

PIM-SM と BGMP を用いたスケーラブルなマルチキャスト QoS ルーティングプロトコルの提案

向山繁喜[†] 植野誠史[†] 加藤聰彦[‡] 鈴木健二[‡]
[†]電気通信大学 [‡]KDD 研究所

インターネットの広帯域化に伴い、マルチキャストを利用した映像や音声の配信への期待が高まっており、それらに対する QoS の保証が望まれている。現在マルチキャストトラフィックに対する QoS の保証を行なうプロトコルが提案されているが、QoS 保証を行なう経路制御のためのメッセージ数が多く、スケーラビリティに問題がある。そこで、PIM-SM と BGMP を用いることでネットワークの階層化を行い、メッセージを広告する範囲を限定することで、スケーラブルな QoS マルチキャストルーティングの提案を行った。本稿では、提案したプロトコルの詳細について述べる。

Proposal of Scalable QoS Multicast Routing Protocol using PIM-SM and BGMP

Shigeki Mukaiyama[†], Seiji Ueno[†], Toshihiko Kato[‡] and Kenji Suzuki[‡]
The University of Electro-Communications[†], KDD R&D Laboratories[‡]

As the Internet bandwidth becomes wider, the demands of video and audio data delivery services using multicast technology are increased. For these services, the guarantee of QoS is required. Although some QoS-aware multicast routing protocols have been proposed so far, they have a scalability problem such that the number of messages for QoS routing is too many. So, we have proposed a scalable QoS multicast routing protocol by introducing the hierarchy to a network by using PIM-SM and BGMP, and by limiting the region to advertise route update messages. This paper describes the details of the proposed protocol.

1. はじめに

インターネットの広帯域化にともない、一対多または多対多の通信を可能とする IP マルチキャストに対する期待が高まっている。これにより、ユニキャストに比べ少ないオーバーヘッドで、音声や画像を含むマルチメディア情報の広範囲への配信が可能になる。しかし、現在の

IP マルチキャストはベストエフォートを基本としているため、音声や動画などの通信をスムーズに行うためには、マルチキャストトラフィックに対する QoS の保証が必要となる。

現在、QoS を考慮したマルチキャストルーティングプロトコルが検討されている。[1]では、最近注目されている PIM-SM[2]をベースとして、リンクの利用可能な帯域の状況を考慮した

経路選択を提案している。しかしこのプロトコルでは、マルチキャストトラフィックの発生・終了により、あるリンクの利用可能帯域が変化すると、ネットワーク内のすべてのマルチキャストルータにその情報を広告しなければならず、スケーラビリティに乏しいという問題がある。

また、マルチキャストルーティングにドメインを導入し、ドメイン内は PIM-SM で、ドメイン間は BGMP[3]で経路制御を行うという検討が行われている。しかし、BGMP に対する QoS の検討は行われていない。

そこで、本稿では、PIM-SM と BGMP を使用してドメインによる階層化を行い、広告すべき利用可能帯域の変更情報を削減する、QoS マルチキャストルーティングプロトコルについて述べる。

2. PIM の QoS 拡張^[1]

2.1. PIM-SM への拡張

PIM-SM では受信者は最初に RP (Rendezvous Point)へ Join メッセージを送信することで Shared-Tree に参加する。その後、Shared-Tree からのトラフィックが閾値を超えると、ルータは発信元に Join メッセージを送信することで Shortest Path Tree (SPT)と呼ばれる最短経路の tree へ切り替える。

[1]はこの SPT への切り替え時に、QoS を保証する (利用可能帯域が十分なリンクから構成される)経路を決定し、その経路へ Join メッセージを送信し、経路上の資源を予約する。その経路上のリンクでは利用可能帯域が変更されるため、OSPF (Open Shortest Path First)を用いて、各マルチキャストルータに変更情報を広告する。

2.2. OSPF への拡張

OSPF に QoS 保証を行なうために必要な拡張を行なう。各リンクの予約可能な帯域情報は、OSPF の Link State Update message を拡張したメッセージを用いて広告する。また、

OSPF ルータは、通常のメトリックを持つデータベースの他に、QoS 保証を行なう経路決定に必要な、各リンクにおける利用可能な帯域情報を持つデータベースを保持しなければならない。Join メッセージを受け取った OSPF ルータは、QoS 保証のためのデータベースを基にルーティングを行なう。

2.3. OSPF 拡張の例

図 1 に OSPF ルータにより構成されるネットワークを示す、矢印上の数字は、その方向に予約可能な帯域である (Mbps)。このとき、各ルータが持つデータベースを表 1 に示す。

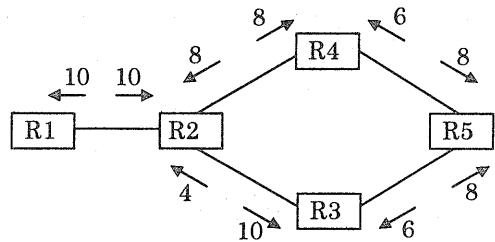


図 1. OSPF ルータ間の利用可能帯域情報

表.1 OSPF ルータの持つ情報

		From				
		R1	R2	R3	R4	R5
To	R1		10			
	R2	10		4	8	
	R3		10			6
	R4		8			6
	R5			8	8	

ルータはネットワーク全体の予約可能な帯域情報を知っており、この情報を基に Join メッセージを送信する経路を決定する。この例で R1 が R5 へ Join メッセージを送信するために選択する経路は、R1/R2/R4/R5 であり、予約可能な最大の帯域は 6Mbps である。

3. BGMP と MBGP

3.1. BGMP の概要

BGMP は、ドメイン間のマルチキャスト経路制御を行うプロトコルであり、PIM-SM と同じく Shared-Tree と SPT の双方をサポートする。BGMP では、PIM-SM の RP に相当するルートドメインが存在する。ルートドメインは、マルチキャストアドレスに応じて静的に割り当てられる。BGMP では PIM-SM と相互運用するために、PIM-SM に対応するメッセージを定義している。また BGMP でのルーティングに必要な情報は、BGP/MBGP (Multi-protocol Extensions for BGP-4) を用いて広告することが可能である。

3.2. BGMP と MBGP への拡張

QoS を保証する経路選択を行なうために MBGP を拡張して利用可能な帯域情報の広告を行なう。また、この時ドメイン内部の帯域情報は考慮しない。

一方、BGP では OSPF と異なり、ネットワーク全体のトポロジを知ることはできないことに注意する必要がある。すなわち、ある BGP

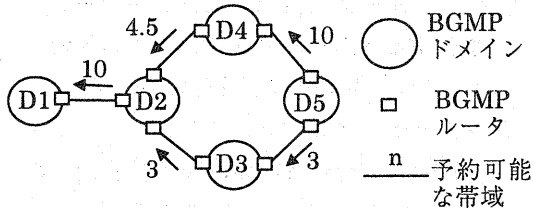


図2 BGMP ルータ間の利用可能帯域情報
表2. BGP/MBGP により広告される経路

D2 が持つ D5 からの経路情報		
宛先	ネクストホップ	予約可能な帯域
D5	D3	4.5
D5	D2	3
D1 が持つ D5 からの経路情報		
宛先	ネクストホップ	予約可能な帯域
D5	D2	4.5

ピアから広告される経路は、そのピアが利用するもののみである。例えば図2において、D5 から D2 までの経路としては D3 経由と D4 経由の二通りがある。しかし、D2 から D1 へ広告される D5 宛ての経路は D4 を経由するもののみである。

このため、利用可能な帯域情報の広告においては、経路上において利用可能な帯域の最小値のみを通知することとする。図2の例で見ると、D5 から D4 を通り D2 まで予約可能帯域 4.5、D5 から D3 を通り D2 まで予約可能帯域 3 の経路が D2 へ広告される。ここで D2 は、D5 からの経路として D4 を通る経路を選択し、D2, D4 を通る D5 への予約可能帯域 4.5 の経路だけを D1 へ広告する。その結果、図2のネットワークにおいて、D2 および D1 が有する D5 からの経路情報は表2に示すとおりとなる。

経路の選択の際には、利用可能な帯域情報の最小値が大きいものを選択する。BGMP ピアへの広告は、自分が利用する最良の経路のみで行われるが、BGMP ルータは 2 番目、3 番目の経路の情報も保持する。これは、Join が失敗した場合の次の候補として利用するためである。図2の例では、D3 を通る D5 への予約可能帯域 3 の経路が 2 番目に選択される経路として保持される。

4. スケーラブルなマルチキャストルーティングプロトコル

4.1. 方針

リンクの帯域変更情報の広告量を削減して、スケーラブルなマルチキャストルーティングプロトコルを検討するにあたり、以下の方針を立てた。

- (1) BGMP と PIM-SM を用いて階層化されたネットワークを対象とする。
- (2) [1]と同様に、SPT に切り替える Join メッセージの転送において、QoS を保証する経路を選択する。さらにそれに伴うリ

- リンクの帯域変更情報の広告を行う。
- (3) ドメイン内のリンクの帯域変更情報はそのドメイン内でのみ広告し、他のドメインには広告しない。ドメイン間ではドメインを接続するリンクの帯域変更情報のみを広告する。これにより帯域変更情報の広告量を制限する。
 - (4) ドメインにまたがって Join メッセージを転送する場合は、1つのドメインでは、そのドメイン内のリンクの情報とドメイン間のリンクの情報という限られた情報を使って、経路を決定し、次のドメインに Join メッセージを転送する。次のドメインでも同様な方法で経路を設定する。
 - (5) 途中で QoS を保証する経路が設定できないことが判明すると、失敗の通知を、QoS 保証が確立された経路を逆に転送し、別の経路を検索させる。これは、BGMP と PIM-SM に Join に対する失敗を通知する Join NACK メッセージを追加することにより実現する。

4.2. QoS を保証する Tree への切り替え

SPT の切り替えは以下の手順で行う。

- (1) 最初は QoS を保証しない Shared-Tree を経由してトラヒックを受信する。
- (2) QoS 保証が要求されると、受信者に接続された PIM-SM ルータが、SPT に切り替えるための Join メッセージを送信する。この PIM-SM ルータは受信者ドメイン D_r において、QoS 保証が実現可能な経路を計算し、BGMP ボーダ ルータ R_r へ向けて Join メッセージを送信する。Join メッセージを受け取った各ルータは、リンクの帯域予約を行い、利用可能帯域の変更をドメイン D_r の中で広告する。
- (3) Join メッセージを受信したボーダ ルータ R_r は、ドメイン間のリンクの利用可能帯

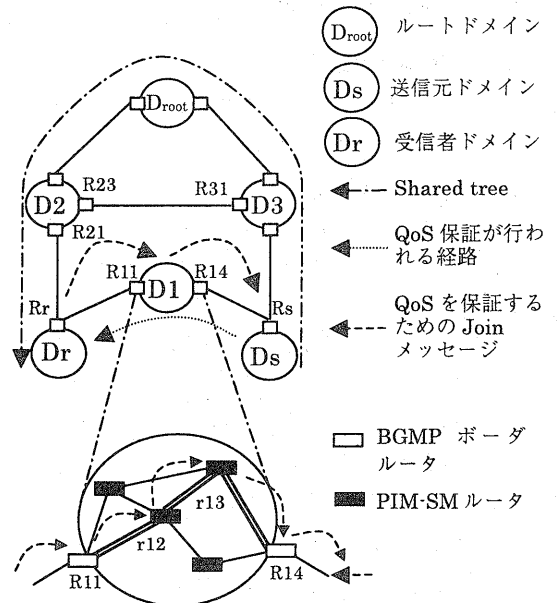


図.3 QoS を考慮した Tree への切り替え

域に基づき、 R_r が保持する QoS を保証することの出来る R_s への最良の経路 (図 3 では $D_r/D1/D_s$ の経路) に従い、Join メッセージをボーダ ルータ R_{11} に送信する。Join メッセージは、通過する AS のリストを保持する。

- (4) ボーダ ルータ R_{11} は、Join メッセージを受信すると、 R_r-R_{11} のリンクの帯域予約を行い、帯域の変更を他の BGMP ボーダ ルータに広告する。さらに、ドメイン D_1 内のリンクの帯域情報から、指定された R_{14} 宛の、QoS を保証できる経路を計算し、Join メッセージを転送する。 D_r と同様に D_1 内において帯域予約と帯域変更の広告が行われる。
- (5) 以下、リンク $R_{14}-R_s$ および送信元ドメイン D_s において同様な手順が繰り返される。なお、Join メッセージが転送される経路上に、すでにそのマルチキャストトラヒックに対して予約されている帯域が存

在する場合は、予約は一つにまとめられる。もし、すでに予約されている帯域が、Joinメッセージによって要求された帯域よりも大きいならば、それ以上のJoinメッセージの転送は、行なわれない。これまでのメッセージの流れの流れを図4に示す。

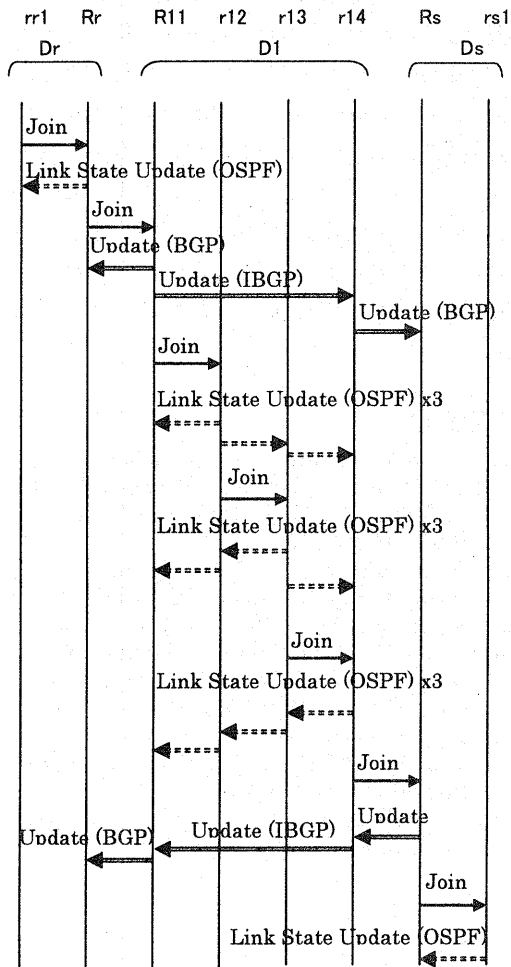


図4. Joinメッセージの流れ

4.3. 経路の再計算

本プロトコルでは、BGMP ボードルータはドメイン間のリンクの情報の広告を行い、それに基づきドメイン間の経路を決定する。このため、Join を転送する経路上のドメイン内部に

おいて QoS の保証が実現できない場合がある。その場合、ボードルータは、Join メッセージを転送した 1 つ前のルータに対して Join NACK メッセージを送信する。Join NACK メッセージを受信したルータは、Join NACK メッセージを受け取った経路の帯域予約を取り消し、Rs への予約可能な経路が他に存在すれば、その経路に Join メッセージを再度送信する。

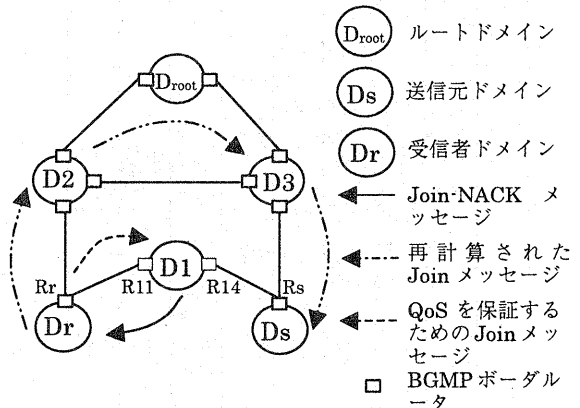


図5 経路の再計算

図5の例ではドメインD1の内部においてQoSが実現できない場合を示しており、R11は、D1内部でのQoS保証が実現できないため、R11, Rr間の帯域予約を取り消し、Join NACKメッセージをRrに送信している。Join NACKメッセージを受け取った、RrはDsへの経路の次の候補であるD2へ向けてJoinメッセージを転送する。D2は、Dsへの最良の経路であるD3へとJoinメッセージを転送する。図5はこのように、Dr/D2/D3/Dsを通る経路にJoinメッセージが再度送信される様子を示し、図6にこの時のメッセージの流れを示す。図6では、D3においてすでに帯域予約が行なわれており、それ以上のJoinメッセージの転送は行なわれない。

このようにJoinメッセージの再送を行なうことにより、帯域予約を行なうことの出来る経路の検索を行なう。

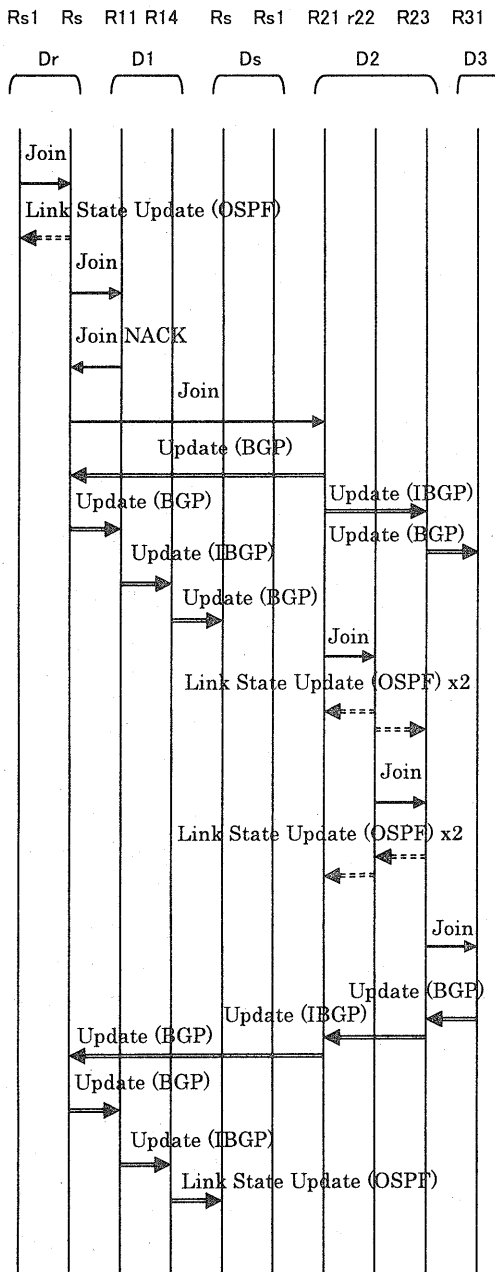


図 6.再送信時のメッセージの流れ
(r22 は、D2内部のルータである)

5. おわりに

本稿では、マルチキャストトラフィックに対する QoS 保証を実現するために、BGMP と PIM-SM を用いて階層化されたネットワークの QoS マルチキャストルーティングプロトコルを提案した。本プロトコルでは、リンクにおいてすでに予約されている帯域に関する情報を考慮していない。これは、ルータで保持しなければならない情報が增大してしまうためである。しかし基本的にサイズが大きくなるものの UPDATE メッセージの数は変わらないと考えられる。このようなリンクにおいてすでに予約されている帯域情報の利用の検討と、本プロトコルを利用することにより減少するメッセージ数の検討などが今後の課題として挙げられる。

参考文献

- [1] S. Biswas and B. Rajagopalan, "A QoS-Aware Routing Framework for PIM-SM Based IP-Multicast," Internet Draft, June 1999.
- [2] D. Estrin, et al., "Protocol Independent Multicast-Sparse Mode (PIM-SM): Protocol Specification," RFC 2362, June 1988.
- [3] D. Thaler, et al., "Border Gateway Multicast Protocol (BGMP): Protocol Specification," Internet Draft, Sep 2000.
- [4] Tony Bates et al., "Multiprotocol Extension for BGP-4": Protocol Specification," Internet Draft, Sep 2000.
- [5] Y. Rekhter et al., "A Border Gateway Protocol 4": Protocol Specification," Mar 1995.