

## 階層型ビデオマルチキャストの受信クライアント構成および制御方式

酒澤 茂之                      渡辺 泰利                      滝嶋 康弘                      和田 正裕

(株) KDD 研究所

〒356-8502 埼玉県上福岡市大原 2-1-15

KDD 研究所画像通信グループ

TEL: 0492-78-7426, FAX: 0492-78-7510

e-mail: sakazawa@kddlabs.co.jp

あらまし 階層型ビデオマルチキャストでは、受信者ごとに異なるネットワーク帯域や CPU 能力に合わせたビデオストリームを効率よく提供することができる。本稿では、階層別に異なるマルチキャストグループで配送されるパケットのフォーマットについて提案しており、受信クライアントにおいて、そのパケットを効率よく、安定的に受信し、復号・表示するためのソフトウェアの構成法とその動作を示す。また、クライアントの CPU 能力に適した階層数だけ受信するために、受信中に CPU 能力不足か否かを検出する判定方式として、I フレームの復号時間のばらつきを利用することを提案し、それに基づいた制御方式の効果を実験により示している。

キーワード 階層型ビデオマルチキャスト、IP マルチキャスト、階層符号化、ヘテロ環境

## Configuration and Control of Receiving Client for Layered Video Multicast

Shigeyuki Sakazawa   Yasutoshi Watanabe   Yasuhiro Takishima   Masahiro Wada

KDD R&D Laboratories Inc.

2-1-15 Ohara Kamifukuoka-shi, Saitama 356-8502

KDD R&D Labs. Visual Communications lab.

Phone: +81 492 78 7426, FAX: +81 492 78 7510

e-mail: sakazawa@kddlabs.co.jp

**Abstract** Layered video multicast can provide receiving clients with appropriate video streams according to their network bandwidth and available CPU resources. First, we will propose a packet format to transport layered data, each of which belongs to separate multicast group. Then, a client configuration and its operation is fully described. As for control of the client, it is also proposed to observe a variance of decoding time for each I-frame, and is shown that the control method based on the observation gives good performance.

**key words** Layered video multicast, IP multicast, Layered coding, Heterogeneous environment

## 1. はじめに

IP ネットワークを用いた放送型のビデオコンテンツ配信の手段として IP マルチキャストが注目されている。IP マルチキャストでは、クラス D の IP アドレスを用いてマルチキャストグループを構成し、そのマルチキャストグループに参加している複数の受信クライアントがパケットを同時受信する。このとき、ビデオサーバは各受信クライアントに個別にパケットを送出する必要はなく、クラス D アドレス宛にパケットを送り出せばよい。そのため、送出負荷が軽くなる。また、ネットワーク上を流れるパケット数も個別にパケットを送り出す場合と比べて大幅に削減される。

しかしながら IP マルチキャストを単純に用いた場合、全受信クライアントが同じビットレートのビデオデータを受信せねばならず、クライアントごとに異なる利用可能なネットワーク帯域に対応することができない。また、ビデオ復号には大きな CPU リソースを必要とするが、クライアントによって異なる CPU 能力にも対応できない。そこで、ビデオデータを階層分割し、各階層ごとに IP マルチキャストを行い、受信クライアントが自身の受信するビデオ階層を選択する受信者主導による階層型ビデオマルチキャスト技術 (RLM) が提案されている [1][2][3]。特に McCanne らはこの分野で先導的な研究を行い、フレーム間引きによる時間方向の階層化とウェーブレット変換を用いた空間方向の階層

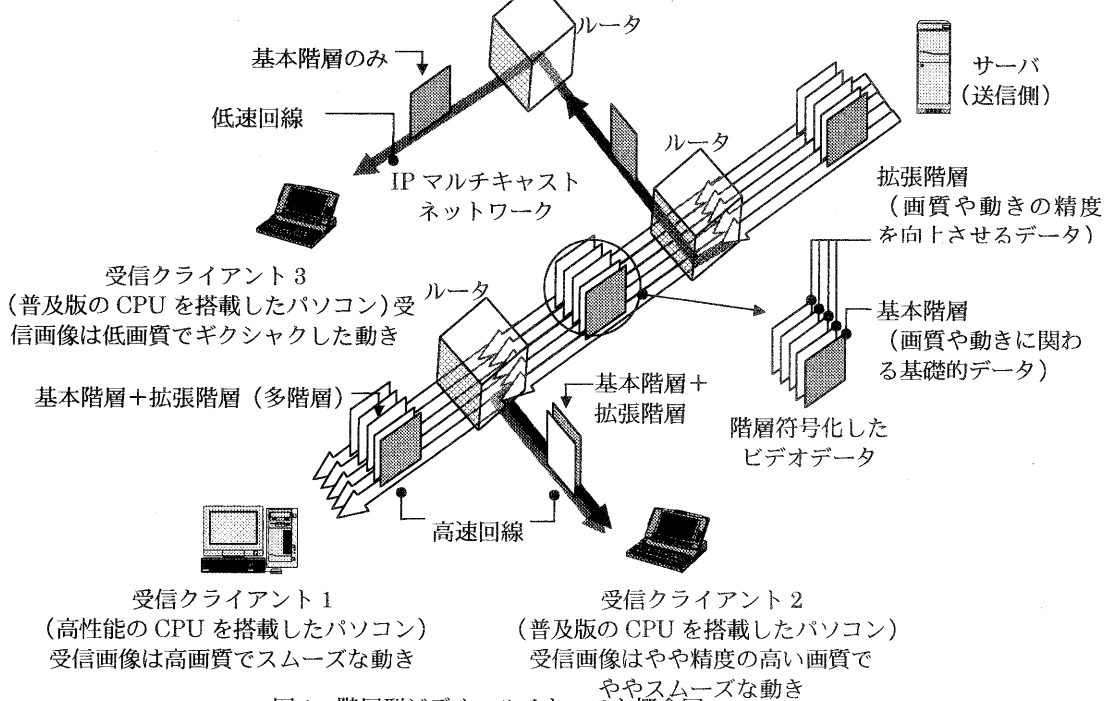
化の組み合わせを提案し、制御自由度の高い RLM を実現している [1]。また Shacham らは輻輳回避の観点から階層制御アルゴリズム [3]、川田らはマルチキャスト受信端末が協調して輻輳回避にあたる方式 [4] をそれぞれ提案しており、優れた特性が示されている。筆者らは、これまでに MPEG-1 を対象として、軽量なアルゴリズムによる階層分割を提案し、最も遅延を少なくするためのネットワークへの送出方法、ならびに階層分割画像の画質評価といったサーバ側の検討結果を報告している [5]。

本稿では、受信クライアントに注目して、通常のビデオストリームとは異なる性質を持つ階層型ストリームのパケット化方式を提案し、それを安定的に受信するためのクライアントソフト構成方法を示す。また、従来の RLM ではあまり検討が行われていなかった CPU リソース不足時にも対応可能な受信階層制御方式を提案する。

## 2. 階層型ビデオマルチキャスト

階層型ビデオマルチキャストは、図 1 に示すように階層符号化されたビデオデータを階層ごとにマルチキャストし、受信クライアントが自ら必要な階層を選択して受信するものである。ここでは文献 [5] に基づいて、その概要を説明するとともに伝送パケットフォーマットを提案する。

送信サーバでは、MPEG-1 で圧縮されたデータを I, P, B ピクチャ (N=15, M=3 が前提) タイプごとに



階層分割する。ついで、各ピクチャタイプ内で DCT 係数を低周波成分(L)と高周波成分(H)とに分割することで、IL, IH, PL, PH, BL, BH という 6 種類の階層データを作成する。さらにオーディオデータ A と合わせて、合計 7 種類のデータをマルチキャストする。受信クライアントは、受信したい階層が所属している IP アドレスに対して IGMP メッセージを発行してマルチキャストグループにジョインする。これにより、その階層データがクライアントのサブネットまで配送されてくる。逆に不要となった階層については、その IP アドレスのマルチキャストグループから離脱することによりパケットの配送が停止される。

このとき、オーディオと各ビデオ階層は独立にストリーム送信されることから、エレメンタリストリームでの取り扱いとなり、システムストリームの機能である PTS, DTS などのタイムスタンプを利用することができなくなる。そこで、時間情報は RTP パケットのシーケンス番号やタイムコード情報により伝送する。各 RTP パケットは UDP/IP で伝送されるが、本稿ではトラフィックのシェーピングのためにネットワークの MTU サイズを考慮して 1500 バイトのパケットとする。したがって各階層の画像データは、1 フレームの情報が複数の RTP パケットに分割されて伝送されることになる。さらに考慮すべきことは、MPEG のビデオ復号においては、1 フレーム中のあるパケットがロスした場合、そのフレーム全体が復号できなくなる<sup>1</sup>。また、フレーム間予測で参照対象の画像フレームが損失した場合も、後続のフレームデータが復号できなくなる。そこで、パケットロスが発生した場合には、新たに到着した RTP パケットが復号可能か否かを判定して、復号不能なパケットを受信側で廃棄する必要がある。これらを考慮すると、RTP パケットに対して、通常のヘッダに加えてビデオフレームに関係したカウンタやフィールドを追加することで、受信クライアントにおける適切な処理を助けると考えられる。そこで、RTP ヘッダおよびユーザ独自のパケットヘッダ部を図 2 のように構成することを提案する。

図 2 において、先頭の 12 バイトは RTP ヘッダそのものであるため説明は省略する。同期化情報識別子以降の 4 バイトについて説明する。

- SEQNO2:SEQNO1 が RTP パケットごとにインクリメントするカウンタであるのに対して、SEQNO2 はビデオフレームごとにインクリメントする。オーディオの場合には、各 GOP の先頭と一番時刻が近いオーディオアクセスユニット(AAU)を先頭とするパケットにおいてインクリメントする。
- STCODE : フレームの先頭、中間、最終のどの

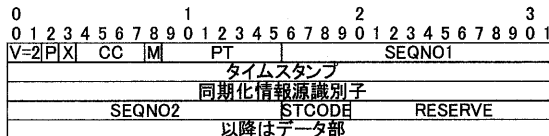


図 2 RTP パケットフォーマット

パケットかの識別情報。

- RESERVE : 将来のための予約。RTP の規定の 4 バイトアライメントに従うためのスタッキングの意味もある。

これらのフィールドに基づくクライアントソフトの動作については、次節で述べる。

### 3. クライアント構成

ストリーミング受信を行うクライアントは一般に、受信バッファにおいてパケット到着の揺らぎを吸収し、後段のビデオ復号が円滑に実施されるように構成される。ところが、提案している階層型ビデオマルチキャストでは、ビデオデータが階層ごとに非同期に到着するため、それらの間の同期関係について考慮しながらデータを取り扱う必要がある。そこで、受信揺らぎ吸収バッファに加えて、デコーダの直前に階層別の復号バッファを設ける。これにより、例えばピクチャタイプ間の依存関係のために、正常に受信したものの復号できないデータを廃棄することが容易となる。また、オーディオ出力・ビデオ表示にあたって、AV 同期をとるために表示用バッファを復号スレッドのあとに設ける。これらの検討に基づいた、クライアントの機能モジュールおよびバッファ構成の提案方式を図 3 に示す。

受信バッファは、階層ごとに異なる IP アドレスで届くため、階層別に準備する必要がある。受信スレッドは、受信バッファがオーバフローしないよう、パケットが到着ししだい復号バッファへ転送する。復号スレッドは、復号バッファからフレームデータを読み出して復号し、復号結果を表示用バッファへ書き出す。表示スレッドは、表示用バッファ内のビデオ・オーディオデータを読み出して、同期を取りつつビデオ表示・オーディオ出力を行う。次に、各スレッドの動作を詳しく述べる。

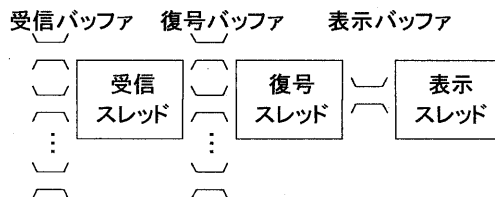


図 3 クライアント構成

<sup>1</sup> 本稿では複雑なエラーリカバリーは行わない

### 3.1 受信スレッド

受信スレッドは、受信を希望する階層のマルチキャストグループアドレスを含んだ IGMP report メッセージをルータへ送信する。その結果、クライアントへパケットが配送されてくるので、受信スレッドは、各受信バッファに対してポーリングを行い、パケットが到着していれば直ちに復号バッファへ移す。ただし、RTP ヘッダの SEQNO1 を監視しており、番号の不連続によりパケットロスが検出された場合には、同じビデオフレームに所属する後続のパケットをすべて廃棄する。このためにユーザヘッダの SEQNO2 と STCODE を利用して廃棄対象のパケットか否かを判定する。

パケットロスは、ネットワーク上で輻輳等のために起きるだけでなく、クライアントの CPU 能力が足りない場合にも起こりうる。これは、他のスレッドにおいて CPU 時間を消費してしまい、受信バッファへのポーリングが間に合わないために受信バッファがオーバフローすることが原因である。あるいは、復号スレッドの処理が間に合っていないために復号バッファがフルの状態が続き、受信スレッドが復号バッファにデータを移せないことも原因となりえる。

### 3.2 復号スレッド

復号スレッドは、MPEG-1 における復号順序に従って復号バッファからデータを取りこんで復号する。ただし、復号バッファ中のデータはエレメンタリストリームフレームごとのデータであるため、時刻情報をもっていない。そこで、RTP ヘッダに付与されていた当該フレームの番号 SEQNO2 をフレームデータのヘッダ情報として保持することで、各階層のフレームデータ間の時間関係を把握する。

復号バッファに必要なデータが存在していない場合には、そのフレームの復号はスキップされる。また、MPEG-1 ではフレーム間に復号上の依存関係があるため、スキップされたフレームに依存するフレームもまた、復号がスキップされる。オーディオに関しては、オーディオアクセスユニット同士に依存関係は無いので、存在するデータを用いて順次復号する。

### 3.3 表示スレッド

主観品質を保つ上で、オーディオの出力が途切れないようにすることが重要である。その一方、PC におけるオーディオ出力はオーディオデバイスが独自のクロックにしたがって行く。そのため、表示スレッドではオーディオデータをデバイスのタイミングにあわせて供給し続けなくてはならない。ビデオ表示については、オーディオと同期させるためにオーディオデバイスのクロックに従属して表示されることになる。このことは、サーバ側における符号化のクロックと、クライアントにおけるクロックが食い違うことにつながり、長時間受信すると必

然的にオーバフローまたはアンダーフローを引き起こす。

表示スレッドの動作は、基本的には、ビデオフレームごとにあらかじめ計算したビデオ表示予定時刻と現在時刻を比較して、遅れている場合にはそのビデオフレームの表示をスキップすることになる。原ビデオ信号が NTSC の場合のビデオ表示予定時刻  $T$  の計算方法を以下に示す。

tc: I ビクチャの先頭のパケットが持つタイムコード

n: GOP の先頭からの表示順でのフレーム番号

T0: ビデオ表示を開始した時刻

tc0: T0 時刻に表示されたビデオフレームの先頭パケットが持っていたタイムコード

$$T = T0 + tc - tc0 + n * 1/29.97$$

オーバフロー時には、この表示予定時刻を早めて（より未来の時間にして）ビデオフレームの表示をスキップしていくことになる。逆にアンダーフロー時には、この表示時刻を遅らせて、描画されている画像をフリーズさせて時刻を合わせる。

## 4. 受信階層の動的制御

クライアントにおける画質低下を引き起こす要因として、CPU 能力の不足とネットワーク輻輳が考えられる。前者については、ビデオデータの受信・復号・表示が間に合わないという形で現れ、後者については、パケットロスが生じるという形で観測される。これらの場合には、受信階層を削減し、適切な階層のみを受信するようにする必要がある。そこで、本節では特に CPU 能力不足の検出方法について検討し、それに基づく自動階層制御アルゴリズムを提案する。ネットワーク輻輳に関しては McCanne らの方式に基づく簡易的なアルゴリズムを用いる。

### 4.1 CPU 能力の不足時の受信階層制御

ビデオ復号は CPU の処理時間を大量に消費する。特に、I フレームの復号では、動ベクトル処理がない代わりに、すべての画素ブロックにおいて可変長符号の復号および逆 DCT 演算を行わなければならないなど、データ処理量が多い。そこで、I フレームの復号に要する時間を観測することで、CPU 能力を推測することが可能になると考えられる。図 4 に 2 種類の PC を用いて全階層を受信したときの I フレーム復号時間を示す。ここで、それぞれ

PC1: 233MHz MMX Pentium

PC2: 400MHz Pentium II

なる CPU を搭載している。

図 4 から、PC2 ではほぼ 10msec 程度で復号が行えているのに対して、PC1 では 30msec 程度時間を要していることがわかる。これは、CPU のクロックが PC2 では 2 倍近いことから単純に復号時間が 1/2 になっているものと思われる。さらに、図 4 か

らは PC1 の場合に復号時間が非常にばらついていることが観測される。これは、CPU 能力が足りない場合には、復号処理を行っている途中で他のスレッドに CPU 時間を占有されてしまい、見かけ上復号時間が増えていると考えられる。PC2 のように十分な CPU 能力がある場合には、復号スレッドは復号に専念できるため復号時間にばらつきが出ていない。

そこで、ばらつき方と受信階層数の関係を見るために、PC1 にて IL+IH (=I), IL+IH+PL+PH (=I+P), IL+IH+PL+PH+BL+BH (=I+P+B) で復号時において、復号時間の 5 秒ごとに移動分散値を計算した結果を図 5 に示す。図 5 から受信階層数を削減するにつれて、分散値が小さくなっていることがわかる。なお、PC1 では復号画質の主観的評価から I+P が適正な受信階層数であることから、分散値が 100 以上の場合には、CPU 能力不足であると判定し、階層数をそれ以下に限定することとする。すなわち、4.2 節で述べる階層追加制御において、分散値 100 以上となる階層は追加しない。

#### 4.2 ネットワーク帯域が不足時の受信階層制御

クライアントにおける観測値はパケットロスであるので、ここでは 5 秒ごとにパケットロス率を求め、ロス率が 5% を超える場合に階層を 1 ずつ減じることとした。逆に、別途定めるタイマ値で示される時間ロス率 0 が続いた場合には階層を 1 つ追加する。このタイマ値は初期値 5 秒で開始し、階層追加後にパケットロスが発生した場合には 2 倍の値に変更する一方、タイマ値の間ロスが発生しなければ初期値に戻される。これにより、ネットワークを不必要に flood させないようにしている。

#### 4.3 実験

サーバからの階層ストリームをネットワークエミュレータ経由で PC1 と PC2 で受信した。サーバからの送出レートは全階層合わせると約 1.3Mbit/s である。ネットワークエミュレータは、簡易なトラフィックポリシング機能を持っており、通過レートがサーバの送出レート以下の場合には、エミュレータ内のバッファがあふれてパケットがロスする。今回のシミュレーションでは、バッファサイズを 1MByte、通過レートを 768kbit/s, 384kbit/s とした。

まず、CPU 適応型制御の効果を見るために、PC1 にて受信実験を行った。ネットワーク輻輳制御は ON の状態で、CPU 制御が ON の場合と OFF の場合を比較した結果を図 6 に示す。CPU 能力制御が働いている場合は、I+P 以上に階層が追加されないのに比べて、制御が無い場合には何度も階層を追加しては削減するという制御を繰り返している。主観的には、このような場合、画像が動いてはフリーズする現象が繰り返し起こることになり、強い悪印象を与える。それに対して、制御ありの場合の画質は安定しているため、主観的な印象は大幅に改善され

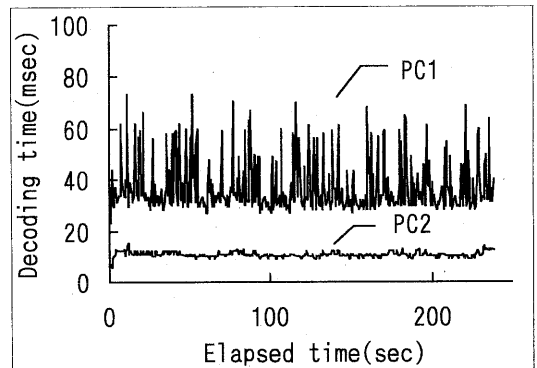


図 4 I フレーム復号時間

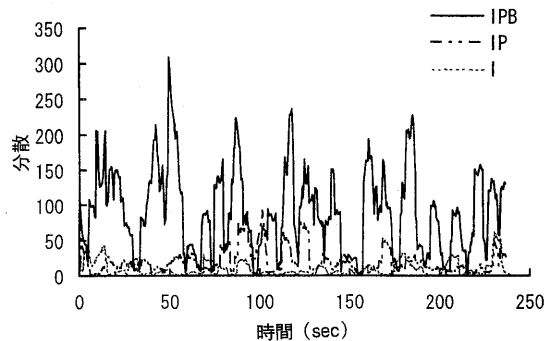


図 5 I フレーム復号時間のばらつき

る。次に、ネットワーク輻輳制御については、PC2 にて受信実験を行った結果を図 7 に示す。図 7 から、768kbit/s の場合ほぼ I+P で、384kbit/s の場合 I 階層で安定して受信していることがわかる。特に 384kbit/s では、I+PL への階層追加の試みの間隔が 2 倍ずつ増えていくことが顕著に観測できる。

#### 5. むすび

多様なネットワーク帯域、CPU リソースに適応して受信者ごとに最適なストリームを提供できる階層型ビデオマルチキャストに関して、受信クライアントの具体的な構成方法および動作について紹介した。また、自動的に最適なストリームを選択するための動的制御アルゴリズムを検討し、その効果を確認した。

#### 謝辞

日頃ご指導いただく KDD 研究所秋葉所長、浅見副所長、松本取締役に感謝いたします。

なお、本研究は通信・放送機構の研究開発課題「IP ネットワーク上の Webcasting における QoS 制御」により行われました。

#### 参考文献

[1] S. McCanne, et al., "Receiver-driven Layered

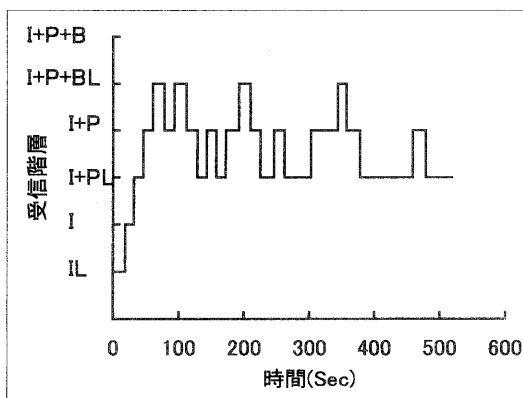
Multicast", Proc. ACM SIGCOMM '96, pp.117-130, August 1996.

[2]Y. Ito, et al., "Rate-Adaptive Real-Time Multicast TV Conference System with Locally Adaptive Packet Flow Control", IEICE Trans. Inf.&Syst., vol.E82-D, no.4, pp.815-821, 1999.

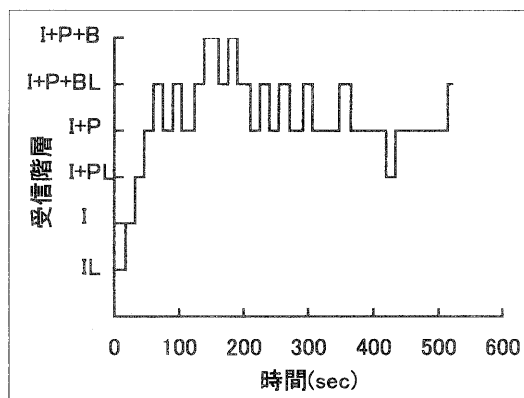
[3]N. Shacham and H. Yokota, "Admission Control Algorithms for Multicast Sessions with Multiple Streams", IEEE J. Sel. Areas Commun., vol.15, no.3, pp.557-566, Aril 1997.

[4]M.Kawada, et al., "Multiple Stream Controller for Layered Multicast", Proceeding of APCC/OECC'99, Beijing, Oct.20, 1999.

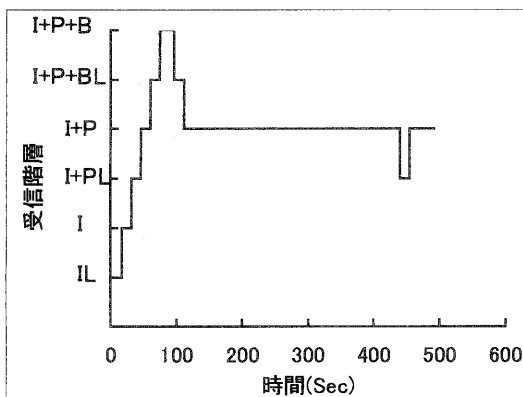
[5]酒澤, 滝嶋, 和田, "ビデオマルチキャストにおける符号化伝送制御方式", 映情メ学会誌, vol.52, no.6, pp.863-870, June 1998.



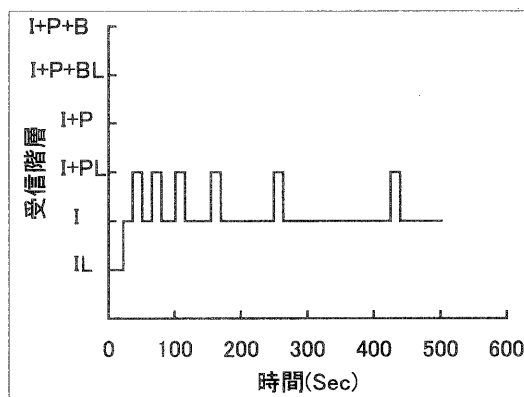
(a) 制御なし



(a) 空き帯域 768kbit/s



(b) 制御あり



(b) 空き帯域 384kbit/s

図6 CPU能力制御の効果

図7 ネットワーク輻輳時の制御結果