

第VIII部

マルチキャスト通信

第 8 部

マルチキャスト通信

第 1 章 はじめに

マルチキャスト通信分科会 (MC-WG) では、JP-MBone コミュニティーと連携し、日本国内における MBone の運用に携わっている。トラフィック監視やさまざまなツールの開発を通じ、その安定運用に寄与してきた。

第 2 章では JP-MBone の現状を述べ、我々が提供している統計情報収集技術とそれによって得られた情報を報告する。

第 3 章では WIDE 大手町 NOC で行っている multicast exchange project について報告する。

第 4 章ではマルチキャストパケット集約機構を提案することで、衛星回線を片方向通信路として含むネットワークにおいて、多数の受信者に対して信頼性を考慮したマルチキャストデータ配信を行う際に、各受信者から重複して送出されるパケットによって大量のトラフィックが発生して衛星回線の帯域資源を無駄に消費する問題点を解決したことを研究報告する。

第 5 章ではマルチキャストネットワークの運用性の向上を目的として、マルチキャスト配送経路を視覚化するシステムを設計・実装したことを研究報告する。

第 2 章 JP MBone

本章は JP-MBone の現状をそのトポロジ、およびトラフィック統計情報を基に記述することを目的とする。

2.1 JP-MBone のトポロジ

図 2.1 に JP-MBone の現在の接続関係を示す。

IMNET、APAN、IIJ、DTI、および WIDE 大手町 NOC はスター状に接続関係が示されているが、これは Multicast IX を示している。その他、特記すべき点としては、WIDE 大手町 NOC、WIDE 京都 NOC、九州大学、名古屋大学、東京大学に渡るループが存在しているということ、また東北大学は学内ネットワークを全て PIM 化しているため、単一組織では国内最大数のマルチキャストルータを有している組織であるという点が挙げられる。

2.2 国内の MBone の現状

まず、日本国内における MBone アクティビティの現状を統計情報によって示す。表 2.1 は、2002 年 5 月 1 日現在の JP-MBone に接続されているルータとそれが所属する組織の数である。

表 2.1 において、etcJP と記されているのは表に出てこない他の*.jp のドメインであり、notJP は*.net 等の*.jp 以外のドメインを持つものである。また、NA は DNS に未登録のルータである。

この調査は、DVMPR_ASK_NEIGHBORS2 の IGMP パケットを用いて、mbone.otemachi.wide.ad.jp から順に接続ルータを辿っていくことによって生成されている。以前は、JP-MBone に所属する組織内では、threshold=64 として、TTL が 64 未満のパケットをフォワーディングしないというポリシーを持って運用されていたため、threshold が 64 のパケットによって到達可能なすべての範囲がすなわち JP-MBone と一致していたが、現状ではこの threshold による制約は外されているため、同様の手法を適用することはできなくなった。

これは、APAN をはじめとする研究プロジェクトによって、以前に比べて海外とのコネクティビティが格段に帯域が増加した結果によるものである。国内、国外という区分けは名義的に存在するに過ぎなくなった。

現在の JP-MBone の海外に対する接続性は、APAN に依存している。表 2.1 における統計情報は APAN 以遠を考慮しないことによって得られている。国内において、APAN に直接接続するサイトもいくつか存在するため、ここでのルータ数の合計

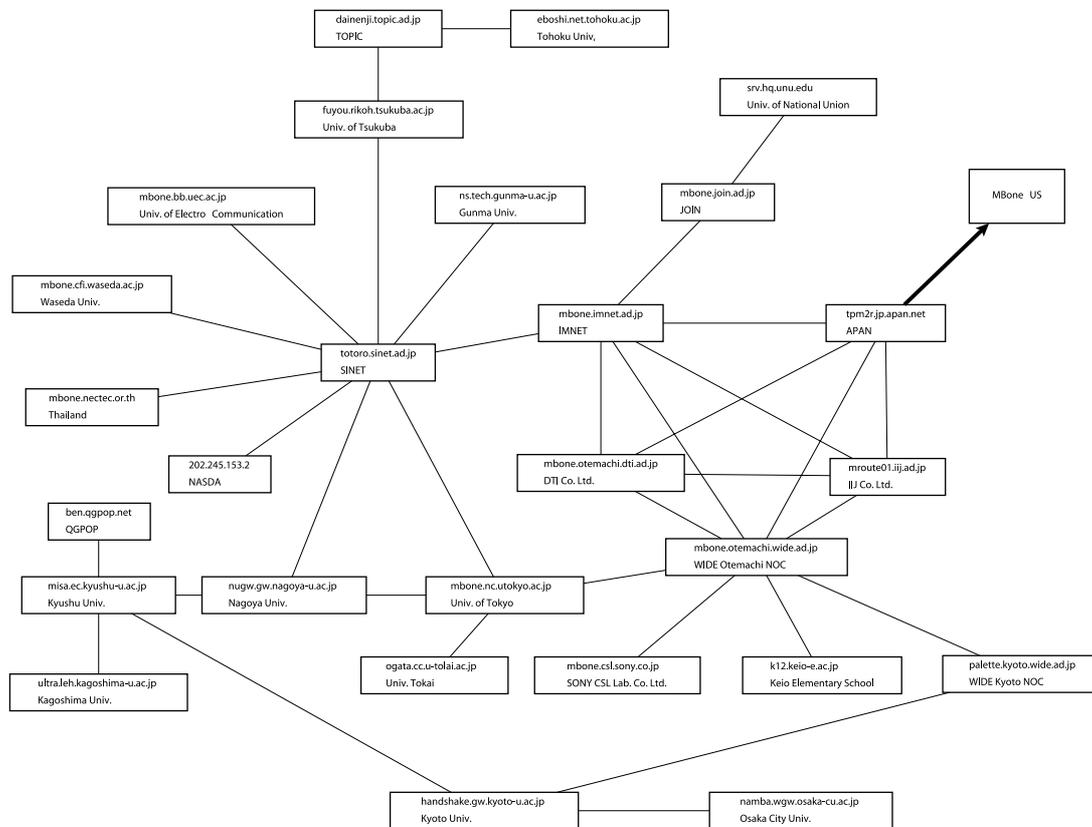


図 2.1. JP-MBone の接続図

表 2.1. JP-MBone におけるルータと接続組織の分布

	合計	ac.jp	ad.jp	co.jp	go.jp	gr.jp	or.jp	ne.jp	etcJP	notJP	NA
組織数	58	42	8	2	0	0	0	0	0	2	3
(active)	21	12	5	1	0	0	0	0	0	1	1
ルータ数	132	42	8	2	0	0	0	0	0	76	3
(active)	21	12	5	1	0	0	0	0	0	2	0

表 2.2. JP-MBone で使用されているルータの種類
の分布

ルータの種類	台数
11.1PM	16
11.2PM	6
11.3PM	2
12.0PM	1
12.1PM	3
12.2PM	1
3.255PGM	23
3.8PGM	5
5.0PGM	1
not active	74
合計	132

は、実際のルータ数よりも若干少なめに見積もったものになっている。

次に JP-MBone で使用されているマルチキャストルータの種類を表 2.2 に示す。

表 2.2 において 11.X と示されているのは CISCO の IOS のバージョンを示す。3.X と示されているのは、mrouted のバージョンである。

現在 JP-MBone において用いられているルータは、CISCO と mrouted がほぼ同数であるということがわかる。ルータ数の合計はここ昨年までの 3 年間ほぼ変動がなかったが、本年は減少傾向にある。

2.3 MBone の状況

MBone 全域において、前章と同様の統計情報を取
得したものを表に整理したものを以下に示す。

表 2.3. MBone に接続されたルータとその分布

ドメイン	ルータ数	ルータ数 (active)	組織数	組織数 (active)
arpa	1	1	1	1
au	20	7	3	3
be	16	11	4	2
br	1	0	1	0
ca	36	17	4	4
ch	82	41	4	4
cn	1	1	1	1
com	469	184	18	11
cz	7	3	2	1
de	6	0	4	0
dk	1	0	1	0
edu	77	35	18	10
es	26	14	2	2
fi	2	1	1	1
gov	60	19	6	4
ie	17	9	1	1
jp	51	22	32	15
kr	9	5	2	1
lu	6	3	2	1
mil	5	3	3	2
net	5264	2309	100	69
nl	1	1	1	1
no	76	38	6	4
org	2	1	2	1
ph	1	1	1	1
pl	7	1	2	1
pt	28	18	3	2
se	305	109	16	12
sg	5	2	1	1
sk	5	0	2	0
th	1	0	1	0
uk	3	2	2	1
us	3	3	2	2
NA	776	112	1	1
合計	7370	2973	250	160

表 2.3 は MBone 全域において接続されているホストを持つトップレベルドメインの一覧が示されている。また、各々のドメインにおける MBone に接続する 2nd レベルドメインの数、ルータの総数を示すものである。

表 2.4 は MBone 全域で使用されているマルチキャ

ストルータの種類分布を示す。

表 2.4 を見てわかるように、圧倒的に CISCO が使用されており、それも IOS バージョン 12.0 および 12.1 であることがわかる。

表 2.4. MBone で使用されているルータの種類
の分布

ルータの種類	台数
0.0G	8
10.2	1
11.0PM	4
11.1PM	26
11.2PM	21
11.3PM	1
11.4PM	2
12.0PM	1910
12.1PM	499
12.2PM	91
3.255PGM	404
3.8PGM	5
5.0PGM	1
NG	4397
合計	7370

2.4 JP-MBone と MBone 間でのトラフィックの
推移

WIDE プロジェクトでは、1997 年より大手町 KDDI NOC において主として ISP 間のマルチキャストトラフィックの IX である Multicast IX を運用している。現状では、JP-MBone とその外部となる MBone とは APAN を経由して接続している。

APAN を経由し、JP-MBone に対して流入してきたマルチキャストのトラフィック量の推移を図 2.2 に示す。これは 2001 年 4 月から 2002 年 4 月までの統計情報である。

図 2.2 をみてわかるように、non-prunning ルータが出現した際には 20~30Mbps 程度のトラフィックが流入している。オレゴン大などを中心にして、恒常的に MPEG2 のストリームを複数(10 程度のチャンネルがある)流しているため、国内においてもま

だコネクションがそれほど太くないサイトに対してはその影響は非常に大きいことがわかる。

第 3 章 Multicast Exchange Project

WIDE プロジェクトでは WIDE 大手町 NOC や NSPIXP-2 がある KDDI 大手町ビル内に、ISP 間のマルチキャストトラフィックを交換するために Multicast Exchange を 1997 年より設置して運用している。

2001 年 12 月に横浜で行われた Internet Week 2001 内での MBone BoF において、JP-Mbone を DVMRP から PIM-SM ベースへ移行することが決定され、Multicast Exchange の部分から順に移行していくことになった。

本章では、Multicast Exchange Project の経緯と現状を述べるとともに、JP-Mbone をどのように移行するのかについての構想を報告する。

3.1 phase I

1996 年秋に NSPIXP-2 が始まったことや MBone へ接続を希望する商用 ISP が当時増えつつあった状況を考慮し、WIDE 及び JP MBone の接続トポロジを整理するために、1997 年 3 月に WIDE 大手町 NOC に mbone.otemachi.wide.ad.jp を設置してトポロジの整理及び NSPIXP2 を通しての ISP からの接続を受け入れていった。これ以前の状況では、WIDE 大手町 NOC にはマルチキャストルータが存在していなく、この設置によりトンネル受け入れも増加して十数本のトンネルを抱えることになった。

これにより JP MBone 接続トポロジの整理をすることができたが、図 3.1 のように多数のトンネル接

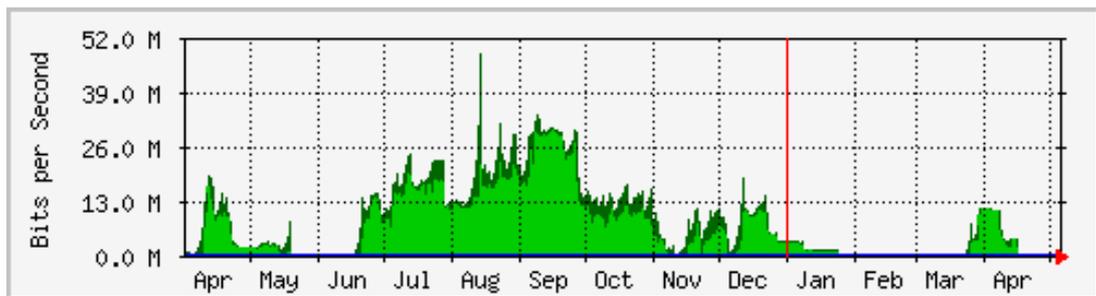


図 2.2. APAN より Multicast IX に流入したトラフィック

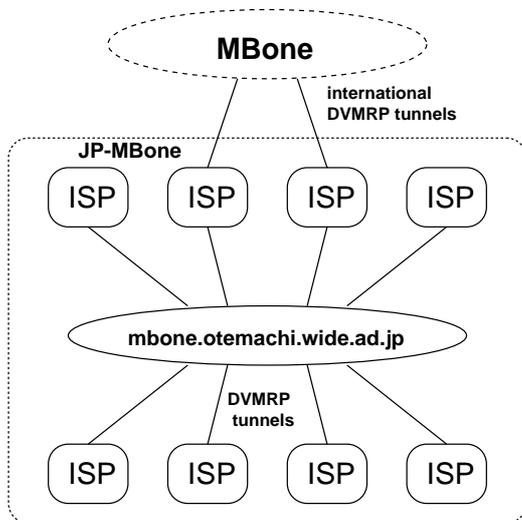


図 3.1. Multicast Exchange Project - phase I

続を抱える `mbone.otemachi.wide.ad.jp` を中心とした放射型の接続トポロジとなっているため、次のような構造上の基本的な問題があった。

- 各 ISP 同士のやりとりが必ず WIDE 経由の間接やりとりになっている。
- 常に `mbone.otemachi.wide.ad.jp` を通過するため、そこが過負荷になる。
- もしも `mbone.otemachi.wide.ad.jp` が落ちると、すべて切断されてしまう。
- すべての接続がトンネル利用であり、また、`mbone.otemachi.wide.ad.jp` はユニキャスト的には各 ISP に対してルータでないため、すべてのトンネリングの packets が出入り口である 1 つの物理インタフェースに集中し、トンネル数が増えるとともに比例する形で通過するトラフィックが増えてしまう。

3.2 phase II

phase I における問題を解決するため、`mbone.otemachi.wide.ad.jp` が設置されている WIDE 大手町 NOC の隣にマルチアクセスのメディアのセグメントを一つ新たに用意し、各 ISP のマルチキャストルータをそのセグメントにすべて接続するという図 3.2 へと進めることになった。

具体的には 100BaseTX のセグメントを新規に設置し、このセグメントにおいて複数のマルチキャストルータが直接マルチキャスト packets をやりとりするようにする。これは、物理層レベルでのマルチキャストを使用することで複数のルータへ渡す時のパケッ

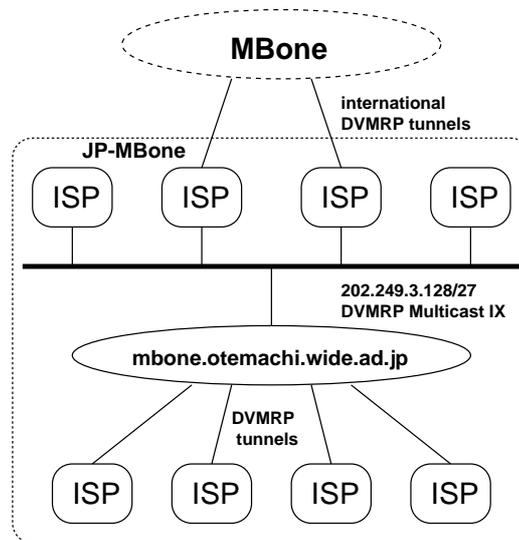


図 3.2. Multicast Exchange Project - phase II

トの複製を避け、トラフィックの削減、ルータ負荷の低減を狙ったものである。このときに 100BaseTX を採用した理由は、多くの ISP のマルチキャストルータが接続して本格的利用が始まると 10Base ではすぐに飽和する可能性があること、および、当時の状況で 100BaseTX が最も安価で適切と判断されたためである。

1998 年 3 月にこの計画を提案した結果、当時 wide、apan、ij、dti、ttnet、ibmnet、imnet、jpix、sinfony、mesh、spin、psi、kddlabs、tokyonet、sphere の 15 の組織が参加表明し、1998 年 7 月より実際の接続を開始した。具体的には WIDE 大手町 NOC に Fast Ethernet のノンスイッチングハブを設け、ちょうど同じ KDD 大手町ビル内に NSPIX-2 があることにより、参加 ISP は各自の保有する場所にある自組織のマルチキャストルータの 1 インタフェースを延ばしてそこに接続する、という形をとった。このセグメントには 202.249.3.128/27 が割り当てられており、現時点での状況を表 3.1 に示す。

3.3 phase III

その後、ドメイン間におけるマルチキャスト経路制御として PIM-SM/MBGP/MSDP の組み合わせが標準になっていった。JP MBone 内でも移行する話が出たが、多用されている PC 系ルータでの対応問題があり、一方、国際接続では早く PIM-SM/MBGP/MSDP に移行することが求められたため、まずは国際接続部分のみ PIM-SM/MBGP/MSDP

表 3.1. Multicast Exchange 参加表明組織の状況

IP アドレス	組織略称名	AS 番号	接続ルータ名
202.249.3.129	wide	2500	mbone.otemachi.wide.ad.jp
202.249.3.130	apan	7660	tpm2r.jp.apan.net
202.249.3.131	iij	2497	mroute01.iij.ad.jp
202.249.3.132	dti	4691	mbone.otemachi.dti.ad.jp
202.249.3.133	ttnet	4716	(tunnel 接続後停止中)
202.249.3.134	ibmnet	4672	(tunnel 接続後停止中)
202.249.3.135	imnet	2513	mbone.imnet.ad.jp
202.249.3.136	jpix	7527	(未接続)
202.249.3.137	sinfony	2527	(tunnel 接続後停止中)
202.249.3.138	mesh	2518	(tunnel 接続後停止中)
202.249.3.139	spin	2915	(未接続)
202.249.3.140	psi	2554	(未接続)
202.249.3.141	kddlabs	7667	(APAN 経由で接続)
202.249.3.142	tokyonet	2521	(tunnel 接続後停止中)
202.249.3.143	sphere	2514	(tunnel 接続後停止中)

へと移行することになった。

その時より現在も国際接続を APAN にお願いする形となっているが、1999 年 12 月にまずは PIM-SM/MBGP/MSDP で対外接続している APAN の内部経路として JP MBone 全体の経路も同様に AS7660 として広告される形で接続した。その後、経路制御上 APAN の内部経路として扱うのは問題があるため、2000 年の 5 月に JP MBone として AS9616 を取得し、図 3.3 に示すように、APAN の

下流 AS として接続する形へ変更された。

3.4 phase IV

phase III では Multicast Exchange において経路制御プロトコルとしては従来と同じく DVMRP を用いているが、DVMRP の利用を続けていくには以下のような問題がある。

- 経路情報交換が安定的に動作しにくい
- 経路数増大に対する耐性の欠如
- 政策的経路制御の困難性
- プロトコルおよび実装上の問題に起因するトラブル (Flooding 等) の多発

特に ISP 間の経路情報交換においてはユニキャストの時と同様に経路選択の自由度を高めるために政策的経路制御ができることが好ましい。このため、Multicast Exchange 含む JP MBone 全体の接続においても ISP 間のマルチキャスト経路情報交換について脱 DVMRP の体制を構築することが求められる。2001 年 12 月に横浜で行われた Internet Week 2001 の MBone BoF において、JP MBone 全体を PIM-SM/MBGP/MSDP 体制へと移行することが決定された。

まずは最初に、図 3.4 に示すような形で、Multicast Exchange 部分を現在の DVMRP から PIM-SM/MBGP/MSDP へと変更する計画になっている。この時点で、Multicast Exchange 部分は、JP MBone

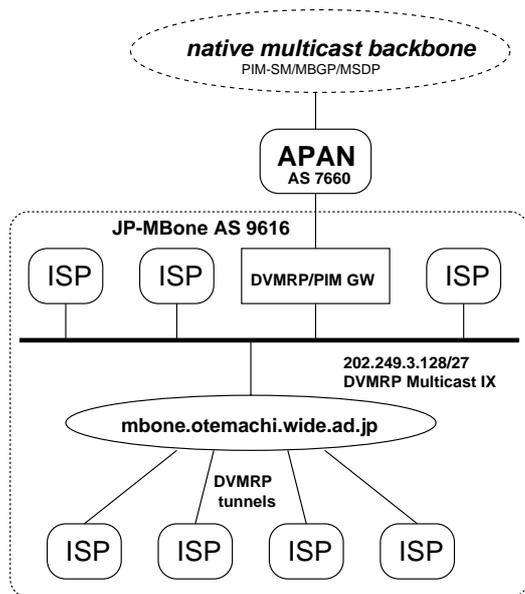


図 3.3. Multicast Exchange Project - phase III

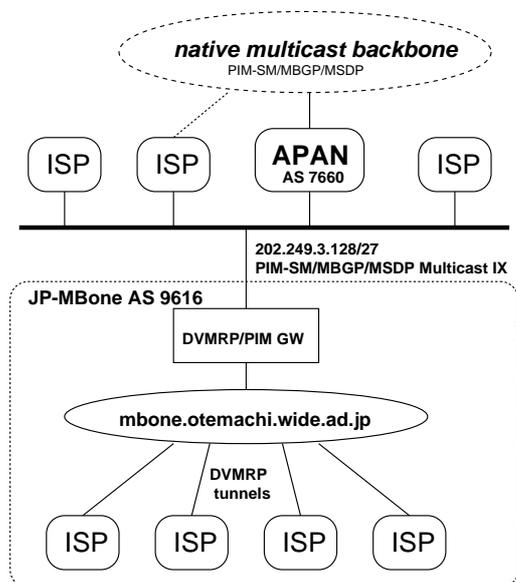


図 3.4. Multicast Exchange Project - phase IV

の AS9616 からははずれ、各接続 ISP などが各自の AS を用いて MBGP によって経路交換をすることになる。一方、DVMRP で構成されている JP MBone 部分については、今までと同様に AS9616 として接続されることになる。

第 4 章 マルチキャストパケット集約

4.1 概要

インターネットにおいて、多数の利用者に対して同一のデータを効率的に配信するための技術として、マルチキャストが広く研究されている。広域同報型の通信路である衛星回線は、1 対多型の通信方式であるマルチキャストとの親和性が高く、衛星回線を利用することにより、広域ネットワーク上に散在する多数の受信者に対して効率的にデータを配信できるインフラとして注目されている。片方向の衛星回線をインターネットのインフラとして利用する技術として UDLR 技術が標準化され、廉価な衛星受信専用装置を用いて衛星回線を利用できる環境が整いつつある。

しかし、UDLR と衛星回線を組み合わせたネットワークにおいて、非常に多数の受信者を対象としてマルチキャストの通信を行う際には、受信者数の規

模性に対応できない問題が生じる。

本研究では、衛星回線を片方向通信路として含むネットワークにおいて、多数の受信者に対して信頼性を考慮したマルチキャストデータ配信を行う際に、各受信者から重複して送出されるパケットによって大量のトラフィックが発生し、衛星回線の帯域資源を無駄に消費する問題点を解決した。

4.2 マルチキャストネットワークにおける問題点

衛星回線を片方向通信路として含む UDLR を用いたネットワークでは、送信ホスト、および各受信ホストは、仮想的に実現する双方向の通信が可能な同一セグメント上に接続する。しかし UDLR では、ネットワークの規模性の問題を議論の対象としていない。ネットワークの双方向性を Broadcast Emulation によって確保するため、Receiver からのトラフィックは、Feed によって衛星回線にそのまま転送される。受信者数が増大した際に、衛星回線の帯域消費も増大することは避けられない。

また、衛星回線を含むネットワークにおいて IGMP を用いた場合、衛星回線が高遅延な通信路であるため、IGMP の送出遅延によるパケット抑制機構が完全には動作しない。マルチキャストルータからの Membership Query を受信したホスト群のうち、最初にタイムアウトしたホストから送信された Membership Report が Feed を経由し、衛星回線を用いてネットワーク上の全ホストに到達する。衛星回線には受信のみ可能なホストが多数接続することが想定されるので、既存のパケット抑制機構によって設定される遅延タイムの多くが、他のホストが送信した Membership Report が到達する前にタイムアウトする。このため、重複した Membership Report が多数送信され、UDLR の Broadcast Emulation により、そのまま衛星回線に転送される。このため、重複した内容のパケットが多数、衛星回線に転送される。必要以上に送出される IGMP や NAK パケットが無駄なトラフィックとして衛星回線に転送される際の負荷が大きく、これを無視できない。

衛星回線が多くの受信者によって共有できる通信路であり、1 対多型の通信方式であるマルチキャストとの親和性が高い。しかし、実用的なサービスを行うには、既存の機構ではこのように抑制しきれなかったトラフィックを如何に扱うかという問題を解決する必要がある。

4.3 解決手法の提案

パケット集約機構 (Packet Aggregator: 以降、PA と記述) は、不必要なパケットが重複して衛星回線に転送されるのを防ぎ、通信路の帯域資源を有効に使用するための機構である。ある特定のプロトコル、またはホストに宛てられたパケットを捕捉し、定められたポリシーによって、それぞれ衛星回線に転送するかを選択する。これにより、衛星回線の帯域資源の無駄な消費を抑制する。

複数送出されるパケットの内容が同一であり、必ずしも全てのホストが送出する必要がないパケットであれば、これらを一に集約できる。これを実現するため機構として、PA を提案した。IGMP や NAK パケットはその内容が重複し、必ずしも全てのホストがパケットを送出する必要があるわけではない点に注目する。重複した IGMP や NAK パケットが衛星回線に転送されると、それらが無駄なトラフィックとなって帯域資源を消費する。このとき PA は、Receiver から UDL に向けて送出されたパケットのうち自分が Feed に転送したパケットの内容を記憶する。そして、IGMP と NAK パケットそれぞれについて、既に UDL に転送されたものと内容を比較し集約する。これにより、IGMP と NAK パケットが衛星回線に転送される総量を必要最低限に抑制できる。PA は内容の重複するパケットを受信した際には、これを破棄する。これにより、各 Receiver から送出される重複パケットによって衛星回線の帯域消費が増大する問題を解決できる。各 Receiver から送出されるパケットのうち、パケット抑制機構が機能するためのトリガーとなるものと、そうでないものがある。IGMP では、マルチキャストルータからの問い合わせである Membership Query を受信すると、ランダムで設定されたタイマが直ちにタイムアウトする Receiver がある。当該 Receiver から送出されるパケットが、他の受信者に到達することでパケット抑制機構が機能し始める。

Membership Query を受信してから直ちに送出された Membership Report が、他の Receiver に到達するまでの間に順次タイムアウトする多くのホストからそれぞれが無駄なパケットが送出される。これらのパケットのうち、パケットの抑制機構が機能し始めるのに必要なパケットのみ Feed によって衛星回線に転送されるのが望ましい。これを図 4.1、および

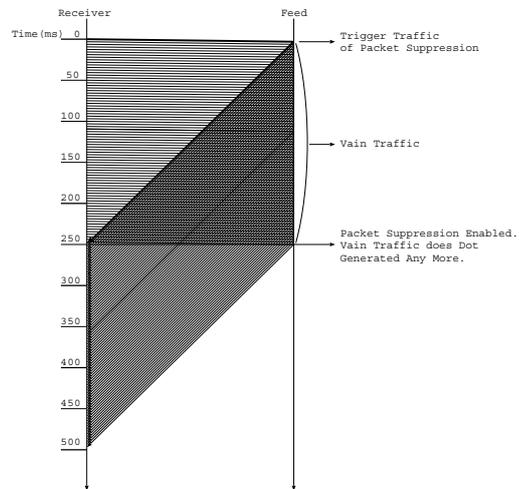


図 4.1. 抑制機構が機能し始めるまでに発生するトラフィック

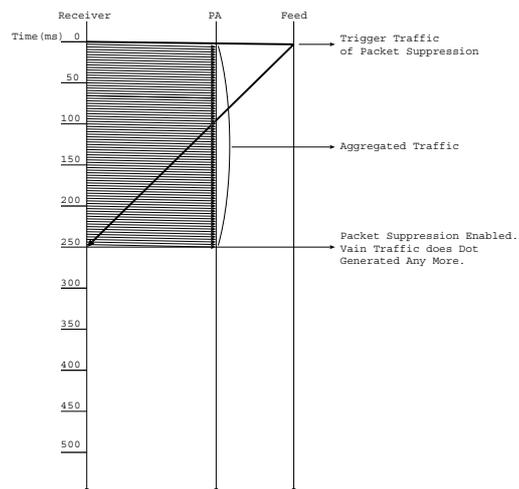


図 4.2. PA によって集約されるトラフィック

び図 4.2 に示す。

NAK パケットについても IGMP と同様に、パケット抑制機構が機能し始めるトリガーとして必要なパケットが Feed によって衛星回線に転送されるのが望ましい。

4.4 PA の構成

PA は、自分が Feed に転送した IGMP、および NAK パケットの内容を、パケットの種類ごとに記録し、保持する。このための媒体をリストと呼びパケットの内容をリストに格納するための媒体をエントリと呼ぶ。パケットのアグリゲートは、PA が保持するリストのエントリと受信するパケットの内容を比較することで行う。パケットの集約は以下に示す処理を経る。

1. PA は Receiver から Feed に向けたパケットを全て捕捉する
2. PA は捕捉したパケットが集約の対象である場合、自分の持つリストを参照し、同様のパケットを既に Feed に転送したかを確認する
 - 自分が既に同一のパケットを Feed に転送していれば、受信したパケットを直ちに破棄する
 - PA が持つリストに無いパケットを受信した場合には、当該パケットを Feed に転送し、パケットの内容をリストに加える
3. PA は保持するリストの内容について、一定期間が過ぎると、これを破棄する

4.5 評価

PA の評価を行うために、以下の図 4.3 に示すネットワークを構築した。M-Router は 2 つのインターフェースを持ち、Network A と Network B とを接続する。M-Router ではマルチキャスト経路制御プロトコルとして DVMRP を動作させる。本トポロジでは、M-Router は Feed を兼ねる。従って、M-Router に到達するパケットは、実ネットワーク上では UDL に転送されることになる。

Network A に接続する M-Client1 と Network B に接続する M-Client2 は、それぞれマルチキャストグループ 233.0.0.1 に参加する。Network A、B に接続する M-Router は、M-Client1 と M-Client2 からの Membership Report を受信する。このとき、M-Router のもつマルチキャストルーティングテーブルには、それぞれのリンクでメンバーが存在するマル

チキャストグループが登録され、当該マルチキャストグループの状態管理を開始する。

4.5.1 PA によるネットワークのスループットへの影響の測定

ネットワークに PA を設置した場合と、そうでない場合のネットワークのスループットについて、FTP によるファイル転送を行ってスループットを測定した。

ネットワークに PA を設置せず、M-Router と M-Client2 を直接接続した。3.5 MByte のファイルを転送し、FTP クライアントが表示するスループットを記録した。その後、ネットワークに PA を設置し、パケットのアグリゲートを行わない状態で同様に測定した。FTP によるデータ転送は、PA の設置前後でそれぞれ 20 回ずつ行った。

PA を設置しない状態でのファイル転送を行った際のスループットの平均は 942.1 Kbps であった。PA を設置し、アグリゲートを行わない状態でファイル転送を行った際のスループットの平均は 937.6 Kbps であった。この結果、PA を設置した状態でのスループットは、設置しない状態でのスループットに比べ、およそ 4 Kbps ほど低下し、その割合はおよそ 0.5% であった。このことから、本実装による PA を設置しても、パケット転送性能にはそれほど影響を与えないことが確認できた。

4.5.2 PA が転送したパケット数の測定

PA でアグリゲーションを行い、PA が転送したパケット数を測定した。測定は 10 回行った。測定結果から、PA が捕捉したパケット総数の平均はおよそ 859 であった。このうち、PA によって転送され

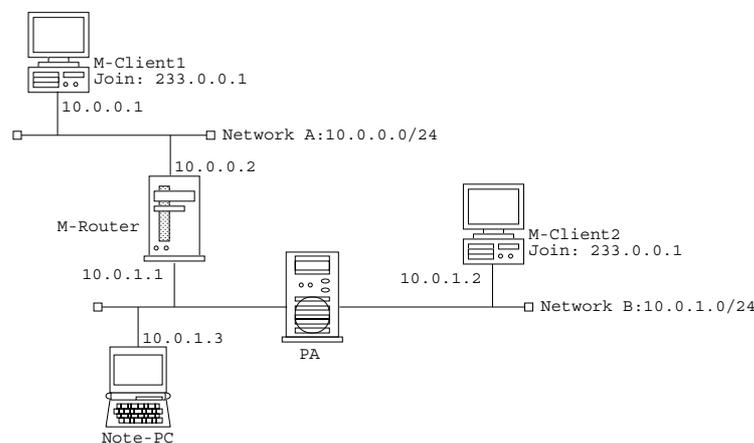


図 4.3. 評価用ネットワーク

たパケット総数の平均はおよそ 6 であった。

また、RP の設定には変更を加えず、PA の設定についてパケットのアグリゲート期間を 2000 ms、10000 ms にそれぞれ変更し、10 回ずつ PA によって転送されるパケット数の測定を行った。その結果、最初に行った測定結果とほぼ同様の結果を得た。

4.5.3 限界性能の測定

PA の限界性能について、PR が送出する重複パケットの数を調節することで測定した。その結果、0.1 秒間におよそ 3800 個のパケットを処理できた。0.1 秒間に集中するパケット数をおよそ 3800 個から 3900 個となるよう負荷を増大させたとき、PA はコアダンプをとともう異常終了が発生した。

4.6 測定結果の考察

1 単位時間あたりに同一のマルチキャスト IP アドレス宛に送出された IGMP パケット数 (p) による帯域消費 (c) は以下の式によって表せる。

$$\begin{aligned} c(\text{Byte}) &= p \times 44 \\ &= 850 \times 44 \\ &= 37400 \\ &\neq 37.4 (\text{KByte}) \end{aligned}$$

このとき、3 単位時間 (t) あたりに 1 つのマルチキャスト IP アドレスに対して送出された Membership Report による帯域消費の合計 ($dc1$) は以下の式で表せる。

$$\begin{aligned} dc(\text{Byte}) &= tc \\ &= 3 \times 850 \times 44 \\ &= 112200 \\ &\neq 112.2 (\text{KByte}) \end{aligned}$$

また、dc からこのマルチキャストグループに参加していると想定されるホスト数 (n) を算出すると、

$$\begin{aligned} dc(\text{Byte}) &= 3 \times \frac{n}{100} \times 44 \\ 3 \times \frac{n}{100} \times 44 &= 112200 \\ n &= 85000 \end{aligned}$$

となる。1 単位時間に PA が M-Router に転送したパ

ケットのうち、集約対象となっていたパケット (pa) が衛星回線の帯域消費 ($dc2$) は

$$\begin{aligned} dc(\text{Byte}) &= 4 \times 44 \\ &= 176 \\ &\neq 0.2 (\text{KByte}) \end{aligned}$$

であるため、実際に送出された Membership Report による帯域消費 ($dc1$) を 1% 以下に集約できる。

4.6.1 まとめ

本研究で開発した PA を用いることにより、不要なマルチキャストパケットが UDL に転送される問題を解決し、多数の受信者が接続する衛星回線を、インターネットの通信路の一部として用いた大規模なマルチキャストデータ配信が行えることを示した。

しかし、大量のパケットが送出された際に、PA によって M-Router に重複したパケットが認められたことより、本実装においては PA が完全に機能しているとは言えない。実装面での改良は今後の課題である。

第 5 章 マルチキャスト経路の視覚化

5.1 概要

マルチキャストネットワークの運用性を高めるためには、運用を支援する仕組みが必要であるが現状では、これが十分であるとはいえない。マルチキャストネットワークの運用には、配送木全体を監視する仕組みが有用である。マルチキャストの経路は、経路の断絶・迂回、グループへの参加・離脱などにより、動的に変動するからである。

本研究では、マルチキャストネットワークの運用性の向上を目的として、マルチキャスト配送経路を視覚化するシステムを設計・実装した。具体的には、マルチキャストグループに加入している複数のルータにおける経路情報、ステータスを収集し、それらを解析して図示することによって、配送経路を図示できるようにした。また、本研究で開発したシステムをテストベッドネットワークにおいて評価し、その有効性を確認した。

表 5.1. マルチキャストツール機能比較表

Tool	データ対象	監視対象	監視時間	出力方法
MantaRay	他のツール	トラフィック量	online	GUI
Mantra	ルータ	トラフィック量 プロトコルメッセージ セッション	online historic	ファイル GUI
Mlisten	アプリケーション	セッション	online historic	ファイル GUI
Mtrace	ルータ	送信者からある受信者への 配送経路	online	テキスト
Mrinfo	ルータ SNMP	ルータの設定	online	テキスト
MultiMon	ルータ アプリケーション	トラフィック量	online historic	GUI
RM	ルータ 他のツール	トラフィック量	online historic	ファイル

5.2 現状のマルチキャスト運用支援ツールに関する考察

マルチキャストネットワークの運用には、配送木全体を監視する仕組みが有用である。その理由は、ネットワーク内の経路が経路迂回・断絶のみならず、グループへの参加・離脱などによって生じるからである。

以下の表 5.1 に既存のマルチキャスト運用支援ツールの機能をまとめた。以下にあげたツールはネットワークの一部を監視するものが主である。これらのツールを使用してもネットワークの局所的な問題しか発見できず、ネットワーク全体の障害に対して適切な処理を行えない。障害に適切に対処するためには、これらのツールに加えてマルチキャスト経路制御プロトコルの動作に注目し、ネットワーク全体でのトラフィックの配送経路を把握できるシステムが必要となる。マルチキャストルータは一つのインターフェイスからパケットを受信し、複数のインターフェイスからパケットを送信する。そのため、一般に管理対象となる経路は複数になる。

また、これらのツールは、mrouted/DVMRP のみに対応したものが多く、現在主流になりつつある PIM-SM などの新しい経路制御プロトコルに対して適用できない。そして現在のところ、IPv6 マルチキャストを支援するシステムは存在しない。

5.3 設計

本システムは経路制御プロトコルにより作られた

配送木の構造を把握し、マルチキャストパケットの配送経路を表示する。そのため、以下の機能別にモジュールを設計した。

- ルータの情報収集
 - ルータ内部情報収集モジュール
- 配送経路の表示
 - 配送経路データベースモジュール
- 描画した配送経路図の管理
 - 配送経路表示モジュール

本設計では、各モジュール間は命令メッセージおよび更新メッセージの 2 つの通信メッセージを用いて、各モジュールのパラメータを送信し合う。

5.3.1 ルータ内部情報収集モジュール

ルータ内部情報収集モジュールは以下の機能を持つ。

- マルチキャストルータの監視
 - ルータの経路情報、ルーティングデーモンの状態、PIM-SM 内部情報を監視する。
- 配送経路データベースへのルータ情報の格納

本システムでは、PIM-SM ルータの内部情報を取得し、Register メッセージを送信・受信しているインターフェイスの IP アドレスを明示するようにした。

図 5.1 は Register メッセージの流れの例である。矢印はマルチキャストパケットの流れを示す。点線は DR から RP へ転送される Register メッセージを意味する。PIM-SM では、PIM Register メッセージをユニキャストで送信しているため、マルチキャストパケットの転送経路を図示しただけではネットワー

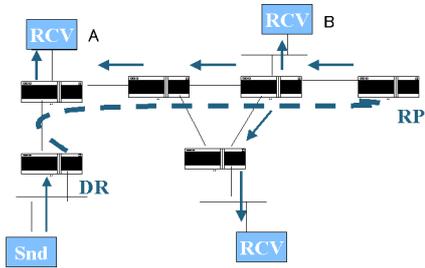


図 5.1. PIM-SM での Register メッセージの例

ク上でマルチキャストパケットの転送経路や、経路制御プロトコルの動作全てを表示できない。Register メッセージの経路も把握する必要がある。

5.3.2 配送経路データベースモジュール

配送経路情報データベースモジュールは、最新の経路情報を格納する経路情報データベースと、過去に描いた配送経路図を格納する配送経路図データベースとの 2 つのデータベースの管理を行う。

経路情報データベースは、管理するルータの最新の経路情報を記録するようにした。各ルータで監視するセッションの経路に変化があった場合、ルータ内部情報収集モジュールは更新メッセージを送信する。経路情報データベースは、更新メッセージが送られてくると、この更新メッセージの管理番号を調べ、該当する ID を持つルータが送受信を行うインターフェイスの IP アドレスを変更する。その結果、データベースには、各ルータの最新の経路情報のみが存在することになる。

配送経路図データベースは、配送経路図表示モジュールが描いた配送経路図を、管理番号を付与して管理する。

システム全体における配送経路データベースモジュールの概観を図 5.2 に示す。

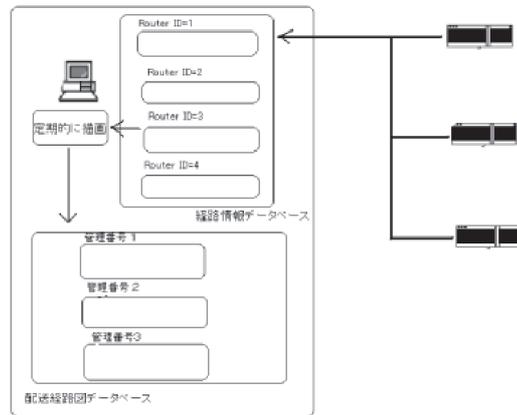


図 5.2. 配送経路データベースモジュール

5.3.3 配送経路表示モジュール

配送経路表示モジュールでは以下項目を含む配送経路図を表示する。

- マルチキャストルータの経路の表示
 - 本システムでは、ルータの機能を「四角」で抽象化し、受信するインターフェイスと送信するインターフェイスを「矢印」で図示する。これにより、配送経路を視覚化する。「矢印」はパケットが受信されるインターフェイスの場合、ルータの方向を指しており、パケットが送信される場合はルータから外部を指す。
- PIM-SM 内部情報の表示
 - PIM-SM のルータより Register メッセージを送受信しているインターフェイスのアドレスを取得し、PIM Register メッセージの流れを、マルチキャストパケットの転送と同様に「矢印」で表示する。また、Register メッセージを送信しているルータを DR と表示する。Register メッセージを受信しているルータは RP と表示する。
- ルーティングデーモンの状態
 - 監視しているルーティングデーモンが起動していない場合のみ、“down” と表示する。
- 時刻
 - 配送経路表示モジュールが配送経路図を作成する時刻を配送経路図に書きこむ。

配送経路図は定期的に描画され、配送経路図データベースに格納される。ネットワーク管理者は、Web インターフェイス (CGI) を用いて、格納された配送経路図を管理番号、時刻等により選択、表示する。

5.4 評価

本研究で開発したシステムを PC をルータとするテストベッドネットワークにおいて評価を実施した。

評価ネットワークは以下のような事項を満足するように構築した。

- マルチキャストパケットの送信者、受信者がそれぞれ一つずつ存在する。
- IPv6 ユニキャストの経路制御は動的に行う (ripng を利用)。
- 送信者、受信者が位置するサブネットワークは複数のルータが存在するブロードキャスト型のリンクである。
- PIM-SM でマルチキャスト経路制御を行う。

本評価ネットワークを用いて実際に作成した配送経路図を図 5.3 に示す。本システムを用いることにより、ネットワーク管理者が視覚的に理解しやすい配送経路図が構築できていることが確認された。

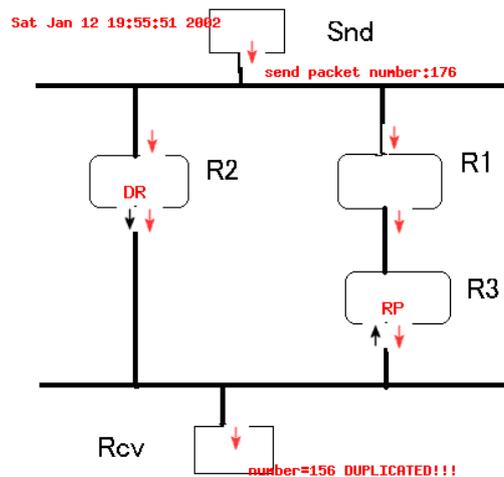


図 5.3. 作成した配送経路図の例

5.5 まとめ

本システムにより、マルチキャストパケットの配送経路経路の変遷を視覚化し、マルチキャストネットワーク管理者が、ネットワーク内におけるマルチキャストパケット転送状況を視覚的に把握できた。今後、より大規模なネットワークでの評価を実施し、実運用を通じた評価ならびに改良を行う予定である。

