

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号
特許第5431467号
(P5431467)

(45) 発行日 平成26年3月5日 (2014.3.5)

(24) 登録日 平成25年12月13日 (2013.12.13)

(51) Int.Cl.

F I

HO 4 L 12/825 (2013.01)

HO 4 L 12/825

HO 4 L 29/08 (2006.01)

HO 4 L 13/00 3 O 7 C

請求項の数 20 (全 13 頁)

(21) 出願番号	特願2011-515721 (P2011-515721)	(73) 特許権者	391030332
(86) (22) 出願日	平成21年6月18日 (2009.6.18)		アルカテルルーセント
(65) 公表番号	特表2011-526115 (P2011-526115A)		フランス国、75 007・パリ、アブニ
(43) 公表日	平成23年9月29日 (2011.9.29)		ユ・オクターブ・グレアール、3
(86) 国際出願番号	PCT/IB2009/053241	(74) 代理人	110001173
(87) 国際公開番号	W02009/156974		特許業務法人川口国際特許事務所
(87) 国際公開日	平成21年12月30日 (2009.12.30)	(72) 発明者	パイク、デイオン
審査請求日	平成23年2月23日 (2011.2.23)		カナダ国、オンタリオ・ケー・2・エス・
(31) 優先権主張番号	12/147, 137		1・ジェイ・6、ステイツツビル、モーニ
(32) 優先日	平成20年6月26日 (2008.6.26)		ングサン・クレセント・14
(33) 優先権主張国	米国 (US)	(72) 発明者	メガリテイ、マーク
			カナダ国、オンタリオ・ケー・2・シー・
			3・ビー・8、オタワ、トンプキンズ・ス
			トリート・125

最終頁に続く

(54) 【発明の名称】 特定のトラフィックフローにバックプレッシャーフロー制御を提供すること

(57) 【特許請求の範囲】

【請求項 1】

データ通信システムのトラフィックフローにバックプレッシャーフロー制御を提供するための装置であって、

トラフィックフローのデータパケットにパケット処理操作を行うためのパケット処理モジュールと、

少なくとも2つの異なるしきい値に基づいて、パケット処理モジュールの輻輳のレベルを判定するバックプレッシャーモジュールと、

輻輳のレベルに応じて、トラフィックフローのどれが一定のフロー制御を必要としているかを判定するフローマッピングエンジンと、

パケット処理モジュールから処理済みデータパケットを受け取り、必要に従って1つまたは複数のトラフィックフローに一定のフロー制御を適用するトラフィック管理モジュールと

を含み、

トラフィック管理モジュールが、トラフィックフローのそれぞれのためのそれぞれの仮想出力キューを含み、仮想出力キューから所与のトラフィックフローに対応するデータパケットが放出されるレートを変えることによって所与のトラフィックフローに一定のフロー制御を適用する、装置。

【請求項 2】

トラフィック管理モジュールが、レートをゼロまで減少させ、それによって仮想出力キ

ユーからのデータパケットの放出を停止することによってレートを変えるように動作可能である、請求項 1 に記載の装置。

【請求項 3】

フローマッピングエンジンが、表を含み、この表を使用して、輻輳のレベルに従って、所与のトラフィックフローの仮想出力キューを判定するように動作可能である、請求項 1 に記載の装置。

【請求項 4】

フローマッピングエンジンがさらに、表を使用して、輻輳のレベルに従って、一定のフロー制御を必要とするパケット処理モジュールへのイングレストラフィックのデータパケットの、それらに適用されるべき、1 つまたは複数の優先度を判定するように動作可能である、請求項 3 に記載の装置。

10

【請求項 5】

さらなるパケット処理のために、トラフィックフローの処理済みデータパケットがトラフィック管理モジュールからパケット処理モジュールに送られることができるように、トラフィック管理モジュールがパケット処理モジュールに通信可能に結合される、請求項 1 に記載の装置。

【請求項 6】

トラフィック管理モジュールが、トラフィックフローのそれぞれのためのそれぞれの仮想出力キューを含み、トラフィック管理モジュールが、仮想出力キューから所与のトラフィックフローに対応するデータパケットが放出されるレートを変えることによって所与のトラフィックフローに一定のフロー制御を適用するように動作可能である、請求項 5 に記載の装置。

20

【請求項 7】

データ通信システムにおいてトラフィックフローのフロー制御を行う方法であって、少なくとも 2 つの異なるしきい値に基づいて、データ通信システムにおけるパケット処理モジュールの輻輳のレベルを検出することと、

輻輳のレベルに応じてパケット処理モジュールによって処理されているトラフィックフローに一定のフロー制御が適用されるべきかどうかを判定することと、

判定が肯定であることに応答してトラフィックフローに一定のフロー制御を適用することと

30

を含み、

仮想出力キューから所与のトラフィックフローに対応するデータパケットが放出されるレートを変えることによって所与のトラフィックフローに一定のフロー制御が適用される方法。

【請求項 8】

検出するステップは、

パケット処理モジュールの入力バッファの現在の充填レベルを少なくとも 2 つの異なるしきい値と比較することと、

比較の結果として、現在の充填レベルと少なくとも 2 つの異なるしきい値との関係に対応する値を有するバックプレッシャーメッセージを形成することと、

40

を含む、請求項 7 に記載の方法。

【請求項 9】

判定するステップは、

パケット処理モジュールへのイングレストラフィックフロー、パケット処理モジュールへの再循環トラフィックフロー、およびデータ通信システムからのイーグレストラフィックフローのうちの、1 つまたは複数について判定を行うこと

をさらに含む、請求項 8 に記載の方法。

【請求項 10】

判定を行うステップは、

バックプレッシャーメッセージの値に従って、パケット処理モジュールへのイングレス

50

トラフィックフローのデータパケットの優先度があればどの優先度がデータパケットに一定のフロー制御を適用すべきか、ならびに再循環トラフィックフローおよびイーグレストラフィックフローのタイプがあればどのタイプがトラフィックフローに一定のフロー制御を適用すべきかを判定すること

をさらに含む、請求項 9 に記載の方法。

【請求項 11】

適用するステップは、

一定のフロー制御を必要とすると判定された優先度を有するデータパケットのパケット処理モジュールへのイーグレストラフィックフローを停止すること、ならびに

一定のフロー制御を必要とすると判断されたトラフィックタイプのイーグレストラフィックフローおよび再循環トラフィックフローを停止することと、

をさらに含む、請求項 10 に記載の方法。

【請求項 12】

適用するステップはさらに、

一定のフロー制御を必要とするトラフィックフローのそれぞれのトラフィックフローに対応する仮想出力キューがデータパケットの放出を停止するように、各仮想出力キューの状態を変更することと

をさらに含む、請求項 11 に記載の方法。

【請求項 13】

先行するステップを継続的に繰り返すことをさらに含む、請求項 7 に記載の方法。

【請求項 14】

バックプレッシャーメッセージの 3 つの異なる値を定義することをさらに含む、請求項 8 に記載の方法。

【請求項 15】

現在の充填レベルが第 1 のしきい値未満のときにバックプレッシャーメッセージの第 1 の値を定義することをさらに含む、請求項 14 に記載の方法。

【請求項 16】

現在の充填レベルが第 2 のしきい値未満であるが、第 1 のしきい値以上のときにバックプレッシャーメッセージの第 2 の値を定義することをさらに含む、請求項 15 に記載の方法。

【請求項 17】

現在の充填レベルが第 2 のしきい値以上のときにバックプレッシャーメッセージの第 3 の値を定義することをさらに含む、請求項 16 に記載の方法。

【請求項 18】

バックプレッシャーメッセージの第 2 の値に 응답して、ベストエフォート (BE) のトラフィックフローのみを停止することをさらに含む、請求項 15 に記載の方法。

【請求項 19】

ベストエフォートのトラフィックフローおよび完全優先転送 (EF) のトラフィックフローを停止し、

バックプレッシャーメッセージの第 3 の値に 응답して、相対的優先転送 (AF) のトラフィックフローのみを送信することをさらに含む、請求項 17 に記載の方法。

【請求項 20】

データ通信システムのトラフィックフローにバックプレッシャーフロー制御を提供するための装置であって、

トラフィックフローのデータパケットにパケット処理操作を行うためのパケット処理モジュールと、

少なくとも 2 つの異なるしきい値に基づいて、パケット処理モジュールの輻輳のレベルを判定するバックプレッシャーモジュールとを含み、バックプレッシャーメッセージの少なくとも 3 つの異なる値がトラフィックフローを選択するのに使用され、さらに、

輻輳のレベルに応じて、トラフィックフローのどれが一定のフロー制御を必要としてい

10

20

30

40

50

るかを判定するフローマッピングエンジンと、

パケット処理モジュールから処理済みデータパケットを受け取り、必要に従って1つまたは複数のトラフィックフローに一定のフロー制御を適用するトラフィック管理モジュールと

を含む、装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、通信ネットワークを対象にし、詳細にはデータ通信システムにおけるフロー制御を対象にする。

【背景技術】

【0002】

データ通信システムで利用されるフロー制御技術では、一般にバックプレッシャーシグナリングが用いられる。こうした技術は、データ通信システム間の通信リンクで使用されることが可能であり、またデータ通信システムのサブシステム間の通信にも使用されることが可能である。一般に、このようなバックプレッシャーフロー制御技術の目的は、受信側のキューにおける、これらのキューをオーバーフローすることによって引き起こされるデータパケットの損失を防ぐことである。

【0003】

従来のバックプレッシャーフロー制御技術は、一般に単純なオンオフシグナリングを使用する。この技術によれば、受信側キューにおいて充填レベルのしきい値を超えることによってバックプレッシャー信号（例えば停止）が生成され、この信号はその後データパケットのソースに送られる。システム間通信の場合、データパケットのソースは通常、別のデータ通信システムの送信バッファである。システム内通信の場合、データパケットのソースは通常、サブシステムのイーグレス（e g r e s s）キューであり、受信側のキューは、別のサブシステムのイングレス（i n g r e s s）キューである。1つのタイプのバックプレッシャー信号（例えば停止）は、このキューへのデータパケットの送信をその後の通知まで一時停止すべきであることをソースに示し、その後の通知は、別のタイプのバックプレッシャー信号（例えば再開）の形式で与えられる。場合によっては2つ以上のデータパケットソースがある可能性があり、その場合、バックプレッシャー信号は、通常それらのソースのすべてに送信されることになる。この技術は実装が簡単であり、処理するために送信元に返送される必要のある情報の量が限られているので、データ通信システムにおいて非常にうまく行っている。

【0004】

単純なオンオフ以上のシグナリングを提供するより高度なバックプレッシャーフロー制御技術が知られている。これらは、キューにおいて連続的に高い充填レベルのしきい値が超過されるにつれて、キューへのデータパケットのフローレートを徐々に下げる技術を含む。他の技術は、単に一定の優先度のデータパケットにバックプレッシャーフロー制御を適用するための手段を有する。しかしながら、優先度のみを使用することは、場合によっては十分ではない可能性があり、特に、1つまたは複数のタイプのトラフィックフローに対して、これらのフロー中のデータパケットの優先度にかかわらず個々のフロー制御を選択的に行うことが好ましいとき、十分ではない可能性がある。

【0005】

データ通信システムのための最先端のトラフィック管理を設計する際の課題の1つは、システムの内部の輻輳状態が、システムのパケット処理およびトラフィック管理エンジンの性能の限界ならびにそのバスインタフェース上の制限によるものであることが多いことである。一般にこうした制限は、このような輻輳状態に柔軟に対処する限られた能力しかないバックプレッシャーフロー制御の解決策をもたらす。

【発明の概要】

【発明が解決しようとする課題】

10

20

30

40

50

【 0 0 0 6 】

したがって、データ通信システムによって搬送される様々なタイプのトラフィックフローに柔軟に適用されることが可能であるバックプレッシャーフロー制御を提供する必要性がある。

【課題を解決するための手段】

【 0 0 0 7 】

データ通信システムにおいて、本発明の諸実施形態は、システムのパケットプロセッサにおける輻輳のレベルに応答してトラフィックフローに一定のフロー制御を適用することによって、トラフィックフローにバックプレッシャーフロー制御を選択的に提供する。

【 0 0 0 8 】

本発明の一態様によれば、データ通信システムのトラフィックフローにバックプレッシャーフロー制御を提供するための装置が提供される。この装置は、トラフィックフローのデータパケットにパケット処理操作を行うためのパケット処理モジュールと、パケット処理モジュールの輻輳のレベルを判定するためのバックプレッシャーモジュールと、輻輳のレベルに基づいて、トラフィックフローのどれが一定のフロー制御を必要とするかを判断するためのフローマッピングエンジンと、パケット処理モジュールから処理されたデータパケットを受け取るための、および必要に従って一定のフロー制御をトラフィックフローの1つまたは複数に適用するためのトラフィック管理モジュールとを含む。

【 0 0 0 9 】

本発明の別の実施形態によれば、データ通信システムにおいてトラフィックフローのフロー制御を行う方法が提供される。この方法は、データ通信システムにおけるパケット処理モジュールの輻輳のレベルを検出するステップと、輻輳のレベルに応じて一定のフロー制御が、パケット処理モジュールによって処理されているトラフィックフローに適用されるべきかどうかを判定するステップと、判定が肯定であることに応答して一定のフロー制御をトラフィックフローに適用するステップとを含む。

【 0 0 1 0 】

有利には、本発明の諸実施形態は、データ通信システムのコストを実質的に追加させることなく実装されることが可能であり、これは、トラフィック管理がデータ通信システム間の主要な差別化要因である可能性があることから、特に有益である可能性がある。

【 0 0 1 1 】

本発明は、図面を参照した次の詳細な説明からさらに理解されるであろう。

【図面の簡単な説明】

【 0 0 1 2 】

【図 1】本発明の一実施形態に従って特定のトラフィックフローにバックプレッシャーフロー制御を提供するための装置のハイレベルブロック図である。

【図 2】図 1 のフロー制御表の一例を示す図である。

【図 3】図 1 の仮想出力キュー（VOQ）割当て表の一例を示す図である。

【図 4】図 1 の例示的バックプレッシャーメッセージおよび対応するVOQフロー制御メッセージを示す図である。

【図 5】本発明の第 2 の実施形態によりトラフィックフロー上でバックプレッシャーフロー制御を行う方法のフローチャートである。

【発明を実施するための形態】

【 0 0 1 3 】

図 1 を参照すると、データ通信システムにおいてバックプレッシャーフロー制御を提供するための装置 10 が、パケット処理モジュール 14、トラフィック管理モジュール 16、およびマイクロプロセッサモジュール 18 と通信している入力/出力（I/O）インタフェースモジュール 12 を含んでいる。またパケット処理モジュール 14 とトラフィック管理モジュール 16 は、互いと通信している。

【 0 0 1 4 】

I/O インタフェースモジュール 12 は、顧客ポートなど、データ通信システム内のソ

10

20

30

40

50

ースからイングレスデータパケット22を受け取り、これらをイングレストラフィック26としてパケット処理モジュール14に転送するための複数の入力ポートを含んでいる。I/Oインタフェースモジュール12はまた、データ通信システムのI/Oポートのような通信システム内の宛先にイーグレスデータパケット24を送信するための複数の出力ポートを含んでいる。I/Oインタフェースモジュール12はさらに、I/Oインタフェースモジュール12がイングレストラフィック26上にバックプレッシャーフロー制御を提供できるように、パケット処理モジュール14から受け取られたバックプレッシャーメッセージをイングレスデータパケット22の優先度に基づいてマッピングするためのフローマッピングエンジン20を含んでいる。フロー制御テーブル46は、後により詳細に説明するように、このようなマッピングを行うためのフローマッピングエンジン20に含まれている。

10

【0015】

パケット処理モジュール14は、イングレストラフィック26としてI/Oインタフェースモジュール12から受け取られたイングレスデータパケット22を格納するための入力バッファ25を含んでいる。それぞれ第1のしきい値T1および第2のしきい値T2として示される、複数の連続的な充填レベルしきい値が、入力バッファ25に対して定義される。灰色の網掛けで示される入力バッファ25の充填レベルは、第1のしきい値と第2のしきい値の間である。パケット処理モジュール14はまた、複数のM個のパケットプロセッサ(PP)を含んでおり、これは、入力バッファ25から受け取られたデータパケットに高スループットのパケット処理操作を行うために並列に動作する。パケット処理モジュール14はさらに、入力バッファ25の現在の充填レベルを複数の充填レベルしきい値と比較することによってバックプレッシャーメッセージ30を作り出すバックプレッシャーモジュール29を含んでいる。バックプレッシャーメッセージ30は、I/Oインタフェースモジュール12に伝えられ、ここでバックプレッシャーメッセージは、イングレスデータパケット22の優先度があれば、どの優先度が一定のフロー制御をパケットに適用すべきかを判定するために使用される。

20

【0016】

インターネットプロトコルDiffServ(Differentiated Service、差別化サービス)のトラフィックの分類を参照して、前述のバックプレッシャーフロー制御の一例を示す。2つの充填レベルしきい値を使用して、本明細書ではタイプとも呼ばれる、3つの異なる値のバックプレッシャーメッセージ30が定義されることが可能である。このようなメッセージの第1のタイプは、入力バッファ25の現在の充填レベルが第1のしきい値T1未満であることに対応する。同様に、このようなメッセージの第2のタイプは、現在の充填レベルが第2のしきい値T2未満であるが、第1のしきい値T1以上であることに対応する。最後にこのようなメッセージの第3のタイプは、現在の充填レベルが第2のしきい値T2以上であることに対応する。第1のタイプのバックプレッシャーメッセージ30を受け取ると、I/Oインタフェースモジュール12は、イングレストラフィック26としてすべての優先度のデータパケットをパケット処理モジュール14に渡し続ける。しかしながら、第2のタイプのバックプレッシャーメッセージ30を受け取ることに対応して、I/Oインタフェースモジュール12は、イングレストラフィック26にこのようなデータパケットを含まないことによって、ベストエフォート(BE)のトラフィックフローを停止させる。さらに、第3のタイプのバックプレッシャーメッセージ30を受け取ることに対応して、I/Oインタフェースモジュール12は、イングレストラフィック26にこのようなデータパケットを含まないことによって、ベストエフォートのトラフィックフローおよび完全優先転送(EF)のトラフィックフローを停止する。この場合、相対的優先転送(AF)タイプのデータパケットのみが、パケット処理モジュールに送信されるイングレストラフィック26に含まれることになる。新しいバックプレッシャーメッセージ30を受け取ると、このようなメッセージの以前のものを破棄し、I/Oインタフェースモジュール12によって最も新しく受け取られたバックプレッシャーメッセージ30の内容に従ってトラフィックフローは停止される、または再開される。この例では、トラフィックフローを停止することが、このようなフローに一定のフロー

30

40

50

制御を適用することを構成する。

【 0 0 1 7 】

パケット処理モジュール 1 4 が、入力バッファ 2 5 からデータパケットを読み取り、複数のパケットプロセッサ 2 8 を介してこれらにパケット処理を行うと、処理済みデータパケット 3 2 がもたらされ、これがトラフィック管理モジュール 1 6 に伝えられる。データパケットに行われる処理のタイプは、基本的なプロトコルの終了操作を含み、ディープパケットインスペクション (D P I) のような高度な操作を含む場合もある。トラフィック管理モジュール 1 6 は、処理済みデータパケット 3 2 にトラフィック管理操作を行うためのキューイングおよびスケジューリングモジュール 3 3 を含み、データパケットの管理されたトラフィックフローをもたらし、キューイングおよびスケジューリングモジュール 3 3 は、管理されたトラフィックフローのデータパケットが、イーグレストラフィック 3 5 として I / O インタフェースモジュール 1 2 に伝えられる、またはさらなる処理を必要とする再循環トラフィック 3 6 としてパケット処理モジュール 1 4 に伝えられる前にこれらを格納するための複数の仮想出力キュー (V O Q) 3 4 を含む。

10

【 0 0 1 8 】

通常、複数の V O Q の各 1 つは、それぞれのトラフィックフローと関連付けられるが、データ通信システムの容量が必要条件を上回る場合、場合によってはある V O Q はどのトラフィックフローにも割り当てられない可能性がある。各トラフィックフローは、トラフィックフロー中のデータパケットのソースアドレスおよび宛先アドレスならびにプロトコルポート識別子によって他のトラフィックフローと区別される。さらに、トラフィックフローのいずれも、トラフィックフローの所有者または顧客、ならびにトラフィックフローによってサポートされるサービスなど他の要素によって識別されることが可能である。

20

【 0 0 1 9 】

フローマッピングエンジン 2 0 の機能に戻ると、エンジン 2 0 はさらに、V O Q フロー制御メッセージ 3 1 を作り出すために、バックプレッシャーメッセージ 3 0 を V O Q 3 4 にマッピングすることを行うように動作可能である。このために、後にさらに詳細に説明するように、フローマッピングエンジン 2 0 に V O Q 割当て表 4 8 が含まれる。V O Q フロー制御メッセージ 3 1 は、トラフィック管理モジュール 1 6 に送信され、パケット処理モジュール 1 4 に送信される再循環トラフィック 3 6 に、および必要に応じてイーグレストラフィック 3 5 に、フロー制御を行う際に使用される。関与するサービスおよびアプリケーションによっては、イーグレストラフィック 3 5 は入力バッファ 2 5 の現在の充填レベルに直接に影響を与えないが、場合によってはイーグレストラフィック 3 5 をフロー制御することは、入力バッファ 2 5 の現在の充填レベルを低減させるという望ましい効果を有することができる。

30

【 0 0 2 0 】

優先度ベースのバックプレッシャーメッセージ 3 0 の 3 つのタイプに関する前述の例を参照して、このようなメッセージを V O Q 3 4 にマッピングして V O Q フロー制御メッセージ 3 1 を作り出す例を次に示す。例えば、4 つの V O Q 3 4 があると仮定し、そのうちの第 1 は再循環トラフィック 3 6 用 (例えば D P I 操作を含む) であり、そのうちの第 2 は、ミラーリングされているイーグレストラフィック 3 5 用 (例えば政府機関による合法の傍受用) であり、そのうちの第 3 は、プレミアムサービスに対する顧客の支払い用であり、そのうちの第 4 はベストエフォートサービスに対する顧客の支払い用である。マッピングエンジン 2 0 が第 1 のタイプのバックプレッシャーメッセージ 3 0 を受け取る (すなわち、現在の充填レベルが第 1 のしきい値 T 1 を下回る) 場合、V O Q 3 4 のどれも一定のフロー制御をこれらに適用しない。したがって、V O Q フロー制御メッセージ 3 1 は、その趣旨で生成される。この例では、単に 1 つの一定のフロー制御が使用され、これは V O Q からの所与のトラフィックフローのすべてのデータパケットの放出を停止することであるが、レート制限など、フロー制御の他の手段が代わりに使用されてもよい。マッピングエンジン 2 0 が第 2 のタイプのバックプレッシャーメッセージ 3 0 を受け取る (すなわち、現在の充填レベルが第 1 のしきい値以下、第 2 のしきい値未満である) 場合、ベスト

40

50

エフォートサービスの顧客のトラフィックは、一定のフロー制御がこれらに適用される。したがって、第4のVOQからのトラフィックフローが停止されるべきであることを示すVOQフロー制御メッセージ31が生成される。マッピングエンジン20が第3のタイプのバックプレッシャーメッセージ30を受け取る（すなわち現在の充填レベルが第2のしきい値以上である）場合、再循環トラフィック36およびベストエフォートサービスの顧客のトラフィックは、一定のフロー制御がこれらに適用される。したがって、第1および第4のVOQからのトラフィックフローが停止されるべきであることを示すVOQフロー制御メッセージ31が生成される。この例では、第2および第3のVOQのトラフィックによって必要とされるサービスのレベルにより、これらのVOQからのトラフィックは、いかなるタイプのバックプレッシャーメッセージ30についても停止されないことに注意

10

【0021】

図2を参照して、次にフロー制御表46の一例について説明する。この例は、バックプレッシャーメッセージ30およびVOQフロー制御メッセージ31に関する前述の例に対応する。フロー制御表46の左列は、前述の例における3つのタイプのバックプレッシャーメッセージ30を含み、各タイプがそれぞれの行にある。残りの列は、イングレストラフィック26のどのデータパケットの優先度がフロー制御されるべきであるか、ならびに再循環トラフィック36およびイーグレストラフィック35のどのタイプのトラフィックフローがフロー制御されるべきかを示し、イングレストラフィック、再循環トラフィック、およびイーグレストラフィックのそれぞれに対してそれぞれの列がある。フローマッピングエンジン20は、パケット処理モジュール14から現在受信済みのバックプレッシャーメッセージ30のタイプによりフロー制御表46中の行にインデックスすること、ならびに行から一定のフロー制御を適用すべきデータパケットの優先度およびトラフィックフローのタイプを読み取ることによって、バックプレッシャーメッセージ30をイングレストラフィック26の優先度に、ならびに再循環36フローおよびイーグレストラフィック35フローのタイプにマッピングする。この例では、適用される一定のフロー制御は、問題のトラフィックフローのデータパケットのフローを停止することである。例えば、図2でタイプ3と示した第3のタイプのバックプレッシャーメッセージ30については、フローマッピングエンジン20は、マッピングテーブル46中の第3の行を読み取り、これは、イングレストラフィック26のベストエフォートおよび完全優先転送の優先度が停止されるべきこと、再循環トラフィック36のDPIタイプが停止されるべきこと、ならびにイーグレストラフィック35のベストエフォートサービスタイプが停止されるべきことを示す。このように、バックプレッシャーフロー制御は、パケット処理モジュール中の輻輳レベルを示すバックプレッシャーメッセージに応答してトラフィックフローに一定のフロー制御を適用することによって、特定のトラフィックフローに選択的に提供される。新しいバックプレッシャーメッセージ30を受け取ると、このようなメッセージの以前に受信したものを破棄し、I/Oインタフェースモジュール12によって最も新しく受け取られたバックプレッシャーメッセージ30の内容に従って様々な手段のフロー制御がトラフィックフローに適用されるようにする。

20

30

【0022】

図3を参照し、次にVOQ割当て表48の一例について説明する。この例は、バックプレッシャーメッセージ30およびVOQフロー制御メッセージ31に関する前述の例に対応する。VOQ割当て表48の左列は、VOQの識別子を含み、各VOQはそれぞれの行に数値識別子を有する。他の列は、そのVOQによって管理されているトラフィックフローのタイプの表示を含んでおり、各行には、その行で特定されるVOQによって管理されるべきトラフィックフローのそれぞれのトラフィックフロー表示が含まれている。フローマッピングエンジン20は、まず一定のフロー制御を適用される再循環トラフィック36およびイーグレストラフィック35のトラフィックフローのタイプを判定することによって、バックプレッシャーメッセージ30をVOQにマッピングする。これは、図2を参照してすでに説明したように、フロー制御表46を使用して行われる。次に、VOQ割当て

40

50

表 4 8 からトラフィックフローの表示を読み取り、これをすでに判定されたように一定のフロー制御を適用されるべきトラフィックフローの表示と比較することによって、フローマッピングエンジン 2 0 は、結果的に一致するこのようなすべての比較について対応する V O Q 識別子を記録する。

【 0 0 2 3 】

図 4 を参照して、次にバックプレッシャーメッセージ 3 0 および対応する V O Q フロー制御メッセージ 3 1 の一例について説明する。この例は、バックプレッシャーメッセージ 3 0 および V O Q フロー制御メッセージ 3 1 に関する前述の例に対応する。前述の例で使用された 3 つの異なるタイプのバックプレッシャーメッセージ 3 0 が、図の左側に示されており、それぞれの実際の形式は、メッセージタイプの序数を示す 2 進数である。図の左側に示された優先度に基づくバックプレッシャーメッセージ 3 0 の各タイプについて、対応する V O Q フロー制御メッセージ 3 1 がこれに隣接して図の右側にある。各 V O Q フロー制御メッセージ 3 1 の形式は、各 V O Q 3 4 に 1 ビットのフィールドを含み、メッセージ 3 1 中のこのような各フィールドの位置が、V O Q 割当て表 4 8 中の V O Q の数値識別子に対応する。V O Q フロー制御メッセージ 3 1 の内容は、各フィールドに 2 進数のインジケータを含み、このような各インジケータは、そのフィールドに対応する V O Q によって管理されているトラフィックフローが一定のフロー制御を適用されるべきかどうかの指示を与える。この例では、適用される一定のフロー制御は、問題の V O Q からの所与のトラフィックフローのすべてのデータパケットを停止することである。しかしながら、この手段は、トラフィックフローをレート制限するなどの他の手段とすることも可能である。図 4 では、「 1 」という値はトラフィックフローが停止されるべきであることを示すのに対し、「 0 」という値はトラフィックフローが停止されるべきではないことを示す。

【 0 0 2 4 】

マイクロプロセッサモジュール 1 8 は、I / O インタフェースモジュール 1 2 を介して制御プレーントラフィック 4 4 を制御プレーンに挿入する。同様に制御プレーントラフィック 4 4 は、I / O インタフェースモジュール 1 2 によって制御プレーンから取り出され、マイクロプロセッサモジュール 1 8 に伝えられる。フロー制御表 4 6 および V O Q 割当て表 4 8 をポピュレートするための制御プレーンからの命令は、マイクロプロセッサモジュール 1 8 を介してフローマッピングエンジン 2 0 へ伝えられる。これらの命令は、オペレータコンソールを介してデータ通信システム上で行われるトラフィックおよびサービスプロビジョニング操作の結果として、ネットワーク管理システム、または他のタイプの操作サポートシステムを起動させる。またフローマッピングエンジン 2 0 のマッピング機能は、特定のタイプの制御プレーントラフィック 4 4 にフロー制御を行うために、制御プレーントラフィックに適用される。

【 0 0 2 5 】

図 5 を参照して、次に本発明の第 2 の実施形態に従ってデータパケットのトラフィックフローのフロー制御を行う方法について説明する。この方法は、データ通信システムにおけるパケット処理モジュールの輻輳のレベルを判定するステップ 1 0 0 に始まる。場合によっては、輻輳のレベルを判定することは、パケット処理モジュールの入力バッファの現在の充填レベルを 1 つまたは複数の充填レベルしきい値と比較し、そのような比較の結果としてバックプレッシャーメッセージを形成するステップを含む。このような場合、バックプレッシャーメッセージは、現在の充填レベルと 1 つまたは複数の充填レベルしきい値との関係に対応する値を有する。ステップ 1 0 0 に続き、この方法は、判定された輻輳のレベルによって、パケット処理モジュールにより処理されている 1 つまたは複数のトラフィックフローに一定のフロー制御が適用されるべきかどうかを判定するステップ 1 0 2 へ進む。場合によっては、一定のフロー制御が適用されるべきかどうかを判定することは、パケット処理モジュールへのイングレストラフィックフロー、パケット処理モジュールへの再循環トラフィックフロー、およびデータ通信システムからのイーグレストラフィックフローのうちの 1 つまたは複数についてこのように判定することを含む。さらに、輻輳のレベルを判定するステップ 1 0 0 においてバックプレッシャーメッセージが形成された場

合、本ステップ102は、このバックプレッシャーメッセージの値に従って、パケット処理モジュールへのイングレストラフィックフローのデータパケットの優先度があればどの優先度が一定のフロー制御をトラフィックフローに適用すべきであるか、および再循環およびイーグレストラフィックフローのタイプがあればどのタイプが一定のフロー制御をトラフィックフローに適用すべきであるかを判定することを含む。このステップ102に続いてこの方法は、判定に従ってトラフィックフローに一定のフロー制御を適用するステップ104に進む。場合によっては、一定のフロー制御を適用するステップ104は、一定のフロー制御を必要とすると判定された優先度を有するデータパケットのために、パケット処理モジュールへのイングレストラフィックのデータパケットのフローを停止すること、ならびに一定のフロー制御を必要とすると判断されたトラフィックタイプのイーグレスおよび再循環トラフィックフローのデータパケットのフローを停止することを含む。場合によっては、イーグレスおよび再循環トラフィックフローのトラフィック管理がVOQを含む場合、このようなトラフィックフローを停止することは、このようなトラフィックフローのそれぞれのフローに対応して各VOQの状態を変更し、VOQがそこからのデータパケットのフローを停止するようにすることを含む。この方法は、一定のフロー制御を適用するステップ104で終了する。しかしながら、通常の動作の下では、この方法は継続的に繰り返され、そのような各繰り返しでは、一定のフロー制御がこの方法の判定によって必要に従ってトラフィックフローに適用され、前の繰り返しから生じる適用から解放される。すなわち、この方法の各繰り返しは前の繰り返しから独立して動作し、一部のトラフィックフローが1つの繰り返しの結果として一定のフロー制御（例えば停止）を適用される場合、同じトラフィックフローは次の繰り返しにおいて一定のフロー制御（例えば再開）を適用されない場合がある。

10

20

【0026】

前述に照らして、パケット処理モジュール14の入力バッファ25が輻輳し始めるとき、特に入力バッファ25の充填レベルが1つまたは複数の充填レベルしきい値を超えると、異なるタイプのバックプレッシャーメッセージ30が生成されることは明らかである。フローマッピングエンジン20は、パケット処理モジュール14から受け取るこのバックプレッシャーメッセージ30を、イングレスデータパケット22の優先度に、およびVOQ34にマッピングする。後者から、フローマッピングエンジン20はVOQフロー制御メッセージ31を生成し、これをトラフィック管理モジュール16に送ってVOQ34の1つまたは複数からのトラフィックフローをフロー制御する。例えば、1つのタイプのバックプレッシャーメッセージは、イーサネット（登録商標）もしくはマルチキャストトラフィックと関連するVOQに、またはデータ通信システムのインターネットプロトコルセキュリティ（IPSEC）回路カードのようなサービスI/Oブレードと関連するVOQにマッピングすることが可能である。このような機能は、ネットワークトラフィックがよりインテリジェントに処理される、および輻輳状態の間分散されることを可能にするなど、トラフィック管理に有利な強化を行う。このような機能は、ネットワークトラフィックが輻輳状態の間によりインテリジェントに処理され、分散されることを可能にするなど、トラフィック管理に有利な強化を行い、これがサービスルータのようなデータ通信システムにおいて重要な差別化要因となりうる。

30

40

【0027】

上述の本発明の諸実施形態には、特許請求の範囲で定義される本発明の範囲を逸脱することなく、数多くの変更、変形、および適応が行われることが可能である。例えば、パケット処理モジュール14の入力バッファ25は、いくつかの充填レベルしきい値がそれぞれに割り当てられた複数の入力バッファに置き換えられることが可能である。例えば、各パケットプロセッサPPに1つのこのような入力バッファがあることが可能である。この場合、バックプレッシャーモジュール29は、このようなすべての入力バッファの充填レベルおよび対応するしきい値によりバックプレッシャーメッセージ30を作り出す。さらに、イングレストラフィック26、処理済みデータパケット32、再循環トラフィック36、およびイーグレストラフィック35を受け取るモジュールの1つまたは複数にリンク

50

バッファが含まれる場合があり、このような各リンクバッファに単純なバックプレッシャーを行うためのそれぞれのオン/オフフロー制御信号を有する。この変更は、装置のより大きいパケットスループット全体に対してモジュールがパケットを処理することができるより高いレートで、このようなモジュールへのデータパケットのパースティングを可能にする。変形は、トラフィック管理モジュール 16 に VOQ 割当て表 48 を含むことである。別の変形は、フロー制御表 46 と VOQ 割当て表 48 を 1 つの表に結合し、一定のフロー制御を適用される必要があるトラフィックフローの VOQ が、バックプレッシャーメッセージのタイプに従って表の行にインデックスすることによって、結合された表から読み取られるようにすることが可能である。例えば、フロー制御表 46 中のトラフィックフロー情報のタイプは、VOQ 割当て表に指定するように、それらのタイプのトラフィックフローを搬送する VOQ の数値識別子に置き換えられることが可能である。この場合、結合された表は、フローマッピングエンジン 20 の中にあることが可能である。

10

【 0 0 2 8 】

前述の例は数個の VOQ 34 を含んでいるだけであるが、一般に高容量のデータ通信システムには数百個のこのような VOQ がある。さらに、細粒度の輻輳レベル検出を提供するために、パケット処理モジュールの入力バッファと関連する多くの充填レベルしきい値がある場合もある。

【 0 0 2 9 】

本発明の諸実施形態の前述の説明から、これらの諸実施形態が、特定のトラフィックフローの特性の 1 つまたは複数に従って特定のトラフィックフローに一定のフロー制御を適用する際に有利に柔軟性を提供することは明らかであろう。

20

【 図 1 】

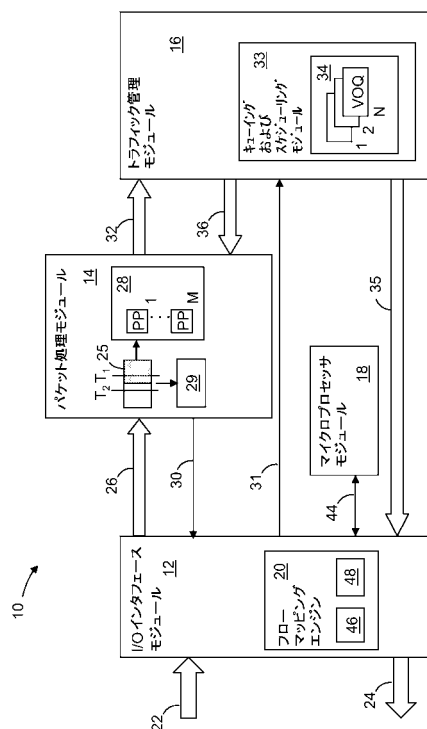


Figure 1

【 図 2 】

46 →

フロー制御表			
優先度による パケットスケジューリング	インGRESS トラフィック	再循環 トラフィック	イーGRESS トラフィック
タイプ 1	なし	なし	なし
タイプ 2	BE	なし	BEサービズ
タイプ 3	BE & EF	DPI	BEサービズ

Figure 2

【図 3】

48 ↗

VOQ 割当て表	
VOQ	トラフィックタイプ
1	再循環トラフィック (DPI)
2	イーグレス(ミラーリング)
3	イーグレス(プレミウムサービス)
4	イーグレス(ベストエフォートサービス)

Figure 3

【図 4】

優先度による パケットレシーケンス:		VOQ フロー制御メッセージ			
		VOQ ₁	VOQ ₂	VOQ ₃	VOQ ₄
タイプ 1	0 1	0	0	0	0
タイプ 2	1 0	0	0	0	1
タイプ 3	1 1	1	0	0	1

Figure 4

【図 5】

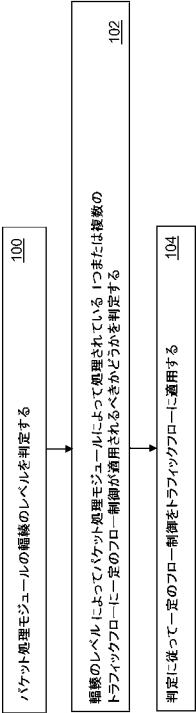


Figure 5

フロントページの続き

審査官 矢頭 尚之

(56)参考文献 特開 2 0 0 0 - 2 4 4 5 2 1 (J P , A)
米国特許第 0 6 9 8 1 0 5 4 (U S , B 1)
特開 2 0 0 0 - 2 0 9 2 5 0 (J P , A)
特開平 1 0 - 2 7 6 2 2 4 (J P , A)

(58)調査した分野(Int.Cl. , D B 名)
H 0 4 L 1 2 / 8 2 5
H 0 4 L 2 9 / 0 8