IP マルチキャスト技術

藤井 直人

（(株)アイアイジェイメディアコミュニケーションズ）

1999 年 12 月 15 日

Internet Week 99 パシフィコ横浜

（社）日本ネットワークインフォメーションセンター編

目次

1　概要

IP マルチキャストは、インターネット上で大量のデータを効率的に配信するための基礎技術です。特に大容量になりがちな動画や音声などのマルチメディアデータを混雑なく配信するための効果が期待されています。この講演では、商用実験が始まった IP マルチキャストについて、基礎技術と現在利用されているプロトコル、IP マルチキャストを用いるアプリケーションを含めて説明します。

2　IP マルチキャストプロトコル

ここでは、IP マルチキャストのプロトコルと基盤技術について説明します。

2.1　IP マルチキャスト概略

通常のインターネットで使われるユニキャストは 1 対 1 の通信です。このため、図 1 のユニキャストの 64Kbps のストリーミングの例のように、4 つのクライアントへストリーミングデータを配信する場合、同じ内容のパケットが 4 つ配信されることになります。このとき、サーバ側では 256Kbps の帯域が消費されます。1000 人が 64Kbps のストリーミングコンテンツを見る場合を想定すると、サーバ側では 64Mbps の帯域が必要になるだけでなく、マシンに対する負荷も増大します。

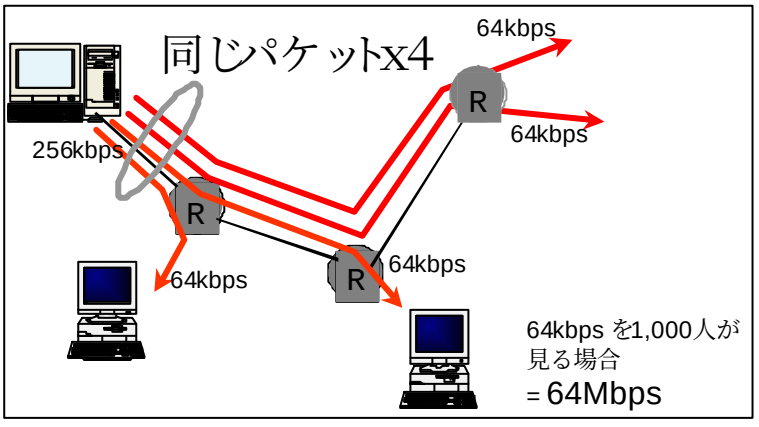

図 1： 64Kbps のストリーミング（ユニキャスト）

これに対し、マルチキャストが優れている点は、1 つのソースから出されたパケットを受信者全員が見ることができるという点です。ルータがその先に受信者がいると判断してパケットをディプリケートしていくという方法です。64Kbps のストリームデータを全部が受信してもサーバ側の帯域は 64Kbps で済みます。

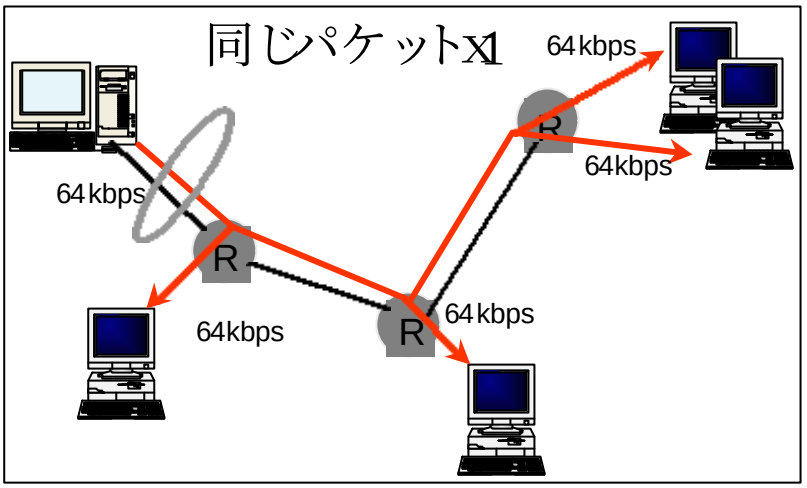

図 2：64Kbps のストリーミング（マルチキャスト）

マルチキャストがブロードキャストと異なる点は、ルータの先にデータを受信するクライアントがない場合には、パケットを転送しないという機能がある点です。

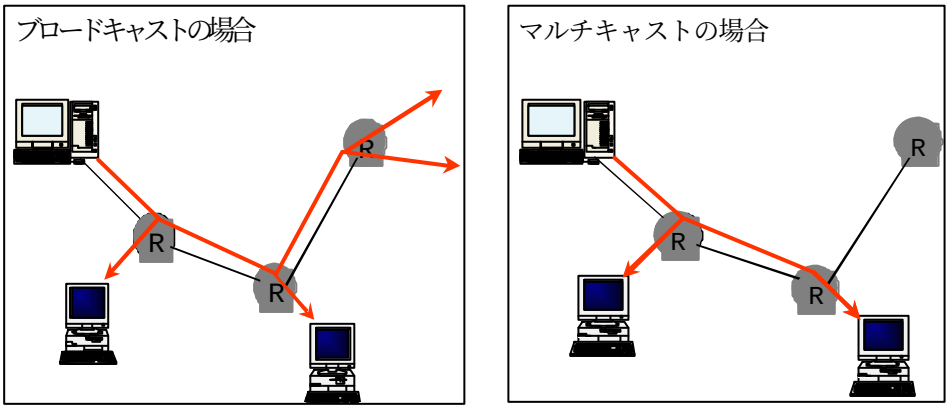

図 3：ブロードキャストとの違い

2.2　マルチキャストアドレス

ユニキャストの場合は、ホストの IP アドレスで送信先を決めました。

これに対しマルチキャストは、マルチキャストアドレスというグループに対して送信します。受信側は、受信したいデータのグループのアドレスを要求する仕組みです。

2.2.1 IPv4 のマルチキャストアドレス

IPv4（Internet Protocol Version4)の場合、マルチキャストアドレスは一般的に Class D と呼ばれている 224.0.0.0～239.255.255.255 の範囲を利用します。

ただ、そのアドレスの範囲が全部自由に利用できるわけではありません。このアドレス範囲には利用方法が決められています。そのアドレスでのパケットの到達範囲を表す Scope は、RFC2365 (Administratively Scoped IP Multicast)で示してあります。

224.0.1.0～228.255.255.255 の範囲は、Global Scope に割り当てています。パケットの到達範囲は全世界という意味です。

239.192.0.0/14 は Organization-local Scope に割り当てています。パケットの到達範囲は組織内という意味です。239.255.0.0/16 は Local Scope に割り当てています。パケットの到達範囲はルータを必要としないローカルのセグメントだけという意味です。

また、固定的なアドレスの管理は IANA（Internet Assigned Numbers Authority)が行っています。詳細は、以下の FTP サイトで入手できます。

ftp://ftp.isi.edu/in-notes/iana/assignments/multicast-addresses

224.0.0.1 は、ALL-SYSTEMS.MCAST.NET です。すべてのマルチキャストの機器はこのグループのアドレスを受信しなければならない約束になっています。

224.0.0.2 は、ALL-ROUTERS.MCAST.NET です。すべてのマルチキャスト対応のルータはこのグループのアドレスを受信しなければなりません。

このほか、どういう目的に使うかを固定的に決められているアドレスもあります。たとえば、マイクロソフトと MSNBC が 224.0.12.0/26、ウォールトディズニー社が 224.0.19.0/26 などの特定の組織から一定の範囲が予約されているアドレススペースもあります。

2.2.2 IPv6 のマルチキャストアドレス

IPv6（Internet Protocol Version6)の場合はパケットの到達範囲を表す Scope もアドレスの中に入っているのが特徴です。アドレスと Scope を組み合わせてアドレスの到達範囲と目的を表します。これらの、IPv6 のマルチキャストアドレスについては RFC2373 と RFC2375 で述べられています。

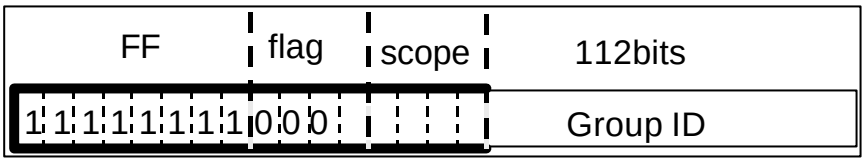

図 4：IPv6 のマルチキャストアドレス

IPv6 のアドレス場合は、先頭の 8 ビットが FF で始まるものをマルチキャストとして使うことになっています。

次の 4 ビットはフラグです。最下位ビットが 0 のときにはアドレスは固定的に、1 のときには動的に割り当てられます。動的に割り当てられるとは、いつかは返さなければならないようなアドレスだという意味です。

さらに次の 4 ビットが Scope です。そのアドレスで到達する範囲を指定しています。Scope が 1 の場合は node-local scope です。そのホストの中でのみ有効という意味です。2 の場合は link-local scope でたとえばイーサネットというセグメントの上でだけ有効です。5 の場合は site-local scope、8 の場合は organization-local scope で、これらはともに論理的な意味での組織に対し有効という意味です。E の場合は global scope で全世界に有効という意味になります。

たとえば、FF02:0:0:0:0:0:0:1 は、マルチキャストアドレス(FF)、固定的な割り当て(0)、link-local scope(2)、All Nodes Address（0:0:0:0:0:0:1）で構成され、同じリンク上にあるすべてのノードに対してこのグループアドレスは有効という意味です。

また、FF02:0:0:0:0:0:0:D は、マルチキャストアドレス(FF)、固定的な割り当て(0)、link-local scope(2)、All PIM Routers（0:0:0:0:0:0:D）で構成され、同じリンク上にあるすべての PIM（Protocol-Independent Multicast）ルータに対してこのグループアドレスは有効という意味です。

また、FF05:0:0:0:0:0:1:3 は、マルチキャストアドレス(FF)、固定的な割り当て(0)、site-local scope(5)、All-dhcp-servers（0:0:0:0:0:1:3）で構成され、自組織内のすべての DHCP(Dynamic Host Configuration Protocol)サーバに対してこのグループアドレスは有効という意味です。

2.2.3 到達範囲の制御

IPv4 で Scope がパケットの到達範囲を制御する方法は 2 つあります。TTL(Time To Live：生存可能時間）を用いる方法とアドレスを用いる方法です。

マルチキャストのパケットの場合も、ユニキャストと同様に TTL によってパケットがネットワーク上にどのくらい存在できるかの生存可能時間を決めることができます。TTL はルータを通過するごとに最初に設定された値から 1 ずつ減らされてきます。この値が 0 になるとそのパケッ

トは転送されません。ユニキャストの場合は、パケットの無限ループを防止する目的で、一定数以上のホップ数（ルータ間の通過回数）に達するとパケットが消滅するという使い方をしています。

マルチキャストが Scope で TTL を用いて伝達範囲を制御するためには、ルータ間のリンクの TTL threshold の値の設定を利用します。

図 5 は TTL threshold を 32 に設定した例です。

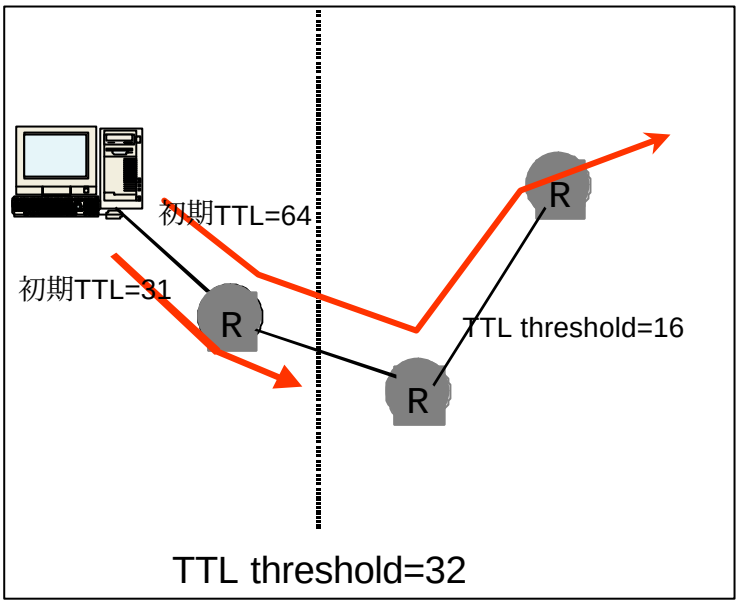

図 5：TTL による Scope の制御

初期 TTL の値を 64 に設定した場合はルータ「R」の通過によって TTL の値は 63 と減りますが、TTL threshold は 32 なのでこれを越えることができます。しかし、初期 TTL の値を 31 に設定した場合はルータ「R」の通過によって TTL の値は 30 と減ります。TTL threshold は 32 なのでこれを越えることができません。

組織間のルータのリンクの TTL threshold を 32 で設定し、組織内のルータのリンクの TTL threshold を 16 で設定した場合、初期 TTL を 32 以上の大きめの値すれば、そのパケットは組織を越えて伝達できます。逆に初期 TTL を 15 に設定すればパケットの到達範囲を組織内ルータすら越えない部署内と制御することが可能になります。日本の国際リンクの TTL threshold は 64 で設定されているので、初期 TTL を 128 とすれば世界中にパケットを伝達させることができます。

アドレスの Scope によるパケット到達範囲の制御は、ルータに設定する boundary を用いて行います。

図 6 はルータ「R」に boundary として 239.0.0.0/8 を設定した例です。

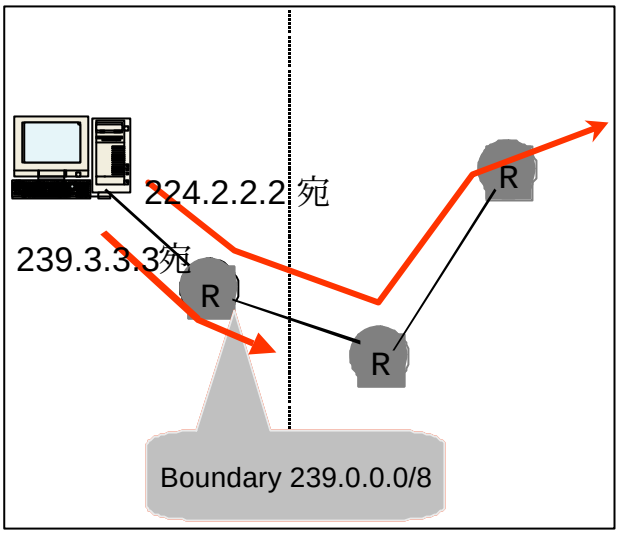

図 6：boundary による Scope の制御

Administrative Scope の範囲に当たる 239.3.3.3 へ宛てたパケットはルータ「R」を越えられますが、破線で示された Administrative Scope の範囲を超えることができません。これに対し、global scope の範囲に当たるアドレスの 224.2.2.2 へ宛てたパケットは、その範囲を越えることができます。

2.2.4 アドレスの割り当て

パケットを送信する際のアドレスの割り当て方法であるアドレスアロケーションをアナウンスする方法として、マルチキャストでは、SAP(Session Announcement Protocol)、GLOP addressing、IETF MALLOC WG（Internet Engineering Task Force Multicast Address Allocation Working Group）の動的マルチキャストアドレス割り当てなどいくつかの試みがなされています。

(1)SAP(Session Announcement Protocol)

実験ネットワークである MBone が採用しているアドレスアロケーションアナウンスのプロトコルは、SAP(Session Announcement Protocol)です。draft-ietf-mmusic-sap-v2-02.txt で規定されています。

SAP では、そのアドレスでのパケットの到達範囲を表す Scope ごとに、マルチキャストアドレスとして使用するグループアドレスをアナウンスするためのアドレスが決められています。global scope の場合は、224.2.127.254/9875 です。Local administrative Scope の場合は、その範囲の一番大きなアドレスを使います。IPv6 については、FF0X:0:0:0:0:0:2:7FFE と決められています。

新たにセッションを開始する場合は、これらの決められたアドレスをし

ばらく受信し、どのアドレスがそのような使われ方をしているかを把握した後に、空いているアドレスの使用をアナウンスすることになります。

セッション情報の記述方法を決めているのは、SDP(Session Description Protocol)です。RFC2327 で規定してあります。以下が記述例です。

```
v=0
o=xxxx 3142894548 3142894629 IN IP4 202.232.2.14
s=xxx Test Channel
i=xxx Test Channel from Osaka branch.
u=http://xxxx.or.jp
e=<xxxxx@xxxx.or.jp>
p=+81-3-xxxx-xxxx
t=3148678800 3151098000
m=audio 29748 RTP/AVP 0
c=IN IP4 239.253.128.81/31
m=video 54210 RTP/AVP 31
c=IN IP4 239.253.128.44/31
```

v はプロトコルのバージョン番号
o はセッションのオーナー
s はセッション名
i はセッション情報
u は URL
e は e メールアドレス
p は電話番号
t は有効期限（セッションが使われる時間）
m は media name（0 ならば PCM オーディオ、31 なら H.261 のビデオ）
c は connection information(マルチキャストアドレス/初期 TTL）です。

(2)GLOP addressing

SAP のようなダイナミックなアロケーションのほかに、実験的な試みとして、アドレスのアロケーションアナウンスの方法として GLOP addressing が提案されています。GLOP addressing は暫定的で、当面の方法として受け入れられています。draft-ietf-mboned-glop-addressing-02.txt で規定しています。

GLOP addressing は、各 AS（Autonomous System：自律システム）が、IANA から、一般的に Class A と呼ばれているアドレスに相当する 233/8 の領域の固定的なアドレスの割り当てを受けるというものです。各 AS の AS 番号からアドレスの真中の 2 オクテット（16bit 分）を計算して割り当てます。各 AS は最後の 8bit 分を自由に使えます。全 AS に対して Class C 相当の範囲のアドレスを渡すという仕組みです。

AS 番号からの計算は 10 進法で表されている AS 番号をいったん 16 進法で表現し直します。その 16 進法の数字を上と下を 2 桁ずつ区切ります。それをもう一度 10 進法に直し、アドレスの真中の 2 オクテットに割り当てるというものです。

IIJ の場合は、
AS2497 = 0x09c1 = 0x09 と 0xc1 = 9 と 193 = 233.9.193/24
となります。

AS 番号からの変換の計算は *http://gigapop.uoregon.edu/glop/*の CGI でもできるようになっています。

(3)動的マルチキャストアドレス割り当て

IETF MALLOC WG で検討しているのは、動的なマルチキャストアドレス割り当てです。IETF MALLOC WG のドラフトは draft-ietf-malloc-arch-03.txt です。その配下に多数のドラフトがあって、動的なマルチキャストアドレス割り当ての全アーキテクチャが述べられています。それらのドラフトに基づくプロトコルが順次検討されています。

IETF MALLOC WG が検討しているのは、世界的なマルチキャストのアドレスを 3 階層に応じた以下の 3 つのプロトコルを用いて動的に割り当てていくという仕組みです。

①MASC
②AAP（Address Allocation Protocol）
③MADCAP（Multicast Address Dynamic Client Allocation Protocol）

の 3 つです。

(a)MASC

MASC は、ドメイン(AS)間の階層に対するプロトコルです。各 AS に対してのマルチキャストアドレスの割り当て方を決めています。ドラフトは draft-ietf-malloc-masc-04.txt です。既にサンプリング・インプリメンテーションが南カリフォルニア大学から出ています（*http://netweb.usc.edu/masc/mascd/*）。最小限の実装例です。

(b)AAP

AAP は、ドメイン(AS)内の階層に対するプロトコルです。MASC から割り当てを受けたマルチキャストアドレスを AS 内で割り当て、AS 内の MAAS（Multicast Address Address Allocation Server）間で交換します。ドラフトは draft-ietf-malloc-aap-02.txt です。

(c)MADCAP

MADCAP は、クライアントに対してマルチキャストアドレスを割り当てるプロトコルです。マルチキャスト DHCP と呼ばれていました。現在、実装するための API まで詳細な検討が進められています。ドラフトは draft-ietf-malloc-madcap-07.txt です。

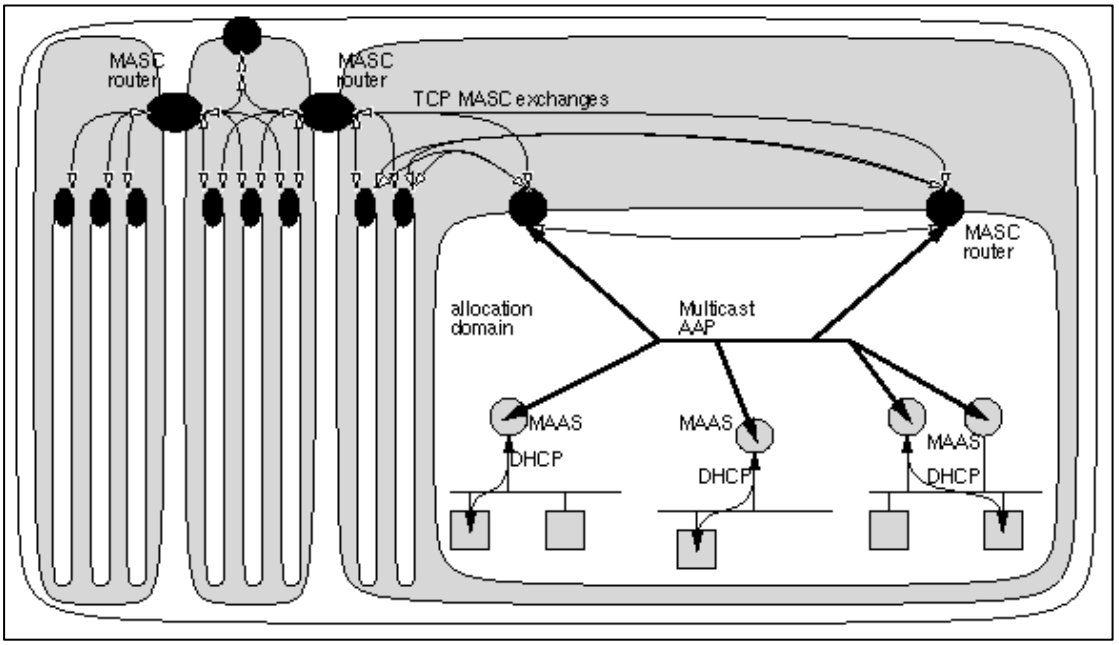

図 7：MASC、AAP、MADCAP の階層構造

ただ、現状では、ホスト側がどのような scope **に属しているのかを分かるような仕組みにはなっていないことから、ホスト側で自組織内だけに届くようにマルチキャストデータパケットの通信経路を制御することができません。このため、**MZAP (Multicast-Scope Zone Announcement Protocol)**というプロトコルが現在、検討されています。**

まだ、MZAP **の仕組みは暫定的です。**MZAP **では、**Zone **に対する権限を持つ** ZBR（Zone border Router）**が** MAAS **の** 239.255.255.252 **に対して、通信経路の告知** Zone Announcement Messages(ZAM)**を送信します。クライアント側は、**MADCAP **によって、**MAAS **へアクセスし、現在どのような** scope **に属していて、データパケットを自組織内だけに届くようにするためにはどのようなマルチキャストアドレスに送出すればいいのか、などの通信経路情報を** ZAM **を通じて把握できるようにします。これによって、動的なマルチキャストアドレス割り当てで生じやすい** misconfiguration **を発見しやすくする効果も期待されています。ドラフトは** draft-ietf-mboned-mzap-05.txt **です。**

IETF **では** MZAP **のほかに** SADP(Scoped Address Discovery Protocol)**も提案されています。**SADP **も** MZAP **と同様にクライアント側が現在属している** scope **を調べるためのプロトコルです。提案だけに終わって**

いる可能性もあり、1999 年ワシントンでの会合では SADP に関して大きな動きはありませんでした。

2.2.5 イーサネットのマルチキャストアドレスマッピング

ここでは、イーサネットレベルでのマルチキャストアドレスへのマッピングについて説明します。

イーサネットのアドレスは 6 オクテットあります。イーサネットレベルでのマルチキャストアドレスは、先頭のフレームが 01（最下位ビットが 1）となっている場合、マルチキャストまたはブロードキャストとして扱われます。さらに、2 番目と 3 番目のフレームが 00、5E となっている場合、IP マルチキャストのためのイーサネットレベルでのマルチキャストアドレスということになります。この、2 番目と 3 番目のフレーム 00、5E と 4 番目のフレームの先頭 1 ビット分を、IETF は IEEE に予約しています。つまり、IP マルチキャストのマッピングには、イーサネットのアドレスは 6 オクテット中、1 番目 01、2 番目 00、3 番目 5E のフレーム 3 オクテット分と 4 番目のフレームの先頭 1 ビット 0 の部分を除く下位 23 ビットが使われることになります。

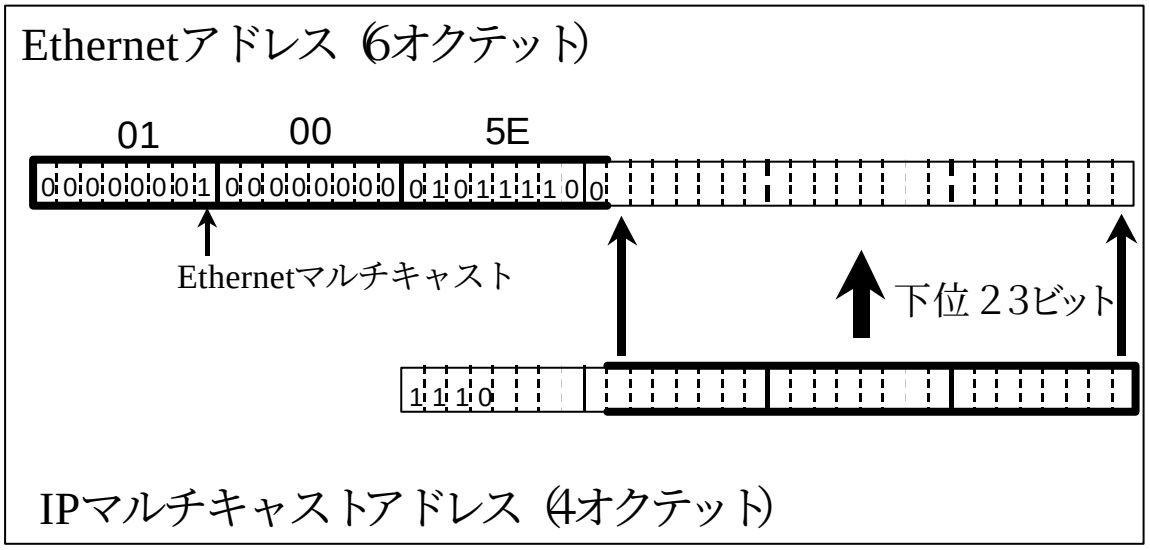

図 8：イーサネットでの IP マルチキャストのフレームマッピング

IETF が IEEE に 4 番目のフレームの先頭 1 ビット分を予約したのは、将来の IP によるデータ配信技術のためのアドレスマッピング領域を確保しているということです。このため、4 番目のフレームの先頭 1 ビット分が 1 のときは、下位 23 ビットは何にも使われていません。

また、イーサネットマルチキャストアドレス 1 個に対して、IP マルチキャストアドレス 32 個が同じアドレスにマッピングされるため、イーサネットマルチキャストアドレスと IP マルチキャストアドレスは 1 対 1 の関係ではありません。

このため、イーサネットのネットワークインタフェースカード（NIC）では、イーサネットマルチキャストアドレスから IP マルチキャストアドレスを判断することはできません。

2.2.6 NIC での対応

イーサネットの NIC には固有の MAC アドレスが付けられています。このため、ユニキャストの場合だと MAC アドレス以外に宛てられたパケットについて CPU への判断を求めずに NIC がハードウェアレベルでフィルタリングしています。

これに対して、マルチキャストの場合は、イーサネットマルチキャストアドレスと IP マルチキャストアドレスは 1 対 1 の関係でないため、NIC レベルで必要なパケットであるかどうかの判断ができません。

このため、マルチキャストに対する NIC レベルでの対応は NIC の種類によって様々です。それらは以下の 4 つのタイプに類型化できます。

①ハッシュテーブルでパケットのフィルタリングを行うタイプ
②マルチキャストパケットを全部受信するタイプ
③マルチキャストパケットを受信するときにそれ以外の全パケットを受信するタイプ
④マルチキャストパケットを受信できないタイプ

最近の NIC は CPU に負荷をかけない①のタイプが増えてきました。

2.3 IP マルチキャストのプロトコル

ここでは、ルータ内のローカルセグメントでのグループ管理プロトコル、ルータ間の主要な転送プロトコル、広域のドメイン間のルーティングプロトコルについてそれぞれ説明します。

2.3.1 グループ管理プロトコル

(1)IGMP

IGMP(Internet Group Membership Protocol)は、ホストがマルチキャストパケットを受信するために、どの IP マルチキャストアドレスのグループに参加しているかをローカルなサブネット上で隣接しているルータに通知する IPv4 用のプロトコルです。IPv6 では、IGMP(Internet Group Membership Protocol)に相当するプロトコルを MLD(Multicast Listener Discovery)と呼んでいます。機能も同じです。ICMPv6 のサブセットに収められています。

IGMP は、RFC1112 で IGMPv1 が、RFC2236 で IGMPv2 がそれぞれ定義されています。IGMPv2 は IGMPv1 に Leave Group 機能を加えたものです。Leave Group とは、ホストがマルチキャストパケットの受信をやめるため IP マルチキャストアドレスのグループへの参加を取りやめるための機能です。IGMPv3 はまだドラフトの段階ですが、詳細がま

とまってきており、近々、RFC として公開されると思います。

IGMP の基本的な動作は、ルータがホストに受信を希望するマルチキャストアドレスを問い合わせる IGMP Query とホストが IGMP Query に答えルータに受信を希望するマルチキャストアドレスを告げる IGMP Report、ホストがマルチキャストの受信を停止するためマルチキャストアドレスのグループへの参加の取りやめをルータに告げる LeaveGroup という 3 つのパケットの送受信によって行われます。

まず、ルータは、オールノードのアドレスを示す 224.0.0.1 に対して General Query として IGMP Query を送信します。これによって、マルチキャストに対応しているローカルなサブネット上のすべてのホストが IGMP Query を受信します。

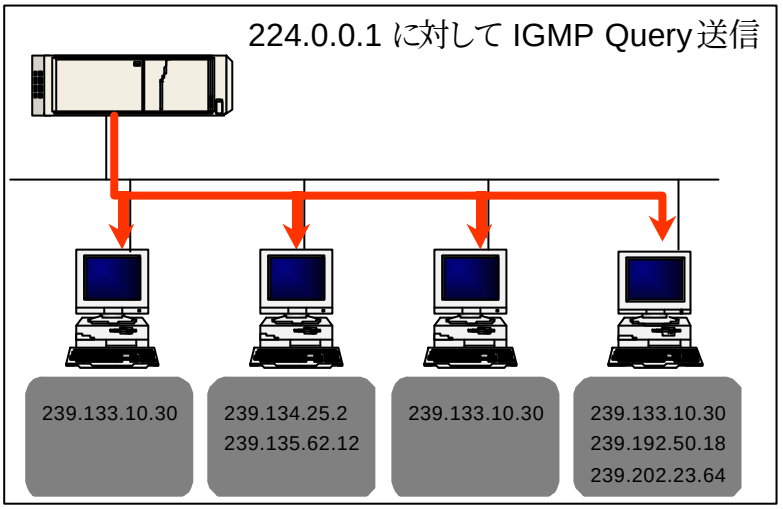

図 9：IGMP Query の送信

次に、IGMP Query を受信した全ホストで、マルチキャストの受信を希望するホストは IGMP Report をルータに対して送信し、マルチキャスト受信のために参加するマルチキャストアドレスを告げます。ルータ側は IGMP Report によって、マルチキャストアドレスグループへの参加を希望するホストの存在を知ることができます。ただ、そのホストが具体的に何台あるのかまでは把握できません。ルータが把握できるのは、ローカルなサブネット上に希望するホストがあるかないかだけです。

また、ホストによる IGMP Report の送信の際、IGMP Query を受信したホストが一斉に IGMP Report を送信するとネットワークがバーストする可能性があります。このため、これを回避するためにホスト側で返答タイマがランダムに起動する仕組みになっています。返答タイマは、受信を希望するマルチキャストアドレスごとにスタートします。ホストは返答タイマがタイムアウトになった順番で IGMP Report を送信します。

さらに、同じマルチキャストアドレスの受信を希望するホストが複数あった場合、既に他のホストが IGMP Report を送信していたら、そのホストは IGMP Report は送信しません。ルータが IGMP Report によって

把握できるのは、ローカルなサブネット上に希望するホストがあるかないかだけですから、ネットワーク上に無駄にパケットを送信しないようにしているわけです。

ホストがマルチキャストパケットの受信を停止するときには、ルータに LeaveGroup というパケットを送信します。

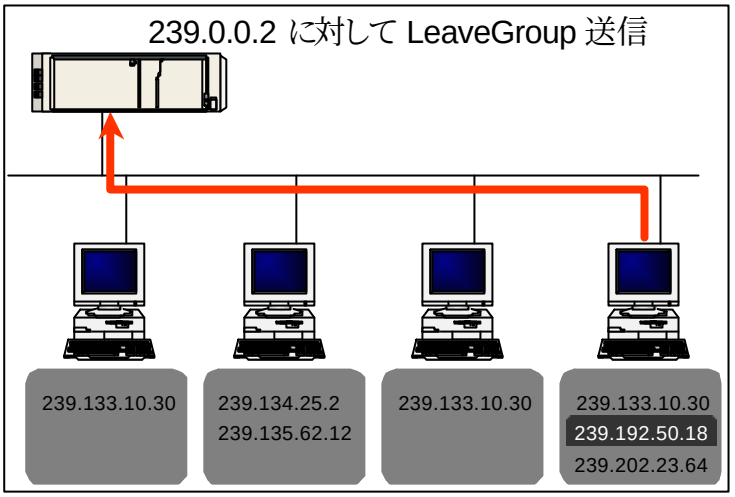

図 10：LeaveGroup の送信

LeaveGroup を受け取ったルータはマルチキャストパケットを受信していたグループに対して GroupSpecificQuery を送信します。

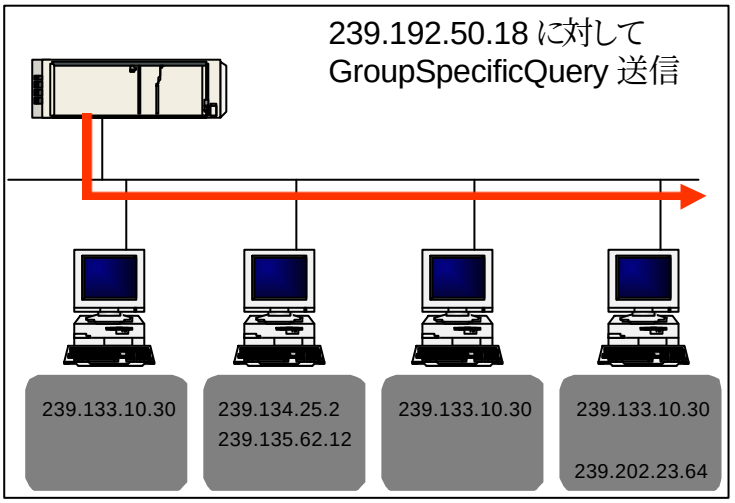

図 11：GroupSpecificQuery の送信

グループの中から、受信停止の要求があったため、マルチキャストパケットの送信を停止していいかを確認するためです。この場合、グループの中から GroupSpecificQuery に対する返信があった場合は、ローカルなサブネット上に受信を希望するホストがあると認識して送信を続けます。逆に、何も返信がなかった場合、グループの中には受信を希望する

ホストは 1 台もないと判断して送信を打ち切ります。

IGMPv1 のように Leave Group 機能がないと、ホストは自分宛のマルチキャストパケットが送られ続けている状態で勝手に受信を停止することになります。ルータはオールノード（ホスト）に対しての Query として General Query を 125 秒間に 1 回発信します。ルータはこの Query に対する応答がないホストへのマルチキャストパケットの転送を停止することになります。

(2)イーサネットスイッチ問題

イーサネットスイッチは、各ポートの先につながっているホストの MAC アドレスによって受け取ったパケットをスイッチングしています。このため、マルチキャストを想定していないイーサネットスイッチは、マルチキャストパケットを受け取るとスイッチング先を判断できません。そのようなイーサネットスイッチは、多くの場合、接続されている全ポートのホストに対し、マルチキャストパケットをコピーして送信してしまいます。これがイーサネットスイッチ問題です。この問題では、イーサネットスイッチがスイッチとしての機能を果たさないというだけでなく、マルチキャストパケットのコピーによって CPU にも大きな負荷をかけてしまいます。

イーサネットスイッチ問題に対処するためのスイッチの対応方法は、①IGMP snooping、②CGMP(Cisco Group Management Protocol)、③IEEE 802.1 GMRP の 3 つがあります。

(a)IGMP snooping

IGMP snooping は、イーサネットスイッチが IGMP Query パケットや IGMP Report パケットでイーサネットレベルのマルチキャストアドレスを覗き見して、転送先を判断する方法です。ただし、この方法はレイヤ 2 スイッチだと snooping 対象のパケットを見分けることができず CPU に負荷がかかりすぎるという欠点があります。これに対し、レイヤ 3 スイッチは IGMP のパケットと通常のマルチキャストパケットを見分けることはでき、IGMP のパケットだけ snooping することはできます。ただ、レイヤ 2 スイッチと比較して高価であるというのが欠点です。

(b)CGMP(Cisco Group Management Protocol)

CGMP は、最寄りのルータ（イーサネットスイッチに対し IGMP Query を送信したルータ）がマルチキャストパケットのスイッチング先を教えるという仕組みです。このため、イーサネットスイッチがレイヤ 2 スイッチであったとしても CPU に負荷はかかりません。ただ、CGMP の欠点は Cisco のルータとスイッチでないと使えないというベンダー依存のプロトコルである点です。

(c)IEEE 802.1 GMRP

IGMP snooping、CGMP にはそれぞれ欠点もあることから、イーサネットレベルで IEEE 802.1 GMRP (Generic attribute registratiojn protocol Multicast Registration Protocol）として標準化を目指す動きもあります。このプロトコルは、各ホスト（NIC）がイーサネットスイ

ッチに対して受信を希望するイーサネットマルチキャストアドレスのグループを宣言するという仕組みです。3Com が積極的に推進しています。ただ、これはまだ新しいプロトコルであるほか、NIC をバージョンアップしなければならないという問題点があります。

2.3.2 ルータ間のマルチキャスト転送プロトコル

ここでは、ルータ間のルーティングプロトコルについて説明します。

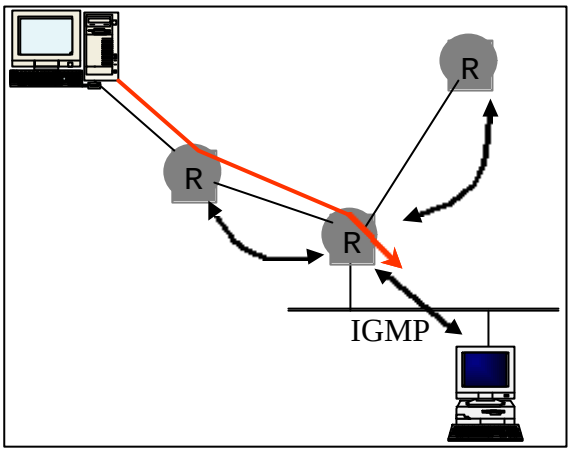

図 12：IP マルチキャストルーティングプロトコル

(1)DVMRP

(a)DVMRP の仕組み

ルータ間のルーティングプロトコルとして最もよく知られているプロトコルは DVMRP（Distance Vector Multicast Routing Protocol）です。ドラフトは、draft-ietf-idmr-dvmrp-v3-09.txt です。

DVMRP は、IGMP に基づき各ルータが接続しているリンクとメトリックを他のルータと相互に交換し、マルチキャストパケットの最適転送経路を決定するプロトコルです。距離ベクトル型、Reverse Path Forwarding、flooding & pruning という特徴があります。

DVMRP では、リンクとメトリックのアナウンスが隣接ルータ間で次々に転送されることで全ルータでルータとメトリックが分かります。また、リンクが切れた場合など定期的にアナウンスされます。ユニキャストの RIP（Routing Information Protocol）とよく似ており、距離ベクトル型の特徴を持っています。

Reverse Path Forwarding（RPF）とは、ルータなどのインタフェースが、受け取ったマルチキャストパケットのソースアドレスから最短経路を通過してきたことが確認できたら、そのパケットを他のインタフェイ

スへ転送し、最短経路でなかったら転送せずに破棄するという仕組みです。

flooding & pruning は、Poison Reverse と Prune パケットによって、マルチキャストパケットを配信する最短経路のツリー構造を確立する仕組みです。

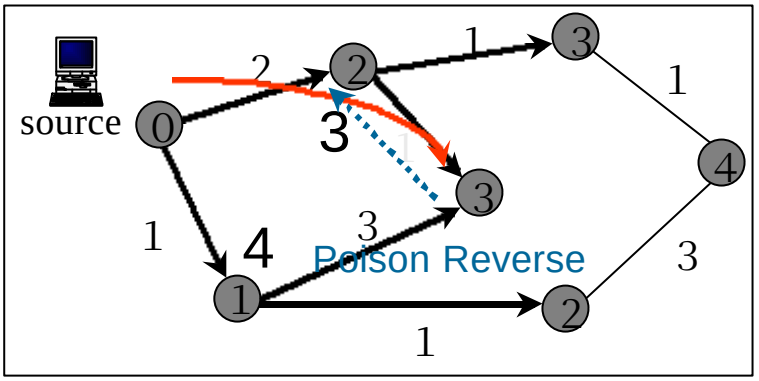

図 13：Poison Reverse

Poison Reverse は、マルチキャストパケットを受け取ったルータ（インタフェース）が上位流として転送元のインタフェースを認識したことを通知する方法です。

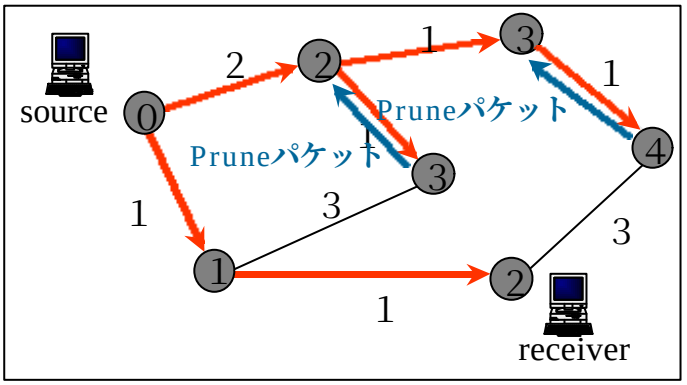

図 14：Prune パケット

Prune パケットは、インタフェースが上位流に対してマルチキャストパケットの送信の停止を求めて送信するパケットです。ローカルなサブネット上にマルチキャストパケットの受信を希望するホストが存在しない場合に送信します。

flooding & pruning では、最短経路のツリー構造を確立するまでは、マ

ルチキャストパケットの受信を希望するホストの有無に関係なく全インタフェースにマルチキャストパケットに配信します。このパケットの流れがいったんは洪水のように溢れるという意味で flooding という言い方をします。

(b)DVMRP の実装例

DVMRP をインプリメンテーションした代表的な例はソフトウェアベースのマルチキャスト対応ルータ mrouted です。mrouted は、マルチキャストに対応したネットワークを相互に結びつけるためマルチキャストに対応していないネットワークにトンネルを作るマルチキャストトンネリングを実現します。

mrouted では、ユニキャストのネットワークを経由してマルチキャストパケットの送受信を可能にするため、マルチキャストパケットをユニキャストパケットの中にカプセリングします。これによって、mrouted 間をトンネルとしてマルチキャスト対応ネットワークを相互に結びつけます。

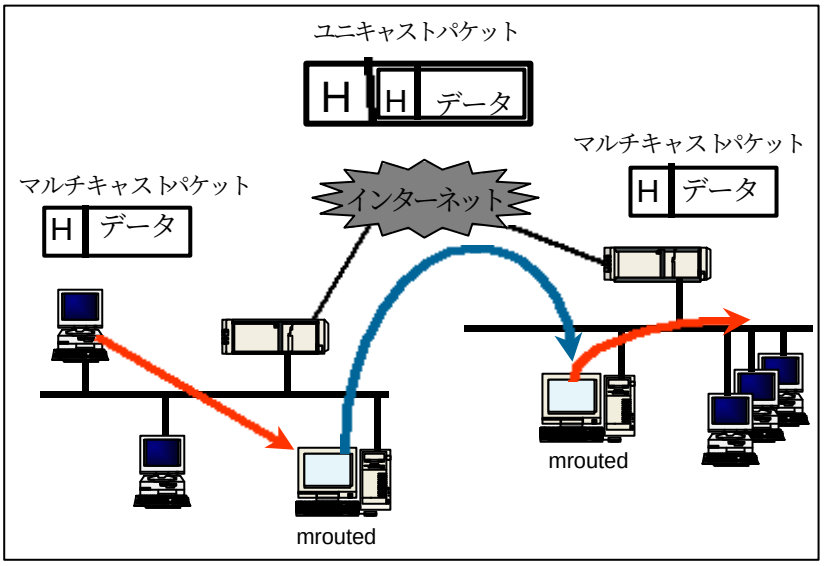

図 15：mrouted のマルチキャストトンネリング

mrouted のトンネルのパラメータは、/etc/mrouted.conf という設定ファイルに、

tunnel 192.168.1.2 192.168.2.3 metric 1 threshold 16 rate_limit 512 boundary 239.255.0.0/16

の 1 行を加えます。その意味は以下のとおりです。

●tunnel で指定しているアドレスは local address、remote address の順です。metric は経路制御のための距離です。通常は 1 を指定します。

●threshold は通過させる TTL の大きさです。組織間なら 32、組織内なら 16 を指定します。

●rate_limit はトンネルに流れる最大の転送レートです。Kbps 単位で転送バイト数の上限を指定します。

●boundary は scope address を設定します。この範囲のアドレスはトンネルを通すなという意味になります。

(c)DVMRP の評価

DVMRP は、DVMRP を実装した mrouted がインターネット上に仮想のマルチキャストネットワークを構築した実験ネットワーク MBone（Multicast backBONE）で広く使われたことで普及しました。しかも、ベンダニュートラルなため、実装例も多くあります。

しかし、広域で使用するには問題も多く難しいでしょう。距離ベクトル型であるため、経路情報の安定に時間がかかるほか、ネットワークが世界的な規模になると不安定な経路による flapping の問題も発生します。最大の問題は、全世界の不必要なリンクや細いリンクにまで flooding し、巨大なトラフィックを発生させてしまうことです。prune しないノードあったりすると、そこが Blackhole となって、全世界からのパケットが流れ込んできてその上位流のサイトまでが迷惑するという問題があります。このため、現在では、DVMRP を使う方向にはなっていません。

(2)PIM

PIM（Protocol Independent Multicast）は、様々なユニキャストのプロトコル上でマルチキャストルーティングを実現させるプロトコルです。密(Dense)モードと疎(Sparse)モードで構成します。

Dense モードは、DVMRP と似ています。異なる点は DVMRP が Distance Vector を使って、自らルーティングテーブルを作っていたのに対して、PIM DM（Dense モード）は全部をユニキャストのルーティングテーブルを使う分だけ負荷が掛からないということです。比較的に狭い地域で、受信者が多くトラフィックも多い場合、例えば全員がマルチキャストパケットを受信するなどブロードキャスト的に使う場合には、flooding & pruning の仕組みでパケットが溢れ出しても構わないというポリシーに基づいています。このため、poison reverse もなく、パケットは全経路に流れることになります。ドラフトは draft-ietf-pim-v2-dm-03.txt です。

これに対し、Sparse モードは、広い地域で、受信者が偏在し、トラフィックも少ない場合に利用するモードです。明示的に RP（ランデブーポイント）を設定し、送信者は RP へ向けてマルチキャストパケットを送信し、受信者は RP へ明示的にマルチキャストパケットをもらいにいくという仕組みです。ドラフトは draft-ietf-pim-v2-sm-01.txt です。

Sparse モードでは、RP に対する受信側の最寄りルータの IGMP Query と IGMP Report、Join message の送信と、送信側の最寄りルータでユニキャスト用にカプセル化した Register Message の RP への送信、RP による送信側の最寄りルータへの Join message の送信の行程を経て、マルチキャストパケットを配信します。

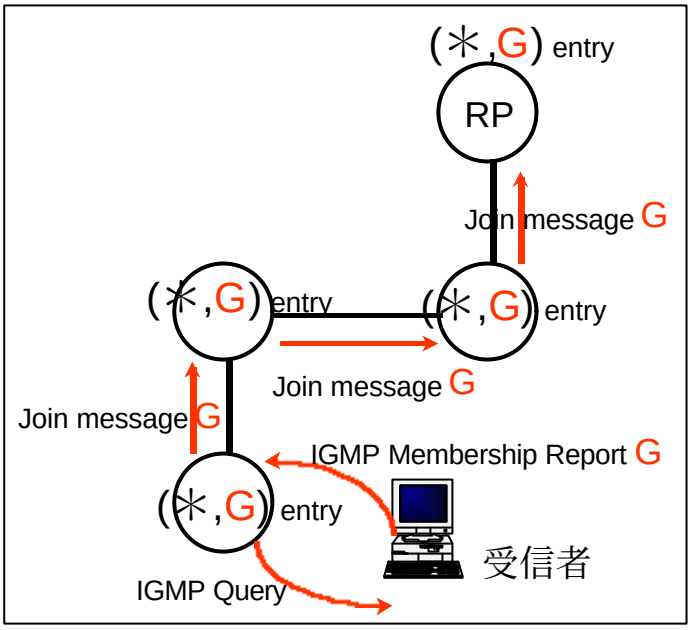

図 16：受信者側最寄りルータと RP

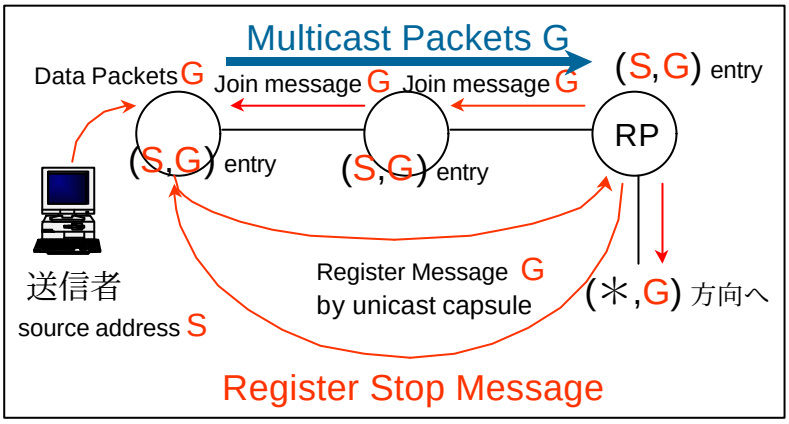

図 17：送信者側最寄りルータと RP

受信側が実際にマルチキャストパケットの受信を開始すると、RP の方向に関係なくマルチキャストパケットを受信する最短経路のツリー構造を確立するため、ルータは送信側方向のルータへの Join message の送信と不要なパケット転送の停止を求める Prune message の送信によって、最終的に最適経路を作り上げます。

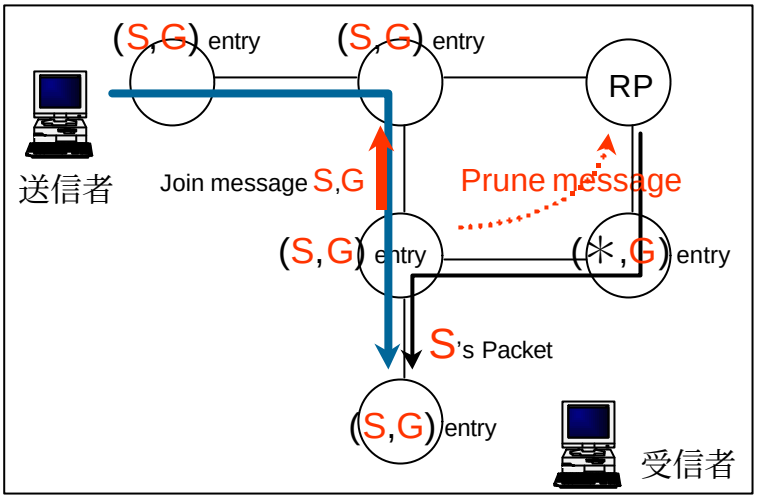

図 18：疎(Sparse)モードでの最適経路

PIM Sparse モードには、明示的に join するため、必要のないリンクに無駄なトラフィックが流れないというメリットがあります。これは、マルチキャストパケットを受信したいホストが join することによって、不必要なルータにまでパケットが溢れることがないためです。

ただ、このプロトコルは、Cisco のプロトコルというベンダ色があったため、かつては他のルータベンダが実装を躊躇している傾向もあったようです。しかし、世界的な潮流として DVMRP をこれ以上、推進していくことは無理という合意も形成されつつあり、ベンダ色が普及の足枷という傾向は最近では薄れてきたようです。

また、RP にトラフィックが集中するという欠点も指摘されています。これに対しては、徐々に複数の RP で負荷分散する技術がプロトコルに装備され始めつつあり解決される方向にあります。

最大の問題は、広域の ISP をまたぐ通信の場合、他組織の RP に依存することになるということです。この Third-party Resource Dependency の問題については、複数の RP をマネジメントする仕組みとして MSDP(Multicast Source Discovery Protocol)などいくつかの解決策が提案されています。

2.3.3 ドメイン間の転送プロトコル

ここでは、広域のドメイン間のルーティングプロトコルについて説明します。将来的には、BGMP(Border Gateway Multicast Protocol)が有力視され、現在開発が進められています。しかし、まだ実装と普及には時間がかかりますので、BGMP が実用段階に入るまでの短期的な解としては、俗称で MBGP と呼ばれている BGP4+を利用しようという取り組みもあります。

(1)BGMP(Border Gateway Multicast Protocol)

BGMP は、MASC で各 AS 割り当てられたアドレスに関しては、その

ドメインが root domain になるという仕組みです。DVMRP や PIM のランデブーポイント(RP)に相当する機能を双方向の Shared Tree を使って domain が行います。マルチキャストで転送するときの Reverse Path Forwarding（RPF）用の経路は BGP4＋の NLRI(Network Layer Reachability Information) に乗せて告知するようにします。

(2)BGP4+

BGP4+は RFC2283(Multiprotocol Extensions for BGP-4)で規定されているマルチプロトコルの拡張で、BGP4 の仕組みだけ利用しようというものです。BGP4 は、Reverse Path Forwarding（RPF）用にユニキャストのプレフィックスをアナウンスするために使っています。

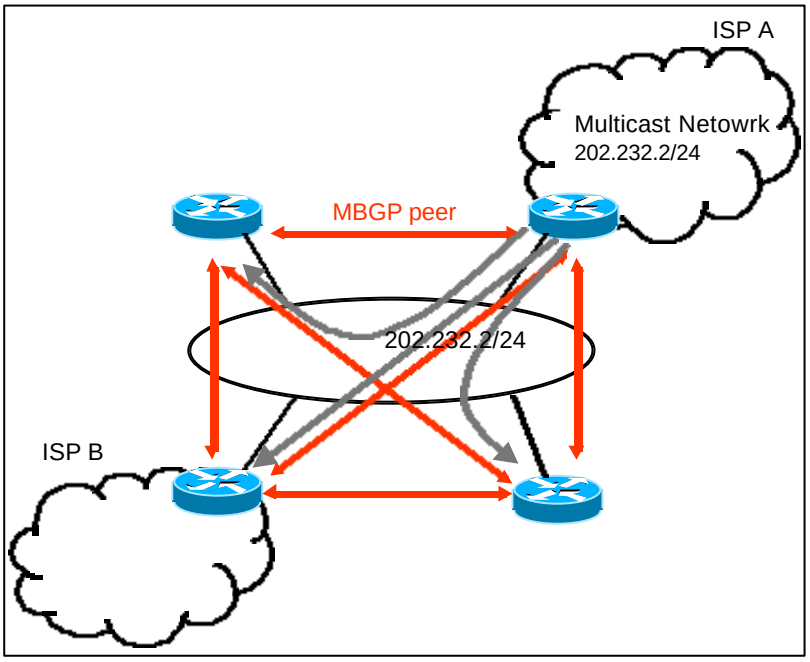

図 19：BGP4+

実際の広域のマルチキャストフォワードには、PIM-SM を使おうとしました。

ただ、PIM-SM を全組織で使うとなると、ルーティングを他の組織の RP に対して完全に依存する Third-party Resource Dependency の問題が浮上します。

それに対する解決策として、MSDP(Multicast Source Discovery Protocol)が提案されています。MSDP では、ISP などのそれぞれのドメインで RP を独自に設置し、RP 間でアクティブなソース情報を交換、共有するプロトコルです。

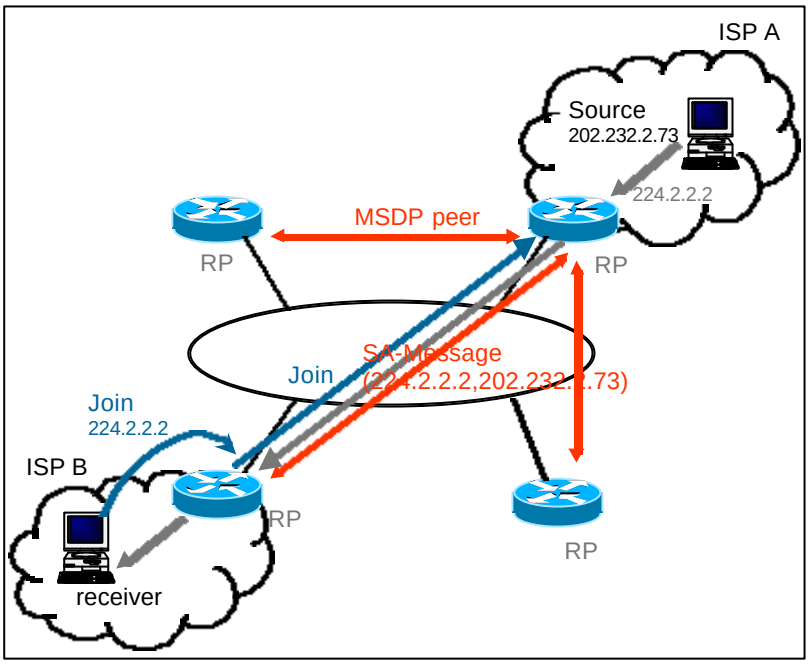

図 20：MSDP

3　マルチキャスト対応機器

ここでは、マルチキャストに対応している機器、OS、ルータ、TA、モデム、ダイアルアップルータ、マルチキャスト対応アプリケーションの設定について説明します。

3.1　マルチキャスト対応ＯＳ

(1)UNIX 系 OS

SunOS は SunOS 4.1.x でも Solaris 2.x のいずれもクライアントになるには何もする必要はありません。ただ、マルチキャストのルータにする場合にはカーネルにパッチが必要です。

●SunOS 4.1.x (ipmulti3.5-sunos41x.tar.gz)
ftp://ftp.iij.ad.jp/pub/multicast/kernel/
●Solaris 2.x (Solaris_mc35+2.x-patch.tar.gz)
ftp://playground.sun.com/pub/multicast/

BSD/OS はデフォルトでカーネルのコンフィグレーションにマルチキャストオプションが付いているので、クライアントにもサーバにもなれます。マルチキャストのルータにする場合には、コメントアウトされている MROUTING のコメントアウトをとって、カーネルにコンパイルすることになります。

```
options         MULTICAST
#options        MROUTING
```

FreeBSD、NetBSD は options MULTICAST すらありません。最初から組み込まれています。マルチキャストのルータにする場合だけ、コメントアウトされている MROUTING のコメントアウトをとって、カーネルにコンパイルすることになります。

```
#options        MROUTING
```

Linux の場合は、設定方法が少し違います。menuconfig で IP multicasting、IP multicast routing、IP tunneling のオプションをオンにします。IP tunneling は、mrouted を使ってトンネリングする場合に必要なオプションです。

IRIX、AIX、Tru64UNIX(DIGITAL UNIX)などの商用 OS は最近のバージョンだとほとんどマルチキャストに対応しています。

(2)PC 系 OS

PC 系 OS はクライアントになるだけだったら、どの OS も対応しています。

Microsoft Windows95 は、IGMP version 1 のレベルで対応しています。Leave Group 機能はありません。マルチキャストのルータにもなれません。Microsoft Windows98 は IGMP version 2のレベルでの対応です。Microsoft WindowsNT version 3.5 以上は、IGMP version 1のレベルで対応です。マルチキャストのルータにはなれないので、mrouted を使うことができません。Windows2000 では対応するということになっているようです。

MacOS は、8.5.1 でイーサネット接続の場合だけ IGMP version1 に対応しています。シリアル接続では IGMP Report をルータに対して送信できないため、マルチキャストパケットの受信ができません。

3.2 マルチキャスト対応ルータ

3.2.1 ソフトウェア

最も有名なのは、mrouted です。ftp://ftp.iij.ad.jp/pub/multicast/mrouted にあります。プロトコルでは DVMRP を採用しています。設定方法は、/etc/mrouted.conf に「tunnel 自分 相手 metric 1 threshold 32 rate_limit 512」と設定すれば、トンネリングが可能になります。トン

ネリング以外でもネイティブにマルチキャスト対応させることができます。

gated も有名なプログラムです。最新バージョンがマルチキャストに対応しています。フリーのバージョンでは対応していません。BGMP、PIM-SM、PIM-DM、MSDP など先進的なプロトコルをインプリメンテーションしています。*http://www.gated.org/*で、コンソーシアムメンバのみに配布しています。

南カリフォルニア大学で開発された pimd は、PIM-SMv2 に対応しています。*http://catarina.usc.edu/pim/pimd/*で入手できます。

3.2.2 ハードウェア

Cisco のルータは、PIM-DM、PIM-SM にほぼ対応しています。単に PIM-DM、PIM-SM を使うだけなら、IOSv11．1 や IOSv10.x で構いません。MGBP や MSDP のなど ISP 間をまたぐプロトコルを利用する場合には、IOSv12.0s などの対応したルータが必要になります。ftp://ftp-eng.cisco.com/ipmulticast.html で詳細を記したドキュメントがあります。

このほか、Baynetworks(Nortel Networks)は、BayRS 13.20 で DVMRP、PIM-SM に対応しています。3Com は、DVMRP、MOSPF に、Cabletron SSR は DVMRP に対応、PIM-DM/SM に対応予定です。また、Torrent IP9000 は DVMRP に、Newbridge VIVID は、DVMRP、MOSPF、PIM に、Juniper は、DVMRP、PIM-SM にそれぞれ対応しています。

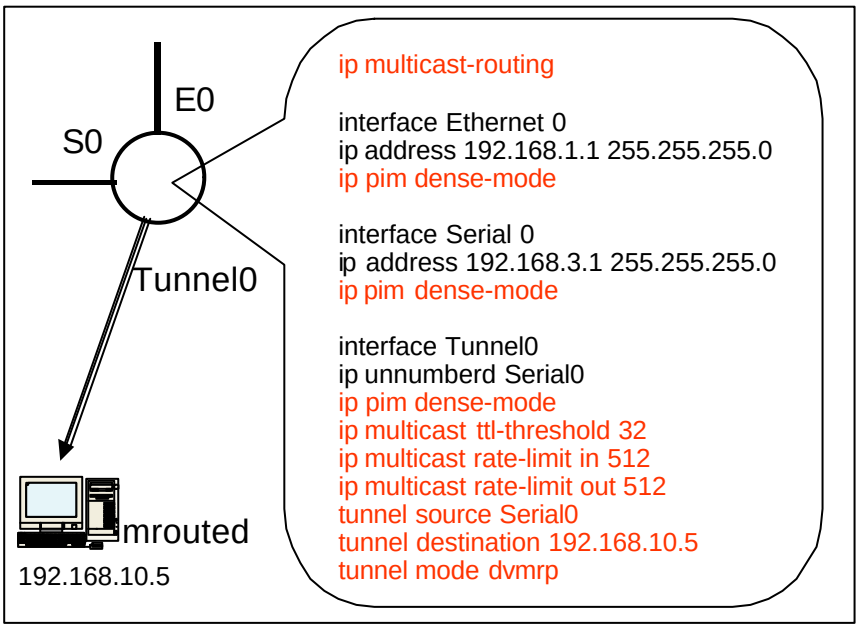

図 21：IOS での PIM-DM

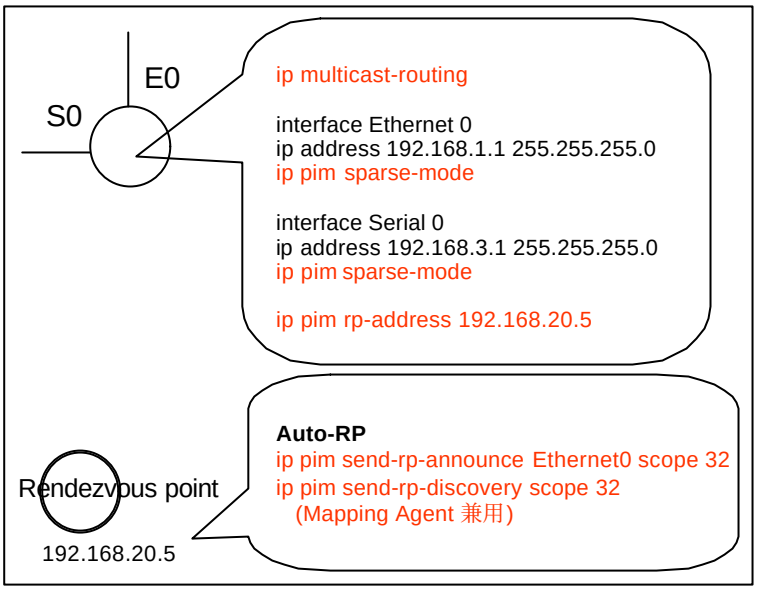

図 22：IOS での PIM-SM

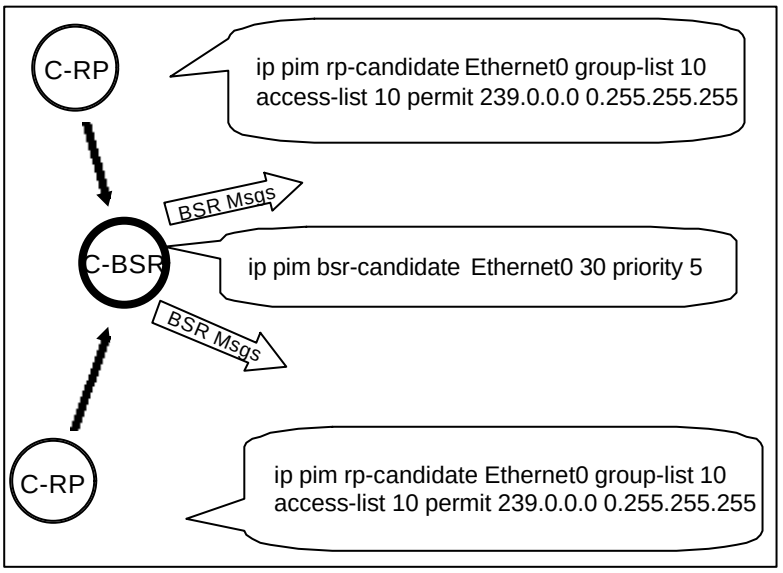

図 23：IOS での PIM-SMv2

3.3 TA、モデム、ダイアルアップルータ

ユーザ側の設備では、TA とモデムはマルチキャストのレイヤとは関係がないので問題なく使用できます。これに対し、ダイアルアップルータの場合は、IGMP ブリッジになっていれば簡単に使用できます。現状では、IIJ SEIL と古河電工 MUCHO の 2 機種が IGMP ブリッジに対応

しています。

ISP 側の設備では、Ascend MAX の場合は以下のような設定が必要です。

Ethernet -> Mod Config -> Multicast Forwarding=Yes
Multicast Client=No
Multicast Rate Limit=0

Radius 側
Ascend-Multicast-Client = Multicast-Yes

3.4 マルチキャスト対応アプリケーション

マルチキャスト対応アプリケーションは以下のようなものがあります。

Audio/Video アプリケーションでは、RealSystemG2、 Windows Medis Technology(WMT)、東芝の MobileMotion、NTT の SoftwareVision がユニキャストのほかマルチキャストで使うこともできます。IP/TV、PrimeCast、 ICAST(I-Station)はマルチキャストを主体にしたアプリケーションです。vic、 vat は MBone で開発されたフリーアプリケーションです。OPTIVISION NAC-3000(MPEG1,2)、 Audioactive(MP3)はハードウェアレベルのアプリケーションです。

Push アプリケーションは、マルチキャストに向いています。PointCast、BackWeb、 Castanet のほか、ミドルウェアの TIBCO/Rendezvous、Java で開発された NTT の RealPush などがマルチキャストに対応しています。

Data Distribution アプリケーションは、大量のデータを配布するときに用いられるアプリケーションです。CAD データの全社設計部門への配布や全国の支店への商品マスタの配布などで活用されます。代表的なアプリケーションは Star Burst Communications 社の OmniCast です。プロトコルには MFTP を採用しています。このほかに、Lucent Technologies 社の e-cast、Global Cast Communications 社の製品など同様な製品があります。

Disk Image Copy アプリケーションは、アプリケーションのバージョンアップのためにプログラムのディスクイメージをマルチキャストを使って一斉に配布するというものです。Ghost、ImageCast などです。

3.4.1 RealG2Server

RealG2Server は、通常のサーバからスプリッティングした先でのマルチキャスト配送や、G2Server 1 台からのユニキャストとマルチキャストの両方の送信も可能です（図 24）。設定例は以下のとおりです。

Configure -> Multicasting -> Back-Channel にて設定

```
<List Name="Multicast">
    <Var DeliveryOnly="0"/>
    <Var TTL="31"/>
    <Var PNAPort="7070"/>
    <Var Resend="1"/>
```

```
        <Var  AddressRange="239.192.200.0-239.192.200.255"/>
        <Var  RTSPPort="554"/>
        <List  Name="ControlList">
            <List  Name="1">
                <Var  Allow="210.130.0.0:255.255.0.0"/>
            </List>
        </List>
    </List>
```

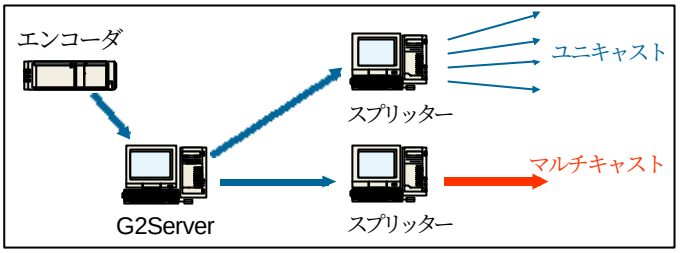

図 24：G2Server からのユニキャスト、マルチキャストの送信

3.4.2 Windows Media

Windows Media の場合は、サーバの論理名ステーションにマルチキャストアドレスや初期 TTL を設定します。

1つのステーションの中に複数のストリームを設定します。その際にエンコーダからのライブストリームを受け取る場合には、そのエンコーダーの IP アドレスを msbd://encoder:7007 と指定します。

また、ディスク上の ASF ファイルを繰り返し再生しながら near on demand のようにストリームすることもできます。

さらに、他のサーバからユニキャストで受信してマルチキャストで再配信するという中継が可能です。msbd://other_server/station1 と指定します。ユニキャストで送信しているデータの一部をマルチキャストで送信するという構成もとれます。

クライアントは、http 経由で取得した.asx ファイルに記述された.nsc ファイルの中のマルチキャストアドレスを参照し、そのアドレスに join します。

3.5 マルチキャストで送信するには

マルチキャストでデータを送信するためには、まずマルチキャスト対応 ISP を選ぶ必要があります。日本では、IIJ がコンシューマ向けの IIJ4U、企業向けにエンタープライズダイアルアップサービスを提供しているほか、NTT サテライトコミュニケーションズの MegaWave、宇宙通信の DirecPC、NTT-ME の XePhion などがあります。

また、ISP へマルチキャストを送信するためにはマルチキャストルータ

による接続が必要になります。ただ、ISP の Head-end まではユニキャストで送信し、ISP のサーバでマルチキャストに変換するという方法もあります。

送信時の注意点は、マルチキャストアドレスは現状では静的に割り当てられているということです。アナウンスは静的もしくは SAP(Session Announcement Protocol)です。

セキュリティ（認証、暗号化）については、アプリケーション層で対応するしかないのが現状です。

信頼性については、基本的に UDP なので、送った後のこと、例えばパケットロスがあったかなどは分かりません。このため、現在、Reliable Multicast Protocol として 20 種類を超える提案がなされています。ただ、アプリケーションによって要求される信頼性のレベルが異なるため、このプロトコルが 1 つに集約されることはないでしょう。

マルチキャストは on demand には向いていません。基本的に live feed のみです。このため、near on demand の仕組みを使って工夫する必要があります。

4. IP マルチキャストへの様々な取り組み

4.1 IPMI の取り組み

IPMI(IP Multicast Initiative)は、Stardust.com の主催で 1996 年に米国で設立された業界団体です。URL は、*http://www.ipmulticast.com/* です。IP マルチキャスト技術の普及啓蒙を目的として、技術ドキュメント発行やセミナーの開催、相互接続検証実験を行っています。年に 1 回 IP Multicast Summit (mCAST2000)を開催し、ハードベンダ、ソフトベンダ、ISP、ICP など 55 社が参加しています。主な参加企業は 3Com、Ascend、Cisco、Extreme、FORE、HP、IBM、Intel、Newbridge、Nortel、Sun、AT&T、BellSouth、C&W、Gilat、Hughes、IIJ、Lucent、PanAmSat、Sprint、UUNET、QwestBroadcast.com、Microsoft、RealNetworks、TIBCO、StarBurst です。

IPMI の日本支部は IPMI-JP（*http://www.iijnet.or.jp/ipmulticast/*）です。運用しているメーリングリストは、ipmulticast@iijnet.or.jp です。参加条件はありません。ドキュメントの翻訳やバイヤーズガイドの作成も予定しています。ISP 間相互接続実験 J/Splash では、MBGP、MSDP、PIM-SM による接続実験を国内 34 社で行っています。

このほかに、Mbone というトンネル技術でのマルチキャスト共同実験もあります。日本では、JP Mbone として活動しています。JP Mbone は、IP マルチキャスト技術の研究開発を目的としています。mbone-jp@wide.ad.jp にて調整します。加入申し込みは mbone-jp-

request@wide.ad.jp で受け付けます。詳細は、下記アドレスにあります。

http://aohakobe.ipc.chiba-u.ac.jp/misc/JP-MBone/

4.2 最新技術動向

最新技術に対する取り組みは以下のテーマごとにさまざまな団体が活動しています。

●IDMR (Inter-domain Multicast Routing)
Routing Protocol、IGMP など広域のマルチキャストルーティングの研究
http://www.cs.ucl.ac.uk/ietf/idmr/
idmr-request@cs.ucl.ac.uk

●MALLOC (Multicast Address Allocation)
MASC、AAP、MADCAP などマルチキャストアドレスの割り当てに関する研究
http://www.aciri.org/malloc/
malloc-request@catarina.usc.edu

●MBoneD (MBONE Deployment)
MBone の普及活動
http://antc.uoregon.edu/MBONED/
mboned-request@ns.uoregon.edu

●PIM(Protocol Independent Multicast)
PIM のプロトコルの標準化
http://netweb.usc.edu/pim/
pim-request@catarina.usc.edu

●MSDP(Multicast Source Discovery Protocol)
プロトコルの開発
http://www.ietf.org/html.charters/msdp-charter.html
msdp@network-services.uoregon.edu

●BGMP(Border-Gateway Multicast Protocol)
BGMP の議論
http://netweb.usc.edu/bgmp/
bgmp-request@catarina.usc.edu

●MMUSIC (Multiparty Multimedia Session Control)
SAP、SDP、SIP、RTSP のほかマルチキャストに限らず IP テ

レフォニなども研究

confctrl-request@isi.edu

●AVT (Audio/Video Transport)

AV protocol format (RTP、RTCP)、Mbone session **情報など**

http://www.cs.columbia.edu/~hgs/rtp/

rem-conf-request@es.net

●RMT (Reliable Multicast Transport)

RMRG(The Reliable Multicast Research Group)**で** RMT **のプロトコルを議論するために最近できた団体**

http://www.east.isi.edu/RMRG/

rm-request@irtf.cs.berkeley.edu