(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号
特許第6731201号
(P6731201)
(45)発行日　令和2年7月29日(2020.7.29)　(24)登録日　令和2年7月8日(2020.7.8)

(51)Int.Cl.　　F I
G06F 13/00 (2006.01)　　G06F 13/00　353B
G06F 17/40 (2006.01)　　G06F 17/40

請求項の数 20　(全 31 頁)

(21)出願番号　特願2017-551678(P2017-551678)
(86)(22)出願日　平成28年3月15日(2016.3.15)
(65)公表番号　特表2018-512678(P2018-512678A)
(43)公表日　平成30年5月17日(2018.5.17)
(86)国際出願番号　PCT/CN2016/076414
(87)国際公開番号　WO2016/155496
(87)国際公開日　平成28年10月6日(2016.10.6)
審査請求日　平成31年3月15日(2019.3.15)
(31)優先権主張番号　201510155909.9
(32)優先日　平成27年4月2日(2015.4.2)
(33)優先権主張国・地域又は機関
中国(CN)

早期審査対象出願

(73)特許権者　510330264
アリババ・グループ・ホールディング・リミテッド
ＡＬＩＢＡＢＡ　ＧＲＯＵＰ　ＨＯＬＤＩＮＧ　ＬＩＭＩＴＥＤ
英国領，ケイマン諸島，グランド・ケイマン，ジョージ・タウン，ワン・キャピタル・プレイス，フォース・フロア，ピー．オー．ボックス　８４７
(74)代理人　100188558
弁理士　飯田　雅人
(74)代理人　100205785
弁理士　▲高▼橋　史生
(74)代理人　100097320
弁理士　宮川　貞二

最終頁に続く

(54)【発明の名称】時間ベースのノード選出方法及び装置

(57)【特許請求の範囲】
【請求項１】
コンピュータにより実施される方法であって：
ホストにより、全てのノードに対する、同一の投票時間と、同一の票集計時間と、同一のリーダーノードの任期とを設定するステップと；
設定された前記同一の投票時間と、前記同一の票集計時間と、前記同一のリーダーノードの任期とを含む時間設定情報を、全ての前記ノードへ送信するステップであって、前記時間設定情報を受信した前記ノードは、前記同一の投票時間中に投票し、前記同一の票集計時間中に票の数を集計し、票集計結果に従ってリーダーノードを選出するよう動作し、且つ前記ノードは、前記同一のリーダーノードの任期に従って周期的なノード選出を実行する、ステップと；を備える、
コンピュータにより実施される方法。
【請求項２】
全ノードのうちにリーダーノードが存在しない場合、前記時間設定情報に含まれる投票時間中に、投票要求を、少なくとも１つのノードにより、前記全ノードへ送信するステップと；
他のノードが前記投票要求に従って生成した投票情報を、ノードのそれぞれにより、受信するステップと；
前記時間設定情報に含まれる票集計時間中に、受信した前記投票情報に含まれる前記票の数を、前記ノードのそれぞれにより、集計するステップと；

特定のサービスを処理するために、集計したそれぞれの投票情報に含まれる前記票の数に従ってリーダーノードを、前記ノードにより、決定するステップと；を更に備える、
請求項１に記載のコンピュータにより実施される方法。
【請求項３】
前記集計するステップを開始する前に、前記投票情報の受信を終了するステップを更に備える、
請求項２に記載のコンピュータにより実施される方法。
【請求項４】
前記集計したそれぞれの投票情報に含まれる前記票の数に従ってリーダーノードを、前記ノードにより、決定するステップは、前記ノードが、集計した前記投票情報に含まれる前記票の数が予め設定された閾値を超えると判断したときに、前記ノードにより、当該ノード自体をリーダーノードとして決定するステップを備え、前記閾値は全ての前記ノードの数の１／２である、
請求項２に記載のコンピュータにより実施される方法。
【請求項５】
全ノードのうちにリーダーノードが存在しない場合、時間設定情報に含まれる投票時間中に、ノードから送信された投票要求を、投票ノードとしての少なくとも２つのノードにより、受信するステップと；
リーダーノードを決定するために、前記ノードが前記票集計時間内に投票情報に含まれる前記票の数を集計できるよう、投票要求に従って前記投票情報を、前記投票ノードにより、生成し、前記投票情報を前記投票時間中に前記ノードへ送信するステップと；を更に備える、
請求項１に記載のコンピュータにより実施される方法。
【請求項６】
前記投票情報を生成し、前記投票情報を前記ノードへ送信するステップは、
受信した前記投票要求に含まれるノード情報を、前記投票ノードにより、決定するステップと；
前記ノードの優先度を前記ノード情報に従って決定するステップと；
最大の優先度を持つノードを選出し、前記ノードのための投票情報を生成し、前記投票情報を前記選出したノードへ送信するステップと；を備える、
請求項５に記載のコンピュータにより実施される方法。
【請求項７】
全ノードのうちにリーダーノードが存在し、前記リーダーノードが、予め設定された基準を自己が満たしている、と判断した場合に、前記投票時間中に、前記時間設定情報に含まれるリーダーノードの任期に従って、再任要求を、前記リーダーノードにより、前記全ノードへ送信するステップと；
前記再任要求に従って、他のノードが生成した投票情報を、前記リーダーノードにより、受信するステップと；
前記票集計時間内に、受信した投票情報に含まれる前記票の数を、前記リーダーノードにより、集計するステップと；
集計した前記投票情報に含まれる前記票の数に従って再任状態を決定し、再任成功後に特定のサービスを、前記リーダーノードにより、処理するステップであって、前記予め設定された基準は、負荷処理基準とデータ更新度とのうち少なくとも一方を備える、ステップと；を更に備える、
請求項１に記載のコンピュータにより実施される方法。
【請求項８】
全ノードのうちにリーダーノードが存在し、前記リーダーノードが、予め設定された基準を自己が満たしていないと判断した場合に、前記時間設定情報に含まれるリーダーノードの任期に従って、後任ノードとして機能するための前記予め設定された基準を満たすノードを、前記リーダーノードにより他のノードの中から選別するステップと；

前記後任ノードは、他のノードが送信した投票情報を受信し、前記票集計時間中に前記投票情報に含まれる前記票の数を集計し、新たなリーダーノードとして決定されるようにするために、投票期間内に、後任ノードについての投票要求を、前記リーダーノードにより、他のノードへ送信するステップと；を更に備える、
請求項１に記載のコンピュータにより実施される方法。
【請求項９】
全ノードのうちにリーダーノードが存在する場合、前記時間設定情報に含まれる投票時間中に、前記リーダーノードが送信した再任要求を、投票ノードとしての少なくとも２つのノードにより、受信するステップと；
前記リーダーノードが票集計時間内に投票情報に含まれる前記票の数を集計して、再任状態を決定できるようにするために、前記投票ノードにより、前記再任要求に従って前記投票情報を生成し、前記時間設定情報に含まれる前記投票時間内に前記投票情報を前記リーダーノードへ送信するステップと；を更に備え、
前記投票ノードは、前記リーダーノードの任期中に前記リーダーノードの前記再任要求のみを受信する、
請求項１に記載のコンピュータにより実施される方法。
【請求項１０】
全ノードのうちにリーダーノードが存在する場合、時間設定情報に含まれる投票時間中に、前記リーダーノードから送信された後任ノードのための投票要求を、投票ノードとしての少なくとも２つのノードにより、受信するステップと；
前記後任ノードが票集計時間中に投票情報に含まれる前記票の数を集計して、新たなリーダーノードを決定できるようにするために、前記投票ノードにより、前記投票要求に従って前記投票情報を生成し、前記時間設定情報に含まれる前記投票時間中に前記投票情報を前記後任ノードへ送信するステップと；を更に備える、
請求項１に記載のコンピュータにより実施される方法。
【請求項１１】
コンピュータシステムによって実行可能な１つまたは複数の命令を格納する非一時的なコンピュータ読取可能媒体であって、前記命令により実行する操作は：
ホストにより、全てのノードに対する、同一の投票時間と、同一の票集計時間と、同一のリーダーノードの任期とを設定するステップと；
設定された前記同一の投票時間と、前記同一の票集計時間と、前記同一のリーダーノードの任期とを含む時間設定情報を、全ての前記ノードへ送信するステップであって、前記時間設定情報を受信した前記ノードは、前記同一の投票時間中に投票し、前記同一の票集計時間中に票の数を集計し、票集計結果に従ってリーダーノードを選出するように動作し、且つ前記ノードは、前記同一のリーダーノードの任期に従って周期的なノード選出を実行する、ステップと；を備える、
非一時的なコンピュータ読取可能媒体。
【請求項１２】
前記操作は：
全ノードのうちにリーダーノードが存在しない場合、前記時間設定情報に含まれる投票時間中に、投票要求を、少なくとも１つのノードにより、前記全ノードへ送信するステップと；
他のノードが前記投票要求に従って生成した投票情報を、ノードのそれぞれにより、受信するステップと；
前記時間設定情報に含まれる票集計時間中に、受信した前記投票情報に含まれる前記票の数を、前記ノードのそれぞれにより、集計するステップと；
特定のサービスを処理するために、集計したそれぞれの投票情報に含まれる前記票の数に従ってリーダーノードを、前記ノードにより、決定するステップと；を更に備える、
請求項１１に記載の非一時的なコンピュータ読取可能媒体。
【請求項１３】

前記操作は、前記集計するステップを開始する前に、前記投票情報の受信を終了するステップを更に備える、
請求項１２に記載の非一時的なコンピュータ読取可能媒体。
【請求項１４】
前記集計したそれぞれの投票情報に含まれる前記票の数に従ってリーダーノードを、前記ノードにより、決定するステップは、前記ノードが、集計した前記投票情報に含まれる前記票の数が予め設定された閾値を超えると判断したときに、前記ノードにより、当該ノード自体をリーダーノードとして決定するステップを備え、前記閾値は全ての前記ノードの数の１／２である、
請求項１２に記載の非一時的なコンピュータ読取可能媒体。
【請求項１５】
前記操作は：
全ノードのうちにリーダーノードが存在しない場合、時間設定情報に含まれる投票時間中に、ノードから送信された投票要求を、投票ノードとしての少なくとも２つのノードにより、受信するステップと；
リーダーノードを決定するために、前記ノードが前記票集計時間内に投票情報に含まれる前記票の数を集計できるよう、投票要求に従って前記投票情報を、前記投票ノードにより、生成し、前記投票情報を前記投票時間中に前記ノードへ送信するステップと；を更に備える、
請求項１１に記載の非一時的なコンピュータ読取可能媒体。
【請求項１６】
コンピュータにより実施されるシステムであって：
１つまたは複数のコンピュータと；
前記１つまたは複数のコンピュータと相互運用可能に結合された１つまたは複数のコンピュータメモリデバイスであって、前記１つまたは複数のコンピュータによって実行されると：
ホストにより、全てのノードに対する、同一の投票時間と、同一の票集計時間と、同一のリーダーノードの任期とを設定するステップと；
設定された前記同一の投票時間と、前記同一の票集計時間と、前記同一のリーダーノードの任期とを含む時間設定情報を、全ての前記ノードへ送信するステップであって、前記時間設定情報を受信した前記ノードは、前記同一の投票時間中に投票し、前記同一の票集計時間中に票の数を集計し、票集計結果に従ってリーダーノードを選出するように動作し、且つ前記ノードは、前記同一のリーダーノードの任期に従って周期的なノード選出を実行する、ステップと；を備える操作を実行させる命令を格納する、有形で非一時的な機械読取可能媒体を備える、前記コンピュータメモリデバイスと；を備える、
コンピュータにより実施されるシステム。
【請求項１７】
前記操作は：
全ノードのうちにリーダーノードが存在しない場合、前記時間設定情報に含まれる投票時間中に、投票要求を、少なくとも１つのノードにより、前記全ノードへ送信するステップと；
他のノードが前記投票要求に従って生成した投票情報を、ノードのそれぞれにより、受信するステップと；
前記時間設定情報に含まれる票集計時間中に、受信した前記投票情報に含まれる前記票の数を、前記ノードのそれぞれにより、集計するステップと；
特定のサービスを処理するために、集計したそれぞれの投票情報に含まれる前記票の数に従ってリーダーノードを、前記ノードにより、決定するステップと；を更に備える、
請求項１６に記載のコンピュータにより実施されるシステム。
【請求項１８】
前記操作は、前記集計するステップを開始する前に、前記投票情報の受信を終了するス

テップを更に備える、
請求項１７に記載のコンピュータにより実施されるシステム。
【請求項１９】
前記集計したそれぞれの投票情報に含まれる前記票の数に従ってリーダーノードを、前記ノードにより、決定するステップは、前記ノードが、集計した前記投票情報に含まれる前記票の数が予め設定された閾値を超えると判断したときに、前記ノードにより、当該ノード自体をリーダーノードとして決定するステップを備え、前記閾値は全ての前記ノードの数の１／２である、
請求項１７に記載のコンピュータにより実施されるシステム。
【請求項２０】
前記操作は：
全ノードのうちにリーダーノードが存在しない場合、時間設定情報に含まれる投票時間中に、ノードから送信された投票要求を、投票ノードとしての少なくとも２つのノードにより、受信するステップと；
リーダーノードを決定するために、前記ノードが前記票集計時間内に投票情報に含まれる前記票の数を集計できるよう、投票要求に従って前記投票情報を、前記投票ノードにより、生成し、前記投票情報を前記投票時間中に前記ノードへ送信するステップと；を更に備える、
請求項１６に記載のコンピュータにより実施されるシステム。
【発明の詳細な説明】
【技術分野】
【０００１】
本願はコンピュータ技術の分野に関し、特に時間ベースのノード選出方法及び装置に関する。
【背景技術】
【０００２】
情報技術の発達にともない、インターネットは多彩で膨大な情報をユーザに提供できるようになってきている。しかし、ネットワークサービスプロバイダは、シングルサーバやデータベースといった従来の処理装置では、膨大なデータ処理の要求にもはや対応できない。そのために登場したのが分散システムである。
【０００３】
一般に、分散システムに含まれる複数の処理装置（サーバ、プロセッサ、データベースなど）は、システム内の様々なデータリソースに対して、スケジューリング、分散処理、連携処理、リモート制御といった複数の種類の管理操作をグローバルに実行してもよい。分散システム内の各処理装置は分散ノードとみなされる（略して「ノード」と呼ぶ）。これにより、分散システムは、膨大なデータをシステム内の各ノードに「分散」して各ノードが対応のデータを処理できるようになるので、データ処理効率は効果的に高まり、データスループットは著しく増加する。
【０００４】
分散システムにおいて、種々のノード内でデータリソースのスケジューリングの効率をさらに高めるには、特殊タイプの情報を同一のノードにて均等に記憶する必要がある。ここで、特殊タイプの情報とは、例えば、データの記憶位置、更新状態、検索キーワード、及びその他の情報を含むメタ情報（メタ情報は、データリソースの記述情報）である。
【０００５】
従来技術では、メタ情報を維持及び管理するために、ノード選出によって適切なノード（リーダーノードとも呼ばれる）が選出される。ノードは具体的に次のようにして選出される。すなわち、種々のノードが他のノードへ投票要求を送信し、各ノードは投票要求を受信すると投票を行い、最終的には、集計された票数に従ってリーダーノードを選出する。
【０００６】

しかし、上記のノード選出方式を採用した場合、種々のノードのそれぞれが大量の処理対象データを保持することになり、或るノードがアイドル状態になり他のノードへ投票要求を送信しても、他のノードはデータ処理中であって、或るノードから送信された投票要求を受信しても即座に投票状態に入れない場合がある。つまり、各ノードの投票及び集計のタイミングが不明確であると、ノード選出をタイムリかつ効率的に実行することは不可能になる。

【発明の概要】

【発明が解決しようとする課題】

【０００７】

本出願の実施の形態は、既存のノード選出における低い効率性の問題を解決するために、時間ベースのノード選出方法及び装置を提供する。

【課題を解決するための手段】

【０００８】

本願の実施の形態は、時間ベースのノード選出方法を提供し、当該方法は分散システム内でリーダーノードを決定するために用いられ、当該方法は：

同一の投票時間と、同一の票集計時間と、同一のリーダーノードの任期とを、ホストにより、全ノードに設定するステップと；

設定された投票時間と、票集計時間と、リーダーノードの任期とを含む時間設定情報を、全てのノードへ送信し、それにより、ノードが投票時間内に投票し、票集計時間内に票の数を集計し、票集計結果に従ってリーダーノードを選出し、そして、ノードはリーダーノードの任期に従って周期的なノード選出を実行できるようになされる、ステップと；を備える。

【０００９】

本願の実施の形態は、時間ベースのノード選出方法を更に提供し、当該方法は分散システム内でリーダーノードを決定するために用いられ、当該方法は：

全ノードのうちにリーダーノードが存在しない場合、時間設定情報に含まれる投票時間内に、投票要求を、キャンペーンノードとしての少なくとも１つのノードにより、全ノードへ送信するステップと；

他のノードが投票要求に従って生成した投票情報を、各キャンペーンノードにより、受信するステップと；

時間設定情報に含まれる票集計時間内に、受信した投票情報の量を、キャンペーンノードのそれぞれにより、集計するステップと；

特定のサービスを処理するために、集計したそれぞれの投票情報の量に従ってリーダーノードを、キャンペーンノードにより、決定するステップと；を備える。

【００１０】

本願の実施の形態は、時間ベースのノード選出方法を更に提供し、当該方法は分散システム内でリーダーノードを決定するために用いられ、当該方法は：

全ノードのうちにリーダーノードが存在しない場合、時間設定情報に含まれる投票時間内に、キャンペーンノードから送信された投票要求を、投票ノードとしての少なくとも２つのノードにより、受信するステップと；

リーダーノードを決定するために、キャンペーンノードが票集計時間内に投票情報の量を集計できるよう、投票要求に従って投票情報を、投票ノードにより、生成し、投票情報を投票時間内にキャンペーンノードへ送信するステップと；を備える。

【００１１】

本願の実施の形態は、時間ベースのノード選出方法を更に提供し、当該方法は分散システム内でリーダーノードを決定するために用いられ、当該方法は：

全ノードのうちにリーダーノードが存在し、リーダーノードが予め設定された基準を自己が満たしている、と判断した場合に、投票時間内に、時間設定情報に含まれるリーダーノードの任期に従って、再任要求を、リーダーノードにより、全ノードへ送信するステップと；

再任要求に従って、他のノードが生成した投票情報を、リーダーノードにより、受信するステップと；
票集計時間内に、受信した投票情報の量を、リーダーノードにより、集計するステップと；
集計した投票情報の量に従って再任状態を決定し、再任成功後に特定のサービスを、リーダーノードにより、処理するステップと；を備え、
予め設定された基準は、負荷処理基準とデータ更新度とのうち少なくとも一方を備える。

【００１２】
本願の実施の形態は、時間ベースのノード選出方法を更に提供し、当該方法は分散システム内でリーダーノードを決定するために用いられ、当該方法は：
全ノードのうちにリーダーノードが存在し、リーダーノードが、予め設定された基準を自己が満たしていないと判断した場合に、時間設定情報含まれるリーダーノードの任期に従って、後任ノードとして機能するための予め設定された基準を満たすノードを、リーダーノードにより他のノードの中から選別するステップと；
後任ノードが、他のノードが送信した投票情報を受信し、票集計時間内に投票情報の量を集計し、新たなリーダーノードとして決定されるように、投票期間内に、後任ノードについての投票要求を、リーダーノードにより、他のノードへ送信するステップと；を備える。

【００１３】
本願の実施の形態は、時間ベースのノード選出方法を更に提供し、当該方法は分散システム内でリーダーノードを決定するために用いられ、当該方法は：
全ノードのうちにリーダーノードが存在する場合、時間設定情報に含まれる投票時間内に、リーダーノードが送信した再任要求を、投票ノードとしての少なくとも２つのノードにより、受信するステップと；
リーダーノードが票集計時間内に投票情報の量を集計して、再任状態を決定できるように、再任要求に従って投票情報を生成し、時間設定情報に含まれる投票時間内に、投票情報を、投票ノードにより、リーダーノードへ送信するステップと；を備え、
投票ノードは、リーダーノードの在任期間中に受信した投票要求を処理しない。

【００１４】
本願の実施の形態は、時間ベースのノード選出方法を更に提供し、当該方法は分散システム内でリーダーノードを決定するために用いられ、当該方法は：
全ノードのうちにリーダーノードが存在する場合、時間設定情報に含まれる投票時間内に、リーダーノードから送信された後任ノードのための投票要求を、投票ノードとしての少なくとも２つのノードにより、受信するステップと；
後任ノードが票集計時間内に投票情報の量を集計して、新たなリーダーノードを決定できるように、投票要求に従って投票情報を生成し、時間設定情報に含まれる投票時間内に、投票情報を、投票ノードにより後任ノードへ送信するステップと；を備える。

【００１５】
本願の実施の形態は、時間ベースのノード選出装置を更に提供し、当該装置は分散システム内でリーダーノードを決定するように構成され、当該装置は：
全ノードに同一の投票時間と、同一の票集計時間と、同一のリーダーノードの任期とを設定するように構成された設定モジュールと；
ノードが、投票時間内に投票し、票集計時間内に票数を集計し、票集計結果に従ってリーダーノードを選出できるように、そして、ノードがリーダーノードの任期に従って周期的なノード選出を実行できるように、投票時間と、票集計時間と、リーダーノードの任期とが含まれる時間設定情報を全ノードへ送信するように構成された送信モジュールと；を備える。

【００１６】
本願の実施の形態は、時間ベースのノード選出装置を更に提供し、当該装置は分散シス

テム内でリーダーノードを決定するように構成され、全ノードのうちにリーダーノードが存在しない場合に、当該装置は：
時間設定情報に含まれる投票時間内に、投票要求を全ノードへ送信するように構成された投票要求モジュールと；
他のノードが投票要求に従って生成した投票情報を受信するように構成された受信モジュールと；
時間設定情報に含まれる票集計時間内に、受信した投票情報の量を集計するように構成された集計モジュールと；
特定のサービスを処理するために、集計したそれぞれの投票情報の量に従って、リーダーノードを決定するように構成された決定モジュールと；を備える。

【００１７】
本願の実施の形態は、時間ベースのノード選出装置を更に提供し、当該装置は分散システム内でリーダーノードを決定するように構成され、及び、投票ノード内に配置され、全ノードのうちにリーダーノードが存在しない場合に、当該装置は：
時間設定情報に含まれる投票時間内に、キャンペーンノードが送信した投票要求を受信するように構成された受信モジュールと；
キャンペーンノードが票集計時間内に投票情報の量を集計し、リーダーノードを決定できるように、投票要求に従って投票情報を生成し、投票時間内に投票情報をキャンペーンノードへ送信するように構成された投票情報モジュールと；を備える。

【００１８】
本願の実施の形態は、時間ベースのノード選出装置を更に提供し、当該装置は分散システム内でリーダーノードを決定するように構成され、及び、リーダーノード内に配置され；全ノードのうちにリーダーノードが存在し、リーダーノードが予め設定された基準を自己が満たしていると判断した場合に、当該装置は：
時間設定情報に含まれるリーダーノードの任期に従って、投票時間内に、再任要求を全てのノードへ送信するように構成された再任要求モジュールと；
他のノードが再任要求に従って生成した投票情報を受信するように構成された受信モジュールと；
票集計時間内に、受信した投票情報の量を集計するように構成された集計モジュールと；
集計した投票情報の量に従って再任状態を決定し、再任の成功後に特定のサービスを処理するように構成された決定モジュールと；を備え、
予め設定された基準は、負荷処理基準とデータ更新度のうち少なくとも一方を備える。

【００１９】
本願の実施の形態は、時間ベースのノード選出装置を更に提供し、当該装置は分散システム内でリーダーノードを決定するように構成され、及び、リーダーノード内に配置され；全ノードのうちにリーダーノードが存在し、リーダーノードが予め設定された基準を自己が満たしていないと判断した場合に、当該装置は：
時間設定情報に含まれるリーダーノードの任期に従って、後任ノードとして機能するために予め設定された基準を満たすノードを他のノードの中から選別するように構成された選別モジュールと；
後任ノードが、他のノードが送信した投票情報を受信し、票集計時間内に投票情報の量を集計し、新たなリーダーノードとして決定されるように、投票期間内に後任ノードのための投票要求を他のノードへ送信するように構成された投票要求モジュールと；を備える。

【００２０】
本願の実施の形態は、時間ベースのノード選出装置を更に提供し、当該装置は分散システム内でリーダーノードを決定するように構成され、及び、投票ノード内に配置され、全ノードのうちにリーダーノードが存在する場合に、当該装置は：
時間設定情報に含まれる投票時間内に、リーダーノードが送信した再任要求を受信する

ように構成された受信モジュールと；
リーダーノードが、票集計時間内に投票情報の量を集計し、再任状態を決定できるように、再任要求に従って投票情報を生成し、時間設定情報に含まれる投票時間内に投票情報をリーダーノードに送信するように構成された投票情報モジュールと；を備える。
【００２１】
本願の実施の形態は、時間ベースのノード選出装置を更に提供し、当該装置は分散システム内でリーダーノードを決定するように構成され、及び、投票ノード内に配置され、全ノードのうちにリーダーノードが存在する場合に、当該装置は：
時間設定情報に含まれる投票時間内に、リーダーノードが送信した後任ノードのための投票要求を受信するように構成された受信モジュールと；
新たなリーダーノードを決定するため、後任ノードが票集計時間内に投票情報の量を集計できるように、投票要求に従って投票情報を生成し、時間設定情報に含まれる投票時間内に投票情報を後任ノードへ送信するように構成され、及び、リーダーノードの在任期間中に受信した投票要求を処理しないように構成された、投票情報モジュールと；を備える。
【００２２】
本出願の実施の形態は、時間ベースのノード選出方法及び装置を提供する。この方法により、ホストは、同一の投票時間及び同一の票集計時間を全ノードに設定し、設定した投票時間及び票集計時間を含む時間設定情報を全ノードへ送信することで、ノードが投票時間内に投票できるようにし、さらに、票集計時間内に票数を集計し、票集計結果に従ってリーダーノードを選出する。キャンペーンノードは、ホストが設定した同一の投票時間及び同一の票集計時間に従って、投票操作及び票集計操作を同期的に実行することで、各ノードが投票及び票集計を行うタイミングが不明確であるという従来技術の欠陥を克服し、これにより、ノード選出処理においてリーダーノードを決定する効率が著しく向上する。
【図面の簡単な説明】
【００２３】
ここで説明する添付の図面は、本願のさらなる理解を提供するために用いられるとともに、本願の一部を構成する。本願の概略的な実施の形態及びその説明は本願を説明するために用いられているが、本願へのいかなる不適切な限定も構成するものではない。
【００２４】
【図１】本出願の実施の形態による時間ベースのノード選出工程を示す。
【図２】本出願の実施の形態による時間ベースのノード選出工程を示す。
【図３】図３ａ、図３ｂ、図３ｃ、図３ｄは、本出願の実施の形態による投票時間及び票集計時間の略図である。
【図４】本出願の一の実施の形態による特定のアプリケーション例において、ノード選出のために各ノードが投票情報を送信する様子を示す略図である。
【図５】本出願の実施の形態による時間ベースのノード選出工程を示す。
【図６】本出願の実施の形態による時間ベースのノード選出工程を示す。
【図７】本出願の実施の形態による時間ベースのノード選出工程を示す。
【図８】本出願の実施の形態による時間ベースのノード選出工程を示す。
【図９】本出願の実施の形態による時間ベースのノード選出工程を示す。
【図１０】本出願の実施の形態による時間ベースのノード選出装置の構造略図である。
【図１１】本出願の実施の形態による時間ベースのノード選出装置の構造略図である。
【図１２】本出願の実施の形態による時間ベースのノード選出装置の構造略図である。
【図１３】本出願の実施の形態による時間ベースのノード選出装置の構造略図である。
【図１４】本出願の実施の形態による時間ベースのノード選出装置の構造略図である。
【図１５】本出願の実施の形態による時間ベースのノード選出装置の構造略図である。
【図１６】本出願の実施の形態による時間ベースのノード選出装置の構造略図である。
【発明を実施するための形態】
【００２５】

本願の目的、技術的な解決策、及び利点を明確にするために、本願の技術的な解決策を本願の特定の実施の形態及びこれと対応する添付図面を参照しながら明瞭かつ完全に以下に説明する。説明する実施の形態は、明らかに、本願の実施の形態の全てではなく単にそのいくつかでしかない。本願の実施の形態に基づき、当業者によって創造的な努力なく得られるその他の全ての実施の形態は、本願の保護範囲に全て含まれるものとする。

【００２６】

図１は、本出願の一の実施の形態による時間ベースのノード選出方法を示す。この方法は分散システム内でリーダーノードを決定するために用いられ、具体的に以下のステップを含む。

【００２７】

ステップＳ１０１：　ホストが全ノードに対して、同一の投票時間と、同一の投票集計時間と、同一のリーダーノードの任期とを設定する。

【００２８】

本出願のこの実施の形態では、分散システム内に１つのホストが配置されており、ホストは分散システム内のノードを管理する（例えば、データのスケジューリング、データの設定、特定指示の送信及びその他を行う）ように構成されている。ホストは、コンピュータ又はサーバといった、監視機能付きネットワーク装置を含むが、これに限定されない。

【００２９】

従来技術では、分散システム内のノードがノード選出を行うとき、種々のノードが投票要求を異なる時間に送信し、また投票要求に応答するタイミングも不明確である。そのため、本発明のこの実施の形態において、分散システム内のホストは、種々のノードが同一タイプの処理操作を同時に実行できるように、分散システム内の種々のノードに時間設定を行う。つまり、ステップＳ１０１では、ホストが種々のノードに同一の投票時間及び同一の票集計時間を設定する。

【００３０】

本出願のこの実施の形態における投票時間及び票集計時間のそれぞれは、（特定の瞬間における時間点ではなく）時間期間である。投票時間及び票集計時間の期間の長さ（持続時間）は、ミリ秒（ｍｓ）、秒（ｓ）、分（ｍｉｎ）、時（ｈ），日（Ｄ）、週、月（ｍｏｎ）のような種々の時間単位で形成された持続時間を含むが、これに限定されない。

【００３１】

加えて、ホストは、全ノードにリーダーノードの任期を設定する。任意のノードがリーダーノードとして選出され、設定された任期を終えると、次のノード選出工程が始められる。

【００３２】

ステップＳ１０２：設定された投票時間と、票集計時間と、リーダーノードの任期とを含む時間設定情報が全ノードへ送信され、これにより、ノードがこの投票時間内に投票し、票集計時間内に票数を集計し、票集計結果に従ってリーダーノードを選出することで、各ノードは、リーダーノードの任期に従って周期的なノード選出を実行できるようになる。

【００３３】

投票時間と、票集計時間と、リーダーノードの任期とを設定したら、ホストは、設定された投票時間と、票集計時間と、リーダーノードの任期とに従って、対応する操作を各ノードが実行できるようにするために、各ノードにその旨を通知する必要がある。そのため、ホストは、投票時間と、票集計時間と、リーダーノードの任期とを含む時間設定情報を各ノードへ送信できる。そうすることで、各ノードが別々に時間設定情報を受信した後に、受信した時間設定情報に従って内蔵クロックを各ノードが調整することにより、種々のノードが投票時間及び票集計時間を同時に起動し、対応する操作を同期的に実行できるようになる。

【００３４】

ここで、ホストが設定した投票時間と、票集計時間と、リーダーノードの任期とは周期

的である点にさらに留意すべきである。その理由は以下の通りである。すなわち、ホストが設定した投票時間と票集計時間とに従って投票操作と票集計操作とをノードが別々に実行して、リーダーノードを決定した後、ノードとリーダーノードが対応のサービスを処理する。しかしながら、実際のアプリケーションの間、分散システム内で処理される膨大なデータが、リーダーノードと他のノードとの処理負荷とデータ更新度とを変化させる可能性があるので、過負荷、データ更新の遅延、リーダーノードのクラッシュなどの状況が引き起こされ、リーダーノードによる特定のサービスの処理の効率に重大な影響を及ぼしかねない。このような場合、ホストが設定した投票時間と、票集計時間と、リーダーノードの任期とが周期的でなければ（つまり、分散システム内のノードがノード選出を１度だけ実行し、リーダーノードが失敗した後はノード選出を実行しない場合）、分散システム全体での処理が停止する。ノード選出は、ホストが投票時間と票集計時間とを再設定するまで再び実行することができないため、これも分散システムのサービス処理効率を低下させる。

【００３５】

したがって、本出願のこの実施の形態では、ホストが設定した投票時間と、票集計時間と、リーダーノードの任期は周期的に繰り返される。例えば、投票時間の間隔周期は３０００ミリ秒である、つまり、第１の投票時間終了から３０００ミリ秒後に第２の投票時間が到来する。別の例では、リーダーノードの任期は１０分間である、つまり、或るノードがリーダーノードとして選出され、リーダーノードとして１０分間機能した後に、次のノード選出処理が始まる。当然ながら、実際のアプリケーションにおいて、投票時間及び票集計時間の時間周期は実際のアプリケーションの必要性に従って調整又は設定され、ここで説明した内容は本出願を限定するものではない。

【００３６】

ホストは、上記ステップＳ１０１乃至Ｓ１０２の方法を用いて、分散システム内の個々のノードに同一の時間を設定することにより、これらのノードが対応する投票操作と票集計操作とを同一時間で実行できるようになり、そしてもともと非同期なノードが同一タイプの操作を同期的に実行できるようになり、その結果、リーダーノードの選出の効率は効果的に高まる。

【００３７】

本出願のこの実施の形態では、ホストが同一の投票時間と、同一の票集計時間と、同一のリーダーノードの任期とを設定した後に、種々のノードが、対応する時間設定情報に従って対応する投票操作又は投票集計操作を実行できる。具体的には、分散システム内のノードに関連して、本出願の一の実施の形態は時間ベースのノード選出方法を提供する。図２に示すように、この方法は以下のステップを含む。

【００３８】

ステップＳ２０１：全ノードのうちにリーダーノードがない場合、キャンペーンノードとしての少なくとも１つのノードが、時間設定情報に含まれる投票時間内に、全ノードに対して投票要求を送信する。

【００３９】

分散システムがメンテナンス又は更新などの状態に入った後には、分散システム内のノードにリーダーノードは存在しない。この場合、ノードはリーダーノードを選出することができる。実際のアプリケーションシナリオでは、ノード選出は少なくとも２つのノードが存在するときに実行でき、ノード選出中には、１つのノードが投票要求を送信できる（ここでは、このノード自体と別のノードへ投票要求が送信される）。したがって、本出願のこの実施の形態のステップＳ２０１では、投票のためのキャンペーンノードとして機能するノードが少なくとも２つ存在する。当然ながら、実際のアプリケーションシナリオでは、ノードが複数ある場合、各ノードはキャンペーンノードとして投票要求を他の各ノードに送信できる。

【００４０】

従来技術との違いは、本出願のこの実施の形態では、キャンペーンノードが、ホストが

設定した投票時間に従って投票する点である。つまり、キャンペーンノードはホストが送信した時間設定情報に従って自己の内蔵クロック信号を設定し、キャンペーンノードの内蔵クロックが設定投票時間に達すると、キャンペーンノードが投票要求の送信操作を実行する。

【００４１】

この投票要求には、他のノードが投票選出を行うための、複数のノード状態（例えば、キャンペーンノードのデータ処理容量及びデータ更新度）を反映するノード情報（例えば、キャンペーンノードのノード識別子及びノード状態）が含まれている。

【００４２】

ステップＳ２０２：各キャンペーンノードは、他のノードが投票要求に従って生成した投票情報を受信する。

【００４３】

キャンペーンノードが送信した投票要求を受信すると、他のノードは、この投票要求に従って対応の投票情報を生成し、この投票情報を対応のキャンペーンノードへ送信できる。

【００４４】

当然ながら、ステップＳ２０２では、各キャンペーンノードによる、投票情報を受信する処理も投票時間内に行われる。

【００４５】

ステップＳ２０３：各キャンペーンノードは、時間設定情報に含まれる票集計時間内に、受信した投票情報の量を集計する。

【００４６】

本出願のこの実施の形態では、各キャンペーンノードは、ホストから受信した時間設定情報に従って自己の内蔵クロックを設定し、キャンペーンノード内のクロックが設定された票集計時間に達すると、キャンペーンノードが票集計操作を実行する。

【００４７】

ステップＳ２０４：キャンペーンノードは、特定のサービスを処理するために、集計したそれぞれの投票情報の量に従ってリーダーノードを決定する。

【００４８】

ステップＳ２０４について、キャンペーンノードがリーダーノードを決定する工程は、票集計時間が終了する前に完了する。さらに、１つのキャンペーンノードがリーダーノードとして決定された後に、このリーダーノードは、リーダーノードが生成された旨を知らせる通知メッセージを他のノードへ送信する。リーダーノードは、例えば分散システムにおけるメタデータの維持及び管理といった特定のサービスの処理を行う。

【００４９】

本出願のこの実施の形態では、最終的に決定したリーダーノードは一意である点に留意されたい。これにより、特定のサービスがリーダーノードを用いて処理されることで、特定のサービスの検索、割り当て、処理効率を向上することができる一方、特定のサービスを処理するために複数のリーダーノード間で生じる競合によって起こるサービスの衝突を回避することもできる。

【００５０】

上記各ステップにより、キャンペーンノードは、ホストが設定した同一の投票時間と同一の票集計時間とに従って投票操作と票集計操作とを同期的に実行し、これにより、各ノードの投票及び票集計のタイミングが不明確であるという従来技術の欠陥が克服され、ノード選出工程においてリーダーノードを決定する効率は著しく向上する。

【００５１】

本出願のこの実施の形態では、ホストが設定した投票時間と票集計時間とが重なることはないため、分散システム内のノードは投票時間内には投票操作のみを実行し、票集計時間内には票集計操作のみを実行することができる点に留意されたい。例えば、図３ａに示す状況で、キャンペーンノードは、時間設定情報に従って、対応する投票時間と票集計時

間とを自己の内蔵クロックに設定する。つまり、図３ａでは、１０００ｍｓから２０００ｍｓまでの時間が投票時間であり、３５００ｍｓから４５００ｍｓまでの時間が票集計時間である。言い換えれば、キャンペーンノードの内蔵クロックが１０００ｍｓに達すると、各キャンペーンノードが投票時間に入り、他のノードに投票情報を送信し、１０００ｍｓの投票時間内に、他のノードが送信した投票情報を受信する。クロックが２０００ｍｓに達すると投票は停止する。同様に、票集計時間（３５００ｍｓから４５００ｍｓ）中に、キャンペーンノードが受信した投票情報の量を集計する。クロックが４５００ｍｓに達すると、キャンペーンノードは票集計を停止し、集計した投票情報の量に従ってリーダーノードを決定する。このような方式により、種々のキャンペーンノードが各設定時間において種々の操作を同期的に実行するようになる。

【００５２】

図３ｂに示す状況では、ホストが設定した投票時間と票集計時間との間には間隔がない。言い換えれば、各キャンペーンノードは、クロックが２０００ｍｓに達すると、投票状態から票集計状態へと直接切り替わる。

【００５３】

しかし、図３ｂに示す状況では、ノード間で投票情報を送信するための期間を設ける必要がある。例えば、ノードは、生成した投票情報をキャンペーンノードへ送信するために１０ｍｓを設ける必要があり、この場合、キャンペーンノードのクロックは投票情報の送信中に既に２０００ｍｓに達する、つまり、キャンペーンノードは票集計状態に入る。キャンペーンノードが票集計状態に入ってからも投票情報を受信してしまうと、キャンペーンノードが投票情報の量を集計する際に、最終的な集計済み投票情報の量の正確性に悪影響を及ぼす。

【００５４】

したがって、本出願のこの実施の形態では、この方法は、キャンペーンノードが、票集計時間内に受信した投票情報の量を集計する前に、票集計時間が到来すると、キャンペーンノードが投票情報の受信を停止することを更に含む。

【００５５】

これにより、図３ｂに示す時間設定において、キャンペーンノードのクロックが２０００ｍｓに達し、票集計状態に入ると、キャンペーンノードは、他のノードが送信した投票情報の受信を停止する。したがって、この方式によれば、投票情報の量を正確に集計することができる。

【００５６】

加えて、本出願のこの実施の形態では、リーダーノードを決定するためにキャンペーンノードが受信した投票情報の量の集計を行う際に、各キャンペーンノードがそれぞれの投票情報の量を個別に集計することを考慮すると、集計した投票情報の量に従ってキャンペーンノード自体がリーダーノードになるかどうかをキャンペーンノードが直接決定できるようにするために閾値を予め設定することができ、また、この閾値及び集計した投票情報の量に従って、キャンペーンノードがリーダーノードになれるかどうかが決定される。

【００５７】

したがって、本出願のこの実施の形態では、具体的には、ステップＳ２０４にて、キャンペーンノードが、集計したそれぞれの投票情報の量に従ってリーダーノードを決定することは、集計した投票情報の量が予め設定された閾値を超える、とキャンペーンノードが判断するときに、キャンペーンノードが自己をリーダーノードとして決定することを含む。閾値は全ノード数の１／２である。

【００５８】

ここでは、分散システム内の各ノードは、分散システム内の全ノードの総数を既に知得している点に留意されたい（本出願の一の実施において、ホストは、分散システム内の全ノードの総数について各ノードに送信してよい）。したがって、ノードの投票情報の量が全ノードのデータ量の１／２を超えた場合に、そのノードをリーダーノードであると一意に決定することができる。

【００５９】
例えば、図４に示すように、分散システムは６つのノード、つまりノードＡ乃至ノードＦを含む。
【００６０】
投票時間内に、分散システム内の６つ全てのノードがキャンペーンノードとして機能して、投票要求を他のノードに送信し（図４では、投票要求送信工程は不図示）、投票情報を受信する。この実施例では、各ノードが投票要求の受信に成功し、各ノードは投票要求を自己へは送信しないと仮定する。
【００６１】
この場合、６つのノードについて、各ノードが自己を除くその他のノードへ投票要求を送信する（実際のアプリケーションにおいて、ノードは投票要求を自己へも送ることができるが、この実施例では、説明を容易にする目的で、各ノードは自己を除くその他のノードへのみ投票要求を送信すると仮定する）。具体的に、ノードＡは投票要求を他の５つのノード（ノードＢ乃至ノードＦ）へ送信する。同様に、ノードＢ乃至ノードＧの各々が、自己を除く他の５つのノードへ投票要求を送信する。これにより、投票要求を受信する各ノードが投票情報を生成し、この投票情報を対応するキャンペーンノードへ送信する。
【００６２】
したがって、この実施例では、各ノードは、投票に参加しているノード数が６つであると決定することができ、閾値を３に設定する。
【００６３】
図４に示すように、ノードＡは生成した投票情報をノードＢへ送信すると仮定する。ノードＢは、生成した投票情報をノードＣへ送信する。ノードＣ、ノードＤ、ノードＥ、ノードＦは全て、それぞれ生成した投票情報をノードＡへ送信する。
【００６４】
票集計時間内に、各ノードは自らが受信した投票情報の量を集計する。具体的には、ノードＡは集計によって受信投票情報の量４を獲得し、ノードＢは集計によって受信投票情報の量１を獲得し、ノードＣは集計によって受信投票情報の量１を獲得し、残りのノードの投票情報の量は０である。
【００６５】
ノードＡ内の投票情報の量は予め設定された閾値を超えるので、ノードＡは自らをリーダーノードとして決定する。次に、ノードＡは、ノードＡがリーダーノードになった旨を他のノードに通知するために、他のノードへ一斉同報を送信する。
【００６６】
しかし、実際のアプリケーションにおいて、各ノードはノード選出中に「自主的に」投票を行うため、各キャンペーンノードが受信した投票情報の量が予め設定された閾値を超えない可能性がある。引き続き上の例について述べる。ノードＦが生成した投票情報はノードＡへ送信されないが、ノードＣへは送信されるとすると、この場合、ノードＡ内の投票情報の量は３であり、ノードＢ内の投票情報の量は１であり、ノードＣ内の投票情報の量は２である。明らかに、３つ各々のノード内の投票情報は予め設定された閾値（つまり３）を超えていない。そのため、どのノードも自らをリーダーノードと決定しない。言い換えれば、どのノードも、特定期間内にリーダーノードの通知メッセージを受信しない。これは、現在のノード選出においてリーダーノードが選出されていないことを示す。
【００６７】
本出願のこの実施の形態では、リーダーノードが選出されない場合、ノードはリーダーノードが決定するまでノード選出を繰り返し実行できる。
【００６８】
前出の実施例を参照すると、本出願のこの実施の形態の好ましい方式として、ホストがノードに設定する投票時間及び票集計時間は周期的であり、票集計時間はさらに詳細に種々のサブ時間に分割することができる。図３ｃに示すように、１周期における投票時間及び票集計時間について、票集計時間は、より詳細には集計時間と通知時間とに分割される

。次に、集計時間内に、各ノードは受信した投票情報の量を集計し、通知時間内にリーダーノードからの通知を受信しなかった場合、ノードは次のノード選出を実行できる。

【００６９】

図３ｄに示すように、本出願のこの実施の形態では、ホストによって割り当てられる複数の選出時間（各選出時間は投票時間及び票集計時間を含む）は周期的である。第１の選出時間内にノードがリーダーノードを決定しない場合、第２の選出時間内に、リーダーノードが選出されるまでノード選出を繰り返し実行する。当然ながら、リーダーノードが選出された後、ノード選出は、時間設定情報に含まれるリーダーノードの任期に従って周期的に実行される。その特定の工程について後に詳細に説明する。そのため、ここでは詳述しない。

【００７０】

図２に示す方法のステップＳ２０１乃至Ｓ２０４は、リーダーノードを決定するために、キャンペーンノードが他の各ノードへ投票要求を送信し、投票情報を受信する工程である。分散システム内の各ノードは、キャンペーンノードが送信した投票要求を受信し、この投票要求に従って対応の投票情報を生成する投票ノードとしても機能できる（或る１つのキャンペーンノードが、他のキャンペーンノードが送信した投票要求を受信することは可能であるが、この場合には、この１つのキャンペーンノードはさらに投票ノードとみなしてもよい）。したがって、投票情報を生成するノードに関して、本出願の一の実施の形態は、時間ベースのノード選出方法をさらに提供する。図５に示すように、この方法は、分散システムにおいてリーダーノードを決定するために用いられ、具体的に以下のステップを含む。

【００７１】

ステップＳ５０１：全ノードの中にリーダーノードが存在しない場合、時間設定情報に含まれる投票時間内に、投票ノードとしての少なくとも２つのノードがキャンペーンノードから送信された投票要求を受信する。

【００７２】

本出願のこの実施の形態では、リーダーノードが決定されない場合、分散システム内の各ノードが投票ノードとして機能して、キャンペーンノードが送信した投票要求を受信することができる。

【００７３】

同様に、投票ノードは、さらに、ホストが送信した時間設定情報を受信し、投票ノード内のクロックを調整し、投票時間及び票集計時間を設定することもできる。

【００７４】

ステップＳ５０２：キャンペーンノードが票集計時間内に投票情報の量を集計し、リーダーノードを決定することができるように、投票ノードは、投票要求に従って投票情報を生成し、投票時間内にこの投票情報をキャンペーンノードへ送信する。

【００７５】

分散システムにおいて、投票ノードは少なくとも１つのキャンペーンノードが送信した投票要求を受信できる（実際のアプリケーションシナリオでは、投票ノードは、通常、複数のキャンペーンノードが送信した投票要求を受信する）。最終的に唯一のリーダーノードが選出されるように、投票ノードは１つのキャンペーンノードに投票する。つまり、投票ノードは、受信した投票要求に従ってキャンペーンノードを１つ選出し、選出されたノードについての投票情報を生成し、この投票情報を選出されたノードへ送信する。

【００７６】

先の一連のステップでは、分散システム内の各投票ノードに同一の投票時間が設定されているため、各投票ノードは同一の投票時間にキャンペーンノードを選出し、対応する投票要求を生成し、投票要求を、選出されたキャンペーンノードへ送信することができる。これにより、投票工程は確実に同期され、ノード選出の効率も向上する。

【００７７】

キャンペーンノードが投票ノードへ投票要求を送信する場合、投票要求には、通常、キ

ャンペーンノードのノード情報含まれることに留意すべきである。これにより、投票ノードが複数の投票要求を受信する際に、投票ノードは、投票要求に含まれるノード情報に従ってキャンペーンノードを選出できる。

【００７８】

したがって、本出願のこの実施の形態では、投票情報を生成し、投票情報をキャンペーンノードへ送信することは、具体的に、投票ノードが、受信した投票要求に含まれているノード情報を判定することと、ノード情報に従って、キャンペーンノードのノード状態重量（node state weight、ノード状態の有力性）を決定することと、最大のノード状態重量を持つキャンペーンノードを選出し、キャンペーンノードの投票情報を生成し、投票情報を、選出されたキャンペーンノードへ送信することとを含む。

【００７９】

ノード情報は、キャンペーンノードのデータ処理容量と処理状態を反映できる情報（例えば、キャンペーンノードのデータ更新情報、負荷処理情報、ハードウェア構成情報）を含むが、これに限定されない。因って、投票ノードは、種々のキャンペーンノードの包括的な状態レベルを反映するために、ノード情報に従って、種々のキャンペーンノードのノード状態重量を決定することができる。そのため、本出願のこの実施の形態では、投票ノードは、最大のノード状態重量を持つキャンペーンノードを選出し、対応する投票情報を生成し、この投票情報をキャンペーンノードへ送信でき、これにより投票工程が完了する。

【００８０】

図２に示す方法のステップＳ２０１乃至Ｓ２０４と図５に示す方法のＳ５０１乃至Ｓ５０２は両方とも、分散システム内にリーダーノードが存在しない場合、ノードが、ホストが設定した投票時間及び票集計時間に従ってリーダーノードを決定する工程である。分散システム内にリーダーノードが既に存在する場合、本出願の一実施の形態は時間ベースのノード選出方法をさらに提供する。図６に示すように、この方法は分散システム内でリーダーノードを決定するために用いられ、具体的に以下のステップを含む。

【００８１】

ステップＳ６０１：全ノードのうちにリーダーノードが存在し、リーダーノードが予め設定された基準を自己が満たしていると判断した場合に、時間設定情報に含まれるリーダーノードの任期に従って投票時間内にこのリーダーノードが全ノードに再任要求を送信する。

【００８２】

本出願のこの実施の形態では、図２に示す方法のＳ２０１乃至Ｓ２０４と、図５に示す方法のＳ５０１乃至Ｓ５０２とによってリーダーノードが決定された後、リーダの任期が終了に近づいたときに、次のノード選出を実行する必要がある。このケースでは、リーダーノード以外のノードは投票要求を送信しないが、リーダーノードだけは他のノードに再任要求を送信できる。

【００８３】

実際のアプリケーションにおいて分散システムが大量の情報を処理する際に、リーダーノードが保持するデータ（メタデータなど）の更新に遅延が生じたり、リーダーノードの処理負荷が連続して増加する状況が生じたりすることにより、結果的にサービス処理効率が低下するなどの可能性がある点に留意されたい。そのため、本出願のこの実施の形態では、次の投票時間が到来すると、現在のリーダーノードはそのノード状態が予め設定された基準を満たしているかどうかを判断する。満たしていれば、それは、リーダーノードの処理負荷及びデータ更新度が特定のサービスの処理の効率に影響を及ぼさないことを意味しており、さらに、この場合に限り、リーダーノードが他のノードに再任要求を送信する。

【００８４】

予め設定された基準には、限定されないものの、負荷処理基準、データ更新度及びそれらの類似が含まれる。種々のノードは種々のハードウェア構成（処理性能、記憶領域など

）を有することができるので、本出願のこの実施の形態では、予め設定されたノードの基準も多様であってよい。この場合、各ノードは、自己のハードウェア構成に従って、対応する予め設定された基準を生成できる。

【００８５】

通常、各ノードの予め設定された基準は、標準値形式で予め設定されてよい。例えば、或るノードについて、負荷処理基準は、中央処理装置（ＣＰＵ）の使用率が８０％を超えないというものであってよい。言い換えれば、ＣＰＵの使用が８０％を超えない場合、ノードは予め設定された基準を満たしているとみなせる。

【００８６】

別の例として、或るノードについて、データ更新度の基準は、データ更新頻度が１回の更新につき１０分以上というものであってよい（データ更新頻度はデータ更新度を反映できる）。言い換えれば、データ更新頻度が１回の更新につき１０分以上である場合、ノードは予め設定された基準を満たしているとみなせる。

【００８７】

当然ながら、上記の内容は、単に、本出願のこの実施の形態で提供される予め設定された基準の一例に過ぎず、実際のアプリケーションシナリオでは、各ノードの予め設定された基準は、実際のアプリケーションの必要性に従って設定可能である。ここで述べる内容は本出願を限定するものではない。

【００８８】

ステップＳ６０２：リーダーノードは、他の各ノードが再任要求に従って生成した投票情報を受信する。

【００８９】

再任要求を受信するノードは、対応する投票情報を再任要求に従って生成し、この投票情報をリーダーノードへ送信することができる。

【００９０】

ステップＳ６０３：票集計時間内に、リーダーノードは受信した投票情報の量を集計する。

【００９１】

Ｓ６０４：リーダーノードは、集計した投票情報の量に従って再任状態を決定し、再任の成功後に、特定のサービスを処理する。

【００９２】

本出願のこの実施形態では、リーダーノードが他のノードへ再任要求を送信する場合、分散システム内に他のキャンペーンノードは存在しない、つまり、各ノードはリーダーノードからのみ再任要求を受信する。したがって、通常、リーダーノードの再任は成功する。

【００９３】

リーダーノードの再任に成功した後に、リーダーノードの再任が成功した旨を各ノードに通知するために、リーダーノードは各ノードへ通知メッセージを送信することができる。そのため、再任に成功したリーダーノードは特定のサービスの処理を継続し、分散システム内の他のノードは通常通りサービスの処理を継続する。

【００９４】

図６に示す方法のＳ６０１乃至Ｓ６０４の場合、本出願の一の実施の形態は、時間ベースのノード選出方法をさらに提供する。図７に示すように、この方法は、分散システム内でリーダーノードを決定するために用いられ、具体的に以下のステップを含む。

【００９５】

ステップＳ７０１：全ノードのうちにリーダーノードが存在する場合、時間設定情報に含まれる投票時間内に、投票ノードとしての少なくとも２つのノードが、リーダーノードから送信された再任要求を受信する。

【００９６】

ステップＳ７０２：リーダーノードが票集計時間内に投票情報の量を集計し、再任状態

を決定できるようにするために、投票ノードが再任要求に従って投票情報を生成し、時間設定情報に含まれる投票時間内にこの投票情報をリーダーノードへ送信し、さらに、リーダーノードの任期中、投票ノードはリーダーノードが送信した投票要求のみを受信する。

【００９７】

図６に示す方法のステップＳ６０１乃至Ｓ６０４、図７に示す方法のステップＳ７０１乃至Ｓ７０２を通して、分散システム内にリーダーノードが既に存在する場合、次の投票時間が到来すると、分散システム内の全てのノードが同期的に投票状態に入ることができ、リーダーノードのみが他のノードに再任要求を送信し、再任要求を受信するノードは生成された投票要求をリーダーノードへ送信することができる。こうすることで、新たなリーダーノードを決定するために他のノードが互いに投票要求を送信し合わないため、各ノードの処理負荷を効率的に低減でき、さらにリーダーノード選出効率も向上する。さらに、リーダーノードの任期中には、投票ノードは他のノードが送信した投票要求を受信しない。これにより、各投票ノードがリーダーノードから送信された投票要求のみを受信できるようになる。

【００９８】

当然ながら、実際のアプリケーションにおいて、分散システム内の各ノードは大量のサービスを処理する。リーダーノードの処理負荷の増加、リーダーノードのデータ更新の遅れなどが生じることで、リーダーノードが予め設定された基準を満たせなくなる可能性がある。この場合、本出願の一の実施の形態は時間ベースのノード選出方法を提供する。図８に示すように、この方法は、分散システムにおいてリーダーノードを決定するために用いられ、具体的に以下のステップを含む。

【００９９】

ステップＳ８０１：全ノードのうちにリーダーノードが存在し、リーダーノードが、予め設定された基準を自己が満たしていないと判断した場合、リーダーノードは、時間設定情報に含まれるリーダーノードの任期に従って、後任ノードとして機能するための予め設定された基準を満たすノードを他のノードの中から選別する。

【０１００】

リーダーノードが、予め設定された基準を自己が満たしていないと判断した場合、それはリーダーノードの処理負荷が過剰に高い又はデータ更新度が過剰に低い可能性があることを示している。この場合、リーダーノードが特定のサービスを処理する効率に大きな影響が及ぶ。つまり、現在のリーダーノードは、その特定のサービスを処理する上でもはや適切でないということである。

【０１０１】

したがって、本出願のこの実施の形態において、リーダーノードは、任期が終了する前に、他のノードから後任ノードを選別できる。本出願のこの実施の形態における一の方式として、リーダーノードは、他のノードのノード情報について問い合わせるために他のノードへクエリ要求を送信し、ノード情報に従って後任ノードを選出できる。本出願のこの実施の形態における別の方式として、ホストは、後任ノードを選出するために、リーダーノードを指定してもよい。つまり新たなリーダーノードが選出される。次に、再選出の指示を受信した後に、現在のリーダーノードが他のノードの中から自己の後任ノードを選ぶことができる。当然ながら、この方式は本出願を限定するものではない。

【０１０２】

ステップＳ８０２：リーダーノードが、後任ノードについての投票要求を投票期間内に他のノードへ送信することで、後任ノードは、他のノードが送信した投票情報を受信し、票集計時間内に投票情報の量を集計し、そして新たなリーダーノードとして決定されることが可能になる。

【０１０３】

投票時間内にリーダーノードが他のノードへ送信した投票要求には後任ノードのノード識別子が含まれており、つまり、投票要求を受信するノードは対応する投票情報を生成し、後任ノードのノード識別子に従ってこの投票情報を後任ノードへ送信することができる

。これにより、各ノードの投票情報が後任ノードのみへ送信されるが、現在のリーダーノードへは送信されない。

【０１０４】

後任ノードのみが各ノードから送信された投票情報を受信することは明白である。そのため、後任ノードは、受信した投票情報の量を票集計時間内に集計し、リーダーノードとして決定される。後任ノードはさらに通知メッセージを各ノードへ送信し、後任ノードが新たなリーダーノードになったことを各ノードに知らせる。

【０１０５】

図８に示す方法のステップＳ８０１乃至Ｓ８０２の場合、本出願の一の実施の形態は、時間ベースのノード選出方法をさらに提供する。図９に示すように、この方法は、分散システムにおいてリーダーノードを決定するために用いられ、具体的に以下のステップを含む。

【０１０６】

ステップＳ９０１：全ノードのうちにリーダーノードが存在する場合、時間設定情報に含まれる投票時間内に、投票ノードとしての少なくとも２つのノードがリーダーノードから送信された後任ノードについての投票要求を受信する。

【０１０７】

ステップＳ９０２：投票ノードが投票要求に従って投票情報を生成し、この投票情報を、時間設定情報に含まれる投票時間内に後任ノードへ送信することで、後任ノードが票集計時間内に投票情報の量を集計し、新たなリーダーノードとして決定されることが可能になる。

【０１０８】

同様に、ステップＳ９０１乃至Ｓ９０２において、リーダーノードが送信した投票要求には後任ノードのノード識別子が含まれている。そのため、投票ノードは、投票要求に従って対応する投票情報を生成し、この投票情報を後任ノードのノード識別子に従って後任ノードへ送信するが、この投票情報を他のノードへは送信しない。これにより、後任ノードがスムーズに新たなリーダーノードになることができる。

【０１０９】

上記の方法及び実施例を通して、分散システム内にリーダーノードが存在しない場合、分散システムのホストは、ノードが投票時間内に同期的に投票できるよう、ノードに同一の投票時間と同一の集計時間とを設定し、票集計時間内に票を同期的に集計する。こうすることで、各ノードが投票又は票集計を行うタイミングの不明確さをなくすことができるので、分散システム内のノードがリーダーノードを選出する効率は大きく向上する。

【０１１０】

さらに、リーダーノードの選出後も、各ノードは周期的な投票時間と票集計時間とに従って操作を実行する。次の投票時間及び票集計時間が到来すると、リーダーノードは自己についての投票要求、又は後任ノードについての投票要求を生成するために自己の状態を判断し、これにより、他のノードが投票要求に従って投票情報を生成し、この投票情報を、対応するノードへ送信できるようになる。このような方式により、特定のサービスを処理するために選出されたリーダーノードが常に高パフォーマンス状態となり、その結果、分散システム全体のサービス処理効率が向上する。

【０１１１】

上記は、本出願の実施の形態において提供される時間ベースのノード選出方法である。これと同じ考えに基づき、本出願のこの実施の形態は、図１０に示す時間ベースのノード選出装置をさらに提供する。

【０１１２】

図１０において、時間ベースのノード選出装置は、設定モジュール１００１と、送信モジュール１００２とを備え：

設定モジュール１００１は、全ノードに同一の投票時間と、同一の票集計時間と、同一のリーダーノードの任期とを設定するように構成され、

送信モジュール１００２は、ノードが、投票時間内に投票し、票集計時間内に票数を集計し、票集計結果に従ってリーダーノードを選出できるように、そして、ノードが、リーダーノードの任期に従って周期的なノード選出を実行できるように、投票時間と、票集計時間と、リーダーノードの任期とが含まれる時間設定情報を全ノードへ送信するように構成される。

【０１１３】

図１０に示すケースでは、本出願の一の実施の形態は時間ベースのノード選出装置をさらに提供する。図１１に示すように、本装置は、分散システム内でリーダーノードを決定するように構成されており、キャンペーンノード内に配置される。

【０１１４】

図１１において、時間ベースのノード選出装置は、投票要求モジュール１１０１と、受信モジュール１１０２と、集計モジュール１１０３と、決定モジュール１１０４とを備え：

全ノードのうちにリーダーノードが存在しない場合に、

投票要求モジュール１１０１は、時間設定情報に含まれる投票時間内に、投票要求を全ノードへ送信するように構成され；

受信モジュール１１０２は、他のノードが投票要求に従って生成した投票情報を受信するように構成され；

集計モジュール１１０３は、時間設定情報に含まれる票集計時間内に、受信した投票情報の量を集計するように構成され；

決定モジュール１１０４は、特定のサービスを処理するために、集計したそれぞれの投票情報の量に従って、リーダーノードを決定するように構成される。

【０１１５】

本願の実施の形態では、受信モジュールは、更に、票集計時間が到来すると投票情報の受信を停止するように構成される。

【０１１６】

決定モジュール１１０４は、具体的に、集計した投票情報の量が予め設定された閾値を超えたと判断した場合に、キャンペーンノードをリーダーノードとして決定するように構成される。

【０１１７】

閾値は全ノードの数の１／２である。

【０１１８】

本出願の実施の形態は、時間ベースのノード選出装置をさらに提供する。図１２に示すように、本装置は、分散システム内でリーダーノードを決定するように構成されており、投票ノード内に配置される。

【０１１９】

図１２では、時間ベースのノード選出装置は、受信モジュール１２０１と、投票情報モジュール１２０２を備え：

全ノードのうちにリーダーノードが存在しない場合に、

受信モジュール１２０１は、時間設定情報に含まれる投票時間内に、キャンペーンノードが送信した投票要求を受信するように構成され；

投票情報モジュール１２０２は、キャンペーンノードが票集計時間内に投票情報の量を集計し、リーダーノードを決定できるように、投票要求に従って投票情報を生成し、投票時間内に投票情報をキャンペーンノードへ送信するように構成される。

【０１２０】

本願の実施の形態では、投票情報モジュール１２０１は、具体的に、受信した投票要求に含まれるノード情報を判定し、ノード情報に従ってキャンペーンノードのノード状態重量を決定し、最大のノード状態重量を持つキャンペーンノードを選出し、キャンペーンノードのための投票情報を生成し、投票情報を選出されたキャンペーンノードへ送信するように構成される。

【０１２１】
本出願の実施の形態は、時間ベースのノード選出装置をさらに提供する。図１３に示すように、本装置は、分散システム内でリーダーノードを決定するように構成されており、リーダーノード内に配置される。
【０１２２】
図１３では、時間ベースのノード選出装置は、再任要求モジュール１３０１と、受信モジュール１３０２と、集計モジュール１３０３と、決定モジュール１３０４とを備え：
全ノードのうちにリーダーノードが存在し、リーダーノードが予め設定された基準を自己が満たしていると判断した場合に、
再任要求モジュールは、時間設定情報に含まれるリーダーノードの任期に従って、投票時間内に、再任要求を全てのノードへ送信するように構成され；
受信モジュールは、他のノードが再任要求に従って生成した投票情報を受信するように構成され；
集計モジュールは、票集計時間内に、受信した投票情報の量を集計するように構成され；
決定モジュールは、集計した投票情報の量に従って再任状態を決定し、再任の成功後に特定のサービスを処理するように構成される。
【０１２３】
予め設定された基準は、負荷処理基準とデータ更新度のうち少なくとも一方を備える。
【０１２４】
本出願の実施の形態は、時間ベースのノード選出装置をさらに提供する。図１４に示すように、本装置は、分散システム内でリーダーノードを決定するように構成されており、リーダーノード内に配置される。
【０１２５】
図１４では、時間ベースのノード選出装置は、選別モジュール１４０１と、投票要求モジュール１４０２とを備え：
全ノードのうちにリーダーノードが存在し、リーダーノードが予め設定された基準を自己が満たしていないと判断した場合に、
選別モジュール１４０１は、「時間設定情報に含まれるリーダーノードの任期に従って、後任ノードとして機能するために予め設定された基準を満たすノードを他のノードの中から選別するように構成され；
投票要求モジュール１４０２は、後任ノードが、他のノードが送信した投票情報を受信し、票集計時間内に投票情報の量を集計し、新たなリーダーノードとして決定されるように、投票期間内に後任ノードのための投票要求を他のノードへ送信するように構成される。
【０１２６】
本出願の実施の形態は、時間ベースのノード選出装置をさらに提供する。図１５に示すように、本装置は、分散システム内でリーダーノードを決定するように構成されており、投票ノード内に配置される。
【０１２７】
図１５では、時間ベースのノード選出装置は、受信モジュール１５０１と、投票情報モジュール１５０２とを備え：
全ノードのうちにリーダーノードが存在する場合に、
受信モジュール１５０１は、時間設定情報に含まれる投票時間内に、リーダーノードが送信した再任要求を受信するように構成され；
投票情報モジュール１５０２は、リーダーノードが、票集計時間内に投票情報の量を集計し、再任状態を決定できるように、再任要求に従って投票情報を生成し、時間設定情報に含まれる投票時間内に投票情報をリーダーノードに送信するように構成され、また、リーダーノードの任期内にリーダーノードの投票要求のみを受信するように構成される。
【０１２８】

本出願の実施の形態は、時間ベースのノード選出装置をさらに提供する。図１６に示すように、本装置は、分散システム内でリーダーノードを決定するように構成されており、投票ノード内に配置される。

【０１２９】

図１６では、時間ベースのノード選出装置は、受信モジュール１６０１と、投票情報モジュール１６０２とを備え：

全ノードのうちにリーダーノードが存在する場合に、

受信モジュール１６０１は、時間設定情報に含まれる投票時間内に、リーダーノードが送信した後任ノードのための投票要求を受信するように構成され；

投票情報モジュール１６０２は、後任ノードが票集計時間内に投票情報の量を集計して、新たなリーダーノードを決定できるように、投票要求に従って投票情報を生成し、時間設定情報に含まれる投票時間内に投票情報を後任ノードへ送信するように構成される。

【０１３０】

典型的な構成では、計算デバイスは１つ以上のプロセッサ（ＣＰＵ）、入／出力インターフェース、ネットワークインターフェース、及びメモリを含む。

【０１３１】

メモリは、揮発性メモリ、ランダムアクセスメモリ（ＲＡＭ）、及び／又は、例えば読出し専用メモリ（ＲＯＭ）又はフラッシュＲＡＭのようなコンピュータで読取り可能な媒体内の不揮発性メモリなどを含んでよい。メモリはコンピュータで読取り可能な媒体の一例である。

【０１３２】

コンピュータで読取り可能な媒体は、不揮発性又は揮発性媒体、可動又は非可動媒体を含み、また、任意の方法あるいは技術によって情報記憶を実行できる。情報はコンピュータで読取り可能な命令、データ構造、及びプログラム又はその他のデータモジュールであってよい。コンピュータの記憶媒体は、例えば、相変化メモリ（ＰＲＡＭ）、スタティックランダムアクセスメモリ（ＳＲＡＭ）、ダイナミックランダムアクセスメモリ（ＤＲＡＭ）、その他のタイプのランダムアクセスメモリ（ＲＡＭ）、読出し専用メモリ（ＲＯＭ）、電気的消去再書込み可能な読出し専用メモリ（ＥＥＰＲＯＭ）、フラッシュメモリ若しくはその他のメモリ技術、コンパクトディスク読取り専用メモリ（ＣＤ－ＲＯＭ）、デジタル多目的ディスク（ＤＶＤ）若しくはその他の光学記憶装置、カセットテープ、磁気テープ／磁気ディスク記憶装置若しくはその他の磁気記憶デバイス、又は任意のその他の非伝送媒体を含むが、これに限定されない、また、アクセス可能な情報を保存するためにコンピューティングデバイスを使用できる。本明細書での定義によれば、コンピュータで読取り可能な媒体は、変調データ信号及び搬送波のような一時的媒体を含まない。

【０１３３】

用語「含む」、「備える」、又はこれらのあらゆる派生形は、非排他的な包含をカバーすることを意図し、一連の要素を含む工程、方法、商品、デバイスは、このような要素を含むだけでなく、明記されていないその他の要素をも含むか、あるいは、その工程、方法、商品、デバイスに固有な要素をさらに含む点にも留意されたい。さらなる制限をせずに、表現「～を含む（include a/an…）」によって限定される要素は、その要素を含む工程、方法、商品、デバイスがさらに有するその他の同じ要素を除外するものではない。

【０１３４】

当業者は、本願の実施の形態を、方法、システム、コンピュータプログラム製品として提供できることを理解すべきである。したがって、本願の実施の形態は、完全なハードウェアの実施の形態、完全なソフトウェアの実施の形態、又はソフトウェアとハードウェアの組み合わせの実施の形態の形態で実施できる。さらに、本願は、１つ以上のコンピュータで使用可能な記憶媒体（磁気ディスクメモリ、ＣＤ－ＲＯＭ、光学メモリなどを含むが、これに限定されない）上で実施できるコンピュータプログラム製品（コンピュータで使用可能なプログラムコードを含む）の形態を採ることができる。

【０１３５】

上記は単なる本願の実施の形態であり、それを用いて本願を制限するものではない。当業者には、本願への様々な変更及び改造が可能であろう。本願の精神及び原理から逸脱せずになされるあらゆる改変、均等物との置き換え、改良等は、本願の請求の範囲に包含されるものである。

［第１の局面］

分散システム内でリーダーノードを決定するために用いられる、時間ベースのノード選出方法であって：

同一の投票時間と、同一の票集計時間と、同一のリーダーノードの任期とを、ホストにより、全ノードに設定するステップと；

前記設定された投票時間と、票集計時間と、リーダーノードの任期とを含む時間設定情報を、全ての前記ノードへ送信し、それにより、前記ノードが前記投票時間内に投票し、前記票集計時間内に前記票の数を集計し、票集計結果に従ってリーダーノードを選出し、そして、前記ノードは前記リーダーノードの任期に従って周期的なノード選出を実行できるようになされる、ステップと；を備える、

時間ベースのノード選出方法。

［第２の局面］

分散システム内でリーダーノードを決定するために用いられる、時間ベースのノード選出方法であって：

全ノードのうちにリーダーノードが存在しない場合、時間設定情報に含まれる投票時間内に、投票要求を、キャンペーンノードとしての少なくとも１つのノードにより、前記全ノードへ送信するステップと；

他のノードが前記投票要求に従って生成した投票情報を、各キャンペーンノードにより、受信するステップと；

前記時間設定情報に含まれる票集計時間内に、受信した前記投票情報の量を、前記キャンペーンノードのそれぞれにより、集計するステップと；

特定のサービスを処理するために、前記集計したそれぞれの投票情報の量に従ってリーダーノードを、前記キャンペーンノードにより、決定するステップと；を備える、

時間ベースのノード選出方法。

［第３の局面］

前記票集計時間内に前記キャンペーンノードが受信した前記投票情報の量を集計する前記ステップの前に、

前記票集計時間が到来したとき、前記キャンペーンノードによる前記投票情報の受信を停止するステップをさらに備える、

第２の局面に記載の方法。

［第４の局面］

前記キャンペーンノードが、前記集計したそれぞれの投票情報の量に従ってリーダーノードを決定するステップは、具体的に、

前記キャンペーンノードが、前記集計した前記投票情報の量が予め設定された閾値を超えると判断したときに、前記キャンペーンノードにより、前記キャンペーンノード自体をリーダーノードとして決定するステップを含み、

前記閾値は全ての前記ノードの数の１／２である、

第２の局面に記載の方法。

［第５の局面］

分散システム内でリーダーノードを決定するために用いられる、時間ベースのノード選出方法であって：

全ノードのうちにリーダーノードが存在しない場合、時間設定情報に含まれる投票時間内に、キャンペーンノードから送信された投票要求を、投票ノードとしての少なくとも２つのノードにより、受信するステップと；

リーダーノードを決定するために、前記キャンペーンノードが票集計時間内に前記投票情報の量を集計できるよう、前記投票要求に従って投票情報を、前記投票ノードにより、

生成し、前記投票情報を前記投票時間内に前記キャンペーンノードへ送信するステップと；を備える、
時間ベースのノード選出方法。
［第６の局面］
前記投票情報を生成し、前記投票情報を前記キャンペーンノードへ送信するステップは、具体的に、
前記受信した投票要求に含まれるノード情報を、前記投票ノードにより、決定するステップと；
前記キャンペーンノードのノード状態重量を前記ノード情報に従って決定するステップと；
最大のノード状態重量を持つキャンペーンノードを選出し、前記キャンペーンノードのための投票情報を生成し、前記投票情報を前記選出したキャンペーンノードへ送信するステップと；を備える、
第５の局面に記載の方法。
［第７の局面］
分散システム内でリーダーノードを決定するために用いられる、時間ベースのノード選出方法であって：
全ノードのうちにリーダーノードが存在し、前記リーダーノードが予め設定された基準を自己が満たしている、と判断した場合に、投票時間内に、時間設定情報に含まれるリーダーノードの任期に従って、再任要求を、前記リーダーノードにより、前記全ノードへ送信するステップと；
前記再任要求に従って、他のノードが生成した投票情報を、前記リーダーノードにより、受信するステップと；
票集計時間内に、前記受信した投票情報の量を、前記リーダーノードにより、集計するステップと；、
前記集計した前記投票情報の量に従って再任状態を決定し、再任成功後に特定のサービスを、前記リーダーノードにより、処理するステップと；を備え、
前記予め設定された基準は、負荷処理基準とデータ更新度とのうち少なくとも一方を備える、
時間ベースのノード選出方法。
［第８の局面］
分散システム内でリーダーノードを決定するために用いられる、時間ベースのノード選出方法であって：
全ノードのうちにリーダーノードが存在し、前記リーダーノードが、予め設定された基準を自己が満たしていないと判断した場合に、時間設定情報含まれるリーダーノードの任期に従って、後任ノードとして機能するための前記予め設定された基準を満たすノードを、前記リーダーノードにより他のノードの中から選別するステップと；
前記後任ノードは、他のノードが送信した投票情報を受信し、票集計時間内に前記投票情報の量を集計し、新たなリーダーノードとして決定されるようにするために、前記投票期間内に、後任ノードについての投票要求を、前記リーダーノードにより、他のノードへ送信するステップと；を備える、
時間ベースのノード選出方法。
［第９の局面］
時間ベースのノード選出方法であって：
全ノードのうちにリーダーノードが存在する場合、時間設定情報に含まれる投票時間内に、前記リーダーノードが送信した再任要求を、投票ノードとしての少なくとも２つのノードにより、受信するステップと；
前記リーダーノードが票集計時間内に前記投票情報の量を集計して、再任状態を決定できるようにするために、前記再任要求に従って投票情報を生成し、前記時間設定情報に含まれる前記投票時間内に、前記投票情報を、前記投票ノードにより、前記リーダーノード

へ送信するステップと；を備え、
前記投票ノードは、前記リーダーノードの任期中に前記リーダーノードの前記投票要求のみを受信する、
時間ベースのノード選出方法。
［第１０の局面］
分散システム内でリーダーノードを決定するために用いられる、時間ベースのノード選出方法であって：
全ノードのうちにリーダーノードが存在する場合、時間設定情報に含まれる投票時間内に、前記リーダーノードから送信された後任ノードのための投票要求を、投票ノードとしての少なくとも２つのノードにより、受信するステップと；
前記後任ノードが票集計時間内に前記投票情報の量を集計して、新たなリーダーノードを決定できるようにするために、前記投票要求に従って投票情報を生成し、前記時間設定情報に含まれる前記投票時間内に、前記投票情報を、前記投票ノードにより前記後任ノードへ送信するステップと；を備える、
時間ベースのノード選出方法。
［第１１の局面］
分散システム内でリーダーノードを決定するように構成された、時間ベースのノード選出装置であって：
全ノードに同一の投票時間と、同一の票集計時間と、同一のリーダーノードの任期とを設定するように構成された設定モジュールと；
前記ノードが、前記投票時間内に投票し、前記票集計時間内に票数を集計し、票集計結果に従ってリーダーノードを選出できるように、そして、前記ノードが前記リーダーノードの任期に従って周期的なノード選出を実行できるようにするために、前記投票時間と、票集計時間と、リーダーノードの任期とが含まれる時間設定情報を前記全ノードへ送信するように構成された送信モジュールと；を備える、
時間ベースのノード選出装置。
［第１２の局面］
分散システム内でリーダーノードを決定するように構成され、キャンペーンノード内に配置される時間ベースのノード選出装置であって、全ノードのうちにリーダーノードが存在しない場合に、
時間設定情報に含まれる投票時間内に、投票要求を前記全ノードへ送信するように構成された投票要求モジュールと；
他のノードが前記投票要求に従って生成した投票情報を受信するように構成された受信モジュールと；
前記時間設定情報に含まれる票集計時間内に、受信した前記投票情報の量を集計するように構成された集計モジュールと；
特定のサービスを処理するために、前記集計したそれぞれの投票情報の量に従って、リーダーノードを決定するように構成された決定モジュールと；を備える、
時間ベースのノード選出装置。
［第１３の局面］
更に、前記受信モジュールは、前記票集計時間が到来すると前記投票情報の受信を停止するように構成された、
第１２の局面に記載の装置。
［第１４の局面］
前記決定モジュールは、具体的に、前記集計した前記投票情報の量が予め設定された閾値を超えたと判断した場合に、前記キャンペーンノードをリーダーノードとして決定するように構成されており、
前記閾値は前記全てのノードの数の１／２である、
第１２の局面に記載の装置。
［第１５の局面］

分散システム内でリーダーノードを決定するように構成され、投票ノード内に配置される、時間ベースのノード選出装置であって、全ノードのうちにリーダーノードが存在しない場合に、
時間設定情報に含まれる投票時間内に、キャンペーンノードが送信した投票要求を受信するように構成された受信モジュールと；
前記キャンペーンノードが票集計時間内に前記投票情報の量を集計し、リーダーノードを決定できるようにするために、前記投票要求に従って投票情報を生成し、前記投票時間内に前記投票情報を前記キャンペーンノードへ送信するように構成された投票情報モジュールと；を備える、
時間ベースのノード選出装置。
［第１６の局面］
前記投票情報モジュールは、具体的に、前記受信した投票要求に含まれるノード情報を判定し、前記ノード情報に従って前記キャンペーンノードのノード状態重量を決定し、最大のノード状態重量を持つキャンペーンノードを選出し、前記キャンペーンノードのための投票情報を生成し、前記投票情報を前記選出されたキャンペーンノードへ送信するように構成された、
第１５の局面に記載の装置。
［第１７の局面］
分散システム内でリーダーノードを決定するように構成され、リーダーノード内に配置される、時間ベースのノード選出装置であって、全ノードのうちにリーダーノードが存在し、前記リーダーノードが予め設定された基準を自己が満たしていると判断した場合に、
時間設定情報に含まれるリーダーノードの任期に従って、投票時間内に、再任要求を全ての前記ノードへ送信するように構成された再任要求モジュールと；
他のノードが前記再任要求に従って生成した投票情報を受信するように構成された受信モジュールと；
票集計時間内に、前記受信した投票情報の量を集計するように構成された集計モジュールと；
前記集計した前記投票情報の量に従って再任状態を決定し、再任の成功後に特定のサービスを処理するように構成された決定モジュールと；を備え、
前記予め設定された基準は、負荷処理基準とデータ更新度のうち少なくとも一方を備える、
時間ベースのノード選出装置。
［第１８の局面］
分散システム内でリーダーノードを決定するように構成され、リーダーノード内に配置される、時間ベースのノード選出装置であって、全ノードのうちにリーダーノードが存在し、前記リーダーノードが予め設定された基準を自己が満たしていないと判断した場合に、
時間設定情報に含まれるリーダーノードの任期に従って、後任ノードとして機能するために予め設定された基準を満たすノードを他のノードの中から選別するように構成された選別モジュールと；
前記後任ノードが、他のノードが送信した投票情報を受信し、票集計時間内に前記投票情報の量を集計し、新たなリーダーノードとして決定されるようにするために、前記投票期間内に前記後任ノードのための投票要求を他のノードへ送信するように構成された投票要求モジュールと；を備える、
時間ベースのノード選出装置。
［第１９の局面］
分散システム内でリーダーノードを決定するように構成され、投票ノード内に配置される、時間ベースのノード選出装置であって、全ノードのうちにリーダーノードが存在する場合に、
時間設定情報に含まれる投票時間内に、前記リーダーノードが送信した再任要求を受信

するように構成された受信モジュールと；

投票情報モジュールであって、前記リーダーノードが、票集計時間内に前記投票情報の量を集計し、再任状態を決定できるようにするために、前記再任要求に従って投票情報を生成し、前記時間設定情報に含まれる前記投票時間内に前記投票情報を前記リーダーノードに送信するように構成されており、また、前記リーダーノードの任期内に前記リーダーノードの前記投票要求のみを受信するように構成された、投票情報モジュールと；を備える、

時間ベースのノード選出装置。

［第２０の局面］

分散システム内でリーダーノードを決定するように構成され、投票ノード内に配置される、時間ベースのノード選出装置であって、全ノードのうちにリーダーノードが存在する場合に、

時間設定情報に含まれる投票時間内に、前記リーダーノードが送信した後任ノードのための投票要求を受信するように構成された受信モジュールと；

前記後任ノードが票集計時間内に前記投票情報の量を集計できるようにするために、前記投票要求に従って投票情報を生成し、前記時間設定情報に含まれる前記投票時間内に前記投票情報を前記後任ノードへ送信するように構成された投票情報モジュールと；を備える、

時間ベースのノード選出装置。

【図１】

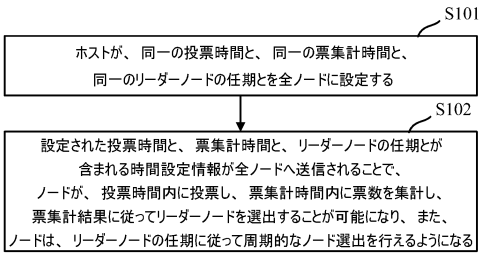

図1

【図２】

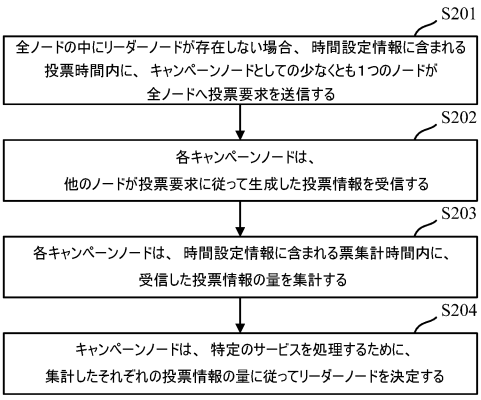

図2

【図３】



図3a

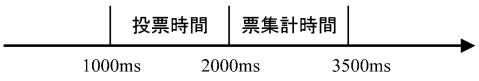

図3b

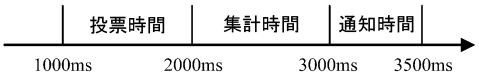

図3c

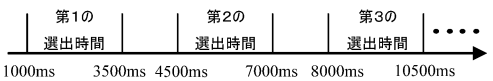

図3d

【図４】

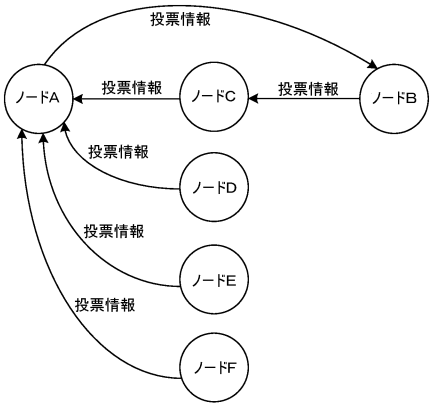

図4

【図５】

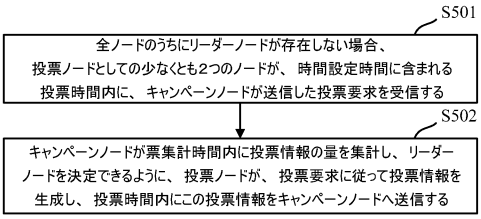

図5

【図６】

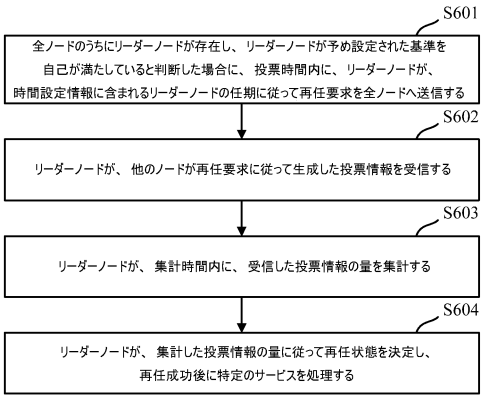

図6

【図７】

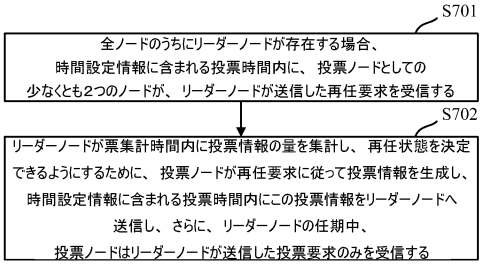

図7

【図８】

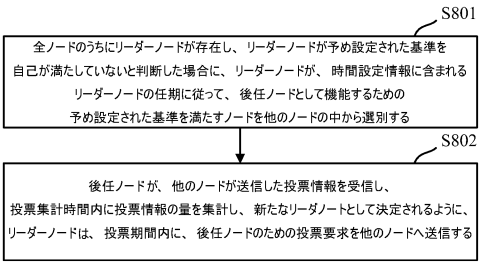

図8

【図９】

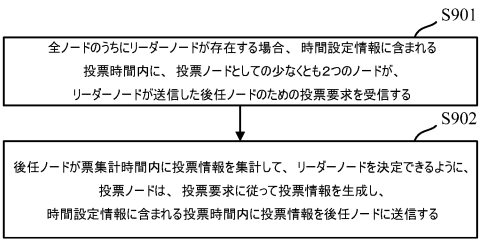

図9

【図１０】

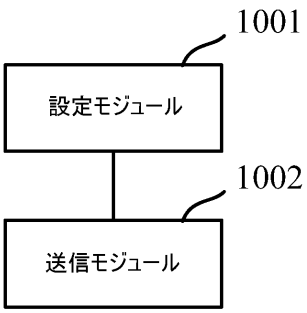

図10

【図１１】

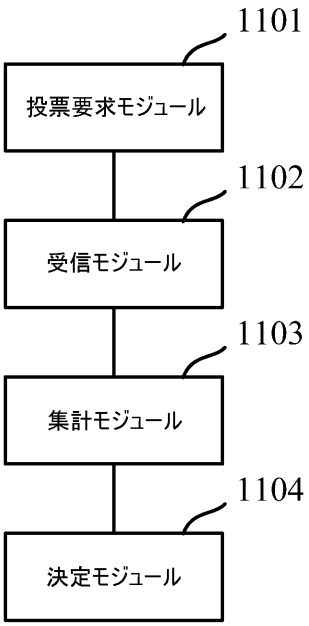

図11

【図１２】

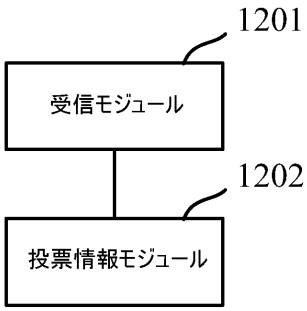

図12

【図１３】

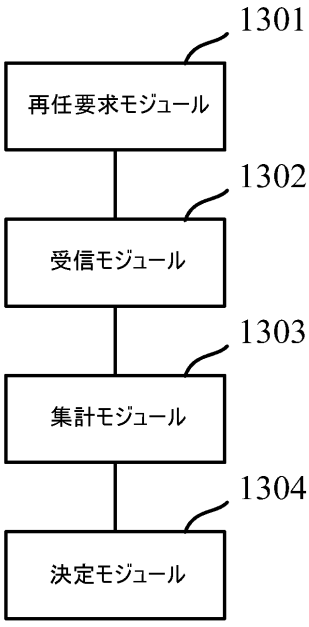

図13

【図１４】

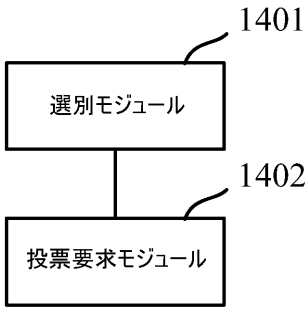

図14

【図１５】

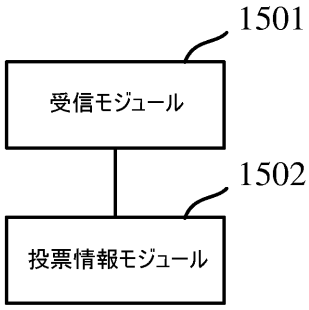

図15

【図１６】

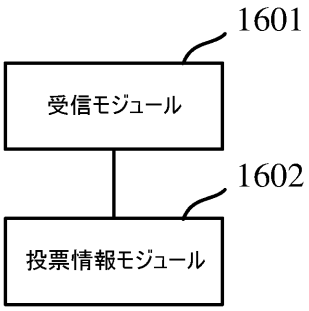

図16

フロントページの続き

(74)代理人 100155192
弁理士 金子 美代子
(74)代理人 100131820
弁理士 金井 俊幸
(74)代理人 100100398
弁理士 柴田 茂夫
(72)発明者 ヤン, ヂェンクン
中華人民共和国 ３１００９９, ハンヂョウ, ナンバー１８ ワンタン ロード, フアンロン タイムズ プラザ, ビルディング ビー １７エフ, アンツ パテント チーム内
(72)発明者 シャオ, ジンリァン
中華人民共和国 ３１００９９, ハンヂョウ, ナンバー１８ ワンタン ロード, フアンロン タイムズ プラザ, ビルディング ビー １７エフ, アンツ パテント チーム内

審査官 木村 雅也

(56)参考文献 米国特許出願公開第２００８／００７１８５３（ＵＳ, Ａ１）
米国特許出願公開第２００８／００７１８７８（ＵＳ, Ａ１）
特開２０１５－０３６８３４（ＪＰ, Ａ）
特開２００６－００４４３４（ＪＰ, Ａ）

(58)調査した分野(Int.Cl., ＤＢ名)
Ｇ０６Ｆ １３／００
Ｇ０６Ｆ １７／４０