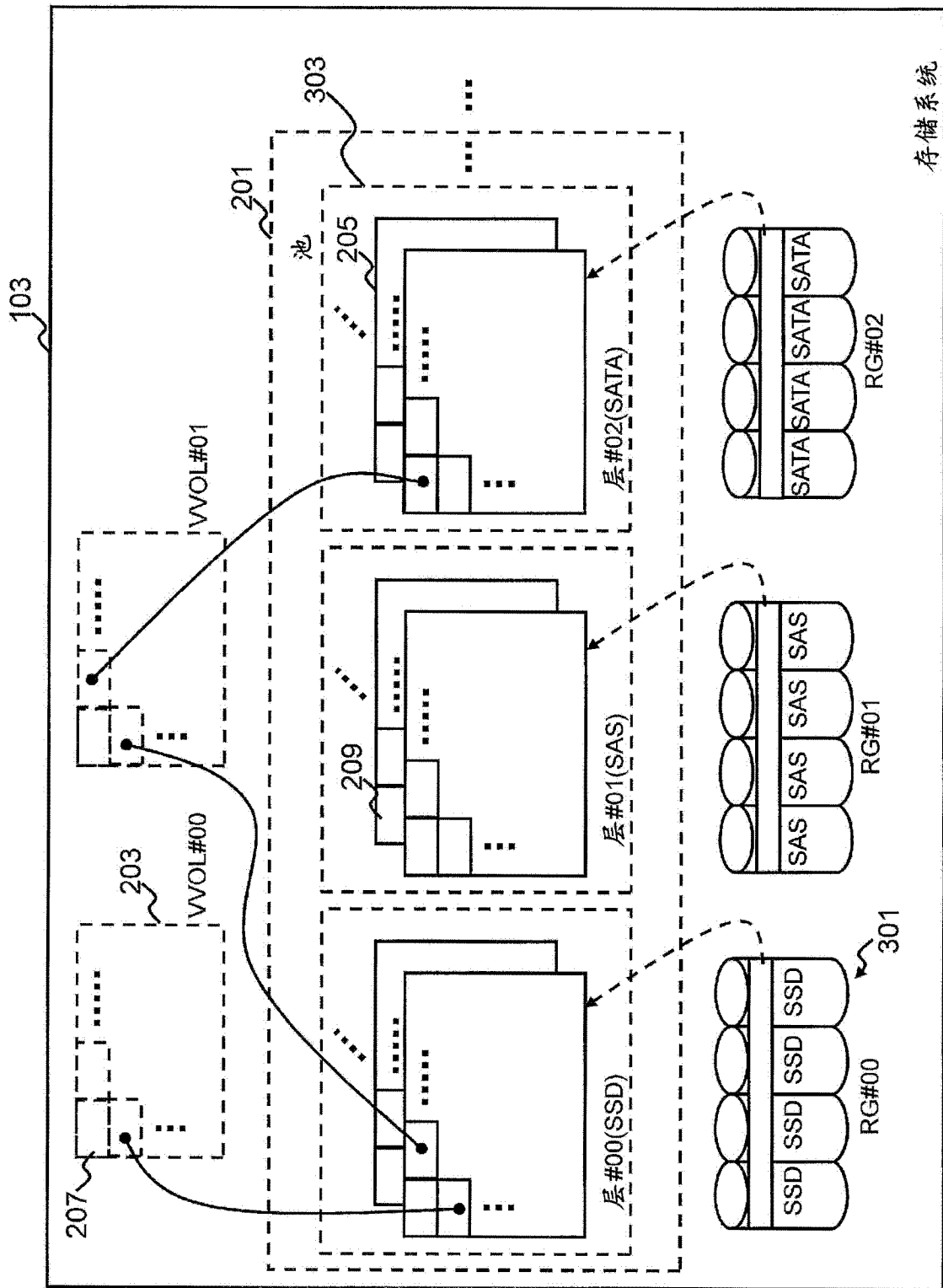


图 3

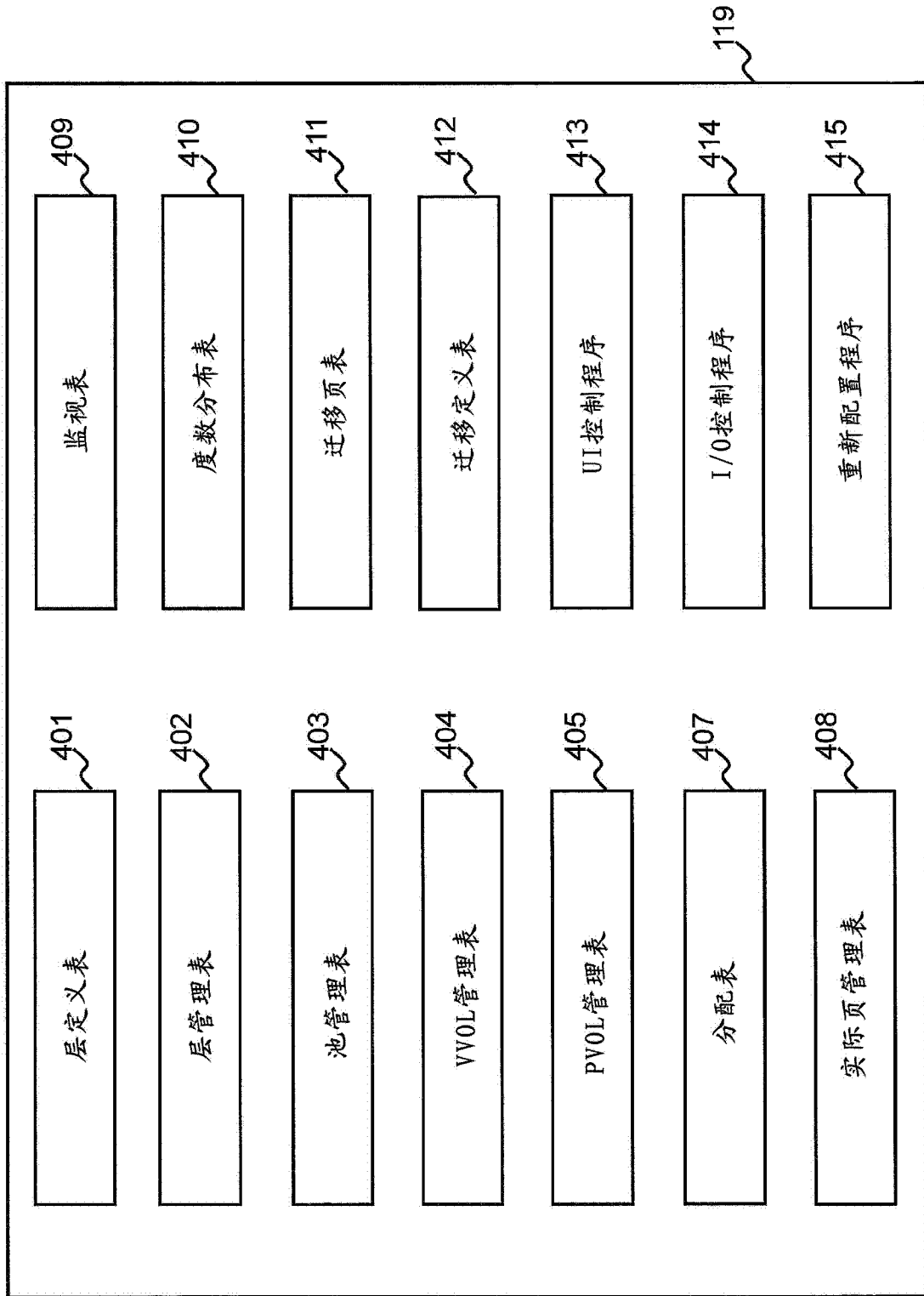


图 4

RVOL种类	层ID
SSD	00
HDD-SAS(15K)	01
HDD-SAS(10K)	02
HDD-SATA	03
...	...

401

图 5

402

池 ID	层 ID	层范围	PIVOL ID	层阙值	使用比例	实际页 ID
00	00	1000-5000	00	90%	75%	00000-00999
			02			
			...			
00	01	0-999	04	80%	75%	01000-02999
			...			
			...			
...
...

图 6

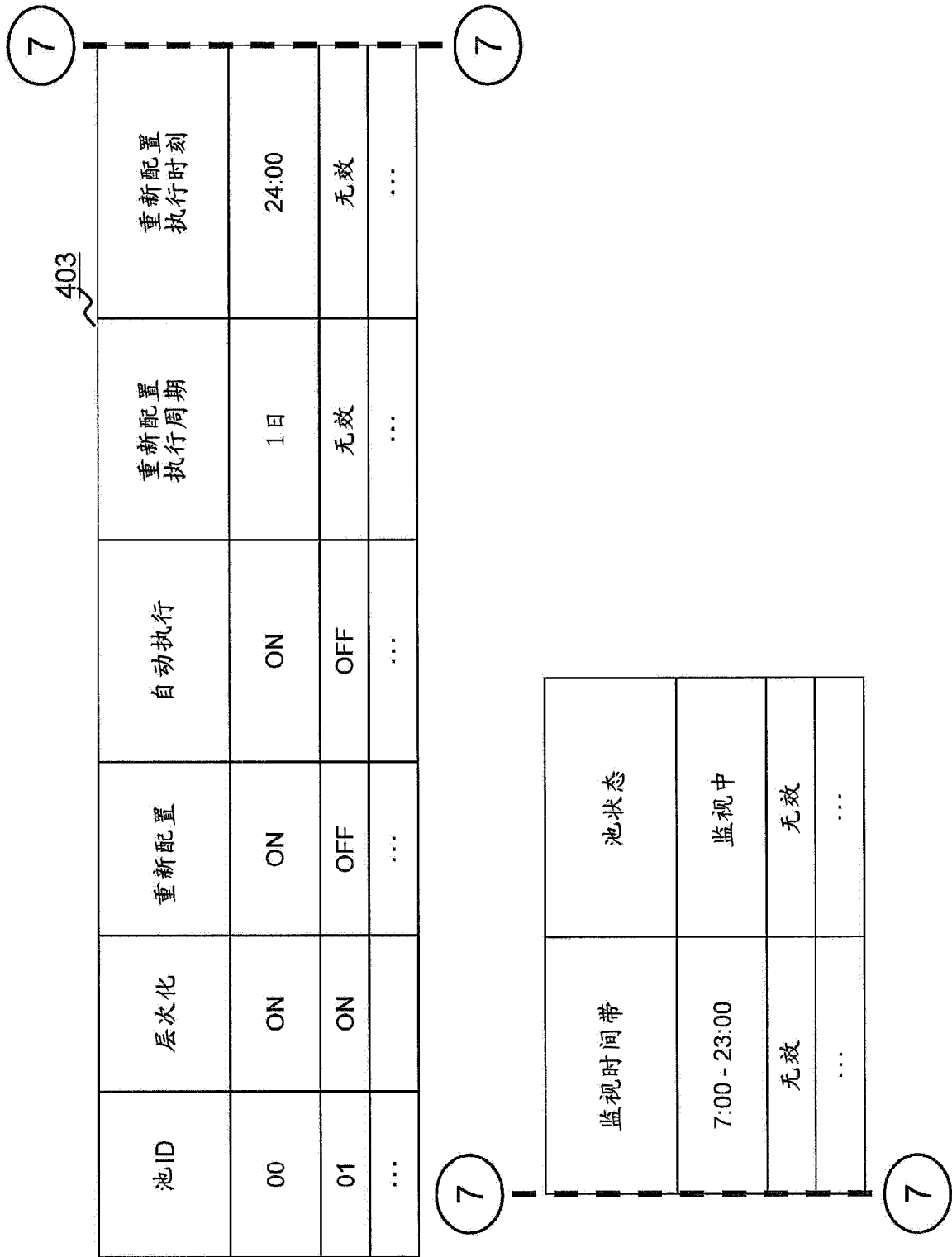


图 7

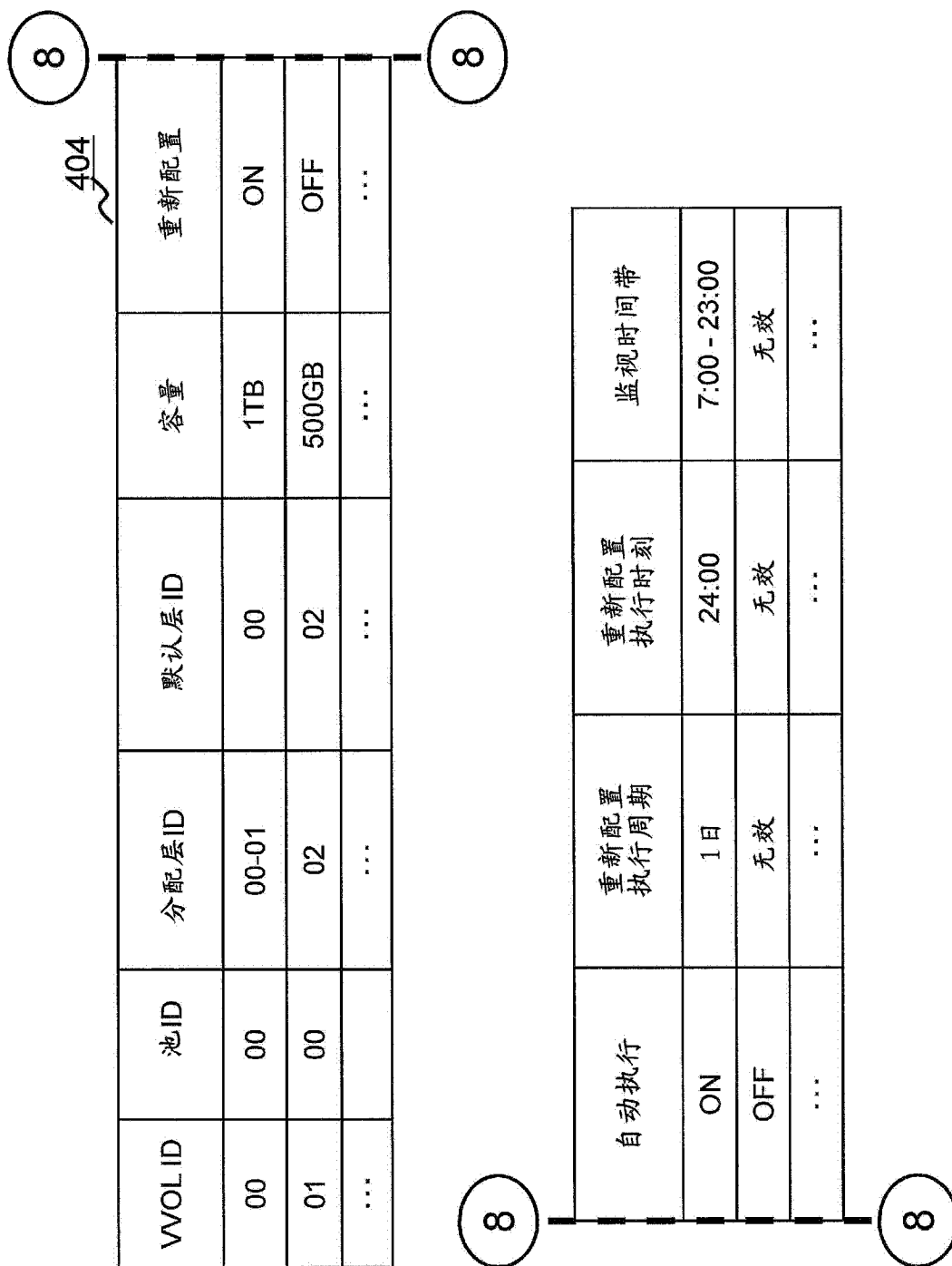


图 8

405

RG ID	PDEV ID	种类	PVOL ID	容量
00	01-04	SSD	01	500GB
			02	500GB
			03	500GB
01	05-08	HDD-SAS(15K)	04	500GB
			05	500GB
...

图 9

407

VVOL ID	虚拟页 ID	池 ID	实际页 ID	层 ID
00	00000	00	00010	01
	00001	00	00100	00
	...	00
01	00000	00	00020	01
	00001	-	-	-
...
...

图 10

408

池ID	实际页ID	状态
00	00000	使用中
	00001	自由

...

图 11

409

WVOL ID	虚拟页 ID	合计 I/O 数	平均 I/O 数	最大 I/O 数	最终 I/O 时刻
00	00000	1000	100	200	17:00
	00001	20	2	5	12:00

01	00000	—	—	—	—
	00001	—	—	—	—

...

图 12

410

VVOL ID	平均I/O数	页数
00	0	561
	1	27
	2	141
	3	57
	4-7	72
	8-15	236
	16-31	0
...

图 13

411

虚拟页 ID	迁移源层 ID	迁移目的地层 ID	迁移状态
10001	00	01	已迁移
10002	01	02	已迁移
10005	02	03	迁移中
10008	03	01	迁移中
10014	01	00	未迁移
...

图 14

412

VVOL ID	层维持	分配层 ID	保存阈值
00	ON	00	80%
		01	无效
		03	无效
01	OFF	01	70%
		02	无效
02	ON	02	无效
		03	70%
...

图 15

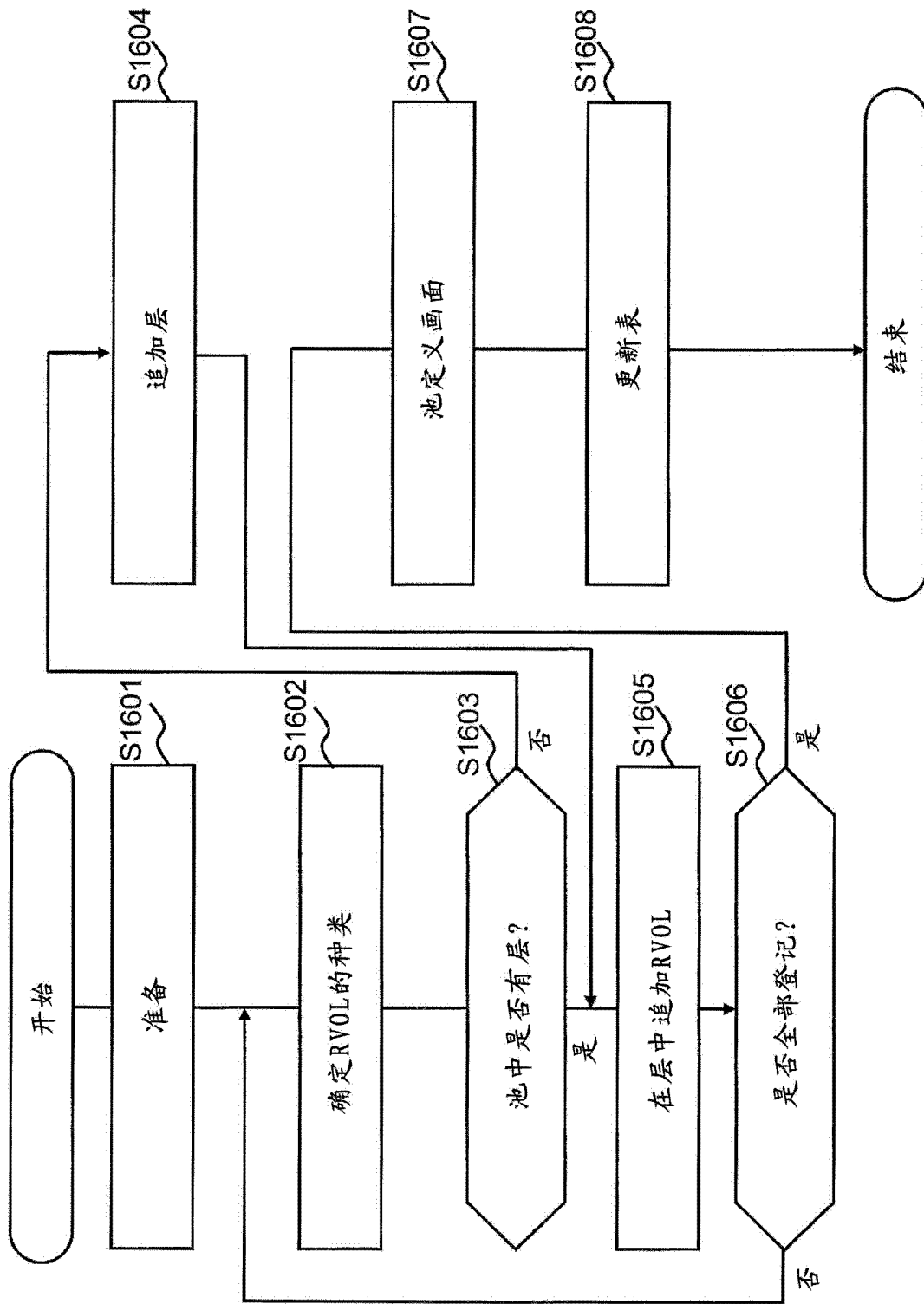


图 16

1700

池定义

池ID ▼

层次化 ON OFF

重新配置 ON OFF

自动执行 ON OFF

执行周期 ▼

执行时刻 ▼

监视时间带 从 ▼ 到 ▼

层阈值 共通 单独

90%	层	00	90%
		01	80%
		03	70%

图 17

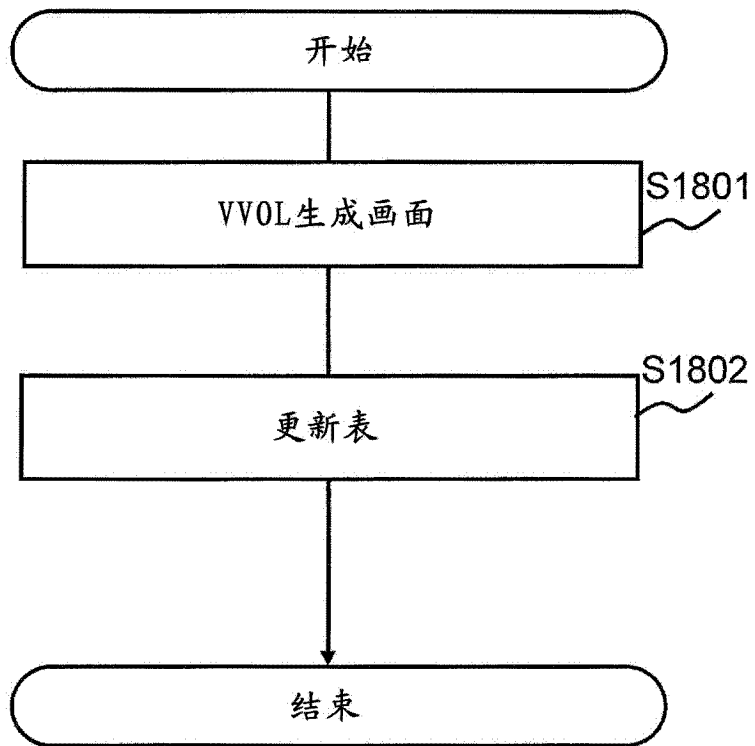


图 18

1900

VVOL生成

VVOL ID	<input type="text" value="00"/>	容量	<input type="text" value="1TB"/>
池 ID	<input type="text" value="00"/>		
重新配置	<input checked="" type="radio"/> ON <input type="radio"/> OFF		
自动执行	<input checked="" type="radio"/> ON <input type="radio"/> OFF		
执行周期	<input type="text" value="1日"/>		
执行时刻	<input type="text" value="24:00"/>		
监视时间带	从 <input type="text" value="07:00"/> 到 <input type="text" value="23:00"/>		
分配层	<input checked="" type="checkbox"/> 层 00 <input type="checkbox"/> 01 <input type="checkbox"/> 03	保存阈值	<input type="text" value="80%"/> <input type="text" value="无效"/> <input type="text" value="无效"/>
默认层	<input type="text" value="00"/>		

图 19

4404

WVOL ID	...	比较方法	比较对象
00	...	绝对值	平均I/O数
01	...	相对值	最大I/O数
...

图 20

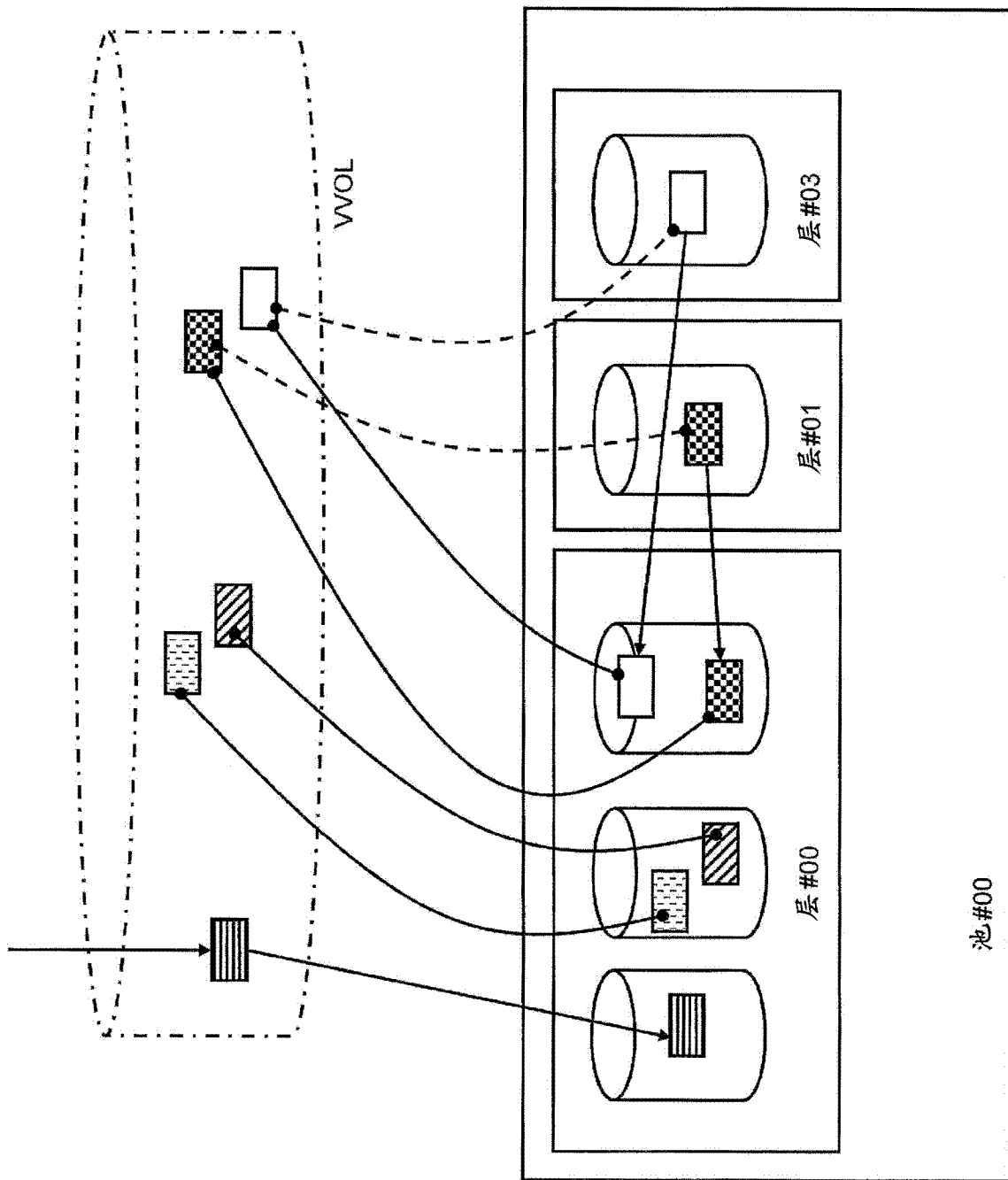


图 21

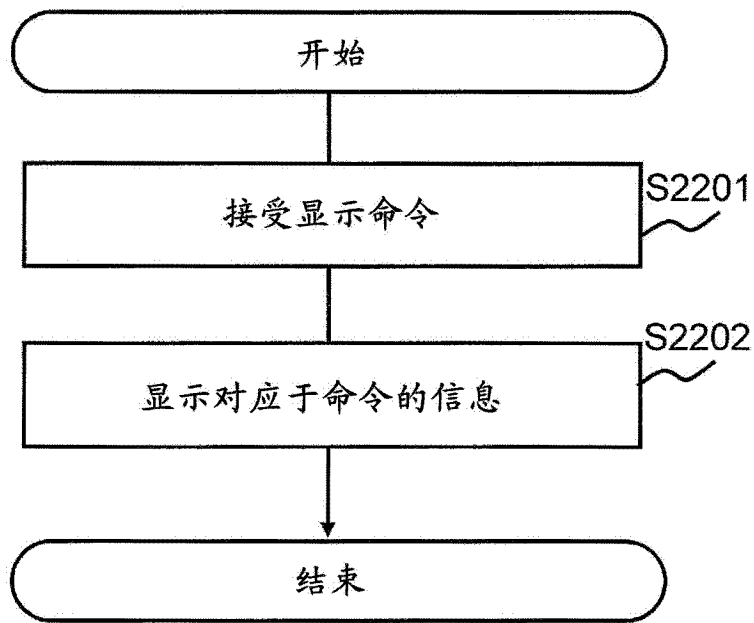


图 22

2300

度数分布显示

WVOL ID 00

度数分布

平均I/O数	页数
0	561
1	27
2	141
3	57
4-7	72
8-15	236
16-31	0

迁移执行

中断

图 23

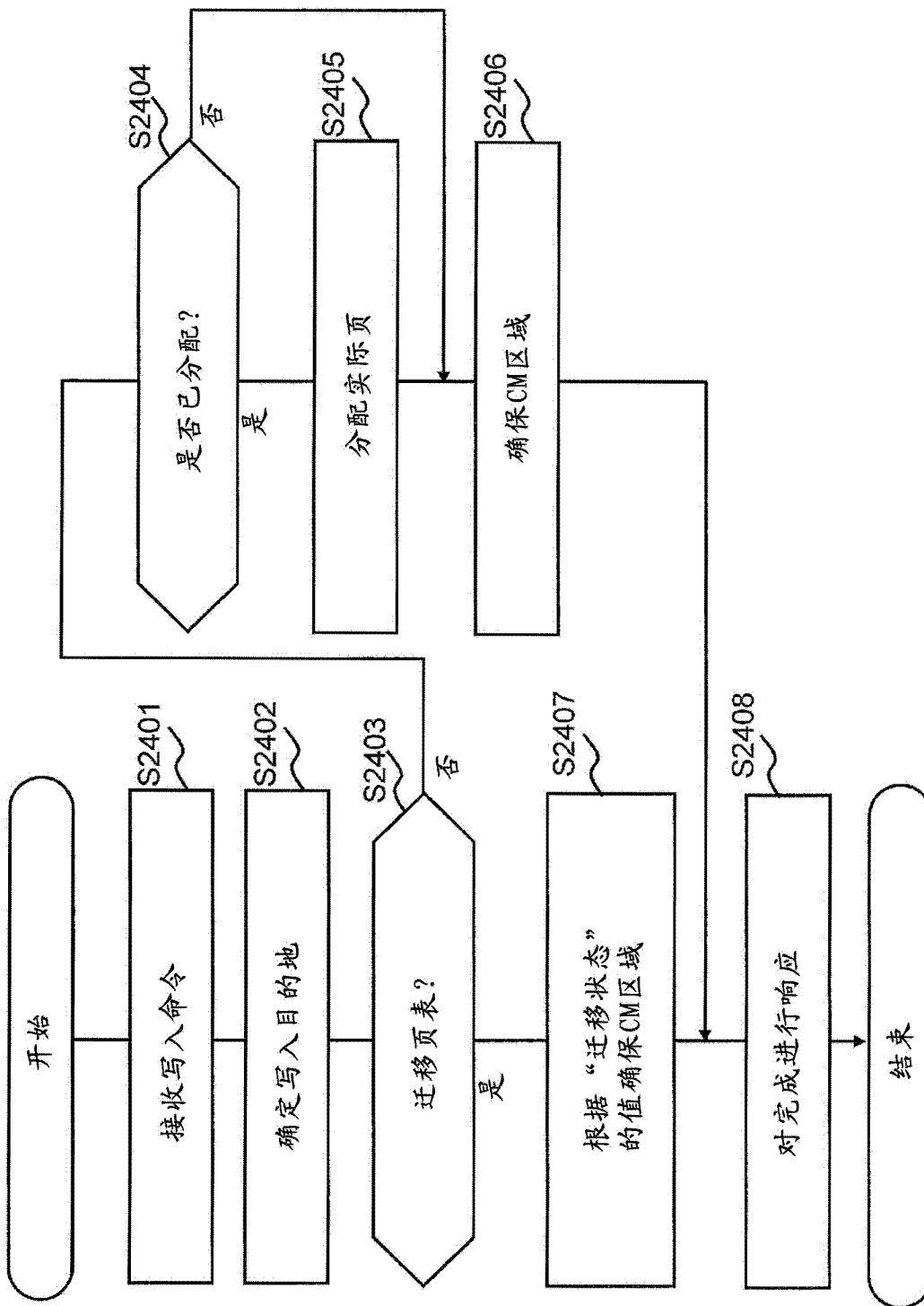


图 24

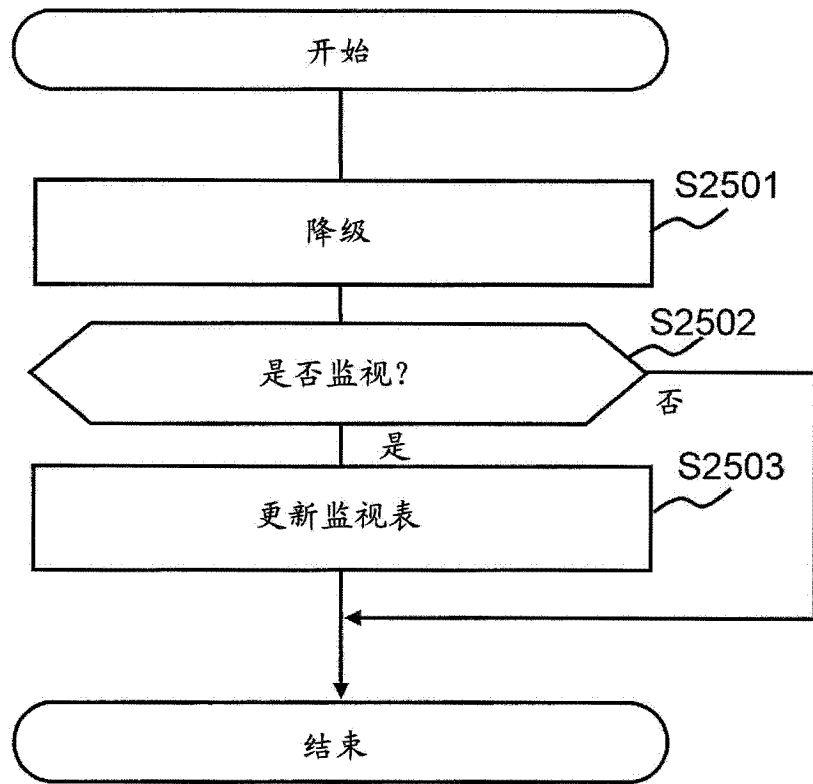


图 25

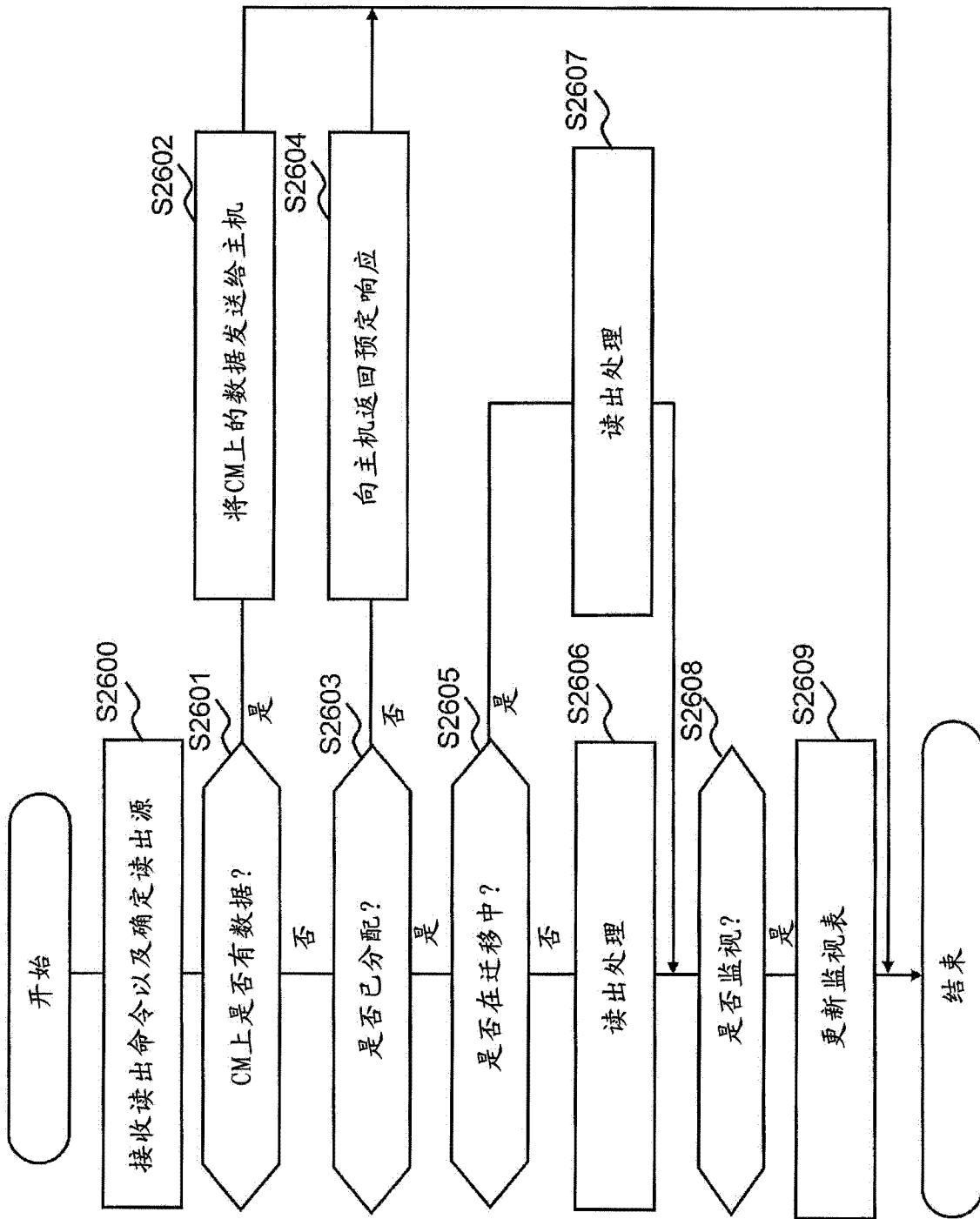


图 26

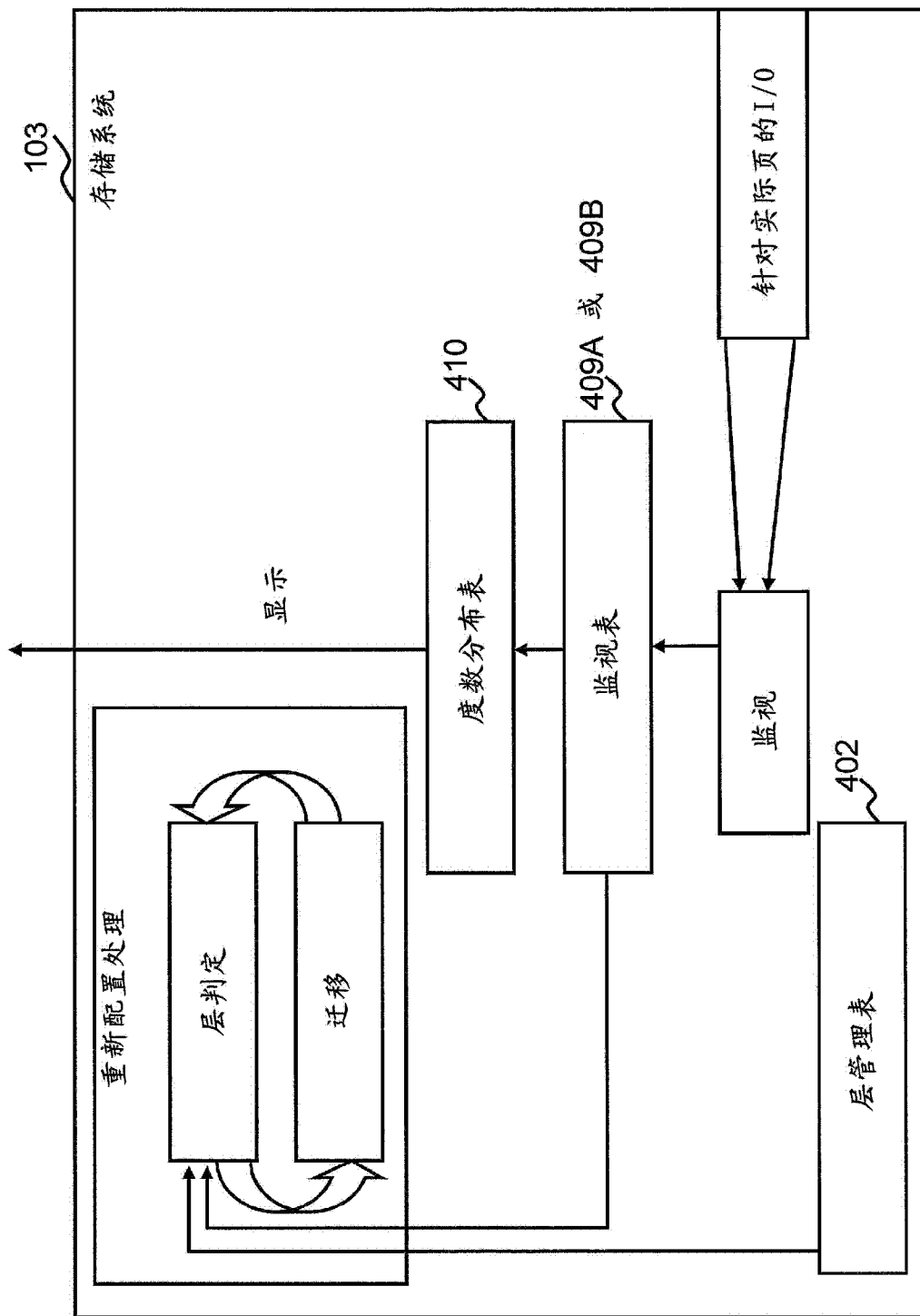


图 27

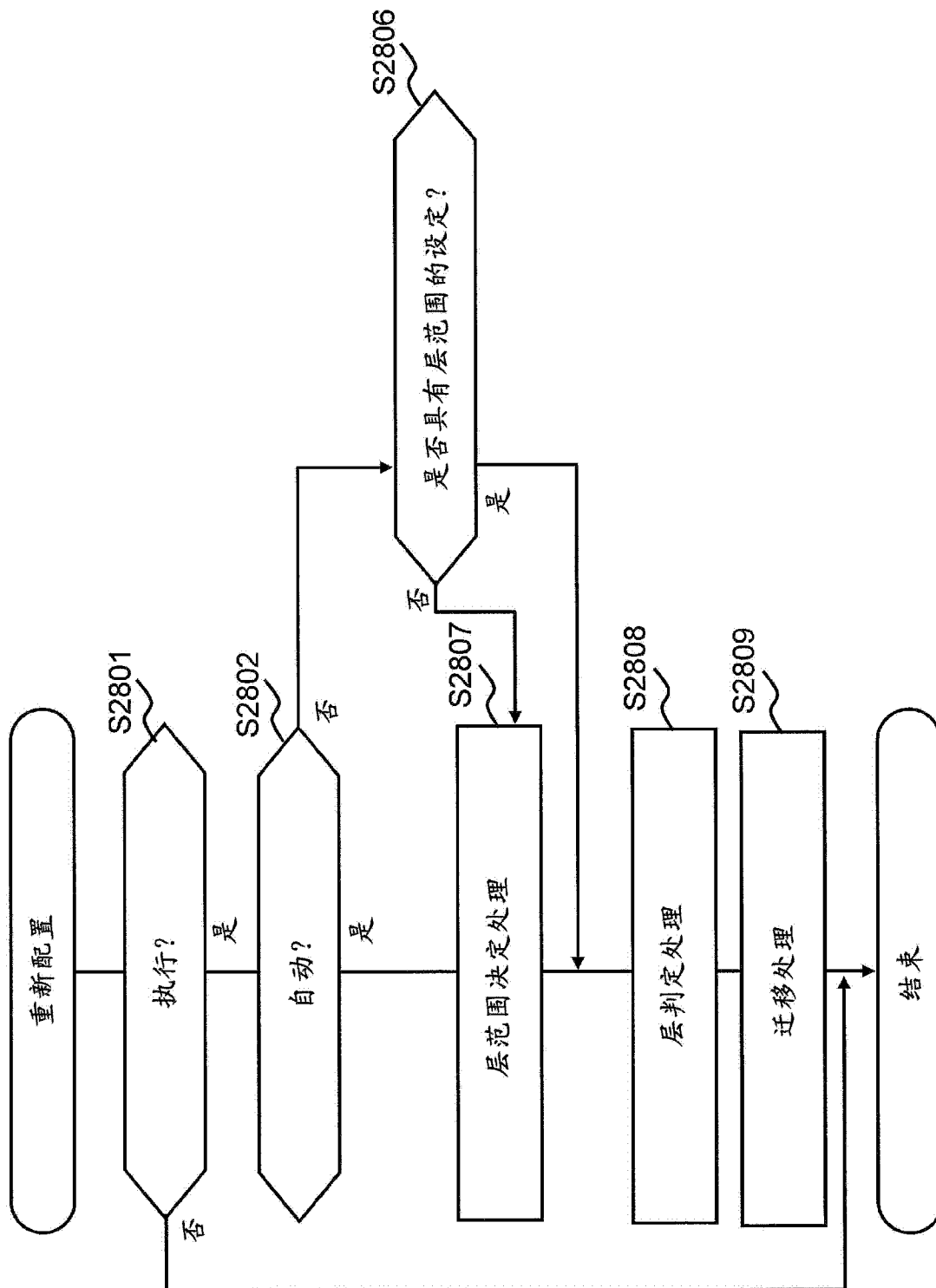


图 28

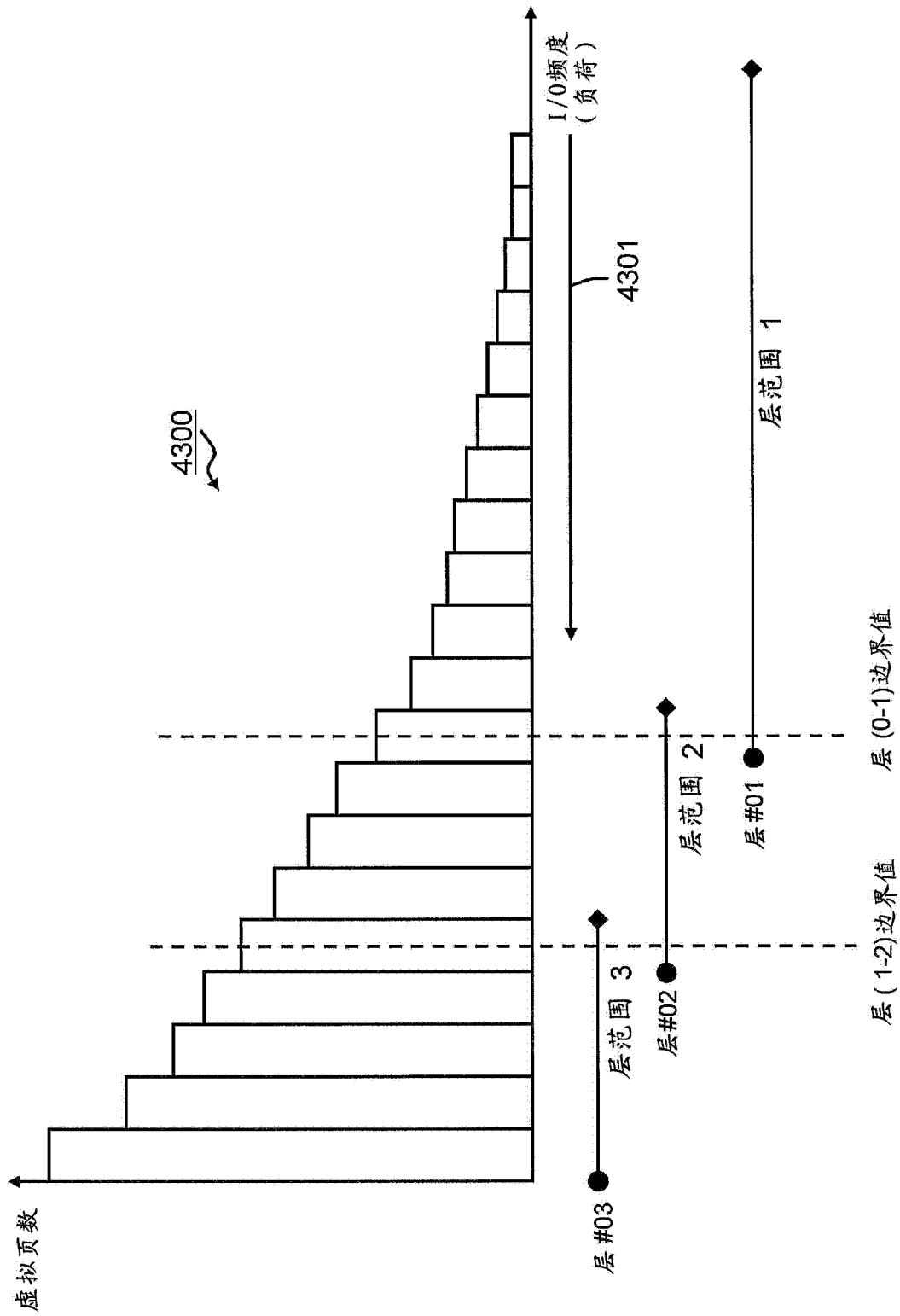


图 29

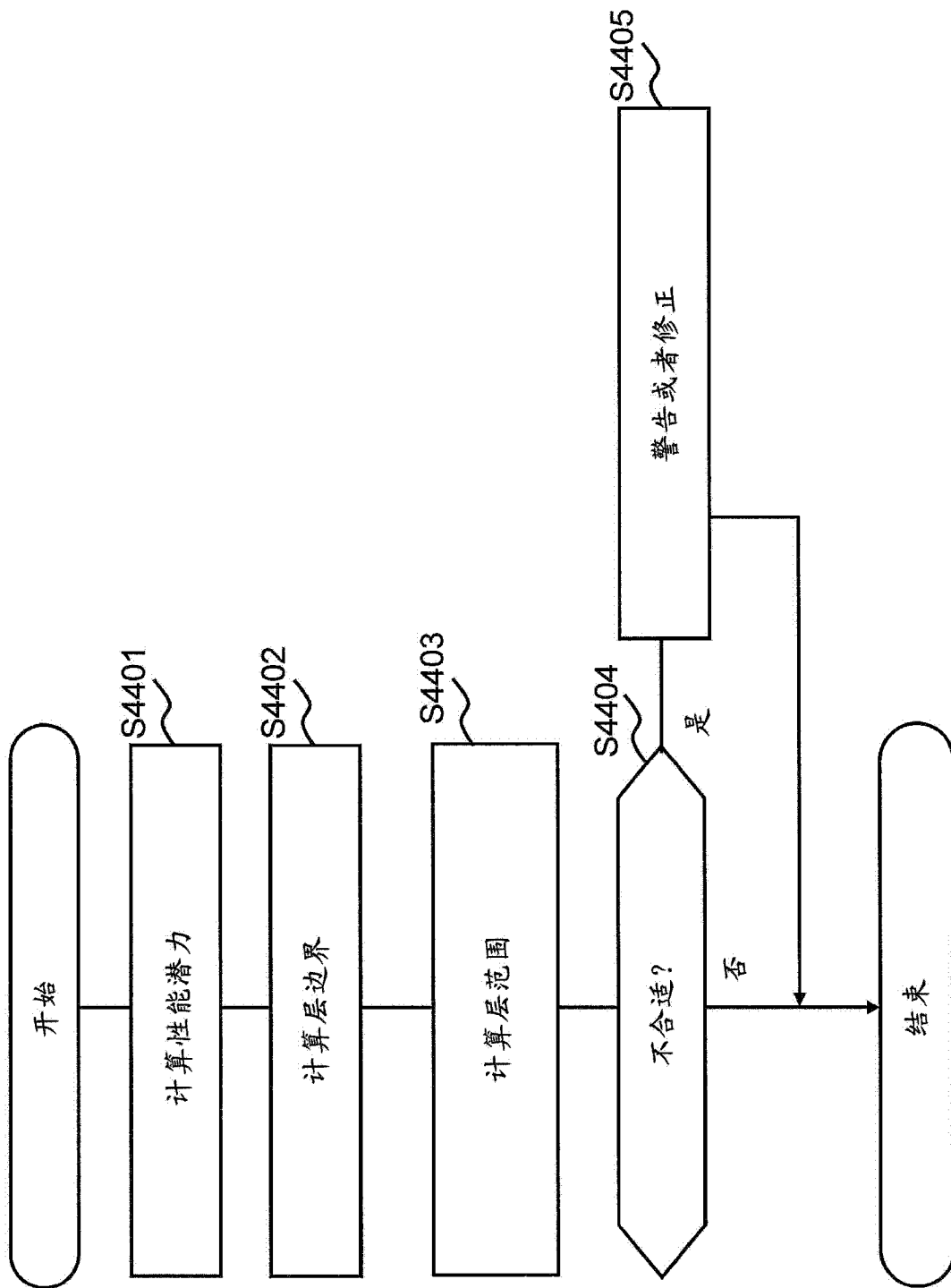


图 30

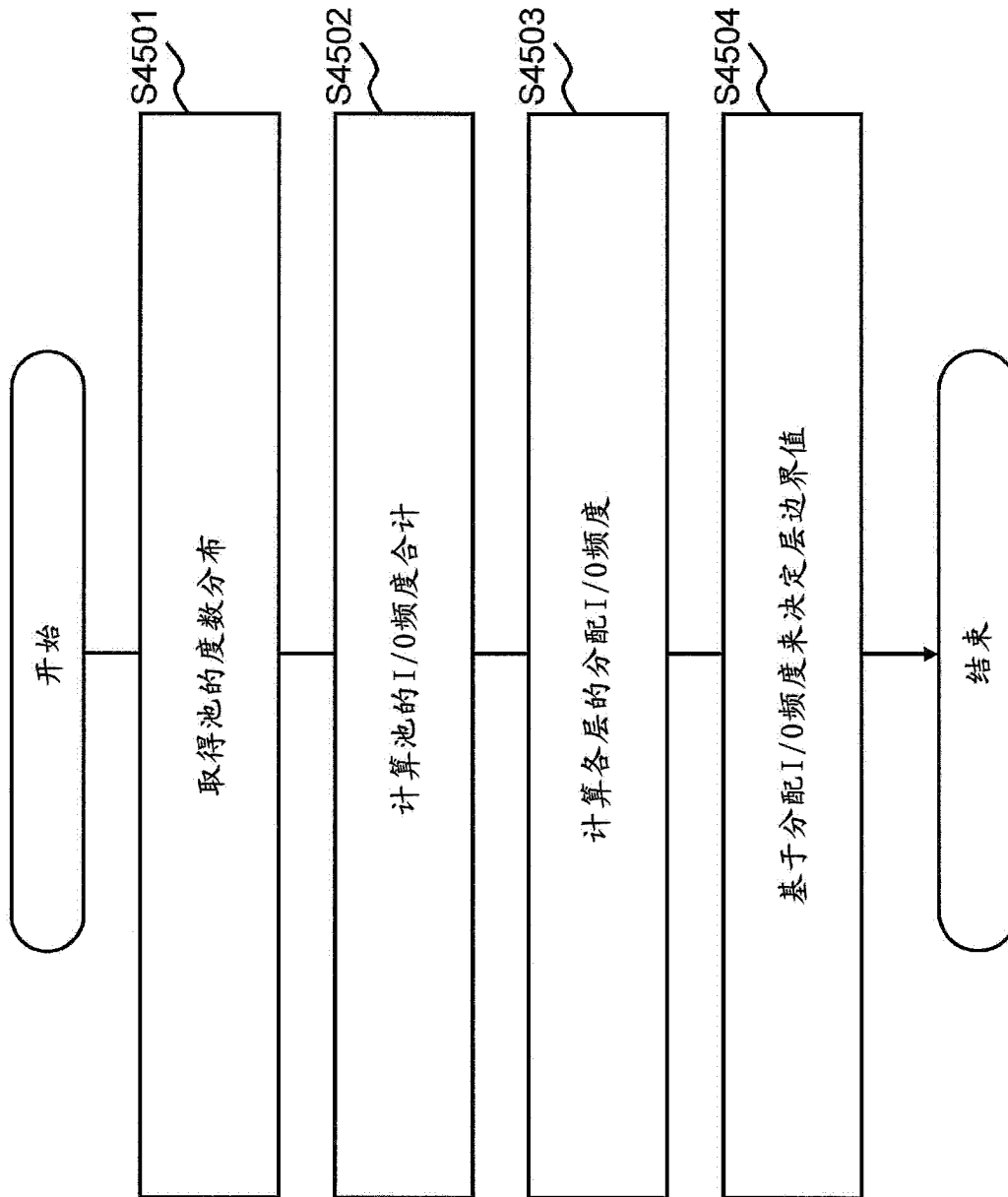


图 31

4403

池 ID	...	比较方法	比较对象
00	...	绝对值	平均 I/O 数
01	...	相对值	最大 I/O 数
...

图 32

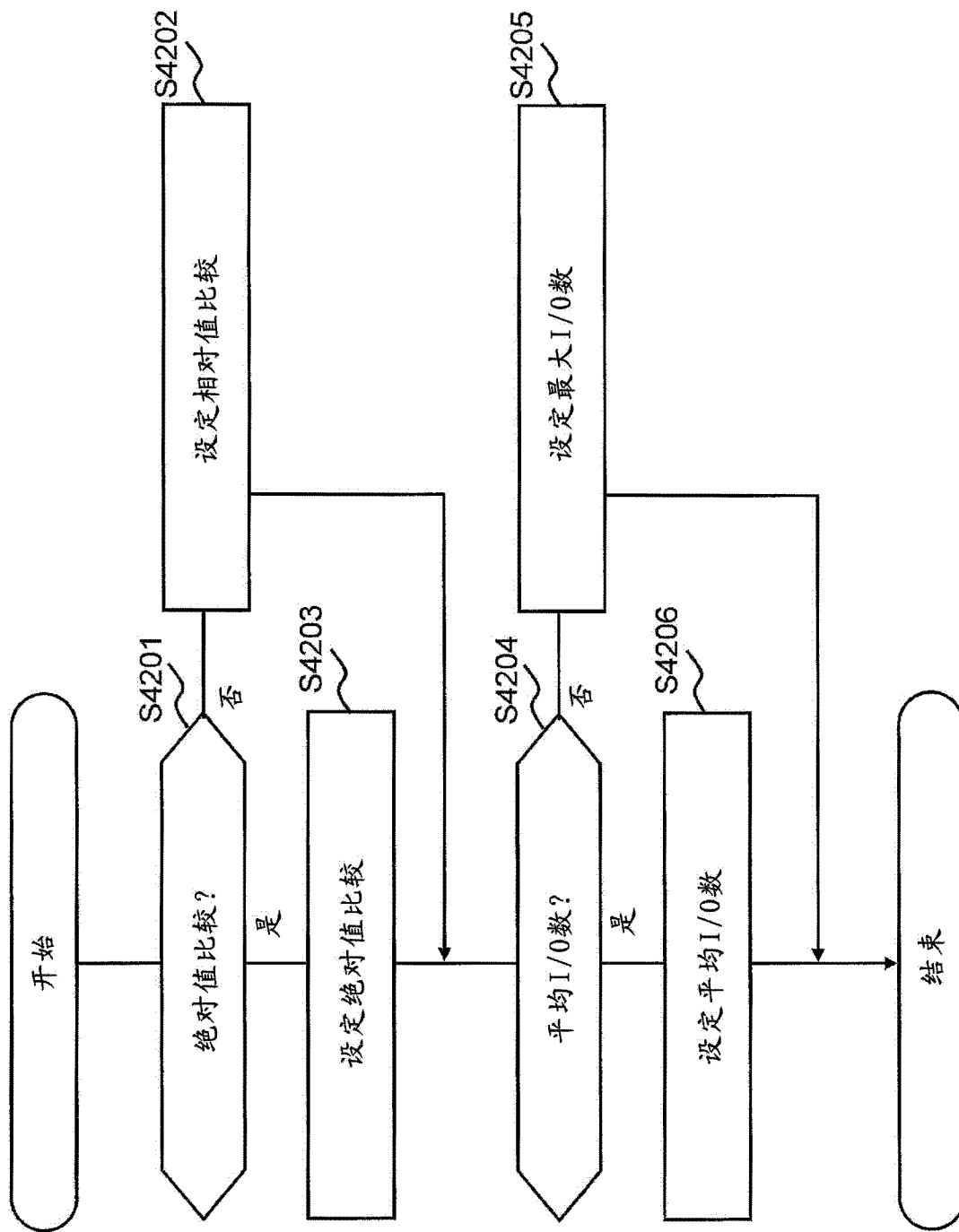


图 33

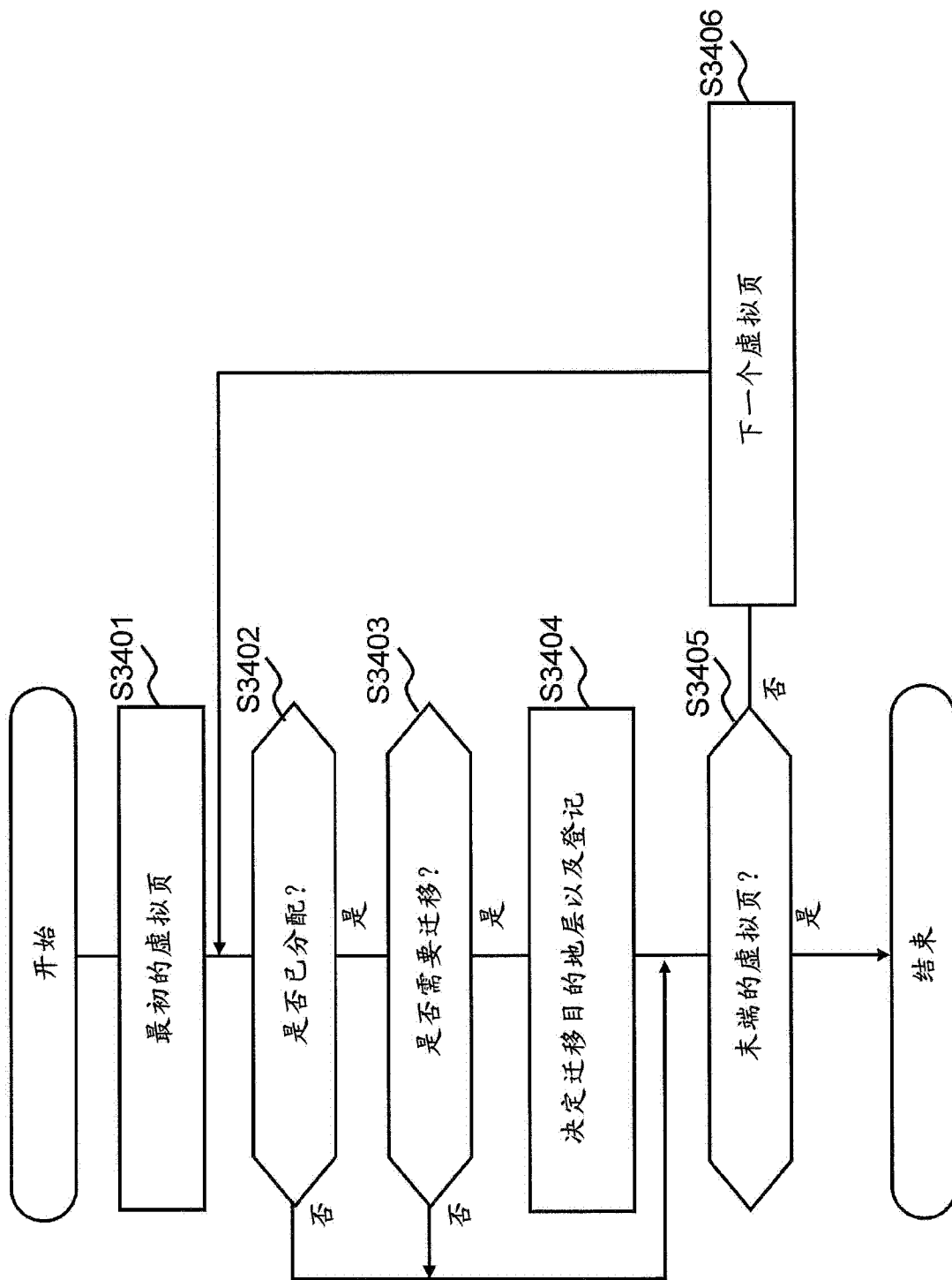


图 34

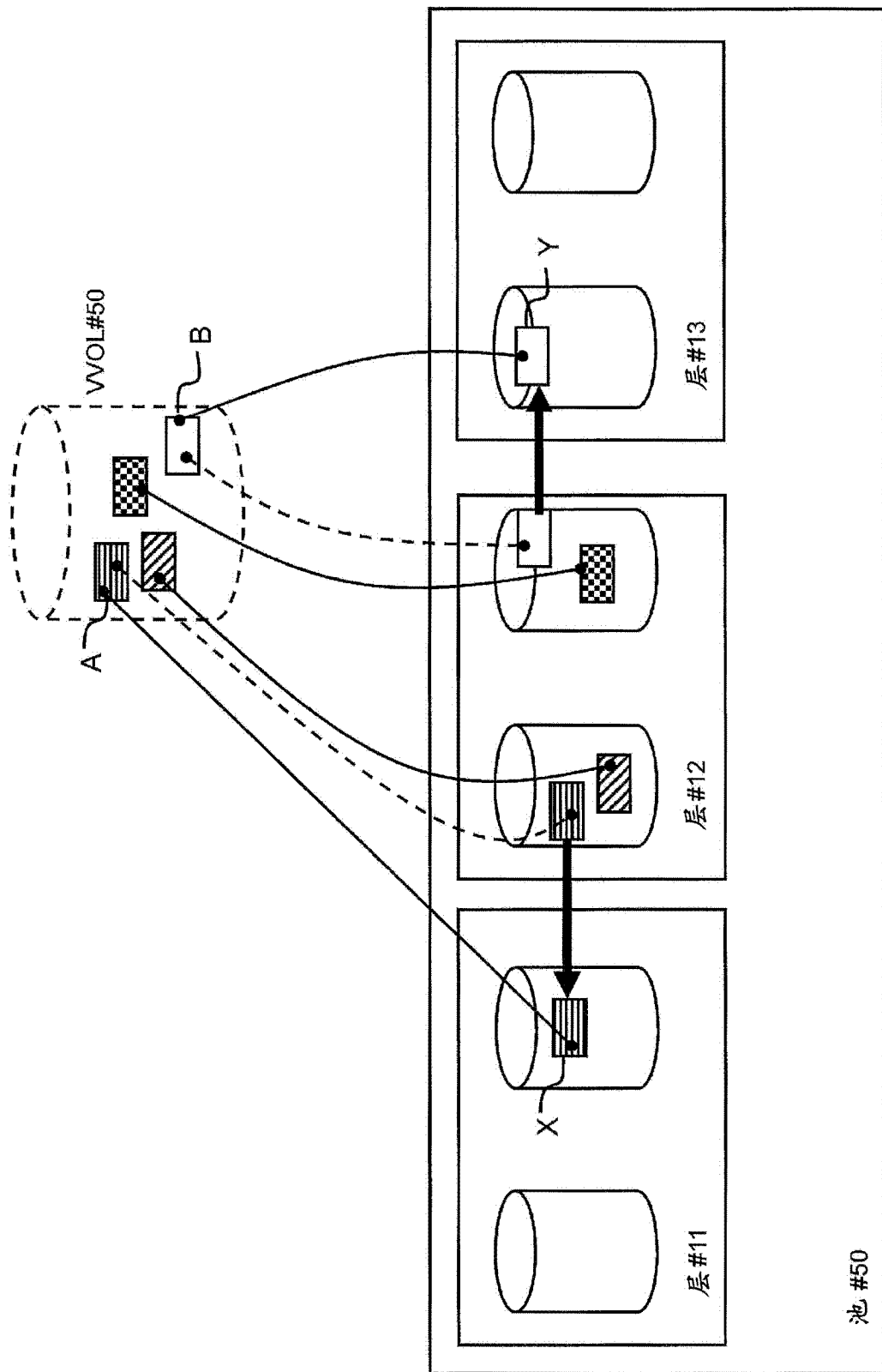


图 35

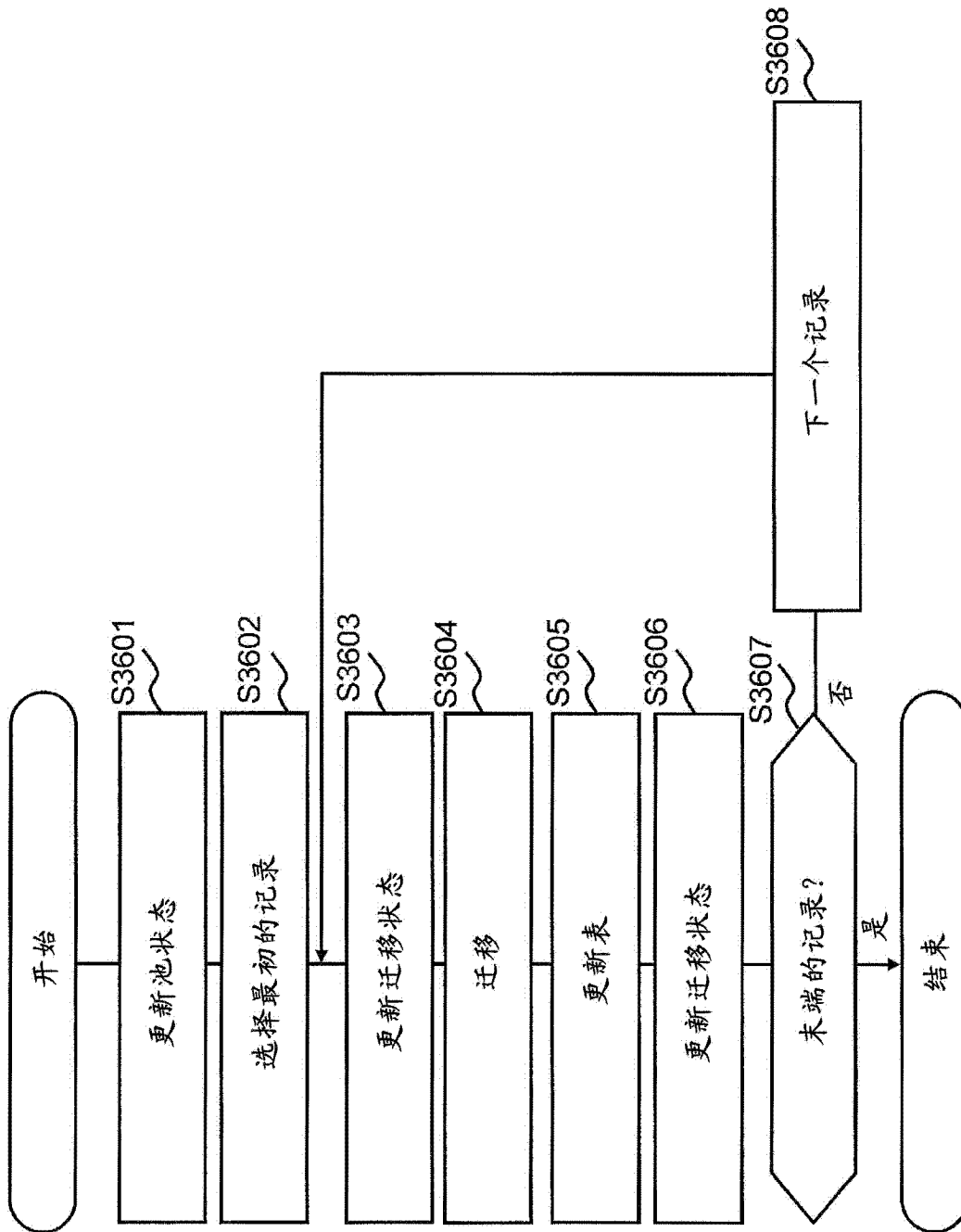


图 36

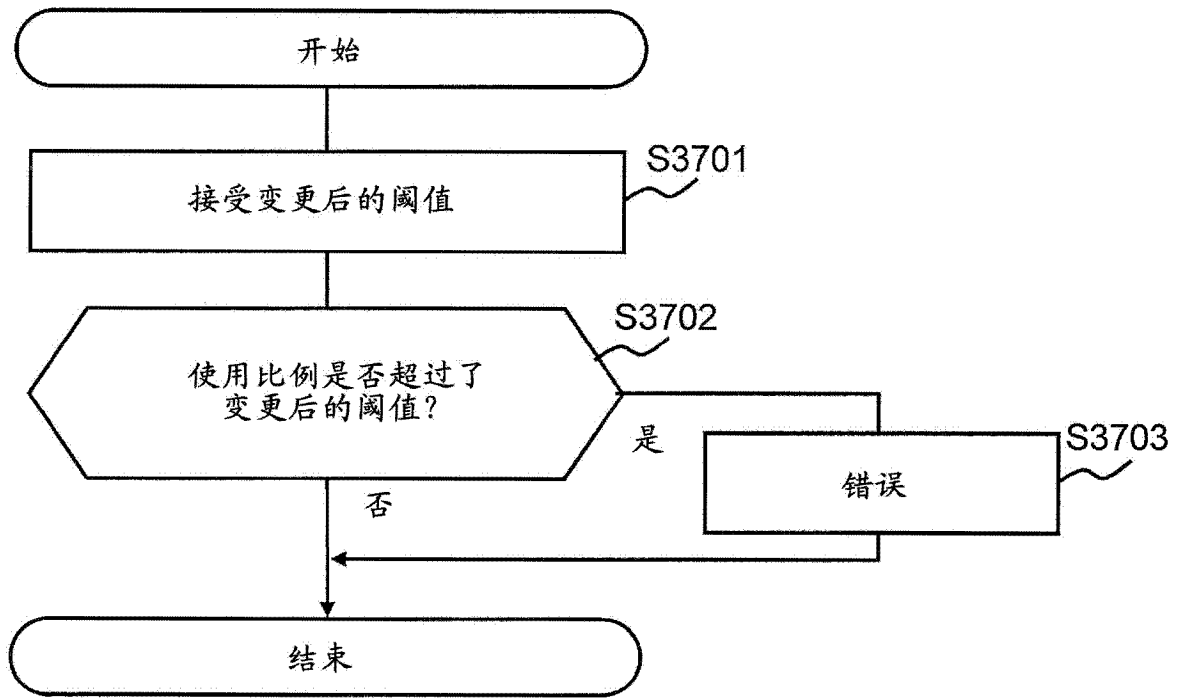


图 37

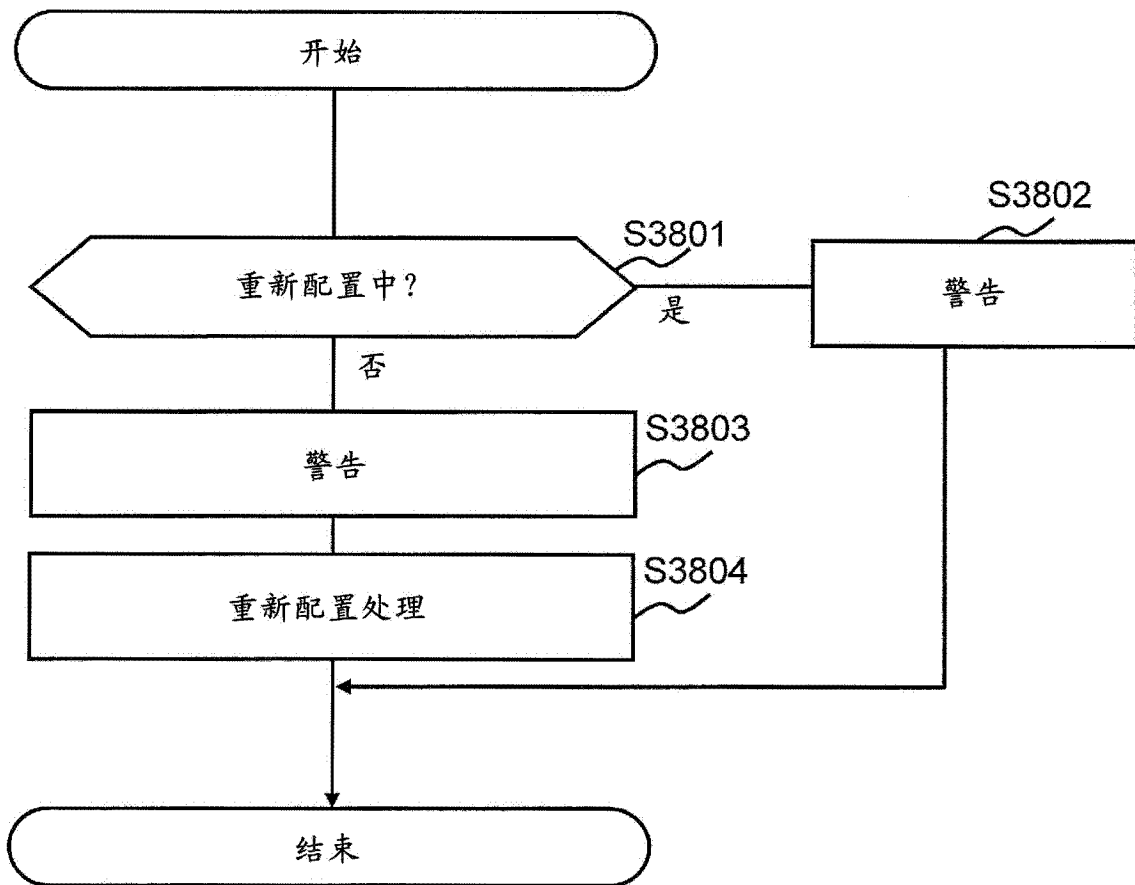


图 38