分散システム特論レポート

RFC 1075 : Distance Vector Multicast Routing Protocol

情報工学専攻 修士1年 荒木研究室

2000年11月6日

焼山 康礼

1 はじめに

インターネットの普及と共に、動画像配送や音楽データ配信等のマルチメディアデータ通信が活発になっている。一般にマルチメディアデータのデータサイズは非常に大きいものになるため、効率の良いデータ配送方式の確立を行うことが重要である。これに対する一つの解決方法として、IP マルチキャストを用いたマルチメディアデータ通信が挙げられる。IP マルチキャストとは、IP 上で複数のグループメンバに対してデータグラムを効率良くに配送するために提案されたものである。テレビ会議やホワイトボード等の数人以上が一つのデータを共有する場合において マルチキャストを用いると、ネットワーク帯域使用を減少させることができる。

しかし、複数のネットワークセグメントにまたがって マルチキャストデータ配送を行う場合には、マルチキャスト配送木を構築し、マルチキャストパケットを転送できるようにすることが必要である。このために マルチキャスト経路制御プロトコルが これらの処理を行うために 提案された。

本輪講では、マルチキャスト経路制御プロトコルとして最初に提案された DVMRP(Distance Vector Multicast Routing Protocol) の紹介を行う。

2 マルチキャストの概要

2.1 マルチキャストとユニキャスト

図1,2は、server から 複数の user に 同一のデータグラムを送信する際の、データの流れを示している。図1ではユニキャストを用いており、図2ではマルチキャストを用いている。ユニキャスト通信の場合では、複数の同一データが server から各 user 宛に送信されているのに対し、マルチキャスト通信の場合では、server から一つのデータが送信され、中間のルータが必要に応じてデータパケットを複製し、各 client に配送するため、ネットワーク中を流れるデータパケット量は減少する。このように、同一データを複数のホストに同時に配送する場合には、ユニキャスト通信よりもマルチキャスト通信の方が、使用されるネットワーク帯域を減少させ、効率的なデータ配送を行なうことができる。

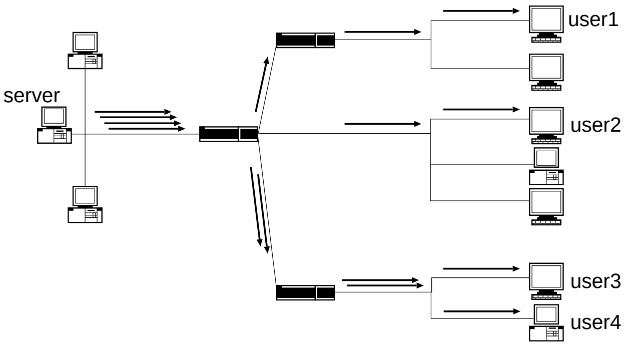

図1: ユニキャスト通信

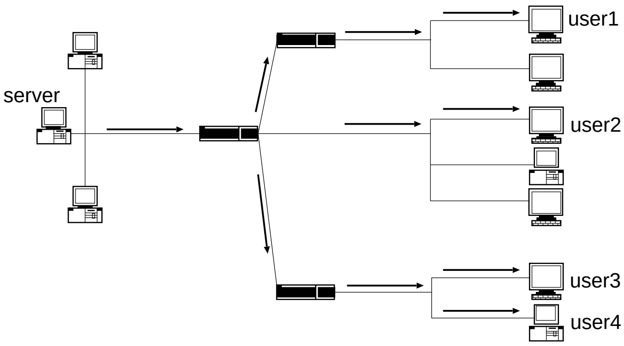

図2: マルチキャスト通信

現在IP マルチキャストは、Mbone(Multicast Backbone) と呼ばれる インターネット上の実験ネットワーク上で動作する。

2.2 マルチキャストの構成要素

1. マルチキャストホストグループ

   IP マルチキャストは、マルチキャストアドレスと呼ばれる特別なアドレスによって取り扱われる。グループは、各マルチキャストセッションを識別する役割を持っている。IPv4 では、上位 4 ビットが 1110(2 進数) で始まるアドレス (クラス D アドレス: 224.0.0.0 ~ 239.255.255.255) が、IPv6 では、上位 1 バイトが ff(16 進数) で始まるアドレスが、マルチキャストアドレスである。

2. 受信ホストのマルチキャストグループへの参加・離脱

   IP マルチキャストでは、マルチキャストパケットを受信するホストをメンバと呼ぶ。受信ホストは、グループに参加 (join) することによってメンバとなり、マルチキャストパケットの受信を行なう。途中からグループに参加 (join) したり、途中で離脱 (leave) することもできる。IP におけるマルチキャストホストグループへの参加・離脱には、IGMP(Internet Group Management Protocol) が用いられる。

   また送信ホストは、グループのメンバになる必要はない。

3. マルチキャストルータ

   複数のネットワークセグメントにまたがってマルチキャスト通信を行うには、中間のルータがマルチキャスト用に経路制御を行い、必要に応じてデータグラムの複製を行う必要がある。これらの一連の動作を定義するものが経路制御プロトコルである。

3 DVMRP(Distance Vector Multicast Routing Protocol)

DVMRP(Distance Vector Multicast Routing Protocol) は、RFC1075 で最初に定義されたマルチキャスト経路制御プロトコルであり、元々ユニキャスト経路制御プロトコル RIP(Routing Information Protocol) から派生したものである。

DVMRP は、Mbone ルータの大部分で使用されている。

3.1 プロトコル

DVMRP メッセージは IGMP メッセージを変更したものである。IGMP メッセージ内のタイプコードを利用し、未使用フィールドをサブタイプとして使用する。DVMRP のメッセージフォーマットを図 4 に示す。

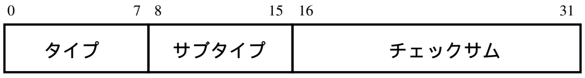

図 3: RFC1075 で定義されている、DVMRP のメッセージフォーマット

- タイプ - 0x13 DVMRP メッセージであることを表す。
- サブタイプ DVMRP メッセージの識別子を格納する。
- チェックサム IGMP と同様の方法で格納される。

RFC1075 で定義されているサブタイプを以下に挙げる。RFC1075 では、ルーティングアルゴリズムとして TRPB(切り詰めリバースパス簿ロードキャスト) を採用するように定義されている。

サブタイプ	説明
1	応答– あて先へのルートを提供する
2	要求– あて先へのルートを要求する
3	非メンバーシップ報告– 非メンバーシップ報告を提供する
4	非メンバーシップの取り消し– 前の非メンバーシップ報告を取り消す

現在の DVMRPv3 では、以下の様に大きく変更されている。これは、ルーティングアルゴリズムとして、RPM(リバースパスマルチキャスト) アルゴリズムを採用したことによる変更である。

サブタイプ	説明
1	Probe – 隣接の発見
2	Report – ルートの交換
7	Prune – マルチキャスト配送木の除去
8	Graft – マルチキャスト配送木の接合
9	Graft Ack – Graft メッセージの Ack

3.2 基本動作

現在の DVMRP では、リバースパスマルチキャスト (RPM) アルゴリズムを導入している。

RPM アルゴリズムに基づいて、任意の 送信元-グループ対 (S,G) についての最初のデータグラムは、パケットの TTL とルータインターフェースのしきい値の許容範囲にあるネットワーク全体に転送される。下流のルータは、直接接続するサブネットにグループメンバが存在しなければ、Prune メッセージを送信元に向けて逆方向に送出する。これにより、メンバが下流にない全ての枝が、マルチキャスト木から削除され、送信元ベースの最短経路木が構築される。一定時間が経過すると、各 (送信元、グループ) 対についてのプルーニング状態は失効し、マルチキャスト木は再構築される。

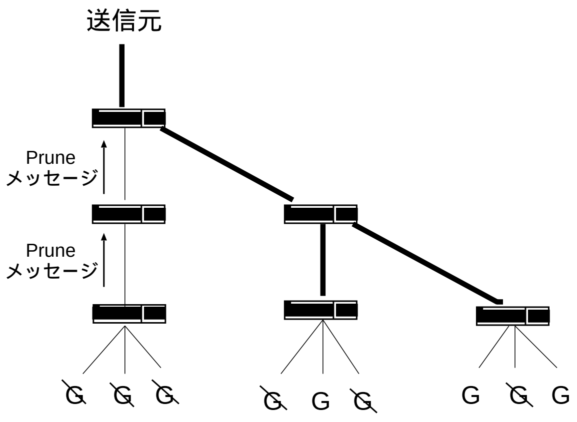

図 4: リバースパスマルチキャスト (RPM)

DVMRP は、直前にプルーニングで削除されたグループの配信木の枝を、接ぎ木 (Graft) する機能も備えている。プルーニングが行われた後で、下流ネットワークに新規メンバが存在することが判明した場合、そのルータは Graft メッセージを送信元方向に送信することで、削除された枝を配信木に戻すことができる。

また、DVMRP はマルチキャストをサポートしていない領域へのマルチキャスト ･データグラム転送技術としてトンネルという技術を利用してる。トンネルは、マルチキャスト転送をサポートしていないネットワーク上にマルチキャストデータグラムを転送する場合に、データグラムをカプセル化し、ユニキャスト ･データグラムとして他の領域にデータグラムを転送する。

3.3 経路情報

ルータは、マルチキャスト配送木を作成するのに必要な経路情報を取得し、保持しておかなければならない。このため ルータは、以下に挙げる二つのテーブルを作成し、必要に応じて更新を行っている。

DVMRP ルーティングテーブル

- Source Subnet(送信元のサブネット)
- Source Netmask(サブネットマスク) 送信元のプレフィックスに対応するサブネットマスク。
- From Gateway(送出ルータ) 特定の送信元プレフィックスの方向にある、直前のルータ。
- Metric(メトリック) 特定のポートに関するコスト。
- Status(状態) 現在のポートの使用可能状況。
- TTL(生存時間) テーブル管理に使用される生存時間の値。
- InPort(受信ポート) ルータの受信トラフィック用インターフェース。
- OutPort(送信ポート) ルータの送信トラフィック用インターフェース。

ルーティングテーブルは、ネットワーク内のすべての可能な送信元プレフィックスまでの最短経路スパニングツリー (リバースパスブロードキャストツ

リー) を表す。ルーティングテーブルには、グループメンバーシップや受信した Prune メッセージについての情報は含まれていないため、DVMRP ルータは転送テーブルを別途作成する必要がある。

DVMRP 転送テーブル

- Source Prefix(送信元プレフィックス) 指定されたグループにマルチキャストデータグラムを送信しているサブネットワーク
- Multicast Group(マルチキャストグループ) マルチキャストデータグラムに宛てられるクラス D IP アドレス。
- InPort(受信ポート) (送信元、グループ) 対における親インターフェース。Prune メッセージを上流に送信した場合には、Pr フラグが与えられる。
- OutPort(送信ポート) 当該の (送信元、グループ) 対についてのマルチキャストデータグラムを転送する際に経由する子インターフェース。ルータがポートの下流ルータから Prune メッセージを受信した場合には、p フラグが与えられる。

3.4 トンネリング

DVMRP では、マルチキャスト ·ルーティングをサポートしていない領域を通して、他のマルチキャスト・ルーティングをサポートしている領域にマルチキャスト ·データグラムを転送するために 「トンネル」 という技術を利用する。トンネルは、LSSR(Loose Source and Record Route) と呼ばれる IP オプションを利用して、マルチキャスト ·データグラムをユニキャスト ·データグラムへカプセル化して転送する。カプセル化を行なうことで、マルチキャストをサポートしていない途中のルータは通常のユニキャスト ·データグラムとして、宛先領域のマルチキャスト ·ルータへデータグラムを転送することができる。

4 DVMRP の実装

DVMRP の実装として、Mbone 上で使用されている mrouted を紹介する。mrouted の最新バージョンは 3.8 である。なお、今回 OS には FreeBSD2.2.8 を利用している。

4.1 ルータの各エントリ情報

mrouted 起動時に、オプションコマンド -d 2 を追加すると、マルチキャストルーティングテーブルや各エントリの内容などを知ることができる。付録に、その実行結果を記載しておく。

4.2 実装とプロトコルの差異

現在の mrouted は RFC1075 で定義された仕様から多くの大きな改良が施されている。RFC1075 では、DVMRP は TRPB(Trancated Reverse Path Broadcasting) アルゴリズムを利用することになっている。しかし、TRPB ではスケーラビリティに問題があるため、現在はよりネットワークを効率的に利用することができる RPM(Reverse Path Multicasting) を利用している。現在の動作している DVMRP は RFC1075 とは多くの点で異なったものになっている。

5 H-DVMRP

マルチキャストネットワークの急速な拡大はルータへの要求を増加させる。DVMRP の現在のバージョンでは、マルチキャストネットワークは一つの平坦なルーティングドメインとみなしている。従って、ルータには Mbone のすべてのサブネットワークの経路制御情報の詳細を保持することが求められる。サブネットワークの数が増え続けると、それに従って経路制御テーブルのサイズや定期的な更新メッセージの交換が増大する。これは、マルチキャストネットワークの拡大に伴い、DVMRP ルータの資源が枯渇し、ルーティングに障害が発生するという問題を引き起こすことになる。このような潜在的な問題を解決するために DVMRP の階層化バージョン H-DVMRP(Hierarchical DVMRP) が現在開発されている。

5.1 マルチキャスト経路制御の階層化の利点

階層的な経路制御では、マルチキャストネットワークはいくつかの領域に分割される。ここで各領域には個別のマルチキャスト経路制御プロトコルを配備できる。これにより、新しいプロトコルを領域単位で利用することができる。

個別のルータの障害が及ぼす影響の範囲は、同一の領域内部で動作するルータに限定される。同様に、領域間の相互接続トポロジーに変更が生じても、その影響は領域間ルータだけに限定される。

また、階層的な経路制御では、ルータは 自分自身の領域内での経路制御の詳細情報のみ保持すれば良いので、ルータに要求される資源を減少させることができる。領域間を接続するルータはお互いの情報を保持していなければならないが、それでも他方の領域の内部の状態まで知る必要はない。

このように、階層化によって ルータが保持する経路情報を削減し、スケーラビリティを高めることができる。

6 おわりに

DVMRP は、良く知られた RIP アルゴリズムに基づいているため 導入が比較的容易である。また、ルータに要求される処理能力やステータス情報が他のマルチキャスト経路制御プロトコルと比べて少ないという利点がある。しかし、DVMRP は RIP の距離ベクトルルーティングに依存するため、経路の収束に時間がかかる。また DVMRP は、ネットワーク全体にマルチキャストトラフィックのフラッドを定期的に行ってマルチキャスト木を際構築する必要がある。従って DVMRP は、大規模ネットワーク上におけるマルチキャスト通信では、適さないという問題がある。

このような問題を改善するために、新たに H-DVMRP やスパースモードマルチキャストプロトコルである、PIM-SM(RFC2362) や CBT(RFC2201,RFC2189) が提案されている。

付録

DVMRP ルーティングテーブル

```
Multicast Routing Table (3 entries)
 Origin-Subnet      From-Gateway    Metric Tmr In-Vif  Out-Vifs
 192.168.2/24                          1     0    2    0* 1*
 192.168.1/24                          1     0    1    0* 2*
 133.5.33/24                           1     0    0    1* 2*
```

DVMRP 転送テーブル

```
vifs_with_neighbors = 0

Virtual Interface Table
Vif  Name  Local-Address                               M  Thr  Rate   Flags
 0   fxp0  133.5.33.102    subnet: 133.5.33/24         1   1      0   querier
                         pkts in : 0
                         pkts out: 0

 1    de0  192.168.1.1     subnet: 192.168.1/24        1   1      0   querier
                         pkts in : 0
                         pkts out: 0

 2    de1  192.168.2.1     subnet: 192.168.2/24        1   1      0   querier
                         pkts in : 0
                         pkts out: 0
```