

(12) 发明专利申请

(10) 申请公布号 CN 102782670 A

(43) 申请公布日 2012. 11. 14

(21) 申请号 201180013647. 2

(51) Int. Cl.

(22) 申请日 2011. 03. 09

G06F 15/16(2006. 01)

(30) 优先权数据

G06F 13/16(2006. 01)

12/723, 153 2010. 03. 12 US

G06F 17/40(2006. 01)

(85) PCT申请进入国家阶段日

2012. 09. 12

(86) PCT申请的申请数据

PCT/US2011/027791 2011. 03. 09

(87) PCT申请的公布数据

W02011/112750 EN 2011. 09. 15

(71) 申请人 微软公司

地址 美国华盛顿州

(72) 发明人 V. 拉文希 B. J. 巴罗斯 M. J. 麦肯

H. 戈迪尔 陈新光 O. V. 奥巴桑乔

P. R. C. 明

(74) 专利代理机构 中国专利代理(香港)有限公司

72001

代理人 李舒 汪扬

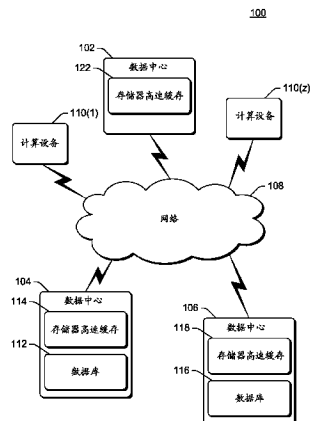
权利要求书 2 页 说明书 14 页 附图 6 页

(54) 发明名称

存储器高速缓存数据中心

(57) 摘要

一种数据中心系统包括耦合到数据中心控制器的存储器高速缓存。存储器高速缓存包括易失性存储器并且存储数据,该数据持久存储于远离数据中心系统设置的不同数据中心系统中的数据库中而不是第一数据中心系统中。数据中心控制器从存储器高速缓存读取数据并且向存储器高速缓存写入数据。



1. 一种在第一数据中心中实施的方法,所述方法包括:

在所述第一数据中心的存储器高速缓存中维护数据,所述第一数据中心是存储器高速缓存数据中心,其中所述数据持久存储于远离第一数据中心设置的第二数据中心中的数据库中,并且其中所述存储器高速缓存包括易失性存储器;

响应于在所述第一数据中心接收的针对数据的请求,如果所述数据在所述存储器高速缓存中,则从所述存储器高速缓存获取所请求的数据,否则从远离第一数据中心设置的不同数据中心获得所请求的数据并且向所述存储器高速缓存添加所请求的数据;

向其接收所述请求的请求者返回所请求的数据;并且

响应于在所述第一数据中心接收的对在所述第一数据中心中存储附加数据的请求,在已经向所述第二数据中心中的所述数据库添加所述附加数据之后在所述存储器高速缓存中存储所述附加数据。

2. 如权利要求 1 所述的方法,其中维护所述数据包括当在所述第一数据中心中无持久存储所述数据的数据库时维护所述数据。

3. 如权利要求 1 所述的方法,其中所述存储器高速缓存包括多个存储器组,并且其中所述维护包括在所述多个存储器组中的每个存储器组中存储所述数据的副本。

4. 如权利要求 3 所述的方法,其中所述多个存储器组中的第一存储器组包括易失性存储器,并且所述多个存储器组中的第二存储器组包括一个或者多个盘驱动。

5. 如权利要求 4 所述的方法,其中所述一个或者多个盘驱动包括一个或者多个固态驱动。

6. 如权利要求 4 所述的方法,还包括:在易失性存储器中的日志中存储所述附加数据,然后随后在所述日志中写入从多个请求接收的数据以在所述一个或者多个盘驱动上依次存储附加数据。

7. 如权利要求 4 所述的方法,其中从所述存储器高速缓存获取所请求的数据包括:在尝试从所述第二存储器组获取所请求的数据之前尝试从所述第一存储器组获取所请求的数据。

8. 如权利要求 3 所述的方法,其中所述多个存储器组中的每个存储器组包括多个服务器计算机中的不同服务器计算机。

9. 如权利要求 3 所述的方法,其中获取所请求的数据包括:

访问所述多个存储器组中的第一存储器组以获取所请求的数据;

如果所请求的数据在所述第一存储器组中,则从所述第一存储器组获取所请求的数据;并且

如果所请求的数据未在所述第一存储器组中,则从所述多个存储器组中的第二存储器组获取所请求的数据并且向所述第一存储器组添加所请求的数据。

10. 如权利要求 3 所述的方法,其中在所述存储器高速缓存中维护数据包括:

比较所述多个存储器组中的第一存储器组中的第一数据集与所述多个存储器组中的第二存储器组中的第二数据集;

向所述第二存储器组复制在所述第一数据集中、未在所述第二数据集中的数据;并且

向所述第一存储器组复制在所述第二数据集中、未在所述第一数据集中的数据。

11. 如权利要求 1 所述的方法,其中所述不同数据中心是第二数据中心,并且其中从

所述不同数据中心获得所请求的数据包括：从所述第二数据中心的数据库获得所请求的数据。

12. 如权利要求 1 所述的方法,还包括：将所述附加数据与来自多个附加请求的数据一起分批以将附加数据存储到单个组合的存储请求中；并且向所述第二数据库发送所述单个组合的存储请求。

13. 如权利要求 1 所述的方法,还包括：

从所述第二数据中心接收新数据,其中响应于对在所述第二数据中心中的所述数据库中存储的数据的更新从所述第二数据中心接收所述新数据；

校验接收的新数据的先前版本是否已经存储于所述存储器高速缓存中；并且

仅在接收的新数据的先前版本已经存储于所述存储器高速缓存中的情况下,才在所述存储器高速缓存中存储所接收的新数据。

14. 一种第一存储器高速缓存数据中心,包括：

存储数据的存储器高速缓存,其中所述数据持久存储于第二数据中心中的数据库中而不是所述第一存储器高速缓存数据中,其中所述第二数据中心远离所述第一存储器高速缓存数据中心设置,并且其中所述存储器高速缓存包括易失性存储器；以及

数据中心控制器,其耦合到所述存储器高速缓存,其中所述数据中心控制器从所述存储器高速缓存读取数据并且向所述存储器高速缓存写入数据。

15. 如权利要求 14 所述的第一存储器高速缓存数据中心,其中所述存储器高速缓存包括多个存储器组,并且其中所述数据中心控制器用于在多个存储器组的每个存储器组中存储所述数据的副本。

## 存储器高速缓存数据中心

### 背景技术

[0001] 向全世界用户提供各种功能的联网服务是可用的。假如有可以访问这样的服务的大量用户,可以构建大型数据中心,这些数据中心包括用于为这些服务的各种用户存储信息的数据库。然而这样的大型数据中心并非没有它们的问题。一个这样的问题在于随着用户数目增加而变得难以让单个数据中心管理服务的所有用户。另一这样的问题在于一些用户可能位于远离数据中心处,从而给这样的用户增加网络延时。

### 发明内容

[0002] 提供这一发明内容以简化形式介绍下文在具体实施方式中进一步描述的概念选集。这一发明内容并非旨在标识要求保护的的主题的关键特征或者必要特征,它也并非旨在用来限制要求保护的的主题的范围。

[0003] 根据一个或者多个方面,在第一数据中心的存储器高速缓存中维护数据,存储器高速缓存包括易失性存储器。数据持久存储于位于远离第一数据中心的第二数据中心中的数据库中。响应于在第一数据中心接收的针对数据的请求,如果数据在存储器高速缓存中,则从存储器高速缓存获取被请求的数据并且将其返回给从其接收请求的请求者。如果数据未在存储器高速缓存中,则从位于远离第一数据中心的的不同数据中心获得数据、将其添加到存储器高速缓存并且返回给请求者。响应于在第一数据中心接收的对在第一数据中心中存储附加数据的请求,在已经向第二数据中心中的数据库添加附加数据之后,在存储器高速缓存中存储附加数据。

[0004] 根据一个或者多个方面,第一数据中心系统可以包括耦合到数据中心控制器的存储器高速缓存。存储器高速缓存可以包括易失性存储器并且存储数据,该数据持久存储于第二数据中心系统中的数据库中而不是第一数据中心系统中。这一第二数据中心系统位于远离第一数据中心系统之处。数据中心控制器从存储器高速缓存读取数据并且向存储器高速缓存写入数据。

### 附图说明

[0005] 相同编号在附图全篇中用来指代相似特征。

[0006] 图 1 图示了实施根据一个或者多个实施例的存储器高速缓存数据中心的例子系统。

[0007] 图 2 图示了根据一个或者多个实施例的例子存储器高速缓存数据中心。

[0008] 图 3 是图示了根据一个或者多个实施例的用于操作存储器高速缓存数据中心的例子过程的流程图。

[0009] 图 4 是图示了根据一个或者多个实施例的用于从存储器高速缓存数据中心读取数据的例子过程的流程图。

[0010] 图 5 是图示了根据一个或者多个实施例的用于向存储器高速缓存数据中心写入数据的例子过程的流程图。

[0011] 图 6 图示了例子计算设备,该计算设备可以被配置成实施根据一个或者多个实施例的存储器高速缓存数据中心的方面。

### 具体实施方式

[0012] 这里讨论存储器高速缓存数据中心。使用多个数据中心来使服务可用于跨越宽泛地理范围的用户。这些多个数据中心中的一个或者多个数据中心包括数据库,该数据库是服务维护的各种数据的持久存储库。此外,这些多个数据中心中的一个或者多个数据中心是包括存储器高速缓存的存储器高速缓存数据中心,该存储器高速缓存存储服务维护的数据中的至少一些数据。然而存储器高速缓存数据中心未包括用于服务的数据库;实际上,存储器高速缓存数据中心与包括数据库的另一数据中心通信以使数据存储于持久存储库中。虽然存储器高速缓存数据中心可以出于其它原因而包括数据库,但是根据这里讨论的技术的存储器高速缓存数据中心未包括作为服务维护的各种数据的持久存储库的数据库。另外,这些多个数据中心中的一个或者多个数据中心可以包括数据库(该数据库是数据的持久存储库)和存储器高速缓存。

[0013] 图 1 图示了例子系统 100,该系统实施根据一个或者多个实施例的存储器高速缓存数据中心。系统 100 包括可以经由网络 108 相互通信的多个数据中心 102、104 和 106。网络 108 可以是多种不同网络(包括因特网、局域网(LAN)、蜂窝或者其它电话网络、内部网、其它公用和 / 或专用网络、其组合等)。

[0014] 数据中心 102、104 和 106 操作用于向跨越宽泛地理范围的多个( $z$  个)计算设备 110 提供服务。计算设备 110 可以处于全世界各种地区、国家和大陆并且经由网络 108 访问数据中心 102、104 和 106。数据中心 102、104 和 106 通常处于不同地理位置并且各自通常对来自地理上与数据中心最近的计算设备 110 的请求做出响应。取而代之,数据中心可以对来自以除了地理距离之外的方式确定(比如基于网络延时、基于网络带宽等确定)为与数据中心最近的计算设备 110 的请求做出响应。数据中心 102、104 和 106 可以由多种不同类型的计算设备 110(比如桌面型计算机、膝上型计算机、移动站、娱乐装置、电视、以通信方式耦合到显示设备的机顶盒、蜂窝或者其它无线电话、游戏控制台、汽车计算机等)访问。

[0015] 数据中心 102、104 和 106 可以向计算设备 110 提供多种不同服务中的一个或者多个服务。例如数据中心 102、104 和 106 可以提供社交联网服务、电子邮件服务、消息接发服务、图像和 / 或视频共享服务、游戏或者其它娱乐服务等中的一个或者多个服务。数据中心 102、104 和 106 维护各种数据作为提供的部分。维护的特定数据可以基于特定服务(比如用于用户的账户信息、用户发送和 / 或接收的电子邮件、用户上传的图像等)而变化。

[0016] 每个数据中心 102、104 和 106 包括操作用于提供一个或者多个服务的一个或者多个服务器计算机。与计算机设备 110 的上文讨论相似,这些服务器计算机可以是多种不同类型的计算设备。每个数据中心 102、104 和 106 通常包括多个服务器计算机(例如大约数以十计、数以百计或者甚至更多服务器计算机)。数据中心 102、104 和 106 也称为服务器群(farm)。然而应当注意任何数目的数据中心 102、104 和 106 可以包括小数目的服务器计算机(例如少于十个服务器计算机)。

[0017] 数据中心 104 包括数据库 112 和存储器高速缓存 114,而数据中心 106 包括数据库 116 和存储器高速缓存 118。数据库 112 和 116 是数据的持久存储库并且可以使用多种不同

数据库技术(例如包括关系或者非关系数据库、大规模非关系存储库等)来实施。例如可以使用 Microsoft SQL Server® 数据库或者 Microsoft Azure™ 表服务来实施数据库 112 和 116。数据库 112 和 116 是用于数据(该数据由数据中心 102、104 和 106 提供的服务维护)的持久存储库并且可以是用于数据(该数据由数据中心 102、104 和 106 提供的服务维护)的权威存储库。数据库 112 和 116 中的不同数据库可以存储不同数据(例如用于不同用户的数据)或者取而代之可以存储相同数据中的至少一些数据。

[0018] 存储器高速缓存 114 和 118 是存储于数据库 112 和 116 中的数据的高速缓存。存储器高速缓存 114 和 118 通常是比数据库 112 和 116 更小并且存储来自数据库 112 和 116 的一些数据的易失性存储设备(例如随机存取存储器(RAM))。通常可以比数据库 112 和 116 更快地访问存储器高速缓存 114 和 118,因此存储器高速缓存 114 和 118 通常存储新近已经访问过的、或者以别的方式被认为更可能在不久的将来被再次访问的数据(因此可以被数据中心 104 和 106 更快访问)。

[0019] 另一方面,数据中心 102 包括存储器高速缓存 122。存储器高速缓存 122 与存储器高速缓存 114 和 118 相似之处在于存储器高速缓存 122 存储已经新近访问的、或者以别的方式被认为更可能在不久的将来再次被访问的数据。然而数据中心 102 未包括作为存储于数据中心 102、104 和 106 中的数据的持久存储库的数据库。数据中心 102 与用于持久存储数据的远程数据中心(数据中心 104 或 106 之一)而不是拥有数据中心 102 本地的数据库。由于数据中心 102 无数据库,所以数据中心 102 也称为存储器高速缓存数据中心。应当注意存储器高速缓存数据中心未包括数据库(该数据库是针对由数据中心 102、104 和 106 提供的服务维护的数据的持久存储库),但是可以出于其它原因而可选地包括维护其它数据的数据库。

[0020] 图示了系统 100 包括两个数据中心 104 和 106,这些数据中心包括存储器高速缓存和数据库二者。然而系统 100 可以包括任何数目的数据中心,每个数据中心包括存储器高速缓存和数据库二者。类似地,虽然系统 100 可以包括任何数目的存储器高速缓存数据中心,但是在系统 100 中图示了存储器高速缓存数据中心 102。此外,系统 100 还可以包括任何数目的数据中心,这些数据中心包括数据库、但是未包括存储器高速缓存。

[0021] 图 2 图示了根据一个或者多个实施例的示例存储器高速缓存数据中心 200。数据中心 200 是未包括本地数据库的数据中心的例子,该本地数据库是针对由数据中心(例如图 1 的数据中心 102)提供的服务维护的数据的持久存储库。存储于数据中心 200 中的数据持久存储于远程数据中心(例如图 1 的数据中心 104 或者 106 之一)的数据库中而不是包括本地数据库。在数据中心 200 中无这样的数据库通过减少数据中心 200 中包括的硬件数量来降低数据中心 200 的成本。另外可以由于在数据中心 200 中无这样的数据库而减少在数据中心之间的通信量,因为数据中心 200 未添加需要与其它数据中心中的其它数据库同步的这样的数据库。

[0022] 数据中心 200 包括数据中心控制器 202 和存储器高速缓存 204。数据中心控制器 202 从计算设备接收针对数据的请求并且响应于这样的请求来获取和返回请求的数据。如下文更具体讨论的那样,可以从存储器高速缓存 204 或者另一数据中心获取这一数据。数据中心控制器 202 也从计算设备接收对存储数据的请求并且响应于这样的请求在存储器高速缓存 204 中存储接收的数据。如下文更具体讨论的那样,也可以向另一数据中心传达

接收的数据。数据中心控制器 202 也在维护存储器高速缓存 204 时执行附加操作(如下文更具体讨论的那样,比如向存储器高速缓存 204 的各种部分中回填数据)。

[0023] 应当注意数据中心 200 可以是包括大量(例如数以百计或者数以千计)服务器计算机的数据中心或者取而代之包括少量服务器计算机的数据中心。包括少量服务器计算机的数据中心可能通常未视为“数据中心”并且可以例如是经由多种不同网络或者耦合技术耦合在一起的多个服务器计算机的汇集、在可以耦合到网络的小外罩(例如存储 2-5 个服务器计算机)中包括的多个服务器计算机的汇集等。

[0024] 数据中心控制器 202 由多个( $m$  个)服务器计算机 206(1)、 $\dots$ 、206( $m$ ) 实施。虽然在图 2 中图示了多个服务器计算机,但是取而代之,数据中心控制器 202 可以由单个服务器计算机实施。每个服务器计算机可以是如上文讨论的多种不同类型的计算设备。

[0025] 存储器高速缓存 204 包括两个或者更多( $n$  个)存储器组 212(1)、 $\dots$ 、212( $n$ )。每个存储器组 212 包括多个服务器计算机 214、216。存储器组 212 内存储特定数据的特定服务器计算机可以由数据中心控制器 202 或者取而代之由存储器组 212 的服务器计算机确定。存储器组 212 内存储特定数据的特定服务器计算机可以用不同方式来确定,并且可以至少部分依赖于数据中心 200 提供的特定服务。例如与特定用户关联的数据可以存储于特定服务器计算机上,与特定日期或者类别关联的数据可以存储于特定服务器计算机上,等等。

[0026] 不同存储器组 212 可以包括相同数目或者取而代之不同数目的服务器计算机。每个服务器计算机 214、216 通常包括易失性存储介质(比如 RAM)和非易失性存储介质(比如只读存储器(ROM)、闪存、光盘、磁盘等)。用于服务器计算机的操作的各种数据和指令可以存储于易失性存储介质和 / 或非易失性存储介质中。然而在一个或者多个实施例中,数据中心 200 存储的用于提供数据中心 200 的服务的数据存储于易失性存储介质中。

[0027] 存储器组 212 作为彼此的备用存储介质来操作,存储相同数据的重复副本。可以响应于多种事件(比如由于服务器计算机的故障、由于安装附加软件等)使存储器组 212 中的一个或者多个服务器计算机掉电。由于数据中心 200 存储的数据存储于易失性存储介质中,所以当使存储器组 212 中的服务器计算机掉电时,存储于该服务器计算机上的数据丢失。然而由于数据的重复副本存储于另一存储器组 212 的服务器计算机上,所以数据中心 200 仍然让数据被存储。在一个或者多个实施例中,预计不同存储器组 212 存储相同数据的重复副本。在替选实施例中,不同存储器组 212 可以存储存储器高速缓存 204 存储的整个数据集的部分。例如,如果有四个存储器组 212,则每个存储器组 212 可以存储存储器高速缓存 204 存储的整个数据集的仅 75%。

[0028] 当存储器组 212 的服务器计算机在已经被掉电之后被重启时,先前存储于该服务器计算机上的数据再次存储于该服务器计算机上。可以用不同方式自动执行数据在服务器计算机上的这一重新存储。在一个或者多个实施例中,数据中心控制器 202 标识服务器计算机先前被掉电、但是后来已经被上电并且当前正在运行。数据中心控制器 202 标识将存储于该服务器计算机上的数据(例如用于特定用户的数据)并且指引另一存储器组 212 的适当一个或者多个服务器计算机向该服务器计算机发送该数据。取而代之,可以用不同方式执行数据在服务器上的重新存储,比如作为如下文更具体讨论的前摄回填(proactive backfilling)过程的部分。

[0029] 回填过程指代从一个存储器组向另一存储器组复制数据,从而相同数据存储于每

个存储器组中。回填过程也可以指代从与存储器组相同的数据中心中的数据库或者取而代之从与存储器组不同的数据中心中的数据库向存储器组复制数据。前摄回填过程指代尝试在请求未在特定存储器组中的数据之前标识该数据而不是响应于请求未在特定存储器组中的数据来标识该数据。

[0030] 在一个或者多个实施例中,根据特定固定排序在服务器计算机 214、216 的易失性存储介质中存储数据。可以使用多种不同固定排序(比如按姓名依字母顺序存储用户的数据、向数据(例如用于特定用户的各种数据)分配特定标识符(ID)和基于 ID 按数值升序存储数据等)。

[0031] 此外,存储于服务器计算机 214、216 上的数据具有关联版本号。这一版本号例如由数据中心控制器 202 或者服务器计算机 214、216 分配。每当更新数据时(比如每当用户请求改变他的或者她的数据时)版本号递增。

[0032] 通过比较来自两个存储器组 212 中的每个存储器组的 ID 和版本号集合以标识在集合之间的任何不同来执行前摄回填。由于数据通常按相同固定排序存储于存储器组 212 的服务器计算机上,所以依次比较集合中的数据以标识在集合之间的任何不同。例如可以获得来自第一存储器组 212 的始于特定位置(例如在存储器组的第一服务器中的第一存储位置)的用于数据的 50 个 ID 和版本号的第一集合并且可以获得来自第二存储器组 212 的始于对应位置的用于数据的 50 个 ID 和版本号的第二集合。如果在第一集合中有未在第二集合中的任何 ID,则从第一存储器组向第二存储器组传送针对该 ID 的数据并且反之亦然。类似地,如果在第一集合中有比第二集合中的对应标识符的版本号更大(指示数据的更新版本)的任何版本号,则从第一存储器组向第二存储器组传送针对该 ID 和版本号的数据并且反之亦然。集合的这一获得和比较可以由数据中心控制器 202 或者取而代之由一个或者多个服务器计算机 214、216 执行。

[0033] 取而代之,可以用不同方式执行前摄回填。例如可以比较数据(例如始于特定位置的针对前 50 个用户的数据)而不是比较 ID 和版本号。应当注意在比较数据本身的情形中, ID 和版本号无需在存储器组 212 中维护。作为另一示例,无需按固定排序存储数据。实际上,可以获得来自第一存储器组的数据并且搜索第二存储器组以确定来自第一存储器组的该数据是否存在于第二存储器组中。

[0034] 从存储器组 212 读取和向存储器组 212 写入数据的粒度可以变化并且由数据中心控制器 202 管理。例如可以作为组从存储器组 212 读取和向存储器组 212 写入针对特定用户的各种数据。因而按特定固定排序(例如按姓名依字母顺序)存储针对用户的数据,并且 50 个 ID 和版本号的每个集合是用于 50 个不同用户的数据的 ID 和版本号。作为另一示例,可以从存储器组 212 读取和向存储器组 212 写入按某一数据量(例如 1 兆字节)分组的数据。因而按特定固定次序(例如按向每 1 兆字节数据组分配的标识符依数值)存储针对用户的数据,并且 50 个 ID 和版本号的每个集合是用于 50 个不同的 1 兆字节数据组的 ID 和版本号。

[0035] 图 3 是图示了根据一个或者多个实施例的用于操作存储器高速缓存数据中心的例子过程 300 的流程图。过程 300 由存储器高速缓存数据中心(比如图 2 的存储器高速缓存数据中心 200)执行并且可以用软件、固件、硬件或者其组合来实施。过程 300 是用于操作存储器高速缓存数据中心的例子过程;这里参照不同图包括对操作存储器高速缓存数据



中心的附加讨论。

[0036] 在过程 300 中,在存储器高速缓存中维护用于数据中心的数据(动作 302)。虽然如上文讨论的那样,存储器高速缓存中的数据持久存储于位于远离实施过程 300 的数据中心的另一数据中心中,但是如上文讨论的那样,所述数据中心未包括本地数据库,该本地数据库作为由数据中心提供的服务所维护的数据的持久储存库。这一维护例如包括对从数据中心读取数据的请求做出响应、对向数据中心写入数据的请求做出响应、执行回填过程等。

[0037] 可以在数据中心接收对读取数据的请求。一般而言,响应于对读取数据的请求,从存储器高速缓存获取所请求的数据(动作 304) 并且将其返回给从其接收请求的计算设备(动作 306)。下文更具体讨论对读取请求做出响应的数据中心。

[0038] 也可以在数据中心接收对写入或者存储数据的请求。一般而言,响应于对写入或者存储数据的请求,向具有持久存储数据的数据库的另一数据中心发送请求的数据(动作 308)。在数据持久存储于数据库中之后,数据存储于存储器高速缓存中(动作 310)。下文更具体讨论对写入请求做出响应的数据中心。

[0039] 回到图 2,数据中心 200 可以从请求者接收针对数据的请求。请求者通常是计算设备(比如用户的计算设备,该用户希望使用数据中心 200 提供的服务)。针对数据的请求由数据中心控制器 202 接收,数据中心控制器 202 又尝试从存储器高速缓存 204 获取请求的数据。如果请求的数据未存储于存储器高速缓存 204 中,则控制器 202 访问一个或者多个其它数据中心以获得请求的数据。如果从另一数据中心获得,则控制器 202 在存储器高速缓存 204 中存储获得的数据并且也向请求者返回获得的数据。

[0040] 在访问另一数据中心以获得请求的数据的情形中,可以用不同方式标识所访问的其它一个或者多个数据中心。在一个或者多个实施例,数据中心控制器 202 包括或者访问记录,该记录指示哪个数据中心具有持久存储特定数据的数据库(例如基于向数据分配的 ID、基于请求数据的计算机设备的用户的姓名等)。包括持久存储该特定数据的数据库的数据中心是被访问的数据中心。取而代之,控制器 202 可以用其它方式标识一个或者多个其它数据中心,例如通过选择与数据中心 200 最近(基于地理距离或者以别的方式确定)的数据中心(并且如果不能从该数据中心获得数据则继续选择下一最近数据中心直至获得数据或者选择所有数据中心)。作为另一示例,控制器 202 可以随机、根据(例如由系统管理员)配置到控制器 202 中的预定顺序、根据一些其它规则或者标准等选择数据中心。

[0041] 取而代之,如果请求的数据未存储于存储器高速缓存 204 中,则所请求的数据的指示被置于队列中。数据中心控制器 202 工作遍历队列(例如按先入先出(FIFO)顺序)从而从另一数据中心获得队列中标识的数据并且向存储器高速缓存 204 添加数据。因此,可以向请求者通知数据暂时不可用而不是从另一数据中心获得请求的数据并且向请求者返回获得的数据。将随后在已经获得数据并且将其添加到存储器高速缓存 204 之后的将来某一时间点使数据可用。

[0042] 在一个或者多个实施例,向另一数据中心发送的针对数据的请求可以一起分批,从而向另一数据中心发送单个组合请求而不是多个个别请求。对单个组合请求中的个别请求的响应由另一数据中心生成并且也分批在一起作为单个组合响应向数据中心 200 返回。个别请求和响应的这一分批在一起减少了在数据中心之间发送的通信次数。

[0043] 可以用不同方式确定多个个别请求是否可以由数据中心控制器 202 分批在一起。

在一个或者多个实施例中,当将向另一数据中心发送请求时,数据中心控制器 202 等待特定时间量(例如 100 毫秒、1 秒等)并且将在特定时间量期间接收的所有请求(或者取而代之至阈值数量的所有请求)分批成单个组合请求。在其它实施例中,使用其它技术(比如等待直至将向另一数据中心发送阈值数目的请求)。其它数据中心可以在确定何时将响应分批在一起以向数据中心 200 返回时使用类似技术。

[0044] 图 4 是图示了根据一个或者多个实施例的用于从存储器高速缓存数据中心读取数据的例子过程 400 的流程图。过程 400 由数据中心(比如图 2 的数据中心 200)执行并且可以用软件、固件、硬件或者其组合来实施。过程 400 是用于从存储器高速缓存数据中心读取数据的例子过程;这里参照不同图包括从存储器高速缓存数据中心读取数据的附加讨论。

[0045] 在过程 400 中,从请求者接收针对数据的请求(动作 402)。请求的存储器组与请求关联。请求者可以标识请求的存储器组(例如先前已经由数据中心通知其中存储数据的存储器组),或者取而代之,另一部件或者设备可以标识请求的存储器组(例如数据中心控制器 202 的部件)。在一个或者多个实施例中,请求的存储器组是默认存储器组(例如在请求者或者其它部件或者设备未标识请求的存储器组的情形中)。应当注意默认存储器组可以对于多个请求而言相同或者取而代之可以变化。如果默认存储器组对于不同请求而言变化,则可以用不同方式(例如随机选择、按照遍历存储器组的轮询过程选择、根据其它规则或者标准选择等)标识用于特定请求的默认存储器组。

[0046] 访问请求的存储器组(动作 404),并且过程 400 基于请求的数据是否存储于请求的存储器组中而继续(动作 406)。如果请求的数据存储于请求的存储器组中,则从请求的存储器组获取(408)请求的数据并且将其返回给请求者(动作 410)。

[0047] 然而如果请求的数据未存储于请求的存储器组中,则关于针对请求的数据是否有尚未被校验的附加存储器组进行校验(动作 412)。如果针对请求的数据有尚未被校验的一个或多个存储器组,则标识和访问另一这样的存储器组(动作 414)。可以用多种不同方式(比如按存储器组标识符的数值顺序(例如存储器组 1、继而存储器组 2、继而存储器组 3 等)、基于随机选择、根据某一其它规则或者标准等标识另一这样的存储器组。

[0048] 过程 400 基于请求的数据是否存储于标识的存储器组中而继续(动作 416)。如果请求的数据未存储于标识的存储器组中,则过程 400 返回以校验针对请求的数据是否有尚未被校验的附加存储器组(动作 412)。

[0049] 然而如果请求的数据未存储于标识的存储器组中,则从标识的存储器组获取请求的数据并且将其回填到请求的存储器组中(动作 418)。因而下次请求数据时预计数据将在请求的存储器组中。除了向请求的存储器组中回填请求的数据之外,还向请求者返回请求的数据(动作 410)。

[0050] 回到动作 412,如果针对请求的数据无尚未被校验的附加存储器组,则从另一数据中心获得请求的数据(动作 420)。向另一数据中心发送针对数据的请求,并且作为响应,从另一数据中心接收请求的数据。如上文讨论的那样,从其获得请求的数据的该另一数据中心可以是多种其它数据中心。

[0051] 向一个或者多个存储器组添加(动作 422)从另一数据中心获得的请求的数据并且也将其返回给请求者(动作 410)。在动作 422 中,可以向一个存储器组(例如请求的存储器

组)添加请求的数据,并且依赖于前摄回填向其它存储器组复制数据。取而代之,可以在动作 422 中向多个存储器组添加请求的数据。

[0052] 过程 400 讨论在动作 406 和 412 中校验存储器组从而造成在从另一数据中心获得请求的数据之前校验所有存储器组。取而代之,可以从另一数据中心获得请求的数据而无须已校验所有存储器组。如果请求的数据未在校验的存储器组中,则过程 400 继续动作 420 以从另一数据中心获得请求的数据。例如可以校验存储器组中阈值数目的存储器组(比如存储器组中固定数目的存储器组(例如存储器组中的三个存储器组)或者相对数目的存储器组(例如存储器组中的 75%))。作为另一示例,另一机制或者技术可以用来在校验存储器组中的至少阈值数目(固定或者可变数目)的存储器组之后确定存储器组最可能未包括请求的数据。

[0053] 应当注意可以出现如下情形,其中请求的数据在存储器组中不可用并且从另一数据中心不可获得。在这样的情形中,向请求者返回请求的数据不可用这样的指示。取而代之,根据数据的性质,请求的数据可以可选地由存储器高速缓存数据中心生成或者创建。

[0054] 回到图 2,数据中心 200 也可以从请求者接收对存储数据的请求。对存储数据的请求由数据中心控制器 202 接收,数据中心控制器又在存储器高速缓存 204 中存储数据。控制器 202 也向包括数据库的另一数据中心传达数据,该数据库持久存储该数据。可以用多种不同方式(类似于如上文讨论的那样标识将从其读取数据的数据中心的方式)标识控制器 202 向其传达数据的数据中心。传达给另一数据中心的数据可以在个别请求基础上(例如控制器 202 可以针对每个对存储数据的请求而发送单独通信)传达,或者取而代之,来自对存储数据的多个请求的数据可以分批在一起并且发送到另一数据中心作为单个组合存储请求(类似于上文关于将针对数据的请求分批的讨论)。

[0055] 图 5 是图示了根据一个或者多个实施例的用于向存储器高速缓存数据中心写入数据的例子过程 500 的流程图。可以用软件、固件、硬件或者其组合实施过程 500。过程 500 的在图 5 的左手侧上图示的动作由第一数据中心(数据中心 1)(比如图 2 的数据中心 200 或者图 1 的数据中心 102)执行。过程 500 的在图 5 的右手侧上图示的动作由第二数据中心(数据中心 2)(比如图 1 的数据中心 104 或者 106)执行。过程 500 是用于向存储器高速缓存数据中心写入数据的例子过程;这里参照不同图包括向存储器高速缓存数据中心写入数据的附加讨论。

[0056] 在过程 500 中,接收待存储的数据(动作 502)。接收所述数据作为写入请求的部分。接收的数据可以是新数据(比如用于新用户的数据)。取而代之,接收的数据可以是对现有数据的改变(比如改变已经存储于存储器高速缓存中的关于特定用户的数据)。

[0057] 向第二数据中心发送接收的数据(动作 504)。如上文讨论的那样,第一数据中心是存储器高速缓存数据中心,因而在动作 504 中向具有持久存储数据的数据库的另一数据中心发送接收的数据。

[0058] 第二数据中心接收数据(动作 506)并且向第二数据中心的数据库添加数据(动作 508)。在向第二数据中心的数据库添加数据之后,第二数据中心向第一数据中心发送已经向第二数据中心的数据库添加数据这样的指示(动作 510)。

[0059] 第一数据中心接收指示(动作 512)并且向存储器高速缓存的一个或者多个存储器组添加接收的数据(来自动作 502)(动作 514)。如果接收的数据是新数据,则向存储器高

速缓存添加这一新数据。如果接收的数据是对现有数据的改变,则可以向存储器高速缓存添加该接收的数据并且无效或者删除先前数据,或者取而代之,该接收的数据可以重写先前数据。

[0060] 在动作 514 中,可以向一个存储器组(例如默认存储器组)添加接收的数据,并且依赖于前摄回填向其它存储器组复制数据。取而代之,可以在动作 514 中向多个存储器组添加接收的数据。

[0061] 第二数据中心在动作 510 中发送的指示可以采用多种不同形式。在一个或者多个实施例中,在动作 510 中发送的指示是在动作 502 中接收的数据。在这样的实施例中,第一数据中心无需保持对在动作 502 中接收的数据的跟踪、但是可以简单地向第二数据中心转发接收的数据、然后响应于在动作 512 中从第二数据中心接收数据而向一个或者多个存储器组添加从第二数据中心接收的数据。在其它实施例中,在动作 510 中发送的指示是特定数据的标识符。例如第一数据中心可以暂时存储在动作 502 中接收的数据并且在动作 504 中将这一暂时存储的数据的标识符与数据一起提供到第二数据中心。第二数据中心在动作 510 中发送的指示是暂时存储的数据的这一标识符。响应于在动作 512 中接收暂时存储的数据的标识符,第一数据中心可以在动作 514 中向一个或者多个存储器组添加暂时存储的数据而无须第二数据中心已在动作 510 中发送数据。

[0062] 第二数据中心也可以可选地向第二数据中心的存储器高速缓存添加数据(动作 516)。通常仅在被请求改变的现有数据已经存储于第二数据中心的存储器高速缓存中的情形中向第二数据中心的存储器高速缓存添加数据。

[0063] 回到图 2,在一个或者多个实施例中,一个或者多个存储器组 212 为非易失性盘存储介质而不是易失性存储介质。多种不同类型的盘和盘驱动可以用作非易失性盘存储介质(比如磁硬盘驱动、光盘和驱动等)。非易失性盘存储介质可以比易失性存储介质成本更少,因而使用非易失性盘存储介质的存储器组可以是一种用于在存储器高速缓存 204 中存储数据的成本有效技术。

[0064] 应当注意存储器组 212(该存储器组为非易失性盘存储介质)不同于数据库。没有管理向非易失性盘存储介质存储器组 212 存储数据和从非易失性盘存储介质存储器组 212 获取数据的数据库应用。实际上,将与存储于易失性盘存储介质存储器组 212 中相同的数据存储于非易失性盘存储介质存储器组 212 中,并且这样的数据无论存储介质为易失性还是非易失性存储介质都按相同顺序存储或者布置于存储介质上。

[0065] 具有非易失性盘存储介质存储器组 212 有助于在存储器组 212 被掉电并且易失性存储介质中的数据丢失的情况下回填易失性存储介质。在这样的情形中,当存储器组 212 再次被上电时,可以从非易失性盘存储介质回填易失性存储介质而不是从具有包括数据的数据库的另一数据中心获得数据。例如假设多个服务器计算机 214、216 在基本上相同时间被掉电并且还假设存储特定数据的重复副本的服务器计算机 214、216 是掉电的服务器计算机。当这些服务器被恢复上电时,存储于易失性存储介质中的特定数据已经丢失并且可以从非易失性盘存储介质回填(如果具有非易失性盘存储介质的服务器未被掉电或者当具有非易失性盘存储介质的服务器被恢复上电时)。因此可以从非易失性盘存储介质回填而不是从另一数据中心获得由于服务器计算机 214、216 的掉电而丢失的数据。从非易失性盘存储介质复制数据通常比从另一数据中心获得数据更快并且也减少了在数据中心之间传

送请求数目和数据量。

[0066] 应当注意可能出现如下情形,其中在从非易失性盘存储介质向存储器组 212 复制数据之时在存储器组 212 接收计算设备请求存储的数据。在这样的情形中,存储器组 212 维护更新记录,该记录包括接收的存储请求。在已经从非易失性盘存储介质向存储器组 212 复制数据之后,向存储器组 212 写入在更新记录中包括的存储请求。

[0067] 对于一些非易失性盘存储介质,用于访问存储介质(包括从存储介质读取数据和向存储介质写入数据)的时间比用于访问易失性存储介质的时间明显更慢。因而在这样的情形中,在尝试从非易失性盘存储介质获取数据之前尝试从一个或者多个易失性存储介质存储器组获取数据。此外,在这样的情形中,非易失性盘存储介质存储器组可以被配置成改进用于访问存储器高速缓存 204 中的存储介质的时间。非易失性盘存储介质存储器组可以被配置成部分通过非易失性盘存储介质中依次存储数据来访问存储介质。例如存储用于特定用户的数据,从而可以从非易失性盘存储介质依次读取数据而不造成发现数据的一个或者多个随机存储部分的寻找时间损失。

[0068] 在一个或者多个实施例,在向非易失性盘存储介质存储器组写入数据时,在易失性存储器中维护对非易失性盘存储介质的改变的日志。当日志满足或者超过阈值大小(例如阈值字节数目、阈值写入次数、用于阈值数目的用户的数据等)时,向非易失性盘存储介质写入日志中的数据。取而代之,可以响应于其它事件或者阈值(比如当数据已经在日志中持续至少阈值时间量时)向非易失性盘存储介质写入日志中的数据。当向非易失性盘存储介质写入日志时,写入日志,从而在非易失性盘存储介质上依次写入数据。在非易失性盘存储介质上依次写入数据指代写入数据使得可以从存储介质读取数据而不必寻找存储介质的新随机位置以读取数据的特定部分。例如数据可以在存储介质上物理上邻近地设置。在非易失性盘存储介质上依次写入数据的具体方式可以基于非易失性盘存储介质的特定设计而变化。

[0069] 向非易失性盘存储介质写入日志可以包括仅从日志向非易失性盘存储介质写入数据并且也可以包括从非易失性盘存储介质读取与日志中的数据组合、然后向非易失性盘存储介质写入的数据。例如可以在日志中包括用于用户的附加数据。如果在非易失性盘存储介质的区域中有充分空间用于写入附加数据并且仍然让用于用户的数据依次存储于非易失性盘存储介质上则才可以向非易失性盘存储介质写入这一附加数据。如果无这样的充分空间,则可以从非易失性盘存储介质读取用于用户的数据,向其添加来自日志的数据,并且可以向非易失性盘存储介质依次写入用于用户的这一新数据(包括来自日志的附加数据)。

[0070] 此外,日志中的每个条目可以具有关联的时间戳(例如指示条目何时被添加到日志、指示数据中心控制器 202 何时接收改变等)。可以按各种规律或者不规律间隔(例如近似按小时)取得易失性存储介质存储器组(该存储器组存储非易失性盘存储介质存储器组中的数据的重副本)的快照。这一快照是易失性存储介质存储器组的副本并且向非易失性盘存储介质写入。从日志删除日志中的任何如下条目,这些条目具有比取得快照的时间更晚的关联时间戳(无需向非易失性盘存储介质写入这样的条目,因为作为向非易失性盘存储介质写入快照的部分已经将它们写入非易失性盘存储介质)。

[0071] 在数据依次存储于非易失性盘存储介质上的情形中,向非易失性盘存储介质依次

读取或者写入分组在一起的数据,并且可以用各种方式将数据分组在一起。例如可以将用于特定用户的数据分组在一起。作为另一示例,可以将非易失性盘存储介质上的数据一起分组成称为桶(bucket)的组,并且每个桶可以包括用于一个或者多个用户的数据(例如桶可以是 200 兆字节)。

[0072] 应当注意可以出现以下情形,其中用于访问非易失性盘存储介质的时间与用于访问易失性存储介质的时间近似相同。例如可以出现如下情形,其中非易失性盘存储介质是固态驱动(例如闪存驱动)。在这样的情形中,可以将非易失性盘存储介质视为如同它是易失性存储介质一样。例如无需在进行尝试以从非易失性盘存储介质获取数据之前尝试从易失性存储介质获取数据,并且非易失性盘存储介质无需被配置成改进用于访问存储介质的时间。

[0073] 在一个或者多个实施例中,存储于存储器高速缓存 204 中的数据具有关联期满时间或者生存时间(TTL)。期满时间关联的数据的粒度可以变化。例如用于不同用户的数据可以具有单独关联期满时间,用于用户的个别数据段(例如电子邮件地址、当前照片等)可以具有单独关联期满时间,更大数量(例如 200 兆字节块)的数据可以具有单独关联期满时间,等等。当接收针对数据的请求并且与请求的数据关联的期满时间已经过去(也称为数据过时)时,即使数据过时仍然可以向请求者返回请求的数据。此外,可以从另一数据中心或者另一存储器组获得请求的数据的新副本(未过时),因此将用未过时的数据针对请求的数据的后续请求做出响应。

[0074] 除了数据中心 200 已经请求的数据之外,数据中心 200 还可以从其它数据中心接收附加数据。这也称为从其它数据中心向数据中心 200 推送数据。可以出于多种不同原因(比如当对数据的改变出现在其它数据中心时)从其它数据中心向数据中心 200 推送数据。

[0075] 向数据中心 200 以及向其它数据中心的这一推送允许对数据库发生的更新将在各存储器高速缓存中存储,无论这些存储器高速缓存所位于的数据中心如何。响应于向在数据中心的数据库写入数据,该数据中心向各其它数据中心通知对数据库的更新,从而各其它数据中心(包括数据中心 200)中的存储器高速缓存可以包括所述更新。向数据库写入的这一数据可以是对先前存储的数据(例如先前为特定用户存储的数据)的改变或者先前版本尚未存储的其它新数据(例如用于新用户的数据)。对数据库的改变因此与各数据中心中的各存储器高速缓存同步。

[0076] 这里关于管理在多个存储器组上和 / 或多个数据中心中存储数据的多个副本讨论各种技术。一般而言,可以用多种不同方式执行数据在不同存储库(无论那些存储库是存储器组还是数据中心)之间的这样的复写并且可以使用多种不同同步或者复写技术或者过程来执行这样的复写。

[0077] 在一个或者多个实施例中,当向数据中心 200 推送数据时,数据中心控制器 202 校验接收的数据的先前版本是否已经存储于存储器高速缓存 204 中。例如可以响应于哪个数据中心控制器 202 校验用于该特定用户的数据是否已经存储于存储器高速缓存 204 中来向数据中心 200 推送用于该特定用户的新数据或者改变的数据。如果接收的数据的先前版本已经存储于存储器高速缓存 204 中,则控制器 202 更新存储器高速缓存 204 中的接收数据。这一更新通常不(虽然取而代之可以)影响近来使用的该数据如何存储于存储器高速缓存 204 中,因而仅仅由于向数据中心 200 推送数据而未将数据反映为已经在存储器高速缓

存 204 中被新近使用。然而如果接收的数据的先前版本未已经存储于存储器高速缓存 204 中,则控制器 202 未在存储器高速缓存 204 中存储接收的数据并且忽略或者丢弃接收的数据。如果接收的数据的先前版本未已经存储于存储器高速缓存 204 中,则控制器 202 认为数据中心 200 近来尚未接收针对该数据的请求,因而控制器 202 无需保持存储器高速缓存 204 中的该数据。

[0078] 也应当注意可以出现如下情形,其中接收在存储器高速缓存 204 中存储数据的请求,但是在存储器高速缓存 204 中用于该数据的存储空间不足。在这样的情形中,可以删除存储器高速缓存 204 的部分以便为新接收的数据腾出空间。可以用多种不同方式(比如通过标识存储器高速缓存 204 的近来最少使用(LRU)的部分、通过标识存储器高速缓存 204 的过时部分等)标识存储器高速缓存 204 的删除部分。作为删除部分的数据的粒度可以变化。例如,可以维护用于不同用户的数据的近来访问信息并且部分可以是用于特定用户的数据。作为另一示例,可以维护用于特定数量的数据(例如 200 兆字节块)的近来访问信息并且部分可以是这一特定的数据量。如果在从存储器高速缓存 204 删除数据之后接收针对删除的数据的请求,则从另一数据中心获得删除的数据。

[0079] 此外,数据中心 200 以及数据中心 200 与之通信的其它数据中心可以使用各种技术来减少传送的数据量和/或在数据中心之间传送数据时消耗的带宽。一种这样的技术是运用压缩算法以压缩通信的数据、由此减少传送的数据量。在一个或者多个实施例中运用的另一技术是标识其它数据中心需要的特定数据段并且仅传送那些特定段。这些特定段例如是新添加的数据段或者已经改变的数据段。例如假设数据中心维护关于不同用户的数据用于社交联网服务,并且这一数据包括用户的姓名、电子邮件地址、最喜爱照片等。如果改变用户的电子邮件地址,则仅需向另一数据中心传送这一数据段(新电子邮件地址)和数据段是什么的指示。无需传送其它数据段(例如最喜爱照片),因为未出现对该数据的改变。在数据中心的控制器(例如数据中心控制器 202)配置有或者以别的方式具有对在存储器高速缓存(例如存储器高速缓存 204)中存储数据的方式的了解。因而控制器可以容易将特定数据段与特定数据段的指示一起传送,并且接收控制器可以容易向存储器高速缓存添加该特定数据段。

[0080] 这里描述的存储器高速缓存数据中心技术也有助于管理数据中心 200 中的服务器计算机。存储器高速缓存数据中心技术可以例如通过允许数据中心 200 在一个或者多个服务器计算机被掉电时继续操作并且在它们被恢复上电时向适当服务器计算机自动复制数据(例如回填)来有助于这样的管理。例如一个或者多个服务器计算机可以在各种时间被掉电以便修复损坏的服务器计算机、更换服务器计算机、在服务器计算机上安装新软件或者软件升级等。数据中心 200 是弹性的并且能够适应这样的情形从而如上文讨论的那样在服务器计算机被恢复上电之后向它们复制数据。

[0081] 此外应当注意通常预计存储于存储器高速缓存 204 中的数据未确切地是存储于其它数据中心(例如图 1 的数据中心 104 和 106)的存储器高速缓存中的相同数据。预计数据中心 20 接收和处置来自特定一组计算机设备(比如(基于地理距离或者别的方式确定)与数据中心 200 接近的计算设备)的针对数据的请求。类似地,预计其它数据中心接收和处置来自不同组计算设备(基于地理距离或者以别的方式确定与那些其它数据中心接近的计算设备)的针对数据的请求。当预计不同计算设备的用户具有和请求对不同数据的访问

时,预计存储于存储器高速缓存 204 中的数据不同于存储于其它数据中心的存储器高速缓存中的数据。

[0082] 此外,上文讨论当在数据中心控制器 202 处接收对存储数据的请求时,数据在存储器高速缓存 204 中存储并且也向具有持久存储数据的数据库的数据中心发送。取而代之,可以出现如下情形,其中未向具有持久存储数据的数据库的这样的数据中心发送数据。数据中心控制器 202 配置有或者以别的方式标识如下信息,该信息指示无需向具有数据库的这样的数据中心发送的数据。在一个或者多个实施例,无需向具有数据库的这样的数据中心发送多种不同类型的数据(比如频繁改变的数据(例如按照比特定阈值更大来接收对存储数据的请求)、如果丢失则可以容易重新计算(例如基于已经存储于数据库中的数据重新计算)的数据等)。

[0083] 因此,这里讨论的存储器高速缓存数据中心允许各自包括多个服务器计算机的数据中心相互组合工作以向跨越宽泛地理范围设置的计算设备提供服务。存储器高速缓存数据中心无数据库,这些数据库是用于由数据中心提供的服务所维护的数据的持久储存库,这减少了存储器高速缓存数据中心的成本和和在数据中心之间通信的数据量(因为存储器高速缓存数据中心未包括需要保持与其它数据中心的数据库同步的数据库)。

[0084] 这里进行如下讨论,这些讨论涉及在可以跨越宽泛地理范围遍布的数据中心内包括存储器高速缓存和数据库。然而将理解这里讨论的技术也可以应用于其它系统。例如这里讨论的技术可以与数据中心内的不同机架而不是不同数据中心系统一起使用。一些机架可以包括具有存储器高速缓存和数据库(这些数据库是用于数据的持久储存库,该数据由机架提供的服务维护)的服务器计算机,其它机架可以包括具有存储器高速缓存、但是无数据库(这些数据库是用于数据的持久储存库,该数据由机架提供的服务维护)的服务器计算机,而更多其它机架可以包括具有数据库(这些数据库是用于数据的持久储存库,该数据由机架提供的服务维护)、但是无存储器高速缓存的服务器计算机。作为另一示例,这里讨论的技术可以使用于数据中心内的不同服务器计算机群集而不是不同数据中心系统内。一些群集可以包括具有存储器高速缓存和数据库(这些数据库是用于数据的持久储存库,该数据由群集提供的服务维护)的服务器计算机,其它群集可以包括具有存储器高速缓存、但是无数据库(这些数据库是用于数据的持久储存库,该数据由群集提供的服务维护)的服务器计算机,而更多其它群集可以包括具有数据库(这些数据库是用于数据的持久储存库,该数据由群集提供的服务维护)、但是无存储器高速缓存的服务器计算机。

[0085] 图 6 图示了例子计算设备 600,该计算设备可以被配置成实施根据一个或者多个实施例的存储器高速缓存数据中心的方面。计算设备 600 可以用来例如实施图 1 的数据中心 102、数据中心 104 或者数据中心 106 的部分或者可以是图 2 的服务器计算机 206、214 或者 216。

[0086] 计算设备 600 包括一个或者多个处理器或者处理单元 602、一个或者多个计算机可读介质 604(其可以包括一个或者多个存储器和 / 或存储部件 606)、一个或者多个输入 / 输出(I/O)设备 608 以及允许各种部件和设备相互通信的总线 610。可以包括计算机可读介质 604 和 / 或一个或者多个 I/O 设备 608 作为计算设备 600 的部分,或者取而代之可以耦合到计算设备 600。总线 610 代表使用多种不同总线架构的若干类型的总线结构(包括存储器总线或者存储器控制器、外围总线、加速图形端口、处理器或者本地总线等)中的一个



或者多个类型。总线 610 可以包括有线和 / 或无线总线。

[0087] 存储器 / 存储部件 606 代表一个或者多个计算机存储介质。部件 606 可以包括易失性介质(比如随机存取存储器(RAM))和 / 或非易失性介质(比如只读存储器(ROM)、闪存、光盘、磁盘等)。部件 606 可以包括固定介质(例如 RAM、ROM、固定硬驱动等)以及可拆卸介质(例如闪存驱动、可拆卸硬驱动、光盘等)。

[0088] 可以用软件实施这里讨论的技术而指令由一个或者多个处理单元 602 执行。将理解不同指令可以存储于计算设备 600 的不同部件中(比如处理单元 602 中、处理单元 602 的各种高速缓存存储器中、设备 600 的其它高速缓存存储器(未示出)中、其它计算机可读介质上,等等)。此外将理解指令在计算设备 600 中存储的位置可以随时间改变。

[0089] 一个或者多个输入 / 输出设备 608 允许用户向计算设备 600 录入命令和信息并且也允许向用户和 / 或其它部件或者设备呈现信息。输入设备的例子包括键盘、光标控制设备(例如鼠标)、麦克风、扫描仪等。输出设备的例子包括显示器设备(例如监视器或者投影仪)、扬声器、打印机、网卡等。

[0090] 这里可以在一般软件背景中描述各种技术。一般而言,软件包括执行特定任务或者实施特定抽象数据类型的例程、程序、对象、部件、数据结构等。可以在某一形式的计算机可读介质上存储或者跨越该计算机可读介质传输这些技术的实施方式。计算机可读介质可以是计算设备可以访问的任何一个或者多个可用介质。举例而言而无限制,计算机可读介质可以包括“计算机存储介质”和“通信介质”。

[0091] “计算机存储介质”包括用任何用于存储信息的方法或者技术实施的易失性和非易失性、可拆卸和不可拆卸介质(所述信息比如计算机可读指令、数据结构、程序模块或者其它数据)。计算机存储介质包括但不限于 RAM、ROM、EEPROM、闪存或者其它存储器技术、CD-ROM、数字多功能盘(DVD)或者其它光学存储器、磁盒、磁带、磁盘存储器或者其它磁存储设备或者可以用来存储所需信息并且可以由计算机访问的任何其它介质。

[0092] “通信介质”通常在调制的数据信号(比如载波或者其它传送机制)中包含计算机可读指令、数据结构、程序模块或者其它数据。通信介质也包括任何信息递送介质。术语“调制的数据信号”意味着如下信号,该信号让它的特性中的一个或者多个特性以在信号中对信息编码这样的方式来设置或者改变。举例而言而未限制,通信介质包括有线介质(比如有线网络或者直接接线连接)和无线介质(比如声学、RF、红外线和其它无线介质)。也在计算机可读介质的范围内包括上述介质中的任何介质的组合。

[0093] 一般而言,可以使用软件、固件、硬件(例如固定逻辑电路)、人工处理或者这些实施方式的组合来实施这里描述的操作或者技术中的任何操作或者技术。在软件实施方式的情况下,使用当在处理器(例如一个或者多个 CPU)上执行时执行指定任务的程序代码来实施所述操作或者技术。程序代码可以存储于一个或者多个计算机可读存储器设备(可以参照图 6 发现其进一步的描述)上。这里描述的存储器高速缓冲数据中心技术的特征是平台独立的,这意味着可以在具有多种处理器的多种商用计算机平台上实施技术。

[0094] 虽然已经用结构特征和 / 或方法动作特有的语言描述了主题内容,但是将理解在所附权利要求中限定的主题未必限于上文描述的特定特征或者动作。实际上,上文描述的特定特征和动作作为实施权利要求的示例形式公开。

100

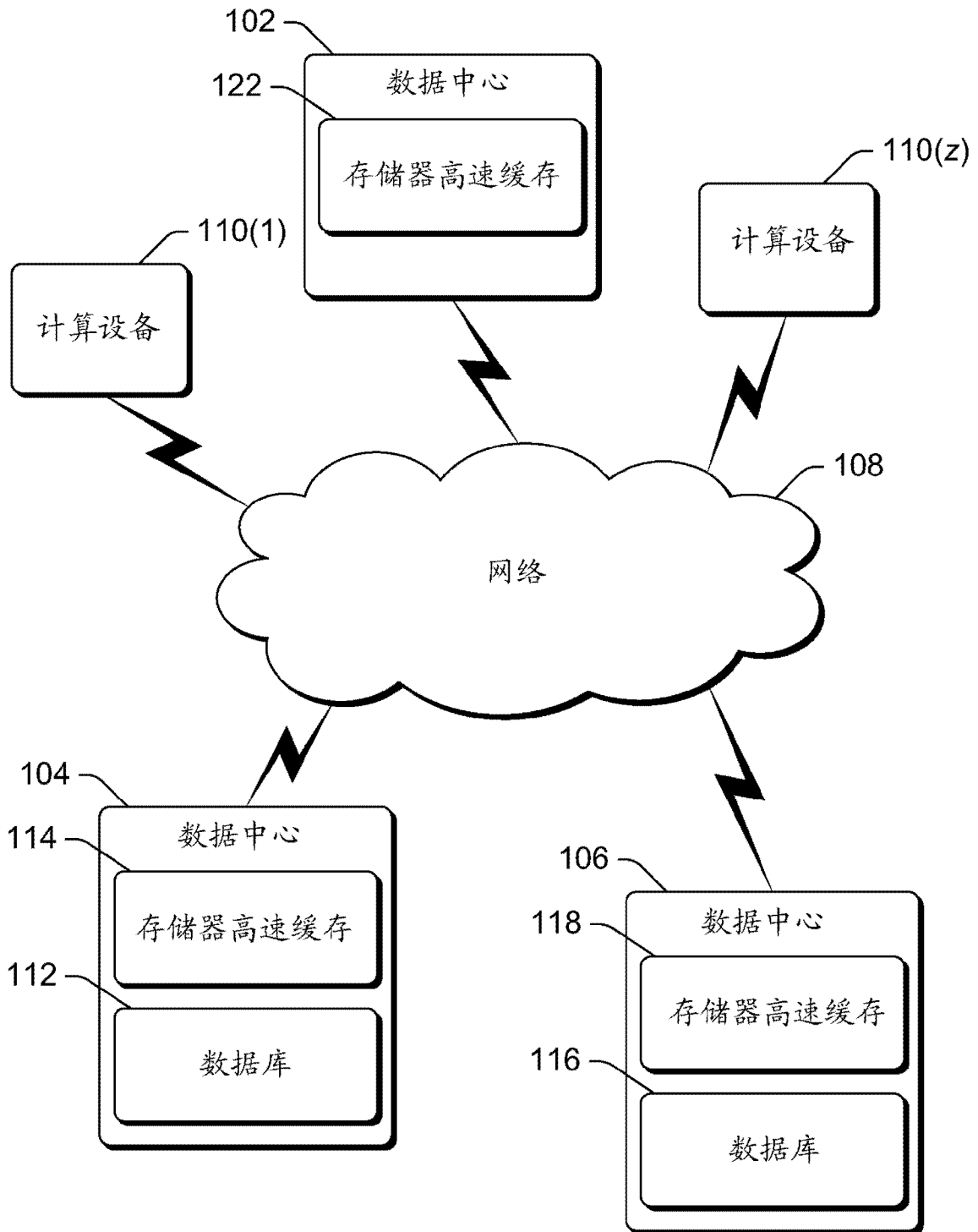


图 1

200

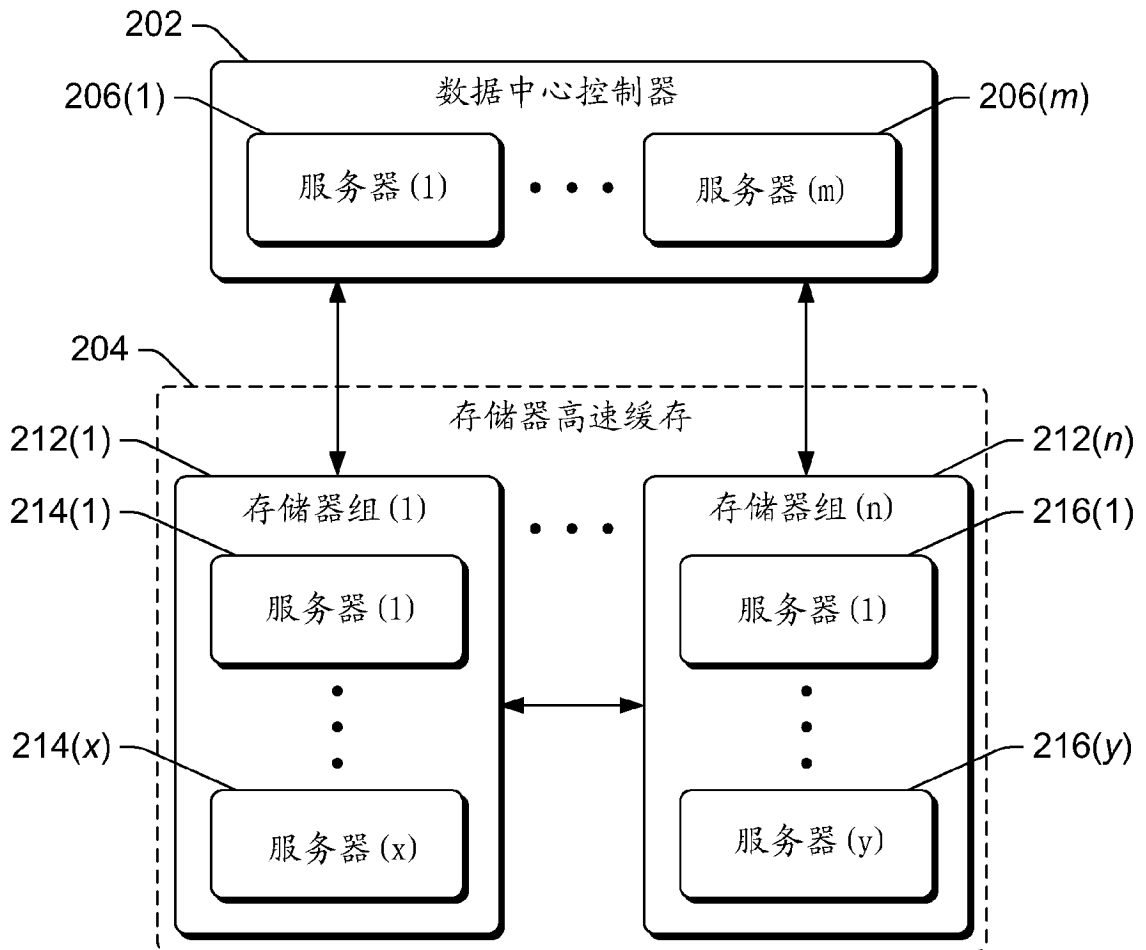


图 2

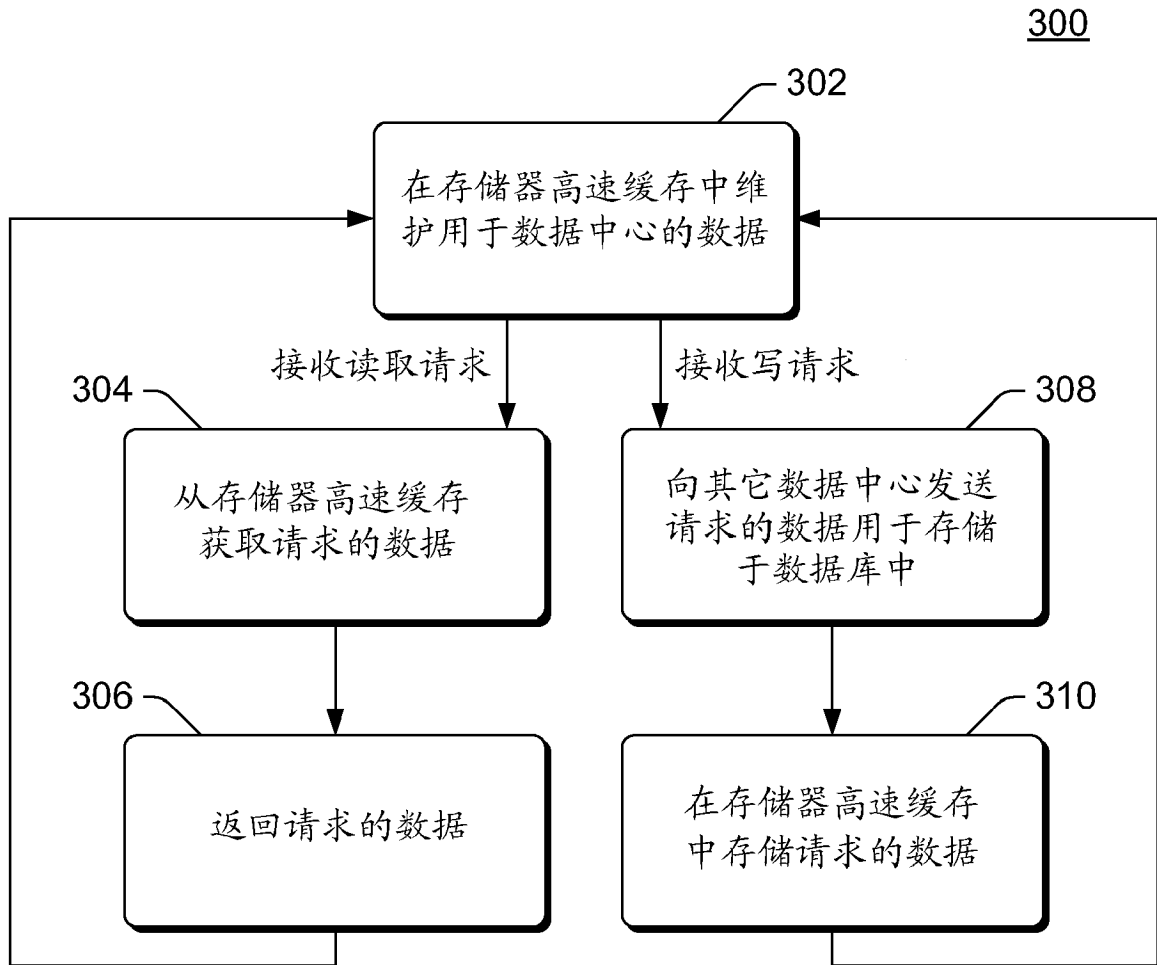


图 3

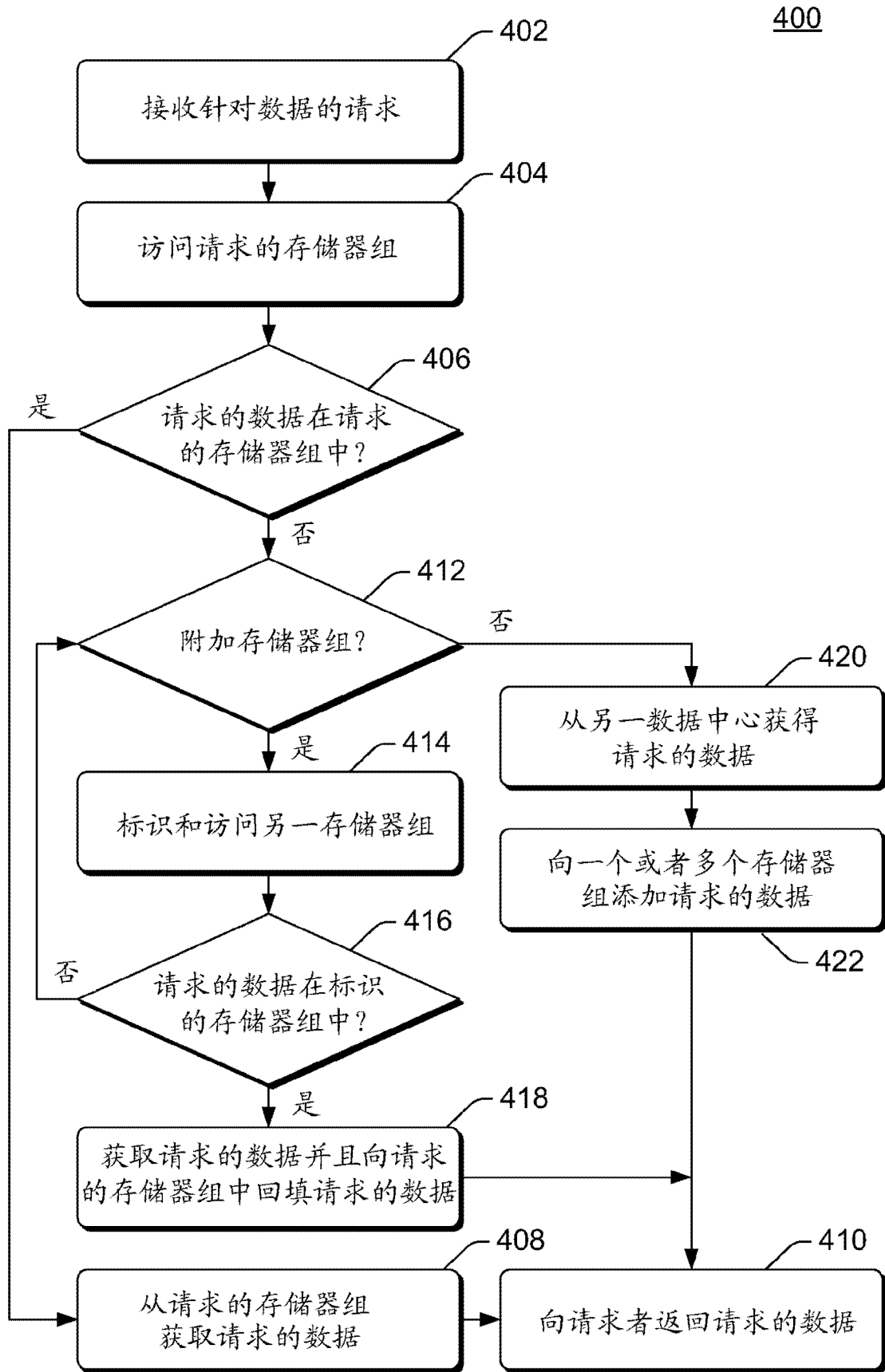


图 4

500

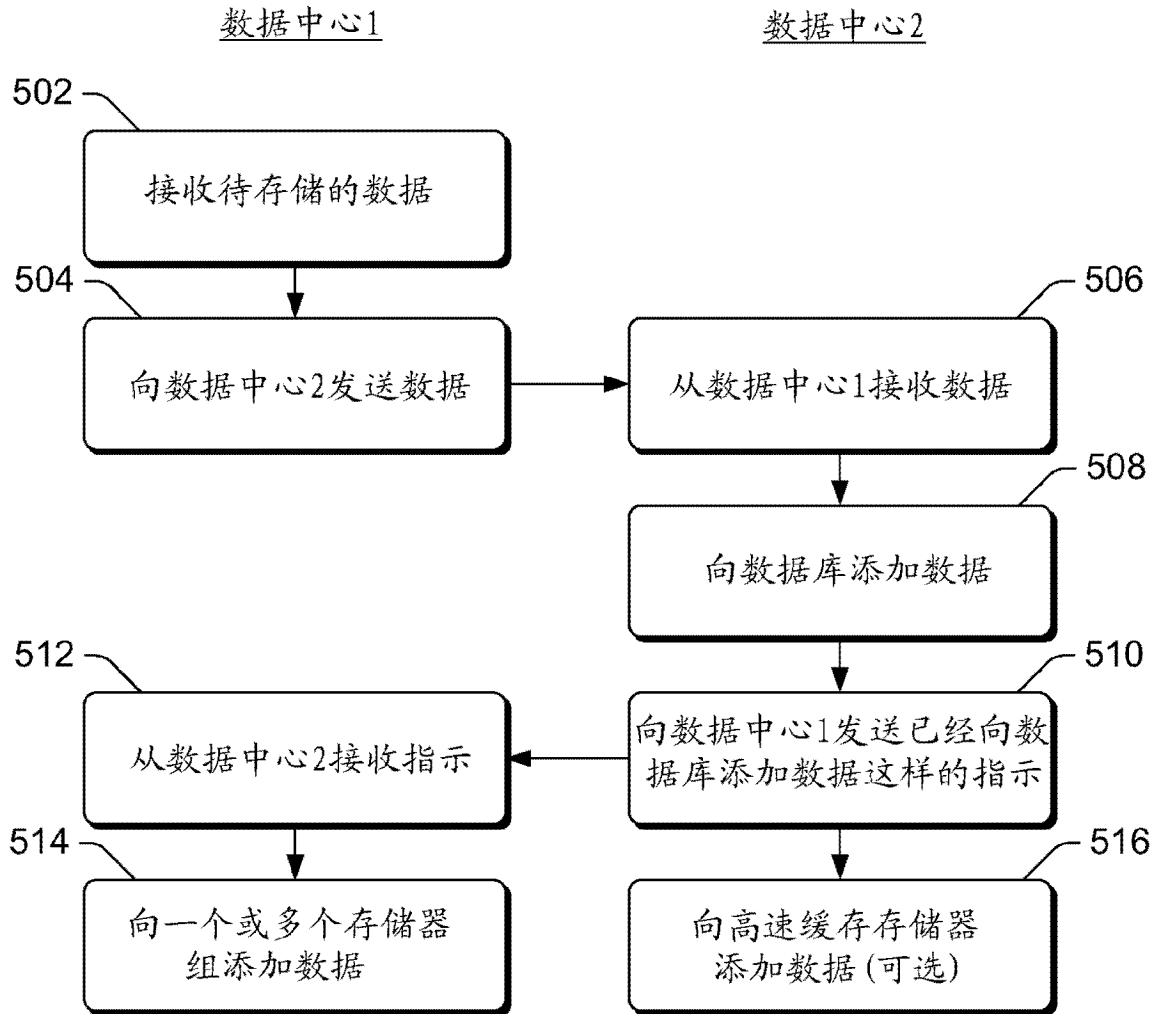


图 5

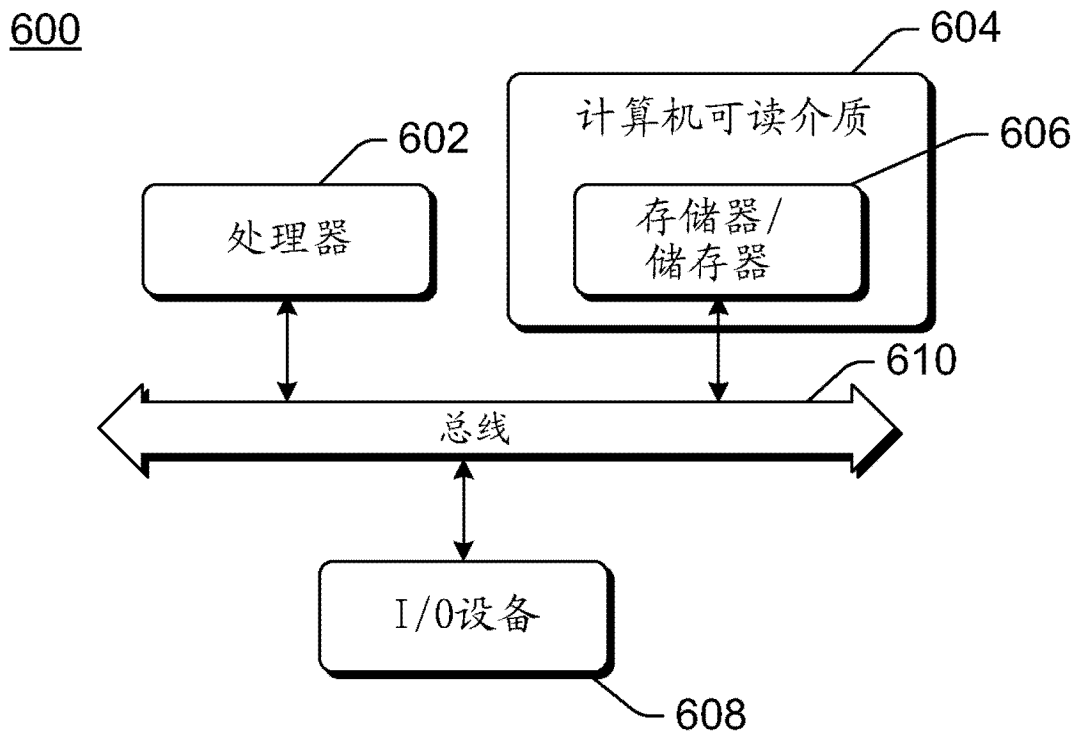


图 6