

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.  
G06F 15/16 (2006.01)  
G06F 15/173 (2006.01)



# [12] 发明专利申请公布说明书

[21] 申请号 200480040843.9

[43] 公开日 2007年2月21日

[11] 公开号 CN 1918557A

[22] 申请日 2004.12.3

[21] 申请号 200480040843.9

[30] 优先权

[32] 2003.12.22 [33] US [31] 10/745,187

[86] 国际申请 PCT/US2004/040475 2004.12.3

[87] 国际公布 WO2005/065111 英 2005.7.21

[85] 进入国家阶段日期 2006.7.24

[71] 申请人 查尔斯·李·麦克耐特

地址 美国加利福尼亚州

[72] 发明人 查尔斯·李·麦克耐特

[74] 专利代理机构 上海专利商标事务所有限公司  
代理人 钱慰民

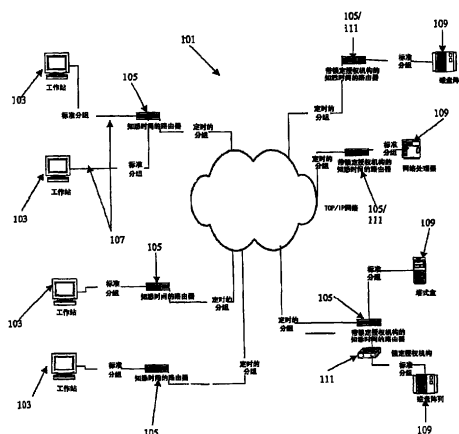
权利要求书 3 页 说明书 20 页 附图 13 页

## [54] 发明名称

使用网络设备中的独立绝对时间值来创建知悉时间的网络的系统和方法

## [57] 摘要

提供了一种知悉时间的网络，它具有固有时间基准，该固有时间基准使得能在不需要传送经同步的时间信息的网络环境中共享网络资源。内部时间基准是使用作为从 GPS 信号提取的时间坐标的绝对时间值来建立的。这些绝对时间值和定义的时钟参数块协调诸知悉时间的设备如何使用内部时间基准，并被进一步用于建立一网络域。能够满足该网络时钟域定时要求的设备可参与该域。一旦加入了一网络时钟域，客户机即可通过维护共享资源的数据一致性的锁定授权机构来访问共享资源。客户机、锁定授权机构和共享资源全部都是知悉时间的设备，它们参与相同的网络时钟域，并使用内部时间基准来通信。



1. 一种知悉时间的网络，包括：

多个知悉时间的设备，所述多个知悉时间的设备中的每一个都具有一绝对时间值，其中当与每个知悉时间的设备相关联的所述绝对时间值满足定义的定时参数时，即创建一固有时间基准，由此消除了与所述多个知悉时间的设备之中同步时间信息的需要。

2. 如权利要求 1 所述的知悉时间的网络，其特征在于，所述固有时间基准在要求所发送和接收的消息的严格排序的任何通信系统中使用。

3. 如权利要求 1 所述的知悉时间的网络，其特征在于，所述多个知悉时间的设备被配置成使用所述固有时间基准来同步地通信。

4. 如权利要求 1 所述的知悉时间的网络，其特征在于，所述多个知悉时间的设备中的每一个的绝对时间值是从 GPS 信号中提取的时间坐标。

5. 如权利要求 1 所述的知悉时间的网络，其特征在于，所述知悉时间的设备是客户机设备。

6. 如权利要求 5 所述的知悉时间的网络，其特征在于，所述客户机设备是工作站、服务器、CPU 或无线设备中的任何一种。

7. 如权利要求 1 所述的知悉时间的网络，其特征在于，所述知悉时间的设备是锁定授权机构。

8. 如权利要求 1 所述的知悉时间的网络，其特征在于，所述知悉时间的设备是共享网络资源。

9. 如权利要求 1 所述的知悉时间的网络，其特征在于，所述知悉时间的设备是路由器。

10. 如权利要求 1 所述的知悉时间的网络，其特征在于，所述定义的定时参数是时钟参数块，所述时钟参数块被配置成阐述所述多个知悉时间的设备的每一个必须能够满足的要求，以使用所述固有时间基准来与其它知悉时间的设备通信。

11. 如权利要求 7 所述的知悉时间的网络，其特征在于，所述锁定授权机构创建一网络时钟域。

12. 如权利要求 11 所述的知悉时间的网络，其特征在于，所述网络时钟域是使用所述时钟参数块来建立的。

13. 如权利要求 11 所述的知悉时间的网络, 其特征在于, 所述时钟参数块是基于所述共享资源的定时要求。

14. 如权利要求 11 所述的知悉时间的网络, 其特征在于, 所述客户机加入一网络时钟域以访问所述共享资源。

15. 如权利要求 14 所述的知悉时间的网络, 其特征在于, 所述锁定授权机构控制所述客户机对所述共享资源的访问。

16. 如权利要求 15 所述的知悉时间的网络, 其特征在于, 所述锁定授权机构通过控制对所述共享资源的访问来提供数据一致性。

17. 如权利要求 15 所述的知悉时间的网络, 其特征在于, 所述锁定授权机构使用多个锁定情景来控制对所述共享资源的访问。

18. 如权利要求 16 所述的知悉时间的网络, 其特征在于, 所述锁定授权机构被配置成基于发送时间戳来重排所接收的消息。

19. 如权利要求 17 所述的知悉时间的网络, 其特征在于, 所述锁定授权机构还被配置成将锁准许给第一客户机, 并将从第二客户机接收的请求重新排队。

20. 如权利要求 8 所述的知悉时间的网络, 其特征在于, 所述共享资源是网络资源。

21. 如权利要求 8 所述的知悉时间的网络, 其特征在于, 所述共享资源是大容量存储机制。

22. 如权利要求 21 所述的知悉时间的网络, 其特征在于, 所述大容量存储机制是数据库。

23. 如权利要求 21 所述的知悉时间的网络, 其特征在于, 所述大容量存储机制是磁盘阵列。

24. 如权利要求 20 所述的知悉时间的网络, 其特征在于, 所述共享资源是大容量存储器系统。

25. 如权利要求 20 所述的知悉时间的网络, 其特征在于, 所述共享资源是流媒体系统。

26. 如权利要求 1 所述的知悉时间的网络, 其特征在于, 所述固有时间基准是由所述锁定授权机构使用所述时钟参数块和所述绝对时间值生成的伪时钟。

27. 如权利要求 11 所述的知悉时间的网络, 其特征在于, 为了使用所述固有时间基准来通信, 所述知悉时间的设备和所述共享资源属于相同的网络时钟域。

28. 如权利要求 11 所述的知悉时间的网络, 其特征在于, 客户机必须加入所

述网络时钟域才能访问所述网络时钟域中的所述共享资源。

29. 如权利要求 2 所述的知悉时间的网络，其特征在于，所述对所接收的消息的重新排序是基于指示所述消息的发送时间的时戳。

30. 一种固有网络时间基准，包括：

绝对时间值，所述绝对时间值是从 GPS 信号提取的时间坐标；

参数块，所述参数块定义定时要求，

其中当所述绝对时间值满足由所述参数块定义的定时要求时，建立所述固有网络时间基准，由此使具有所述绝对时间值的设备无需同步时间信息即能够通信。

31. 如权利要求 30 所述的固有网络时间基准，其特征在于，所述固有时间基准在要求所发送和接收的消息的严格排序的任何通信系统中被使用。

32. 如权利要求 30 所述的固有网络时间基准，其特征在于，所述设备是知悉时间的设备。

33. 如权利要求 30 所述的固有网络时间基准，其特征在于，所述固有网络时间基准用作一知悉时间的设备的伪时钟。

34. 如权利要求 32 所述的固有网络时间基准，其特征在于，所述知悉时间的设备包括用于提取所述时间坐标的 GPS 接收器。

35. 如权利要求 32 所述的固有网络时间基准，其特征在于，所述定时要求是由共享资源阐述的。

使用网络设备中的独立绝对时间值来创建知悉时间的网络的系统和方法

## 发明背景

### 1. 发明领域

本发明一般涉及计算机网络，尤其涉及创建允许在不需要传送经同步的时间信息的分布式网络环境中共享网络资源的固有网络时间。

### 2. 背景

迄今为止，计算机网络还不能像通常在独立计算机系统中那样实现复杂的通信和资源共享能力。给计算机网络带来特定问题的两个领域包括尚未解决的时钟分布问题、以及不能在共享的资源之间维护数据一致性的问题。给计算机网络设计者带来主要问题的这些相同的领域一度也曾给计算机设计者带来了问题。但是，数十年前人们就已开发出对于关于计算机的这些问题的解决方案。不幸的是，当在网络级实现时，尚未能成功开发出对计算机起效的解决方案。

#### a. 时钟分布问题

在任何计算机系统中，对集中的和公共的时基的需要是不可避免的。时钟分布对于计算机设计来说是至关重要的，因为所有子系统和组件必须能够看到相同的时钟信号才能正确地工作。这对于独立计算机和计算机网络都是如此。时钟指示所有子系统和组件之间的高级协调。由此，所有组件都必须与计算机时钟同步才能工作。

常规的计算机网络试图通过使用时间服务器在网络上的计算机之间协调和同步时间来向网络引入定时，从而来解决时钟分布问题。但是，因为网络固有地受到等待时间问题的影响，所以时间服务器不能够提供计算机网络中所有组件都能依赖的准确和精确的时间基准。

计算机网络上的数据传输使用自计时波形。尽管这允许在接收器处能对所发送的信号进行准确的重建，但是没有办法来确定关于发送时钟对在网络上进行传输时的相位、相移或是延迟的定时的信息。因此等待时间变成未知量。对等待时间无

所知的话，就不可能在计算机网络上提供精确的时间基准。由此，网络时钟分布问题绝大多数的解决方案固有地受到与向网络上的每个客户机和资源传送时间相关联的带宽和等待时间的问题的限制。因此，网络中固有的局限所导致的结果是，任何最终需要“时间同步”或是“时间协调”的解决方案注定会失败。

#### b. 资源共享

使用多 CPU 系统的计算机能够通过允许每个 CPU 能够排它地且受控地访问所选择的资源来共享 I/O 资源和存储器资源。这种排它的访问称为锁。锁定处理确保在 CPU 可访问共享资源期间，该资源的数据将不会被除对该资源具有排它访问或者对该资源加锁的 CPU 以外的任何其它 CPU 所改变。经由锁定概念来准许排它访问可确保被共享的资源的数据一致性。

迄今为止，在网络级上使用资源共享概念的尝试都失败了。具体而言，允许一台网络计算机中的 CPU 共享位于另一台网络计算机上的资源和子系统的尝试业已证明是不成功的，其原因是众多的。这些原因凸显了网络环境与单机环境之间的差异。例如，网络在数据传递方面往往是不可靠的，它不能检测高数据传输误码率，并且不能纠正数据传输失败。此类缺陷使得要在共享资源时确保数据一致性非常困难。此外，与能同步地操作的计算机不同，计算机网络是以异步方式操作的。因此，子系统、组件和软件的设计者不能就从其它网络组件发来的消息的顺序或次序作出假定。此外，网络不能够区分所接收的消息的次序的优先级妨碍了它正确地建立、准许和释放对网络上所共享的资源的锁定的能力。

计算机网络缺少精确时间同步迫使计算机网络使用复杂的握手来共享资源。握手为一个客户机提供对共享资源的最小形式的受限和不受限访问，同时向其它客户机通知该共享资源的状态（即，被锁定或未被锁定）。但是在对共享资源的访问被准许时，所有客户机必须并发。握手技术固有地是薄弱的，因为如果与任何客户机的通信被丢失或被延迟，则所有客户机对共享资源的访问都将被拒绝。由此，尽管已证明在计算机级的资源共享是无价的，但是计算机与计算机网络之间的差异妨碍了资源共享在维持数据一致性同时的网络级实现。

最终，网络应当以与当前独立计算机所实现的能力相当的方式来执行任务。因此，需要一种能够消除网络时钟分布问题，由此允许在众多客户机之间有效地共享计算机资源，而同时在被共享的资源之间维持数据一致性的系统和方法。如以下详细记载的本发明通过给出一种建立固有网络时间基准的方法，并进一步给出一种

用于不需要传送时间信息的共享网络资源的系统来解决这一问题。

### 发明概述

为了克服上述问题，本文中所记载的系统和方法为知悉时间的设备提供一种固有的网络时间基准，它能消除在知悉时间的设备之间传送经同步的时间信息来进行通信的必要。更具体地，该固有的网络时间基准在与知悉时间的设备相关联的绝对时间值满足定义的定时要求时创建的。具体而言，知悉时间的设备通过从所接收的 GPS 信号提取时间坐标来获得绝对时间值。如果该绝对时间值满足在参数块中所阐述的定义的定时要求，则可建立使这些知悉时间的设备能够像在它们之间生成并分布了公共时钟那样进行交互的合成网络时钟或者伪网络时钟。

在本发明的另一个实施例中，本文中所描述的系统和方法还提供了一种知悉时间的网络。具体而言，知悉时间的网络使用与每个知悉时间的设备相关联的绝对时间值来创建与计算机时钟等效的时间基准。构成基本的知悉时间的网络的知悉时间的设备包括，但不限于，客户机、路由器、锁定授权机构和共享资源。锁定授权机构负责创建和维护网络时钟域，它提供客户机与共享资源之间受控的锁定服务。具体而言，锁定授权机构使用时钟参数块来创建网络时间域。时钟参数块与绝对时间值一起被用来生成伪时钟。时钟参数块定义绝对时间值的定时要求和精确值，客户机必须保证这些才能够参与其中。一旦客户机加入了网络时钟域，该客户机即使用伪时钟来与域中其它知悉时间的设备通信。具体而言，客户机请求对共享资源的访问，而锁定授权机构根据基于发送的时间戳来将排它锁准许给客户机。

在又一个实施例中，提供了一种在无级别网络中控制共享资源并同时确保与该共享资源相关联的数据的完整性和一致性的方法。

在另一个实施例中，提供了一种根据基于发送时间的的时间戳来对传入的分组进行排序、定序和分类的方法。类似地，还提供了一种使用绝对时间值来对所接收的分组进行排序和定序的方法。还提供了一种使用绝对时间值来创建发送时间戳的系统和方法。

在又一个实施例中，提供了一种用于实现可在许多设备之中访问、锁定和共享网络资源的健壮网络的系统和方法。

在再一个实施例中，提供了一种使用与知悉时间的设备相关联的绝对时间值来建立网络时钟域的方法。

## 附图简述

在附图中通过示例方式而非限定方式示出了本文中所讲授的本发明的优选实施例，附图中：

图 1 是根据本发明来使用的基本的知悉时间的网络；

图 2 是使用三个单独的网络时钟域来向至少三个网络资源提供经同步的访问的知悉时间的网络；

图 3 是描述锁定授权机构所采取的用来初始化网络时钟域的过程的流程图；

图 4 和 4a 都是示出客户机加入网络时钟域的过程的流程图；

图 5 是示出在锁定授权机构标识到被连上的网络资源时所执行的枚举和验证过程的流程图；

图 6 是示出确定在图 5 中所描述的枚举过程中所标识的网络资源的事务能力的过程的流程图；

图 7 是示出使用网络时钟域参数和绝对时间值来生成伪时钟的定时图；

图 8 是示出锁定授权机构的逻辑结构的框图；

图 9 是示出基本的无序拒绝锁定情景的示意图；

图 10 是示出在无序地接收到请求时锁定授权机构将锁定请求和事务重新定序的锁定情景的示意图；

图 11 是示出在客户机能够回卷已在进行中的事务时锁定授权机构如何能将无序地接收的请求重新定序的示意图；以及

图 12 是描述客户机的回卷请求被拒绝的锁定情景的示意图。

## 优选实施例的详述

在以下对示例性实施例的描述中，将会尽可能地指出与不同类型的系统相关的实现差异或是独一性问题。但是应当理解，本文中所描述的系统和方法适用于任何类型的网络系统。

### I. 绝对时基

全球定位系统（“GPS”）是便于为导航目的使用的巨大的卫星网络。GPS 通常被描述为以三角测量或是三维空间的概念为基础。但是，GPS 实际上在四维中进行三角测量。由此，启用 GPS 的设备使用四个坐标，即经度坐标、纬度坐标、海拔高度坐标和时间坐标来知悉其位置。因为每个 GPS 接收器都使用时间坐标来

计算其导航位置，所以 GPS 固有地创建出对地球上任何地方的所有设备都可用的绝对时间值。

有利的是，GPS 信号所实现的固有时间基准实质上解决了很难处理的网络时钟分布问题。并且，GPS 通过根本上忽略此问题的存在来解决这一问题。GPS 不是把焦点集中在时钟分布和时间同步上，而是可被用来定义一类知悉时间的设备。每个知悉时间的设备都包含其自己的 GPS 接收器，以及用来从所接收的 GPS 信号中提取时间坐标的时间解码电路。因为每个知悉时间的设备都具有得自 GPS 接收器的其自己的内部时间值，所以无需建立公共的时基。实质上，GPS 系统已经提供了所需的时间平台。由此，当知悉时间的设备通过从所接收的 GPS 信号提取时间坐标来创建其自己的“伪时钟”时，就建立了绝对时间值。绝对时间值消除了常规的对中央时钟或是时间服务器的需要。

但是知悉时间的设备仍然能够像是在它们之中生成和分布了公共时钟那样来进行交互。为了在不传送和同步时间信息的情况下进行通信，各个知悉时间的设备要遵守关于其内部地生成的“伪时钟”的使用的确立的协议。如以下将详细描述，与网络时钟域结合的知悉时间的网络被用来协调知悉时间的设备如何使用其内部地生成的“伪时钟”来与其它知悉时间的设备交互和访问共享的网络资源。

## II. 知悉时间的设备

一般而言，知悉时间的设备可以是包括 GPS 接收器的任何设备。在基于位置信息来确定时间坐标时可任选地使用逻辑。当然，这些要素不是限定性的，因此从 GPS 信号中提取时间坐标时可使用任何硬件或软件方法。各家制造商制造用于从 GPS 信号中提取精确时间坐标的设备。由此，从 GPS 信号中提取精确时间的技术对于本领域技术人员而言是公知的。因此，这些技术中的任何一种都可被用来获得时间坐标。此外，GPS 接收器中的某些部分（诸如天线和 RF 部分等）也可位于设备的外部，或甚至在位于通常与 GPS 系统相关联的误差范围内的数个设备之间共享。

就知悉时间的设备而言，重要的是要强调每个设备都是完全独立的，并且不需要与其它设备协调或同步也能建立公共时间基准。每个知悉时间的设备都能够基于所提取的 GPS 信号坐标来以已知的最差准确性和派生的当前时间准确性（以信号完整性、误差测量和设备能力为基础）来维持时间。

或者，知悉时间的设备可被构造成具有专用的总线接口，该接口使设备能够

获得具有准确且已知的延迟的 GPS 信号。精确的定时坐标可通过使用算法调整对延迟进行补偿来从通过总线接收的 GPS 信号中提取。总线可以是广播总线或是请求-响应系统。更具体地，总线提供一种将数个知悉时间的设备组合为单个设备的方法。这可通过将数个接口卡插入到公共底板中来实现。在此情形中，GPS 接收器将为总线提供连接到该总线的每块接口卡的绝对时间值。优选的是，总线的电特性和定时特性应被严格控制。此外，可将设备放置在物理上非常靠近的位置。此外，这些设备还应被并置在诸如计算机或路由器底板等已具有内部时基（诸如计算机时钟等）的设备中，以使总线能够完全同步。或者，一组设备可共享单个 GPS 接收器以降低成本和空间要求。

### III. 知悉时间的网络

因为 GPS 提供了一种获取绝对时间值的手段，所以无需在各个计算机和网络外围设备之间对时间进行同步或协调。这一绝对时基的建立允许知悉时间的设备能够共享资源而同时又能维护所共享的资源的完整性。这是通过使用知悉时间的网络来实现的。

图 1 描绘了根据本发明的基本的知悉时间的网络 101。具体而言，知悉时间的网络 101 由知悉时间的设备构成，这些设备包括但不限于如下所述的客户机 103、资源 109、知悉时间的路由器 105 以及知悉时间的锁定授权机构 111。

#### a. 客户机：

客户机 103 是知悉时间的设备，它能够与知悉时间的网络 101 中的其它知悉时间的设备通信。客户机 103 的示例性和非限定性实施例包括工作站、服务器、中央处理单元以及任何无线设备。此外，因为知悉时间的网络与所使用的网络传输的实际类型是不相关的，所以使用卫星传输的设备也可构成客户机。软件驱动程序被加载到客户机 103 上，以使客户机 103 能够与包括如图 1 中所示的知悉时间的路由器 105 在内的其它知悉时间的设备通信。客户机 103 通过如下所述的标准网络传输 107 连到知悉时间的路由器 105。

#### b. 网络传输

如图 1 中所示，客户机 103 经由网络传输 107 连到知悉时间的路由器 105。网络传输本身可以是任何类型的网络传输。例如，网络传输 107 可以是诸如光线信道

网络等既包括物理层又包括逻辑传输层的传输标准。或者，网络传输 107 可以是 TCP/IP 网络传输，它包括由各种物理传输层（包括但不限于以太网和 ATM 物理网络拓扑结构）支持的逻辑层（TCP/IP）。

简言之，任何数据传输网络可用作在知悉时间的设备之间传递的加盖了时间戳的数据的传输。此传输还包括利用卫星传输来进行连接的网络。由此，本文中所描述的方法和技术是与传输介质不相关的，由此允许包括无线、有线或是光纤在内的任何网络传输都可被使用。

### c. 知悉时间的路由器

如图所示，在网络传输 107 上发送的来自客户机 103 的分组被路由通过知悉时间的路由器 105。所发送的通过知悉时间的路由器 105 的分组由如上所述的知悉时间的客户机 103 用得自 GPS 信号的绝对时间值来加盖时间戳。与常规路由器不同，知悉时间的路由器 105 不在已连接的设备中间传达、同步或交换时间信息。相反，知悉时间的路由器 105 依靠每个知悉时间的客户机设备 103 内本地可用的绝对时间值。或者，因为每个知悉时间的路由器 105 都有 GPS 接收器，所以知悉时间的路由器 105 可使用其自己的绝对时间值来对所发送的分组盖戳。

更具体地，知悉时间的路由器 105 所使用的时间戳是与分组从客户机 103 发送的时间点相关联的。使用基于时间戳的发送时间的时间值就允许知悉时间的路由器 105 和接收器能基于发送时间而不是接收时间来确定顺序或次序。这是在共享网络资源时维护数据一致性的至关重要的要素。

知悉时间的路由器 105 接受标准网络分组，并以不影响分组在网络传输上的路由分组的方式对分组加盖时间戳。分组时间戳加盖可在物理层级或在更高的网络协议级实现。

#### 1) 网络层时间戳加盖

时间戳加盖可通过在标准的 20 字节外扩展 TCP 层报头或是 IP 层报头以容纳时间戳来实现。根据网络配置，时间戳可被插在 IP 层以免影响使用 TCP（传输控制协议）、ICMP（因特网控制消息协议）、UDP（用户数据报协议）或是其它分组协议的分组的传输。

当然，将时间戳插在 TCP 层仍然允许诸如 HTTP（超文本传输协议）、FTP（文件传输协议）、Telnet 等较高层协议的使用。但是，将时间戳插在 TCP 层就

仅允许指定的 TCP 端口参与知悉时间的网络，由此使其它端口不具有知悉时间的特征。当然，系统管理员可配置网络以根据个人需要来提供时间戳和知悉时间的特征。对于在因特网上实现的知悉时间的分组和网络，时间戳被插在 IP 层，由此允许所有 TCP 端口的参与。

## 2) 物理层时间戳加盖

在某些情形中，媒体访问控制 (MAC) 层报头可被配置为接受时间戳。具有可选的网络报头的光纤信道网络可使用加盖了时间戳的 MAC 层报头。根据网络配置，MAC 层报头可被配置成为每个帧或序列保持时间戳。

就以太网 MAC 报头而言，可对比特模式使用边带调制以提供时间戳编码。使用新硬件的知悉时间的网络可选择在边带调制中将时间戳编码，因为边带调制在不牺牲传输带宽的情况下每分组提供一个时间戳。

### d. 锁定授权机构:

锁定通过在指定的一段时间里给予一个客户机对一资源排它的访问来确保事务与该共享资源的同步和一致性。最优地，给予单个客户机的排它权持续很短的一段时间，由此来给每个客户机提供对该共享资源具有排它的和连续的访问的表象。锁定功能为每个客户机提供了在其可访问共享资源期间，该资源的状态处于该客户机的排它控制之下的保证。

如图 1 中所示，知悉时间的网络 101 还包括锁定授权机构 111。锁定授权机构 111 被配置成串行化和同步客户机 103 对共享资源 109 的访问。共享资源 109 可以是可由一个或多个客户机访问的任何组件或设备。例如，图 1 描绘了多个共享资源 109，包括，但不限于，诸如磁盘阵列和数据库等大容量存储机制、大容量存储器系统、以及流媒体（视频和声音）系统等。重要的是要认识到锁定授权机构 111 也是知悉时间的装置。锁定授权机构及其相关的锁定情景（在维持数据完整性的同时允许访问共享资源）将在以下更加详细地讨论。

概述之，图 1 中所示的知悉时间的网络 101 使用与每个网络元素（即，客户机 103、路由器 105、锁定授权机构 111 以及共享资源 109）相关联的绝对时间值来创建与计算机时钟等效的时标，由此消除了在网络单元之间同步时间信息的必要。由此，当客户机 103 通过网络传输 107 来发送分组时，到达知悉时间的路由器 105 的是具有显示出发送时间的时间戳的分组。知悉时间的路由器在网络 101 上将

分组转发给锁定授权机构 111。锁定授权机构 111 使用基于发送时间的时戳来区分所接收的分组的优先次序。一建立起准确的次序，锁定授权机构 111 就准许单个客户机 103 对共享资源 109 的访问。

不使用网络时钟并且不在诸设备之间交换时间信息的同步是通过创建网络时钟域来实现的。如以下将详细描述，网络时钟域是使用定义每个知悉时间的设备如何为联络和访问网络时钟域中的资源而使用其绝对时间值的一组参数来建立的。

#### IV. 网络时钟域

在知悉时间的网络内，锁定授权机构负责创建网络时钟域和初始化网络时钟域。此外，锁定授权机构使用网络时钟域来控制对网络上的共享资源的访问。有利的是，锁定授权机构可创建和控制多个网络域。

常规的锁定授权机构只是基于所要接收的第一个分组来简单地管制对共享资源的访问，由此缺乏确定哪个分组首先被发送的能力。这导致客户机请求被无序地处理。相反，本文中所描述的锁定授权机构被配置成检查从知悉时间的网络上的各个知悉时间的客户机接收的分组时戳。该锁定授权机构基于分组的发送时间戳以及锁定授权机构中缓冲存储器的可用性来重排分组。具有最早的发送时间戳的分组由此被传递到共享资源。本锁定授权机构还标记无序接收的分组。

在锁定授权机构提供客户机与共享资源之间的锁定服务之前，锁定授权机构创建网络时钟域。应当注意，锁定授权机构可创建和支持多个网络时钟域以处理具有不同能力的客户机。例如，一些客户机可能无法达到高性能网络时钟域所要求的时间精确性。然而同一个客户机可能能够参与具有最小准确性和等待时间要求的网络时钟域。此外，客户机可使用不同的方法来响应并重排锁定授权机构无序接收的分组。因此，将具有相似能力的所有客户机分组到匹配那些客户机能力的网络时钟域是合乎需要的。简言之，网络时钟域是基于共享资源的定时要求的定时协议，其中所有设备（包括但不限于，路由器，锁定授权机构和客户机）都必须能够满足由定时协议阐述的定时要求才能与彼此和与共享资源交互。

参考图 2 可最好地描述知悉时间的网络上的知悉时间的设备（包括锁定授权机构、共享资源、客户机和网络时钟域）之间的相互关系。图 2 描绘了使用三个单独的网络时钟域 210、220 和 230 的知悉时间的网络 200。每个网络时钟域 210、220、230 提供对包括磁盘阵列 214、234、网络处理器 224 和塔式盒（tower box）236 的共享网络资源的经同步的访问。知悉时间的网络 200 还对每个网络时钟域包括至少

一个锁定授权机构。如图 2 中所示，锁定授权机构 212 创建和维护网络时钟域 210。类似地，网络时钟域 220 和 230 各自由锁定授权机构 222/232 维护。重要的是要注意，锁定授权机构可与知悉时间的路由器一体化，如网络时钟域 210、220 中所示。或者，锁定授权机构可存在于知悉时间的路由器的外部，如网络时钟域 230 中所示。

如图 2 中进一步示出的，每个锁定授权机构与至少一个共享资源相关联。例如，网络时钟域 210 中的锁定授权机构 212 控制对磁盘阵列 214 的访问。另一方面，维护网络时钟域 230 的锁定授权机构 232 控制对两个共享资源（即，塔式盒 236 和磁盘阵列 234）的访问。

每个锁定授权机构 212、222、232 负责维护至少一个已建立的网络域 210、220、230。客户机加入网络时钟域以访问由锁定授权机构控制的共享资源（例如，磁盘阵列 214、234，网络处理器 224 以及塔式盒 236）。锁定授权机构可建立一个以上的网络时钟域，但图 2 中为简单起见每个锁定授权机构仅示出一个网络时钟域。

如图 2 中进一步描绘的，网络时钟域 210 中当前有两个客户机 215、216 参与该域。类似地，网络时钟域 220 也具有两个客户机 216、226 参与该域。显然，单个客户机可参与多个网络时钟域，如客户机 216 所示，它当前正在使用两个网络时钟域 210、220。

### **a. 初始化网络时钟域**

简言之，在网络时钟域的初始化期间，锁定授权机构必须确定有多少个网络资源被连到该锁定授权机构、所连的资源类型、以及这些网络资源在事务锁定方面的能力。图 3 是描绘了由锁定授权机构实现的初始化网络时钟域的动态过程。

该过程以锁定授权机构本身的初始化为开始。在步骤 301，锁定授权机构接受由网络管理员提供的配置参数。这些参数包括，但不限于，所支持的时钟域的数目，所控制的资源的类型、以及每个所支持的网络时钟域的特性和能力。

一旦锁定授权机构初始化 301 完成，锁定授权机构就循环遍历并查询其端口以标识已连接的网络资源，如步骤 303 中所示。在步骤 305，确定是否发现了已连接的网络资源。如将参考图 5 更加详细地讨论的，步骤 307 表示一旦成功标识网络资源，该网络资源即被枚举和验证。另一方面，如果没有发现已连接的网络资源，则该过程返回步骤 303，由此允许锁定授权机构继续查询端口以寻找可能的已连接网络资源。一旦完成步骤 307，锁定授权机构就被置为空闲状态以等待客户机查询，如步骤 309 中所示。

图 5 是示出在锁定授权机构标识出已连接的网络资源所执行的枚举和验证过程（见图 3 中的步骤 307）的流程图。在步骤 503，锁定授权机构向所标识的已连接到端口的网络资源发送地址请求。此请求的形式取决于底层网络的类型。例如，在 TCP/IP 网络中，锁定授权机构可向设备发送 ARP（地址解析协议）消息。然后就所请求的地址的接收作出确定，如步骤 505 中所示。如果在锁定授权机构接收到所请求的地址之前已经过去了预定量的时间，则请求超时，如步骤 506 中所示。锁定授权机构随即退出枚举和验证过程，并前进至下一个端口，如步骤 507 中所示。但是，如果从已连接的网络资源接收到所请求的地址，则锁定授权机构使用接收到的地址来向该网络资源发送附加信息请求，如步骤 509 中所示。

继续图 5，在步骤 511 已连接的网络资源接收附加信息请求。在步骤 513，就已连接的网络资源是否是共享资源（相对于客户机、服务器或专用资源而言）作出确定。如果已连接的网络资源是共享网络资源，则该过程前进至步骤 515 以确定该网络资源类型是否可由锁定授权机构控制。如果已连接的设备是共享资源，并且如果已连接的设备是可由锁定授权机构控制的资源，则该过程前进至步骤 517。如果就步骤 513 和 515 作出了否定的确定，则该过程退出枚举和验证过程并前进至下一个端口，如步骤 507 中所示。

回到步骤 517，一旦确定已连接的网络资源是共享资源并且可由锁定授权机构控制，锁定授权机构即评估已连接网络资源的事务能力。已连接网络资源的能力的评估将结合图 6 来讨论。使用通过枚举和验证过程并通过网络资源能力确定过程而获得的信息，锁定授权机构为该网络资源建立起网络时钟域参数块，如步骤 519 所示。时钟参数块的重要性和用途将在以下更加详细地讨论。

现在转到图 6，提供了示出确定网络资源的事务能力的过程的流程图。在步骤 601，锁定授权机构发送一组测试事务以确定资源的事务响应时间能力。测试事务可根据网络资源的类型而定。例如，对于存储网络资源，测试可以是查询数据命令、读取容量命令，接着是测试单元就绪命令。接下来可以是一组读和写命令。一旦测试完成，在步骤 603，锁定授权机构即确定该网络设备响应这些测试事务或命令的平均、最小和最大响应时间。用来测试响应时间的命令数目可由系统管理员配置，但通常是在 100 个事务的范围内。这提供了统计平均值而又不会过度延迟初始化过程。如在步骤 605 中所示，网络资源的事务响应时间能力被存储在存储器中。

一旦确定了资源的事务响应时间能力，锁定授权机构就向该网络资源发送回卷测试事务，如在步骤 607 中所示。回卷测试事务确定网络资源是否能够对事务执

行回卷。一旦网络资源完成回卷测试事务，锁定授权机构就向网络资源发送事务回卷命令，如在步骤 609 中所示。如果在步骤 611 网络资源拒绝回卷命令，则可确定该网络资源不能执行资源方的事务回卷（见步骤 613）。一旦如步骤 613 中所示，确定网络资源不能够执行资源回卷，则在步骤 617 锁定授权机构即退出资源事务能力处理，由此来强制锁定授权机构对所有事务都使用客户机方回卷。另一方面，如果网络资源接受回卷命令，并正确地回卷测试事务，如在步骤 615 中所示，则锁定授权机构将使用资源方回卷处理来重排客户机锁定请求。无论资源事务能力的确定是肯定的还是否定的，该过程总是退出（如步骤 617 所示），并返回步骤 519（在此如参考图 5 所描绘和描述地为资源建立起网络时钟参数块）。

### b. 时钟参数块和网络时钟域

时钟参数块在创建和加入网络时钟域时是有用的，因为时钟参数块定义了客户机为参与网络时钟域所必须满足的要求。由此，这些参数实质上定义了网络资源的要求，这些要求通过使用绝对时间值来使客户机和网络资源能够通信。在时钟参数块中阐述的参数的示例包括，但不限于，最小时钟准确性、最小时钟精确性、时钟值（通常以纳秒计）、时钟相位数以及时钟时间开始值。

时钟参数块中的参数还定义客户机将如何生成在网络通信定时中所使用的内部“伪时钟”。建立网络时钟域的锁定授权机构生成完全相同的伪时钟，从而客户机与锁定授权机构可执行时间已同步的操作，而无需提供主时钟源或者交换实际时钟信息或信号。

图 7 示出使用网络时钟域参数和绝对时间值来生成伪时钟的方法。具体而言，网络时钟域参数块 701 定义时钟参数块中所提供的所有时间的时钟单位（例如，纳秒、皮秒）。换言之，时钟参数块中的第一个参数定义了该时钟参数块中其它参数所使用的时间单位。时钟参数块包含客户机必须使用其内部 GPS 时间恢复电路来保证的绝对时间值的最小准确性和精确性值。如果客户机不能保证这些值，则它可能不可加入该网络时钟域。

客户机以等于或大于时钟参数块 701 中所阐述的最小绝对时间更新速率的速率来更新存储在 GPS 绝对时间寄存器 703 中的时间值。每当该值被更新时，要将其与时钟报时信号模数和相位 0 时钟报时信号模数相比较（如果时钟具有不止一个相位）。如果模数值小于时钟参数块 701 中的绝对时间精确性，则生成时钟报时信号 705。报时信号 705 被用来定义伪时钟的下一个边沿。由此，一系列的报时信号

705 生成满足所需准确性和精确性规范的一个或多个相位的伪时钟 707。所生成的伪时钟 707 被监视以确保所生成的伪时钟 707 的抖动小于时钟参数块 701 中所指定的最大抖动。监视所生成的伪时钟 707 的抖动确保伪时钟继续提供所需的准确性和精确性以便于利用锁定授权机构。

### c. 加入网络时钟域

客户机可动态地或是通过使用先验知识来加入网络时钟域。客户机通过向锁定授权机构发送请求以找出加入该网络时钟域需要什么能力和配置来加入网络时钟域。换言之，客户机向锁定授权机构请求时钟参数块。时钟参数块从锁定授权机构被发送到客户机。

如果客户机能够满足在时钟参数块中所阐述的要求，则客户机向锁定授权机构发送确认消息。一旦发送了确认消息，即视该客户机为已加入了该网络时钟域。一旦客户机加入了网络时钟域，客户机即可获得对由锁定授权机构控制的网络资源的锁和排它性访问。

更具体地，图 4 和 4a 是示出客户机加入在图 3 中初始化的网络时钟域的过程的流程图。具体而言，将客户机添加到网络时钟域是由客户机发起的。客户机通过向等待着的锁定授权机构发送网络资源查询消息来启动加入网络时钟域的过程，如步骤 401 中所示。在步骤 403，就是否已接收到该客户机查询作出确定。如果没有接收到该请求，则该过程将锁定授权机构返回到其由步骤 401 表示的空闲状态。另一方面，如果成功地接收了客户机的查询，则在步骤 405 锁定授权机构确定其是否控制如在步骤 403 中所请求的类型的资源。如果锁定授权机构未控制所请求的资源，则在步骤 406 向请求客户机发送拒绝消息。或者，如果锁定授权机构未控制所请求的资源，则该过程前进至步骤 407，在此锁定授权机构用查询响应消息来响应客户机的消息。查询响应消息包括要求客户机发送描述客户机所能够支持的网络时钟域参数的时钟参数块的请求。

在步骤 409，锁定授权机构接收对查询响应消息的响应，该响应描述客户机能力。如果锁定授权机构没有从客户机接收到响应，则锁定授权机构保持等待状态，如步骤 408 中所示。在步骤 411，锁定授权机构将客户机的能力与网络时钟域的要求相比较。如果客户机的能力在网络时钟域所要求可接受的范围之内，则客户机即被允许加入网络时钟域，如步骤 413 中所示。一旦被接纳加入网络时钟域，在步骤 415，确认消息被发送到客户机，由此允许客户机请求对也参与该网络时钟域的

共享资源的访问。另一方面，如果客户机的能力不匹配网络时钟域所阐述的要求，则客户机加入该域的请求被拒绝，如在步骤 417 中所示。

一旦加入网络时钟域，客户机即可请求对共享资源的访问并参与下述锁定方案。此时，锁定授权机构继续等待其它客户机请求访问受控的资源，其中重复参考图 4 所描述的过程。

网络域的一般原理适用于所有的客户机。例如，客户机加入或离开网络时钟域的时间不影响该网络时钟域的工作。此外，客户机离开域不需要任何条件。如果多个客户机想要加入该域，则锁定授权机构将按如分组发送时间戳所确定的次序来排列所接收的分组请求。重要的是，知悉时间的设备、或客户机可参与多个网络时钟域，但是，共享资源仅可参与单个网络时钟域。

如以上简要提及的，客户机还可通过先验知识变为网络时钟域的成员。如果客户机是通过先验知识而变成的成员，则加入该域所需的知识只需由系统或网络管理员在配置锁定授权机构时提供。

## V. 锁定授权机构

一旦客户机成功地加入了网络时钟域，该客户机即可开始经由锁定授权机构来请求对共享资源的访问。锁定授权机构是确保向每个客户机提供串行化的和经同步的资源访问的重要手段。锁定功能还通过确保分组请求是按根据与该分组相关联的发送时间戳的正确顺序处理来维持共享资源的数据的一致性。

### a. 锁定授权机构描述

锁定授权机构可用许多不同的物理形式来实现。例如，锁定授权机构可被结合到诸如存储设备等网络资源中。它还可被结合到知悉时间的路由器中，或可在知悉时间的网络上作为单独和独立的物理单元而存在。

图 8 示出锁定授权机构的逻辑结构。如图所示，锁定授权机构包括四个主要逻辑块：网络时钟域逻辑 801；客户机接口逻辑 803；资源接口逻辑 805；以及事务控制逻辑 807。

#### 1. 网络时钟域逻辑

网络时钟域逻辑 801 执行两个主要功能，即，域参数控制功能 809 和域成员资格控制功能 811。第一个功能，即域参数控制功能 809 建立适合资源和网络能力

的网络时钟域参数。此外，域参数控制功能 809 可被配置成建立多个网络时钟域。建立多个网络时钟域的主要原因是允许具有不同能力的客户机访问共享资源。例如，一些客户机可能在定时方面能够实现较高的准确性或精确性，而其它可能使用轮询方法来获得可作出锁定请求的时隙。由此，除了客户机试图加入该网络时钟域所需的能力以外，域参数控制功能 809 还确定网络时钟域所支持的锁定方法的类型。

网络时钟域逻辑 801 中的第二个功能是域成员资格控制功能 811。此功能响应对网络时钟域参数的客户机请求，并从客户机接收加入网络时钟域的请求。该功能包含用于确保试图加入该域的任何客户机都具有参与所需的能力的逻辑。

## 2. 客户机接口逻辑

客户机接口逻辑 803 起到知悉时间的网络的接口的作用。客户机接口逻辑 803 包含在分组从网络上的各个客户机到来时检查和比较时间戳所需的逻辑。客户机接口逻辑还从每个客户机获得该客户机的事务回卷能力。

## 3. 资源接口逻辑

资源接口逻辑 805 提供到正被控制的资源的专用接口。资源接口是正被访问的类型的资源所特有的。此外，资源接口逻辑还确定共享资源的事务排队和回卷能力。

## 4. 事务控制逻辑

事务控制逻辑 807 包括三个主要功能：锁定控制功能 813；事务控制功能 815；以及事务回卷功能 817。这三个功能一起工作就可实现在图 9-12 中示出并在以下详细讨论的各种锁定情景梯级图的逻辑。

锁定控制功能 813 从时钟域成员资格控制功能 811 获得参与该网络时钟域的客户机的列表。锁定控制功能 813 还从客户机接口功能 803 获得锁定请求分组的到达次序。使用此信息，锁定控制功能即可根据资源的当前状态以及客户机的能力来准许、拒绝或将客户机的锁定请求排队。

事务控制功能 815 将来自客户机的事务请求和响应连到共享资源。此外，事务控制功能 815 建立从共享资源回到当前对该共享资源持有锁的客户机的数据通路。根据共享资源的能力和锁定授权机构的具体实现，事务控制功能 815 还为锁的

持续时间维护事务历史 819。如果客户机的资源能够实现回卷，则此事务历史 815 可被事务回卷控制器 817 用来回卷事务。事务历史 819 还可被提供给锁定被拒绝或延迟的其它客户机，以使该客户机能够保持数据和事务状态一致性。

### c. 锁定情景

重建和保持从知悉时间的设备发来的传入分组的发送次序的能力对于在无级别网络中保持数据一致性是关键的。除了控制和状态消息以外，对分组排序还为客户机和服务器两者提供了关于两个数据序列的次序的适当信息。这允许诸如锁定授权机构等从多个源接收多个分组的接收器能基于发送的次序来将分组分类，并据此对所接收的命令采取行动。这一将分组排序、定序和分类的概念是使用如以下将详细描述在许多锁定情景来实现的。重要的是要注意，锁定情景可能涉及多个客户机，但是，为简单起见，下述示例性实施例被约束在两个客户机。

下述锁定情景涉及两个客户机、锁定授权机构和共享资源，所有这些都是由锁定授权机构所建立的网络时间域的成员。使用时间梯级图来说明事务的顺序和依存关系，以及左边的客户机、中间的锁定授权机构和右边的网络资源之间的通信。

每个锁定情景描述了客户机和锁定授权机构用来将锁准许给一客户机、以及在客户机请求被无序接收时对来自另一客户机的对同一个锁的请求进行重排的一种不同的方法。在这些示例性锁定情景中，锁请求用于对共享资源的排它性访问。此类请求的一个示例可以是读/写锁。一些锁请求无需是排它性的，诸如只读锁。可同时将多个只读锁准许给诸客户机。因此，当客户机请求诸如读/写锁等排它性锁时就产生了锁定问题。如果读/写锁请求是在读锁被准许以后发送的，则读/写锁可作为待决请求来保持，直至读锁被释放。此外，如在梯级图中所描绘的术语响应是指客户机所请求的数据或事务信息。一般而言，一旦响应完成，网络资源上的锁即被自动释放。

#### 1. 基本的无序拒绝

如图 8 中所示，事务控制功能连接来自客户机的事务请求和响应。但是，图 9 是示出基本的无序拒绝锁定情景的示意图。更具体地，图 9 示出了锁定授权机构如何使用对来自客户机的锁定请求的拒绝来将客户机的请求重新排队并确保维持了数据一致性。

如图所示，客户机 1 首先向锁定授权机构发出对共享资源的锁请求 901。接着，

客户机 2 也发出对同一共享资源的锁请求 903。

通常，客户机 1 将会接收到锁并执行其事务。来自客户机 2 的请求 903 将由锁定授权机构排队，并在客户机 1 所持有的锁被释放后立即被准许。根据请求是阻塞还是非阻塞请求，客户机 2 在等待准许锁之际可挂起也可不挂起操作。

但是，在此情形中，锁定授权机构在接收到由客户机 1 发送的锁定请求 901 之前就接收到来自客户机 2 的请求 903。客户机 1 的请求 901 的延迟可能是由诸如网络传输延迟等若干因素引起的。简言之，锁定授权机构无序地接收请求。在此时，锁定授权机构不知道客户机 1 已经发出了前一锁请求，因此锁定授权机构将共享网络资源上的锁 905 准许给客户机 2。

当来自客户机 1 的锁请求 901 到达时，锁定授权机构从请求分组中的绝对时间戳确定请求 901 首先被发出并无序地到达。锁定授权机构不能确定是否要准许该请求，因为它没有关于客户机 1 和 2 的与其请求相关联的事务的信息。

事务请求 903 由此被完成，如从网络资源发送到锁定授权机构的响应 906 以及从锁定授权机构 907 发送到客户机 2 的响应 908 所证明的。锁定授权机构通过拒绝来自客户机 1 的请求 901 来处理此歧义性。由此客户机 1 被通知锁定授权机构没有准许请求 907。

在此特定配置中，没有向客户机 1 提供请求被无序接收这一事实以外的关于拒绝的原因的信息。客户机 1 由此将具有新的时间戳的新请求 909 重新排队，并向锁定授权机构发出该新请求 909。因为该请求现在被定序在来自客户机 2 的请求 903 之后，所以锁定授权机构保持该请求 909 为待决，直至客户机 2 完成其事务并释放锁（见响应消息 906）。在此时，来自客户机 1 的锁请求 909 被准许 913，由此可保持资源处的数据的一致性以及由客户机 1 和客户机 2 两者执行的事务的一致性。一旦完成其事务，客户机 1 即释放共享资源 915 上的锁。此外，响应 917 从锁定授权机构发送到客户机 1 以确认作为完成事务的结果的锁的释放。

## 2. 具有资源回卷的基本重排

图 10 是示出当请求被无序接收时锁定授权机构如何重排锁定请求和事务的示意图。在此给定的锁定情景中，被锁定的共享资源能够回卷事务以保持数据一致性。但是，回卷能力在从事务完成起已过去多少时间方面，或者在事务的大小和内容方面是偶然的。

在一些资源回卷和重排的情形中，共享资源可拒绝回卷请求。在此类情况下，

锁定授权机构将恢复到使用如参考图 9 所描述的基本的无序锁请求的拒绝。

现在参考图 10, 客户机 1 发出锁请求 1001。接着, 客户机 2 也发出锁请求 1003。因为锁定授权机构首先接收到请求 1003, 所以锁定授权机构准许 1005 客户机 2 的锁请求。但是, 由于网络延迟和等待时间, 来自客户机 1 的请求 1001 是在来自客户机 2 的请求 1003 之后接收到的。锁定授权机构使用客户机 1 的锁请求 1001 中的基于发送的时间戳来确定请求 1001 是无序地接收的。

由此, 锁定授权机构挂起对来自客户机 2 的命令和事务的处理。它然后向共享资源发送回卷请求 1007。共享资源将回卷从锁 1005 被准许起与客户机 2 相关联的所有事务和命令。一旦锁定授权机构回卷了事务 1007, 客户机 2 在网络资源上的锁被释放 1010。如果由于时间量或是事务特性使得回卷不可能, 则该资源将拒绝回卷请求 (未示出)。

假定回卷请求 1007 是成功的, 锁定授权机构将把客户机 2 的锁请求以及与客户机 2 相关联的所有事务重新排队。锁定授权机构随后将锁 1011 准许给客户机 1, 并允许客户机 1 继续进行其命令和事务。一旦事务完成, 该网络资源上的锁即被释放 1013, 并且锁定授权机构向客户机 1 通知已完成的事务 1014。但是, 该资源将被再次锁定 1015 到客户机 2, 并且客户机 2 的事务将继续进行。一旦客户机 2 的事务完成, 该网络资源上的锁即被释放 1016, 并且锁定授权机构向客户机 2 通知 1017 已完成的事务。

此过程的优点是回卷和重新准许对客户机 2 是透明的。客户机 2 仅仅知道了比通常预期的要长的一段时间来满足客户机 2 的事务。但是, 不需要客户机 2 的任何其它动作或响应就能以保持数据一致性和完整性的方式来服务其事务。

### 3. 使用客户机回卷的重排

图 11 示出在客户机能够回卷已在进行中的事务的情况下锁定授权机构如何重排无序地接收的请求。在此特定锁定情景中, 客户机 1 向控制网络资源的锁定授权机构发送锁请求 1101。接着, 客户机 2 发送对同一资源的锁请求 1103。由于网络延迟和等待时间, 锁定授权机构首先接收到来自客户机 2 的锁请求 1103, 并准许了锁请求 1105。当来自客户机 1 的锁请求 1101 到达时, 锁定授权机构通过检查传输来自客户机 1 的锁请求 1101 的分组中所包含的基于发送的时间戳来确定请求 1101 是被无序地接收的。此外, 客户机 2 的请求已被完成, 如从网络资源到锁定授权机构然后到客户机 2 的响应 1104 所证明的。

因为在接收到客户机 1 的请求 1101 之前客户机 2 已经接收到响应 1104, 所以可执行客户机方的回卷。在配置网络时钟域时, 客户机 2 指示其能够回卷进行中的事务。由此, 锁定授权机构挂起客户机 2 的事务, 并向客户机 2 发送回卷请求 1106。客户机 2 接收请求 1107, 并且锁定授权机构回卷事务 1108。网络资源随即通过向锁定授权机构发送消息 1109 来确认回卷事务。

锁定授权机构随后准许客户机 1 的锁请求, 并继续进行客户机 1 事务 1113。一旦完成, 锁即被释放 1114, 并且锁定授权机构通知客户机 1 该事务完成 1115。锁定授权机构随即将锁 1116 准许给客户机 2, 并继续执行客户机 2 的事务。类似地, 锁最终被释放 1117, 并且锁定授权机构通知客户机 2 该事务完成 1118。

使用客户机回卷的优点在于, 客户机可保留对回卷事务的方式的完全控制。此外, 客户机可基于回卷以及来自客户机 1 的其它活动来决定修改或删除事务。缺点是需要更多的网络话务, 并且要求客户机具有更加复杂的软件和事务维护。

#### 4. 客户机回卷请求的拒绝

图 12 是描绘客户机的回卷请求被拒绝的锁定情景的示意图。更具体地, 图 12 描述在客户机拒绝来自锁定授权机构的回卷请求时所经历的一系列事件。

如前, 客户机 1 发送锁请求 1201, 接着客户机 2 发送锁请求 1203。来自客户机 2 的锁请求 1203 首先被接收到, 并且锁定授权机构准许了锁 1205。当来自客户机 1 的请求 1201 到达时, 锁定授权机构通过检查传输来自客户机 1 的锁请求的分组中的发送时间戳而确定它是被无序地接收的。但是, 如响应消息 1204 所示, 在锁定授权机构接收到来自客户机 1 的请求 1201 以前, 响应已被发送给客户机 2。因此, 锁定授权机构必须试图回卷与客户机 2 有关的事务, 实质上就是要求客户机 2 忽略从锁定授权机构发来的响应消息 1206。

在此情形中, 客户机 2 先前已指示了它能够回卷事务以保持次序和一致性。因此, 锁定授权机构挂起客户机 2 的事务并向客户机 2 发送回卷请求 1206。但是, 在此情形中, 客户机 2 因为某种原因而无法回卷该事务。可能是因为已经过去了太长时间或者是客户机 2 已经将该事务的结果用于其它处理并且不能保证回卷事务与数据的一致性。客户机 2 因此向锁定授权机构发送回卷拒绝 1209。锁定授权机构随即继续进行客户机 2 的事务, 并向客户机 1 发送锁拒绝 1213, 由此来强制客户机 1 将其请求重新排队 1215。

一旦客户机 2 的事务完成并且锁被释放 1214, 客户机 1 即将具有较晚的时间

戳的新的锁请求 1215 发送到锁定授权机构。因为此请求 1215 不是无序的，所以一旦客户机 2 的事务完成，它即被准许 1216。一旦完成，锁即被释放 1217，并且锁定授权机构通知客户机 1 事务完成 1218。

#### ***d. 处理锁定出错***

知悉时间的网络及其相关联的网络时钟域提供对在网络上发送的各种锁定出错的健壮且简单的处理。不幸的是，出错状况是由于许多情况而发生的。例如，出错状况可能因为分组在网络上丢失或是过度延迟而发生。最坏的情形是分组的丢失，因为延迟的分组最终将会到达，并允许拆解可能已经产生的死锁状况。但是，当分组被丢失时，除了通过很长的超时时段以外，发送设备没有任何办法来验证丢失，因此也没有任何办法来知道分组没有被接收到。

当请求或持有锁的客户机崩溃或是经历某种类型的灾难性的操作出错时，就会发生出错状况情况的另一种示例。锁定授权机构和被锁定的资源没有任何方法知道客户机不再正常工作。如果此时客户机正持有或请求锁，则该锁必须被释放，并且来自该客户机的任何其它或待决的请求必须被拒绝，否则就将产生死锁情况。类似地，资源或锁定授权机构可能会经历某种灾难性的操作出错。有待决锁的客户机没有任何方法可得到该情况的通知，由此导致死锁。

客户机与锁定授权机构之间的时间的正确管理是对上述情况的解决方案。所有锁请求和锁准许都是使用指定的超时值来执行的。因为超时值是基于绝对时间（使用网络时钟域的参数来诠释绝对时间值），所以在给定网络的固有传输延迟和等待时间的前提下，超时可以尽可能地短。不存在锁被请求或准许，然后被抛弃，使得系统不能检测和纠正抛弃并恢复正常操作这样的情形。

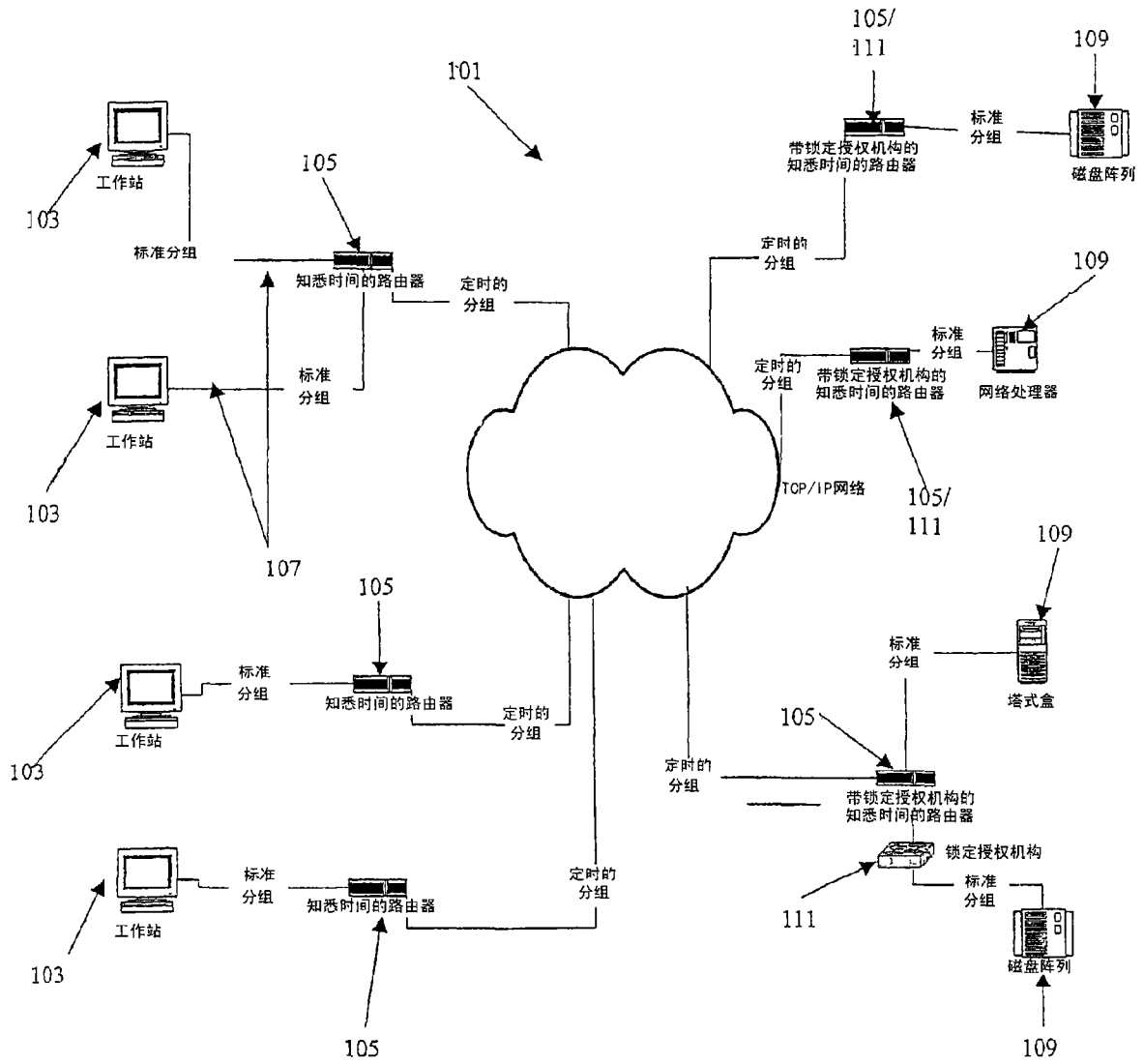


图 1

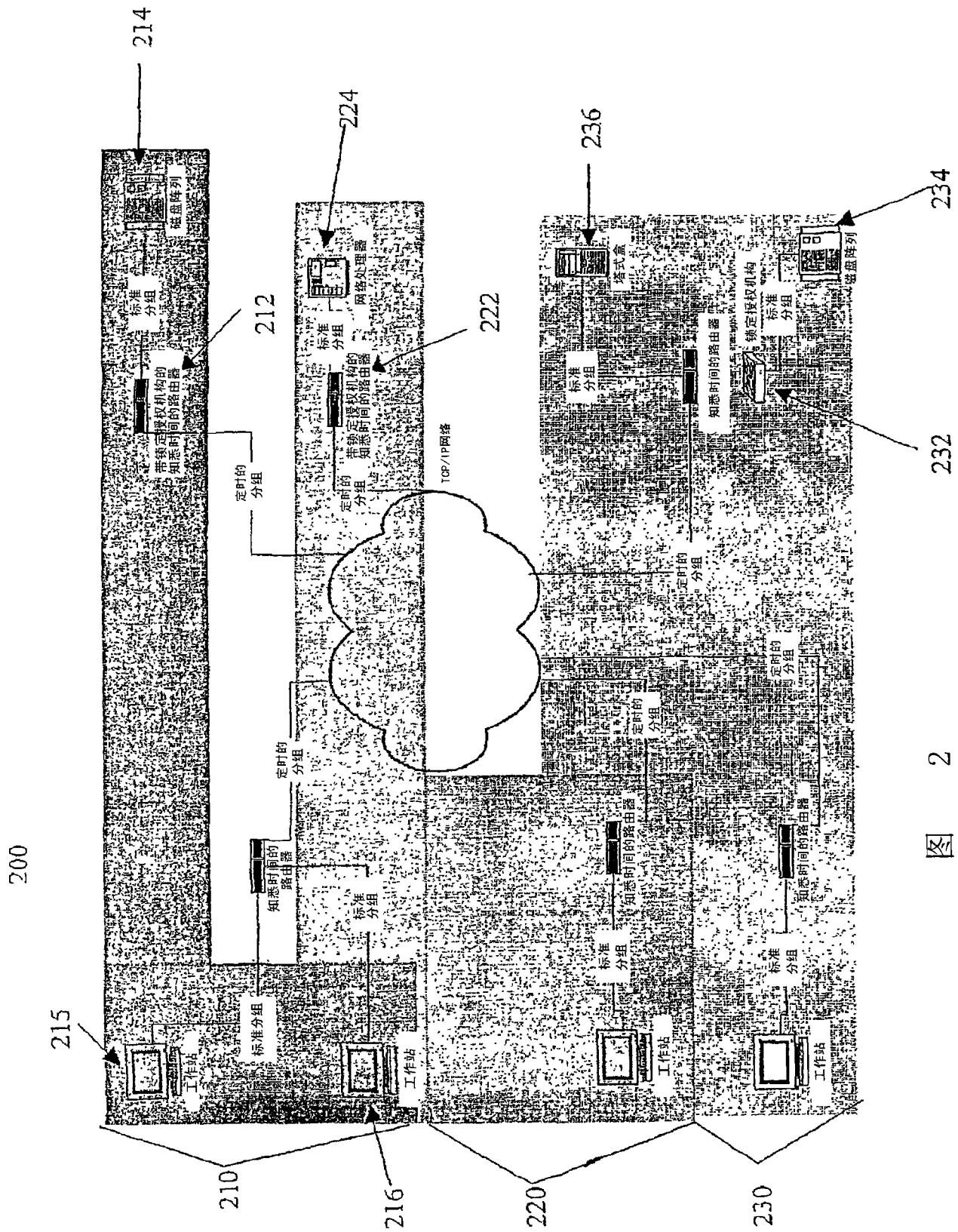


图 2

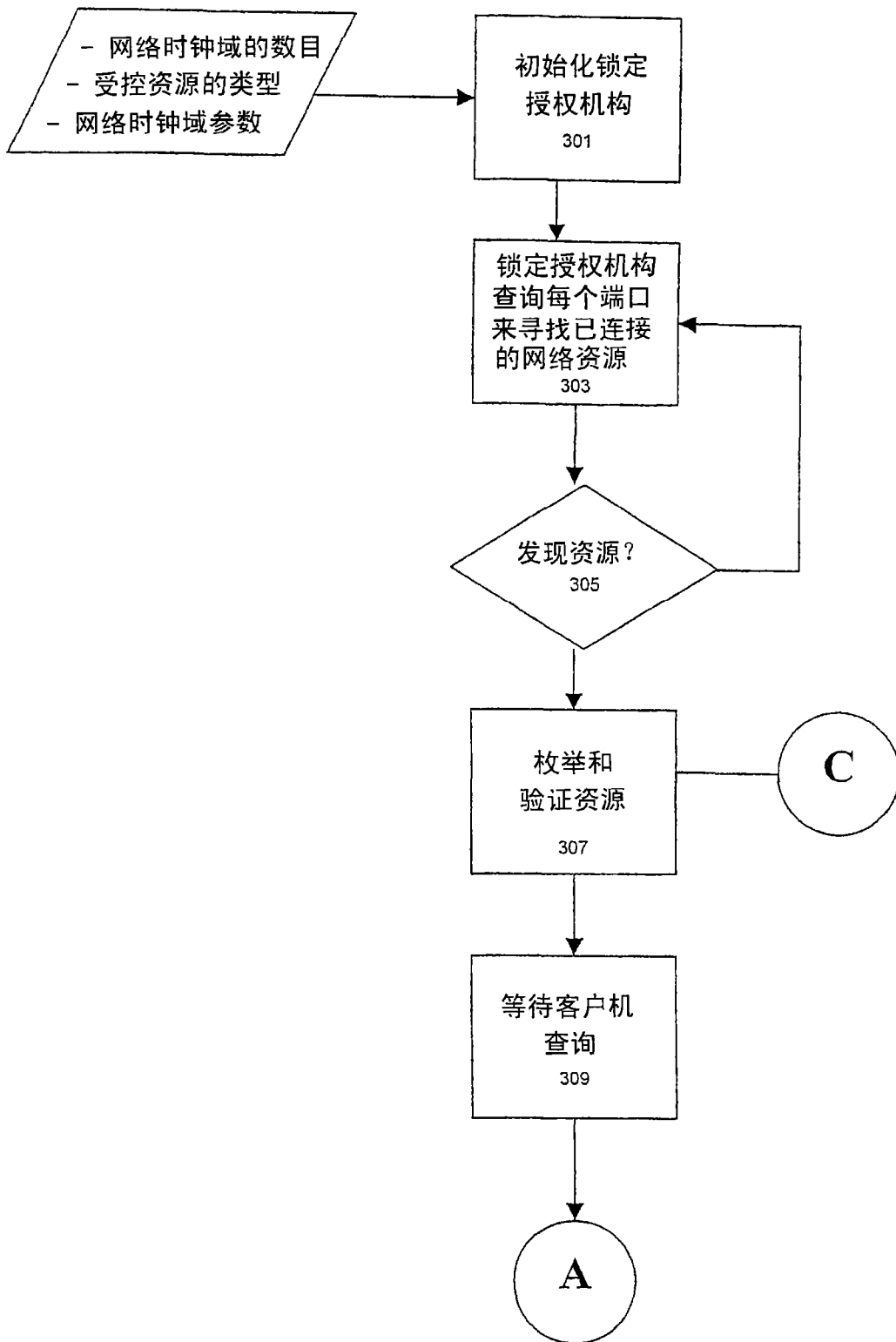


图 3

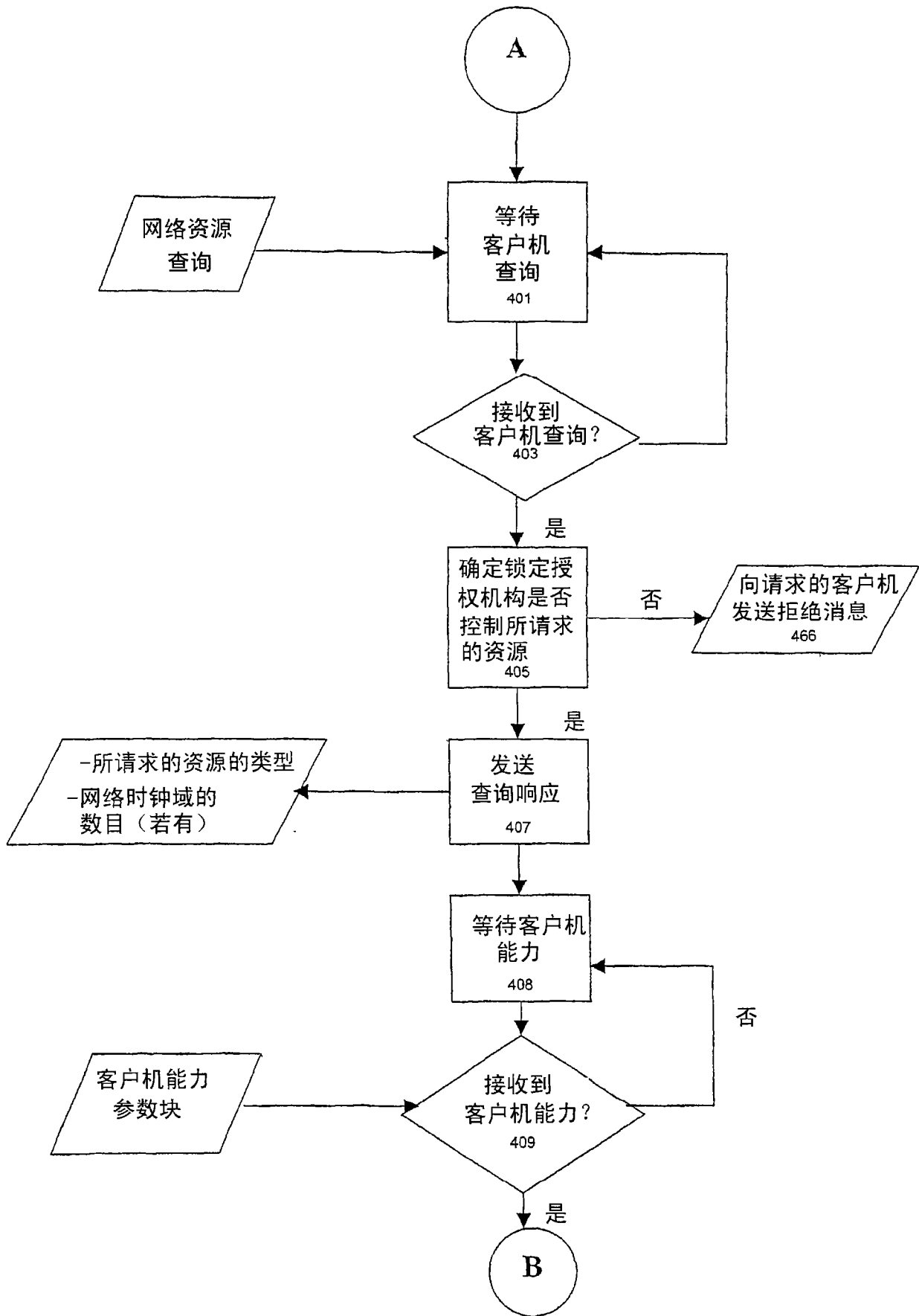


图 4

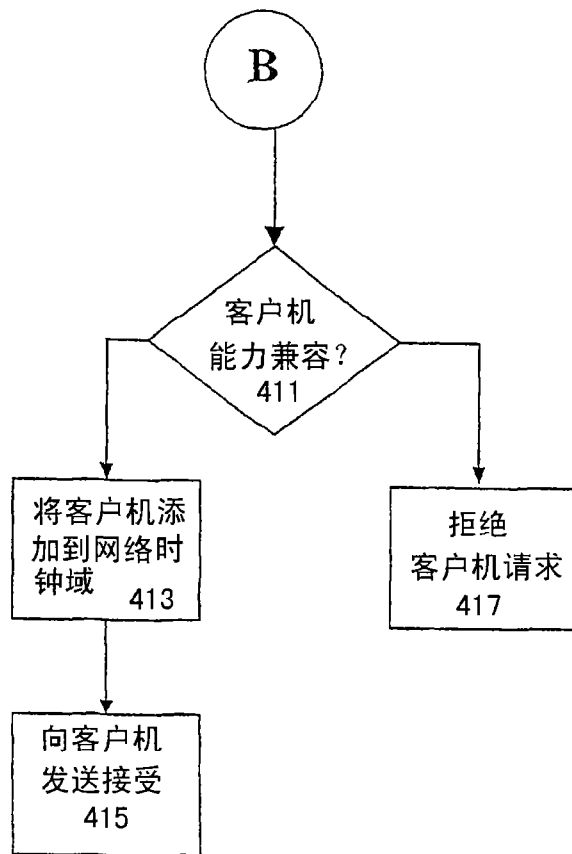


图 4A

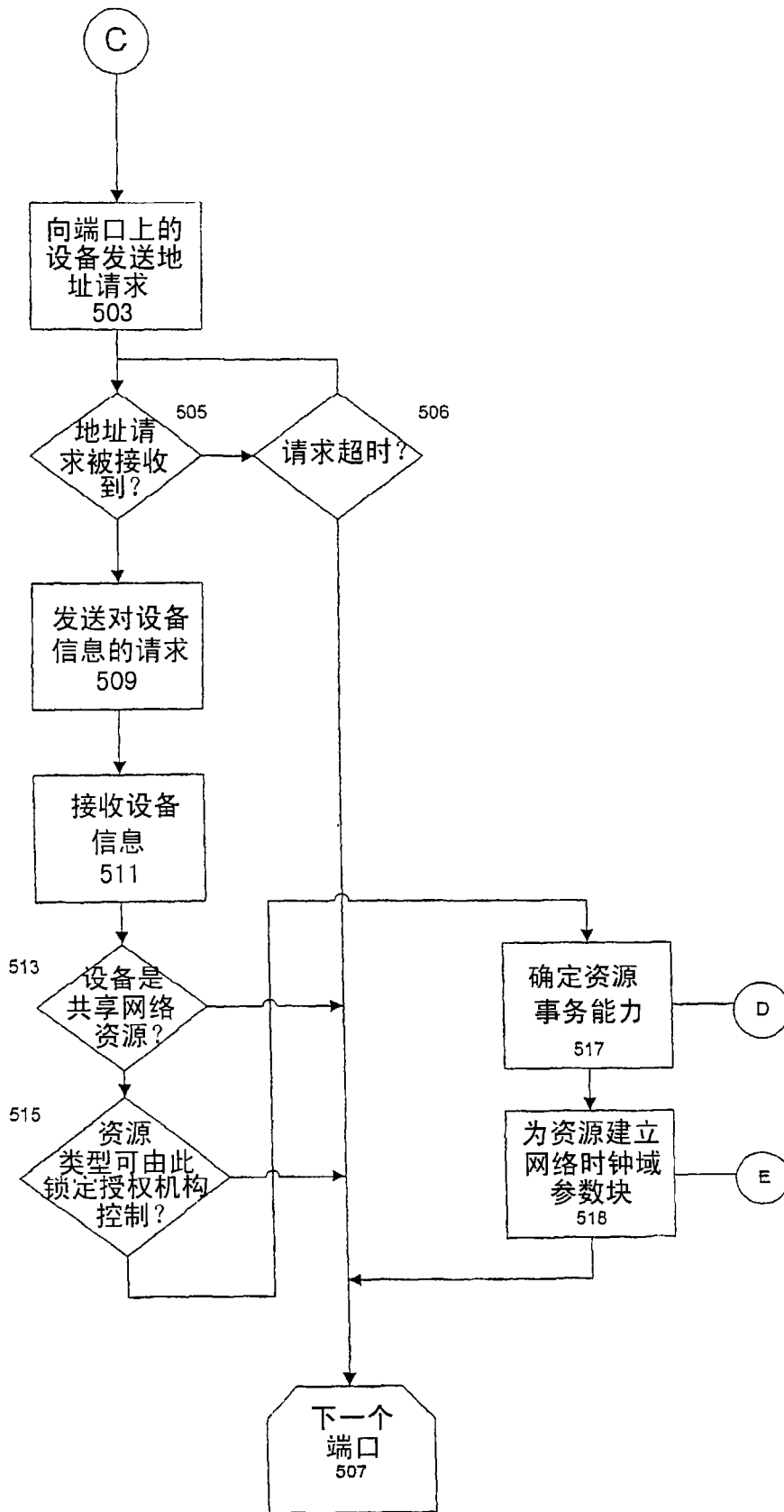


图 5

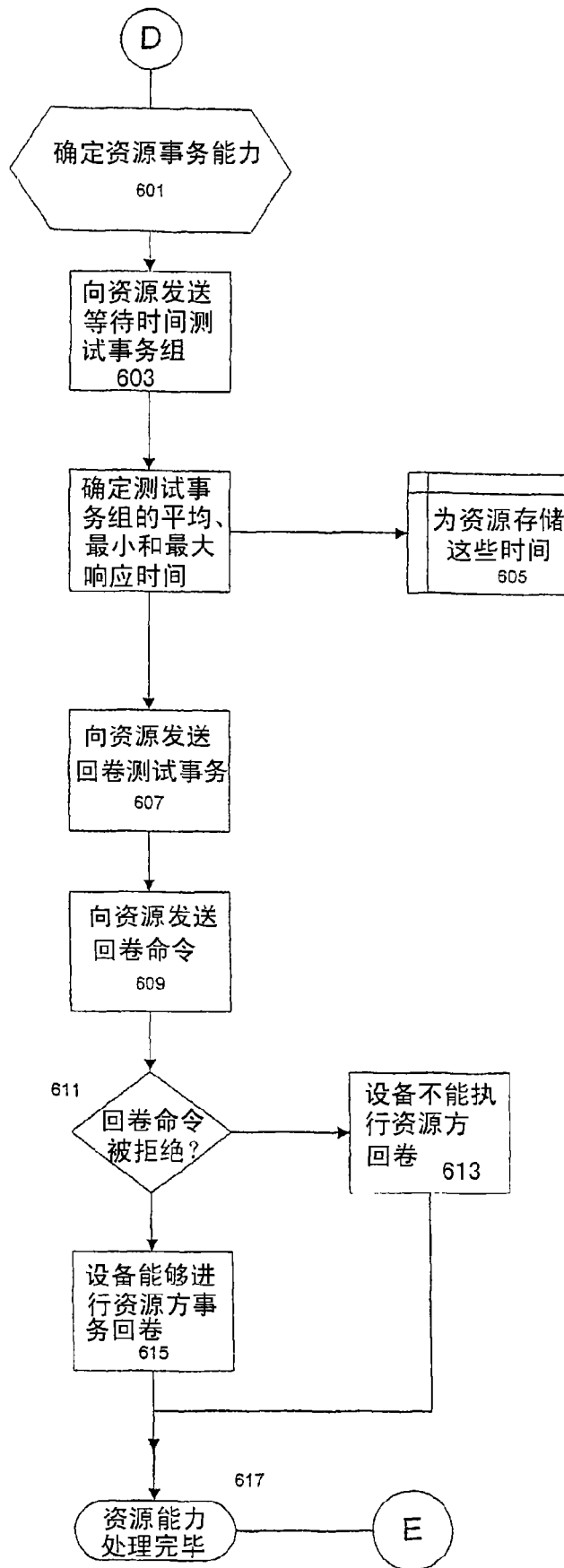


图 6

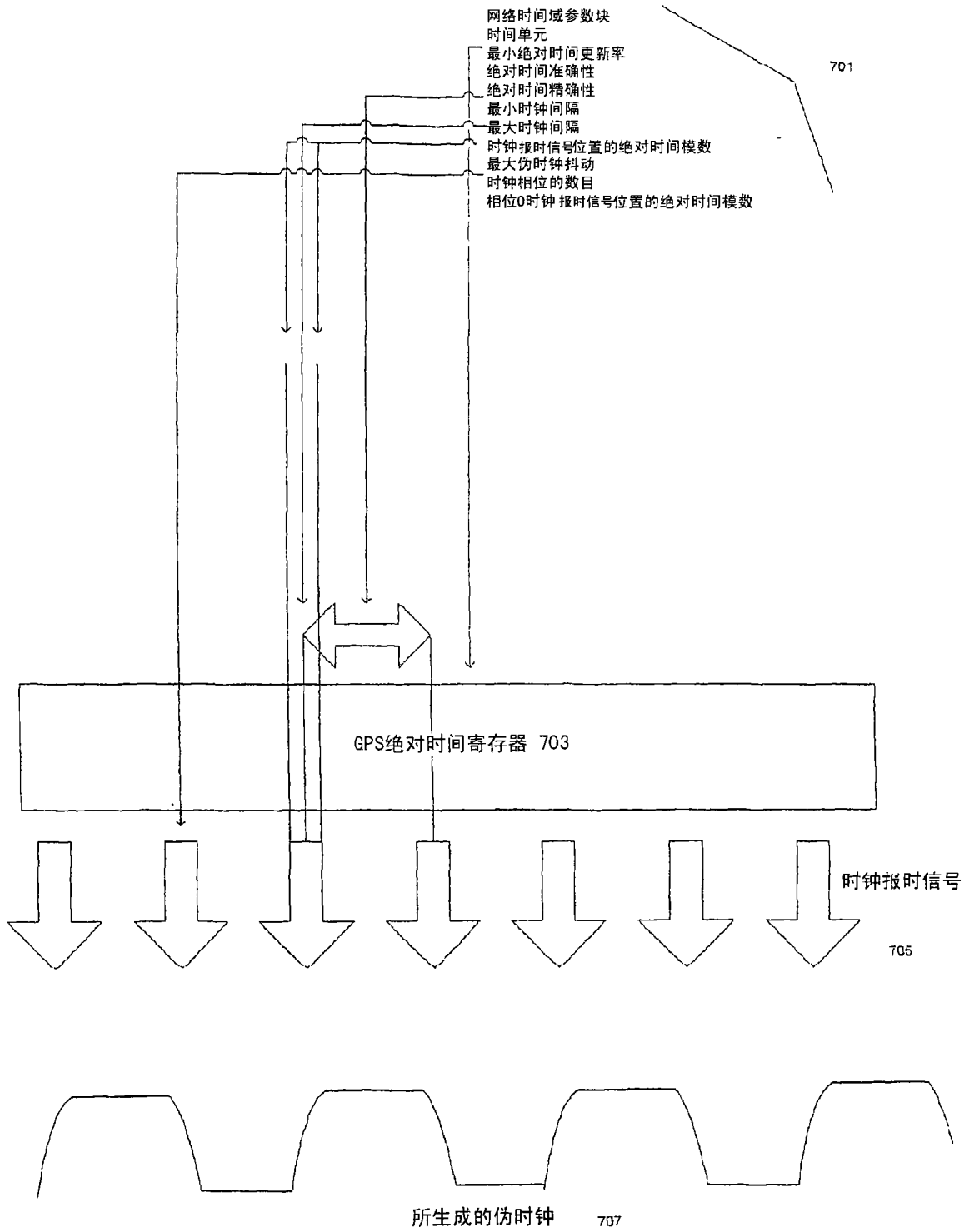


图 7

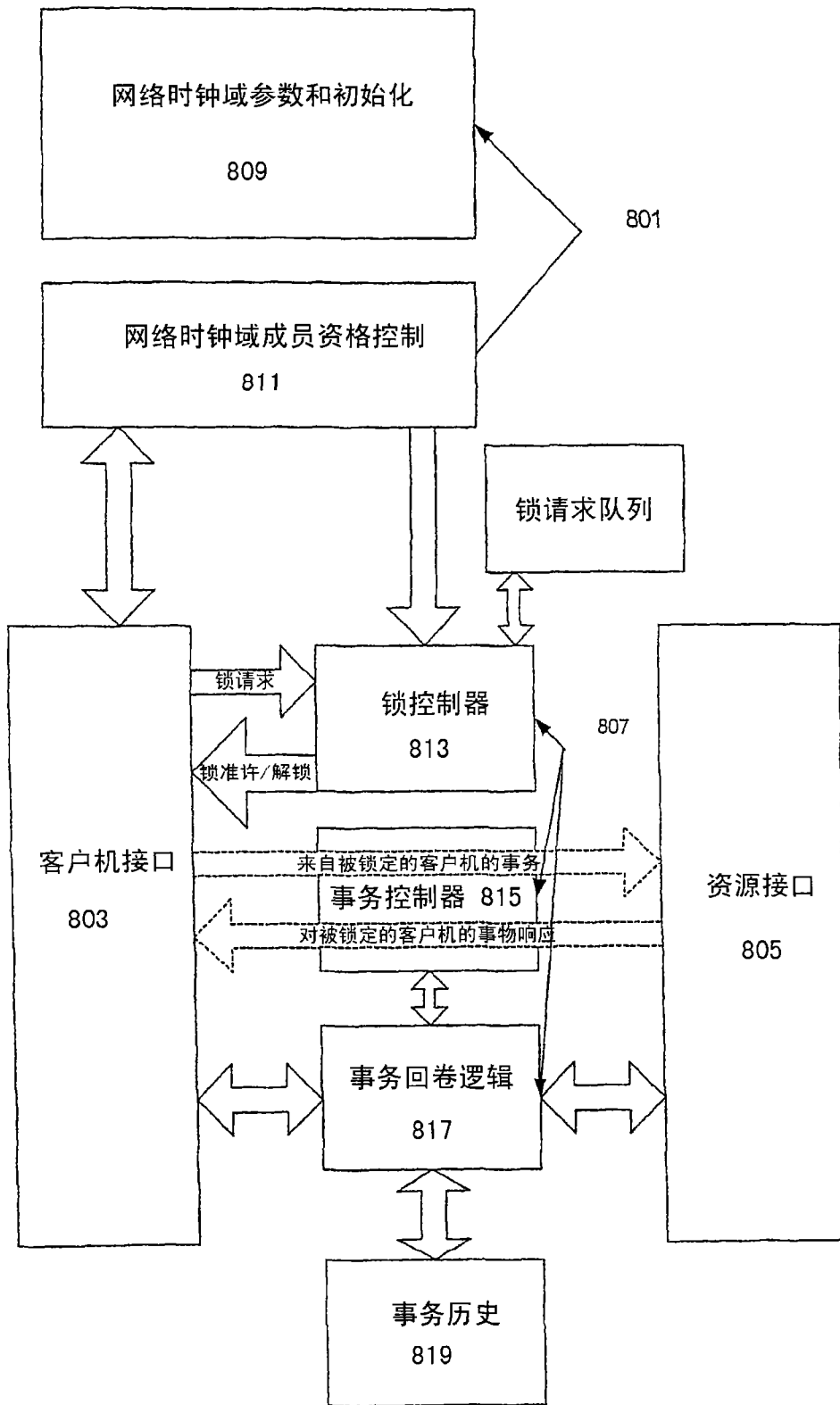


图 8

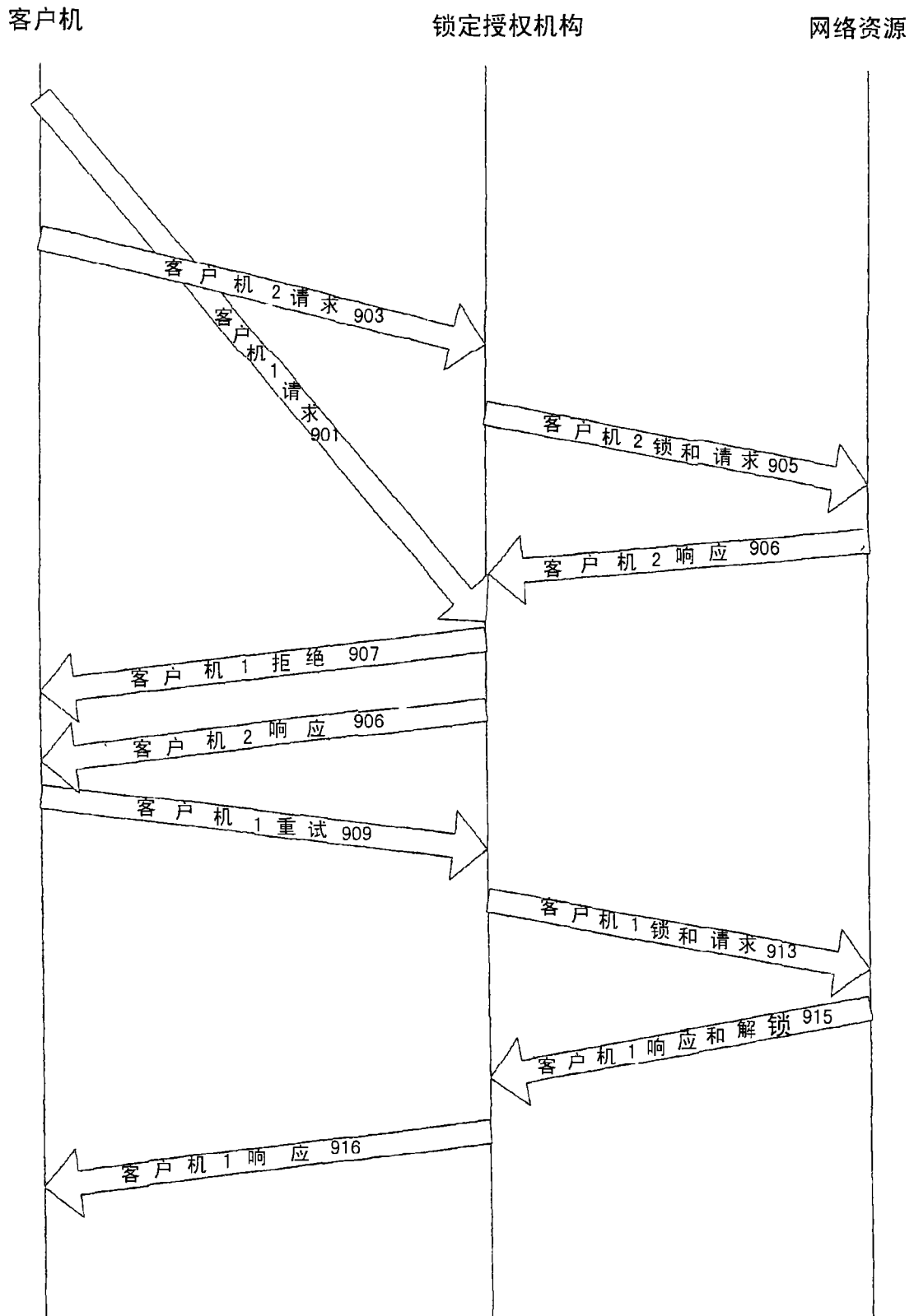


图 9

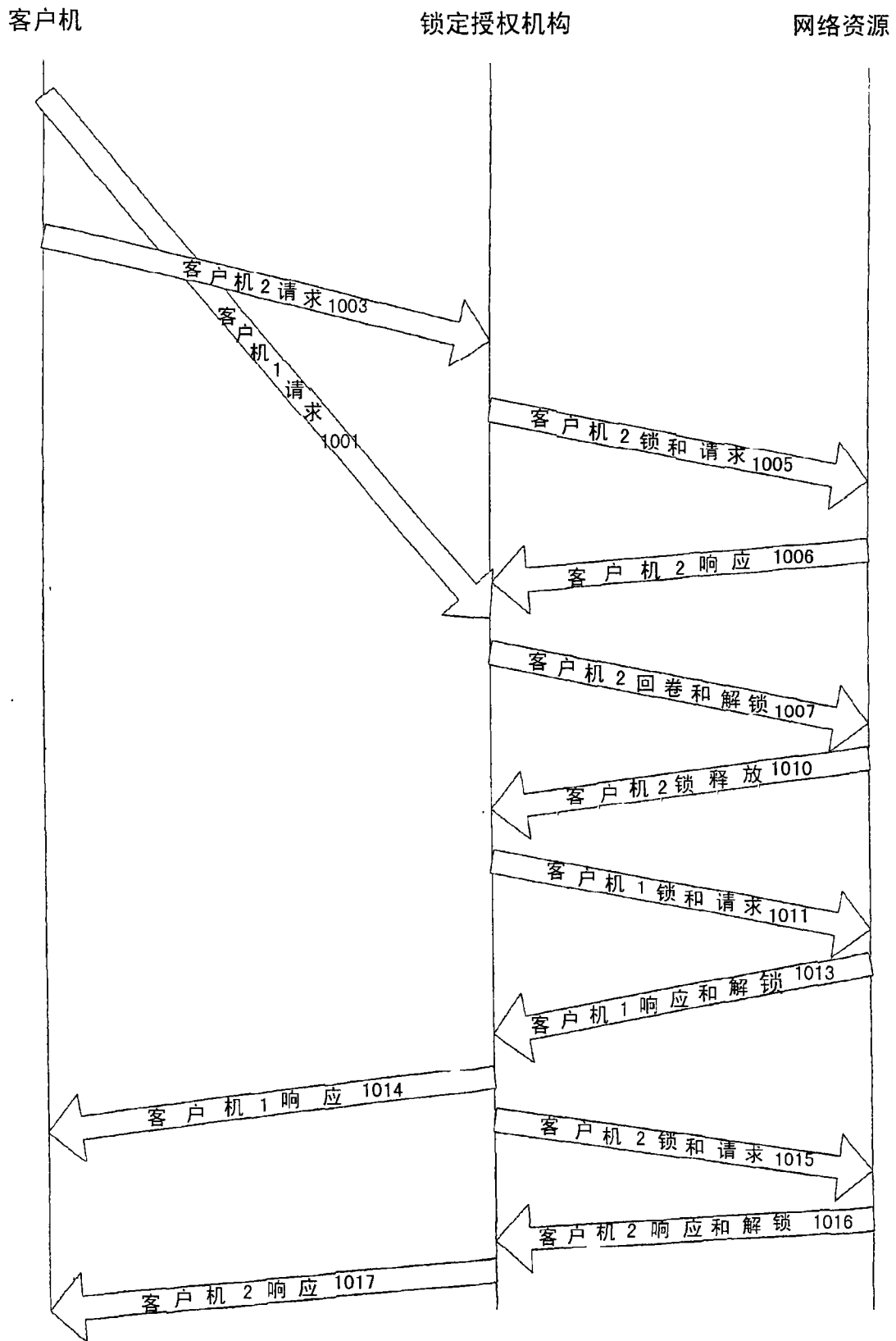


图 10

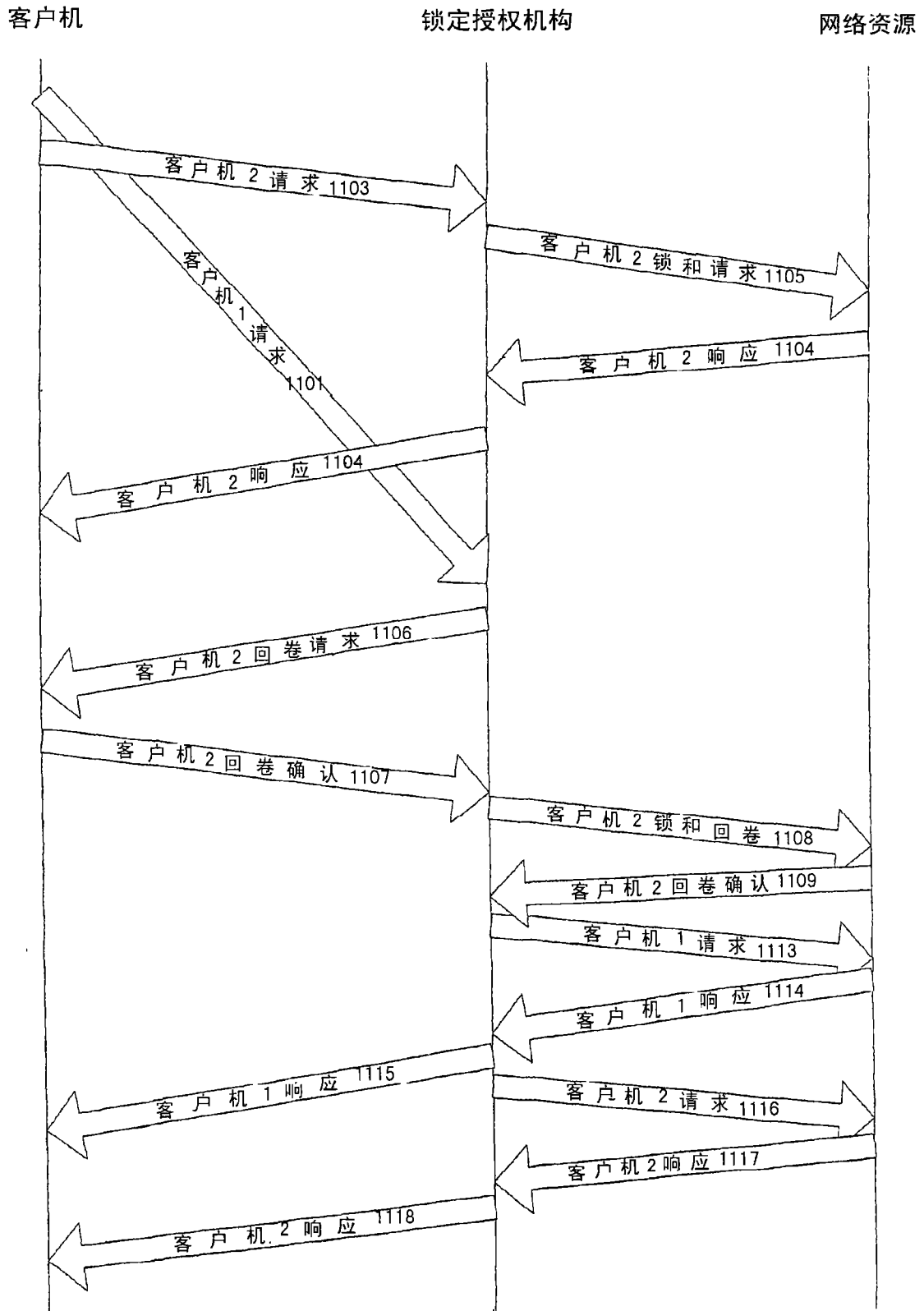


图 11

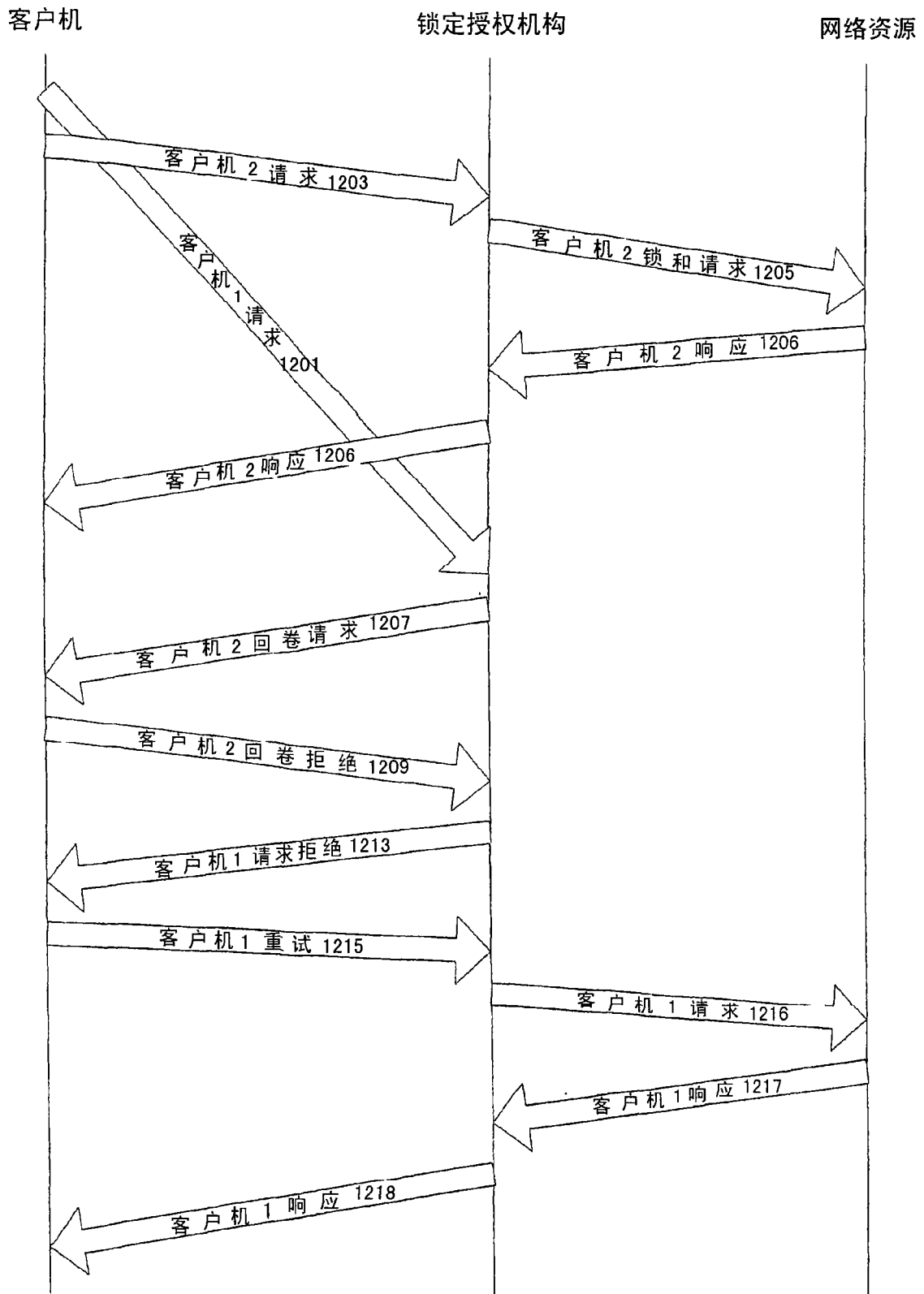


图 12