

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6777650号  
(P6777650)

(45) 発行日 令和2年10月28日 (2020. 10. 28)

(24) 登録日 令和2年10月12日 (2020. 10. 12)

(51) Int. Cl.	F I
HO 4 L 12/891 (2013. 01)	HO 4 L 12/891
HO 4 L 29/00 (2006. 01)	HO 4 L 13/00 S

請求項の数 18 (全 15 頁)

(21) 出願番号	特願2017-553093 (P2017-553093)	(73) 特許権者	597149146
(86) (22) 出願日	平成28年4月8日 (2016. 4. 8)		ドイッチェ テレコム アーゲー
(65) 公表番号	特表2018-511275 (P2018-511275A)		ドイツ国. 5 3 1 1 3 ボン, フリードリ
(43) 公表日	平成30年4月19日 (2018. 4. 19)		ヒ - エベルト - アレエ 1 4 0
(86) 国際出願番号	PCT/EP2016/057795	(74) 代理人	100094112
(87) 国際公開番号	W02016/162501		弁理士 岡部 譲
(87) 国際公開日	平成28年10月13日 (2016. 10. 13)	(74) 代理人	100106183
審査請求日	平成31年3月25日 (2019. 3. 25)		弁理士 吉澤 弘司
(31) 優先権主張番号	15163169. 4	(74) 代理人	100114915
(32) 優先日	平成27年4月10日 (2015. 4. 10)		弁理士 三村 治彦
(33) 優先権主張国・地域又は機関	欧州特許庁 (EP)	(74) 代理人	100120363
			弁理士 久保田 智樹
		(74) 代理人	100125139
			弁理士 岡部 洋

最終頁に続く

(54) 【発明の名称】 TCPトンネル及びネイティブTCP情報に基づくバンドリングシナリオにおけるパケットのスケジューリングのための方法及びシステム

(57) 【特許請求の範囲】

【請求項 1】

伝送制御プロトコルTCPセッション又はユーザデータグラムプロトコルUDPセッションに対するパケットベースのスケジューリングのためのシステムであって、第1のスケジューリングモジュールと、少なくとも1つのトランスポートネットワークに接続された少なくとも2つのリンクとを備えた第1の結合モジュールであって、前記第1の結合モジュールはユーザデバイスに接続可能であり、前記第1の結合モジュールを前記少なくとも1つのトランスポートネットワークに接続する各リンクに対して、各TCPトンネルは前記少なくとも1つのトランスポートネットワークと第2の結合モジュールとの間で構成され、前記第1のスケジューリングモジュールは前記第2の結合モジュールに向けて前記TCPトンネルを介して、好ましくはインターネット上で、データパケットをスケジューリング及び分配するように構成された、第1の結合モジュールと、

第2のスケジューリングモジュールを備える前記第2の結合モジュールであって、

前記第2の結合モジュールは、前記TCPトンネルを介して前記少なくとも一つのトランスポートネットワークに接続され、前記第2のスケジューリングモジュールは前記第1の結合モジュールに向けて前記TCPトンネルを介して前記データパケットをスケジューリング及び分配するように構成された、前記第2の結合モジュールとを備え、

前記データパケットの再順序付けは前記ユーザデバイス又はインターネットによって実行され、

前記第2の結合モジュールは、受信したパケットをデカプセル化して、前記再順序付け

10

20

することなく前記ユーザデバイス又はインターネットへ転送し、

前記ユーザデバイス又はインターネットは、カプセル化する前の元のデータパケットの TCP ヘッダー情報に基づいて前記再順序付けする、システム。

【請求項 2】

前記第 1 の結合モジュールに有線又は無線接続を介して接続された第 1 のルータをさらに備え、好ましくは前記第 1 のルータが、データを前記ユーザデバイスから受信 / 前記ユーザデバイスに送信するように構成された、請求項 1 に記載のシステム。

【請求項 3】

前記第 2 の結合モジュールに有線又は無線接続を介して接続された第 2 のルータをさらに備え、好ましくは前記第 2 のルータが、データをサーバから受信 / サーバに送信するように構成されたネットワーク・アドレス・トランスレーションモジュールを備える、請求項 2 に記載のシステム。

10

【請求項 4】

前記第 1 及び / 又は第 2 のスケジューリングモジュールが、前記 TCP トンネルからのネイティブ情報に基づいて前記第 1 のルータと前記第 2 のルータの間の前記 TCP トンネルの容量を取得するように構成された、請求項 3 に記載のシステム。

【請求項 5】

前記 TCP トンネルの前記ネイティブ情報が、輻輳ウィンドウ  $C_{cwnd}$  のサイズ、TCP スロースタート閾値  $S_{thresh}$ 、平滑化往復時間  $T_{RTT}$ 、送出されたパケット数  $P_{out}$ 、肯定応答されたパケット数  $P_{sacked}$ 、再送信されたパケット数  $P_{retrans}$ 、及び喪失したとみなされるパケット数  $P_{lost}$  のうちの少なくとも 1 つを備える、請求項 4 に記載のシステム。

20

【請求項 6】

前記第 1 及び / 又は第 2 のスケジューリングモジュールが、前記第 1 のルータから前記第 2 のルータに現在伝搬しているパケット数を、

$$P_{fly} = P_{out} - P_{sacked} + P_{retrans} - P_{lost}$$

として取得するように構成された、請求項 5 に記載のシステム。

【請求項 7】

前記第 1 及び / 又は第 2 のスケジューリングモジュールが、残容量を、

$$C_{left} = C_{cwnd} - P_{fly}$$

として取得するように構成された、請求項 6 に記載のシステム。

30

【請求項 8】

前記第 1 及び / 又は第 2 のスケジューリングモジュールが、規定間隔内で前記残容量の変化  $C_{left}$  を取得するように構成された、請求項 7 に記載のシステム。

【請求項 9】

前記第 1 及び / 又は第 2 のスケジューリングモジュールが、スケジューリングアルゴリズムを用いて前記第 1 のルータと前記第 2 のルータの間の前記 TCP トンネルに対するスケジューリング重み付けを取得するように構成された、請求項 3 から 8 のいずれか一項に記載のシステム。

【請求項 10】

前記スケジューリング重み付けが、前記 TCP トンネルの各々を介して送信されるパケットの量の割合及び / 又は前記 TCP トンネルの各々を介して連続して送信されるパケットの量で構成された、請求項 9 に記載のシステム。

40

【請求項 11】

前記第 1 及び / 又は第 2 のスケジューリングモジュールが、前記スケジューリング重み付けを連続的に又はタイムスロットにおいて適合するように構成され、好ましくは前記第 1 及び / 又は第 2 のスケジューリングモジュールが、前タイムスロット中に収集されたパラメータに基づいて後続のタイムスロットの前記スケジューリング重み付けを適合するように構成された、請求項 10 に記載のシステム。

【請求項 12】

50

伝送制御プロトコルセッション又はユーザデータグラムプロトコルUDPセッションについてのパケットベースのスケジューリングのための方法であって、

(a) 第1のスケジューリングモジュールと、少なくとも1つのトランスポートネットワークに接続された少なくとも2つのリンクとを備えた第1の結合モジュールを接続するステップであって、前記第1の結合モジュールはユーザデバイスに接続可能である、ステップ、

(b) 前記第1の結合モジュールを前記少なくとも1つのトランスポートネットワークに接続する各リンクに対して、前記少なくとも1つのトランスポートネットワークと第2の結合モジュールとの間で、各TCPトンネルを構成するステップ、

(c) 前記第1の結合モジュールから前記第2の結合モジュールに向けて前記TCPトンネル及び前記少なくとも1つのトランスポートネットワークを介してデータパケットをスケジューリング及び分配するステップ、

(d) 第2のスケジューリングモジュールを備える前記第2の結合モジュールを前記TCPトンネルを介して前記少なくとも1つのトランスポートネットワークに接続するステップであって、前記第2の結合モジュールがサーバに接続可能である、ステップ、及び

(e) 前記第2の結合モジュールから前記第1の結合モジュールに向かう前記TCPトンネルを介して前記データパケットをスケジューリング及び分配するステップを備え、

前記データパケットの再順序付けは前記ユーザデバイス又はインターネットによって実行され、

前記第2の結合モジュールは、受信したパケットをデカプセル化して、前記再順序付けすることなく前記ユーザデバイス又はインターネットへ転送し、

前記ユーザデバイス又はインターネットは、カプセル化する前の元の前記データパケットのTCPヘッダー情報に基づいて前記再順序付けする、方法。

#### 【請求項13】

第1のルータを前記第1の結合モジュールに有線又は無線接続を介して接続するステップを備え、好ましくは前記第1のルータが、データを前記ユーザデバイスから受信/前記ユーザデバイスに送信するように構成された、請求項12に記載の方法。

#### 【請求項14】

第2のルータを前記第2の結合モジュールに有線又は無線接続を介して接続するステップを備え、好ましくは第2のルータが、データを前記サーバから受信/前記サーバに送信するように構成されたネットワーク・アドレス・トランスレーションモジュールを備える、請求項13に記載の方法。

#### 【請求項15】

前記TCPトンネルからのネイティブ情報に基づいて前記第1のルータと前記第2のルータの間の前記少なくとも2つのTCPトンネルの各々の容量を取得するステップを備える、請求項14に記載の方法。

#### 【請求項16】

データパケットをスケジューリング及び分配するステップが、前記第1のルータと前記第2のルータの間の前記TCPトンネルに対するスケジューリング重み付けを取得するようにスケジューリングアルゴリズムに基づく、請求項14又は15に記載の方法。

#### 【請求項17】

前記スケジューリング重み付けが、前記TCPトンネルの各々を介して送信されるパケットの割合及び/又は前記TCPトンネルの各々を介して連続して送信されるパケットの量で構成される、請求項16に記載の方法。

#### 【請求項18】

スケジューリング重み付けが連続的に又はタイムスロットにおいて適合され、好ましくは、後続のタイムスロットのスケジューリング重み付けが、前タイムスロット中に収集されたパラメータに基づく、請求項17に記載の方法。

#### 【発明の詳細な説明】

10

20

30

40

50

## 【技術分野】

## 【0001】

本発明は、伝送制御プロトコル(TCP)セッション又はユーザデータグラムプロトコル(UDP)セッションについてのパケットベースのスケジューリングを用いるリソースバンドリングのための方法及びシステムに関する。

## 【背景技術】

## 【0002】

過去において、セッション/アプリケーションレイヤ、トランスポートレイヤ、ネットワークレイヤ及びリンクレイヤ手法に大別され得るリソースバンドリングのための手法に多大な努力が払われてきた。以下において、これらのグループの各々についての最も重要な例を簡単に説明する。

10

## 【0003】

HTTPレンジリクエストは、リソースのバンドリングを行うのにも使用され得るセッション/アプリケーションレイヤ構成である。この構成は、ダウンロードするコンテンツを、異なるインターフェースを介する個別のTCPセッションにおいてその後にダウンロードされるスライスに分割する。この構成は、サムスンギャラクシーS5内に実装され、「ネットワークブースタ」といわれる(S.E. Corporation「Network booster-for enhanced data performance」、Techpaper)。

## 【0004】

20

トランスポートレイヤによるソリューションの最も顕著な例は、リクエスト・フォー・コメント(RFC)6824においてインターネット・エンジニアリング・タスク・フォース(IETF)によって規定されるマルチパスTCP(MPTCP)である(A. Ford、C. Raiciu、M. Handley及びO. Bonaventure、「Tcp extensions for multipath operation with multiple addresses」、RFC、No. 6824、2013年1月)。MPTCPは、ピア間の複数の経路の同時使用がアプリケーションに対してトランスペアレントとなることを可能とするTCPの拡張としてみることができる。MPTCPは、iPhone(登録商標)のような市販製品において既に使用されている。Siriusサービスは、例えば、MPTCPを用いて信頼性を高める。iPhone(登録商標)がセルラ及びWi-Fi接続を有する場合、Siriusは2つのセッションを、1つをWi-Fiを介して、1つをセルラネットワークを介して開始する。接続の一方が利用不可となり、又は信頼性がなくなると、バックアップ接続が直ちに利用可能となる。

30

## 【0005】

他のトランスポートレイヤの例は、ストリーム制御伝送プロトコル(SCTP)(R. Stewart他、「Stream Control Transmission Protocol」、RFC2960(提案された標準)、インターネット・エンジニアリング・タスク・フォース、2000年10月、RFC4960によって使用されなくなり、RFC3309によって更新された)。SCTPの効果の1つは、リンクのバンドリングを原則として可能ともし、そのマルチホーミング構成である。

40

## 【0006】

原則として、セッション/アプリケーションレイヤ手法及びトランスポートレイヤ手法は、ネットワークの何らかの特殊な関与なしにエンドデバイスによって実行されるバンドリングとしてエンドトゥエンド(E2E)シナリオに適するということができる。E2E手法の欠点は、それがユーザデバイス及びインターネットサービスによってサポートされなければならないので、その技術を導入する労力である。

## 【0007】

ネットワークレイヤ手法は、それらがエンドデバイスを変更する必要なく、容易に使用され得るので、ネットワークオペレータに対する好適なソリューションである。その代わり、オペレータの制御下にあり、エンドデバイスに対してトランスペアレントなバンドリ

50

ングを実行するプロキシサーバが使用される。

【0008】

リンクレイヤバンドリングとは、同等技術の複数チャネルのバンドリングのことをいう。例として、IEEE 802.3ad (WO 2009/019258 A1) において規定されるイーサネットリンクのバンドリング、及び例えば Atheros Super G モードにおいて実施される Wi-Fi チャネルのバンドリングがある。(E. G. Llaïro 及び D. Giustinianno, 「Smartap: Practical wlan backhaul aggregation」、Wireless Days, IEEE, 2013, pp. 1-7) において提案される SmartAP ソリューションは、単一无線マルチチャネル可視化を用いてワイヤレス・ローカル・エリア・ネットワーク (WLAN) バックホールの集約に対処する。これによって、異なるチャネルが使用される場合であっても、複数の隣接アクセスポイント (AP) との接続が可能となる。D. Giustinianno, E. Goma, A. Lopez Toledo, I. Dangerfield, J. Morillo 及び P. Rodriguez, 「Fair wlan backhaul aggregation」、Proceedings of the Sixteenth Annual International Conference on Mobile Computing and Networking, ser. MobiCom '10, New York, NY, USA: ACM, 2010, pp. 269-280 は、記載のシステムにおける正当性を調査している。

【発明の概要】

【0009】

本発明の課題は、向上した性能を有する伝送制御プロトコル (TCP) セッション又はユーザデータグラムプロトコル (UDP) セッションのためのパケットベースのスケジューリングのための方法及びシステムを提供することである。この課題は、独立請求項の構成で達成される。従属請求項は、発明の更なる態様に関する。

【0010】

本発明は、ネットワークオペレータに好ましく適する包括的なバンドリングアーキテクチャを提案する。一方、本発明は、オーバーザトップ (OTT) コンテンツシナリオにも適する。他の手法に対して、提示される手法は TCP トンネルを用いてシステムの複雑さを軽減し、高いバンドリングゲインを与える。旧来的なバンドリングソリューションにおけるような経路モニタは不要である。その代わりに、TCP トンネルから到来するネイティブ情報が、最も効率的なパケットスケジューリングについて決定するために、パケットスケジューラに対する入力として使用される。すなわち、バンドリングゲインを最大化するためにスケジューリング手法に対する重み付けを決定することができる革新的な解決手段が記載される。言い換えると、TCP トンネル及び特に TCP トンネルからのネイティブ情報が、各トンネルを介して利用可能な帯域幅を推定するのに使用され得る。この情報は、利用可能なトンネルに対するデータパケットの効率的なマッピングのために使用される。本発明によると、異なるトンネルが異なるネットワークインターフェース、例えば、Wi-Fi 及び 3GPP を通過してもよく、各トンネルを介した利用可能な帯域幅は大きく異なり得る。本発明は、トンネルの各々を介して利用可能な帯域幅を考慮することによって効率的な帯域幅集約を与える。

【0011】

本発明の第1の態様によると、伝送制御プロトコル TCP セッション又はユーザデータグラムプロトコル UDP セッションに対するパケットベースのスケジューリングのためのシステムが提供される。システムは、第1のスケジューリングモジュールと、少なくとも1つのトランスポートネットワークに接続された少なくとも2つのアクセスインターフェースとを備える第1の結合モジュールを備え、第1の結合モジュールはユーザデバイスに接続可能であり、TCP トンネルは第2の結合モジュールを終端するアクセスインターフェースの各々を介して構成され、第1のスケジューリングモジュールは第2の結合モジュールに向けて TCP トンネルを介して、好ましくはインターネット上で、データパケット

をスケジューリング及び分配するように構成される。システムは、単独で又は組合せにおいて、第2のスケジューリングモジュールと、少なくとも1つのネットワークの各々に接続された少なくとも1つのアクセスインターフェースとを備える第2の結合モジュールをさらに備えていてもよく、第2の結合モジュールはサーバに接続され、第2のスケジューリングモジュールは第1の結合モジュールに向けてTCPトンネルを介してデータパケットをスケジューリング及び分配するように構成される。

【0012】

データパケットがネットワークレイヤ（レイヤ3）上で分配されることが好ましい。これによって、標準TCPプロトコルを利用することが可能となる。したがって、新たなTCPプロトコルは不要となり得る。

10

【0013】

システムは、第1の結合モジュールに有線又は無線接続を介して接続される第1のルータをさらに備えていてもよく、好ましくは、第1のルータはデータをユーザデバイスから受信し、及びユーザデバイスに送信するように構成される。

【0014】

システムはまた、第2の結合モジュールに有線又は無線接続を介して接続された第2のルータを備え、好ましくは、第2のルータはデータをサーバから／サーバに受信及び送信するように構成されたネットワーク・アドレス・トランスレーションモジュールを備える。

【0015】

20

第1のルータはマルチホーミング化されていてもよく、第1のルータが少なくとも2つのインターフェースを介してインターネットに接続されることを意味する。両ルータ（第1及び第2のルータ）は、異なるネットワークに属していてもよく、TCPトンネルを介して相互接続される。両ルータ間のTCPトンネル数は、第1のルータがインターネットに接続されるインターフェース数に基づく。これらのインターフェースの各々を介して、1つのトンネルが使用される。

【0016】

用語「接続される」は、あるモジュールが、それが接続されるそれぞれのモジュールの一部であること、すなわち、モジュールが内部接続され得ることも意味し得ることが認識される。例えば、第1及び／又は第2の結合モジュールは、それぞれユーザデバイス及び／又はサーバの集積部分となり得る。他の例として、第1及び／又は第2の結合モジュールは、それぞれ第1及び／又は第2のルータの集積部分となり得る。

30

【0017】

好ましくは、第1及び／又は第2のスケジューリングモジュールは、TCPトンネルからのネイティブ情報に基づいて第1のルータと第2のルータの間のTCPトンネルの容量を取得するように構成される。この情報は、TCPトンネルの各々に対するスケジューリング重み付けについて決定するのに使用され得る。

【0018】

またさらに、TCPトンネルのネイティブ情報は、以下の：輻輳ウインドウCwndのサイズ、TCPスロースタート閾値Ssthresh、平滑化往復時間RTT、送出されたパケット数Pout、肯定応答されたパケット数Packed、再送信されたパケット数Retrans、及び喪失したものとみなされるパケット数Plost、のうちの少なくとも1つを備え得る。代替的又は追加的に、TCPトンネルインターフェースの送信者キューサイズが、更なるパラメータとして考慮され得る。

40

【0019】

本発明の更なる態様によると、第1及び／又は第2のスケジューリングモジュールは、 $P_{fly} = P_{out} - P_{packed} + P_{retrans} - P_{lost}$ として第1のルータから第2のルータに現在伝搬しているパケット数を各トンネルについて取得するように構成される。

【0020】

50

より好ましくは、第 1 及び / 又は第 2 のスケジューリングモジュールは、 $C_{left} = C_{cwnd} - P_{fly}$  として残容量を取得するように構成される。

【0021】

より一層好ましくは、第 1 及び / 又は第 2 のスケジューリングモジュールは、規定された間隔内で残容量  $C_{left}$  の変化を取得するように構成される。

【0022】

一方、当業者に理解されるはずの、本発明と同じ効果を実現するために他のパラメータが取得及び使用され得る点で、パラメータの上記列挙は網羅的ではないことが理解されるべきである。

【0023】

本発明の更なる態様によると、第 1 及び / 又は第 2 のスケジューリングモジュールは、スケジューリングアルゴリズムを用いて第 1 のルータと第 2 のルータの間の TCP トンネルに対するスケジューリング重み付けを計算するように構成される。

【0024】

好ましくは、スケジューリング重み付けは、トンネルの各々を介して送信されるパケットの割合及び / 又はトンネルの各々を介して連続して送信されるパケットの量で構成される。

【0025】

さらに、第 1 及び / 又は第 2 のスケジューリングモジュールは、スケジューリング重み付けを連続的に及び / 又はタイムスロットに適合させるように構成され得る。

【0026】

好ましくは、第 1 及び / 又は第 2 のスケジューリングモジュールは、後続のタイムスロットのスケジューリング重み付けを前タイムスロット中に収集されたパラメータに基づいて適合させるように構成され得る。

【0027】

本発明の第 2 の態様によると、好ましくは上述したシステムを用いて伝送制御プロトコルセッションについてのパケットベースのスケジューリングのための方法が提供される。方法は、(a) 第 1 のスケジューリングモジュール及び少なくとも 1 つのネットワーク接続された少なくとも 2 つのアクセスインターフェースを備える第 1 の結合モジュールを接続するステップであって、第 1 の結合モジュールはユーザデバイスに接続され得る、ステップ、(b) 第 2 の結合モジュールで終端するアクセスインターフェースの各々を介して TCP トンネルを構成するステップ、(c) 第 1 の結合モジュールから第 2 の結合モジュールに向けて TCP トンネル及び少なくとも 1 つのネットワークを介してデータパケットをスケジューリング及び分配するステップ、及び / 又は (d) 第 2 のスケジューリングモジュール及び少なくとも 1 つのアクセスインターフェースを備えた第 2 の結合モジュールを少なくとも 1 つのネットワークの各々に接続するステップであって、第 2 の結合モジュールがサーバに接続され得る、ステップ、並びに (e) 第 2 の結合モジュールから第 1 の結合モジュールに向けて TCP トンネルを介してデータパケットをスケジューリング及び分配するステップを備える。

【0028】

方法は、第 1 のルータを第 1 の結合モジュールに有線又は無線接続を介して接続するステップをさらに備えていてもよく、好ましくは、第 1 のルータはデータをユーザデバイスから / ユーザデバイスに受信及び送信するように構成される。

【0029】

方法はまた、第 2 のルータを第 2 の結合モジュールに有線又は無線接続を介して接続するステップを備えていてもよく、好ましくは、第 2 のルータは、データをサーバから / サーバに受信及び送信するように構成されたネットワーク・アクセス・トランスレーションモジュールを備える。

【0030】

ここでも、接続するステップは、2 つの別個のデバイス / モジュールを接続することを

10

20

30

40

50

意味し、又は第1のデバイス/モジュールを第2のデバイス/モジュールに内部接続すること、すなわち、第2のデバイス/モジュールが第1のデバイス/モジュールに集積され得ることを意味し得る。

【0031】

好ましくは、方法は、TCPトンネルからのネイティブ情報に基づいて第1のルータと第2のルータの間の少なくとも2つのTCPトンネルの各々の容量を取得するステップを備える。

【0032】

さらに、データパケットをスケジューリング及び分配するステップは、第1のルータと第2のルータの間のTCPトンネルに対するスケジューリング重み付けを取得するスケジューリングアルゴリズムに基づいていてもよい。

10

【0033】

スケジューリング重み付けは、トンネルの各々を介して送信されるパケットの割合及び/又はトンネルの各々を介して連続して送信されるパケットの量で構成され得る。

【0034】

好ましくは、スケジューリング重み付けは、連続的に及び/又はタイムスロットにおいて適合される。

【0035】

より一層好ましくは、後続のタイムスロットのスケジューリング重み付けは、前タイムスロット中に収集されたパラメータに基づいていてもよい。

20

【0036】

本解決手段のすべてに共通するのは、利用可能なリンクを介したパケットのスケジューリングの課題である。異なるアクセスネットワークのリソースをバンドリングすることは、スループット及びネットワーク故障に対する弾力性を向上する対策となる。一方、アクセスネットワークリソースのバンドリングは、特にTCPトラフィックの場合に難しい課題となる。特に、アクセスリソースのバンドリングを行う場合の課題は、以下の通りである。

【0037】

・第1の課題は、今日のネットワークで使用される最も多いトランスポートプロトコルは伝送制御プロトコル(TCP)及びユーザデータグラムプロトコル(UDP)であること並びにこれらのプロトコルが複数のリンクを介して動作するようには設計されていないことから来ている。これらのプロトコルは、同じセッションに属するパケットが1つのインターネットプロトコル(IP)アドレスからしか受信されないことを想定している。しかし、複数のアクセスリンクが使用される場合、各インターフェースは異なるIPアドレスを得ることになる。この問題を克服するために、端末によって使用されるすべてのアクセスリンクの代わりにネットワーク・アドレス・トランスレーション(NAT)を実行する追加のボックスが導入されてきた。

30

【0038】

・第2の課題は、利用可能なリンクを介したパケットのスケジューリングに関する。まず、それは、2つの基本タイプのスケジューリング手法：セッションベースとパケットベースの間で区別され得る。セッションベースのスケジューリングが実行される場合、特定セッションのパケットは常に同じリンクを介して送信される一方で、セッションは利用可能なリンクに分配されることになる。例えば、セッションAのパケットがリンク $L_1$ で送信される一方で、セッションBのパケットがリンク $L_2$ を介して送信される。セッションベースの手法には、幾つかの欠点がある。例えば、単一のセッションは複数の利用可能なリンクの利益を受けることができない。同じ端末への/同じ端末からの複数のセッションの場合でさえも、異なるリンクの性能が大幅に異なる場合にはバンドリングゲインが保証可能とならない。この発明の焦点は、単一のセッションが利用可能なリンクを介して分配されるパケットベースのスケジューリング手法にある。異なるリンクは、利用可能な帯域幅(B)、遅延(D)、遅延ジッタ(J)及びパケット損失(L)の観点で異なる特性を

40

50



有し得るので、利用可能なリンクを介してパケットが均等に分配されるラウンドロビンのスケジューリングを実行するのでは充分でない。むしろ、重み付けスケジューリングが好ましい。特にTCPセッションについては、これらの重み付けはエンドトゥエンド性能に大きな影響を有し、したがってこれらの重み付けは好ましくは慎重に決定される。

【0039】

・最後に、再順序付けパケットの課題は、パケットベースのスケジューリングが適用される場合に作用し始める。異なるリンクの異なる特性に起因して、及びパケットがインターネットを介して完全に異なるルートを介して伝搬し得ることに起因して、パケットは、誤った順序で宛先に到着し得る。アプリケーションに応じて、これはサービス品質の低下をもたらす得る。

【0040】

いくつかの好ましい実施形態を、図面を参照してここに説明する。説明の目的のため、請求項に記載される本発明の範囲から逸脱することなく、種々の具体的詳細を説明する。

【図面の簡単な説明】

【0041】

【図1】図1は、従来技術によるマルチホーミング化された端末の例を示す。

【図2(a)】本発明の4つの好適な実施形態を示す。

【図2(b)】本発明の4つの好適な実施形態を示す。

【図2(c)】本発明の4つの好適な実施形態を示す。

【図2(d)】本発明の4つの好適な実施形態を示す。

【図3】図3は、本発明の一実施形態によるスロットにおけるスケジューリング重み付けの適合を示す。

【図4】図4は、本発明の一実施形態による一例のスケジューリングアルゴリズムを示す。

【発明を実施するための形態】

【0042】

原則として、インターネット11におけるサーバ12に接続された端末10である、いわゆるマルチホーミング化された端末10に対するリソースバンドリングは、少なくとも2つの別個の、すなわち異なるリンク( $L_1$ 、 $L_2$ 、 $\dots$ 、 $L_n$ )を用いて行われる。一例を図1に示す。

【0043】

端末10は、複数のライン( $L_1$ 、 $L_2$ 、 $\dots$ 、 $L_n$ )を用いてインターネット11に接続される。今日の端末10は通常、インターネット11におけるサーバ12との通信のためにラインの1つしか使用しない。通常は、使用されるインターフェースの選択は、階層的ポリシーに基づく。スマートフォンは、例えば、3GPPインターフェースよりもWi-Fiインターフェースを愛好し得る一方で、ラップトップはWi-Fiインターフェースよりもイーサネットインターフェースを愛好し得る。結果として、端末によって感受される最大帯域幅は、インターネット11への他のリンクが存在する場合であっても選択ネットワークインターフェースによって利用可能な帯域幅に制限される。すべての利用可能なインターフェース( $N$ )の同時使用の可能性を想定すると、端末10毎に感受される帯域幅は大幅に増加することになる。理想の場合、端末10毎の全体帯域幅( $BW_{ideal}$ )は、単に各インターフェース( $BW_i$ )によって与えられる帯域幅の合計となる。しかし、現実の実施では、もう少し難しく、100%のバンドリング効率( $E$ )はほとんど達成不可能である。したがって、端末10によって感受される帯域幅( $BW$ )は、以下のように表現され得る。

【数1】

$$BW = \left( \sum_{i=1}^N BW_i \right) E$$

【0044】

以降において、上記課題の１つ以上に対する解決手段を与えるアクセスリソースのバンドリングのためのアーキテクチャが提示される。さらに、バンドリングゲインを最大化するためにスケジューリング手法に対して重み付けを決定することができる革新的な解決手段が提示される。まず、提案される全般的なアーキテクチャを説明してから、ネイティブＴＣＰ情報に基づくパケットの重み付けスケジューリングの主な課題の詳細を説明する。

#### 【００４５】

本発明の好適な実施形態による解決手段の概観を図２（ａ）に示す。左側のノード（Ｃ）はユーザデバイス１００、例えばスマートフォン又はラップトップを表し、右側のデバイス（Ｓ）はインターネット（Ｉ）１０７におけるサーバ１０８、例えばアプリケーションサーバである。ユーザデバイス１００及びサーバ１０８は、標準的デバイスである。第１のルータ（Ｒ<sub>１</sub>）１０１及び第２のルータ（Ｒ<sub>２</sub>）１０４においてそれらにトランスペアレントな結合知能が実装されてる。第１のルータ１０１は、ユーザデバイス１００に対してインターネット１０７へのゲートウェイ（例えばＤＳＬルータ）であり、いくつかのアクセスインターフェース（ここでは、説明の目的のため、２つのアクセスインターフェースＬ<sub>１</sub>及びＬ<sub>２</sub>で示す）を有する。第２のルータ１０４は、ネットワークオペレータの１以上のトランスポートネットワーク１０３内でルータを表す。ルータ１０１及び１０４の双方は、ネットワークオペレータのトランスポートネットワーク（ＴＲ）１０３を介して相互に接続する。それは、第１のルータ１０１によって使用されるアクセスインターフェースＬ<sub>１</sub>及びＬ<sub>２</sub>の各々を介して、アクセスインターフェースＬ<sub>１</sub>及びＬ<sub>２</sub>の各々に対する第２のルータ１０４へのそれぞれのＴＣＰトンネル（ $tun_1$ 及び $tun_2$ ）を構成する。

#### 【００４６】

これらのトンネル $tun_1$ 及び $tun_2$ の目的は、データパケットをユーザデバイス１００とサーバ１０８の間で搬送することである。ルータ１０１及び１０４の各々はまた、実装されるそれぞれのスケジューリングモジュール（ＳＣ）１０２及び１０５を有する。なお、スケジューリングモジュール１０２及び１０５の各々は、それぞれのルータ１０１及び１０４に集積されて（内部接続されて）すべての関係する情報へのリアルタイムでのアクセスを取得し得る。一方、スケジューリングモジュールはまた、ルータに有線又は無線接続を介して単に取り付けられる（外部接続される）外部モジュールとして構成されてもよい。以降において、用語「接続され」は、内部接続される（すなわち集積される）こと及び／又は外部接続される（すなわち取り付けられる）ことの双方について使用されている。言い換えると、図２（ａ）はスケジューリングモジュール１０２及び１０５が第１及び第２のルータ１０１及び１０４にそれぞれ集積されていることを示すが、これらのモジュール１０２及び１０５を別個のデバイス（不図示）として提供することも可能である。スケジューリングモジュール１０２及び１０５は、異なるトンネルインターフェース $tun_1$ 及び $tun_2$ を介してパケットのスケジューリング及び分配を担う。最後に、第２のルータ１０４は、インターネット１０７に向けてインターフェース上でＮＡＴモジュール１０６を用いている。スケジューリングモジュール１０２及び１０５の各々並びに対応するインターフェースの組合せを、それぞれの結合モジュールともいう。

#### 【００４７】

図２（ｂ）～（ｄ）は、本発明による更なる代替の解決手段を示す。構成要素には、図２（ａ）のものに応じて対応の参照符号が付される。

#### 【００４８】

ここで、図２（ｂ）は代替の実施形態を示し、第１及び第２のルータ１０１及び１０４は省略される。すなわち、スケジューリングモジュール１０２及び１０５は、それぞれユーザデバイス１００及びサーバ１０８に集積される（接続される）。この構成では、第１のバンドリングは、第１のルータ１０１においてはもはや実行されないが、ユーザのユーザデバイス１００においては直接実行され、それはインターネットへの幾つかの経路を構成する。またさらに、第２のバンドリングは第２のルータ１０４においてはもはや実行されないが、インターネットにおけるサーバ１０８において直接実行される。

## 【 0 0 4 9 】

図 2 ( c ) 及び ( d ) は、図 2 ( a ) 及び ( b ) によって示す実施形態の更なる変形例を示す。特に、図 2 ( a ) の実施形態と比較して、図 2 ( c ) では第 2 のルータ 1 0 4 のみが省略され、図 2 ( d ) では第 1 のルータ 1 0 1 のみが省略される。第 1 及び第 2 のスケジューリングモジュール 1 0 2 及び 1 0 5 は、ユーザデバイス 1 0 0 及びサーバ 1 0 8 にそれぞれ集積され、又は別個の外部装置 ( 不図示 ) に提供され得ることが理解されなければならない。

## 【 0 0 5 0 】

例えば、D . K a s p e r、*「Multipath aggregation of heterogeneous access networks」*、SIGMultimedia Rec . Vol . 4 No . 1 pp . 2 7 - 2 8、2 0 1 2 年 3 月、で提示されるような周知の手法と比較して、図 1 の実施形態による本発明は 2 つの大きな簡素化を有する。第 1 に、パケットの再順序付けを担うデジッタバッファは使用されない。その代わりに、パケットの再順序付けは、TCP のネイティブ再順序付けアルゴリズムに残される。すなわち、再順序付けは、ユーザデバイス 1 0 0 又はインターネットサーバ 1 0 8 によって行われる。したがって、システムは、それは追加の再順序付けモジュールを必要としないので、ユーザデバイス 1 0 0 又はインターネットサーバ 1 0 8 によって行われた再順序付けの利益を受ける。これによって、他の解決手段と比較してシステムが簡素化される。第 2 に、TCP トンネル  $tun_1$  及び  $tun_2$  が第 1 のルータ 1 0 1 と第 2 のルータ 1 0 4 の間で使用されることに起因して、リンク能力の判定を担う経路モニタモジュールが不要となる。これらのトンネル  $tun_1$  及び  $tun_2$  からの情報に基づいて、リンク  $L_1$  及び  $L_2$  の容量が推定され得る。

## 【 0 0 5 1 】

したがって、本発明は、上述した課題の 1 つ以上に対する解決手段を提供することができる。NAT が第 2 のルータ 1 0 4 において実行され、これによって、トンネル  $tun_1$  及び  $tun_2$  の使用とともに、パケットが送信された経路から独立してサーバ 1 0 8 がクライアントの同じ IP アドレスを常時確認することが確実となる。サーバ 1 0 8 によって確認されるユーザデバイス 1 0 0 の IP アドレスは、第 2 のルータ 1 0 4 の公開 IP アドレスである。したがって、サーバ 1 0 8 からユーザデバイス 1 0 0 へのパケットは、常に第 2 のルータ 1 0 4 に宛てられることになる。TCP トンネル  $tun_1$  及び  $tun_2$  の使用は、経路モニタモジュールの必要性を回避し、同時にスケジューリング処理のために必要な情報を提供する。TCP が輻輳ウインドウ (  $C_{cwnd}$  ) の適合に基づいて常に各  $tun_1$  及び  $tun_2$  での帯域幅使用率を最大化しようとするのが、ここで利用される。TCP トンネル  $tun_1$  及び  $tun_2$  からの利用可能な情報の例は以下の通りである：

- ・  $C_{cwnd}$  のサイズの観点でのリンクの容量
- ・ TCP スロースタート閾値 (  $S_{thresh}$  )
- ・ TCP によって決定 / 計算された RTT である平滑化往復時間 (  $T_{RTT}$  )
- ・ 送出されたパケット数である  $P_{out}$
- ・ 肯定応答されたパケット数である  $P_{sacked}$
- ・ 再送信されたパケット数である  $P_{retrans}$
- ・ 喪失したとみなされるパケット数である  $P_{lost}$

## 【 0 0 5 2 】

これらの基本パラメータに基づいて、追加のパラメータが導出され得る。例えば：

- ・ 送信者によって送信されたが未だ肯定応答されていないパケットであるフライ状態にあるパケット (  $P_{fly}$  )。したがって、これらは送信者から受信者に現在伝搬しているパケットである。

$$P_{fly} = P_{out} - P_{sacked} + P_{retrans} - P_{lost}$$

- ・ 単に  $C_{cwnd} - P_{fly}$  であり、その容量に達するまで他の肯定応答を待つことなくリンクに送信され得るパケット数を表現する残容量 (  $C_{left}$  )
- ・ 規定された間隔内の残容量 (  $C_{left}$  ) の変化

## 【 0 0 5 3 】

スケジューリング重み付け (  $SW$  ) は、一般に関係式、例えば (  $A : B : C : D$  ) で表現される。以降の 4 つの利用可能なリンクを用いる例では、第 1 のリンクは  $A$  の重み付けを有し、次のリンクの packets を送信する前に  $A$  個の packets がこのリンクを介して送信され、 $B$  個の packets が第 2 のリンクを介して送信され、 $C$  個の packets が第 3 のリンクを介して送信され、最後に  $D$  個の packets が第 4 のリンクを介して送信されることを意味する。その後、第 1 のリンクが再度スケジューリングなどされる。スケジューリング重み付け (  $SW$  ) は 2 つの特性を有する。

- ・リンク間で送信される packets の割合を表す重み付け割合 (  $W_R$  )。例えば、2 つのリンクの場合、割合 2 は、リンク 2 と比較してリンク 1 を介して 2 倍多い packets が送信されることを意味する。

- ・1 つのリンクを介して連続して送信される packets の量を表す重み付け高さ (  $W_H$  )。高さ 10 は、連続 10 個の packets がリンク 1 を介して送信されることを意味する。

## 【 0 0 5 4 】

例 :  $W_R = 2$ 、 $W_H = 10$        $SW = ( 10 : 5 )$ 。

## 【 0 0 5 5 】

性能上の理由により、スケジューリング重み付けの適合は、図 3 に示すように、連続的ではなくタイムスロット (例えば、スロット  $n$ 、スロット  $n+1$ 、スロット  $n+2$ ) において行われ得る。これらのタイムスロットの継続時間は、パラメータ (  $D_T$  ) である。タイムスロット (  $X_n$  ) 中に収集されたパラメータのセットは、次のタイムスロット  $X_{n+1}$  についてのスケジューリングアルゴリズム (  $SW_{n+1} = f(X_n)$  ) に対する入力として用いられる。すなわち、これらのパラメータ (タイムスロット中に収集されたパラメータ)、及び選択的な入力又は構成パラメータに基づいて、スケジューリングアルゴリズムは次のスロットなどについてのスケジューリング重み付けを計算する。

## 【 0 0 5 6 】

本発明によるスケジューリングアルゴリズム例を図 4 に示す。ここで、説明の目的のため、2 つのトンネル (トンネル  $a$  及びトンネル  $b$ ) が利用可能であるものとする。スケジューリングアルゴリズムは、スリープ  $D_T$  で、すなわちスロット時間  $D_T$  だけ待機してステップ  $S10$  において開始する。packets が送信されるべき場合、スケジューリングアルゴリズムはステップ  $S11$  においてトンネル  $a$  の残容量  $C_{left, a}$  が所定の閾値 (  $THR$  ) 以上であるかを確認する。トンネル  $a$  の残容量  $C_{left, a}$  が所定の閾値以上である場合、スケジューリングアルゴリズムはステップ  $S12$  においてトンネル  $a$  を介して packets を送信し、この例では  $SW(1 : 0)$  のスケジューリング重み付けを有し、スケジューリングアルゴリズムは開始点 (ステップ  $S10$ ) に戻り、手順は他のスロットについて反復される。ここで、 $SW(1 : 0)$  は、トンネル  $a$  を介して 1 つの packets が送信され、トンネル  $b$  を介して packets は送信されないことを意味する。すなわち、トンネル  $a$  を介する packets のみが送信され、この例については  $W_R$  が無限であり、 $W_H$  が 1 である。一方、トンネル  $a$  の残容量  $C_{left, a}$  が所定の閾値以上でない場合、スケジューリングアルゴリズムは、ステップ  $S13$  においてトンネル  $b$  の残容量  $C_{left, b}$  が所定の閾値 (  $THR$  ) 以上でないかを判定する。トンネル  $a$  に対する閾値とトンネル  $b$  に対する閾値は必ずしも等しい必要はなく、異なる閾値となり得ることが理解されるべきである。

## 【 0 0 5 7 】

トンネル  $b$  の残容量  $C_{left, b}$  が所定の閾値以上である場合、スケジューリングアルゴリズムはステップ  $S14$  においてトンネル  $b$  を介して packets を送信し、この例では  $SW(0 : 1)$  のスケジューリング重み付けを有し、スケジューリングアルゴリズムは開始点 (ステップ  $S10$ ) に戻り、手順は他のスロットについて反復される。一方、トンネル  $b$  の残容量  $C_{left, b}$  が所定の閾値以上でない場合、スケジューリングアルゴリズムはいずれのトンネルも残容量を有しないと判定し、スケジューリングアルゴリズムはステップ  $S15$  において packets を送信し、この例では  $SW(1 : 1)$  のスケジューリング

重み付けを有し、スケジューリングアルゴリズムはステップ S 1 0 における開始点に戻り、手順は他のスロットについて反復される。

【 0 0 5 8 】

言い換えると、トンネル a は優先トンネルであり、それが空き容量を有する場合にはパケットはこのトンネル a のみを介して送信される。トンネル a が空き容量を有さない場合、トンネル b が空き容量を有する場合にパケットはトンネル b を介して送信される。両リンクが空き容量を有さない場合、ラウンドロビンスケジューリングが実行される。

【 0 0 5 9 】

ネットワークレイヤ手法の背景で記載されたが、当業者であれば、本発明は他のバンドリング手法、例えば M P T C P にも適用され得ることが分かるはずである。ただし、上記ネットワークレイヤ手法（レイヤ 3）を用いることによって、新たな T C P プロトコルが不要となり得る。したがって、標準 T C P プロトコルが使用され得る。

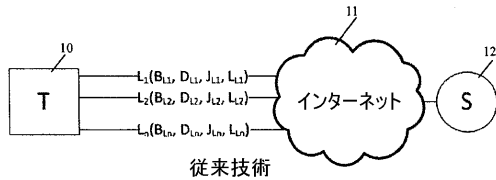
【 0 0 6 0 】

本発明はその範囲又は本質的特徴を逸脱することなくいくつかの形態で具現化され得るので、上記実施形態は、特に断りがない限り、前述の詳細のいずれによっても限定されず、付随する特許請求の範囲において規定される範囲内で広く解釈されるべきであり、したがって本発明に含まれるすべての変更及び変形は付随する特許請求の範囲によって包含されることが意図されることが理解されるべきである。

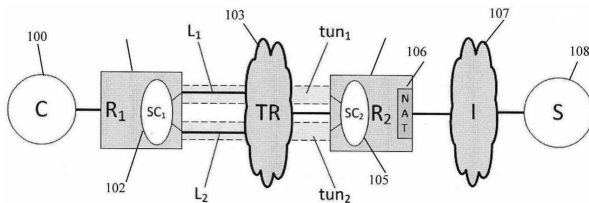
【 0 0 6 1 】

またさらに、特許請求の範囲において、文言「備える」は他の要素又はステップを除外せず、不定冠詞「a」又は「an」は複数形を除外しない。単一の部は、特許請求の範囲に記載されるいくつかの構成の機能を充足し得る。属性又は値との関連での用語「本質的に」、「約」、「概ね」などは、特に、属性又は値をそれぞれ正確に規定もする。特許請求の範囲におけるいずれの参照符号も、範囲を限定するものとして解釈されるべきではない。

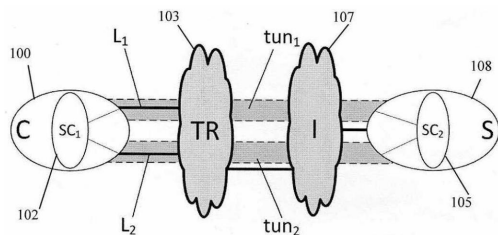
【 図 1 】



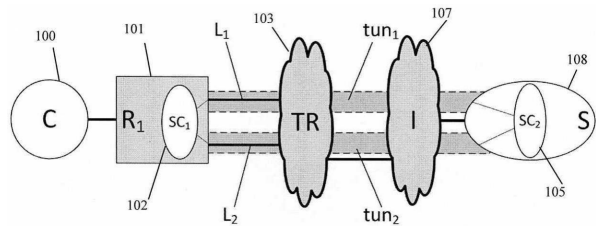
【 図 2 ( a ) 】



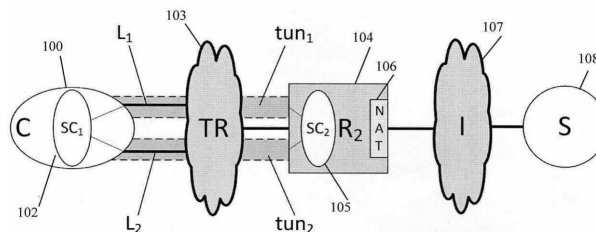
【 図 2 ( b ) 】



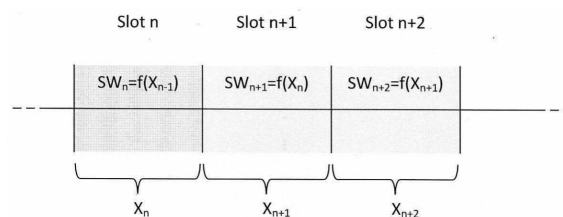
【 図 2 ( c ) 】



【 図 2 ( d ) 】



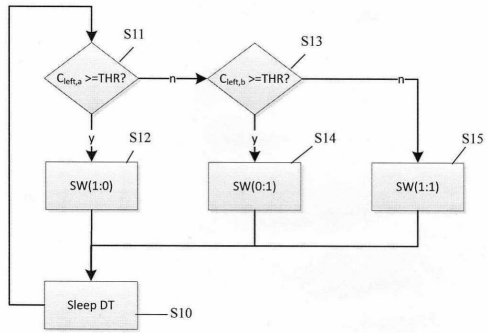
【 図 3 】



10

20

【 図 4 】



---

フロントページの続き

(72)発明者 バイヤー, ニコ

ドイツ 6 1 2 3 1 バート ナウハイム, ツム ザウアーブルンネン 4 3

(72)発明者 ガーザヴィ, アマー

ドイツ 6 4 3 9 0 エルツハウゼン, バーンシュトラーセ 1 5 7

審査官 中川 幸洋

(56)参考文献 米国特許出願公開第 2 0 1 1 / 0 1 7 0 5 4 5 ( U S , A 1 )

特開 2 0 1 5 - 0 5 0 7 4 6 ( J P , A )

特開 2 0 0 7 - 2 4 3 4 4 7 ( J P , A )

(58)調査した分野(Int.Cl. , D B 名)

H 0 4 L 1 2 / 8 9 1

H 0 4 L 2 9 / 0 0