

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号
特許第4252003号
(P4252003)

(45) 発行日 平成21年4月8日 (2009.4.8)

(24) 登録日 平成21年1月30日 (2009.1.30)

(51) Int.Cl.
H04L 12/56 (2006.01)

F I
H04L 12/56 I O O A

請求項の数 4 (全 20 頁)

(21) 出願番号	特願2004-61912 (P2004-61912)	(73) 特許権者	000005223
(22) 出願日	平成16年3月5日 (2004.3.5)		富士通株式会社
(65) 公開番号	特開2005-252766 (P2005-252766A)		神奈川県川崎市中原区上小田中4丁目1番1号
(43) 公開日	平成17年9月15日 (2005.9.15)	(74) 代理人	100090516
審査請求日	平成18年8月2日 (2006.8.2)		弁理士 松倉 秀実
		(74) 代理人	100113608
			弁理士 平川 明
		(74) 代理人	100105407
			弁理士 高田 大輔
		(74) 代理人	100089244
			弁理士 遠山 勉

最終頁に続く

(54) 【発明の名称】 伝送装置

(57) 【特許請求の範囲】

【請求項 1】

ネットワークを介してパケットを伝送する際に複数の方路への負荷分散を行う伝送装置であって、

受信したパケットから宛先アドレス及び送信元アドレスを抽出するアドレス抽出部と、
各方路から受信される制御情報に基づいて輻輳状況を検出し、この輻輳状況を示す輻輳情報を求める輻輳検出部と、

負荷分散を行う複数の方路に対応した複数のルーティング情報をそれぞれ複数の領域に格納したルーティングテーブルと、

前記ルーティングテーブルの各領域にハッシュ値を割り当て、宛先アドレス及び送信元アドレスのハッシュ値と対応する領域を選択し、この選択した領域のルーティング情報から前記宛先アドレスに対応する方路を索出し、その索出した方路を決定する方路決定部と

、
前記方路決定部で決定した方路にパケットを出力する送信制御部とを備え、
前記方路決定部が、前記輻輳情報に基づいて前記各領域のハッシュ値の割り当てを変更し、

前記パケットを複数の受信ポートから受信し、これと対応する複数の送信ポートから送信する場合に、前記送信制御部が、前記各送信ポートにおいてどの領域のルーティング情報に基づく方路への送信パケットがどれだけ送信されたかを示す情報である送信統計情報を前記輻輳検出部に通知し、

10

20

前記輻輳検出部が、各受信ポートから得られる制御情報を常時又は定期的に監視し、この制御情報に基づいて負荷の高い受信ポートを求め、前記送信統計情報に基づいてこの負荷の高い受信ポートと対応する送信ポートにおいて送信量の多いパケットの方路を求め、この方路を示す情報を輻輳情報として方路決定部に通知することを特徴とする伝送装置。

【請求項 2】

前記方路決定部が、この輻輳情報に示された領域の前記割り当てを減らすことを特徴とする請求項 1 記載の伝送装置。

【請求項 3】

ネットワークを介してパケットを伝送する際に複数の方路への負荷分散を行う伝送装置にて運用される方法であって、

受信したパケットから宛先アドレス及び送信元アドレスを抽出するステップと、

各方路から受信される制御情報に基づいて輻輳状況を検出し、この輻輳状況を示す輻輳情報を求めるステップと、

前記パケットを複数の受信ポートから受信し、これと対応する複数の送信ポートから送信する場合に、

前記各送信ポートにおいてどの領域のルーティング情報に基づく方路への送信パケットがどれだけ送信されたかを示す情報である送信統計情報を通知するステップと、

前記輻輳情報を求めるステップにおいて、各受信ポートから得られる制御情報を監視し、この制御情報に基づいて負荷の高い受信ポートを求め、前記送信統計情報に基づいてこの負荷の高い受信ポートと対応する送信ポートにおいて送信量の多いパケットの方路を求め、この方路を示す情報を輻輳情報として通知するステップと、

負荷分散を行う複数の方路に対応した複数のルーティング情報をそれぞれ複数の領域に格納した各ルーティングテーブルの各領域に割り当てるハッシュ値を前記輻輳情報に基づいて変更するステップと、

前記送信元アドレスと前記宛先アドレスのハッシュ値と前記輻輳情報とに対応した方路をルーティングテーブルから索出し、索出した方路を決定するステップと、

前記決定した方路にパケットを出力するステップと、有ることを特徴とする伝送方法。

【請求項 4】

前記方路を決定するステップにおいて、前記輻輳情報に示された領域の前記割り当てを減らすことを特徴とする請求項 3 記載の伝送方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ネットワークを介してパケットを転送する技術に関する。

例えば、本発明は、IP ネットワーク網における IP パケットルーター装置の IP パケットフォワーディング処理に関する。

【背景技術】

【0002】

近年、IP パケットのトラヒックの急増に伴い、IP ネットワークの拡大やネットワークリソースの効率的な活用、伝送品質の向上が要求されている。

【0003】

従来の技術では、IP ネットワークにおけるトラヒックの経路は、ネットワーク設計により定まっていた。従って IP ネットワーク網のトラヒックが変化し、ネットワークパス間でのバランスが崩れてくると、ネットワーク管理者によるネットワークパス見直しや、再構築の必要が生じていた。

【0004】

図 3 に、従来のトラヒックフローの例を示す。

この図に示す IP スイッチ（ルーター）は、IP ネットワークのユーザ拠点との境界に設けたエッジルータである。

10

20

30

40

50

【0005】

ここで、ルーターαに接続された端末A、B、CからIPネットワークを介してルーターβに接続されている端末Xに対して、大量のデータ送信を実施したとする。この場合、ルーターαでは、宛先アドレス(DA: Destination Address)を用いて方路検索を行い、接続ポートを決定する。従って同じ相手先IPアドレスへの全てのパケットが同じ経路を介してルーターβに伝送される事となる。即ち、ルーターαとルーターβ間には、複数のパスが存在するにも関わらず、1本のパスに対してトラヒックが集中し、パケットの転送遅延の発生、もしくはパケットロスが発生する原因となっていた。

【0006】

そして、このトラヒックの集中による問題を解決するため、同一の相手先アドレスへのパケットを複数ルートに分散して伝送する所謂負荷分散処理(Load balancing 処理)も用いられる様になってきた。

10

【0007】

この負荷分散処理を行ったトラヒックフローを図12に示す。図11と同様にルーターαには端末A、B、Cから端末Xに対しての大量のデータ転送が要求されている。この場合、ルーターαでは負荷分散処理が実施される為に、其々の端末A、B、Cからのパケットは、異なる経路を用いてルーターβに伝送される。つまり、複数のパスに負荷を分散して転送する事により、ネットワーク効率を高めることができる。

【0008】

この負荷分散処理を実現する為には、例えば図13に示す様な構成のルーター90が用いられている。同図に示す様に従来のルーター90は、複数のルーティングテーブルを記憶した記憶部95や、ルーティングテーブルインターフェイス94、ハッシュ演算部93等からなる方路決定部を備えている。

20

【0009】

ルーター90は、パケットを受信すると、IPヘッダ抽出部91にて宛先アドレス(DA)と送信元アドレス(SA: Source Address)を抽出し、ハッシュ演算部93へと渡す。

【0010】

ハッシュ演算部93では、DA及びSAを元にハッシュ演算(例えばCRC演算)を行い、演算結果として求めたハッシュ値をルーティングテーブルインターフェイス94に送信する。該ルーティングテーブルインターフェイス94は、記憶部95のルーティングテーブルを検索して前記ハッシュ値と対応する出力方路情報を求める。

30

【0011】

ルーティングテーブルインターフェイス94により得られた出力方路情報に基づいて送信制御部96は、パケットバッファ92からのパケットを対応する送信ポート97から送信する。

【0012】

本願発明に関連する先行技術として、例えば、下記の特許文献1及び2に開示される技術がある。

【特許文献1】特開2003-218914号公報

【特許文献2】特開平9-270801号公報

40

【発明の開示】

【発明が解決しようとする課題】

【0013】

上述した技術では、入力されるパケットのDA及びSAにのみ基づいて出力方路を求めている為、選択された回線が輻輳状態にある場合も想定できる。

このため、DA及びSAのみに基づいた負荷分散処理では、選択したルートが輻輳状態であってもこの回線にパケットを送出する事になり、輻輳を助長する場合があった。

【0014】

そこで本発明では、回線の輻輳情報を負荷分散処理条件に加えることにより、トラヒックを適切に分散できるルート選択処理を可能とする技術の提供を目的とする。

50

【課題を解決するための手段】

【0015】

本発明の伝送装置は、ネットワークを介してパケットを伝送する際に複数の方路への負荷分散を行う伝送装置であって、

受信したパケットから宛先アドレス及び送信元アドレスを抽出するアドレス抽出部と、
各方路から受信される制御情報に基づいて輻輳状況を検出し、この輻輳状況を示す輻輳情報を求める輻輳検出部と、

前記宛先アドレス、送信元アドレス及び輻輳情報に基づいて方路を決定する方路決定部と、

前記方路決定部で決定した方路にパケットを出力する送信制御部とを備える。

10

前記方路決定部は、送信元アドレスと宛先アドレスのハッシュ値を求め、該ハッシュ値と輻輳情報とに対応した方路をルーティングテーブルから索出しても良い。

【0016】

前記ルーティングテーブルは、負荷分散を行う複数の方路に対応した複数のルーティング情報をそれぞれ複数の領域に格納しており、

前記方路決定部は、該ルーティングテーブルの各領域にハッシュ値を割り当て、前記宛先アドレス及び送信元アドレスのハッシュ値と対応する領域を選択し、この選択した領域のルーティング情報から前記宛先アドレスに対応する方路を索出しても良い。

【0017】

前記方路決定部が、前記輻輳情報に基づいて前記各領域のハッシュ値の割り当てを変更することにより、輻輳状態に応じて選択される領域を変えても良い。

20

【0018】

前記方路決定部は、各方路から受信した輻輳情報に応じ、輻輳が生じている方路の前記割り当てを減らすことで、輻輳している方路へのルーティング量を減らし、負荷バランスをとっても良い。

【0019】

また、前記パケットを複数の受信ポートから受信し、これと対応する複数の送信ポートから送信する場合、

前記送信制御部が、前記各送信ポートにおいてどの領域のルーティング情報に基づく方路への送信パケットがどれだけ送信されたかを示す情報である送信統計情報を前記輻輳検出部に通知し、

30

前記輻輳検出部が、各受信ポートから得られる制御情報を常時又は定期的に監視し、この制御情報に基づいて負荷の高い受信ポートを求め、前記送信統計情報に基づいてこの負荷の高い受信ポートと対応する送信ポートにおいて送信量の多いパケットの方路を求め、この方路を示す情報を輻輳情報として方路決定部に通知しても良い。

【0020】

前記パケットを複数の受信ポートから受信し、これと対応する複数の送信ポートから送信する場合、

前記送信制御部が、前記各送信ポートにおいてどの領域のルーティング情報に基づく方路への送信パケットがどれだけ送信されたかを示す情報である送信統計情報を前記輻輳検出部に通知し、

40

前記輻輳検出部が、各受信ポートから得られる制御情報を常時又は定期的に監視し、この制御情報に基づいて負荷の高い受信ポートを求め、前記送信統計情報に基づいてこの負荷の高い受信ポートと対応する送信ポートにおいて送信量の多いパケットの方路を索出したルーティングテーブルの領域を求め、この領域を示す情報を輻輳情報として方路決定部に通知し、

前記方路決定部が、この輻輳情報に示された領域の前記割り当てを減らすことで、輻輳している方路へのルーティング量を減らし、負荷バランスをとっても良い。

【0021】

また、本発明の伝送方法は、

50

ネットワークを介してパケットを伝送する際に複数の方路への負荷分散を行う伝送装置にて運用される際、

受信したパケットから宛先アドレス及び送信元アドレスを抽出するステップと、

各方路から受信される制御情報に基づいて輻輳状況を検出し、この輻輳状況を示す輻輳情報を求めるステップと、

前記宛先アドレス、送信元アドレス及び輻輳情報に基づいて方路を決定するステップと、

前記決定した方路にパケットを出力するステップを含む。

【0022】

前記方路を決定するステップにおいて、送信元アドレスと宛先アドレスのハッシュ値を求め、該ハッシュ値と輻輳情報とに対応した方路をルーティングテーブルから索出しても良い。

【0023】

また、前記伝送方法において、前記ルーティングテーブルが、負荷分散を行う複数の方路に対応した複数のルーティング情報をそれぞれ複数の領域に格納しており、前記各ルーティングテーブルの各領域にハッシュ値を割り当てておき、

前記方路を決定するステップにおいて、前記宛先アドレス及び送信元アドレスのハッシュ値と対応する領域を選択し、この選択した領域のルーティング情報から宛先アドレスに対応する方路を索出しても良い。

【0024】

前記方路を決定するステップにおいて、前記輻輳情報に基づき前記各領域のハッシュ値の割り当てを変更することにより、輻輳状態に応じて選択される領域を変えても良い。

【0025】

また、前記パケットを複数の受信ポートから受信し、これと対応する複数の送信ポートから送信する場合、

前記各送信ポートにおいてどの領域のルーティング情報に基づく方路への送信パケットがどれだけ送信されたかを示す情報である送信統計情報を前記輻輳検出部に通知するステップを更に備え、

前記輻輳情報を求めるステップにおいて、各受信ポートから得られる制御情報を常時又は定期的に監視し、この制御情報に基づいて負荷の高い受信ポートを求め、前記送信統計情報に基づいてこの負荷の高い受信ポートと対応する送信ポートにおいて送信量の多いパケットの方路を求め、この方路を示す情報を輻輳情報として方路決定部に通知しても良い。

【0026】

前記パケットを複数の受信ポートから受信し、これと対応する複数の送信ポートから送信する場合、

前記各送信ポートにおいてどの領域のルーティング情報に基づく方路への送信パケットがどれだけ送信されたかを示す情報である送信統計情報を前記輻輳検出部に通知するステップを更に備え、

前記輻輳情報を求めるステップにおいて、各受信ポートから得られる制御情報を常時又は定期的に監視し、この制御情報に基づいて負荷の高い受信ポートを求め、前記送信統計情報に基づいてこの負荷の高い受信ポートと対応する送信ポートにおいて送信量の多いパケットの方路を索出したルーティングテーブルの領域を求め、この領域を示す情報を輻輳情報として方路決定部に通知し、

前記方路を決定するステップにおいて、この輻輳情報に示された領域の前記割り当てを減らすことで、輻輳している方路へのルーティング量を減らし、負荷バランスをとっても良い。

【0027】

また、本発明のネットワークシステムは、ネットワークを介してパケットを伝送する際

10

20

30

40

50

に複数の方路への負荷分散を行う伝送装置をノードとして含み、

該伝送装置が、

受信したパケットから宛先アドレス及び送信元アドレスを抽出するアドレス抽出部と、

各方路から受信される制御情報に基づいて輻輳状況を検出し、この輻輳状況を示す輻輳情報を求める輻輳検出部と、

前記宛先アドレス、送信元アドレス及び輻輳情報に基づいて方路を決定する方路決定部と、

前記方路決定部で決定した方路の次のノードにパケットを出力する送信制御部とを備える。

【0028】

10

図3は、本発明の原理説明図である。同図に示す伝送装置10は、アドレス抽出部16や、輻輳検出部25、方路決定部17、送信制御部18を備えている。

【0029】

加入者側からパケットが入力されると伝送装置10は、アドレス抽出部16にて該パケットのヘッダ情報から方路検索で使用する宛先アドレス及び送信元アドレスの情報を抽出する。パケットデータとの関連付けを行う為にフローIDを付加し、方路決定部17へ通知する。伝送装置10は、方路検索結果が求まるまでの間、パケットデータを方路決定部17に渡したアドレス情報と同一のフローIDと共に、パケットバッファ6に格納させる。

【0030】

20

方路決定部17には、負荷分散に応じたルーティング情報を格納した複数の領域を有するルーティングテーブルを記憶部5に記憶している。

【0031】

輻輳検出部25は、送信制御部が送信ポート22から各方路へ出力したパケットの送信統計情報と、ポーズ監視部8が各受信ポート14にて受信される制御情報（バックプレッシャ情報）とに基づいて各方路の輻輳状況を求め、この輻輳状況に応じた輻輳情報を方路決定部17に通知する。

【0032】

方路決定部17は、テーブル選択部15が、DA及びSAのハッシュ演算値と、輻輳情報とに基づいて方路検索に用いるルーティングテーブルの領域を決定し、この領域のルーティング情報についてDAによる方路検索を行い、当該パケットを出力する方路を決定し、この方路へ出力するための出力方路情報と前記フローIDを送信制御部18に通知する。

30

【0033】

送信制御部18のスケジューラ24は、パケットバッファ6を介して共有メモリ23に格納されたパケットデータのうち前記フローIDと対応するパケットデータを前記出力方路情報に基づいて対応する送信ポート22から送信する。

このように本発明の構成であれば、以下の2つのポイントにより、課題解決の作用が期待できる。

【0034】

40

(1) 図3の方路決定部に示す如く、IPヘッダ情報のDA情報だけで方路検索動作を行うのではなく、DA、SAに対するハッシュ演算値を用い、複数の前記領域の中から使用する領域を選択し、この選択した領域のルーティング情報によって出力方路検索を行っている。従って、IPネットワーク上の一つのサーバーに対して、複数の端末より一斉にアクセスした場合においても、全てのパケットデータが同一のパスを使用する事を避け、複数のパスに振り分けて通信すること（負荷分散）が可能となる。

【0035】

(2) 方路決定部による方路検索処理は、受信ポートでのバックプレッシャ計測値と送信ポートでの送信統計情報に基づく輻輳情報によって選択される領域に基づくものであり、実際のネットワーク負荷に基づく負荷分散処理を行う事が可能となる。

50

【発明の効果】

【0036】

本発明は、回線の輻輳情報を負荷分散処理条件に加えることにより、トラヒックを適切に分散できるルート選択処理を可能とする技術を提供できる。

【発明を実施するための最良の形態】

【0037】

以下、図面を参照して本発明を実施するための最良の形態について説明する。以下の実施の形態の構成は例示であり、本発明は実施の形態の構成に限定されない。

図1は、本発明に係る伝送装置（IPネットワークにおけるIPスイッチ、例えばルーター）を用いたネットワークシステムの概略構成図、図2はこのパケット伝送システムの機能ブロック図である。

【0038】

この図に示す伝送装置10、20は、IPネットワークのユーザ拠点との境界に設けたエッジルータである。

【0039】

ここで、ルーター10に接続された端末2a、2b、2cからIPネットワークを介してルーター20に接続されている端末2xに対して、大量のパケットを送信した場合について説明する。この場合、ルーター10は、従来から用いられるIPヘッダー中の宛先アドレス(DA:Destination Address)によりルーティングテーブルを検索して、送信ポートを決定するフォワーディング機能に加え、パケットのIPヘッダーに含まれる宛先アドレス(DA)及び送信元アドレス(SA:Source Address)に対する演算値、ならびに回線輻輳状態に基づく輻輳情報を用いて、ルーティングテーブルに含まれる複数の領域から前記検索に用いる領域を選択する機能を有する。これによりルーター10は、各回線の負荷バランスを図りながら出力方路決定を行う。

【0040】

図4は、本発明に係る伝送装置10の要部説明図であり、パケットを最大8方路へ選択的に転送することにより負荷分散を行うIPフォワーディング処理部分の機能ブロックを示している。なお、伝送装置20については、伝送装置10と同一の構成であるので説明を省略する。

【0041】

同図に示すように、この伝送装置10は、IPフォワーディング処理のため、テーブル検索部（方路決定部に相当）17、パケットバッファ6、ヘッダ抽出部（アドレス抽出部に相当）16、送信制御部18、受信制御部19、輻輳検出部25を備えている。

【0042】

この受信制御部19は、受信バッファ13や、ポーズ監視部8を有している。

受信バッファ13-1～13-nは、複数の受信ポート14-1～14-nで各方路から受信したパケットをそれぞれ格納している。

【0043】

ポーズ監視部8は、各方路から受信される制御情報（バックプレッシャ情報）を常時又は定期的に監視し、輻輳検出部25に通知する。例えば、輻輳が発生して滞留パケットが増えた場合、対向装置（各方路においてリンクを挟んで装置10と対向する次のノード）は、ポーズパケットを伝送装置10に送信して送信の一次停止を要求する。ポーズ監視部8は、この送信停止要求を受信する時間（フロー制御時間）・頻度をポート毎に監視して統計をとり、この制御情報に基づく情報（制御統計情報）を定期的に輻輳検出部25へと通知する。

【0044】

ヘッダ抽出部16は、前記受信バッファ13を介して受信したパケットのヘッダ情報から宛先アドレス及び送信元アドレスを抽出する。

輻輳検出部25は、後述の如くポーズ監視部8からの制御統計情報と、送信制御部18からの送信統計情報とに基づいて輻輳情報を求め、この輻輳情報をテーブル検索部17に

通知する。この輻輳情報は、後述の如く、領域判定部 15 における閾値変更情報として用いられる。

【0045】

テーブル検索部 17 は、前記宛先アドレス、送信元アドレス及び輻輳情報に基づいて方路を決定する。このとき、テーブル検索部 17 のハッシュ演算部 21 が、送信元アドレスと宛先アドレスのハッシュ値を求め、領域判定部 15 がこのハッシュ値と輻輳情報とに対応した方路をルーティングテーブルから索出する。このルーティングテーブルは、記憶部 5 に記憶されており、負荷分散を行う方路に応じた複数のルーティング情報をそれぞれ複数（本例では 8 個）の領域に格納している。この複数の領域は、一つのルーティングテーブルが有する複数の領域として記憶されているものでも良いし、独立して複数のテーブルとして記憶されているものでも良い。また、この複数の領域は、物理的に特定の位置に存在するものに限らず、アドレスやラベル、格納の順序等によって一連の情報であること（どの領域の情報であるか）が特定できるように格納されたものでも良い。

10

【0046】

領域判定部 15 は、この各領域に予めハッシュ値の範囲を割り当て、前記ハッシュ演算部 21 が求めたハッシュ値と対応する領域を選択し、この選択した領域のルーティング情報を参照して前記宛先アドレスと対応する方路を索出する。このとき、領域判定部 15 は、輻輳情報に基づいて前記各領域のハッシュ値の割り当てを変更することにより、輻輳状態に応じて選択される領域を変えるようにしている。これにより、輻輳状態に応じた負荷分散が行われるようにしている。例えば、輻輳が生じた方路に係る領域の割り当てを減らし、輻輳が生じていない方路に係る領域の割り当てを増やすことにより、輻輳している方路へのルーティング量を減らし、負荷バランスをとる。

20

【0047】

パケットバッファ 6 は、前記テーブル検索部 17 でテーブル検索を行い、情報を得る間パケットを待たせる為のパバッファである。

【0048】

図 5 は、本実施形態の伝送装置によるパケット伝送の手順（伝送方法）の説明図である。

伝送装置 10 は、対向装置から送られたパケットを受信ポート 14-1～14-n で受信し（ステップ 1、以下 S1 のように略記する）、これが転送対象のパケットであれば（S2）、このパケットデータを受信バッファ 13-1～13-n に格納し（S3）、このパケットデータから IP ヘッダ抽出部にて IP ヘッダ情報（DA, SA）を抽出する（S4）。このとき伝送装置 10 は、抽出元のパケットデータとの関連付けを行う為に該 IP ヘッダ情報にフロー ID を付加し、この IP ヘッダ情報をテーブル検索部 17 へ通知する（S5）。また伝送装置 10 は、方路検索結果が求まるまでの間、抽出元のパケットデータをテーブル検索部 17 に渡したアドレス情報と同一のフロー ID と共にパケットバッファ 6 に格納させる（S6）。

30

【0049】

一方、ポーズ監視部 8 は、ステップ 2 で受信したパケットがポーズパケットか否かを監視し、ポーズパケットであれば制御統計情報を求めてテーブル検索部 17 に通知する（S7）。輻輳検出部 25 は、この制御統計情報と送信制御部からの送信統計情報とに基づいて輻輳情報を求め、テーブル検索部 17 に通知する（S8）。

40

【0050】

また、テーブル検索部 17 のハッシュ演算部は、前記 DA と SA のハッシュ値を求め、領域判定部 15 に通知する（S9）。領域判定部 15 は、この DA 及び SA のハッシュ演算値と、輻輳情報とに基づいてルーティングテーブルの領域を選択する（S10）。そして、インターフェイス 4 を介してこの選択した領域のルーティング情報を参照し、DA による方路検索を行い、当該パケットを出力する方路を決定し、この方路へ出力するための情報（出力方路情報）と前記ヘッダ情報に付加されていたフロー ID を送信制御部 18 に通知する（S11）。一方、これと同時にパケットバッファ 6 からパケットデータが送信

50

制御部 18 に入力される (S 12)。

【0051】

送信制御部 18 は、インターフェイス 9 を介してフロー ID が一致する出力方路情報とパケットデータとを対応付けて共有メモリ 23 に保存する (S 13)。

【0052】

送信制御部 18 のスケジューラ 11 は、共有メモリ 23 に格納されたパケットデータに対応する出力方路情報に示される送信ポート 22 から送信する (S 14)。このときスケジューラ 11 による制御は、特に限定されるものではなく、既知のスケジューラ制御 (例えば、RR: Round Robin, WRR: Weighted Round Robin) を用いてパケットを送信バッファ 12-1 ~ 12-n に振り分け、送信ポート 22-1 ~ 22-n からそれぞれ送信する。

10

【0053】

また、送信制御部 18 は、各送信ポートにおいてどの領域のルーティング情報に基づく方路への送信パケットがどれだけ送信されたかを示す情報である送信統計情報を求め、これを領域判定部 15 へ通知する (S 15)。

【0054】

以上説明した流れにおける主要な機能の詳細について以下で説明する。

(I) テーブル検索部の機能。

図 6 は、テーブル検索部 17 の説明図である。

記憶部 5 に記憶されたルーティングテーブルの各領域 a1 ~ a8 には、それぞれ負荷分散のために異ならせたルーティング情報を蓄えられており、テーブル検索部 17 は、この領域 a1 ~ a8 の何れかから選択的に DA に対する方路を検索する事により、出力すべき方路、即ち出力方路情報を得る事が出来る。

20

【0055】

このテーブルを検索する為のキー情報 (RAM のアドレス情報) は、領域判定部 15 と、IP アドレステーブル 3 により生成される。

IP アドレステーブル 3 は、IP ヘッダー中の宛先アドレス (DA) と、この DA に係る出力方路情報を記憶しているルーティングテーブル上のポインタ値とを対応付けて記憶している。なお、本実施形態の IP アドレステーブル 3 は、CAM (Content Addressable Memory) を用いて構成している。

【0056】

30

本実施形態のテーブル検索部 17 では、DA とルーティング情報のフレキシブルな関連付けを行う事ができるようにしている。

【0057】

一般的には、DA に対応したポインタ値を用いてルーティングテーブルを検索する事により出力方路情報、付加ラベル情報を得るが、その場合、DA に対して 1:1 のルーティング情報となり、負荷分散処理ができない。

【0058】

そこで、本構成では、領域判定部 15 での判定結果も用いて方路検索を行う。領域判定部 15 は、ハッシュ演算部 21 から宛先アドレス (DA) 及び送信元アドレス (SA) のハッシュ値を受け、これに基づいて領域判定を行い、0h から 7h までの 8 通りの判定結果を出力する。

40

【0059】

この領域判定部 15 における領域判定の仕組みを図 7 に示す。

図 7 (A) に示すように、ハッシュ演算部 21 では、入力される宛先アドレス (DA) 32 bit と送信元アドレス (SA) 32 bit の計 64 bit について、ハッシュ演算 (CRC-16 演算) を行い、16 bit の演算結果 (0 ~ 65535 / 10 進) を得る。なお、本実施形態では、ハッシュ演算として CRC-16 を行い 16 bit の演算結果を得たが、ハッシュ演算に CRC-8 等を用い、異なる精度の結果を得るようにしても良い。この精度は、負荷分散する方路数等に応じて任意に設定できる。また、演算方法として CRC 演算を用いず、SA 32 bit の MSB から LSB の文字の並びを逆にしたもの

50

と、D A 3 2 b i t とで E O R をとって 3 2 b i t の値にし、更に上位 1 6 b i t と下位 1 6 b i t との E O R をとって 1 6 b i t の値を得る等、他の演算方法を採用する事も可能である。

【0060】

これらの演算により得られた値は、領域判定部 1 5 にて 0 h ~ 7 h のどの領域に属する値かを判定する。

この領域判定部 1 5 では、0 ~ 6 5 5 3 5 までの値（即ち 1 6 b i t ）を 8 つの範囲 X に分割する 7 つの閾値 A ~ G を設定（記憶）している。領域判定部 1 5 は、この 7 つの閾値 A ~ G を、後述する送信制御部、受信制御部から通知される情報に基づいて変化させることで、この範囲 X を変更する。

10

【0061】

領域判定部 1 5 による判定は、図 7（B）の表に示す通り、ハッシュ演算部 2 1 からのハッシュ演算値が、どの範囲 X に属するかを求め、求めた範囲の判定値を出力する。即ち、ハッシュ演算値が、0 以上閾値 A 未満であれば判定値 0 h を出力し、閾値 A 以上閾値 B 未満であれば 1 h を出力し、・・・閾値 G 以上 6 5 5 3 5 以下であれば 7 h を出力する。領域判定部 1 5 は、この判定値に基づいて方路検索に用いるテーブルを選択する。例えば、判定値 0 h ~ 7 h とルーティングテーブルの領域 a 1 ~ a 8 とがそれぞれ対応付けられている場合、即ち領域 a 1 ~ a 8 にそれぞれ前記ハッシュ値の範囲 X が割り当てられている場合、ハッシュ演算値と対応する領域が選択される。

20

【0062】

本実施形態において、記憶部 5 のルーティングテーブルは、図 1 0 に示すように D A に対する出力方路情報が複数の領域にそれぞれ記憶されている。例えば、ある D A に対する出力方路情報が # 1 - 1 ~ # 1 - 8 の 8 個記憶されており、このうち # 1 - 1 が領域 a 1 の出力方路情報、# 1 - 2 が領域 a 2 の出力方路情報、・・・# 1 - 8 が領域 a 8 の出力方路情報となっている。

【0063】

そしてテーブル検索部 1 7 は、I P アドレステーブル 3 から得た D A に対するポインタ値に前記判定値を加えてアドレス情報を生成し、記憶部 5 のルーティングテーブルを検索してこのアドレス情報に対応する出力方路情報を得る。テーブル検索部 1 7 は、この結果得られた出力方路情報と判定値を該 D A のフロー情報と共に送信制御部 1 8 に通知する。

30

【0064】

（II）送信制御部の機能

図 8 に送信制御部 1 8 の構成を示す。

送信制御部 1 8 の共有メモリアンタフェイス部 9 では、テーブル検索部 1 7 から通知されるフロー I D に対応するパケットデータをパケットバッファ 6 から読み出し、共有メモリに書き込む。一方、該インタフェイス部 9 は、出力方路情報とエリア判定情報（判定値）をスケジューラ部 2 4 に通知する。通知を受けたスケジューラ部 2 4 は、既知の技術であるスケジューリング手法により対応する送信バッファ 1 2 - 1 ~ 1 2 - n にそれぞれ、共有メモリ 2 3 のパケットデータをエリア判定情報と共に転送する。なお、大量のデータを扱う場合、この部分をクロスバースイッチ等の大規模スイッチにて構成する事も可能である。

40

【0065】

送信バッファ 1 2 - 1 ~ 1 2 - n は、それぞれ送信ポート 2 2 - 1 ~ 2 2 - n に対応している。パケットデータは、この送信バッファ 1 2 - 1 ~ 1 2 - n を介して、各ポート 2 2 - 1 ~ 2 2 - n に出力される。また、各送信バッファ 1 2 - 1 ~ 1 2 - n では、パケットデータと共に通知されるエリア判定情報を終端し、送出するパケットがどの領域のルーティングテーブルの検索によるものかの統計処理を行う。各送信ポートの統計処理結果は、輻輳検出部 2 5 へと通知される。

【0066】

（III）受信制御部の機能

50

受信制御部 19 では、図 4 に示す通り、各受信ポートにて検出される対向装置からのフロー制御時間を監視・計測するポーズ監視部 8 を持つ。

【0067】

対向装置からは、回線が輻輳した滞留パケットが増えた場合、ポーズパケットを用いて送信の一次停止を要求してくる。ポーズ監視部 8 では、この送信停止要求を受信する時間・頻度をポート毎に統計をとり監視する。得られた情報は、定期的に輻輳検出部 25 へと通知される。

【0068】

(IV) 輻輳検出部における輻輳検出処理

図 9 に示す通り、送信制御部 18 は、送信パケットと共に転送されてくるエリア判定情報を元に、周期的にどのテーブル領域にてルート検索されたパケットが多いか、即ち各送信ポートにおける前記領域別のパケット数の情報（送信統計情報）を輻輳検出部 25 に通知する。

【0069】

また、受信制御部 19 は、回線輻輳時に受信されるポーズパケットにより対向装置から制御されるポーズ時間を周期的に監視し、各受信ポートにおける単位時間当たりのポーズ制御発生頻度の情報（制御統計情報）を輻輳検出部 25 に通知する。

【0070】

そして輻輳検出部 25 では、これら送信統計情報と制御情報とから輻輳状況を認識する。つまり、輻輳検出部 25 は、送信統計情報からポーズ制御発生頻度の高い受信ポートを検出すると、対応する送信ポート（同じ対向装置と接続しているポート）において送信パケット量の多い領域を求める。これによりこの領域に基づく方路に輻輳が発生していることが認識できるので、この領域を示す情報を輻輳情報としてテーブル検索部 17 の領域判定部 15 に通知する。

【0071】

(V) 領域判定部内での閾値処理

領域判定部 15 は、輻輳検出部 25 から輻輳情報を受信した場合、この輻輳状態に示される領域を狭めるように閾値を変更する。即ち、該領域のハッシュ値の割り当てを減らし、該領域が選択される確率を減少させる。この制御により、他の領域の割り当てが増える、即ち他の領域が選択される確率が増えることになり、輻輳した方路へのパケットが他の方路へルーティングされる。この結果、負荷の高いポートからのパケットの送信を減らし、それ以外のポートへのパケット送信を増やすことができる。

【0072】

以上のように本実施形態によれば、IP フォワーディング処理において、IP ヘッダー（SA, DA）と輻輳情報を用い、回線輻輳状態を加味した複数方路へのトラヒックバランシング処理（負荷分散処理）が可能となる。

【0073】

ここで言う、回線輻輳状態を加味したバランシング処理とは、各ポートに対応するネットワーク上のバックプレッシャ発生頻度監視を行い、輻輳情報をテーブル検索部にフィードバックする事によるものである。即ち、輻輳が検出されたポートへの送信を単に停止し他のポートからの送信を行うのでは無く、回線の輻輳状況に応じてテーブル検索部で輻輳回線が選択される確率を落とす事により、パケットを送信する方路を制御するものである。これにより刻々と変化する回線負荷を考慮した負荷分散処理を伴うフォワーディング処理を行う事が可能となり、ネットワークの効率的な使用（各ネットワークパス間でのトラヒックの均質化）を可能とする。これにより、IP ネットワークのパフォーマンス向上に寄与するところが多い。

【0074】

従って、回線の輻輳状態監視結果情報を負荷分散処理条件に加える事により、複数のネットワークパスのトラヒック負荷を分散し、転送効率を上げつつネットワークの負荷変動に追従した柔軟なルート選択処理を行う事が可能となる。

【0075】

〈その他の実施形態〉

本発明は、上述の図示例にのみ限定されるものではなく、本発明の要旨を逸脱しない範囲内において種々変更を加え得ることは勿論である。

【0076】

例えば、以下に付記した構成であっても上述の実施形態と同様の効果が得られる。また、これらの構成要素は可能な限り組み合わせることができる。

【0077】

(付記1)

ネットワークを介してパケットを伝送する際に複数の方路への負荷分散を行う伝送装置 10
であって、

受信したパケットから宛先アドレス及び送信元アドレスを抽出するアドレス抽出部と、
各方路から受信される制御情報に基づいて輻輳状況を検出し、この輻輳状況を示す輻輳
情報を求める輻輳検出部と、

前記宛先アドレス、送信元アドレス及び輻輳情報に基づいて方路を決定する方路決定部
と、

前記方路決定部で決定した方路にパケットを出力する送信制御部とを備える伝送装置。

(1)

【0078】

(付記2)

前記方路決定部が、送信元アドレスと宛先アドレスのハッシュ値を求め、該ハッシュ値
と輻輳情報とに対応した方路をルーティングテーブルから索出する付記1に記載の伝送装
置。(2)

【0079】

(付記3)

前記ルーティングテーブルが、負荷分散を行う複数の方路に対応した複数のルーティン
グ情報をそれぞれ複数の領域に格納しており、

前記方路決定部が、該ルーティングテーブルの各領域にハッシュ値を割り当て、前記宛
先アドレス及び送信元アドレスのハッシュ値と対応する領域を選択し、この選択した領域
のルーティング情報から前記宛先アドレスに対応する方路を索出する付記2に記載の伝送
装置。(3)

【0080】

(付記4)

前記方路決定部が、前記輻輳情報に基づいて前記各領域のハッシュ値の割り当てを変更
することにより、輻輳状態に応じて選択される領域を変える付記3に記載の伝送装置。

【0081】

(付記5)

前記方路決定部が、各方路から受信した輻輳情報に応じ、輻輳が生じている方路の前記
割り当てを減らすことで、輻輳している方路へのルーティング量を減らし、負荷バランス
をとる付記4に記載の伝送装置。

【0082】

(付記6)

前記パケットを複数の受信ポートから受信し、これと対応する複数の送信ポートから送
信する場合に、

前記送信制御部が、前記各送信ポートにおいてどの領域のルーティング情報に基づく方
路への送信パケットがどれだけ送信されたかを示す情報である送信統計情報を前記輻輳検
出部に通知し、

前記輻輳検出部が、各受信ポートから得られる制御情報を常時又は定期的に監視し、こ
の制御情報に基づいて負荷の高い受信ポートを求め、前記送信統計情報に基づいてこの負
荷の高い受信ポートと対応する送信ポートにおいて送信量の多いパケットの方路を求め、

10

20

30

40

50

この方路を示す情報を輻輳情報として方路決定部に通知する付記 5 に記載の伝送装置。

【0083】

(付記 7)

前記パケットを複数の受信ポートから受信し、これと対応する複数の送信ポートから送信する場合に、

前記送信制御部が、前記各送信ポートにおいてどの領域のルーティング情報に基づく方路への送信パケットがどれだけ送信されたかを示す情報である送信統計情報を前記輻輳検出部に通知し、

前記輻輳検出部が、各受信ポートから得られる制御情報を常時又は定期的に監視し、この制御情報に基づいて負荷の高い受信ポートを求め、前記送信統計情報に基づいてこの負荷の高い受信ポートと対応する送信ポートにおいて送信量の多いパケットの方路を索出したルーティングテーブルの領域を求め、この領域を示す情報を輻輳情報として方路決定部に通知し、

10

前記方路決定部が、この輻輳情報に示された領域の前記割り当てを減らすことで、輻輳している方路へのルーティング量を減らし、負荷バランスをとる付記 4 に記載の伝送装置。

【0084】

(付記 8)

ネットワークを介してパケットを伝送する際に複数の方路への負荷分散を行う伝送装置にて運用される方法であって、

20

受信したパケットから宛先アドレス及び送信元アドレスを抽出するステップと、
各方路から受信される制御情報に基づいて輻輳状況を検出し、この輻輳状況を示す輻輳情報を求めるステップと、

前記宛先アドレス、送信元アドレス及び輻輳情報に基づいて方路を決定するステップと、

前記決定した方路にパケットを出力するステップを含む伝送方法。(4)

【0085】

(付記 9)

前記方路を決定するステップにおいて、送信元アドレスと宛先アドレスのハッシュ値を求め、該ハッシュ値と輻輳情報とに対応した方路をルーティングテーブルから索出する付記 8 に記載の伝送方法。

30

【0086】

(付記 10)

前記ルーティングテーブルが、負荷分散を行う複数の方路に対応した複数のルーティング情報をそれぞれ複数の領域に格納しており、前記各ルーティングテーブルの各領域にハッシュ値を割り当てておき、

前記方路を決定するステップにおいて、前記宛先アドレス及び送信元アドレスのハッシュ値と対応する領域を選択し、この選択した領域のルーティング情報から宛先アドレスに対応する方路を索出する付記 9 に記載の伝送方法。

40

【0087】

(付記 11)

前記方路を決定するステップにおいて、前記輻輳情報に基づき前記各領域のハッシュ値の割り当てを変更することにより、輻輳状態に応じて選択される領域を変える付記 10 に記載の伝送方法。

【0088】

(付記 12)

前記パケットを複数の受信ポートから受信し、これと対応する複数の送信ポートから送信する場合に、

前記各送信ポートにおいてどの領域のルーティング情報に基づく方路への送信パケット

50

がどれだけ送信されたかを示す情報である送信統計情報を前記輻輳検出部に通知するステップを更に備え、

前記輻輳情報を求めるステップにおいて、各受信ポートから得られる制御情報を常時又は定期的に監視し、この制御情報に基づいて負荷の高い受信ポートを求め、前記送信統計情報に基づいてこの負荷の高い受信ポートと対応する送信ポートにおいて送信量の多いパケットの方路を求め、この方路を示す情報を輻輳情報として方路決定部に通知する付記 11 に記載の伝送方法。

【0089】

(付記 13)

前記パケットを複数の受信ポートから受信し、これと対応する複数の送信ポートから送信する場合に、

前記各送信ポートにおいてどの領域のルーティング情報に基づく方路への送信パケットがどれだけ送信されたかを示す情報である送信統計情報を前記輻輳検出部に通知するステップを更に備え、

前記輻輳情報を求めるステップにおいて、各受信ポートから得られる制御情報を常時又は定期的に監視し、この制御情報に基づいて負荷の高い受信ポートを求め、前記送信統計情報に基づいてこの負荷の高い受信ポートと対応する送信ポートにおいて送信量の多いパケットの方路を索出したルーティングテーブルの領域を求め、この領域を示す情報を輻輳情報として方路決定部に通知し、

前記方路を決定するステップにおいて、この輻輳情報に示された領域の前記割り当てを減らすことで、輻輳している方路へのルーティング量を減らし、負荷バランスをとる付記 10 に記載の伝送方法。

【0090】

(付記 14)

ネットワークを介してパケットを伝送する際に複数の方路への負荷分散を行う伝送装置をノードとして含むネットワークシステムであって、

前記伝送装置が、

受信したパケットから宛先アドレス及び送信元アドレスを抽出するアドレス抽出部と、
各方路から受信される制御情報に基づいて輻輳状況を検出し、この輻輳状況を示す輻輳情報を求める輻輳検出部と、

前記宛先アドレス、送信元アドレス及び輻輳情報に基づいて方路を決定する方路決定部と、

前記方路決定部で決定した方路の次のノードにパケットを出力する送信制御部とを備えるネットワークシステム。(5)

【図面の簡単な説明】

【0091】

【図1】本発明に係るネットワークシステムの概略構成図

【図2】ネットワークシステムの機能ブロック図

【図3】本発明の原理説明図

【図4】伝送装置の機能を示すブロック図

【図5】伝送方法の説明図

【図6】テーブル検索部の説明図

【図7】領域判定部における領域判定の説明図

【図8】送信制御部の説明図

【図9】輻輳検出の説明図

【図10】ルーティングテーブルの説明図

【図11】従来のIPネットワークの説明図

【図12】従来の負荷分散処理の説明図

【図13】従来のルータの構成を示すブロック図

【符号の説明】

10

20

30

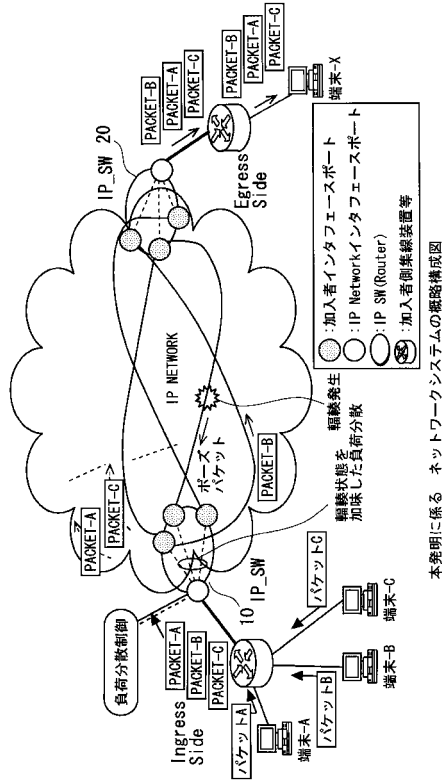
40

50

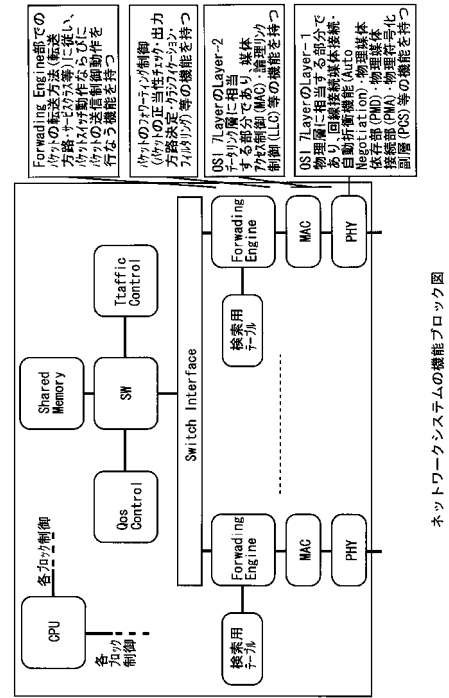
【0092】

2 a, 2 b, 2 c, 2 x	端末	
3	アドレステーブル	
4	インターフェイス	
5	記憶部	
6	パケットバッファ	
8	ポーズ監視部	
9	共有メモリアンターフェイス部	
10, 20	伝送装置（ルーター）	
11	スケジューラ	10
12 (12-1~12-n)	送信バッファ	
13 (13-1~13-n)	受信バッファ	
14 (14-1~14-n)	受信ポート	
15	領域判定部	
16	ヘッダ抽出部（アドレス抽出部）	
17	テーブル検索部（方路決定部）	
18	送信制御部	
19	受信制御部	
21	ハッシュ演算部	
22 (22-1~22-n)	送信ポート	20
22 a	送信ポート	
22	送信ポート	
23	共有メモリ	
24	スケジューラ部	
25	輻輳検出部	

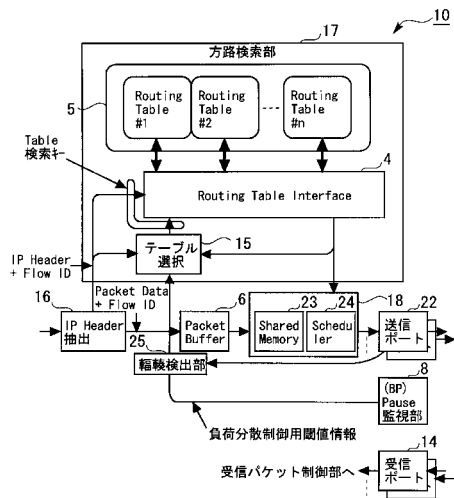
【図 1】



【図 2】

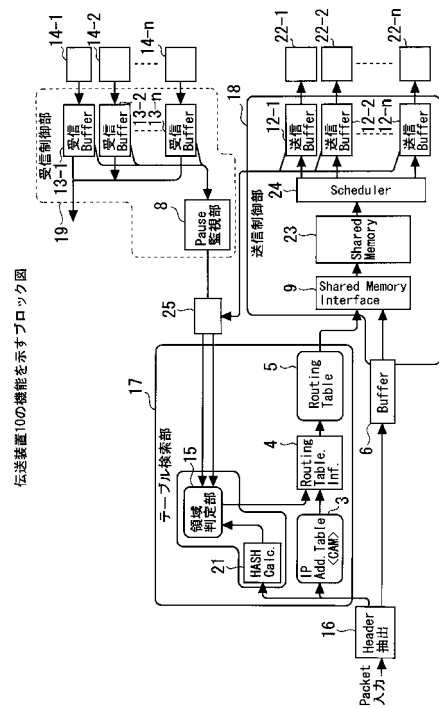


【図 3】

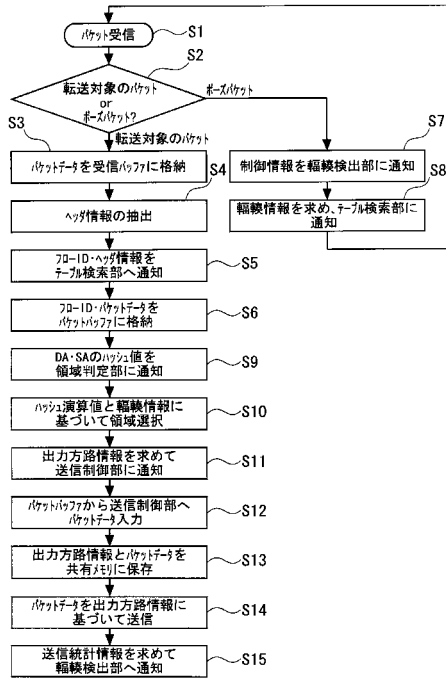


本発明の原理説明図

【図 4】

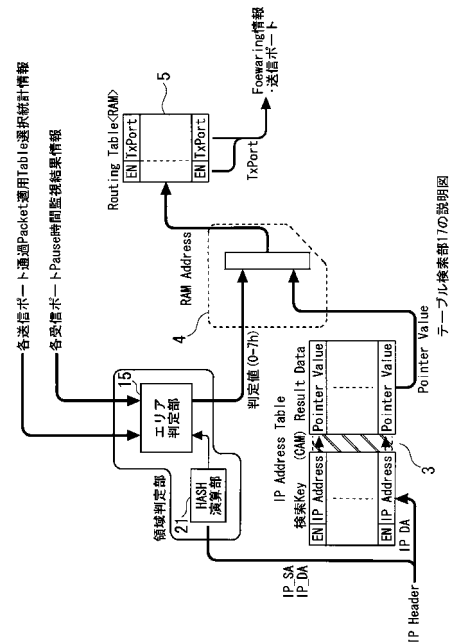


【図 5】



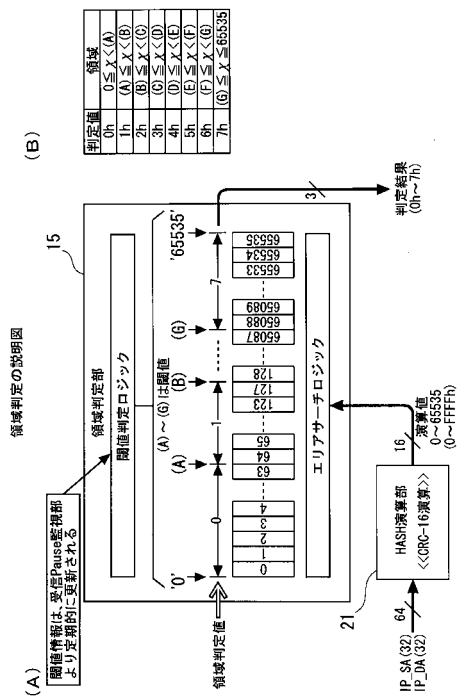
伝送方法の説明図

【図 6】

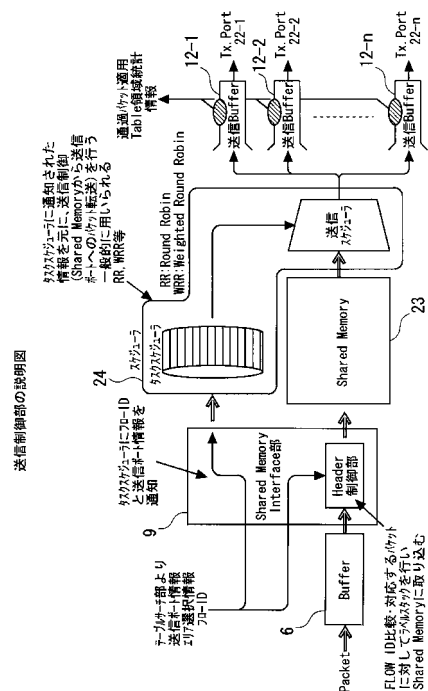


テーブル検索部17の説明図

【圖 7】

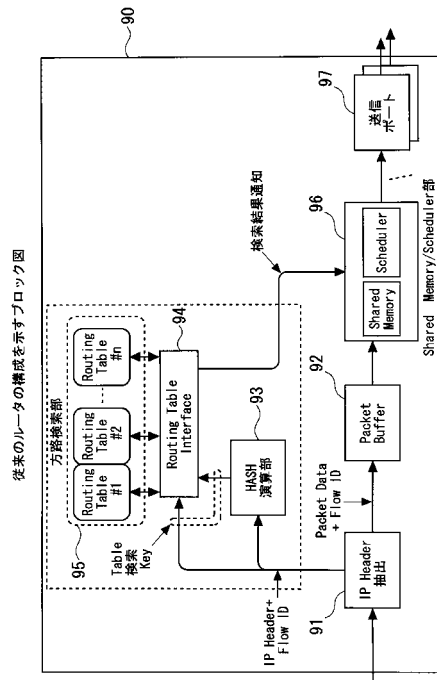


【图 8】



FLOW ID比較・対応するバケット
に対してパルスタックを行い
Shared Memoryに取り込む

【図 13】



フロントページの続き

- (72)発明者 河村 英俊
神奈川県横浜市港北区新横浜三丁目9番18号 富士通ネットワークテクノロジーズ株式会社内
- (72)発明者 御幡 昭博
神奈川県横浜市港北区新横浜三丁目9番18号 富士通ネットワークテクノロジーズ株式会社内
- (72)発明者 橋永 健次
神奈川県横浜市港北区新横浜三丁目9番18号 富士通ネットワークテクノロジーズ株式会社内
- (72)発明者 諸井 貫一
神奈川県横浜市港北区新横浜三丁目9番18号 富士通ネットワークテクノロジーズ株式会社内

審査官 玉木 宏治

- (56)参考文献 特開平10-257071 (JP, A)
特開2004-007361 (JP, A)
宮本 正和 他, 動的負荷分散型IP網の構成法と経路制御アルゴリズムの性能に関する研究
, 電子情報通信学会技術研究報告 (信学技報) NS2002-113 IN2002-57 CS2002-68, 2002年
9月23日

- (58)調査した分野(Int.Cl., DB名)
H04L 12/00-66