(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 110178340 A
(43)申请公布日 2019.08.27

(21)申请号 201880005441.7

(22)申请日 2018.12.13

(85)PCT国际申请进入国家阶段日
2019.06.26

(86)PCT国际申请的申请数据
PCT/CN2018/120870 2018.12.13

(87)PCT国际申请的公布数据
WO2019/072295 EN 2019.04.18

(71)申请人 阿里巴巴集团控股有限公司
地址 英属开曼群岛大开曼资本大厦一座四层847号邮箱

(72)发明人 林鹏

(74)专利代理机构 北京博思佳知识产权代理有限公司 11415
代理人 艾佳

(51)Int.Cl.
H04L 9/32(2006.01)
H04L 29/06(2006.01)

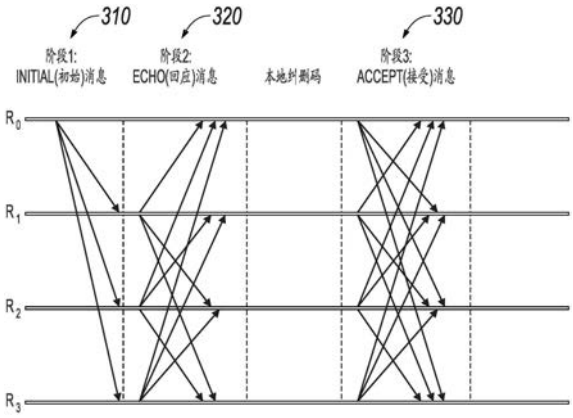
权利要求书4页 说明书30页 附图15页

(54)发明名称

在分布式系统中进行网络节点的恢复处理

(57)摘要

一种用于对区块链网络的网络节点进行恢复处理的方法包括:区块链网络的网络节点将状态请求消息广播至所述区块链网络的其他网络节点待恢复目标交易具有目标序列号,从所述其他网络节点接收状态回复消息,所述状态回复消息各自包括序列号,基于所述状态回复消息中的所述序列号识别所述目标序列号,向所述其他网络节点发送请求消息以从每个所述其他网络节点请求ECHO消息,确定所述其他网络节点发送的有效ECHO消息的数量,基于所述多个有效ECHO消息恢复所述目标交易,以及向所述其他网络节点发送表示所述网络节点已被恢复的消息。



CN 110178340 A

1.一种计算机实现的方法,用于对区块链网络的网络节点进行恢复处理,所述方法包括:

区块链网络的网络节点将状态请求消息广播至所述区块链网络中的多个其他网络节点,其中,所述网络节点待恢复目标交易具有目标序列号;

所述网络节点从所述多个其他网络节点接收多个状态回复消息,其中,所述多个状态回复消息各自包括序列号;

响应于确定各自包括相同的序列号的状态消息的数量超过了预设阈值,所述网络节点基于所述相同的序列号识别所述目标序列号;

所述网络节点向所述多个其他网络节点发送请求消息,其中,所述请求消息从所述多个其他网络节点分别请求ECHO消息,所述ECHO消息是由所述多个其他网络节点分别传输并用于在所述多个其他网络节点之间就所述目标交易具有所述目标序列号达成共识的消息,并且所述ECHO消息包括所述目标交易的一部分以及所述多个其他网络节点各自的签名;

所述网络节点从所述多个其他网络节点接收多个ECHO消息;

所述网络节点确定所述多个ECHO消息中的有效ECHO消息的数量,其中,所述有效ECHO消息各自包括所述目标序列号;

响应于确定所述有效ECHO消息的数量超过预设阈值,所述网络节点基于所述多个有效ECHO消息在所述网络节点处恢复所述目标交易具有所述相同的序列号;以及

所述网络节点向所述多个其他网络节点发送表示所述网络节点已被恢复的消息。

2.如权利要求1所述的方法,其中,所述多个网络节点包括主节点和一个或多个备节点。

3.如权利要求1所述的方法,其中,所述网络节点是主节点或备节点。

4.如权利要求1所述的方法,其中,所述请求消息包括所述目标序列号。

5.如权利要求4所述的方法,其中,所述方法还包括:

所述多个其他网络节点中除所述网络节点外的每个其他网络节点在向所述网络节点发送所述ECHO消息之前,验证所述请求消息。

6.如权利要求1所述的方法,其中,所述方法还包括:

所述网络节点验证每个所述ECHO消息是否有效,

其中,验证每个所述ECHO消息是否有效,包括:利用梅克尔树验证每个所述ECHO消息是否有效。

7.如权利要求6所述的方法,其中,验证每个所述ECHO消息是否有效还包括:

验证所述ECHO消息中的所述签名是否有效。

8.如权利要求1所述的方法,其中,

每个所述ECHO消息还包括与所述目标交易相关联的多个纠删码EC区块中的至少一个区块,

所述多个EC区块是利用所述目标交易根据EC码生成的。

9.如权利要求1所述的方法,其中,基于所述多个有效ECHO消息在所述网络节点处恢复所述目标交易具有所述相同的序列号,包括:

利用所述多个有效ECHO消息中的所述多个EC区块的子集重建所述目标交易。

10.如权利要求1所述的方法,其中,发送至所述多个其他网络节点的表示所述网络节

点已被恢复的所述消息包括：

所述多个有效ECHO消息中的一组签名，和

所述目标序列号。

11.一种非暂态计算机可读存储介质，其耦接到一个或多个计算机并配置有可由所述一个或多个计算机执行的指令，用于：

区块链网络的网络节点将状态请求消息广播至所述区块链网络中的多个其他网络节点，其中，所述网络节点待恢复目标交易具有目标序列号；

所述网络节点从所述多个其他网络节点接收多个状态回复消息，其中，所述多个状态回复消息各自包括序列号；

响应于确定各自包括相同的序列号的状态消息的数量超过了预设阈值，所述网络节点基于所述相同的序列号识别所述目标序列号；

所述网络节点向所述多个其他网络节点发送请求消息，其中，所述请求消息从所述多个其他网络节点分别请求ECHO消息，所述ECHO消息由所述多个其他网络节点分别传输并用于在所述多个其他网络节点之间就所述目标交易具有所述目标序列号达成共识的消息，并且所述ECHO消息包括所述目标交易的一部分以及所述多个其他网络节点各自的签名；

所述网络节点从所述多个其他网络节点接收多个ECHO消息；

所述网络节点确定所述多个ECHO消息中的有效ECHO消息的数量，其中，所述有效ECHO消息各自包括所述目标序列号；

响应于确定所述有效ECHO消息的数量超过预设阈值，所述网络节点基于所述多个有效ECHO消息在所述网络节点处恢复所述目标交易具有所述相同的序列号；以及

所述网络节点向所述多个其他网络节点发送表示所述网络节点已被恢复的消息。

12.如权利要求11所述的非暂态计算机可读存储介质，其中，所述多个网络节点包括主节点和一个或多个备节点。

13.如权利要求11所述的非暂态计算机可读存储介质，其中，所述网络节点是主节点或备节点。

14.如权利要求11所述的非暂态计算机可读存储介质，其中，所述请求消息包括所述目标序列号。

15.如权利要求14所述的非暂态计算机可读存储介质，还配置有可由所述一个或多个计算机执行的指令，用于：

所述多个其他网络节点中除所述网络节点外的每个其他网络节点在向所述网络节点发送所述ECHO消息之前，验证所述请求消息。

16.如权利要求11所述的非暂态计算机可读存储介质，还配置有可由所述一个或多个计算机执行的指令，用于：

所述网络节点验证每个所述ECHO消息是否有效，其中，所述验证每个所述ECHO消息是否有效包括：利用梅克尔树验证每个所述ECHO消息是否有效。

17.如权利要求16所述的非暂态计算机可读存储介质，其中，验证每个所述ECHO消息是否有效还包括：验证所述ECHO消息中的所述签名是否有效。

18.如权利要求11所述的非暂态计算机可读存储介质，其中，每个所述ECHO消息还包括与所述目标交易相关联的多个纠删码EC区块中的至少一个区块，所述多个EC区块是利用所

述目标交易根据EC码生成的。

19.如权利要求11所述的非暂态计算机可读存储介质,其中,基于所述多个有效ECHO消息在所述网络节点处恢复所述目标交易具有所述相同的序列号包括:利用所述多个有效ECHO消息中的所述多个EC区块的子集重建所述目标交易。

20.如权利要求11所述的非暂态计算机可读存储介质,其中,发送至所述多个其他网络节点的表示所述网络节点已被恢复的所述消息包括:

所述多个有效ECHO消息中的一组签名,和

所述目标序列号。

21.一种系统,包括:

一个或多个计算机;以及

一个或多个计算机可读存储器,其耦接到所述一个或多个计算机并配置有可由所述一个或多个计算机执行指令,用于:

区块链网络的网络节点将状态请求消息广播至所述区块链网络中的多个其他网络节点,其中,所述网络节点待恢复目标交易具有目标序列号;

所述网络节点从所述多个其他网络节点接收多个状态回复消息,其中,所述多个状态回复消息各自包括序列号;

响应于确定各自包括相同的序列号的状态消息的数量超过了预设阈值,所述网络节点基于所述相同的序列号识别所述目标序列号;

所述网络节点向所述多个其他网络节点发送请求消息,其中,所述请求消息从所述多个其他网络节点分别请求ECHO消息,所述ECHO消息是由所述多个其他网络节点分别传输并用于在所述多个其他网络节点之间就所述目标交易具有所述目标序列号达成共识的消息,并且所述ECHO消息包括所述目标交易的一部分以及所述多个其他网络节点各自的签名;

所述网络节点从所述多个其他网络节点接收多个ECHO消息;

所述网络节点确定所述多个ECHO消息中的有效ECHO消息的数量,其中,所述有效ECHO消息各自包括所述目标序列号;

响应于确定所述有效ECHO消息的数量超过预设阈值,所述网络节点基于所述多个有效ECHO消息在所述网络节点处恢复所述目标交易具有所述相同的序列号;以及

所述网络节点向所述多个其他网络节点发送表示所述网络节点已被恢复的消息。

22.如权利要求21所述的系统,其中,所述多个网络节点包括主节点和一个或多个备节点。

23.如权利要求21所述的系统,其中,所述网络节点是主节点或备节点。

24.如权利要求21所述的系统,其中,所述请求消息包括所述目标序列号。

25.如权利要求24所述的系统,其中,所述计算机存储器还配置有可由所述一个或多个计算机执行的指令,用于:

所述多个其他网络节点中除所述网络节点外的每个其他网络节点在向所述网络节点发送所述ECHO消息之前,验证所述请求消息。

26.如权利要求21所述的系统,其中,所述计算机存储器还配置有可由所述一个或多个计算机执行的指令,用于的:

所述网络节点验证每个所述ECHO消息是否有效,其中,验证每个所述ECHO消息是否有

效包括:利用梅克尔树验证每个所述ECHO消息是否有效。

27.如权利要求26所述的系统,其中,所述验证每个所述ECHO消息是否有效还包括:验证所述ECHO消息中的所述签名是否有效。

28.如权利要求21所述的系统,其中,每个所述ECHO消息还包括与所述目标交易相关联的多个纠删码EC区块中的至少一个区块,所述多个EC区块是利用所述目标交易根据EC码生成的。

29.如权利要求21所述的系统,其中,基于所述多个有效ECHO消息在所述网络节点处恢复所述目标交易具有所述相同的序列号,包括:利用所述多个有效ECHO消息中的所述多个EC区块的子集重建所述目标交易。

30.如权利要求21所述的系统,其中,发送至所述多个其他网络节点的表示所述网络节点已被恢复的所述消息包括:

所述多个有效ECHO消息中的一组签名,和

所述目标序列号。

在分布式系统中进行网络节点的恢复处理

背景技术

[0001] 分布式账本系统(DLS),又称共识网络和/或区块链网络,使得参与的实体能够安全地且不可篡改地存储数据。在不引用任何特定用例的情况下,DLS通常被称为区块链网络。区块链网络的示例可以包括:公有区块链网络、私有区块链网络和联盟区块链网络。公有区块链网络向所有实体开放DLS的使用,并开放共识处理的参与。私有区块链网络针对特定实体提供,该实体集中控制读写权限。联盟区块链网络针对选择的实体组群提供,该实体组群控制共识处理,并包括访问控制层。

[0002] 共识机制是分布式区块链系统的主要组成部分。共识机制是计算机科学用在分布式处理或系统中就单一数据值达成共识的处理。共识机制是为了在涉及多个不可靠节点的网络中实现可靠性而设计的。解决该问题(所谓共识问题)在分布式计算及多代理系统中非常重要。

[0003] 区块链依赖于共识机制以在节点之间达成协议。区块链是在点对点(P2P)网络中由分布式计算机管理的去中心化数据库。每个节点维护一份账本的副本以防止单点故障(SPOF)。更新和验证同时反映于所有副本中。

[0004] 尽管可以使用多个现有技术在区块链系统的网络节点中执行共识,但用于执行共识的更有效的解决方案将是有优势的。

发明内容

[0005] 本申请的实施方式包括用于解决分布式系统(例如,区块链网络)中的共识问题的计算机实施方法。更具体地,本申请的实施方式涉及在分布式系统中对网络节点进行恢复处理。

[0006] 在一些实施方式中,动作包括:区块链网络的网络节点将状态请求消息广播至所述区块链网络中的多个其他网络节点,其中,所述网络节点待恢复目标交易具有目标序列号;所述网络节点从所述多个网络节点接收多个状态回复消息,其中,所述多个状态消息各自包括序列号;响应于确定各自包括相同的序列号的状态回复消息的数量超过了预设阈值,所述网络节点基于所述相同的序列号识别所述目标序列号;所述网络节点向所述多个网络节点发送请求消息,其中,所述请求消息从所述多个网络节点分别请求ECHO消息,所述ECHO消息是由所述多个网络节点分别传输并用于在所述多个网络节点之间就所述目标交易具有所述目标序列号达成共识的消息,并且所述ECHO消息包括所述目标交易的一部分以及所述多个网络节点各自的签名;所述网络节点从所述多个网络节点接收多个ECHO消息;所述网络节点确定所述多个ECHO消息中的有效ECHO消息的数量,其中,所述有效ECHO消息各自包括所述目标序列号;响应于确定所述有效ECHO消息的数量超过预设阈值,所述网络节点基于所述多个有效ECHO消息在所述网络节点处恢复所述目标交易具有所述相同的序列号;所述网络节点向所述多个网络节点发送表示所述网络节点已被恢复的消息。

[0007] 其他实施方式包括对应的系统、装置和计算机程序,该计算机程序被配置为实现该方法的动作并被编码在计算机存储设备上。

[0008] 这些和其他实施方式可各自可选地包括以下特征中的一个或多个：

[0009] 第一特征，可与以下任何特征组合，其中，所述多个网络节点包括主节点和一个或多个备节点。

[0010] 第二特征，可与以下任何特征组合，其中，所述网络节点是主节点或备节点。

[0011] 第三特征，可与以下任何特征组合，其中，所述请求消息包括所述目标序列号。

[0012] 第四特征，可与以下任何特征组合，其中，所述多个网络节点中除所述网络节点外的每个网络节点在向所述网络节点发送所述ECHO消息之前，验证所述请求消息。

[0013] 第五特征，可与以下任何特征组合，其中，所述网络节点利用梅克尔树验证每个所述ECHO消息是否有效。

[0014] 第六特征，可与以下任何特征组合，其中，所述网络节点还验证所述ECHO消息中的所述签名是否有效。

[0015] 第七特征，可与以下任何特征组合，其中，每个所述ECHO消息还包括与所述目标交易相关联的多个纠删码(erasure code，EC)区块中的至少一个区块，所述多个EC区块是利用所述目标交易根据EC码生成的。

[0016] 第八特征，可与以下任何特征组合，其中，所述网络节点利用所述多个有效ECHO消息中的所述多个EC区块的子集重建所述目标交易。

[0017] 第九特征，可与以下任何特征组合，其中，发送至所述多个网络节点的表示所述网络节点已被恢复的所述消息包括：所述多个有效ECHO消息中的一组签名，和所述目标序列号。

[0018] 本申请还提供了耦接到一个或多个处理器并且其上存储有指令的一个或多个非暂态计算机可读存储介质，当所述指令由一个或多个处理器执行时，所述指令促使一个或多个处理器按照本文提供的方法的实施方式进行操作。

[0019] 本申请还提供了用于实施本文提供的方法的系统。该系统包括一个或多个处理器以及耦接到一个或多个处理器并且其上存储有指令的计算机可读存储介质，当所述指令由一个或多个处理器执行该指令时，所述指令促使一个或多个处理器根据本文提供的方法的实施方式进行操作。

[0020] 本申请公开了改良的共识机制，包括用于以下的技术：在分布式系统内的网络节点之间达成共识，在分布式系统中实现主节点的变更，以及在分布式系统中实现网络节点的恢复处理。所述共识机制可在不同应用中实现各种优点。

[0021] 例如，下述共识处理包括许多特征，这些特征改良了区块链系统的操作并且帮助缓解了网络瓶颈。例如，所述共识处理包括根据纠删码(EC)将交易请求转换为多个EC区块，并向每个网络节点发送其中一个EC区块。所述EC区块的大小比初始交易请求的小。因此，向所述网络节点发送所述EC区块而不是完整交易请求，减小了在区块链网络中的网络节点间传输的数据区块的大小，从而节省了网络带宽并减少了网络负载。这进一步减少了写入网络节点的存储空间及从网络节点的存储空间读取的数据大小，从而减轻了网络节点的存储空间的负担并提高了整个区块链系统的效率。

[0022] 此外，本申请描述了时元改变处理，该处理包括将相应权重分配给所述共识处理的多个阶段，基于所述多个阶段的所述相应权重确定权重和，并基于所述权重和确定新主节点。基于所述权重和而不是轮询法(round robin method)的时元改变处理有助于及时选

择非错误新主节点。与所述轮询法不同，本申请中的时元改变处理依赖于所述权重和来选择该新主节点，这可以减少寻找无错误的该新主节点时的延时或延迟。此方法可以进一步提高该整个区块链系统在提供该区块链服务时的效率。

[0023] 另外，本申请讨论了恢复处理，所述恢复处理包括诸如由网络节点发送申请成为新主节点的状态请求消息、并从其他网络节点接收状态回复消息的操作。这些操作被执行使得所述错误网络节点的所述恢复处理不干扰其他非错误网络节点之间的所述共识处理的正常操作。通过降低该恢复处理的复杂性来恢复错误网络节点，这有助于节省计算和网络资源。

[0024] 应当确认的是，根据本申请的方法可以包括本文描述的方面和特征的任何组合。也即，根据本申请的方法不限于本文所详述的方面和特征的组合，还包括任何方面和特征的任何组合。

[0025] 以下在附图和描述中阐述了本申请的一个或多个实施方式的细节。根据说明书和附图以及权利要求，本申请的其他特征和优点将显而易见。

附图说明

[0026] 图1描绘了可用于执行本申请实施方式的环境的示例。

[0027] 图2描绘了根据本申请实施方式的概念性架构的示例。

[0028] 图3描绘了可根据本申请实施方式执行的共识处理的示例。

[0029] 图4描绘了可根据本申请实施方式执行的共识处理的示例。

[0030] 图5描绘了根据本申请实施方式的哈希树的示例。

[0031] 图6描绘了根据本申请实施方式的分布式系统的网络节点之间通信的消息的示例。

[0032] 图7描绘了根据本申请实施方式的在分布式系统中进行主节点变更的处理的示例。

[0033] 图8描绘了根据本申请实施方式的在分布式系统中进行主节点变更处理的示例。

[0034] 图9描绘了根据本申请实施方式的分布式系统的网络节点之间通信的消息的示例。

[0035] 图10描绘了根据本申请实施方式的在分布式系统中进行网络节点恢复处理的处理的示例。

[0036] 图11描绘了根据本申请实施方式的在分布式系统中进行网络节点恢复处理的处理的示例。

[0037] 图12描绘了根据本申请实施方式的分布式系统的网络节点内通信的消息的示例。

[0038] 图13描绘了根据本申请实施方式示出了共识装置的模块的示例图。

[0039] 图14描绘了根据本申请实施方式示出了共识装置的模块的示例图。

[0040] 图15描绘了根据本申请实施方式示出了主节点变更装置的模块的示例图。

[0041] 图16描绘了根据本申请实施方式示出了主节点变更装置的模块的示例图。

[0042] 图17描绘了根据本申请实施方式示出了恢复装置的模块的示例图。

[0043] 各附图中相同的参考符号表示相同的元件。

具体实施方式

[0044] 本申请的实施方式包括用于解决分布式系统(例如区块链网络)中的共识问题的计算机实施方法。更具体地,本申请的实施方式涉及在分布式系统中对网络节点进行恢复处理。

[0045] 在一些实施方式中,动作包括:区块链网络的网络节点将状态请求消息广播至所述区块链网络中的多个其他网络节点,其中,所述网络节点待恢复目标交易具有目标序列号;所述网络节点从所述多个网络节点接收多个状态回复消息,其中,所述多个状态消息各自包括序列号;响应于确定各自包括相同的序列号的状态回复消息的数量超过了预设阈值,所述网络节点基于所述相同的序列号识别所述目标序列号;所述网络节点向所述多个网络节点发送请求消息,其中,所述请求消息从所述多个网络节点分别请求ECHO消息,所述ECHO消息是由所述多个网络节点分别传输并用于在所述多个网络节点之间就所述目标交易具有所述目标序列号达成共识的消息,并且所述ECHO消息包括所述目标交易的部分以及所述多个网络节点各自的签名;所述网络节点从所述多个网络节点接收多个ECHO消息;所述网络节点确定所述多个ECHO消息中的有效ECHO消息的数量,其中,所述有效ECHO消息各自包括所述目标序列号;响应于确定所述有效ECHO消息的数量超过预设阈值,所述网络节点基于所述多个有效ECHO消息在所述网络节点处恢复所述目标交易具有所述相同的序列号;所述网络节点向所述多个网络节点发送表示所述网络节点已被恢复的消息。

[0046] 为本申请实施方式提供进一步的背景,并且如上所述,分布式账本系统(DLS),又可称为共识网络(例如,由点对点节点组成)或区块链网络,使得参与的实体能够安全地且不可篡改地进行交易和存储数据。在不参考任何特定用例的情况下,本文中所用术语“区块链”通常是指DLS。如上所述,区块链网络的示例可以包括:公有区块链网络、私有区块链网络和联盟区块链网络。

[0047] 区块链是一种数据结构,其以一种允许验证未来交易以与存储在链中的所有先前交易一致的方式存储交易。区块链包括一个或多个区块。链中的每个区块通过包含在该链中紧邻其之前的前一区块的加密哈希值(cryptographic hash)来连接到该前一区块。每个区块还包括时间戳,其自身的加密哈希值以及一个或多个交易。已经被区块链网络中的节点验证过的交易经哈希处理并编入梅克尔(Merkle)树中。梅克尔树是一种数据结构,其中在该树的叶节点处的数据是经哈希处理的,并且在该树的每个分支中的所有哈希值在该分支的根处级联(concatenated)。此处理沿树持续至整个树的根,该整个树的根部存储了代表树中所有数据的哈希值。通过确定级联的哈希值是否与树的结构一致,可快速验证声称是存储在该树中的交易的哈希值。

[0048] 尽管区块链是用于存储交易的数据结构,但区块链网络是管理、更新以及维护一个或多个区块链的计算节点的网络。如上所述,区块链网络可被提供为公有区块链网络、私有区块链网络或联盟区块链网络。

[0049] 在公有区块链网络中,共识处理由共识网络中的节点控制。例如,数百、数千甚至数百万个实体可以协同运作公有区块链网络,每个实体操作该公有区块链网络中的至少一个节点。因此,公有区块链网络可被认为是关于参与实体的公有网络。在一些示例中,大多数实体(节点)必须对每个区块签名以使该区块有效,并被添加至区块链网络的区块链(分布式账本)中。示例性公有区块链网络包括利用分布式账本(称为区块链)的特定点对点支

付网络。然而，如上所述，术语“区块链”通常用于指代分布式账本，而不特指任何特定的区块链网络。

[0050] 通常，公有区块链网络支持公开交易。公开交易被公有区块链网络内的所有节点共享，并存储于全局区块链中。全局区块链是跨所有节点复制的区块链。也即，所有节点相对于全局区块链都处于完全共识状态。为了达成共识(例如，同意向区块链添加区块)，在该公有区块链网络内实施共识协议。共识协议的示例包括，但不限于，工作证明(POW)、权益证明(POS)以及权限证明(POA)。本文进一步引用POW作为非限制性示例。

[0051] 通常，为特定实体提供私有区块链网络，该特定实体集中控制读写权限。实体控制哪些节点能参与到该区块链网络中。因此，私有区块链网络通常被称为权限网络，其限制允许谁参与网络，以及它们的参与级别(例如，仅在特定交易中)。可以使用各种类型的访问控制机制(例如，现有参与者对添加新实体进行投票，管理机构可以控制权限)。

[0052] 通常，联盟区块链网络在参与实体之间是私有的。在联盟区块链网络中，该共识处理由一组经授权的节点控制，一个或多个节点由相应实体(例如，金融机构、保险公司)操作。例如由十(10)个实体(例如，金融机构、保险公司)组成的联盟可以运作联盟区块链网络，每个实体操作该联盟区块链网络中的至少一个节点。因此，联盟区块链网络可被认为是对于参与实体的私有网络。在一些示例中，每个实体(节点)必须对每个区块签名，以使该区块有效并被添加到该区块链中。在一些示例中，至少实体(节点)的子集(例如，至少7个实体)必须对每个区块签名，以使该区块有效并被添加到区块链中。

[0053] 本文参考联盟区块链网络进一步详尽描述了本申请的实施方式。然而，可以预期，本申请的实施方式可以在任何适当类型的区块链网络中实现。

[0054] 鉴于以上背景，本文进一步详尽描述了本申请的实施方式。更具体地，且如上所述，本申请的实施方式涉及在分布式系统中的网络节点执行恢复处理。

[0055] 区块链是防篡改的、共享的数字账本，其在公有或私有点对点网络中记录交易。该账本被分发至网络中的所有成员节点，且网络中发生的资产交易历史被永久地记录于区块中。

[0056] 共识机制保证分布式区块链网络中的所有网络节点以相同顺序执行交易并随后写入该同一账本中。共识模型旨在解决的一个问题是克服拜占庭(Byzantine)错误。在拜占庭错误中，分布式区块链网络中诸如服务器或网络节点的组件可不一致地出现错误并且对错误检测系统起作用，向不同的观察者显示不同症状。其他网络节点很难宣称它错误并将其从网络中关闭，因为它们需要首先就哪个网络节点最先出现错误达成共识。

[0057] 在分布式系统的上下文中，拜占庭容错(BFT)是分布式计算机网络根据需要发挥作用并正确地达到足够共识的能力，尽管系统的恶意组件(即，区块链网络中的网络节点)出错或将错误信息传播给其他节点。目的是通过减轻这些恶意节点对网络的正确功能和由系统中的诚实节点达成正确共识的影响来防止灾难性系统错误。

[0058] 然而，现存BFT机制已被证实在许多方面低效。例如，当尝试克服拜占庭错误时，现有的BFT机制增加了实施分布式区块链网络的复杂性，从而增加了分布式区块链网络中的网络节点之间的通信的延时。实用的拜占庭容错(Practical Byzantine Fault Tolerance，PBFT)是旨在改进现有的BFT共识机制的优化方法之一。PBFT模型侧重于提供实用的拜占庭状态机复制，其通过假设存在独立节点错误和由特定独立节点传播的操作消息

来容忍拜占庭错误(恶意节点)。

[0059] 在PBFT模型中,所有节点都按顺序排列,其中一个节点作为主节点(领导者),而其他节点被称为备节点。系统内的所有节点彼此通信,且目标是使大多数诚实节点就系统状态达成一致。节点彼此通信,且不仅需要证明消息来自特定的对等节点,还需要证明消息在传输中未被修改。

[0060] 为了使PBFT模型起作用,假设在给定的漏洞窗口中,网络中的恶意节点数量不能同时等于或超过系统中的全部节点的1/3。系统中的节点越多,那么在数学上越不太可能使全部节点的接近1/3的数量成为恶意。只要最多(n-1)/3个节点同时是恶意或错误,该算法就有效地提供了活跃性和安全性,其中n代表总节点。

[0061] 每轮PBFT共识(称为视图)包括四个阶段:

[0062] (1)客户端向领导节点发送请求以调用服务操作;

[0063] (2)领导节点向备节点组播该请求;

[0064] (3)节点执行该请求并然后发送回复至客户端;以及

[0065] (4)客户端等待来自不同节点的具有相同结果的f+1(f代表可能错误的节点的最大数)个回复。

[0066] 最终结果是:所有诚实节点就记录顺序达成一致,并且他们或者接受或者拒绝该结果。领导节点在每个视图期间以轮询方案改变,并且如果领导节点没有组播请求的情况经过了特定时间量,则领导节点甚至可以替换为名为视图改变的协议。大多数诚实节点也可以决定领导者是否错误,并以队列中的下一领导者作为替代移除他们。

[0067] 然而,PBFT共识机制存在一些局限性。例如,因为节点间需要繁琐的通信量,PBFT模型可以在具有相对小尺寸的共识组的传统形式中很好地工作。在网络节点间传输庞大的区块数据导致了网络负载问题并导致网络瓶颈。此外,使用方法认证码(MAC)作为PBFT模型中认证消息的格式,对于诸如加密货币网络和MAC之类的大共识组中的节点之间所需的通信量而言可能是低效的。可能存在无法向第三方证明消息的真实性的固有缺陷。

[0068] 此外,若在利用PBFT使用的轮询法变更领导节点时遇到连续的恶意节点,将因延时或延迟了寻找诚实的领导节点而影响区块链服务。例如,当选择第一网络节点作为新的领导节点时,第一网络节点可能是恶意节点,因此不能被选为新的领导节点。在轮询法中,可以选择队列中的第二网络节点作为新的领导节点。然而,若第二网络节点也是恶意节点,将验证队列中的另一网络节点是否适合作为领导节点。继续这一处理直到识别出新的领导节点是诚实的。领导节点的这种频繁变更在区块链服务中引入了显著的延迟。

[0069] 此外,区块链网络中的网络节点可能在任意时间经历拜占庭错误或崩溃错误。例如,网络节点可能被恶意网络攻击者破坏并且行为不当。如果受到攻击的网络节点没有及时恢复,则恶意网络攻击者可能在未被检测到的情况下破坏超过1/3的网络节点,从而危及区块链网络和服务。

[0070] 为解决上述与现存BFT共识机制和PBFT共识机制相关的问题和关注,本申请公开了改良的共识机制,包括:用于在分布式系统内的网络节点之间达成共识的技术、用于在分布式系统中执行主节点变更的技术、以及用于对分布式系统中的网络节点执行恢复处理的技术。所描述的共识机制可以在不同应用中展现各种优点。

[0071] 例如,以下所讨论的共识处理包括许多特征,这些特征改进了区块链系统的操作

且帮助缓解了网络瓶颈，例如，所描述的共识处理包括根据纠删码(EC)将交易请求转换为多个EC区块，并向每个网络节点发送其中一个区块。EC区块的大小比原始交易请求的小。因此，向网络节点发送EC区块而不是完整交易请求，减少了在区块链网络中的网络节点之间传输的数据区块的大小，从而节省了网络带宽并减少了网络负载。这进一步减少写入并从网络节点的存储器空间读取的数据大小，从而减轻了网络节点的存储器空间的负担并提高了整个区块链系统的效率。

[0072] 此外，本申请描述了时元改变处理，该处理包括：将相应权重分配给共识处理的多个阶段、基于所述多个阶段的相应权重确定权重和、以及基于权重和确定新的主节点。基于权重和而不是轮询法(round robin method)进行的时元改变处理可以有助于及时选择无错误的新主节点。主节点可以是领导节点，其具有在包括该主节点的多个网络节点之间发起一轮共识处理的权限。区块链网络中的其他网络节点可以称为备节点。时元改变处理可以帮助解决轮询法的以下问题：当新主节点的队列中的多个网络节点出现错误时，该轮询方法会导致主节点频繁变更。与轮询法不同，本申请中的时元改变处理依赖于权重和来选择新主节点，这可以减少寻找无错误的新主节点的延时或延迟。此方法可以进一步提高整个区块链系统在提供区块链服务时的效率。

[0073] 另外，本申请讨论了恢复处理，该恢复处理包括诸如由适用于新主节点的网络节点发送状态请求消息以及从其他网络节点接收状态应答消息的操作。执行这些操作使得错误网络节点的恢复处理不干扰其他非错误网络节点之间的共识处理的正常操作。通过降低恢复处理的复杂性，这有助于节省用于恢复错误网络节点的计算和网络资源。

[0074] 图1描绘了可用于执行本申请实施方式的环境100的示例。在一些示例中，环境100使得实体能够参与到联盟区块链网络102中。环境100包括计算设备或系统106、108和网络110。在一些示例中，网络110包括局域网(LAN)、广域网(WAN)、因特网或其组合，并且连接网络站点、用户设备(例如，计算设备)以及后台系统。在一些示例中，可以通过有线和/或无线通信连接来访问网络110。在一些示例中，网络110使得与联盟区块链网络102通信以及在联盟区块链网络102内部进行通信成为可能。通常，网络110代表一个或多个通信网络。在一些情况下，计算设备106、108可以是云计算系统(未示出)的节点，或每个计算设备106、108可以是单独的云计算系统，该系统包括通过网络互联的多个计算机并用作分布式处理系统。

[0075] 在所描述的示例中，计算系统106、108可以各自包括能够作为节点参与到联盟区块链网络102中的任何合适的计算机系统。示例性计算设备包括，但不限于服务器、台式计算机、膝上型计算机、平板计算设备以及智能电话。在一些示例中，计算系统106、108承载一个或多个由计算机实施的服务用于与联盟区块链网络102交互。例如，计算系统106可以承载第一实体(例如，用户A)的由计算机实施的、例如交易管理系统的服务，第一实体使用该交易管理系统管理其与一个或更多其他实体(例如，其他用户)的交易。计算系统108可以承载第二实体(例如，用户B)的由计算机实施的、例如交易管理系统的服务，第二实体使用该交易管理系统管理其与一个或更多其他实体(例如，其他用户)的交易。在图1的示例中，联盟区块链网络102被表示为节点的点对点网络，且计算系统106、108分别提供参与到联盟区块链网络102中的第一实体和第二实体的节点。

[0076] 图2描绘了根据本申请实施方式的概念性架构200的示例。概念性架构200的示例包括分别对应于参与者A、参与者B和参与者C的参与者系统202、204和206。每个参与者(例

如,用户、企业)参与到包括多个节点214的、被提供为点对点网络的区块链网络212中,至少一些节点将信息不可篡改地记录在区块链216中。如本文中进一步详述,尽管在区块链网络212中示意性地描绘了单个区块链216,但是在区块链网络212上提供并维护了区块链216的多个副本。

[0077] 在所描述的示例中,每个参与者系统202、204、206分别由或代表参与者A、参与者B和参与者C提供,并且在区块链网络内作为各自的节点214发挥作用。如本文所使用的,节点通常是指连接至区块链网络212的独立系统(例如,计算机、服务器),且使得相应的参与者能够参与到区块链网络中。在图2的示例中,参与者对应于每个节点214。然而,可以预测,一个参与者可以操作区块链网络212内的多个节点214,和/或多个参与者可以共享一个节点214。在一些示例中,参与者系统202、204、206使用协议(例如,超文本传输安全协议(HTTPS))和/或使用远程过程调用(RPC)与区块链网络212通信或通过区块链网络212进行通信。

[0078] 节点214可以在区块链网络212中具有不同的参与程度。例如,一些节点214可以参与共识处理(例如,作为将区块添加至区块链216的监视节点),而其他节点214不参与此共识处理。作为另一示例,一些节点214存储区块链216的完整的副本,而其他节点214仅存储区块链216的一部分的副本。例如,数据访问权限可以限制相应参与者在其相应系统内存储的区块链数据。在图2的示例中,参与者系统202、204、206存储区块链216相应的完整副本216'、216"、216"'。

[0079] 区块链(例如,图2中的区块链216)由一系列区块构成,每个区块存储数据。数据的示例包括表示两个或更多个参与者之间的交易的交易数据。虽然本文通过非限制性示例使用了交易,但是可以预期,任何适当的数据都可以存储在区块链中(例如,文档、图像、视频、音频)。交易的示例可以包括,但不限于有价物的交换(例如,财产、产品、服务和货币)。交易数据被不可篡改地存储在区块链中。也即,交易数据不能被改变。

[0080] 在将交易数据存储于区块中之前,对交易数据进行哈希处理。哈希处理是将交易数据(作为字符串数据提供)转换为固定长度哈希值(也作为字符串数据提供)的处理。无法通过将哈希值进行去哈希处理来获得交易数据。哈希处理确保即使交易数据中的轻微改变也会导致完全不同的哈希值。此外,如上所述,哈希值具有固定长度。也即,无论交易数据的大小如何,哈希值的长度都是固定的。哈希处理包括通过哈希函数处理交易数据以生成哈希值。哈希函数的示例包括但不限于输出256位哈希值的安全哈希算法(SHA)-256。

[0081] 多个交易的交易数据被哈希处理并存储在区块中。例如,提供两个交易的哈希值,并对它们自身进行哈希处理以提供另一个哈希值。重复该处理,直到为所有要存储在区块中的交易提供单个哈希值为止。该哈希值被称为梅克尔根哈希,并存储于区块的头中。任何交易中的变更都将导致其哈希值发生变化,并最终导致梅克尔根哈希值发生变化。

[0082] 通过共识协议将区块添加至区块链。区块链网络内的多个节点参与共识协议,并竞相将区块添加到区块链中。这样的节点被称为矿工(或监视节点)。上文介绍的POW用作非限制性示例。

[0083] 矿工节点执行共识处理以将交易添加至区块链中。虽然多个矿工节点参与共识处理,但只有一个矿工节点可以将区块写入区块链。也即,矿工节点在共识处理中竞相将它们的区块添加至区块链中。具体而言,矿工节点周期性地从交易池中收集待处理的交易(例

如,最多达到区块中可包括的交易数量的预定限制,如果有的话)。交易池包括来自区块链网络中的参与者的交易消息。矿工节点构建区块,并将交易添加至区块中。在将交易添加至区块之前,矿工节点检查是否有任一交易已经包含在区块链的区块中。如果该交易已经包含在另一区块中,则该交易将被丢弃。

[0084] 矿工节点生成区块头,对该区块中的所有交易进行哈希处理,并将所得的哈希值成对组合以生成进一步的哈希值,直到为区块中的所有交易提供单一哈希值(梅克尔根哈希值)。将该哈希值添加至区块头中。矿工还确定区块链中最近的区块(即,添加至区块链中的最后的区块)的哈希值。矿工节点还向区块头添加随机数(nonce value)和时间戳。在挖矿处理中,矿工节点尝试查找满足所需参数的哈希值。矿工节点不断变更随机数直到找到满足所需参数的哈希值。

[0085] 区块链网络中的每个矿工都试图找到满足所需参数的哈希值,并以这种方式彼此竞争。最终,矿工节点之一找到满足所需参数的哈希值,并将其通告给区块链网络中的所有其他矿工节点。其他矿工节点验证哈希值,若确定是正确的,则验证区块中的每个交易,接受该区块,并将该区块附加至它们的区块链的副本中。以此种方式,区块链的全局状态在区块链网络内的所有矿工节点上是一致的。上述处理是POW共识协议。

[0086] 参考图2提供了非限制性示例。在此示例中,参与者A想要向参与者B发送一定数额的资金。参与者A生成交易消息(例如,包括发送方、接收方和资金值字段),并将交易消息发送至区块链网络,从而将交易消息添加到了交易池中。区块链网络中的每个矿工节点创建一个区块,并从交易池中取出所有交易(例如,最多达到可添加到区块的交易数量的预定限制,如果有的话),并将交易添加至区块中。以此种方式,参与者A发布的交易被添加至矿工节点的区块中。

[0087] 在一些区块链网络中,用密码学来维护交易的隐私。例如,若两个节点想保持交易隐私,使得区块链网络中的其他节点不能看出交易的细节,则这两个节点可以对交易数据进行加密处理。加密方法的示例包括,但不限于对称加密和非对称加密。对称加密是指使用单一密钥进行加密(从明文生成密文)和解密(从密文中生成明文)的加密处理。在对称加密中,同一密钥可用于多个节点,因此每个节点都可以对交易数据进行加密/解密。

[0088] 非对称加密使用密钥对,每个密钥对包括私钥和公钥,私钥仅对相应节点是已知的,而公钥对于区块链网络中的任何或所有其他节点是已知的。一个节点可以使用另一节点的公钥来加密数据,并且该加密的数据可以使用所述另一个节点的私钥解密。例如,再次参考图2,参与者A可以使用参与者B的公钥来加密数据,并将该加密数据发送给参与者B。参与者B可以使用它的私钥来解密该加密的数据(密文)并提取原始数据(明文)。使用节点的公钥加密的消息只能使用节点的私钥解密。

[0089] 非对称加密用于提供数字签名,这使得交易中的参与者能够确认交易中的其他参与者以及交易的有效性。例如,节点可以对消息进行数字签名,而另一个节点可以基于参与者A的数字签名来确认该消息是由该节点发送的。数字签名还可以用于确保消息在传输中不被篡改。例如,再次参考图2,参与者A要向参与者B发送消息。参与者A生成该消息的哈希值,然后使用其私钥加密该哈希值以提供为该加密哈希值的数字签名。参与者A将该数字签名附加到消息上,并将该具有数字签名的消息发送至参与者B。参与者B使用参与者A的公钥解密该数字签名,并提取哈希值。参与者B对该消息进行哈希处理并比较哈希值。若哈希值

相同，参与者B可以确认该消息确实是来自参与者A，且未被篡改。

[0090] 图3描绘了可以根据本申请实施方式执行的、用于在分布式系统(例如，区块链网络102和212)的网络节点(例如，节点214)之间达成共识的处理300的示例。具体地，图3示出了根据本申请示例性实施例呈现的用于在正常情况下达成共识的方法300的图。如图3所示，共识处理300包括下述三个阶段或阶段310、阶段320和阶段330。

[0091] 在共识处理300的第一阶段310，区块链网络中的主节点(或领导节点)向其他网络节点(即，备节点)发送第一消息。该第一消息表示主节点正在发起共识处理。例如，如图3中所示的，主节点R_0向区块链网络中的其他网络节点R_1、R_2、R_3发送INITIAL(初始)消息。注意，仅为了说明目的，处理300被示为包括四个网络节点R_0、R_1、R_2和R_3，处理300可以包括任何合适数量的网络节点。第一阶段和INITIAL消息的格式将在下文参考图4-6更详细地讨论。

[0092] 在共识处理300的第二阶段320，每个备节点接收由主节点发送的第一消息，响应于该第一消息准备第二消息，以及将该第二消息组播至其他网络节点，该第二消息表示备节点已经接收到来自主节点的第一消息，并正在发送回复以响应第一消息。例如，如图3所示的，备节点R_1接收由主节点R_0发送的INITIAL消息，并以ECHO消息作为第二消息的示例回复主节点R_0。同时，备节点R_1还将ECHO消息组播至其他备节点，例如，备节点R_2和R_3。类似地，备节点R_2和R_3各自将ECHO消息组播至包括主节点R_0的其他网络节点。

[0093] 例如，当诸如主节点或备节点的网络节点从其他网络节点接收ECHO消息时，该网络节点可以验证ECHO消息中的信息。第二阶段和ECHO消息的格式将在下文参考图4-6更详细地讨论。

[0094] 在共识处理300的第三阶段330，每个网络节点向其他网络节点组播第三消息。第三消息表示网络节点已经接受预设数量的第二消息。在一些实施方式中，该第三消息可以表示网络节点已经准备好执行交易。在一些实施方式中，该第三消息可以表示已经在网络节点成功重建了交易。例如，如图3所示，主节点R_0将ACCEPT消息组播至备节点R_1、R_2和R_3。类似地，备节点R_1、R_2和R_3各自将ACCEPT消息组播至其他网络节点。在本申请的一些实施方式中，在组播ACCEPT消息之前，网络节点根据纠删码(EC)以及ECHO消息中的信息是否是第二阶段接收的信息，来确定是否发送ACCEPT消息。第三阶段、EC码和ACCEPT消息的格式将在下文参考图4-6更详细地讨论。

[0095] 当网络节点从其他网络节点接收到足够的ACCEPT消息时，网络节点确定已经达成共识。例如，若主节点R_0或备节点R_1、R_2或R_3接收到法定数量(quorum，例如，$2f+1$，其中f代表错误网络节点的数量)的ACCEPT消息，则在网络节点中自动达成共识。

[0096] 图4描绘了可以根据本申请实施方式执行的、用于在分布式系统(例如，区块链网络102或212)的网络节点(例如，节点214或节点R_0、R_1、R_2和R_3)之间达成共识的处理400的示例。在一些实施方式中，处理400可以利用一个或多个计算设备执行一个或多个计算机可执行程序来进行。为了清楚地呈现，以下描述总体上描述了在本说明书其他附图的背景中的方法400。应当理解，方法400可以由诸如任何合适的系统、环境、软件和硬件，或系统、环境、软件和硬件的组合合理进行。在一些实施方式中，方法400的各步骤可以并行、组合、循环或以任何顺序进行。

[0097] 首先，处理400可以结合如图1-3所示的系统100-300进行。在本申请的一些实施方式中，区块链网络102和/或212包括主节点404和一个或多个备节点406。区块链网络102和/

或212与计算系统106和/或108通信，例如，客户端节点402通过网络110提供区块链服务。客户端节点402、主节点404和备节点406中的每个可以是专用计算机或配置成执行本文所讨论的处理的其他数据处理装置。例如，客户端节点402也可以指与区块链网络交互的客户端终端或客户端设备。客户端节点402可以安装诸如用于访问区块链网络并与区块链网络通信的、与区块链网络连接的客户端应用或客户端软件开发套装(SDK)。主节点404和一个或多个备节点406也可以被称为共识节点或网络节点，其在区块链网络中达成共识并不可篡改地记录信息。

[0098] 处理400始于408，其中客户端节点402生成交易请求。在本申请的一些实施方式中，交易请求可以包括请求从区块链网络102和/或212请求区块链服务。

[0099] 在410，客户端节点402将交易请求组播至区块链网络102和/或212的主节点404。在本申请的一些实施方式中，主节点404给交易请求分配序列号，以在从客户端节点402接收到交易请求之后跟踪交易请求。

[0100] 在412，主节点402在从客户端节点402接收交易请求之后生成多个EC区块。在本申请的一些实施方式中，主节点404利用交易请求根据EC码生成多个EC区块。例如，参考图5，主节点404在交易请求502上应用EC码504，并使用EC码504将交易请求502转换为EC消息506。EC码504是在比特纠删(bit erasures)的假设下的前向纠错(FEC)码。EC码504将初始交易请求502转换为更长的EC消息506，使得可以从EC消息506的部分或片段中恢复初始交易请求502。

[0101] 在本申请的一些实施方式中，EC码504是近乎最优的纠删码，诸如，旋风码(Tornado code)或低密度奇偶校验码。在本申请的可选实施方式中，EC码504是近乎最优的喷泉码(fountain code)，诸如喷泉码、在线码、卢比变换码(Luby transform(LT)code)或速龙码(raptor code)。在本申请的可选实施方式中，EC码504是最优纠删码，诸如，奇偶校验码、校验档案码(Parchive code)、里德-所罗门码(Reed-Solomon code)或再生成码(regenerating code)。在本申请的一些实施方式中，EC码504可以是任何合适类型的纠删码。

[0102] 在将交易请求502转换为EC消息506之后，主节点404利用EC消息506生成多个EC区块508。例如，如图5所示，主节点404通过划分EC消息506来生成EC区块A、EC区块B、EC区块C、EC区块D这四个EC区块508。注意，为作说明目的，图5中示出的EC区块508包括四个区块，EC区块508可以生成为包括任何合适数量的EC区块508。EC区块508将通过INITIAL消息被发送至相应的备节点406。

[0103] 在本申请的一些实施方式中，EC区块508具有相同大小。然而，在可选实施方式中，EC区块508可以具有彼此不同的大小。

[0104] 在本申请的一些实施方式中，主节点404利用EC区块508生成哈希树500(如梅克尔树)。哈希树500包括标记有数据区块的哈希值的多个叶节点和标记有子节点标记的加密哈希值的多个非叶节点。例如，如图5所示，哈希树500被配置为包括四个叶节点510：哈希值A、哈希值B、哈希值C和哈希值D，其被生成为它们相应EC区块508的加密哈希值；四个非叶节点512被生成为它们相应子节点510级联的哈希值，以及一个非叶节点514被生成为其子节点512的哈希值且是哈希树500的根哈希值。

[0105] 哈希树500允许对大型数据结构的内容进行有效且安全的验证。哈希树500可以用

于验证在计算机内和计算机之间存储、处理和传输的任何类型的数据。它们可以帮助确保从P2P网络中其他对等点接收的数据区块是未受损且未被篡改接收的，并且甚至检查其他对等点没有发送虚假区块。下面将参考共识处理400的如下步骤以更详细地讨论利用哈希树500验证数据区块。

[0106] 返回参考图4，主节点404在生成EC区块508和哈希树500后生成第一消息(例如，INITIAL消息)。第一消息表示主节点正在发起共识处理。在一些实施方式中，作为第一消息的示例的INITIAL消息，是利用EC区块508和哈希树500中的信息生成的。在本申请的一些实施方式中，参考图6，INITIAL消息具有为<epoch,tx_root_hash,ec_block_hash,ec_block,seq,j>的格式，其中“epoch”代表正在发送的消息的一轮共识，“tx_root_hash”代表哈希树500中的根哈希值514，“ec_block_hash”代表哈希树500中的哈希值510和/或512，“ec_block”代表哈希树500中的EC区块508，“seq”代表与交易请求502关联的序列号，以及“j”代表生成并发送INITIAL消息的网络节点。在一些实施方式中，INITIAL消息可以具有不同格式，例如，包括附加或不同的字段。

[0107] 返回参考图4，在416，在共识处理的第一阶段，主节点404将INITIAL消息组播至其他网络节点(例如，备节点406)。在一些实施方式中，被发送至备节点406的INITIAL消息具有<epoch,tx_root_hash,ec_block_hash,ec_block,seq,j>的格式。例如，主节点404可以向第一备节点406发送第一INITIAL消息<epoch 1,Hash ABCD,{Hash B,Hash C,Hash D},EC block A,1,0>，以及向第二备节点406发送第二INITIAL消息<epoch 1,Hash ABCD,{Hash A,Hash C,Hash D},EC block B,1,0>等。注意，INITIAL消息中的信息，例如“ec_block”可与“ec_block_hash”一起使用以重建哈希树500。例如，在第一INITIAL消息<epoch 1,Hash ABCD,{Hash B,Hash C,Hash D},EC block A,1,0>中，可以将EC区块508“EC block A”进行哈希处理以生成加密哈希值510“Hash A”，其进一步与其他哈希值510“{Hash B、Hash C、Hash D}”一起使用来重建哈希树500。参考下述共识处理的步骤，如下文所详述的，所重建的哈希树500将被用于验证ECHO消息。

[0108] 在418，每个备节点406在从主节点404接收INITIAL消息后，在共识处理的第二阶段生成第二消息(例如，ECHO消息)。该第二消息表示备节点已经从主节点接收到第一消息。响应于第一消息，第二消息作为回复被发送。在本申请的一些实施方式中，ECHO消息由备节点406生成，包括INITIAL消息或INITIAL消息的一部分以及备节点406的与INITIAL消息关联的签名。例如，备节点406可以通过利用私钥对INITIAL消息或INITIAL消息的摘要进行签名来生成签名。该私钥签名可由使用与该私钥配对的公钥的其他网络节点使用，来认证包括该私钥签名的ECHO消息。

[0109] 在本申请的一些实施方式中，参考图6，ECHO消息具有为<epoch,tx_root_hash,ec_block_hash,ec_block,seq,sign_proof,j>的格式，其中“epoch”代表正在发送的消息的一轮共识，“tx_root_hash”代表哈希树500中的根哈希值514，“ec_block_hash”代表哈希树500中的哈希值510和/或512，“ec_block”代表哈希树500中由相应备节点406接收的EC区块508，“seq”代表与交易请求502相关联的序列号，“sign proof”代表备节点406的与INITIAL消息关联的签名，以及“j”代表生成并发送ECHO消息的网络节点。在一些实施方式中，ECHO消息可以具有不同格式，例如，包括附加或不同的字段。

[0110] 返回参考图4，在420，备节点406向主节点404发送ECHO消息。在421，每个备节点

406向其他备节点406发送ECHO消息。在423，每个备节点406可以从其他备节点406接收ECHO消息。

[0111] 在422，主节点404验证由备节点406发出的ECHO消息。在本申请的一些实施方式中，主节点404根据哈希树500验证ECHO消息是否有效。例如，主节点可以从第一备节点406接收第一ECHO消息<epoch1,Hash ABCD,{Hash B,Hash C,Hash D},EC block A,1,1>。主节点404可以从消息中检索EC区块508“EC block A”并将其哈希处理以生成加密哈希值510“Hash A”。主节点404进一步使用生成的哈希值510“Hash A”与消息中的其他哈希值510“{Hash B、Hash C、Hash D}”以重建哈希树500。然后，主节点404确定重建的哈希树500的根哈希值514并将其与ECHO消息中的例如“Hash ABCD”的根哈希值514比较。若两个根哈希值514匹配，则主节点404确定ECHO消息有效。主节点404可以存储有效的ECHO消息并丢弃确定为无效ECHO消息。

[0112] 在424，主节点404确定有效ECHO消息的数量是否超过预设阈值。在本申请的一些实施方式中，主节点404确定有效ECHO消息的数量是否达到法定数量n-f或2f+1，其中n是网络节点的总数，且f是网络可容忍的错误节点的最大数量。

[0113] 在426，响应于确定有效ECHO消息的数量达到了法定数量，主节点404重建交易请求502。在本申请的一些实施方式中，主节点404根据EC码基于有效ECHO消息的至少一个子集来重建交易请求502。例如，主节点404可以检索有效ECHO消息的法定数量(例如，n-f或2f+1个)中的n-2f或f+1数量个EC区块508，并使用检索到的EC区块508以根据EC码504重建交易请求502。

[0114] 在428，在共识处理的第三阶段，响应于确定已经成功重建交易请求502，主节点404生成第三消息(例如，ACCEPT消息)。该第三消息表示网络节点已经接受预设数量的第二消息。在一些实施方式中，该第三消息可以表示网络节点已经准备好执行交易。在一些实施方式中，该第三消息可以表示已经成功地在网络节点中重建交易。例如，ACCEPT消息可以用于向其他网络节点表示已经成功重建交易请求502。若主节点404未能重建交易请求502，则主节点404可能不生成ACCEPT消息。

[0115] 在本申请的一些实施方式中，参考图6，ACCEPT消息具有为<epoch,tx_root_hash,seq,sign_proofs,j>的格式，其中“epoch”代表正在发送的消息的一轮共识，“tx_root_hash”代表哈希树500中的根哈希值514，“seq”代表与交易请求502关联的序列号，“sigh_proofs”代表有效ECHO消息中的一组签名，以及“j”代表生成并发送ACCEPT消息的网络节点。在一些实施方式中，ACCEPT消息可以具有不同的格式，例如，包括附加或不同的字段。

[0116] 返回参考图4，在430，主节点404向备节点406发送ACCEPT消息。

[0117] 类似于主节点404，每个备节点406可以重建交易请求，例如，通过执行与主节点404的步骤422-428类似的步骤。在432，响应于确定交易请求502已经由备节点406成功重建，每个备节点406生成ACCEPT消息。在一些实施方式中，主节点404和备节点406可以以并行方式执行步骤422-428，例如，如图3所示。

[0118] 在434，备节点406向主节点404发送ACCEPT消息。同时每个备节点406可以向其他备节点406发送ACCEPT消息。

[0119] 在436，响应于确定ACCEPT消息的数量超过了预设阈值，主节点404执行交易请求502。在本申请的一些实施方式中，主节点404确定接收的ACCEPT消息是否相同以及相同的

ACCEPT消息的数量是否达到法定数量(例如,2f+1)。若相同的ACCEPT消息的数量达到了法定数量,则主节点404确定已在所有网络节点之间达成共识,然后在本地执行交易请求502。在本申请的一些实施方式中,若主节点404确定相同的ACCPET消息的数量未达到法定数量,则主节点404确定在所有网络节点中尚未达成共识,然后避免执行交易请求502。

[0120] 在本申请的一些实施方式中,在执行交易请求502之前,每个备节点406可以执行如上述436的由主节点404执行的相同操作。若备节点406确定其接收的ACCEPT消息超过了预设阈值,则备节点406确定网络节点之间已经达成共识并在本地执行交易请求502。在本申请的一些实施方式中,若备节点406确定相同的ACCEPT消息的数量未达到法定数量,则备节点406确定尚未在所有网络节点之间达成共识,然后制止执行交易请求502。

[0121] 在438,主节点404在执行交易请求502之后向客户端节点402发送交易结果。已经在本地成功执行交易请求502的备节点406也可以将它们相应的交易结果发送至客户端节点402。

[0122] 如上所讨论的共识处理包括许多特征,这些特征改进了整个区块链系统的操作并帮助缓解了网络瓶颈。例如,本申请中的共识处理包括根据EC码利用交易请求生成多个EC区块,并向每个网络节点发送其中一个EC区块。该EC区块的大小比初始交易请求的小。因此,向网络节点发送EC区块而不是交易请求,减小了在区块链网络中的网络节点之间传输的数据区块的大小,从而节省了网络带宽并减少了网络负载。这进一步减少写入并从网络节点的存储器空间读取的数据大小,从而降低了网络节点的存储器空间的负载并提高了整个区块链系统的效率。

[0123] 在共识处理期间,备节点正在等待来自主节点的请求。然而,主节点可能遭遇拜占庭错误或崩溃错误,使得主节点不能在预设时间窗口内广播请求。当主节点没有组播请求的情况持续了特定时间量时,需要选择新主节点以防止备节点无限期地等待执行请求。

[0124] 图7描绘了可根据本申请实施方式执行的、用于在分布式系统(例如,区块链网络102或212)中进行主节点(例如,节点214或404)变更的处理700的示例。具体地,图7示出了根据本申请示例性实施例呈现的用于执行主节点变更的方法700的图。在一些实施方式中,主节点与包括将主节点作为领导者的共识处理的时元相关联。主节点的变更可导致时元的改变。

[0125] 在一些实施方式中,响应于确定当前时元的主节点需要被变更,区块链网络中的备节点向其他网络节点发送第一消息。该第一消息表示备节点希望成为新时元内的新主节点。例如,如图7所示,响应于备节点R_0确定当前主节点错误且需要进行时元改变,备节点R_0向区块链网络中的其他网络节点R_1、R_2和R_3发送EPOCH_CHANGE消息。EPOCH_CHANGE消息是表示备节点R_0申请成为新主节点的第一消息的示例。时元改变可以促使从具有当前主节点的当前时元变更为具有新主节点的新时元。注意仅用于说明目的,处理700被示出为结合四个网络节点实施。处理700可以结合任何合适数量的网络节点实施。

[0126] 然后,每个网络节点接收由备节点发送的第一消息,响应于该第一消息准备第二消息,并将第二消息组播至其他网络节点。例如,如图7所示,网络节点R_1接收由备节点R_0发送的EPOCH_CHANGE消息,并向备节点R_0回复NEW_EPOCH消息,该NEW_EPOCH消息表示了确认该备节点R_0可以成为新主节点。同时,网络节点R_1还将NEW_EPOCH消息组播至其他网络节点,例如,网络节点R_2和R_3。类似地,网络节点R_2和R_3各自将NEW_EPOCH消息组播至其他网络节

点。

[0127] 上述的时元改变处理、EPOCH_CHANGE消息的格式、以及NEW_EPOCH消息的格式将在下文参考图8-9详细讨论。

[0128] 图8描绘了可根据本申请实施方式执行的、用于在分布式系统(例如,区块链网络102或212)中进行主节点变更的处理800的示例。在一些实施方式中,示例性处理800可以使用一个或多个计算机设备执行一个或多个计算机可执行程序来执行。为了清楚地呈现,以下描述总体上描述了在本申请中的其他附图的执行背景中的方法800。可以理解的是,方法800可以由诸如任何合适的系统、环境、软件和硬件,或系统、环境、软件和硬件的组合合理进行。在一些实施方式中,方法800的各步骤可以并行、组合、循环或以任何顺序进行。

[0129] 处理800始于806,其中备节点802确定需要执行时元改变,本文讨论的时元改变促使从具有当前主节点的当前时元变更为具有新主节点的新时元。如上面参考图3-6所讨论的,示例性时元可以包括用于使用主节点在多个网络节点之间达成共识的共识处理(例如,共识处理300或400)。

[0130] 在本申请的一些实施方式中,响应于确定备节点802在未收到来自当前主节点的请求的情况持续了特定时间量之后仍在等待来自当前主节点的请求,备节点802确定需要进行时元改变。例如,当前主节点可能遭遇拜占庭错误或崩溃错误使得当前主节点不能在预设时间窗口内组播请求。因此,由超时触发时元改变,以防止备节点无限期地等待执行请求。本文讨论的时元改变处理,通过允许系统在主节点发生错误时取得进展来提供活跃性并减少了网络延迟。

[0131] 在808,备节点802确定备节点802在当前时元内共识处理的每个阶段相应的权重。在一些实施方式中,共识处理包括上文参考图3-6所述的三个阶段。权重是对备节点802成为新时元中的新主节点的资格的度量。

[0132] 在本申请的一些实施方式中,备节点802将备节点802在当前时元内的共识处理的第一阶段的权重确定为第一值。例如,若备节点802已经进入共识处理的第一阶段(例如,共识处理300的第一阶段310),则可以为备节点802分配10%的初始权重。在本申请的可选实施方式中,备节点802可以针对当前共识处理的第一阶段给备节点802分配任何合适的权重值。

[0133] 在本申请的一些实施方式中,备节点802基于Quorum验证处理确定备节点802在当前时元内共识处理的第二阶段(例如,共识处理300的第一阶段320)的权重。通过确定备节点802是否在共识处理的第二阶段从其他网络节点中接收到预设数量(例如,2f+1)的ECHO消息来执行Quorum验证处理。

[0134] 在本申请的一些实施方式中,若备节点802的Quorum验证失败(例如,备节点802接收的ECHO消息的数量低于预设阈值),则备节点802可以将备节点802在共识处理的第二阶段的权重确定为第一值。若备节点802通过了Quorum验证(例如,备节点802接收的ECHO消息的数量等于或超过了预设阈值),则备节点802可以将备节点802在共识处理的第二阶段的权重确定为第二值。在本申请的一些实施方式中,设定第二值大于第一值。例如,若备节点802的Quorum验证失败,则可以针对共识处理的第二阶段给备节点802分配为零的权重值。若备节点802通过了Quorum验证,则备节点802可以针对共识处理的第二阶段给备节点802分配45%的权重值。然而,在本申请的可选实施方式中,备节点802可以针对当前时元内共

识处理的第二阶段给备节点802分配任何合适数值。

[0135] 在本申请的一些实施方式中，Quorum验证进一步包括验证ECHO消息是否有效，该ECHO消息由备节点802在共识处理第二阶段期间从其他网络节点接收。例如，备节点802可以利用公钥来认证ECHO消息中的私钥签名，以确定ECHO消息是否有效。

[0136] 与确定第二阶段的权重类似，在一些实施方式中，备节点802基于Quorum验证处理，确定备节点802在当前时元内共识处理的第三阶段(例如，共识处理300的第三阶段330)的权重。通过确定备节点802在当前时元内共识处理的第三阶段中从其他网络节点中是否接收到预设数量(如，2f+1个)的接受消息来执行Quorum验证处理。来自其他网络节点的每个接受消息表示每个其他网络节点已接受预定数量的ECHO消息。该接受消息可以是，例如，上文参考共识处理300的第三阶段330描述的ACCEPT消息。

[0137] 在本申请的一些实施方式中，若备节点802的Quorum验证失败(例如，备节点802接收的ACCEPT消息的数量低于预设阈值)，则备节点802可以将备节点802在共识处理的第三阶段的权重确定为第一值。若备节点802通过了Quorum验证(例如，备节点802接收的ACCEPT消息的数量等于或超过了预设阈值)，则备节点802可以将备节点802在共识处理的第三阶段的权重确定为第二值。在一些实施方式中，设定第二值大于第一值。例如，若备节点802的Quorum验证失败，则备节点802可以针对共识处理的第二阶段给备节点802分配为零的权重值。若备节点802通过了Quorum验证，备节点802可以针对共识处理的第二阶段给备节点802分配为45％的权重值。然而，在本申请的可选实施方式中，备节点802可以针对当前时元内共识处理的第三阶段给备节点802分配任何合适的值。

[0138] 在810，在确定备节点802在当前时元内共识处理的各阶段的相应的权重后，备节点802基于相应的权重确定备节点802在共识处理的权重和。在本申请的一些实施方式中，权重和是备节点在当前时元内共识处理的每个阶段的相应权重的加和。例如，若备节点802已经确定备节点802在第一阶段的第一权重值为10％、备节点802在第二阶段的第二权重值为45％、备节点802在第三阶段的第三权重值为45％，则备节点802确定权重和为100％。作为另一示例，若备节点802已经确定备节点802在第一阶段的第一权重值为10％、备节点802在第二阶段的第二权重值为45％、备节点802在第三阶段的第三权重值为0，则备节点802确定权重和为55％。

[0139] 在812，若备节点802确定在810处确定的权重和达到或超过了预设阈值，则备节点802向其他网络节点804发送EPOCH_CHANGE消息。例如，若在810处确定的权重和达到100％，则备节点802可以向其他网络节点804发送EPOCH_CHANGE消息。EPOCH_CHANGE消息表示请求将具有当前主节点的当前时元变更为以备节点作为新主节点的新时元。

[0140] 在本申请的一些实施方式中，参考图9，EPOCH_CHANGE消息具有为<weight,epoch+1,ECHO{},ACCEPT{},j>的格式，其中“weight”代表先前在810针对共识处理确定的备节点802的权重和，“epoch+1”代表与新主节点关联的新一轮共识(即新时元)，“ECHO{}”代表在共识处理的第二阶段期间备节点802接收的一组ECHO消息，“ACCEPT{}”代表在共识处理的第三阶段期间备节点802接收的一组ACCEPT消息，以及“j”代表生成并发送EPOCH_CHANGE消息的网络节点(例如，备节点802)。在一些实施方式中，EPOCH_CHANGE消息可以具有不同的格式，例如通过包括附加或不同的字段。

[0141] 返回参考图8，在814，除备节点802之外的网络节点804验证由备节点802发送的

EPOCH_CHANGE消息。在一些实施方式中，每个网络节点804通过验证EPOCH_CHANGE消息中的权重和是否有效来验证EPOCH_CHANGE消息是否有效。在一些实施方式中，验证EPOCH_CHANGE消息中的权重和是否有效，包括：验证包含在EPOCH_CHANGE消息内的ECHO消息中的该组签名是否有效。例如，每个网络节点804可以使用公钥来认证包括在EPOCH_CHANGE消息内的ECHO消息中的该组私钥签名。

[0142] 在816，响应于验证由备节点802发送的EPOCH_CHANGE消息是有效的，每个网络节点804向备节点802发送NEW_EPOCH消息。NEW_EPOCH消息表示确认备节点将成为新主节点。例如，由网络节点804发送的NEW_EPOCH消息包括网络节点804确认备节点802将成为新时元内新主节点的指示。同时，每个网络节点804还向其他网络节点804发送NEW_EPOCH消息。

[0143] 参考图9，NEW_EPOCH消息生成具有<epoch+1,i,j,seq,ec_digest>的格式，其中“epoch+1”代表与新主节点关联的新一轮共识(即，新时元)，“i”代表新时元内的新主节点，“j”代表发送NEW_EPOCH消息的网络节点804，以及“ec_digest”代表EPOCH_CHANGE消息的摘要。在一些实施方式中，EPOCH_CHANGE消息的摘要包括EPOCH_CHANGE消息的哈希值。在一些实施方式中，NEW_EPOCH消息可以具有不同的格式，例如，包括附加或不同的字段。

[0144] 返回参考图8，在818，备节点802验证由网络节点804发送的NEW_EPOCH消息是否有效。在一些实施方式中，备节点802通过验证NEW_EPOCH消息中的EPOCH_CHANGE消息的摘要是否有效来验证NEW_EPOCH消息。由于摘要包括EPOCH_CHANGE消息的信息，所以摘要还包括EPOCH_CHANGE消息中的签名。备节点802可以通过验证EPOCH_CHANGE消息中的该组签名是否有效来验证EPOCH_CHANGE消息的摘要。

[0145] 在820，备节点802确定在818处确定的有效NEW_EPOCH消息的数量是否超过了预设阈值。在一些实施方式中，预设阈值是法定数量(例如，2f+1)

[0146] 在822，响应于确定所确定的有效NEW_EPOCH消息的数量超过了预设阈值，备节点802确定备节点802成为新时元中的新主节点。注意到每个网络节点804与备节点802相同地执行步骤818-822，并且网络节点804和备节点802可以以并行方式进行步骤818-822。例如，每个网络节点804可以验证从其他网络节点804发来的一组NEW_EPOCH消息，确定有效NEW_EPOCH消息的数量是否超过预设阈值，并确定新主节点。

[0147] 如上所讨论的时元改变处理(例如，处理700或800)包括许多特征，这些特征可以改进整个区块链系统的操作并帮助缓解网络瓶颈。例如，本申请中的时元改变处理包括：给共识处理的三个阶段分配相应的权重，基于三个阶段的相应权重确定权重和，并基于权重和确定新主节点。基于权重和而不是轮询法的时元改变处理可以有助于及时选择无错误的新主节点。当用于新主节点队列中的多个网络节点发生错误时，轮询法可能导致主节点的频繁变更。这将因延时或延迟找到非错误主节点而显著影响区块链服务。和轮询法不同，本申请中的时元改变处理依赖于权重和来选择新主节点，这可以减少寻找非错误新主节点的时间。这可以进一步提高整个区块链系统在提供区块链服务时的效率。

[0148] 在区块链网络运行期间，由于网络波动、突然断电、磁盘故障等原因，一些网络节点的执行速度可能落后于大多数网络节点的执行速度。在此情境下，系统中超过1/3的网络节点可能发生错误。若在系统的生命周期内，少于1/3的网络节点发生错误，则BFT可提供安全性和活跃度。然而这些保障对于长寿命系统来说是不够的，因为在上述情景中很可能超过上限。因此，需要一个恢复处理，使错误网络节点再次正常运行，并继续参与后续的共识

处理，以允许系统在其生命周期内容忍多于f个错误。此外，所描述的恢复处理可以恢复仍在进行共识处理(例如，共识处理300或400)的一个或多个网络节点，且无需等到所有网络节点之间达成共识。这样，所描述的恢复处理可以进一步减少系统延时并提高区块链网络的效率。

[0149] 图10描绘了可根据本申请实施方式执行的、用于执行分布式系统(例如，区块链网络102或212)的网络节点(例如，节点214或404)的恢复处理的处理1000的示例。具体地，图10示出了根据本申请示例性实施例呈现的用于执行网络节点的恢复处理的方法1000的图。如图10所示，处理1000包括几个阶段和时期。

[0150] 在第一阶段1010中，想要恢复目标交易具有目标序列号R_0的网络节点(例如，网络节点R_0)将状态请求消息(例如，QUERY_STATE消息)组播至其他网络节点，以表示网络节点待恢复。状态请求消息可以包括网络节点R_0想要恢复的目标序列号。在第二阶段1020中，其他网络节点接收状态请求消息并向网络节点R_0发送状态回复消息(例如，REPLY_STATE消息)。在第三阶段1030中，网络节点R_0向其他网络节点发送请求消息(例如，FETCH_ECHO消息)以从每个其他网络节点请求ECHO消息。如上述参考图3-6所述，ECHO消息可以与相应的其他网络节点在共识处理300的第二阶段320发送的ECHO消息相同。在第四阶段1040中，响应于FETCH_ECHO消息，每个其他网络节点向网络节点R_0发送ECHO消息。在第五阶段1050，网络节点R_0根据EC码验证ECHO消息并恢复目标交易，例如，根据上述参考图4的示例性重建技术。在第六阶段1060中，网络节点R_0向其他网络节点发送ACCEPT消息以表示网络节点已经被恢复。

[0151] 注意，仅用于说明性目的，处理1000被示为结合四个网络节点实施。处理1000可以结合任何合适数量的网络节点实施。处理1000、QUERY_STATE消息的格式、以及REPLY_STATE消息的格式将在下文参考图11-12更详细地讨论。

[0152] 图11描绘了可根据本申请实施方式执行的、用于在分布式系统中(例如，区块链网络102或212)执行网络节点恢复处理的处理1100的示例。在一些实施方式中，处理1100可以使用一个或多个计算机设备执行一个或多个计算机可执行程序来进行。为了清楚地呈现，以下描述总体上描述了在本说明书中的其他附图的背景中的方法1100。可以理解的是，方法1100可以由诸如任何合适的系统、环境、软件和硬件，或系统、环境、软件和硬件的组合合理进行。在一些实施方式中，方法1100的各步骤可以并行、组合、循环或以任何顺序进行。

[0153] 处理1100始于1106，其中网络节点1102将状态请求消息组播至其他网络节点1104。状态请求消息包括网络节点1102将恢复目标交易具有目标序列号的指示。网络节点1102可以是主节点或备节点。

[0154] 在本申请的一些实施方式中，参考图12，QUERY_STATE消息，作为状态请求消息的示例，具有为<j,seq>的格式，其中“j”代表发送QUERY_STATE消息的网络节点1102，以及“seq”代表网络节点1102在当前共识处理中的最大序列号或最新序列号。在一些实施方式中，QUERY_STATE消息可以具有不同的格式，例如，包括附加或不同的字段。

[0155] 通过将QUERY_STATE消息广播至其他网络节点1104，网络节点1102请求其他网络节点1104将它们的最新序列号发送至网络节点1102，从而获取区块链系统的最新区块信息。并且，通过获取整个区块链系统的最新区块信息，网络节点1102可以能够同步到整个系统的最新状态，从而将其自身恢复以继续参与至共识处理中。

[0156] 再次参考图11,在1108,响应于接收到状态请求消息,每个其他网络节点1104向网络节点1102发送状态回复消息(例如,REPLY_STATE消息)。在一些实施方式中,状态回复消息包括与网络节点1104关联的先前序列号。

[0157] 在一些实施方式中,参考图12,REPLY_STATE消息作为状态回复消息的示例,具有为<j,last_seq>的格式,其中“j”代表发送REPLY_STATE消息的网络节点1104,以及“last_seq”代表网络节点1104在当前共识处理中的先前序列号。在一些实施方式中,REPLY_STATE消息可以具有不同格式,例如,包括附加或不同的字段。

[0158] 返回参考图11,在1110,网络节点1102确定接收到的状态回复消息的数量是否超过预设阈值。例如,网络节点1102可以确定接收到的REPLY_STATE的数量是否超过法定数量(例如,2f+1或n-f个)。在一些实施方式中,网络节点1102还确定接收到的法定数量的REPLY_STATE消息是否包括相同的序列号。接收到的法定数量的REPLY_STATE消息包括相同的序列号,意味着大部分的网络节点1104就整个系统的公共状态达成共识。

[0159] 在1112,若网络节点1102确定从网络节点1104接收到的包括该相同的序列号的状态回复消息的数量超过预设阈值,则网络节点1102基于该相同的序列号确定目标序列号。例如,网络节点1102可以确定该目标序列号为该相同的序列号(例如,“last_seq”)的增量(例如,“last_seq+1”)。

[0160] 在1114,网络节点1102向其他网络节点1104发送请求消息(例如,FETCH_ECHO消息)。FETCH_ECHO消息由网络节点1102发送,以从每个其他网络节点1104请求ECHO消息。如上参考图3-6所讨论的,ECHO消息是由网络节点1104传输的消息,用于在网络节点1104之间就目标交易达成共识。ECHO消息包括目标交易的一部分(例如,EC区块)以及发送ECHO消息的网络节点1104的签名。

[0161] 在一些实施方式中,参考图12,作为请求消息示例的FETCH_ECHO消息,具有为<j,last_seq+1>的格式,其中“j”代表发送FETCH_ECHO消息的网络节点1102,以及“last_seq+1”代表与网络节点1102正从其他网络节点1104请求的ECHO消息相关联的目标序列号。在一些实施方式中,FETCH_ECHO消息可以具有不同格式,例如,包括附加或不同的字段。

[0162] 本文所讨论的FETCH_ECHO消息是由网络节点1102发送,以从其他网络节点1104请求包括最新序列号或最大序列号的ECHO消息。通过从其他网络节点1104收集包括最新序列号或最大序列号的ECHO消息,网络节点1102可能恢复与最新序列号关联的最新状态。

[0163] 返回参考图11,在1116,响应于接收到的FETCH_ECHO消息,每个网络节点1104向网络节点1102发送ECHO消息。在一些实施方式中,每个网络节点1104在向网络节点1102发送ECHO消息之前,验证FETCH_ECHO消息。在一些实施方式中,每个网络节点1104通过确定包含于FETCH_ECHO消息中的序列号是否超过与网络节点1104关联的最新序列号来验证FETCH_ECHO消息。若包含于FETCH_ECHO消息中的序列号等于与网络节点1104关联的最新序列号,则网络节点1104确定FETCH_ECHO有效,且将ECHO消息发送至网络节点1102。若包含于FETCH_ECHO消息中的序列号超过与网络节点1104关联的最新序列号,则网络节点1104确定FETCH_ECHO无效,且不会将ECHO消息发送至网络节点1102。

[0164] 在1118,网络节点1102验证由网络节点1104发送的ECHO消息是否有效。在一些实施方式中,网络节点1102利用梅克尔树验证ECHO消息。例如,网络节点1102可以利用包含于ECHO消息中的数据来重建梅克尔树并确定重建的梅克尔树的重建的根哈希值。然后网络节

点1102可以将重建的根哈希值与包含于ECHO消息中的根哈希值进行比较。若重建的根哈希值与包含于ECHO消息中的根哈希值匹配，则网络节点1102确定ECHO消息有效。若重建的根哈希值与包含于ECHO消息中的根哈希值不匹配，则网络节点1102确定ECHO消息无效并可以丢弃无效的ECHO消息。

[0165] 在一些实施方式中，网络节点1102通过进一步验证ECHO消息中的签名是否有效来验证ECHO消息是否有效。例如，网络节点1102可以利用与该私钥配对的公钥来验证ECHO消息的私钥签名，以验证签名。

[0166] 在1120，网络节点1102确定从其他网络节点1104接收的有效ECHO消息的数量是否超过预设阈值，例如，网络节点1102可以确定从其他网络节点1104接收的有效ECHO消息的数量是否超过法定数量(例如，2f+1)。

[0167] 在1122，响应于确定有效ECHO消息的数量超过预设阈值，网络节点1102恢复目标交易具有目标序列号。在一些实施方式中，网络节点1102利用包含于多个有效ECHO消息中的数据来恢复目标交易。例如，网络节点1102可以检索包含于ECHO消息中的EC区块的子集，以根据EC码重建目标交易。

[0168] 在1124，网络节点1102在恢复目标交易后将ACCEPT消息组播至其他网络节点1104。例如，网络节点1102可以在成功重建目标交易后将ACCPET消息组播至其他网络节点1104。在一些实施方式中，ACCEPT消息包括ECHO消息中的一组签名和目标序列号。通过向其他网络节点1104发送包括签名和目标序列号的ACCEPT消息，网络节点1102向其他网络节点1104表示网络节点1102已被恢复并被同步到系统的最新状态。

[0169] 本申请中的上面所讨论的恢复处理包括许多特征，这些特征改进了实施恢复处理的计算机操作并帮助缓解了网络瓶颈。例如，本申请中的恢复处理包括以下操作：由申请成为新主节点的网络节点发送状态请求消息；从其他网络节点接收状态回复消息；以及由该网络节点发送FETCH_ECHO消息，以从其他网络节点请求ECHO消息。在进行这些操作时，错误网络节点的恢复处理不会干扰其他非错误网络节点中的共识处理的正常运行。通过降低恢复处理的复杂性来恢复错误网络节点，这有助于节省计算和网络资源。

[0170] 参考图13，图13是根据本申请实施方式示出的共识装置1300的模块图示。用于达成共识的装置1300可以应用于基于区块链技术的共识系统。例如，装置1300可以对应于图1-6所示的实施方式。装置1300可以在区块链网络中的主节点实现。装置1300包括：接收器或接收单元1302，配置为接收交易请求；生成单元1304，配置为利用所述交易请求根据纠删码EC码生成多个EC区块；发送器或发送单元1306，配置为分别向一个或多个备节点发送多个第一消息，其中，所述多个第一消息各自包括与所述多个EC区块关联的复合哈希值；接收器或接收单元1302，进一步配置为从至少所述一个备节点接收至少一个第二消息，其中，所述至少一个第二消息包括所述多个第一消息中的一个第一消息和与所述多个第一消息中的该个第一消息相关联的至少一个备节点的签名；验证单元1306，配置为响应于从至少一个备节点接收至少一个第二消息，验证至少一个第二消息是否有效；确定单元1310，配置为确定有效第二消息的数量是否超过预设阈值；重建单元1312，配置为响应于确定有效第二消息的数量超过了预设阈值，根据EC码，基于多个有效第二消息的子集重建交易请求；发送器或发送单元1306，进一步配置为响应于确定已经成功重建交易请求，向其他网络节点发送第三消息，其中所述第三消息包括有效第二消息中的一组签名；接收器或接收单元1302，

进一步配置为从至少一个所述备节点接收至少一个第三消息;以及执行单元1314,配置为响应于接收到预设数量的相同的第三消息,执行交易请求。

[0171] 在一可选实施方式中,交易请求与序列号相关联。

[0172] 在一可选实施方式中,根据EC码生成多个EC区块包括以下:利用EC码将交易请求转换为EC消息,并将EC消息划分为所述多个EC区块。

[0173] 在一可选实施方式中,利用哈希树生成多个EC区块的所述复合哈希值。

[0174] 在一可选实施方式中,哈希树包括梅克尔(Merkle)树,且其中复合哈希值是梅克尔树的根哈希值。

[0175] 在一可选实施方式中,与多个第一消息中的一个第一消息相关联的所述至少一个备节点的签名包括:与所述多个第一消息中的该个第一消息相关联的至少一个备节点的私钥签名。

[0176] 在一可选实施方式中,所述至少一个第二消息还包括所述多个EC区块中的至少一个区块。

[0177] 在一可选实施方式中,验证所述至少一个第二消息是否有效,包括以下:利用所述至少一个第二消息中的多个EC区块中的至少一个区块生成重建的哈希树;确定重建的哈希树的重建的复合哈希值;以及确定重建的复合哈希值是否与所述至少一个第二消息中的复合哈希值匹配。

[0178] 在一可选实施方式中,确定单元1310进一步配置为:响应于确定重建的复合哈希值与至少一个第二消息中的所述复合哈希值匹配,验证所述第二消息有效。

[0179] 在一可选实施方式中,预设数量的相同的第三消息包括:所述预设数量的所述第三消息具有相同的签名组。

[0180] 图13为示出了内部功能模块和共识装置1300的结构的示意图。实际执行主体可以为电子设备,该电子设备包括以下:至少一个处理器;和配置为存储至少一个处理器可执行的指令的存储器。

[0181] 至少一个处理器配置为:接收交易请求;利用交易请求根据纠删码(EC)生成多个EC区块;将多个第一消息分别发送至一个或多个备节点,其中,所述多个第一消息中的每个包括与多个EC区块关联的复合哈希值;从至少一个所述备节点接收至少一个第二消息,其中,所述至少一个第二消息包括所述多个第一消息中的一个第一消息以及与所述多个第一消息中的该个第一消息相关联的至少一个备节点的签名;响应于从至少一个备节点接收的至少一个第二消息,验证所述至少一个第二消息是否有效;确定有效第二消息的数量是否超过预设阈值;响应于确定有效第二消息的数量超过预设阈值,根据EC码基于多个有效第二消息的子集重建交易请求;响应于确定已经成功重建交易请求,向其他网络节点发送第三消息,其中,第三消息包括有效第二消息中的一组签名;从至少一个所述备节点接收至少一个第三消息;以及响应于接收到预设数量的相同的第三消息,执行交易请求。

[0182] 可选地,交易请求与序列号关联。

[0183] 可选地,根据EC码生成多个EC区块包括:利用EC码将交易请求转换为EC消息,并将EC消息划分为所述多个EC区块。

[0184] 可选地,利用哈希树生成多个EC区块的所述复合哈希值。

[0185] 可选地,哈希树包括梅克尔树,并且复合哈希值是梅克尔树的根哈希值。

[0186] 可选地，与多个第一消息中的一个关联的至少一个备节点的签名包括：与多个第一消息中的一个关联的至少一个备节点的私钥签名。

[0187] 可选地，所述至少一个第二消息还包括所述多个EC区块中的至少一个区块。

[0188] 可选地，验证所述至少一个第二消息是否有效，包括以下：利用所述至少一个第二消息中的多个EC区块中的至少一个区块生成重建的哈希树；确定重建的哈希树的重建的复合哈希值；以及确定重建的复合哈希值与至少一个第二消息中的复合哈希值是否匹配。

[0189] 可选地，至少一个处理器还配置为：响应于确定重建的复合哈希值与至少一个第二消息中的复合哈希值匹配，确定所述第二消息有效。

[0190] 可选地，预设数量的相同的第三消息包括：所述预设数量的所述第三消息具有相同的签名组。

[0191] 参考图14，图14是根据本申请实施方式示出的共识装置1400的模块的示意图。用于达成共识的装置1400可以应用于基于区块链技术的共识系统。装置1400可以对应于图1-6所示的实施方式，装置1400可以在区块链网络中的备节点实施。装置1400包括：接收器或接收单元1402，配置为从主节点接收第一消息，其中，所述第一消息包括与多个EC区块关联的复合哈希值，所述多个EC区块是主节点根据EC码利用交易请求生成的；发送器或发送单元1404，配置为响应于接收到的第一消息，备节点向其他网络节点发送第二消息，其中，所述第二消息包括第一消息和与第一消息关联的备节点的签名；接收器或接收单元1402，进一步配置为从除该备节点外的至少一个备节点接收至少一个第二消息；验证单元1406，配置为响应于从至少一个备节点接收至少一个第二消息，验证至少一个第二消息是否有效；确定单元1408，配置为确定有效第二消息的数量是否超过预设阈值；重建单元1410，配置为响应于确定有效第二消息的数量超过了预设阈值，根据EC码，基于多个有效第二消息的子集重建交易请求；发送器或发送单元1404，配置为响应于确定已经成功重建所述交易请求，向其他网络节点发送第三消息，其中，所述第三消息包括有效第二消息中的一组签名；接收器或接收单元1402，进一步配置为从至少一个备节点接收至少一个第三消息；以及执行单元1412，配置为响应于接收到预设数量的相同的第三消息，执行交易请求。

[0192] 在一可选实施方式中，根据EC码生成多个EC区块包括以下：利用EC码将交易请求转换为EC消息，并将EC消息划分为多个EC区块。

[0193] 在一可选实施方式中，利用哈希树生成多个EC区块的复合哈希值。

[0194] 在一可选实施方式中，哈希树包括梅克尔树，并且其中复合哈希值是梅克尔树的根哈希值。

[0195] 在一可选实施方式中，与第一消息关联的备节点的签名包括：与第一消息关联的备节点的私钥签名。

[0196] 在一可选实施方式中，所述至少一个第二消息还包括多个EC区块中的至少一个。

[0197] 在一可选实施方式中，验证至少一个第二消息是否有效，包括以下：利用在至少一个第二消息中的多个EC区块中的至少一个生成重建的哈希树；确定重建的哈希树的重建的复合哈希值；对比重建的复合哈希值与至少一个第二消息中的复合哈希值；以及确定重建的复合哈希值是否与至少一个第二消息中的复合哈希值匹配。

[0198] 在一可选实施方式中，确定单元1408进一步配置为：响应于确定重建的复合哈希值与至少一个第二消息中的复合哈希值匹配，确定所述第二消息有效。

[0199] 在一可选实施方式中，预设数量的相同的第三消息包括预设数量的第三消息具有相同的签名组。

[0200] 图14是示出了共识装置的1400的内部功能模块和结构的示意图。实际的执行主体可以为电子设备，该电子设备包括以下：至少一个处理器；和配置为存储至少一个处理器可执行的指令的存储器。

[0201] 所述至少一个处理器配置为：从主节点接收第一消息，其中所述第一消息包括与多个EC区块关联的复合哈希值，其中所述多个EC区块是由主节点利用交易请求根据EC码生成的；响应于接收到第一消息，备节点向其他网络节点发送第二消息，其中，所述第二消息包括第一消息和与第一消息关联的备节点的签名；从除该备节点外的至少一个备节点接收至少一个第二消息；响应于从所述至少一个备节点接收至少一个第二消息，验证至少一个第二消息是否有效；确定有效第二消息的数量是否超过预设阈值；响应于确定有效第二消息的数量超过了预设阈值，根据EC码基于多个有效第二消息的子集重建交易请求；响应于确定已经成功重建交易请求，向其他网络节点发送第三消息，其中，第三消息包括有效第二消息中的一组签名；从至少一个备节点接收至少一个第三消息；以及响应于接收到预设数量的相同的第三消息，执行交易请求。

[0202] 可选地，根据EC码生成多个EC区块包括：利用EC码将交易请求转换为EC消息，并将EC消息划分为多个EC区块。

[0203] 可选地，利用哈希树生成多个EC区块的复合哈希值。

[0204] 可选地，哈希树包括梅克尔树，并且复合哈希值是梅克尔树的根哈希值。

[0205] 可选地，与第一消息关联的备节点的签名包括：与第一消息关联的备节点的私钥签名。

[0206] 可选地，至少一个第二消息还包括所述多个EC区块中的至少一个。

[0207] 可选地，验证至少一个第二消息是否有效，包括：利用在至少一个第二消息中的多个EC区块中的至少一个生成重建的哈希树；确定重建的哈希树的重建的复合哈希值；对比重建的复合哈希值与至少一个第二消息中的复合哈希值；以及确定重建的复合哈希值是否与至少一个第二消息中的复合哈希值匹配。

[0208] 可选地，至少一个处理器还配置为：响应于确定重建的复合哈希值与至少一个第二消息中的复合哈希值匹配，确定所述第二消息有效。

[0209] 可选地，预设数量的相同的第三消息包括预设数量的第三消息具有相同的签名组。

[0210] 参考图15，图15是根据本申请实施方式示出的主节点变更装置1500的模块的图。用于变更主节点的装置1500可以应用于基于区块链技术的共识系统。装置1500可以对应于图7-9所示的实施方式，装置1500可以在区块链网络中的备节点实施。装置1500包括以下：确定单元1502，配置为确定需要进行时元改变，其中所述时元改变促使具有当前主节点的当前时元变更为具有新主节点的新时元，其中当前时元包括利用主节点在多个网络节点中达成共识的共识处理，共识处理包括三个阶段；确定单元1502，进一步配置为确定该备节点在当前时元内的共识处理的三个阶段中的每个阶段的相应的权重，其中，所述权重是对该备节点成为新时元中的新主节点的资格的度量；确定单元1502，还配置为基于该备节点在当前时元内的共识处理的三个阶段中的每个阶段的相应的权重确定该备节点的权重和；发

送器或发送单元1504，配置为响应于确定权重和达到了第一预设阈值，向多个网络节点中除该网络节点外的其他网络节点发送EPOCH_CHANGE消息，其中，EPOCH_CHANGE消息表示请求将具有当前主节点的当前时元变更为该备节点作为新主节点的新时元，以及EPOCH_CHANGE消息包括该备节点的权重和；接收器和接收单元1506，配置为从多个网络节点中除该备节点外的至少一个其他网络节点接收至少一个NEW_EPOCH消息，其中，所述NEW_EPOCH消息表示该备节点将成为新主节点的确认；验证单元1508，配置为验证至少一个NEW_EPOCH消息是否有效；确定单元1502，还配置为确定所述至少一个NEW_EPOCH消息中的有效NEW_EPOCH消息的数量是否超过第二预设阈值；以及确定单元1502，还配置为响应于确定有效NEW_EPOCH消息的数量超过第二预设阈值，确定该备节点将成为新时元中的新主节点。

[0211] 在一可选实施方式中，确定该备节点在当前时元内的共识处理的三个阶段中的每个阶段的相应的权重包括：将该备节点在共识处理第一阶段的权重确定为第一值。

[0212] 在一可选实施方式中，确定该备节点在当前时元内的共识处理的三个阶段中的每个阶段的相应的权重包括：响应于确定当前时元内的共识处理第二阶段的Quorum验证失败，将该备节点在共识处理第二阶段的权重确定为第一值；以及响应于确定当前时元内的共识处理第二阶段的Quorum验证成功，将该备节点在共识处理第二阶段的权重确定为第二值，其中第一值小于第二值。

[0213] 在一可选实施方式中，针对网络节点的在第二阶段中的Quorum验证包括从其他网络节点接收预设数量的ECHO消息。

[0214] 在一可选实施方式中，确定该备节点在当前时元内的共识处理的三个阶段中的每个阶段的相应的权重包括：响应于确定当前时元内的共识处理第三阶段的Quorum验证失败，将该备节点在共识处理第三阶段的权重确定为第三值；以及响应于确定当前时元内的共识处理第三阶段的Quorum验证成功，将该备节点在共识处理第三阶段的权重确定为第四值，其中第三值小于第四值。

[0215] 在一可选实施方式中，网络节点在第三阶段中的Quorum验证包括从其他网络节点接收预设数量的接受消息，其中，来自其他网络节点的每个接受消息表示每个其他网络节点已经接受了预设数量的ECHO消息。

[0216] 在一可选实施方式中，EPOCH_CHANGE消息进一步包括与多个网络节点中的一组网络节点关联的一组签名，且NEW_EPOCH消息包括EPOCH_CHANGE消息的摘要。

[0217] 在一可选实施方式中，验证所述至少一个有效NEW_EPOCH消息是否有效，包括：验证所述至少一个NEW_EPOCH中的EPOCH_CHANGE消息的摘要是否有效；以及，验证所述至少一个NEW_EPOCH中的EPOCH_CHANGE消息的摘要是否有效，包括：验证EPOCH_CHANGE消息中的该组签名是否有效。

[0218] 在一可选实施方式中，确定需要进行时元改变包括：响应于确定在旧时元中在预设时间内尚未达成共识，确定需要进行时元改变。

[0219] 在一可选实施方式中，主节点变更装置1500还包括：操作单元1510，配置为在具有新主节点的新时元中操作，其中所述新时元包括利用新主节点在多个网络节点中达成共识的共识处理。

[0220] 图15是示出了主节点变更装置1500的内部功能模块和结构的示意图。实际执行主体可以为电子设备，该电子设备包括以下：至少一个处理器；和配置为存储至少一个处理器

可执行的指令的存储器。

[0221] 该至少一个处理器被配置为:确定需要进行时元改变,其中所述时元改变促使从具有当前主节点的当前时元变更为具有新主节点的新时元,其中当前时元包括利用主节点在多个网络节点中达成共识的共识处理,所述共识处理包括三个阶段;确定该备节点在当前时元内的共识处理的三个阶段中的每个阶段的相应的权重,其中所述权重是对该备节点成为新时元中的新主节点的资格的度量;基于该备节点在当前时元的三个阶段中的每个阶段的相应的权重确定该备节点的权重和;响应于确定所述权重和达到了第一预设阈值,向多个网络节点中除该网络节点外的其他网络节点发送EPOCH_CHANGE消息,其中所述EPOCH_CHANGE消息表示请求将具有当前主节点的当前时元变更为以该备节点作为新主节点的新时元,以及所述EPOCH_CHANGE消息包括该备节点的权重和;从多个网络节点中除该备节点外的至少一个其他网络节点接收至少一个NEW_EPOCH消息,其中,所述NEW_EPOCH消息表示确认该备节点将成为新主节点;验证至少一个NEW_EPOCH消息是否有效;确定至少一个NEW_EPOCH消息中的有效NEW_EPOCH消息的数量是否超过第二预设阈值;以及响应于确定有效NEW_EPOCH消息的数量超过第二预设阈值,确定所述备节点将成为新时元中的新主节点。

[0222] 可选地,确定该备节点在当前时元内的共识处理的三个阶段中的每个阶段的相应的权重包括:将该备节点在共识处理第一阶段的权重确定为第一值。

[0223] 可选地,确定该备节点在当前时元内的共识处理的三个阶段中的每个阶段的相应的权重包括:响应于确定当前时元内的共识处理第二阶段的Quorum验证失败,将该备节点在共识处理第二阶段的权重确定为第一值;以及响应于确定当前时元内的共识处理第二阶段的Quorum验证成功,将该备节点在共识处理第二阶段的权重确定为第二值,其中所述第一值小于第二值。

[0224] 可选地,针对网络节点第二阶段中的Quorum验证包括从其他网络节点接收预设数量的ECHO消息。

[0225] 可选地,确定该备节点在当前时元内的共识处理的三个阶段中的每个阶段的相应的权重包括:响应于确定当前时元内的共识处理第三阶段的Quorum验证失败,确定该备节点在共识处理第三阶段的权重作为第三值;以及响应于确定当前时元内的共识处理第三阶段的Quorum验证成功,确定该备节点在共识处理第三阶段的权重为第四值,其中所述第三值小于第四值。

[0226] 可选地,网络节点在第三阶段中的Quorum验证包括从其他网络节点接收预设数量的接受消息,其中来自其他网络节点的每个接受消息表示每个其他网络节点已经接受了预设数量的ECHO消息。

[0227] 可选地,EPOCH_CHANGE消息进一步包括与多个网络节点中的一组网络节点关联的一组签名,且NEW_EPOCH消息包括EPOCH_CHANGE消息的摘要。

[0228] 可选地,验证所述至少一个有效NEW_EPOCH消息是否有效,包括:验证所述至少一个NEW_EPOCH中的EPOCH_CHANGE消息的摘要是否有效;以及,验证至少一个NEW_EPOCH中的EPOCH_CHANGE消息的摘要是否有效,包括:验证EPOCH_CHANGE消息中的该组签名是否有效。

[0229] 可选地,确定需要进行时元改变包括:响应于确定在旧时元中在预设时间内尚未达成共识,确定需要进行时元改变。

[0230] 可选地,所述至少一个处理器进一步配置为利用新主节点在新时元中操作,其中

所述新时元包括利用所述新主节点在多个网络节点中达成共识的共识处理。

[0231] 参考图16,图16是根据本申请实施方式示出的主节点变更装置1600的模块的图。用于变更主节点的装置1600可以应用于基于区块链技术的共识系统。装置1600可以对应于图7-9所示的实施方式。例如,装置1600可以在区块链网络中的网络节点实施。装置1600包括以下:接收器或接收单元1602,配置为从除该网络节点外的备节点接收EPOCH_CHANGE消息,其中,所述EPOCH_CHANGE消息包括需要进行时元改变的指示,所述时元改变促使从具有当前主节点的当前时元变更为具有新主节点的新时元;验证单元1604,配置为验证EPOCH_CHANG消息是否有效;发送器或发送单元1606,配置为响应于验证EPOCH_CHANGE消息有效,向其他网络节点发送NEW_EPOCH消息,其中,所述NEW_EPOCH消息包括EPOCH_CHANGE消息的摘要;接收器或接收单元1602,进一步被配置为从多个网络节点中除该网络节点外的至少一个其他网络节点接收至少一个NEW_EPOCH消息;验证单元1604,进一步被配置为验证所述至少一个NEW_EPOCH消息是否有效;确定单元1608,被配置为确定至少一个NEW_EPOCH消息中的有效NEW_EPOCH消息的数量是否超过预设阈值;以及确定单元1608,还被配置为响应于有效NEW_EPOCH消息的数量超过预设阈值,确定该备节点将成为新时元内的新主节点。

[0232] 在一可选实施方式中,EPOCH_CHANGE消息包括与备节点关联的权重和以及与多个网络节点中的一组网络节点关联的一组签名。

[0233] 在一可选实施方式中,验证EPOCH_CHANGE消息是否有效,包括:验证EPOCH_CHANGE消息中的权重和是否有效;以及,验证EPOCH_CHANGE消息中的权重和是否有效,包括:验证该组签名是否有效。

[0234] 在一可选实施方式中,验证至少一个NEW_EPOCH消息是否有效,包括:验证至少一个NEW_EPOCH消息中的EPOCH_CHANGE消息的摘要是否有效;以及,验证至少一个NEW_EPOCH消息中的EPOCH_CHANGE消息的摘要是否有效,包括:验证EPOCH_CHANGE消息中的该组签名是否有效。

[0235] 图16是示出了主节点变更装置1600的内部功能模块和结构的示意图。实际执行主体可以为电子设备,该电子设备包括以下:至少一个处理器;和配置为存储至少一个处理器可执行的指令的存储器。

[0236] 至少一个处理器配置为:从除所述网络节点外的备节点接收EPOCH_CHANGE消息,其中,EPOCH_CHANGE消息包括需要进行时元改变的指示,所述时元改变促使从具有当前主节点的当前时元变更为具有新主节点的新时元;验证EPOCH_CHANG消息是否有效;响应于验证EPOCH_CHANGE消息有效,向其他网络节点发送NEW_EPOCH消息,其中NEW_EPOCH消息包括EPOCH_CHANGE消息的摘要;从多个网络节点中除该网络节点外的至少一个其他网络节点接收至少一个NEW_EPOCH消息;验证至少一个NEW_EPOCH消息是否有效;确定至少一个NEW_EPOCH中的有效NEW_EPOCH的数量是否超过了预设阈值;以及响应于确定有效NEW_EPOCH消息的数量超过预设阈值,确定该备节点成为新时元内的新主节点。

[0237] 可选地,EPOCH_CHANGE消息包括与备节点关联的权重和以及与所述多个网络节点中的一组网络节点关联的一组签名。

[0238] 可选地,验证EPOCH_CHANGE消息是否有效,包括:验证EPOCH_CHANGE消息中的权重和是否有效;以及,验证EPOCH_CHANGE消息中的权重和是否有效,包括:验证该组签名是否有效。

[0239] 可选地，验证至少一个NEW_EPOCH消息是否有效，包括：验证至少一个NEW_EPOCH消息中的EPOCH_CHANGE消息的摘要是否有效；以及，验证至少一个NEW_EPOCH消息中的EPOCH_CHANGE消息的摘要是否有效，包括：验证EPOCH_CHANGE消息中的该组签名是否有效。

[0240] 参考图17，图17是根据本申请实施方式示出的恢复装置1700的模块的图。恢复装置1700可以应用于基于区块链技术的共识系统。装置1700可以对应于图10-12所示的实施方式。例如，装置1700可以在区块链网络中的网络节点实现。装置1700包括以下：广播单元1702，配置为由区块链网络中的网络节点将状态请求消息广播至区块链网络中的多个其他网络节点，其中，该网络节点待恢复目标交易具有目标序列号；接收器1704或接收单元1704，配置为从多个其他网络节点接收多个状态回复消息，其中，所述多个状态消息各自包括序列号；识别单元1706，配置为响应于确定各自包括相同的序列号的状态回复消息的数量超过了预设阈值，基于相同的序列号识别目标序列号；发送器1708或发送单元1708，配置为向多个其他网络节点发送请求消息，其中，请求消息从多个其他网络节点分别请求ECHO消息，ECHO消息是由多个其他网络节点分别传输并用于在多个其他网络节点之间就目标交易具有目标序列号达成共识的消息，并且ECHO消息包括目标交易的一部分以及多个其他网络节点各自的签名；接收器1704或接收单元1704，进一步配置为从多个其他网络节点接收多个ECHO消息；确定单元1710，配置为确定多个ECHO消息中的有效ECHO消息的数量，其中，所述有效ECHO消息各自包括所述目标序列号；恢复单元1712，配置为响应于确定有效ECHO消息的数量超过预设阈值，基于多个有效ECHO消息在该网络节点处恢复目标交易具有相同序列号；以及发送器1708，进一步配置为向多个其他网络节点发送表示该网络节点已被恢复的消息。

[0241] 在一可选实施方式中，多个网络节点包括主节点和一个或多个备节点。

[0242] 在一可选实施方式中，该网络节点是主节点或备节点。

[0243] 在一可选实施方式中，请求消息包括目标序列号。

[0244] 在一可选实施方式中，恢复装置1700还包括：验证单元1714，配置为多个其他网络节点中除该网络节点外的每一个其他网络节点在向该网络节点发送ECHO消息之前，验证所述请求消息。

[0245] 在一可选实施方式中，验证单元1714还配置为验证每个ECHO消息是否有效；其中，验证每个ECHO消息是否有效，包括利用梅克尔树验证每个ECHO消息是否有效。

[0246] 在一可选实施方式中，验证每个ECHO消息是否有效还包括：验证ECHO消息中的签名是否有效。

[0247] 在一可选实施方式中，每个ECHO消息进一步包括与目标交易关联的多个纠删码(EC)区块中的至少一个区块，其中，多个EC区块是利用目标交易根据EC码生成的。

[0248] 在一可选实施方式中，基于多个有效ECHO消息在该网络节点处恢复目标交易具有相同序列号，包括：利用多个有效ECHO消息中的多个EC区块的子集重建目标交易。

[0249] 在一可选实施方式中，发送至多个其他网络节点的表示该网络节点已被恢复的消息包括：多个有效ECHO消息中的一组签名，和目标序列号。

[0250] 先前实施方式中示出的系统、装置、模块或单元可以由计算机芯片或实体实现，或可以通过利用具有特定功能的产品实现。典型的实施设备为计算机，且计算机可以为个人计算机、膝上型计算机、蜂窝电话、照相电话、智能电话、个人数字助手、媒体播放器、导航设

备、电子邮件接收和发送设备、游戏控制台、平板电脑、可穿戴设备或任何这些设备的组合。

[0251] 对于装置中每个单元的功能和角色的实现处理，可参考前述方法中对应的步骤的实现处理。为简单起见，在此不做赘述。

[0252] 由于装置的实施方式基本对应于方法实施方式，对于相关部分，可参考方法实施方式中的相关描述。先前描述的装置实施方式仅为示例。单独描述的单元可以是或可以不是物理分离的，且示出的作为单元的部分可以是或可以不是物理单元，可以位于一个位置或可以分布在多个网络单元。可以基于实际需求选择某些或全部模块以达成本申请解决方案的目标。本领域普通技术人员在不付出创造性劳动的情况下就可以理解并实现本申请的实施方式。

[0253] 图17是示出了恢复装置1700的内部功能模块和结构的示意图。实际执行主体可以为电子设备，该电子设备包括以下：至少一个处理器；和配置为存储至少一个处理器可执行的指令的存储器。

[0254] 所述至少一个处理器配置为：区块链网络中的网络节点将状态请求消息广播至区块链网络中的多个其他网络节点，其中，该网络节点待恢复目标交易具有目标序列号；从多个其他网络节点接收多个状态回复消息，所述多个状态消息各自包括序列号；响应于确定各自包括相同的序列号的状态回复消息的数量超过了预设阈值，基于相同的序列号识别目标序列号；向所述多个其他网络节点发送请求消息，其中，所述请求消息从所述多个其他网络节点分别请求ECHO消息，所述ECHO消息是由所述多个其他网络节点分别传输并用于在多个其他网络节点之间就目标交易具有目标序列号达成共识的消息，并且ECHO消息包括目标交易的一部分以及多个其他网络节点各自的签名；从多个其他网络节点接收多个ECHO消息；确定多个ECHO消息中的有效ECHO消息的数量，其中，所述有效ECHO消息各自包括目标序列号；响应于确定有效ECHO消息的数量超过预设阈值，基于多个有效ECHO消息在该网络节点处恢复目标交易具有相同序列号；以及向多个其他网络节点发送消息，表示该网络节点已被恢复。

[0255] 可选地，多个网络节点包括主节点和一个或多个备节点。

[0256] 可选地，该网络节点是主节点或备节点。

[0257] 可选地，所述请求消息包括目标序列号。

[0258] 可选地，所述至少一个处理器进一步配置为，所述多个其他网络节点中除该网络节点外的每个其他网络节点在向所述网络节点发送ECHO消息之前，分别验证所述请求消息。

[0259] 可选地，验证每个ECHO消息是否有效；其中，验证每个ECHO消息是否有效，包括：利用梅克尔树验证每个ECHO消息是否有效。

[0260] 可选地，验证每个ECHO消息是否有效还包括：验证ECHO消息中的签名是否有效。

[0261] 可选地，每个ECHO消息进一步包括与目标交易关联的多个纠删码(EC)区块中的至少一个区块，其中，所述多个EC区块是利用目标交易根据EC码生成的。

[0262] 可选地，基于所述多个有效ECHO消息在该网络节点处恢复目标交易具有相同序列号，包括：利用多个有效ECHO消息中的多个EC区块的子集重建目标交易。

[0263] 在一可选实施方式中，发送至多个其他网络节点的表示网络节点已被恢复的消息包括：多个有效ECHO消息中的一组签名，和目标序列号。

[0264] 本申请中描述的主题、动作和操作的实施方式可在数字电子电路中，有形体现的计算机软件或固件、计算机硬件中实现，包括本申请中公开的结构及其结构等同物、或它们中的一个或多个的组合。本申请中描述的主题的实施方式的实施可以实现一个或多个计算机程序，例如，一个或多个计算机程序指令模块，编码在计算机程序载体上，用于由数据处理装置执行或控制数据处理装置的操作。载体可以是有形的非暂态计算机存储介质。可选地或附加地，载体可以是人为生成的传播信号，例如，机器生成的电信号、光信号或电磁信号，其被生成来编码信息用于传输到合适的接收装置上以供数据处理设备执行。计算机存储介质可以是或部分是机器可读存储设备、机器可读存储基板、随机或串行存取存储设备或它们中的一个或多个的组合。计算机存储介质不是传播信号。

[0265] 术语“数据处理装置”包括用于处理数据的所有类型的装置、设备和机器，包括如：可编程处理器、计算机或多个处理器或计算机。数据处理装置可包括专用逻辑电路，例如FPGA(现场可编程门阵列)、ASIC(专用集成电路)或GPU(图形处理单元)。除了硬件，该装置还可包括为计算机程序创建执行环境的代码，例如，构成处理器固件、协议栈、数据库管理系统、操作系统或者它们之中一个或多个的组合的代码。

[0266] 计算机程序也可被称为或描述为程序、软件、软件应用、应用程序、app、模块、软件模块、引擎、脚本或代码，可以以任何形式的编程语言编写，包括编译或演绎性语言、说明或程序性语言；它可以配置为任何形式，包括作为独立程序，或作为模块、组件、引擎、子程序或适合在计算环境中执行的其他单元，该环境可包括通过在一个或多个位置由数据通信网络互连的一个或多个计算机。

[0267] 计算机程序可以但非必须对应于文件系统中的文件。计算机程序可存储在：存有其他程序或数据的文件的一部分中，例如，存储在标记语言文档中的一个或多个脚本；专用于所讨论的程序的单个文件；或者多个协调文件，例如，存储一个或多个模块、子程序或部分代码的文件。

[0268] 本申请中描述的处理和逻辑流程可由一台或多台计算机执行一个或多个计算机程序进行，以进行通过对输入数据进行运算并生成输出的操作。处理和逻辑流程也可由例如FPGA、ASIC、GPU等的专用逻辑电路或专用逻辑电路与一个或多个编程计算机的组合来执行。

[0269] 适合于执行计算机程序的计算机可以是基于通用和/或专用微处理器，或任何其他种类的中央处理单元。通常，中央处理单元将从只读存储器和/或随机存取存储器接收指令和数据。计算机的元件可包括用于执行指令的中央处理单元和用于存储指令和数据的一个或多个存储设备。中央处理单元和存储器可以补充有专用逻辑电路或集成在专用逻辑电路中。

[0270] 通常，计算机将耦接到至少一个非暂态计算机可读存储介质(也称为计算机可读存储器)上。耦接到计算机的存储介质可以是计算机的内部组件(例如，集成硬盘驱动器)或外部组件(例如，通用串行总线(USB)硬盘驱动器或通过网络访问的存储系统)。存储介质的示例可包括：例如，磁盘、磁光盘或光盘、固态硬盘、诸如云存储系统的网络存储资源或其他类型的存储介质。但是，计算机不需要这样的设备。此外，计算机可嵌入在另一个设备中，例如移动电话、个人数字助理(PDA)、移动音频或视频播放器、游戏控制台、全球定位系统(GPS)接收器或便携式存储设备，例如通用串行总线(USB)闪存驱动器，这里仅举几例。

[0271] 为了提供针对与用户的交互，本申请实施方式中描述的主题可在计算机上实现或被配置为与该计算机通信，该计算机具有：显示设备，例如，LCD(liqiud crystal display)监视器，用于向用户显示信息；以及输入设备，用户可以通过该输入设备向该计算机提供输入，例如键盘和例如鼠标、轨迹球或触摸板等的指针设备。其他类型的设备也可用于提供与用户的交互；例如，提供给用户的反馈可是任何形式的感官反馈，例如视觉反馈、听觉反馈或触觉反馈；并且可以接收来自用户的任何形式的输入，包括声学、语音或触觉输入。此外，计算机可通过向用户使用的设备发送文档并从用户使用的设备接收文档来与用户交互；例如，通过向客户设备上的网络浏览器发送网络页面以响应从网络浏览器接收到的请求，或通过与例如智能电话或电子平板电脑等的用户设备上运行的应用程序(app)进行交互。此外，计算机可以通过向个人设备(例如，运行消息应用的智能电话)轮流发送文本消息或其他形式的消息并接收来自用户的响应消息来与用户交互。

[0272] 本申请结合系统、装置和计算机程序组件使用术语“配置为”。对于被配置为执行特定操作或动作的一个或多个计算机的系统，意味着该系统已经在其上安装了在运行中促使该系统执行所述操作或动作的软件、固件、硬件或它们的组合。对于被配置为执行特定操作或动作的一个或多个计算机程序，意味着一个或多个程序包括当由数据处理装置执行时使该装置执行操作或动作的指令。对于被配置为执行特定操作或动作的专用逻辑电路，意味着该电路具有执行操作或动作的电子逻辑。

[0273] 虽然本申请包含许多具体实施细节，但这些实施细节不应被解释为对权利要求本身所限定的要求保护的范围的限制，而是作为对特定实施例的具体特征的描述。在本申请多个单独实施方式的上下文中描述的多个特定特征也可在单个实施方式中组合实现。相反，在单个实施方式的上下文中描述的各种特征也可以单独地或以任何合适的子组合在多个实施方式中实现。此外，尽管可以将上述特征描述为以某种组合方式实现，此外，尽管上面的特征可以描述为以某些组合起作用并且甚至最初如此要求保护，但是在一些情况下，可以从要求保护的组合中删除来自该组合的一个或多个特征，并且可以要求保护指向子组合或子组合的变体。

[0274] 类似地，虽然以特定顺序在附图中描绘了操作并且在权利要求中叙述了操作，但是这不应该被理解为：为了达到期望的结果，要求以所示的特定顺序或依次执行这些操作，或者要求执行所有示出的操作。在一些情况下，多任务和并行处理可能是有利的。此外，上述实施例中的各种系统模块和组件的划分不应被理解为所有实施例中都要求如此划分，而应当理解，所描述的程序组件和系统通常可以一起集成在单个软件产品或者打包成多个软件产品。

[0275] 已经描述了特定主题的实施方式。其他实施方式在以下权利要求的范围内。例如，权利要求中记载的动作可以按照不同的顺序执行并且仍然实现期望的结果。作为一个示例，附图中所描绘的处理无需要求所示的特定顺序或次序来实现期望的结果。在一些情况下，多任务和并行处理可能是有利的。

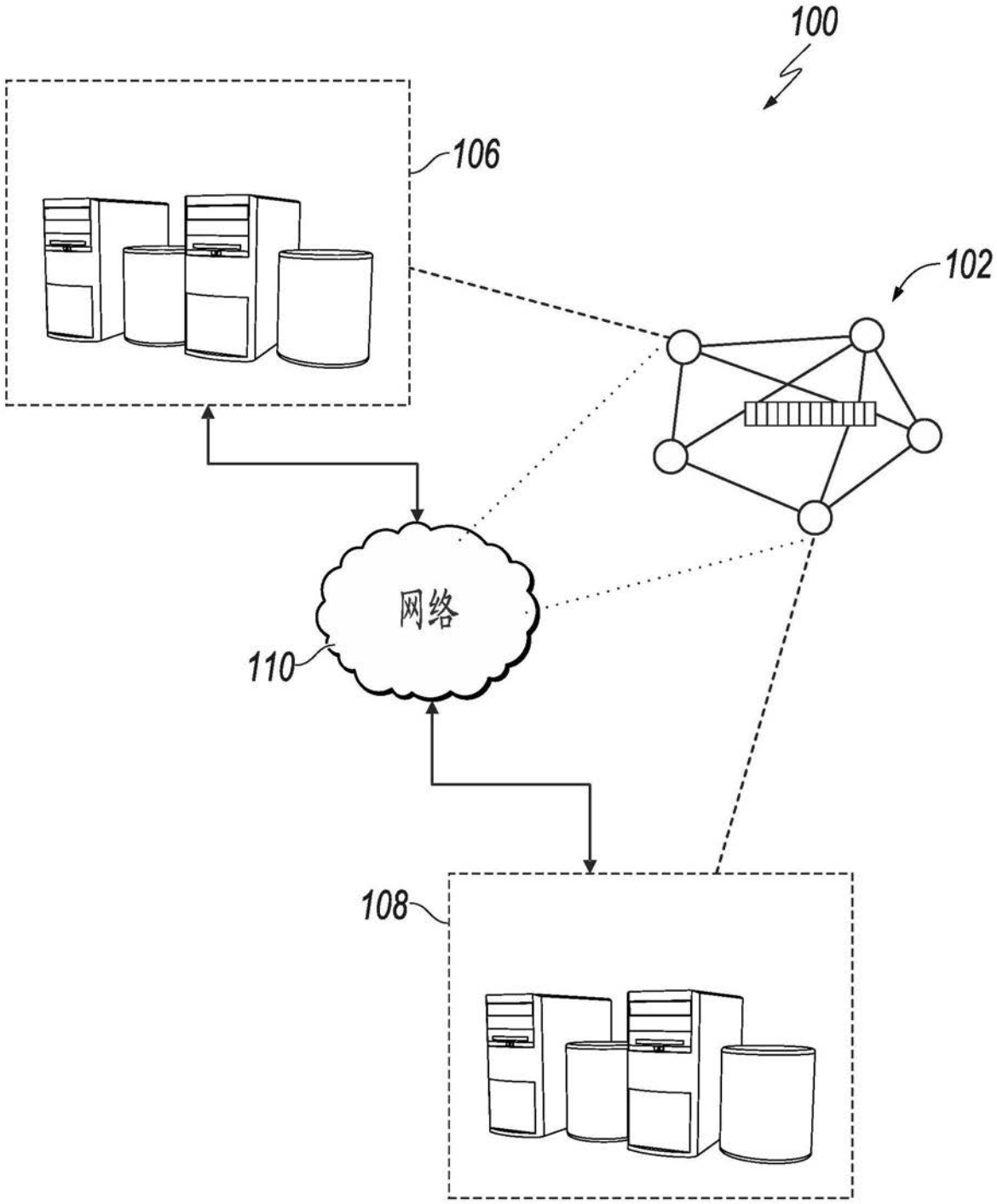


图1

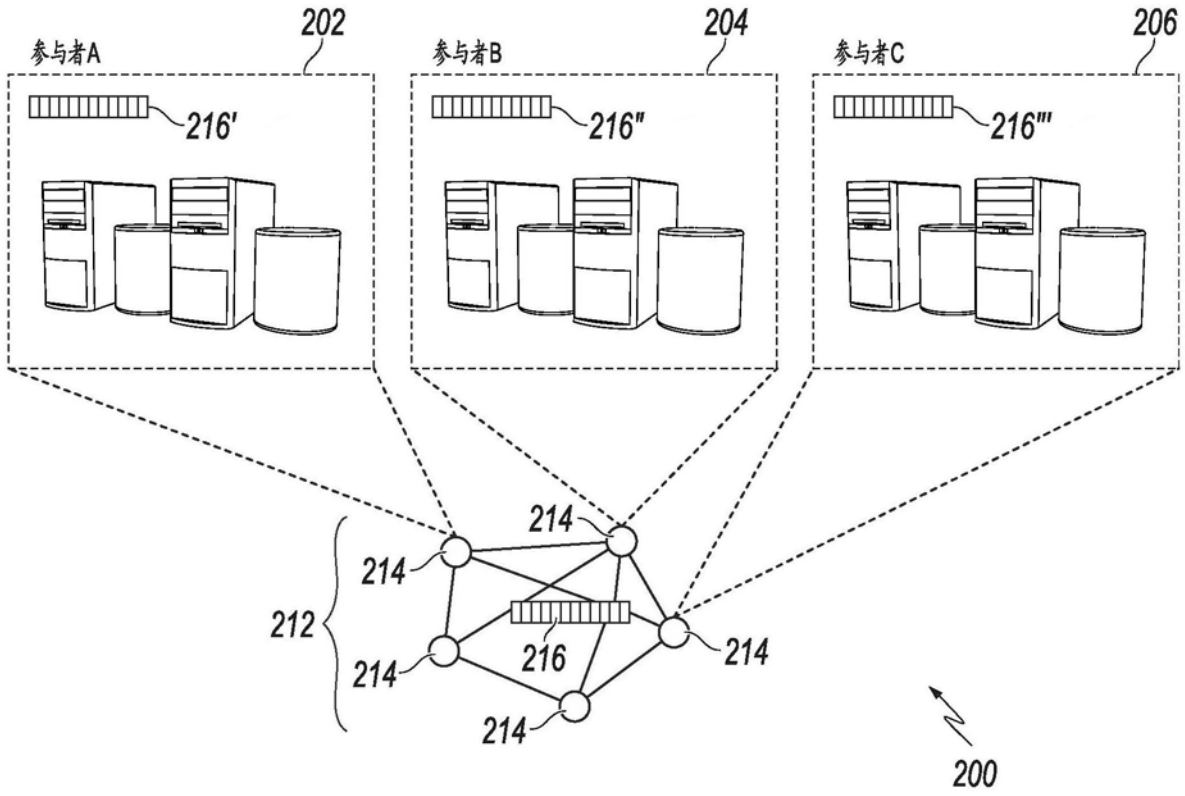

参与者A
202
216'
参与者B
204
216"
参与者C
206
216'''
214
214
212
214
216
214
214
200

图2

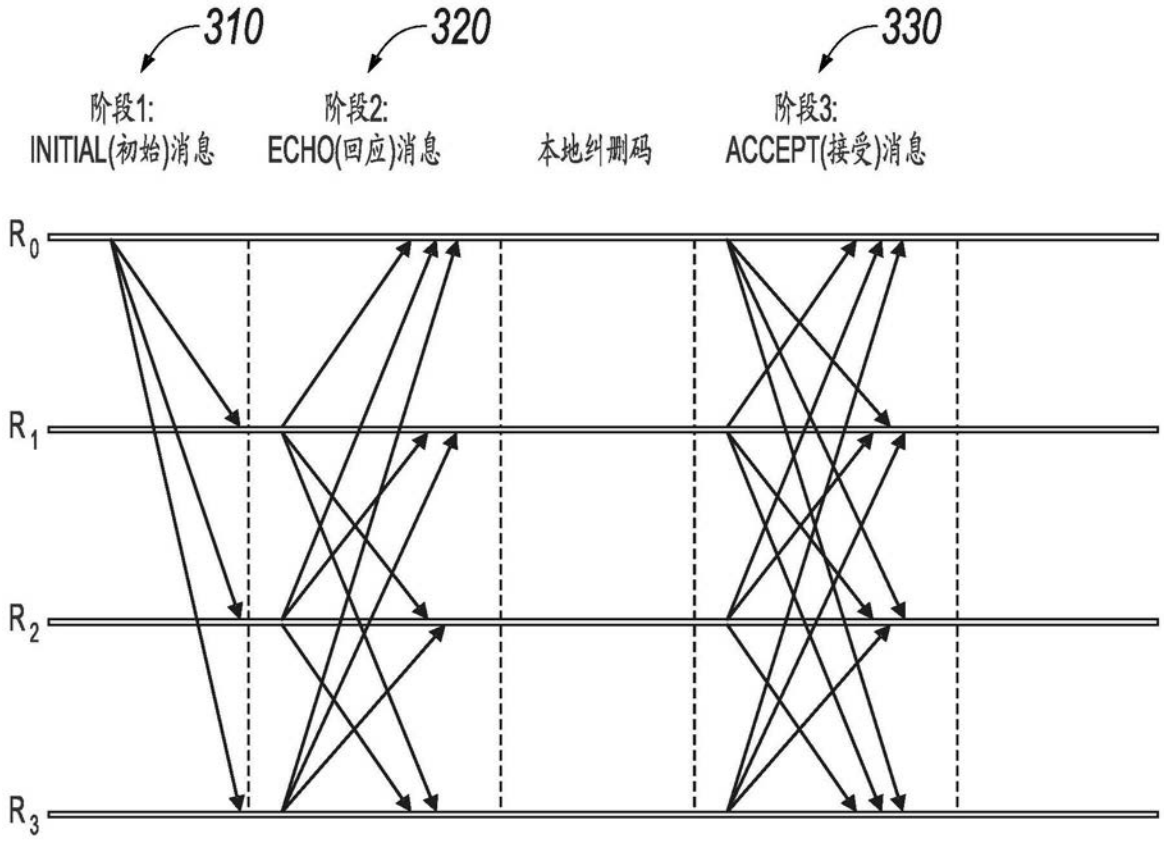

310
320
330
阶段1:
INITIAL(初始)消息
阶段2:
ECHO(回应)消息
本地纠删码
阶段3:
ACCEPT(接受)消息
R0
R1
R2
R3

图3

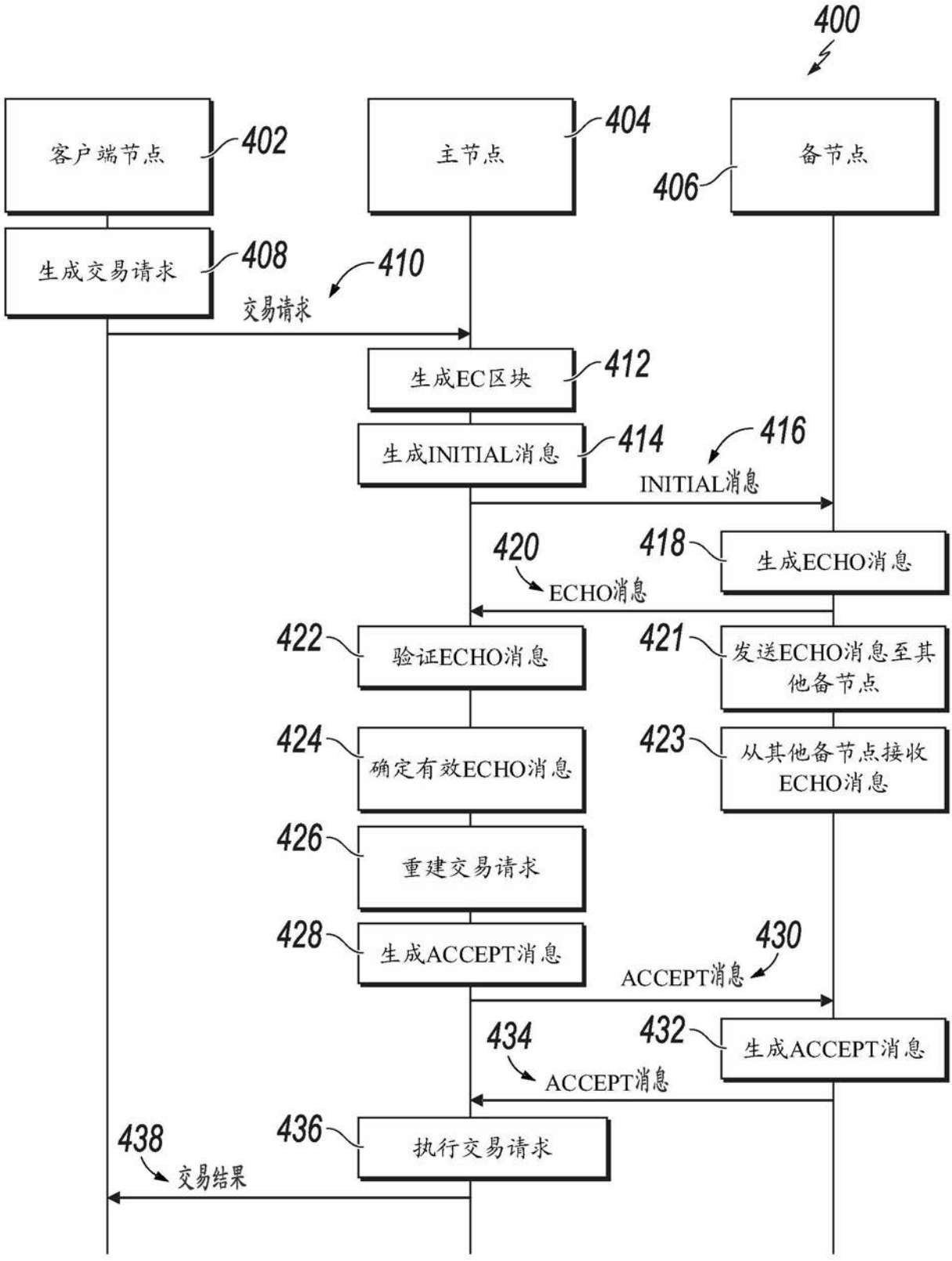

400
客户端节点
402
主节点
404
备节点
406
生成交易请求
408
410
交易请求
生成EC区块
412
生成INITIAL消息
414
416
INITIAL消息
420
418
生成ECHO消息
ECHO消息
422
验证ECHO消息
421
发送ECHO消息至其他备节点
424
确定有效ECHO消息
423
从其他备节点接收ECHO消息
426
重建交易请求
428
生成ACCEPT消息
430
ACCEPT消息
434
432
生成ACCEPT消息
ACCEPT消息
436
执行交易请求
438
交易结果

图4

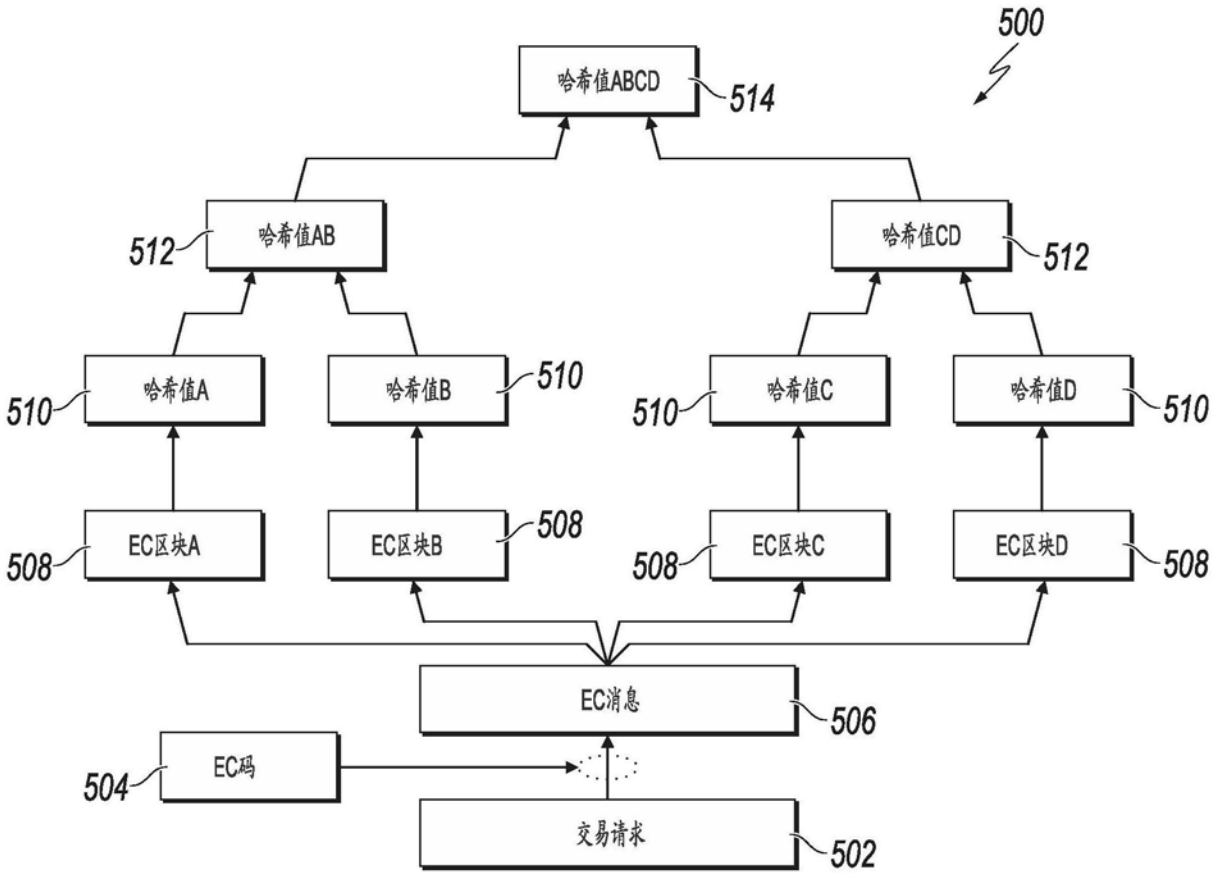


图5



Initial<epoch, tx_root_hash, ec_block_hash, ec_block, seq, j>
Echo<epoch, tx_root_hash, ec_block_hash, ec_block, seq, sign_proof, j>
Accept<epoch, tx_root_hash, seq, sign_proofs, j>

图6

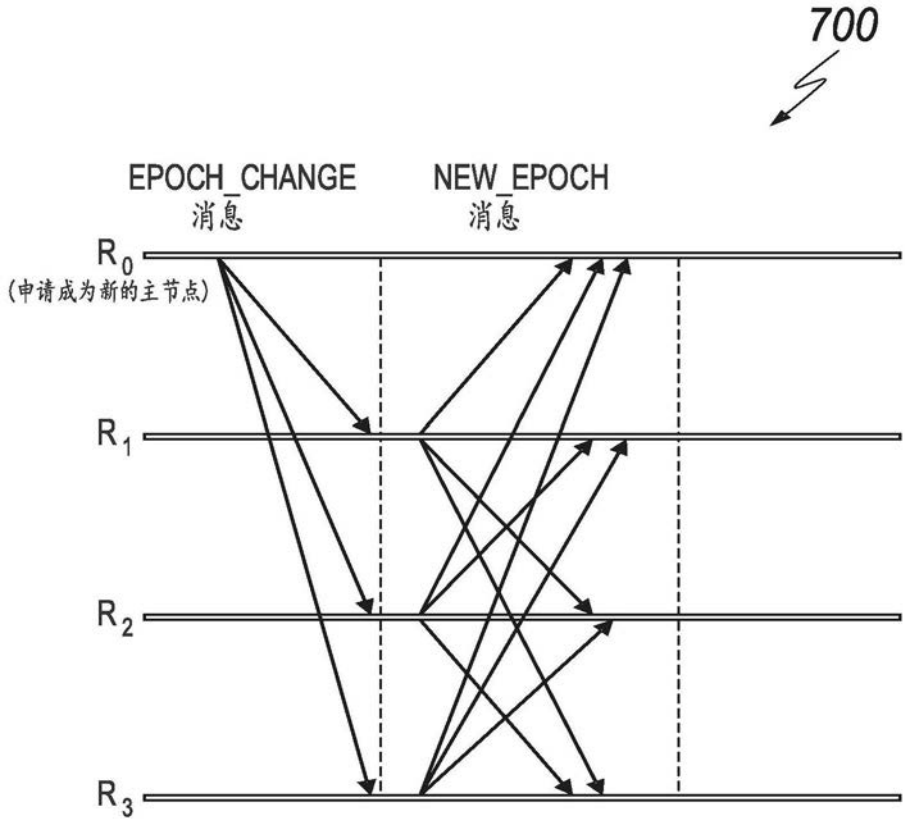


图7

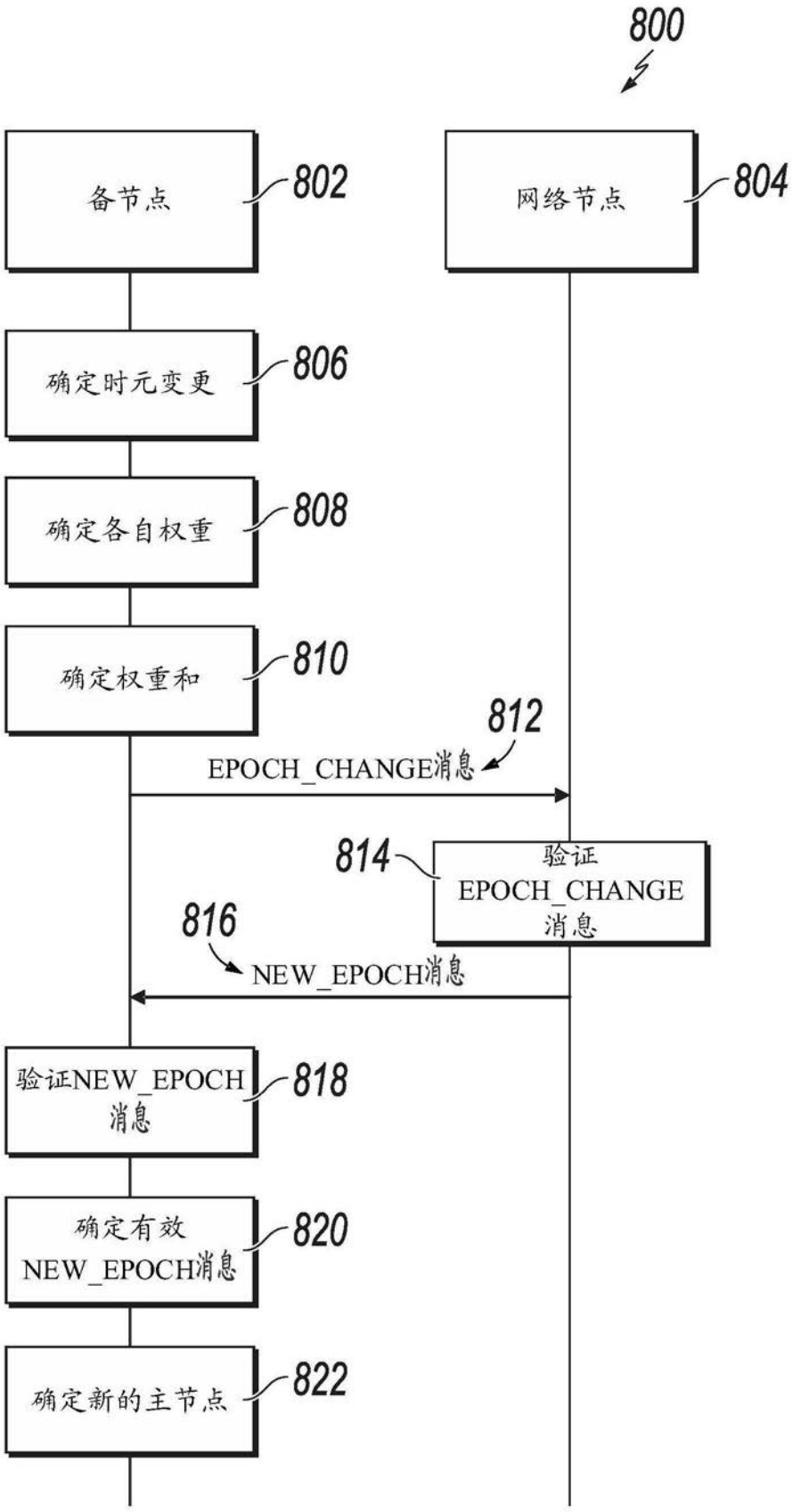

800
备节点
802
网络节点
804
确定时元变更
806
确定各自权重
808
确定权重和
810
812
EPOCH_CHANGE消息
814
验证
EPOCH_CHANGE
消息
816
NEW_EPOCH消息
验证NEW_EPOCH
消息
818
确定有效
NEW_EPOCH消息
820
确定新的主节点
822

图8



Epoch_change <weight, epoch+1, Echo{}, Accept{}, j>
Echo {<epoch, seq, sign_proof, j>}
Accept {<epoch, tx_root_hash, sign_proofs, seq, j>}
New_epoch <epoch+1, i, j, seq, ec_digest>

图9

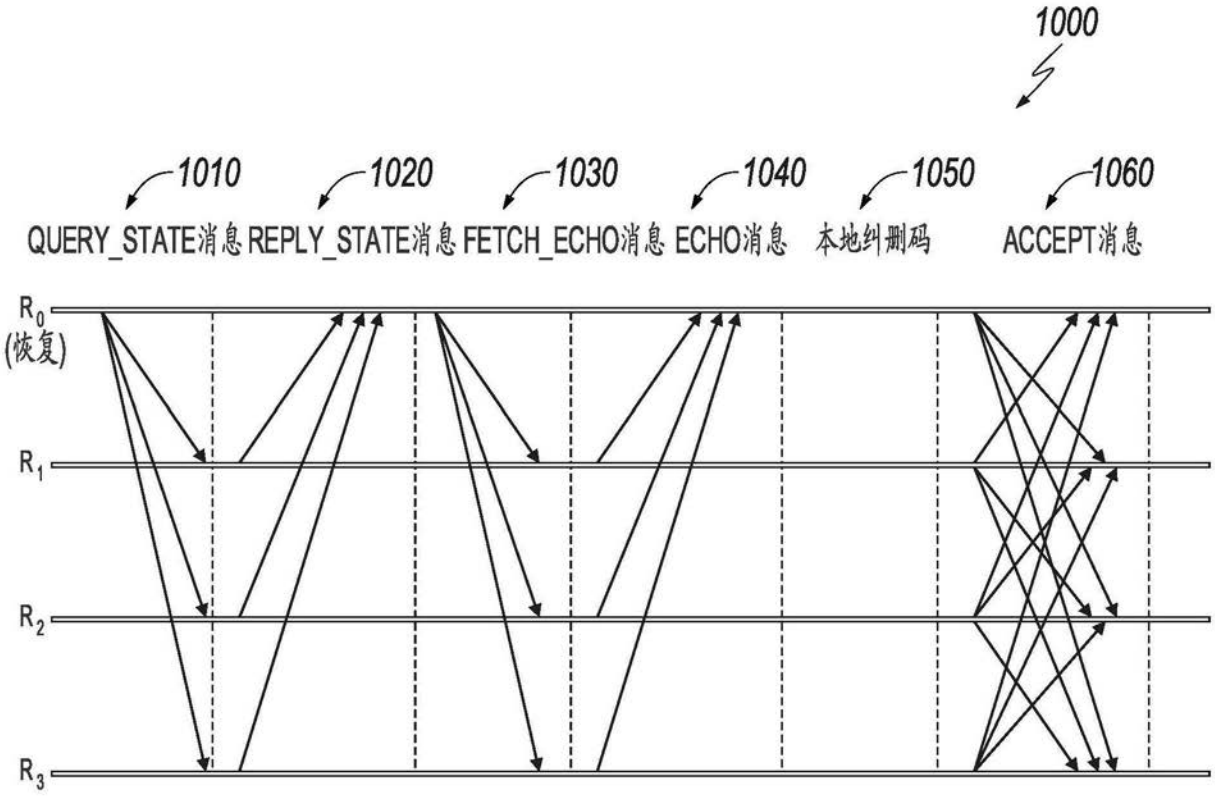


图10

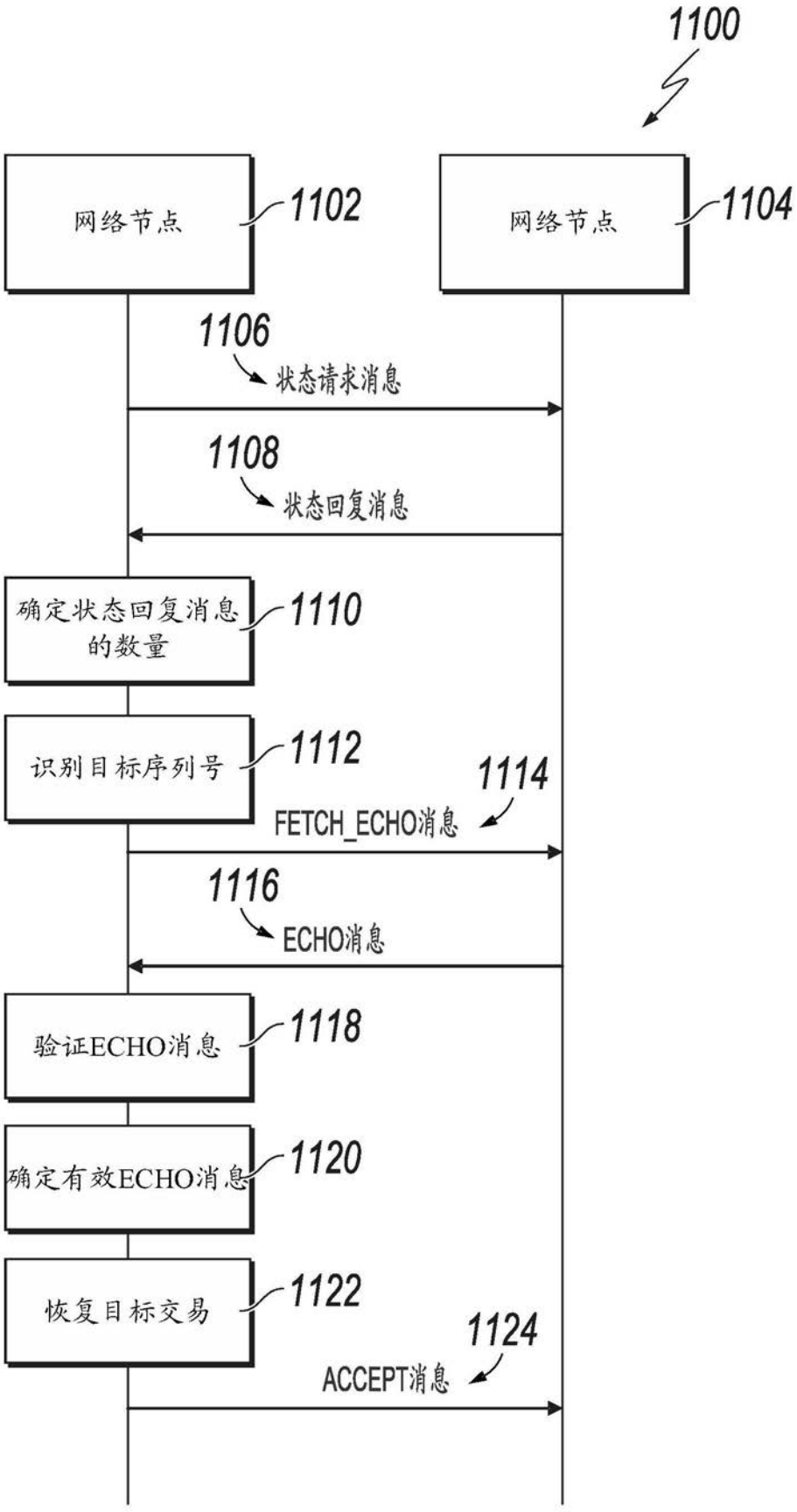

1100
网络节点
1102
网络节点
1104
1106
状态请求消息
1108
状态回复消息
确定状态回复消息的数量
1110
识别目标序列号
1112
1114
FETCH_ECHO消息
1116
ECHO消息
验证ECHO消息
1118
确定有效ECHO消息
1120
恢复目标交易
1122
1124
ACCEPT消息

图11

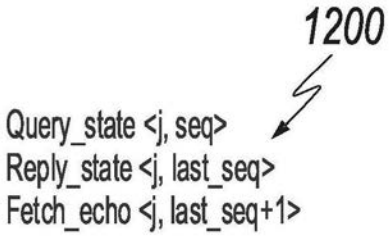

1200
Query_state <j, seq>
Reply_state <j, last_seq>
Fetch_echo <j, last_seq+1>

图12

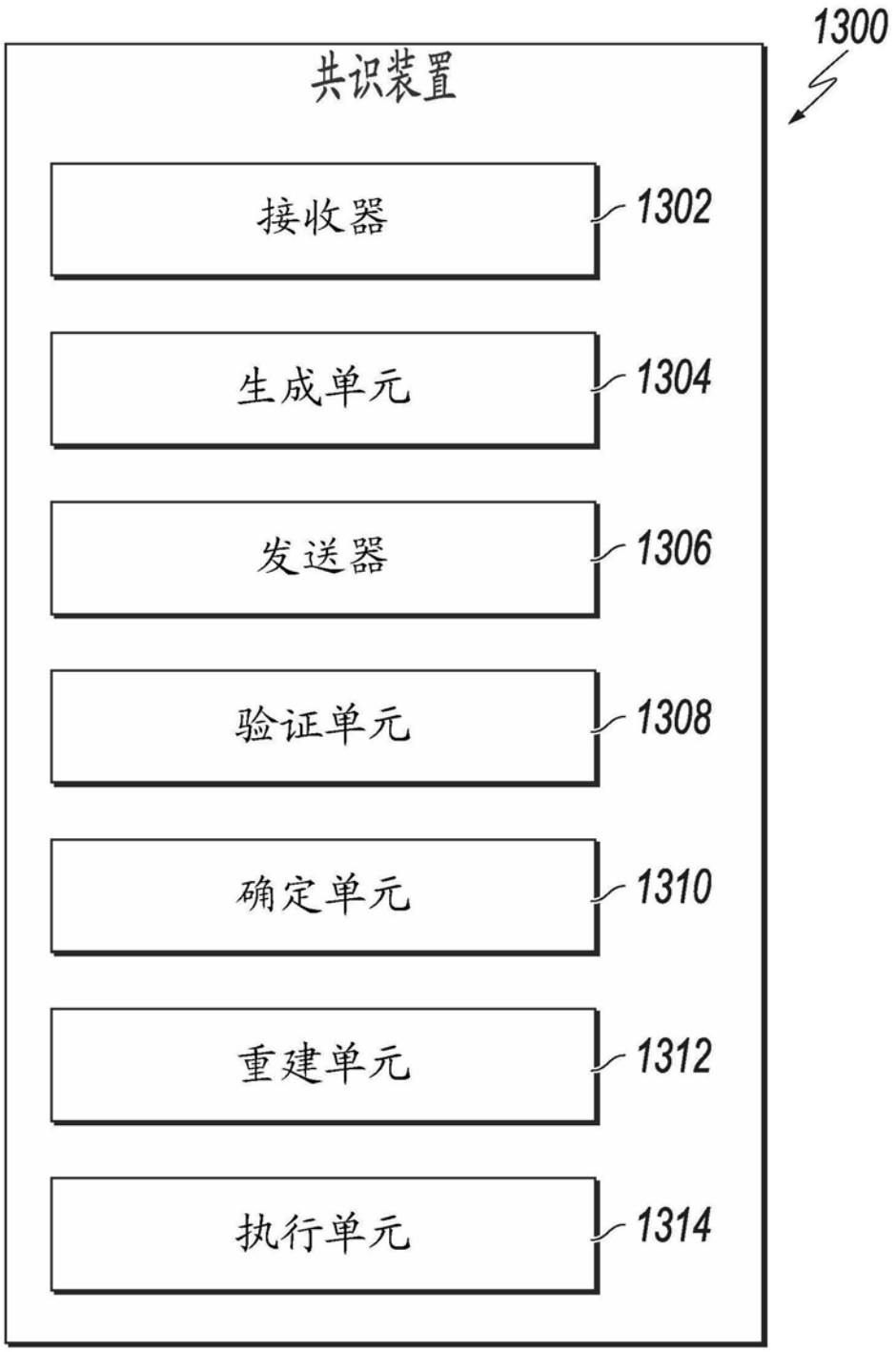

1300
共识装置
接收器
1302
生成单元
1304
发送器
1306
验证单元
1308
确定单元
1310
重建单元
1312
执行单元
1314

图13

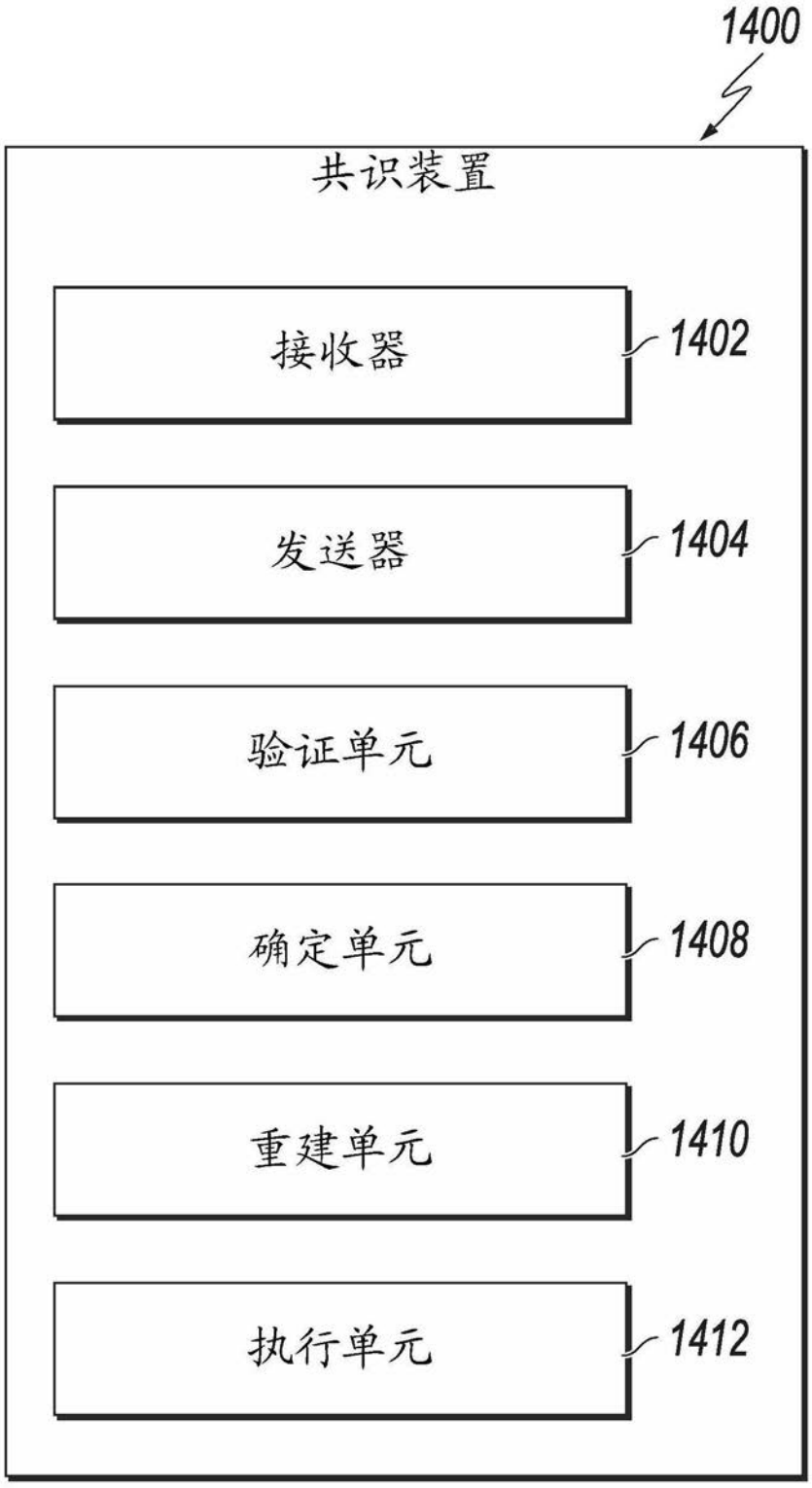

1400
共识装置
接收器
1402
发送器
1404
验证单元
1406
确定单元
1408
重建单元
1410
执行单元
1412

图14

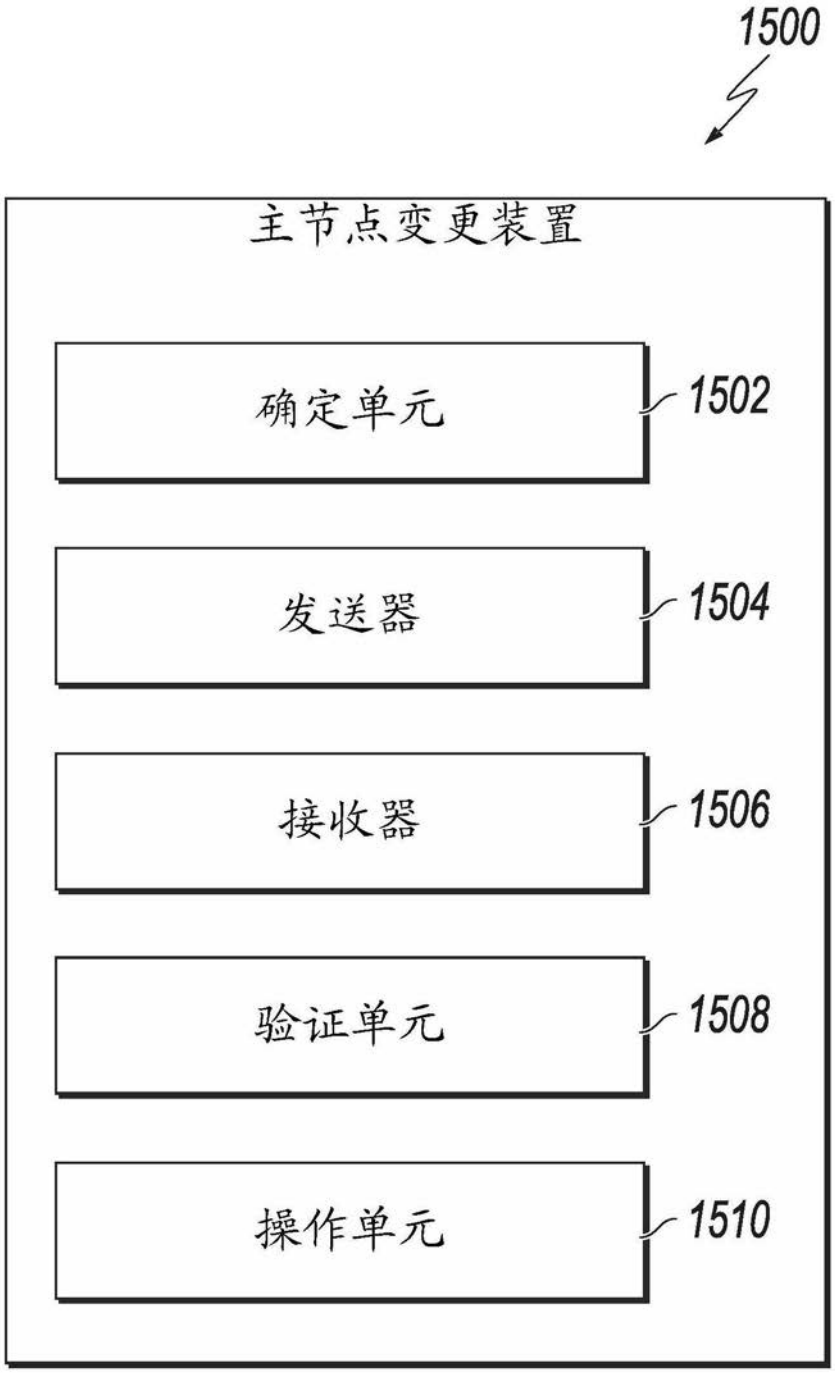


图15

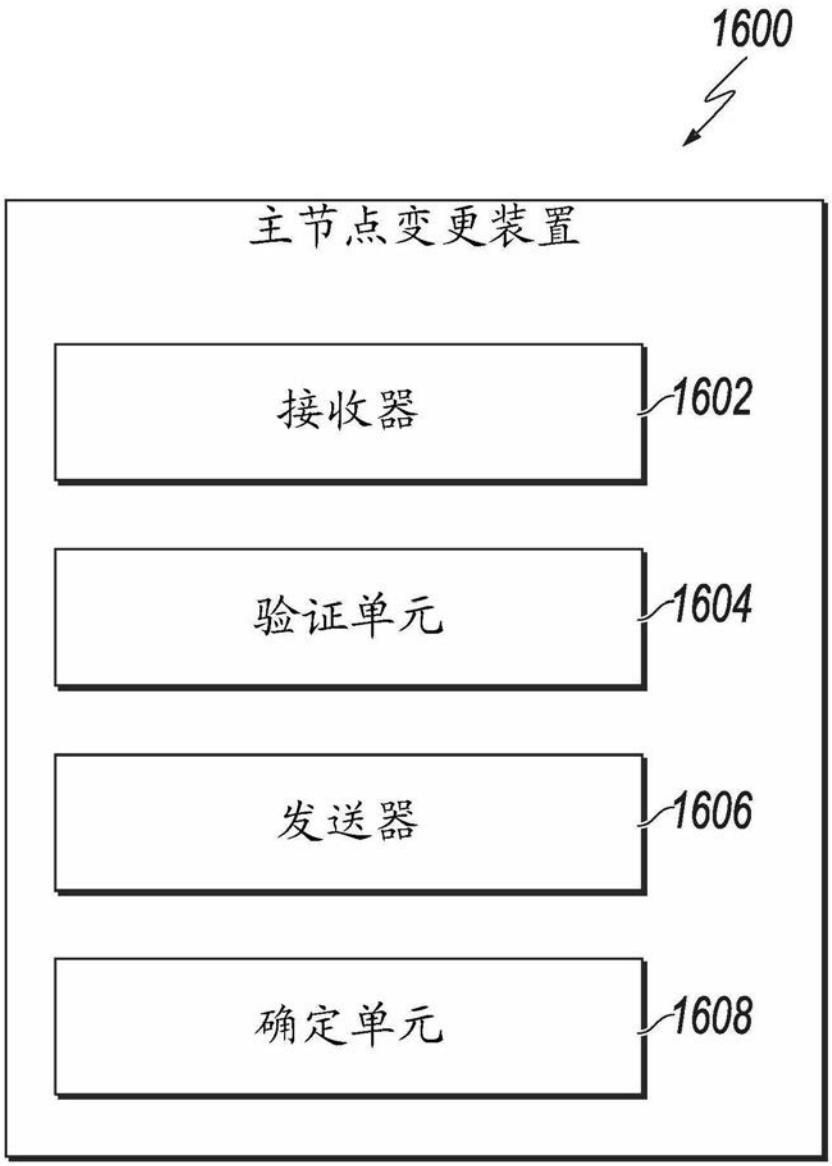


图16

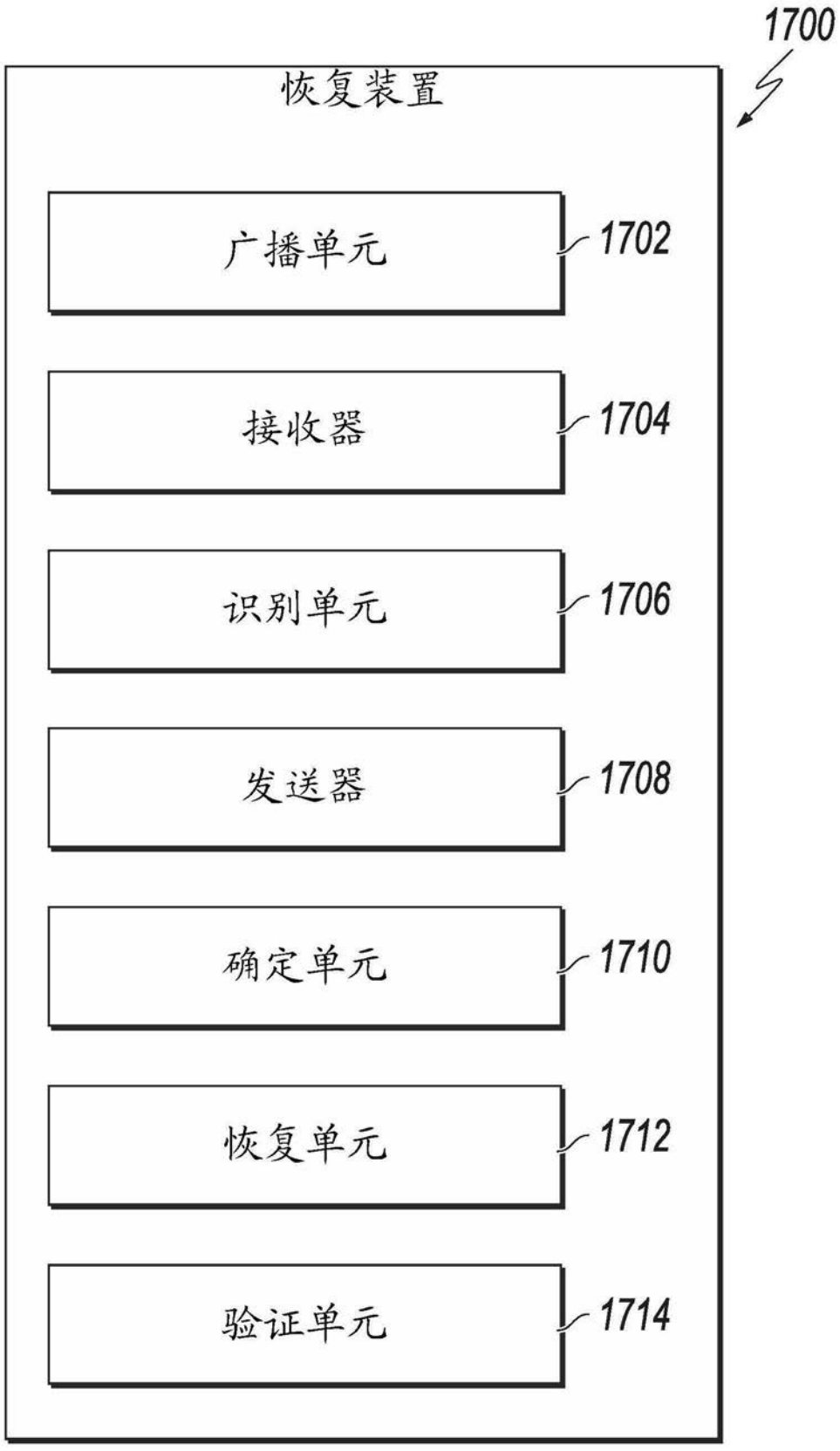

1700
恢复装置
广播单元
1702
接收器
1704
识别单元
1706
发送器
1708
确定单元
1710
恢复单元
1712
验证单元
1714

图17