[19] 中华人民共和国国家知识产权局

[51] Int. Cl.
G06F 12/02 (2006.01)
G06F 3/06 (2006.01)

[12] 发明专利说明书

专利号 ZL 200610168702.6

[45] 授权公告日 2009 年 10 月 28 日

[11] 授权公告号 CN 100555244C

[22] 申请日 2006.12.19
[21] 申请号 200610168702.6
[30] 优先权
[32] 2006. 3.29 [33] JP [31] 2006 - 092217
[73] 专利权人 株式会社日立制作所
地址 日本东京都
[72] 发明人 田中胜也 岛田健太郎
[56] 参考文献
US2001023472A1 2001.9.20
CN1248334A 2000.3.22
CN1701513A 2005.11.23
JP2001318829A 2001.11.16
审查员 邓 隽

[74] 专利代理机构 北京银龙知识产权代理有限公司
代理人 许 静

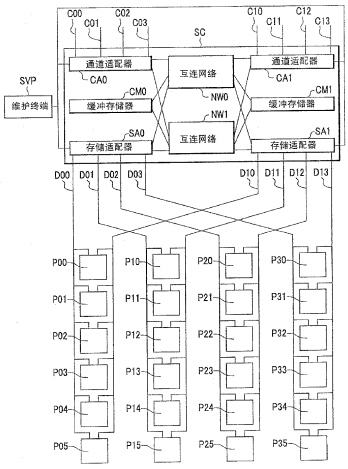
权利要求书 5 页 说明书 25 页 附图 23 页

[54] 发明名称

使用闪存的存储系统及其平均读写方法

[57] 摘要

一种使用闪存存储器的存储系统，包括存储控制器和作为存储介质的多个闪存存储器模块。每个闪存存储器模块包括至少一个闪存存储器芯片，以及用于对属于该闪存存储器芯片的存储块的擦除次数进行平均的存储器控制器。存储控制器把多个闪存存储器模块组合成第一逻辑组，把用于访问属于第一逻辑组的闪存存储器模块的第一地址转变成用于在存储控制器中指示第一地址的第二地址，并且把多个第一逻辑组组合成第二逻辑组。



1. 一种使用闪存存储器的存储系统，包括存储控制器和作为存储介质的多个闪存存储器模块，

每个闪存存储器模块包括至少一个闪存存储器芯片，以及用于对属于该闪存存储器芯片的存储块的擦除次数进行平均的存储器控制器，

所述存储控制器：

把多个闪存存储器模块组合成第一逻辑组，

把属于第一逻辑组的多个闪存存储器模块相对应的多个第一地址转换成用于在存储控制器中指示第一地址的第二地址，并且

把多个第一逻辑组组合成第二逻辑组，

其中，所述存储控制器：

在所述第一地址限定的逻辑页面地址区域中，在具有最大平均擦除次数的闪存存储器模块中选择具有最大总写入量的逻辑页面地址区域，并且在具有最小平均擦除次数的闪存存储器模块中选择具有最小总写入量的逻辑页面地址，并且

在两个所选择的逻辑页面地址区域之间执行数据交换，然后

改变第一地址与第二地址之间的映射，

从而在属于第一逻辑组的多个闪存存储器之间对每个闪存存储器的平均擦除次数进行平均，并且

所述存储器控制器对每个闪存存储器模块中的属于闪存存储器芯片的块的擦除次数进行平均。

2. 根据权利要求 1 所述的使用闪存存储器的存储系统，其中

第二逻辑组包括用于在闪存存储器模块中的任一个上出现故障时重建所记录的数据的冗余信息。

3. 根据权利要求 1 所述的使用闪存存储器的存储系统，其中

第二逻辑组是 RAID 级 0、RAID 级 1、RAID 级 1+0、RAID 级 3、RAID 级 5 或者 RAID 级 6 中任何一个上的逻辑组，并且为构成第二逻辑组的每一个第一逻辑组提供相等的容量。

4. 根据权利要求 1 或 2 所述的使用闪存存储器的存储系统，其中

第一值是存储系统操作期间每个闪存存储器模块的有效写入速度与存储系统耐用性的乘积除以闪存存储器耐用性而得到的商；

第二值是每个闪存存储器模块的持续的写入速度与存储系统耐用性的乘积除以闪存存储器耐用性而得到的商；并且

将第一逻辑组的容量设置为不小于第一值且不大于第二值。

5. 根据权利要求 1 或 2 所述的使用闪存存储器的存储系统，

其中

第二逻辑组是 RAID 级 2 或者 RAID 级 4 上的逻辑组；并且

在构成第二逻辑组的第一逻辑组中，如果用于存储冗余信息的第一逻辑组的数目是 m，而用于存储数据的第一逻辑组的数目是“n”，

则将用于存储冗余信息的第一逻辑组的容量设置为用于存储数据的第一逻辑组的容量的至少一倍且不大于“n/m”倍。

6. 根据权利要求 1 或 2 所述的使用闪存存储器的存储系统，其中将经由第一地址访问的存储区域设置为大于经由第二地址访问的存储区域。

7. 根据权利要求 1 或 2 所述的使用闪存存储器的存储系统，其中存储控制器存储每个闪存存储器模块的第一地址与第二地址之间的映射信息，以及每个闪存存储器模块中存储块的平均擦除次数。

8. 根据权利要求 1 或 2 所述的使用闪存存储器的存储系统，其中存储控制器构建多个第二逻辑组。

9. 根据权利要求 1 或 2 所述的使用闪存存储器的存储系统，其中存储控制器构建不同的 RAID 级上的多个第二逻辑组。

10. 根据权利要求 1 或 2 所述的使用闪存存储器的存储系统，其中当激活存储系统或者当存储介质连接于存储系统时，存储控制器判断该存储介质是否是闪存存储器。

11. 一种用于使用闪存存储器的存储系统的平均读写方法，该存储系统包括：

闪存存储器模块，包括至少一个闪存存储器芯片和用于对属于该闪存存储器芯片的存储块的擦除次数进行平均的存储器控制器；以及

存储控制器，用于

把闪存存储器模块组合成第一逻辑组，

把用于访问属于第一逻辑组的闪存存储器模块的第一地址转变成用于在存储控制器中指示第一地址的第二地址，以及

把多个第一逻辑组组合成第二逻辑组，

该方法包括：

允许存储控制器对于闪存存储器模块中的每个预定存储区域的写入量提供次数管理的步骤；

允许存储控制器计算平均擦除次数的步骤，该平均擦除次数是通过把经过预定时段每个闪存存储器模块的总写入量除以闪存存储器模块的容量而得到的；以及

允许存储控制器判断平均擦除次数的最大值和最小值之间的差是否不小于预定值的第一判断步骤，

其中

在第一判断步骤，如果平均擦除次数的差不小于预定值，则该方法进一步包括：允许存储控制器在具有平均擦除次数的最大差值的闪存存储器模块之中，在具有最大写入量的存储区域和具有最小写入量的存储区域之间交换数据，并且改变第一地址与第二地址之间的映射信息的步骤。

12. 根据权利要求 11 所述的用于使用闪存存储器的存储系统的平均读写方法，进一步包括：

允许存储控制器在执行数据交换步骤之后判断是否存在没有交换数据的多个闪存存储器模块的第二判断步骤，

其中

在第二判断步骤，如果判断存在没有交换数据的多个闪存存储器模块，则该方法进一步包括允许存储控制器回到第一判断步骤的步骤。

13. 根据权利要求 11 或 12 所述的用于使用闪存存储器的存储系统的平均读写方法，其中

如果每个第一逻辑组的总写入量达到预定值，存储控制器执行第一判断步骤。

14. 根据权利要求 11 或 12 所述的用于使用闪存存储器的存储系统的平均读写方法，其中

将第一地址的存储区域设置为大于第二地址的存储区域，并且

数据交换步骤包括通过在第一地址的存储区域中重写数据来交换数据的步骤。

15. 根据权利要求 11 或 12 所述的用于使用闪存存储器的存储系统的平均读写方法，其中

如果每个第一逻辑组的总写入量达到预定值，则存储控制器执行第一判断步骤，

将第一地址的存储区域设置为大于第二地址的存储区域，并且

数据交换步骤包括通过在第一地址的存储区域中重写数据来交换数据的步骤。

16. 根据权利要求 11 所述的用于使用闪存存储器的存储系统的平均读写方法，其中为闪存存储器模块提供自由区域，该自由区域的大小等于存储器控制器对于闪存存储器模块中的每个预定存储区域可以传输的传输数据的大小，为该自由区域提供关于写入量的次数管理。

17. 根据权利要求 11 所述的用于使用闪存存储器的存储系统的平均读写方法，其中为闪存存储器模块提供与闪存存储器模块中的预定存储区域大小相同的自由区域，为该自由区域提供关于写入量的次数管理。

18. 根据权利要求 11 所述的用于使用闪存存储器的存储系统的平均读写方法，其中

如果用新的闪存存储器模块替换原有闪存存储器模块，

则将替换前的原有闪存存储器模块中进行了次数管理的每个预定存储区域的写入量设置为替换后的闪存存储器模块中的每个预定存储区域的写入量。

19. 一种用于使用闪存存储器的存储系统的平均读写方法，该存储系统包括：

闪存存储器模块，包括至少一个闪存存储器芯片和用于对属于该闪存存储器芯片的存储块的擦除次数进行平均的存储器控制器；以及

存储控制器，用于

把闪存存储器模块组合成第一逻辑组，

把用于访问属于第一逻辑组的闪存存储器模块的第一地址转变成用于在存储控制器中指示第一地址的第二地址，以及

把多个第一逻辑组组合成第二逻辑组，

该方法包括：

允许存储控制器对于闪存存储器模块中的每个预定存储区域的写入量提供次数管理的步骤；

允许存储控制器将作为预定时间处记录的原有平均擦除次数的第一平均擦除次数加到第二平均擦除次数上从而计算每个闪存存储器模块的第三平均擦除次数的步骤，其中该第二平均擦除次数是将所述预定时间以后的总写入量除以闪存存储器模块的容量而得到的，

允许存储控制器判断平均擦除次数的最大值和最小值之间的差是否不小于预定值的第一判断步骤，

其中

在第一判断步骤，如果判定差不小于预定值，则该方法进一步包括：允许存储控制器在具有擦除次数的最大差值的闪存存储器模块之中，在具有最大写入量的存储区域和具有最小写入量的存储区域之间交换数据，并且改变第一地址与第二地址之间的映射信息的步骤。

20. 根据权利要求 19 所述的用于使用闪存存储器的存储系统的平均读写方法，进一步包括允许存储控制器用所述第一平均擦除次数和所述第二平均擦除次数的和替代第一平均擦除次数的步骤。

使用闪存的存储系统及其平均读写方法

技术领域

本发明涉及一种能够在多个闪存存储器模块之间进行平均读写的使用闪存存储器的存储系统，用于该存储系统的平均读写方法，以及用于该存储系统的平均读写程序。

背景技术

本申请请求 2006 年 3 月 29 日提交的、申请号为 2006-092217 的日本专利申请的利益，在此处通过参考而引入其公开。

用于存储数据的系统（以下简称“存储系统”）通常包括随机存取非易失性存储介质。随机存取非易失性存储介质包括，例如磁盘或光盘。近来，常见的存储系统具有许多光盘驱动器。

当各种半导体技术进一步提高时，开发了诸如闪存存储器之类的非易失性半导体存储器，在这种非易失性半导体存储器上能够对数据进行擦除。闪存存储器是一种半导体存储器，其是用作只读存储器（ROM）以及既可读又可写的随机存取存储器（RAM）的非易失性存储器。与具有许多小光盘驱动器的存储系统相比，把闪存存储器作为存储介质的存储系统在使用寿命、功耗节约以及存取时间上比较出色。

此处将给出对闪存存储器的说明。

通常，由于其属性的原因，不能把数据直接重写到闪存存储器上。也就是说，为了把数据重写到闪存存储器上，不得不把存储在闪存存储器上的有效数据转移到其他位置。因此，在逐个存储块的基础上（on block by block basis）对存储数据进行擦除。此后，把其他数据写入到已经擦除了数据的每个存储块中。存储块表示每次擦除数据的的单位存储区域。

在闪存存储器中，例如，其中擦除了数据的存储区域总是设置为“1”。因此，当重写数据时有可能通过二进制比特转换把“1”重写为“0”。然而，除非擦除了存储数据，否则不可能直接把“0”重写为“1”。为了把数据重写

到闪存存储器上，擦除了闪存存储器的整个存储块。因此，当把数据重写到闪存存储器上时，闪存存储器总是需要进行存储块擦除。

闪存存储器有存储块擦除次数的限制。例如，保证了存储块擦除次数高达每个存储块 100,000 次。如果特定存储块由于密集地进行数据重写的原因而经历了过多的擦除次数，那么就会变得再也不可能对该存储块中的数据进行擦除，这会引起问题。因此，在使用闪存存储器作为存储介质的存储系统中，必须准备平均读写处理以防止在特定存储块上出现密集的擦除次数。

JP-A-8-16482 中公开了一种平均读写方法，其中存储系统采用映射管理方法以便提供主机与闪存存储器之间的存储块联系关系（block association relationship）的灵活性，以解决当计算机访问逻辑存储块时由逻辑存储块单方面地（one-sidedly）选择闪存存储器的物理存储块的问题。在这个方案中，这种常规存储系统对主机访问的每个逻辑存储块的写入次数，以及由存储系统擦除的每个物理存储块的擦除次数进行管理。如果存在写入次数过大的逻辑存储块和擦除次数过大的物理存储块；以及写入次数较小的逻辑存储块和擦除次数较小的物理存储块，则以下述方式来提供映射，即，允许写入次数过大的逻辑存储块对应于擦除次数较小的物理存储块，并且允许写入次数较小的逻辑存储块对应于擦除次数过大的物理存储块。

通常，闪存存储器模块（以下简称“PDEV”）是由存储器控制器和多个闪存存储器芯片构成的，并且该存储器控制器提供与上述常规方案相同的平均读写处理。在大规模存储系统中，可以想到作为存储介质的许多闪存存储器彼此相连接以建立大容量存储器。在这种情况下，通过利用控制器来为每个闪存存储器提供平均读写。然而，在其中特定闪存存储器模块经历了密集的重写次数的情况中，当闪存存储器模块的擦除次数愈加增大时，模块损耗越快。为了防止特定模块上的擦除次数增大，需要在多个闪存存储器模块之间提供平均读写。

如果把上述平均读写方案施加于具有与之相连接的许多闪存存储器的存储系统，则会存在闪存存储器模块中的存储器控制器会遮蔽闪存存储器芯片中的物理存储块的问题，这将防碍存储系统中的存储控制器管理每个物理存储块的擦除次数。

此外，如果把常规平均读写方案应用于在闪存存储器模块中不使用存储器控制器（那就是说，不为每个闪存存储器模块提供平均读写）的整个存储系统上，则存储系统必须整体地管理非常多的物理存储块的擦除次数，导致管理负担的增加和存储系统性能的恶化。

鉴于上述问题，需要提供一种使用闪存存储器的存储系统，其能够在多个闪存存储器模块之间进行平均读写而不必使用闪存存储器物理存储块上的映射信息，以及需要提供一种用于该存储系统的平均读写方法和用于该存储系统的平均读写程序。

发明内容

在本发明的一个方面，提供了一种使用闪存存储器的存储系统，包括存储控制器和作为存储介质的多个闪存存储器模块。每个闪存存储器模块包括至少一个闪存存储器芯片，以及用于对属于该闪存存储器芯片的存储块的擦除次数进行平均的存储器控制器。存储控制器把多个闪存存储器模块组合成第一逻辑组，把用于访问属于第一逻辑组的闪存存储器模块的第一地址转变成用于在存储控制器中指示第一地址的第二地址，并且把多个第一逻辑组组合成第二逻辑组。

在本发明的另一个方面，提供了一种用于使用闪存存储器的存储系统的平均读写方法，该存储系统包括：闪存存储器模块，包括至少一个闪存存储器芯片和用于对属于该闪存存储器芯片的存储块的擦除次数进行平均的存储器控制器；以及存储控制器，用于把闪存存储器模块组合成第一逻辑组，把用于访问属于第一逻辑组的闪存存储器模块的第一地址转变成用于在存储控制器中指示第一地址的第二地址，以及把多个第一逻辑组组合成第二逻辑组。

该方法包括：允许存储控制器对于闪存存储器模块中的每个预定存储区域的写入量提供次数管理的步骤；允许存储控制器计算平均擦除次数的步骤，该平均擦除次数是通过把经过预定时段每个闪存存储器模块的总写入量除以闪存存储器模块的容量而得到的；以及允许存储控制器判断平均擦除次数的最大值和最小值之间的差是否不小于预定值的第一判断步骤。在第一判断步骤，如果平均擦除次数的差不小于预定值，则该方法进一步包括：允许存储控制器在具有平均擦除次数的最大差值的闪存存储器模块之中，在具有最大

写入量的存储区域和具有最小写入量的存储区域之间交换数据，并且改变第一地址与第二地址之间的映射信息的步骤。

在本发明的又一个方面中，提供了一种用于使用闪存存储器的存储系统的平均读写程序，该存储系统包括：闪存存储器模块，包括至少一个闪存存储器芯片和用于对属于该闪存存储器芯片的存储块的擦除次数进行平均的存储器控制器；以及存储控制器，用于把多个闪存存储器模块组合成第一逻辑组，把用于访问属于第一逻辑组的闪存存储器模块的第一地址转变成用于在存储控制器中指示第一地址的第二地址，以及把多个第一逻辑组组合成第二逻辑组。

该程序执行：允许计算机对于闪存存储器模块中的每个预定存储区域的写入量提供次数管理的流程；允许计算机计算平均擦除次数的流程，该平均擦除次数是通过把经过预定时段每个闪存存储器模块的总写入量除以闪存存储器模块的容量而得到的；计算机判断平均擦除次数的最大值和最小值之间的差是否不小于预定值的第一判断的流程。如果差值不小于预定值，则该程序进一步包括用于允许计算机改变第一地址和第二地址之间的映射信息的流程。

当结合附图一起阅读以下的发明的详细说明，本发明的其他特征和优点会变得更加明显。

附图说明

图 1 是显示根据本发明实施例的存储系统的结构的框图。

图 2 是显示通道适配器的结构的框图。

图 3 是显示存储适配器的结构的框图。

图 4 是显示闪存存储器模块的结构的框图。

图 5 是显示闪存存储器模块的存储块的结构的框图。

图 6 是显示根据本发明的该实施例的存储系统的逻辑组结构和地址转换层次（hierarchy）的框图。

图 7 是显示根据本发明该实施例的存储系统的 RAID 组的结构的框图。

图 8 是显示其中闪存存储器模块和硬盘驱动器与存储控制器相连接的例子的框图。

图 9 是显示用于在多个闪存存储器模块之间进行平均读写的方法的流程图。

图 10 显示了在根据本发明实施例的平均读写流程所伴随的数据交换流程之前，虚拟页面地址和逻辑页面地址之间的地址转换表。

图 11 显示了在根据本发明实施例的平均读写流程所伴随的数据交换流程之后，虚拟页面地址和逻辑页面地址之间的地址转换表。

图 12 显示了存储控制器中管理的用于每个闪存存储器模块的擦除次数管理表。

图 13 是用于说明平均读写流程所伴随的数据交换流程之前的虚拟页面地址和逻辑页面地址之间的映射的框图。

图 14 是用于说明平均读写流程所伴随的数据交换流程之后的虚拟页面地址和逻辑页面地址之间的映射的框图。

图 15 显示了数据交换流程之前的初始状态。

图 16 显示了数据交换流程期间的状态。

图 17 显示了另一个数据交换流程期间的状态。

图 18 显示了另一个数据交换流程期间的状态。

图 19 显示了另一个数据交换流程期间的状态。

图 20 显示了另一个数据交换流程期间的状态。

图 21 显示了另一个数据交换流程期间的状态。

图 22 显示了另一个数据交换流程期间的状态。

图 23 显示了另一个数据交换流程期间的状态。

图 24 显示了数据交换流程之后的最终状态。

图 25 是显示在数据交换流程之前/之后偏移量值如何转变的表。

图 26 是显示如图 15 到图 24 所述的、偏移量值为“0”的逻辑页面地址区域与偏移量值为“1”的逻辑页面地址区域之间的数据交换流程的流程图。

图 27 是显示偏移量值为“0”的逻辑页面地址区域与偏移量值为“0”的逻辑页面地址区域之间的数据交换流程的流程图。

图 28 是显示偏移量值为“1”的逻辑页面地址区域与偏移量值为“1”的逻辑页面地址区域之间的数据交换流程的流程图。

图 29 是用于说明数据交换流程之前虚拟页面地址和逻辑页面地址之间的映射的框图。

图 30 是用于说明数据交换流程之后虚拟页面地址和逻辑页面地址之间的映射的框图。

图 31 是用于说明数据交换流程之前虚拟页面地址与逻辑页面地址之间的地址转换表的一个表格。

图 32 是用于说明数据交换流程之后虚拟页面地址与逻辑页面地址之间的地址转换表的一个表格。

图 33 是用于说明数据交换流程之前的自由区域管理表的一个表格。

图 34 是用于说明数据交换流程之后的自由区域管理表的一个表格。

图 35 是显示如何替换闪存存储器模块的步骤的流程图。

图 36 显示了当闪存存储器模块上出现故障时的情况。

图 37 是用于说明闪存存储器模块替换之后的状态的框图。

图 38 是显示在闪存存储器模块替换之后如何重建数据的框图。

图 39 是显示其中将备份组中的闪存存储器模块替换为新模块的情况的框图。

具体实施方式

在下文中参考附图，提供关于本发明的实施例的说明。

<概述>

根据本发明实施例的一种使用闪存存储器的存储系统，包括存储控制器和作为存储介质的多个闪存存储器模块。每个闪存存储器模块（例如，闪存存储器模块 P0）包括至少一个闪存存储器芯片（例如，闪存存储器芯片 405），以及用于对属于该闪存存储器芯片的存储块（例如，存储块 406）的擦除次数进行平均的存储器控制器（例如，控制器 MC）。存储控制器（例如，存储控制器 SC）把多个闪存存储器模块组合成第一逻辑组（例如，平均读写组 W00），并且把用于访问属于第一逻辑组的闪存存储器模块的第一地址（例如，逻辑页面地址 600）转换成用于在存储控制器中表示第一地址的第二地址（例如，虚拟页面地址 604），并且将多个第一逻辑组进行组合以构成第二逻辑组（例如，RAID（Redundant Array of Independent Disks，独立磁盘冗余阵列）

组）。

图 1 是显示根据本实施例的存储系统的结构的框图。

存储系统 100 包括存储控制器 SC 和闪存存储器模块 P00 到 P35。

存储控制器 SC 包括通道适配器 CA0、CA1，缓冲存储器 CM0、CM1，存储适配器 SA0、SA1，以及互连网络 NW0、NW1。虽然通道适配器 CA0、CA1，缓冲存储器 CM0、CM1，存储适配器 SA0、SA1 在附图中是分别成对显示的，但是那些组件不限制于成对提供，而是可以以任意数量提供。

互连网络 NW0、NW1 可以是构成存储控制器 SC 的部分的交换器（Switch）以及互连设备。具体地说，互连网络 NW0、NW1 将通道适配器 CA0、缓冲存储器 CM0 以及存储适配器 SA0 相互连接。互连网络 NW0、NW1 还将通道适配器 CA1、缓冲存储器 CM1 以及存储适配器 SA1 相互连接。

如随后在图 2 中所示的通道适配器 CA0，经由通道 C00、C01、C02、C03 而与外部主机系统（未显示）相连接。通道适配器 CA1 经由通道 C10、C11、C12、C13 而与外部主机系统（未显示）相连接。主机系统表示用于对根据本实施例的存储系统 100 进行数据读取和写入的计算机。存储系统 100 经由光纤通道交换器、FC-AL（光纤频通道仲裁环）、SAS（串行附属 SCSI）扩展器等等而与主机系统或其他存储系统相连接。

缓冲存储器 CM0 暂时地存储从通道适配器 CA0 和存储适配器 SA0 所接收到的数据。缓冲存储器 CM1 暂时地存储从通道适配器 CA1 和存储适配器 SA1 中所接收的数据。

存储适配器 SA0 与闪存存储器模块 P00 等等（随后在图 3 中描述）相连接。具体地说，存储适配器 SA0 经由通道 D00 而与闪存存储器模块 P00 到 P05 相连接。存储适配器 SA0 还经由通道 D01 与闪存存储器模块 P10 到 P15 相连接。存储适配器 SA0 更进一步经由通道 D02 与闪存存储器模块 P20 到 P25 相连接。此外，存储适配器 SA0 经由通道 D03 而与闪存存储器模块 P30 到 P35 相连接。

存储适配器 SA1 与闪存存储器模块 P00 等等相连接。具体地说，存储适配器 SA1 经由通道 D10 而与闪存存储器模块 P00 到 P05 相连接。存储适配器 SA1 还经由通道 D11 与闪存存储器模块 P10 到 P15 相连接。存储适配器 SA1

更进一步经由通道 D12 与闪存存储器模块 P20 到 P25 相连接。此外，存储适配器 SA1 经由通道 D13 而与闪存存储器模块 P30 到 P35 相连接。具体地说，存储适配器和闪存存储器模块经由光纤通道交换器、FC-AL、SAS 扩展器等等而相互连接。

通道适配器 CA0、CA1 和存储适配器 SA0、SA1 与维护终端 SVP 相连接。维护终端 SVP 向通道适配器 CA0、CA1 和/或存储适配器 SA0、SA1 发送由存储系统 100 的管理员所输入的设置信息。代替采用存储适配器 SA0 和通道适配器 CA0，存储系统 100 可装有单个适配器。在这种情况下，这个适配器执行由存储适配器 SA0 和通道适配器 CA0 所执行的流程。

图 2 是显示通道适配器的结构的框图。通道适配器 CA0 包括主机通道接口 21、缓冲存储器接口 22、网络接口 23、处理器 24、本地存储器 25 以及处理器外围设备控制单元 26。

主机通道接口 21 经由通道 C00、C01、C02、C03 而与外部主机系统（未显示）相连接。主机通道接口 21 在通道 C00、C01、C02、C03 上的数据传送协议和存储控制器 SC 内部的数据传送协议之间进行相互转换。

缓冲存储器接口 22 与互连网络 NW0、NW1 相连接。网络接口 23 与维护终端 SVP 相连接。主机通道接口 21 和缓冲存储器接口 22 经由信号线 27 相互连接。

处理器 24 通过运行存储在本地存储器 25 上的每一个程序而执行各种流程。具体地说，处理器 24 控制主机系统和互连网络 NW0、NW1 之间的数据传送。

本地存储器 25 存储由处理器 24 所运行的程序。本地存储器 25 存储处理器 24 所要查阅的表。该表可以由管理员来设置或改变。

在这种情况下，管理员输入关于设置或改变该表的信息。维护终端 SVP 经由网络接口 23 把管理员输入的信息发送给处理器 24。处理器 24 根据所接收的信息而生成或改变该表。然后，处理器 24 将该表存储在本地存储器 25 上。

处理器外围设备控制单元 26 控制主机接口通道 21、缓冲存储器接口 22、网络接口 23、处理器 24 以及本地存储器 25 之间的数据传送。处理器外围设

备控制单元 26 例如是芯片组等等。通道适配器 CA1 具有与通道适配器 CA0 相同的结构。因此，此处省去了对通道适配器 CA1 的说明。

图 3 是显示根据本实施例的存储适配器的框图。存储适配器 SA0 包括缓冲存储器接口 31、存储器通道接口 32、网络接口 33、处理器 34、本地存储器 35 以及处理器外围设备控制单元 36。

缓冲存储器接口 31 与互连网络 NW0、NW1 相连接。存储器通道接口 32 与通道 D00、D01、D02、D03 相连接。存储器通道接口 32 在通道 D00、D01、D02、D03 上的数据传送协议和存储控制器 SC 内部的数据传送协议之间进行相互转换。缓冲存储器接口 31 和存储器通道接口 32 经由信号线 37 相互连接。网络接口 33 与维护终端 SVP 相连接。

处理器 34 通过运行存储在本地存储器 35 上的每一个程序而执行各种流程。

本地存储器 35 存储由处理器 34 所运行的程序。本地存储器 35 同样存储由处理器 34 所查阅的表。该表可以由管理员来设置或改变。

在这种情况下，管理员把关于设置或改变该表的信息输入到维护终端 SVP 中。维护终端 SVP 经由网络接口 33 把管理员输入的信息发送给处理器 34。处理器 34 根据所接收的信息而生成或改变该表。然后，处理器 34 将该表存储在本地存储器 35 上。

处理器外围设备控制单元 36 控制缓冲存储器接口 31、存储器通道接口 32、网络接口 33、处理器 34 和本地存储器 35 之间的数据传送。处理器外围设备控制单元 36 可以是芯片组等等。存储适配器 SA1 具有与存储适配器 SA0 相同的结构。因此，此处省去了对存储适配器 SA1 的说明。

图 4 是显示根据本发明的闪存存储器模块的结构的框图。闪存存储器模块 P00 包括存储器控制器 MC 和闪存存储器 MEM。闪存存储器 MEM 存储数据。存储器控制器 MC 读取/写入或擦除存储在闪存存储器 MEM 上的数据。

存储器控制器 MC 包括处理器（μP）401、接口单元（I/F）402、数据传送单元（HUB）403、存储器（RAM）404 和存储器（ROM）407。

闪存存储器 MEM 包括多个闪存存储器芯片 405。每个闪存存储器芯片 405 包括多个存储块 406 以便在其上存储数据。每个存储块 406 是存储器控

制器 MC 擦除数据的一个单位，如随后在图 5 中所描述的那样。

存储块 406 包括多个页面。页面是存储器控制器 MC 读/写数据的单位，如随后在图 5 中所描述的那样。把每个页面分类为有效页面、无效页面、未使用页面或者坏页面。有效页面是存储有效数据的页面。无效页面是存储无效数据的页面。未使用页面是没有存储数据的页面。坏页面是物理上不可用的页面，例如，因为页面包含有损坏的存储元件。

接口单元 402 经由通道 D00 而与存储控制器 SC 中的存储适配器 SA0 相连接。接口单元 402 也经由通道 D10 而与存储控制器 SC 中的存储适配器 SA1 相连接。

接口单元 402 从存储适配器 SA0 和存储适配器 SA1 接收指令。来自存储适配器 SA0 和存储适配器 SA1 的指令诸如是 SCSI 命令。

具体地说，接口单元 402 从存储适配器 SA0 和存储适配器 SA1 接收数据。然后接口单元 402 把所接收的数据存储在存储器 404 上。接口单元 402 也把存储在存储器 404 上的数据发送给存储适配器 SA0 和存储适配器 SA1。

存储器 404 例如是能够以高速来读/写数据的动态随机存取存储器。存储器 404 暂时地存储由接口单元 402 发送或接收的数据。存储器 407 是用于存储要由处理器 401 来运行的程序的非易失性存储器。当激活闪存存储器模块 P00 时，把程序从存储器 407 拷贝到存储器 404 上从而处理器 401 能够运行该程序。存储器 404 存储处理器 401 所要查阅的表。该表可包括例如闪存存储器 MEM 的逻辑页面地址和物理页地址之间的地址转换表。逻辑页面地址是当从闪存存储器模块外部（例如，从存储适配器 SA0）访问作为在闪存存储器上读/写数据的单位的页面时使用的地址。物理页面地址是当存储器控制器 MC 访问作为在闪存存储器上读/写数据的单位的页面时所使用的地址。

数据传送单元 403 可以是例如用于将处理器 401、接口单元 402、存储器 404、存储器 407 以及闪存存储器 MEM 相互连接的交换器，并控制这些组件之间的数据传送。

处理器 401 通过运行存储在存储器 404 上的每一个程序而执行各种流程。例如，处理器 401 查阅闪存存储器的逻辑页面地址和物理页面地址之间的地址转换表（存储在存储器 404 上），然后根据该表在闪存存储器 MEM 上读/

写数据。处理器 401 为闪存存储器模块中的存储块 406 提供回收流程（存储块回收流程）和平均读写流程。

回收流程（存储块回收流程）是把存储块 406 中的无效页面重建成未使用页面的流程，以便能够把含有较少未使用页面的存储块重建为再次有效。此处假定作为回收流程的目标的存储块 406 包括有效页面、无效页面和未使用页面，其中很多是无效页面。在这种情况下，需要擦除无效页面以便增加未使用页面。然而，擦除流程不是在逐页面基础上进行的，而是在逐存储块基础上进行的。因此，需要以这样一种方式来把存储块重建为有效，即，把目标存储块的有效页面拷贝到空存储块上，然后擦除目标存储块。具体地说，处理器 401 把作为回收流程的目标的存储块 406（即目标存储块）中有效页面上所存储的数据拷贝到未使用的存储块上。处理器 401 将拷贝了数据的该未使用的存储块的逻辑存储块号码修改为目标存储块的逻辑存储块号码。然后，擦除目标存储块上的所有数据，从而完成回收流程。

例如，随着处理器 401 把更多数据写入到存储块 406 上，存储块 406 中更多的未使用页面被减少。然后，如果存储块 406 变得缺少未使用页面，则处理器 401 不能再把数据写到存储块 406 上。因此，处理器 401 通过在存储块 406 上执行回收流程而把无效页面回收成未使用页面。

平均读写流程是用于对存储块 406 的擦除次数进行平均的流程，从而能够提高闪存存储器 MEM 耐用性。闪存存储器 MEM 经历的数据擦除次数越多，闪存存储器 MEM 最终达到其耐用性的速度就越快。通常，保证闪存存储器 MEM 的耐用性高达 10,000 到 100,000 次。

现在，其他闪存存储器模块 P01 到 P35 具有与闪存存储器模块 P00 相同的结构。因此，省去了对这些模块 P01 到 P35 的说明。

图 5 是显示闪存存储器模块的存储块的结构的图示。闪存存储器模块 P00 的存储块 406 包括多个页面 501。存储块 406 通常包括几十个页面 501（例如，32 个页面、64 个页面）。

每个页面 501 是存储器控制器 MC 之类读/写数据的一个单位。例如，在 NAND 型闪存存储器中，存储器控制器 MC 之类以 20 到 30 μs 或更低/页面的速度来读数据，并且以 0.2 到 0.3 ms /页面的速度写数据。存储器控制器 MC

之类以 2 到 4 ms /存储块的速度来擦除数据。

页面 501 包括数据段 502 和冗余段 503。例如，数据段 502 包含 512 字节，冗余段 503 包含 16 字节。数据段 502 存储顺序的数据（ordinal data）。

冗余段 503 存储关于页面 501 的管理信息和错误校正码。管理信息包括偏移量地址和页面状态。偏移量地址是页面 501 所属的存储块 406 中的相对地址。页面状态显示了页面 501 是否是有效页面、无效页面、未使用页面、或正在进行处理的页面。错误校正码是用于检测和校正页面 501 上的错误的码，诸如 Humming 码。

图 6 是显示逻辑组的结构和地址转换的层次的框图。图 6 中的存储系统具有与图 1 中的存储系统相同的硬件配置。为了方便起见，仅显示了作为与闪存存储器模块 P00 到 P35 连接的存储控制器 SC 的通道的通道 D00、D01、D02、D03，并且在图中省略了通道 D10、D11、D12、D13。

在根据本实施例的存储系统 100 中，在同一个通道上相互连接的闪存存储器模块构成了一个平均读写组（WDEV）。例如，通道 D00 上的闪存存储器模块 P00 到 P03 构成了平均读写组 W00。类似地，通道 D01 上的闪存存储器模块 P10 到 P13 构成了平均读写组 W10；通道 D02 上的闪存存储器模块 P20 到 P23 构成了平均读写组 W20；以及通道 D03 上的闪存存储器模块 P30 到 P33 构成了平均读写组 W30。

从存储控制器 SC 可经由每个闪存存储器模块相应的逻辑页面地址访问每个闪存存储器模块。例如，可经由模块的每个相应的逻辑页面地址 600 而访问通道 D00 上的闪存存储器模块 P00 到 P03。类似地，可经由模块的每个相应的逻辑页面地址 601 而访问通道 D01 上的闪存存储器模块 P10 到 P13；可经由模块的每个相应的逻辑页面地址 602 而访问通道 D02 上的闪存存储器模块 P20 到 P23；以及可经由模块的每个相应的逻辑页面地址 603 而访问通道 D03 上的闪存存储器模块 P30 到 P33。

存储控制器 SC 把属于相同的平均读写组的闪存存储器模块的多个逻辑页面地址放到一起成为一组并且把该组转换成单个虚拟页面地址。例如，存储控制器 SC 把属于平均读写组 W00 的闪存存储器模块 P00 到 P03 的逻辑页面地址 600 放在一起，并且把该组转换成虚拟页面地址 604。类似地，把属

于平均读写组 W10 的闪存存储器模块 P10 到 P13 的逻辑页面地址 601 放在一起并转换成虚拟页面地址 605；把属于平均读写组 W20 的闪存存储器模块 P20 到 P23 的逻辑页面地址 602 放在一起并转换成虚拟页面地址 606；以及把属于平均读写组 W30 的闪存存储器模块 P30 到 P33 的逻辑页面地址 603 放在一起并转换成虚拟页面地址 607。

如上所述，存储控制器 SC 把逻辑页面地址转换成虚拟页面地址。以这种方式，即使为了平均读写的目的而在闪存存储器模块之间传递数据并且改变了相关的逻辑页面地址，作为较高层次设备的存储控制器 SC 也能够改变逻辑页面地址与对应于该逻辑地址的虚拟页面地址之间的映射，从而能够不相冲突地访问数据。

在根据本实施例的存储系统 100 中，将多个平均读写组进行组合从而构成单个 RAID 组（VDEV）。在图 6 中，将四个平均读写组 W00 到 W30 组合成一个 RAID 组 V00。构成单个 RAID 组的每一个平均读写组中的每个虚拟页面地址页面区域具有相同的存储容量。将一个或多个 RAID 组中的区域组合成单个逻辑卷 608，其是存储控制器 SC 向主机系统显示的存储区域。

通道 D00 上的闪存存储器模块 P04、P05 构成备份组（YDEV）Y00。类似地，通道 D01 上的闪存存储器模块 P14、P15 构成备份组 Y10；通道 D02 上的闪存存储器模块 P24、P25 构成备份组 Y20；以及通道 D03 上的闪存存储器模块 P34、P35 构成备份组 Y30。随后将描述如何替换模块。

图 7 是显示根据本实施例的存储系统 100 的 RAID 组的结构的框图。RAID 组 720 是处于 RAID 级 5 的 RAID 组，由平均读写组 700 到 703 构成。例如，平均读写 700 由闪存存储器模块 730、731 构成。需要指出的是，根据功能把 RAID 分类到一个级别中，诸如 RAID 级 0 或 RAID 级 1 等等。

RAID 组 721 是由平均读写组 704、705 构成的处于 RAID 级 1 的 RAID 组。类似地，RAID 组 722 是由平均读写组 706、707 构成的处于 RAID 级 1 的 RAID 组。

在存储系统 100 中，如果把 RAID 分类到级别 0、1、3、5、6 或 1+0 中，则为构成相同 RAID 组的每一个平均读写组的逻辑页面地址区域提供相等的容量。平均读写组容量的上限是由公式 1 定义的，其下限是由公式 2 定义的。

具体地说，“闪存存储器模块的持续的写入速度”和“系统耐用性”的乘积除以“闪存存储器耐用性”得到第二值（上限）。“对系统进行操作时的闪存存储器模块的有效写入速度”与“系统耐用性”的乘积除以“闪存存储器耐用性”得到第一值（下限）。接着，将每个平均读写组的逻辑页面地址区域的容量设置成不小于第一值且不大于第二值。例如，系统耐用性通常为 5 到 10 年，闪存存储器耐用性通常为 10,000 到 100,000 次。公式 2 中的有效写入速度表示考虑从主机系统到存储系统 100 的写访问率（write access ratio）的有效写入速度。

[公式 1]

$$\text{平均读写组容量值（上限）} = \frac{\text{每个模块持续的写入速度} \times \text{系统耐用性}}{\text{闪存存储器耐用性}}$$

[公式 2]

$$\text{平均读写组容量值（下限）} = \frac{\text{每个模块的有效写入速度} \times \text{系统耐用性}}{\text{闪存存储器耐用性}}$$

将平均读写组的容量设置为落入由公式 1 和公式 2 所定义的范围。通过为平均读写组中的闪存存储器模块提供平均读写，能够保证闪存存储器模块耐用性在存储系统 100 的系统耐用期之内。

RAID 组 723 是由平均读写组 708 到 711 构成的处于 RAID 级 4 的 RAID 组；平均读写组 708 到 710 是用于存储数据的平均读写组；而平均读写组 711 是用于存储奇偶校验（parity）的平均读写组。用于存储奇偶校验的平均读写组的更新次数比用于存储数据的其他平均读写组的更新次数更多。因此，为了提供处于 RAID 级 4 的 RAID 组中的平均读写流程，将用于存储奇偶校验的平均读写组中的逻辑页面地址区域的容量设置为大于用于存储数据的平均读写组中的逻辑页面地址区域的容量。例如，如果构成 RAID 组的平均读写组的数量是“n”，则将用于存储奇偶校验的平均读写组的逻辑页面地址区域的容量设置为不小于用于存储数据的平均读写组的逻辑页面地址区域的容量的一倍且不大于用于存储数据的平均读写组的逻辑页面地址区域的容量的（n-1）倍。

图中未示出，在 RAID 级 2 上，用于存储冗余信息的平均读写组比用于存储数据的平均读写组具有更多的更新次数。例如，在 RAID 级 2 上，如果存在 10 个用于存储数据的平均读写组以及 4 个用于存储奇偶校验的平均读写组（10D4P），则将用于存储冗余信息的平均读写组中的逻辑页面地址区域的容量设置为不小于用于存储数据的平均读写组中的逻辑页面地址区域的容量的一倍且不大于用于存储数据的平均读写组中的逻辑页面地址区域的容量的 10/4=2.5 倍。对于 25D5P，将用于存储冗余信息的平均读写组中的逻辑页面地址区域的容量设置为不小于用于存储数据的平均读写组中的逻辑页面地址区域的容量的一倍且不大于该容量的 25/5 = 5 倍。

换句话说，在 RAID 级 2 或 RAID 级 4 上，如果用于存储数据的平均读写组的数量是“n”，而用于存储冗余信息的平均读写组的数量是“m”，则将用于存储冗余信息的平均读写组中的逻辑页面地址区域的容量设置为不小于用于存储数据的平均读写组中的逻辑页面地址区域的容量的一倍且不大于该容量的“n/m”倍。

以这种方式，通过组合平均读写组而构成了存储控制器 SC 中的每个 RAID 组。具体地说，存储控制器 SC 考虑每个 RAID 组的平均读写组而对其进行管理。因此，每个平均读写组的虚拟页面地址被视为是独立的，而与每个平均读写组中的逻辑页面地址及虚拟页面地址之间的映射无关。因此，存储控制器 SC 能够把处于不同级别的多个 RAID 组相互连接。

图 8 是显示其中闪存存储器模块和硬盘驱动器与存储控制器 SC 相连接的例子的框图。闪存存储器模块 810 到 812 构成了平均读写组 830。闪存存储器模块 813 到 815 构成了平均读写组 831，并且平均读写组 830、831 构成了 RAID 组 840。

类似于图 6，存储控制器 SC 把逻辑页面地址 800 转换成虚拟页面地址 802 以便访问闪存存储器模块 810 到 812 中的任何一个。存储控制器 SC 还把逻辑页面地址 801 转换成虚拟页面地址 803 以便访问闪存存储器模块 813 到 815 中的任何一个。

将硬盘驱动器 820 和 823 组合成 RAID 组 841。类似地，将硬盘驱动器 821 和 824 组合成 RAID 组 842；并且将硬盘驱动器 822 和 825 组合成 RAID

组 843。存储控制器 SC 经由逻辑存储块地址 804 或 805 而访问每个硬盘驱动器。在由硬盘驱动器构成的 RAID 组中，因为不需要平均读写所以没有定义平均读写组。存储控制器 SC 仅在由闪存存储器模块构成的 RAID 组中定义平均读写组，并且把逻辑页面地址转换成虚拟页面地址。

当激活系统或把存储介质与系统相连接时，存储控制器 SC 根据是否需要任何地址转换的判断或者如何配置 RAID 组的判断等等而改变控制，所述判断取决于存储介质是否是闪存存储器或硬盘驱动器。

存储控制器 SC 通过利用由闪存存储器模块构成的 RAID 组 840 或者由硬盘驱动器构成的 RAID 组 841 到 843 中任何一个的区域，或者通过将 RAID 组 840 和 RAID 组 841 到 843 的区域组合而构成逻辑卷 808。闪存存储器模块的存储区域或硬盘驱动器的存储区域的选择可以是这样的，即在闪存存储器模块上存储读取访问较多并且更新次数较少的数据，并且在硬盘驱动器上存储更新次数较多的数据。闪存存储器模块能够利用潜规则（law latency）访问硬盘驱动器。因此，如果根据存储介质的访问属性而选择存储区域，如上所述，则可实现存储系统的高性能。

将参考附图，给出关于根据本实施例的存储系统 100 的操作的说明。

将参考图 9 到图 14，给出关于用于根据本实施例的存储系统 100 的平均读写方法的说明。这个方法在多个闪存存储器模块之间提供平均读写。

图 9 是显示多个闪存存储器模块之间的平均读写流程的流程图。为了简便，假定目标平均读写组 W00 具有两个闪存存储器模块 P00、P04。

图 10 显示了根据本实施例的平均读写流程伴随的数据交换流程之前的虚拟页面地址和逻辑页面地址之间的地址转换表。

图 11 显示了根据本实施例的平均读写流程伴随的数据交换之后的虚拟页面地址和逻辑页面地址之间的地址转换表。

参考图 10 和图 11，表示了虚拟页面地址和逻辑页面地址之间的映射，以及这些映射的偏移量值。在根据本实施例的存储系统中，把逻辑页面地址区域（数据长度）设置为大于相应的虚拟页面地址区域（数据长度）。在逻辑页面地址区域中，如果在起始地址端写有有效数据并且在结束地址端存在自由区域，则指示偏移量值“0”。如果在结束地址端写有有效数据并且在起始

地址端存在自由区域，则指示偏移量值“1”。在这种情况下，自由区域的大小是闪存存储器页面的数据段的整数倍（至少一倍)，并且等于闪存存储器模块中的存储器控制器一次在闪存存储器上读/写的数据量。

图 12 显示了由存储控制器 SC 管理的每个闪存存储器模块的擦除次数管理表。存储控制器 SC 记录作为闪存存储器模块内的数据交换单位的每个区域的总写入量。如公式 3 所示，通过把前一平均擦除次数值与平均处理次数相加可以获得闪存存储器模块中的闪存存储器的平均擦除次数，所述平均处理次数是通过把独立的预定时间段中模块的每个逻辑页面地址区域的单独的总写入量的和，除以模块的整个逻辑页面地址区域容量（模块容量）而得到的。

[公式 3]

$$\text{平均擦除次数} = \text{前一值} + \frac{\sum \begin{matrix}\text{独立的预定时间段中模块的} \\ \text{每个逻辑页面地址区域的} \\ \text{单独的总写入量的和}\end{matrix}}{\text{模块容量}}$$

在图 12 的管理表中，记录了两个平均擦除次数值。一个是最后一次运行平均读写流程时所记录的前一平均擦除次数（f00、f04），另一个是到目前为止的当前平均擦除次数（e00、e04）。从执行前一平均读写流程的上一次的时间直到现在，记录用于管理每个逻辑页面地址区域的写入次数的总写入量。通过公式 3 可容易地计算出当前平均擦除次数。通过管理每个独立的预定时间段中逻辑页面地址区域的总写入量，可以获得对于逻辑页面地址区域最近的访问频率。在该管理表中，用这样的方式来设置移动标志，即，在数据交换流程之前把该标志设置为“0”，在数据交换流程之后把该标志设置为“1”。在公式 3 中，在每个独立的时间段中对总写入量进行管理。假定没有独立的预定时间段，则可以由公式 4 来表示经过整个时间段的总写入量。公式 3 的结果或公式 4 的结果得到相同的平均擦除次数值。

[公式 4]

$$\text{平均擦除次数}=\frac{\sum \begin{matrix}\text{经过整个预定时间段模块的}\\ \text{每个逻辑页面地址区域的}\\ \text{单独的总写入量的和}\end{matrix}}{\text{模块容量}}$$

即使当诸如电源故障之类的故障出现时或当不在系统服务时间内时，也需要保持图 10 或图 11 所示的地址转换表以及图 12 中的平均擦除次数管理表。因此，存储控制器 SC 在每个闪存存储器模块的每个预定区域上存储与每个模块的地址转换表以及平均擦除次数管理表有关的数据。

在图 9 中，当事件出现时（例如每当任何平均读写组（WDEV）的总写入量达到预定值时），或者以每个预定的时间段，存储控制器 SC 执行平均读写流程。此时，存储控制器 SC 把平均读写组中的闪存存储器模块的移动标志设置为“0”（S901）。

接下来，存储控制器 SC 通过查寻图 12 的平均擦除次数管理表而检验移动标志设置为“0”，并检验平均擦除次数的最大值和最小值（S902）。

存储控制器 SC 判断平均擦除次数的最大值和最小值之间的差是否不小于预定值（S903）。如果擦除次数的差不小于预定值，则存储控制器 SC 进行到 S904。如果擦除次数的差小于预定值，则存储控制器 SC 完成该流程。

然后，存储控制器 SC 从图 12 的管理表中，在具有最大平均擦除次数的闪存存储器模块（PDEV）中选择具有最大总写入量的逻辑页面地址区域；并且在具有最小平均擦除次数的闪存存储器模块（PDEV）中选择具有最小总写入量的逻辑页面地址（S904）。

接下来，存储控制器 SC 把图 10 的虚拟页面与逻辑页面之间的地址转换表的状态字段设置为“交换”。具体地说，存储控制器 SC 把表示“交换”的值输入到图 10 的转换表的状态字段中，以便在两个所选择的逻辑页面地址区域之间执行数据交换并且改变与相应的虚拟页面地址的映射（S905）。在访问其上指示了“交换”的存储区域中，存储控制器 SC 暂时地停留在等待状态，并且在数据交换操作和映射改变操作完成之后，再次尝试访问该存储区域。在这个流程期间，将从主机系统写入的数据存储在存储控制器 SC 中的缓冲

存储器上。

接下来，存储控制器 SC 在上述两个逻辑页面地址区域之间交换数据（S906）。随后将给出关于数据交换流程的详细说明。

数据交换流程之后，存储控制器 SC 对数据交换目标区域所属于的闪存存储器模块（PDEV）的字段（其中记录了前一平均擦除次数值）中的当前平均擦除次数值进行更新，并且将总写入量清零（S907）。因此，紧接在数据交换操作之后的平均擦除次数值与前一平均擦除次数值相同。

如图 11 所示，存储控制器 SC 改变虚拟页面地址与逻辑页面地址之间的映射和偏移量值，并且清除状态字段，然后把移动标志设置为“1”（S908）。在图 12 中，要成为平均读写的目标的平均读写组 W00 包括两个闪存存储器模块 P00、P04。因此，如果在闪存存储器模块 P00、P04 之间执行了任何数据交换，则将闪存存储器模块 P00、P04 的移动标志都设置为“1”。

存储控制器 SC 判断是否存在移动标志为“0”的多个闪存存储器模块（PDEV)（S909）。如果判断没有移动标志为“0”的多个闪存存储器模块，则存储控制器 SC 完成平均读写流程。如果判断存在移动标志为“0”任意多个闪存存储器模块，则存储控制器 SC 返回到 S902。在图 12 中，如果闪存存储器模块 P00、P04 的移动标志都设置为“1”，则存储控制器 SC 完成平均读写流程。例如，如果平均读写组包括四个或更多闪存存储器模块，则存储控制器 SC 在 S902 更进一步检验是否能够在其余的闪存存储器模块之间执行数据交换。

每次执行平均读写流程，都改变虚拟页面地址与逻辑页面地址之间的映射，并且同样更新平均擦除次数。因此，每次流程平均读写流程都需要对存储在闪存存储器模块的预定区域上的地址管理表（图 10 或图 11）与图 12 的平均擦除次数管理表进行更新。

现在，如下描述 S906 的数据交换流程。

图 13 是为了说明平均读写流程伴随的数据交换流程之前的虚拟页面地址与逻辑页面地址之间的映射的框图。为了提供例子，此处要说明如何交换数据以及如何改变虚拟页面地址区域的数据区域 1301 与虚拟页面地址区域的数据区域 1302 之间的映射。虚拟页面地址区域的数据区域 1301 与逻辑页

面地址区域的数据区域 1303 相对应。虚拟页面地址区域的数据区域 1302 与逻辑页面地址区域的数据区域 1304 相对应。例如，假定在逻辑页面地址区域上，在数据区域之间存在任何自由区域。如果自由区域位于相关的数据区域之后，则把偏移量值设置为“0”；以及如果自由区域位于相关的数据区域之前，则把偏移量值设置为“1”。例如，在图 13 中，因为数据区域 1303 具有在其之前的自由区域（虚线所示），从而其偏移量设置为“1”；并且数据区域 1304 具有在其之后的自由区域（虚线所示），从而其偏移量设置为“0”。因此，设置了逻辑页面地址区域的总存储量大于虚拟页面地址区域的总存储量。

图 14 是用于说明执行了平均读写流程的伴随的数据交换之后，虚拟页面地址和逻辑页面地址之间的映射的框图。虚拟页面地址区域的数据区域 1401 与逻辑页面地址区域的数据区域 1404 相对应；并且虚拟页面地址区域的数据区域 1402 与逻辑页面地址区域的数据区域 1403 相对应。数据区域 1403 的偏移量值是“0”，且数据区域 1404 的偏移量值是“0”。

为了提供数据交换流程的例子，图 15 到图 24 以一步接一步的基础，显示了如何在偏移量值为“0”的数据区域与偏移量值为“1”的数据区域之间交换数据。在图 15 到图 24 中，在左边是偏移量值为“0”的逻辑页面地址区域。如图所示，逻辑页面地址区域被分成五段，如分别以“E”、“F”、“G”、“H”和“-”所示。“E”到“H”表示其上写入了有效数据的区域，而“-”表示自由区域。

在图 15 到图 24 中，在右边是偏移量值为“1”的逻辑页面地址区域。逻辑页面地址区域被分成五段，如分别以“A”、“B”、“C”、“D”和“-”所示。“A”到“D”表示其上写入了有效数据的区域，而“-”表示自由区域。

图 15 显示了数据交换流程之前的初始状态。在左边偏移量值是“0”，而在右边偏移量值是“1”。

图 16 显示了数据交换流程期间的状态，把左边（偏移量值为“0”）的逻辑页面地址区域 E 上的数据重写到右边（偏移量值为“1”）的自由区域上。

图 17 显示了数据交换流程期间的状态，把右边的逻辑页面地址区域 A 上的数据重写到左边的原逻辑页面地址区域 E 上。

图 18 显示了数据交换流程期间的状态，把左边的逻辑页面地址区域 F 上

的数据重写到右边的原逻辑页面地址区域 A 上。

图 19 显示了数据交换流程期间的状态，把右边的逻辑页面地址区域 B 上的数据重写到左边的原逻辑页面地址区域 F 上。

图 20 显示了数据交换流程期间的状态，把左边的逻辑页面地址区域 G 上的数据重写到右边的原逻辑页面地址区域 B 上。

图 21 显示了数据交换流程期间的状态，把右边的逻辑页面地址区域 C 上的数据重写到左边的原逻辑页面地址区域 G 上。

图 22 显示了数据交换流程期间的状态，把左边的逻辑页面地址区域 H 上的数据重写到右边的原逻辑页面地址区域 C 上。

图 23 显示了数据交换流程期间的状态，把右边的逻辑页面地址区域 D 上的数据重写到左边的原逻辑页面地址区域 H 上。

图 24 显示了数据交换流程之后的最终状态。偏移量值在左边是“0”，且偏移量值在右边也是“0”。

基本上，闪存存储器是这样的一种半导体器件，其中不能在物理地址区域上执行重写流程。具体地说，为了交换物理地址区域上的数据，实际上将数据拷贝到未使用页面上，然后把存储了数据的原始页面设置成无效页面。因此，在这个原始页面上没有执行实际的重写过程。

根据本实施例，上述流程全部在逻辑页面地址区域基础上执行，因此，能够在逻辑页面上重写数据。用这种方式，通过执行这种重写流程可以基于数据交换来执行平均读写流程。

图 25 是显示数据交换流程之前/之后偏移量值如何转变的表。

如果在偏移量值都为“0”的两个逻辑页面地址区域之间执行数据交换，则数据交换流程之后偏移量值分别变为“0”和“1”。如果在偏移量值为“0”的逻辑页面地址区域与偏移量值为“1”的逻辑页面地址区域之间执行数据交换，则数据交换流程之后偏移量值分别变为“0”和“0”。如果在偏移量值都为“1”的两个逻辑页面地址区域之间执行数据交换，则数据交换流程之后偏移量值分别变为“1”和“0”。

将参考流程图，给出关于数据交换流程的详细说明。

图 26 是显示偏移量值为“0”的逻辑页面地址区域与偏移量值为“1”的

逻辑页面地址区域之间的数据交换流程（如图 15 到图 24 所述）的流程图。此处，存储控制器 SC 在偏移量值为“0”的逻辑页面地址区域与偏移量值为“1”的逻辑页面地址区域之间设置数据交换目标（S2601）。

存储控制器 SC 把目标逻辑页面地址区域分成“n”段；并设置为“i=1”（S2602）。就图 15 而言，例如，设置为“n=5”，并且把有效数据写入到所分割的（n-1）个段上，剩余的一个段用作自由区域。存储控制器 SC 把数据从偏移量值为“0”的第 i 个逻辑页面地址区域中移动到偏移量值为“1”的第 i 个逻辑页面地址区域中（S2603），然后同样地把数据从偏移量值为“1”的第 i+1 个逻辑页面地址区域中移动到偏移量值为“0”的第 i 个逻辑页面地址区域中（S2604）；并且使“i”递增“1”（S2605）。然后，存储控制器 SC 判断是否“i=n”（S2606）。如果判断不是“i=n”，那么存储控制器 SC 返回到 S2603。如果判断为“i=n”，那么存储控制器 SC 完成数据交换流程。

图 27 是显示偏移量值为“0”的逻辑页面地址区域与偏移量值为“0”的逻辑页面地址区域之间的数据交换流程的流程图。此处，存储控制器 SC 在偏移量值为“0”的逻辑页面地址区域与偏移量值为“0”的逻辑页面地址区域之间设置数据交换目标（S2701）。

存储控制器 SC 把目标逻辑页面地址区域分成“n”段，并设置为“i=n”（S2702）。有效数据写入到所分割的（n-1）个段上，剩余的一个段是自由区域。直到到达“i=1”之前，以逐段的基础重复数据交换流程（S2703 到 S2706）。在 S2706，如果判断为“i=1”，则完成数据交换流程。

图 28 是显示偏移量值为“1”的逻辑页面地址区域与偏移量值为“1”的逻辑页面地址区域之间的数据交换流程的流程图。此处，存储控制器 SC 在偏移量值为“1”的逻辑页面地址区域与偏移量值为“1”的逻辑页面地址区域之间设置数据交换目标（S2801）。

存储控制器 SC 把目标逻辑页面地址区域分成“n”段；并设置为“i=2”（S2802）。有效数据写入到所分割的（n-1）个段上，剩余的一个段是自由区域。直到变成“i>n”之前，以逐段的基础重复数据交换流程（S2803 到 S2806）。在 S2806，如果判断为“i>n”，则完成数据交换流程。

图 29 到图 34 是说明根据本发明另一个实施例的平均读写流程的附图。

根据上述实施例，在闪存存储器模块中分布有用于数据交换的自由区域。在这个实施例中，将说明把自由区域处理为每个模块中的一个组的方法。

图 29 是用于说明执行数据交换流程之前的虚拟页面地址和逻辑页面地址之间的映射的框图。

图 30 是用于说明执行数据交换流程之后的虚拟页面地址与逻辑页面地址之间的映射的框图。为了简便，假定目标平均读写组 W00 具有两个闪存存储器模块 P00、P04。在数据交换之前的状态（图 29）中，闪存存储器模块 P00、P04 的逻辑页面地址区域具有从地址 AC0 到地址 AC4 之前的数据区域。地址 AC4 的区域或更多区域用作用于数据交换的自由区域（2903、2904）。这个自由区域的大小（数据长度）与用于数据交换的数据区域的大小相同以执行平均读写流程。

图 31 是用于说明执行数据交换流程之前的用于虚拟页面地址与逻辑页面地址之间的数据交换的地址转换表的一个表格。图 32 是用于说明执行数据交换流程之后的用于虚拟页面地址与逻辑页面地址之间的数据交换的地址转换表的一个表格。在这个实施例中，由于把用于数据交换的自由区域处理为一个组，从而图 10 和图 11 的地址转换表所需要的偏移量值管理就不必要了。代之以需要管理自由区域位置。

图 33 是用于说明执行数据交换流程之前的自由区域管理表的一个表格，而图 34 是用于说明执行数据交换流程之后的自由区域管理表的一个表格。自由区域管理表管理每个闪存存储器模块中的自由区域的起始逻辑页面地址和大小（数据长度）。

将参考图 29，给出如何在虚拟页面地址区域的数据区域 2901 和数据区域 2902 之间交换数据以及如何改变虚拟页面地址与逻辑页面地址之间的映射的说明。如图 31 中的逻辑页面地址与虚拟页面地址之间的地址转换表所示，可以理解虚拟页面地址区域中的数据区域 2901 与逻辑页面地址区域中的数据区域 2905 相对应；而虚拟页面地址区域中的数据区域 2902 与逻辑页面地址区域中的数据区域 2906 相对应。参考图 33 中的自由区域管理表，可以理解闪存存储器模块 P00 中用于数据交换的自由区域是区域 2903；而闪存存储器模块 P04 中用于数据交换的自由区域是区域 2904。

接下来，将数据区域 2905 上的数据写入到自由区域 2904 上，将数据区域 2906 上的数据写入到自由区域 2903 上。如图 30 所示，将虚拟页面地址区域中的数据区域 3001 设置为与逻辑页面地址区域中的数据区域 3004 相对应；并将虚拟页面地址区域中的数据区域 3002 设置为与逻辑页面地址区域中的数据区域 3003 相对应。完成上述数据交换流程之后，更新虚拟页面地址与逻辑页面地址之间的转换表，如图 32 所示。如图 34 所示，可以理解闪存存储器模块 P00 中用于数据交换的自由区域是区域 3005；而闪存存储器模块 P04 中用于数据交换的自由区域是区域 3006。

根据本实施例，将用于数据交换的自由区域处理为每个模块一个组，而不是象其他实施例所述的那样在闪存存储器模块中分配用于数据交换的自由区域，借此消除了对偏移量值的管理，结果产生了较容易的数据交换控制。

接下来，此处将对在闪存存储器模块（PDEV）上出现故障的情况给出说明。

参考图 35 到图 39，给出了对当闪存存储器模块上出现故障时如何替换闪存存储器模块的方法的说明。

图 35 是显示如何替换模块的步骤的流程图。

图 36 到图 39 是说明图 35 的流程图的每个步骤的框图。

图 36 显示了当闪存存储器模块上出现故障时的情况。图 36 显示了 RAID 组（VDEV）V00，以及构成 RAID 组 V00 的平均读写组（WDEV）W00、W10、W20、W30。备份组（YDEV）Y00 连接于连接了平均读写组 W00 的同一通道 D01 上。

现在，假定在平均读写组 W00 中的闪存存储器模块（PDEV）P01 上出现故障（S3501）。

然后，选择其中平均读写（WDEV）W00 可用的备份组（YDEV）。选择了连接在平均读写组 W00 的相同通道 D01 上的备份组 Y00（S3502）。然后，从属于该备份组 Y00 的闪存存储器模块中，选择了闪存存储器模块 P04 用于替换闪存存储器模块 P01（S3503）。

图 37 是用于说明闪存存储器模块替换之后的状态的框图。如图 37 所示，把闪存存储器模块 P01 替换为平均读写组 W00 与备份组 Y00 之间的闪存存

储器模块 P04。在替换中，故障的模块 P01 停留在待机状态。

接下来，图 38 是显示在闪存存储器替换之后如何重建数据的框图。如图 38 所示，将写在闪存存储器模块 P01 上的数据重建并写入到新合并到平均读写组 W00 中的闪存存储器模块 P04 上（S3504）。要指出的是，此时，由于平均读写，用于数据重建的数据存储并分配在不同平均读写组中的闪存存储器模块之间。换句话说，在同一虚拟页面地址中不同的平均读写组中所存储的数据上进行数据重建。

图 39 是显示其中将备份组中的闪存存储器替换为新的闪存存储器模块的情况的框图。如图 39 所示，把替换中处于待机状态的闪存存储器模块 P01 替换为新的闪存存储器模块 P06，并且将模块 P06 合并到备份组 Y00 中（S3505）。然后，完成闪存存储器模块替换。

根据模块替换之前旧模块的总写入量，判断在模块替换之后是否能够立即执行平均读写流程。要指出的是，除了伴随着数据重建的一些写入之外，新替换的模块在相关的逻辑页面地址区域的整个区域上没有总写入量，因此，不可能知道逻辑页面地址区域的每个预定区域的写入频率。可以使用模块替换之前旧模块的总写入量来获知逻辑页面地址区域的写入频率，以便执行平均读写流程。

本发明提供了一种在多个闪存存储器模块之间提供平均读写的方法，其适用于提高闪存存储器模块耐用性的目的，具体地说适用于利用具有多个闪存存储器模块的大容量闪存存储器的存储系统；一种为此的平均读写方法；以及用于运行上述方法的平均读写程序。

已经如上所述说明了根据本发明的实施例。然而，本发明的实施例不局限于那些说明，本领域技术人员可确定本发明的基本特征并且在不脱离权利要求的精神和范围的情况下可以对本发明进行各种改进和变化以使它适应于各种应用和条件。

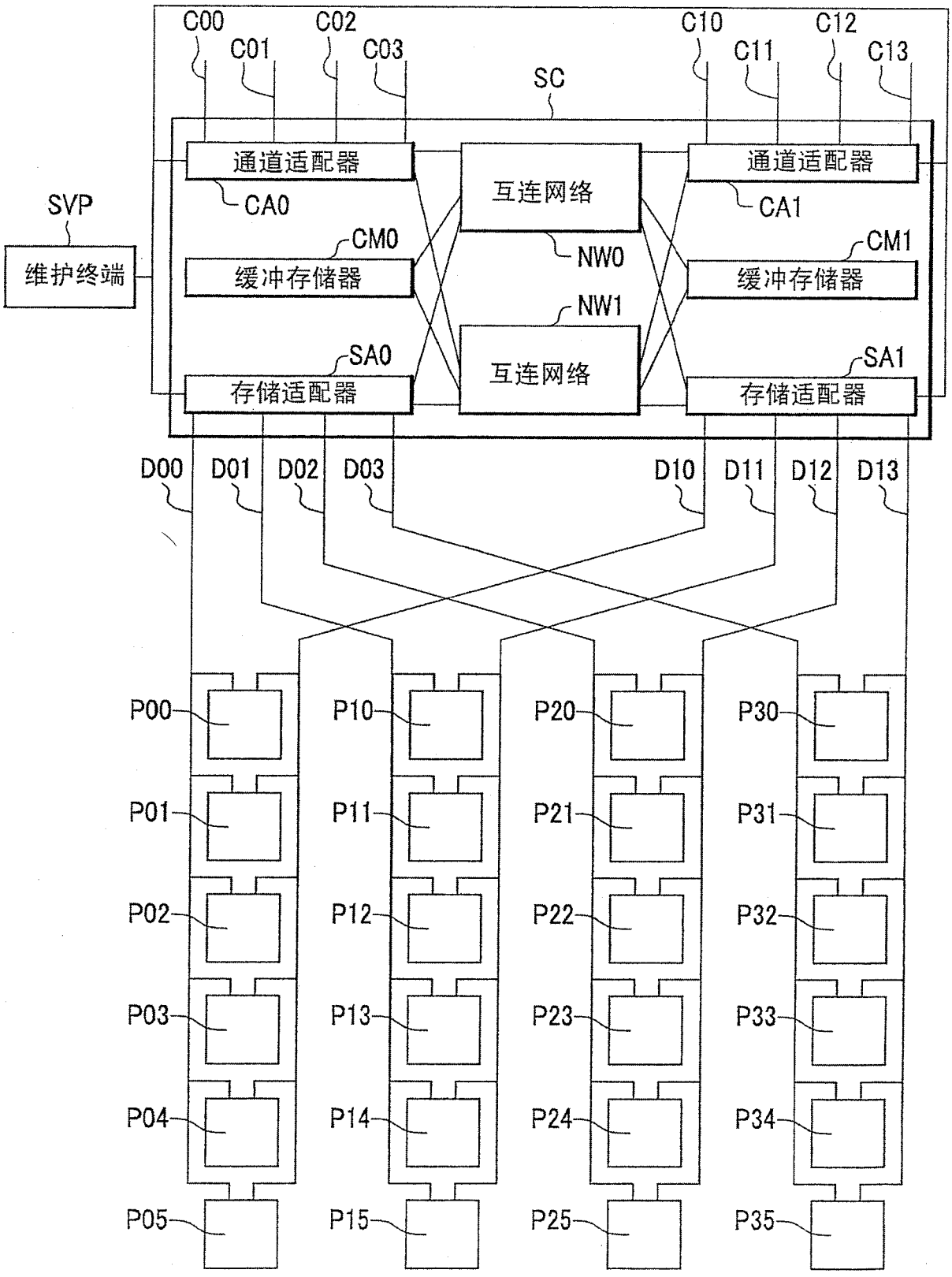

C00
C01
C02
C03
C10
C11
C12
C13
SC
通道适配器
CA0
互连网络
NW0
NW1
通道适配器
CA1
SVP
维护终端
CM0
缓冲存储器
CM1
缓冲存储器
SA0
存储适配器
互连网络
SA1
存储适配器
D00
D01
D02
D03
D10
D11
D12
D13
P00
P10
P20
P30
P01
P11
P21
P31
P02
P12
P22
P32
P03
P13
P23
P33
P04
P14
P24
P34
P05
P15
P25
P35

图 1

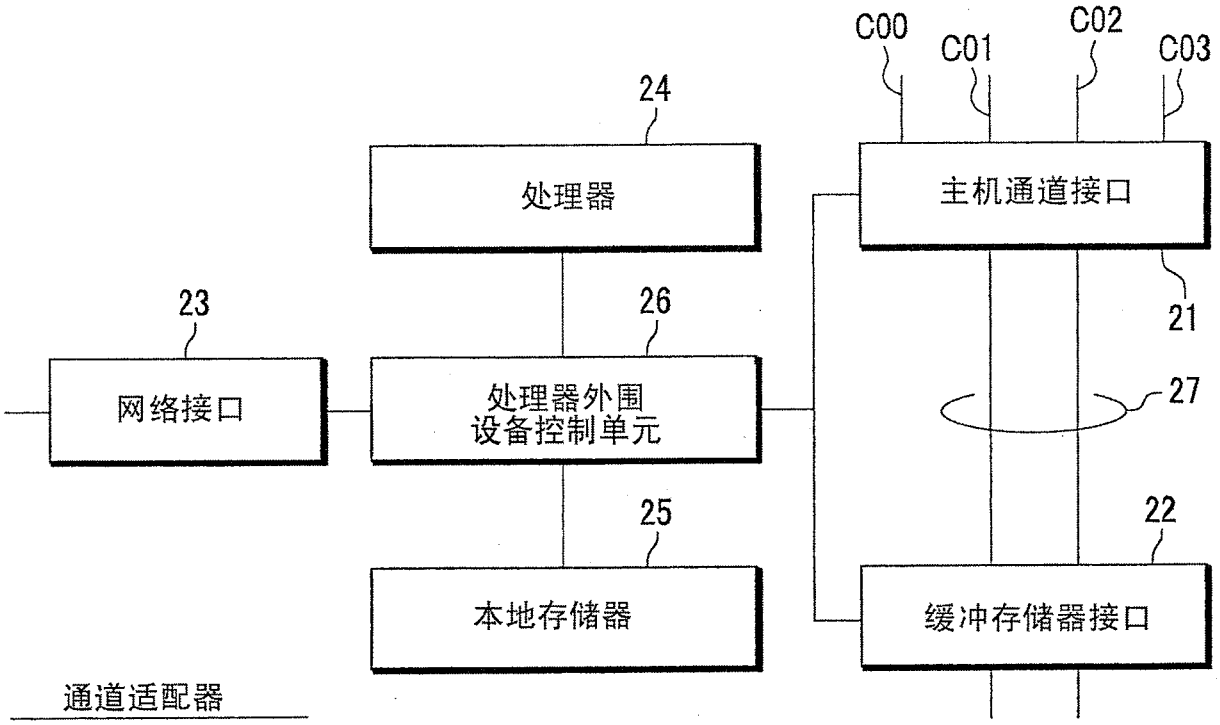

C00
C01
C02
C03
24
处理器
主机通道接口
23
26
21
网络接口
处理器外围
设备控制单元
27
25
22
本地存储器
缓冲存储器接口
通道适配器

图 2

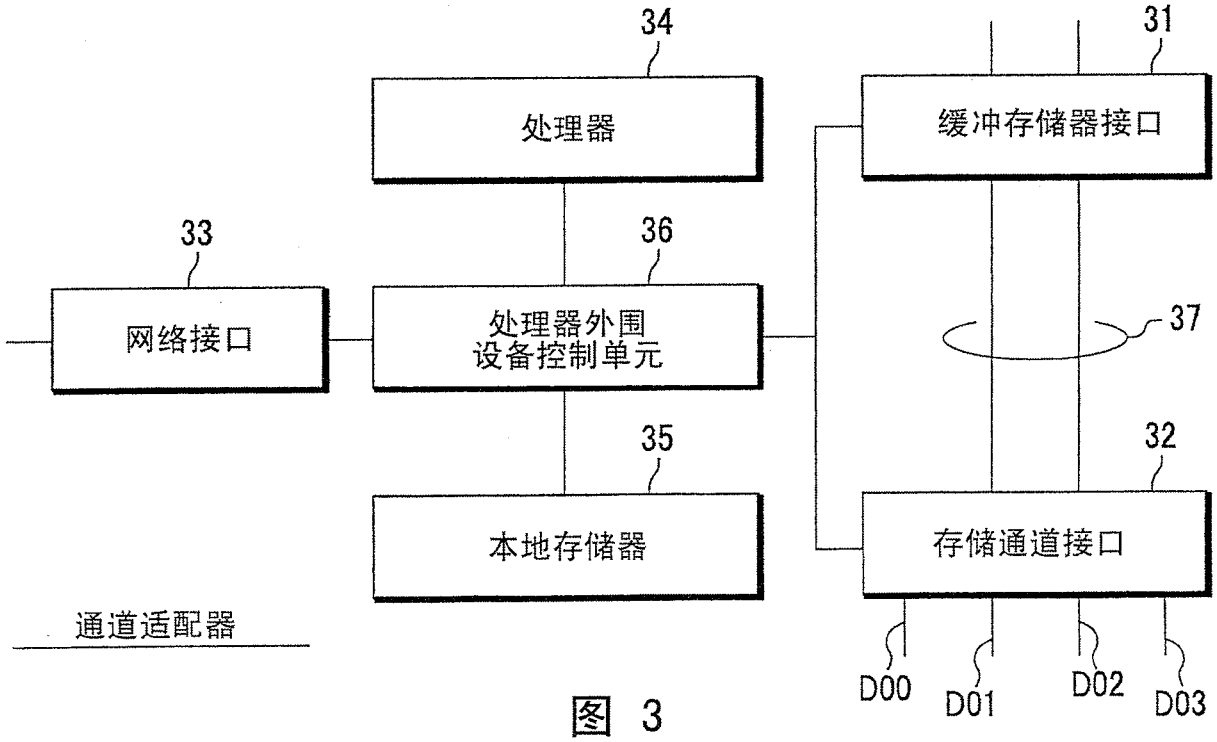

34
31
处理器
缓冲存储器接口
33
36
网络接口
处理器外围
设备控制单元
37
35
32
本地存储器
存储通道接口
通道适配器
D00
D01
D02
D03

图 3

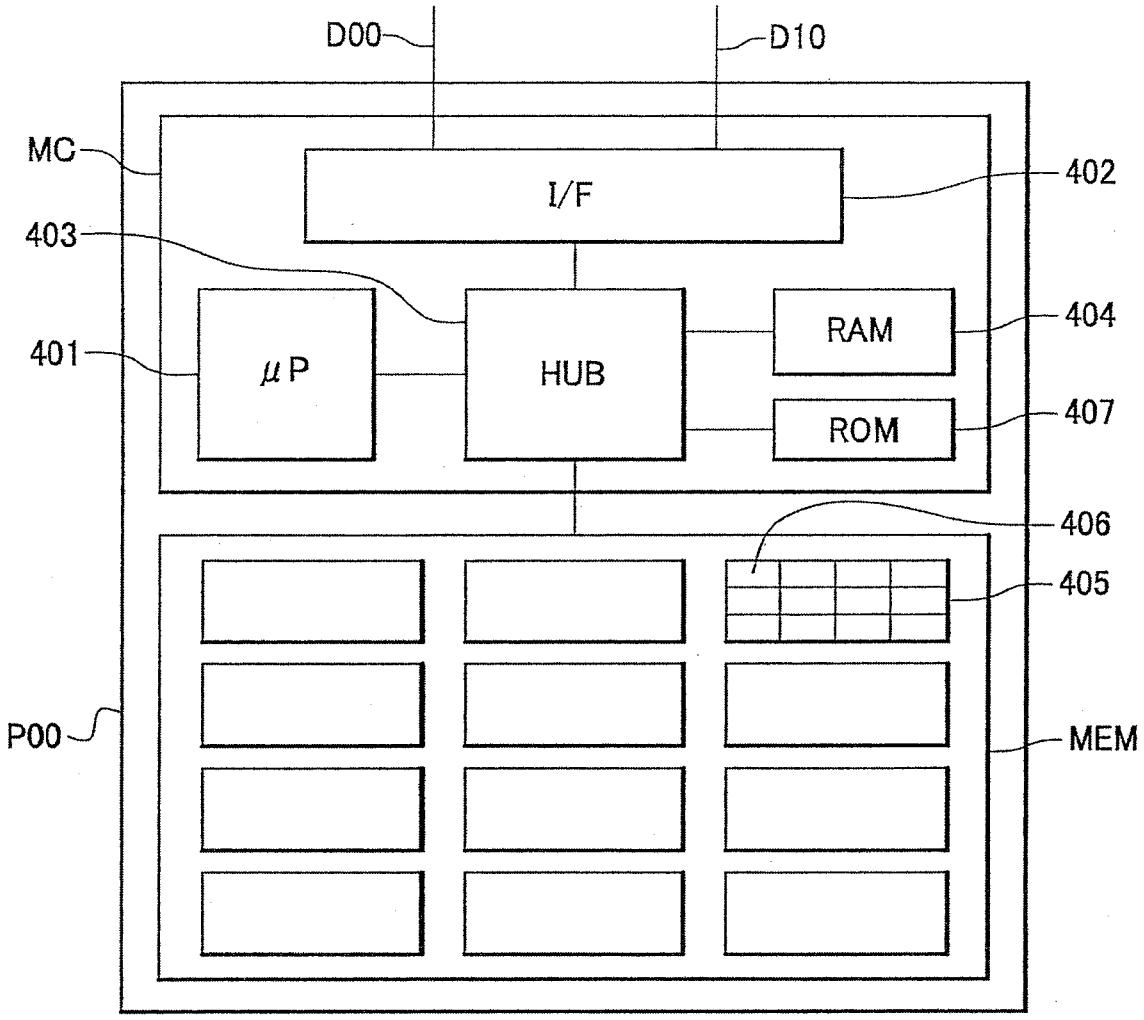

D00
D10
MC
I/F
402
403
401
μP
HUB
RAM
404
ROM
407
406
405
P00
MEM

图 4

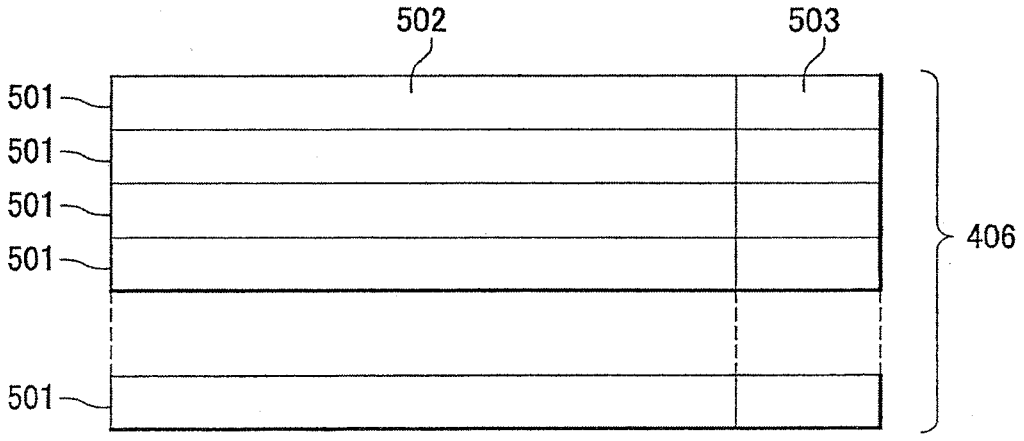

502
503
501
501
501
501
501
406

图 5

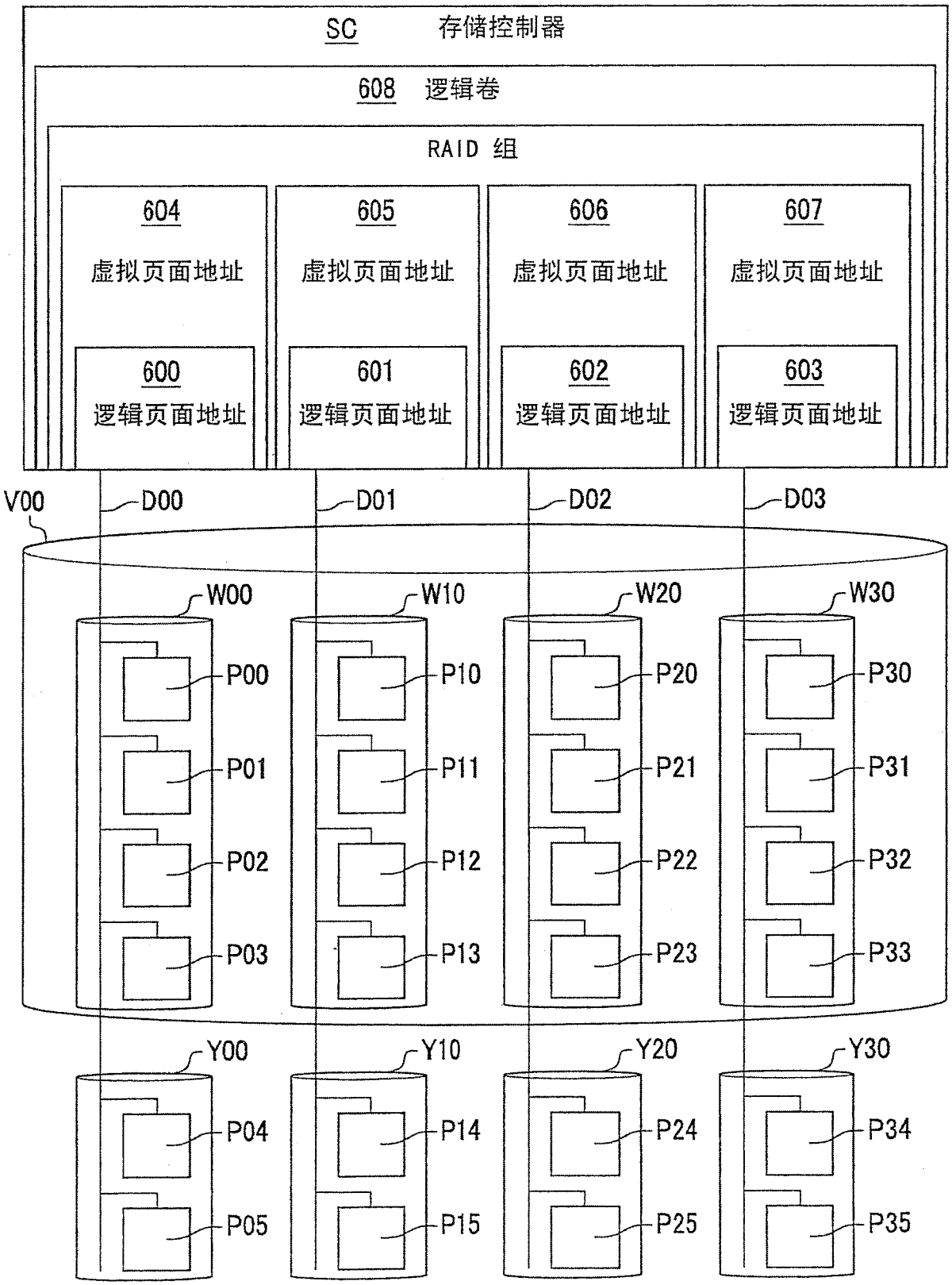
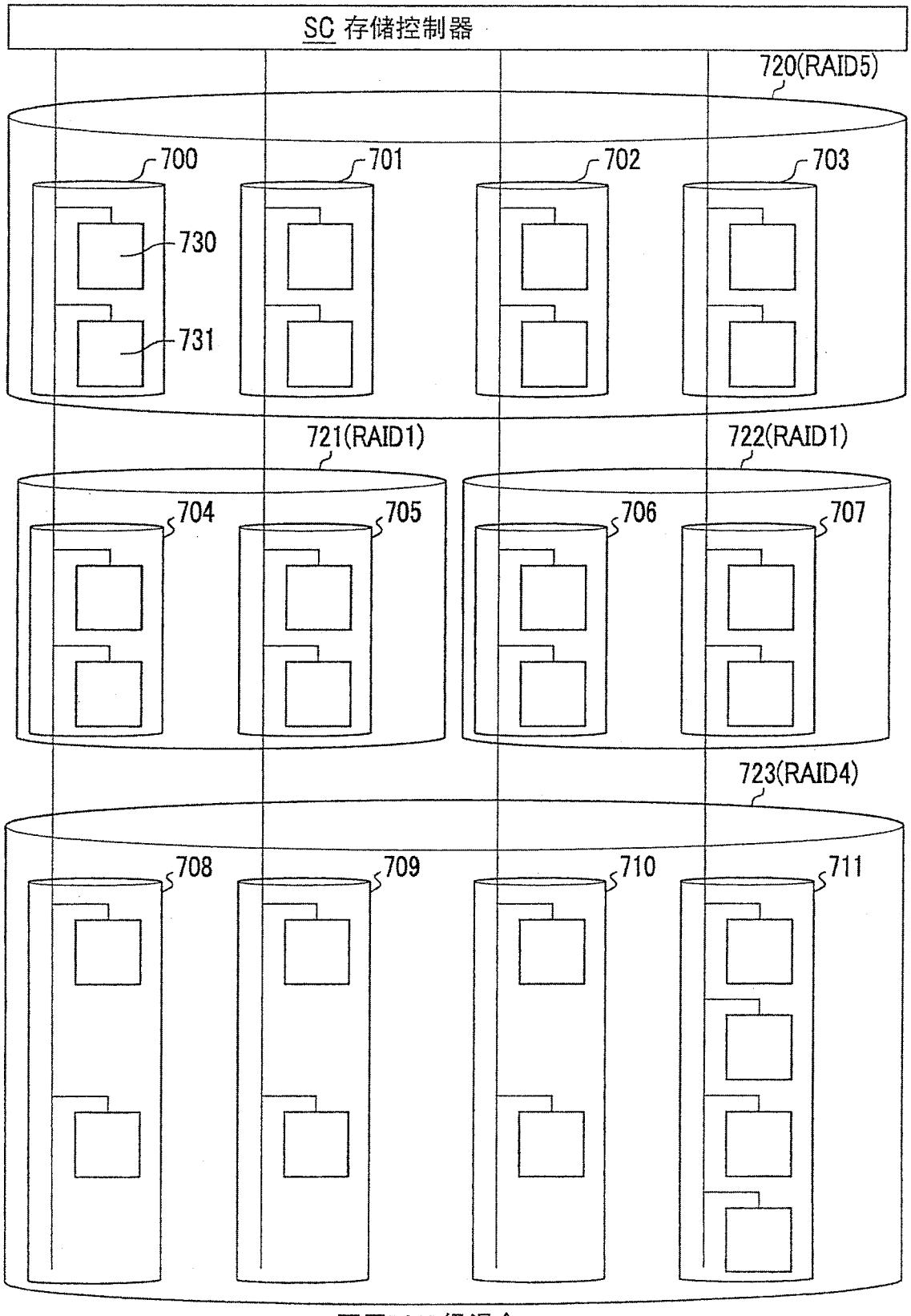

SC 存储控制器
608 逻辑卷
RAID 组
604
虚拟页面地址
605
虚拟页面地址
606
虚拟页面地址
607
虚拟页面地址
600
逻辑页面地址
601
逻辑页面地址
602
逻辑页面地址
603
逻辑页面地址
V00
D00
D01
D02
D03
W00
W10
W20
W30
P00
P01
P02
P03
P10
P11
P12
P13
P20
P21
P22
P23
P30
P31
P32
P33
Y00
Y10
Y20
Y30
P04
P05
P14
P15
P24
P25
P34
P35

图 6



不同RAID级混合 图 7

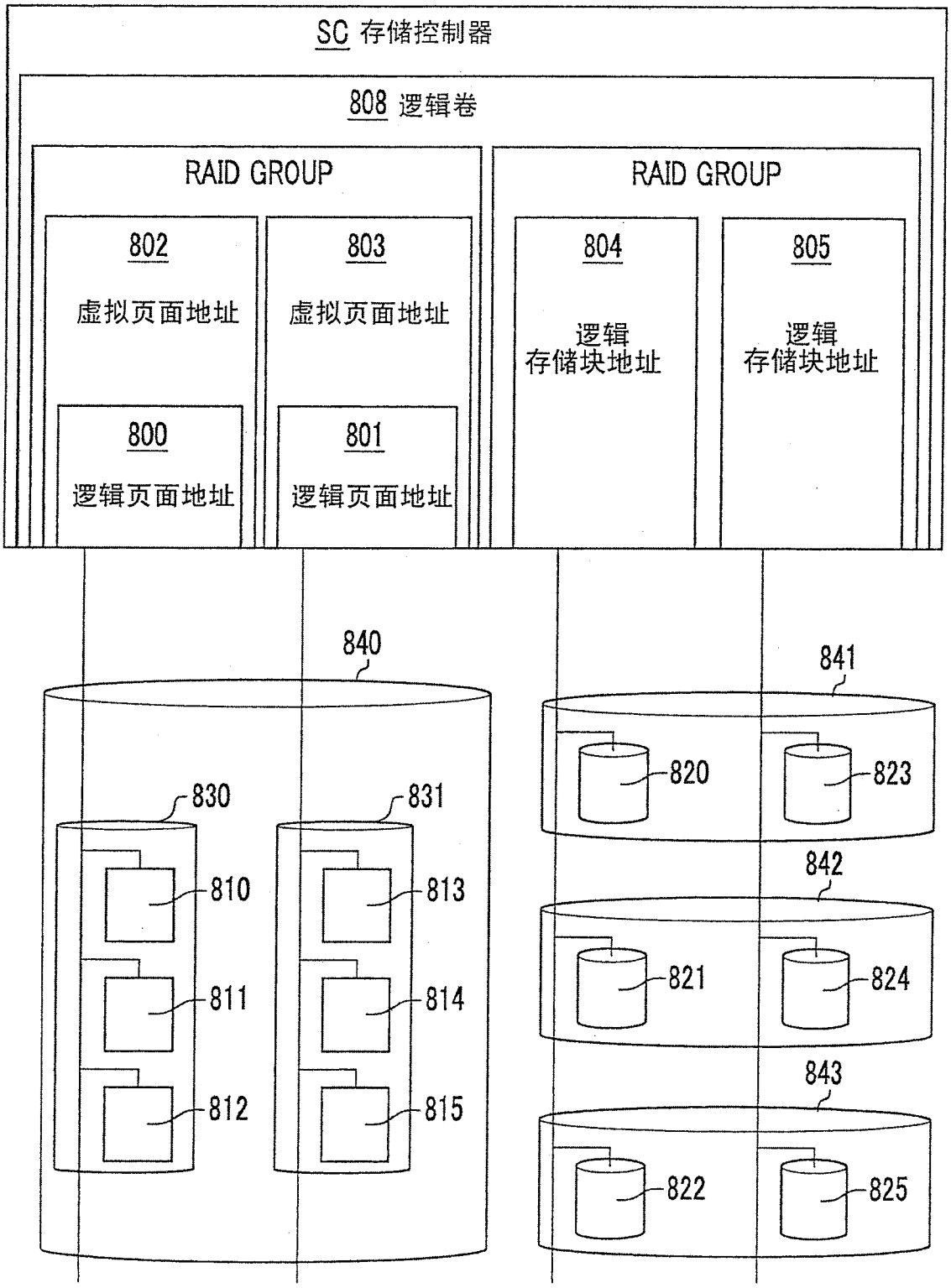

SC 存储控制器
808 逻辑卷
RAID GROUP
RAID GROUP
802
虚拟页面地址
803
虚拟页面地址
804
逻辑
存储块地址
805
逻辑
存储块地址
800
逻辑页面地址
801
逻辑页面地址
840
841
830
831
810
813
811
814
812
815
820
823
842
821
824
843
822
825

图 8

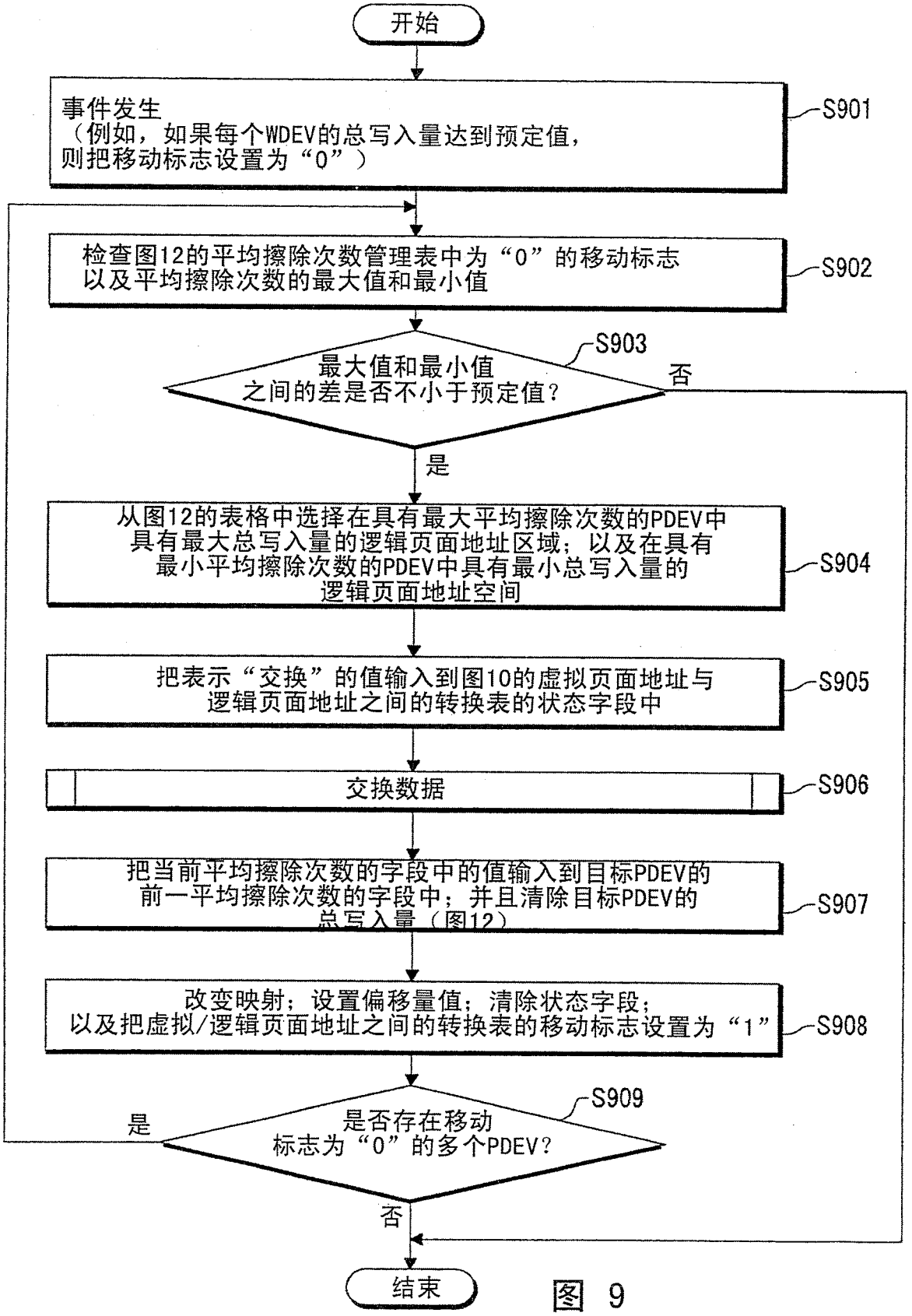

开始
事件发生
（例如，如果每个WDEV的总写入量达到预定值，
则把移动标志设置为“0”）
S901
检查图12的平均擦除次数管理表中为“0”的移动标志
以及平均擦除次数的最大值和最小值
S902
S903
最大值和最小值
之间的差是否不小于预定值？
否
是
从图12的表格中选择在具有最大平均擦除次数的PDEV中
具有最大总写入量的逻辑页面地址区域；以及在具有
最小平均擦除次数的PDEV中具有最小总写入量的
逻辑页面地址空间
S904
把表示“交换”的值输入到图10的虚拟页面地址与
逻辑页面地址之间的转换表的状态字段中
S905
交换数据
S906
把当前平均擦除次数的字段中的值输入到目标PDEV的
前一平均擦除次数的字段中；并且清除目标PDEV的
总写入量（图12）
S907
改变映射；设置偏移量值；清除状态字段；
以及把虚拟/逻辑页面地址之间的转换表的移动标志设置为“1”
S908
S909
是否存在移动
标志为“0”的多个PDEV？
是
否
结束

图 9

WDEV	虚拟页面地址	数据长度	PDEV	逻辑页面地址	偏移量	状态
W00	AA0		P00	AB0	0	
	AA1		P00	AB1	1	1
	AA2		P00	AB2	0	
	AA3		P00	AB3	0	
	AA4		P04	AB0	1	
	AA5		P04	AB1	1	
	AA6		P04	AB2	0	1
	AA7		P04	AB3	1	

图 10

WDEV	虚拟页面地址	数据长度	PDEV	逻辑页面地址	偏移量	状态
W00	AA0		P00	AB0	0	
	AA1		P04	AB2	0	
	AA2		P00	AB2	0	
	AA3		P00	AB3	0	
	AA4		P04	AB0	1	
	AA5		P04	AB1	1	
	AA6		P00	AB1	0	
	AA7		P04	AB3	1	

图 11

PDEV	PDEV容量	逻辑页面地址	总写入量	平均擦除次数	平均擦除次数（前一）	移动标志
P00	s00	AB0	a000	e00=f00+(a000+a001+a002+a003)/s00	f00	0
		AB1	a001			
		AB2	a002			
		AB3	a003			
P04	s04	AB0	a040	e04=f04+(a040+a041+a042+a043)/s04	f04	0
		AB1	a041			
		AB2	a042			
		AB3	a043			

图 12

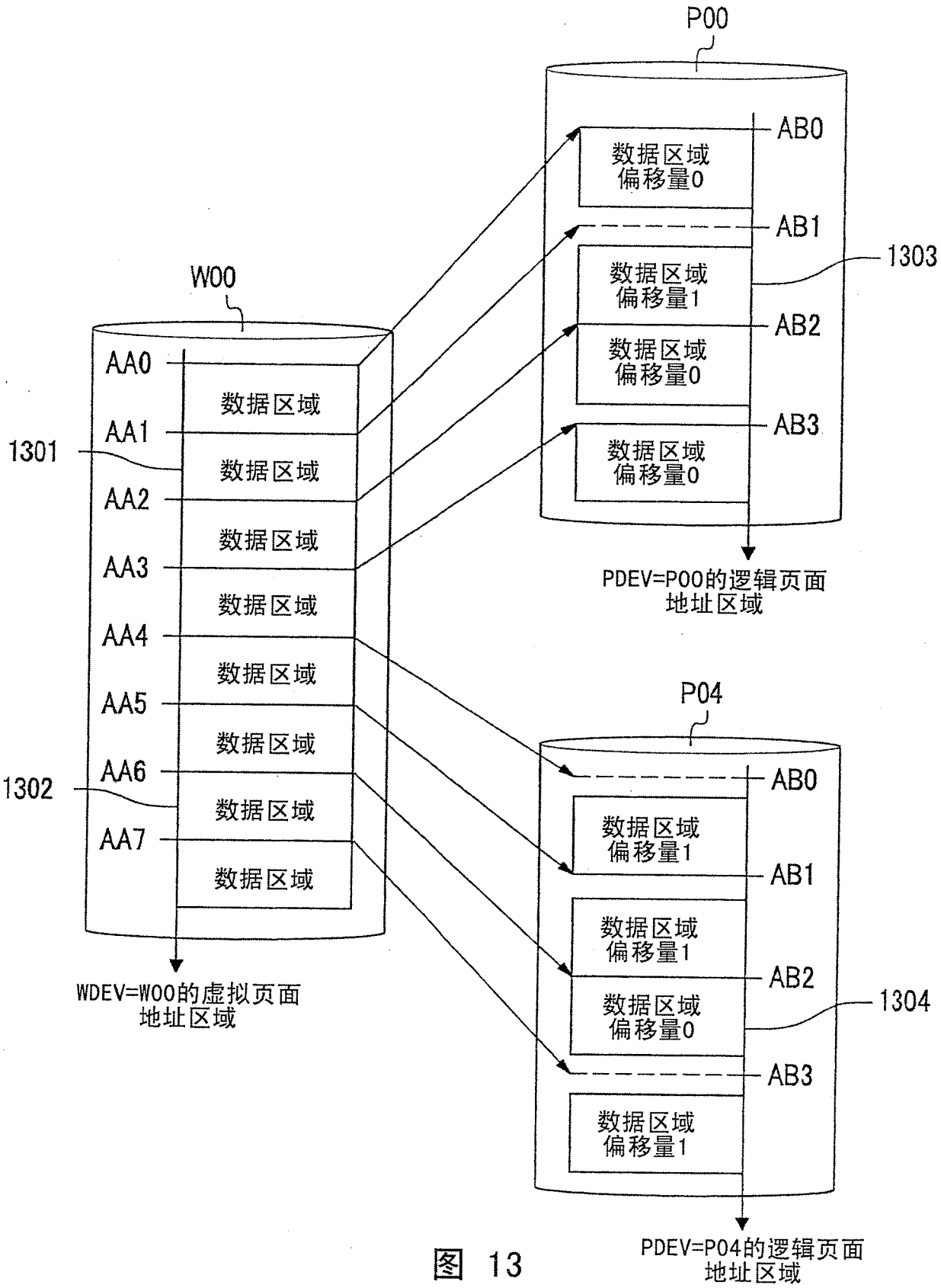

P00
AB0
数据区域
偏移量0
AB1
数据区域
偏移量1
1303
AB2
数据区域
偏移量0
AB3
数据区域
偏移量0
PDEV=P00的逻辑页面
地址区域
W00
AA0
数据区域
AA1
1301
数据区域
AA2
数据区域
AA3
数据区域
AA4
数据区域
AA5
数据区域
AA6
1302
数据区域
AA7
数据区域
WDEV=W00的虚拟页面
地址区域
P04
AB0
数据区域
偏移量1
AB1
数据区域
偏移量1
AB2
数据区域
偏移量0
1304
AB3
数据区域
偏移量1
PDEV=P04的逻辑页面
地址区域

图 13

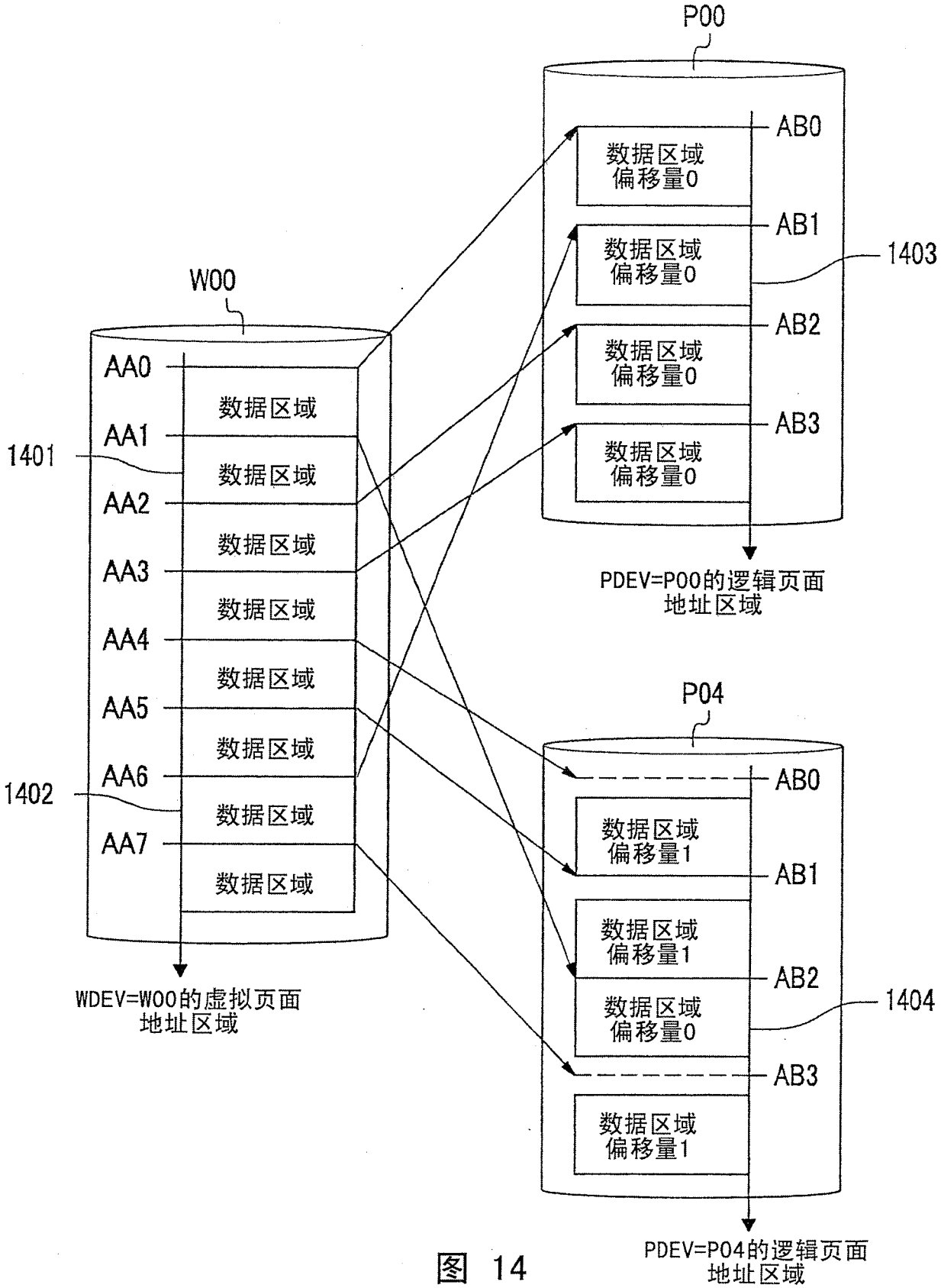

P00
AB0
数据区域
偏移量0
AB1
数据区域
偏移量0
1403
AB2
数据区域
偏移量0
AB3
数据区域
偏移量0
PDEV=P00的逻辑页面
地址区域
W00
AA0
数据区域
AA1
数据区域
1401
AA2
数据区域
AA3
数据区域
AA4
数据区域
AA5
数据区域
AA6
数据区域
1402
AA7
数据区域
WDEV=W00的虚拟页面
地址区域
P04
AB0
数据区域
偏移量1
AB1
数据区域
偏移量1
AB2
数据区域
偏移量0
1404
AB3
数据区域
偏移量1
PDEV=P04的逻辑页面
地址区域

图 14

图 15

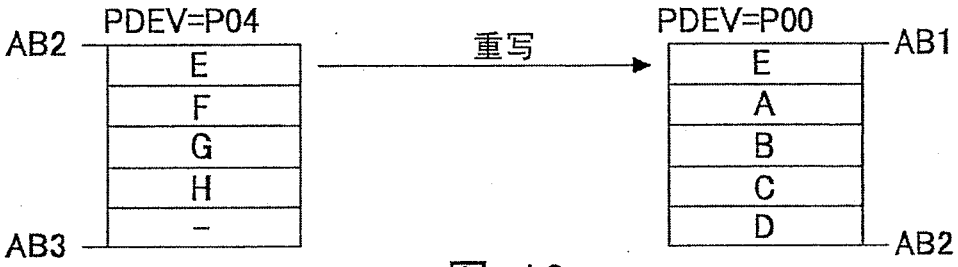


图 16

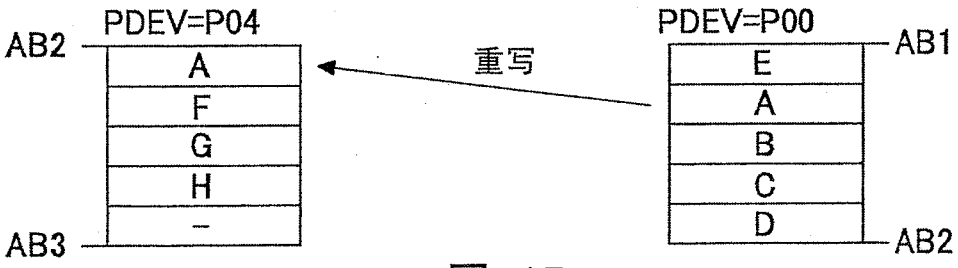


图 17

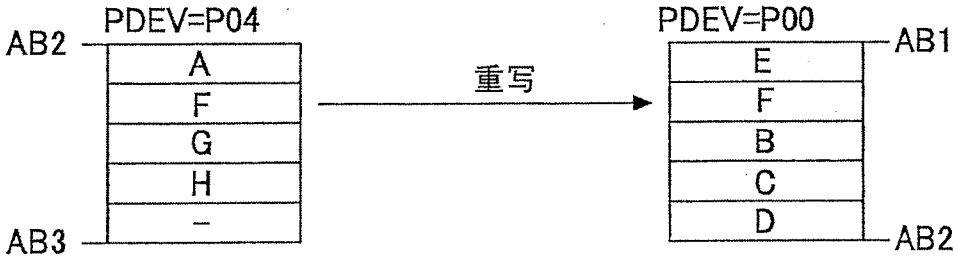


图 18

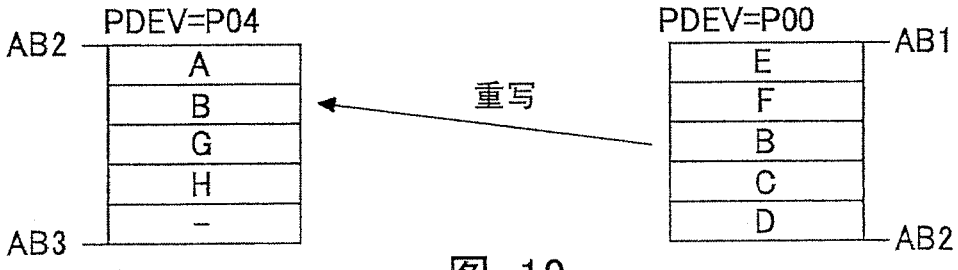


图 19

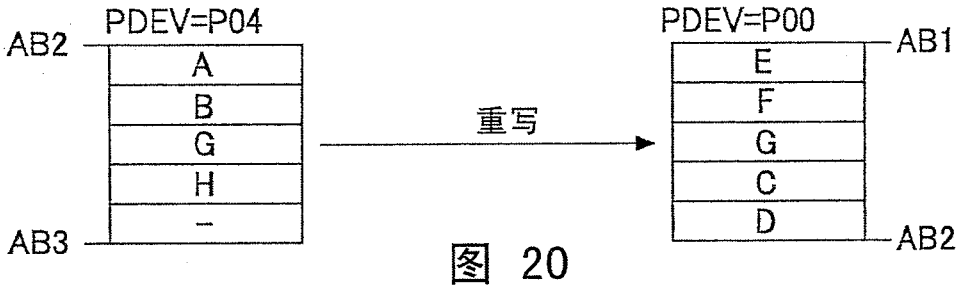


图 20

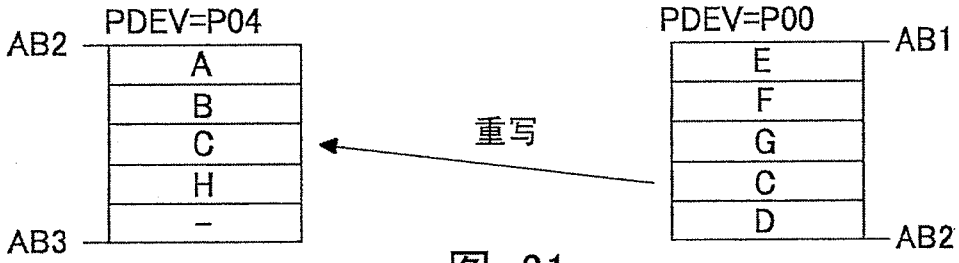


图 21

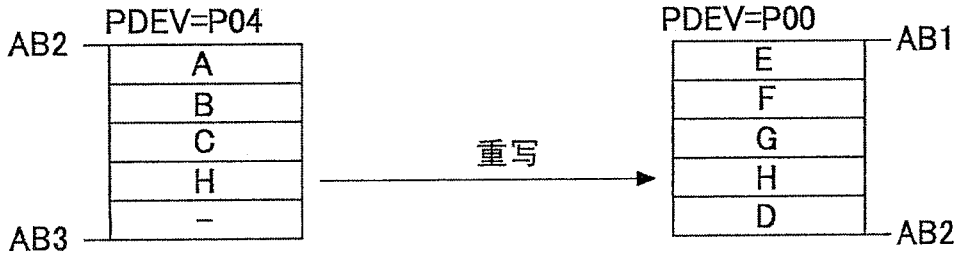


图 22

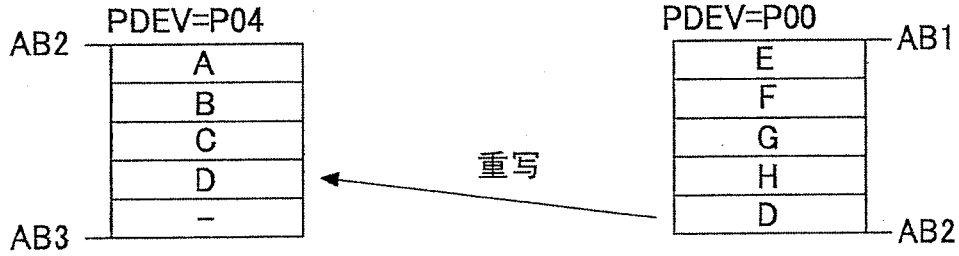


图 23

图 24

数据交换前的偏移量值组合		数据交换后的偏移量值组合	
0	0	0	1
0	1	0	0
1	1	1	0

图 25

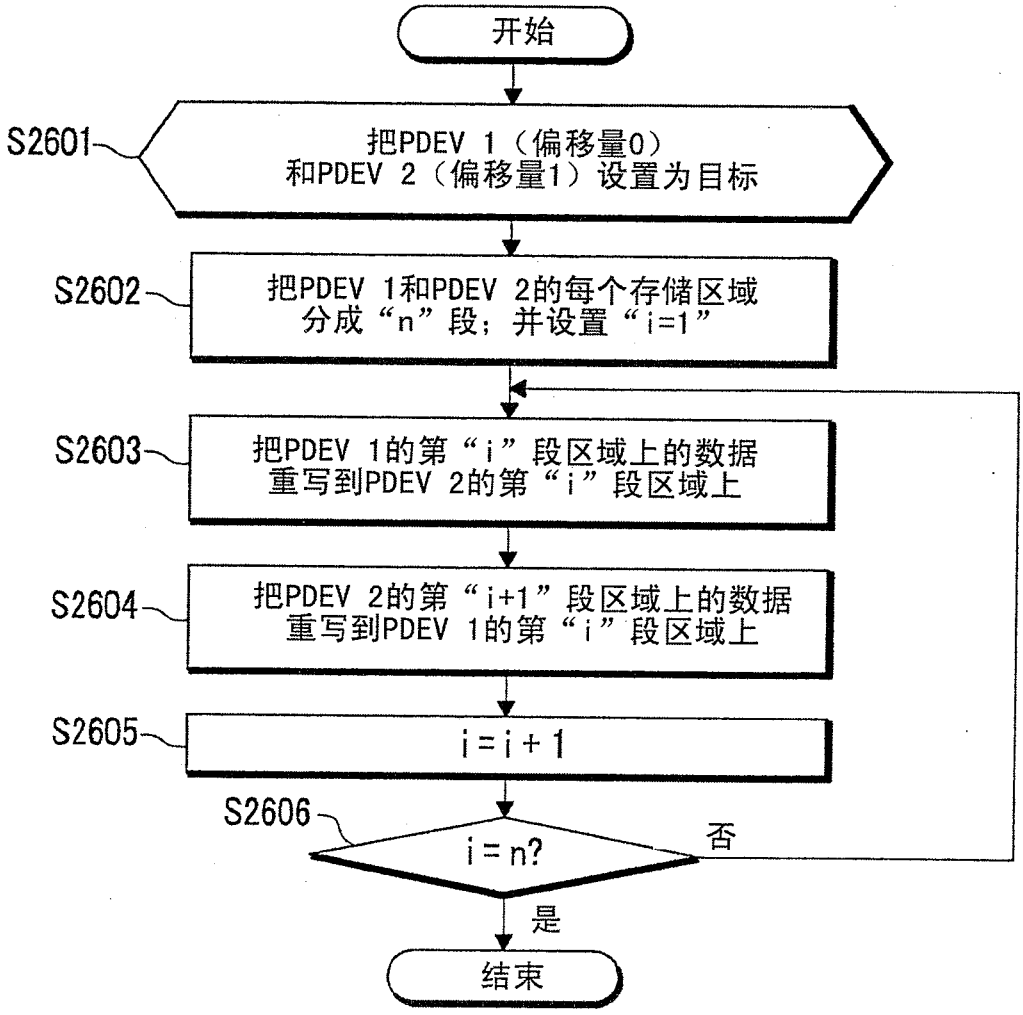


图 26

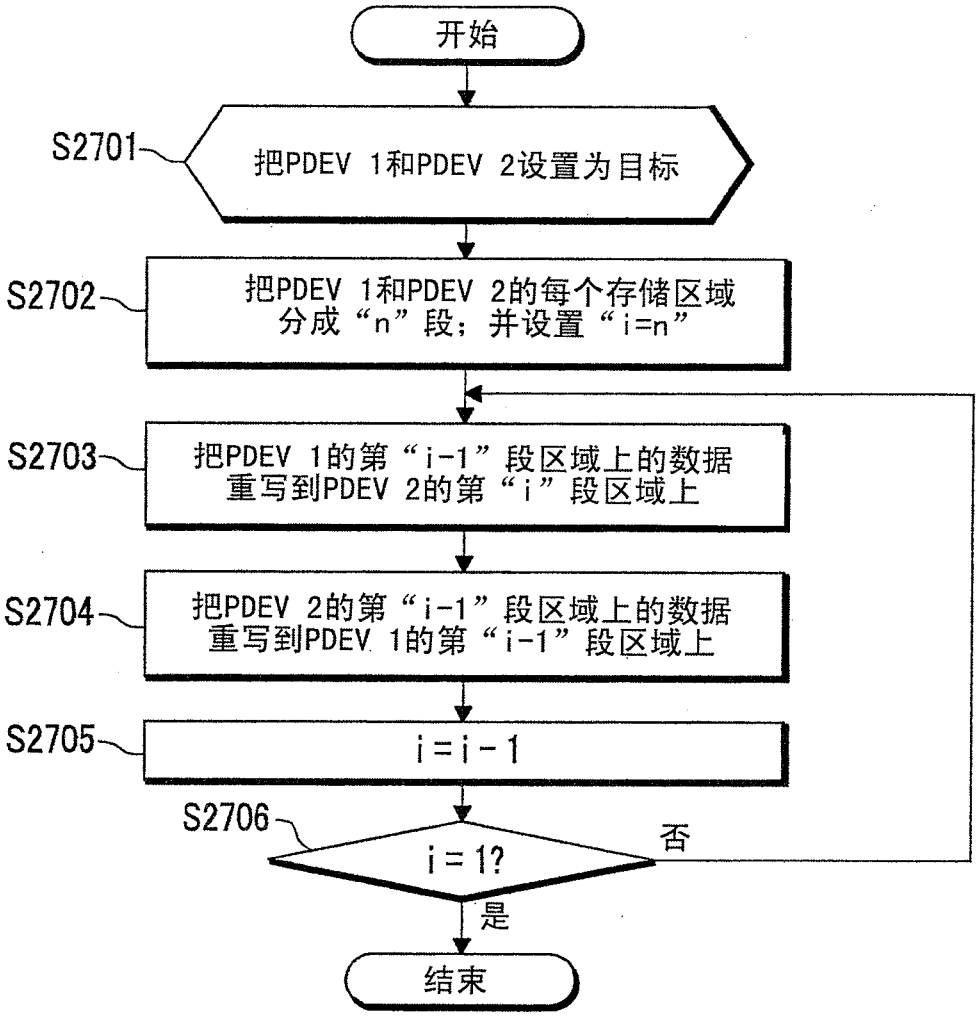

开始
S2701
把PDEV 1和PDEV 2设置为目标
S2702
把PDEV 1和PDEV 2的每个存储区域
分成“n”段；并设置“i=n”
S2703
把PDEV 1的第“i-1”段区域上的数据
重写到PDEV 2的第“i”段区域上
S2704
把PDEV 2的第“i-1”段区域上的数据
重写到PDEV 1的第“i-1”段区域上
S2705
i=i-1
S2706
i=1?
否
是
结束

图 27

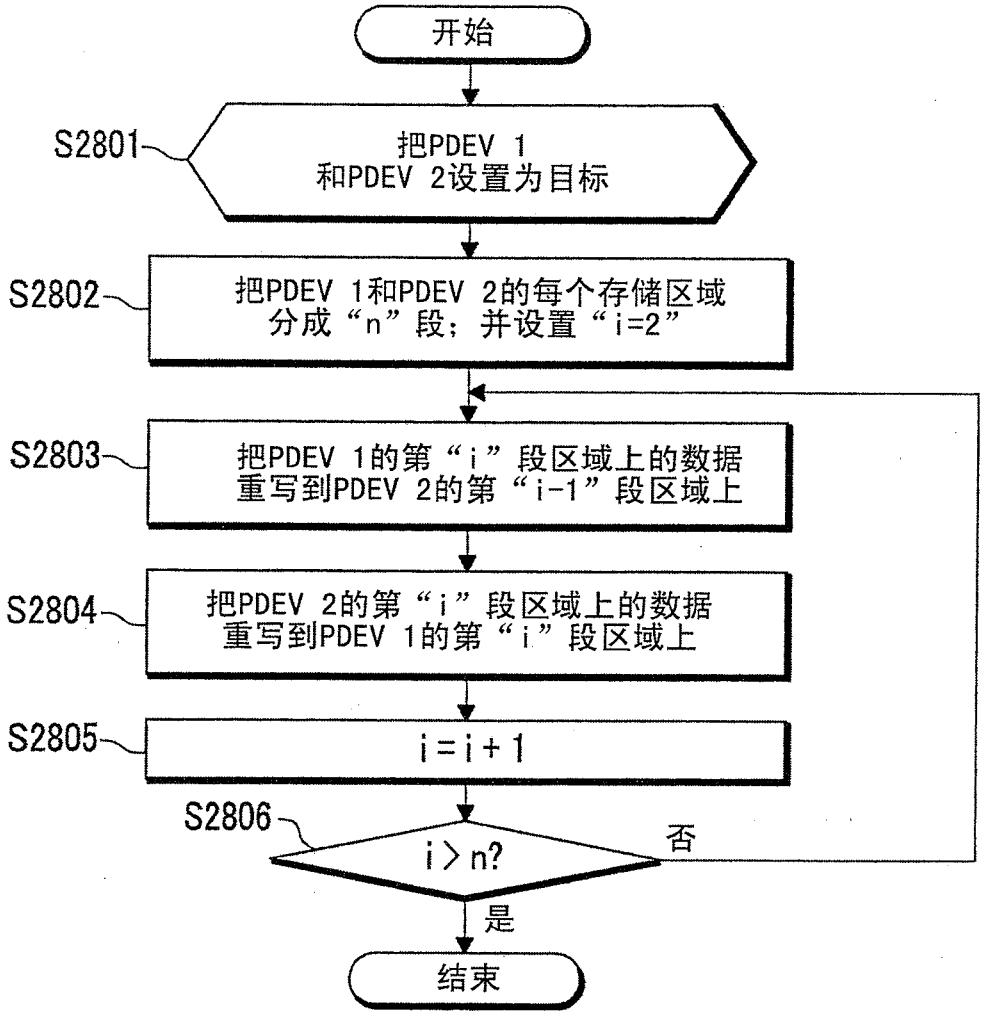

开始
S2801
把PDEV 1
和PDEV 2设置为目标
S2802
把PDEV 1和PDEV 2的每个存储区域
分成“n”段；并设置“i=2”
S2803
把PDEV 1的第“i”段区域上的数据
重写到PDEV 2的第“i-1”段区域上
S2804
把PDEV 2的第“i”段区域上的数据
重写到PDEV 1的第“i”段区域上
S2805
i=i+1
S2806
i>n?
否
是
结束

图 28

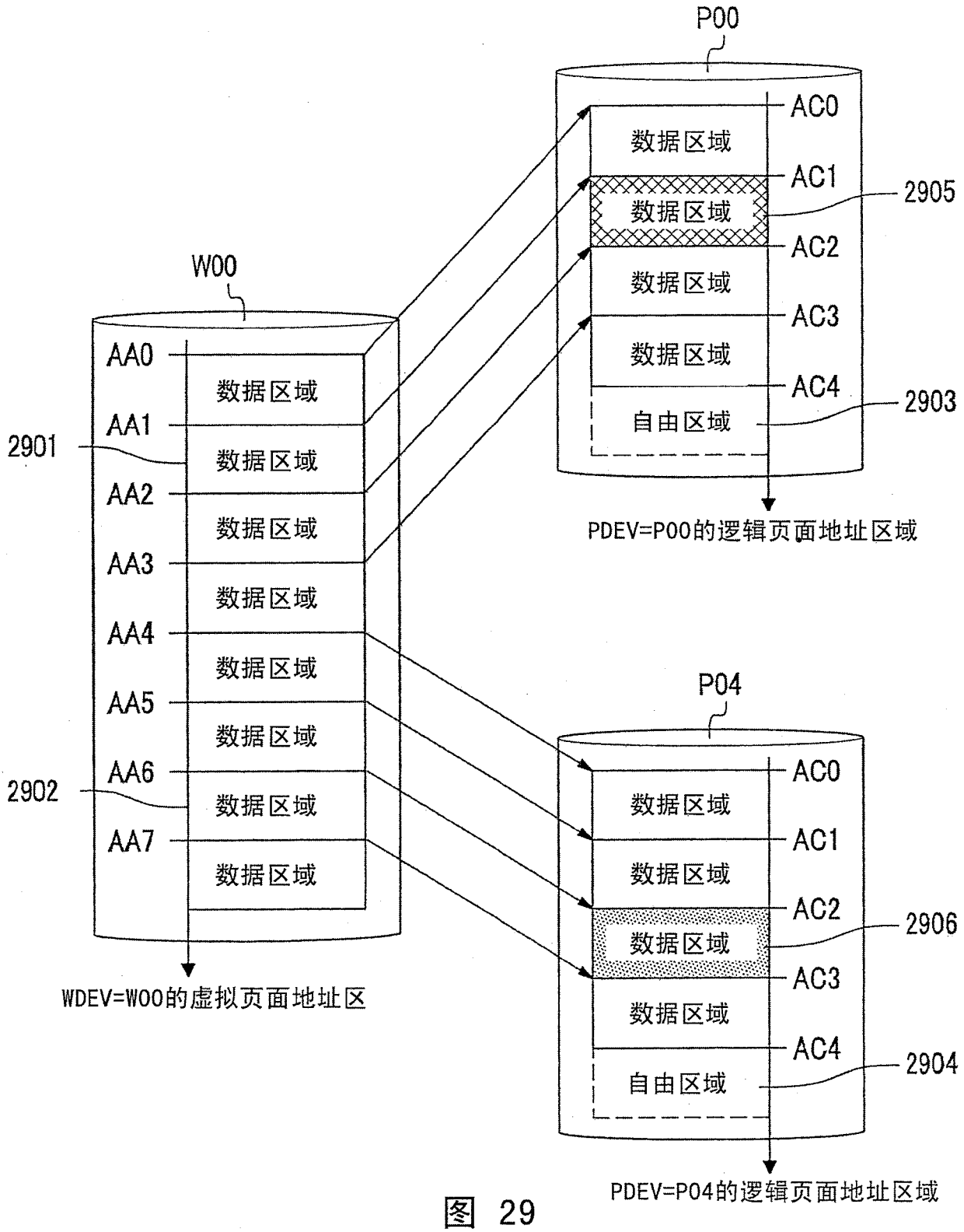

P00
AC0
数据区域
AC1
数据区域
2905
AC2
数据区域
AC3
数据区域
AC4
自由区域
2903
PDEV=P00的逻辑页面地址区域
W00
AA0
数据区域
AA1
2901
数据区域
AA2
数据区域
AA3
数据区域
AA4
数据区域
AA5
数据区域
AA6
2902
数据区域
AA7
数据区域
WDEV=W00的虚拟页面地址区
P04
AC0
数据区域
AC1
数据区域
AC2
数据区域
2906
AC3
数据区域
AC4
自由区域
2904
PDEV=P04的逻辑页面地址区域

图 29

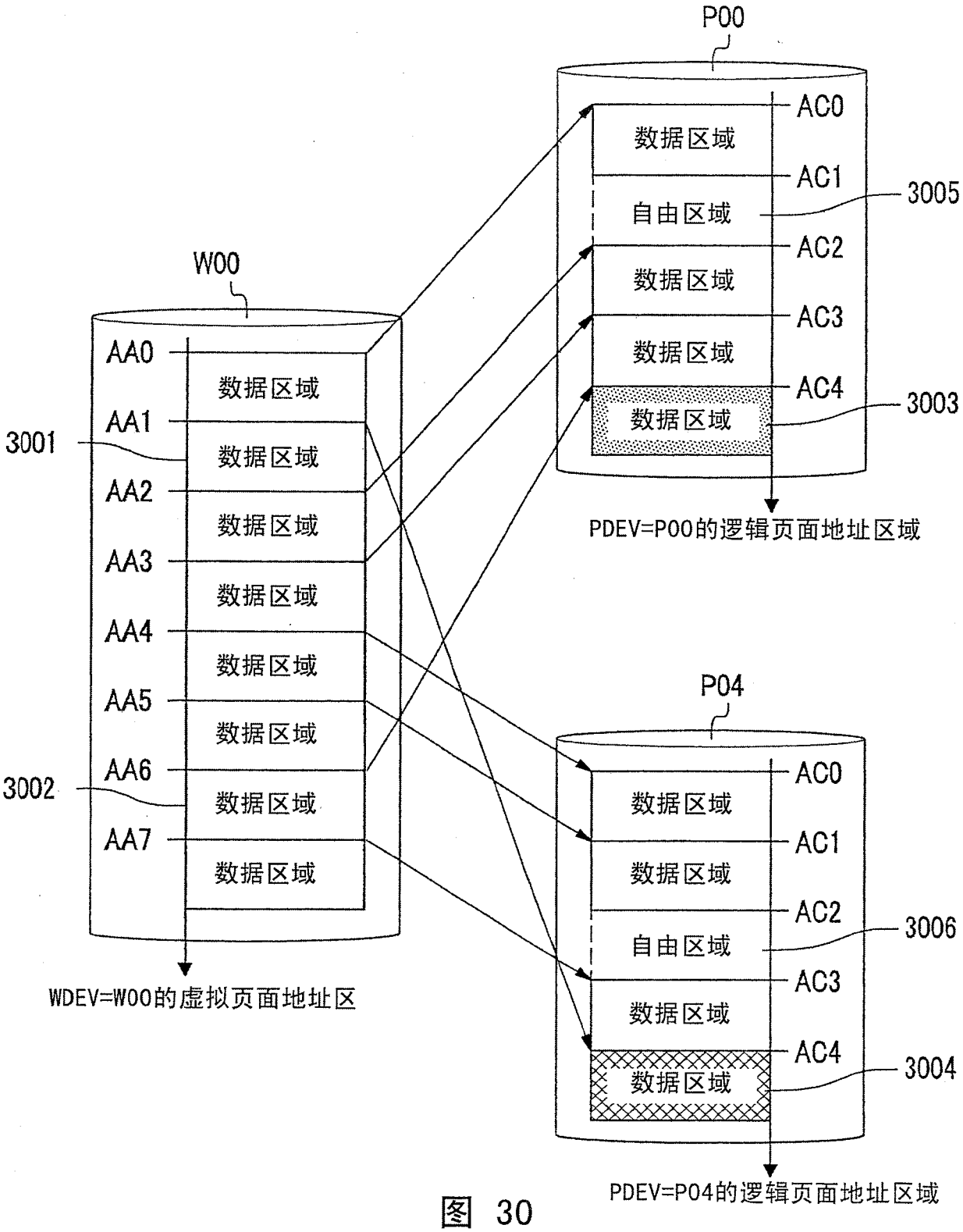

P00
AC0
数据区域
AC1
自由区域
3005
AC2
数据区域
AC3
数据区域
AC4
数据区域
3003
PDEV=P00的逻辑页面地址区域
W00
AA0
数据区域
AA1
3001
数据区域
AA2
数据区域
AA3
数据区域
AA4
数据区域
AA5
数据区域
AA6
3002
数据区域
AA7
数据区域
WDEV=W00的虚拟页面地址区
P04
AC0
数据区域
AC1
数据区域
AC2
自由区域
3006
AC3
数据区域
AC4
数据区域
3004
PDEV=P04的逻辑页面地址区域

图 30

WDEV	虚拟页面地址	数据长度	PDEV	逻辑页面地址	状态
W00	AA0		P00	AC0	
	AA1		P00	AC1	
	AA2		P00	AC2	
	AA3		P00	AC3	
	AA4		P04	AC0	
	AA5		P04	AC1	
	AA6		P04	AC2	
	AA7		P04	AC3	

图 31

WDEV	虚拟页面地址	数据长度	PDEV	逻辑页面地址	状态
W00	AA0		P00	AC0	
	AA1		P04	AC4	
	AA2		P00	AC2	
	AA3		P00	AC3	
	AA4		P04	AC0	
	AA5		P04	AC1	
	AA6		P00	AC4	
	AA7		P04	AC3	

图 32

PDEV	逻辑页面地址	数据长度
P00	AC4	
P04	AC4	

图 33

PDEV	逻辑页面地址	数据长度
P00	AC1	
P04	AC2	

图 34

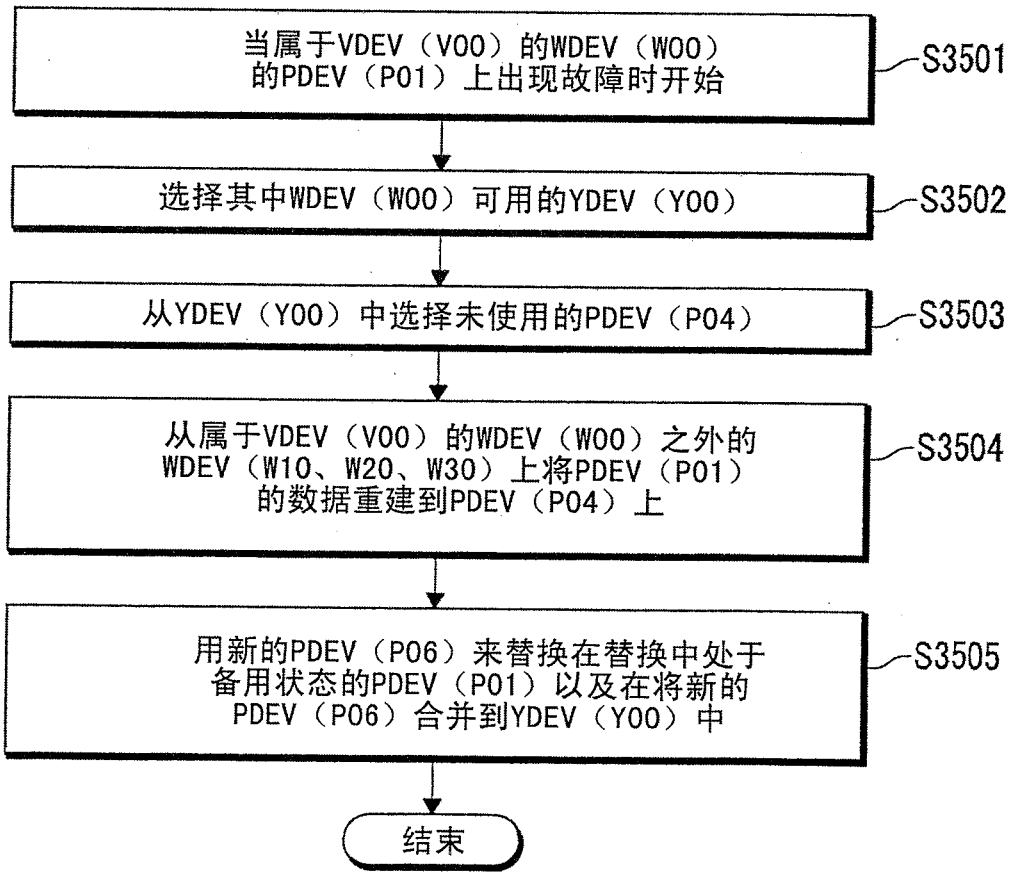


图 35

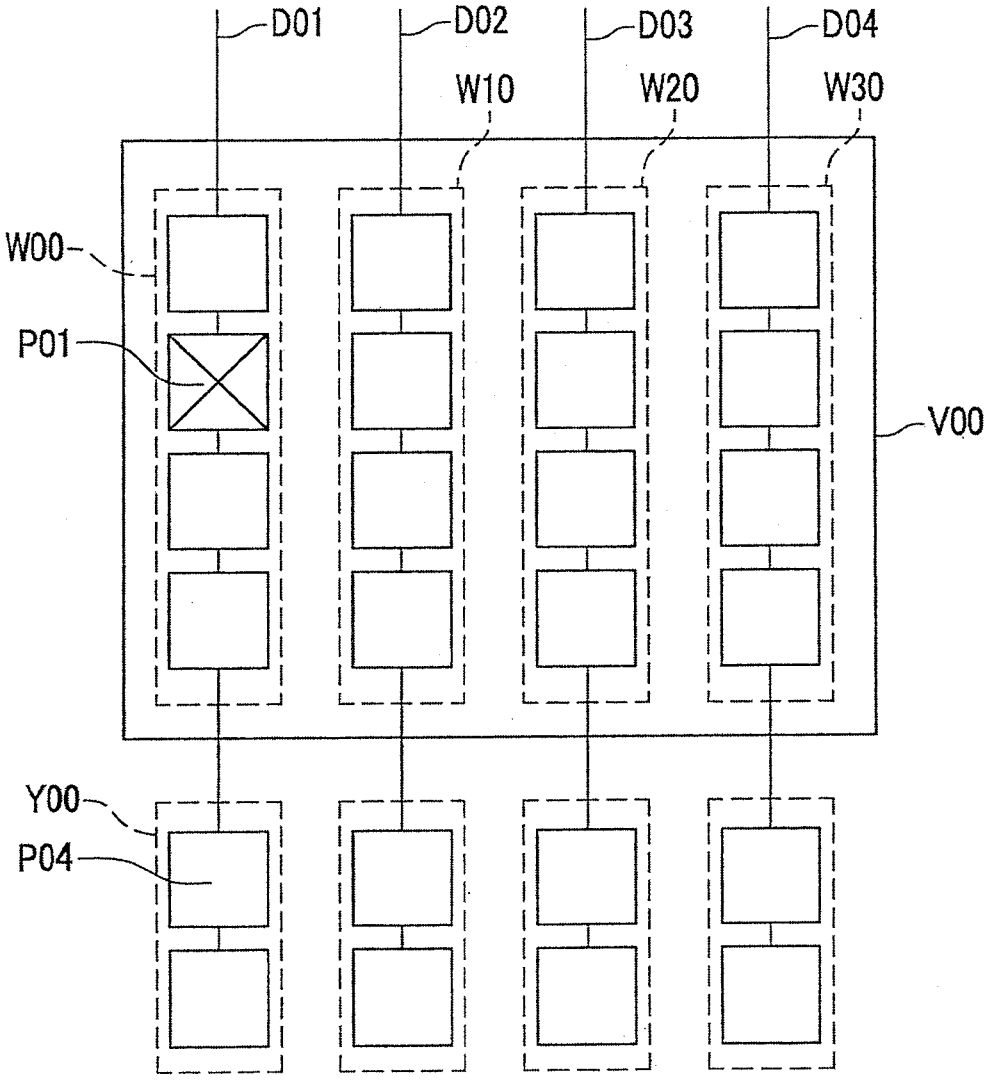


图 36

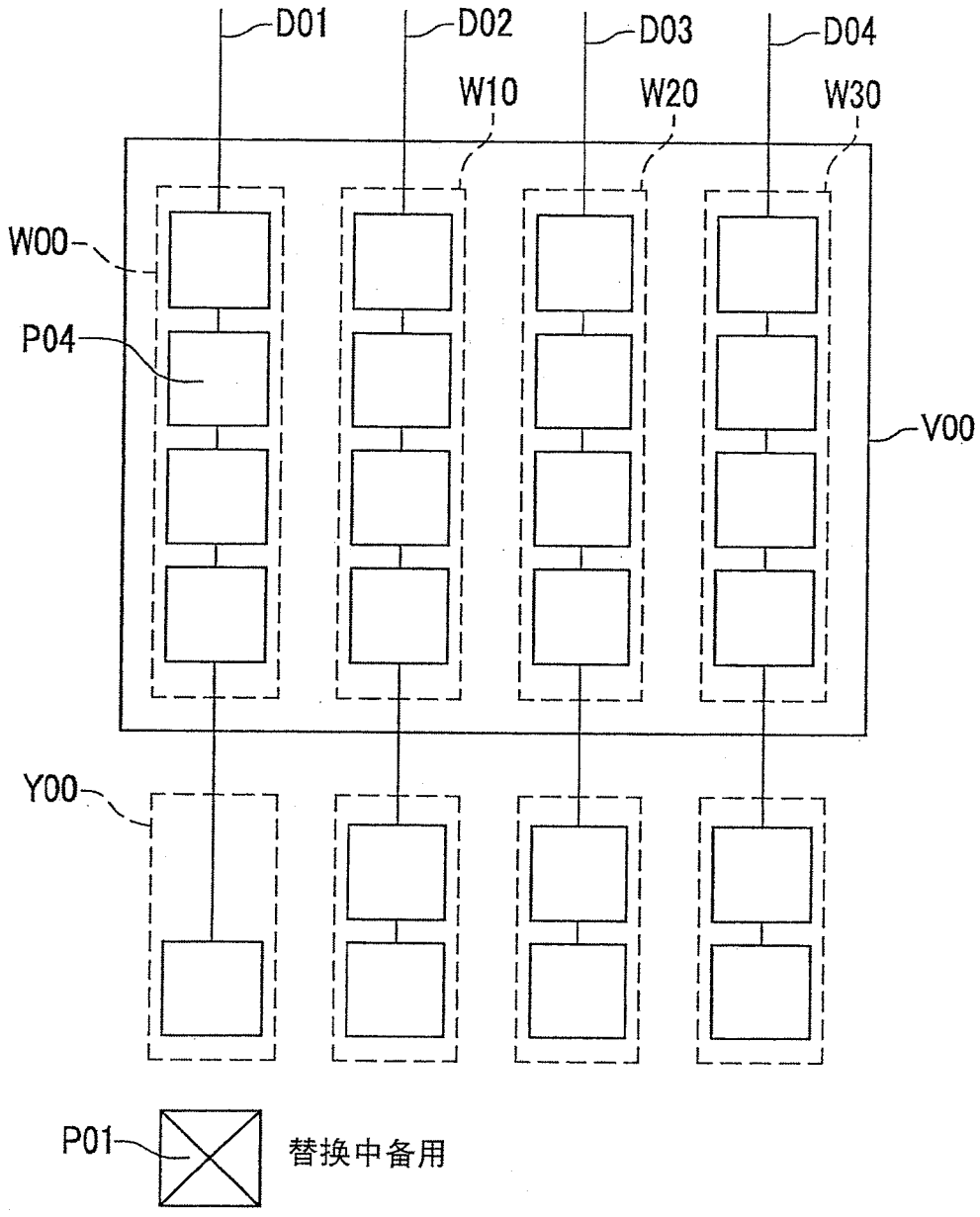


图 37

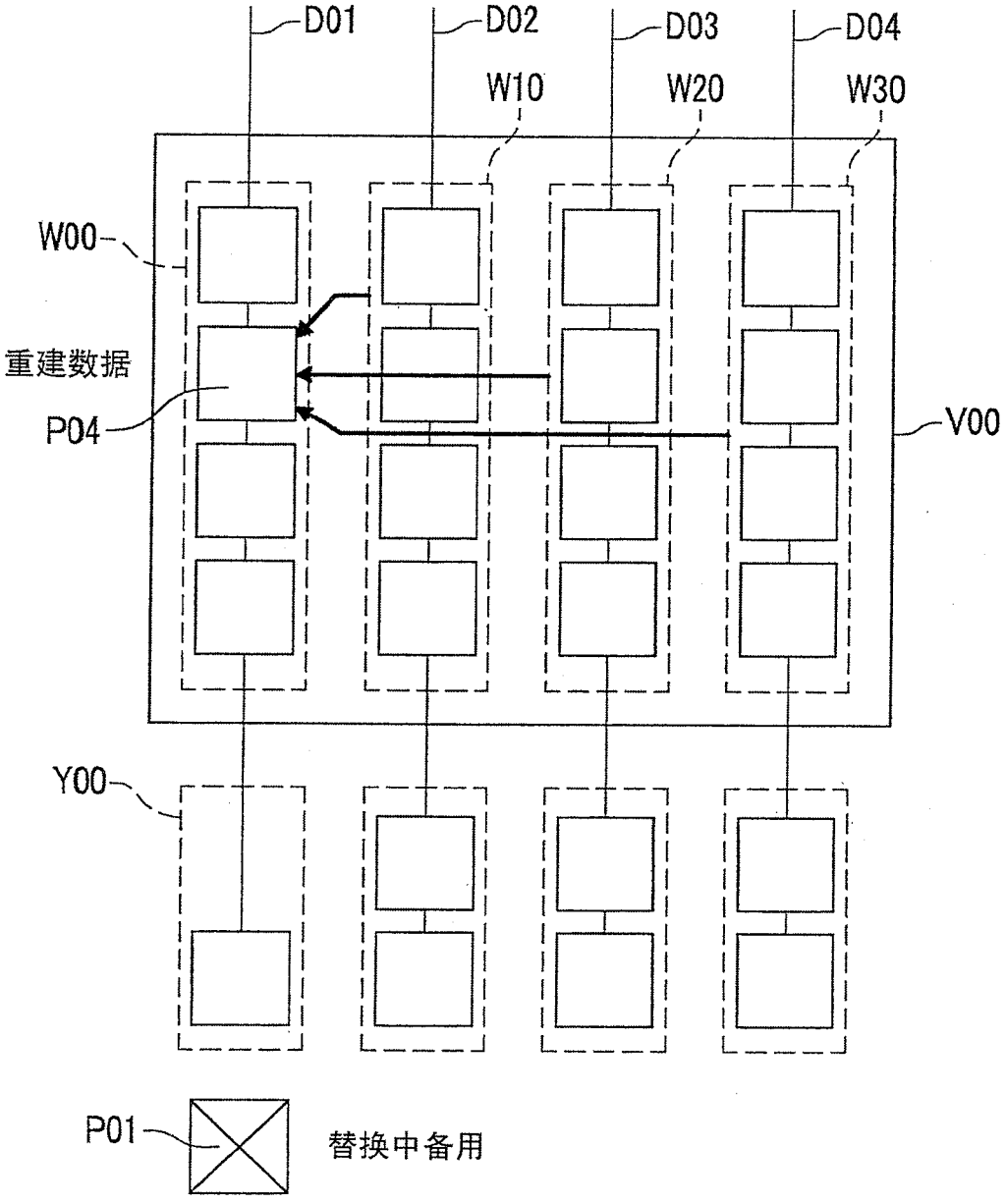


图 38

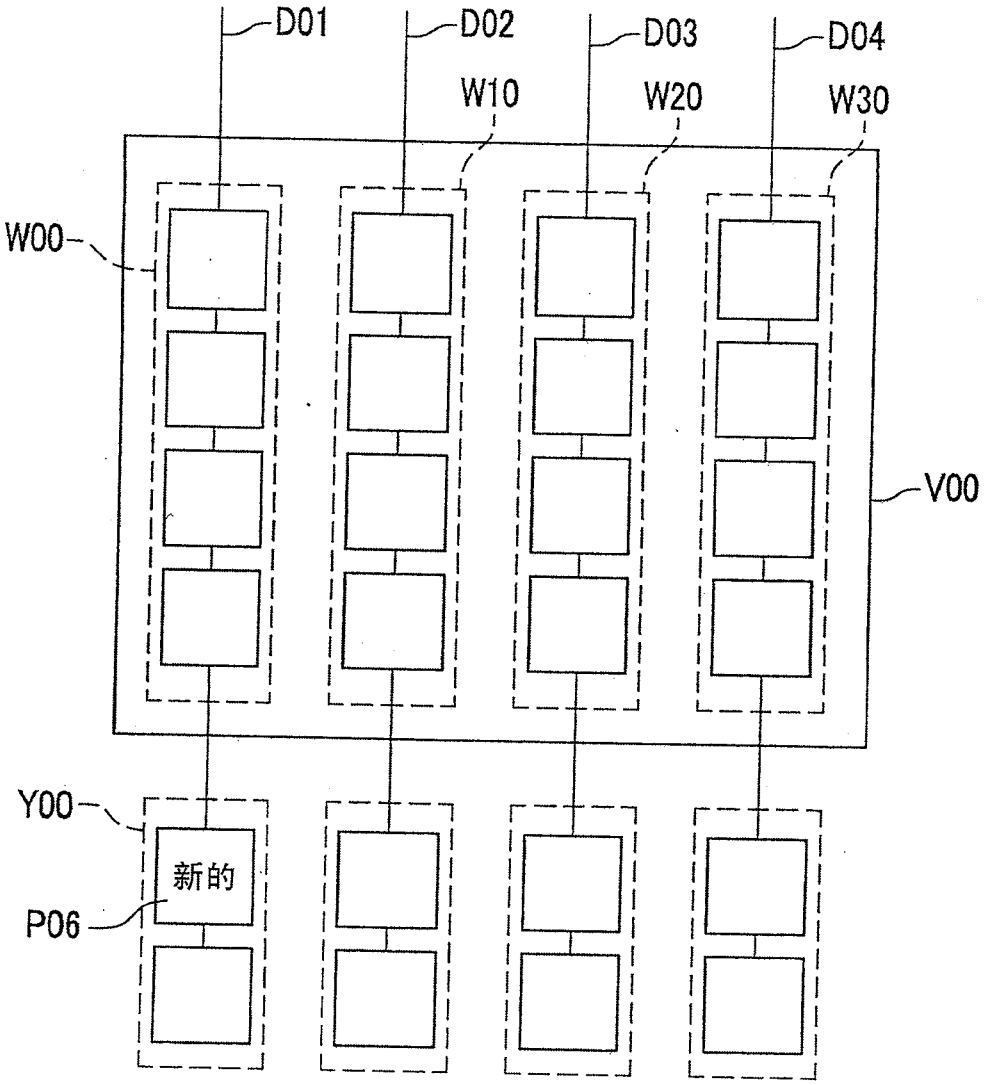


图 39