



(12) 发明专利

(10) 授权公告号 CN 1517885 B

(45) 授权公告日 2013.03.06

(21) 申请号 200410003501.1

(22) 申请日 2004.01.29

(30) 优先权数据

10/352,599 2003.01.28 US

(73) 专利权人 微软公司

地址 美国华盛顿州

(72) 发明人 D·C·布朗 M·V·利昂诺夫

M·M·伯德

(74) 专利代理机构 上海专利商标事务所有限公

司 31100

代理人 陈炜

(51) Int. Cl.

G06F 12/08 (2006.01)

(56) 对比文件

US 5809543 A, 1998.09.15, 全文.

CN 1195817 A, 1998.10.14, 全文.

US 6216212 B1, 2001.04.10, 全文.

审查员 尹剑峰

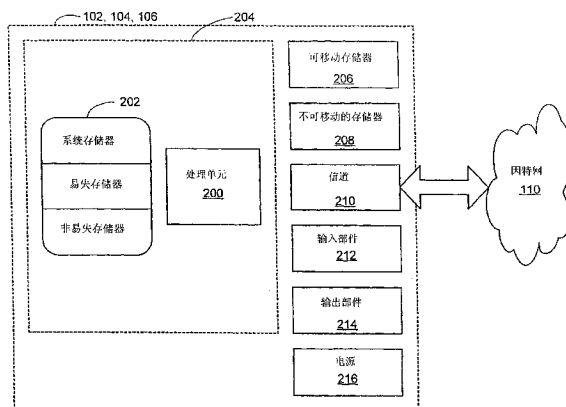
权利要求书 5 页 说明书 10 页 附图 13 页

(54) 发明名称

关于利用原子性更新的中央高速缓冲存储器
的方法和系统

(57) 摘要

揭示了一种中央高速缓冲存储器,该中央高速缓冲存储器的更新无需锁定管理费用。更新是“原子性的”,这体现在:这些更新无法被中途打断。应用程序始终自由地读取该高速缓冲存储器中的数据,从而通过参考表格来存取该数据。应用程序不是直接更新该高速缓冲存储器,相反,它们将更新请求发送到服务程序。为了更新该高速缓冲存储器,该服务程序在两个阶段中进行。在第一个阶段中,该服务程序准备该新数据,并将其加入该高速缓冲存储器,而不会更新该参考表格。在第一个阶段期间,存取该高速缓冲存储器的应用程序无法“看见”该新数据,因为还没有更新该参考表格。在第一个阶段完成之后,该服务程序执行该更新过程的第二个阶段:利用原子性来更新该参考表格。该双相更新过程使该高速缓冲存储器始终处于一致状态。



1. 一种在具有应用程序、具有中央高速缓冲存储器和具有与所述应用程序不同的高速缓冲存储器服务程序的计算环境中,供所述高速缓冲存储器服务程序用来尝试将有关的数据加入所述中央高速缓冲存储器的方法,所述方法包括:

从所述应用程序接收请求,该请求要求将所述有关的数据加入所述中央高速缓冲存储器;

确定是否将所述有关的数据加入所述中央高速缓冲存储器;其中如果确定将有关的数据加入到所述中央高速缓冲存储器:

首先,将所述有关的数据存储在所述中央高速缓冲存储器;以及,

其次,将对所述有关的数据的参考加入所述中央高速缓冲存储器中的参考表格;

其中将对所述有关的数据的参考加入所述中央高速缓冲存储器中的参考表格是原子性操作。

2. 如权利要求 1 所述的方法,其特征在于,所述应用程序在第一计算设备上运行,其中所述高速缓冲存储器服务程序在第二计算设备上运行,并且其中所述第一和第二计算设备彼此不同。

3. 如权利要求 2 所述的方法,其特征在于,从所述应用程序接收请求包括通过从以下构成的组中选择的通信介质来接收:串行通信线路、并行通信线路、局域网、内联网、调制解调器连接和因特网。

4. 如权利要求 1 所述的方法,其特征在于,所述有关的数据包括字体图示符。

5. 如权利要求 1 所述的方法,其特征在于,确定是否将所述有关的数据加入所述中央高速缓冲存储器至少部分地基于从以下构成的组中选择的操作:

确定所述有关的数据是否将适合所述中央高速缓冲存储器;

确定是否已经在所述中央高速缓冲存储器中找到所述有关的数据;

收集有关该请求和有关从所述应用程序或其他应用程序接收的其他请求的统计数字;以及,

参考从所述应用程序接收的收集的统计数字。

6. 如权利要求 1 所述的方法,其特征在于,将所述有关的数据加入所述中央高速缓冲存储器包括:

在所述中央高速缓冲存储器中的参考表格中选择项目,该选择至少部分基于所述有关的数据的散列。

7. 如权利要求 6 所述的方法,还包括:

如果对所选项目的检查显示所选项目没有提及所述中央高速缓冲存储器中的数据元素,那么:

用所述有关的数据来填充所述中央高速缓冲存储器中的新的数据元素;以及,

加入对所述参考表格中的所选项目中的有关的数据的参考。

8. 如权利要求 7 所述的方法,其特征在于,加入对有关的数据的参考包括:计算到所述新的数据元素的整数偏移量。

9. 如权利要求 6 所述的方法,还包括:

如果对所选项目的检查显示所选项目提及所述中央高速缓冲存储器中的数据元素,那么:

如果对所述数据元素的检查显示该数据元素提及进一步的数据元素,那么,重复检查所述数据元素的步骤,其中检查所述进一步的数据元素,直到达到没有提及另一个数据元素的数据元素为止;以及,

当达到没有提及另一个数据元素的数据元素时,那么:

用所述有关的数据来填充所述中央高速缓冲存储器中的新的数据元素;以及,

加入对所述数据元素中的有关的数据的参考。

10. 如权利要求 9 所述的方法,其特征在于,加入对有关的数据的参考包括:计算到所述新的数据元素的整数偏移量。

11. 如权利要求 1 所述的方法,还包括:

确定是否创建新的中央高速缓冲存储器;以及,

根据创建所述新的中央高速缓冲存储器的确定,来创建所述新的中央高速缓冲存储器。

12. 如权利要求 11 所述的方法,其特征在于,确定是否创建所述新的中央高速缓冲存储器至少部分基于从以下构成的组中选择的操作:

确定所述有关的数据是否将适合该中央高速缓冲存储器;

确定该中央高速缓冲存储器中的参考表格是否装满;以及

确定已使用所述中央高速缓冲存储器中的数据有多新近。

13. 如权利要求 11 所述的方法,其特征在于,创建所述新的中央高速缓冲存储器包括:创建所述新的中央高速缓冲存储器;

在所述中央高速缓冲存储器中加入对所述新的中央高速缓冲存储器的参考;以及,

将所述中央高速缓冲存储器标记为过时。

14. 如权利要求 13 所述的方法,还包括:

用从该过时的中央高速缓冲存储器中选择的数据元素来填充所述新的中央高速缓冲存储器,该选择至少部分基于被应用于所收集的统计数字的试探法,所收集的统计数字是关于该过时的中央高速缓冲存储器中的数据元素的使用。

15. 如权利要求 14 所述的方法,其特征在于,从以下构成的组中选择该试探法:最新近使用的、最频繁使用的。

16. 如权利要求 13 所述的方法,还包括:

标记该过时的中央高速缓冲存储器,以便当没有应用程序提及该过时的中央高速缓冲存储器时,自动删除该过时的中央高速缓冲存储器。

17. 一种在具有应用程序、具有中央高速缓冲存储器和具有与所述应用程序不同的高速缓冲存储器服务程序的计算环境中,供所述应用程序用于尝试存取有关的数据的方法,该方法包括:

搜索所述中央高速缓冲存储器中的所述有关的数据;

确定是否在所述中央高速缓冲存储器中找到所述有关的数据;

如果在所述中央高速缓冲存储器中找到所述有关的数据,那么在所述中央高速缓冲存储器中存取所述有关的数据;以及,

如果在所述中央高速缓冲存储器中没有找到所述有关的数据,那么:

尝试在除所述中央高速缓冲存储器以外的其他地方存取所述有关的数据;以及,

向所述高速缓冲存储器服务程序发出请求,要求将所述有关的数据加入所述中央高速缓冲存储器;

其中如果确定将所述有关的数据加入所述中央高速缓冲存储器;

首先,将所述有关的数据存储在所述中央高速缓冲存储器;以及,

其次,将对所述有关的数据的参考加入所述中央高速缓冲存储器中的参考表格;

其中将对所述有关的数据的参考加入所述中央高速缓冲存储器中的参考表格是原子性操作;

其中尝试在除所述中央高速缓冲存储器中以外的其他地方存取所述有关的数据包括:

在本地的高速缓冲存储器中搜索对于所述应用程序是有关的数据;

如果在所述本地高速缓冲存储器中找到所述有关的数据,那么在本地高速缓冲存储器中存取所述有关的数据;以及,

如果在所述本地高速缓冲存储器中没有找到所述有关的数据,那么,调用例行程序来创建有关的数据。

18. 如权利要求 17 所述的方法,其中,所述应用程序在第一计算设备上运行,其中所述高速缓冲存储器服务程序在第二计算设备上运行,并且其中,所述第一和第二计算设备彼此不同。

19. 如权利要求 18 所述的方法,其特征在于,向所述高速缓冲存储器服务程序发出请求包括通过从以下构成的组中选择的通信介质来发出:串行通信线路、并行通信线路、局域网、内联网、调制解调器连接和因特网。

20. 如权利要求 17 所述的方法,其特征在于,所述有关的数据包括字体图示符。

21. 如权利要求 17 所述的方法,其特征在于,在所述中央高速缓冲存储器中搜索有关的数据包括:

在所述中央高速缓冲存储器中的参考表格中选择项目,该选择至少部分基于所述有关的数据的散列。

22. 如权利要求 21 所述的方法,还包括:

如果对所选项目的检查显示所选项目没有提及所述中央高速缓冲存储器中的数据元素,那么,决定所述有关的数据不在所述中央高速缓冲存储器中。

23. 如权利要求 21 所述的方法,还包括:

如果对所选项目的检查显示所选项目提及所述中央高速缓冲存储器中的数据元素,并且如果对所述数据元素的检查显示所述数据元素包含所述有关的数据,那么决定所述有关的数据在所述中央高速缓冲存储器中。

24. 如权利要求 23 所述的方法,其特征在于,所选项目借助于所述中央高速缓冲存储器中的整数偏移量来提及数据元素。

25. 如权利要求 21 所述的方法,还包括:

如果对所选项目的检查显示所选项目提及所述中央高速缓冲存储器中的数据元素,并且,如果对所述数据元素的检查显示所述数据元素不包含所述有关的数据,那么:

如果所述数据元素没有提及进一步的数据元素,那么决定所述有关的数据不在所述中央高速缓冲存储器中;以及,

如果所述数据元素提及进一步的数据元素,那么重复检查所述数据元素的步骤,其中检查所述进一步的数据元素,直到达到包含所述有关的数据的数据元素或者达到没有提及另一个数据元素的数据元素为止。

26. 如权利要求 17 所述的方法,还包括:

如果在所述本地高速缓冲存储器中没有找到所述有关的数据,那么将所创建的有关的数据加入所述本地高速缓冲存储器。

27. 如权利要求 17 所述的方法,还包括:

如果在所述本地高速缓冲存储器中没有找到所述有关的数据,那么向所述高速缓冲存储器服务程序发出请求,要求将所创建的有关的数据加入所述中央高速缓冲存储器。

28. 如权利要求 17 所述的方法,其特征在于,向所述高速缓冲存储器服务程序发出请求包括:

存储该请求;

存储一个或多个其他的请求;以及,

向所述高速缓冲存储器服务程序发出所存储的请求。

29. 如权利要求 17 所述的方法,还包括:

检验所述中央高速缓冲存储器,以了解所述中央高速缓冲存储器是否被标记为过时;以及,

如果所述中央高速缓冲存储器被标记为过时,那么使用该过时的中央高速缓冲存储器中的参考来对另一个中央高速缓冲存储器进行定位。

30. 如权利要求 17 所述的方法,还包括:

收集有关所述应用程序对所述中央高速缓冲存储器的使用的统计数字;以及,

将所收集的统计数字发给所述高速缓冲存储器服务程序。

31. 一种在具有应用程序、具有中央高速缓冲存储器和具有与所述应用程序不同的高速缓冲存储器服务程序的计算环境中,供所述应用程序用于尝试存取有关的数据的系统,该系统包括:

用于搜索所述中央高速缓冲存储器中的所述有关的数据的装置;

用于确定是否在所述中央高速缓冲存储器中找到所述有关的数据的装置;

用于如果在所述中央高速缓冲存储器中找到所述有关的数据,那么在所述中央高速缓冲存储器中存取所述有关的数据的装置;以及,

如果在所述中央高速缓冲存储器中没有找到所述有关的数据,那么:

用于尝试在除所述中央高速缓冲存储器以外的其他地方存取所述有关的数据的装置;以及,

用于向所述高速缓冲存储器服务程序发出请求,要求将所述有关的数据加入所述中央高速缓冲存储器的装置;

其中如果确定将所述有关的数据加入所述中央高速缓冲存储器:

用于将所述有关的数据存储在所述中央高速缓冲存储器的装置;以及,

用于将对所述有关的数据的参考加入所述中央高速缓冲存储器中的参考表格的装置;

其中将对所述有关的数据的参考加入所述中央高速缓冲存储器中的参考表格是原子

性操作；

其中用于尝试在除所述中央高速缓冲存储器中以外的其他地方存取所述有关的数据的装置包括：

用于在本地的高速缓冲存储器中搜索对于所述应用程序是有关的数据的装置；

用于如果在所述本地高速缓冲存储器中找到所述有关的数据，那么在本地高速缓冲存储器中存取所述有关的数据的装置；以及，

用于如果在所述本地高速缓冲存储器中没有找到所述有关的数据，那么，调用例行程序来创建有关的数据的装置。

关于利用原子性更新的中央高速缓冲存储器的方法和系统

技术领域

[0001] 本发明通常涉及计算机记忆存储技术,更具体地说,涉及高速缓冲存储器。

背景技术

[0002] 要创建或访问计算机应用程序所需要的某些数据,其费用昂贵。这些费用可以包括计算资源,用于计算在网络上访问该数据的数据费用和输送费用(包括带宽和时间)。一旦在花费资源来创建或访问这些数据之后,计算设备就经常会将该数据存储在“高速缓冲”存储器中。然后,如果该计算设备再次需要该数据,则可以廉价地从该高速缓冲存储器中存取这些数据。

[0003] 该高速缓冲存储器对于该原始应用程序或原始计算设备而言可以是本地的,也可以在几个应用程序和设备之中共享该高速缓冲存储器。后一种类型的高速缓冲存储器经常被称作“中央”高速缓冲存储器。在一些环境中,每个应用程序支持供其自己使用的本地高速缓冲存储器,同时与其他应用程序共享中央高速缓冲存储器。对该中央高速缓冲存储器进行优化,用于存储对一个以上的应用程序有用的数据;同时,这些本地高速缓冲存储器可以为对于每个应用程序而言是特定的那些数据提供超高速缓存的各种好处。

[0004] 在中央高速缓冲存储器中管理数据很少是简单明了的。尝试从该中央高速缓冲存储器中读取数据的多个应用程序很少会引起难题,但当至少一个应用程序希望将数据加入该高速缓冲存储器时,就不可以这样说了。如果在一个应用程序正对该中央高速缓冲存储器进行写入的同时,允许其他应用程序从该高速缓冲存储器进行读取,那么,这些阅读器会获得过时或甚至是曲解的数据。当一个以上的应用程序希望将数据加入该高速缓冲存储器时,这种存取协调问题会加剧。

[0005] 缓解这种存取协调问题的一种普通的方法被称作“高速缓冲存储器锁定”。只要一个应用程序希望通过增加、删除或修改其内容来更改该高速缓冲存储器的内容,它就可以寻求对“锁”数据结构的单独存取。当复写器应用程序具有该锁时,它可以修改该高速缓冲存储器,并且,只要复写器具有该锁,就可以防止其他应用程序存取该高速缓冲存储器。这样,可以防止阅读器获得过时或曲解的数据。如果两个应用程序都希望修改该高速缓冲存储器,那么,它们中的一个应用程序必须等候,直到另一个放弃该锁为止。

[0006] 在协调对中央高速缓冲存储器的存取的过程中,锁会非常有用。但是,显而易见,只要一个应用程序希望修改该高速缓冲存储器,这些锁就会延迟所有应用程序的存取。对于一些中央高速缓冲存储器而言,应用程序容易容忍这种关闭。但是,对于其他高速缓冲存储器而言,这真的会是一件麻烦事。例如,考虑字体图示符高速缓冲存储器。计算机屏幕上所显示的字符由被称作“图示符”的单独元素构成。由于这些图示符中的一些图示符包含重大数量的数据,并且由于这些图示符中的一些图示符在其生成过程中消耗重要的计算资源,因此,它们是中央高速缓冲存储器的理想主题。但是,如果在将新的图示符加入该字体图示符高速缓冲存储器的同时锁定该字体图示符高速缓冲存储器,则会在写入该计算机屏幕的应用程序中引起显著的延迟。

[0007] 当限制中央高速缓冲存储器可用的这些内存资源时,产生了另一个高速缓冲存储器管理问题。希望将数据加入该高速缓冲存储器的多个应用程序彼此独立地进行操作。这样,通常关于应该将哪些数据加入该中央高速缓冲存储器,以便改善该操作环境,这些应用程序没有“全局”观点。如果该中央高速缓冲存储器增长得太大,并且用较小的尺寸再形成,以便虑及进一步的增加,则会产生同样的问题。没有一个应用程序可以决定:哪些数据应该被保留在该中央高速缓冲存储器中,以及哪些数据最好被除去,以便释放内存,供高速缓冲存储器将来增长。

发明内容

[0008] 鉴于前述内容,本发明提供了一种中央高速缓冲存储器,该中央高速缓冲存储器可以在没有锁定的延迟管理费用的情况下被加以更新,并且对该高速缓冲存储器内的数据的重要性具有全局观点。“原子性”更新提供了存取协调的各种好处,而不会招致锁定的延迟管理费用。高速缓冲存储器更新是“原子性的”,这体现在:它们被设计成无法被中途打断。通过这些高速缓冲存储器更新,当被应用程序存取时,高速缓冲存储器内总是保持最新、一致。

[0009] 应用程序总是自由地读取该中央高速缓冲存储器中的数据,从而通过参考表格来存取该数据。但是,这些应用程序没有直接更新该高速缓冲存储器,相反,它们将更新请求发送到服务程序。为了更新该高速缓冲存储器,该高速缓冲存储器服务程序在两个阶段中进行。在第一个阶段中,该高速缓冲存储器服务程序准备该新数据并将它们加入该高速缓冲存储器,而不会更新该参考表格。这第一个阶段可能会花费一些时间,但应用程序仍然可以充分地对该高速缓冲存储器进行存取,这是因为该高速缓冲存储器没有被锁定。在这第一个阶段期间,存取该高速缓冲存储器的应用程序无法“看见”该新数据,因为该参考表格还没有被更新。只有在该高速缓冲存储器数据被充分加以准备并被载入该高速缓冲存储器之后,该高速缓冲存储器服务程序才执行该更新过程的第二个阶段:更新该参考表格。该更新(包括只改变一个指针)利用原子性来加以执行,而不会锁定该高速缓冲存储器。这样,该双相更新过程不要求锁定该高速缓冲存储器,并且一直使该高速缓冲存储器处于存取应用程序的有效状态。由于一个高速缓冲存储器服务程序执行所有的更新,因此,不需要锁在多个高速缓冲存储器复写器之中进行协调。

[0010] 该高速缓冲存储器服务程序收集关于如何使用该高速缓冲存储器中的数据的统计数字。当该高速缓冲存储器增长得太大时,该高速缓冲存储器服务程序使用这些统计数字来决定应该将哪些数据拷贝到新的高速缓冲存储器中。再次在两个阶段中,利用原子性来创建这个新的高速缓冲存储器。在第一个阶段期间,该高速缓冲存储器服务程序创建这个新的高速缓冲存储器,并且用来自老的高速缓冲存储器的所选择的数据来填充这个新的高速缓冲存储器。迄今,应用程序还不具备有关这个新的高速缓冲存储器的知识。当这个新的高速缓冲存储器准备好使用时,该高速缓冲存储器服务程序在老的高速缓冲存储器的头部中为其加入参考。然后,在第二个阶段中,并且通过使用另一项原子性操作,该高速缓冲存储器服务程序将老的高速缓冲存储器标记为“过时的”。一注意到这个老的高速缓冲存储器被标记为过时,应用程序就遵循对该新高速缓冲存储器的参考,并且开始只使用这个新的高速缓冲存储器。如在高速缓冲存储器内的更新中,按这样一种方式来执行用于替换

整个高速缓冲存储器的这种机制,以便应用程序始终可以看见一致的高速缓冲存储器。

[0011] 应用程序可以继续使用过时的高速缓冲存储器,直到它们注意到过时的标志并转到这个新的高速缓冲存储器。一旦已转换所有应用程序,就自动删除该过时的高速缓冲存储器。

[0012] 在一些实施例中,高速缓冲存储器内的参考表格包括偏移量,这些偏移量相对于该高速缓冲存储器内的另一个位置来规定该高速缓冲存储器中所存储的数据位置。这具有一个优点:该高速缓冲存储器可以被存储为文件,并且在主宿该高速缓冲存储器的计算设备重新启动之后被立即再使用。

[0013] 该中央高速缓冲存储器可以由一个计算设备来主宿,并可以由那个计算设备和其他计算设备上的应用程序来使用。每个应用程序也可以具有结合该中央高速缓冲存储器来加以使用的其自己的本地高速缓冲存储器。如果该本地高速缓冲存储器具有与该中央高速缓冲存储器相同的数据结构,那么,可以为这两种高速缓冲存储器使用相同的超高速缓存访问代码。

[0014] 该高速缓冲存储器服务程序将试探法应用于其所搜集的关于高速缓冲存储器使用的统计数字,以便决定当替换该高速缓冲存储器时要保存哪些数据。在一些实施例中,提供了用户界面,从而允许改变该试探法,并允许监控该高速缓冲存储器的操作。

[0015] 附图说明

[0016] 所附权利要求书详细陈述了本发明的各个特点,但通过以下详细的描述并结合附图,可以最佳程度地理解本发明及其目的和优点。在这些附图中:

[0017] 图 1a 是框图,表现了经由局域网 (LAN) 来共享超高速缓存数据的三个计算设备;

[0018] 图 1b 是框图,表现了在图 1a 的这些计算设备之中被共享的中央高速缓冲存储器;

[0019] 图 2 是示意图,通常展示了支持本发明的示范计算机系统;

[0020] 图 3a ~ 3d 一起构成流程图,展示了根据本发明的示范方法,应用程序使用该示范方法来尝试从中央高速缓冲存储器存取数据;

[0021] 图 4a 是示意图,表现了根据本发明的中央高速缓冲存储器的示范数据结构;

[0022] 图 4b 是示意图,表现了中央高速缓冲存储器中的数据元素的示范数据结构;

[0023] 图 5a ~ 5d 一起构成根据本发明的示范方法的流程图,该示范方法是关于维持中央高速缓冲存储器的例行程序;以及,

[0024] 图 6 是示范方法的流程图,该示范方法用于配置中央高速缓冲存储器服务程序。

[0025] 具体实施方式

[0026] 参考这些附图(其中,类似的参考数字提及类似的元件),将本发明展示为在合适的计算环境中被加以执行。下文基于本发明的实施例,不应该被视作在这里未作明确描述的选择性实施例方面限制本发明。

[0027] 在下文中,除非另有指示,参照由一个或多个计算设备执行的动作和操作的符号表示,来描述本发明。照此,将会理解:这类动作和操作(有时被称作“是计算机执行的”)包括用组织形式表示数据的电信号计算设备的处理单元所执行的操作。这项操作变换该数据或将其保存在该计算设备的存储系统中的位置,这样,用精通该技术领域的人所熟悉的方式来重新配置或改变该设备的操作。其中保存有数据的这些数据结构是具有由该数据的

格式定义的特殊属性的存储器的物理位置。但是,前述上下文中正在描述本发明,而这并不意味着起限制的作用,因为精通该技术领域的人将会理解:也可以在硬件中执行下文中所描述的各种动作和操作。

[0028] 本发明提供了一种利用原子性更新的中央高速缓冲存储器。该中央高速缓冲存储器可以由运行于主宿该高速缓冲存储器的计算设备上的应用程序来专用,或者,如在图 1a 中,该中央高速缓冲存储器可以由几个计算设备上的应用程序来使用。在图 1a 中,三个计算设备——A102、B104 和便携式计算机 106 处于超高速缓存数据共享环境 100 中。中央高速缓冲存储器(未示出)驻留在计算设备 A102 上,并且可以经由 LAN 108 由应用程序来加以存取。现存的标准通信协议用于在共享环境 100 内的各个计算设备之中传送高速缓冲存储器请求和响应。

[0029] 与因特网 110 的连接被示出指出:即使远程计算设备也可以加入超高速缓存数据共享环境 100。实际上,这类远程设备存取该中央高速缓冲存储器所需的通信时间增加了,这样,为了满足该高速缓冲存储器提供对数据的快速存取的目的,可运行计数器。大多数中央高速缓冲存储器的情况将只涉及一个或至多几个位置接近的计算设备。

[0030] 图 1b 呈现了图 1a 中的超高速缓存数据共享环境 100 的示范实施例的结构细节。应用程序 A 112 正在计算设备 A 102 上运行。在一起构成应用程序 A 112 的各种例行程序之中,是本地和中央高速缓冲存储器搜索程序 114。当应用程序 A 112 需要可以驻留在高速缓冲存储器中的数据时,这个程序 114 在当前的本地高速缓冲存储器 116 和当前的中央高速缓冲存储器 118 中寻找该数据。本地高速缓冲存储器 116 是应用程序 A 112 的一部分,并且在应用程序 A 112 的控制之下。中央高速缓冲存储器 118 不是应用程序 A 112 的一部分,但应用程序 A 112 和其他应用程序可以对中央高速缓冲存储器 118 进行存取。

[0031] 如果中央高速缓冲存储器搜索程序 114 在本地高速缓冲存储器 116 中或在中央高速缓冲存储器 118 中找到所请求的数据,则它将该数据返回到应用程序 A 112。如果没有找到所请求的数据,则在其他地方搜寻该数据,或者创建该数据。一旦找到或创建该数据,高速缓冲存储器搜索程序 114 就要求:通过调用本地高速缓冲存储器更新程序 120,将该数据加入本地高速缓冲存储器 116。也将请求发送到中央高速缓冲存储器服务程序 124,以便将该数据加入中央高速缓冲存储器 118。

[0032] 当将数据加入本地高速缓冲存储器和中央高速缓冲存储器时,这两种高速缓冲存储器都会增长。当它们变得太大时,可创建新的高速缓冲存储器,并用先前的高速缓冲存储器的部分数据来填充这个新的高速缓冲存储器。为了选择将哪些数据传送到这个新的高速缓冲存储器,可以应用任何数量的方法。例如,选择最新近使用的数据或最频繁使用的数据。图 1b 表现了先前的本地高速缓冲存储器 122 以及当前的本地高速缓冲存储器 116。

[0033] 当创建新的中央高速缓冲存储器时,注意使用该中央高速缓冲存储器来防止应用程序的中断。当新的中央高速缓冲存储器 118 准备就绪时,将较老的中央高速缓冲存储器 126 标记为“过时的”,以便警告应用程序:具备较新的高速缓冲存储器。但是,那些应用程序不需要立即切换到较新的高速缓冲存储器 118,而可以选择暂时继续存取较老的中央高速缓冲存储器 126。一旦不再有应用程序存取较老的中央高速缓冲存储器 126,就删除那个高速缓冲存储器。

[0034] 应用程序 B 128 在超高速缓存数据共享环境 100 中的另一个计算设备上运行。该

应用程序被示出没有本地高速缓冲存储器,但它具有中央高速缓冲存储器搜索程序 130。那个例行程序仍然存取较老的中央高速缓冲存储器 126,但还没有注意到这个高速缓冲存储器被标记为过时。高速缓冲存储器搜索程序 130 可以要求将数据加入该中央高速缓冲存储器,但中央高速缓冲存储器服务程序 124 将把那个数据加入当前的中央高速缓冲存储器 118,而不是加入过时的高速缓冲存储器 126。

[0035] 图 1 中的计算设备 102、104 和 106 可能属于任何构造。图 2 是框图,通常展示了支持本发明的示范计算机系统。图 2 中的计算机系统只是合适的环境的一个例子,并且,它并不意在提出有关本发明的使用或功能性的范围的任何限制。不应该将计算设备 102 解释为具有涉及图 2 中所展示的任何部件或部件组合的任何从属性或要求。本发明可用于其他许多通用或专用的计算环境或配置。适合用于本发明的众所周知的计算系统、环境和配置的例子包括(但不局限于)个人计算机、服务器、手持设备或便携式设备、写字板设备、多处理器系统、基于微处理器的系统、置顶盒、可编程消费电子设备、网络 PCs、小型计算机、大型计算机以及包括以上任何系统或设备的分布式计算环境。在其最基本的配置中,计算设备 102 通常包括至少一个处理单元 200 和存储器 202。存储器 202 可能是易失的(例如, RAM)、非易失的(例如, ROM 或快闪存储器),也可能是两者的某种组合。在图 2 中,用虚线 204 展示了该最基本的配置。计算设备 102 可能具有额外的特点和功能性。例如,计算设备 102 可以包括额外的存储器(可移动的和不可移动的),该额外的存储器包括(但不局限于)磁盘、光盘和磁带。在图 2 中,由可移动存储器 206 和不可移动的存储器 208 来展示这类额外的存储器。计算机存储介质包括易失的和非易失的可移动和不可移动的介质,该介质用于信息(例如,计算机可读指令、数据结构、程序模块或其他数据)存储的任何方法或技术来加以执行。存储器 202、可移动存储器 206 和不可移动的存储器 208 都是计算机存储介质的例子。计算机存储介质包括(但不局限于)RAM、ROM、EEPROM、快闪存储器、其他存储技术、CD-ROM、数字通用磁盘、其他光学存储器、盒式磁带、磁带、磁盘存储器、其他磁性存储设备,以及可以被用来存储所需信息并可以由设备 102 来存取的其他任何介质。任何这类计算机存储介质都可能是设备 102 的一部分。设备 102 也可能包含允许该设备与其他设备进行通信的信道 210。信道 210 是通信介质的例子。通信介质通常具体表现计算机可读指令、数据结构、程序模块或调制数据信号(例如,载波或其他传送机制)中的其他数据,并且,它包括任何信息传递介质。术语“调制数据信号”意味着一种信号,该信号的一个或多个特征用这样一种方式来加以设置或更改,以便为该信号中的信息编码。举例来讲(不作限制),通信介质包括有线介质(例如,有线网络和直线连接)和无线介质(例如,声音、RF、红外线和其他无线介质)。如这里所使用的术语“计算机可读介质”包括存储介质和通信介质。计算设备 102 也可能具有输入设备 212(例如,键盘、鼠标、笔、语音输入设备、写字板、触摸输入设备等)。也可能包括诸如显示器(可能与触摸输入设备结合)、扬声器和打印机的输出设备 214。所有这些设备在该技术领域中都众所周知,这里不需要详细讨论。

[0036] 图 3a ~ 3d 呈现了关于图 1b 中的本地和中央高速缓冲存储器搜索程序 114 的示范方法。该流程图包括不需要被包括在高速缓冲存储器搜索程序 114 的每个实施例中的许多选项。

[0037] 在该流程图始于图 3a 中的步骤 300 之前,向高速缓冲存储器搜索程序 114 请求存取与应用程序 A 112 有关的数据。在步骤 300 中,通过检查是否可以存取中央高速缓冲存

存储器,来开始高速缓冲存储器搜索程序 114。如果没有中央高速缓冲存储器可存取,那么,高速缓冲存储器搜索程序 114 进行到图 3d 中的步骤 326,并在本地高速缓冲存储器中寻找有关的数据。高速缓冲存储器搜索程序 114 的其他实施例根据这些情况来尝试在该中央高速缓冲存储器之前存取本地高速缓冲存储器。

[0038] 如果中央高速缓冲存储器可存取,那么,高速缓冲存储器搜索程序 114 在步骤 302 中进行核对,来了解该中央高速缓冲存储器是否被标记为“过时的”。注意,这个步骤和随后的步骤 304 是任选的:高速缓冲存储器搜索程序 114 可以继续使用过时的高速缓冲存储器。但是,为了保持当前状态,高速缓冲存储器搜索程序 114 应该定期检验该中央高速缓冲存储器的状况,以便一旦它注意到老的高速缓冲存储器过时了,它就会存取并放弃这个老的高速缓冲存储器。

[0039] 在中央高速缓冲存储器被标记为过时之前,中央高速缓冲存储器服务程序 124 创建新的中央高速缓冲存储器(见图 5a 和 5b 中的步骤 512 ~ 518)。这样,在步骤 304 中,高速缓冲存储器搜索程序 114 可以遵循在该过时的中央高速缓冲存储器的头部中对该当前的中央高速缓冲存储器的参考。由于请求到达时与处理那个请求时之间的延迟会很重大,因此,在该延迟期间,一个以上的高速缓冲存储器可以被标记为过时。在那种情况下,重复步骤 302 和 304,直到达到非过时的高速缓冲存储器。

[0040] 在步骤 306 中,高速缓冲存储器搜索程序 114 存取该中央高速缓冲存储器,来了解它是否包含有关的数据。根据该中央高速缓冲存储器的结构的细节,步骤 306 的细节会有很大的变化。如图 3a ~ 3c 中所示的步骤 306 呈现了一项可能的实施,并且意在结合图 4a 和 4b 中所示的示范高速缓冲存储器数据结构来加以观察。

[0041] 图 4a 展示了中央高速缓冲存储器 400 的结构,中央高速缓冲存储器 400 可以结合步骤 306 中所展示的程序来使用。尺寸固定的头部 402 包含该过时的标志,并且,如果设置该标志,则参考更当前的中央高速缓冲存储器 400。接下来是尺寸固定的参考表格 404。参考表格 404 中的每个项目要么是对数据元素存储区 406 中的数据元素(见图 4b 中的 410)的参考,要么是特殊值(例如, NULL),该特殊值指出:该项目没有提到数据元素。在一些实施例中,这些参考作为整数偏移量来加以提供。这允许:在主机计算设备 102 重新启动之后,将中央高速缓冲存储器 400 作为磁盘上的文件来加以存储,并且对其进行恢复。

[0042] 实际的超高速缓存数据作为单独的数据元素 410 来加以存储。被分配给这些数据元素 406 的空间按需要变成不受约束空间 408 的区域。当用完不受约束空间 408 时,高速缓冲存储器服务程序 124 就该创建新的中央高速缓冲存储器 400 了。

[0043] 步骤 306 中所展示的特殊实施通过散列有关的数据,在步骤 308 中开始。然后,在步骤 310 中,使用该散列,作为进入中央高速缓冲存储器 400 中的参考表格 404 的索引。在图 3b 的步骤 312 中,检查参考表格 404 中的所选择的项目。如果那个项目没有提到任何数据元素,那么,有关的数据不在中央高速缓冲存储器 400 中(步骤 314)。然后,高速缓冲存储器搜索程序 114 进行到图 3d 中的步骤 326,以便在除中央高速缓冲存储器 400 以外的其他地方寻找有关的数据。

[0044] 另一方面,如果参考表格 404 中的所选择的项目提到数据元素,那么,那个数据元素可能或可能不包含有关的数据。其中的原因是:不同的数据值可能会散列到相同的值,并导致在参考表格 404 中选择相同的项目。

[0045] 查阅图 4b 中的示范数据元素结构 410。每个数据元素 410 始于参考 412, 参考 412 要么提到具有相同的散列值的另一个数据元素 410, 要么提到特殊值 (例如, NULL), 该特殊值指出: 该项目没有提到另一个数据元素 410。就利用参考表格 404 中的这些项目, 该参考可以作为整数偏移量来加以提供。参考 412 后面是数据标识符领域 414, 该数据标识符领域专门识别这个数据元素 410 中所包含的数据。最后是包含该数据本身的领域 416。这个领域 416 的尺寸和结构针对所存储的数据的性质。在一些实施例中, 数据元素 410 包含一个领域, 该领域指出其尺寸或 (同样) 特定元素数据区 416 的末端的位置。在其他实施例中, 中央高速缓冲存储器 400 的头部 402 中的领域指出所分配的数据存储区 406 的末端。无论如何, 使用这些长度字段来指出: 在哪里可以将另一个数据元素 410 加入高速缓冲存储器 400 (在图 5d 的步骤 534 中, 以下将加以讨论)。

[0046] 注意数据元素 410 的这个示范结构, 返回到图 3b 中的步骤 316。步骤 316 中的数据元素 410 的散列匹配有关数据的散列。为了了解这个数据元素 410 是否包含有关的数据, 可以将数据元素 410 的数据标识符领域 414 和有关的数据进行比较。如果它们匹配, 那么, 该搜索完成。有关的数据在步骤 318 中被加以检索, 并被传递到应用程序 A 112。这成功地结束了图 3a ~ 3d 中的高速缓冲存储器搜索程序 114。

[0047] 另一方面, 如果步骤 316 中的比较揭示这个数据元素 410 不包含有关的数据, 那么, 高速缓冲存储器搜索程序 114 进行到图 3c 中的步骤 320。在步骤 320 中, 检查数据元素 410 的参考域 412。如果这个参考域没有提到另一个数据元素 410, 那么, 中央高速缓冲存储器 400 不包含有关的数据 (步骤 322)。高速缓冲存储器搜索程序 114 进行到图 3d 中的步骤 326, 以便在其他地方搜索有关的数据。

[0048] 在步骤 320 中, 如果数据元素 410 的参考域 412 提到进一步的数据元素, 那么, 检查那个进一步的数据元素, 来了解它是否包含有关的数据。步骤 324 捕获这个过程: 通过重复步骤 316 ~ 322 来跟随数据元素链 410, 直到找到并检索到有关的数据 (步骤 318), 或者直达到该链的末端而没有找到有关的数据 (步骤 322)。

[0049] 如果中央高速缓冲存储器 400 不包含有关的数据, 那么, 在图 3d 的步骤 326 中, 高速缓冲存储器搜索程序 114 可以搜索对于应用程序 A 112 而言是本地的高速缓冲存储器 116。在一些实施中, 步骤 326 的细节监控步骤 306 的细节。这允许再使用该高速缓冲存储器搜索代码。无论如何, 如果本地高速缓冲存储器 116 包含有关的数据, 那么, 在步骤 332 中检索那些数据。如果不包含有关的数据, 并且假定没有其他的高速缓冲存储器搜索有关的数据, 那么, 在步骤 328 中创建那些数据。然后, 在步骤 330 中将那些数据加入本地高速缓冲存储器 116, 以促进将来的存取。可以在步骤 334 中请求将所创建的数据加入中央高速缓冲存储器 400。

[0050] 图 3a ~ 3d 展示了高速缓冲存储器搜索程序 114 如何从中央高速缓冲存储器 400 中或从其他地方检索有关的数据。图 5a ~ 5d 展示了高速缓冲存储器 400 的另一个方面: 高速缓冲存储器服务程序 124, 当目前的高速缓冲存储器变满时, 高速缓冲存储器服务程序 124 将数据加入高速缓冲存储器 400 并创建新的中央高速缓冲存储器。由于高速缓冲存储器服务程序 124 是将数据加入高速缓冲存储器 400 的唯一的例行程序, 因此, 不需要锁定高速缓冲存储器 400, 以防止多个复写器之中的冲突。也由于所有的写请求都流过高速缓冲存储器服务程序 124, 因此, 它可以收集有关高速缓冲存储器使用的统计数字并形成对将数据

加入高速缓冲存储器 400 的特殊请求的重要性的“全局”观点,并且,可以决定:当创建新的高速缓冲存储器时,应该传送哪些数据。图 5a ~ 5d 中的程序只是起示范的作用。它使用在图 4a 和 4b 中所介绍的各种中央高速缓冲存储器数据结构。

[0051] 高速缓冲存储器服务程序 124 始于图 5a 中的步骤 500,这时,它接收到将数据加入中央高速缓冲存储器 400 的请求。当高速缓冲存储器搜索程序 114 执行图 3d 中的步骤 334 时,这个请求可能已发生。

[0052] 在步骤 502 中,高速缓冲存储器服务程序 124 的一些实施例收集有关所接收的请求和有关其他请求的统计数字。在步骤 504 中,高速缓冲存储器服务程序 124 决定是否服从这个请求。一些实施例完全服从所有请求,而其他实施例在决定有关的数据是否“值得”被加入中央高速缓冲存储器 400 之前先参考所收集的统计数字。如果与高速缓冲存储器 400 中的不受约束空间 408 比较,有关数据的尺寸太大,则一些实施例拒绝该请求。将有关的数据和高速缓冲存储器 400 中的数据进行比较(可能通过执行图 3a ~ 3c 中的步骤 306),并且,如果已经存在有关的数据,则拒绝该请求。发生这种情况的原因是:排列在步骤 500 中所接收的请求,并且请求到达时与将有关的数据加入高速缓冲存储器 400 时(在图 5b ~ 5d 的步骤 522 中)之间会有重大的延迟。在该延迟期间,请求的高速缓冲存储器服务程序 114 或另一个例行程序会再存取高速缓冲存储器 400,无法找到有关的数据,并且再提出该请求。到处理稍后的请求的时候,依照较早的请求,有关的数据已经被加入队列 400。无论如何,如果在步骤 506 中决定不服从该请求,那么,高速缓冲存储器服务程序 124 在步骤 508 中等待进一步的请求。

[0053] 如果决定服从该请求,那么,在步骤 510 中,高速缓冲存储器服务程序 124 决定是否需要新的中央高速缓冲存储器 400。例如,如果有关的数据将不适合高速缓冲存储器 400 中的不受约束空间 408,那么,在图 5a 和 5b 的步骤 512 ~ 520 中创建新的高速缓冲存储器 400。如果参考表格 404 变满,或者,如果高速缓冲存储器 400 具有太多暂时还没有被使用的数据元素 410,则也可以创建新的高速缓冲存储器 400。如果不需要新的高速缓冲存储器 400,那么,高速缓冲存储器服务程序 124 在图 5b 的步骤 522 中直接开始加入有关的数据。

[0054] 如果需要新的中央高速缓冲存储器 400,则在步骤 512 中创建空核。高速缓冲存储器通常被存储在 RAM 中,以便它们可以提供对其数据的快速存取。一些操作系统允许 RAM 的区域由该文件系统来映射。这提供了某些优点,以后将对这些优点进行描述。

[0055] 在图 5b 的步骤 514 中,高速缓冲存储器服务程序 124 用从现存的高速缓冲存储器中选择的数据元素 410 来填充新近被创建的高速缓冲存储器 400。可以通过几种可能的方法来选择数据元素 410。高速缓冲存储器服务程序 124 可以收集有关已如何使用数据元素 410 的统计数字。然后,选择最新近使用的元素或最频繁使用的元素。一些实施例使用该选择中的随机性元素。无论如何,一旦选择数据元素 410,就将它们拷贝到新的中央高速缓冲存储器 400 的数据存储区 406 中,并且填充新的高速缓冲存储器 400 的参考表格 404。注意,步骤 514 会需要一些时间来加以执行,但是,在这个步骤期间,老的高速缓冲存储器总是可存取的。

[0056] 在老的高速缓冲存储器的头部 402 中,在步骤 516 中写入对新的高速缓冲存储器 400 的参考,并且,在步骤 518 中,设置标志,以便将老的高速缓冲存储器标记为“过时的”。现在,新的高速缓冲存储器 400 准备就绪,可以使用了。高速缓冲存储器搜索程序 114 一看

见老的高速缓冲存储器中的该过时标志,就遵循对新的高速缓冲存储器 400 的参考(见图 3a 中的步骤 302 和 304)。

[0057] 在步骤 520 中,高速缓冲存储器服务程序 124 要求:一没有提到该过时的高速缓冲存储器的应用程序,也就是说,所有这些应用程序一放弃其对该过 时的高速缓冲存储器的参考,该操作系统就自动删除该过时的高速缓冲存储器。

[0058] 不管是否刚刚创建新的高速缓冲存储器 400,都在步骤 522 中将有关的数据加入当前的高速缓冲存储器 400。图 5b ~ 5d 中的步骤 522 监控图 3a ~ 3c 中的步骤 306 的复杂化,因为步骤 522 创建在步骤 306 中被搜索的各种数据结构。附图中所提供的、关于步骤 306 和 522 的细节只说明了一些实施例,而在大多数实施例中,这些步骤彼此监控。

[0059] 步骤 522 的特殊实施例从在步骤 524 中散列有关的数据开始(图 3a 中的监控步骤 308)。在步骤 526 中,使用该散列,在高速缓冲存储器 400 的参考表格 404 中选择项目(监控步骤 310)。如果在步骤 528 中,所选择的项目还没有提到数据元素 410,那么,在图 5d 的步骤 534 中,在高速缓冲存储器 400 的数据存储区 406 中创建数据元素 410。然后,用有关的数据来填充这个新的数据元素。注意,当这个填充步骤(population step)进行时,存取高速缓冲存储器 400 的高速缓冲存储器搜索程序 114 将不会看见该新数据,所以可能会提出加入这些相同的数据的额外请求。一旦填充步骤 534 完成并且新的数据元素 410 准备就绪,高速缓冲存储器服务程序 124 就利用原子性来更新参考表格 404 中的所选择的指针,以指向新的数据元素 410(步骤 536)。该参考的写操作只占据一个计算机指令,以便它本质上是不可中断的。这样,在这项操作期间,不需要锁定高速缓冲存储器 400,以保持其内部的一致性。

[0060] 返回到步骤 528,如果参考表格 404 中的所选择的项目已经提到数据元素 410,那么,在步骤 530 中检查数据元素 410 的参考域 412。在步骤 532 中,遵循该参考链,直到它结束为止。(在图 3b ~ 3c 的步骤 316 ~ 324 中,这沿相同的参考链来监控该搜索。)一旦找到该链的末端,该高速缓冲存储器服务程序就进行到以上所描述的步骤 534,其中,新的数据元素 410 被定位并用有关的数据来填充。在这种情况下,步骤 536 利用原子性将对新的数据元素 410 的参考加入曾经处于该链末端的现存的数据元素 410,这样,将该参考链延长又一个“链环”。

[0061] 总而言之,无论是将新的数据元素 410 加入现存的中央高速缓冲存储器 400,还是创建新的中央高速缓冲存储器 400,高速缓冲存储器服务程序 124 都在两个阶段中进行。首先,准备该数据,并且完成所有耗时的的工作。在这个阶段期间,对于高速缓冲存储器搜索程序 114 而言,中央高速缓冲存储器 400 的这些变化不可见。所以,不需要锁定中央高速缓冲存储器 400,来保持其内部的一致性。也不需要按高优先级来运行高速缓冲存储器服务程序 124。第二个阶段包括:将单一指针写入新的数据元素 410 或写入新的中央高速缓冲存储器 400。写入单一指针是本质上不可中断的程序,因此,又不需要在此阶段期间锁定中央高速缓冲存储器 400。一旦第二个阶段完成,高速缓冲存储器搜索程序 114 就可以存取这个新的数据或新的高速缓冲存储器。

[0062] 图 6 表现了一种用于监控和配置高速缓冲存储器服务程序 124 的操作的方法。在步骤 600 中,显示了该中央高速缓冲存储器的最大尺寸和当前尺寸。在步骤 602 中,显示了当填充新的中央高速缓冲存储器时在图 5b 的步骤 514 中被用来从现存的中央高速缓冲存

存储器中选择数据元素的试探法。在步骤 604 中,显示了在步骤 504 中被用来决定是否服从将数据加入该中央高速缓冲存储器的请求的试探法。记住:这种试探法可以只是“始终服从”。步骤 606 显示了已被搜集的、关于高速缓冲存储器使用的各种统计数字。网络管理员可以分析这些统计数字,并决定该中央高速缓冲存储器没有正在最佳地进行操作。该网络管理员在步骤 608 中输入某种变化,该变化在步骤 610 中被反映回来。图 6 意在仅仅提供用户界面的风味,而无法公平对待该技术领域中众所周知的众多不同的高速缓冲存储器分析工具。

[0063] 鉴于可以应用本发明的原理的这许多可能的实施例,应该认识到:这里根据附图而描述的实施例意在只是起说明性的作用,而不应该被视作限制本发明的范围。例如,精通该技术领域的人将会认识到:在不脱离本发明的精神的前提下,可以在布置和细节方面修改所展示的实施例,尤其是基于这些实施例的各种数据结构和程序。虽然按照软件模块或部件来描述本发明,但是,精通该技术领域的人将会认识到:这样的软件模块或部件可以同等地由硬件部件来取代。所以,如这里所描述的本发明计划可将所有这类实施例包括在以下权利要求书及其相等物的范围以内。

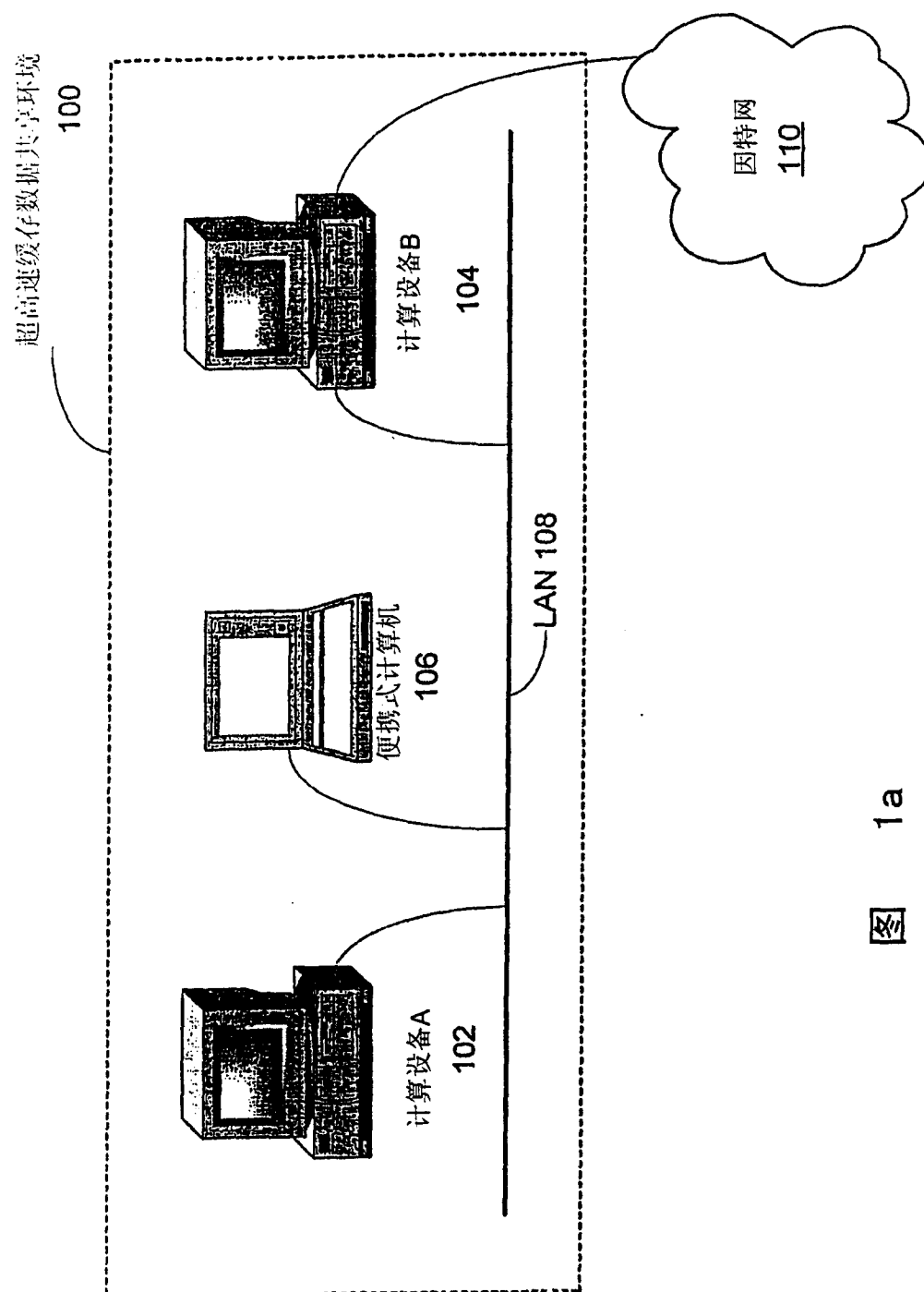


图 1a

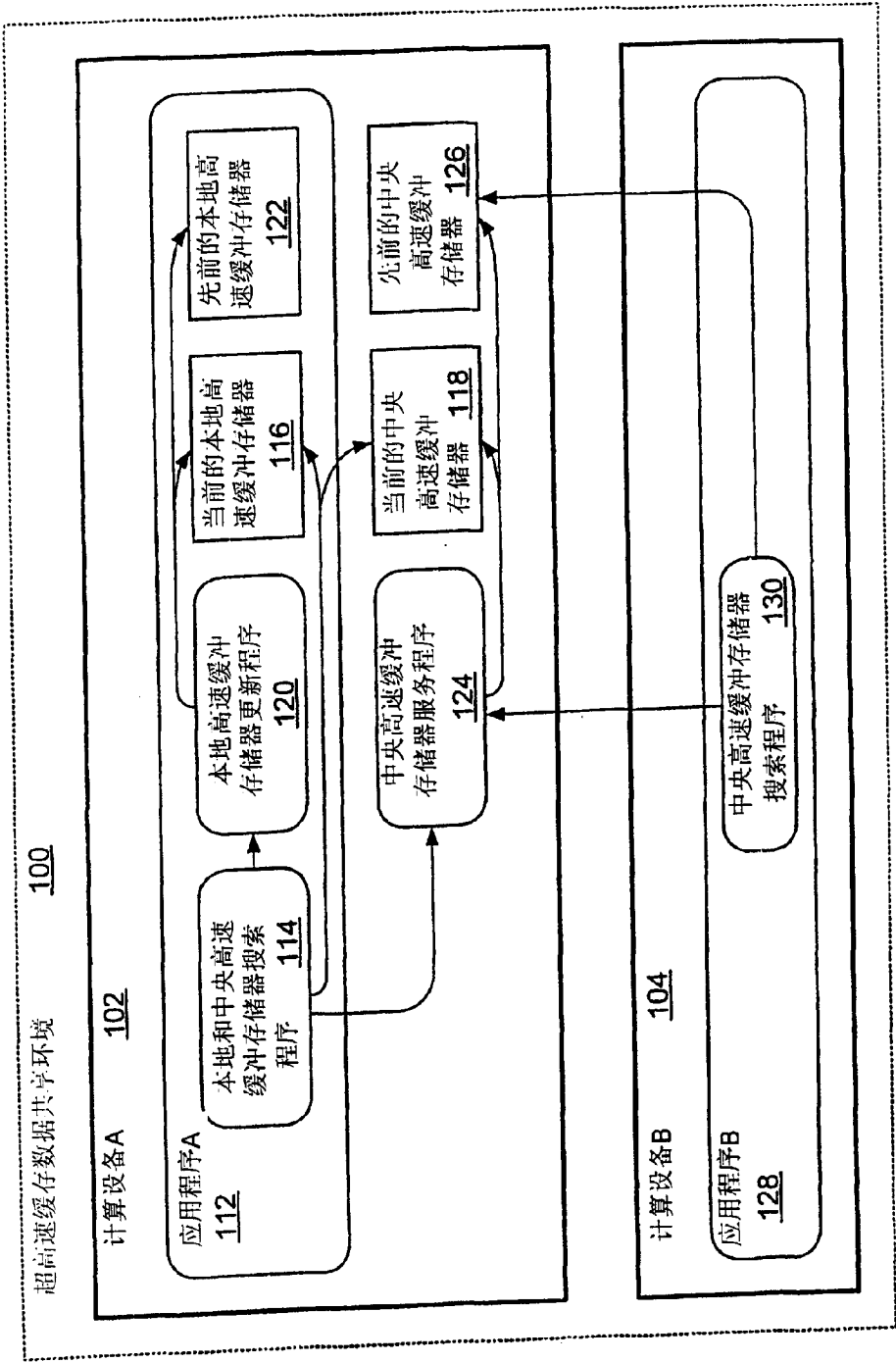


图 1b

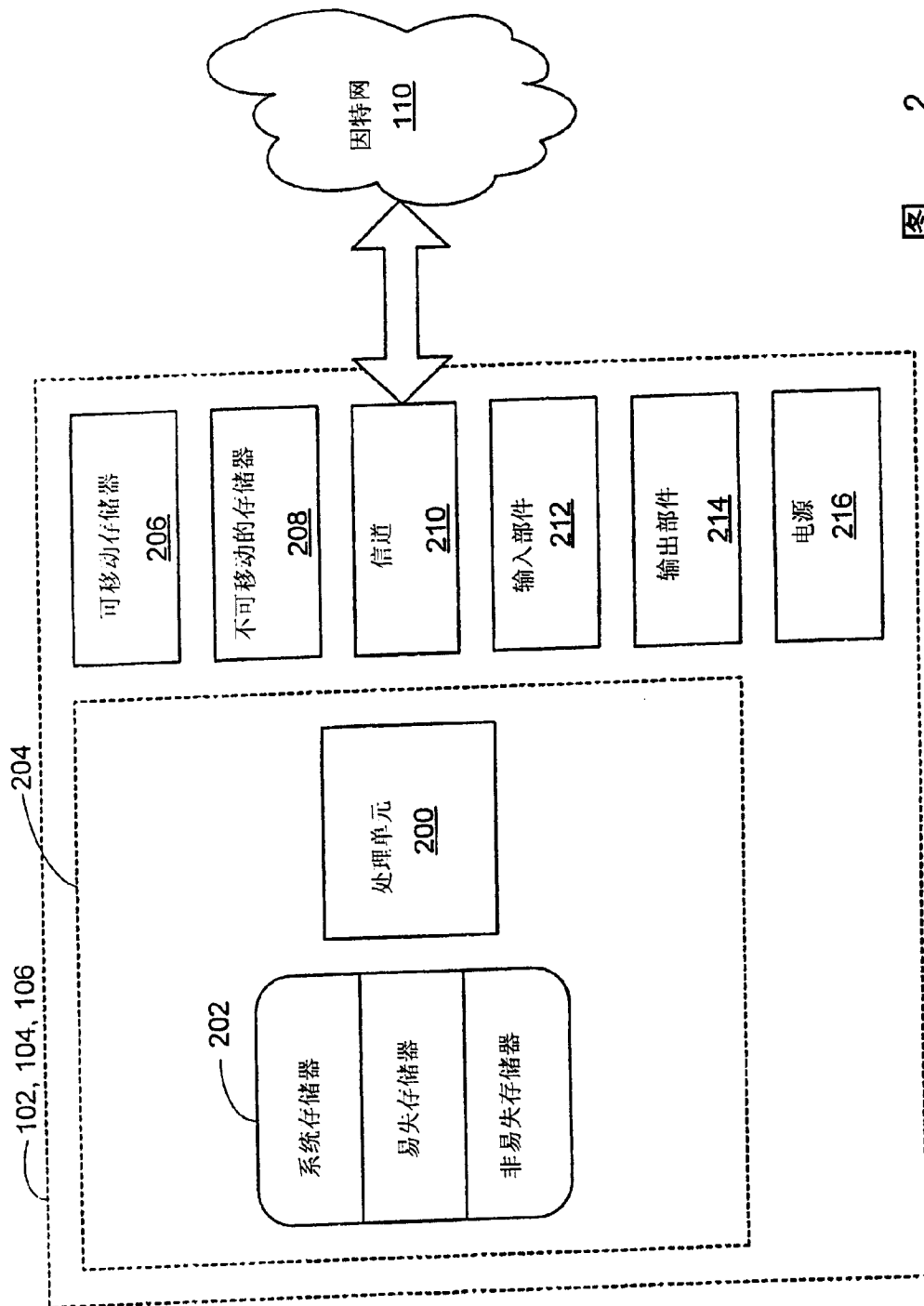


图 2

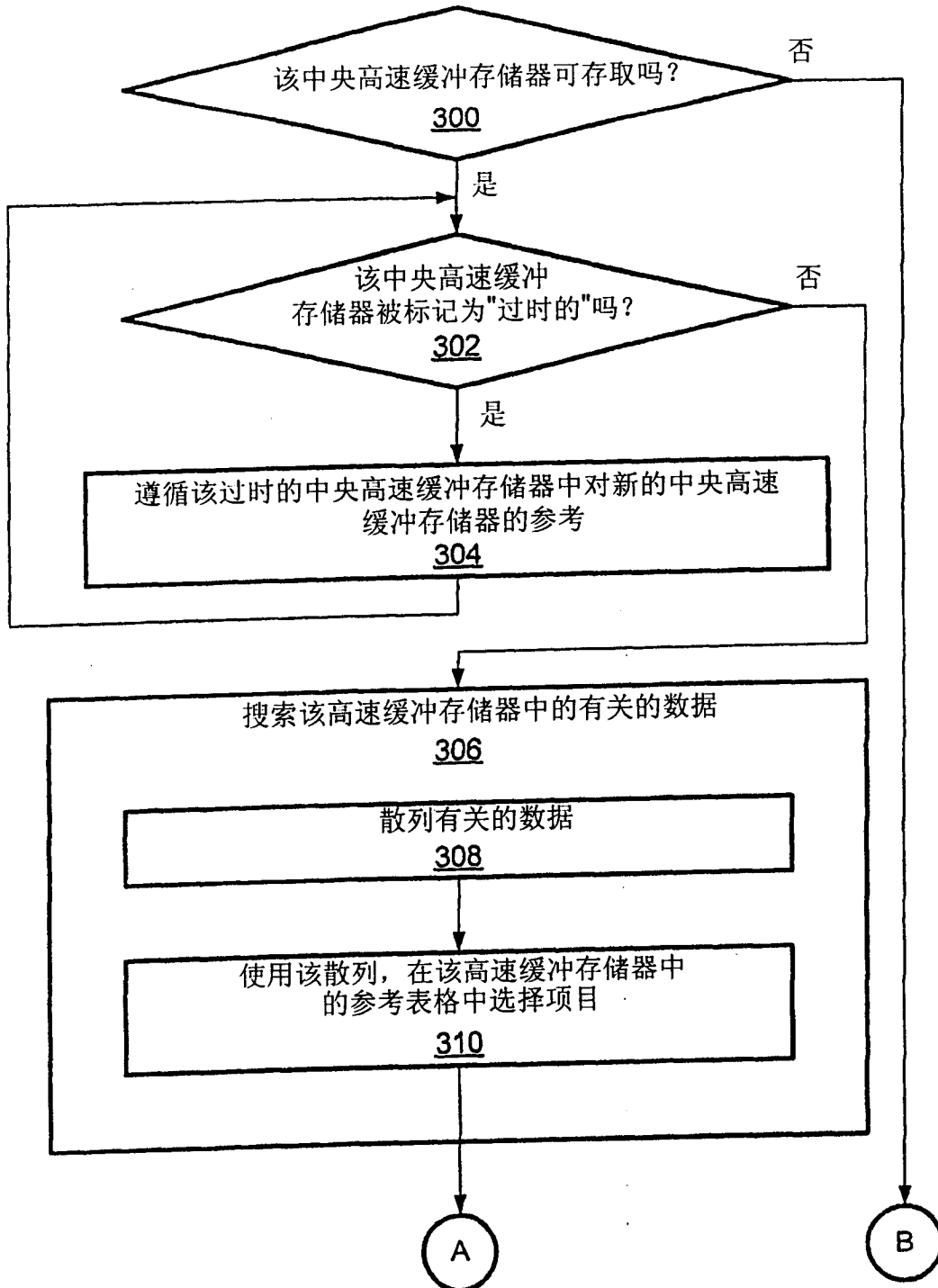


图 3a

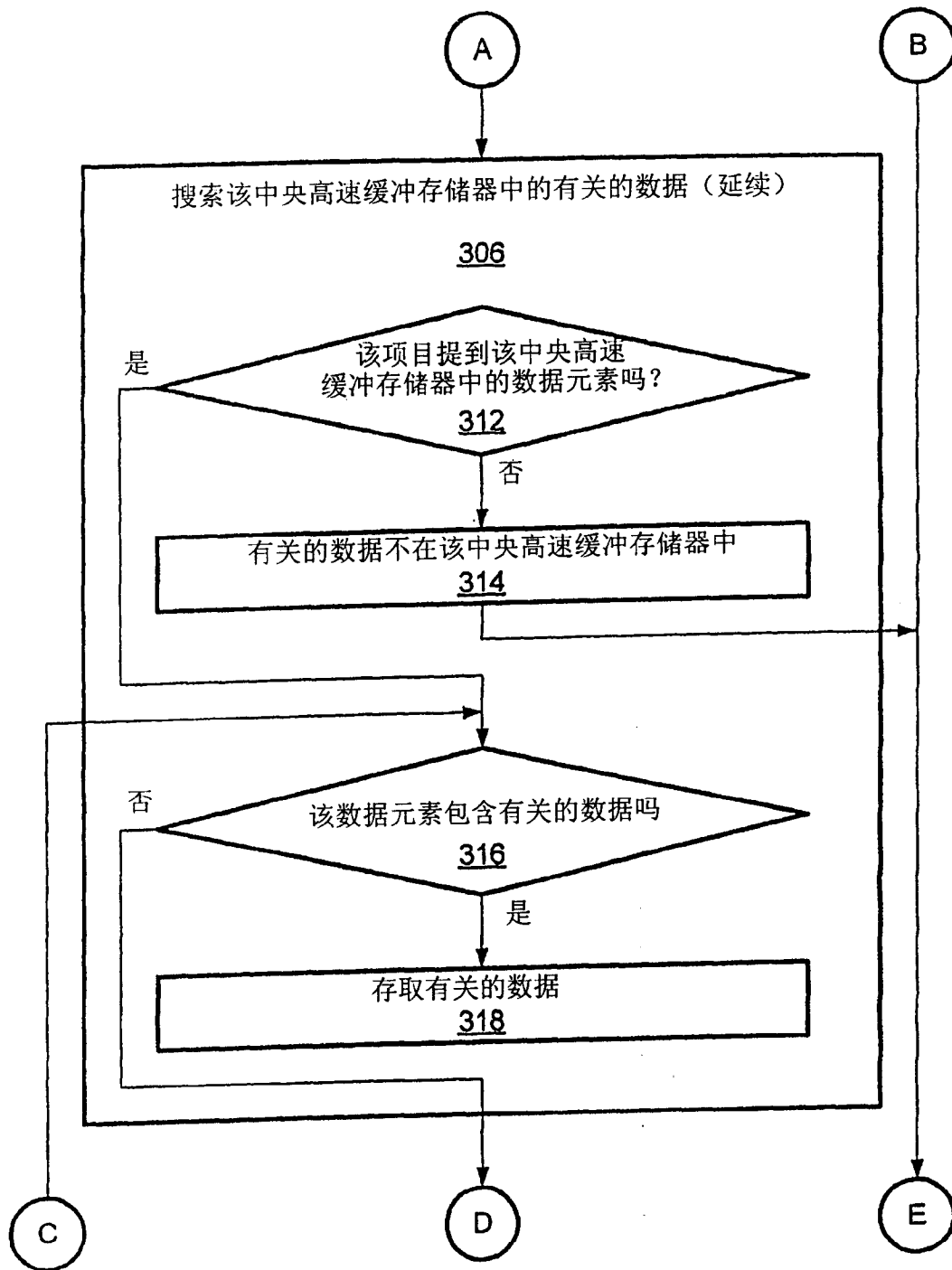


图 3b

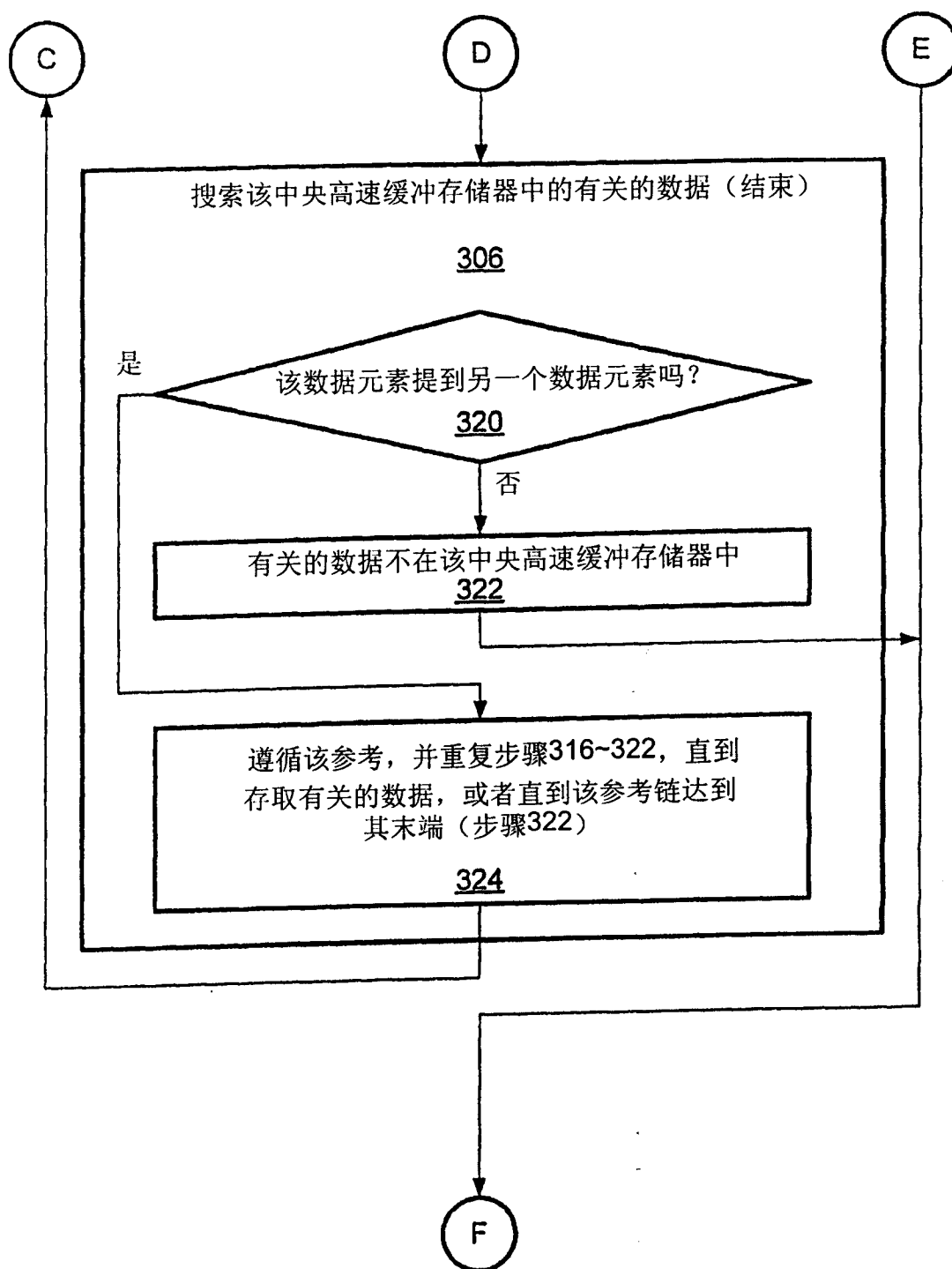


图 3c

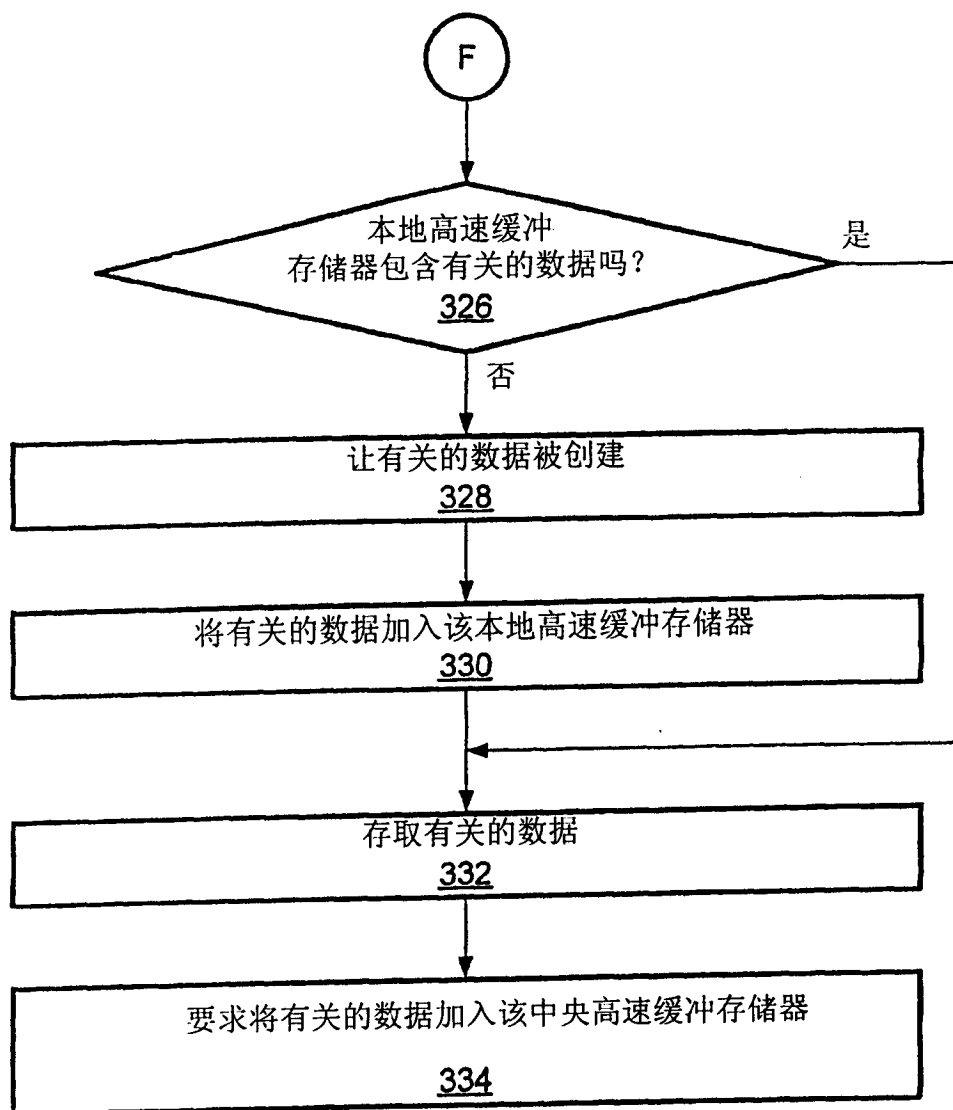


图 3d

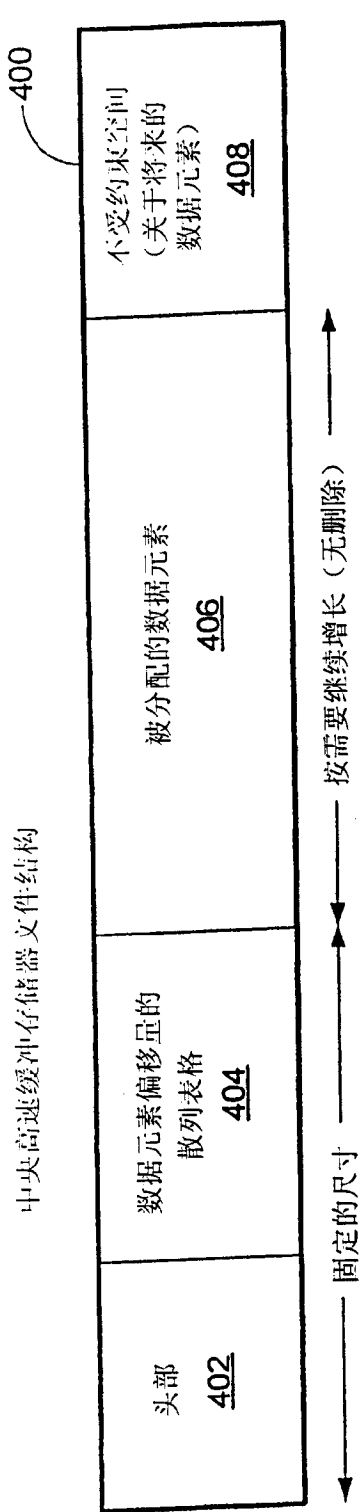


图 4a

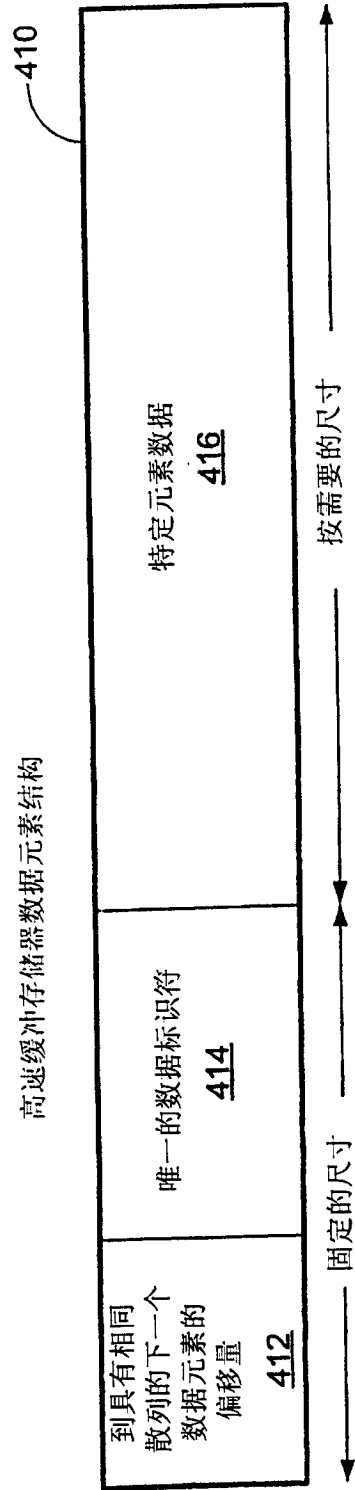


图 4b

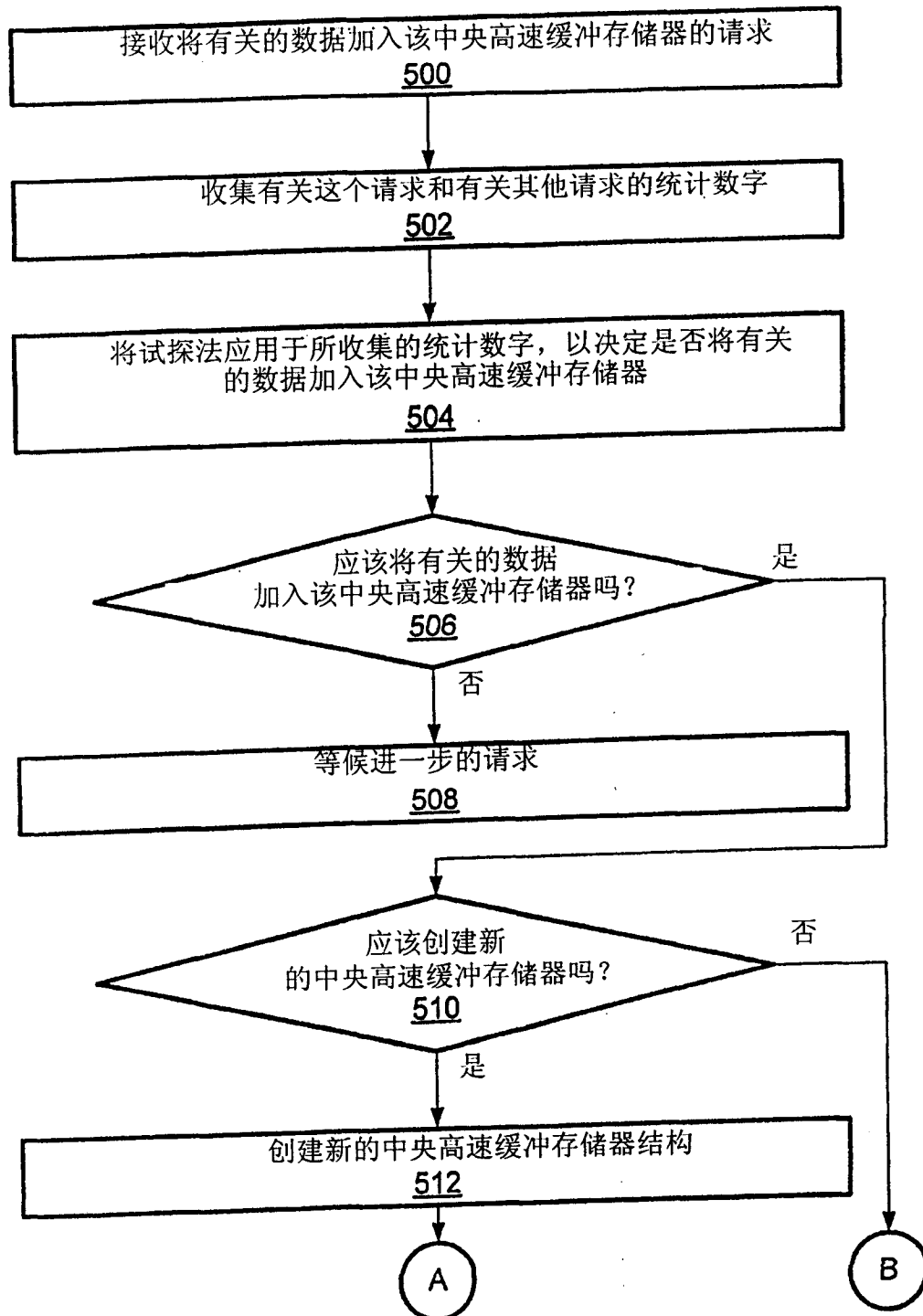


图 5a

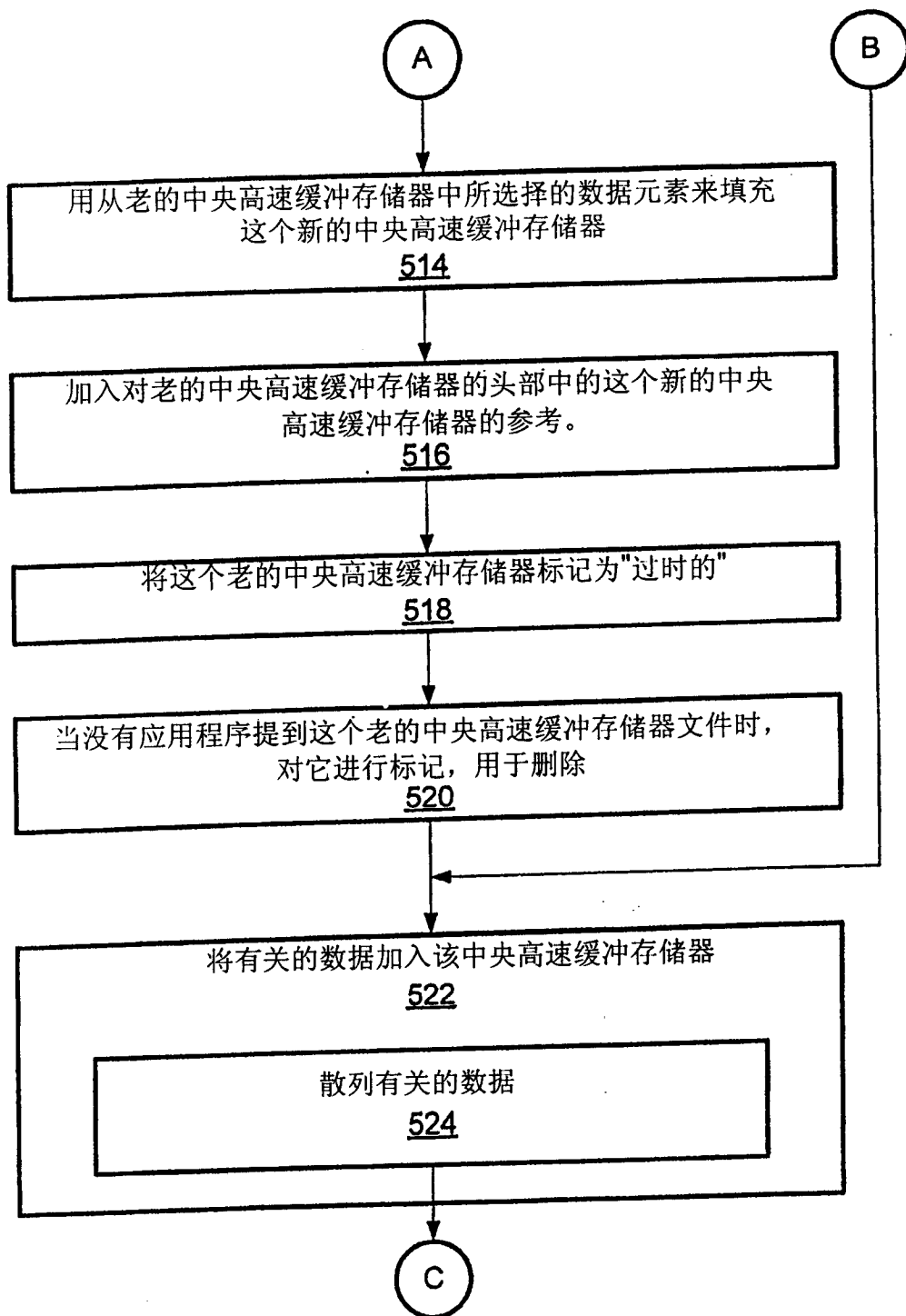


图 5b

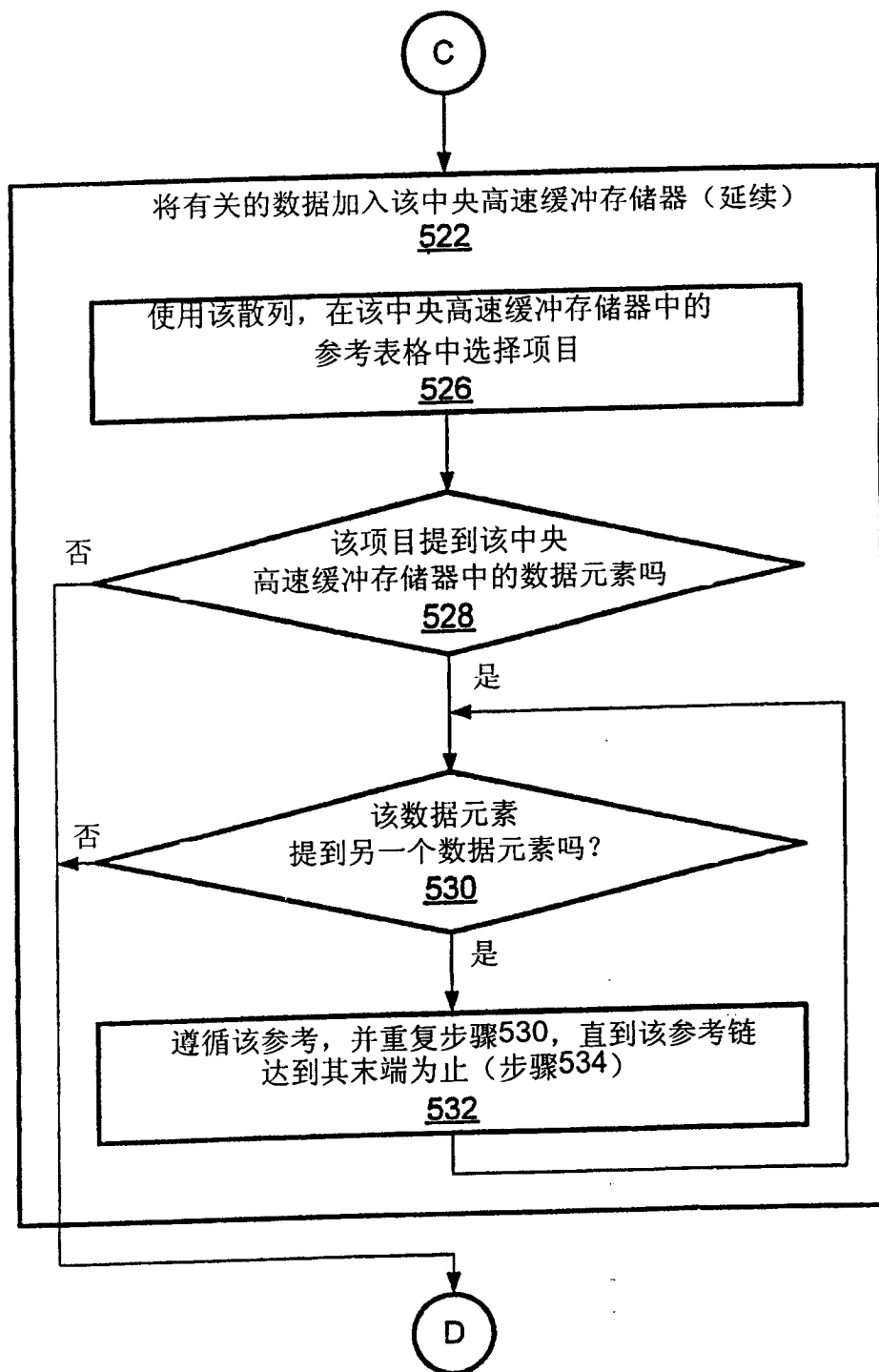


图 5c

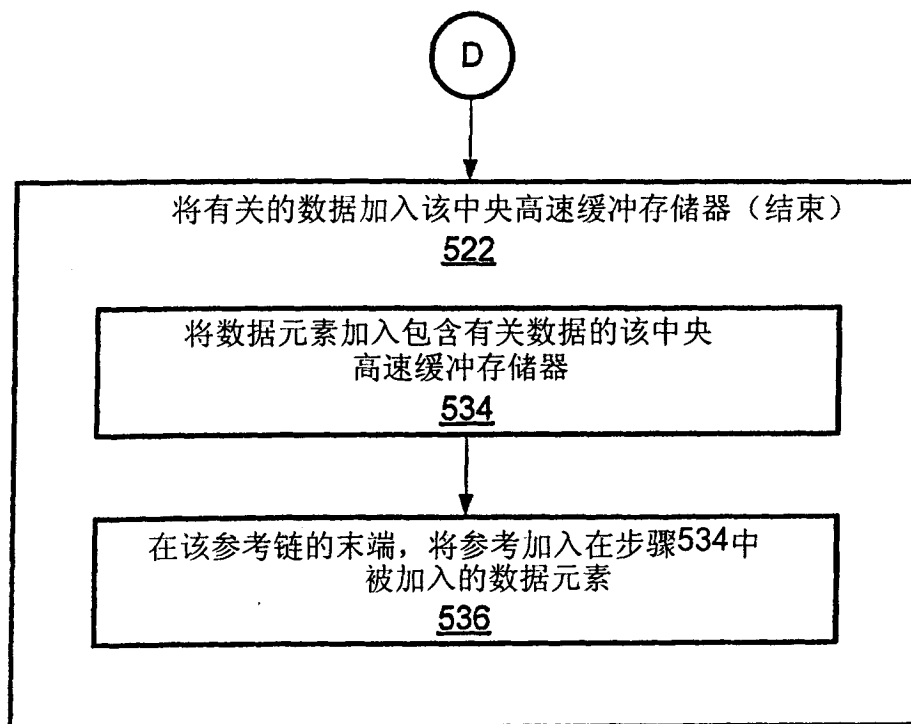


图 5d

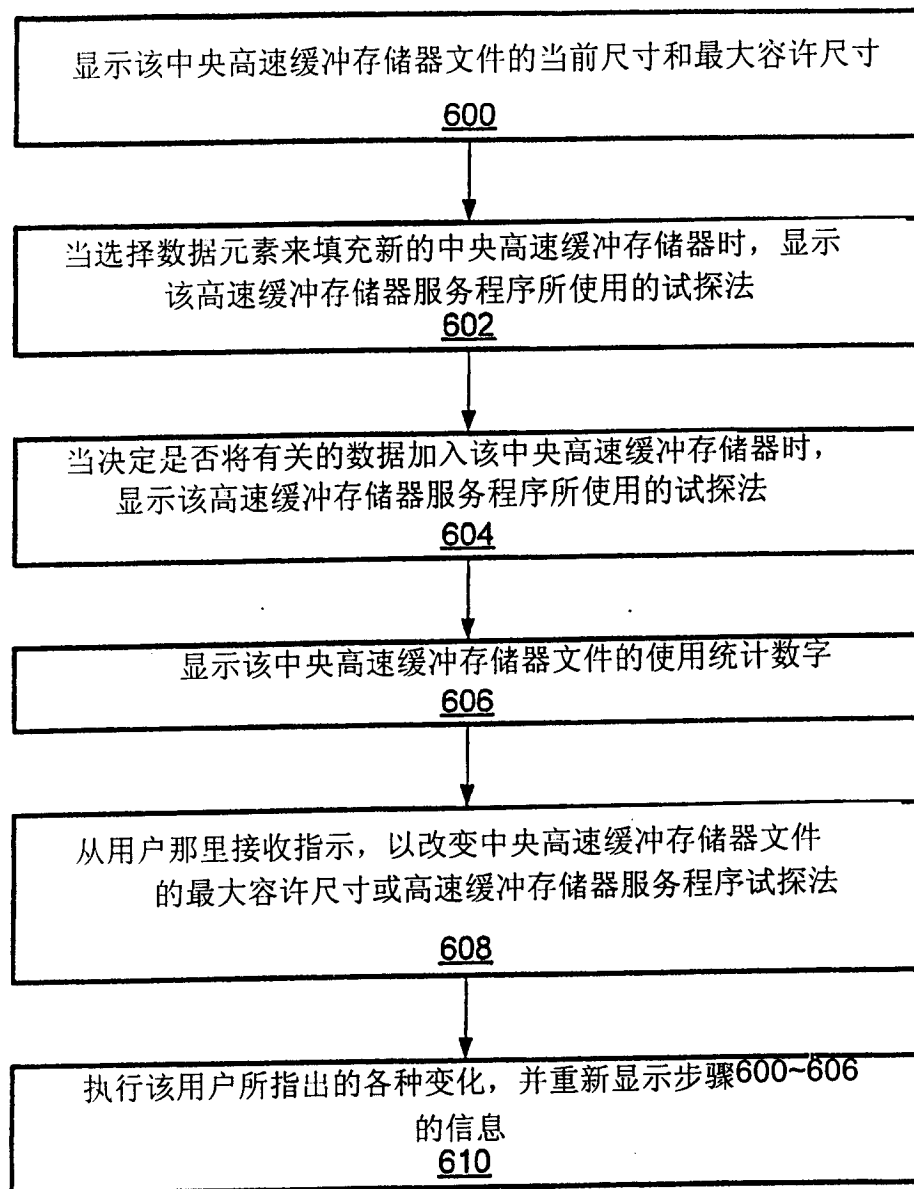


图 6