JAIST Repository
https://dspace.jaist.ac.jp/

Title	多重ルーティング型マルチホームアーキテクチャの提案
Author(s)	宇多，仁 ；小柏，伸夫；永見，健一；近藤，邦昭；中川，郁夫；篠田，陽一；江崎，浩
Citation	電子情報通信学会論文誌 B，J87-B(10)：1564-1573
Issue Date	2004-10-01
Type	Journal Article
Text version	publisher
URL	http://hdl.handle.net/10119/4668
Rights	Copyright (C)2004 IEICE. 宇多 仁，小柏 伸夫，永見 健一，近藤 邦昭，中川 郁夫，篠田 陽一，江崎 浩，電子情報通信学会論文誌 B，J87-B(10)，2004，1564-1573. http://www.ieice.org/jpn/trans_online/
Description	

JAIST
JAPAN
ADVANCED INSTITUTE OF
SCIENCE AND TECHNOLOGY

Japan Advanced Institute of Science and Technology

多重ルーチング型マルチホームアーキテクチャの提案

宇多　仁[†*]　小柏　伸夫[†**]　永見　健一[††]　近藤　邦昭[††]
中川　郁夫[††]　篠田　陽一[†††]　江崎　浩[††††]

A New Multi-Homing Architecture Based on Overlay Network

Satoshi UDA[†*], Nobuo OGASHIWA[†**], Kenichi NAGAMI[††], Kuniaki KONDO[††], Ikuo NAKAGAWA[††], Yoichi SHINODA[†††], and Hiroshi ESAKI[††††]

あらまし　インターネットの発展に伴い，重要度の高いネットワークへの到達性の向上を目指したマルチホーム接続が広く行われるようになった．しかし，現在用いられているマルチホーム方式は，インターネット経路制御に対するインパクトが大きく，更なるマルチホーム利用者の増加に耐えられない．また，障害時の到達性向上は実現できるが，流入トラヒックの制御が困難であるために，正常時にマルチホームによる複数の接続を有効に使い分けることが困難である．我々は，本論文で，これらの問題点を解決する新たなマルチホームアーキテクチャとして多重ルーチングを用いたマルチホームアーキテクチャを提案する．本方式により，マルチホームの主目的である障害時の到達性の向上を実現しつつ，利用者数に対する対規模性の向上と，流入トラヒックの柔軟な制御を可能とする．

キーワード　インターネット，マルチホーム，多重ルーチング

1. まえがき

インターネットは急激に拡大を遂げ，現在では，情報通信における社会基盤の一つとなっている．このように成長したインターネットにおいては，接続されるネットワークの重要度も様々である．重要度の高いネットワークにおいては，接続回線の障害やふくそうなどによるサーバへの到達性の喪失などの問題を回避するために，複数の接続事業者（ISP）等と複数の回線で接続するマルチホーム接続 [1], [2] が広く利用されている．

現在広く用いられているマルチホーム接続は，利用者網に対する経路到達性を経路制御システムの作用で冗長化し実現するものが一般的である．この方式では，特に，利用者網あてトラヒックの制御が難しいために，複数の接続をより柔軟かつ効率的に利用したいといった利用者の要求にこたえることが困難である．更には，インターネット上の経路情報数の増加といった大きな問題を抱えており，更に多くの利用者がマルチホーム接続を利用することは困難である．

本論文では，既存マルチホーム技術の問題点をまとめた上で，指摘した問題点を解決する新たなマルチホームアーキテクチャとして多重ルーチング技術を用いたマルチホーム接続を提案する．そのアーキテクチャに基づいてプロトタイプ実装及び運用実験を行った経験をまとめる．

2. マルチホーム技術の現状と問題点

現在のマルチホーム技術は，利用者網を複数の ISP へ接続し，経路制御システムの作用により利用者網に対する経路到達性を向上させるものが一般的である．

† 北陸先端科学技術大学院大学情報科学研究科，石川県
School of Information Science, Japan Advanced Institute of Science and Technology (JAIST), 1–1 Asahidai, Tatsunokuchi, Ishikawa-ken, 923–1292 Japan
†† (株) インテック・ネットコア，東京都
Intec NetCore Inc., 1–3–3 Shinsuna, Koto-ku, Tokyo, 136–0075 Japan
††† 北陸先端科学技術大学院大学情報科学センター，石川県
Center for Information Science, Japan Advanced Institute of Science and Technology (JAIST), 1–1 Asahidai, Tatsunokuchi, Ishikawa-ken, 923–1292 Japan
†††† 東京大学情報理工学系研究科，東京都
Graduate School of Information Science and Technology, The University of Tokyo, 7–3–1 Hongo, Bunkyo-ku, Tokyo, 113–8656 Japan
* 現在，北陸先端科学技術大学院大学情報科学センター
** 現在，(株) インテック・ネットコア

マルチホーム接続利用者網は，そのアドレス空間を経路制御プロトコルを用いインターネットへ広告する．該当利用者網あてのパケットは，この経路情報に基づき，いずれかの接続 ISP を介し該当利用者網へ到達する．

インターネットの適用領域が拡大し，情報化社会の社会基盤となりつつあることから，エンドユーザ環境においてもインターネット接続の頑健性が重要となりつつある．しかし，既存マルチホーム技術は，次のような課題を抱えており，そのままの形でエンドユーザ環境を含む広範囲に展開することは困難である．

- インターネット上の経路情報の増加
- 複数接続の柔軟な制御と有効利用の難しさ
- マルチホーム利用者網の運用の難しさ

2.1 インターネット上の経路エントリ数

2003 年末にはインターネット上の経路数は 12 万経路を超え，今なお増加を続けている．経路数の増加は，バックボーンルータのメモリ空間を圧迫するため，大きな問題となっている．このような経路数の増加を抑制するため，CIDR（Classless Inter-Domain Routing）を用いたプロバイダ集約が広く用いられている [3], [4]．マルチホーム接続では，複数の接続を介しての利用者網への経路到達性を確保するため，利用者網のアドレス空間を利用者網が接続するすべての ISP から広告する必要がある．つまり，多くの場合，各 ISP は，マルチホーム接続する利用者のアドレス空間を経路集約することなくインターネットへ広告する必要があり（図 1），マルチホーム接続利用者ごとにインターネット上の経路数が 1 つずつ増加することになる(注1)．

現在では，インターネット上の経路数の半数以上である約 6 万経路を，プレフィックス長 24 ビット以上の小さなアドレス空間の経路が占めるに至っている（表 1）．これは，マルチホームをはじめとする小さなアドレス空間の利用者の経路がインターネットの経路の半数以上を占めていることを示す．

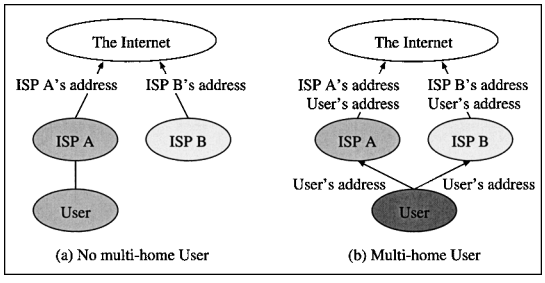


図 1　既存マルチホーム接続における経路広告
Fig. 1　Route advertisements on traditional multi-homing.

2.2 複数接続の柔軟な制御と有効利用

既存マルチホーム技術の大きな問題として，定常時に複数の回線を効率的に運用することが困難なことが挙げられる．マルチホーム接続の主目的は，障害時の到達性の確保であるが，定常時に保持する複数の回線を効率的に利用したいという要求は強い．

既存マルチホーム技術では，トラヒックの回線選択は，経路制御システムの作用によりなされる．このため，経路制御の枠内での制御に限られ，複数のリンクの使い分けは，パケットの始点とあて先の情報のみに基づいたものに限られる．例えば，マルチホーム接続において各接続 ISP の特性に違いがある場合（一方が広帯域・低品質，もう一方が狭帯域・高品質など），利用アプリケーションごとにこれらのリンクを使い分けたいといった要求が生じるが，既存マルチホーム技術では，このような要求にこたえることは困難である．

このような問題に対し，利用者網からの流出トラヒックを制御し効率的な回線選択を行う取組みがなされており，RouteScience PathControl(注2)，BIG-IP Link Controller(注3)等が製品化されている．これらは，受

表 1　インターネット上の経路数（プレフィックス長ごと）
Table 1　A number of route entries on the Internet (by each prefix length).

Prefix 長	エントリ数	Prefix 長	エントリ数
1	0	17	1,666
2	0	18	2,991
3	0	19	8,446
4	0	20	8,458
5	0	21	6,011
6	0	22	8,997
7	0	23	8,498
8	21	24	59,378
9	5	25	78
10	8	26	87
11	13	27	13
12	53	28	19
13	98	29	31
14	256	30	1
15	469	31	0
16	7,400	32	20

(注1)：利用者網の接続先 ISP が単一 ISP であれば該当 ISP における経路集約が可能であるが，冗長性確保の点で不安が残るため，このような形態は極めてまれである．
(注2)：RouteScience 社：http://www.rootscience.com/
(注3)：F5 Networks 社：http://www.f5networks.com/

信者との間の経路特性を計測しつつ，動的に最適な経路を選択し，パケット送出リンクを制御するものである．しかし，これらは，CDN（Contents Distribution Network）事業者等を主な対象とする製品であり，利用者網からの流出トラヒックの制御を目的としており，流入トラヒックの制御においては，既存のマルチホーム方式と同等である．エンドユーザ環境での利用を意識すると，流入トラヒックの制御が重要となるが，これらの手法では流入トラヒックの制御の実現は困難である．

2.3 利用者側ルータの運用

既存マルチホーム技術においては，マルチホーム利用者網と ISP との間の経路制御には AS（Autonomous System）間経路制御プロトコルである BGP-4（Border Gateway Protocol version 4）[5] が用いられる．BGP-4 による経路制御は，インターネットのバックボーン上で用いられる高機能で複雑な経路制御プロトコルであり，この運用には高い技術を要する．

利用者網側ルータにおいて，複雑な経路制御プロトコルである BGP-4 を運用する必要があることはマルチホームの広範囲への適用において高い障壁となる．更に，複数の接続のより有効な利用には，この BGP-4 による経路交換において繊細なパラメータ調整が必要となる．エンドユーザ環境へのマルチホーム接続の導入を意識すると，より簡易な運用が可能な方式が求められる．

3. 多重ルーチング型マルチホーム

前章で述べた既存マルチホーム技術の問題点を解決するために，我々は経路制御技術だけに頼らない，多重ルーチング技術に基づいたマルチホームアーキテクチャを提案する．

3.1 提案アーキテクチャ概要

既存のマルチホーム方式において，流入トラヒックの制御が困難な点に着目すると，その原因として，利用者網へパケットを送出する側のルータが利用者の制御可能範疇にないことが挙げられる．そこで，提案方式では，トラヒック分散制御装置（DR: Distribution Router）と呼ぶ装置をインターネットバックボーン（任意の ISP）上に導入し，利用者網あてのトラヒックを DR を介して転送することとした．DR では，利用者網あてパケットを解析し利用者のポリシに応じたパケット配送経路選択等を実現する（図 2）．

利用者網と DR の間は，マルチホーム接続ごとにトンネル等の仮想接続を用いて接続する．また，利用者網側では利用者側終端装置（UR: User Router）と呼ぶ装置を導入し，仮想接続を終端する．例えば，二つの ISP（ISP-A，ISP-B）を用いてマルチホーム接続する利用者網は，2 種類（ISP-A 経由及び ISP-B 経由）の仮想接続で DR と結ばれることとなる．UR においてこの仮想接続を終端する端点アドレスには，各接続 ISP から割り当てられたアドレスを用いる．つまり，ISP-A 側のリンクに割り当てられたアドレス，及び，ISP-B 側のリンクに割り当てられたアドレスである．このように ISP-A 及び ISP-B から見ると利用者網は，それぞれの ISP のアドレスのみを用いた通信を行っているように見え，また，インターネット上のノードから見て，利用者網は DR の先に接続されているように見える．つまり，DR の運用事業者は仮想接続により利用者にインターネット接続を提供する ISP となる．

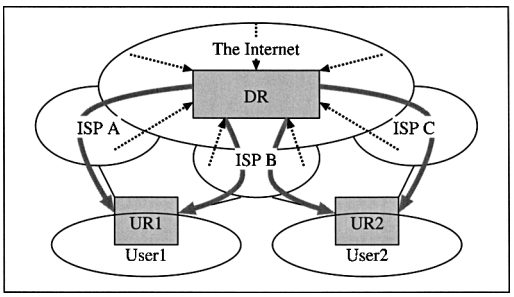


図 2 提案マルチホームアーキテクチャの概要
Fig. 2 Overview of our proposed multi-homing architecture.

利用者網あてトラヒックは，DR においてパケット解析とポリシに応じたパケット配送経路選択が行われ，利用者あて仮想接続のいずれかを用いて UR へと転送される．DR–UR 間の仮想接続は，利用者網の物理的な接続に対応しており，この DR の選択機能により，利用者網あて各パケットの通るべき物理回線等を選択することが可能である．この経路選択機構においては，一般的な経路選択手法（つまり，あて先アドレスのみに基づいた経路選択）を超えて，より詳細なパケットの特性まで解析した経路選択が可能である．例えば，リアルタイムアプリケーションの通信に対しては低ジッタの経路選択を行うようなことも可能となる．また，各物理リンクのトラヒック状況を計測し，その結果を回線選択ポリシにフィードバックすることで，複数のリンクのトラヒックを同等にし回線の利用率を

高めるような制御も可能である．

もちろん，マルチホーム接続において一部接続の障害時にも利用者網への到達性を維持することは最も重要なことである．本アーキテクチャにおいては，経路選択機構と利用者網を接続する仮想接続の状態を，生存確認パケットの利用などの手法で常時監視する．経路選択機構では，障害による仮想接続の切断を検出し，該当仮想接続を利用しない経路選択へと動的に経路選択ルールを変更する．これにより，障害のある経路をパケット転送に用いなくなり，利用者網への接続性は維持される．

3.2 利用者網経路情報の広告

提案方式においては，利用者網が利用するアドレス空間は，DR を運用する事業者に対して割り当てられた CIDR ブロックから割り当てる．利用者網の経路情報は，複数の利用者網の経路情報を経路集約した上で広告する．

これは，提案方式では，各利用者網が物理的にどのような ISP 等を介して接続されているかにかかわらず，DR の先に同等に接続されているかのように見えるため，DR における経路集約が可能となるのである．これにより，既存マルチホーム技術の問題点の一つとして挙げた，経路数の増加による経路制御システムへの影響を大幅に改善することが可能である．

また，利用者網の物理接続先に変更が生じた場合においても，本方式では，仮想接続の利用者網側端点が変更となるのみで，アドレス・リナンバリングなどの必要はない．これも，仮想接続の利用者網側端点の変更後も，利用者網は変更前と同等に DR の先に接続されているかのように見えるためである．

3.3 DR の複数化

上述のアーキテクチャにより，利用者網への流入トラヒックの制御は可能となる．また，マルチホーム接続の本来の目的である，一部接続の障害時における利用者網への到達性の維持も可能である．しかし，利用者網あてトラヒックがすべて DR を介して配送されるアーキテクチャをとっているため，DR が単一点障害（a single point of failure）の要因となる．

この回避のために，複数の DR をインターネット上に分散配置することとする．全 DR は利用者網の UR との間に仮想接続を確立し，すべての DR から隣接 ISP 等に対し利用者網の経路情報を広告する．これにより，利用者網あてのトラヒックは，インターネット上の経路制御システムの作用により，発信点から最も近い DR を経由し，仮想接続を通して各利用者網に設置された UR へと配送されることとなる．つまり，分散配置した DR は通常時は全 DR が利用者網あてのパケット配送を行うこととなる（図 3）．このように，全 DR が同等なパケットの中継処理等を行うために，DR 間で利用者網などに関する情報を常に共有する．

DR の障害時，あるいは，DR の保守作業時には，DR からの利用者網経路情報広告を停止することで，該当 DR による利用者網あてパケットの中継処理を安全に停止することが可能である．該当 DR からの経路広告が停止すると，それまで該当 DR で処理されていたトラヒックは，経路制御システムの作用により自動的に次候補の DR において処理されることとなる．

また，この構成により，DR の障害に対する耐故障性の向上のみならず，パケット配送遅延の低減と並列処理によるトラヒック集中の回避といった利点も挙げられる．先に述べたとおり，提案アーキテクチャでは，利用者網あてトラヒックは常に DR を介して配送されるが，この影響で利用者網に対して直接配送を行った場合に対しての配送遅延が生じる．ここで，複数の DR がインターネット上に分散配置されることで，この配送遅延を小さなものとできる．また，単一 DR 構成では利用者網あてのトラヒックがすべて単一の DR を介して配送されるために，DR の負荷増大が懸念されるが，DR の複数化により DR でのパケット転送処理を分散化できる．

3.4 利用者網発トラヒックの取扱い

提案アーキテクチャでは，DR から UR の間に仮想接続を確立し，利用者網あてパケットは DR を介し仮想接続により利用者網へ配送される．本節では，逆方向，つまり，利用者網発トラヒックの取扱いに関して述べる．

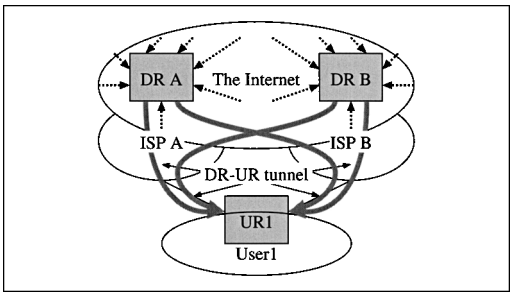


図 3　複数の DR を用いた構成
Fig. 3　Using multiple DRs for increasing reliability.

利用者網発のトラヒックに対しては，そのインターネットへの送出手法として，仮想接続を用いDRを介して送出する手法と，仮想接続を用いず隣接ISPを介して直接送出する手法が挙げられる．

ところが，ISPにおいては，始点IPアドレスを詐称した攻撃などを予防する目的で，自らが割り当てたアドレス以外を始点とするパケットを拒絶する運用が広く行われている．このようなことから，直接隣接ISPへと送出する手法は実運用上現実的ではない．

そこで，本提案アーキテクチャにおいては，利用者網発のトラヒックも仮想接続を用いDRを介して配送する．具体的には，URにおいてパケットには仮想接続のためのカプセル化が施され，Anycast技術を用いることで，利用者網から最近傍のDRへと導き，該当DRを介してインターネットへ配送する．

4. 評　　価

4.1 既存方式との比較

本節では，我々の提案する多重ルーチング型マルチホームアーキテクチャの評価として既存方式に基づくマルチホーム方式との比較を行う．ここでは，マルチホームを行う主目的である障害時の利用者網（保護対象ネットワーク）への到達性向上とともに，**2.** で述べた既存マルチホームの問題点として挙げられる以下の項目に着目し比較した．

- マルチホーム利用者の増加がバックボーン上の経路数に与える影響
- 複数のリンクの有効な使い分けの容易さ
- マルチホーム接続ネットワークの運用の難しさ

まず，マルチホーム接続の主目的である，障害時の対象ネットワークへの到達性の向上効果に関して述べる．既存マルチホームアーキテクチャでは，障害時に障害個所でのBGP-4による経路交換が停止することにより障害個所を経由する利用者ネットワークあての経路情報が消失し，他のネットワークを経由するものに変わることで実現されている．ところが，そもそもBGP-4は頻繁な経路変更を想定したプロトコルではないために，障害を検出した後，経路が安定し到達性が回復するまでにはネットワークの構成にもよるが10秒から数分程度の時間を要する．我々の提案方式では，DR–UR間の生存確認により障害を検出し，該当トンネルを向いたDRの経路表エントリを非障害トンネル側に変更するのみで経路切換が完了する．このため，生存確認間隔にも依存するが，1〜数秒程度の時間で経路切換が完了し到達性が回復する．

次に，マルチホーム利用者数の増加がインターネット上の経路数に与える影響について述べる．既存方式では，先にも述べたとおり，通常，利用者一つの接続ごとにインターネット上の経路エントリが一つ増加する．提案方式では，インターネットバックボーンへの経路広告は経路選択機構が行い，この広告時に経路集約が可能である．これにより，利用者数の増加はバックボーン上の経路数にほとんど影響を与えない．

また，マルチホーム接続の複数の接続を有効に活用するという点では，BGP-4の広告コスト調整のみでしか流入トラヒックを制御できない既存方式は，複数のリンクにトラヒックを均等に割り振るといった制御や，アプリケーションの要求に応じた品質のリンクを使用する制御などは不可能である．提案方式では，経路選択機構による柔軟な経路選択（ポリシルーチング）により利用者の要求に柔軟に応じられ，UR側トラヒック監視機能との連携により先に挙げたようなトラヒックの割振りなどが実現可能である．更に，この応用により，パケット転送の拒否，あるいは，一定帯域幅に制限するといった処理も可能である．これは，利用者網あてのDoS（Deny of Service）攻撃などに対して有効である．

更に，マルチホーム利用者側の運用コストに関しては，既存方式に高度な経路制御知識が必要であったのに比べ，提案方式では，経路制御などの複雑な運用はDRの運用者（マルチホームサービスの提供者）が行うこととなる．利用者側は，URの設置・運用を行うことになるが，URに実装されるユーザインタフェース次第で利用者の作業は簡素なものとできる．特に，ポリシに基づくトラヒック制御などといった高度な機能を必要としないエンドユーザにおいては，提案方式では「DRのアドレス」及び「利用者識別情報」のみの設定でマルチホーム接続の恩恵を得られる．

ここまで述べた項目を含め，表2にまとめた．これから，我々の提案手法は，既存手法の問題点を解決し利用者の要求にこたえつつ基本的機能である障害時の到達性回復時間の短縮も実現していることがいえる．

ここまでは，主にマルチホーム利用者の要求に基づいて比較を行ったが，インターネット全体の視点から網に及ぼす影響に関しても考察する必要がある．特に，インターネット全体に大きな影響を及ぼすものとして，下記の2点に関して比較する．

- トラヒック増加によるインターネットへの負荷

の増大

- 制御情報の増加によるインターネットへの影響

まず，トラヒック増加によるインターネットへの負荷増加に着目すると，提案方式ではマルチホーム利用者網あてのパケットがすべて DR を経由して配送されることから，常に最短の経路によりトラヒックが配送されるとは限らず，これにより全体として見た負荷の増大につながる．しかし，DR の配置場所を検討することで，その影響を抑えることは可能である．つまり，多くのトラヒックの集まるインターネットエクスチェンジ近傍に DR を配置することで大きな効果が期待できる．

次に，制御情報の増加によるインターネットへの影響に着目すると，ここまでも述べてきたとおり，既存方式ではマルチホーム利用者数の増加に伴うインターネット全体の経路情報の増加が深刻な問題となっている．経路数の増大は，インターネットの広い範囲のルータに大きな負荷を強いる．このことから，既存方式により，更に多くのマルチホーム利用者を収容することは，大きな危険性をはらんでいる．

4.2 経路集約の効果

前節でも述べたとおり，提案方式では，経路広告時に経路集約が可能である．本節では，現状のインターネット上の経路数にマルチホーム接続が与えている影響をまとめ，提案方式の経路集約効果について既存マルチホーム方式と比較しその効果を評価する．

2.1 で示したように，現在では，インターネット上の経路数の半数以上である約 6 万経路を，プレフィックス長 24 ビット以上の小さなアドレス空間の経路が占めるに至っている．その一方で，RIR（Regional Internet Registries）から ISP 等へのアドレス空間の割当は，プロバイダ集約による経路集約効果を期待し，一般に 20 ビット長前後で行われている．この結果から，ISP が RIR から割り当てられたブロックについて，ブロック全体の経路の広告に加えて，該当ブロック中のより細かな経路を広告するパンチングホールが多数行われていることが予想できる．

パンチングホールによる経路広告の多くの用途はマルチホームの実現である．マルチホーム接続では，各 ISP は利用者網のアドレス空間を集約することなくインターネットへ広告する必要があるためである．

表 3 は，より詳細な解析である．この表は，AS 17932 において受信した全経路情報から CIDR に従った割当が行われているアドレス空間を抽出し，RIR が割当を行う際の最小割当単位と，実際にインターネット上に流通している経路情報を比較している．この表から，最小割当単位が 20 ビット長程度のブロックでは，最小割当単位より細かな経路が多数存在しパンチングホールが多数生じていることが分かる．また，最小割当単位が 24 ビット長以上の空間では，最小割当単位

表 2　既存方式と提案方式の比較
Table 2　Comparison the proposed architecture with the traditional method.

	既存方式	提案方式
障害時に対する頑健性	良	良
トラヒック分割	困難	良
ポリシに応じた制御	困難	良
構造	単純	やや複雑
セッション維持性	良	良
パケットフィルタリング	困難	良
利用者数増大に対する耐性	低	高
ルータへの影響	大	小
ホストへの影響	無	無
運用・管理	難解	容易
ISP との協調	必要	不要

表 3　インターネット上の経路数（CIDR ブロックごと）
Table 3　Routing entries in the Internet in each CIDR block.

ブロック	RIR	最小割当	経路数	過小経路広告
24.0.0.0/8	ARIN	20	0	0
60.0.0.0/7	APNIC	20	1,169	442 (37%)
62.0.0.0/8	RIPE	19	1,248	611 (48%)
63.0.0.0/8	ARIN	20	2,835	2,516 (88%)
64.0.0.0/6	ARIN	20	13,703	10,703 (78%)
68.0.0.0/7	ARIN	20	2,990	1,831 (61%)
80.0.0.0/7	RIPE	20	1,656	715 (43%)
82.0.0.0/8	RIPE	20	31	5 (16%)
193.0.0.0/8	RIPE	29	3,974	0 (0%)
194.0.0.0/7	RIPE	29	5,396	0 (0%)
196.0.0.0/8	ARIN	24	694	0 (0%)
198.0.0.0/7	ARIN	24	8,146	7 (0%)
200.0.0.0/8	LACNIC	24	4,747	1 (0%)
201.0.0.0/8	LACNIC	20	0	0
202.0.0.0/7	APNIC	24	13,395	77 (0%)
204.0.0.0/6	ARIN	24	14,268	10 (0%)
208.0.0.0/7	ARIN	20	8,775	7,637 (87%)
210.0.0.0/7	APNIC	20	3,665	1,951 (53%)
212.0.0.0/7	RIPE	19	4,755	2,876 (60%)
216.0.0.0/8	ARIN	20	6,197	4,982 (80%)
217.0.0.0/8	RIPE	20	1,514	843 (55%)
218.0.0.0/7	APNIC	20	1,301	360 (27%)
220.0.0.0/7	APNIC	20	473	212 (44%)
222.0.0.0/8	APNIC	20	0	0

ARIN:　American Registry for Internet Numbers
RIPE:　Reseau IP Europeens
APNIC:　Asia-Pacific Network Information Center
LACNIC:　Latin American and Caribbean Internet Address Registry

より細かな経路はほとんど存在しない．しかし，これらの空間は，マルチホーム利用者への割当を前提とし最小割当単位を 24 ビット長程度に小さく設定されたものも多く，マルチホームの利用が細かな経路情報を増やしインターネット上の経路エントリ数を増大させていることに変わりはない．

提案方式によるマルチホームの実現では，複数のマルチホーム利用者間の経路集約が可能である．これにより，マルチホーム利用者ごとにパンチングホールが生じることにはならず，通常のプロバイダ集約と同程度の集約が期待できる．現在，マルチホーム利用者の増加がインターネット上の経路数の増加に大きな影響を与えている．提案方式により，今後のマルチホーム利用者増加による経路エントリ数の増加を抑制できる．

4.3 仮想接続の利用による影響

4.1 で触れたが，提案方式によるマルチホーム接続では，トンネル等の仮想接続を介して転送を行うことによる影響が生じる．これは，パケットがトンネル中を配送される際，トンネル自体を示す IP ヘッダとトンネルヘッダがもとの IP パケットに付加されることによる．例えば，トンネル方式として GRE（Generic Routing Encapsulation）[6] トンネルを用いた場合，最低 28 バイト（IP ヘッダ 20 バイトと GRE ヘッダ 8 バイト）が付加される．

インターネット上の各リンクには，そのデータリンクの特性等に応じて MTU（Maximise Transfer Unit）が定められている．ここで，トンネル両端点間の MTU を 1500 と仮定すると，上記 GRE トンネルによる仮想リンクの MTU は 28 バイトの縮退により 1472 バイトとなる．ところが，実際のインターネット上には，パケットフィルタリング等の設定の誤りから MTU 超過の際の動作を正常に行えないネットワークが数多く存在する．これは大きな問題であるが，MTU 超過処理を正しく行えないネットワークとの通信を維持するために，通常，構築するネットワークの MTU を 1500 未満にしないという運用がなされている．先の GRE トンネルの例では，仮想リンクの MTU として 1500 を維持すると，1472 バイトを超過するサイズのパケットをカプセル化した場合，トンネルの両端点間では IP 断片化された状態で配送される．この場合，1500 バイト長の一つのパケットをトンネルを介して送達する際には，1500 バイト長のパケットと 48 バイト長のパケットの二つのパケットに断片化し送達される（図 4）．

表 4 は，インターネットバックボーン上でトラヒックを 24 時間観測し，パケット総数と消費帯域をそれぞれ 100 として，パケット長ごとの比率をまとめた．更に，このトラヒックを提案手法におけるトンネルを介して送った場合，つまり，48 バイト長のカプセル化ヘッダを付加し必要に応じて IP 断片化処理を行った場合のパケット数及び消費帯域を示している．

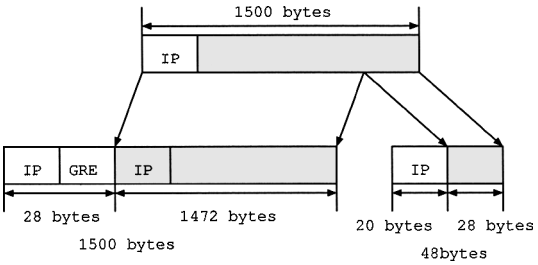


図 4　カプセル化によって生じる IP 断片化
Fig. 4　IP fragment arising in encapsulation.

表 4　カプセル化前後のパケット長分布の比較
Table 4　Comparison a distribution of packet lengthes between before/after encapsulation.

転送されるパケット長	カプセル化前 構成比	（帯域）	カプセル化後 構成比	（帯域）
20 〜 47	24.15	(2.16)	1.92	(0.13)
48	2.89	(0.27)	20.57	(1.94)
49 〜 255	32.23	(5.28)	58.30	(10.66)
256 〜 511	7.31	(5.72)	6.24	(4.77)
512 〜 1023	6.03	(7.73)	7.42	(9.32)
1024 〜 1499	7.29	(19.67)	5.51	(14.60)
1500	20.10	(59.17)	22.06	(64.94)
計	100.00	(100.00)	122.02	(106.36)

この結果から，カプセル化処理により，パケット数が約 22%，帯域が約 6%追加消費されていることが分かる．更に，パケット長ごとのパケット数分布にも変化が見られる．まず，パケット長 48 バイトのパケットの増加が顕著だが，これはパケット長 1500 バイトのパケットがカプセル化前により断片化された結果生じたものである．更に，パケット長 1500 バイトのパケットの増加，パケット長 48 バイト未満のパケットの急減が見られる．

このように，提案アーキテクチャがトンネルを用いたインターネットと利用者網の接続を行っていることによりある程度の影響が生じることが観測された．この影響は，パケット数にしてみると 20%を超える増加と大きなものであるが，帯域に対する影響はたかだか 6%程度である．このたかだか 6%の帯域増に関しては，提案アーキテクチャが提供するリンクの有効利用とト

ラヒック制御に関する利点と比較するとその影響は限定的なものであるといえる．

5. プロトタイプ実装と動作実験

本提案の多重ルーチング型マルチホームアーキテクチャの動作実証のため，プロトタイプ実装を行った．本プロトタイプは，既存方式と比較し提案方式で新たに実現可能となった機能である利用者網への流入トラヒックの細かな制御の実証を目的とした．本プロトタイプで実現した主要機能は，次のとおりである．

- DR–UR 間の仮想接続の確立
- DR での利用者網あてパケット中継
- DR からの利用者網経路広告
- DR における細粒度トラヒック制御

本実装は，UR 及び DR の実装から構成される．DR–UR 間の仮想接続技術としては，IP over IP トンネリング技術の一つである GRE トンネルを採用した．DR では，利用者網アドレス空間の経路広告機能，トラヒック制御に用いる仮想接続の管理，複数の仮想接続へのトラヒック振分け，仮想接続の状態監視機能を実現する機能を実装した．UR では，仮想接続の確立処理，仮想接続を介したパケットの受信処理を実現する機能を実装した．

UR 及び DR の実装は，ソフトウェアルータとしても広く用いられている NetBSD オペレーティングシステムに対して行った．この実装は，NetBSD 1.6 RELEASE に対するカーネル拡張と制御アプリケーションの実装からなる．

まずこの実装の機能試験として，実験網を用いた動作実験を行った．実験網は二つの ISP（AS1，AS2）とその両 ISP に接続する利用者網を模し，各 AS 内に DR（dr1，dr2）を，マルチホーム利用者網に UR（ur）を配置した（図 5）．この網上で，各 ISP 網内のノード（user）からマルチホーム利用者網内のノードへのパケットが，提案アーキテクチャの設計どおり，DR を介し，仮想接続（ここでは GRE トンネル）を通り，UR へ伝送される基本動作の確認ができた．更に，DR において，利用者網あてトラヒックの処理ポリシを設定し，アプリケーション（ポート番号）ごとに異なった接続を介してパケットを配送することが可能となっている点も確認した．

また，上記実験網において一部の接続の障害時における障害検出と経路切換に関しても実験を行った．ここでは，利用者網の対外リンク 2 本を各々切断し，AS1 内のノード（user）と利用者網内のノード（user）との間での接続性を確認した．利用者網の接続のうち実際にトラヒックが流れている側の接続を切断したところ，約 200 ms で仮想接続の生存確認機能が障害を検知しそのトラヒックが他方の接続を介しての配送へと自動変更されることを確認した(注4)．更に，実アプリケーションにおける実験として，DV（Digital Video）による動画伝送アプリケーション DVTS(注5)を用いた利用者網への動画配信中の回線切断実験を行ったが，回線切換の影響は目視上ほとんど気にならない範囲内（8 フレーム程度の脱落）であった．また，アプリケーションのセッション断なども生じないことを確認した．

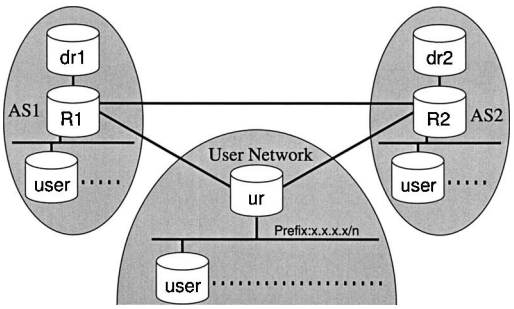


図 5　実験ネットワークトポロジー
Fig. 5　Topology of experimental network.

更に，本提案システムの実運用における知見の獲得を目的とし，本実装を用いた実運用網上での広域運用実験を開始している．

6. 考察と課題

本アーキテクチャでは，利用者網あてのパケットは，常に DR を経由して配送される．このため，利用者網への最短経路による配送と比較すると，DR へと迂回するためにパケット配送時間の増加が懸念される．しかし，DR はインターネットバックボーン上に分散配置されるため，十分な数の DR を分散配置することで，その送達遅延の増加は軽微なものとできると考える．この送達遅延時間の増加問題に関しては，DR の配置状況と遅延時間に関し，予定している実ネットワーク

(注4)：今回の実験では，AS1 内の DR と利用者網の間の遅延が約 20 ms であった．これに対し，仮想接続の生存確認間隔を 100 ms と静的に設定し実験を行った．

(注5)：DVTS は DV フレームを UDP を用いて配送するアプリケーションであり，比較的高品質な網を想定しており，パケット脱落時の再送や受信側アプリケーションでのバッファリングは行われない．映像のフレームレートは 30 フレーム/秒で，約 30 Mbit/s の帯域を必要とする．

上での広域運用実験を通して評価する予定である．

マルチホームにおいて，障害時にも利用者網への到達性を確保することは重要であり，提案アーキテクチャでは生存確認パケットによる障害検出が大きな役割を担っている．現状のプロトタイプでは，**5.** で述べたように生存確認間隔は静的に設定する必要がある．しかし，この確認間隔等は，DR–UR 間の帯域幅や送達遅延・ジッタなどに応じて適正値が異なることは明らかである．そこで，DR–UR 間の帯域幅や送達遅延・ジッタなどに応じて，確認動作のパラメータを動的に更新する拡張を施す必要がある．

本アーキテクチャは，既存のマルチホームアーキテクチャの完全な代替をねらったものではない．DR は隣接 ISP に利用者網の経路を BGP-4 により広告することとなる．実際，DR を複数の ISP に接続するような形態で運用する場合には，既存マルチホームアーキテクチャに基づいた経路広告を行うこととなる．しかし，ここでの広告は利用者網経路が集約された状態で行われることから，マルチホーム利用者の増加がインターネット上の経路数に与える影響は軽微なものとなる．このように，比較的大きなアドレス空間をマルチホーム接続する用途においては，既存アーキテクチャも有効な手段であるといえる．その一方で，提案アーキテクチャは，既存アーキテクチャではなし得なかった，エンドユーザを含めた広範囲へのマルチホームの適用を可能とするものである．

エンドユーザへの本マルチホームの適用を考慮すると，対規模性に関する考察が不可避である．先に述べたように，提案アーキテクチャでは複数ユーザ間での経路集約が可能となり，利用者数の増大が経路制御システムに与える影響は回避できた．更に，**3.3** で述べた DR の複数設置により，DR でのパケット転送処理を分散化することは可能である．しかし，全 DR が各利用者あての仮想接続を保持しているために，利用者数の増加により利用者や仮想接続の管理負荷が増大する．このような利用者数の増加に対して，前章で述べた DR 群を複数用意することで対規模性の向上が可能である．最も単純な手法としては，図 6 のように，あらかじめ利用者を複数のグループに分け（1 群，2 群，...），各グループに対応する DR 群（1 群の利用者を収容する DR 群として DR_1A, DR_1B, ...）を設置する．この場合，各 DR 群は群間で完全に独立して動作可能であり，利用者の増加に対し，利用者グループ及びそれに対応する DR 群を追加設置することでより多くの利用者の収容が可能となる．実際には，インターネット上に分散して DR 設置個所を確保し，各個所に利用者グループ数ずつの DR を設置することとなる．

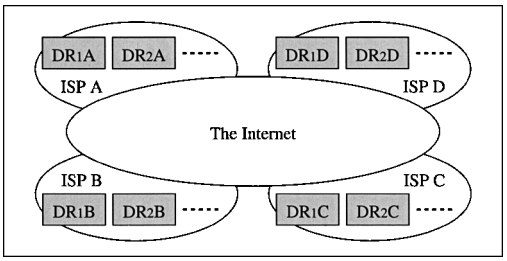


図 6　複数の DR 群を用いた構成
Fig. 6　A structure based on multiple DRs.

また，**4.3** においては，トンネリングによる MTU 縮退の問題に関して述べた．筆者らは，トンネリングによる MTU 縮退の影響を回避するために，MTU 値に影響を与えないトンネリング手法の開発を試みている．これは，利用者網のアドレス空間が比較的狭域であることなどを利用し，カプセル化の際，IP ヘッダ圧縮技術等を用いることで MTU 縮退を防ぐものである．

7. むすび

本論文では，既存マルチホーム技術の問題点をまとめ，その解決として多重ルーチング型マルチホームアーキテクチャを提案した．このアーキテクチャは，インターネット上の経路数の増加といった既存マルチホームの大きな問題を解決するばかりでなく，利用者アプリケーションの要求特性に応じた回線選択といった新たなサービス体系の実現への可能性をもっている．

更に，特に利用者網への流入トラヒック制御を実現するプロトタイプ実装を行い，その動作を確認した．また，広域運用実験へ向けて，現在進めている実装の拡張や制御プロトコルの設計に関して議論した．

謝辞　本研究は，総務省戦略的情報通信研究開発推進制度「国際技術獲得型研究開発」の助成によって行われている．関係者各位に深謝する．

文　献

[1] T. Bates and Y. Rekhter, “RFC 2260: Scalable support for multi-homed multi-provider connectivity,” IETF, Jan. 1998.
[2] J. Hagino and H. Snyder, “RFC 3178: IPv6 multihoming support at site exit routers,” IETF, Oct. 2001.
[3] Y. Rekhter and T. Li, “RFC 1518: An architecture

for IP address allocation with CIDR," IETF, Sept. 1993.
[4] V. Fuller, T. Li, J. Yu, and K. Varadhan, "RFC 1519: Classless inter-domain routing (CIDR)," IETF, Sept. 1993.
[5] Y. Rekhter and T. Li, "RFC 1771: A border gateway protocol 4 (BGP-4)," IETF, March 1995.
[6] D. Farinacci, T. Li, S. Hanks, D. Meyer, and P. Traina, "RFC 2784: Generic routing encapsulation (GRE)," IETF, March 2000.

（平成 16 年 1 月 14 日受付，5 月 7 日再受付）

宇多　仁　（正員）

1997 東京理科大・理工卒．1999 北陸先端科学技術大学院大学情報科学研究科博士前期課程了．2004 同博士（情報科学）了．同情報科学センター助手．現在，インターネットの高機能化及び運用技術に関する研究に従事．

小柏　伸夫　（正員）

1999 芝浦工大・システム工卒．2001 北陸先端科学技術大学院大学情報科学研究科博士前期課程了．2004 同博士（情報科学）了．現在，(株）インテック・ネットコア勤務．インターネット経路制御に関する研究開発に従事．

永見　健一

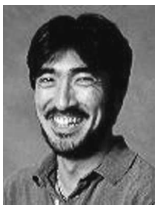
1992 東工大理工学研究科了．（株）東芝入社．CSR/MPLS 等の研究開発に従事．2001 東京工業大学工学部博士（工学）．2002（株）インテック・ネットコア入社．現在，インターネット経路制御とネットワーク監視技術の研究開発に従事．

近藤　邦昭

1992 神奈川工大情報工学科了．1997（株）インターネットイニシアティブ入社．BGP-4 の監視運用システムの開発等に従事．2002（株）インテック・ネットコア入社．現在，インターネット経路制御とネットワーク監視技術の研究開発に従事．

中川　郁夫

1993 東工大理工学研究科了．（株）インテック入社．同社研究所にて，ネットワーク管理，大規模経路制御技術，次世代インターネットに関する研究に従事．2002（株）インテック・ネットコア取締役．現在，ネットワークの信頼性に関する研究に従事．



篠田　陽一

1983 東工大・工卒．1985 同大工学部博士前期課程了．1988 東工大工学部助手．1989 東京工業大学工学部（工博）．1991 北陸先端科学技術大学院大学情報科学研究科助教授．現在，同大学情報科学センター教授．

江崎　浩　（正員）

1987 九州大学大学院修士課程了．同年，（株）東芝入社．CSR/MPLS 等の研究開発に従事．1998 東京大学博士（工学）．1998 東京大学計算機センター助教授．現在，同大学情報工学系研究科助教授．