

経路帯域測定に基づくストリーム配送品質推定法の評価 A study of the path bandwidth measurement for the stable Audio-Visual stream distribution

小松 義幸[†]
Yoshiyuki Komatsu

島村 和典[†]
Kazunori Shimamura

1. はじめに

現在、インターネットは大きく発展し、個人のインターネット環境のブロードバンド化が進み、ストリーミングによる動画配信やIP電話などの新しいアプリケーションが登場してきた。大容量のデータがリアルタイムにやり取りされるようになり、ネットワークへの信頼性や安定性への要求が高まっている。

しかし、現在のインターネットはベストエフォート型の通信で成り立っており、通信の品質保証はされていない。ネットワークの経路上でエンドホストの通信品質を管理する仕組みが無いため、必要な帯域が確保できず、通信品質を長時間一定に保つことが難しい。特にストリームなどリアルタイム性の必要な通信ではその影響が大きく、途中で途切れたりするといった障害が発生する。

この問題を回避する方法としては、予めユーザが自分の通信帯域に応じてストリームのビットレートを選ぶ方法がある。しかし、この方法では転送されるデータ量は一定であり、急にネットワークが混雑して帯域が狭くなった場合には対処できない。

したがって、この問題を解決するためには、ストリームの配送品質をネットワークのトラフィック状況に応じて調整する方法が望ましい。また、バス上のボトルネックリンクの帯域を通信に利用するホスト数に応じて適切に分配して利用させ、輻輳が発生するのを防ぐことも重要である。

本論では、ネットワーク上のエンドホストに対して、通信に利用するネットワーク経路の利用可能帯域をトラフィック状況に応じて公平に分配し、リアルタイムかつ低負荷に通知できるシステムを提案し、その性能評価を実験により確認した。

2. 利用可能帯域の定義

本論では利用可能帯域を以下のように定義する。利用可能帯域は複数のリンクからなるパスに対して定義するものとする。送信ホスト(Tx)、受信ホスト(Rx)間のパス上に n 個のルータが存在する場合を考える。ルータ間のリンクをそれぞれ L_1, L_2, \dots, L_n とし、それぞれのリンクの利用可能帯域を A_1, A_2, \dots, A_n とする。このときのTx, Rx間のパスの利用可能帯域 A は、パス上のリンクの中で、最も未使用の帯域が少ないリンクの未使用帯域であり、以下ようになる。

$$A = \min \{A_i\} (i = 1, 2, \dots, n)$$

3. 提案方式

提案する測定法ではまず、ネットワーク上の自律システム (Autonomous System, 以下 AS) ごとにモニタリ

ングサーバ (Monitoring Server, 以下 MS) を設置し、AS内にあるルータからAS内のパスのトラフィック情報をSNMPを使って常時モニタリングする。そして、その情報からそれぞれのパスの利用可能帯域を保持する。このサーバはエンドホストから送られてきたプローブパケット (Probe Packet, 以下 PP) に対し、必要に応じて自分の管轄するAS内のパスの利用可能帯域を書き込んで次のASまたは、相手ホストへと中継する機能を持つ。いくつかの中継を経て、相手ホストに到達したPPを相手ホストが送信元のホストに送り返すことで利用可能帯域を通知する。

このシステムの全体構成を、図1に示す。また、AS内のルータから帯域情報を収集する機能をモニタリング、エンドホストがPPを送受信する機能をプローブと呼ぶ。以降それぞれの機能について解説する。

3.1 モニタリング

AS内に設置されたMSのモニタリング法を図2に示す。MSは、予め登録されたAS内のルータからSNMPを使ってインターフェース情報とルーティング情報を収集し、そこからAS内の全経路パターンを把握する。そして、一定時間ごとにそれぞれのルータから各インターフェースのトラフィック量をSNMPで収集し、経路ごとにその経路上で最も利用可能帯域の狭いリンクの帯域をその経路の利用可能帯域として保持する。

モニタリングの動作手順は次のようになっている。

1. 予め与えられたAS内のルータリストを読み込む。このリストにはIPアドレスやコミュニティ名などの情報が含まれている。
2. ルータリストを基にそれぞれのルータからSNMPを使ってインターフェースの数と速度、ルーティングテーブルを収集する。
3. 収集した情報を基にルータ間のリンクの一覧を作成する。この一覧にはリンク速度や利用可能帯域などの情報も書き込む。

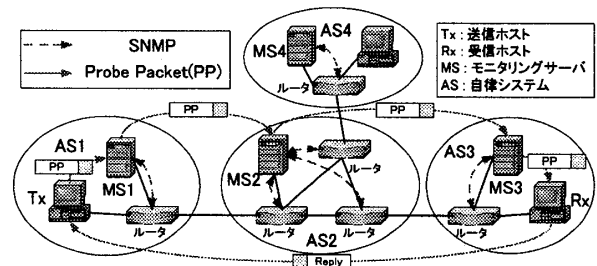


図1: システム全体構成

[†]高知工科大学, KUT

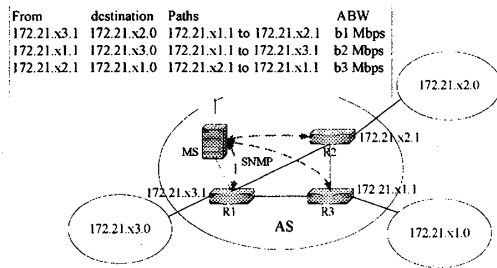


図 2: モニタリング法

- AS 内にある境界ルータ全てを 2 つずつ組み合わせたペアを両端とするパスの一覧を作成する (図?? 上部)。この一覧には AS への入り口となる境界ルータから出口となる境界ルータまでの全てのルータ IP アドレスと、その間にある全てのリンク情報が含まれる。また、ルーティングテーブルからそれぞれのパスを使う宛先 IP アドレスを割り出して列挙する。
- パスごとに、リンクの一覧からそのパス上のリンクの利用可能帯域を調べ、最も小さい帯域をそのパスの利用可能帯域とする。
- 定期的に各リンクのトラフィック量を SNMP で収集し、パスの帯域情報を更新する。

モニタリング手順 4 で列挙されるルータは、出口側ルータのルーティングテーブルから外部 AS ヘルティンクされている IP アドレスを抜き出し、入り口となるルータから順に各ルータのルーティングテーブルで、どのルータヘルティンクされているかを辿って調べる。

なお、モニタリングでのトラフィック量収集時には、1 台の MS につき、UDP による SNMP の要求と応答の 2 つのパケットがルータのインターフェースの数分発生する。ただし、このとき発生するパケットは送信した MS の属する AS 内でのみやり取りされる上、全てのパケットが同じ経路を通るわけではない。よって、ネットワークに対する負荷は小さいといえる。

3.2 プロープ

プロープは、エンドホストが他のホストに対して通信を開始して終了するまでの間、定期的に MS に対して PP を送信することで行う。このプロープの方法を図 3 に示す。

具体的な動作手順は次のようになっている。

- Tx が送信開始時に Tx と Rx の IP アドレスと Tx の利用可能帯域 (b0) を PP に書き込み、そのホストが所属する AS の MS に送信する。

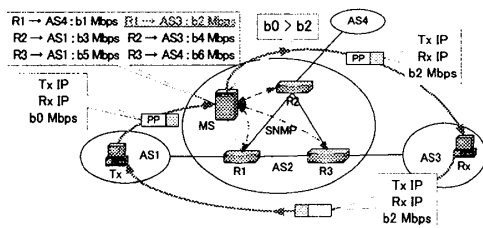


図 3: プロープの方法

- PP を受信した MS は、前の MS で書き込まれた自 AS の始点ルータの IP アドレスと Rx の IP アドレスを、モニタリングで作成したパスの一覧と照らし合わせて使用するパスを特定する。
- 特定したパスの利用可能帯域 (b2) と b0 を比較し、b0 より小さければ b2 を PP に書き込む。
- Rx の IP アドレスを見て、PP が次に通る AS の入り口となるルータの IP アドレスを PP に書き込む。
- PP を次の AS の MS へと転送する。MS には予め近隣の AS にある MS の情報が登録されている。
- 2 から 5 の処理を PP が Rx の属する AS に到達するまで繰り返す。
- Rx の所属する AS の MS は PP 2 の処理を行った後 Rx に転送する。
- PP を受け取った Rx はこれを Tx に返信する。
- 返信されたパケットから Tx - Rx 間の利用可能帯域を知る。

3.2.1 プロープパケット

プロープパケットは通常の IP パケットとして送信され、トランスポート層は UDP を使用する。データ部分に格納されるデータは次の通りである。

- Tx の IP アドレス — 32bit
- Rx の IP アドレス — 32bit
- AS の入り口となるルータの IP アドレス — 32bit
- 通信開始、継続、終了のフラグ — 2bit
- MS で書き込まれるパスの利用可能帯域 — 32bit

3.3 特徴

この方式では、モニタリングによる SNMP パケットを除けば、1 回のプロープ時に送信されるパケットが 2 つであるため、ネットワークへの負荷が小さく、1 回のプロープにかかる時間もパケットの往復時間と MS での処理時間だけなので極めて短い。そのため、連続的に任意の間隔で測定し続けることが可能である。また、システム全体がアプリケーション層で動作するため、既存のプロトコルへの変更も必要ない。

4. 性能評価実験

この実験では、提案方式の性能を把握し、他の方式との比較を行うため、実験環境上で提案方式の実装と他の方式の計測ツールをそれぞれ動作させ、計測時間、通知帯域、クライアント負荷、ネットワーク負荷、実帯域への追従性について調査し、その評価を行った。この提案方式の実装プログラムは MS 上で動作させるサーバプログラムと、エンドホスト上で動作させるクライアントプログラムの 2 つから構成される。プログラム言語には Perl を用い、SNMP を使用する部分においては既存の Perl モジュールを組み込んで使用している。なお、現段階では利用可能帯域を測定し、エンドホストに通知する機能のみ実装している。

4.1 実験環境

この実験で用いた実験環境を図4に示す。それぞれのホスト、ルータ間は全て100MbpsのEthernetで接続されており、それぞれのルータではNet-SNMP 5.1が動作している。

4.2 比較対象

本実験では、現在利用可能な帯域測定ツールであるPathload-1.1とIperf-1.7を提案方式との比較対象として同じ環境で動作させてその性能を計測した。

4.3 実験

提案した方式と比較対象のツールについて以下の項目に関して調査を行い、それぞれの結果をまとめた。

4.3.1 計測時間

それぞれのツールがTx, Rx間の帯域を1回計測するのに必要な所要時間を測定した。提案方式の測定については、送信ホスト上でPPをキャプチャし、その到着時間の差を測定している。なお、Iperfについては測定時間を指定できるため、デフォルト(10秒)のままとした。

実験結果は、表1の通りとなった。この結果はそれぞれのツールごとに、10回の計測の平均計測時間とその標準偏差を出している。提案方式は1回の計測がTx-Rx間のパケットのRTTとMSでの処理時間だけになっており、実験結果からも他のツールに比べて10の4乗オーダーの差で短時間での計測が可能であることがわかる。なお、この実験環境でのTx, Rx間のパケットのRTTは平均0.7msec程度になっており、計測時間平均値の3.1msecからこのRTTを引いた2.4msecがMSでの処理時間といえる。今回の実験で用いた環境は途中経由するMSが1台のみであるが、複数のネットワークを経由し、数台のMSを経由する場合を考えると、処理時間はこの整数倍になると考えられる。しかし、1台辺り2.4msec程度の増加なので0.1秒間隔でプローブを行うとしても十分無視できるレベルであるといえる。また、計測時間の標準偏差も $7.34e^{-05}$ と小さく、安定して計測できている。

表1: 計測時間測定結果

	提案方式	Pathload	Iperf
平均	0.00312	8.14	10.0
標準偏差	$7.34e^{-05}$	1.96	0

4.3.2 通知帯域

この実験ではTx, Rx間のパス上に他のトラフィックが発生していない状態と、Iperfを用いてルータ1, ルータ3の間でUDPによる30Mbpsと50Mbpsのトラフィックを発生させ続けた場合の、それぞれのツールの通知帯域を調べている。なお、提案方式のモニタリングとプローブはそれぞれ1秒間隔で行っている。

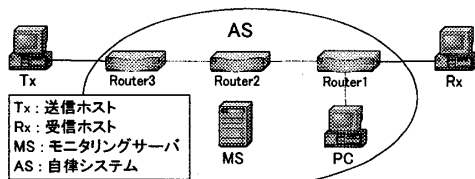


図4: 実験環境

実験結果はそれぞれ、表2、表3、表4の通りとなった。これはそれぞれのパターンで5回ずつ計測した結果の平均と標準偏差をまとめたものである。

表2: 通知帯域(負荷無し)

	提案方式	Pathload(下限-上限)	Iperf
平均	100	95.836	96.172
標準偏差	0	0.34	0.08

表3: 通知帯域(負荷30Mbps)

	提案方式	Pathload(下限-上限)	Iperf
平均	67.962	37.922	63.410
標準偏差	0.23	0.72	16.21

表4: 通知帯域(負荷50Mbps)

	提案方式	Pathload(下限-上限)	Iperf
平均	47.762	38.566	43.322
標準偏差	0.26	0.68	3.54

まず、表2のネットワーク負荷が無い状態での通知帯域については、提案方式はルータから通過したバイト数を収集するため、負荷が無いのでそのままのリンク速度が通知されている。そして、Pathloadはパケットペアによる到着間隔、Iperfは直接TCPのスループットを計測しているため、それぞれの違いからこのような結果になったと考えられる。

次に、表3の30MbpsのUDPトラフィックを発生させた場合の結果については、提案方式とIperfについては妥当な計測結果が得られたのに対して、Pathloadが約38~63Mbpsという広い帯域値の計測結果しか得ることができなかった。この原因についてははっきりしていないが、おそらくPathload独自の計測方法による問題であると思われる。また、このときの標準偏差はPathloadの上限が16.21と大きな値になっている。

次に、表4の50MbpsのUDPトラフィックを発生させた場合の結果については、それぞれのツールで多少の違いはあるものの、全て利用可能帯域としては妥当な計測結果が得られた。また、このときの標準偏差についてはPathloadの上限値が3.54と他と比べて若干大きい結果となった。

そして、全ての計測結果を見ると、提案方式の通知帯域は100MbpsのEthernetでの実際の利用可能帯域から考えるとやや高い帯域を通知している。これは提案方式が、リンク速度から直接通過バイト数を引いた値を通知しているためであるが、実際の帯域にあわせるためにはこれをある程度減じて通知する必要があると考えられる。しかし、この点を除けば、提案方式はそれぞれの帯域において標準偏差が小さく、安定して正確な値が通知されているといえる。

4.3.3 クライアントの処理負荷

この実験では、それぞれのツールが動作している時のクライアントのCPU使用率を測定している。

実験結果は、表5の通りとなった。この表は、それぞれのツールで計測を10回行った際の平均値をまとめている。この実験結果では、PathloadとIperfはクライアントCPUのおよそ半分またはそれ以上を使用しているのに対して、提案方式はほとんどクライアントのCPUに負荷をかけていないことが分かる。

表 5: クライアント処理負荷測定結果

	提案方式 (Tx,Rx)	Pathload	Iperf	
平均	2.4	2.1	63.3	49.0

この実験結果より、提案方式は他のツールをクライアント上で動作させるよりもはるかに低負荷で、他のアプリケーションの動作にほとんど影響の無いレベルで動作することが確認できた。

4.3.4 ネットワーク負荷

この実験では、それぞれのツールが1回の計測を行う間に Tx - Rx 間で送受信されるパケットをキャプチャし、その数を測定した。

実験結果は、表6の通りとなった。この表は、それぞれのツールで5回計測を行った際のパケット数の平均値を示している。

表 6: ネットワーク負荷測定結果

	提案方式	Pathload	Iperf
平均	2	7609	65191

この結果では、提案方式は1つのパケットを往復させるだけなので計測パケットは全て2つとなっており、ネットワークへの負荷は極めて低く抑えられている。Iperfに関しては直接 TCP でデータを転送を行っているため最も多いパケットが発生するという結果を得た。

4.3.5 通知帯域の実帯域への追従性

この実験では、提案方式が Tx - Rx 間の測定を行っている間に複数の TCP フローによって、利用可能帯域が変動した場合のプロープ及びモニタリングの間隔による通知帯域の変化を測定した。具体的には、まず提案方式で Tx - Rx 間の利用可能帯域を測定し続け、その間 60 秒に渡ってルータ 1, ルータ 3 の間にランダムな TCP フローを発生させた時の通知帯域をプロープ及びモニタリングの間隔をそれぞれともに 0.5 秒, 1 秒, 2 秒と変えて記録した。

実験結果はそれぞれ、図 5, 図 6, 図 7 の通りとなっている。これらのグラフはそれぞれ、60 秒の間に変動した実際の利用可能帯域とそのときの提案方式による通知帯域を表している。

この結果から、どのプロープ間隔でも、そのプロープ間隔による遅延を除けば、細かく上下する TCP による帯域変動に概ね追従して通知できていることがわかる。ただし、プロープ間隔が 1 秒以上になった場合は実際の帯域に対して通知される帯域の遅延が大きくなり、リアルタイムに通知できているとは言い難くなっている。よって、プロープ間隔は 0.5 秒以下程度にしておく必要があると言える。そしてこれに伴って、MS が行うモニタリングの間隔は 0.2 秒程度が妥当であると推測される。また、一部利用可能帯域が非常に小さくなっている時に、通知される帯域が 0Mbps を下回っている場合がある。これはルータで高負荷時に、通過バイトが次の秒に多くカウントされてしまっていると予想される。ただし、この現象は利用可能帯域が 0 に近い状態で現れているため、利用可能帯域は 0 であると判断しても差し支えないと考えられる。

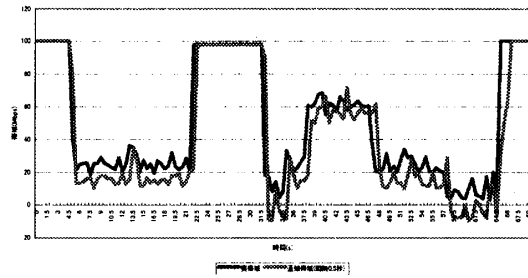


図 5: 通知帯域 (プロープ間隔 0.5 秒)

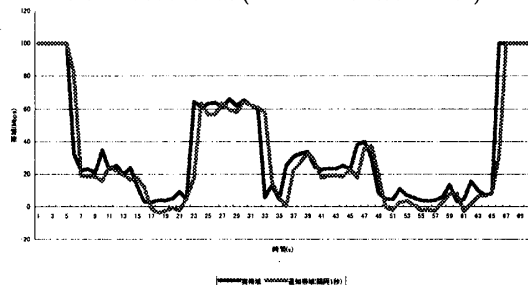


図 6: 通知帯域 (プロープ間隔 1 秒)

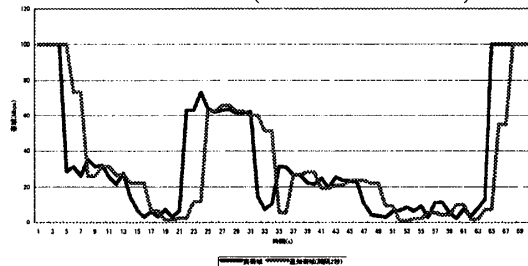


図 7: 通知帯域 (プロープ間隔 2 秒)

4.4 性能評価実験まとめ

以上の実験より、提案方式は計測時間、クライアントの処理負荷、ネットワークへの負荷において他のツールと比較して優れており、計測時間も一定で短時間、低負荷で測定が可能という結果が得られた。また、プロープ間隔を 0.5 秒以下、モニタリング間隔を 0.2 秒程度にすることで細かい帯域変動にも追従できることが確認できた。

5. まとめ

本論では、ネットワーク上のエンドホストに対し、通信時にエンドホスト間のパスの利用可能帯域を低負荷でリアルタイムに通知するシステムを提案した。また、その性能評価実験を行い、他のツールと帯域計測性能の比較を行ってその有効性を確認した。今後は、得られた利用可能帯域を通信するエンドホスト数に応じて、公平に再分配して通知する機能の実装を行っていく。

参考文献

- [1] 中平拓司, 島村和典, “ICMP を応用した End-to-End 利用可能帯域測定法”, FIT2002 情報技術レターズ vol.1, 2002
- [2] 小松義幸, 中平拓司, 島村和典, “モニタリングサーバを用いたエンドホスト間利用可能帯域測定法の提案”, 平成 15 年度 電気関係学会 四国支部連合大会, 2003