

キャンパスネットワークのネイティブマルチキャスト化

戸田 洋三

千葉大学 総合情報処理センター

〒 263-8522 千葉市稲毛区弥生町 1-33

E-mail: yozo@ipc.chiba-u.ac.jp

あらまし IP マルチキャスト技術はキャンパスネットワーク程度の規模の LAN ではじゅうぶん実用となるレベルに達し、現在はプロバイダ間接続における経路制御技術など広域ネットワークへの適用を目指した研究がすすめられている。しかし現状キャンパスネットワークを構成するルータでは IP マルチキャストパケットのルーティングは行わない・行なえない設定となっている例が多く、IP マルチキャスト技術の一般ユーザーへの普及が進まない原因となってしまうている。IP マルチキャスト技術の研究を進めていくには、幹線レベルでマルチキャスト対応を行なって一般利用者がもっと簡単にマルチキャスト機能を活用できる環境を提供する必要がある。

本稿では千葉大学キャンパスネットワークを例として、幹線を構成するルータを IP マルチキャストに対応させ全学的な IP マルチキャスト環境を実現する過程でどのような検討を行なったかについて報告する。

Making native multicast campus network

Yozo TODA

Information Processing Center, Chiba University

1-33, Yayoicho, Inage, Chiba 263-8522 JAPAN

E-mail: yozo@ipc.chiba-u.ac.jp

ABSTRACT IP multicasting technology is matured enough to utilise within campus networks, and research for wide-area multicast, such as inter-provider multicast routing, is on-going. But we are facing difficulty in using IP multicast in campus networks, because in most cases routers of the backbone network are not configured as multicast-enabled, or even they are not multicast-capable ones. For more research activity and making broad utilisation of IP multicasting technology, we should make the backbone network multicast-enabled and provide the ready-to-use environment for multicast users.

This article describes how we are considering about making Chiba University campus network multicast-enabled; which routing protocol to be used, router configurations, and remaining problems.

1 IP マルチキャスト

多数の受信者に向けてデータを送る際、通常のユニキャストの形態では個々の受信者に対してパケットを送る必要があり、ネットワーク帯域・送信ホストの能力が無駄に消費されてしまう。IP マルチキャストは送信ホストのかわりにルータが必要に応じて複数の方向にパケットを配送することにより効率よく通信を行なおうという仕組みである [1][2][3]。ホスト側の機能としては、最近の Unix 系 OS やパソコン用 OS などでは最初からマルチキャスト機能が実装されるようになってきた [4]。これに対してルータ側の機能としてはまだ実装されていないか不十分なものが多く、ユニキャストルータとは別に Unix ホストをマルチキャストルータに仕立てて対応する例が多い。この場合マルチキャスト経路情報の伝搬によく使われているのが **mrouted** というプログラムであるが、このプログラムをインストールしたり他組織との間にトンネルを設定して MBone に接続したりといった作業は一般利用者には難しく、マルチキャスト技術利用の障害となっている。遠隔講義・ビデオオンデマンド・NTP による時刻情報配布などマルチキャスト環境ならではのアプリケーションを普及させるには普段から一般利用者が気軽にマルチキャスト対応アプリケーションを利用できるよう幹線部分からマルチキャストに対応しておくことが重要である。

本論文では千葉大学ネットワーク幹線のマルチキャスト化のために検討した事項について報告する。同様の規模のネットワークにおけるマルチキャスト環境構築を考える際の参考となれば幸いである。以下、第2節ではいくつかのマルチキャスト経路制御プロトコルの特徴についてまとめ、第3節では千葉大学ネットワーク幹線をマルチキャスト対応にするために検討した事項について述べる。第4節ではマルチキャスト環境を維持・管理していくために必要となるトラフィックなどの監視体制について述べる。

2 IP マルチキャスト経路制御

マルチキャストパケットの配送を行なうルータは以下のような動作によりマルチキャストパケットを転送する：

1. マルチキャストパケットを受信する
2. マルチキャスト経路表に照らして正しいインタフェースから来たものであることを確認する (Reverse Path Checking)

3. フォワーディングキャッシュを参照して転送すべきインタフェースのリストを得る

4. 上記リストのインタフェースからパケットを送信する

通常のユニキャストルータでは宛先アドレスと対応するゲートウェイを組にした経路表を使ってルーティングを行なう。マルチキャストルータでは送信元アドレスと対応するゲートウェイを組にした経路表 (マルチキャスト経路表と呼ぶことにする) を持つ他に、グループアドレス毎にパケットを転送する必要があるインタフェースのリストを持つ (フォワーディングキャッシュと呼ぶことにする)。

Reverse-Path-Checking とは、受けとったパケットに記録されている送信元アドレスをマルチキャスト経路表から探しだし、そこに記録されている (incoming) インタフェースがパケットを受けとったインタフェースと一致している場合のみパケットの転送を行なうということである。これが一致していない場合、マルチキャストパケットの配送経路がループするのを防ぐため転送は行なわずにパケットを廃棄する。また、フォワーディングキャッシュは各グループアドレス毎に転送すべき (outgoing) インタフェースのリストを対応付けたものである。パケットを転送する場合にはフォワーディングキャッシュを参照し、ここに登録されているインタフェースのみにパケットを送り出す。このようにマルチキャスト経路表とフォワーディングキャッシュという二つのデータ構造を使って無駄なトラフィックを抑え効率的なパケット配送を行なうわけである。ルータが持っている二つのデータ構造をネットワークトポロジーおよびグループメンバの所在にあわせて自動的に更新するためにルータ間で情報を交換する手順をマルチキャスト経路制御プロトコルと呼ぶ。ここでは **DVMRP** [6]、**PIM-DM** [7]、および **PIM-SM** [8] の3種類のプロトコルをとりあげる。

DVMRP (Distance Vector Multicast Routing Protocol) は現在もっとも広く利用されているプロトコルである。IP マルチキャストの実験が始まった当初から開発・利用されてきた **mrouted** というプログラムが主流で、さまざまな種類の Unix 系マシンにインストールして使うことができ、「トンネリング」によりマルチキャストに対応していないルータを飛び越えてべつの **mrouted** との間を接続していくことができる。この仕組みによりインターネット上でマルチキャスト環境の実験を行なう組織を相互接続したものを **MBone** と呼んでいる。

DVMRP では送信元アドレスの経路情報を独自に(ユニキャスト経路とは独立して)扱うことでユニキャスト経路とマルチキャスト配送経路とを異なった状態に保つことができる。これはほとんどのルータがマルチキャスト未対応でトンネリングによりマルチキャスト環境の接続性を確保しなければならない状況では必須の機能であった。また、ファイアウォールを構築している環境でユニキャストパケットとマルチキャストパケットの通り道を別にしておきたいというような要求に対応できる便利なプロトコルでもある。しかし、MBone 接続サイトが増加するにつれて DVMRP で交換する経路情報は増加しており現在はおよそ数千程度になる。MBone に接続するとこれらの経路をすべて受けることになるため mrouterd を用いて MBone に接続するにはそれなりの能力を持ったマシンが必要となっている。

キャンパスネットワークで広く使われているルータは Cisco の製品だと思われるが、最近の Cisco ルータは PIM (Protocol-Independent Multicast) と呼ばれるマルチキャスト経路制御プロトコルを実装している(実際には PIM-DM と PIM-SM の 2 種類がある)。PIM では DVMRP とは異なり、送信元アドレスの経路情報としてユニキャスト経路情報をそのまま流用する。ユニキャスト配送経路とマルチキャスト配送経路が常に一致している環境では mrouterd を立ちあげるかわりにルータをそのまま使った方が手軽である。

DMVRP や PIM-DM では各グループアドレス毎に最初に流れてきたマルチキャストパケットはとりあえずすべてのルータに配送してしまい、配送不要(そのグループに join しているメンバがない)ルータからは prune メッセージをもらうことで枝刈りを行なう。この仕組みを broadcast-and-prune と呼ぶ。prune メッセージには有効時間が定められており、定期的にパケット再配送と prune メッセージ応答が繰り返される。

これにたいして PIM-SM では、マルチキャストパケットの配送を受けたいルータが明示的に自分の存在を送信ホストに通知する必要がある(explicit-join)。この通知の仲介役として前もって特定のルータを RP(Rendezvous Point, ランデブーポイント)としてルーティングドメイン中のすべてのルータに教えておく。送信ホストから送り出されたマルチキャストパケットを受信した(first-hop)ルータはこの RP に向けてパケットを転送しグループメンバへの配送を依頼する。あるグループに属するメンバを持っている(last-hop)ルータは

あらかじめ RP にその存在を通知してそのグループに関する RP からルータまでの配送木を確立する。このように RP は送信側と受信側がコンタクトする場所として機能するわけである。

パケット配送経路が構成する木構造に着目してプロトコルを分類すると、送信元ホストを根とする source-based tree をつくるものと RP(ランデブーポイント, Rendezvous Point)を中心にして配送木 shared-tree を構成するものがある。DVMRP と PIM-DM は source-based tree 型、PIM-SM は shared-tree 型のプロトコルである¹。ただし、PIM-SM は設定によりある程度多量のトラフィックが流れるようになる source-based tree に切り替える仕組みを持っている。

shared tree 型プロトコルの場合には RP を中心としてトラフィックが集中する傾向があるので、太いリンクと細いリンクが混在している環境での利用には注意する必要がある²。もっとも、キャンパスネットワークの範囲では帯域幅に余裕がある場合が多いとも考えられるので backbone の部分を shared-tree として使うというようなデザインもよいかもしれない。

3 キャンパスネットワークでのマルチキャスト経路制御

現在千葉大学キャンパスネットワーク幹線のルータはマルチキャスト対応にはなっていない。総合情報処理センターで mrouterd を立ちあげ、MBone との接続および学内接続用に使っている。MBone に接続したい場合には個々に mrouterd を立ちあげてセンターの mrouterd との間にトンネルを張ってもらっている。この方法では IP マルチキャストを利用したいところは自力で Unix マシンおよび mrouterd の設定を行なう必要があり、初心者にはなかなか難しい。そこでマルチキャスト環境を普及させるためには以下のような点に注意しつつ、もっと利用者側に負担のかからない方法を考える必要がある。

- ユーザ側で mrouterd を立ち上げる手間はかけさせたくない
- ただし必要に応じてユーザ側にあるルータ(mrouterd, pimd, cbtd など)と接続できたほうがよい(MBone との接続も含む)
- マルチキャスト経路情報を調整できるようにし学外に迷惑をかけない

¹ちなみに MOSPF は source-based tree 型, CBT[9][10] は shared-tree 型のマルチキャスト経路制御プロトコルである。

²ちなみに PIM-SM では配送経路は一方向であるが、CBT での配送経路は双方向に使われるため PIM-SM よりも集中の度合は緩和されるものと思われる。

- 不正なパケットは filtering できるようにし学外に迷惑をかけない (flooding 対策も含む)
- boundary および threshold 設定の確認

MBone への接続においてはまだしばらくは DVMRP, 具体的には mrouted が使われると思われるので DVMRP との相互接続性を確保しておくことは重要である。また、ユーザ側で新たなプロトコルを使うルータを立ちあげて接続したいという希望が出てきた場合にも対応できるとよい。いまのところ考えられるのは mrouted, pimd, cbtd といったプログラムを使ってみたいというような状況だろうか。MBone との接続においても, mrouted にかわって pimd, cbtd などとの接続という可能性は今後出てくると思われる。

結局 2 通りの選択枝を考えた。

- (a) 経路制御プロトコルとしては DVMRP を利用し, mrouted をインストールした Unix マシンを必要な数だけ配置する
- (b) PIM を利用し, 幹線接続用ルータをマルチキャスト対応させる

(a) 案では, いまも改良が続けられている mrouted の最新版を利用できる。別のプロトコルを採用する場合にもプログラムさえあればすぐに対応できる, などの利点があるが予算と管理の手間を考えると実現は難しい。一方 (b) 案ではすでに利用しているルータの設定を変更すれば済むので簡単に対応できるのが利点であるが, 逆に運用ポリシー上既存ルータ設定の変更が許されるのか・実作業部隊に変更後の運用サポートを任せられるのかという点で問題でもある。

千葉大学の場合には (b) 案を採用する予定である。学内幹線ルータの設定作業を行なっているのは総合情報処理センターのスタッフ数名ほどなのでこれらのなかで相談しておけばよい。また, DVMRP と比較して PIM-DM の仕組みはごく単純なので, なにか問題が起こっても対応はそれほど困難ではないと考えている。

なお, Cisco ルータでは pruning 機能などに問題のあるバージョンも存在するようだが, 我々が使っているものは問題ないようだ。

PIM には PIM-SM と PIM-DM の 2 種類があるが, 我々は PIM-DM を選択したい。その理由は次のとおり。

- PIM-SM では個々のルータに RP は誰かという情報を設定しなければならず, RP となるルータが故障した場合やアドレスを変更する場合に対応作業が面倒

- RP 情報を自動的に設定する auto-rp という仕組みがあるが, 分かりづらそう
- ネットワーク幹線にそれほど帯域幅の細いところがあるわけではなく, broadcast-and-prune 型でも問題ないだろう
- ユーザ側のルータを幹線に接続している部分もあり, この場合マルチキャストに関しても設定をやらなければならないのでできるだけ設定項目は簡単にしたい

千葉大学の学内幹線は FDDI ネットワークと ATM ネットワークの 2 つの部分からなっており, ATM ネットワークに部局 LAN を接続する場合は自分でルータを用意して接続するようお願いしている。すなわちこの場合のルータはユーザ側の所有なのでマルチキャスト関連の設定も部局側管理者にやらなければならない。設定に関するドキュメントをつくったり実際の作業を手伝うとしてもやはり必要な設定項目はできるだけ少なくしておいた方がよいであろう。

4 ネットワーク監視

MBone 上でいままでに経験されている問題には

- なんらかの原因で pruning がうまくいかずに flooding を起こす
- ユニキャスト経路情報を DVMRP の経路情報に突っ込んで MBone 上に流してしまい他組織の mrouted をハングアップさせてしまう
- トンネルの経路が想定していたリンクとは違った部分を通り回線が飽和してしまう
- ローカルな範囲に収めておくべきセッションが外部に流れてしまい, 大量のトラフィックで回線を埋めてしまう

などがある。このような問題が起こらないよう初期設定をしっかりと行なうとともに普段からネットワークの状況を監視して, 問題が発生したら済みやかに対応できるよう準備しておく必要がある。MRTG[11] によるトラフィックの定期観測は必須であろう。mrouted の場合には SNMP 対応バージョンを利用することによりトンネルも含めてインタフェース毎のトラフィックを観測することができる。Cisco ルータについても SNMP 機能をきちんと設定して同様の観測を行なわなければならない。トラフィック全体の観測の他に, セッション毎のトラフィックを調べる・マルチキャスト経路

情報の総量を調べる・特定のグループにjoinしているホストが学内にあるかどうかをリアルタイムで調べるなど、問題の早期発見や利用状況の把握に役立てるためにはいろいろなデータが考えられる。

マルチキャストの通信に関して運用上必要なデータの種類、あるいはユーザの利用状況を調査するために必要なデータの種類の観測技術については今後も検討をすすめる必要がある。

5 まとめ

本稿ではキャンパスネットワークの幹線をマルチキャスト対応にするための検討項目について述べた。我々の環境ではCiscoルータを使っている関係上PIM-DMを利用するのがもっとも適していると考えた。しかし、プロトコルを調べただけではあるが、CBTなどをキャンパスネットワーク上で使うのも悪くないという印象を持っている。

第3節では(a)案は実現が困難としたが、最近ではall-in-oneタイプのサーバや弁当箱タイプのDHCPサーバなどが出てきている。同様にしてマルチキャストルータとしてだけはたらくとかトラフィック観測と合わせた単機能製品があると、(a)案もじゅうぶん現実味を持ってくるはずだ。

マルチキャスト経路制御に関しては、大学・企業などの単位ではすでに問題なく使える使える段階にあり、実際閉じたネットワークのなかでのサービスとしてはいろいろ活用されているようである。しかし、広域ネットワーク上でのサービスとなると現在もいろいろな議論がなされており、まだまだ枯れた技術とはいえない。PIM-SMが提案されすでに利用されつつあるが、一方でCBTやstatic multicast, simple multicastなどといった提案も活発

に議論されているところである。今後、どのようにして広域プロバイダのネットワークのマルチキャスト化が進んでいくか興味あるところである。

参考文献:

- [1] "MBone", V. Kumar, New Riders Publishing, 1995 (「インターネットマルチキャスト MBone」, 楠本博之 監訳, 1996)
- [2] 「インターネット参加の手引き1996年度版(bit別冊)」, 1996
- [3] "Deploying IP Multicast in the Enterprise", T. Maufer, Prentice-Hall, 1998
- [4] RFC1112: Host extensions for IP multicasting, S. Deering, 1989
- [5] RFC1812: Requirements for IP Version 4 Routers, F. Baker, 1995
- [6] Distance Vector Multicast Routing Protocol (draft-ietf-idmr-dvmrp-v3-08), T. Pusateri, 1999
- [7] Protocol Independent Multicast Version 2 Dense Mode Specification (draft-ietf-pim-v2-dm-01.txt), S. Deering et.al., 1998
- [8] Protocol Independent Multicast-Sparse Mode (PIM-SM): Protocol Specification (RFC2362), D. Estrin, et.al., 1998
- [9] Core Based Trees (CBT) Multicast Routing Architecture(rfc2201), A. Ballardie, 1997
- [10] Core Based Trees (CBT version 3) Multicast Routing - Protocol Specification -, A. Ballardie and et.al., 1998
- [11] Multi Router Traffic Grapher (<http://ee-staff.ethz.ch/oetiker/webtools/mrtg/>)
- [12] JP-MBone Information Page (<http://aohakobe.ipc.chib-u.ac.jp/misc/JP-MBone/>)