(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 112425120 A
(43)申请公布日 2021.02.26

(21)申请号 201980045977.6

(22)申请日 2019.06.24

(30)优先权数据
62/716,680 2018.08.09 US
62/722,754 2018.08.24 US
16/433,928 2019.06.06 US

(85)PCT国际申请进入国家阶段日
2021.01.08

(86)PCT国际申请的申请数据
PCT/US2019/038724 2019.06.24

(87)PCT国际申请的公布数据
WO2020/033048 EN 2020.02.13

(71)申请人 赫尔实验室有限公司
地址 美国加利福尼亚州

(72)发明人 A·诺金 J·D·兰姆金斯

(74)专利代理机构 北京三友知识产权代理有限公司 11127
代理人 王万影 王小东

(51)Int.Cl.
H04L 12/18(2006.01)
H04L 29/06(2006.01)

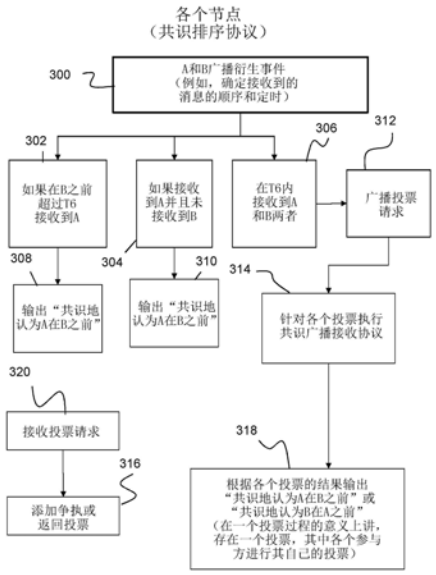
权利要求书3页 说明书11页 附图5页

(54)发明名称

用于广播消息的共识排序的系统和方法

(57)摘要

系统涉及网络中的多个节点，以及节点实现广播消息的共识排序的过程。例如，一个或更多个节点通过就两个或更多个广播消息衍生事件A和B的顺序达成协定而进行操作。如果节点在看到事件B之前的预定时间段(T6)之前看到事件A，则该节点输出“共识地认为A在B之前”作为共识广播排序。如果节点看到了事件A而在至少等待了T6之后未看到事件B，则该节点输出“共识地认为A在B之前”作为共识广播排序。然而，如果节点在T6内看到事件A和事件B两者，则该节点广播对消息排序的投票请求、针对投票执行共识广播接收协议以及基于接收到的投票做出排序决策。



CN 112425120 A

1.一种用于广播消息的共识排序的系统,所述系统包括:

网络中的多个节点,各个节点具有存储器和一个或更多个处理器,所述存储器是编码有可执行指令的非暂时性计算机可读介质,使得在执行所述指令时,所述多个节点中的一个或更多个节点执行以下操作:

在一个或更多个网络上就两个或更多个广播消息衍生事件A和B的顺序达成协定;

其中,如果节点在看到事件B之前的预定时间段(T6)之前看到事件A,则输出"共识地认为A在B之前"作为共识广播排序;

其中,如果所述节点看到了事件A而在至少等待了T6之后未看到事件B,则输出"共识地认为A在B之前"作为共识广播排序;

其中,如果所述节点在T6内看到事件A和事件B两者,则:

广播对消息排序的投票请求;

针对所述投票执行共识广播接收协议;以及

基于接收到的所述投票做出排序决策。

2.根据权利要求1所述的系统,所述系统还包括以下操作:所述网络中的各个节点以某一顺序观察一个或更多个广播消息衍生事件A和B,使得在接收后,所述系统继续执行就两个或更多个消息衍生事件A和B的顺序达成协定的操作。

3.根据权利要求2所述的系统,所述系统还包括以下操作:广播争执消息,使得所述争执消息由第一节点发送,以指示所述第一节点控告所述第二节点未遵循协议。

4.根据权利要求3所述的系统,所述系统还包括以下操作:基于与所述网络中的节点有争执的其它节点的数目来将给定节点指定为已知恶意节点。

5.根据权利要求4所述的系统,其中,所述已知恶意节点是网络中的传感器,其中,按照使来自所述传感器的信号被所述网络中的诚实节点丢弃的方式将所述传感器与所述网络隔离。

6.根据权利要求4所述的系统,其中,所述共识广播接收协议是基于接收到消息衍生事件A的时间与所述消息衍生事件A的发送方变为已知恶意的时间或相对于这种时间的偏移之间的共识排序的,使得:

如果共识地认为所述消息衍生事件A是在所述发送方变为已知恶意的或相对于这种时间的偏移之前接收到的,则在诚实节点当中共识地接收所述消息衍生事件A的实际值;或者

如果共识地认为所述发送方节点变为已知恶意的或者相对于这种时间的偏移在接收到所述发送方节点的网络广播之前,则在所述诚实节点当中共识地接收所述消息衍生事件A的零值。

7.根据权利要求1所述的系统,其中,所述两个或更多个广播消息衍生事件A和B的排序对共享状态进行修改,使得共识排序协议用于维护所述共享状态的一致见解,以确保修改的排序无论何时能够导致不同最终状态,都会向诚实节点通知要使用何种顺序。

8.一种用于广播消息的共识排序的计算机实现的方法,所述方法包括以下动作:使网络中的多个节点中的一个或更多个节点执行编码在非暂时性计算机可读介质上的指令,使得在执行所述指令时,所述多个节点中的一个或更多个节点执行以下操作:

在一个或更多个网络上就两个或更多个广播消息衍生事件A和B的顺序达成协定;

其中,如果节点在看到事件B之前的预定时间段(T6)之前看到事件A,则输出"共识地认

为A在B之前”作为共识广播排序；

其中，如果所述节点看到了事件A而在至少等待了T6之后未看到事件B，则输出“共识地认为A在B之前”作为共识广播排序；

其中，如果所述节点在T6内看到事件A和事件B两者，则：

广播对消息排序的投票请求；

针对所述投票执行共识广播接收协议；以及

基于接收到的所述投票做出排序决策。

9.根据权利要求8所述的方法，所述方法还包括以下操作：所述网络中的各个节点以某一顺序观察一个或更多个广播消息衍生事件A和B，使得在接收后，所述方法继续执行就两个或更多个消息衍生事件A和B的顺序达成协定的操作。

10.根据权利要求9所述的方法，所述方法还包括以下操作：广播争执消息，使得所述争执消息由第一节点发送，以指示所述第一节点控告所述第二节点未遵循协议。

11.根据权利要求10所述的方法，所述方法还包括以下操作：基于与所述网络中的节点有争执的其它节点的数目来将给定节点指定为已知恶意节点。

12.根据权利要求11所述的方法，其中，所述已知恶意节点是网络中的传感器，其中，按照使来自所述传感器的信号被所述网络中的诚实节点丢弃的方式将所述传感器与所述网络隔离。

13.根据权利要求11所述的方法，其中，所述共识广播接收协议是基于接收到消息衍生事件A的时间与所述消息衍生事件A的发送方变为已知恶意的时间或相对于这种时间的偏移之间的共识排序的，使得：

如果共识地认为所述消息衍生事件A是在所述发送方变为已知恶意的或相对于这种时间的偏移之前接收的，则在诚实节点当中共识地接收所述消息衍生事件A的实际值；或者

如果共识地认为所述发送方节点变为已知恶意的或者相对于这种时间的偏移在接收到所述发送方节点的网络广播之前，则在所述诚实节点当中共识地接收所述消息衍生事件A的零值。

14.根据权利要求8所述的方法，其中，所述两个或更多个广播消息衍生事件A和B的排序对共享状态进行修改，使得共识排序协议用于维护所述共享状态的一致见解，以确保修改的排序无论何时导致不同最终状态，都会向诚实节点通知要使用何种顺序。

15.一种用于广播消息的共识排序的计算机程序产品，所述系统包括：

非暂时性计算机可读介质，所述非暂时性计算机可读介质编码有可执行指令，使得一个或更多个处理器执行所述指令时，所述一个或更多个处理器使网络中的多个节点中的一个或更多个节点执行以下操作：

在一个或更多个网络上就两个或更多个广播消息衍生事件A和B的顺序达成协定；

其中，如果节点在看到事件B之前的预定时间段(T6)之前看到事件A，则输出“共识地认为A在B之前”作为共识广播排序；

其中，如果所述节点看到了事件A而在至少等待了T6之后未看到事件B，则输出“共识地认为A在B之前”作为共识广播排序；

其中，如果所述节点在T6内看到事件A和事件B两者，则：

广播对消息排序的投票请求；

针对所述投票执行共识广播接收协议;以及

基于接收到的所述投票做出排序决策。

16.根据权利要求15所述的计算机程序产品,所述计算机程序产品还包括导致以下操作的指令:所述网络中的各个节点以某一顺序观察一个或更多个广播消息衍生事件A和B,使得在接收后,所述计算机程序产品继续执行就两个或更多个消息衍生事件A和B的顺序达成协定的所述操作。

17.根据权利要求16所述的计算机程序产品,所述计算机程序产品还包括导致以下操作的指令:广播争执消息,使得所述争执消息由第一节点发送,以指示所述第一节点控告所述第二节点未遵循协议。

18.根据权利要求17所述的计算机程序产品,所述计算机程序产品还包括导致以下操作的指令:基于与所述网络中的节点有争执的其它节点的数目来将给定节点指定为已知恶意节点。

19.根据权利要求18所述的计算机程序产品,其中,所述已知恶意节点是网络中的传感器,其中,按照使来自所述传感器的信号被所述网络中的诚实节点丢弃的方式将所述传感器与所述网络隔离。

20.根据权利要求18所述的计算机程序产品,其中,所述共识广播接收协议是基于接收到消息衍生事件A的时间与所述消息衍生事件A的发送方变为已知恶意的时间或相对于这种时间的偏移之间的共识排序的,使得:

如果共识地认为所述消息衍生事件A是在所述发送方变为已知恶意的或相对于这种时间的偏移之前接收到的,则在诚实节点当中共识地接收所述消息衍生事件A的实际值;或者

如果共识地认为所述发送方节点变为已知恶意的或者相对于这种时间的偏移在接收到所述发送方节点的网络广播之前,则在所述诚实节点当中共识地接收所述消息衍生事件A的零值。

21.根据权利要求15所述的计算机程序产品,其中,所述两个或更多个广播消息衍生事件A和B的排序对共享状态进行修改,使得共识排序协议用于维护所述共享状态的一致见解,以确保修改的排序无论何时导致不同最终状态,都会向诚实节点通知要使用何种顺序。

用于广播消息的共识排序的系统和方法

[0001] 政府权益

[0002] 本发明是在由国土安全部签发的合同号为HSHQDC-13-C-B0026的政府支持下完成的。政府拥有本发明的某些权益。

[0003] 相关申请的交叉引用

[0004] 本申请是2019年6月6日提交的美国申请No.16/433,928的部分继续申请，该美国申请No.16/433,928是2018年8月9日提交的美国临时申请No.62/716,680的非临时专利申请，该美国临时申请No.62/716,680的全部内容通过引用并入于此。

[0005] 本申请还是2018年8月24日提交的美国临时申请No.62/722,754的非临时专利申请，该美国临时申请No.62/722,754的全部内容通过引用并入于此。

[0006] 发明背景

(1)技术领域

[0007] 本发明涉及用于多方网络协议的系统和方法，并且更具体地，涉及用于实现如下协议的系统和方法，在该协议中，基于该协议中的节点之间的共识排序(consensus ordering)来处理消息。

(2)背景技术

[0008] 本公开涉及解决在实现用于网络上的多个节点(例如，计算机或其它装置)的联网协议时出现的问题。在一个示例中，假设节点使用安全的多方计算(MPC)协议执行一些常见任务。在任何时间点，特定参与者可能检测到另一参与者正在偏离协议并因此是恶意的。应注意，术语“恶意的”是指未能遵循协议的任何原因，诸如，软件错误、硬件故障、网络攻击等。此时，MPC实现方式中的常见方法是所述方广播指示该方与恶意方有争执的“争执(dispute)”消息。注意，争执消息必然指示至少一个参与方是恶意的，但不一定指示哪一方是恶意的，因为恶意方可以选择广播指向诚实一方的争执消息。

[0009] 一些研究人员试图解决与识别恶意方或恶意节点相关联的问题。举例来说，Tapus等人试图通过提供一种与MPC实现方式相反的串行化机制(参见并入的参考文献的列表，参考文献2)来按照非常不同的设置解决该问题。值得注意的是，Tapus等人的工作使用的协议的复杂度要比大多数实现方式中所需的复杂度高。其它小组通信工作集中在开发可靠的多播层来实现总顺序(参见参考文献3和4)。例如，出于文件系统的目的，由于该方法的较低级别以及在这种系统中存在大量多播组的可能性，因此这种方法不能很好地工作。理想地，这种排序的控制应该按照更高级别。

[0010] 换句话说，存在解决了应用层的总顺序和多组进程成员关系的系统(参见参考文献5和6)。这些系统中的一些系统假定它们用于获得子组中全局总排序的节点的层次结构。这种方法是负担较重的，因为即使没有节点是恶意的，也需要大量的通信和显著的延迟。

[0011] 因此，仍然需要与现有技术相比需要非常少的计算开销的简单且有效的协议。

发明内容

[0012] 本公开涉及用于广播消息的共识排序的系统和方法。在各方面中，所述系统包括网络中的多个节点。各个节点具有存储器和一个或更多个处理器。所述存储器是编码有可执行指令的非暂时性计算机可读介质，使得在执行所述指令时，所述多个节点中的一个或更多个节点执行多个操作，诸如：

[0013] 在一个或更多个网络上就两个或更多个广播消息衍生事件A和B的顺序达成协定；

[0014] 其中，如果节点在看到事件B之前的预定时间段(T6)之前看到事件A，则输出“共识地认为A在B之前”作为共识广播排序；

[0015] 其中，如果所述节点看到了事件A而在至少等待了T6之后未看到事件B，则输出“共识地认为A在B之前”作为共识广播排序；

[0016] 其中，如果所述节点在T6内看到事件A和事件B两者，则：

[0017] 广播对消息排序的投票请求；

[0018] 针对所述投票执行共识广播接收协议；以及

[0019] 基于接收到的所述投票做出排序决策。

[0020] 在另一方面，一个或更多个节点还执行以下操作：所述网络中的各个节点以某一顺序观察一个或更多个广播消息衍生事件A和B，使得在接收后，所述系统继续执行就两个或更多个消息衍生事件A和B的顺序达成协定的操作。

[0021] 在又一方面，一个或更多个节点还执行以下操作：广播争执消息，使得所述争执消息由第一节点发送，以指示所述第一节点控告所述第二节点未遵循协议。

[0022] 在又一方面，一个或更多个节点还执行以下操作：基于与所述网络中的节点有争执的其它节点的数目来将给定节点指定为已知恶意节点。

[0023] 在另一方面，所述已知恶意节点是网络中的传感器，其中，按照使来自所述传感器的信号被所述网络中的诚实节点丢弃的方式将所述传感器与所述网络隔离。

[0024] 在另一方面，所述共识广播接收协议是基于接收到消息衍生事件A的时间与所述消息衍生事件A的发送方变为已知恶意的时间或相对于这种时间的偏移之间的共识排序的，使得：

[0025] 如果共识地认为所述消息衍生事件A是在所述发送方变为已知恶意的或相对于这种时间的偏移之前接收的，则在诚实节点当中共识地接收所述消息衍生事件A的实际值；或者

[0026] 如果共识地认为所述发送方节点变为已知恶意的或者相对于这种时间的偏移在接收到所述发送方节点的网络广播之前，则在所述诚实节点当中共识地接收所述消息衍生事件A的零值。

[0027] 在另一方面，所述两个或更多个广播消息衍生事件A和B的排序对共享状态进行修改，使得共识排序协议用于维护所述共享状态的一致见解，以确保修改的排序无论何时导致不同最终状态，都会向诚实节点通知要使用何种顺序。

[0028] 最后，本发明还包括计算机程序产品和计算机实现的方法。所述计算机程序产品包括在非暂时性计算机可读介质上存储的计算机可读指令，所述计算机可读指令能够由具有一个或更多个处理器的计算机执行，使得在执行所述指令时，一个或更多个节点的所述一个或更多个处理器执行本文列出的操作。另选地，计算机实现的方法包括使计算机执行

这种指令并且执行所得到的操作的动作。

附图说明

[0029] 结合参考以下附图，本发明的目的、特征以及优点将从本发明的各个方面的以下详细描述变得显而易见，其中：

[0030] 图1是描绘了根据本发明的各种实施方式的系统的部件的框图；

[0031] 图2是体现本发明的一方面的计算机程序产品的例示图；

[0032] 图3是例示了根据本发明的各方面的共识排序的过程的流程图；

[0033] 图4是例示了根据本发明的各方面的共识广播接收协议的流程图；以及

[0034] 图5是例示了根据本发明的各方面的投票计数协议的流程图。

具体实施方式

[0035] 本发明涉及用于多方网络协议的系统和方法，并且更具体地，涉及用于实现如下协议的系统和方法，在该协议中，基于该协议中的节点之间的共识排序来处理消息。可以实现所述系统的这种协议的非限制性示例是多方网络(例如，多方计算(MPC))。在不考虑共识排序的情况下广播消息。一旦广播消息到达接收方，则该接收方需要按照共识顺序处理该广播消息(即，系统的一些其它部分的正确功能取决于确保广播消息集的所有“诚实”接收方按照相同顺序使用所述广播消息集，即使消息最初到达的顺序在每一个接收方处可能不同)。

[0036] 呈现以下描述以使本领域普通技术人员能够作出和使用本发明并将其结合到特定应用的上下文中。多种修改以及不同应用中的多种用途对于本领域技术人员来说将是显而易见的，并且本文限定的总体原理可以应用于广泛方面。因此，本发明不旨在限于所呈现的各个方面，而是涵盖与本文所公开的原理和新颖特征相一致的最广范围。

[0037] 在下面的详细说明中，阐述了许多具体细节，以使得能够更加彻底地理解本发明。然而，本领域技术人员将明白，本发明可以在不限于这些具体细节的情况下实施。在其它情况下，公知结构和装置按框图形式示出而不被详细示出，以免模糊本发明。

[0038] 读者应留意与本说明书同时提交的所有文件和文档，这些文件和文档与本说明书一起公开以供公众查阅，所有这些文件和文档的内容通过引用并入于此。本说明书(包括任何所附权利要求、摘要以及附图)中公开的所有特征可以由用于相同、等同或相似目的的替代特征来代替，除非另有明确说明。因此，除非另有明确说明，否则所公开的各个特征仅是典型系列的等同或相似特征的一个示例。

[0039] 此外，权利要求中的未明确陈述用于执行指定功能的“装置”或用于执行特定功能的“步骤”的任何要素不被解释为在35U.S.C.第112节第6款中指定的“装置”或“步骤”条款。具体地，在本文的权利要求中使用“…的步骤”或“…的动作”不旨在援引35 U.S.C.第112节第6款的规定。

[0040] 在详细描述本发明之前，首先提供了引用参考文献的列表。接下来，提供了对本发明的各个主要方面的说明。随后，介绍部分为读者提供了本发明的总体理解。最后，提供本发明的各种实施方式的具体细节，以给出具体方面的理解。

[0041] (1)所并入的参考文献的列表

[0042] 在本申请中引用以下参考文献。为了清楚和方便起见，这些参考文献在本文中被列为读者的中心资源。下列参考文献通过引用并入于此，就像在本文中完全陈述一样。这些参考文献通过参照如下对应文献参考号而在本申请中加以引用：

[0043] 1.G.Bracha.An asynchronous[(n-1)/3]-resilient consensus protocol.In T.Kameda,J.Misra,J.Peters,and N.Santoro,editors,Proceedings of the Third Ammal ACM Symposium on Pruciples of Distributed Computing,Vancouver,B.C.,Canada,August 27-29,1984,pages 154-162.ACM,1984.

[0044] 2.Cristian Tapus,Aleksey Nogin,Jason Hickey,and Jerome White.A Simple Serializability Mechanism for a Distributed Objects System.In David A.Bader and Ashfaq A.Khokhar,editors,Proceedings of the 17th International Conference on Paralleladd Distributed Computing Systems (PDCS-2004).International Society for Computers and Their Applications (ISCA),2004.

[0045] 3.G.V.Chockler,N.Huleihel,and D.Dolev.An adaptive totally ordered multicast protocol that tolerates partitions.In Proceedings of the seventeenth ammal ACM symposium on Principles of distrihuted computing,pages 237-246.ACM Press,1998.

[0046] 4.George Coulouris,Jean Dollimore,and Tim Kindberg.Distributed Systems:Concepts and Design.Addison-Wesley,fifthedition,Chapters 15-17(2012).

[0047] 5.Paul D.Ezhilchelvan,Raimundo A.Macˆedo,and Santosh K.Shrivastava.Newtop:a fault-tolerant group communication protocol In Proceedings of the 15th International Conference on Distributed Computing Systems (ICDCS'95),page 296.IEEE Computer Society,1995.

[0048] 6.L.E.Moser,P.M.Melliar-Smith,D.A.Agarwal,R.K.Budhia,and C.A.Lingley-Papadopoulos.Totem:a fault-tolerant multicast group communication system.Commun.ACM,39(4):54-63,1996.

[0049] (2)主要方面

[0050] 本发明的各种实施方式包括三个“主要”方面。第一方面是一种系统。该系统通常采用计算机系统或网络操作软件中的多个计算机系统的形式或采用“硬编码”指令集的形式。该系统可以结合到提供不同功能的各种各样的装置中。第二主要方面是使用数据处理系统(计算机)运行的通常采用软件形式的方法。第三主要方面是计算机程序产品。所述计算机程序产品通常表示存储在诸如光学存储装置(例如，光盘(CD)或数字通用盘(DVD))或磁存储装置(诸如，软盘或磁带)的非暂时性计算机可读介质上的计算机可读指令。计算机可读介质的其它非限制性示例包括硬盘、只读存储器(ROM)以及闪存型存储器。这些方面将在下文进行更详细说明。

[0051] 图1提供了描绘本发明的系统中的至少一个计算机的示例的框图。例如，当在具有多个节点的网络中实现时，各个节点是与网络中其它节点进行通信的独立计算机系统。因此，图1提供了那些计算机系统100中的至少一个计算机系统的非限制性示例。注意，本文所述的系统和方法可以在云以及台式机中的服务器上实现。计算机系统100可以是典型的计算机，或者在其它方面，可以是移动装置以及物联网装置(例如，传感器网络)，甚至可以是

飞机或出于容错和网络安全目而使用该协议(例如,多方计算协议等)的其它平台上的控制计算机集。

[0052] 在各种实施方式中,计算机系统100可以包括被配置成传送信息的地址/数据总线102。另外,一个或更多个数据处理单元(诸如处理器104(或多个处理器))与地址/数据总线102联接。处理器104被配置成处理信息和指令。在一个方面中,处理器104是微处理器。另选地,处理器104可以是不同类型的处理器,诸如并行处理器、专用集成电路(ASIC)、可编程逻辑阵列(PLA)、复杂可编程逻辑器件(CPLD)或现场可编程门阵列(FPGA)。

[0053] 计算机系统100被配置成利用一个或更多个数据存储单元。计算机系统100可以包括与地址/数据总线102联接的易失性存储器单元106(例如,随机存取存储器(“RAM”)、静态RAM、动态RAM等),其中,易失性存储器单元106被配置成存储用于处理器104的信息和指令。计算机系统100还可以包括与地址/数据总线102联接的非易失性存储器单元108(例如,只读存储器(“ROM”)、可编程ROM(“PROM”)、可擦除可编程ROM(“EPROM”)、电可擦除可编程ROM(“EEPROM”)、闪存存储器等),其中,非易失性存储器单元108被配置成存储用于处理器104的静态信息和指令。另选地,计算机系统100可以执行诸如在“云”计算中从在线数据存储单元取回的指令。在一个方面中,计算机系统100还可以包括与地址/数据总线102联接的一个或更多个接口(诸如,接口110)。所述一个或更多个接口被配置成使得计算机系统100能够与其它电子装置和计算机系统对接。由所述一个或更多个接口实现的通信接口可以包括有线通信技术(例如,串行电缆、调制解调器、网络适配器等)和/或无线通信技术(例如,无线调制解调器、无线网络适配器等)。

[0054] 在一个方面中,计算机系统100可以包括与地址/数据总线102联接的输入装置112,其中,输入装置112被配置成将信息和命令选择传送至处理器100。根据一个方面,输入装置112是可以包括字母数字键和/或功能键的字母数字输入装置(诸如键盘)。另选地,输入装置112可以是除字母数字输入装置之外的输入装置。在一个方面中,计算机系统100可以包括与地址/数据总线102联接的光标控制装置114,其中,光标控制装置114被配置成将用户输入信息和/或命令选择传送至处理器100。在一个方面中,光标控制装置114使用诸如鼠标、轨迹球、触控板、光学跟踪装置或触摸屏的装置来实现。尽管如此,但在一个方面中,诸如响应于使用与输入装置112相关联的特殊键和键序列命令,光标控制装置114经由来自输入装置112的输入被引导和/或激活。在另选方面中,光标控制装置114被配置成由语音命令来引导或指导。

[0055] 在一个方面中,计算机系统100还可以包括与地址/数据总线102联接的一个或更多个可选计算机可用数据存储装置(诸如存储装置116)。存储装置116被配置成存储信息和/或计算机可执行指令。在一个方面中,存储装置116是诸如磁或光盘驱动器(例如,硬盘驱动器(“HDD”)、软盘、光盘只读存储器(“CD-ROM”)、数字通用盘(“DVD”))的存储装置。依据一个方面,显示装置118与地址/数据总线102联接,其中,显示装置118被配置成显示视频和/或图形。在一个方面中,显示装置118可以包括阴极射线管(“CRT”)、液晶显示器(“LCD”)、场发射显示器(“FED”)、等离子体显示器或适于显示视频和/或图形图像以及用户可识别的字母数字字符的任何其它显示装置。

[0056] 本文所提出的计算机系统100是根据一个方面的示例计算环境。然而,计算机系统100的非限制性示例并不严格限于是计算机系统。例如,一个方面规定了计算机系统100表

示可以根据本文所述各个方面使用的一种数据处理分析。此外，还可以实现其它计算系统。实际上，本技术的精神和范围不限于任何单个数据处理环境。因此，在一个方面中，使用通过计算机执行的计算机可执行指令(诸如程序模块)来控制或实现本技术的各个方面的一个或更多个操作。在一个实现中，这样的程序模块包括被配置成执行特定任务或实现特定抽象数据类型的例程、程序、对象、部件和/或数据结构。另外，一个方面规定了通过利用一个或更多个分布式计算环境来实现本技术的一个或更多个方面，诸如，在分布式计算环境中，由通过通信网络链接的远程处理装置执行任务，或者诸如，在分布式计算环境中，各种程序模块位于包括存储器-存储装置的本地和远程计算机存储介质中。

[0057] 图2示出了具体实现本发明的计算机程序产品(即，存储装置)的示图。计算机程序产品被示出为软盘200或诸如CD或DVD的光盘202。然而，如先前提到的，计算机程序产品通常表示存储在任何兼容的非暂时性计算机可读介质上的计算机可读指令。关于本发明所使用的术语“指令”通常指示要在计算机上执行的一组操作，并且可以表示整个程序的片段或者个体可分离的软件模块。“指令”的非限制性示例包括计算机程序代码(源代码或目标代码)和“硬编码”电子器件(即，编码到计算机芯片中的计算机操作)。“指令”被存储在任何非暂时性计算机可读介质上，诸如存储在计算机的存储器中或软盘、CD-ROM以及闪存驱动器上。在任一种情况下，这些指令被编码在非暂时性计算机可读介质上。

[0058] (3)介绍

[0059] 本公开提供了用于支持广播操作的网络上的多个节点的联网协议(例如，能够实现这种协议的计算机系统或其它装置)。广播操作是如下传输：如果至少一个接收方接收到消息，那么即使发送方试图“作弊”，所有参与交换的接收方也将接收同一消息。可以通过网络硬件(诸如，Infiniband结构)或经由分离的基础广播子协议(诸如，Bracha的异步广播协议)来确保该属性(参见参考文献1)；其中总是按照发送来自给定发送方的广播的顺序接收所述来自给定发送方的广播，并且其中在特定广播操作完成之前该特定广播操作可以花费多长时间是有限制的。

[0060] 此外，假定如下设置：最多t个参与者(t是公共参数，并且对于该协议而言，t可以与(n-1)/2一样大)可能是恶意的并且以任意方式违反该协议。剩余节点(称为“诚实的”)将遵循本公开中限定的协议。

[0061] 协议被设计成允许关于两个或更多个广播事件之间的排序的全局协定(在非恶意节点或“诚实”节点之间)。例如，如果两个不同发送方在大约同一时间广播，并且不同接收方以不同顺序看到该广播，则该协议将导致某种共识排序，其中所有诚实节点都同意将这两个事件视为以特定顺序发生。当所有非恶意节点以相同顺序看到事件时(例如，如果所述非恶意节点碰巧离得很远)，则共识排序将与此一致。该协议能够对任何基于广播的事件进行排序，而不仅是单独的广播(例如，对“在全部接收到广播A、B和C之后10秒”与“接收到广播D”进行排序)。

[0062] 这种共识排序的重要性在于使诚实节点能够保持一致的全局状态而无需具有任何进一步的协调。仅作为一个示例，如果所有诚实节点以它们不一定直接可见的某种全局状态的相同表示开始，并且各个广播操作暗示对该状态的特定更改，则确保所有诚实节点都对这些更改的顺序达成协定将足以使所述诚实节点保持对该全局状态的一致见解而无需任何进一步的同步。然后，附加协议可以进一步利用该一致全局状态。这样的全局状态的

示例将包括分布式数据库、分布式文件系统和关于哪些节点是恶意的并且将被忽略和/或与相关网络隔离的分布式协定。

[0063] 本公开的该协议至少是针对以下场景而开发的，并且可以在包括共享状态、传感器网络等的各种应用中实现。因此，应理解，尽管本文描述了多方计算(MPC)协议，但该MPC协议被用于一个示例实施方式的例示性目的，并且本发明不旨在限于此。假设节点使用安全的MPC协议执行一些常见任务。在任何时间点，特定参与者可以检测到另一参与者正在偏离协议并因此是恶意的。在这一点上，MPC实现方式中的一种常见方法是使该方广播指示该方与恶意方有争执的“争执”消息。争执消息是来自一方的指示该方与另一方有争执的某种消息，诸如，“[我与X有争执]”。典型的网络协议将具有“报头”(其中报头的一个字段将被指定为“消息类型”)和消息主体的概念。一个示例实现方式将发送一个消息，其中报头中的消息类型字段将包含由协议设计者指派的数字值，以指示“争执”消息类型，并且主体将只是与发送方有争执的节点的身份。

[0064] 注意，争执消息必然指示至少一个参与方是恶意的，但不一定指示哪一参与方是恶意的，因为恶意方可以选择广播指向诚实参与方的争执消息。然而，一旦单方P与至少t+1个其它方存在争执，则知道t+1中的至少一个一定是诚实的(由于阈值t的假定)，因此P一定是恶意的。由于所有参与者可以看到广播争执，因此，此时所有参与者知道P一定是恶意的这一事实，并且P被指定为“已知恶意的”。作为MPC实现方式的一部分，通常不允许已知恶意方对MPC计算做贡献；在该方通常提供了输入的那些计算步骤中，将代替使用空(零)输入。这消除了由于试图从恶意方获取输入而导致的潜在干扰和潜在延迟。然而，如果一方P广播消息M，则MPC协议的正确性通常要求所有诚实方接受消息M，或者所有诚实方将M替换为空/零消息，但无论哪种情况，都必须是一致的。这需要诚实方之间就在P变为已知恶意的之前还是之后广播M达成共识。本协议解决了这个问题并且包括附加属性，即，除了每次新实体变为已知恶意的时(这最多只能发生t次)都会有1次有界延迟(bounded delay)之外，不会向基础MPC处理中引入任何延迟。

[0065] 本文描述的系统和方法在分布式协议领域提供了技术上的改进，并且允许需要这种一致性并且可以在各种应用中使用的系统的显著加速。例如，本公开的协议可以用于允许分布式载具/飞机以安全的方式有效地提交工作或以其它方式与分布式服务器进行通信。

[0066] (4)各种实施方式的具体细节

[0067] 如上所述，本公开提供一种通过广播消息的共识排序实现用于多个节点的联网协议的系统。在详细描述协议之前，对与本协议相关联的概念有一个初步了解是有帮助的。此外，以下还概述了共识排序协议的用法、共识排序协议的细节、共识广播接收协议以及可以被实现以提高协议特定用途性能的优化。下面依次且进一步详细描述这些项中的各项。

[0068] (4.1)概念概述

[0069] 该协议适用于各方可以向所有其它方(可能包括其自身)发送组消息的网络。尽管“广播”术语广泛用于这种消息传输，但也包含其它相关技术(诸如，多播)。该协议还适用于广播传播延迟(发送消息的时刻与所有参与者接收到该消息的时刻之间的时间)有界的网络。广播的最大传播延迟在本文中表示为“T4”。此外，“T6”是预定时间段，并且表示为如下时间段：如果给定节点超过预定时间段(即，T6)间隔接收到两个广播消息，则知道所有节点

按照同一顺序接收到了这些消息。很容易看出，T6最多为2*T4(为简单起见，可以使用2*T4，但是使用单独的T6常数允许更高的通用性并有助于明确描述本发明)。如上所述，假定网络具有类似广播的操作。进一步假定，尽管执行了该操作，但是相关定时属性是已知的。具体地，一旦开始，则已知完成这种操作可以花费多长时间。换句话说，将最坏情况的广播完成时间表示为T4。这意味着，如果两个广播消息(可能来自不同发送方)间隔至少为2*T4被接收到，则保证了所有接收方将以同一顺序看到消息。可以确定小于2*T4的时间常数，该时间常数仍将具有这种属性。具有这种属性的最小已知时间段都表示为T6。两者是构成特定系统中的广播实现方式的最坏情况定时属性的保守估计(保守＝可以大于但不能小于可能的未知真实值)的预定时间段。还很容易看出，如果T6界限适用于所有广播，那么它也适用于衍生于广播的事件，诸如：

[0070] 1.当各个节点接收到针对X的第t+1个争执广播时就会在所述各个节点上发生的“X变为已知恶意的”(这是衍生事件(derived event)，因为不同节点按不同顺序看到该争执，例如，如果超过t+1个节点广播其与X有争执，那么诚实节点可以接收到在特定节点处接收到的前t+1个争执的子集，该子集跨所有节点可能都不相同)

[0071] 2.“自特定广播事件以来已经经过了某一特定时间”

[0072] (4.2)共识排序协议的使用概述

[0073] 共识排序协议的起点是网络中固定节点集(各个节点知道该集是什么)中的各个节点以某一顺序接收相同的广播消息;遵循消息的顺序在接收方之间可能会有所不同的T6约束。当需要包含作为本公开内容核心的共识排序协议的较大系统“一致地”限定广播衍生事件A与B之间的排序时，所述共识排序协议由所述较大系统执行。

[0074] 这里，“一致地”表示在所有诚实节点处，确定结果应该相同。注意，只有当排序实际上很重要时(例如，当A和B两者正修改全局共享状态的相同部分时，或者当A和B中的一者是“X变为已知恶意的”，另一者是“X发送了广播”并且需要就接受由X发送的实际消息还是忽略X并代替接受0达成协定时)，这种较大系统才需要这种一致性。为了实现这种一致性，节点将执行以下指定的共识排序协议。协议结束时，各个节点输出“共识地认为A在B之前”或“共识地认为B在A之前”，并且该协议保证所有诚实节点将输出同一结果，而不管不诚实节点的行为如何。例如，实现安全的MPC协议的系统可以使用如上所述的该协议。另一示例可以是在分布式文件系统或分布式数据库参考同一数据时，其会对写/写或读/写竞争(write races)进行排序。下面进一步详细介绍了一对相互依赖的算法;共识排序协议和共识广播接收协议。

[0075] (4.3)共识排序协议

[0076] 当实现共识排序协议并且如图3所示时，协议通过A和B的任何广播衍生事件300开始。例如，各个诚实节点通过确定至少两个消息A和B的接收的顺序和定时开始。此后，各个诚实节点执行以下步骤：

[0077] 1.如果在B之前超过T6接收到A302:如果诚实节点在看到B之前超过T6看到A，则输出308“共识地认为A在B之前”(“共识地认为B在A之前”类似)。例如，在广播并接收了至少A和(在除2以外的所有情况下)B之后，执行协议。这些输出被传送到较大系统，然后所述较大系统决定其结果(重点是确保所有诚实节点都将得出同一答案)。

[0078] 2.如果接收到A并且未接收到B 304:如果看到了A，并且在A之后至少等待T6之后

没有看到B,则输出310“共识地认为A在B之前”(“共识地认为B在A之前”类似);

[0079] 3.接收到A和B两者306:如果在彼此的T6内看到两个事件,则:

[0080] i广播投票请求312;

[0081] ii.执行下面描述的共识广播接收协议314,以共识地接收n个投票,各个投票是从0到2的数字(投票“0”是从已知恶意节点接收到的投票,投票“1”指定“A在B之前”的投票,投票“2”指定“B在A之前”的投票);

[0082] 1.如果(0的数目+“A在B之后”投票的数目)<t+1,则输出“共识地认为A在B之前”(“共识地认为B在A之前”类似);

[0083] 2.如果(0的数目+“A在B之后”投票的数目)>=t+1以及(0的数目+“B在A之后”投票的数目)>=t+1:

[0084] a.论点(Lemma)(共识排序协议这部分的属性,可以更容易地看出总体共识排序协议具有所需的属性):如果发生这种情况(更普遍的是,如果两侧都有诚实的投票),则所有诚实节点已发送了投票请求;

[0085] b.只要跨所有诚实节点一致,这里任意平局决胜是OK的。(即,本公开描述了一系列实施方式;在该步骤中用于平局决胜的任何一致方法将是本发明的可行实施方式);

[0086] c.当将共识排序协议用作共识广播接收协议的一部分时,以下是所需的平局决胜过程:考虑到“共识地认为”在接收到消息之前发送方变为已知恶意的(即,所有诚实节点决定接收0而不是来自变为已知恶意的节点的实际消息的平局决胜);

[0087] 4.接收投票请求320并添加争执或投票316:在任何节点(包括用户可能正在操作的节点)变为已知恶意的之前(局部地之前,而不是“之前”)从该任何节点接收投票请求之后:

[0088] i.如果在接收到请求之前局部地超过2*T6看到至少一个事件(A或B),则添加与请求方的争执;

[0089] ii.如果在接收到请求之后局部地超过T4仍未看到至少一个事件(A或B),则添加与请求方的争执;

[0090] iii.如果局部地看到至少一个事件,则发送两个事件中的哪个事件先发生的投票;

[0091] 1.论点:将在T4内或2*T4+T6内看到两个事件,请求方将被指定为已知恶意的;以及

[0092] iv.适当超时之后,添加与未针对来自在超时期满的时候并不已知是恶意的节点的请求进行投票的任何节点的争执。

[0093] 5.输出共识排序318:在各个参与者进行其自己的投票之后,协议根据各个投票的结果在节点之间分配共识排序,诸如,“共识地认为A在B之前”或“共识地认为B在A之前”。

[0094] 注意,步骤4是支持步骤,因为其它步骤涉及特定节点如何决定其自己的输出,而步骤4涉及帮助其它节点做出如图3所示的独立序列的决定。在步骤3ii中,协议需要等待,直到获得n个投票为止。当然,恶意节点可能拒绝投票,但是如果节点恶意地过多拖延其投票,则将导致:1)在步骤4.iv中添加与该节点的争执(除非已经存在争执),2)随着各个诚实节点进入步骤4.iv,最终将存在t+1个争执(因为至少有t+1个诚实节点),并且该节点将变为已知恶意的(如上所述),以及3)现在可以将来自已知恶意节点的预期投票替换为0(再次

如上所述),最终使投票总数达到n。

[0095] (4.4)共识广播接收协议

[0096] 共识广播接收协议用于保证网络中的所有节点共识地接收同一值。在操作中并且如图4所示,共识广播接收协议过程如下进行:

[0097] 1.每次从发送方400接收到网络广播时,在接收到广播的时间与该广播的发送方变为已知恶意的(这是广播衍生事件)时间之间利用共识排序;

[0098] i.402:如果共识地认为在发送方变为已知恶意的之前接收到网络广播,则共识地接收该网络广播的实际值,或者

[0099] ii.404:如果共识地认为在接收到发送方节点的网络广播之前该发送方节点变为已知恶意的,则共识地接收0值。

[0100] (4.5)优化以提高协议特定用途的性能

[0101] 本公开还提供了一种替代的移位共识广播接收协议,其针对恶意节点的发现是罕见事件的用例进行了优化。罕见事件是取决于特定实现方式和开发人员视情况而定的。作为非限制性示例,可以将超过Z(例如,每月超过一次等)的恶意节点的发现预定义为罕见事件。在这种场景中,系统实现上述基本共识广播接收协议,将“发送方变为已知恶意的”替换为“在发送方变为已知恶意的之后T6”。

[0102] 通过将“不确定”时段(网络广播消息的接收将导致需要由共识广播协议调用的共识排序协议执行其投票步骤3的时段)从发送方变为已知恶意的附近的间隔[-T6;T6]移位至其附近的[0;2*T6],系统允许立即处理来自尚未变为已知恶意的发送方的所有消息。“不确定”时段是网络广播消息的接收将导致需要由共识广播协议调用的共识排序协议来执行其上面的投票步骤3的时段。在网络或系统可以开始完全忽略来自刚刚变为已知恶意的节点的广播(例如,阻止来自节点的广播,将广播与该节点隔离开,以仅供管理员节点查看,同时终止到网络中的其它节点的广播)之前,对不确定时段进行移位的成本使系统可能愿意等待并接受来自刚刚变为已知恶意的节点的广播的时段长度翻倍。

[0103] 类似地,在共识排序协议中,有用的优化是使用在另一方向上移位的共识广播接收协议。这将具有以下效果:能够不接受来自已知恶意节点的任何投票,而必须等待2*T6才能接受来自非已知恶意节点的任何投票。好处是,这将限制不必要的递归(即,在可以成功完成两个协议中的任何一个协议之前,两个协议可能递归地相互调用结束的次数),只有在另一节点在起初触发投票的初始节点之后很快变为已知恶意的后,才会发生任何递归。为了阐明上述内容,只要共识排序协议在步骤3(ii)中调用共识广播接收协议,就会发生递归,因为这又会反过来调用共识排序协议。

[0104] 一旦采取了上述两个优化,则最终的优化是对共识排序协议用于共识广播接收协议的服务的投票进行批处理,这导致了如图5所示的以下优化的共识广播协议:

[0105] 1. 500:当接收到新的网络广播并且发送方不是已知恶意的时,立即共识地接收该新的网络广播;

[0106] 2. 502:一旦发送方变为已知恶意的,则启动定时器,直到不确定时段(发送方变为已知恶意的开始并在此之后2×T6结束的时间段)在2×T6后结束为止,该定时器一直继续。

[0107] 3. 504:当接收到新的网络广播并且发送方处于不确定时段时,将消息排队以便

以后投票；

[0108] 4．506：当不确定时段关于特定发送方结束时，要求对该发送方的所有排队消息进行投票(可以使用单个投票消息请求来要求对一系列消息进行投票)，然后遵循剩余的共识排序协议(步骤3-4)，然后是剩余的共识广播接收协议(步骤1(i)-1(ii))。

[0109] 5．508：当接收到新的网络广播并且发送方是已知恶意的长达超过不确定时段时，共识地接收0值。(例如，如果消息是投票，则“0”、“1”和“2”是合法值，并且在这种情况下，来自已知恶意发送方的投票将被替换为“0”，而不管其实际是何值)。

[0110] (4.6)示例实现方式

[0111] 如本领域技术人员可以理解的，存在可以实现本文描述的协议的多种应用。例如，如果在载具(例如，从传感器接收数据的汽车或飞机)中存在多个冗余电子控制单元，则对于控制单元来说，具有一致的传感器数据见解以使得其自己的状态是一致的是非常重要的。在该示例中，控制单元可以使用该协议来创建传感器数据的共识排序、考虑特定传感器故障的决策的共识排序以及考虑冗余控制器故障之一的决策的共识排序。这允许多个控制单元维护整个系统状态的相同见解。在该示例中，如果控制单元一致认为传感器中的一个传感器有故障，则系统可以隔离或以其它方式停止从相关传感器(在该上下文中被视为“恶意的”)接收数据。作为又一示例，如果控制单元一致认为冗余控制器中的一个冗余控制器有故障，则该特定故障控制器可以被隔离或以其它方式从与剩余控制单元的通信中移除。

[0112] 最后，虽然已经根据多个实施方式对本发明进行了说明，但本领域普通技术人员应当容易地认识到，本发明可以在其它环境中具有其它应用。应注意，可以有许多实施方式和实现。此外，所附权利要求绝不旨在将本发明的范围限于上述特定实施方式。另外，“用于…的装置”的任何用语旨在引发要素和权利要求的装置加功能的解读，而未特别使用“用于…的装置”用语的任何要素不应被解读为装置加功能要素，即使权利要求以其它方式包括了“装置”一词。此外，虽然已经按特定顺序陈述了特定方法步骤，但这些方法步骤可以按任何期望的顺序进行，并且落入本发明的范围内。

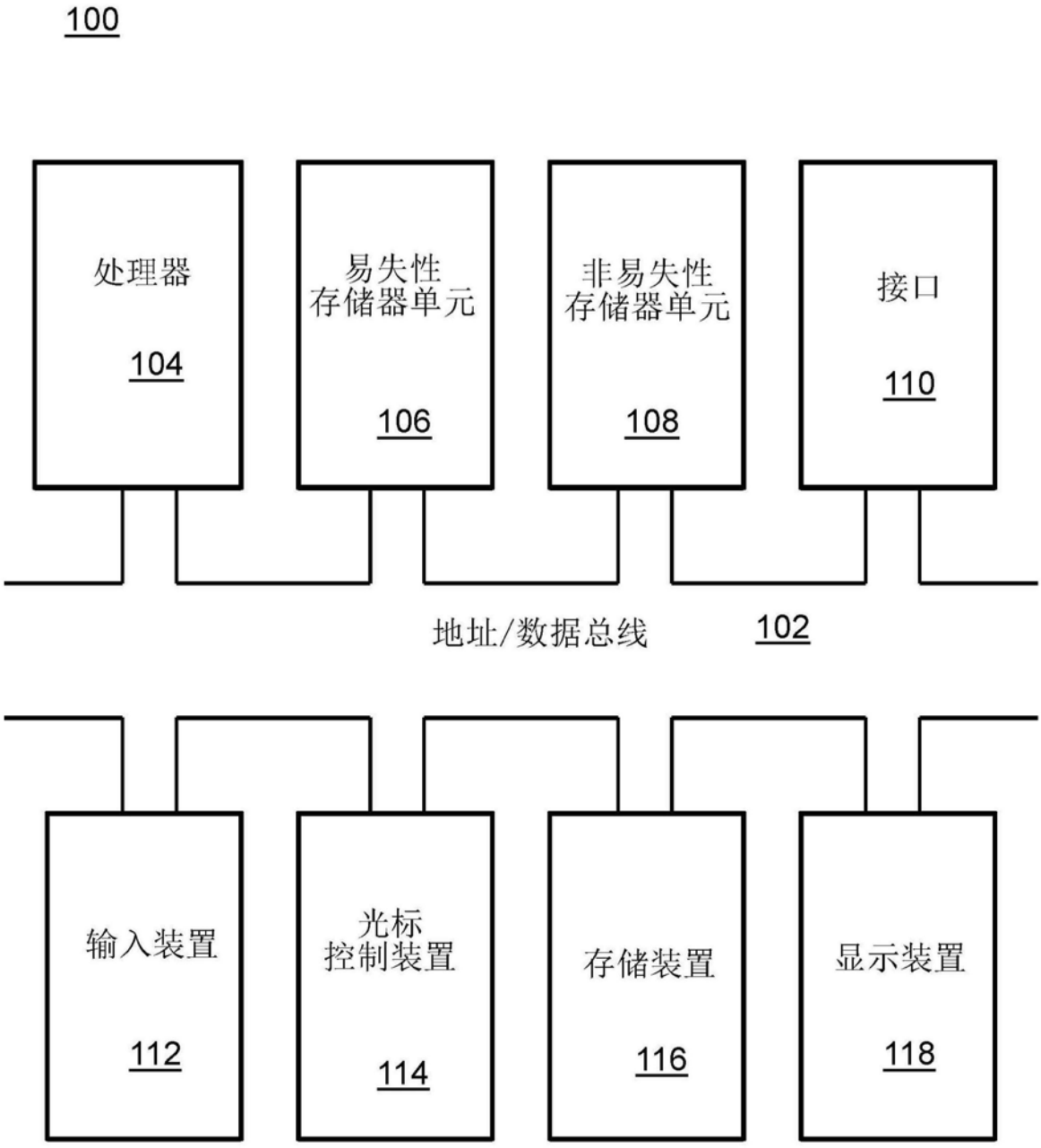

100
处理器
104
易失性
存储器单元
106
非易失性
存储器单元
108
接口
110
地址/数据总线
102
输入装置
112
光标
控制装置
114
存储装置
116
显示装置
118

图1

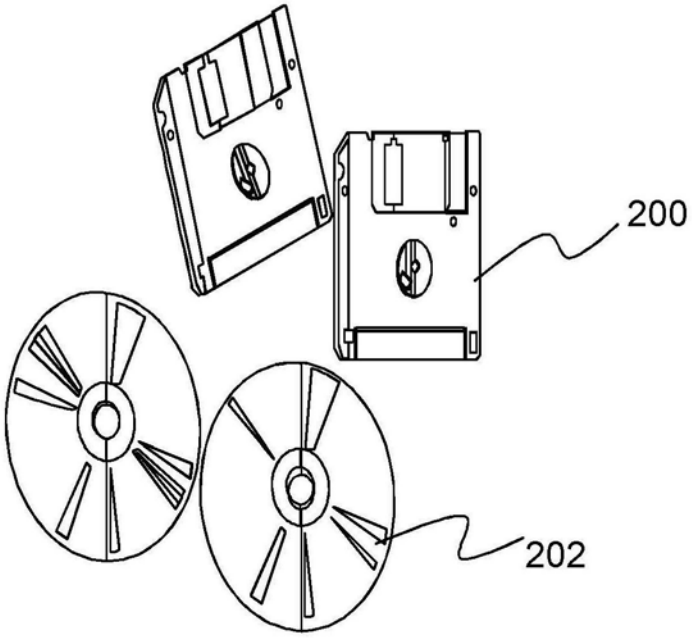


图2

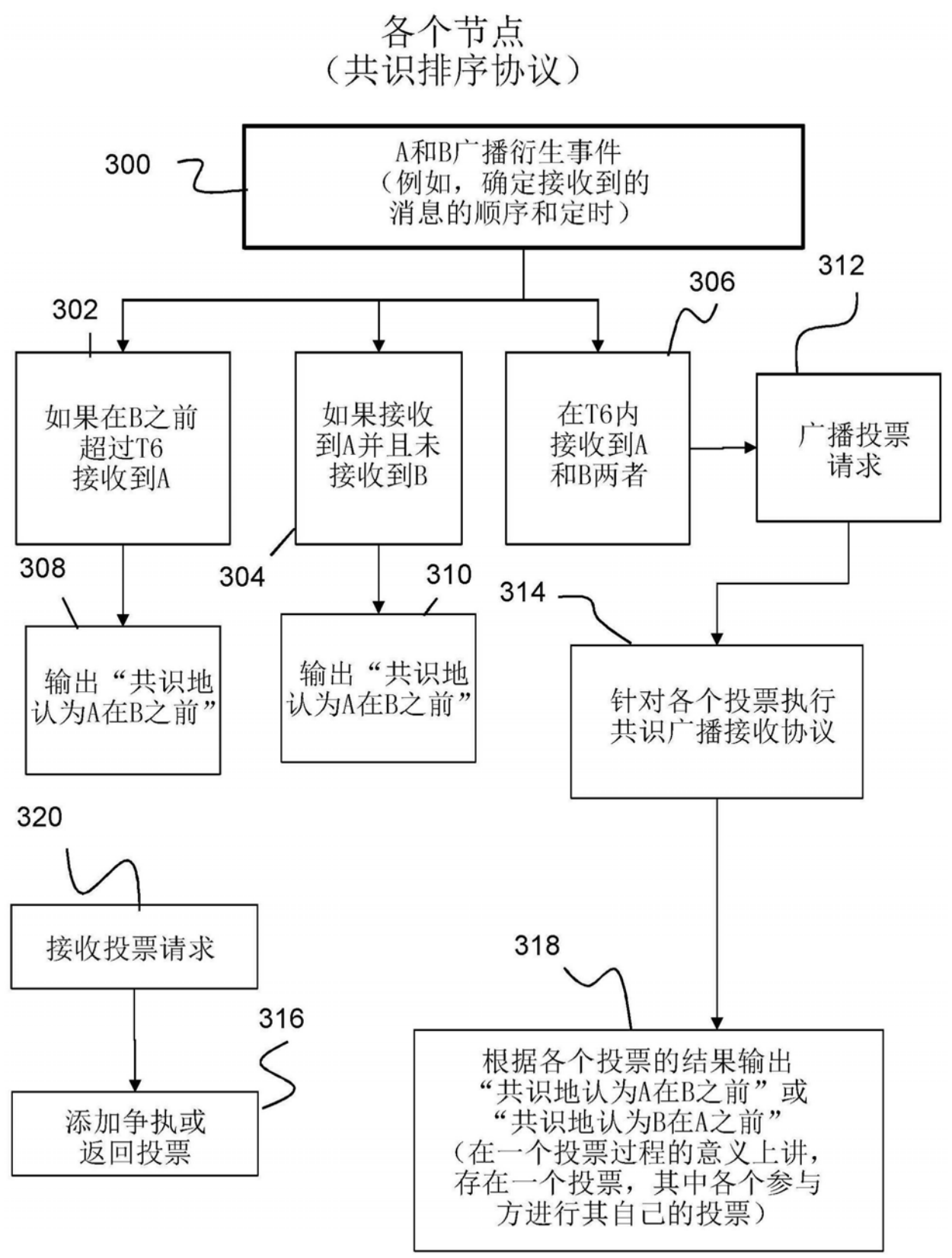

各个节点
（共识排序协议）
300
A和B广播衍生事件
（例如，确定接收到的
消息的顺序和定时）
302
306
312
如果在B之前
超过T6
接收到A
如果接收
到A并且未
接收到B
在T6内
接收到A
和B两者
广播投票
请求
308
304
310
314
输出“共识地
认为A在B之前”
输出“共识地
认为A在B之前”
针对各个投票执行
共识广播接收协议
320
接收投票请求
318
316
添加争执或
返回投票
根据各个投票的结果输出
“共识地认为A在B之前”或
“共识地认为B在A之前”
（在一个投票过程的意义上讲，
存在一个投票，其中各个参与
方进行其自己的投票）

图3

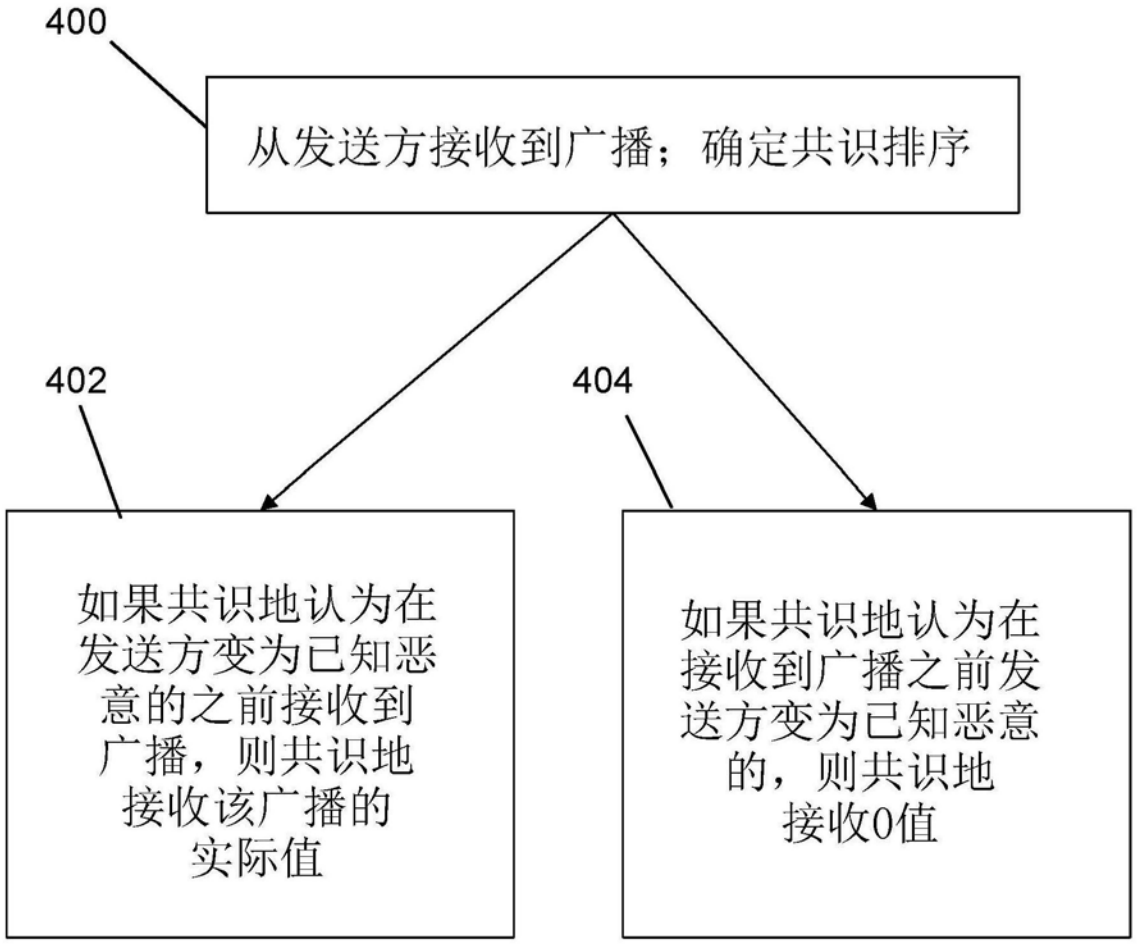

400
从发送方接收到广播；确定共识排序
402
404
如果共识地认为在发送方变为已知恶意的之前接收到广播，则共识地接收该广播的实际值
如果共识地认为在接收到广播之前发送方变为已知恶意的，则共识地接收0值

图4

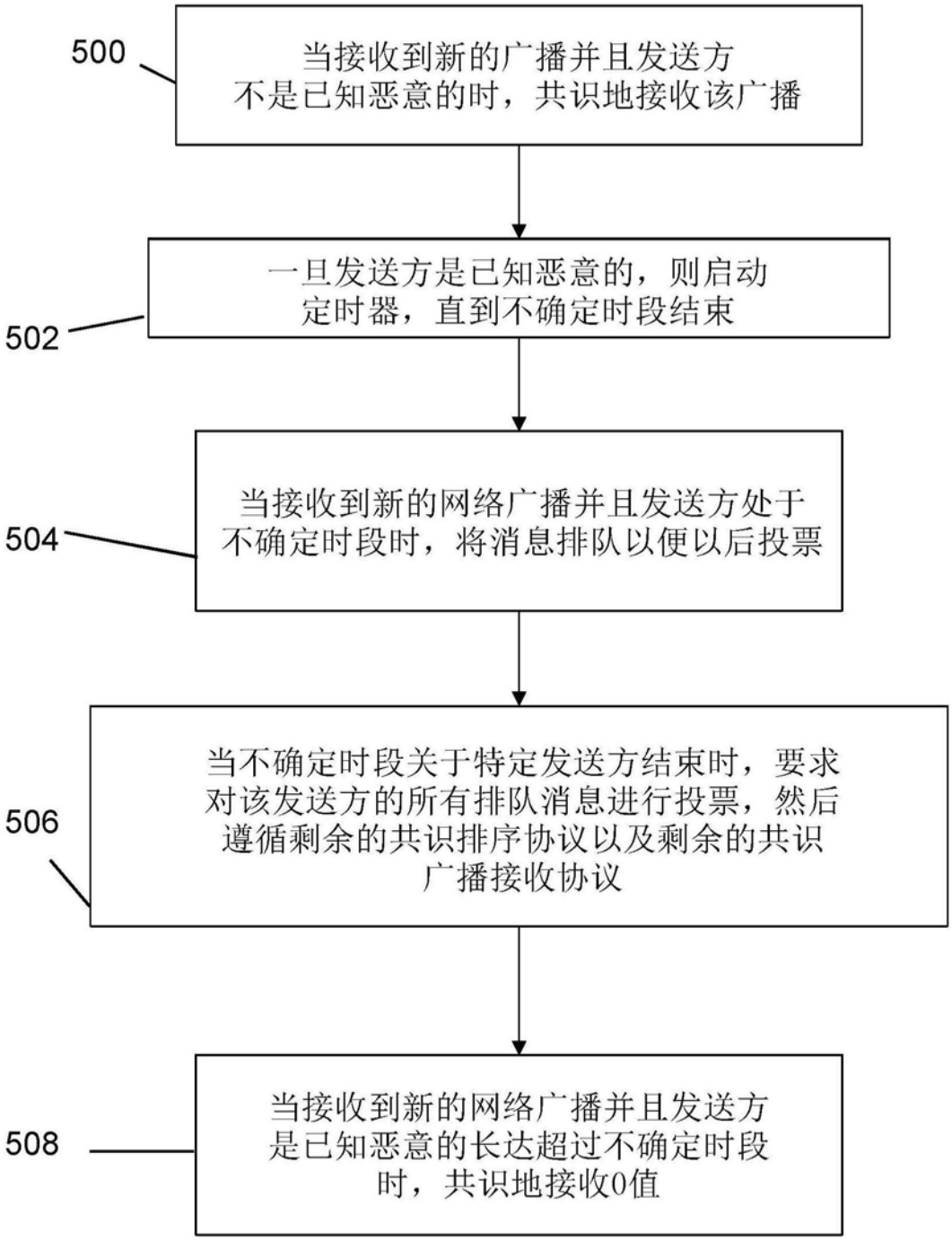

500
当接收到新的广播并且发送方不是已知恶意的时，共识地接收该广播
502
一旦发送方是已知恶意的，则启动定时器，直到不确定时段结束
504
当接收到新的网络广播并且发送方处于不确定时段时，将消息排队以便以后投票
506
当不确定时段关于特定发送方结束时，要求对该发送方的所有排队消息进行投票，然后遵循剩余的共识排序协议以及剩余的共识广播接收协议
508
当接收到新的网络广播并且发送方是已知恶意的长达超过不确定时段时，共识地接收0值

图5