

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第5372247号  
(P5372247)

(45) 発行日 平成25年12月18日 (2013. 12. 18)

(24) 登録日 平成25年9月27日 (2013. 9. 27)

(51) Int. Cl. F I  
H04L 12/851 (2013.01) H04L 12/851

請求項の数 14 (全 17 頁)

(21) 出願番号	特願2012-515434 (P2012-515434)	(73) 特許権者	391030332
(86) (22) 出願日	平成22年6月10日 (2010. 6. 10)		アルカテルルーセント
(65) 公表番号	特表2012-530439 (P2012-530439A)		フランス国、75007・パリ、アブニ
(43) 公表日	平成24年11月29日 (2012. 11. 29)		ユ・オクターブ・グレアール、3
(86) 国際出願番号	PCT/EP2010/058120	(74) 代理人	110001173
(87) 国際公開番号	W02010/145982		特許業務法人川口国際特許事務所
(87) 国際公開日	平成22年12月23日 (2010. 12. 23)	(72) 発明者	ドウ・フレエースハウウエル、ダニー
審査請求日	平成24年2月15日 (2012. 2. 15)		ベルギー国、ペー9940・エーフエル
(31) 優先権主張番号	09290462.2		ゲム、ベルゼーレ・ドルブ・31
(32) 優先日	平成21年6月18日 (2009. 6. 18)	(72) 発明者	ラーフエンス、クーンラート
(33) 優先権主張国	欧州特許庁 (EP)		ベルギー国、ペー9000・ヘント、ル
		(72) 発明者	ーイスパーンストラート・45
			バクセリ、フランソワ
			フランス国、エフ92190・ムドン、
			リュ・ドウ・パリ・83

最終頁に続く

(54) 【発明の名称】 輻輳制御方法および装置

(57) 【特許請求の範囲】

【請求項 1】

ネットワークノード内のトラフィックの輻輳の制御方法であって、前記トラフィックはレイヤの数 (L) によって分類されており、このノードへの入来する流れの数 (n(t)) を監視し、それによって、この入来する流れの数 (n(t)) に基づいて、および許可されたレイヤの現在の数 (l(t)) に基づいて、前記ノード内に入るための許可されたレイヤの次の数 (l(t+1)) が動作テーブル (T) の参照によって判断されるステップを含む、方法。

【請求項 2】

前記動作テーブルが、前記ノードに向かう観察されたトラフィックに基づいて動的に調節される、請求項 1 に記載の方法。

10

【請求項 3】

前記動作テーブルが、観察期間中に観察されたトラフィックに基づいて定期的なインターバルで動的に調節され、それによってマルコフ決定プロセスを使用して最適な動作テーブルを判断する、請求項 2 に記載の方法。

【請求項 4】

動作テーブルが、前記ノード内にローカルで記憶される、請求項 1 から 3 のいずれかに記載の方法。

【請求項 5】

前記動作テーブルが、ネットワーク輻輳コントローラによって集中的に算出され、さら

20

に前記ノードに通信される、請求項 1 から 3 のいずれかに記載の方法。

【請求項 6】

前記方法を実施する異なるノードの間で、前記動作テーブルの参照によって判断して得られる動作を通信し、それによって、競合する動作が生じた場合に、ノードごとに次の数の許可されたレイヤを適応させる別の発見的制御が前記ノード内で実行されるステップをさらに含む、請求項 1 から 4 のいずれかに記載の方法。

【請求項 7】

前記動作テーブルが、ネットワーク輻輳コントローラによって集中的に更新される、請求項 1 から 4 のいずれかに記載の方法。

【請求項 8】

ネットワークノード (IM ; AN1 ; AN2 ; AN3) に入るレイヤの数 (L) によって分類化されたトラフィックの輻輳の制御のための輻輳制御デバイス (CCIM ; CCIM' ; CCAN1 ; CCAN2 ; CCAN3) であって、このノードへの入来する流れの数 (n(t)) を監視し、前記許可されたレイヤの数に関連するレイヤ数によって識別されたパケットが前記ノードに入ることを可能にするだけであるように、この入来する流れの数 (n(t)) に基づいて、動作テーブル (T) に基づいて、および許可されたレイヤの現在の数 (l(t)) に基づいて、許可されたレイヤの次の数 (l(t+1)) を判断するように構成されている、輻輳制御デバイス (CCIM ; CCIM) 。

【請求項 9】

観察期間中に前記ノードに向かう観察されたトラフィックに基づいて、前記動作テーブルを動的に調節するようにさらに構成されている、請求項 8 に記載の輻輳制御デバイス (CCIM) 。

【請求項 10】

マルコフ決定プロセスを使用し、それにより最適な動作テーブルを判断することにより、前記観察期間中に観察されたトラフィックに基づいて、定期的なインターバルで前記動作テーブルを動的に調節するように構成されている、請求項 9 に記載の輻輳制御デバイス (CCIM) 。

【請求項 11】

前記動作テーブルの参照によって判断して得られた動作を、別のノード内の別の輻輳制御デバイスと通信するようにさらに構成され、競合する動作が生じた場合に、次の数の許可されたレイヤを適応させる別の発見的制御を実施するようにさらに構成されている、請求項 9 から 10 のいずれかに記載の輻輳制御デバイス (CCIM) 。

【請求項 12】

前記ネットワークノードが、前述の請求項 8 から 11 のいずれかに記載の輻輳制御デバイス (CCIM ; CCIM' ; CCAN2 ; CCAN3 ; CCIM) からなっていることを特徴とする、ネットワークノード (AN1 ; AN2 ; AN3 ; IM) 。

【請求項 13】

前記ネットワーク輻輳デバイス (NC) が、前記動作テーブル (T) を判断し、前記輻輳制御デバイス (CCIM) に前記動作テーブルを通信するように構成されている、請求項 8 に記載の輻輳制御デバイス (CCIM) と通信する、ネットワーク輻輳制御デバイス (NC) 。

【請求項 14】

さらに、観察されたトラフィックに基づいて前記動作テーブルを更新し、前記輻輳制御デバイス (CCIM) に前記更新された動作テーブルを提供するように構成されている、請求項 13 に記載のネットワーク輻輳制御デバイス (NC) 。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、輻輳制御方法および装置に関する。

【背景技術】

10

20

30

40

50

## 【0002】

オンデマンドストリーミングサービスでは、輻輳制御は普通、リソース許可制御を介して行われる。このようなシステムは、特徴が知られていると想定される進行中の流れ数を監視すること、または流れの集合の(瞬間)ビットレートを直接監視することのいずれかによって、特徴が知られていると想定される新しい流れが、それを通して流れが移動する全てのリンクにまだ適合するかどうかをチェックする。このチェックによりイエスの回答が得られた場合、流れは許可され、そうでない場合は、流れは拒絶される。この決定は、アプリケーションレベルで、セッションをセットアップしないことによって、またはセッションがセットアップされた場合でも、拒絶された流れから生じるトラフィックを遮断するネットワークの縁部にあるポリシーエンフォースャによってのいずれかで実施される。このようなアーキテクチャでは、ユーザは、映像を完全な品質で得る、またはサービスが拒否されるいずれかである。

10

## 【0003】

輻輳制御の代替方法は、拡張可能コーデックを介してである。その中で、各マルチメディア流れは、重要性が低くなっているレイヤ内で符号化される。

## 【0004】

リソース許可制御によるものに対して、拡張可能コーデックに基づく方法は、サービスへのユーザアクセスを決して拒否しないが、その品質はときどき、ユーザが目指しているものより低い。

## 【発明の概要】

20

## 【発明が解決しようとする課題】

## 【0005】

これらの知られている方法全ての欠点は、より高い重要度のパケットが失われたり廃棄したりしないことの絶対的な保証がないということである。

## 【課題を解決するための手段】

## 【0006】

欠点を克服するために、本発明による方法は、このノードへの入来する流れの量を監視し、それによって、この入来する流れの量に基づいて、および許可されたレイヤの現在の数に基づいて、前記ノード内に入るための許可されたレイヤの次の数が動作テーブルの参照によって判断されるステップを含んでいる。

30

## 【0007】

このように、許可されたレイヤの数を前の数、入来する流れの数、および動作テーブルに依存させることによって、より正確な輻輳制御方法が得られ、同時に、より重要なパケットが保存されることが保証される。

## 【0008】

改善した実施形態では、この動作テーブルは、前記ノードに向かう観察されたトラフィックに基づいて動的に調節される。

## 【0009】

これは、決定がトラフィック状態を反映して、許可されたレイヤのより正確な判断を可能にし、さらには品質を改善するという利点を有する。

40

## 【0010】

観察期間中に観察されたトラフィックに基づいて定期的なインターバルで動作テーブルを動的に調節することによって、マルコフ決定プロセスを使用して最適な動作テーブルを判断し、さらにより良い品質が得られる。

## 【0011】

動作テーブルは、ネットワーク輻輳コントローラによって集中的に算出することができ、さらに前記ノードに通信することができる、またはノード自体の中でローカルで記憶させることができる。

## 【0012】

動作テーブルは、ノード内でローカルで更新する、またはネットワーク輻輳コントロー

50

ラによって集中的に更新することができる。

【0013】

動作テーブルがノード内で更新される場合、前記動作テーブルの参照によって判断されるように、得られる動作の前記方法を実施する異なるノード間の通信を行うことができ、それによって、競合する動作が起こる場合、ノード毎に次の数の許可されたレイヤを適合させるためのさらに発見的な制御が前記ノード内で実行される。このように、隣接するノード間に生じる潜在的に競合する状態を解消することができる。

【0014】

本発明は、本方法を実施するための輻輳制御デバイスと、特定の実施形態の輻輳制御デバイスに動作テーブルおよびその更新を通信するためのネットワーク輻輳コントローラにも関する。

10

【0015】

特許請求の範囲で使用される「結合された」という用語は、直接接続のみに限定するものとして解釈されるべきではないことに留意されたい。したがって、「デバイスBに結合されたデバイスA」という表現の範囲は、デバイスAの出力がデバイスBの入力に直接接続されているデバイスまたはシステムに限定すべきではない。他のデバイスまたは手段を含む経路である可能性がある、Aの出力およびBの入力の間の経路が存在するということを意味する。

【0016】

特許請求の範囲で使用される「備えた」という用語は、この後に挙げる手段に限定するものとして解釈されるべきではないことに留意されたい。したがって、「手段AおよびBを備えたデバイス」という表現の範囲は、構成要素AおよびBのみからなるデバイスに限定するべきものではない。本発明に関して、デバイスの唯一の関連構成要素がAおよびBであることを意味する。

20

【0017】

添付の図面と合わせて実施形態の以下の説明を参照することによって、本発明の上記および他の目的および特徴は、より明らかになり、発明自体が最もよく理解されるであろう。

【図面の簡単な説明】

【0018】

【図1】方法の一実施形態を実現するための高レベルアーキテクチャを示す図である。

【図2】方法の別の実施形態を実現するための別の高レベルアーキテクチャを示す図である。

【図3】拡張可能なレイヤの原理を示す図である。

【図4】ノード実施の一実施形態の輻輳制御デバイスの実施形態を示す略図である。

【図5a】輻輳制御デバイスのいくつかの実施形態で使用される決定テーブルの例を示す図である。

【図5b】輻輳制御デバイスのいくつかの実施形態で使用される決定テーブルの例を示す図である。

【図6】方法の一実施形態を実施する第1のフローチャートである。

40

【図7】改良された方法のステップを実施するフローチャートである。

【図8】方法のこのような改良された実施形態の効果を示すタイミング図である。

【図9】輻輳制御デバイスの改良バージョンの実施形態を示す略図である。

【図10】方法の変形実施形態を実現するための別の高レベルアーキテクチャを示す図である。

【図11】方法の別の実施形態を実現するための別の高レベルアーキテクチャを示す図である。

【発明を実施するための形態】

【0019】

説明および図面は単に、本発明の原理を示しているだけである。したがって、当業者は

50

本明細書には明示的に記載または図示されていないが、本発明の原理を実施し、その趣旨および範囲内に含まれる様々な構成を工夫することが可能であることが分かるであろう。さらに、本明細書で言及されている全ての例は基本的に、本発明の原理、および当技術分野を促進するために（１人または複数人の）発明者が貢献した概念を理解する際に読者を助けるための単に教育的な目的であることを明示的に意図したものであり、このように特に言及された例および状態に限定するものとして解釈すべきではない。さらに、本発明の原理、態様および実施形態と、その特別な例に言及する、本明細書での全てのステートメントは、その同等物を含むことを意図している。

#### 【 0 0 2 0 】

本明細書のあらゆるブロック図は、本発明の原理を実施する例示的な回路の概念図を示すことを、当業者は理解すべきである。同様に、このようなコンピュータまたはプロセッサは明示的に示されているかどうかに関わらず、あらゆるフローチャート、フロー図、状態遷移図、擬似コードなどは、コンピュータ読取可能な媒体で実質的に、コンピュータまたはプロセッサによってそのように実施することができる様々なプロセスを示すことが分かるであろう。

10

#### 【 0 0 2 1 】

方法の一実施形態は、チョーキングベース輻輳制御と呼ばれる、拡張可能コーデックを介した新しいタイプの輻輳制御を提案している。この方法は、図 1 に略図的に示されているものなどの多くのネットワークで使用することができる。以下では、本発明の一実施形態は、例えば DSL アクセスネットワークでの使用に対して説明されているが、本発明の他の実施形態は、固定または移動に関わらず、全ての他のタイプのネットワークで可能である。

20

#### 【 0 0 2 2 】

図 1 では、マルチメディアサーバ MM は、サービスアグリゲータなどの中間ノード IM を介して少なくとも 1 つのアクセスノードに結合されている。図 1 では、これらの 3 つだけが図示されており、AN 1 から AN 3 と示されている。リアルネットワークでは、アクセスノードの数は容易に 10 から 100 の範囲内であってもよい。普通、これらのアクセスノードはそれぞれ、U 1 から U n で示された個別のユーザに結合させることができ、1 つのアクセスノードは典型的には 100 から 1000 のユーザにサービスを提供することができる。

30

#### 【 0 0 2 3 】

本発明の一実施形態によると、ノードの少なくとも 1 つは輻輳制御デバイスを備えている。図 1 では、全てのノードが、それぞれのアクセスノード AN 1 から AN 3 内のそれぞれの輻輳制御デバイスに対して CCAN 1 から CCAN 3、および中間ノード IM 内に含まれている輻輳制御デバイスに対して CCI M と示される、このような輻輳制御デバイスを備えている実施形態が示されている。アクセスノードまたは中間ノードであるかに関わらず、1 つのノードだけが、このような輻輳制御デバイスを備えている他の実施形態が可能である。中間ノードだけが、本発明による輻輳制御デバイスを備えているこのような実施形態が図 2 に示されており、さらに以下のパラグラフでより詳細に説明する。

#### 【 0 0 2 4 】

本発明による輻輳制御デバイスのほとんどの実施形態は、マルチメディア流れが拡張可能な方法で符号化されることを想定している。拡張可能な映像符号化が、例えば ITU-T Rec. H. 264 の付録 G 「汎用視聴覚サービスのための先進映像符号化」内で標準化され、符号化された映像流れは、ベースレイヤおよび少なくとも 1 つの拡張 (enhancement) レイヤから構築されていることを暗示している。多くの標準コーデック、例えば MPEG 2、4 は拡張可能な拡張子を有するが、他の専有スキームを使用することもできる。パケットベース輸送では、各レイヤに関連付けられたビットストリームは、各パケットのヘッダ内の識別子に基づいて、どのレイヤに属するかが分かるようにパケット化されている。例としては、この目的で使用することができる、IP ヘッダ内の DiffServ Code Point (DSCP) またはサービスタイプ (ToS) バイト

40

50

を挙げることができ、どのレイヤにパケットが属するかを識別する。しかし、他のタイプのヘッダ内の他の識別子を使用することもできる。

【 0 0 2 5 】

ユーザの宅内での復号器がベースレイヤを受けるだけである場合、映像を基本的品質で復号化することができる。より多くの層を復号器で受ければ受けるほど、復号化された映像の品質がより良くなる。図 3 に示すように、特定のレイヤの重要性を判断する、レイヤ内の順序付けがあり、レイヤ 0 はベースレイヤで、最も重要なレイヤであり、一方、この場合は  $L$  である最も高い数を有するレイヤは、最も重要でないレイヤである。任意選択のレイヤ  $l$  では、全てのレイヤ 0 から  $l - 1$  までが受けられない限り、このレイヤ  $l$  が無用であるという規則がある。

10

【 0 0 2 6 】

本発明による輻輳制御デバイスのほとんどの実施形態はまた、その一部である個別ノード内のスケジューラによって行われる、スケジューリング技法に依存している。図 4 では、 $S$  で示されたスケジューラを備えたノード  $IM$  が示されている。このスケジューラは、異なる流れのパケット間の区別を行わないが、このレイヤを識別するための識別子に基づいて、異なるレイヤに関連付けられたパケットを区別する。このようなスケジューリング技法は、ノードが図 4 に  $B$  で示された入力バッファを備えていることを暗示している可能性があるが、これは必然ではない。ノードはさらに、所定のタイムスロットで、例えば毎秒、入力パケットが例えばバッファ  $B$  を介してノードに入ることを可能にするように、スケジューラが  $L + 1$  レイヤからどのレイヤ  $l$  まで適合されているかを判断する、およびノードまたは調節するようになっているコントローラ  $C$  をそれ自体が備えている、輻輳制御デバイス  $CCIM$  の一実施形態を含んでいる。レイヤ  $l$  まで（これを含む）が許可されている場合、これは範囲  $[ 0, l ]$  内のレイヤを示す識別子を有する次のタイムスロット内の全てのパケットが許可されており、残りの入力パケットが拒否されるということを意味している。標準的  $SV C$  符号化が使用される場合、たくさんのノードは既にスケジューリング機構を含んでいるが、パケットを許可するまたは拒否する決定は、例えば流れの宛先に基づいて行われる。しかし、この機構は輻輳の問題を解消しない。他の既存の機構では、他の欠点は、より高いレイヤ数を有するあまり重要でないレイヤも、より低いレイヤに対するバッファ空間を犠牲にして、いくつかのバッファ空間を占める可能性があるという事実に関連している。加えて、これらの他の機構では、許可されたレイヤの数は時間の経過とともに急速に変動して、悪いユーザ経験につながる可能性がある。他の欠点は、このような従来技術のシステムが発振する可能性があるということである。

20

30

【 0 0 2 7 】

本発明によると、輻輳制御デバイスの一実施形態はしたがって、所定の例では時間  $t$  で、 $l(t)$  の現在の値に基づいて、ノードへの入来する流れの現在の数  $n(t)$  に基づいて、また決定テーブル  $T$  に基づいて、次の特定の時間インスタンス  $t + 1$  に対する  $l(t + 1)$  の値を判断するように構成されたコントローラを備えている。このコントローラはしたがって、次のタイムスロット  $t + 1$  内でどのレイヤ  $l$  までサポートするかを判断することが可能になる。図 4 では、コントローラ  $C$  は、監視デバイス  $M$  から、時間インスタンス  $t$  でのノードへの入力流れ  $F_1$  から  $F_n$  の数  $n(t)$  を受け、さらに、 $l(t + 1)$  を判断することができるように、 $T$  で示された決定テーブルを調べるようになっている。この値はスケジューラ  $S$  に与えられ、これに応じて、そのレイヤ識別子に基づいて入力パケットをフィルタリングし、それによって、レイヤ  $l(t + 1)$  までのパケットだけが、通過させることが可能であり、時間インスタンス  $t + 1$  でバッファリングされる。

40

【 0 0 2 8 】

この決定テーブルは図 4 では  $T$  で示されている。このような決定テーブルの例が図 5 a および 5 b に示されている。図 5 a の決定テーブルは、入力変数として、最も高いレイヤ数  $l(t)$ 、およびそのノードへの入来する流れの合計数  $n(t)$  を有する。その最終列は、これらの入力の関数として、関連動作を示しており、この動作は「0」、「-」または「+」のいずれかによって示されている。「0」で示される動作は、次のスロット  $t +$

50

1では、同じ数のレイヤ(0から最大1以下)がバッファに入ることが可能であるということの意味している。「-」で示された動作は、次のスロット $t+1$ では、1つ少ないレイヤがバッファに入ることが可能であり、その結果、0から $1-1$ の間でこれらを含むレイヤ識別子を備えた全ての packets が許可されることになる。「+」で示された動作は、次のスロット $t+1$ では、例えば、1つ多いレイヤがバッファに入ることが可能であり、その結果、レイヤ0から $1+1$ の間でレイヤ識別子を含むすべての packets が許可されることになる。したがって、テーブル5 a上の第1行を読むと、これは、現在のスロットで、6つのレイヤがバッファに入ることが可能である場合、170の入力流れがこのスロットでノードに入り、次のスロットでバッファに入ることが可能なレイヤの数を考慮する限り、変更を行う必要はなく、したがって、この数は6のままであることを意味している。このテーブル上の第2行は、現在のスロットで、6つのレイヤがバッファに入ることが可能である場合、180の入来する流れがこのスロットでノードに入り、次のスロットでの記憶のためにバッファに入ることができるレイヤの数は、例えば、次のスロットで許可されている5つのレイヤに対応する、現在の値に対して1だけ下げなければならないということを示している。このテーブルの第3行は、現時点で、7つのレイヤがバッファに入ることが可能であり、ノードが現在130の入来する流れを受ける場合、次の時間インスタンスの間に、8つのレイヤがバッファに入ることが可能であることを示している。

10

## 【0029】

1だけ増減させる代わりに、これらの動作「+」または「-」はまた、他の実施形態では、例えば2または3つ少ないまたは多いレイヤが許可されるということを示している。

20

## 【0030】

テーブルの別の例が、図5 bに示されている。絶対数の入来する流れの代わりに、数 $n(t)$ の流れが区分にグループ化される。このテーブルの最初の3つの行は、現在のスロットでは、 $1(t) = 6$ つのレイヤがバッファに入ることが可能である場合、

1) 動作が行われず、(テーブルの表示された第2行にしたがって、)流れの数が140から170の境界の間である場合に、次のスロット内の許可されたレイヤの数を6、例えば $1(t+1) = 6$ にしておく。

2) (テーブル上の表示された第3行にしたがって、)170を超える流れがある場合に、次のスロットでは、許可されたレイヤの数を、例えば6から5に減らす動作が行われる。

30

3) (テーブル上の表示された第1行にしたがって、)140未満の流れがある場合、許可されたレイヤの数を6から7に増やす動作が行われる。

## 【0031】

前のテーブル5 aと同様の考察を行うことができ、したがって、「+」動作は、1以上である所定の数での増加を示し、「-」動作は、増加の所定の数と同じである可能性があるが、これと異なってもよい1以上である別の所定の数での減少を示すことができる。

## 【0032】

全ての可能な状態が挙げられている図5 aのタイプのテーブルはしたがって、多くの数の行を含んでおり、レイヤの数および流れの数の全ての可能な組み合わせを示している。より詳細には、ノードが現在サポートしている流れの数 $n(t)$ が区分されていない場合、行の数はレイヤの最大数掛ける流れの最大数、例えば $10 \times 1000 = 10000$ に等しい。区分により、テーブルをはるかに短くすることが可能になる、例えば、3入力/レイヤがテーブル5 bで示されている例では必要である。

40

## 【0033】

単純な実施形態では、このような決定テーブルを、ノード内にまたは輻輳制御デバイス自体内にローカルで記憶することができる。図4は、このテーブルが輻輳制御デバイスCIM内に記憶されている一実施形態を示している。図はさらに、全ての流れからの、レイヤ1までの識別子を備えた packets がしたがって、バッファ内に通ることが可能である

50

という原理を略図的に示している。バッファ自体は、この図によって提案されるように、異なるレイヤに関連するいくつかのサブバッファに分割する必要はなく、これは、全ての流れからのパケットが通過することができるが、そのレイヤ数に基づいてフィルタリングされているという原理をより良く示すために描かれているだけである。

#### 【 0 0 3 4 】

記載した方法の異なるステップを実施するための詳細な実施形態を示すフローチャートが、図 6 に示されている。この方法は、現在のタイムスロットでの入来する流れ  $F_1$  から  $F_n$  の数を監視することで始まる。これは、図 4 に示すように、監視デバイス  $M$  によって行うことができる。次のステップでは、この入来する流れの数と、現在の最も高い許可されたレイヤ  $l(t)$  を使用して、テーブル内への適当な入力を判断し、そのステップで、得られた動作をテーブルから読み取ることもできる。これらのステップは、コントローラ  $C$  内で行うことができる。得られる「動作」値によって、 $l(t+1)$  の値が適合され、この値はスケジューラ  $S$  にフィードバックされて、どのレイヤに属するかを全ての入力パケットから判断し、それに応じて、レイヤ数が  $l(t+1)$  以下である場合にバッファまで通過させる、またはレイヤ数が  $l(t+1)$  より大きい場合にこれらを破棄するように構成されている。パケットのレイヤ識別は、このプロトコルが輸送に使用される場合に、例えば IP ヘッダ内で見られる。これらのヘッダ内の識別子に基づいてパケットを分類する方法は、当業者によく知られており、さらに詳細には説明しない。

#### 【 0 0 3 5 】

より複雑な実施形態では、輻輳制御デバイスは、例えば  $l(t)$  および  $l(t+1)$  を判断するための時間インスタンスが、約数秒の大きさである前の例では、20分毎にこのテーブルを定期的に更新するように構成されている。この更新は、観察期間中に、ネットワークノードに向かって観察されたトラフィックに基づく可能性があるが、これは必然ではない。このテーブルを更新するために、輻輳制御デバイス  $CCIM$  は、したがって、レイヤ許可数を更新するタイムスロットよりはるかに長い、観察期間と呼ばれる、特定の期間にわたってトラフィックをモデリングするように構成されている。より正確には、事前に選択したトラフィックモデルのいくつかのパラメータを判断する。このようなモデルはマルコフモデルであるが、次のパラグラフで説明するように、別のタイプであってもよい。

#### 【 0 0 3 6 】

$l(t+1)$  を判断するために使用される決定テーブルを更新するために、方法は例えば、以下の手順からなる可能性がある：現在アクティブな決定テーブルに加えて、多くの予め選択されたまたは所定の代替的な決定テーブルが維持される。これらは例えば、図 5 b のタイプの区分したテーブルの閾値を代表的セットの値に設定することによって得ることができ、これらの値自体は、テーブルを完全に特定するためにレベル  $l$  ごとに 2 つの閾値だけが必要であり、レイヤ  $l+1$  に関連する閾値がレイヤ  $l$  に関連するものより高いテーブルが論理的でないという原則に基づいて選択することができる。したがって、これらの原則を念頭に置いて、代表的なテーブルのセットを容易に選択することができ、これらは最初に  $IM$  または輻輳コントローラ内に記憶させることができる。これらの代替決定テーブルそれぞれでは、その後、この代替動作テーブルが使用される場合に、値関数が何を生じさせるかが観察期間で算出される。このような値関数は例えば、以下の式 (1) によって与えられる：

$$V[l(t), n(t)] = R[l(t), n(t), a(l(t), n(t))] + V[l(t+1), n(t+1)] \quad (1)$$

#### 【 0 0 3 7 】

式中、 $V[l(t), n(t)]$  (および、 $V[l(t+1), n(t+1)]$ ) は、時間  $t$  で許可された  $l$  のレイヤ、および  $n$  の入来する流れがある場合の値関数 (および、それぞれ時間  $t+1$  での将来の値関数) を示しており、

$R[l(t), n(t), a(l(t), n(t))]$  は、瞬時報酬関数を示しており、

10

20

30

40

50

$a(l(t), n(t))$  は、決定テーブルで与えられたように、時間  $t$  で許可された  $l$  のレイヤおよび  $n$  の入来する流れがある場合に行われる動作を示している。

【0038】

瞬時報酬関数の例が、以下の式(2)によって与えられる：

$$R[l(t), n(t), a(l(t), n(t))] = \alpha \cdot G(l(t)) \cdot n(t) - \beta \cdot n(t) \cdot \max\{(F - C) / F, 0\} - \gamma \cdot n(t) \cdot l_{\{a(l(t), n(t)) > 0\}} \quad (2)$$

式中、 $\alpha$ 、 $\beta$  および  $\gamma$  は以下の解釈の正の定数である。 $\alpha$  は、オペレータがサポートされた流れごとに得る時間単位ごとの報酬であり、 $\beta$  はオペレータが損失パケットごとに支払わなければならないペナルティであり、 $\gamma$  はオペレータが品質変化ごとに支払わなければならないペナルティである。

$\alpha \cdot G(l(t))$  は、スロット  $t$  内でのレイヤ  $l$  までの流れの輸送に関連する報酬、例えば、単一のユーザが、映像およびレイヤ  $l$  に対応する品質を受けるために喜んで支払う価格である。

$F$  は、スロット  $t$  内のトラフィック量である。 $n(t)$  では、流れ  $F$  は  $n(t) \cdot l(t)$  に対応しており、

$C$  はリンク容量、すなわち、スロット  $t$  ごとに伝達することができる情報量である。別の方法では、 $C$  は、オーバーフローをより十分避けるために、リンク容量より僅かに小さく選択することができる。

【0039】

観察期間の後、最も高い値を蓄積した代替決定テーブルは、アクティブな決定テーブルに関連する値を超える場合に、アクティブな決定テーブルまで進められる。この場合、アクティブな決定テーブルは代替決定テーブルの1つとなるように降格される。この方法を効率的にするために、多くの代替動作テーブルを評価する必要がある。これを避ける代替方法を次に記載する。

【0040】

この代替方法では、最初に遷移マトリックスを作り出さなければならず、これはトラフィックの観察に基づいている。このような遷移マトリックスは、流れの数が、この入力に対するマトリックスの行によって特性化された特定の値から、この入力に対するマトリックスの列によって特性化された別の値まで増減される可能性を示す入力を含むことができる。この場合、遷移マトリックス  $TRM[n(t), n(t+1)]$  はしたがって、現在のタイムスロット  $t$  および次のタイムスロット  $t+1$  の絶対数の流れの観察された差に基づいて構築される。この遷移マトリックスを入力として、 $l(t)$  を判断するために使用される決定テーブルは、例えば、これ以下MDP理論と省略するマルコフ決定プロセス理論に基づいて更新することができ、3つの可能な動作だけを前のパラグラフで説明した動作として行うことができる前提で値関数の平均を最適化することができる：

前に記載した決定テーブルに関連するような動作に対応するように、

1) 動作「+」によって示されるように、次のスロットで1つまたは複数のレイヤだけ余分にすることを可能にする、

2) 動作「0」によって示されるように、次のスロットで同じ量のレイヤを可能にする、または、

3) 動作「-」によって示されるように、次のタイムスロットで1つまたは複数のレイヤだけ少なくすることを可能にする。

【0041】

MDP論理により、このような遷移マトリックス  $TRM[n(t), n(t+1)]$  および所与の値関数によって記載された、マルコフプロセスに対して最適な動作テーブルを選択することが可能になる。値関数  $V(l(t), n(t))$  は、以下の式(3)に示すように、瞬時報酬、およびそれ自体、遷移マトリックスおよび行われる動作に依存する予測将来値の合計からなることができる：

$$V[l(t), n(t)] = R[l(t), n(t), a(l(t), n(t))] +$$

10

20

30

40

50

$$j \text{ TRM} [ n ( t ) , j ] \cdot V [ l ( t + 1 ) , j ] \quad ( 3 )$$

式中、 $V [ l ( t ) , n ( t ) ]$ は、時間  $t$  で許可された  $l$  のレイヤ、および  $n$  の入来する流れがある場合の値関数を示しており、

$R [ l ( t ) , n ( t ) , a ( l ( t ) , n ( t ) ) ]$ は、瞬時報酬関数を示しており、

$a ( l ( t ) , n ( t ) )$ は、決定テーブルで与えられているように、時間  $t$  での許可された  $l$  のレイヤおよび  $n$  の入来する流れがある場合に行われる動作を示しており、

$j \text{ TRM} [ n ( t ) , j ] \cdot V [ l ( t + 1 ) , j ]$ は、状態  $( l ( t ) , n ( t ) )$  から  $( l ( t + 1 ) , n ( t + 1 ) )$  までの移動に関連する平均将来値を示しており、  
 式中、 $l ( t + 1 )$ は、動作  $a ( l ( t ) , n ( t ) )$  および許可されたレイヤ  $l ( t )$  の現在の数によって判断される、次のタイムスロットの許可されたレイヤの数である。

10

【 0 0 4 2 】

実際、 $l ( t + 1 ) = l ( t ) + 1_{\{ a ( l , n ) = " + " \}} - 1_{\{ a ( l , n ) = " . " \}}$  であり、式中  $1_A$  は、ステートメント  $A$  が当てはまる、あるいは  $0$  である場合に、値  $1$  をとる特性関数である。

【 0 0 4 3 】

瞬時報酬関数  $R [ l ( t ) , n ( t ) , a ( l ( t ) , n ( t ) ) ]$  の例は、既にご与えた式 ( 2 ) によって与えることができる。このような報酬関数の別の例は、以下の式 ( 4 ) によって与えることができる：

$$R [ l ( t ) , n ( t ) , a ( l ( t ) , n ( t ) ) ] = \cdot G ( l ( t ) ) - \cdot m a x \{ ( F - C ) / F , 0 \} - \cdot l_{\{ a ( l ( t ) , n ( t ) ) = 0 \}} \quad ( 4 )$$

20

式中、 $\cdot$  および  $\cdot$  は正の定数である。

$\cdot G ( l ( t ) )$  は、スロット  $t$  でのレイヤ  $l$  までの流れの輸送に関連する報酬、例えば、映像およびレイヤ  $l$  に対応する品質を受けるためにユーザが支払いたい価格である。

$F$  は、スロット  $t$  内のトラフィック量である。  $n$  の流れでは、  $F$  は  $n \cdot l$  に対応しており、

$C$  は、リンク容量である。

【 0 0 4 4 】

両方の式 ( 2 ) および ( 4 ) 内で、  $m a x \{ ( F - C ) / F , 0 \}$  は、スロット  $t$  の間に損失したパケットにほぼ等しく、それによってこの第 2 項は損失したパケットに対する割引に等しい。最終項は、変動に関する割引であり ( 式中、  $l_{\{ a ( l , n ) = 0 \}}$  は、  $a ( l , n ) = " 0 "$  である場合、値  $1$  をとり、そうでない場合は  $0$  をとる。特性関数であり)、スロットごとにあまりにしばしば、サポートされた最大レイヤ  $l$  を変更するのを妨げる。

30

【 0 0 4 5 】

このプロセスは、図 7 に略図的に示されている。本実施形態では、第 1 のブロックは、次のタイムスロット内の流れの数  $n ( t + 1 )$  を測定することによって、前に測定した流れの数  $n ( t )$  とこれを比較することによって、遷移マトリックス  $\text{TRM}$  がどのように次第に蓄積されるかを示している。より正確には、始めに、 $\text{TRM}$  の全ての入力が  $0$  にされる。各タイムスロット  $t$  では、測定した流れの数が実際  $n ( t )$  から  $n ( t + 1 )$  まで変化したことが分かった場合に、数  $n ( t )$  の行、および数  $n ( t + 1 )$  の列の  $\text{TRM}$  マトリックスへの入力が  $1$  だけ増加される。したがって、合計観察期間の後に、 $\text{TRM}$  マトリックス内の各入力値は普通、  $n ( t )$  の入来する流れが何回  $n ( t + 1 )$  に変更されたかを含んでおり、  $n ( t )$  は現在のスロットで行を特性化する変数であり、  $n ( t + 1 )$  は、次のスロットで列を特性化する別の変数を示している。

40

【 0 0 4 6 】

最後に、遷移マトリックスは正規化され、それによって、行の全ての入力が同じ数によって掛けられ、その結果、各行の入力の合計が  $1$  に等しくなる。これにより、最適な決定テーブル  $T$  を、例えば、 $\text{MDP}$  理論のアルゴリズムを使用して、値関数の反復により得る

50

ことが可能になる。

【0047】

この最適な決定テーブルは、図7で説明されるように、(2)を前提として等式(1)を最大化する決定テーブルである。この最適なテーブルがこのMDP理論に基づいて見つけられると、この最適なテーブルはその後、また図7に示すように、次のサイクルの間に使用されるようにインストールされる。

【0048】

遷移マトリックスを構築し、最適な決定テーブルを判断するこのようなプロセスは、全ての観察期間Tで行うことができ、好ましい実施形態では、 $T \gg$  タイムスロットtである。トラフィックが観察される期間は、Tに等しい必要はないが、もっと長くてもよい、例えば2Tであってもよく、プロセス中に適応させることもできることに留意されたい。この点で、観察期間およびテーブル適応期間は異なってもよく、それぞれプロセスの間に調節可能である。

【0049】

この学習されたモデルに基づいて、このような改良型コントローラの一実施形態は、例えば、各ノードに対して個別に、マルコフ決定プロセス(MDP)理論を使用して、最適な決定テーブルを算出するように構成されている。最適な決定テーブルは入来する流れの現在の量 $n(t)$ に左右されるだけであるが、行われる動作は、トラフィックの可能性のある将来の進化を考慮することに留意されたい。例えば、MDP理論により、観察期間が、全ての可能なイベントが起こるのに十分長く、知られている日ごとの進化に対して十分短く選択される場合に、最も起こりそうな将来の進化を予測することが可能になる。

【0050】

この学習プロセスの効果が図8に示されている。この図は、時間tの関数として、入来する流れnの量の進化を示している。この進化は、太い黒線で示されている。時間インスタンス0では、この数はむしろ低く、次の時間インスタンス(1)では、この数は増加されるなどである。この図はまた、1の1つの特定の値に対して、図5(b)に示されたタイプの(区分された)決定テーブルの閾値を示している。したがって、現在のスロットtでバッファに入ることが可能である最も高いレイヤであった1のこの特定の値では、テーブル5bにしたがって2つの関連する閾値がある。 $n(t)$ が最も低い閾値より低い場合、動作はこのテーブルにしたがって「+」であり、 $n(t)$ が2つの閾値の間である場合、動作は「0」であり、 $n(t)$ が最も高い閾値を超える場合、その動作は「-」である。1つの特定の1に対する最も高い閾値のみが図に示されている。図は、「デルタ調節」と示されているテーブル調節期間にわたってトラフィックを観察した後に、新しい決定テーブルが算出されロードされることを示している。図5bのタイプの決定テーブルを参照すると、これは、高い閾値の値の増加によって、図8に示されている閾値の値を調節するということになる。トラフィックの合計量がこの閾値を超えていることがほとんど起こらなかったため、MDPアルゴリズムの結果はしたがって、この観察に基づいて、この閾値を増加させるのが安全であるということになる。

【0051】

1つのノードに含まれるこのようなより複雑な輻輳制御デバイスCCIMに対するより詳細な実施が、図9に示されている。図4に関して、この中間ノードは、本実施形態では、余分な処理または他のデバイスを備えており、1つのデバイスは遷移マトリックスを構築するためのものでTRMと示されており、別のデバイスは、等式(1)および(2)によって与えられる報酬関数にしたがって最良の報酬を生み出す決定テーブルを判断するためのものでMDPと示されており、決定テーブルTの更新につながる。もちろん、方法ステップのすべては、コントローラC'内でも、または1つの処理デバイス上で実行することができる。

【0052】

これに関して、「コントローラ」と記された任意の機能ブロックを含む、図に示された様々な要素の機能は、専用ハードウェアと、適当なソフトウェアに関してソフトウェアを

10

20

30

40

50

実行することが可能なハードウェアの使用により与えることができることが記載されている。プロセッサによって与えられる場合、機能は単一の専用プロセッサによって、単一の共有プロセッサによって、またはそのいくつかを共有することができる複数の個別プロセッサによって与えることができる。さらに、「プロセッサ」または「コントローラ」という用語の明示的な使用は、ソフトウェアを実行することが可能なハードウェアのみのことを言うものと解釈すべきではなく、これに限らないが、デジタル信号プロセッサ(DSP)ハードウェア、ネットワークプロセッサ、特定用途向け集積回路(ASIC)、フィールドプログラマブルゲートアレイ(FPGA)、ソフトウェアを記憶するための読み取り専用メモリ(ROM)、ランダムアクセスメモリ(RAM)、および不揮発性記憶装置を暗示的に挙げるることができる。他のハードウェア、従来のおよび/またはカスタムハードウェアを含めることもできる。

10

## 【0053】

前に記載した実施形態とは別に、それぞれ説明した方法を実施することができる、隣接したノード間の潜在的な矛盾する決定を避けるために、協調方式が使用される他の実施形態が可能である。局所決定テーブルを備えた木ベースネットワークアーキテクチャでは、このような協調方式の最も単純な例は、積極的な減少および注意深い増加に基づくことができる。より詳細には、図1のネットワークを再び参照すると、これは以下の方針を実施する必要がある可能性がある：

1) ANからの全ての「-」動作を実施し、IMからの「+」動作を無視する、例えば、少なくとも1つのANが動作「-」を有する場合に、IM内で「+」動作を「0」に変化させる。

20

2) IMだけが「-」動作を有し、全てのANが「+」または「0」動作のいずれかを有する場合、最も高い流れの数をサポートするAN上で「-」動作を実施する。

3) ANが「-」動作を要求せず、1つのANだけが「+」動作を要求する場合、IMが「-」動作を要求しなかった場合のみにこれを可能にする。

4) ANが「-」動作を要求せず、2つ以上のANが「+」動作を要求する場合、IMが「-」動作を要求しなかった場合に、1つのANだけが「+」動作例えば、最小1および最大nを有するものを可能にする。

## 【0054】

多数のノードへの拡張はしたがって、各局所状態  $l_k(t)$ ,  $n_k(t)$  を、ノードk内の適当な動作に関連させる、ノード毎に局所決定テーブル、プラス協調または均衡を破る方式を必要とし、ここでkはノード毎のインデックスを示している。これは、図10の個別の輻輳制御デバイス間の破線矢印により示された、ノード間の異なる輻輳制御デバイス間のいくつかの通信を必要とする。別の方法ではまた、グローバル協調方式を実施することができる。この場合、ノード毎の決定テーブルは、局所情報だけでなく、グローバル情報によるものであり：グローバルな中央実施動作テーブルは、 $(l_1(t), n_1(t), \dots, l_k(t), n_k(t), \dots, l_k(t), n_k(t))$  の各可能な組合せに関連した動作を含んでいる。この場合、このグローバル動作テーブルが中に記憶される中央またはグローバルネットワークコントローラと、個別のノードの間で情報を交換する必要があり、それによってそれぞれ各ノードに対して独自の異なる動作テーブルを有することができる。非局所である  $(l_1(t), n_1(t), \dots, l_k(t), n_k(t), \dots, l_k(t), n_k(t))$  の全ての情報はその後、グローバルネットワークコントローラから個別のノードまで通信されるものである。

30

40

## 【0055】

このような情報の交換とは別に、このグローバルネットワーク輻輳コントローラの別の機能は、テーブル更新を判断することである可能性がある。その目的で、流れの数  $n_k(t)$  が、各ノード  $k (= 1 \dots K)$  に対して時間の経過と共にどのように進化するかが観察される。これは、そのネットワークトポロジーの知識、およびアプリケーションプロバイダとの情報の交換により行うことができる。その後、観察期間Tにわたる観察の結果、(どのように  $n_k$  がノードkで進化するかを捕捉するために) 各ノードに対して個別に遷

50

移マトリックス、または(どのように  $n_1(t), \dots, n_k(t), \dots, n_K(t)$  が進化するかを捕捉するために)ネットワークに対してグローバル遷移マトリックスのいずれかを構築するように構成されている。さらに、(同じ局所報酬関数および同じセットの可能な動作で)個別に、または合計報酬が局所報酬の加重和である場合にグローバルに、各ノードに対してこの遷移マトリックスに対するMDPを解くようになっている。ノード毎の決定テーブルが得られ、その後さらに、局所ノードに通信される。

【0056】

ノードが個別にその動作テーブルを判断および適応させる前述の場合、および均衡を破る手順が必要である場合でさえも、ネットワーク輻輳コントローラは、可能性のある均衡が破られた動作自体を算出し、これらを個別のノードに通信することができる、またはネットワーク輻輳コントローラは、局所ノードの決定を互いに中継することができ、その結果、独自のさらなる決定を行うことができる。但し、これらは独自の輻輳状態にアクセスでき、いつ協調動作が必要であるかが分かっているものとする。

10

【0057】

中央またはグローバルネットワーク輻輳コントローラNCを備えたこのような実施形態が、図11に示されている。したがって、異なるノード内の個別の輻輳制御デバイス全てと通信するように構成されている。

【0058】

別の実施形態では、トラフィックは、流れの数を測定する代わりに、リンク上のビットレートを介して観察することができる。

20

【0059】

本発明の原理を特定の装置に関連して上に記載したが、添付の特許請求の範囲に規定されているように、この記載は例としてのみ行われたものであり、本発明の範囲を限定するものではないことを明らかに理解されたい。

【図1】

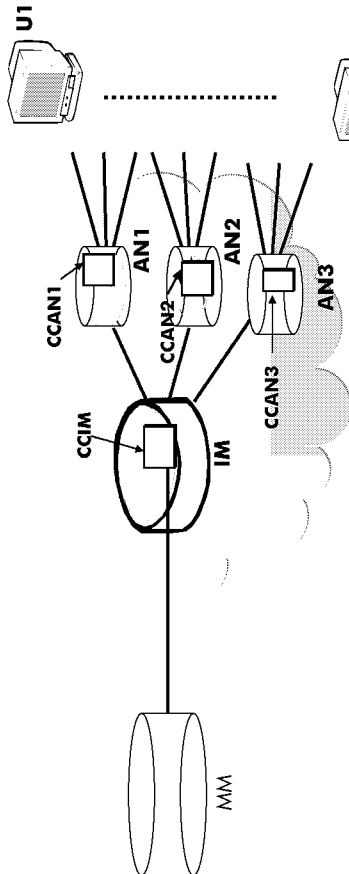


Figure 1

【図2】

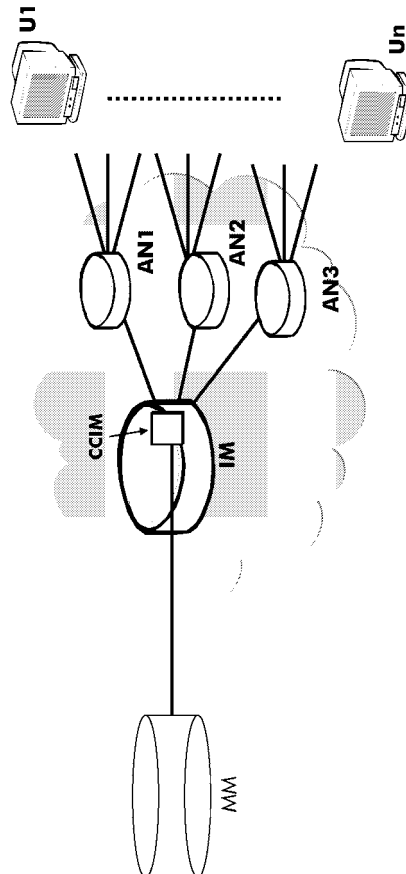


Figure 2

【 図 3 】

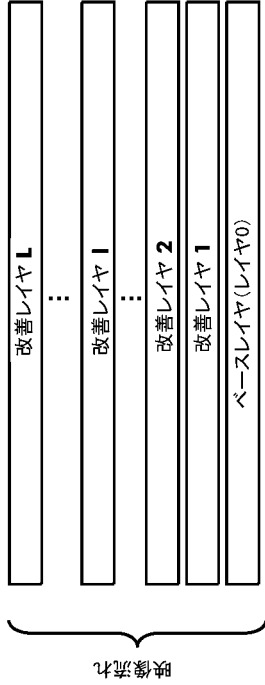


Figure 3

【 図 4 】

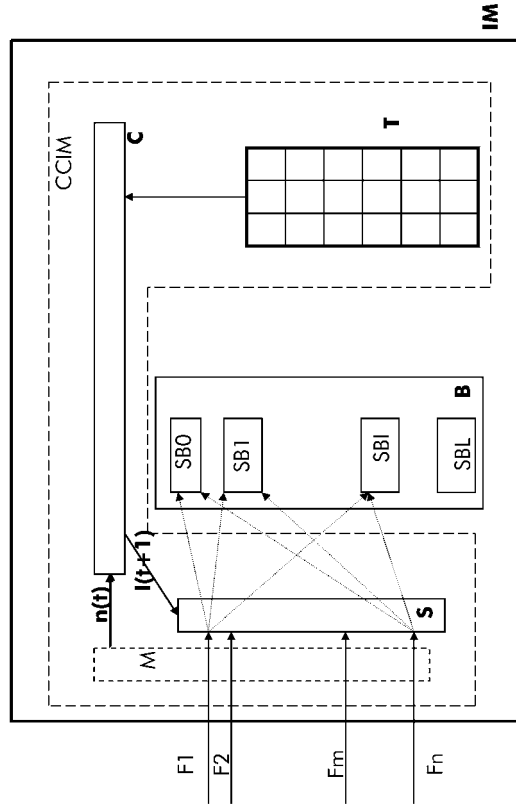


Fig. 4

【 図 5 a ）】

最も高いレイヤ $I(t)$	流れの数 $n(t)$	動作
6	170	0
6	180	-
7	130	+
7	140	0
7	150	-
8	130	-

Fig. 5a

【 図 5 b ）】

最も高いレイヤ $I(t)$	流れの数 $n(t)$	動作
$I(t)=6$	$n(t) < 140$	+
$I(t)=6$	$140 \leq n(t) \leq 170$	0
$I(t)=6$	$n(t) > 170$	-
$I(t)=7$	$n(t) < 130$	+
$I(t)=7$	$130 \leq n(t) \leq 150$	0
$I(t)=7$	$n(t) > 150$	-
$I(t)=8$	$n(t) < 100$	+
$I(t)=8$	$100 \leq n(t) \leq 130$	0
$I(t)=8$	$n(t) > 130$	-

Fig. 5b

【 図 6 】

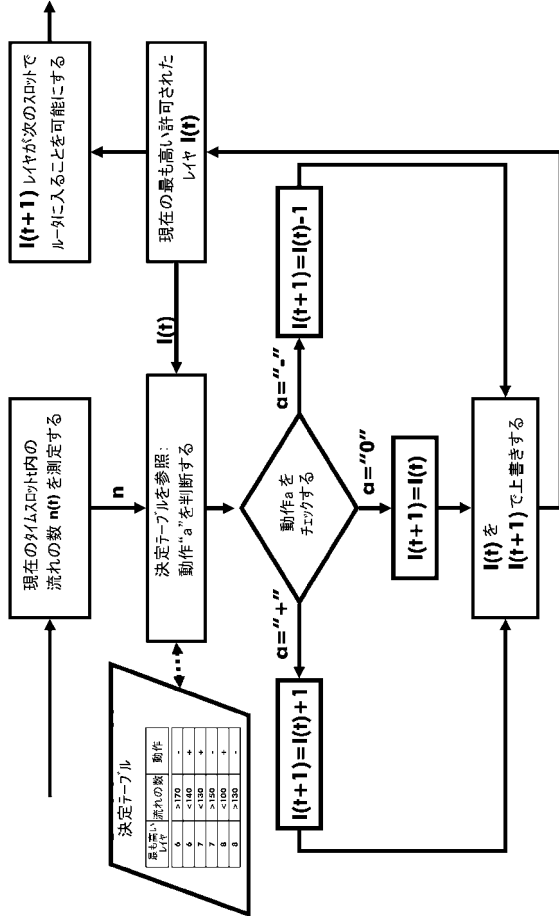


Fig. 6

【 図 7 】

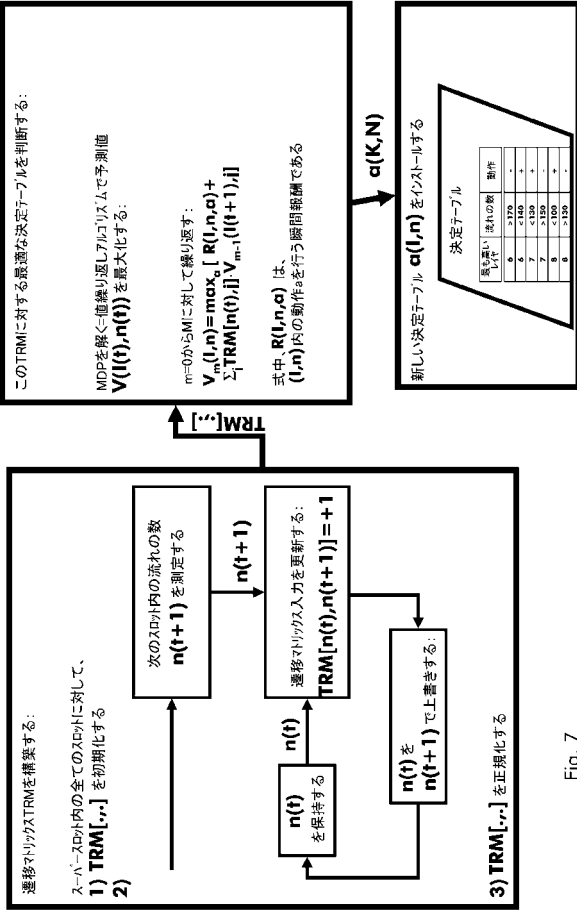


Fig. 7

【 図 8 】

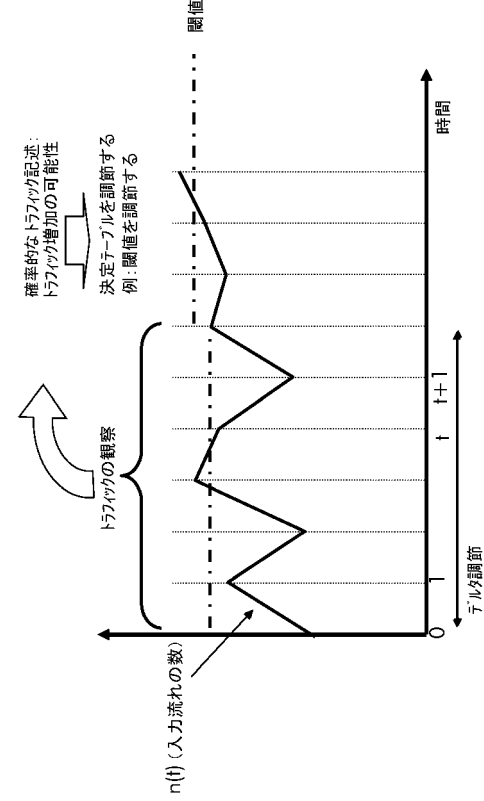


Figure 8

【 図 9 】

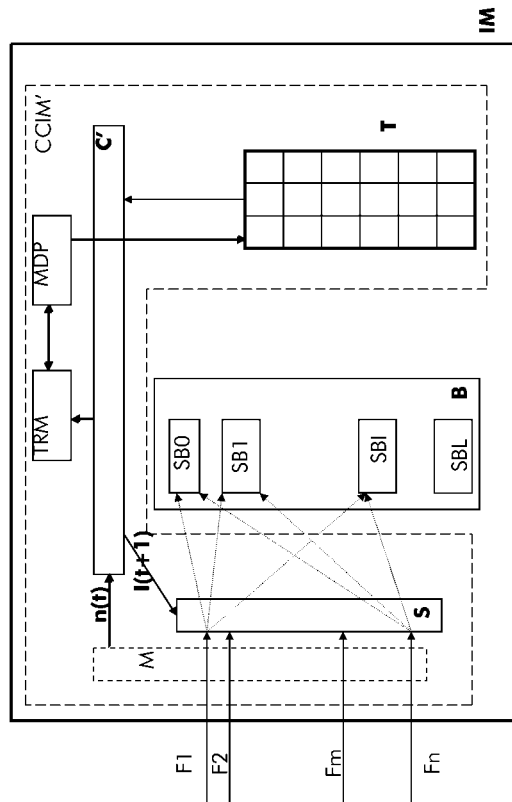


Fig. 9

【 図 10 】

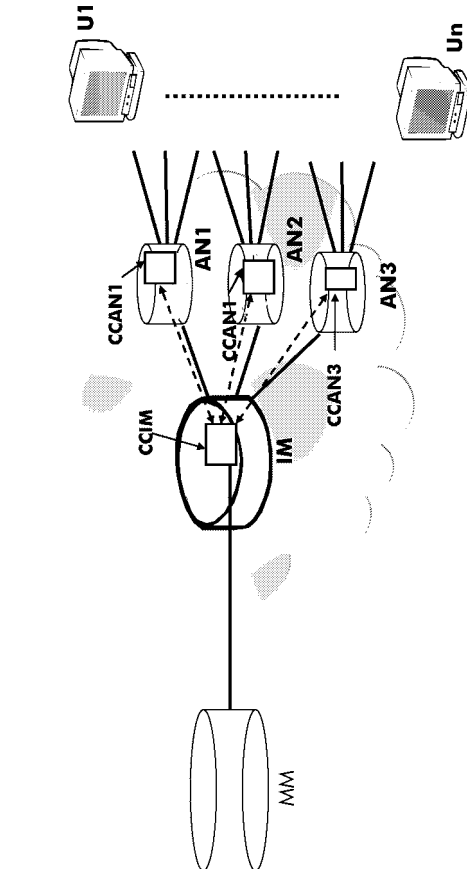


Figure 10

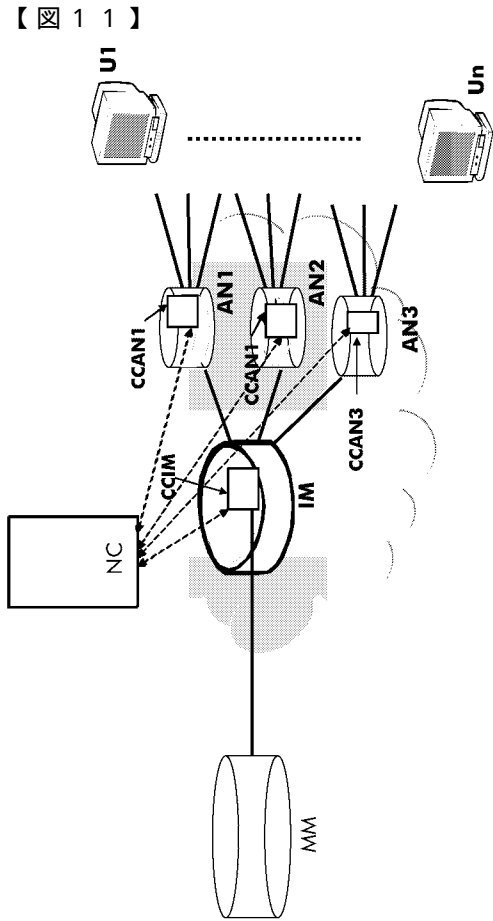


Figure 11

---

フロントページの続き

(72)発明者 ホン, ドユイ

フランス国、エフ - 9 4 1 2 0 ・ フォントネイ・ス・ボワ、リュ・デ・アルエット・1 5

審査官 中木 努

(56)参考文献 特開2005 - 204157 (JP, A)

特開2005 - 064970 (JP, A)

特開平07 - 264197 (JP, A)

特開2006 - 060448 (JP, A)

特表2007 - 504694 (JP, A)

特開2003 - 078561 (JP, A)

(58)調査した分野(Int.Cl., DB名)

H04L 12/00 - 12/955