

(12) 发明专利

(10) 授权公告号 CN 101802815 B

(45) 授权公告日 2013. 07. 17

(21) 申请号 200880107014. 6

(22) 申请日 2008. 09. 12

(30) 优先权数据

11/855, 864 2007. 09. 14 US

(85) PCT申请进入国家阶段日

2010. 03. 11

(86) PCT申请的申请数据

PCT/US2008/076262 2008. 09. 12

(87) PCT申请的公布数据

W02009/036345 EN 2009. 03. 19

(73) 专利权人 微软公司

地址 美国华盛顿州

(72) 发明人 D·R·戴维斯 M·C·托马斯

M·R·克拉克

(74) 专利代理机构 上海专利商标事务所有限公

司 31100

代理人 顾嘉运 钱静芳

(51) Int. Cl.

G06F 17/00 (2006. 01)

G06F 17/40 (2006. 01)

G06F 15/16 (2006. 01)

(56) 对比文件

WO 2005/096187 A1, 2005. 10. 13, 说明书第 0017 段至第 0058 段, 附图 1-2、7-9.

US 2002/0070965 A1, 2002. 06. 13, 全文.

WO 2006/118823 A2, 2006. 11. 09, 全文.

CN 1656468 A, 2005. 08. 17, 说明书第 7 页第 26 行至第 25 页第 10 行, 附图 1-3.

审查员 赵传海

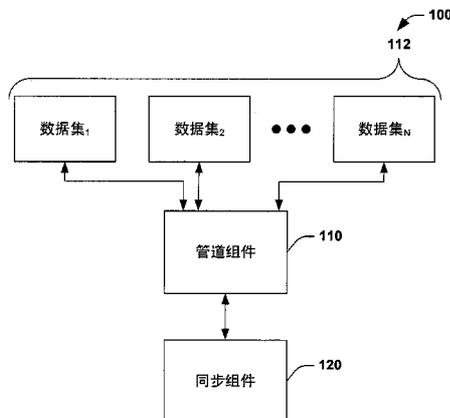
权利要求书2页 说明书11页 附图11页

(54) 发明名称

数据 - 驱动同步

(57) 摘要

与特定应用程序和 / 或端点相比, 同步被绑定到数据和可任选的一组同步行为。结果, 端点之间的同步不需要创建定制的同步解决方案或依赖于供应商和 / 或应用程序专用机制。在一种情况下, 可以预订标准数据和 / 或同步方案并将其用来动态地配置跨端点的同步。



1. 一种数据同步系统(100),包括:
启用与两个或更多数据集(112)的交互的管道组件(110);以及
便于所述数据集(112)的不由相关联的应用程序和 / 或端点绑定的数据 - 驱动同步的同步组件(120);

其中所述同步组件包括变更检测组件,所述变更检测组件被配置为描述所述两个或更多数据集的至少一个中的变更被标识的方式,并且所述同步组件被进一步配置为根据所描述的方式来执行所述两个或更多数据集的同步。

2. 如权利要求 1 所述的系统,其特征在于,还包括定义用于同步的数据集数据结构的数据方案组件(230)。

3. 如权利要求 2 所述的系统,其特征在于,所述数据方案组件(230)标识变更单元。

4. 如权利要求 3 所述的系统,其特征在于,所述数据方案组件(230)指定内容类型。

5. 如权利要求 2 所述的系统,其特征在于,还包括定义如何交换数据来用于同步目的的同步模型组件(240)。

6. 如权利要求 5 所述的系统,其特征在于,所述同步模型组件(240)指定数据的当前状态。

7. 如权利要求 5 所述的系统,其特征在于,还包括比较所述数据集并标识差异的组件(210)。

8. 如权利要求 7 所述的系统,其特征在于,还包括更新一个或多个数据集以解决所述差异并从而同步所述数据集的组件(220)。

9. 如权利要求 1 所述的系统,其特征在于,所述数据集或其片段符合可扩展内容类型方案。

10. 一种跨 - 存储同步方法,包括:

确定数据集的内容类型和变更单元;

基于相关联的端点能力和 / 或偏好来从多种预定义标识方法中选择变更检测方法,所述变更检测方法用于描述所述数据集中的变更要被标识的方式;以及

根据所述内容类型、变更单元以及所选择的变更检测方法来跨两个或更多数据存储同步所述数据集。

11. 如权利要求 10 所述的方法,其特征在于,还包括利用同步状态信息来同步所述数据集。

12. 如权利要求 10 所述的方法,其特征在于,还包括利用标识相同数据结构的上下文专用标识符来同步所述数据集。

13. 如权利要求 10 所述的方法,其特征在于,还包括利用与所述变更单元相关联的内容不可知论版本信息来同步所述数据集。

14. 如权利要求 10 所述的方法,其特征在于,还包括从公开方案获取所述内容类型和变更单元。

15. 如权利要求 10 所述的方法,其特征在于,还包括携带不支持的同步数据。

16. 一种便于同步的系统(500),包括:

用于根据定义如何交换数据的同步模型和描述要交换的数据的数据方案来在两个或更多端点之间协商同步协议的装置(532);

用于作为协商的一部分,在两个或更多个端点之间交换变更检测模式的装置;以及用于根据所协商的同步协议来配置同步的装置(534)。

17. 如权利要求 16 所述的系统,其特征在于,还包括用于根据所协商的协议跨所述两个或更多个端点来同步数据的装置(120)。

数据 - 驱动同步

[0001] 背景

[0002] 现今存在多种应用程序,它们利用两种通用模型中的一种来在两端点之间同步数据,即将数据储存库从一个位置复制到另一个位置或在主机存储和客户机存储之间传输数据。第一种情况的经典示例涉及跨数据库集群的或基于 web 的主机和基于客户机的应用程序之间的数据库复制。该模型的关键是从软件开发的观点看,数据通常在处于数据管理系统的完全控制之下的参与者之间在明确定义的模式内传输。第二种情况(在该上下文中称为跨 - 存储同步)更脆弱,因为它通常涉及在不具有对数据集内的项的相同支持级别的端点之间传输数据。一个经典示例是在台式计算机和诸如移动电话或便携式媒体播放器等便携式设备之间进行同步。

[0003] 为应对跨 - 存储同步的挑战,常规解决方案是在连接主机和客户机设备的自定义协议的顶层创建应用程序栈。同步栈供应商通常还为主机和客户机设备编写一个或多个同步提供程序以使该管道能够携带来自诸如电子邮件程序等应用程序的数据。例如,一旦这些提供程序被安装在台式计算机和便携式设备上,则通过该管道的同步成为可能的。当然,这些解决方案所及的范围限于栈供应商愿意执行的工作量。例如,如果栈供应商不提供对主机上的特定数据源的支持,则将来自该源的内容同步到客户机就是不可能的。

[0004] 栈供应商有时使系统对第三方可用以允许他们也创建在管道的两端工作的提供程序并从而鼓励更宽的同步生态系统。然而,因为该管道的所有方面都由栈供应商编写,所以解决方案被锁定于特定供应商并且被限于供应商使该栈可用的场合。

[0005] 为此,希望安装用于将他们的应用程序与任何便携式设备进行同步的解决方案的大多数供应商或希望支持应用程序的同步的便携式设备制造商,往往购买第三方解决方案,开发他们自己的解决方案,或简单地决定不提供这一功能。

[0006] 概述

[0007] 以下呈现了简化的概述,以提供对所公开的主题的某些方面的基本理解。本概述不是广泛性的概观。它并不旨在标识关键 / 重要元素,也不旨在描绘所要求保护的题目的范围。其唯一的目的是以简化的形式来介绍一些概念,作为稍后提出的更为详细的描述的序言。

[0008] 简言之,本发明涉及数据 - 驱动同步。提供了一种不由应用程序和或 / 端点绑定的宽泛的同步系统和相关联的方法。不同于在同步管道的两侧需要来自单个供应商的专用代码的应用程序栈或基于 - 筒仓 (silo) 的方法,提出了一种定义如何交换信息来实现跨数据集和 / 或数据集源的基于模型的方法。以此方式,任何源可以与任何目的地进行同步。

[0009] 根据本发明的一个方面,可以根据端点所预订的同步模型的功能来执行同步。为了同步目的,该同步模型使用特定协议来定义数据和如何交换该数据,各端点协商来建立启用通信和 / 或交互的、相对于数据同步的特定同步模型。随后,可以配置各同步组件来实现关于同步和数据交换的指定行为。

[0010] 为实现上述及相关目的,在此结合以下描述和附图描述了所要求保护的题目的某些说明性方面。这些方面指示可实践本主题的各种方式,它们均落在所要求保护的题目的

范围之内。当结合附图阅读以下详细描述时,本发明的其它优点和新颖特征将变得清楚。

[0011] 附图简述

[0012] 图 1 是根据所公开的主题的一方面的数据 - 驱动同步系统的框图。

[0013] 图 2 是启用数据同步的典型的同步组件的框图。

[0014] 图 3 是典型的数据方案组件或其输出的框图。

[0015] 图 4 是典型的同步模型 / 方案组件或其输出的框图。

[0016] 图 5 是根据所公开的主题的一方面的用于在两端点之间同步数据的系统的框图。

[0017] 图 6 是根据所公开的主题的一方面的同步模型的框图。

[0018] 图 7 是根据所公开的主题的一方面的数据同步方法的流程图。

[0019] 图 8 是根据所公开的主题的一方面的同步方法的流程图。

[0020] 图 9 是根据所公开的主题的一方面的利用数据方案的同步方法的流程图。

[0021] 图 10 是根据本发明的利用数据和同步方案的同步方法的流程图。

[0022] 图 11 是示出用于本发明各方面的合适的操作环境的示意性框图。

[0023] 图 12 是示例计算环境的示意性框图。

[0024] 详细描述

[0025] 在下文中提供了实现源和目的地不可知 (agnosticism) 的数据同步的系统和方法。结果,应用程序和端点能够在不绑定到应用程序和 / 或端点专用解决方案的情况下提供同步功能。相反,同步被绑定到数据和一组同步行为。更具体地,同步可以根据数据方案来执行并用同步模型来优化。使用在丰富但基本的通信协议上层构建的附加层来定义数据以及如何交换该数据以用于同步目的。此外,可以在任何源和任何目的地之间协商数据方案和同步模型,以便跨不同的存储和 / 或端点来动态地配置最优同步。

[0026] 现在参考附图描述本发明的各个方面,在全部附图中用相同的标号来指示相同的或相应的元素。然而应该了解,附图及其相关详细描述不旨在将所要求保护的主题限于所公开的具体形式。相反,其意图是覆盖落在所要求保护的题目的精神和范围内的所有修改、等效和替换方案。

[0027] 最初参考图 1,提供了根据所要求保护的题目的一方面的数据 - 驱动同步系统 100。系统 100 包括允许多个数据集 112(1 到 N,其中 N 是大于 1 的整数)之间的通信的管道组件 110。作为示例,管道组件 110 可对应于,包括、和 / 或实现允许从数据集 112 中提取信息的数据传输机制或协议(例如,媒体传输协议 (MTP))。数据集 112 可以嵌入在多个相关联的端点或设备内,这些端点或设备诸如移动设备、台式计算机、和服务器、以及其他。管道组件 110 通信耦合到同步组件 120 并可操作用于向其供应或提供数据。同步组件 120 使用管道组件 110 来便于两个或更多数据集 112 的同步。如在下文进一步描述的,可以在被标识为主机和客户机的端点之间执行同步,其中主机可以跨一个或多个客户机来驱动同步。

[0028] 根据所要求保护的题目的一方面,系统 100 可以同步不由相关联的应用程序和 / 或端点绑定的数据集 112。换言之,数据同步不限于相关应用程序或设备供应商所提供的功能。相反,同步是数据 - 驱动的或由数据和可任选的一组同步行为绑定。

[0029] 考虑跨 - 存储场景中通常使用的常规筒仓 (silo) 同步系统,其中对于数据集内的各项,存储不包括相同的支持级别(例如,准确性损失)。在该系统中,提供程序提供可插入

到筒仓顶层和筒仓底层的代码,以确保所有必需的同步功能在其之间发生。因为代码在两侧(例如,源和目的地)运行,所以确保正确地处理所有准确性问题是提供程序的责任。

[0030] 该常规系统有几个问题。首先,通用应用程序完全受同步筒仓作者支配,以跨这些应用程序同步数据。附加地或另选地,端点同样由给定同步筒仓绑定。作为响应,可以开发自定义同步解决方案或可以不提供同步。在此,系统 100 提供实现源应用程序和目的地不可知的、用于同步数据的机制。结果,应用程序和/或便携式设备供应商有机会以最低成本提供所需功能,而非创建他们自己的解决方案或特许的限制性解决方案。

[0031] 如根据下文进一步描述将理解的,为实现应用程序和/或端点独立性的丰富级别并鼓励来自多个不同供应商的强支持,可以使用通用的良好-定义的标准解决方案。此外,为与各种应用程序和或端点合作,将传统上构建的筒仓同步栈作为模型一般是适当的。或者,并非聚焦于被同步的应用程序和/或端点(例如,筒仓的顶层和底层),同步可以聚焦于端点通信的中途,以允许丰富同步体验(例如,筒仓的中间层)。在一个实现中,可以提供标准化通信机制以鼓励在顶层和底层进行开发来提供丰富的同步选项。

[0032] 转而关注图 2,描绘了根据所要求保护的的主题的一方面的典型的同步组件 120。该同步组件包括数据分析组件 210 和更新组件 220。这两个组件 210 和 220 负责基本同步功能。更具体地,数据分析组件 210 分析两个或更多个数据集来确定需要如何更改这些集合(如果需要的话)来同步这些集合。更新组件 220 便于根据数据分析组件 210 所提供的分析来更改一个或多个数据集。

[0033] 另外,同步组件 120 包括影响数据分析组件 210 的数据方案组件 230 和同步模型/方案组件 240。数据方案组件 230 定义要交换的数据。同步模型组件 240(在此也被称为同步方案组件)指定要如何交换数据以用于同步目的。换言之,可以向数据分析组件 210 提供数据和同步方案来定制其操作。具体地,这些方案可以控制交换什么数据以及如何交换该数据来同步两个或更多个数据集。

[0034] 在一种情况下,数据方案组件 230 和同步模型组件 240 可以例如从网站或其他网络位置获取方案来提供给数据分析组件 210。或者,组件 230 和 240 可对应于方案本身。在任一种情况下,这些方案可以用多种不同的方式来表示,包括但不限于 XML(可扩展标记语言)或编译时间生成的二进制结构。

[0035] 图 3 示出根据所要求保护的的主题的一方面的典型的数据方案组件 230 或其输出。数据集不总是包括关于同步的相同信息或准确性级别。例如,与个人信息管理器(PIM)应用程序相关联的移动设备存储通常不包括与台式存储相同的数据,尤其是在这些应用程序由不同的制造商创建的情况下。为解决这一问题,内容组件 310 可以描述有待同步的数据。这使得同步主机和客户机能够讲相同的数据交换语言。

[0036] 作为示例,讲联系人对象项目的同步主机不能够与讲日历对象的客户机进行同步。然而,如果主机和客户机讲同一内容类型(如联系人对象),则它们应该能够参与同步关系。注意,这里假定了内容类型暗示该数据的具体格式。换言之,如果两个端点都指示它们讲“联系人”,则假定它们引用同一联系人格式(例如,名、姓、街道地址、城市、州、以及邮政编码)。这不是严格需要的但简化了同步。在必要的情况下可以执行转换来将不同的格式映射到特定标准化和/或公布的格式。

[0037] 数据方案组件 230 还包括用于标识与内容组件 310 所描述的数据相关联的变更的

单元的变更组件 320。足以简单地说,因为同步主机和客户机都讲同一内容类型所以可以执行同步。然而,为最大化同步性能,只有被变更过的最小量数据应当从一个端点传送到另一个端点。在某些情况下,将数据集之中的特定项分为变更单元是有意义的,因为它们往往频繁变更,同时在分组的情况下可更高效地管理其他项。例如,电子邮件地址可频繁变更并且因此被指定为分开的变更单元,而包括街道名称、城市、州、以及邮政编码的物理地址可被分组成单个变更单元。

[0038] 图 4 描绘根据所要求保护的主题的一方面的典型的同步模型 / 方案组件 240 或其输出。如上所述,同步模型组件 240 标识要如何交换数据来实现同步。典型的组件 240 包括状态组件 410 和变更检测组件 420。

[0039] 状态组件 410 是用与以内容 - 类型无关的方式描述同步的当前状态的模型。考虑以所示次序执行以下操作的一般化同步操作:

[0040] 1. 确定从与第二端点的最后同步操作以来而在第一端点上变更了什么。

[0041] 2. 通过检查该第二端点上的对象来确定该第二端点是否具有第一端点上已有的变更。

[0042] i. 如果数据是相同的,则忽略变更。

[0043] ii. 如果数据不是相同的,并且只变更了一侧,则从一侧向另一侧同步最新近的变更。

[0044] iii. 如果不是相同的并且变更单元在两侧都已变更,则使用某种形式的查看器来解决这两个变更之间的冲突。

[0045] 3. 将变更从第一端点复制到第二端点。

[0046] 在没有关于一端点上的同步状态的丰富信息的情况下,执行动作 2 的唯一方式通常是列举变更集合(来自动作 1)中的每一项,发现什么已变更并随后在第二端点上定位对应的项(在不共享共同同步 ID 的情况下,这可退化成列举并查找匹配),并且随后检查变更单元中的每一属性以查看它们是否有不同。对于简单的属性或数据集,该比较可能不是性能的考虑,但对于较大属性或数据集,该比较能是非常昂贵的。通过引入同步的当前状态,图 2 的数据分析组件 210 可以查询状态存储而非直接查询对象,以确定是否需要进行变更。

[0047] 变更检测组件 420 描述在一侧或另一侧标识被变更的方式。标识变更存在着许多可能性。例如,无论哪一个端点驱动同步操作都可以利用全列举模型来全面地询问另一存储。在基于 - 锚的方法中,存储必须管理或实际上提供变更或增量(delta),但这些增量可能不是用于要发生的同步操作的最优形式或语言。或者,与基于 - 锚的模型类似,可以提供增量,但以最优或基本上最优的方式来提供,这是相对于尤其是管理同步元数据的方式而言。还有,在服务模型中,一切事物都可被打包并发送到进行所有工作并返回结果的服务(例如,web 服务)。

[0048] 同步模型组件 240 还包括用于进一步便于同步的标识符组件 430 和版本组件 440。标识符组件 430 为诸如对象等数据的集合提供上下文专用标识符。在没有标识符的情况下定位对象的成本相当昂贵。添加标识符使其更容易识别,这样所引用的对象是相同的而不必比较附加数据。可以明白,这样的同步标识符(即 ID)可以是方案中已经存在的特定一栏或需要映射功能来基于所提供的数据确定 ID。

[0049] 版本组件 440 描述内容 - 类型不可知论共同格式,用于交换与同每一数据结构相

关联的每一变更单元的版本的信息有关。版本的引入,就能通过比较版本而非比较数据极大地简化了确定是否对同步操作的两侧应用了变更的过程。具有版本暗示了至少两个概念。第一,存在用于版本信息的共同格式,以及第二,该格式能在每个对象基础上以变更单元粒度而在同步客户端点上存储,并可以从中检索。

[0050] 图 5 描绘根据所要求保护的主题的一方面的用于在两端点之间同步数据的系统 500。系统 500 包括端点 510 和 520。此外,端点 510 和 520 可对应于两个移动设备或一个移动设备和一个台式计算机或服务器。每一端点都包括可利用同步组件 120 来同步的本地数据存储 512。

[0051] 每一同步组件 120 包括用于协商同步协议和 / 或建立同步关系的协商组件 532。诸如基于 COM 的对象交换接口或媒体传输协议 (MTP) 等数据传输协议,此外可用作端点通信的基础。协商组件 532 便于标识关于同步的特定规则或过程。因此,同步模型可以在支持所需同步的丰富的其他通信协议的顶层构建。

[0052] 同步协议协商在本质上是动态的。并非利用用于同步的固定或严格的通信方案和接口集合 (常规情况下),各端点可以协商协议或关系。此外当各端点在对话中显现时,它们相对于同步来通信的方式可以在此时根据能力和 / 或性能来确定。

[0053] 一旦经由协商组件 532 建立了同步协议,则同步组件 120 可以用配置组件 534 来配置以便以特定方式交互。因此,同步组件 120 可包括通用功能,该通用功能适用于或可配置成根据模型而非取决于特定预编程同步行为来执行同步。

[0054] 协商和后续配置可以取决于端点 510 和 520 所预订的一个或多个方案 540。根据一个方面,可以公布方案 540 或使其公开,例如作为 XML 文档。方案 540 可包括数据方案 542 和同步方案 544 子组件。数据方案组件 542 描述要在同步过程中交换的数据,包括内容类型和相关联的变更单元。同步方案组件 544 描述要如何交换数据来同步数据集。此外,例如,同步方案可包括描述当前同步状态和变更检测的一个或多个模型。

[0055] 在协商和后续同步期间,端点 510 和 520 可以交换不同类型的信息。如图所示,可以交换与数据的内容类型或与该数据所符合的方案的标识有关的信息。还传送与端点支持的变更检测模式和当前状态模型有关的信息。该信息可用于配置端点之间的同步。在必要的情况下,可以在端点 510 和 520 之间交换对象或其他数据结构。另外,可以传递指令或操作来便于在设备上执行。

[0056] 同步效率可至少部分地取决于同步操作中的每一参与端点的能力水平。端点只需标识定义内容类型的数据方案和要参与同步操作的变更信息的粒度。如果端点只能提供这一信息,则创建一个同步解决方案是可能的,其中另一端点自己列举该受限端点内的项、发现并应用变更。在该 (发明) 范围的另一头,在两个端点能够例如基于状态信息独立地标识变更并确定需要传送哪一些的情况下,它们可以只交换变更。在此,同步可以很快速并且高效,因为所传送的额外数据量由变更单元定义而非同步操作来绑定。在该 (发明) 范围的两头之间,存在多个混合模型。对于能够参与同步关系的每一端点,应能定义它自己的能力应对它从另一端点需要的能力,以便可以创建同步伙伴关系,以使工作尽可能高效。

[0057] 根据所公开的主题的一方面,可以使方案 540 标准化。通过标准化数据或对象格式和相关联的同步机制,可以启用从任何源到任何目的地的同步。开发商、制造商、供应商等可以就多个数据和 / 或同步方案达成一致,以允许更广泛的执行同步的方式。例如,标准

联系人对象可包括用于名、姓、街道名称、州、城市、以及邮政编码的字段。如果所有或许多供应商等利用该用于联系人对象的格式，则将启用这些对象跨应用程序和 / 或体系结构的同步。

[0058] 可以表述附加方案项以供参与同步关系。开发标准化解决方案的最大挑战之一是一就交换的项的总体达成一致。并非相信可以定义满足想要参与同步关系的每一方的定义的数据方案，而是该方案能够支持可扩展性以允许添加新项，例如来支持自定义同步模型。例如，商务联系人管理器可能不具有即时消息收发地址的概念，因为即时消息收发在商务通信中一般不使用；然而，聚焦于消费者的联系人管理器通常包含这样的字段。如果用于联系人的原始数据方案在没有定义这一字段，则面向消费者的版本的、指示其支持标准方案和添加即时消息收发字段的能力使得理解该即时消息收发字段的其他应用程序能够与它进行同步。此外，某些客户机可被配置成数据出租车 (taxi)，它能够携带诸如即时消息收发字段而不必实际上理解该字段的附加同步元素。

[0059] 还可以明白，可以经由标准化过程更强地定义数据传输机制或同步协议的一个或多个方面。例如，台式计算机平台可以定义特定模型，用于如何注册不同同步组件来指示它们都支持相同的方案 540。在该模型内，可以强定义变更检测和状态管理的概念来启用最优交换模式。然而，尽管可以通过定义用于各个对象的默认数据传输机制来实施基本级别的互操作性，但对该系统内的两个端点而言，协商基于对附加方案项的共同知识的更丰富的数据交换模式或更丰富的数据传输模型也是可能的。

[0060] 该模型还具有允许引入代理的好处。返回到图 1，注意到，管道组件 110 能够附连到多个数据集 112。在交换模式固定于台式的系统中，出于在所描述的台式系统内同步的目的，可以引入能够从多个数据集 112 拉 (pull) 数据的数据不可知论代理组件 514，这些数据集之中的每一个都可能代表不同的方案集合 540。如图所示，数据存储 512 可以有选择地与同步组件 120 直接耦合或经由代理组件 514 耦合。

[0061] 图 6 示出根据所要求保护的的主题的一方面的纲领性同步模型 600。模型 600 宽泛地定义一组功能，该组功能在被实现时，经由一个或多个组件启用存储在多个端点上的不由相关联的应用程序和 / 或体系结构绑定的数据的同步。从底层开始并向上移动，模型 600 包括数据传输层 610、同步方案层 620、以及数据方案层 630。

[0062] 数据传输层 610 启用端点之间的通信。作为示例而非限制，可以使用启用对象以及各种其他设备和内容信息的交换的诸如 COM 接口或 MTP 等丰富的通信协议。因此，层 610 提供用于跨各端点或端点设备的数据交换的基础或基底。

[0063] 下一层是定义关于实现同步的数据交换的同步方案或模型层 620。因此，层 620 提供针对关于通过原始数据传输层 610 交换的数据的限制的附加同步。同步方案层可以定义和 / 或支持各种数据交换模型和机制。例如，同步层 620 可以提供用于描述当前状态的内容 - 类型无关模型、用于数据结构的上下文专用标识符、用于交换每一变更单元的版本信息的内容 - 类型无关格式、用于表达附加方案项已参与同步关系的方式、用于指示准确性的可用其维护对象方案中的特定项的方法、和 / 或用于指示每一参与端点的同步操作能力的水平的方式。

[0064] 最后，数据方案层 630 描述要在同步操作中进行交换的数据。针对这一点存在两个方面：内容 - 类型或对象方案以及变更单元粒度。换言之，数据的格式和与该数据相关联

的变更的单元。

[0065] 为帮助进一步清晰关于模型的利用,考虑其中从移动电话到台式计算机来同步联系人数据的同步场景。每一设备包括能支持或不支持同步的联系人存储。这些设备能够跨对象交换协议来交换根据对象方案定义的对象。交换对象的方式可以取决于移动设备所实现的同步方案。如果该设备实现变更检测模型和 / 或其他同步功能,则它可以根据这一功能检测并向台式组件返回变更。或者,移动设备可不实现任何同步方案,例如,该设备供应商不想投资这一技术并强制台式设备完全列举它。

[0066] 已经关于一些组件之间的交互描述了上述系统、体系结构等。应该明白,此类系统和组件可以包括在此所指定的那些组件或子组件、所指定组件或子组件中的一部分、和 / 或另外的组件。子组件也可以被实现为在通信上被耦合到其他组件而不是被包括在父组件中的组件。此外,一个或多个组件和 / 或子组件可以结合成提供聚集功能的单个组件。系统、组件、和 / 或子组件之间的通信可以根据推 (push) 和 / 或拉模型来实现。各组件也可以与一个或多个其他组件进行交互,出于简要考虑在此未具体描述该组件但本领域的技术人员均已知。

[0067] 此外,应该明白,以上公开的系统以及以下方法的不同部分可以包括或包含基于人工智能、机器学习或知识或规则的组件、子组件、进程、装置、方法、或机制 (例如,支持向量机、神经网络、专家系统、贝叶斯信任网络、模糊逻辑、数据融合引擎、分类器……)。此类组件和其他组件可以自动化地执行特定机制或进程,由此使得系统和方法的各部分变得更加自适应、高效及智能。作为示例而非限制,这些机制可以相对于同步协议的协商和 / 或同步组件的配置来使用。例如,可以根据显式或隐式收集的设备类型或其他上下文信息来推断端点能力和 / 或偏好,并相应地配置同步。

[0068] 考虑到以上描述的示例性系统,参考图 7-10 的流程图将可以更好地理解依照所公开的主题实现的方法。尽管出于说明简单的目的,各方法被显示和描述为一系列框,但应该理解和领会,所要求保护的主题不受框次序的限制,因为一些框能够以不同的次序和 / 或与在此描绘和描述的其它框并发地发生。而且,并非所有示出的框都是实现以下描述的方法所必需的。

[0069] 参考图 7,描绘了根据所要求保护的主题的一方面的数据同步方法 700。在参考标号 710 处,从两个或更多个源获取数据。在 720 处,随后这些源被同步,并不由相关联的应用程序、端点、体系结构等绑定。焦点在于要同步的数据而非要同步的应用程序或端点。换言之,同步是数据 - 驱动的。

[0070] 常规上并如上所述,同步限于范围受限的自定义应用程序同步解决方案。如果这样的解决方案不提供对特定数据源或数据集之中的元素的支持,则不能同步该数据。同步供应商有时使其系统对第三方可用以允许他们创建在管道的两端工作的提供程序来鼓励更宽的同步生态系统。然而,这将同步锁定于特定供应商并限制了这样的同步的范围。因此,设备供应商必须放弃同步功能或招致生成他们自己的自定义解决方案的成本。

[0071] 方法 700 提供在某种意义上与这些因素无关的达到同步的更宽的方法。在一个实现中,使用定义如何交换信息以实现跨数据集同步的基于协议的方法。结果,例如与特许来自第三方的软件相比,进入的唯一障碍是协议和一组标准的简单实现。换言之,与特定应用程序或体系结构相比,同步被绑定到数据和一组同步行为。

[0072] 图 8 是根据所要求保护的的主题的一方面的同步方法 800 的流程图。在参考标号 810 处,例如利用诸如 MTP 等基本通信协议来通信地连接各数据源。接着,在标号 820 处,协商同步协议。在一种情况下,这可对应于标识对话中的每一端点实现了的多个标准或预设方案、模型、和 / 或机制中的哪一个,以及基于此来确定用于高效地同步数据的总体模型。可以单独地基于选择端点数据集所支持的最高层或最丰富的同步模型和 / 或利用诸如端点偏好和上下文(例如,负载、处理能力……)等附加信息来选择该模型。在标号 830 处,可以根据所协商的通信协议来交换数据并跨两个或更多个源来同步。

[0073] 图 9 描绘根据所要求保护的的主题的一方面的利用数据方案的同步方法 900。在标号 910 处,出于同步目的,使用预定或所协商的通信协议,数据被获取或以其他方式被访问。在参考标号 920 处,确定数据内容类型。可以从与该数据相关联的端点和 / 或外部方案直接获取该信息。此外,根据一个方面,将内容 - 类型标准化以便于与应用程序和 / 或体系结构无关的数据交换。例如,内容类型可以是与具有定义要同步的信息的各方面的相关联的方案文档的联系人对象或媒体文件。在参考标号 930 处,确定与内容类型相关联的变更单元,例如是以与内容类型本身相同或相似的方式。作为示例,邮件地址可构成单个变更单元,这就使得街道、城市、州、以及邮政编码总是作为一个组来交换,而电子邮件地址可以是分开的变更单元。在标号 940 处,利用数据方案来同步数据。例如,同步端点可以列举该数据、标识变更、并相对于其数据来实现这些变更。

[0074] 图 10 是根据所要求保护的的主题的一方面的利用数据和同步方案来同步数据的方法 1000 的流程图。在标号 1010 处,获取或以其他方式访问数据。在参考标号 1020 处,确定数据方案,例如包括内容类型的标识和变更单元粒度。在标号 1030 处,确定描述当前同步状态的内容类型不可知论模型。这可通过无需分析变更集合中的每一项而相反通过比较状态以发现变更来帮助同步过程。在标号 1040 处,确定用于在变更单元级标识数据版本的内容类型不可知论版本格式。在 1050 处,在不比较数据的情况下,上下文标识符格式被标识,以便于定位同一数据结构(如对象)。在标号 1060 处,诸如完全列举或基于一锚的方法以及其他等变更标识方法被标识。从动作 1030 到 1060 所确定的所有信息可以从与数据相关联的端点直接传递到其他端点,或通过外部方案的标识来间接地传递,等等。此外,该信息可对应于多个标准或公布方案、模型、类型等中的一种的标识。在参考标号 1070 处,根据所收集的数据方案和同步信息来同步数据。例如,仔细审查根据数据方案定义的数据,来确定根据变更检测方法以及状态、版本、和 / 或标识符信息和被交换的变更,以确保数据已被同步。

[0075] 在此所用的术语“组件”、“系统”等等意指与计算机相关的实体,其可以是硬件、硬件和软件的组合、软件、或执行中的软件。例如,组件可以是但不限于:在处理器上运行的进程、处理器、对象、实例、可执行代码、执行的线程、程序和 / 或计算机。作为说明,在计算机上运行的应用程序和计算机都可以是组件。一个或多个组件可以驻留在进程和 / 或执行的线程中,并且组件可以位于一个计算机内和 / 或分布在两个或更多的计算机之间。

[0076] 在本文中使用的词语“示例性”或其各种形式意味着用作示例、实例或说明。在此被描述为“示例性”的任何方面或设计并不一定要被解释为相比其它方面或设计更优选或有利。此外,各示例只是出于清楚和理解的目的来提供的并且并不意味着以任何方式限制或约束所要求保护的的主题或本发明的相关部分。可以理解,本可呈现不同范围的多个其它

或替换示例,但已出于简明的目的而省略了。

[0077] 如此处所使用,术语“推论”或“推断”通常指的是从经由事件和 / 或数据捕捉的一组观察结果来推理或推断系统、环境、和 / 或用户状态的过程。例如,推论可用于标识特定的上下文或动作,或可生成状态的概率分布。推论可以是概率性的,即,基于数据和事件的考虑计算感兴趣的状态的概率分布。推论也可以指用于从一组事件和 / 或数据组成更高级事件的技术。这类推论导致从一组观察到的事件和 / 或存储的事件数据中构造新的事件或动作,而无论事件是否在相邻时间上相关,也无论事件和数据是来自一个还是若干个事件和数据源。各种分类方案和 / 或系统(例如,支持向量机、神经网络、专家系统、贝叶斯信任网络、模糊逻辑、数据融合引擎……)可结合执行关于本发明的自动化和 / 或推断的动作来采用。

[0078] 此外,本发明的所有或部分可实现为方法、装置或制品,这是使用标准编程和 / 或工程设计技术而生产的软件、固件、硬件或其任何组合并用以控制计算机实现所公开的发明。如在此使用的术语“制品”旨在涵盖可以从任何计算机可读设备或介质访问的计算机程序。例如,计算机可读介质可以包括但不限于磁存储设备(例如,硬盘、软盘、磁带……)、光盘(例如,紧致盘(CD)、数字多功能盘(DVD)……)、智能卡和闪存设备(例如,卡、棒、键驱动器……)。另外应该明白,可以使用载波来承载计算机可读电子数据,例如那些用于发送和接收电子邮件或用于访问如因特网或局域网(LAN)等网络的数据。当然,本领域的技术人员将会认识到,在不背离所要求保护的的主题的范围或精神的前提下可以对这一配置进行许多修改。

[0079] 为给所公开的主题的各方面提供上下文,图 11 和 12 以及以下讨论旨在提供可以在其中实现所公开的主题的各方面的合适的环境的简要、概括的描述。尽管以上在运行于一台或多台计算机上的程序的计算机可执行指令的一般上下文中描述了本主题,但本领域的技术人员将认识到,本发明也可结合其它程序模块实现。一般,程序模块包括执行特定任务和 / 或实现特定抽象数据类型的例程、程序、组件、数据结构等。此外,本领域的技术人员可以理解,系统 / 方法可用其他计算机系统配置实现,包括单处理器、多处理器或多核处理器计算机系统、小型计算设备、大型计算机、以及个人计算机、手持式计算设备(例如,个人数字助理(PDA)、电话、手表……)、基于微处理器或可编程消费者或工业电子设备等。所示各方面也可在任务由通过通信网络链接的远程处理设备中执行的分布式计算环境中实现。然而,所要求保护的的主题的一些方面,如果不是全部方面,可以在独立计算机上实施。在分布式计算环境中,程序模块可以位于本地和远程存储器存储设备中。

[0080] 参考图 11,用于实现此处所公开的各方面的示例性环境 1110 包括计算机 1112(例如,台式机、膝上型计算机、服务器、手持式、可编程消费或工业电子产品……)。计算机 1112 包括处理单元 1114、系统存储器 1116 和系统总线 1118。系统总线 1118 将包括但不限于系统存储器 1116 的系统组件耦合到处理单元 1114。处理单元 1114 可以是各种可用处理器中的任意一种。可以理解,双微处理器、多核、和其它多处理器体系结构也可用作处理单元 1114。

[0081] 系统存储器 1116 包括易失性和非易失性存储器。基本输入 / 输出系统(BIOS)包含诸如在启动期间在计算机 1112 的元件之间传送信息的基本例程,它被存储在非易失性存储器中。作为示例而非限制,非易失性存储器可包括只读存储器(ROM)。易失性存储器包

括随机存取存储器 (RAM), 它可被用作为外部高速缓冲存储器以帮助处理。

[0082] 计算机 1112 还包括可移动 / 不可移动、易失性 / 非易失性计算机存储介质。例如, 图 11 示出了大容量存储 1124。大容量存储 1124 包括但不限于, 诸如磁盘或光盘驱动器、软盘驱动器、闪存或记忆棒之类的设备。另外, 大容量存储 1124 可以包括独立的或者与其它存储介质结合的存储介质。

[0083] 图 11 提供用作用户和 / 或其它计算机之间以及在合适的操作环境 1110 中描述的基本计算机资源之间的中介的软件应用程序 1128。这种软件应用程序 1128 包括系统和应用程序软件的一个或两者。系统软件可以包括存储在大容量存储 1124 上的操作系统, 该操作系统用于控制并分配计算机系统 1112 的资源。应用程序软件通过存储在系统存储器 1116 和大容量存储 1124 的任一或两者之上的程序模块和数据来利用系统软件对资源的管理。

[0084] 计算机 1112 还包括通信地耦合到总线 1118 并帮助与计算机 1112 交互的一个或多个接口组件 1126。作为示例, 接口组件 1126 可以是端口 (例如, 串行、并行、PCMCIA、USB、火线……) 或接口卡 (例如, 声音、视频、网络……) 等等。接口组件 1126 可以接收输入和提供输出 (有线地或无线地)。例如, 输入可以从以下设备接收, 这些设备包括但不限于, 诸如鼠标、跟踪球、指示笔、触摸垫等定点设备, 键盘, 话筒, 操纵杆, 游戏手柄, 圆盘式卫星天线, 扫描仪, 照相机, 其它计算机等。还可以由计算机 1112 经由接口组件 1126 向输出设备提供输出。输出设备可以包括显示器 (例如, CRT、LCD、等离子……)、扬声器、打印机和其它计算机等等。

[0085] 图 12 是本发明可与其交互的示例计算环境 1200 的示意框图。系统 1200 包括一个或多个客户机 1210。客户机 1210 可以是硬件和 / 或软件 (例如, 线程、进程、计算设备)。系统 1200 还包括一个或多个服务器 1230。因此, 系统 1200 可对应于两层客户机服务器模型或多层模型 (例如, 客户机、中间层服务器、数据服务器) 以及其它模型。服务器 1230 也可以是硬件和 / 或软件 (例如, 线程、进程、计算设备)。服务器 1230 可以容纳各线程以通过例如利用本发明的各方面来执行转换。客户机 1210 和服务器 1230 之间的一种可能的通信可以是在两个或多个计算机进程之间传输的数据分组的形式。

[0086] 系统 1200 包括可以用来使客户机 1210 和服务器 1230 之间通信更容易的通信框架 1250。客户机 1210 可在操作上连接至一个或多个客户机数据存储 1260, 这可以用来存储对客户机 1210 本地的信息。同样地, 服务器 1230 可在操作上连接到一个或多个服务器数据存储 1240, 这可以用来存储对服务器 1230 本地的信息。

[0087] 客户机 / 服务器交互可以相对于此处描述的同步一起使用。作为示例, 包括数据存储 1260 的诸如移动设备 (例如, 电话、个人数字助理、膝上型计算机、媒体播放器……) 等客户机 1210 可利用通信框架 1250 来寻求与服务器 1230 和相关联的数据存储 1240 进行同步。例如, 对个人的膝上型计算机上的数据进行的改变可以与业务服务器上的对应数据进行同步。而且, 客户机 1210 可以在一个或多个服务器数据存储 1240 上发布和 / 或访问标准方案、模型、协议等以便于同步引擎 / 组件配置。

[0088] 以上所已经描述的内容包括所要求保护的主题的各方面的示例。当然, 出于描绘所要求保护的主题的目的而描述每一个可以想到的组件或方法的组合是不可能的, 但本领域内的普通技术人员应该认识到, 所要求保护的主题的许多进一步的组合和排列都是可能

的。从而,所公开的主题旨在涵盖落入所附权利要求书的精神和范围内的所有这样的变更、修改和变化。而且,对于在详细描述或权利要求书中使用术语“包括”、“包含”、“具有”、“含有”或其它形式的变型而言,这样的术语旨在包含在类似于术语“包括”的一种方式中,如同“包括”在用作权利要求书中的过渡词时所解释的那样。

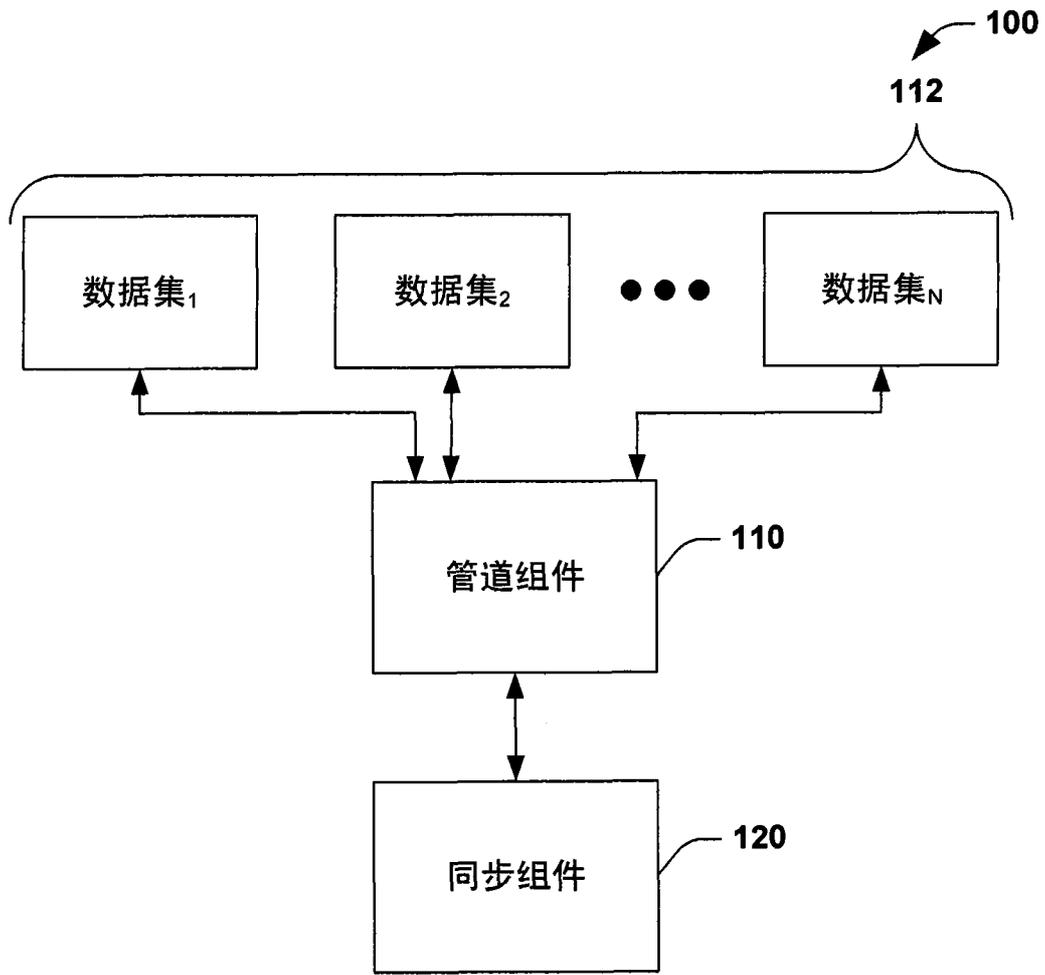


图 1

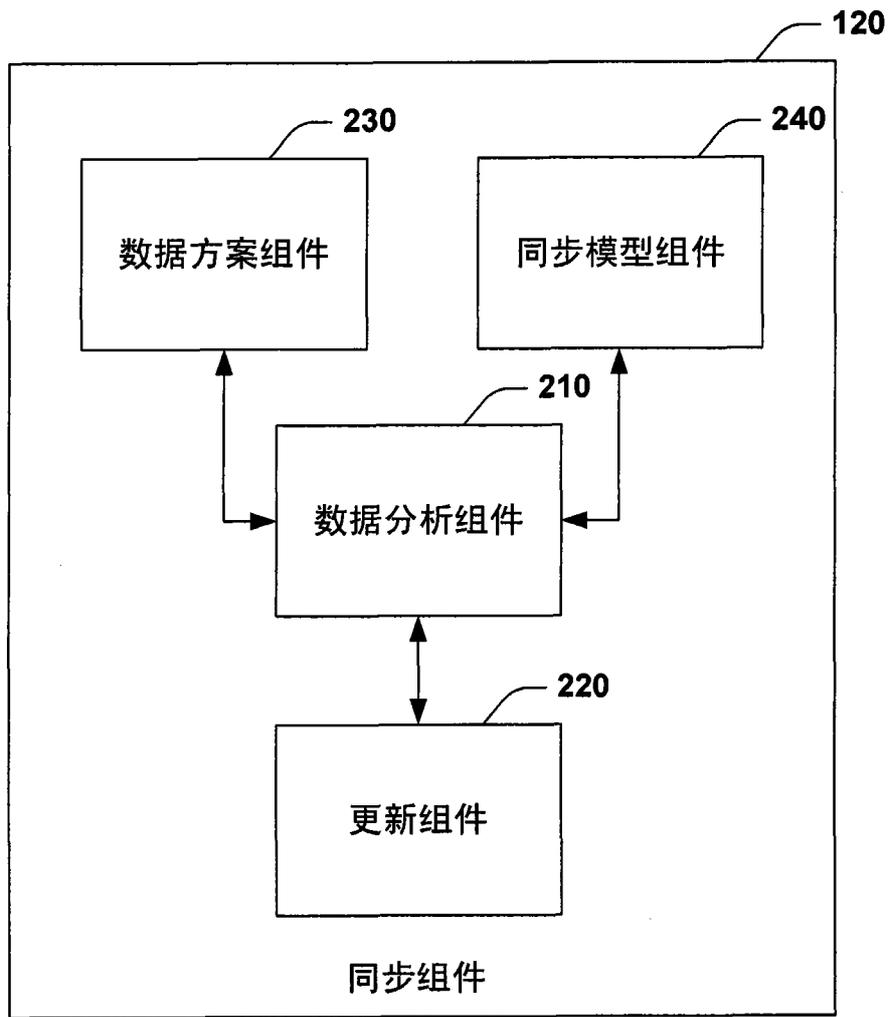


图 2

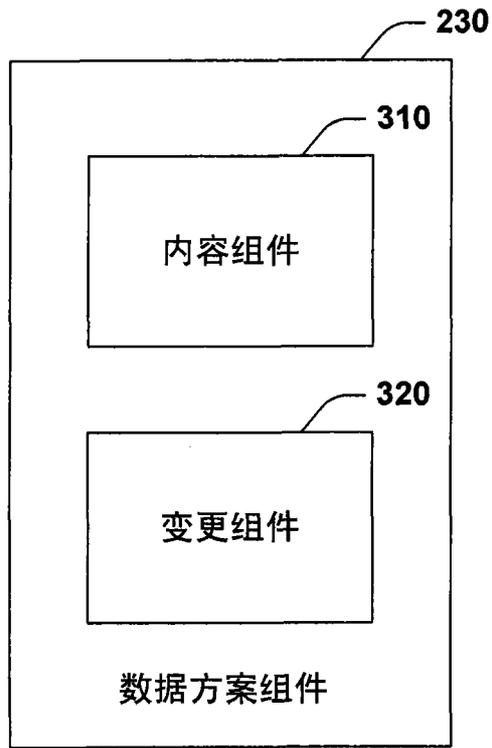


图 3

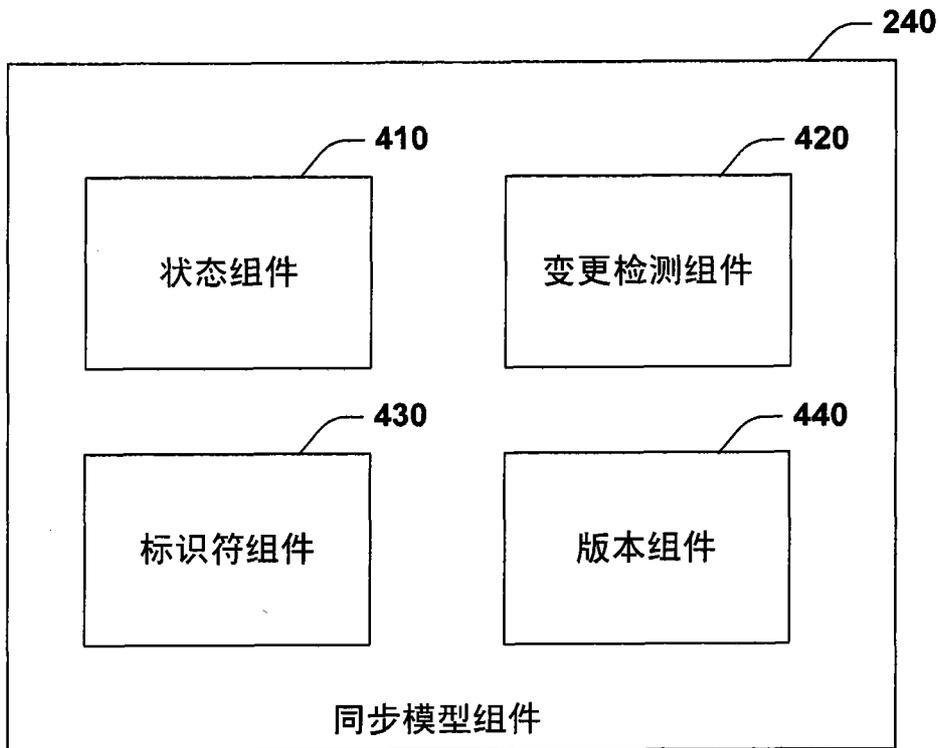


图 4

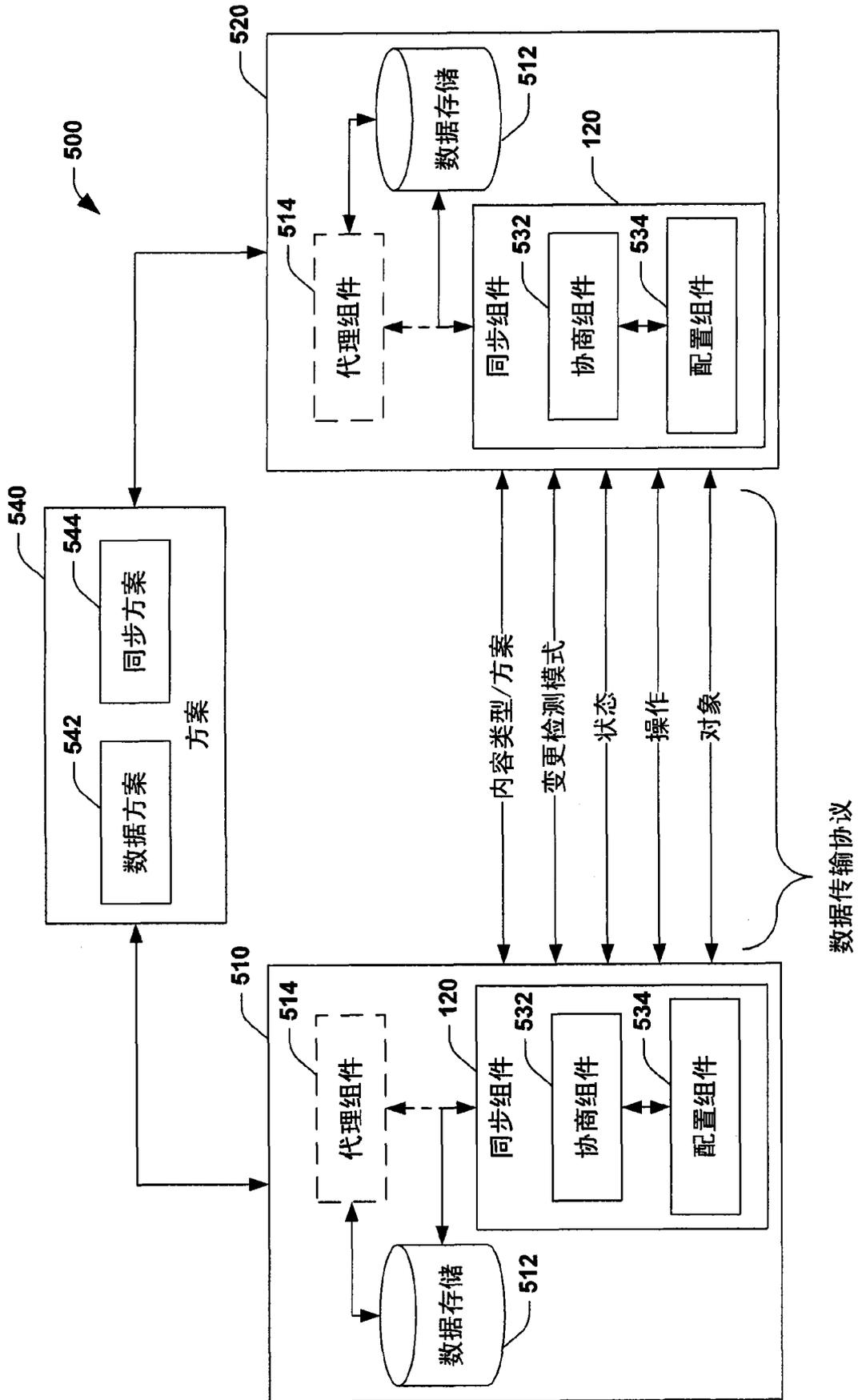


图 5

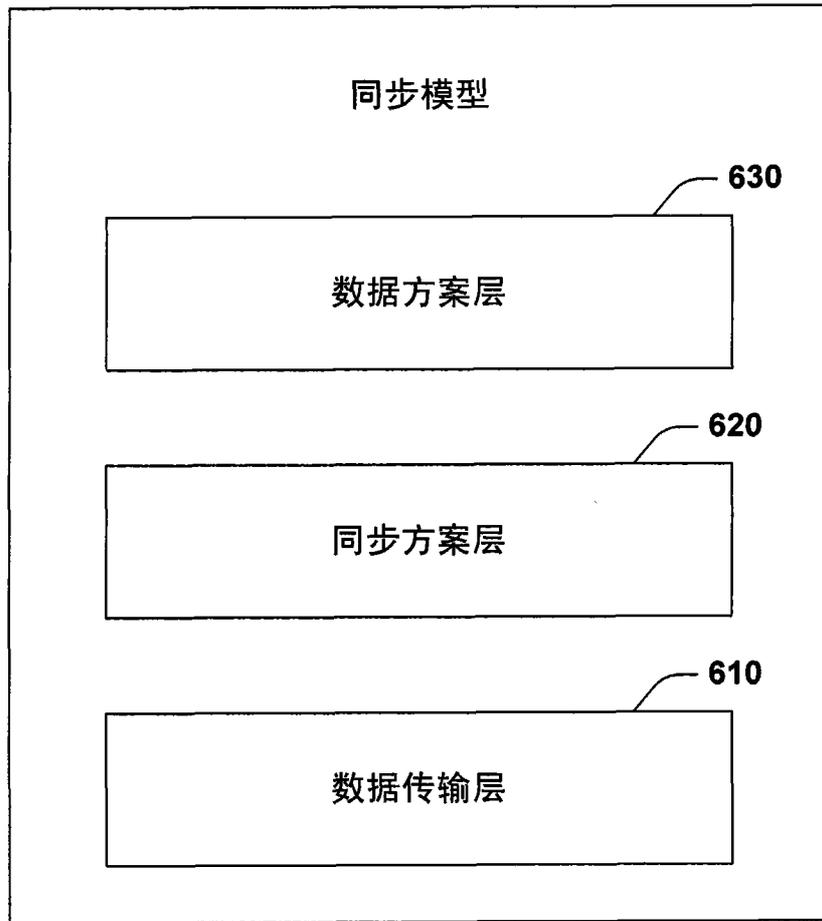


图 6

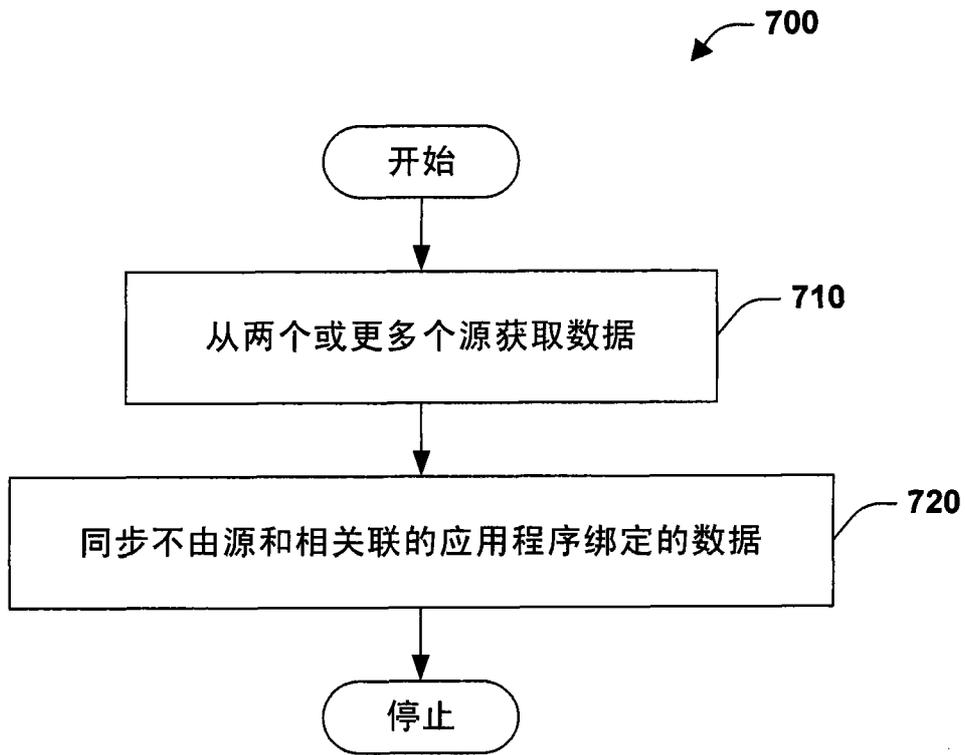


图 7

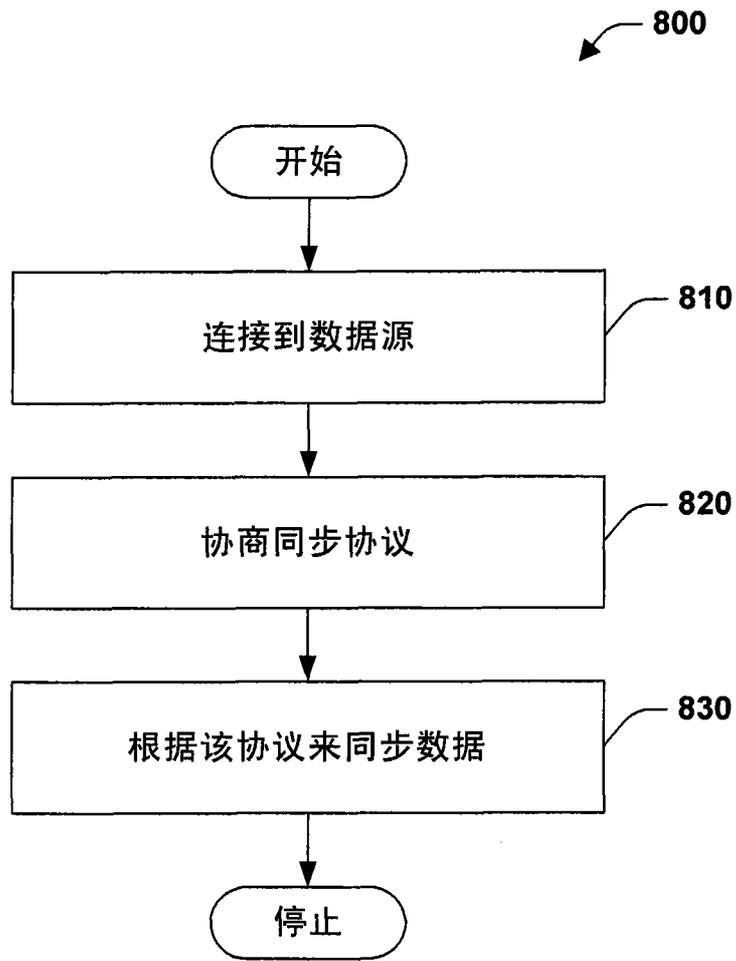


图 8

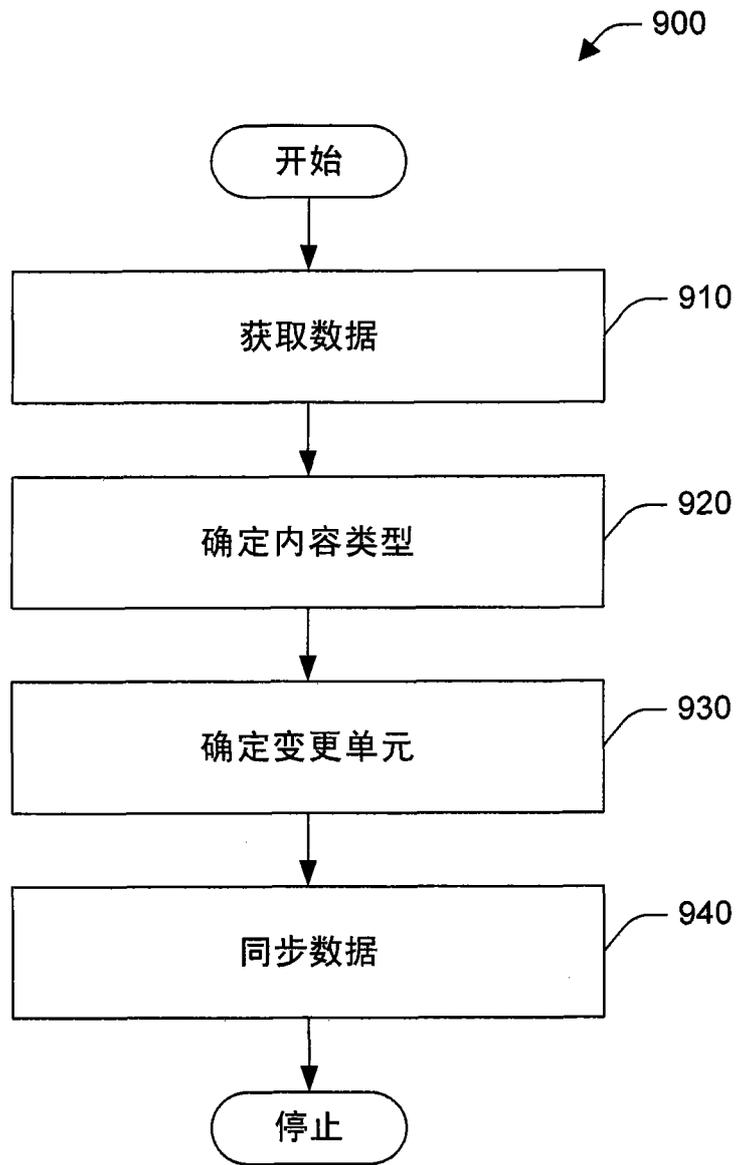


图 9

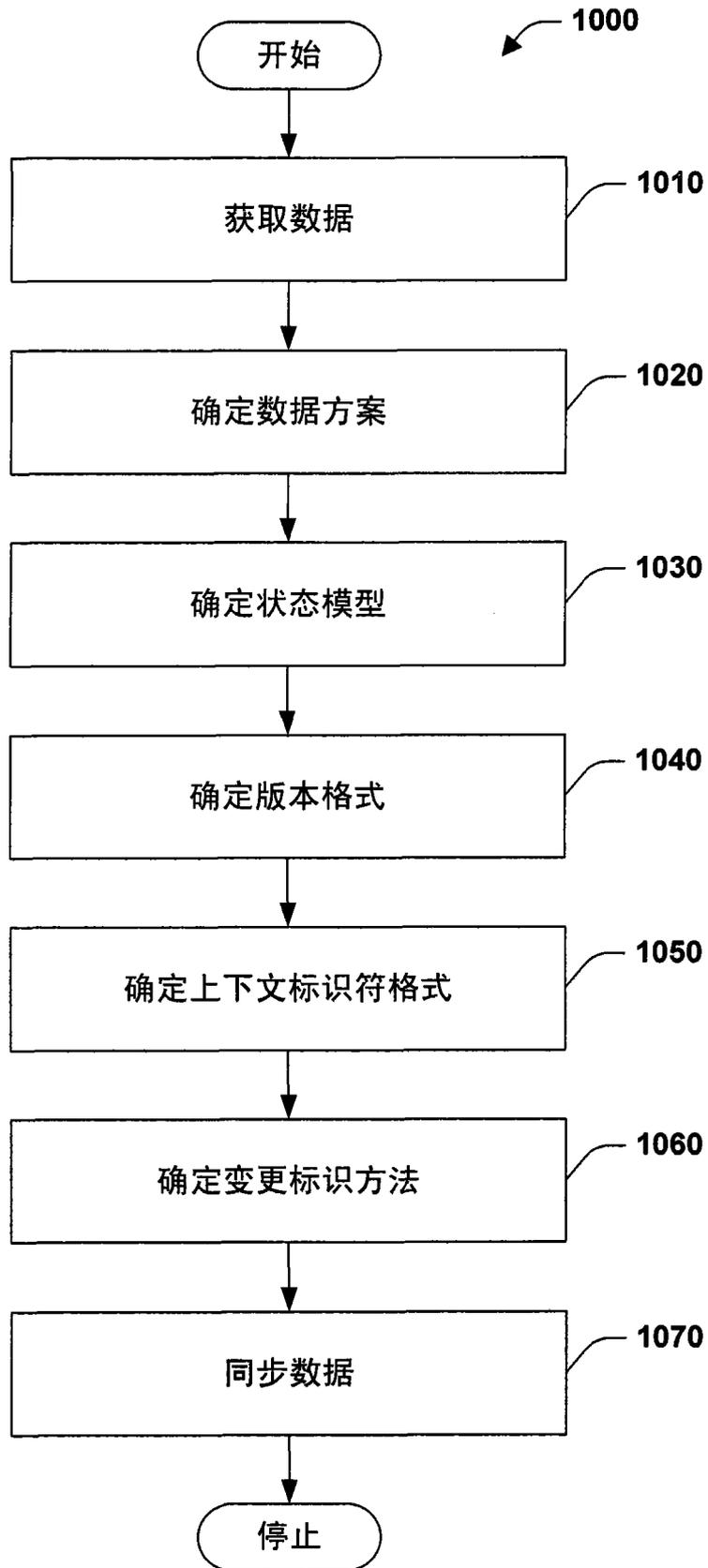


图 10

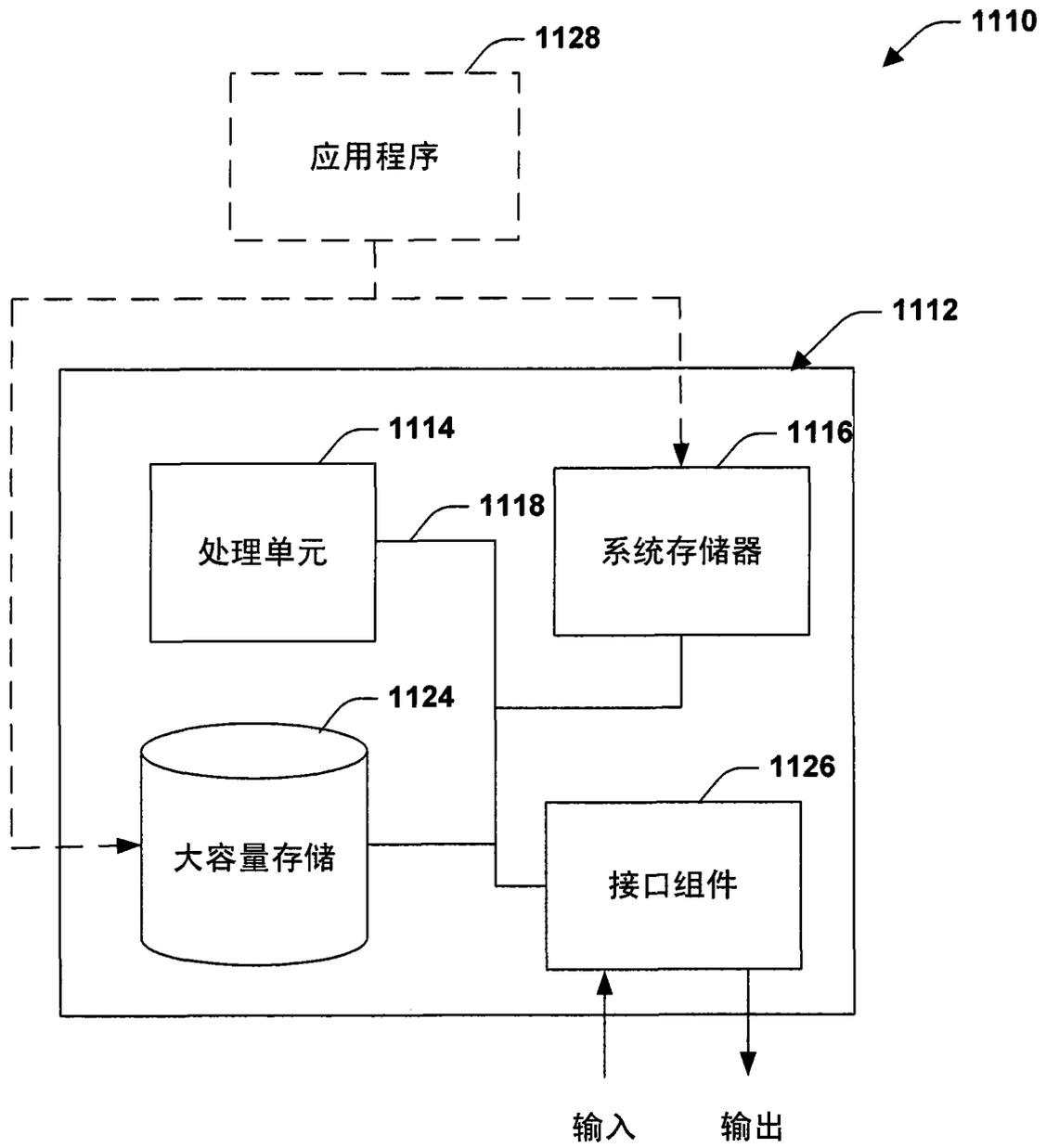


图 11

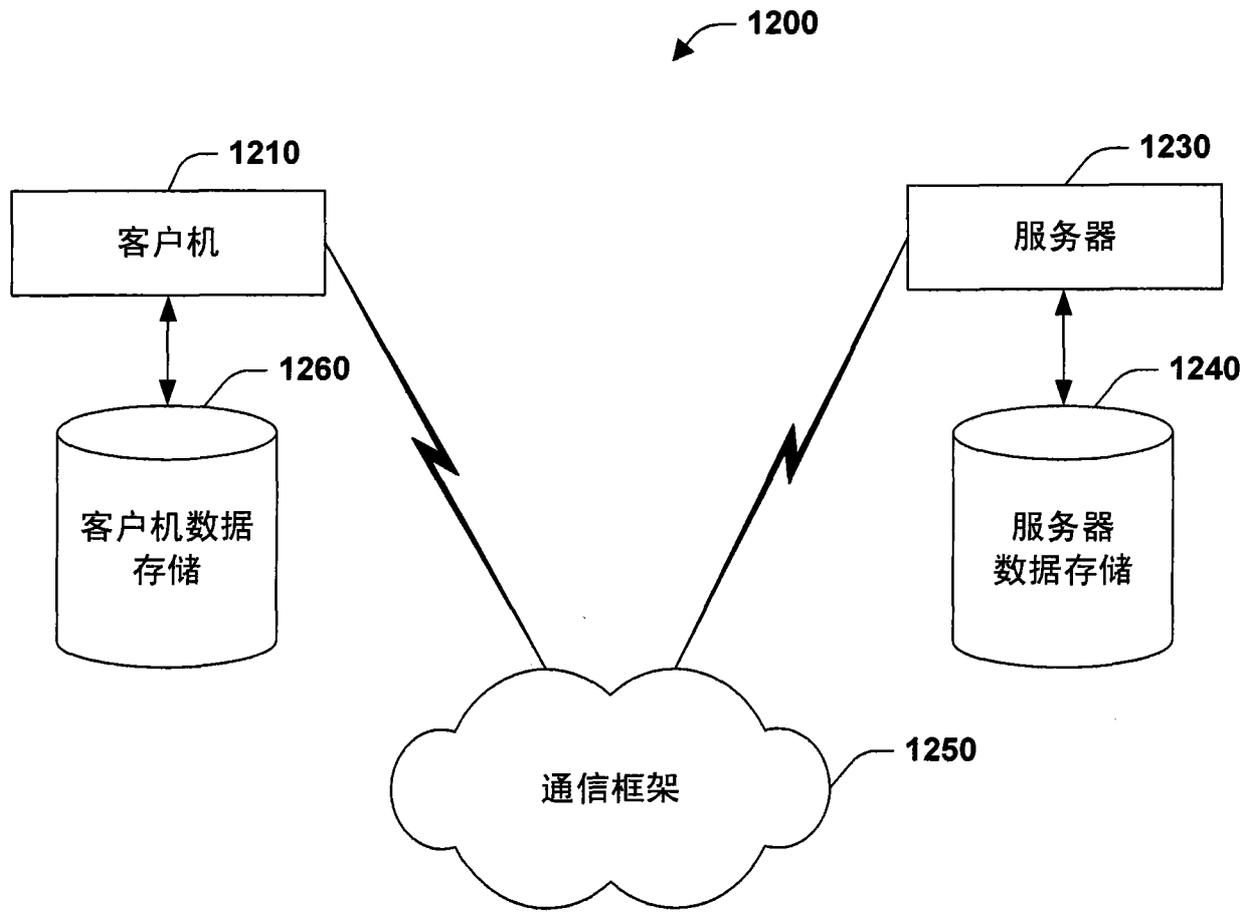


图 12