



(12) 发明专利

(10) 授权公告号 CN 112073250 B

(45) 授权公告日 2022. 03. 22

(21) 申请号 202010981645.3

(22) 申请日 2020.09.17

(65) 同一申请的已公布的文献号  
申请公布号 CN 112073250 A

(43) 申请公布日 2020.12.11

(73) 专利权人 新华三信息安全技术有限公司  
地址 230001 安徽省合肥市高新区创新大道2800号创新产业园二期H2栋541室

(72) 发明人 陈南飞

(74) 专利代理机构 北京柏杉松知识产权代理事务所(普通合伙) 11413  
代理人 丁芸 马敬

(51) Int. Cl.

H04L 41/0663 (2022.01)

H04L 67/10 (2022.01)

(56) 对比文件

US 2020028750 A1, 2020.01.23

CN 106487611 A, 2017.03.08

US 2020042410 A1, 2020.02.06

CN 109728941 A, 2019.05.07

CN 101217402 A, 2008.07.09

US 2017344444 A1, 2017.11.30

审查员 刘媛

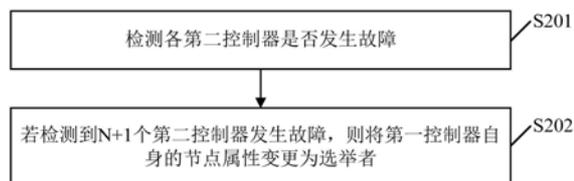
权利要求书2页 说明书10页 附图4页

(54) 发明名称

控制器集群故障处理方法、装置、控制器及控制器集群

(57) 摘要

本发明实施例提供了一种控制器集群故障处理方法、装置、控制器及控制器集群,控制器集群中包括:第一控制器和2N+1个第二控制器;第一控制器的节点属性为学习者,禁止该控制器参与主控制器的选举和投票;第二控制器的节点属性为选举者,允许该控制器参与主控制器的选举和投票。第一控制器检测各第二控制器是否发生故障;如果第一控制器检测到N+1个第二控制器发生故障,则将第一控制器自身的节点属性变更为选举者。在有N+1个第二控制器发生故障的情况下,第一控制器将自身的节点属性变更为选举者,此时,有N+1个控制器的节点属性为选举者,超过半数,实现了在有控制器故障的情况下,控制器集群正常提供服务的目的。



1. 一种控制器集群故障处理方法,其特征在于,应用于控制器集群中的第一控制器;所述控制器集群包括:所述第一控制器和 $2N+1$ 个第二控制器;所述第一控制器的节点属性为学习者,所述第二控制器的节点属性为选举者;所述 $N$ 为正整数;若控制器的节点属性为学习者,则表征禁止该控制器参与主控制器的选举和投票;若控制器的节点属性为选举者,则表征允许该控制器参与主控制器的选举和投票;

所述方法包括:

检测各第二控制器是否发生故障;

若检测到 $N+1$ 个第二控制器发生故障,则将所述第一控制器自身的节点属性变更为选举者。

2. 根据权利要求1所述的方法,其特征在于,各第二控制器记录有:经检测得到的其他第二控制器是否可达的检测信息;

所述检测各第二控制器是否发生故障的步骤,包括:

针对任一第二控制器,若在预设周期内,未收到所述任一第二控制器发送的心跳消息,则从各指定第二控制器中获取针对所述任一第二控制器的检测信息,其中,所述指定第二控制器为经检测确定未发生故障的第二控制器;

若获取到的所有检测信息中存在所述任一第二控制器不可达,则确定所述任一第二控制器发生故障。

3. 根据权利要求2所述的方法,其特征在于,所述控制器集群中包括第一站点和第二站点;所述第一站点中设置有 $N+1$ 个第二控制器,所述第二站点中设置有 $N$ 个第二控制器和所述第一控制器;所述指定第二控制器为所述第二站点中的 $N$ 个第二控制器。

4. 根据权利要求1所述的方法,其特征在于,所述第一控制器存储有集群成员列表,所述集群成员列表用于记录所述控制器集群中每个控制器的控制器标识和每个控制器的节点属性;

在所述将所述第一控制器自身的节点属性变更为选举者的步骤之后,所述方法还包括:

将所述集群成员列表中所述第一控制器的节点属性更新为选举者,并将任一发生故障的第二控制器的节点属性更新为学习者,得到更新后的集群成员列表;

将所述更新后的集群成员列表发送至各第二控制器,以使所述各第二控制器根据所述更新后的集群成员列表设置自身的节点属性。

5. 一种控制器集群故障处理装置,其特征在于,应用于控制器集群中的第一控制器;所述控制器集群包括:所述第一控制器和 $2N+1$ 个第二控制器;所述第一控制器的节点属性为学习者,所述第二控制器的节点属性为选举者;所述 $N$ 为正整数;若控制器的节点属性为学习者,则表征禁止该控制器参与主控制器的选举和投票;若控制器的节点属性为选举者,则表征允许该控制器参与主控制器的选举和投票;

所述装置包括:

检测模块,用于检测各第二控制器是否发生故障;

设置模块,用于若检测到 $N+1$ 个第二控制器发生故障,则将所述第一控制器自身的节点属性变更为选举者。

6. 根据权利要求5所述的装置,其特征在于,各第二控制器记录有:经检测得到的其他

第二控制器是否可达的检测信息；

所述检测模块,具体用于:针对任一第二控制器,若在预设周期内,未收到所述任一第二控制器发送的心跳消息,则从各指定第二控制器中获取针对所述任一第二控制器的检测信息,其中,所述指定第二控制器为经检测确定未发生故障的第二控制器;若获取到的所有检测信息中存在所述任一第二控制器不可达,则确定所述任一第二控制器发生故障。

7. 根据权利要求6所述的装置,其特征在于,所述控制器集群中包括第一站点和第二站点;所述第一站点中设置有 $N+1$ 个第二控制器,所述第二站点中设置有 $N$ 个第二控制器和所述第一控制器;所述指定第二控制器为所述第二站点中的 $N$ 个第二控制器。

8. 根据权利要求5所述的装置,其特征在于,所述第一控制器存储有集群成员列表,所述集群成员列表用于记录所述控制器集群中每个控制器的控制器标识和每个控制器的节点属性;

所述装置还包括:

更新模块,用于将所述集群成员列表中所述第一控制器的节点属性更新为选举者,并将任一发生故障的第二控制器的节点属性更新为学习者,得到更新后的集群成员列表;

发送模块,用于将所述更新后的集群成员列表发送至各第二控制器,以使所述各第二控制器根据所述更新后的集群成员列表设置自身的节点属性。

9. 一种控制器,其特征在于,包括处理器和机器可读存储介质,所述机器可读存储介质存储有能够被所述处理器执行的机器可执行的指令,所述指令由所述处理器加载并执行:以实现权利要求1-4任一项所述的方法。

10. 一种控制器集群,其特征在于,包括:第一控制器和 $2N+1$ 个第二控制器;所述第一控制器的节点属性为学习者,所述第二控制器的节点属性为选举者;所述 $N$ 为正整数;若控制器的节点属性为学习者,则表征禁止该控制器参与主控制器的选举和投票;若控制器的节点属性为选举者,则表征允许该控制器参与主控制器的选举和投票;

所述第一控制器,用于检测各第二控制器是否发生故障;若检测到 $N+1$ 个第二控制器发生故障,则将所述第一控制器自身的节点属性变更为选举者。

11. 根据权利要求10所述的控制器集群,其特征在于,所述控制器集群中包括第一站点和第二站点;所述第一站点中设置有 $N+1$ 个第二控制器,所述第二站点中设置有 $N$ 个第二控制器和所述第一控制器。

## 控制器集群故障处理方法、装置、控制器及控制器集群

### 技术领域

[0001] 本发明涉及网络技术领域,特别是涉及一种控制器集群故障处理方法、装置、控制器及控制器集群。

### 背景技术

[0002] 随着网络控制器的发展,以及用户网络规模的不断增大,单个控制器已经无法满足用户的需求,并且单个控制器存在单点故障的隐患,可能导致用户网络失去控制而出现问题,因此,控制器需支持分布式集群部署。为满足用户需求,控制器集群支持部署在多个不同的站点,当某一站点因为地震、火灾、停电等灾难性事故,导致该站点部署的控制器不可用时,希望其他站点部署的控制器可以正常组建集群,继续给用户提供服务。

[0003] 控制器集群的服务功能主要由主控制器提供,控制器集群选举主控制器的算法中,要求选举胜出的主控制器获得半数以上的控制器投票。如果控制器集群中部署偶数(即 $2N$ , $N$ 为正整数)台控制器,在选举主控制器时,可能有两台控制器同时获得半数(即 $N$ 台)控制器投票,导致无法选出主控制器,使得控制器集群无法提供服务。因此,在实际部署控制器集群时,一般使用奇数(即 $2N+1$ , $N$ 为正整数)台控制器组建集群。

[0004] 对于部署了 $2N+1$ 台控制器的控制器集群,如果有 $N$ 台控制器发生故障,另外 $N+1$ 台控制器可以正常选举主控制器,因为在选举主控制器时有控制器可以获得半数以上(即 $N+1$ )的投票,则控制器集群可以继续提供服务。但是,如果有 $N+1$ 台控制器发生故障,则剩下的 $N$ 台控制器不符合获得半数以上(即 $N+1$ )投票的选举条件,因此无法选出主控制器,控制器集群也就无法提供服务。

### 发明内容

[0005] 本发明实施例的目的在于提供一种控制器集群故障处理方法、装置、控制器及控制器集群,以保证在有控制器故障的情况下,控制器集群能够正常提供服务。具体技术方案如下:

[0006] 第一方面,本发明实施例提供了一种控制器集群故障处理方法,应用于控制器集群中的第一控制器;控制器集群包括:第一控制器和 $2N+1$ 个第二控制器;第一控制器的节点属性为学习者,第二控制器的节点属性为选举者; $N$ 为正整数;若控制器的节点属性为学习者,则表征禁止该控制器参与主控制器的选举和投票;若控制器的节点属性为选举者,则表征允许该控制器参与主控制器的选举和投票;

[0007] 该方法包括:

[0008] 检测各第二控制器是否发生故障;

[0009] 若检测到 $N+1$ 个第二控制器发生故障,则将第一控制器自身的节点属性变更为选举者。

[0010] 第二方面,本发明实施例提供了一种控制器集群故障处理装置,应用于控制器集群中的第一控制器;控制器集群包括:第一控制器和 $2N+1$ 个第二控制器;第一控制器的节点

属性为学习者,第二控制器的节点属性为选举者; $N$ 为正整数;若控制器的节点属性为学习者,则表征禁止该控制器参与主控制器的选举和投票;若控制器的节点属性为选举者,则表征允许该控制器参与主控制器的选举和投票;

[0011] 该装置包括:

[0012] 检测模块,用于检测各第二控制器是否发生故障;

[0013] 设置模块,用于若检测到 $N+1$ 个第二控制器发生故障,则将第一控制器自身的节点属性变更为选举者。

[0014] 第三方面,本发明实施例提供了一种控制器,包括处理器和机器可读存储介质,机器可读存储介质存储有能够被处理器执行的机器可执行的指令,指令由处理器加载并执行:以实现本发明实施例第一方面所提供的方法。

[0015] 第四方面,本发明实施例提供了一种机器可读存储介质,机器可读存储介质内存储有机器可执行的指令,指令由处理器加载并执行,以实现本发明实施例第一方面所提供的方法。

[0016] 第五方面,本发明实施例提供了一种控制器集群,包括:第一控制器和 $2N+1$ 个第二控制器;第一控制器的节点属性为学习者,第二控制器的节点属性为选举者; $N$ 为正整数;若控制器的节点属性为学习者,则表征禁止该控制器参与主控制器的选举和投票;若控制器的节点属性为选举者,则表征允许该控制器参与主控制器的选举和投票;

[0017] 第一控制器,用于检测各第二控制器是否发生故障;若检测到 $N+1$ 个第二控制器发生故障,则将第一控制器自身的节点属性变更为选举者。

[0018] 本发明实施例提供一种控制器集群故障处理方法、装置、控制器及控制器集群,控制器集群中包括:第一控制器和 $2N+1$ 个第二控制器;第一控制器的节点属性为学习者,第二控制器的节点属性为选举者;若控制器的节点属性为学习者,则表征禁止该控制器参与主控制器的选举和投票;若控制器的节点属性为选举者,则表征允许该控制器参与主控制器的选举和投票。第一控制器检测各第二控制器是否发生故障;如果第一控制器检测到 $N+1$ 个第二控制器发生故障,则将第一控制器自身的节点属性变更为选举者。

[0019] 在本发明实施例中,针对每个控制器设置了不同的节点属性,节点属性为选举者的控制器允许参与主控制器的选举和投票,节点属性为学习者的控制器禁止参与主控制器的选举和投票,在正常情况下,控制器集群部署有 $2N+1$ 个节点属性为选举者的第二控制器,也就是说,参与主控制器选举和投票的控制器数目是奇数,在有 $N+1$ 个第二控制器发生故障的情况下,节点属性为学习者的第一控制器将自身的节点属性变更为选举者,此时,有 $N+1$ 个控制器的节点属性为选举者,能够参与主控制器的选举和投票,也就是说,有超过半数以上的控制器可以参与主控制器的选举和投票,从而实现了在有控制器故障的情况下,控制器集群能够正常提供服务的目的。

## 附图说明

[0020] 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实施例。

- [0021] 图1a为现有技术的ADWAN控制器集群的结构示意图；
- [0022] 图1b为现有技术的ADWAN控制器集群中站点B发生灾难性事故时的结构示意图；
- [0023] 图1c为现有技术的ADWAN控制器集群中站点A发生灾难性事故时的结构示意图；
- [0024] 图2为本发明实施例的控制器集群故障处理方法的流程示意图；
- [0025] 图3为本发明实施例的ADWAN控制器集群的结构示意图；
- [0026] 图4为本发明实施例的节点属性更新的ADWAN控制器集群的结构示意图；
- [0027] 图5为本发明实施例的ADWAN控制器集群恢复后的结构示意图；
- [0028] 图6为本发明实施例的控制器集群故障处理装置的结构示意图；
- [0029] 图7为本发明实施例的控制器结构示意图；
- [0030] 图8为本发明实施例的控制器集群的结构示意图。

### 具体实施方式

[0031] 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。

[0032] ADWAN(Application Driven Wide Area Network,应用驱动广域网)控制器是基于在广域网领域的技术积累和丰富经验,采用SDN(Software Defined Network,软件定义网络)思想及其相关技术开发的新一代广域网控制器。ADWAN控制器的集群支持部署在两个不同的地点,在本发明实施例中,分别称这两个地点为站点A、站点B。站点A、站点B部署的ADWAN控制器组成异地集群,给用户提供服务。

[0033] 当站点A因为地震、火灾、停电等灾难性事故,导致站点A部署的ADWAN控制器不可用时,用户希望站点B部署的ADWAN控制器可以正常组建集群,继续给用户提供服务。

[0034] 在ADWAN控制器集群中,控制器的状态包括如下三种类型:

[0035] (1) Follower(跟随者):控制器加入集群后的初始类型。

[0036] (2) Candidate(候选者):当选举计时器超时后,控制器由Follower变为Candidate,并发起Leader(主控制器)选举。若本轮Leader选举失败,Candidate发起新一轮Leader选举。

[0037] (3) Leader:集群的主控制器。当某Candidate获得半数以上控制器的选举投票,它就由Candidate变为Leader,并向其他控制器发送心跳消息,宣告它已成为Leader,其他Candidate控制器将由Candidate变为Follower。

[0038] 在对ADWAN控制器进行集群部署时,使用 $2N+1$ 台ADWAN控制器组建集群。如图1a所示,站点A、站点B中的一方(假设为站点A)部署 $N+1$ 台(例如2台)控制器,另一方(假设为站点B)部署 $N$ 台(例如1台)控制器。如图1b所示,当站点B发生灾难性事故时,控制器C3不可用,控制器C1、C2可用,控制器集群中可用控制器的数量为2,满足选举Leader的条件,控制器C1、C2可以组建集群,继续为用户提供服务。如图1c所示,当站点A发生灾难性事故时,控制器C1、C2不可用、C3可用,控制器集群中可用控制器的数量为1,不满足选举Leader的条件,无法为用户提供服务。可见,当站点A发生灾难性事故时,站点B中的控制器无法组建集群,导致集群无法恢复,不能为用户提供服务。

[0039] 为了保证在有控制器故障的情况下,控制器集群能够正常提供服务,本发明实施例提供了一种控制器集群故障处理方法、装置、控制器、机器可读存储介质及控制器集群。下面首先对本发明实施例所提供的控制器集群故障处理方法进行介绍。

[0040] 本发明实施例所提供的控制器集群故障处理方法可以应用于控制器集群中的第一控制器,控制器集群包括第一控制器和 $2N+1$ 个第二控制器, $N$ 为正整数。本实施例中以第一控制器和第二控制器有不同的节点属性为例,对本实施例中的方法进行说明。具体的,以第一控制器当前的节点属性为学习者、以第二控制器当前的节点属性为选举者为例进行说明。其中,节点属性为学习者表征禁止该控制器参与主控制器的选举和投票,节点属性为选举者表征允许该控制器参与主控制器的选举和投票。其中第一控制器和第二控制器的节点属性会随着满足不同的条件而发生变化。具体的会在后面的实施例中加以说明。实现本发明实施例所提供的控制器集群故障处理方法的方式可以为上述第一控制器上的软件、硬件电路和逻辑电路中的至少一种方式。

[0041] 如图2所示,本发明实施例所提供的一种控制器集群故障处理方法,应用于上述第一控制器,该控制器集群故障处理方法可以包括如下步骤。

[0042] S201,检测各第二控制器是否发生故障。

[0043] S202,若检测到 $N+1$ 个第二控制器发生故障,则将第一控制器自身的节点属性变更为选举者。

[0044] 应用本发明实施例,针对每个控制器设置了不同的节点属性,节点属性为选举者的控制器允许参与主控制器的选举和投票,节点属性为学习者的控制器禁止参与主控制器的选举和投票,在正常情况下,控制器集群部署有 $2N+1$ 个当前的节点属性为选举者的第二控制器,也就是说,参与主控制器选举和投票的控制器数目是奇数,在有 $N+1$ 个第二控制器发生故障的情况下,节点属性为学习者的第一控制器将自身的节点属性变更为选举者,此时,有 $N+1$ 个控制器的节点属性为选举者,能够参与主控制器的选举和投票,也就是说,控制器集群中有超过半数以上的控制器可以参与主控制器的选举和投票,从而实现了在有控制器故障的情况下,控制器集群能够正常提供服务的目的。

[0045] 本发明实施例中,以第一控制器当前的节点属性为Learner(学习者)为例进行说明,当前作为Learner(学习者)的第一控制器不参与主控制器选举投票,不计入选举投票总票数,当选举计时器超时时,第一控制器不会从Follower变为Candidate,不会发起选举;本实施例中,以第二控制器当前的节点属性为Voter(选举者)为例进行说明,当前作为Voter(选举者)的第二控制器参与主控制器选举投票,计入选举投票总票数,当选举计时器超时时,第二控制器从Follower变为Candidate,然后发起选举。

[0046] 第一控制器可以检测各第二控制器是否发生故障。具体检测各第二控制器是否发生故障的方式可以是:向各第二控制器发送探测报文,如果第二控制器未故障,则会回复响应报文,如果在一定时间内,没有收到某个第二控制器回复的响应报文,则说明该第二控制器故障。检测各第二控制器是否发生故障的方式还可以是:各第二控制器之间以及各第二控制器与第一控制器之间会定时发送心跳消息,如果第一控制器在一段时间内没有收到某个第二控制器发送的心跳消息,则说明该第二控制器故障。当然,检测各第二控制器是否故障的方式还可以通过事件检测、硬件状态检测等方式实现,这里不再一一赘述。

[0047] 在本申请实施例的一种实现方式中,S201具体可以为:针对任一第二控制器,若在

预设周期内,未收到该第二控制器发送的心跳消息,则确定该第二控制器发生故障。

[0048] 如上述,各第二控制器之间以及各第二控制器与第一控制器之间会定时发送心跳消息,也就是说,针对某一个第二控制器,第一控制器知道每间隔预设周期就可以收到该第二控制器发送的心跳消息,如果第一控制器在预设周期内没有收到心跳消息,则说明该第二控制器与第一控制器不通,可以初步判断该第二控制器故障。进一步的,为了提高检测的准确性,还可以判断是否连续多个周期都没有收到该第二控制器发送的心跳信息,如果连续多个周期都没有收到该第二控制器发送的心跳信息,则确定第二控制器与第一控制器不通,确定该第二控制器发生故障。

[0049] 在本申请实施例的一种实现方式中,各第二控制器记录有:经检测得到的其他第二控制器是否可达的检测信息。S201具体可以为:针对任一第二控制器,若在预设周期内,未收到该第二控制器发送的心跳消息,则从各指定第二控制器中获取针对该第二控制器的检测信息,其中,指定第二控制器为经检测确定未发生故障的第二控制器;若获取到的所有检测信息中存在该第二控制器不可达,则确定该第二控制器发生故障。

[0050] 如上述,除了各第二控制器与第一控制器之间会定时发送心跳消息以外,各第二控制器之间也会定时发送心跳消息,各第二控制器可以采用与第一控制器判断第二控制器是否故障相同的检测方式,来检测其他第二控制器是否故障,各第二控制器可以在本地记录经检测得到的其他第二控制器是否可达的检测信息,具体可以利用第二控制器可达列表来对检测信息进行记录,可达的第二控制器用Reachable标识,不可达的第二控制器用Unreachable标识。针对某一个第二控制器,如果第一控制器在预设周期内没有收到心跳消息,则说明该第二控制器与第一控制器不通,可以初步判断该第二控制器故障,然后,第一控制器从各指定第二控制器中获取针对该第二控制器的检测信息,若获取到的所有检测信息中存在该第二控制器不可达,则说明该第二控制器确实故障,其他第二控制器也无法与其交互心跳信息。具体的,第一控制器从各指定第二控制器中获取针对该第二控制器的检测信息的方式可以是:第一控制器向各指定第二控制器发送获取指令,各指定第二控制器收到获取指令后,将收集的关于该第二控制器的检测信息发送给第一控制器。本实施例中所述的指定第二控制器是指经检测确定未发生故障的第二控制器,这些指定第二控制器都能够和第一控制器正常交互心跳消息。

[0051] 在本申请实施例的一种实现方式中,控制器集群中包括第一站点和第二站点;第一站点中设置有N+1个第二控制器,第二站点中设置有N个第二控制器和第一控制器;指定第二控制器为第二站点中的N个第二控制器。

[0052] 本申请实施例具体可以应用于ADWAN控制器集群,ADWAN控制器集群支持部署在两个不同的站点(即第一站点和第二站点),第一站点中设置有N+1个第二控制器,第二站点中设置有N个第二控制器和第一控制器,通常情况下,地震、火灾、停电等灾难性事故发生某一个站点,即如果第二站点发生灾难性事故时,第二站点中的N个第二控制器不可用,第一站点中的N+1个控制器可用,ADWAN控制器集群中可用控制器的数量为N+1,可用控制器的数量过半;而如果第一站点发生灾难性事故时,第一站点中的N+1个第二控制器不可用,第二站点中的N个第二控制器可用,ADWAN控制器集群中可用控制器的数量为N,可用控制器的数量未到一半,此时,第二站点中的第一控制器接收不到第一站点中各第二控制器的心跳信息,则初步判断第一站点中的各第二控制器故障,然后从第二站点中的各第二控制器(即指定

第二控制器)获取针对第一站点中各第二控制器的检测信息,如果基于检测信息确定第一站点中各第二控制器都是不可达的,则可以确定第一站点中的各第二控制器都发生了故障。此时,第一控制器将自身的节点属性变更为选举者,也就是说,控制器集群中能够参与主控制器的选举和投票的第二控制器仍然有N+1个,控制器集群中可用控制器的数量为N+1,可用控制器的数量过半。

[0053] 在本发明实施例中,如果第一控制器检测到N+1个第二控制器发生故障,此时将第一控制器自身的节点属性变更为选举者。原本有N个节点属性为选举者的第二控制器正常、此时第一控制器的节点属性也变更为了选举者,则节点属性为选举者的控制器的数量仍然为N+1个,也就是说,有超过半数以上的控制器可以参与主控制器的选举和投票,从而实现了在有控制器故障的情况下,控制器集群能够正常提供服务的目的。

[0054] 在本申请实施例的一种实现方式中,第一控制器存储有集群成员列表,集群成员列表用于记录控制器集群中每个控制器的控制器标识和每个控制器的节点属性。

[0055] 在S202之后,该方法还可以包括:将集群成员列表中第一控制器的节点属性更新为选举者,并将任一发生故障的第二控制器的节点属性更新为学习者,得到更新后的集群成员列表;将更新后的集群成员列表发送至各第二控制器,以使各第二控制器根据更新后的集群成员列表设置自身的节点属性。

[0056] 第一控制器本地可以记录一个集群成员列表,该集群成员列表用于记录控制器集群中每个控制器的控制器标识和每个控制器的节点属性,例如,可以以Cluster={ (C1, Voter), (C2, Voter), (C3, Voter), (C4, Learner) }来表示控制器C1的节点属性为选举者、控制器C2的节点属性为选举者、控制器C3的节点属性为选举者、控制器C4的节点属性为学习者。这样,第一控制器在变更自身的节点属性以外,还会更新集群成员列表,具体更新的内容包括:将第一控制器的节点属性更新为选举者,将任一发生故障的第二控制器的节点属性更新为学习者,例如控制器C1和C2发生故障,则第一控制器C4可以选择控制器C1或者C2,将集群成员列表中其节点属性更新为学习者。第一控制器在更新集群成员列表之后,将更新后的集群成员列表发送给各第二控制器,各第二控制器在接收到更新后的集群成员列表后,则可根据更新后的集群成员列表设置自身的节点属性。

[0057] 为了便于理解,下面以两个站点分别部署两台控制器的ADWAN控制器集群为例,控制器集群的结构如图3所示,控制器C1和C2部署在站点A中,控制器C3和C4部署在站点B中,控制器C1、C2和C3的节点属性为Voter(选举者),控制器C4的节点属性为Learner(学习者)。虽然控制器集群中包含4台控制器,但参与主控制器选举投票的控制器只有3台,即选举投票总票数是3。因此,在主控制器选举投票时,控制器至少需要获得2张投票(含控制器自身投出的1票),即可成为本次选举的主控制器。

[0058] 在4个控制器组成的控制器集群中,节点属性为Voter的第二控制器最多只能有3个。第二控制器之间、第二控制器与第一控制器之间定期发送心跳消息时,均携带自身的节点属性。当某个第二控制器出现故障时,其他第二控制器和第一控制器将这个第二控制器标记为不可达。

[0059] 节点属性为学习者的第一控制器根据心跳消息,维护各第二控制器是否可达信息。当第一控制器发现2个第二控制器不可达、只有1个第二控制器可达时,控制器集群可能无法完成主控制器的选举,第一控制器根据具体情况决定是否将节点属性变更为Voter。举

例说明如下：

[0060] 使用Reachable代表可达、Unreachable代表不可达列表，使用 $\text{Cluster}_{(\text{Term})}$ 代表集群成员列表，其中，集群成员以控制器标识、节点属性组成的二元组表示，Term表示主控制器的任期（代表选举出一个控制器作为主控制器的期限，如果达到期限，则重新进行主控制器选举），初始值为0，Candidate发起主控制器选举时，当前任期值需加1。假设图3中的控制器C4发现控制器C1和C2不可达、C3可达，即

[0061]  $\text{Cluster}_{(\text{Term})} = \{ (C1, \text{Voter}), (C2, \text{Voter}), (C3, \text{Voter}), (C4, \text{Learner}) \}$

[0062]  $\text{Reachable}_{(C4)} = \{ C3 \}, \text{Unreachable}_{(C4)} = \{ C1, C2 \}$

[0063] 控制器C4与C3发起探测（从控制器C3获取检测信息），确认C1、C2是否均不可达。

[0064] 若 $\text{Reachable}_{(C3)} = \{ C3, C1 \}, \text{Unreachable}_{(C3)} = \{ C2 \}$ ，说明C1与C3可达，但C1与C4之间的网络可能出现故障，符合选举条件；若 $\text{Reachable}_{(C3)} = \{ C3, C2 \}, \text{Unreachable}_{(C3)} = \{ C1 \}$ ，说明C2与C3可达，但C2与C4之间的网络可能出现故障，符合选举条件；若 $\text{Reachable}_{(C3)} = \{ C3, C1, C2 \}, \text{Unreachable}_{(C3)} = \{ \}$ ，说明C1、C2与C3可达，但C1、C2与C4之间的网络可能出现故障，符合选举条件。

[0065] 在上述三种情况下，控制器集群可以选举主控制器，控制器C4的节点属性无需变更为Voter。

[0066] 若 $\text{Reachable}_{(C3)} = \{ C3 \}, \text{Unreachable}_{(C3)} = \{ C1, C2 \}$ ，说明C1、C2与C3也不可达，不符合主控制器的选举条件，无法组建控制器集群。

[0067] 并且，控制器C4与C3进行周期性探测，假设探测周期为1分钟，若控制器C4连续探测设定次数（设定为5次）后发现，在每次探测中C1、C2都不可达，则C4与C3确认C1、C2不可达是有效的，于是C4的节点属性由学习者变更为Voter。

[0068] C4的节点属性由Learner变更为Voter的同时，从不可达的控制器C1和C2中，随机选择一个的节点属性变更为Learner。假设当前任期为N，并选择控制器C1的节点属性变更为Learner，则有：

[0069]  $\text{Cluster}_{(N)} = \{ (C1, \text{Voter}), (C2, \text{Voter}), (C3, \text{Voter}), (C4, \text{Learner}) \}$

[0070]  $\text{Cluster}_{(N+1)} = \{ (C1, \text{Learner}), (C2, \text{Voter}), (C3, \text{Voter}), (C4, \text{Voter}) \}$

[0071] 再次选举主控制器时，控制器集群中各控制器的节点属性以最新的 $\text{Cluster}_{(N+1)}$ 为准。

[0072] 当控制器C4与C3再次互传心跳消息后， $\text{Reachable}_{(C4)} = \{ C4, C3 \}, \text{Unreachable}_{(C4)} = \{ C2 \}, \text{Reachable}_{(C3)} = \{ C3, C4 \}, \text{Unreachable}_{(C3)} = \{ C2 \}$ ，符合主控制器选举条件，因为投票总票数为3，C3、C4最多均可以获得2张投票，超过总票数的半数，它们均可以被选举为主控制器，因此，C3、C4可以恢复控制器集群，如图4所示。

[0073] 若控制器C1、C2文件系统未损坏，可以恢复运行，则控制器C1、C2在启动后加入控制器集群，将会检测到集群成员列表已经发生变更，并且控制器C1、C2本地维护的集群成员列表较旧，因此根据最新的集群成员列表进行更新，控制器C1的节点属性为Learner，控制器C2的节点属性为Voter。然后，控制器C1以Learner的节点属性、C2以Voter的节点属性加入控制器集群。

[0074] 若控制器C1、C2文件系统已损坏，无法恢复运行，则使用两台全新的控制器C1' 和C2' 替换C1和C2，根据最新的集群成员列表对新的控制器的节点属性进行配置。恢复后的控

制器集群结构如图5所示。

[0075] 综上,当站点A发生灾难性事故导致控制器C1、C2不可用时,站点B的控制器C3、C4可以恢复控制器集群,继续为用户提供服务,并且当站点A从灾难性事故中恢复后,控制器C1、C2可以重新加入控制器集群。

[0076] 相应于上述方法实施例,本发明实施例提供了一种控制器集群故障处理装置,应用于控制器集群中的第一控制器;控制器集群包括:第一控制器和 $2N+1$ 个第二控制器;第一控制器的节点属性为学习者,第二控制器的节点属性为选举者; $N$ 为正整数;若控制器的节点属性为学习者,则表征禁止该控制器参与主控制器的选举和投票;若控制器的节点属性为选举者,则表征允许该控制器参与主控制器的选举和投票;如图6所示,该装置可以包括:

[0077] 检测模块610,用于检测各第二控制器是否发生故障;

[0078] 设置模块620,用于若检测到 $N+1$ 个第二控制器发生故障,则将第一控制器自身的节点属性变更为选举者。

[0079] 可选的,各第二控制器记录有:经检测得到的其他第二控制器是否可达的检测信息;

[0080] 检测模块610,具体可以用于:针对任一第二控制器,若在预设周期内,未收到该第二控制器发送的心跳消息,则从各指定第二控制器中获取针对该第二控制器的检测信息,其中,指定第二控制器为经检测确定未发生故障的第二控制器;若获取到的所有检测信息中存在该第二控制器不可达,则确定该第二控制器发生故障。

[0081] 可选的,控制器集群中包括第一站点和第二站点;第一站点中设置有 $N+1$ 个第二控制器,第二站点中设置有 $N$ 个第二控制器和第一控制器;指定第二控制器为第二站点中的 $N$ 个第二控制器。

[0082] 可选的,第一控制器存储有集群成员列表,集群成员列表用于记录控制器集群中每个控制器的控制器标识和每个控制器的节点属性;

[0083] 该装置还可以包括:

[0084] 更新模块,用于将集群成员列表中第一控制器的节点属性更新为选举者,并将任一发生故障的第二控制器的节点属性更新为学习者,得到更新后的集群成员列表;

[0085] 发送模块,用于将更新后的集群成员列表发送至各第二控制器,以使各第二控制器根据更新后的集群成员列表设置自身的节点属性。

[0086] 应用本发明实施例,针对每个控制器设置了不同的节点属性,节点属性为选举者的控制器允许参与主控制器的选举和投票,节点属性为学习者的控制器禁止参与主控制器的选举和投票,在正常情况下,控制器集群部署有 $2N+1$ 个节点属性为选举者的第二控制器,也就是说,参与主控制器选举和投票的控制器数目是奇数,在有 $N+1$ 个第二控制器发生故障的情况下,节点属性为学习者的第一控制器将自身的节点属性变更为选举者,此时,有 $N+1$ 个控制器的节点属性为选举者,能够参与主控制器的选举和投票,也就是说,有超过半数以上的控制器可以参与主控制器的选举和投票,从而实现了在有控制器故障的情况下,控制器集群能够正常提供服务的目的。

[0087] 本发明实施例还提供了一种控制器,如图7所示,包括处理器701和机器可读存储介质702,机器可读存储介质702存储有能够被处理器701执行的机器可执行的指令,指令由处理器701加载并执行:以实现本发明实施例所提供的控制器集群故障处理方法。

[0088] 上述机器可读存储介质可以包括RAM(Random Access Memory,随机存取存储器),也可以包括NVM(Non-volatile Memory,非易失性存储器),例如至少一个磁盘存储器。可选的,机器可读存储介质还可以是至少一个位于远离前述处理器的存储装置。

[0089] 上述处理器可以是通用处理器,包括CPU(Central Processing Unit,中央处理器)、NP(Network Processor,网络处理器)等;还可以是DSP(Digital Signal Processor,数字信号处理器)、ASIC(Application Specific Integrated Circuit,专用集成电路)、FPGA(Field-Programmable Gate Array,现场可编程门阵列)或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件。

[0090] 机器可读存储介质702与处理器701之间可以通过有线连接或者无线连接的方式进行数据传输,并且控制器与其他控制器或者网络设备之间可以通过有线通信接口或者无线通信接口进行通信。图7所示的仅为通过总线进行数据传输的示例,不作为具体连接方式的限定。

[0091] 本实施例中,处理器701通过读取机器可读存储介质702中存储的机器可执行的指令,能够实现:针对每个控制器设置了不同的节点属性,节点属性为选举者的控制器允许参与主控制器的选举和投票,节点属性为学习者的控制器禁止参与主控制器的选举和投票,在正常情况下,控制器集群部署有 $2N+1$ 个节点属性为选举者的第二控制器,也就是说,参与主控制器选举和投票的控制器的数目是奇数,在有 $N+1$ 个第二控制器发生故障的情况下,节点属性为学习者的第一控制器将自身的节点属性变更为选举者,此时,有 $N+1$ 个控制器的节点属性为选举者,能够参与主控制器的选举和投票,也就是说,有超过半数以上的控制器可以参与主控制器的选举和投票,从而实现了在有控制器故障的情况下,控制器集群能够正常提供服务的目的。

[0092] 另外,本发明实施例还提供了一种机器可读存储介质,机器可读存储介质内存储有机器可执行的指令,指令由处理器加载并执行,以实现本发明实施例所提供的控制器集群故障处理方法。

[0093] 本实施例中,机器可读存储介质存储有处理器在运行时执行本发明实施例所提供的方法的机器可执行的指令,因此能够实现:针对每个控制器设置了不同的节点属性,节点属性为选举者的控制器允许参与主控制器的选举和投票,节点属性为学习者的控制器禁止参与主控制器的选举和投票,在正常情况下,控制器集群部署有 $2N+1$ 个节点属性为选举者的第二控制器,也就是说,参与主控制器选举和投票的控制器的数目是奇数,在有 $N+1$ 个第二控制器发生故障的情况下,节点属性为学习者的第一控制器将自身的节点属性变更为选举者,此时,有 $N+1$ 个控制器的节点属性为选举者,能够参与主控制器的选举和投票,也就是说,有超过半数以上的控制器可以参与主控制器的选举和投票,从而实现了在有控制器故障的情况下,控制器集群能够正常提供服务的目的。

[0094] 本发明实施例提供了一种控制器集群,如图8所示,包括:第一控制器810和 $2N+1$ 个第二控制器820;第一控制器810的节点属性为学习者,第二控制器820的节点属性为选举者; $N$ 为正整数;若控制器的节点属性为学习者,则表征禁止该控制器参与主控制器的选举和投票;若控制器的节点属性为选举者,则表征允许该控制器参与主控制器的选举和投票;

[0095] 第一控制器810,用于检测各第二控制器820是否发生故障;若检测到 $N+1$ 个第二控制器820发生故障,则将第一控制器810自身的节点属性变更为选举者。

[0096] 可选的,控制器集群中包括第一站点和第二站点;第一站点中设置有 $N+1$ 个第二控制器,第二站点中设置有 $N$ 个第二控制器和第一控制器。

[0097] 应用本发明实施例,针对每个控制器设置了不同的节点属性,节点属性为选举者的控制器允许参与主控制器的选举和投票,节点属性为学习者的控制器禁止参与主控制器的选举和投票,在正常情况下,控制器集群部署有 $2N+1$ 个节点属性为选举者的第二控制器,也就是说,参与主控制器选举和投票的控制器数目是奇数,在有 $N+1$ 个第二控制器发生故障的情况下,节点属性为学习者的第一控制器将自身的节点属性变更为选举者,此时,有 $N+1$ 个控制器的节点属性为选举者,能够参与主控制器的选举和投票,也就是说,有超过半数以上的控制器可以参与主控制器的选举和投票,从而实现了在有控制器故障的情况下,控制器集群能够正常提供服务的目的。

[0098] 第一控制器810与第二控制器820之间、第二控制器820与第二控制器820可以通过有线连接或者无线连接的方式进行数据传输。图8所示的仅为通过总线进行数据传输的示例,不作为具体连接方式的限定。

[0099] 对于控制器、机器可读存储介质、控制器集群实施例而言,由于其涉及的方法内容基本相似于前述的方法实施例,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。

[0100] 需要说明的是,在本文中,诸如第一和第二等之类的关系术语仅仅用来将一个实体或者操作与另一个实体或操作区分开来,而不一定要求或者暗示这些实体或操作之间存在任何这种实际的关系或者顺序。而且,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、物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、物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、物品或者设备中还存在另外的相同要素。

[0101] 本说明书中的各个实施例均采用相关的方式描述,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,对于装置、控制器、机器可读存储介质、控制器集群实施例而言,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。

[0102] 以上所述仅为本发明的较佳实施例,并非用于限定本发明的保护范围。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,均包含在本发明的保护范围内。

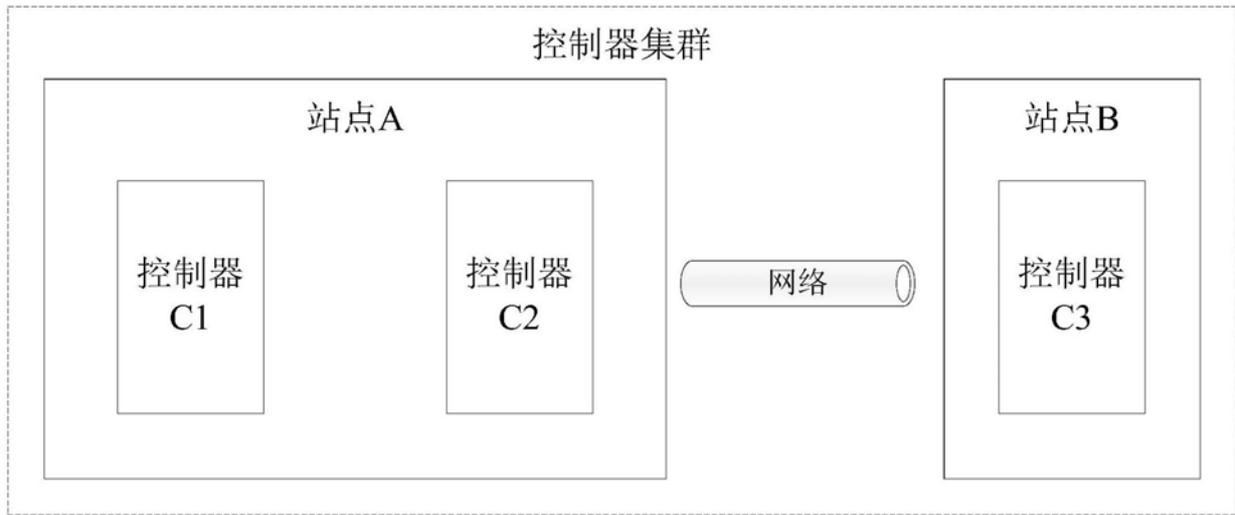


图1a

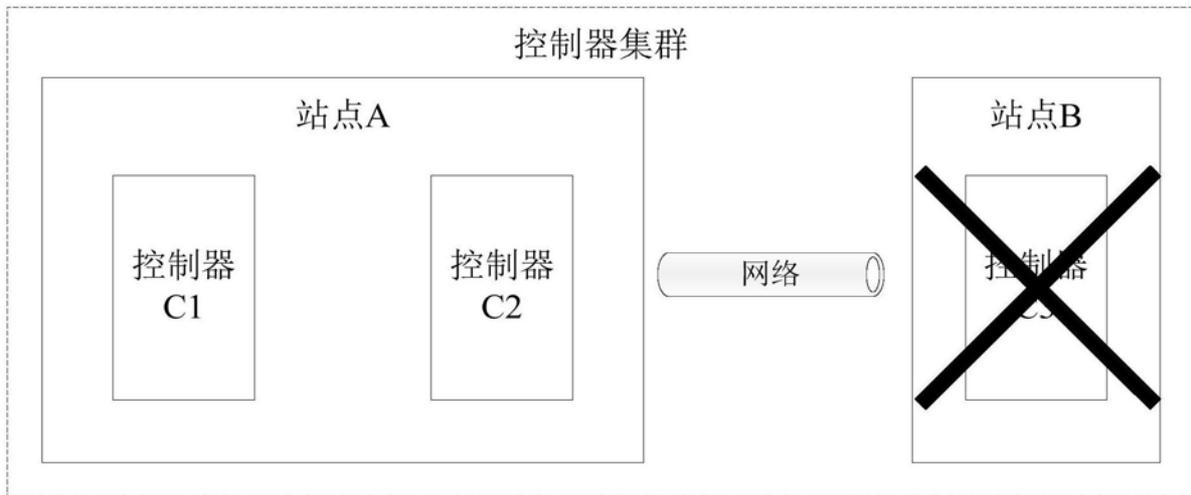


图1b

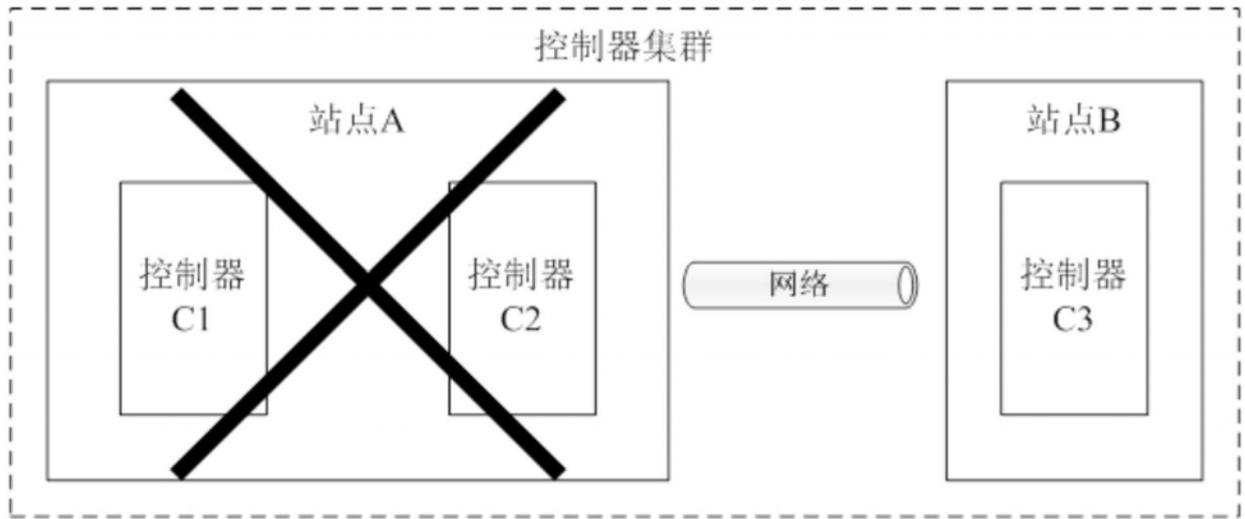


图1c

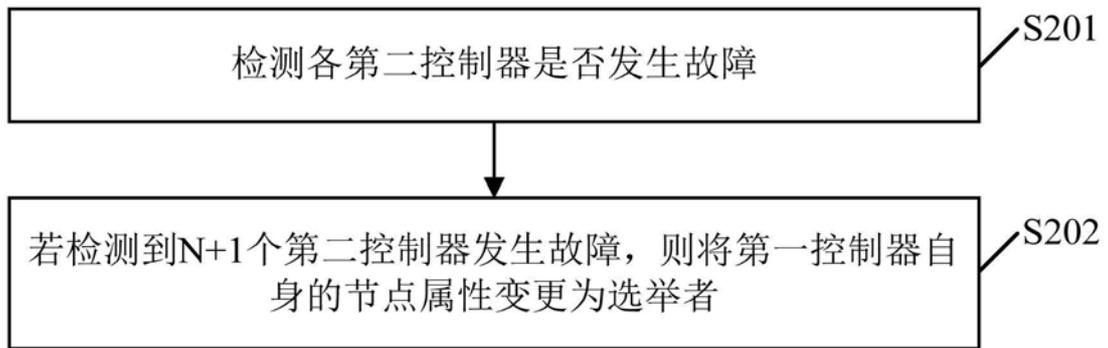


图2

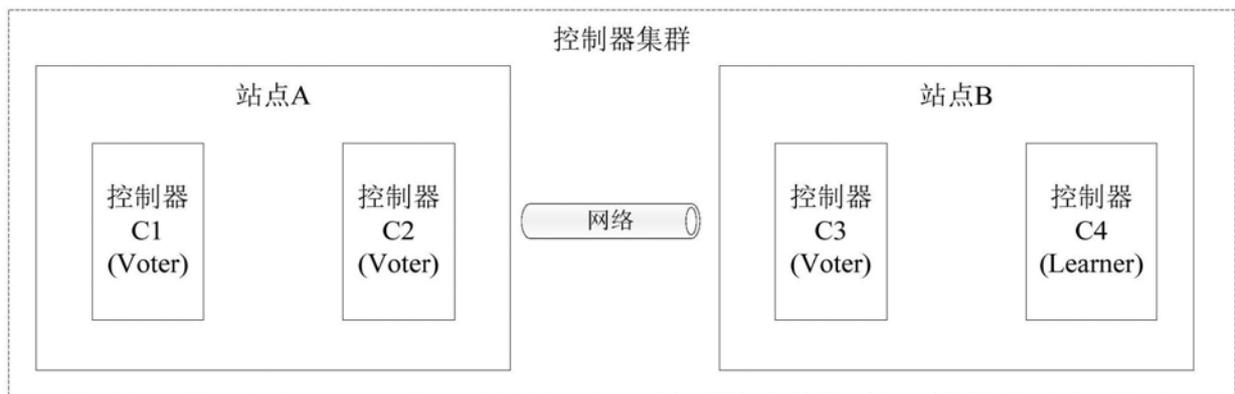


图3

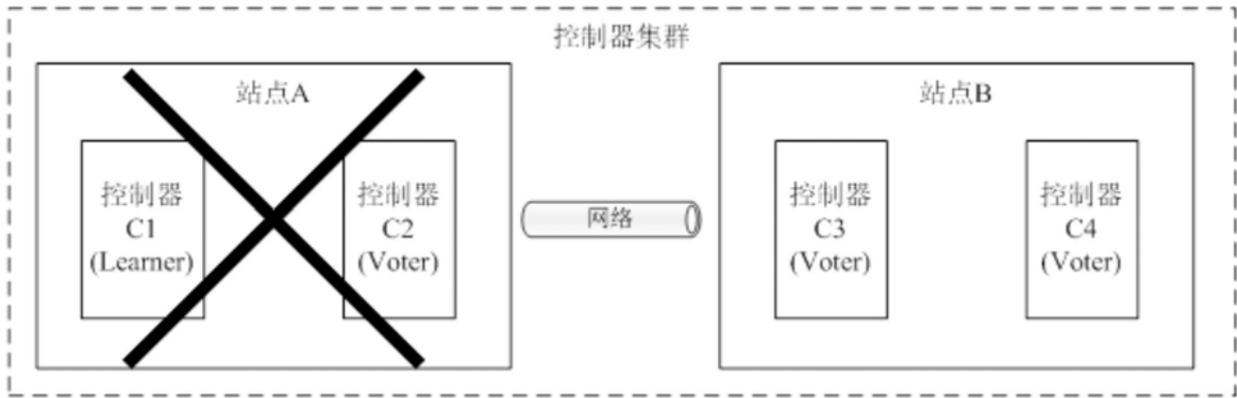


图4

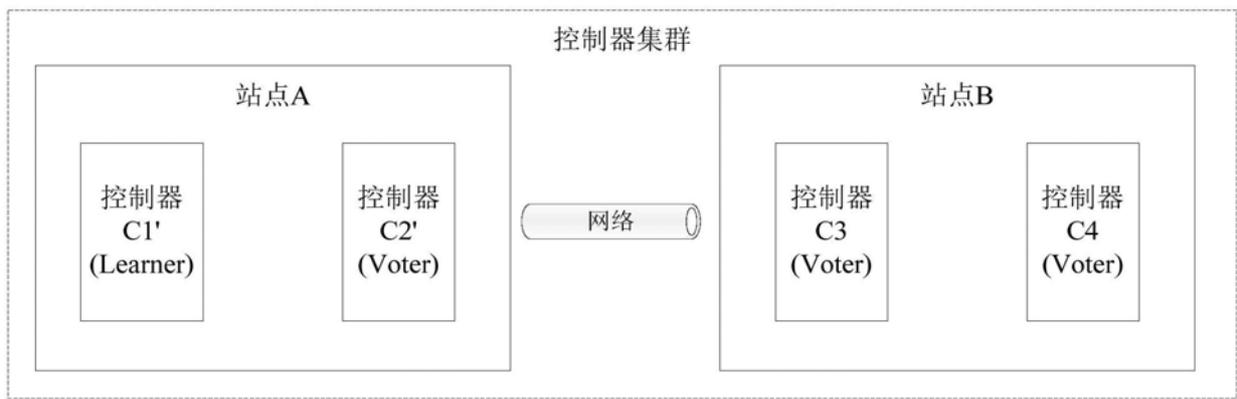


图5

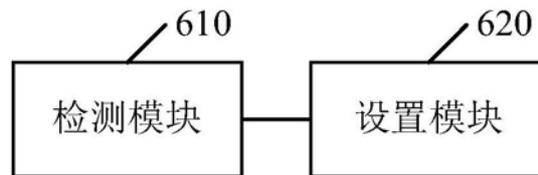


图6

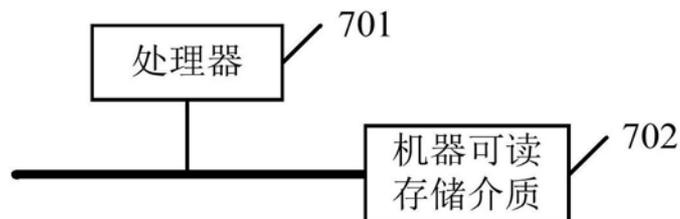


图7

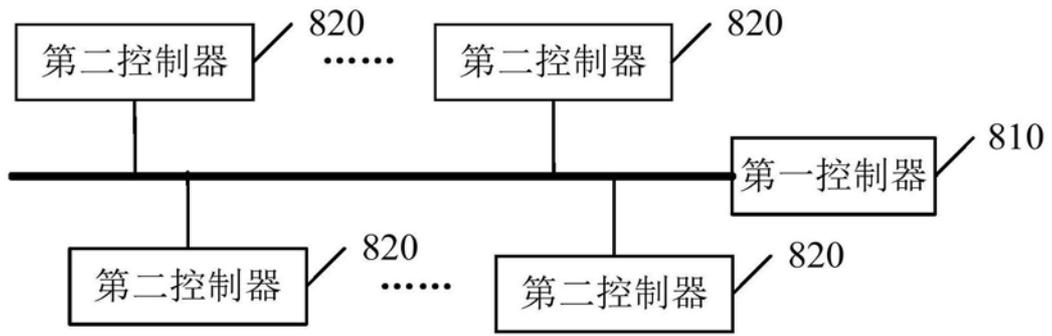


图8