

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.
G06F 3/06 (2006.01)



[12] 发明专利说明书

专利号 ZL 200610000587.1

[45] 授权公告日 2008 年 5 月 14 日

[11] 授权公告号 CN 100388178C

[22] 申请日 2006.1.11

[21] 申请号 200610000587.1

[30] 优先权

[32] 2005.1.13 [33] US [31] 11/036,628

[73] 专利权人 国际商业机器公司

地址 美国纽约

[72] 发明人 M·D·史密斯

[56] 参考文献

EP1158397A2 2001.11.28

US6145066A 2000.11.7

CN1495612A 2004.5.12

US5333315A 1994.7.26

US2004/0054656A1 2004.3.18

US5276867A 1994.1.4

审查员 高 银

[74] 专利代理机构 北京市中咨律师事务所

代理人 于 静 李 峥

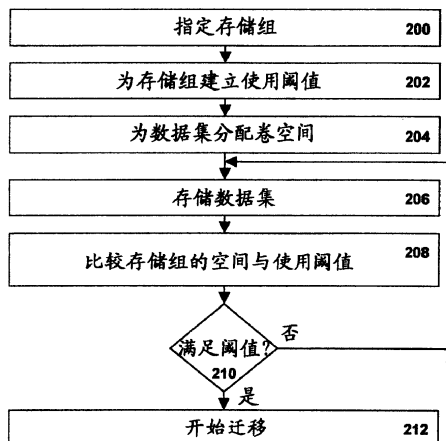
权利要求书 3 页 说明书 6 页 附图 3 页

[54] 发明名称

管理层次存储系统中的数据的方法和系统

[57] 摘要

提供了用于管理层次存储子系统中的数据的方法 and 装置。将多个卷指定为等级 0 的存储组；为该存储组建立高阈值；将用于一数据集的空间分配给该存储组的一卷，该将数据集存储在该卷中；比较该高阈值与该存储组各卷中的所有数据集消耗的空间总量；且如果该高阈值小于或等于该总量，则将数据集从该存储组迁移到等级 1 存储设备。可选地，向每个存储组分配高阈值，并且当在一存储组中使用的空间达到或超过该高阈值时，将从该存储组中具有最小自由空间的卷开始数据迁移。又可选地，当选择一卷以进行迁移时，设置标志以防止该卷中的空间被分配给新的数据集。在完成迁移时，清除标志以允许分配。这样，避免了迁移和空间分配之间的争用。



1. 一种用于管理层次存储子系统中的数据的方法，包括：

将多个卷指定为用于等级 0 存储设备的存储组；

为该存储组建立一高阈值；

比较该高阈值与由存储到该存储组中的各卷中的所有数据集所使用的空间的总量；

如果已达到该高阈值，则识别所述存储组中的所有卷中具有最小自由空间量的卷，并从存储在该识别的卷中的数据集开始，将数据集从该存储组迁移到等级 1 存储设备；以及

将用于新数据集的空间分配给所述存储组的所有卷中具有最大自由空间的卷，并将该新数据集存储在该卷中，借此降低了所述用于新数据集的空间分配将是分配到数据集当前正被从其迁移的卷的空间的风险。

2. 根据权利要求 1 的方法，还包括：

为该存储组中的每个卷建立一低阈值；以及

所述迁移步骤将数据集从所述识别的卷迁移，直到由存储到该识别的卷中的所有数据集使用的空间的总量小于或等于用于该识别的卷的低阈值。

3. 根据权利要求 1 的方法，还包括：

为所述存储组建立一低阈值；以及

所述识别和迁移步骤被重复执行，直到由存储在该存储组中的各卷中的所有数据集使用的空间的总量小于或等于用于该存储组的所述低阈值，或直到数据集已从该存储组中的所有卷迁移，以两者中先发生者为准。

4. 根据权利要求 1 的方法，还包括：

当数据集从所述识别的卷的迁移开始时设置标志；

如果设置了标志，则防止用于新数据集的空间被分配给该识

别的卷；以及

当数据集从该识别的卷的迁移完成时，清除该标志，借此允许将用于新数据集的空间分配给该识别的卷。

5. 一种层次存储管理系统，包括：

主计算机，其附接有至少第一和第二等级的数据存储设备，该第一等级的数据存储设备被配置为至少第一存储组，该第一存储组具有被分配用于存储数据集的多个卷的空间；

用于为该第一存储组建立一高阈值的装置；

用于比较该高阈值与由存储到该第一存储组中的各卷中的所有数据集消耗的空间的总量的装置；

用于如果该高阈值小于或等于由存储到该第一存储组中的各卷中的所有数据集所使用的空间的总量，则识别所述第一存储组中的所有卷中具有最小自由空间量的卷，并从存储在该识别的卷中的数据集开始将数据集从该第一存储组迁移到该第二等级存储设备的装置；以及

用于将用于新数据集的空间分配给所述第一存储组的所有卷中具有最大自由空间的卷，并将该新数据集存储在该卷中，借此降低了所述用于新数据集的空间分配将是分配到数据集当前正被从其迁移的卷的空间的风险的装置。

6. 根据权利要求 5 的系统，还包括：

用于为所述第一存储组中的每个卷建立一低阈值的装置；以及

所述迁移装置用于将数据集从所述识别的卷迁移，直到由存储在该识别的卷中的所有数据集使用的空间的总量小于或等于用于该识别的卷的所述低阈值。

7. 根据权利要求 5 的系统，还包括：

用于为所述第一存储组建立一低阈值的装置；以及

所述识别和迁移装置用于重复所述识别和迁移，直到由存储

到该第一存储组中的各卷中的所有数据集所使用的空间的总量小于或等于用于该第一存储组的所述低阈值，或直到数据集已被从该第一存储组中的所有卷迁移，以两者中先发生者为准。

8. 根据权利要求 5 的系统，还包括：

用于当数据集从所述识别的卷的迁移开始时设置标志的装置；

用于如果设置了该标志则防止用于新数据集的空间被分配该识别的卷的装置；以及

用于当数据集从该识别的卷的迁移完成时清除该标志，借此允许将用于新数据集的空间分配给该识别的卷的装置。

9. 根据权利要求 7 的系统，其中所述第一等级的存储设备包括直接存取存储设备。

10. 根据权利要求 7 的系统，其中所述第二等级的存储设备包括磁带存储设备。

管理层次存储系统中的数据的方法和系统

技术领域

本发明涉及在层次存储子系统中的存储等级之间高效的数据迁移。

背景技术

数据处理设施 (facility) 通常包括一个或多个主机设备, 存储设备或直接地或通过网络附接在所述主机设备上。随着用户的业务增长, 对增加的存储设备的需要也增长了。然而, 安装附加的高速存储设备例如 DASD (直接存取存储设备) 可能是昂贵的。此外, 并非用户的所有存储设备会都需要是昂贵、高速类型的。而是可将这样的存储设备分配给经常被访问的数据, 而可将较不经常访问的数据存储在较慢和较低廉的存储设备例如磁带盒中。称具有多个等级的存储设备被配置为层次存储设备。

主机的操作系统可包括一层次存储管理器。在 IBM®S/390®服务器上运行的 OS/390®或 z/OS®操作系统可包括一数据设施存储管理子系统 (DFSMS™), 该子系统自动化了存储管理策略和过程以将数据在存储层次结构中上下移动。应理解其他厂商可提供也管理层次存储空间的其他产品。而且, 尽管将在来自 IBM 的 DFSMS 的情境中描述本发明, 本发明并不限于被并入该特定产品中。

DFSMS 的一个组件是层次存储管理器 (DFSMSHsm™, 在这里也被称为 “HSM”), 该管理器监督备份、恢复、迁移和空间管理操作。由 DFSMSHsm 的空间管理功能提供的一种功能是 “间隔迁移”, 该功能指导将数据集定期迁移到其存储较为低廉或其存储形式节省空间的位置。HSM 间隔迁移基于用户指定的结构 (construct) 自动地管理 DASD 存储空间 (也被称为等级 0) 中的卷。可将存储有具有共同的由用户定义的存储要求的

数据集的各卷组合为“存储组”。用户可为一卷选择的参数之一是“高阈值”。以预定的间隔，HSM 按字母顺序检查各 DASD 卷，并确定由存储到一卷中的数据集所使用的空间是否超过该高阈值。如果是这样，则该卷中的数据集就适合于从等级 0 迁移到等级 1（例如 DASD/磁带介质）。然后在所有这样的卷上执行迁移，这种迁移是从每个卷中最大的数据集开始，并以越来越小的数据集继续进行，直到达到另一个用户定义参数，即该卷的“低阈值”。

DFSMS 的另一个功能是为新数据集分配空间。通常，用于新数据集的空间将被分配给具有最大自由空间量的卷。然而，经常地，空间被分配到的卷是数据集正在从其迁移的同一个卷。这种冲突造成对共同资源包括卷目录表（VTOC）、目录（catalog）条目等的争用，从而导致数据迁移和/或用户工作负荷的速度减慢。

因此，存在着减少数据迁移和空间分配之间的争用以及增加迁移速度的需要。

发明内容

本发明提供了用于管理层次存储子系统中的数据的方法和装置。该方法包括将多个卷指定为用于等级 0 存储设备的存储组，为该存储组建立高阈值，将用于一数据集的空间分配给该存储组的一卷，该将数据集存储在该卷中，比较该高阈值与由存储在该存储组中的各卷中的所有数据集所消耗的空间的总量，以及如果该高阈值小于或等于由存储在该存储组中的各卷的所有数据集所使用的空间的总量，则将数据集从该存储组迁移到等级 1 存储设备。

在一个实施例中，向每个存储组分配高阈值，并且当由一存储组使用的空间达到或超过该高阈值时，将开始从该存储组中的卷的数据迁移，并且是从具有最小自由空间的卷开始。这样，就减少了迁移和空间分配之间的争用。

在另一个实施例中，当选择一卷以进行迁移时，设置一标志，该标志

防止该卷中的空间被分配给新的数据集。在完成了该迁移之后，清除该标志，从而允许分配。这样，就避免了迁移和空间分配之间的争用。

根据本发明的一个方面，提供了一种用于管理层次存储子系统中的数据的方法，该方法包括以下步骤：将多个卷指定为用于等级 0 存储设备的存储组；为该存储组建立高阈值；比较该高阈值与由存储到该存储组中的各卷中的所有数据集所使用的空间的总量；如果已达到该高阈值，则识别所述存储组中的所有卷中具有最小自由空间量的卷，并从存储在该识别的卷中的数据集开始，将数据集从该存储组迁移到等级 1 存储设备；以及将用于新数据集的空间分配给所述存储组的所有卷中具有最大自由空间的卷，并将该新数据集存储在该卷中，借此降低了所述用于新数据集的空间分配将是分配到数据集当前正被从其迁移的卷的空间的风险。

根据本发明的另一个方面，提供了一种层次存储管理系统，该层次存储管理系统包括：主计算机，其附接有至少第一和第二等级的数据存储设备，该第一等级的数据存储设备被配置为至少第一存储组，该第一存储组具有被分配用于存储数据集的多个卷的空间；用于为该第一存储组建立高阈值的装置；用于比较该高阈值与由存储到该第一存储组中的各卷中的所有数据集消耗的空间的总量的装置；用于如果该高阈值小于或等于由存储到该第一存储组中的各卷中的所有数据集所使用的空间的总量，则识别所述第一存储组中的所有卷中具有最小自由空间量的卷，并从存储在该识别的卷中的数据集开始将数据集从该第一存储组迁移到该第二等级存储设备的装置；以及用于将用于新数据集的空间分配给所述第一存储组的所有卷中具有最大自由空间的卷，并将该新数据集存储在该卷中，借此降低了所述用于新数据集的空间分配将是分配到数据集当前正被从其迁移的卷的空间的风险的装置。

附图说明

图 1 是其中可包含本发明的层次存储系统的框图；

图 2 是本发明的方法的流程图；

图 3 是本发明的一实施例的流程图；

图 4 是本发明的另一实施例的流程图；以及

图 4 是本发明的又一实施例的流程图。

具体实施方式

图 1 是其中可包含本发明的层次存储系统 100 的框图。系统 100 包括主机设备 110，多个等级的数据存储设备 120 和一管理接口 130 附接在该主机设备上。主机 110 包括处理器 112、存储器 114 和操作系统 116。操作系统 116 包括层次存储管理功能 118。为了本说明的目的，该层次存储管理器 118 被示为操作地连接到存储设备 120。然而，操作系统 116 和/或主机 110 中的其他过程也可访问存储设备 120。因此，如这里所使用的术语“连接”可指两个组件、设备或子系统之间的间接关系，以及两者之间的直接电连接。

存储设备 120 包括高速（和昂贵）的等级 0 存储设备 122、以及较慢（和较低廉）的等级 1 存储设备 124。在图 1 中被表示为更慢（和更低廉）的等级 2 存储设备 126 的其他等级的存储设备也可连接到层次存储管理器 118。如将在以下描述的，各卷 132 被组合为存储组 130 以便存储数据集 134。

作为示例而非限制，所述主机 110 可以是 IBM S/390，操作系统 116 可以是 z/OS，并且层次存储管理功能可由 DFSMSHsm 提供。等级 0 存储设备可以是 DASD 设备，例如 IBM TotalStorageTM Model 2107。等级 1 存储设备可以是 DASD/磁带介质，例如 IBM TotalStorage Model 2105 企

业存储服务器或 IBM TotalStorage 3592 企业磁带系统。

图 2 是本发明的方法的流程图。通过使用 DFSMShsm 或其等价物, 设施管理员将一个或多个卷指定为等级 0 存储组 (步骤 200); 也可类似地指定其他的存储组。对于每个存储组, 由设施管理员建立一高阈值, 并将其存储在存储器 114 中 (步骤 202), 该存储器可包括数据库、表或类似物。然后可将用于新数据集的空间分配给该存储组的卷 (步骤 204), 并且将数据集存储在该卷中 (步骤 206)。在由管理员规定的时间, DFSMShsm 比较任何存储组的空间与高阈值 (步骤 208)。如果已满足该阈值 (步骤 210), 就是说, 该高阈值小于或等于该存储组中的已使用空间量, 则将数据集从该卷迁移到等级 1 存储设备 (步骤 212)。

参照图 3 的流程图, 在一个实施例中, 检查该存储组的各卷 (步骤 300), 并且识别具有最小自由空间的卷 (步骤 302)。然后从存储在该识别出的卷中的数据集开始数据集从该存储组的迁移 (步骤 304)。

除了为每个存储组建立的高阈值外, 可为每个存储组中的每个卷建立一低阈值 (步骤 306)。当建立了卷低阈值时, 数据集从一卷的迁移将继续下去, 直到达到该低阈值 (步骤 308)。

也可指导 DFSMS 将用于新数据集的空间分配给存储组的所有卷中具有最大自由空间的卷 (步骤 310)。相反地, 具有最小自由空间的卷更可能是迁移的对象。因此, 降低了空间将被分配到的卷是数据集正在从其被迁移的同一卷的风险。

在另一个实施例 (图 4) 中, 为每个存储组建立一低阈值 (步骤 400), 以指示可允许的自由空间量 (或最小的使用空间量)。当迁移开始时, 识别存储组的所有卷中具有最小自由空间量的卷 (步骤 402), 并且迁移从来自所识别的卷的数据集开始 (步骤 404)。接着以随后识别的具有下一个最小自由空间量的卷的数据集进行迁移 (步骤 406)。该过程继续进行, 直到已满足所述低阈值 (步骤 408), 或直到已在该存储组中的所有卷上执行了迁移 (步骤 410), 以两者中先发生者为准。

在又一个实施例中 (图 5), 当一卷的数据集的迁移开始时, 设置一

标志（步骤 500），该标志防止用于新数据集的空间被分配给该卷（步骤 502），这样就防止了迁移和分配功能之间的争用。当从该卷的迁移完成（步骤 504）时，清除该标志（步骤 506），从而允许分配（步骤 508）。

重要的是注意到，尽管已在全功能的数据处理系统的情境中描述了本发明，本领域的技术人员将认识到本发明的过程能够以包含指令的计算机可读介质的形式和多种形式分发，并且不管实际用来进行该分发的信号承载介质的特定类型本发明都适用。计算机可读介质的示例包括可记录类型的介质例如软盘、硬盘驱动器、RAM、和 CD-ROM，以及传输类型的介质例如数字和模拟通信链路。

对本发明的描述是为了说明和描述的目标呈现的，而非旨在是穷尽性的或限于本发明的所公开的形式。对于本领域的普通技术人员来说，很多修改和改变将是显然的。所选择和描述的实施例是为了最好地解释本发明的原理和实际应用，并使本领域的其他普通技术人员能够理解本发明，以实现具有适合于所考虑的特定应用的各种修改的实施例。此外，尽管以上针对方法和系统进行了描述，本领域中的上述需要也可以包含用于管理层次存储设备的指令的计算机程序产品来满足，或以这样一种用于部署计算基础设施的方法来满足，该方法包括将计算机可读代码集成到计算系统中以便管理层次存储设备。

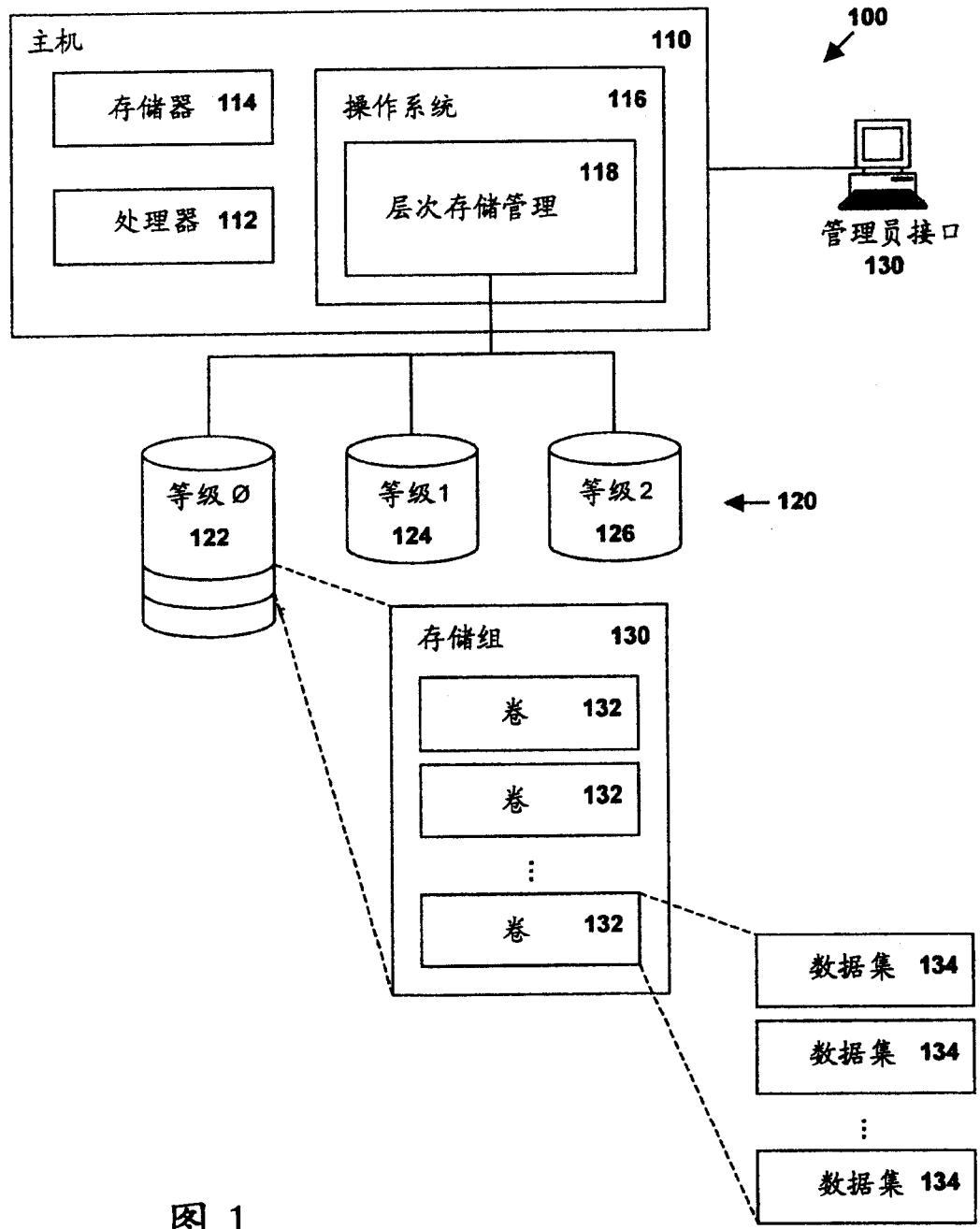


图 1

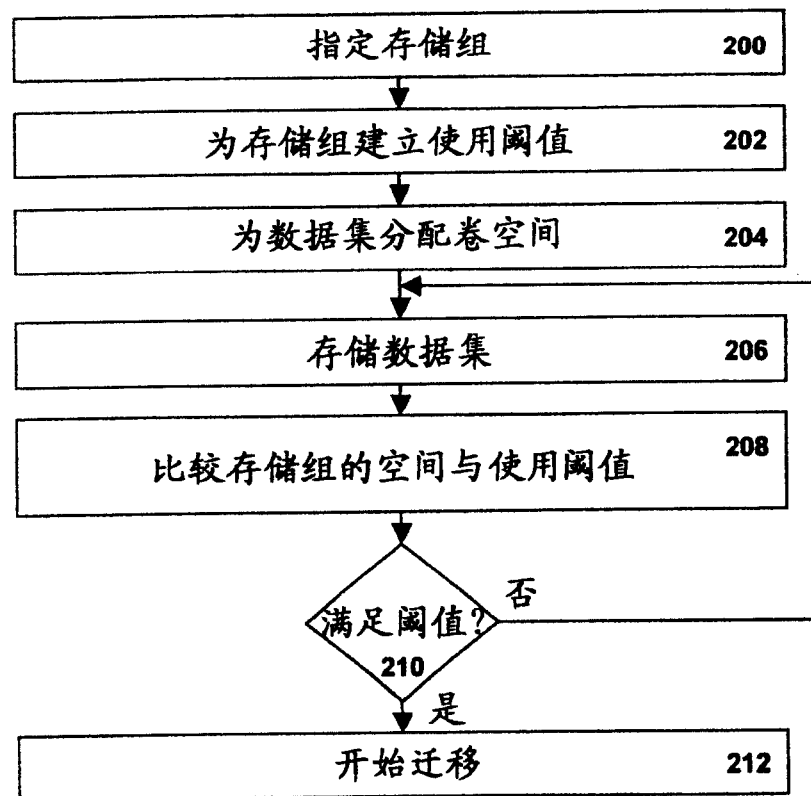


图 2

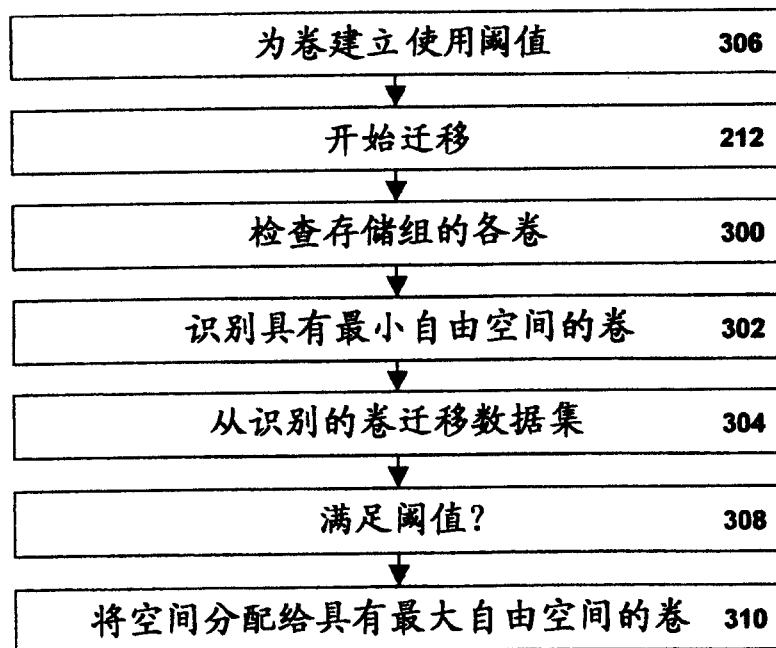


图 3

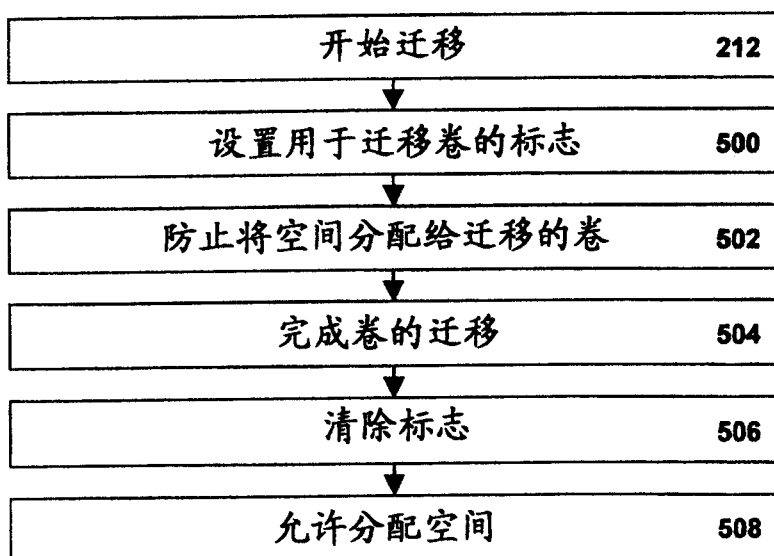
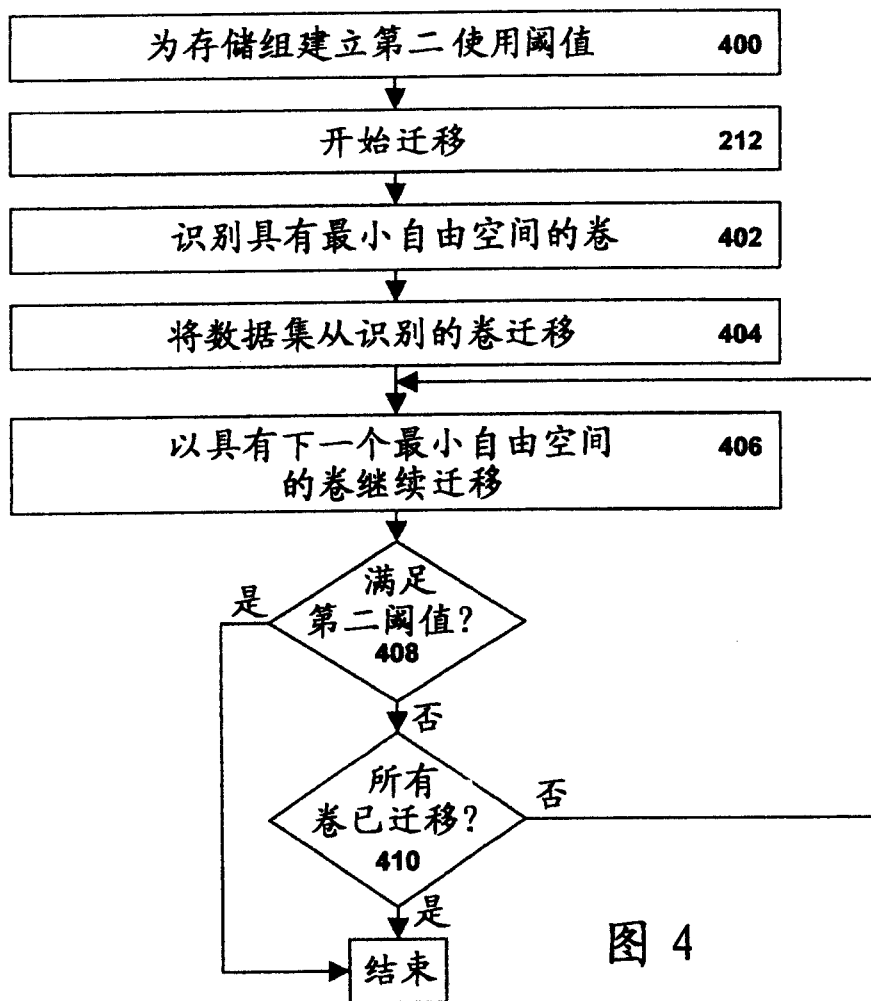


图 5