

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4685672号
(P4685672)

(45) 発行日 平成23年5月18日(2011.5.18)

(24) 登録日 平成23年2月18日(2011.2.18)

(51) Int.Cl.

H04L 12/56 (2006.01)

F 1

H04L 12/56 400Z
H04L 12/56 200Z

請求項の数 12 (全 14 頁)

(21) 出願番号 特願2006-75333 (P2006-75333)
 (22) 出願日 平成18年3月17日 (2006.3.17)
 (65) 公開番号 特開2006-270953 (P2006-270953A)
 (43) 公開日 平成18年10月5日 (2006.10.5)
 審査請求日 平成21年3月17日 (2009.3.17)
 (31) 優先権主張番号 11/088,130
 (32) 優先日 平成17年3月23日 (2005.3.23)
 (33) 優先権主張国 米国(US)

(73) 特許権者 500046438
 マイクロソフト コーポレーション
 アメリカ合衆国 ワシントン州 9805
 2-6399 レッドモンド ワン マイ
 クロソフト ウェイ
 (74) 代理人 100077481
 弁理士 谷 義一
 (74) 代理人 100088915
 弁理士 阿部 和夫
 (72) 発明者 ジテンドラ ディー・パディエ
 アメリカ合衆国 98052 ワシントン
 州 レッドモンド ワン マイクロソフト
 ウェイ マイクロソフト コーポレーション内

最終頁に続く

(54) 【発明の名称】利用可能な帯域幅推定

(57) 【特許請求の範囲】

【請求項 1】

少なくとも送信元コンピューティングデバイスと受信側コンピューティングデバイスの間におけるネットワークバス上の利用可能な帯域幅を測定する方法であって、前記ネットワークバスは、少なくとも1つが、前記ネットワークバスにおける他のいずれかのリンクよりも少ない量の利用可能な帯域幅を有するタイトリンクである複数のリンクを含み、前記方法は、

前記タイトリンクを通過した後で且つ前記受信側コンピューティングデバイスに到達する前に相対的に大きいパイロットパケットを伝送するステップと、

前記タイトリンクを通過した後で且つ前記受信側コンピューティングデバイスに到達する前に前記大きいパイロットパケットをドロップするステップと、前記大きいパイロットはタイム・ツー・ライブ (Time-to-live) パラメータを有し、前記大きいパイロットパケットをドロップするステップは、そのタイム・ツー・ライブの有効期限消滅時に前記大きいパイロットパケットをドロップすることを含み、

前記大きいパイロットパケットの前記伝送に続いて直ちに、パケットペアを前記送信元コンピューティングデバイスから前記受信側コンピューティングデバイスに伝送するステップであって、前記パケットペアは、初期遅延によって隔てられた少なくとも2つのパケットを含むが別の大きいパイロットパケットを含まず、前記少なくとも2つのパケットは前記大きいパイロットパケットよりも相対的に小さいサイズを有するステップと、

前記送信元コンピューティングデバイスにおいて、前記受信側デバイスから、前記パケ

10

20

ットペアのパケット間の、該パケットが前記受信側コンピューティングデバイスに到着した際の相対遅延の指示を獲得するステップと、

前記送信元コンピューティングデバイスにおいて、前記ネットワークパスに関する利用可能な帯域幅値を、前記パケットペアのパケット間の、該パケットが前記受信側コンピューティングデバイスに到着した際の相対遅延に基づいて計算するステップと
を備えることを特徴とする方法。

【請求項 2】

前記パイロットパケットの前記伝送に続いて直ちに、前記送信元コンピューティングデバイスから前記受信側コンピューティングデバイスに伝送するステップであって、前記パケットペアは、初期遅延によって隔てられた少なくとも 2 つのパケットを含む前記ステップは、前記タイリンクに向かう途中で前記パケットペアによって経験されることが予期される拡大の量を計算し、前記予期される拡大を考慮して前記ペアの前記パケット間の前記初期遅延を設定するステップをさらに含むことを特徴とする請求項 1 に記載の方法。 10

【請求項 3】

前記送信元コンピューティングデバイスにおいて、前記受信側デバイスから、前記パケットペアのパケット間の、該パケットが前記受信側コンピューティングデバイスに到着した際の相対遅延の指示を獲得する前記ステップは、前記受信側コンピューティングデバイスから、前記パケットが前記受信側コンピューティングデバイスに到着したそれぞれの時刻を示すタイムスタンプ情報を獲得するステップを含むことを特徴とする請求項 1 に記載の方法。 20

【請求項 4】

前記ネットワークパスに関する利用可能な帯域幅値を、前記パケットペアのパケット間の、該パケットが、前記受信側コンピューティングデバイスに到着した際の相対遅延の指示に基づいて計算する前記ステップは、前記パケットペアのパケット間の、該パケットが前記受信側コンピューティングデバイスに到着した際の相対遅延と、前記パケット間の、該パケットが前記タイリンクに到着した際の遅延との差を計算するステップをさらに含むことを特徴とする請求項 1 に記載の方法。

【請求項 5】

所定の回数、前記ステップセットを繰り返すステップをさらに備えることを特徴とする請求項 1 に記載の方法。 30

【請求項 6】

所定の回数、前記ステップセットを繰り返す前記ステップは、前記ネットワークパスに関する複数の利用可能な帯域幅値をもたらし、前記方法は、

前記複数の利用可能な帯域幅値を平均して、平均の利用可能な帯域幅をもたらすステップをさらに備えることを特徴とする請求項 5 に記載の方法。

【請求項 7】

少なくとも送信元コンピューティングデバイスと受信側コンピューティングデバイスの間ににおけるネットワークパス上の利用可能な帯域幅を測定する方法であって、前記ネットワークパスは、少なくとも 1 つが、前記ネットワークパスにおける他のいずれかのリンクよりも少ない量の利用可能な帯域幅を有するタイリンクである複数のリンクを含む方法を実行する、コンピュータ読み取り可能な命令をその上に有するコンピュータ可読記録媒体であって、前記命令は、 40

前記タイリンクを通過した後で且つ前記受信側コンピューティングデバイスに到達する前に相対的に大きいパイロットパケットを传送すること、

前記タイリンクを通過した後で且つ前記受信側コンピューティングデバイスに到達する前に前記大きいパイロットパケットをドロップすること、前記大きいパイロットはタイム・ツー・ライブ (Time-to-live) パラメータを有し、前記大きいパイロットパケットをドロップするステップは、そのタイム・ツー・ライブの有効期限消滅時に前記大きいパイロットパケットをドロップすることを含み、

前記大きいパイロットパケットの前記伝送に続いて直ちに、パケットペアを前記送信元 50

コンピューティングデバイスから前記受信側コンピューティングデバイスに伝送すること、前記パケットペアは、初期遅延によって隔てられた少なくとも2つのパケットを含むが別の大きいパイロットパケットを含まず、前記少なくとも2つのパケットは前記大きいパイロットパケットよりも相対的に小さいサイズを有し、

前記送信元コンピューティングデバイスにおいて、前記受信側デバイスから、前記パケットペアのパケット間の、該パケットが前記受信側コンピューティングデバイスに到着した際の相対遅延の指示を獲得すること、

前記送信元コンピューティングデバイスにおいて、前記ネットワークバスに関する利用可能な帯域幅値を、前記パケットペアのパケット間の、該パケットが前記受信側コンピューティングデバイスに到着した際の相対遅延に基づいて計算すること、

10

のための命令を含むことを特徴とするコンピュータ可読記録媒体。

【請求項8】

前記パイロットパケットの前記伝送に続いて直ちに、前記送信元コンピューティングデバイスから前記受信側コンピューティングデバイスに伝送することであって、前記パケットペアは、初期遅延によって隔てられた少なくとも2つのパケットを含む前記伝送とのための前記命令は、前記タイトリンクに向かう途中で前記パケットペアによって経験されることが予期される拡大の量を計算し、前記予期される拡大を考慮して前記ペアの前記パケット間の前記初期遅延を設定することのための命令をさらに含むことを特徴とする請求項7に記載のコンピュータ可読記録媒体。

【請求項9】

20

前記送信元コンピューティングデバイスにおいて、前記受信側デバイスから、前記パケットペアのパケット間の、該パケットが前記受信側コンピューティングデバイスに到着した際の相対遅延の指示を獲得することのための前記命令は、前記受信側コンピューティングデバイスから、前記パケットが、前記受信側コンピューティングデバイスに到着したそれぞれの時刻を示すタイムスタンプ情報を獲得することのための命令を含むことを特徴とする請求項7に記載のコンピュータ可読記録媒体。

【請求項10】

前記ネットワークバスに関する利用可能な帯域幅値を、前記パケットペアのパケット間の、該パケットが、前記受信側コンピューティングデバイスに到着した際の相対遅延の指示に基づいて計算することのための前記命令は、前記パケットペアのパケット間の、該パケットが前記受信側コンピューティングデバイスに到着した際の相対遅延と、前記パケット間の、該パケットが前記タイトリンクに到着した際の遅延との差を計算することのための命令をさらに含むことを特徴とする請求項7に記載のコンピュータ可読記録媒体。

30

【請求項11】

所定の回数、前記命令を繰り返すための命令をさらに備えることを特徴とする請求項7に記載のコンピュータ可読記録媒体。

【請求項12】

所定の回数、前記命令を繰り返すことは、前記ネットワークバスに関する複数の利用可能な帯域幅値をもたらし、前記コンピュータ可読記録媒体は、

前記複数の利用可能な帯域幅値を平均して、平均の利用可能な帯域幅をもたらすための命令をさらに備えることを特徴とする請求項11に記載のコンピュータ可読記録媒体。

40

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、一般に、コンピュータネットワーキングに関し、より詳細には、異なるタイトリンク(tight link)およびナローリンク(narrow link)を有するネットワークバス上の帯域幅を推定する方法およびシステムに関する。

【背景技術】

【0002】

ネットワーキング環境では、ネットワーキングタスクに関わるエンティティが、特定の

50

用途のためにどれだけの帯域幅が利用可能であるか、または利用可能である可能性があるかを推定できることがしばしば、重要である。例えば、不十分な帯域幅しか利用可能でない場合、不十分に実行されなければならない、または後で途中停止されなければならないようなタスクを開始することは意味をなさないことがある。しばしば、オーディオ情報およびビデオ情報のストリーミング伝送は、満足の行くユーザ体験を提供するために、あるサービス品質を必要とし、そのサービス品質が利用できない場合、ユーザ体験は乏しいものとなる。

【0003】

近年、ネットワーキング技術の人気の高まりが見られ、このため、利用可能な帯域幅を算出することを目的とする技術の増加が見られている。残念ながら、それらの技術の多くは、現実において裏付けのない想定を行い、このため、利用可能な帯域幅の不正確な推定をもたらす。例えば、パケットギャップ方法（PGM）は、インターネットパス上のナローリンクとタイトリンクが同一である場合にだけ、正しく機能する。

【0004】

以下の参照文献（非特許文献1、非特許文献2）は、帯域幅測定のいくつかの初步的な様子を理解するのにユーザの役に立ちうる。

【0005】

【非特許文献1】K. Lai and M. Baker, "Measuring Link Bandwidths Using a Deterministic Model of Packet Delay", ACM SIGCOMM, August 2000

【非特許文献2】J. Strauss and D. Katabi and F. Kaashoek, "A Measurement Study of Available Bandwidth Estimation Tools", ACM/USENIX IMC, October 2003

【発明の開示】

【課題を解決するための手段】

【0006】

ネットワークパスにおいて、タイトリンクが、ナローリンクと同一ではない場合、利用可能な帯域幅推定の既存のPGM技術に固有の不可欠な仮定が崩れ、それらの技術は、もはや機能しない。本発明の諸実施形態では、固有のパケットプローブシリーズ構造（unique packet probe series structure）および処理を使用して、利用可能な帯域幅を推定する。詳細には、プローブパケット間の間隔が、タイトリンクまでに至るリンク群によって生じさせられる拡大（dilation）を考慮して送信元においてまず調整されて、プローブがタイトリンクに到着した際に、間隔が適切であるようになる。さらに、本発明の一実施形態において、マルチパケットプローブが使用される。マルチパケットプローブは、本明細書で、ときとして、「パイロット」パケットと呼ばれる大きいパケットと、その後に続く2つのはるかに小さいパケットとを含む。大きいパケットは、次に、タイトリンクを通過すると、ドロップされる。その後のナローリンクによってほとんど影響されない2つの小さいパケットは、タイトリンクによって設定された間隔を（基本的に、タイトリンクによって誘発された遅延を符号化して）宛先までずっと保つ。

【0007】

本発明のさらなる特徴および利点は、添付の図を参照して進められる、例示的な諸実施形態の以下のより詳細な説明から明白となろう。

【0008】

添付の特許請求の範囲を、本発明の諸特徴を詳細に記載するが、本発明、および本発明の利点は、添付の図面と併せて解釈される、以下の詳細な説明から最もよく理解されよう。

【発明を実施するための最良の形態】

【0009】

本発明は、添付の図面と併せて読まれるべき、以下の詳細な説明を介して、より完全に理解されよう。この説明では、本発明の様々な実施形態の中で、同様の符号は、同様の要素を指す。ブロードバンドアクセスネットワーク、およびその他のネットワークにおいて

10

20

30

40

50

帯域幅を推定するための方法およびシステムを以下に説明する。本明細書で説明する方法およびシステムは、単に例示的であり、本発明の趣旨および範囲を逸脱することなく、変形形態を実施できることが、当業者には容易に理解されよう。

【0010】

本発明の諸実施形態は、適切なコンピューティング環境において実施されているものとして示される。必須ではないが、本発明は、パーソナルコンピュータによって実行される、手続きなどの、コンピュータ実行可能命令の一般的なコンテキストで説明する。一般に、手続きには、特定のタスクを実行する、または特定の抽象データ型を実装する、プログラムモジュール、ルーチン、関数、プログラム、オブジェクト、コンポーネント、データ構造体などが含まれる。さらに、本発明は、ハンドヘルドデバイス、マルチプロセッサシステム、およびマイクロプロセッサベースの家庭用電化製品、またはプログラマブル家庭用電化製品を含め、様々なコンピュータシステム構成において実施できることが当業者には理解されよう。また、本発明は、通信ネットワークを介してリンクされたリモート処理デバイス群によってタスクが実行される、分散コンピューティング環境において実施することもできる。分散コンピューティング環境では、プログラムモジュール群は、ローカルメモリ記憶装置とリモートメモリ記憶装置の両方の中に配置することができる。コンピュータシステムという用語は、分散コンピューティング環境において見ることができるもの、コンピュータ群のシステムを指すのに使用することができる。

【0011】

図1は、本発明を実施することができる適切なコンピューティングシステム環境100の実施例を示す。コンピューティングシステム環境100は、適切なコンピューティング環境の一実施例に過ぎず、本発明の用法または機能の範囲について何ら限定を示唆するものではない。また、コンピューティング環境100が、例示的な動作環境100に示したコンポーネントのいずれの1つ、または組合せに関連する依存関係または要件も有すると解釈してはならない。本発明の少なくとも1つの実施形態の中で使用するのに適したコンピューティングデバイスは、例示的な動作環境100に示した各コンポーネントを含むが、本発明の別の実施形態は、一部、またはすべての不可欠ではないコンポーネント、例えば、ネットワーク通信に要求されるもの以外の入出力デバイス群を除外するデバイスを利用する。

【0012】

とは言え、本発明を実施するための1つの例示的なシステムは、コンピュータ110の形態で汎用コンピューティングデバイスを含む。コンピュータ110のコンポーネントには、処理装置120、システムメモリ130、ならびにシステムメモリを含む様々なシステムコンポーネントを処理装置120に結合するシステムバス121が含まれるが、それらに限定されない。システムバス121は、様々なバスアーキテクチャのいずれかを使用する、メモリバスまたはメモリコントローラ、周辺バス、およびローカルバスを含め、いくつかのタイプのバス構造のいずれであってもよい。

【0013】

コンピュータ110は、通常、様々なコンピュータ可読媒体を含む。コンピュータ可読媒体は、コンピュータ110がアクセスすることができる任意の利用可能な媒体とすることができる、揮発性媒体と不揮発性媒体、取り外し可能媒体と取り外し不能媒体がともに含まれる。例として、限定としてではなく、コンピュータ可読媒体は、コンピュータ記憶媒体および通信媒体を含みうる。コンピュータ記憶媒体には、コンピュータ可読命令、データ構造、プログラムモジュール、またはその他のデータなどの情報を記憶するための任意の方法または技術で実装された、揮発性媒体と不揮発性媒体、取り外し可能媒体と取り外し不能媒体が含まれる。コンピュータ記憶媒体には、RAM、ROM、EEPROM、フラッシュメモリまたは他のメモリ技術、光ディスクストレージ、磁気カセット、磁気テープ、磁気ディスクストレージまたは他の磁気記憶装置、あるいは所望の情報を格納するのに使用することができ、コンピュータ110がアクセスすることができる他の任意の媒体が含まれるが、以上には限定されない。通信媒体は、通常、搬送波または他のトランスポ

10

20

30

40

50

ート機構などの変調されたデータ信号中に、コンピュータ可読命令、データ構造、プログラムモジュール、またはその他のデータを具現化するものであり、任意の情報伝達媒体が含まれる。「変調されたデータ信号」という用語は、信号内に情報を符号化するような形で特性の1つまたは複数が設定、または変更されている信号を意味する。例として、限定としてではなく、通信媒体には、有線ネットワークまたは直接有線接続などの有線媒体、ならびに音響媒体、RF媒体、赤外線媒体、およびその他の無線媒体などの無線媒体が含まれる。以上の媒体のいずれかの媒体の組合せも、コンピュータ可読媒体の範囲内に含まれる。

【0014】

システムメモリ130は、読み出し専用メモリ(ROM)131やランダムアクセスメモリ(RAM)132などの揮発性メモリおよび/または不揮発性メモリの形態で、コンピュータ記憶媒体を含む。例として、限定としてではなく、図1は、オペレーティングシステム134、アプリケーションプログラム群135、その他のプログラムモジュール群136、およびプログラムデータ137を示す。

10

【0015】

また、コンピュータ110は、他の取り外し可能コンピュータ記憶媒体と取り外し不能コンピュータ記憶媒体、揮発性コンピュータ記憶媒体と不揮発性コンピュータ記憶媒体も含むことが可能である。単に例として、図1は、取り外し不能不揮発性磁気媒体に対して読み取りまたは書き込みを行うハードディスクドライブ141、取り外し可能不揮発性の磁気ディスク152に対して読み取りまたは書き込みを行う磁気ディスクドライブ151、およびCD-ROMなどの取り外し可能不揮発性光ディスク156に対して読み取りまたは書き込みを行う光ディスクドライブ155を示す。例示的な動作環境で使用することができる他のコンピュータ記憶媒体には、磁気テープカセット、フラッシュメモリカード、DVD、デジタルビデオテープ、ソリッドステートRAM、ソリッドステートROMなどが含まれるが、以上には限定されない。ハードディスクドライブ141は、通常、インターフェース140のような取り外し不能メモリインターフェースを介してシステムバス121に接続され、磁気ディスクドライブ151および光ディスクドライブ155は、通常、インターフェース150のような取り外し可能メモリインターフェースでシステムバス121に接続される。

20

【0016】

コンピュータシステムは、さらなる種類の取り外し可能不揮発性の記憶装置のためのインターフェース群を含むことが可能である。例えば、コンピュータは、USBフラッシュドライブ(UFD)154を受け入れることができるUSBポート153、またはセキュアデジタル(SD)メモリカード158を受け入れることができるSDカードスロット157を有することが可能である。USBフラッシュドライブは、様々なコンピューティングデバイス上のUSBポートに挿入することができるUSBコネクタが取り付けられたフラッシュメモリデバイスである。SDメモリカードは、切手サイズのフラッシュメモリデバイスである。USBフラッシュドライブとSDカードはともに、小さいパッケージで大きい記憶容量を提供し、高いデータ転送速度を提供する。

30

【0017】

前述し、図1に示したドライブ群、および関連するコンピュータ記憶媒体により、コンピュータ可読命令、データ構造、プログラムモジュール、およびその他のデータの記憶が、コンピュータ110に提供される。図1では、例えば、ハードディスクドライブ141が、オペレーティングシステム144、アプリケーションプログラム群145、その他のプログラムモジュール群146、およびプログラムデータ147を格納しているのを示す。以上のコンポーネントは、オペレーティングシステム134、アプリケーションプログラム群135、その他のプログラムモジュール群136、およびプログラムデータ137と同一であることも、異なることも可能であることに留意されたい。オペレーティングシステム144、アプリケーションプログラム群145、その他のプログラムモジュール群146、およびプログラムデータ147に、本明細書では、少なくとも、それらが異なる

40

50

コピーであることを示すために異なる符号を与えている。ユーザは、タブレット、または電子デジタイザ 164、マイク 163、キーボード 162、ならびにマウス、トラックボール、またはタッチパッドと一般に呼ばれるポインティングデバイス 161などの入力デバイス群を介して、コマンドおよび情報をコンピュータ 110に入力することができる。以上、およびその他の入力デバイス群は、しばしば、システムバスに結合されたユーザ入力インターフェース 160を介して処理装置 120に接続されるが、パラレルポート、ゲームポート、またはユニバーサルシリアルバス(USB)などの、他のインターフェースおよびバス構造で接続してもよい。また、モニタ 191、または他のタイプのディスプレイデバイスも、ビデオインターフェース 190のようなインターフェースを介してシステムバス 121に接続される。また、モニタ 191は、タッチスクリーンパネルなどと一体化されてもよい。モニタおよび/またはタッチスクリーンパネルは、タブレット型のパソコン用コンピュータの場合のように、コンピューティングデバイス 110が組み込まれた筐体に物理的に結合することができることに留意されたい。加えて、コンピューティングデバイス 110のようなコンピュータは、出力周辺装置インターフェース 195などを介して接続することができる、スピーカ 197やプリンタ 196のような、他の周辺出力デバイス群も含むことが可能である。10

【0018】

コンピュータ 110は、好ましくは、リモートコンピュータ 180のような、1つまたは複数のリモートコンピュータに対する論理接続を使用するネットワーク化された環境において動作するか、または動作するように適応可能である。リモートコンピュータ 180は、パソコン用コンピュータ、サーバ、ルータ、ピアデバイス、またはその他のネットワークノードであることが可能であり、通常、コンピュータ 110に関連して前述した諸要素の一部またはすべてを含むが、メモリ記憶装置 181だけが、図 1に示されている。図 1に示した論理接続には、LAN 171およびWAN 173が含まれるが、その他のネットワークも含まれることが可能である。20

【0019】

LAN 環境で使用される場合、コンピュータ 110は、ネットワークインターフェースまたはネットワークアダプタ 170を介して LAN 171に接続可能である。また、コンピュータ 110は、WAN 173を介して通信を確立するためのモデム 172、または他の手段も含むことが可能である。モデム 172は内蔵または外付けであってよく、ユーザ入力インターフェース 160、または他の適切な機構を介してシステムバス 121に接続することができる。ネットワーク化された環境では、コンピュータ 110に関連して示したプログラムモジュール群、またはこのプログラムモジュール群の諸部分は、リモートメモリ記憶装置の中に格納することができる。例として、限定としてではなく、図 1は、リモートアプリケーションプログラム群 185が、メモリデバイス 181上に存在するのを示す。図示したネットワーク接続は、例示的であり、コンピュータ間で通信リンクを確立する他の手段も使用することができることが理解されよう。30

【0020】

本明細書で使用するネットワーキングの「容量」(C)、または帯域幅とは、ネットワーキングリンクが、持続的にサポートすることができる最大データ転送速度である。容量は、ときとして、リンクの生の帯域幅と厳密でない呼ばれ方をされるが、これは、完全に正確ではない。例えば、トークンバケット(token bucket)などのトラフィックシェーバ(shaper)が、リンクの帯域幅を「管理する」のに使用される場合、生の帯域幅と容量は相異なる。リンクの利用可能な帯域幅(A)は、リンクの未使用の容量、すなわち、リンクの容量 C から、リンクが伝送しているトラフィックの時間平均量(time-averaged volume)を引いた値である。40

【0021】

インターネットバスの「ナローリング」は、最小の容量を有するリンクである。タイトリンクは、利用可能な帯域幅の量が最も少ないリンクである。このため、インターネットバスの容量は、そのバス上のナローリングの容量である。インターネットバスの利用可能50

な帯域幅は、そのパス上のタイトリンクの利用可能な帯域幅である。一般に、パスは、同一の容量、または利用可能な帯域幅を有する複数のナローリンクおよび／またはタイトリンクを有する可能性がある。このため、本明細書の説明は、単一のナローリンク、および単一のタイトリンクを有するネットワークパスに関するが、説明する技術は、複数のナローリンクおよび／またはタイトリンクを有するネットワークパスにも容易に等価に拡張できることが当業者には理解されよう。

【0022】

前述したとおり、ネットワーク通信アプリケーションに関する主要な関心事は、インターネットパスの利用可能な帯域幅を正確に測定できることである。インターネットパス上で利用可能な帯域幅を測定するための現行の方法は、2つの一般化されたタイプ、(1) プローブレートモデル (probe rate model) (PRM)、および(2) プローブギャップモデル (probe gap model) (PGM) である。

【0023】

PRM法は、プローブパケットのトレイン (train) を所与の転送速度 R_i で、送信元コンピューティングデバイスから宛先コンピューティングデバイスに送信する。プローブパケットは、宛先において、ある転送速度 R_o で受信される。プローブが行われているリンク上の利用可能な帯域幅は、 R_o が、その速度までは、ほぼ R_i に等しい速度と一致する。次に、利用可能な帯域幅と合致するこの速度を超えて送信速度 R_i が増加されるにつれ、パケットは、パケットが送信されている速度では、もはや受信されない。

【0024】

他方、PGM法は、一対の等しいサイズのプローブパケットを送信元コンピューティングデバイスから宛先コンピューティングデバイスに送信し、ただし、パケットは、ナローリンク上、およびタイトリンク上のいずれのプローブパケットの予期される伝送時間（別々に計算される）よりも超えない遅延量 d_i だけ、時間的に離間される。パケットは、 d_o という、通常、より大きい間隔で、宛先に現れる。パケット間の間隔は、通常、2つのパケットの立ち下りエッジの間で測定される。 d_i と d_o の差に基づき（すなわち、ネットワークパスによって誘発された遅延に基づき）、関心対象のインターネットパスの利用可能な帯域幅を推定することができる。

【0025】

PGM法は、重大な欠点を有する。これらの方法は、インターネットパス上のナローリンクとタイトリンクが同一である場合にだけ、正しく機能する。本明細書の主要な目的は、この欠点を直す技術を説明することである。以下の説明は、PGM法をより詳細に説明し、次いで、前述した欠点を直す方法を説明する。

【0026】

PGMモデルに対応するパケットパスの図を図2に示す。通常、2つの等しいサイズのプローブパケット201、203が、遅延 d_i によって隔てられて送信される。以下に説明するとおり、PGM法は、第2のプローブパケット203が、タイトリンク205のキューの中に、第1のプローブパケット201が出発する前までに到着することを要する。これが、PGM法が、タイトリンクとナローリンクが同一であることを要する理由である。

【0027】

t_0 を、送信側によって送信された第1のプローブパケット201が、タイトリンク205のキューの中に到着する時点であるものとする。何らかの時点 t_i で、そのパケットは、キューの先頭に到着し、その時点で、そのパケットの传送が始まる。传送は、時刻 t_2 で完了する。次に、第2のパケット203が、時刻 t' にタイトリンク205のキューの中に到着するものと仮定する。 $t_0 < t' < t_2$ である場合、第2のパケット203の传送は、 t_0 と t' の間にタイトリンク205のキューの中に到着するいずれかのパケット（すなわち、クロストラフィック）によって遅延させられ、クロストラフィックの量に相当する、出力遅延 d_o の増加がもたらされる。詳細には、 d_i が、ナローリンク上、およびタイトリンク上のいずれかのプローブパケットの传送時間と等しい ($t' - t_2$ であ

10

20

30

40

50

ること、すなわち、第2のプローブパケットが、第1のプローブパケットが伝送を終える前までに到着することを意味する)場合、 $d_0 - d_i$ は、クロストラフィックを伝送する時間とまったく等しい。したがって、利用可能な帯域幅は、 $A = C \times (1 - (d_0 - d_i) / d_i)$ として推定することができる。

【0028】

しかし、 $t' > t_2$ である(すなわち、第2のプローブパケットが、第1のプローブパケットが出発した後に到着する)場合、PGM法は、破綻する。というのは、その場合、 t_2 と t' の間にキューに到着した可能性があるトラフィック、およびキューを出発した可能性があるトラフィックについての情報が、まったく存在しないからである。このため、PGM法は、第2のプローブパケット203が、第1のプローブパケット201が伝送を終える前までに、タイトリンク205のキューに到着することを要する。条件 $t' < t_2$ は、タイトリンクとナローリンクが同一である場合、保証するのは容易である。ナローリンクの容量C(例えば、パケットペアなどの終端間技術を使用して推定される)の知識を使用して、 d_i が、単に、ナローリンク上のいずれかのプローブパケット210、203の伝送時間を超えないことを保証にするだけでよい。

10

【0029】

しかし、タイトリンクがナローリンクと同一ではない場合、PGM法は、破綻する。ナローリンクがタイトリンクに先立つ場合、プローブパケット間の間隔は、プローブパケットがナローリンクを出ると、拡大される可能性がある。結果のギャップは、 t' を t_2 より大きくするだけ十分に、すなわち、第2のプローブパケットが、第1のプローブパケットが出発した後にタイトリンクに到着するようにするだけ十分に大きい可能性がある。他方、タイトリンクがナローリンクに先立つ場合、タイトリンクにおけるパケット間の間隔は、所望されるとおりである。しかし、ナローリンクが、次に、パケット間の間隔を拡大させて、タイトリンクによって決められた真の間隔が、宛先において保たれない可能性がある。

20

【0030】

以上の問題が、本発明の諸実施形態では、固有のパケットプローブシリーズ構造、およびその構造の処理を使用することによって解決される。2つの仮定を行う。すなわち、(1)パスにおけるタイトリンクのID(identity)を知っていること、および(2)例えば、知られているpathchar技術などの別個のツールを使用して、タイトリンクに至るまでのリンク群の容量を算出することができることがある。概要を述べると、プローブパケット間の間隔が、タイトリンクに至るまでのリンク群によって生じさせられる拡大を考慮して、送信元においてまず調整されて、プローブがタイトリンクに達した際に、間隔が所望されるとおりであるようになる。さらに、本発明の実施形態では、マルチパケットプローブが使用され、マルチパケットプローブは、大きいパケットと、その後に続く2つのはるかに小さいパケットとを含む。大きいパケットは、次に、本発明の一実施形態では、タイトリンクを通過すると、ドロップされる。その後のナローリンクによって影響されるとしてもわずかしか影響されない2つの小さいパケットは、タイトリンクによって設定された間隔を宛先までずっと保つ。本発明の一実施形態では、前もって必要とされる知識は、パスにおけるいずれのリンクがタイトリンクであるか、およびタイトリンクに先立つリンク群の帯域幅だけである。これは、pathcharやパケットテールゲーティング(tailgating)などの、既存の終端間(end-to-end)技術を使用して推定することができる。

30

【0031】

本発明の一実施形態による問題解決法を図3に示す。詳細には、図3は、いくつかのリンクを順次に通過する、本発明の実施形態によるパケットトレインの時間図(temporal diagram)を示す。このトレインは、3つのパケット、すなわち、大きいパイロットパケット301と、その後に続く2つの小さいパケット303、305から成る。パケットトレインが、一連のリンク307、309、311、および313を通過しているのが示されている。本発明の一実施形態では、大きいパケット301は、パスMT

40

50

U（例えば、多くのネットワークバスの場合、1500バイト）と等しいサイズであり、小さいパケット303、305は、最小サイズのIPパケット（40バイト）である。しかし、当業者には理解されるとおり、その他のパケットサイズを使用してもよい。大きいパケット301、および第1の小さいパケット303は、連続して（back-to-back）、または可能な限り接近した間隔で送信される。第2の小さいパケット305は、 d_0 というギャップの後に送信され、ここで、 d_0 は、タイトリンク311までの、タイトリンク311を含むすべてのリンク上における大きいパケット301の伝送時間の合計と等しい。それぞれの小さいパケットの伝送時間は、無視できるほど小さいと仮定することができるが、この仮定は、必須ではない。本発明の一実施形態では、大きいパケット301のTime to Live (TTL) プロパティは、大きいパケット301が、タイトリンク311を通過した後、即時にドロップされるように設定される。10

【0032】

このパケットトレイン301、303、305が、様々なリンク307、309、311、313を通ってネットワークバスを伝送されるにつれ、それぞれの相次ぐリンクにより、2つの小さいパケット303と305の間のギャップが、第1のパケット301の伝送時間と等しい量だけ縮むようにされる。タイトリンク311では、ギャップは、タイトリンク311上の第1のパケット301の伝送時間と等しい。第2のパケット303は、第1のパケット301の直ぐ後に続くので、第3のパケット305は、第2のパケット303が、タイトリンク311のキューを出る前に、そのキューに入る。このため、タイトリンク311の入力における2つの小さいパケット303、305の間の間隔は、利用可能な帯域幅の正確な推定のためにPGM法によってちょうど要求されるとおりの間隔になっている。タイトリンク311の出力における間隔は、クロストラフィックの伝送時間を反映する。大きいパケット301は、タイトリンク311を通過するとドロップされ、小さいパケット303、305だけが、宛先まで伝送されるので、タイトリンクの下流の帯域幅が制約されたリンク群、例えば、リンク313は、間にほとんど影響を与えない。20

【0033】

本発明の実施形態によるプロセスを示す流れ図を図4に示す。プロセス400のステップ401で、送信側デバイスは、関心対象のバスにおけるタイトリンクを識別する。送信側デバイスは、ステップ403で、タイトリンクに至るまでの、関心対象のバスにおけるネットワーキング群によって生じさせられるものと予期される拡大の量、および大きいパケット（例えば、MTUサイズの）がタイトリンクを通過するのにかかると予期される時間の量を計算する。その後、ステップ405で、送信側デバイスは、拡大を考慮して、プローブパケット間に入れる間隔を計算する。次に、送信側デバイスは、ステップ407で、MTU（例えば、1500バイト）などの、大きいサイズの第1のパケットをネットワーク上に伝送し、関心対象のバスを介して、受信側デバイスに送る。パケットのTTLは、ステップ403で計算された、パケットがタイトリンクを通過するのにかかると予期される時間の量に設定される。30

【0034】

ステップ409で、送信側デバイスが、小さいパケット（例えば、40バイトという最小サイズのIPパケット）をネットワーク上に伝送し、関心対象のバスを介して、受信側デバイスに送る。次に、パケットトレインを完了させるために、送信側デバイスは、ステップ405で先に計算された遅延時間だけ待ち、次いで、ステップ411で、場合により、ただし、必須ではなく第1の小さいパケットと同一サイズの、別の小さいパケットをネットワーク上に伝送し、関心対象のバスを介して、受信側デバイスに送る。40

【0035】

以上に概要を述べたステップの後、パケットトレインは、関心対象のバス、およびそのバスのリンク群を通過する。小さいパケット間の初期間隔は、予期されるとおり、ステップ413において、拡大を通じて変更されて、拡大の効果が、初期間隔におけるオフセットによって消されるようになる。トレイン、または少なくとも大きいパケットが、タイトリンクを通過すると、大きいパケットのTTLは、有効期限が切れて、大きいパケットは50

、ステップ415でドロップされる。

【0036】

ステップ417で、受信側デバイスにおいて、受信パケット（現時点では、小さいパケットだけ）が受信され、タイムスタンプが付けられて、パケット間の相対遅延を記録することができるようになる。ステップ419で、受信側デバイスが、タイムスタンプ情報を示す1つまたは複数の概要メッセージを送信側デバイスに転送する。最後に、送信側デバイスが、ステップ421で、パケット間の相対遅延 d_{final} を計算し、計算された遅延 d_{final} から、タイトリンクへの入力における既知のパケット間隔を引く。その差は、タイトリンクによって導入された遅延の量を正確に表し、そのため、関心対象のリンク上の利用可能な帯域幅が、ステップ423で、送信側デバイスによって標準的な方法で計算されることが可能になる。10

【0037】

前述した3つのパケットの単一のトレインの使用は、特定の時点に関する利用可能な帯域幅の推定値をもたらすことができることに留意されたい。しかし、利用可能な帯域幅は、上昇する可能性も、低下する可能性もあり、本発明の一実施形態では、特定の値ではなく、平均値が使用される。詳細には、図4で以上に示したプロセスは、特定の環境の安定性または不安定性を考慮して、適宜、複数回、繰り返され、利用可能な帯域幅値は、それぞれのそのような測定に関して計算される。計算された利用可能な帯域幅値は、次に、様々な形で一緒に平均されて、平均の利用可能な帯域幅値を形成する。

【0038】

ネットワーク帯域幅推定のための新たな有用なシステムが説明されたことが理解されよう。本発明を実施するための、本発明者らが知っている最良の形態を含む、本発明の好ましい諸実施形態を、本明細書で説明した。以上の説明を読むことで、それらの好ましい諸実施形態の変形形態が、当業者には明白となりうる。本発明者らは、当業者が、そのような変形形態を適宜、使用することを予期しており、本発明者らは、本発明が、本明細書で具体的に説明したのとは異なる形で実施されることを企図している。したがって、本発明は、適用される法律によって許されるとおり、本明細書に添付した特許請求の範囲に記載する本件のすべての変形形態および均等形態を含む。さらに、前述した諸要素の、可能なすべての変種におけるいずれの組合せも、本明細書で特に明記しない限り、または文脈と明らかに矛盾しない限り、本発明によって包含される。20

【0039】

本発明を説明する文脈における「ある(a)」、「ある(an)」、および「その、前記(the)」という語、ならびに類似の指示語の使用は（特に請求項の文脈において）、本明細書で特に明記しない限り、または文脈と明らかに矛盾しない限り、単数形と複数形をともに範囲に含むと解釈されるものとする。「含む、備える(comprising)」、「有する(having)」、「含む(including)」、および「含む(containing)」という語は、特に明記しない限り、制限のない語（すなわち、「含むが、限定されない」を意味する）と解釈されるものとする。本明細書における値の範囲のいずれの列挙も、本明細書で特に明記しない限り、単に、その範囲内に含まれるそれぞれの別個の値を個々に指す略記法の役割をすることを意図しており、それぞれの別個の値が、あたかも本明細書で個々に列挙されているかのように、本明細書に組み込まれている。本明細書で説明したすべての方法は、本明細書で特に明記しない限り、または文脈と明らかに矛盾しない限り、任意の適切な順序で実行することができる。本明細書で提供するあらゆる実施例、または例示的な言い回し（例えば、「などの」）の使用は、単に、本発明をよりよく説明することを目的とし、特に主張しない限り、本発明の範囲に限定を課すものではない。本明細書におけるいずれの言い回しも、特許請求していない要素が、本発明の実施に不可欠であることを示すものと解釈してはならない。40

【図面の簡単な説明】

【0040】

【図1】本発明の実施形態による通信構成プロセスを実施する無線ネットワークの一部と50

して使用することができるコンピューティングデバイスの例示的なアーキテクチャを示す概略図である。

【図2】プロープギャップモデルに対応するパケットパスを示す図である。

【図3】本発明の実施形態による、利用可能な帯域幅を推定するのにパケットトレインが使用される形を示す理想化されたパケットタイミングプロットである。

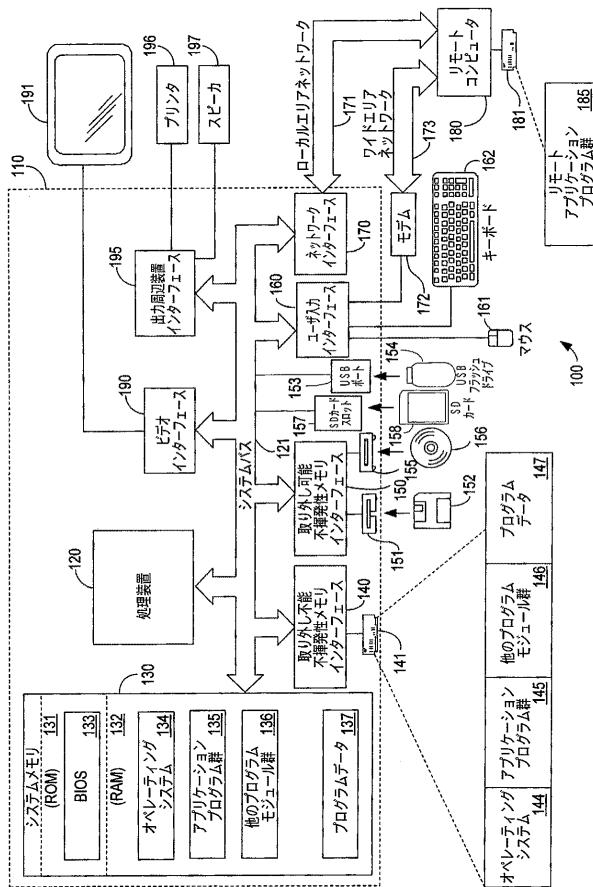
【図4】本発明の実施形態による、タイトリンクを有するネットワークバス上の利用可能な帯域幅を推定するプロセスを示す流れ図である。

【符号の説明】

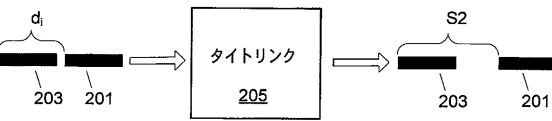
【0041】

1 2 0	処理装置	10
1 2 1	システムバス	
1 3 0	システムメモリ	
1 3 4	オペレーティングシステム	
1 4 4	オペレーティングシステム	
1 3 5	アプリケーションプログラム群	
1 4 5	アプリケーションプログラム群	
1 3 6	他のプログラムモジュール群	
1 4 6	他のプログラムモジュール群	
1 3 7	プログラムデータ	
1 4 7	プログラムデータ	20
1 4 0	取り外し不能不揮発性メモリインターフェース	
1 5 0	取り外し可能不揮発性メモリインターフェース	
1 5 3	U S B ポート	
1 5 4	U S B フラッシュドライブ	
1 5 7	S D カードスロット	
1 5 8	S D カード	
1 6 0	ユーザ入力インターフェース	
1 6 1	マウス	
1 6 2	キーボード	
1 7 0	ネットワークインターフェース	30
1 7 1	ローカルエリアネットワーク	
1 7 2	モデム	
1 7 3	ワイドエリアネットワーク	
1 8 0	リモートコンピュータ	
1 8 5	リモートアプリケーションプログラム群	
1 9 0	ビデオインターフェース	
1 9 5	出力周辺装置インターフェース	
1 9 6	プリンタ	
1 9 7	スピーカ	

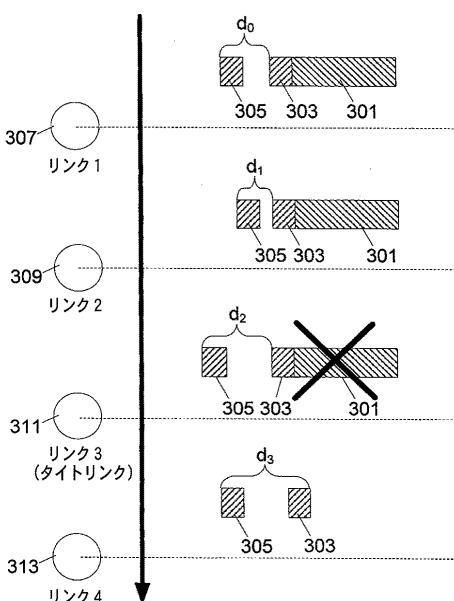
【図1】



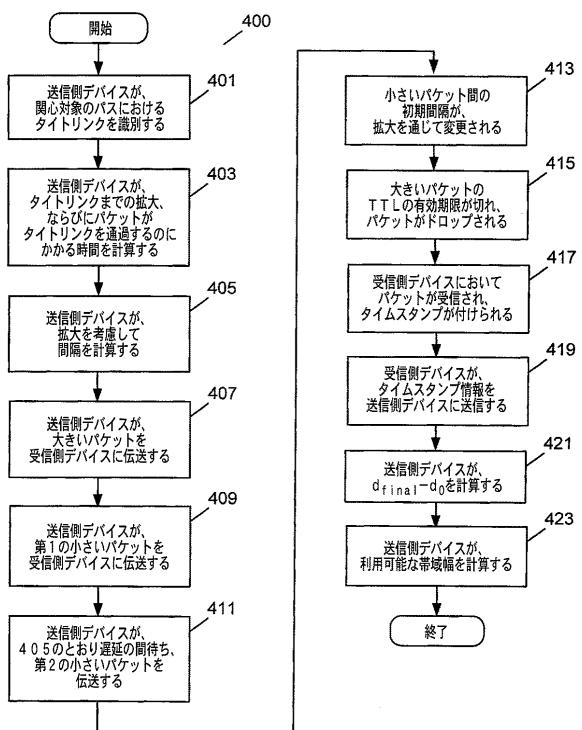
【図2】



【図3】



【図4】



フロントページの続き

(72)発明者 ヴェンカタ エヌ . パドマナバン
アメリカ合衆国 98052 ワシントン州 レッドモンド ワン マイクロソフト ウェイ マ
イクロソフト コーポレーション内

審査官 斎藤 浩兵

(56)参考文献 特開2000-307651(JP,A)
特開2003-249960(JP,A)
特開2000-224172(JP,A)
国際公開第02/025878(WO,A1)
米国特許出願公開第2003/0117959(US,A1)
特表平06-508008(JP,A)
特開2005-039724(JP,A)

(58)調査した分野(Int.Cl., DB名)

H04L 12/56