

集中経路制御ブリッジネットワークの実現と応用*

3 F-6

平中幸雄 松本慎平 櫻井啓嗣† 伊藤悟史
山形大学工学部

1. はじめに

バックボーンネットワークの高速化には、その中継装置の単純化が必要である。その効果的な実現方法として、集中経路制御を導入したブリッジネットワークを、昨年の秋季大会で提案した[1]。今回、この方式がインターネットバックボーン中継方式として優れていること、基本動作実証実験の結果[2]、経路最適化の方法[3]、実用的な応用として、階層化による大規模化の方法、マルチキャスト実現方法、マルチホーム接続での経路制御、移動 IP への対応、広域 VLAN の実現法などを述べ、その有効性を示す。

2. 次世代インターネットの条件

2. 1 高速化

従来のインターネット接続は、特に WAN 部分を中心に、ルータ間パケット転送をベースにしていた。しかし、WAN 部分の高速化とともに、転送遅延の問題や経済性からルータを出来るだけ排除する傾向にある。ラベルスイッチング[4]などの技術は、経路設定のみルータに依存するが、経路設定がすめば、ルータを経由しないで ATM 層など下位層で転送することによって、転送遅延の削減とボトルネック解消を図っている。従来のルータは、転送と経路制御を一体で行ってきたが、両者を分離するのが当然と考えられる時期に来たと言ふべきである。

2. 2 ネットワークの構造の柔軟性

現状のルータ接続中心のネットワークでもうひとつ問題なのは、経路制御が物理的なネットワークに制約を加えるという点である。ルータのポートごとにネットワークを規定するという基本構造は、明快ではあるが、CIDR がネットワーク接続変更の実質的な制約を与えている[5]ように、経路の論理的な制御は、単なる論理操作ではなく、物理的な構成の前提を要求することが多い。自由に経路制御をしたいとすると、それを可能にする物理的なネットワークを組む必要があるというのが、現状である。VLAN の可能な装置によってこの問題を回避することはできるが、ネットワーク構造がさらに管理困難な構造になるであろう。

2. 3 経路制御ポリシー

インターネットはもともと分散経路制御であったが、経路制御ポリシーの実現のため、AS ごとの実質的な集中経路管理に移行してきている。さらに、AS ごとに独立な制御をしていると、到達不能や経路決定不能が起り得る。この解決のため、routing registry が設けられた[6]が、これは、インターネット全体の集中経路制御の必要性を示していると考えられることができる。

2. 4 帯域保証

近年、音声・動画のリアルタイム通信が拡大するにつれ、インターネットにおいても帯域保証の要求が強くなってきている。動的経路変更をしながら帯域保証するのは困難である。帯域保証の要求に対しては、静的経路制御と一体で帯域確保を行う必要がある。ネットワーク経路上の利用可能帯域を確認しながら、経路設定を行う必要がある。これは論理的には、ひとつの経路設定主体が経路設定作業を行う、集中型経路制御であり、実際にも一箇所で処理を行う集中制御で実現すべきである。

2. 5 マルチキャスト

インターネットは、従来のラジオ・テレビ放送も吸収してしまう通信インフラストラクチャーとして考えて

* Implementation and Applications of Centralized Route Control Bridge Network

Yukio HIRANAKA, Shinpei MATSUMOTO, Keiji SAKURAI and Satoshi ITO, Yamagata University

4-3-16 Jonan, Yonezawa, 992-8510 Japan, corresponding e-mail: zioi@ieee.org

† currently, NTT East Corp.

も、おかしくない状況になってきた。しかし、放送型通信を経済的、効率的に行うには、伝送媒体・方式自体が放送型に適している必要がある。たとえば、電話型ユニキャストを前提に構成された ATM を放送型マルチキャストに使うには、ATM スイッチにやや複雑な仕組みを導入せざるを得ず、その制御方法を考えるとかなり厄介な問題が発生する。LAN エミュレーションのように、物理層でのマルチキャストをあきらめる方法では、既存の放送が可能な通信網に経済的に取って変われる可能性は少ない。放送型通信は、放送型ネットワーク媒体で伝送すべきである。

3. 集中経路制御型ブリッジネットワーク

提案の集中経路制御型ネットワークは、以上の問題を解決すべく検討中のもので、転送と経路制御を分離する。転送には、マルチキャストに適した通信方式である MAC 層ブリッジネットワークとし、従来のネットワークとの整合性を基本的に確保する。帯域保証などの高度の機能は、経路設定を行う装置の機能として実現する。経路制御を含むネットワーク制御装置を設定することで、さまざまな機能の実現方法が明確になる。ブリッジネットワークとすることで、IP 層に限られないさまざまなプロトコルの共存が容易に実現できることになり、より簡易な、もしくは新しいプロトコルの実装や、筆者らが検討中の汎用データフォーマット [7] による、プロトコルをほとんど必要としない通信方式などの展開が可能となる。

ブリッジネットワークはブロードキャストを前提に構成されており、大規模化での効率の低下が問題で、従来も検討されたが [8]、実用的なものではなかった。しかし、自分で通信相手を見つける (arp) ことをあきらめ、とりあえず集中経路サーバに任せることにすれば、つまり通常ブロードキャストで解決することを、特定のホストに送ることで処理可能にすれば、ブロードキャストの無いブリッジネットワークが実現できる。

4. ネットワーク構成と基本動作

集中経路制御型ブリッジネットワークの基本構成を図 1 に示す。ネットワークはルートサーバ (RS)、ブリッジ (B)、エッジルータ (ER) から構成される。エッジルータ以下に非幹線系のネットワークが接続することになる。

ルートサーバは、各エッジルータから送られてくるネットワーク情報とブリッジからの情報を収集し、ネットワークの全体の情報をここで保有し、ネットワークの経路制御を集中して行う。また、Proxy ARP として動作し、すべてのアドレス解決をここで行う。さらに、エッジルータ間の通信の際、ICMP redirect によって、相手先確認後は通信がルートサーバを介さず、エッジルータ間で直接通信を行えるようにする。ブリッジは、ブロードキャストや送信ポートが不明なパケットをルートサーバへのデフォルト・ポートへ送信する。エッジルータは、ルートサーバへ定期的にネットワーク情報を含んだ RIP パケットを送信する。非幹線系ネットワークに対しては通常のルータとしての動作をする。

ネットワークの基本動作の流れは以下のようなになる(図 1)。

1. 経路サーバ(RS)直下のブリッジ(B)を起点に spanning tree 化を行う。
2. spanning tree 化の過程で各ブリッジはルートサーバ方向へのデフォルト・ポートの設定を行う。
3. 各エッジルータにはルートサーバをデフォルトルータとして登録する。
4. 各エッジルータは定期的にルートサーバ宛にネットワーク情報を含んだ RIP パケット転送する。RIP パケットが流れる経路上のブリッジはエッジルータの存在を知る。特定相手への default 経路通信に含めてそれを行ってもよい。

これで、ルートサーバから、全部の ER と接続ネットワークへの経路ができ、これ以降通信が可能となる。各ブリッジにルートサーバ方向へのデフォルト・ポートを設定するが、一般のブリッジではデフォルト・ポートを設定することは出来ないので、ブリッジの機能を拡張することによりこれを実現する。送信すべきポートが不明なパケットやブロードキャストパケットをブリッジが受け取った場合、それらのパケットをすべてルートサーバへ送る。

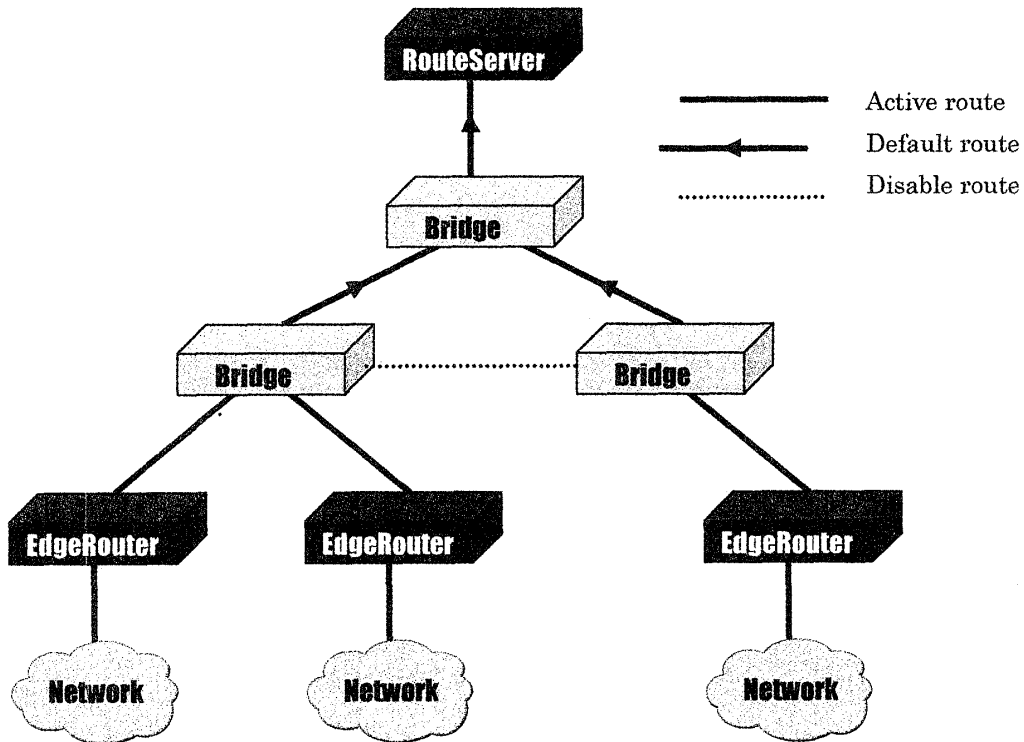


図1 ブリッジバックボーンネットワークの構成

エッジルータからの RIP パケットが経由したブリッジは、発信元のエッジルータの物理アドレスとその方向を学習することになる。RIP パケット送信の間隔を、ブリッジのステーションキャッシュのリフレッシュ間隔より短くすることで、ブリッジは常にエッジルータの位置を把握することができる。

一度ルートサーバにパケットを送ると、ルートサーバは ICMP redirect で相手の IP アドレスを通知してくるので、MAC アドレスをルートサーバの Proxy ARP により解決すれば、以降は直接通信ができる。Spanning tree の構成上、ルートサーバに到達する前に、必ず宛先の MAC アドレスを知っているブリッジを通過することができる。

5. 基本動作実証実験の結果

考案したネットワークを実際に構成し、接続性の検査や動作の確認といった基礎的な実験を行った。100Base-TX Ethernet adapter を複数持たせた PC/AT 互換機 (CPU:AMD K6-2 350MHz、OS: FreeBSD 2.2.8R) を 9 台使用してルートサーバ、ブリッジ、エッジルータなどの機能を実現した。中継マシンを、ブリッジにする場合とルータにする場合とで伝搬遅延を測定したところ、1 台あたりそれぞれ、 $20\mu s/50\mu s$ となり、大きな差があった。先頭部分のみで判断するオンザフライ処理をすれば、大幅な高速化が期待できる。

実験の結果、エンドステーション間での接続性が確認された。ルートサーバの ProxyARP、エッジルータからの RIP の送信とルートサーバでの受信もうまく動作した。しかし、エンドステーション間の通信では ICMP redirect の通知がうまく動作せず、すべてのパケットがルートサーバを介しての通信になった。また、ICMP redirect と arp の 2 段階処理は 1 回で済むように工夫すべきと考えられる。

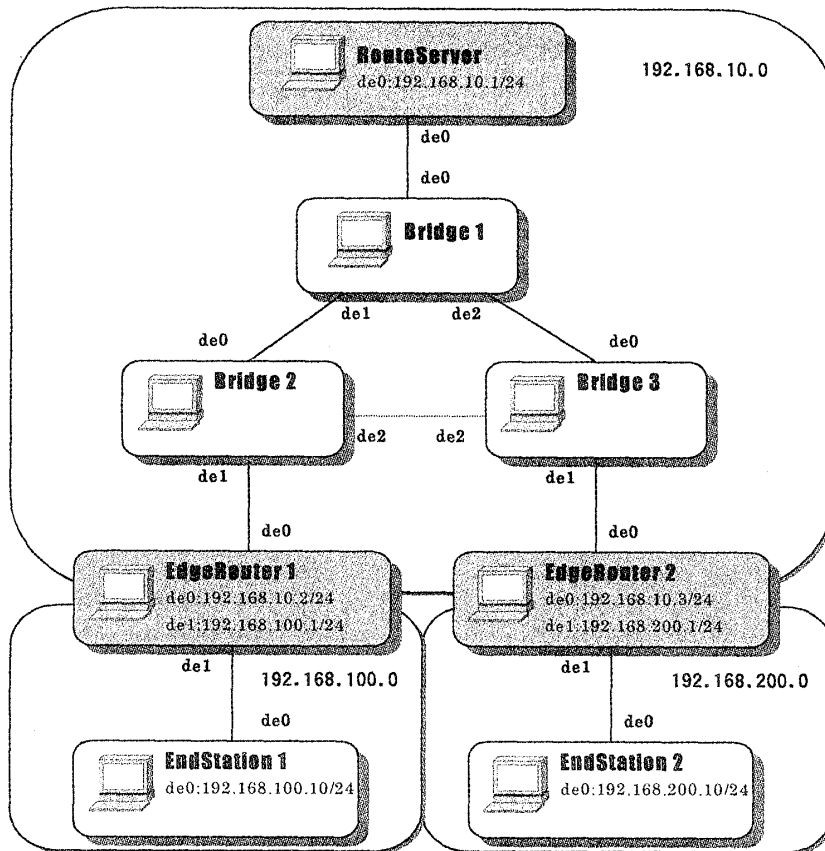


図 2 実験用ネットワーク構成

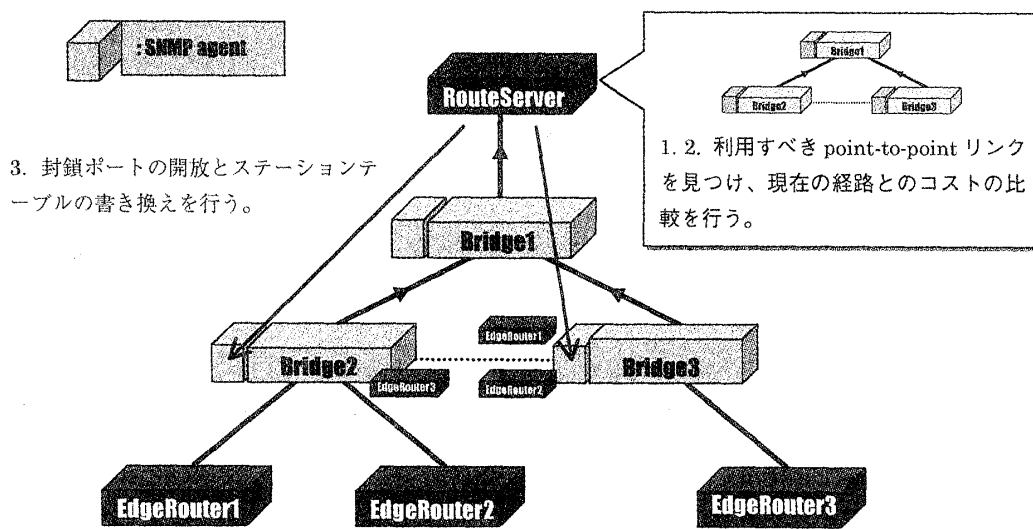


図 3 spanning tree 封鎖ポートの利用

6. 経路最適化の方法

経路計算と経路の決定を一台の装置（ルートサーバ）で行う集中方式では、情報を一元管理するために、厳密に経路を管理することができる。図3では、各ブリッジの状態をSNMPを利用して監視することを示している。これにより、たとえば、spanning tree 化により利用できなくなった封鎖ポートがあった場合、それを利用した経路最適化制御を行うことができる。ショートカットした方がよい、たとえば、エッジルータ ER1 の MAC アドレスを、ブリッジ B3 の近道となるポートの MAC アドレステーブルにルートサーバから書きこむ。すると、ER3 から ER1 へのパケットは B3 から B2 へ直接届くことになる。

7. マルチキャストの実現

マルチキャストでは、中継装置が必要に応じてパケットを複製して配送するのが望ましい。これによりユニキャストと比べネットワークに負担の少ない一対多の通信が実現できる。本ネットワークでは以下のようにして実現することを考えている。

1. マルチキャストグループ A または B に入りたいホストは、それぞれが接続しているエッジルータに対し、その意思を伝えるメッセージを送信する。ホストからエッジルータへは一般的な IP マルチキャストの方法を利用する。
2. それを受け取った各エッジルータはその情報をルートサーバに伝える。このメッセージは IP マルチキャストプロトコルを代用することを考えている。図4ではエッジルータ 3、4 以下のホストがグループ A、エッジルータ 1、2、4 以下のホストがグループ B に入ろうとしている。
3. メッセージを受け取ったルートサーバはグループ情報を収集し、マルチキャストグループとそのグループ宛のデータを送るべきエッジルータのペアの情報を保有する。
4. ルートサーバはこの情報を基に、各マルチキャストグループのパケットが通るべき経路上のブリッジのステーションキャッシュに、マルチキャストグループに対応させた MAC アドレスを書き込む。この時、あるグループ宛のパケットを停止するための MAC アドレスも指定する。これによりパケットが不要なネットワークに流れるのを防ぐ事ができる。

これ以降、マルチキャストによる通信が可能になる。マルチキャストアドレス宛のパケットを受け取ったブリッジは、ステーションキャッシュを見て、それを送るべきポートとデフォルトポートに送るようにする。この場合、ブリッジの拡張機能として、同じマルチキャスト MAC アドレスを2つのポートに書き込めるようする機能と、送るべきポートが2以上ある時はパケットを複製し必要なポートに送れるようにする機能が必要となる。

8. ルートサーバの階層化による大規模ネットワークへの適用

インターネット全体など、大規模ネットワークに適用するには、経路サーバを階層化し、経路サーバ間の通信は、下位経路サーバをエッジルータと同じように扱えばよい。既存ネットワークとの接続は、エッジルータを介して行える。図5のように、ローカルサーバが管理する各ブリッジネットワークを、ルートサーバ直下のブリッジで上位ブリッジネットワークに接続する。spanning tree 形成は、ローカルルートサーバ直下のブリッジを基点におこない、ルートサーバごとにひとつのブリッジネットワークを形成し、図6の手順でルートサーバ同士の接続を行い、ネットワークの結合をする。

9. マルチホームにおける経路選択

ルートサーバの集中経路制御により、あるホストからマルチホーム化されたホストへのアクセスの際、複数ある経路から意図的に経路を選択して送信できるようにする。その方法は次のようになる。

1. あるネットワークからマルチホーム化されたサイトへ通信する場合、どちらの経路を通るべきかといった情報をルートサーバに持たせる。たとえば図7のネットワーク3にあるホストがマルチホーム化されたネットワーク0のホストへ通信する時はエッジルータ2を通る経路を利用し、ネットワー

ク 4 にあるホストからネットワーク 0 にあるホストへ通信する時はエッジルータ 1 を通る経路を利用するといった情報を持たせる。これは対象となるサイトがルートサーバへその意思を伝えるメッセージを送信することで登録する。

2. 対象となるサイトからのアクセスがあった場合、適切な経路を通して通信が行われるようにルートサーバが宛先のエッジルータを選択してパケットを送信する。例えばネットワーク 3 にあるホストがネットワーク 0 のホスト宛にパケットを送信した場合、ルートサーバは受け取ったパケットをエッジルータ 2 宛に送信することになる。

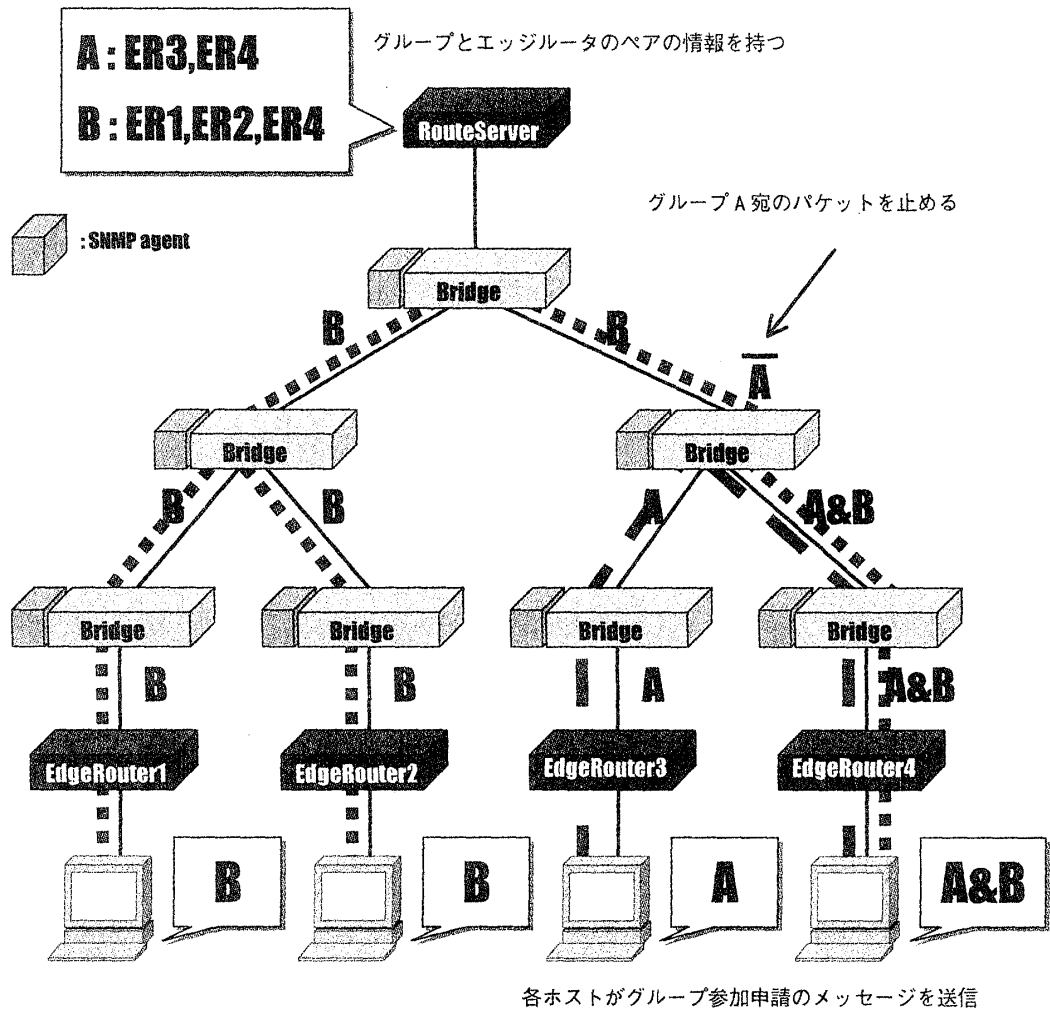


図 4 マルチキャストの実現

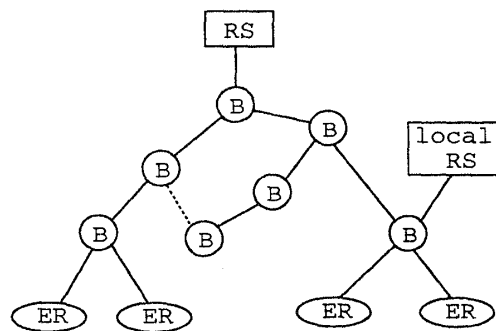
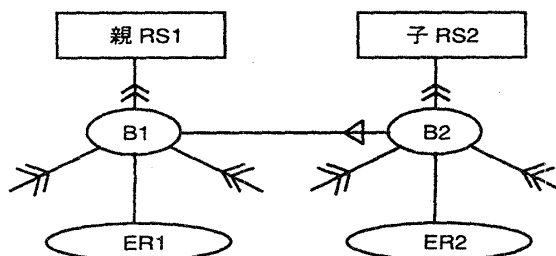


図 5 All bridge backbone network

RSの階層化:



1. 子RSに親RSのIP MACを設定する。
2. 中継B2に親RSのMACとportを設定する。
3. 子RSから親RSへパケットを送りRS間通信路を作る。
4. ER1からER2に接続する。

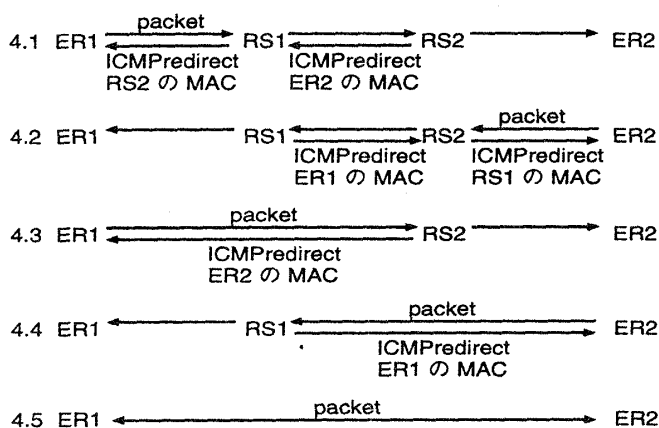


図6 ルートサーバの階層化手順

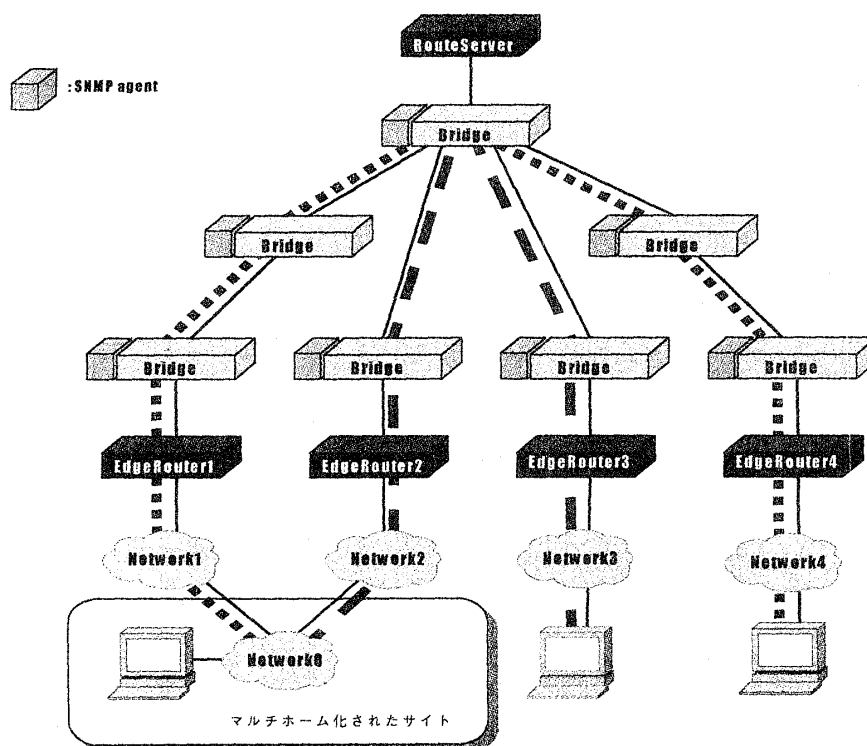
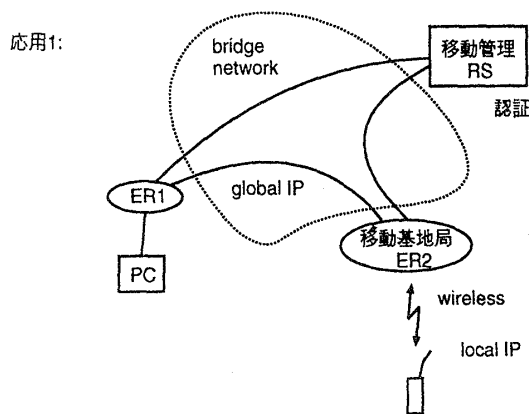


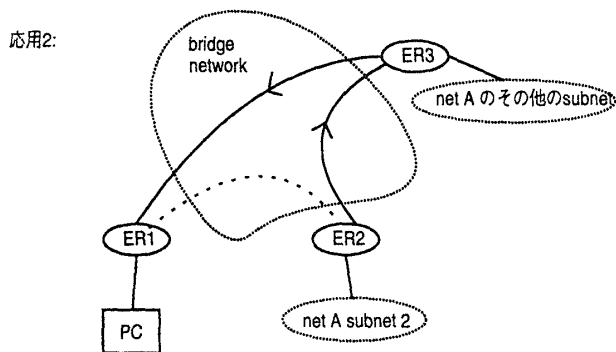
図7 マルチホームにおける経路選択

10. 移動 IP への対応と広域 VLAN の実現

図 8、図 9 には移動 IP や VLAN への応用も示す。



移動基地局では管理局から移動機のMACに対するglobal IPをもらい、以降アドレス変換する。
最初はRSを通すが、以降はER1 <-> ER2の最適経路で通信する。



ER1 ER2 は 最初ER 3を経由して接続をするが、2回目からは直接通信する。
VLAN と同様論理的な結合が可能になる。

図 8 移動 IP 利用の実現

図 9 VLAN の実現

11. まとめ

集中経路制御型ブリッジネットワークが実際に動作することを示し、現状のインターネットの多くの問題を解決する可能性があることを示した。実装上の課題は多くはなく、既存の機器の設定変更や組合せでかなりの部分が実現できると考えている。経路サーバについては、最近の3層スイッチに、MAC アドレス通知機能、エッジルータ監視機能を付加すればよいであろう。

この研究の一部は通信・放送機構平成十年度新規事業創出型研究開発制度による受託研究として行った。

参考文献：

- [1]平中幸雄、バックボーンネットワークのブリッジによる高速化と集中経路制御の考え方、情報処理学会第 57 回全国大会,2G-01,1998.10.
- [2]松本慎平、ブリッジによるバックボーンネットワークの高速化、情報処理学会東北支部 1998 年度第 4 回研究会,資料番号 98-4-19,1999.3.
- [3]桜井啓嗣、ブリッジネットワークにおける集中経路制御,情報処理学会東北支部 1998 年度第 4 回研究会,資料番号 98-4-20,1999.3.
- [4]Tom Noll,MPLS: The New Order in IP Networking?, Network Magazine, vol.14, no.4, pp.48-52, 1999.
- [5]Tim Bass, Internet Exterior Routing Protocol Development:Problems, Issues, and Misconceptions,IEEE Network,July/August,50-55,1997
- [6]Ramesh Govindan et.al., An Architecture for Stable, Analyzable Internet Routing, IEEE Network, vol.13, no.1, pp.29-35, 1999.
- [7]平中幸雄・嘉藤雅文、汎用データフォーマットによるマルチメディアデータの表現、情報処理学会第 59 回全国大会,4V-09,1999.9.
- [8]Riaz Ahmall and Fred Halsall, Interconnecting High-Speed LANs and Backbones, IEEE Network,vol.17, no.5, pp36-43 1993.