

ATM上のIPマルチキャスト通信のための アーキテクチャとその実装

石川 憲 洋†

インターネット上の仮想的なマルチキャスト通信ネットワークであるMBone上で、マルチメディアツールを用いた様々なマルチメディア通信実験が実施されている。その経験は、IPマルチキャスト通信が、インターネット上でのマルチメディア通信のための基盤技術であることを示している。しかしながら、マルチメディア通信の観点からみた場合、現在のインターネットは、帯域不足、サービス品質(QOS)の未保証などの問題点がある。一方、ATMは、QOSの保証が可能な広帯域ネットワークのための基本技術として、幅広く認められている。したがって、インターネット上でマルチメディア通信を成功させるためには、ATM上でのIPマルチキャスト通信の実現が非常に重要である。本論文では、ATM上のIPマルチキャスト通信のための新しいスケーラブルなアーキテクチャを提案する。我々は、IPマルチキャスト通信に適したATMの機能であるポイント・ツー・マルチポイントコネクションを効率的に利用するために、共有木アーキテクチャを提案する。我々は、ATM上のIPマルチキャスト通信のために、IGMPを拡張して、IGMP-ATMと呼ぶプロトコルを新たに開発した。我々は本アーキテクチャに基づくプロトタイプを開発した。既存のマルチメディアツール(vat, vicなど)が、我々のプロトタイプ上で十分な性能で動作している。

An Architecture for IP Multicast Communications over ATM Networks and Its Implementation

NORIHIRO ISHIKAWA†

The deployment of IP multicast has been realized by MBone, a virtual multicast network on the Internet. Various multimedia tools have been developed and tested on MBone. The experience gained shows that IP multicast is a core technology for multimedia communications over the Internet. However, from the multimedia communications point of view, the Internet has serious drawbacks (i.e. low bandwidth and lack of QOS guarantee). On the other hand, ATM has been widely accepted as the key technology for high speed networking. ATM can provide high bandwidth as well as QOS guarantee. Therefore, the solid integration of IP multicast with ATM is essential for the success of multimedia communications over the Internet. In this paper, we propose a new architecture for scalable IP multicast communications over ATM. We propose a shared tree architecture for IP multicast routing over ATM, in order to efficiently use a point-to-multipoint connection which is the ATM function suitable for IP multicasting. We newly defined a protocol called IGMP-ATM by extending IGMP for IP multicast communications over ATM. We have implemented a prototype system based on our architecture. Existing multimedia tools (e.g. vat and vic) are now running over our prototype with acceptable performance.

1. はじめに

近年、IPマルチキャスト通信のためのプロトコル^{1),2)}が、幅広いプラットフォーム上でサポートされるようになり、インターネット上の仮想的なマルチキャスト通信ネットワークであるMBone³⁾上で、マルチメディアツールを用いた様々なマルチメディア通信実験が実施

されている。MBone上で使用されているマルチメディアツールとしては、ビデオ会議のためのvic⁴⁾、オーディオ会議のためのvatなどが代表的である。その経験は、IPマルチキャスト通信が、インターネット上のマルチメディア通信のための基盤技術であることを示している。

しかしながら、イーサネット、ISDN、専用線などの既存のネットワークを使用して構築された現在のインターネットは、マルチメディア通信の観点から見た場合、帯域不足、サービス品質(QOS)の未保証など

† NTT 情報通信研究所
NTT Information and Communication Systems Laboratories

の問題点がある。

一方、ATMは、構内網(LAN)から広域網(WAN)までシームレスに適用可能で、マルチメディア通信に必要なQOSの保証が可能な広帯域ネットワークング技術として、幅広く認められている。したがって、インターネット上でマルチメディア通信を成功させるためには、ATM上でのIPマルチキャスト通信の実現が非常に重要である。

本論文では、ATM上のスケーラブルなIPマルチキャスト通信のための新しいアーキテクチャ⁵⁾を提案する。本アーキテクチャにより、ATM LAN上での小規模なIPマルチキャスト通信から、ATM広域網を含む大規模なIPマルチキャスト通信までが実現可能となる。我々は、本アーキテクチャに基づくプロトタイプをワークステーション(WS)上で開発した。既存のマルチメディアツール(vic, vatなど)が、我々のプロトタイプ上で十分な性能で動作している。

IPマルチキャスト通信の新しいアプリケーションとして、電子ニュース、ソフトウェアなどのデジタル情報を同時に誤りなく多数のユーザに分配する高信頼マルチキャスト通信が注目を集めている。高信頼マルチキャスト通信のためのプロトコルは、IPマルチキャスト通信を利用してUDP上で実現されているものが大部分である⁶⁾。したがって、これらのプロトコルを本アーキテクチャ上で動作させることにより、ATM上の高信頼マルチキャスト通信が実現可能となる。

本論文の構成は、以下のとおりである。2章では、ATM上のIPマルチキャスト通信に対する要求条件について述べる。3章では、提案するATM上のスケーラブルなIPマルチキャスト通信のアーキテクチャについて述べる。4章では、本アーキテクチャのWS上での実装について述べる。5章では、本アーキテクチャに関する残された課題について述べる。6章では、関連する研究に言及する。

2. 要求条件

ATM上のIPマルチキャスト通信に対する要求条件を以下に述べる。

- スケーラビリティ：インターネットは急速に発展を続けているため、スケーラビリティは非常に重要である。ATM上のIPマルチキャスト通信のためのアーキテクチャは、ATM LAN上の小規模なIPマルチキャスト通信から、ATM LAN、ATM広域網を含む大規模なIPマルチキャスト通信までを効率的にサポートできる必要がある。
- ATMの有効使用：IGMP¹⁾などの既存のIPマ

ルチキャスト通信のためのプロトコルは、イーサネットなどのブロードキャストLAN上で使用することを想定して設計されているため、ATM上のIPマルチキャスト通信のためには新しいアーキテクチャが必要となる。特に、ATMが提供する機能で、マルチキャスト通信に適しているポイント・ツー・マルチポイントコネクションを有効に利用する必要がある。

- ロバストネス：ATM上のIPマルチキャスト通信に参加しているあるノードで障害が発生した場合でも、IPマルチキャスト通信を継続して実行できる必要がある。
- 相互接続：MBoneは、IPマルチキャスト通信のための実験ネットワークとして急速に発展しているため、MBoneと相互接続できる必要がある。そのためには、MBoneで使用されているIPマルチキャストルーティングプロトコルであるDVMRP²⁾と相互接続できることが要求される。
- マルチキャスト通信アプリケーションのサポート：既存のIGMPの実装では、RFC 1112で定義しているIPサービスインタフェースに基づいて、APIが定義されている。既存のマルチメディアツール(vic, vatなど)は、このAPI(たとえば、ソケットインタフェース)を使用して開発されている。これらのツールが、ATM上でも容易に動作可能とするために、本アーキテクチャは、RFC 1112と互換性があるIPサービスインタフェースを提供することが要求される。

3. アーキテクチャ

2章で述べた要求条件を満足するために、本論文では、ATM上のIPマルチキャスト通信のための新しいアーキテクチャを提案する。

3.1節では、ATMマルチキャスト論理IPサブネットワーク(ATM MLIS: ATM Multicast Logical IP Subnetwork)上でのIPマルチキャスト通信のためのアーキテクチャを提案する。3.2節では、3.1節で提案したアーキテクチャを拡張して、ATM上でのIPマルチキャストルーティングのためのスケーラブルなアーキテクチャを提案する。

3.1 ATM MLIS上のIPマルチキャスト通信

3.1.1 モデル

本論文では、ATM MLISを以下のとおり定義する。

- ATM MLIS：1台のATMマルチキャストルータにより管理されるホストの集合
ATM MLISは、RFC 1577⁷⁾で定義しているATM

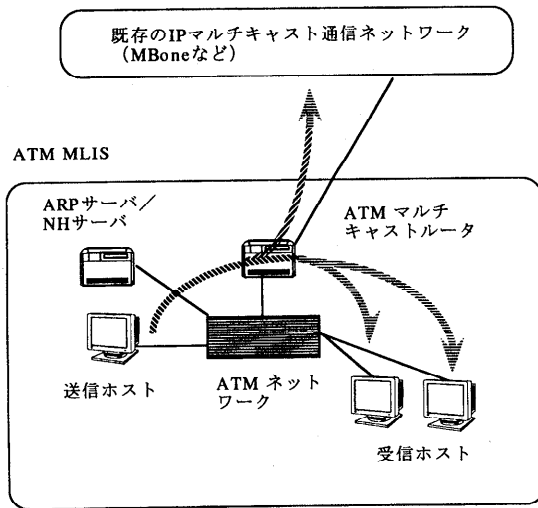


図1 ATM MLIS 上の IP マルチキャスト通信のネットワークモデル
Fig. 1 Network model for IP multicasting over ATM MLIS.

LIS と類似の概念であるが、以下の点で異なる。

- (1) ネットワークの運用管理などの理由により、1つの ATM LIS を複数の ATM MLIS に分割してもよい。
- (2) 異なる ATM LIS に属するホスト間で、ルータを介することなく直接 IP 通信を行うことを可能とする、NHRP⁸⁾ のようなプロトコルを利用すれば、複数の ATM LIS をまとめて1つの ATM MLIS と定義することも可能である。

提案するアーキテクチャのネットワークモデルを図1に示す。

送信ホストは、ATM のポイント・ツー・ポイントコネクションを使用して、IP マルチキャストデータグラムを ATM マルチキャストルータに送信する。ATM マルチキャストルータは、ATM のポイント・ツー・マルチポイントコネクションを使用して、受信した IP マルチキャストデータグラムを複数の受信ホストに分配する。ATM マルチキャストルータは、受信した IP マルチキャストデータグラムを既存の IP マルチキャスト通信ネットワーク (MBone など) にルーティングしてもよい。

本アーキテクチャの受信ホスト上での実装モデルを図2に示す。このモデルは説明のためのものであり、実際の実装を制約するものではない。

ATM の UNI としては、ATM Forum UNI 3.1⁹⁾ で規定しているポイント・ツー・マルチポイント接続を含む SVC を使用することを前提とする。ATM 上で

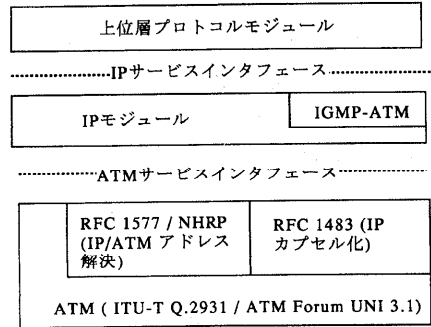


図2 受信ホスト上での実装モデル
Fig. 2 Implementation model on a receiving host.

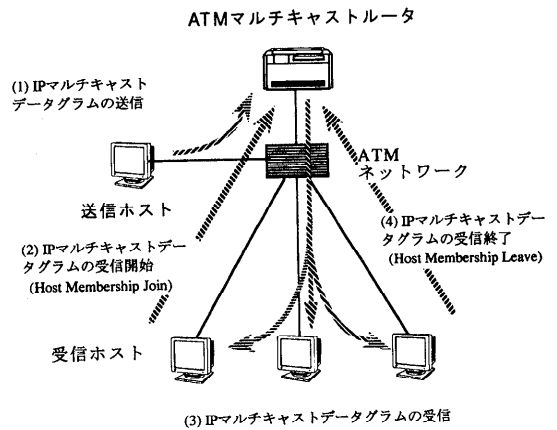


図3 ATM MLIS 上の IP マルチキャスト通信のシナリオ
Fig. 3 Scenario for IP multicasting over ATM MLIS.

の IP カプセル化方式としては、RFC 1483¹⁰⁾ で規定している LLC/SNAP カプセル化方式を使用する。IP アドレスから ATM アドレスへのアドレス解決には、RFC 1577 で規定している ARP サーバまたは NHRP で規定している NH サーバを使用する。IGMP-ATM は、IGMP と同じ機能を ATM 上で実現するために、IGMP を拡張して設計したプロトコルである。

3.1.2 プロトコル手順

提案するアーキテクチャのプロトコル手順を、送信ホスト、受信ホスト、ATM マルチキャストルータの観点から述べる (図3)。

3.1.2.1 送信ホストの動作

IP マルチキャストデータグラムを送信する送信ホストの動作は、以下のとおりである。

- (1) 送信ホストは、ARP サーバまたは NH サーバにアクセスして、ATM マルチキャストルータの IP アドレスから ATM アドレスを導出する。
- (2) 導出した ATM マルチキャストルータの ATM アドレスを使用して、送信ホストと ATM マル

チキャストルータとの間で、データ転送用 VCC として、ATM ポイント・ツー・ポイントコネクションを設定する。

- (3) 設定したデータ転送用 VCC を使用して、IP マルチキャストデータグラムを送信する。

上記の手順から分かるように、IP マルチキャストデータグラムの送信手順は、宛先アドレスがホストグループを識別するクラス D の IP アドレスであることを除いて、IP ユニキャストデータグラムの送信手順と同じである。

3.1.2.2 受信ホストの動作

受信ホストの上位層プロトコルモジュールから見た場合、IP マルチキャストデータグラムの受信手順は、IP ユニキャストデータグラムの受信手順と同じである。しかしながら、IP マルチキャストデータグラムの受信を開始する前に、上位層プロトコルモジュールは、そのホストグループに参加することを、IP モジュールに通知する必要がある。同様に、IP マルチキャストデータグラムの受信を終了するとき、そのホストグループから離脱することを、IP モジュールに通知する必要がある。

IP サービスインタフェースレベルで RFC 1112 と互換性を保持するために、本アーキテクチャでは、上記の目的を実現するために、RFC 1112 で規定している以下の 2 つのオペレーションを変更することなく使用する。

- JoinHostGroup
(group-address, ATM interface)
- LeaveHostGroup
(group address, ATM interface)

JoinHostGroup オペレーションを受信した受信ホストの動作は以下のとおりである。

- (1) 受信ホストは、ARP サーバまたは NH サーバにアクセスして、ATM マルチキャストルータの IP アドレスから ATM アドレスを導出する。
- (2) 導出した ATM マルチキャストルータの ATM アドレスを使用して、受信ホストと ATM マルチキャストルータとの間で、制御用 VCC として、ATM ポイント・ツー・ポイントコネクションを設定する。
- (3) 設定した制御用 VCC を使用して、Host Membership Join メッセージを送信する。

LeaveHostGroup オペレーションを受信した受信ホストの動作は以下のとおりである。

- (4) 制御用 VCC を使用して、Host Membership Leave メッセージを送信する。

3.1.2.3 ATM マルチキャストルータの動作

ATM マルチキャストルータでは、IP マルチキャストデータグラムの宛先アドレス（すなわち、クラス D の IP アドレス）と、その宛先アドレスを持つ IP マルチキャストデータグラムを複数の受信ホストに対して分配するためのデータ転送用 VCC（すなわち、ATM ポイント・ツー・マルチポイントコネクション）を 1 対 1 に対応づけて管理する。

受信ホストから Host Membership Join メッセージを受信した ATM マルチキャストルータの動作は、以下のとおりである。

- (1) 受信ホストが参加することを要求するホストグループ宛ての IP マルチキャストデータグラムを分配するためのデータ転送用 VCC が存在するか否かを確認する。
- (2) 存在する場合は、ATM UNI の ADD PARTY 手順を使用して、受信ホストをデータ転送用 VCC のリーフとして追加する。存在しない場合は、ATM UNI の SETUP 手順を使用して、新たにデータ転送用 VCC を設定する。

送信ホストから IP マルチキャストデータグラムを受信した ATM マルチキャストルータの動作は、以下のとおりである。

- (3) 受信した IP マルチキャストデータグラムを受信ホストに分配するためのデータ転送用 VCC が存在するか否かを確認する。
- (4) 存在する場合は、そのデータ転送用 VCC を使用して、受信した IP マルチキャストデータグラムを分配する。存在しない場合は、受信した IP マルチキャストデータグラムを廃棄する。

受信ホストから Host Membership Leave メッセージを受信した ATM マルチキャストルータの動作は、以下のとおりである。

- (5) ATM UNI の DROP PARTY 手順を使用して、受信ホストが離脱することを要求するホストグループ宛ての IP マルチキャストデータグラムを分配するためのデータ転送用 VCC から、受信ホストへのリーフを削除する。

3.1.3 IGMP-ATM メッセージ

IGMP-ATM で定義しているメッセージを表 1 に示す。

Host Membership Join メッセージは、受信ホストの正当性を確認するための認証パラメータをオプションとして追加していることを除いて、IGMP の Host Membership Report メッセージと同じである。Join Ack メッセージ、Join Nak メッセージ、Host Mem-

表 1 IGMP-ATM のメッセージ
Table 1 IGMP-ATM messages.

| メッセージ | パラメタ |
|-----------------------|--------------|
| Host Membership Join | グループアドレス, 認証 |
| Join Ack | グループアドレス |
| Join Nak | グループアドレス, 理由 |
| Host Membership leave | グループアドレス |

bership Leave メッセージは、本プロトコルで新たに定義した。IGMP の Host Membership Query メッセージは使用しない。

3.2 ATM 上の IP マルチキャストルーティング 3.2.1 モデル

3.1 節で述べた ATM MLIS 上の IP マルチキャスト通信のためのアーキテクチャは、1 台の ATM マルチキャストルータですべての送信ホスト、受信ホストの管理を行うため、スケーラビリティに問題があった。たとえば、1 台の送信ホストから 50 台の受信ホストにビデオを分配する放送型アプリケーションの場合、ATM マルチキャストルータは受信ホストとの間で、50 本の制御用 VCC と 50 個のリーフを持つデータ転送用ポイント・ツー・マルチポイントコネクションを設定する必要がある。このように、3.1 節で述べたアーキテクチャでは、スケーラビリティは 1 台の ATM マルチキャストルータの性能、機能に依存することになり、たとえば、1000 台以上の受信ホストを対象とした大規模な IP マルチキャスト通信を実現することは困難である。

上記の課題を解決するために、3.1 節で提案したアーキテクチャを拡張して、ATM 上の IP マルチキャストルーティングのためのスケーラブルなアーキテクチャを提案する。提案するアーキテクチャは ATM マルチキャストルータ間が木構造で接続される構成とした。木構造を構成する ATM マルチキャストルータ間には上下関係がある。各ルータから見て、上位のルータを親ルータ、下位のルータを子ルータと呼ぶ。最上位の ATM マルチキャストルータは子ルータのみを持つ。最下位の ATM マルチキャストルータは親ルータのみを持つ。

送信ホストは、任意のレベルの ATM マルチキャストルータと接続し、IP マルチキャストデータグラムを送信することができる。受信ホストは、通常最下位の ATM マルチキャストルータと接続し、IP マルチキャストデータグラムを受信する。送信ホスト、受信ホスト、ATM マルチキャストルータ間は ATM VCC により接続される。ATM マルチキャストルータは、アーキテクチャ上、子ルータと受信ホストを区別する必要

はなく、ATM の有効利用のために、ATM ポイント・ツー・マルチポイントコネクションを使用して子ルータおよび受信ホストと接続する。

IP マルチキャストデータグラムの宛先アドレス（すなわち、クラス D の IP アドレス）ごとに、マルチキャストルーティングツリー（MRT: Multicast Routing Tree）を定義する。特定の ATM マルチキャストルータが、各 MRT のルートとして指定される。

本アーキテクチャのスケーラビリティを、上記の 1 台の送信ホストから 1000 台以上の受信ホストにビデオを分配する放送型アプリケーションの例を用いて説明する。ルートの下で子ルータとして 50 台の ATM マルチキャストルータが接続され、さらに各子ルータに 50 台の受信ホストが接続される構成とする。この場合、各 ATM マルチキャストルータが 50 台までの受信ホストに対してビデオを分配する性能を持つならば、結果として 2500 台までの受信ホストに対してビデオを分配することが可能となる。特に、ATM マルチキャストルータから受信ホストへの分配は ATM ポイント・ツー・マルチポイントコネクションを使用するため、受信ホストの台数が増加しても ATM マルチキャストルータの負荷には影響を及ぼさない。

このように、木構造のアーキテクチャを採用することにより、ATM 上のスケーラブルな IP マルチキャスト通信が実現できる。

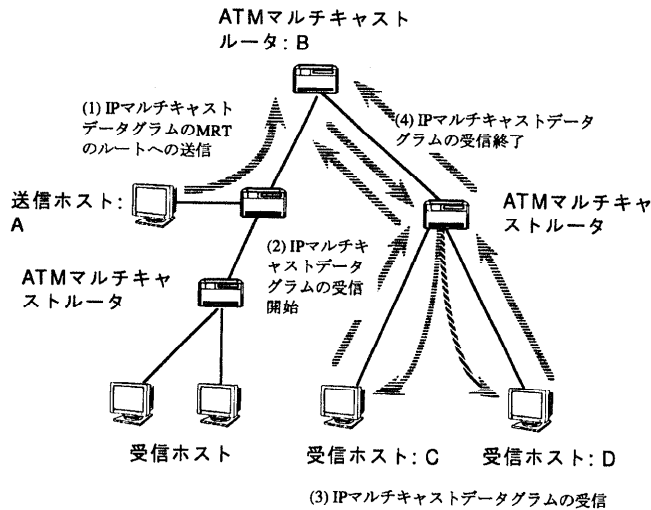
3.2.2 プロトコル手順

提案するアーキテクチャのプロトコル手順を、送信ホスト、受信ホスト、ATM マルチキャストルータの観点から述べる（図 4）。最後に、IP マルチキャストルーティングの性能を改善するためのショートカット手順について述べる。

3.2.2.1 送信ホストの動作

送信ホストは、IP マルチキャストデータグラムを、その宛先アドレスで識別されるホストグループに対して定義された MRT のルートに送信する。

送信ホストがルートとの間で、直接データ転送用 VCC を設定できる場合は、3.1.2.1 項で述べた手順が適用される。送信ホストがルートとの間で、直接データ転送用 VCC を設定できない場合は、IP マルチキャストデータグラムを、宛先アドレスにルートの IP アドレスを設定した IP ユニキャストデータグラムとして送信する。この場合、IP マルチキャストデータグラムの宛先アドレスは、ルーズソースルートオプションを使用して、IP ユニキャストデータグラム内に設定される。すなわち、IP マルチキャストデータグラムは、ルートまで IP ユニキャストデータグラムとして



備考 このシナリオでは、送信ホストAが、ホストグループに対して定義されたMRTのルートであるATMマルチキャストルータBを介して、受信ホストB、Cに対して、IPマルチキャストデータグラムを送信している。

図4 ATM 上の IP マルチキャストルーティングのシナリオ

Fig. 4 Scenario for IP multicast routing over ATM.

ルーティングされ、ルートで IP マルチキャストデータグラムに展開される。

3.2.2.2 受信ホストの動作

受信ホストの動作は、3.1.2.2 項で述べた ATM MLIS 上の IP マルチキャスト通信における受信ホストの動作と同じである。

3.2.2.3 ATM マルチキャストルータの動作

ATM マルチキャストルータが MRT のルートである場合の動作は、3.1.2.3 項で述べた ATM MLIS 上の IP マルチキャスト通信における ATM マルチキャストルータの動作と同じである。以下では、ATM マルチキャストルータが MRT のルートでない場合の動作について述べる。

受信ホストまたは子ルータから Host Membership Join メッセージを受信した ATM マルチキャストルータの動作は以下のとおりである。

- (1) ATM マルチキャストルータがすでに親ルータに対して Host Membership Join メッセージを送信して、その応答として Join Ack メッセージを受信済みの場合は、3.1.2.3 項で述べた手順が適用される。
- (2) ATM マルチキャストルータがすでに親ルータに対して Host Membership Join メッセージを送信して、その応答を未受信の場合は、親ルータからの応答を待つ。その後、親ルータから Join Ack メッセージを受信した場合は、受信ホストまたは子ルータとの間にデータ転送用 VCC を

設定する。

- (3) ATM マルチキャストルータが親ルータに対して Host Membership Join メッセージを未送信の場合は、親ルータに対して Host Membership Join メッセージを送信する。その後、親ルータから Join Ack メッセージを受信した場合は、受信ホストまたは子ルータとの間にデータ転送用 VCC を設定する。

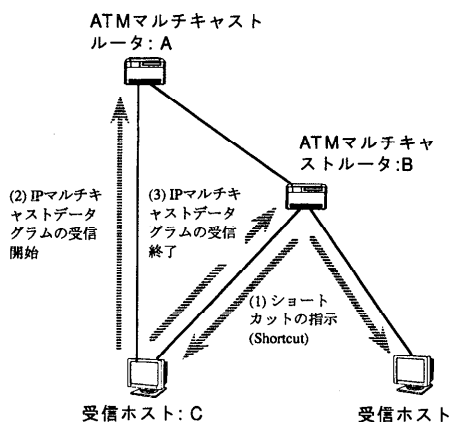
このように、Host Membership Join メッセージは、受信ホストから MRT のルートまで、順次伝搬される。

送信ホストまたは親ルータから IP マルチキャストデータグラムを受信した ATM マルチキャストルータの動作は、3.1.2.3 項で述べた ATM MLIS 上の IP マルチキャスト通信における ATM マルチキャストルータの動作と同じである。

受信ホストまたは子ルータから Host Membership Leave メッセージを受信した ATM マルチキャストルータの動作は以下のとおりである。

- (4) ATM UNI の DROP PARTY 手順を使用して、データ転送用 VCC から、受信ホストまたは子ルータへのリーフを削除する。
- (5) ATM マルチキャストルータが MRT のルートではなく、受信ホストまたは子ルータへのリーフを削除した結果、データ転送用 VCC が解放された場合は、親ルータに対して Host Membership Leave メッセージを送信する。

このように、Host Membership Leave メッセージ



備考 このシナリオでは、受信ホストCが、ATMマルチキャストルータBからのショートカットの指示に基づいて、ATMマルチキャストルータAに対して、ショートカットのためのATM VCCの設定を要求している。

図5 ショートカット手順

Fig. 5 Shortcut procedures.

は、受信ホストからMRTのルートまで、順次伝搬される。

3.2.2.4 ショートカット手順

本アーキテクチャは、ATMマルチキャストルータを木構造で接続することによりスケーラビリティを確保しているが、中継するATMマルチキャストルータの数(ホップ数)が増えると遅延時間が増大し性能が低下する。特に、ビデオ会議などの実時間性を要求するアプリケーションでは遅延時間などに対し強い要求条件がある。

遅延時間などに関して本アーキテクチャの性能を改善するために、ATMマルチキャストルータをバイパスするショートカット手順を提案する(図5)。

ATMマルチキャストルータは、下記のような場合に、受信ホストおよび子ルータに対してショートカットの指示を行う。

- 送信ホストからあるホストグループに対して、大量のデータが長時間にわたり送信されている。

MRTのルートではない場合、ATMマルチキャストルータは、ショートカットアドレスとして親ルータのIPアドレスを設定したShortcutメッセージを送信することにより、受信ホストおよび子ルータに対してショートカットを指示することができる。

Shortcutメッセージを受信した受信ホストおよび子ルータは下記の手順によりショートカットを実行する。

- (1) Shortcutメッセージのショートカットアドレスで指定されたルータに対して、Host Membership Joinメッセージを送信する。
- (2) 応答としてJoin Ackメッセージを受信した場合

は、親ルータに対してHost Membership Leaveメッセージを送信する。

- (3) 応答としてJoin Nakメッセージを受信した場合は、なにもしない。

提案するショートカット手順は各ATMマルチキャストルータが独立して実行することが可能であり、実装も容易である。しかしながら、すべてのATMマルチキャストルータがショートカット手順を実行すると、結果として1台のATMマルチキャストルータがすべての受信ホストに対してIPマルチキャストデータを分配することになり、スケーラビリティが失われる。

どのような判断基準に基づいてショートカット手順を適用するかについてはATMマルチキャストルータの運用の問題であるが、下記のようにATMマルチキャストルータがショートカットを指示する基準を明確にすることにより、資源の有効利用を考慮した効率的なショートカットを実現できると考える。

- ショートカット手順を適用するホストグループを限定する。
- ショートカットを指示する基準(一定時間内に受信したデータ量など)を定める。

また、ショートカット手順を実行した結果、資源が不足すると判断した場合は、ATMマルチキャストルータは受信ホストまたは下位のルータからのショートカット要求を拒否してもよい。

このように、我々が提案する方式では、各ATMマルチキャストルータが、資源の有効利用を考慮してショートカット手順を実行することを可能としている。

3.2.3 IGMP-ATMメッセージ

3.2.2.4項で述べたショートカット手順のためにShortcutメッセージが追加されたことを除き、3.1.3項で述べたIGMP-ATMのメッセージがそのまま適用される。

4. 実装

4.1 プロトタイプの開発

我々は、本アーキテクチャに基づくプロトタイプを、下記の環境で、OSのカーネルを改造して開発した。

WS: SunSPARCstation 20

(SuperSPARC+: 60 MHz)

OS: SunOS 4.1.4

ATM Switch: FORE ASX-200

ATM SBus Adapter: FORE SBA-200 (155 Mbps)

ATMマルチキャストルータ機能は、WS上のソフトウェアルータとして開発した。OSのカーネルを拡

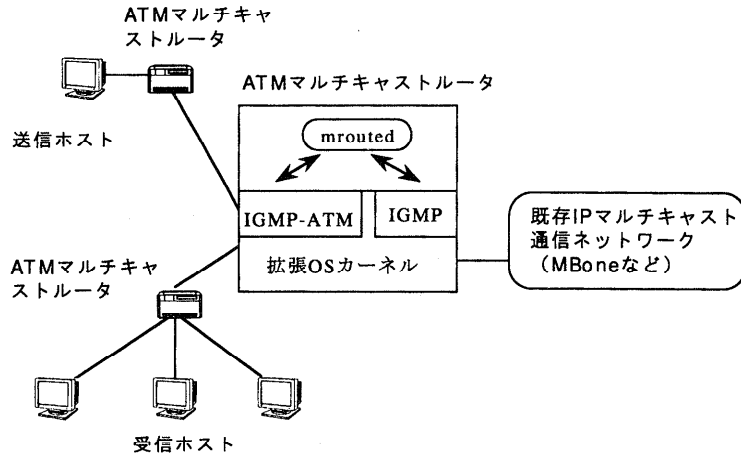


図6 既存IPマルチキャスト通信ネットワークとの相互接続
Fig. 6 Interworking with existing IP multicast Networks.

表2 プロトタイプの性能評価
Table 2 Performance measurement of the prototype.

| パケット長 (バイト) | 256 | 1024 | 4096 | 8192 |
|-----------------------|-----|------|------|-------|
| UDP/IP マルチキャスト (Mbps) | 8 | 34.9 | 88.9 | 108.7 |
| UDP/IP ユニキャスト (Mbps) | 9.4 | 39.4 | 92.2 | 116.6 |

張して、送信ホスト、受信ホスト、ATM マルチキャストルータの機能を同時に実装しているので、1台のWSを同時に、送信ホスト、受信ホスト、ATM マルチキャストルータとして動作させることも可能である。

2章で述べた「マルチキャスト通信アプリケーションのサポート」に関する要求条件を満足するために、SunOSのIPマルチキャスト通信に関するソケットインタフェースの仕様には変更を加えていない。結果として、vic, vatなどの既存のマルチキャスト通信アプリケーションが、ソースコードを修正することなく、我々のプロトタイプ上で動作している。

提案するアーキテクチャに基づいて開発したプロトタイプの性能評価を下記の形態で実施した。1台の送信ホストから1台のATMマルチキャストルータを介して受信ホストが接続される形態で、1台の受信ホスト上でスループットの測定を行った。スループットの測定は、UDP/IPマルチキャストとUDP/IPユニキャストの場合について行った(表2)。

表2に示すように、UDP/IPマルチキャストのスループットは、UDP/IPユニキャストのスループットの約90%であり、本アーキテクチャに基づくマルチキャストルーティングのオーバーヘッドがわずかであることを確認した。また、データのコピーはATMSwitch内で行われるため、受信ホストの数を5台程度まで増

やしても、受信ホストのスループットには影響を及ぼさなかった。

複数台のATMマルチキャストルータが木構造で接続された場合でも、個々のATMマルチキャストルータの処理は同じなので、受信ホストのスループットには大きな影響を及ぼさないと考えられる。ただし、遅延時間などの評価が必要である。より大規模なネットワーク環境における評価は今後の課題である。

4.2 既存IPマルチキャスト通信ネットワークとの相互接続

2章で述べた「相互接続」に関する要求条件を満足するために、我々のプロトタイプでは、Mboneなどの既存IPマルチキャスト通信ネットワークとの相互接続を実現している。その実現方式を以下に述べる。

Mboneでは、DVMRPを実装しているmroutedと呼ぶWS上のソフトウェアルータによりIPマルチキャストルーティングを実現している。我々のプロトタイプ上で、mroutedをデーモンとして起動し、それと拡張したOSカーネルを連動させることにより、既存IPマルチキャスト通信ネットワークとの相互接続を実現した(図6)。mroutedは、各サブネット上に受信ホストが存在するか否かを調べるために、各ネットワークインタフェースに対して、定期的にIGMPのHost Membership Queryメッセージを送信している。一方、3.1.3項で述べたように、我々のアーキテクチャではHost Membership Queryメッセージを使用していない。我々のプロトタイプでは、mroutedからHost Membership Queryメッセージの送信要求を受けると、OSカーネルが、自分自身で管理テーブル

を検索して、各 ATM サブネット上に受信ホストが存在するか否かを調べ、その結果を mrouted に応答している。mrouted には修正を加えていない。このように、我々のプロトタイプでは、既存 IP マルチキャスト通信ネットワークとの相互接続を比較的単純に実現することができた。

5. 残された課題

我々のアーキテクチャは、2章で述べた要求条件の中で、「ロバストネス」に関する要求条件を満足していない。MRT を構成するある ATM マルチキャストルータで障害が発生した場合、その ATM マルチキャストルータより下位の ATM マルチキャストルータ、受信ホストは以降正しく動作できない。

この課題を解決する最も単純な方法は、各 ATM マルチキャストルータに対して、バックアップ用の ATM マルチキャストルータを用意することである。送信ホスト、受信ホストまたは子ルータは、ある ATM マルチキャストルータの障害を検出した場合、バックアップ用の ATM マルチキャストルータとの間で ATM コネクションを設定し、IP マルチキャスト通信を再開する。バックアップ用の ATM マルチキャストルータは、必要に応じて親ルータとの間で、ATM コネクションを設定する。

また、本アーキテクチャは ATM と IP を独立したレイヤとして扱うことを前提として設計した。近年、ATM と IP のルーティングを統合し、より効率的なルーティングを実現するレイヤ 3 スイッチングの研究開発が盛んである。本アーキテクチャのレイヤ 3 スイッチングへの適用も今後の課題である。

6. 関連研究

ATM 上の IP マルチキャスト通信についてはこれまでに様々な方式が提案されている。本章ではこれまでに提案された主要な方式をサーベイし、特に重要と考える方式については我々が提案する方式との詳細な比較評価を行う。

IETF では、ATM 上の IP マルチキャスト通信を実現するために、マルチキャストアドレス解決サーバ¹¹⁾ (MARS: Multicast Address Resolution Server) と呼ぶ方式を開発している。MARS は、IP マルチキャスト通信に適用するために、RFC 1577 で規定している ARP サーバを拡張している。MARS は我々が提案する方式と同様の狙いを持ち、ほぼ同じ環境で実現されているため、6.1 節で詳細な比較評価を行う。

ATM Forum では、ATM 上のコンピュータ通信

のために、LAN エミュレーション¹²⁾ (LANE: LAN Emulation over ATM) と呼ぶ方式を開発している。LANE は、Ethernet/IEEE 802.3, IEEE 802.5 などの既存 LAN を ATM 上でエミュレーションする方式で、IP のみならず既存 LAN 上で動作するレイヤ 3 プロトコルならば、その仕様を変更することなく ATM 上で利用できる利点を持つ。ATM 上でエミュレートされた LAN 上で、RFC 1112 で規定している IGMP を動作させることにより、LANE 上で IP マルチキャスト通信を実現できる。しかしながら、本方式は既存 LAN を ATM 上でエミュレーションするためオーバーヘッドが大きい。加えて、IGMP などのレイヤ 3 プロトコルから直接 ATM にアクセスすることができないため、ATM の提供するマルチキャスト通信に適した機能であるポイント・ツー・マルチポイントコネクションを有効に利用することができない。

ATM Forum では、さらに MPOA¹³⁾ (Multi-Protocol Over ATM) と呼ぶ方式を開発している。MPOA は、LANE と IETF で開発された NHRP を組み合わせることにより、LANE では実現できなかったサブネットを越えたショートカットを可能としている。しかしながら、MPOA がショートカットの対象としているのはユニキャスト通信の場合のみであり、マルチキャスト通信は対象としていない。

6.1 MARS との比較評価

MARS では、ATM 上の IP マルチキャスト通信の実現方式として、各送信ホストからすべての受信ホストに対して直接 ATM ポイント・ツー・マルチポイントコネクションを設定する VC メッシュ方式と、マルチキャストサーバが各送信ホストから受信した IP マルチキャストデータグラムを ATM ポイント・ツー・マルチポイントコネクションを利用してすべての受信ホストに中継するマルチキャストサーバ方式の 2 方式を規定している。

MARS のアーキテクチャは、マルチキャストアドレス解決機能、マルチキャストサーバ機能、マルチキャストルータ機能を独立した構成要素としてモデル化しているため、きわめて複雑である。一方、我々のアーキテクチャは、上記の各機能を統合して ATM マルチキャストルータとしてモデル化しているため、MARS と比較して遥かにシンプルである。

本論文では、VCC 消費量の観点から、我々が提案する方式と MARS との定量的な比較評価を行った (図 7)¹⁴⁾。n 台のホストが参加するビデオ会議で、各ホストが送受信を行う場合を想定して評価を行った。データ転送用 VCC はビデオ用およびオーディオ用の

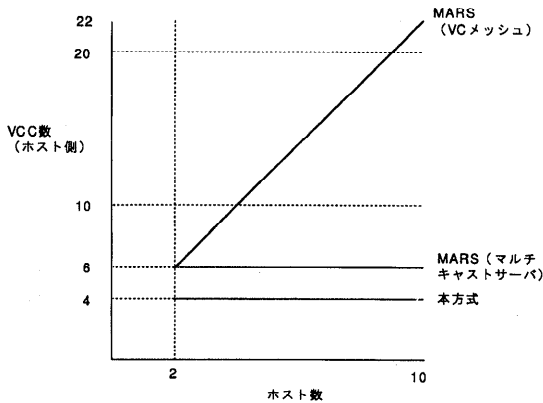


図 7 MARS との比較

Fig. 7 Comparison with MARS.

2本とした。評価は利用可能な資源の制限が厳しいホスト側で行った。我々が提案する方式では、ホスト側で設定されるVCC数はビデオ会議に参加するホスト数とは無関係につねに4本（データ送信用VCC:1本，データ受信用VCC:2本，制御用VCC:1本）である。

MARSのVCメッシュ方式では、設定されるVCC数は、 $2n + 2$ 本（データ送信用VCC:2本，データ受信用VCC: $2n - 2$ 本，制御用VCC:2本）である。ただし、VCメッシュ方式の場合は、ホスト間で直接データ転送用VCCを設定するため、データ転送がATMマルチキャストルータを経由して2ホップで実現される我々が提案する方式と比較すると性能面では優れている。

MARSのマルチキャストサーバ方式では、マルチキャストサーバを1台と仮定すると、設定されるVCC数は、つねに6本（データ送信用VCC:2本，データ受信用VCC:2本，制御用VCC:2本）である。加えて、マルチキャストサーバとマルチキャストアドレス解決サーバの間で、2本の制御用VCCが設定される。マルチキャストサーバ方式の場合は、データ転送がマルチキャストサーバを経由して2ホップで実現されるため、性能面では我々が提案する方式と同等である。

図7に示すように、我々が提案する方式は、ATMの代表的な資源であるVCCの消費量の観点から評価した場合、MARSよりも優れている。MARSのVCメッシュ方式は、性能面では我々が提案する方式よりも優れているが、ホスト数の増加に比例してVCC消費量が増加するためスケーラビリティに欠け、大規模なATM上のIPマルチキャスト通信に適していない。

また、MARSはATM LIS上のIPマルチキャスト通信方式であり、本論文で提案したようなATM上

で大規模なIPマルチキャスト通信を実現するためのスケーラブルなIPマルチキャストルーティング方式は、MARSの対象外としている。加えて、MARSの枠組みでは、効率的にショートカット手順を実現することができないことをMARSの開発者自身が認めている¹⁵⁾。

7. まとめ

本論文では、ATM上のスケーラブルなIPマルチキャスト通信のための新しいアーキテクチャを提案した。我々のアーキテクチャとそのプロトタイプは、下記の特徴がある。

- (1) ATM LAN上の小規模なIPマルチキャスト通信からATM広域網を含む大規模なIPマルチキャスト通信までを効率的にサポート可能である。
- (2) ATMポイント・ツー・マルチポイントコネクションを有効に利用している。
- (3) ショートカット手順を効率的に実現している。
- (4) Mboneなどの既存IPマルチキャスト通信ネットワークとの相互接続をサポートしている。
- (5) vic, vatなどの既存のマルチキャスト通信アプリケーションが、ソースコードを修正することなく動作可能である。

今後の研究課題としては、障害に対するロバスタネスの確保、RSVPなどを使用したATM上のQoS制御（帯域保証など）との連動、本アーキテクチャのレイヤ3スイッチングへの適用などがある。

参考文献

- 1) Deering, S.: Host Extensions for IP Multicasting, RFC 1112 (1989).
- 2) Waitzman, D., Partridge, C. and Deering, S.: Distance Vector Multicast Routing Protocol, RFC 1075 (1988).
- 3) Kumar, V.: *MBone: Interactive Multimedia on the Internet*, New Riders Publishing (1996).
- 4) McCanne, S. and Jacobson, V.: vic: A Flexible Framework for Packet Video, *ACM Multimedia'95* (1995).
- 5) Ishikawa, N.: An Architecture for Scalable IP Multicast Routing over ATM Networks, *3rd International Workshop on Protocols for Multimedia Systems* (1996).
- 6) 城下, 高橋, 山下, 山内, 串田: 高信頼マルチキャスト通信(RMTP)の各種ネットワークへの適用性, 信学技報, SSE95-196/IN95-140 (1996).
- 7) Laubach, M.: Classical IP and ARP over ATM, RFC 1577 (1994).

- 8) Luciani, J., Katz, D., Piscitello, D., Cole, B. and Doraswamy, N.: NBMA Next Hop Resolution Protocol (NHRP), RFC 2332 (1998).
- 9) ATM Forum: ATM User Network Interface (UNI) Specification Version 3.1 (1995).
- 10) Heinanen, J.: Multiprotocol Encapsulation over ATM AAL5, RFC 1483 (1993).
- 11) Armitage, G.J.: Support for Multicast over UNI 3.0/3.1 based ATM Networks, RFC 2022 (1996).
- 12) ATM Forum: LAN Emulation over ATM, Version 1.0 (1995).
- 13) ATM Forum: Multi-Protocol over ATM, Version 1.0 (1997).
- 14) Armitage, G.J.: Issues affecting MARS Cluster Size, RFC 2121 (1997).
- 15) Armitage, G.J.: VENUS-Very Extensive Non-Unicast Service, RFC 2191 (1997).
(平成 9 年 9 月 2 日受付)
(平成 10 年 6 月 5 日採録)



石川 憲洋 (正会員)

1978年京都大学工学部情報工学科卒業。1980年同大学大学院工学研究科情報工学専攻修士課程修了。同年日本電信電話公社(現NTT)入社。現在、NTT情報通信研究所に所属。

OSI, ATM, インターネットプロトコル, マルチメディア通信, マルチメディア情報端末などの研究開発に従事。