



[12] 发明专利申请公布说明书

[21] 申请号 200610153942.9

[43] 公开日 2007年4月25日

[11] 公开号 CN 1952942A

[22] 申请日 2006.9.12

[21] 申请号 200610153942.9

[30] 优先权

[32] 2005.10.21 [33] US [31] 11/255,506

[71] 申请人 国际商业机器公司

地址 美国纽约

[72] 发明人 查德·S·米特切尔

大卫·M·诺顿

厄尔·埃尔斯沃思

奥玛·B·瓦加斯

[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 陈 炜

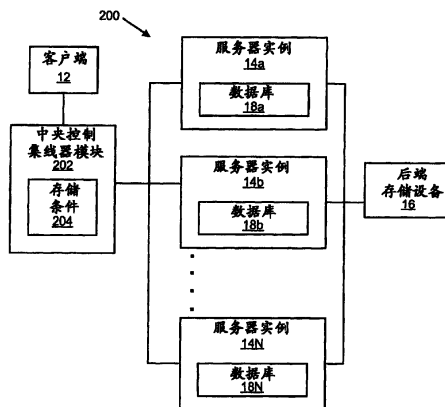
权利要求书4页 说明书17页 附图5页

[54] 发明名称

用于数据存储服务器的自动虚拟化的设备、系统和方法

[57] 摘要

本发明公开了一种用于自动地虚拟化数据存储服务器的设备、系统和方法。所述设备、系统和方法包括：中央控制集线器模块，被配置为服务于客户端和一个或多个存储服务器之间的存储服务器操作；检测模块，被配置为确定是否满足存储条件集合；以及创建模块，被配置为对满足存储条件集合作出响应，创建附加存储服务器实例。所述中央控制集线器模块还可以包括映象模块，被配置为维护客户端和服务器之间的存储映象并且向客户端提供存储映象。当创建新的服务器实例时，它可以包括新的维护数据库实例。



1. 一种用于自动地虚拟化数据存储服务器的设备，所述设备包括：

中央控制集线器模块，被配置为服务于客户端和一个或多个存储服务器之间的存储服务器操作；

检测模块，被配置为确定是否满足存储条件集合；以及

创建模块，被配置为对满足存储条件集合作出响应，创建附加存储服务器实例。

2. 如权利要求 1 所述的设备，其中，所述中央控制集线器模块还包括映象模块，被配置为维护客户端和一个或多个服务器之间的存储映象，并且向客户端提供所述存储映象。

3. 如权利要求 1 所述的设备，其中，所述中央控制集线器模块还包括响应模块，被配置为对从所述客户端接收的存储操作命令作出响应，所述响应模块包括：

翻译模块，被配置为翻译所述存储操作命令；以及

交互模块，被配置为响应于所述存储操作命令而与一个或多个服务器进行交互。

4. 如权利要求 1 所述的设备，其中，所述中央控制集线器模块还包括：

编译模块，被配置为编译来自一个或多个存储服务器的返回数据；以及

结果模块，被配置为基于存储操作命令来对所述返回数据进行操作，并且把结果数据返回给客户端。

5. 如权利要求 1 所述的设备，其中，所述中央控制集线器模块还被配置为使用标准存储操作命令来操作，以便使由中央控制集线器模块服务的存储服务器操作对于客户端和一个或多个服务器中的至少一个来说是透明的。

6. 如权利要求 1 所述的设备，其中，所述附加存储服务器实例

是从现有服务器克隆的，并且使用基本上相同的策略、用户和管理信息。

7. 如权利要求 1 所述的设备，其中，所述附加存储服务器实例包括与存储服务器相关联的新的存储维护数据库实例，并且所述存储条件集合包括超出存储维护数据库阈值。

8. 如权利要求 1 所述的设备，其中，所述创建模块还被配置为创建对应于附加存储服务器实例的附加存储代理实例。

9. 一种用于自动地虚拟化数据存储服务器的系统，所述系统包括：

一个或多个数据存储服务器；

与一个或多个服务器进行电子通信的客户端；

中央控制集线器模块，被配置为服务于客户端和一个或多个存储服务器之间的存储服务器操作；

检测模块，被配置为确定是否满足存储条件集合；以及

创建模块，被配置为对满足存储条件集合作出响应，创建附加存储服务器实例。

10. 如权利要求 9 所述的系统，其中，所述创建模块还被配置为创建对应于附加存储服务器实例的附加存储代理实例。

11. 如权利要求 9 所述的系统，其中，所述中央控制集线器模块还包括映象模块，被配置为维护客户端和一个或多个服务器之间的存储映象，并且向客户端提供所述存储映象。

12. 如权利要求 9 所述的系统，其中，所述中央控制集线器模块还包括响应模块，被配置为对从所述客户端接收的存储操作命令作出响应，所述响应模块包括：

翻译模块，被配置为翻译所述存储操作命令；

交互模块，被配置为响应于所述存储操作命令而与一个或多个服务器进行交互；

编译模块，被配置为编译来自一个或多个存储服务器的返回数据；以及

结果模块,被配置为基于存储操作命令来对所述返回数据进行操作,并且把结果提供给客户端。

13. 一种用于自动地虚拟化数据存储服务器的方法,所述方法包括:

服务于客户端和一个或多个存储服务器之间的存储服务器操作;
确定是否满足存储条件集合; 并且
对满足存储条件集合作出响应, 创建附加存储服务器实例。

14. 如权利要求 13 所述的方法, 其中还包括如下操作: 维护所述客户端和一个或多个服务器之间的存储映象, 并且向客户端提供所述存储映象。

15. 如权利要求 13 所述的方法, 其中还包括对从所述客户端接收的存储操作命令作出响应的操作, 所述操作包括:

翻译所述存储操作命令;
响应于所述存储操作命令而与一个或多个服务器进行交互;
编译来自一个或多个存储服务器的返回数据; 并且
基于所述存储操作命令来对所述返回数据进行操作, 并且向客户端提供结果数据。

16. 如权利要求 13 所述的方法, 其中, 存储服务器操作包括标准存储操作命令, 以便使所述存储服务器操作对于客户端和一个或多个服务器中的至少一个来说是透明的。

17. 如权利要求 13 所述的方法, 其中, 所述附加存储服务器实例是从现有服务器克隆的, 并且使用基本上相同的策略、用户和管理信息。

18. 如权利要求 13 所述的方法, 其中, 所述附加存储服务器实例包括与所述附加存储服务器实例相关联的新的存储维护数据库实例, 并且所述存储条件集合包括超出存储维护数据库的大小阈值。

19. 一种用于采用被配置为自动地虚拟化数据存储服务器的计算基础设施的方法, 所述方法包括如下步骤:

采用包括多个模块的软件, 所述模块被配置为:

服务于客户端和一个或多个存储服务器之间的存储服务器操作;

确定是否满足存储条件集合; 以及

对满足存储条件集合作出响应, 创建附加存储服务器实例;

以及

依照用户要求来配置所述存储条件集合。

20. 如权利要求 19 所述的方法, 还包括: 分析服务器存储系统, 所述服务器存储系统包括通过存储网络连接的一个或多个服务器和客户端; 确定所述系统是否可以得益于虚拟的数据存储服务器; 并且测试所采用的软件以确保正确的功能。

用于数据存储服务器的自动 虚拟化的设备、系统和方法

技术领域

本发明涉及数据存储系统，并且尤其涉及数据存储服务器的自动扩展和虚拟化。

背景技术

现代技术已经导致商业对电子数据存储系统的过度依赖。这些存储系统用来备份和存档必须被迅速地并且可靠地存储和检索的重要信息。通常，存储系统包括可由任意数目的客户端跨越公用网络访问的远程服务器或者服务器组。存储服务器可以利用诸如磁盘、光盘和磁带子系统之类的存储设备来为客户端存储大量的操作数据、备份数据和存档数据。管理软件通常用于配置一个维护数据库，所述维护数据库记录用于描述存储在后端存储设备上的文件的元数据。通常，这些维护数据库识别文件位置、文件属于哪个客户端、以及存储、创建或者修改文件的日期。这些维护数据库可以具有由正在使用的管理软件所定义的有限大小。通常，当现有维护数据库之一达到其容量时，必须创建新的维护数据库。这通常要求人工配置新的服务器并且重新配置每一个可能会导致高货币成本和产量损失的受影响的客户端。

在一个实施例中，使用由纽约的 Armonk 的 IBM 公司提供的 TSM (Tivoli Storage Manager) 软件来管理备份或者存档系统，其是通过维护存储在特定服务器上的文件的维护数据库来实现的。TSM 服务器维护数据库可以具有 530 千兆字节 (GB) 的上限大小，并且对于最佳性能和可维护性而言，其可以被限制为大约 100 吉字节。维护数据库的大小是通过确定将填充给定配置的文件数目并且将该值乘以每一条目的大小而计算的。假定不会要求存储池备份并且不聚集

文件，那么典型的维护数据库条目的大小大约是 600 字节，不过此数目可以因文件大小或者文件名而改变。因此，具有 12 兆兆字节 (TB) 的存储量并且平均文件大小为 10 千字节的服务器将会要求一个 840 千兆字节的维护数据库，其超出了 530 千兆字节的限制。

在另一个实施例中，把维护数据库大小表示为可利用的总存储空间的百分比。使用 TSM 软件，维护数据库的大小限制通常是总存储容量的百分之一到五。下面提供的表列出了对于各种存储容量而言的维护数据库的大小需求。

使用 600 字节/文件的维护数据库大小需求计算

平均 文件 大小	总的后端存储设备				
	3 TB	6 TB	12 TB	21 TB	48 TB
10 KB	210 GB*	420 GB*	840 GB**	1.47 TB**	3.36 TB**
100 KB	21 GB	42 GB	84 GB	147 GB*	336 GB*
250 KB	8.4 GB	16.8 GB	33.6 GB	58.8 GB	134.4 GB*
500 KB	4.2 GB	8.4 GB	16.8 GB	29.4 GB	67.2 GB
1 MB	2.1 GB	4.2 GB	8.4 GB	14.7 GB	33.6 GB
2 MB	1.05GB	2.1 GB	4.2 GB	7.35 GB	16.8 GB

使用存储容量的百分数的维护数据库大小需求计算

	总的后端存储设备				
	3 TB	6 TB	12 TB	21 TB	48 TB
1%	30 GB	60 GB	120 GB*	210 GB*	480 GB*
5%	150 GB	300 GB*	600 GB**	1.2 TB**	2.4 TB**

* 超出 TSM 的 100 GB 的实际维护数据库大小限制

** 超出 TSM 的 530 GB 的最大维护数据库大小限制

如上表中表明的那样，在很多情况下，单个维护数据库不足以管理较大存储区域。在典型的存档环境下，这些维护数据库可能达到其最大容量或者超出其理想的操作容量。特别是在数据保持环境下，数

据可以在较长时间段内被保留，由此不会释放维护数据库空间并且导致大小和效率问题。目前，对这些问题的补救措施是用新的维护数据库人工地配置新的服务器，然后使这些新的服务器在线，但就金钱和生产而言，这是非常昂贵的。

图 1 是举例说明常规的数据存储系统 10 的示意性框图。客户端 12 与服务器实例 (server instance) 14 连接。服务器实例 14 访问后端存储设备 16，并且包括维护数据库 18。在各种实施例中，客户端 12 可以包括个人计算机、工作站、膝上型电脑或者其它设备，正如本领域技术人员将意识到的那样。后端存储设备 16 可以包括磁盘存储设备、磁带存储设备、存储子系统或者其它存储设备或者存储设备组合，正如本领域技术人员将会意识到的那样。维护数据库 18 用来存储对应于存储在后端存储设备 16 中的数据的元数据。在一个实施例中，所述元数据可以包括诸如文件名、位置和大小、以及存储或者访问文件的日期之类的信息。

服务器实例 14 可以服务于多个客户端，并且可以服务于它们每一个的存储操作。客户端 12 可以经由局域网 (LAN) 连接至服务器实例 14，并且服务器实例 14 可以经由存储区域网络 (SAN) 连接至后端存储设备 16。如上所述，采用当前配置的一个问题是，维护数据库 18 可能会达到或者超出要求人工实现新的服务器和维护数据库的容量阈值。

根据先前的讨论，应该清楚的是，存在对于这样一种设备、系统和方法的需要，其中所述设备、系统和方法通过根据需要创建新的服务器实例和维护数据库来自动地虚拟化数据存储服务器。有益的是，这种设备、系统和方法将会增强系统的性能和可靠性，消除维护数据库的大小限制问题，并且明显减少扩展存储系统所需的人工配置步骤的数目。

发明内容

为适应所属技术领域的目前状态，特别是为适应所属领域中通过

目前可利用的数据存储方案仍未彻底解决的问题和需要，而开发了本发明。因此，开发了本发明，以便提供一种用于虚拟化数据存储服务器的设备、系统和方法，它们能够克服所属领域中的许多或全部上述缺陷。

用于自动虚拟化数据存储服务器的设备具有包含多个模块的逻辑部件，所述多个模块被配置为在功能上执行自动服务器创建的必要步骤。在所述实施例中的这些模块包括：中央控制集线器模块，被配置为服务于客户端和一个或多个存储服务器之间的存储服务器操作；检测模块，被配置为确定是否满足存储条件集合；以及创建模块，被配置为对满足存储条件集合作出响应，创建附加存储服务器实例。

在一个实施例中，所述设备被配置为维护客户端和一个或多个服务器之间的存储映象，并且向客户端提供所述存储映象。在另一个实施例中，所述设备被配置为对从客户端接收的存储操作命令作出响应，包括：翻译存储操作命令，响应于所述操作命令而与一个或多个服务器交互，编译来自一个或多个存储服务器的返回数据，基于存储操作命令对所述返回数据进行操作，并且把结果数据返回客户端。

在一个实施例中，所述设备被进一步配置为使用标准存储操作命令来进行操作，以便使由中央控制集线器模块服务的存储服务器操作对于客户端和一个或多个服务器中的至少一个来说是透明的。在另一个实施例中，附加存储服务器实例包括与所述存储服务器相关联的新的存储维护数据库实例，并且所述条件集合包括超出存储维护数据库阈值。

在一个进一步的实施例中，所述设备可以被配置为创建对应于附加存储服务器实例的附加存储代理实例。

还提供了本发明的用于自动虚拟化数据存储服务器的系统。特别的是，在一个实施例中，所述系统包括一个或多个数据存储服务器；与一个或多个服务器电子通信的客户端；中央控制集线器模块，被配置为服务于客户端和一个或多个服务器之间的存储服务器操作；检测模块，被配置为确定是否满足存储条件集合；以及创建模块，被配置

为对满足存储条件集合作出响应，创建附加存储服务器实例。

所述系统还可以包括对应于附加存储服务器实例的附加存储代理实例。所述系统还可以包括映象模块，被配置为维护客户端和一个或多个服务器之间的存储映象，并且向客户端提供所述存储映象；以及响应模块，被配置为对从客户端接收的存储操作命令作出响应。所述响应模块可以包括：翻译模块，被配置为翻译所述存储操作命令；交互模块，被配置为响应于所述存储操作命令来与一个或多个服务器交互；编译模块，被配置为编译来自一个或多个存储服务器的返回数据；以及结果模块，被配置为对所述返回数据进行操作并且向客户端提供结果。

还提供了本发明的、用于自动虚拟化数据存储服务器的方法。所公开的实施例中的方法基本上包括用于实现上文中就所述设备和系统的操作所给出的功能所需的步骤。

还提供了本发明的、用于采用被配置为自动虚拟化数据存储服务器的计算基础设施的方法。在所公开的实施例中，所述方法包括：使用包括多个模块的软件，所述模块被配置为服务于客户端和一个或多个存储服务器之间的存储服务器操作；确定是否满足存储条件集合；并且对满足存储条件集合作出响应，创建附加存储服务器实例。所述方法还可以包括依照用户要求来配置存储条件集合。

在一个进一步的实施例中，所述方法包括：分析服务器存储系统，所述服务器存储系统包括通过存储网络连接的一个或多个服务器和客户端；确定所述系统是否可以得益于虚拟数据存储服务器；并且测试所采用的软件以确保正确的功能。

在整个说明书中，对特征、优点的描述或类似的语言并不意味着可以采用本发明来实现的所有特征和优点应该是或者的确是在本发明的任何单个实施例中。相反，应该把描述了特征和优点的语言理解为这样的意思，即：结合实施例所描述的具体特征、优点或者特性被包括在本发明的至少一个实施例中。因此，在整个说明书中，对特征和优点的描述或类似的语言可以但未必涉及相同的实施例。

此外，在一个或多个实施例中，所描述的本发明的特征、优点和特性可以依照任何适当的方式进行组合。所属领域技术人员将会意识到的是，本发明可以在没有特定实施例的一个或多个具体特征或优点的情况下实施。在其它情况下，在某些实施例中可能会出现未必存在于本发明的所有实施例中的额外的特征和优点。

通过下列描述和所附权利要求书，本发明的这些特征和优点将变得更加明显，或者可以通过如此后所述那样实践本发明来获知。

附图说明

为了使本发明的优点易于理解，将参考在附图中所图示的具体实施例来呈现对上文简述的本发明的更加具体的描述。应理解的是，这些附图只描述了本发明的典型实施例，因此不应将其视为对本发明范围的限制，将通过使用附图更加具体和详细地描述和解释本发明，其中：

图 1 是举例说明常规的数据存储系统的示意性框图；

图 2 是举例说明依照本发明的数据存储系统的一个实施例的示意性框图；

图 3A 是举例说明依照本发明的中央控制集线器模块的一个实施例的示意性框图；

图 3B 是举例说明依照本发明的响应模块的一个实施例的示意性框图；

图 4 是举例说明依照本发明的映象模块的一个实施例的示意性框图；

图 5 是举例说明依照本发明在 LANfree 环境下的数据存储系统的一个实施例的示意性框图；并且

图 6 是举例说明依照本发明的数据存储服务器虚拟化方法的一个实施例的示意性流程图。

具体实施方式

此说明书中所描述的许多功能部件都被称为模块,以便更加特别地强调其实现方式的独立性。例如,模块可以作为硬件电路来实现,所述硬件电路包括常规的 VLSI 电路或者门阵列,诸如逻辑芯片、晶体管之类的现有半导体或者是其它分立的元件。模块还可以用可编程硬件设备、诸如现场可编程门阵列、可编程阵列逻辑、可编程逻辑设备等实现。

模块还可以用软件实现,以便由各种类型的处理器执行。举例来说,一个标识的可执行代码模块可以包括计算机指令的一个或多个物理或者逻辑块,举例来说,其可以被构建为对象、程序或函数。尽管如此,所标识模块的可执行代码无需物理地位于一起,而是可以包括存储在不同位置上的不同的指令,当这些指令逻辑上结合在一起时,其构成模块并且实现该模块的规定目的。

实际上,可执行代码模块可以是单条指令或者是许多指令,并且甚至可以分布在多个不同代码段上,分布在不同程序当中,以及跨越多个存储器设备分布。同样地,操作数据可以在模块内被识别并且图示,并且可以依照任何适当的形式实现并且被组织在任何适当类型的数据结构内。所述操作数据可以作为单个数据集被收集,或者可以分布在不同位置上(包括在不同存储设备上),并且至少部分地可以仅作为电子信号存在于系统或网络上。

在整个说明书中所提及的“一个实施例”、“实施例”或者类似语言是指把结合实施例描述的特定特征、结构或者特征包含在本发明的至少一个实施例中。因此,在整个说明书中,短语“在一个实施例中”、“在一实施例中”以及类似语言的出现可以但也未必全部涉及相同的实施例。正如此处使用的那样,术语“自动”指的是在几乎没有用户干预的情况下可能响应于激励而发生。术语“存储服务器”指的是构成服务器的软件和硬件,其中所述服务器通常用于存储并且维护电子数据。术语“存储服务器实例”指的是存储服务器的具体出现事件,并且可以是另一个存储服务器的克隆。

所提及的信号承载介质可以采用任何形式,只要能够生成信号、

导致信号生成、或者使机器可读指令程序在数字处理设备上执行即可。信号承载介质可以由传输线、光盘、数字视频盘、磁带、Bernoulli驱动器、磁盘、穿孔卡、闪存、集成电路或者其它数字处理设备的存储器设备来实现。

此外，在一个或多个实施例中，所描述的本发明的特征、结构和特性可以依照任何适当的方式进行组合。在随后的描述中，提供了很多具体细节，诸如编程示例、软件模块、用户选择、网络事务、维护数据库查询、维护数据库结构、硬件模块、硬件电路、硬件芯片等，以便提供对本发明实施例的彻底的了解。然而，所属领域技术人员将会意识到的是，本发明可以在没有一个是或多个该具体细节的情况下实施，或者利用其它的方法、组件、材料等来实施。在其它情况下，没有示出或详细说明众所周知的结构、材料或操作以免模糊本发明的各个方面。

图2举例说明了用于自动虚拟化数据存储服务器的系统200的一个实施例。在一个实施例中，客户端12与中央控制集线器模块202相连，所述中央控制集线器模块202与一个或多个服务器实例14a-N相连，每一服务器实例14a-N包括维护数据库18a-N。一个或多个服务器实例14a-N与后端存储设备16相连。

所述系统200包括中央控制集线器模块202。所述中央控制集线器模块202可以包含多个逻辑模块，用于服务于一个或多个客户端12和一个或多个服务器实例14a-N之间的存储服务器操作。在一个实施例中，所述中央控制集线器模块202允许客户端12与多个服务器实例14a-N通信，就好像客户端12正与单个服务器通信一样，以便使通信对于客户端12或者服务器实例14a-N中的至少一个而言是透明的。透明是指例如客户端12不知道客户端12正在与一个以上的服务器实例14a-N通信。在替代性的实施例中，中央控制集线器模块202只是分别地便于客户端12和服务器实例14a-N之间的通信。在此实施例中，中央控制集线器202将通信从客户端12引导到适当的服务器实例14a。另外，在一个实施例中，所述中央控制集线器模块202

可以在独立的一块硬件上实现，或者在另一个实施例中，可以与服务器实例 14a-N 或者其它软件模块共享硬件。

在一个实施例中，所述中央控制集线器模块 202 包含用于自动创建附加服务器实例 14a 所必需的逻辑。例如，所述中央控制集线器模块 202 可以依照存储条件集合 204 来确定是否需要附加服务器实例 14b。倘若如此，则中央控制集线器模块 202 创建第二服务器实例 14b。在一个实施例中，中央控制集线器 202 对服务器实例 14a 进行克隆，以便创建服务器实例 14b。优选的是，所克隆的服务器实例 14b 使用与原始服务器实例 14a 基本上相同的策略、用户和管理信息。在另一个实施例中，中央控制集线器模块 202 可以生成基本上不同于服务器实例 14a 的服务器实例 14b。所述处理可以重复无限的次数，正如通过创建服务器实例 14a-N 及其相关联的维护数据库 18a-N 而描述的那样。

存储条件集合 204 表明对附加服务器实例 14b 的需要。在各种实施例中，所述存储条件集合 204 可以包括但不限于：包括维护数据库 18a 中所使用的存储空间量、维护数据库 18a 中剩余的存储空间量、系统中检测出的错误，存储处理负载或者由用户所定义的其他条件在内的因素。

在一个实施例中，附加服务器实例 14b 的创建是由诸如维护数据库 18a 达到或者超出大小阈值之类的存储条件来触发的。例如，维护数据库 18a 可能正接近其最大容量，或者可能已经超过了其最佳操作大小。新创建的附加服务器实例 14b 包括附加维护数据库 18b，所述中央控制集线器模块 202 可以使用该附加维护数据库 18b 来减少数据库 18a 上的负载。例如，在一个实施例中，所述中央控制集线器模块 202 把存储在服务器实例 14a 上的数据的一部分移至新的服务器实例 14b，或者在替代性的实施例中，把所有将来的数据全部引导至新的服务器实例 14b，以允许客户端 12 拥有存储在两个服务器实例 14a-b 上的数据。在另一个实施例中，来自所有新客户端 12 的通信被引导至新的服务器实例 14b 而不是服务器实例 14a。

在一个实施例中,所述中央控制集线器模块 202 在局域网(LAN) 环境内操作,其中客户端 12 跨越 LAN 与服务器实例 14a-N 进行通信。客户端 12 与中央控制集线器模块 202 建立通信,以便在开始和验证期间可以并入握手以便获取涉及客户端 12 的主要服务器 14a 以及任何辅助服务器 14b-N 的信息,其中所述任何辅助服务器 14b-N 包含客户端的元数据。这样做允许客户端 12 经由中央通信集线器 202 依照透明方式与一个或多个服务器 14a-N 进行通信,或者作为选择,允许客户端 12 直接与服务器实例 14a-N 进行通信。

图 3A 举例说明了中央控制集线器模块 202 的一个实施例,所述中央控制集线器模块 202 包括检测模块 302、创建模块 304、响应模块 305、映象模块 306 和存储条件集合 204。优选的是,所述中央控制集线器模块 202 服务于一个或多个客户端 12 和一个或多个服务器实例 14a-N 之间的存储服务器操作。另外,所述中央控制集线器模块 202 被配置为产生附加服务器实例 14b-N,并且便于一个或多个客户端 12 和新的服务器实例 14b-N 之间的通信。

在一个实施例中,检测模块 302 确定存储条件集合 204 内的存储条件是否得以满足。满足存储条件集合 204 中的条件之一,通常表明需要附加服务器实例 14b。例如,一个存储条件 204 可以表明维护数据库 14a 的大小阈值。另一存储条件 204 可以定义服务器的负载阈值。作为选择,可以要求相关存储条件 204 的组合来触发附加服务器实例 14b 的创建。在某些实施例中,存储条件集合 204 可以由用户或者通过计算机输入来改变或者重新配置。

为了检测存储条件 204 何时被满足,检测模块 302 监控所述服务器实例 14a-N。优选的是,所述映象模块 306 在中央控制集线器 202 上定期地存储已更新的存储条件信息,诸如维护数据库大小。检测模块 302 访问所述存储条件信息,以便确定存储条件 204 之一是否已经被满足。在另一个实施例中,检测模块 302 定期地测试存储条件 204 的出现。例如,中央控制集线器模块 202 可以定期地查询服务器实例 14a-N,以便提示来自每一服务器实例 14a-N 的、表明存储条件 204

是否已被满足的响应。在又一个实施例中，每一服务器实例 14a-N 均被配置为当存储条件 204 之一得以满足时、通知检测模块 302。依照此方式，对存储条件 204 进行自动检查，以便确定何时要求新的服务器实例 14a-N。

创建模块 304 响应于来自检测模块 302 的、表明一个或多个存储条件 204 已经被满足的信号，创建附加服务器实例 14b-N。在一个实施例中，所述创建模块 304 响应于来自管理控制台（未示出）的用户输入，创建附加服务器实例 14b-N。在另一个实施例中，利用基本上相同的策略、用户以及管理信息，从另一现有的服务器实例 14a 中克隆新的服务器实例 14b。优选的是，使用中央控制集线器 202 可以利用的软件或者存储在其内的软件来生成新的服务器实例 14b。一旦新的服务器实例 14b 被生成，创建模块 304 就通过复制来自另一服务器实例 14a 的设定，或者在另一个实施例中通过从文件或者脚本中读取配置设定，来对它进行配置。另外，所述创建模块 304 启动新的服务器实例 14b、客户端 12 和中央控制集线器 202 之间的通信。在一个实施例中，新的服务器实例 14b 用由用户所定义的其自身策略、用户和管理信息来创建。

下面详细说明响应模块 305 和映象模块 306(包括存储映象 310)。

图 3B 是举例说明依照本发明的响应模块 305 的一个实施例的示意性框图。所述响应模块 305 包括翻译模块 314、交互模块 316、编译模块 318 和结果模块 320。优选的是，所述响应模块 305 被配置为通过对从客户端 12 接收的存储操作命令 308 作出响应来便于客户端 12 和服务器实例 14a-N 之间的通信。存储操作命令 308 包括客户端 12 和存储服务器实例 14a 之间的典型通信，诸如查询、恢复或者其它存储操作命令，正如本领域技术人员将会意识到的那样。

翻译模块 314 被配置为翻译存储操作命令 308。例如，客户端 12 可以向中央控制集线器模块 202 发送查询请求或者恢复命令。翻译模块 314 接收所述命令，并且把对应的信号发送到交互模块 316。所述交互模块 316 响应于存储操作命令 308 而与一个或多个服务器实例

14a-N 进行交互。例如，如果中央控制集线器模块 202 接收到查询请求或者其它标准存储操作命令 308，那么交互模块 316 向每个适当的服务器实例 14a-N 发送对应的命令，以允许客户端 12 依照透明方式与多个服务实例 14a-N 进行通信。

所述编译模块 318 编译来自服务器实例 14a-N 的返回数据，并且在 一个实施例中，它可以向客户端 12 提供所述数据，就好像客户端 12 正直接与单个服务器实例 14a 进行通信一样。所述结果模块 320 基于存储操作命令 308 对返回数据进行操作，并且把经操作后的结果数据返回到客户端 12。例如，结果模块 312 可以对结果数据进行排序，除去重复的有效对象，或者只向客户端 12 返回最近的结果数据。所述结果数据还可以包括元数据，诸如哪个服务器实例 14a-N 维护所请求的数据。所述元数据可以允许客户端 12 发布其它操作命令，诸如直接对服务器实例 14a 进行恢复或者检索。

在另一个实施例中，所述中央控制集线器模块 202 接收来自客户端 12 的查询命令，但并不编译并且也不把结果数据返回到客户端 12，相反，中央控制集线器 202 可以通过把查询和客户端信息引导至每一服务器实例 14a-N 来充当代理。因此，所涉及的各个服务器实例 14a-N 可以直接向客户端 12 发送数据结果。然后，客户端 12 可以编译并处理所述返回数据。

图 4 是举例说明依照本发明的映象模块 306 的一个实施例的示意性框图。所述映象模块 306 被配置为维护客户端 12 和一个或多个服务器实例 14a-N 之间的存储映象 310，并且向客户端 12 提供存储映象。所述映象模块 306 可以包括客户端至服务器的映象 402，其包含诸如哪些客户端 404 拥有哪些服务器实例 406 上的数据、哪些文件空间 408 被存储在哪些服务器实例 406 上之类的信息，并且包含用于标识映象数据何时被存储以及映象数据上一次何时被修改的时间点信息 410。

存储映象 310 中还包括表明硬件可用性 412 和存储空间可用性 414 的信息，所述信息可以用于创建附加服务器实例 14a-N。然而，存储映象 310 不局限于这些类别，而是可以包括更多或更少的与客户

端和服务器的关系有关的信息。如图所示，存储映象 310 示出了两个客户端和两个服务器的信息。客户端 A 在服务器 1 上具有两个文件空间并且在服务器 2 上具有一个文件空间。客户端 B 具有两个位于服务器 2 上的文件空间，其中一个用于长期存储，而另一个用于由文件空间名称表明的系统状态备份。此信息由中央控制集线器 202 使用，以服务于客户端 12 和服务器实例 14a-N 之间的标准存储操作命令 308。

在一个实施例中，客户端 12 可以在中央控制集线器模块 202 中查询客户端 12 希望从中进行查询的文件空间 408（包括时间点信息 410）。在一个实施例中，映象模块 306 搜索客户端至服务器的映象 402，并且向客户端 12 返回文件空间至服务器的响应的列表。然后，客户端 12 能够向文件空间至服务器的响应列表中所包括的每个服务器实例 14a-N 分别发布各自的查询，并且对结果进行编译以便向用户提供透明视图。

在一个实施例中，在启动时客户端 12 可以查询映象模块 306，以便识别客户端 12 所拥有的数据被存储在哪里。由映象模块 306 返回的元数据可以表明多个服务器的服务器信息，诸如服务器名称、通信方法、通信信息等。客户端 12 可以在存储器中存储这种元数据。当需要时，客户端 12 可以参考此元数据，以便直接与服务器实例 14a 进行通信。当由创建模块 304 创建了新的服务器实例 14b-N 时，可以把新的信息添加到存储器中。

对于各种实施例中的服务器验证来说，可以把密码存储在中央控制集线器模块 202 中，使之跨越服务器实例 14a-N 同步，或者将其移至在服务器实例 14a-N 外部的某些实体。如果把用于访问多个服务器实例 14a-N 的密码存储在中央控制集线器模块 202 上，那么客户端 12 可以向中央控制集线器 202 进行验证，就好像验证是使用常规的服务器管理软件、诸如 Tivoli Storage Manager 进行处理的一样。在一个实施例中，客户端 12 向中央控制集线器 202 进行验证，就好像中央控制集线器 202 是服务器实例 14a 一样。然后，中央控制集线器 202 向所有适当的服务器实例 14a-N 发送会话密钥，并且把验证成功/失败

返回到客户端 12。然后，将会话密钥用于加密在客户端 12 和服务器实例 14a-N 之间直接传递的数据。

图 5 是举例说明在 LANfree 环境下的数据存储系统 500 的一个实施例的示意性框图。存储代理 504a-b 便于在 LANfree 路径上进行通信，正如本领域技术人员将会意识到的那样。LANfree 路径 508 是这样一种通信路径，其绕过 LAN 506，并且允许数据在客户端 12 和服务器实例 14a 之间直接传送。有益的是，LANfree 路径减少了 LAN 506 上的负载。LAN 路径 510 是包括 LAN 506 的组件和成员的通信路径。在一个实施例中，中央控制集线器模块 202 被类似地包括在存储代理中，但服务于客户端 12 与服务器实例 14a 和 14b 之间的 LAN 路径通信。

通常，在 LANfree 环境下，客户端 12 通过执行握手协议与存储代理 504a 连接，以便获得适当的服务器实例信息。一旦获得服务器实例信息，客户端 12 就与服务器实例 14a 连接，而无需来自存储代理 504a 的帮助。在本发明的一个实施例中，客户端的数据可以被存储在多个服务器实例 14a-N 上，其中每个服务器实例均具有相应的存储代理实例 504a-N。映象模块 306（参见图 3）可以包括客户端至存储代理的映象信息，在一个实施例中，该信息允许客户端 12 启动与中央控制集线器模块 202 的通信并且查询适当的存储代理信息。例如，所述存储映象 310 可以包括对应于每一服务器实例 406 的存储代理实例 504a，以便把每一客户端 402 映射至一个或多个存储代理 504a-b 和一个或多个服务器实例 406。一旦把查询结果返回客户端 12 后，客户端 12 就可以启动与适当的存储代理实例 504a 的通信，并且存储操作命令 308 可以照常继续。在 LANfree 环境下，当创建附加服务器实例 14b-N 时，中央控制集线器 202 还可以创建对应于附加服务器实例 14b-N 的相关存储代理 504b。

随后的示意性流程图总体上是作为逻辑流程图提出的。因而，所描述的顺序和所标记的步骤表示所给出的方法的一个实施例。可以设想出在功能、逻辑或效果上等效于所示方法的一个或多个步骤或者其

一部分的其它步骤和方法。另外，所采用的格式和符号是为解释所述方法的逻辑步骤而提供的，并且应当理解为不对所述方法的范围作出限制。虽然在流程图中采用了各种箭头类型和线类型，但是应理解这些不是对相应方法的范围的限制。实际上，某些箭头或者其它连接符仅用来表明方法的逻辑流程。举例来说，箭头可以表明所述方法的所列举步骤之间的未规定持续时间的等待或者监控时段。另外，执行特定方法的顺序可以或者未必严格地遵循所示出的相应步骤的顺序。

图 6 是举例说明依照本发明的用于数据存储服务器 14 的虚拟化方法 600 的一个实施例的示意性流程图。所述方法 600 开始（步骤 602），并且对中央控制集线器模块 202 进行初始化（步骤 604）。在各种实施例中，初始化所述中央控制集线器模块 202 的处理可以包括但不局限于：向中央控制集线器 202 供电，执行与中央控制集线器 202 的握手协议，或者在客户端 12 和中央控制集线器模块 202 之间建立通信。在一个实施例中，所述中央控制集线器模块 202 服务于客户端 12 和一个或多个服务器实例 14a-N 之间的存储服务器操作，以便使所述操作对于客户端 12 和一个或多个服务器 14a-N 中的至少一个而言是透明的。有益的是，这种透明性动态地提供了多个附加存储服务器 14a-N，并且无需对客户端 12 或者服务器实例 14a-N 进行软件改变。因而，受维护数据库 18a-N 的大小限制而强加的限制得以克服。

在一个实施例中，所述中央控制集线器模块 202 通过登记与连接到中央控制集线器模块 202 的每一客户端 12 和每一服务器实例 14a-N 之间的关系有关的信息，来维护客户端至服务器的映象 402（步骤 606）。所述客户端至服务器的映象 402 可以包括诸如哪些客户端 404 具有存储在哪些服务器 406 上的文件、客户端的文件空间 408 位于哪里、以及客户端的文件何时被存储（410）或上一次何时被访问（410）之类的信息。映象 402 还可以包括与存储代理 504a、可用硬件 412 和可用存储空间 414 有关的信息。

接下来，中央控制集线器模块 202 确定是否已经接收了存储操作命令 308（步骤 608）。倘若如此，则中央控制集线器模块 202 服务

于存储操作命令 308 而不是服务于服务器实例 14a (步骤 610), 或者把存储操作命令 308 引导至适当的服务器实例 14a-N 以便进行处理。服务于存储操作命令 308 的处理可以包括对查询、恢复/检索命令、备份/存档命令、服务器验证请求或者其它存储操作命令 308 作出响应, 正如本领域技术人员将会意识到的那样。因此, 所述中央控制集线器模块 202 可以翻译存储操作命令, 与一个或多个服务器实例 14a-N 进行交互, 编译来自一个或多个服务器实例 14a-N 的返回数据, 并且对所述返回数据进行操作, 以便向客户端 12 提供适当的结果。在另一个实施例中, 所述中央控制集线器模块 202 可以直接向客户端 12 提供映象信息, 以便允许客户端 12 直接与适当的服务器实例 14a-N 进行通信。

接下来, 所述中央控制集线器模块 202 确定是否需要新的服务器实例 14b-N (步骤 612)。中央控制集线器模块 202 评估存储条件集合 204。如果存储条件 204 得以满足, 那么中央控制集线器模块 202 创建新的服务器实例 14b (步骤 614)。在一个实施例中, 可以通过执行定义了可由中央控制集线器模块 202 利用的新服务器实例 14b 的软件, 来创建新的服务器实例 14b。在另一个实施例中, 新的服务器实例 14b 可以从现有服务器 14a 来克隆, 并且可以利用基本上相同的策略、用户和管理信息。新的服务器实例 14b 还可以包括与新的服务器实例 14b 相关联的新的存储维护数据库实例 18b。接下来, 方法 600 返回到维护客户端至服务器的映象 402 (步骤 606)。如果不需要新的服务器实例 (步骤 612), 那么中央控制集线器模块 202 继续维护客户端至服务器的映象 (步骤 606)。

本发明还包括用于采用被配置为自动虚拟化数据存储服务器的计算基础设施的方法。在一个实施例中, 采用包括多个模块的软件, 以便服务于客户端 12 和一个或多个存储服务器 14a-N 之间的存储服务操作 308, 确定存储条件集合 204 是否得以满足, 并且对满足存储条件集合 204 作出响应而创建附加存储服务器实例 14b-N。所述方法还可以包括依照用户请求来配置存储条件集合 204。在各种实施例

中，所述存储条件集合 204 可以包括当前存储维护数据库 18a 超出大小阈值或者服务器工作负荷超出服务器负载阈值。

在一个实施例中，所述方法还包括：分析服务器存储系统 200，其中所述服务器存储系统 200 包括通过存储网络连接的一个或多个服务器 14a-N 和客户端 12；确定所述系统 200 是否可以得益于虚拟的数据存储服务器；并且测试所采用的软件以确保正确的功能。

在不脱离本发明的实质或基本特性的情况下，可以依照其它具体形式来实现本发明。所描述的实施例在各方面都应仅仅被视为是例证性的而非限制性的。因此，本发明的范围由所附权利要求书表明而不是由以上的描述来表明。落入权利要求的含义和等效范围内的所有变化都被包含在本发明的范围内。

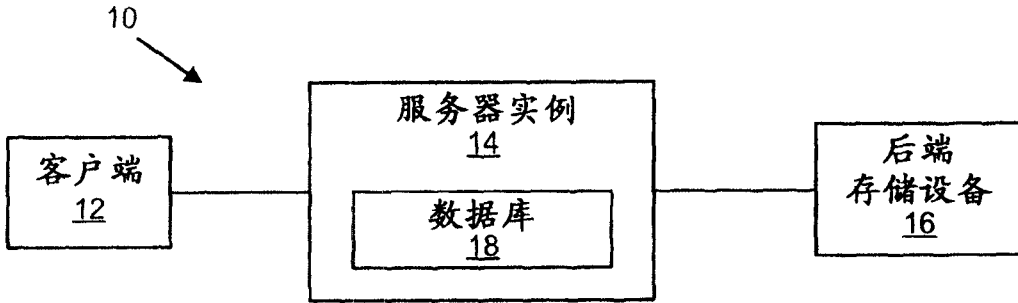


图1(现有技术)

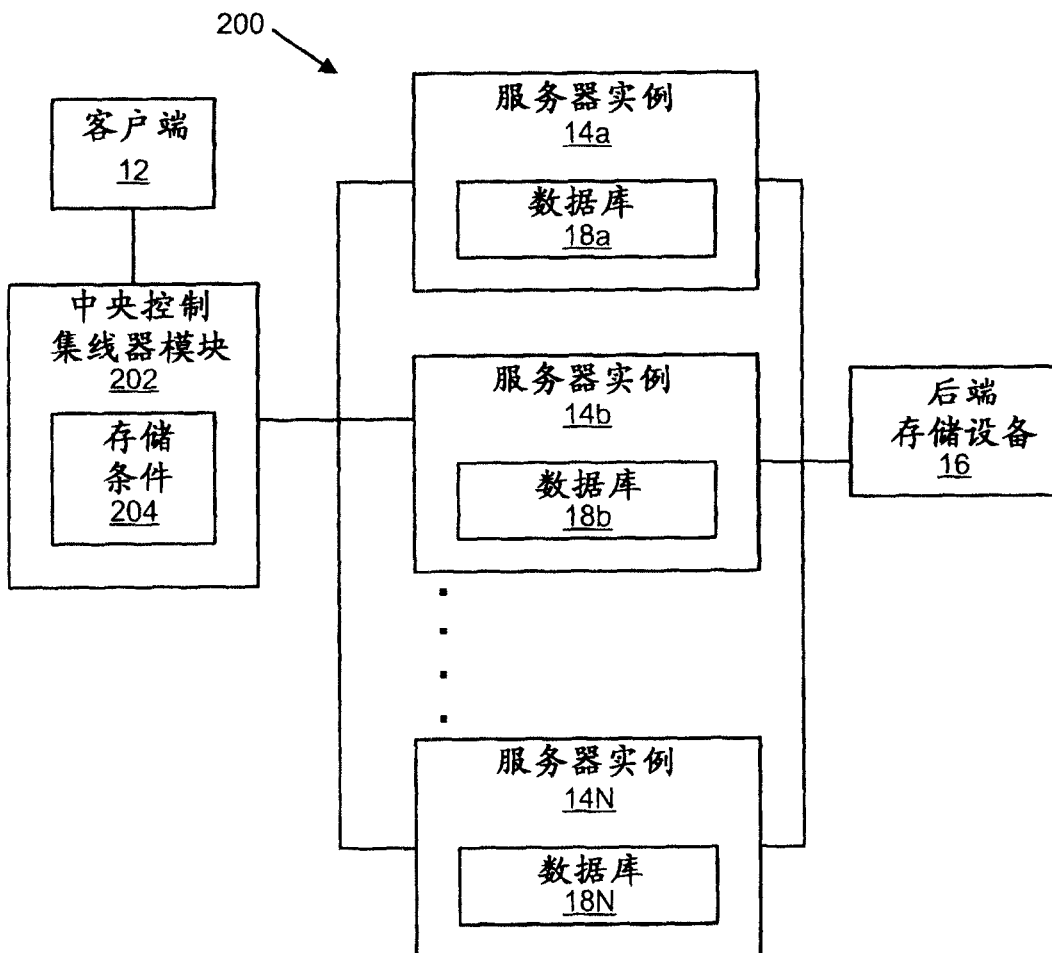


图2

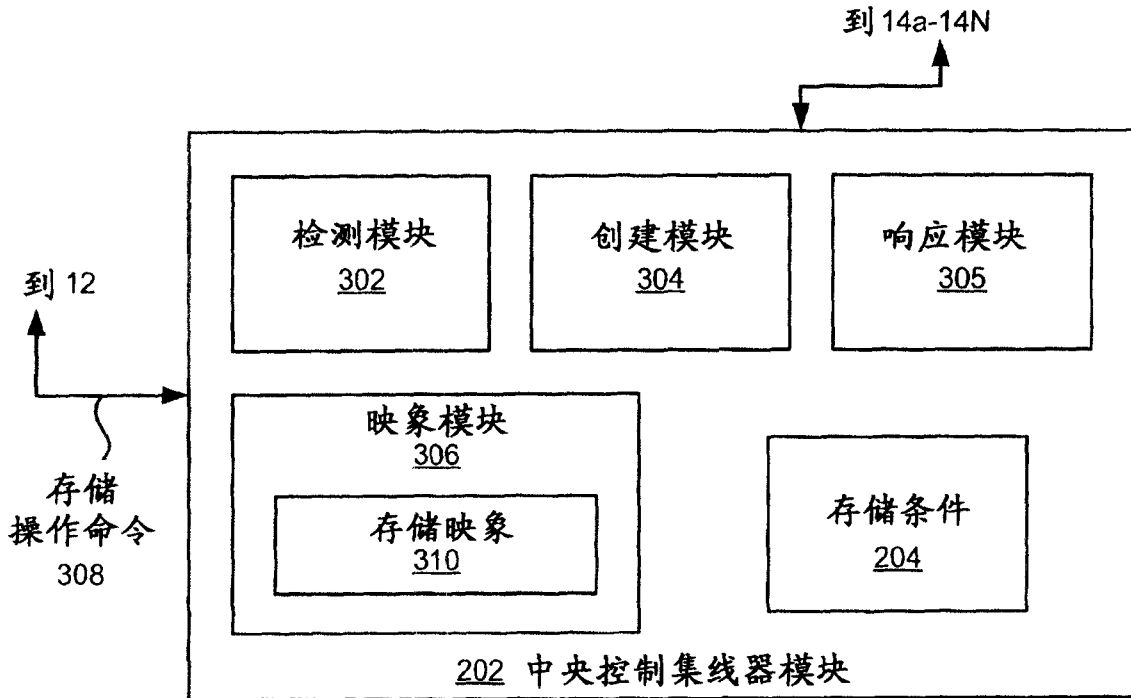


图 3A

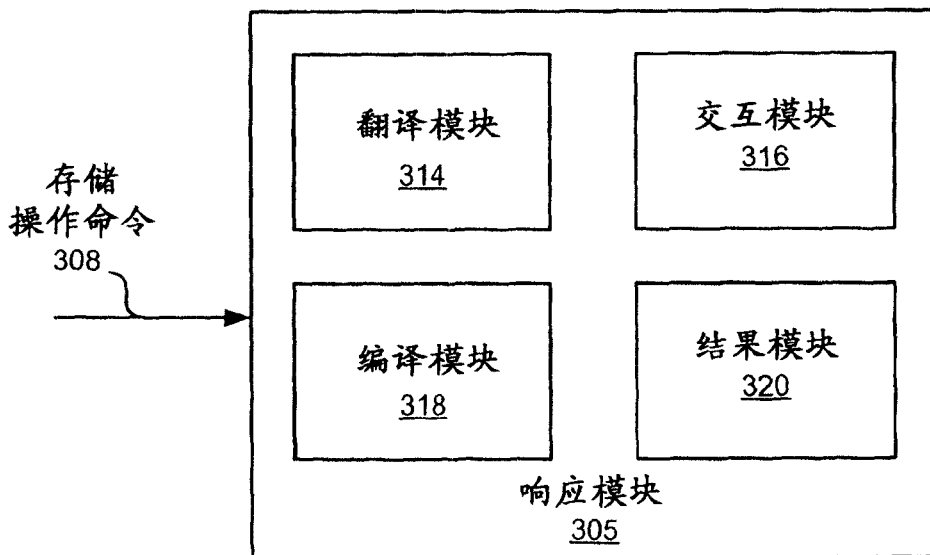


图 3B

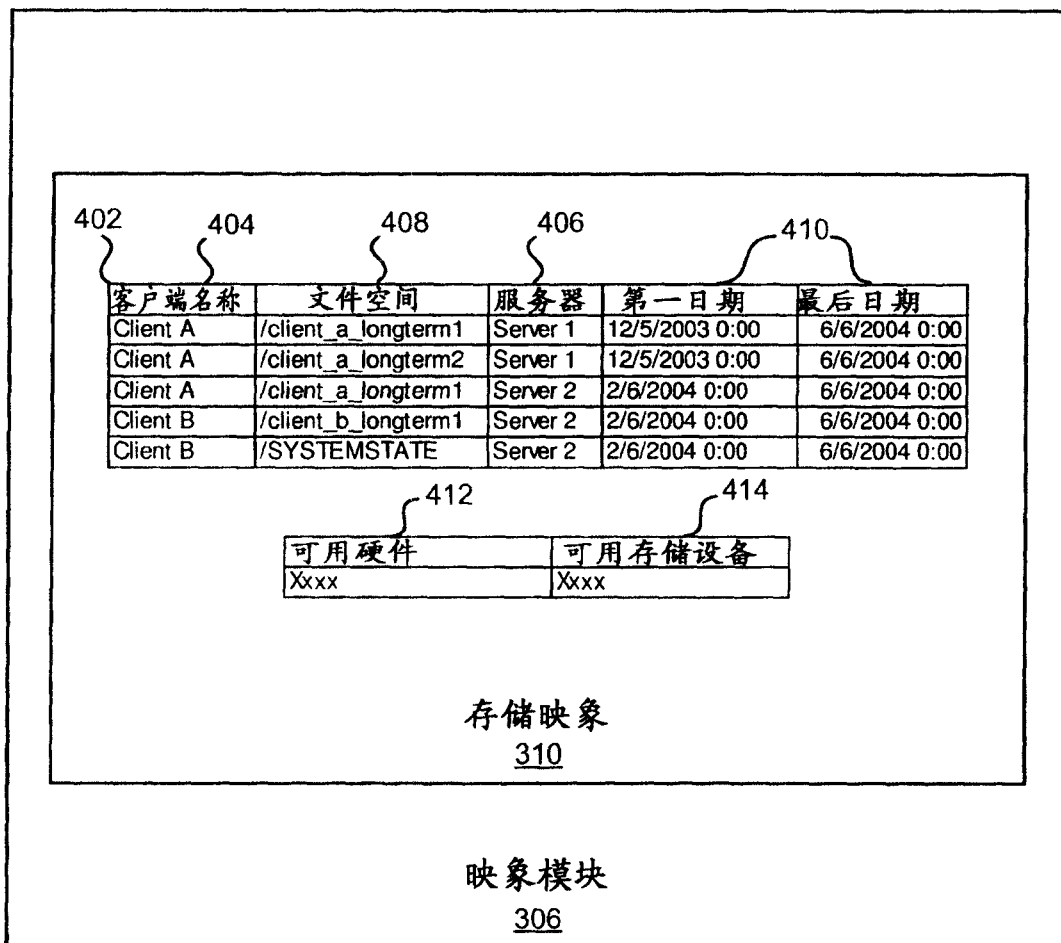


图 4

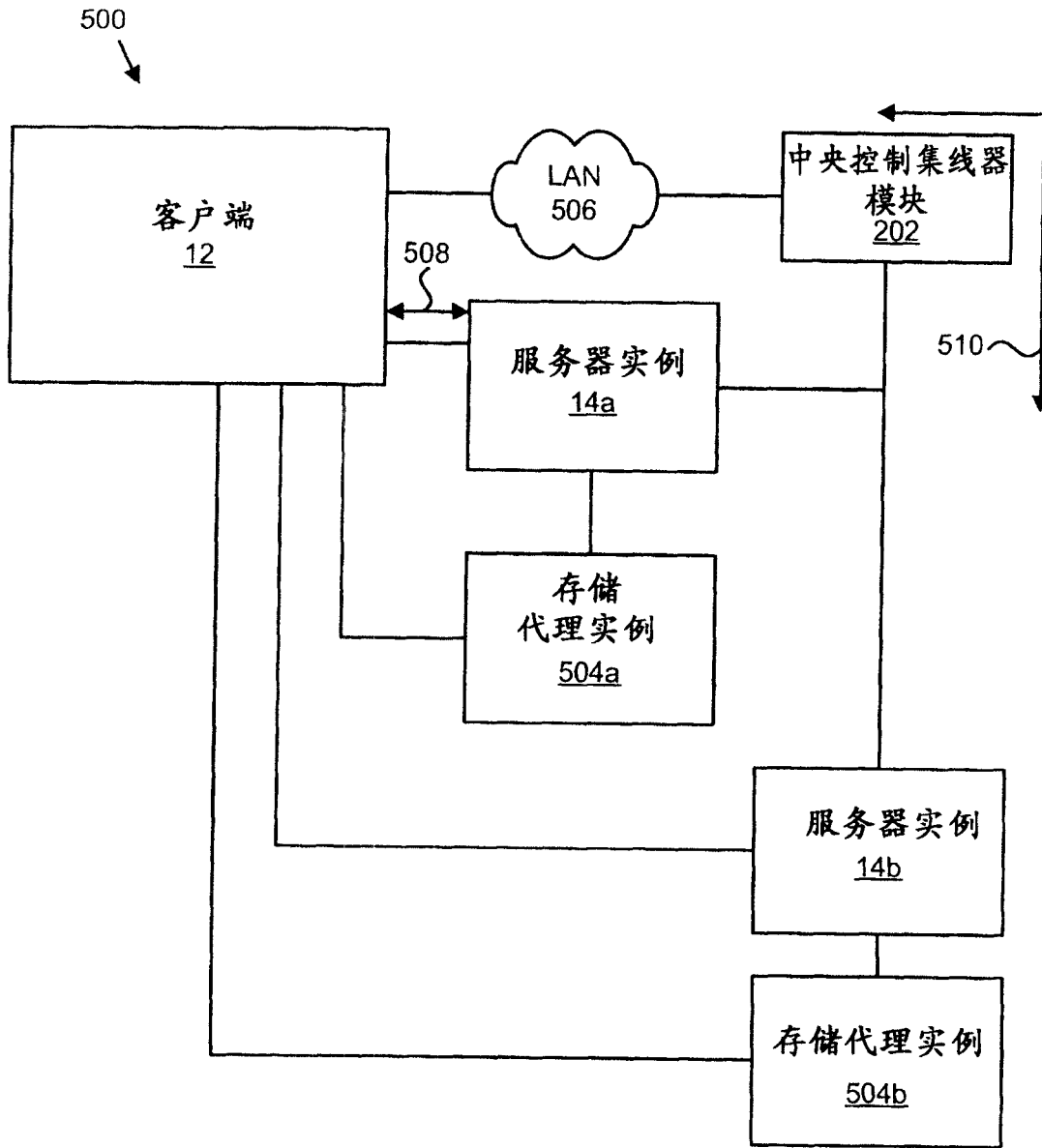


图 5

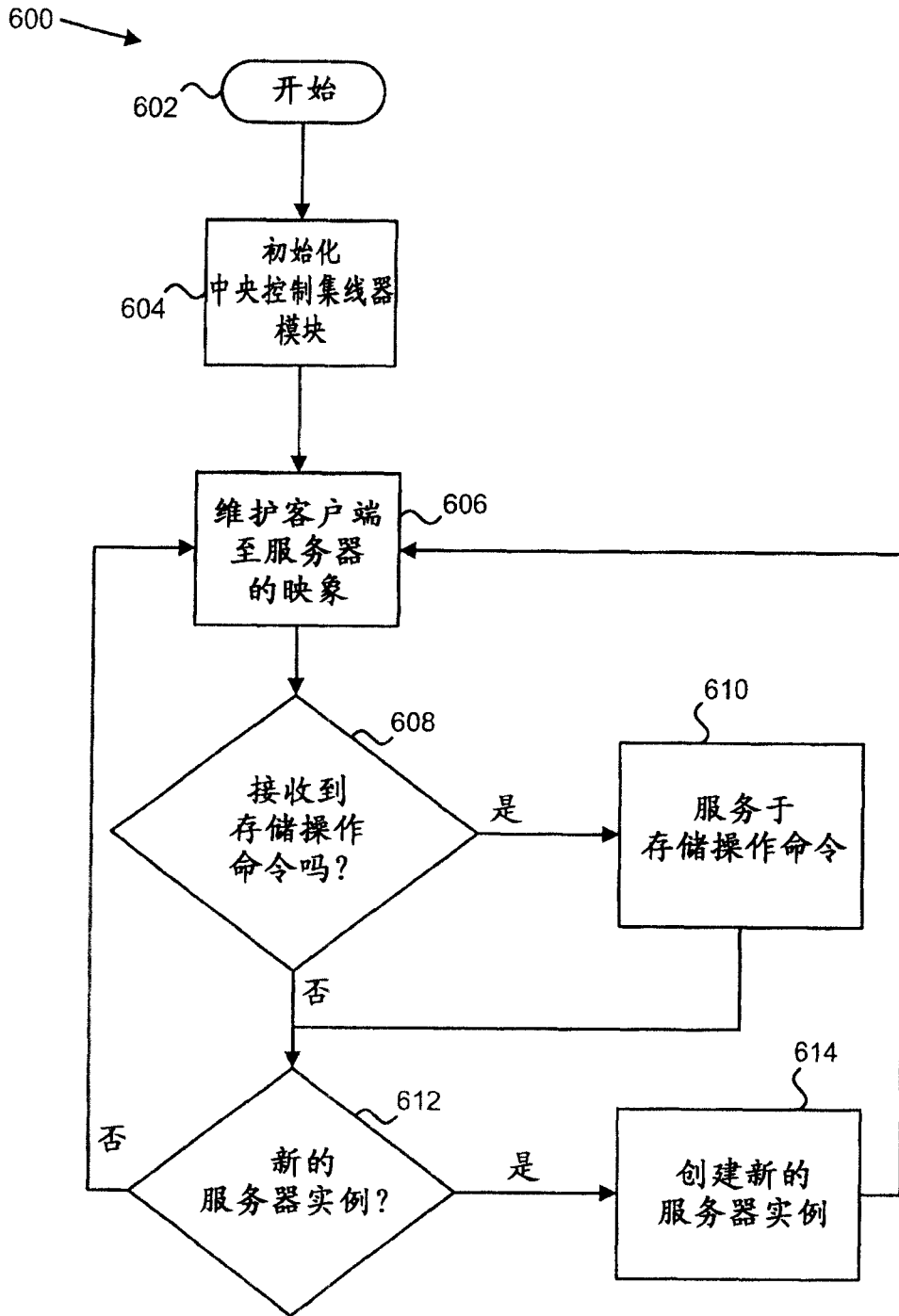


图6