



(12)发明专利申请

(10)申请公布号 CN 107124305 A

(43)申请公布日 2017.09.01

(21)申请号 201710263587.9

(22)申请日 2017.04.20

(71)申请人 腾讯科技(深圳)有限公司

地址 518057 广东省深圳市南山区高新区  
科技中一路腾讯大厦35层

(72)发明人 郭锐 李茂材 梁军 屠海涛  
赵琦 王宗友 张建俊 朱大卫  
刘斌华

(74)专利代理机构 北京三高永信知识产权代理  
有限责任公司 11138

代理人 朱雅男

(51)Int. Cl.

H04L 12/24(2006.01)

H04L 12/26(2006.01)

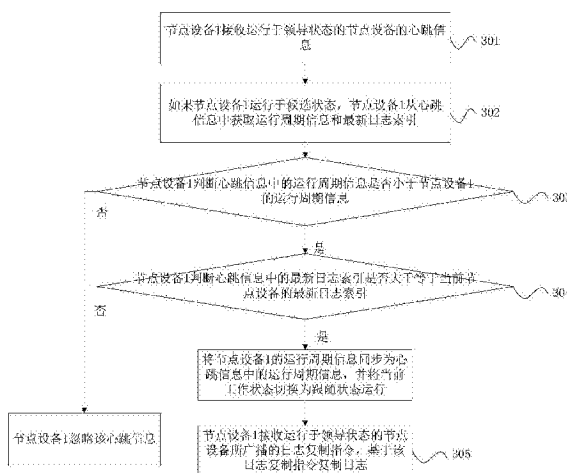
权利要求书2页 说明书7页 附图3页

(54)发明名称

节点设备运行方法及节点设备

(57)摘要

本发明公开了一种节点设备运行方法及节点设备,属于网络技术领域。该方法包括:接收运行于领导状态的节点设备的心跳信息;如果当前节点设备运行于候选状态,从心跳信息中获取运行周期信息和最新日志索引;如果心跳信息中的运行周期信息小于当前节点设备的运行周期信息,且心跳信息中的最新日志索引大于等于当前节点设备的最新日志索引,将当前节点设备的运行周期信息同步为心跳信息中的运行周期信息,并将当前工作状态切换为跟随状态运行或保持候选状态。本发明通过将当前节点设备的运行周期信息同步为心跳信息中的运行周期信息,解决了现有技术中系统中的子集群恢复网络连接时不能合为一个系统工作的问题,提高了系统的工作可靠性。



1. 一种节点设备运行方法,其特征在于,所述方法包括:

接收运行于领导状态的节点设备的心跳信息;

如果当前节点设备运行于候选状态,从所述心跳信息中获取运行周期信息和最新日志索引;

如果所述心跳信息中的运行周期信息小于所述当前节点设备的运行周期信息,且所述心跳信息中的最新日志索引大于等于所述当前节点设备的最新日志索引,将所述当前节点设备的运行周期信息同步为所述心跳信息中的运行周期信息,并将当前工作状态切换为跟随状态运行或保持所述候选状态。

2. 根据权利要求1所述的方法,其特征在于,所述如果当前节点设备运行于候选状态,从所述心跳信息中获取运行周期信息和最新日志索引,包括:

如果所述当前节点设备运行于候选状态,从所述心跳信息中获取多个签名;

如果所述多个签名的数量大于系统中节点设备数量的半数,且所述多个签名均验证通过,从所述心跳信息中获取运行周期信息和最新日志索引。

3. 根据权利要求1所述的方法,其特征在于,所述如果所述心跳信息中的运行周期信息小于所述当前节点设备的运行周期信息,且所述心跳信息中的最新日志索引大于等于所述当前节点设备的最新日志索引,将所述当前节点设备的运行周期信息同步为所述心跳信息中的运行周期信息,并将当前工作状态切换为跟随状态运行或保持所述候选状态之后,所述方法还包括:

接收所述运行于领导状态的节点设备所广播的日志复制指令,基于所述日志复制指令复制日志。

4. 根据权利要求1所述的方法,其特征在于,所述如果所述心跳信息中的运行周期信息小于所述当前节点设备的运行周期信息,且所述心跳信息中的最新日志索引大于等于所述当前节点设备的最新日志索引,将所述当前节点设备的运行周期信息同步为所述心跳信息中的运行周期信息,并将当前工作状态切换为跟随状态运行或保持所述候选状态之后,所述方法还包括:

基于所述当前节点设备的最新日志索引和所述心跳信息中的最新日志索引,确定所述当前节点设备需要添加的日志索引;

向所述运行于领导状态的节点设备发送日志添加请求。

5. 根据权利要求1-4中任一项所述的方法,其特征在于,所述心跳信息或所述日志复制指令携带系统中各个节点设备在响应所述运行于领导状态的节点设备的投票请求时的签名。

6. 一种节点设备,其特征在于,所述节点设备包括:

接收模块,用于接收运行于领导状态的节点设备的心跳信息;

获取模块,用于如果当前节点设备运行于候选状态,从所述心跳信息中获取运行周期信息和最新日志索引;

运行模块,用于如果所述心跳信息中的运行周期信息小于所述当前节点设备的运行周期信息,且所述心跳信息中的最新日志索引大于等于所述当前节点设备的最新日志索引,将所述当前节点设备的运行周期信息同步为所述心跳信息中的运行周期信息,并将当前工作状态切换为跟随状态运行或保持所述候选状态。

7. 根据权利要求6所述的节点设备,其特征在于,所述获取模块用于:  
如果所述当前节点设备运行于候选状态,从所述心跳信息中获取多个签名;  
如果所述多个签名的数量大于系统中节点设备数量的半数,且所述多个签名均验证通过,从所述心跳信息中获取运行周期信息和最新日志索引。
8. 根据权利要求6所述的节点设备,其特征在于,所述接收模块还用于:  
接收所述运行于领导状态的节点设备所广播的日志复制指令,基于所述日志复制指令复制日志。
9. 根据权利要求6所述的节点设备,其特征在于,所述节点设备还包括:  
确定模块,用于基于所述当前节点设备的最新日志索引和所述心跳信息中的最新日志索引,确定所述当前节点设备需要添加的日志索引;  
发送模块,用于向所述运行于领导状态的节点设备发送日志添加请求。
10. 根据权利要求6-9中任一项所述的节点设备,其特征在于,所述心跳信息或所述日志复制指令携带系统中各个节点设备在响应所述运行于领导状态的节点设备的投票请求时的签名。

## 节点设备运行方法及节点设备

### 技术领域

[0001] 本发明涉及网络技术领域,特别涉及一种节点设备运行方法及节点设备。

### 背景技术

[0002] 随着网络技术的发展,基于集群为客户端提供服务的方式越来越普遍。为了保证集群中各个节点设备保持一致性,节点设备运行时一般可以应用BFT-Raft (Byzantine Fault Tolerance algorithm-Raft,拜占庭容错筏算法)。

[0003] 根据BFT-Raft,节点设备的工作状态可以分为三种:跟随状态follower、候选状态candidate和领导状态leader。当任一节点设备a运行于跟随状态时,可以根据该集群中运行于领导状态的节点设备b所广播的心跳信息,确定该节点设备b运行正常,并基于节点设备b的指示复制日志。当节点设备a在一段时间内未接收到节点设备b的心跳信息,可以确定节点设备b运行故障,并切换为候选状态运行,将投票请求广播至集群中的各个节点设备,一旦接收到该集群中半数以上的节点设备的投票,节点设备a可以切换为领导状态运行,并将心跳信息广播至集群中的各个节点设备、基于和客户端的交互存储日志、指示各个节点设备复制日志。需要说明的是,在运行于候选状态的节点设备a接收到心跳信息时,如果该心跳信息携带的运行周期信息小于该节点设备a的运行周期信息,则节点设备a会忽略该心跳信息。

[0004] 在实现本发明的过程中,发明人发现现有技术至少存在以下问题:

[0005] 由于一个集群可能分裂成网络相隔离的两个子集群,如,子集群A和子集群B,该子集群B中包括该集群中运行于领导状态的节点设备a,且子集群A的节点设备数量小于子集群B的节点设备数量,则子集群A中的节点设备不能通过投票选出一个新的运行于领导状态的节点设备,因此子集群A的节点设备的运行周期信息随时间递增,当子集群A与子集群B恢复网络连接后,子集群A中的任一节点设备b可以接收到节点设备a的心跳信息,由于节点设备a的运行周期信息小于节点设备b的运行周期信息,则节点设备b会忽略该心跳信息,导致节点设备b无法加入原来的集群。

### 发明内容

[0006] 为了解决现有技术的问题,本发明实施例提供了一种节点设备运行方法及节点设备。所述技术方案如下:

[0007] 一方面,提供了一种节点设备运行方法,所述方法包括:

[0008] 接收运行于领导状态的节点设备的心跳信息;

[0009] 如果当前节点设备运行于候选状态,从所述心跳信息中获取运行周期信息和最新日志索引;

[0010] 如果所述心跳信息中的运行周期信息小于所述当前节点设备的运行周期信息,且所述心跳信息中的最新日志索引大于等于所述当前节点设备的最新日志索引,将所述当前节点设备的运行周期信息同步为所述心跳信息中的运行周期信息,并将当前工作状态切换

为跟随状态运行或保持所述候选状态。

[0011] 另一方面,提供了一种节点设备,所述节点设备包括:

[0012] 接收模块,用于接收运行于领导状态的节点设备的心跳信息;

[0013] 获取模块,用于如果当前节点设备运行于候选状态,从所述心跳信息中获取运行周期信息和最新日志索引;

[0014] 运行模块,用于如果所述心跳信息中的运行周期信息小于所述当前节点设备的运行周期信息,且所述心跳信息中的最新日志索引大于等于所述当前节点设备的最新日志索引,将所述当前节点设备的运行周期信息同步为所述心跳信息中的运行周期信息,并将当前工作状态切换为跟随状态运行或保持所述候选状态。

[0015] 本发明实施例通过在接收到心跳信息时,获取心跳中的运行周期信息和最新日志索引,如果获取的运行周期信息小于自身的运行周期信息,且获取的最新日志索引不小于自身的最新日志索引,则将自身的运行周期信息同步为该心跳信息中的运行周期信息,并将自身的工作状态切换为跟随状态或保持候选状态,使得切换为跟随状态的节点设备可以直接根据本次心跳信息与第二子集群合为一个系统工作,或者保持候选状态的节点设备可以当再次接收到心跳信息时,依据自身的运行周期信息和最新日志索引均不大于心跳信息中的对应两项,切换为跟随状态运行,进而也能和第二子集群合为一个系统工作,提高了系统的工作可靠性。

## 附图说明

[0016] 为了更清楚地说明本发明实施例中的技术方案,下面将对实施例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。

[0017] 图1是本发明实施例提供的一种节点设备运行的实施环境示意图;

[0018] 图2是本发明实施例提供的一种节点设备工作状态的切换示意图;

[0019] 图3是本发明实施例提供的一种节点设备运行方法的流程图;

[0020] 图4是本发明实施例提供的一种节点设备的模块示意图;

[0021] 图5是本发明实施例提供的一种节点设备的模块示意图;

[0022] 图6是本发明实施例提供的一种节点设备结构示意图。

## 具体实施方式

[0023] 为使本发明的目的、技术方案和优点更加清楚,下面将结合附图对本发明实施方式作进一步地详细描述。

[0024] 图1是本发明实施例提供的一种节点设备运行的实施环境示意图。参见图1,该实施环境为一个由多个节点设备构成的系统,该系统也相当于一个集群,节点设备2为该系统中运行于领导状态的节点设备,在节点设备2运行正常时,可以定时地向各个运行于跟随状态的节点设备广播心跳信息,如,节点设备3、节点设备4,每个运行于跟随状态的节点设备在接收到心跳信息时可以确定节点设备1运行正常,并重置定时器(一般为0.5-1秒之间的随机值,这样可以避免各个节点设备的定时器的计时时长相同可能造成反复选举的情况),

等待下一次心跳信息。

[0025] 事实上,系统中各个节点设备的工作状态是可以动态切换的,参见图2,本发明实施例提供了一种节点设备工作状态的切换示意图。一旦运行于跟随状态(follower)的节点设备在定时器超时的情况下没有接收到心跳信息,可以确定运行于领导状态的节点设备运行故障,并切换为候选状态(candidate)运行;进而,节点设备可以重置定时器,并广播投票请求,直到接收到该系统中半数以上的投票确认消息切换为领导状态(Leader)运行,或者接收到运行于领导状态的节点设备的心跳信息时切换为跟随状态运行,或者定时器超时的情况下保持候选状态开始新一轮选举;运行于领导状态的节点设备可以在发现比自身具有更高运行周期信息(term)的节点设备时切换为跟随状态运行。

[0026] 在该系统为客户端提供服务时,当该系统中的任一节点设备接收到客户端的服务命令时,可以将该服务命令重定向至节点设备2,由节点设备2向各个节点设备广播日志添加请求,该日志添加请求用于请求将该服务命令添加到日志中,如果节点设备2可以接收到各个节点设备对日志添加请求的确认消息,可以响应该客户端的服务命令,将该服务命令添加到日志中,并向各个节点设备广播日志复制指令,使得各个节点设备将该服务命令复制到日志中。在实际的应用场景中,该系统可以是底层基于区块链技术的交易系统,该服务命令可以为客户端的交易信息,每个节点设备所存储的日志可以对应一条区块链,当添加交易信息到日志中时,实际是将该交易信息存储到当前区块的下一区块中,由于已存储至区块链中的数据不可更改,可以有效地防止交易信息被篡改,提高交易信息的安全性。

[0027] 由于网络中断等原因,该系统中的各个节点设备可能分裂形成两个网络相隔的子集群,即第一子集群和第二子集群,且第一子集群的节点设备数量小于第二子集群的节点设备数量。该第二子集群中包括该系统中运行于领导状态的节点设备2。进而,该第二子集群中运行于跟随状态的节点设备可以依据该节点设备2定时广播的心跳信息继续正常工作;第一子集群由于和节点设备2的网络中断,其中运行于跟随状态的节点设备在定时器超时的情况下也不能接收到节点设备2的心跳信息,依据bft-raft的超时选举机制,运行于跟随状态的节点设备会切换为候选状态运行,重置定时器,将自身的运行周期信息加一,并广播投票请求。然而,由于该第一子集群的节点设备数量小于系统中节点设备数量的一半,因此该第一子集群中的任一节点设备都不能接收到大于该系统中节点设备数量的一般的投票请求,导致直到各个节点和设备的计时器超时,该第一子集群中也不能选出领导状态的节点设备,则第一子集群中的节点设备会保持候选状态,但在定时器超时的情况下再次重置定时器,将自身的运行周期信息加一,并广播投票请求,如果第一子集群和第二子集群恢复网络连接,依照现有技术,即使该第一子集群中的节点设备接收到节点设备2的心跳信息,但由于该心跳信息中的运行周期信息小于该第二子集群中的节点设备的投票请求,该第一子集群中的节点设备也会忽略节点设备2的心跳信息,并继续等待其他节点设备的投票或者接收到有资格成为领导状态的节点设备的心跳信息,直到定时器再次超时,如此循环。

[0028] 图3是本发明实施例提供的一种节点设备运行方法的流程图,参见图3,该方法可以应用于图1所示实施例的节点设备1,具体包括以下步骤:

[0029] 301、节点设备1接收运行于领导状态的节点设备的心跳信息。

[0030] 其中,该节点设备1运行于候选状态,可以在选举开始时基于自身的运行周期信

息、最新日志索引(last log index)和节点设备标识生成投票请求,并广播该投票请求。该运行于领导状态的节点设备可以为图1所示实施例中的第二子集群中的节点设备2。由于该第一子集群的节点设备数量小于系统中节点设备数量的一半,该第一子集群中的任一节点设备都不能接收到超过系统中节点设备数量的一半的投票,则第一子集群中始终不能选出领导状态的节点设备,进而导致节点设备1始终运行于候选状态,并在每次定时器超时时递增自身的运行周期信息,开始一轮新的选举。心跳信息由第二子集群中的节点设备2定时广播,可以携带该节点设备2的运行周期信息、最新日志索引和节点设备标识等。在第一子集群和第二子集群恢复网络连接时,该节点设备2所广播的心跳信息可以被该运行于候选状态的节点设备接收到。

[0031] 302、如果节点设备1运行于候选状态,节点设备1从心跳信息中获取运行周期信息和最新日志索引。

[0032] 其中,运行周期信息是指发送该心跳信息的节点设备当前所处的运行周期号。每次进行选举时,由跟随状态切换为候选状态的节点设备的运行周期信息会加一,最终成为领导状态的节点设备可以将运行周期信息携带在心跳信息中,并广播给其他节点设备,接收到心跳信息的候选状态的节点设备可以切换为跟随状态运行,并将自身的运行周期信息同步为该心跳信息中的运行周期信息,因此,该运行周期信息可以表征一个节点设备是否始终与运行于领导状态的节点设备保持同步且运行正常。最新日志索引是指发送该心跳信息的节点设备最新存储的日志的索引,每次运行于领导状态的节点设备添加新的日志后,该最新日志索引加一,且该运行于领导状态的节点设备可以将日志复制指令广播给其他节点设备,使得接收到日志复制指令的节点设备可以同步该领导状态的节点设备的日志和最新日志索引,因此,该最新日志索引可以表征一个节点设备的日志完整性,显然,运行于领导状态的节点设备为在其系统中日志完整性最好的节点设备。

[0033] 该步骤中,节点设备1可以分别按照运行周期信息和最新日志索引在心跳信息中的协议位置,从心跳信息中分别提取出对应协议位置的运行周期信息和最新日志索引。

[0034] 在实际的应用场景中,为了避免有的网络设备伪装成领导状态的节点设备来发送心跳信息,提高系统的安全性,该心跳信息需携带系统中各个节点设备在响应运行于领导状态的节点设备的投票请求时的签名。当接收到该心跳信息时,如果节点设备1运行于候选状态,可以从心跳信息中获取多个签名;如果多个签名的数量大于系统中节点设备数量的半数,且多个签名均验证通过,从心跳信息中获取运行周期信息和最新日志索引。该系统中的每个节点设备可以配置有自身的私钥以及各个节点设备的公钥。节点设备1可以从心跳信息中提取出各个节点设备的签名作为该多个签名,并采用已配置的任一节点设备的公钥对该节点设备的签名进行验证,如果各个节点设备的签名均验证通过,且验证通过的签名数量大于该系统中节点设备数量的一半,说明该心跳信息确实来自运行于领导状态的节点设备,则可以从心跳信息中获取运行周期信息和最新日志索引,并继续执行步骤302。

[0035] 303、节点设备1判断心跳信息中的运行周期信息是否小于节点设备1的运行周期信息,如果是,执行步骤304,如果不是,忽略该心跳信息。

[0036] 该步骤中,为了印证该系统目前处于分裂后的子集群之间已恢复网络连接,且第一子集群内的节点设备正在进行选举的实施场景,考虑到第一子集群和第二子集群在初始分裂时系统中各个节点设备的运行周期信息相同,则在分裂后该第一子集群不断进行选举

的过程中节点设备1的运行周期信息会不断递增,而第二子集群由于节点设备2运行正常,其中的节点设备的运行周期信息会保持不变,因此运行周期信息可以作为上述实施场景的印证依据之一,如果心跳信息中的运行周期信息小于该节点设备1的运行周期信息,上述实施场景得到印证,则继续执行步骤304,如果该心跳信息中的运行周期信息不小于节点设备1的运行周期信息,说明不符合上述实施场景,可以忽略该心跳信息,但由于该条件符合选举出新的领导状态的节点设备的情况之一,则节点设备1也可以继续执行步骤304(事实上,该情况属于系统中正常选举的情况,本发明实施例不限定节点设备1如何处理)。

[0037] 304、节点设备1判断心跳信息中的最新日志索引是否大于等于当前节点设备的最新日志索引,如果是,将节点设备1的运行周期信息同步为心跳信息中的运行周期信息,并将当前工作状态切换为跟随状态运行,如果否,忽略该心跳信息。

[0038] 考虑到在系统分裂之前,该系统中的各个节点设备的日志理应与节点设备2的日志同步,因此第二子集群经过为客户端服务的一段时间,在两个子集群恢复网络连接之后,节点设备2所存储的日志应该不少于该节点设备1所存储的日志,也因此可以将最新日志索引作为印证上述实施场景的依据之一,如果多个节点设备的投票请求中的最新日志索引均不小于节点设备1的最新日志索引,说明该投票请求对应的节点设备已存储的日志量等于或多于节点设备1的日志量,上述实施场景最终得到各项印证,因此节点设备1可以将节点设备1的运行周期信息同步为心跳信息中的运行周期信息,并切换为跟随状态运行,如果该心跳信息中的最新日志索引小于节点设备1的最新日志索引,上述实施场景没有得到印证,则可以忽略该心跳信息。

[0039] 当然,如果节点设备1切换为跟随状态,还需要重置定时器,并且基于当前节点设备1的最新日志索引和心跳信息中的最新日志索引,确定节点设备1需要添加的日志索引;向运行于领导状态的节点设备发送日志添加请求。该日志添加请求可以携带需要添加的日志索引,使得节点设备2接收到日志添加请求时,可以将需要添加的日志索引对应的日志返回给节点设备1。

[0040] 事实上,节点设备1也可以将节点设备1的运行周期信息同步为心跳信息中的运行周期信息,并保持候选状态运行。该实现方式中,节点设备1可以直到再次接收到节点设备2的心跳信息时切换为跟随状态运行,并同步该节点设备2的日志。

[0041] 需要说明的是,本发明实施例对节点设备1执行步骤303和304的时序不做具体限定,事实上,节点设备1也可以先对最新日志索引进行判断,再对运行周期信息进行判断,或者,为了提高判断效率,并尽快使得第一子集群和第二子集群合为一个系统工作,节点设备1也可以同时对最新日志索引和运行周期信息进行判断,只要二者分别满足上述各自的判断条件,节点设备1即可将当前工作状态切换至跟随状态(或保持候选状态)。

[0042] 基于上述节点设备运行方法,原来为该第一子集群中的运行于候选状态的任一节点设备均能与该第二子集群合为一个系统工作。

[0043] 本发明实施例通过在接收到心跳信息时,获取心跳中的运行周期信息和最新日志索引,如果获取的运行周期信息小于自身的运行周期信息,且获取的最新日志索引不小于自身的最新日志索引,则将自身的运行周期信息同步为该心跳信息中的运行周期信息,并将自身的工作状态切换为跟随状态或保持候选状态,使得切换为跟随状态的节点设备可以直接根据本次心跳信息与第二子集群合为一个系统工作,或者保持候选状态的节点设备可以



当再次接收到心跳信息时,依据自身的运行周期信息和最新日志索引均不大于心跳信息中的对应两项,切换为跟随状态运行,进而也能和第二子集群合为一个系统工作,提高了系统的工作可靠性。

[0044] 305、节点设备1接收运行于领导状态的节点设备所广播的日志复制指令,基于该日志复制指令复制日志。

[0045] 为了保证系统中各个节点设备都能存储完整的日志,从而保证系统的一致性,该运行于领导状态的节点设备可以在系统中的各个节点设备确定添加新的日志后广播日志复制指令,使得该节点设备1可以接收到该日志复制指令,从而将该系统最新接收到的服务指令添加到日志中。当然,基于bft-raft不仅解决节点设备一致性,而且解决了节点设备欺诈,数据被篡改、丢失或顺序错乱的问题,该日志复制指令需携带系统中各个节点设备在响应运行于领导状态的节点设备的投票请求时的签名,使得节点设备1可以对该日志复制指令进行验证,并在验证通过后进行日志复制。

[0046] 需要说明的是,在第一子集群和第二子集群在恢复网络连接后,如果该系统中任一节点设备在定时器超时的情况下没有接收到心跳信息,可以确定运行于领导状态的节点设备运行故障,则切换为候选状态,重置定时器,并广播投票请求,直到接收到大于该系统中节点设备数量的半数的投票请求时成为新的领导状态的节点设备,或者直到接收到新的领导状态的节点设备的心跳信息时切换为跟随状态。

[0047] 图4是本发明实施例提供的一种节点设备的模块示意图。参见图4,该节点设备具体包括:

[0048] 接收模块401,用于接收运行于领导状态的节点设备的心跳信息;

[0049] 获取模块402,用于如果当前节点设备运行于候选状态,从心跳信息中获取运行周期信息和最新日志索引;

[0050] 运行模块403,用于如果心跳信息中的运行周期信息小于当前节点设备的运行周期信息,且心跳信息中的最新日志索引大于等于当前节点设备的最新日志索引,并将当前节点设备的运行周期信息同步为心跳信息中的运行周期信息,并将当前工作状态切换为跟随状态运行或保持候选状态。

[0051] 本发明实施例通过在接收到心跳信息时,获取心跳中的运行周期信息和最新日志索引,如果获取的运行周期信息小于自身的运行周期信息,且获取的最新日志索引不小于自身的最新日志索引,则将自身的运行周期信息同步为该心跳信息中的运行周期信息,并将自身的工作状态切换为跟随状态或保持候选状态,使得切换为跟随状态的节点设备可以直接根据本次心跳信息与第二子集群合为一个系统工作,或者保持候选状态的节点设备可以当再次接收到心跳信息时,依据自身的运行周期信息和最新日志索引均不大于心跳信息中的对应两项,切换为跟随状态运行,进而也能和第二子集群合为一个系统工作,提高了系统的工作可靠性。

[0052] 在一种可能实现方式中,获取模块402用于:

[0053] 如果当前节点设备运行于候选状态,从心跳信息中获取多个签名;

[0054] 如果多个签名的数量大于系统中节点设备数量的半数,且多个签名均验证通过,从心跳信息中获取运行周期信息和最新日志索引。

[0055] 在一种可能实现方式中,接收模块401还用于:

[0056] 接收运行于领导状态的节点设备所广播的日志复制指令,基于日志复制指令复制日志。

[0057] 在一种可能实现方式中,基于图4的节点设备组成,参见图5,节点设备还包括:

[0058] 确定模块504,用于基于当前节点设备的最新日志索引和心跳信息中的最新日志索引,确定当前节点设备需要添加的日志索引;

[0059] 发送模块504,用于向运行于领导状态的节点设备发送日志添加请求。

[0060] 在一种可能实现方式中,心跳信息或日志复制指令携带系统中各个节点设备在响应运行于领导状态的节点设备的投票请求时的签名。

[0061] 上述所有可选技术方案,可以采用任意结合形成本发明的可选实施例,在此不再一一赘述。

[0062] 需要说明的是:上述实施例提供的节点设备在执行节点设备运行方法时,仅以上述各功能模块的划分进行举例说明,实际应用中,可以根据需要而将上述功能分配由不同的功能模块完成,即将节点设备的内部结构划分成不同的功能模块,以完成以上描述的全部或者部分功能。另外,上述实施例提供的节点设备与节点设备运行方法实施例属于同一构思,其具体实现过程详见方法实施例,这里不再赘述。

[0063] 图6是本发明实施例提供的一种节点设备结构示意图。参照图6,该节点设备可以被提供为一个服务器,该节点设备600包括处理组件622,其进一步包括一个或多个处理器,以及由存储器632所代表的存储器资源,用于存储可由处理部件622的执行的指令,例如应用程序。存储器632中存储的应用程序可以包括一个或一个以上的每一个对应于一组指令的模块。此外,处理组件622被配置为执行指令,以执行上述节点设备运行方法。

[0064] 节点设备600还可以包括一个电源组件626被配置为执行节点设备600的电源管理,一个有线或无线网络接口650被配置为将节点设备600连接到网络,和一个输入输出(I/O)接口658。节点设备600可以操作基于存储在存储器632的操作系统,例如Windows Server™,Mac OS X™,Unix™,Linux™,FreeBSD™或类似。

[0065] 本领域普通技术人员可以理解实现上述实施例的全部或部分步骤可以通过硬件来完成,也可以通过程序来指令相关的硬件完成,所述的程序可以存储于一种计算机可读存储介质中,上述提到的存储介质可以是只读存储器,磁盘或光盘等。

[0066] 以上所述仅为本发明的较佳实施例,并不用以限制本发明,凡在本发明的精神和原则之内,所作的任何修改、等同替换、改进等,均应包含在本发明的保护范围之内。

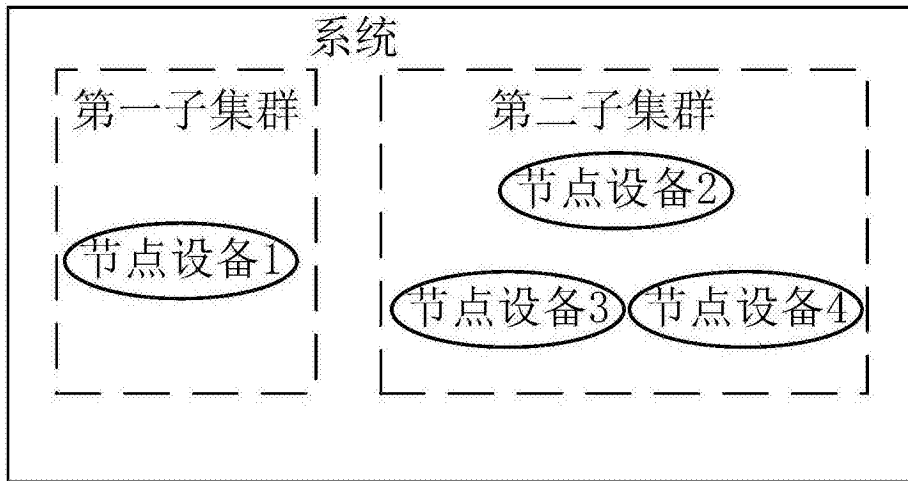


图1

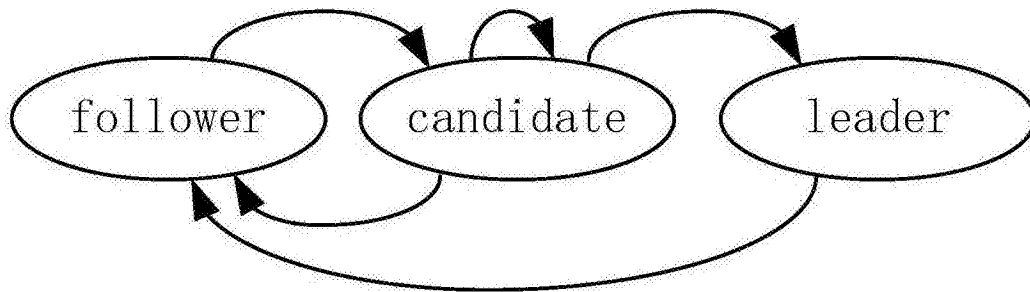


图2

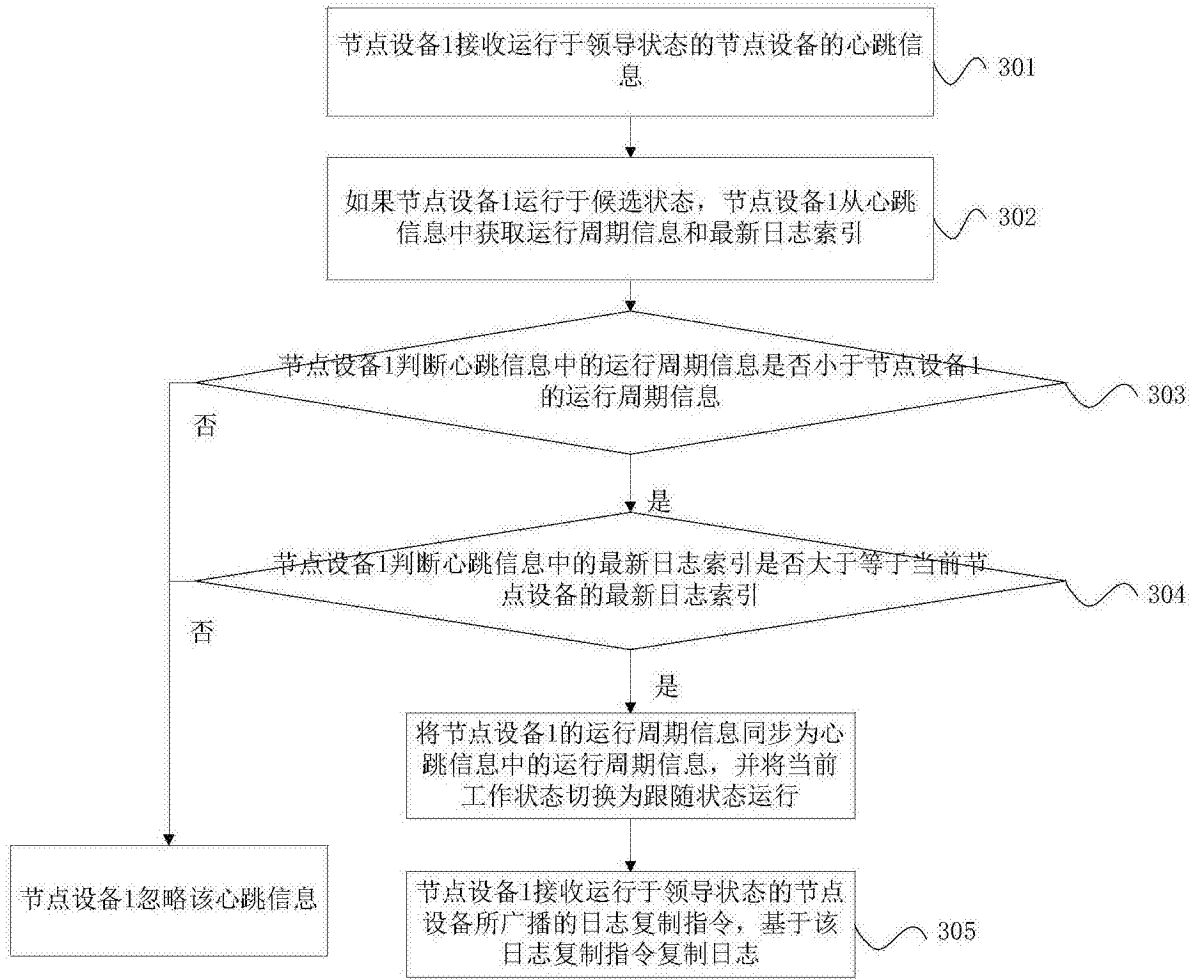


图3

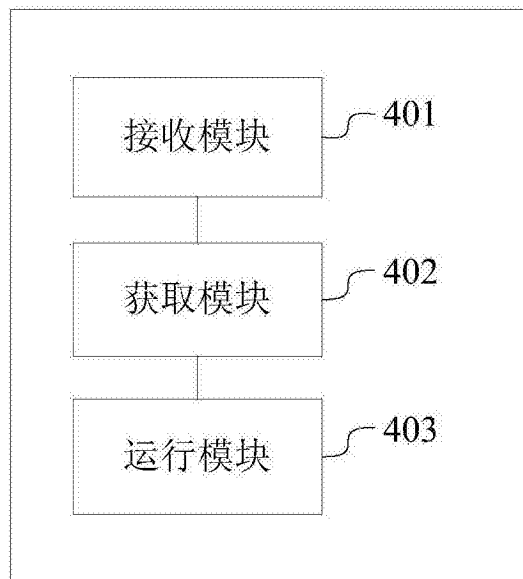


图4

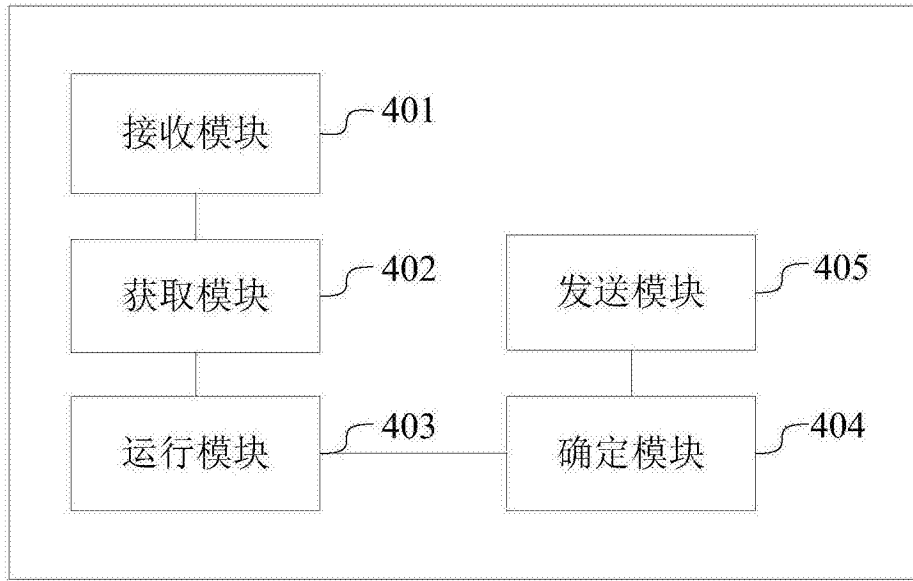


图5

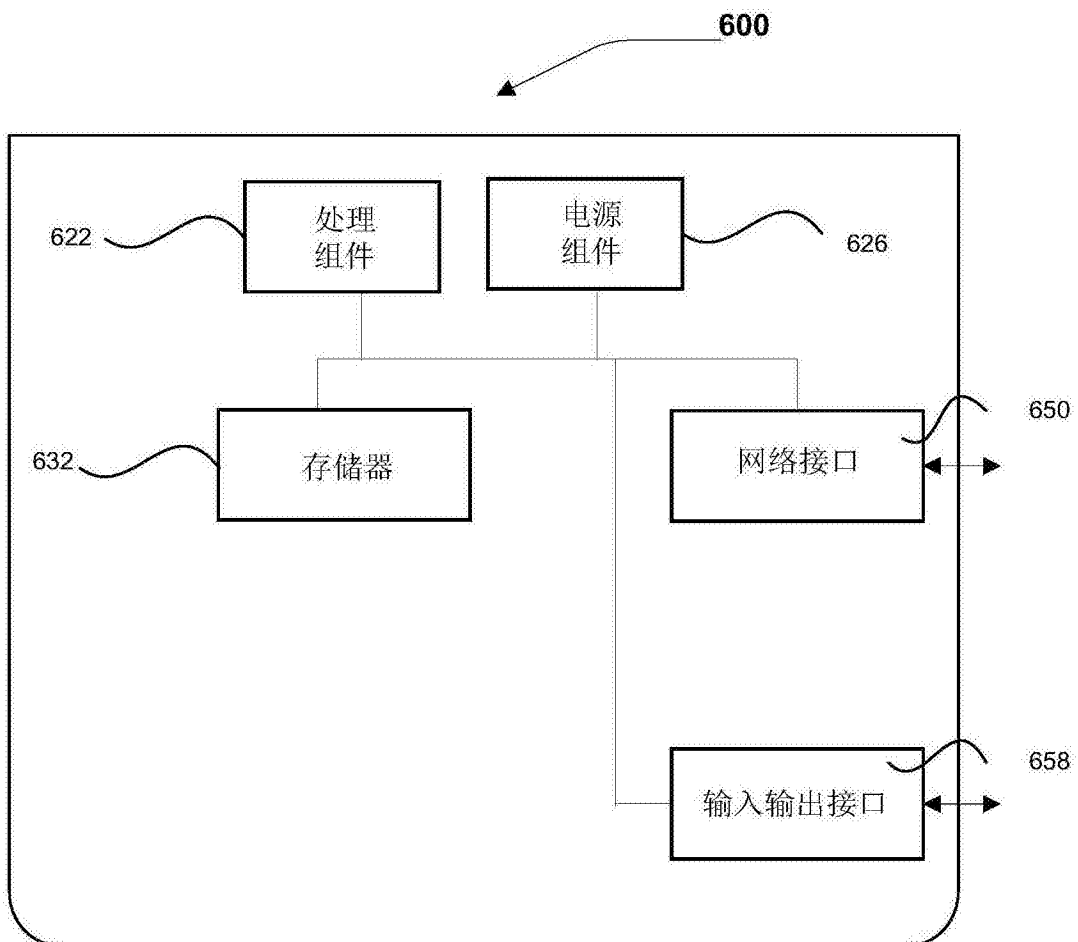


图6