

複数経路を用いた IP パケットの高速伝送方式の提案と試作

M2004MM011 家田 直幸

指導教員 長谷川 利治

1 はじめに

現在のインターネットでは、情報の発信元から目的地までの通信経路は同時に1つである。変わりゆくトラフィックにより使用する通信経路の利用可能帯域幅は常に変化しており、これに伴い End-to-End の通信速度も変化している。通信速度が安定しない理由として、経路途中のネットワークで利用可能帯域幅が小さいボトルネックや高負荷なトラフィックによるパケットの損失、遅延、輻輳制御などが起きていることが挙げられる。

これまでに TCP Reno や TCP Vegas[1] といった輻輳制御そのものの動作方式を変える手法で高スループットを得たり、TFRC(TCP-Friendly Rate Control Protocol)によるトラフィック制御[2]によって、TCP、UDP の通信に対して以前より送信レートの変動が小さい安定した通信をすることが可能となった。

ブロードバンドサービスの普及に伴って、高スループットかつ安定した通信を可能とする技術が重要な課題(高画質のストリーミングやリアルタイムストリーミングにおいて)となってきた。これまでの議論では通信経路が1つであることが前提であり、このような手段では通信経路のネットワークキャパシティを超えることができない。加えて高スループットを得ようとすれば、ネットワークに負荷を掛けることになり、他の通信を圧迫してしまう。

そこで本研究では他の通信を極力圧迫することなく高スループットを得るために、同時に複数の経路を用いる通信手法を検討する。複数の通信経路の利用可能帯域幅を少しずつ利用することで、結果的に1つの経路を使用する時よりも高スループットを実現し、同時に通信経路の負荷を分散することで輻輳を避ける安定した通信を実現する。複数の経路を用いる送信の問題として、使用する経路の遅延の格差によって送信側のパケット送り出し順と目的地でのパケット到着順が入れ替わること、使用する各経路の利用可能帯域幅が異なることが挙げられる。そこでこの2点を考慮に入れた送信振り分け手法を考案し、実験で評価した。実験はシミュレーションによる数値演算と実機実験の2通り行った。シミュレーションでは振り分けルータに到着したパケットが、考案した経路振り分けアルゴリズムを通じて目的地に着く状況を模倣させた。送信時とのパケット到着順の入れ替わり頻度、入れ替わり距離の情報をもとにアルゴリズムの有効性を評価し・問題点の改良に役立てた。また、考案したアルゴリズムを実験ネットワーク内で実装し、実験を行った。

2 複数経路の利用モデル

現在のインターネットでも複数の経路を利用した通信モデルは存在する。これらの通信モデルと本研究の提案する複数経路を使用した通信モデルとの違いを解説する。

2.1 現在の複数経路モデル

複数経路を使用可能な環境には

1. BGP4
2. 複数 ISP を別の IP アドレスで利用する
3. Overlay (P2P などの応用で構成される上位のネットワーク)

がある。1はBGP4を利用して接続先ごとに自ASを持ち、入出力のトラフィックコントロールを行う際に代替経路として複数の経路を使用する。2はマルチホームミングともいうが、接続が維持できなくなった際に他ISPに切り替えてインターネットに接続することで通信の安全性を高めたり、負荷分散などに用いられる。いずれにしても、代替経路としては使えるが、複数経路に振り分けて使う一般的な方法はない。3のOverlay Networkにおける複数経路は[3]で述べられているように、ノードをアプリケーションレベルのルータと考えることができるので、同時に複数の経路を使った通信を行うことが可能である。

複数経路利用の研究は上記のようにいくつか見られるが、ほとんどがOverlayやad hocネットワークのもので通常のIP接続においての実現方法はあまり述べられていない。本研究では、前提として複数経路の利用は片方、もしくは両端のホスト間に特殊なルータを置くことでIPレベルでの複数経路への送信を実現する。

2.2 本研究の複数経路モデル

本研究で提案する複数経路利用モデルは振り分けルータで、振り分け先のDstIPアドレスを指定することによって実現することができる[4]。接続形態によって以下の環境で利用することができる。

1. ISPに依存しないポータブルIPアドレス*1を使っての複数の経路にマルチホーム接続
2. ISPから割り当てられた別のIPアドレスで複数の経路にマルチホーム接続

1の場合はルータのルーティングテーブルを使用せずに直接DstIPアドレスを指定することで、送信を振り分けることができる(図1)。2の場合、振り分けルータ側

*1 プロバイダ非依存アドレス「Provider Independent Address(PIアドレス)」とも言う

で NAT を利用することで ISP から割り当てられた別の IP アドレスを同時に使うことが可能となる。それと同時に、受信側でもこれらを同じ SrcIP アドレスから送られてきたものとして受け取らなければならないので、さらに手前でアドレスを変換する (図 2)。

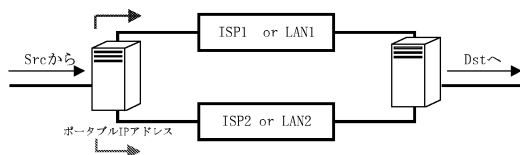


図1 複数経路モデル1

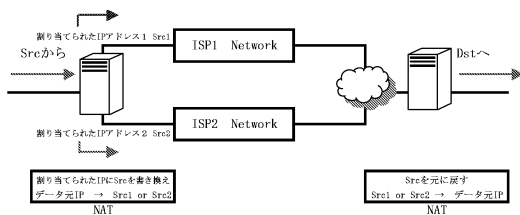


図2 複数経路モデル2

3 複数経路伝送の問題と課題

複数の経路を利用して通信することは容易ではない。経路として利用できるネットワークの状況を適切に判断して、送信を振り分ける必要がある。本研究では経路振り分けルータが次の情報に基づいて送信を複数経路に振り分ける。

- 経路の利用可能帯域幅
- 宛先 IP までの遅延

これらの情報の収集方法については本研究では議論しない。経路の利用可能帯域幅と遅延についての情報収集論は [5] で述べられている。

3.1 複数経路の問題点

複数の経路を使用するためには、単一の経路を使用する場合と異なる処理が必要である。その主な原因は次の2点である。

- 各経路の利用可能帯域幅が異なる
- 各経路の遅延が異なるため到着順が入れ替わる

利用可能帯域幅が異なるので、ある程度の重みを付けた経路の振り分け制御やトラフィック制御を行う必要がある。これはネットワークのキャパシティが一様でなく、同じ割合でトラフィックを流すと、利用可能帯域幅が小さい経路に対して大きな負荷が掛かることを回避する。過度の負荷によりパケット損失率が上がり、特に TCP の場合、輻輳制御が働くとその経路のスループットが低下するだけでなく、送信フロー全体のスループットが低下してしまう可能性がある。また、遅延が異なるために、振り分けルータでの送り出しと同じ順序でパケットが到着する保証がない。到着順の入れ替わりは、

TCP では修正されるが、UDP ではこのような機構は基本的にはない (RTP ではタイムスタンプがあるので修正可能)。

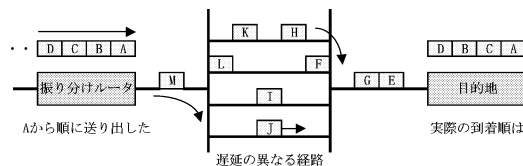


図3 到着順入れ替わり

これまでの研究では TCP についての複数経路はいくつか議論されてきたが UDP を含む IP レベルでは複数経路はほとんど議論されていない。その理由としては、非 TCP は再送機構がないことやこのようなパケットの入れ替わりを修正する機能に対応していないことがある。しかし UDP を用いるストリーミングや RTP を用いる IP 電話、ビデオ会議でも複数経路の利用が有効だと考えられる。

3.2 非 TCP での複数経路実現に対する課題

UDP のような再送機構のないサービスでは

1. データの損失を極力避ける。
2. 到着順序の入れ替わりをできるだけ少なくする。

必要がある。UDP の複数経路通信では本研究で前提とする特殊なルータからの情報を利用し各経路の利用可能帯域幅を見ながら送信経路の振り分けるので、輻輳状態に陥る可能性のある経路に流すデータを他経路に回すことができる。これは輻輳制御機構、再送機構がない UDP にとって送信レートを下げることなく通信することを可能にし、かつ輻輳下でのパケット損失の発生を減らすことによる通信品質の向上といった大きなメリットとなる。また、輻輳時の TCP フローとの公平性について議論されているが、複数経路環境では輻輳状態に陥る可能性のある経路へのパケットの流入を抑制し、UDP フローによる輻輳の悪化を回避する。利用可能帯域幅と遅延が異なる環境下で出来るだけ順序の入れ替わりを抑制することは、複数経路を用いて通信を可能にするための重要な課題である。本研究では入れ替わりをなるべく回避するような送信アルゴリズムの考案に重点を置く。

4 送信振り分けアルゴリズム

複数経路を用いる通信の実現が困難な理由は、各経路の性能が違うことである。均一な経路が並んでいる場合、ラウンドロビンやランダムで経路を決定する送信アルゴリズムで十分な性能を発揮することができる。これは到着順序の入れ替わりの要因となる遅延がほぼ同じであり、通信経路の帯域幅も同じためである。一般的にこのような経路を探すのは困難で、あまり実用性がない。本研究ではこの2点

1. 遅延に重点を置いたアルゴリズム

2. 利用可能帯域幅に重点を置いたアルゴリズム

に着目し、その振り分けアルゴリズムを考案・検討する。本研究で考案した2つのアルゴリズムを次に詳しく述べる。このアルゴリズムの片方を用いる、もしくは2つ併用することで高スループットかつ安定した通信を実現する。

4.1 Delay Based Algorithm

複数の経路を使用してIPパケットの送信をすると、配送経路の遅延によってパケット到着順序の入れ替わりが発生する。それを極力減らすために遅延ベースでの送信制御を行う。

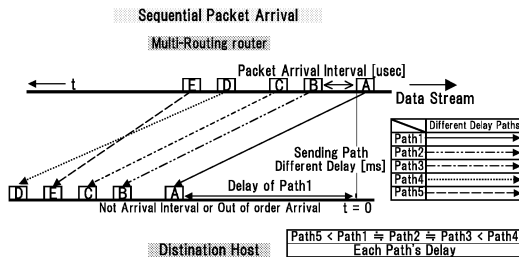


図4 RoundRobin (Before)

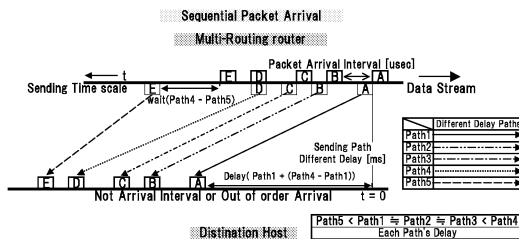


図5 RoundRobin (After)

単純なRoundRobinを経路の切り替えに利用すると、図4のように目的地でパケット到着順の入れ替わりが起こる。到着順が図5になるように入れ替わりが抑制できるアルゴリズムを考えた。連続したパケットを振り分け配送する場合、任意のデータ量、またはパケット数だけ送り出したら配送する経路を変える。遅延はそのタイミング調整に使用する。遅延の差分だけタイミングを遅らせて送り出すことで、End-to-Endでは遅延が同じ経路を使用したようにみせる。このアルゴリズムの問題は、遅延の分散が大きいと到着順の入れ替わり距離が大きくなってしまふことである。振り分けに使用する最良の経路を選択する工夫が必要だが、1経路当たりの負荷を均等に分散させて高速の通信が見込める。その一方、経路を切り替えるデータ量についての重みを考えなければ、各経路を均等に使用するために帯域の細い回線により負荷が掛かる。

4.2 Available Bandwidth Based Algorithm

目的地への経路における利用可能帯域幅は常に変動し、それ以上のスループットは出すことができない。そのため、同時に複数の経路でそれぞれの利用可能帯域幅

に応じた速度でパケットを配送し、全体のスループットを向上させた安定した通信速度を実現する。

図6のようにA, Bの経路が使用可能な状況下で、それぞれの利用可能帯域幅がaMbps, bMbpsとする。また経路を切り替えるタイミングを利用可能帯域幅の $ax[0 < x < 1]$ とする。ルータに到着するIPパケットを連続してある経路に送り出し続けていき、配送パケットサイズ(datasize)と配送を始めてからの時間(sendtime)を利用して、式を満たす限り振り分けルータはその経路にパケットを送り出す。式を満たさなくなったら、別の任意の経路に対して同じことを繰り返す。以下は経路Aに対する条件式

$$\frac{ax}{8} > \frac{\text{datasize}}{\text{sendtime}} \quad (1)$$

この条件を満たさなくなったらsendtimeを初期化して次の経路に送り出し始める。

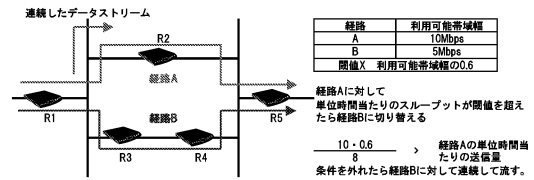


図6 Available Bandwidth Based Algorithm

この場合、利用可能帯域幅に応じた負荷分散を行うことができるので、各経路の能力に応じた「公平」な経路使用が期待できる。また、前述したDelay Based Algorithmと組み合わせて使用することで双方のメリットを生かすことができると推測できる。

5 シミュレーション

シミュレーションでは振り分けルータにアルゴリズムを組み込んださいに、到着したパケットがどのように次のノードに送り出されるのか、またどのような到着順で目的地のホストに到達するのかを模倣する。これによって実際にルータにアルゴリズムを組み込んださいの問題点、課題点などを考察しアルゴリズムを改良できる。ここでシミュレーションの結果から着目すべき点は

- パケット到着順序の入れ替わり頻度
- 順番が入替わったパケットの移動距離
- スループット

である。これらからアルゴリズムの有効性を議論する。

以下に示すシミュレーション結果は、次の想定を基に実験した。使用した入力トラフィックモデルは実インターネットを想定し、到着間隔とパケットサイズは指数分布に従うように変動させた。これに従い振り分けルータに到着する仮想パケットデータは到着間隔の平均が0.5ms、パケットサイズは平均が500バイトの指数乱数より生成した。生成した仮想パケットデータは振り分けアルゴリズムによって経

路が決まり、経路のパラメータに従った目的地への到着を模倣する。入れ替わりを左右する遅延に関しては、 $\text{遅延} = \text{前パケットの遅延} \cdot C + \text{指数乱数} \cdot (1-C)$ とした。前のパケットとの相関係数を C とし、指数乱数の平均を経路の平均遅延にすることで経路ごとの遅延を変動させた。

5.1 Delay Based Algorithm

シミュレーションでは合計 3 経路を同時に使用してパケットがどのような到着順で目的地に着くのかを観測した。設定した経路情報は表 1、経路の切り替えは 1 パケット毎とした。結果として導かれたデータは表 2 である。評価目標と照らし合わせると、順番の入れ替わりが

	遅延 [ms]	帯域幅 [kbps]	
		Delay Based	Available Bandwidth
Path1	20	4800	4800
Path2	30	3600	2400
Path3	40	2400	1200

表 1 Information of each path parameters

	RoundRobin	Delay	Available Bandwidth
	タイミング調整なし	タイミング調整あり	タイミング調整あり
Dist/Packet	13.79734	4.23116	0.04318
Fraction	0.97428	0.91577	0.017
Distance/Point	14.161576	4.62033	2.54
UDP Throughput	6411.88kbps	6411.87kbps	6132.50kbps

表 2 Simulation results

約 92 % 発生し、平均で 1 パケット当たり平均 5 個分到着順が前後する。経路の遅延差によるタイミング制御のあり、なしを比べると、パケットの入れ替わり距離を抑制できていることが分かる。このアルゴリズムの問題点として浮き上がってきたのは、パケット到着順の入れ替わりである。入れ替わりが発生する主な要因としてはパケットの到着間隔とランダム遅延の格差にある。パケットが伝送路に乗るときに掛かるシリアル化遅延はリンク速度に依存しマイクロ秒の範囲であるのに対し、伝送遅延はミリ秒である。送り出し経路を 1 パケットずつ変えることで生じる各経路の伝送遅延の差が、振り分けルータに到着するパケットの到着間隔よりかなり大きいことが推測された。

5.2 Available Bandwidth Based Algorithm

シミュレーションでは合計 3 経路を同時に使用してパケットがどのような到着順で目的地に着くのかを観測した。設定したパラメータ・経路情報は (表 1)。ルータが送り出し経路を切り替える閾値は転送容量が利用可能帯域幅の 0.25 を超えたときとし、0.1 秒ごとに転送量を確認した。結果として導かれたデータは表 2 となった。パケット到着順の入れ替わりは頻度、距離ともに Delay Based Algorithm を下回る結果となった。作成した仮想パケットデータの UDP スループットは 6416.75kbps であり、若干スループットが落ちている。これは経路切り替え時に送信のタイミング遅らせることの累積がオーバーヘッドになっていると推測される。しかし、シミュレーション終了時の各経路の転送量は利用可能帯域幅の比になり、能力に「公平」な送信を実現した。

6 おわりに

本研究では高速で安定な通信を実現するために複数経路を用いる IP データグラム の伝送方式を検討した。複数経路を用いる上で問題となるパケット到着順の入れ替わりをなるべく抑制するような送信アルゴリズムを考案し、また利用可能帯域幅に応じた負荷を分散させた安定した通信を実現するような送信アルゴリズムも考案した。

これらの考案したアルゴリズムの有効性を評価するために、シミュレーションによる実験を行った。実験結果から Delay Based Algorithm は使用経路の遅延差による UDP スループットへの影響はなく、遅延の変動によるパケット到着順の影響が大きいたことが導かれた。通信速度に関しては、複数経路を使用することで単一の経路のみの UDP スループットよりも高速となった。Available Bandwidth Based Algorithm は経路を切り替えるまで入れ替わりが発生しないため、到着順の入れ替わりを抑制した送信を可能にする。加えて、利用可能帯域幅に応じた「公平」な経路使用ができる。どちらのアルゴリズムでも単一の経路で送信する時よりも、複数の経路を使用することで高速化かつ負荷を分散させることができた。また、実験ネットワークを構築し、考案したアルゴリズムの実装も行った。

本研究は、UDP 特に RTP を用いたマルチメディアアプリケーションの通信に効果を発揮することが予測される。考案したアルゴリズムを用いることで、高速・安定した通信品質の送信を可能にし、かつネットワークへの負荷を分散させる送信を実現する。

今後の課題としては、シミュレーションで評価できなかった TCP スループットを実験ネットワークで測定すること、パケット到着順が入れ替わることでの UDP ストリーミングの感覚評価などが挙げられる。

参考文献

- [1] Mo, J., La, R.J., Anantharam, V., Walrand, J.: "Analysis and comparison of TCP Reno and Vegas," INFOCOM(3), 1556-1563, 1999.
- [2] Handley, M., Floyd, S., Padhye, J., Widmer, J.: "Equation-Based Congestion Control for Unicast Applications," SIGCOMM 2000, August 2000.
- [3] Han, H., Shakkottai, S., Hollot, V.C., Srikant, R., Towsley, D.: "Overlay TCP for Multi-Path Routing and Congestion Control," Proc. of IEEE INET 2000, July 2000.
- [4] 元井郁, 柳田陽平: "複数経路を用いた IP データグラムの配送方法とその性能評価," 南山大学情報通信学科卒業論文 (2006.1).
- [5] 吉田秀考: "インターネットにおける可用帯域幅の推定と遅延の評価," 南山大学大学院数理工学情報研究科修士論文 (2006.1).