

(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 1

H04L 12/891 (2013.01)
H04L 29/00 (2006.01)H04L 12/891
H04L 13/00

S

請求項の数 18 (全 15 頁)

(21) 出願番号 特願2017-553093 (P2017-553093)
 (86) (22) 出願日 平成28年4月8日 (2016.4.8)
 (65) 公表番号 特表2018-511275 (P2018-511275A)
 (43) 公表日 平成30年4月19日 (2018.4.19)
 (86) 國際出願番号 PCT/EP2016/057795
 (87) 國際公開番号 WO2016/162501
 (87) 國際公開日 平成28年10月13日 (2016.10.13)
 審査請求日 平成31年3月25日 (2019.3.25)
 (31) 優先権主張番号 15163169.4
 (32) 優先日 平成27年4月10日 (2015.4.10)
 (33) 優先権主張国・地域又は機関
歐州特許庁 (EP)

(73) 特許権者 597149146
ドイツ テレコム アーゲー
ドイツ国. 53113 ボン, フリードリ
ヒ - エベルト - アレエ 140
(74) 代理人 100094112
弁理士 岡部 譲
(74) 代理人 100106183
弁理士 吉澤 弘司
(74) 代理人 100114915
弁理士 三村 治彦
(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SPORT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SPORT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SPORT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SPORT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・トランスマネージメント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、輻輳ウインドウC_{wind}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、

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

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續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】

10

20

30

40

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SPORT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SPORT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SPORT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SPORT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SPORT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 1 3】

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 1 2 に記載の方法。

【請求項 1 4】

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ранスレーションモジュールを備える、請求項 1 3 に記載の方法。

【請求項 1 5】

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 1 4 に記載の方法。

【請求項 1 6】

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⁴⁰にスケジューリングアルゴリズムに基づく、請求項 1 4 又は 1 5 に記載の方法。

【請求項 1 7】

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 1 6 に記載の方法。

【請求項 1 8】

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 1 7 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」、Tech paper)。

【0004】

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。20

【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する、そのマルチホーミング構成である。30

【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くなるので、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40

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 2009/019258A1)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 SuperG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.Llairo及びD.Giustiniiano、「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、单一無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.Giustiniiano, E.Goma, A.Lopez Toledo, I.Dangerfield, J.Morillo及びP.Rodriguez、「Fair wlan backhaul aggregations」、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。10

【発明の概要】

【0009】20

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、オーバーレイトップ(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、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。30

【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40

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つのトランスポート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できるようになる。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。

20

【0015】

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}、及び喪失したもの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、 $P_{fly} = P_{out} - P_{sacked} + P_{retrans} - P_{lost}$ 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、 $C_{1\text{eff}} = C_{\text{cwnd}} - P_{f1y}$ 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 $C_{1\text{eff}}$ 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つのトランスポート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、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つのトランスポート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。異なるアクセスマッシュ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、アクセスマッシュ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、アクセスマッシュ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と想定している。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ран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L₁で送信される一方で、セッションBのパケットがリンクL₂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4050

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し得る。

10

【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。

20

【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。

30

【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選好し得る一方で、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選好し得る。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。理想的な場合、端末10毎の全体帯域幅($BW_{dealing}$)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。

40

【数1】

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

【0044】

50

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(R₁)101及び第2のルータ(R₂)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L₁及びL₂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L₁及びL₂の各々を介して、アクセスインターフェースL₁及びL₂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(tun₁及びtun₂)を構成する。

【0046】

これらのトンネルtun₁及びtun₂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tun₁及びtun₂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。

【0049】

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104のみが省略され、図2(d)では第1のルータ101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102及び105は、ユーザデバイス100及びサーバ108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。

【0050】

例えば、D. Kasper、「Multipath aggregation of heterogeneous access networks」、SIGMultiMedia Rec. Vol. 4 No. 1 pp. 27-28、2012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、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100又はインターネットサーバ108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100又はインターネットサーバ108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101と第2のルータ104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。
10
20

【0051】

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104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108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108によって確認されるユーザデバイス100のIPアドレスは、第2のルータ104の公開IPアドレスである。したがって、サーバ108からユーザデバイス100へのパケットは、常に第2のルータ104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(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ことが、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：
30

- ・ 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($S_{threshold}$)
- ・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}

【0052】

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

10

20

30

40

50

【0053】

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、次のリンクのパケットを送信する前に A 個のパケットがこのリンクを介して送信され、B 個のパケットが第 2 のリンクを介して送信され、C 個のパケットが第 3 のリンクを介して送信され、最後に D 個のパケットが第 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パケット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 倍多いパケットが送信されることを意味する。10

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

【0054】

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

【0055】

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。20

【0056】

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。パケット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 を介してパケットを送信し、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 つのパケットが送信され、トンネル b を介してパケットは送信されないことを意味する。すなわち、トンネル a を介するパケットのみが送信され、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。3040

【0057】

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 を介してパケットを送信し、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さないと判定し、スケジューリングアルゴリズムはステップ S15 においてパケットを送信し、この例では SW (1 : 1) のスケジューリング50

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。

10

【 0 0 6 0 】

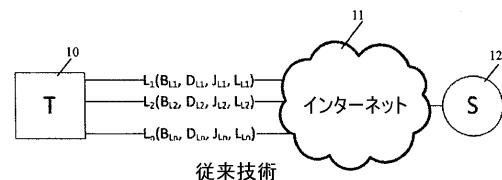
本発明はその範囲又は本質的特徴を逸脱することなくいくつかの形態で具現化され得るので、上記実施形態は、特に断りがない限り、前述の詳細のいずれによても限定されず、付随する特許請求の範囲において規定される範囲内で広く解釈されるべきであり、したがって本発明に含まれるすべての変更及び変形は付随する特許請求の範囲によって包含されることが意図されることが理解されるべきである。

【 0 0 6 1 】

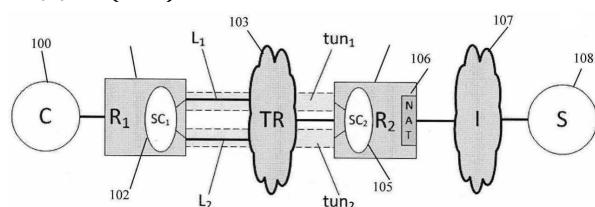
またさらに、特許請求の範囲において、文言「備える」は他の要素又はステップを除外せず、不定冠詞「 a 」又は「 a n 」は複数形を除外しない。单一の部は、特許請求の範囲に記載されるいくつかの構成の機能を充足し得る。属性又は値との関連での用語「本質的に」、「約」、「概ね」などは、特に、属性又は値をそれぞれ正確に規定もする。特許請求の範囲におけるいずれの参照符号も、範囲を限定するものとして解釈されるべきではない。

20

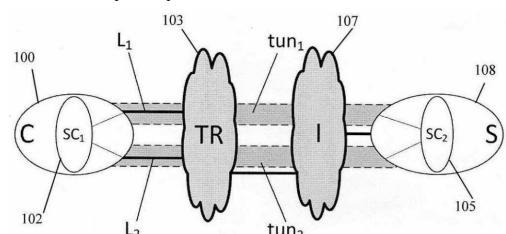
【 図 1 】



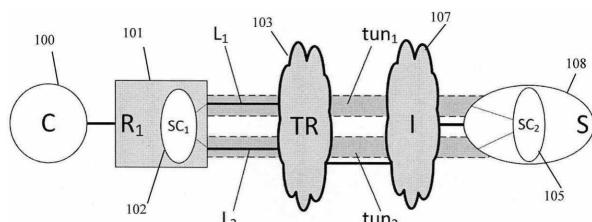
【 図 2 (a) 】



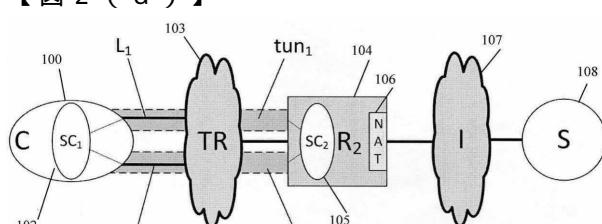
【 図 2 (b) 】



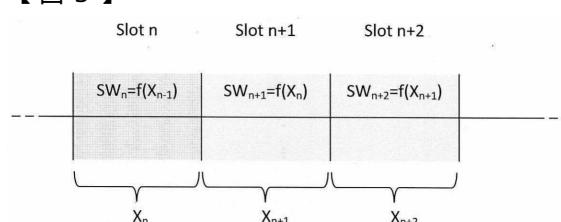
【 図 2 (c) 】



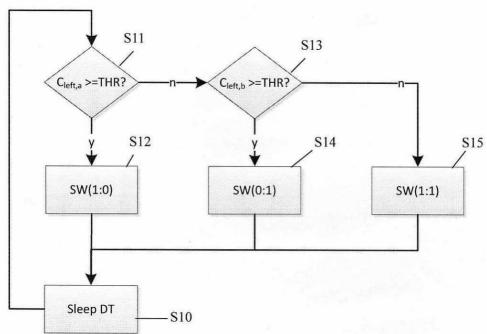
【 図 2 (d) 】



【 図 3 】



【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.CI. , D B 名)

H 0 4 L 1 2 / 8 9 1

H 0 4 L 2 9 / 0 0