■3 群（コンピュータネットワーク）- 3 編（ネットワーク層）

6 章 Multicast Routing

（執筆者：安川 健太）

■概要■

本章では，マルチキャストルーティングのプロトコルについて解説を行う．

マルチキャストルーティングのプロトコルは，マルチキャストツリーの構成法の違いと，対象とする受信者の密度の違いにより分類される．一つ目の分類項目であるマルチキャストツリーの構成法としては，送信元ツリー（Source Tree）型と呼ばれる，あるマルチキャストグループの送信者から受信者までの最短の Shortest Path Tree を用いる方法と，共有ツリー（Shared Tree）型と呼ばれる，一つのマルチキャストグループに対して共有パスを含む配送経路を利用する方法の二つが存在する．二つ目の，対象とするマルチキャスト受信者の密度の違いによる分類としては，マルチキャスト受信者が密集している場合を対象とするデンスモード（Dense Mode）プロトコルと，離散的に存在している場合に通信を効率的に行うことを目的としたスパースモード（Sparse Mode）プロトコルの二つが存在する．

以上の分類軸に留意し，既存の IP ネットワークにおけるマルチキャストルーティングプロトコルを分類した結果を表に示す．本章では，表のうち，代表的なプロトコルである DVMRP と PIM を取り上げ，解説を行う．

表 6·1　マルチキャストルーティングプロトコルの分類

	Source Tree	Shared Tree
Dense Mode	DVMRP (Distance Vector Multicast Routing Protocol), PIM-DM (Protocol Independent Multicast-Dense Mode)	
Sparse Mode	PIM-SSM (PIM-Source Specific Multicast), MOSPF (Multicast Open Shortest Path First)	PIM-SM (PIM-Sparse Mode), CBT (Core Based Trees)

【本章の構成】

6-1 節において DVMRP を，6-2 節において PIM を取り上げ，それぞれ解説を行う．なお，PIM には，表にあるように，デンスモードの PIM-DM と，スパースモードの PIM-SM が存在する（PIM-SSM は PIM-SM に内包される）．6-2 節では，それぞれを取り上げて説明するが，PIM-DM は，6-1 節で述べる DVMRP と類似点が多いことから，6-2 節の説明は，6-1 節を参照しつつ行っており，PIM-SM に重きを置いている．したがって，6-2 節の PIM-DM の部分を読む前に，6-1 節を読むことを推奨する．

■3 群 - 3 編 - 6 章

6-1 DVMRP

（執筆者：安川 健太）

DVMRP（Distance Vector Multicast Routing Protocol）は，マルチキャストの経路制御専用の距離ベクトル型ルーティングプロトコルとして，RFC1075 で定義された．ユニキャスト用の距離ベクトル型ルーティングプロトコルである RIP（Routing Information Protocol）を元に開発されたため，両者の動作には類似点がある．RIP や OSPF（Open Shortest Path First）と同様，Interior Gateway Protocol（IGP）であり，AS（Autonomous System）内のルーティングにのみ用いられる（3 章参照）．

RIP を元に開発が行われたものの，DVMRP と RIP には重要な相違点がある．それは，RIP が送信者から受信者への経路制御を目的とするのに対し，DVMRP は，受信者から特定の送信元アドレスへの経路を把握することを目的とする点である．なぜなら，送信元からの配信ツリー（送信元ツリー）を構成することが，DVMRP の役割であるからである．すなわち，各ルータが，自身の相対的位置から送信元への最短経路を DV アルゴリズムにより把握することで，各マルチキャストグループごとの送信元ツリーを構築することが，DVMRP の目的である．

RFC1075 に定義された DVMRP v1 では，このような送信元ツリーの構築を，TRPB（Truncated Reverse Path Broadcasting）というアルゴリズムにより実現する[2)]．その後，DVMRP v3 において，TRPB よりも適切なツリー構築を可能とする RPM（Reverse Path Multicasting）を利用するよう拡張されている [2)]（RFC 1075 でも RPM の利用は示唆されているが，実験的（Experimental）とされている）．なお，DVMRP v3 は，本稿執筆時点でも RFC にはなっておらず，最新の Internet Draft である[3)]についても期限切れとなっているが，DVMRP v3 の実装は製品レベルでも多く存在する．

DVMRP のもう一つの特徴は，プロトコル自体がトンネリングをサポートしていることである．この機能により，DVMRP ルータは，マルチキャストをサポートしないネットワークをバイパスし，孤立する複数のマルチキャストドメインを接続することが可能である．この機能は，マルチキャストをサポートしないネットワークを通じて MBone に接続する際に重宝された（DVMRP では，トンネリングで用いる仮想インタフェースも，物理インタフェースも同等に扱われる．以降で単にインタフェースと書いた場合には，両方を含むことに注意されたい）．

本節ではまず, DVMRP の動作の中核である TRPB について, その元となった RPB（Reverse Path Broadcasting）から始めて解説を行った後，DVMRP v3 で採用されている RPM を紹介する．その後，TRPB 及び RPM が動作するために必要なルーティングテーブルと，その構築のためのメッセージ交換について解説を行う．最後に，DVMRP の経緯と今後について簡潔に述べ，本節を閉じる．

6-1-1 DVMRP におけるマルチキャストツリーの構築

以降では，DVMRP のマルチキャストツリー構築アルゴリズムを，RPB, TRPB, RPM の順に説明する．図 6･1 に，説明を補助するネットワークトポロジの例を示したので，適宜参照されたい．なお，経路情報交換については後述することとし，ここでは各 DVMRP ルータ

は，必要な送信元への経路情報を取得済みであるものとする．また，以降で単にルータとある場合には，DVMRP ルータを指すものとする．

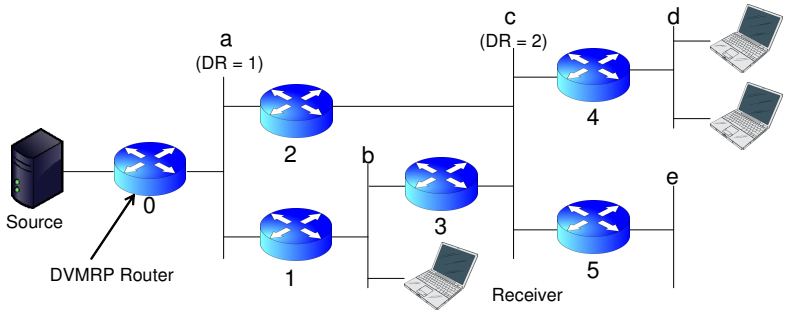

図 6･1　ネットワークトポロジの例．

（1）TRPB の元となる RPB

一般に，マルチキャストパケットを受け取ったマルチキャストルータは，

1. そのパケットを転送するか否か
2. 転送すべきネットワークインタフェース

の二つの判断を行う必要がある．

RPB では，まず 1. の判断を，パケットを受信したインタフェースが，そのパケットの送信元への最短経路上にあるかどうかを判定する．この判定は，Reverse Path Forwarding（RPF）チェックと呼ばれ，RPF チェックの結果，パケットの送信元への最短経路上にあると判断されたインタフェースは，RPF インタフェースと呼ばれる．RPB では，パケットを受信したインタフェースが，そのパケットの送信元にとっての RPF インタフェースがである場合にのみ，ルータはそのパケットを転送する．すなわち，これは送信元に戻る経路を考えることに相当し，これがアルゴリズムの名前の由来である．

2. の判断は，各インタフェースに接続されるネットワークが，自身にとって Child ネットワークであり，自身がその Child ネットワークにおける Designated Router（DR, 代表ルータ）であるかどうかを判定することで行われ，両者が共に真であるインタフェースについてのみ，ルータはパケットを転送する．

まず，Child ネットワークとは，あるマルチキャストツリーにおいて，自身よりも 1 段下位に位置するネットワークを指す．図 6･1 において，ネットワーク c は，ルータ 2, 3 にとっての Child ネットワークである．逆に，パケットを受信したインタフェース，すなわち，RPF インタフェースに接続されるネットワークは Parent ネットワークといえ，当然パケットの転送は行わない．

DR とは，同じネットワークに接続されたルータのうち，当該マルチキャストツリー上，最も送信元に近いルータを指す．各ルータは，それぞれのもつルーティングテーブルを参照し，次のように DR を選出する．

1. 接続されたネットワークにおいて，送信元に最も小さいメトリックで辿り着けるルータを見つける
2. 上記手順で複数のルータが見つかった場合，アドレスが最も小さいルータを選択する

図 6･1 において，各ネットワークのメトリックが等しいとすると，ネットワーク c の DR はルータ 2，ネットワーク a の DR はルータ 1 である．以上の手順で，各ルータはネットワークごとの DR を各自判断可能であることから，先述のパケットを転送するインタフェースの決定は各ルータで独自に実行できる．

上記の判定を行った上でパケットを転送することで，送信元ツリーに従ったマルチキャストルーティングが実現される．また，2. の判断に際して，DR のみが Child ネットワークへのパケット転送を行うことから，同じネットワークに二つ以上同じパケットが転送されることはない．これが RPB の特徴である．図 6･2 に，赤色の点線で RPB においてパケットが転送されるリンクを示した．

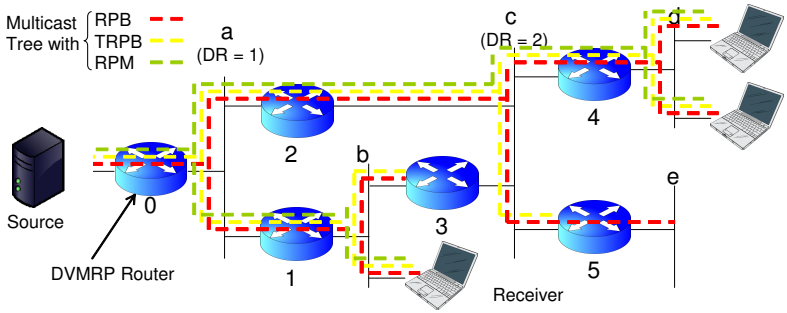

図 6･2　RPB, TRPB, RPM により構成されるマルチキャストツリーの比較．

(2) DVMRP の基本となるアルゴリズム TRPB

RPB では，受信者が存在するか否かにかかわらず，パケットが転送される．このため，受信者のいないマルチキャストパケットがツリーの末端まで配送されるという無駄が生じ得る（図 6･2 のネットワーク e 参照）．この点を改良し，受信者の存在しないネットワークにはパケットを転送しないように拡張した RPB，すなわち，マルチキャストツリーの葉（Leaf）の部分の切り詰めを行うアルゴリズムが，DVMRP v1 で採用された TRPB である．

ここではまず，Leaf ネットワークという用語の定義を確認する．RFC1075 では，ある送信元に対し，どのマルチキャストルータからも Parent ネットワークとはならないネットワークを Leaf と呼ぶことが述べられていることから，本節でもこの定義に従う．図 6･1 において，ネットワーク d, e は Leaf ネットワークといえる．

前述の葉の切り詰めを行うため，TRPB では，RPB で行った Child ネットワークの認識に加え，Leaf ネットワークの検出が必要となる．DVMRP では，Leaf ネットワークの検出を，経路情報の交換において，各ルータに，Parent ネットワークに対して，そのネットワークを送信元への次ホップとして利用していることを広告させることで実現する．すなわち，各ルー

タは，あるインタフェースについて，どの隣接ルータからも Parent ネットワークとして利用している旨の通知がなければ，そのインタフェースに接続されたネットワークは Leaf であると判断できる．

DVMRP では，この通知を，各ルータに，次ホップとして利用しているルータに対して，その次ホップへの経路をメトリック無限大として広告することで行う．この動作は，Poison Reverse と呼ばれ，本来 RIP における，Split Horizon 問題を解決する一つの手段として用いられる手法である（RIP 及び，Split Horizon 問題については，マルチキャストルーティングから離れるため，3 章 1 節及び，参考文献 4, 5) を参照されたい）．

図 6･1 のネットワークでは，ルータ 0, 1, 2 は他のルータから見て次ホップになっているため，Poison Reverse を受ける．これに対し，ルータ 3, 4, 5 は他のルータの次ホップになっていないため，Poison Reverse を受け取らず，自身の Child ネットワークが Leaf ネットワークであると判断できる．なお，図 6･1 において，ルータ 3 はルータ 4, 5 にとっての次ホップのように見えるが，ネットワーク c において DR ではないため，実際には次ホップではないことに注意されたい．

以上のようにして検出した情報を元に，受信者のいない Leaf ネットワークへのパケット転送を防ぐのが TRPB の特徴である．図 6･2 に，黄色の点線で TRPB で構築されるマルチキャストツリーを示した．

（3）TRPB をより効率化した RPM

TRPB では，Leaf ネットワークに受信者が存在しない場合においても，Leaf ネットワークに接続するルータまでは，パケットが配信されていた．言い換えれば，送信元から末端のルータ（枝）までは常にパケットが配信されてしまう（図 6･2 のルータ 3, 5 参照）．これは，受信者の少ない状況，あるいは偏った状況では通信リソースが無駄に使用されることを意味する．これを改善したアルゴリズムが，RPM である．

RPM では，最初のツリー構築は TRPB に従って行われる．その後，Child ネットワークがすべて Leaf ネットワーク，かつ，それらのネットワーク上に受信者がいないルータは，Prune（刈り取り）メッセージ（RFC 1075 では NMR（Non-Membership Report））を Parent ネットワークに発し，当該マルチキャストパケットの転送停止を依頼する．Prune メッセージを受け取ったルータは，自身を送信元への次ホップとするすべてのルータから Prune メッセージを受信した時点で，その Parent ネットワークに Prune メッセージを送るとともに，当該マルチキャストパケットの転送を停止する（図 6･3）．なお，自身を次ホップとするルータのリストは各ルータからの Poison Reverse により構築可能である．この手順を再帰的に繰り返すことで，マルチキャストツリーはある時点での受信者の存在状況に合わせて最適化される．これが RPM の特徴である．図 6･2 に，緑色の点線で RPM により構築されるマルチキャストツリーを示した．

なお, 各 Prune メッセージには有効期限が設けられており, 有効期限が切れた時点で, パケットの転送は再開され，受信者の有無によらないパケット転送，すなわち，枝までの Broadcast が起きる．よって，その時点においても受信者がいないルータは，再度 Prune メッセージを送信することとなる．RPM では，このような Broadcast と Pruning のサイクルを繰り返すことで，マルチキャストツリーの更新を行う．また，Prune メッセージを送った後，その有効期限が切れる前に新たな受信者が現れた場合には，受信開始までの遅延を回避するため，ルー

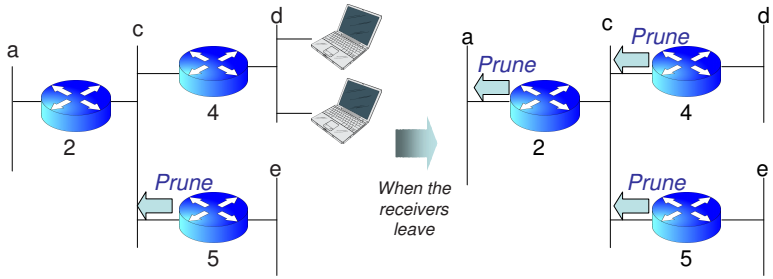

図 6･3　RPM における Prune メッセージ送信の例．

タは Graft（接ぎ木）メッセージ（RFC1075 では NMR Cancel）を送信し，パケット転送を再開させる．なお，Graft メッセージの損失による遅延を防ぐため，Graft メッセージ受信時には Graft Ack メッセージによる確認応答が行われ，必要に応じて Graft メッセージは再送される．これらの仕組みを用いることで，受信者の増減に対応した，動的なマルチキャストツリーの管理を実現する．

6-1-2 アルゴリズムの動作に必要なルーティングテーブルとその構築

これまで，DVMRP が実装するマルチキャストツリー構築アルゴリズムについて述べてきた．次に，それらのアルゴリズムが動作するために必要なルーティングテーブルと，その構築について述べる．

以下は，RFC1075 に示された，DVMRP の動作に用いるルーティングテーブルが含むべき情報のリストである．

- 目的地アドレス（マルチキャストパケットの送信元アドレス）
- 目的地アドレスのサブネットマスク
- 目的地アドレスへの次ホップルータ
- 次ホップルータへのインタフェース
- Child ネットワークに接続されたインタフェースのリスト
- Leaf ネットワークに接続されたインタフェースのリスト
- 各インタフェースの DR のアドレス
- 各インタフェースの DR でないルータのアドレス
- エントリの経過時間を示すタイマ
- エントリの状態を表すフラグのセット
- メトリック
- 無限大を意味する値

DVMRP では，上記の経路表の構築を，RIP と同様の定期的な経路情報交換と，DV 型経路制御アルゴリズムにより行う．RIP と同様，各ルータは，自身が知る各マルチキャスト送信

元ネットワークについて，ネットワークアドレス，ネットマスク，メトリック値などの情報を交換しあう．また，Leaf ネットワークの検出及び，自身を次ホップとするルータのリストの取得を可能とするために，各ルータは Poison Reverse を行う．これが RIP から継承された部分である．DVMRP では，上記の情報の交換を，IGMP データグラムを用いて行う．

IGMP は，5 章 2 節にあるとおり，TCP や UDP と並んで IP の直上に位置するプロトコルである．その IGMP データグラムにおいて，タイプヘッダが 0*x*13 にセットされたものが，DVMRP メッセージを意味する．すなわち，DVMRP のメッセージは，タイプヘッダが 0*x*13 にセットされた IGMP データグラムのペイロードとして運ばれる．

上記のようにして，各ルータはルーティング情報の交換を行うが，そのタイミングと内容は，下記のように定められている．

- Route Report Interval（RFC1075 では FULL_UPDATE_RATE）というパラメータで指定された一定時間の経過時 – すべてのルーティング情報を交換する．
- 経路変更検出時 – このときは，変更があった経路についてのみ情報を送信する．
- ルータが再起動あるいは新たに加わったとき – すべてのルーティング情報を交換する．
- ルータ終了時 – 自身がもつ全経路のメトリックを無限大に設定し，全インタフェースに広告する．

以上のようにして，各ルータは先述のルーティングテーブルの構築を行うことができ，それを基にした TRPB 及び RPM による送信元ツリーの構築が可能となる．

6-1-3 DVMRP のまとめ

DVMRP の登場は 80 年代後半と非常に早く，インターネットの初期から存在するプロトコルであり，その後のマルチキャストプロトコルにも大きな影響を与えた．特に，次節で紹介する PIM-DM はその影響を強く受けている．

DVMRP は，文中で述べたとおり，Broadcast と Pruning を繰り返すオーバーヘッドを伴うことや，RIP と同様，経路収束速度が遅いなどの問題を抱えていることは確かである．しかし，トンネリングをサポートしていることや，独自のルーティングテーブルを作成するため，ユニキャストとマルチキャストで異なる配信経路を提供できることなど，実用的な利点は確かに存在する．以上のことから，RFC は古く，Internet Draft にその仕様を頼らなければならないものの，依然として利用され続けるプロトコルだといえる．IETF のメーリングリスト中で Deering 氏が，「DVMRP の RFC は死んだが，プロトコル自体は生きている」と述べているが，この言葉がこれを表わしている．

■参考文献

1) D.Waitzman, C. Partridge, and S.E. Deering, “Distance Vector Multicast Routing Protocol,” RFC 1075 (Experimental), November 1988.
2) S.E. Deering and D. Cheriton, “Multicast Routing in Datagram Internetworks and Extended LANs,” ACM Transactions on Computer Systems, Vol. 8, No. 2, pp. 85–110, May 1990.
3) T. Pusateri, “Distance Vector Multicast Routing Protocol,” Internet Draft draft-ietf-idmr-dvmrp-v3-11, October 2003. Expired at April 2004.

4) C.L. Hedrick, “Routing Information Protocol,” RFC 1058 (Historic), June 1988. Updated by RFCs 1388, 1723.
5) G. Malkin, “RIP Version 2,” RFC 2453 (Standard), November 1998. Updated by RFC 4822.

■3 群 - 3 編 - 6 章

6-2 PIM

（執筆者：安川 健太）

本節では，Protocol Independent Multicast（PIM）について解説を行う．本章冒頭で述べた通り，PIM には，PIM-Dense Mode（DM）と PIM-Sparse Mode（SM）が存在する．このうち，PIM-DM は RFC3973 [1]で，PIM-SM は RFC4601 [2]で，それぞれ定義されている．

PIM の特徴として，プロトコル内部にネットワークトポロジを把握するための経路情報交換メカニズムを持っていない点が挙げられる．これは，前節で述べた DVMRP が，専用の経路情報交換を行うことと対照的である．PIM は，ユニキャストのルーティングプロトコルにより得られた経路情報を参照して，マルチキャストツリーの構築を行う．すなわち，経路情報の結果のみ参照するため，特定のプロトコルに依存しない．この点が，PIM という名称の由来である．このため，DVMRP のように，ユニキャストとマルチキャストで異なる配信経路を利用できないことや，プロトコルで参照できる情報が限られることなどの制限は課されるが，経路情報交換メカニズムの実装や，そのためのオーバーヘッドは不要だという利点がある．

なお，PIM-DM と PIM-SM は，上記の特徴を共有するが，両者の動作及び原理は大きく異なる．本章冒頭で述べたように，PIM-DM は，DVMRP と同じ送信元ツリー / デンスモードに分類されるプロトコルであり，類似点も多いことから，本節ではまず PIM-DM について解説を行い，その後，PIM-SM について解説を行う．

6-2-1 PIM-DM

PIM-DM は，DVMRP から強く影響を受けたプロトコルであり，両者は非常によく似た特徴をもっている．PIM-DM は，DVMRP と同様，Broadcast と Pruning に基づく動作を行うえ，RPF チェックや，Graft メッセージによる Prune の解除を行う点なども DVMRP と同様である．

一番の違いは，PIM-DM は，独自の経路情報交換メカニズムを内包しないため，DVMRP に比べ，利用可能な情報が制限される点にある．特に重要なのは，PIM-DM では，各ルータが，自身を次ホップとして利用する下位ルータの存在を経路情報から取得できない点である（DVMRP では，これを Poison Reverse を用いて経路情報の一部として交換する）．この点を PIM-DM では Prune メッセージを打ち消す Join メッセージを用いることで解決している．また，ルータ間で同じ経路情報を共有しないため，DVMRP と同じ仕組みでの DR 選出ができないため，一つの Child ネットワーク上に複数の PIM ルータがいた場合，PIM Assert という特別なメッセージ交換を行い，そのネットワークにパケットを転送するルータを決定する．

以降で，PIM-DM の動作概要を説明する．なお，説明のため，(S, G) という表記を持って，マルチキャストグループ G とその送信元 S というペアを表すものとする．

各 PIM-DM ルータは，マルチキャストパケット受信時，RPF チェックを行った後，そのパケットの (S, G) ペアについて，マルチキャストパケットを受信するか否かを意味する，“Prune 状態” を作成し，保持する．Prune 状態の初期状態は”Forwarding”（転送中）である．その後，ルータは，自身の Child ネットワークにそのグループのメンバがいるか否かを判断し，メンバがいればその状態のままパケットの転送を続けるが，メンバがいない場合，RPF インター

フェースから Prune メッセージを送信するとともに，(S, G) ペアの Prune 状態を”Pruned”（Pruning 中）に設定する．

Child ネットワークをもつ PIM-DM ルータは，自身の RPF インタフェース以外のインタフェースから Prune メッセージを受け取った場合，Prune 状態を”Prune Pending” に設定し，Prune Pending Timer（PPT）を設定する．PPT が切れるまでの間に：

- **Child ネットワークから Join メッセージを受け取った場合**：(S, G) ペアのメンバが下位にいることから，Prune 状態を “Fowarding” に設定し，PPT を解除，パケット転送を続ける．
- **Child ネットワークから Join メッセージを受け取らなかった場合**：Prune 状態を”Pruned” に移行し，Parent ネットワークに Prune メッセージを送信する．

上記手順を，図 6･3 と同じ状況において行った場合の例を，図 6･4 に示す．上記手順を再帰的に繰り返すことで，マルチキャストツリーは最適化される．なお，Prune メッセージを送信した後，同じインタフェースで Join メッセージを受信した場合，そのルータは Prune Limit Timer を設定し，一定時間 Prune メッセージの送信を停止する．

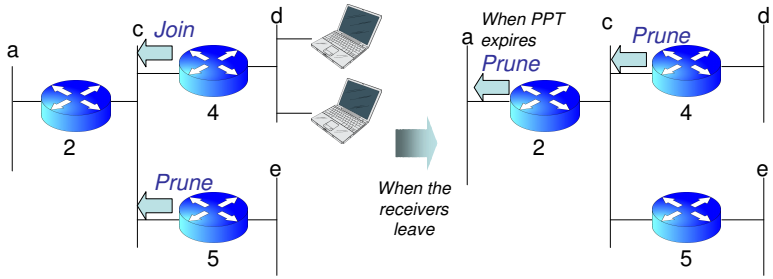

図 6･4　PIM-DM における Prune メッセージ送信の例

PIM-DM でも，DVMRP と同様，Prune 状態は限られた有効期限をもち，その有効期限が切れた際には，Broadcast が再開される．Prune 状態が “Pruned” の時に新たに (S, G) ペアの受信者が現れた場合，それを検出した末端の PIM-DM ルータは，RPF インタフェースに向けて Graft メッセージを送信し，Pruning を解除する．Graft メッセージには，Graft Ack メッセージを用いた受信確認応答と再送が行われ，Graft Ack メッセージの受信までは Prune 状態は”AckPending” に留まり，転送は保留される．

以上が，DVMRP との違いに注目した，PIM-DM の概要である．以上からわかるとおり，PIM-DM は，DVMRP の経路情報交換を省き，それによって不足する制御を Join メッセージや関連するタイマを用いて補ったプロトコルであると考えることもできる．

6-2-2 PIM-SM

PIM-SM は，本章冒頭で述べたとおり，受信者の密度が希薄なスパース型の状況に適した

プロトコルとして設計されている．そのため，DVMRP や PIM-DM とは異なり，Broadcast と Pruning を繰り返すのではなく，明示的なマルチキャストグループへの Join メッセージが送信されたネットワークにのみマルチキャストパケットの配信を行う（PIM-DM の Join メッセージとは意味が異なることに注意）．すなわち，PIM-DM のような配信方法を Push 型と呼ぶのであれば，PIM-SM は Pull 型の配信を行う仕組みだといえる．

また，PIM-SM は，PIM-DM や DVMRP とは異なり，共有ツリー型に分類されるプロトコルである．共有ツリー型と言われる理由は，PIM-SM では，ネットワーク中に Randezvous Point（RP）と呼ばれるルータが存在し，その RP でマルチキャストストリームを束ね，そこを起点に受信者に配信する場合があるためである．図 6･5 に，ルータ 2 が RP となる場合の，PIM-SM におけるマルチキャストツリーを示す．

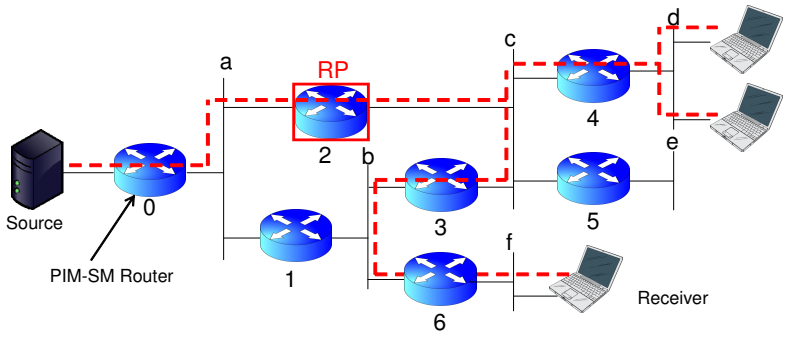

図 6･5　PIM-SM における RP を用いたマルチキャストツリー

RP を用いた共有ツリーを利用する理由を述べる前に，まず，マルチキャストグループへの参加の仕方として，送信元指定がある場合とない場合の 2 通りが存在することについて述べておく．

マルチキャストが生まれた当初は，多対多の通信を許容するため，送信元を指定しないマルチキャストグループへの参加，（以下，(*, G) Join）も含めて検討が行われていた．すなわち，宛先アドレスが G であるマルチキャストパケットであれば，送信元によらず受信するという参加の仕方である．主に，多地点遠隔会議などの用途が例として挙げられる．これに対し，1 対多の通信を行う，送信元を指定したマルチキャストグループへの参加，（以下，(S, G) Join）も昨今では利用例が多い．マルチキャストを用いた放送型サービスなどが例として挙げられる．

前者のような (*, G) Join を行う場合，受信者及びマルチキャストルータは，あらかじめ送信者のアドレスを知らないことから，Join メッセージを送ろうにも，送るべき宛先を特定できない．そのような仮定の元でも，明示的な Join メッセージによるマルチキャストツリーの構築を行うために，RP が用いられる．すなわち，送信者側ルータ，受信者側ルータともに，グループアドレス G について，RP となるルータを知っていれば，送信者は RP に向けて G 宛

のマルチキャストパケットを送信し，受信者は RP に向けてグループアドレス　G への Join メッセージを送ることで，(*, G) Join が実現できる．以降でこの手順をより詳しく説明する．

（1）PIM-SM における (*, G) Join

説明に入る前に，Designated Router （DR, 代表ルータ）という言葉を再定義をしておく．PIM-SM における DR とは，一つのサブネット上に複数の PIM-SM ルータが存在した場合に，それらの代表として PIM-SM プロトコルの処理を行うルータを指す（DVMRP における DR とは異なり，実際の転送処理を行うルータとは異なる場合がある．実際の転送処理を行うルータは，PIM-DM と同様，PIM Assert メッセージの交換により定まる）．また，以降の説明に共通する前提条件として，各ルータは，マルチキャストグループアドレスに対応する RP のアドレスを知っている，サブネット上に (*, G) に Join している受信者はいないことの 2 点を仮定する．以上の緒準備の元，PIM-SM における (*, G) Join の流れを説明する．

1. マルチキャスト受信者がグループ G の受信希望を発する（受信希望の通知は，IGMP や MLD の利用が考えられるが，他の方法でも構わない）．

2. 受信希望が発されたサブネットの DR は，マルチキャストグループ G を担当する RP を宛先として，(*, G) Join メッセージを RPF インタフェースから送信する．

3. (*, G) Join メッセージは，各ルータによって Hop-by-Hop で伝送されていく．この過程で，各ルータは，(*, G) Join メッセージを受け取ったインタフェースと，(*, G) に Join していることを記録する．

4. (*, G) Join メッセージが RP に到達するか，すでに (*, G)Join しているルータに受信される時点で (*, G) Join は完了される．

以上の手順の後は，受信者のローカルサブネットの DR から，RP に至る経路上の各ルータが (*, G) Join の状態と，受信者のいるインタフェースを記録しているため，RP からのネイティブマルチキャスト配信が可能となる．なお，動作のうえでは，(*, G) Join は，RP に対して (S, G) Join を行うこと（言わば，(RP, G) Join）と等しいことを付記しておく．

なお，これまで，送信元 S から RP へのパケット配送の実現方法について説明していないが，S から RP への経路は，マルチキャストツリーを根から葉に向けてたどる必要があり，S には不可能であるため，何らかの方法でこれを実現する必要がある．PIM-SM では，RP Tree（RPT）と Register-stop というフェーズを経て，送信元から RP へのパケット配送を実現する．以降でこれらについて説明する．

また，この時点では，配信経路は RP を介した共有ツリーになっている．共有ツリー型の配信は，送信元から RP までの経路が共有されるため，スケーラビリティの観点で優位であるが，原理的に各 (S, G) ペアについて Shortest Path Tree（SPT）を提供しないため，不必要な遅延やトラヒックの増加が問題となる可能性がある．図 6･5 の例においても，図の下部に位置する受信者は，ルータ 1 を経由した方が最短経路でパケットを受信できる．

そのような非効率を検出した場合，PIM-SM では，各ルータが送信元からの SPT を用いた配信経路に切り替えるという選択肢を持っている．以降では，そのような SPT への切り替えを行う，SPT フェーズについても解説する．

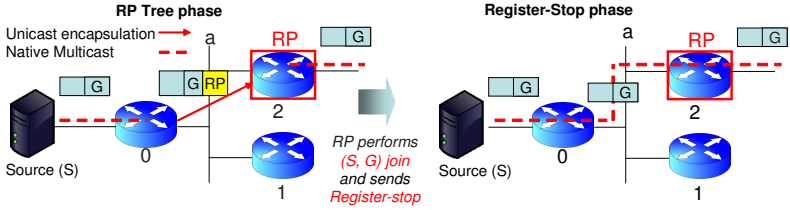

図 6·6　RP Tree フェーズから Register-Stop フェーズへの移行

（2）PIM-SM におけるマルチキャストツリーの更新

ここでは，RPT，Register-Stop，SPT フェーズについて，それぞれ説明する．なお，下記の三つのフェーズは，送信元 S が特定された時点で独立に実行可能であることから，実際には同時に，独立して実行され得ることを付記しておく．

1. **RPT フェーズ （図 6·6 左参照）**
   このフェーズは，送信元 S がマルチキャスト配信を開始した直後に対応する．この時点では，RP から S への経路が不明なうえ，経路上のルータも Join 状態を記録していないため，S からのマルチキャストパケットは RP に到達できない．そのため，送信者 S から送られたグループ G 宛のマルチキャストパケットは，S の接続されたサブネットの DR により，RP に宛てたユニキャストパケットにカプセル化（Encapsulation）され，RP まで送信される．RP では，受け取ったユニキャストパケットから元のマルチキャストパケットを取り出し（Decapsulation），Join しているメンバに向けてネイティブマルチキャスト配信する．以上が，RPT フェーズの概要である．

   なお，このフェーズでの，S から RP へのカプセル化パケット送信のプロセスは Registering，カプセル化されたパケットは PIM Register パケットと呼ばれる．当然ながら，このフェーズでは，配信に際して，ユニキャストカプセル化のオーバーヘッドが伴う．

2. **Register-Stop フェーズ（図 6·6 右参照）**
   このフェーズは，RP が S からのカプセル化された Register パケットに代わり，ネイティブマルチキャストパケットを受け取ることを試みるフェーズである．なお，一般に RP はこの試みを行うが，RPT フェーズを続けてもよい．

   RP は S からのユニキャストパケット（PIM Register パケット）を受け取った時点で，送信元 S を特定できることから，送信元を指定した，(S, G) Join を行うことができる．すなわち，この (S, G) Join が成功すれば，ネイティブマルチキャストの受信が実現できる．(S, G) Join の動作は，先述の (*, G) Join を，RP から S に向けて行う場合と同じであるため，省略する．

   RP からの (S, G) Join が完了すると，RP は，S から，カプセル化されたユニキャストパケットと，ネイティブマルチキャストパケットの両方を受け取る状態になる．その

状態になった場合は，届くユニキャストパケットを廃棄しつつ，S のローカルサブネットに位置する DR に向けて，Register-Stop メッセージを送る．Register-Stop メッセージが届いた時点で，DR はユニキャストカプセル化を中止し，RP はネイティブマルチキャストのみを受け取る状態になる．

3. **SPT フェーズ （図 6･7 参照）**
   このフェーズは，各 DR が，RPT の利用が非効率だと判断した場合に実行されるオプションで，フェーズ完了後は当該ルータは，RPT の代わりに，S からの SPT を利用するようになる．

   このフェーズの実行をする DR は，送信元 S への (S, G) Join を試みる（RP Tree フェーズ以降，マルチキャスト配信は行われていることから，S の特定は可能）．(S, G) Join の動作は，先述の (*, G) Join を，DR から S に向けて行う場合と同じであるため，省略する．

   (S, G) Join が完了すると，DR は，S からの SPT と RPT 両方からパケットを受信するようになる．その場合，DR は RPT からのパケットを廃棄しつつ，(S, G) マルチキャストの送信停止を意味する，(S, G) Prune メッセージを，RP に向けて送信する．これを (S, G, rpt) Prune と呼ぶ．(S, G, rpt) メッセージは，(*, G) Join メッセージと同様に RP に向けて Hop-by-Hop 伝送されるが，経路上の PIM-SM ルータは，Join とは逆に，(*, G) Join 状態をキャンセルしながらメッセージを伝搬させていく．

以上の手順をそれぞれのルータが独立に行うことで，必要に応じたマルチキャストツリーの効率化が図られる．

なお，(*, G) Join は複数の送信元からの受信を想定しており，送信元が複数存在する場合には，送信元ごとに上記三つのフェーズが実行されることに注意されたい．

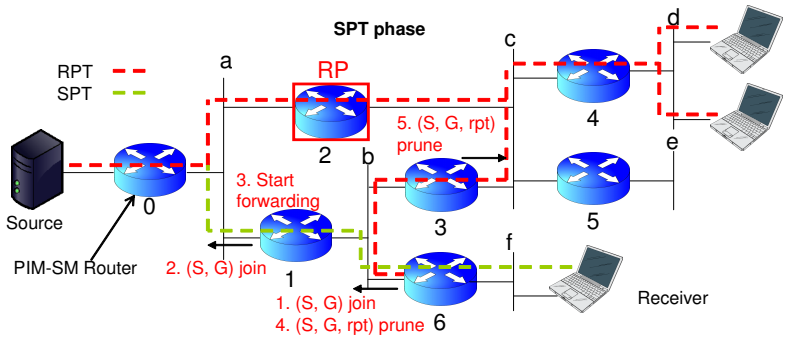

図 6･7　SPT フェーズ完了直前の状態

（3）RP の発見

これまで，各ルータが，それぞれのマルチキャストグループに対応する RP を知っている

ことを前提として説明を行ったが，実際には各ルータは何らかの方法で RP を知る必要がある．これには，静的なコンフィギュレーションの投入と，Bootstrap Router (BSR) を用いた動的コンフィギュレーションの 2 通りがある．

動的コンフィギュレーションを行う場合，まず，各 PIM ドメインごとに，一つのルータを BSR として選出する．そして，その他の RP の候補となるルータは自身の能力を BSR に定期的に通知する．BSR はその情報をもとに，マルチキャストグループごとに RP となるルータを選択し，選択した RP のアドレスを同 PIM ドメイン中の全ルータに広告する．なお，RP の選択は，グループアドレスをハッシュ関数にかけ，RP 群にマッピングすることで行われる（BSR については，RFC4601 にも記載があるが，RFC5059 により更新されている）．

(4) PIM-SSM

先述のように，マルチキャストは当初から多対多通信を見据えて検討がなされてきたが，放送サービスへの利用などが普及し，マルチキャストの用途として，1 対多通信としての重要性が増している．1 対多通信であれば，マルチキャストを受信する際，送信元も既知であるという仮定を置くことに無理はない．また，IGMP v3 では，従来の (*, G) Join に加え，送信元を指定した (S, G) Join を明示的に行うことが可能になっている．このように，送信元を指定したマルチキャストのモデルのことを Source Specific Multicast (SSM) と呼ぶ [3]．

これまでの説明から分かるように，PIM-SM のサブセットにより，SSM を実現することが可能である．すなわち，受信者は，(*, G) Join を行う代わりに (S, G) Join を行い，PIM ルータも (S, G) Join のみを用いるようにすればよい．そのような PIM-SM のサブセットによる SSM の実現は，PIM-SSM と呼ばれる．

6-2-3 PIM のまとめ

PIM は DVMRP に比べ，標準化がよく行われているものの，一部のベンダへの依存があることが，普及の妨げになっているとの見方がある．しかし，広域で大規模なマルチキャストネットワークを構築する際は，スパースモードのプロトコルが必要になること，同モードのプロトコルとして最も普及しているのは PIM-SM であることなどから，その重要性は高く，今後の動向は興味深い．また，RP に負荷が集中しやすいこと，プロトコルが複雑であることなどが障害の原因になりがちだという指摘もあるが，今後利用頻度が高いと考えられる 1 対多通信であれば，PIM-SSM を利用できることから，そのような場合には，これらは問題になりにくい．

以上より，PIM は今後も有力なマルチキャストルーティングプロトコルとして存在していくものと考えられる．

■参考文献

1) A. Adams, J. Nicholas, and W. Siadak, “Protocol Independent Multicast - Dense Mode (PIM-DM): Protocol Specification (Revised),” RFC 3973 (Experimental), January 2005.
2) B. Fenner, M. Handley, H. Holbrook, and I. Kouvelas, “Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification,” (Revised). RFC 4601 (Proposed Standard), August 2006. Updated by RFC 5059.
3) H. Holbrook and B. Cain, “Source-Specific Multicast for IP. RFC 4607 (Proposed Standard), August 2006.

4) N. Bhaskar, A. Gall, J. Lingard, and S. Venaas, “Bootstrap Router (BSR) Mechanism for Protocol Independent Multicast (PIM),” RFC 5059 (Proposed Standard), January 2008.