(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 106155780 B
(45)授权公告日 2020.01.31

(21)申请号 201510155909.9

(22)申请日 2015.04.02

(65)同一申请的已公布的文献号
申请公布号 CN 106155780 A

(43)申请公布日 2016.11.23

(73)专利权人 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 阳振坤 肖金亮

(74)专利代理机构 北京晋德允升知识产权代理有限公司 11623
代理人 杨移

(51)Int.Cl.
G06F 9/48(2006.01)
G06F 9/50(2006.01)

(56)对比文件
CN 102682572 A,2012.09.19,说明书第6-17段及附图2-4.
CN 102682572 A,2012.09.19,说明书第6-17段及附图2-4.
US 7139790 B1,2006.11.21,全文.
US 2008071853 A1,2008.03.20,全文.
CN 104933132 A,2015.09.23,全文.
CN 101252603 A,2008.08.27,全文.
CN 104301971 A,2015.01.21,全文.
CN 103634375 A,2014.03.12,全文.
cszhouwei.Raft一致性算法.《https://blog.csdn.net/cszhouwei/article/details/38374603》.2014,第2-4页.

审查员 崔芳婷

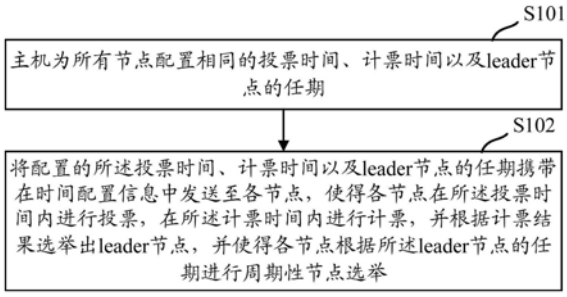
权利要求书5页 说明书15页 附图8页

(54)发明名称

一种基于时间的节点选举方法及装置

(57)摘要

本申请公开了一种基于时间的节点选举方法及装置，该方法包括：主机为所有节点配置相同的投票时间和计票时间，将配置的所述投票时间和计票时间携带在时间配置信息中发送至各节点，使得各节点在所述投票时间内进行投票，在所述计票时间内进行计票，并根据计票结果选举出leader节点。各竞选节点根据主机配置的相同的投票时间和计票时间，同步执行投票操作和计票操作，改善了现有技术中各节点投票、计票时机不能确定的缺陷，从而使得对节点的选举过程中，极大提升了确定leader节点的效率。



CN 106155780 B

1.一种基于时间的节点选举方法，其特征在于，所述方法用于分布式系统中确定leader节点，所述方法包括：

主机为所有节点配置相同的投票时间、计票时间以及leader节点的任期；

将配置的所述投票时间、计票时间以及leader节点的任期携带在时间配置信息中发送至各节点，不同的节点均同步触发所述投票时间和所述计票时间，使得当各节点的内置时钟到达配置的所述投票时间时，各节点在所述投票时间内同步进行投票，当各节点的内置时钟到达配置的所述计票时间时，各节点在所述计票时间内同步进行计票，并根据计票结果选举出leader节点，并使得各节点根据所述leader节点的任期进行周期性节点选举；

其中，各节点在所述投票时间内同步进行投票具体包括各节点根据接收到的投票请求中携带的节点信息来进行投票。

2.一种基于时间的节点选举方法，其特征在于，所述方法用于分布式系统中确定leader节点，分布式系统中不同的节点均根据内置时钟同步触发时间配置信息所携带的投票时间和计票时间，所述方法包括：

当所有节点中不存在leader节点时，至少一个节点作为竞选节点，在所述投票时间内，向所有节点发送投票请求，以使各节点在所述投票时间内进行投票；

各竞选节点接收其他节点根据所述投票请求生成的投票信息，其中，所述其他节点根据所述投票请求生成的投票信息具体为其他节点根据接收到的投票请求中携带的节点信息生成的投票信息；

各竞选节点在所述时间配置信息所携带的计票时间内，同步统计各自接收到的投票信息的数量；

各竞选节点根据统计出的各自的投票信息的数量确定leader节点，对指定业务进行处理。

3.如权利要求2所述的方法，其特征在于，在所述竞选节点在计票时间内，统计接收到的所述投票信息的数量之前，还包括：

所述竞选节点在所述计票时间到来时，停止接收所述投票信息。

4.如权利要求2所述的方法，其特征在于，各竞选节点根据统计出的各自的投票信息的数量确定leader节点，具体包括：

当所述竞选节点判断统计出的所述投票信息的数量，超过预设的阈值时，该竞选节点将自身确定为leader节点；

其中，所述阈值为所有节点数量的二分之一。

5.一种基于时间的节点选举方法，其特征在于，所述方法用于分布式系统中确定leader节点，分布式系统中不同的节点均根据内置时钟同步触发时间配置信息所携带的投票时间和计票时间，所述方法包括：

当所有节点中不存在leader节点时，至少两个节点作为投票节点，在所述投票时间内，接收竞选节点发送的投票请求；

所述投票节点根据所述投票请求，在投票时间内同步生成投票信息发送至竞选节点，使得所述竞选节点在计票时间内同步统计所述投票信息的数量，确定leader节点，其中，所述生成投票信息具体包括根据接收到的投票请求中携带的节点信息来生成投票信息。

6.如权利要求5所述的方法，其特征在于，生成投票信息发送至竞选节点，具体包括：

所述投票节点确定接收到的所述投票请求中所携带的节点信息；

根据所述节点信息，确定所述竞选节点的节点状态权重；

选定节点状态权重最大的竞选节点，并生成针对该竞选节点的投票信息，发送至选定的竞选节点。

7.一种基于时间的节点选举方法，其特征在于，所述方法用于分布式系统中确定leader节点，分布式系统中不同的节点均根据内置时钟同步触发时间配置信息所携带的投票时间和计票时间，所述方法包括：

当所有节点中存在leader节点，且所述leader节点确定自身符合预设标准时，所述leader节点根据时间配置信息中的leader节点的任期，在所述投票时间内向所有节点发送连任请求，以使各节点在所述投票时间内同步进行投票；

所述leader节点接收其他节点根据所述连任请求生成的投票信息；

所述leader节点在计票时间内，统计接收到的所述投票信息的数量；

所述leader节点根据统计出的所述投票信息的数量确定连任状态，并在成功连任后对指定业务进行处理；

其中，所述预设标准，包括处理负载标准、数据更新程度中的至少一种。

8.一种基于时间的节点选举方法，其特征在于，所述方法用于分布式系统中确定leader节点，分布式系统中不同的节点均根据内置时钟同步触发时间配置信息所携带的投票时间和计票时间，所述方法包括：

当所有节点中存在leader节点，且所述leader节点确定自身不符合预设标准时，所述leader节点根据时间配置信息中的leader节点的任期，在其他节点中筛选出符合所述预设标准的节点，作为继承节点；

所述leader节点在所述投票时间内，向其他节点发出针对该继承节点的投票请求，以使各节点在所述投票时间内同步进行投票，使得所述继承节点接收其他节点发送的投票信息，并在计票时间内统计所述投票信息的数量确定为新任leader节点。

9.一种基于时间的节点选举方法，其特征在于，所述方法用于分布式系统中确定leader节点，包括：

不同的节点均根据内置时钟同步触发时间配置信息中所携带的投票时间和计票时间，当所有节点中存在leader节点，且所述leader节点确定自身符合预设标准时，至少两个节点作为投票节点，在所述投票时间内，接收所述leader节点发送的连任请求；

所述投票节点根据所述连任请求，在时间配置信息中所携带的投票时间内，同步生成投票信息发送至所述leader节点，使得所述leader节点在计票时间内统计所述投票信息的数量确定连任状态；

并且，所述投票节点在所述leader节点的任期内只接收所述leader节点的投票请求。

10.一种基于时间的节点选举方法，其特征在于，所述方法用于分布式系统中确定leader节点，不同的节点均根据内置时钟同步触发时间配置信息中所携带的投票时间和计票时间，所述方法包括：

当所有节点中存在leader节点，且所述leader节点确定自身不符合预设标准时，至少两个节点作为投票节点，在所述投票时间内，接收所述leader节点发送的针对继承节点的投票请求；

所述投票节点根据所述投票请求，在时间配置信息中所携带的投票时间内，同步生成投票信息发送至所述继承节点，使得所述继承节点在计票时间内统计所述投票信息的数量，确定新任leader节点。

11.一种基于时间的节点选举装置，其特征在于，所述装置用于分布式系统中确定leader节点，所述装置包括：

配置模块，用于为所有节点配置相同的投票时间、计票时间以及leader节点的任期；

发送模块，用于将配置的所述投票时间、计票时间以及leader节点的任期携带在时间配置信息中发送至各节点，不同的节点均同步触发所述投票时间和所述计票时间，使得当各节点的内置时钟到达配置的所述投票时间时，各节点在所述投票时间内同步进行投票，当各节点的内置时钟到达配置的所述计票时间时，各节点在所述计票时间内同步进行计票，根据计票结果选举出leader节点，并使得各节点根据所述leader节点的任期进行周期性节点选举；

其中，各节点在所述投票时间内同步进行投票具体包括各节点根据接收到的投票请求中携带的节点信息来进行投票。

12.一种基于时间的节点选举装置，其特征在于，所述装置用于分布式系统中确定leader节点，设置在竞选节点中，分布式系统中不同的节点均根据内置时钟同步触发时间配置信息所携带的投票时间和计票时间，且当所有节点中不存在leader节点时，所述装置包括：

投票请求模块，用于在所述投票时间内，向所有节点发送投票请求，以使各节点在所述投票时间内进行投票；

接收模块，用于接收其他节点根据所述投票请求生成的投票信息，其中，所述其他节点根据所述投票请求生成的投票信息具体为其他节点根据接收到的投票请求中携带的节点信息生成的投票信息；

统计模块，用于在所述时间配置信息所携带的计票时间内，同步统计各自接收到的投票信息的数量；

确定模块，用于根据统计出的各自的投票信息的数量确定leader节点，对指定业务进行处理。

13.如权利要求12所述的装置，其特征在于，所述接收模块，还用于在所述计票时间到来时，停止接收所述投票信息。

14.如权利要求12所述的装置，其特征在于，所述确定模块，具体用于当判断统计出的所述投票信息的数量，超过预设的阈值时，将所述竞选节点确定为leader节点；

其中，所述阈值为所有节点数量的二分之一。

15.一种基于时间的节点选举装置，其特征在于，所述装置用于分布式系统中确定leader节点，设置在投票节点中，分布式系统中不同的节点均根据内置时钟同步触发时间配置信息所携带的投票时间和计票时间，且当所有节点中不存在leader节点时，所述装置包括：

接收模块，用于在所述投票时间内，接收竞选节点发送的投票请求；

投票信息模块，用于根据所述投票请求，在投票时间内同步生成投票信息发送至竞选节点，使得所述竞选节点在计票时间内同步统计所述投票信息的数量，确定leader节点，其

中，所述生成投票信息具体包括根据接收到的投票请求中携带的节点信息来生成投票信息。

16.如权利要求15所述的装置，其特征在于，所述投票信息模块，具体用于确定接收到的所述投票请求中所携带的节点信息，根据所述节点信息，确定所述竞选节点的节点状态权重，选定节点状态权重最大的竞选节点，并生成针对该竞选节点的投票信息，发送至选定的竞选节点。

17.一种基于时间的节点选举装置，其特征在于，所述装置用于分布式系统中确定leader节点，设置在leader节点中，分布式系统中不同的节点均根据内置时钟同步触发时间配置信息所携带的投票时间和计票时间，当所有节点中存在leader节点，且所述leader节点确定自身符合预设标准时，所述装置包括：

连任请求模块，用于根据时间配置信息中的leader节点的任期，在所述投票时间内向所有节点发送连任请求，以使各节点在所述投票时间内同步进行投票；

接收模块，用于接收其他节点根据所述连任请求生成的投票信息；

统计模块，用于在计票时间内，统计接收到的所述投票信息的数量；

确定模块，用于根据统计出的所述投票信息的数量确定连任状态，并在成功连任后对指定业务进行处理；

其中，所述预设标准，包括处理负载标准、数据更新程度中的至少一种。

18.一种基于时间的节点选举装置，其特征在于，所述装置用于分布式系统中确定leader节点，设置在leader节点中，分布式系统中不同的节点均根据内置时钟同步触发时间配置信息所携带的投票时间和计票时间，当所有节点中存在leader节点，且所述leader节点确定自身不符合预设标准时，所述装置包括：

筛选模块，用于根据时间配置信息中的leader节点的任期，在其他节点中筛选出符合所述预设标准的节点，作为继承节点；

投票请求模块，用于在所述投票时间内，向其他节点发出针对该继承节点的投票请求，以使各节点在所述投票时间内同步进行投票，使得所述继承节点接收其他节点发送的投票信息，并在计票时间内统计所述投票信息的数量确定为新任leader节点。

19.一种基于时间的节点选举装置，其特征在于，所述装置用于分布式系统中确定leader节点，设置在投票节点中，不同的节点均根据内置时钟同步触发时间配置信息中所携带的投票时间和计票时间，当所有节点中存在leader节点，且所述leader节点确定自身符合预设标准时，所述装置包括：

接收模块，用于在所述投票时间内，接收所述leader节点发送的连任请求；

投票信息模块，用于根据所述连任请求，在时间配置信息中所携带的投票时间内，同步生成投票信息发送至所述leader节点，使得所述leader节点在计票时间内统计所述投票信息的数量确定连任状态，并用于在所述leader节点的任期内只接收所述leader节点的投票请求。

20.一种基于时间的节点选举装置，其特征在于，所述装置用于分布式系统中确定leader节点，设置在投票节点中，不同的节点均根据内置时钟同步触发时间配置信息中所携带的投票时间和计票时间，当所有节点中存在leader节点，且所述leader节点确定自身不符合预设标准时，所述装置包括：

接收模块，用于在所述投票时间内，接收所述leader节点发送的针对继承节点的投票请求；

投票信息模块，用于根据所述投票请求，在时间配置信息中所携带的投票时间内，同步生成投票信息发送至所述继承节点，使得所述继承节点在计票时间内统计所述投票信息的数量，确定新任leader节点。

一种基于时间的节点选举方法及装置

技术领域

[0001] 本申请涉及计算机技术领域，尤其涉及一种基于时间的节点选举方法及装置。

背景技术

[0002] 随着信息技术的发展，互联网可以为用户提供多元化、海量化的信息，对于网络服务商而言，传统的单一服务器或数据库等处理设备，已经不能满足对海量数据进行处理的要求，因此分布式系统应运而生。

[0003] 通常，分布式系统可以对其中包含多个处理设备(如：服务器、处理器或者数据库等)，以全局方式对该系统中的各类数据资源进行调度、分布式处理、协同处理、远程控制等多种管理操作。对于分布式系统中的每一个处理设备而言，均可以看作是一个分布式节点(简称为节点)，这样一来，分布式系统便可以将巨量数据“分布”到系统中的每一个节点中，使得每一个节点都对相应的数据进行处理，从而有效提升对数据的处理效率并显著增加数据的吞吐量。

[0004] 在分布式系统中，为了进一步提升对不同节点中的数据资源的调度效率，需要将某些特殊类型的信息统一存储在同一节点中，如：元信息(元信息是对数据资源的描述性信息)，包括：数据的存储位置、更新状态、查找关键词等信息。

[0005] 现有技术中，采用节点选举的方式，选出合适的节点(也可称为leader节点)对元信息进行维护管理。其中，对节点的选举方式具体为：不同的节点均会向其他节点发送投票请求，各节点接收到投票请求后进行投票，最后根据统计出的票数，选出leader节点。

[0006] 但是，采用上述节点选举的方式时，不同的节点均有各自的大量数据需要处理，当某一节点处于空闲并向其他节点发送投票请求后，其他的各节点可能正在处理数据，那么，其他的节点即使接收到该节点发送的投票请求后，也不会立即进入投票状态，也就是说，各节点对投票、统计的时机均是不确定，不能及时有效地进行节点选举。

发明内容

[0007] 本申请实施例提供一种基于时间的节点选举方法及装置，用以解决目前进行节点选举时效率较低的问题。

[0008] 本申请实施例提供的一种基于时间的节点选举方法，所述方法用于分布式系统中确定leader节点，所述方法包括：

[0009] 主机为所有节点配置相同的投票时间、计票时间以及leader节点的任期；

[0010] 将配置的所述投票时间、计票时间以及leader节点的任期携带在时间配置信息中发送至各节点，使得各节点在所述投票时间内进行投票，在所述计票时间内进行计票，并根据计票结果选举出leader节点，并使得各节点根据所述leader节点的任期进行周期性节点选举。

[0011] 本申请实施例还提供的一种基于时间的节点选举方法，所述方法用于分布式系统中确定leader节点，所述方法包括：

[0012] 当所有节点中不存在leader节点时，至少一个节点作为竞选节点，在时间配置信息所携带的投票时间内，向所有节点发送投票请求；

[0013] 各竞选节点接收其他节点根据所述投票请求生成的投票信息；

[0014] 各竞选节点在所述时间配置信息所携带的计票时间内，统计各自接收到的投票信息的数量；

[0015] 各竞选节点根据统计出的各自的投票信息的数量确定leader节点，对指定业务进行处理。

[0016] 本申请实施例还提供的一种基于时间的节点选举方法，所述方法用于分布式系统中确定leader节点，所述方法包括：

[0017] 当所有节点中不存在leader节点时，至少两个节点作为投票节点，在时间配置信息所携带的投票时间内，接收竞选节点发送的投票请求；

[0018] 所述投票节点根据所述投票请求，在投票时间内生成投票信息发送至竞选节点，使得所述竞选节点在计票时间内统计所述投票信息的数量，确定leader节点。

[0019] 本申请实施例还提供的一种基于时间的节点选举方法，所述方法用于分布式系统中确定leader节点，所述方法包括：

[0020] 当所有节点中存在leader节点，且所述leader节点确定自身符合预设标准时，所述leader节点根据时间配置信息中的leader节点的任期，在投票时间内向所有节点发送连任请求；

[0021] 所述leader节点接收其他节点根据所述连任请求生成的投票信息；

[0022] 所述leader节点在计票时间内，统计接收到的所述投票信息的数量；

[0023] 所述leader节点根据统计出的所述投票信息的数量确定连任状态，并在成功连任后对指定业务进行处理；

[0024] 其中，所述预设标准，包括处理负载标准、数据更新程度中的至少一种。

[0025] 本申请实施例还提供的一种基于时间的节点选举方法，所述方法用于分布式系统中确定leader节点，所述方法包括：

[0026] 当所有节点中存在leader节点，且所述leader节点确定自身不符合预设标准时，所述leader节点根据时间配置信息中的leader节点的任期，在其他节点中筛选出符合所述预设标准的节点，作为继承节点；

[0027] 所述leader节点在所述投票时段内，向其他节点发出针对该继承节点的投票请求，使得所述继承节点接收其他节点发送的投票信息，并在计票时间内统计所述投票信息的数量确定为新任leader节点。

[0028] 本申请实施例还提供的一种基于时间的节点选举方法，所述方法用于分布式系统中确定leader节点，所述方法包括：

[0029] 当所有节点中存在leader节点时，至少两个节点作为投票节点，在时间配置信息中所携带的投票时间内，接收所述leader节点发送的连任请求；

[0030] 所述投票节点根据所述连任请求，在时间配置信息中所携带的投票时间内，生成投票信息发送至所述leader节点，使得所述leader节点在计票时间内统计所述投票信息的数量确定连任状态；

[0031] 并且，所述投票节点在所述leader节点的任期内接收到的投票请求不进行处理。

[0032] 本申请实施例还提供的一种基于时间的节点选举方法，所述方法用于分布式系统中确定leader节点，所述方法包括：

[0033] 当所有节点中存在leader节点时，至少两个节点作为投票节点，在时间配置信息中携带的投票时间内，接收所述leader节点发送的针对继承节点的投票请求；

[0034] 所述投票节点根据所述投票请求，在时间配置信息中所携带的投票时间内，生成投票信息发送至所述继承节点，使得所述继承节点在计票时间内统计所述投票信息的数量，确定新任leader节点。

[0035] 本申请实施例提供的一种基于时间的节点选举装置，所述装置用于分布式系统中确定leader节点，所述装置包括：

[0036] 配置模块，用于为所有节点配置相同的投票时间、计票时间以及leader节点的任期；

[0037] 发送模块，用于将配置的所述投票时间、计票时间以及leader节点的任期携带在时间配置信息中发送至各节点，使得各节点在所述投票时间内进行投票，在所述计票时间内进行计票，并根据计票结果选举出leader节点，并使得各节点根据所述leader节点的任期进行周期性节点选举。

[0038] 本申请实施例另提供的一种基于时间的节点选举装置，所述装置用于分布式系统中确定leader节点，设置在竞选节点中，且当所有节点中不存在leader节点时，所述装置包括：

[0039] 投票请求模块，用于在时间配置信息所携带的投票时间内，向所有节点发送投票请求；

[0040] 接收模块，用于接收其他节点根据所述投票请求生成的投票信息；

[0041] 统计模块，用于在所述时间配置信息所携带的计票时间内，统计各自接收到的投票信息的数量；

[0042] 确定模块，用于根据统计出的各自的投票信息的数量确定leader节点，对指定业务进行处理。

[0043] 本申请实施例另提供的一种基于时间的节点选举装置，所述装置用于分布式系统中确定leader节点，设置在投票节点中，且当所有节点中不存在leader节点时，所述装置包括：

[0044] 接收模块，用于在时间配置信息所携带的投票时间内，接收竞选节点发送的投票请求；

[0045] 投票信息模块，用于根据所述投票请求，在投票时间内生成投票信息发送至竞选节点，使得所述竞选节点在计票时间内统计所述投票信息的数量，确定leader节点。

[0046] 本申请实施例另提供的一种基于时间的节点选举装置，所述装置用于分布式系统中确定leader节点，设置在leader节点中，当所有节点中包含leader节点，且所述leader节点确定自身符合预设标准时，所述装置包括：

[0047] 连任请求模块，用于根据时间配置信息中的leader节点的任期，在投票时间内向所有节点发送连任请求；

[0048] 接收模块，用于接收其他节点根据所述连任请求生成的投票信息；

[0049] 统计模块，用于在计票时间内，统计接收到的所述投票信息的数量；

[0050] 确定模块，用于根据统计出的所述投票信息的数量确定连任状态，并在成功连任后对指定业务进行处理；

[0051] 其中，所述预设标准，包括处理负载标准、数据更新程度中的至少一种。

[0052] 本申请实施例另提供的一种基于时间的节点选举装置，所述装置用于分布式系统中确定leader节点，设置在leader节点中，当所有节点中包含leader节点，且所述leader节点确定自身不符合预设标准时，所述装置包括：

[0053] 筛选模块，用于根据时间配置信息中的leader节点的任期，在其他节点中筛选出符合所述预设标准的节点，作为继承节点；

[0054] 投票请求模块，用于在所述投票时段内，向其他节点发出针对该继承节点的投票请求，使得所述继承节点接收其他节点发送的投票信息，并在计票时间内统计所述投票信息的数量确定为新任leader节点。

[0055] 本申请实施例另提供的一种基于时间的节点选举装置，所述装置用于分布式系统中确定leader节点，设置在投票节点中，当所有节点中包含leader节点时，所述装置包括：

[0056] 接收模块，用于在时间配置信息中所携带的投票时间内，接收所述leader节点发送的连任请求；

[0057] 投票信息模块，用于根据所述连任请求，在时间配置信息中所携带的投票时间内，生成投票信息发送至所述leader节点，使得所述leader节点在计票时间内统计所述投票信息的数量确定连任状态。

[0058] 本申请实施例另提供的一种基于时间的节点选举装置，所述装置用于分布式系统中确定leader节点，设置在投票节点中，当所有节点中存在leader节点时，所述装置包括：

[0059] 接收模块，用于在时间配置信息中携带的投票时间内，接收所述leader节点发送的针对继承节点的投票请求；

[0060] 投票信息模块，用于根据所述投票请求，在时间配置信息中所携带的投票时间内，生成投票信息发送至所述继承节点，使得所述继承节点在计票时间内统计所述投票信息的数量，确定新任leader节点，并用于对在所述leader节点的任期内接收到的投票请求不进行处理。

[0061] 本申请实施例提供一种基于时间的节点选举方法及装置，通过本方法，主机为所有节点配置相同的投票时间和计票时间，将配置的所述投票时间和计票时间携带在时间配置信息中发送至各节点，使得各节点在所述投票时间内进行投票，在所述计票时间内进行计票，并根据计票结果选举出leader节点。各竞选节点根据主机配置的相同的投票时间和计票时间，同步执行投票操作和计票操作，改善了现有技术中各节点投票、计票时机不能确定的缺陷，从而使得对节点的选举过程中，极大提升了确定leader节点的效率。

附图说明

[0062] 此处所说明的附图用来提供对本申请的进一步理解，构成本申请的一部分，本申请的示意性实施例及其说明用于解释本申请，并不构成对本申请的不当限定。在附图中：

[0063] 图1、图2，以及图5至图9为本申请实施例提供的基于时间的节点选举过程；

[0064] 图3a至3d为本申请实施例提供的投票时间和计票时间的示意图；

[0065] 图4为本申请实施例提供的具体应用实例中各节点发送投票信息进行节点选举的

示意图；

[0066] 图10至图16为本申请实施例提供的基于时间的节点选举装置结构示意图。

具体实施方式

[0067] 为使本申请的目的、技术方案和优点更加清楚，下面将结合本申请具体实施例及相应的附图对本申请技术方案进行清楚、完整地描述。显然，所描述的实施例仅是本申请一部分实施例，而不是全部的实施例。基于本申请中的实施例，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，都属于本申请保护的范围。

[0068] 图1为本申请实施例提供的基于时间的节点选举方法，该方法用于分布式系统中确定leader节点，该方法具体包括以下步骤：

[0069] S101，主机为所有节点配置相同的投票时间、计票时间以及leader节点的任期。

[0070] 在本申请实施例中，分布式系统中设置有主机，用于对该分布式系统中的各节点进行管理，如：调度数据、配置数据、发出特定指令等。其中，所述主机包括但不限于：计算机、服务器等具有监控功能的网络设备。

[0071] 考虑到现有技术中，处于分布式系统中的各节点在进行节点选举时，不同节点发出投票请求，以及对该投票请求进行响应的时机均不确定，因此，在本发明实施例中，将通过分布式系统中的主机，对处于该分布式系统中的所述各节点进行时间配置，使得各节点在相同的时间中进行同类型的处理操作。也即，在上述步骤S101中，由主机为不同的节点配置相同的投票时间和计票时间。

[0072] 其中，本申请实施例中的投票时间和计票时间均是一段时间区间(并非是某时刻的时间点)，投票时间和计票时间的时间区间的长度(时长)，包括但不限于：毫秒(ms)、秒(s)、分钟(min)、小时(h)、天(D)、星期(week)、月(mon)等不同时间单位构成的时长。

[0073] 另外，主机还将为所有节点配置leader节点的任期，当任一节点选举为leader节点且经过了配置好的任期后，则开始下一次的节点选举过程。

[0074] S102，将配置的所述投票时间、计票时间以及leader节点的任期携带在时间配置信息中发送至各节点，使得各节点在所述投票时间内进行投票，在所述计票时间内进行计票，并根据计票结果选举出leader节点，并使得各节点根据所述leader节点的任期进行周期性节点选举。

[0075] 主机配置完成投票时间、计票时间以及leader节点的任期后，就需要通知各节点，以便使各节点按照配置的投票时间、计票时间以及leader节点的任期执行相应的操作，因此，主机会将配置好的投票时间、计票时间以及leader节点的任期携带在时间配置信息中，发送至各节点。这样一来，当各节点分别接收到时间配置信息后，则各节点会根据接收到的时间配置信息，调整自身的内置时钟，实现不同的节点均同步触发所述投票时间和计票时间，以便同步执行相应的操作。

[0076] 这里还需要说明的是，主机配置的投票时间、计票时间以及leader节点的任期具有周期性。这是因为：当各节点根据主机配置的投票时间和计票时间分别进行投票和计票操作，确定出leader节点之后，各节点与leader节点将对相应的业务进行处理，而由于实际应用中，分布式系统处理的数据量巨大，会使得leader节点和其他各节点的处理负载、数据更新程度发生变化，就有可能导致leader节点出现负荷超标、数据更新滞后、宕机等情况，

从而严重影响leader节点对指定业务的处理效率。在这样的情况下，若主机配置的投票时间、计票时间以及leader节点的任期不具有周期性(也即，分布式系统中的各节点只进行一次节点选举的过程，在leader节点出现故障后，各节点不会进行节点选举)，那么，就会导致分布式系统整体处理停滞。若要重新进行节点选举，只能等待主机重新配置投票时间和计票时间，也会造成分布式系统对业务处理的效率下降。

[0077] 因此，在本申请实施例中，由主机配置的投票时间、计票时间以及leader节点的任期将周期性的重复。例如：投票时间的间隔周期为3000ms，也就是说，当第一次投票时间结束的3000ms后，将进入第二次投票时间。又例如：leader节点的任期为10min，也即，当某一节点选举为leader节点并持续10min后，开始进行下一次节点选举过程。当然，在实际应用时，投票时间和计票时间的时间周期将根据实际应用的需要进行调整或设置，这里并不构成对本申请的限定。

[0078] 采用上述步骤S101～S102的方法，主机为分布式系统中各自独立的各节点配置相同的时间，使得这些节点可以在相同的时间中执行相应的投票、计票操作，也使得原本时间异步的各节点可以同步执行同一类型的操作，有效提升了对leader节点进行选举的效率。

[0079] 在本申请实施例中，在上述主机配置了相同的投票时间、计票时间以及leader节点的任期后，不同的节点将根据相应的时间配置信息执行相应的投票或计票操作，具体地，对于分布式系统中的节点而言，本申请实施例中提供了一种基于时间的节点选举方法，如图2所示，所述方法包括步骤：

[0080] S201，当所有节点中不存在leader节点时，至少一个节点作为竞选节点，在时间配置信息所携带的投票时间内，向所有节点发送投票请求。

[0081] 当分布式系统进行维护、更新等状态之后，该分布式系统中的各节点中不存在leader节点，此时，各节点将选举出leader节点。考虑到在实际应用场景中，最少在两个节点的情况下，便可以进行节点选举，在节点选举时，可由一个节点发出投票请求(投票请求发送给该节点自身以及另一个节点)，故在本申请实施例的上述步骤S201中，将有至少两个节点作为竞选节点进行投票。当然，在实际应用场景中，在有多个节点的情况下，各节点均会作为竞选节点，向其他节点发送投票请求。

[0082] 与现有技术不同的是，在本申请实施例中，竞选节点将按照主机所配置完成的投票时间进行投票，也即，竞选节点将根据主机发送的时间配置信息，对自身的内置时钟信号进行配置，当竞选节点中内置的时钟到达配置的投票时间时，则该竞选节点将执行发送投票请求的操作。

[0083] 其中，所述投票请求中包含有该竞选节点的节点标识、节点状态等节点信息，以反映出该竞选节点的数据处理能力、数据更新程度等多种节点状态，以便与其他节点进行投票选择。

[0084] S202，各竞选节点接收其他节点根据所述投票请求生成的投票信息。

[0085] 当其他节点接收到所述竞选节点发送的投票请求后，就会根据所述投票请求，生成相应的投票信息，并发送至对应的竞选节点中。

[0086] 当然，上述步骤S202中，各竞选节点接收投票信息的过程，也在投票时间内。

[0087] S203，各竞选节点在所述时间配置信息所携带的计票时间内，统计各自接收到的投票信息的数量。

[0088] 在本申请实施例中，各竞选节点将根据从主机接收到的时间配置信息，对自身内置的时钟进行配置，当竞选节点中的时钟到达配置的计票时间时，竞选节点将进行计票操作。

[0089] S204，各竞选节点根据统计出的各自的投票信息的数量确定leader节点，对指定业务进行处理。

[0090] 对于步骤S204而言，各竞选节点确定出leader节点的过程，将在计票时间结束之前完成。并且，某一竞选节点确定为leader节点后，该leader节点将会向其他节点发送通知消息，以告知其他节点leader节点已经产生。该leader节点将对指定的业务进行处理，如：对分布式系统中的元数据进行维护管理等。

[0091] 需要说明的是，在本申请实施例中，最终确定出的leader节点是唯一的，这样一来，通过一个leader节点实现对指定业务进行处理，可以增加对指定业务的调取、分配、处理的效率，同时，也可以避免因多个leader节点争抢处理指定业务，而导致业务冲突的情况。

[0092] 通过上述步骤，各竞选节点根据主机配置的相同的投票时间和计票时间，同步执行投票操作和计票操作，改善了现有技术中各节点投票、计票时机不能确定的缺陷，从而使得对节点的选举过程中，极大提升了确定leader节点的效率。

[0093] 需要说明的是，在本申请实施例中，主机所配置的投票时间和计票时间不会发生重叠，使得分布式系统中的各节点只在在投票时间内进行投票操作，并只在计票时间内进行计票操作。例如：如图3a所示的情况下，竞选节点根据时间配置信息，对该竞选自身内置的时钟配置相应的投票时间和计票时间。也即，在图3a中，在1000ms至2000ms的时间内为投票时间，在3500ms至4500ms的时间内为计票时间。也就是说，当竞选节点内置的时钟到达1000ms时，各竞选节点进入投票时间，并向其他节点发送投票信息，并且在时长为1000ms的投票时间内，接收其他节点发送的投票信息。并在时钟到达2000ms时，停止投票。同样地，在计票时间(3500ms至4500ms)内，竞选节点将统计接收到的投票信息。而在时钟到达4500ms时，各竞选节点停止计票，并根据统计出的投票信息的数量，确定出leader节点。这样的方式保证了不同的竞选节点在不同的配置时间内，同步执行不同的操作。

[0094] 而在如图3b所示的情况下，由主机配置的投票时间和计票时间之间没有时间间隔。也就是说，对于各竞选节点而言，当时钟到达2000ms时，将投票状态直接转换为计票状态。

[0095] 但是，在如图3b所示的情况下，节点之间发送投票信息需要消耗一定的时间，例如：假设某一节点将生成的投票信息发送至某竞选节点中需要耗费10ms，此时，在该投票信息传输的过程中，竞选节点的时钟已到达2000ms，也就是说，该竞选节点进入计票状态。若该竞选节点已经进入计票状态后，又接收到该投票信息，那么，在该竞选节点对投票信息进行统计时，将影响最终统计的投票信息的数量的准确性。

[0096] 因此，在本申请实施例中，在所述竞选节点在计票时间内，统计接收到的所述投票信息之间，还包括：所述竞选节点在所述计票时间到来时，停止接收所述投票信息。

[0097] 通过这样的方式，在如图3b所示的时间配置的情况下，当竞选节点的时钟到达2000ms进入计票状态时，该竞选节点将停止接收其他节点发送的投票信息。从而，这样的方式将保证对统计出的投票信息的数量的准确性。

[0098] 另外，在本申请实施例中，考虑到各竞选节点对其接收到的投票信息进行统计确定leader节点时，各竞选节点均是独立统计各自的投票信息的数量，为了保证竞选节点可以根据统计出的投票信息直接确定出自身是否成为leader节点，那么，就可以预先设置阈值，并根据该阈值和统计出的投票信息的数量，来确定竞选节点是否可以成为leader节点。

[0099] 因此，在本申请实施例中，上述步骤S204中，各竞选节点根据统计出的各自的投票信息的数量确定leader节点，具体为：当所述竞选节点判断统计出的所述投票信息的数量，超过预设的阈值时，该竞选节点将自身确定为leader节点。其中，所述阈值为所有节点数量的二分之一。

[0100] 这里需要说明的是，分布式系统中的每一节点，均已知处于该分布式系统中的所有节点的总数(作为本申请的一种方式，可由主机将分布式系统中的节点总数分别发送给各节点)，从而，当某一节点的投票信息的数量超过所有节点数据量的二分之一时，那么，就可以唯一确定出该节点为leader节点。

[0101] 例如：如图4所示，分布式系统中包含6个节点，即，节点A～F。

[0102] 在投票时间内，分布式系统中的6个节点均作为竞选节点，均向其他节点发送投票请求(发送投票请求的过程并未在图4中示出)，并接收投票信息。假设在本示例中，各节点均可成功接收到投票请求，并且，各节点不向自身发送投票请求。

[0103] 那么，对于这6个节点而言，每一个节点都会向除自身以外的各节点发送投票请求(实际应用中，节点也会向自身发出投票请求，本例中为了方便描述，只假设各节点向除自身以外的其他节点发送投票请求的情况)，具体来说，对于节点A而言，节点A向另外5个节点(节点B～节点F)发送投票请求。同样地，节点B～节点G也向除自身之外的5个节点发送投票请求。使得接收到投票请求的各节点生成投票信息发送至相应的竞选节

[0104] 从而在本示例中，对于各节点而言，均可确定参与投票的节点的数量为6。且设置阈值为3。

[0105] 假设如图4所示，节点A将生成的投票信息发送给节点B。节点B将生成的投票信息发送给节点C。节点C、节点D、节点E、节点F均将各自生成的投票信息发送给节点A。

[0106] 在计票时间内，各节点将统计其自身所接收到的投票信息的数量，具体地，节点A统计出其接收到的投票信息的数量为4，节点B统计出其接收到的投票信息的数量为1，节点C统计出其接收到的投票信息的数量为1，其余节点中的投票信息的数量为0。

[0107] 节点A中的投票信息的数量大于预设的阈值，因此，节点A可以确定其自身为leader节点。之后，该节点A向其他节点发送广播，以通知其他节点该节点A已成为leader节点。

[0108] 但是，在实际应用中，由于节点选举的过程是各节点“自主”进行投票，就有可能出现各竞选节点所接收到的投票信息的数量，均未超过预设阈值的情况。沿用上例：假设节点F所生成的投票信息并未发送给节点A，而是发送给节点C，这样一来，节点A中的投票信息的数量为3，节点B中的投票信息的数量为1，节点C中的投票信息的数量为2，显然，这三个节点中的投票信息的数量均为超过预设阈值(也即，3)。从而，不会有节点将自身确定为leader节点，也就是说，各节点在指定的一段时长内，均未接收到leader节点的通知消息，这就表明本次节点选举并未能选举出leader节点。

[0109] 在本申请实施例中，一旦出现未能选举出leader节点的情况，各节点会立即重新

进行节点选举的过程，直到确定出leader节点。

[0110] 结合上述示例，作为本申请实施例的一种优选方式，主机为各节点配置的投票时间和计票时间具有周期性，并且计票时间可以更具体地划分为不同的子时间。如图3c所示，针对一个周期的投票时间和计票时间而言，计票时间具体分为统计时间和通知时间。那么，各节点在统计时间内统计各自接收的投票信息的数量，若在通知时间内，各节点未接收到任何leader节点的通知，则会进行下一次的节点选举。

[0111] 如图3d所示，本申请实施例中由主机分配的多次选举时间(每一次选举时间包括投票时间和计票时间)具有周期性。当在第一次选举时间内，各节点未确定出leader节点的情况下，将在第二次选举时间内再次进行节点选举的过程，直到选举出leader节点。当然，选举出leader节点后，则会按照时间配置信息中的leader节点的任期，周期性的进行节点选举的过程，其具体过程将在后续进行详细描述，在此不作过多赘述。

[0112] 上述如图2所示的方法S201～S204，是基于竞选节点向其他节点发送投票请求，并接收投票信息以确定leader节点的过程。对于分布式系统中的各节点，也可作为投票节点，接收竞选节点发出的投票请求，并根据投票请求生成相应的投票信息(即使对于竞选节点而言，也会接收到其他竞选节点发送的投票请求，此时，该竞选节点也可以看作是投票节点)。从而，生成投票信息的节点而言，本申请实施例还提供一种基于时间的节点选举方法，如图5所示，该方法用于分布式系统中确定leader节点，该方法具体包括以下步骤：

[0113] S501，当所有节点中不存在leader节点时，至少两个节点作为投票节点，在时间配置信息所携带的投票时间内，接收竞选节点发送的投票请求。

[0114] 在本申请实施例中，在未确定出leader节点时，分布式系统内的各节点均可以作为投票节点，以接收竞选节点发送的投票请求。

[0115] 同样地，所述投票节点也将接收主机发送的时间配置信息，调整该投票节点内部的时钟，配置投票时间和计票时间。

[0116] S502，所述投票节点根据所述投票请求，在投票时间内生成投票信息发送至竞选节点，使得所述竞选节点在计票时间内统计所述投票信息的数量，确定leader节点。

[0117] 在分布式系统中，投票节点将接收至少一个竞选节点发送的投票请求(在实际应用场景下，投票节点通常会接收到多个竞选节点发出的投票请求)，而由于最终只能选举出一个leader节点，从而，投票节点将选定一个竞选节点进行投票，也即，投票节点将根据其接收到的投票请求，选定一个竞选节点，并生成针对选定的节点的生成投票信息并发送至选定的节点中。

[0118] 通过上述步骤，正是由于分布式系统中的各投票节点配置了相同的投票时间内，使得各投票节点可以在相同的投票时间内，选定出竞选节点，并生成相应的投票请求发送至选定的竞选节点中，从而保证了投票过程的同步性，也提升了对节点进行选举的效率。

[0119] 需要说明的是，当竞选节点向投票节点发送投票请求时，该投票请求中通常会携带该竞选节点的节点信息。这样一来，在投票节点接收了多个投票请求的情况下，投票节点就可以根据投票请求中所携带的节点信息，选定竞选节点。

[0120] 故在本申请实施例中，生成投票信息发送至竞选节点，具体为：所述投票节点确定接收到的所述投票请求中所携带的节点信息，根据所述节点信息，确定所述竞选节点的节点状态权重，选定节点状态权重最大的竞选节点，并生成针对该竞选节点的投票信息，发送

至选定的竞选节点。

[0121] 其中，所述节点信息，包括但不限于竞选节点的数据更新信息、处理负载信息、硬件配置信息等反映竞选节点数据处理能力、处理状态的信息。从而，投票节点就可以根据节点信息，确定出不同竞选节点的节点状态权重，用以反映不同竞选节点的综合状态水平。从而，在本申请实施例中，投票节点会选定节点状态权重最高的竞选节点，并生成相应的投票信息发送至该竞选节点。完成投票过程。

[0122] 以上如图2所示的方法S201～S204，以及如图5所示的方法S501～S502，均是在分布式系统中还不存在leader节点的情况下，各节点根据主机配置的投票时间和计票时间，确定leader节点的过程。而对于分布式系统中已经存在leader节点的情况，本申请实施例还提供一种基于时间的节点选举方法，如图6所示，所述方法用于分布式系统中确定leader节点，具体包括以下步骤：

[0123] S601，当所有节点中包含leader节点，且所述leader节点确定自身符合预设标准时，所述leader节点根据时间配置信息中的leader节点的任期，在投票时间内向所有节点发送连任请求。

[0124] 在本申请实施例中，通过如图2所示的方法S201～S204，以及如图5所示的方法S501～S502，确定出leader节点之后，当该leader的任期即将结束前，就需要进行下一次节点选举，此时，除leader节点之外的其他节点不会发送投票请求，而只有leader节点可以向其他节点发送连任请求。

[0125] 需要说明的是，由于在实际应用，分布式系统所处理的信息量巨大，就可能会出现leader节点所维护的数据(如：元数据)更新滞后，或者，leader节点的处理负荷不断增加，导致其处理业务的效率下降等情况。从而，在本申请实施例中，当下一次投票时间到来时，当前的leader节点会判断自身的节点状态是否符合预设标准，如果符合，则说明该leader节点的处理负荷、数据更新程度不会影响对指定业务处理的效率，因此，该leader节点才会向其他节点发送连任请求。

[0126] 其中，所述预设标准，包括但不限于：处理负载标准、数据更新程度等。由于不同节点的硬件配置(处理性能、存储空间等)可能不同，故在本申请实施例中，各节点的预设标准也可能不相同，那么，各节点会根据其自身的硬件配置生成相应的预设标准。

[0127] 通常，各节点的预设标准，会以标准值的方式预先设定。如：对于某一节点而言，处理负载标准为：中央处理器(Central Processing Unit，CPU)的使用率不超过80％。也就是说，当CPU的使用了未超过80％时，则可认为该节点符合预设标准。

[0128] 又如：对于某一节点而言，数据更新程度标准为：数据更新频率超过10min/次(数据更新频率可以反映出数据更新的程度)。也就是说，当节点的数据更新频率超过10min/次时，则可认为该节点符合预设标准。

[0129] 当然，上述内容仅是本申请实施例提供的预设标准的示例，在实际应用场景下，各节点的预设标准将根据实际应用的需要进行设置，这里并不构成对本申请的限定。

[0130] S602，所述leader节点接收其他节点根据所述连任请求生成的投票信息。

[0131] 接收到连任请求的节点，会根据该连任请求生成对应的投票信息，发送至leader节点中。

[0132] S603，所述leader节点在计票时间内，统计接收到的所述投票信息的数量。

[0133] S604,所述leader节点根据统计出的所述投票信息的数量确定连任状态,并在成功连任后对指定业务进行处理。

[0134] 在本申请实施例中,由于在leader节点向其他节点发送连任请求的情况下,分布式系统中没有其他竞选节点,也就是说,各节点只会接收到来自leader节点的连接请求,故leader节点通常会成功连任。

[0135] 在leader节点成功连任之后,该leader节点会向各节点发送通知消息,通知各节点该leader节点已成功连任,从而,成功连任的leader节点将继续处理指定业务,分布式系统中的其他节点将继续正常处理业务。

[0136] 在上述如图6所示的方法S601～S604的情况下,本申请实施例另提供一种基于时间的节点选举方法,如图7所示,所述方法用于分布式系统中确定leader节点,具体包括以下步骤:

[0137] S701,当所有节点中存在leader节点时,至少两个节点作为投票节点,在时间配置信息中所携带的投票时间内,接收leader节点发送的连任请求。

[0138] S702,所述投票节点根据所述连任请求,在时间配置信息中所携带的投票时间内,生成投票信息发送至所述leader节点,使得所述leader节点在计票时间内统计所述投票信息的数量确定连任状态,并且,所述投票节点在所述leader节点的任期内只接收所述leader节点发送的投票请求。

[0139] 通过上述如图6所示的方法S601～S604,以及如图7所示的方法S701～S702,在分布式系统中已经存在leader节点的情况下,当下一次投票时间到来时,分布式系统中的所有节点均会同步进入投票状态,并只由leader节点向其他节点发送连任请求,接收到连任请求的节点会将生成的投票请求发送至leader节点,这样一来,其他节点便不会互相发送投票请求重新确定leader节点,从而有效降低了各节点的处理负荷,也提升了对leader节点的选举效率。并且,任一投票节点在leader节点的任期内,均不会接收其他节点发送的投票请求,这样,就可以保证各投票节点只接收leader节点发出的投票请求。

[0140] 当然,在实际应用中,分布式系统中的各节点处理大量的业务,leader节点就有可能出现处理负载增加、数据更新滞后等情况,使得该leader节点不符合预设的标准,那么,在这样的情况下,本申请实施例提供一种基于时间的节点选举方法,如图8所示,所述方法用于分布式系统中确定leader节点,具体包括以下步骤:

[0141] S801,当所有节点中包含leader节点,且所述leader节点确定自身不符合预设标准时,所述leader节点根据时间配置信息中的leader节点的任期,在其他节点中筛选出符合所述预设标准的节点,作为继承节点。

[0142] leader节点确定自身不符合预设标准,就表明该leader节点可能出现处理负责过高,或者数据更新程度过低,在这样的情况下,该leader节点处理指定业务的效率将收到严重影响。也即,当前该leader节点已不适合再处理指定业务。

[0143] 从而,在本申请实施例中,当leader节点的任期结束前,所述leader节点将从其他节点中筛选出继承节点。作为本申请实施例中的一种方式,可由leader节点向其他节点发送查询请求,以查询其他节点的节点信息,根据节点信息选定继承节点。作为本申请实施例中的另一种方式,也可以由主机指定该leader节点选定继承节点,也即,进行leader节点的改选,那么,当前的leader节点接收到改选指令后,就会从其他节点中选定一个节点作为该

leader节点的继承节点。当然，上述方式并不作为对本申请的限定。

[0144] S802，所述leader节点在所述投票时段内，向其他节点发出针对该继承节点的投票请求，使得所述继承节点接收其他节点发送的投票信息，并在计票时间内统计所述投票信息的数量确定为新任leader节点。

[0145] 在投票时间内，leader节点将向其他节点发送的投票请求中，携带有继承节点的节点标识，也就是说，接收到该投票请求的节点会生成相应的投票信息，并根据继承节点的节点标识，将投票信息发送至该继承节点中。这样一来，各节点的投票信息只发送给继承节点，而不会发送给当前的leader节点。

[0146] 显然，只有继承节点接收到各节点发送的投票信息，故在计票时间内，继承节点将统计接收到的投票信息，确定为leader节点。并向各节点发送通知消息，以通知各节点，该继承节点已成为新任leader节点。

[0147] 在上述如图8所示的方法S801～S802的情况下，本申请实施例另提供一种基于时间的节点选举方法，如图9所示，所述方法用于分布式系统中确定leader节点，具体包括以下步骤：

[0148] S901，当所有节点中存在leader节点时，至少两个节点作为投票节点，在时间配置信息中携带的投票时间内，接收所述leader节点发送的针对继承节点的投票请求。

[0149] S902，所述投票节点根据所述投票请求，在时间配置信息中所携带的投票时间内，生成投票信息发送至所述继承节点，使得所述继承节点在计票时间内统计所述投票信息的数量，确定新任leader节点。

[0150] 同样，对于上述步骤S901～S902，由于leader节点发送的投票请求中，携带有继承节点的节点标识，所以，投票节点会根据投票请求生成相应的投票信息，并根据继承节点的节点标识，将投票信息发送至该继承节点中，而不会发送到其他节点中。这就保证了继承节点可以顺利成为新任的leader节点。

[0151] 通过以上方法和实例，在分布式系统中不存在leader节点的情况下，分布式系统的主机为各节点配置相同的投票时间和计票时间，使得各节点在投票时间内同步进行投票，并且在计票时间内同步进行计票。这样一来，就不会出现各节点进行投票或计票的时机不确定的情况，从而极大提升了分布式系统中各节点选举leader节点的效率。

[0152] 并且，在选举出leader节点之后，各节点仍根据周期性的投票时间和计票时间，在下一次投票时间和计票时间到来时，leader节点会判断自身的状态，以生成针对自身的投票请求或者是针对继承节点的投票请求，使得其他节点根据该投票请求生成投票信息发送至对应节点中。这样的方式保证选举出的用于对指定业务进行处理的leader节点，始终处于高性能状态，可提升分布式系统整体对业务处理的效率。

[0153] 以上为本申请实施例提供的基于时间的节点选举方法，基于同样的思路，本申请实施例还提供一种基于时间的节点选举装置，如图10所示。

[0154] 在图10中，所述基于时间的节点选举装置包括：配置模块1001和发送模块1002，其中，

[0155] 所述配置模块1001，用于为所有节点配置相同的投票时间、计票时间以及leader节点的任期。

[0156] 所述发送模块1002，用于将配置的所述投票时间、计票时间以及leader节点的任

期携带在时间配置信息中发送至各节点，使得各节点在所述投票时间内进行投票，在所述计票时间内进行计票，并根据计票结果选举出leader节点，并使得各节点根据所述leader节点的任期进行周期性节点选举。

[0157] 在如图10所示的情况下，本申请实施例还提供一种基于时间的节点选举装置，如图11所示，所述装置用于分布式系统中确定leader节点，设置在竞选节点中。

[0158] 在图11中，所述基于时间的节点选举装置包括：投票请求模块1101、接收模块1102、统计模块1103和确定模块1104，其中，当所有节点中不存在leader节点时，

[0159] 所述投票请求模块1101，用于在时间配置信息所携带的投票时间内，向所有节点发送投票请求。

[0160] 所述接收模块1102，用于接收其他节点根据所述投票请求生成的投票信息。

[0161] 所述统计模块1103，用于在所述时间配置信息所携带的计票时间内，统计各自接收到的投票信息的数量。

[0162] 所述确定模块1104，用于根据统计出的各自的投票信息的数量确定leader节点，对指定业务进行处理。

[0163] 在本申请实施例中，所述接收模块1102，还用于在所述计票时间到来时，停止接收所述投票信息。

[0164] 所述确定模块1104，具体用于当判断统计出的所述投票信息的数量，超过预设的阈值时，将所述竞选节点确定为leader节点。

[0165] 其中，所述阈值为所有节点数量的二分之一。

[0166] 本申请实施例还提供一种基于时间的节点选举装置，如图12所示，所述装置用于分布式系统中确定leader节点，设置在投票节点中。

[0167] 在图12中，所述基于时间的节点选举装置包括：接收模块1201和投票信息模块1202，其中，当所有节点中不包含有leader节点时，

[0168] 所述接收模块1201，用于在时间配置信息所携带的投票时间内，接收竞选节点发送的投票请求。

[0169] 所述投票信息模块1202，用于根据所述投票请求，在投票时间内生成投票信息发送至竞选节点，使得所述竞选节点在计票时间内统计所述投票信息的数量，确定leader节点。

[0170] 在本申请实施例中，所述投票信息模块1202，具体用于确定接收到的所述投票请求中所携带的节点信息，根据所述节点信息，确定所述竞选节点的节点状态权重，选定节点状态权重最大的竞选节点，并生成针对该竞选节点的投票信息，发送至选定的竞选节点。

[0171] 本申请实施例还提供一种基于时间的节点选举装置，如图13所示，所述装置用于分布式系统中确定leader节点，设置在leader节点中。

[0172] 在图13中，所述基于时间的节点选举装置包括：连任请求模块1301、接收模块1302、统计模块1303以及确定模块1304，其中，当所有节点中包含leader节点，且所述leader节点确定自身符合预设标准时，

[0173] 所述连任请求模块1301，用于根据时间配置信息中的leader节点的任期，在投票时间内向所有节点发送连任请求。

[0174] 所述接收模块1302，用于接收其他节点根据所述连任请求生成的投票信息。

[0175] 所述统计模块1303，用于在计票时间内，统计接收到的所述投票信息的数量。

[0176] 所述确定模块1304，根据统计出的所述投票信息的数量确定连任状态，并在成功连任后对指定业务进行处理。

[0177] 其中，所述预设标准，包括处理负载标准、数据更新程度中的至少一种。

[0178] 本申请实施例还提供一种基于时间的节点选举装置，如图14所示，所述装置用于分布式系统中确定leader节点，设置在leader节点中。

[0179] 在图14中，所述基于时间的节点选举装置包括：筛选模块1401和投票请求模块1402，其中，当所有节点中包含leader节点，且所述leader节点确定自身不符合预设标准时，

[0180] 所述筛选模块1401，用于根据时间配置信息中的leader节点的任期，在其他节点中筛选出符合所述预设标准的节点，作为继承节点。

[0181] 所述投票请求模块1402，用于在所述投票时段内，向其他节点发出针对该继承节点的投票请求，使得所述继承节点接收其他节点发送的投票信息，并在计票时间内统计所述投票信息的数量确定为新任leader节点。

[0182] 本申请实施例还提供一种基于时间的节点选举装置，如图15所示，所述装置用于分布式系统中确定leader节点，设置在投票节点中。

[0183] 在图15中，所述基于时间的节点选举装置包括：接收模块1501和投票信息模块1502，其中，当所有节点中包含leader节点时，

[0184] 所述接收模块1501，用于在时间配置信息中所携带的投票时间内，接收所述leader节点发送的连任请求。

[0185] 所述投票信息模块1502，用于根据所述连任请求，在时间配置信息中所携带的投票时间内，生成投票信息发送至所述leader节点，使得所述leader节点在计票时间内统计所述投票信息的数量确定连任状态，并用于在所述leader节点的任期内只接收所述leader节点发送的投票请求。

[0186] 本申请实施例还提供一种基于时间的节点选举装置，如图16所示，所述装置用于分布式系统中确定leader节点，设置在投票节点中。

[0187] 在图16中，所述基于时间的节点选举装置包括：接收模块1601和投票信息模块1602，其中，当所有节点中存在leader节点时，

[0188] 所述接收模块1601，用于在时间配置信息中携带的投票时间内，接收所述leader节点发送的针对继承节点的投票请求。

[0189] 所述投票信息模块1602，用于根据所述投票请求，在时间配置信息中所携带的投票时间内，生成投票信息发送至所述继承节点，使得所述继承节点在计票时间内统计所述投票信息的数量，确定新任leader节点。

[0190] 在一个典型的配置中，计算设备包括一个或多个处理器(CPU)、输入/输出接口、网络接口和内存。

[0191] 内存可能包括计算机可读介质中的非永久性存储器，随机存取存储器(RAM)和/或非易失性内存等形式，如只读存储器(ROM)或闪存(flash RAM)。内存是计算机可读介质的示例。

[0192] 计算机可读介质包括永久性和非永久性、可移动和非可移动媒体可以由任何方法

或技术来实现信息存储。信息可以是计算机可读指令、数据结构、程序的模块或其他数据。计算机的存储介质的例子包括，但不限于相变内存(PRAM)、静态随机存取存储器(SRAM)、动态随机存取存储器(DRAM)、其他类型的随机存取存储器(RAM)、只读存储器(ROM)、电可擦除可编程只读存储器(EEPROM)、快闪记忆体或其他内存技术、只读光盘只读存储器(CD-ROM)、数字多功能光盘(DVD)或其他光学存储、磁盒式磁带，磁带磁磁盘存储或其他磁性存储设备或任何其他非传输介质，可用于存储可以被计算设备访问的信息。按照本文中的界定，计算机可读介质不包括暂存电脑可读媒体(transitory media)，如调制的数据信号和载波。

[0193] 还需要说明的是，术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含，从而使得包括一系列要素的过程、方法、商品或者设备不仅包括那些要素，而且还包括没有明确列出的其他要素，或者是还包括为这种过程、方法、商品或者设备所固有的要素。在没有更多限制的情况下，由语句“包括一个.....”限定的要素，并不排除在包括所述要素的过程、方法、商品或者设备中还存在另外的相同要素。

[0194] 本领域技术人员应明白，本申请的实施例可提供为方法、系统或计算机程序产品。因此，本申请可采用完全硬件实施例、完全软件实施例或结合软件和硬件方面的实施例的形式。而且，本申请可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质(包括但不限于磁盘存储器、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。

[0195] 以上所述仅为本申请的实施例而已，并不用于限制本申请。对于本领域技术人员来说，本申请可以有各种更改和变化。凡在本申请的精神和原理之内所作的任何修改、等同替换、改进等，均应包含在本申请的权利要求范围之内。

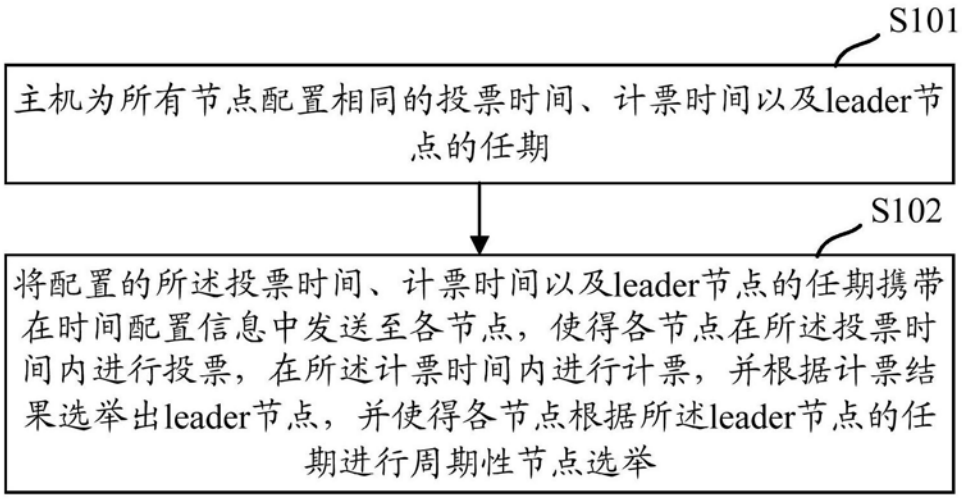

S101
主机为所有节点配置相同的投票时间、计票时间以及leader节点的任期
S102
将配置的所述投票时间、计票时间以及leader节点的任期携带在时间配置信息中发送至各节点，使得各节点在所述投票时间内进行投票，在所述计票时间内进行计票，并根据计票结果选举出leader节点，并使得各节点根据所述leader节点的任期进行周期性节点选举

图1

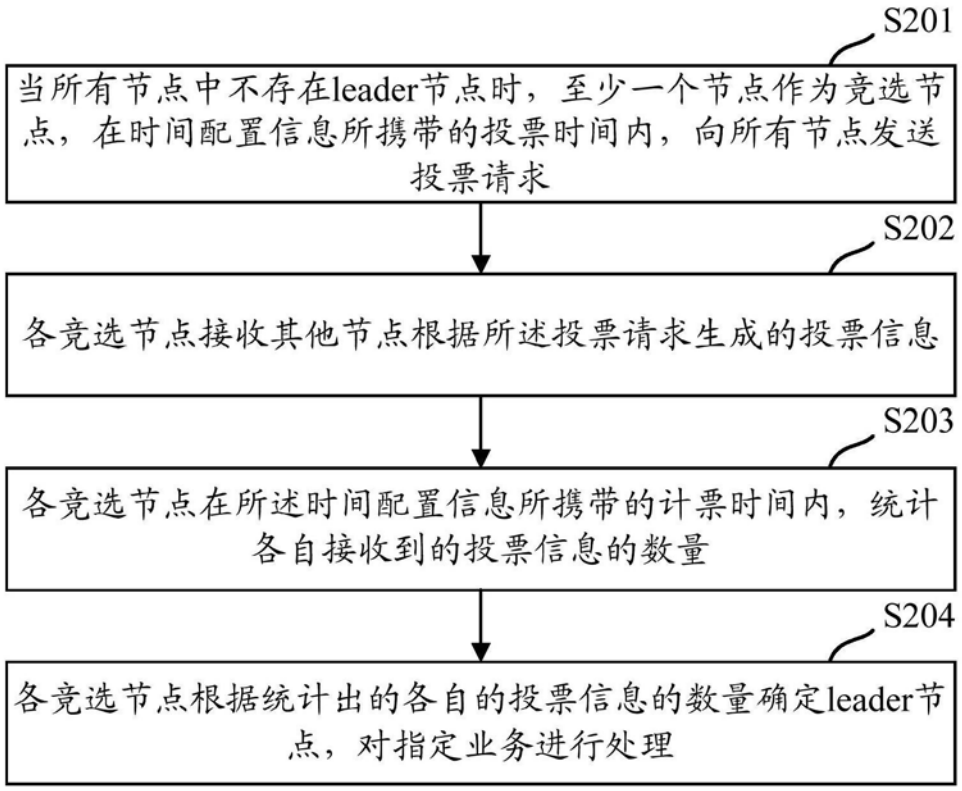

S201
当所有节点中不存在leader节点时，至少一个节点作为竞选节点，在时间配置信息所携带的投票时间内，向所有节点发送投票请求
S202
各竞选节点接收其他节点根据所述投票请求生成的投票信息
S203
各竞选节点在所述时间配置信息所携带的计票时间内，统计各自接收到的投票信息的数量
S204
各竞选节点根据统计出的各自的投票信息的数量确定leader节点，对指定业务进行处理

图2

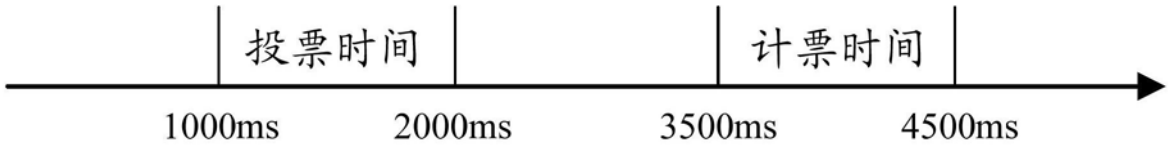

投票时间
计票时间
1000ms
2000ms
3500ms
4500ms

图3a

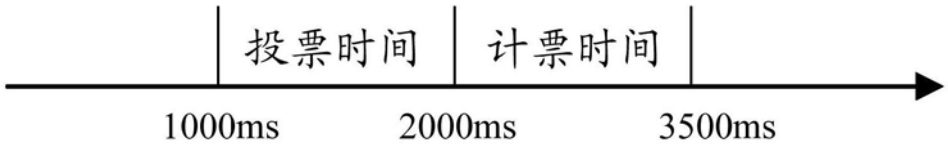


图3b

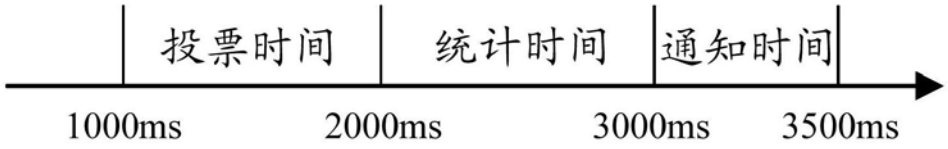


图3c

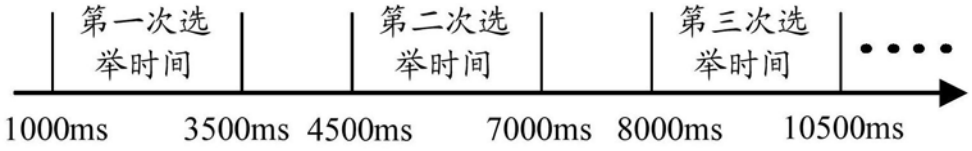


图3d

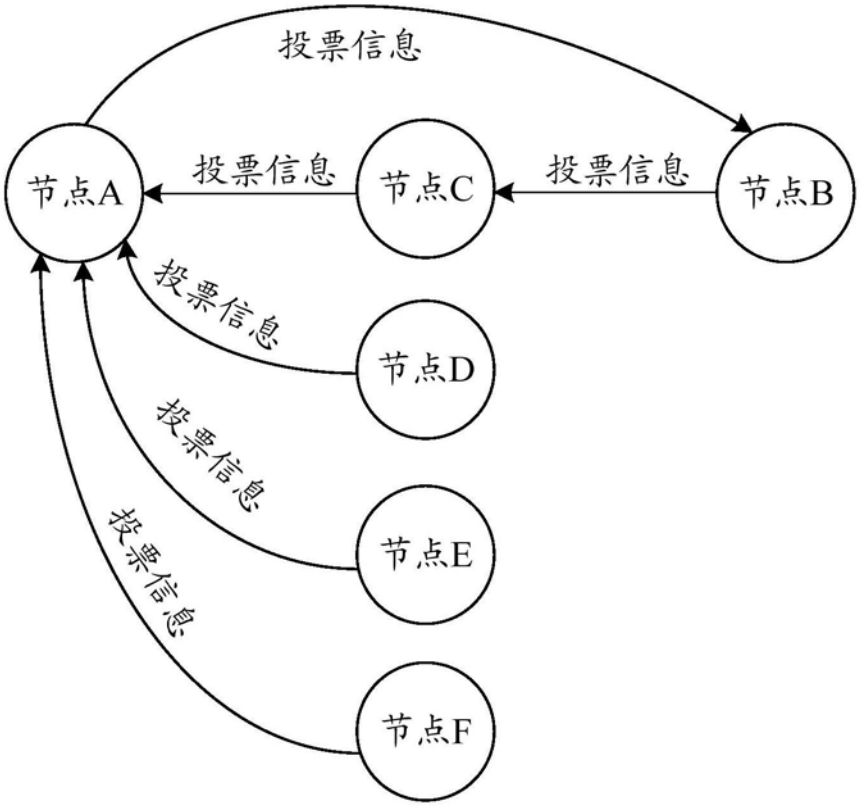


图4

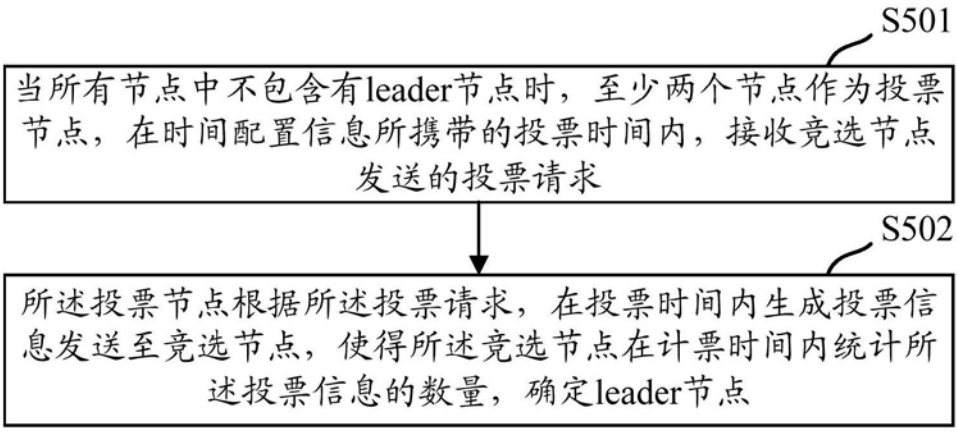

S501
当所有节点中不包含有leader节点时，至少两个节点作为投票节点，在时间配置信息所携带的投票时间内，接收竞选节点发送的投票请求
S502
所述投票节点根据所述投票请求，在投票时间内生成投票信息发送至竞选节点，使得所述竞选节点在计票时间内统计所述投票信息的数量，确定leader节点

图5

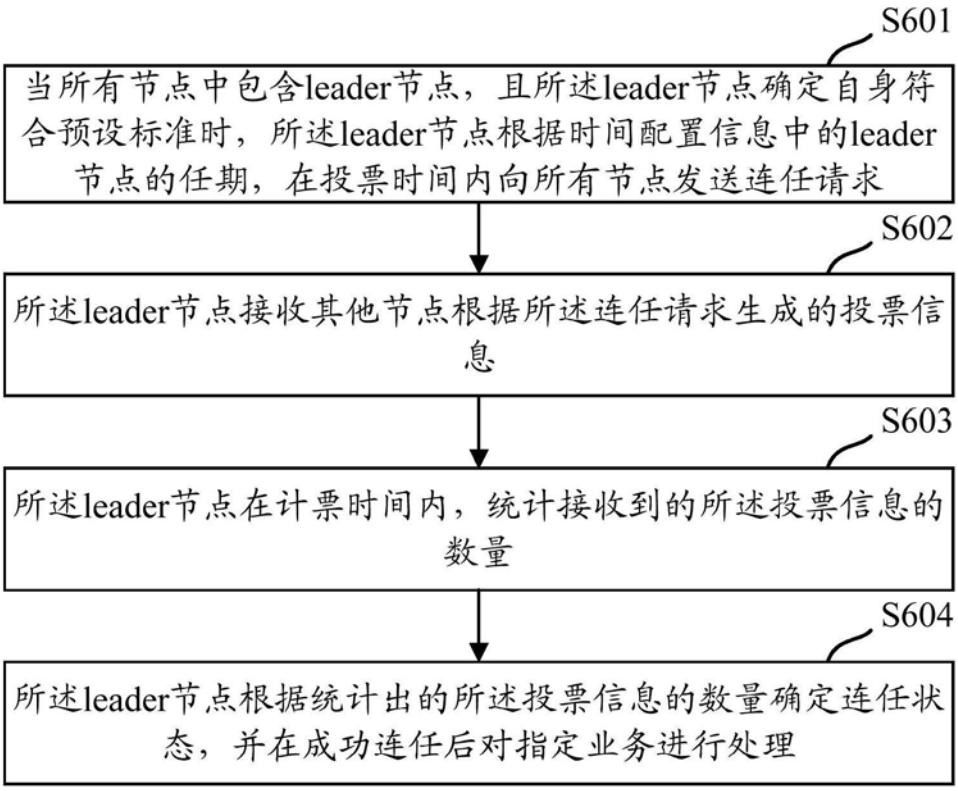

S601
当所有节点中包含leader节点，且所述leader节点确定自身符合预设标准时，所述leader节点根据时间配置信息中的leader节点的任期，在投票时间内向所有节点发送连任请求
S602
所述leader节点接收其他节点根据所述连任请求生成的投票信息
S603
所述leader节点在计票时间内，统计接收到的所述投票信息的数量
S604
所述leader节点根据统计出的所述投票信息的数量确定连任状态，并在成功连任后对指定业务进行处理

图6

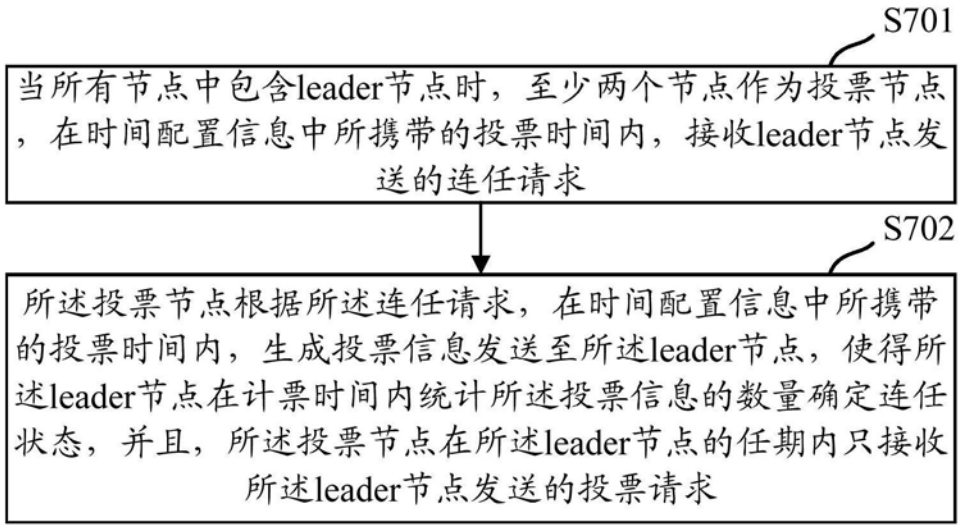

S701
当所有节点中包含leader节点时，至少两个节点作为投票节点，在时间配置信息中所携带的投票时间内，接收leader节点发送的连任请求
S702
所述投票节点根据所述连任请求，在时间配置信息中所携带的投票时间内，生成投票信息发送至所述leader节点，使得所述leader节点在计票时间内统计所述投票信息的数量确定连任状态，并且，所述投票节点在所述leader节点的任期内只接收所述leader节点发送的投票请求

图7

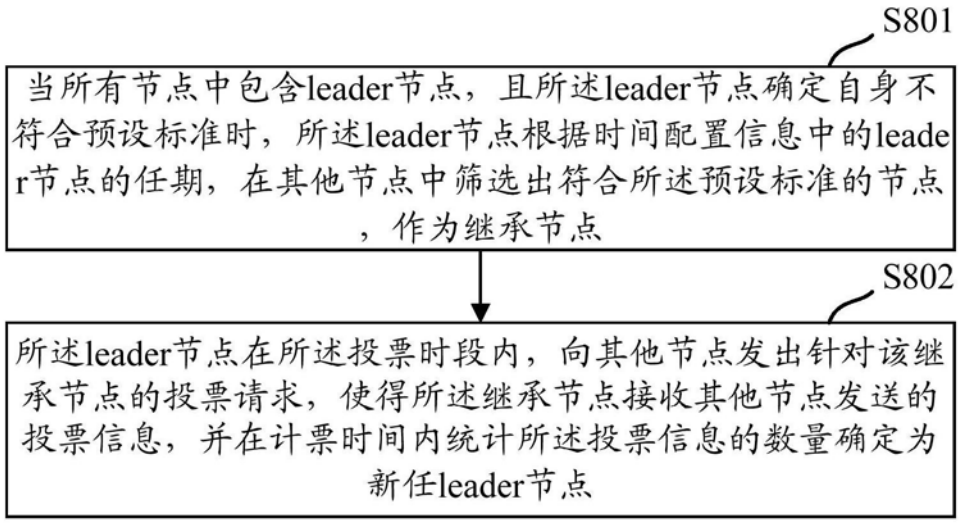

S801
当所有节点中包含leader节点，且所述leader节点确定自身不符合预设标准时，所述leader节点根据时间配置信息中的leader节点的任期，在其他节点中筛选出符合所述预设标准的节点，作为继承节点
S802
所述leader节点在所述投票时段内，向其他节点发出针对该继承节点的投票请求，使得所述继承节点接收其他节点发送的投票信息，并在计票时间内统计所述投票信息的数量确定为新任leader节点

图8

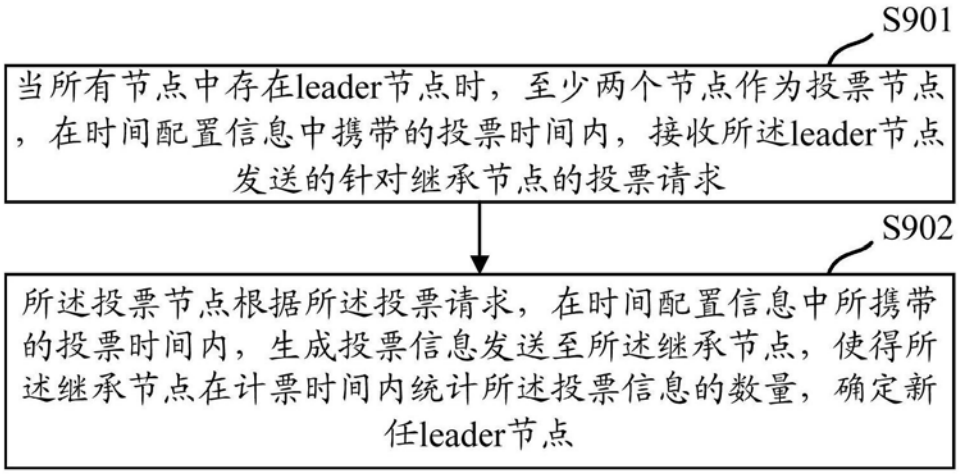

S901
当所有节点中存在leader节点时，至少两个节点作为投票节点，在时间配置信息中携带的投票时间内，接收所述leader节点发送的针对继承节点的投票请求
S902
所述投票节点根据所述投票请求，在时间配置信息中所携带的投票时间内，生成投票信息发送至所述继承节点，使得所述继承节点在计票时间内统计所述投票信息的数量，确定新任leader节点

图9

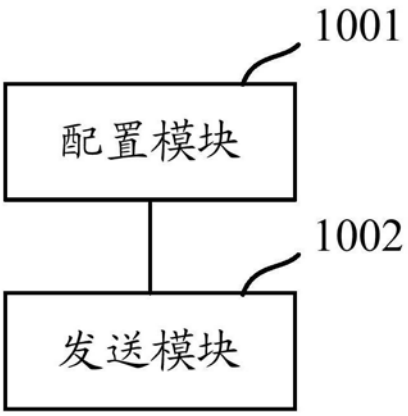

1001
配置模块
1002
发送模块

图10

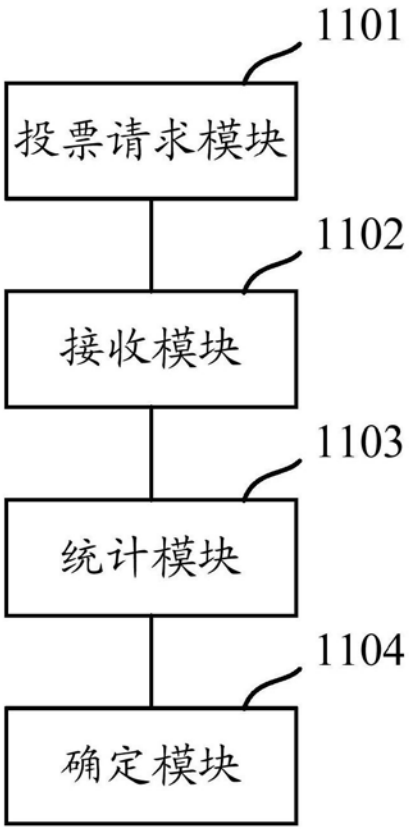

1101
投票请求模块
1102
接收模块
1103
统计模块
1104
确定模块

图11

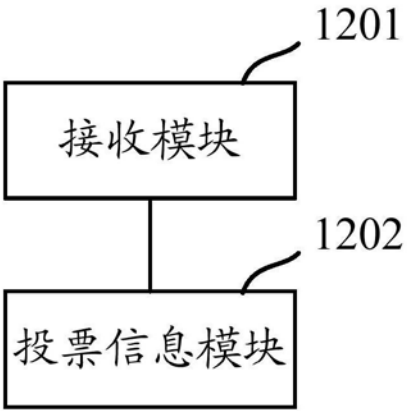

1201
接收模块
1202
投票信息模块

图12

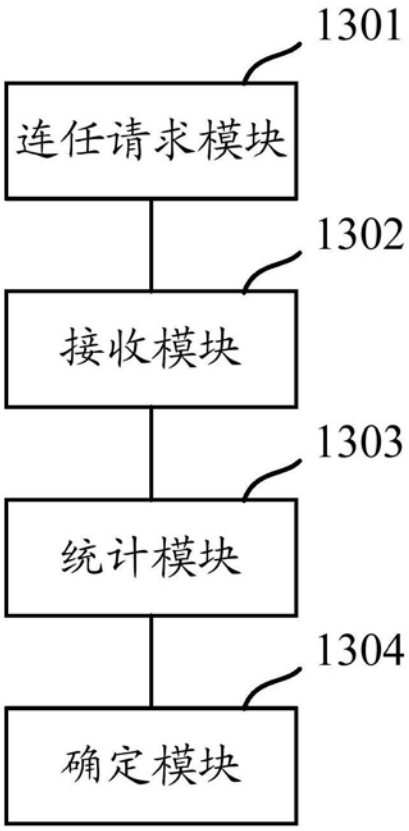

1301
连任请求模块
1302
接收模块
1303
统计模块
1304
确定模块

图13

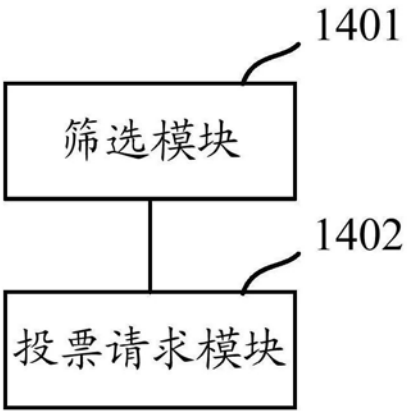

1401
筛选模块
1402
投票请求模块

图14

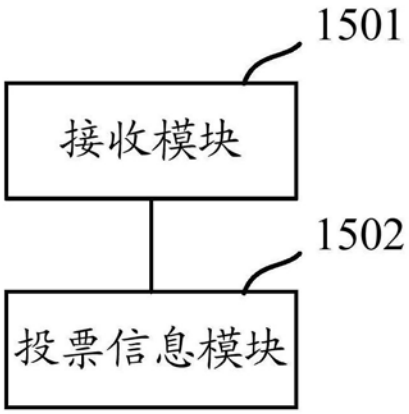


图15

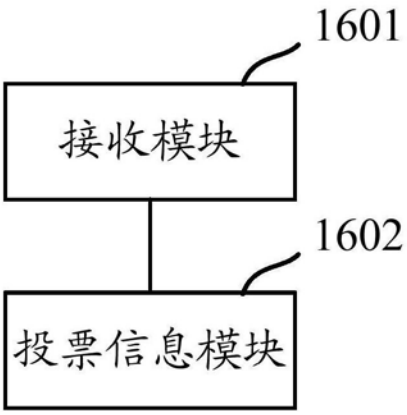


图16