



[12] 发明专利说明书

专利号 ZL 03800377.5

[45] 授权公告日 2007 年 12 月 12 日

[11] 授权公告号 CN 100354813C

[22] 申请日 2003.4.10 [21] 申请号 03800377.5

[30] 优先权

[32] 2002. 4. 10 [33] US [31] 10/120,016

[86] 国际申请 PCT/US2003/011272 2003. 4. 10

[87] 国际公布 WO2003/088045 英 2003. 10. 23

[85] 进入国家阶段日期 2003. 12. 4

[73] 专利权人 伊姆西公司

地址 美国马萨诸塞州

[72] 发明人 纳坦·维什利茨科

哈伊姆·科佩洛维茨

哈纳·莫雷塞特 阿迪·奥弗

[56] 参考文献

US6101497A 2000. 8. 8

US5241670A 1993. 8. 31

审查员 吴 敏

[74] 专利代理机构 北京金信立方知识产权代理有

限公司

代理人 张金海

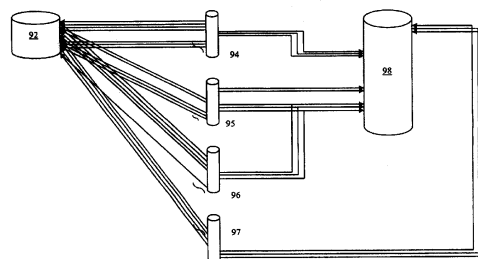
权利要求书 2 页 说明书 21 页 附图 14 页

[54] 发明名称

虚拟存储装置

[57] 摘要

访问所存储的数据包括提供一具有一指针表的虚拟存储区，所述指针指向至少两个其他存储区的段，其中，虚拟存储区不包含数据段；响应于访问虚拟存储区的数据的请求，确定其他存储区的哪一特定区包含数据；并使用指针表访问其他存储区的特定区上的数据。访问所存储的数据还可包括将其他存储区的第一区与虚拟存储区关联，其中虚拟存储区代表其他存储区的第一区的数据的备份。访问所存储的数据还可包括，当虚拟存储区被初始与其他存储区的第一区关联时，使表的所有指针初始指向其他存储区的第一区的段。存储区可以是存储装置。段可以是磁道。



1、一种提供已存储数据的及时点备份的方法，包括
提供具有指针表的虚拟存储区，所述指针指向已存储数据的段及至少一其他存储区的段，其中，虚拟存储区不包含数据段；

响应于将新数据写到虚拟存储区的指针表的指针指向的已存储数据的段，在写之前将已存储数据的段复制到至少一其它存储区的段，使得虚拟存储区的指针指向至少一其它存储区的段，及将新数据写到已存储数据的段；

响应于访问虚拟存储区的数据的请求，确定其他存储区的哪一特定区包含数据；及

使用指针表访问其他存储区的特定区上的数据。

2、如权利要求1的方法，进一步包括：

将已存储的数据与虚拟存储区关联，其中虚拟存储区代表已存储数据的备份。

3、如权利要求2的方法，进一步包括：

当虚拟存储区被初始与已存储的数据关联时，使表的所有指针初始指向已存储的数据。

4、如权利要求1的方法，其中：在拷贝数据到至少一其它存储区的段之前，至少一其它存储区的段被保持为不包含数据的空闲段。

5、如权利要求3的方法，进一步包括：

将数据指示器与已存储数据的段进行关联，其中，特定段的数据指示器指出：在已存储的数据已与虚拟存储区关联后，是否已经对特定段执行写操作。

6、如权利要求1的方法，其中每一数据段是数据磁道。

7、如权利要求1的方法，其中每一存储区是存储装置。

8、一种对数据存储区中存储的数据进行及时点备份的方法，包括：

提供具有第一指针表的第一虚拟存储区，所述指针指向数据存储区的段及至少一其他存储区的段，其中，第一虚拟存储区不包含数据

段；

响应于将新数据写到第一虚拟存储区的第一指针表的指针指向的数据存储区的段，在写之前将数据段复制到至少一其它存储区的段，使得第一虚拟存储区的指针指向至少一其它存储区的段，及将新数据写到数据存储区的段；

将数据存储区与第一虚拟存储区关联，其中第一虚拟存储区代表数据存储区的数据在第一时间点的备份；

提供具有第二指针表的第二虚拟存储区，所述指针指向数据存储区的段及至少一其他存储区的段，其中，第二虚拟存储区不包含数据段；

响应于将新数据写到第二虚拟存储区的第二指针表的指针指向的数据存储区的段，在写之前将数据段复制到至少一其它存储区的段，使得第二虚拟存储区的指针指向至少一其它存储区的段；

将数据存储区与第二虚拟存储区关联，其中第二虚拟存储区代表数据存储区的数据在第二时间点的备份；

响应于访问虚拟存储区之一的数据的请求，确定存储区的哪一特定区包含数据；及

使用指针表之一访问特定存储区上的数据。

9、如权利要求8的方法，进一步包括：

当第一虚拟存储区被初始与数据存储区关联时，使第一表的所有指针初始指向数据存储区的段；

当第二虚拟存储区被初始与数据存储区关联时，使第二表的所有指针初始指向数据存储区的段。

虚拟存储装置

技术背景

1. 技术领域

本申请涉及计算机存储装置，特别是涉及提供存储在计算机存储装置上的数据的一部分的备份的领域。

2. 相关现有技术

主处理机系统使用存储装置可存储和检索数据，存储装置包含：多个主机接口单元（主机适配器）、磁盘驱动器、以及磁盘接口单元（磁盘适配器）。这样的存储装置由，例如，麻萨诸塞州霍普金顿的EMC 有限公司提供，并在授权给 Yanai 等人的美国专利 5,206,939、授权给 Galtzur 等人的美国专利 5,778,394、授权给 Vishlitzky 等人的美国专利 5,845,147 以及授权给 Ofek 的美国专利 5,857,208 中公开。主机系统通过随存储装置提供的多个通道访问存储装置。主机系统通过存储装置的通道提供数据和访问控制信息，存储装置也通过这些通道提供数据给主机系统。主机系统不直接寻址存储装置的磁盘驱动器，而是更愿访问呈现给主机系统的多个逻辑卷。逻辑卷可以、也可不对应于实际的磁盘驱动器。

在一些实例中，想要提供一逻辑卷的备份，该备份接着被其他的进程访问。例如，在实际所存储的数据上测试新软件，可能会做一包含数据的逻辑卷的备份，与原始数据相对，该备份用于测试新软件。一旦测试完成，备份可被消除。从而新软件在不影响实际数据的情形下而在实际数据上得以测试。这就减少了因测试新软件和/或新功能而破坏实际数据的可能性。

进行这样的备份的一个困难在于：该备份要求与从其获得数据的逻辑卷一样多的存储空间，因为备份处理简单地创立了一包含原始卷的所有数据的新卷。此外，在一些实例中，原始卷与备份卷之间的差是最小的。因而，逻辑卷的这样的备份所需要的额外的存储空间使用

得有点没有效率，因为其仅复制已经存在的数据。因此，需要提供一有效使用存储空间的拷贝数据的设备。

发明概述

根据本发明，访问所存储的数据包括：提供一具有指针表的虚拟存储区，所述指针指向至少两个其他存储区的段，其中虚拟存储区不包含数据段，响应于访问虚拟存储区的数据的请求，确定其他存储区的哪一特定区包含数据，并访问使用指针表的其他存储区的特定区上的数据。访问所存储的数据还可包括：将其他存储区的第一区与虚拟存储区关联，虚拟存储区代表其他存储区的第一区的数据的备份。访问所存储的数据还可包括：当虚拟存储区与其他存储区的第一区初始关联时，使得表的所有指针初始指向其他存储区的第一区的段。访问所存储的数据还可包括：响应于对其他存储区的第一区上的第一段的写，拷贝第一段的数据到其他存储区的第二区上的第二段并使得表指针的对应指针指向第二段。在从第一段拷贝数据到第二段之前，第二段应保持为不包含数据的空段。访问所存储的数据还可包括：保持一其他存储区的第二区的所有空段的双链表。访问所存储的数据还可包括：将数据指示器与其他存储区的第一区的段关联，用于特定段的数据指示器指示在其他存储区的第一区已经与虚拟存储区关联后，对特定段是否已执行写操作。访问所存储的数据还可包括：响应于对其他存储区的第一区上的第一段的写，以及数据指示器指示在其他存储区的第一区已经与虚拟存储区关联后还没有对第一段执行写操作，拷贝第一段的数据到其他存储区的第二区上的第二段，并使得表指针的对应指针指向第二段。访问所存储的数据还可包括：发送状态信息到一装置，该信息使得在从第一段拷贝数据到第二段后执行写操作。每一数据段可以是一数据磁道。每一存储区可以是存储装置。

进一步地根据本发明，访问所存储的数据包括：提供具有第一指针表的第一虚拟存储区，所述指针指向至少两个其他存储区的段，第一虚拟存储区不包含数据段，其他存储区的第一区与第一虚拟存储区

关联, 在这里第一虚拟存储区代表在时间的第一点上其他存储区的第一区的数据的备份; 提供具有第二指针表的第二虚拟存储区, 所述指针指向至少两个其他存储区的段, 第二虚拟存储区不包含数据段, 其他存储区的第一区与第二虚拟存储区关联, 这里第一虚拟存储区代表在时间的第二点上其他存储区的第一区的数据的备份, 响应于访问虚拟存储区之一的数据的请求, 确定其他存储区的哪一特定区包含数据, 使用指针表之一访问其他存储区的特定区上的数据。访问所存储的数据还可包括: 在第一虚拟存储区初始与其他存储区的第一区关联时, 使得第一表的所有指针初始指向其他存储区的第一区的段, 且在第二虚拟存储区初始与其他存储区的第一区关联时, 使得第二表的所有指针初始指向其他存储区的第一区的段。访问所存储的数据还可进一步包括: 将第一数据指示器与其他存储区的第一区的段关联, 在这里, 用于第一其他存储区的特定段的第一数据指示器指示在其他存储区的第一区已经与第一虚拟存储区关联后, 是否已向特定段执行写操作; 将第二数据指示器与其他存储区的第一区的段关联, 在这里, 用于第一其他存储区的特定段的第二数据指示器指示: 在其他存储区的第二区已经与第二虚拟存储区关联后, 是否已向特定段执行写操作。访问所存储的数据也可包括: 响应于向其他存储区的第一区上的第一段进行写, 以及数据指示器指示在其他存储区的第一区已与第一虚拟存储区关联后还没有对第一段执行写操作, 拷贝第一段的数据到其他存储区的第二区上的第二段, 并使第一表的指针中的对应指针指向第二段。访问所存储的数据也可包括: 响应于向其他存储区的第一区上的第一段进行写, 以及数据指示器指示在其他存储区的第一区已与第一及第二虚拟存储区关联后还没有对第一段执行写操作, 拷贝第一段的数据到其他存储区的第二区上的第二段, 并使第一表的指针中的对应指针及第二表的指针中的对应指针指向第二段。访问所存储的数据也可包括: 响应于对应第二段的第一虚拟存储区的写, 从第二段拷贝数据到第三段。访问所存储的数据也可包括: 使得第一表的指针中的对应指针指向第三段。访问所存储的数据也可包括: 使得第二表的指

针中的对应指针指向第三段。第一时间可与第二时间相同。第一时间也可不同于第二时间。每一数据段可以是一数据磁道。每一存储区可以是存储装置。

进一步根据本发明，一计算机程序产品包括：可执行编码提供具有指针表的虚拟存储区，所述指针指向至少两个其他存储区的段，虚拟存储区不包含数据段，可执行编码响应于访问虚拟存储区的数据的请求而确定其他存储区的哪一特定区包含数据，且可执行编码使用指针表访问其他存储区的特定区上的数据。计算机程序产品还可包括：可执行编码还将虚拟存储区与其他存储区的第一区关联，在这里，虚拟存储区代表其他存储区的第一区的数据的拷贝。计算机程序产品也可包括：当虚拟存储区初始与其他存储区的第一区关联时，可执行编码使得指针表的所有指针初始指向其他存储区的第一区的段。计算机程序产品还可进一步包括：响应于向其他存储区的第一区的第一段进行写，可执行编码拷贝第一段的数据到其他存储区的第二区上的第二段，并使表指针的对应指针指向第二段。计算机程序产品还可包括：可执行编码将数据指示器与其他存储区的第一区的段关联，在这里，用于特定段的数据指示器指示在其他存储区的第一区已经与虚拟存储区关联后，是否已向特定段执行写操作。计算机程序产品也可包括：响应于向其他存储区的第一区上的第一段进行写，以及数据指示器指示在其他存储区的第一区已与虚拟存储区关联后还没有对第一段执行写操作，可执行编码拷贝第一段的数据到其他存储区的第二区上的第二段，并使表指针的对应指针指向第二段。计算机程序产品也可包括：可执行编码发送状态信息给在从第一段拷贝数据到第二段之后使得写操作被执行的装置。每一数据段可以是一数据磁道。每一存储区可以是存储装置。

进一步地根据本发明，虚拟存储装置包括至少一用于将虚拟存储装置与标准存储装置关联的表，用于虚拟存储装置的数据段的存储器，在此，用于段的存储器的第一部分是标准存储装置的数据段，以及由至少一表提供的第一多个指针，在此至少部分指针指向对应于第

一部分的标准存储装置的段。虚拟存储装置也可包括：用于段的存储器的第二部分，所述段包含不同于标准存储装置的映射于第二段的对应段上的数据的数据；以及由至少一表提供的第二多个指针，在此第二多个指针指向不同于标准存储装置的装置的段。每一第二部分的段可能包含标准存储装置的对应段上的数据的早先版本。每一数据段可以是一数据磁道。

附图简要说明

图 1 为与在此描述的系统连同使用的存储装置的简图。

图 2 为一存储器的简图，其示出了与在此描述的系统连同使用的不同逻辑卷。

图 3 为根据于此描述的系统，表示虚拟装置的使用的简图。

图 4 为根据于此描述的系统，表示多个虚拟装置的使用的简图。

图 5 示出了与于此描述的系统连同使用的装置表的简图。

图 6 为根据于此描述的系统，读与虚拟装置连同使用的表的流程图。

图 7 为根据于此描述的系统，向与虚拟装置连同使用的表写的流程图。

图 8 为根据于此描述的系统，更改虚拟装置表及建立虚拟装置的流程图。

图 9 为根据于此描述的系统，更改用于处理日志装置的磁道的数据结构的流程图。

图 10 为根据于此描述的系统，所执行的与读虚拟装置有关的步骤的流程图。

图 11 为根据于此描述的系统，由磁盘适配器执行的与向标准逻辑装置写有关的步骤的流程图，其中对该标准逻辑装置已建立一虚拟装置。

图 12 为根据于此描述的系统，由主机适配器执行的与向标准逻辑装置写有关的步骤的流程图，其中对该标准逻辑装置已建立一虚拟装置。

图 13 为根据于此描述的系统，所执行的与向虚拟装置写有关的步骤的流程图。

图 14 为所执行的与除去虚拟装置有关的步骤的流程图。

具体实施例

参见图 1，存储装置 30 包括多个主机适配器 (HA) 32-34、多个磁盘适配器 (DA) 36-38、及多个磁盘驱动器 42-44。磁盘驱动器 42-44 的每一个均连接到磁盘适配器 36-38 的对应适配器。存储装置 30 还包括一可被主机适配器 32-34 与磁盘适配器 36-38 访问的全局存储器 46。存储装置 30 还包括一可访问全局存储器 46 的 RDF 适配器 (RA) 48。RA48 可经由一数据链路 52 与一个或多个另外的远程存储装置(未示出) 和/或一个或多个其他远程装置(未示出) 进行通信。主机适配器 32-34、磁盘适配器 36-38、全局存储器 46 及 RDF 适配器 48 均连接到总线 54，总线 54 提供来帮助其间的通信。

每一主机适配器 32-34 可连接到一个或多个访问存储装置 30 的主机(未示出)。主机读存储在磁盘驱动器 42-44 上的数据并写数据到磁盘驱动器 42-44。全局存储器 46 包含一保存来自磁盘驱动器 42-44 的数据磁道的高速缓存，以及用于可被主机适配器 32-34、磁盘驱动器 36-38 及 RDF 适配器 48 访问的表的存储器。注意，对于在此的讨论，数据块均被描述为一磁道或多磁道数据。然而，本领域的普通技术人员将感激，于此描述的系统可与任何适当增加的数据量或数据段一起工作，包括可能的不定增加的数据量和/或固定增加的数据量。

参看图 2，所示出的存储装置 30 包括多个标准逻辑装置 61-68。每一标准逻辑装置 61-68 可对应于可被一个或多个连接到存储装置 30 的主机访问的卷。每一标准逻辑装置 61-68 可以、也可不对应于

一磁盘驱动器 42-44。因而，例如，标准逻辑装置 61 可对应于磁盘驱动器 42、可对应于磁盘驱动器 42 的一部分、或可对应于磁盘驱动器 42 的一部分及磁盘驱动器 43 的一部分。对主机，每一标准逻辑装置 61-68 表现为磁盘存储器的一邻近块（contiguous block），即使每一标准逻辑装置 61-68 可、或不对应于磁盘驱动器 42-44 的实际邻近的物理存储器。

存储装置 30 也可包括多个虚拟装置 71-74。对连接到存储装置 30 的主机，虚拟装置 71-74 表现为包含数据存储器的一邻近块的卷。虚拟装置 71-74 的每个可表示在一时间点上标准逻辑装置 61-68 之一的整体的备份、标准逻辑装置 61-68 之一的一部分、及标准逻辑装置 61-68 之一的整体及标准逻辑装置 61-68 之一的一部分的结合。然而，如在别处的更详细的描述那样，虚拟装置 71-74 不包含来自标准逻辑装置 61-68 的跟踪数据。作为替代，虚拟装置 71-74 的每个连接到一存贮部分或所有跟踪数据的日志装置 76 或日志装置 78，如在别处的更详细的描述那样。虚拟装置 71-74 包含指向或标准逻辑装置 61-68 或日志装置 76、78 上的数据磁道的表。

虚拟装置 71 可表示一时间点上标准逻辑装置 61 的备份。如在别处的更详细的描述，虚拟装置 71 连接到日志装置 76，日志装置 76 包含跟踪数据以使虚拟装置 71 对主机容易表现为一时间点上标准逻辑装置 61 的备份。对于一个以上的虚拟装置使用单个日志装置是可能的。因而，虚拟装置 72-74 被示出并连接到日志装置 78。类似地，对于一个以上的虚拟装置代表时间点上的单个标准逻辑装置的多个备份也是可能的。因而，虚拟装置 72、73 被示出并作为时间点上标准逻辑装置 64 的备份。虚拟装置 72、73 可代表同一时间点标准逻辑装置 64 的备份，或者，做为选择，也可代表发生在不同时间的时间点上的标准逻辑装置 64 的备份。注意，仅部分标准逻辑装置 61-68 被示出并与虚拟装置 71-74 的对应装置关联，同时其他的标准逻辑装置 61-68 未被示出。

在一些实施例中，使用存储区代替存储装置实现在此描述的系统是可能的。因而，例如，虚拟装置 71-74 可以是虚拟存储区，标准逻辑装置 61-68 可以是标准逻辑区，及日志装置 76、78 可以是日志区。在一些例子中，这样的实现可允许混合逻辑/虚拟装置，其中，一单一逻辑装置的部分表现为标准逻辑装置，部分表现为虚拟装置，和/或部分表现为日志装置。因此，应该理解的是，在适当的例子中，对装置于此论述的基准也可应用于存储区，其可以、也可不与存储装置直接一致。

参考图 3，该简图示出了标准逻辑装置 82、虚拟装置 84 及日志装置 86。如上所述，虚拟装置 84 可代表一时间点上所有或部分标准逻辑装置 82 的备份。连接到存储装置的访问虚拟装置 84 的主机可能已主机访问标准逻辑装置 82 同样的方式访问虚拟装置 84。然而，虚拟装置 84 不包含任何来自标准逻辑装置 82 的跟踪数据。代替地，虚拟装置 84 包括多个指向或标准逻辑装置 82 或日志装置 86 上的磁道的表项目。

当虚拟装置被确立 84（例如，当一时间点备份源自标准逻辑装置 82 时），虚拟装置 84 被创建并被提供以适当的表项目，这些表项目在确立时指向标准逻辑装置 82 的磁道。访问虚拟装置 84 以读取磁道的主机基于指向标准逻辑装置 82 的磁道的虚拟装置 84 的表项目，从标准逻辑装置 82 读取适当的磁道。

在虚拟装置 84 被建立后，主机写数据到标准逻辑装置 82 是可能的。在那种情况下，存储在标准逻辑装置 82 上的在前的数据被拷贝到日志装置 86，且先前指向标准逻辑装置 82 的磁道的虚拟装置 84 的表项目改为指向日志装置 86 的新磁道。因而，访问虚拟装置 84 的主机或读取自虚拟装置 84 建立后还未变化的标准逻辑装置 82 的磁道，做为选择，也可读取在虚拟装置 84 建立后包含从标准逻辑装置 82 拷贝的数据的日志装置 86 的对应磁道。调整与读和写标准逻辑装置 82 及虚拟装置 84 有关的数据及指针将在此文的别处进行详细讨论。

在于此描述的实施例中，主机不直接访问日志装置 86。即，日志装置 86 专门连同虚拟装置 84（和在此文别处详细描述的其他虚拟装置）一起使用。此外，对于在此描述的实施例，标准逻辑装置 82、虚拟装置 84 及日志装置 86 可在单个存储装置 30 上提供。当然，在分开的存储装置上提供不同的逻辑装置及日志装置也是可能的，分开的存储装置使用，例如，RDF 协议或其他远程通信协议互连。此外，使一个或多个标准逻辑装置 82、虚拟装置 84 和/或日志装置 86 的部分在适当互连的分开的存储装置上提供也是可能的。

参考图 4，虚拟装置的使用的另一例子示出了一标准逻辑装置 92、多个虚拟装置 94-97 及一日志装置 98。在图 4 的例子中，虚拟装置 94 代表发生在上午 10 点这一时间点的标准逻辑装置 92 的备份。类似地，虚拟装置 95 代表中午 12 点标准逻辑装置 92 的备份，虚拟装置 96 代表下午 2 点标准逻辑装置 92 的备份，及虚拟装置 97 代表下午 4 点标准逻辑装置 92 的备份。注意，所有虚拟装置 94-97 可共享日志装置 98。此外，对于虚拟装置 94-97 的一个以上的表项目或虚拟装置 94-97 的表项目的一子集，指向日志装置 98 的同样的磁道是可能的。例如，虚拟装置 95 和虚拟装置 96 被示出具有指向日志装置 98 的同一磁道的表项目。

在于此讨论的一实施例中，日志装置 98 及于此讨论的其他日志装置由一组由存储装置 30 管理的日志装置提供。在这种情况下，由于虚拟装置要求额外的日志装置磁道，虚拟装置使用日志装置合用机制可使得更多的日志装置存储区被创建。这种方式的合用存储装置资源在现有技术中是公知的。不使用合用技术的其他技术也可用于提供日志装置存储区。

参考图 5，简图 100 图解了用于保持对装置信息的跟踪的表。第一表 102 对应于存储装置或存储装置的一元件所使用的所有装置，如 HA 和/或 DA。表 102 包括多个对应于由存储装置（或存储装置的部分）使用的所有逻辑装置的逻辑装置项目 106-108。表 102 中的项目包括对标准逻辑装置、虚拟装置、日志装置及其他类型的逻辑装置的描述。

表 102 的项目 106-108 中的每个对应于包含用于每一逻辑装置的信息的另一表。例如，项目 107 可对应于表 112。表 112 包括一首部，首部包含附加信息。表 112 还包括用于逻辑装置的每一柱面的项目 116-118。在于此公开的实施例中，逻辑装置可包含任何数量的柱面，取决于逻辑装置被怎样初始化。当然，在其他实施例中，逻辑装置可包含固定数量的柱面。

表 112 被示作包括一用于额外的磁道字节 119 的段。额外的磁道字节 119 以本文别处讨论的方式连同日志装置使用。在于此公开的实施例中，对日志装置的每一磁道有八个额外的磁道字节。对不是日志装置的装置，可以不用额外的磁道字节 119。

每一柱面项目 116-118 对应于一磁道表。例如，项目 117 可对应于包括一首部 124 的磁道表 122，首部 124 具有附加信息。磁道表 122 还包括用于每一磁道的项目 126-128。在于此公开的实施例中，对于每一柱面有 15 个磁道。然而，对其他的实施例，对每一磁道，有不同数量的磁道甚至是变化数量的磁道也是可能的。对标准逻辑装置及日志装置，项目 126-128 的每一项目的信息包括一指向（或直接的、或间接的）存储装置 30（或远程存储装置，如果系统是那样的结构）的磁盘驱动器 42-44 之一上的物理地址的指针。因而，磁道表 122 可用于将对应于表 102、112、122 的逻辑装置的逻辑地址定位至存储装置 30 的磁盘驱动器 42-44 上的物理地址。对虚拟装置，表 122 的项目 126-128 中的每一项目指向对应标准逻辑装置或对应日志装置的一磁道。然而，对其他的实施例，使用不同的机制是可能的，其中，表 102、112、122 仅用于包含跟踪数据的标准逻辑装置，而另一类型的表，如简单的磁道阵列，被虚拟装置使用以定位虚拟装置的磁道至对应的标准逻辑装置或日志装置的磁道。

日志装置的每一磁道或者是空闲的，意味着其未被虚拟装置使用或被分配，意味着磁道被一个或多个虚拟装置中的表项目指向。在于此公开的一实施例中，日志装置的磁道通过第一创建一日志装置的所有空闲磁道的双链表而被管理。用于双链表的指针由表 112 的额外磁

道字节 119 提供,从而用于日志装置的额外磁道字节 119 包含 8 个用于日志装置的每一磁道的字节。对日志装置的空闲的每一磁道,额外的 8 个字节包括一指向日志装置的下一空闲磁道的向前的指针及一指向日志装置的前一空闲磁道的向后的指针。以这种方式使用双链表使访问日志装置的空闲磁道变得容易。

此外,如果日志装置的一磁道被分配(即,由一个或多个虚拟装置使用),用于磁道的对应的额外磁道字节 119 可被用于指回标准逻辑装置的对应磁道。因而,在虚拟装置被建立后,当向标准逻辑装置执行一写时,来自标准逻辑装置的数据被拷贝到日志装置的新磁道且对应于日志装置的新磁道的额外磁道字节转而指回从其拷贝数据的标准逻辑装置的磁道。使日志装置的每一磁道指回标准逻辑装置的对应磁道是有用的,例如,在数据恢复情况时。

此外,对于此公开的一实施例,用于对被分配的磁道每磁道额外 8 字节的指针与数据存储在一起。即,当日志装置的特定磁道被分配时,指回标准逻辑装置的指针与额外的磁道字节 119 存储在一起,此外,在日志装置的磁道上,指针与磁道数据本身一起存储。对 CKD 格式的磁道,额外的 8 字节可被存储在块零位(block zero)中。对于 FBA 格式的磁道,额外的 8 字节可被存储在追加在磁道末端上的附加的块中。在于此公开的实施例中,一块是 512 字节,一 FBA 磁道包含 40 块,当附加的块被追加时其增长为 41 块。不同的磁道格式被公开,例如,授权给 Yanai 等人的美国专利号 5,206,939,在此其被组合于此用于参考。

图 5 的表 102、112、122 可存储在存储装置 30 的全局存储器 46 中。此外,对应于由特定主机访问的装置的表可存储在 HA32-36 的对应适配器的局部存储器中。此外,RA48 和/或 DA36-38 也可使用和局部地存储表 102、112、122 的部分。

参考图 6,流程图 140 图解,当主机从对应于可通过虚拟装置访问的磁道的装置表读取数据时所执行的步骤。即,流程图 140 图解获得关于磁道的信息,一用于虚拟装置的表项目指向该磁道。

处理在测试步骤 142 开始，如果感兴趣的磁道（即，对应于被读的表项目的磁道）在标准逻辑装置或日志装置上，则其被确定。其被确定是通过访问用于虚拟装置的装置表项目并确定用于感兴趣的磁道的表项目是否指向或标准逻辑装置或日志装置。如果在测试步骤 142 确定用于虚拟装置的表中的指针指向标准逻辑装置，接着，控制从步骤 142 转到步骤 148，在步骤 148 感兴趣的表项目被读取。在步骤 148 之后，处理完成。

如果在测试步骤 142 确定对于感兴趣的磁道的用于虚拟装置的装置表中的指针指向日志装置，接着，控制从步骤 142 转移到步骤 158，在步骤 158 感兴趣的日志表项目被读取。在步骤 158 之后，处理完成。

注意，在某些情况中，对数据的访问可被一标记或一锁控制以阻止有权同时使用数据的多进程。这在装置表正被读取或修改的情况特别有用。于此公开的系统期望用于控制多进程访问数据的多种机制中的任何一种，包括常规的软件和/或硬件锁的组合，也就是公知的“标记”或“信息量”。在一些情况下，访问数据的进程可能需要等待直到另一进程释放该数据。在一实施例中，硬件锁控制到软件锁（标记）的访问，因而进程首先获得硬件锁的控制、测试软件锁，接着，如果软件锁被清除，进程置位软件锁并接着释放硬件锁。如果该进程获得硬件锁并确定软件锁未被清除，该进程接下来释放硬件锁以使已经置位软件锁的进程在迟一点的时间可清除软件锁。进一步应注意，在某些情况下，首先读取对应于特定磁道的表项目、将磁道读进高速缓存插槽（如果磁道尚未在高速缓存中）、锁上高速缓存插槽、并接着重读对应的表项目是有用的。

参考图 7，流程图 170 图解了连同写信息到装置表所执行的步骤，其中装置表用于一对应于标准逻辑装置或日志装置的虚拟装置。处理在第一步骤 172 开始，该步确定对应于被写的装置表项目的特定磁道是否在标准逻辑装置或日志装置上。如果确定感兴趣的特定磁道在标准逻辑装置上，控制从步骤 172 转到步骤 178，在步骤 178 对应于被

写的装置表项目的磁道被锁定。当当前进程正在修改对应于磁道的装置表项目时，在步骤 178 锁定磁道阻止了其他进程访问该磁道，并阻止修改对应的表项目。在步骤 178 之后是步骤 182，写操作在该步被执行。在步骤 182 之后是步骤 184，在该步磁道被解锁。在步骤 184 之后，处理完成。

如果在测试步骤 172 确定对应于用于虚拟装置的表项目的磁道被修改指向日志装置，控制从测试步骤 172 转到步骤 194，在步骤 194 对应于被写的装置表的项目的日志装置的磁道被锁定。在步骤 194 之后是步骤 196，写操作在该步被执行。在步骤 196 之后是步骤 198，在该步磁道被解锁。在步骤 198 之后，处理完成。

参考图 8，流程图 210 图解了所执行的与修改对应于虚拟装置的装置表有关的步骤。这可与图 7 的流程图 170 对比，图 7 图解了修改用于标准逻辑装置或日志装置的装置表，所述装置被用于虚拟装置的装置表的磁道的项目所指向。在流程图 210 中，用于虚拟装置的装置表被修改，与用于标准逻辑装置的装置表或用于日志装置的装置表成对比。

处理在第一步骤 212 开始，在该步确定对表的修改是否与建立虚拟装置相关联。正如此文别处讨论的那样，建立虚拟装置包括，在虚拟装置被创建后，使得虚拟装置可用于主机的访问。建立虚拟装置使得虚拟装置与标准逻辑装置相关联（从而，代表在建立时的时间点上标准逻辑装置的备份）。在与标准逻辑装置关联之前，虚拟装置没有被建立并不可由主机访问。在被建立后，虚拟装置可被主机访问。

在步骤 212 确定对表的修改与建立虚拟装置关联，接着控制从步骤 212 转到步骤 214，在该步设置用于虚拟装置的装置锁以阻止其他进程访问该表。装置锁与此文别处讨论的高速缓存插槽锁相似。

在步骤 214 之后是步骤 216，在该步虚拟装置表的指针被调整指向标准逻辑装置的磁道，且置位对应于被建立的虚拟装置的标准逻辑装置的每一磁道的保护位。在于此公开的实施例中，标准逻辑装置的每一磁道有 16 位可被置为保护位，一个用于每一虚拟装置被确立到

标准逻辑装置。在一些实施例中，保护位可具有不与虚拟装置关联的用途。当用于标准逻辑装置的每一磁道的位可被置位时，建立的一个新虚拟装置在 16 位字段中可被分配一新数位位置。正如在此文别处详细讨论的那样，保护位被置位且其随后的向标准逻辑装置的写表明需要特殊的处理以提供确立到标准逻辑装置的虚拟装置。在此文别处对特殊处理进行了详细描述。还是在步骤 216，对于虚拟装置的装置表的磁道项目被全部修改为指向标准逻辑装置的对应磁道。因而，当虚拟装置被首先确立时，虚拟装置的装置表的所有指针指向标准逻辑装置的磁道。

在步骤 216 之后是步骤 217，在该步虚拟装置被设置为就绪状态，从而使得虚拟装置可被主机访问。在步骤 217 之后是步骤 218，在该步虚拟装置被解锁，从而允许其他进程访问。在步骤 218 之后，处理完成。

如果测试步骤 212 确定虚拟装置尚未被确立（即，一些其他的操作正被执行），控制从测试步骤 212 转到步骤 222 以锁定对应于正被修改的虚拟装置的装置表的项目的磁道。注意，在步骤 222 锁定的磁道也可能是标准逻辑装置（如果虚拟装置的装置表中的所感兴趣的项目指向标准逻辑装置）上的磁道或日志装置（如果感兴趣的项目指向日志装置）的一磁道。在步骤 222 之后是步骤 224，在该步对虚拟装置的装置表的修改被执行。在步骤 224 之后是步骤 226，在该步磁道被解锁。在步骤 226 之后，处理完成。

参考图 9，流程图 230 图解了所执行的与操作日志装置的磁道有关的步骤。如上所述，日志装置的磁道通过创建一空闲的日志装置的磁道（即，磁道可用于接受新数据）的双链表而得以维持。因而，如果需要一个或多个磁道连同对应的虚拟装置一起使用，则从双链表获得空闲的磁道，其被以常规方式修改以表明由虚拟装置所提供来使用的磁道不再空闲。相反地，如果一个或多个由一个或多个虚拟装置使用的磁道不再被需要，磁道被以常规方式送回到双链表，以表明磁道

是空闲的。图 9 的流程图 230 图解了所执行的与控制操作磁道的多进程访问磁道（和磁道指针）有关的步骤。

处理开始于测试步骤 232，在该步确定被执行的操作是否正修改空闲列表上的最适合的磁道。注意，仅修改空闲列表上的磁道，其通过，例如，将一空闲磁道从列表的一部分转移到另一部分或从一空闲列表转移到另一空闲列表（在多空闲列表的情况下）而实现修改，其不包括对对应于任何数据的磁道的修改。如果在测试步骤 232 确定被执行的修改不涉及空闲列表上的任何磁道，控制从步骤 232 转到步骤 234，在该步磁道被锁定以阻止其他进程访问。

在步骤 234 之后或步骤 232 之后（如果未到达步骤 234）是测试步骤 236，在该步确定操作是否涉及单独分配的磁道。对于任何涉及单独分配的磁道的操作，则不必锁住空闲磁道的日志装置表。如果在步骤 236 确定被执行的操作没有操作单独分配的磁道，则控制从步骤 236 转到步骤 238，在该步空闲磁道的日志装置表被锁住以阻止其他进程访问。在步骤 238 之后或步骤 236 之后（如果步骤 238 未被执行）是步骤 242，在该步执行修改。

在步骤 242 之后是测试步骤 244，其确定操作是否涉及单独分配的磁道。如果在测试步骤 244 确定被执行的修改不涉及单独分配的磁道，则控制从步骤 244 转到步骤 246，在该步日志装置空闲表被解锁。在步骤 246 之后或步骤 244 之后（如果未到达步骤 246）是测试步骤 248，其确定被执行的操作是否正修改空闲表上的单独磁道。如果在步骤 248 确定被执行的操作正修改空闲表上的单独磁道，则控制从步骤 248 转到步骤 252，在该步，在步骤 234 被锁住的磁道被解锁。在步骤 252 之后或步骤 248 之后（如果步骤 252 未被执行），处理结束。

参考图 10，流程图 280 图解了所执行的与从虚拟装置读取数据有关的步骤。处理开始于测试步骤 282，其确定对于虚拟装置的所感兴趣的磁道的装置表项目是否指向标准逻辑装置或执行日志装置。如果在步骤 282 确定表指向标准逻辑装置，则控制从步骤 282 转到步骤 284，在该步磁道被从标准逻辑装置读取。在步骤 284 之后，处理结

束。二选一地，如果在测试步骤 282 确定虚拟装置的装置表指向日志装置，则控制从步骤 282 转到步骤 286，在该步感兴趣的磁道被从日志装置读取。在步骤 286 之后，处理结束。

注意，在一些情况下，在测试步骤 282 之前，可能先确定被读取的感兴趣的磁道已经在高速缓存存储器（全局存储器）中。在这种情况下，磁道可从高速缓存存储器获得，而无须执行步骤 282、284、286 中的任何步骤。

参考图 11，流程图 300 图解了由 DA 执行的、与向标准逻辑装置的一磁道写有关的步骤，其中虚拟装置已在先前被确立到该标准逻辑装置。处理开始于第一步骤 302，其确定用于被写在标准逻辑装置上的磁道的任何保护位是否已经被置位。如果在测试步骤 302 确定保护位未被置位，则控制从步骤 302 转到步骤 304，在该步执行正常的写操作。即，在步骤 304，数据被以常规方式写到标准逻辑装置，不用关心已在先前被确立到标准逻辑装置的虚拟装置的存在。在步骤 304 之后，处理结束。

如果在测试步骤 302 确定一个或多个保护位已经被置位在正被写的标准逻辑装置的磁道上，控制从步骤 302 转到步骤 306，在该步获得日志装置的空闲磁道。日志装置的空闲磁道是需要的，以从标准逻辑装置的磁道拷贝数据。同样，如此文别处的更详细的描述，日志装置的空闲磁道可通过使用空闲磁道的双链表而被管理。因而，在步骤 306，其可能通过来回移动日志装置的空闲磁道的列表并适当地修改指针以移动空闲磁道之一而获得空闲磁道。

在步骤 306 之后是步骤 308，在该步，对对应于已在测试步骤 302 确定被置位的保护位的每一虚拟装置，虚拟装置的指针，其最初指向标准逻辑装置上的被写的磁道，在步骤 308 被修改指向在步骤 306 获得的日志装置的空闲磁道。如上所述，可能有一个以上虚拟装置被确立到标准逻辑装置。对已被确立到特定标准逻辑装置的每一虚拟装置，对标准逻辑装置的每一磁道，特殊的保护位将被置位。因此，在步骤 308，用于对应于在步骤 302 检测到的置位保护位的所有虚拟装

置的磁道指针均被改变。虚拟装置的装置表中的磁道指针被修改指向在步骤 306 获得的新磁道。

在步骤 308 之后是步骤 312，在该步数据被从标准逻辑装置拷贝到在步骤 306 获得的日志装置上的新磁道上。在于此公开的一实施例中，拷贝数据可通过，从磁盘存储器移动数据到存储装置的全局存储器（例如，进入一高速缓存插槽）、接着置位一写未决指示器以使数据被拷贝到在步骤 306 获得的日志装置的磁道上。步骤 312 表示从正被写的标准逻辑装置的磁道拷贝数据到在步骤 306 获得的日志装置的新磁道。由于所有指针在步骤 308 均被修改，在磁道被写之前任何已确立到标准逻辑装置的虚拟装置现在均指向旧数据（即，当虚拟装置被确立时存在于标准逻辑装置的磁道上的数据）。同样要注意的，与拷贝磁道有关，标准逻辑装置磁道的保护位被拷贝到用于日志装置上的磁道的虚拟装置映射位，这在本文别处有详细阐述。

在步骤 312 之后是步骤 314，在该步，在步骤 306 获得的日志装置的磁道被修改，从而表中的额外字节（另处讨论）被调整指回正被写的标准逻辑装置的磁道。使日志装置的磁道指向对应的提供数据的标准逻辑装置的磁道在许多情况下都是有用的。例如，其在与数据恢复有关时有用。在步骤 314 之后是步骤 316，在该步，正被写的标准逻辑装置的磁道的保护位被清除。在步骤 316 之后是步骤 318，在该步状态被发送到 HA。在步骤 318 之后，处理结束。

注意，一旦 HA 接收状态，HA 可执行一正常的写操作，且在这种情况下，在测试步骤 302，保护位将不被置位，因为保护位在步骤 316 被清除。执行写操作的 HA 看见在步骤 302 被置位的保护位并发送保护请求给适当的 DA。HA 接着可从 DA 断开并等待来自 DA 的、表明正常的写被执行的状态到达。当 HA 被断开并等待来自 DA 的状态时，DA 可执行在流程图 300 中公开的步骤。这将在下面详细描述。

参考图 12，流程图 320 图解了由 HA 执行的、与向标准逻辑装置写有关的步骤，其中一个或多个虚拟装置已被确立到该标准逻辑装置。处理开始于第一测试步骤 322，其确定用于正被写的标准逻辑装

置的磁道的保护位是否被置位。如果在测试步骤 322 确定没有保护位被置位，则控制从步骤 322 转到步骤 324，在该步执行正常的写。在步骤 324 之后，处理结束。

如果在测试步骤 322 确定用于正被写的标准逻辑装置的磁道的一个或多个保护位被置位，控制从步骤 322 转到步骤 326，在该步，HA 发送一请求给 DA 表明用于磁道的保护位被置位。当 DA 接收在步骤 326 发送的请求时，DA 执行在上面讨论的、在图 11 的流程图 300 中提出的操作。在步骤 326 之后是步骤 328，在该步，HA 从 DA 断开，以允许其他的进程和/或其他 HA 与 DA 一起执行操作。

在步骤 328 之后是步骤 332，在该步，HA 等待 DA 执行图 11 的流程图 300 中提出的操作并发送状态给 HA 以表明适当的步骤已被执行以处理已置位的保护位。在步骤 332 之后，处理转回步骤 322，用于标准逻辑装置的磁道的保护位再次被测试。注意，在第二次反复上，期望正被写的标准逻辑装置的磁道的保护位已在步骤 322 被清除，因为 DA 应在执行流程图 300 的步骤时清除了保护位。当然，在 DA 清除保护位与步骤 322 被再次执行之间，新虚拟装置被确立到标准逻辑装置总是可能的。然而，通常所期望的是对标准逻辑装置的特定磁道的步骤 322 的第二次反复将确定所有保护位均被清除，且控制将从步骤 322 转到步骤 324 以执行一正常的写。

参考图 13，流程图 340 图解了所执行的与向虚拟装置写有关的步骤。流程图 340 表示由 HA 及 DA 执行的步骤，因而可被提供为两个流程图，类似于图 11 的流程图 300 及图 12 的流程图 320。然而，可由那些本领域普通技术人员理解的是，流程图 340 可表示类似于那些在流程图 300、320 中提出的、并在说明书的对应部分中描述的步骤的步骤的一部分。

处理开始于第一步骤 342，其确定虚拟装置是否指向标准逻辑装置。如果“是”，则控制从测试步骤 342 转到步骤 344，在该步获得日志装置的空闲磁道。在步骤 344 之后是步骤 346，来自对应于被写的磁道的标准逻辑装置的数据被从标准逻辑装置拷贝到在步骤 344

获得的日志装置的磁道上。在步骤 346 之后是步骤 348，在该步，用于磁道的虚拟装置指针被调整指向在步骤 344 获得的磁道。在步骤 348 之后是步骤 352，对应于虚拟装置的保护位在标准逻辑装置的磁道数据中被清除，因而表明，当向标准逻辑装置的磁道写时，不需要任何代表虚拟装置的特殊处理。在步骤 352 之后是步骤 354，在该步写被执行。在步骤 354，被写的数据可以是被写到在步骤 344 获得的磁道的一磁道或一磁道的一部分。在步骤 354 之后，处理结束。如果数据对应于整个磁道，则可能排除步骤 346（该步从标准逻辑装置的磁道拷贝数据到日志装置的新磁道），因为在步骤 354 写所有磁道的数据会盖写在步骤 346 拷贝的所有数据。

如果在测试步骤 342 确定用于被写的虚拟装置的磁道的指针没有指向标准逻辑装置，则控制从步骤 342 转到测试步骤 356，其确定是否一个以上的虚拟装置已被确立到标准逻辑装置。如果没有，则控制从步骤 356 转到步骤 358，在该步向日志装置的磁道执行一正常的写操作。如果在测试步骤 356 确定有一个以上的虚拟装置被确立到标准逻辑装置，则控制从步骤 356 转到步骤 362，在该步从日志装置获得空闲磁道。

在步骤 362 之后是步骤 364，对应于被写的虚拟装置的磁道的数据被拷贝到在步骤 362 获得的磁道。在步骤 364 之后是步骤 366，虚拟装置指针被调整指向新磁道。在一实施例中，用于正被写的虚拟装置的指针被使得指向新磁道。供选择地，也可能不改变用于正被写的虚拟装置的指针，而是调整所有其他虚拟装置的所有指针指向在步骤 366 的磁道。

在步骤 366 之后是步骤 368，在该步，对于日志装置的磁道的虚拟装置映射位被修改。对日志装置磁道，虚拟装置映射位可被用于表明哪些虚拟装置指向每一磁道，在一实施例中，有 16 个虚拟装置映射位且每位对应于一特定的虚拟装置。因而，在步骤 356 的测试可检查对于磁道的虚拟装置映射位。

在步骤 368 之后是步骤 369，在该步执行写。注意，是否向在步骤 362 获得的磁道执行写或由被写的虚拟装置最初指向的磁道执行写取决于在步骤 366 指针是怎样被调整的。然而，在所有情况下，数据被写到由正被写数据的虚拟装置所指向的磁道。在步骤 369 之后，处理结束。

参考图 14，流程图 370 图解了所执行的与消除（即，除去）一虚拟装置有关的步骤。一旦虚拟装置已被确立并已用于其特定的目的，就可能想要消除虚拟装置。处理开始于第一步骤 372，指针被设定指向虚拟装置的第一磁道。虚拟装置通过检查对应于虚拟装置的每一磁道而被消除。

在步骤 372 之后是步骤 374，其确定正被检查的虚拟装置的磁道是否指向标准逻辑装置。如果是，则控制从步骤 374 转到步骤 376 以清除对应于被消除的虚拟装置的标准逻辑装置的磁道上的保护位。在步骤 376 之后是步骤 378，为了通过检查下一磁道而连续处理，指针指向虚拟装置的下一磁道。在步骤 378 之后是步骤 382，其确定处理是否结束（即，虚拟装置的所有磁道已经被处理）。如果没有，则控制从步骤 382 转回到上述的测试步骤 374。

如果在测试步骤 374 确定被检查的虚拟装置的磁道没有指向标准逻辑装置，则控制从步骤 374 转到步骤 384，在该步，对应于被消除的虚拟装置的日志装置的磁道上的虚拟装置映射位被清除。日志装置的每一磁道可有一组表明哪些虚拟装置使用日志装置的磁道的虚拟装置映射位。因而，在步骤 384，对应于被消除的虚拟装置的虚拟装置映射位被清除。

在步骤 384 之后是测试步骤 386，其确定在步骤 384 被清除的位是否是为磁道设定的虚拟装置映射位的最后一位。换言之，测试步骤 386 确定是否有其他的虚拟装置正使用日志装置上的磁道。如果在测试步骤 386 确定最后的虚拟装置映射位在步骤 384 已被清除（从而，没有其他虚拟装置使用该磁道），则控制从步骤 386 转到步骤 388，在该步，日志装置的磁道被送回日志装置磁道的空闲列表，如此文别

处所讨论的那样。在步骤 388 之后或在步骤 386 之后（如果确定在步骤 384 被清除的位不是日志装置的磁道的最后一位虚拟装置映射位）是步骤 378，如上所述，虚拟装置的下一磁道被指向以用于随后的检查。一旦对应于虚拟装置的所有磁道已经被处理，与虚拟装置关联的表及其他数据结构也可被消除，尽管在一些实施例中，自虚拟装置的表及其他数据结构可被保持，只要在虚拟装置被解除确立后，虚拟装置还没有调整为可为主机使用。

在本发明结合不同的实施例被公开的同时，在其上的修改对那些本领域技术人员是显而易见的。因此，本发明的精神和范围在下述的权利要求中被提出。

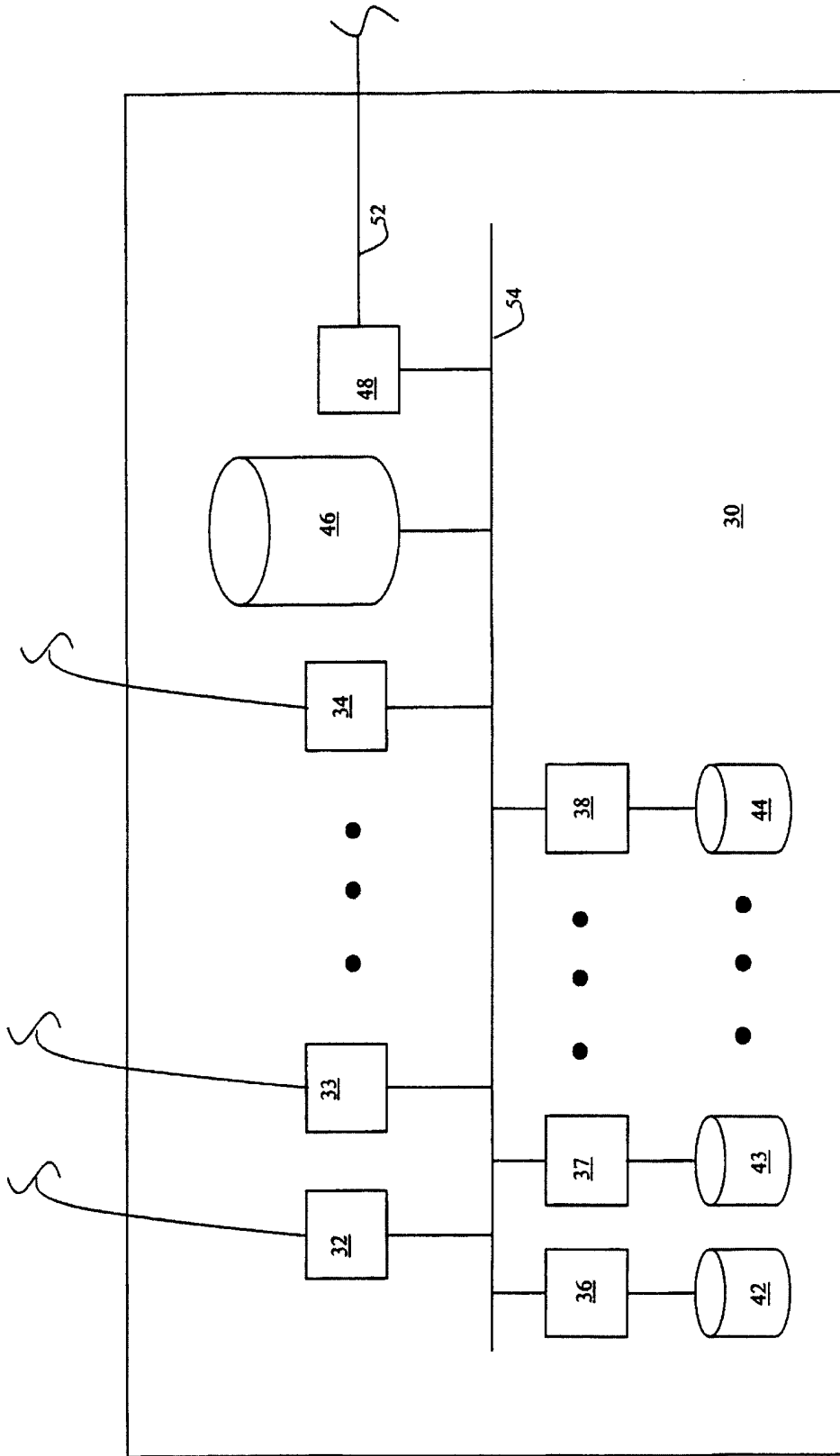


图 1

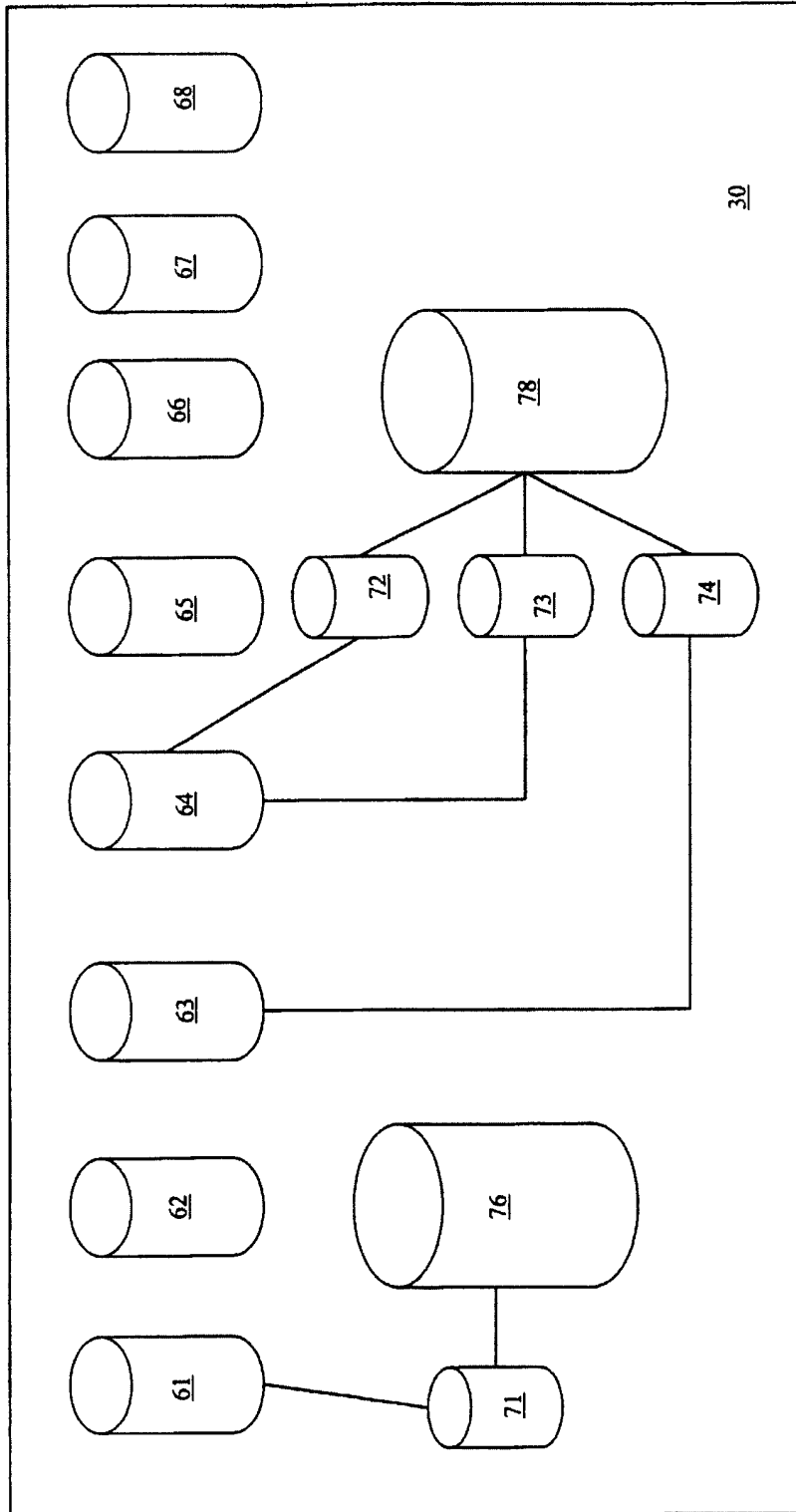


图 2

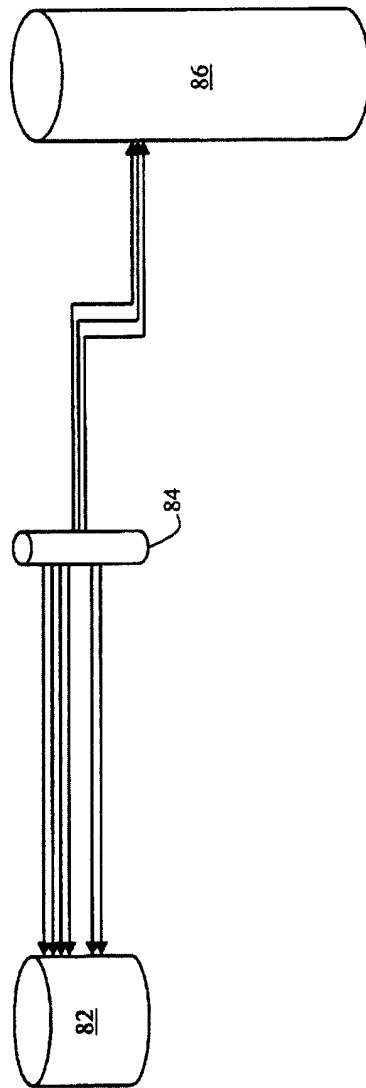


图 3

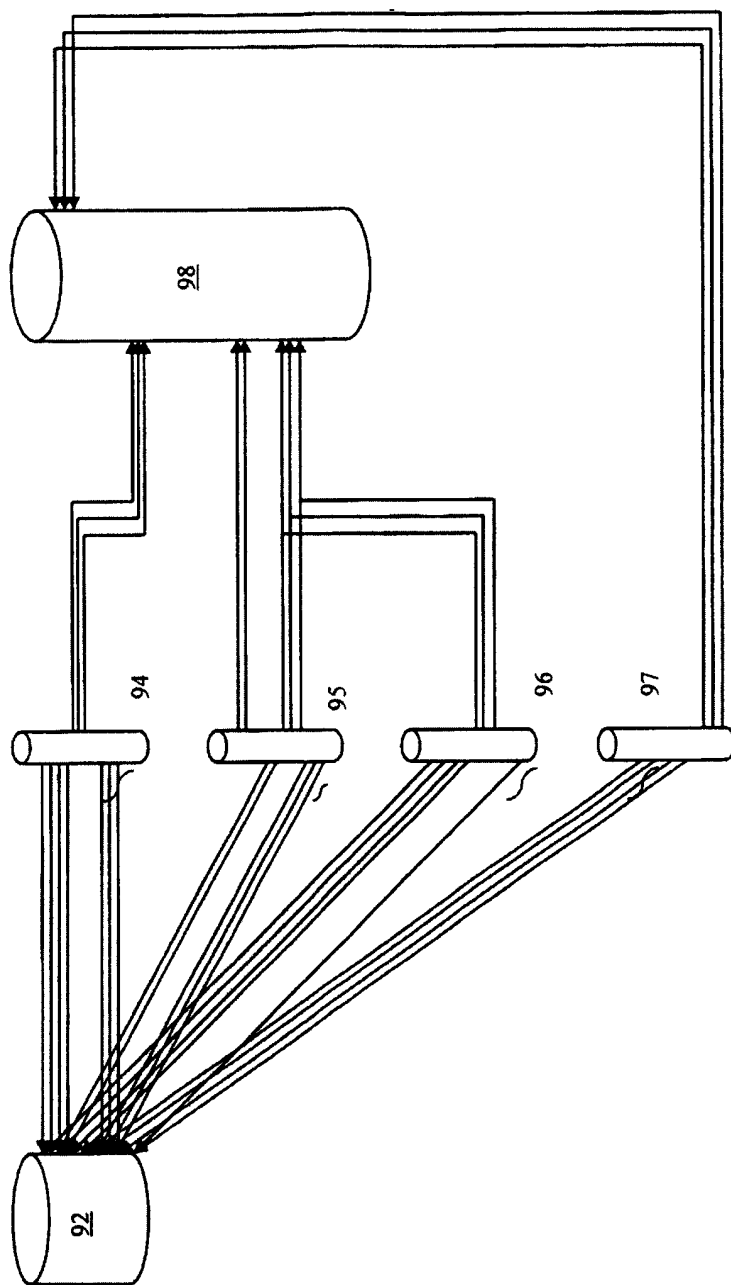


图 4

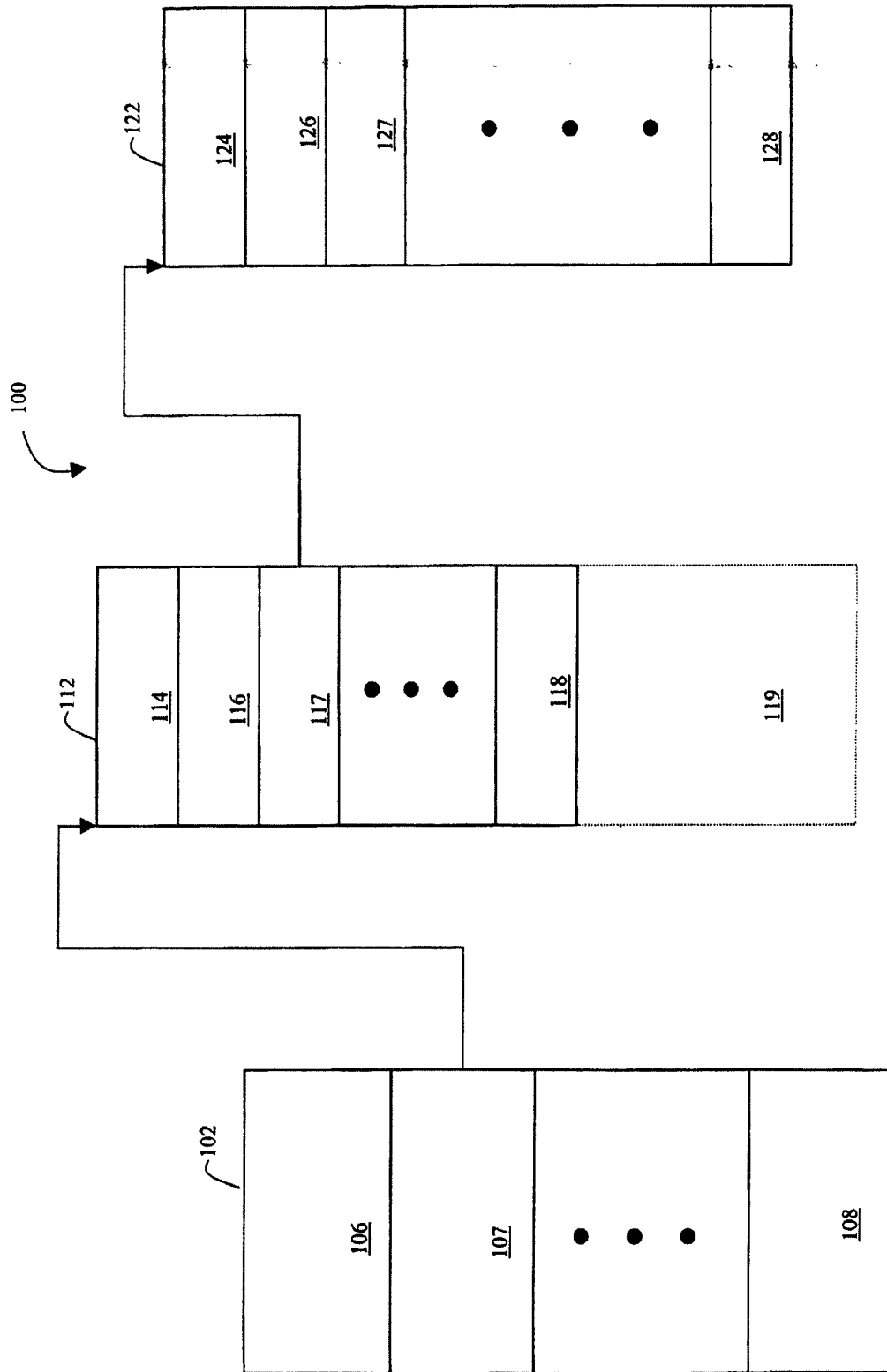


图 5

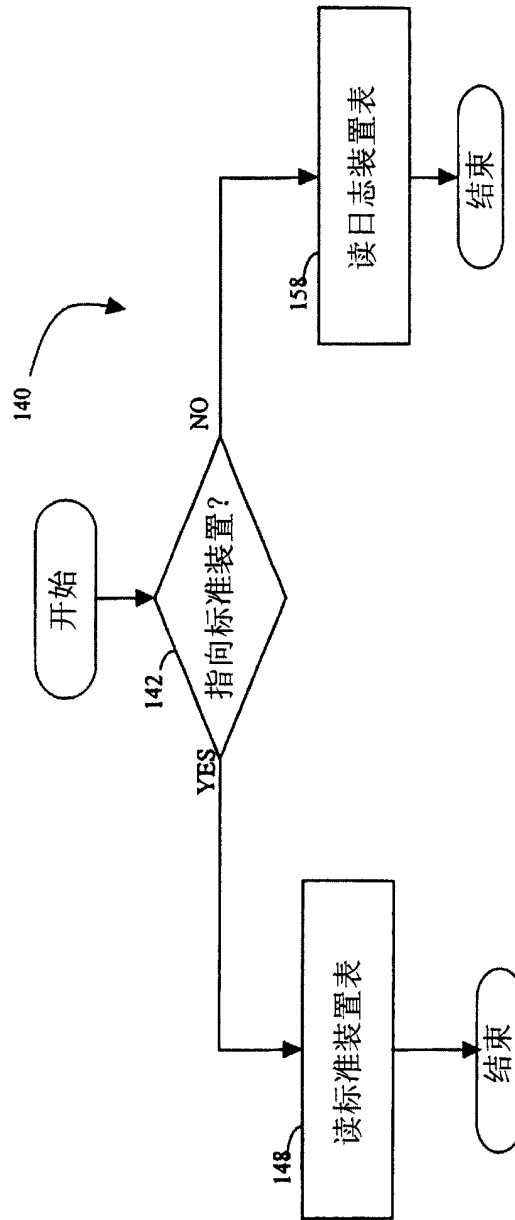


图 6

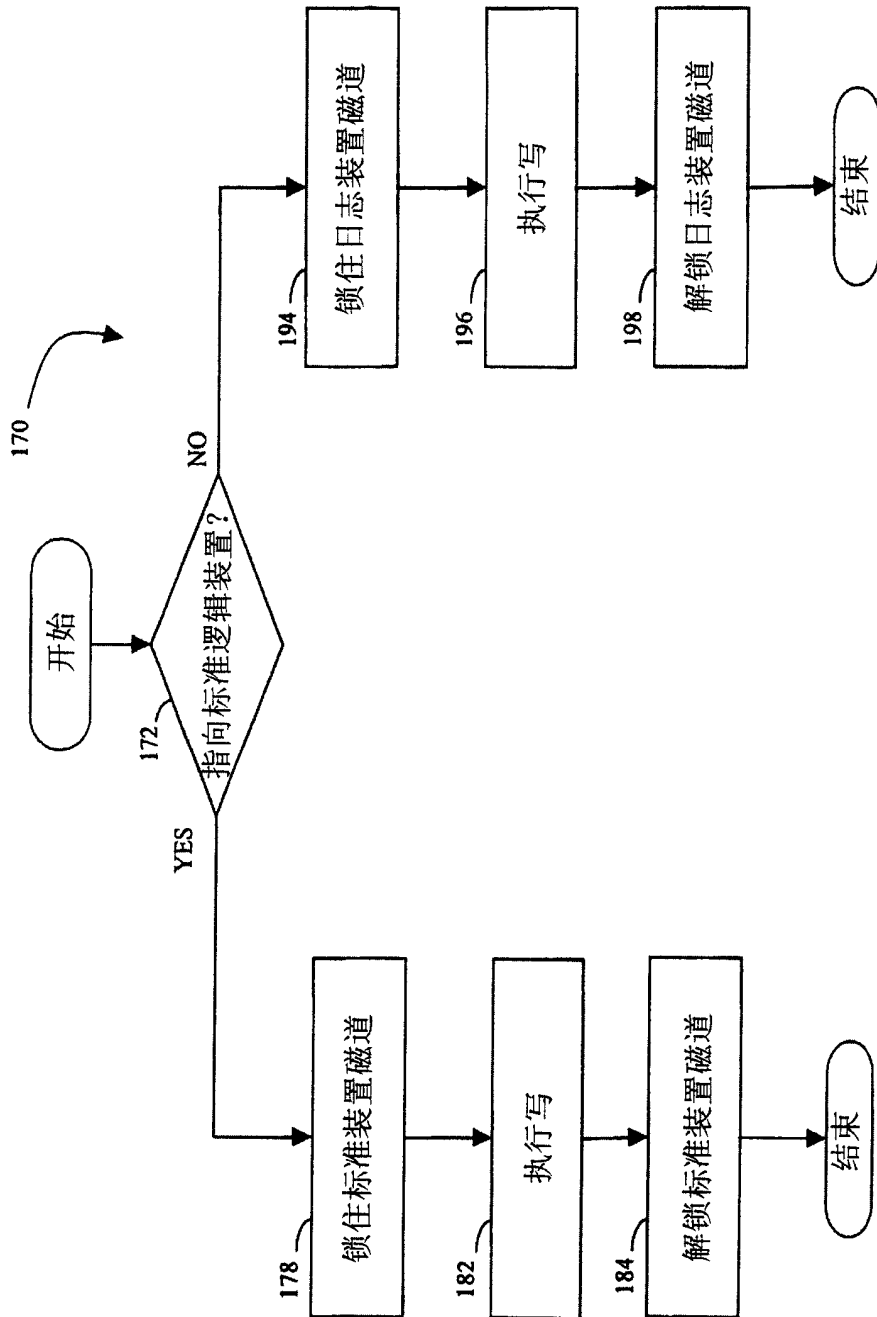


图 7

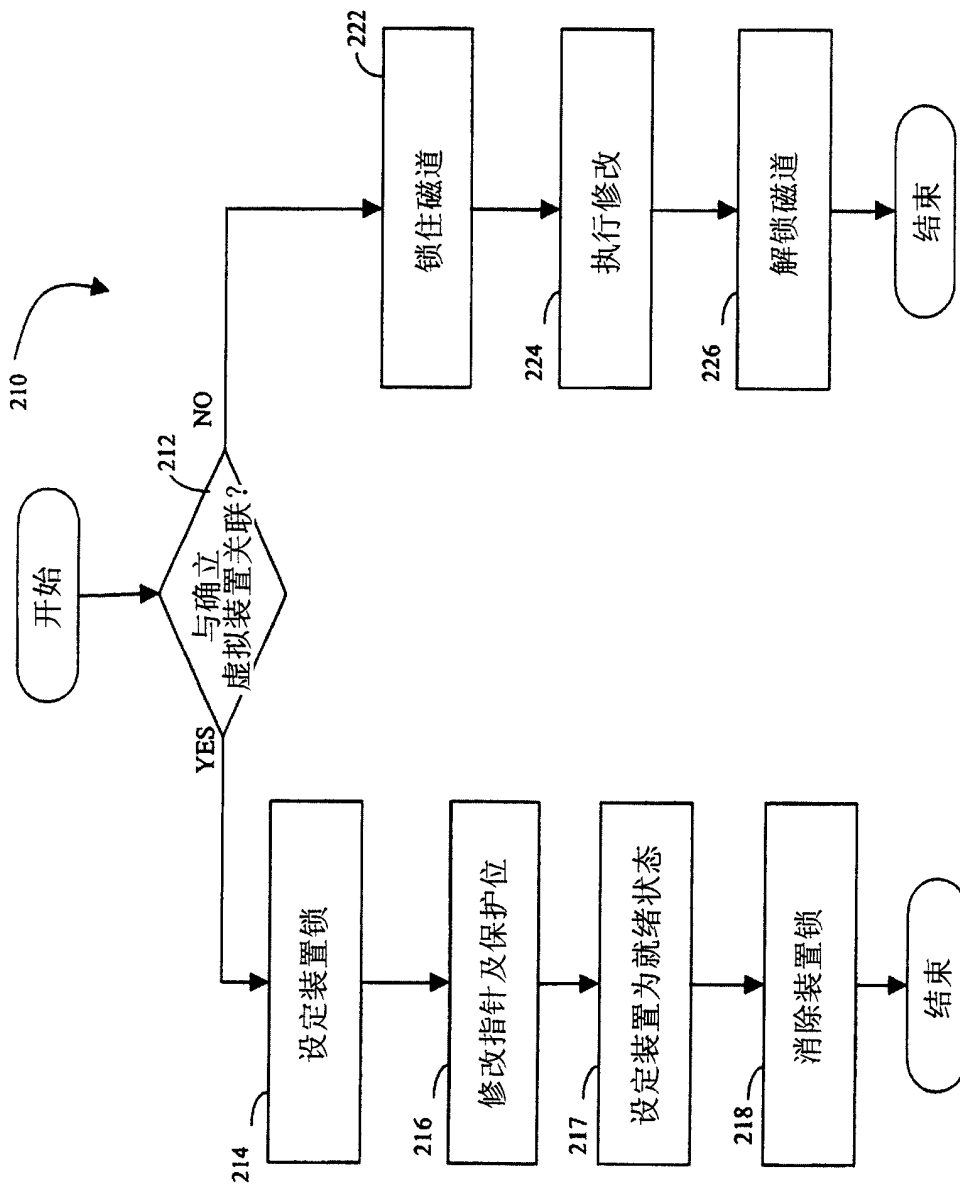


图 8

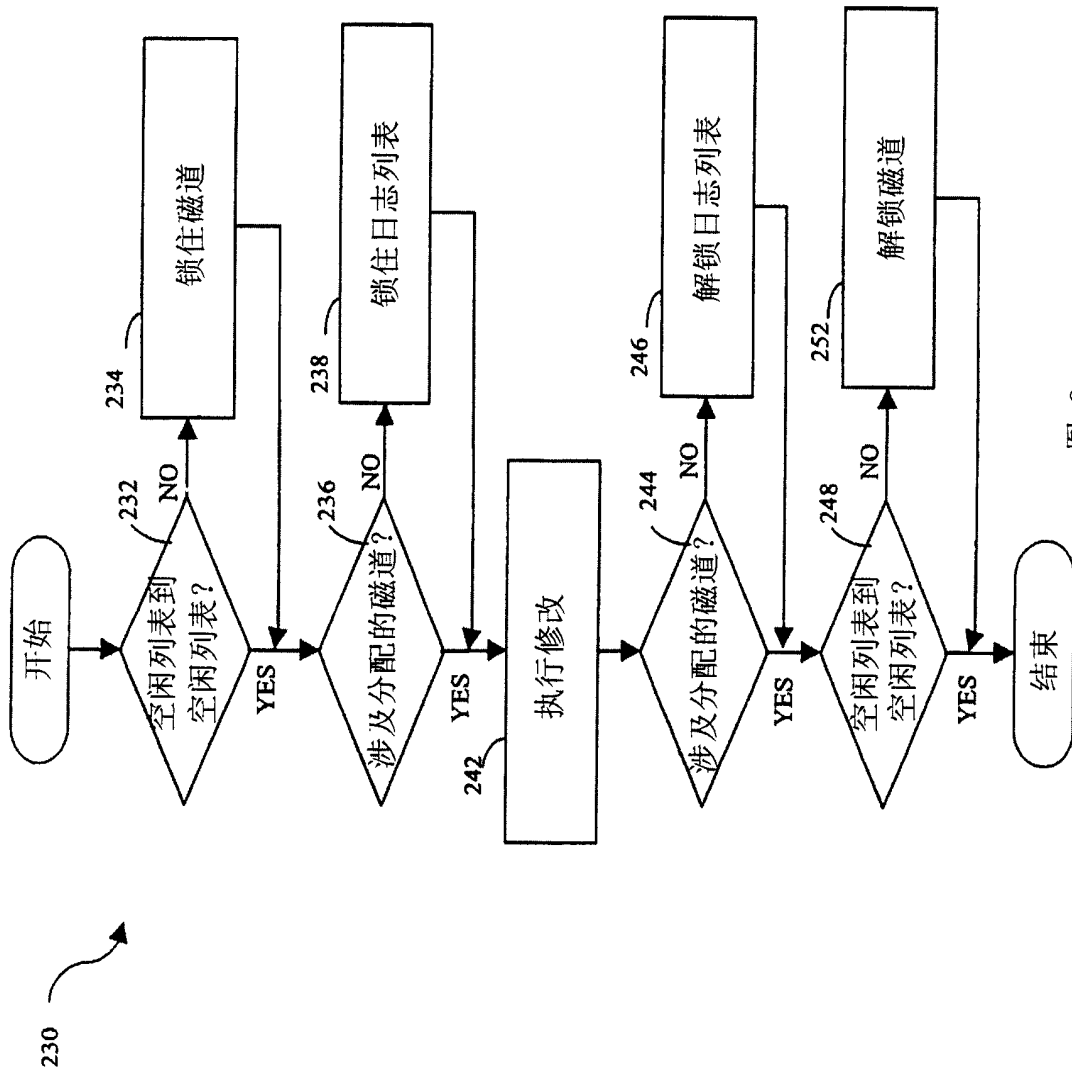


图 9

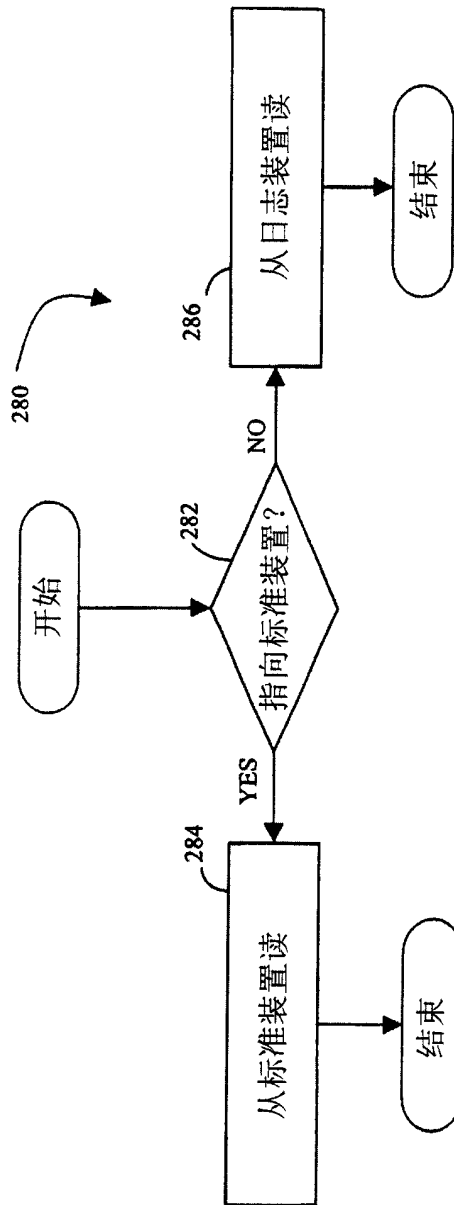


图 10

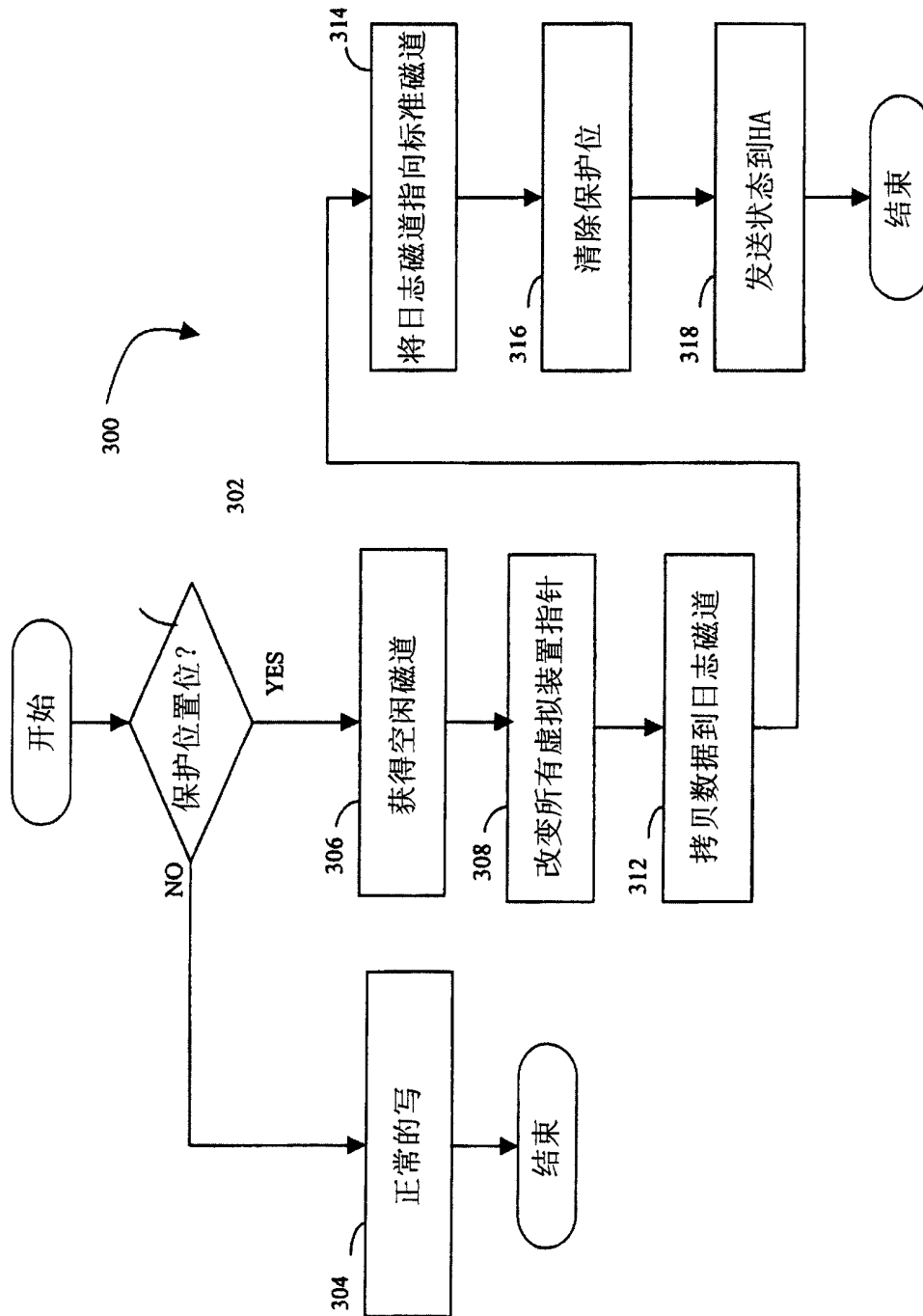


图 11

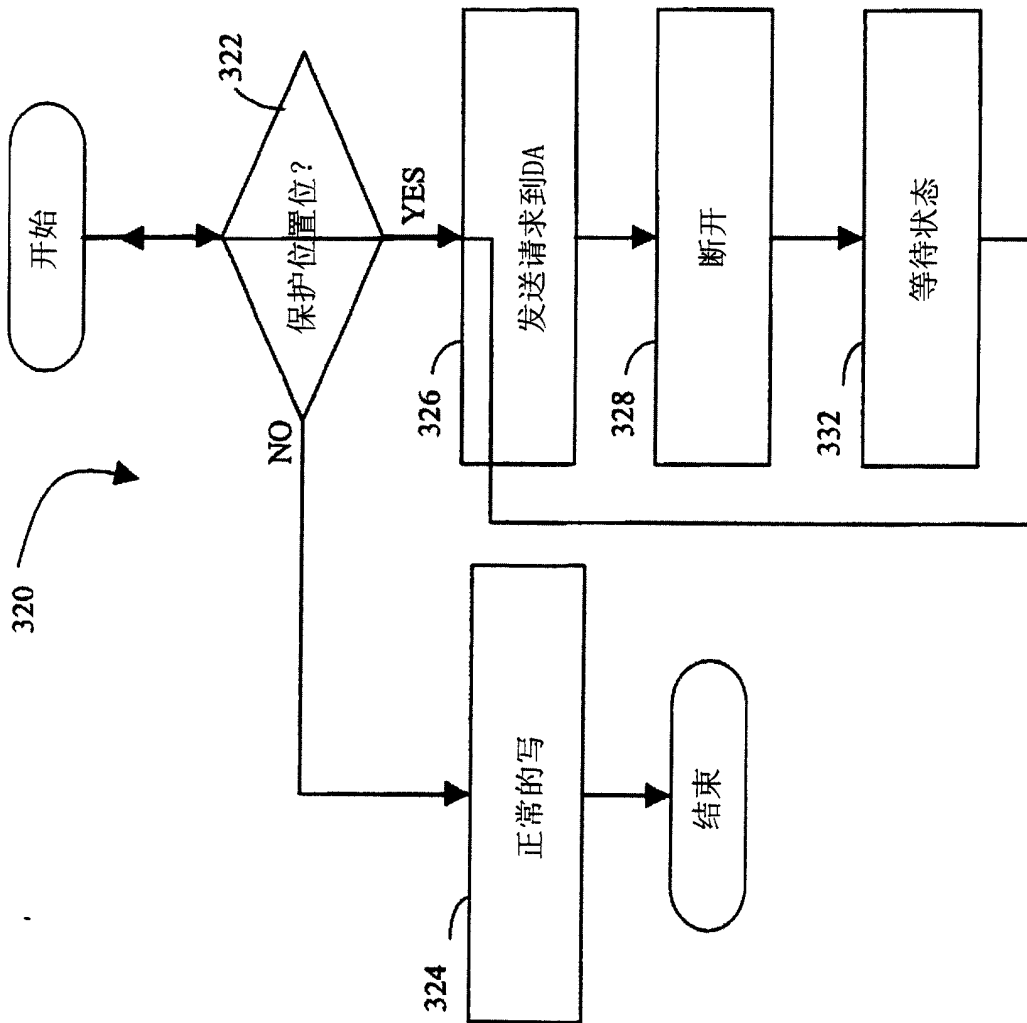


图 12

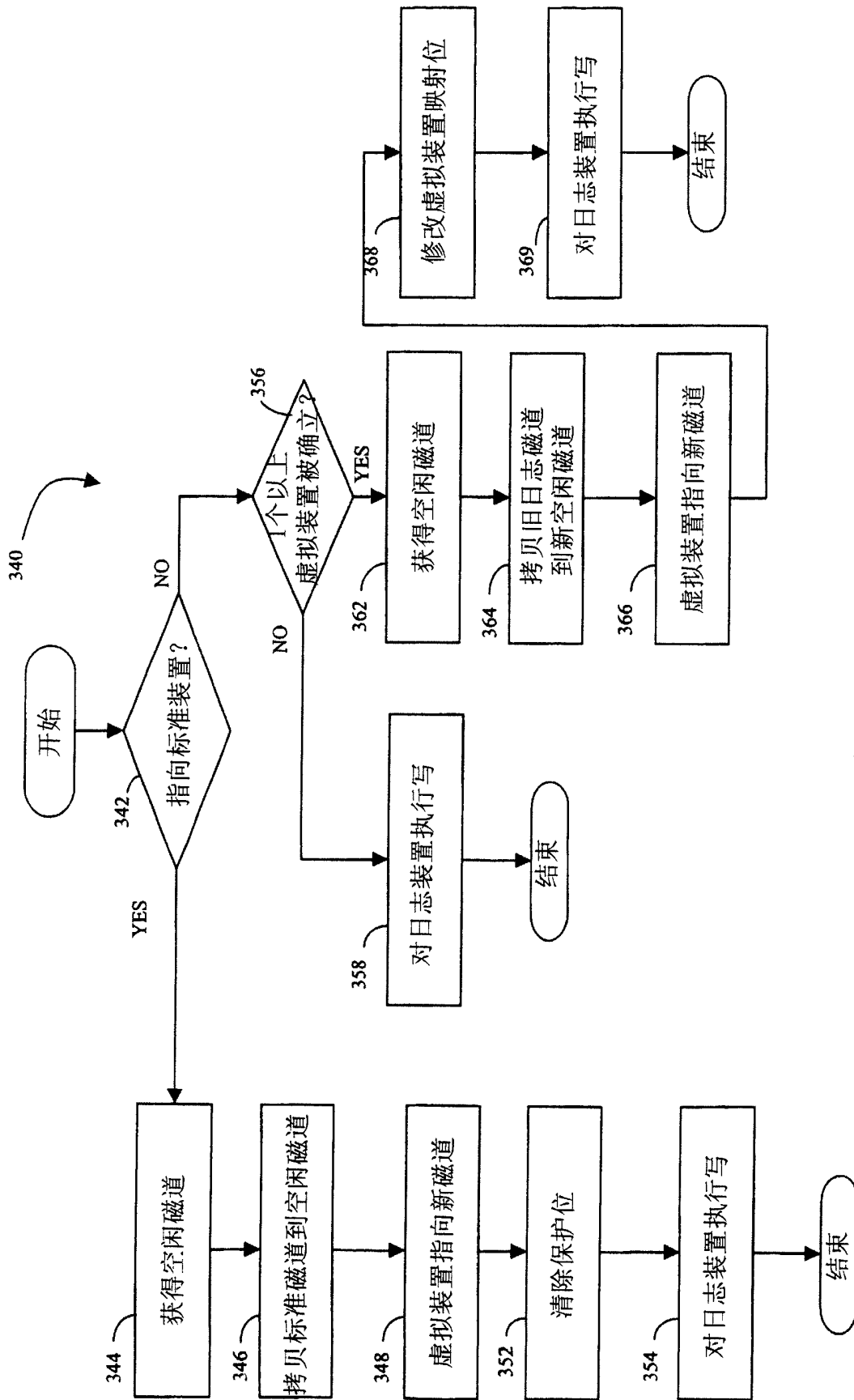


图 13

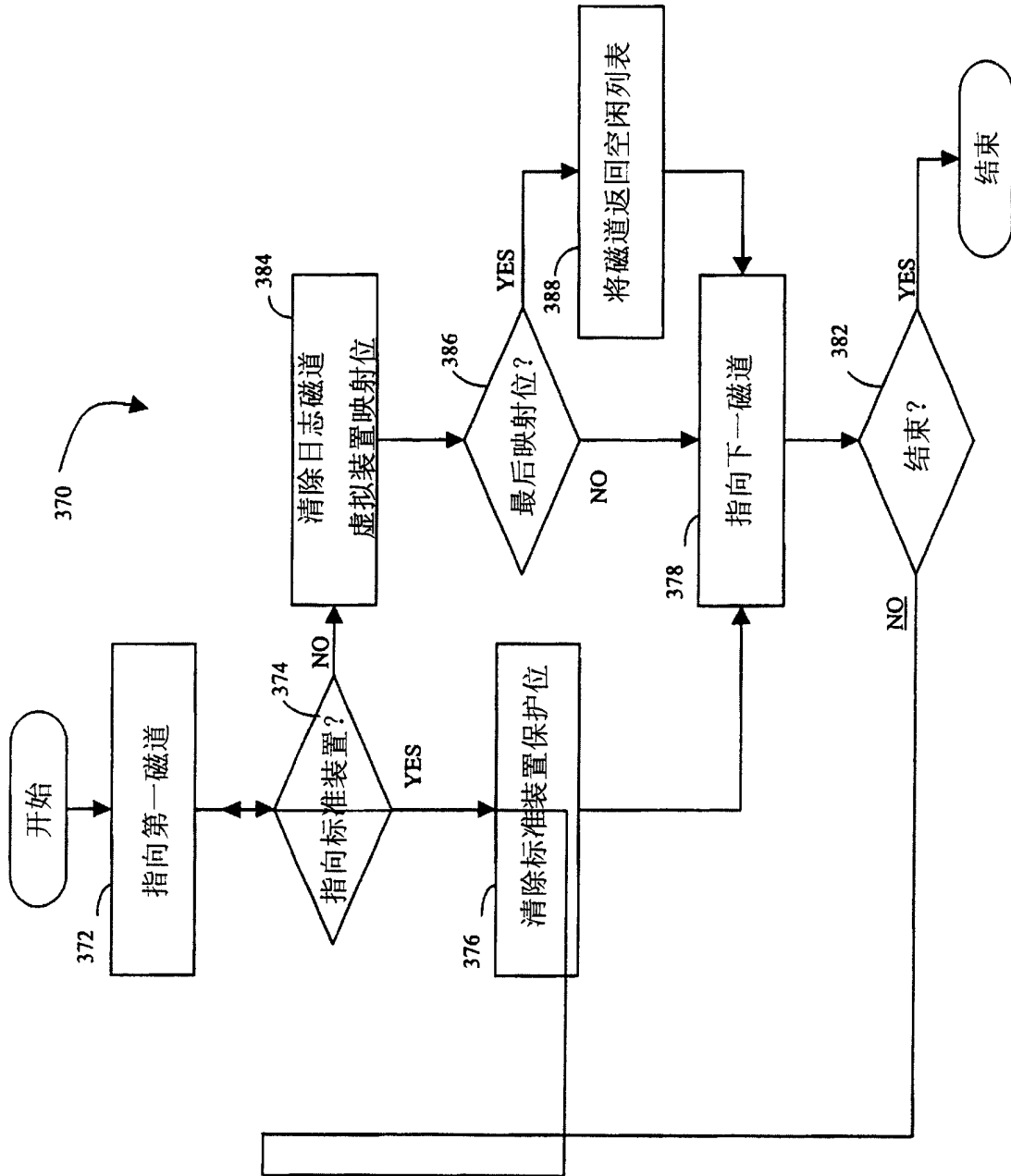


图 14