

## 網の輻輳によるセル廃棄が IP over ATM のスループット特性に対して与える影響の評価

石橋 勇人、岡部 寿男、金澤 正憲

京都大学 大型計算機センター

### 概要

ここ数年、ATM の高速性に着目して ATM を用いた LAN が多数の大学等で構築されてきた。現在の LAN では、IP がもっとも主要なプロトコルであるため、ATM LAN では IP over ATM ないし LAN エミュレーションによって IP を利用することになる。

本稿では、より ATM の高速性を活かすことができ、スケーラビリティにも優れた IP over ATM (RFC 1577) 方式による計算機間の通信に関して、UDP および TCP についてスループット特性を実測し、評価を行った。

特に、ATM 特有の現象である輻輳時のセル廃棄が通信のスループットに与える影響を調べるため、CBR トラフィックによって輻輳条件を作り出し、そこでのスループット特性に着目して評価を行っている。

## A Throughput Analysis of IP over ATM Traffic on a Congested ATM Link

Hayato Ishibashi, Yasuo Okabe and Masanori Kanazawa

Data Processing Center, Kyoto University

### Abstract

ATM is one of key technologies for very high speed computer networks. Many campus networks adopt ATM LANs as their backbone networks in these days. IP is a dominant protocol in such networks.

In this paper, we have tried to reveal throughput characteristics of IP over ATM traffic, especially on a congested ATM link. An ATM link is a lossy link under such conditions and UDP/TCP over ATM traffic is largely reduced its performance in some cases. We have used CBR continuous traffic as interference traffic and made up a cell loss condition. We also described how traffic shaping works well and not so well by its parameter.

## 1 はじめに

近年、高速ネットワークとしての ATM (Asynchronous Transfer Mode) が注目され、多数の大学等において ATM LAN が導入されている。現在の LAN における支配的なプロトコルは IP (Internet Protocol) であり、これは ATM LAN であるからといって変わるものではない。したがって、IP を ATM LAN において運用することが重要となる。

ATM では、セルと呼ばれる比較的短い固定長パケットにデータを格納して伝送するため、IP パケットを ATM セルを用いて伝送するためには、1つのパケットを多数のセルに分割して送ることになる。

また、ATM では、セルの到達を必ずしも保証していないため、輻輳時にセル廃棄が起こる場合があり、これらのことが IP 通信の振る舞いを従来のネットワークと異なったものとしている。

そこで、本稿では、ATM LAN における IP 通信のスループット特性を実験によって実測し、解析する。特に、ATM に特有の性質であるセル廃棄が生ずるような環境下において、IP 通信がどのような特性を示すかについて測定している。

## 2 IP over ATM の特質

### 2.1 IP と ATM

ATM LAN 上で IP による通信を行おうとする場合、LAN エミュレーションと呼ばれる方式と、IP over ATM<sup>1</sup> (以下では IPoA と表記する) と呼ばれる方式の2つが主流である。

LAN エミュレーションは、その名が示すとおり、従来型のネットワークである Ethernet やトークンリングを ATM でエミュレートする方式であり、IP に限らず任意の上位プロトコルを使用することが可能となる。しかし、方式上サーバがボトルネックとなるためにスケラビリティに欠け、また、それほどスループットが期待できない。

一方、IPoA 方式は基本的に IP に特化した方式であり、LAN エミュレーションと比較して高いスループットが期待できる。本稿では、この IPoA 方式についてのみ取り扱う。

<sup>1</sup>ここでいう IP over ATM とは、RFC 1577[1] ならびに関連 RFC に準拠した方式を指すものとする。

### 2.2 IP over ATM

IP パケットが ATM ネットワーク上を流れる場合、IPoA では、まず LLC (Logical Link Control)/SNAP (SubNetwork Attachment Point) フォーマットによってカプセル化し、AAL5 (ATM Adaptation Layer 5) を用いてセルに分割している [2](図 1)。

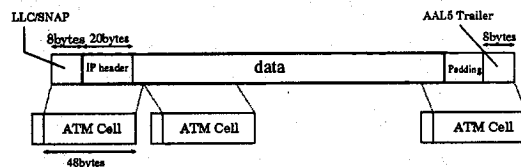


図 1: IPoA のカプセル化

したがって、IP 1 パケットに対するオーバーヘッドは、

$$\begin{aligned} &8 \text{バイト}(\text{LLC/SNAP}) + 8 \text{バイト}(\text{AAL5}) \\ &= 16 \text{バイト} \end{aligned}$$

である (正確には、これにパディングが加わる)<sup>2</sup>。

IP の最大のパケット長 (Maximum Transmission Unit; MTU) はデータリンクによって異なるが、Ethernet では 1500 バイト、ATM の場合には 9180 バイトが標準となっている。

これに対して ATM のデータ転送単位であるセルでは、ペイロードサイズが 48 バイトと定められているため、IP の 1 パケットは最大 192 個のセルに分割されてネットワーク上を流れることになる (MTU サイズがデフォルトの場合)。

このために、1つのセルが廃棄されることによって 1パケット全体の損失へとつながることになり、セル廃棄の影響が強く現れる。

### 3 スループット特性の測定

IPoA トラフィックのスループット特性を評価するために、2台の Unix ワークステーション (以下 WS) とトラフィックジェネレータを ATM スイッチ

<sup>2</sup>本稿では、1バイト=1オクテット=8ビットとする。

を介して接続し、いくつかの条件の下でスループットを測定した(図2)。SVCのセットアップタイムの影響を除くため、各機器間はPVCで接続している。

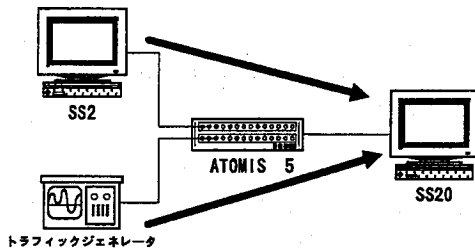


図2: 実験環境

実験に使用したWSはSun SPARCstation20(SS20)、SPARCstation 2(SS2)の2台であり、それぞれのスペックは表1に示すとおりである。

また、ATMスイッチには2.5Gbpsの交換能力を持つ入出力バッファ型ATMスイッチであるNEC ATOMIS5を使用した。各機器とATMスイッチ間の接続はOC-3c(155.52Mbps)である。

実験にあたっては、受信側のWSの能力がネックとなってスループットが低く測定されることがないように、より処理能力の高いSS20を受信側、SS2を送信側としている。

スループットの測定は、

1. WS間の通信以外にトラフィックがない場合
2. WS間の通信以外にトラフィックジェネレータによる妨害トラフィックがあり、それによって輻輳が生じる場合

の2つの条件の下で、それぞれ送信メッセージサイズを1KB~64KBまで1KBずつ増加させながら行った。MTUサイズは標準の9180バイトである。

スループットの値は、“Netperf”[3]を使用して10秒間トラフィックを発生させ、5回計測したスループットを平均して求めたものである。

## 4 測定結果とその考察

### 4.1 UDPトラフィックのスループット特性

UDPパケットの送信を行った場合に得られたスループット特性を図3に示す。

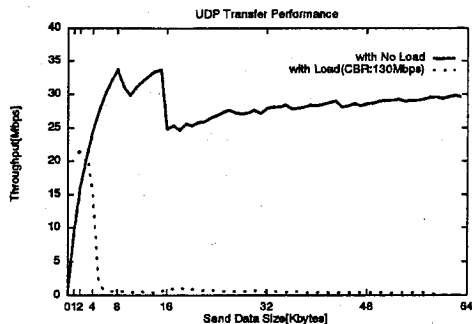


図3: UDPのスループット

図において、実線で示されているのが、無負荷時、すなわちWS間のnetperfによるトラフィックのみが存在する場合のスループットである。この最大値は、送信側のWSであるSS2の最大の送信能力を示していると考えられ、その値は33.6Mbpsであった。

送信メッセージサイズがMTUサイズ以下である場合には、送信メッセージサイズが大きくなるにしたがってパケット長に対してデータの占める割合が増え、ヘッダによるオーバーヘッドが減少する。また、パケットの組み立て/分解に要するオーバーヘッドも減少する(OSにとっては、パケット処理のための割り込みの回数が減る)。このため転送効率が増え、スループットは単調に増大する。

図3では、その後送信メッセージサイズが10KBでスループットが減少し、さらに16KBで大きな落ち込みが見られるが、これらの時点ではUDPパケットの破棄は見られず、また条件を変化させても再現するため、実験に使用したシステムのものであると考えられる。

次に、図3において、破線で示されているのが高負荷時、すなわちトラフィックジェネレータによって発生させたCBRによる130Mbpsの妨害トラフィックが存在する場合である。

この場合にUDPとして利用可能な帯域幅は、

$$(155.52 - 130) \times \frac{260}{270} \times \frac{48}{53} \times \frac{9180}{9216} \times \frac{9180 - 20 - 8}{9180} = 22.10 \text{ Mbps}$$

であるから、WS間のトラフィックがこれ以下の場

表 1: 実験に使用したシステム

	SS20	SS2
ハードウェア	Sun SPARCstation 20 (SuperSPARC 60MHz)	Sun SPARCstation 2 (SPARC 25MHz)
OS	Solaris 2.5	Solaris 2.5
主記憶	64MB	48MB
NIC	Adaptec ANA-5240 (512KB)	Adaptec ANA-5240 (512KB)
NIC ドライバ	Ver. 2.04b	Ver. 2.04b

合には輻輳は生じないが、これを越えた時点で輻輳によるセル廃棄が生じ始めることになる(図4)。無負荷時の結果から SS2 は最大約 34Mbps のデータ送信能力を持つことがわかっているため、トラフィックジェネレータからのトラフィック 130Mbps との合計は出力側のラインの持つ帯域幅を明らかに超えている。したがって、この場合には輻輳が生じ、セル廃棄が起こることになる。

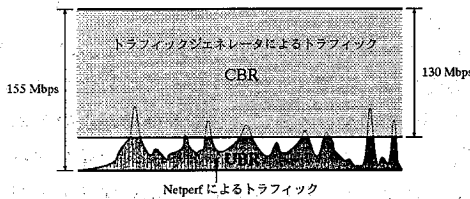


図 4: 輻輳によるセル廃棄

図3において輻輳がないときにスループットが 22.10Mbps を越えるのは、データサイズが 4KB を越えたあたりであるが、輻輳時の実測結果を見ると、ちょうどこの付近からスループットが急激に低下し、ほぼ 0(0.5Mbps 程度)となっていることがわかる。

したがって、スループットが落ち込むのは、利用可能な帯域幅を使い尽くしてしまったことによって輻輳によるセル廃棄が発生したためであると考えられる。UDP はフロー制御機構を持たないために、送信側は輻輳の如何に関わらず同様にパケットを送出し続ける。このため、一度輻輳が生じて ATM スイッチのバッファが溢れ始めると、その後は継続的にバッファが溢れ続けることになり、受信側がパ

ケットを組み立てられるだけのセルが到達しないためにスループットがほとんど 0 となるものと考えられる。

#### 4.2 UDP トラフィックに対するトラフィックシェーピングの効果

ATM では、トラフィック制御のためのトラフィックシェーピングと呼ばれる機能がある。これは、セルの送出間隔を指定された一定の間隔にそろえることによって、“行儀の良い”トラフィックを生成するものである(図5)。

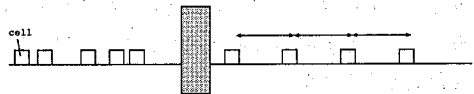


図 5: トラフィックシェーピング

トラフィックシェーピングによってセル間隔が広がる(= 送出レートが下がる)ため、輻輳が回避でき、セル廃棄が避けられると期待される。

図6に実線で示したのが送信側でシェーピングを行い、送出レートを一定にしたときのスループット特性である。ここでは、シェーピングのレートを 16% としている。16% という値は、CBR によるトラフィックを除いた残り、すなわち UBR トラフィックが利用可能な帯域幅である。 $((155.52 - 130) / 155.52 = 0.16)$

図からわかるように、MTU サイズ以上のデータを送信する場合には、一定のスループットが得られている。すなわち、送信されるデータが常に利用可

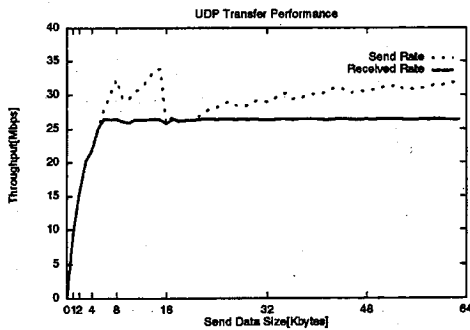


図 6: UDP のスループット (16% シェーピング)

能な帯域に収まっているため、輻輳が生じず、セル廃棄が回避できていると考えられる。

次に、シェーピングレートを若干上げて、18% (28 Mbps 相当) としてみたのが図 7 である。

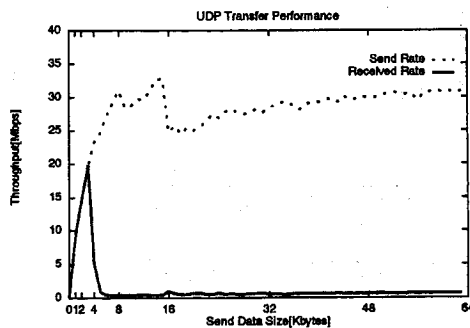


図 7: UDP のスループット (18% シェーピング)

図のように、今度はシェーピングの効果が全く見られず、シェーピングを行わない場合と同様の結果となっている。これは、少しでも入力レートが出力レートよりも大きければ、結局はバッファ溢れによるセル廃棄が生じるためである。先に述べたように、一度バッファ溢れが生じてしまうと、フィードバックを持たない UDP では回復は望めない。

本節の実験結果から、

1. UDP において、適切なトラフィックシェーピングは極めて効果的である。
2. しかしながら、シェーピングのレートが利用可

能な帯域幅を少しでも越えていると意味がない。

ことがわかる。

実際のネットワークにおける輻輳回避の観点から考えると、1の性質により一見シェーピングは非常に有効であるように見える。しかしながら、利用可能な帯域が常に変動するようなネットワークでは、2の性質が致命的である。そのようなネットワークでは利用可能な帯域を予測することが困難なので、それに合わせてきちんとシェーピングを行うことはほとんど不可能だからである。したがって、有効な輻輳回避のためには、何らかの別の手段を組み合わせることが必要であろう。

### 4.3 TCP トラフィックのスループット特性

UDP と同様にして、TCP における無負荷時のスループット特性をプロットしたのが図 8 である。

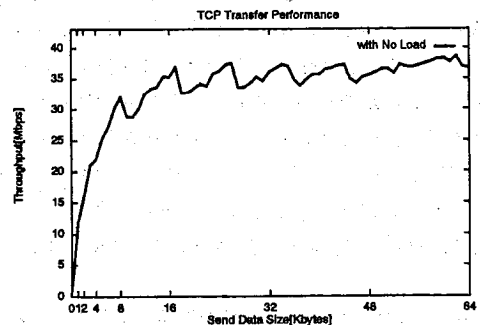


図 8: TCP のスループット (輻輳なし)

TCP の場合にもやはりデータサイズが大きくなるにつれてパケットの組み立て / 分解に必要なオーバーヘッドが減少するため、当初スループットは単純に増加する。その後はパケット化する際の端数の出方により、MTU サイズを周期として振動する形となる。

次に、CBR による 130Mbps の妨害トラフィックを加えたときのスループット特性が図 9 である。

TCP の場合の有効帯域を UDP の場合と同様に計算すると、

$$(155.52 - 130) \times \frac{260}{270} \times \frac{48}{53} \times \frac{9180}{9216} \times \frac{9180 - 20 - 20}{9180}$$

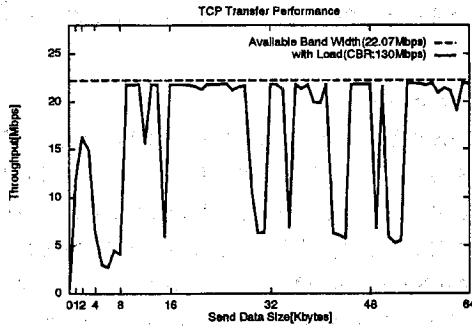


図 9: TCP のスループット (輻轉時)

$$= 22.07Mbps$$

となる。無負荷の状態においてこのスループットに達するのは、やはり送信データサイズが 4KB を越えた時点である。したがって、このあたりから輻轉によるスループットの低下が始まる。

しかし、UDP とは異なって TCP はフロー制御機構を有するので、パケットが失われると送信側は送出レートを減少させる。このために輻轉状態から回復することができるのだが、それでもやはりデータサイズによってはスループットが極端に低下する現象が見られる。これは、TCP のフロー制御アルゴリズムによる帯域制御とセルのバースト的送出によるセル廃棄が相互作用した結果であると考えられる。

#### 4.4 TCP トラフィックに対するトラフィックシェーピングの効果

TCP トラフィックに対して 16% のシェーピングをかけると、図 10 のようになる。

シェーピングの効果によって、見事にスループット特性が改善されていることがわかる。

さらに、シェーピングのレートを 20%, 24%, 32% と順次上げてみたときのスループット特性をそれぞれ図 11, 12, 13 に示す。

TCP の場合には、帯域の 2 倍 (32%) まで送出レートを上げて概ね均一なスループット特性が得られるが、やはり部分的にスループットが落ち込む減少が見られる。この傾向は送出レートを上げるほど顕著となり、シェーピングなしの状態へと近づいて行く。したがって、スループットの落ち込みは輻轉に

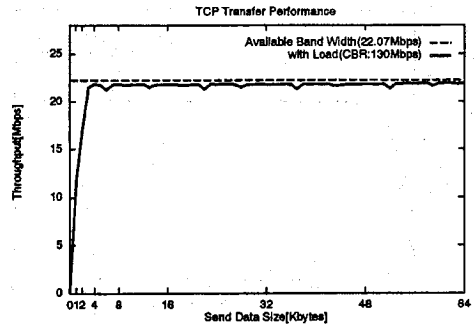


図 10: TCP のスループット (16% シェーピング)

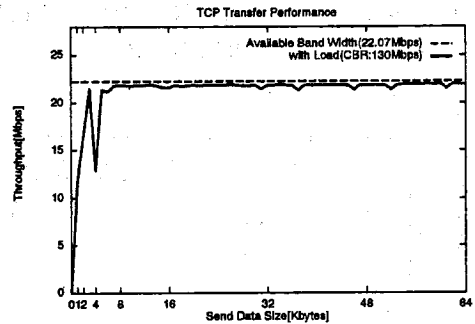


図 11: TCP のスループット (20% シェーピング)

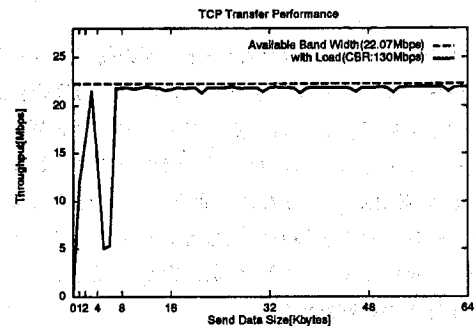


図 12: TCP のスループット (24% シェーピング)

よるセル廃棄が引き起こすパケット損失が原因であると考えられる。

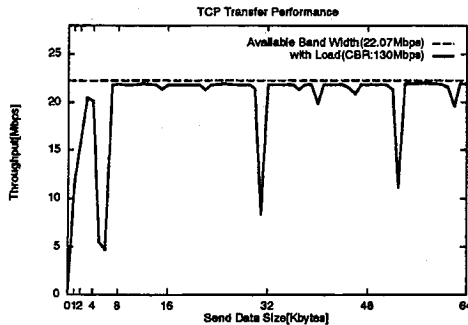


図 13: TCP のスループット (32% シェーピング)

これらの実験からわかるように、フロー制御機構を有する TCP であっても、やはりトラフィックシェーピングはトラフィック特性の改善のために有効に働く。シェーピングのレートを設定するために、利用可能な帯域幅を正確に知ることが困難であるという点は UDP の場合と変わりはないが、TCP の場合にはシェーピングのレートに対してそれほど敏感ではないので、実用的なレベルでシェーピングを設定することが可能ではないかと考えられる。そこで、その時点で使用可能と推定される帯域に合わせてシェーピングをかける機構を用意することによって、TCP のフロー制御を補なってスループット特性を改善できるのではないかと考えている。

## 5 おわりに

本稿では、輻輳が生じた場合の IP over ATM のトラフィック特性を実測し、シェーピングによる輻輳の回避がトラフィック特性の改善に有効であることを確認した。また、UDP ではシェーピングのレートに極めて敏感であること、TCP はシェーピングのレートに対してそれほど敏感ではないが、やはりパラメータによってはスループットの大きな低下を招く場合があることを示した。この改善のためには、シェーピング機構とフロー制御機構を組み合わせることが有効ではないかと考えられる。

現在、ソースが公開された FreeBSD による実験環

境を整えており、IPoA 環境において TCP のフロー制御とシェーピング機構を相乗させることによってトラフィック特性を安定させるためのアルゴリズムの考案と実際のドライバの改善を試みているところである。今後は改良を加えたコードによる評価を行うとともに、各種のパケット棄却アルゴリズムとの相互作用などについても調べてゆきたい。

**謝辞** 本稿の執筆にあたり、実験に多大なる協力をいただいた京都大学工学研究科応用システム専攻 (現 日本電気) の高崎喜紀氏にこの場を借りて感謝いたします。

## 参考文献

- [1] M. Laubach: Classical IP and ARP over ATM, RFC 1577(1994).
- [2] J. Heinanen: Multiprotocol Encapsulation over ATM Adaptation Layer 5, RFC 1483(1993).
- [3] Information Networks Division, Hewlett-Packard Co.: Netperf: A Network Performance Benchmark Revision 2.1, <http://www.cup.hp.com/netperf/NetperfPage.html> (1996).
- [4] 高崎喜紀: IP over ATM におけるトラフィック特性に関する研究, 京都大学修士論文 (1997).