## (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 111611319 A (43)申请公布日 2020.09.01

(21)申请号 202010511010.7

(22)申请日 2020.06.08

(71)申请人 杭州复杂美科技有限公司 地址 310000 浙江省杭州市西湖区文三路 90号东部软件园6号楼7层702室

(72)发明人 袁新强 王志文 吴思进

(51) Int.CI.

*G06F* 16/27(2019.01) *G06F* 16/22(2019.01)

HO4L 12/733(2013.01)

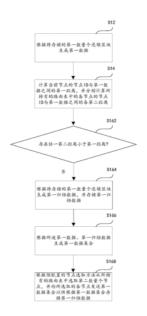
权利要求书2页 说明书6页 附图3页

#### (54)发明名称

分布式数据存储方法、设备和存储介质

#### (57)摘要

本发明提供一种分布式数据存储方法、设备和存储介质,该方法包括:根据待存储的第一数量个连续区块生成第一数据;计算当前节点的节点ID与第一数据之间的第一距离,并分别计算所持有的路由表中的各节点的节点ID与第一数据之间的各第二距离;判断是否存在任一第二距离小于第一距离:否,则根据待存储的第一数量个连续区块生成第一归档数据,并存储第一归档数据;以及,根据第一数据、第一归档数据生成第一数据集合;以及,根据预配置的节点选取方法从所持有的路由表中选取第二数量个节点,并向所选取的各节点发送第一数据集合以供根据第一数据集合存储第一归档数据。本申请保证分布式存储的区块数据不会丢失。



CN 111611319 A

1.一种分布式数据存储方法,其特征在于,适用于区块链节点,所述方法包括:

根据待存储的第一数量个连续区块生成第一数据;

计算当前节点的节点ID与所述第一数据之间的第一距离,并分别计算所持有的路由表中的各节点的节点ID与所述第一数据之间的各第二距离;

判断是否存在任一所述第二距离小于所述第一距离:

否,则根据所述待存储的第一数量个连续区块生成第一归档数据,并存储所述第一归档数据:以及,

根据所述第一数据、所述第一归档数据生成第一数据集合:以及,

根据预配置的节点选取方法从所持有的路由表中选取第二数量个节点,并向所选取的各节点发送所述第一数据集合以供根据所述第一数据集合存储所述第一归档数据。

2.根据权利要求1所述的方法,其特征在于,所述根据预配置的节点选取方法从所持有的路由表中选取第二数量个节点包括:

根据所述第二距离从小到大的顺序从所持有的路由表中选取第二数量个节点。

3.根据权利要求1所述的方法,其特征在于,所述根据待存储的第一数量个连续区块生成第一数据包括:

根据待存储的第一数量个连续区块的区块哈希生成第一数据。

4.根据权利要求1-3任一项所述的方法,其特征在于,区块链网络为kad网络,所存储的第一归档数据配置有过期时刻,所述方法还包括:

接收其它区块链节点发送的第二数据集合;其中,所述第二数据集合为其它区块链节点根据第二数据、所述第一归档数据所生成,所述第二数据为其它区块链节点根据所存储的所述第一归档数据所生成;

根据接收所述第二数据集合的时刻更新所述过期时刻;

判断当前时刻是否为预配置的时刻:

是,则检查当前时刻是否超过所述过期时刻:

是,则删除所存储的所述第一归档数据;

否,则返回所述根据预配置的节点选取方法从所持有的路由表中选取第二数量个节点。

5.一种分布式数据存储方法,其特征在于,适用于区块链节点,所述方法包括:

接收第一区块链节点发送的第一数据集合;其中,所述第一数据集合包括第一数据和第一归档数据,所述第一数据由所述第一区块链节点根据待存储的第一数量个连续区块所生成,所述第一数据集合由所述第一区块链节点在判断第一距离不小于任一第二距离时所生成,所述第一距离根据所述第一区块链节点的节点id与所述第一数据计算得到,各所述第二距离根据所述第一区块链节点所持有的路由表的各节点的节点ID与所述第一数据计算得到,所述第一归档数据由所述第一区块链节点根据待存储的第一数量个连续区块所生成,所述第一区块链节点还用于将所述第一数据集合发送至根据预配置的节点选取方法从所持有的路由表中选取的第二数量个节点;

根据所述第一数据集合存储所述第一归档数据。

6.根据权利要求5所述的方法,其特征在于,所述将所述第一数据集合发送至根据预配 置的节点选取方法从所持有的路由表中选取的第二数量个节点包括: 将所述第一数据集合发送至根据所述第二距离从小到大的顺序从所持有的路由表中选取的第二数量个节点。

- 7.根据权利要求5所述的方法,其特征在于,所述第一数据由所述第一区块链节点根据 待存储的第一数量个连续区块的区块哈希所生成。
- 8.根据权利要求5-7任一项所述的方法,其特征在于,区块链网络为kad网络,所存储的 所述第一归档数据配置有过期时刻,所述方法还包括:

接收所述第一数据集合或其它区块链节点发送的第二数据集合;其中,所述第二数据 集合为其它区块链节点根据第二数据、第二归档数据所生成,所述第二数据和第二数据集 合为其它区块链节点根据所存储的所述第一归档数据所生成;

根据接收所述第一数据集合或所述第二数据集合的时刻更新所述过期时刻;

判断当前时刻是否为预配置的时刻:

是,则检查当前时刻是否超过所述过期时刻:

是,则删除所存储的所述第一归档数据;

否,则根据所存储的所述第一归档数据生成第三数据,并根据所述第三数据、所述第一 归档数据生成第三数据集合;以及,

根据预配置的节点选取方法从所持有的路由表中选取第二数量个节点;以及,

向所选取的各节点发送所述第三数据集合。

- 9.一种设备,其特征在于,所述设备包括:
- 一个或多个处理器:

存储器,用于存储一个或多个程序,

当所述一个或多个程序被所述一个或多个处理器执行时,使得所述一个或多个处理器执行如权利要求1-8中任一项所述的方法。

10.一种存储有计算机程序的存储介质,其特征在于,该程序被处理器执行时实现如权利要求1-8中任一项所述的方法。

# 分布式数据存储方法、设备和存储介质

#### 技术领域

[0001] 本申请涉及分布式存储和区块链技术领域,具体涉及一种分布式数据存储方法、设备和存储介质。

## 背景技术

[0002] 在现有技术中,假设每个节点都存有block (1)~block (10999),回滚深度为10000;在当前节点生成block (11000)时,向其它区块链节点广播block (11000),具体的广播、共识机制在此不做赘述;当block (11000)执行成功时即认为block (1)~block (1000)不会被改变了;当前节点根据block (1)~block (1000) 计算第一数据chunkhash,节点与chunkhash的距离的计算方法为:xor (节点id,chunkhash);当前节点通过迭代查询的方式最终找到D1为全局最近节点 (即xor (节点id (D1),chunkhash)最小),当前节点向D1发送数据集合 {chunkhash、chunk},chunk为block (1)~block (1000)的归档数据 (归档数据可以为block (1)~block (1000) 本身,也可以配置为block (1)~block (1000)的对称加密数据、block (1)~block (1000)的压缩数据等可以还原出block (1)~block (1000)的数据,也可以配置为区块高度区间 [1,1000],D1收到数据集合时,在本地P2P模块存储chunk(当归档数据为区块高度区间 [1,1000],D1根据 [1,1000]生成chunk,在本地P2P模块存储所生成的chunk);在一段时间后,区块链中所有节点的blockchain模块都会删除block (1)~block (1000);

[0003] 在上述机制中,当前节点寻找D1的过程需要经过较长时间,若在此期间当前节点 宕机,则可能造成D1没有收到chunkhash,D1就不会在P2P模块额外打包一份block (1)~block (1000),在一段时间后,区块链中所有节点的blockchain模块都会删除block (1)~block (1000),整个区块链网络中没有地方可以查询到block (1)~block (1000)的区块数据。

#### 发明内容

[0004] 鉴于现有技术中的上述缺陷或不足,期望提供一种不会丢失区块数据的分布式数据存储方法、设备和存储介质。

[0005] 第一方面,本发明提供一种适用于区块链节点的分布式数据存储方法,上述方法包括:

[0006] 根据待存储的第一数量个连续区块生成第一数据;

[0007] 计算当前节点的节点ID与第一数据之间的第一距离,并分别计算所持有的路由表中的各节点的节点ID与第一数据之间的各第二距离;

[0008] 判断是否存在任一第二距离小于第一距离:

[0009] 否,则根据待存储的第一数量个连续区块生成第一归档数据,并存储第一归档数据;以及,

[0010] 根据第一数据、第一归档数据生成第一数据集合:以及,

[0011] 根据预配置的节点选取方法从所持有的路由表中选取第二数量个节点,并向所选取的各节点发送第一数据集合以供根据第一数据集合存储第一归档数据。

[0012] 第二方面,本发明提供一种适用于区块链节点的分布式数据存储方法,上述方法包括:

[0013] 接收第一区块链节点发送的第一数据集合;

[0014] 根据第一数据集合存储第一归档数据。

[0015] 第三方面,本发明还提供一种设备,包括一个或多个处理器和存储器,其中存储器包含可由该一个或多个处理器执行的指令以使得该一个或多个处理器执行根据本发明各实施例提供的分布式数据存储方法。

[0016] 第四方面,本发明还提供一种存储有计算机程序的存储介质,该计算机程序使计算机执行根据本发明各实施例提供的分布式数据存储方法。

[0017] 本发明诸多实施例提供的分布式数据存储方法、设备和存储介质通过根据待存储的第一数量个连续区块生成第一数据;计算当前节点的节点ID与第一数据之间的第一距离,并分别计算所持有的路由表中的各节点的节点ID与第一数据之间的各第二距离;判断是否存在任一第二距离小于第一距离:否,则根据待存储的第一数量个连续区块生成第一归档数据,并存储第一归档数据;以及,根据第一数据、第一归档数据生成第一数据集合;以及,根据预配置的节点选取方法从所持有的路由表中选取第二数量个节点,并向所选取的各节点发送第一数据集合以供根据第一数据集合存储第一归档数据的方法,保证区块数据被分布式存储。

### 附图说明

[0018] 通过阅读参照以下附图所作的对非限制性实施例所作的详细描述,本申请的其它特征、目的和优点将会变得更明显:

[0019] 图1为本发明一实施例提供的一种分布式数据存储方法的流程图。

[0020] 图2为图1所示方法的一种优选实施方式的流程图。

[0021] 图3为本发明一实施例提供的另一种分布式数据存储方法的流程图。

[0022] 图4为本发明一实施例提供的一种设备的结构示意图。

#### 具体实施方式

[0023] 下面结合附图和实施例对本申请作进一步的详细说明。可以理解的是,此处所描述的具体实施例仅仅用于解释相关发明,而非对该发明的限定。另外还需要说明的是,为了便于描述,附图中仅示出了与发明相关的部分。

[0024] 需要说明的是,在不冲突的情况下,本申请中的实施例及实施例中的特征可以相互组合。下面将参考附图并结合实施例来详细说明本申请。

[0025] 图1为本发明一实施例提供的一种分布式数据存储方法的流程图。如图1所示,在本实施例中,本发明提供一种适用于区块链节点的分布式数据存储方法,上述方法包括:

[0026] S12:根据待存储的第一数量个连续区块生成第一数据:

[0027] S14: 计算当前节点的节点ID与第一数据之间的第一距离,并分别计算所持有的路由表中的各节点的节点ID与第一数据之间的各第二距离;

[0028] S162:判断是否存在任一第二距离小于第一距离:

[0029] 否,则执行步骤S164:根据待存储的第一数量个连续区块生成第一归档数据,并存储第一归档数据:以及,

[0030] S166:根据所述第一数据、所述第一归档数据生成第一数据集合:以及,

[0031] S168:根据预配置的节点选取方法从所持有的路由表中选取第二数量个节点,并向所选取的各节点发送第一数据集合以供根据第一数据集合存储第一归档数据。

[0032] 具体地,以将"根据预配置的节点选取方法从所持有的路由表中选取第二数量个节点"配置为"根据第二距离从小到大的顺序从所持有的路由表中选取第二数量个节点";将"根据待存储的第一数量个连续区块生成第一数据"配置为"根据待存储的第一数量个连续区块生成第一数据为1000,第二数量为10;第一数据集合包括第一数据和待存储的第一数量个连续区块;生成的归档数据为第一数量个连续区块本身为例;假设待存储的第一数量个连续区块为block(1)~block(1000);节点N0、M0、P0、Q0为第一距离不小于任一第二距离的节点;

[0033] 假设当前节点为节点NO:

[0034] N0执行步骤S12,根据blockhash(1)~blockhash(1000)生成第一数据chunkhash;

[0035] N0执行步骤S14,计算当前节点的节点ID与第一数据之间的第一距离,并分别计算 所持有的路由表中的各节点的节点ID与第一数据之间的各第二距离:

[0036] N0执行步骤S162,判断是否存在任一第二距离小于第一距离:

[0037] 由于第一距离不小于任一第二距离,则执行步骤S164,根据待存储的block(1)~block(1000)生成归档数据chunk,在本地P2P模块存储chunk;以及,

[0038] N0执行步骤S166,根据chunkhash生成第一数据集合 {chunkhash,chunk};以及,

[0039] N0执行步骤S168,根据第二距离从小到大的顺序从所持有的路由表中选取10个节点(假设为 $N1\sim N10$ ),并向 $N1\sim N10$ 发送 {chunkhash,chunk)};

[0040] N1~N10根据{chunkhash,chunk}在本地P2P模块存储chunk。

[0041] 同理,M0、P0、Q0执行步骤S12-S168,假设M0所选取的10个节点为M1~M10;P0所选取的10个节点为P1~P10;Q0所选取的10个节点为Q10;则在区块链网络中,在本地P2P模块存储有Chunk的节点为:M00M100,M00M100,M00M100,M00M100.

[0042] 在一段时间(假设为12h)后,区块链网络中的所有节点将删除blockchain模块中的block(1)~block(1000),即最终区块链网络中只有N0~N10、M0~M10、P0~P10、Q0~Q10的本地P2P模块存储有block(1)~block(1000)。

[0043] 用户的用户端或其它区块链节点可以从N0~N10、M0~M10、P0~P10、Q0~Q10处获取chunk,在获取到chunk后,在本地的blockchain模块还原出block(1)~block(1000)以获得block(1)~block(1000)的区块数据。

[0044] 在更多实施例中,"根据预配置的节点选取方法从所持有的路由表中选取第二数量个节点"还可以根据实际需求进行配置,例如配置为:根据物理距离从小到大的的顺序从所持有的路由表中选取第二数量个节点;可实现相同的技术效果。

[0045] 在更多实施例中,"根据待存储的第一数量个连续区块生成第一数据"还可以根据实际需求进行配置,例如配置为:根据待存储的第一数量个连续区块的区块父哈希生成第一数据;可实现相同的技术效果。

[0046] 在更多实施例中,第一数量,第二数量均可根据实际需求进行配置,可实现相同的技术效果。

[0047] 在更多实施例中,第一归档数据还可以根据实际需求进行配置,例如配置为: block(1)~block(1000)的压缩数据,可实现相同的技术效果。

[0048] 上述实施例不会丢失区块数据。

[0049] 优选地,根据预配置的节点选取方法从所持有的路由表中选取第二数量个节点包括:

[0050] 根据第二距离从小到大的顺序从所持有的路由表中选取第二数量个节点。

[0051] 上述实施例的分布式数据存储原理可参考图1所示的方法,此处不再赘述。

[0052] 优选地,根据待存储的第一数量个连续区块生成第一数据包括:

[0053] 根据待存储的第一数量个连续区块的区块哈希生成第一数据。

[0054] 上述实施例的分布式数据存储原理可参考图1所示的方法,此处不再赘述。

[0055] 优选地,待存储的第一数量个连续区块由当前区块高度和区块回滚深度决定。

[0056] 图2为图1所示方法的一种优选实施方式的流程图。如图2所示,在一优选实施例中,区块链网络为kad网络,所存储的第一归档数据配置有过期时刻,上述方法还包括:

[0057] S17:接收其它区块链节点发送的第二数据集合;其中,第二数据集合为其它区块链节点根据第二数据、第一归档数据所生成,第二数据为其它区块链节点根据所存储的第一归档数据所生成;

[0058] S18:根据接收第二数据集合的时刻更新过期时刻;

[0059] S191:判断当前时刻是否为预配置的时刻:

[0060] 是,则执行步骤S192:检查当前时刻是否超过过期时刻:

[0061] 是,则执行步骤S1921:删除所存储的第一归档数据;

[0062] 否,则执行步骤S1922:返回根据预配置的节点选取方法从所持有的路由表中选取第二数量个节点。

[0063] 在现有的kad网络中配置有动态平衡策略,具体为:全局最优节点(假设为图1所示实施例中的N0)向N1~N10发送第一数据集合{chunkhash,chunk},N1~N10在本地P2P模块存储chunk;以当前节点为N1为例,假设N1收到{chunkhash,chunk}时为2020年5月1日的2:00:00;N1将chunk的过期时刻更新为2020年5月2日的2:00:00;假设预配置的时刻为收到{chunkhash,chunk}时的每隔10小时,在5月1日的2:00:00至5月2日的07:59:59,N1一直未收到其它区块链节点广播的{chunkhash,chunk};N1在5月2日的08:00:00,检测到当前时刻已超过更新的过期时刻,则删除所存储的chunk;假设N1在5月2日的4:00:00又接收到{chunkhash,chunk},N1将过期时刻更新为5月3日的4:00:00;由于kad网络中的k桶机制,N0~N10都大概率在彼此所持有的路由表中,N1将{chunkhash,chunk}发送给N0、N2~N10,以供更新所存储的chunk的过期时刻。

[0064] 由于节点存在宕机的可能,在极端情况下,N0~N10全部宕机,则区块链网络中也会失去blockhash(1)~blockhash(1000);配置动态平衡策略在于,假设N1突然宕机,N0也会从所持有的路由表中选出10个第二距离最小的节点发送第一数据集合,使得区块链网络中一定有11个节点持有chunk以供还原出blockhash(1)~blockhash(1000)。

[0065] 本申请的区块链网络为kad网络,配置有同样的动态平衡策略,步骤如S17~

S1922,此处不再赘述。

[0066] 上述实施例减少了区块链网络中分布式数据的备份数量。

[0067] 图3为本发明一实施例提供的另一种分布式数据存储方法的流程图。如图3所示, 在本实施例中,本发明提供一种适用于区块链节点的分布式数据存储方法,上述方法包括:

[0068] S22:接收第一区块链节点发送的第一数据集合;其中,第一数据集合包括第一数据和第一归档数据,第一数据由第一区块链节点根据待存储的第一数量个连续区块所生成,第一数据集合由第一区块链节点在判断第一距离不小于任一第二距离时所生成,第一距离根据第一区块链节点的节点id与第一数据计算得到,各第二距离根据第一区块链节点所持有的路由表的各节点的节点ID与第一数据计算得到,第一归档数据由第一区块链节点根据待存储的第一数量个连续区块所生成,第一区块链节点还用于将第一数据集合发送至根据预配置的节点选取方法从所持有的路由表中选取的第二数量个节点;

[0069] 根据第一数据集合存储第一归档数据。

[0070] 上述实施例的分布式数据存储原理可参考图1所示的方法,此处不再赘述。

[0071] 优选地,将第一数据集合发送至根据预配置的节点选取方法从所持有的路由表中 选取的第二数量个节点包括:

[0072] 将第一数据集合发送至根据第二距离从小到大的顺序从所持有的路由表中选取的第二数量个节点。

[0073] 上述实施例的分布式数据存储原理可参考图1所示的方法,此处不再赘述。

[0074] 优选地,第一数据由第一区块链节点根据待存储的第一数量个连续区块的区块哈希所生成。

[0075] 上述实施例的分布式数据存储原理可参考图1所示的方法,此处不再赘述。

[0076] 优选地,待存储的第一数量个连续区块由当前区块高度和区块回滚深度决定。

[0077] 优选地,区块链网络为kad网络,所存储的第一数量个连续区块配置有相应的过期时刻,方法还包括:

[0078] 区块链网络为kad网络,所存储的第一归档数据配置有过期时刻,方法还包括:

[0079] 接收第一数据集合或其它区块链节点发送的第二数据集合;其中,第二数据集合为其它区块链节点根据第二数据、第二归档数据所生成,第二数据和第二数据集合为其它区块链节点根据所存储的第一归档数据所生成;

[0080] 根据接收第一数据集合或第二数据集合的时刻更新过期时刻;

[0081] 判断当前时刻是否为预配置的时刻:

[0082] 是,则检查当前时刻是否超过过期时刻:

[0083] 是,则删除所存储的第一归档数据;

[0084] 否,则根据所存储的第一归档数据生成第三数据,并根据第三数据、第一归档数据生成第三数据集合;以及,

[0085] 根据预配置的节点选取方法从所持有的路由表中选取第二数量个节点;以及,

[0086] 向所选取的各节点发送第三数据集合。

[0087] 上述实施例的分布式数据存储原理可参考图2所示的方法,此处不再赘述。

[0088] 图4为本发明一实施例提供的一种设备的结构示意图。

[0089] 如图4所示,作为另一方面,本申请还提供了一种设备500,包括一个或多个中央处

理单元 (CPU) 501,其可以根据存储在只读存储器 (ROM) 502中的程序或者从存储部分508加载到随机访问存储器 (RAM) 503中的程序而执行各种适当的动作和处理。在RAM503中,还存储有设备500操作所需的各种程序和数据。CPU501、ROM502以及RAM503通过总线504彼此相连。输入/输出 (I/0) 接口505也连接至总线504。

[0090] 以下部件连接至I/0接口505:包括键盘、鼠标等的输入部分506;包括诸如阴极射线管(CRT)、液晶显示器(LCD)等以及扬声器等的输出部分507;包括硬盘等的存储部分508;以及包括诸如LAN卡、调制解调器等的网络接口卡的通信部分509。通信部分509经由诸如因特网的网络执行通信处理。驱动器510也根据需要连接至I/0接口505。可拆卸介质511,诸如磁盘、光盘、磁光盘、半导体存储器等等,根据需要安装在驱动器510上,以便于从其上读出的计算机程序根据需要被安装入存储部分508。

[0091] 特别地,根据本公开的实施例,上述任一实施例描述的方法可以被实现为计算机软件程序。例如,本公开的实施例包括一种计算机程序产品,其包括有形地包含在机器可读介质上的计算机程序,所述计算机程序包含用于执行上述任一方法的程序代码。在这样的实施例中,该计算机程序可以通过通信部分509从网络上被下载和安装,和/或从可拆卸介质511被安装。

[0092] 作为又一方面,本申请还提供了一种计算机可读存储介质,该计算机可读存储介质可以是上述实施例的装置中所包含的计算机可读存储介质;也可以是单独存在,未装配入设备中的计算机可读存储介质。计算机可读存储介质存储有一个或者一个以上程序,该程序被一个或者一个以上的处理器用来执行描述于本申请提供的方法。

[0093] 附图中的流程图和框图,图示了按照本发明各种实施例的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上,流程图或框图中的每个方框可以代表一个模块、程序段、或代码的一部分,该模块、程序段、或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意,在有些作为替换的实现中,方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如,两个接连地表示的方框实际上可以基本并行地执行,它们有时也可以按相反的顺序执行,这根据所涉及的功能而定。也要注意的是,框图和/或流程图中的每个方框、以及框图和/或流程图中的方框的组合,可以通过执行规定的功能或操作的专用的基于硬件的系统来实现,或者可以通过专用硬件与计算机指令的组合来实现。

[0094] 描述于本申请实施例中所涉及到的单元或模块可以通过软件的方式实现,也可以通过硬件的方式来实现。所描述的单元或模块也可以设置在处理器中,例如,各所述单元可以是设置在计算机或移动智能设备中的软件程序,也可以是单独配置的硬件装置。其中,这些单元或模块的名称在某种情况下并不构成对该单元或模块本身的限定。

[0095] 以上描述仅为本申请的较佳实施例以及对所运用技术原理的说明。本领域技术人员应当理解,本申请中所涉及的发明范围,并不限于上述技术特征的特定组合而成的技术方案,同时也应涵盖在不脱离本申请构思的情况下,由上述技术特征或其等同特征进行任意组合而形成的其它技术方案。例如上述特征与本申请中公开的(但不限于)具有类似功能的技术特征进行互相替换而形成的技术方案。

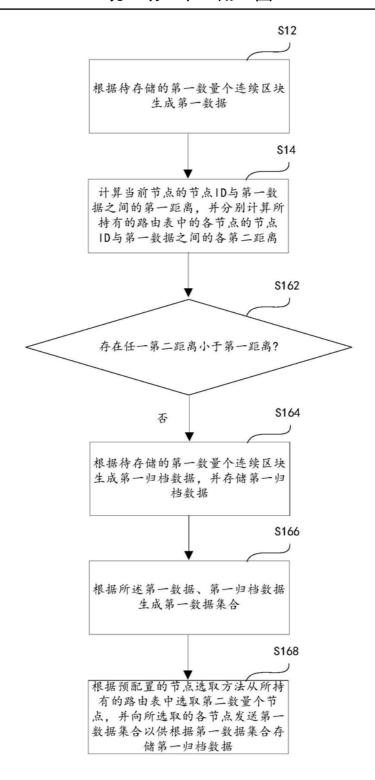


图1

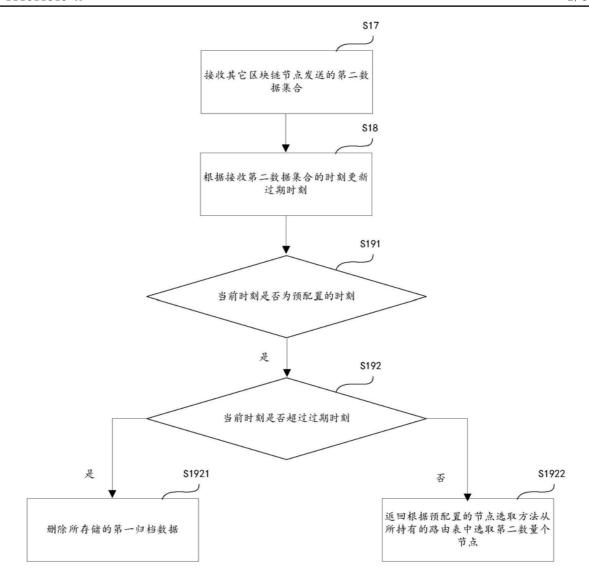


图2

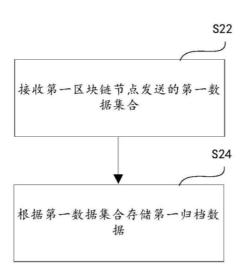


图3

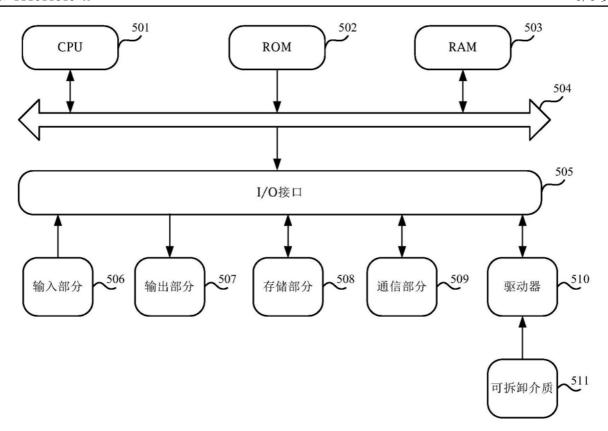


图4