



(12)发明专利申请

(10)申请公布号 CN 108134706 A
(43)申请公布日 2018.06.08

(21)申请号 201810001668.6

(22)申请日 2018.01.02

(71)申请人 中国工商银行股份有限公司
地址 100140 北京市西城区复兴门内大街
55号

(72)发明人 赖春晖 陈满才 赵开山 刘朝伟
彭顺求 陈法山

(74)专利代理机构 北京三友知识产权代理有限公司 11127
代理人 乔媛 汤在彦

(51)Int.Cl.
H04L 12/24(2006.01)
H04L 29/08(2006.01)

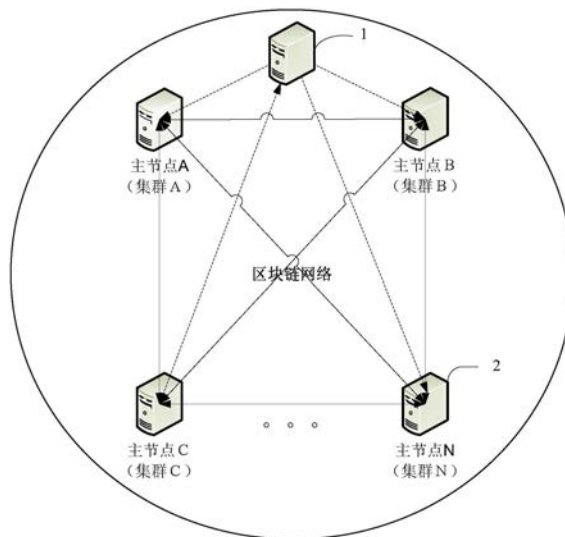
权利要求书4页 说明书17页 附图16页

(54)发明名称

区块链多活高可用系统、计算机设备以及方法

(57)摘要

本发明提供了一种区块链多活高可用系统、方法、计算机设备以及计算机可读存储介质,涉及区块链技术领域。所述系统包括一认证节点以及多个区块链节点,每个区块链节点为一集群,集群中的主节点参与区块链,区块链网络由多个集群组成,区块链之间的通讯由主节点完成,当区块链达成共识之后,每个集群内部的主节点将区块链数据同步更新至从节点以及普通节点中,当集群中的一故障集群的主节点失效或共识出错时,通过主从选举规则进行主从选举,将从节点切换为主节点提供区块链服务,认证节点通知故障集群之外的其他集群更新故障集群的主节点信息。本发明大大提高了整个区块链的可靠性,满足银行7*24小时的高可用性要求。



1. 一种区块链多活高可用系统,其特征是,所述系统包括一认证节点以及多个区块链节点,每个所述区块链节点为一集群,每个所述集群内部通过主从选举规则从多个普通节点中选举出一主节点以及一从节点,所述集群中的主节点参与区块链,

其中,区块链网络由多个集群组成,区块链之间的通讯由所述主节点完成,当区块链达成共识之后,每个集群内部所述主节点将区块链数据同步更新至所述从节点以及所述普通节点中,当所述集群中的一故障集群的主节点失效或共识出错时,通过主从选举规则进行主从选举,将所述从节点切换为主节点提供区块链服务,所述认证节点通知所述故障集群之外的其他集群更新所述故障集群的主节点信息。

2. 根据权利要求1所述的系统,其特征是,所述认证节点包括身份认证单元、主节点数据单元、主节点认证单元、主节点保活单元、主节点更新单元、主节点切换单元和主节点查询单元,

其中,所述身份认证单元,用于对节点进行身份验证,当所述节点通过身份验证后,将所述节点加入集群;

所述主节点数据单元,用于维护和存储所述集群中的主节点信息,所述主节点信息包括集群ID、主节点IP地址和主节点老化计时器;

所述主节点认证单元,用于当通过身份验证且所述集群的主节点未曾注册时,添加主节点信息至所述主节点数据单元,并返回注册结果给所述集群的主节点,注册成功时,所述注册结果包含集群ID信息,所述集群ID信息为所述认证节点通过认证节点公钥加密得到;

所述主节点保活单元,用于接收所述主节点发送的主节点保活数据包,根据所述主节点老化计时器判断所述主节点是否活动;

所述主节点更新单元,用于接收所述主节点发送的主节点更新数据包;

所述主节点切换单元,用于接收所述主节点发送的故障数据包,所述故障数据包满足预设条件时,向所述主节点所在的所述集群发送强制切换主节点通知;

所述主节点查询单元,用于接收所述主节点发送的主节点查询数据包,根据所述主节点查询数据包查询所述主节点数据单元,并返回查询结果给所述主节点。

3. 根据权利要求2所述的系统,其特征是,所述主节点包括主节点外控单元、主节点内控单元和主节点区块链管理单元,

其中,所述主节点外控单元,用于与所述认证节点进行数据交互;

所述主节点内控单元,用于负责所述集群中主节点、从节点以及普通节点之间的数据交互;

所述主节点区块链管理单元,用于代表所述集群参与区块链通讯,当区块链达成共识后,将区块链数据同步更新至所述从节点以及所述普通节点。

4. 根据权利要求3所述的系统,其特征是,所述主节点外控单元包括主节点注册单元、更新单元、保活单元、查询单元、注册数据单元、集群身份数据单元、集群身份管理单元、共识出错通知单元和切换单元,

其中,所述主节点注册单元,用于当通过主从节点选举规则选举为集群初始主节点后,代表所述集群向所述认证节点发起主节点注册数据包请求,注册成功后,获取所述认证节点通过认证节点公钥加密后的集群ID信息,并将所述集群ID信息存储在注册数据单元中;

所述更新单元,用于当通过主从节点选举规则选举为集群更新主节点后,代表所述集

群向所述认证节点发起主节点更新数据包请求,主节点更新数据包请求中包含所述认证节点通过公钥加密后的集群ID信息,所述集群ID信息从所述注册数据单元中获得;

所述保活单元,用于当所述主节点注册单元注册成功或所述更新单元更新成功后,代表所述集群通过保活定时器定时向所述认证节点发起主节点保活数据包,定时刷新所述集群的主节点的老化计时器,以保持认证节点信息可用;

所述查询单元,用于当所述主节点需要通过其他集群的主节点信息加入区块链网络时,向所述认证节点发送主节点查询数据包,从所述认证节点中查询相应集群的主节点的信息;

所述注册数据单元,用于存储所述主节点注册单元向所述认证节点发起注册请求后的认证节点返回信息,并同时存储保活定时器,当所述保活定时器到期时重置保活定时器,并向所述认证节点发送主节点保活数据包的请求;

所述集群身份数据单元,用于存储所述集群身份管理单元通过所述认证节点发起加入集群后认证节点返回的信息;

所述集群身份管理单元,用于向所述认证节点发起加入集群的认证请求,并将认证节点返回的信息存储于所述集群身份数据单元;

所述共识出错通知单元,用于当区块链共识发现一集群的主节点计算出错时,称发生出错的集群为故障集群,向认证节点发送所述故障集群的主节点出错的通知;

所述切换单元,用于接收所述认证节点发出的强制主节点切换信息,并通过所述主节点内控单元进行集群内部的主从选举。

5. 根据权利要求4所述的系统,其特征是,所述主节点内控单元包含邻居数据单元、主从选举单元、节点保活单元、注册数据同步和区块链数据同步单元,

其中,所述邻居数据单元,用于存储所述集群内部所有节点邻居的信息,所述信息包含集群ID、节点优先级、节点IP地址、节点角色以及节点老化计时器;

所述主从选举单元,用于负责所述集群的集群主从选举运算,选举完成后更新所述邻居数据单元中的节点角色信息;

所述节点保活单元,用于维护所述集群的保活定时器,当所述保活定时器到期时,重置所述保活定时器,并向所述集群内的从节点以及普通节点定时发送节点广播保活数据包;

所述注册数据同步单元,用于当所述主节点向所述认证节点注册成功或者收到所述集群内所述从节点或普通节点的注册数据更新请求后,同步更新注册数据到集群内的从节点或普通节点;

所述区块链数据同步单元,用于当达成共识或收到所述集群内从节点或普通节点的区块链数据更新请求后,负责同步更新区块链数据到集群内的从节点或普通节点中。

6. 一种根据区块链多活高可用系统进行主节点故障恢复的方法,其特征是,所述方法包括:

一认证节点确定出一故障集群的主节点处于失效状态时,删除所述认证节点中所述故障集群的主节点对应的信息;

所述故障集群内的从节点以及普通节点确定出所述主节点处于失效状态时,删除所述主节点对应的信息;

当所述从节点、普通节点确定出所述主节点中存在所述故障集群的从节点时,将所述

从节点选举为所述集群的新主节点；

所述新主节点向所述认证节点发送所述故障集群的主节点更新数据包；

所述认证节点根据所述主节点更新数据包将所述新主节点对应的信息添加至主节点数据单元中，标识为所述故障集群的主节点；

所述认证节点向所述故障集群内的普通节点、新主节点和其他集群的主节点发送所述故障集群的主节点更新数据包；

所述故障集群的普通节点以及所述新主节点根据所述主节点更新数据包更新主节点信息。

7. 根据权利要求6所述的方法，其特征是，所述方法还包括：

所述认证节点检查所述主节点更新数据包，判断所述新主节点的身份是否合法；

当判断为是时，执行所述认证节点根据所述主节点更新数据包将所述新主节点对应的信息添加至主节点数据单元中，标识为所述故障集群的主节点的步骤；

否则，所述认证节点拒绝更新所述主节点数据单元。

8. 根据权利要求7所述的方法，其特征是，所述方法还包括：

所述新主节点、普通节点根据邻居数据单元中各邻居节点信息通过主从选举规则进行选举，确定出新从节点。

9. 一种根据区块链多活高可用系统进行主从节点故障恢复的方法，其特征是，所述方法包括：

一认证节点确定出一故障集群的主节点处于失效状态时，删除所述认证节点中所述故障集群的主节点对应的信息；

所述故障集群的普通节点确定出所述主节点和从节点处于失效状态时，删除所述主节点对应的信息以及所述从节点对应的信息；

当所述普通节点确定出主节点、从节点不存在时，根据邻居数据单元中的邻居节点信息过主从选举规则进行选举，确定出新主节点以及新从节点；

所述新主节点向所述认证节点发送所述故障集群的主节点更新数据包；

所述认证节点根据所述主节点更新数据包将所述新主节点对应的信息添加至主节点数据单元中，标识为所述故障集群的主节点；

所述认证节点向所述故障集群内的普通节点、新主节点和其他集群的主节点发送所述故障集群的主节点更新数据包；

所述故障集群的普通节点以及所述新主节点根据所述主节点更新数据包更新主节点信息。

10. 根据权利要求9所述的方法，其特征是，所述方法还包括：

所述认证节点检查所述主节点更新数据包，判断所述新主节点的身份是否合法；

当判断为是时，执行所述认证节点根据所述主节点更新数据包将所述新主节点对应的信息添加至主节点数据单元中，标识为所述故障集群的主节点的步骤；

否则，所述认证节点拒绝更新所述主节点数据单元。

11. 一种根据区块链多活高可用系统进行主节点共识错误恢复的方法，其特征是，所述方法包括：

一认证节点统计一故障集群的主节点的集群错误数量；

当所述集群错误数量达到一阈值时,所述认证节点向所述故障集群内的主节点、从节点以及普通节点发送主节点强制切换数据包;

所述故障集群根据所述主节点强制切换数据包调整所述主节点为不可选举状态;

故障机群的从节点以及普通节点确定出邻居数据单元中主节点不存在、从节点存在,将所述从节点选举为所述故障集群的新主节点;

所述新主节点向所述认证节点发送所述故障集群的主节点更新数据包;

所述认证节点根据所述主节点更新数据包将所述新主节点对应的信息添加至主节点数据单元中,标识为所述故障集群的主节点;

所述认证节点向所述故障集群内的普通节点、新主节点和其他集群的主节点发送所述故障集群的主节点更新数据包;

所述故障集群的普通节点以及所述新主节点根据所述主节点更新数据包更新主节点信息。

12. 根据权利要求11所述的方法,其特征是,所述方法还包括:

认证节点接收一主节点共识错误数据包,所述主节点共识错误数据包由一集群在进行区块链共识的过程中检测到仅有故障集群的主节点的计算结果不相同发送。

13. 根据权利要求11所述的方法,其特征是,所述方法还包括:

所述认证节点检查所述主节点更新数据包,判断所述新主节点的身份是否合法;

当判断为是时,执行所述认证节点根据所述主节点更新数据包将所述新主节点对应的信息添加至主节点数据单元中,标识为所述故障集群的主节点的步骤;

否则,所述认证节点拒绝更新所述主节点数据单元。

14. 一种计算机设备,包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序,其特征在于,所述处理器执行所述计算机程序时实现权利要求6至13任一所述方法。

15. 一种计算机可读存储介质,其特征在于,所述计算机可读存储介质存储有执行权利要求6至13任一所述方法的计算机程序。

区块链多活高可用系统、计算机设备以及方法

技术领域

[0001] 本发明关于区块链技术领域,具体的讲是一种区块链多活高可用系统、方法、计算机设备以及计算机可读存储介质。

背景技术

[0002] 本部分旨在为权利要求书中陈述的本发明的实施方式提供背景或上下文。此处的描述不因为包括在本部分中就承认是现有技术。

[0003] 随着区块链的发展,区块链从公有链逐步转到私有链或联盟链中,但在这些私有链或联盟链的区块链产品中,面对如银行等关键行业,对区块链网络可用性要求非常严格,需要保持业务一直可用,需要提供7*24小时区块链服务,特别是在跨机构或者跨中心部署区块链网络时,需要实现区块链网络多活的部署。除此之外,当前区块链容错功能都有限,如当前普遍通过PBFT实现的区块链,由于该PBFT算法($F \leq (N-1) / 3$,N为部署节点数量,F为容错的节点数量),在N节点情况下,最多只能容错F个节点,而无法突破F个节点的容错。另外在节点出现故障时,区块链没有机制实现节点的快速恢复,当业务频繁区块链数据量大时,节点恢复需要很长的时间。

[0004] 当前区块链存在容错功能有限和节点无法快速恢复的问题,无法满足关键业务特别是银行业务多活的网络系统要求,因此,需要一种区块链多活的方案。

发明内容

[0005] 有鉴于此,本发明提供了一种区块链多活高可用系统、方法、计算机设备以及计算机可读存储介质,通过采取集群节点代替原来区块链节点,正常情况下由集群中的主节点参与区块链工作,在主节点出现异常时由其他集群积极实时代替,大大提高了整个区块链的可靠性,满足银行7*24小时的高可用性要求。

[0006] 为了实现上述目的,本发明提供了一种区块链多活高可用系统,所述系统包括一认证节点以及多个区块链节点,每个所述区块链节点为一集群,每个所述集群内部通过主从选举规则从多个普通节点中选举出一主节点以及一从节点,所述集群中的主节点参与区块链,

[0007] 其中,区块链网络由多个集群组成,区块链之间的通讯由所述主节点完成,当区块链达成共识之后,每个集群内部所述主节点将区块链数据同步更新至所述从节点以及所述普通节点中,当所述集群中的一故障集群的主节点失效或共识出错时,通过主从选举规则进行主从选举,将所述从节点切换为主节点提供区块链服务,所述认证节点通知所述故障集群之外的其他集群更新所述故障集群的主节点信息。

[0008] 本发明的目的之一是,提供了一种根据区块链多活高可用系统进行主节点故障恢复的方法,所述方法包括:

[0009] 一认证节点确定出一故障集群的主节点处于失效状态时,删除所述认证节点中所述集群的主节点对应的信息;

- [0010] 所述故障集群内的从节点以及普通节点确定出所述主节点处于失效状态时,删除所述主节点对应的信息;
- [0011] 当所述从节点、普通节点确定出所述主节点中存在所述故障集群的从节点时,将所述从节点选举为所述集群的新主节点;
- [0012] 所述新主节点向所述认证节点发送所述故障集群的主节点更新数据包;
- [0013] 所述认证节点根据所述主节点更新数据包将所述新主节点对应的信息添加至主节点数据单元中,标识为所述故障集群的主节点;
- [0014] 所述认证节点向所述故障集群内的普通节点、新主节点和其他集群的主节点发送所述故障集群的主节点更新数据包;
- [0015] 所述故障集群的普通节点以及所述新主节点根据所述主节点更新数据包更新主节点信息。
- [0016] 本发明的目的之一是,提供了一种根据区块链多活高可用系统进行主从节点故障恢复的方法,所述方法包括:
- [0017] 一认证节点确定出一故障集群的主节点A处于失效状态时,删除所述认证节点中所述集群的主节点对应的信息
- [0018] 所述故障集群的普通节点确定出所述主节点和从节点处于失效状态时,删除所述主节点对应的信息以及所述从节点对应的信息;
- [0019] 当所述普通节点确定出主节点、从节点不存在时,根据邻居数据单元中的邻居节点信息过主从选举规则进行选举,确定出新主节点以及新从节点;
- [0020] 所述新主节点向所述认证节点发送所述故障集群的主节点更新数据包;
- [0021] 所述认证节点根据所述主节点更新数据包将所述新主节点对应的信息添加至主节点数据单元中,标识为所述故障集群的主节点;
- [0022] 所述认证节点向所述故障集群内的普通节点、新主节点和其他集群的主节点发送所述故障集群的主节点更新数据包;
- [0023] 所述故障集群的普通节点以及所述新主节点根据所述主节点更新数据包更新主节点信息。
- [0024] 本发明的目的之一是,提供了一种根据区块链多活高可用系统进行主节点共识错误恢复的方法,所述方法包括:
- [0025] 一认证节点统计一故障集群的主节点的集群错误数量;
- [0026] 当所述集群错误数量达到一阈值时,所述认证节点向所述故障集群内的主节点、从节点以及普通节点发送主节点强制切换数据包;
- [0027] 所述故障集群根据所述主节点强制切换数据包调整所述主节点为不可选举状态;
- [0028] 故障机群的从节点以及普通节点确定出邻居数据单元中主节点不存在、从节点存在,将所述从节点选举为所述故障集群的新主节点;
- [0029] 所述新主节点向所述认证节点发送所述故障集群的主节点更新数据包;
- [0030] 所述认证节点根据所述主节点更新数据包将所述新主节点对应的信息添加至主节点数据单元中,标识为所述故障集群的主节点;
- [0031] 所述认证节点向所述故障集群内的普通节点、新主节点和其他集群的主节点发送所述故障集群的主节点更新数据包;

[0032] 所述故障集群的普通节点以及所述新主节点根据所述主节点更新数据包更新主节点信息。

[0033] 本发明的目的之一是,提供了一种计算机设备,包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序,所述处理器执行所述计算机程序时实现上述方法。

[0034] 本发明的目的之一是,提供了一种计算机可读存储介质,所述计算机可读存储介质存储有执行上述方法。

[0035] 本发明的有益效果在于,提供了一种区块链多活高可用系统、方法、计算机设备以及计算机可读存储介质,通过采取集群节点代替原来区块链节点,正常情况下由集群中的主节点参与区块链工作,在主节点出现异常时由其他集群积极实时代替,大大提高了整个区块链的可靠性,满足银行7*24小时的高可用性要求。

[0036] 为了让本发明的上述和其他目的、特征和优点能更明显易懂,下文特举较佳实施例,并配合所附图式,作详细说明如下。

附图说明

[0037] 为了更清楚地说明本发明实施例或现有技术中的技术方案,下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本发明的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

[0038] 图1为本发明实施例提供的一种区块链多活高可用系统的结构示意图;

[0039] 图2为本发明实施例提供的一种区块链多活高可用系统中认证节点的结构框图;

[0040] 图3为本发明实施例提供的一种区块链多活高可用系统中一集群的结构示意图;

[0041] 图4为本发明实施例提供的一种区块链多活高可用系统中的主节点的结构框图;

[0042] 图5为本发明实施例提供的一种区块链多活高可用系统中从节点的结构框图;

[0043] 图6为本发明实施例提供的一种区块链多活高可用系统中普通节点的结构框图;

[0044] 图7为本发明实施例提供的一种区块链多活高可用系统中主节点外控单元图;

[0045] 图8为本发明实施例提供的一种区块链多活高可用系统中主节点内控单元图;

[0046] 图9为本发明实施例提供的一种区块链多活高可用系统中从节点外控单元图;

[0047] 图10为本发明实施例提供的一种区块链多活高可用系统中从节点内控单元图;

[0048] 图11为本发明实施例提供的一种区块链多活高可用系统中普通节点外控单元图;

[0049] 图12为本发明实施例提供的一种区块链多活高可用系统中普通节点内控单元图;

[0050] 图13为本发明提供的具体实施例中主节点故障恢复流程图;

[0051] 图14为本发明提供的具体实施例中主从节点故障恢复流程图;

[0052] 图15为本发明提供的具体实施例中主节点共识错误恢复流程图;

[0053] 图16为本发明提供的具体实施例中正常时基于集群的两公司之间多活网络部署图;

[0054] 图17为本发明提供的具体实施例中故障时基于集群的两公司之间多活网络部署图;

[0055] 图18为本发明提供的具体实施例中正常时基于集群的四公司之间多活网络部署图;

图；

[0056] 图19为本发明提供的具体实施例中故障时基于集群的四公司之间多活网络部署图；

[0057] 图20为本发明提供的根据区块链多活高可用系统进行主节点故障恢复的方法流程图；

[0058] 图21为本发明提供的根据区块链多活高可用系统进行主从节点故障恢复的方法流程图；

[0059] 图22为本发明提供的根据区块链多活高可用系统进行主节点共识错误恢复的方法流程图。

具体实施方式

[0060] 下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0061] 本领域技术人员知道,本发明的实施方式可以实现为一种系统、装置、方法或计算机程序产品。因此,本发明公开可以具体实现为以下形式,即:完全的硬件、完全的软件(包括固件、驻留软件、微代码等),或者硬件和软件结合的形式。

[0062] 下面参考本发明的若干代表性实施方式,详细阐释本发明的原理和精神。

[0063] 针对现有技术中区块链存在容错功能有限和节点无法快速恢复的问题,无法满足关键业务特别是银行业务多活的网络系统要求,本发明实施例提供了一种区块链多活高可用系统,通过集群的方式实现区块链多活。区块链网络由若干个集群组成,而每个集群内部通过主从选举规则,选举出一个主节点、一个从节点和若干普通节点,主节点代表集群向认证节点进行注册、更新和查询主节点等功能,所有集群的主节点通过认证节点的服务组成一个P2P全互连的区块链网络,所有的区块链之间的通讯均由集群的主节点负责完成;当区块链达成共识之后,每个集群内部,主节点都会将区块链数据同步更新到其他非主节点中;当某集群主节点失效或者共识出错时,将通过主动或者被动的主从更新方式进行集群内部的主从选举,由于从节点的存在,能够实现主节点的快速切换,并通过认证节点通知其他集群快速更新该集群主节点信息,对外代表集群提供区块链服务,从而实现区块链故障节点的快速更换,实现区块链网络的多活。

[0064] 图1为该系统的结构示意图,请参阅图1,所述系统包括一认证节点1以及多个区块链节点,每个所述区块链节点为一集群2,每个所述集群内部通过主从选举规则从多个普通节点中选举出一主节点以及一从节点,所述集群中的主节点参与区块链,

[0065] 其中,区块链网络由多个集群组成,区块链之间的通讯由所述主节点完成,当区块链达成共识之后,每个集群内部所述主节点将区块链数据同步更新至所述从节点以及所述普通节点中,当所述集群中的一故障集群的主节点失效或共识出错时,通过主从选举规则进行主从选举,将所述从节点切换为主节点提供区块链服务,所述认证节点通知所述故障集群之外的其他集群更新所述故障集群的主节点信息。

[0066] 图1是区块链多活示意图,为联盟链或私有链,包含认证节点、多个区块链节点,区

区块链节点是一个集群,参与区块链的是所述集群中的主节点,在所述联盟链或私有链中,所述主节点就是普通的区块链节点,参与联盟链或私有链的处理工作。但是,这些主节点又属于各自集群中的主节点。如果当前主节点工作异常时,由所属集群中的其它节点成为新的主节点,取代原来主节点参与区块链处理的工作。

[0067] 图2为本发明实施例提供的一种区块链多活高可用系统中认证节点的结构框图,请参阅图2,认证节点1为区块链各集群节点提供身份验证和主节点数据维护等功能,包括:身份认证单元11、主节点数据单元12、主节点外控单元13、主节点保活单元14,主节点更新单元15、主节点切换单元16和主节点查询单元17。

[0068] 其中,所述身份认证单元11,用于对节点进行身份验证,当所述节点通过身份验证后,将所述节点加入集群。

[0069] 在具体的实施方式中,所述身份认证单元11负责区块链节点加入集群时的身份验证,只有通过认证节点身份验证后的节点才能加入集群,并返回验证结果给区块链节点。

[0070] 所述主节点数据单元12,用于维护和存储所述集群中的主节点信息,所述主节点信息包括集群ID、主节点IP地址和主节点老化计时器。

[0071] 其中,集群ID表示主节点所属的集群唯一标识;主节点IP为集群ID中当前的主节点IP地址;主节点老化计时器为该集群ID中主节点老化计时器,表示认证节点多久未收到该集群ID主节点发送的保活数据包,当计时器达到最大老化时间时,该集群主节点信息将从主节点数据单元12中删除。

[0072] 所述主节点认证单元13,用于当通过身份验证且所述集群的主节点未曾注册时,添加主节点信息至所述主节点数据单元,并返回注册结果给所述集群的主节点,注册成功时,所述注册结果包含集群ID信息,所述集群ID信息为所述认证节点通过认证节点公钥加密得到;

[0073] 所述主节点保活单元14,用于接收所述主节点发送的主节点保活数据包,根据所述主节点老化计时器判断所述主节点是否活动;

[0074] 所述主节点更新单元15,用于接收所述主节点发送的主节点更新数据包,只有通过认证节点身份校验,并且从主节点更新数据包中用认证节点自身私钥解密后的集群ID和声明要加入的集群ID一致时,认证节点才会更新集群主节点信息到主节点数据单元12。

[0075] 所述主节点切换单元16,用于接收所述主节点发送的故障数据包,所述故障数据包满足预设条件时,向所述主节点所在的所述集群发送强制切换主节点通知。在本发明的一种实施方式中,诸如认证节点统计一故障集群的主节点的集群错误数量,当所述集群错误数量达到一阈值。

[0076] 所述主节点查询单元17,用于接收所述主节点发送的主节点查询数据包,根据所述主节点查询数据包查询所述主节点数据单元,并返回查询结果给所述主节点。

[0077] 图3为本发明实施例提供的一种区块链多活高可用系统中一集群的结构示意图,请参阅图3,集群N2对内进行集群内部的主从节点选举,对外通过选举出的主节点进行认证节点和集群之间的通讯。每个集群内部包括N个集群节点,其中 $N \geq 2$;集群内部的节点角色由一个主节点N21、一个从节点n22和若干普通节点nn23组成。其中主节点N21和从节点n22必须存在,并且有且只有一个,普通节点nn23为可选。节点角色通过集群主从选举规则选举得出。

[0078] 所述主节点N21,主要负责集群之间P2P全互联通信,代表集群N提供区块链节点服务,如参与区块链各主节点之间共识过程。非主节点(非主节点,包括从节点和普通节点,下同)不参与集群之间的通讯。

[0079] 所述从节点n22,作为主节点的备用节点,当发现集群内的主节点N21失效时,从节点n22可以快速切换为主节点N21。

[0080] 所述普通节点nn23,当Master节点N21或者从节点n22失效时,普通节点nn23将在集群内部进行主从节点的选举,并有机会成为主节点N21或者从节点n22。

[0081] 图4为本发明实施例提供的一种区块链多活高可用系统中的主节点的结构框图,如图4所示,主节点21主要负责集群之间P2P全互联通信,代表集群提供区块链节点服务,包括主节点外控单元211、主节点内控单元212和主节点区块链管理单元213。

[0082] 其中,所述主节点外控单元211,用于与所述认证节点进行数据交互,包括主节点的注册、更新、保活、查询等功能。

[0083] 所述主节点内控单元212,用于负责所述集群中主节点、从节点以及普通节点之间的数据交互,包括集群内部邻居节点数据存储、主从选举、节点保活、数据同步等功能。

[0084] 所述主节点区块链管理单元213,用于代表所述集群参与区块链通讯,当区块链达成共识后,将区块链数据同步更新至所述从节点以及所述普通节。

[0085] 图5为本发明实施例提供的一种区块链多活高可用系统中从节点的结构框图,如图5所示,从节点22作为主节点的备用节点,当发现主节点出现故障时,可以立即转换为主节点,包括从节点外控单元221、从节点内控单元222和从节点区块链管理单元223。

[0086] 所述从节点外控单元221,负责与认证节点进行数据交互,包括从节点的注册、主节点切换等功能。

[0087] 所述从节点内控单元222,负责集群内部各节点之间的数据交互,包括集群内部邻居节点数据存储、主从选举、节点保活、数据同步等功能。

[0088] 所述从节点区块链管理单元223,负责从集群内的主节点同步更新区块链数据到本地,不参与其他区块链集群主节点的之间通讯。

[0089] 图6为本发明实施例提供的一种区块链多活高可用系统中普通节点的结构框图,如图6所示,当发现主节点或者从节点出现故障时,普通节点通过主从选举规则,可能会成为主节点或者从节点,包括普通节点外控单元231、普通节点内控单元232和普通节点区块链管理单元233。

[0090] 所述普通节点外控单元231,负责与认证节点进行数据交互,包括其他节点的注册、主节点切换等功能。

[0091] 所述普通节点内控单元232,负责集群内部各节点之间的数据交互,包括集群内部邻居节点数据存储、主从选举、节点保活、数据同步等功能。

[0092] 所述普通节点区块链管理单元233,负责从集群内的主节点同步更新区块链数据到本地,不参与其他区块链集群主节点的之间通讯。

[0093] 图7为本发明实施例提供的一种区块链多活高可用系统中主节点外控单元图,如图7所示,主节点外控单元211负责与认证节点进行数据交互,包括主节点注册单元2111、更新单元2112、保活单元2113、查询单元2114、注册数据单元2115、集群身份数据单元2116、集群身份管理单元2117、共识出错通知单元2118和切换单元2119。

[0094] 其中,所述主节点注册单元2111,用于当通过主从节点选举规则选举为集群初始主节点后,代表所述集群向所述认证节点发起主节点注册数据包请求,注册成功后,获取所述认证节点通过认证节点公钥加密后的集群ID信息,并将所述集群ID信息存储在注册数据单元2115中;

[0095] 所述更新单元2112,用于当通过主从节点选举规则选举为集群更新主节点后,代表所述集群向所述认证节点发起主节点更新数据包请求,主节点更新数据包请求中包含所述认证节点通过公钥加密后的集群ID信息,所述集群ID信息从所述注册数据单元2115中获得;

[0096] 所述保活单元2113,用于当所述主节点注册单元2111注册成功或所述更新单元2112更新成功后,代表所述集群通过保活定时器定时向所述认证节点发起主节点保活数据包,定时刷新所述集群的主节点的老化计时器,以保持认证节点信息可用;

[0097] 所述查询单元2114,用于当所述主节点需要通过其他集群的主节点信息加入区块链网络时,向所述认证节点发送主节点查询数据包,从所述认证节点中查询相应集群的主节点的信息;

[0098] 所述注册数据单元2115,用于存储所述主节点注册单元2111向所述认证节点发起注册请求后的认证节点返回信息,并同时存储保活定时器,当所述保活定时器到期时重置保活定时器,并向所述认证节点发送主节点保活数据包的请求;

[0099] 所述集群身份数据单元2116,用于存储所述集群身份管理单元2117通过所述认证节点发起加入集群后认证节点返回的信息;

[0100] 所述集群身份管理单元2117,用于向所述认证节点发起加入集群的认证请求,并将认证节点返回的信息存储于所述集群身份数据单元2116;

[0101] 所述共识出错通知单元2118,用于当区块链共识发现一集群的主节点计算出错时,称发生出错的集群为故障集群,向认证节点发送所述故障集群的主节点出错的通知;

[0102] 所述切换单元2119,用于接收所述认证节点发出的强制主节点切换信息,并通过所述主节点内控单元2122进行集群内部的主从选举。

[0103] 图8为本发明实施例提供的一种区块链多活高可用系统中主节点内控单元图,如图8所示,主节点内控单元212负责集群内部各节点之间的数据交互,包含邻居数据单元2121、主从选举单元2122、节点保活单元2123、注册数据同步2124和区块链数据同步单元2125。

[0104] 其中,所述邻居数据单元2121,用于存储所述集群内部所有节点邻居的信息,所述信息包含集群ID、节点优先级、节点IP地址、节点角色以及节点老化计时器。集群ID表示节点所属的集群唯一标识;节点优先级为邻居节点参与集群内部主从选举时的优先级;节点IP地址为该邻居节点的IP地址;节点角色为该邻居节点在当前集群的所属角色;节点老化计时器,表示多久未收到该邻居节点发送的保活数据包。除老化时间外,集群内部的所有邻居数据单元2121信息均一致。

[0105] 所述主从选举单元2122,用于负责所述集群的集群主从选举运算,由于除老化时间外,集群内部的所有邻居数据单元2121信息均一致,因此集群内部所有邻居节点在完成主从选举后,选举的主从结果也将一致,选举完成后更新所述邻居数据单元中的节点角色信息;

[0106] 所述节点保活单元2123,用于维护所述集群的保活定时器,当所述保活定时器到期时,重置所述保活定时器,并向所述集群内的从节点以及普通节点定时发送节点广播保活数据包,同时也接收从其他节点定时发送而来的节点广播保活数据包,其中节点保活数据包中包含节点的集群ID,优先级、IP地址等信息,从而可以添加、更新或删除邻居数据单元2121的节点信息。

[0107] 所述注册数据同步单元2124,用于当所述主节点向所述认证节点注册成功或者收到所述集群内所述从节点或普通节点的注册数据更新请求后,同步更新注册数据到集群内的从节点或普通节点;

[0108] 所述区块链数据同步单元2125,用于当达成共识或收到所述集群内从节点或普通节点的区块链数据更新请求后,负责同步更新区块链数据到集群内的从节点或普通节点中。

[0109] 图9为本发明实施例提供的一种区块链多活高可用系统中从节点外控单元图,如图9所示,从节点外控单元221负责与认证节点进行数据交互,包括主节点切换单元2211,注册数据单元2212,集群身份数据2213和集群身份管理2214。

[0110] 所述主节点切换单元2211,负责接受从认证节点的主节点切换单元16发送集群被动进行主从选举的数据包,收到数据包后将开始集群内部的主从选举。

[0111] 所述注册数据单元2212,负责存储从集群内主节点同步更新的注册数据单元2115的数据。

[0112] 所述集群身份数据2213,负责存储集群身份管理2214通过认证节点发起加入集群后的认证节点返回信息。

[0113] 所述集群身份管理2214,负责向认证节点发起加入集群的认证请求,并将认证节点返回信息存储于集群身份数据2213。

[0114] 图10为本发明实施例提供的一种区块链多活高可用系统中从节点内控单元图,如图10所示,从节点内控单元222负责集群内部各节点之间的数据交互,包含邻居数据单元2221、主从选举单元2222、节点保活单元2223、注册数据同步2224和区块链数据同步2225。

[0115] 所述邻居数据单元2221,与邻居数据单元2121相同。

[0116] 所述主从选举单元2222,与主从选举单元2122相同。

[0117] 所述节点保活单元2223,与节点保活单元2123相同。

[0118] 所述注册数据同步2224,负责接收从集群内主节点同步更新注册数据或者负责向集群内的主节点发起注册数据更新请求。

[0119] 所述的区块链数据同步2225,负责接收从集群内主节点同步更新区块链数据或者负责向集群内的主节点发起区块链数据更新请求。

[0120] 图11为本发明实施例提供的一种区块链多活高可用系统中普通节点外控单元图,如图11所示,普通节点外控单元231,负责与认证节点进行数据交互,包括主节点切换单元2311,注册数据单元2312,集群身份数据2313和集群身份管理2314。

[0121] 所述的主节点切换单元2311,与主节点切换单元2211相同。

[0122] 所述的注册数据单元2312,与注册数据单元2212相同。

[0123] 所述的集群身份数据2313,与集群身份数据2213相同。

[0124] 所述的集群身份管理2314,与集群身份管理2214相同。

[0125] 图12为本发明实施例提供的一种区块链多活高可用系统中普通节点内控单元图,如图12所示,普通节点内控单元232,负责集群内部各节点之间的数据交互,包含邻居数据单元2231、主从选举单元2322、节点保活单元2323、注册数据同步2324和区块链数据同步2325。

[0126] 所述的邻居数据单元2321,与邻居数据单元2121相同。

[0127] 所述的主从选举单元2322,与主从选举单元2122相同。

[0128] 所述的节点保活单元2323,与节点保活单元2123相同。

[0129] 所述的注册数据同步2324,与注册数据同步单元2124相同。

[0130] 所述的区块链数据同步2325,与区块链数据同步单元2125相同。

[0131] 如上即是本发明提供的一种区块链多活高可用系统,通过采取集群节点代替原来区块链节点,正常情况下由集群中的主节点参与区块链工作,在主节点出现异常时由其他集群积极实时代替,大大提高了整个区块链的可靠性,满足银行7*24小时的高可用性要求。

[0132] 此外,尽管在上文详细描述中提及了系统的若干单元模块,但是这种划分仅仅并非强制性的。实际上,根据本发明的实施方式,上文描述的两个或更多单元的特征和功能可以在一个单元中具体化。同样,上文描述的一个单元的特征和功能也可以进一步划分为由多个单元来具体化。以上所使用的术语“模块”和“单元”,可以是实现预定功能的软件和/或硬件。尽管以下实施例所描述的模块较佳地以软件来实现,但是硬件,或者软件和硬件的组合的实现也是可能并被构想的。

[0133] 在介绍了本发明示例性实施方式的区块链多活高可用系统之后,接下来,参考附图对本发明示例性实施方式的方法进行介绍。该方法的实施可以参见上述整体的实施,重复之处不再赘述。

[0134] 图20为本发明实施例提供的一种根据区块链多活高可用系统进行主节点故障恢复的方法的流程图,请参阅图20,该方法包括:

[0135] S11:一认证节点确定出一故障集群的主节点处于失效状态时,删除所述认证节点中所述故障集群的主节点对应的信息。

[0136] 在本发明的一种实施方式中,当故障集群的主节点发生故障,无法对外提供任何服务,无法向外发送保活数据包时,认证节点发现主节点在主节点数据单元12的老化计数器已到期,认为该主节点已失效。

[0137] S12:所述故障集群内的从节点以及普通节点确定出所述主节点处于失效状态时,删除所述主节点对应的信息;

[0138] 故障集群内的从节点、普通节点发现主节点在主节点数据单元的老化计数器已到期,认为主节点已失效。

[0139] S13:当所述从节点、普通节点确定出所述主节点中存在所述故障集群的从节点时,将所述从节点选举为所述集群的新主节点。

[0140] S14:所述新主节点向所述认证节点发送所述故障集群的主节点更新数据包;

[0141] 在本发明的其他实施方式中,当故障集群内的从节点、普通节点发现邻居数据单元中不存在故障集群的从节点时,根据邻居数据单元中各邻居节点信息(除新主节点外),通过主从选举单元进行选举,确定出新从节点。

[0142] S15:所述认证节点根据所述主节点更新数据包将所述新主节点对应的信息添加

至主节点数据单元中,标识为所述故障集群的主节点。

[0143] 在本发明的其他实施方式中,所述认证节点检查所述主节点更新数据包,判断所述新主节点的身份是否合法;当判断为是时,执行S15;否则,所述认证节点拒绝更新所述主节点数据单元。

[0144] S16:所述认证节点向所述故障集群内的普通节点、新主节点和其他集群的主节点发送所述故障集群的主节点更新数据包;

[0145] S17:所述故障集群的普通节点以及所述新主节点根据所述主节点更新数据包更新主节点信息。

[0146] 下面结合具体的实施例,详细介绍本发明提供的主节点故障恢复方法,用以解决主节点出现故障时区块链系统自行恢复的方法。如图13所示,在该具体实施例中该方法包括:

[0147] S1301:集群A的主节点A分别通过主节点保活单元2113向认证节点发送保活数据包和通过节点保活单元2123向集群内普通节点发送保活数据包。

[0148] S1302:认证节点通过主节点保活单元14收到集群A主节点A发送的保活数据包,并且该主节点A在主节点数据单元12的老化计数器未到期,则认为该主节点A状态是正常,同时重置该主节点A在主节点数据单元12的老化计数器;集群A内的从节点a、节点a1和节点a2通过节点保活单元收到主节点A的保活数据包,并且该主节点A在邻居数据单元的老化计数器未到期,则认为该主节点A状态是正常,同时重置该主节点A在邻居数据单元的老化计数器。

[0149] S1303:集群A的主节点A发生故障,无法对外提供任何服务,无法向外发送保活数据包。

[0150] S1304:认证节点发现主节点A在主节点数据单元12的老化计数器已到期,认为节点A已失效;集群A内的节点a,节点a1和节点a2发现主节点A在主节点数据单元的老化计数器已到期,认为节点A已失效。

[0151] S1305:认证节点发现主节点A已失效,将主节点数据单元12中删除集群A主节点A的信息;集群A内的节点a,节点a1和节点a2发现主节点A已失效,将主节点数据单元中删除主节点A的信息。

[0152] S1306:集群A内的节点a,节点a1和节点a2发现邻居数据单元中存在从节点a,则立即将该从节点a选举为集群A新的主节点。

[0153] S1307:集群A内的节点a发现自己成为集群A新的主节点,立即向认证节点发送集群A主节点更新的数据包通知。

[0154] S1308:集群A内的节点a,节点a1和节点a2发现邻居数据单元中不存在集群A的从节点,根据邻居数据单元中各邻居节点信息(除新主节点a外),通过主从选举单元进行选举,选择结果:节点a2为新的从节点。

[0155] S1309:认证节点通过主节点更新单元16收到集群A新主节点a发送的主节点更新数据包,通过身份认证单元11和主节点数据单元12检查数据包的信息是否符合新主节点a身份是否合法。如果不合法将拒绝更新。

[0156] S1310:认证节点检查主节点数据单元11中的集群A主节点信息,发现不存在集群A主节点信息。

[0157] S1311:认证节点将集群A新主节点a信息添加到主节点数据单元11中,标识为集群A的主节点。

[0158] S1312:认证节点通过主节点更新单元16分别向集群A内所有节点和其他集群N主节点N(包括所有其他集群的主节点)发送集群A主节点更新的通知。

[0159] S1313:集群A所有节点(节点a,节点a1和节点a2)收到认证节点发送的主节点更新数据包后,立即根据该数据包内容比较当前邻居数据单元中主节点信息,如果不相同则更新邻居数据单元中主节点信息;如果相同则忽略该数据包。

[0160] 图21为本发明实施例提供的一种根据区块链多活高可用系统进行主从节点故障恢复的方法的流程图,请参阅图21,该方法包括:

[0161] S21:一认证节点确定出一故障集群的主节点处于失效状态时,删除所述认证节点中所述故障集群的主节点对应的信息。

[0162] 在本发明的一种实施方式中,当故障集群的主节点、从节点发生故障,无法对外提供任何服务,无法向外发送保活数据包时,认证节点发现主节点在主节点数据单元12的老化计数器已到期,认为该主节点已失效。

[0163] S22:所述故障集群的普通节点确定出所述主节点和从节点处于失效状态时,删除所述主节点对应的信息以及所述从节点对应的信息;

[0164] 故障集群内的普通节点发现主节点和从节点在其他节点数据单元的老化计数器已到期,认为主节点和从节点已失效。

[0165] S23:当所述普通节点确定出主节点、从节点不存在时,根据邻居数据单元中的邻居节点信息过主从选举规则进行选举,确定出新主节点以及新从节点。

[0166] S24:所述新主节点向所述认证节点发送所述故障集群的主节点更新数据包;

[0167] S25:所述认证节点根据所述主节点更新数据包将所述新主节点对应的信息添加至主节点数据单元中,标识为所述故障集群的主节点。

[0168] 在本发明的其他实施方式中,所述认证节点检查所述主节点更新数据包,判断所述新主节点的身份是否合法;当判断为是时,执行S25;否则,所述认证节点拒绝更新所述主节点数据单元。

[0169] S26:所述认证节点向所述故障集群内的普通节点、新主节点和其他集群的主节点发送所述故障集群的主节点更新数据包;

[0170] S27:所述故障集群的普通节点以及所述新主节点根据所述主节点更新数据包更新主节点信息。

[0171] 下面结合具体的实施例,详细介绍本发明提供的主从节点故障恢复方法,如图14所示,主从节点故障恢复流程,实现当集群主从节点同时发生故障时,通过集群内主从选举规则,选举出新的主节点和从节点,并通过新的主节点对外提供区块链服务,详细流程如下:

[0172] S1401:集群A的主节点A分别通过主节点保活单元2113向认证节点发送保活数据包和通过节点保活单元2123向集群内其他节点发送保活数据包。

[0173] S1402:认证节点通过主节点保活单元14收到集群A主节点A发送的保活数据包,并且该主节点A在主节点数据单元12的老化计数器未到期,则认为该主节点A状态是正常,同时重置该主节点A在主节点数据单元12的老化计数器;集群A内的从节点a、节点a1和节点a2

通过节点保活单元收到主节点A的保活数据包,并且该主节点A在邻居数据单元的老化计数器未到期,则认为该主节点A状态是正常,同时重置该主节点A在邻居数据单元的老化计数器。

[0174] S1403:集群A的从节点a分别通过节点保活单元2223向集群内其他节点发送保活数据包。

[0175] S1404:集群A内的主节点A、节点a1和节点a2通过节点保活单元收到从节点a的保活数据包,并且该主节点A在邻居数据单元的老化计数器未到期,则认为该从节点a状态是正常,同时重置该从节点a在邻居数据单元的老化计数器。

[0176] S1405:集群A的主节点A和从节点a发生故障,无法对外提供任何服务,无法向外发送保活数据包。

[0177] S1406:认证节点发现主节点A在主节点数据单元12的老化计数器已到期,认为节点A已失效;集群A内的节点a1和节点a2发现主节点A和从节点a在其他节点数据单元的老化计数器已到期,认为节点A和节点a已失效。

[0178] S1407:认证节点发现主节点A已失效,将主节点数据单元12中删除集群A主节点A的信息;集群A内的节点a1和节点a2发现主节点A和节点a已失效,将主节点数据单元中删除主节点A和节点a的信息。

[0179] S1408:集群A内的节点a1和节点a2发现主从节点不存在,立即根据邻居数据单元中的邻居节点信息(当前邻居只有节点a1和节点a2)进行主从选举,选举结果:节点a2为新的主节点,节点a1为新的从节点。

[0180] S1409:集群A内的节点a2发现自己成为集群A新的主节点,立即向认证节点发送集群A主节点更新的数据包通知。

[0181] S1410:认证节点通过主节点更新单元16收到集群A新主节点a2发送的主节点更新数据包,通过身份认证单元11和主节点数据单元12检查数据包的信息是否符合新主节点a2身份是否合法。如果不合法将拒绝更新。

[0182] S1411:认证节点检查主节点数据单元11中的集群A主节点信息,发现不存在集群A主节点信息。

[0183] S1412:认证节点将集群A新主节点a2信息添加到主节点数据单元11中,标识为集群A的主节点。

[0184] S1413:认证节点通过主节点更新单元16分别向集群A内所有节点和其他集群N主节点N(包括所有其他集群的主节点)发送集群A主节点更新的通知。

[0185] S1414:集群A所有节点(节点a1和节点a2)收到认证节点发送的主节点更新数据包后,立即根据该数据包内容比较当前邻居数据单元中主节点信息,如果不相同则更新邻居数据单元中主节点信息;如果相同则忽略该数据包。

[0186] 图22为本发明实施例提供的一种根据区块链多活高可用系统进行主节点共识错误恢复的方法的流程图,请参阅图22,该方法包括:

[0187] S31:一认证节点统计一故障集群的主节点的集群错误数量;

[0188] S32:当所述集群错误数量达到一阈值时,所述认证节点向所述故障集群内的主节点、从节点以及普通节点发送主节点强制切换数据包;

[0189] S33:所述故障集群根据所述主节点强制切换数据包调整所述主节点为不可选举

状态;

[0190] S34:故障机群的从节点以及普通节点确定出邻居数据单元中主节点不存在、从节点存在,将所述从节点选举为所述故障集群的新主节点;

[0191] S35:所述新主节点向所述认证节点发送所述故障集群的主节点更新数据包;

[0192] S36:所述认证节点根据所述主节点更新数据包将所述新主节点对应的信息添加至主节点数据单元中,标识为所述故障集群的主节点;

[0193] S37:所述认证节点向所述故障集群内的普通节点、新主节点和其他集群的主节点发送所述故障集群的主节点更新数据包;

[0194] S38:所述故障集群的普通节点以及所述新主节点根据所述主节点更新数据包更新主节点信息。

[0195] 下面结合具体的实施例,详细介绍本发明提供的主节点共识错误恢复方法,用以解决当主节点在区块链共识出现错误时区块链系统自行恢复的方法。如图15所示,该方法包括:

[0196] S1501:集群N主节点N在通过区块链管理单元213进行区块链共识时,发现只有集群A的主节点A和自己的计算结果不相同。

[0197] S1502:集群N主节点N通过共识错误通知单元2118向认证节点发送主节点共识错误数据包。

[0198] S1503:认证节点通过主节点切换单元16统计一段时间内关于集群A主节点A的所有集群错误数量。

[0199] S1504:认证节点通过主节点切换单元16检查发现一段时间内,集群A主节点A的所有集群错误数量已达到设置的阈值,将通过主节点切换单元16向该集群A内的所有节点(节点A,节点a,节点a1和节点a2)发送集群A主节点强制切换的数据包。

[0200] S1505:集群A内的所有节点(节点A,节点a,节点a1和节点a2)通过主节点切换单元接收到集群A主节点强制切换的数据包,在邻居数据单元中调整原主节点A为不可选举状态,节点角色调整为其他节点角色。

[0201] S1506:集群A内的所有节点(节点A,节点a,节点a1和节点a2)发现邻居数据单元中主节点已不存在,但存在从节点a,立即将该从节点a选举为集群A新的主节点。

[0202] S1507:集群A内的节点a发现自己成为集群A新的主节点,立即向认证节点发送集群A主节点更新的数据包通知。

[0203] S1508:集群A内的所有节点(节点A,节点a,节点a1和节点a2)发现邻居数据单元中不存在集群A的从节点,根据邻居数据单元中各邻居节点信息(包括节点a1和节点a2,排除新主节点a外和不可选举状态的节点A),通过主从选举单元进行选举,选举结果:节点a2为新的从节点。

[0204] S1509:认证节点通过主节点更新单元16收到集群A新主节点a发送的主节点更新数据包,通过身份认证单元11和主节点数据单元12检查数据包的信息是否符合新主节点a身份是否合法。如果不合法将拒绝更新。

[0205] S1510:认证节点检查主节点数据单元11中的集群A主节点信息,发现不存在集群A主节点信息。

[0206] S1511:认证节点将集群A新主节点a信息添加到主节点数据单元11中,标识为集群

A的主节点。

[0207] S1512:认证节点通过主节点更新单元16分别向集群A内所有节点和其他集群N主节点N(包括所有其他集群的主节点)发送集群A主节点更新的通知。

[0208] S1513:集群A所有节点(节点A,节点a,节点a1和节点a2)收到认证节点发送的主节点更新数据包后,立即根据该数据包内容比较当前邻居数据单元中主节点信息,如果不相同则更新邻居数据单元中主节点信息;如果相同则忽略该数据包。

[0209] 本发明还提供了一种计算机设备,包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序,所述处理器执行所述计算机程序时实现上述方法。

[0210] 本发明还提供了一种计算机可读存储介质,所述计算机可读存储介质存储有执行上述方法。

[0211] 本发明还提出一种基于所述区块链多活系统的主节点共识错误恢复方法,具体的:

[0212] 甲乙两公司通过基于PBFT区块链构建彼此之间的联盟链。如果使用普通区块链构建方案,在甲乙两公司的机房各自放置两个区块链节点,该方案存在任何一个机房出现故障时,整个区块链网络将无法提供服务。

[0213] 本发明还提出一种基于区块链多活系统的两公司物理区域部署区块链方法,用以解决当某一公司部署出现整个物理区域出现故障时区块链系统自行恢复的方法。该方法包括:

[0214] 如图16所示,正常时基于集群的两公司多活部署,通过如下方式部署实现:除在甲乙两公司的机房各自放置两个区块链集群主节点外,还将在对方机房放置两个集群的从节点,在正常时只有主节点之间参与共识,在区块链达成共识后,主节点将区块链数据同步更新到集群内部的从节点。当任何一个机房发生故障时,另外一个机房均可以立即接管,如图示17所示,故障时基于集群的两公司多活部署,具体实现:这里假设公司乙机房出现故障,所有节点无法提供区块链服务,则此时在公司甲机房内,可以快速通过集群内的主从选举,将集群B的原从节点b提升为主节点b,将集群D的原从节点d提升为主节点d,组成一个区块链数据和之前一致网络,提供对方服务。

[0215] 甲乙丙丁四公司通过基于PBFT区块链构建彼此之间的联盟链。如果使用普通区块链构建方案,在甲乙丙丁四公司的机房各自放置一个区块链节点,该方案存在任何两个或两个以上机房出现故障时,整个区块链网络将无法提供服务。

[0216] 本发明还提出一种所述区块链多活系统的四公司物理区域部署区块链方法,用以解决当某两公司部署出现整个物理区域出现故障时区块链系统自行恢复的方法。该方法包括:

[0217] 如图18所示,正常时基于集群的四公司多活部署,通过如下方式部署实现:除在甲乙丙丁四公司的机房各自放置一个区块链集群主节点外,还将在其他共识机房放置集群的从节点和other节点,即在甲机房放置集群A的主节点A、集群B的从节点b和集群C的other节点c1;在乙机房放置集群B的主节点B、集群D的从节点d和集群A的other节点a1;在丙机房放置集群C的主节点C、集群A的从节点a和集群D的other节点d1;在丁机房放置集群D的主节点D、集群C的从节点c和集群B的other节点b1。

[0218] 在正常时只有主节点之间参与共识,在区块链达成共识后,主节点将区块链数据

同步到集群内部的非主节点;当任何两个机房发生故障时,另外两个机房均可以立即接管,如图19所示,这里假设公司丙丁两个机房同时出现故障,两个机房内的所有节点无法提供区块链服务,则此时在公司甲,乙两个机房内,可以快速通过集群内的主从选举,在公司甲机房内将集群C的原other节点c1提升为主节点c1,在公司乙机房内将集群D的原从节点d提升为主节点d,组成一个区块链数据和之前一致网络,提供对外区块链服务。

[0219] 对于使用PBFT算法的区块链,此方案解决三个机房同时出现物理故障的问题:剩下一个机房仍然可以通过提升此机房内的其他两个节点为对应集群的主节点,从而组成一个三节点的区块链网络,仍然满足PBFT算法对外提供区块链服务。

[0220] 基于集群的区块链部署而且容易进行扩展,对于实施实例五,为保证任何三个机房同时宕机时,另外一个机房均存在四个节点同时运作,对外提供区块链服务,此扩展非常灵活,具体实现时:分别在甲乙丙丁机房分别增加集群A、集群B、集群C、集群D的一个other节点即可,加入过程对正在运行的区块链网络透明。

[0221] 综上所述,本发明提供了一种区块链多活高可用系统、方法、计算机设备以及计算机可读存储介质,通过采取集群节点代替原来区块链节点,正常情况下由集群中的主节点参与区块链工作,在主节点出现异常时由其他集群积极实时代替,大大提高了整个区块链的可靠性,满足银行7*24小时的高可用性要求。

[0222] 本发明提供的一种区块链多活的系统及方法,具有以下效果和优点:

[0223] 1、本发明提供了一种基于集群的区块链网络系统,在不改变共识算法的同时,可以给区块链网络提供更多节点的容错功能。

[0224] 2、本发明提供的节点自修复功能,提高了区块链网络的可用性。

[0225] 3、本发明提供的集群内部节点动态加入的方法,可方便对集群内的节点在逻辑和物理部署提供了进行动态扩展,实现区块链网络部署进行灵活部署,实现区块链节点多活。

[0226] 4、本发明提供的集群内部多节点功能,可方便对集群内部的节点数据备份,在保持数据一致性的同时,提高节点数据恢复的效率。

[0227] 对于一个技术的改进可以很明显地区分是硬件上的改进(例如,对二极管、晶体管、开关等电路结构的改进)还是软件上的改进(对于方法流程的改进)。然而,随着技术的发展,当今的很多方法流程的改进已经可以视为硬件电路结构的直接改进。设计人员几乎都通过将改进的方法流程编程到硬件电路中来得到相应的硬件电路结构。因此,不能说一个方法流程的改进就不能用硬件实体模块来实现。例如,可编程逻辑器件(Programmable Logic Device,PLD)(例如现场可编程门阵列(Field Programmable Gate Array,FPGA))就是这样一种集成电路,其逻辑功能由用户对器件编程来确定。由设计人员自行编程来把一个数字系统“集成”在一片PLD上,而不需要请芯片制造厂商来设计和制作专用的集成电路芯片。而且,如今,取代手工地制作集成电路芯片,这种编程也多半改用“逻辑编译器(logic compiler)”软件来实现,它与程序开发撰写时所用的软件编译器相类似,而要编译之前的原始代码也得用特定的编程语言来撰写,此称之为硬件描述语言(Hardware Description Language,HDL),而HDL也并非仅有一种,而是有许多种,如ABEL(Advanced Boolean Expression Language)、AHDL(Altera Hardware Description Language)、Confluence、CUPL(Cornell University Programming Language)、HDCal、JHDL(Java Hardware Description Language)、Lava、Lola、MyHDL、PALASM、RHDL(Ruby Hardware Description

Language)等,目前最普遍使用的是VHDL (Very-High-Speed Integrated Circuit Hardware Description Language)与Verilog2。本领域技术人员也应该清楚,只需要将方法流程用上述几种硬件描述语言稍作逻辑编程并编程到集成电路中,就可以很容易得到实现该逻辑方法流程的硬件电路。

[0228] 控制器可以按任何适当的方式实现,例如,控制器可以采取例如微处理器或处理器以及存储可由该(微)处理器执行的计算机可读程序代码(例如软件或固件)的计算机可读介质、逻辑门、开关、专用集成电路(Application Specific Integrated Circuit, ASIC)、可编程逻辑控制器和嵌入微控制器的形式,控制器的例子包括但不限于以下微控制器:ARC625D、Atmel AT91SAM、Microchip PIC18F26K20以及Silicone Labs C8051F320,存储器控制器还可以被实现为存储器的控制逻辑的一部分。

[0229] 本领域技术人员也知道,除了以纯计算机可读程序代码方式实现控制器以外,完全可以通过将方法步骤进行逻辑编程来使得控制器以逻辑门、开关、专用集成电路、可编程逻辑控制器和嵌入微控制器等的形式来实现相同功能。因此这种控制器可以被认为是一种硬件部件,而对其内包括的用于实现各种功能的装置也可以视为硬件部件内的结构。或者甚至,可以将用于实现各种功能的装置视为既可以是实现方法的软件模块又可以是硬件部件内的结构。

[0230] 上述实施例阐明的系统、装置、模块或单元,具体可以由计算机芯片或实体实现,或者由具有某种功能的产品来实现。

[0231] 为了描述的方便,描述以上装置时以功能分为各种单元分别描述。当然,在实施本申请时可以把各单元的功能在同一个或多个软件和/或硬件中实现。

[0232] 通过以上的实施方式的描述可知,本领域的技术人员可以清楚地了解到本申请可借助软件加必需的通用硬件平台的方式来实现。基于这样的理解,本申请的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来,该计算机软件产品可以存储在存储介质中,如ROM/RAM、磁碟、光盘等,包括若干指令用以使得一台计算机系统(可以是个人计算机,服务器,或者网络系统等)执行本申请各个实施例或者实施例的某些部分所述的方法。

[0233] 本说明书中的各个实施例均采用递进的方式描述,各个实施例之间相同相似的部分互相参见即可,每个实施例重点说明的都是与其他实施例的不同之处。尤其,对于系统实施例而言,由于其基本相似于方法实施例,所以描述的比较简单,相关之处参见方法实施例的部分说明即可。

[0234] 本申请可用于众多通用或专用的计算机系统环境或配置中。例如:个人计算机、服务器计算机、手持系统或便携式系统、平板型系统、多处理器系统、基于微处理器的系统、置顶盒、可编程的消费电子系统、网络PC、小型计算机、大型计算机、包括以上任何系统或系统的分布式计算环境等等。

[0235] 本申请可以在由计算机执行的计算机可执行指令的一般上下文中描述,例如程序模块。一般地,程序模块包括执行特定任务或实现特定抽象数据类型的例程、程序、对象、组件、数据结构等等。也可以在分布式计算环境中实践本申请,在这些分布式计算环境中,通过通信网络而被连接的远程处理系统来执行任务。在分布式计算环境中,程序模块可以位于包括存储系统在内的本地和远程计算机存储介质中。

[0236] 虽然通过实施例描绘了本申请,本领域普通技术人员知道,本申请有许多变形和变化而不脱离本申请的精神,希望所附的权利要求包括这些变形和变化而不脱离本申请的精神。

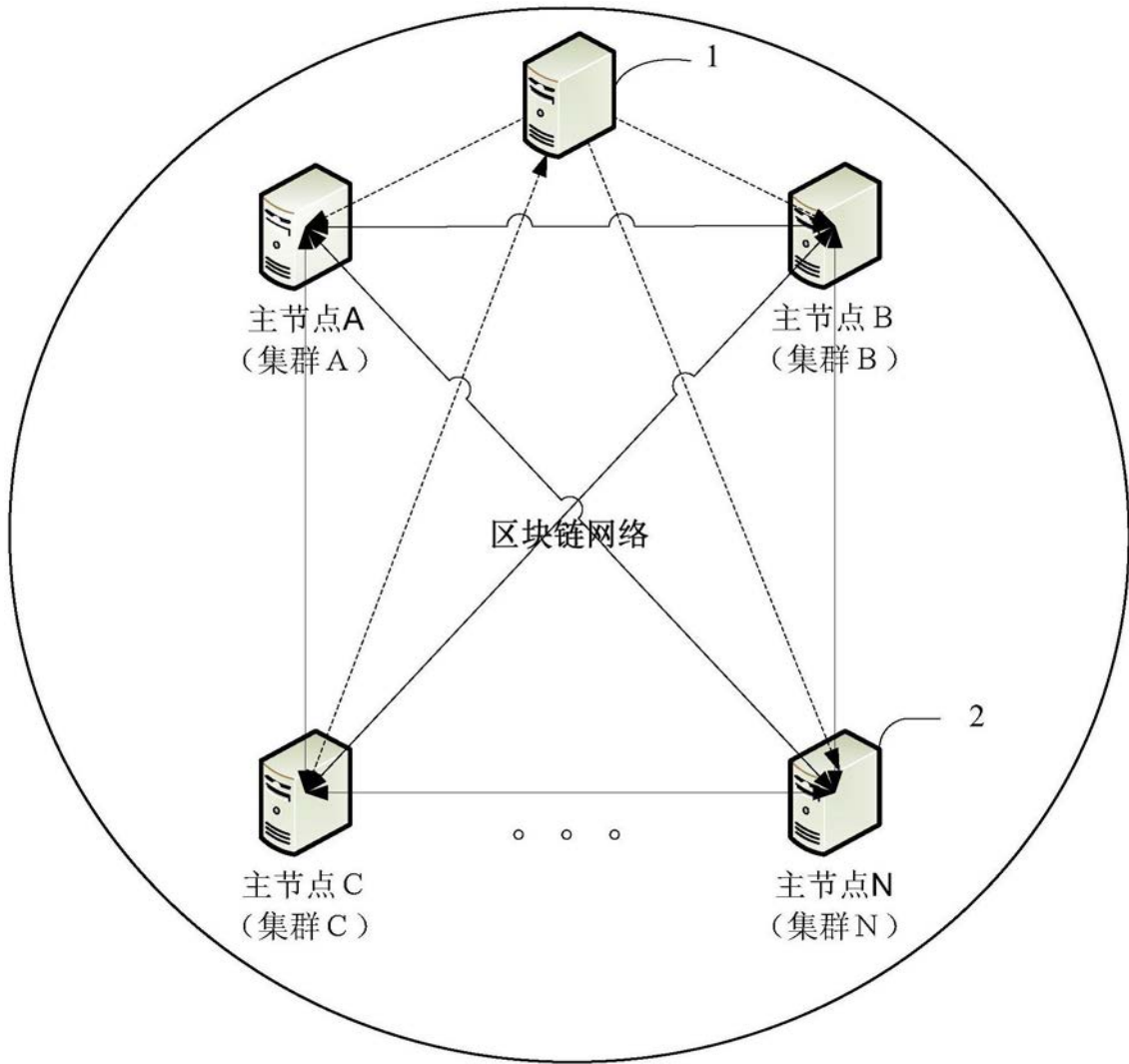


图1

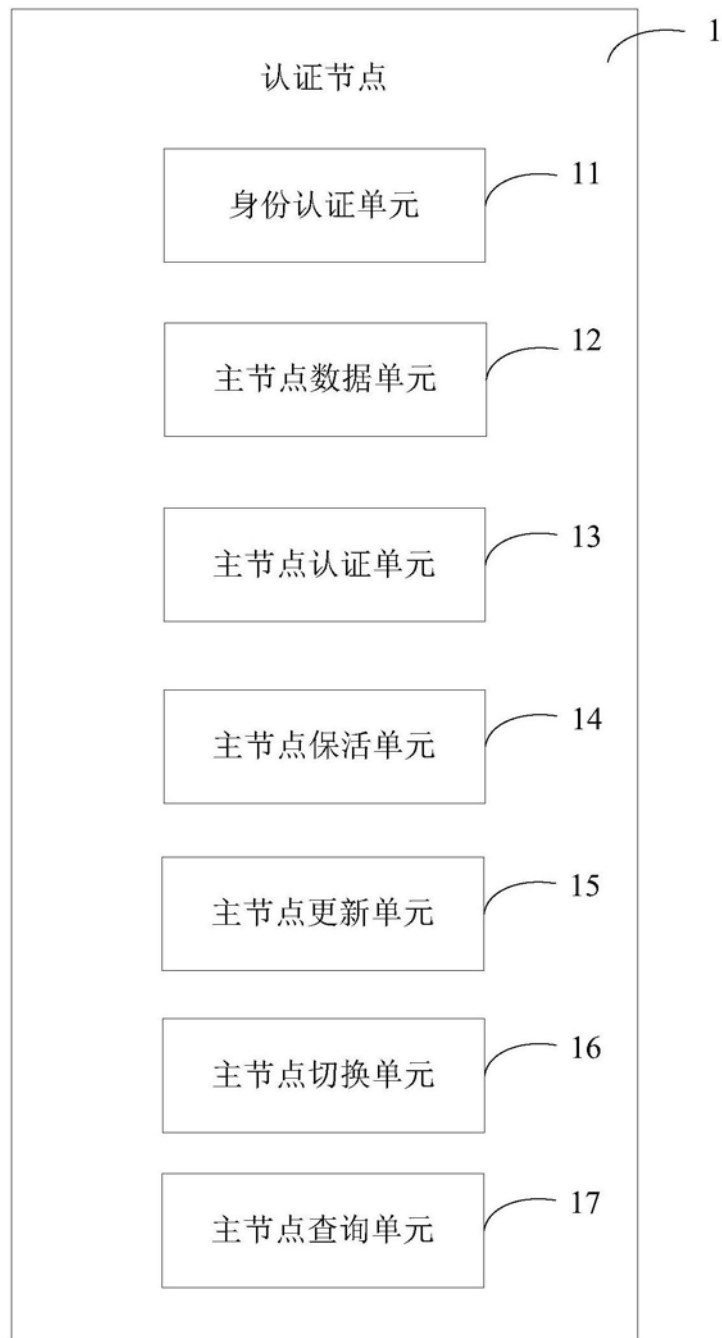


图2

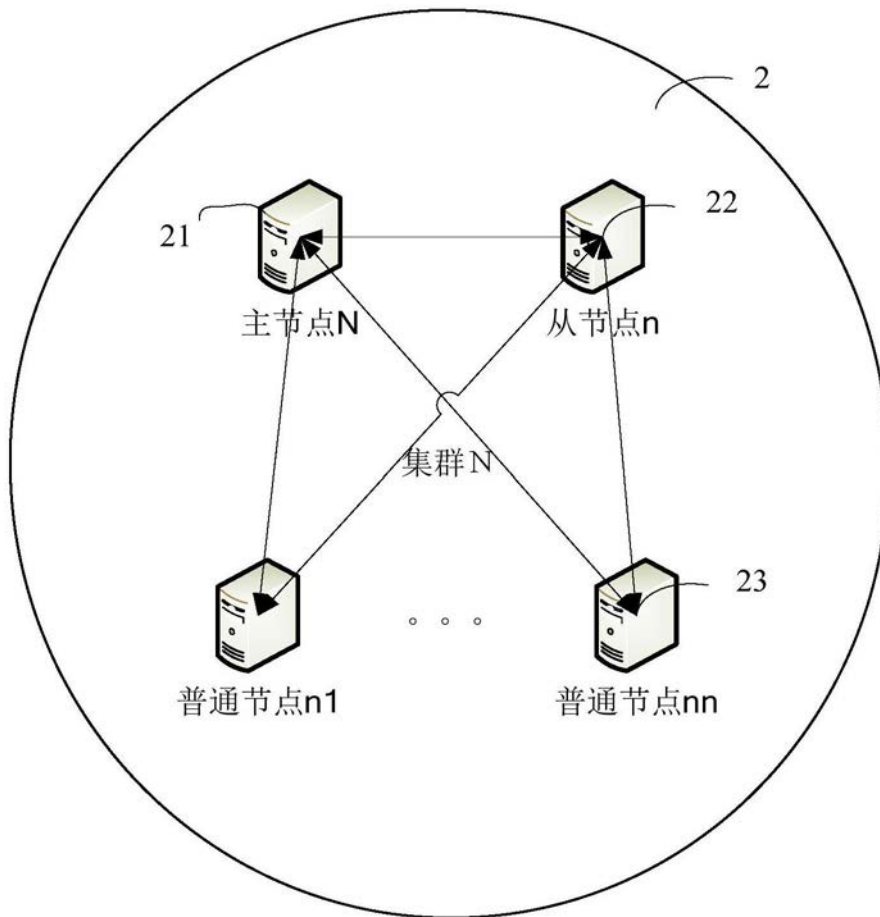


图3

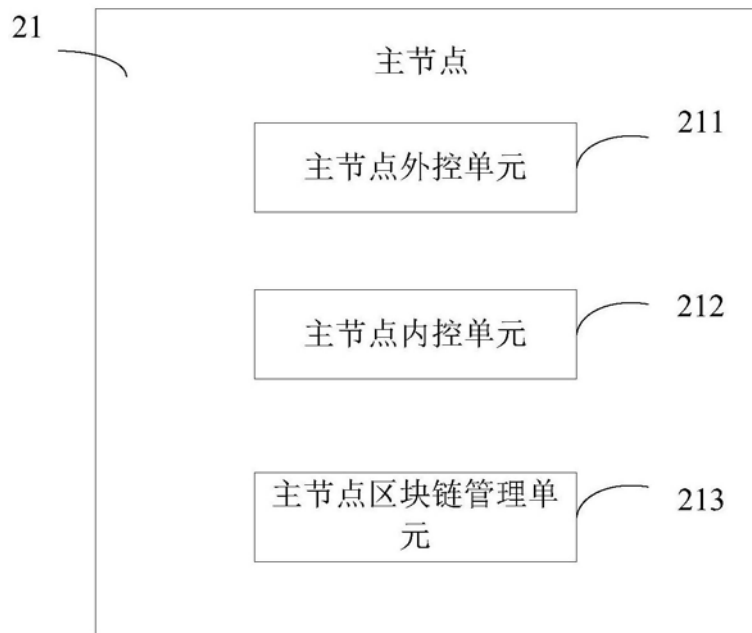


图4

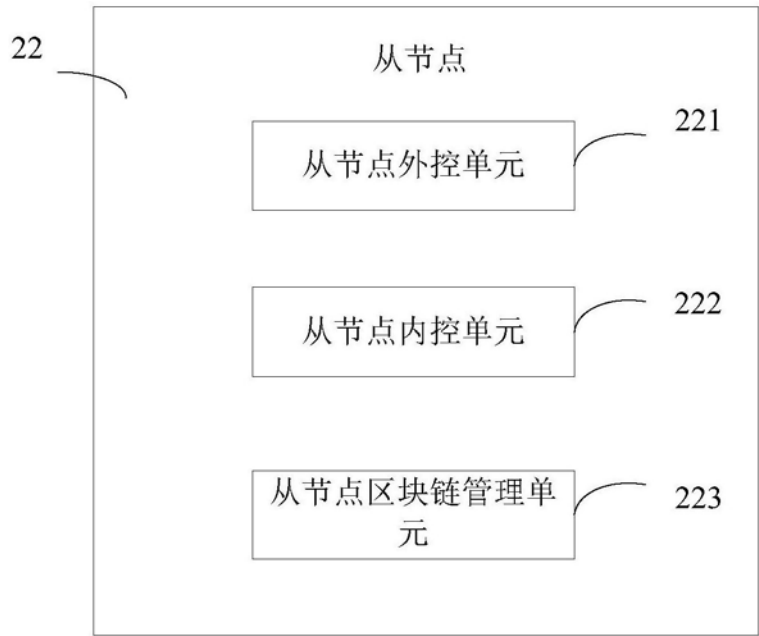


图5

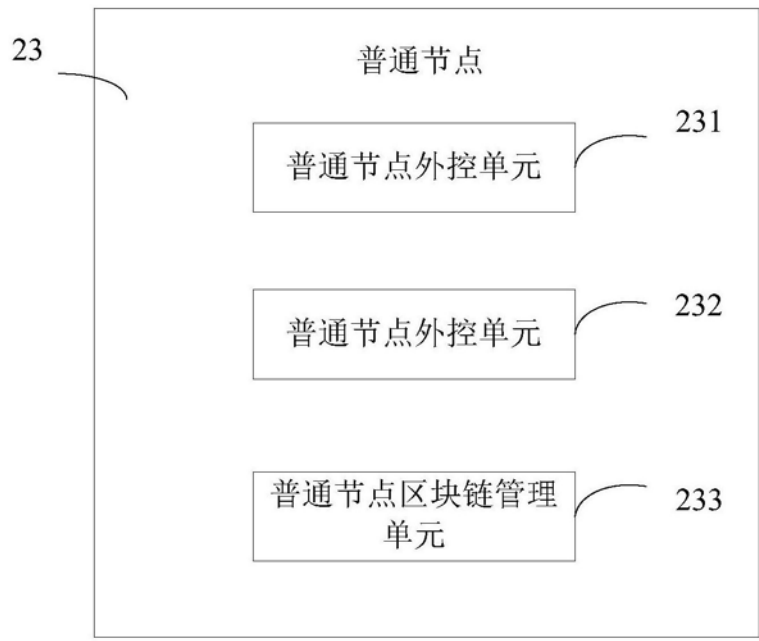


图6



图7

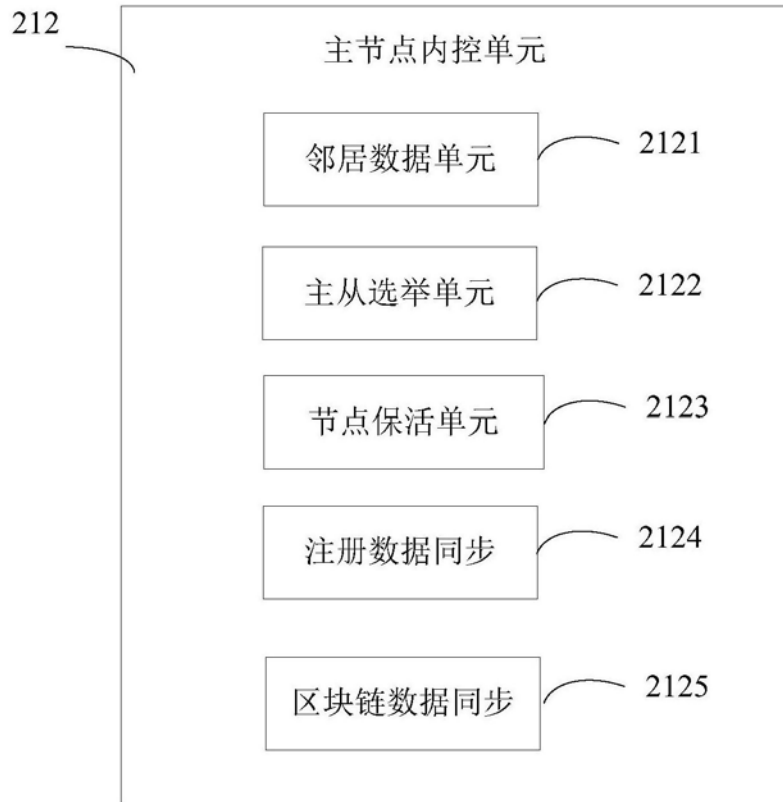


图8

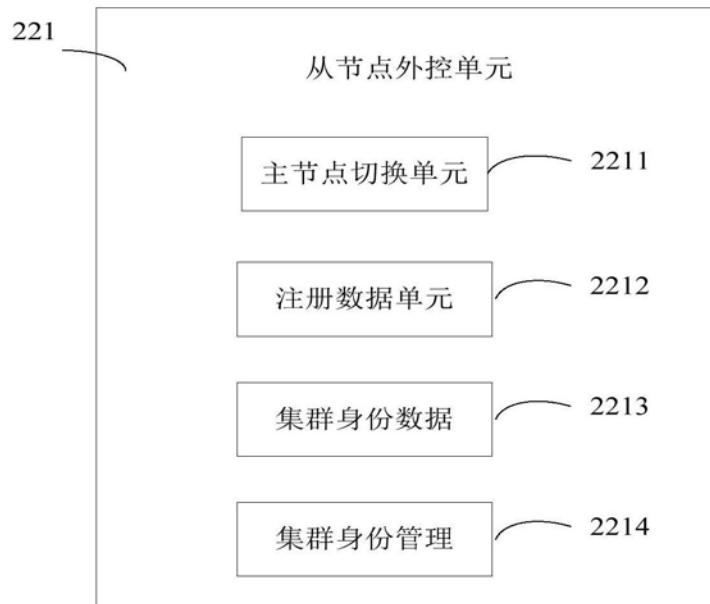


图9

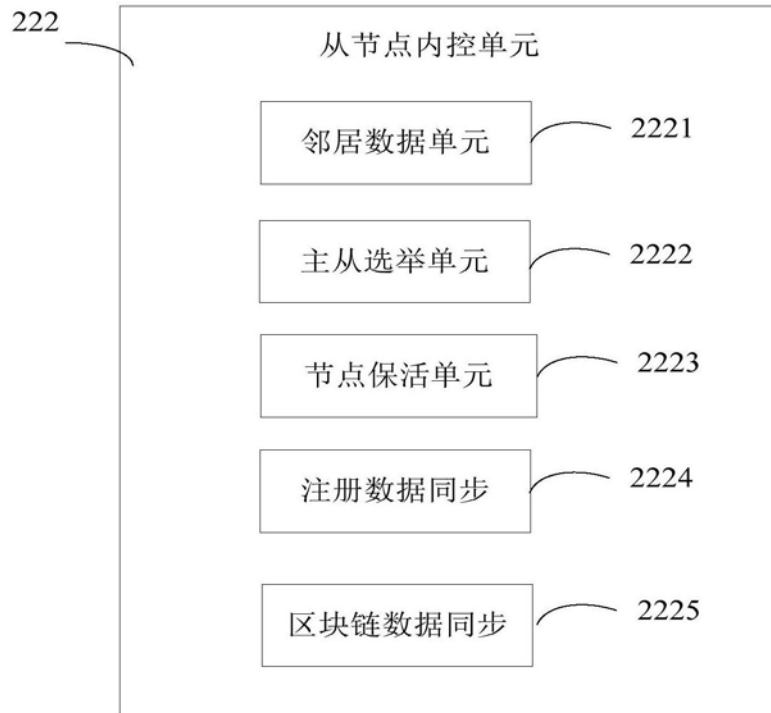


图10

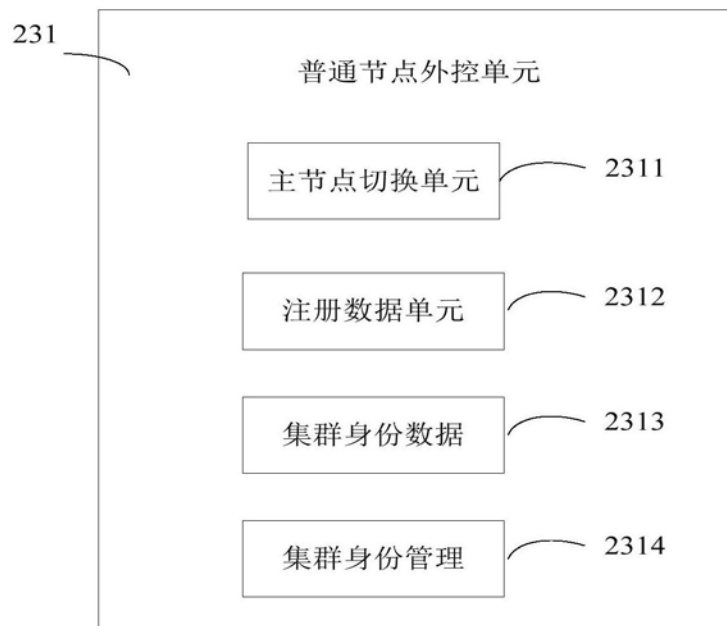


图11

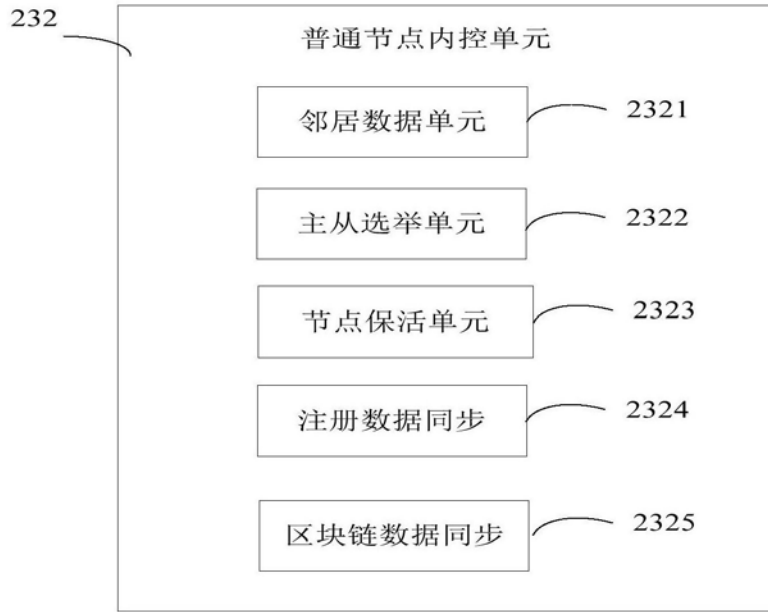


图12



图13

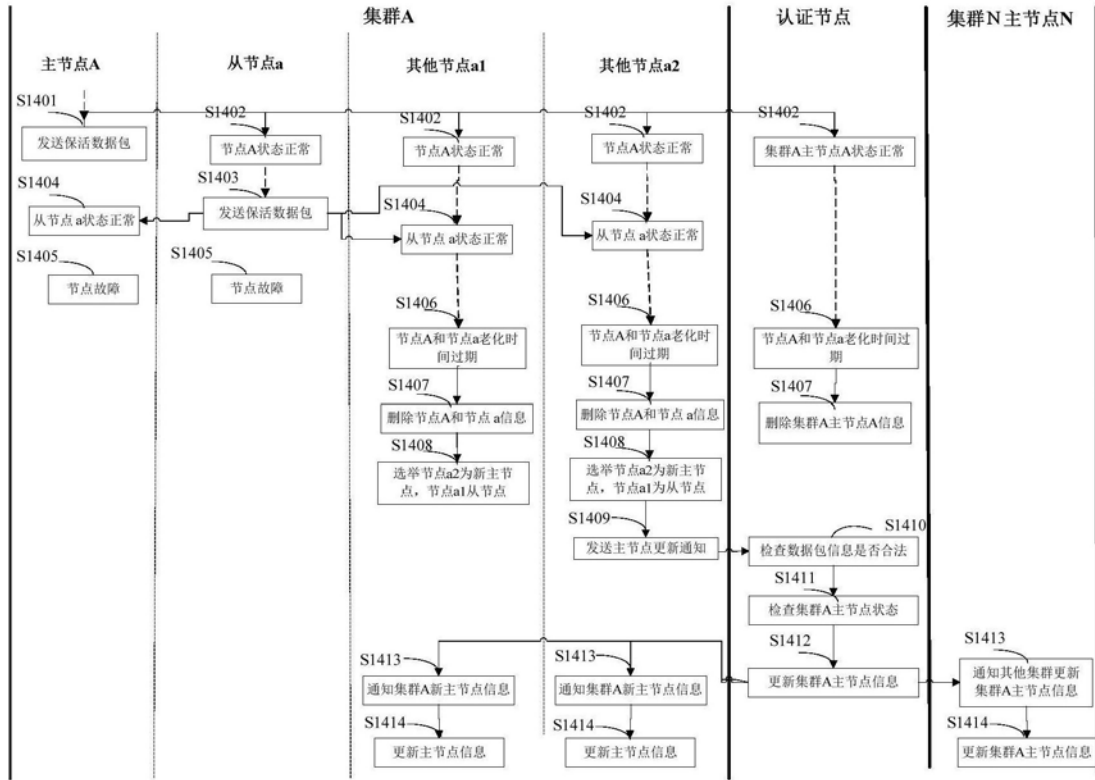


图14

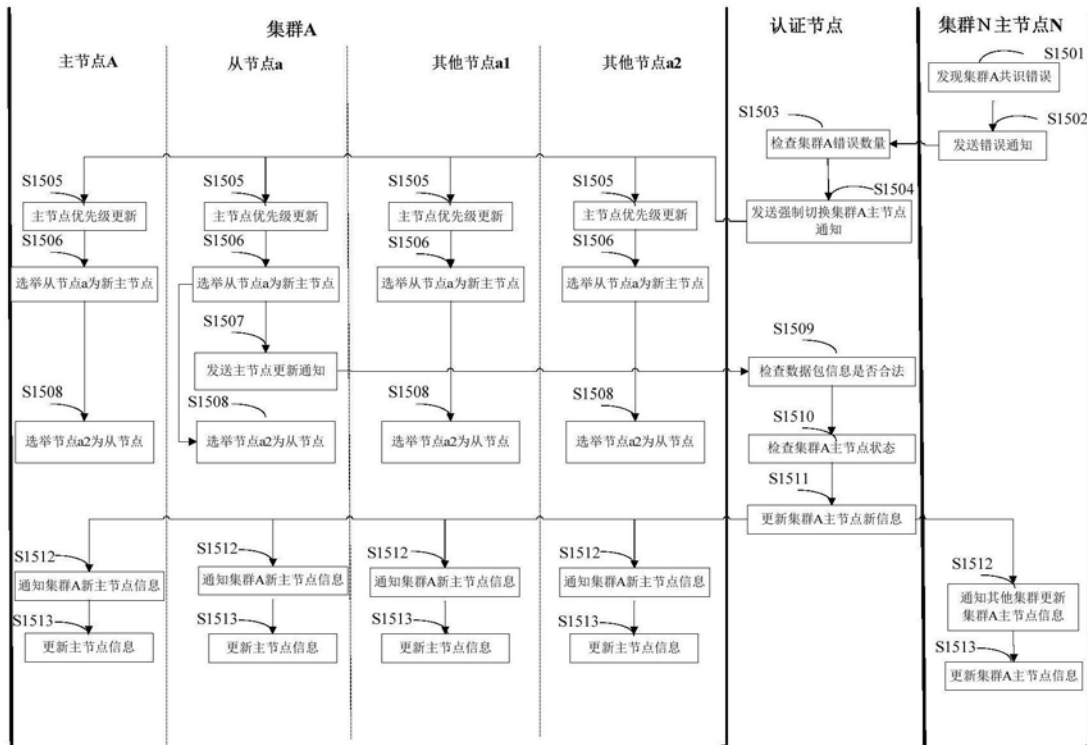


图15

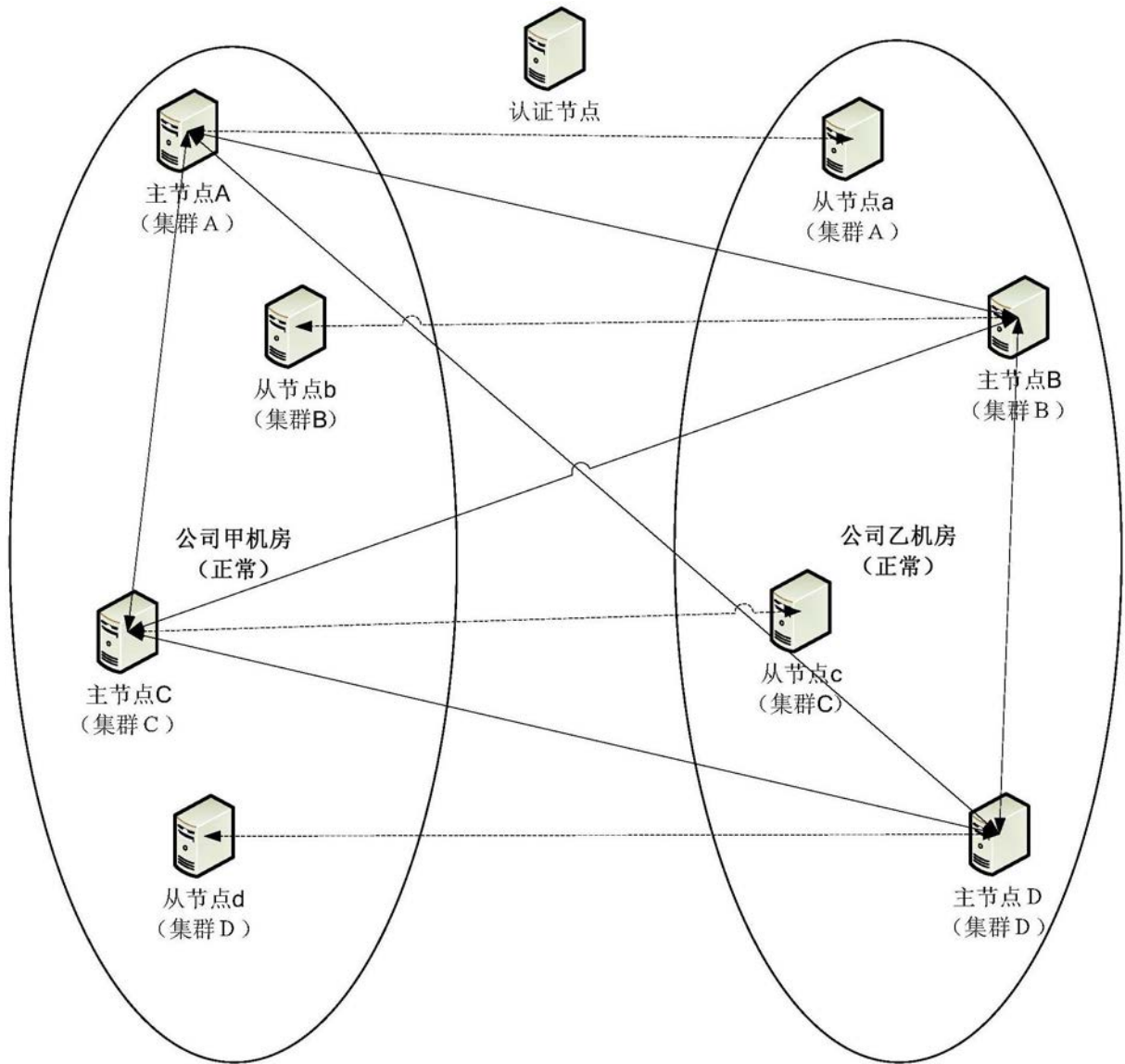


图16

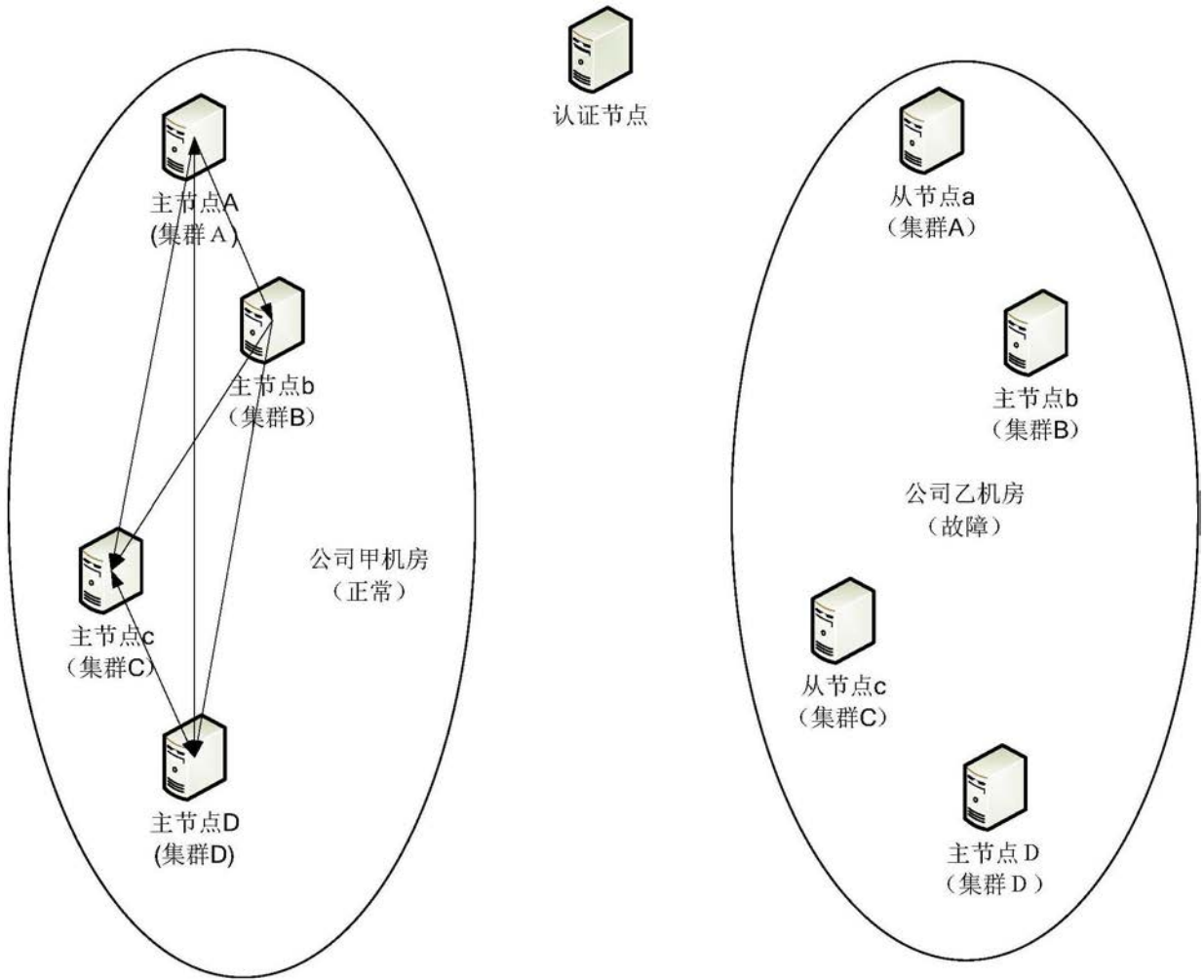


图17

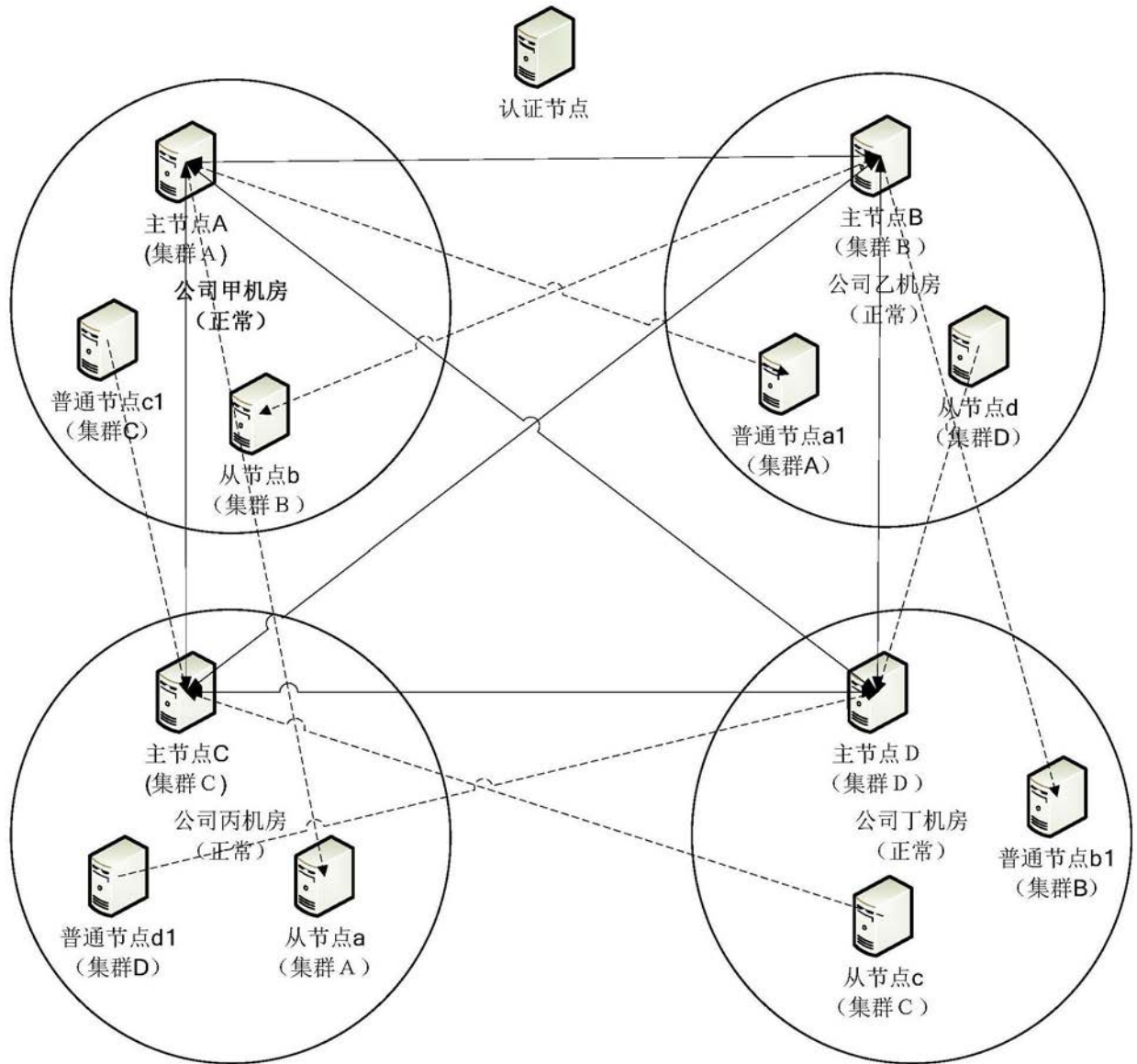


图18

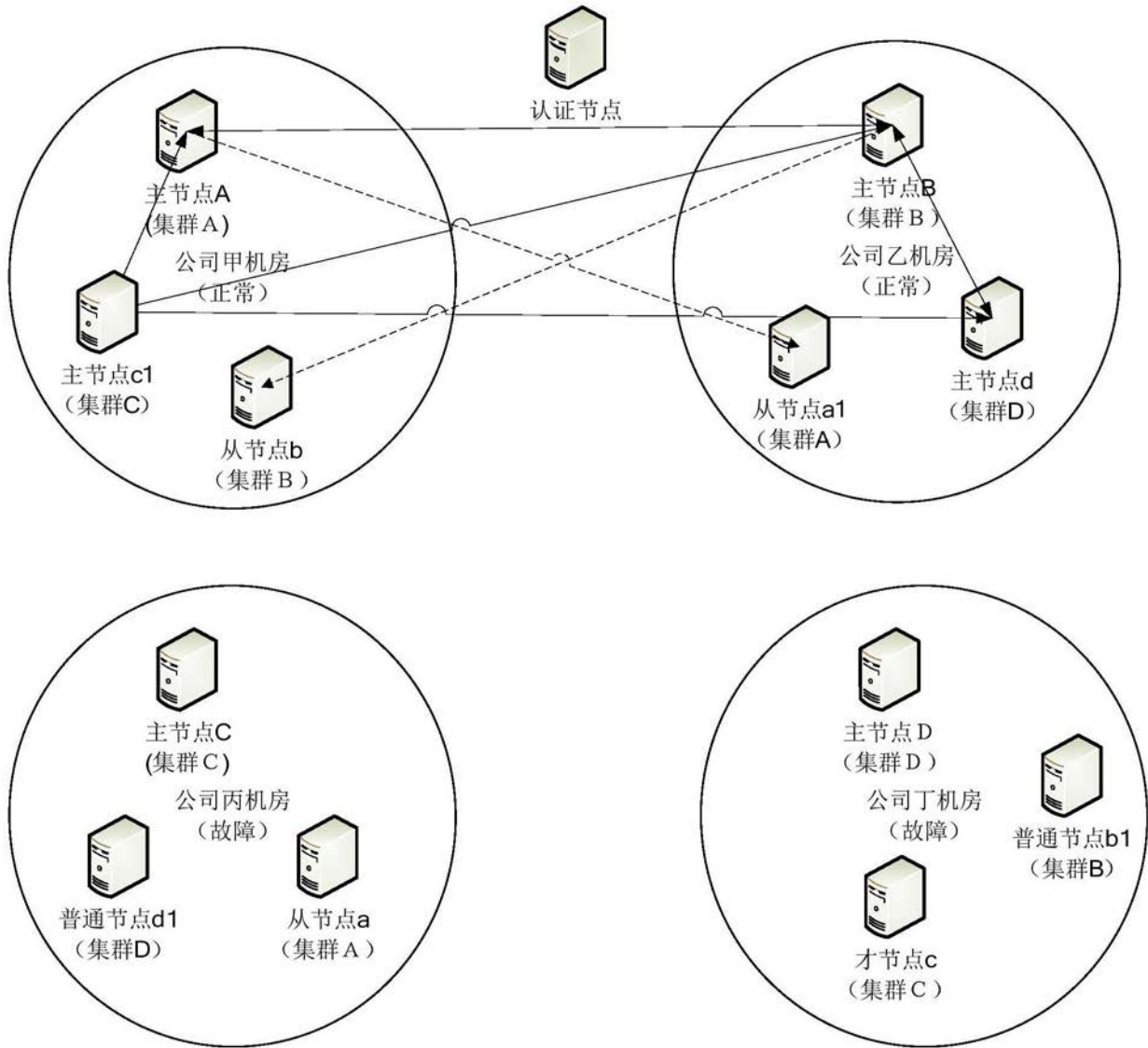


图19

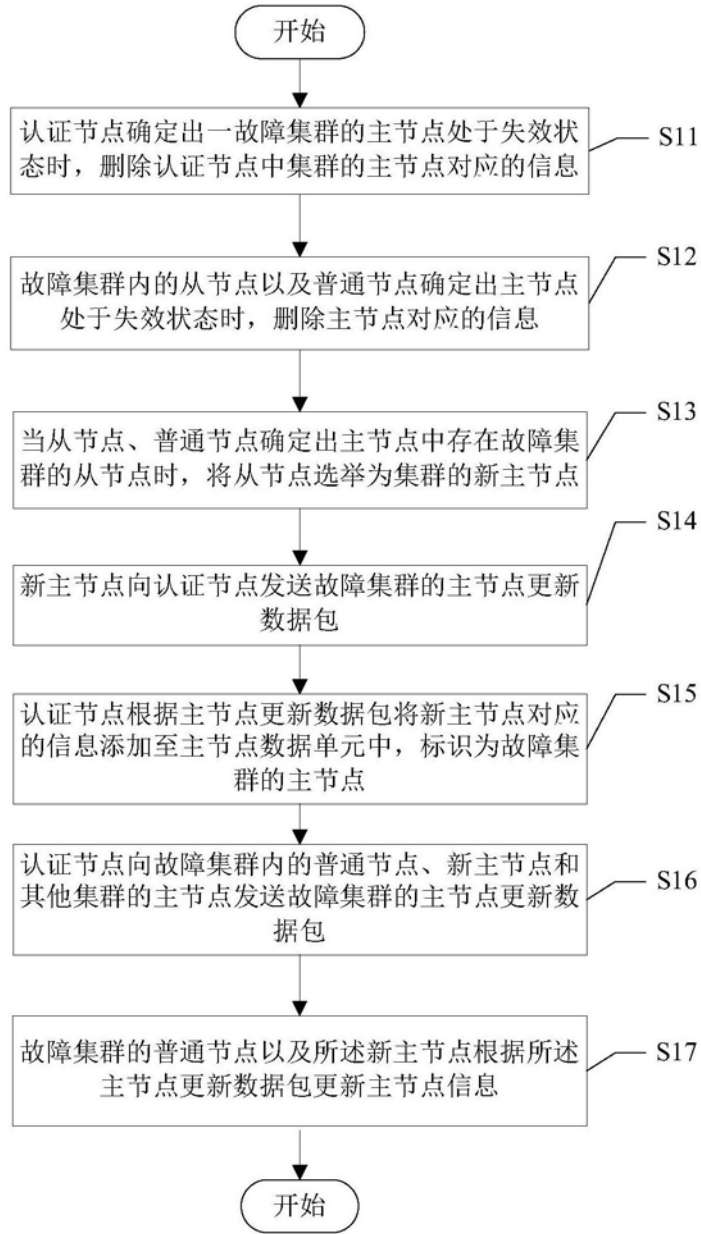


图20

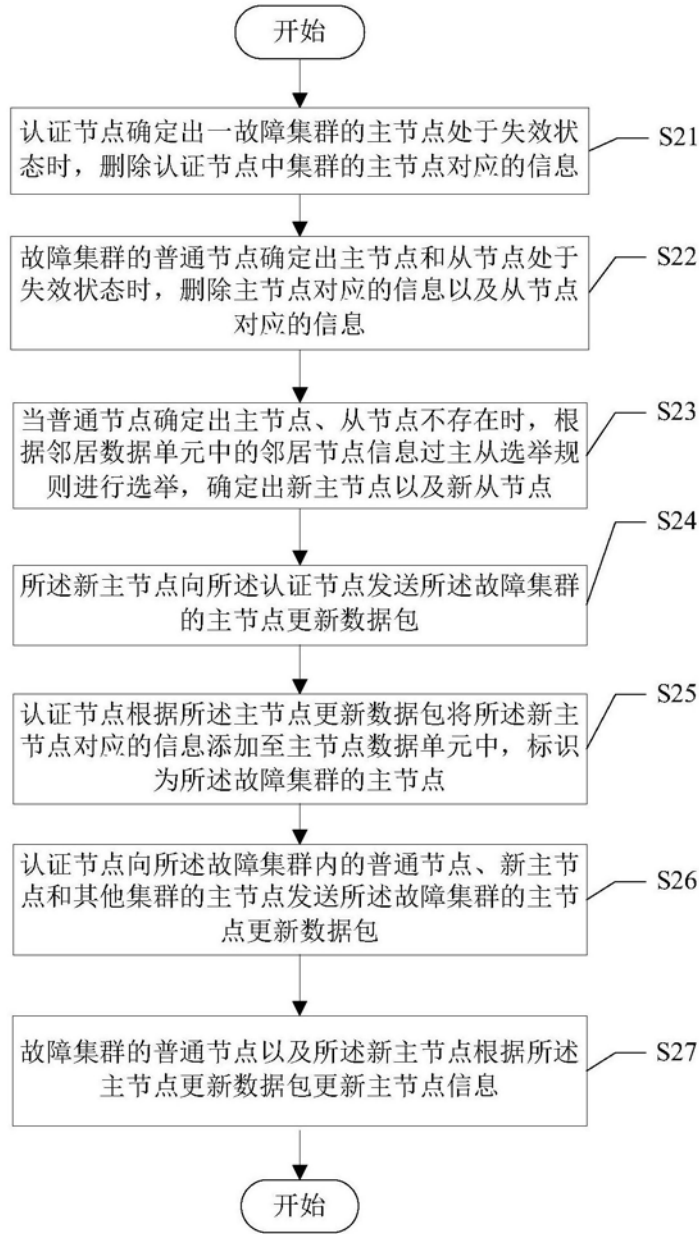


图21

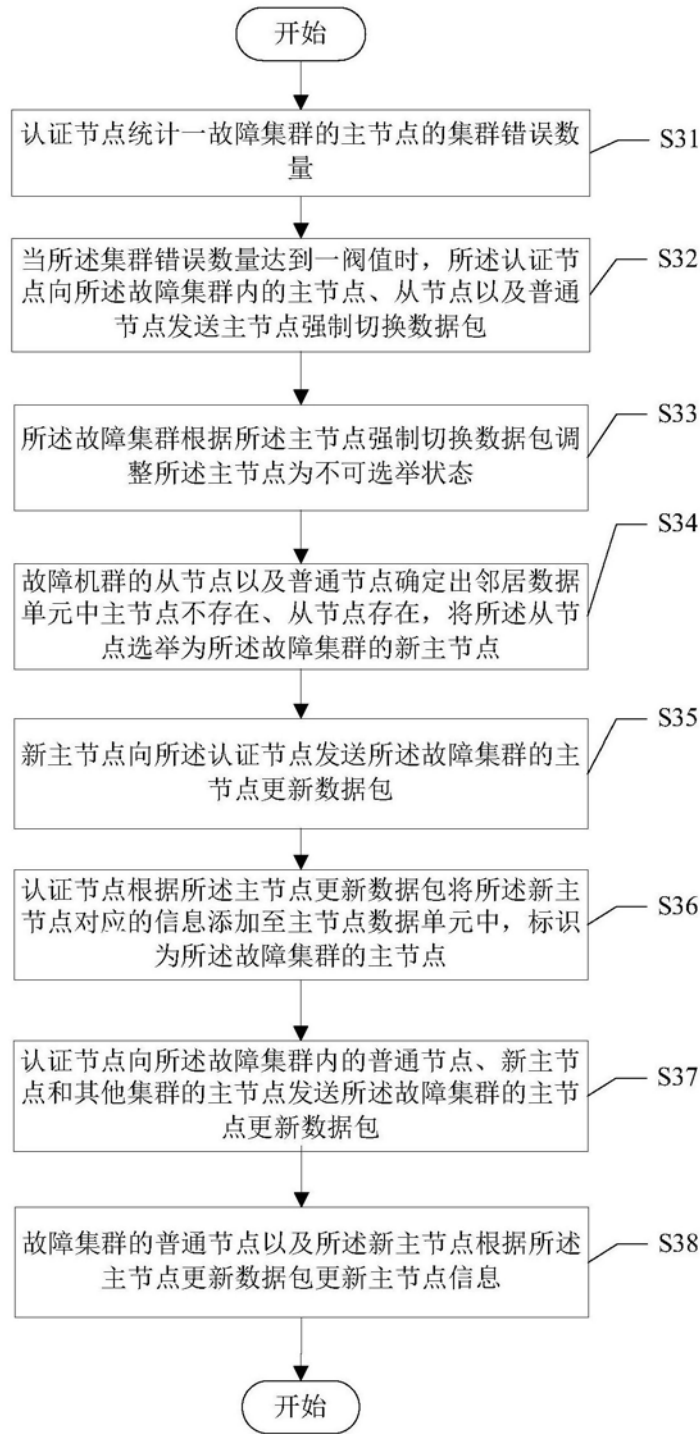


图22