

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.  
G06F 11/20 (2006.01)



# [12] 发明专利申请公布说明书

[21] 申请号 200780033241.4

[43] 公开日 2009年8月19日

[11] 公开号 CN 101512495A

[22] 申请日 2007.8.22

[21] 申请号 200780033241.4

[30] 优先权

[32] 2006.9.7 [33] US [31] 11/470,970

[86] 国际申请 PCT/EP2007/058712 2007.8.22

[87] 国际公布 WO2008/028803 英 2008.3.13

[85] 进入国家阶段日期 2009.3.6

[71] 申请人 国际商业机器公司

地址 美国纽约

[72] 发明人 R·F·巴尔特福伊

N·M·克莱顿 S·菲恩布利特

O·格卢克 E·马鲁尔

G·A·斯皮尔

[74] 专利代理机构 北京市中咨律师事务所

代理人 于静 杨晓光

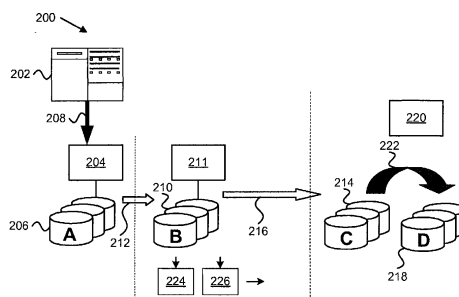
权利要求书 7 页 说明书 25 页 附图 11 页  
按照条约第 19 条的修改 4 页

## [54] 发明名称

用于改进的同步数据镜像交换的装置、系统和方法

## [57] 摘要

本发明披露了一种用于激活同步镜像作为主存储卷的装置、系统和方法。所述装置、系统和方法包括：引导第三存储卷将从第二存储卷发送到第三存储卷的更新存储在所述第一存储卷上以响应所述第一存储卷在故障之后变得可运行；终止将更新从所述第二存储卷发送到所述第三存储卷以响应所述第一存储卷变得与所述第三存储卷基本一致；以及将所述第二存储卷接收的更新同步地存储在所述第一存储卷上并且将所述第一存储卷接收的更新异步地存储在所述第三存储卷上，以响应在所述第一和第三存储卷基本一致之后终止将更新从所述第二存储卷发送到所述第三存储卷。



1. 一种包括计算机可读介质的计算机程序产品，所述计算机可读介质具有针对激活同步镜像作为主存储卷而编程的计算机可用程序代码，所述计算机程序产品的操作包括：

引导第三存储卷将从第二存储卷发送到所述第三存储卷的更新存储在所述第一存储卷上以响应所述第一存储卷在故障之后变得可运行；

终止将更新从所述第二存储卷发送到所述第三存储卷以响应所述第一存储卷变得与所述第三存储卷基本一致；以及

将所述第二存储卷接收的更新同步地存储在所述第一存储卷上并且将所述第一存储卷接收的更新异步地存储在所述第三存储卷上，以响应在所述第一和第三存储卷基本一致之后终止将更新从所述第二存储卷发送到所述第三存储卷。

2. 如权利要求中所述的计算机程序产品，还包括指定所述第二存储卷作为主存储卷以响应在所述第一存储卷处检测到故障，其中所述第一存储卷被指定为所述主存储卷，所述第二存储卷包括所述第一存储卷的数据的同步镜像副本，并且所述第三存储卷包括所述故障之前的所述第二存储卷的数据的异步镜像副本。

3. 如权利要求中所述的计算机程序产品，还包括终止所述第二存储卷上的存储关系，所述存储关系包括所述第二存储卷用于同步地存储来自所述第一存储卷的数据的信息。

4. 如权利要求中所述的计算机程序产品，还包括：

指定所述第一存储卷作为所述主存储卷，以响应将所述第二存储卷接收的更新同步地存储在所述第一存储卷上并且将所述第一存储卷接收的更新异步地存储在所述第三存储卷上；

引导所述第三存储卷将从所述第一存储卷发送到所述第三存储卷的更新存储在所述第二存储卷上以响应所述第一存储卷被指定为所述主存储卷；

终止将更新从所述第一存储卷发送到所述第三存储卷以响应所述第二存储卷变得与所述第三存储卷基本一致；以及

将所述第一存储卷接收的更新同步地存储在所述第二存储卷上并且将所述第二存储卷接收的更新异步地存储在所述第三存储卷上，以响应终止将更新从所述第一存储卷发送到所述第三存储卷。

5. 如权利要求中所述的计算机程序产品，其中引导所述第三存储卷将从所述第二存储卷发送到所述第三存储卷的更新存储在所述第一存储卷上进一步包括：引导所述第三存储卷创建存储关系并且跟踪所述第三存储卷后续接收的更新，所述存储关系包括所述第三存储卷用于将更新异步地存储在所述第一存储卷上的信息。

6. 如权利要求中所述的计算机程序产品，其中终止将更新从所述第二存储卷发送到所述第三存储卷进一步包括：从所述第二存储卷和所述第三存储卷删除存储关系，所述存储关系包括所述第二存储卷用于将数据存储在所述第三存储卷上的信息。

7. 如权利要求中所述的计算机程序产品，其中终止将更新从所述第二存储卷发送到所述第三存储卷进一步包括：跟踪在终止将更新从所述第二存储卷发送到所述第三存储卷之后所述第二存储卷接收的更新。

8. 如权利要求中所述的计算机程序产品，其中将更新异步地存储在所述第三存储卷上进一步包括：使用闪速复制过程从所述第三存储卷来更新第四存储卷，其中在所述第四存储卷上按时间顺序组织更新。

9. 如权利要求中所述的计算机程序产品，其中所述第四存储卷包括适于重新初始化主存储卷的数据。

10. 如权利要求中所述的计算机程序产品，其中终止将更新从所述第二存储卷发送到所述第三存储卷进一步包括终止到所述第四存储卷的更新。

11. 如权利要求中所述的计算机程序产品，其中将所述第一存储卷接收的更新异步地存储在所述第三存储卷上进一步包括重新启动到所述第四存储卷的更新。

12. 如权利要求中所述的计算机程序产品，其中将所述第一存储卷接收的更新异步地存储在所述第三存储卷上进一步包括跟踪所述第一存储卷接收的要复制到所述第三存储卷的更新。

13. 如权利要求中所述的计算机程序产品，其中将所述第二存储卷接收的更新同步地存储在所述第一存储卷上进一步包括引导所述第二存储卷在所述第二存储卷上创建存储关系，所述存储关系包括所述第二存储卷用于将数据同步地存储在所述第一存储卷上的信息。

14. 一种用于激活同步镜像作为主存储卷的装置，所述装置包括：

主交换模块，其配置为指定第二存储卷作为主存储卷以响应在第一存储卷处检测到故障，其中所述第一存储卷被指定为所述主存储卷，所述第二存储卷包括所述第一存储卷的数据的同步镜像副本，并且第三存储卷包括所述故障之前的所述第二存储卷的数据的异步镜像副本；

C-A 故障回复模块，其配置为引导所述第三存储卷将从所述第二存储卷发送到所述第三存储卷的更新存储在所述第一存储卷上以响应所述第一存储卷在所述故障之后变得可运行；

B-C 终止模块，其配置为终止将更新从所述第二存储卷发送到所述第三存储卷以响应所述第一存储卷变得与所述第三存储卷基本一致；以及

B-A-C 存储模块，其配置为将所述第二存储卷接收的更新同步地存储在所述第一存储卷上并且将所述第一存储卷接收的更新异步地存储在所述第三存储卷上，以响应所述 B-C 终止模块终止将更新从所述第二存储卷发送到所述第三存储卷。

15. 如权利要求中所述的装置，还包括 C-A 强制故障转移模块，其配置为引导所述第三存储卷创建存储关系并且跟踪在创建所述存储关系之后所述第三存储卷接收的更新，所述存储关系包括所述第三存储卷用于将更新异步地存储在所述第一存储卷上的信息。

16. 如权利要求中所述的装置，其中所述主交换模块还包括 A-B 终止模块，其配置为终止所述第二存储卷和所述第三存储卷上的存储关系以响应在所述第一存储卷处检测到故障并且使所述第二存储卷成为主存储卷，

所述存储关系包括所述第二存储卷用于同步地存储来自所述第一存储卷的数据的信息。

17. 如权利要求中所述的装置，其中所述 B-C 终止模块还被配置为从所述第二存储卷删除存储关系，所述存储关系包括所述第二存储卷用于将数据存储在其上的信息。

18. 如权利要求中所述的装置，其中所述 B-C 终止模块还包括跟踪更改模块，其配置为跟踪在终止将更新从所述第二存储卷发送到所述第三存储卷之后所述第二存储卷接收的更新。

19. 如权利要求中所述的装置，还包括第四存储卷，并且其中所述第三存储卷被配置为使用闪速复制过程将更新存储到所述第四存储卷，其中在所述第四存储卷上按时间顺序组织所述更新。

20. 如权利要求中所述的装置，其中所述 B-C 终止模块还被配置为终止到所述第四存储卷的更新。

21. 如权利要求中所述的装置，其中所述 B-A-C 存储模块还被配置为重新启动到所述第四存储卷的更新。

22. 如权利要求中所述的装置，其中所述 B-A-C 存储模块还包括 A-C 跟踪模块，其配置为引导所述第一存储卷跟踪所述第一存储卷接收的要复制到所述第三存储卷的更新。

23. 如权利要求中所述的装置，其中所述 B-A-C 存储模块还包括 B-A 增量重新同步模块，其配置为引导所述第二存储卷在所述第二存储卷上创建存储关系，所述存储关系包括所述第二存储卷用于将数据同步地存储在所述第一存储卷上的信息。

24. 一种用于激活同步镜像作为主存储卷的系统，所述系统包括：

第一存储卷，其在所述第一存储卷出现故障之前被指定为主存储卷；

第二存储卷，其包括所述故障之前的所述第一存储卷的数据的同步镜像副本；

第三存储卷，其包括所述故障之前的所述第二存储卷的数据的异步镜像副本；

主机，其配置为将更新存储到所述第一存储卷和所述第二存储卷，所述主机包括

主交换模块，其配置为指定所述第二存储卷作为所述主存储卷以响应在所述第一存储卷处检测到故障；

C-A 故障回复模块，其配置为引导所述第三存储卷将从所述第二存储卷发送到所述第三存储卷的更新存储在所述第一存储卷上以响应所述第一存储卷在所述故障之后变得可运行；

B-C 终止模块，其配置为终止将更新从所述第二存储卷发送到所述第三存储卷以响应所述第一存储卷变得与所述第三存储卷近乎一致；以及

B-A-C 存储模块，其配置为将所述第二存储卷接收的更新同步地存储在所述第一存储卷上并且将所述第一存储卷接收的更新异步地存储在所述第三存储卷上，以响应终止将更新从所述第二存储卷发送到所述第三存储卷。

25. 如权利要求中所述的系统，还包括与所述第三存储卷位于一起的第四存储卷，所述第三存储卷配置为使用闪速复制过程将更新存储到所述第四存储卷，其中在所述第四存储卷上按时间顺序组织所述更新。

26. 如权利要求中所述的系统，其中所述主机还配置为使用所述第三存储卷的内容来初始化所述第一存储卷和所述第二存储卷之一以响应出现故障，其中从所述第二存储卷复制所述第三存储卷的内容。

27. 一种用于在存储卷上创建新的存储关系的方法，所述方法包括：

在第一存储卷上创建存储关系，其中所述存储关系包括所述第一存储卷用于将所述第一存储卷接收的更新存储到第二存储卷的信息，并且所述第一存储卷与所述第二存储卷没有已有的存储关系；

跟踪从创建所述存储关系以来到数据结构中的所述第一存储卷的更新，其中所述更新旨在被复制到所述第二存储卷；以及

根据所述数据结构中的所跟踪更新将所述更新从所述第一存储卷复制到所述第二存储卷。

28. 如权利要求中所述的方法，其中存储关系是第一存储关系，并且

所述第一存储卷与第三存储卷具有已有的第二存储关系。

29. 如权利要求中所述的方法，其中在不重新初始化所述第一存储卷的情况下创建所述存储关系。

30. 一种用于部署软件以激活同步镜像作为主存储卷的方法，所述方法包括：

确定客户同步镜像和异步镜像配置以及可靠性要求；以及

执行同步镜像激活计算机代码，所述计算机代码配置为

指定第二存储卷作为主存储卷以响应在第一存储卷处检测到故障，其中所述第一存储卷被指定为所述主存储卷，所述第二存储卷包括所述第一存储卷的数据的同步镜像副本，并且第三存储卷包括所述故障之前的所述第二存储卷的数据的异步镜像副本；

引导所述第三存储卷将从所述第二存储卷发送到所述第三存储卷的更新存储在所述第一存储卷上以响应所述第一存储卷在所述故障之后变得可运行；

终止将更新从所述第二存储卷发送到所述第三存储卷以响应所述第一存储卷变得与所述第三存储卷基本一致；以及

将所述第二存储卷接收的更新同步地存储在所述第一存储卷上并且将所述第一存储卷接收的更新异步地存储在所述第三存储卷上，以响应在所述第一和第三存储卷基本一致之后终止将更新从所述第二存储卷发送到所述第三存储卷。

31. 一种用于激活同步镜像作为主存储卷的装置，所述装置包括：

用于指定第二存储卷作为主存储卷以响应在第一存储卷处检测到故障的设备，其中所述第一存储卷被指定为所述主存储卷，所述第二存储卷包括所述第一存储卷的数据的同步镜像副本，并且第三存储卷包括所述故障之前的所述第二存储卷的数据的异步镜像副本；

用于引导所述第三存储卷将从所述第二存储卷发送到所述第三存储卷的更新存储在所述第一存储卷上以响应所述第一存储卷在所述故障之后变得可运行的设备；

用于终止将更新从所述第二存储卷发送到所述第三存储卷以响应所述第一存储卷变得与所述第三存储卷基本一致的设备；以及

用于将所述第二存储卷接收的更新同步地存储在所述第一存储卷上并且将所述第一存储卷接收的更新异步地存储在所述第三存储卷上，以响应在所述第一和第三存储卷基本一致之后终止将更新从所述第二存储卷发送到所述第三存储卷的设备。



## 用于改进的同步数据镜像交换的装置、系统和方法

### 技术领域

本发明涉及同步数据镜像，具体地说，涉及在包括异步数据镜像的系统中有效地从主存储卷交换到同步数据镜像。

### 背景技术

数据存储系统用户可能需要高度可靠性以及从故障快速恢复。有些数据存储系统用户可能需要保护数据以免由于自然灾害而丢失。例如，诸如金融机构、零售商之类的数据存储系统用户可能针对金融交易、库存控制等需要高度可靠性。金融机构和其他机构可以使用诸如 IBM®提供的 TotalStorage® Enterprise Storage Server®之类的企业存储系统（“ESS”）来存储可靠性非常重要的数据。此类系统可以包括被组织为数据存储卷的高度可伸缩数据存储设备，例如廉价/独立磁盘冗余阵列（“RAID”）。

数据存储系统可以包括连接数据存储卷的存储区域网络（“SAN”）。通常，存储控制器用于通过 SAN 访问数据存储卷。主机通常将写入请求发送到将数据存储在本地的主存储卷上的存储控制器。在级联数据系统布置中，存储控制器还将存储在主存储卷上的数据发送到辅助数据存储卷的存储控制器。数据存储卷还可以位于主存储卷和关联存储控制器的远程，以便允许在影响主存储卷的自然灾害之后恢复。数据存储卷还可以相对接近主存储卷，这降低了数据访问率。

为了最小化故障之后的恢复时间，辅助存储卷可以是主数据卷的镜像。可以使用同步写入操作将数据写入镜像数据卷。通常，同步写入操作需要在将成功消息返回到请求写入操作的主机之前将数据写入主存储卷和镜像数据卷。同步写入操作需要更多的开销，这是成功地将数据写入主存储卷

和镜像存储卷需要的时间所致。

通常，主存储卷和镜像数据卷之间的距离会影响性能，使得位于主存储卷和关联存储控制器远程的数据存储卷可能将性能降低到使同步写入操作无法实行。可以使用异步写入操作将数据写入远程数据卷。异步写入操作通常允许在成功地将数据写入所述写入操作被导向的存储卷之后返回针对所述写入操作的成功消息。然后，在完成将数据写入所述写入操作被导向的存储卷存储装置之后，将所述数据写入远程存储卷。

在级联存储系统中，可以将数据写入主存储卷并且同步写入镜像数据卷。然后，将数据从镜像数据卷异步复制到远程数据卷。远程存储卷中将要存储的数据通常被存储在非易失性存储器或写入高速缓存以及存储控制器的常规高速缓存中，以便如果在可以将数据异步写入远程存储卷之前发生错误，数据将不会丢失。

通常使用一个或多个位图来跟踪要异步写入远程存储卷的数据。通常，使用位阵列（通常称为活动磁道阵列、更改的磁道阵列或更改记录（“CR”）位图），按照自上次同步、或自上次启动异步传输的存储卷与远程存储卷是一致的以来，启动所做更改的异步传输的存储卷上的磁道地址来保持实时记录。使用第二位图（有时称为恢复磁道阵列、复制磁道阵列或不同步“OOS”位图）指定哪些磁道已作为异步写入操作的一部分被写入主存储卷并且仍需要被复制到远程存储卷。

通常，当更新成功地被写入远程存储卷时，OOS位图的相应位将被清除。在一个实施例中，将CR位图的内容定期与OOS位图合并。在另一个实施例中，通常当OOS位图指示没有更多要写入远程存储卷的更新时，CR位图的内容将定期改写OOS位图。

图1A中示出了常用数据存储系统体系结构100。数据存储系统100包括在将数据发送到第一存储卷106的第一存储控制器104的主机102中。第一存储卷106初始被指定为主存储卷，并且通过本地连接108从第一存储控制器104接收数据。第二存储卷110是主存储卷的镜像数据卷，并且借助同步连接112通过关联的第二存储控制器111接收数据。在一个实施

例中，同步连接 112 是城域镜像（Metro Mirror）连接。通常第一存储控制器 104 和第一存储卷 106 位于本地位置，并且还可以相隔某一距离。

第二存储卷 110 接收的更新还被写入通常位于第一和第二存储卷 106、110 远程的第三存储卷 114。更新被异步发送到第三存储卷 114。第二和第三存储卷 110、114 通过可以包括光纤电缆、铜质电缆、无线连接等的连接 116 通信。在一个实施例中，异步连接 116 是全球镜像（Global Mirror）连接。在一个实施例中，第四存储卷 118 与第三存储卷 114 位于一起。第三存储控制器 120 便于将数据复制到第三存储卷 114 以及从第三存储卷 114 复制到第四存储卷 118。通常，第三存储控制器 120 使用闪速复制过程 122 将更新复制到第四存储卷 118，并且在第四存储卷 118 上按时间顺序汇集更新。

通常，如果数据丢失、损坏或不可访问，则使用第四存储卷 118（如果在存储系统 100 中存在）将数据恢复到第一或第二存储卷 106、110。在一个实施例中，使用第四存储卷 118 的内容使第三卷与第三存储卷 114 变为同步状态，并且使用第三存储卷 114 重新初始化第一和/或第二存储卷 106、110。由于复制过程需要的时间以及主机 102 不可访问存储卷 106、110、118 的时间，替换第一或第二存储卷 106、110 上的数据（而不是增量地同步存储卷 106、110、118）是所不期望的。

第二存储卷 110 可以具有关联的 CR 位图 124，以便跟踪当第一和第二存储卷 106、110 一致之后发生的更新。第二存储卷 110 还可以具有关联的 OOS 位图，以便跟踪将要异步复制到第三存储卷 114 的更新何时被成功复制。OOS 位图通常跟踪通过接收 CR 位图内容的更新或第二存储卷 110 接收的更新。

图 1B 示出了第一存储卷 106 上出现故障的情况。主机 102 通常交换主存储卷，使得在修复第一存储卷 106 之前第二存储卷 110 是主存储卷。主机 102 通常建立与第二存储卷 110 的连接 128。当修复第一存储卷 106 时，如图 1C 中所示，第二存储卷 110 的第二存储控制器 111 通常暂停到第三存储卷 114 的更新并且建立与第一存储卷 106 的连接 130 以重新同步。

在一个实施例中，建立连接 130 以将数据从第二存储卷 110 发送到第一存储卷 106 是故障转移连接。

一旦第一存储卷 106 与第二存储卷 110 一致，如图 1D 中所示，主机 102 就将主存储卷交换为第一存储卷 106，并且第一存储卷 106 重新建立与第二存储卷 110 的同步连接 112。在一个实施例中，将数据从第一存储卷 106 发送到第二存储卷 110 的重新建立是故障回复 (failback) 连接。第二存储卷 110 再次开始将更新异步发送到第三存储卷 114，第三存储卷 114 开始更新第四存储卷 118。故障转移和故障回复连接需要第一存储卷 106 和第二存储卷 110 之间预先存在的关系，例如存储关系。

在第二存储卷 110 与第一存储卷 106 同步并且已暂停到第三存储卷 114 的更新 (参见图 1C) 期间，第三和第四存储卷 114、118 开始变旧。如果第一存储卷 106 长时间无法运行，则第三和第四存储卷 114、118 将需要很长时间才能变为与第一和第二存储卷 106、110 一致的状态。重新同步第一、第三和第四存储卷 106、114、118 的过程可能花费数小时或甚至数天。任何需要在第三存储卷 114 变为几乎一致的状态之前将第三存储卷 114 的内容重新加载到第一或第二存储卷 106、110 的灾害都将由于第三存储卷 114 未接近于一致有问题。

从以上讨论，应显而易见的是需要一种有效地交换同步镜像卷的装置、系统和方法。有利的是，此类装置、系统和方法在将第二、第三和第四存储卷 110、114、118 维持在几乎同步的状态的同时重新同步第一存储卷 106，并且将减少重新同步系统 100 需要的时间。所述装置、系统和方法将在先前不存在预先存在的存储关系的存储卷之间强制故障转移。

## 发明内容

提供了本发明的一种用于激活同步镜像作为主存储卷的方法。所披露的实施例中的方法主要包括针对所述装置和系统的操作执行以下提供的功能所需的步骤。在一个实施例中，所述方法包括引导第三存储卷将从第二存储卷发送到所述第三存储卷的更新存储在所述第一存储卷上以响应所述第一

存储卷在故障之后变得可运行。所述方法包括终止将更新从所述第二存储卷发送到所述第三存储卷以响应所述第一存储卷变得与所述第三存储卷基本一致。所述方法包括将所述第二存储卷接收的更新同步地存储在所述第一存储卷上，并且将所述第一存储卷接收的更新异步地存储在所述第三存储卷上，以响应在所述第一和第三存储卷基本一致之后终止将更新从所述第二存储卷发送到所述第三存储卷。

所述方法还可以包括指定所述第二存储卷作为主存储卷以响应在所述第一存储卷处检测到故障。所述第一存储卷被指定为主存储卷，所述第二存储卷包括所述第一存储卷的数据的同步镜像副本，并且所述第三存储卷包括所述故障之前的所述第二存储卷的数据的异步镜像副本。在另一个实施例中，所述方法包括终止所述第二存储卷上的存储关系。所述存储关系包括所述第二存储卷用于同步地存储来自所述第一存储卷的数据的信息。

在备选实施例中，所述方法包括指定所述第一存储卷作为主存储卷以响应将所述第二存储卷接收的更新同步地存储在所述第一存储卷上以及将所述第一存储卷接收的更新异步地存储在所述第三存储卷上。所述备选实施例的方法包括引导所述第三存储卷将从所述第一存储卷发送到所述第三存储卷的更新存储在所述第二存储卷上以响应所述第一存储卷被指定为主存储卷。所述备选实施例的方法包括终止将更新从所述第一存储卷发送到所述第三存储卷以响应所述第二存储卷变得与所述第三存储卷基本一致。所述备选实施例的方法包括将所述第一存储卷接收的更新同步地存储在所述第二存储卷上并且将所述第二存储卷接收的更新异步地存储在所述第三存储卷上，以响应终止将更新从所述第一存储卷发送到所述第三存储卷。

在一个实施例中，引导所述第三存储卷将从所述第二存储卷发送到所述第三存储卷的更新存储在所述第一存储卷上还包括引导所述第三存储卷创建存储关系并且跟踪所述第三存储卷后续接收的更新。所述存储关系包括所述第三存储卷用于将更新异步地存储在所述第一存储卷上的信息。在另一个实施例中，终止将更新从所述第二存储卷发送到所述第三存储卷还包括从所述第二存储卷和所述第三存储卷删除存储关系。所述存储关系包

括所述第二存储卷用于将数据存储在上述第三存储卷上的信息。在另一个实施例中，终止将更新从所述第二存储卷发送到所述第三存储卷还包括跟踪在终止将更新从所述第二存储卷发送到所述第三存储卷之后所述第二存储卷接收的更新。

在一个实施例中，将更新异步地存储在所述第三存储卷上还包括使用闪速复制过程更新所述第三存储卷的第四存储卷，其中在所述第四存储卷上按时间顺序组织更新。通常，所述第四存储卷包括适于重新初始化主存储卷或恢复所述第三卷的数据。在一个实施例中，终止将更新从所述第二存储卷发送到所述第三存储卷还包括终止到所述第四存储卷的更新。在另一个实施例中，将所述第一存储卷接收的更新异步地存储在所述第三存储卷上还包括重新启动到所述第四存储卷的更新。

在一个实施例中，将所述第一存储卷接收的更新异步地存储在所述第三存储卷上还包括跟踪所述第一存储卷接收的要复制到所述第三存储卷的更新。所述存储关系包括所述第一存储卷用于将更新异步地存储在所述第三存储卷上的信息。在另一个实施例中，将所述第二存储卷接收的更新同步地存储在所述第一存储卷上还包括引导所述第二存储卷在所述第二存储卷上创建存储关系。所述存储关系包括所述第二存储卷用于将数据同步地存储在所述第一存储卷上的信息。

还提供了一种用于激活同步镜像作为主存储卷的装置，所述装置具有多个模块，所述模块被配置为在功能上执行在维持到第二和第三存储卷的更新的同时，从第三存储卷同步已修复的第一存储卷所需的步骤。所述实施例中的这些模块包括指定第二存储卷作为主存储卷以响应在所述第一存储卷处检测到故障的主交换模块。所述第一存储卷被指定为主存储卷，所述第二存储卷包括所述第一存储卷的数据的同步镜像副本，并且所述第三存储卷包括所述故障之前的所述第二存储卷的数据的异步镜像副本。

所述装置包括引导所述第三存储卷将从所述第二存储卷发送到所述第三存储卷的更新存储在所述第一存储卷上以响应所述第一存储卷在故障之后变得可运行的 C-A 故障回复模块。所述装置包括终止将更新从所述第二

存储卷发送到所述第三存储卷以响应所述第一存储卷变得与所述第三存储卷基本一致的 B-C 终止模块。所述装置包括将所述第二存储卷接收的更新同步地存储在所述第一存储卷上并且将所述第一存储卷接收的更新异步地存储在所述第三存储卷上，以响应所述 B-C 终止模块终止将更新从所述第二存储卷发送到所述第三存储卷的 B-A-C 存储模块。

在一个实施例中，所述装置被配置为包括引导所述第三存储卷创建存储关系并且跟踪在创建所述存储关系之后所述第三存储卷接收的更新的 C-A 强制故障转移模块。所述存储关系包括所述第三存储卷用于将更新异步地存储在所述第一存储卷上的信息。在另一个实施例中，所述主交换模块还包括被配置为终止所述第二存储卷和所述第三存储卷上的存储关系以响应在所述第一存储卷处检测到故障并且使所述第二存储卷成为主存储卷的 A-B 终止模块。所述存储关系包括所述第二存储卷用于同步地存储来自所述第一存储卷的数据的信息。

在一个实施例中，所述 B-C 终止模块还被配置为从所述第二存储卷删除存储关系。所述存储关系包括所述第二存储卷用于将数据存储在所述第三存储卷上的信息。在另一个实施例中，所述 B-C 终止模块还包括被配置为跟踪在终止将更新从所述第二存储卷发送到所述第三存储卷之后所述第二存储卷接收的更新的跟踪更改模块。

在一个实施例中，所述装置包括第四存储卷，并且所述第三存储卷被配置为使用闪速复制过程将更新存储到所述第四存储卷。在所述第四存储卷上按时间顺序组织更新。在一个实施例中，所述 B-C 终止模块还被配置为终止到所述第四存储卷的更新。在另一个实施例中，所述 B-A-C 存储模块还被配置为重新启动到所述第四存储卷的更新。

在一个实施例中，所述 B-A-C 存储模块还包括引导所述第一存储卷跟踪所述第一存储卷接收的要复制到所述第三存储卷的更新的 A-C 跟踪模块。在另一个实施例中，所述 B-A-C 存储模块还包括引导所述第二存储卷在所述第二存储卷上创建存储关系的 B-A 增量重新同步模块。所述存储关系包括所述第二存储卷用于将数据同步地存储在所述第一存储卷上的信

息。

还提供了本发明的一种用于激活同步镜像作为主存储卷的系统。所述系统可以体现为在第一存储卷出现故障之前被指定为主存储卷的第一存储卷、包括故障之前的所述第一存储卷的数据的同步镜像副本的第二存储卷、包括故障之前的所述第二存储卷的数据的异步镜像副本的第三存储卷，以及被配置为将更新存储到所述第一存储卷和所述第二存储卷的主机。具体地说，在一个实施例中，所述主机包括指定所述第二存储卷作为主存储卷以响应在所述第一存储卷处检测到故障的主交换模块。所述主机包括引导所述第三存储卷将从所述第二存储卷发送到所述第三存储卷的更新存储在所述第一存储卷上以响应所述第一存储卷在故障之后变得可运行的 C-A 故障回复模块。

所述主机包括终止将更新从所述第二存储卷发送到所述第三存储卷以响应所述第一存储卷变得与所述第三存储卷几乎一致的 B-C 终止模块。所述主机包括将所述第二存储卷接收的更新同步地存储在所述第一存储卷上并且将所述第一存储卷接收的更新异步地存储在所述第三存储卷上，以响应终止将更新从所述第二存储卷发送到所述第三存储卷的 B-A-C 存储模块。所述系统还可以包括与所述第三存储卷位于一起的第四存储卷，其中所述第三存储卷被配置为使用闪速复制过程将更新存储到所述第四存储卷，并且在所述第四存储卷上按时间顺序组织更新。

提供了本发明的一种用于在存储卷上创建新的存储关系的方法。所披露的实施例中的方法主要包括针对所述装置和系统的操作执行以上提供的功能所需的步骤。在一个实施例中，所述方法包括在第一存储卷上创建存储关系。所述存储关系包括所述第一存储卷用于将所述第一存储卷接收的更新存储到第二存储卷的信息，并且所述第一存储卷与所述第二存储卷没有已有的存储关系。所述方法包括跟踪从创建所述存储关系以来到数据结构中的第一存储卷的更新。更新将要被复制到所述第二存储卷。所述方法包括根据所述数据结构中的所跟踪更新将更新从所述第一存储卷复制到所述第二存储卷。在一个实施例中，存储关系是第一存储关系，并且所述第



一存储卷与所述第三存储卷具有已有的第二存储关系。在另一个实施例中，在不重新初始化所述第一存储卷的情况下创建所述存储关系。

提供了本发明的一种用于部署软件以激活同步镜像作为主存储卷的方法。所披露的实施例中的方法主要包括针对所述装置和系统的操作执行以下提供的功能所需的步骤。在一个实施例中，所述方法包括确定客户同步镜像和异步镜像配置以及可靠性要求，并且执行同步镜像激活计算机代码。所述计算机代码被配置为指定第二存储卷作为主存储卷以响应在第一存储卷处检测到故障。所述第一存储卷被指定为主存储卷，所述第二存储卷包括所述第一存储卷的数据的同步镜像副本，并且第三存储卷包括所述故障之前的所述第二存储卷的数据的异步镜像副本。

所述计算机代码被配置为引导所述第三存储卷将从所述第二存储卷发送到所述第三存储卷的更新存储在所述第一存储卷上以响应所述第一存储卷在故障之后变得可运行。所述计算机代码被配置为终止将更新从所述第二存储卷发送到所述第三存储卷以响应所述第一存储卷变得与所述第三存储卷基本一致。所述计算机代码被配置为将所述第二存储卷接收的更新同步地存储在所述第一存储卷上并且将所述第一存储卷接收的更新异步地存储在所述第三存储卷上，以响应在所述第一和第三存储卷基本一致之后终止将更新从所述第二存储卷发送到所述第三存储卷。

本说明书中对功能、优点的参考或类似语言并非暗示可以与本发明一起实现的所有功能和优点应在本发明的任何单个实施例中。相反，应当理解，引用所述功能和优点的语言指与实施例一起描述的特定功能、优点或特性包括在本发明的至少一个实施例中。因此，本说明书中对功能、优点的讨论和类似语言可以（但并不一定）指相同的实施例。

此外，所述功能、优点和特性可以以任何适当的方式组合在一个或多个实施例中。相关领域的技术人员将认识到，可以在没有特定实施例的一个或多个特定功能或优点的情况下实现本发明。在其他情况下，可以在特定实施例中认识到可能不存在于本发明的所有实施例中的其他功能和优点。

## 附图说明

将仅通过实例的方式参考以下附图描述本发明的优选实施例，这些附图是：

图 1A 是示出主存储卷出现故障之前的典型三位置数据存储系统的一个实施例的示意性方块图；

图 1B 是示出主存储卷出现故障之后的其中第二存储卷成为主存储卷的典型三位置数据存储系统的一个实施例的示意性方块图；

图 1C 是示出重新同步第一存储卷同时暂停到第三存储卷的更新的典型三位置数据存储系统的一个实施例的示意性方块图；

图 1D 是示出重新建立第一存储卷作为主存储卷、重新建立到第二存储卷的同步更新以及到第三存储卷的异步更新之后的典型三位置数据存储系统的一个实施例的示意性方块图；

图 2 是示出主存储卷出现故障之前的根据本发明的优选实施例的用于改进的同步镜像交换的数据系统的一个实施例的示意性方块图；

图 3 是示出根据本发明的优选实施例的用于改进的同步镜像交换的装置的一个实施例的示意性方块图；

图 4 是示出根据本发明的优选实施例的用于改进的同步镜像交换的装置的备选实施例的示意性方块图；

图 5 是示出根据本发明的优选实施例的用于改进的同步镜像交换的方法的一个实施例的示意性流程图；

图 6 是示出根据本发明的优选实施例的用于改进的同步镜像交换的方法的备选实施例的示意性流程图；

图 7A 是示出主存储卷出现故障之后的根据本发明的优选实施例的用于改进的同步镜像交换的数据系统的一个实施例的示意性方块图；

图 7B 是示出其中第二存储卷成为主存储卷的根据本发明的优选实施例的用于改进的同步镜像交换的数据系统的一个实施例的示意性方块图；

图 7C 是示出具有为了将更新从第三存储卷发送到第一存储卷而建立

的链接的根据本发明的优选实施例的用于改进的同步镜像交换的数据系统的一个实施例的示意性方块图；

图 7D 是示出根据显示增量重新同步更改记录位图的本发明的优选实施例的用于改进的同步镜像交换的数据系统的一个实施例的示意性方块图；

图 7E 是示出根据显示终止到第四存储卷的更新的本发明的优选实施例的用于改进的同步镜像交换的数据系统的一个实施例的示意性方块图；

图 7F 是示出根据显示终止从第二存储卷到第三存储卷的更新的本发明的优选实施例的用于改进的同步镜像交换的数据系统的一个实施例的示意性方块图；

图 7G 是示出根据显示终止从第三存储卷到第一存储卷的更新的本发明的优选实施例的用于改进的同步镜像交换的数据系统的一个实施例的示意性方块图；

图 7H 是示出根据显示用于将更新从第一存储卷发送到第三存储卷的连接的本发明的优选实施例的用于改进的同步镜像交换的数据系统的一个实施例的示意性方块图；

图 7I 是示出根据显示用于将更新从第二存储卷发送到第三存储卷的同步连接的本发明的优选实施例的用于改进的同步镜像交换的数据系统的一个实施例的示意性方块图；

图 7J 是示出根据显示等待卷变得同步的本发明的优选实施例的用于改进的同步镜像交换的数据系统的一个实施例的示意性方块图；以及

图 7K 是示出根据显示在第一、第三和第四卷之间重新建立异步写入操作的本发明的优选实施例的用于改进的同步镜像交换的数据系统的一个实施例的示意性方块图。

## 具体实施方式

将本说明书中描述的多个功能单元标记为模块，以便更具体地强调它们的实现无关性。例如，模块可以被实现为包括定制 VLSI 电路或门阵列、

诸如逻辑芯片、晶体管或其他分离组件之类的现用半导体的硬件电路。模块还可以在诸如现场可编程门阵列、可编程阵列逻辑、可编程逻辑器件之类的可编程硬件设备中实现。

模块还可以在软件中实现，以便由各种类型的处理器执行。例如，标识的可执行代码的模块可以包括一个或多个物理或逻辑的计算机指令块，所述块可以例如组织为对象、过程或功能。然而，标识的模块的可执行代码不需要在物理上位于一起，但是可以包括存储在不同位置的不同指令，当所述指令被逻辑地结合时，将包括所述模块并实现模块的所述目的。

实际上，可执行代码的模块可以是单个指令或多个指令，并且甚至可以分布在数个不同的代码段上、不同的程序中，以及跨数个存储设备。同样，操作数据可以在模块中被标识和在此示出，并且可以包括在任何适当的形式中并组织在任何适当类型的数据结构中。操作数据可以被收集为单个数据集，或可以分布在包括不同存储设备的不同位置上，并且可以至少部分地仅作为电子信号存在于系统或网络中。

对信号承载介质的引用可以采取任何能够生成信号、导致信号生成或导致在数字处理装置上执行机器可读指令的节目的形式。信号承载介质可以体现为传输线路、光盘、数字视盘、磁带、伯努利驱动器、磁盘、穿孔卡、闪存、集成电路或其他数字处理装置存储器件。

本发明的实施例的所述功能、结构或特性可以以任何适当的形式组合在一个或多个实施例中。在以下说明中，提供了大量特定的详细信息，如编程、软件模块、用户选择、网络事务、数据库查询、数据库结构、硬件模块、硬件电路、硬件芯片等的实例，以便彻底理解本发明的实施例。但是，相关领域的技术人员将认识到，可以在没有—个或多个特定的细节的情况下实现本发明，或者可以通过其他方法、组件、材料等实现本发明。

通常作为逻辑流程图说明在此包括的示意性流程图。因此，所示顺序和标记的步骤指示所提供方法的一个实施例。可以构想其他步骤和方法，它们在功能、逻辑或效果上与所示方法的一个或多个步骤或其各部分等效。另外，提供所使用的格式和符号以说明所述方法的逻辑步骤，并且应理解

它们并非限制所述方法的范围。尽管可以在流程图中使用不同的箭头类型和线条类型，但是应理解，它们并非限制相应方法的范围。实际上，某些箭头或其他连接符可用来仅指示所述方法的逻辑流。例如，箭头可以指示所示方法的列举步骤之间的未指定持续时间的等待或监视周期。另外，出现特定方法的顺序可以或不严格地依照所示的相应步骤的顺序。

图 2 是示出主存储卷出现故障之前的根据本发明的优选实施例的用于改进的同步镜像交换的数据系统 200 的一个实施例的示意性方块图。数据系统 200 包括主机 202、第一存储控制器 204、第一存储卷 206、主机 202 和第一存储控制器 204 之间的连接 208、第二存储卷 210、第二存储控制器 211、第一和第二存储卷 206、210 之间的同步连接 212、第三存储卷 214、第二和第三存储卷 210、214 之间的异步连接 216、第四存储卷 218、第三存储控制器 220、第三和第四存储卷 214、218 之间的连接 222，以及与第二存储卷 210 关联的 CR 位图 224 和 OOS 位图 226，其将在下面描述。在一个实施例中，除了数据系统 200 包括根据本发明的优选实施例的用于改进的同步镜像交换的模块、组件、连接等之外，数据系统 200 的组件类似于针对图 1A 描述的数据存储系统 100。

数据系统 200 包括一个或多个将写入请求发送到存储控制器 204、211、220 的主机 202。主机 202 通常是服务器，但可以是工作站、个人计算机、便携式计算机等。在一个实施例中，主机 202 包括计算机、服务器、打印机、路由器等的网络。在另一个实施例中，主机 202 与第一存储控制器 204 结合。主机 202 通常通过 SAN 与存储控制器 204、211、220 通信，但可以通过局域网、广域网、因特网、无线网络等通信。主机 202 还可以被配置为将读取请求、复制请求或其他数据相关的请求发送到存储控制器 204、211、220。本领域的技术人员将认识到能够将数据写入请求发送到存储控制器 204、211、220 的主机 202 的其他实施例。

数据系统 200 包括被配置为从主机 202 接收写入请求并且将文件、文件更新、文件信息等（此后称为“更新”）存储到一个或多个存储卷 206、210、214、218 的第一存储控制器 204。在一个实施例中，第一存储控制器

204 是企业存储系统 (“ESS”) 控制器。在另一个实施例中, 第一存储控制器 204 是 IBM System Storage® DS8000® 存储服务器。第一存储控制器 204 可以是服务器、工作站、数据移动器, 或具有多个组件和计算机的系统。通常, 第一存储控制器 204 包括非易失性存储器或写入高速缓存以及用于暂时存储更新的易失性存储器或高速缓存。本领域的技术人员将认识到能够从主机 202 接收写入请求并且将与写入请求相关的更新存储在存储卷 206、210、214、218 上的其他第一存储控制器 204 配置。

系统 200 包括存储从第一存储控制器 204 接收的更新的第一存储卷 206。通常, 在正常运行条件下, 第一存储卷 206 被指定为主存储卷并且直接从第一存储控制器 204 接收更新。第一存储卷 206 包括至少一个硬盘驱动器 (“HDD”), 但是还可以包括一起形成卷的盘阵列。第一存储卷 206 可以包括 RAID 阵列、简单磁盘捆绑 (“JBOD”) 等。第一存储卷 206 可以包括磁带驱动器、光盘驱动器等。在一个实施例中, 第一存储卷 206 包括形成一个卷的一个盘或盘阵列。在另一个实施例中, 第一存储卷 206 包括多个存储卷。

通常, 第一存储控制器 204 可访问第一存储卷 206 的一个或多个存储卷。在一个实施例中, 第一存储卷 206 包括数据移动器、控制器或其他具有处理器和存储器的设备, 以便控制第一存储卷 206 的一个或多个存储卷并且从第一存储控制器 204 或其他存储控制器接收更新。本领域的技术人员将认识到能够从第一存储控制器 204 接收更新的其他第一存储卷 206。

系统 200 包括从主机 202 到第一存储控制器 204 的数据连接 208。主机 202 通过连接 208 与存储控制器 204、211、220 通信, 连接 208 通常包括 SAN, 但可以包括光纤连接、小型计算机系统接口 (“SCSI”)、局域网、路由器、交换机等。在级联数据存储系统 200 中, 通常当第一存储卷 206 被指定为主存储卷时, 第一存储控制器 204 将来自主机 202 的更新直接发送到第一存储卷 206, 而不通过第二或第三存储卷 210、214。

数据系统 200 包括基本类似于第一存储卷 206 的第二存储卷 210。通常, 第二存储卷 210 初始是第一存储卷 206 的同步镜像数据卷, 并且包括

与第一存储卷 206 相同的数据。通常，第二存储卷 210 具有与第一存储卷 206 相同的容量、相同的磁道数等。第二存储卷 210 可以位于与第一存储卷 206 足够接近的中间位置，以便允许适合同步连接 212 的连接而不会产生不可接受的性能降低。在一个实施例中，第二存储卷 210 与第一存储卷 206 位于一起。

数据系统 200 包括第一存储卷 206（通过关联的第一存储控制器 204）到第二存储卷 210（通过关联的第二存储控制器 211）之间的数据连接 212。数据连接 212 基本类似于主机 202 到第一存储控制器 204 的连接 208。数据连接 212 可以是第一和第二存储卷 206、210 之间的同步连接。数据系统 200 包括与第二存储卷关联并且基本类似于第一存储控制器 204 的第二存储控制器 211。

系统 200 包括通常位于第一和第二存储卷 206、210 远程的第三存储卷 214 和第四存储卷 218。系统 200 包括与第三和第四存储卷 214、218 关联并且基本类似于第一和第二存储控制器 204、211 的第三存储控制器 220。第三存储卷 214 初始从第二存储卷和控制器 210、211 异步接收更新。在异步操作模式期间，第三存储控制器 220 使用闪速复制过程 222 将第三存储卷 214 接收的更新复制到第四存储卷 218。第三存储控制器 220 通常根据第三存储卷 214 与第二存储卷 210 一致时的时间点或其他适合的重要时间点，按时间顺序重新排列被复制到第四存储卷 218 的更新。

系统 200 包括第二和第三存储卷 210、214（以及关联的控制器 211、220）之间的数据连接 216。数据连接 216 可以是广域网连接、光纤网络连接等，并且可以包括光纤电缆、铜质电缆、无线连接、集线器、路由器等。第三和第四存储卷 214、218 通常位于第一和第二存储卷 206、210 的远程，以便第二和第三存储卷 210、214 之间的数据连接 216 适合于此类距离。

系统 200 包括与第二存储卷 210 关联的 CR 位图 224 和 OOS 位图 226。位图 224、226 通常是数据存储并且由第二存储控制器 211 管理。在一个实施例中，CR 位图 224 跟踪当卷一致时，在与第三存储卷 214 同步之后发生的到第二存储卷 210 的更新。OOS 位图跟踪要被写入第三存储卷 214 的

更新。通常，位图 224、226 都标识到存储卷 210、214 的更新的磁道地址。

例如，当 OOS 位图 226 耗尽之后，将 CR 位图 224 的内容定期与 OOS 位图 226 合并。当 OOS 位图 226 耗尽时，第二和第三卷 210、214 将被视为一致或同步。通过将 OOS 位图 226 跟踪的更新复制到第三存储卷 214 而耗尽 OOS 位图 226。当将更新从第二存储卷 210 写入第三存储卷 214 时，将从 OOS 位图 226 清除与所复制的更新关联的位。CR 位图 224 跟踪在第二和第三存储卷 210、214 一致时的时间点之后发生的更改。OOS 位图 226 还可以跟踪第二存储卷 210 接收的更新。在一个实施例中，在 OOS 位图 226 被完全清除之后，将 CR 位图 224 复制到 OOS 位图 226。当 CR 位图 224 的内容被复制到 OOS 位图 226 或与 OOS 位图 226 合并之后，还可以清除 CR 位图 224 以跟踪后续更新。

图 3 是示出根据本发明的优选实施例的用于改进的同步镜像交换的装置 300 的一个实施例的示意性方块图。装置 300 包括主机 202，主机 202 具有 C-A 故障回复模块 302、B-C 终止模块 304 和 B-A-C 存储模块 306，将在下面描述它们。在一个实施例中，装置 300 的模块 302、304、306 包括在主机 202 中。在另一个实施例中，装置 300 的模块 302、304、306 包括在存储管理计算机（未示出）中。在再一个实施例中，装置 300 的模块 302、304、306 包括在存储控制器 204、211、220 中。本领域的技术人员将认识到装置 300 的模块 302、304、306 的其他位置以激活异步镜像。

装置 300 包括引导第三存储卷 214 将从第二存储卷 210 发送到第三存储卷 214 的更新存储在第二存储卷 210 上以响应第一存储卷 206 在故障之后变得可运行的 C-A 故障回复模块 302。当第一存储卷 206 出现故障之后，第二存储卷 210 可以被指定为主存储单元，并且主机 202 可以将更新发送到第二存储卷 210。通常，第二存储卷 210 继续将更新发送到第三存储卷 214。当第一存储卷 206 未运行时，第三存储卷 214 跟踪 OOS 位图中的更改。C-A 故障回复模块 302 认识到第一存储卷 206 可运行，并且开始将与第三存储卷 214 关联的 OOS 位图所跟踪的更新发送到第一存储卷 206。在一个实施例中，更新被异步地发送。在另一实施例中，更新被同步地发送。



装置 300 包括终止将更新从第二存储卷 210 发送到第三存储卷 214 以响应第一存储卷 206 变得与第三存储卷 214 基本一致的 B-C 终止模块 304。当第一存储卷 206 开始从第三存储卷 214 接收更新之后，第一存储卷 206 最终达到它与第三存储卷 214 几乎一致的点。在一个实施例中，当第一和第三存储卷 206、214 一致时，B-C 终止模块 304 终止发送更新。在另一个实施例中，当第一和第三存储卷 206、214 基本一致时，B-C 终止模块 304 终止发送更新。当 C-A 故障回复模块 302 将更新异步发送到第一存储卷 206 时，后一实施例是优选的。在一个实施例中，B-C 终止模块 304 终止到第四存储卷 218 的更新。

装置 300 包括将第二存储卷 210 接收的更新同步地存储在第一存储卷 206 上并且将第一存储卷 206 接收的更新异步地存储在第三存储卷 214 上，以响应在第一和第三存储卷 206、214 基本一致之后终止将更新从第二存储卷 210 发送到第三存储卷 214 的 B-A-C 存储模块 306。在一个实施例中，B-A-C 存储模块 306 重新启动到第四存储卷 218 的更新。有利地，B-A-C 存储模块 306 将数据存储系统 200 恢复到更新可以从主机 202 流到主存储卷（第二存储卷 210），同步流到镜像数据卷（第一存储卷 206），然后异步地流到远程存储卷（第三和第四存储卷 214、218）的状态。

装置 300 克服了现有技术系统（参见图 1A 至 1D）的许多缺点。装置 300 维持从主机 202 到主存储卷（第二存储卷 210）的数据流，同时还在出现故障的存储卷（第一存储卷 206）恢复联机和同步时实质上维持到远程存储卷（第三和第四存储卷 214、218）的数据流。装置 300 还可以用于重复与装置 300 关联的方法，以便将第一存储卷 206 恢复为主存储卷，并且将第二存储卷 210 恢复为镜像数据卷。

图 4 是示出根据本发明的优选实施例的用于改进的同步镜像交换的装置 400 的备选实施例的示意性方块图。装置 400 包括主机 202，主机 202 具有 C-A 故障回复模块 302、B-C 终止模块 304 和 B-A-C 存储模块 306，它们基本类似于针对图 3 的装置 300 描述的模块。此外，装置 400 包括主交换模块 402 和 C-A 强制故障转移模块 404。在一个实施例中，主交换模

块 402 包括 A-B 终止模块 406。在另一个实施例中，B-C 终止模块 304 包括跟踪更改模块 408。在另一个实施例中，B-A-C 存储模块 306 包括 A-C 跟踪模块 410 和 B-A 增量重新同步模块 412。在一个实施例中，装置 400 包括 C-B 强制故障回复模块 414、B-C 终止模块 416 和 A-B-C 存储模块 418。将在下面描述这些模块。

装置 400 包括指定第二存储卷 210 作为主存储卷以响应在第一存储卷 206 处检测到故障的主交换模块 402。初始地，将第一存储卷 206 指定为主存储卷，第二存储卷 210 包括第一存储卷的数据的同步镜像副本，并且第三存储卷 214 包括故障之前的第二存储卷的数据的异步镜像副本。

在一个实施例中，主交换模块 402 包括 IBM 的 HyperSwap 程序。在另一个实施例中，在检测到第一存储卷 206 中的故障之后，主交换模块 402 自动指定第二存储卷 210 作为主存储卷。在另一个实施例中，主交换模块 402 通过操作员干预来指定第二存储卷 210 作为主存储模块。在一个实施例中，在检测到第二存储卷 210 中的故障之后，主交换模块 402 指定第一存储卷 206 作为主存储卷。本领域的技术人员将认识到主交换模块 402 可以在检测到故障之后指定主存储卷的其他方法。

在一个实施例中，装置 400 包括引导第三存储卷 214 创建存储关系并且跟踪在创建存储关系之后到存储模块的更新的 C-A 强制故障转移模块 404。所述存储关系包括第三存储卷 214 用于将更新异步地存储在第一存储卷 206 上的信息。在一个实施例中，C-A 强制故障转移模块 404 位于第二存储控制器 211 上。在其他实施例中，C-A 强制故障转移模块 404 位于主机 202 或其他存储控制器 204、220 上。

通常，本发明的实施例包括强制故障转移功能和强制故障回复功能，这些功能是对现有技术的故障转移和故障回复命令的改进。强制故障转移功能允许创建先前不存在的与存储卷关联的存储关系。例如，在第一和第二存储卷 206、210 之间可能存在典型的存储关系，以便将来自第一存储卷 206 的更新同步地存储在第二存储卷 210 上。

可以发出故障转移命令以便在第一存储卷 206 出现故障后被联机之

后，指示第二存储卷 210 将更新发送到第一存储卷 206。一旦第一存储卷 206 被同步，就发出故障回复命令以便在第一存储卷 206 再次被指定为主存储卷之后，引导第一存储卷 206 将更新存储在第二存储卷 210 上。使用现有技术的故障转移或故障回复命令，引导第二存储卷 210 跟踪第二存储卷 210 接收的更新并且然后将所跟踪的更新存储在第三存储卷 214 上的存储关系将不可能存在。

强制故障转移功能允许存储卷与其他存储卷建立初始化期间未预定义的存储关系。在一个实施例中，强制故障转移功能与其他存储卷建立存储关系但不发送更新。实际上，所述存储卷创建 OOS 位图以跟踪将被存储到其他存储卷的更改。然后，强制故障回复功能启动从跟踪 OOS 位图上的更改的存储卷流到其他存储卷的更新。在一个实施例中，强制故障转移功能终止所述存储卷的任何在先存储关系。

C-A 强制故障转移模块 404 是强制故障转移功能的实例。例如，第三存储卷 214 与第二存储卷 210 具有异步存储关系，并且在正常操作期间接收更新。C-A 强制故障转移模块 404 在第三存储卷 214 上与第一存储卷 206 建立存储关系。一旦建立存储关系，第三存储卷 214 便开始跟踪第三存储卷 214 接收的将被发送到第一存储卷 206 的更改。在一个实施例中，所述更改存储在 OOS 位图中。OOS 位图可以存储在第三存储控制器 220 上。在没有强制故障转移功能的情况下，第三存储卷 214 将被限制为接收和发送到第二存储卷 210 的更新，并且不能与第一存储卷 206 建立新的存储关系以及跟踪到第一存储卷 206 的更新。

在一个实施例中，主交换模块 402 包括终止第一存储卷 206 上的存储关系以响应在第一存储卷 206 处检测到故障并且使第二存储卷 210 成为主存储卷的 A-B 终止模块 406。所述存储关系包括第二存储卷 210 用于同步地存储来自第一存储卷 206 的数据的信息。在一个实施例中，终止第二存储卷 210 上的存储关系允许主机 202 指定第二存储卷 210 作为主存储卷，并且将更新发送到第二存储卷 210。

在一个实施例中，B-C 终止模块 304 包括跟踪在终止将更新从第二存

储卷 210 发送到第三存储卷 214 之后第二存储卷 210 所接收的更新的跟踪更改模块 408。通常，跟踪更改模块 408 跟踪与第二存储卷 210 关联的增量重新同步 CR 位图中的更改。在一个实施例中，增量重新同步 CR 位图位于第二存储控制器 211 上。跟踪更改模块 408 创建的增量重新同步 CR 位图允许跟踪更改模块 408 跟踪在 B-C 终止模块 304 暂停从第二存储卷 210 到第三存储卷 214 的更新之后接收的更新。通常，在创建新的增量重新同步 CR 位图之后，将不会使用与第二存储卷 210 关联并且用于跟踪第一存储卷 206 出现故障之前的更新的 CR 位图。

在一个实施例中，B-A-C 存储模块 306 包括引导第一存储卷 206 在第一存储卷 206 上创建存储关系的 A-C 跟踪模块 410。A-C 跟踪模块 410 引导第一存储卷 206 创建与第一存储卷 206 关联的 OOS 位图，以便跟踪将被复制到第三存储卷 214 的更新。通常，OOS 位图位于第一存储控制器 204 上。

在另一个实施例中，B-A-C 存储模块 306 包括引导第二存储卷 210 在第二存储卷 210 上创建存储关系的 B-A 增量重新同步模块 412。所述存储关系包括第二存储卷 210 用于将数据同步地存储在第一存储卷 206 上的信息。B-A 增量重新同步模块 412 在第二存储卷 210 上重新创建出现故障之前存在并且从第二存储卷 210 删除的存储关系，以便允许主机 102 将更新发送到第二存储卷 210。

通常，在 B-C 终止模块 304 终止从第二存储卷 210 到第三存储卷 214 的更新之前，将耗尽跟踪到第三存储卷 214 的更新并与第二存储卷 210 关联的 OOS 位图。在 OOS 位图耗尽之后，当 B-A 增量重新同步模块 412 引导第二存储卷 210 建立存储关系时，通常不会使用此位图。在一个实施例中，B-A 增量重新同步模块 412 引导第二存储控制器 211 将跟踪更改模块 408 所创建的增量重新同步 CR 位图的内容合并到跟踪被复制到第一存储卷 206 的更新的新 OOS 位图中。在另一个实施例中，B-A 增量重新同步模块 412 引导第二存储控制器 211 将跟踪更改模块 408 所创建的增量重新同步 CR 位图的内容合并到与第二存储卷 210 关联的现有 OOS 位图中，并且

该 OOS 位图跟踪被复制到第一存储卷 206 的更新。在一个实施例中，一旦第一和第二存储卷 206、210 一致，将不使用 OOS 位图。

在一个实施例中，装置 400 包括引导第三存储卷 214 将从第一存储卷 206 发送到第三存储卷 214 的更新存储在第二存储卷 210 上以响应第一存储卷 206 被指定为主存储卷的 C-B 强制故障回复模块 414。在另一个实施例中，装置 400 包括终止将更新从第一存储卷 206 发送到第三存储卷 214 以响应第二存储卷 210 变得与第三存储卷 214 基本一致的 B-C 终止模块 416。在再一个实施例中，装置 400 包括将第一存储卷 206 接收的更新同步地存储在第二存储卷 210 上并且将第二存储卷 210 接收的更新异步地存储在第三存储卷 214 上，以响应终止将更新从第一存储卷 206 发送到第三存储卷 214 的 A-B-C 存储模块 418。

图 5 是示出根据本发明的优选实施例的用于改进的同步镜像交换的方法 500 的一个实施例的示意性流程图。方法 500 开始 (502)，C-A 故障回复模块 302 引导 (504) 第三存储卷 214 将从第二存储卷 210 发送到第三存储卷 214 的更新存储在第二存储卷 210 上以响应第一存储卷 206 在故障之后变得可运行。在一个实施例中，C-A 故障回复模块 302 引导第三存储卷 214 发出在第三存储卷 214 上创建与第一存储卷 206 的存储关系的强制故障转移功能。

B-C 终止模块 304 终止 (506) 将更新从第二存储卷 210 发送到第三存储卷 214 以响应第一存储卷 206 变得与第三存储卷 214 基本一致。通常，当第一存储卷 206 变为一致或几乎一致的状态时，主机 202 继续将更新发送到第二存储卷 210，并且第二存储卷 210 继续将更新发送到第三存储卷 214。

B-A-C 存储模块 306 将第二存储卷 210 接收的更新同步地存储 (508) 在第一存储卷 206 上以响应 B-C 终止模块 304 终止将更新从第二存储卷 210 发送到第三存储卷 214。此外，B-C 终止模块 304 将第一存储卷 206 接收的更新异步地存储 (508) 在第三存储卷 214 上以响应 B-C 终止模块 304 终止将更新从第二存储卷 210 发送到第三存储卷 214，方法 500 结束 (510)。

在一个实施例中，方法 500 提供了一种方法以便从主卷交换到同步镜像卷，同时维持接收到第二存储卷 210（先前是出现故障的第一存储卷 206 的镜像副本）的更新。第二存储卷 210 是新指定的主卷。方法 500 还允许在第一存储卷 206 返回操作之后同步第一存储卷 206，同时维持第四存储卷 218 几乎一致（作为第三存储卷 214 继续接收更新的结果）。当第一和第二存储卷 206、210 的位置相当接近时，或使第二存储卷 210 作为指定的主存储卷性能降低很小或不降低时，可能希望维持第二存储卷 210 作为主存储卷。

图 6 是示出根据本发明的优选实施例的用于改进的同步镜像交换的方法 600 的备选实施例的示意性流程图。方法 600 提供了使用针对图 5 中描述的方法 500 所用的相同技术的相反过程，以便将级联数据存储系统 200 返回到其原始体系结构，其中第一存储卷 206 作为主存储卷，第二存储卷 210 同步接收来自第一存储卷 206 的更新，并且第三存储卷 214 异步接收来自第二存储卷 210 的更新。当包含第二存储卷 210 的中间存储站点比第一存储卷 206 更远离主机 202 时，或具有比第一存储卷 206 更慢的到主机 202 的连接时，可能需要方法 600。

方法 600 开始于 (602) 主交换模块 402，主交换模块指定 (604) 第二存储卷 210 作为主存储卷以响应在第一存储卷 206 处检测到故障。第一存储卷 206 被指定为主存储卷，第二存储卷 210 包括故障之前的第一存储卷的数据的同步镜像副本，并且第三存储卷 214 包括故障之前的第二存储卷的数据的异步镜像副本。C-A 故障回复模块 302 引导 (606) 第三存储卷 214 将从第二存储卷 210 发送到第三存储卷 214 的更新存储在第二存储卷 206 上以响应第一存储卷 206 在故障之后变得可运行。

B-C 终止模块 304 终止 (608) 将更新从第二存储卷 210 发送到第三存储卷 214 以响应第一存储卷 206 变得与第三存储卷 214 基本一致。B-A-C 存储模块 306 将第二存储卷 210 接收的更新同步地存储 (610) 在第一存储卷 206 上，并且 B-C 终止模块 304 将第一存储卷 206 接收的更新异步地存储 (610) 在第三存储卷 214 上，以响应 B-C 终止模块 304 终止将更新从

第二存储卷 210 发送到第三存储卷 214。

主交换模块 402 判定 (612) 第一存储卷 206 是否将是主存储卷。如果否, 则方法 600 结束 (622)。如果主交换模块 402 判定 (612) 第一存储卷 206 将是主存储卷, 则主交换模块 402 指定 (614) 第一存储卷 206 作为主存储卷。C-B 强制故障回复模块 414 引导 (616) 第三存储卷 214 将从第一存储卷 206 发送到第三存储卷 214 的更新存储在第二存储卷 210 上以响应第一存储卷 206 被指定为主存储卷。

B-C 终止模块 416 终止 (618) 将更新从第一存储卷 206 发送到第三存储卷 214 以响应第二存储卷 210 变得与第三存储卷 214 基本一致。A-B-C 存储模块 418 将第一存储卷 206 接收的更新同步地存储 (620) 在第二存储卷 210 上, 并且将第二存储卷 210 接收的更新异步地存储在第三存储卷 214 上, 以响应终止将更新从第一存储卷 206 发送到第三存储卷 214, 方法 600 结束 (622)。

图 7A 至 7K 是示出根据本发明的优选实施例的用于改进的同步镜像交换的数据系统 200 的一个实施例的示意性方块图。图 7A 至 7K 示出了根据本发明的优选实施例的用于重新配置数据系统 200 以实现在主存储卷出现故障之后将镜像数据卷指定为主存储卷的步骤。

图 7A 示出了针对图 2 描述的系统 200, 其中第一存储卷 206 处于故障状态。通常, 故障将中断沿着第一存储卷 206 和第二存储卷 210 之间的连接 212 的同步数据流。主交换模块 402 在检测到故障之后指定第二存储卷 210 作为主存储卷 (图 7B)。在一个实施例中, A-B 终止模块 406 终止第二存储卷 210 上与第一存储卷 206 的存储关系。此实施例允许主交换模块 402 指定第二存储卷 210 作为主存储模块, 并且允许主机 202 通过主机 202 和第二存储卷 210 之间的连接 228 将更新发送到第二存储卷 210。

C-A 强制故障转移模块 404 引导第三存储卷 214 在第三存储卷 214 上建立存储关系。故障转移命令在第三存储卷 214 上建立第三和第一存储卷 214、206 之间的关系, 以便第三存储卷 214 将更新发送到第一存储卷 206。通常, 第一存储卷 206 此时未运行, 因此 C-A 强制故障转移模块引导第三

存储卷 214 (通过第三存储控制器 220) 创建 OOS 位图 230, 以便跟踪第三存储卷 214 上将被复制到第一存储卷 206 的更改。

C-A 强制故障转移模块 404 引导第三存储控制器 220 发出强制故障回复功能以便启动从第三存储卷 214 到第一存储卷 206 的更新(参见图 7C)。一旦已恢复第一存储卷 206, 就通过第三和第一存储卷 214、206 之间的连接 232 发送更新。通常, 更新被异步复制。在一个实施例中, 获取与第一存储卷 206 (未示出) 关联的 OOS 位图的内容(如果可能), 以便重新同步可能在出现故障期间发生的更改。然后, 系统 200 可以继续运行, 更新从主机 202 流到第二存储卷 210, 并且第二存储卷 210 将更新异步发送到第三存储卷 214, 同时第三存储卷 214 更新第一存储卷 206。

在一个实施例中, 当第一存储卷 206 几乎同步时, 跟踪更改模块 408 在第二存储控制器 211 上创建增量重新同步 CR 位图 234, 以便跟踪在停止到第三存储卷 214 的更新之后第二存储卷 210 将接收的更新(参见图 7D)。最终, 与第二存储卷 210 关联的原始 CR 位图 224 上的更新将被传输到第二存储卷的 OOS 位图 226, 并且原始 CR 位图 224 是不活动的(在图 7E 中示为不存在)。在准备停止从第二存储卷 210 到第三存储卷 214 的更新时, 第三存储控制器 220 将暂停到第四存储卷 218 的更新, 并且第四存储卷 218 开始变旧。

然后, B-C 终止模块 304 终止从第二存储卷 210 到第三存储卷 214 的更新(参见图 7F)。B-C 终止模块 304 还终止第三存储卷 214 中的第二到第三存储卷关系。在清除与第三存储卷 214 关联的 OOS 位图 230 之前, 更新继续流到第一存储卷 206。第一和第三存储卷 206 和 214 通常在此时是一致的。然后暂停从第三存储卷 214 到第一存储卷 206 的更新(参见图 7G)。

A-C 强制故障转移模块 410 引导第一存储卷 206 创建 OOS 位图 236 以跟踪被复制到第三存储卷 214 的更新(参见图 7H)。B-A-C 存储模块 306 发出故障回复命令, 并且第一存储卷 206 开始异步更新第三存储卷 214 (参见图 7H)。



B-A 增量重新同步模块 412 引导第二存储卷 210 在第二存储卷 210 上建立同步存储关系以将更新发送到第一存储卷 206 (参见图 7I)。如果第一存储卷 206 与第二存储卷 210 没有关系, 则 B-A 增量重新同步模块 412 引导第一存储卷 206 建立关系以从第二存储卷 210 接收更新。在一个实施例中, 第二存储控制器 211 将增量重新同步 CR 位图 234 的内容合并到新的 OOS 位图 238 中以跟踪被复制到第一存储卷 206 的更新。在另一个实施例中, 第二存储控制器 211 将增量重新同步 CR 位图 234 的内容合并到现有的未使用的 OOS 位图 226 (未在图 7I 中示出) 内。

一旦清除与第二存储卷 210 关联的 OOS 位图 238 的内容, 通常第一和第二存储卷 206、210 将是一致的, 并且 OOS 位图 238 是不活动的 (在图 7J 中示为不存在)。将在第一存储卷 206 上创建 (或重新激活) CR 位图 240, 并且 B-A-C 存储卷引导第一存储卷 206 将更新异步发送到第三存储卷 214。当接收更新时, 第三存储控制器 220 使用闪速复制过程更新第四存储卷 218 (参见图 7K)。

现在, 已在级联数据存储配置中配置系统 200, 其中数据从主机 202 流到第二存储卷 210, 同步地流到第一存储卷 206, 异步地流到第三存储卷 214。在第四存储卷 218 上按时间顺序排列更新。如果需要第一存储卷 206 成为主存储卷, 则重复此过程, 并且将主卷从第二存储卷 210 交换到第一存储卷 206。有利地, 图 7 中所示的过程以及在此所述的方法 500、600 和装置 300、400 允许主卷交换和返回, 其中第三/第四存储卷 214、218 保持近似同步, 同时重新同步第一存储卷 206。此外, 更新继续从主机 202 流到第二存储卷 210, 并且从第二存储卷 210 流到第三存储卷 214。

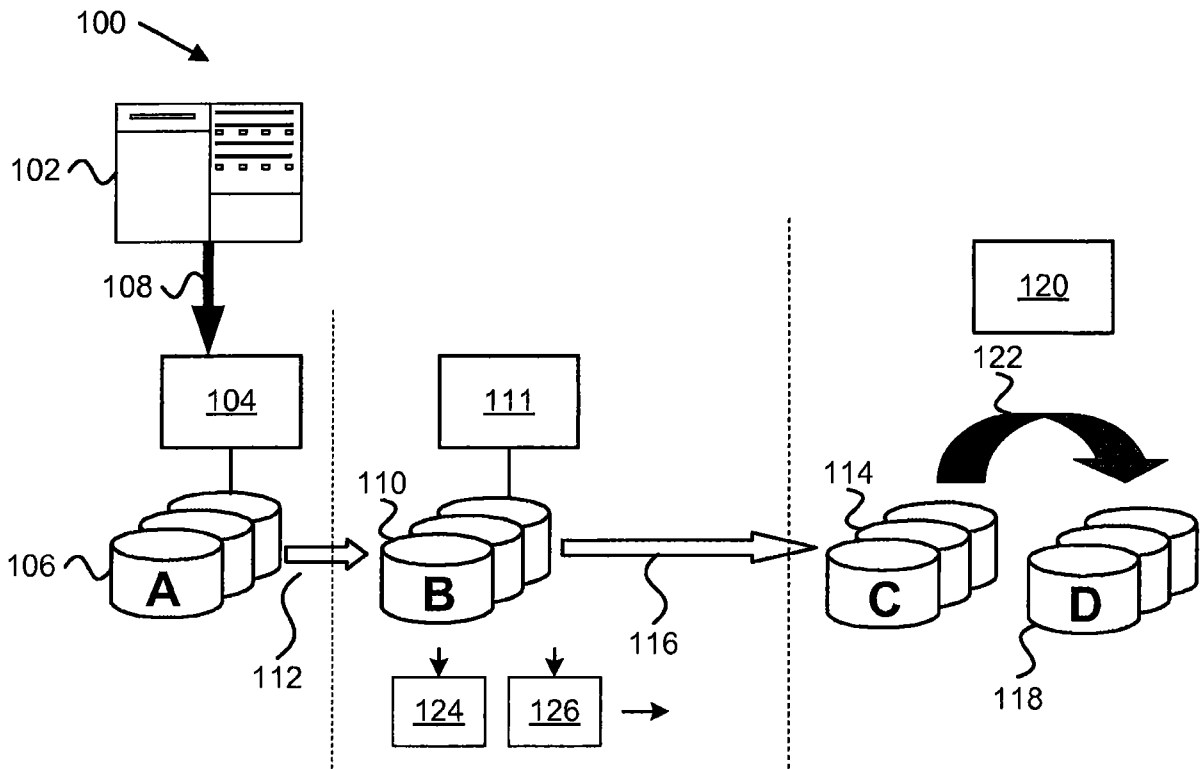


图 1A  
(现有技术)

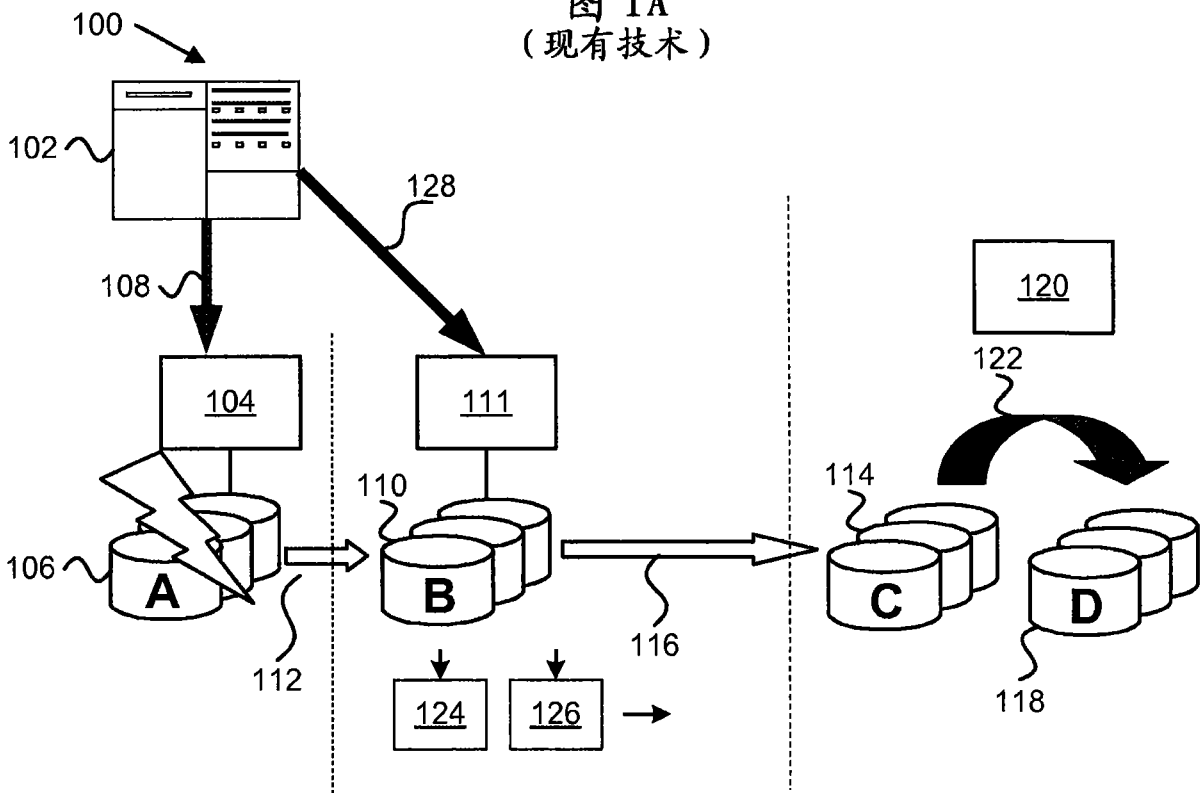


图 1B  
(现有技术)

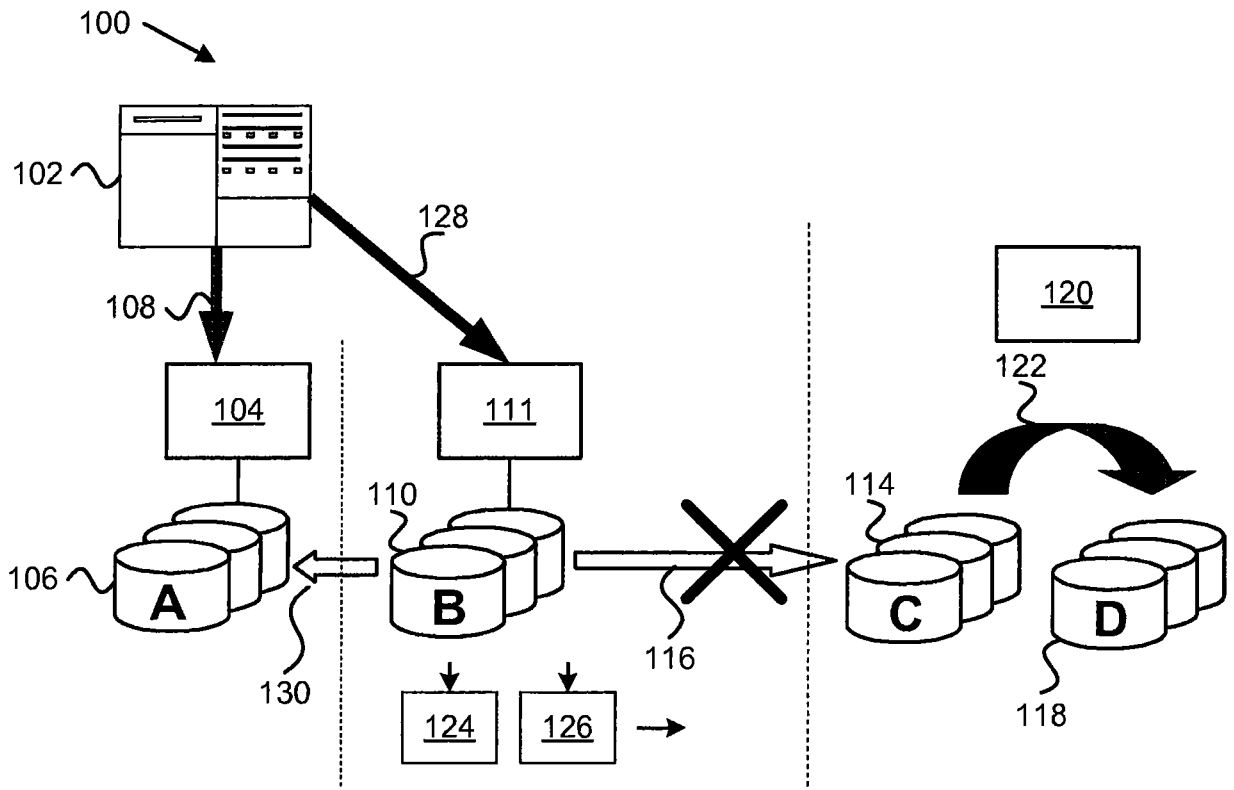


图 1C  
(现有技术)

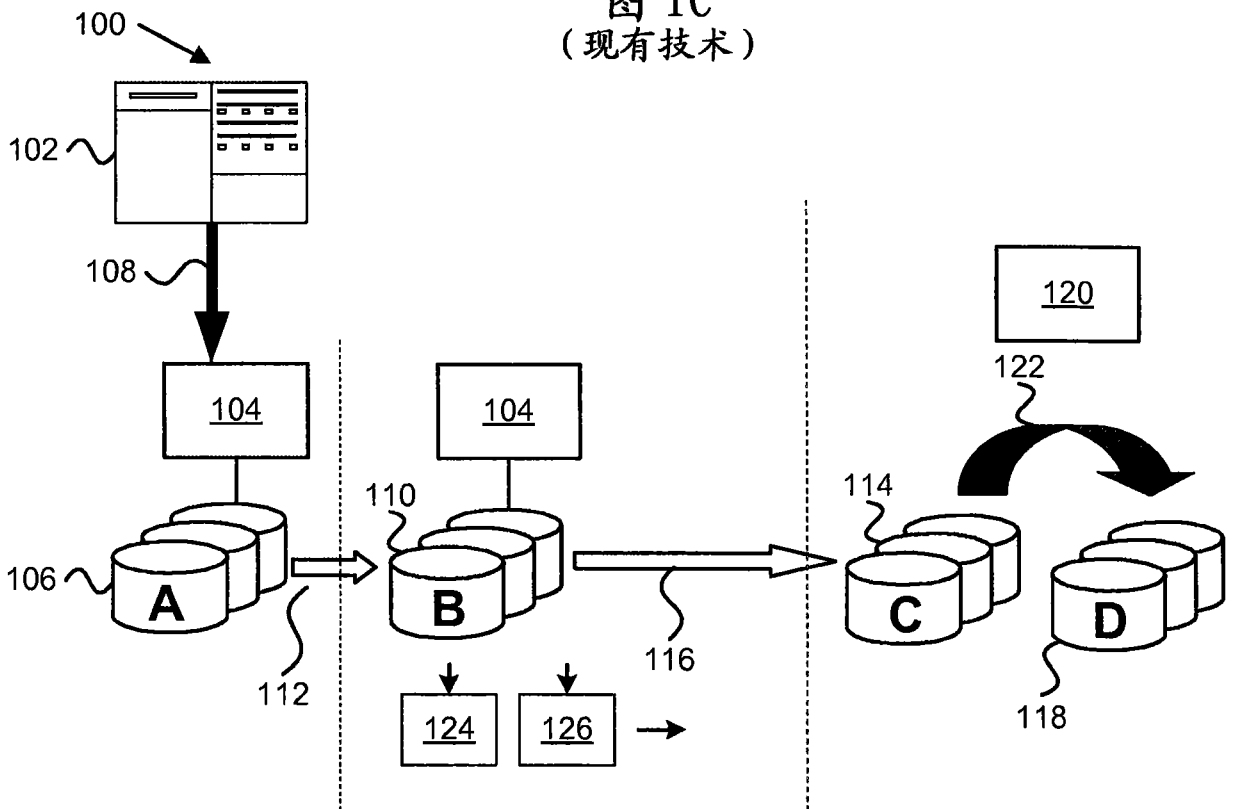


图 1D  
(现有技术)

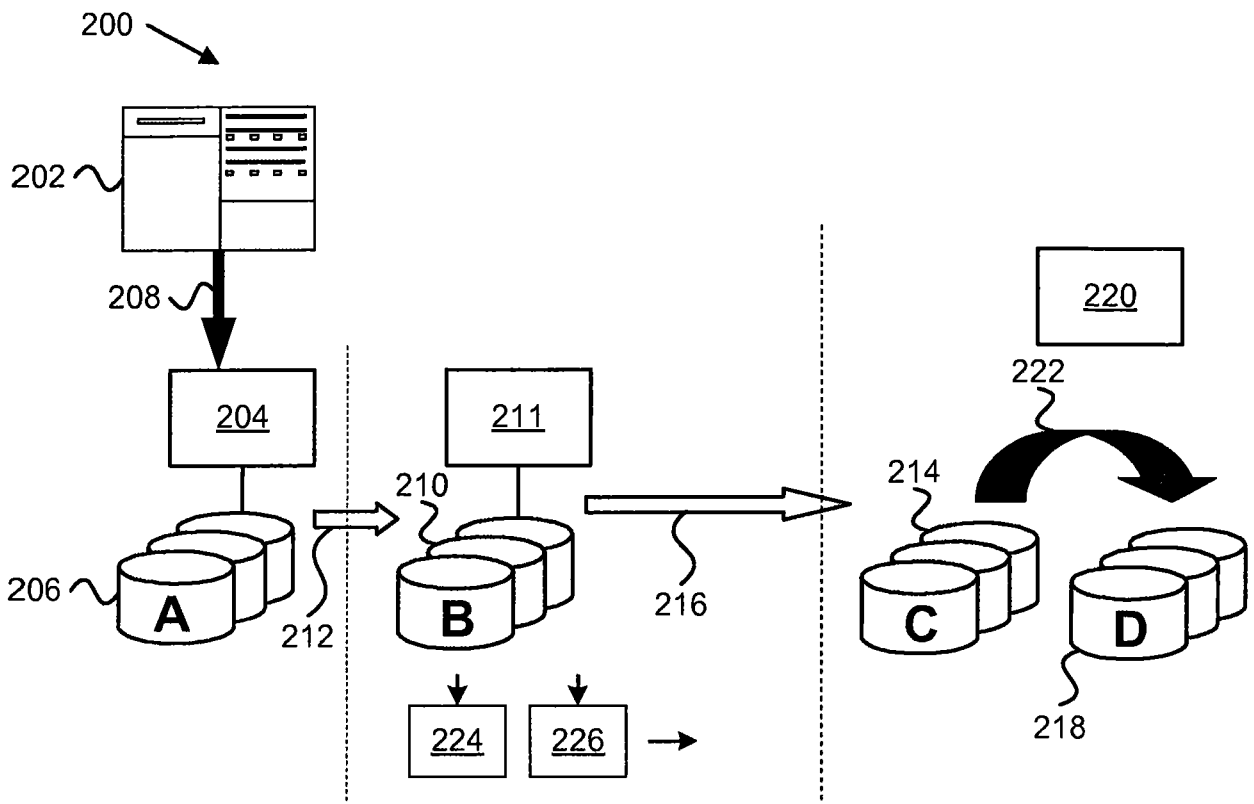


图 2

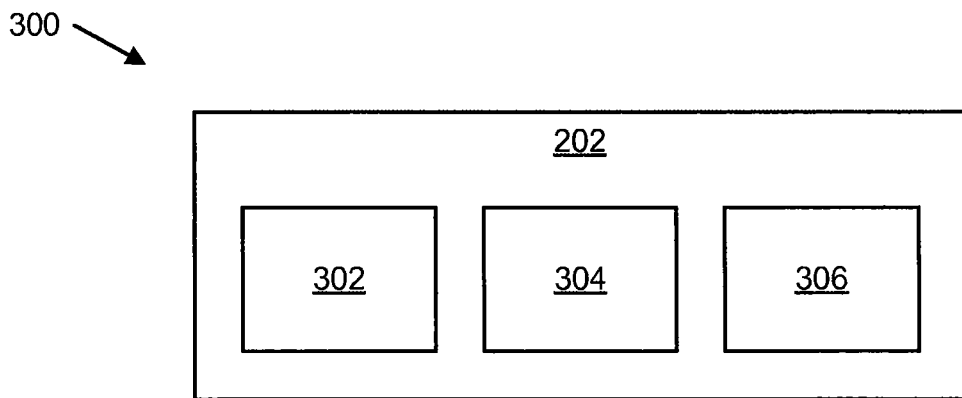


图 3

400 →

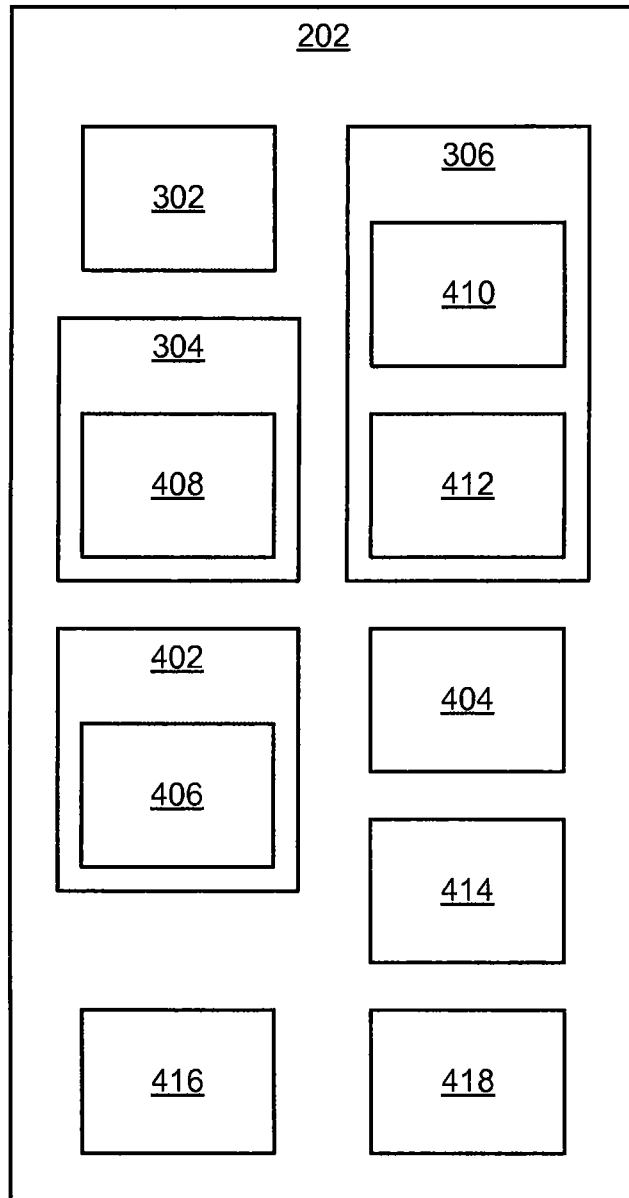


图 4

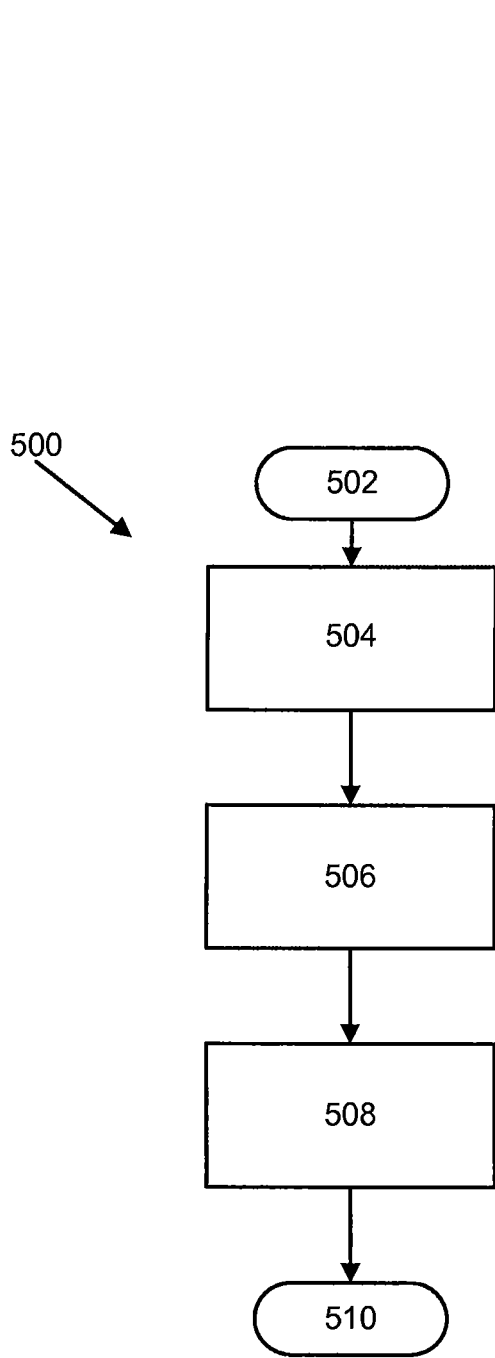


图 5

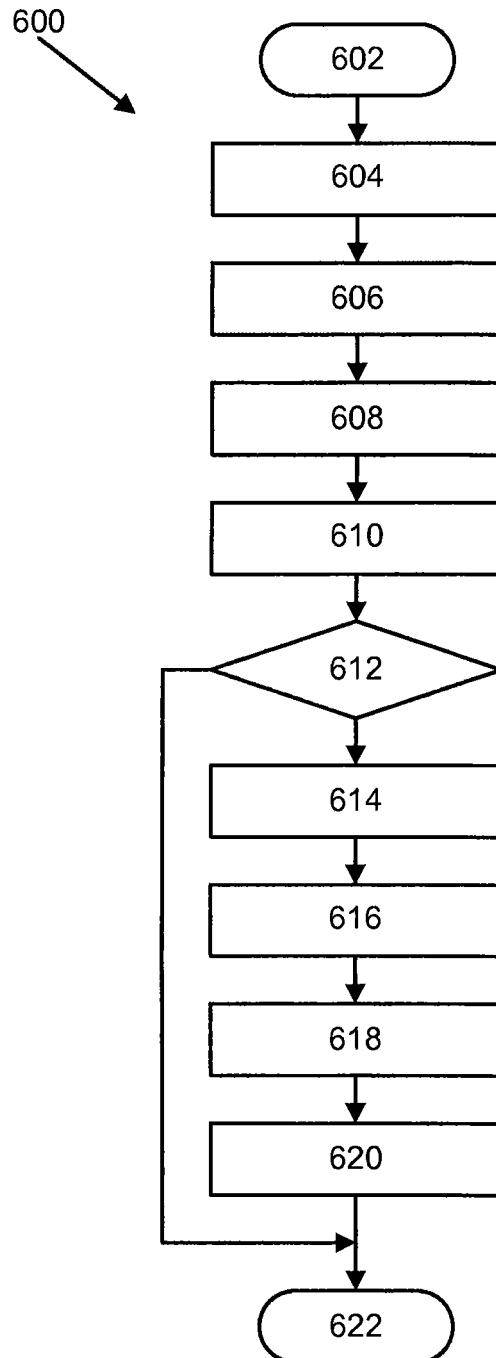


图 6

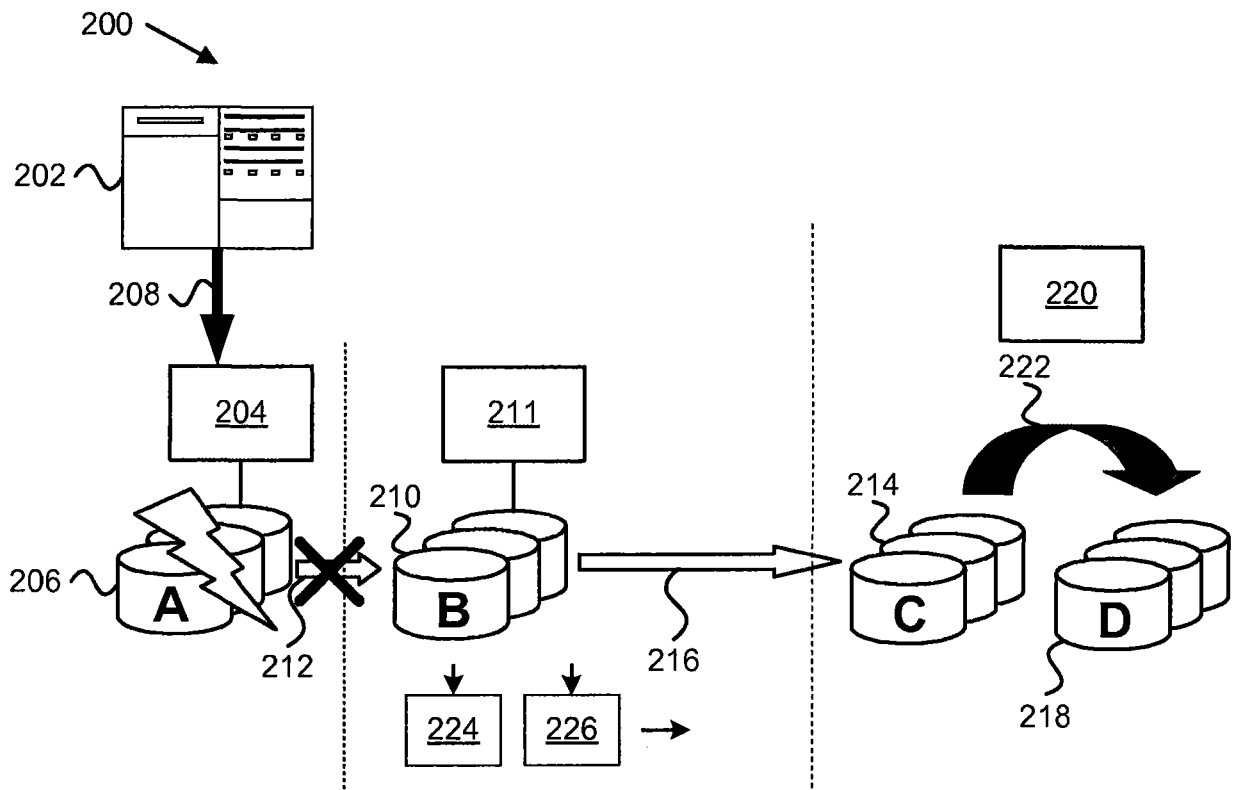


图 7A

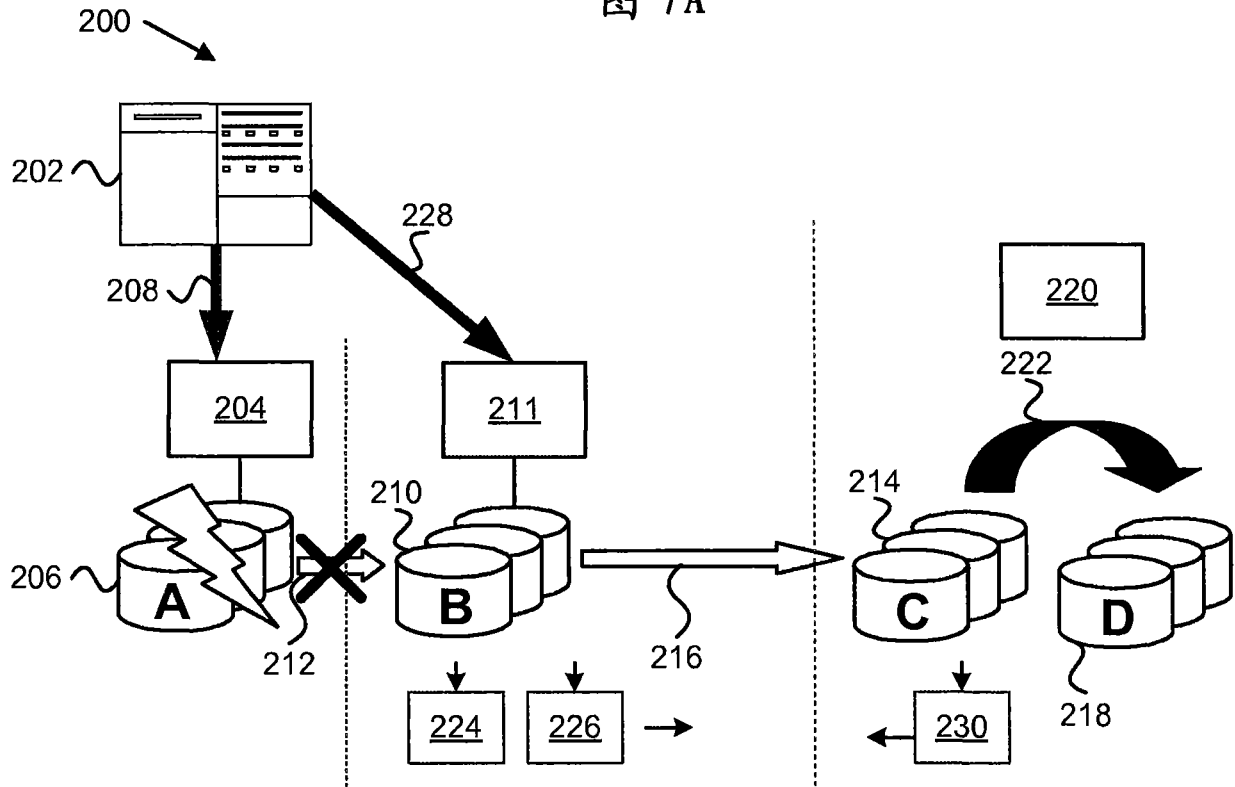


图 7B

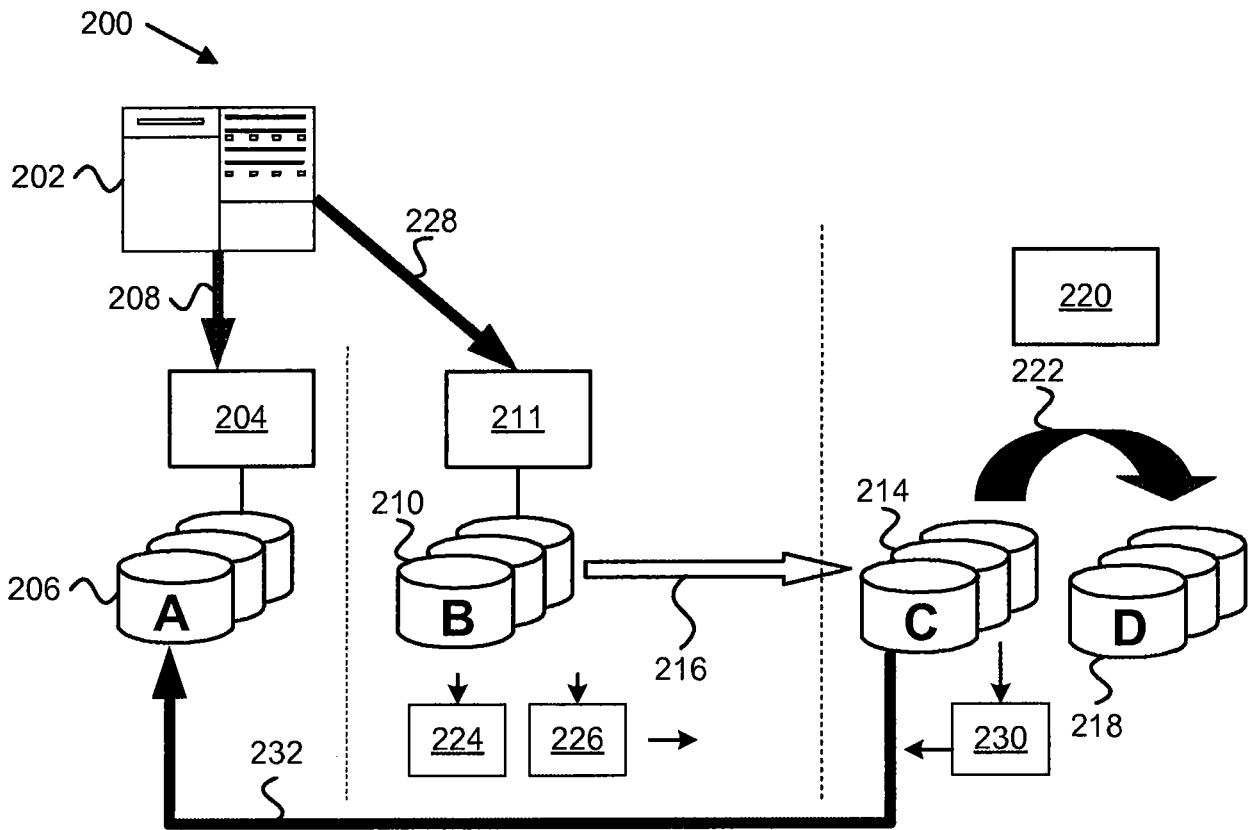


图 7C

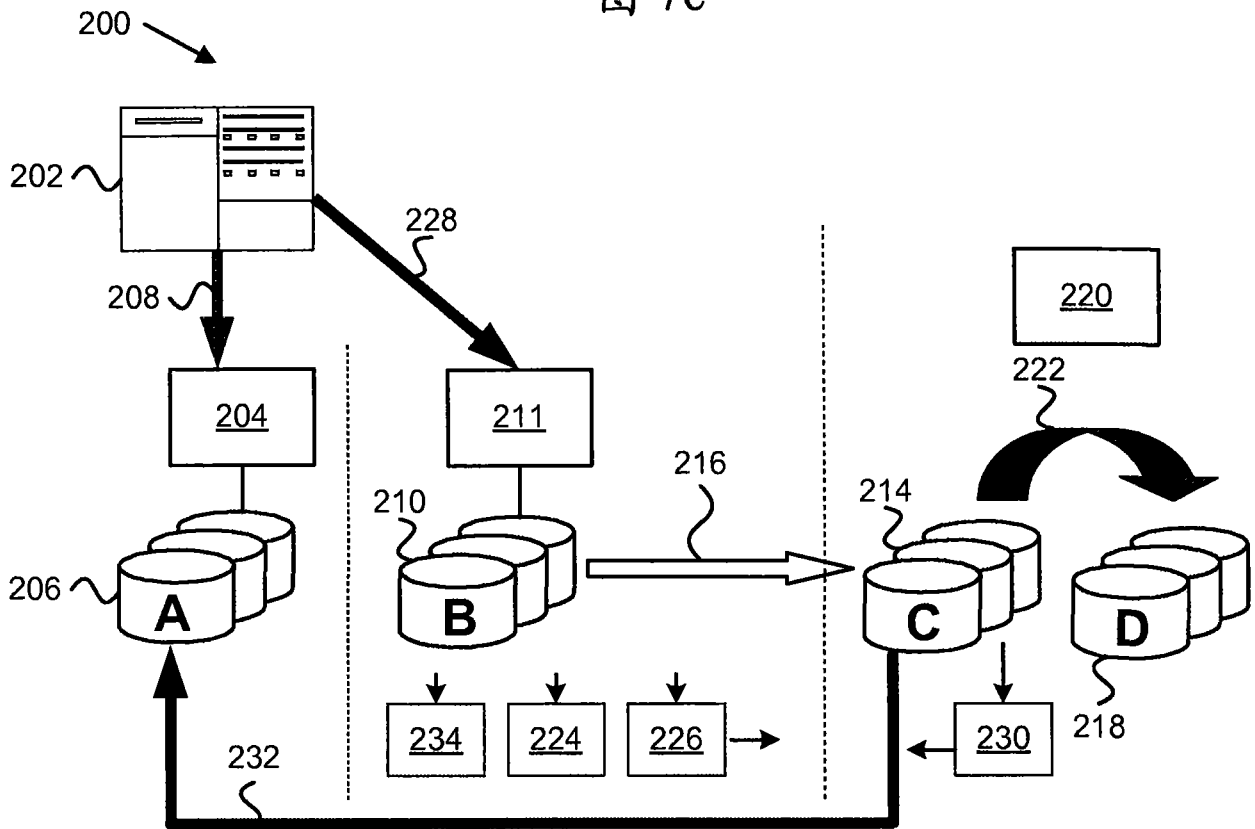


图 7D



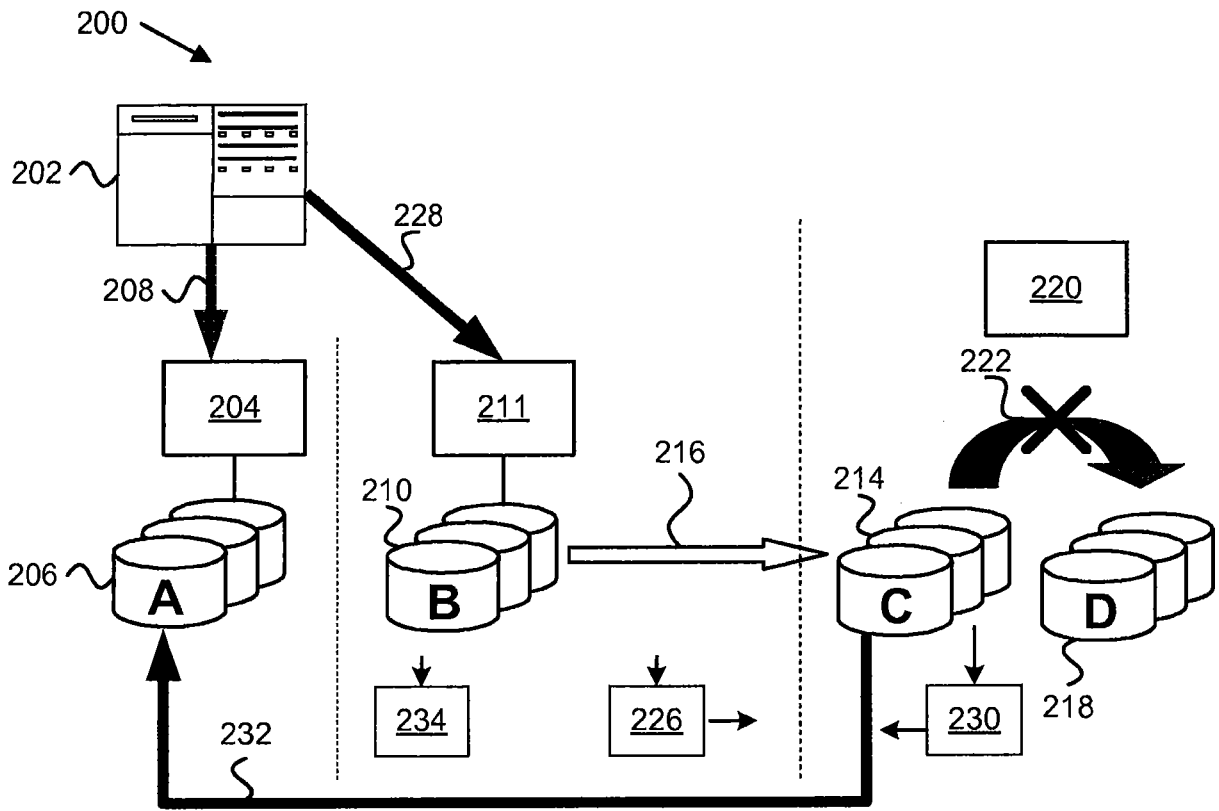


图 7E

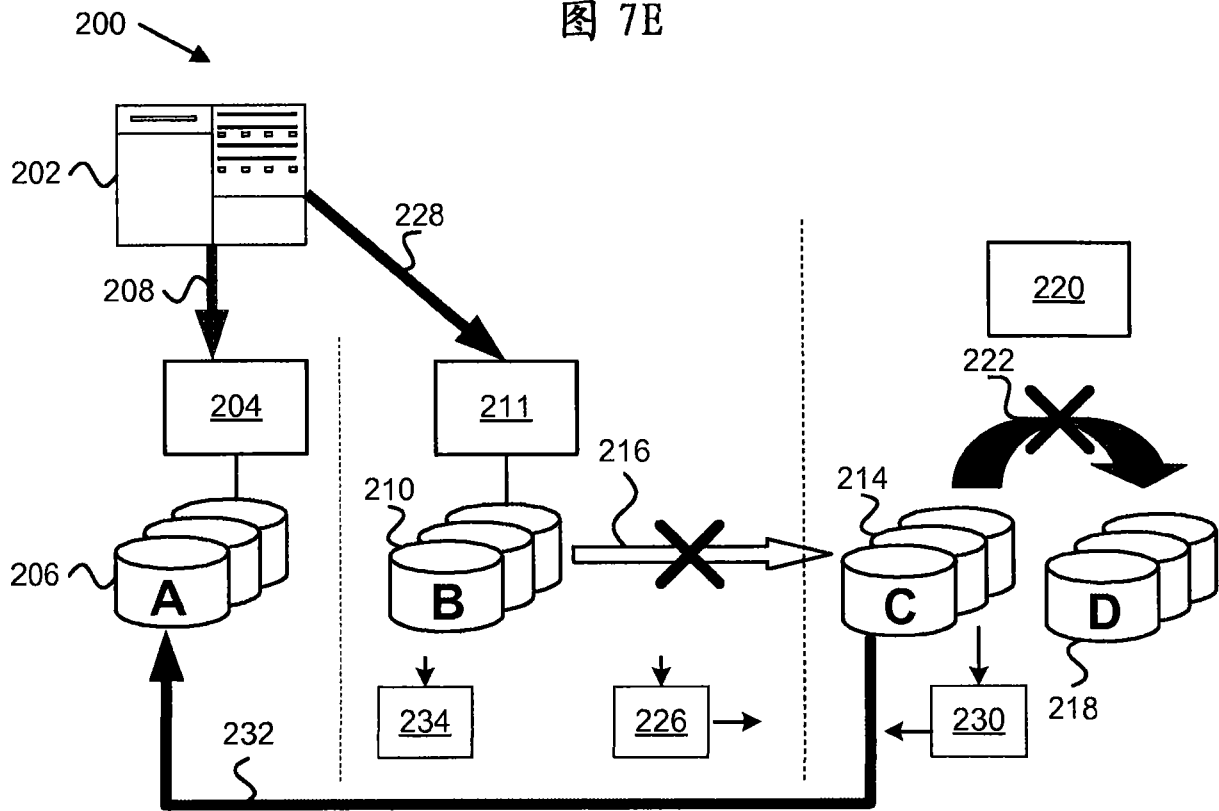


图 7F

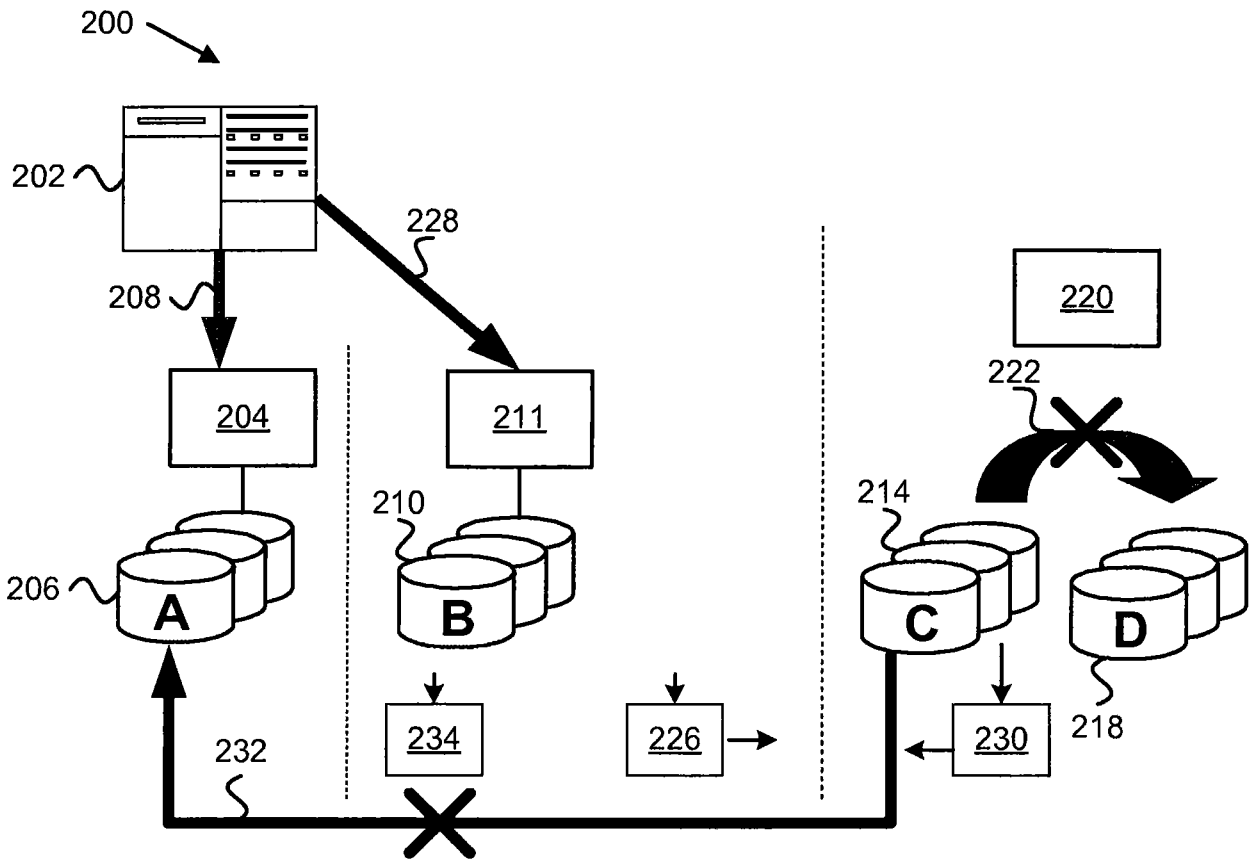


图 7G

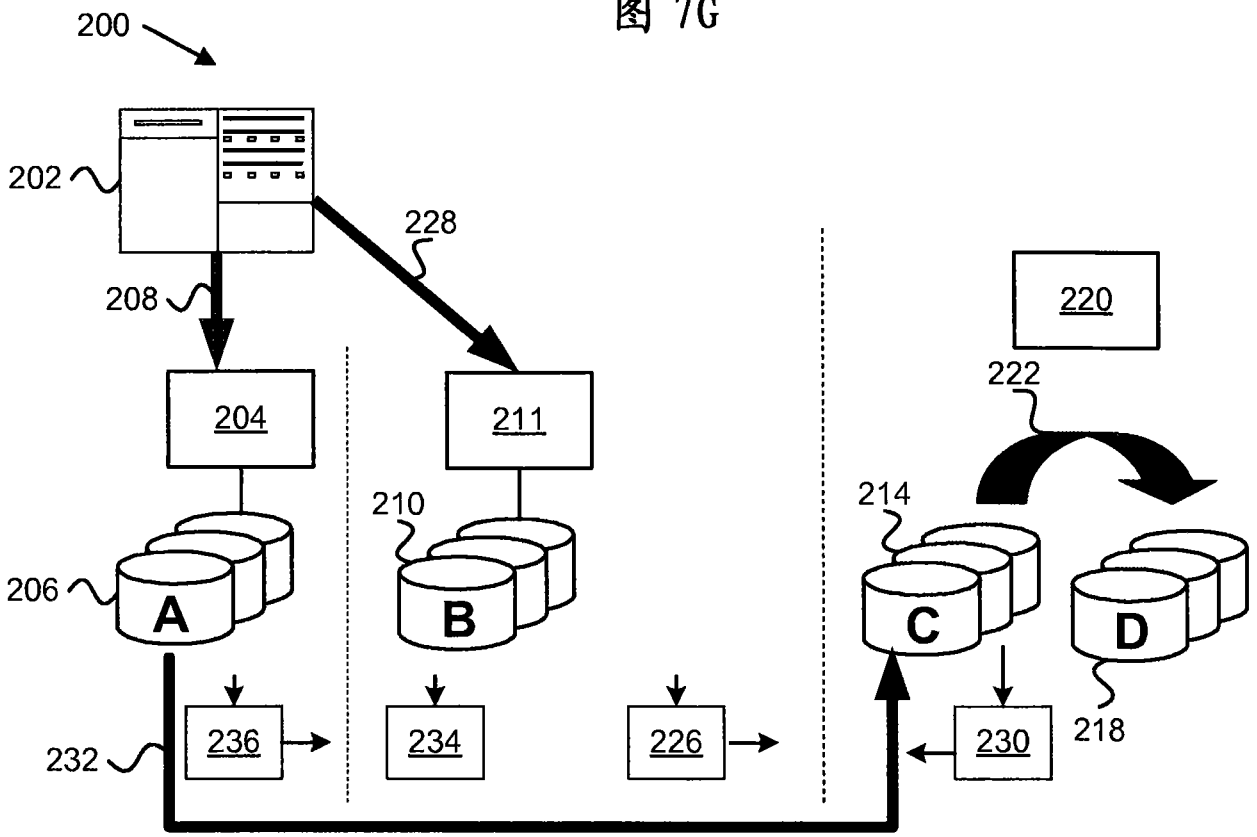


图 7H

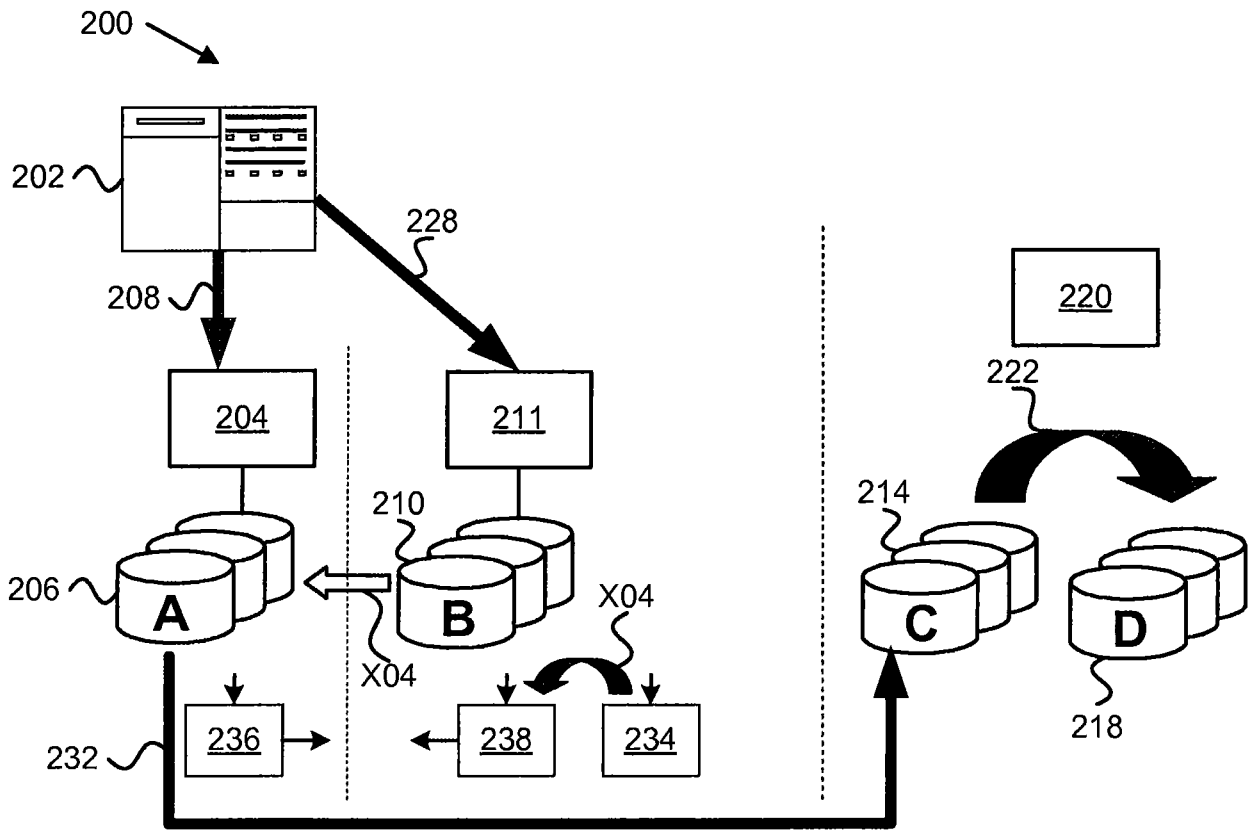


图 7I

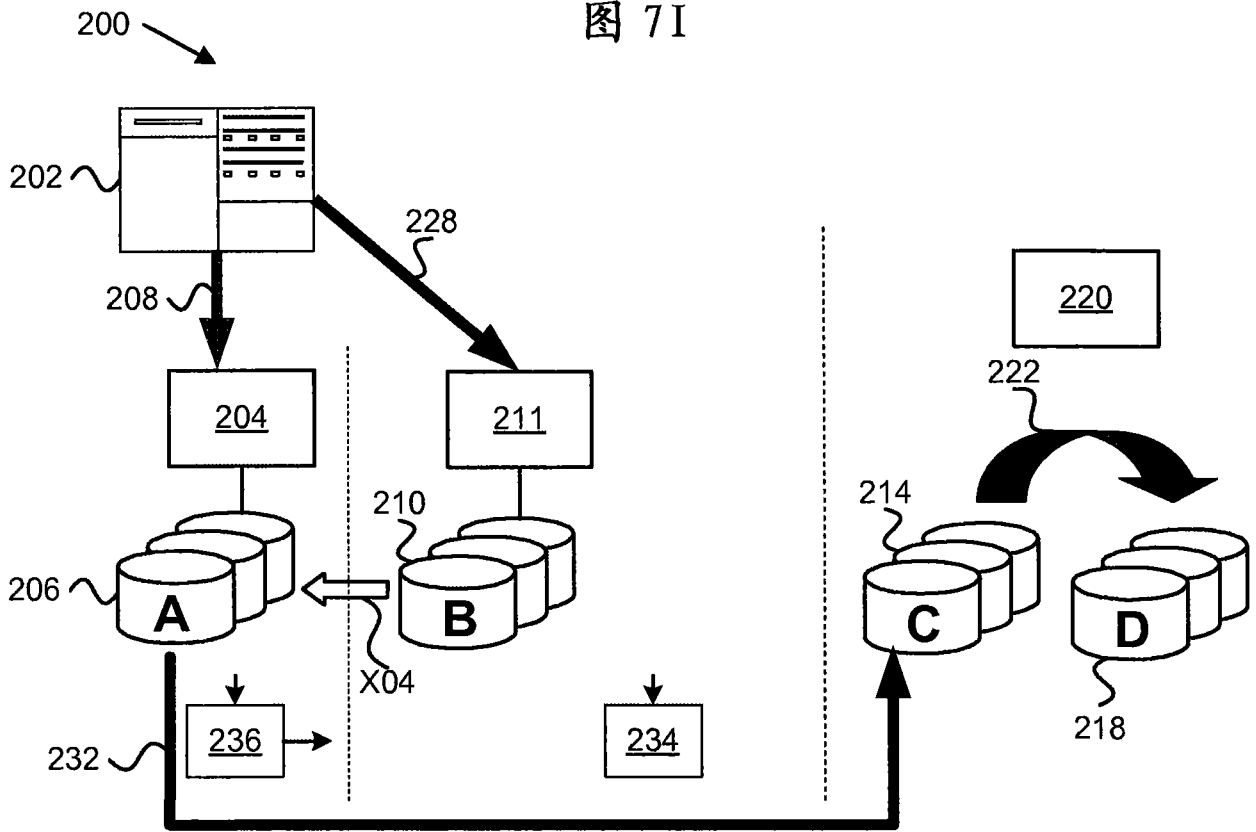


图 7J

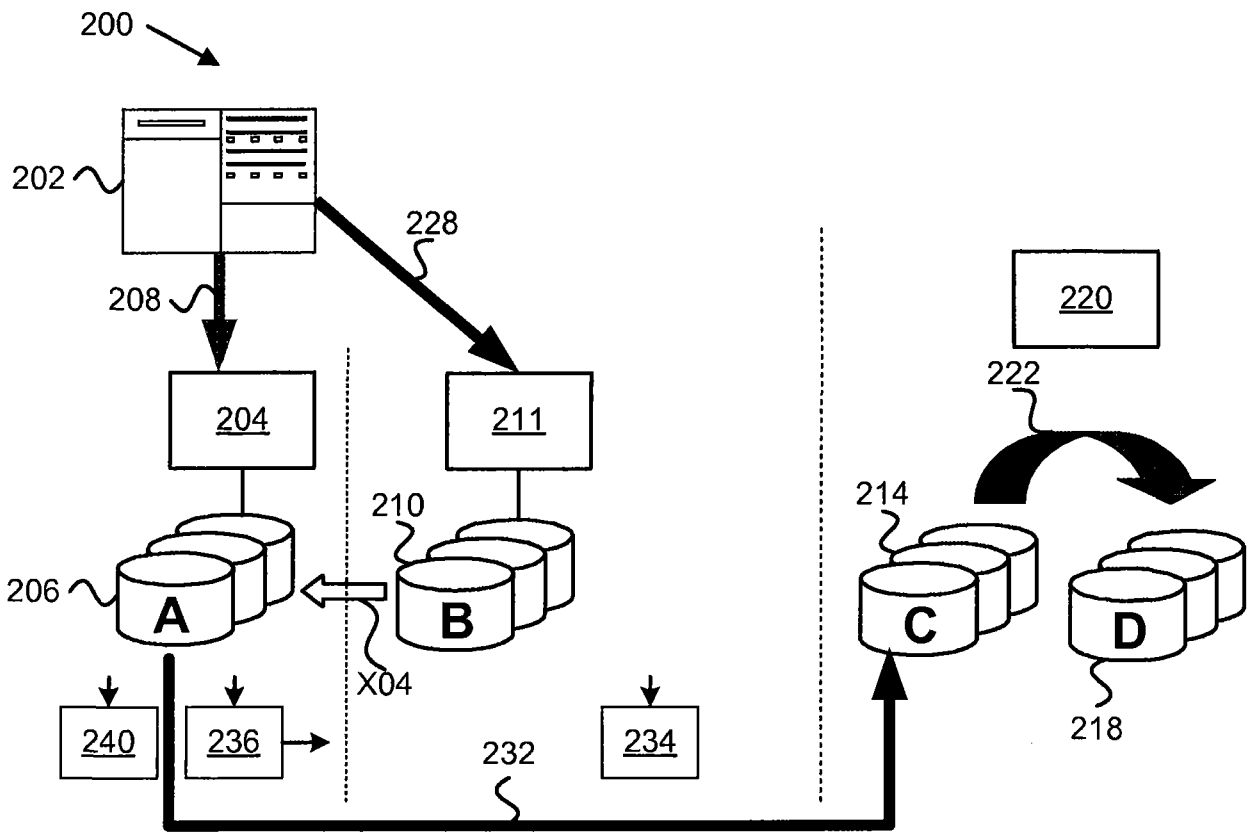


图 7K

1. 一种用于系统中的计算机程序产品, 所述系统包括:

第一存储卷, 其在出现故障之前被指定为主存储卷;

第二存储卷, 其包括所述故障之前的所述第一存储卷的数据的同步镜像副本; 以及

第三存储卷, 其包括所述故障之前的所述第二存储卷的数据的异步镜像副本, 所述计算机程序产品包括计算机可读介质, 所述计算机可读介质具有针对激活同步镜像作为主存储卷而编程的计算机可用程序代码, 所述计算机程序产品的操作包括:

引导所述第三存储卷将从所述第二存储卷发送到所述第三存储卷的更新存储在所述第一存储卷上以响应所述第一存储卷在所述故障之后变得可运行;

终止将更新从所述第二存储卷发送到所述第三存储卷以响应所述第一存储卷变得与所述第三存储卷基本一致; 以及

将所述第二存储卷接收的更新同步地存储在所述第一存储卷上并且将所述第一存储卷接收的更新异步地存储在所述第三存储卷上, 以响应在所述第一和第三存储卷基本一致之后终止将更新从所述第二存储卷发送到所述第三存储卷。

2. 如权利要求 1 中所述的计算机程序产品, 还包括指定所述第二存储卷作为主存储卷以响应在所述第一存储卷处检测到故障, 其中所述第一存储卷被指定为所述主存储卷, 所述第二存储卷包括所述故障之前的所述第一存储卷的数据的同步镜像副本, 并且所述第三存储卷包括所述故障之前的所述第二存储卷的数据的异步镜像副本。

3. 如权利要求 2 中所述的计算机程序产品, 还包括终止所述第二存储卷上的存储关系, 所述存储关系包括所述第二存储卷用于同步地存储来自所述第一存储卷的数据的信息。

4. 如权利要求 2 中所述的计算机程序产品, 还包括:

指定所述第一存储卷作为所述主存储卷,以响应将所述第二存储卷接收的更新同步地存储在所述第一存储卷上并且将所述第一存储卷接收的更新异步地存储在所述第三存储卷上;

引导所述第三存储卷将从所述第一存储卷发送到所述第三存储卷的更新存储在所述第二存储卷上以响应所述第一存储卷被指定为所述主存储卷;

终止将更新从所述第一存储卷发送到所述第三存储卷以响应所述第二存储卷变得与所述第三存储卷基本一致;以及

将所述第一存储卷接收的更新同步地存储在所述第二存储卷上并且将所述第二存储卷接收的更新异步地存储在所述第三存储卷上,以响应终止将更新从所述第一存储卷发送到所述第三存储卷。

5. 如权利要求1中所述的计算机程序产品,其中引导所述第三存储卷将从所述第二存储卷发送到所述第三存储卷的更新存储在所述第一存储卷上进一步包括:引导所述第三存储卷创建存储关系并且跟踪所述第三存储卷后续接收的更新,所述存储关系包括所述第三存储卷用于将更新异步地存储在所述第一存储卷上的信息。

6. 如权利要求1中所述的计算机程序产品,其中终止将更新从所述第二存储卷发送到所述第三存储卷进一步包括以下操作之一:

a) 从所述第二存储卷和所述第三存储卷删除存储关系,所述存储关系包括所述第二存储卷用于将数据存储在所述第三存储卷上的信息;以及

b) 跟踪在终止将更新从所述第二存储卷发送到所述第三存储卷之后所述第二存储卷接收的更新。

7. 如权利要求1中所述的计算机程序产品,其中将更新异步地存储在所述第三存储卷上进一步包括:使用闪速复制过程从所述第三存储卷更新第四存储卷,其中在所述第四存储卷上按时间顺序组织更新。

8. 如权利要求1中所述的计算机程序产品,其中将所述第二存储卷接收的更新同步地存储在所述第一存储卷上进一步包括:引导所述第二存储卷在所述第二存储卷上创建存储关系,所述存储关系包括所述第二存储卷

用于将数据同步地存储在所述第一存储卷上的信息。

9. 一种用于激活同步镜像作为主存储卷的装置, 所述装置包括:

主交换模块, 其配置为指定第二存储卷作为主存储卷以响应在第一存储卷处检测到故障, 其中所述第一存储卷被指定为所述主存储卷, 所述第二存储卷包括所述故障之前的所述第一存储卷的数据的同步镜像副本, 并且第三存储卷包括所述故障之前的所述第二存储卷的数据的异步镜像副本;

C-A 故障回复模块, 其配置为引导所述第三存储卷将从所述第二存储卷发送到所述第三存储卷的更新存储在所述第一存储卷上以响应所述第一存储卷在所述故障之后变得可运行;

B-C 终止模块, 其配置为终止将更新从所述第二存储卷发送到所述第三存储卷以响应所述第一存储卷变得与所述第三存储卷基本一致; 以及

B-A-C 存储模块, 其配置为将所述第二存储卷接收的更新同步地存储在所述第一存储卷上并且将所述第一存储卷接收的更新异步地存储在所述第三存储卷上, 以响应所述 B-C 终止模块终止将更新从所述第二存储卷发送到所述第三存储卷。

10. 如权利要求 9 中所述的装置, 还包括 C-A 强制故障转移模块, 其配置为引导所述第三存储卷创建存储关系并且跟踪在创建所述存储关系之后所述第三存储卷接收的更新, 所述存储关系包括所述第三存储卷用于将更新异步地存储在所述第一存储卷上的信息。

11. 如权利要求 9 中所述的装置, 其中所述主交换模块还包括 A-B 终止模块, 其配置为终止所述第二存储卷和所述第三存储卷上的存储关系以响应在所述第一存储卷处检测到故障并且使所述第二存储卷成为主存储卷, 所述存储关系包括所述第二存储卷用于同步地存储来自所述第一存储卷的数据的信息。

12. 如权利要求 9 中所述的装置, 其中所述 B-C 终止模块还配置为从所述第二存储卷删除存储关系, 所述存储关系包括所述第二存储卷用于将数据存储在第一存储卷上的信息。

13. 如权利要求 9 中所述的装置, 其中所述 B-C 终止模块还包括跟踪更改模块, 其配置为跟踪在终止将更新从所述第二存储卷发送到所述第三存储卷之后所述第二存储卷接收的更新。

14. 如权利要求 9 中所述的装置, 还包括第四存储卷, 并且其中所述第三存储卷被配置为使用闪速复制过程将更新存储到所述第四存储卷, 其中在所述第四存储卷上按时间顺序组织所述更新。

15. 如权利要求 9 中所述的装置, 其中所述 B-A-C 存储模块还包括以下各项之一:

A-C 跟踪模块, 其配置为引导所述第一存储卷跟踪所述第一存储卷接收的要复制到所述第三存储卷的更新; 以及

B-A 增量重新同步模块, 其配置为引导所述第二存储卷在所述第二存储卷上创建存储关系, 所述存储关系包括所述第二存储卷用于将数据同步地存储在所述第一存储卷上的信息。