

# マルチホームネットワークにおける帯域を考慮したバックボーン選択手法

岡山 聖彦<sup>†</sup> 山井 成良<sup>††</sup> 宮下 卓也<sup>††</sup>

## 概要

マルチホームネットワークに接続された複数のバックボーンを利用して動的なトラフィック分散を実現するために、各バックボーンから通信先までのネットワークの状態を反映した選択基準として、TCP のコネクション確立時間を利用して適切なバックボーンを選択する方法が提案されている。しかし、この方法はRTT(ラウンドトリップ時間)の計測と同義であるので、必ずしもスループットの向上に結びつかない。そこで本論文では、バックボーンを選択基準として、連続したパケットを各バックボーン経由で送付することにより、利用可能な帯域を推定する方法を提案する。

## A Backbone Selection Method Based on Bandwidth for Multihomed Networks

Kiyohiko Okayama<sup>†</sup> Nariyoshi Yamai<sup>††</sup> Takuya Miyashita<sup>††</sup>

### Abstract

To achieve dynamic traffic ballancing on multihomed networks, which are connected to the Internet with more than one backbone, the backbone selection method based on measurement of TCP connection setup time via each backbone has been proposed. However, because this method is equivalent to measurement of the round trip time(RTT), it does not necessarily improve the throughput. In this paper, we propose a backbone selection method which can estimate available bandwidth of each backbone by sending sequence of packets via each backbone.

## 1 はじめに

近年、インターネット利用の急激な増加により、WWW、FTPなどの広域ネットワークサービスにおける応答時間の悪化が深刻な問題となってきた。これに対処する一つの方法として、自組織のネットワークを複数のバックボーンネットワーク(以下、単にバックボーンと呼ぶ)と接続し、通信先に応じて利用するバックボーンを使い分けることにより応答時間の改善を図るマルチホームネットワークが注目されている。

マルチホームネットワークでトラフィックを分散する場合、従来の経路制御方法ではバックボーンから入手した経路情報と通信先アドレスのみで利用するバックボーンが一意に定まるため、通信先に偏りが生じると効率的なトラフィック分散が行われず、特定のバックボーンにトラフィックが集中する危険性がある。

さらに、一般にマルチホームネットワークでは接続先バックボーンから経路情報を入手できるようにバックボーン管理者と協調して設定作業を行う必要があり、導入や管理にかなりの技術レベルと管理コストが要求される点も問題であった。

これらの問題を解消するため、複数のバックボーンと自組織のネットワークとの接続を受け持つルータ(以下、特に説明がない限りルータと表記する)において、すべてのバックボーンを用いてコネクション確立を試みることにより、コネクション単位で最も応答の速いバックボーンを選択する方法[1](以下、従来手法という)が提案されている。

しかし、従来手法はラウンドトリップ時間(RTT)の計測と同義であるため、ネットワークサービスの応答時間は改善されるものの、各バックボーンを経由する経路で利用可能な帯域を正確に反映しておらず、必ずしもスループットの向上に結びつかない。このため、ファイル転送などのスループットを重視するサービスにおいては、効率的なトラフィック分散が行えない可能性があるという問題がある。

そこで本論文では、連続したパケットを各バック

<sup>†</sup>岡山大学工学部, Faculty of Engineering, Okayama University

<sup>††</sup>岡山大学総合情報処理センター, Computer Center, Okayama University

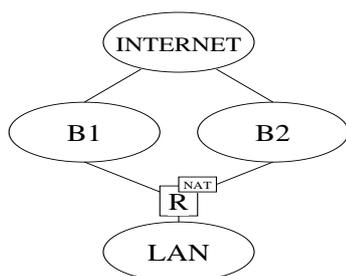


図 1: マルチホームネットワークの構成

バックボーン経由で送出し、応答パケットの到着間隔を計測することにより、利用可能な帯域を推定する方法と、RTTと帯域推定とを併用する方法を提案する。以下、従来手法の問題点について述べた後、提案手法の概要と、その有効性を検証するために行なった性能評価実験について述べる。

## 2 従来の動的トラフィック分散手法の概要と問題点

まず、本論文が前提とするマルチホームネットワークの構成を図1に示す。図において、LANは自組織のネットワークであり、ルータRにより、2つのバックボーンB1、B2に接続されている。自組織のネットワークアドレスはB1側から与えられているものとし、バックボーンB2とはNAT[2][3]を経由して通信するように設定されているものとする。なお、ここでは簡単化のため、バックボーン数を2としているが、3以上の場合でも対応することができる。この場合、3番目以降のバックボーンは、B2と同様にNAT経由で通信するように設定すればよい。

従来手法において、ルータは内部から外部へのコネクションを監視し、内部から外部に向けて新たなコネクション確立要求が発生すると以下のように動作する。

1. Rは内部から外部へのSYNパケットを受け取ると、このパケットを複製して全てのバックボーンに送出する。このとき、B2を経由するSYNパケットについては、送信元のアドレスをB2のアドレスに変換する。
2. Rはあるバックボーンから上記のSYNパケットに対する最初のSYN+ACKパケットを受け取ると、そのパケットを(B2経由の場合にはアドレス変換を行った上で)本来の送信先に中継する。また、今後このコネクションに対してこのバックボーンを利用するように記録する。

3. Rは他のバックボーンから2番目以降のSYN+ACKパケットを受け取ると、そのパケットを中継せずに破棄し、代わりにそのバックボーンを用いてRSTパケットを送出する。

従来手法において、復路、すなわち、外部から内部への応答パケットの経路は、NATを利用して制御している。具体的には、往路においてB1を選択した場合にはアドレス変換を行わず、B2を選択した場合にはNATを利用して送信元のアドレスをB2から与えられたアドレスに変換してパケットを送出する。これにより、復路のパケットは往路と同じバックボーンを経由してルータに届くので、往路と復路を一括して制御することにより、効率的なトラフィック分散を実現している。

以上のように、従来手法では、自組織ネットワーク内のクライアントがインターネット上のサーバと通信する場合、コネクション確立要求を受け取ったルータがすべてのバックボーンを用いてサーバへのコネクション確立要求を試み、最も応答の速いバックボーンを利用する。このため、通信先に偏りが生じた場合でもコネクション毎に個別のバックボーンを用いて効率的にトラフィックを分散することが可能であり、しかも、各バックボーンとの間で経路情報を交換する必要がないため、導入や管理が容易である。

しかし、バックボーンを選択基準としてコネクション確立要求を用いることは、RTTの計測と同じ意味を持つ。RTTは、通信先までの距離(ホップ数)や経路上にあるルータの性能や負荷、経路上の各リンクの帯域や利用率などの要素により影響を受けて変化する。ただし、通信先までのホップ数による影響が大きいことから、利用可能な帯域が狭くても、ホップ数が小さければそのバックボーンを選択するケースが多くなると考えられる。このため、従来手法は応答時間が重視される対話型のサービスに対しては有効である反面、スループットが重視されるファイル転送などのサービスにおいては、必ずしも効率的なトラフィック分散が行なえないという問題がある。

## 3 帯域を考慮したバックボーン選択手法の提案

### 3.1 帯域推定によるバックボーン選択手法

2で示した問題点を解決するため、通信先の経路までの利用可能な帯域を推定することによって適切

なバックボーンを選択する方法(以下、帯域推定方式という)を提案する。

ネットワークの帯域を測る1つの方法として、連続パケットの到着間隔の測定によるもの[4]がある。この方法では、ネットワークに連続して複数のパケットを送出し、それぞれのパケットの到着間隔が、ネットワーク内の実際の帯域や他の通信パケットによって変わることを利用して帯域を推定する。

例えば、ネットワークの途中にボトルネックとなるリンクがある場合、この部分で連続パケットの到着間隔が広がるため、これらのパケットに対する応答パケットの到着間隔も広がることになる。同様に、経路途中のネットワークが混雑している場合、他の通信パケットが連続パケットの間に割り込むことになり、パケットの到着間隔は割り込んだ他のパケットの送信時間の間だけ広がることになる。

したがって、帯域推定方式ではこの原理を利用し、マルチホームネットワークに接続された各バックボーンに対して連続したパケットを同じ間隔で送出し、その到着間隔が小さいものを利用可能帯域が広いバックボーンとみなして選択する。ただし、文献[4]の方法では測定の精度を高めるために多数のパケットを連続して送出するので、コネクション要求ごとに測定を行なおうとすると、コネクション確立に要する時間が増大すると共に、ネットワークに多大の負荷がかかることになる。このため、本論文で提案する帯域推定方式では、必要最低限の数である2つの連続したパケットを用いることにする。

帯域推定方式では、従来手法と同様に、TCPのコネクション確立手順である3ウェイハンドシェイクを利用する。ルータは全てのバックボーンから、サーバに対して連続した2つのコネクション要求を送る。それに対する応答を全て受け取った後に、それぞれのバックボーンにおいて、連続したコネクション確立要求に対する応答の到着間隔を調べ、それが最も小さいバックボーンを選択する。選択されなかったバックボーンと、帯域推定のために新たに生成したコネクション確立要求に対しては、それぞれ対応するRSTフラグ付きパケットをサーバに送信して、コネクションを破棄する。

なお、ルータでは、全ての応答が返ってきたかどうかという情報と、応答を受け取った時刻とを記録し、比較するためのコネクション確立待ちテーブルを用意する。そして、コネクション確立要求が出されるたびにエントリを追加、また、コネクションが確立されるたびにエントリを削除する。

しかし、このような手順では、応答パケットが全て揃わなければ比較が行なえないため、いずれかの

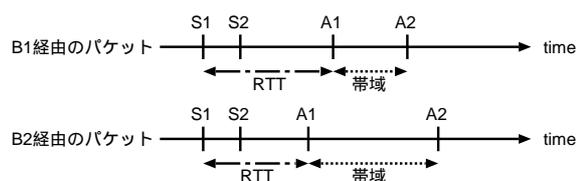


図 2: ルータにおけるパケットの送受信間隔の例(併用方式)

バックボーンに障害が発生して通信できなくなっている場合、バックボーンを選択することができないという問題がある。また、あまりにもRTTが長い場合は、利用できる帯域を有効に活用できないとも考えられる。したがって、ルータがコネクション確立要求を受け取ってからバックボーン選択するまでにタイムアウトを設けて、タイムアウトした場合には、既に返ってきている応答のみで利用するバックボーンを選択を行うことにする。

### 3.2 RTTと帯域推定併用によるバックボーン選択手法

3.1で述べた帯域推定方式では、そのコネクションのデータ転送におけるスループットは向上すると考えられるが、従来手法に比して、バックボーン選択に時間がかかり、コネクション確立に要する時間やタイムアウトの実装が問題になると考えられる。そこで、もう1つの方法として、RTTと帯域推定の併用によるバックボーン選択手法(以下、併用方式という)を提案する。

具体的には、帯域推定方式を簡略化することで実現する。すなわち、全てのバックボーンからの応答を待つのではなく、同一バックボーンからの2つのコネクション確立要求に対する応答が到着した時点で、そのバックボーンを選択するものとする。

例として、ルータにおけるバックボーン選択時のパケット送受信タイミングを図2に示す。組織内部のクライアントから外部のサーバに対してコネクション確立要求(SYNパケット)が発生した場合、ルータはこれを複製し、まず、1番目のSYNパケットをB1およびB2経由でほぼ同時にサーバに送出し(S1)、続けて2番目のSYNパケットをB1およびB2経由でほぼ同時にサーバに送出する(S2)。図2において、S1に対する応答(SYN+ACK)パケットの到着がA1、S2に対する応答パケットの到着がA2にそれぞれ対応しており、S1~A1の間隔がRTT、A1~A2の間隔が利用可能帯域を意味している。したがって、それぞれの間隔の和が小さい方、すなわち、A2

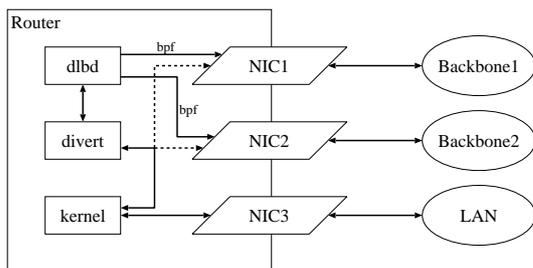


図 3: 試作ルータ内部の構成

の時刻が早い方のバックボーンを選択すれば、RTTと利用可能帯域の両方を考慮したバックボーン選択を行なっていることになる。

このように、併用方式ではすべての応答パケットを待つ必要がないので、帯域推定方式に比して、バックボーン選択に要する時間、すなわち、コネクション確立に要する時間が改善されると共に、障害などの理由によって応答パケットが返ってこないバックボーンに対しても、特別にタイムアウトを設けることなく対処できる。さらに、帯域だけでなく、RTTも考慮して、遅延の大きすぎるバックボーンを排除することができる。また、実装においても、タイムアウトなどの煩雑な処理を必要としないので、帯域推定方式に比べ、比較的容易に実装が可能であると考えられる。

## 4 実装と性能評価

### 4.1 実装方法

試作ルータは、OSとしてFreeBSD4.1-RELEASEを搭載したAT互換機に、3枚のネットワークインタフェースを装着したものをを用いた。本来、FreeBSDでは経路制御がカーネルで行われるが、試作ルータでは実装を容易にするためにカーネルを全く変更せず、ユーザプロセスで経路制御処理を行うようにした。このため、各バックボーンから到着するパケットはすべてdivert[5]機能を用いてカーネルが処理する前に横取りし、また、バックボーンへのパケット送出手は、従来のソケットを用いて行った場合カーネルによって処理されてしまうため、bpf(Berkeley Packet Filter)[6]を用いて直接ネットワークインタフェースにパケットフレームを書き出す処理を行うようにした。また、内部ネットワークから外部ネットワークへの通信パケットは、デフォルトのネットワークインタフェースへ送られるところでdivert機能によりユーザプロセスへ渡されるようにした。

図3に試作ルータの内部構成とパケットの経路を示

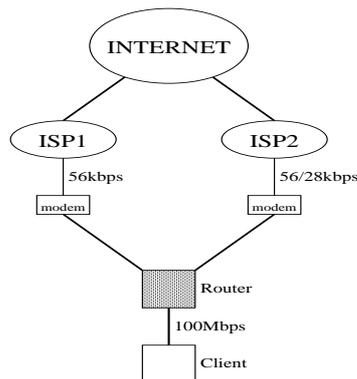


図 4: 実験環境

す。試作ルータは3枚のネットワークインタフェースカードNIC1~NIC3を持ち、NIC1とNIC2はそれぞれバックボーン1、バックボーン2と接続し、NIC3は自組織ネットワーク(LAN)に接続している。まず、内部から外部へのパケットは、NIC3からカーネルに渡され、デフォルトのインタフェースであるNIC1へと送られる。このときdivert機能によりパケットは横取りされて、動的トラフィック分散を行うユーザプロセスdlbdにパケットが渡される。dlbdではNATのコネクション管理テーブルから利用するバックボーンに従い、必要ならばアドレス変換を施し、bpfを用いて選択されたインタフェースに出力する。次に、外部から内部へのパケットは、NIC1あるいはNIC2からカーネルに渡される前にdivert機能によりdlbdに渡される。パケットは、dlbdにおいて必要ならばアドレス逆変換を施され、再びdivertを通じてカーネルに渡され、カーネルからNIC3へと出力される。

### 4.2 性能評価実験の方法

性能評価実験は、図4に示すように、クライアントとインターネットの間に試作ルータを配置し、実際にクライアントからインターネット上のサーバとHTTPによるデータ通信を行うことによって実施した。試作ルータとインターネットの間はダイヤルアップ接続された2種類のISPによって接続され、各モデムからISPまでのダイヤルアップ回線をバックボーンとみなした。また、それぞれのダイヤルアップ回線の速度として、56Kbpsあるいは28Kbpsの2種類を切り替えて用いた。

この実験環境において、クライアントからインターネット上のHTTPサーバに対して同時に複数のデータ転送要求を発生させ、試作ルータにおいてトラフィック分散を行った。

また、バックボーン選択手法として、従来手法(以下、RTT方式という)、帯域推定方式、併用方式の3つを用い、それぞれについて、平均スループットとバックボーン選択率を求めた。

なお、HTTPリクエストの発生には webjamma[7]を一部改造したものをを用い、webjamma実行時のURLリストには、岡山大学総合情報処理センターに設置されたHTTPプロキシサーバのアクセスログから連続した1800個のURLを取り出して用いた。本来 webjamma では、親プロセスが一定数の子プロセスを生成し、各子プロセスがHTTPリクエストを親プロセスから受け取りHTTPプロトコルによるアクセスを行い、データ転送が終了次第、次のリクエストを親プロセスから受け取りデータ転送を行う。したがって、常時、子プロセスの数だけ並行してアクセスが発生する。しかし、このようなアクセスパターンは現実とは大きく異なるため、本実験では親プロセスがHTTPリクエストを渡す間隔が指数分布に従うように webjamma を改造した。HTTPリクエストの平均発生間隔については、小さすぎるとバックボーンが常に飽和状態になり、大きすぎると次のコネクション要求が発生するまでに既存のコネクションが終了してしまい、いずれの場合も有意な計測結果を得ることができない。このため、予備実験を数回を行い、計測時のHTTPリクエスト平均発生間隔を0.6秒とした。

実際の計測は、バックボーンの回線速度(対称および非対称)と、バックボーン選択手法との全ての組合せで行った。なお、いくつかのURLにおいて、URL自体が存在しなかったり、計測時にHTTPサーバがダウンしていた、データ転送がタイムアウトをおこした等の理由で正常にデータ転送が行えなかったURLがいくつか存在したが、このようなURLは実験結果から除外している。

## 4.3 実験結果と考察

### 4.3.1 平均スループット

実験結果として、平均スループットおよびバックボーンの平均選択率を、それぞれ、表1および表2に示す。

まず、回線速度が対称の場合について述べる。表2からわかるように、全ての選択手法においてISP1の利用率が高くなっていることがわかる。さらに表1から、全体のスループット、1コネクション当りのスループット共に、帯域推定方式が最も高い値を示しており、続いて、併用方式、RTT方式の順に低くなっている。RTT方式では効果的にトラフィック分

表 1: 平均スループット

	選択手法	平均スループット (Kbps)	
		56k/56k	56k/28k
全体の スループット	帯域推定	48.0	49.5
	併用	45.3	52.2
	RTT	40.8	43.0
1コネクショ ンあたりの スループット	帯域推定	6.5	3.7
	併用	5.3	3.9
	RTT	3.8	2.4

表 2: 平均バックボーン選択率

		ISP1	ISP2
ISP1 56Kbps ISP2 56Kbps	帯域推定	73.99%	26.01%
	併用	76.59%	23.41%
	RTT	57.86%	42.14%
ISP1 56Kbps ISP2 28Kbps	帯域推定	81.48%	18.52%
	併用	73.95%	26.05%
	RTT	71.29%	28.71%

散が行なわれておらず、平均スループットが低い。これに対して、帯域推定方式や併用方式では、ISP1の利用率が高く、平均スループットが高くなっている。この結果から、ISP2よりISP1の方が利用可能な帯域が広く、スループットと推定される帯域との間に強い相関関係があることがわかる。

次に、回線速度が非対称の場合について述べる。表1および表2からわかる通り、いずれの選択手法もネットワーク状態の変化に対応し、回線速度の速いISP1の経路を選択して平均スループットが高くなっている。特に、併用方式によるバックボーン選択手法における平均スループットは最も高くなっている。

以上のことから、平均スループットを向上させるには、帯域推定方式と併用方式が適していると考えられる。また、スループットと帯域との間に強い相関関係があることが明らかになり、スループットの向上にはRTTよりも帯域が重要であることが確認された。

### 4.3.2 信頼性

次に、各バックボーン選択手法の信頼性について述べる。表3に正常に転送の行なわれたURL数を示す。

表3から、いずれの回線速度の場合もRTT方式が最も正常に転送されたURL数が多いことがわかる。RTT方式では、ISP1とISP2両方のバックボーンを経由してコネクション確立を試み、応答があった時点でそのバックボーンが選択され、選択されたバック

表 3: 正常に転送の行なわれた URL 数

回線速度	選択手法	URL 数
ISP1 56Kbps	帯域推定	1192
ISP2 56Kbps	併用	1237
	RTT	1773
ISP1 56Kbps	帯域推定	1308
ISP2 28Kbps	併用	1295
	RTT	1775

クボーン経由でコネクションが確立される。つまり、通常のコネクション確立と同じ時間で、コネクションを確立することができ、さらに、片方の ISP 経由の経路に障害があっても問題がないため、正常にコネクションを確立できた URL 数が最も多くなったと考えられる。これに対し、帯域推定方式と併用方式では、バックボーン選択に必要なパケット数が多いため、コネクションが確立されるまでの時間が長くなり、また、コネクション確立要求のパケットが連続して送られるため、途中でパケットが失われる確立も高くなっていると考えられる。その結果、webjamma におけるタイムアウトによって HTTP リクエストが破棄され、正常にコネクションを確立できた URL 数が低下したと考えられる。

以上の結果から、RTT 方式が障害に強く、最も信頼性の高いバックボーン選択手法であると言える。帯域推定方式および併用方式は、RTT 方式に比して正常に転送された URL 数が大幅に減少しているが、ルータの実装において、タイムアウト処理やコネクション確立要求の再送処理等に改善の余地があると考えられる。特に、併用方式については、各バックボーンに送出する 1 番目のコネクション要求パケットはほぼ同時であるため、1 番目のコネクション要求パケットに対する応答パケットは無視して、2 番目のコネクション要求パケットに対する応答パケットのみを考慮するように実装を変更することにより、信頼性の大幅な向上が期待できる。

## 5 おわりに

本論文では、マルチホームネットワークで効率的な動的トラフィック分散を行なう手法として、従来のコネクション確立時間、すなわち、RTT に基づいたバックボーン選択手法に加え、帯域推定に基づいたバックボーン選択手法と、RTT と帯域推定を併用したバックボーン選択手法を提案した。実際に実験ネットワークを用いて性能評価実験を行なった結果、スループットについては、提案手法が有効であ

ることが確認された。

ただし、バックボーンを選択を行なう仕組みの違いにより、信頼性については RTT に基づいたバックボーン選択手法が優れていたため、今後は提案手法の実装を改善すると共に、サービスの種類に応じてこれらのバックボーン選択手法を使い分けるための仕組みについて検討を行なう予定である。

## 参考文献

- [1] 岡山聖彦, 山井成良, 島本裕志, 宮下卓也, 岡本卓爾: “マルチホームネットワークにおける透過的な動的トラフィック分散”, 情報処理学会論文誌, Vol.41, No.12, pp.3255-3264 (2000).
- [2] Egevang, K. and Francis, P: “The IP Network Address Translator (NAT)”, RFC1631 (1994).
- [3] Srisuresh, P. and Holdrege, M: “The IP Network Address Translator (NAT) Terminology and Considerations”, RFC2663 (1999).
- [4] Carter, R. L. and Crovella, M. E.: “Measuring Bottleneck Line Speed in Packet-Switched Networks”, Tech. Rep. BU-CS-96-006, Boston University (1996).
- [5] Cobbs, A: “divert - kernel packet diversion mechanism”, <http://www.FreeBSD.org/cgi/man.cgi>.
- [6] McCanne, S: “bpf - Berkeley Packet Filter” <http://www.FreeBSD.org/cgi/man.cgi>.
- [7] R.P. Wooster and M. Abrams: “Proxy Caching that Estimates Page Load Delays”, Proc. 6th World Wide Web Conference, pp.325-334 (1997).