



(12) 发明专利申请

(10) 申请公布号 CN 105611979 A

(43) 申请公布日 2016. 05. 25

(21) 申请号 201480000790. 1

(51) Int. Cl.

(22) 申请日 2014. 03. 18

A99Z 99/00(2006. 01)

(85) PCT国际申请进入国家阶段日
2014. 08. 28

(86) PCT国际申请的申请数据
PCT/JP2014/057380 2014. 03. 18

(87) PCT国际申请的公布数据
W02015/140931 JA 2015. 09. 24

(71) 申请人 株式会社东芝
地址 日本东京都
申请人 东芝解决方案株式会社

(72) 发明人 浅山敦史

(74) 专利代理机构 永新专利商标代理有限公司
72002

代理人 徐殿军

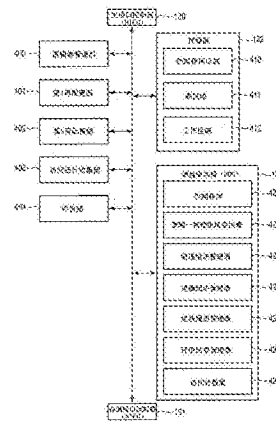
权利要求书3页 说明书17页 附图11页

(54) 发明名称

具备试验区域的层级化存储系统、存储控制器及程序

(57) 摘要

根据实施方式,层级化存储系统具备第1存储装置、与上述第1存储装置相比访问速度较低速而存储容量较大的第2存储装置、和存储控制器。上述存储控制器具备第1再配置部和第2再配置部。上述第1再配置部将上述第2存储装置的第3存储区域被分配了的第1逻辑组块的数据的配置目标、与上述第1存储的第1存储区域被分配且与上述第1逻辑组块相比访问频度统计值小的第2逻辑组块的数据的配置目标掉换。第2再配置部对于上述第3存储区域被分配了的逻辑组块的组,再分配上述第1存储装置的第2存储区域。



1. 一种层级化存储系统，

具备：

第 1 存储装置，具备包含多个物理组块的存储区域，该多个物理组块具有第 1 尺寸；

第 2 存储装置，具备包含多个物理组块的存储区域，并且与上述第 1 存储装置相比访问速度低速而存储容量大，该多个物理组块具有上述第 1 尺寸；以及

存储控制器，控制上述第 1 存储装置及上述第 2 存储装置；

上述存储控制器具备逻辑卷管理部、访问统计收集部、第 1 再配置部、第 2 再配置部和评价部；

上述逻辑卷管理部管理逻辑卷，所述逻辑卷是包含具有上述第 1 尺寸的多个逻辑组块且被向主机计算机提供的逻辑卷，以上述物理组块为单位被分配上述第 1 存储装置的第 1 存储区域、上述第 1 存储装置的第 2 存储区域及上述第 2 存储装置的第 3 存储区域；

上述访问统计收集部收集访问频度统计值，该访问频度统计值表示向上述逻辑卷的上述多个逻辑组块分别访问的状况；

上述第 1 再配置部选择访问频度统计值大的第 1 逻辑组块，在对上述第 1 逻辑组块分配了上述第 3 存储区域内的第 1 物理组块的情况下，将被分配给与上述第 1 逻辑组块相比访问频度统计值小的第 2 逻辑组块的上述第 1 存储区域内的第 2 物理组块向上述第 1 逻辑组块再分配，并且将上述第 1 物理组块向上述第 2 逻辑组块再分配；

上述第 2 再配置部以对上述第 3 逻辑组块的组分配了上述第 3 存储区域为条件，选择上述第 2 存储区域应被再分配的第 3 逻辑组块的组，对上述第 3 逻辑组块的组再分配上述第 2 存储区域；

上述评价部评价上述第 2 存储区域的上述再分配的效果。

2. 如权利要求 1 所述的层级化存储系统，

上述第 2 再配置部改变所选择的第 3 逻辑组块的组而重复上述第 3 逻辑组块的组的选择，在选择了上述第 3 逻辑组块的组的情况下，对于被分配了上述第 2 存储区域的第 4 逻辑组块的组再分配上述第 3 存储区域内的原来的区域，并且对于上述选择的第 3 逻辑组块的组再分配上述第 2 存储区域内的物理组块的组；

上述评价部将有关上述第 2 存储区域的上述再分配的效果的评价结果的第 1 履历、和物理组块一览的第 2 履历向上述存储控制器内的存储装置保存，所述物理组块一览表示对于被再分配了上述物理组块的组的上述第 3 逻辑组块的组在上述再分配之前分配的上述第 3 存储区域内的物理组块的组。

3. 如权利要求 2 所述的层级化存储系统，

上述第 2 再配置部基于上述第 1 履历及上述第 2 履历，将评价结果最高的第 1 再分配的状态复原。

4. 如权利要求 3 所述的层级化存储系统，

上述第 2 再配置部基于上述第 2 履历，确定对于在上述第 1 再分配中上述物理组块的组被再分配了的上述第 3 逻辑组块的组在上述第 1 再分配之前分配的上述第 3 存储区域内的物理组块的组，通过将上述确定的物理组块的组的数据向上述第 2 存储区域拷贝，将上述第 1 再分配的状态复原。

5. 如权利要求 2 所述的层级化存储系统，

上述第 2 再配置部从上述第 3 存储区域中, 以与上述第 2 存储区域对应的数量的物理组块为单位顺序地选择物理组块的组, 在从上述第 3 存储区域选择了物理组块的组的情况下, 选择被分配给上述被选择的物理组块的组的逻辑组块的组作为上述第 3 逻辑组块的组。

6. 如权利要求 2 所述的层级化存储系统,

上述第 2 再配置部依次选择被分配给上述第 1 存储区域内的物理组块的逻辑组块在该分配之前分别被分配了的上述第 3 存储区域内的物理组块, 在从上述第 3 存储区域选择了物理组块的情况下, 选择上述被选择的物理组块的周边的物理组块的组被分配了的逻辑组块的组作为上述第 3 逻辑组块的组。

7. 如权利要求 2 所述的层级化存储系统,

上述评价部对应于上述第 2 存储区域的上述再分配, 执行对于上述逻辑卷的随机读访问及对于上述逻辑卷的第 4 存储区域的顺序读访问的至少一方, 测量上述随机读访问及上述顺序读访问的上述至少一方的吞吐量, 取得该测量出的吞吐量作为有关上述第 2 存储区域的上述再分配的效果的评价结果。

8. 一种存储控制器, 在具备第 1 存储装置和第 2 存储装置的层级化存储系统中, 控制上述第 1 存储装置及上述第 2 存储装置, 所述第 1 存储装置具备包含多个物理组块的存储区域, 该多个物理组块具有第 1 尺寸; 所述第 2 存储装置具备包含多个物理组块的存储区域, 并且与上述第 1 存储装置相比访问速度低速而存储容量大, 该多个物理组块具有上述第 1 尺寸; 该存储控制器,

具备:

逻辑卷管理部, 管理逻辑卷, 所述逻辑卷是包含具有上述第 1 尺寸的多个逻辑组块且被向主机计算机提供的逻辑卷, 以所述物理组块为单位被分配上述第 1 存储装置的第 1 存储区域、上述第 1 存储装置的第 2 存储区域及上述第 2 存储装置的第 3 存储区域;

访问统计收集部, 收集访问频度统计值, 该访问频度统计值表示向上述逻辑卷的上述多个逻辑组块分别访问的状况;

第 1 再配置部, 选择访问频度统计值大的第 1 逻辑组块, 在对上述第 1 逻辑组块分配了上述第 3 存储区域内的第 1 物理组块的情况下, 将被分配给与上述第 1 逻辑组块相比访问频度统计值小的第 2 逻辑组块的上述第 1 存储区域内的第 2 物理组块向上述第 1 逻辑组块再分配, 并且将上述第 1 物理组块向上述第 2 逻辑组块再分配;

第 2 再配置部以对上述第 3 逻辑组块的组分配了上述第 3 存储区域为条件, 选择上述第 2 存储区域应被再分配的第 3 逻辑组块的组, 对上述第 3 逻辑组块的组再分配上述第 2 存储区域; 以及

评价部, 评价上述第 2 存储区域的上述再分配的效果。

9. 一种程序, 在具备第 1 存储装置和第 2 存储装置的层级化存储系统中, 所述第 1 存储装置具备包含多个物理组块的存储区域, 该多个物理组块具有第 1 尺寸; 所述第 2 存储装置具备包含多个物理组块的存储区域, 并且与上述第 1 存储装置相比访问速度低速而存储容量大, 该多个物理组块具有上述第 1 尺寸, 该程序用来使控制上述第 1 存储装置及上述第 2 存储装置的存储控制器执行:

由上述逻辑卷管理部管理逻辑卷, 所述逻辑卷是包含具有上述第 1 尺寸的多个逻辑组

块且被向主机计算机提供的逻辑卷,以上述物理组块为单位被分配上述第 1 存储装置的第 1 存储区域、上述第 1 存储装置的第 2 存储区域及上述第 2 存储装置的第 3 存储区域;

收集访问频度统计值,该访问频度值表示向上述逻辑卷的上述多个逻辑组块分别访问的状况;

选择访问频度统计值大的第 1 逻辑组块,在对上述第 1 逻辑组块分配了上述第 3 存储区域内的第 1 物理组块的情况下,将被分配给与上述第 1 逻辑组块相比访问频度统计值小的第 2 逻辑组块的上述第 1 存储区域内的第 2 物理组块向上述第 1 逻辑组块再分配,并且将上述第 1 物理组块向上述第 2 逻辑组块再分配;

以对上述第 3 逻辑组块的组分配了上述第 3 存储区域为条件,选择上述第 2 存储区域应被再分配的第 3 逻辑组块的组对上述第 3 逻辑组块的组再分配上述第 2 存储区域;以及评价上述第 2 存储区域的上述再分配的效果。

具备试验区域的层级化存储系统、存储控制器及程序

技术领域

[0001] 本发明的实施方式涉及具备试验区域的层级化存储系统、存储控制器及程序。

背景技术

[0002] 近年,开发了具备访问速度不同的第1及第2存储装置的存储系统。这里,假设第1存储装置的访问速度及存储容量是高速小容量,第2存储装置与第1存储装置相比是低速大容量。存储系统通过将第1存储装置(以下,称作高速存储装置)及第2存储装置(以下,称作低速存储装置)层级性地组合来实现。所以,这样的存储系统也被称作层级化存储系统。在层级化存储系统中,一般高速存储装置被分配给上位层级,低速存储装置被分配给下位层级。

[0003] 此外,近年来,在存储系统全部中,存储区域的虚拟化变得普遍。这在层级化存储系统中也这样。在这样的层级化存储系统中,高速存储装置及低速存储装置的存储区域(以下,称作物理卷)的至少一部分被分配(即被映射)到逻辑卷中。所谓逻辑卷,是指由利用存储系统(层级化存储系统)的主计算机(以下,称作主机)识别为逻辑性的存储驱动器(即,逻辑盘或逻辑单元)的逻辑性的存储区域。

[0004] 在主机上动作的应用那样的上位的程序使用逻辑地址对逻辑卷访问。逻辑地址例如被主机的操作系统(OS)或存储系统的存储控制器具有的地址变换功能变换为高速存储装置或低速存储装置的物理地址。基于该物理地址,访问映射在逻辑卷内的存储区域中的物理卷内的存储区域。由此,上位的程序能够不意识存储装置内的存储区域而对该存储装置内的存储区域访问。

[0005] 在层级化存储系统中,将逻辑卷及物理卷为了管理而分割为具有一定尺寸的多个小区域。将这样的小区域称作组块(chunk)。在需要将逻辑卷内的组块及物理卷内的组块区别的情况下,将前者称作逻辑组块,将后者称作物理组块。

[0006] 存储控制器按照组块(例如,逻辑组块),监视向对应的组块的访问的状况。作为表示该访问的状况的指标,一般使用一定期间(监视期间)中的按照组块的访问次数(即,读/写的输入输出次数)或访问数据量的总和那样的访问频度统计值。

[0007] 存储控制器基于按照逻辑组块的访问频度统计值,分别对于被分配了上位层级的物理组块的逻辑组块再分配下位层级的物理组块、对于被分配了下位层级的物理组块的逻辑组块再分配上位层级的物理组块。即,存储控制器分别将上位层级的物理组块的数据向下位层级的物理组块再配置,将下位层级的物理组块的数据向上位层级的物理组块再配置。

[0008] 对应于上述再分配(再配置),将逻辑组块的地址变换信息更新。即,将与逻辑组块的逻辑地址建立了对应的物理地址更新,以使其表示再分配给该逻辑组块的物理组块。上位程序在这样的向逻辑组块的物理组块的再分配后,也能够不意识该再分配(更详细地讲,伴随着该再分配的地址变换信息的更新),而使用该逻辑组块的逻辑地址向对应的物理组块(即,再分配后的物理组块)访问。这样,层级化存储系统与存储区域的虚拟化的相容

性良好。

[0009] 现有技术文献

[0010] 专利文献

[0011] 专利文献 1 :特开平 10 - 154101 号公报

发明内容

[0012] 发明所要解决的课题

[0013] 如上述那样,在以往的层级化存储系统中,按照组块取得访问频度统计值,将访问频度统计值较大(即,访问次数或总访问数据量较多)的组块(逻辑组块)与上位层级(即,高速存储装置)的物理组块建立对应。即,在以往技术中,按照监视期间,对被判定为访问频度统计值(访问频度)较大(较高)的逻辑组块分配(映射)高速存储的存储区域(即,高速区域)。通过这样的分配(层级化),能够期待层级化存储系统整体的性能的提高。

[0014] 在以往技术中,对访问频度统计值较小的逻辑组块分配低速存储装置的存储区域(即,低速区域)。但是,在访问频度统计值较小的逻辑组块的集合中,有可能包含通过被分配高速区域能够使层级化存储系统整体的性能(即,系统性能)进一步提高的组块。

[0015] 本发明的目的是提供一种给访问频度统计值较小的逻辑组块也给与被分配高速区域的机会、能够评价该分配给系统性能带来的效果的具备试验区域的层级化存储系统、存储控制器及程序。

[0016] 用于解决课题的方法

[0017] 根据技术方案,层级化存储装置具备第 1 存储装置、与上述第 1 存储装置相比访问速度较低速而存储容量较大的第 2 存储装置、和存储控制器。上述第 1 存储装置具备包含具有第 1 尺寸的多个物理组块的存储区域。上述第 2 存储装置具备包含具有上述第 1 尺寸的多个物理组块的存储区域。上述存储控制器控制上述第 1 存储装置及上述第 2 存储装置。上述存储控制器具备逻辑卷管理部、访问统计收集部、第 1 再配置部、第 2 再配置部和评价部。上述逻辑卷管理部以上述物理组块为单位被分配上述第 1 存储装置的第 1 存储区域、上述第 1 存储装置的第 2 存储区域及上述第 2 存储装置的第 3 存储区域的逻辑卷。上述逻辑卷包含具有上述第 1 尺寸的多个逻辑组块,且被向主机计算机提供。上述访问统计收集部收集访问频度统计值,该访问频度统计值表示向上述逻辑卷的上述多个逻辑组块分别访问的状况。上述第 1 再配置部选择访问频度统计值较大的第 1 逻辑组块,在对上述第 1 逻辑组块分配了上述第 3 存储区域内的第 1 物理组块的情况下,将被分配给与上述第 1 逻辑组块相比访问频度统计值小的第 2 逻辑组块的上述第 1 存储区域内的第 2 物理组块向上述第 1 逻辑组块再分配,并且将上述第 1 物理组块向上述第 2 逻辑组块再分配。上述第 2 再配置部以对上述第 3 逻辑组块的组分配了上述第 3 存储区域为条件,选择上述第 2 存储区域应被再分配的第 3 逻辑组块的组对上述第 3 逻辑组块的组再分配上述第 2 存储区域。上述评价部评价上述第 2 存储区域的上述再分配的效果。

附图说明

[0018] 图 1 是表示有关实施方式的计算机系统的典型的硬件结构的框图。

- [0019] 图 2 是用来说明图 1 所示的层级化存储系统中的存储区域的典型的层级构造的图。
- [0020] 图 3 是用来说明图 2 所示的逻辑卷的典型的结构的图。
- [0021] 图 4 是主要表示图 1 所示的存储控制器的典型的功能结构的框图。
- [0022] 图 5 是表示图 4 所示的逻辑—物理组块变换表的数据构造的例子的图。
- [0023] 图 6 是表示图 4 所示的物理组块管理表的数据构造的例子的图。
- [0024] 图 7 是表示图 4 所示的逻辑组块管理表的数据构造的例子的图。
- [0025] 图 8 是表示图 4 所示的组块履历管理表的数据构造的例子的图。
- [0026] 图 9 是表示图 4 所示的评价结果管理表的数据构造的例子的图。
- [0027] 图 10 是用来说明在该实施方式中采用的第 1 再配置处理的典型的次序的流程图。
- [0028] 图 11 是用来说明在该实施方式中采用的第 2 再配置处理的典型的次序的流程图。
- [0029] 图 12 是用来说明在图 11 所示的第 2 再配置处理中执行的评价处理的典型的次序的流程图。
- [0030] 图 13 是用来说明第 2 再配置处理中的初次的再配置的概要的图。
- [0031] 图 14 是用来说明第 2 再配置处理中的初次以外的再配置的概要的图。
- [0032] 图 15 是用来说明在该实施方式的变形例中采用的第 2 再配置处理的典型的次序的流程图。
- [0033] 图 16 是用来说明该变形例中的第 2 再配置处理的概要的图。

具体实施方式

- [0034] 以下,参照附图对各种实施方式进行说明。
- [0035] 图 1 是表示有关一实施方式的计算机系统的典型的硬件结构的框图。计算机系统由层级化存储系统 10 及 n 台主机计算机(以下,称作主机)20_1 至 20_n 构成。在本实施方式中,n 是比 1 大的整数。即,计算机系统具备多个主机。但是,n 也可以是 1。即,计算机系统也可以具备单一的主机。
- [0036] 主机 20_1 至 20_n 将层级化存储系统 10 作为自身的外部存储装置使用。主机 20_1 至 20_n 例如经由交换式集线器(网络交换机)那样的交换机 30 与层级化存储系统 10(更详细地讲,层级化存储系统 10 的存储控制器 13)连接。在以下的说明中,在不需要将主机 20_1 至 20_n 特别区别的情况下,将主机 20_1 至 20_n 分别表述为主机 20_*。
- [0037] 层级化存储系统 10 具备高速存储装置(第 1 存储装置)11、低速存储装置(第 2 存储装置)12 和存储控制器 13。高速存储装置 11 及低速存储装置 12 经由与交换机 30 同样的交换机 14 与存储控制器 13 连接。
- [0038] 高速存储装置 11 例如由闪存阵列那样的访问应答性能良好的(即,访问速度为高速的)闪存存储装置构成。闪存存储装置使用多个闪存存储器板实现。闪存存储器板分别包括闪存存储器的集合。在本实施方式中,假设多个闪存存储器板分别不具有与硬盘驱动器(HDD)的互换性。但是,也可以代替多个闪存存储器板而使用具有与 HDD 的互换性的多个固态驱动器(SSD)。
- [0039] 另一方面,低速存储装置 12 例如由相比高速存储装置 11(闪存存储装置)访问应答性能较低(即,访问速度为低速)但存储容量较大的硬盘驱动器(HDD)阵列构成。HDD 阵

列例如是具备多个 HDD 的 RAID (Redundant Arrays of Inexpensive Disks 或 Redundant Arrays of Independent Disks) 结构的存储装置。另外, 高速存储装置 11 及低速存储装置 12 并不一定需要具有阵列结构。

[0040] 如上述那样, 在图 1 所示的层级化存储系统 10 中, 存在访问速度 (即访问应答性能) 不同的两个 (两种) 存储装置 (即, 高速存储装置 111 及低速存储装置 112)。在此情况下, 高速存储装置 111 被分配给上位层级 (高速层级、第 1 层级), 低速存储装置 112 被分配给下位层级 (低速层级、第 2 层级)。另外, 层级化存储系统 10 也可以具备例如磁带装置那样的比低速存储装置 112 更低速 (低层级、第 3 层级) 的存储装置 (第 3 层级的存储装置)。

[0041] 存储控制器 13 接受从主机 20_* 给出的、使用逻辑地址的数据访问 (读访问或写访问) 的请求 (输入输出请求), 执行被请求的输入输出 (I/O)。在该 I/O 的执行时, 存储控制器 13 使用周知的地址变换功能, 将逻辑地址变换为物理卷识别码 (ID) 及物理地址。逻辑地址是指由主机 20_* 识别的逻辑卷 (逻辑盘) 内的地址。物理地址表示与逻辑地址建立了对应的存储装置内的物理位置, 物理卷 ID 表示该存储装置的物理卷 (即, 物理的存储区域) 的 ID。存储控制器 13 基于该物理卷 ID 及物理地址对高速存储装置 11 或低速存储装置 12 访问。在本实施方式中, 高速存储装置 11 或低速存储装置 12 能够以第 2 尺寸的区域为最小单位进行访问。将该第 2 尺寸的区域称作块 (或扇区)。在本实施方式中, 第 2 尺寸是 512 字节 (B)。

[0042] 存储控制器 13 具备主接口控制器 (以下称作 HIFC) 130、存储接口控制器 (以下称作 SIFC) 131、存储器 132、HDD133 和 CPU134。

[0043] HIFC130 控制该 HIFC130 与主机 20_* 之间的数据转送 (数据转送协议)。HIFC130 接收从主机 20_* 对逻辑卷的 I/O 请求 (读访问请求或写访问请求), 返送对于该 I/O 请求的应答。HIFC130 如果从主机 20_* 接收到 I/O 请求, 则将该 I/O 请求向 CPU134 传递。接收到 I/O 请求的 CPU134 处理该 I/O 请求。

[0044] SIFC131 从 CPU134 接收基于 CPU134 接收到的来自主机 20_* 的 I/O 请求 (对于逻辑卷的写访问请求或读访问请求) 的访问请求 (对于物理卷的写访问请求或读访问请求)。SIFC131 根据接收到的访问请求, 执行向高速存储装置 111 或低速存储装置 113 的访问。

[0045] 存储器 132 是 DRAM 那样的可改写的易失性存储器。存储器 132 的一部分区域被用于保存从 HDD133 装载的存储控制程序 420 (图 4)。存储器 132 的另一部分区域被作为用于 CPU134 的工作区域使用。存储器 132 的再另一部分区域被用于保存各种表。它们的存储器映射图的详细情况后述。

[0046] 在 HDD133 中保存有存储控制程序 420。CPU134 通过在存储控制器 13 被起动机时执行初始化程序装载机 (IPL), 将保存在 HDD133 中的存储控制程序 420 向存储器 132 装载。IPL 保存在读出专用存储器 (ROM) 或闪存 ROM (FROM) 那样的非易失性存储器中。此外, 保存在存储器 132 中的各种表被适当保存在 HDD133 中。将保存在 HDD133 中的各种表根据需要 (例如在该存储控制器 13 的起动机时) 而从 HDD133 向存储器 132 装载。

[0047] CPU134 按照装载在存储器 132 中的存储控制程序 420, 作为后述的逻辑卷管理部 400、第 1 再配置部 401、第 2 再配置部 402、访问统计收集部 403 及评价部 404 发挥功能。即, CPU134 通过将保存在存储器 132 中的存储控制程序 420 执行, 控制层级化存储系统 10

整体（特别是存储控制器 13 内的各部）。

[0048] 在本实施方式中存储控制器 13 如图 1 所示,从主机 20_1 至 20_n 独立地装备。但是,存储控制器 13 也可以内置在主机 20_1 至 20_n 的某个（主机 20_*）中。在此情况下,存储控制器 13（更详细地讲,存储控制器 13 的功能）也可以使用主机 20_* 具有的操作系统 (OS) 的功能的一部分实现。

[0049] 此外,存储控制器 13 也可以装备到安装在主机 20_* 的卡插槽中而使用的卡中。此外也可以是,存储控制器 13 的一部分内置在主机 20_* 中,该存储控制器 13 的其余装备在卡中。此外,主机 20_*、存储控制器 13 和高速存储装置 11 及低速存储装置 12 的一部分或全部也可以收存在 1 个箱体中。

[0050] 图 2 是用来说明图 1 所示的层级化存储系统 10 的存储区域的典型的层级构造的图。高速存储装置 11 及低速存储装置 12 分别包括物理卷 110 及 120。在以下的说明中,将高速存储装置 11 的物理卷 110 称作高速物理卷 110,将低速存储装置 12 的物理卷 120 称作低速物理卷 120。

[0051] 高速物理卷 110 及低速物理卷 120 为了管理而被存储控制器 13 分割为第 1 尺寸的组块（即物理组块）。即,高速物理卷 110 及低速物理卷 120 分别具备多个组块。在本实施方式中,第 1 尺寸（即,物理组块的尺寸）是 4 千字节 (KB),是第 2 尺寸 (512B) 的 8 倍。即,1 个组块由 8 个扇区构成。当然,第 1 尺寸除了 4KB 以外,也可以是例如 4 兆字节 (MB)。

[0052] 高速物理卷 110 内的组块的群及低速物理卷 120 内的组块的群如图 2 所示那样被使用卷池 15 管理。卷池 15 用于将高速物理卷 110 内的组块的群及低速物理卷 120 内的组块的群逻辑地收容（集中）。在本实施方式中,卷池 15 由高速卷组 151 及低速卷组 152 构成。高速物理卷 110 内的组块的群被集中到卷池 15 的高速卷组 151 中,低速存储装置 12 内的组块的群被集中到卷池 15 的低速卷组 152 中。

[0053] 被集中到卷池 15 中的组块的群的例如一部分被分配给逻辑卷 16。逻辑卷 16 为了管理而被存储控制器 13 分割为与上述物理组块相同尺寸（第 1 尺寸）的组块（即逻辑组块）。即,逻辑卷 16 具备多个组块。

[0054] 逻辑卷 16 由映射区域 161 及自由空间 162 构成。逻辑卷 16 的映射区域 161 如在图 2 中用箭头 171 及 172 表示那样,用于将从卷池 15 的高速卷组 151 及低速卷组 152 选择的（集中的）组块的集合映射（分配）。自由空间 163 是指没有被分配物理组块的区域。

[0055] 逻辑卷 16 的状态能够从主机 20_* 经由存储控制器 13 识别。即,主机 20_* 能够将分配了高速存储装置 11 及低速存储装置 12 各自的物理组块的逻辑卷 16 识别为 1 个逻辑盘。并且,主机 20_* 在使用逻辑地址对逻辑卷 16 访问的情况下,能够不意识映射到由该逻辑地址指定的逻辑的存储区域中的物理的存储区域而对该物理的存储区域访问。

[0056] 为此,存储控制器 13（更详细地讲,存储控制器 13 的逻辑卷管理部 400）使用逻辑—物理组块变换表 421（图 4）管理逻辑卷 16 内的各个逻辑组块与映射在该各个逻辑组块中的物理组块的对应。保持在逻辑—物理组块变换表 421 的各条目中的信息是一种地址变换信息。存储控制器 13 在被从主机 20_* 送来用于读访问或写访问的 I/O 请求的情况下,将由该 I/O 请求指定的逻辑地址基于逻辑—物理组块变换表 421（进而,后述的物理组块管理表 422）向物理卷 ID 及物理地址变换。存储控制器 13 基于该物理卷 ID 及物理地址,对高速存储装置 11（高速物理卷 110）或低速存储装置 12（低速物理卷 120）即物理层的存储

(物理的存储区域)访问。

[0057] 在本实施方式中,在分配给逻辑卷 16 的映射区域 161 的第 1 集合及第 2 集合之间,以组块为单位移动数据。即,在高速存储装置 11(高速物理卷 110)与低速存储装置 12(低速物理卷 120)之间,以组块为单位移动数据。该数据移动在存储控制器 13 的控制下进行。即使进行这样的数据移动,只要逻辑-物理组块变换表 421 被正确地更新,存储控制器 13 就能够不使主机 20_* 注意该数据移动而执行该数据移动。

[0058] 图 3 是用来说明图 2 所示的逻辑卷 16 的典型的结构的图。另外,在图 3 中,将逻辑卷 16 的自由空间 162 省略。如图 3 所示,对逻辑卷 16 的映射区域 161 分配的组块的集合的一部分(以下称作第 1 集合)存在于高速物理卷 110 中,该组块的集合的其余(以下称作第 2 集合)存在于低速物理卷 120 中。

[0059] 在以往技术中,对逻辑卷 16 分配的高速物理卷 110 内的存储区域(即,第 1 集合的区域)被用于保存例如访问频度统计值(访问频度)较大(较高)的逻辑组块的数据。将这样的高速物理卷 110 内的存储区域称作高访问频度区域。

[0060] 相对于此,在本实施方式中,对逻辑卷 16 分配的高速物理卷 110 内的存储区域如图 3 所示,包括试验区域(以下称作 T 区域)111 及高访问频度区域(以下称作 HAF 区域)112 的两个存储区域。HAF 区域(第 1 存储区域)112 与以往技术同样,用于向访问频度统计值较大的逻辑组块的集合的分配。即,被分配 HAF 区域 112 的逻辑组块的条件(以下称作第 1 条件)是访问频度统计值更大。

[0061] 另一方面,T 区域(第 2 存储区域)111 被用于向基于与访问频度统计值不同的尺度选择的逻辑组块的集合的再分配和该再分配的效果的评价。在本实施方式中,从被分配了低速物理卷 120 内的存储区域(即,第 2 集合的区域)121 的逻辑卷 16 内的逻辑组块的群中选择被再分配 T 区域 111 的逻辑组块的集合。

[0062] 即,T 区域 111 被再分配的逻辑组块的条件(以下称作第 2 条件)是至少被分配低速物理卷 120 内的存储区域(第 3 存储区域)121。更详细地讲,T 区域 111 被再分配的逻辑组块的条件(以下称作第 2 条件)被分配低速物理卷 120 内的存储区域 121,并且 T 区域 111 是未分配。该低速物理卷 120 内的存储区域 121 用于将逻辑卷 16 的映射区域 161 内的逻辑组块群中的访问频度统计值(访问频度)较小(较低)的逻辑组块的数据保存。所以,将该存储区域 121 称作低访问频度区域(以下称作 LAF 区域)121。

[0063] 存储控制器 13(更详细地讲,存储控制器 13 的逻辑卷管理部 400)基于来自主机 20_* 的逻辑卷制作请求或经由存储控制器 13 的用户接口从用户给出的逻辑卷制作请求制作逻辑卷 16。在该逻辑卷 16 的制作时,存储控制器 13 制作包括高速存储装置 11 的高速物理卷 110 及低速存储装置 12 的低速物理卷 120 的卷池 15。如上述那样,卷池 15 由高速卷组 151 及低速卷组 152 构成。

[0064] 存储控制器 13 基于逻辑卷制作请求,从高速卷组 151 及低速卷组 152 分别切出第 1 组块的集合及第 2 组块的集合,向逻辑卷 16 的映射区域 161 分配。

[0065] 逻辑卷 16 中的映射区域 161 及自由空间 162 的比例能够从主机 20_*(或用户)指定。此外,对映射区域 161 分配的第 1 集合及第 2 集合的比例也能够从主机 20_*(或用户)指定。此外,高速物理卷 110 中的 T 区域 111 及 HAF 区域 112 的比例也能够从主机 20_*(或用户)指定。进而,逻辑卷 16 整体的存储容量也能够在不预先设定的上限值的范围内

从主机 20_*(或用户)指定。

[0066] 图 4 是主要表示图 1 所示的存储控制器 13 的典型的功能结构的框图。存储控制器 13 具备逻辑卷管理部 400、第 1 再配置部 401、第 2 再配置部 402、访问统计收集部 403 及评价部 404。这些功能要素 400 至 404 是通过图 1 所示的存储控制器 13 的 CPU134 执行存储控制程序 420 而实现的软件模组。但是,也可以将功能要素 400 至 404 的至少 1 个通过硬件模组实现。

[0067] 存储器 132 包括控制程序区域 410、表区域 411 及工作区域 412。控制程序区域 410 用于保存由 CPU134 执行的存储控制程序 420 的至少一部分。该存储控制程序 420 如上述那样预先保存在 HDD133 中,该存储控制程序 420 的至少一部分在存储控制器 13 的启动时被从该 HDD133 向控制程序区域 410 装载。

[0068] 表区域 411 用于将保存在 HDD133 中的各种表的至少一部分保存。工作区域 412 用于将 CPU134 在执行存储控制程序 420 时利用的临时的数据保存。

[0069] HDD133 除了上述的存储控制程序 420 以外,还保存逻辑—物理组块变换表 421、物理组块管理表 422、逻辑组块管理表 423、组块履历管理表 424、评价结果管理表 425 及访问计数表 426。这些表 421 至 426 例如在存储控制器 13 的启动时被从 HDD133 装载到存储器 132 的表区域 411 中而使用。因此,在装载在存储器 132 中的表被更新的情况下,该更新被适当反映到 HDD133 内的对应的表中。但是,在以后的说明中,为了简略化,设为表 421 至 426 在保存在 HDD133 中的状态下被使用。

[0070] 图 5 表示图 4 所示的逻辑—物理组块变换表 421 的数据构造的例子。逻辑—物理组块变换表 421 具有与图 2 所示的逻辑卷 16 的逻辑组块建立了对应的条目。逻辑—物理组块变换表 421 的各条目被用于保持逻辑组块号码与物理组块号码的对。逻辑组块号码表示对应的逻辑组块(即逻辑卷 16 内的逻辑组块)。物理组块号码表示分配给对应的逻辑组块的物理组块。

[0071] 在本实施方式中,层级化存储系统 10 具备包括逻辑卷 16 的多个逻辑卷。但是,在图 2 中,逻辑卷 16 以外的逻辑卷为了作图的方便而被省略。这样,在层级化存储系统 10 具备多个逻辑卷的本实施方式中,与多个逻辑卷分别建立对应而准备包括图 5 所示的逻辑—物理组块变换表 421 的多个逻辑—物理组块变换表。但是,只要在逻辑—物理组块变换表的条目信息中包含逻辑卷 ID,也可以代替多个逻辑—物理组块变换表而使用 1 个逻辑—物理组块变换表。逻辑卷 ID 是指包含保持在对应的条目中的逻辑组块号码表示的逻辑组块的逻辑卷。另外,层级化存储系统 10 也可以仅具备逻辑卷 16。

[0072] 图 6 表示图 4 所示的物理组块管理表 422 的数据构造的例子。物理组块管理表 422 具有与集中在卷池 15 中的物理组块建立了对应的条目。物理组块管理表 422 的各条目被用于保持物理组块号码、物理卷 ID、扇区号码和状态(状态标志)的组。

[0073] 物理组块号码表示对应的物理组块。物理卷 ID 表示包含对应的物理组块的物理卷。在本实施方式中,物理组块号码与对应的物理组块属于哪个物理卷(存储装置)无关,是唯一的。但是,物理组块号码也可以仅在对应的物理组块属于的物理卷内是唯一的。在此情况下,只要在逻辑—物理组块变换表 421 的条目信息中除了逻辑组块号码及物理组块号码以外还包含物理卷 ID 就可以。物理卷 ID 表示分配给对应的逻辑组块的物理组块属于的物理卷。

[0074] 接着,扇区号码表示构成对应的物理组块的扇区的列的开头扇区。状态表示有对于被分配了对应的物理组块的逻辑组块分配 T 区域 111 内的物理组块的情况(状态=1 的情况)还是没有(状态=0 的情况)。

[0075] 图 7 表示图 4 所示的逻辑组块管理表 423 的数据构造的例子。逻辑组块管理表 423 具有与构成层级化存储系统 10 具备的多个逻辑卷(更详细地讲,包括逻辑卷 16 的多个逻辑卷)的全部的逻辑组块分别建立了对应的条目。逻辑组块管理表 423 的各条目被用于保持逻辑组块号码、逻辑卷 ID 和扇区号码的组。逻辑组块号码表示对应的逻辑组块。逻辑卷 ID 表示包含对应的逻辑组块的逻辑卷。扇区号码表示构成对应的逻辑组块的扇区的列的开头扇区。

[0076] 图 8 表示图 4 所示的组块履历管理表 424 的数据构造的例子。组块履历管理表 424 被用于管理第 2 再配置处理的履历(以下称作组块履历)。所谓第 2 再配置处理,是指对于与低速物理卷 120 的 LAF 区域 121 内的物理组块建立了对应的逻辑组块再分配(建立对应于)高速物理卷 110 的 T 区域 111 内的物理组块的处理。第 2 再配置处理包括评价该再分配的效果的处理。

[0077] 组块履历管理表 424 具有与组块履历(第 2 履历)建立了对应的条目。组块履历管理表 424 的各条目被用于保持组块履历号码、日时信息和物理组块一览的组。组块履历号码是指组块履历的取得顺序。在本实施方式中,组块履历号码越小,表示是越新的组块履历。日时信息表示对应的组块履历被取得的日时。物理组块一览表示对于由对应的第 2 再配置处理再分配了 T 区域 111 内的物理组块的集合的逻辑组块,在该再分配前分别建立了对应的 LAF 区域 121 内的物理组块的集合。在本实施方式中,T 区域 111 由 4000 的物理组块构成。因而,物理组块一览表示 4000 的物理组块的集合。

[0078] 图 9 表示图 4 所示的评价结果管理表 425 的数据构造的例子。评价结果管理表 425 被用于管理在组块履历取得时的第 2 再配置处理中被执行的评价处理的结果(评价结果)的履历。

[0079] 评价结果管理表 425 具有与评价结果的履历(第 1 履历)建立了对应的条目。评价结果管理表 425 的各条目被用于保持评价结果号码、日时信息和吞吐量信息的组。评价结果号码是指评价结果(评价结果的履历)的取得顺序,与对应的组块履历的组块履历号码一致。在本实施方式中,评价结果号码越小,表示越新的评价结果。日时信息表示对应的评价结果(组块履历)被取得的日时。吞吐量信息表示对应的评价结果。在本实施方式中,作为对应的评价结果而使用吞吐量。关于该吞吐量在后面叙述。

[0080] 接着,参照图 10 对在本实施方式中采用的第 1 再配置处理进行说明。图 10 是用来说明由存储控制器 13(更详细地讲,存储控制器 13 的第 1 再配置部 401)执行的第 1 再配置处理的典型的次序的流程图。第 1 再配置处理例如包括在第 1 期间(监视期间)中按照组块(例如逻辑组块)取得访问频度统计信息(访问频度统计值)的处理。第 1 再配置处理还包括基于所取得的访问频度统计值向访问频度较高的逻辑组块再分配高速物理卷 110 的 HAF 区域 112 内的物理组块的处理。第 1 再配置处理例如以与第 1 期间对应的周期定期地执行。

[0081] 在本实施方式中,将第 1 再配置处理及第 2 再配置处理独立地执行。因此,第 1 再配置处理及第 2 再配置处理相互被排他地执行。即,在某一方的再配置处理的执行中,另一

方的再配置处理为等待状态,在执行中的再配置处理的完成后,另一方的再配置处理的等待状态被解除。

[0082] 所以,存储控制器 13 的第 1 再配置部 401 首先判定第 2 再配置处理是否是执行中(步骤 S1)。如果第 2 再配置处理是执行中(步骤 S1 的 Yes),则第 1 再配置部 401 等待第 2 再配置处理的完成。

[0083] 相对于此,如果第 2 再配置处理不是执行中(步骤 S1 的 No),则第 1 再配置部 401 将访问统计收集部 403 起动。于是,访问统计收集部 403 监视第 1 期间中的向各个逻辑组块的访问,取得该向各个逻辑组块的例如访问的频度(步骤 S2)。如果更具体地记述,则访问统计收集部 403 使用访问计数表 426 收集向各个逻辑组块的访问的频度,作为访问频度统计信息(访问频度值)。该向各个逻辑组块的访问的频度指示表示向该各个逻辑组块的访问的状况的指标。访问计数表 426 的各条目被用于保持表示向对应的逻辑组块的访问的次数(例如,数据读的次数与数据写的次数的合计)的访问计数值。该访问计数值每当对应的逻辑组块被访问而增加 1。另外,访问统计收集部 403 也可以收集向各个逻辑组块的访问的数据量的总和作为访问频度统计信息。

[0084] 如果监视期间经过,即如果由访问统计收集部 403 取得访问频度统计信息(步骤 S2),则第 1 再配置部 401 基于该访问频度统计信息(访问计数表 426),选择访问频度最高的逻辑组块,即访问计数值(访问频度统计值)最大的逻辑组块(步骤 S3)。接着,第 1 再配置部 401 如以下这样判定对于所选择的逻辑组块(以下称作第 1 逻辑组块)是否分配了低速物理卷 120(LAF 区域 121)内的物理组块(步骤 S4)。

[0085] 首先,第 1 再配置部 401 基于第 1 逻辑组块的逻辑组块号码参照逻辑-物理组块变换表 421。并且,第 1 再配置部 401 确定与第 1 逻辑组块建立了对应的物理组块的物理组块号码。接着,第 1 再配置部 401 基于所确定的物理组块号码参照物理组块管理表 422。并且,第 1 再配置部 401 确定与所确定的物理组块号码成组的物理卷 ID。基于该确定的物理卷 ID,第 1 再配置部 401 判定是否对第 1 逻辑组块分配了低速物理卷 120 内(即,低速区域侧)的物理组块。

[0086] 如果对第 1 逻辑组块分配了低速物理卷 120 内的物理组块(步骤 S4 的 Yes),则第 1 再配置部 401 向步骤 S5 前进。在步骤 S5 中,第 1 再配置部 401 从被分配了 HAF 区域 112(更详细地讲,HAF 区域 112 内的物理组块)的未选择的逻辑组块中,如以下这样选择访问频度最低、即访问计数值(访问频度统计值)最小的逻辑组块。首先,第 1 再配置部 401 基于包含 HAF 区域 112 的高速物理卷 110 的物理卷 ID,参照物理组块管理表 422。由此,第 1 再配置部 401 确定与高速物理卷 110 的物理卷 ID 成组的物理组块号码。接着,第 1 再配置部 401 参照逻辑-物理组块变换表 421,确定与所确定的物理组块号码成组的未选择的逻辑组块号码。接着,第 1 再配置部 401 参照访问计数表 426,从所确定的逻辑组块号码表示的未选择的逻辑组块中选择访问频度最低的逻辑组块(以下称作第 2 逻辑组块)。

[0087] 接着,第 1 再配置部 401 将第 1 逻辑组块及第 2 逻辑组块(即所选择的两组块)的数据的配置目标掉换(步骤 S6)。更具体地讲,第 1 再配置部 401 将分配给第 1 逻辑组块的 LAF 区域 121 内的物理组块(以下称作第 1 物理组块)的数据与分配给第 2 逻辑组块的 HAF 区域 112 内的物理组块(以下称作第 2 物理组块)的数据掉换。即,第 1 再配置部 401 将第 1 逻辑组块的数据再配置给第 2 物理组块,将第 2 逻辑组块的数据再配置给第 1 物理

组块。

[0088] 接着,第 1 再配置部 401 基于上述的数据配置目标的掉换(再配置),将与第 1 逻辑组块及第 2 逻辑组块分别建立了对应的逻辑-物理组块变换表 421 的条目的信息更新(步骤 S7)。即,第 1 再配置部 401 将与第 1 逻辑组块及第 2 逻辑组块分别建立了对应的逻辑-物理组块变换表 421 的条目的信息,以反映上述数据配置目标的掉换。更详细地讲,第 1 再配置部 401 将与第 1 逻辑组块建立了对应的逻辑-物理组块变换表 421 的条目中的物理组块号码从表示第 1 物理组块的号码更新为表示第 2 物理组块的号码。同样,第 1 再配置部 401 将与第 2 逻辑组块建立了对应的逻辑-物理组块变换表 421 的条目中的物理组块号码从表示第 2 物理组块的号码更新为表示第 1 物理组块的号码。

[0089] 第 1 再配置部 401 如果执行步骤 S7,则向步骤 S8 前进。此外,在没有对第 1 逻辑组块分配低速物理卷 120 内的物理组块的情况下(步骤 S4 的 No),第 1 再配置部 401 也向步骤 S8 前进。在步骤 S8 中,第 1 再配置部 401 判定是否选择了与 HAF 区域 112 对应的数量的逻辑组块。所谓与 HAF 区域 112 对应的数量的逻辑组块,是指与构成该 HAF 区域 112 的物理组块的数量相同数量的逻辑组块。

[0090] 如果没有选择与 HAF 区域 112 对应的数量的逻辑组块(步骤 S8 的 No),则第 1 再配置部 401 向步骤 S9 前进。在步骤 S9 中,第 1 再配置部 401 基于访问计数表 426,选择访问频度接着较高的逻辑组块作为第 2 逻辑组块。并且,第 1 再配置部 401 向步骤 S4 返回。

[0091] 这样,第 1 再配置部 401 将被分配了 LAF 区域 121 内的物理组块的逻辑组块中的访问频度较高的逻辑组块的数据的配置目标、与配置在 HAF 区域 112 中的逻辑组块中的访问频度较低的逻辑组块的数据的配置目标掉换。如果最终选择了与 HAF 区域 112 对应的数量的逻辑组块(步骤 S8 的 Yes),则第 1 再配置部 401 向步骤 S10 前进。在步骤 S10 中,第 1 再配置部 401 判定是否由用户指定了访问频度统计信息(即访问计数表 426)的初始化。

[0092] 如果没有指定初始化(步骤 S10 的 Yes),则第 1 再配置部 401 将访问频度统计信息初始化(步骤 S10),结束第 1 再配置处理。相对于此,如果没有指定初始化(步骤 S10 的 No),则第 1 再配置部 401 将步骤 S10 跳过,结束第 1 再配置处理。将第 1 再配置处理如上述那样以对应于第 1 期间的周期定期地执行。

[0093] 在上述说明中,为了简略化,设想了第 1 逻辑组块的访问频度总是比第 2 逻辑组块的访问频度高的情况。但实际上,也可能有第 1 逻辑组块的访问频度比第 2 逻辑组块的访问频度低的情况。在这样的情况下,第 1 再配置部 401 只要将步骤 S6 至 S8 跳过而向步骤 S10 前进就可以。

[0094] 接着,参照图 11 及图 12,对在本实施方式中采用的第 2 再配置处理进行说明。图 11 是用来说明由存储控制器 13(更详细地讲,存储控制器 13 的第 2 再配置部 402)执行的第 2 再配置处理的典型的次序的流程图。图 12 是用来说明在图 11 所示的第 2 再配置处理中执行的评价处理的典型的次序的流程图。

[0095] 存储控制器 13 的第 2 再配置部 402 首先判定第 1 再配置处理是否是执行中(步骤 S21)。如果第 1 再配置处理是执行中(步骤 S21 的 Yes),则第 2 再配置部 402 等待第 1 再配置处理的完成。相对于此,如果第 1 再配置处理不是执行中(步骤 S21 的 No),则第 2 再配置部 402 向步骤 S22 前进。在步骤 S22 中,第 2 再配置部 402 判定是否是第 2 再配置处理中的初次的再配置(再分配)。

[0096] 如果是第 2 再配置处理中的初次的再配置 (步骤 S22 的 Yes), 则第 2 再配置部 402 向步骤 S23 前进。在这样是初次的再配置的情况下, 在本实施方式中, T 区域 111 内的物理组块都没有被分配给逻辑组块。在步骤 S23 中, 第 2 再配置部 402 例如从 LAF 区域 121 的开头选择与 T 区域 111 对应的数量 Q 的物理组块。所谓与 T 区域 111 对应的数量 Q, 是指构成该 T 区域 111 的物理组块的数量。这里, T 区域 111 具备 4,000 个物理组块, 因而, 假设 Q 是 4,000。

[0097] 接着, 第 2 再配置部 402 将所选择的 Q 个物理组块 (以下称作第 3 物理组块) 的数据向 T 区域 111 拷贝 (步骤 S24)。接着, 第 2 再配置部 402 基于该拷贝, 将逻辑-物理组块变换表 421 的条目的信息更新 (步骤 S25)。即, 第 2 再配置部 402 在逻辑-物理组块变换表 421 中, 将与所选择的 Q 个第 3 物理组块建立对应的逻辑组块号码变更。更详细地讲, 第 2 再配置部 402 在逻辑-物理组块变换表 421 中, 将与所选择的 Q 个第 3 物理组块到此为止建立了对应的 Q 个逻辑组块 (以下称作第 3 逻辑组块) 建立对应的物理组块号码, 从表示第 3 物理组块的号码更新为表示 T 区域 111 内的 Q 个物理组块 (以下称作第 4 物理组块) 的号码。这样, 第 2 再配置部 402 通过对于与所选择的 Q 个第 3 物理组块分别建立了对应的 Q 个第 3 逻辑组块, 将 T 区域 111 内的 Q 个第 4 物理组块建立对应, 进行物理组块的再分配。进而, 在步骤 S25 中, 第 2 再配置部 402 将与第 3 物理组块建立了对应的物理组块管理表 422 的条目中的状态设定为 1。

[0098] 第 2 再配置部 402 如果执行步骤 S25, 则将评价部 404 起动。于是, 评价部 404 评价从 LAF 区域 121 向 T 区域 111 的物理组块的数据的拷贝 (步骤 S24) 的效果。即, 评价部 404 按照图 12 所示的流程图, 如以下这样评价将对所选择的逻辑组块分配的物理组块从 LAF 区域 121 内的物理组块变更为 T 区域 111 内的物理组块的再分配的效果 (步骤 S26)。

[0099] 首先, 评价部 404 设定块 (I/O 块) 的尺寸 (第 2 尺寸) (步骤 S41)。如上述那样, 块是向高速存储装置 11 及低速存储装置 12 的访问的最小单位。接着, 评价部 404 从在层级化存储系统 10 (存储控制器 13) 内准备的系统时钟取得当前的时刻 (即系统时刻) (步骤 S42)。

[0100] 接着, 评价部 404 执行向逻辑卷 16 的随机读访问 (步骤 S43)。评价部 404 基于从该随机读访问的开始时点到结束时点的时间 [s] 和由该随机读访问读出的数据的量 [MB], 计算该随机读访问的吞吐量 [MB/s] (步骤 S44)。并且, 评价部 404 向组块履历管理表 424 及评价结果管理表 425 如以下这样追加条目信息 (步骤 S45)。

[0101] 首先, 评价部 404 生成包含组块履历号码 (= 1)、表示在步骤 S42 中取得的系统时刻 (日时) 的日时信息、在步骤 S23 中选择的 Q 个第 3 物理组块的物理组块号码的条目信息 (组块履历信息)。此时, 假设在组块履历管理表 424 中存在 1 个以上有效的条目 (例如多个条目)。在此情况下, 第 1 再配置部 401 将多个条目各自的组块履历信息向后续的条目移动, 并将各组块履历信息中的组块履历号码增加 1。并且, 评价部 404 将所生成的组块履历信息设定到组块履历管理表 424 的开头条目中。另外, 也可以是将组块履历号码根据对应的组块履历信息的生成而顺序地生成、将第 i 个 ($i = 1, 2, 3, \dots$) 生成的组块履历信息保持在组块履历管理表 424 的第 i 个条目中的结构。

[0102] 此外, 评价部 404 生成包含评价结果号码 (= 1)、表示在步骤 S42 中取得的系统时刻 (日时) 的日时信息、和表示在步骤 S44 中计算出的吞吐量的吞吐量信息的条目信息

(评价结果信息)。评价部 404 与将组块履历信息向组块履历管理表 424 追加同样,将所生成的评价结果信息向评价结果管理表 425 追加。评价部 404 如果了执行步骤 S45,则结束评价处理(步骤 S26)。

[0103] 于是,第 2 再配置部 402 判定在 LAF 区域 121 内是否有未选择的物理组块(即,在步骤 S23 中没有被选择的物理组块)(步骤 S27)。如果有未选择的物理组块(步骤 S27 的 Yes),则第 2 再配置部 402 向步骤 S22 返回。在步骤 S22 中,第 2 再配置部 402 如上述那样,判定是否是第 2 再配置处理中的初次的再配置。

[0104] 如果如该例那样不是初次的再配置(步骤 S22 的 No),则为了下个再配置,第 2 再配置部 402 向步骤 S28 前进。在步骤 S28 中,第 2 再配置部 402 将 T 区域 111 内的全部物理组块(即 Q 个第 4 物理组块)的数据向低速存储装置 12 的原来的区域移动。所谓原来的区域,是指在最近的步骤 S24 中数据被拷贝到 T 区域 111 内的 Q 个物理组块中时、该数据曾被保存的 LAF 区域 121 内的 Q 个物理组块(以下称作第 5 物理组块)的区域,即在最近的步骤 S24 中执行的拷贝的拷贝源。该 LAF 区域 121 内的原来的区域(即,在步骤 S28 中执行的移动的移动目标)基于组块履历管理表 424 的开头条目(即,组块履历号码为 1 的条目)的组块一览确定。这样,在步骤 S28 中,第 2 再配置部 402 对于被分配了当前 T 区域 111 内的 Q 个第 4 物理组块的 Q 个逻辑组块(以下称作第 4 逻辑组块)再分配 LAF 区域 121 内的原来的区域。

[0105] 接着,第 2 再配置部 402 从 LAF 区域 121 的后续位置选择 Q 个物理组块(步骤 S29)。所谓 LAF 区域 121 的后续位置,是指对在步骤 S28 中 T 区域 111 的数据被移动的区域后续的位置。接着,第 2 再配置部 402 与在初次的再配置中执行步骤 S23 的情况同样,如以下这样执行步骤 S24 及 S25。

[0106] 首先,第 2 再配置部 402 将在步骤 S29 中选择的 Q 个物理组块(以下称作第 3 物理组块)的数据向 T 区域 111 拷贝(再配置)(步骤 S24)。即,第 2 再配置部 402 向与在步骤 S29 中选择的 Q 个第 3 物理组块分别建立了对应的 Q 个逻辑组块(以下称作第 3 逻辑组块)再分配 T 区域 111 内的 Q 个物理组块(第 4 物理组块)。

[0107] 接着,第 2 再配置部 402 基于上述拷贝(再分配),将与所选择的 Q 个第 3 物理组块建立了对应的逻辑—物理组块变换表 421 的条目的信息(即与 Q 个第 3 逻辑组块建立了对应的逻辑—物理组块变换表 421 的条目的信息)更新(步骤 S25)。更详细地讲,第 2 再配置部 402 将与 Q 个第 3 逻辑组块分别建立了对应的逻辑—物理组块变换表 421 的条目中的物理组块号码从表示第 3 物理组块的号码更新为表示第 4 物理组块的号码。此外,在步骤 S25 中,第 2 再配置部 402 将与被分配了 Q 个第 4 物理组块的 Q 个第 4 逻辑组块建立了对应的逻辑—物理组块变换表 421 的条目的信息也更新。更详细地讲,第 2 再配置部 402 将与 Q 个第 4 逻辑组块分别建立了对应的逻辑—物理组块变换表 421 的条目中的物理组块号码从表示第 4 物理组块的号码更新为表示第 5 物理组块的号码。

[0108] 由此,例如如果是第 4 次再配置,则首先,将 T 区域 111 内的全部物理组块的数据向由图 8 的组块履历管理表 424 的开头条目中的物理组块一览表示的 LAF 区域 121 内的物理组块 8,000 至 11,999(即,物理组块号码为 8,000 至 11,999 的物理组块)移动(步骤 S28)。即,对于被分配了 T 区域 111 内的 4,000 个物理组块的 4,000 个逻辑组块,分别再分配物理组块 8,000 至 11,999。

[0109] 接着,将LAF区域121内的物理组块12,000至15,999(即,物理组块号码为12,000至15,999的物理组块)的数据向T区域111拷贝(步骤S29及S24)。即,对于被分配了LAF区域121内的物理组块12,000至15,999的4,000个逻辑组块,分别再分配T区域111内的4,000个物理组块。接着,再次执行步骤S26及S27。

[0110] 第2再配置部402重复上述步骤S22、S28、S29及S24至27。如果最终选择了LAF区域121内的全部物理组块的结果是不再有未选择的物理组块(步骤S27的No),则第2再配置部402向步骤S30前进。在步骤S30中,第2再配置部402基于组块履历管理表424及评价结果管理表425,如以下这样选择评价最高的(即吞吐量最大的)物理组块的组(步骤S31)。首先,第2再配置部402参照评价结果管理表425,确定表示吞吐量为最大的评价结果的评价结果号码。接着,第2再配置部402参照组块履历管理表424,确定包含与所确定的评价结果号码一致的组块履历号码的条目。并且,第2再配置部402选择设定在所确定的组块履历管理表424的条目中的物理组块一览表表示的物理组块的组(Q个物理组块),作为评价最高的物理组块的组。

[0111] 接着,第2再配置部402将所选择的Q个物理组块的数据向T区域111再配置(拷贝)(步骤S31)。即,第2再配置部402通过将分配了所选择的Q个物理组块的Q个逻辑组块的数据向T区域111再配置,将评价最高的再配置的状态复原。虽然在图11的流程图中被省略,但第2再配置部402在该再配置(步骤S31)的前后,分别执行与步骤S28及S25对应的处理。另外,第2再配置部402也可以将评价结果管理表425表示的评价结果的一览经由用户接口向用户提示,使用户从该评价结果的一览中选择希望的评价结果。在此情况下,第2再配置部402只要选择能够得到由用户选择的评价结果的物理组块的组就可以。

[0112] 第2再配置部402如果执行步骤S31,则判定是否由用户指定了关于第2再配置处理的状态(更详细地讲,组块履历管理表424及评价结果管理表425)(步骤S32)。如果指定了状态的初始化(步骤S32的Yes),则第2再配置部402将该状态初始化(步骤S33),结束第2再配置处理。相对于此,如果没有指定状态的初始化(步骤S32的No),则第2再配置部402将步骤S33跳过,结束第2再配置处理。在此情况下,在接着起动了第2再配置处理的情况下,继续使用当前的组块履历管理表424及评价结果管理表425。

[0113] 在本实施方式中,第2再配置部402将第2再配置处理中的上述再配置(例如,步骤S22、S28、S29及S24至27)定期地(例如每1小时)执行。但是,也可以将执行该再配置的周期基于该评价处理中的评价结果是否比前次好来变更。具体而言,在评价结果比前次好的情况下,也可以为了将该状态较长地维持,第2再配置部402使开始下个再配置的时期延迟。相反,在评价结果比前次差的情况下,第2再配置部402也可以将开始下个再配置的时期提早。此外,第2再配置处理既可以例如根据用户的指示适当执行,或者也可以定期地执行。

[0114] 图13及图14是用来说明本实施方式的第2再配置处理的概要的图。图13是用来说明第2再配置处理的初次的再配置的概要的图,图14是用来说明第2再配置处理中的初次以外的再配置的概要的图。

[0115] 在初次的再配置中(步骤S22的Yes),从低速物理卷120的LAF区域121的开头选择Q个物理组块(步骤S23)。并且,将所选择的Q个物理组块的数据如在图13中用箭头A1表示那样向高速物理卷110的T区域111拷贝(再配置)(步骤S24)。另外,在图13及

14 中,为了作图的方便而表示 $Q = 5$ 的情况。

[0116] 在初次以外的再配置中(步骤 S22 的 No),首先,将 T 区域 111 的数据如在图 14 中用箭头 A11 表示那样,向 LAF 区域 121 的原来的区域移动(再配置)(步骤 S28)。接着,从 LAF 区域 121 的后续位置的开头选择 Q 个物理组块(步骤 S29)。并且,将所选择的 Q 个物理组块的数据如在图 14 中用箭头 A12 表示那样,向 T 区域 111 拷贝(再配置)(步骤 S24)。

[0117] 这样,在本实施方式中,第 2 再配置部 402 将在基于访问频度(访问频度统计值)的再配置中不会被配置到高速物理卷 110 侧那样的物理组块的组(低速物理卷 120 侧的物理组块的组)的数据向该高速物理卷 110 内的 T 区域 111 再配置(步骤 S24)。在此状态下,评价部 404 评价该再配置(即,从 LAF 区域 121 向 T 区域 111 的数据的再配置)的效果(步骤 S26)。即,根据本实施方式,能够将在基于访问频度的再配置中不会被选择那样的(即,不能确认效果那样的)物理组块的组的数据向高速物理卷 110 内的 T 区域 111 再配置,评价其再配置的效果。此外,根据本实施方式,将这样的再配置及评价按照 LAF 区域 121 内的连续的 Q 个物理组块执行,将其再配置及评价的履历分别向组块履历管理表 424 及评价结果管理表 425 记录。由此,根据本实施方式,能够基于组块履历管理表 424 及评价结果管理表 425 确定与评价最高的配置对应的 Q 个物理组块,将该确定的 Q 个物理组块的数据再配置(复原)到 T 区域 111 中。由此,能够进一步提高层级化存储系统 10 的性能。

[0118] 另外,评价部 404 为了评价从 LAF 区域 121 向 T 区域 111 的逻辑组块的数据的再配置的效果,也可以代替向逻辑卷 16 的随机读访问,而执行向逻辑卷 16 内的任意的存储区域(第 4 存储区域)的顺序读访问。在此情况下,评价部 404 只要取得顺序读访问中的吞吐量作为评价结果就可以。这里,作为任意的存储区域,也可以指定特别被要求顺序读访问的性能提高的存储区域、例如保存有特定的备份数据的存储区域。此外,上述任意的存储区域也可以是逻辑卷 16 整体。进而,评价部 404 也可以分别执行上述随机读访问和顺序读访问,取得该两访问中的吞吐量的例如平均值作为评价结果。

[0119] < 变形例 >

[0120] 接着,对上述实施方式的变形例进行说明。在上述实施方式中,在第 2 配置处理中,从低速物理卷 120 的 LAF 区域 121,以 Q 个物理组块为单位,顺序地选择物理组块。并且,每当选择 Q 个物理组块,就将该选择的 Q 个物理组块的数据向高速物理卷 110 的 T 区域 111 再配置。相对于此,在本变形例中,通过与上述实施方式不同的方法选择向 T 区域 111 再配置数据的 Q 个物理组块。

[0121] 首先,假设基于访问频度,从低速物理卷 120 的 LAF 区域 121 内的物理组块向高速物理卷 110 的 HAF 区域 112 内的物理组块配置(拷贝)数据。将这样的情况下的 LAF 区域 121 内的物理组块称作高访问频度物理组块。本变形例的特征在于,按照新从 HAF 区域 112 对 T 区域 111 再配置了数据的高访问频度物理组块,将该周边的 Q 个物理组块的数据向 T 区域 111 再配置。

[0122] 以下,参照图 15 对在本变形例中采用的第 2 再配置处理(即,第 2 再配置处理的变形例)进行说明。图 15 是用来说明在本变形例中由第 2 再配置部 402 执行的第 2 再配置处理的典型的次序的流程图。在本变形例中,假设将第 2 再配置处理对应于第 1 再配置处理的结束而执行。此外,在第 1 再配置处理中,第 1 再配置部 401 生成从 LAF 区域 121 新将数据配置到 HAF 区域 112 中的物理组块的一览。

[0123] 在第 2 再配置处理中,第 2 再配置部 402 取得由第 1 再配置部 401 生成的物理组块的一览(步骤 S51)。接着,第 2 再配置部 402 将变量 p 及 q 设定为初始值 1(步骤 S52)。变量 p 表示 LAF 区域 121 中的以第 i 个物理组块 CHK_i 为基准的物理组块的相对位置。例如,物理组块 CHK_{i-p} 及 CHK_{i+p} 表示 LAF 区域 121 中的第 $i-p$ 个及第 $i+p$ 个物理组块。即,物理组块 CHK_{i-p} 及 CHK_{i+p} 相对于物理组块 CHK_i 分别先行及后续 p 个组块。变量 q 表示 T 区域 111 中的物理组块的相对位置。例如,物理组块 CHK_q 表示 T 区域 111 中的第 q 个物理组块。

[0124] 接着,第 2 再配置部 402 基于所取得的物理组块的一览,选择物理组块(第 6 的物理组块) CHK_i (步骤 S53)。接着,第 2 再配置部 402 基于所选择的物理组块 CHK_i 和变量 p,从 LAF 区域 121 中选择物理组块 CHK_{i-p} (步骤 S54)。接着,第 2 再配置部 402 参照与所选择的物理组块 CHK_{i-p} 建立了对应的物理组块管理表 422 的条目中的状态,判定该状态(即,物理组块 CHK_{i-p} 的状态)是否表示未配置(步骤 S55)。

[0125] 如果是未配置(步骤 S55 的 Yes),则第 2 再配置部 402 从 T 区域 111 中选择第 q 个物理组块 CHK_q (步骤 S56)。接着,第 2 再配置部 402 将所选择的物理组块 CHK_q 的数据与上述实施方式的步骤 S28 同样,向低速存储装置 12(低速物理卷 120)内的原来的区域移动(步骤 S57)。

[0126] 接着,第 2 再配置部 402 将所选择的物理组块 CHK_{i-p} 的数据向所选择的物理组块 CHK_q 拷贝(步骤 S58)。这样,第 2 再配置部 402 将物理组块(第 3 物理组块) CHK_{i-p} 的数据向 T 区域 111 内的物理组块(第 4 物理组块) CHK_q 再配置。即,第 2 再配置部 402 向被分配了物理组块 CHK_{i-p} 的逻辑组块(第 3 逻辑组块)再分配物理组块 CHK_q 。

[0127] 接着,第 2 再配置部 402 将逻辑—物理组块变换表 421 更新,以反映上述数据移动(步骤 S57)及数据拷贝(步骤 S58)(步骤 S59)。在步骤 S59 中,第 2 再配置部 402 将与物理组块 CHK_{i-p} 建立了对应的物理组块管理表 422 的条目中的状态设定为 1。接着,第 2 再配置部 402 将变量 q 增加 1(步骤 S60)。并且,第 2 再配置部 402 判定增加后的变量 q 是否超过 Q(步骤 S61)。Q 如上述那样表示 T 区域 111 中的物理组块的数量。

[0128] 如果增加后的变量 q 不超过 Q(步骤 S61 的 No),则第 2 再配置部 402 为了选择下个物理组块而向步骤 S62 前进。此外,在物理组块 CHK_{i-p} 的状态表示已配置的情况下(步骤 S55 的 No),第 2 再配置部 402 也向步骤 S62 前进。在步骤 S62 中,第 2 再配置部 402 从 LAF 区域 121 选择物理组块 CHK_{i+p} 。接着,第 2 再配置部 402 参照与所选择的物理组块 CHK_{i+p} 建立了对应的物理组块管理表 422 的条目中的状态,判定该状态(即,物理组块 CHK_{i+p} 的状态)是否表示未配置(步骤 S63)。

[0129] 如果是未配置(步骤 S63 的 Yes),则第 2 再配置部 402 从 T 区域 111 中选择第 q 个物理组块 CHK_q (步骤 S64)。接着,第 2 再配置部 402 将所选择的物理组块 CHK_q 的数据向低速存储装置 12 内的原来的区域移动(步骤 S65)。

[0130] 接着,第 2 再配置部 402 将所选择的物理组块 CHK_{i+p} 的数据向物理组块 CHK_q 拷贝(步骤 S66)。这样,第 2 再配置部 402 将物理组块 CHK_{i+p} 的数据向 T 区域 111 内的物理组块 CHK_q 再配置。即,第 2 再配置部 402 对被分配了物理组块 CHK_{i+p} 的逻辑组块再分配物理组块 CHK_q 。

[0131] 接着,第 2 再配置部 402 将逻辑-物理组块变换表 421 更新,以反映上述数据移动(步骤 S65)及数据拷贝(步骤 S66)(步骤 S67)。在步骤 S67 中,第 2 再配置部 402 将与物理组块 CHK_i+p 建立了对应的物理组块管理表 422 的条目中的状态设定为 1。接着,第 2 再配置部 402 将变量 q 增加 1(步骤 S68)。并且,第 2 再配置部 402 判定增加后的变量 q 是否超过 Q (步骤 S69)。

[0132] 如果增加后的变量 q 没有超过 Q (步骤 S69 的 No),则第 2 再配置部 402 为了选择下个物理组块而向步骤 S70 前进。此外,在物理组块 CHK_i+p 的状态表示已配置的情况下(步骤 S63 的 No),第 2 再配置部 402 也向步骤 S70 前进。在步骤 S70 中,第 2 再配置部 402 将变量 p 增加 1。并且,第 2 再配置部 402 回到步骤 S54,基于在步骤 S43 中选择的物理组块 CHK_i 和增加后的变量 p ,从 LAF 区域 121 中选择物理组块 CHK_i-p 。这样,第 2 再配置部 402 重复步骤 S54 至 S70,直到增加后的变量 q 超过 Q (步骤 S61 或 S69 的 Yes)。

[0133] 假设最终增加后的变量 q 超过了 Q (步骤 S61 或 S69 的 Yes)。在此情况下,第 2 再配置部 402 判断为包含在 LAF 区域 121 中的物理组块 CHK_i 的周边(即前后)的 Q 个物理组块的数据被再配置到 T 区域 111 中。所以,第 2 再配置部 402 为了评价该再配置的效果,执行与上述实施方式的评价处理(步骤 S26)同样的评价处理(步骤 S71)。

[0134] 第 2 再配置部 402 如果执行评价处理(步骤 S71),则判定为在由步骤 S51 取得的物理组块的一览中有未选择的物理组块(步骤 S72)。如果有未选择的物理组块(步骤 S72 的 Yes),则第 2 再配置部 402 回到步骤 S53,选择未选择的物理组块 CHK_i 。这样,第 2 再配置部 402 重复步骤 S53 至 S72,直到没有未选择的物理组块(步骤 S72 的 No)。

[0135] 假设最终不再有未选择的物理组块(步骤 S72 的 No),在此情况下,第 2 再配置部 402 执行图 11 的步骤 S30 以后的处理。另外,在最初选择了物理组块 CHK_i 的情况下,第 2 再配置部 402 不需要在将该物理组块 CHK_i 的周边的 Q 个物理组块的数据向 T 区域 111 再配置前,将该 T 区域 111 的数据向低速存储装置 12 的原来的区域移动。但是,在图 15 的流程图中,为了简略化,最初选择了物理组块 CHK_i 的情况下的处理次序被省略。

[0136] 此外,虽然在图 15 的流程图中被省略,但在物理组块 CHK_i-p 不在 LAF 区域 121 的情况下,第 2 再配置部 402 将步骤 S55 至 S61 跳过。同样,在物理组块 CHK_i+p 不在 LAF 区域 121 的情况下,第 2 再配置部 402 将步骤 S63 至 S69 跳过。此外,在物理组块 CHK_i-p 及 CHK_i+p 都不在 LAF 区域 121 的情况下,第 2 再配置部 402 仅将 T 区域 111 内的已再配置的数据向 LAF 区域 121 的原来的区域送回,执行图 11 的步骤 S30 以后的处理。

[0137] 图 16 是用来说明本变形例的第 2 再配置处理的概要的图。在图 16 中,假设选择了 LAF 区域 121 内的物理组块 CHK_i (步骤 S53)。物理组块 CHK_i 的数据在先行于此次第 2 再配置处理的第 1 配置处理中,如在图 16 中用箭头 A21 表示那样被与 HAF 区域 112 内的物理组块 CHK_j 的数据掉换(步骤 S6)。

[0138] 将 T 区域 111 内的 Q 个物理组块的数据如在图 16 中用箭头 A22 表示那样向 LAF 区域 121 内的原来的区域移动(再配置)(步骤 S57 及 S65)。另一方面,将所选择的物理组块 CHK_i 的周边的 Q 个物理组块的数据如在图 16 中用箭头 A23 及 A24 表示那样向 T 区域 111 内的 Q 个物理组块拷贝(再配置)(步骤 S58 及 S66)。

[0139] 这样,在本变形例中,第 2 再配置部 402 按照由第 1 再配置部 401 新从 LAF 区域 121 将数据再配置到 HAF 区域 112 中的高访问频度物理组块,选择其周边的 Q 个物理组块。

并且,第 2 再配置部 402 将所选择的 Q 个物理组块的数据向 T 区域 111 再配置。即,第 2 再配置部 402 按照高访问频度物理组块,将其周边的 Q 个物理组块的数据向 T 区域 111 再配置。在此状态下,评价部 404 评价该再配置的效果(步骤 S71)。由此,根据本变形例,与从 LAF 区域 121 以 Q 个物理组块为单位顺序地选择物理组块的上述实施方式相比,能够选择更加期待再配置的效果的物理组块的组。

[0140] 根据以上说明的至少 1 个实施方式,对于访问频度较低的组块也给与被分配高速区域的机会,能够评价其分配给系统性能带来的效果。

[0141] 说明了本发明的一些实施方式,但这些实施方式是作为例子提示的,并不是要限定发明的范围。这些新的实施方式能够以其他各种各样的形态实施,在不脱离发明的主旨的范围内能够进行各种省略、替代、变更。这些实施方式及其变形包含在发明的范围或主旨中,并且包含在权利要求书所记载的发明和其等价的范围中。

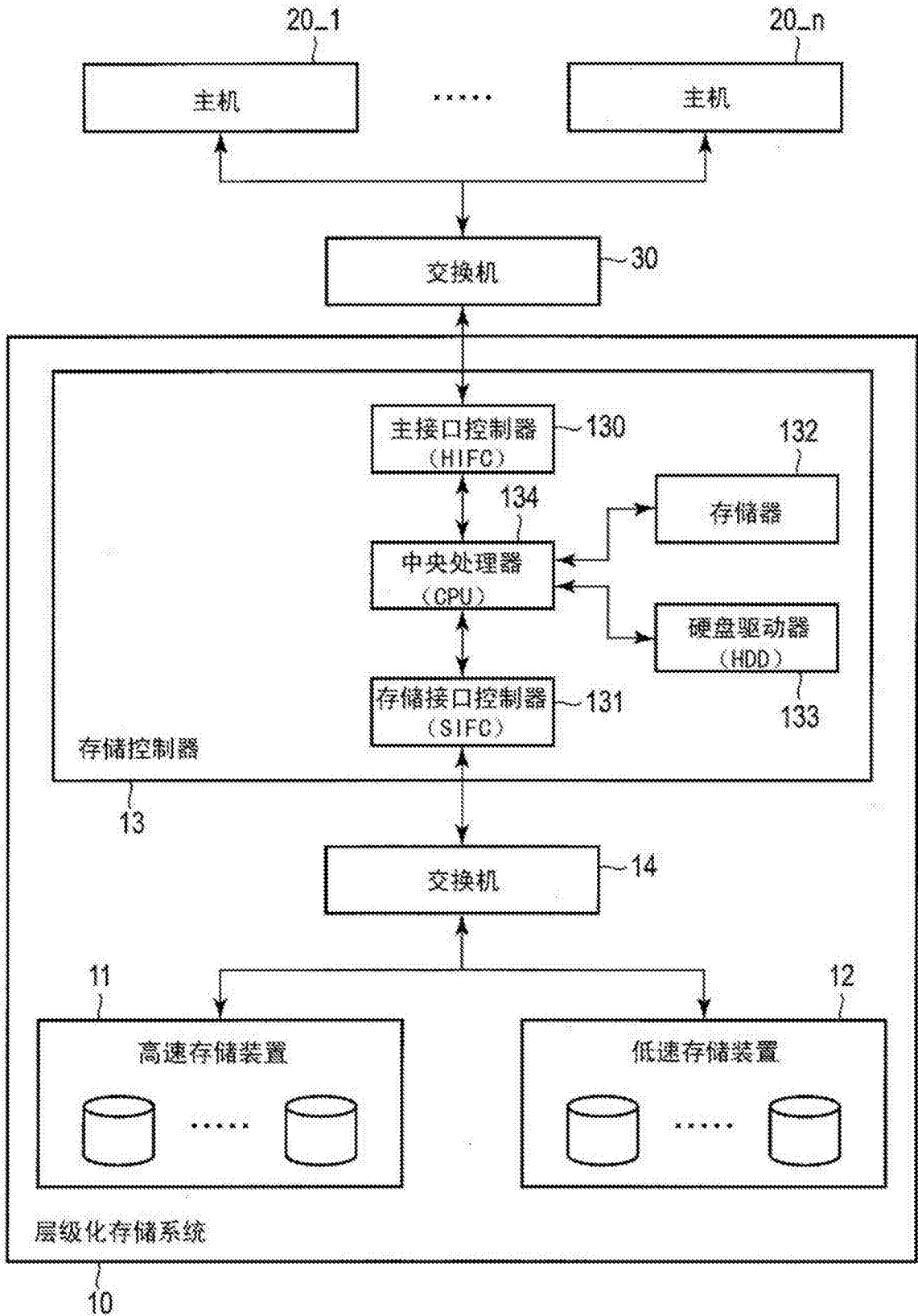


图 1

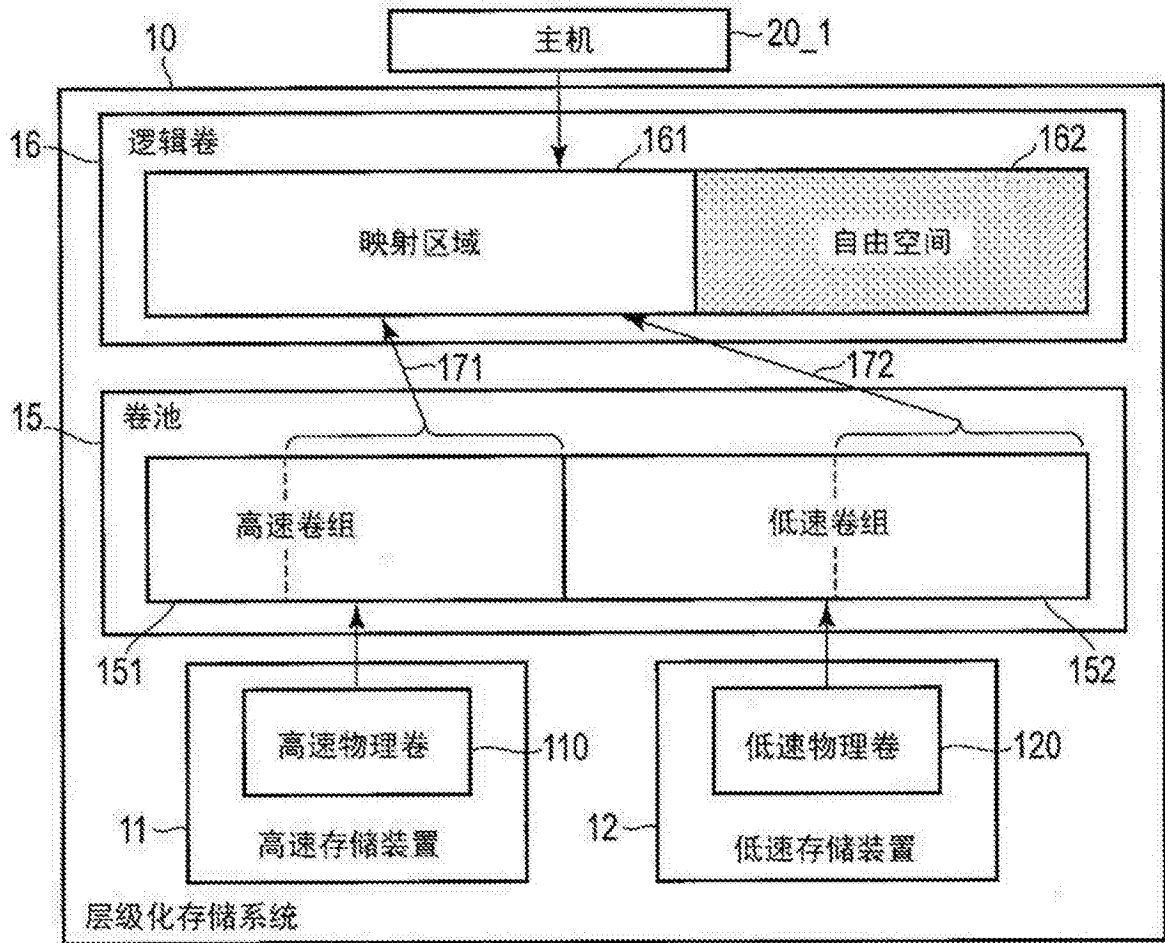


图 2

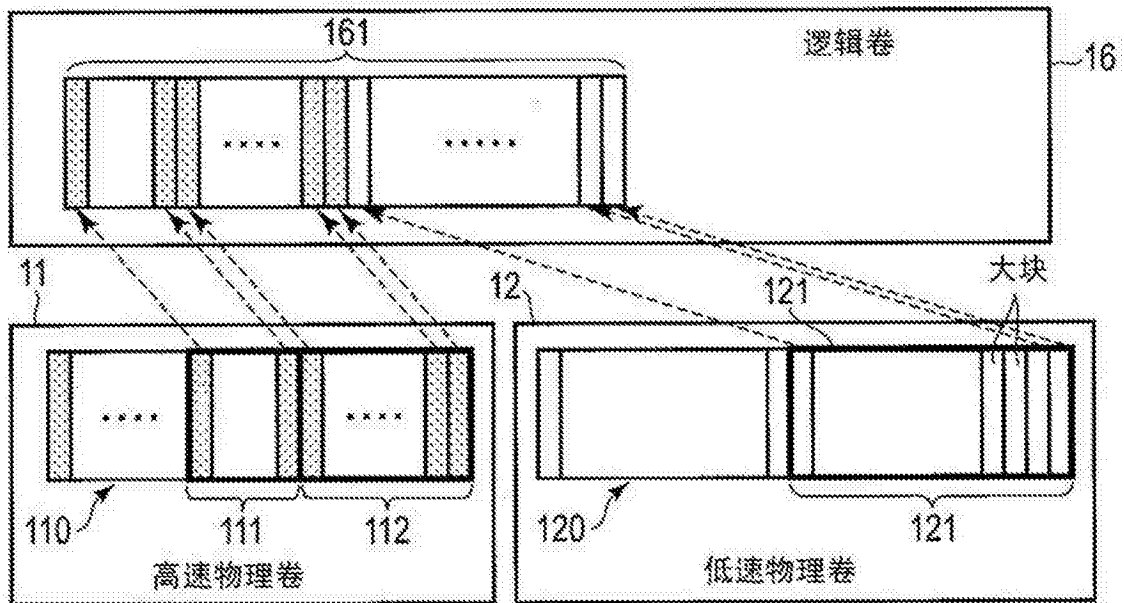


图 3

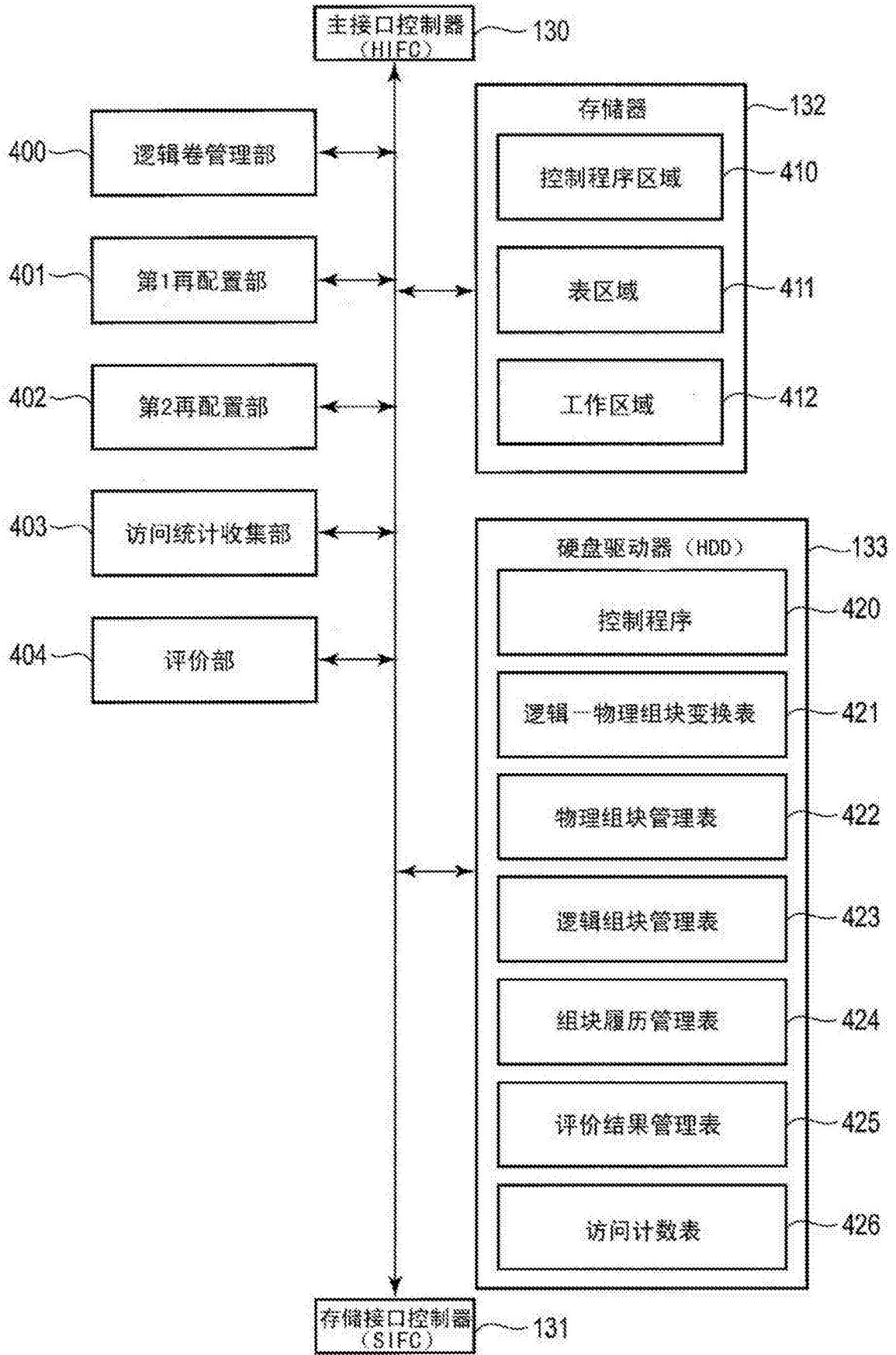


图 4

421

逻辑组块号码	物理组块号码
0	1
1	10
2	12
⋮	⋮

图 5

422

物理组块号码	物理卷 ID	扇区号码	状态
1	1	0	1
2	1	8	1
3	1	16	0
⋮	⋮	⋮	⋮
10	2	0	0
11	2	8	1
12	2	16	0
⋮	⋮	⋮	⋮

图 6

423

逻辑组块号码	逻辑卷ID	扇区号码
1	1	0
2	1	8
3	1	16
⋮	⋮	⋮
10	2	0
11	2	8
12	2	16
⋮	⋮	⋮

图 7

424

No	日時	物理组块一览			
		8000	8001	11999
1	2013/10/31 12	8000	8001	11999
2	2013/10/31 11	4000	4001	7999
3	2013/10/31 10	0	2	3999
⋮	⋮	⋮	⋮	⋮	⋮

图 8

425

No	日時	吞吐量 [MB/s]
1	2013/10/31 12	XXXX
2	2013/10/31 11	XXXX
3	2013/10/31 10	XXXX
⋮	⋮	⋮

图 9

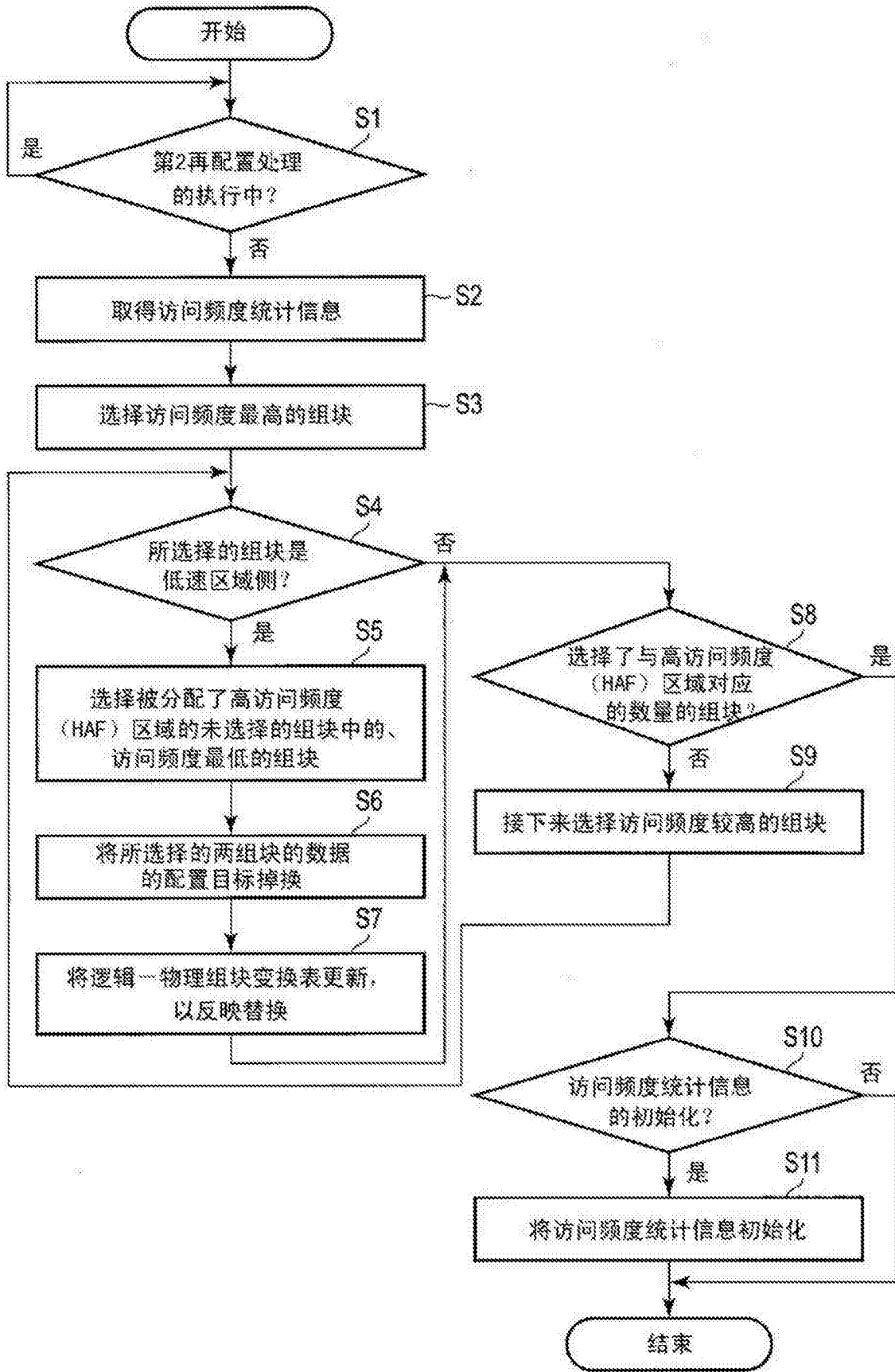


图 10

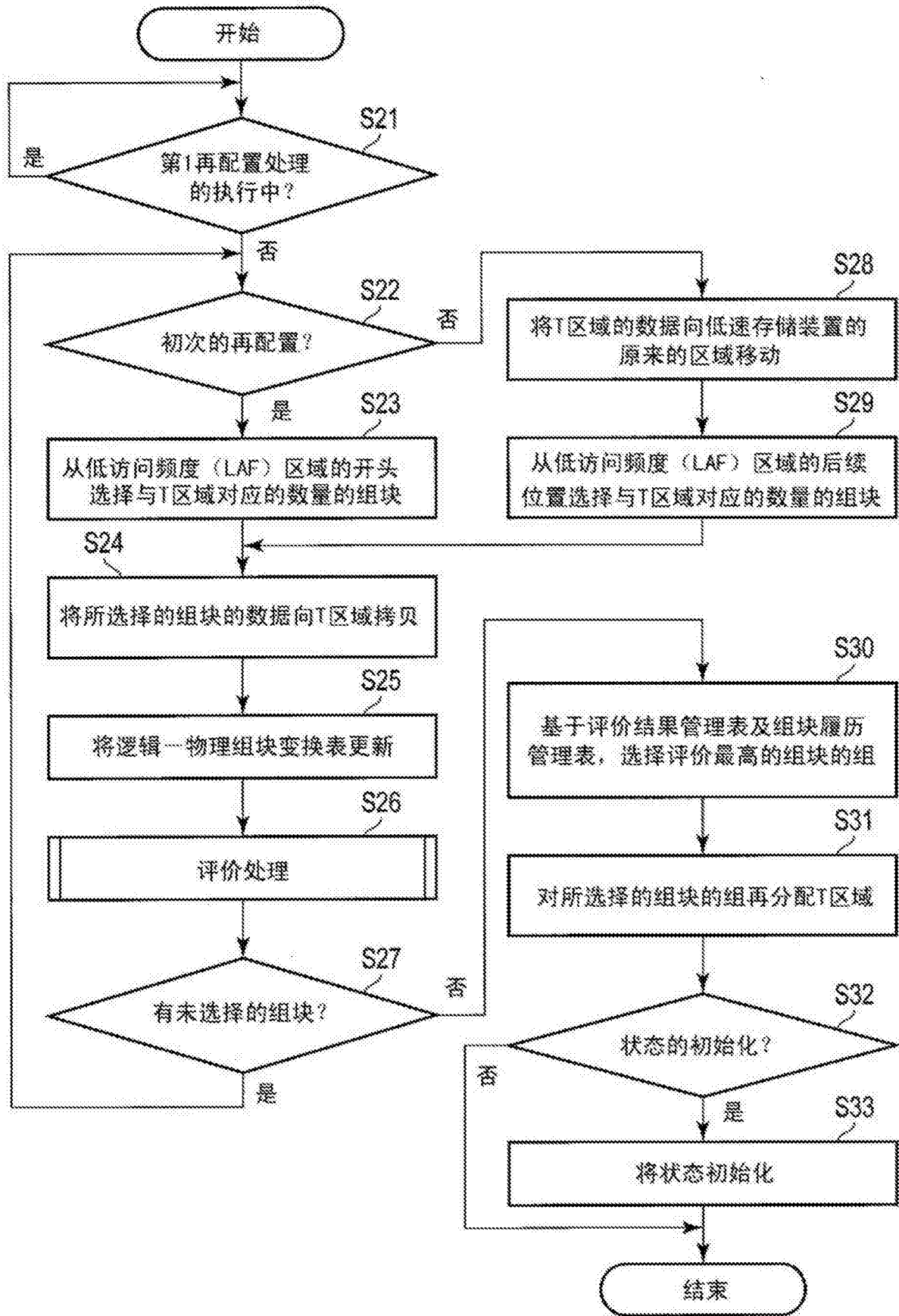


图 11

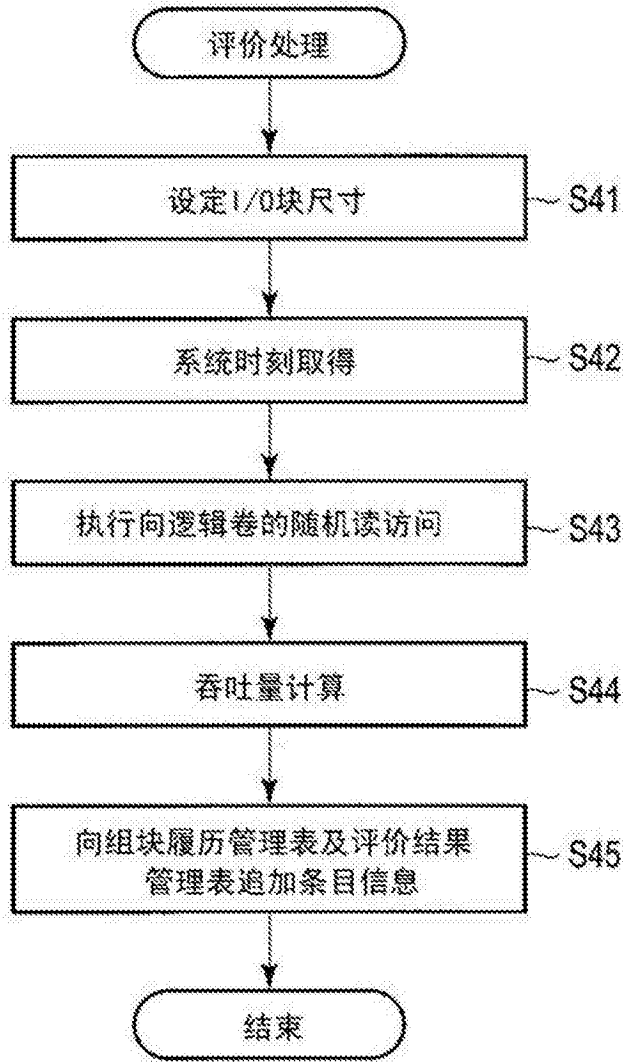


图 12

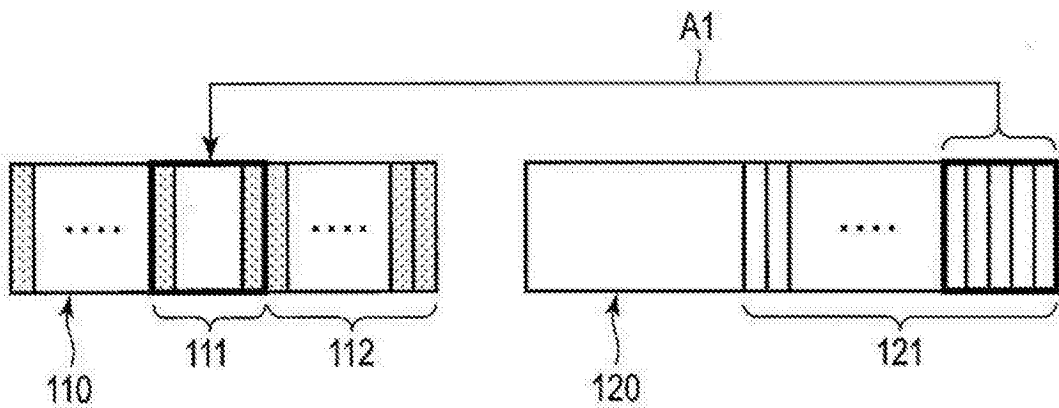


图 13

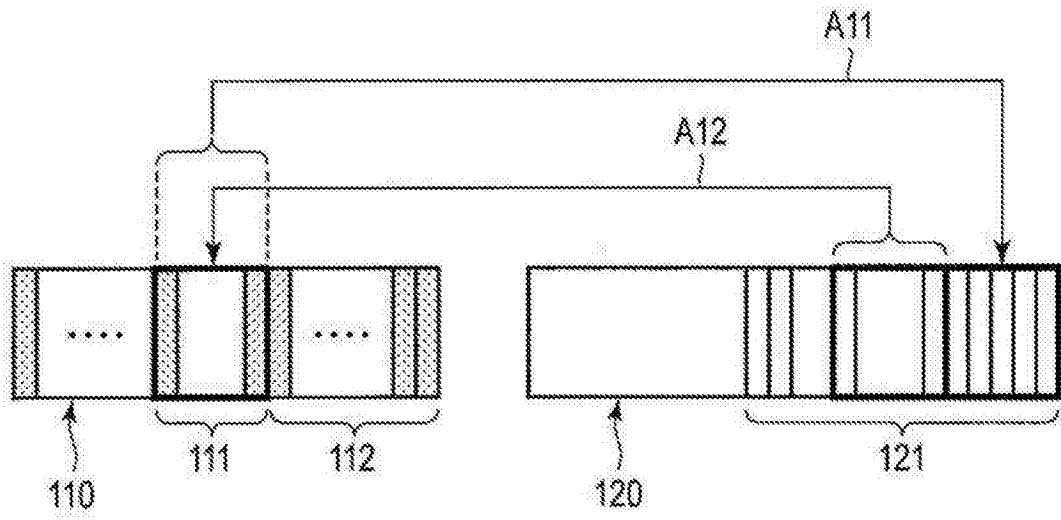


图 14

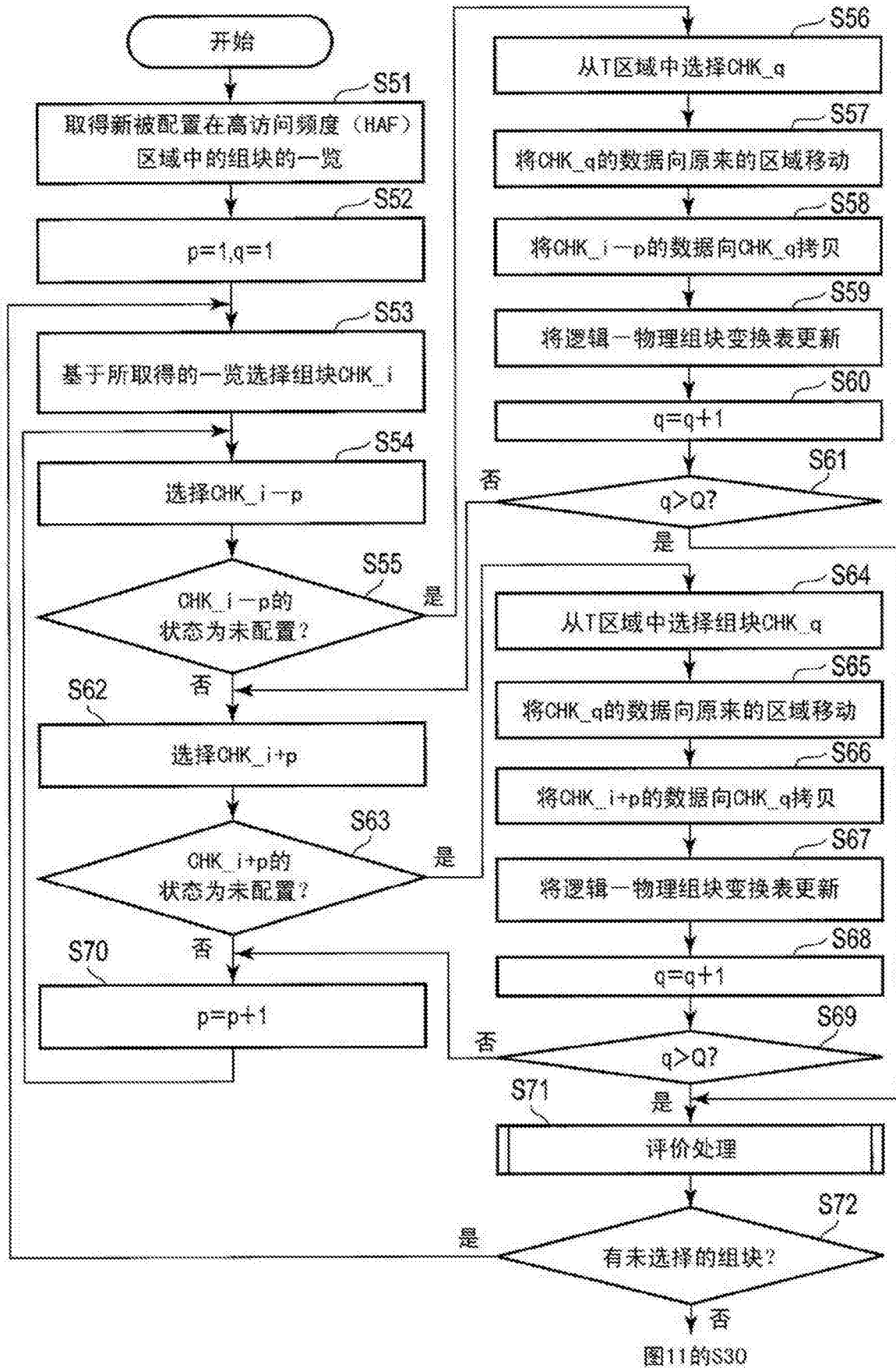


图 15

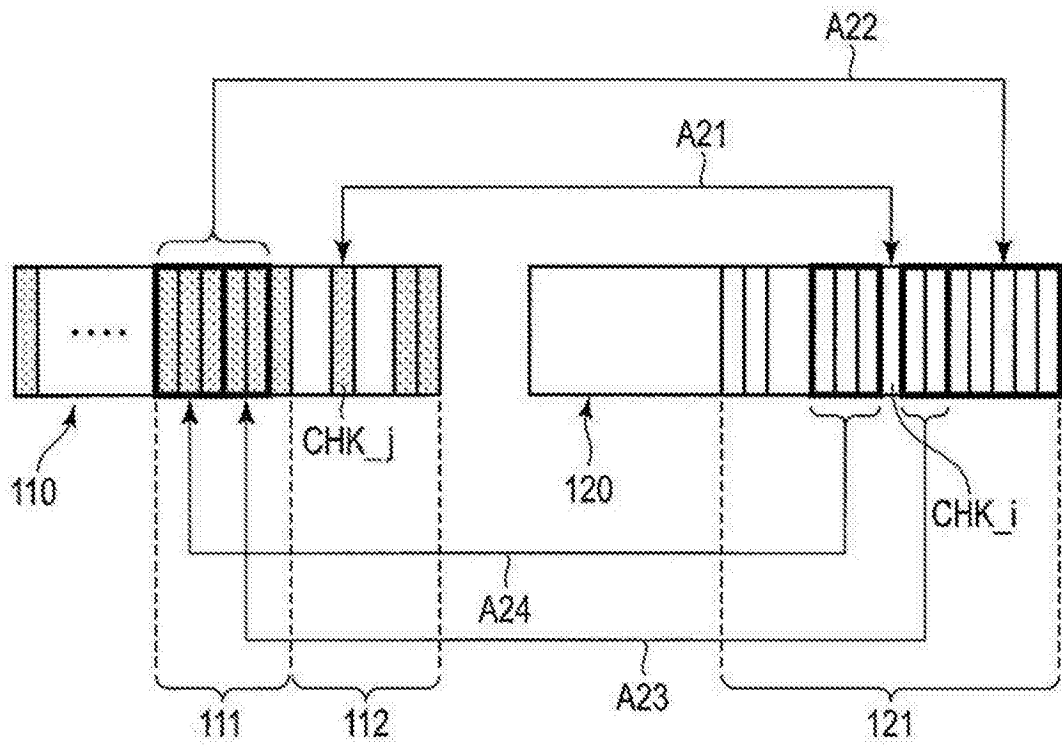


图 16