

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第3813696号

(P3813696)

(45) 発行日 平成18年8月23日(2006.8.23)

(24) 登録日 平成18年6月9日(2006.6.9)

(51) Int. Cl. F I
 HO4L 12/28 (2006.01) HO4L 12/28
 HO4Q 3/00 (2006.01) HO4Q 3/00

請求項の数 1 (全 20 頁)

(21) 出願番号	特願平9-172615	(73) 特許権者	596170170
(22) 出願日	平成9年6月27日(1997.6.27)		ゼロックス コーポレーション
(65) 公開番号	特開平10-215258		XEROX CORPORATION
(43) 公開日	平成10年8月11日(1998.8.11)		アメリカ合衆国 コネチカット州 スタン
審査請求日	平成16年6月28日(2004.6.28)		フォード、ロング・リッジ・ロード 80
(31) 優先権主張番号	020643		O
(32) 優先日	平成8年6月27日(1996.6.27)	(74) 代理人	100079049
(33) 優先権主張国	米国 (US)		弁理士 中島 淳
(31) 優先権主張番号	872906	(74) 代理人	100084995
(32) 優先日	平成9年6月11日(1997.6.11)		弁理士 加藤 和詳
(33) 優先権主張国	米国 (US)	(72) 発明者	ジョセフ ビー. ライルズ
			アメリカ合衆国 94043 カリフォル
			ニア州 マウンテン ビュー ローラ
			ー 2403

最終頁に続く

(54) 【発明の名称】 パケット交換通信システム

(57) 【特許請求の範囲】

【請求項 1】

パケット交換通信システムにおいて、パケットの複数の時間多重直列フローを各ネットワークトラフィックコントラクトに成形するトラフィック成形装置は、前記フローの各々に対する関連するトレランスを有するピークパケット出力率を確立する少なくとも一つのコントラクトを含み、前記トラフィック成形装置は、

アクティブフローのパケットを構成するキューイングメカニズムを有し、前記キューイングメカニズムは、キュー順序の先頭の最も古い未決パケットに従った各キューの前記一つのフローを含み、

作業維持キュー及び非作業維持カレンダーキューを有するスケジューリングメカニズムを有し、前記スケジューリングメカニズムは、前記キューイングメカニズムに結合し、(i) 前記一つのフローのパケットが前記トラフィックコントラクトに適合するレートで前記トラフィック成形装置によって出力される場合に、前記作業維持キューに前記一つのフローのパケットをキューに入れ、(ii) 前記一つのフローのパケットが前記トラフィック成形装置によってエミットされるレートがコントラクトに確立されたピークレートに対してトレランスを越えた場合、前記フローのコントラクトに確立されたピークパケットレートよりも大きくないレートでエミッションの前記非作業維持カレンダーキューに前記一つのフローのパケットをスケジューリングする、
 パケット交換通信システム。

【発明の詳細な説明】

10

20

【 0 0 0 1 】

【 発明の属する技術分野 】

本発明は、パケット交換式通信ネットワーク、特に指定されたトラフィック記述子に適するために、そのようなネットワークまたはネットワーク要素内のキューイングポイントにおいて時間多重パケットのフローを生じさせるトラフィック成形に関する。

【 0 0 0 2 】

【 従来技術 】

A . トラフィックコントラクト / 定義

パケット交換式通信ネットワーク上で現在動いている殆どのアプリケーションは、ネットワークから偶然得る帯域幅が如何なるものであっても満足して動作できる。何故なら、それらのアプリケーションでは「弾性」帯域幅が条件となっているからである。これらのアプリケーションをサポートするこの種のサービスは、インターネットの世界では「ベストイフォート」サービスとして、または広帯域 I S D N / A T M の世界では「アベイラブルビットレート」(A B R) として知られている。

10

【 0 0 0 3 】

しかし、拘束 (bounded) ジッター、即ち拘束パケット遅延変動 (一般に A T M の関連ではセル遅延変動と呼ばれる) を提供するネットワークサービスに対しての要求が増している。例えば、この種のサービスは、回路エミュレーションおよび映像などのリアルタイムなアプリケーションにおいて必要となる。インターネットがこのような要求に答えるか否かまたどの様にして応えるかは明確ではないが、広帯域 I S D N / A T M の分野では、ユーザー - ネットワーク折衝型トラフィックコントラクトの概念を導入することで応えている。

20

【 0 0 0 4 】

知られているように、ユーザー - ネットワーク A T M コントラクトは、トラフィックのパラメータ、トレランス、およびサービス要求の質などを含むトラフィック記述子によって定義される。関連するトラフィックパラメータの各々について適合定義 (conformance definition) が特定される。したがって、A T M サービスは、サービスの質 (Q o S) の目標値と多重方式の異なる組み合わせをサポートするために、これらのトラフィックパラメータとそれらに対応する適合仕様を用いることができる。

【 0 0 0 5 】

部分的にオーバーラップする A T M トラフィッククラスは、国際標準化機構 (I T U - T) の通信標準化部門と、A T M フォーラムによって定義されている。幾つかの例では、本質的に同一の属性 (attribute) を有するトラフィッククラスに対してこれらの 2 つのグループは異なる名称を付けており、下記の名称変換テーブルは現在使われている等価の対応名称を示す。

30

【 0 0 0 6 】

ITUトラフィッククラス	ATMフォーラムトラフィッククラス
ABR	ABR
決定ビットレート (Deterministic Bit Rate: DBR)	一定ビットレート (Constant Bit Rate: CBR)
統計的ビットレート (Statistical Bit Rate: SBR)	可変ビットレート (Variable Bit Rate: VBR)
対応する名称はないが、検討中	リアルタイム可変ビットレート (Real time Variable Bit Rate: rt-VBR)
対応する名称はないが、検討中	非指定ビットレート (Unspecified Bit Rate: UBR)

【 0 0 0 7 】

仮想回線 (V C) コネクションまたは仮想パス (V P) コネクションのための A T M サービスコントラクトは、コネクションのサービスレートを記述するパラメータを複数含むことができる。これには、ピークセルレート (P C R)、維持可能 (sustainable) セルレート (S C R)、固有 (intrinsic) バーストトレランス (I B T)、最少セルレート (M C R) が含まれる。これらのパラメータの全てが個々のコネクションまたは個々のサービスのクラスに関連があるわけではないが、それらがサービスコントラクトの黙示的または明示的に指定された要素である場合、それらは重視しなければならない。V C コネクションが、下記の議論の中心となっているが、V P コネクションもそのように特定できることが分かるであろう。A T M コネクションのためのデータ転送単位は、通常「セル」と呼ばれる。しかし、本明細書では、データ転送単位を表すのに時として「パケット」という語が使用されている。これは、「パケット」という語は、本革新技術のより広い概念の幾つかにふさわしいより一般的な用語であるからである。

【 0 0 0 8 】

I T U - T 勧告 1 . 3 7 1 で指定されたジェネリックセルレートアルゴリズム (G C R A) は、トラフィック記述子との適合についてパケットまたはセルのフローをテストするのに非常に適している。このようなテストを行うために、G C R A は、出力 (emission) 間隔 (即ち、フローレートの逆数) とトレランス についての仕様を必要とする。実際には、このトレランスは、コネクション、コネクションセットアップパラメータ、またはサービスの種類などの種々の要因に依存するであろう。後で分かるように、G C R A はブル関数として採用でき、コネクション上の固定サイズのパケットまたはセルのフローの場合、G C R A (出力間隔、トレランス) は、もしフローがピークレートに従っていれば偽 (false) であり、フローが最少レートに従っていれば真 (true) である。例えば、もし G C R A (1 / P C R , P C R) が偽であれば、セルのソースは P C R に従う。同様に、もし G C R A (1 / M C R , M C R) が偽であれば、コネクションまたはフローは M C R に従う。後で分かるように、「出力間隔」は「セルレート」の逆数である。

【 0 0 0 9 】

D B Rトラフィックコントラクトは、コネクションが行われている間、一定量の帯域幅をコネクションに対して継続的に利用できることを予測してコネクションを確立するソースには適切である。従って、ネットワークがD B Rコネクションに委ねる帯域幅は、P C R値によって特徴づけられる。さらに、そのようなコネクションでのセルまたはパケットのフローは、もしそれが $G C R A (1 / P C R , P C R)$ に適合していれば、トラフィックコントラクトを遵守している。一方、S B Rトラフィックコントラクトは、S C Rと $I_{B T}$ およびP C Rと $p_{C T}$ の通知された選択を考慮に入れた既知のトラフィック特性を有するアプリケーションに適している。S B Rまたは $r t - S B R$ フローは、もしそのフローが $G C R A (1 / P C R , P C R)$ だけでなく、 $G C R A (1 / S C R , I_{B T})$ にも適合する場合、そのトラフィックコントラクトを遵守している。

10

【 0 0 1 0 】

前述したように、A B Rトラフィックコントラクトは、予約されていない帯域幅を使用した結果として生じる情報転送レートにおける動的な変化を許容できるアプリケーションに適している。P C RとM C Rは、そのようなコネクションを確立するソースによって指定され、これらのパラメータはネットワークとの折衝の対象となる。したがって、A B Rコネクションについて利用できる帯域幅はM C R（零でもよい）と、定義された割り当てポリシーにより、予約されていない帯域幅をA B Rコネクションの間で共用することによって生じる可変セルレートとの和である（即ち、ソースがその指定されたM C Rを上回って受信する帯域幅は、折衝されたP C Rだけでなく、ネットワークポリシーにも依存する）。

ネットワークからのフィードバックにより、ソースアプリケーションは、A B Rコネクションに対してセルまたはパケットを送るレートを動的に調整することができる。A B Rフローは、 $G C R A (1 / M C R , M C R)$ に適合していれば、常にトラフィックコントラクトを遵守しており、 $G C R A (1 / P C R , P C R)$ に適合していなければ、常に非遵守である。M C RとP C Rの間の領域における適合性は、A B Rフィードバックに依存し、従って動的に決定される。

20

【 0 0 1 1 】

U B Rトラフィックコントラクトは、A B Rコントラクトに類似しているが、U B Rコントラクトは、M C Rの仕様に順応できず、また動的な適合定義を有しない。したがって、U B Rフローは、もしそれが $G C R A (1 / P C R , P C R)$ に適合していれば、トラフィックコントラクトを遵守している。

30

【 0 0 1 2 】

B . トラフィック成形

I T U - T勧告 1 . 3 7 1 は、トラフィックを下記の点に関し、トラフィック記述子に適合させるためにネットワーク要素でのトラフィックの再成形の可能性について言及している：

「トラフィック成形は、Q o Sの目標値を満たしながらネットワーク効率を改善するため、または後続のインターフェースでの適合を確保するために、V C CまたはV P C上のセルストリームのトラフィック特性を変更して、これらのトラフィック特性の要求される変更を達成するメカニズムである。トラフィック成形は、A T Mコネクション上のセルシーケンスの完全さを維持しなければならない。成形はセルフローの特性を変更し、その結果、セルの平均転送遅れを増加させるものである。

40

【 0 0 1 3 】

トラフィック成形の例は、ピークセルレートの低減、バースト長さの制限、セルの時間間隔を適切にすることによるC D Vの低減、およびキューサービス方式である。

【 0 0 1 4 】

トラフィック成形を行うか否か、またそれをどこで行うかはネットワークオペレータの選択による。例えば、ネットワークオペレータは好適なU P C / N P C機能とともにトラフィック成形を行うことを選択できる。

【 0 0 1 5 】

50

分離したセルフローまたは集合的セルフローに関してトラフィック成形を行うことはオペレータが選べるオプションである。

【 0 0 1 6 】

したがって、如何なる A T M コネクションに対してもトラフィック成形を行うことができる。

【 0 0 1 7 】

ネットワークオペレータ/サービスプロバイダが利用できるオプションには以下のものがある：

a . 成形なし

・如何なる成形機能も用いることなく、ネットワーク出口における適合性を確保しながらネットワーク入口における適合セルの任意のフローを許容できるようにネットワークの寸法又は次元(dimension)を決める。

b . 成形あり

・入口における適合セルの任意のフローが、Q o S 目標値を満たしながらネットワークまたはネットワークセグメントによって運ばれるようにネットワークの寸法を決めるとともに動作させ、出口における適合テストを満たすようにトラフィックの出力成形を行う。

【 0 0 1 8 】

・ネットワークまたはネットワークセグメントの入口におけるトラフィックを成形し、Q o S 目標値およびネットワークまたはネットワークセグメントの出口におけるその後の適合テストを満たすとともに、成形によって得られたトラフィック特性にしたがってリソースを割り当てる。

【 0 0 1 9 】

また、ソースによってまたは U N I において発生されるセルが、使用される A T C に関する折衝トラフィックコントラクトに適合していることを確実にするために、顧客の装置内、またはソースにおいてトラフィック成形を使用してもよい(5 . 5 章参照)。」 I T U - T 勧告 1 . 3 7 1、 6 . 2 . 5 章。

【 0 0 2 0 】

C . リアルタイムおよび非リアルタイムコネクションのためのスケジューリング/現存のツールおよび技術

知られているように、もし「ベストイフォート」のインターネットサービスまたは A B R A T M サービスを採用するアプリケーションの間で帯域幅が「フェア」に分割されない場合、種々の好ましくない現象が生じるであろう。

【 0 0 2 1 】

現在、殆どの A T M スイッチでは F I F O キューイングが実施されている。しかし、F I F O キューイングは、A B R トラフィックに使用した場合、異常な挙動を示す。また、F I F O は、正しく行動しているユーザーを誤って行動しているユーザから保護できない(アイソレーションを提供しない)。このような欠点のため、重み付けフェアキューイングなどの非 F I F O キューイングメカニズム、またはラウンドロビンのようなフェアキューイングに近似した方式がしばしば提案される。

【 0 0 2 2 】

非弾性的帯域幅を要求する種類のサービスは、時として、データが拘束ジッター(即ち、セルまたはパケットの遅延変動が制限されていること)を伴ってネットワークを介して伝送されることを要求する。知られているように、リアルタイムストリームに対して拘束ジッターを提供するために重み付けフェアキューイングを使用することができる。さらに、仮想クロックおよび自己計時型フェアキューイングの密接に関連したメカニズムを使用するシステムのために試験済の遅延制限を提供できる。

【 0 0 2 3 】

したがって、弾性(ベストイフォート/A B R)サービスと非弾性(またはリアルタイム)サービスの何れもがフェアキューイングおよび関連するアルゴリズムを使用することによって利益を受けることが知られている。

10

20

30

40

50

【 0 0 2 4 】

1. 重み付けフェアキューイングおよび仮想クロック

フェアキューイングおよび関連のアルゴリズム（例えば、フレームベースのフェアキューイング、欠陥ラウンドロビンなど）はパケットまたは他のデータ転送単位（例えば、ATMセルはここでの議論においてはパケットである）のシーケンスに対して作用する。ATMの場合、これらのシーケンスはVCIまたはVPIの何れかによって認識され、この認識に適したインターネットのプロトコルは<IPアドレス、プロトコル、ポート>トリプル（IPv4）またはフロー識別子（IPv6）に基づくものである。自己計時型重み付けフェアキューイングおよび仮想クロックの何れにおいても、パケットはタイムスタンプによって順序づけられる（ソートされる）（ラウンドロビンなどの方式はタイムスタンプによってパケットの順序づけに対して近似を行う）。このタイムスタンプは仮想終了時刻を表し（またはパケットの仮想開始時刻を同様に表す）、開始時刻を得るとともに、パケット長に、特定のパケットシーケンスに対しての帯域幅の割り当てを表す重みを掛けることによって得られるオフセットを加えて計算される。

10

【 0 0 2 5 】

より詳細には、仮想クロックの場合、仮想終了時刻は下記のように計算される。

【 0 0 2 6 】

$$VT(f, 0) = 0$$

$$VT(f, j+1) = \max\{\text{到着}(f, j+1), VT(f, j)\} + \text{長さ}(f, j+1)/\text{レート}(f) \quad (1)$$

但し、 $VT(f, j)$ はフロー（仮想回線） f のパケット j に関連する仮想終了時刻であり、

20

到着 (f, j) はフロー f のパケット j の到着時刻であり、

長さ (f, j) はフロー f のパケット j の長さである。

【 0 0 2 7 】

自己計時型重み付けフェアキューイングは、下記式に従って仮想終了時刻を指定する。

【 0 0 2 8 】

$$VT(f, 0) = 0$$

$$VT(f, j+1) = \max\{\text{システム仮想時刻}, VT(f, j)\} + \text{長さ}(f, j+1) * \text{重み}(f)$$

(2)

30

但し、システム仮想時刻は、パケット $(f, j+1)$ が到着する時刻において提供される（出力される）パケットに関する仮想時刻である。

【 0 0 2 9 】

ATMの場合、セルが固定サイズ（例えば、53バイト長）であるため、パケット長は一定である。したがって、式(1)および式(2)の最右項は、フロー毎の定数になる。仮想クロックの場合、式を下記のように単純化できる。

【 0 0 3 0 】

$$VT(f, j+1) = \max\{\text{到着}(f, j+1), VT(f, j)\} + \text{定数}(f) \quad (3)$$

【 0 0 3 1 】

自己計時型重み付けフェアキューイングの場合、式を下記のように単純化できる。

40

【 0 0 3 2 】

$$VT(f, j+1) = \max\{\text{システム仮想時刻}, VT(f, j)\} + \text{定数}(f) \quad (4)$$

【 0 0 3 3 】

即ち、仮想クロックまたは自己計時型重み付け型フェアキューイングを実施するATMキューイングポイントは下記のステップを行う。

【 0 0 3 4 】

1) VCについての現在の仮想時刻 (a) と、セルの到着時刻 $i)$ とシステム仮想時刻 $ii)$ の何れか (b) との間の最大値を計算し、

2) 上記ステップ1の結果に、帯域幅のVCの占有率を表すVC毎の定数を加え、

3) ステップ1と2によって指定した仮想タイムスタンプの値が増加する順序でセルを提

50

供（それらを送信）する。

【 0 0 3 5 】

2. 優先権

一つのトラフィッククラスに他のものより高い優先権を与えるということは、もし高い優先順位を有するトラフィッククラスが転送の準備のできたセル群を有していれば、これらのセル群は、常に優先順位の低いトラフィッククラスのセルに優先して伝送される。

【 0 0 3 6 】

優先権のメカニズムは先制的 (preemptive) であることも非先制的であることもできる。この用語はオペレーティングシステムの文献から由来している。非先制的優先権メカニズムは、スケジューリング時刻において対象物 (オペレーティングシステムの世界ではプロセスであり、ATMの世界ではVC) に対して優先順位を与え、対象物はそれが提供されるまでこの優先順位を維持する。他方、先制的優先権メカニズムは、対象物が提供されるのを待っている間に対象物の優先順位を変えることができる。例えば、先制システムでは、「このVCを優先順位3でスケジュールする」と言うことはできるが、もしそれが200マイクロ秒の間に提供されない場合には、その優先順位が2に上げられる。

【 0 0 3 7 】

3. 作業維持および非作業維持キューイング

「作業維持 (work conserving)」という用語は、作業が作られることもないし破壊されることもない任意のキューイングシステムを示すために時々使用される。この用語と一致して、待機セルが与えられた時、常にセルを出力リンクに伝送するスイッチは、「作業維持スイッチ」である。純粋なFIFOを採用するスイッチ、重み付けフェアキューイングまたは仮想クロックスケジューリングアルゴリズムは全て作業維持型である。これに対し、非作業維持型スイッチは、セルが転送のために待機していても、セルを送信しないことを選択できる。後で分かるように、これを行う方法は、特定のセルを伝送する前に現時刻がそのセルに関連したタイムスタンプに等しいか大きくなるまで待つようにスイッチをプログラムすることである。

【 0 0 3 8 】

作業維持スイッチは、伝送リンクを完全に利用しようと試みるが、必ずしもバーストを除いたり、阻止したりしない。これに対し、非作業維持スイッチは、より厳しい適合テスト (即ち、より小さい を有するGCRA) を満たすようにトラフィックを再成形するためにセルを戦略的に遅延させることができる。さらに、所定のコネクションに指定された量のバッファリングが割り当てられるだけの非作業維持スイッチは、割り当てられたバッファ空間をオーバーフローするセルを捨てるかもしくはそれにタグを付けることによってポリシー (policing) 機能 (ITUではUPC/NPCと呼ぶ) を行う。非作業維持キューイングシステムの例は、仮想時間は実時間よりも速く過ぎる (立ち往生するか、非作業維持状態になる) ことが許されない仮想クロックアルゴリズムを改造したストールド (stalled) 仮想クロックである。

【 0 0 3 9 】

4. カレンダーキュー

カレンダーキューは、アクションを時間順に並べたリストで、各アクションは、実時刻がそのアクションに関連する時刻と等しいかそれを越えた時にキューから取り出されて実行される。時間間隔が制限されたカレンダーキューは、「タイムホイール (time-wheel)」または「タイムライン (time-line)」として知られている線型アレーとして表すことができる。タイムホイールはバケットに対する事象をポインタに対して相対的に割り当てるもので、ホイールサイズを算術モジュロとして使用してバケットインデックスが計算される。これらのデータ構造は、キューイングメカニズムとして文献で良く知られている。タイムホイールでは、絶対時刻は、現時刻 (実時刻) に対するオフセットとして表され、アレー中の各要素は、それらが属するバケットに対して指定された時刻に実行すべき一つ以上のアクションを (典型的には、リンクされたリストの形態で) 含むバケットである。そのようなタイムホイールのバケットの何れのものも空であることができる、即ちそれに関連する事

10

20

30

40

50

象を有しなくてもよい。

【0040】

各タイムホイールについて、二つの重要な時間 $t_{earliest}$ と t_{latest} があり、これらはアレー中のアクティブエントリーについてのヘッドポイントおよびテイルポイントに対応し、 $t_{earliest}$ は、提供すべき次のエントリー（例えば、パケットまたはセル）の時刻であり、 t_{latest} は、スケジュールされた事象を含む最新の（時間的に最も遠い）パケットに関連する時刻である。 $t_{earliest}$ と t_{latest} の間の差はタイムホイールの長さ b から 1 を引いた長さを越えることはできない。これは、時間がモジュロ b を維持していることにより、そして、 $b - 1$ を越えるオフセット（仮想クロックまたは重み付けフェアキューイングにおいてそれぞれレートまたは重みの何れかを掛けたパケット長）が無いことを 10
 確実にすることによって確実にされる。OC-3 の速度（149.76 mbps - SONET ペイロードレート）で動作する ATM リンクの場合、リンク上に秒当たり約 353208 セルが存在する。したがって、もし 64 Kbps（音声電話の速度）のフロー（AAL タイプ 1 が使用される場合には、毎秒約 174 セル）が、サポートすべき最も遅い速度のコネクションで、最も遅い速度に対する最も早いサポートされた速度の割合は 2029、切上げると 2^{11} となる。この割合は、実時刻を計算する間に加算される最大オフセットである。したがって、長さが 2030（2 の冪乗への切上げを考慮にいれると 2048）のタイムホイールは、64 Kbps から完全な OC-3 リンクの速度までの間の速度の回線に関連する実時刻をコード化するのに充分である。

【0041】

タイムホイールアレーの長さは、アレー要素が 2 以上のタイムオフセットを含むことができるようにすることによって減少できる。例えば、もし上記のタイムホイールが 2048 要素から 256 要素に減少された場合、各パケットはその中にマップ化した 8 個のタイムオフセットを有するであろう。複数オフセットに及ぶ単一のパケット内のアクションは、狂った順番で行われるかも知れないが、パケットの間ではアクションは順番どおりになるであろう。これにより、カレンダーキュー内のアクションの順序付けの精度が低下するという犠牲を払って、そのようなタイムホイールに割り当てる必要があるメモリの量を減少できる。

【0042】

D. 複数出力チャネル上での時間多重フローのためのトラフィック成形 30
 時間多重化されたパケットまたはセルのフローをそれらのトラフィックコントラクトに適合させるのに必要な何れのトラフィック成形は、異なった出力チャネルのためのフローを相互に分離するのに必要な切り替えまたはルーティング動作が完了した後で実施するのが好ましい。これにより、マルチプレクサのスループット効率を最適化できる。

【0043】

しかし、以前の出力待機 ATM スイッチは一般に、FIFO（先入れ、先だし）出力バッファを採用していた。これらのバッファはそれらを通ずるフローの何れのものの再成形の制御に関与できない。その代わりに、これらのバッファから出力される VC 毎の時間多重フロー（per-VC time multiplexed flows）は、本質的にそれらの中にロードされる入力フローの時間多重複合物である。勿論、これらの出力フローは、バッファの固有待ち時間 40
 のために入力フローに対して時間的に遅延される。また、もし異なるフローのデータ転送制限の間でスケジューリングの衝突が発生した場合にはこれらの出力フローの一つ以上のもののセル遅延変動（CDV）が増加するであろう。何故なら、これらの干渉は、所謂「伝送衝突」を引き起こすからである。

【0044】

後で分かるように、CDV の増加は、比較的トレランスが厳しい DBR トラフィックのようなトラフィックの場合、特にやっかいなものである。したがって、もしソースとデスティネーションの間の各ホップが、上記の形式の単純な FIFO 出力キューを含む場合、その指定されたトレランス内での適合性を確保するために、この CDV に敏感なトラフィックが作ることができるホップの数を制限することが必要になるであろう。 50

【 0 0 4 5 】

したがって、時間多重出力の送出のために複数入力から複数出力へトラフィックを送る、ATMスイッチまたはその他のルーターのためのより効率的でより効果的なトラフィック成形メカニズムとプロセスに対するニーズがあることは明らかである。

【 0 0 4 6 】

【 発明が解決しようとする課題 】

本発明は、識別されないビットレートサービスのフロー毎出力待機ルーチングメカニズムにおけるレート成形を提供する。

【 0 0 4 7 】

【 課題を解決するための手段 】

本発明の請求項 1 の態様では、パケット交換通信システムにおいて、パケットの複数の時間多重直列フローを各ネットワークトラフィックコントラクトに成形するトラフィック成形装置は、前記フローの各々に対する関連するトレランスを有するピークパケット出力速度を確立する少なくとも一つのコントラクトを含み、前記トラフィック成形装置は、アクティブフローのパケットを構成するキューイングメカニズムを有し、前記キューイングメカニズムは、キュー順序の先頭の最も古い未決パケットに従った各キューの前記一つのフローを含み、作業維持キュー及び非作業維持カレンダーキューを有するスケジューリングメカニズムを有し、前記スケジューリングメカニズムは、前記キューイングメカニズムに結合し、(i) 前記一つのフローのパケットが前記トラフィックコントラクトに適合する速度で前記トラフィック成形装置によって出力される場合に、前記作業維持キューに前記一つのフローのパケットをキューに入れ、(i i) 前記一つのフローのパケットが前記トラフィック成形装置によってエミットされる速度がコントラクトに確立されたピーク速度に対してトレランスを越えた場合、前記フローのコントラクトに確立されたピークパケット速度よりも速くない速度でエミッションの前記非作業維持カレンダーキューに前記一つのフローのパケットをスケジューリングする、ことを含む。

【 0 0 4 8 】

【 発明の実施の形態 】

A . 代表的な環境

図面、特に図 1 を参照すると、ATMスイッチ 2 1 の入力および出力ポートは、典型的には、U t o p i a 2 インターフェースを介して 1 個以上の物理層に結合され、第 2 の好適なインターフェースを介してスイッチ制御プロセッサモジュール 2 2 に接続されている。これにより、スイッチ 2 1 は、接続された任意の物理層とデータおよびコントロールセルを交換し、また制御プロセッサモジュール 2 2 と制御セルを交換する。この標準的なプラクティスと調和して、通信チャネルは一方向性であり、したがって、双方向通信には一組のチャネルが必要である。

【 0 0 4 9 】

スイッチ 2 1 は、入力キューから V C 毎出力キューへデータおよび制御セルをスイッチングするために、スイッチングファブリック 2 4、ファブリック制御モジュール 2 5 およびリザベーションリング 2 6 を含んでいる。これらのキュー中のセルはデータメモリ 2 7 中のデータパスに記憶され、これらの入力および出力キューはキュー制御モジュール 2 8 によって管理される。典型的には、データメモリ 2 7 は、最大で約 1 2 0 0 0 セルを記憶できる容量である。データおよび制御セルフローのコネクションレコードは、制御プロセッサモジュール 2 2 へのルーチングのための、レート基準エンジン / トラフィックマルチプレクサ 3 1 によって捕らえられる所定の形式の制御セルと共に制御メモリ 2 9 内の制御パス中に記憶される。制御 R A M 2 9 は、最大で約 8 2 0 0 のコネクションレコードと 6 4 K のセルレコードを記憶できることが好ましい。制御プロセッサモジュール 2 2 とスイッチ 2 1 との間の相互作用は本発明の範囲を越えているので、本明細書では記載しない。しかし、ATMスイッチの設計に詳しい人であれば、制御プロセッサは主にコネクションの確立および終了、並びに O A M (Operation and Maintenance) 機能を実施する責任がある。

【 0 0 5 0 】

スイッチ 2 1 のデータパスは所定のレート、例えば 4 0 M H z で同期的にクロック動作される。しかし、従来の同期パイプライン設計のプラクティスに合わせ、このクロック信号の位相は、一つのパイプラインのステージから次のステージへと転送される前に落ちつくために適切な時間をデータに与えるために、データ経路に沿った異なった点で異なった量だけ（図示しない手段により）遅延される。

【 0 0 5 1 】

標準的なプラクティスに従い、デスティネーションと通信することを望むソースは、スイッチ 2 1 が存在する A T M ネットワークとの折衝を、ネットワークに対して S E T U P メッセージを送ることによって開始する。このメッセージは、デスティネーションを識別し、要求された接続のための関連トラフィックパラメータの全てを明示的または黙示的に指定する。もしネットワークが、これらのトラフィックパラメータ（または、ソースが受け入れることを希望しているパラメータの変形版）によって定義されるトラフィックコントラクトに委ねるための準備ができている場合には、ネットワークは S E T U P メッセージをデスティネーションへ送る。そして、デスティネーションが、トラフィックコントラクトの条件に従ってソースからメッセージトラフィックを受信する準備ができていれば、デスティネーションは C O N N E C T メッセージをソースに返す。この C O N N E C T メッセージは、トラフィックコントラクトに適合するセルフフローのための指定された仮想パス（V P）内の指定された仮想回線（V C）上に接続が確立されたことを確認するものである。I T U - T 勧告 0 . 2 3 9 1 および A T M フォーラム U N I 4 . 0 仕様を参照のこと。これらの信号伝達プロトコルを発生させることなく、供給することにより「永久」仮想接続を確立できる。

【 0 0 5 2 】

接続が確立された後データセルが流れ始める。図 2 に示すように、セルの形は、スイッチ 2 1 が行う動作のためにスイッチ 2 1 を通過するにつれ変化する。マルチキャストのためにセルを複製してもよいが、下記の議論は、不必要に複雑になることを避けるためにユニキャストの動作に限定する。

【 0 0 5 3 】

図 2 に示すように、スイッチ 2 1 が受信する各上りのセル 4 1 は、V P インデックスと V C インデックスを含むヘッダを有している。これらのインデックスは、接続の一つのホップについての固有アドレスを定義するために組合わさる。接続は複数のホップから構成することができ、したがって次のフープのための V P および V C インデックスは、図 2 において 4 2 で示されているように、スイッチ 2 1 を通過する時にセルのヘッダに書き込まれる。

【 0 0 5 4 】

スイッチ 2 1 は、関連するフローのための接続レコードが存在する制御 R A M 2 9 内のアドレスを計算するために上りのセル（図 2 の 4 1）の V P および V C インデックスを採用する。典型的には、この接続レコードは、フローがスイッチ 2 1 から出ていく出力ポート（即ち、スイッチレベルでのデスティネーション）を認識するためのビットベクトル、粒状の優先権スケール（granular priority scale）上でのフローの相対的な優先順位を識別するための優先権インデックス、およびスイッチ 2 1 内のフローを唯一的に識別する回線インデックス（「C i r c u i t I n d e x」）を含んでいる。図 2 において 4 3 で示されているように、これらの接続パラメータはセルヘッダ内に書き込まれる。そして、セルがデータ R A M メモリ 2 9 へ書き込まれ、セルに対してのポインタが複数の F I F O 入力キューの適切な一つとリンクされる。なお、キューの選択は関連するフローの優先順位に基づいて行われる。

【 0 0 5 5 】

入力キュー内のキューセルのヘッドの相対的な優先順位が、各セルタイムの間に調べられ、最も高い優先順位を有するキューセルのヘッドが、次のアービトレーション（arbitration）セッションの間のアービトレーションのために選択される。さらに、キューセルの

10

20

30

40

50

優先順位のより低いヘッド（即ち、キューセルの任意の選択されていないヘッド）の優先順位は（図略の手段により）一定量ずつ増加され、これによって、次のアービトレーションセッションにおけるアービトレーションのためにセルが選択される可能性を増加させる。したがって、優先権が高い入力キューは、優先権が低いキューよりも単位時間当たりのスループットが大きい、優先権が低いキューは、キューセルのヘッドの優先順位が時間とともに増加するため、遅れが制限される。

【0056】

各アービトレーションサイクルは、スイッチ21の一セル時間を必要とし、したがってアービトレーションのために選択されるセルのルーティング情報は、アービトレーションに勝ったセルまたはセル群のペイロードがスイッチングファブリック24へ放出されるより1セル時間だけ前にリザベーションリング26に供給される。即ち、図2において44で示すように、リザベーションリング26とスイッチングファブリック24によって受信されるセルは、次のアービトレーションサイクルのためのセル（即ち「現在セル」）のヘッドから構成されており、これに前のアービトレーションサイクルでアービトレーションに成功したセル（即ち「前のセル」）のボディまたはペイロードが続いている。したがって、セルボディがファブリック24に到着すると、これらのセルをそれらの出力ポートデスティネーションにそれぞれ送るためにファブリックがファブリック制御装置25によって既に構成されている。

【0057】

図示の実施の形態では、アービトレーションおよびルーティングのためにセルが4ビット幅のニブルに分解される。この後（スイッチングプロセスのテストを行うために提供されるであろう「アイドルセル」を除き）、セルが再び組み立てられ、(a)スイッチ21の適切な出力ポートへの時間的にスケジュールされた転送のため、および/または(b)制御プロセッサモジュール22への転送のために、データパス、制御パス、またはその両方においてキューに入れられる。スイッチ21の出力ポートへのセルの時間的にスケジュールされた転送が本発明の中心であり、以下これについて更に述べる。一方、セルの分解と再組立、テストプロセス、および制御プロセッサ22のRM (Resource Management) およびOAM (Operation and Maintenance) セルとの相互作用は深く考える必要のない付随的なトピックである。

【0058】

図3を参照すると、スイッチ21はスイッチングファブリック24の出力側、即ち送信側でファンアウトしていることが分かるであろう。したがって、スイッチ21の一つの出力チャンネルのみが示されているが、このチャンネルは他のチャンネルを代表するものであることが分かるであろう。

【0059】

図示のように、スイッチングファブリック24からのセルボディおよびそれらの関連する回線インデックスを受け入れるためにフィルセルモジュール51を設けることが好ましい。スイッチングファブリック24の出力側の「有効セル時間」は、通常のセルタイムの「k」増速ファクターに対する割合によって決定される。したがって、例えば、もし通常のセル時間が113クロックサイクル/セルである場合、スイッチングファブリック24の出力側での有効セル時間は、 $k = 2$ であれば、56.5サイクル/セルである。

【0060】

有効なセルが受信された時、フィルセルモジュール51は、典型的に、セルをデータメモリ27へ書き込むためのデータ構造についてのリンクされ番号の付けられたフリーリスト52から出力されるセル構成を用いる。このため、フィルセルモジュール51は、要求に応じて、フリーリスト52の一番上からセル構造を読み出すためのフッチステートマシン53を含んでいることが好ましい。これにより、フィルセルモジュール51は、セルフロー制御ユニット55にセルの到着を知らせるために送る「到着」メッセージへ、セルについての回線インデックスと、データメモリ27内のセルの位置についてのポインタを挿入することができる。回線インデックスにより、フロー制御ユニット55は、制御メモリ

10

20

30

40

50

29内のコネクションレコードからセルが属するVCまたはフローを突き止めることができる。そして、セルフロー制御ユニット55はフローのトラフィック成形状態をチェックできる。フロー制御ユニット55がこれらの制御セルを認識し、またそれらがデータパス、制御パス、またはその両方においてキューに入れられているか否かを決定できるようにするためにOAM/RMリコグナイザ57を設けるのが有利である。

【0061】

トラフィックコントラクトコンプライアントフローのセルのためのメモリポインタは、セルフロー制御ユニット55がキュー制御ユニット58へ送る「addCell」メッセージに回答してVC毎のキュー内に入れられる。各addCellメッセージは、それが属するセルおよび関連するフローまたはVCのための回線インデックスを識別する。また、addCellメッセージは、セルがデータパス、制御パス、またはその両方においてキューに入れられるべきか否かを示す。セルが適切にキューに入れられると、新たにキューに入れられたセルを、それが属するVCについて将来行われるレート成形演算の間に考慮すべきことをフロー制御ユニット55に通知するために、キュー制御ユニット58がフロー制御ユニット55に「added」メッセージを返す。

10

【0062】

キュー制御ユニット58が、それぞれのキューにセットされている深さ制御の制限に関するVC毎キューの長さをモニターするのが有利である。これにより、制御ユニット58は、VC毎のキューが過剰に長くなった時、ABRフローに対してのコンゲッション制御動作を開始することができる。また、制御ユニット58は、トラフィックコントラクトを越えているフローを特定し、そのような適合しないフローのセルを削除するかまたは記録するための適切なポーリング機能(図示略)を動作させることができる。

20

【0063】

アドミッションコントローラ61は、キュー制御ユニット58から返される「added」メッセージを監視し、スケジューラ62に、空でないVC毎キューのためのキューセルのヘッドがスケジュールされた時刻に転送されるようにカレンダーキュー63上に組み入れる。スケジューラ62は、カレンダーキュー63上へのキューセルのヘッドの組入れを、それらのためにスケジューラ62が計算するそれぞれの仮想終了時刻 $VT(f, j+1)$ (または「仮想開始時刻」)に従って行うために、VC毎の仮想クロックを採用している。

【0064】

カレンダーキュー63は、スケジュールされたセルがスケジュールされた時刻よりも前に伝送のために放出されることを阻止するため、システムの「実時刻」と「現在時刻」を追跡する。即ち、スケジューラ62とカレンダーキュー63は、ストールド仮想クロックを提供し、このため伝送することがスケジュールされたセルは、システムの実時刻が各々の伝送時刻に達した時のみに、伝送のために放出される。図示のように、カレンダーキュー63によって伝送のために放出されたセルを有するコネクションは、伝送リスト65上の伝送の準備ができているセルを有するコネクションのリンクリストにリンクされる。

30

【0065】

カレンダーキュー63はフロー制御ユニット55に対して、任意の所定のコネクション上での伝送のためにセルを放出するか否かを知らせる。すると、フロー制御ユニット55は、もし所定のコネクションのためのVC毎のキュー上の次のセル(もしあれば)についてのリファレンス(即ち、キューセルの新しいヘッド)を要求し、アドミッションコントロール61に、このリファレンスをスケジューリングのためにスケジューラ62に対して認めるべきことを通知する。これにより、アドミッションコントロール61はカレンダーキュー63との閉ループ通信に効果的に関与し、スケジューリングを認めたキューセルのヘッドが、VC毎のキュー内の他の全てのセルの除外に対して認められることを確実にする。したがって、カレンダーキュー63は、一つ以上の時間制限されたタイムホイール、または「タイムライン」66を採用することによって実施できる。これらのタイムホイールのタイムスパンは、少なくともシステムがサポートするように設計された最低周波数のフローの周期と同じ程度にして時間ラップ(time wrap)によって引き起こされる曖昧さを阻止し

40

50

なければならず、好ましくは2倍とし、これにより相対時刻を2の補数演算を使用して比較できる。

【0066】

B. 固定ビット長データ転送リミットのフローの成形による指定されたピークフローレート

ストール仮想クロック伝送制御は、出力待機ルーチングメカニズムからのATMセルなどの固定ビット長データ転送単位の時間多重フローを、DBR/CBR ATMサービスのためのPCRなどの指定されたピークのデータ単位のフローレートに成形するのに非常に適していることが分かる。前述したように、所定の出力ポートに送られるフローのデータ転送単位は、送りだされた後、フロー毎キューに入れられる。これらのキューの先頭にあるデータ転送単位は、スケジューラ62によってタイムラインカレンダーキュー63に組み込むことがアドミッションコントローラ61によって認められる（他の全ての転送単位は除いて）。そして、スケジューラ62は、これらのキュー転送単位のヘッドに関してフロー毎仮想クロック演算を行い、それらの各々の理論的終了時刻 $V_T(f, j+1)$ または各々の理論的開始時刻に従ってカレンダーキュー63から放出するようにスケジュールする。

10

【0067】

実時刻は、成形された時間多重出力トラフィックが出力リンク71の帯域幅を本質的に満たすことができる速度でタイムライン63上で増加させるのが有利である。思い出されるように、スケジューラ62が各フローのキューメンバーのヘッドをマップ化することができる分解可能なタイムスロットの最大数は、これらのフローの最低許容周波数に対する最高許容周波数の比に基づく。したがって、実時刻がバケット毎にセルレートの論理的倍数で増加される。

20

【0068】

タイムライン63についてのリファレンスである現在の実時刻より前であるか等しい時刻を表すタイムスロットに存在するデータ転送単位は、伝送のために選ばれる資格があり、したがって、前述のようにして転送リスト65にリンクされる。しかし、タイムライン63の遅いタイムスロットに関連するタイムスロットに存在するデータ転送単位は、システムの実時刻がこれらのタイムスロットに到達するのに充分なだけ進むまで待ちの状態になる。ロールオーバーによる曖昧さを避けるために、タイムライン63は、以前にスケジュールされたデータ転送リミットに対する全てのリファレンスを、次のスキャンを見越して後からスケジュールされた転送単位に対しての任意のリファレンスが挿入される前に各タイムスロットから除去される。

30

【0069】

上記の構成は、適合するDBR/CBR ATMフローをそれらのトラフィックコントラクトによって指定されるPCRに効果的に成形するが、これらのフローのセル遅延変動(CDV)をそれらのトラフィックコントラクトの P_{CR} パラメータに適合させる上での助けにはならない。

【0070】

C. 相対CDVを最小化するための複数優先レベル

40

本発明によれば、異なった周波数のフローによってATMスイッチの出力ポートなどの多重化ポイントへ供給されるデータ転送単位に優先権を与え、周波数が高いフローのデータ転送単位に、これと衝突する可能性がある周波数の低いフローのデータ転送単位よりも高い伝送優先順位が与えられる。図3に示すように、この伝送優先権は、ストール仮想クロックスケジューリングメカニズム63などによるスケジュールが許可されたデータ転送単位を、これらのデータ転送単位が属するフローの周波数に基づき、優先順位で並べられた複数のタイムライン66a-66eまたは出力FIFOキューの一つまたは他のものへと導くことによって実施される。例えば、ATMスイッチの場合、5つの異なった周波数依存/サービス種類依存の出力優先順位で実施することが推奨される。この優先順位には、(1)出力リンクの最大レートの少なくとも1/16の折衝された出力レートを有する

50

フローからのセルに対しての最も高い優先順位、(2)出力リンクレートの $1/16 \sim 1/256$ の範囲の折衝された出力レートを有するフローからのセルに対しての2番目の優先順位、および(3)出力リンクレートの $1/256 \sim 1/4096$ の範囲の折衝された出力レートを有するフローからのセルに対しての3番目の優先順位が含まれる。低い2つの優先権は、ゼロでない折衝されたMCRレートを有するABRコネクションに対して、およびMCRレートが零のURBコネクションとABRコネクションに対してそれぞれ設定するのが好ましい。

【0071】

後から分かるように、本発明は、周波数の低いフローのCDVを実質的に増加させることなく、周波数の高いフローのCDVを効果的に減少させる。一般論として、許容できるCDVは、フローの折衝されたレートに依存して変化する。例えば、100個のセルタイムのCDVは、予期されるセル10個あたり1個の出力間隔に関しては非常に大きいですが、もし折衝された出力間隔がセル2029個に対して1個にすぎない場合には一般的に問題にならない。

10

【0072】

周波数が異なるフローのデータ転送単位またはセルの伝送をスケジューリングするためにカレンダーキューメカニズムを採用する場合、高周波数で優先順位の高いフローは、許容可能な低いCDVを達成するようにスケジューリングされ、かつ単一のセルタイムの精度に分解される必要があるが、低周波数/低優先順位の高いフローは、例えば、16セルタイムの精度により粗く分解できる。これは、カレンダーキュー63上のタイムスロットの数を減少できることを意味する。これにより、周波数の低いフローのキューセルのヘッドのスケジューリングには一般に必要な精度を多少犠牲にして、カレンダーキュー63を実現するのに必要なメモリの量を減少できる。

20

【0073】

名目上相違する一定周波数のフロー間での多重化ポイントにおける伝送の衝突の問題を解消するために提供される本発明による周波数に基づく優先順序付与の技術は、そのような多重化ポイントへセルまたはその他のデータ転送単位を供給するためのフロー毎出力キューを維持する作業を有するアプリケーションなど、フローの相対ジッターを減少させるための多くの異なるアプリケーションにおいて採用できる。

【0074】

D. ソース・ツー・デスティネーションABR制御ループを有するネットワークにおけるABRサービスについてのMCRおよびPCRパラメータのトラフィック成形
思い出されるように、標準ABRトラフィックコントラクトは明示的または黙示的に指定されたPCRおよびMCRパラメータを予期しており、MCRはゼロでもよい。 $GCR A(1/MCR, MCR)$ に適合するABRフローは有効であり提供される資格があるが、ネットワークはそのフローに対して、それに許可された帯域幅を与えていない。これに対し、 $GCR A(1/PCR, PCR)$ に違反するフローは適合しないものである。したがって、ATMスイッチ21などによってネットワークまたはネットワーク要素の出力に送られるABRコネクションまたはフローは、(1)MCRがゼロでない保証(non-zero MCR guarantee)を有するABRフローは、これらの保証を効果的に満足させるための適切な出力帯域幅を受け入れること、および(2)ABRの何れもそのような出力においてそれらのPCRコミットメントに違反しないことを確実にするよう成形することが有利である。複数の多重化ポイントを有するネットワークでは、この成形を多重化ポイントの各行うことが有用である。

30

40

【0075】

図4は、ABRフローまたはVCコネクションをネットワークまたはネットワーク要素の出力においてそれらのトラフィックコントラクトに適合させるために成形する一つの技術を示している。図示のように、それぞれのフローが出力サービスを要求するレートを監視し、最初に85においてフローが $GCR A(1/MCR, MCR)$ に適合しているか否かを判断し、86において $GCR A(1/MCR, MCR)$ に適合していないフローがG C

50

RA (1 / PCR , PCR) に適合しているか否かを判断する。これらの一連のテスト 85 と 86 の結果を反映する制御信号がスケジューラ 62 およびそれに関連するステアリングロジック 87 にフィードバックされ、各フローのデータ転送単位 (例えば、セル) が出力に供給される速度に対して適切なレート成形調整が行えるようにする。後で分かるように、この要求される状態は、全ての ABRV フローが、GCRA (1 / MCR , MCR) に適合しておらず (またはかろうじて適合している)、かつ GCRA (1 / PCR , PCR) に偽 / 適合であることである。なぜなら、これらのアルゴリズムは、それぞれ最低および最大の許容フローレートについてテストするからである。

【 0076 】

より詳しく述べると、GCRA (1 / MCR , MCR) に適合するフローは、その MCR 保証を満たすのに必要な出力帯域幅を受け取っていないであろう。したがって、そのようなフローが 85 において識別されたとき、スケジューラ 62 とそのステアリングロジック 87 は、その保証された MCR を幾分上回った速度での出力のために比較的優先順位が高いストール仮想クロックカレンダーキュー 88 上にそのフローのデータ転送単位 (例えば、セル) に対しての引き続いて受信されるリファレンスをスケジュールするためにセットアップされる。例えば、これらの引き続いて受信されるセルまたはパケットのリファレンスは、所定の加速ファクタによってフローについての MCR を乗算して決定される速度での出力のためにカレンダーキュー 88 上に組み込まれる。これにより、ネットワークは、これらの追加の多分「遅れている」フローに、それらの MCR 保証を満足するのに必要な追加の出力帯域幅を与える。

【 0077 】

GCRA (1 / MCR , MCR) 非適合 ABR フローは、それぞれ許容できる程度に低いかまたは許容できない程高い GCRA (1 / PCR , PCR) 適合または非適合レートでの出力サービスを要求しているか否かを判断するために 86 において更にテストされる。このテストの結果が他の制御信号によって捕らえられ、スケジューラ 62 とそのステアリングロジック 87 にフィードバックされ、これらの GCRA (1 / MCR , MCR) 非適合フローに対するその後受信されるリファレンスを優先順位が低いキュー 89 上に速度に感応してキューに入れるためにスケジューラ 62 とそのステアリングロジック 87 をセットアップする。即ち、86 において GCRA (1 / PCR , PCR) 適合であることが分かったフローについてのその後受信されるリファレンスは、MCR 保証を有する ABR フローについての伝送リストの最後に付けられる。何故なら、それはラウンドロビンの順番で提供されるからである。これに対し、86 においてフローが GCRA (1 / PCR , PCR) 非適合であることが分かった場合には、そのフローについてのその後受信されるリファレンスは、そのフローについての PCR に従い、非作業維持型ストール仮想クロックカレンダーキュー 89 上にスケジューラ 62 によって組み込まれ、これにより PCR リミットをフローに対して課す。

【 0078 】

より一般的には、図 4 の実施の形態は下記の ABR トラフィック成形アルゴリズムを実施することは明らかであろう：

もし GCRA (1 / MCR , MCR) 適合であれば、MCR * 加速係数の速度で優先順位の高いキューに入れ、

そうでなく、もし GCRA (1 / PCR , PCR) 適合であれば、優先順位の低い作業維持型のキューに入れ、

そうでなければ、PCR 出力間隔に基づき、優先順位の低い非作業維持型のキューに入れる。

【 0079 】

しかし、類似または異なる原理に基づく代替手段を使用できることは明らかである。

【 0080 】

例えば、図 5 に示すように、一つの魅力のある代替手段は、85 と 86 (図 4) において GCRA (1 / MCR , MCR) 適合でかつ GCRA (1 / PCR , PCR) 適合である

10

20

30

40

50

ことが見出されたフローのデータ転送単位、即ちセルに対するリファレンスを、(1) MCR スケジューリング間隔で優先順位の高いカレンダーキュー 88 と、(2) 優先順位の低いキューの作業維持領域との両方に入れることである。これは「競合」状態を作り出す。何故なら、そのような二重にキューに入れられたリファレンスの一つのインスタンスがキューの何れかにおいてサービスのために取り出される度に、デリンカー 91 は他のキューからそのリファレンスの他のインスタンスを除去するからである。これは、トレランスに関してのより厳しい限界が維持されるであろうことを意味する。この変形例でのトラフィック成形アルゴリズムは：

キューへ入れるとき、

速度 MCR で優先順位の高いキューに入れ、

もし GCRA (1 / PCR, PCR) 適合であれば、優先順位の低い作業維持型のキューに入れ、

そうでなければ、PCR 出力間隔に基づき、優先順位の低い非作業維持型のキューに入れ、

キューから出すとき、

もし優先順位の高いキューから出すのであれば、優先順位の低いキューからリンクを外し、そうでなければ、優先順位の低いキューからリンクを外す。

【0081】

図 6 は、ABR コネクションまたはフローを成形するための更に他の実施の形態を示す。この実施の形態では、MCR が零でないことの保証を有する MCR 非適合 ABR フローの各々に対して提供される出力帯域幅が、出力リンク（即ち、ラインインターフェースの転送側）の ABR セル間の平均間隔と ABR 「レディーキュー」の現在の長さに基づいて動的に調整される。この ABR 「レディーキュー」の現在長 Len は、93 において ABR 伝送リスト 65 e および 65 f 上の作業維持キューの長さを測定することで判断される。一方、カレンダーキュー上の ABR セル間の平均間隔 S は、任意の非 ABR フローに対して許可が与えられた後で ABR フローが利用できる帯域幅を分割することによってリンクの割合として決定できる。これらの非 ABR フローによって要求される帯域幅は、94 において、非 ABR フローのためのエントリーを伝送リスト 65 に提供するのに必要な全体的な帯域幅を計算することによって決定される。

【0082】

例えば、もし 94 において、出力リンクの帯域幅の 50% が ABR フローのために利用できると判断された場合、当然その時の出力リンク上の ABR セルまたはデータ転送単位の間隔 S が 2 セルタイムとなる。この情報と ABR 「レディーキュー」の計算長 Len とにより、スケジューラ 62 に、非作業維持型カレンダーキュー 89 上にこれらのフローについての到着するリファレンスを下記の各出力間隔 T_t でスケジュールさせることにより MCR 非適合 ABR フローのサービスを最適化できる。

【0083】

$$T_t = \text{Min} (S_t , \text{Max} (P_t , R)) \quad (3)$$

但し、R = 予測ラウンドロビン ABR

サービスタイム (R = 現時刻 + S * Len)、

P_t = フロー t の次のセルのための最も早いサービスコントラクトコンプライアントデバチャー (P_t = 現時刻 + 1 / PCR)、

S_t = フロー t の次のセルのための最も遅いサービスコントラクトコンプライアントデバチャー (S_t = 現時刻 + 1 / MCR)

である。

【0084】

E. 分割 ABR コネクションのトラフィック成形

知られているように、ABR コネクションは、図 7 に 101 として示したようなソース・ツー・デスティネーション制御ループまたは図 8 に 102 として示したような分割制御ループを用いて動作させることができる。分割制御ループを採用する場合、個々に制御され

10

20

30

40

50

るセグメントの各々は次のセグメントに対して仮想ソースVSとして機能する。したがって、各ABR制御セグメント(最初のもは除く)に対しては、ABRソースエンドポイントとして作用する仮想ソースから供給が行われる。制御ループが仮想ソースによって受けられる場合には、制御ループから逆向きリソースマネジメント(RM)セルは除去される。しかし、各ソース、即ち仮想ソースは、デスティネーションまたは仮想デスティネーションVDへの利用可能なネットワークの帯域幅を、このフィードバックに基づいて計算でき、セグメント毎に、許可セルレート(ACR)として表すことができる。したがって、分割制御ループを有するこれらのABRコネクションについては、許可される最大フロー毎データ転送単位またはセルレート、 $\text{Max}(MCR, \text{Min}(ACR, PCR))$ を使用してABRトラフィック成形をセグメント毎に行うのが有利である。図4~図6の実施の形態では、 $\text{Max}(MCR, \text{Min}(ACR, PCR))$ に等しいPCR'を設定し、そして計算においてPCRの代わりにPCR'を使用することで上記のようなトラフィック成形を実施できる。

10

【0085】

G. UBRサービスのためのPCRへのトラフィック成形

図4の実施の形態は、トラフィックをUBRサービスのための指定されたPCRに成形するために単純化できる。UBRサービスは明示的または黙示的に指定されたPCRを予測しているが、MCRは予測していない。したがって、ABRのMCR保証を実施するために作られた条項(provision)は削除でき、これはUBRフローまたはコネクションのデータ転送単位またはセルに対する到着リファレンスが下記のキュー入力ルールに従ってキュー89に入れられるであろうことを意味する。

20

【0086】

もしGCRA($1/PCR, PCR$)適合であれば、作業維持領域の最後のバケット内に入れ、

そうでなければ、PCR出力間隔に基づき、非作業維持領域に入れる。

【図面の簡単な説明】

【図1】本発明を使用することが有利なATMスイッチの単純化したブロック図である。

【図2】図1に示すスイッチを横切る間にATMセルがとることが好ましい種々の形を追跡した図である。

【図3】図1に示すチップの出力側、即ち送信側における代表的なチャンネルのより詳細なブロック図である。

30

【図4】本発明のストールド仮想カレンダーキュー実施の概略図である。

【図5】本発明の別のストールド仮想クロックカレンダーキュー実施の概略図である。

【図6】本発明のさらに別のストールド仮想クロックカレンダーキュー実施の概略図である。

【図7】ソース・ツー・デスティネーション制御ループを有するABRコネクションの概略図である。

【図8】分割制御ループを有するABRコネクションの概略図である。

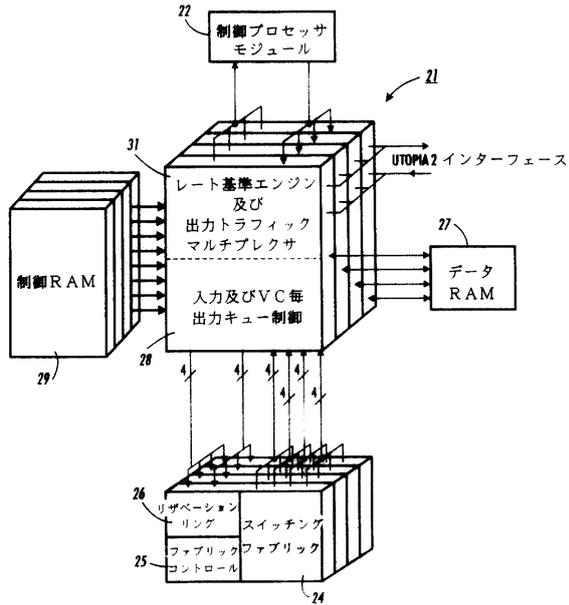
【符号の説明】

62 スケジューラ

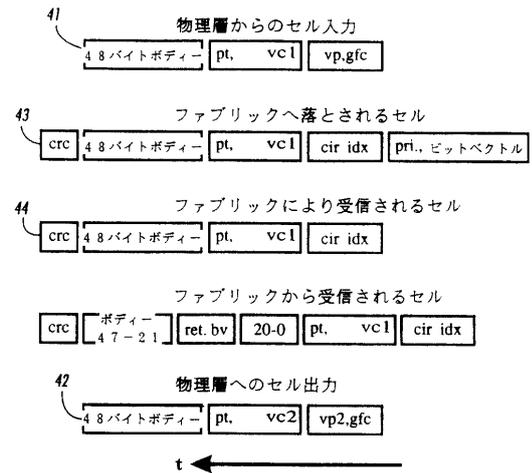
40

87 ステアリングロジック

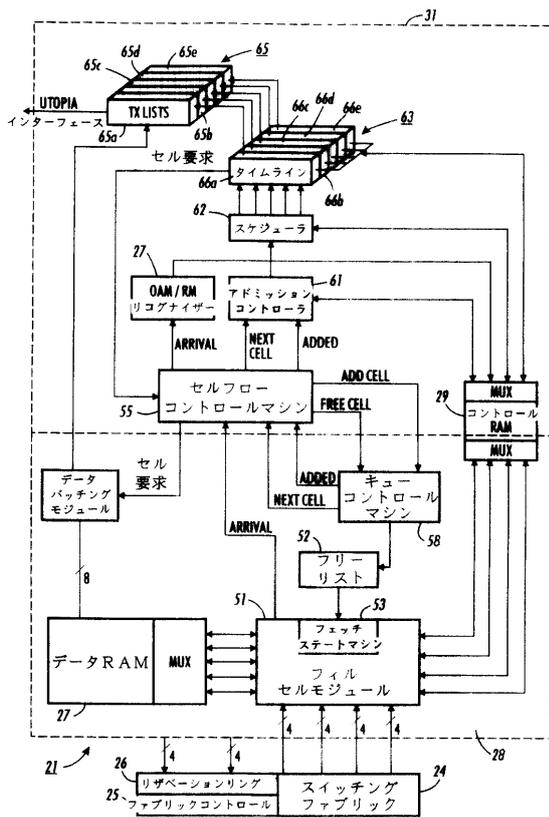
【図1】



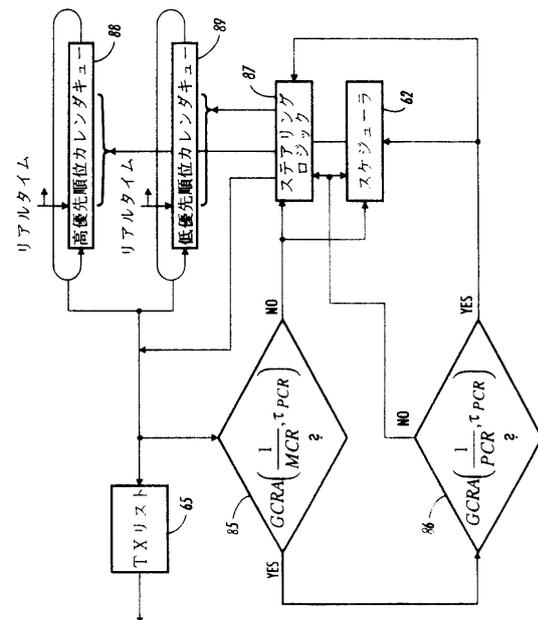
【図2】



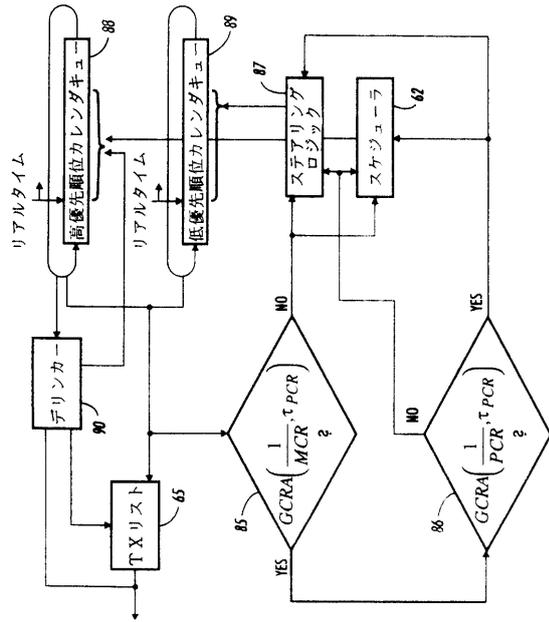
【図3】



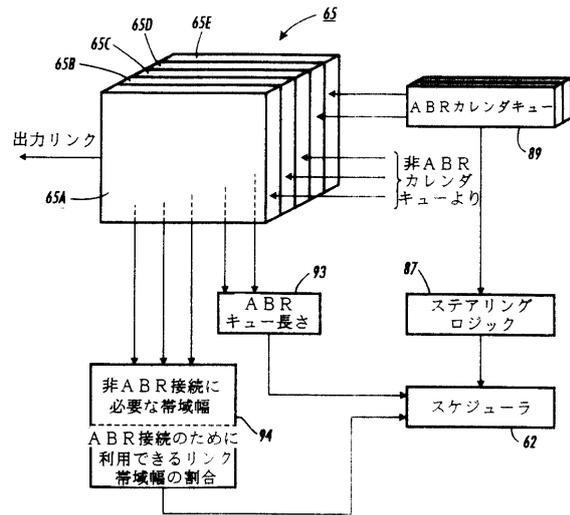
【図4】



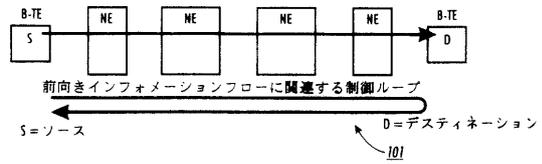
【 図 5 】



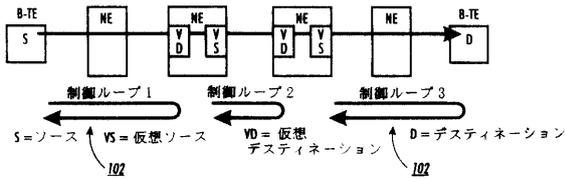
【 図 6 】



【 図 7 】



【 図 8 】



フロントページの続き

- (72)発明者 クリストファー ジェイ・カブラー
アメリカ合衆国 02129 マサチューセッツ州チャールズタウン サーティーンズ ストリート 106 アpartment 230
- (72)発明者 ランディス シー・ロジャーズ
アメリカ合衆国 03848 ニューハンプシャー州 キングストン ロックリモン ロード 61

審査官 石井 研一

- (56)参考文献 特開平09-238159(JP,A)
特開平09-186692(JP,A)
特開平07-170267(JP,A)
特開平07-170266(JP,A)
特開平07-095211(JP,A)
特開平07-058778(JP,A)

- (58)調査した分野(Int.Cl., DB名)
H04L 12/28
H04Q 3/00