

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5296803号  
(P5296803)

(45) 発行日 平成25年9月25日(2013.9.25)

(24) 登録日 平成25年6月21日(2013.6.21)

(51) Int.Cl.

F I

H04L 12/70 (2013.01)

H04L 12/70 I O O A

請求項の数 21 (全 19 頁)

(21) 出願番号 特願2010-537549 (P2010-537549)  
 (86) (22) 出願日 平成20年11月4日(2008.11.4)  
 (65) 公表番号 特表2011-507378 (P2011-507378A)  
 (43) 公表日 平成23年3月3日(2011.3.3)  
 (86) 国際出願番号 PCT/IB2008/054587  
 (87) 国際公開番号 W02009/074902  
 (87) 国際公開日 平成21年6月18日(2009.6.18)  
 審査請求日 平成23年9月15日(2011.9.15)  
 (31) 優先権主張番号 11/955,933  
 (32) 優先日 平成19年12月13日(2007.12.13)  
 (33) 優先権主張国 米国(US)

(73) 特許権者 598036300  
 テレフオンアクチーボラゲット エル エ  
 ム エリクソン (パブル)  
 スウェーデン国 ストックホルム エスー  
 1 6 4 8 3  
 (74) 代理人 100095957  
 弁理士 亀谷 美明  
 (74) 代理人 100096389  
 弁理士 金本 哲男  
 (74) 代理人 100101557  
 弁理士 萩原 康司  
 (74) 代理人 100128587  
 弁理士 松本 一騎

最終頁に続く

(54) 【発明の名称】 処理メトリック情報のノード間における交換

(57) 【特許請求の範囲】

【請求項 1】

ノードの処理キャパシティに関する情報を取得する方法であって、  
 前記ノードから第1のレートで第1の信号を受信するステップと、  
 前記ノードから第2のレートで第2の信号を受信するステップと、  
 前記第2のレートと前記第1のレートとの間の比率を算出するステップと、  
 前記比率のプリセット値との比較に基づいて前記ノードの前記処理キャパシティを決定  
 するステップと、

を含み、

前記第1の信号の連続するインスタンスの各々は、前記第1の信号の前のインスタ  
 スのシーケンス番号に基づいて算出されるシーケンス番号を含み、

前記第1の信号の最新のインスタンスの直後に受信される前記第2の信号のインスタン  
 スは、前記第1の信号の前記最新のインスタンスのシーケンス番号に基づくシーケンス番  
 号を含み、

前記第1の信号の次のインスタンスの前に受信される前記第2の信号の連続するイン  
 スタンスの各々は、前記第2の信号の前のインスタンスのシーケンス番号に基づくシー  
 ケンス番号を含む、

方法。

【請求項 2】

前記第1のレートは一定であり、及び

10

20

前記第 2 のレートは可変である、  
請求項 1 に記載の方法。

【請求項 3】

前記プリセット値は、前記ノードが通常の処理キャパシティを有する場合の前記第 2 及び第 1 のレートの間の比率に等しい、  
請求項 1 に記載の方法。

【請求項 4】

前記ノードの前記処理キャパシティを決定する前記ステップに基づいてルーティングテーブルを更新するステップと、  
前記ノードの前記処理キャパシティが低い場合に、前記ノードに向けてメッセージを送信することを控えるステップと、  
をさらに含む、請求項 1 に記載の方法。

10

【請求項 5】

前記比率における変化を監視するステップと、  
前記比率が変化する場合に、前記ノードの処理キャパシティの変化を検出するステップと、  
をさらに含む、請求項 1 に記載の方法。

【請求項 6】

前記第 2 のレートの減少は、前記処理キャパシティの低下を示す、  
請求項 5 に記載の方法。

20

【請求項 7】

前記第 1 の信号の連続するインスタンスの各々における前記シーケンス番号は、前記第 1 の信号の前記前回のインスタンスの前記シーケンス番号に前記プリセット値を加えたものに等しく、

前記第 1 の信号の前記最新のインスタンスの直後に受信される前記第 2 の信号の前記インスタンスにおける前記シーケンス番号は、前記第 1 の信号の前記最新のインスタンスの前記シーケンス番号に 1 を加えたものに等しく、

前記第 1 の信号の前記次のインスタンスの前に受信される前記第 2 の信号の連続するインスタンスの各々における前記シーケンス番号は、前記第 2 の信号の前記前回のインスタンスの前記シーケンス番号に 1 を加えたものに等しい、

30

請求項 1 に記載の方法。

【請求項 8】

前記第 1 の信号の前記最新のインスタンスの前記シーケンス番号に基づいて、又は前記第 2 の信号の前記前回のインスタンスの前記シーケンス番号に基づいて、前記第 2 の信号の次のインスタンスの期待されるシーケンス番号を算出するステップと、

前記期待されるシーケンス番号に等しくないシーケンス番号を含む前記第 2 の信号の前記次のインスタンスを受信することにより、キャパシティの問題を検出するステップと、  
をさらに含む、請求項 1 に記載の方法。

【請求項 9】

前記第 1 の信号の前記最新のインスタンスの前記シーケンス番号に基づいて、又は前記第 2 の信号の前記前回のインスタンスの前記シーケンス番号に基づいて、前記第 2 の信号の次のインスタンスの期待されるシーケンス番号を算出するステップ、をさらに含む、

40

前記第 2 のレートと前記第 1 のレートとの間の前記比率を算出することは、前記期待されるシーケンス番号に合致しないシーケンス番号を含む前記第 2 の信号の所与のインスタンスを無視することを含む、

請求項 1 に記載の方法。

【請求項 10】

ノードの処理キャパシティに関する情報を発信する方法であって、

前記ノードの通常の処理キャパシティを示すプリセット値以下である提供キャパシティ

50

値を決定するステップと、

前記ノードから基本レートで周期的に第 1 の信号を送信するステップと、

前記ノードから前記提供キャパシティ値を示す第 2 のレートで周期的に第 2 の信号を送信するステップと、

を含み、

前記第 2 のレートは、前記基本レートに前記提供キャパシティ値を乗じたものに等しく

、  
前記第 1 の信号及び前記第 2 の信号は、前記基本レートと前記第 2 のレートとの間の比率に基づいて前記ノードの処理キャパシティを決定するピアノードへ送信される、

方法。

10

【請求項 1 1】

前記ノードは複数のインタフェースを含み、

周期的な第 1 の信号の各々は全てのインタフェース上で送信され、及び

周期的な第 2 の信号の各々は全てのインタフェース上で送信される、

請求項 1 0 に記載の方法。

【請求項 1 2】

前記提供キャパシティ値は、周期的に決定される、

請求項 1 0 に記載の方法。

【請求項 1 3】

前記提供キャパシティ値は、前記ノードの実際の処理キャパシティに従って設定される

20

、  
請求項 1 0 に記載の方法。

【請求項 1 4】

前記提供キャパシティ値は、前記ノードの前記通常の処理キャパシティの所定の割合に従って設定される、

請求項 1 0 に記載の方法。

【請求項 1 5】

自身の処理キャパシティに関する情報を発信するためのノードであって、

プロセッサと、

前記プロセッサの負荷レベルを示すよう構成される負荷インジケータと、

30

ピアノードに向けて信号を送信するよう構成される 1 つ以上のインタフェースと、

第 1 のレートで第 1 の信号を送信するよう前記 1 つ以上のインタフェースに要求し、

前記負荷レベルと前記ノードの処理キャパシティの最大量との比較に基づいて第 2 のレートを算出し、並びに

前記第 2 のレートで第 2 の信号を送信するよう前記 1 つ以上のインタフェースに要求する、

よう構成される、制御ロジックと、

を含み、

前記負荷レベルと処理キャパシティの前記最大量との比較に基づいて第 2 のレートを算出することは、前記負荷レベルが低い場合に前記第 2 のレートを最大値に設定することを含み、

40

前記第 1 の信号及び前記第 2 の信号は、前記第 1 のレートと前記第 2 のレートとの間の比率に基づいて前記ノードの処理キャパシティを決定する前記ピアノードへ送信される、

ノード。

【請求項 1 6】

前記第 1 のレートで前記制御ロジックのシーケンスを開始するよう構成されるクロックをさらに含む、請求項 1 5 に記載のノード。

【請求項 1 7】

自身の処理キャパシティに関する情報を発信するためのノードであって、

プロセッサと、

50

前記プロセッサの負荷レベルを示すよう構成される負荷インジケータと、  
 ピアノードに向けて信号を送信するよう構成される１つ以上のインタフェースと、  
 第１のレートで第１の信号を送信するよう前記１つ以上のインタフェースに要求し、  
 前記負荷レベルと前記ノードの処理キャパシティの最大量との比較に基づいて第２のレートを算出し、並びに

前記第２のレートで第２の信号を送信するよう前記１つ以上のインタフェースに要求する、

よう構成される、制御ロジックと、  
 を含み、

前記負荷レベルと処理キャパシティの前記最大量との比較に基づいて第２のレートを算出することは、前記負荷レベルが低い場合に前記第２のレートを最大値に設定することを含み、

10

前記第１の信号の連続するインスタンスの各々は、前記第１の信号の前のインスタンスのシーケンス番号に基づいて算出されるシーケンス番号を含み、

前記第１の信号の最新のインスタンスの直後に送信される前記第２の信号のインスタンスは、前記第１の信号の前記最新のインスタンスのシーケンス番号に基づくシーケンス番号を含み、及び

前記第１の信号の次のインスタンスの前に送信される前記第２の信号の連続するインスタンスの各々は、前記第２の信号の前のインスタンスのシーケンス番号に基づくシーケンス番号を含む、

20

ノード。

#### 【請求項１８】

ピアノードの処理キャパシティに関する情報を取得するためのノードであって、  
 前記ピアノードから信号を受信するよう構成される１つ以上のインタフェースと、  
 前記１つ以上のインタフェースから第１の信号を複数回受信し、  
 前記第１の信号の第１のレートを決定し、  
 前記１つ以上のインタフェースから第２の信号を複数回受信し、  
 前記第２の信号の第２のレートを決定し、  
 前記第２及び第１のレートの間の比率を算出し、並びに  
 前記比率に基づく前記ピアノードの処理キャパシティをピアテーブルに記憶する、  
 よう構成される、制御ロジックと、  
 クロックと、  
 を含み、

30

前記制御ロジックは、前記クロックからのタイミング情報を使用して前記第１及び第２のレートを決定し、並びに前記第１及び第２のレートを前記ピアテーブルに記憶する、  
 ノード。

#### 【請求項１９】

前記第１及び第２の信号は、シーケンス番号を含み、  
 前記制御ロジックは、最新の第１の信号のシーケンス番号に基づいて、又は前回の第２の信号のシーケンス番号に基づいて、次の第２の信号の期待されるシーケンス番号を計算するようさらに構成され、並びに

40

前記制御ロジックは、前記期待されるシーケンス番号を前記ピアテーブルに記憶するようさらに構成される、

請求項１８に記載のノード。

#### 【請求項２０】

前記制御ロジックは、前記期待されるシーケンス番号に等しくないシーケンス番号を含む次の第２の信号を受信することにより、キャパシティの問題を検出するようさらに構成される、

請求項１９に記載のノード。

#### 【請求項２１】

50

前記第 2 及び第 1 のレートの間比率を算出することは、前記期待されるシーケンス番号に合致しないシーケンス番号を含む所与の第 2 の信号を無視することを含む、請求項 19 に記載のノード。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、計算キャパシティに関するメトリック (metric) 情報をノード間で交換するための方法及びノードに関する。

【背景技術】

10

【0002】

計算ノード (computing nodes) のネットワークにおいて、ノード間の通信は、予め定義されたルールに従って行われてもよく、又はネットワークの現在の状態に基づくメトリックに従って行われてもよい。ネットワーク、特にアドホックなネットワークにおけるノードの調整は、問題を孕み得る。例えば、複数のノードが全て同じ処理パワーを有しているわけではない場合、ノード間のリンクの既知の又は期待されるキャパシティのみに従ってルーティングすることは、効率的でない可能性がある。ネットワークには、可変レベルの柔軟性をもって設計されるものもあり、そのノードのうちのいくつかはメッセージ又は要求を送信する際にピアノード間における宛先を選択することができる。これらネットワークを効率的な手法で運用できるようにするために、ノードがそのような選択をすることを可能とする改良されたメトリックを提供することが必要となる。

20

【0003】

2つのコプロセッサ上に負荷をバランシングする単純な手法は、単純に一方のプロセッサ上に一定のタスクを、もう一方のプロセッサ上に他のタスクを割り当てることである。例えば、パーソナルコンピュータ上の1つの中央処理装置 (CPU) は、もう一方のCPUが入力データの計算を処理する傍ら、画像表示を処理し得る。このような所定の構成 (pre-arranged) は、例えば出力を画像で表示できるようにする前に大量の入力データを処理する必要がある場合、非常に非効率的である。

【0004】

処理負荷をバランシングする他の方法は、様々なプロセッサをスキャンすることにより、各プロセッサにおいて処理を待つタスクキューの長さを取得することである。この方法は、様々なプロセッサが同一の場所に配置されている場合のみ効果があり、例えばコプロセッサの場合、少なくともある部分スキャン処理の実行に供される1つの個別のプロセッサに依存し得る。

30

【0005】

並行スケジューリングは、負荷をバランシングするための別のアプローチである。並行スケジューリングの方法において、全てのプロセッサは協同して仕事をスケジューリングする。並行スケジューリングは、コンパイル時間又は実行時間において全体的な負荷情報を使用することにより負荷をバランシングする。これらの方法は一般に、プロセッサの相対的な処理キャパシティを考慮した、様々なプロセッサへの予め決定されたタスクの割り当てに依存する。並行スケジューリングの方法は、ノードにかかる負荷の観点についての時間的に変化する要件のような、変化する条件にほとんど適応できない。

40

【0006】

2つのノード間における最適な計算パスを見つけるためにルーティングプロトコルと共に通常使用されるメトリックは、最小ホップカウントと呼ばれる。2つのノード間における様々なパスが比較され、離散リンク、リピータ、ルータ、及び/又はゲートウェイの数がカウントされ、パス毎のホップの数を生じさせる。ホップの数が最小となる所与のパスが、2つのノードを接続するための最適なパスとして選択される。第1のノードが2つ以上のピアノードとの間で通信する選択肢を有する場合、ピアノードのうち最小数のホップを使用して到達され得る1つを選択してもよい。故に、最小ホップカウントを使用して2

50

つのノード間における複数のパスのうちの1つを選択してもよく、又は複数のピアノードのうちの1つを選択してもよい。

【0007】

ノードが異なる能力を有する環境において、ホップカウントのルーティングは適切ではないことが示されている。例えば“A High-Throughput Path Metric for Multi-Hop Wireless Routing”(D.S.J. De Couto et al., M.I.T. Computer Science and Artificial Intelligence Laboratory, International Conference on Mobile Computing and Networking, San Diego CA, September 2003)において、最小ホップカウントのルーティング技術は、ある環境において貧弱な性能を有することを著者は実証している。彼らは、2つのノード間に配置されるリンクの前方及び後方への配信比率(forward and reverse delivery ratio)を使用して算出される期待される送信カウントに基づいて、別のメトリックを提案している。その一方、“Comparison of Routing Metrics for Static Multi-Hop Wireless Networks”(R.Draves et al., Microsoft Research, Proceedings of the 2004 conference on Applications, technologies, architectures, and protocols for computer communications, Portland OR, August-September 2004)の他の著者は、メッセージの送信者が移動体である場合、ホップカウントのメトリックは他のリンク品質(link-quality)のメトリックよりも優れた性能を有することを主張している。

10

【0008】

それにも関わらず、ホップカウント又はより高度なメトリックにより最善のルーティングパスを決定する先行技術の方法は、ノードにはトラフィック処理能力が非常に高いものもあれば、非常に低いものもあり得るという事実を考慮に入れていない。ノードがネットワークのキャパシティを最適に使用できるように、その個々の処理キャパシティに関するメトリック情報をノード間において交換することに関する問題は、キャパシティの限界又はノード自身の現在の負荷を考慮に入れていないノード間の通信手段のみに基づくノードの調整によっては解決されない。

20

【発明の概要】

【発明が解決しようとする課題】

【0009】

通信するノードが、そのピアノードの現在の処理能力に関する正確で最新の情報を取得できるようにするための手段を有することは明らかに利点があるであろう。そのため、本発明の広範な目的は、1つのノードの処理キャパシティに関するメトリック情報をノード間において交換するための方法及びノードを提供することである。

30

【課題を解決するための手段】

【0010】

本発明の第1の観点は、ノードの処理キャパシティに関する情報を取得する方法を対象とする。上記方法は、上記ノードから第1のレートで受信される第1の信号を含む。上記ノードから第2のレートで第2の信号も受信される。上記第2のレートと上記第1のレートとの間の比率が算出される。上記比率のプリセット値との比較に基づいて上記ノードの上記処理キャパシティが決定される。

40

【0011】

本発明の第2の観点は、ノードの処理キャパシティに関する情報を取得する上記方法の変形を対象とする。連続する第1の信号の各々は、前回の第1の信号のシーケンス番号に基づいて算出されるシーケンス番号を含む。最新の第1の信号の直後に第2の信号が受信される場合、それは上記最新の第1の信号のシーケンス番号に基づくシーケンス番号を含む。その後、次の第1の信号の前に受信される連続する第2の信号の各々は、前回の第2の信号のシーケンス番号に基づくシーケンス番号を含む。上記最新の第1の信号の、又は上記前回の第2の信号の上記シーケンス番号に基づいて、次の第2の信号の期待されるシーケンス番号が算出される。上記期待されるシーケンス番号に等しくないシーケンス番号を含む上記次の第2の信号を受信することにより、生じ得るキャパシティの問題が

50

検出される。

【 0 0 1 2 】

本発明の第3の観点は、ノードの処理キャパシティに関する情報を発信する方法を対象とする。上記方法は、提供キャパシティ値を決定するステップをもって開始される。上記提供キャパシティ値は、上記ノードの通常の処理キャパシティを示すプリセット値以下である。上記ノードは、基本レートで周期的に第1の信号を送信する。上記ノードは、上記提供キャパシティ値を示す第2のレートで周期的に第2の信号も送信する

【 0 0 1 3 】

本発明の第4の観点は、自身の処理キャパシティに関する情報を発信するためのノードを対象とする。上記ノードは、プロセッサと、上記プロセッサの負荷レベルを示す負荷インジケータとを含む。1つ以上のインタフェースは、ピアノードに向けて信号を送信するよう構成される。制御ロジックは、第1のレートで第1の信号を送信するよう上記1つ以上のインタフェースに要求する。上記制御ロジックは、上記プロセッサの上記負荷レベルを読み取り、及び上記負荷レベルの上記ノードの処理キャパシティの最大量との比較に基づいて第2のレートを算出する。上記制御ロジックは、上記第2のレートで第2の信号を送信するよう上記1つ以上のインタフェースに要求する。

10

【 0 0 1 4 】

本発明の第5の観点は、ピアノードの処理キャパシティに関する情報を取得するためのノードを対象とする。上記ノードは、上記ピアノードから信号を受信する1つ以上のインタフェースを含む。上記ノードは、制御ロジックも含む。上記制御ロジックは、上記1つ以上のインタフェースから第1の信号を複数回受信する。それは、上記第1の信号の第1のレートを決定する。上記制御ロジックはまた、上記1つ以上のインタフェースから第2の信号を複数回受信し、及び上記第2の信号の第2のレートを決定する。上記制御ロジックは、それから上記第2及び第1のレートの間の比率を算出し、並びに上記比率に基づく上記ピアノードの処理キャパシティをピアテーブルに記憶する

20

【図面の簡単な説明】

【 0 0 1 5 】

ここで本発明のより詳細な理解のため、さらなる目的及びその利点について、次の添付の図面と併せて、以下で説明する。

【 0 0 1 6 】

30

【図1】ノードの処理キャパシティに関する情報を発信(propagate)する一例となる方法を示す。

【図2】本発明の他の観点に従って、ピアノードにおいて対象ノード(node of interest)の処理キャパシティに関する情報を取得する一例となる方法を示す。

【図3】本発明の方法の追加的観点のステップを示す。

【図4】対象ノードの処理キャパシティを決定する処理においてより精度を上げるためのいくつかの例となる観点を示す。

【図5】本発明の観点に係る一例となるノードを示す。

【図6】本発明のいくつかの観点に係る一例となるピアテーブルを示す。

【発明を実施するための形態】

40

【 0 0 1 7 】

本発明の革新的な教示を、例となる様々な使用及び好適な実施形態の観点を特に参照して説明する。一方、この実施形態が本発明の革新的な教示の多くの有利な使用のうち数例しか提供していないことは、理解されるべきである。概して、本アプリケーションの本明細書における記述は、本発明の主張される様々な観点のいずれをも必ずしも限定しない。さらにいくつかの記述は、発明の特徴のいくつかに適用してもよく、他には適用しなくてもよい。図の説明において、同様の番号は本発明の同様の要素を表す。

【 0 0 1 8 】

本発明は、1つのノードの処理キャパシティに関するメトリック情報をノード間において交換するための方法及びノードを提供する。ノードのピアは、このメトリック情報を使

50

用して、例えばネットワークの全体キャパシティをより良好に使用するように、そのルーティングテーブルを調節することができる。対象処理ノードは、例えば10秒毎といった定期的な間隔でHELLOメッセージをブロードキャストする。その後処理ノードは、例えば毎秒1回といったより高いレートでALIVEメッセージをブロードキャストする。ピアノードは、ALIVE及びHELLOメッセージの間の比率を計算し、現在の例では10:1となる。HELLOメッセージは一定のレートで送信される。処理ノードが高い負荷若しくは輻輳を経験している場合、又は何らかの理由でノードにかかるトラフィックの負荷を削減することが望まれている場合、処理ノードはALIVEメッセージをいくらか遅れさせてもよい。処理ノードがダウンしている場合、又は処理ノード及びピアノードのインタフェースの間のパスがダウンしている場合、比率は無限大に近づく。

10

#### 【0019】

ピアノードは、比率を使用してそのルーティングテーブルを調節してもよく、それは例えば1より大きい比率を示す処理ノードに向けてのトラフィックの送信を削減し又は回避することによる。いかなる数のピアノードの間においても、ノードのいくつか又は全てを本発明の文脈において対象ノードと考えてもよく、いずれのピアも他のピアの処理キャパシティに関する情報を受信する必要がある。そのようなものとして、第1のノードは第2のノードから周期的なHELLO及びALIVEメッセージを受信する一方、それはまた同時に自身の周期的なHELLO及びALIVEメッセージを当該第2のノードに向けて送信し得る。

#### 【0020】

20

本発明の文脈において、プロセッサを含み、及びピア要素(peer element)との間で通信するための手段を有する任意の計算要素(computing element)をノードは含み得る。本発明を実装し得るノードの例は、コンピュータ内部のコプロセッサ、ルータ、サーバ、スイッチ等を含む。ピア要素は、例えば同一コンピュータ上の2つのコプロセッサであってもよいという点において互いに類似する要素であってもよく、又は例えばピアとしてのルータを有するサーバ若しくはピアとしての無線基地局を有するセルラ移動体の交換局を含むという点において互いに類似しない要素であってもよい。

#### 【0021】

ここで図面を参照すると、図1は、本発明のいくつかの観点に従って、ノードの処理キャパシティに関する情報を発信する一例となる方法を示している。シーケンス100のステップは対象ノードにおいて実装されてもよく、当該ノードは例えば通常の又は名目の(nominal)処理キャパシティを有し、ピアノードからの代わりにいくつかの機能の処理を実行するという要求を処理してもよく、及びときどきその処理負荷が予め決定された限界に到達しつつあることを検出してもよい。ステップ110において、対象ノードは提供キャパシティ値(OC)を決定する。提供キャパシティ値は、対象ノードがピアノードに提供し得る処理の最大量を表すプリセット値を超えることはない。対象ノードは、様々なやり方で提供キャパシティを決定し得る。ある実施形態においては、プロセッサの負荷がプリセットの限界の範囲内にとどまる限り、提供キャパシティは名目上の値であるとされてよく、例えばプロセッサの負荷が70パーセント未満であれば、名目の提供キャパシティが導かれ得る。またある実施形態においては、提供キャパシティはプロセッサの負荷の線形表現であってもよく、例えばプロセッサの負荷が10パーセントであるときには提供キャパシティは90パーセントとなり、又はプロセッサの負荷が60パーセントであるときには提供キャパシティは40パーセントとなる。提供キャパシティは、処理の要求で起こり得る急上昇からそのノードを保護するため、対象ノードの実際の処理キャパシティよりも低い値を表してもよい。実際の処理キャパシティと提供キャパシティとの間の非線形な関係を含む他の変化も、本発明から逸脱することなく使用されてもよい。

30

40

#### 【0022】

ステップ120において、対象ノードはノードの1つ以上のインタフェース上で第1の信号、例えばHELLOメッセージを送信する。HELLOメッセージは、対象ノードのID(Identity)を好適に含むが、1つの対象ノードしかその処理キャパシティに関する

50



情報を発信しない場合には当該IDは省略されてもよい。HELLOメッセージは、対象ノードの任意の数のピアノードに向けて送信され、対象ノードのハートビートとして使用される。以下に詳述されるように、シーケンス100は、好適に一定のレートである第1のレート、又は基本レートでHELLOメッセージが確実に送信されるようにする手法で継続的に繰り返される。その後対象ノードは、ステップ140と150と160とを含むループ130において、複数の第2の信号、例えばALIVEメッセージを送信する。ALIVEメッセージも、対象ノードのIDを好適に含む。ステップ150においてALIVEメッセージは送信され、対象ノードの同じインタフェース上で全てのピアノードに向けても送信される。提供キャパシティがない場合(OCは0に等しい)には、もちろんステップ150においてALIVE信号は送信されない。ステップ140と160とにより制御されるループ130における繰り返しは、各HELLOメッセージに続いてALIVEメッセージが確実にOCの回数(OC times)送信されるようにする。その結果、ALIVEメッセージは、提供キャパシティ値の関数である第2のレートで送信される。例えば、ALIVEメッセージのレートは、提供キャパシティ値により乗じられる基本レートに等しくてもよい。ループ130が完了すると、シーケンス100は代わりにステップ110又は120において継続してもよい。ある実施形態においては、毎回HELLOメッセージを送信する前に一度、提供キャパシティ値がステップ110において決定され得る。またある実施形態においては、対象ノード又はピアノードのアプリケーションに応じて、より少ない頻度で提供キャパシティ値がステップ110において決定され得る。シーケンス100と同等の、第2のレートが提供キャパシティと関連している他のシーケンスを使用して、第1及び第2のレートでHELLO及びALIVEメッセージを送信してもよい。例えば、HELLOメッセージを送信後、ALIVEメッセージの第1のインスタンスの前に、提供キャパシティが算出されてもよい。

#### 【0023】

HELLOメッセージを送信する基本レートが一定となるように、シーケンス100は一定のレートで好適に実行される。ピアノードが対象ノードの処理キャパシティをHELLO及びALIVEメッセージのレートを比較することにより評価できるようにする手段として、ALIVEメッセージのレートは可変である。基本レートがいくらか変化する実施形態において、ピアノードは毎回のHELLOメッセージに続いて受信されるALIVEメッセージの数を比較することによりやはり対象ノードの処理キャパシティを評価し得る。一方、基本レートが実質的に一定である場合、ピアノードは長期間にわたるALIVEメッセージのレートをHELLOメッセージの固定された基本レートと比較することにより、対象ノードの処理キャパシティの算出結果(calculations)を有利に平均化してもよい。

#### 【0024】

HELLO及びALIVEメッセージの周期性は、それらメッセージを送信し及び受信する様々なノードによりサービス提供されるアプリケーションに従って、ノードの性質に従って、又はノードが接続されている手法に従って、決定され得る。例えば同一のイーサネットリンクに、又は処理ラック(processing rack)内の同一のバックプレーン上で接続されることにより例えば対象ノードとそのピアノードとがISOレイヤ2レベルでリンクされる場合、HELLOメッセージは例えば毎秒1回のレートで送信されてもよく、ALIVEメッセージは、処理キャパシティが名目的であるときには毎秒数回までのレートで送信されてもよい。ノードがISOレイヤ3レベルで接続される場合、並びにHELLO及びALIVEメッセージがIPパケット又はATMフレームとして送信される必要がある場合、HELLOメッセージのレートは数秒毎に1回であってもよく、一方ALIVEメッセージのレートは、処理キャパシティが名目的であるときには毎秒1回のオーダであってもよい。1GHzオーダの処理クロックのスピードを有する同一のパーソナルコンピュータの2つ以上のコプロセッサは、毎ミリ秒ほどのオーダでHELLO及びALIVEメッセージを送信し得る。あるアプリケーションにおいては、メッセージが送信されるレートはさらにもっと高い又は低くてもよい。

## 【 0 0 2 5 】

図 1 が対象ノードの処理キャパシティに関する情報を発信する一例となる方法を示す一方、図 2 は、本発明の他の観点に従って、ピアノードにおいて対象ノードの処理キャパシティに関する情報を取得する一例となる方法を示している。シーケンス 2 0 0 のステップは、例えば対象ノードのピアノードにおいて実装されてもよい。これらのステップは、ピアノードにおいて継続的な手法で好適に実行されるが、不定期間隔で、例えばピアノードが対象ノードの処理キャパシティを評価する必要があるときに実行されてもよい。ピアノードが次のイベントが起こるのを待っているときに、ステップ 2 1 0 において方法が開始する。次のイベントは、ステップ 2 1 5 において第 1 のタイプの信号を受信することであってもよく、当該第 1 の信号は、例えば対象ノードにより送信される H E L L O メッセージである。ステップ 2 2 0 においてピアノードは第 1 の信号の受信の時間を記憶する (note)。ステップ 2 2 5 において、ピアノードはさらに進むための十分なデータが受信されているか否かを判定する。第 1 の信号が 1 回しか受信されていない場合、又はシーケンス 2 0 0 が再開されてから 1 回しか受信されていない場合、データは不十分であるとみなされ、次のイベントが期待されるステップ 2 1 0 へ処理は戻る。そうではなく、ステップ 2 1 5 において第 1 の信号が既に少なくとも 2 回は受信されている場合、ステップ 2 2 5 においてデータは十分であるとみなされ、第 1 の信号の到着レートが算出される次のステップ 2 3 0 へ進む。第 1 の信号が既に複数のインスタンスにおいて受信されている場合、ピアノードは全てのこれらインスタンスをカバーする時間にわたって第 1 の到着レートを任意に平均化してもよく、又は単純に第 1 の信号の最後の 2 つのインスタンスに基づいて到着レートを考慮してもよい。第 1 の到着レートは、対象ノードの処理キャパシティを評価するための基礎として使用されることになるため、原則として一定で稀にしか変化しない。一方、対象ノード又は対象ノードとピアノードとの間の接続リンクが、ダウンしているか、再起動しているか、又はそうでなければそれらの通常の機能を担うことができない場合には、第 1 の到着レートは変化する可能性がある。ステップ 2 5 5 においてシーケンスは継続し、それについては以下で詳述される。

## 【 0 0 2 6 】

ステップ 2 1 0 においてピアノードが待っているときに起こる別の次のイベントは、ステップ 2 3 5 において第 2 のタイプの信号を受信することであってもよく、当該第 2 の信号は、例えば対象ノードにより送信される A L I V E メッセージである。ステップ 2 4 0 においてピアノードは第 2 の信号の受信の時間を記憶する。ステップ 2 4 5 において、ピアノードはさらに進むための十分なデータが受信されているか否かを判定する。第 2 の信号が 1 回しか受信されていない場合、シーケンス 2 0 0 が再開されてから 1 回しか受信されていない場合、又は設定可能な値よりも少ない数のインスタンスにおいてしか受信されていない場合、データは不十分であるとみなされ、次のイベントが期待されるステップ 2 1 0 に処理は戻る。そうではなく、ステップ 2 3 5 において第 2 の信号が既に少なくとも 2 回受信されている場合、及び好適には第 2 の信号が設定可能な値に少なくとも等しい回数受信されている場合、ステップ 2 4 5 においてデータは十分であるとみなされ、第 2 の信号の到着レートが算出される次のステップ 2 5 0 へ進む。ピアノードは、そのような信号の数をカバーする時間にわたって第 2 の到着レートを好適に平均してもよい。好適な実施形態において、最後の第 1 の信号が受信されてからの期間にわたる第 2 の信号の到着レートが算出される。第 2 の到着レートは、対象ノードの現在の処理キャパシティの指標として使用されることになるため、変化してもよい。ステップ 2 5 0 における算出に続いて、ステップ 2 5 5 においてシーケンスは継続する。

## 【 0 0 2 7 】

ステップ 2 3 0 又はステップ 2 5 0 のいずれかに続き得るステップ 2 5 5 において、第 1 及び第 2 の信号の到着レートの比率が算出される。第 2 の信号は、第 1 の信号と比べてより高い到着レートを通常有していると期待されるため、好適な実施形態は通常の場合において、第 2 の到着レートを第 1 の到着レートで除して、1 よりも大きい比率を取得する。当業者は、第 1 のレートの比率を第 2 のレートで除する別の実施形態が同等の結果をも

たらずことを容易に観察するであろう。もちろん、ステップ 255 において、第 1 及び第 2 の到着レートの中の 1 つしかステップ 230 及び 250 において算出されていない場合、結果として生じる比率はゼロ又は無限大であり得る。そのような結果は、シーケンス 200 の文脈において非論理的ではなく、対象ノードの処理キャパシティに関する情報を取得する際に実際に使用され得る。ステップ 260 において、ステップ 255 において取得される比率がプリセット値と比較される。プリセット値は、ピアノードにおいて前もって設定され、対象ノードが通常の処理キャパシティのレベルである場合は第 1 及び第 2 の信号の到着レートの期待される比率に等しい。対象ノードの処理キャパシティは、比較から決定される。例えば、第 1 の信号の期待されるレートが毎 10 秒に 1 回である場合、及び第 2 の信号の期待されるレートが毎秒 1 回である場合、プリセット値は 10 に等しく、当該比率は第 2 の到着レートを第 1 の到着レートで除することにより取得される（比率が第 1 の到着レートを第 2 の到着レートで除することにより取得される実施形態においては、プリセット値は 10 分の 1（0.1）に等しくなる）。比率がプリセット値の 10 よりも小さい場合、これは対象ノードが低い処理キャパシティを示すために第 2 の信号を送信する速度を落としていることを示す。比率がゼロである場合、対象ノードは使用可能な処理キャパシティを有していないとみなされる。

#### 【0028】

シーケンス 200 がごく最近初期化されているとき、特別な場合が起こり得る。ステップ 230 において第 1 の到着レートが算出され、一方でステップ 245 において十分なデータが欠如していることにより第 2 の到着レートが算出できない場合、ステップ 255 においてゼロの比率が取得され得る。この場合ピアノードは、対象ノードが使用可能な処理キャパシティを有していないと好適にみなしてよい。対照的にステップ 255 において無限大の比率が取得されることは、第 1 の到着レートがまだ算出されていないことを示し得る。他の変形において、ステップ 225 及び 245 において十分なデータについて検証することなく、ステップ 230 及び 250 において第 1 及び第 2 の到着レートをそれぞれ算出することができる。所与のタイプの信号が 1 回しか受信されていない場合、この信号のタイプについての到着レートは極度に遅いとみなされ、またステップ 255 において極値の比率を生じさせるであろう。これらの場合のいずれにおいても、シーケンス 200 の継続的な実行により、程なくステップ 225 及び 245 において十分なデータが利用可能となり、ステップ 260 において対象ノードの処理キャパシティを適切に決定できることを保証するであろう。

#### 【0029】

他の場合において、対象ノードに通信の問題があるためステップ 225 又は 245 若しくはその両方において十分なデータが取得されない可能性がある。これにより比率がゼロ又は無限大のいずれに等しくなるとしても、これはピアノードに対して対象ノードの処理キャパシティが利用可能でないことを示している。

#### 【0030】

図 3 は、本発明の方法の追加的観点のステップを示す。シーケンス 300 は、図 2 のシーケンス 200 と並行して実行されてもよい。特に、シーケンス 300 は図 2 のステップ 260 に続いて開始してもよく、シーケンス 200 の継続中の処理において対象ノードの処理キャパシティが再評価されるときはいつでも実行されてもよい。代わりにシーケンス 300 は、不定期間隔で、第 1 及び第 2 の信号のレートよりも低いレートで周期的に、又はピアノードの特定の必要性に基づいて必要に応じて、実行されてもよい。

#### 【0031】

ステップ 260 において対象ノードの処理キャパシティが決定された後、ピアノードはさらなるアクションを取り得る。これらアクションのうちのいくつかは補完的であってもよく、ピアノードは、これらアクションのうちのいくつかを他のアクションとは独立してとってもよい。ピアノードは、比率における変化を監視することにより、対象ノードの処理キャパシティにおける変化を監視し得る。ステップ 310 において、ステップ 255 において取得される比率が、前もって記憶されている比率の値と比較される。ステップ 31

5 において、この比較の結果は3つの異なる値を有し得る。ステップ315における結果が比率の増加を示している場合、ピアノードはステップ320において、対象ノードの処理キャパシティが増加したことを記録 (record) し、又は別の方法でメモリに記憶 (take note) する。ステップ315における結果が比率の減少を示している場合、ピアノードはステップ325において、対象ノードの処理キャパシティが減少したことを記録する。ステップ315における結果が比率が変化していないことを示している場合、特定のアクションは取られない。

#### 【0032】

ステップ310～325が実行されるか否かに関わらず、ステップ330においてシーケンスは継続されてよく、ステップ260において決定された対象ノードの処理キャパティは、予め決定された閾値と比較して評価される。処理キャパティが閾値よりも低い場合、ステップ335においてピアノードはそのルーティングテーブルから対象ノードを削除してもよい。その結果、ピアノードは対象ノードに向けて信号、メッセージ又は要求を送信することを控えるであろう。もちろん、対象ノードのルーティングテーブルからの削除は、場合によってはトラフィック及びペイロードの交換にのみ適用可能であって、保守 (maintenance) シグナリングには適用されなくてもよい。例えば、ピアノードは対象ノードに向けての自身の周期的なHELLO及びALIVE信号の送信を継続してもよい。その代わりに、ピアノードは、そのルーティングテーブルにおいて対象ノードに与えられるコストを、対象ノードに向けてより少ない数の信号、メッセージ又は要求が送信されることを保証するように、より高いコストに調整してもよい。ステップ330において、対象ノードの処理キャパティが高く又は名目のレベルであるとみなされ得る。例えば、閾値を超える処理キャパティのレベルは、高いとみなされ得る。代わりに、ステップ255において算出される比率がプリセット値に等しいとわかった場合にのみ、キャパティは名目上の値であるとされてもよい。処理キャパティを高い値又は名目上の値とするように決定するためのどのように精緻な基準を使用するとしても、ステップ340において、対象ノードは、ピアノードのルーティングテーブル内に追加され又は維持される。対象ノードの処理キャパティが高いレベル又は名目のレベルにごく最近戻った場合、ノードはルーティングテーブルに追加される。その代わりに、処理キャパティが高く又は名目上の値であるとされる場合、ルーティングテーブルにおいて対象ノードに与えられるコストは調整され得る。処理キャパティがしばらくの間継続的に名目上の値である場合、ノードは既にルーティングテーブル内に存在しており、そこに維持され、ノードに与えられるコストは低い値に維持される。

#### 【0033】

図2のシーケンス200に示されている方法は、対象ノードの処理キャパティを決定する処理において精度を上げるための追加的観点を含み得る。第1及び第2の信号は、そのインスタンスの全てにおいて、シーケンス番号を伴って送信されてもよい。第1の信号が第1のインスタンスにおいて送信される場合、それは基本シーケンス番号を含む。第1のインスタンスにおいて送信される基本シーケンス番号の実際の値は、多くの実施形態において関連がない。それから第1の信号の次のインスタンスは、式(1)に従ったシーケンス番号を伴う：

#### 【0034】

#### 【数1】

$$SQN_{firstsignal} = F_1(SQN_{previousfirstsignal}) \quad (1)$$

#### 【0035】

ここで  $F_1$  は、第 1 の信号のシーケンス番号と前回の第 1 の信号のシーケンス番号との関係である。

【 0 0 3 6 】

最新の第 1 の信号の直後に送信される第 2 の信号の第 1 のインスタンスは、式 ( 2 ) に従ったシーケンス番号を伴う：

【 0 0 3 7 】

【 数 2 】

$$SQN_{secondsignal} = F_2(SQN_{latestfirstsignal}) \quad (2)$$

10

【 0 0 3 8 】

ここで  $F_2$  は、第 2 の信号のシーケンス番号と最新の第 1 の信号のシーケンス番号との関係である。

【 0 0 3 9 】

その後次回の第 1 の信号の前に送信される第 2 の信号の他のインスタンスは、式 ( 3 ) に従ったシーケンス番号をそれぞれ伴う：

20

【 0 0 4 0 】

【 数 3 】

$$SQN_{secondsignal} = F_3(SQN_{previoussecondsignal}) \quad (3)$$

【 0 0 4 1 】

ここで  $F_3$  は、第 2 の信号のシーケンス番号と前回の第 2 の信号のシーケンス番号との関係である。

30

【 0 0 4 2 】

これらの関数が一義的なシーケンスの計算を可能としている限り、及び対象ノードとピアノードとが同じ関数を使用している限り、第 1 及び第 2 の信号に伴われるシーケンス番号を計算するために、様々な関数  $F_1$ 、 $F_2$ 、及び  $F_3$  の組が使用され得る。

【 0 0 4 3 】

関数  $F_1$ 、 $F_2$ 、及び  $F_3$  を使用する単純な一例において、第 1 の信号の第 1 のインスタンスにおいて伴われる基本シーケンス番号は、ゼロに等しい。関数  $F_1$  は、前回の第 1 の信号のシーケンス番号にプリセット値に等しい数を加算し、ここで例えばプリセット値は 1 0 に等しい。関数  $F_2$  は、最新の第 1 の信号のシーケンス番号に 1 を加算する。関数  $F_3$  は、前回の第 2 の信号のシーケンス番号に 1 を加算する。この例において、第 1 の信号は第 1 のインスタンスにおいて、0 に等しい基本シーケンス番号と共に送信される。一連の第 2 の信号は、それから  $F_2$  を使用して算出される 1 に等しいシーケンス番号と共に、それから  $F_3$  を使用して算出される 2、3、・・・、1 0 に等しいシーケンス番号と共に送信される。それから第 1 の信号の次のインスタンスは、 $F_1$  を使用して算出される 1 0 に等しいシーケンス番号を伴って送信される。

40

【 0 0 4 4 】

他の実施形態において、より複雑な関数が使用されてもよい。一例として、 $F_2$  及び  $F_3$  はシーケンス番号に作用する一方向性のハッシュ法の手続きを含んでもよく、一方  $F_1$

50

は、同じ一方向性のハッシュ法をプリセット値に等しい回数シーケンス番号に作用させてもよい。

【 0 0 4 5 】

図 4 は、対象ノードの処理キャパシティを決定する処理においてより精度を上げるためのいくつかの例となる観点を示す。シーケンス 4 0 0 及び 4 5 0 は、図 2 のシーケンス 2 0 0 のステップのうちのいくつかと並行してピアノードにおいて動作し、シーケンス 2 0 0 のオプションとしてのサブルーチンとみなされ得る。サブルーチン 4 0 0 では、図 2 の前述の説明において示されているように、ステップ 2 1 5 においてピアノードにより第 1 の信号が受信される。第 1 の信号は、任意にシーケンス番号を含み得る。そうである場合、ステップ 4 1 7 において次回の第 2 の信号の期待されるシーケンス番号が、関数  $F_2$  を使用して算出される。ステップ 4 1 9 において、サブルーチンは図 2 におけるステップ 2 1 5 に続くステップ、即ちステップ 2 2 0 に戻る。サブルーチン 4 5 0 では、図 2 に示されているように、ステップ 2 3 5 において第 2 の信号がピアノードにより受信される。第 2 の信号もシーケンス番号を含み得る。ステップ 4 5 2 において、第 2 の信号に含まれるシーケンス番号が期待されるシーケンス番号に合致するか否かがチェックされる。ある実施形態においては、受信されるシーケンス番号と期待されるシーケンス番号とが等しい場合に、合致しているとされる。またある実施形態においては、受信されるシーケンス番号が期待されるシーケンス番号以上である場合に、合致しているとされ得る。合致している場合には、ステップ 4 5 4 において次回の第 2 の信号に適用可能な別の期待されるシーケンス番号が算出され、ステップ 4 5 9 においてサブルーチンは図 2 におけるステップ 2 3 5 に続くステップ、即ちステップ 2 4 0 に戻る。ステップ 4 5 2 において第 2 の信号に含まれるシーケンス番号が期待されるシーケンス番号に合致しないとみなされた場合、ピアノードはこの第 2 の信号又は別の第 2 の信号が失われ若しくは遅延していることを検出する。ステップ 4 5 6 において、ピアノードは、キャパシティの問題が起こっていることを検出する。キャパシティの問題は、対象ノードにおける処理キャパシティの問題、又はピアノードと対象ノードとの間の通信リンク上の問題であり得る。ステップ 4 5 8 において、ピアノードは図 2 のさらなる処理において無視されるべき第 2 の信号にマークを付ける。それからサブルーチンは、ステップ 4 5 9 において図 2 のシーケンスに戻る。

【 0 0 4 6 】

ある実施形態においては、類似のロジックを使用して次回の第 1 の信号の期待されるシーケンス番号が算出され得る。同等の手法において、次回の第 1 の信号について期待されるシーケンス番号に合致しないシーケンス番号を含む次回の第 1 の信号を受信することは、対象ノードにおける問題又はピアノードと対象ノードとの間の通信リンク上における問題のいずれかを示すであろう。

【 0 0 4 7 】

図 5 は本発明のある観点到に係る一例となるノードを示しており、ここで図 5 を参照しながら一例となるノードの構成を説明する。例となるノード 5 0 0 は、自身の処理キャパシティに関する情報を発信することと、ピアノードの処理キャパシティに関する情報を取得することとの両方を可能とする特徴を含む。他の例となるノードは、これら特徴のいくつか、即ち処理キャパシティの情報の発信又は取得のいずれかのみを可能とすることを含み得ることは理解されるべきである。これら他の例となるノードは、それらが使用されるコンピューティングネットワークアーキテクチャの必要性に従って、例となるノード 5 0 0 の特徴のサブセットを含んでもよい。

【 0 0 4 8 】

ノード 5 0 0 は、メモリ 5 1 0、プロセッサ 5 2 0、プロセッサの負荷インジケータ 5 3 0、クロック 5 4 0、制御ロジック 5 5 0、及び 1 つ以上のインタフェース 5 6 0 を含む。インタフェース 5 6 0 の各々は、1 つの単独の装置として、又はシグナリング、メッセージ、及びデータを受信（入力）し及び送信（出力）するための異なる装置として実装されてもよい。ノードは、複数のピアノードに向けて接続されており；ピアノードに向けてノードを接続するための手段は、例えばあるピアノードに向けての接続はイーサネット

リンク上にあってもよく、一方またあるピアノードに向けての接続は非同期転送モード（ＡＴＭ）リンク上にあってもよいというように、変わってもよい。そのためノード５００は、異なるタイプの複数のリンク上で接続するための複数のインタフェース５６０を含み得る。他の機能をさらに有し得るノードは、そのため当該技術分野において良く知られているようなより多くのコンポーネントを含んでもよい。ある実施形態においては、ノード５００は完全なシステム、例えばコンピュータである。またある実施形態においては、ノード５００は基本的に、例えばプロセッサ５２０、負荷インジケータ５３０、クロック５４０、及び制御ロジック５５０を含む単一の処理要素から成る。

#### 【００４９】

メモリ５１０は、いくつかのその内容のために、例えばランダムアクセスメモリ（ＲＡＭ）を使用して実装されるテンポラリメモリであってもよい。いくつかの他の観点のために、それは好適に、電氣的に消去することができ、再プログラミング可能であり、及び例えばフラッシュメモリ若しくはデータ記憶モジュールとして実装され得る、不揮発性メモリ又は持続性メモリである。メモリ５１０は、例えば図２のシーケンス２００におけるステップ２６０で使用されるノード５００において利用可能な処理の最大量を表すプリセット値、及び図３のシーケンス３００におけるステップ３３０において使用される予め決定された閾値を含むいくつかの定数５１２を記憶する。

#### 【００５０】

メモリ５１０はまた、ピアノードのテーブル６００をこれらピアノードの処理キャパシティに関する情報を記憶するために記憶してもよい。図６は、本発明のいくつかの観点に係る一例となるピアテーブル６００を示す。テーブル６００は、複数の行６２０～６２８を含み、これらはヘッダ行である第１の行６２０を除いて、テーブルを有するノードのピアに関する。テーブルは、ピアのリストにおけるノードの処理キャパシティに関する情報を取得し及び記憶するためにノードにより使用される。ピアテーブル６００は、複数の列６０２～６１６を含み、これらのうちのいくつかはオプションである。第１の列６０２は、ピアのＩＤ  $P_n$  を含む。第２の列は、第１の信号、例えばＨＥＬＬＯ信号を受信した最新のインスタンスの時間  $T_{1n}$  を記憶する。第１の信号の新たなインスタンスが受信されるたびに、受信の時間を前回の受信の時間  $T_{1n}$  と比較して、列６０６に記憶される第１の信号のレート  $R_{1n}$  を計算する。列６０４において新たな受信の時間は前回の値を上書きする。もちろん、他の同等の実施形態を使用して第１の信号の受信のレートを追跡することもできる。ある実施形態においては、列６０６における第１の信号のレートは、単純に第１の信号の全ての受信で上書きされる。またある実施形態においては、列６０６の第１の信号のレート  $R_{1n}$  は、時間にわたって平均されてもよい。またある実施形態においては、列６０６に記憶するためのより正確なレート  $R_{1n}$  を計算するために、列６０４において毎行１つの  $T_{1n}$  の値を記憶するのではなく、第１の信号の時間の値の配列を記憶することもできる。列６０８及び６１０は同様に、第２の信号、例えばＡＬＩＶＥ信号の受信の時間  $T_{2n}$  及びレート  $R_{2n}$  をそれぞれ記憶する。列６１２において、２つの信号のレートの比率が各ピアについて記憶される。例となるピアテーブル６００において、比率は第２の信号のレート  $R_{2n}$  を第１の信号のレート  $R_{1n}$  で除することにより取得され、１０となる比率は名目のレートであることを示す。任意的な列６１４は、所与の行のピアについてのルーティングのコストを記憶する。コストは数値であり得、低いコストはピアへ向けてのルーティングが好適であることを示す。それは単純に高いか低いかの表示でもあり得る。コストは、ピアへ向かうルーティングが許可されるか否かを示す２値表示であってもよい。ある実施形態においては、ルーティングのコストは、ピアテーブル６００とは異なる別のルーティングテーブルにおいて記憶されてもよい。別の任意的な列６１６は、次の第２の信号についての期待されるシーケンス番号（Ｅ－ＳＱＮ） $E_n$  を記憶する。ピアテーブル６００は、ノードのピアに関する追加的情報を維持するために他の列（図示しない）も含んでもよい。

#### 【００５１】

ピアテーブル６００において、行６２２～６２８上の例となるピア  $P_1 \sim P_4$  は、これ

10

20

30

40

50

らピアの処理キャパシティに関連する様々な状況を示している。行622上において、10に等しい比率により示されるように、 $P_1$ の処理キャパシティはその名目のレベルである。トラフィックは $P_1$ に向けてルーティングされてよく、ルーティングのコストは低い。行624上において、 $P_2$ から第1の信号は少なくとも1回受信されたが、第2の信号はまだ受信されていない。 $P_2$ について第2の信号のレートは算出されなくてもよく、従って比率はゼロに等しい。トラフィックは $P_2$ に向けてルーティングされ得ず、ルーティングのコストは高い。行626上において、 $P_3$ から第2の信号しか受信されていない。第1の信号のレート $R_{13}$ は設定されていないため、第1の信号のレートについての空値でレート $R_{23}$ を割った比率は無限大という結果を生じさせる。これもまたルーティングの問題を示しているため、ルーティングのコストは高い。行628上において、 $P_4$ につ

10

#### 【0052】

図6においてテーブルの形でピアテーブルの内容が表されているが、その内容がこれに限定されないことは留意されるべきである。ピアテーブルは、関連のデータベース、ハードコードされたコンポーネント、マイクロプロセッサ、プログラミングライブラリ等から構成され得る。

#### 【0053】

プロセッサ520は、ノード500において実装される特徴に係るタスクを実行する。負荷インジケータ530は、プロセッサ520上の負荷レベルを継続的に反映する。クロック540は、制御ロジック550におけるシーケンスを周期的に開始する。各シーケンスの始めに、制御ロジック550は負荷インジケータ530からプロセッサの負荷を読み取る。制御ロジック550は、メモリ510における定数512からプリセット値を読み取る。制御ロジック550は、第1の信号、例えばHELLOメッセージをシーケンス毎に1回送信するよう1つ以上のインタフェース560に要求する。制御ロジック550は、第1の信号に含めるためのシーケンス番号を任意に生成してもよく、1回のシーケンスの間に送信される第1の信号の1つのインスタンスのシーケンス番号は、先行するシーケンスにおける第1の信号のシーケンス番号に基づいて好適に生成される。それから制御ロジック550は、プロセッサの負荷及びプリセット値からシーケンスの間に送信されるべき第2の信号、例えばALIVEメッセージの数 $N$ を決定する。プロセッサの負荷が低い場合、数 $N$ はプリセット値に等しく設定される。プロセッサの負荷が高いとみなされるべきか低いとみなされるべきかの判定は、メモリ510に記憶されているいくつかの定数512に依存し得る。制御ロジック550は、ALIVEメッセージに含まれるべきシーケンス番号を算出してもよく、連続するシーケンス番号の各々は、前回送信された信号のシーケンス番号に基づいて好適に算出される。制御ロジック550は、決定された数 $N$ の第2の信号を、第1の信号のレートの $N$ 倍に相当するレートで送信するよう1つ以上のインタフェース560に要求する。

20

30

#### 【0054】

インタフェース560のうちの1つは、現在自身の処理キャパシティに関する情報を発信している対象のピアノード $P_n$ から一連の第1及び第2の信号を受信し得る。当該インタフェースは、制御ロジック550に情報を提供する。制御ロジック550は、クロック540からタイミング情報を読み取ることにより第1又は第2の信号の受信の時間を読み取り、及びピアテーブル600においてピアノードのID  $P_n$ を含む行上に、第1の信号の時間 $T_{1n}$ 又は第2の信号の時間 $T_{2n}$ を記憶する。第1又は第2の信号が十分な回数受信されている限り、プロセッサは信号のレート $R_{1n}$ と $R_{2n}$ とを算出し、及びピアノード $P_n$ についてそれらを記憶する。両方のレートがピアノード $P_n$ について算出されると、制御ロジック550はピアノード $P_n$ の処理キャパシティを示す比率を計算し、及びテーブルにおいて列612にそれを記憶する。制御ロジック550は、ピアノード $P_n$

40

50



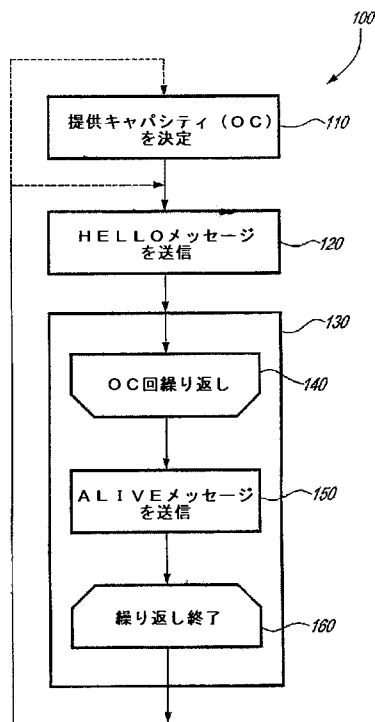
についてのルーティングのコストの値を任意的に記憶してもよい。第 1 及び第 2 の信号は、シーケンス番号を含んでもよい。制御ロジック 550 は、ピアノード  $P_n$  から受信される次回の第 2 の信号に適用可能な期待されるシーケンス番号 ( $E - SQN$ ) を計算し、及びそれを  $E_n$  として記憶する。次回の第 2 の信号がピアノード  $P_n$  から受信される際、当該次回の第 2 の信号に含まれるシーケンス番号が  $E_n$  に適合しない場合、制御ロジック 550 は当該第 2 の信号を無視してもよい。

【0055】

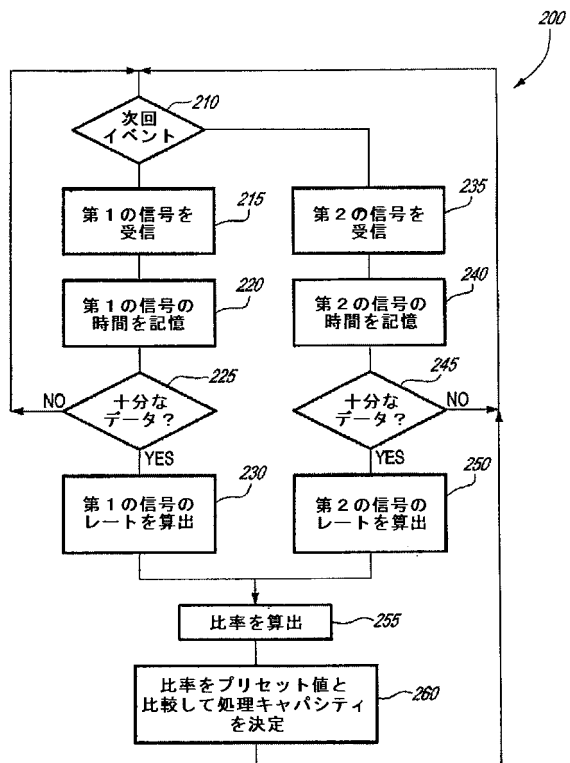
本発明の方法及びノードの好適な実施形態のいくつかの観点が付図の図面において示され、及び前述の詳細な説明において説明されているが、本発明は開示される実施形態に限定されず、次の特許請求の範囲により説明され及び定義される本発明の精神から逸脱することなく多数の再構成、変更、及び置換が可能であることは理解されるであろう。

10

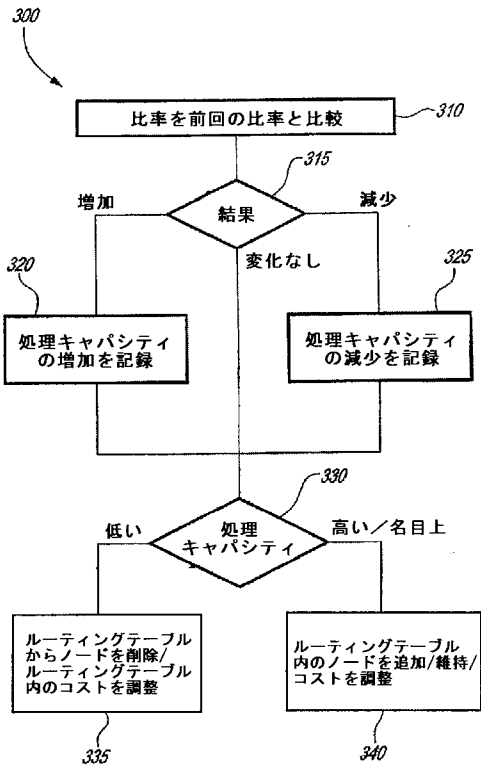
【図 1】



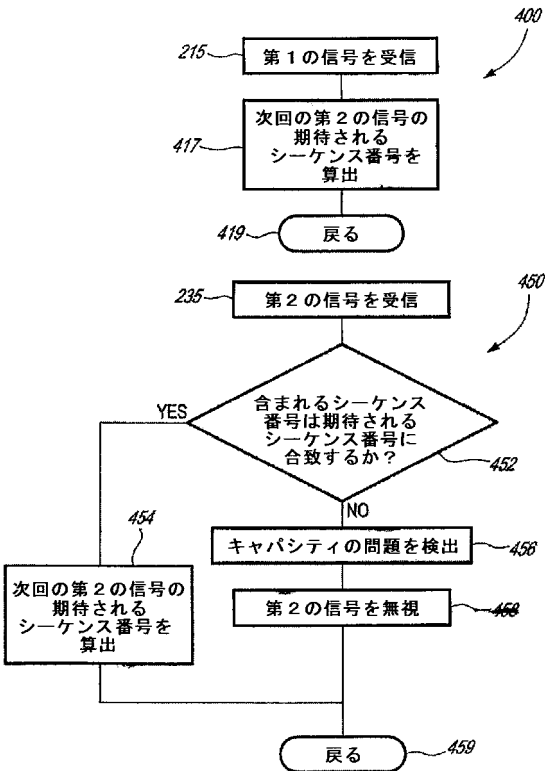
【図 2】



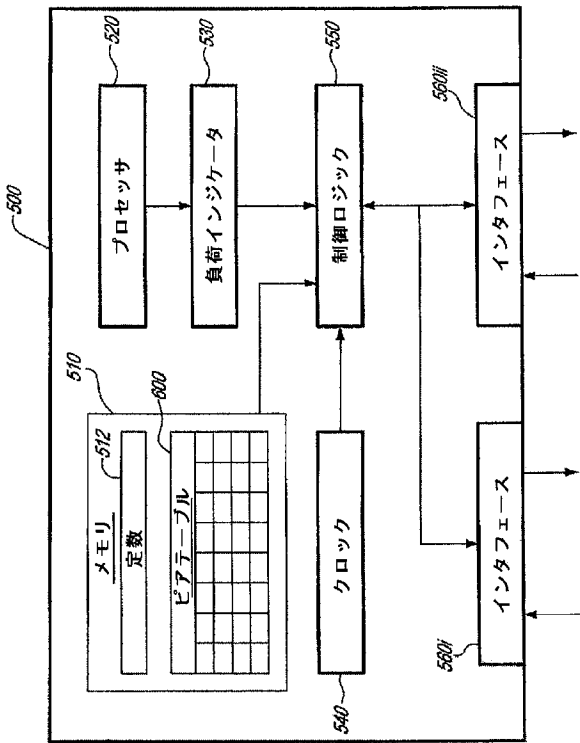
【図 3】



【図 4】



【図 5】



【図 6】

600					
602					
604					
606					
608					
610					
612					
614					
616					
620					
622					
624					
626					
628					
第1の信号第1のレート の時間	T <sub>11</sub>	T <sub>12</sub>	—	T <sub>14</sub>	...
第1の信号第2のレート の時間	R <sub>11</sub>	R <sub>12</sub>	0	R <sub>14</sub>	...
第2の信号第2のレート の時間	R <sub>21</sub>	0	R <sub>23</sub>	R <sub>24</sub>	...
比率	10	0	∞	7	...
ルーティング テーブル 内のコスト	低い	高い	高い	中程度	...
E-SON	E <sub>1</sub>	E <sub>2</sub>	E <sub>3</sub>	E <sub>4</sub>	...

---

フロントページの続き

(72)発明者 ロッシ、フレデリック

カナダ ケベック州 エイチ３ティー １エム３ モントリオール ３７９３ ラコンベ

審査官 永井 啓司

(56)参考文献 特開２００１－０７５９３４（ＪＰ，Ａ）

特開２００４－１７１２２６（ＪＰ，Ａ）

(58)調査した分野(Int.Cl.，ＤＢ名)

G 0 6 F １ 3 / 0 0、

H 0 4 L １ 2 / 0 0 - 1 2 / 2 6、1 2 / 5 0 - 1 2 / 9 5 5