



(12) 发明专利

(10) 授权公告号 CN 101578586 B

(45) 授权公告日 2012. 12. 12

(21) 申请号 200880001877. 5

(51) Int. Cl.

(22) 申请日 2008. 01. 03

G06F 11/20(2006. 01)

(30) 优先权数据

(56) 对比文件

11/622, 996 2007. 01. 12 US

US 2005081091 A1, 2005. 04. 14, 说明书第 1 页第 2 段至第 2 页第 4 段、权利要求 1, 12, 15.

(85) PCT 申请进入国家阶段日

CN 1581091 A, 2005. 02. 16, 全文.

2009. 07. 08

US 6694447 B1, 2004. 02. 17, 说明书第 6 栏第 19 行至第 7 栏第 62 行、附图 4 - 5.

(86) PCT 申请的申请数据

PCT/EP2008/050032 2008. 01. 03

US 6694447 B1, 2004. 02. 17, 说明书第 6 栏第 19 行至第 7 栏第 62 行、附图 4 - 5.

(87) PCT 申请的公布数据

W02008/084007 EN 2008. 07. 17

US 2005081091 A1, 2005. 04. 14, 说明书第 1 页第 2 段至第 2 页第 4 段、权利要求 1, 12, 15.

(73) 专利权人 国际商业机器公司

审查员 董洪梅

地址 美国纽约

(72) 发明人 B·C·比尔德斯雷

R·F·巴尔特福伊

G·E·麦克布赖德

(74) 专利代理机构 北京市中咨律师事务所

11247

代理人 于静 杨晓光

权利要求书 5 页 说明书 9 页 附图 9 页

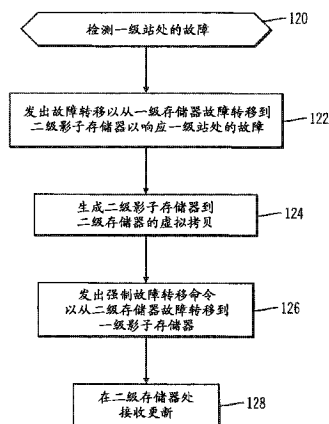
(54) 发明名称

在故障转移和故障回复环境中使用虚拟拷贝

(57) 摘要

在一种方法中,在故障转移和故障回复环境中使用虚拟拷贝,在系统运行期间,从一级站处的一级第一存储器向二级站处的二级第一存储器复制更新。在一级和二级站中的至少一个处维护第二存储器。在所述一级站出现故障后执行从所述一级站到所述二级站的故障转移。在所述一级站恢复后使用至少一个第二存储器来使所述二级站与所述一级站同步。响应于所述一级站的恢复,向所述至少一个第二存储器或从所述至少一个第二存储器仅复制在所述故障转移期间对所述二级站做出的更新。

CN 101578586 B



1. 一种用于同步一级站与二级站之间的数据的方法,包括:

在系统运行期间从所述一级站处的一级第一存储器向所述二级站处的二级第一存储器复制更新;

在所述二级站中维护二级第二存储器;

在所述一级站出现故障后执行从所述一级站到所述二级站的故障转移;

在所述一级站恢复后使用所述二级第二存储器来使所述二级站与所述一级站同步;以及

响应于所述一级站的恢复,从所述二级第二存储器复制在故障转移期间更新的数据单元到所述一级第一存储器,

所述方法还包括:

响应于所述故障转移,生成所述二级第一存储器到所述二级第二存储器的虚拟拷贝;以及

响应于生成所述虚拟拷贝,发出从所述二级第二存储器到所述一级第一存储器的故障转移命令,其中在所述故障转移后相对于所述二级第二存储器来执行读取和写入操作。

2. 根据权利要求1的方法,其中生成所述虚拟拷贝包括:

在所述二级第一和第二存储器中的数据单元之间建立时间点复制关系;以及

从所述二级第一存储器向所述二级第二存储器复制在所述时间点复制关系中指示的数据,其中如果在将所述二级第一存储器处的数据单元复制到所述二级第二存储器之前已更新了所述二级第二存储器中的相应数据单元,则不将所述二级第一存储器处的所述数据单元复制到所述二级第二存储器。

3. 根据权利要求1的方法,还包括:

发出强制故障转移命令以从所述二级第二存储器故障转移到所述一级第一存储器,从而导致记录在所述故障转移期间对所述二级第二存储器的更新;以及

从所述二级第二存储器向所述一级第一存储器发出强制故障回复命令,以导致在所述强制故障转移之后发生的所记录的对所二级第二存储器的更新被复制到所述一级第一存储器。

4. 根据权利要求1的方法,其中重写所述一级第一存储器中的在故障转移之前被更新并且未被复制到所述二级第一存储器的任何数据单元中的更新发生在从所述二级第一存储器向所述一级第一存储器复制数据单元之后。

5. 根据权利要求4的方法,还包括:

从所述一级第一存储器向所述二级第一存储器发出故障转移命令以响应所述一级站处的故障,其中在执行所述故障转移命令之后生成所述虚拟拷贝;以及

从所述二级第一存储器向所述一级第一存储器发出故障回复命令以响应所述一级站的恢复,以导致从所述二级第一存储器复制所述数据单元,以便重写在到所述二级站的故障转移之前发生并且未被复制到所述二级第一存储器的对所一级第一存储器中的数据单元的更新。

6. 根据权利要求1的方法,其中在所述一级站处存在一级第二存储器,其中在所述一级站出现故障后执行从所述一级站到所述二级站的故障转移包括执行从所述一级第一存储器到所述二级第一存储器的第一故障转移,并且所述方法还包括:

生成所述二级第一存储器到所述二级第二存储器的虚拟拷贝以响应所述故障转移；

执行从所述二级第二存储器到所述一级第二存储器的第二故障转移，其中在所述故障转移后相对于所述二级第二存储器来执行读取和写入操作；

生成所述一级第一存储器到所述一级第二存储器的虚拟拷贝以响应所述一级站的恢复；以及

仅将所述故障转移之后对所述二级第二存储器做出的更新复制到所述一级第二存储器，以响应到所述一级第二存储器的虚拟拷贝和响应所述一级站的恢复。

7. 根据权利要求 6 的方法，还包括：

从所述二级第二存储器向所述一级第二存储器发出强制故障转移命令以执行所述二级第二存储器的故障转移并且进一步导致记录在所述故障转移之后对所述二级第二存储器做出的更新；以及

从所述二级第二存储器向所述一级第二存储器发出强制故障回复命令以导致在所述故障转移之后的所记录的对所述二级第二存储器的更新被复制到所述一级第二存储器。

8. 根据权利要求 6 的方法，还包括：

在所述一级站恢复后，从所述二级第一存储器向所述一级第一存储器复制数据单元，以重写在所述故障转移之前发生并且未被复制到所述二级第二存储器的对所述一级第一存储器中的数据单元的更新，其中发生所述一级第一存储器到所述一级第二存储器的虚拟拷贝操作，以响应在所述一级站恢复后，从所述二级第一存储器向所述一级第一存储器复制数据单元。

9. 根据权利要求 8 的方法，还包括：

从所述一级第一存储器向所述二级第一存储器发出故障转移命令以响应所述一级站处的故障，其中在执行所述故障转移命令之后生成从所述二级第一存储器到所述二级第二存储器的虚拟拷贝；以及

从所述二级第一存储器向所述一级第一存储器发出故障回复命令以响应所述一级站的恢复，以导致从所述二级第一存储器复制所述数据单元，以便重写在到所述二级第一存储器的故障转移之前发生并且未被复制到所述二级站的对所述一级第一存储器中的数据单元的更新。

10. 根据权利要求 1 的方法，其中在所述一级站处存在一级第二存储器，其中执行从所述一级站到所述二级站的故障转移包括：从所述一级第一存储器故障转移到所述二级第一存储器；其中复制对所述二级站做出的更新以响应所述恢复包括：将系统运行期间在从一级站处的一级存储器故障转移到二级站之后对所述二级第一存储器做出的更新复制到所述一级第一存储器；还包括：

生成所述一级第一存储器到所述一级站处的一级第二存储器的虚拟拷贝，以响应将所述更新复制到所述一级存储器。

11. 根据权利要求 10 的方法，还包括：

将生成所述虚拟拷贝期间对所述二级存储器做出的更新复制到所述一级第二存储器以响应到所述一级第二存储器的虚拟拷贝。

12. 根据权利要求 11 的方法，还包括：

从所述二级第一存储器向所述一级第一存储器发出故障回复命令以响应所述一级站

的恢复,以导致从所述二级第一存储器复制所述数据单元,以便重写在到所述二级第一存储器的故障转移之前发生并且未被复制到所述二级第一存储器的对所述一级第一存储器中的数据单元的更新;以及

发出强制故障回复命令以响应完成从所述一级第一存储器到所述一级第二存储器的虚拟拷贝,以便在生成所述虚拟拷贝期间执行所述更新的复制。

13. 一种用于同步一级站与二级站之间的数据的系统,其中所述一级站具有一级第一存储器,其中所述二级站具有二级第一存储器,并且其中在所述一级和二级站中的至少一个处维护第二存储器,包括:

复制管理器组件,其进一步被配置为包括:

使在系统运行期间从所述一级站处的一级第一存储器向所述二级站处的二级第一存储器复制更新的装置;

使在所述二级站中维护二级第二存储器的装置;

使在所述一级站出现故障后执行从所述一级站到所述二级站的故障转移的装置;

在所述一级站恢复后使用所述二级第二存储器来使所述二级站与所述一级站同步的装置;以及

响应于所述一级站的恢复,从所述二级第二存储器复制在故障转移期间更新的数据单元到所述一级第一存储器的装置,

所述复制管理组件还包括:

响应于所述故障转移,生成所述二级第一存储器到所述二级第二存储器的虚拟拷贝的装置;以及

响应于生成所述虚拟拷贝,发出从所述二级第二存储器到所述一级第一存储器的故障转移命令的装置,其中在所述故障转移后相对于所述二级第二存储器来执行读取和写入操作。

14. 根据权利要求 13 的系统,其中生成所述虚拟拷贝的装置包括:

在所述二级第一和第二存储器中的数据单元之间建立时间点复制关系的装置;以及

从所述二级第一存储器向所述二级第二存储器复制在所述时间点复制关系中指示的数据的装置,其中如果在将所述二级第一存储器处的数据单元复制到所述二级第二存储器之前已更新了所述二级第二存储器中的相应数据单元,则不将所述二级第一存储器处的所述数据单元复制到所述二级第二存储器。

15. 根据权利要求 13 的系统,还包括:

发出强制故障转移命令以从所述二级第二存储器故障转移到所述一级第一存储器,从而导致记录在所述故障转移期间对所述二级第二存储器的更新的装置;以及

从所述二级第二存储器向所述一级第一存储器发出强制故障回复命令,以导致在所述强制故障转移之后发生的所记录的对所述二级第二存储器的更新被复制到所述一级第一存储器的装置。

16. 根据权利要求 13 的系统,

其中重写所述一级第一存储器中的在故障转移之前被更新并且未被复制到所述二级第一存储器的任何数据单元中的更新发生在从所述二级第一存储器向所述一级第一存储器复制数据单元之后。

17. 根据权利要求 16 的系统,还包括:

从所述一级第一存储器向所述二级第一存储器发出故障转移命令以响应所述一级站处的故障的装置,其中在执行所述故障转移命令之后生成所述虚拟拷贝;以及

从所述二级第一存储器向所述一级第一存储器发出故障回复命令以响应所述一级站的恢复,以导致从所述二级第一存储器复制所述数据单元,以便重写在到所述二级站的故障转移之前发生并且未被复制到所述二级第一存储器的对所述一级第一存储器中的数据单元的更新的装置。

18. 根据权利要求 13 的系统,其中在所述一级站处存在一级第二存储器,其中使在所述一级站出现故障后执行从所述一级站到所述二级站的故障转移的装置进一步执行从所述一级第一存储器到所述二级第一存储器的第一故障转移,所述系统还包括:

生成所述二级第一存储器到所述二级第二存储器的虚拟拷贝以响应所述故障转移的装置;

执行从所述二级第二存储器到所述一级第二存储器的第二故障转移的装置,其中在所述故障转移后相对于所述二级第二存储器来执行读取和写入操作;

生成所述一级第一存储器到所述一级第二存储器的虚拟拷贝以响应所述一级站的恢复的装置;以及

仅将所述故障转移之后对所述二级第二存储器做出的更新复制到所述一级第二存储器,以响应到所述一级第二存储器的虚拟拷贝和响应所述一级站的恢复的装置。

19. 根据权利要求 18 的系统,还包括:

从所述二级第二存储器向所述一级第二存储器发出强制故障转移命令以执行所述二级第二存储器的故障转移并且进一步导致记录在所述故障转移之后对所述二级第二存储器做出的更新的装置;以及

从所述二级第二存储器向所述一级第二存储器发出强制故障回复命令以导致在所述故障转移之后的所记录的所述二级第二存储器的更新被复制到所述一级第二存储器的装置。

20. 根据权利要求 18 的系统,还包括:

在所述一级站恢复后,从所述二级第一存储器向所述一级第一存储器复制数据单元,以重写在所述故障转移之前发生并且未被复制到所述二级第二存储器的对所述一级第一存储器中的数据单元的更新的装置,其中发生所述一级第一存储器到所述一级第二存储器的虚拟拷贝操作,以响应在所述一级站恢复后,从所述二级第一存储器向所述一级第一存储器复制数据单元。

21. 根据权利要求 20 的系统,还包括:

从所述一级第一存储器向所述二级第一存储器发出故障转移命令以响应所述一级站处的故障的装置,其中在执行所述故障转移命令之后生成从所述二级第一存储器到所述二级第二存储器的虚拟拷贝;以及

从所述二级第一存储器向所述一级第一存储器发出故障回复命令以响应所述一级站的恢复,以导致从所述二级第一存储器复制所述数据单元,以便重写在到所述二级第一存储器的故障转移之前发生并且未被复制到所述二级站的对所述一级第一存储器中的数据单元的更新的装置。

22. 根据权利要求 13 的系统,其中在所述一级站处存在一级第二存储器,其中执行从所述一级站到所述二级站的故障转移的装置包括:从所述一级第一存储器故障转移到所述二级第一存储器的装置;其中复制对所述二级站做出的更新以响应所述恢复的装置包括:将系统运行期间在从一级站处的一级存储器故障转移到二级站之后对所述二级第一存储器做出的更新复制到所述一级第一存储器的装置;

所述系统还包括:生成所述一级第一存储器到所述一级站处的一级第二存储器的虚拟拷贝,以响应将所述更新复制到所述一级存储器的装置。

23. 根据权利要求 22 的系统,还包括:

将生成所述虚拟拷贝期间对所述二级存储器做出的更新复制到所述一级第二存储器以响应到所述一级第二存储器的虚拟拷贝的装置。

24. 根据权利要求 23 的系统,还包括:

从所述二级第一存储器向所述一级第一存储器发出故障回复命令以响应所述一级站的恢复,以导致从所述二级第一存储器复制所述数据单元,以便重写在到所述二级第一存储器的故障转移之前发生并且未被复制到所述二级第一存储器的对所述一级第一存储器中的数据单元的更新的装置;以及

发出强制故障回复命令以响应完成从所述一级第一存储器到所述一级第二存储器的虚拟拷贝,以便在生成所述虚拟拷贝期间执行所述更新的复制的装置。

在故障转移和故障回复环境中使用虚拟拷贝

技术领域

[0001] 本发明涉及在故障转移和故障回复环境中使用虚拟拷贝的方法、系统和程序。

背景技术

[0002] 在典型灾难恢复解决方案中,数据容纳在一级站处以及容纳在一个或多个二级站处。这些二级站维护数据的同步副本,以便在一级站出现灾难的情况下尽可能减少数据丢失。如果发生灾难,处理或者“故障转移”到二级站之一,或者将数据从二级站复制回一级站。为了实现灾难恢复,二级站典型地在地理上远离一级站,即,位于不同的城市、州等,以便同一灾难不会影响两个站点。

[0003] 灾难恢复系统典型地解决两类故障:在单个时刻的意外灾难性故障和在一段时间期间的数据丢失。在第二类渐进的灾难中,对卷的更新可能丢失。对于任意一类故障,都可在远程位置提供数据的副本。典型地在应用系统向一级站处的一级存储设备写入新的数据时生成此类双重副本或影子副本。

[0004] 在镜像备份系统中,在卷对中维护数据。卷对包括一级存储设备中的卷和二级存储设备中的相应卷,后者包括在一级卷中维护的数据的一致副本。典型地,将在一级存储控制单元中维护所述对的一级卷,并且在位于不同于所述一级存储控制单元的物理位置处的二级存储控制单元中维护所述对的二级卷。存储控制单元是物理硬件单元,其包括与一个或多个存储设备集成的存储服务器以向主机提供存储能力。存储服务器是通过提供一个或多个逻辑子系统的功能来提供一个或多个存储设备与主机之间的接口的物理单元。存储服务器可以提供存储设备所未提供的功能。存储服务器包括一个或多个存储设备群集。可以提供一级存储控制单元以控制对一级存储器的访问,以及提供二级存储控制单元以控制对二级存储器的访问。

[0005] 当使用两个在地理上分散的服务器场来远程镜像数据以用于灾难恢复能力时,在站点之一已停机并且现在恢复后,将出现重新建立镜像的性能问题。在此类情况下,两个站点之间的大部分数据是完全相同的,但在一个站点停机期间,少部分数据已在一个站点处被更改而未在另一站点处被更改。在历史上,为了重新建立站点之间的同步(重新建立镜像),一个站点被选择为是当前的,然后将所有数据复制到另一站点。由于要移动的数据量,此镜像操作是非常耗时的过程(在周的量级)。

[0006] 此外,在二级站处维护镜像副本的同时,客户可能希望生成二级镜像副本到二级虚拟副本的虚拟拷贝,然后脱离二级虚拟拷贝站运行生产以在虚拟拷贝上进行测试和练习,以便测试二级站的运行。

[0007] 在某些镜像实施方式中,用户可以使二级站处的二级卷镜像一级站处的一级卷的数据。在此情况下,在故障期间,二级卷用于生产和运行,并且在故障转移到二级卷期间记录更改。在一级卷的恢复期间,执行故障回复以仅复制在故障转移到一级卷之后对二级卷的更改。在实现中,用户可以创建二级卷的虚拟拷贝并在二级卷的虚拟拷贝上进行练习,并且仍然使用二级卷用于恢复目的。在此情况下,在恢复以后,仍从主二级卷恢复更新而不管

二级卷的虚拟拷贝。此实施方式要求恢复站具有两种配置,一种用于练习,另一种用于恢复。此类配置增加了复杂性并且增大了在需要恢复操作时引入错误的可能。

[0008] 在其他镜像实施方式中,为了从虚拟拷贝二级卷恢复,用户可能将二级卷的整个虚拟拷贝复制到一级卷。

[0009] 本领域中存在持续改进一级与二级站之间的故障和恢复处理的需要。

发明内容

[0010] 提供了一种方法、系统和制品,用于在故障转移和故障回复环境中使用虚拟拷贝。在系统运行期间,从一级站处的一级第一存储器向二级站处的二级第一存储器复制更新。在一级和二级站中的至少一个处维护第二存储器。在所述一级站出现故障后执行从所述一级站到所述二级站的故障转移。在所述一级站恢复后使用至少一个第二存储器来使所述二级站与所述一级站同步。响应于所述一级站的恢复,向所述至少一个第二存储器或从所述至少一个第二存储器仅复制在所述故障转移期间对所述二级站做出的更新。

[0011] 在其他实施例中,在所述二级站处存在二级第二存储器并且响应于所述故障转移,生成所述二级第一存储器到所述二级第二存储器的虚拟拷贝。响应于生成所述虚拟拷贝,执行从所述二级第二存储器到所述一级第一存储器的故障转移,其中在所述故障转移后相对于所述二级第二存储器来执行读取和写入操作,并且其中仅复制对所述二级站做出的更新包括:从所述一级第一存储器向所述二级第二存储器复制所述故障转移期间的更新。

[0012] 在其他实施例中,生成所述虚拟拷贝包括:在所述二级第一和第二存储器中的数据单元之间建立时间点复制关系;以及从所述二级第一存储器向所述二级第二存储器复制在所述时间点复制关系中指示的数据。如果在将所述二级第一存储器处的数据单元复制到所述二级第二存储器之前已更新了所述二级第二存储器中的相应数据单元,则不将所述二级第一存储器处的所述数据单元复制到所述二级第二存储器。

[0013] 在其他实施例中,从所述一级第一存储器处的所述二级第二存储器发出强制故障转移命令以执行所述故障转移,从而导致记录在所述故障转移期间对所述二级第二存储器的更新。从所述二级第二存储器向所述一级第一存储器发出强制故障回复命令,以导致在所述强制故障转移之后发生的所记录的所述二级第二存储器的更新被复制到所述一级第一存储器。

[0014] 在其他实施例中,在所述一级站恢复后,从所述二级第一存储器向所述一级第一存储器复制数据单元,以重写在到所述二级站的故障转移之前发生并且未被复制到所述二级第一存储器的对所述一级第一存储器中的数据单元的更新。向所述二级第二存储器或从所述二级第二存储器复制对所述一级第一存储器的更新发生在从所述二级第一存储器向所述一级第一存储器复制数据单元之后。

[0015] 在其他实施例中,从所述一级第一存储器向所述二级第一存储器发出故障转移命令以响应所述一级站处的故障,其中在执行所述故障转移命令之后生成所述虚拟拷贝。从所述二级第一存储器向所述一级第一存储器发出故障回复命令以响应所述一级站的恢复,以导致从所述二级第一存储器复制所述数据单元,以便重写在到所述二级站的故障转移之前发生并且未被复制到所述二级第一存储器的对所述一级第一存储器中的数据单元的更

新。

[0016] 在其他实施例中,在所述一级站处存在一级第二存储器并且在所述二级站处存在二级第二存储器,其中执行所述故障转移包括执行从所述一级第一存储器到所述二级第一存储器的第一故障转移。此外,生成所述二级第一存储器到所述二级第二存储器的虚拟拷贝以响应所述故障转移。执行从所述二级第二存储器到所述一级第二存储器的第二故障转移,其中在所述故障转移后相对于所述二级第二存储器来执行读取和写入操作。生成所述一级第一存储器到所述一级第二存储器的虚拟拷贝以响应所述一级站的恢复。仅将所述故障转移之后对所述二级第二存储器做出的更新复制到所述一级第二存储器,以响应到所述一级第二存储器的虚拟拷贝和响应所述一级站的恢复。

[0017] 在其他实施例中,从所述二级第二存储器向所述一级第二存储器发出强制故障转移命令以执行所述二级第二存储器的故障转移并且进一步导致记录在所述故障转移之后对所述二级第二存储器做出的更新。从所述二级第二存储器向所述一级第二存储器发出强制故障回复命令以导致在所述故障转移之后的所记录的对所二级第二存储器的更新被复制到所述一级第二存储器。

[0018] 在其他实施例中,在所述一级站恢复后,从所述二级第一存储器向所述一级第一存储器复制数据单元,以重写在所述故障转移之前发生并且未被复制到所述二级第二存储器的对所一级第一存储器中的数据单元的更新。发生所述一级第一存储器到所述一级第二存储器的虚拟拷贝操作,以响应在所述一级站恢复后,从所述二级第一存储器向所述一级第一存储器复制数据单元。

[0019] 在其他实施例中,从所述一级第一存储器向所述二级第一存储器发出故障转移命令以响应所述一级站处的故障。在执行所述故障转移命令之后生成从所述二级第一存储器到所述二级第二存储器的虚拟拷贝。从所述二级第一存储器向所述一级第一存储器发出故障回复命令以响应所述一级站的恢复,以导致从所述二级第一存储器复制所述数据单元,以便重写在到所述二级第一存储器的故障转移之前发生并且未被复制到所述二级站的对所一级第一存储器中的数据单元的更新。

[0020] 在其他实施例中,在所述一级站处存在一级第二存储器。执行从所述一级站到所述二级站的故障转移包括:从所述一级第一存储器故障转移到所述二级第一存储器。复制对所二级站做出的更新以响应所述恢复包括:将系统运行期间在从一级站处的一级存储器故障转移到二级站之后对所述二级第一存储器做出的更新复制到所述一级第一存储器。生成所述一级第一存储器到所述一级站处的一级第二存储器的虚拟拷贝,以响应将所述更新复制到所述一级存储器。

[0021] 在其他实施例中,将生成所述虚拟拷贝期间对所述二级存储器做出的更新复制到所述一级第二存储器以响应到所述一级第二存储器的虚拟拷贝。

[0022] 在其他实施例中,从所述二级第一存储器向所述一级第一存储器发出故障回复命令以响应所述一级站的恢复,以导致从所述二级第一存储器复制所述数据单元,以便重写在到所述二级第一存储器的故障转移之前发生并且未被复制到所述二级第一存储器的对所一级第一存储器中的数据单元的更新。发出强制故障回复命令以响应完成从所述一级第一存储器到所述一级第二存储器的虚拟拷贝,以便在生成所述虚拟拷贝期间执行所述更新的复制。

附图说明

- [0023] 以下将仅通过实例的方式参考附图描述本发明的实施例,这些附图是:
- [0024] 图 1 示出了数据镜像计算环境的实施例;
- [0025] 图 2 示出了一级和二级站处的存储器的布置的实施例;
- [0026] 图 3 示出了建立一级和二级站之间的复制关系的操作的实施例;
- [0027] 图 4 示出了根据图 2 的存储器布置处理一级站处的故障的操作的实施例;
- [0028] 图 5 示出了根据图 2 的存储器布置处理一级站处的恢复的操作的实施例;
- [0029] 图 6 示出了一级和二级站处的存储器的布置的实施例;
- [0030] 图 7 示出了根据图 6 的存储器布置处理一级站处的故障的操作的实施例;
- [0031] 图 8 示出了根据图 6 的存储器布置处理一级站处的恢复的操作的实施例;
- [0032] 图 9 示出了一级和二级站处的存储器的布置的实施例;
- [0033] 图 10 示出了根据图 9 的存储器布置处理一级站处的故障的操作的实施例;
- [0034] 图 11 示出了根据图 9 的存储器布置处理一级站处的恢复的操作的实施例。

具体实施方式

[0035] 图 1 示出了网络计算环境的实施例。一级站 2a 包括连接到一级存储器 6a 的一级服务器 4a 并且二级站 2b 包括连接到二级存储器 6b 的二级服务器 4b。在正常运行期间,一级服务器 4a 通过网络 12 从主机系统 10 中的应用 8 接收更新和输入/输出(I/O)请求(可能存在多个向一级服务器 4a 提供更新的主机系统 10 和应用 8)。同时,一级服务器 4a 通过网络 12 将一级存储器 6a 处的数据镜像到二级存储器 6b 中存储该数据的二级服务器 4b。一级 4a 和二级 4b 服务器分别包括复制管理器程序 14a 和 14b,后者管理镜像复制操作以允许在二级站 2b 处镜像一级站 2a 处的数据。复制管理器程序 14a 和 14b 建立和维护各站之间的复制关系,如国际商业机器公司的存储产品中提供的 Metro Mirror 或 Global Mirror 复制关系,或其他供应商复制关系技术。在一个实施例中,存储器 6a、6b 维护组织成逻辑单元(例如,卷、逻辑单元号(LUN),它们包括诸如轨道或块的数据单元)的数据。此外,每个服务器 4a、4b 包括高速缓存 18a、18b,后者包括一个或多个易失性和/或非易失性存储器件。高速缓存 18a、18b 缓存从主机 10 接收的更新,直到数据被离台到存储器 6a、6b,并且缓存从主机 10 接收的请求。

[0036] 服务器 4a、4b 可以包括企业存储服务器,如 IBM DS8000™ 服务器或本领域公知的其他适合的存储控制器。(DS8000 是国际商业机器公司的商标。)复制管理器 14a 和 14b 包括设计为协调一级 2a 与二级 2b 站之间的数据的镜像、同步以及重新同步的应用。在所述实施例中,复制管理器 14a、14b 还被描述为执行通用存储管理操作,如从高速缓存 18a、18b 登台和离台数据。因此,在所述实施例中,复制管理器 14a、14b 可以包括用于执行与镜像明确无关的存储操作的代码。存储器 6a、6b 可以包括存储设备的阵列,如简单磁盘捆绑(JBOD)、直接存取存储设备(DASD)、独立磁盘冗余阵列(RAID)、虚拟化设备、磁带存储、闪存等。网络 12 可以包括存储区域网络(SAN)、局域网(LAN)、内联网、互联网、广域网(WAN)、对等网络、无线网络、仲裁回路网络等。网络 12 可以包括一个或多个交换机以提供不同网络 12 元件之间的一个或多个通信路径。

[0037] 图 1 示出了一级 2a 和二级 2b 站均包括连接到存储器 6a、6b 的一个服务器 4a、4b。在附加实施例中，站 2a、2b 可以包括连接到多个存储系统的多个服务器，从而提供服务器场。

[0038] 图 2 示出了包括一级存储器 50 和一级影子存储器 52 的一级站 2a 处的存储器 6a 以及包括二级存储器 54 和二级影子存储器 56 的二级站 2b 处的存储器 6b 的布置。在正常运行期间，在一级存储器 50 与二级影子存储器 54 之间建立复制关系 58，如对等或扩展远程复制关系，以便将对一级存储器 50 的更新镜像到二级影子存储器 54。在二级存储器 56 与一级影子存储器 52 之间建立另一复制关系 60。虚拟拷贝操作 64 创建一级影子存储器 52 并且虚拟拷贝操作 62 创建二级存储器 56。复制操作 58 和 60 不能同时处于活动状态。当生产位于一级站 2a 并且数据正在被镜像到二级站 2b 时，复制操作 58 可以是活动的。在一级站 2a 到二级站 2b 的故障转移期间，复制操作 58 暂停。一旦一级站 2a 恢复并且再次运行，就可以启动复制操作 60。

[0039] 虚拟拷贝操作 62 和 64 可以包括逻辑的时间点复制操作。时间点复制涉及将所有数据从源卷物理地复制到目标卷，以便目标卷具有数据的时间点副本。还可以通过逻辑地生成数据的副本并且然后在需要时或在后台复制数据（实际上推迟了物理复制）来生成时间点副本。执行此逻辑复制操作以尽量减小不可访问目标卷和源卷的时间。

[0040] 一种此类逻辑复制操作被称为 **FlashCopy®** (FlashCopy 是国际商业机器公司或“IBM”的注册商标)。**FlashCopy®** 涉及在不同设备上的一级和二级卷之间建立逻辑时间点复制关系。一旦建立了逻辑关系，主机然后就可以立即访问一级和二级卷上的数据，并且可以作为后台操作的一部分复制数据。作为后台操作的一部分将数据从一级卷复制到二级卷。在复制数据时，对一级高速缓存中尚未使用来自一级存储器的数据更新的任何轨道的读取将导致源轨道被登台到二级高速缓存，然后再将访问提供给来自二级高速缓存的轨道。对二级轨道上尚未复制的数据的任何读取将导致数据被从一级设备复制到二级高速缓存，使得二级目标具有来自源的在 **FlashCopy®** 操作的时刻存在的副本。此外，对一级存储器上尚未被复制的轨道的任何写入将导致一级存储器上的轨道被复制到二级存储器。

[0041] 图 3 示出了在复制管理器 14a 和 / 或 14b 中或在一级 4a 和二级 4b 服务器外部的复制管理程序中实现将一级存储器 50 处的数据镜像到二级影子存储器 56 的操作的实施例。响应于启动备份操作（块 100），在一级存储器 50 与二级影子存储器 54 之间建立（块 102）复制关系 58 以镜像数据和对一级存储器 50 的更新。

[0042] 图 4 示出了在复制管理器 14a 和 / 或 14b 中或在一级 4a 和二级 4b 服务器外部的复制管理程序中实现响应检测到一级站 2a 处的故障（块 120）的操作的实施例。作为响应，发出（块 122）故障转移命令以从一级存储器 50 故障转移到二级影子存储器 54 以响应一级站 2a 处的故障。作为响应，生成（块 124）二级影子存储器 54 到二级存储器 56 的虚拟拷贝 62。此外，发出（块 126）强制故障转移命令以从二级存储器 56 故障转移到一级影子存储器 52。块 122 和 126 处的两个故障转移命令导致二级站存储器 54 和 56 维护更改记录位图以指示在故障转移后更新的数据单元（例如，轨道）。在故障转移后使用二级存储器 56 启动（块 128）二级站 2b 上的主机操作，其中此类更改反映在二级存储器 56 的更改记录位图中。这样，在虚拟拷贝二级存储器 56 上练习生产。

[0043] 图 5 示出了在复制管理器 14a 和 / 或 14b 中或在一级 4a 和二级 4b 服务器外部的

复制管理程序中实现处理一级站 2a 处的恢复的操作的实施例。响应于检测到 (块 150) 一级站 2a 处的恢复,发出 (块 152) 从二级影子存储器 54 到一级存储器 50 的故障回复命令,以便重写一级存储器 50 中的在故障转移之前被更新并且未被复制到二级影子存储器 54 的任何数据单元,以使一级存储器 50 和二级影子存储器 54 同步。在同步从二级影子存储器 54 到一级存储器 50 的所有对之后,生成 (块 154) 一级存储器 50 到一级影子存储器 52 的虚拟拷贝。发出 (块 156) 从二级存储器 56 到一级影子存储器 52 的强制故障回复命令,以复制在故障转移后并且在建立复制关系 60 之前由二级存储器 56 记录的在二级存储器 56 处更新的数据单元。终止 (块 158) 一级存储器 50 与二级影子存储器 54 之间的复制关系 58。

[0044] 在备选恢复实施例中,操作可以在一级站 2a 恢复后切换到二级站 2b。可以通过使一级影子存储器 52 与恢复后的二级影子存储器 56 同步来完成此操作。在一级站 2a 返回同步后,一级站 2a 处的操作继续而不受影响,并且二级存储器 56 上的练习会话可以继续而不影响有关一级站的主机操作。此外,从一级存储器 50 到二级影子存储器 54 的镜像继续。

[0045] 通过图 2-5 的所述实施例,客户可以在两个站点上运行生产并且在两个位置提供了保护。此外,通过所述实施例,仅从二级存储器 56 复制作为虚拟拷贝操作的结果产生的更改后的更新,其中二级影子存储器 54 维护故障转移时的数据。虚拟拷贝卷 (即,二级存储器 56) 用于记录更改并且作为故障回复的结果,仅将记录的更改从虚拟拷贝卷复制回一级站。可以通过在建立复制关系 60 之前执行从虚拟拷贝卷 (二级存储器 56) 到一级影子存储器 52 的强制故障转移来实现此结果,这导致虚拟拷贝卷 (二级存储器 56) 记录故障转移期间更改的更新,以便在故障回复期间仅复制记录的更改。执行从二级影子存储器 54 到一级存储器 50 的故障回复,以便删除在一级存储器 50 处接收的在故障转移前未被复制到二级影子存储器 54 的任何更新,以使一级存储器 50 与二级影子存储器 54 同步。将二级虚拟拷贝卷 56 处的记录的更改应用于一级虚拟拷贝 (一级影子存储器 52) (包括同步到一级存储器 50 的数据) 确保了到一级影子存储器 52 的故障回复将二级站 2b 与一级站 2a 同步。此外,图 2-5 的实施例提供了对称配置,从而就图 3-5 的操作而言,二级站 2b 可以用作一级站并且一级站 2a 可以用作二级站,即,主机操作可以在站 2b 上运行、从站 2b 镜像到 2a、从站 2b 故障转移到站 2a,然后从站 2b 恢复和重新同步到站 2a。

[0046] 图 6 示出了包括一级存储器 180 和一级影子存储器 182 的一级站 2a 处的存储器 6a 以及包括二级存储器 184 的二级站 2b 处的存储器 6b 的实施例。在正常运行期间,在一级存储器 180 与二级存储器 184 之间建立复制关系 186,如对称或扩展远程复制关系,以便将一级存储器 180 的更新镜像到二级存储器 184。在一级站 2a 处的恢复之后,在二级存储器 56 与一级影子存储器 58 之间建立其他复制关系 188,以便将数据从二级存储器 184 复制到一级影子存储器 182。虚拟拷贝操作 190 创建一级影子存储器 182。虚拟拷贝操作 190 可以包括如根据图 2 所述的逻辑时间点复制操作。

[0047] 图 7 示出了在复制管理器 14a 和 / 或 14b 中或在一级 4a 和二级 4b 服务器外部的复制管理程序中实现响应检测到一级站 2a 处的故障 (块 200) 的操作的实施例。作为响应,发出 (块 202) 从一级存储器 180 到二级存储器 184 的故障转移命令。作为故障转移的结果,在二级存储器 184 处执行读取和写入,二级存储器 184 记录故障转移期间与复制关系 186 一致的任何更改。

[0048] 图 8 示出了在复制管理器 14a 和 / 或 14b 中或在一级 4a 和二级 4b 服务器外部的复制管理程序中实现处理一级站 2a 处的恢复的操作的实施例。响应于检测到 (块 230) 一级站 2a 处的恢复, 发出 (块 232) 从二级存储器 184 到一级存储器 180 的故障回复, 以复制故障转移之后的任何更新以及重写一级存储器 180 中在故障转移之前更新并且未被复制到二级存储器 184 的任何数据单元。在二级存储器 184 被同步到一级存储器 180 之后, 终止 (块 234) 复制关系 186。生成 (块 236) 同步后的一级存储器 180 到一级影子存储器 182 的虚拟拷贝 190。然后发出 (块 238) 从二级存储器 184 到一级影子存储器 182 的强制故障回复, 以便复制在暂停之后并且在生成虚拟拷贝 182 期间在二级存储器 184 处更新的数据单元。在更新一级影子存储器 182 之后, 终止 (块 240) 二级存储器 184 与一级存储器 180 之间的复制关系 188。

[0049] 通过图 6-8 的操作, 到一级存储器 180 的故障回复将作为故障转移和复制关系 184 的结果记录的更改的更新从二级存储器 184 复制到一级存储器 180。此外, 在执行虚拟拷贝 190 操作时, 二级存储器 184 继续记录对二级存储器 184 的更改和活动。这些更新然后被复制到一级影子存储器 182。在终止复制关系 188 之后, 可以在一级影子存储器 182 上运行生产。

[0050] 图 9 示出了包括一级存储器 250 的一级站 2a 处的存储器 6a 以及包括二级影子存储器 252 和二级存储器 254 的二级站 2b 处的存储器 6b 的实施例。在正常运行期间, 在一级存储器 250 与二级影子存储器 252 之间建立复制关系 256, 如对等或扩展远程复制关系, 以便将对一级存储器 250 的更新镜像到二级存储器 252。建立从二级存储器 254 到一级存储器 252 的其他复制关系 258。虚拟拷贝操作 260 创建二级存储器 254。虚拟拷贝操作 260 可以包括如根据图 2 所述的逻辑时间点复制操作。

[0051] 图 10 示出了在复制管理器 14a 和 / 或 14b 中或在一级 4a 和二级 4b 服务器外部的复制管理程序中实现响应检测到一级站 2a 处的故障的操作的实施例。响应于故障 (块 280), 发出 (块 282) 故障转移命令以从一级存储器 250 故障转移到二级影子存储器 252。这导致二级影子存储器 252 维护任何更改的记录位图。生成 (块 284) 二级影子存储器 151 到二级存储器 254 的虚拟拷贝, 并且发出 (块 286) 从二级存储器 254 到一级存储器的强制故障转移命令, 以便记录故障转移之后对二级存储器 254 的更新。强制故障转移建立复制关系 258, 后者使得二级存储器 254 能够记录故障转移之后的更改。在故障转移之后, 就二级存储器 254 而言, 生产在二级站 2b 处继续。

[0052] 图 11 示出了在复制管理器 14a 和 / 或 14b 中或在一级 4a 和二级 4b 服务器外部的复制管理程序中实现处理一级站 2a 处的恢复的操作的实施例。响应于检测到 (块 300) 一级站 2a 处的恢复, 发出 (块 302) 从二级影子存储器 252 到一级存储器 254 的故障回复, 以便重写一级存储器 250 中在故障转移之前更新并且未被复制到二级影子存储器 252 的任何数据单元。在一级存储器 250 被同步到二级影子存储器 252 之后, 终止 (块 304) 一级存储器 250 与二级影子存储器 252 之间的复制关系 256。发出 (块 306) 从二级存储器 254 到一级存储器 250 的强制故障回复, 以便复制在故障转移之后在二级存储器 254 处更新并且被记录为强制故障转移所建立的复制关系 258 的一部分的数据单元。

[0053] 通过图 9-11 的操作, 到一级存储器 250 的故障回复将作为故障转移和复制关系 258 的结果记录的更改的更新从二级存储器 254 复制到一级存储器 250。在一级存储器 250

与虚拟拷贝二级存储器 254 同步后,二级存储器 254 可以成为图 6 中的一级存储器 180,并且二级影子存储器 252 用作图 6 中的一级影子存储器 182,其中一级存储器 250 将用作图 6 中的二级存储器 184。因此,在同步之后,通过使图 9 中的二级站 2b 用作图 6 中的一级站 2a,生产可以在二级存储器 254 上继续。

[0054] 其他实施例详细信息

[0055] 使用生产软件、固件、硬件或它们的任意组合的标准编程和 / 或工程技术,上述操作可以被实现为方法、装置或制品。所述操作可以被实现为在“计算机可读介质”中维护的代码,其中处理器可以从所述计算机可读介质读取并执行所述代码。计算机可读介质可以包括诸如磁存储介质(例如,硬盘驱动器、软盘、磁带等)、光存储装置(CD-ROM、DVD、光盘等)、易失性和非易失性存储设备(例如,EEPROM、ROM、PROM、RAM、DRAM、SRAM、闪存、固件、可编程逻辑等)之类的介质。实现所述操作的代码还可以在硬件逻辑(例如,集成电路芯片、可编程门阵列(PGA)、专用集成电路(ASIC)等)中实现。此外,实现所述操作的代码可以在“传输信号”中实现,其中传输信号可以通过空间或通过传输介质(如光纤、铜线等)传播。其中编码所述代码或逻辑的传输信号还可以包括无线信号、卫星传输、无线电波、红外线信号、蓝牙等。其中编码所述代码或逻辑的传输信号能够通过发射站发射并且通过接收站接收,其中在传输信号中编码的代码或逻辑可以被解码并存储在接收和发射站或设备处的硬件或计算机可读介质中。“制品”包括其中可以实现代码的计算机可读介质、硬件逻辑、和 / 或传输信号。其中编码实现所述实施例操作的代码的设备可以包括计算机可读介质或硬件逻辑。当然,本领域的技术人员将认识到,可以对此配置做出许多修改而不偏离本发明的范围,并且所述制品可以包括本领域公知的适合的信息承载介质。

[0056] 术语“一个实施例”、“实施例”、“多个实施例”、“所述实施例”、“所述多个实施例”、“一个或多个实施例”、“某些实施例”和“某一实施例”指“本发明(多个)的一个或多个(但不是所有)实施例”,除非另外明确指出。

[0057] 术语“包含”、“包括”、“具有”及其变型指“包括但不限于”,除非另外明确指出。

[0058] 列举的项目的列表并非暗示任何或所有的项目互相排斥,除非另外明确指出。

[0059] 术语“一”、“一个”和“所述”指“一个或多个”,除非另外明确指出。

[0060] 相互通信的设备和组件不必持续地相互通信,除非另外明确指出。此外,相互通信的设备可以直接通信或通过一个或多个媒介间接地通信。

[0061] 具有若干相互通信的组件的实施例的描述并非暗示所有此类组件都是必需的。相反,描述了各种可选的组件以例示本发明的多种可能的实施例。

[0062] 此外,尽管可以按照连续的顺序来描述处理步骤、方法步骤、算法或类似步骤,但是此类处理、方法和算法可以被配置为以交替顺序工作。换言之,所描述的步骤的任何序列或顺序并不一定指示要求按此顺序执行步骤。实际可以按任何顺序执行在此描述的处理的步骤。此外,可以同时地执行某些步骤。

[0063] 当在此描述单个设备或物品时,将显而易见的是,可以使用多个设备 / 物品(无论它们是否协作)来代替单个设备 / 物品。同样,当在此描述了多个设备或物品(无论它们是否协作)的情况下,将显而易见的是,可以使用单个设备或物品来代替所述多个设备或物品,或者可以使用不同数量的设备 / 物品代替所示数量的设备或程序。设备的功能和 / 或特性可以替代地由一个或多个其他未明确描述为具有此类功能 / 特性的设备来体现。因

此,本发明的其他实施例不必包括设备本身。

[0064] 图 3-5、7-9、10 和 11 的所示操作示出特定事件按照特定顺序发生。在备选实施例中,可以以不同的顺序执行特定操作、修改或删除特定操作。此外,可以将步骤添加到上述逻辑而仍然与所述实施例一致。此外,在此说明的操作可以顺序地发生或可以并行地处理特定操作。进而,可以由单个处理单元或分布式处理单元执行操作。

[0065] 出于示例和说明目的提供了本发明的各个实施例的上述说明。其并非旨在是穷举的或将本发明限于所公开的精确形式。根据上述教导,许多修改和变化都是可能的。其旨在本发明的范围并非由此详细说明来限制,而是由此后所附的权利要求来限制。以上说明、实例和数据提供了对本发明的组成部分的制造和使用的完整说明。由于可以在不偏离本发明的精神和范围的情况下做出本发明的许多实施例,所以本发明存在于此后所附的权利要求之内。

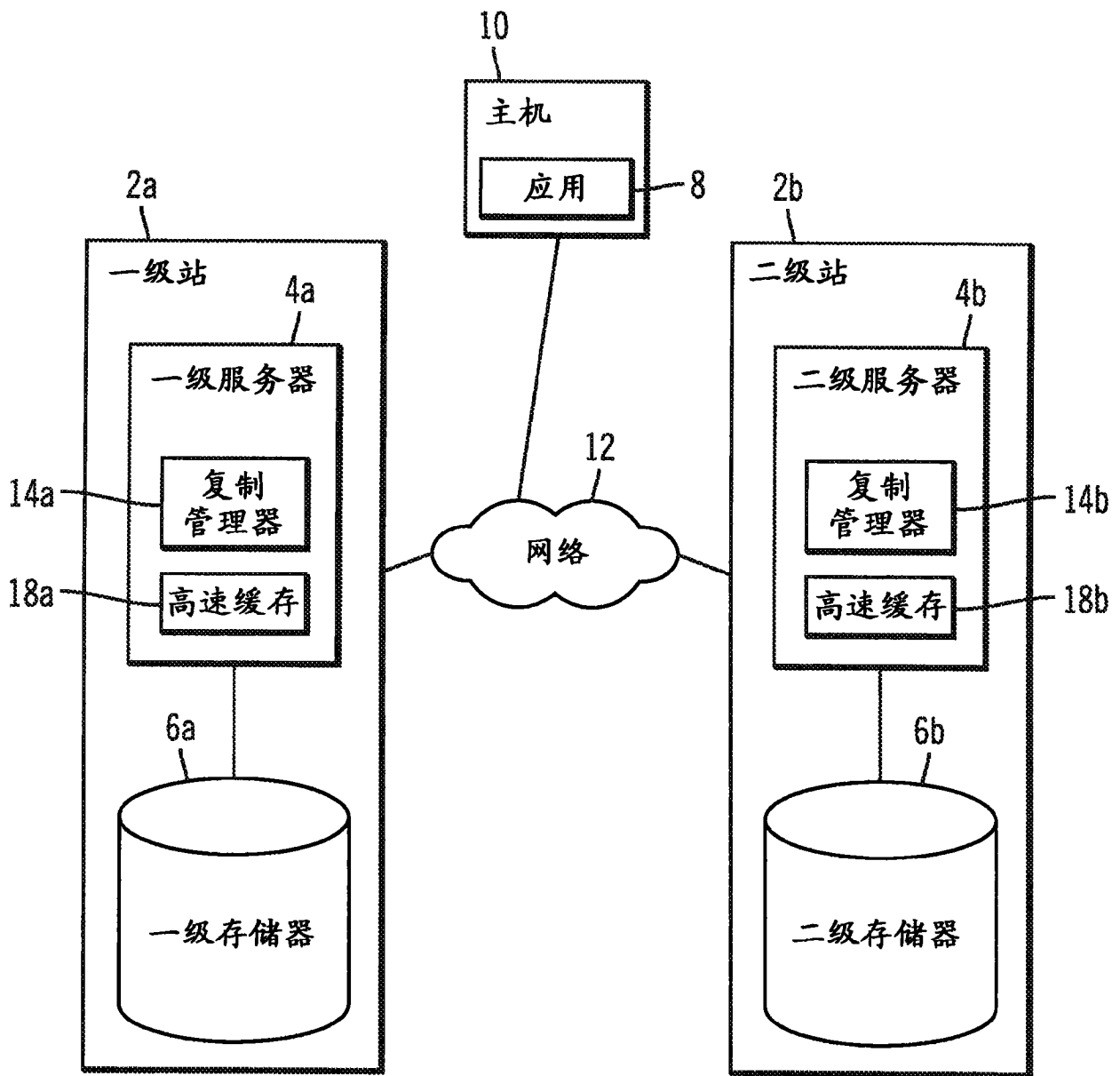


图 1

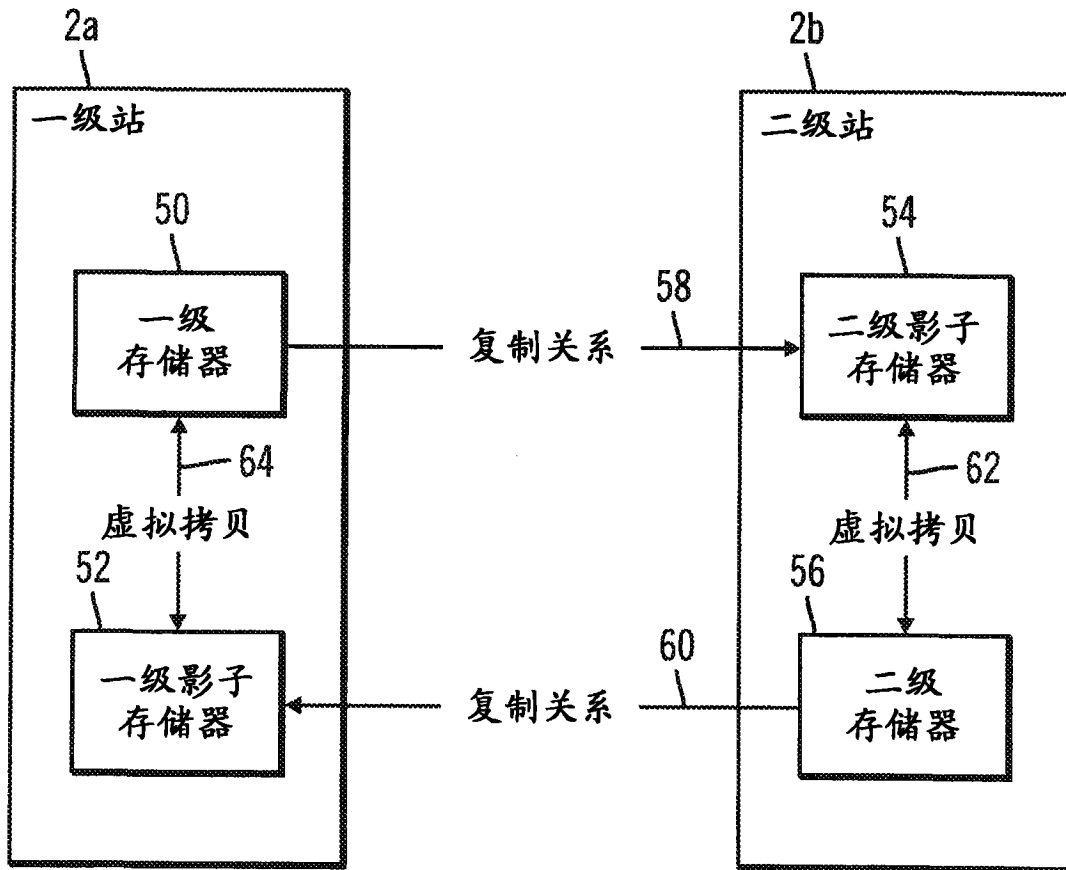


图 2

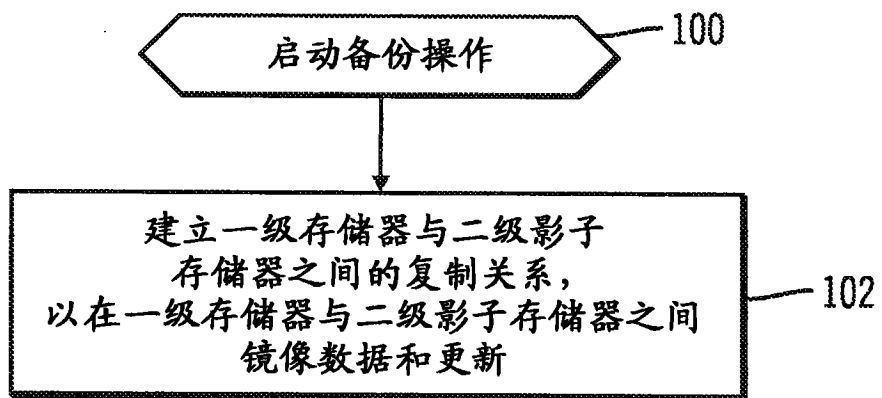


图 3

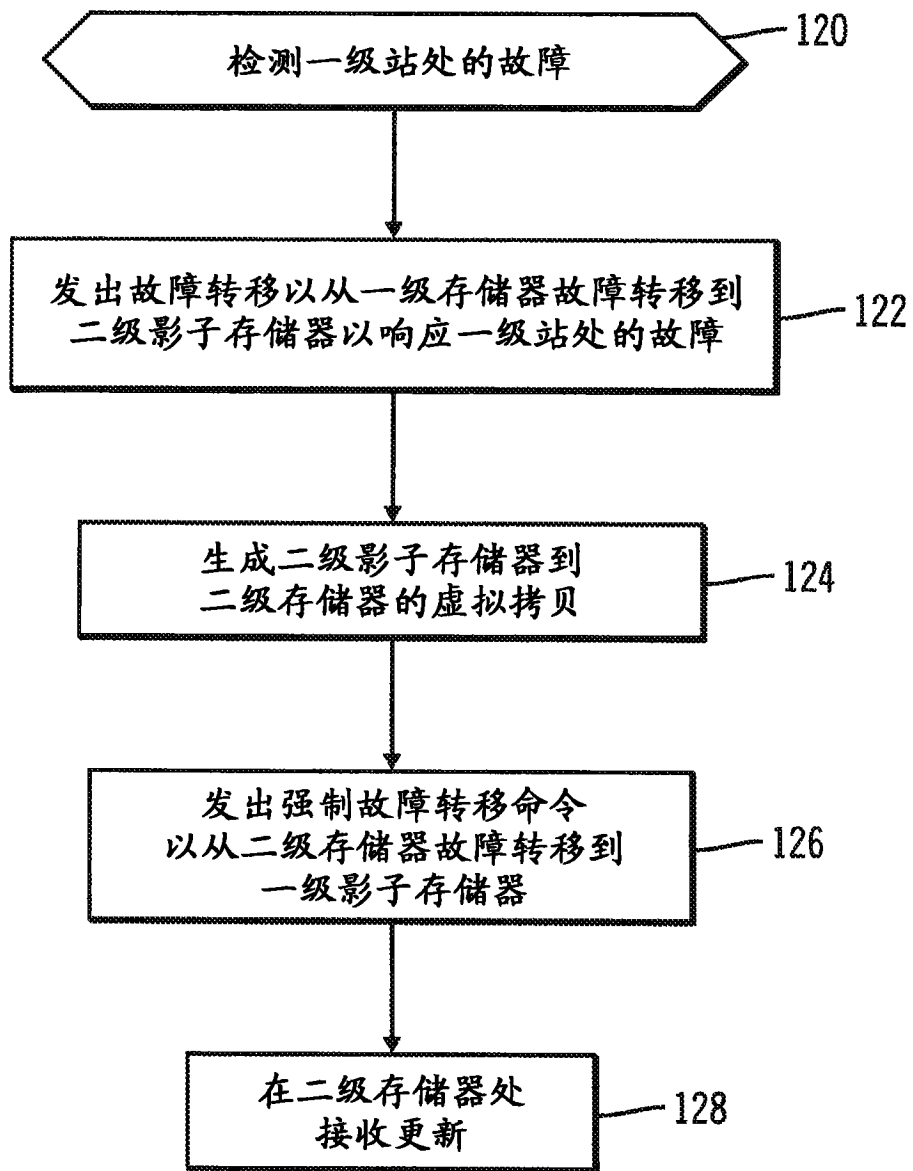


图 4

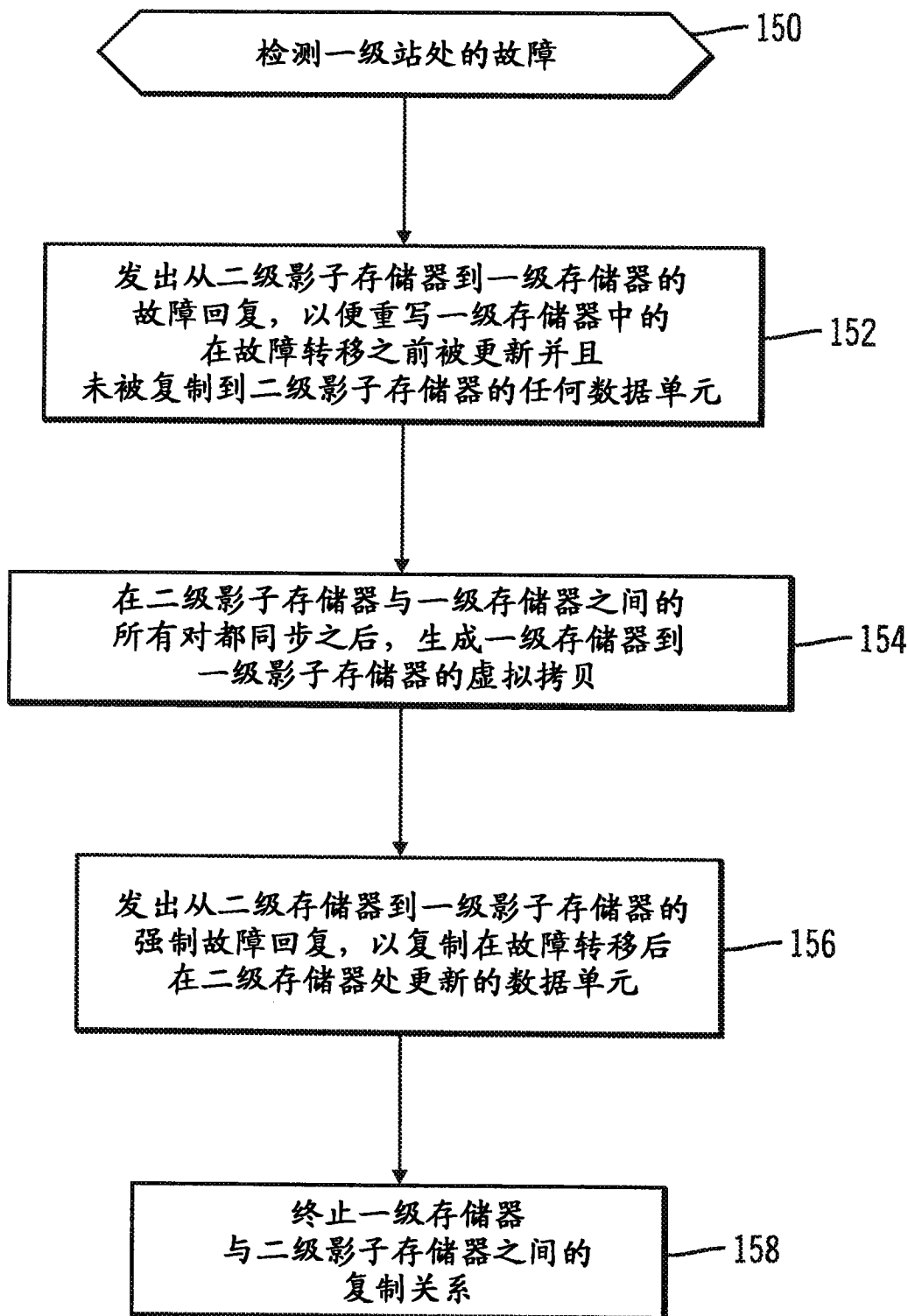


图 5

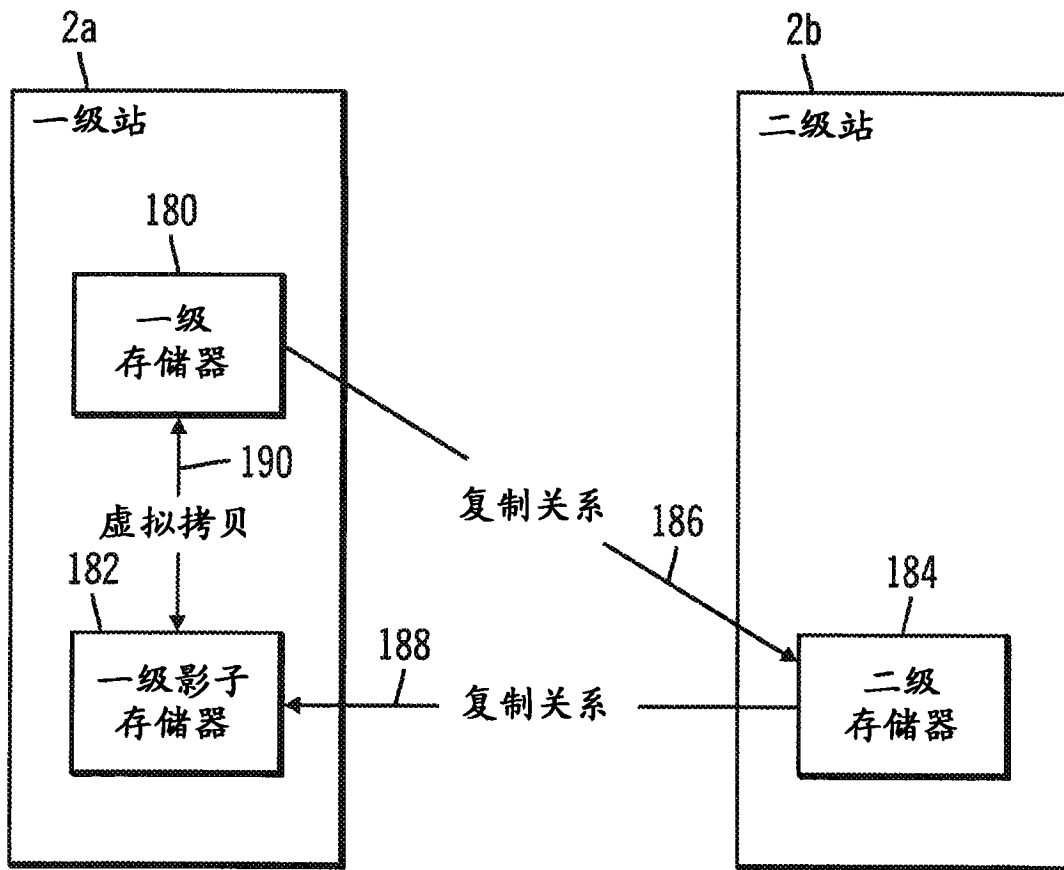


图 6

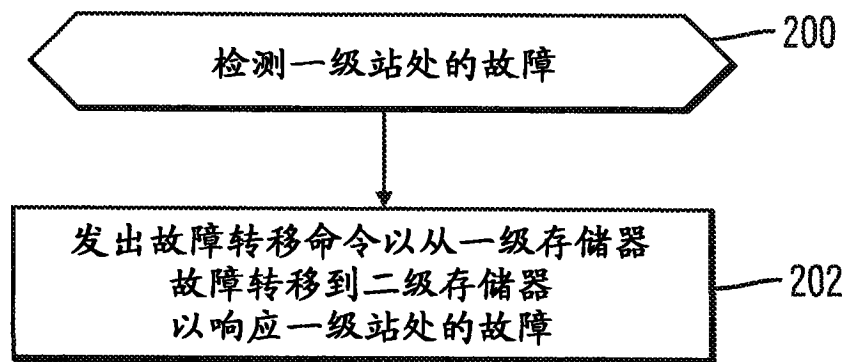


图 7

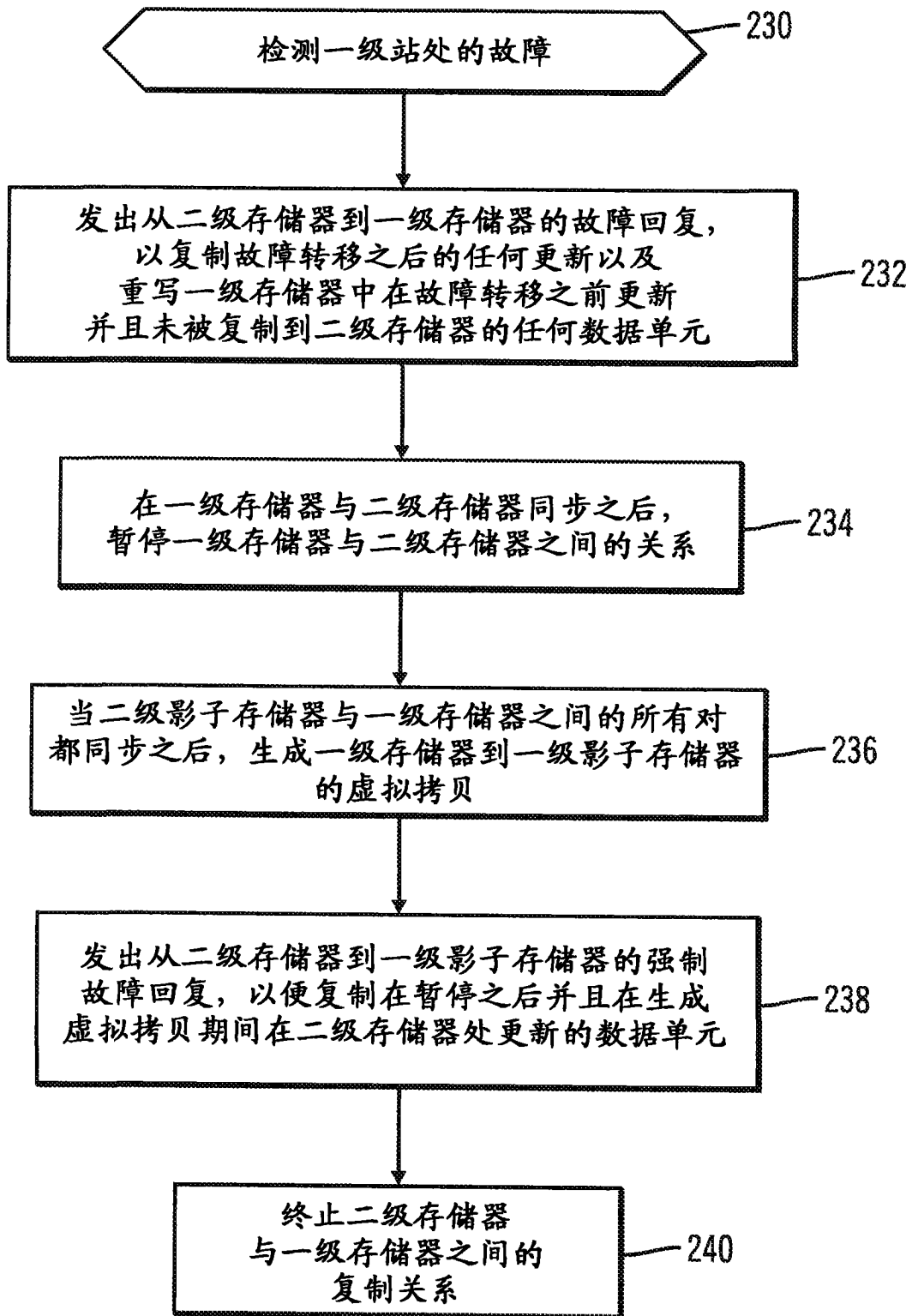


图 8

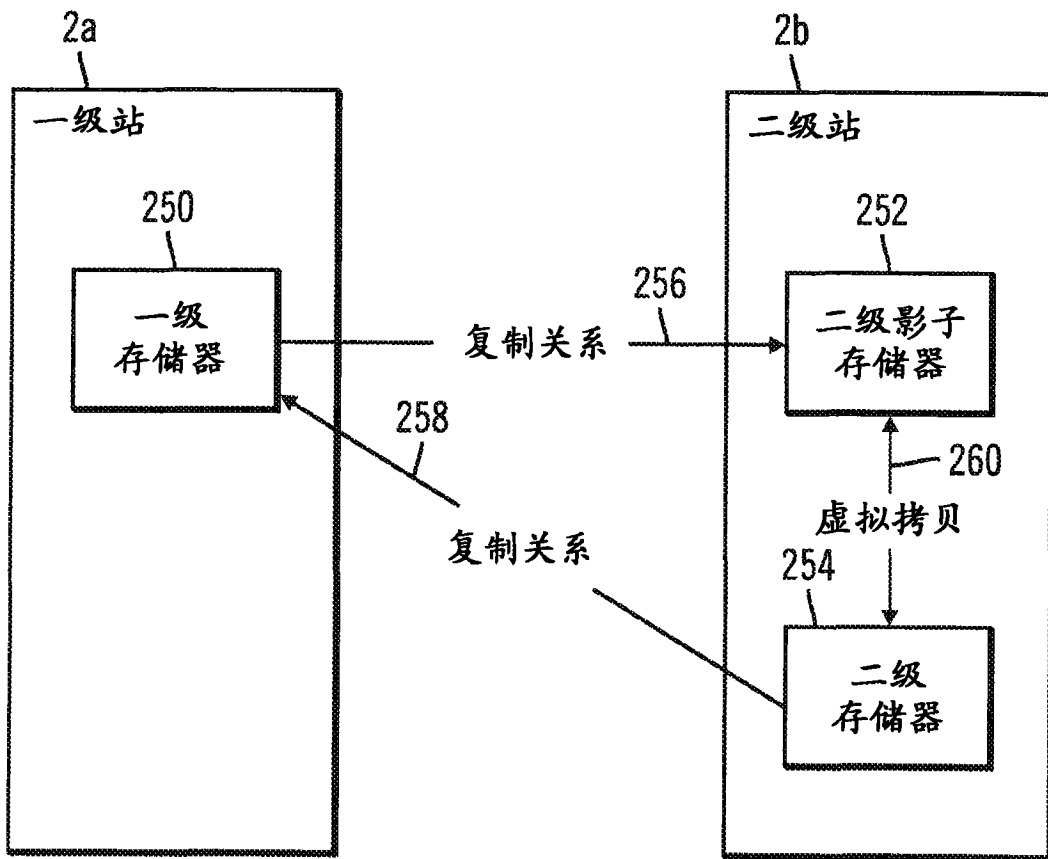


图 9

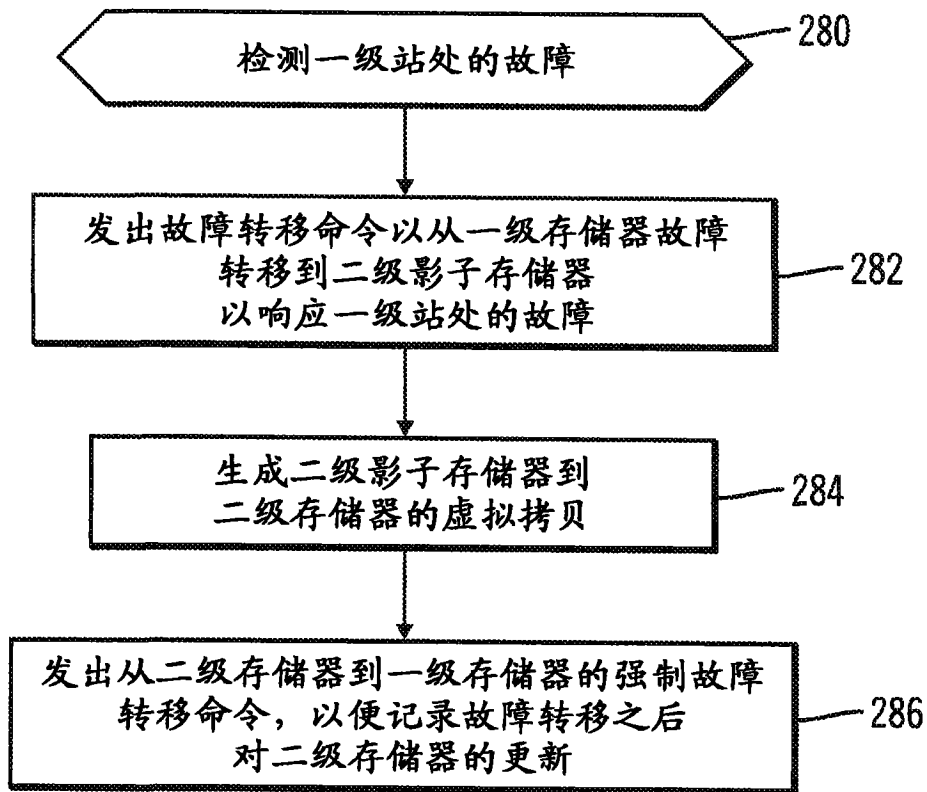


图 10

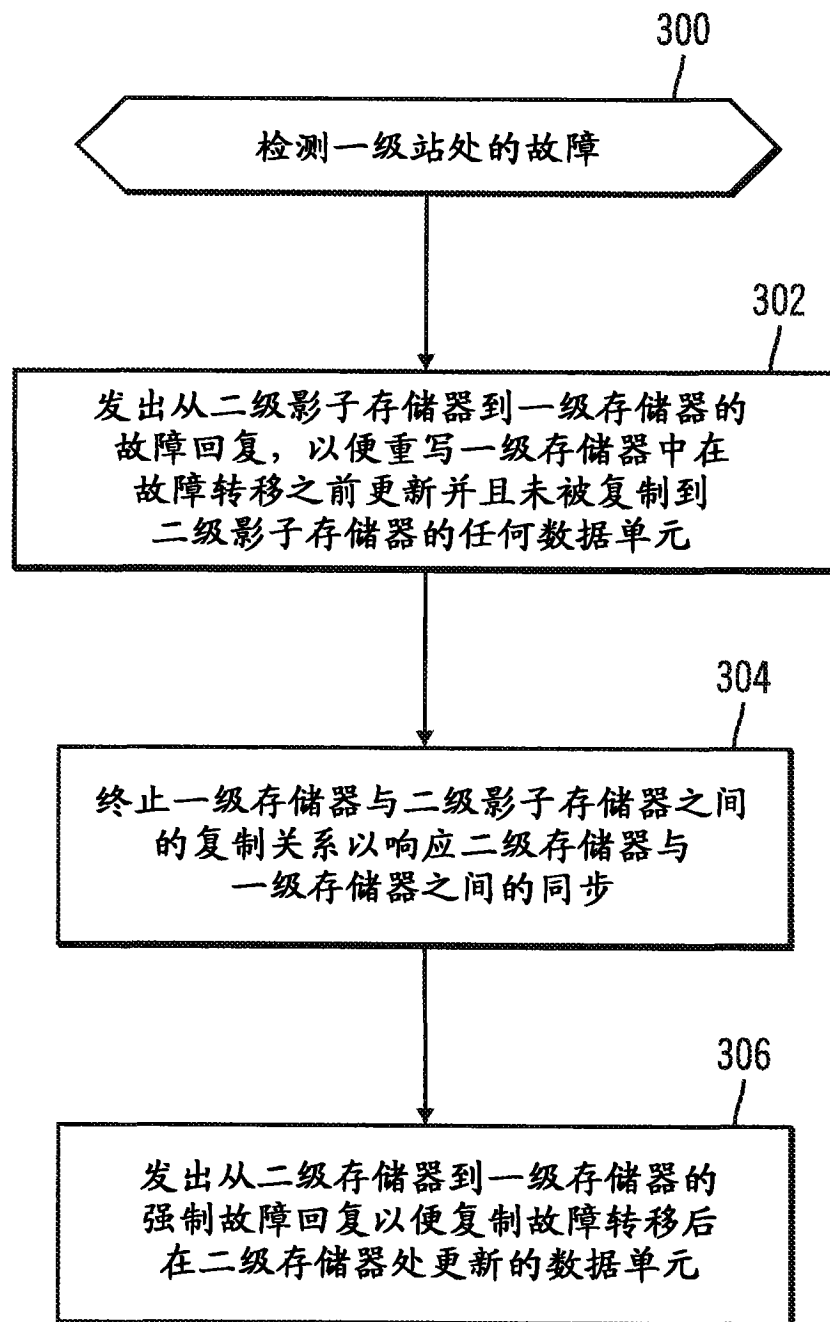


图 11