

(12) 发明专利

(10) 授权公告号 CN 101501652 B

(45) 授权公告日 2013. 03. 20

(21) 申请号 200680003967. 9

(51) Int. Cl.

(22) 申请日 2006. 02. 23

G06F 12/00(2006. 01)

(30) 优先权数据

(56) 对比文件

11/072, 141 2005. 03. 04 US

US 6154852 A, 2000. 11. 28,

11/072, 128 2005. 03. 04 US

CN 1227644 A, 1999. 09. 01,

11/072, 612 2005. 03. 04 US

US 6611901 B1, 2003. 08. 26,

(85) PCT申请进入国家阶段日

审查员 李东

2007. 08. 03

(86) PCT申请的申请数据

PCT/US2006/006474 2006. 02. 23

(87) PCT申请的公布数据

W02006/096339 EN 2006. 09. 14

(73) 专利权人 伊姆西公司

地址 美国马萨诸塞州

(72) 发明人 肯尼思·J·加利波

罗伯特·K·卡斯特 埃兰·奥加德

(74) 专利代理机构 北京金信立方知识产权代理

有限公司 11225

代理人 黄威

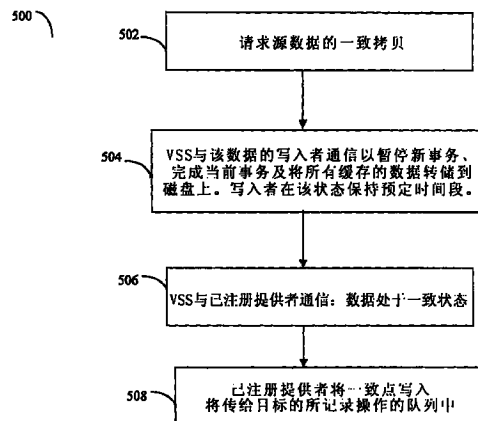
权利要求书 5 页 说明书 11 页 附图 14 页

(54) 发明名称

检查点及一致性标记符

(57) 摘要

本发明描述了用于获取处于一致状态的源数据拷贝的方法、计算机程序产品和系统。具有相应时序的、修改所述源数据的一个或多个文件操作被记录。接收请求一致状态的源数据拷贝的请求。确定在相应的时序中所述源数据在应用一部分文件操作之后在哪一点处于一致状态。标记相应的时序中所述源数据处于一致状态的那一点。被确定使源数据处于一致状态的那部分文件操作应用于源数据的拷贝。



1. 用于获取一致状态的源数据拷贝的方法,包括:  
记录一个或多个具有相应时序的文件操作,其修改所述源数据;  
接收请求一致状态的源数据拷贝的请求;  
当应用一部分所述文件操作之后确定在所述相应时序中所述源数据在哪一点处于一致状态;  
在所述相应时序中标记所述源数据处于一致状态的所述点;及  
将确定使所述源数据处于一致状态的那部分文件操作应用于所述源数据的所述拷贝。
2. 根据权利要求 1 的方法,其中所述标记所述相应时序中的所述点包括:  
确定所述文件操作中所述一致状态的相应文件操作。
3. 根据权利要求 1 的方法,还包括:  
实时记录应用于源数据的所述部分文件操作,其中所述部分文件操作被记录为时间有序的一系列记录,每一所述记录对应于所述部分中的所述文件操作之一;  
将标记符记录插入所述一系列记录中,其指明所述源数据何时处于一致状态;及  
将对应于所述一系列记录的文件操作应用于源数据的所述拷贝,直到标记符记录被确定为止。
4. 根据权利要求 3 的方法,还包括:  
当使用所述源数据的所述拷贝执行操作时暂停进一步将所记录的文件操作应用于源数据的所述拷贝,同时允许新输入的文件操作应用于所述源数据。
5. 根据权利要求 4 的方法,其中所述暂停进一步将所记录的文件操作应用于源数据的所述拷贝通过在包括所述源数据的所述拷贝的目标系统上运行目标脚本进行控制。
6. 根据权利要求 5 的方法,还包括:  
将对应于所述部分文件操作的记录传输给所述目标系统,其中所述传输所述记录根据包括所述源数据的源系统上的处理而被暂停。
7. 根据权利要求 6 的方法,还包括:  
在所述源系统上运行源脚本以控制所述记录从所述源系统到所述目标系统的传输。
8. 根据权利要求 7 的方法,其中所述标记符记录包括对应于至少下述之一的指示符:  
所述源脚本、所述目标脚本、指示是否暂停将所述部分数据操作应用于所述源数据的所述拷贝的第一标志值、指示是否暂停将所述记录从所述源系统传输给所述目标系统的第二标志值。
9. 根据权利要求 4 的方法,其中使用所述源数据的所述拷贝执行的所述操作是所述源数据的备份操作。
10. 根据权利要求 1 的方法,其中所述请求处于一致状态的源数据的拷贝的请求使用连接到协调对所述源数据的访问的组成部分的应用程序编程接口提出,响应于接收到所述请求,所述组成部分执行以下处理:  
使所述源数据的写入者完成现有的事务、不开始任何新的事务、及转储数据高速缓存保存的所述源数据,所述写入者在预定时间段内暂停对所述源数据开始新事务;  
与已注册组成部分通信:所述源数据处于一致状态;及  
与所述写入者通信以重新开始新事务。
11. 根据权利要求 10 的方法,其中所述已注册组成部分通过将标记符记录插入对应于

所述部分文件操作的时间有序的一系列记录中而实现所述在所述相应的时序中标记所述点,所述方法进一步包括:

将所述记录从包括所述源数据的源位置传输给包括所述源数据的所述拷贝的目标位置,其中不管是否有另外的数据操作应用于所述源数据,所述部分文件操作均被应用于所述拷贝。

12. 一种获取一致状态的源数据拷贝的系统,包括:

包括数据源的第一数据存储系统;

包括数据目标的第二数据存储系统;

包括在连接到所述第一数据存储系统的第一主机,其包括:

镜像驱动器,记录一个或多个具有相应时序的文件操作,所述文件操作修改所述源数据;

服务组件,接收请求一致状态的源数据拷贝的请求,当应用一部分所述文件操作之后确定在所述相应时序中所述源数据在哪一点处于一致状态;及

复制服务组件,在所述相应时序中标记所述源数据处于一致状态的所述点;及

包括在连接到所述第二数据存储系统的第二主机,其包括:

复制服务组件,将确定使所述源数据处于一致状态的那部分

文件操作应用于所述源数据的所述拷贝。

13. 根据权利要求 12 的系统,其中所述请求处于一致状态的源数据的拷贝的请求使用包括在所述第一主机中的、连接到所述第一主机中包括的协调对所述源数据的访问的组成部分的应用程序编程接口提出,所述组成部分执行以下处理:

使所述源数据的写入者完成现有的事务、不开始任何新的事务、及转储数据高速缓存保存的所述源数据,所述写入者根据至少一条件暂停开始新事务;

与已注册组成部分通信:所述源数据处于一致状态;及

与所述写入者通信以根据所述至少一条件重新开始新事务。

14. 用于捕获应用于数据集的文件操作的方法,包括:

记录一个或多个具有相应时序的文件操作,其修改所述数据集;

接收请求一致状态的数据集的拷贝的请求;

当应用一部分所述文件操作之后确定在所述相应时序中所述数据集在哪一点处于一致状态;及

在所述相应时序中标记所述数据集处于一致状态的所述点。

15. 根据权利要求 14 的方法,其中所述在所述相应时序中标记所述点包括:

确定所述文件操作中获得所述一致状态的相应文件操作。

16. 根据权利要求 14 的方法,还包括:

实时记录应用于数据集的所述部分文件操作,其中所述部分文件操作被记录为时间有序的一系列记录,每一所述记录对应于所述部分中的所述文件操作之一;及

将标记符记录插入所述一系列记录中,其指明所述数据集何时处于一致状态。

17. 根据权利要求 16 的方法,还包括:

从与所述数据集相关联的系统传输对应于所述部分文件操作的记录,其中所述传输所述记录根据所述系统上执行的而被暂停。

18. 根据权利要求 17 的方法,还包括:

在所述系统上运行源脚本以控制所述记录到目标的传输。

19. 根据权利要求 18 的方法,其中所述标记符记录包括对应于至少下述之一的指示符:所述源脚本、在所述记录传输给目标之后连同处理所述记录运行的目标脚本、指示是否暂停将所述部分文件操作应用于在所述目标处的所述数据集的拷贝的第一标志值、指示是否暂停从所述系统传输所述记录的第二标志值。

20. 根据权利要求 14 的方法,其中所述请求处于一致状态的数据集的拷贝的请求使用连接到协调对所述数据集的访问的组成部分的应用程序编程接口提出,响应于接收到所述请求,所述组成部分执行以下处理:

使所述数据集的写入者完成现有的事务、不开始任何新的事务、及转储数据高速缓存保存的所述数据集,所述写入者在预定时间段内暂停对源数据开始新事务;

与已注册组成部分通信:所述数据集处于一致状态;及

与所述写入者通信以重新开始新事务。

21. 根据权利要求 20 的方法,其中所述已注册组成部分通过将标记符记录插入对应于所述部分文件操作的时间有序的一系列记录中而实现所述在所述相应的时序中标记所述点。

22. 用于捕获应用于数据集的文件操作的系统,包括:

镜像驱动器部件,其在成功完成感兴趣的文件操作基础上,在一位置相对于其它已记录的感兴趣文件操作在时间有序中表示中记录对应于所述感兴趣的文件操作应用于所述数据集时的所述文件操作;及

复制服务组件,其在所述数据集处于一致状态的通知基础上,在所述位置的下一条目中记录一致点标记符,所述一致点标记符指明当到对应于所述一致点标记符的时间点所记录的感兴趣文件操作被应用于所述数据集时所述数据集处于一致状态。

23. 根据权利要求 22 的系统,还包括:

协调对所述数据集的访问的组成部分;及

连接到所述组成部分的应用程序编程接口,其中记录所述数据集处于一致状态的时间点的请求使用所述应用程序编程接口做出,响应于接收所述请求,所述组成部分执行以下处理:

使所述数据集的写入者完成现有的事务、不开始任何新的事务、及转储数据高速缓存保存的所述数据集,所述写入者在预定时间段内暂停对源数据开始新事务;

与已注册组成部分通信:所述数据集处于一致状态;及

与所述写入者通信以重新开始新事务。

24. 根据权利要求 23 的系统,还包括:

使用所述数据集执行文件操作并是所述数据集的写入者的应用程序。

25. 根据权利要求 22 的系统,其中所述镜像驱动器是当处理文件操作时包括在以内核模式运行的多个组成部分的调用链中的驱动器。

26. 根据权利要求 22 的系统,还包括:

以用户模式运行的代码,其从所述位置读信息,所述信息对应于已记录的文件操作和一致点标记符,所述代码还将所述信息传输给目标。

27. 产生源数据的时间点拷贝的方法,包括:

接收对应于应用于所述源数据的时间有序的一系列已记录文件操作的一组记录,所述一组记录包括一个或多个一致点标记符,每一一致点标记符指示当到所述时间点出现的所述已记录文件操作部分应用于所述源数据时所述源数据处于一致状态的时间点;

将对应于所述一组记录的已记录文件操作应用于源数据的所述拷贝,直到对应于一致点标记符的标记符记录被确定为止。

28. 根据权利要求 27 的方法,还包括:

当使用所述源数据的所述拷贝执行操作时暂停进一步将已记录的文件操作应用于源数据的所述拷贝。

29. 根据权利要求 28 的方法,其中所述暂停进一步将所记录的文件操作应用于源数据的所述拷贝通过在包括所述源数据的所述拷贝的目标系统上运行目标脚本进行控制。

30. 根据权利要求 29 的方法,其中所述标记符记录包括对应于至少下述之一的指示符:所述目标脚本、指示是否暂停将所记录文件操作应用于所述源数据的所述拷贝的第一标志值。

31. 根据权利要求 28 的方法,其中使用所述源数据的所述拷贝执行的所述操作是所述源数据的备份操作。

32. 根据权利要求 27 的方法,其中所述一个或多个一致点标记符中的第一一致点标记符对应于所述源数据在第一时间点的第一版本,所述一个或多个一致点标记符中的第二一致点标记符对应于所述源数据在第二时间点的第二版本,所述方法还包括:

在检测到所述第一一致点标记符的基础上,暂停进一步应用所记录文件操作第一时间段,同时所述源数据的所述拷贝的当前状态对应于所述第一版本;

重新开始处理所述记录组,直到检测到所述第二一致点标记符为止;

在检测到所述第二一致点标记符的基础上,暂停进一步应用所记录文件操作第二时间段,同时所述源数据的所述拷贝的当前状态对应于所述第二版本;及

重新开始处理所述记录组。

33. 根据权利要求 32 的方法,其中所述暂停步骤和所述重新开始步骤均根据标志值和包括所述源数据的所述拷贝的目标系统上运行的脚本进行控制。

34. 根据权利要求 33 的方法,其中所述脚本的位置包括在对应于所述第一和第二一致点标记符中的每一个的标记符记录的字段中,及所述标志值包括在每一标记符记录的另一字段中。

35. 用于产生源数据的时间点拷贝的目标系统,包括:

包括所述源数据的所述拷贝的数据存储装置;

复制服务组件,执行以下处理:

接收对应于应用于所述源数据的时间有序的一系列已记录文件操作的一组记录,所述记录组包括一个或多个一致点标记符,每一一致点标记符指示当到所述时间点出现的所述已记录文件操作部分应用于所述源数据时所述源数据处于一致状态的时间点;以及

将对应于所述记录组的已记录文件操作应用于源数据的所述拷贝,直到对应于一致点标记符的标记符记录被确定为止。

36. 根据权利要求 35 的目标系统,进一步配置为执行以下处理:

当使用所述源数据的所述拷贝执行操作时暂停进一步将已记录的文件操作应用于源数据的所述拷贝,其中暂停进一步将所记录的文件操作应用于源数据的所述拷贝通过在包括所述源数据的所述拷贝的目标系统上运行目标脚本进行控制。

37. 根据权利要求 36 的目标系统,其中所述标记符记录包括对应于至少下述之一的指示符:所述目标脚本、指示是否暂停将所记录文件操作应用于所述源数据的所述拷贝的第一标志值。

38. 根据权利要求 35 的目标系统,其中所述一个或多个一致点标记符中的第一标记符对应于所述源数据在第一时间点的第一版本,所述一个或多个一致点标记符中的第二标记符对应于所述源数据在第二时间点的第二版本,所述目标系统还配置为执行以下处理:

在检测到所述第一一致点标记符的基础上,暂停进一步应用所记录文件操作第一时间段,同时所述源数据的所述拷贝的当前状态对应于所述第一版本;

重新开始处理所述记录组,直到检测到所述第二一致点标记符为止;

在检测到所述第二一致点标记符的基础上,暂停进一步应用所记录文件操作第二时间段,同时所述源数据的所述拷贝的当前状态对应于所述第二版本;及

重新开始处理所述记录组。

## 检查点及一致性标记符

[0001] 发明背景

### 技术领域

[0002] 本申请总体上涉及数据存储系统,特别是使用数据存储系统中的一致点及相关联的一致数据拷贝的技术。

### 背景技术

[0003] 计算机系统可包括不同的资源,其由一个或多个主机处理器使用。计算机系统资源和主机处理器可由一个或多个通信连接互连。例如,这些资源可包括数据存储装置如包括在伊姆西公司制造的数据存储系统中的数据存储装置。这些数据存储系统连到一个或多个主机处理器并向每一主机处理器提供存储服务。来自一个或多个不同厂家的多个数据存储系统可被连接并可为计算机系统中的一个或多个主机处理器提供公共数据存储。

[0004] 结合数据存储系统可执行不同的任务。例如,可在数据存储系统中执行处理从而在目标系统创建和维护来自源系统的数据的镜像拷贝。当执行导致对源系统进行数据修改的文件操作时,所述修改可被记录且之后应用于数据的目标系统拷贝。在执行数据操作时,当数据处于一致状态时,可能希望使用目标系统上的数据拷贝。例如,当所有未完成数据库事务均被提交时,源或目标系统可表征为在该时间点处于一致状态,任何新引入的事务均被暂停或“排队”,且任何数据库缓冲器均在所选时间点转储清除。然而,很难建立和确定目标系统上的数据拷贝何时处于这样的一致状态。

[0005] 因而,需要有效的技术用于在目标系统上提供数据的一致拷贝并用于确定目标系统上的数据拷贝何时处于一致状态,同时使对源系统的数据及访问源系统上的数据的应用的负面影响最小。

### 附图说明

[0006] 本发明的特征和优点从下面结合附图对本发明示例性实施例进行的详细描述可明显看出,其中:

[0007] 图 1 为可使用在此描述的技术的计算机系统的实施例。

[0008] 图 2A 为数据存储系统的实施例。

[0009] 图 2B 所示为图 2A 的数据存储系统的一实施例中包括的指挥站和存储器之间的逻辑内部通信。

[0010] 图 3 为用以执行在此所述的技术以在目标上获得数据的一致拷贝的组成部分的例子。

[0011] 图 4 为可连同文件操作的实时复制一起使用以维护源数据的镜像拷贝的组成部分的例子。

[0012] 图 5 为可包括在连同将一致点标记符插入所捕获的文件操作流中一起使用的框架中的组成部分的例子。

[0013] 图 6 为所捕获的文件操作流的例子。

[0014] 图 7 为可包括在一致点标记符记录中的组成部分的例子。

[0015] 图 8 为确定源数据何时处于一致状态所执行的处理步骤的流程图。

[0016] 图 9 为源系统处理对应于所捕获文件操作的记录及转发给目标系统的一致点标记符所执行的处理步骤的流程图。

[0017] 图 10 为目标系统处理对应于所捕获文件操作的记录及从源系统接收的一致点标记符所执行的处理步骤的流程图。

[0018] 图 11 和 12 为其中可使用在此描述的技术和组成部分的其它结构的例子。及

[0019] 图 13 更详细地示出了可连同在此描述的技术一起使用的数据存储系统结构的例子。

### 具体实施方式

[0020] 现在参考图 1, 其示出了可连同实现在此所述的技术一起使用的计算机系统的实施例。计算机系统 10 包括通过通信媒质 18 连到主机系统 14a-14n 的数据存储系统 12。在计算机系统 10 的该实施例中, N 个主机 14a-14n 可在执行输入 / 输出 (I/O) 操作或数据请求时访问数据存储系统 12。通信媒质 18 可以是本领域技术人员已知的多种网络或其它类型通信连接中的任一或多个。通信媒质 18 可以是网络连接、总线、构架、和 / 或其它类型的数据链路, 如本领域已知的硬连线或其它连接。例如, 通信媒质 18 可以是因特网、内联网、网络或其它无线或其它硬连线连接, 主机系统 14a-14n 借助其可访问和与数据存储系统 12 通信, 并还可与计算机系统 10 中包括的其它组成部分通信。

[0021] 计算机系统 10 中包括的每一主机系统 14a-14n 和数据存储系统 12 均可通过根据通信媒质 18 的类型提供和支持的多种连接中的任何一种连接到通信媒质 18。主计算机系统 14a-14n 中包括的处理器可以是多种专有或可从市场上购得的单或多处理器系统中的任何一种, 如基于 Intel 的处理器、或能够根据每一特定实施例和应用支持流量的、可从市场上购得的其它类型处理器。

[0022] 应注意, 可包括在数据存储系统 12 中的硬件和软件的特定例子在此进行详细描述, 其可随每一特定实施例变化。每一主计算机 14a-14n 和数据存储系统可全部位于同一物理场所, 或者, 也可位于不同的物理位置。可用于在计算机系统 10 的主计算机系统和数据存储系统之间提供不同类型的连接的通信媒质的例子可使用多种不同的通信协议, 如 SCSI、光纤通道、iSCSI 等。主机、管理组件及数据存储系统借助其可连到通信媒质的部分或所有连接可通过其它通信设备, 如 Connectrix 或其它可能存在的交换设备, 如电话线、中继器、多路复用器、甚至人造卫星。

[0023] 每一主计算机系统可根据不同类型的任务执行不同类型的数据操作。在图 1 的实施例中, 主计算机 14a-14n 中的每一个可向数据存储系统 12 发出数据请求以执行数据操作。例如, 在主计算机 14a-14n 之一上运行的应用程序可执行导致一个或多个数据请求发给数据存储系统 12 的读或写操作。

[0024] 现在参考图 2A, 其示出了可包括在图 1 的计算机系统 10 中的数据存储系统 12 的实施例。图 2A 的数据存储系统 12 包括可由一个或多个不同厂家制造的一个或多个数据存储系统 20a-20n。每一数据存储系统 20a-20n 可以相互连接 (未示出)。另外, 数据存储系



统也可通过任一或多个通信连接 31 连到主机系统,所述通信连接可根据特定实施例中使用的不同协议而随每一特定实施例和设备变化。所使用的通信连接的类型可随某些系统参数和要求变化,如与根据主计算机系统发给数据存储系统 12 的 I/O 请求速率所需要的带宽和吞吐量有关的参数和要求。在下面详细描述的例子中,参考元件 20a 的详图。应注意,类似的详述也可应用于任一或多个其它元件,如 20n,但为了简单说明已省略这些描述。应注意,实施例可包括来自一个或多个厂家的数据存储系统。20a-20n 中的每一个可以是图 1 的计算机系统 10 的实施例中包括的、向主计算机系统提供存储服务的资源。应注意,数据存储系统 12 可独立运行,或也可包括为包括其它组成部分的存储区网络 (SAN) 的一部分。

[0025] 每一数据存储系统如 20a 可包括多个磁盘设备或卷,如由 n 排磁盘或卷 24a-24n 组成的排列 24。在该排列中,每一排磁盘或卷可连到负责与磁盘或卷 24 的一部分之间的操作的后端管理的磁盘适配器 (DA) 或指挥站。在系统 20a 中,单一 DA 如 23a 可负责一排磁盘或卷如排 24a 的管理。

[0026] 系统 20a 还可包括一个或多个主机适配器 (HA) 或指挥站 21a-21n。这些 HA 中的每一个可用于管理一个或多个主机系统和全局存储器之间的通信和数据操作。在实施例中,HA 可以是光纤通道适配器或有助于主机通信的其它适配器。

[0027] 一个或多个内部逻辑通信路径可能存在于 DA、RA、HA 和存储器 26 之间。例如,实施例可使用一个或多个内部总线和 / 或通信模块。例如,全局存储器部分 25b 可用于促进数据存储系统中的 DA、HA 和 RA 之间的数据传送及其它通信。在一实施例中,DA23a-23n 可使用高速缓存器执行数据操作如与其它磁盘适配器或指挥站或系统 20a 的其它组成部分通信,所述高速缓存器可包括在全局存储器 25b 中。另一部分 25a 是存储器的可用于与随每一实施例变化的其它目的地连接的部分。

[0028] 如该实施例中所述的特定数据存储系统或其特定装置如磁盘不应视为限制。其它类型的可从市场上购得的数据存储系统及控制对这些特定装置的访问的处理器和硬件也可包括在实施例中。

[0029] 在存储系统 20a 中还示出了 RA 或远程适配器 40。RA 可以是包括用于促进数据存储系统之间如两个相同或不同类型的数据存储系统之间的通信的处理器硬件。

[0030] 主机系统通过通道向存储系统提供数据和访问控制信息,存储系统也可通过通道向主机系统提供数据。主机系统不直接寻址存储系统的磁盘驱动器,而是对数据的访问可从主机系统看作多个逻辑装置或逻辑卷 (LV) 的地方提供给一个或多个主机系统。LV 可以也可不对应于实际的磁盘驱动器。例如,一个或多个 LV 可驻留在单一物理磁盘驱动器上。单一存储系统中的数据可由多个主机寻址从而使主机能共享其中驻留的数据。HA 可用于数据存储系统和主机系统之间的通信。RA 可用于促进两个数据存储系统之间的通信。DA 可用于帮助到驻留于其上的相关磁盘驱动器和 LV 的通信。

[0031] DA 在磁盘驱动器上执行 I/O 操作。在下面的描述中,在其它指挥站发起的与 I/O 操作有关的数据请求之后,驻留在 LV 上的数据可由 DA 寻址。

[0032] 现在参考图 2B,其示出了数据存储系统中包括的指挥站和存储器之间的逻辑内部通信的表示。图 2B 中包括连到存储器 26 的多个指挥站 37a-37n。指挥站 37a-37n 中的每一个代表可包括在数据存储系统中的 HA、RA 或 DA 之一。在于此公开的实施例,可以有高达 16 个指挥站连到存储器 26。其它实施例可使用更高或更低数量的指挥站。

[0033] 图 2B 的表示还包括可选通信模块 (CM) 38, 其在指挥站 37a-37n 之间提供备选通信路径。指挥站 37a-37n 中的每一个均可连到 CM38, 使得指挥站 37a-37n 中的任一个均可发送消息和 / 或数据给指挥站 37a-37n 中的任何其它指挥站, 而无需通过存储器 26。CM38 可使用常规 MUX/ 路由器技术实现, 其中指挥站 37a-37n 中的发送指挥站提供适当的地址以使消息和 / 或数据能被指挥站 37a-37n 中的预定接收指挥站接收。此外, 指挥站 37a-37n 中的发送指挥站能够将消息同时通信给指挥站 37a-37n 的所有其它指挥站。

[0034] 现在参考图 3, 其示出了可用于执行在此描述的技术的系统 10 的组成部分的例子。例子 100 包括如前结合其它附图所述的主机 14a 和数据存储系统 12。主机 14a 包括在其上运行的数据库应用程序 102。数据库应用程序 102 可读和写数据, 例如从可在主机 14a 本地的数据存储装置 104。与执行某些操作如备份操作一起, 来自主机本地的装置如装置 104 的数据可被复制到另一中心位置, 如数据存储系统 12 上的装置 106。之后, 数据存储系统 12 的装置 106 可被备份到其它数据存储装置上。连同操作装置 104 中包括的文件, 数据库应用 10 可执行基于文件的命令, 如读、写、截取文件操作。根据特定实施例的文件系统和结构, 也可在文件系统级对数据执行备份操作, 其中将被备份的指定组成部分是特定文件、目录等。作为备份操作的主题的数据可以是包括在装置 106 中的一部分数据。

[0035] 在获得和维护来自源装置如 104 的源数据的拷贝时, 可执行初始同步操作。同步操作可开始将源数据的文件复制到目标, 如包括在装置 106 中的目标数据。在同步之后, 源和目标数据拷贝完全一样。之后, 随后对源数据的更新和修改可通过对源数据执行的镜像文件操作而被异步地复制到目标数据上。换言之, 如果特定文件操作导致源数据的修改或变化, 该文件操作可被记录并继而应用于目标数据。当基于目标数据执行备份操作时, 目标数据可表征为源数据在时间点的一致拷贝。如在此所描述的, 源数据的拷贝在时间点处于一致状态, 例如, 当所有未完成数据库事务均被提交时, 任何新进入的事务暂停或“排队”, 及任何数据库缓冲器或其它缓存的数据相对于所选时间点转储清除。

[0036] 下面将描述的是用于从目标如数据存储系统 12 的装置 106 上的装置 104 获得源数据的一致拷贝。目标数据在其为源数据的一致拷贝的时间点时可用于执行其它操作和任务如备份操作。

[0037] 尽管例子 100 仅示出了单一主机 14a, 应注意, 系统 10 中包括的任一其它主机也可连同在此所述的技术使用。主机 14a、特定应用程序如数据库应用程序 102、及执行备份操作时存储装置 106 上包括的目标数据的使用均被选择用于说明在此描述的技术。除数据库应用程序以外, 其它应用程序也可用于执行源数据装置复制到目标数据存储装置的操作。另外, 除备份操作以外, 目标数据存储装置还可连同多种不同操作中的任一操作使用。

[0038] 图 3 中包括的组成部分仅代表实施例中实际包括的组成部分的一部分。已包括在图 3 中的特定组成部分仅用于说明和示例的目的。本文别处描述了多种结构。例如, 在实施在此描述的技术时, 一实施例可使用主机 - 主机通信、数据存储系统 - 主机、或数据存储系统 - 数据存储系统通信。

[0039] 在下面的段落中, 将首先描述用于将源数据复制在目标上的技术。随后将描述复制技术的另外的组成以帮助确定目标数据何时可被表征为源数据的一致拷贝。

[0040] 现在将描述的是用于将源数据如源装置 104 的数据复制到目标装置如装置 106 的技术。在此所述的复制技术可根据文件系统级命令在文件系统级执行。根据将结合图 4

描述的一种技术,多个文件系统级命令可被捕获和保存为实时应用于源装置 104 的文件操作。当这些文件操作被捕获和记录在源系统如主机 14a 上的队列或存储位置中时,捕获的操作可转发给目标系统或服务器以应用于目标装置 106 中包括的目标数据。

[0041] 应注意,尽管在此详细描述的是使用异步复制的特定复制技术,在此描述的技术可连同本领域一般技术人员已知的其它复制技术一起使用,如多种同步复制模式。

[0042] 现在参考图 4,其示出了可用于实时复制文件操作以在目标装置 106 上保持源装置 104 中包括的源数据的镜象拷贝的组成部分的例子 200。例子 200 中包括的是源系统 230 和目标系统 240。例如,源系统 230 可以是如前结合图 3 所示的主机 14a。例如,目标系统 240 可以是数据存储系统 12 中包括的或连接到数据存储系统 12 的服务器。应用程序 102 可执行文件操作如文件写操作或导致源数据修改的其它操作。在应用程序 102 运行期间,由应用程序 102 发出的这些文件操作(步骤 1)可使用一个或多个组成部分捕获。部分组成部分在内核空间 202 中运行。在内核空间 202 中运行的组成部分可包括 I/O 管理器 204、镜象驱动器 206 和文件系统 208。应注意,实施例可包括不同于结合该例子所示的其它组成部分。例如,实施例可在文件系统 208 下面包括一个或多个与磁盘或其它装置 104 接口的其它驱动器。镜象驱动器 206 可表征为过滤驱动器,其在备份调用链过程中捕获感兴趣的文件操作(步骤 2)并将感兴趣的文件操作记录在内核高速缓存 210 中(步骤 3)。在内核高速缓存 210 存满或溢出时,镜象驱动器 206 可将捕获的数据操作记录在溢出位置 212(步骤 4)。内核高速缓存和/或溢出位置 212 中包括的捕获数据操作的消费者或读者为复制服务 214,用户模式的过程。复制服务 214 从内核高速缓存 210(步骤 5)和溢出位置 212(步骤 6)读捕获的数据操作。应注意,镜象驱动器 206 持续写到溢出位置 212,直到复制服务处理足够量的来自内核高速缓存的捕获文件操作时为止。当这发生时,镜象驱动器可重新开始向内核高速缓存 210 写。当镜象驱动器在内核高速缓存 210 或溢出位置 212 排队内核操作时,复制服务 214 从内核高速缓存或溢出位置 212 读文件操作。复制服务 214 将从 210 和/或 212 读的文件操作传给适当的目标系统如服务器 240(步骤 7)。例如,位于目标服务器上的代码可包括复制服务 216,其处理所接收的、传自源系统 230 的服务 214 的文件操作(步骤 8)。之后,目标上的复制服务 216 可将所接收的文件操作应用于目标存储装置 106 中包括的目标数据的文件(步骤 9)并向源系统发回确认(步骤 10)。应注意,在复制服务 216 被阻止执行或应用文件操作时,文件操作可被记录在封锁数据容器 220 中(步骤 11)。应注意,例如,如果磁盘存满、文件正在使用中等,将文件操作应用于文件可被阻止。一旦文件被标记为有被阻止的操作,则对该特定文件的所有随后的文件操作均被保存在封锁数据容器 220 中,直到所述文件与非封锁文件操作状态相关联为止。复制服务 216 可定期尝试重新应用封锁数据容器 220 中包括的文件操作(步骤 12)。

[0043] 如刚才所述的,镜象可表征为应用于源数据的文件操作的实时镜象。例如,可使用镜象驱动器 206 和服务 214 帮助镜象。镜象可由应用程序如数据库应用程序 102 的文件操作触发。例如,应用程序也可以是任何其它用户模式过程,如 Microsoft Word。文件操作也可由例如通过映射的驱动器连到所述源的远程客户计算机引起。刚才结合图 4 的例子 200 所述的操作可表征为具有包括三个不同的阶段:镜象阶段,其中文件操作在内核中捕获并排队发送给目标。第二阶段可表征为转发阶段,其中文件操作从所述源转发给目标。第三阶段可表征为更新阶段,其中文件操作应用于目标数据。应注意,源系统 230 和目标系统 240

可位于相同或不同计算机系统内。换言之,230 可包括在第一计算机系统中,240 可包括在第二不同的计算机系统中。或者,组件 230 和 240 可物理上紧邻定位或位于地理上不同的位置。例子 200 的这些及其它变化很容易由本领域一般技术人员意识到。

[0044] 如每一请求的返回状态所指示的,当文件操作已被完成时,镜象驱动器 206 可捕获和记录文件操作,所述返回状态由 202 中包括的组件的返回调用链的向上箭头指示。在完成文件操作的基础上,镜象驱动器 206 确定是否应捕获特定文件操作。在该例子中,如果操作导致文件的修改如文件创建、文件写操作、文件截断操作等,文件操作可以是感兴趣的操作。应注意,其它条件可与确定特定文件操作是否为感兴趣的操作相关联,例如,文件操作的成功状态。如果镜象驱动器 206 根据实施例指定的一个或多个条件确定特定文件操作是感兴趣的操作,则数据操作可被捕获并保存在内核高速缓存 210 中。应注意,在该例子中,内核高速缓存可表征为充作所有镜象的文件操作的队列的共享存储器的区域。内核高速缓存可以是固定大小。如果内核高速缓存存满,则文件操作可记录在溢出区域 212。在该例子中,源上的复制服务 214 和目标上的复制服务 216 可表征为用户模式应用程序,且不要要求在特权模式如内核模式 202 下运行。

[0045] 关于记录感兴趣的文件操作,镜象驱动器 206 产生将应用于目标装置 106 中包括的目标数据的文件操作流。前述内容描述了异步数据移动,其中在另外的对源数据的数据操作被允许继续之前文件系统变化不必提交给目标。而是,源系统被允许在数据复制处理的同时正常运行。前述例子中使用的异步复制使能立即捕获源数据的任何变化,并继而缓存在源系统本地。之后,当网络和其它资源允许时所述变化转发给目标系统或服务器。

[0046] 现在将描述的是可用于在前述所记录的文件操作流中插入另外的一致点标记符的技术。一致点标记符指明复制在目标数据中的源数据的拷贝处于一致状态的各个点。换言之,一致点标记符可插入文件操作流中以用作标识点,使得如果到该特定标记符的数据操作应用于目标装置,目标装置可表征为源数据拷贝的一致时间点。如将结合图 4 的组成部分在此描述的一实施例中,复制服务 214 的修改版本也可用于将一致点标记符插入将应用于数据的目标拷贝的文件操作流或队列中。

[0047] 应注意,RepliStor 是由伊姆西公司制造的可从市场上购得的产品,其连同如在此所述的数据复制和镜象一起使用。可从市场上购得的 RepliStor 产品包括复制服务 214 和 216 及按结合图 4 的实施例 200 所述进行运行的镜象驱动器。现在将描述的是其中如图 4 中所示的那些组成部分的修改版本可用于在目标上获得源数据的一致时间点拷贝。包括在目标中的源数据拷贝可用于执行本领域一般技术人员将意识到的其它操作。

[0048] 现在参考图 5,示出了可包括在框架中的组成部分的例子 300,其用于将一致点标记符插入在源系统上捕获的文件操作流中。例子 300 中包括写入者 302、提供者 304、及请求者 306。例子 300 中还另外包括组成部分 VSS310。在该例子中,VSS310 是微软公司的卷映射拷贝服务 (VSS)。VSS 是可通过商业途径获得的产品,其可表征为提供具有应用程序编程接口 (API) 的框架,该框架可连同创建数据的一致时间点拷贝一起使用。如更常使用的,VSS 允许系统管理员创建一卷文件的快照或映射拷贝,其可共享为网络资源。VSS 与不同的组成部分通信,这些组成部分可相对于特定数据集表征为不同的类,如写入者 302、提供者 304 和请求者 306,从而可产生数据的时间点拷贝。写入者 302 可表征为第一组组成部分,其将如图 4 的源系统 230 中包括的源数据写到数据拷贝。例如,写入者可以是结合图 3 和 4

所示的数据库应用程序 102、微软交换服务器、或其它应用程序。通常，写入者 302 可执行更新或导致数据集修改的其它类型的操作。提供者 304 可表征为提供数据集的软件和 / 或硬件组成部分，如基础存储系统。例如，数据集的提供者可包括数据存储系统如伊姆西公司的 Symmetrix 数据存储系统或其它数据存储系统。提供者 304 通常提供可由写入者 302 访问的资源或数据集。请求者 306 可表征为提出获取时间点拷贝的请求的那些组成部分。一般地，例如，请求者可包括用于请求对提供者 304 之一提供的数据执行备份复制的备份软件应用程序。在运行中，请求者 306 可向 VSS 组成部分 310 发出请求以获得数据集的时间点拷贝。之后，VSS310 与该数据集的写入者 302 通信以暂停任何新事务、完成任何当前事务、及将任何缓存的数据转储到磁盘上。一旦写入者 302 已对所请求的数据集完成这组操作，VSS 组成部分 310 与适当的提供者 304 通信以开始所请求数据集的映射拷贝过程。一旦映射拷贝已被创建，则备份软件（如该例子中的请求者）可从该映射拷贝将数据复制到磁带上，而不涉及特定数据集的写入者。因而，VSS 用作为获得数据的时间点拷贝而促进写入者 302、提供者 304 和请求者 306 之间的通信的框架。

[0049] 前述框架可用于获得源数据的一致时间点拷贝。与在此所述的技术有关，基于主机的用户模式进程如图 4 的例子 200 的复制服务 214，可注册为数据集如源数据的提供者 304 之一。因而，使用图 4 的框架，复制服务 214 可注册为提供者并在所请求的数据集（如源数据）处于一致状态时可由 VSS310 通知。当需要产生一致点标记符时，当请求者 306 之一使用应用程序编程接口（API）时进程可发起请求。应注意，充作用于一致点标记符产生的请求者和提供者的进程或组成部分可以是相同的进程或组成部分，也可以是不同的进程或组成部分。API 可根据特定实施例如 VSS 和其它组成部分进行指定。

[0050] 作为例子，调度程序可以是在源系统中运行并作为请求者 306 请求源数据的进程。复制服务 214 可被注册为源数据的提供者 304。VSS310 与数据库应用程序和源数据的任何其它写入者 302 通信以完成现有的事务、不开始任何新的事务、及转储任何缓存的源数据，使得源数据处于一致状态。之后，VSS310 通知源数据的注册提供者 304，其导致源数据当前处于一致状态的通知将发送给服务 214。对源数据的任何进一步的操作均被暂停预定的一小段时间。之后，服务 214 可将一致点标记符产生为源数据的所捕获文件操作流中的下一组成部分。然后，源数据的各个写入者 302 可重新开始与源数据有关的正常操作。

[0051] 现在参考图 6，其示出了可保存在内核高速缓存 210 中的所捕获文件操作流的实施例。例子 400 中包括记录列表。每一记录或对应于已被捕获的特定文件操作，如记录 404，或对应于一致点标记符，如记录 402。每一捕获的文件操作记录 404 中包括的数据根据特定文件操作可包括特定文件名称、位置、数据值等。一致点标记符记录 402 可以是写入文件操作流的特殊记录。可包括在一致点标记符记录中的数据元的实施例将在下面详细描述。

[0052] 现在参考图 7，其示出了可包括在一致点标记符记录 402 中的组成部分的例子。在图 7 的例子 402 中，一致点标记符记录可包括暂停转发标志 404、暂停更新标志 406、源脚本的标识符 408、目标脚本的标识符 410、及其它数据 412。在一实施例中，字段 404 和 408 可用于在源系统上处理一致点标记符，字段 406 和 410 可用于在目标系统上处理一致点标记符。

[0053] 源系统上的服务 214 可根据内核高速缓存中的特定记录检查和执行处理。当记录对应于一致点标记符时，可使用字段 404 和 408 的值执行特殊处理。当一致点标记符记录

402 被检测时, 字段 408 可指明在源系统上运行的脚本。该脚本可包括根据暂停转发标志 404 的值检查和执行处理的处理步骤。暂停转发标志 404 可以是用于在源系统上处理所捕获的文件操作流的二进位值。在一实施例中, 如果暂停转发标志 404 被置位 (如等于 1), 则标志 404 指示对应于所捕获文件操作的记录不应从源发送给目标系统, 从而导致文件操作的待办事项将保存在源系统上而不是目标系统上。如果暂停转发标志 404 为关 (如等于 0), 其表明对应于所捕获文件操作的记录应从源系统转发给目标系统, 使得任何待办事项或数据的内建将在目标系统上发生, 而不是源系统。暂停转发标志 404 可用于字段 408 指示的源脚本的处理步骤。应注意, 源脚本字段 408 是可选字段, 因为对于其中没有脚本运行因而暂停转发标志值可被忽略情况下的记录 402, 未指定值。在这种情况下, 只要标志 402 的值为 0, 文件操作即可有效地转发给目标系统。在一实施例中, 标志 404 可导致记录在源系统上排队预定时间段。在该预定时间段之后, 复制服务 214 可重新开始处理和转发记录给目标系统。在另一实施例中, 脚本可包括导致记录转发给目标系统的处理步骤以暂停直到发生特定事件为止。当该特定事件已出现时, 可发信号通知服务 214 重新开始处理记录。

[0054] 对于 400 的从源系统转发的所接收记录, 处理也可由服务 216 在目标系统上执行。当检测到对应于一致点标记符的记录时, 可使用字段 406 和 410 的值执行特殊处理。暂停更新标志 406 可以是二进位值, 当捕获文件操作的应用应暂停或不应用于目标系统上的目标数据拷贝时其值为 1。否则, 当暂停更新标志值为 0 时其表明如 400 的所传输记录指明的所捕获文件操作应用于数据的目标拷贝。标志 406 可用于字段 410 指示的目标系统脚本执行的处理。字段 410 可选地确定将在目标系统上运行的脚本。在目标系统处理时标志 406 的值可被忽视的情况下, 特定一致点标记符的记录也可能未被指定。字段 408 和 410 可包括标识符如包含将在相应源或目标系统上运行的脚本的目录位置和文件名称。

[0055] 应注意, 实施例可包括另外的可用于处理源和 / 或目标系统上的记录的字段。

[0056] 在一实施例中, 源系统的复制服务可创建记录 402 的每一例图因而初始化记录 402 的字段。作为例子, 如果备份将使用目标数据的拷贝每天执行, 源系统上的调度程序进程可发起请求以按每 24 小时间隔产生一致点标记符。相应的一致点标记符记录的特定值可指该操作的脚本。除了为源数据的备份操作一致拷贝产生标记符以外, 另外的产生标记符的请求也可按预定时间间隔或基于与其它操作有关的某些事件的发生执行。这样, 为一致点标记符创建记录的服务因而可用相应操作的特定脚本初始化字段, 如脚本字段 408 和 410。类似地, 产生记录 402 的服务 214 也可根据将要执行的特定操作初始化每一记录例图的其它字段。

[0057] 现在参考图 8, 其示出了在确定源数据何时处于一致状态的实施例中执行的处理步骤的流程图 500。500 的步骤概括了上面结合图 5 所述的处理。在步骤 502, 提出获取源数据的一致拷贝的请求。例如, 这样的请求可连同调度程序任务执行定期备份操作等一起提出。在步骤 504, VSS 组成部分与写入源数据的写入者通信以暂停新的事务、完成当前事务、及将所有缓存的数据转储到磁盘上。写入者可在该状态暂停预定时间段。之后, VSS 与数据处于一致状态的源数据的注册提供者通信。在该例子中, 提供者可以是例如图 4 的复制服务 214。在步骤 508, 已注册提供者将对应于一致点标记符的记录写入将被处理的所记录操作的队列从而转发给目标系统。

[0058] 现在参考图 9, 其示出了在处理来自将被转发给目标系统的记录的输出队列的记

录时由源系统执行的处理步骤的流程图。如本文别处结合图 4 的组成部分所述,这些记录可保存在内核高速缓存和 / 或溢出位置。在步骤 552,读输出队列中的下一记录。在步骤 554,确定该记录是否对应于一致点标记符。如果否,控制进行到步骤 556 以执行其它处理。应注意,步骤 556 的其它处理可包括如本文别处所述的将记录转发给目标系统的步骤。另外,实施例还可根据不同于在此所述的文件操作及每一记录中包括的其它字段执行其它处理步骤。之后,控制进行到步骤 552 读下一记录。如果步骤 554 结果为是,控制进行到步骤 558,读当前记录的字段。在步骤 560 确定是否已指定源脚本。如果否,控制进行到步骤 552 以继续处理下一记录。如果步骤 560 结果为是,即表明源脚本已被指定,则在步骤 561 获取源脚本并运行该脚本。在源脚本的该例子中包括使用暂停转发标志及基于所述标志的值有条件地执行处理步骤的语句。换言之,在该例子中,源脚本包括用于执行步骤 562、564 和 566 的语句。在步骤 562,确定暂停转发标志是否被置位。如果是,控制进行到步骤 564,源系统暂停将输出队列中的记录发送给目标系统。如果步骤 562 的结果为否,则不暂停从输出队列转发进一步的记录给目标系统。控制进行到步骤 566,根据源脚本可执行其它处理步骤。在该例子中,源脚本可规定:如果先前被暂停,则从输出队列转发记录可被重新开始。还应注意,对前述的变化可在脚本中说明。控制进行到步骤 552,处理输出队列中的下一记录。流程图 550 的步骤可由源系统的复制服务 214 及源系统脚本执行。

[0059] 现在参考图 10,其示出了在处理从源系统接收的记录时由目标系统执行的处理步骤的流程图。在步骤 602,读输入队列中的下一记录。在步骤 604,确定该记录是否对应于一致点标记符。如果否,控制进行到步骤 610 以确定这是否是将要应用于已经处于封锁状态的文件的操作。如果否,控制进行到步骤 606 以将所述操作应用于目标数据。在步骤 606 之后,控制进行到步骤 602 以处理输入队列中的下一记录。否则,如果文件操作将应用于处于封锁状态的文件,控制进行到步骤 612 以将记录保存在已保存的块文件中,如图 4 的组成部分 220。如本说明书别处所述的,在特定文件不再处于封锁状态的一些随后的时间点,封锁的文件操作可应用于目标数据。在步骤 612 之后,控制进行到步骤 602 以处理输入队列中的下一记录。如果步骤 604 确定当前记录对应于一致点标记符,则控制进行到步骤 608 以读当前记录的字段。在步骤 614,确定是否已指定目标脚本。如果否,控制进行到步骤 602。否则,控制进行到步骤 616 以读所述脚本并开始运行所述脚本。在目标脚本的该例子中包括使用暂停应用更新标志及基于所述标志的值有条件地执行处理步骤的语句。换言之,在该例子中,目标脚本包括用于执行步骤 618、620 和 624 的语句。在步骤 618,确定暂停应用更新标志是否被置位。如果否,控制进行到步骤 624。否则,控制进行到步骤 620 以暂停处理输入队列中的任何进一步的记录,因而暂停应用所述记录指明的任何进一步的文件操作。在步骤 624,另外的处理可根据脚本执行,如果先前在步骤 620 暂停,脚本还可导致输入队列中的记录的处理重新开始。例如,在步骤 624 根据脚本的处理可包括执行目标数据的备份或执行目标数据的拆分,同时目标系统当前正暂停对目标数据的任何进一步的更新。随后,在备份或其它操作执行之后,目标系统可重新开始输入队列中的记录的处理。还应注意,对前述的变化可在脚本中说明。控制进行到步骤 602 以处理下一记录。流程图 600 的步骤可由目标系统的复制服务 216 和目标系统脚本执行。

[0060] 前面描述了使用 VSS 的实施例。然而,实施例可在协调访问特定数据集的组成部分之间的通信时使用其它技术,而不使用 VSS。例如,实施例可使用脚本和 / 或其它编程技

术作为 VSS 的替代。

[0061] 应注意,图 4 的组成部分,当修改为包括获取和利用一致点标记符的功能时,可被包括在多种不同结构的任一结构中。例如,第一结构可在主机上包括源 230 的组成部分。源的复制服务 214 可驻留在主机中并在其中运行。数据源、装置 104 可以在主机本地,也可连接到主机。目标 240 的一种结构可包括复制服务 216 驻留于其上并在其上运行的目标主机。目标装置 106 可以在目标主机本地或连接到目标主机。前述可表征为基于主机-主机的通信系统,连同在此所述的技术使用的复制服务、镜象驱动器和其它组成部分驻留在主机系统上。连同在此所述的技术使用的组成部分的一部分也可驻留在数据存储系统内并在其中运行。例如,目标系统 240 的实施例可包括目标数据存储系统,如 Symmetrix 或 Clarion 数据存储系统,复制服务 216 驻留于其上并在其上运行。源系统 230 可包括使用适当通信连接与目标数据存储系统通信的主机。在另一实施例中,源系统 230 可包括源主机,应用程序 102 及镜象驱动器 206 在其上运行。源系统 230 还可包括源数据存储系统,复制服务 214 基于其运行并与源主机通信以从 210 和 212 获取记录的文件操作和标记符。230 的源数据存储系统可使用在此所述的技术与目标 240 通信(按适于目标 240 的方式使用基于主机的或基于数据存储的通信)。现在将说明的是可连同前述技术使用的不同结构的一些例子。

[0062] 现在参考图 11,示出了其中可使用在此所述的组成部分和技术的结构的例子 1000。例子 1000 包括连接到两个源主机 1014a-b 的源数据存储系统 1012。两个主机 1014a-b 可使用通信媒质 1020 与目标主机 1014c-d 通信。两个源主机 1014a-b 可相互通信及在通信媒质 1018a 上与源数据存储系统 1012 通信。类似地,目标主机 1014c-d 可相互通信及在通信媒质 1018b 上与目标数据存储系统 1012a 通信。前述可表征为使用在此所述的技术的基于主机-主机的通信,其中源 230 和目标 240 的用于执行在此所述的处理步骤的组成部分驻留在相应的源和目标主机系统上并在其上运行。前述例子 1000 还示出了实施例中也可能存在的、数据存储系统的分布特性。源和目标数据存储系统可以是例如结合图 2 所述的存储系统。

[0063] 现在参考图 12,示出了其中可使用在此所述的组成部分和技术的另一结构的例子 1100。在例子 1100 中,源主机 1102 与目标数据存储系统 1012a 通信。目标数据存储系统 1012a 可包括一个或多个数据存储系统,如 Symmetrix 或 Clarion 数据存储系统,源主机 1102 在 1020 上与其通信。源主机 1102 可从源数据存储系统 1012 获得源数据。源数据存储系统 1012 可以是例如结合图 2 所述的存储系统。

[0064] 现在参考图 13,示出了可包括在源和/或目标数据存储系统中的组成部分的另一结构的例子 1200。具体地,例子 1200 可以是用于图 12 的目标数据存储系统的结构。例子 1200 包括组成部分 1050,其可表征为用于帮助通过公共点的一个或多个数据存储系统 20a-20n 之间的通信的前端处理器或计算机。例如,组成部分 1050 可以是可包括在 SAN(存储区网络)中的服务器系统或其它用于促进与组成部分 20a-20n 的通信的处理器。结合图 4 的组成部分可使用一致标记符根据在此所述的技术进行修改,在例子 1200 是目标系统的实施例中,复制服务 216 可驻留在组成部分 1050 上并于其上运行。

[0065] 图 11-13 为选择来图示可包括在使用在此所述的技术的实施例中的不同结构的一些例子。在使用在此所述的技术进行基于主机-主机、数据存储-主机及主机-数据存储的通信时,本领域一般技术人员将可意识到其它变化。



[0066] 前面描述了用于将一致点标记符从源系统传给目标系统的技术,其中数据自动从源系统的源数据复制到目标系统的目标数据。源数据可以是基于文件系统的组成部分如文件、目录等。如上所述的一实施例可使用微软的 VSS(卷映射拷贝服务)框架以指定时间间隔或在一致点获取源数据的一致“快照”,并将相关联的一致点标记符写入将应用于目标中的拷贝的所捕获文件操作队列中。不同的组成部分可使用所定义的应用程序编程接口(API)通信。所捕获的文件操作队列及一致点标记符可持续发送给目标系统,在那里所捕获的文件操作被读并应用于目标数据拷贝,直到达到一致点标记符为止。在该点,目标系统知道目标数据相对于源数据的快照处于一致状态。在将来自队列的任何更多的所捕获文件操作应用于目标数据之前,如执行多个任务的任一任务如备份操作、一致拆分操作等时在目标数据中所反射的,目标系统可使用源数据的一致拷贝。同样,在此所述的是在源和目标系统之间传输时可用于控制队列中的所捕获文件操作流的特征。脚本可在源系统上运行以控制源系统的输出队列中包括的文件操作和一致点标记符发送给目标系统的流速(如暂停转发从而不发送对目标系统的任何进一步的更新/写)。脚本也可在目标系统上运行以暂停将任何进一步的所捕获文件操作应用于目标数据。这样,代表源数据的时间点一致拷贝的目标数据可被冻结在该一致状态一定时间,使得可执行其它操作(如备份操作、拆分操作等)。备选实施例可使用脚本或其它编程技术控制所有已注册数据写入者(如数据库应用程序)、提供者(如数据存储系统如 Symmetrix 数据存储系统)和请求者之间的协调,从而获取数据源的一致快照及将一致点标记符插入将作为数据复制的一部分应用于目标数据的文件操作记录队列中,而不是使用 VSS 框架。如在此根据不同结构所述的及本领域一般技术人员将意识到的其它结构,不同的组成部分可驻留在主机和/或数据存储系统上并在其上运行。

[0067] 在本发明已结合详细图示和描述的优选实施例公开的同时,其修改和改进对本领域技术人员而言将显而易见。因此,本发明的精神和范围应仅由下述权利要求限制。

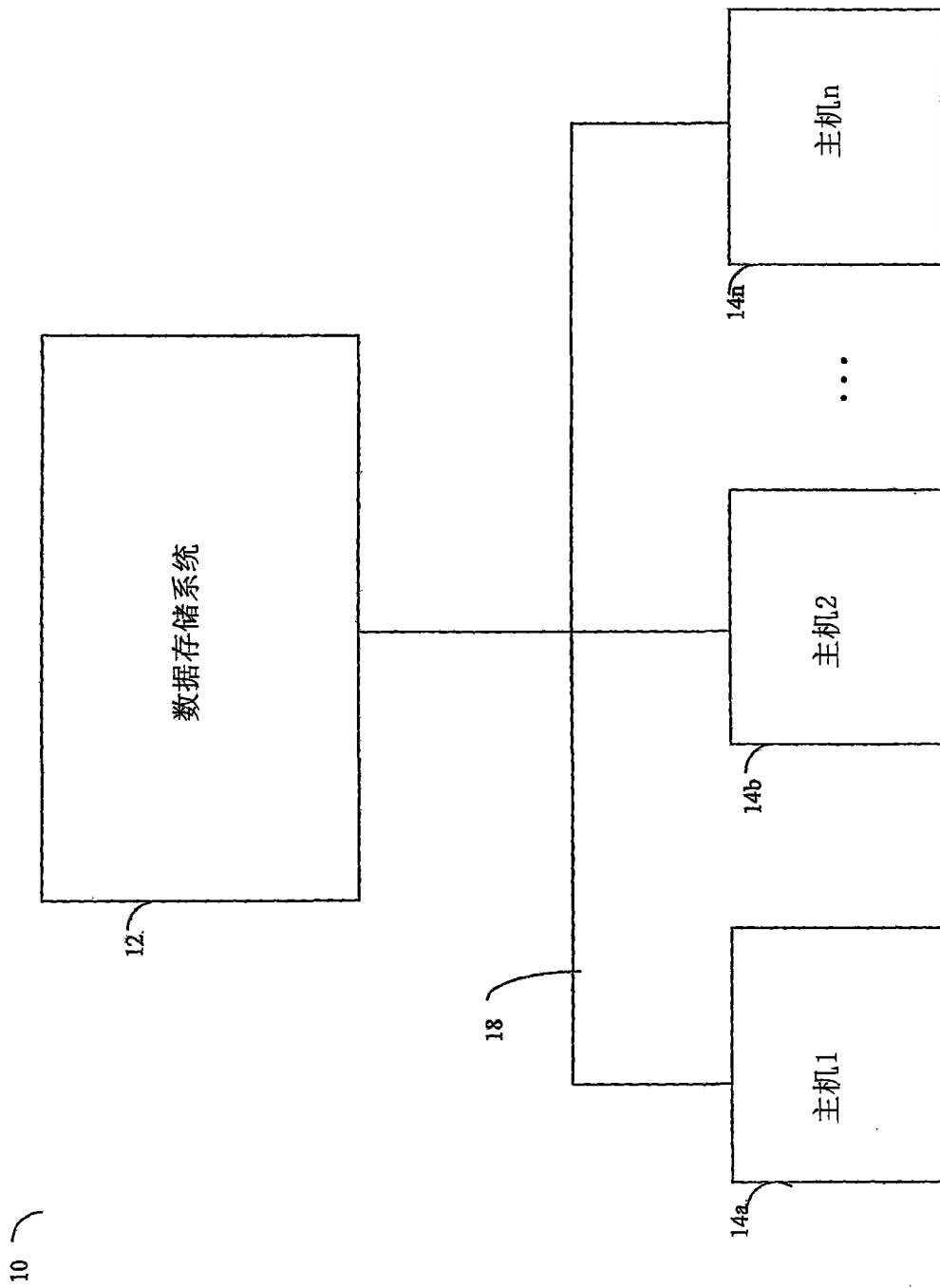


图1

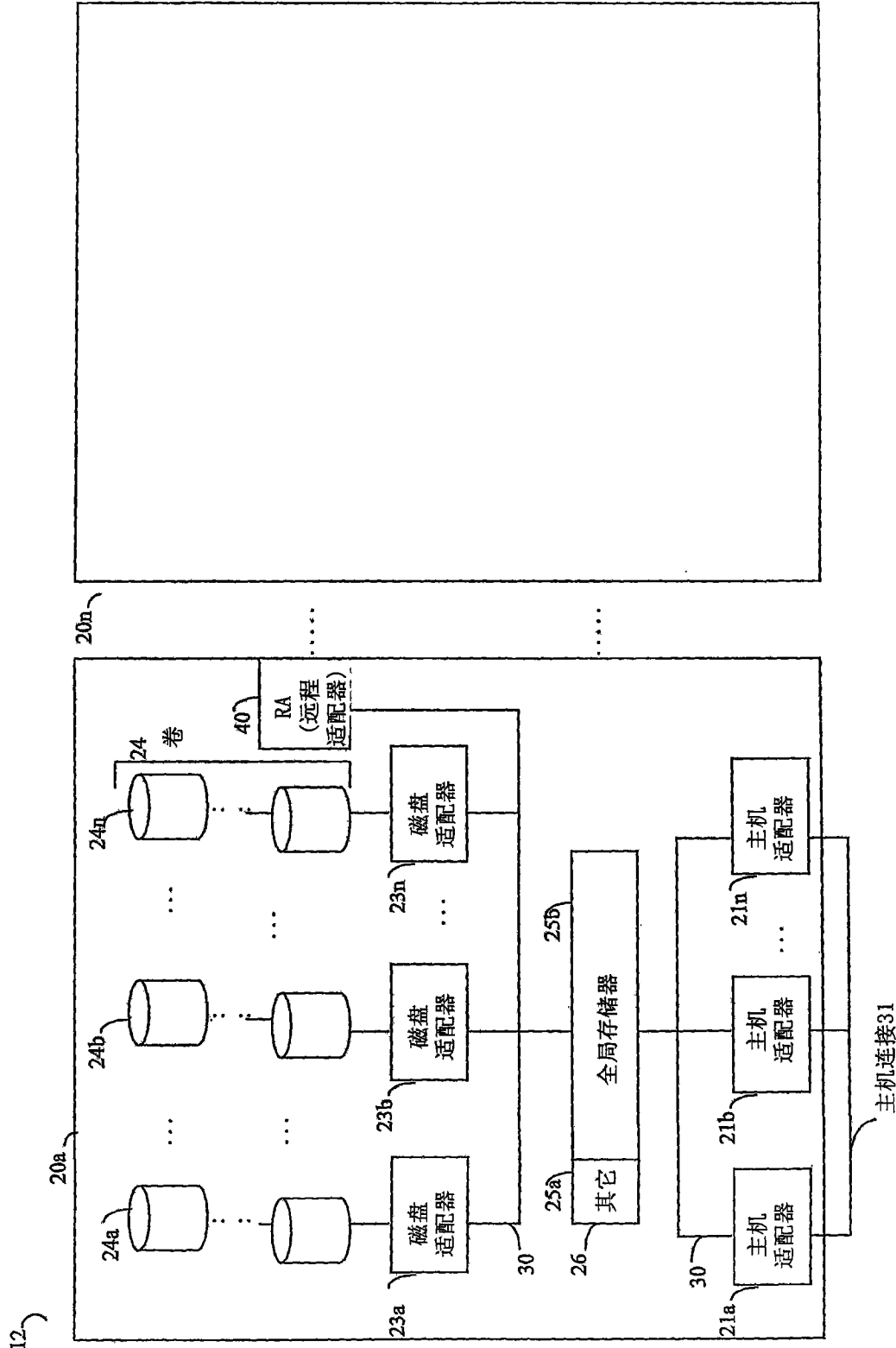


图2A

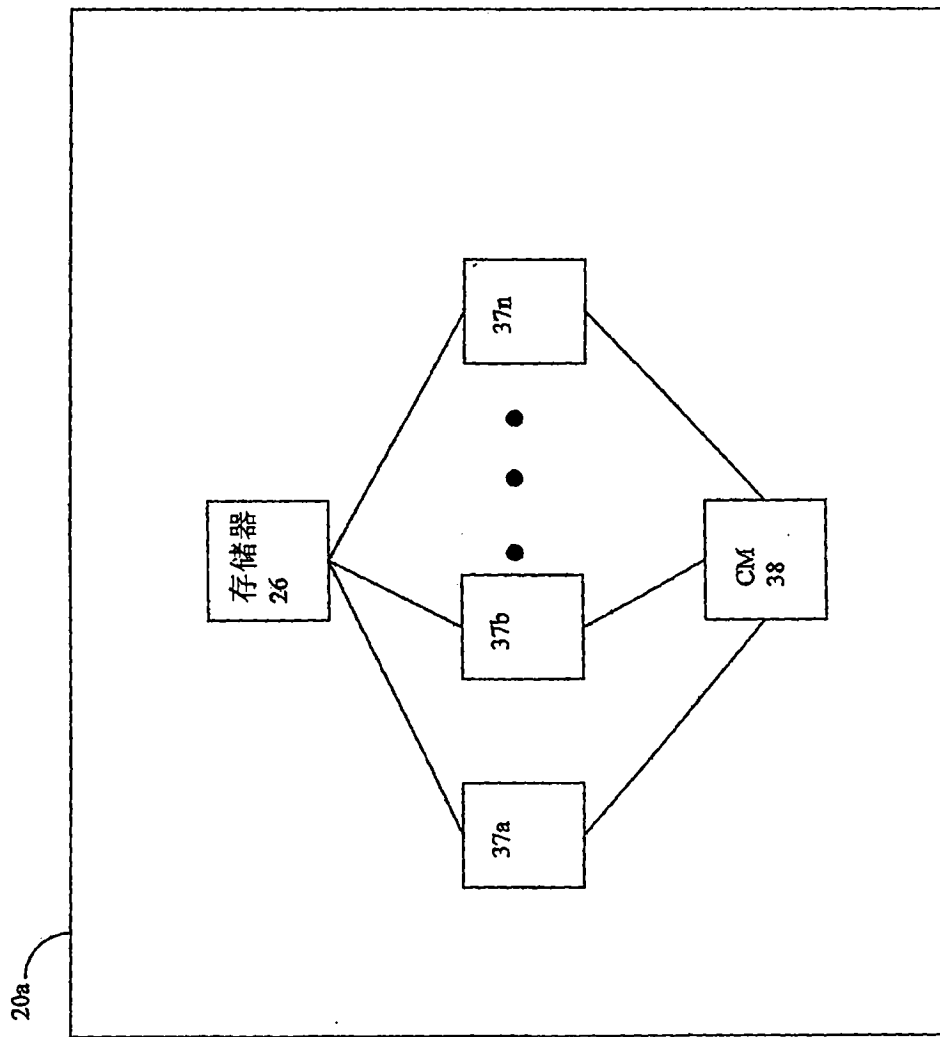


图2B

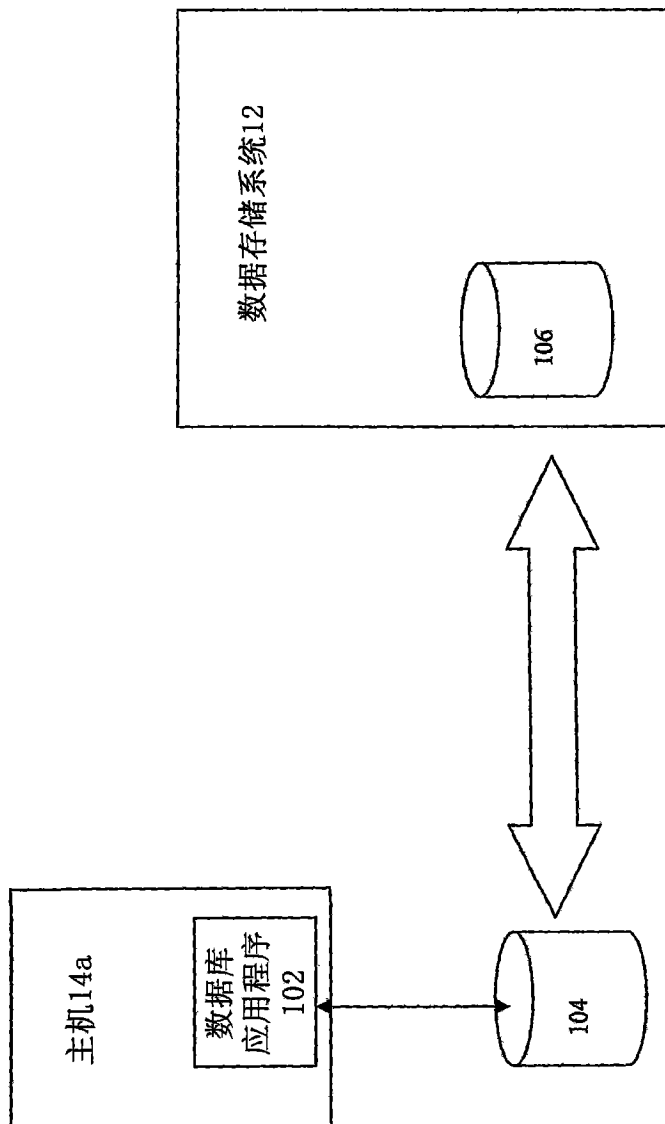


图3

200 ~

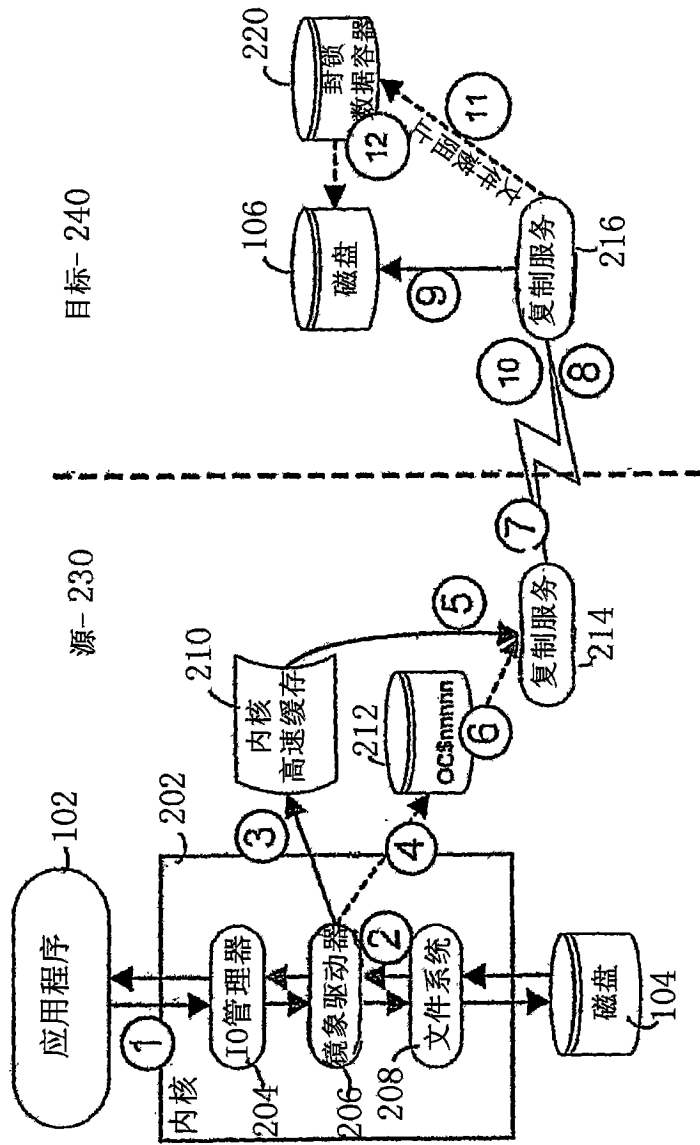


图4

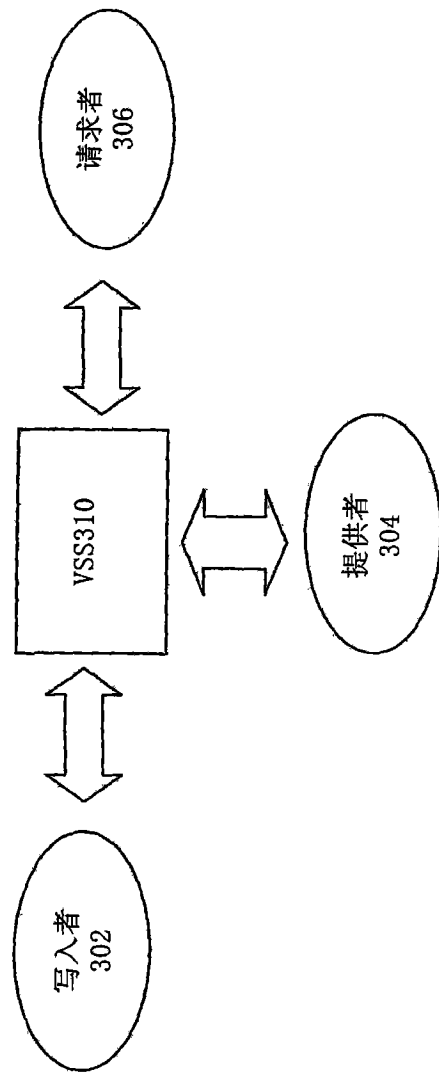


图5

400

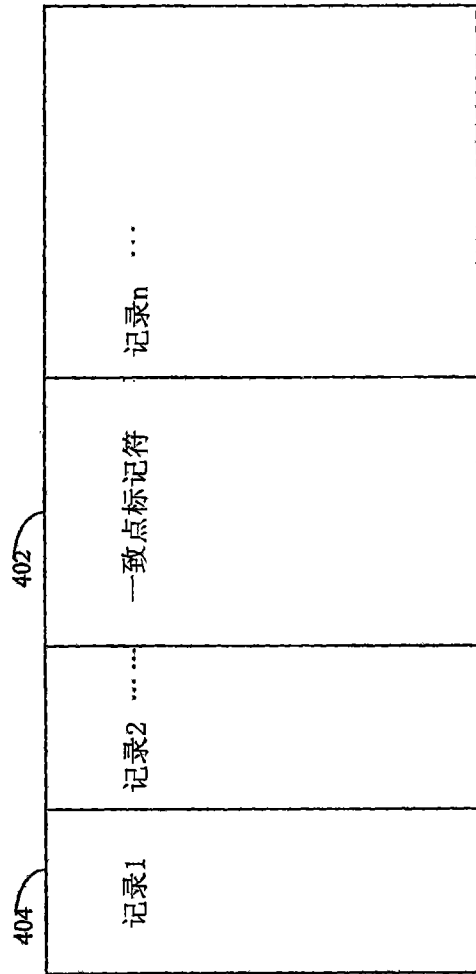


图6



402

暂停转发标志404
暂停更新标志406
用于源脚本的标识符408
用于目标脚本的标识符410
其它数据412

图7

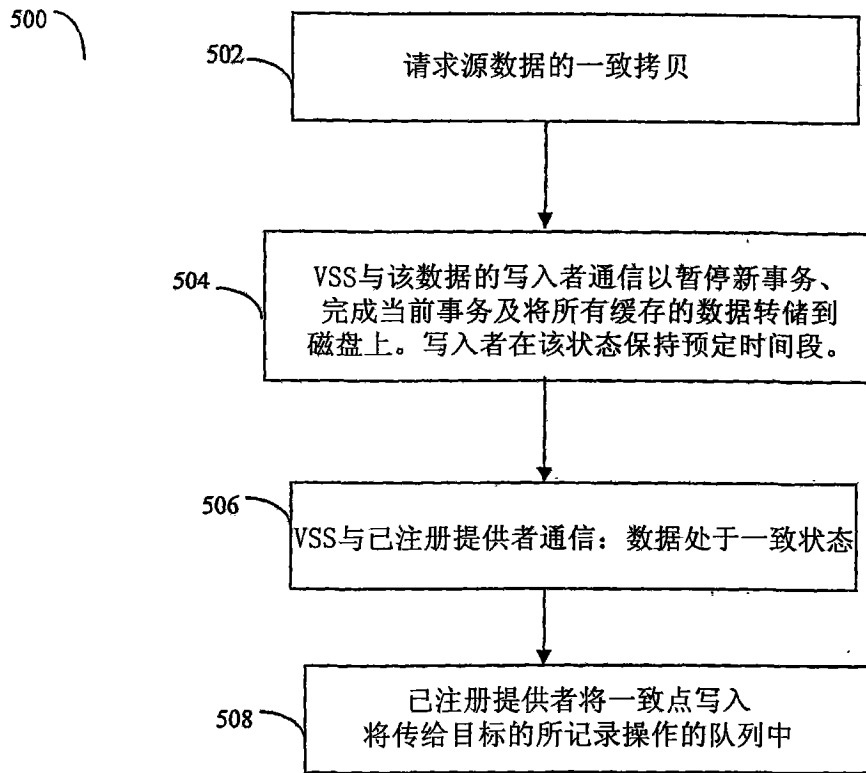


图8

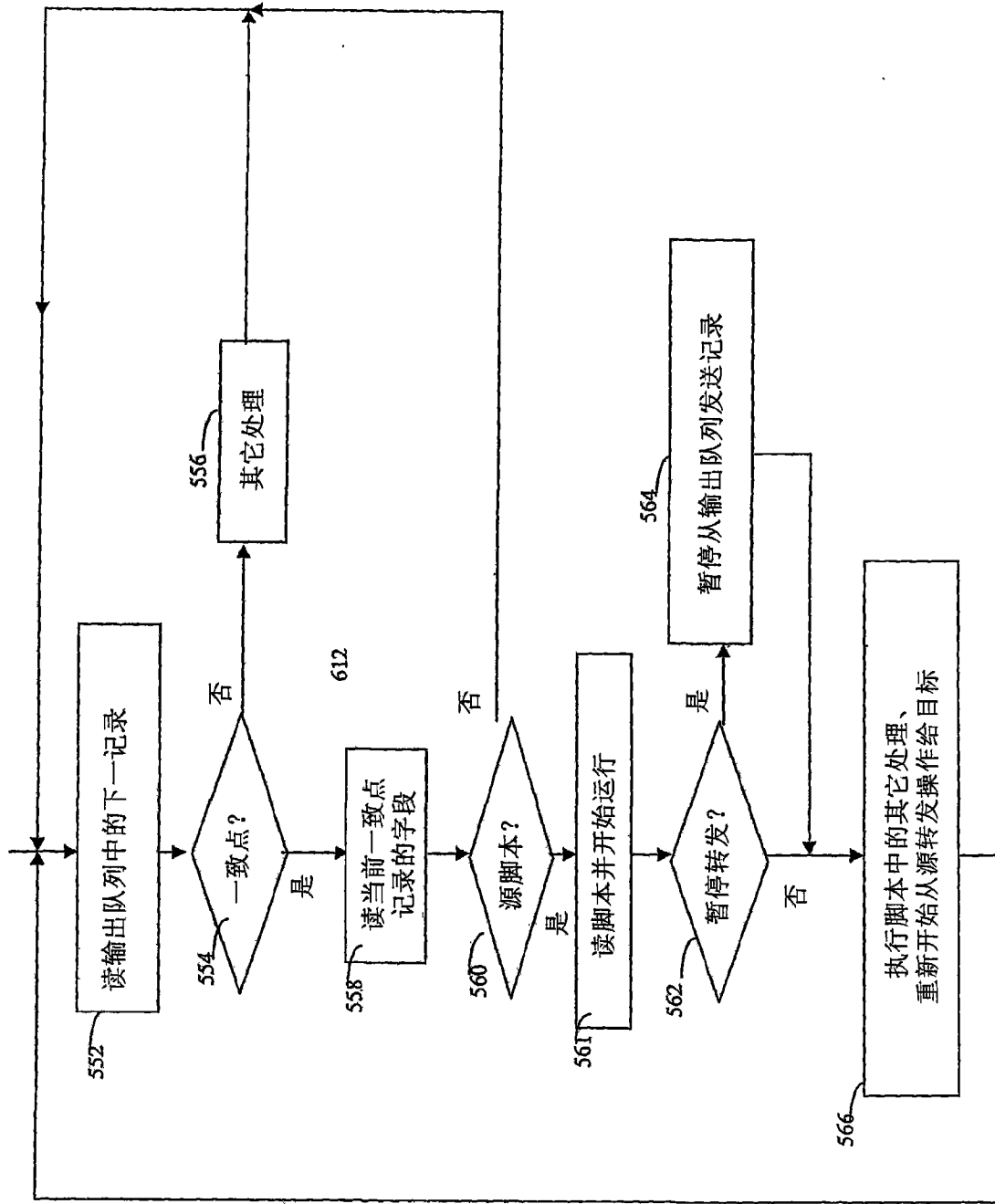


图9

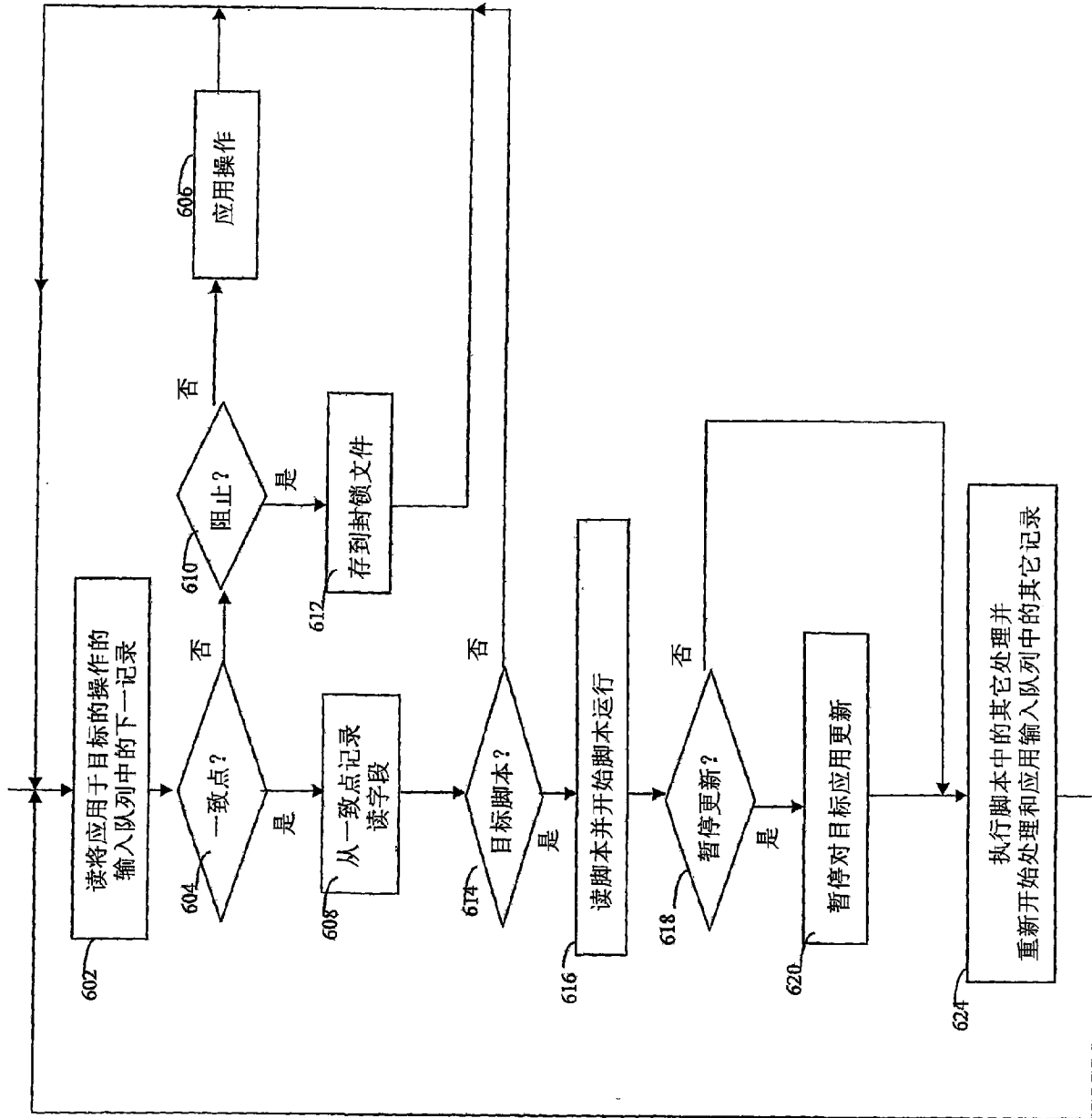


图10

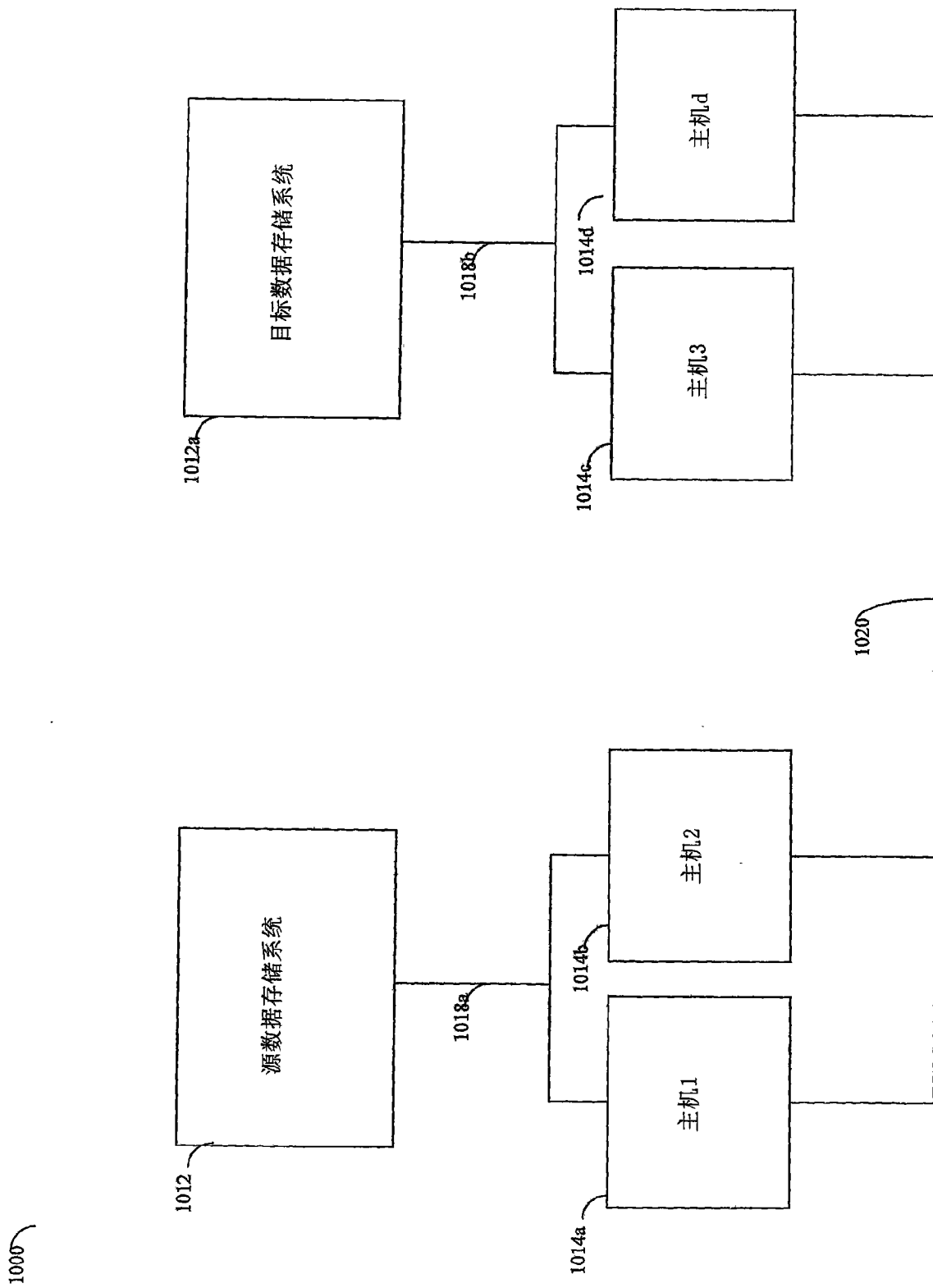


图11

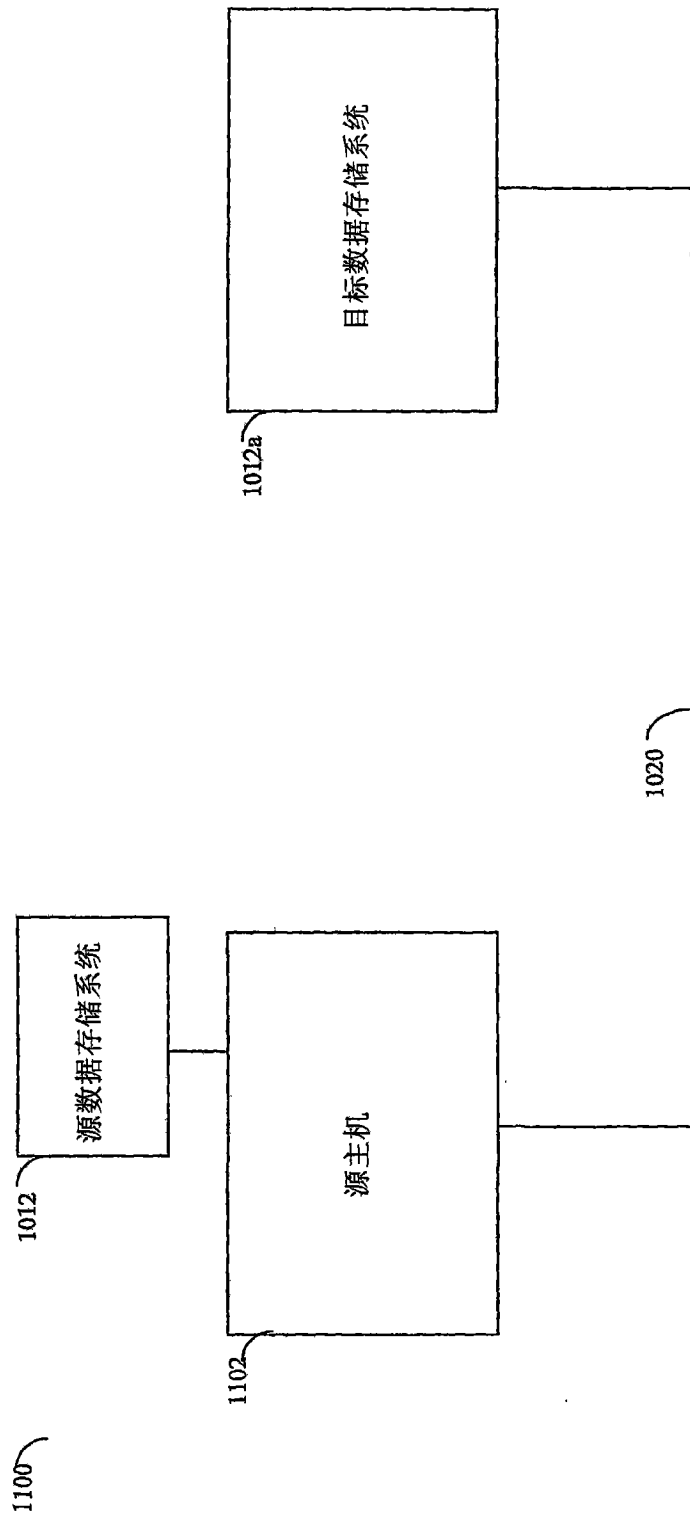


图12

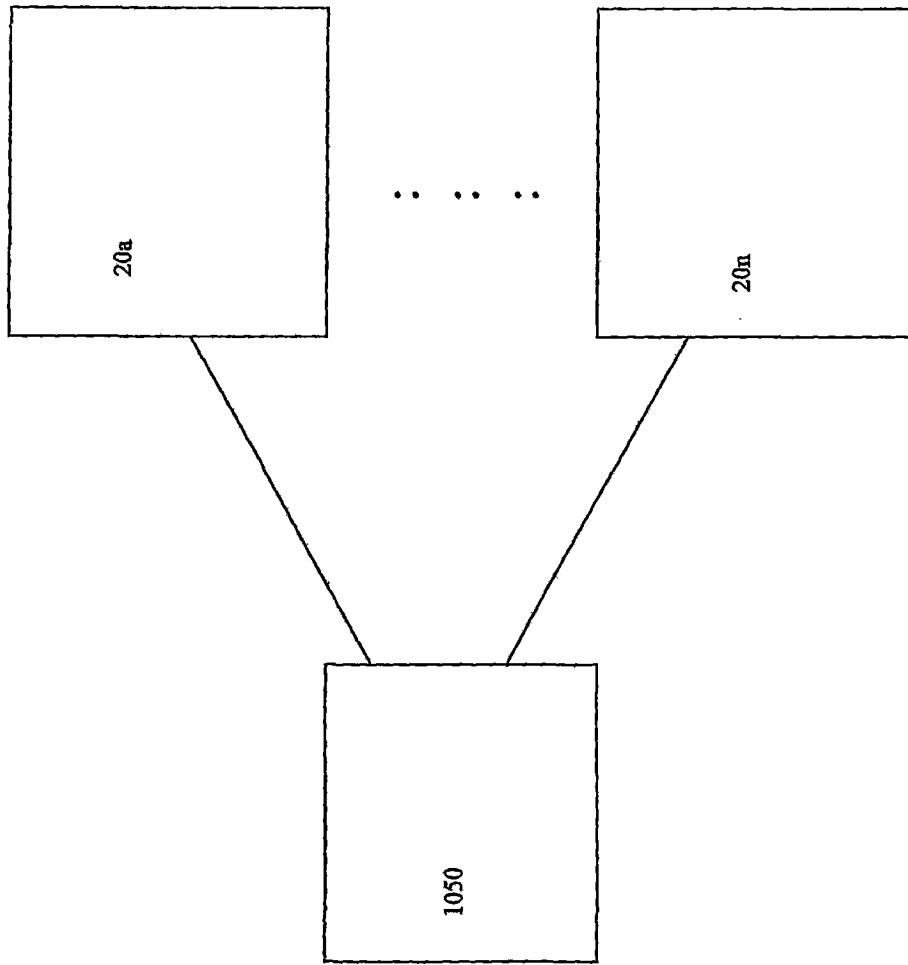


图13

1200