(19)国家知识产权局

(12)发明专利

(10)授权公告号 CN 111902817 B
(45)授权公告日 2024.04.09

(21)申请号 201980004448.1
(22)申请日 2019.08.20
(65)同一申请的已公布的文献号
申请公布号 CN 111902817 A
(43)申请公布日 2020.11.06
(85)PCT国际申请进入国家阶段日
2020.03.10
(86)PCT国际申请的申请数据
PCT/CN2019/101575 2019.08.20
(87)PCT国际申请的公布数据
WO2019/228550 EN 2019.12.05
(73)专利权人 创新先进技术有限公司
地址 开曼群岛大开曼岛乔治镇医院路27号开曼企业中心
(72)发明人 卓海振
(74)专利代理机构 北京博思佳知识产权代理有限公司 11415
专利代理师 艾佳
(51)Int.Cl.
G06F 21/64(2006.01)
(56)对比文件
CN 109274717 A,2019.01.25
CN 109359096 A,2019.02.19
CN 109933592 A,2019.06.25
WO 2019101230 A2,2019.05.31
审查员 张琳琳

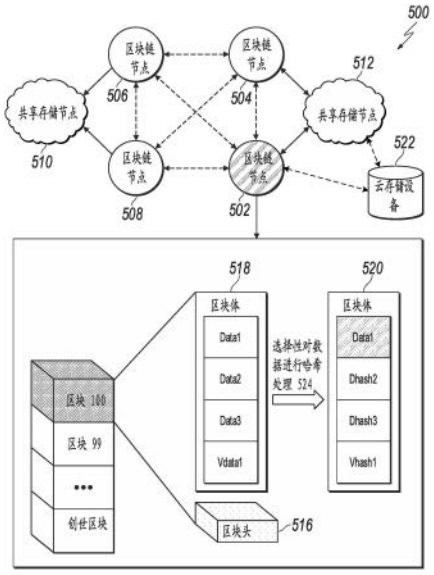
权利要求书2页 说明书18页 附图8页

(54)发明名称

基于共享节点和纠错编码的区块链数据存储

(57)摘要

本文公开了用于存储区块链数据的方法、系统和装置,包括编码在计算机存储介质上的计算机程序。方法之一包括:确定与区块链的当前区块相关联的区块数据和当前状态数据;将所述当前状态数据发送到所述区块链网络的一个或多个共享存储节点;对所述区块数据执行纠错编码以生成编码区块数据;基于一个或多个预定规则将所述编码区块数据划分为多个数据集;基于所述一个或多个预定规则存储所述多个数据集中的一个或多个数据集;对所述多个数据集中的每个剩余数据集进行哈希处理,以生成与所述多个数据集中的剩余数据集相对应的一个或多个哈希值;存储所述一个或多个哈希值和所述当前状态数据。



CN 111902817 B

1.一种计算机实现的用于存储区块链数据的方法,所述方法包括:

区块链网络中的区块链节点确定与区块链的当前区块相关联的区块数据和当前状态数据;

所述区块链节点将所述当前状态数据发送到所述区块链网络的一个或多个共享存储节点,其中,所述区块链节点存储所述当前状态数据,并且所述一个或多个共享存储节点存储与所述区块链的每个区块相关联的历史状态数据;

所述区块链节点对所述区块数据执行纠错编码以生成编码区块数据;

所述区块链节点基于一个或多个预定规则将所述编码区块数据划分为多个数据集;

所述区块链节点基于所述一个或多个预定规则存储所述多个数据集中的一个或多个数据集;

所述区块链节点对所述多个数据集中的每个剩余数据集进行哈希处理,以生成与所述多个数据集中的所述剩余数据集相对应的一个或多个哈希值;以及

所述区块链节点存储所述一个或多个哈希值和所述当前状态数据。

2.如权利要求1所述的方法,还包括:

所述区块链节点向所述一个或多个共享存储节点之一发送哈希值,以检索包括在所述历史状态数据中的账户状态;

所述区块链节点接收针对发送所述哈希值的响应中的所述账户状态;以及

所述区块链节点基于所述哈希值验证所述账户状态是所述区块链的一部分。

3.如权利要求1所述的方法,其中,所述区块链网络包括至少f+1个共享存储节点和不超过2f+2个共识节点,f是在所述区块链网络中能够容忍的故障共享存储节点和共识节点的最大数量。

4.如前述任一项权利要求所述的方法,其中,所述一个或多个共享存储节点是通过从所述区块链网络中的所有3f+1、3f+2或3f+3个节点接收到2f+1个投票选举出的,f是所述区块链网络中能够容忍的故障共享存储节点和共识节点的最大数量。

5.如权利要求1所述的方法,其中,所述当前状态数据和所述历史状态数据被存储为固定深度默克尔树。

6.如权利要求1所述的方法,其中,所述纠错编码是纠删编码。

7.如权利要求1所述的方法,其中,所述区块链节点是第一区块链节点,并且所述方法还包括:

确定所述第一区块链节点未存储执行智能合约所需的至少部分所述区块数据;

基于所述一个或多个预定规则识别存储所述多个数据集中的至少一个所述剩余数据集的第二区块链节点以及与所述多个数据集中的所述至少一个剩余数据集相对应的至少一个哈希值;以及

将所述至少一个哈希值发送至所述第二区块链节点,以检索所述多个数据集中的所述至少一个剩余数据集。

8.如权利要求7所述的方法,其中,所述至少一个哈希值是至少一个第一哈希值,所述方法还包括:

响应于发送所述至少一个第一哈希值,从所述第二区块链节点接收至少一个数据集;

对该至少一个数据集进行哈希处理以生成至少一个第二哈希值;以及

如果所述至少一个第一哈希值与所述至少一个第二哈希值相同,则确定所述至少一个数据集是真实的。

9.如权利要求8所述的方法,还包括:

基于所述一个或多个预定规则识别存储所述多个数据集中的至少一个所述剩余数据集的第三区块链节点以及与所述多个数据集中的所述至少一个剩余数据集相对应的至少一个第三哈希值;

将所述至少一个第三哈希值发送至所述第三区块链节点,以检索所述多个数据集中的所述至少一个剩余数据集;

响应于发送所述至少一个第三哈希值,从所述第三区块链节点接收至少一个数据集;

对该至少一个数据集进行哈希处理以生成至少一个第四哈希值;以及

如果所述至少一个第三哈希值与所述至少一个第四哈希值相同,则确定该至少一个数据集是真实的。

10.如权利要求1所述的方法,其中,所述一个或多个预定规则包括:

一个或多个指令,用于基于区块链网络中的区块链节点的数量将所述编码区块数据划分为所述多个数据集,和

与所述多个数据集中的一个或多个数据集要被所述区块链节点中的每个区块链节点存储或进行哈希处理有关的分配。

11.如权利要求1所述的方法,还包括:

将所述当前区块存储为预写日志WAL文件;以及

报告所述当前区块被存储并准备用于达成共识。

12.如权利要求1所述的方法,还包括:

将所述区块数据存储在所述区块链节点的缓冲器中,并且响应于识别出所述缓冲器已满而执行所述纠错编码。

13.一种通信共享区块链数据的系统,包括:

一个或多个处理器;以及

耦接到所述一个或多个处理器且其上存储有指令的一个或多个计算机可读存储器,所述指令能由所述一个或多个处理器执行以执行权利要求1至12中任一项所述的方法。

14.一种用于通信共享区块链数据的装置,所述装置包括用于执行权利要求1至12中任一项所述的方法的多个模块。

基于共享节点和纠错编码的区块链数据存储

技术领域

[0001] 本文涉及基于共享节点和纠错编码的区块链数据存储。

背景技术

[0002] 分布式账本系统(DLS),也可称为共识网络和/或区块链网络,使得参与的实体能够安全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。区块链网络类型的示例可以包括公共区块链网络、私有区块链网络和联盟区块链网络。为选定的实体群组提供联盟区块链网络,所述实体控制共识处理,并且所述联盟区块链网络包括访问控制层。

[0003] 基于区块链的程序可以由诸如以太坊的分布式计算平台执行。例如,以太坊虚拟机(EVM)为以太坊中的智能合约提供运行环境。以太坊区块链可以被视为基于交易的状态机。以太坊中的状态数据可以集合成一个被称为世界状态的全局共享状态。世界状态包括以太坊账户地址和账户状态之间的映射。世界状态可以存储在诸如默克尔帕特里夏树(Merkle Patricia tree,MPT)的数据结构中。

[0004] 除了状态数据,区块链网络还可以存储其他类型的数据,例如区块数据和索引数据。区块数据可以包括区块头和区块体。区块头可以包括特定区块的身份信息,并且区块体可以包括用该区块确认的交易。当越来越多的交易进入区块链时,状态数据和区块数据的大小可能会变得非常大。在一些DLS中,即使不频繁访问某些旧的区块数据或状态数据,每个节点也会存储整个区块链副本,这会占用大量存储空间。在一些DLS中,一些共享节点存储整个区块链副本,并与其他可能导致“数据不对等”的区块链节点共享区块链数据。

[0005] 因此,期望在保持数据对等性和数据处理效率的同时减少存储在DLS中的节点上的数据量。

发明内容

[0006] 本文描述了用于通信和共享区块链数据的技术。这些技术一般涉及:区块链网络中的区块链节点确定与区块链的当前区块相关联的区块数据和当前状态数据;区块链节点将当前状态数据发送到区块链网络的一个或多个共享存储节点,其中,区块链节点存储当前状态数据,并且一个或多个共享存储节点存储与区块链的每个区块相关联的历史状态数据;区块链节点执行区块数据的纠错编码以生成编码区块数据;区块链节点基于一个或多个预定规则将编码区块数据划分为多个数据集;区块链节点基于一个或多个预定规则存储多个数据集中的一个或多个数据集;区块链节点对多个数据集中的每个剩余数据集进行哈希处理,以生成与多个数据集中的剩余数据集相对应的一个或多个哈希值;以及区块链节点存储一个或多个哈希值和当前状态数据。

[0007] 本文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,当所述指令由所述一个或多个处理器执行时,所述指令将促使所述一个或多个处理器按照本文提供的方法的实施例执行操作。

[0008] 本文还提供了用于实施本文提供的所述方法的系统。该系统包括一个或多个处理器以及耦接到所述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,当所述指令由所述一个或多个处理器执行时,所述指令将促使所述一个或多个处理器按照本文提供的方法的实施例执行操作。

[0009] 应了解,依据本文的方法可以包括本文描述的方面和特征的任意组合。也就是说,根据本文的方法不限于本文具体描述的方面和特征的组合,还包括所提供的方面和特征的任意组合。

[0010] 以下在附图和描述中阐述了本文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,本文的其他特征和优点将显现。

附图说明

[0011] 图1描绘了可用于执行本文实施例的环境的示例。

[0012] 图2描绘了根据本文实施例的架构的示例。

[0013] 图3描绘了根据本文实施例的固定深度默克尔树(fixed depth Merkle tree,FDMT)数据结构的示例。

[0014] 图4描绘了根据本文实施例的用于存储区块链数据的数据库的示例。

[0015] 图5描绘了根据本文实施例的状态数据编码和哈希处理的示例。

[0016] 图6描绘了根据本文实施例的数据存储布置的示例。

[0017] 图7描绘了可以根据本文实施例执行的处理的示例。

[0018] 图8描绘了根据本文实施例的装置的模块的示例。

[0019] 在各个附图中,相同的附图标记和名称表示相同的元件。

具体实施方式

[0020] 本文描述了用于通信和共享区块链数据的技术。这些技术一般涉及:区块链网络中的区块链节点确定与区块链的当前区块相关联的区块数据和当前状态数据;区块链节点将当前状态数据发送到区块链网络的一个或多个共享存储节点,其中,区块链节点存储当前状态数据,并且一个或多个共享存储节点存储与区块链的每个区块相关联的历史状态数据;区块链节点对区块数据执行纠错编码以生成编码区块数据;区块链节点基于一个或多个预定规则将编码区块数据划分为多个数据集;区块链节点基于一个或多个预定规则存储多个数据集中的一个或多个数据集;区块链节点对多个数据集中的每个剩余数据集进行哈希处理,以生成与多个数据集中的剩余数据集相对应的一个或多个哈希值;以及区块链节点存储一个或多个哈希值和当前状态数据。

[0021] 本文中描述的技术产生若干技术效果。例如,主题的实施例可以允许节省区块链节点的存储资源,而不显著地降低计算效率。通过仅保存ECC编码的区块数据的一部分和与剩余的ECC编码的区块数据相对应的哈希值,可以减少整个区块链网络的存储空间消耗。

[0022] 在一些实施例中,代替存储整个区块,区块链节点可以存储ECC编码的区块的所选部分和与剩余的编码区块相对应的哈希值。即使区块链节点从故障的区块链节点接收到非真实数据,只要非真实数据的百分比小于或等于ECC允许的错误位或丢失位的最大比例,该区块数据也可以被恢复。

[0023] 此外,由于大多数历史状态数据是不频繁被使用的“冷”数据,因此通过仅将“冷”状态数据保存在共享存储节点中,可以显著提高整个区块链网络的存储空间使用率。对于具有N个节点的区块链网络,其中,N等于3f+1、3f+2或3f+3,f是故障共识节点的最大数量,(N-f-1)/N个区块链节点仅需要将“热”数据存储为当前状态树,而无需将“冷”和“热”数据都存储为历史状态树。对于将f+1个节点用作共享存储节点来存储历史状态树的N个节点的区块链网络,最多可以容忍f个故障共识节点。换句话说,节省存储空间不会损害数据可靠性。由于f+1个共享存储节点确保了系统的可靠性,因此可以提高数据安全性,并且相对独立于底层服务平台的安全级别。

[0024] 为了提供本文的实施例的进一步的背景,并且如上所述,分布式账本系统(DLS),也可以被称为共识网络(例如,由点对点节点组成)和区块链网络,使得参与实体能够安全地并且不可篡改地进行交易,并存储数据。尽管术语区块链通常与特定网络和/或用例相关联,但是在不参考任何特定用例的情况下,本文中使用的区块链通常是指DLS。

[0025] 区块链是以交易不可篡改的方式存储交易的数据结构。因此,区块链上记录的交易是可靠且可信的。区块链包括一个或多个区块。链中的每个区块通过包含在链中紧邻其之前的前一区块的加密哈希值(cryptographic hash)链接到该前一区块。每个区块还包括时间戳、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的交易经哈希处理并编入默克尔(Merkle)树中。Merkle树是一种数据结构,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连接。此过程沿着树持续一直到整个树的根,在整个树的根处存储了代表树中所有数据的哈希值。声称是存储在树中的交易的哈希值可以通过确定其是否与树的结构一致而被快速验证。

[0026] 区块链是用于存储交易的去中心化或至少部分去中心化的数据结构,而区块链网络是通过广播、验证和确认交易等来管理、更新和维护一个或多个区块链的计算节点的网络。如上所述,区块链网络可作为公有区块链网络、私有区块链网络或联盟区块链网络被提供。在本文中参考联盟区块链网络进一步详细描述本文的实施例。然而,可以预期,本文的实施例可以在任何适当类型的区块链网络中实现。

[0027] 通常,联盟区块链网络在参与的实体中是私有的。在联盟区块链网络中,共识处理由授权的节点集控制,该节点集可以被称为共识节点,一个或多个共识节点由相应实体(例如,金融机构、保险公司)操作。例如,由10个实体(例如,金融机构、保险公司)组成的联盟可以操作联盟区块链网络,每个实体操作联盟区块链网络中的至少一个节点。

[0028] 在一些示例中,在联盟区块链网络内,提供全局区块链作为跨所有节点复制的区块链。也就是说,所有的共识节点相对于全局区块链处于完全共识状态。为了达成共识(例如,同意将区块添加到区块链),在联盟区块链网络内实施共识协议。例如,联盟区块链网络可以实施实用拜占庭容错(PBFT)共识,下面将进一步详细描述。

[0029] 图1是示出了可用于执行本文实施例的环境100的示例的示图。在一些示例中,环境100使得实体能够参与联盟区块链网络102。环境100包括计算系统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,并且连接网站、用户设备(例如,计算设备)和后端系统。在一些示例中,可以通过有线和/或无线通信链路来访问网络110。在一些示例中,网络110使得能够与联盟区块链网络102通信或在联盟区块链网络102内部通信成为可能。通常,网络110表示一个或多个通信网络。在一些情况下,计算

系统106、108可以是云计算系统(未示出)的节点,或者每个计算系统106、108可以是单独的云计算系统,其包括通过网络互连并且用作分布式处理系统的多个计算机。

[0030] 在所描绘的示例中,计算系统106、108可以各自包括能够作为节点参与至联盟区块链网络102中的任何适当的计算设备。计算设备的示例包括(但不限于)服务器、台式计算机、笔记本电脑、平板电脑和智能手机。在一些示例中,计算系统106、108承载用于与联盟区块链网络102交互的一个或多个由计算机实施的服务。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,例如,第一实体使用该交易管理系统管理其与一个或多个其他实体(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,例如,第二实体使用该交易管理系统管理其与一个或多个其他实体(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(Peer-to-Peer network),并且计算系统106、108分别提供参与联盟区块链网络102的第一实体和第二实体的节点。

[0031] 图2描绘了根据本文的实施例的架构200的示例。示例性概念架构200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、206。每个参与者(例如,用户、企业)参与到作为点对点网络提供的区块链网络212中,该点对点网络包括多个节点214,至少一些节点将信息不可篡改地记录在区块链216中。如图中进一步详述,尽管在区块链网络212中示意性地描述了单个区块链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。

[0032] 在所描绘的示例中,每个参与者系统202、204、206分别由参与者A、参与者B和参与者C提供或代表参与者A、参与者B和参与者C,并且在区块链网络中作为各自的节点214发挥作用。如这里所使用的,节点通常是指连接到区块链网络212且使相应的参与者能够参与到区块链网络中的个体系统(例如,计算机、服务器)。在图2的示例中,参与者对应于每个节点214。然而,可以预期,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,超文本传输协议安全(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。

[0033] 节点214可以在区块链网络212内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,作为将区块添加到区块链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,数据访问特权可以限制相应的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和206存储区块链216的相应的完整副本216’、216’’和216”’。

[0034] 区块链(例如,图2的区块链216)由一系列区块组成,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。尽管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(例如,资产、产品、服务、货币)的交换。交易数据不可篡改地存储在区块链中。也就是说,交易数据不能改变。

[0035] 在将交易数据存储在区块中之前,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)的过程。不可

能对哈希值进行去哈希处理(un-hash)以获取交易数据。哈希处理可确保即使交易数据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。也就是说,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。

[0036] 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,并对它们本身进行哈希处理以提供另一个哈希值。重复此过程,直到针对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Merkle根哈希值,并存储在区块的头中。任何交易中的更改都会导致其哈希值发生变化,并最终导致Merkle根哈希值发生变化。

[0037] 通过共识协议将区块添加到区块链。区块链网络中的多个节点参与共识协议,并竞相将区块添加到区块链中。这种节点称为共识节点。上面介绍的PBFT用作共识协议的非限制性示例。共识节点执行共识协议以将交易添加到区块链,并更新区块链网络的整体状态。

[0038] 更详细地,共识节点生成区块头,对区块中的所有交易进行哈希处理,并将所得的哈希值成对地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个哈希值(Merkle根哈希值)。将此哈希值添加到区块头中。共识节点还确定区块链中最新区块(即,添加到区块链中的最后一个区块)的哈希值。共识节点还向区块头添加随机数(nonce)和时间戳。

[0039] 通常,PBFT提供容忍拜占庭故障(例如,故障节点、恶意节点)的实用拜占庭状态机复制。这通过在PBFT中假设将发生故障(例如,假设存在独立节点故障和/或由共识节点发送的操纵消息)而实现。在PBFT中,以包括主共识节点和备共识节点的顺序提供共识节点。主共识节点被周期性地改变。通过由区块链网络内的所有共识节点对区块链网络的全局状态达成一致,将交易添加到区块链中。在该处理中,消息在共识节点之间传输,并且每个共识节点证明消息是从指定的对等节点(peer node)接收的,并验证在传输期间消息未被篡改。

[0040] 在PBFT中,共识协议是在所有共识节点以相同的状态开始的情况下分多个阶段提供的。首先,客户端向主共识节点发送调用服务操作(例如,在区块链网络内执行交易)的请求。响应于接收到请求,主共识节点将请求组播到备共识节点。备共识节点执行请求,并且各自向客户端发送回复。客户端等待直到接收到阈值数量的回复。在一些示例中,客户端等待直到接收到f+1个回复,其中f是区块链网络内可以容忍的错误共识节点的最大数量。最终结果是,足够数量的共识节点就将记录添加到区块链的顺序达成一致,并且该记录或被接受或被拒绝。

[0041] 在一些区块链网络中,用密码学来维护交易的隐私。例如,如果两个节点想要保持交易隐私,以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。加密处理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个密钥既进行加密(从明文生成密文)又进行解密(从密文生成明文)的加密过程。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。

[0042] 非对称加密使用密钥对,每个密钥对包括私钥和公钥,私钥仅对于相应节点是已

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。节点可以使用另一个节点的公钥来加密数据,并且该加密的数据可以使用其他节点的私钥被解密。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,并将加密数据发送给参与者B。参与者B可以使用其私钥来解密该加密数据(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该节点的私钥解密。

[0043] 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,而另一个节点可以根据参与者A的该数字签名来确认该消息是由该节点发送的。数字签名也可以用于确保消息在传输过程中不被篡改。例如,再次参考图2,参与者A将向参与者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作为加密哈希值的数字签名。参与者A将该数字签名附加到该消息上,并将该具有数字签名的消息发送给参与者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。如果哈希值相同,参与者B可以确认该消息确实来自参与者A,且未被篡改。

[0044] 如前所述,区块链网络可以存储不同类型的数据,例如状态数据、区块数据和索引数据。状态数据通常存储为内容寻址状态树,例如MPT或固定深度默克尔树(FDMT)。在FDMT数据结构下,当前状态数据可以与历史状态数据分离。在以太坊类型的系统中,与当前区块相关联的状态信息可以被视为频繁被虚拟机检索以执行智能合约的“热”数据。历史状态数据可以存储为历史状态树,该树可以包括始于创世区块的区块链账户状态的完整副本。与存储在历史状态树中的先前区块关联的状态信息可以被视为“冷”数据,访问这些数据以执行智能合约的频率较低。

[0045] 因为FDMT下的历史状态树是内容寻址的状态树,其本质上是增量式的,历史状态树的大小会因新区块的产生而变得非常大。由于历史状态树中的大多数数据是不频繁被使用的“冷”数据,因此就存储资源的使用而言,将这些数据存储在每个区块链节点中可能效率很低。为了节省存储资源而不实质影响计算效率,可以将历史状态数据存储在一个或多个可信存储位置或通过投票选举出的一个或多个共享存储节点上。然后可以由区块链网络的其他节点共享对所述历史状态数据的访问。

[0046] 与历史状态数据类似,区块数据包括区块链网络中的所有交易,这可能占用大量存储空间。对于存储资源有限的区块链节点,期望区块链网络中的每个区块链节点仅存储区块数据的一部分,并且可以从其他节点检索剩余的区块数据以减少存储消耗。但是,如果区块链网络中存在故障节点或不可靠节点,则接收到的数据可能不可信或可能发生数据丢失。

[0047] 在一些实施例中,诸如纠删编码的ECC可以用于对区块数据进行编码。通过共享ECC编码的区块数据而不是原始区块数据,即使存在非真实数据或发生数据丢失,只要非真实数据或数据丢失小于或等于可通过ECC纠正的错误位(bit)或丢失位的最大比例,原始区块数据就可以被恢复。

[0048] 图3描绘了根据本文实施例的FDMT数据结构300的示例。在FDMT下,账户状态可以作为键值对(KVP)存储在历史状态树302和当前状态树304的结构中。键对应于唯一标识区块链账户值的地址。历史状态树302可以包括区块链的可用状态信息的完整副本。当前状态树304可以包括当前区块的状态信息。因此,当前状态树304的大小可以明显小于历史状态

树302的大小。

[0049] 在一些实施例中,当前状态树304可以是位置寻址状态树。对于位置寻址状态树,当前状态树304的节点值可以基于唯一地标识该节点的键(即,节点ID)来检索。当新节点被添加到当前状态树304时,节点值可以与其唯一的节点ID(例如,当前状态树304的ID 1-1、ID 2-1等)相关联,而不必考虑其内容。在一些情况下,当前状态树304的KVP可以表示为<node ID,node value>。在一些情况下,KVP中的键还可以包括节点值的对应区块ID。在这种情况下,节点ID可以充当键的前缀,而区块ID可以充当键的后缀。然后可以将当前状态树304的KVP表示为<node ID+block ID,node value>。

[0050] 在一些实施例中,历史状态树302可以是内容寻址状态树。对于内容寻址状态树,每个账户值都可以具有与信息内容本身的值唯一关联的内容地址。为了从历史状态树302检索信息,可以提供内容标识,从中可以确定和检索账户值的位置。类似于MPT,历史状态树302的每个节点可以包括指向树的下一节点的指针的哈希值(例如,历史状态树302下的Hash1、Hash2和Hash3)。沿着指针的路径,最后一个元素存储键的末端部分的哈希值(例如,历史状态树302下的Hash4、Hash5、Hash6和hash7)以及与这些键配对的值。历史状态树302的KVP可以表示为<hash(node value),node value>。

[0051] 由于内容寻址树的节点地址取决于节点值,因此可以将新状态信息作为附加树结构添加到历史状态树302中而不是对现有树进行更改,从而保留原始树结构并提高数据存储/检索效率。

[0052] 图4描绘了根据本文实施例的用于存储区块链数据的数据库400的示例。数据库400可以是例如LevelDB或RocksDB的键值数据库。数据库400可以在FDMT数据结构下存储数据,该结构包括用于存储历史状态树的历史数据库410和用于存储当前状态树的当前数据库412。对于图4中描绘的四个区块,区块i-2 402、区块i-1 404和区块i 406是先前完成的区块。区块i+1 408是当前区块。每个区块可以具有区块头和区块体。区块头可以包括诸如世界状态的根哈希值的信息。根哈希值可以用作状态树的安全且唯一的标识。换句话说,根哈希值可以加密地取决于账户状态。区块体可以包括相应区块的已确认交易。

[0053] 历史数据库410可以将历史状态数据存储为历史状态树。历史状态数据可以是与区块链的先前区块关联的任何状态数据。当前数据库412可以将当前状态数据存储为当前状态树(例如,如图4的框412所示)。当前状态数据可以是与当前区块相关联的状态数据。在区块链上创建当前区块后,将创建新区块,从而当前状态数据成为历史状态数据。通过创建新区块而生成的状态数据将成为新的当前状态数据。

[0054] 历史状态数据和当前状态数据可以包括历史账户状态和当前账户状态。以太坊区块链账户可以包括外部拥有的账户和合约账户。外部拥有的账户可以由私钥控制,并且不与任何用于执行智能合约的代码关联。合约账户可以通过其与执行智能合约的代码相关联的合约代码来控制。

[0055] 以太坊账户的状态可以包括四个组分:随机值、余额、代码哈希值和存储根。如果该账户是外部拥有的账户,则随机数可以表示从该账户地址发送的交易的数量。余额可以代表该账户拥有的数字资产。代码哈希值可以是空字符串的哈希值。存储根可以为空。如果该账户是合约账户,则随机数可以表示该账户创建的合约数量。余额可以代表该账户拥有的数字资产。代码哈希值可以是与账户关联的虚拟机代码的哈希值。存储根可以存储与存

储树关联的根哈希值。存储树可以通过对账户的存储内容的哈希值进行编码来存储合约数据。

[0056] 历史状态树可以包括始于创世区块的区块链账户状态的完整副本,并且可以根据交易执行进行更新。例如,存储在先前区块i-1 404中的根哈希值是在区块i-1 404完成时世界状态的根哈希值。世界状态与存储在区块i-1 404以及在区块i-1 404之前的区块中的所有交易相关联。类似地,存储在当前区块i+1 408中的根哈希值是与存储在区块i+1 408和区块i+1 408之前的区块中的所有交易关联的世界状态的根哈希值。

[0057] 当前状态树可以包括由于新添加到当前区块i+1 408的交易而被更新或添加的状态信息。如在图3的描述中所讨论的,历史状态树可以将状态信息存储为表示为内容可寻址的<hash(node value),node value>的KVP。在一些实施例中,可以基于一个或多个与位置相关的ID来对当前状态树进行位置寻址。例如,当前状态树可以将状态信息存储为表示为<node ID,node value>的KVP,其中节点值可以基于其相应节点ID来寻址。作为另一示例,KVP中的键可以是节点ID和与节点值的对应的区块ID的组合。节点ID可以用作键的前缀,而区块ID可以用作键的后缀,以用于遍历FDMT或MPT的值。

[0058] 图5描绘了根据本文实施例的区块链网络500的另一示例。在高层级,区块链网络500包括多个区块链节点502、504、506和508、多个共享存储节点510和512,以及可通信地耦接到多个共享存储节点502和504或者区块链节点502、504、506和508中的一个或多个的云存储设备522。在一些情况下,共享存储节点510和512可以是具有POA的节点,诸如由区块链网络500的部署者管理的节点。在这种情况下,共享存储节点510和512可以在区块链网络500之外。在一些情况下,区块链节点可以是区块链网络500的一部分,在这种情况下,共享存储节点510和512的POA可以通过投票获得。例如,假设区块链网络包括3f+1个节点(在图5所示的示例中f＝1,当共享存储节点510和512均不参与区块链网络500的共识时,)、3f+2个节点(当共享存储节点510和512之一参与区块链网络500的共识时)或3f+3个节点(当共享存储节点510和512两者参与区块链网络的共识时),其中f是拜占庭节点的最大数量,如果2f+1个节点投票(通过其各自的数字签名背书)选举一个区块链节点作为共享存储节点,则2f+1个投票可以用作信任共享存储节点的POA。

[0059] 如在对图4的描述中所讨论的,在FDMT数据结构下,当前状态数据可以与历史状态数据分离。当前状态数据可以被存储为当前状态树,其包括与当前区块相关联的状态信息,诸如根据新添加到当前区块的交易而更新或添加的状态数据。在以太坊类型的系统中,与当前区块相关联的状态信息可以被视为频繁被虚拟机检索以执行智能合约的“热”数据。历史状态数据可以存储为历史状态树,该树可以包括始于创世区块的区块链的账户状态的完整副本。与存储在历史状态树中的先前区块关联的状态信息可以被视为“冷”数据,访问这些数据以执行智能合约的频率较低。

[0060] 为了节省存储资源而不显著降低计算效率,可以将历史状态树存储在与共享存储节点510和512相关联的历史数据库(例如图4中描述的历史数据库410)上或可通信地耦接至共享存储节点510和512的云存储设备522上。共享存储节点510和512可以将对历史状态树的访问共享给区块链节点502、504、506和508。云存储设备522可以是可以在云上提供存储服务的存储设备,例如网络附加存储(NAS)或对象存储服务(OSS)。

[0061] 当交易被处理进入当前区块时,与交易相关联的状态数据可以由区块链节点502、

504、506和508中的一个或多个发送到共享存储节点510和512以进行存储。在一些实施例中,区块链节点502、504、506和508中的一个或多个可以将状态数据和状态数据的哈希值作为KVP发送至共享存储节点510和512。在接收到状态数据之后,共享存储节点510和512可以验证接收到的状态数据或KVP是否已经在本地存储或存储在云存储设备522中。如果是,则共享存储节点510和512可以拒绝或放弃接收到的状态数据。否则,共享存储节点510和512可以计算状态数据的哈希值或验证接收到的哈希值是状态数据的哈希值,并将哈希值和状态数据存储到历史状态树。

[0062] 在一些实施例中,共享存储节点510和512可以验证状态数据是否是区块链的有效状态数据。如前所述,共享存储节点510和512可以存储历史状态树,该历史状态树是内容寻址的并且包括区块链的状态信息的完整副本。共享存储节点510和512可以计算接收到的状态数据的哈希值。然后,计算出的哈希值可以用于基于区块链的世界状态根哈希值(例如,使用Merkle证明)来验证状态数据是否是区块链的一部分。如果是,则可以将状态数据确定为内容寻址的。

[0063] 当区块链节点502、504、506和508中的任何一个需要从共享存储节点510或512检索状态数据时,可以将对应的哈希值发送到与区块链节点通信的共享存储节点。如图5中描绘的示例所示,区块链节点502和504可以将哈希值发送到共享存储节点512,区块链节点506和508可以将哈希值发送到共享存储节点510。区块链节点可以基于地理位置、网络状况、已建立的通信协议、安全性考虑等来选择从中检索状态数据的共享存储节点。应当理解,区块链节点502、504、506和508中的任何一个都可以根据本文的不同实施例来选择与共享存储节点510和512中的任何一个通信。

[0064] 由于存储在共享存储节点510和512中的历史状态树是内容寻址的,因此哈希值可以用作用于寻址相应状态数据的键。在基于哈希值识别出相应的状态数据之后,相应的共享存储节点510或512可以将识别出的状态数据发送回区块链节点。接收状态数据的区块链节点可以对接收到的状态数据进行哈希处理,以验证状态数据是否为内容寻址的。如果是,则将状态数据确定为真实的。否则,状态数据是不真实的。如果状态数据不真实,则区块链节点可以选择将共享存储节点报告为故障节点(或拜占庭节点)。如果区块链网络500中存在存储历史状态树的其他节点,则区块链节点可以将哈希值发送到一个或多个其他节点以检索相应的状态数据。

[0065] 除了通过共享存储节点510和512共享历史数据之外,区块链节点502、504、506和508还可以根据区块数据编码和哈希处理来共享区块数据。以区块100为例,区块100可以包括区块头516和区块体518。在区块数据被存储在区块100中之后,区块链节点502可以参与到与其他区块链节点502、504、506和508的共识处理中。在共识处理期间,区块链节点502可执行共识算法,例如工作量证明(PoW)或权益证明(PoS),以在区块链上创建相应的区块。

[0066] 当区块数据被写入区块的区块体时,区块链节点502可以对区块数据执行ECC。这样,区块链节点502不需要存储整个区块,而是可以基于一个或多个预定规则存储ECC编码的区块数据的所选部分和与剩余的编码区块数据相对应的哈希值。该编码和哈希处理500会尤其适合于区块链节点502具有低磁盘空间的情况。

[0067] 在一些实施例中,代替将数据存储为区块,区块链节点502可以存储WAL文件或其他类似的前滚日志文件。WAL文件可以记录区块链节点502已提交但尚未存储的区块数据。

使用WAL文件,原始区块链数据可以保存在数据库文件中,而对区块链数据的更改可以写入单独的WAL文件中。可以利用更改提交前滚而无需写入原始区块链数据。这种布置允许在将更改提交到WAL文件的同时继续进行区块链数据的操作。通过使用WAL文件存储通过编码和哈希处理500进行的更改,区块链节点502可以指示其具有用于共识的区块数据,同时在适当的时候在后台执行ECC。这样,可以在区块链节点302的计算资源的利用率较低时执行ECC,以减少对共识处理的计算效率或等待时间的影响。

[0068] 在一些实施例中,区块链节点502可以将区块数据存储在缓冲器中。当数据的大小大于预定阈值或当缓冲器已满时,区块链节点502可以对存储在缓冲器中的区块数据执行ECC。在执行ECC之后,区块链节点502可以遵循编码和哈希处理500以存储编码区块数据和哈希值,如以下描述中所讨论的。

[0069] 通过将冗余位添加到数据,ECC可用于控制不可靠传输上的错误或数据丢失。冗余可以允许纠正错误或丢失的数据而无需重新传输数据。ECC的一个示例可以是纠删编码。使用纠删编码,可以将k个符号的消息编码为具有n个符号的代码字,其中k和n是自然数,并且k<n。该消息可以从n个符号的代码字的子集中被恢复。比例r=k/n是纠删编码的码率。

[0070] 通过使用ECC,每个区块链节点可以存储编码区块数据的一部分,并在需要时从其他区块链节点检索剩余的编码区块数据。在一些实施例中,可以在区块链节点502的计算资源的利用率低于预定值(例如40%)时执行ECC。这样,可以减少对区块链节点502上的其他计算操作的干扰。在一些实施例中,可以在区块链节点502的存储空间的使用率大于或等于预定百分比时执行ECC,使得在ECC之后,可以删除编码区块数据的某些部分以释放存储空间。

[0071] 再次以区块100为例,在执行ECC之后,可以基于一个或多个预定规则将编码区块数据划分为多个数据集。在图5所示的示例中,存储在区块100的区块体518中的编码区块数据被分为四个数据集,分别是Data1、Data2、Data3和VData1,每个数据集将由区块链节点502、504、506和508之一保存。VData1可以表示ECC的用于纠错的冗余位。根据一个或多个预定规则选择要由区块链节点352存储的Data1。选择Data2、Data3和VData1以分别进行哈希处理524从而分别产生哈希值Dhash2、Dhash3和Vhash1。

[0072] 现在转向图6,图6描绘了根据本文实施例的数据存储布置600的示例。如前所述,根据一个或多个预定规则选择要由区块链节点502存储的Data1。基于数据存储布置600,区块链节点504存储Data2,并分别对Data1、Data3和VData1进行哈希处理以分别生成哈希值Dhash1、Dhash3和Vhash1。区块链节点506存储Data3并分别对Data1、Data2和VData1进行哈希处理以分别生成哈希值Dhash1、Dhash2和Vhash1。区块链节点508存储VData1并分别对Data1、Data2和Data3进行哈希处理以分别生成哈希值Dhash1、Dhash2和Dhash3。

[0073] 再次参考图5,由于哈希值与同一区块的编码数据集对应,因此它们可以通过该区块的区块ID来索引。如,区块链节点502可以索引与具有区块ID 100的区块100相关联的Data1、Dhash1、Dhash2和Vhash1。这样,区块链节点502可以使用索引的区块ID来将哈希值映射到其对应的区块。在一些情况下,索引数据可以存储在区块链节点502或可通信地耦接到区块链节点502的云存储设备522中。

[0074] 应当理解,根据一个或多个预定规则,可以为一个或多个区块链节点502、504、506和508制订其他数据存储布置。通常,一个或多个预定规则可以包括用于基于区块链网络中

的区块链节点的数量将编码区块数据划分为多个数据集的一个或多个指令。一个或多个预定规则还可以包括与多个数据集中的一个或多个数据集要由每个区块链节点存储或进行哈希处理有关的分配。为了确保数据的对等性,一个或多个预定规则可以包括与至少一个数据集要由区块链网络的每个区块链节点存储的有关的分配。例如,当区块链网络具有四个以上的节点时,基于预定规则,可以将区块100的编码区块数据划分为四个以上的数据集。每个区块链节点可以存储一个以上的数据集,并对其他节点存储的剩余数据集进行哈希处理。

[0075] 在生成Dhash2、Dhash3和Vhash1之后,区块链节点502的区块体520可以存储Data1、Dhash2、Dhash3和Vhash1。对于区块链的每个区块,区块链节点502仅存储ECC编码的区块数据的一个数据集(即,Data1)和三个哈希值(即,Dhash2、Dhash3和Vhash1)而不是原始区块数据以节省储存空间。当区块链节点502确定执行智能合约需要区块100的区块数据时,它可以根据一个或多个预定规则分别从区块链节点504、506和508检索Data2、Data3和Vdata1。

[0076] 为了从其他区块链节点504、506和508检索数据集,区块链节点502可以根据一个或多个预定规则发送与要检索的数据集相对应的哈希值。例如,为了检索Data2,区块链节点502可以将Dhash2发送到区块链节点504。如果区块链节点504存储有Data2,则其可以响应于接收到Dhash2而将Data2发送回区块链节点502。在从区块链节点504接收到Data2之后,区块链节点502可以对接收到的数据集进行哈希处理并将哈希值与Dhash2进行比较。如果哈希值与Dhash2相同,则区块链节点502可以确定接收到的数据集是真实的。否则,可以将接收到的数据集确定为不真实。当接收到的数据集被确定为不真实时,区块链节点502可以将区块链节点504报告为故障节点(或拜占庭节点)。如果由区块链节点502接收的非真实数据的百分比小于或等于可通过ECC纠正的错误位或丢失位的最大比例,则区块100可以从本地存储和接收的数据集中被恢复。

[0077] 图7是用于通信和共享区块链数据的处理600的示例的流程图。为了方便起见,处理700将被描述为由位于一个或多个位置的一个或多个计算机的系统执行,并且根据本文被适当地编程。例如,例如图1中的计算系统106、108的计算系统中适当编程的计算设备可以执行处理700。

[0078] 在702,区块链网络中的区块链节点确定与区块链的当前区块相关联的区块数据和当前状态数据。

[0079] 在704,区块链节点将当前状态数据发送到区块链网络的一个或多个共享存储节点,其中,区块链节点存储当前状态数据,并且一个或多个共享存储节点存储与区块链的每个区块相关联的历史状态数据。

[0080] 在706,区块链节点对区块数据执行纠错编码以生成编码区块数据。

[0081] 在708,区块链节点基于一个或多个预定规则将编码区块数据划分为多个数据集。

[0082] 在710,区块链节点基于一个或多个预定规则存储多个数据集中的一个或多个数据集。

[0083] 在712,区块链节点对多个数据集中的每个剩余数据集进行哈希处理,以生成与多个数据集中的剩余数据集相对应的一个或多个哈希值。

[0084] 在714,区块链节点存储一个或多个哈希值和当前状态数据。

[0085] 在一些情况下,处理700还包括:区块链节点向一个或多个共享存储节点之一发送哈希值,以检索包括在历史状态数据中的账户状态;区块链节点接收针对发送该哈希值的响应中的账户状态;以及区块链节点基于哈希值验证账户状态是区块链的一部分。

[0086] 在一些情况下,区块链网络包括至少f+1个共享存储节点和不超过2f+2个共识节点,其中f是区块链网络中可以容忍的故障共享存储节点和共识节点的最大数量。

[0087] 在一些情况下,一个或多个共享存储节点是通过从区块链网络中的所有3f+1、3f+2或3f+3个节点接收到2f+1个投票选举出的,其中f是区块链网络中可以容忍的故障共享存储节点和共识节点的最大数量。

[0088] 在一些情况下,当前状态数据和历史状态数据被存储为固定深度默克尔树。

[0089] 在一些情况下,纠错编码是纠删编码。

[0090] 在一些情况下,区块链节点是第一区块链节点,并且处理700还包括:确定第一区块链节点未存储执行智能合约所需的至少部分区块数据;存储多个数据集中的至少一个剩余数据集的第二区块链节点以及与多个数据集中的至少一个剩余数据集相对应的至少一个哈希值;将至少一个哈希值发送至第二区块链节点,以检索多个数据集中的至少一个剩余数据集。

[0091] 在一些情况下,所述至少一个哈希值是至少一个第一哈希值,处理700还包括:响应于发送至少一个第一哈希值,从第二区块链节点接收至少一个数据集;以及对该至少一个数据集进行哈希处理以生成至少一个第二哈希值;如果至少一个第一哈希值与至少一个第二哈希值相同,则确定所述至少一个数据集是真实的。

[0092] 在一些情况下,处理700还包括:基于一个或多个预定规则识别存储多个数据集中的至少一个剩余数据集的第三区块链节点以及多个数据集中的所述至少一个剩余数据集相对应的至少一个第三哈希值;将至少一个第三哈希值发送至第三区块链节点,以检索多个数据集中的至少一个剩余数据集;响应于发送至少一个第三哈希值,从第三区块链节点接收至少一个数据集;对所述至少一个数据集进行哈希处理以生成至少一个第四哈希值;如果至少一个第三哈希值与至少一个第四哈希值相同,则确定该至少一个数据集是真实的。

[0093] 在一些情况下,一个或多个预定规则包括用于基于区块链网络中的区块链节点的数量将编码区块数据划分为多个数据集的一个或多个指令,以及与所述多个数据集中的一个或多个数据集要被所述区块链节点中的每个区块链节点存储或进行哈希处理有关的分配。

[0094] 在一些情况下,处理700还包括:将当前区块存储为预写日志(WAL)文件;并报告当前区块被存储并准备用于达成共识。

[0095] 在一些情况下,处理700还包括:将区块数据存储在区块链节点的缓冲器中,并且响应于识别出缓冲器已满而执行纠错编码。

[0096] 图8是根据本文的实施例的装置800的模块的示例的示图。

[0097] 装置800可以是被配置为通信和共享区块链数据的区块链节点的实施例的示例。装置800可以对应于上述实施例,装置800包括:确定模块802,其确定与区块链的当前区块相关联的区块数据和当前状态数据;发送模块804,其将当前状态数据发送至区块链网络的一个或多个共享存储节点,其中,所述区块链节点存储当前状态数据,并且所述一个或多个

共享存储节点存储与所述区块链的每个区块相关联的历史状态数据;编码模块806,其对区块数据执行纠错编码以生成编码区块数据;划分模块808,其将编码区块数据划分为多个数据集;存储模块810,其存储多个数据集中的一个或多个数据集;哈希模块812,其对多个数据集中的每个剩余数据集进行哈希处理,以生成与多个数据集中的剩余数据集相对应的一个或多个哈希值;存储模块810,其存储所述一个或多个哈希值和当前状态数据。

[0098] 在可选实施例中,装置800还包括:发送子模块,其向一个或多个共享存储节点之一发送哈希值,以检索包括在历史状态数据中的账户状态;接收子模块,接收针对发送哈希值的响应中的账户状态;并且由区块链节点基于所述哈希值验证所述账户状态是区块链的一部分。

[0099] 在可选实施例中,区块链网络包括至少f+1个共享存储节点,以及不超过2f+2个共识节点,其中f是区块链网络中可以容忍的故障共享存储节点和共识节点的最大数量。

[0100] 在可选实施例中,一个或多个共享存储节点是通过从区块链网络中的所有3f+1、3f+2或3f+3个节点接收到2f+1个投票选举出的,其中f是区块链网络中可以容忍的故障共享存储节点和共识节点的最大数量。

[0101] 在可选中,当前状态数据和历史状态数据被存储为固定深度默克尔树。

[0102] 在可选实施例中,纠错编码是纠删编码。

[0103] 在可选实施例中,区块链节点是第一区块链节点,并且装置800还包括:确定子模块,其确定第一区块链节点未存储执行智能合约所需的至少部分区块数据;识别子模块,其识别存储多个数据集中的至少一个剩余数据集的第二区块链节点以及与多个数据集中的所述至少一个剩余数据集相对应的至少一个哈希值;以及发送模块,其将所述至少一个哈希值发送至第二区块链节点,以检索多个数据集中的至少一个剩余数据集。

[0104] 在可选实施例中,所述至少一个哈希值是至少一个第一哈希值,装置800还包括:接收子模块,其响应于发送至少一个第一哈希值,接收至少一个数据集;以及对所述至少一个数据集进行哈希处理以生成至少一个第二哈希值;以及确定子模块,如果至少一个第一哈希值与至少一个第二哈希值相同,则该确定子模块确定所述至少一个数据集是真实的。

[0105] 在可选实施例中,装置800还包括:识别子模块,其识别存储所述多个数据集中的至少一个所述剩余数据集的第三区块链节点以及与所述多个数据集中的所述至少一个剩余数据集相对应的至少一个第三哈希值;发送子模块,其将至少一个第三哈希值发送至第三区块链节点,以检索所述多个数据集中的至少一个剩余数据集;接收子模块,其响应于发送至少一个第三哈希值,接收至少一个数据集;哈希子模块,其对所述至少一个数据集进行哈希处理以生成至少一个第四哈希值;以及确定子模块,如果至少一个第三哈希值与至少一个第四哈希值相同,则该确定子模块确定该至少一个数据集是真实的。

[0106] 在可选实施例中,一个或多个预定规则包括用于基于区块链网络中的区块链节点的数量将编码区块数据划分为多个数据集的一个或多个指令,以及与多个数据集中的一个或多个数据集要被区块链节点中的每个区块链节点存储或进行哈希处理有关的分配。

[0107] 在可选实施例中,装置800还包括:存储子模块,其将当前区块存储为预写日志(WAL)文件;以及报告子模块,其报告当前区块被存储并准备用于达成共识。

[0108] 在可选实施例中,装置800还包括:存储子模块,其将区块数据存储在区块链节点的缓冲器中,并且响应于识别出缓冲器已满而执行纠错编码。

[0109] 在先前实施中示出的系统、装置、模块或单元可以通过使用计算机芯片或实体来实现,或者可以通过使用具有特定功能的产品来实现。典型的实现设备是计算机,计算机可以是个人计算机、膝上型计算机、蜂窝电话、相机电话、智能电话、个人数字助理、媒体播放器、导航设备、电子邮件收发设备、游戏控制台、平板电脑、可穿戴设备或这些设备的任意组合。

[0110] 对于装置中各个单元的功能和角色的实施过程,可以参考前一方法中相应步骤的实施过程。为简单起见,这里省略了细节。

[0111] 由于装置实施例基本上与方法实施例相对应,因此对于相关部分,可以参考方法实施例中的相关描述。前述装置实施例仅仅是示例。被描述为单独部分的模块可以或可以不是物理上分离的,并且显示为模块的部分可以是或可以不是物理模块,可以位于一个位置,或者可以分布在多个网络模块上。可以基于实际需求来选择一些或所有模块,以实现本文方案的目标。本领域普通技术人员无需付出创造性努力就能理解和实现本申请的实施例。

[0112] 再次参考图8,它可以被解释为示出了区块链节点的内部功能模块和结构。本质上,执行主体实质上可以是电子设备,该电子设备包括以下:一个或多个处理器;以及被配置为存储一个或多个处理器的可执行指令的一个或多个计算机可读存储器。在一些实施例中,一个或多个计算机可读存储器耦接至所述一个或多个处理器且其上存储有程序指令的,所述指令能由所述一个或多个处理器执行以执行本文描述的算法、方法、函数、处理、流程和进程。

[0113] 本文中描述的技术产生若干技术效果。例如,主题的实施例可以允许节省区块链节点的存储资源,而不显著地降低计算效率。通过仅保存ECC编码的区块数据的一部分和与剩余ECC编码的区块数据相对应的哈希值,可以减少整个区块链网络的存储空间消耗。

[0114] 在一些实施例中,代替存储整个块,区块链节点可以存储ECC编码的区块的所选部分和与剩余的编码区块相对应的哈希值。即使区块链节点从故障的区块链节点接收到非真实数据,只要非真实数据的百分比小于或等于ECC允许的错误位或丢失位的最大比例,该区块数据就可以被恢复。

[0115] 此外,由于大多数历史状态数据是不频繁被使用的“冷”数据,因此通过仅将“冷”状态数据保存在共享存储节点中,可以显著提高整个区块链网络的存储空间使用率。对于具有N个节点的区块链网络,其中N等于3f+1、3f+2或3f+3,其中f是故障共识节点的最大数量,(N-f-1)/N个区块链节点仅需要将“热”数据存储为当前状态树,而无需将“冷”和“热”数据都存储为历史状态树。对于将f+1个节点用作共享存储节点来存储历史状态树的N个节点的区块链网络,最多可以容忍f个故障共识节点。换句话说,节省存储空间不会损害数据可靠性。由于f+1个共享存储节点确保了系统的可靠性,因此可以提高数据安全性,并且相对独立于底层服务平台的安全级别。

[0116] 所描述的主题的实施例可单独地或组合地包括一个或多个特征。

[0117] 例如,在第一实施例中,一种计算机实现的用于通信共享区块链数据的方法,所述方法包括:区块链网络中的区块链节点向区块链网络的一个或多个共享存储节点发送与区块链的当前区块相关联的当前状态信息以及一个或多个交易,其中,该区块链节点存储与区块链的每个区块相关联的区块头和当前状态信息,该一个或多个共享存储节点存储与区

块链的每个区块相关联的区块体和历史状态信息,并且其中历史状态信息被存储为历史状态树的值为与区块链网络相关联的账户的账户状态,键为对应账户状态的哈希值的键值对(KVP),;区块链节点验证一个或多个交易是由一个或多个共享存储节点存储的;以及区块链节点从一个或多个共享存储节点之一接收与区块链节点相关联的交易的通知。

[0118] 前述和其它描述的实施例可以各自可选地包括一个或多个以下特征:

[0119] 第一特征,可与以下任意特征组合,还包括:区块链节点向一个或多个共享存储节点之一发送哈希值,以检索包括在所述历史状态数据中的账户状态;区块链节点接收针对发送所述哈希值的响应中的所述账户状态;以及区块链节点基于哈希值验证账户状态是区块链的一部分。

[0120] 第二特征,可与以下任意特征组合,其中,区块链网络包括至少f+1个共享存储节点和不超过2f+2个共识节点,其中,f是在区块链网络中可以容忍的故障共享存储节点和共识节点的最大数量。

[0121] 第三特征,可与以下任意特征组合,其中,一个或多个共享存储节点是通过从所述区块链网络中的所有3f+1、3f+2或3f+3个节点接收到2f+1个投票选举出的,其中f是区块链网络中可以容忍的故障共享存储节点和共识节点的最大数量。

[0122] 第四特征,可与以下任意特征组合,其中,当前状态数据和历史状态数据被存储为固定深度默克尔树。

[0123] 第五特征,可与以下任意特征组合,其中,所述纠错编码是纠删编码。

[0124] 第六特征,可与以下任意特征组合,其中,区块链节点是第一区块链节点,并且所述方法还包括:确定第一区块链节点未存储执行智能合约所需的至少部分区块数据;基于一个或多个预定规则识别存储多个数据集中的至少一个剩余数据集的第二区块链节点以及与多个数据集中的所述至少一个剩余数据集相对应的至少一个哈希值;将所述至少一个哈希值发送至第二区块链节点,以检索所述多个数据集中的至少第一个剩余数据集。

[0125] 第七特征,可与以下任意特征组合,其中,所述至少一个哈希值是至少一个第一哈希值,所述方法还包括:响应于发送至少一个第一哈希值,从第二区块链节点接收至少一个数据集;对该至少一个数据集进行哈希处理以生成至少一个第二哈希值;如果至少一个第一哈希值与至少一个第二哈希值相同,则确定所述至少一个数据集是真实的。

[0126] 第八特征,可与以下任意特征组合,其中,所述方法还包括:基于一个或多个预定规则识别存储多个数据集中的至少一个剩余数据集的第三区块链节点以及与多个数据集中的所述至少一个剩余数据集相对应的至少一个第三哈希值;将至少一个第三哈希值发送至第三区块链节点,以检索多个数据集中的至少一个剩余数据集;响应于发送至少一个第三哈希值,从第三区块链节点接收至少一个数据集;对所述至少一个数据集进行哈希处理以生成至少一个第四哈希值;如果至少一个第三哈希值与至少一个第四哈希值相同,则确定该至少一个数据集是真实的。

[0127] 第九特征,可与以下任意特征组合,其中,一个或多个预定规则包括用于基于区块链网络中的区块链节点的数量将编码区块数据划分为所述多个数据集的一个或多个指令,以及与所述多个数据集中的一个或多个数据集要被所述区块链节点中的每个区块链节点存储或进行哈希处理有关的分配。

[0128] 第十特征,可与以下任意特征组合,其中,所述方法还包括:将当前区块存储为预

写日志(WAL)文件;以及报告当前区块被存储并准备用于达成共识。

[0129] 第十一特征,可与以下任意特征组合,其中,所述方法还包括:将区块数据存储在区块链节点的缓冲器中,并且响应于识别出缓冲器已满而执行纠错编码。

[0130] 本文中描述的主题、动作和操作的实施例可以在数字电子电路中、有形体现的计算机软件或固件中、包括本文中公开的结构及其结构等同物的计算机硬件中,或者它们中的一个或多个的组合中实现。本文中描述的主题的实施例可以实现为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的操作。例如,计算机程序载体可以包括具有一个或多个计算机可读存储介质,其上编码或存储有指令。载体可以是有形的非暂态计算机可读介质,例如磁盘、磁光盘或光盘、固态驱动器、随机存取存储器(RAM)、只读存储器(ROM)或其他媒体类型。可选地或附加地,载体可以是人工生成的传播信号,例如,机器生成的电、光或电磁信号,其被生成以编码用于传输到合适的接收器装置以供数据处理装置执行的信息。计算机存储介质可以是或可以部分是机器可读存储设备、机器可读存储基板、随机或串行访问存储器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。

[0131] 计算机程序也可以被称为或描述为程序、软件、软件应用程序、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,或者作为模块、组件、引擎、子程序或适合在计算环境中执行的其他单元,该环境可包括由数据通信网络互连的在一个或多个位置一台或多台计算机。

[0132] 计算机程序可以但非必须对应于文件系统中的文件。计算机程序可以存储在:保存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、子程序或代码部分的多个文件。

[0133] 举例来说,用于执行计算机程序的处理器包括通用和专用微处理器,以及任何类型的数字计算机的任何一个或多个处理器。通常,处理器将接收用于执行的计算机程序的指令、以及接收来自耦接到处理器的非暂态计算机可读介质的数据。

[0134] 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,包括例如可编程处理器、计算机或多个处理器或计算机。数据处理设备可以包括例如FPGA(现场可编程门阵列),ASIC(专用集成电路)或GPU(图形处理单元)的专用逻辑电路。除了硬件,该装置还可以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、或者它们一个或多个的组合的代码。

[0135] 本文中描述的处理和逻辑流程可以由一台或多台计算机或处理器执行一个或多个计算机程序进行,以通过对输入数据进行运算并生成输出来执行操作。过程和逻辑流程也可以由例如FPGA、ASIC或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。

[0136] 适合于执行计算机程序的计算机可以基于通用和/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。

[0137] 通常,计算机还将包括或可操作地耦接至一个或多个大容量存储设备,以从一个或多个存储设备接收数据或将数据传输到一个或多个大容量存储设备。存储设备可以是例如磁盘、磁光盘或光盘、固态驱动器或任何其他类型的非暂态计算机可读介质。然而,计算机不需要具有这样的设备。因此,计算机可以耦接到本地和/或远程的例如一个或多个存储器的一个或多个存储设备。例如,计算机可以包括作为计算机的集成组件的一个或多个本地存储器,或者计算机可以耦接到云网络中的一个或多个远程存储器。此外,计算机可以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)接收器或例如通用串行总线(USB)闪存驱动器的便携式存储设备,仅举几例。

[0138] 组件可以通过诸如直接地连接、或通过一个或多个中间组件彼此电学连接或光学连接可通信地连接而彼此“耦接”,。如果其中一个组件被集成到另一个组件中,组件也可以彼此“耦接”。例如,集成到处理器(例如,L2高速缓存组件)中的存储组件“耦接到”处理器。

[0139] 为了提供与用户的交互,本文中所描述的主题的实施例可以在计算机上实现或配置为与该计算机通信,该计算机具有:显示设备,例如,LCD(液晶显示器)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以是任何形式的感觉反馈,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,包括声音、语音或触觉输入。此外,计算机可以通过向用户使用的设备发送文档和从用户使用的设备接收文档来与用户交互;例如,通过向用户设备上的web浏览器发送web页面以响应从web浏览器收到的请求,或者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手机)轮流发送文本消息或其他形式的消息来并接收来自用户的响应消息来与用户交互。

[0140] 本文使用与系统,装置和计算机程序组件有关的术语“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被数据处理装置执行时促使该装置执行所述操作或动作的指令。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所述操作或动作的电子逻辑。

[0141] 尽管本文包含许多具体实施细节,但这些不应被解释为由权利要求本身限定的对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。在本文单独实施例的上下文中描述的某些特征也可以在单个实施例中组合实现。相反,在单个实施例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例中实现。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,但是在可选实施例中,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,并且要求保护可以指向子组合或子组合的变体。

[0142] 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:为了达到期望的效果,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。在某些情况下,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中或打包成多个软件产品。

[0143] 已经描述了主题的特定实施例。其他实施例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以不同的顺序执行并且仍然实现期望的结果。作为一个示例,附图中描绘的过程无需要求所示的特定顺序或次序来实现期望的结果。在可选实施例中,多任务和并行处理可能是有利的。

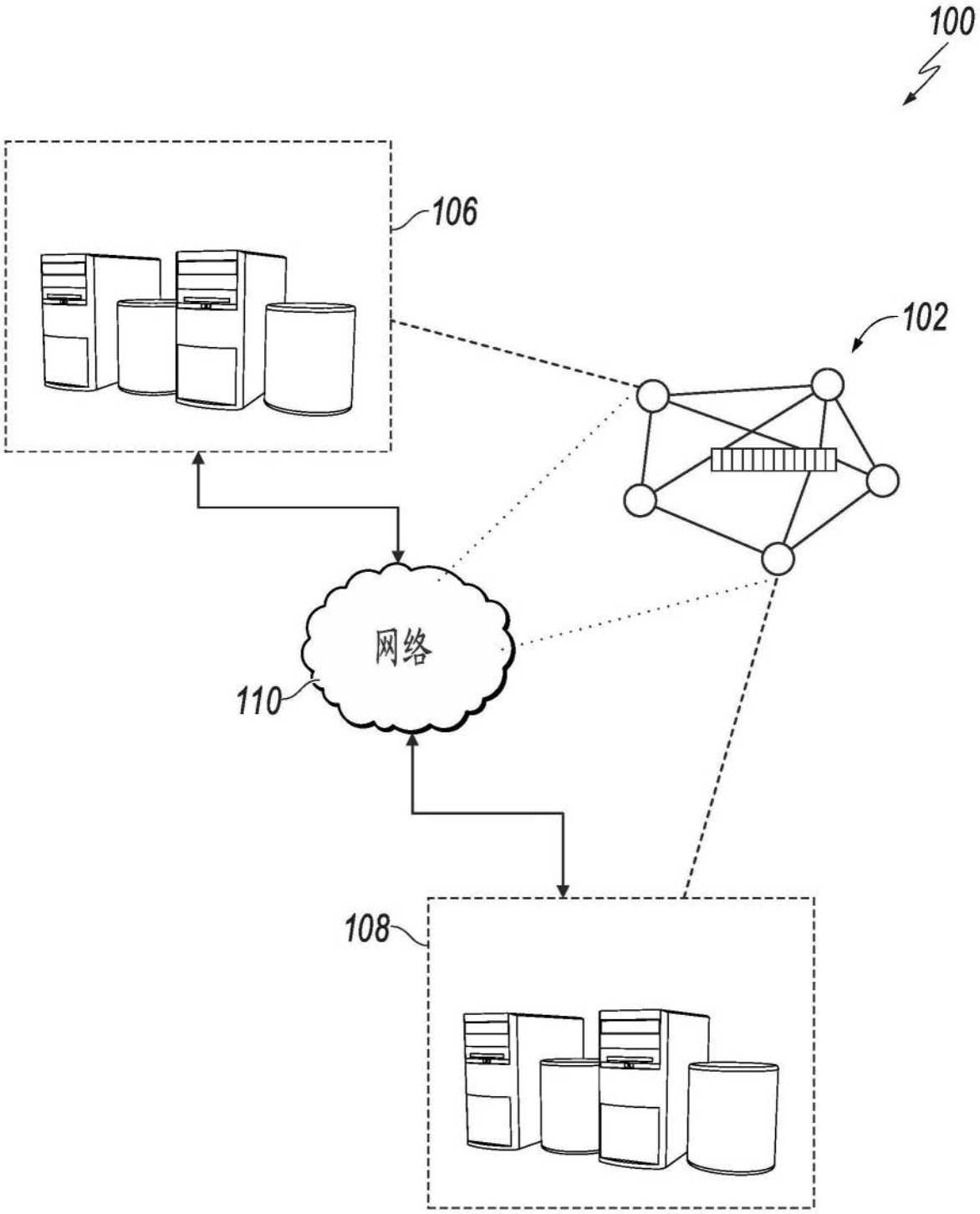


图1

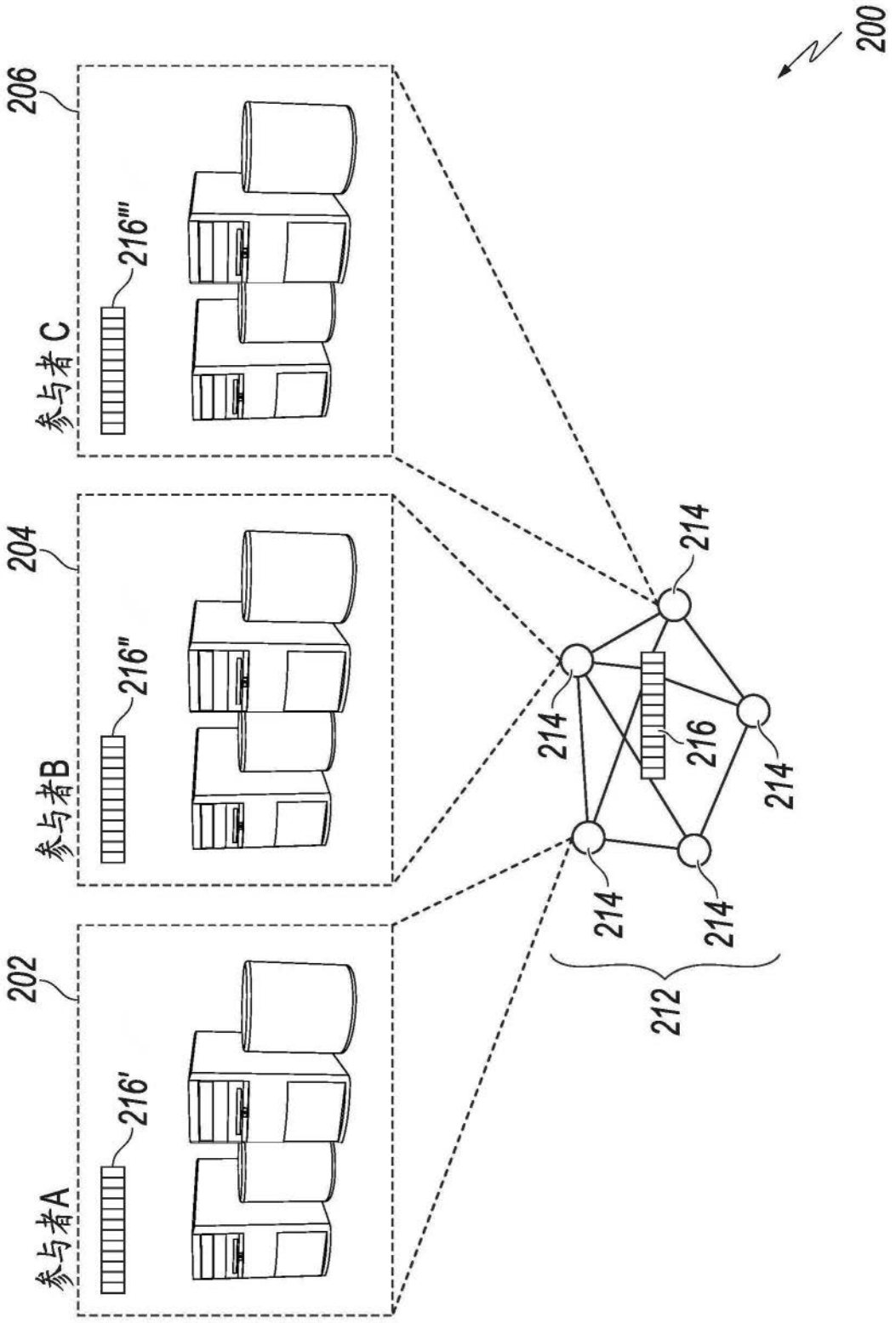

200
206
参与者C
216'''
204
参与者B
216''
202
参与者A
216'
214
214
214
214
214
216
212

图2

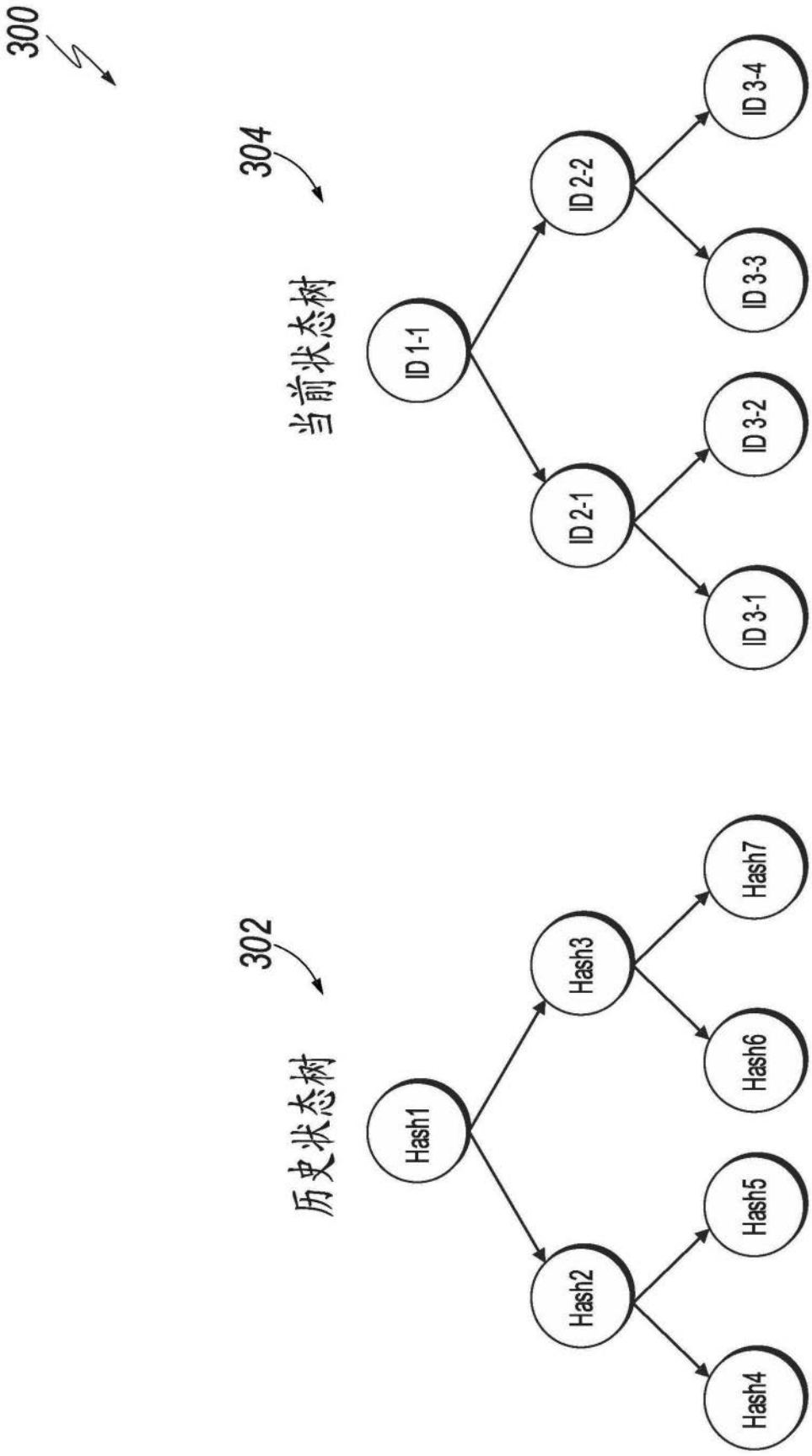

300
304
当前状态树
ID 1-1
ID 2-2
ID 2-1
ID 3-4
ID 3-3
ID 3-2
ID 3-1
302
历史状态树
Hash1
Hash3
Hash2
Hash7
Hash6
Hash5
Hash4

图3

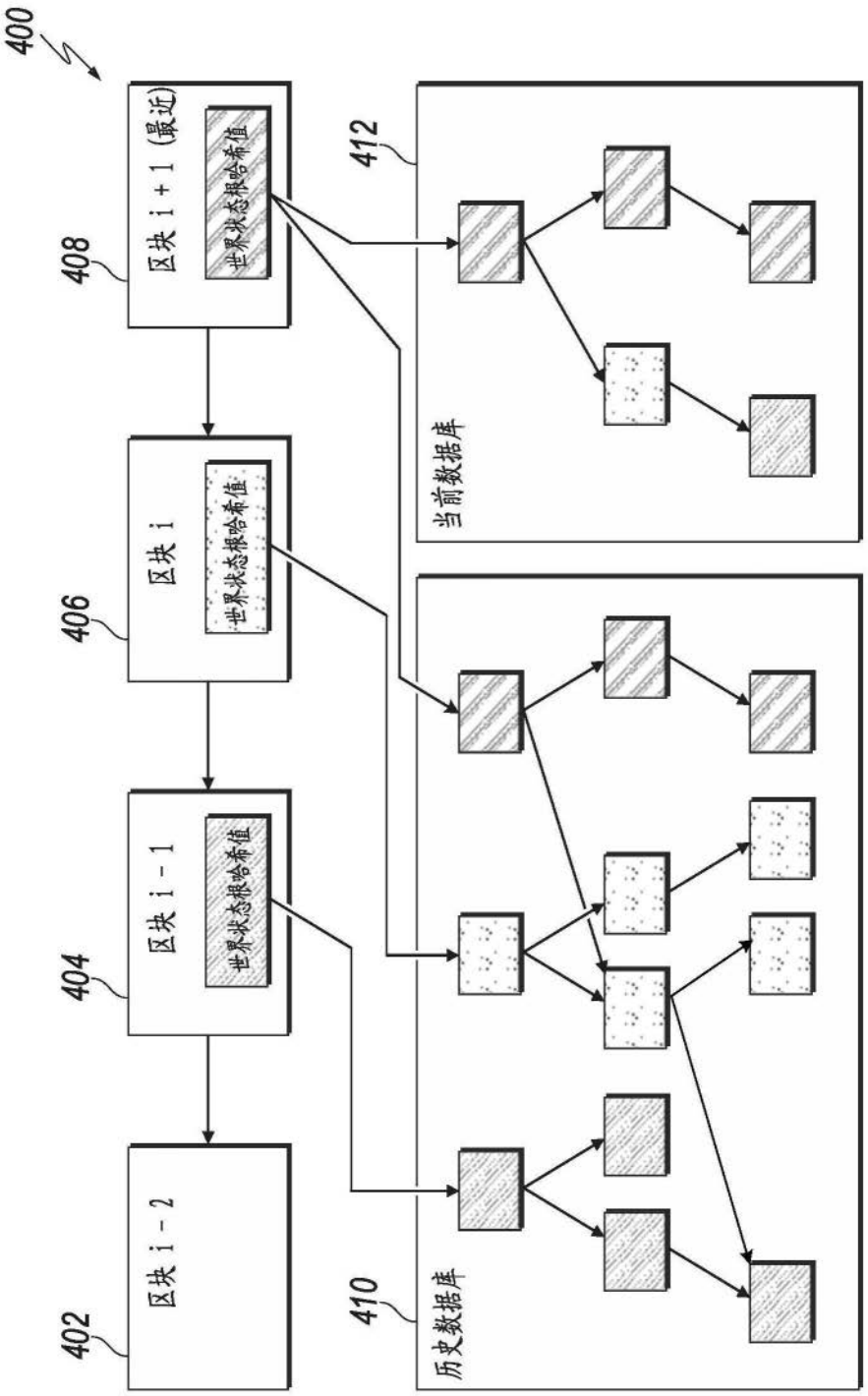

400
区块 i + 1 (最近)
世界状态根哈希值
408
区块 i
世界状态根哈希值
406
区块 i - 1
世界状态根哈希值
404
区块 i - 2
402
412
当前数据库
410
历史数据库

图4

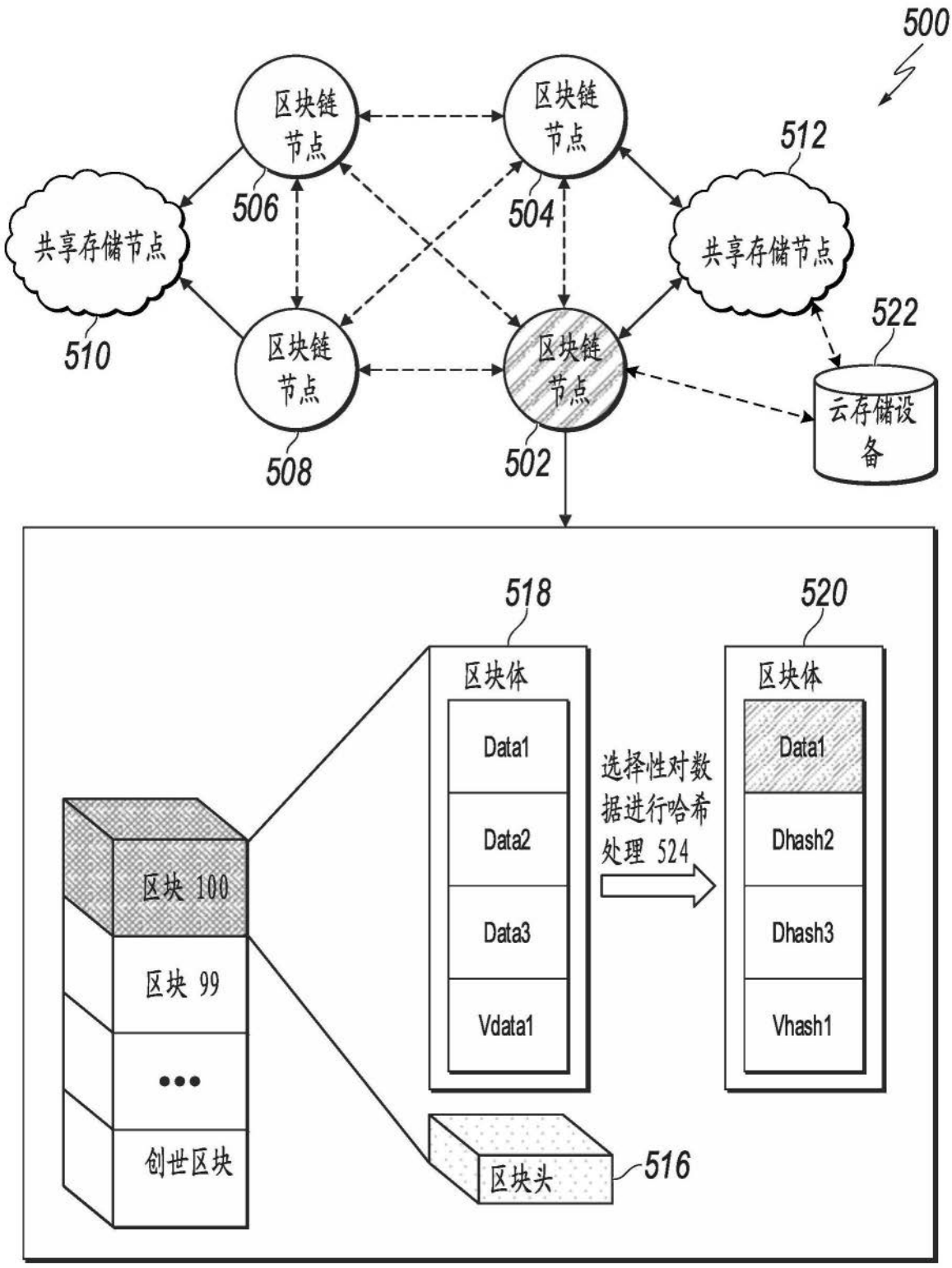

500
区块链节点
区块链节点
512
506
504
共享存储节点
共享存储节点
510
区块链节点
区块链节点
522
云存储设备
508
502
518
520
区块体
区块体
Data1
Data2
Data3
Vdata1
选择性对数据进行哈希处理 524
Data1
Dhash2
Dhash3
Vhash1
区块 100
区块 99
•••
创世区块
区块头
516

图5

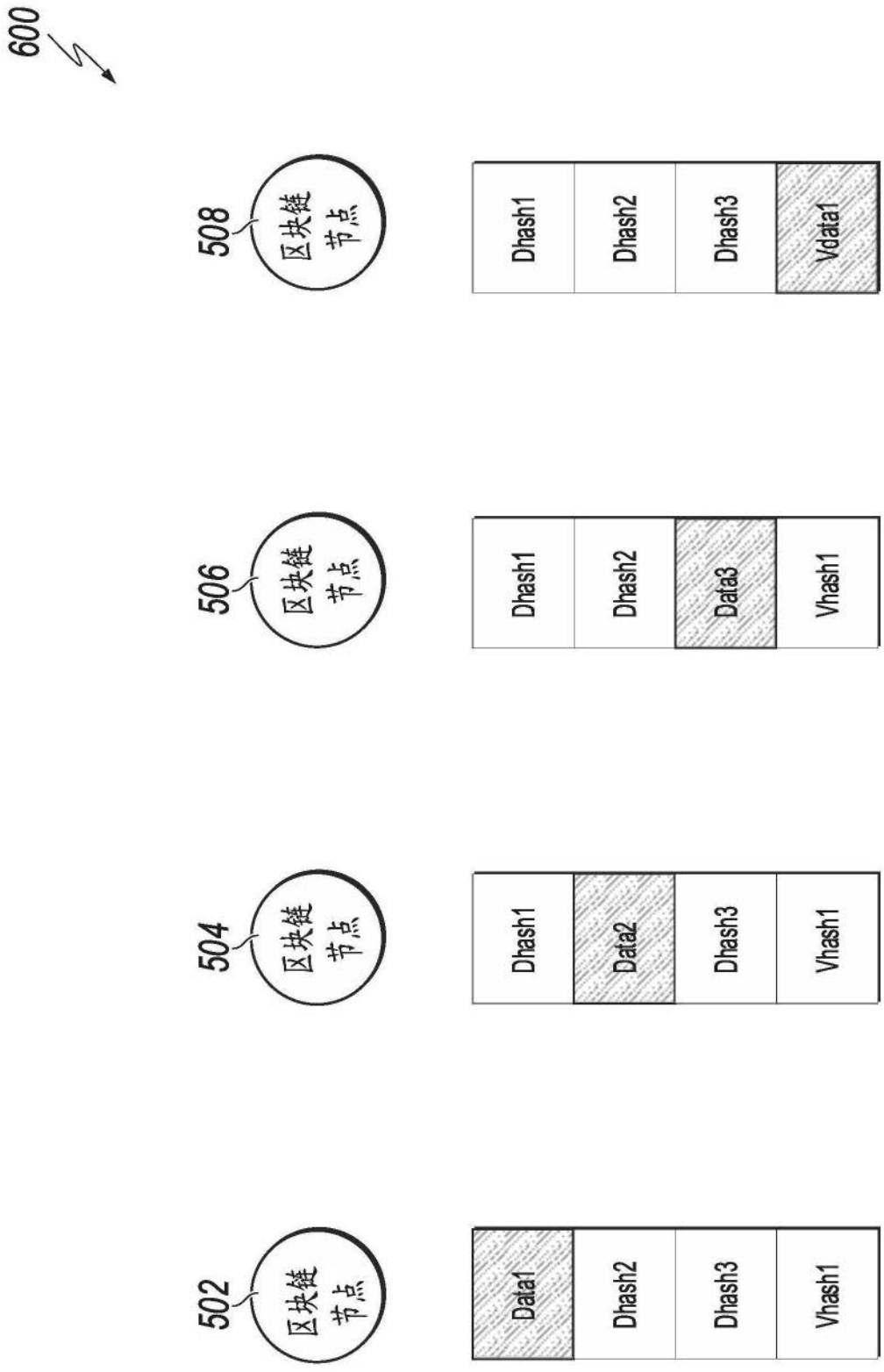

600
508
区块链节点
Dhash1
Dhash2
Dhash3
Vdata1
506
区块链节点
Dhash1
Dhash2
Data3
Vhash1
504
区块链节点
Dhash1
Data2
Dhash3
Vhash1
502
区块链节点
Data1
Dhash2
Dhash3
Vhash1

图6

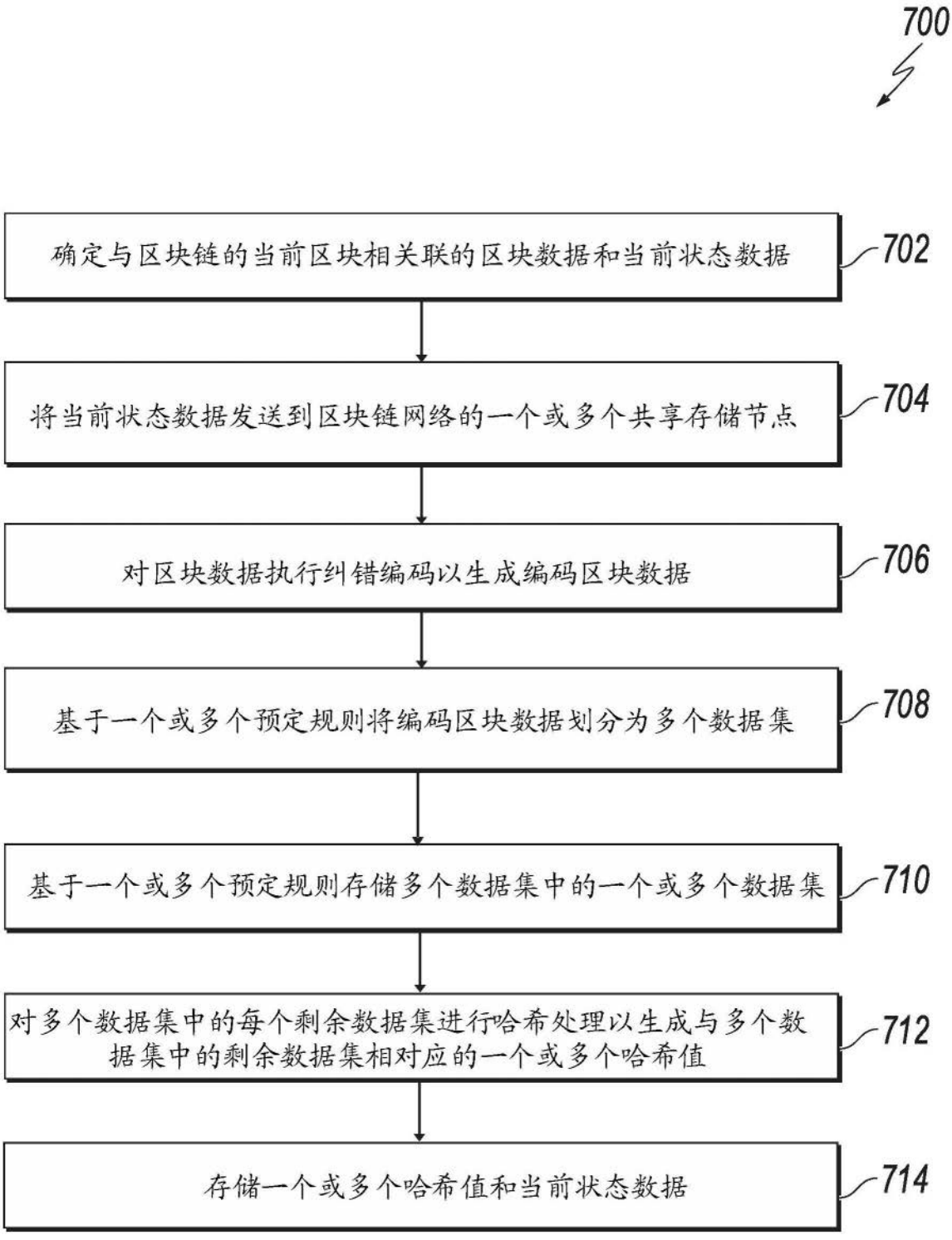

700
确定与区块链的当前区块相关联的区块数据和当前状态数据
702
将当前状态数据发送到区块链网络的一个或多个共享存储节点
704
对区块数据执行纠错编码以生成编码区块数据
706
基于一个或多个预定规则将编码区块数据划分为多个数据集
708
基于一个或多个预定规则存储多个数据集中的一个或多个数据集
710
对多个数据集中的每个剩余数据集进行哈希处理以生成与多个数据集中的剩余数据集相对应的一个或多个哈希值
712
存储一个或多个哈希值和当前状态数据
714

图7

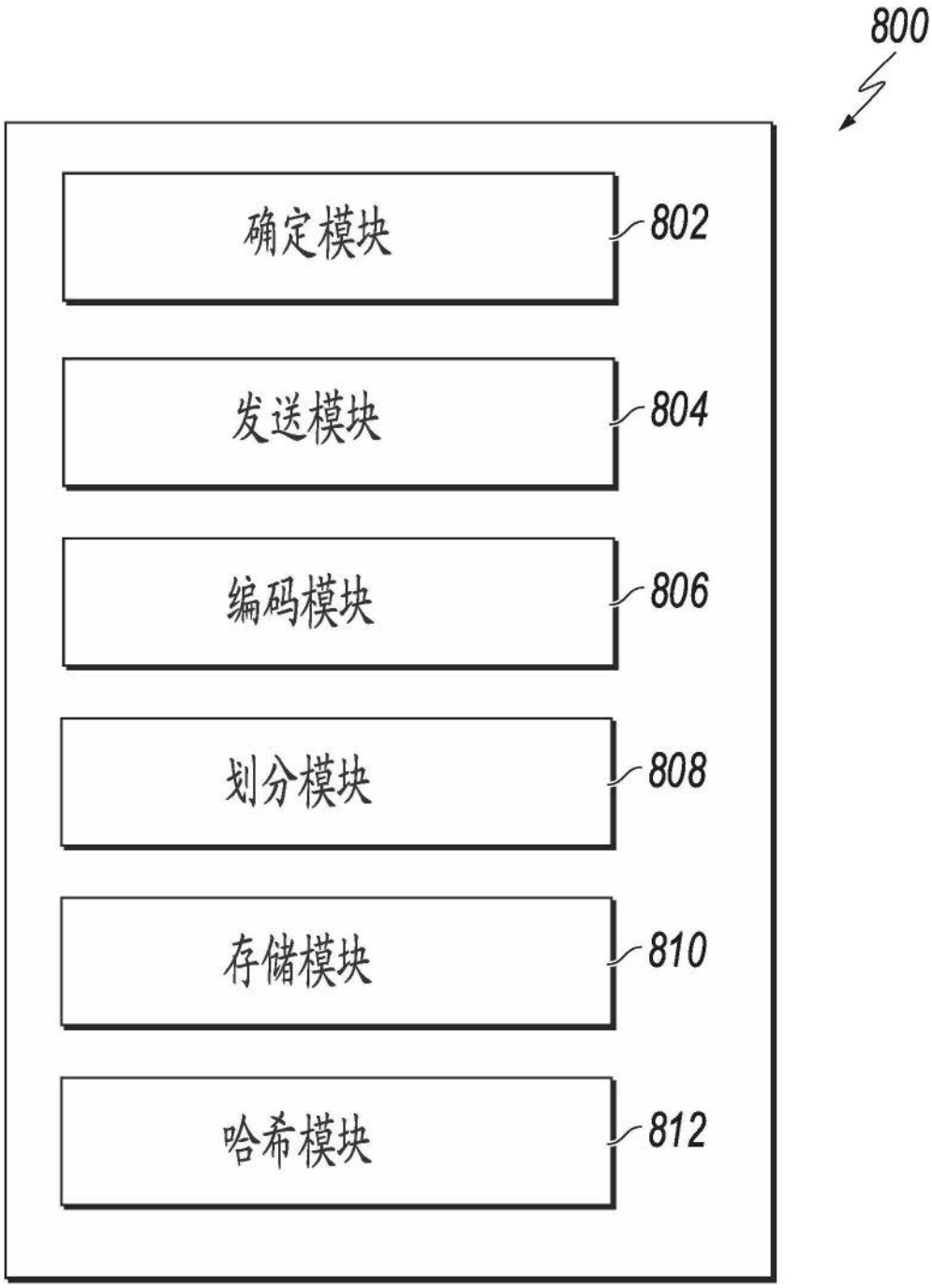


图8