

マルチホーム環境における経路とMTAの状態を考慮したメール受信経路動的選択手法

實藤 翔[†] 金 勇[†] 清家 巧^{†,*1} 岡山 聖彦^{††} 山井 成良^{††}

[†]岡山大学 大学院自然科学研究科

^{††}岡山大学 総合情報基盤センター

E-mail: †{shou-j,yongkim,seike}@dist.cne.okayama-u.ac.jp, ††{okayama,yamai}@cc.okayama-u.ac.jp

概 要

電子メールシステムを安定的に運用するための方法として、マルチホームネットワークがよく用いられる。その運用方法として我々の研究グループでは、受信経路のネットワークの状態を考慮した、メール受信経路動的選択手法を提案した。しかし、この手法ではMTA(メールサーバ)の状態を考慮していないため、高負荷なMTAが選択される恐れがあった。そこで本論文では受信経路のネットワークの状態だけでなく、MTAの状態も考慮したメール受信経路動的選択手法を提案する。本手法では、その2つの状態に応じて適切な受信経路を選択することでMTAの負荷分散を実現し、またMTAに障害が発生した場合でも、そのMTAを経由する受信経路を自動的に回避することが可能となる。

A Dynamic Routing Method for Inbound E-mail Delivery Considering Route and MTA Conditions on Multihomed Environment

Shou JITSUTOU[†], Yong JIN[†], Takumi SEIKE^{†,*1}, Kiyohiko OKAYAMA^{††}, and Nariyoshi YAMAI^{††}

[†]Graduate School of Natural Science and Technology, Okayama University

^{††}Information Technology Center, Okayama University

E-mail: †{shou-j,yongkim,seike}@dist.cne.okayama-u.ac.jp, ††{okayama,yamai}@cc.okayama-u.ac.jp

Abstract

To operate e-mail system stably, multihomed networks are well used. As an operating method on such networks, we have so far proposed a dynamic route selection method considering status of networks. However, status of Mail Transfer Agents(MTAs) have not been considered in this method so that a heavily loaded MTA would be possibly selected. In this paper, we propose a dynamic route selection method considering not only status of networks but also that of MTAs. With this method, an appropriate route could be selected based on status of networks and MTAs so that load balancing of MTAs could be realized. Furthermore in case breakdown MTAs exist, such MTAs could be avoided automatically.

*1現在、アドソル日進株式会社
Presently with Ad-sol Nissin Corporation

1 はじめに

電子メールは WWW と同じくインターネットにおいて最も普及しているサービスの一つであり、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。特に、近年のインターネットの普及と広帯域化に伴い、より重要な情報あるいはサイズの大きなメッセージが頻繁に交換されるようになり、電子メールシステムを安定に稼働させることが重要である。このような安定性の観点から、複数のバックボーンを用いてインターネットに接続するマルチホームネットワークの利用が考えられる。これにより、1つのバックボーンに障害が発生しても、他のバックボーンを経由して電子メールを配送ができるため、サービスの耐障害性の向上が期待できる。しかし、単純に複数のバックボーンと接続しただけでは、送信者にとって最も適切な経路を利用できないという問題があった。

この問題に対して我々の研究グループでは、受信経路を決定する機能を持つ応答を多重化することによって、マルチホーム環境において受信経路のネットワークの状態を考慮したメール受信経路動的選択手法を提案した [1][2]。この手法では自組織宛メールの配送において、ネットワーク遅延時間が最も短い経路を選択することができる。しかし、この手法では MTA(メールサーバ)の状態について考慮していないため、MTA が高負荷であっても遅延時間の短いバックボーンに接続されていれば、この MTA が選択されることになるという問題がある。

そこで本論文では上記の問題を解決するため、受信経路のネットワークの状態だけではなく、MTA の状態も考慮したメール受信経路動的選択手法を提案する。本手法では MTA の負荷が大きい場合、その MTA を経由する受信経路を選ばれにくくすることで MTA の負荷分散を実現し、また MTA に障害が発生している場合でも、その MTA を経由する受信経路を自動的に回避することが可能となる。

2 従来のメール受信経路動的選択手法

本章では、我々の研究グループが提案しているメール受信経路動的選択手法(以下、従来手法という)と、その問題点について述べる。

2.1 従来のメール受信経路動的選択手法

従来手法では、組織外からの MX 問合せに対して組織内では各バックボーンに対してそれぞれの経路を優先経路とする MX 応答を返すように施している。この時、リゾルバは複数の MX 応答を受け取ることになるが、リゾルバは最も速く到着した MX 応答のみ受取する。しかし、一般的なゾーンサーバには 1 つの MX 問合せに対して複数の MX 応答を返すような機能はないので、そのような動作を行うシステムを構築した。従来手法のシステムの構成例を図 1 に示す。

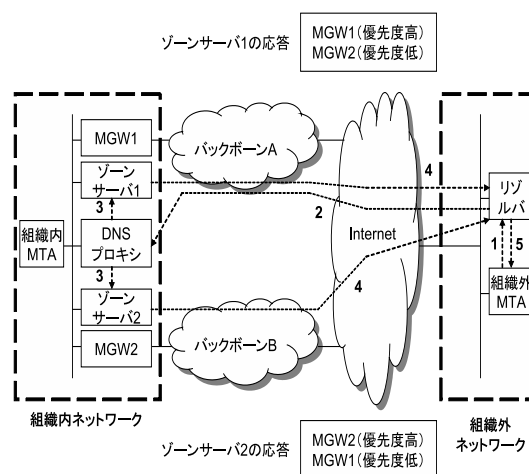


図 1: 従来手法のシステムの構成例

図 1 の構成について説明する。MX 問合せを複製する機能を持つ DNS プロキシを設置し、各バックボーンの接続点にそれぞれメールゲートウェイ (MGW1 及び MGW2) と、ゾーンサーバ (ゾーンサーバ 1 及びゾーンサーバ 2) を設置する。ここで各ゾーンサーバでは、同一のバックボーンに接続するメールゲートウェイ (以下、MGW と呼ぶ) の優先度が高くなるように、予め組織内ドメインの MX レコードを設定しておく。ここで、従来手法を用いた MX 問合せ及び MX 応答の手順について説明する。なお、以下の数字は図 1 中の数字と対応している。

1. 組織外 MTA はリゾルバに対して宛先メールアドレスのドメイン名に基づいた MX 問合せを行う
2. リゾルバは DNS プロキシに対して、MX 問合せを行う
3. DNS プロキシは MX 問合せを受け取ると、その MX 問合せを 2 つに複製し各バックボーンの接続点に置かれたゾーンサーバへ転送する

4. 各ゾーンサーバは MX 問合せを受け取ると、各バックボーン経由でリゾルバに MX 応答を行う
5. リゾルバは先に到着した応答のみを、組織外 MTA に返す

このように DNS プロキシと既存の DNS サーバプログラムを組み合わせることで、ネットワークの状態を考慮した電子メール受信経路の動的選択が可能となる。

2.2 従来のメール受信経路の動的選択手法の問題点

従来手法では、複数の MX 応答を用いることでネットワークの状態を考慮しているが、複数の MGW の状態については考慮していない。つまり MGW に障害が発生していたり、高負荷であったりする場合でも、遅延時間の短いバックボーンに接続されていれば、その MGW を経由する受信経路が選択されてしまう、という問題がある。また、組織外からの MX 問合せは組織内に存在する単一の DNS プロキシで受け付けているため、その DNS プロキシに障害が発生した場合、MX 問合せ及び MX 応答ができなくなってしまう問題もある。

3 経路と MTA の状態を考慮したメール受信経路動的選択手法

本章では、MTA の状態を考慮し、耐障害性も向上した経路動的選択手法を提案する。

3.1 提案手法の概要

2.2 節で述べたとおり、従来手法では 1 つの MGW に障害が発生した場合、障害の発生している MGW を経由する経路が選択されてしまう可能性がある。そこで提案手法では、ゾーンサーバと MGW を同一のホストに設置するようにする。これにより MGW に障害が発生した場合、同一のホストに存在するゾーンサーバが MX 応答を返さなくなるため、障害の発生している MGW を経由する受信経路が決定されてしまう問題を回避することが可能となる。また、従来手法ではバックボーンの遅延時間に基づいて受信経路を選択していたが、それに加えて MX 応答に MGW の状態を考慮した遅延を付加するこ

とで、より適切な経路選択が可能となる。そのために MGW の負荷の計測を行い、その負荷量に応じて MX 応答に遅延を付加する機能を持つバランサをバックボーンごとに設置する。

さらに提案手法では、バックボーンごとに DNS プロキシも設置する。もし DNS プロキシに障害が発生したような場合でも、他の DNS プロキシを用いて MX 問合せ及び MX 応答ができるため、耐障害性も向上することが期待できる。

3.2 経路と MTA の状態を考慮した経路選択手法

リゾルバが複数の MX 応答を受信する際、各応答は各バックボーンの影響だけでなく、MGW の負荷量に応じた遅延の影響も受けることによって MX 応答の到着時間に差が生じる。そして最も応答の速いバックボーンから到着した MX 応答が有効となる。このようにして MGW の状態とネットワークの状態を考慮した経路選択が可能になる。このことを踏まえたシステム構成を図 2 に示す。

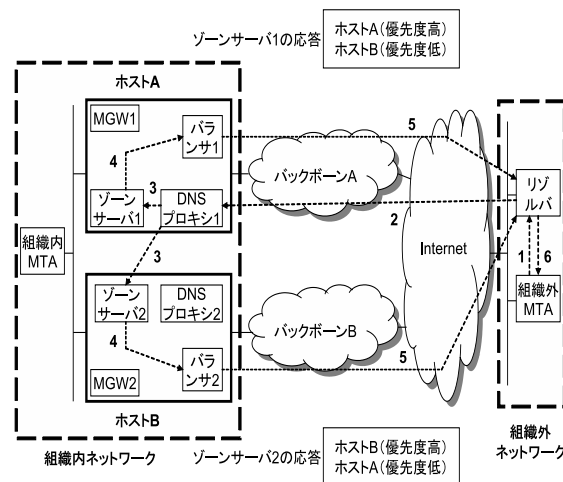


図 2: 提案手法のシステム構成

図 2 の構成について説明する。ゾーンサーバ、MGW、DNS プロキシ、バランサを持つホストを、各バックボーンの接続点にホスト A、ホスト B として設置する。また各 MGW の優先度については従来手法と同じく、各ゾーンサーバと同一ホストに存在する MGW の優先度が高くなるように、予め組織内ドメインの MX レコードを設定しておく。以下に、図 2 を用いて提案システムにおける組織外 MTA から組織内 MTA へメールを配送する場合の MX 問合せ及

びMX 応答の手順について説明する．なお，以下の数字は図 2 中の数字と対応している．

1. 組織外 MTA はリゾルバに対して宛先メールアドレスのドメイン名に基づいた MX 問合せを行う
2. リゾルバはいずれかの DNS プロキシに，MX 問合せを行う
3. DNS プロキシは MX 問合せを複製し，各バックボーン接続点に置かれたゾーンサーバへ転送する
4. 各ゾーンサーバは MX 問合せを受け取ると，同一のホストに存在するバランサに MX 応答を行う
5. バランサで負荷量に応じた遅延を加え，各バックボーン経由でリゾルバに MX 応答を行う
6. リゾルバは先に到着した応答のみを，組織外 MTA に返す

このように複数の DNS プロキシを組み合わせ，さらに MX 応答に MGW のシステム負荷量に応じた遅延を加えるシステムを従来のシステムに組み合わせることで，提案手法の動作が実現可能となる．

4 提案手法の実装と動作確認

本章では，提案手法に基づいたシステムの実装と，システムを用いた動作確認実験と結果について述べる．

4.1 実装

実装したシステムの構成を図 3 に示す．

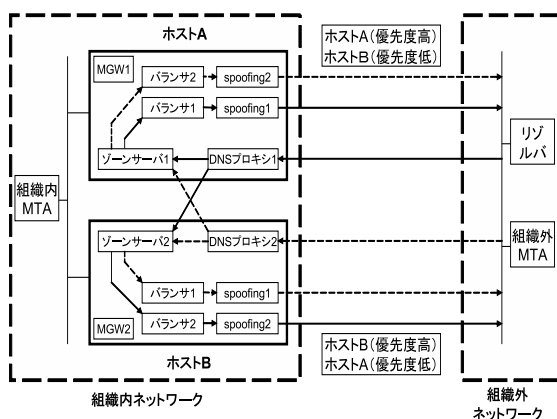


図 3: 実装したシステムの構成

図 3 の構成について説明する．MGW，ゾーンサーバ，DNS プロキシ，バランサ，spoofing を持つホストを各バックボーン接続点にホスト A，ホスト B として組織内ネットワークに設置した．リゾルバは，宛先アドレスに基づいていずれかの DNS プロキシに MX 問合せを行うように設定し，DNS プロキシは MX 問合せをゾーンサーバ 1 とゾーンサーバ 2 に転送するように設定した．そして，ゾーンサーバ 1 は MX レコードにホスト A とホスト B を設定し，ホスト A の優先度を高く設定した．ゾーンサーバ 2 も同様に MX レコードにホスト A とホスト B を設定したが，こちらはホスト B の優先度を高く設定した．

spoofing とは，ゾーンサーバの MX 応答を DNS プロキシの応答として偽装するプログラムのことである．これは MX 応答の送信元 IP アドレスと送信元ポート番号を，MX 問合せに対応した送信元 IP アドレスと送信元ポート番号に偽装する機能を持つ．MX 問合せに対して他のホストで応答した場合，または MX 問合せの宛先ポート番号と，その問合せに対する MX 応答の送信元ポート番号が異なる場合，MX 応答はリゾルバが受け取った時点で破棄されてしまう [3]．そのため，このプログラムを利用して MX 応答を偽装する必要がある．

しかし，DNS プロキシからゾーンサーバへ MX 問合せを転送するとき，その MX 問合せがどの DNS プロキシから転送されたか区別をつけないと，MX 応答の送信元 IP アドレスの偽装処理ができなくなる．その点については宛先ポート番号を変更することで区別を可能にした．また MX 応答をバランサと spoofing に渡すために，FreeBSD の ipfw[4] と divert[5] という機能を利用した．そしてバランサは Perl モジュールの，BSD::Sysctl[6] と Time::HiRes[7] を利用した．

図 3 中に書かれている実線の矢印は，DNS プロキシ 1 に MX 問合せを行ったときの MX 問合せ及び MX 応答の流れを示しており，点線の矢印は DNS プロキシ 2 に MX 問合せを行ったときの MX 問合せ及び MX 応答の流れを示している．

4.2 動作確認実験

本節では，作成した提案システムを設置した実験ネットワークにおいて行った動作確認実験と，それに対する結果について述べる．また，実験環境は図 3 とほぼ同様であるが，MGW の性能による受信メー

ルの偏りを調査するため、ホストの性能を異なるものとした。2つのホストの性能を比較したものを表1に示す。

表 1: ホストの性能

	ホスト A	ホスト B
CPU	Intel Celeron (564.84 MHz)	Pentium 4 3.00GHz
メモリ	128MB	1024MB
OS	FreeBSD 6.4-RELEASE	FreeBSD 6.4-RELEASE
MTA	Sendmail8.14.3[8]	Sendmail8.14.3
NIC	100Mbps/s <full-duplex>	100Mbps/s <full-duplex>
ゾーンサーバ	BIND9.5.0-P1[9]	BIND9.5.0-P1

表1から、ホストAよりもホストBの方が処理能力が高いことが分かる。なお、組織外MTAにはSendmail8.14.3を利用した。また、この動作確認実験において大量のメールの生成するために、postfix 付属のソフトウェアであるsmtp-source[10]を利用した。またロードアベレージに対してMX 応答を遅延させる時間について表2に示す。

表 2: ロードアベレージと MX 応答遅延時間

ロードアベレージ	MX 応答遅延時間 (ms)
0.50 未満	0
0.50 以上 1.00 未満	5
1.00 以上 2.00 未満	10
2.00 以上 3.00 未満	15
3.00 以上 4.00 未満	20
4.00 以上	25

上述した実験環境において、MGW にかかる負荷による受信メールの偏りを調査するため、以下に示す3パターンの実験を行った。

実験 1

実験1では、双方のホストに急激な負荷がかからないように一定間隔で電子メールを受信した場合、電子メールがどちらホストの受信サーバに配送されるか確認した。

この実験では、ホストA及びホストBに対して電子メールを送信する組織外ホスト1を用意した。そしてその組織外ホスト1から2秒間隔で1通ずつ合計1000通の電子メールを配

送し、どのホストを経由して受信するか確認した。

実験 2

実験2では、双方のホストにある程度負荷がかかるように電子メールを連続的に受信した場合、電子メールがどれくらい分散するか確認した。なお、実験1と同様に組織外ホスト1を用意したが、今度は組織外ホスト1から1000通の電子メールを連続的に配送し、どのホストを経由して受信するか確認した。

実験 3

実験3では、ホストBが高負荷である場合に電子メールを受信した場合、電子メールがどれくらい分散するか確認した。

この実験では、ホストBに対して多数の電子メールを、5つのSMTPセッションを並列に走らせて集中的に配送する組織外ホスト2を新たに用意した。それによってホストBに対してメール受信の負荷を与えつつ、組織外ホスト1から1000通の電子メールを連続的に配送し、どのホストを経由して受信するか確認した。

以上3つの動作確認実験を行った結果を表3に示す。

表 3: 動作確認実験の結果

条件	ホスト A	ホスト B
実験 1(2 秒間隔で配送)	11 通	989 通
実験 2(連続的に配送)	139 通	861 通
実験 3(ホスト B に負荷を与えつつ連続的に配送)	638 通	362 通

表3から、それぞれのホストに急激な負荷がかかっていない場合、処理能力の高いホストを経由する受信経路経路が優先的に決定されるが、その処理能力の高いホストが高負荷になるにつれて、受信経路が分散されていることが分かる。つまり、MTAの状態を考慮したメール受信経路選択について提案手法は正しく動作していると言える。

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電子メールシステムを安定に運用するために、以前

我々が提案したメール受信経路動的選択手法の問題点を考察し、これを解決するためにMTAの状態を考慮したメール受信経路の動的選択手法を提案した。そして、提案手法を実装して動作確認実験を行い、高負荷であるMTAを経由する受信経路が選択されにくくなることを確認した。

今後の課題として、実運用を通じた性能評価の検討が挙げられる。また本論文ではMTAの状態としてロードアベレージのみを参照していたので、CPUの処理能力などの状態を考慮した検討と、遅延時間の再検討も行う必要がある。

参考文献

- [1] 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,” FIT2007, LL-004, 2007 .
- [2] 金勇, 清家巧, 岡山聖彦, 中村素典, 山井成良: “ALG を用いたマルチホーム環境における自組織宛メール配送の動的経路選択手法,” IPSJ Symposium Series Vol.2008, No.13, Dec. 2008.
- [3] R. Elz, R. Bush: “Clarifications to the DNS Specification,” RFC 2181, July 1997 .
- [4] Ugen J. S. Antsilevich, Poul-Henning Kamp, Alex Nash, Archie Cobbs, Luigi Rizzo: “ipfw - IP firewall and traffic shaper control program,” FreeBSD System Manager’s Manual, January 16 2006 .
- [5] A. Cobbs: “divert - kernel packet diversion mechanism,” FreeBSD Kernel Interface Manual, December 17 2004 .
- [6] David Landgren.: “BSD::Sysctl - Manipulate kernel sysctl variables on BSD-like systems,” User Contributed Perl Documentation, 2007-01-17.
- [7] D. Wegscheid, R. Schertler, J. Hietaniemi, G. Aas: “Time::HiRes - High resolution alarm, sleep, gettimeofday, interval timers,” Perl Programmers Reference Guide, 2001-09-22.
- [8] Sendmail Inc.:
“Sendmail Home,”
<http://www.sendmail.org>.
- [9] Internet Systems Consortium, Inc.:
“ISC BIND,”
<http://www.isc.org/index.pl?/sw/bind> .
- [10] “The Postfix Home Page,”
<http://www.postfix.org> .