

## マルチホーム環境における DNS 応答の多重化による自組織宛メール配送の動的経路選択手法

金 勇<sup>†1</sup> 山井 成良<sup>†2</sup> 岡山 聖彦<sup>†2</sup>  
清 家 巧<sup>†3</sup> 中 村 素典<sup>†4</sup>

電子メールにおいて処理能力や耐障害性の向上を実現する方法として、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を効率良く利用するにはネットワークの構成や利用状況に応じて適切な経路を選択する必要がある。しかし、従来の経路選択手法ではネットワークの構成を考慮しないため、必ずしも適切な経路が選択されなかったり、ネットワークの利用状況に応じた動的なトラフィック分散ができなかったりするなどの問題があった。そこで本論文では自組織宛メール配送を対象とした動的経路選択手法を提案する。本手法では配送時に行われる DNS (Domain Name System) サーバへの問合せに対して内容の異なる応答を複数の経路から同時に返し、最初に応答が到着した経路を選択させる。これによりネットワークの構成や利用状況に応じた適切な経路を選択でき、また障害が発生している経路を自動的に迂回することが可能になる。

### A Dynamic Route Selection Method Using Multiple DNS Replies for Inbound E-mail Delivery on Multihomed Environment

YONG JIN,<sup>†1</sup> NARIYOSHI YAMAI,<sup>†2</sup> KIYOHICO OKAYAMA,<sup>†2</sup>  
TAKUMI SEIKE<sup>†3</sup> and MOTONORI NAKAMURA<sup>†4</sup>

As a way to improve throughput and fault tolerance of E-mail system, multihomed network is taken into account. To take full advantage of multihomed network, it is required to conduct proper route selection based on topology and utilization statistics of network. However, most conventional operation schemes have some issues such that do not perform appropriate route selection nor dynamic traffic balancing since they do not take network topology into account. In this paper, we propose a dynamic route selection method for inbound e-mail delivery. In this method, different replies corresponding to a DNS (Domain Name System) query launched right before e-mail delivery are sent back simul-

taneously through multiple routes and the one through which the first arrived reply was delivered is selected for e-mail delivery. With this method, an appropriate route based on the topology and utilization statistics of network can be selected for inbound e-mail delivery. Furthermore, unavailable routes with failure can be avoided automatically as well.

#### 1. はじめに

電子メールは WWW と同様にインターネットにおいて最も普及しているサービスの 1 つであり、社会的な活動を支える通信手段としてもはや必要不可欠な存在となっている。本来、電子メールは実時間性を必要とするサービスではないが、最近では緊急情報や安否情報のような重要な情報が頻繁に交換されるため、受信側にとってはたとえば MTA (Mail Transfer Agent) やネットワーク障害により電子メールの配送が大きく遅延するような状況はできる限り回避することが望ましい。

一方、広告や詐欺などを目的に不特定多数の利用者に一方的に送信される spam メールは電子メール全体の流通量の 85%以上を占め、また画像を含む比較的大きい spam メールが増加している状況<sup>1)</sup> にあり、電子メールの利用を阻害する要因になっている。特に、Joe job や backscatter など、spam メールの配送に起因する大量の配送不能通知メールの集中は事実上サービス不能攻撃と同等であり、通常の電子メールの配送に大きな影響を及ぼす<sup>2)</sup>。このような状況(以下、高負荷時)に対処するには、受信側組織において大量の電子メールを速やかに受信して処理できるようにすることが必要である。

上記のような観点から、受信側組織では耐障害性や処理能力の向上を目的として MTA やネットワークの多重化が図られることが多い。このうち MTA の多重化は比較的容易に実現できるため、以降ではネットワークの多重化について考える。

ネットワーク多重化を実現する代表的な手法として、自組織のネットワークを複数のパツ

<sup>†1</sup> 岡山大学大学院自然科学研究科  
Graduate School of Natural Science and Technology, Okayama University

<sup>†2</sup> 岡山大学総合情報基盤センター  
Information Technology Center, Okayama University

<sup>†3</sup> アドソル日進株式会社  
Ad-Sol Nissin Corporation

<sup>†4</sup> 国立情報学研究所  
National Institute of Informatics

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マルチホームネットワークが最近注目されている。マルチホームネットワークの構成方法にはいくつかの方法があるが、電子メールでは各バックボーンに ALG (Application Level Gateway)<sup>3)</sup> に相当するメールゲートウェイ（以下、MGW）を設置する方法がよく用いられる。

自組織のネットワークを複数のバックボーンに単に接続しただけでは、これらの使い分けができず、耐障害性の向上やトラフィック分散による性能の向上を図ることができない。そこで、マルチホームネットワークの利点を活かすためには適切な経路選択<sup>\*1</sup>が必要になる。すなわち、送信側から受信側への各バックボーンを経由する経路のうち、障害が発生している経路を迂回させ、また高負荷時にはネットワークの利用状況に応じてトラフィックを分散させる必要がある。また、トラフィックが少ない場合（以下、低負荷時）でも、全体のトラフィックの削減や通信時間の短縮を図るため、他の条件が同じであれば利用可能な経路のうち遅延時間が小さい（以下、近い）、帯域が広い（以下、速い）、あるいはトラフィックが少ない経路を選択させることが望ましい。以下では、特に断りのない限り、このような経路をまとめて適切な経路と呼ぶ。

上記の目的のため、種々の経路選択方法がこれまでに提案されており、代表的なものとしては DNS ラウンドロビン<sup>4)</sup>、DistributedDirector<sup>5)</sup> などがある。これらはいずれも本来は同一の役割を果たす複数のサーバ間での負荷分散を目的としているが、ALG を用いたマルチホームネットワーク環境では経路選択にも利用できる。ところが、これらの方法では、経路中の特定の箇所にトラフィックが集中したり、必ずしも適切な経路を選択させられなかったりするなどの問題があった。

そこで、本論文では MGW を用いたマルチホームネットワークにおいて上記の問題を解決するため、他組織から自組織宛に送られてくる電子メールの配送（以下、自組織宛メール配送）における動的経路選択手法を提案する。本提案手法では、送信者が電子メールを送る際に行う MX (Mail eXchange) レコードの問合せ（以下、MX 問合せ）に対して、自組織側から複数の経路を経由して内容の異なる MX レコードを含む応答（以下、MX 応答）を同時に返すようにする。これにより、最初に送信者側に到着した MX 応答が有効になるため、輻輳の発生により遅延時間が増加した経路を回避させてネットワークの利用状況に応じたトラフィック分散を図ることが可能になる。また、ある経路上に障害が発生した場合、そ

の経路を経由する MX 応答は届かず、利用可能な経路を経由した MX 応答のみが有効になるため、自動的な障害回避も可能となる。さらに、ネットワークの構成や利用状況に応じて送信者に適切な経路を選択させることが可能となる。

以下、2 章では、従来の ALG を用いたマルチホーム環境における経路選択手法とその問題点について述べる。次に、3 章では本論文で提案するネットワークの状態を考慮した動的経路選択手法の設計について述べ、4 章では提案手法の実装と有効性を確認するために行った動作確認実験と結果、また提案手法に潜在する問題点と対策について述べる。最後に、5 章では結論と今後の課題について述べる。

## 2. 従来の自組織宛メール配送の経路選択手法とその問題点

### 2.1 対象となるネットワーク構成

マルチホームネットワークは 1 つのネットワークを複数のバックボーンによりインターネットに接続する形態で、耐故障性の向上やトラフィック分散による通信性能の改善などを図る方法として注目されている。マルチホームネットワークの構成方法には AS (Autonomous System) 番号を取得する方法（方法 1<sup>6)</sup>）、NAT (Network Address Translation)<sup>7)</sup> を用いる方法（方法 2<sup>8)</sup>）、ALG を用いる方法（方法 3<sup>9)</sup>）などがある。これらのうち、方法 1 は同一の自組織 IP アドレスを用い、ネットワーク層での経路制御によりバックボーンを選択する方法である。この方法では BGP (Border Gateway Protocol)<sup>10)</sup> を用いて経路制御を行う必要があり、導入や運用にかなりの技術レベルと管理コストが必要となる点が問題となる。一方、方法 2 および 3 ではいずれもバックボーンごとに異なる IP アドレスを用い、トランスポート層あるいはアプリケーション層でバックボーンを選択する方法である。これらの方法では利用可能なアプリケーションが NAT や ALG に対応できるものに限定されるが、導入や運用が比較的容易であるという利点がある。

電子メールシステムでは、多くの組織がセキュリティ対策などを目的として ALG に相当する MGW を導入し、自組織 MTA と他組織 MTA との間でメッセージを中継するような構成を採用している。そこで、以降の議論では、特に断りのない限り図 1 に示すような方式 3 に基づくマルチホームネットワークを想定し、また特に自組織宛メール配送を対象とする。なお、自組織から他組織へのメール配送については、文献 8)、11)、12) で提案している手法などを適用できるため、本論文では対象としない。

### 2.2 従来の経路選択手法と問題点

図 1 のようなマルチホームネットワーク環境では、各 MGW のバックボーン側 IP アド

\*1 通常の hop-by-hop の経路制御とは異なるため、本論文では経路選択と呼ぶ。

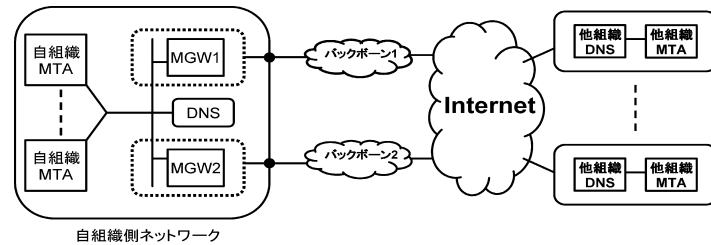


図 1 ネットワーク構成  
Fig. 1 Network configuration.

レスは各バックボーンより割り当てられるため、自組織宛メール配送では MGW の選択がその MGW に対応するバックボーンを経由する経路の選択（以下、経路選択）と等価になる。したがって、複数のバックボーン間でのトラフィック分散のためには、分散配置された複製サーバ間での負荷分散手法が適用できる。

複製サーバ間での従来の負荷分散手法としては、DNS ラウンドロビン、DistributedDirector などが知られている。以下では、これらの手法とその問題点を示す。

### 2.2.1 DNS ラウンドロビン

DNS ラウンドロビンは、1 つのホスト名やドメイン名に複数の DNS 資源レコードを割り当て、クライアント側 DNS キャッシュサーバ（以下、キャッシュサーバ）からの問合せごとに資源レコードの順番を変更しながら応答する手法である。電子メールの場合には、同一優先度の MX レコードが複数存在する場合に適用可能である。この場合、クライアントではどの資源レコードを優先してもかまわないが、多くの実装では最初の資源レコードを優先するため、結果として各サーバおよび各バックボーンが同じ割合で選択されると期待できる。この手法は BIND<sup>13)</sup> などほとんどの DNS サーバの実装で利用可能であるため、広範囲に用いられている。

しかし、この手法ではネットワークの構成や利用状況を考慮していないため、速い経路や高負荷時に輻輳が少ない経路が選択されず、適切なトラフィック分散が行えないという問題がある。また、低負荷時にもこの手法ではネットワークの構成がまったく考慮されていないため、単に MGW の負荷やバックボーンのトラフィックが一定の割合で分散される効果しか期待できない。

### 2.2.2 DistributedDirector

DistributedDirector（以下、DD）はクライアントやサーバに変更を加えることなく、各

クライアントに最も近いサーバを選択させることができるシステムである。具体的には、DD は DNS サーバとして動作し、キャッシュサーバからの問合せを受信した場合、各サーバの近傍のルータと通信してそれらのルータからキャッシュサーバまでの距離を取得し、BGP で用いられる AS パス長、帯域、リンクの負荷、あるいは RTT（Round Trip Time）などに基づいて適切なサーバをキャッシュサーバに返す。これにより DD は複数の経路のうちネットワークの帯域やリンクの利用状況を勘案した適切な経路を選択させることが可能になる。

しかし、距離の尺度のうち、AS パス長を用いる方法は、これだけでは高負荷時にネットワークの利用状況に応じた動的なトラフィック分散が行えず、また 2.1 節で述べたように BGP の導入や運用・管理などにはかなりの技術とコストが要求されるという問題がある。一方、RTT を用いる方法では高負荷時にも対応可能であるが、RTT 測定のコストが大きいという別の問題が生じる。すなわち、DD では DNS 問合せを受けるたびに RTT 測定を数回行うため、測定をすべて完了するのに時間を要するだけでなく、特にネットワークの障害やフィルタリングなどにより一部あるいはすべての RTT 測定が失敗するとタイムアウトが発生するまで待ち続け、経路選択に要する時間が大きくなる点が問題になる。さらに、帯域など、その他の尺度を用いる方法はこれらの情報を収集できる特殊な経路制御プロトコルを用いる必要があり、事実上利用できない。

このように、DD はいずれの距離の尺度を用いる場合でも問題があり、自組織宛メール配送の経路選択にそのまま適用するには適していない。

## 3. ネットワークの状態を考慮した自組織宛メール配送の動的経路選択手法

2 章で述べたように、従来の自組織宛メール配送の経路選択手法にはいくつかの問題があった。そこで本章では、それらの問題を解決するために高負荷時には適切にトラフィック分散を行い、また低負荷時にも適切な経路を選択させる動的経路選択手法（以下、提案手法）を提案する。

### 3.1 使用経路の動的選択

これまで述べたように、図 1 の構成では自組織宛メールの配送に用いられる経路（以下、配送経路）の選択は MGW の選択と等価で、適切な MGW を MX レコードとしてキャッシュサーバに返すことにより実現できる。したがって、自組織宛メールの配送経路をネットワークの構成や利用状況に応じて動的に選択するには、DNS サーバが MX 問合せを受け、各 MGW からキャッシュサーバまでの経路の状態を計測し、返すべき MX レコードをその結果に応じて決定する必要がある。しかし、DD における RTT の測定のように専用の

パケットを送出してネットワークの状態を測定する方法では、前章で述べたように特に測定が失敗した場合に経路選択に要する時間が大きくなる点が問題である。

そこで、本論文では MX 応答パケットを利用してネットワークの状態を測定する手法を提案する。この手法では 1 つの MX 問合せに対して、複数の MX 応答を用意し、それぞれ各バックボーンを経由して同時にキャッシュサーバに返す。各 MX 応答は同一の内容ではなく、その応答が経由するバックボーン、すなわち応答が送出される接続点の MGW を優先する MX レコードを含む。ただし、各 MX 応答の送信元 IP アドレスは、DNS 問合せの宛先アドレスと同じになるように書き換える。DNS では通常は 1 つの問合せに対して最初に返された応答だけが有効になり、それ以外の応答は破棄されるため、これにより最も近い MGW が選択されることになる。また、この手法では少なくとも 1 つの経路が有効であれば MX 応答の受信と同時に経路選択が完了するため、一部の経路に障害が発生している場合でも DD のようなタイムアウト待ちによる経路選択時間の増加は発生しない。耐障害性の向上は DNS ラウンドロビンでも実現できるが、送信側が経路上に障害が発生している MGW を選択するとタイムアウトにより失敗するまで SMTP セッションの確立を待ち続けるのに対して、提案手法ではこのようなタイムアウト待ちが生じない<sup>\*1</sup>点で優れているといえる。

上記の動作は遅延時間に基づく経路選択と等価であり、経路の距離やホップ数、経路上の各リンクの容量や利用率などの要素による影響を受けて変化する。一般に、これらの要素は遅延時間とスループットの間に負の相関関係をもたらす要因になるため、遅延時間が短い経路を選択すると長期的には電子メール配送のスループットも向上するものと思われる。ただし、この手法では 1 つのパケットだけを使って遅延時間を測定するため精度がそれほど高くはなく、また DD における RTT 測定と同様に通信開始後のネットワーク状態の変化を考慮できないため、個々の経路選択が必ずしも適切であるとは限らないことに注意する。

なお、提案手法は、たとえば WWW など、ALG に対応した他のアプリケーションにも適用可能である。その場合には MX レコードの代わりに A レコードを用い、優先する A レコードを DNS 応答の最初に置けばよい。ただし、その有効性が一部制限される点に注意する。すなわち、電子メールにおける MX レコードとは異なり、複数の A レコードのうちどれを優先するかはクライアントの実装に依存するため、必ずしも近い経路が選択されるとは限らない。また、最も近い経路を使用した通信に失敗した場合に必ずしも他の経路を使用して再び通信を試みるとは限らない。

\*1 ただし、MGW に障害が発生している場合などにはタイムアウト待ちが生じる。

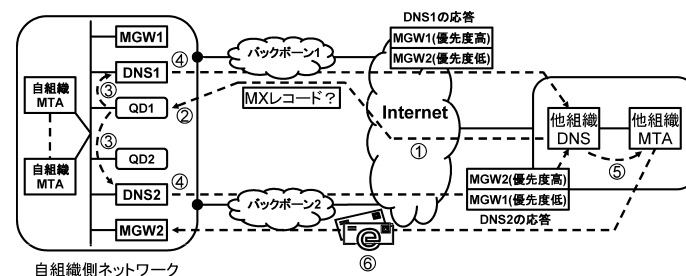


図 2 提案手法のシステム構成と動作

Fig. 2 System configuration and process of the proposed method.

### 3.2 提案手法のシステム構成

以下では、説明の簡単化のため、自組織ネットワークが 2 つのバックボーンにより接続されている環境を想定する。なお、3 つ以上のバックボーンに接続されている環境においても提案手法の適用は容易である。

提案手法のシステム構成を図 2 に示す。この図に示すように、各バックボーンの接続点にそれぞれ MGW (MGW1 および MGW2)、自組織側 DNS コンテンツサーバ (DNS1 および DNS2、以下、DNS サーバ)、および DNS 問合せ複製器 (QD1 および QD2、以下、複製器) を設置する。これらはすべて 2 つのネットワークインタフェースを有し、そのうち一方は自組織内ネットワークに、また他方はバックボーンに接続され、それぞれ接続先から IP アドレスを割り当てられているものとする。

構成要素のうち、複製器はキャッシュサーバ (他組織 DNS) からの MX 問合せを受信するとその問合せを複製して各 DNS サーバに中継する役割を果たす。また、自組織宛メールのドメインに対応する NS (Name Server) レコードにはすべての複製器を指定し、複製装置が単一障害点にならないようにする。各 DNS サーバでは、自組織ドメインに対する MX レコードとして MGW1 と MGW2 の両方を返すように設定しておくが、その DNS サーバと同一バックボーンの接続点に設置されている MGW の優先度が高くなるように設定する。すなわち、DNS1 では MGW1 の優先度を高く、DNS2 では MGW2 の優先度を高く設定する。また、遅延時間の測定頻度を増やして特に高負荷時や他の通信によるネットワーク利用状況の変化の際に適切にトラフィック分散を行えるようにするため、MX レコードの TTL (Time To Live) は小さい値 (たとえば 0) とする。

### 3.3 提案手法の動作

次に、図 2 において他組織 MTA から自組織宛メール配送を行う場合の動作を説明する。なお、以下の説明では図中の数字と手順の番号は対応し、括弧内は図中の例を示している。

- (1) 他組織 MTA は自組織宛メール配送を行う際、キャッシュサーバに宛先ドメインの MX 問合せを行う。キャッシュサーバは宛先ドメインに対する NS レコードに従って、いずれかの複製器 (QD1) に MX 問合せを行う。
- (2) 複製器はキャッシュサーバからの MX 問合せを受信する。
- (3) 複製器は受信した MX 問合せを 2 つに複製し、各 DNS サーバへ中継する。
- (4) 各 DNS サーバは複製器から中継された MX 問合せを受信すると、同一バックボーンに接続されている MGW の優先度が高いような MX 応答をキャッシュサーバに送信する。このとき、MX 応答の送信元 IP アドレスは最初に MX 問合せを受信した複製器 (QD1) のアドレスと同じになるように変換する。
- (5) キャッシュサーバは最初の MX 応答 (DNS2 からの MX 応答) を受信して MX レコードを他組織 MTA に返す。後から到着した MX 応答 (DNS1 からの MX 応答) は、対応する MX 問合せが存在しなくなったため破棄する。
- (6) 他組織 MTA は優先度の高い MGW (MGW2) に電子メールを配送する。

### 3.4 BGP による経路制御との併用

これまでの議論では、図 1 のような ALG を用いたマルチホーム環境を前提としていた。しかし、提案手法は AS 番号を取得して BGP により経路制御を行うようなマルチホーム環境においても適用可能である。すなわち、すべてのバックボーンに障害が発生していない状態において、任意のキャッシュサーバから見た各 MGW への経路が、当該 MGW に対応するバックボーンを経由するものであれば、提案手法は ALG を用いたマルチホーム環境と同様に動作する。また、DD のような AS パス長に基づく経路選択基準と組み合わせ、たとえば AS パス長に大きな差が見られない場合に提案手法を適用することも可能である。これにより、BGP による経路制御の利点を生かしながら、輻輳などにより遅延時間が増加した経路を避けて通信を行うなど、動的トラフィック分散の効果も期待できる。

### 3.5 DNSSEC との併用

DNS は問合せパケットや応答パケットの偽装が行いやすく、セキュリティ的に問題が多いプロトコルである。特に最近では DNS 応答の偽装による攻撃 (DNS cache poisoning) の危険性が指摘されるようになり、その対策として暗号技術による偽装パケットの検出を目的とした DNSSEC<sup>14)</sup> の導入が徐々に進んでいる。そこで提案手法を DNSSEC と併用した

場合の影響について考察する。

提案手法では 1 つの MX 問合せに対して内容の異なる複数の MX 応答を返すが、各 MX 応答に対して署名を作成できるため、いずれの MX 応答も偽装とは見なされず有効である。ただし、MX 応答には署名が含まれるため、そのパケットサイズはその分だけ大きくなり、標準的な上限サイズである 512 オクテットを超えてキャッシュサーバで受信や処理を行えなくなる可能性がある。この場合、提案手法が有効に機能するためには、少なくともキャッシュサーバが EDNS0<sup>15)</sup> に対応し、512 オクテットを超える UDP パケットの送受信が可能である必要がある。一方、DNS ラウンドロビンでは、たとえ 512 オクテットを超える UDP パケットの送受信が行えない場合でも、TCP での再問合せが行われる場合には対応することが可能である。提案手法は単独では TCP での MX 問合せには対応できないが、これに対応するには、UDP での MX 問合せに対しては提案手法の動作を行い、TCP で再問合せがあった場合には DNS ラウンドロビンとして動作する方法が有効であると思われる。

### 3.6 提案手法の問題点と対策

残念ながら提案手法は適用可能な環境に制約がある。以下ではその理由と対策について議論する。

3.1 節で述べたように、提案手法では同一の送信元 IP アドレスを持つ複数の MX 応答を作成し、これらを各バックボーンを経由してキャッシュサーバに送信する。したがって、DNS 問合せが経由しなかったバックボーンから見ると、そのバックボーンから割り当てられていない IP アドレスを送信元とするパケットが自組織側ネットワークから送られてきたことになる。ところが、最近ではセキュリティ対策のためこのようなパケットを送信元 IP アドレスの詐称と見なし、流入防止用フィルタリング (ingress filtering)<sup>16)</sup> を導入するバックボーンが増加している。このようなバックボーンが存在すると、提案手法ではそのバックボーンを経由する経路は選択されず、有効性が減少する。ただし、少なくとも MX 問合せと同じ経路を逆方向に送信される MX 応答は流入防止用フィルタリングを通過できることから、1 つのバックボーンしか存在していない場合と同等の機能・性能は維持できる。

この問題への対処方法としては、AS 番号を取得して BGP により経路制御を行う方法が考えられる。すなわち、自組織ネットワークに対して AS 番号を取得してマルチホームネットワークを構築すると、どのバックボーンに対しても自組織ネットワークで使われる IP アドレスの範囲が広告され、流入防止用フィルタリングの対象外となる。この場合、BGP の導入により運用・管理などにはかなりの技術とコストが要求されるが、これは DD などの従来手法とも共通の問題点であり、提案手法の優位性は残されている。

また、BGP を用いなくても、MX 応答として使われる可能性のある IP アドレスに関して各バックボーンの管理者にフィルタリングを解除するように依頼する方法も考えられる。このフィルタリング解除は特定のアドレスだけを対象としているため、これによるセキュリティ上の問題は事実上無視でき、また技術的にも容易であるため、学術系ネットワークや地域 IX (Internet eXchange) などを中心に、提案手法が適用可能なバックボーンは少なからず存在すると思われる。

なお、フィルタリングが解除されたバックボーンとそうでないバックボーンが混在する場合、MX 問合せが経由したバックボーンに応じてキャッシュサーバに返される MX 応答の個数に差が生じることに注意する。たとえば、図 2 においてバックボーン 1 だけがフィルタリングを解除している場合、バックボーン 2 を経由する MX 問合せに対してはキャッシュサーバには 2 つの MX 応答が返されるが、バックボーン 1 を経由する MX 問合せに対してはバックボーン 2 を経由する MX 応答がフィルタリングされるため、キャッシュサーバには 1 つの MX 応答しか返されない。しかし、このような場合でも前者の問合せについては提案手法は有効に動作するため、ある程度の効果が期待できる。

#### 4. システム実装と動作確認

本章では、提案手法に基づいた試作システムの実装と、試作システムを用いた動作確認実験および実ネットワークでの機能確認実験について述べる。

##### 4.1 試作システムの実装

前章で述べた構成に基づき、我々は試作システムの実装を行った。試作システムの構成を図 3 に示す。

このシステムでは機材の都合上、NIC (Network Interface Card) を 2 枚 (それぞれを NIC1, NIC2 とする) 装備した 1 台の計算機 (OS: FreeBSD 7.0) 上に図 2 における自組織側ネットワークのすべての構成要素を実装した。その際、複数存在する構成要素はできる限り集約し、DNS サーバのみ 2 つのプロセス (DNS1, DNS2) を稼働させ、MGW MGW (自組織 MTA と兼用)、複製器 (QD) は 1 つのプロセスでいずれの NIC を経由した通信でも処理できるようにした。2 つの DNS サーバには BIND 9.5.0 をそのまま用いたが、MX 問合せを待ち受けるポート番号は UDP 53 番以外で互いに重ならないものを選択した。これは DNS の標準ポート番号である UDP 53 番ポートは複製器が使用するためである。また、3.1 節で述べた送信元 IP アドレスの変換機能は MX 問合せの受信と MX 応答の送信が同一計算機上で行われるため不要になるが、その代わりに各 DNS サーバからの応答の送信元

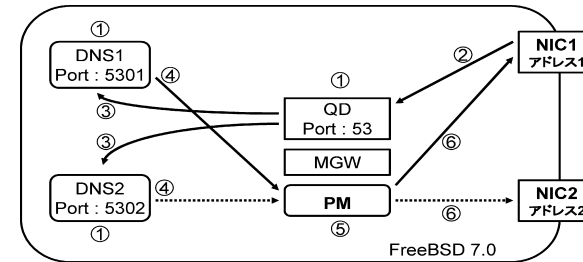


図 3 試作システムの構成と動作

Fig. 3 Configuration and process of the prototype system.

ポート番号を UDP 53 番に戻す必要があるため、FreeBSD の持つ ipfw<sup>17)</sup> および divert<sup>18)</sup> 機能を用いてポート番号変換器 (PM: 以下、変換器) で処理した後に各バックボーンに送出するようにした。

図 3 において MX 問合せを受信した場合の動作を説明する。なお、以下の説明では図中の数字と手順の番号は対応し、括弧内は図中の例を示している。

- (1) 初期状態では DNS1 および DNS2 はそれぞれ異なるポート番号 (DNS1 は UDP 5301 番, DNS2 は UDP 5302 番) で MX 問合せを待ち受ける。また、複製器は UDP 53 番で DNS 問合せを待ち受ける。
- (2) 複製器はキャッシュサーバからの MX 問合せをいずれかの NIC (NIC1) を経由して受信する。その際、MX 問合せの宛先 IP アドレスも得るようにする。
- (3) 複製器は受信した MX 問合せを 2 つに複製し、各 DNS サーバが待ち受ける UDP ポートに中継する。その際、MX 問合せの宛先 IP アドレスは元のアドレス (アドレス 1) を保持するようにする。
- (4) 各 DNS サーバは複製器から中継された MX 問合せを受信すると 3.3 節で述べたような MX 応答を出力する。この応答の送信元 IP アドレスは MX 問合せの宛先 IP アドレス (アドレス 1) と等しく、また送信元ポート番号は各 DNS サーバの待受け用ポート番号 (DNS1 からは 5301 番, DNS2 からは 5302 番) と等しくなる。
- (5) ifpw 機能および divert 機能により、送信元ポート番号が各 DNS サーバの待受け用ポート番号と等しいパケットは変換器に渡される。変換器は送信元ポート番号を 53 番に変更する。
- (6) 変更後の MX 応答は変更前のポート番号に応じて中継先 (next hop) が定められ、中

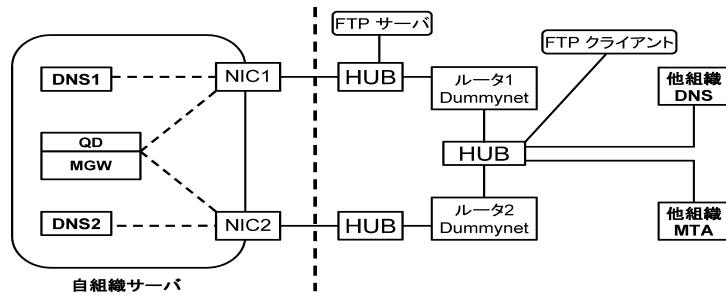


図 4 実験ネットワークの構成  
Fig. 4 Configuration of the experimental network.

継先に対応する NIC を介してバックボーンに送出される．すなわち，変更前のポート番号が 5301 番の場合には NIC1 に，5302 番の場合には NIC2 に出力される．

#### 4.2 実験ネットワークにおける性能評価

提案手法の有効性を評価するため，前節で述べた試作システムを用いた実験ネットワークを構築して実験を行った．実験ネットワークの構成を図 4 に示す．この図において破線より左側が自組織ネットワークの部分を表し，各 NIC とルータとの間はバックボーンを表している．2 つのルータには FreeBSD を搭載した計算機を使用し，dummynet<sup>19)</sup> により伝送遅延や帯域を設定できるようにした．また，FTP クライアントと FTP サーバを図のように配置し，各バックボーンにトラフィックを発生できるようにした．なお，図中のリンクの帯域はすべて 100 Mbps であるが，ルータで帯域を制限している場合がある．

##### 4.2.1 高負荷時における性能評価

この実験ネットワークにおいて，まず高負荷時におけるトラフィック分散の有効性を評価するために，様々な条件で他組織 MTA から自組織 MTA に対して 25 KB のメッセージ 100 通を 20 個のプロセスを用いて連続して配送し，各経路の使用回数と配送完了までの時間を 1 つの条件下につき 10 回測定して平均を求めた．その際，ネットワークを高負荷状態にするため各バックボーンの帯域は 10 Mbps に制限した．具体的な測定の条件は，(1) 両方のバックボーンとも同一の状態である場合，(2) バックボーン 1 に往復とも 2 ms の遅延を付加した場合，(3) バックボーン 1 の帯域をさらに制限 (1 Mbps) した場合，(4) バックボーン 1 上で FTP により他組織から自組織の方向へのトラフィック (3 セッションで合計約 550 KB/s) を発生させた場合，(5) バックボーン 1 上で FTP により自組織から他組織の方向へのトラフィック (3 セッションで合計約 550 KB/s) を発生させた場合の 5 通りであ

表 1 高負荷時における送信時間と各経路の選択回数

Table 1 The transmission time and the number of selection for each backbone network on high load situation.

条 件	評価項目	提案手法	従来手法
(1) 同一状態	経路選択回数	44:56	51:49
	送信時間 (秒)	11.3	11.1
(2) 遅延付加	経路選択回数	26:74	50:50
	送信時間 (秒)	11.6	11.0
(3) 帯域制限	経路選択回数	19:81	51:49
	送信時間 (秒)	14.7	26.4
(4) 同方向負荷	経路選択回数	3:97	50:50
	送信時間 (秒)	29.7	37.0
(5) 逆方向負荷	経路選択回数	2:98	48:52
	送信時間 (秒)	30.9	38.0

経路選択回数は小数点以下四捨五入

る．また，比較の対象として，従来手法 (DNS ラウンドロビン) を用いた場合についても測定を行った．その結果を表 1 に示す．

この結果より，条件 (1) ではどちらの選択手法でもほとんど差がないが，一方のバックボーンにおいて条件 (3) のように帯域が大きく制限されていたり，条件 (4)，(5) のようにトラフィックが発生していたりする場合には，提案手法ではそのバックボーンを迂回し，比較的短時間で送信を終了できるのに対して，DNS ラウンドロビンでは各バックボーンをほぼ均等に利用し送信時間が長くなっていることが分かる．一方，条件 (2) では各経路の選択回数には選択手法により違いが生じたが，送信時間にはあまり差が見られなかった．これは提案手法では遅延時間の小さい経路に SMTP コネクションがある程度集中するが，送信時間に影響を及ぼすほどには集中せず，遅延時間の大きい経路も利用していることを意味する．これらの結果から，提案手法は DNS ラウンドロビンと比較して高負荷時に同等か，より適切にトラフィック分散を行っているといえる．

##### 4.2.2 低負荷時における性能評価

次に，低負荷時における適切な経路の選択機能の有効性を評価するため，同じ実験ネットワークにおいて 1 秒間隔で 100 通の電子メールを配送し，各経路の使用回数を測定した．具体的な条件は，(1) 両方のバックボーンとも同一の状態である場合，(2) バックボーン 1 に往復とも 2 ms の遅延を付加した場合，(3) バックボーン 1 の帯域を 10 Mbps に制限した場合，(4) バックボーン 1 の帯域を 1 Mbps に制限した場合，(5) バックボーン 1 上で FTP により他組織から自組織の方向へのトラフィック (3 セッションで合計約 9.8 MB/s) を発生さ

表 2 低負荷時における各経路の選択回数

Table 2 The number of selection for each backbone network on low load situation.

条 件	提案手法	従来手法
(1) 同一状態	50:50	48:52
(2) 遅延付加	2:98	48:52
(3) 帯域制限 (10 Mbps)	50:50	48:52
(4) 帯域制限 (1 Mbps)	2:98	48:52
(5) 同方向負荷	18:82	48:52
(6) 逆方向負荷	11:89	52:48

小数点以下四捨五入

せた場合, (6) バックボーン 1 上で FTP により自組織から他組織の方向へのトラフィック (3 セッションで合計約 10.2 MB/s) を発生させた場合の 6 通りである. なお, バックボーンの帯域の標準値は 100 Mbps である. また, 比較の対象として, 従来手法 (DNS ラウンドロビン) を用いた場合についても測定を行った. その結果を表 2 に示す.

この結果より, DNS ラウンドロビンではどの条件でも各バックボーンがほぼ均等に利用されるのに対して, 提案手法では条件 (3) を除いて適切な経路が高い率で選択されていることが分かる. 条件 (3) では提案手法においても各バックボーンが均等に利用されているが, これは 10 Mbps への帯域制限では MX 応答の到着順序にほとんど影響を与えなかったことを意味する. これらの結果から, 提案手法は DNS ラウンドロビンと比較して低負荷時の多くの場合に適切な経路選択を行っているといえる. なお, DNS ラウンドロビンにおいて両方のバックボーンの選択回数に差が生じているのは, MX 問合せの再送が発生している場合があるためと思われる.

#### 4.3 実ネットワーク環境での機能確認実験と結果

次に, 実際のネットワーク環境で試作システムが正しく動作するかどうか確認するため, 鳥取県三朝町に位置する岡山大学三朝地区 (以下, 三朝地区) で動作確認実験を行った.

三朝地区は従来は鳥取情報ハイウェイと岡山情報ハイウェイ (OKIX) を経由して岡山大学津島地区 (以下, 津島地区) と接続されており, 学外とのアクセスは津島地区に設置されている SINET 岡山大学ノード経由で行われていたが, 鳥取大学内設置されている SINET 鳥取大学ノード経由でのアクセスが可能になると, こちらを経由する方が高速かつ低遅延で通信できることになる. そこで平成 20 年 8 月に三朝地区のマルチホーム接続を行い, 実運用に向けた準備を行っている<sup>20)</sup>. 本実験はこのネットワーク環境を用いて実施した.

図 5 に実験に用いた実ネットワーク環境を示す. 試作システムは三朝地区に設置され, 2

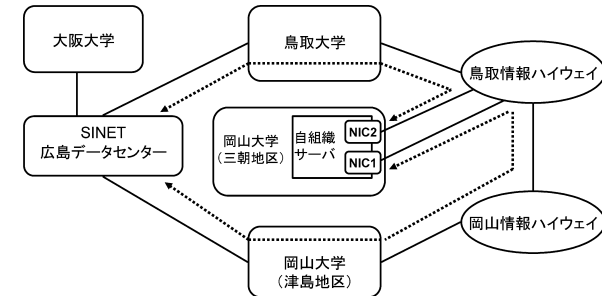


図 5 実ネットワーク環境

Fig. 5 Real network environment.

つの NIC を異なるネットワークに接続し, NIC1 は SINET 岡山大学ノード経由で, NIC2 は SINET 鳥取大学ノード経由で通信できるように構成されている. このうち NIC1 には岡山大学に従来より割り当てられているアドレスを, また NIC2 には SINET から新たに割り当てられたアドレスを用いた. また, SINET 広島データセンター (以下, 広島 DC) では試作システムが出力するパケットを送信元 IP アドレスにかかわらず, 岡山大学ノード, 鳥取大学ノードのいずれから受信した場合でも中継するように, フィルタリングの解除を依頼した. なお, 図 5 で示されるリンクのうち, 最も帯域が狭いのは鳥取情報ハイウェイと三朝地区との間を接続するものであり, その帯域は 2 本の合計で 100 Mbps である.

この環境において, 津島地区および大阪大学から試作システムに向けて電子メールを 1 秒間隔で 100 通送信し, 各配送経路の使用回数を調査する実験を行った. その際, 実験の条件として, (1) 試作システムへの他のトラフィックがない場合, (2) 試作システムへの他のトラフィックとして, 送信元 MTA と同じ場所から試作システムに対して近いほうの経路を指定して連続的に電子メールを送信した場合, の 2 種類の場合について実験した. その結果を表 3 に示す. この表において鳥取経由は SINET 鳥取大学ノードを経由する経路を表す.

なお, この実験では DNS ラウンドロビンとの比較は行わなかった. これは前節の実験結果で示されるように, DNS ラウンドロビンでは両方の経路を均等に利用することが明らかであるためである.

この表から, 条件 (1) の場合には, 送信元が津島地区あるいは大阪大学のいずれであっても近い経路が選択されていることが分かる. 一方, 条件 (2) の場合は一部のメール配送において近い経路が選択されず, 迂回経路が選択されていることが分かる. 実ネットワーク環



表 3 実ネットワーク環境における各経路の選択回数

Table 3 The number of selection for each route on the actual network environment.

他のトラフィック	送信元	鳥取経由	OKIX 経由
(1) なし	津島地区	0 回	100 回
(1) なし	大阪大学	100 回	0 回
(2) OKIX 経由	津島地区	5 回	95 回
(2) 鳥取経由	大阪大学	97 回	3 回

境での実験では一般の通信が混在し、また我々の管理が及ばないネットワークを経由するため、条件 (2) で発生させたトラフィックがどの程度の影響を与えたか不明であるが、このネットワーク環境では発生させたトラフィックによる影響よりも経路による遅延時間の差の影響の方が大きいといえる。しかし、一部のメール配送がトラフィックを発生させた経路を迂回していることから、試作システムは少なくともある程度のトラフィック分散機能を有するといえる。

## 5. む す び

本論文では、マルチホームネットワークにおいて自組織宛メール配送の性能や耐障害性の向上を図り、また適切な経路を選択させるために、内容の異なる MX 応答を複数の経路を経由して同時に返送し、最初に MX 応答が到着した経路を選択させる、自組織宛メール配送の動的経路選択手法を提案した。また、提案手法に基づく試作システムを実装し、実験ネットワークでの性能評価および実ネットワーク環境での機能確認実験を行い、提案手法の有効性を示した。

今後の課題としては、実ネットワークでの運用を通じた性能評価や WWW など他のアプリケーションへ適用した場合の評価があげられる。また、現在の手法では一部のバックボーンには適用できないため、適用範囲の拡大も重要な課題である。さらに、MGW の負荷など、他の要因を考慮した選択基準も今後検討を進めていきたい。

## 参 考 文 献

- 1) Symantec Corporation: State of Spam A Monthly Report (online). [http://eval.symantec.com/mktginfo/enterprise/other\\_resources/b-state\\_of\\_spam\\_report\\_05-2009.en-us.pdf](http://eval.symantec.com/mktginfo/enterprise/other_resources/b-state_of_spam_report_05-2009.en-us.pdf) (accessed 2008-06-10)
- 2) 山井成良: spam メールの現状と対策の動向: 2. 技術的側面から見た spam メール対策: 2.4 バウンスメール対策, 情報処理, Vol.46, No.7, pp.362-766 (2005).

- 3) Srisuresh, P. and Holdrege, M.: IP Network Address Translator (NAT) Terminology and Considerations, RFC 2663, IETF (1999).
- 4) Brisco, T.: DNS Support for Load Balancing, RFC 1794, IETF (1995).
- 5) CISCO Systems Inc.: DistributedDirector (online). <http://www.cisco.com/> (accessed 2008-06-10)
- 6) Hawkinson, J. and Bates, T.: Guidelines for creation, selection and registration of an Autonomous System (AS), RFC 1930, IETF (1996).
- 7) Srisuresh, P. and Egevang, K.: Traditional IP Network Address Translator (Traditional NAT), RFC 3022, IETF (2001).
- 8) 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).
- 9) 中川郁夫, 上谷 一, 鍋島公章, 樋地正浩, 今野幸典: マルチホーム環境におけるアプリケーションルーティング技術の提案, 情報処理学会分散システム/インターネット運用技術研究会研究報告, 98-DSM, No.12-7, pp.37-42 (1999).
- 10) Rekhter, Y. and Li, T.: A Border Gateway Protocol 4, RFC 1771, IETF (1995).
- 11) 岡山聖彦, 山井成良, 久保武志, 宮下卓也: マルチホームネットワークにおけるアプリケーションプロトコルの性質を考慮した動的トラフィック分散, 情報処理学会論文誌, Vol.46, No.4, pp.1007-1016 (2005).
- 12) 山井成良, 土居正行, 岡山聖彦, 中村素典: マルチホームネットワークにおける電子メールシステムの高信頼化運用手法, 情報科学技術フォーラム情報技術レターズ, Vol.6, pp.373-376 (2007).
- 13) Internet Systems Consortium, Inc: ISC BIND (online). <http://www.isc.org/index.pl?sw/bind> (accessed 2008-06-10)
- 14) Eastlake, D.: Domain Name System Security Extensions, RFC 2535, IETF (1999).
- 15) Vixie, P.: Extension Mechanisms for DNS (EDNS0), RFC 2671, IETF (1999).
- 16) Baker, F. and Savola, P.: Ingress Filtering for Multihomed Networks, RFC 3704, IETF (2004).
- 17) Antsilevich, U.J.S., Kamp, P.-H., Nash, A., Cobbs, A. and Rizzo, L.: ipfw-IP firewall and traffic shaper control program, *FreeBSD System Manager's Manual* (2006).
- 18) Cobbs, A.: divert - kernel packet diversion mechanism, *FreeBSD Kernel Interface Manual* (2004).
- 19) Rizzo, L.: dummynet - traffic shaper, bandwidth manager and delay emulator, *FreeBSD Kernel Interface Manual* (2002).
- 20) 山井成良, 岡山聖彦, 金 勇, 河野圭太, 大隅淑弘: 岡山大学における地域 IX と SINET を利用したネットワーク冗長化, 情報処理学会インターネットと運用技術研究会研究報告, No.2009-IOT-4, pp.113-118 (2009).

(平成 21 年 6 月 15 日受付)

(平成 21 年 12 月 17 日採録)



金 勇 (学生会員)

平成 21 年岡山大学大学院自然科学研究科電子情報システム工学専攻博士前期課程修了。現在、同大学院自然科学研究科産業創成工学専攻博士後期課程に在学中。主に管理の容易なマルチホームネットワーク構築技術に関する研究に従事。分散システム、ネットワークアーキテクチャ等に興味を持つ。



山井 成良 (正会員)

昭和 59 年大阪大学工学部電子工学科卒業。昭和 61 年同大学大学院博士前期課程修了。昭和 63 年同大学院基礎工学研究科 (物理系専攻情報工学分野) 博士後期課程退学。同年奈良工業高等専門学校情報工学科助手。同講師、大阪大学情報処理教育センター助手、同大学大型計算機センター講師、岡山大学総合情報処理センター (現、総合情報基盤センター) 助教授を経て、平成 18 年より同教授。分散システム、マルチメディアシステム、マルチメディアネットワークの研究に従事。IEEE、電子情報通信学会各会員。博士 (工学)。



岡山 聖彦 (正会員)

平成 2 年大阪大学基礎工学部情報工学科卒業。平成 4 年同大学大学院基礎工学研究科博士前期課程修了。同年同大学院基礎工学研究科博士後期課程を退学し、同大学工学部助手。平成 6 年奈良先端科学技術大学院大学情報科学研究科助手。平成 10 年岡山大学工学部助手。平成 17 年同大学総合情報基盤センター助手。平成 19 年同助教。博士 (工学)。インターネットアーキテクチャ、ネットワーク管理、ネットワークセキュリティの研究に従事。電子情報通信学会会員。



清家 巧

平成 19 年岡山大学工学部通信ネットワーク工学科卒業。平成 21 年同大学大学院自然科学研究科 (電子情報システム工学専攻) 博士前期課程修了。同年アドソル日進株式会社入社。迷惑メール対策、分散システム運用管理等に興味を持つ。



中村 素典 (正会員)

1994 年京都大学大学院工学研究科博士後期課程単位取得退学。立命館大学理工学部助手、京都大学経済学部助教授、京都大学学術情報メディアセンター助教授を経て、2007 年より国立情報学研究所教授、現在に至る。博士 (工学)。日本ソフトウェア科学会、電子情報通信学会各会員。コンピュータネットワーク、ネットワークコミュニケーション、認証連携等の研究に従事。