



(12)发明专利

(10)授权公告号 CN 107533499 B

(45)授权公告日 2020.09.08

(21)申请号 201680025475.3

(22)申请日 2016.05.03

(65)同一申请的已公布的文献号
申请公布号 CN 107533499 A

(43)申请公布日 2018.01.02

(30)优先权数据
14/704,893 2015.05.05 US

(85)PCT国际申请进入国家阶段日
2017.11.02

(86)PCT国际申请的申请数据
PCT/IB2016/052502 2016.05.03

(87)PCT国际申请的公布数据
W02016/178138 EN 2016.11.10

(73)专利权人 国际商业机器公司
地址 美国纽约

(72)发明人 J·G·汤普森 D·彼得森
G·A·斯皮尔 A·G·麦克卢尔
D·罗曼 M·弗兰肯贝格
M·布兰德纳

(74)专利代理机构 北京市中咨律师事务所
11247
代理人 于静 张亚非

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

(56)对比文件
US 2005/0071708 A1,2005.03.31
US 2006/0015764 A1,2006.01.19
US 2007/0288710 A1,2007.12.13
US 2014/0258659 A1,2014.09.11

审查员 赵小娟

权利要求书5页 说明书11页 附图5页

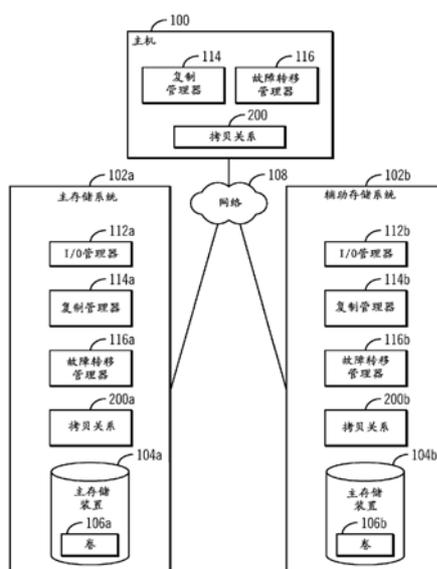
(54)发明名称

用于在存储系统之间执行故障转移的方法和系统

(57)摘要

提供一种用于在第一存储系统(102a)与第二存储系统(102b)之间执行故障转移的计算机程序产品、系统和方法。在所述第一存储系统(102a)与所述第二存储系统(102b)之间同步数据。响应于当同步所述数据时所述第一存储系统处(102a)的故障转移事件,执行从所述第一存储系统(102a)到所述第二存储系统(102b)的故障转移。响应于所述故障转移事件,确定所述第一存储系统(102a)的第一存储单元不能够操作,并且确定所述第一存储系统(102a)的第二存储单元能够操作。响应于确定所述第二存储单元能够操作,通过将所述第二存储系统(102b)的第二存储单元镜像到所述第一存储系统(102a)的所述第二存储单元,启动用于将更新拷贝到所述第一存储系统(102a)的所述第二存储单元的重新同步。

CN 107533499 B



1. 一种用于在第一存储系统与第二存储系统之间执行故障转移的计算机可读存储介质,所述计算机可读存储介质具有随其包含的程序指令,所述程序指令能由处理器执行以便导致操作,所述操作包括:

在所述第一存储系统与所述第二存储系统之间同步数据;

响应于当同步所述数据时所述第一存储系统处的故障转移事件,执行从所述第一存储系统到所述第二存储系统的故障转移,其中作为所述故障转移的一部分,到所述第一存储系统的输入/输出(I/O)请求被重定向到所述第二存储系统;

响应于所述故障转移事件,确定所述第一存储系统的第一存储单元不能够操作;

响应于所述故障转移事件,确定所述第一存储系统的第二存储单元能够操作;

响应于确定所述第一存储系统的所述第二存储单元能够操作,当I/O请求正在被重定向到所述第二存储系统时,通过将所述第二存储系统的第二存储单元镜像到所述第一存储系统的所述第二存储单元,启动用于将更新拷贝到所述第一存储系统的所述第二存储单元的第一重新同步;

在执行所述故障转移后的预定时段之后,向所述第一存储系统发出运行状况查询以便判定所述第一存储系统是否能够完全操作,其中当所述第一存储系统能够完全操作时,所述第一存储系统的所述第一和第二存储单元都能够操作;以及

响应于确定对所述运行状况查询的响应指示所述第一存储系统能够完全操作,启动用于将对所述第二存储系统的所述第一存储单元的更新重新同步到所述第一存储系统的所述第一存储单元的第二重新同步。

2. 根据权利要求1所述的计算机可读存储介质,其中以同步拷贝模式执行所述数据在所述第一存储系统与所述第二存储系统之间的同步,并且其中以异步拷贝模式执行当I/O请求正在被重定向到所述第二存储系统时在所述重新同步期间的所述更新的拷贝。

3. 根据权利要求1所述的计算机可读存储介质,其中确定所述第一存储单元不能够操作包括:

启动所述第一存储系统的所述第一存储单元的时间点拷贝;以及

确定所述第一存储单元的所述时间点拷贝失败,其中响应于确定所述第一存储单元的所述时间点拷贝失败,所述第一存储单元被确定为不能够操作;

其中确定所述第二存储单元能够操作包括:

启动所述第一存储系统的所述第二存储单元的时间点拷贝;以及

确定所述第二存储单元的所述时间点拷贝成功,其中响应于确定所述第二存储单元的所述时间点拷贝成功,所述第二存储单元被确定为能够操作。

4. 根据权利要求3所述的计算机可读存储介质,其中所述操作进一步包括:

响应于所述故障转移事件,针对所述第一存储系统启动软隔离状态以便防止对所述第一存储系统处的所述存储单元的I/O访问,其中使用具有用于允许所述时间点拷贝操作在所述第一存储系统的所述软隔离状态期间继续的参数的命令来启动所述时间点拷贝。

5. 根据权利要求3所述的计算机可读存储介质,其中所述第一存储单元包括所述第一和第二存储系统的第一卷,并且其中所述第二存储单元包括所述第一和第二存储系统的第二卷。

6. 根据权利要求1所述的计算机可读存储介质,其中所述第一存储单元包括所述第一

和第二存储系统的第一卷,并且其中所述第二存储单元包括所述第一和第二存储系统的第二卷,其中被确定为不能够操作的所述第一卷包括经历数据丢失的磁道的子集,其中响应于指示所述第一存储系统能够完全操作的所述运行状况查询,所述重新同步进一步执行:

将所述第二存储系统的所述第一卷中的磁道的子集拷贝到所述第一存储系统中的所述第一卷中的磁道的对应子集,其中所述第一存储系统的所述第一卷中未经历数据丢失并且不与所述第二存储系统的所述第一卷中当I/O请求正在被重定向到所述第二存储系统时被更新的磁道对应的磁道不进行重新同步。

7. 根据权利要求1所述的计算机可读存储介质,其中以异步拷贝模式执行在对所述运行状况查询的指示所述第一存储系统能够完全操作的响应之前的所述第二存储系统的所述第二存储单元到所述第一存储系统的所述重新同步,其中所述操作进一步包括:

响应于指示所述第一存储系统能够完全操作的查询,将所述第二存储系统的所述第二存储单元到所述第一存储系统的所述重新同步转变为同步拷贝模式。

8. 根据权利要求7所述的计算机可读存储介质,其中所述操作进一步包括:

响应于指示所述第一存储系统能够操作的运行状况查询,启动所述第一存储系统的所述第一存储单元的时间点拷贝,其中响应于确定所述时间点拷贝成功,执行同步的重新同步操作。

9. 根据权利要求1所述的计算机可读存储介质,其中所述操作进一步包括:

响应于指示所述第一存储系统不能完全操作的所述运行状况查询,收集用于所述第一存储系统的修复的诊断信息;以及

响应于基于所收集的诊断信息对所述第一存储系统执行修复,接收从所述第一存储系统的管理员发出的重新同步命令,其中所接收的重新同步命令启动将对所述第二存储系统的所述第一存储单元的更新拷贝到所述第一存储系统的所述第一存储单元。

10. 根据权利要求1所述的计算机可读存储介质,其中当I/O请求正在被重定向到所述第二存储系统时,对是所述第一存储系统的所述第一存储单元的镜像的所述第二存储系统的第一存储单元的更新不被拷贝到所述第一存储系统的所述第一存储单元。

11. 一种用于在第一存储系统与第二存储系统之间执行故障转移的系统,所述系统包括:

至少一个处理器;以及

计算机可读存储介质,其具有程序指令,所述程序指令由所述至少一个处理器执行以便执行操作,所述操作包括:

在所述第一存储系统与所述第二存储系统之间同步数据;

响应于当同步所述数据时所述第一存储系统处的故障转移事件,执行从所述第一存储系统到所述第二存储系统的故障转移,其中作为所述故障转移的一部分,到所述第一存储系统的输入/输出(I/O)请求被重定向到所述第二存储系统;

响应于所述故障转移事件,确定所述第一存储系统的第一存储单元不能够操作;

响应于所述故障转移事件,确定所述第一存储系统的第二存储单元能够操作;

响应于确定所述第一存储系统的所述第二存储单元能够操作,当I/O请求正在被重定向到所述第二存储系统时,通过将所述第二存储系统的第二存储单元镜像到所述第一存储系统的所述第二存储单元,启动用于将更新拷贝到所述第一存储系统的所述第二存储单元

的第一重新同步；

在执行所述故障转移后的预定时段之后，向所述第一存储系统发出运行状况查询以便判定所述第一存储系统是否能够完全操作，其中当所述第一存储系统能够完全操作时，所述第一存储系统的所述第一和第二存储单元都能够操作；以及

响应于确定对所述运行状况查询的响应指示所述第一存储系统能够完全操作，启动用于将对所述第二存储系统的所述第一存储单元的更新重新同步到所述第一存储系统的所述第一存储单元的第二重新同步。

12. 根据权利要求11所述的系统，其中以同步拷贝模式执行所述数据在所述第一存储系统与所述第二存储系统之间的同步，并且其中以异步拷贝模式执行当I/O请求正在被重定向到所述第二存储系统时在所述重新同步期间的所述更新的拷贝。

13. 根据权利要求11所述的系统，其中确定所述第一存储单元不能够操作包括：

启动所述第一存储系统的所述第一存储单元的时间点拷贝；以及

确定所述第一存储单元的所述时间点拷贝失败，其中响应于确定所述第一存储单元的所述时间点拷贝失败，所述第一存储单元被确定为不能够操作；

其中确定所述第二存储单元能够操作包括：

启动所述第一存储系统的所述第二存储单元的时间点拷贝；以及

确定所述第二存储单元的所述时间点拷贝成功，其中响应于确定所述第二存储单元的所述时间点拷贝成功，所述第二存储单元被确定为能够操作。

14. 根据权利要求13所述的系统，其中所述操作进一步包括：

响应于所述故障转移事件，针对所述第一存储系统启动软隔离状态以便防止对所述第一存储系统处的所述存储单元的I/O访问，其中使用具有用于允许所述时间点拷贝操作在所述第一存储系统的所述软隔离状态期间继续的参数的命令来启动所述时间点拷贝。

15. 根据权利要求11所述的系统，其中所述第一存储单元包括所述第一和第二存储系统的第一卷，并且其中所述第二存储单元包括所述第一和第二存储系统的第二卷，其中被确定为不能够操作的所述第一卷包括经历数据丢失的磁道的子集，其中响应于指示所述第一存储系统能够完全操作的所述运行状况查询，所述重新同步进一步执行：

将所述第二存储系统的所述第一卷中的磁道的子集拷贝到所述第一存储系统中的所述第一卷中的磁道的对应子集，其中所述第一存储系统的所述第一卷中未经历数据丢失并且不与所述第二存储系统的所述第一卷中当I/O请求正在被重定向到所述第二存储系统时被更新的磁道对应的磁道不进行重新同步。

16. 根据权利要求11所述的系统，其中以异步拷贝模式执行在对所述运行状况查询的响应指示所述第一存储系统能够完全操作的响应之前的所述第二存储系统的所述第二存储单元到所述第一存储系统的所述重新同步，其中所述操作进一步包括：

响应于指示所述第一存储系统能够完全操作的查询，将所述第二存储系统的所述第二存储单元到所述第一存储系统的所述重新同步转变为同步拷贝模式。

17. 一种用于在第一存储系统与第二存储系统之间执行故障转移的方法，所述方法包括：

在所述第一存储系统与第二存储系统之间同步数据；

响应于当同步所述数据时所述第一存储系统处的故障转移事件，执行从所述第一存储

系统到所述第二存储系统的故障转移,其中作为所述故障转移的一部分,到所述第一存储系统的输入/输出(I/O)请求被重定向到所述第二存储系统;

响应于所述故障转移事件,确定所述第一存储系统的第一存储单元不能够操作;

响应于所述故障转移事件,确定所述第一存储系统的第二存储单元能够操作;

响应于确定所述第一存储系统的所述第二存储单元能够操作,当I/O请求正在被重定向到所述第二存储系统时,通过将所述第二存储系统的第二存储单元镜像到所述第一存储系统的所述第二存储单元,启动用于将更新拷贝到所述第一存储系统的所述第二存储单元的第一重新同步;

在执行所述故障转移后的预定时段之后,向所述第一存储系统发出运行状况查询以便判定所述第一存储系统是否能够完全操作,其中当所述第一存储系统能够完全操作时,所述第一存储系统的所述第一和第二存储单元都能够操作;以及

响应于确定对所述运行状况查询的响应指示所述第一存储系统能够完全操作,启动用于将对所述第二存储系统的所述第一存储单元的更新重新同步到所述第一存储系统的所述第一存储单元的第二重新同步。

18. 根据权利要求17所述的方法,其中以同步拷贝模式执行所述数据在所述第一存储系统与所述第二存储系统之间的同步,并且其中以异步拷贝模式执行当I/O请求正在被重定向到所述第二存储系统时在所述重新同步期间的所述更新的拷贝。

19. 根据权利要求17所述的方法,其中确定所述第一存储单元不能够操作包括:

启动所述第一存储系统的所述第一存储单元的时间点拷贝;以及

确定所述第一存储单元的所述时间点拷贝失败,其中响应于确定所述第一存储单元的所述时间点拷贝失败,所述第一存储单元被确定为不能够操作;

其中确定所述第二存储单元能够操作包括:

启动所述第一存储系统的所述第二存储单元的时间点拷贝;以及

确定所述第二存储单元的所述时间点拷贝成功,其中响应于确定所述第二存储单元的所述时间点拷贝成功,所述第二存储单元被确定为能够操作。

20. 根据权利要求19所述的方法,进一步包括:

响应于所述故障转移事件,针对所述第一存储系统启动软隔离状态以便防止对所述第一存储系统处的所述存储单元的I/O访问,其中使用具有用于允许所述时间点拷贝操作在所述第一存储系统的所述软隔离状态期间继续的参数的命令来启动所述时间点拷贝。

21. 根据权利要求20所述的方法,其中所述第一存储单元包括所述第一和第二存储系统的第一卷,并且其中所述第二存储单元包括所述第一和第二存储系统的第二卷,其中被确定为不能够操作的所述第一卷包括经历数据丢失的磁道的子集,其中响应于指示所述第一存储系统能够完全操作的所述运行状况查询,所述重新同步进一步执行:

将所述第二存储系统的所述第一卷中的磁道的子集拷贝到所述第一存储系统中的所述第一卷中的磁道的对应子集,其中所述第一存储系统的所述第一卷中未经历数据丢失并且不与所述第二存储系统的所述第一卷中当I/O请求正在被重定向到所述第二存储系统时被更新的磁道对应的磁道不进行重新同步。

22. 根据权利要求20所述的方法,其中以异步拷贝模式执行在对所述运行状况查询的响应指示所述第一存储系统能够完全操作的响应之前的所述第二存储系统的所述第二存储单

元到所述第一存储系统的所述重新同步,所述方法进一步包括:

响应于指示所述第一存储系统能够完全操作的查询,将所述第二存储系统的所述第二存储单元到所述第一存储系统的所述重新同步转变为同步拷贝模式。

用于在存储系统之间执行故障转移的方法和系统

技术领域

[0001] 本发明涉及一种用于在故障转移到镜像第一存储系统的第二存储系统之后重新同步到所述第一存储系统的计算机程序产品、系统和方法。

背景技术

[0002] 在存储环境中,存储控制器可以维护镜像拷贝关系,其中镜像拷贝关系中的主卷包括数据从中被物理拷贝到辅助卷的存储装置或卷。故障转移程序(例如国际商业机器公司(“IBM”)的**HyperSwap®**,其是**z/OS®**操作系统中的一项功能)通过维护镜像拷贝关系针对磁盘故障提供持续可用性,以便向一个或多个目标(或辅助)存储系统提供一个或多个主存储系统上的所有主磁盘卷的同步副本。(HyperSwap和z/OS是IBM在全世界各个国家的注册商标)。当检测到磁盘故障时,操作系统中的代码标识HyperSwap管理的卷而不是使I/O请求失败,HyperSwap在内部控制块中切换(或交换)信息以使得针对镜像拷贝关系的辅助卷驱动I/O请求。故障转移的执行非常迅速,并且仅对主机应用造成很小的影响。不向主机应用通知主磁盘故障,主机应用不知道它们的访问已被交换到数据的辅助副本。因为辅助卷是故障之前的主卷的相同副本,所以I/O请求将成功而不影响发出I/O请求的程序,该程序可以是应用程序或操作系统的一部分。因此,这向程序屏蔽了磁盘故障并且避免了应用和/或系统中断。

[0003] 当发生主磁盘故障时,故障转移功能自动地将主机系统的访问从发生故障的主磁盘控制单元交换到包含数据的辅助副本的辅助控制单元。在发生故障转移之后,两个存储系统对之间的镜像被暂停,这意味着由应用对当前主副本进行的更新未被镜像到辅助副本。当处于这种暂停状态时,其他转移操作是不可能的。这使得客户面临影响剩余的数据的唯一良好副本的任何类型的其他故障。

[0004] 在当前技术中,为了将主卷和辅助卷返回到启用故障转移的状态,管理员或用户从已发生故障的主存储系统收集和分析诊断信息,并且在必要时进行修复。管理员/用户然后可以启动将数据从当前辅助站点重新同步回经历故障转移的主存储装置。管理员/用户可以通过启动主存储设备的时间点(“PiT”,point-in-time)副本来执行重新同步操作,以便提供数据的一致副本。在重新同步完成之前,主存储系统设备将不一致,并且因此不对恢复有用。时间点拷贝防止在重新同步期间发生的辅助存储系统故障。

[0005] 如果任何主存储设备已发生数据丢失(由于故障),则通过从对应于已发生故障的主存储系统设备的辅助存储系统设备执行完全拷贝,重新同步已发生故障的设备。对于没有经历数据丢失的主存储系统设备,在重新同步期间,仅将辅助存储系统处的已更新的磁道拷贝回主存储系统。当重新同步操作完成时,镜像对将再次回到彼此同步。

[0006] 管理员/用户管理的重新同步过程能够需要许多小时或甚至几天。在该过程完成之前,客户面临第二故障。

发明内容

[0007] 提供一种用于在第一存储系统与第二存储系统之间执行故障转移的计算机程序产品、系统和方法。在所述第一存储系统与所述第二存储系统之间同步数据。响应于当同步所述数据时所述第一存储系统处的故障转移事件,执行从所述第一存储系统到所述第二存储系统的故障转移,其中作为所述故障转移的一部分,到所述第一存储系统的输入/输出(I/O)请求被重定向到所述第二存储系统。响应于所述故障转移事件,确定所述第一存储系统的第一存储单元不能够操作,并且响应于所述故障转移事件,确定所述第一存储系统的第二存储单元能够操作。响应于确定所述第二存储单元能够操作,当I/O请求正在被重定向到所述第二存储系统时,通过将所述第二存储系统的第二存储单元镜像到所述第一存储系统的所述第二存储单元,启动用于将更新拷贝到所述第一存储系统的所述第二存储单元的重新同步。

[0008] 在到所述第二存储系统的故障转移期间,通过作为所述重新同步的一部分使辅助复制管理器自动地将所述第二存储单元处的更新异步拷贝到那些能够操作的第一存储单元,所述第一存储单元处的数据被尽可能保持最新,以使得一旦所述第一存储系统被修复并能够完全操作,能够更快地完成所述第一存储系统的重新同步。

[0009] 在进一步的实施例中,以同步拷贝模式执行所述数据在所述第一存储系统与所述第二存储系统之间的同步,并且以异步拷贝模式执行当I/O请求正在被重定向到所述第二存储系统时在所述重新同步期间的所述更新的拷贝。

[0010] 在进一步的实施例中,确定所述第一存储单元不能够操作包括:启动所述第一存储系统的所述第一存储单元的时间点拷贝;以及确定所述第一存储单元的所述时间点拷贝失败。响应于确定所述第一存储单元的所述时间点拷贝失败,所述第一存储单元被确定为不能够操作。此外,确定所述第二存储单元能够操作包括:启动所述第一存储系统的所述第二存储单元的时间点拷贝;以及确定所述第二存储单元的所述时间点拷贝成功,其中响应于确定所述第二存储单元的所述时间点拷贝成功,所述第二存储单元被确定为能够操作。

[0011] 获取能够操作的所述第二存储单元的时间点拷贝维护了所述第一存储系统中的可访问数据的副本。

[0012] 在进一步的实施例中,响应于所述故障转移事件,针对所述第一存储系统建立软隔离状态以便防止对所述第一存储系统处的所述存储单元的I/O访问,其中使用具有用于允许所述时间点拷贝操作在所述第一存储系统的所述软隔离状态期间继续的参数的命令来启动所述时间点拷贝。

[0013] 在进一步的实施例中,所述重新同步包括第一重新同步,并且向所述第一存储系统发出运行状况查询以便判定所述第一存储系统是否能够完全操作。当所述第一存储系统能够完全操作时,所述第一存储系统的所述第一和第二存储单元都能够操作。响应于确定对所述运行状况查询的响应指示所述第一存储系统能够完全操作,将对所述第二存储系统的所述第一存储单元的更新重新同步到所述第一存储系统的所述第一存储单元。

[0014] 在进一步的实施例中,所述第一存储单元包括所述第一和第二存储系统的第一卷,并且所述第二存储单元包括所述第一和第二存储系统的第二卷。被确定为不能够操作的所述第一卷包括经历数据丢失的磁道的子集。响应于指示所述第一存储系统能够完全操作的所述运行状况查询,所述重新同步进一步执行:将所述第二存储系统的所述第一卷中

的磁道的子集拷贝到所述第一存储系统中的所述第一卷中的磁道的对应子集。所述第一存储系统的所述第一卷中未经历数据丢失并且不与所述第二存储系统的所述第一卷中当I/O请求正在被重定向到所述第二存储系统时被更新的磁道对应的磁道不进行重新同步。

[0015] 在所述修复和恢复之后,所恢复的不能够操作的第一卷中仅在所述故障转移期间被更新或经历数据丢失的磁道才需要被同步。此外,通过在能够操作的卷的故障转移期间启动重新同步,在所述第一存储系统被恢复之后的同步的完成快得多,因为大部分数据可能已经被重新同步。

[0016] 在进一步的实施例中,以异步拷贝模式执行在对所述运行状况查询的指示所述第一存储系统能够完全操作的响应之前的所述第二存储系统的所述第二存储单元到所述第一存储系统的所述重新同步。响应于指示所述第一存储系统能够完全操作的查询,将所述第二存储系统的所述第二存储单元到所述第一存储系统的所述重新同步转变为同步拷贝模式。

附图说明

[0017] 现在将仅通过示例的方式参考附图描述本发明的实施例,这些附图是:

[0018] 图1示出存储复制环境的一个实施例;

[0019] 图2示出拷贝关系的一个实施例;

[0020] 图3示出时间点拷贝命令的一个实施例;

[0021] 图4示出用于执行故障转移操作的操作的一个实施例;

[0022] 图5示出用于重新同步回已恢复的主存储系统的操作的一个实施例;

[0023] 图6示出其中可以实现图1的组件的计算环境。

具体实施方式

[0024] 所描述的实施例提供了用于在从主存储系统到辅助存储系统的故障转移的情况下重新同步数据的技术。响应于故障转移,对于能够操作的那些主卷或存储单元,在主存储系统已恢复之前,对相应辅助卷的更新被重新同步到能够操作的主卷。对于不能够操作或经历数据丢失的那些主存储系统卷,对相应辅助卷的更新例如在更改记录位图中被指示,并且然后在主存储系统已被恢复或修复之后被重新同步回主卷。

[0025] 使用所描述的实施例,数据立即开始回到已发生故障的主存储系统的能够操作的卷的自动程序控制的重新同步,以便最小化使主存储系统完全同步并可用于处理辅助存储系统处的第二故障转移所需的时间。以这种方式,在恢复主存储系统之后,因为大部分能操作的主卷已经根据自动重新同步过程被重新同步,因此减少了用于在恢复主存储系统之后重新同步的时间。

[0026] 图1示出数据镜像和故障转移环境的一个实施例,该环境具有连接到主存储系统102a和辅助存储系统102b的主机系统100。主存储系统102a包括具有卷106a的主存储装置104a,卷106a以镜像对或拷贝关系被镜像到辅助存储系统102b中的对应卷106b。主机100以及存储系统102a和102b可以通过网络108通信。可以存在向主106a和辅助106b卷提供输入/输出(I/O)请求的其它主机(未示出)。

[0027] 主102a和辅助102b存储系统包括输入/输出(I/O)管理器112a、112b,它们分别用

于管理针对主106a和辅助106b卷的I/O操作。主机系统100包括复制管理器114,其用于在不同卷106a、106b之间建立镜像拷贝关系200。存储系统102a、102b包括复制管理器114a、114b,它们用于管理主106a与辅助106b卷之间的数据的复制或镜像,并且与主机复制管理器114协调复制。

[0028] 主机100包括故障转移管理器116,其用于响应于故障转移事件(例如主存储系统102a处的存储或组件故障),管理从主存储系统102a到辅助存储系统102b的I/O操作的故障转移,并且用于当辅助存储系统102b管理I/O操作并遇到错误时,管理从辅助存储系统102b到主存储系统102a的故障转移。主102a和辅助102b存储系统分别包括故障转移管理器116a、116b,它们用于与主机故障转移管理器116协调来实现故障转移操作。

[0029] 主102a和辅助102b存储系统维护关于由主机复制管理器114建立的拷贝关系200a、200b的信息。当主存储系统102a正在运行并管理针对卷106a的I/O操作时,拷贝关系200、200a、200b指示数据正在从主卷106a被镜像或同步到对应的辅助卷106b。在一个实施例中,来自主卷106a的数据作为一致性组被镜像,其中以同步拷贝模式将数据拷贝到对应的辅助卷106b,在同步拷贝模式中,直到写入被确认为已存储在辅助存储装置104b处,写入才完成。

[0030] 响应于故障转移事件,故障转移管理器116、116a、116b将故障转移协调到辅助存储系统102b,其中辅助存储系统102b接手管理对辅助卷106b(其具有从主卷106a复制的数据)的I/O请求。在故障转移之后,主机复制管理器114可以在辅助存储系统102b处创建拷贝关系200b,以便将对辅助卷106b的更新重新同步回主卷106a。因此,当生产站点已移动到辅助存储装置104b时,辅助复制管理器114b重新同步对与主存储系统102a处的能够操作的主卷106a相对应的辅助卷106b的更新。

[0031] 在一个实施例中,拷贝关系200由主机复制管理器114创建,并且作为拷贝关系200a、200b的本地副本被提供给主102a和辅助102b存储系统。以这种方式,主机100管理主102a和辅助102b存储系统的复制、故障转移和重新同步操作。在一个备选实施例中,主存储系统102a和辅助存储系统102b自身可以管理复制和故障转移而无需主机100参与。

[0032] 存储系统102a和102b可以包括适合于管理对附接存储设备的访问的企业存储控制器/服务器,例如但不限于国际商业机器公司(“IBM”)的**DS8000®**存储系统或所属技术领域中公知的其它供应商存储服务器。(DS8000是IBM在全世界各个国家的注册商标)。在一个实施例中,复制管理器114、114a、114b包括用于管理跨系统的卷的镜像的程序,例如但不限于IBM镜像程序Geographically Dispersed Parallel Sysplex(**GDPS®**)和**Tivoli®**Storage Productivity Center for Replication(TPC-R),它们定义复制会话和拷贝对200。可以选择不同类型的技术来拷贝数据,例如同步镜像、异步镜像或时间点拷贝、或者这些不同镜像类型中的多种的组合。故障转移管理器116、116a、116b可以包括适合于处理从存储系统102a、102b中的一个到另一个的故障转移的程序,例如但不限于**IBM®**HyperSwap产品,其从已建立的拷贝对来建立故障转移会话。(IBM、GDPS、Tivoli、以及HyperSwap是IBM在全世界各个国家的注册商标)。

[0033] 网络108可以包括存储区域网络(SAN)、局域网(LAN)、内联网、因特网、广域网(WAN)、对等网络、无线网络、仲裁环路网络等。存储装置104a、104b可以各自在以下项中实

现：一个或多个存储设备、或者被配置为简单磁盘捆绑 (JBOD)、直接存取存储设备 (DASD)、独立磁盘冗余阵列 (RAID) 阵列、虚拟化设备、磁带存储装置、闪存等的存储设备阵列。存储设备可以包括硬盘驱动器、包括固态电子器件的固态存储设备 (SSD)、EEPROM (电可擦式可编程只读存储器)、闪存、闪存盘、随机存取存储器 (RAM) 驱动器、存储级存储器 (SCM) 等、相变存储器 (PCM)、电阻随机存取存储器 (RRAM)、自旋转移力矩存储器 (STM-RAM)、导电桥接 RAM (CBRAM)、硬盘驱动器、光盘、磁带等。尽管示出某一数量的元件实例 (例如节点组、受管组件、邮箱等), 但可以存在任何数量的这些组件。

[0034] 图2将拷贝关系200_i的实例 (其可以包括拷贝关系200、200a、200b的一个实例) 的一个实施例示为包括: 拷贝对标识符 (ID) 202; 从中拷贝数据的主卷204 (其可以包括卷160a或106b之一); 数据被镜像到的辅助卷206 (其可以包括卷106a或106b之一); 以及更改记录位图208, 其指示主卷204中需要被拷贝或同步到辅助卷206的数据单元或磁道。当更改记录位图208中指示的所有更新被拷贝到辅助卷206时, 拷贝关系200_i达到双工或同步状态。更改记录位图208可以被初始化以指示磁道不需要被同步。当更新磁道时, 位图208中的对应位被设置为指示磁道需要被拷贝到辅助卷206。

[0035] 图3示出时间点拷贝命令300的一个实施例, 时间点拷贝命令300包括: 时间点拷贝302命令操作符; 进行PiT拷贝操作的源卷304; 以及软隔离超控 (soft fence override) 306, 其指示I/O管理器112a、112b即使在源卷304被与I/O操作隔离 (fenced-off) 的情况下也允许PiT拷贝命令继续。如果在源卷304处存在到另一个存储系统处的对应卷的故障转移, 则源卷304可以被隔离。隔离状态阻止对处于该状态的卷304的读取和写入。即使源卷304处于隔离状态, 参数306也允许时间点拷贝命令继续。如果参数306未指示当卷304处于隔离状态时允许时间点命令继续, 则时间点命令可能被隔离状态所阻止。

[0036] 时间点拷贝以看似即时的方式复制数据, 并且允许主机在到拷贝卷的实际数据传输被推迟到稍后时间时继续访问源卷。时间点拷贝看似即时, 因为响应于生成关系数据结构, 向拷贝操作返回输入/输出 (“I/O”) 完成而不将数据从源卷复制到目标卷。时间点拷贝技术通常在时间点拷贝关系建立时推迟源卷中的数据到拷贝目标卷的传输, 直到请求对源卷上的该数据块的写入操作。数据传输还可以作为后台拷贝过程继续且对系统性能具有最小影响。响应于时间点拷贝命令而立即建立的时间点拷贝关系包括位图或其它数据结构, 所述位图或其它数据结构指示块在源卷或复制卷处的卷中的位置。时间点拷贝包括源卷中的数据与要由传输到目标卷的更新盖写的数据的组合。

[0037] 图4示出由主机100、主存储系统102a、以及辅助存储系统102b组件中的组件为了执行从主存储系统102a到辅助存储系统102b的故障转移而执行的操作的一个实施例。系统102b可以用作主存储系统, 并且系统102a可以用作辅助存储系统以便进行故障转移。控制开始于主机复制管理器114启动 (在方框400) 镜像拷贝操作以便在主卷106a与辅助卷106b之间同步数据。主机复制管理器114可以在主存储卷106a与对应辅助存储卷106b之间建立 (在方框402) 拷贝关系200, 以便以同步拷贝模式将数据和更新从主卷106a镜像到对应辅助卷106b。拷贝关系200被提供给主复制管理器114a以便用于镜像/同步数据。在同步拷贝模式下, 直到辅助存储系统102a确认所拷贝的数据已被存储在对应辅助卷106b中, 该数据才被指示为已完成。在一个实施例中, 可以以一致性组模式将数据从主卷106a镜像到辅助卷106b, 以使得辅助卷106b处截至某一时间点的数据与主卷106a处的数据保持一致。

[0038] 故障转移管理器116a可以检测(在方框404)主存储系统102a处的故障事件,例如主存储系统102a的一个或多个组件(包括主存储装置104a的存储设备)的故障。在检测到(在方框404)故障转移事件时,主故障转移管理器116a或主机故障转移管理器116a可以确定(在方框406)主卷106a中经历数据丢失的磁道,并且记录主存储系统102a中的那些已发生故障的磁道。主存储系统102a随后可以将关于已发生故障的磁道的信息报告给主机116或辅助116b故障转移管理器,以便在从辅助存储系统102b重新同步回主存储系统102a期间使用。不能够操作的主卷106a中的磁道可以包括所述卷中的磁道的子集,其中其它磁道可能未经历数据丢失。

[0039] 主机故障转移管理器116然后可以与主116a和辅助116b故障转移管理器协调以从主存储系统102a处的主卷106a启动(在方框408)故障转移,以便将I/O请求重定向到辅助存储系统102b处的对应辅助卷106b。作为故障转移的一部分,主机复制管理器114可以暂停将数据从主卷106a同步到辅助卷106b的拷贝关系200a。主机故障转移管理器116可以进一步针对主卷106a建立(在方框410)软隔离状态。软隔离状态防止对交换到辅助卷106b的一个或多个主卷106a进行I/O操作,以便防止在故障转移之后对这些卷106a的任何意外访问。

[0040] 主机复制管理器114或故障转移管理器116然后可以针对主存储系统102a中的每个主卷i执行方框412到422的操作循环。对于主卷i,主机复制管理器114或故障转移管理器116可以针对卷i启动(在框414)PiT拷贝操作,且软隔离超控参数306被设置为指示即使软隔离状态针对卷i有效,PiT拷贝操作也应该继续。主机复制管理器114或故障转移管理器116判定(在方框416)卷i是否能够操作。在一个实施例中,可以通过启动设置了软隔离超控参数306的时间点拷贝命令300进行该判定,参数306被设置为指定即使卷i处于软隔离状态,也允许时间点命令300继续,这种情况将与方框406的操作的结果相同。如果时间点拷贝操作对于卷i成功,则卷i被确定为能够操作,否则如果时间点拷贝操作失败,则卷i被确定为不能够操作。通过获得能够操作的主卷106a的时间点拷贝,这些主卷106a的截至故障转移事件的时间点的数据被保留。如果其它恢复选项失败,则这些时间点拷贝可以用于数据恢复。作为方框414的操作的一部分,可以使用其它技术判定主卷106a是否能够操作。

[0041] 如果(在方框416)主卷i能够操作,则主机复制管理器114(或故障转移管理器116)创建(在方框418)拷贝关系200_i以从对应于主卷i的辅助存储装置102b处的辅助卷i启动重新同步操作,以便针对更改记录位图中指示的更新,以异步拷贝模式将对辅助卷i的更新传输到主卷i。当接收对辅助卷106b的更新时,设置更改记录位图208中用于辅助卷i的对应位。辅助复制管理器114b扫描更改记录位图208,从而查找辅助卷106b处的更新以便重新同步或拷贝到在故障转移期间保持能够操作的对应主卷106a。在异步拷贝模式下,复制完成而不必从主卷i接收数据已被接收的确认。在一个实施例中,异步模式可以包括非一致性组模式(CGM),以使得数据未被确保在辅助106b卷与主106a卷之间一致。作为在到辅助存储系统102b的故障转移期间的重新同步的一部分,通过使辅助复制管理器114自动地将辅助卷106b处的更新异步地拷贝到那些能够操作的主卷106a,主卷106a处的数据被尽可能保持最新,以使得在主存储系统102a被修复并且能够完全操作之后,能够更快地完成主存储系统102a的重新同步。使用所描述的实施例,通过由故障转移管理器116或116b控制的故障转移过程来自动启动重新同步回能够操作的卷。

[0042] 此外,与同步地执行拷贝相比,异步地执行重新同步拷贝显著降低对访问辅助卷

的主机应用的性能影响。此外,在执行任何故障分析和设备修复(如果需要)之前,能够在故障转移之后立即启动异步拷贝。

[0043] 如果(在方框416)卷*i*不能够操作,则主机复制114或故障转移116管理器针对从辅助卷*i*到主卷*i*的重新同步创建(在方框420)暂停的拷贝关系200_{*i*},并且针对卷*i*更新更改记录位图208以便将报告为具有数据丢失的任何磁道指示为已更新。在更改记录位图208中将经历数据丢失的卷*i*中的磁道指示为已更新将导致辅助复制管理器114b在卷*i*被修复并能够操作之后的后续重新同步期间复制具有数据丢失的磁道的数据,无论该数据在故障转移期间是否已被更新。

[0044] 图4的操作结果是在故障转移之后,辅助存储系统102b立即开始将对辅助卷106b的更新重新同步到保持能够操作的主卷106a。对于不能够操作的主卷106a,对应辅助卷106b的更改记录位图208被设置为将经历数据丢失的磁道指示为已更新,并且重新同步被暂停,直到不能够操作的主卷被确定为能够操作。通过重新同步在故障转移期间能够操作的那些主卷,主存储系统将在主存储系统102a变得能够完全操作之后很快地准备好进行故障转移,因为在辅助存储系统102b处的故障转移期间,辅助存储系统102b处的大量更新可能已经被重新同步到主卷106a。

[0045] 在图4的操作中,针对存储系统的卷确定可操作性。在备选实施例中,能够针对除卷以外的存储单元(例如存储设备、逻辑分区、物理分区、逻辑驱动器等)确定镜像和可操作性。

[0046] 图5示出由主机复制管理器114和故障转移管理器116为了在故障事件之后恢复主存储系统102a而执行的操作的一个实施例。主机故障转移管理器114在从故障转移开始的预定延迟之后,向主存储系统102a发出(在方框500)运行状况查询以便判定主存储系统102a是否能够操作。对于许多故障类型,在经过某一时间之后,主存储系统102a可能已通过内部恢复操作和程序而恢复。在从主存储系统102a接收(在方框502)对运行状况查询的响应之后,如果(在方框504)响应指示主存储系统102a能够完全操作,则主机复制管理器114可以执行(在方框506)与先前不能够操作的主卷106a对应的辅助卷106b中的的每一者的时间点拷贝操作,以便提供可以用于恢复(如果需要)的一致副本。主机复制管理器114启动(在方框508)重新同步,以便针对与先前指示为不能够操作的主卷106a对应的每个辅助卷106b,同步地拷贝在更改记录位图208中被指示为已更新的数据。可以通过取消暂停拷贝关系200b来启动重新同步,以便将辅助卷重新同步到先前不能够操作的主卷106a。用于重新同步到先前不能够操作的主卷106a的更改记录位图208将对应辅助卷106b中被更新的磁道指示为已更新,并且将不能够操作的主卷中被报告的磁道指示为经历数据丢失。以这种方式,恢复后的不能够操作的主卷106a的重新同步仅拷贝已更新的数据和经历数据丢失的磁道的数据,并且不拷贝整个恢复后的不能够操作的卷。这通过避免拷贝辅助卷中与恢复后的主卷106a中的在故障转移期间未被更新且未经历数据丢失的磁道对应的磁道而加快了重新同步。

[0047] 辅助复制管理器114b进一步将能够操作的卷的更新(当I/O请求被重定向到辅助存储系统102b时,这些更新自故障转移开始以来已被异步地拷贝)的重新同步转变(在方框514)为同步拷贝模式,以使得在恢复主存储系统102a之后,同步地拷贝要重新同步的剩余数据。

[0048] 如果(在方框504)对运行状况查询的响应指示主存储系统102a不能完全操作,则主机故障转移管理器114收集(在方框510)用于主存储系统102a的修复的诊断信息以便报告给管理员。管理员然后可以基于所收集的诊断信息,例如通过更换和/或修复主存储系统102a中的单元、组件和存储设备,继续修复已发生故障的主存储系统102。在完成修复之后,管理员可以启动用于执行重新同步的重新同步命令。在方框512,主机复制管理器114或辅助复制管理器114b可以从管理员接收手动重新同步命令,并且然后进行到方框506以便针对先前指示为不能够操作的卷启动重新同步。

[0049] 使用所描述的实施例,在修复和恢复之后,所恢复的不能够操作的主卷106a中仅在故障转移期间被更新或经历数据丢失的磁道(这些磁道在针对已发生故障的卷106a建立的拷贝关系200的更改记录位图中被指示)才需要被同步。此外,通过在能够操作的主卷106a的故障转移期间启动重新同步,在主存储系统102a被恢复之后的同步的完成快得多,因为大部分数据可能已经被重新同步。

[0050] 本发明可以是系统、方法和/或计算机程序产品。计算机程序产品可以包括计算机可读存储介质,其上载有用于使处理器实现本发明的各个方面的计算机可读程序指令。

[0051] 计算机可读存储介质可以是可以保持和存储由指令执行设备使用的指令的有形设备。计算机可读存储介质例如可以是一但不限于一电存储设备、磁存储设备、光存储设备、电磁存储设备、半导体存储设备或者上述的任意合适的组合。计算机可读存储介质的更具体的例子(非穷举的列表)包括:便携式计算机盘、硬盘、随机存取存储器(RAM)、只读存储器(ROM)、可擦式可编程只读存储器(EPROM或闪存)、静态随机存取存储器(SRAM)、便携式压缩盘只读存储器(CD-ROM)、数字多功能盘(DVD)、记忆棒、软盘、机械编码设备、例如其上存储有指令的打孔卡或凹槽内凸起结构、以及上述的任意合适的组合。这里所使用的计算机可读存储介质不被解释为瞬时信号本身,诸如无线电波或者其它自由传播的电磁波、通过波导或其它传输介质传播的电磁波(例如,通过光纤电缆的光脉冲)、或者通过电线传输的电信号。

[0052] 这里所描述的计算机可读程序指令可以从计算机可读存储介质下载到各个计算/处理设备,或者通过网络、例如因特网、局域网、广域网和/或无线网下载到外部计算机或外部存储设备。网络可以包括铜传输电缆、光纤传输、无线传输、路由器、防火墙、交换机、网关计算机和/或边缘服务器。每个计算/处理设备中的网络适配卡或者网络接口从网络接收计算机可读程序指令,并转发该计算机可读程序指令,以供存储在各个计算/处理设备中的计算机可读存储介质中。

[0053] 用于执行本发明操作的计算机可读程序指令可以是汇编指令、指令集架构(ISA)指令、机器指令、机器相关指令、微代码、固件指令、状态设置数据、或者以一种或多种编程语言的任意组合编写的源代码或目标代码,所述编程语言包括面向对象的编程语言—诸如Java、Smalltalk、C++等,以及常规的过程式编程语言—诸如“C”语言或类似的编程语言。计算机可读程序指令可以完全地在用户计算机上执行、部分地在用户计算机上执行、作为一个独立的软件包执行、部分在用户计算机上部分在远程计算机上执行、或者完全在远程计算机或服务器上执行。在涉及远程计算机的情形中,远程计算机可以通过任意种类的网络—包括局域网(LAN)或广域网(WAN)—连接到用户计算机,或者,可以连接到外部计算机(例如利用因特网服务提供商来通过因特网连接)。在一些实施例中,通过利用计算机可读

程序指令的状态信息来个性化定制电子电路,例如可编程逻辑电路、现场可编程门阵列(FPGA)或可编程逻辑阵列(PLA),该电子电路可以执行计算机可读程序指令,从而实现本发明的各个方面。

[0054] 这里参照根据本发明实施例的方法、装置(系统)和计算机程序产品的流程图和/或框图描述了本发明的各个方面。应当理解,流程图和/或框图的每个方框以及流程图和/或框图中各方框的组合,都可以由计算机可读程序指令实现。

[0055] 这些计算机可读程序指令可以提供给通用计算机、专用计算机或其它可编程数据处理装置的处理器,从而生产出一种机器,使得这些指令在通过计算机或其它可编程数据处理装置的处理器执行时,产生了实现流程图和/或框图中的一个或多个方框中规定的功能/动作的装置。也可以把这些计算机可读程序指令存储在计算机可读存储介质中,这些指令使得计算机、可编程数据处理装置和/或其它设备以特定方式工作,从而,存储有指令的计算机可读存储介质则包括一个制品,其包括实现流程图和/或框图中的一个或多个方框中规定的功能/动作的各个方面的指令。

[0056] 也可以把计算机可读程序指令加载到计算机、其它可编程数据处理装置、或其它设备上,使得在计算机、其它可编程装置或其它设备上执行一系列操作步骤,以产生计算机实现的过程,从而使得在计算机、其它可编程装置、或其它设备上执行的指令实现流程图和/或框图中的一个或多个方框中规定的功能/动作。

[0057] 附图中的流程图和框图显示了根据本发明的不同实施例的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上,流程图或框图中的每个方框可以代表一个模块、程序段或指令的存储单元,所述模块、程序段或指令的存储单元包含一个或多个用于实现规定的逻辑功能的可执行指令。在有些作为替换的实现中,方框中所标注的功能可以以不同于附图中所标注的顺序发生。例如,两个连续的方框实际上可以基本并行地执行,它们有时也可以按相反的顺序执行,这依所涉及的功能而定。也要注意的,框图和/或流程图中的每个方框、以及框图和/或流程图中的方框的组合,可以用执行规定的功能或动作的专用的基于硬件的系统来实现,或者可以用专用硬件与计算机指令的组合来实现。

[0058] 图1的计算组件(包括主机100和存储系统102a、102b)可以在一个或多个计算机系统(例如图6中所示的计算机系统602)中实现。计算机系统/服务器602可以在由计算机系统执行的计算机系统可执行指令(诸如程序模块)的一般语境下描述。通常,程序模块可以包括执行特定的任务或者实现特定的抽象数据类型的例程、程序、目标程序、组件、逻辑、数据结构等。计算机系统/服务器602可以在通过通信网络链接的远程处理设备执行任务的分布式云计算环境中实施。在分布式云计算环境中,程序模块可以位于包括存储设备的本地和远程计算机系统存储介质上。例如,在云计算环境中,存储系统102a、102b可以包括云计算节点,它们提供作为基于云的服务的存储服务。

[0059] 如图6中所示,计算机系统/服务器602以通用计算设备的形式表现。计算机系统/服务器602的组件可以包括但不限于:一个或者多个处理器或者处理单元604;系统存储器606;将包括系统存储器606的不同系统组件连接到处理器604的总线608。总线608表示几类总线结构中的一种或多种,包括存储器总线或者存储器控制器,外围总线,图形加速端口,处理器或者使用多种总线结构中的任意总线体系架构的局域总线。举例来说,这些体系架构包括但不限于工业标准体系架构(ISA)总线,微通道体系架构(MCA)总线,增强型ISA

(EISA) 总线、视频电子标准协会 (VESA) 局域总线以及外围组件互连 (PCI) 总线。

[0060] 计算机系统/服务器602典型地包括多种计算机系统可读介质。这些介质可以是能够被计算机系统/服务器602访问的任意可获得的介质,包括易失性和非易失性介质、可移动的和不可移动的介质。

[0061] 系统存储器606可以包括易失性存储器形式的计算机系统可读介质,例如随机存取存储器 (RAM) 610和/或高速缓冲存储器612。计算机系统/服务器602可以进一步包括其它可移动/不可移动的、易失性/非易失性计算机系统存储介质。仅作为举例,存储系统613可以用于读写不可移动的、非易失性磁介质(图6中未显示,通常称为“硬盘驱动器”)。尽管图6中未示出,可以提供用于对可移动非易失性磁盘(例如“软盘”)读写的磁盘驱动器,以及对可移动非易失性光盘(例如CD-ROM、DVD-ROM或者其它光介质)读写的光盘驱动器。在这些情况下,每个组件可以通过一个或者多个数据介质接口与总线608相连。如下面进一步示出和描述的,存储器606可以包括至少一个程序产品,该程序产品具有一组(例如至少一个)程序模块,这些程序模块被配置以执行本发明实施例的功能。

[0062] 具有一组(至少一个)程序模块616的程序/实用工具614,可以存储在存储器606中,这样的程序模块616包括但不限于操作系统、一个或者多个应用程序、其它程序模块以及程序数据,这些示例中的每一个或某种组合中可能包括网络环境的实现。计算机602的组件可以实现为程序模块616,这些程序模块通常执行在此描述的本发明实施例的功能和/或方法。图1的系统可以在一个或多个计算机系统602中实现,其中如果这些系统在多个计算机系统602中实现,则计算机系统可以通过网络通信。

[0063] 计算机系统/服务器602也可以与一个或多个外部设备618(例如键盘、指点设备、显示器620等)通信,还可与一个或者多个使得用户能与该计算机系统/服务器602交互的设备通信,和/或与使得该计算机系统/服务器602能与一个或多个其它计算设备进行通信的任何设备(例如网卡,调制解调器等等)通信。这种通信可以通过输入/输出(I/O)接口622进行。并且,计算机系统/服务器602可以通过网络适配器624与一个或者多个网络(例如局域网(LAN),广域网(WAN)和/或公共网络,例如因特网)通信。如图所示,网络适配器624通过总线608与计算机系统/服务器602的其它组件通信。应当理解,尽管图中未示出,其它硬件和/或软件组件可以与计算机系统/服务器602一起操作,包括但不限于:微代码、设备驱动器、冗余处理单元、外部磁盘驱动阵列、RAID系统、磁带驱动器以及数据备份存储系统等。

[0064] 这里使用的诸如*i*之类的参考字符在此用于表示元素的实例的可变数量,它们可以表示相同或不同值,并且当与不同所述实例中的不同或相同元素使用时,可以表示相同或不同值。

[0065] 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“本发明的一个或多个(但不是所有)实施例”,除非另外明确指出。

[0066] 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。

[0067] 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。

[0068] 术语“一”、“一个”和“所述”指“一个或多个”,除非另外明确指出。

[0069] 相互通信的设备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通过一个或多个媒介直接或间接地通信。

[0070] 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,描述了多种可选的组件以例示本发明的更多可能的实施例。

[0071] 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,当在此描述多个设备或物品(无论它们是否协作)时,将显而易见的是,可以使用单个设备/物品来代替多个设备或物品,或者可以使用不同数量的设备/物品而不是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它未显式描述为具有此类功能/特性的设备来体现。因此,本发明的其它实施例不必包括设备本身。

[0072] 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。

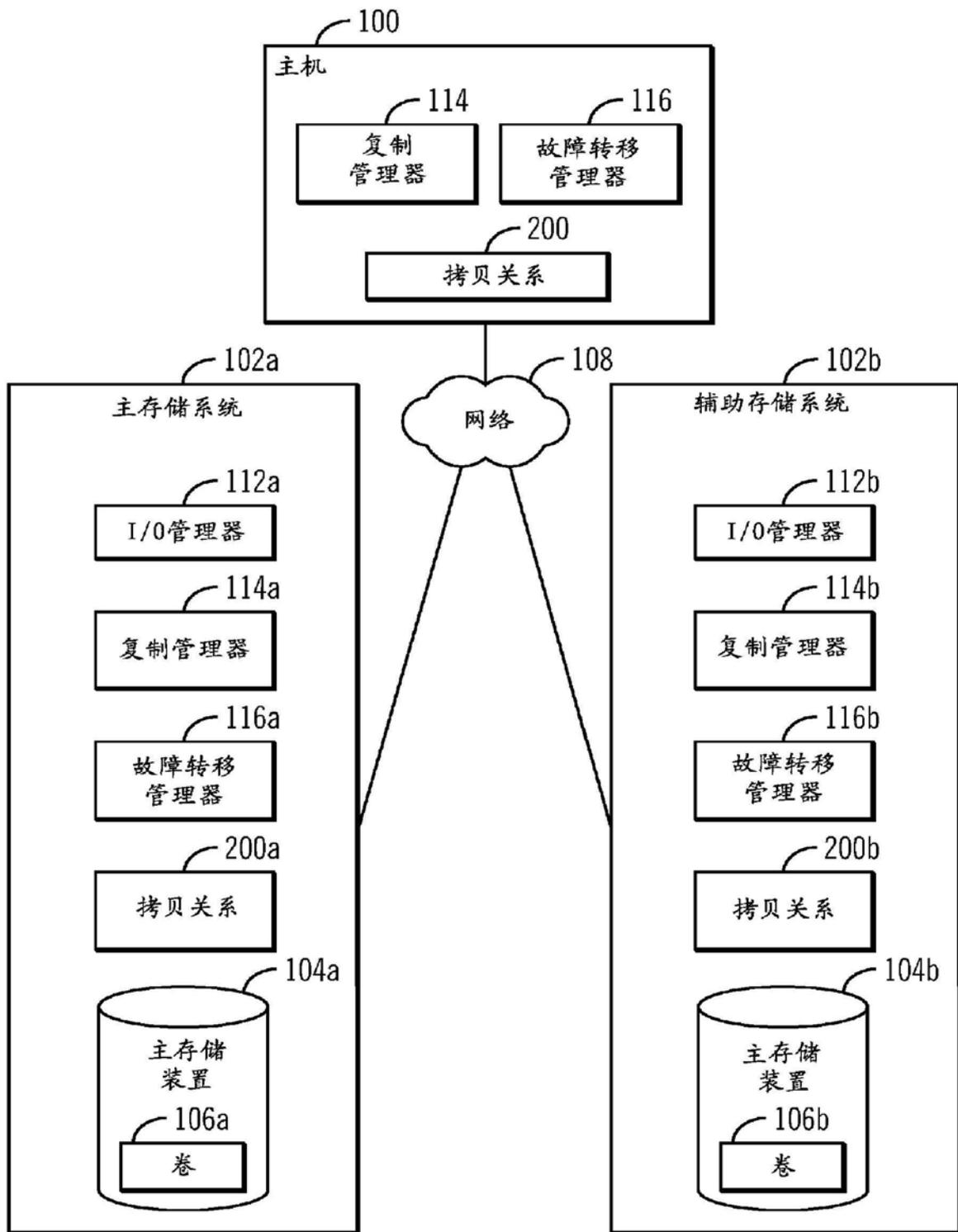
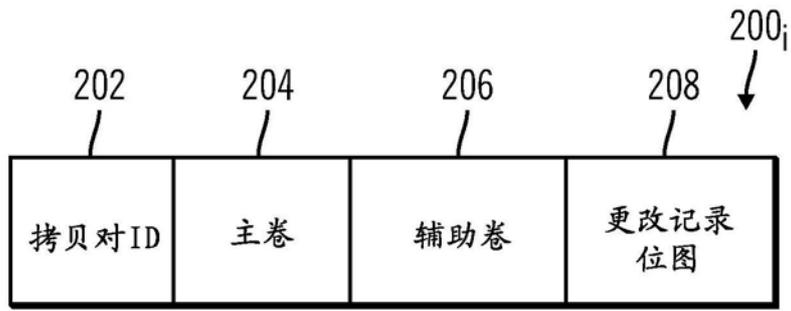
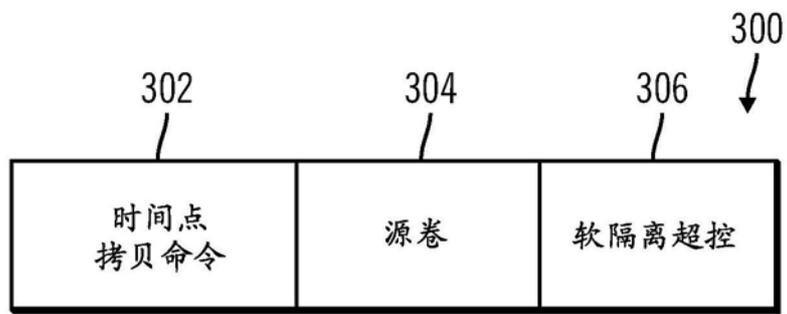


图1



拷贝关系

图2



时间点拷贝命令

图3

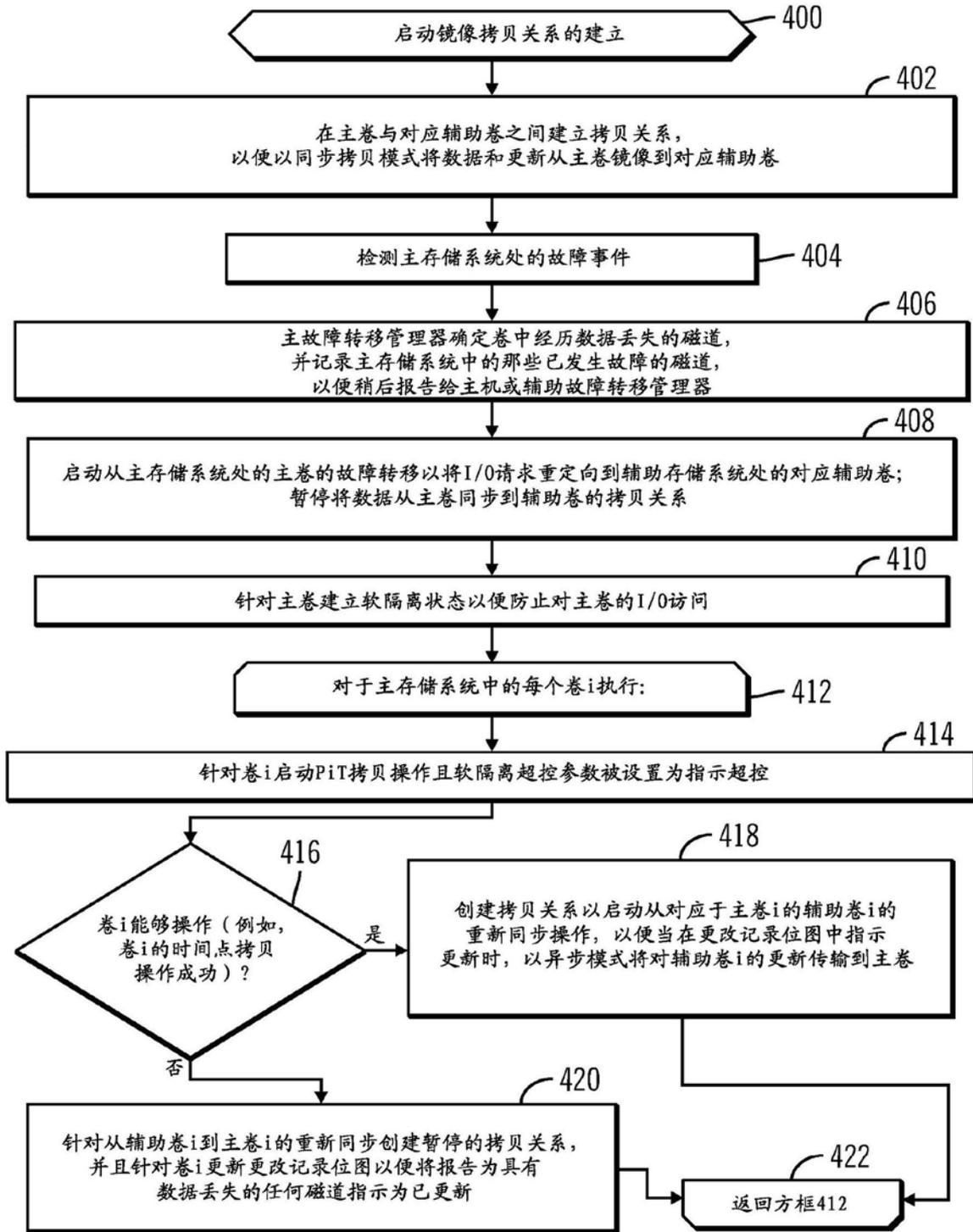


图4

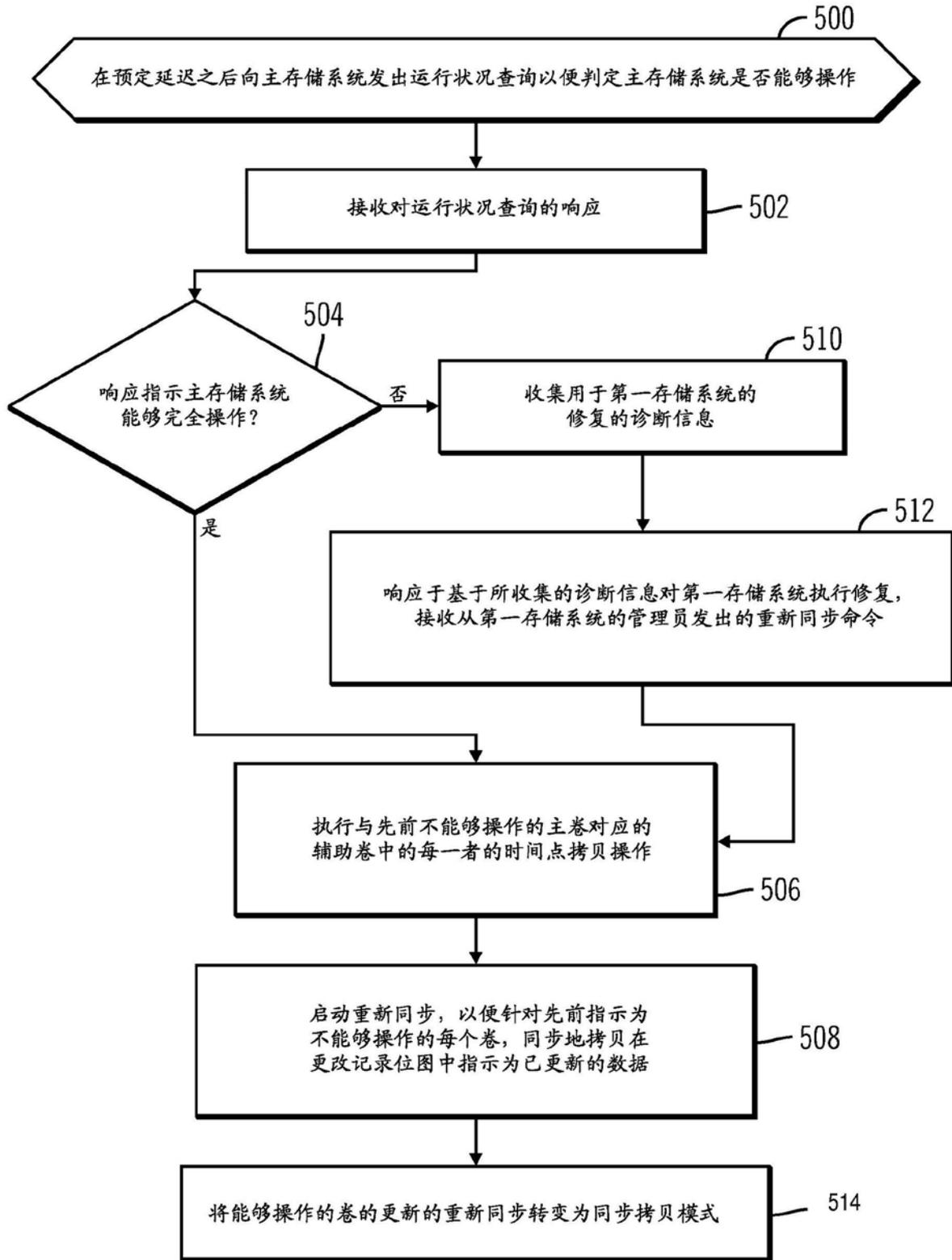


图5

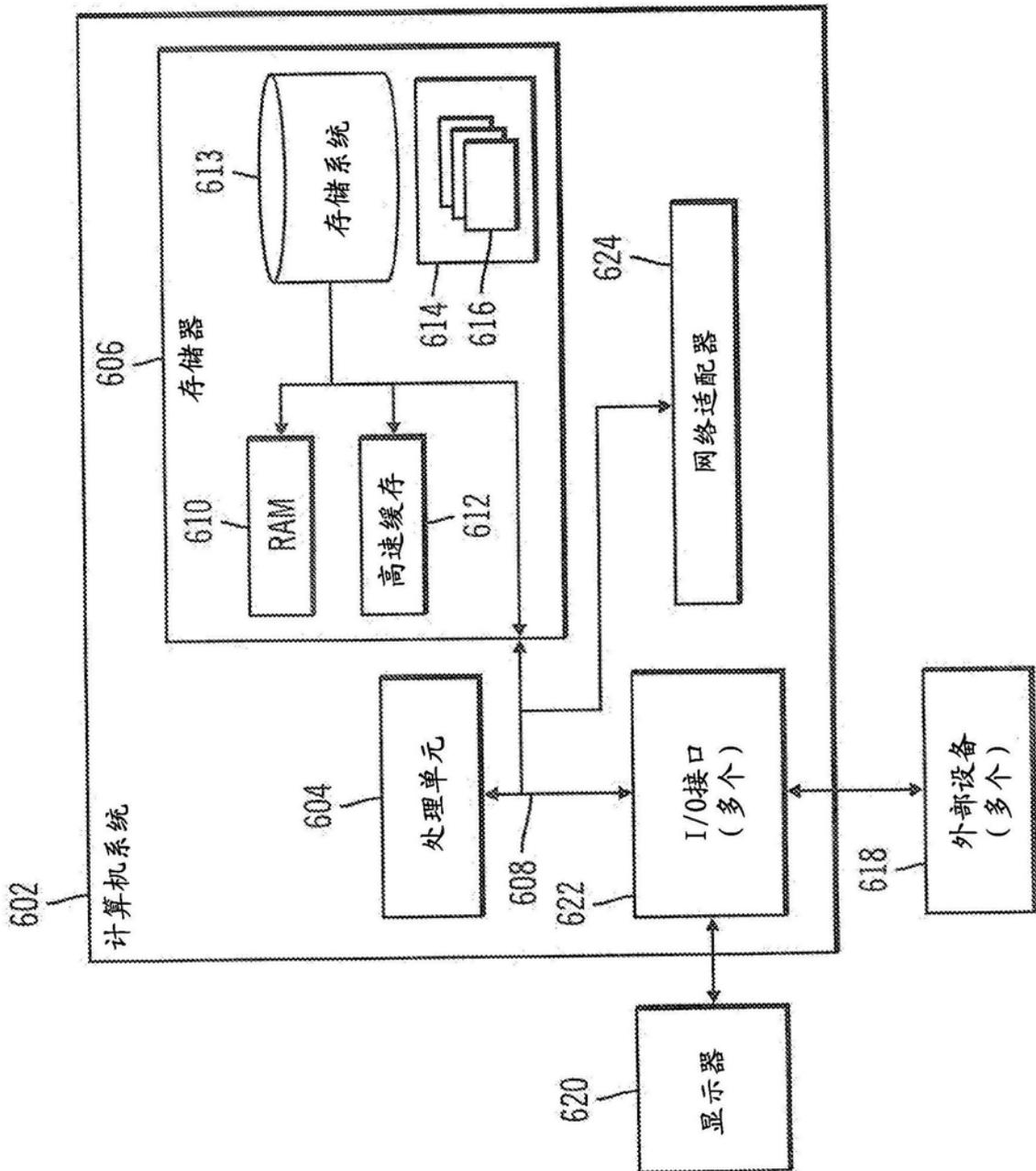


图6