



[12]发明专利申请公开说明书

[21]申请号 96196941.5

[43]公开日 1998年10月14日

[11]公开号 CN 1196130A

[22]申请日 96.9.13

[30]优先权

[32]95.9.14 [33]FI[31]954333

[86]国际申请 PCT / FI96 / 00484 96.9.13

[87]国际公布 WO97 / 10552 英 97.3.20

[85]进入国家阶段日期 98.3.13

[71]申请人 诺基亚电信公司

地址 芬兰埃斯波

[72]发明人 皮卡·考伊弗尼耶米

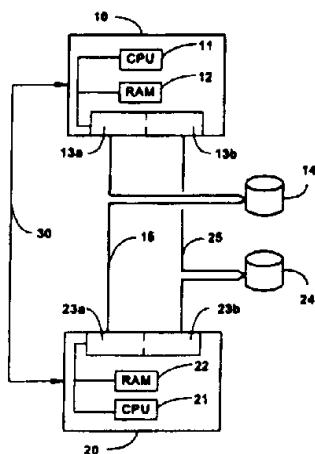
[74]专利代理机构 中国国际贸易促进委员会专利商标
事务所
代理人 鄂 迅

权利要求书 3 页 说明书 4 页 附图页数 1 页

[54]发明名称 在双工计算机部件中控制共享磁盘数据

[57]摘要

在容错系统中，已经知道将数据写入两个单独的磁盘驱动器。也知道对计算机进行双工操作，使得当一个计算机发生故障时，另一个计算机可以继续工作。在这样的系统中，活动部件通常控制着两个磁盘驱动器。也已经知道在计算机的主存储器中存储磁盘驱动器中最近所用的区域，因为这些区域（例如目录）可能必被再次读取。问题在于双工计算机的利用，在由于高速缓存于计算机（10, 20）的主存储器 12 和 22 中而提升操作时，每个计算机能够利用磁盘驱动器（14, 24）。如果一个计算机修改了该磁盘驱动器上的数据，另一计算机不必知道这种修改，但是使用其存储器中的过期数据。根据本发明，希望控制磁盘驱动器（14）的计算机（10）对其进行保留。如果保留成功，则执行磁盘任务。如果改变了磁盘驱动器（14）中的数据，则向另一计算机（20）通知这种改变，并且从其存储器中擦除与所改变的数据相应的高速缓存数据。



权 利 要 求 书

1. 在包括至少两个计算机（10，20）、一个共享磁盘驱动器（14）和将计算机（10，20）与磁盘驱动器（14）连接起来的磁盘驱动器总线（15）的系统中控制共享磁盘驱动器（14）的方法，其特征在于：使用共享磁盘驱动器（14）的第一计算机（10）

- 在一个磁盘任务之前，检查第二计算机（20）是否已经发出它已在共享磁盘驱动器（14）中做了修改的通知；

- 响应于所检测到的通知，从其高速缓冲存储器中擦除与在共享磁盘驱动器中已被改变的区域相应的数据；

- 执行磁盘任务；以及

- 向第二计算机（20）通知在磁盘驱动器（14）已被改变的数据。

2. 根据权利要求1的方法，其特征在于：在磁盘任务之前，第一计算机（10）确信第二计算机（20）不能与第一计算机（10）同时使用共享磁盘驱动器。

3. 根据权利要求2的方法，其特征在于：如果磁盘驱动器（14）尚未为第一计算机（10）保留，则第一计算机（10）通过将磁盘驱动器（14）保留一段它为了执行磁盘任务所需的时间，从而避免同时使用磁盘驱动器（14）。

4. 根据权利要求3的方法，其特征在于：在磁盘任务之后，使用共享磁盘驱动器（14）的第一计算机（10）释放磁盘驱动器（14）。

5. 根据权利要求3或4的方法，其特征在于：磁盘驱动器（14）的保留和/或释放是利用对磁盘驱动器（14）的保留命令执行的。

6. 根据权利要求3或4的方法，其特征在于：磁盘驱动器（14）的保留和/或释放是利用向第二计算机（20）提供的信号执行的。

7. 根据权利要求4的方法，其特征在于：磁盘驱动器（14）的释放响应于第二计算机（20）指示其需要磁盘驱动器（14）。

8. 根据权利要求3的方法，其特征在于：如果磁盘驱动器（14）不能被保留，则使用共享磁盘驱动器（14）的第一计算机（10）向第二计

算机（20）指示它需要该磁盘驱动器（14）。

9. 根据权利要求8的方法，其特征在于：如果第二计算机（20）未对第一计算机（10）发送的指示提供关于需要磁盘驱动器（14）的预定回答，则使用共享磁盘驱动器（14）的第一计算机释放对磁盘驱动器（14）的保留。

10. 根据权利要求1的方法，其特征在于：第一计算机（10）与第二计算机（20）之间的信号传送是经过磁盘驱动器总线（15）执行的。

11. 根据权利要求1的方法，其特征在于：第一计算机（10）与第二计算机（20）之间的信号传送是经过独立于磁盘驱动器总线（15）的消息总线（30）执行的。

12. 用于从两台计算机（10，20）控制一个共享磁盘驱动器（14）的系统，至少一台计算机（10）包括一个用于高速缓存磁盘驱动器（14）中数据的存储器（12），该系统还包括一条连接计算机（10，20）和磁盘驱动器（14）的磁盘驱动器总线（15），以及一种磁盘控制软件，其特征在于：磁盘控制软件包括以下功能：

- 限制某时间仅由一台计算机（10，20）使用磁盘驱动器（14）；
- 由第一计算机（10）从第二计算机（20）接收第一信号，并且对第一信号进行检查，以便检测第二计算机（20）是否已经改变了磁盘驱动器（14）中的数据；
- 从第一计算机（10）的存储器（12）中擦除与第二计算机（20）在磁盘驱动器（14）的数据中所做的修改相应的数据；
- 由第一计算机（10）产生第二信号，并且将其发送到第二计算机（20），使得第二信号向第二计算机（20）指示磁盘驱动器（14）中的哪些数据已被第一计算机（10）修改。

13. 根据权利要求12的系统，其特征在于限制某时间仅由一台计算机（10，20）使用磁盘驱动器（14）的步骤包括以下功能：

- 产生一个指示保留磁盘驱动器（14）的信号；
- 产生一个指示释放磁盘驱动器（14）的信号；以及
- 确定是否成功地保留磁盘驱动器（14）。

14. 根据权利要求13的系统，其特征在于计算机（10，20）之间的

信号在磁盘驱动器总线（ 15 ）上传播。

15. 根据权利要求 13 的系统，其特征在于计算机（ 10 ， 20 ）之间的至少某些信号经过独立于磁盘驱动器总线（ 15 ）的消息总线（ 30 ）传播。

16. 根据权利要求 13 的系统，其特征在于有关保留磁盘驱动器（ 14 ）的信息被存储在磁盘驱动器（ 14 ）中并从其中读取。

17. 根据权利要求 15 的系统，其特征在于有关保留磁盘驱动器（ 14 ）的信息经过消息总线（ 30 ）发送给其他计算机（ 10 ， 20 ）。

说 明 书

在双工计算机部件中控制共享磁盘数据

本发明涉及双工计算机系统，尤其涉及两个计算机部件能够用来控制同一磁盘驱动器上的数据的机构。

在例如交换中心中使用的容错系统中，已经知道将数据写入两个单独的磁盘驱动器。也知道对计算机进行双工操作，使得当一个计算机部件发生故障时，另一个计算机可以继续工作。在这样一个系统中，活动部件控制着两个磁盘驱动器。非活动部件或者被保留，或者发生故障，但是它通常不控制各磁盘驱动器。也已经知道在计算机的主存储器中存储磁盘驱动器的最后读取的区域，因为至少某些这些区域（例如目录）可能必被再次读取。

参照图 1，双工计算机通常是以这样一种方式设计的，使得只有计算机 10 和 20 中的一台（例如 10）使用磁盘驱动器 14 和/或磁盘驱动器 24。另一计算机（该情形下为 20）被保留或发生故障。第一磁盘驱动器总线 15 将两个计算机部件与第一磁盘驱动器 14 连接起来。相应地，第二磁盘驱动器总线 25 将两个计算机部件与第二磁盘驱动器 24 连接起来。图 1 示出每个计算机部件 10 和 20 分别包括两个磁盘驱动器控制器 13a 和 13b 以及 23a 和 23b 的一种实施方式。另一方案是，每个计算机部件 10 和 20 可以包括一个磁盘驱动器控制器，该控制器具有到磁盘驱动器总线 15 和 25 的单独接口。磁盘驱动器总线 15 和 25 可以例如是 SCSI（小型计算机系统接口）总线，并且磁盘驱动器 14 和 24 可以相应地是具有 SCSI 接口的磁盘驱动器。图 1 也示出一个将计算机 10 和 20 连接起来的双向消息总线 30。计算机 10 和 20 经过消息总线 30 交换维护消息，例如，它们需要相互的状态等。

图 1 的系统示出具有两个磁盘驱动器 14 和 24，因为在双工系统中，磁盘驱动器通常是重复的。在这个例子中，磁盘驱动器总线也被做成重复的。就本发明所相关的，磁盘驱动器 14 和 24 也可以连接到一个公共磁盘驱动器总线 15 或 25。然而，本发明涉及对各磁盘驱动器进行控制，而不管系统中可能包括的任何其他磁盘驱动器。因而，磁盘驱动器和磁盘驱动

器总线的数目对于本发明不是重要的。

将磁盘驱动器 14 的数据高速缓存到一个计算机 10 的主存储器 12 中对于本领域内熟练的技术人员而言是公知的技术。这种高速缓存是基于这种实际观察：当对磁盘驱动器 14 的某些部分执行一个磁盘任务时，可能不久执行与相同部分相关的另一个磁盘任务。因此，最好将与磁盘驱动器 14 的某些最近使用区域相应的数据存储在主存储器 12 中。为了对一个磁盘驱动器进行高速缓存，通常从主存储器 12 中保留某个区域。另一方案是，磁盘驱动器控制器 13a 和 13b 可以含有单独的高速缓冲存储器。

问题在于双工计算机的利用，在由于将磁盘驱动器高速缓存于计算机 10 和 20 的主存储器 12 和 22 中而提升操作时，每个计算机具有至少利用一个磁盘驱动器 14 和 24 的能力。如果一个计算机部件（例如 10）修改了磁盘驱动器 14 上的数据，另一计算机部件（该情形下为 20）不必知道。如果该计算机 20 依赖于其主存储器中的高速缓存数据而不从磁盘驱动器 14 读最近更新的数据，该系统产生混乱。由于容错，计算机部件 10 和 20 具有单独的主存储器，使得计算机 10 和 20 不能直接向相互的主存储器写已被改变的磁盘驱动器 14 的区域。

根据以上所述，本发明的目的在于提供一种两个计算机部件能够用来以可靠方式在共享磁盘驱动器中高速缓存数据的方法和系统。本发明的目的是利用独立权利要求的公开为特征的方法和系统实现的。最佳实施方式在从属权利要求中提出。

本发明是基于这种构思：需要使用一个磁盘驱动器的计算机部件确信另一个计算机部件不同时使用该磁盘驱动器，然后该计算机检查另一计算机是否已经修改了该磁盘驱动器中的任何数据。如果磁盘驱动器中的数据已被改变，则从其主存储器中擦除与该数据相应的高速缓存数据。之后，执行一个磁盘任务，并且向另一计算机通知在磁盘驱动器中的被改变的数据。

根据本发明的方法和系统的主要优点是，共享的磁盘数据不能由两个计算机控制，使得数据仍被高速缓存于两个计算机的主存储器中。另外，为便于实现，本发明几乎不对磁盘控制软件进行增加。考虑到由于复制两个计算机系统而将在任何情形下需要这种改变，通常不必向系统进行报

告。

将参照图 1 并结合最佳实施方式更详细地描述本发明，图 1 示出包括双工磁盘驱动器的双工计算机部件。

根据本发明的用于控制由两个计算机 10 和 20 共享的磁盘驱动器 14 的方法能够在图 1 所示的系统中执行，例如以下述步骤进行。假设第一计算机 10 想要使用第一磁盘驱动器 14。由于系统对称性，该方法由另一计算机以相应方式工作。

(i) 检查另一计算机 20 是否已经发出对在磁盘驱动器 14 中已做修改的通知；

(ii) 如果另一计算机 20 已经发出对在磁盘驱动器 14 中已做修改的通知，则擦除与磁盘驱动器 14 中改变区域相应的数据的高速缓存；

(iii) 执行一个磁盘任务；

(iv) 如果在该磁盘任务中改变磁盘驱动器 14 中的数据，则向另一计算机通知磁盘驱动器 14 中的改变区域。

该磁盘任务可以包括几条对磁盘驱动器 14 的命令。在该处理中，必须以某些方式避免另一计算机 20 使用磁盘驱动器 14。这是非常有利地以这样一种方式执行的，使得在该磁盘任务之前，第一计算机 10 例如利用 SCSI 命令 RESERVE 为其自身保留该磁盘驱动器。于是，在该磁盘任务之后，可以利用 SCSI 命令 RELEASE 释放该磁盘驱动器。在每个磁盘任务之后能够释放该磁盘驱动器。另一方案是，仅当另一计算机指示它需要访问该磁盘驱动器时才可释放。这种对磁盘驱动器需要的指示或者经过消息总线 30 或者经过磁盘驱动器总线 15 传送。

如果计算机 10 未能设法保留磁盘驱动器 14，则另一计算机 20 已保留了该磁盘驱动器。第一计算机 10 于是可以向第二计算机 20 发送它需要该磁盘驱动器 14 的指示。此后，第一计算机 10 可以尝试再次保留磁盘驱动器 14。

最好将新尝试的次数和总的时间限制为预定的值。选择这样的预定时间，使得另一计算机 20 将不能够在该时间内完成其磁盘任务并且应当释放磁盘驱动器 14。另一方案是，可以有两个预定时间限制。较短的时间限制（例如 5 秒）用于短保留，较长的时间限制（可以是几分钟）用于长保留，

例如在格式化和封装磁盘驱动器期间。

如果另一计算机 20 在与预定时间限制相应的时间内未释放磁盘驱动器 14，则向该计算机 20 发送一个询问，问其是否已保留了磁盘驱动器 14。如果对于该询问未收到预期的回答，则认为另一计算机 20 在执行磁盘任务时出现故障。在这种情形下，向磁盘驱动器 14 发送复位命令，并进行新的保留尝试。

根据本发明的一个实施方式，计算机 10 和 20 之间的以及从计算机到磁盘驱动器 14 的信号传输都经过磁盘驱动器总线 15 进行。当使用具有 SCSI 接口的磁盘驱动器时该实施方式尤其适用，因为 SCSI 包括定义很好的协议，利用该协议除了传统上与磁盘驱动器相关的命令之外，还能传送其他命令。SCSI 版本 II 在 ANSI 标准 X3.141-1994 中定义。

根据本发明另一实施方式，计算机 10 和 20 之间的信号传输经过独立于磁盘驱动器总线的消息总线进行。这种独立的消息总线 30 示于图 1。

根据本发明的方法最好实现为双工计算机系统的磁盘控制软件的一部分。根据本发明对磁盘控制软件增加的部分存储在存储磁盘控制软件其他部分的存储器中。如果根据本发明在计算机之间的信号传送是经过独立于磁盘驱动器总线的并连接两台计算机的消息总线进行的，则该消息总线不必构成已知技术。根据以上所述，对于本领域内熟练的技术人员显然知道如何实现在消息总线上的信号传送以便实现根据本发明的方法。

从以上所述，对于本领域内熟练的技术人员显然知道，作为技术改进，可以以许多不同的方式实现本发明的基本构思。因此，本发明及其实施方式不受上述例子的限制，而可以在权利要求的范围内变化。

说 明 书 附 图

