

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.
G06F 9/46 (2006.01)



[12] 发明专利申请公布说明书

[21] 申请号 200780003270.6

[43] 公开日 2009 年 2 月 25 日

[11] 公开号 CN 101375250A

[22] 申请日 2007.1.23

[21] 申请号 200780003270.6

[30] 优先权

[32] 2006. 2. 3 [33] US [31] 11/346,822

[86] 国际申请 PCT/US2007/002020 2007.1.23

[87] 国际公布 WO2007/092167 英 2007.8.16

[85] 进入国家阶段日期 2008.7.17

[71] 申请人 甲骨文国际公司

地址 美国加利福尼亚州

[72] 发明人 尤金·霍 威尔逊·外·顺·禅

王德峰 安杰洛·普鲁西诺

托尔加·于雷克

[74] 专利代理机构 北京律盟知识产权代理有限责任公司
代理人 刘国伟

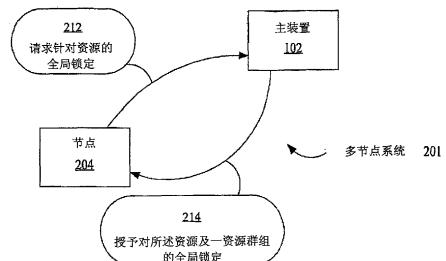
权利要求书 2 页 说明书 10 页 附图 6 页

[54] 发明名称

自适应区域锁定

[57] 摘要

本发明提供一种用于管理数据库系统中的锁定的方法及设备。主节点向第一请求者节点授予对第一资源及包括所述第一资源的一资源群组的锁定。所述请求者节点接收对应于所述资源群组的可指示针对所述群组中的第二资源的锁定已经存在的映射。如果所述请求者节点需要对位于所述群组中的资源的锁定，那么所述请求者节点授予其自身所述锁定而通知所述主节点。第二请求者节点请求针对所述资源群组中特定资源的锁定。所述第一请求者节点授予对所述特定资源的锁定且更新所述映射以指示不同的节点保持针对所述特定资源的锁定。



1、一种用于管理数据库系统中的锁定请求的方法，所述方法包含：

传输表示对资源的锁定的请求的消息；及

接收以下确认：

已授予对所述资源的所述锁定；及

已授予对包括所述资源的一资源群组的锁定。

2、如权利要求 1 所述的方法，其中接收所述确认包括接收识别所述资源群组的一个或一个以上资源被一个或一个以上其它节点锁定的映射。

3、如权利要求 1 所述的方法，其中所述消息包括对所述群组的锁定的请求。

4、一种用于管理数据库系统中的锁定请求的方法，所述方法包含：

在主节点处从第一节点接收对第一资源的锁定的请求；及

向所述第一节点发送以下确认：

已向所述第一节点授予对所述第一资源的所述锁定；及

已授予对包括所述第一资源的一资源群组的锁定。

5、如权利要求 4 所述的方法，其中发送所述确认包括发送识别所述资源群组的一个或一个以上资源被一个或一个以上其它节点锁定的映射。

6、如权利要求 4 所述的方法，其中所述请求还包括对所述资源群组的所述锁定的请求。

7、如权利要求 4 所述的方法，其中对所述第一资源的所述锁定与对所述群组的所述锁定是不同类型的锁定。

8、如权利要求 4 所述的方法，其中所述主节点确定所述资源群组。

9、一种用于管理数据库系统中的锁定请求的方法，所述方法包含：

传输表示对资源的锁定的请求的消息；及

接收：

已授予对所述资源的所述锁定的确认；

已授予对包括所述资源的一资源群组的锁定的确认；及

识别所述资源群组的一个或一个以上资源被一个或一个以上其它节点锁定的映射。

10、如权利要求 9 所述的方法，其中所述步骤包括：

所述第一节点确定其需要对所述资源群组中所包括的第二资源的第二锁定；

基于所述映射，所述第一节点确定无其它节点具有对所述第二资源的非兼容性锁定；及

所述第一节点授予对所述第二资源的所述第二锁定。

11、如权利要求 9 所述的方法，其中所述步骤包括：

所述第一节点确定其需要对所述资源群组中所包括的第二资源的锁定；

基于所述映射，所述第一节点确定另一节点具有对所述第二资源的非兼容性锁定；

所述第一节点从所述主节点请求对所述第二资源的所述锁定。

12、如权利要求 9 所述的方法，其中所述步骤包括：

在所述第一节点处接收指示第二节点请求对第二资源的锁定的消息；

响应于接收所述消息，在所述映射中指示另一节点保持对所述第二资源的所述锁定。

13、如权利要求 9 所述的方法，其中所述步骤包括：

在所述第一节点处接收指示第二节点请求对第二资源的锁定的消息；

响应于接收所述消息，在所述第一节点处确定释放对所述资源群组的所述锁定；及

通知所述主节点对所述资源群组的所述锁定的所述释放。

14、一种用于管理数据库系统中的锁定请求的方法，所述方法包含：

从第一节点传输表示对一资源群组的锁定的第一请求的第一消息；

在所述第一节点处接收已授予对所述资源群组的所述锁定的确认；及

当所述第一节点保持针对所述资源群组的所述锁定时，在所述第一节点处接收指示第二节点请求对资源的锁定的第二消息，其中所述资源群组包括所述资源；

向所述第二节点传输授予对所述资源的所述锁定的第三消息；

维持对所述资源群组的所述锁定；及

在所述第一节点上，致使识别所述资源群组的映射指示对所述资源的所述锁定由所述第二节点保持。

15、一种机器可读媒体，其携载一个或一个以上指令序列，当由一个或一个以上处理器执行时所述指令序列致使所述一个或一个以上处理器执行权利要求 1-14 中的任一者中所陈述的方法。

自适应区域锁定

技术领域

本发明涉及锁定资源，且更特定来说涉及聚合锁定。

背景技术

此部分中所说明的方法是可推行的方法，但未必是先前已构思或推行的方法。因此，除非另外指明，否则不应假设此部分中所说明的任何方法仅由于其被包括在此部分中而被看作是现有技术。

在多处理系统上运行的多个进程可存取共享资源，例如磁盘块。这些共享资源中的某些可一次仅由一个进程存取，而其它可由多个进程同时存取。因此，已研发“同步机制”来控制多个进程对共享资源的存取。所述同步机制向进程授予锁定。锁定授予所述锁定的保持者以特定方式存取特定资源的权利。一旦进程被授予锁定，那么所述进程保持或拥有所述锁定直到所述锁定被放弃、撤销或另外终止。锁定由数据结构（例如，旗语、读取/写入锁存及状态变量）表示。存在许多种类型的锁定。某些类型的锁定允许共享资源由许多进程同时共享（例如，共享读取锁定），而其它类型的锁定阻止授予对同一资源的任何类型的锁定（排他性写入锁定）。

负责授予锁定的实体在本文中称作锁定管理器。在单节点多处理系统中，锁定管理器通常是由存取共享资源的节点上进程执行及调用的软件组件。

与单节点系统相比，多节点系统由计算装置或节点的网络组成，其每一者可以是多处理系统。所述节点中的每一者可存取一组共享资源。多节点系统使用同步机制（称作全局同步机制）来控制所述多节点系统中的节点对所述组共享资源的存取。

全局锁定机制包括负责向多节点系统上的进程发出锁定的全局锁定管理器。为使节点存取共享资源，其由全局锁定管理器授予“全局锁定”。全局锁定是可由多节点系统中的节点上的全局锁定管理器向另一节点上的一个或一个以上进程授予的锁定以在多节点系统中的任一节点上执行的进程之间协调对所述共享资源的存取。

一种类型的全局锁定管理器是分布式锁定管理器，其由分布在多节点系统的节点上的局部锁定管理器组成，其中所述局部锁定管理器中的一者或一者以上在多节点系统中的每一节点上运行。每一锁定管理器负责协调所述局部锁定管理器的节点上的进程的全局锁定。局部锁定管理器是指与其在其上驻存的节点相关的局部锁定管理器；所述节点及在所述节点上运行的进程是指局部节点及与所述局部锁定管理器及所述节点相关的局部进程。

驻存在节点上的局部锁定管理器向其它节点上的锁定管理器及在与所述局部锁定管理器相同的节点上运行的进程发出全局锁定。需要对资源的由非局部锁定管理器管理的全局锁定的进程从其局部锁定管理器请求所述全局锁定。如果局部锁定管理器已经保持兼容性全局锁定，那么所述局部锁定管理器向所述局部进程发出全局锁定。如果所述局部锁定管理器不保持兼容性全局锁定，那么所述局部锁定管理器首先从所述非局部锁定管理器获得一个。一旦获得，那么所述局部锁定管理器向所述局部进程发出所述全局锁定。

为便于表达，局部锁定管理器向局部进程发出的全局锁定在本文中称作局部锁定。因此，局部锁定管理器从另一锁定管理器获得全局锁定且向局部进程发出兼容性局部锁定。

而且，为便于表达，在本文中将节点说明为执行动作且是动作的对象。然而，此只是一种表达节点上的一个或一个以上进程正在执行动作或是动作的对象的方便方式。举例来说，可将请求、获得及发出全局锁定或局部锁定的锁定管理器说明为请求、获得及发出全局锁定或局部锁定的节点。

获取全局锁定可比仅获取局部锁定获取起来更昂贵。此是因为全局锁定可招致局部锁定与另一节点上的局部锁定管理器之间的节点间通信及交互作用。

所述交互作用可招致特别昂贵的操作形式，称作兵（ping）。当必须将驻存在一个服务器的高速缓存器中的资源版本供应到不同服务器的高速缓存器时，兵发生。因此，当在节点 A 修改其高速缓存器中的资源 x 之后另一节点 B 请求资源 x 时，兵发生。

高速缓存器融合

一种执行兵的方式称作高速缓存器融合。执行节点之间的资源的高速缓存器拷贝的传送来加速锁定机制。图 1 是图解说明根据本发明的实施例的多节点系统 101 及用于请求及传送经高速缓冲存储的资源的高速缓存器融合协议。系统 101 中的节点可直接或经由网络（例如，LAN 或因特网）而彼此进行通信。为获取对共享资源的全局锁定，当请求节点 104 从特定共享资源的锁定管理器驻存于其上的主节点 102 请求对所述特定共享资源的锁定（步骤 112）时，所述高速缓存器融合协议开始。

主节点 102 接收请求且确定任何其它节点是否保持对所述共享资源的非兼容性锁定。在无节点保持对所述共享资源的非兼容性锁定的简单情形下，所述主节点直接向所述请求节点授予所述锁定。如果所述主节点自身保持对所述共享资源的非兼容性锁定，那么所述主节点将最终直接向所述请求节点授予对所述共享资源的锁定。否则，另一节点（即，保持节点 106）保持对所述共享资源的非兼容性锁定。

所述主节点向保持节点 104 发送指示请求节点 104 请求对所述特定共享资源（保持节点 106 保持其锁定）的锁定的消息（步骤 114）。锁定节点 106 授予所述锁定且可直接向请求节点 104 发送所述共享资源的拷贝。在某些情形下，即使保持节点 106 保持对所述共享资源的兼容性锁定（例如，请求节点 104 请求对所述资源的共享锁定

且保持节点 106 保持对所述资源的共享锁定) 将触发从保持节点 106 到请求节点 104 的互连消息, 因为保持节点 106 可能保持所述共享资源的脏的或已修改版本。

最后, 一旦请求节点 104 接收到所述共享资源及所述锁定, 请求节点 104 通知主节点 102 (步骤 118) 请求节点 104 具有对所述共享资源的锁定。因此, 对共享资源的锁定的每一请求可导致产生四个节点间消息。因此, 需要降低获取全局锁定的成本的技术。

一种降低获取全局锁定的成本的技术使用“主控技术”, 其基于对共享资源的存取的图案向一子组共享资源指派主节点。(共享资源的主节点支配其它节点对所述共享资源的存取。) 举例来说, 如果对一部分共享数据的大部分存取由特定节点执行, 那么将所述节点指派为所述部分共享数据的主节点。此降低节点之间的消息传送开销, 因为将必须获取更少的全局锁定(由于所述特定节点需要对所述部分共享数据的大部分存取)。对所述部分共享数据的未来存取将仅需要所述特定节点授予局部锁定。然而, 主控不消除执行更多指令来获取全局锁定的成本。

用以降低获取全局锁定的成本的另一种技术是使用粗粒锁定。在此方案中, 在较高级的粒度(例如, 表格或文件)而不是在较精细级的粒度(例如, 行或磁盘块)处获取锁定。当在较高级的粒度处获取锁定时, 隐含地授予较精细级的粒度处的共享数据级。举例来说, 如果针对整个表格获取全局锁定, 那么暗指且没有必要获取所述表格的行或块的个别全局锁定, 从而避免获得每一行及块的全局锁定的成本。

此技术的优点是其不依赖于主节点的指派。然而, 显著缺点是此技术可导致假争用。具体来说, 如果节点需要修改已由另一节点在冲突模式中锁定的表格中的行, 那么所述节点必须放弃对所述表格的锁定, 虽然所述两个节点可存取不同的行或甚至不同的块。

另一种降低获取全局锁定的成本的技术是使用层级锁定。在此方案中, 首先在所述层级中的较高级(例如, 表格)处获取锁定。如果在所述层级的较高级处获取全局锁定, 那么隐含地在所述层级的较低级处授予全局锁定。当另一节点随后需要在冲突模式中存取所述层级的较低级(例如, 行或块)中的数据时, 所述第一节点降级其锁定且获取所述层级中的所述较低级处的锁定。

此技术的缺点是获得全局锁定的成本高涨且转向其锁定请求触发所述降级的请求节点。为实现所述请求, 执行用以获取所述层级的所述较低级处的所有共享数据的全局锁定的工作。执行此工作, 尽管所述请求节点仅请求对所述共享数据的一小部分的锁定。

如明确显示, 需要降低获取全局锁定的避免伴随上文所说明的用于降低全局锁定的成本的技术的缺陷的技术。

发明内容

附图说明

本文以实例的方式而非限定的方式在附图的图式中图解说明本发明，且在附图中相同的参考编号指代相同的元件，且附图中：

图 1 是图解说明根据本发明的实施例的多节点系统及用于请求及传送经高速缓冲存储的资源的方法的框图；

图 2 是图解说明根据本发明的实施例的多节点系统及用于获得对一群组共享资源的锁定的方法的框图；

图 3 是图解说明根据本发明的实施例的特定节点上的一群组共享资源的映射；

图 4 是图解说明根据本发明的实施例的多节点系统及用于节点获得对一资源群组中的资源的锁定的方法的框图，其中所述节点维持对所述资源群组的锁定且另一节点维持对所述资源的锁定；

图 5 是图解说明根据本发明的实施例的多节点系统及用于节点获得对一资源群组中的资源的锁定的方法的框图，其中另一节点维持对所述资源群组的锁定；及

图 6 是可用于实施本发明的实施例的计算机系统的框图。

具体实施方式

在以下说明中，为便于解释，论述了许多具体细节以提供对本发明的透彻理解。然而，将明了，不需要这些具体细节的情况下也可以实践本发明。在其它例示中，在框图形成中显示众所周知的结构及装置，以避免对本发明造成不必要的遮蔽。

概述

本发明提供一种用于管理数据库系统中的锁定的方法及设备。第一请求者请求对第一资源的全局锁定且主节点授予对所述第一资源以及包括所述第一资源的一资源群组的锁定。授予对所述资源群组的全局锁定是基于所述第一请求者将倾向于存取所述群组中的其它资源的假设来执行的。所述请求者节点接收对应于所述资源群组的映射，其可指示另一节点所保持的全局锁定对于所述群组中的第二资源已经存在。所述请求者可授予其自身与对所述资源群组的锁定兼容的局部锁定而通知所述主节点。

第二请求者节点可请求对所述资源群组中的特定资源的全局锁定。然后，所述第一请求者节点可放弃或向下转换对所述特定资源的锁定且向所述第二请求者授予对所述特定资源的所述锁定。所述第一请求者可更新所述第一请求者节点的映射以指示不同的节点（即，所述第二请求者节点）保持对所述特定资源的锁定。如果所述第一请求者节点已修改或弄脏所述特定资源，那么所述第一请求者节点还将发送对应于所述修改的数据。

请求对资源的锁定

图 2 是图解说明根据本发明的实施例的多节点系统 101 及用于获得对一群组共享资源的锁定的方法。大体来说，资源是可由所述多节点系统中的节点共享的任何条目。资源的普通实例是数据库数据块。因此，一资源群组可对应于存储在永久性存储器中的一组两个或两个以上邻接数据块。

主节点 102 负责向其它节点授予对数据库系统中的共享资源的全局锁定。最初，节点 204 请求对资源的特定类型或一资源群组的全局锁定（步骤 212）。所述请求可包括对特定资源的全局锁定及对包括所述资源的特定资源群组的全局锁定两者的请求。来自节点 204 的所述请求还可以是仅对所述特定资源或所述特定资源群组的全局锁定的请求。

主装置 102 向节点 204 发送授予对至少一资源群组的全局锁定（例如，共享锁定）的请求（步骤 214）的消息。如果来自节点 204 的请求是针对特定的资源，那么主装置 102 动态地确定（或已静态地确定）所述特定资源是在特定资源群组（例如，连续数据块组）中。主装置 102 发送对应于所述资源群组的映射，其指示所述群组中所有资源的锁定及锁定类型。如果节点 204 上的局部进程请求所述映射中所识别的所述资源群组中的任何资源的全局锁定，其中所请求的全局锁定与对所述资源群组的全局锁定兼容，那么节点 204 的局部锁定管理器简单地授予所述局部进程所述锁定且更新所述映射而无需针对所述锁定请求主装置 102。

举例来说，如果所述资源群组的全局锁定是共享锁定，那么节点 204 将无需针对所述群组中的任何资源从主装置 102 请求共享锁定来向局部进程授予局部共享锁定。实质上，节点 204 仅在一个消息中接收多个资源的锁定。

针对所述资源群组授予的全局锁定可与所述特定资源的全局锁定相同或不同。举例来说，如果节点 204 请求对所述特定资源的排他性锁定，那么主装置 102 可授予对所述特定资源的所述排他性锁定且授予对包括所述特定资源的所述资源群组的共享锁定。在所述情形下，所述映射自身还可指示存在授予给节点 204 的对所述特定资源的排他性锁定。

如果节点 204 最初请求（即，步骤 1）对资源的至少一个全局锁定且所述系统中的另一节点保持对所述资源的非兼容性锁定，那么主装置 102 授予对包括所述资源的一资源群组的全局锁定且向另一节点发送指示节点 204 请求对所述资源的全局锁定的消息。那么，除对所述资源群组的锁定的类型以外，所述映射可指示另一节点保持非兼容性全局锁定。

取决于所述数据库服务器网络中的节点如何配置，另一节点可不向节点 204 授予对所述资源的全局锁定直到预定的时间周期流过，尤其在所述另一节点具有对所述资源的排他性锁定时，因为排他性锁定的获得在节点间消息传送方面相对“昂贵”。如果已从主装置 102 接收到对所述资源的某数量的请求，那么不管所述预定时间周期，所述另一节点还可经配置以释放其对所述资源的全局锁定。

所述另一节点最终将对所述资源的全局锁定向下转换为兼容性锁定模式或完全

放弃所述锁定且向节点 204 发送所述资源的最新版本的拷贝，类似于相对于图 1 所说明的过程。然后，节点 204 通知主装置 102 节点 204 保持对所述资源的锁定。因此，在其中节点 204 请求对已锁定资源的非兼容性全局锁定且被授予对包括所述已锁定资源的一资源群组的全局锁定的情形下，节点 204 将接收两个消息：一个来自主装置 102 的授予对所述资源群组的锁定的消息，及一个来自所述另一节点的授予对所述资源的锁定的消息。

区域映射

除所述群组的全局锁定以外，发送到节点 204 的映射还可指示所述系统中的另一节点保持对所述群组中的特定资源的锁定。图 3 是图解说明根据本发明的实施例的一群组共享资源的映射 300 的图表，所述映射在特定节点上，例如节点 204。映射 300 可以是保持节点 204 的关于所述资源群组及可能所述群组中的个别资源的锁定模式的信息的任何形式的数据结构（例如，字符阵列）或数据文件（例如，包含经组织上下文的简单上下文文件），如下文所说明。

在此实例中，映射 300 指示已授予对对应于映射 300 的所述资源群组的共享全局锁定 302。而且，映射 300 指示另一节点保持对对应于块 7（“资源 7”）的资源的排他性全局锁定 304。最后，映射 300 指示节点 204 已向自身授予共享锁定（即，节点 204 的局部锁定管理器向与节点 204 相关的局部进程授予所述共享锁定）。

映射（例如，映射 300）可维持与所述资源群组中的资源相关的其它信息，例如，所述系统中的哪一节点具有排他性锁定，（举例来说）及所述锁定何时被授予。而且，虽然图 3 显示映射 300 对应于十六个资源，但映射 300 可对应于任何数量的资源，此可由用户（例如数据库管理员）静态地确定，或由主装置 102 基于（举例来说）哪些资源请求地最频繁，冲突发生频率，冲突何时发生等动态地确定。

来自群组锁定保持者的随后锁定请求

图 4 是图解说明根据本发明的实施例的多节点系统 401 及用于节点（例如，请求节点 404）获得对一资源群组中的资源的全局锁定的方法，其中请求节点 404 维持对所述资源群组的全局锁定且另一节点（例如，保持节点 406）维持对所述资源的全局锁定。

如果请求节点 404 需要对映射 300 中的资源 7 的全局锁定且所述锁定与保持节点 406 所保持的对资源 7 的当前锁定不兼容，那么根据高速缓存器融合协议向主装置 102 及保持节点 406 发送节点间消息。举例来说，请求节点 404 向主装置 102 发送消息以请求对资源 7 的全局锁定（步骤 412）。主装置 102 确定哪一节点保持对资源 7 的排他性锁定且然后向保持节点 406 发送消息（步骤 414）以通知保持节点 406 请求节点 404 请求对资源 7 的排他性全局锁定。保持节点 406 将其全局锁定向下转换为空模式或完全放弃其全局锁定并向请求节点 404 发送消息（步骤 416）以授予对资源 7 的锁定。保持节点 406 还发送反应可能已发生的对资源 7 的任何修改的资源 7 的最新版本

的拷贝。请求节点 404 通知主装置 102（步骤 418）请求节点 404 保持对资源 7 的排他性全局锁定。主装置 102 更新其自己的指示请求节点 404 现在保持对资源 7 的排他性锁定的信息的表格。

在其中请求节点 404 请求映射 300 所识别的所述资源群组中资源的全局锁定且无其它节点保持对所述资源的全局锁定或无其它节点保持对所述资源的非兼容性全局锁定的情形下，不需要产生多个节点间消息。然而，如果请求节点 404 想要映射 300 所识别的所述资源群组中的资源的排他性锁定，其在共享模式中保持区域锁定，那么节点 404 将此全局锁定扩展为融合锁定且向主装置 102 提交所述升级请求。请求节点 404 可在映射 300 中指示所述锁定扩展。

第二节点请求对群组中的资源的锁定

图 5 是图解说明根据本发明的实施例的多节点系统 501 及用于节点（例如，请求节点 504）获得对一资源群组中的资源的锁定的方法的框图，其中另一节点（例如，保持节点 506）维持对所述资源群组的锁定。

在保持节点 506 获得对应于一资源群组的映射 300 之后，另一节点（例如，请求节点 504）可请求对对应于映射 300 的所述资源群组中所包括的资源的全局锁定。举例来说，假设请求节点 504 请求对与映射 300 中的块 3（即，“资源 3”）相关联的资源的排他性全局锁定且保持节点 506 在共享模式中保持所述群组锁定。请求节点 504 向主装置 102 发送对资源 3 的排他性全局锁定的请求（步骤 512）。主装置 102 基于（例如）主装置 102 自己的映射确定保持节点 506 具有对包括资源 3 的所述资源群组的共享全局锁定。

此时至少存在两种可能情形：1)保持节点 506 保持对资源 3 的个别共享全局锁定（即，除群组锁定以外恰好保持对所述资源的全局锁定）或 2)保持节点 506 不保持对资源 3 的个别全局锁定。如果保持节点 506 保持排他性锁定，那么主装置 102 将具有所述信息。

如果保持节点 506 保持对资源 3 的个别共享全局锁定，那么后跟用于释放及授予锁定的高速缓存器融合协议。具体来说，一旦保持节点 506 由主装置 102 通知请求节点 504 请求排他性全局锁定，那么保持节点 506 将其对资源 3 的锁定从共享模式向下转换为空模式，或放弃对资源 3 的锁定。然后，保持节点 506 向节点 504 发送资源 3 的拷贝且向请求节点 504 授予对资源 3 的排他性全局锁定。请求节点 504 通知主装置 102 请求节点 504 保持对资源 3 的排他性全局锁定。保持节点 506 上的映射 300 经更新以指示对资源 3 的排他性全局锁定由另一节点保持。

如果保持节点 506 不保持对资源 3 的个别全局锁定，那么主装置 102 向保持节点 506 发送另一节点需要对资源 3 的排他性锁定的消息（步骤 514）。保持节点 506 通知主装置 102 其不保持个别全局锁定（步骤 516）。主装置 102 更新其对应于包括资源 3 的所述资源群组的映射或主装置 102 更新指示请求节点 504 现在保持对资源 3 的排他

性全局锁定的另一记录。主装置 102 向请求节点 504 授予最初请求的所述排他性锁定（步骤 518）。

其它锁定情形

可存在其中需要将对应于特定映射的整个资源群组的锁定升级或降级为不同锁定（例如，从排他性锁定降级为共享锁定）或其中需要废除所述整个映射的情况。

举例来说，假设图 2 中的节点 204 维持映射 300 且节点 204 在共享模式中保持对一资源群组的全局锁定。节点 204 可随后获得映射 300 中所识别的资源的显著部分的多个排他性锁定。将所述群组锁定升级到排他性模式可更有效，此取决于多节点系统 201 中的节点之间的现存业务及可影响升级的其它因素。

更普通的情形可是一个或一个以上其它节点请求映射 300 中所识别的所述资源群组中的资源的多个非兼容性全局锁定发生。节点 204 可基于一个或一个以上因素（例如，多少对所述资源群组中的资源的个别锁定由其它节点保持，未来的更多冲突请求的可能性及节点 204 是否请求对所述群组中的任何资源的锁定）来确定需要将所述群组锁定降级或废除。在所述群组锁定的降级的情形中，节点 204 通知主装置 102 将主装置 102 上对应于映射 300 的映射（例如）从排他性锁定降级为共享锁定或从共享锁定降级为空锁定。

节点 204 可改为通知主装置 102 废除所述群组锁定。其中可需要废除的情形是当所述锁定模式被转换为不同类型的锁定模式（例如，主要-读取锁定）时。为转换为不同类型的锁定，必须首先将所述群组锁定废除且然后获得新的锁定类型。

可需要群组锁定的传递的另一情况是主节点的故障。如果主装置 102 出故障，那么在所述多节点系统中的节点中选择新的主节点。维持一资源群组的映射的每一非主节点将其相应的映射发送到所述新主节点使得所述新主节点能够授予对个别资源的锁定及对资源群组的锁定。

至此，已在分布式锁定环境中说明所论述的本发明的实施例。然而，本发明的实施例并不限于分布式环境。其还可实施于非分布式锁定环境中。

硬件概述

图 6 是图解说明可在其上实施本发明的实施例的计算机系统 600 的框图。计算机系统 600 包括总线 602 或用于传递信息的其它传递机构，及与总线 602 耦合的用于处理信息的处理器 604。计算机系统 600 还包括主存储器 606，例如随机存取存储器（RAM）或其它动态存储装置，其耦合到总线 602 以用于存储将由处理器 604 执行的信息及指令。主存储器 606 也可以用来存储临时变量或执行将由处理器 604 执行的指令期间的其它中间信息。计算机系统 600 进一步包括只读存储器（ROM）608，或耦合到总线 302 的用于存储处理器 604 的静态信息及指令的其它静态存储装置。提供存储装置 610（例如，磁盘或光盘）并将其耦合到总线 602 以用于存储信息及指令。

计算机系统 600 可经由总线 602 耦合到显示器 612（例如，阴极射线管（CRT）），

以向计算机用户显示信息。输入装置 614（其包括字母数字键及其它键）耦合到总线 602，以将信息及命令选择传递到处理器 604。另一类型的用户输入装置是光标控制 616（例如，鼠标、轨迹球或光标方向键），以将方向信息及命令选择传递到处理器 604 并控制显示器 612 上的光标移动。此输入装置通常在两个轴（第一轴（例如，x）及第二轴（例如，y））上具有两个自由度，其允许装置在平面中规定位置。

本发明涉及用于实施本文所说明的技术的计算机系统 600 的使用。根据本发明的一个实施例，所述技术由计算机系统 600 响应于处理器 604 执行主存储器 606 中所包含的一个或一个以上指令的一个或一个以上序列而执行。可以从另一机器可读媒体（例如，存储装置 610）将所述指令读入主存储器 606 中。主存储器 606 中所包含的指令序列的执行致使处理器 604 执行本文所说明的过程步骤。在替代性实施例中，可使用硬接线电路来替代软件指令或与软件指令组合使用来实施本发明。因此，本发明的实施例并不限于硬件电路与软件的任何具体组合。

本文所用术语“机器可读媒体”是指参与提供致使机器以特定方式操作的数据的任何媒体。在使用计算机系统 600 实施的实施例中，各种机器可读媒体涉及（举例来说）将指令提供到处理器 604 供执行。所述媒体可采用许多形式，包括但不限于：非易失性媒体、易失性媒体及传输媒体。非易失性媒体包括（举例来说）光盘或磁盘，例如存储装置 610。易失性媒体包括动态存储器，例如主存储器 606。传输媒体包括同轴电缆、铜线及光纤，其中包括包含总线 602 的线。传输媒体也可以采用声波或光波的形式，例如在无线电波及红外线数据通信期间产生的那些声波或光波。

机器可读媒体的常见形式包括（举例来说）软盘、柔性盘、硬盘、磁带或任何其它磁性媒体、CD-ROM、任何其它光学媒体、穿孔卡片、纸带、任何其它具有孔图案的物理媒体、RAM、PROM 及 EPROM、快闪 EEPROM、任何其它存储器芯片或盒式磁盘、如下文所说明的载波或可由计算机进行读取的任何其它媒体。

各种形式的机器可读媒体可涉及将一个或一个以上指令的一个或一个以上序列携载到处理器 604 供执行。举例来说，所述指令可最初携载于远端计算机的磁盘上。所述远端计算机可将所述指令载入到其动态存储器内并经由电话线使用调制解调器发送所述指令。计算机系统 600 本地的调制解调器可在电话线上接收数据并使用红外线传输器将所述数据转换为红外线信号。红外线检测器可接收携载于所述红外线信号中的数据且适当电路可将所述数据置于总线 602 上。总线 602 将所述数据携载到主存储器 606，且处理器 604 从主存储器 606 检索所述指令并执行所述指令。主存储器 606 接收的指令可视需要地在由处理器 604 执行之前或之后存储在存储装置 610 上。

计算机系统 600 还包括耦合到总线 602 的通信接口 618。通信接口 618 提供耦合到网络链路 620 的双向数据通信，网络链路 620 连接到局域网 622。举例来说，通信接口 618 可以是综合服务数字网络（ISDN）卡或调制解调器，以将数据通信连接提供到对应类型的电话线。作为另一实例，通信接口 618 可以是局域网（LAN）卡，以将数据通信连接提供到兼容的 LAN。也可以实施无线链路。在任何所述实施方案中，通

信接口 618 发送并接收电信号、电磁信号或光信号，所述信号携载表示各种类型的信息的数字数据流。

网络链路 620 通常经由一个或一个以上网络向其它数据装置提供数据通信。举例来说，网络链路 620 可通过局域网 622 向主机计算机 624 或因特网服务提供商（ISP）626 操作的数据装备提供连接。ISP 626 又经由世界范围的包数据通信网络 628（目前一般称作“因特网”）提供数据通信服务。局域网 622 及因特网 628 两者使用携载数字数据流的电信号、电磁信号或光信号。穿过各种网络的信号及网络链路 620 上并穿过通信接口 618 的信号（其携载往来于计算机系统 600 的数字数据）是输送信息的载波的实例性形式。

计算机系统 600 可经由网络、网络链路 620 及通信接口 618 发送消息及接收数据（包括程序码）。在因特网的实例中，服务器 630 可经由因特网 628、ISP 626、局域网 622 及通信接口 618 传输应用程序的所请求码。

所接收码可随其接收而由处理器 604 处理，及/或存储在存储装置 610 中或其它非易失性存储器中供随后执行。以此方式，计算机系统 600 可获得载波形式的应用程序码。

在上述说明中，已参照可根据不同实施方案而改变的许多具体细节说明了本发明的实施例。因此，本发明及本申请者打算作为本发明实质的唯一及排他性标识是以发出此权利要求书的具体形式从本申请案发出的一组包括任何后续修正的权利要求项。针对所述权利要求书中所包含的术语所明确论述的任何定义应如权利要求书中所使用地那样来支配所述术语的意义。因此，权利要求书中未明确陈述的任何限制、元件、性质、特征、优点或属性不应以任何方式限制本权利要求书的范围。因此，须将本说明书及图式视为仅具有例证意义而非限制意义。

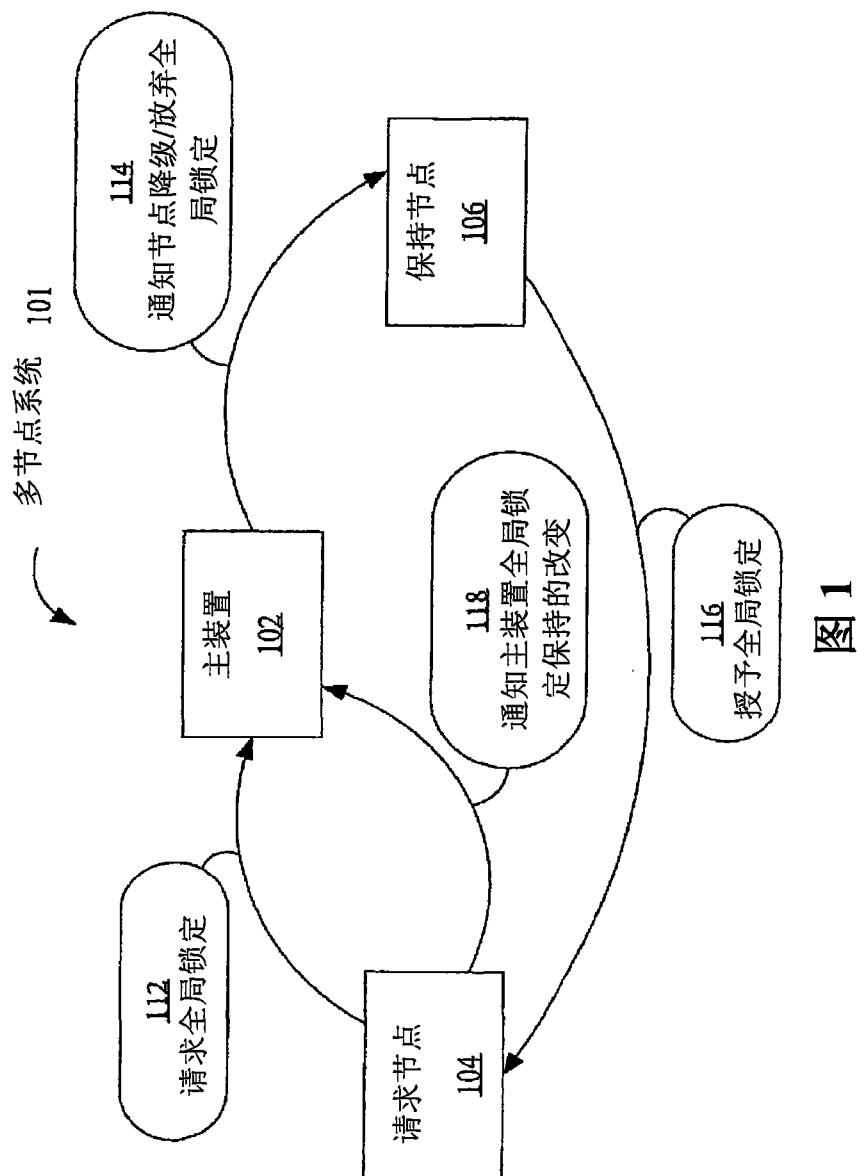


图 1

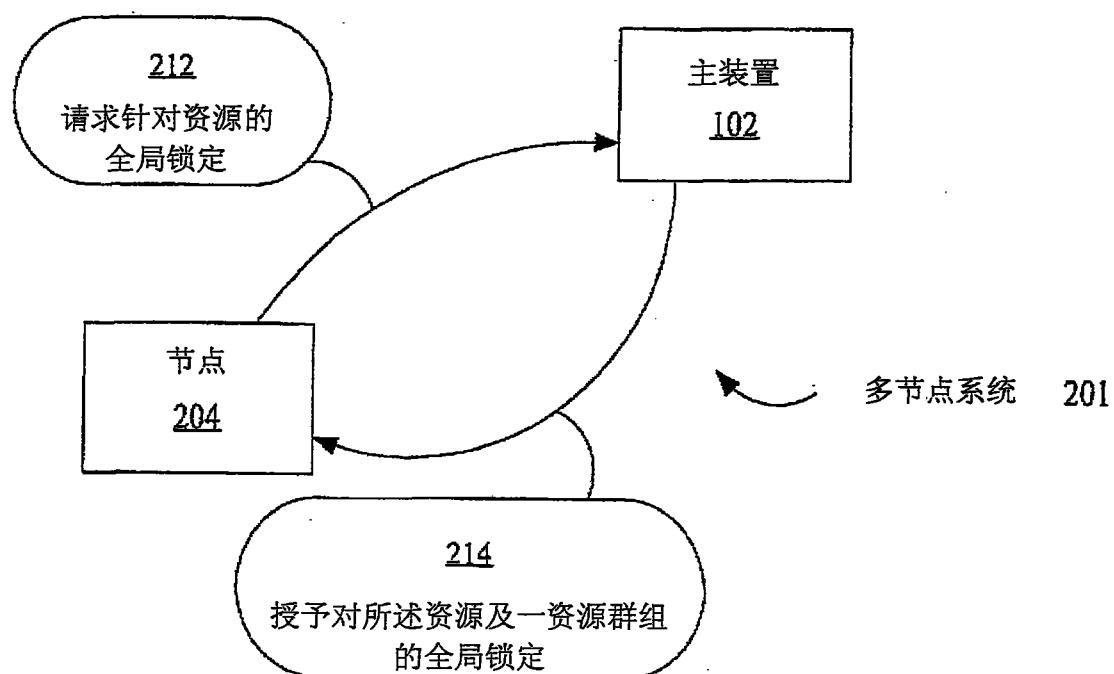


图 2

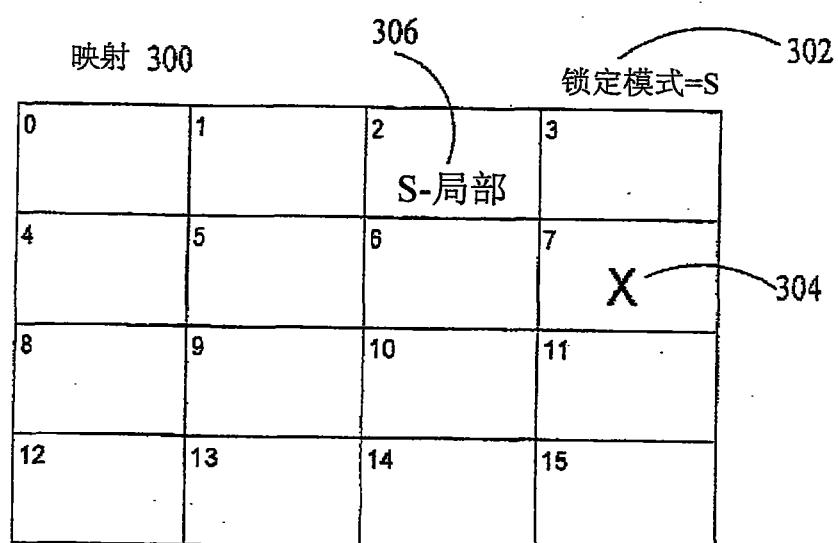
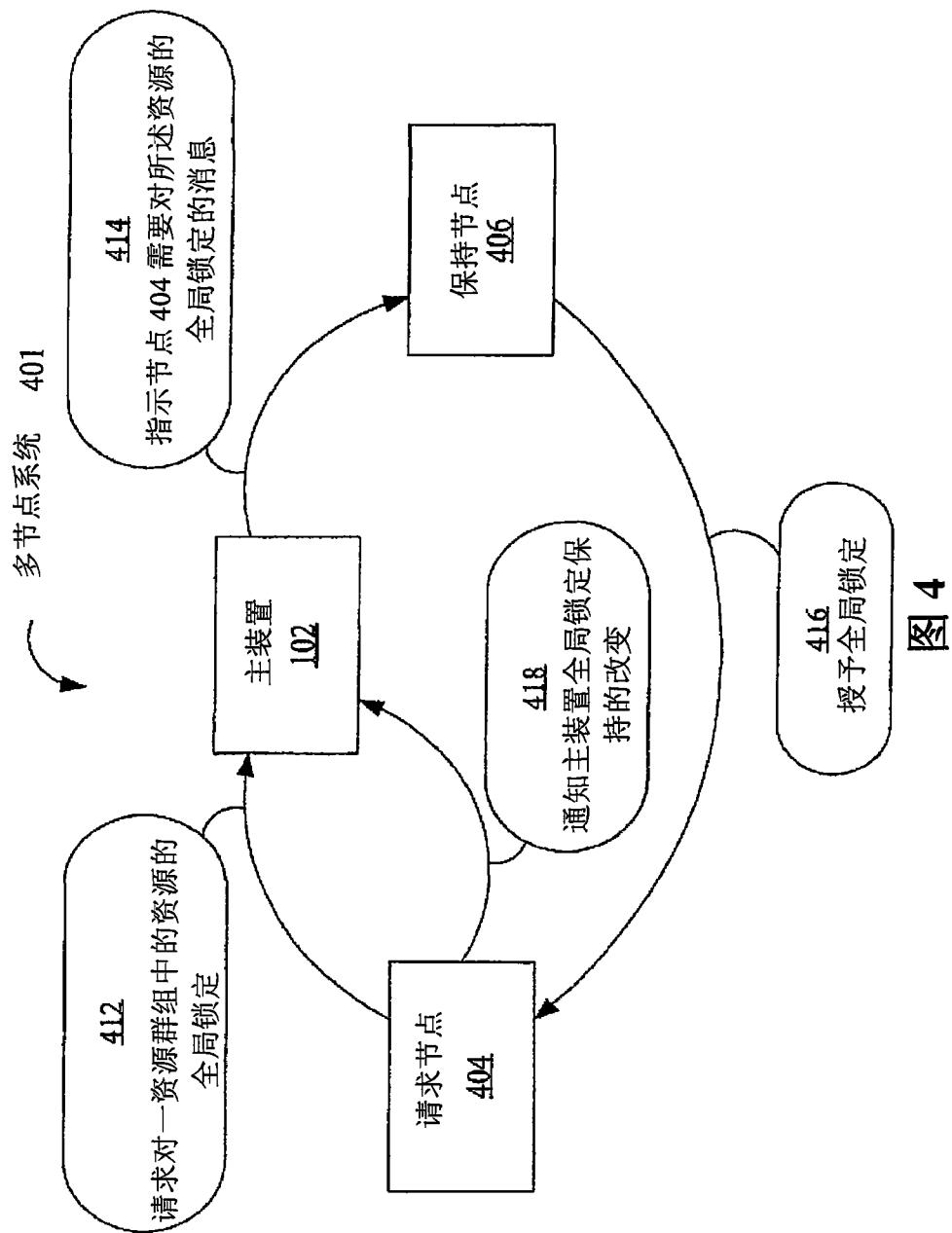


图 3



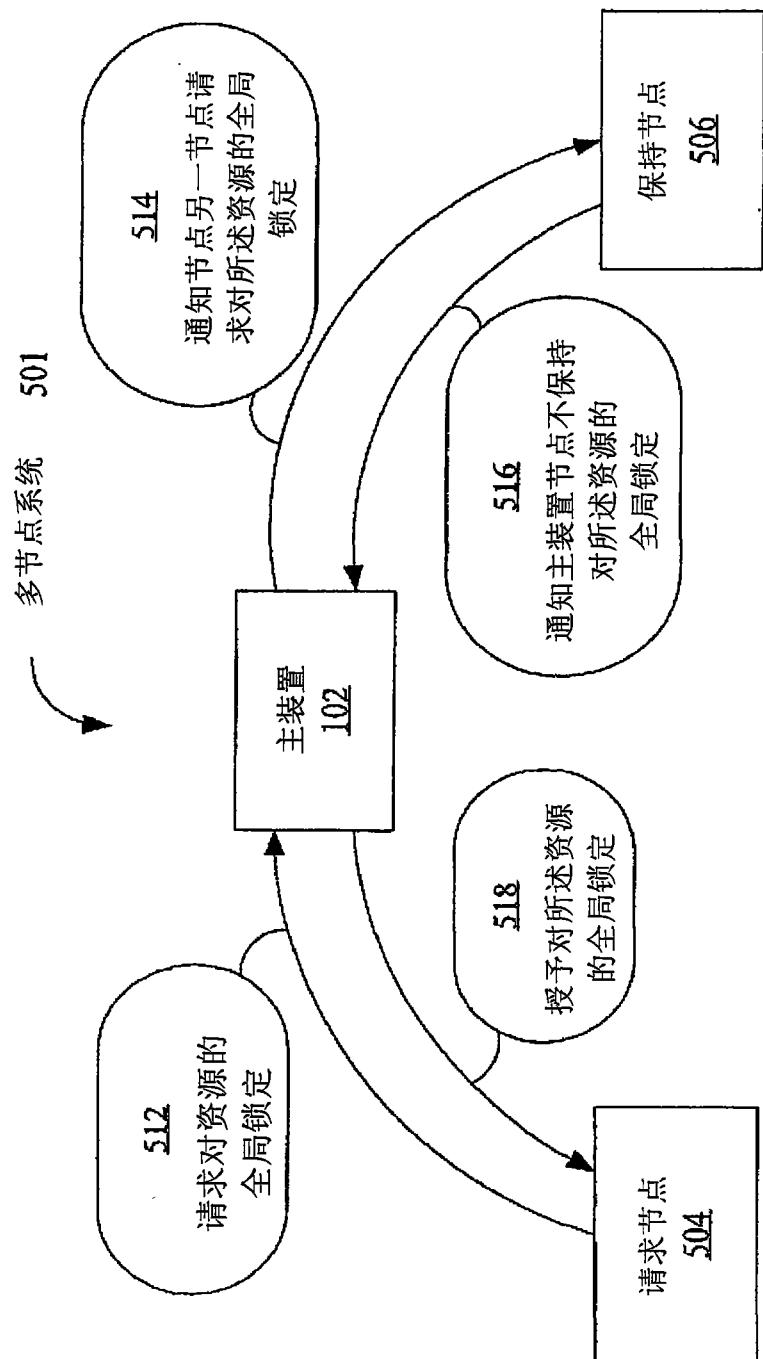


图 5

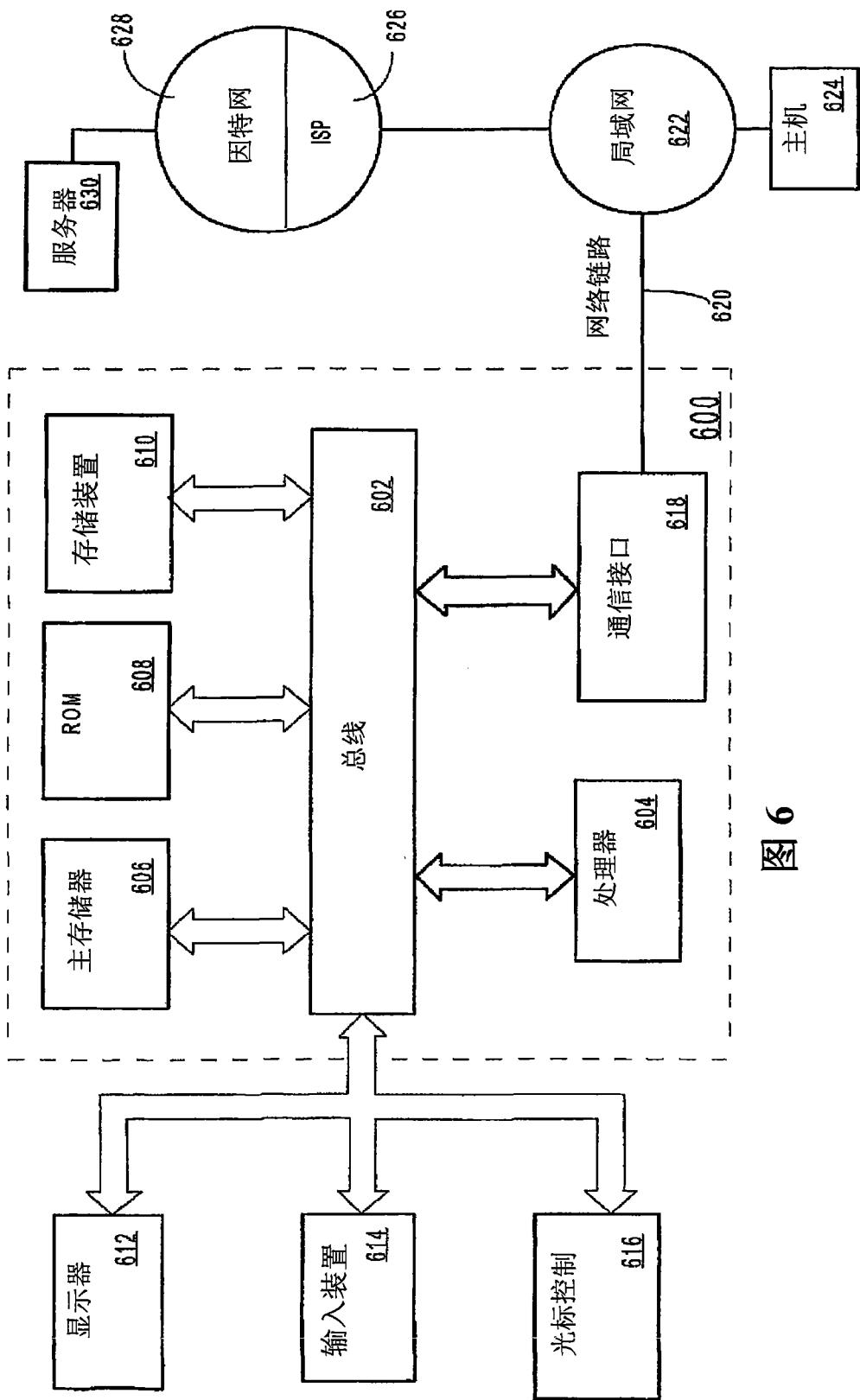


图 6