



[12] 发明专利说明书

专利号 ZL 98810177.7

[45] 授权公告日 2005 年 10 月 19 日

[11] 授权公告号 CN 1223945C

[22] 申请日 1998.10.5 [21] 申请号 98810177.7

[30] 优先权

[32] 1997.10.16 [33] US [31] 08/951,644

[86] 国际申请 PCT/US1998/021017 1998.10.5

[87] 国际公布 WO1999/021093 英 1999.4.29

[85] 进入国家阶段日期 2000.4.14

[71] 专利权人 M-系统快闪盘开拓者公司

地址 以色列特拉维夫

[72] 发明人 艾米尔·班 尤金·齐尔班南

审查员 邹 斌

[74] 专利代理机构 中国国际贸易促进委员会专利
商标事务所

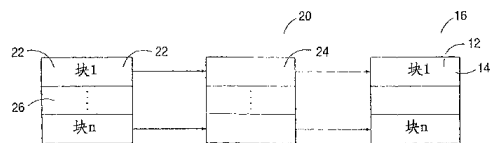
代理人 王以平

权利要求书 10 页 说明书 15 页 附图 6 页

[54] 发明名称 改进的闪速文件系统

[57] 摘要

一种闪速存储器设备以及对该设备写入并且重新组织该设备的方法。该闪速存储器设备(20)包括物理设备(10)、虚拟设备(22)以及将虚拟设备的虚拟地址与物理设备的物理地址联系起来的虚拟映射图(24)。



1.一种存储器组织方法,在该存储器中数据只能被写入还没有被写过的部分,这样,该存储器的被写过部分必须被擦除以便成为没有被写过的,并且其中用于读或写数据的存储器部分的大小不同于用于擦除的最小存储器部分的大小,该方法包括步骤:

(a) 提供该存储器的若干个物理单元,每个所述物理单元的大小分别等于用于擦除的最小存储器部分的大小,每个所述物理单元都由一个物理单元号指定,而且每个所述物理单元都被分解为若干物理块,每个所述物理块的大小分别等于用于读或写数据的存储器部分的大小,并且每个所述物理块都由所述物理单元中的一个物理块偏移指定;

(b) 提供该存储器的若干个虚拟单元,每个虚拟单元都由一个虚拟单元号所指定,并且每个所述虚拟单元体现了若干个虚拟块,每个所述虚拟块由所述虚拟单元中的一个虚拟块偏移指定;

(c) 将每个虚拟单元映射到至少一个物理单元,以形成一个虚拟映射图;

(d) 根据所述虚拟映射图,将所述虚拟单元中的每个虚拟块映射到所述至少一个物理单元中的一个物理块;

(e) 接收在某个虚拟块上写入数据的写命令;

(f) 寻找包含所述虚拟块的一个虚拟单元;

(g) 在被映射到所述虚拟单元的物理单元中寻找一个可写入的块;

(h) 将所述数据写入所述可写入物理块以形成一个被写过的物理单元;

(i) 如果在所述被写过的物理单元中找不到一个没有被写过的物理块,则寻找第二个物理单元;

(j) 将所述数据写入所述第二个物理单元中的一个可写入物理块;
并且

(k) 通过另外将所述虚拟单元映射到所述第二个物理单元来更新所述虚拟映射图,以使所述虚拟单元既对应所述第二个物理单元,又对应

所述写过的物理单元，所述第二个物理单元和所述被写过的物理单元形成一个物理单元链。

2.权利要求1的方法，其中所述第二个物理单元的所述可写入物理块有一个物理块偏移，并且所述物理块偏移对应所述被映射的虚拟单元的所述虚拟块偏移。

3.权利要求1的方法，其中所述第二个物理单元的所述可写入物理块具有一个物理块偏移，并且所述物理块偏移不同于所述被映射的虚拟单元的所述虚拟块偏移。

4.权利要求1的方法，进一步包括步骤：

(l)如果在任何物理单元中找不到一个没有被写过的物理块，则寻找对应某个链中多个物理单元的第二个虚拟单元；

(m)寻找所述链中的所述最后物理单元；

(n)将数据从所述被写过的物理单元的每个所述物理块中移动到所述最后物理单元的一个可写入物理块中，所述可写入物理块具有与所述被写过物理单元的所述物理块相同的块偏移；并且

(o)通过将所述虚拟单元映射到所述最后物理单元来更新所述虚拟映射图，以使所述虚拟单元只对应所述最后物理单元。

5.权利要求4的方法，进一步包括步骤：

(p)除所述最后物理单元外，擦除所述链中所有所述被写过物理单元。

6.权利要求1的方法，进一步包括步骤：

(l)如果在所述链上找不到可分配的没有写过的物理单元，则分配一个没有被写过的物理单元以便重新组织；

(m)将数据从所述被写过物理单元的每个所述物理块中移动到所述没有被写过的物理单元的一个可写入物理块中，并且

(n)通过将所述虚拟单元映射到所述没有被写过的物理单元以更新所述虚拟映射图，以使所述虚拟单元只对应所述没有被写过的物理单元。

7.一种存储器组织方法，在该存储器中数据只能被写入存储器中还没有被写过的部分，这样，该存储器的被写过部分必须被擦除以便成为没

有被写过的，并且其中用于读或写数据的存储器部分的大小不同于用于擦除的最小存储器部分的大小，该方法包括步骤：

(a) 提供该存储器的若干个物理单元，所述每个物理单元的大小分别等于擦除的最小存储器部分的大小，每个所述物理单元都由一个物理单元号指定，而且每个所述物理单元都被分解为若干物理块，每个所述物理块的大小分别等于用于读或写数据的存储器部分，并且每个所述物理块都由所述物理单元中的一个物理块偏移指定；

(b) 提供该存储器的若干个虚拟单元，每个虚拟单元都由一个虚拟单元号所指定，并且每个所述虚拟单元体现了若干个虚拟块，每个所述虚拟块由所述虚拟单元中的一个虚拟块偏移指定；

(c) 提供一个虚拟映射图，用于将每个虚拟单元映射到至少一个物理单元；并且

(d) 将所述虚拟单元中的每个虚拟块映射到所述至少一个物理单元中的一个物理块；

(e) 接收在某个虚拟块上写入数据的写命令；

(f) 寻找包含所述虚拟块的一个虚拟单元；

(g) 在被映射到所述虚拟单元的物理单元中寻找一个可写入的块；

(h) 将所述数据写入所述可写入物理块；

(i) 如果在一个写过的物理单元中找不到没有被写过的物理块，则分配一个没有被写过的物理单元来形成一条物理单元链，以使所述没有被写过的物理单元成为所述链中的最后一个物理单元；

(j) 将所述数据写入所述最后物理单元中的一个没有被写过的物理块；

(k) 除了在步骤(j)中被写入的数据外将数据从每个所述物理块移到所述最后物理单元的一个可写入物理块中；并且

(l) 将所述虚拟单元映射到所述写过的物理单元以使所述虚拟单元就对应所述写过的物理单元，从而更新所述虚拟映射图。

8.一种存储器数据写入方法，其中数据只能被写入存储器的某个没有被写过的部分，这样，存储器的被写过部分就必须被擦除以便成为没有

被写过的，该方法包括步骤：

(a) 提供若干个物理单元，而且每个物理单元又被分成若干个物理块，每个所述物理单元具有一个物理单元号并且每个所述物理块有一个在所述物理单元内的物理块偏移；

(b) 提供若干个虚拟单元，而每个虚拟单元又被分成若干个虚拟块，每个所述虚拟单元有一个虚拟单元号并且每个所述虚拟块有一个在所述虚拟单元内的虚拟块偏移，每个虚拟单元被映射到至少一个物理单元上；

(c) 接收在某个虚拟块上写入数据的写命令；

(d) 确定包含带有虚拟块偏移的所述虚拟块的一个虚拟单元；

(e) 寻找对应所述虚拟单元的一个物理单元；

(f) 在所述物理单元内寻找一个物理块；

(g) 判定所述物理块是否没有被写过；

(h) 仅当所述物理块没有被写过时，将所述数据写入所述物理块；

(i) 反之，如果所述物理块已经被写过，分配第二个物理单元；

(j) 在所述第二个物理单元中寻找一个可写入的物理块，所述可写入的物理块具有物理块偏移；

(k) 将所述数据写入所述可写入的物理块中；

(l) 另外将所述虚拟单元映射到包含所述可写入物理块的所述第二物理单元，这样所述虚拟单元被另外映射到所述第二个物理单元，以便形成一个物理单元链；

(m) 如果在写过的物理单元中找不到没有被写过的物理块，则寻找对应链中若干个物理单元的第二个虚拟单元；

(n) 寻找所述链中的最后一个物理单元；

(o) 将所述写过的物理单元的所述物理块中的所有数据传送到所述最后物理单元的所述物理块中；并且

(p) 更新所述虚拟映射图，使得所述虚拟单元只对应所述最后物理单元。

9. 权利要求 8 的方法，其中所述可写入物理块的所述物理块偏移具有与所述虚拟块偏移相同的块偏移号。

10.权利要求 8 的方法，进一步包括步骤：

(q) 擦除所有的所述写过物理单元。

11.一种存储器数据写入方法，其中数据只能被写入存储器的某个没有被写过的部分，这样，存储器的被写过部分就必须被擦除以便成为没有被写过的，该方法包括步骤：

(a) 提供若干个物理单元，而且每个物理单元又被分成若干个物理块，每个所述物理单元具有一个物理单元号并且每个所述物理块有一个在所述物理单元内的物理块偏移；

(b) 提供若干个虚拟单元，而每个虚拟单元又被分成若干个虚拟块，每个所述虚拟单元有一个虚拟单元号并且每个所述虚拟块也有一个在所述虚拟单元内的虚拟块偏移，每个虚拟单元被映射到至少一个物理单元上；

(c) 接收在某个虚拟块上写入数据的写命令；

(d) 确定包含带有虚拟块偏移的所述虚拟块的一个虚拟单元；

(e) 寻找对应该虚拟单元的一个物理单元；

(f) 在所述物理单元内寻找一个物理块；

(g) 判定所述物理块是否没有被写过；

(h) 仅当所述物理块没有被写过时，将所述数据写入所述物理块；

(i) 反之，如果所述物理块已经被写过，分配第二个物理单元；

(j) 在所述第二个物理单元中寻找一个可写入的物理块，所述可写入的物理块具有物理块偏移；

(k) 将所述数据写入所述可写入的物理块中；并且

(l) 另外将所述虚拟单元映射到包含所述可写入物理块的所述第二物理单元，以使所述虚拟单元被另外映射到所述第二个物理单元，以便形成一个物理单元链；

(m) 如果没有可供分配的被写过的物理单元，则寻找所述链中的最后一个物理单元；

(n) 将所述写过物理单元的所述物理块中的所有数据都传送到所述最后物理单元的所述物理块中；并且

(o)更新所述虚拟映射图以使所述虚拟单元只对应所述最后物理单元。

12.权利要求 11 的方法，进一步包括步骤：

(p)除了所述最后物理单元之外，擦除所有的所述写过物理单元。

13.一种存储器组织方法，在该存储器中数据只能被写入还没有被写过的部分，这样，该存储器的被写过部分必须被擦除以便成为没有被写过的，并且其中用于读或写数据的存储器部分的大小不同于用于擦除的最小存储器部分的大小，该方法包括步骤：

(a)提供该存储器的若干个物理单元，所述物理单元中的每个的大小分别等于用于擦除的最小存储器部分的大小，每个所述物理单元都由一个物理单元号指定，而且每个所述物理单元都被分解为若干物理块，每个所述物理块的大小分别等于用于读或写数据的存储器部分的大小，并且每个所述物理块都由所述物理单元中的一个物理块偏移指定；

(b)提供该存储器的若干个虚拟单元，每个虚拟单元都由一个虚拟单元号所指定，并且每个所述虚拟单元体现了若干个虚拟块，每个所述虚拟块由所述虚拟单元中的一个虚拟块偏移指定；

(c)提供一个虚拟映射图，用于将每个虚拟单元映射到至少一个物理单元；

(d)将所述虚拟单元中的每个虚拟块映射到所述至少一个物理单元中的一个物理块；

(e)接收在某个虚拟块上写入数据的写命令；

(f)寻找包含所述虚拟块的一个虚拟单元；

(g)在被映射到所述虚拟单元的物理单元中寻找一个可写入的块；

(h)将所述数据写入所述可写入物理块，形成一个被写过的物理单元；

(i)如果在所述被写过的物理单元中找不到一个没有被写过的物理块，则寻找具有可写入物理块的第二个物理单元；

(j)将所述数据写入所述第二个物理单元中的所述可写入物理块；

(k)通过另外将所述虚拟单元映射到所述第二个物理单元来更新所

述虚拟映射图，这样所述虚拟单元既对应所述第二个物理单元，又对应所述写过的物理单元，所述第二个物理单元和所述被写过的物理单元形成一个物理单元链；

(l) 如果在任何物理单元中找不到一个没有被写过的物理块，则寻找对应某个链中多个物理单元的第二个虚拟单元；

(m) 寻找所述链中的最后物理单元；

(n) 将数据从所述被写过的物理单元的每个所述物理块中移动到所述最后物理单元的一个可写入物理块中，所述可写入物理块事实上具有与所述被写过物理单元的所述物理块相同的块偏移；并且

(o) 通过将所述虚拟单元映射到所述最后物理单元来更新所述虚拟映射图，这样，所述虚拟单元只对应所述最后物理单元。

14. 一种存储器组织方法，在该存储器中数据只能被写入还没有被写过的部分，这样，该存储器的被写过部分必须被擦除以便成为没有被写过的，并且其中用于读或写数据的存储器部分的大小不同于用于擦除的最小存储器部分的大小，该方法包括步骤：

(a) 提供该存储器的若干个物理单元，每个所述物理单元的大小分别等于用于擦除的最小存储器部分的大小，每个所述物理单元都由一个物理单元号指定，而且每个所述物理单元都被分解为若干物理块，每个所述物理块的大小分别等于用于读或写数据的存储器部分的大小，并且每个所述物理块都由所述物理单元中的一个物理块偏移指定；

(b) 提供该存储器的若干个虚拟单元，每个虚拟单元都由一个虚拟单元号所指定，并且每个所述虚拟单元体现了若干个虚拟块，每个所述虚拟块由所述虚拟单元中的一个虚拟块偏移指定；

(c) 将每个虚拟单元映射到至少一个物理单元，以形成一个虚拟映射图；并且

(d) 根据所述虚拟映射图，将所述虚拟单元中的每个虚拟块映射到所述至少一个物理单元中的一个物理块，使得所述每个虚拟块和被映射所述每个虚拟块的所述物理块具有基本上相等的块偏移。

15. 一种存储器组织方法，在该存储器中数据只能被写入还没有被写

过的部分，这样，该存储器的被写过部分必须被擦除以便成为没有被写过的，并且其中用于读或写数据的存储器部分的大小不同于用于擦除的最小存储器部分的大小，该方法包括步骤：

(a) 提供该存储器的若干个物理单元，每个所述物理单元的大小分别等于用于擦除的最小存储器部分的大小，每个所述物理单元都由一个物理单元号指定，而且每个所述物理单元都被分解为若干物理块，每个所述物理块的大小分别等于用于读或写数据的存储器部分的大小，并且每个所述物理块都由所述物理单元中的一个物理块偏移指定；

(b) 提供该存储器的若干个虚拟单元，每个虚拟单元都由一个虚拟单元号所指定，并且每个所述虚拟单元体现了若干个虚拟块，每个所述虚拟块由所述虚拟单元中的一个虚拟块偏移指定；

(c) 将每个虚拟单元映射到最多两个物理单元，以形成一个虚拟映射图；以及

(d) 根据所述虚拟映射图，将所述虚拟单元中的每个虚拟块映射到所述至少一个物理单元中的一个物理块。

16. 一种存储器组织方法，在该存储器中数据只能被写入还没有被写过的部分，这样，该存储器的被写过部分必须被擦除以便成为没有被写过的，并且其中用于读或写数据的存储器部分的大小不同于用于擦除的最小存储器部分的大小，该方法包括步骤：

(a) 提供该存储器的若干个物理单元，每个所述物理单元的大小分别等于用于擦除的最小存储器部分的大小，每个所述物理单元都由一个物理单元号指定，而且每个所述物理单元都被分解为若干物理块，每个所述物理块的大小分别等于用于读或写数据的存储器部分的大小，并且每个所述物理块都由所述物理单元中的一个物理块偏移指定；

(b) 提供该存储器的若干个虚拟单元，每个虚拟单元都由一个虚拟单元号所指定，并且每个所述虚拟单元体现了若干个虚拟块，每个所述虚拟块由所述虚拟单元中的一个虚拟块偏移指定；

(c) 将每个虚拟单元映射到至少一个物理单元，以形成一个虚拟映射图；并且

(d) 根据所述虚拟映射图, 将所述虚拟单元中的每个虚拟块映射到所述至少一个物理单元中的一个物理块; 以及

(e) 通过改变该存储器的至少一个所述多个虚拟单元和该存储器的至少一个所述多个物理单元之间的对应关系而更改所述虚拟映射图, 以反映在存储器中的改变。

17. 一种存储器组织方法, 在该存储器中数据只能被写入还没有被写过的部分, 这样, 该存储器的被写过部分必须被擦除以便成为没有被写过的, 并且其中用于读或写数据的存储器部分的大小不同于用于擦除的最小存储器部分的大小, 该方法包括步骤:

(a) 提供该存储器的若干个物理单元, 每个所述物理单元的大小分别等于用于擦除的最小存储器部分的大小, 每个所述物理单元都由一个物理单元号指定, 而且每个所述物理单元都被分解为若干物理块, 每个所述物理块的大小分别等于用于读或写数据的存储器部分的大小, 并且每个所述物理块都由所述物理单元中的一个物理块偏移指定;

(b) 提供该存储器的若干个虚拟单元, 每个虚拟单元都由一个虚拟单元号所指定, 并且每个所述虚拟单元体现了若干个虚拟块, 每个所述虚拟块由所述虚拟单元中的一个虚拟块偏移指定;

(c) 将每个虚拟单元映射到至少一个物理单元, 以形成一个虚拟映射图; 并且

(d) 接收在某个虚拟块上写入数据的写命令;

(e) 寻找包含所述虚拟块的一个虚拟单元;

(f) 在被映射到所述虚拟单元的物理单元中寻找一个可写入的块;

(g) 将所述数据写入所述可写入物理块以形成一个被写过的物理单元;

(h) 如果在所述被写过的物理单元中找不到一个没有被写过的物理块, 则寻找第二个物理单元;

(i) 将所述数据写入所述第二个物理单元中的一个可写入物理块; 并且

(j) 把所述虚拟单元映射到所述第二物理单元, 以及映射到所述被

写过的物理单元，所述第二物理单元和所述被写过的物理单元形成一个物理单元链。

改进的闪速文件系统

本发明涉及一种处理闪速设备上数据存储的系统，具体来说，涉及一种管理分页方式闪速设备上信息存储和检索的系统，使其性能犹如闪速磁盘。

闪速设备包括由闪速类型的浮栅晶体管构成的电可擦除和可编程只读存储器（EEPROM），这是一种非易失性存储器，在功能和性能上类似EPROM存储器，只不过增加了一种功能，允许线路内可编的操作擦除存储器的页。与传统的磁盘比较，闪速设备的优点是价格相当便宜，而且只需要相当小的功率。但是在闪速设备中，如果不对存储器的某个已经被写入的区域预先进行页面擦除，就不能对该区域重新写入。闪速设备的这种局限性使之与典型的现有操作系统程序不兼容，因为数据不能再被写入到闪速设备中已经写入数据的那些存储器区域上，除非预先对该区域进行擦除。

当前的技术已经提供了这样的软件产品，允许现有操作系统程序在不作修改的情况下对闪速设备进行管理。然而，这些当前技术的程序都具有缺陷。例如，一种程序将闪速存储器作为“写一次读多次”的设备来操作。这种现有技术的软件产品不能反复使用已经被写入的存储器位置。当所有的位置都最终被写入时，该存储器就不能再进一步使用，如果用户不进行特殊处理的话。其他当前技术的程序，例如SanDisk公司提供的程序，每当新数据将要被写入某个页时，该程序能擦除并且重写整个存储器的页。该系统的缺点是要求多个擦除周期，这是相当缓慢和低效的而且将导致物理介质本身更快的退化。

为了克服现有技术的这些缺陷，美国5404485号专利提供了一种闪速文件系统（FFS），这里将加以引用。FFS提供了一种闪速设备上数据存储和操纵的系统，该系统允许这些设备仿真磁的基于盘的数据存储器。如上所述，闪速设备相当便宜的费用和低功率消耗使其成为数

据存储器的一种令人喜爱的选择，尤其是对于膝上型的便携式计算机。FFS提高了闪存设备代替磁盘存储器的能力。实际上，美国5404485号专利提供的FFS已经被证明是如此有用，以至其数据设计规格已经被PCMCIA（个人计算机存储器卡国际协会）和JEIDA（日本电子工业开发协会）采用作为一种标准，称之为闪存转换层（FTL）。

FFS主要描述了闪存EEPROM设备的虚拟映射系统。该虚拟映射图是一个表，将闪存设备中读/写块的物理地址和该块的虚拟地址联系起来。由于每个这样的块都很小，为512字节，而该虚拟映射图本身的规模很大。FFS还包括在闪存EEPROM上存储和维护大量虚拟映射图的方法，最大程度地减少存储虚拟映射图所需的其他存储量。

如上所述，FFS对于将闪存设备转换为磁盘存储的仿真器已经被证明是相当成功的，由于如此成功使之已经被采用作为一种工业标准。然而，FFS并不能满足更新型闪存设备技术的所有需求。具体来说，FFS对NAND和AND闪存技术并不成功。

企图克服当前技术闪存存储器结构的某些缺陷、具体来说就是写之前先擦除的系统的另一个例子是由美国5479638号专利提供的。在该专利的系统中，如果需要对某个被写块进一步编程，则某个特定读/写块的物理位置被移位。但是，该系统的缺点是只能对每次能够擦除单个512字节读/写块的闪存设备进行操作。因为这样的一个要求是在硬件级上实现的，因此，该系统也不能用于较新的NAND和AND闪存技术。

NAND和AND在以下几个方面不同于较老的闪存设备技术。首先，与较老闪存设备的64KB相比，NAND和AND的可擦除单位规模较小，约为8KB。其次，对于NAND和AND来说，擦除时间要快得多，即使按擦除单个字节所要求的时间来计算也如此。第三，闪存存储器被分成NAND和AND的若干页，长度为256或512字节，这是硬件设备本身的一个固定特性。应该注意到，这里所用到的术语“页”大概相当于较老闪存技术中用到的术语“块”，虽然“页”和“块”的具体特性有些不同。这些特征对基于NAND和AND技术的闪存设备的操作

具有若干含义。

首先，分页方式存储器在写一页或其中的任何部分时具有一个固定的开销。而用于当前闪速技术中写操作的开销是与被写入的字节数成比例的。其次，NAND和AND中的闪速存储器被配置使得每一页都有几个可特殊寻址的备用字节。这些备用字节是有关闪速存储器系统信息存储的方便位置。最后，在某个页被擦除之前，对该页可以被写的次数有一个限定。这个限定是相当低的，8次或10次，在此之后，如果没有预先擦除，进一步写入是不可靠的。因此，分页方式存储器对于成功的数据存储和检索具有突出的优点和新的挑战。

不幸的是，如上所述，当前可用的现有技术数据处理系统FFS对于采用分页方式的闪速存储器的操作具有明显的缺点。具体来说，由于分页方式编程所强加的限制，使得FFS对例如NAND和AND这样的分页方式闪速技术表现了非优化的性能。另外，美国5479638号专利中提供的系统也不能用于这样的闪速技术，因为它要求逐块的擦除操作。

因此，需要提供一种处理NAND和AND闪速设备上数据存储的系统，它将具有很大的优势，该系统不但能优化分页方式闪速技术的性能，而且也能用于较老的、非分页方式的设备。

本发明提供了一种存储器组织方法，在该存储器中，数据只能被写入还没有被写过的部分，这样，该存储器的被写过部分必须被擦除以便成为没有被写过的，并且其中，用于读或写数据的存储器部分的规模，不同于擦除的最小存储器部分的规模，该方法包括步骤：（a）提供该存储器的若干个物理单元，每个物理单元都是擦除的最小存储器部分，每个物理单元都由一个物理单元号指定，而且每个物理单元都被分解为若干物理块，每个物理块都是用于读或写数据的存储器部分，并且每个物理块都由物理单元中的一个物理块偏移指定；（b）提供该存储器的若干个虚拟单元，每个虚拟单元都由一个虚拟单元号指定，并且每个虚拟单元体现了若干个虚拟块，每个虚拟块由虚拟单元中的一个虚拟块偏移指定；（c）提供一个虚拟映射图，将每个虚拟单元至少映射到一个物理单元；并且（d）将虚拟单元中的每个虚拟块映

射到至少一个物理单元中的一个物理块。

该方法最好还包括步骤：(e)接收在虚拟块上写入数据的写命令；(f)寻找包含该虚拟块的虚拟单元；(g)在被映射到该虚拟单元的物理单元中寻找一个可写入的块；并且(h)将数据写入该可写入物理块。该方法最好还包括步骤：(i)如果在一个被写过的物理单元中找不到一个没有被写过的物理块，则寻找一个没有被写过的物理单元；(j)将数据写入该没有被写过的物理单元中的一个可写入物理块；并且(k)通过另外将该虚拟单元映射到该没有被写过的物理单元来更新该虚拟映射图，这样，该虚拟单元既对应没有被写过的物理单元，又对应写过的物理单元，没有被写过的物理单元和被写过的物理单元形成一个物理单元链。没有写过的物理单元的可写入物理块最好有一个物理块偏移，并且该物理块偏移对应被映射的虚拟单元的虚拟块偏移。该方法最好还进一步包括步骤：(l)如果在某个写过的物理单元中找不到一个没有被写过的物理块，则寻找对应某个链中多个物理单元的第二个虚拟单元；(m)寻找该链中的最后一个物理单元；(n)将数据从被写过的物理单元的每个物理块中移动到最后一物理单元的一个可写入物理块中，可写入物理块确实具有与被写过物理单元的物理块相同的块偏移；并且(o)通过将该虚拟单元映射到最后一物理单元来更新该虚拟映射图，这样，该虚拟单元实际上只对应最后的物理单元。该方法最好还进一步包括步骤：(p)真正擦除该链中所有被写过的物理单元，只留下最后的物理单元。

另外，该方法还可以进一步包括步骤：(1)如果在该链上找不到可分配的没有写过的物理单元，则分配一个没有被写过的物理单元以便重新组织；(m)将数据从被写过物理单元的每个物理块中移动到没有被写过的物理单元的一个可写入物理块中，并且(n)通过将虚拟单元映射到没有被写过的物理单元以更新该虚拟映射图，这样，该虚拟单元实际上只对应没有被写过的物理单元。该方法最好还进一步包括步骤：(o)擦除所有的被写过物理单元。

根据本发明的其他最佳实施例，该方法进一步包括步骤：(i)如

果在写过的物理单元中找不到没有被写过的物理块，则分配一个没有被写过的物理单元来形成一条物理单元链，这样，这个没有被写过的物理单元就是该链中的最后一个物理单元；（j）将数据写入最后物理单元中的一个没有被写过的物理块；（k）除了在步骤（j）中被写入的数据将数据从每个物理块移到没有被写过的物理单元的一个可写入物理块中；并且（m）更新该虚拟映射图，即将虚拟单元映射到写过的物理单元，这样虚拟单元就对应写过的物理单元。

根据本发明的另一个实施例，提供了为某个存储器写入数据的一种方法，在这种方法中，数据只能写入存储器的没有被写过的部分，这样，存储器的被写过部分必须被擦除以便成为没有被写过的，该方法包括步骤：（a）提供若干个物理单元而且每个物理单元又被分成若干个物理块，每个物理单元具有一个物理单元号并且每个物理块也有一个在其物理单元内的物理块偏移；（b）提供若干个虚拟单元而且每个虚拟单元又被分成若干个虚拟块，每个虚拟单元有一个虚拟单元号并且每个虚拟块也有一个在其虚拟单元内的虚拟块偏移，每个虚拟单元被映射到至少一个物理单元上；（c）接收在虚拟块上写入数据的写命令；（d）确定包含带有虚拟块偏移的虚拟块的一个虚拟单元；（e）寻找对应该虚拟单元的一个物理单元；（f）在该物理单元内寻找一个物理块；（g）判定该物理块是否没有被写过；（h）仅当该物理块没有被写过时，将数据写入该物理块；（i）反之，如果该物理块已经被写过，分配一个没有被写过的物理单元；（j）在该没有被写过的物理单元中寻找一个可写入的物理块，可写入的物理块具有物理块偏移；（k）将数据写入该可写入物理块中；并且（l）另外将该虚拟单元映射到包含该可写入物理块没有被写过的物理单元，这样，该虚拟单元被另外映射到没有被写过物理单元，以便形成一个物理单元链。可写入物理块的物理块偏移最好具有与该虚拟块偏移相同的块偏移号。

该方法最好进一步包括步骤：（m）如果在写过的物理单元中找不到没有被写过的物理块，则寻找对应链中若干个物理单元的第二个虚拟单元；（n）寻找该链的最后一个物理单元；（o）将写过的物理

单元的物理块中的所有数据传送到最后物理单元的物理块中；并且(p)更新虚拟映射图，使得该虚拟单元只对应最后的物理单元。该方法最好进一步包括步骤：(q)擦除所有的写过物理单元。

另外，该方法还可以进一步包括步骤：(l)如果找不到没有被写过的物理单元则寻找该链中的最后一个物理单元；(m)将写过的物理单元的物理块中的所有数据都传送到最后物理单元的物理块中；并且(n)更新虚拟映射图，使得该虚拟单元只对应最后的物理单元。该方法最好进一步包括步骤：(o)实际上除了最后的物理单元之外擦除所有的写过的物理单元。

本发明提供了组织某种闪速存储器的方法，其中，用于读或写数据的存储器部分的规模，例如一个块，不同于擦除的最小部分的规模，例如一个单元。闪速存储器类型的例子可以根据本发明的方法来组织，包括但不限于由NAND和AND技术示例的分页方式设备。并且提供了对闪速存储器读写数据的方法，以及当找不到更多的没有被写过的物理单元时，重新组织该闪速存储器的方法。

此后，术语“物理单元”被定义为物理介质或存储器的硬件上的一个单元，是可以被擦除的存储器的最小部分，或者是其整数倍。这是存储器中相连的、规模固定的并且是可擦除的一个部分。术语“物理块”被定义为用于读或写数据的存储器的部分。此后，术语“虚拟单元”被定义为与物理单元具有相同的规模。对于例如NAND和AND这样的分页存储技术，可擦除的存储器的最小部分大于其页的规模，通常约为8KB。正如这里所用到的，术语“物理块”等于分页方式的存储技术中的术语“页”。因此，虚拟单元和物理单元一样大。

此后，术语“虚拟映射图”指的是一张表，该表将一个虚拟单元与至少一个对应的物理单元联系起来。如上所述，每个单元，虚拟的或物理的，都由若干个块组成。单元内块的准确位置可根据一个或多个预置的规则来确定，下面进一步描述。

每个物理单元由一个物理单元号指定。每个物理块的位置由物理块偏移给出。同样，每个虚拟单元由一个虚拟单元号指定。每个虚拟

块位置由虚拟块偏移给出。应该注意到，每个虚拟单元号可以对应一个或多个物理单元号。因此，虚拟单元和物理单元之间的映射可以是一对一，也可以是一对多。

此后，术语“写数据”描述在闪速存储器上存储数据的动作。术语“读数据”说明从闪速存储器中检索数据的动作。此后，术语“没有被写过”表示可以将数据写入的存储器的某个部分，例如一个物理块。因此，术语“没有被写过”包括但不局限于刚被擦除的存储器的某个部分。

在具有根据本发明进行组织的闪速存储器的计算机或其他电子设备中，该设备的操作系统与虚拟单元和虚拟块交互作用以便读写数据。包括虚拟单元和虚拟块的虚拟介质作为操作系统与闪速存储设备交互作用的一个界面。例如，操作系统发出一个写命令，要求将数据写入具有某个虚拟块偏移的一个虚拟块。接着寻找包含该虚拟块的虚拟单元。虚拟映射图寻找在存储器某个物理单元内对应的物理块，数据被实际存储在这里。虽然操作系统发出读写命令，似乎虚拟单元和虚拟块是闪速存储器的实际硬件，事实上，真正的硬件是闪速存储器的物理单元和物理块。因此操作系统仅仅知道虚拟单元和虚拟块，而不是直接和硬件本身交互作用。

这样一个界面的优点是闪速存储器的固有缺点，例如在可以进一步写入之前需要进行一次擦除，已经通过操作系统和虚拟存储器的交互作用来加以克服。此外，电子设备的操作系统并不需要组织闪速存储器的地址。另外，操作系统可以和多种不同的闪速存储器技术交互作用而不需要更多的修改，因为一个界面可以用于多种类型的闪速存储设备。因此，本发明的方法允许对闪速存储设备和使用它们的电子设备具有最大的灵活性。

这里仅仅通过示例的方式，参考以下的附图来描述本发明，其中：
图1是根据本发明的物理闪速存储设备的一个示意图；
图2是根据本发明组织闪速存储设备的基本系统的示意图；
图3A和3B分别表示根据本发明的AND和FMAX系统；

图4A和4B分别表示根据本发明的AND和FMAX系统的写算法;以及

图5A和5B分别表示根据本发明的AND和FMAX系统的重新组织算法。

本发明提供了一种用于组织分页方式闪速存储设备的系统,例如NAND和AND闪速设备。该系统包括对这样的闪速设备读写的方法。此外,该系统还提供一个界面使得闪速设备能够成功地仿真磁盘存储器。这样的闪速存储器设备可以安装在各种主设备中,例如个人计算机或膝上计算机。

虽然本发明将在闪速存储器的环境中描述,但熟悉这一技术的人应该理解,其原理同样适用于具有与闪速存储器相同的写、读和单元擦除特性的数据存储设备。

结合附图和相关的描述,可以更好地理解根据本发明用于组织分页式闪速存储设备的系统的原理和操作。

现在参看附图,图1根据例如NAND这样的技术,示意性地说明当前技术的物理页式闪速存储设备。术语“物理设备”此后被定义为实际的硬件本身,包括闪速存储设备的物理介质。这样的物理介质通常由闪速EEPROM器件组成,虽然任何合适的非易失性、可编程的存储设备都可以代替。术语“可编程”此后被定义为可例如通过将数据写入存储设备修改的。

闪速存储器物理设备10被表示为至少带有一个物理单元12。物理单元12是物理设备10可以被擦除的最小段。物理单元12包括整数个块,独立地被表示为块1到块n,其中n为整数,并且集中地被表示为块14。块14是由一组相连的、固定长度的物理字节地址组成,并且是其硬件的一个特征。具体来说,块14的规模是物理设备10的一个特性。用户数据可以被存放在块14的主数据区域16中。每一块14也有一个控制数据区18。控制数据区18可由来自块14主部分的独立算法实现寻址,并且不被包含在块14规模的计算中。正如以下将进一步描述的,控制数据区18便于存储与闪速文件系统本身有关的信息。每个物理单元12可

以是被分配的单元，也可以是一个非分配的单元。每个非分配单元是自由的，不包括有用的数据，并且随时准备被定位和分配。每个分配的单元已经被定位并且包含数据。

图2说明一种用于组织基本闪速存储设备的系统。系统20控制虚拟介质22和物理设备10，并且通过虚拟映射图24将虚拟介质22和物理设备10联系起来。虚拟介质22包含若干个虚拟单元26。每个虚拟单元26具有多个虚拟块28。每个虚拟单元26由虚拟地址指定。虚拟地址包括虚拟单元号，指定某个特定的虚拟单元26。每个虚拟块28由虚拟块偏移指定。同样，每个物理单元12都有一个物理地址。物理地址包括物理单元号，指定某个特定的物理单元12。每个物理块14有一个物理块偏移。

虚拟映射图24将一个被分配的虚拟单元26，映射至至少一个被分配的物理单元12。对于物理单元，如果某个虚拟单元已经被映射到至少一个物理单元，则这是一个被分配的虚拟单元。然而，一个虚拟单元26可以被映射到多个物理单元12。因此，虚拟单元26和物理单元12之间的对应关系可以是一对一，也可以是一对多。

系统20操作如下。带有闪速存储设备（没有表示）的电子设备的操作系统对某个特定的虚拟单元26中的某个特定的虚拟块28发出命令，例如读命令或写命令。虚拟映射图24随后寻找某个物理单元12中对应的物理块14。

如果该映射的对应关系是如图所示的一对一关系，每个虚拟单元26被映射到一个物理单元12，这种情况似乎显得相当简单。然而，如上所述，物理设备10对读写数据具有特定的物理限制，包括在另一个数据可以被写入之前执行周期性擦除的要求。对于这些限制，有两个可能的解决方案，不包括对物理设备10或其中的部分进行经常重复的擦除。

图3A给出第一种解决方案，其中虚拟单元26和物理单元12之间的对应关系是一对多关系，因此，每个虚拟单元26对应多个物理单元12。虚拟映射图必须具有执行这样一种映射所需的信息。图3A给出虚拟映

射图24部分的一个例子，并且可用于本发明的AND系统。

虚拟图24的这个部分表示物理单元12中的物理块14，以及虚拟单元26中的虚拟块28。在该例子中，一个特定的虚拟单元30对应两个物理单元12。第一个物理单元12是主单元32。第二个物理单元12是替换单元34。每个虚拟单元26只能有一个主单元34。但是，与每个虚拟单元26相关的可能没有、也可能有多个替换单元34。作为一个例子，虚拟单元36只对应一个主单元38而没有替换单元34，因此，虚拟单元36是非替换虚拟单元的一个例子。

虚拟块28的组织将取决于对应某个特定虚拟单元26的物理块14的数目。对于虚拟单元30，某些虚拟块28对应主单元32中的物理块14，而其他虚拟块28则对应替换单元34中的物理块14。对于虚拟单元36，实际上所有的虚拟块28都对应主单元38中的物理块14。

在这种最简单的情况下，其中的虚拟单元是一个非替换单元，定位某个特定物理块14的过程如下。虚拟单元36带有指定该单元的虚拟单元号44，以及指定虚拟块42的虚拟块偏移46。注意，虚拟块偏移46也是一个号。物理单元号50指定主单元38。物理块偏移52指定主单元38中的物理块54。为了实现读或写数据操作而定位物理块54，第一个规则就是将所需虚拟块的偏移46除以每个虚拟单元中的块数，以便确定虚拟单元号44。虚拟映射图24随后将虚拟单元号44映射到物理单元号50。第二个规则是根据物理块偏移52必须与虚拟块偏移46具有相同的号，从而定位物理单元38中所需的物理块14，在这种情况下是物理块54。因此，虚拟映射图24只包含有关虚拟和物理单元的信息，但规则被用来确定合适的块偏移。

在更复杂的情况下，每个虚拟单元对应多个物理单元。在这种情况下，两个或多个物理单元的组被称为一个“链”。例如，虚拟单元号72指定虚拟单元30，而虚拟块偏移74指定虚拟块70。物理单元号78指定替换单元34并且物理块偏移80指定替换单元34中的物理块82。因此，虚拟单元30的虚拟块70对应替换单元34的物理块82。

为了实现读或写数据操作而定位物理块82，第一个规则也是将所

需虚拟块的偏移74除以每个虚拟单元中的块数，以便确定虚拟单元号72。虚拟映射图24随后将虚拟单元号72映射到物理单元号78。但是有一个问题。如上所述，第二个规则是根据物理块偏移（必须与虚拟块偏移具有相同的号），定位物理单元中所需的物理块。在这种情况下，该链中有若干个物理块14。为了确定哪个物理块14有数据，第三个规则是检查该链的每个物理单元中具有与虚拟块70相同的块偏移的每个物理块14。最后的非自由物理块14，在这种情况下是替换单元34的物理块82，包括供读出的所需数据。反之，对于写数据，第一个自由的物理块14是所需的块。

由于物理块是按照它们在所属的链中的物理单元的次序被写入的，术语“最后的非自由物理块”指的是在该链中最远的、但仍然是非自由的单元中的物理块。或者在该链中没有更多的单元，或者具有与该链中的下一个单元相同的块偏移的物理块是自由的。同样，为了找到第一个自由的物理块，对该链中的每个物理单元中具有所需块偏移的每个物理块进行检查，从主单元开始，并且依次通过每个替换单元继续下去，直到发现一个自由块。

反之，FMAX虽然使用相同的虚拟映射图和寻址系统，但每个主单元只有一个替换单元，如图3B所示。为此，FMAX使用简单和组合替换（物理）单元。简单替换单元是这样单元，其中物理单元的所有物理块偏移实际上都与对应虚拟单元的虚拟块偏移直接相关。组合替换单元是这样的一个单元，其中并不需要存在虚拟块偏移和物理块偏移之间的这样一种直接对应关系。反之，如果具有对应物理块偏移的物理块不能用于写入，则选择另一个物理块。随后将控制信息写入控制数据区以便确定虚拟块和物理块之间的实际对应关系。

如图3B所示，主单元97有一个简单替换单元98，该单元具有若干个物理块100，每个块对应虚拟单元104中的一个虚拟块102。每个物理块偏移对应一个虚拟块偏移，该偏移是相同的偏移号。

然而，如果得不到具有所需物理块偏移的物理块，则在相同物理单元中的另一个物理块必须被写入，并且该替换单元成为一个组合替

换单元。第二个主单元109有一个组合的物理单元110，该单元也有若干个物理块112，其中的每个对应虚拟单元116中的一个虚拟块114。然而，一个物理块偏移可以对应一个相同的偏移号的虚拟块偏移，而第二个物理块偏移可以对应第二个不是相同的偏移号的虚拟块偏移。为了找到一个特定的物理块，必须检查被写入控制数据区的控制信息。正如下面将进一步描述的，这对于写数据以及必要时重新组织FMAX系统都具有重要的意义。

图4A给出了处理图3A中虚拟映射图的流程图，而图4B则给出处理图3B中虚拟映射图的流程图。在这种最简单的情况下，所有的替换单元或者是简单单元，或者是只有一个替换单元的主单元，相同的步骤可以被用于AND和FMAX。首先，计算虚拟单元号和虚拟块偏移，即把将被定位的虚拟块的数目除以每个虚拟单元中的块数从而给出虚拟单元号。而模数，或者除法的余数就是虚拟块偏移。

接着，检查虚拟映射图以便找到对应该虚拟单元的物理单元。如果找不到对应该虚拟单元的物理单元，则物理存储器所要的部分并不存在该闪速设备上。如上所述，这种简单的模式仅在所有的替换单元都是简单单元、或者主单元只有一个替换单元时起作用。然而，如果将要写入数据的物理块已经被其他数据编程或者写入，该模式将不起作用。在这种情况下，需要一种替换模式，该模式能够找到可以写入数据的另一个物理块。

图4A（ANAND）和图4B（FMAX）说明两个不同的算法。这两个算法都以相同的方式开始。在步1中，所需的物理单元被定位。在步2中对应所标识的块偏移的物理块在该物理单元中被定位。在步3中，如果该块是没有被写过的，则数据被写入该块中。如果找不到所要的物理块，则本发明的这两个系统，AND和FMAX中，每个技术处理所要的物理块已经被写入的情况的方式就不同。

如图4A所示，AND系统将通过检查替换单元处理这种情况。在步4中，检查第x个替换物理单元，这里x是初始值为1的一个整数。如果该物理单元有一个具有所要的物理块偏移的没有被写过的物理块，则

数据被写入该物理块。如果找不到这样的块，则如步5所示， x 增1并且重复步4。步4和步5被重复直到该数据或者被写入某个块中，或者在该链中找不到其他的替换单元。在步6中，一个没有分配的物理单元被分配作为一个替换单元，并且该数据被写入到具有所要块偏移的块中。

FMAX系统采用不同的方式处理这种情况，如图4B所示。在步4中，替换单元中具有相同物理块偏移的一个物理块被定位。如果该物理块是没有写过的，则数据被写入该物理块中。否则，如步5所示，该替换单元中具有不同物理块偏移的一个物理块被定位。步5被重复，直到一个没有写过的物理块被找到。该替换单元现在是一个组合单元，因为虚拟块偏移不再与物理块偏移相同。在步6中，控制信息被加到控制数据区中，使得该映射模式能够发现组合单元中任何物理块的正确位置。

然而，即使这些替换算法也不能满足对闪速设备所有的不同需要的处理。AND和FMAX系统最终都将遇到这样的一种情况，进一步的数据不能被写入物理单元的一个块中，因为不存在这样的物理块。

在这种情况下，虚拟单元必须被重新组织以便重新把数据构造为其最简单的状态，这是一个非替换的主单元。在这个重新组织的过程中，以前属于该虚拟单元表示的物理替换单元被释放，因此成为没有分配的或自由的物理单元。对于AND替换单元以及简单FMAX替换单元，这个重新组织过程被称为折叠，如以下图5A所示。

折叠要求物理块按照与本来应该被写入主单元相同的替换单元中的相同物理块偏移被写入，其理由将随着该过程的描述而更加清楚。在折叠的第一步，该链中最后一个物理单元物理单元 x ，被识别，其中， x 是一个整数，取值范围从1到某个预定的取决于执行情况的限定值。注意，其中的 x 等于1的替换单元实际上是主单元，并且该算法的其余部分没有被执行。还要注意，对于FMAX， x 等于1或2。

在步2中，单元 x 的块 n 被检查，其中 n 是个整数。如果数据被写入块 n ， n 增1。否则，则在步3中 x 减1。步2和步3被重复，直到或者 x 等于0，或者找到一个写过的块 n 。如果找到一个写过的块 n ，则在步4中将数据移动到该链中最后一个替换单元的块 n 中。步2到步4被重复，直到

所有的数据已经被传送到最后的替换单元，该单元随后成为主单元。该链中的所有其他单元，如果有的话，包括以前的主单元，随后都被释放并且可以用于分配。该虚拟映射图也被更新以便反映该虚拟单元现在对应一个物理单元的事实。

不幸的是，折叠并不适用于组合FMAX替换单元，因为该替换单元中的块并不是总有一个等于该虚拟块偏移的物理块偏移。图5B中给出了对这样组合物理单元的一个不同的重新分配过程。在步1中，一个新的、没有被分配的物理单元被指定作为新的主物理单元。在步2中，检查组合物理单元的块 n 。如果数据被写入组合物理单元的块 n ，则在步3中将数据拷贝到新的主单元。否则，原来的主单元的块 n 中的数据被写入到新的主单元。在步4中对 n 增1。重复步2到步4，直到所有的块已经被拷贝。一旦所有的块已经被拷贝，以前的替换单元，以及原来的主单元，都被释放并且可以分配。象在以前的过程中一样，虚拟映射图被更新以便反映该虚拟单元现在只对应一个物理单元这样的事实。

一个大大简化的重新组织的实施例也是可能的。在这个简化的实施例中，替换单元被分配后，立即进入该重新组织的过程。因此，替换单元只是系统的一个过渡特征，并且处于静止状态，其中物理存储器没有经受写入处理，数据全部都存在非替换的主单元中。替换单元的存在仅仅是为了写入过程的处理。在该过程结束时所有的信息都被传送到一个新的单元，因此该替换单元消失了。这种方法的优点是实施起来简单，而且管理这个过程所需的控制结构也简单。但是，其缺点是效率不高，降低了该系统的写入性能。

本发明的所有方法都必须能够将控制信息记录在物理闪速设备本身上，以便描述被存储数据的状态。具体来说，单元和块的控制信息最好被保存起来，虽然这些数据也可以另外从其他类型的数据中重新构造出来。单元控制信息描述了被分配给该物理单元的物理单元号、该物理单元本身作为主单元或替换单元的状态、以及它与其他单元相对的位置。块控制信息说明该物理块是否被占用、是自由的或被驻留

在某个不同物理块中的信息所取代。

这些不同类型信息的一种或两种可以被记录在该物理设备的某个特定部分上。如上述图1中所提到的，AND和FMAX系统最好将每个物理单元12分成存放记录在物理闪速设备上的实际用户数据的主数据区16以及存放控制信息的控制数据区18。虽然这些区域被表示为块16的子分区，但物理单元12也可以被分成主数据区和控制数据区，实际上与分块无关。应该注意，控制数据区18没有被包含在主数据区16的块定位模式中，并且当计算物理闪速盘的全部规模时也没有被包含。

由于NAND和AND闪速技术对每个存储块都有备用区，因此控制信息一般都记录在该块的备用区中，而用户数据则是位于主块区中。

对于那些没有提供备用区的闪速技术，每个物理单元可以被分成用于存放用户数据的主区以及用于存放所需控制信息的一个附加区。

应该理解，上面的描述仅仅是用作示例，而在本发明的实质和范围内还可以有许多其他的实施例。

图1

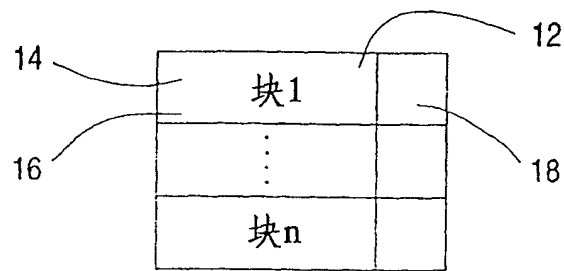


图2

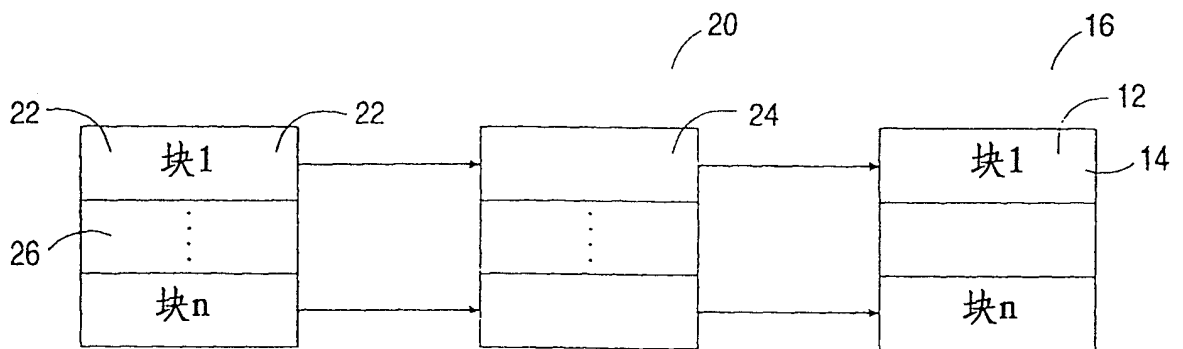


图 3A

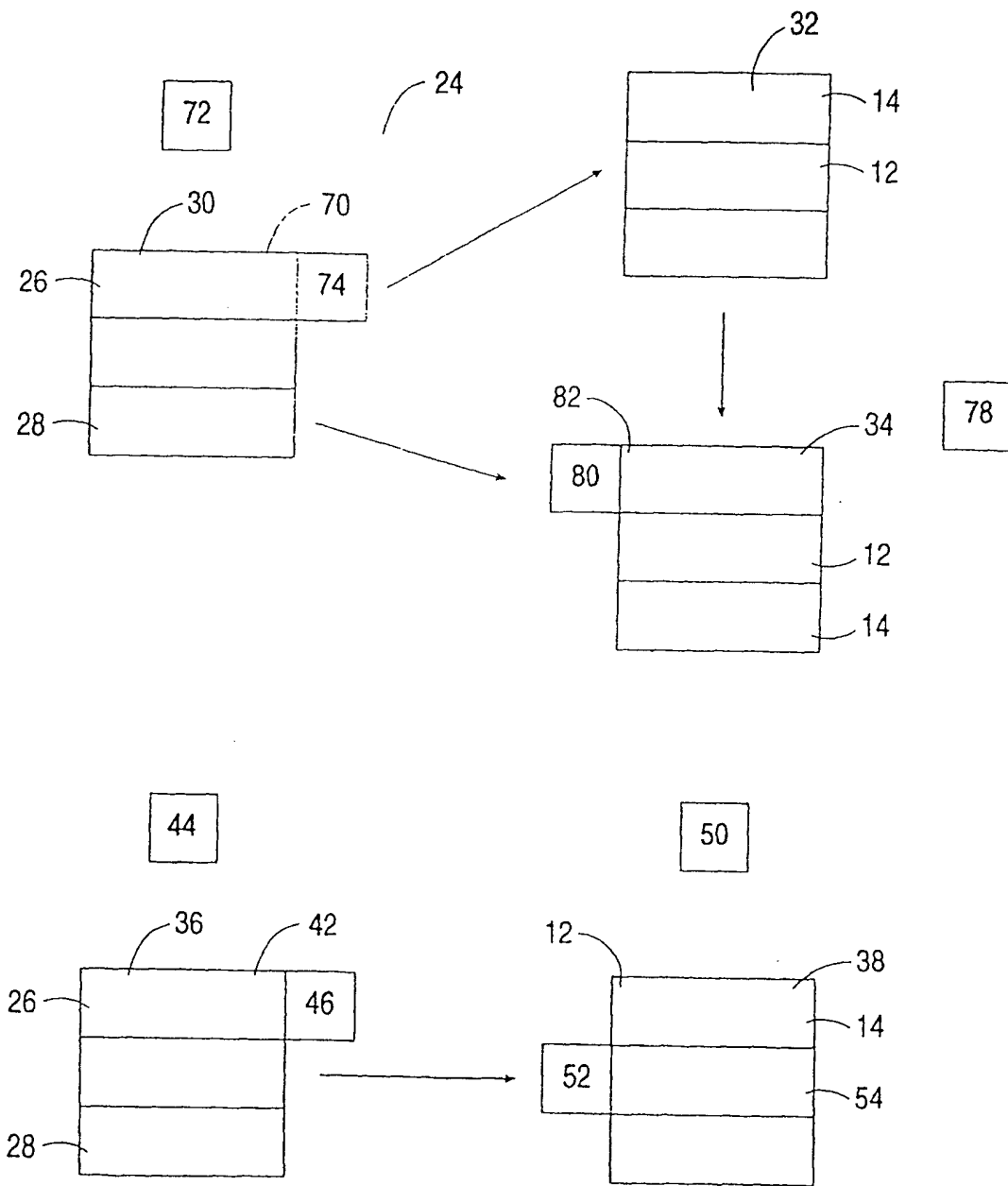


图 3B

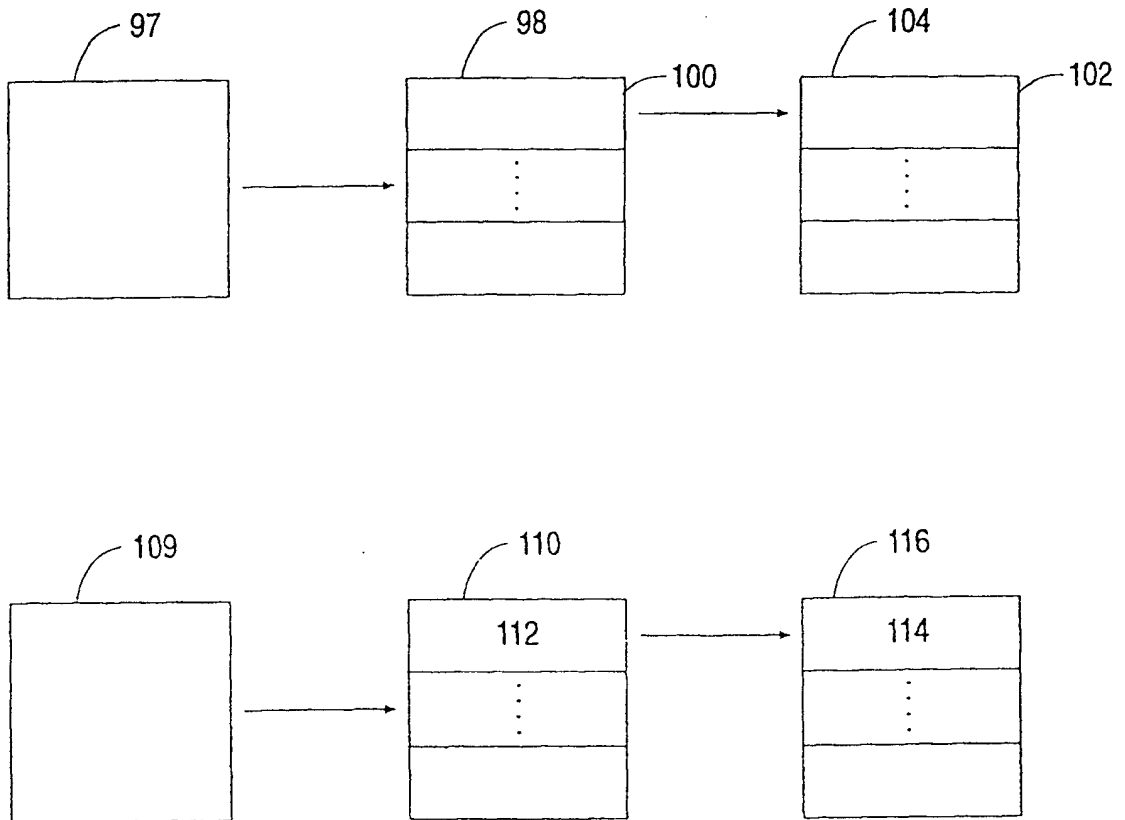


图 4A

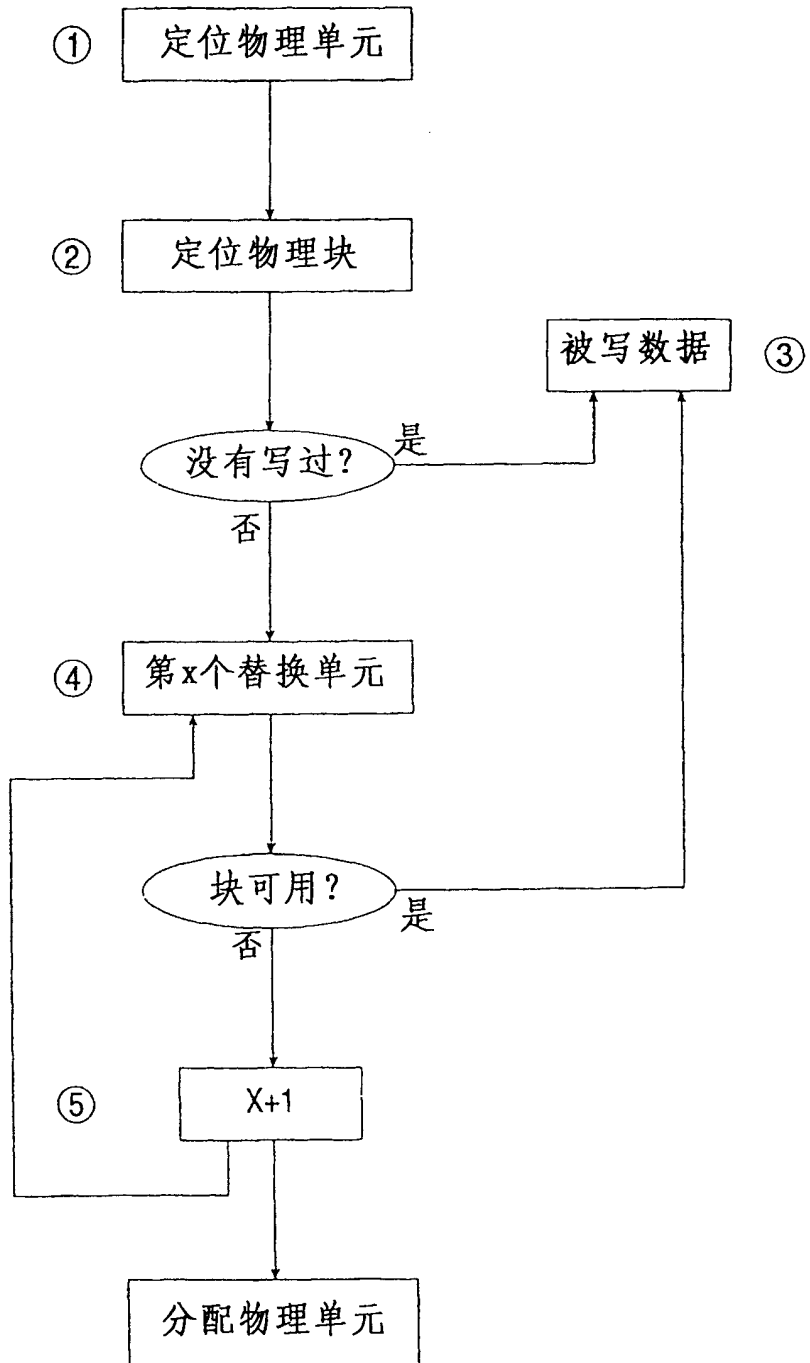


图4B

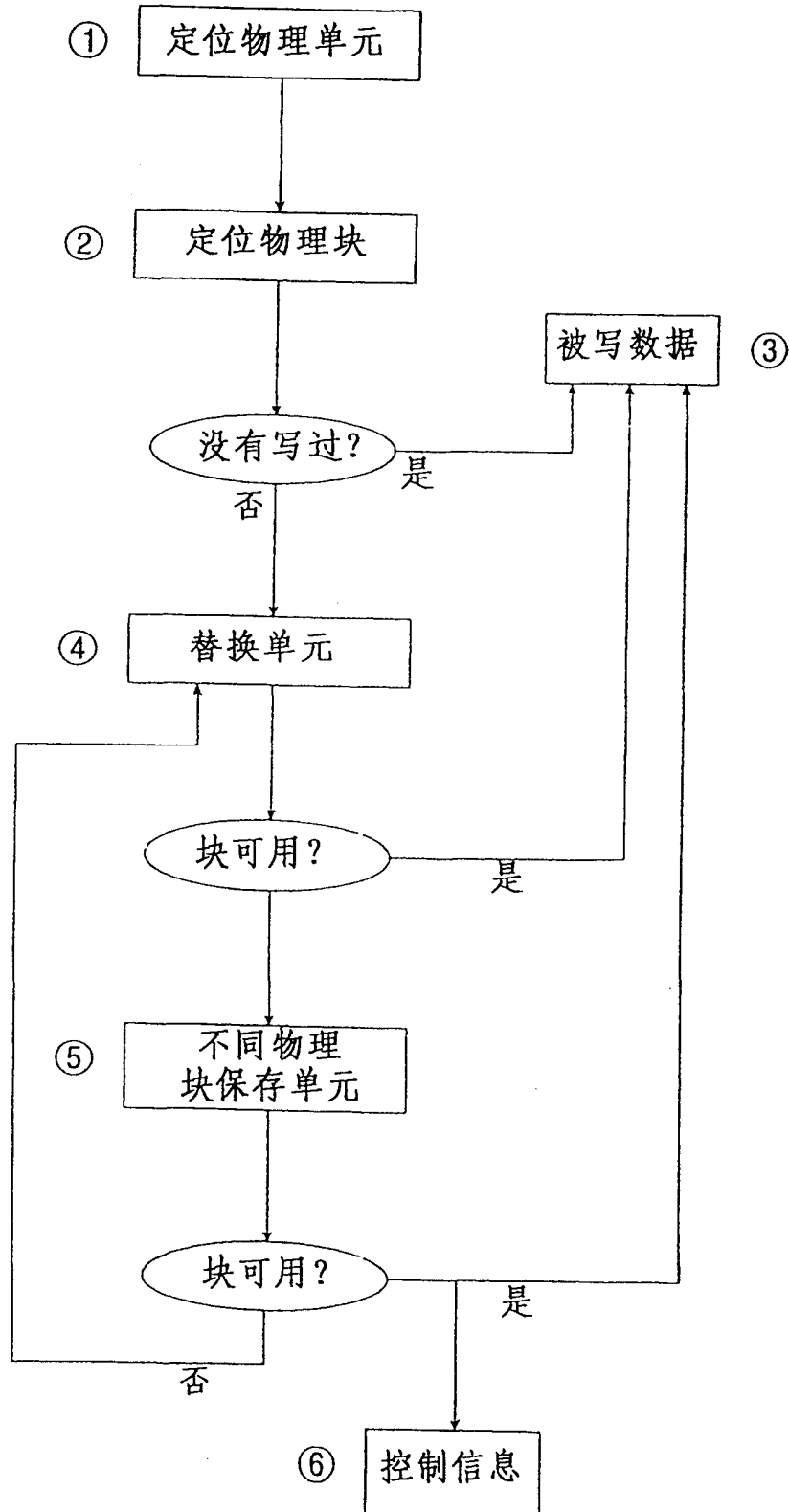


图 5A

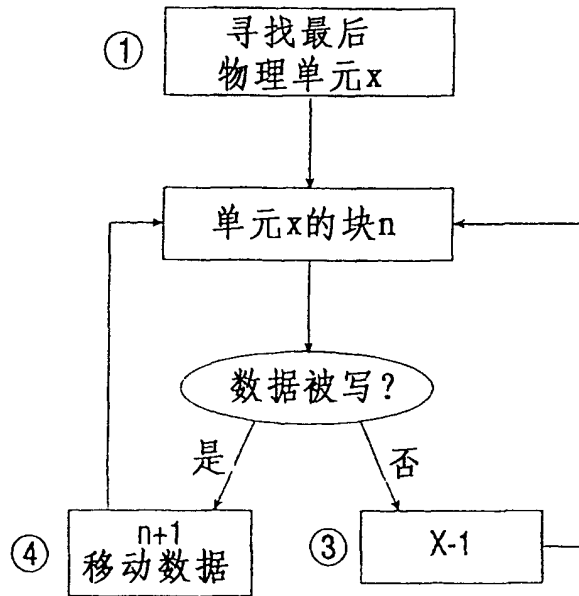


图 5B

