衛星回線を用いたマルチキャストデータ配信における パケット集約機構の研究

慶應義塾大学 環境情報学部
片岡 広太郎
kotaro@sfc.wide.ad.jp

指導教員
村井 純
徳田 英幸
楠本 博之
中村 修
南 政樹

平成14年2月18日

概 要

本研究では、衛星回線を片方向通信路として含むネットワークにおいて、多数の受信者に対して信頼性のあるマルチキャストデータ配信を行う際に、各受信者から重複して送出されるIGMPやNAKパケットによって大量のトラフィックが発生し、衛星回線の帯域資源を無駄に消費する問題を解決した。

広域同報型の通信路である衛星回線は、1対多型の通信方式であるマルチキャストとの親和性が高い。衛星回線をマルチキャストの通信路として用いれば、広域ネットワーク上に散在する多数の受信者に対して効率的にデータを配信できる。

衛星回線に対して受信の接続性しか持たないホストであっても、送信可能なホストが受信のみ可能なホストに代行して衛星回線にパケットを転送する技術として、UDLRが研究されている。安価に衛星回線を利用できる環境が整い、多くのホストが衛星回線を含むネットワークに接続できるようになった。

衛星回線を含むネットワークにおいて、信頼性のあるマルチキャストデータ配信を行う際には、マルチキャストグループ管理やデータの欠損を知らせる確認応答を行う。しかし、既存の遅延タイマを用いたパケット送出抑制機構が衛星回線の遅延によって完全には機能しないため、パケットが重複して送出される。ネットワークに接続するホスト数が多大である場合、重複して送出されるパケットがUDLRによって衛星回線に転送され、衛星回線の帯域消費が無視できない程に増大する問題が生じる。

本研究では、ホストから送出されるIGMPとNAKパケットは、各々内容が重複する点に着目した。これらのパケットを集約すれば、衛星回線に転送されるパケットの量を低減し、帯域資源の無駄な消費を防げる。

これを実現するための機構として、衛星回線を片方向通信路として含むネットワークにおいて有効に機能するパケット集約機構を提案し、設計と実装を行った。本機構は、重複した内容のパケットが多数送出された際に、これらを捕捉する。そして、定められたポリシに従って、衛星回線に転送するかを選択する。

本機構を多数の受信者を対象とする信頼性のあるマルチキャストデータ配信に適用すれば、IGMPやNAKにもとづく信頼性保証を動作させる際に、遅延タイマを用いたパケット送出抑制機構が正しく機能するのに必要なパケットのみを衛星回線に転送し、それ以降順次到達する重複パケットを破棄できる。これにより、UDLR環境において、衛星回線の無駄な帯域消費を最小限に抑えてデータを配信できる。

本研究では、内容の重複したパケットが一定方向に多数送出される環境において本機構を動作させ、衛星回線に転送されるパケット数を測定し、受信者数の規模性への対応について評価した。この結果、本機構が、多数重複して送出されるパケットによって衛星回線の帯域消費が増大する問題を解決していることを示した。

本研究により、衛星回線の特性を活かして規模性のあるデータ配信ができる。多数の受信者が接続する衛星回線をインターネットの通信路の一部として用いた、大規模なマルチキャストデータ配信が実現できる。

Abstract

This paper describes a packet aggregation architecture for the network that includes uni-directional satellite links. Multicast data delivery system for a large number of receivers requires enough scalability in managing multicast groups and keeping reliability for the delivery data.

Satellite Links are useful for multicast communication. But satellite links are used as uni-directional links because of cost performance for its operation and equipments. UDLR (Unidirectional Link Routing protocol) enables UDL network to work as virtual Bi-directional Link network that many protocols suppose as working environment. Many receive-only hosts can connect to satellite network by using UDLR.

However, to run the multicast data delivery system in a large scale UDL Network, many duplicate packets for multicast group management and reliable data transfer go into satellite link because of long delay that is a characteristic of satellite link, and of Broadcast Emulation in UDLR. The traffic generated by those packets consumes satellite link's band resources heavily.

Aggregation of control packets for multicast data delivery is efficient for reduce waste of band resources. Packet Aggregator focuses on IGMP (Internet Group Management Protocol) packets and NAK (Negative Acknowledge) packets have duplicate contents, when they are transmitted by many receivers respectively. Packet Aggregator selects and forward the pakets that are just necessary for group management and data delivery.

Packet Aggregator makes satellite links work effectively to deliver data by multicast. This architecture realizes multicast data delivery with scalability and reliability.

目 次

図 目 次

表 目 次

第1章　はじめに

1.1　研究の背景

インターネットの利用者数は爆発的に増加し、社会生活の様々な面で利用されるようになった。インターネットへの接続性を提供する通信路は、ケーブルテレビ、xDSL 等の各種通信技術の進歩と普及により、利用者の接続する通信路の選択肢が広がった。

現在では、新聞や雑誌、テレビ等と同様の情報伝達媒体として、インターネットが利用されている。World Wide Web 等を用いることで、既存の情報伝達媒体だけでは提供できなかった、より細かなサービスを利用者に提供するようになった。

新聞を例に考える。新聞は1日の配送回数が限られているため、新聞という情報伝達媒体のみでは、最新の情報を継続して利用者に伝達できない。しかし、新聞社が設けているWEB サイトでは、常に新しい情報が公開されている。利用者は定期的に受け取る新聞からだけでなく、新聞社の WEB サイトを参照することで、幅広い情報を随時取得できる。既存の情報伝達媒体である新聞とインターネットとを組み合わせることで、利用者に多様な情報を提供している。

インターネットで提供されるサービスの多様化にともない、サービスが提供するコンテンツには、映像や音声を組み合わせたマルチメディアデータが広く利用されるようになった。インターネットの利用者数が増加しているのと同様に、ネットワーク上で伝達されるデータも飛躍的に増大している。

これにともない、複数の利用者に同一のデータを効率的に配信する要求が高まっている。しかし、それぞれの利用者がインターネットに接続する通信路の品質は一様でなく、利用者はネットワーク上に散在している。このため、多数の利用者に同一データを配信する際に、異なる品質の通信路を用いた広域ネットワークにおいて効率的にデータの配信を行える技術が期待されている。

1.2　現在のインターネットにおける問題点

現在、多数の利用者に対して同一のデータを効率的に配信する要求が高まっているが、これを実現する有効な手法が実現されていない。

インターネットにおいて、多数の利用者に対して同一のデータを効率的に配信するための技術として、マルチキャストが広く研究されている。マルチキャストを用いて配送されるデータは、多くのマルチキャストルータを経由する。しかし、既存のネットワークでは、マルチキャストの通信に対応していないルータが存在する。このため、マルチキャストによるデータ配信をネットワーク上の全てのホストに対して行うのは困難である。また、受信者に向かう通信路の帯域に大きな差異がある場合、通信のボトルネックが発生する。こ

のとき、マルチキャストデータはフロー制御を行わない UDP を用いて配送されるのが一般的であるため、多量の配送データが狭帯域な通信路を通る際に、他の通信に悪影響を及ぼす可能性がある。これを図 1.1 に示す。

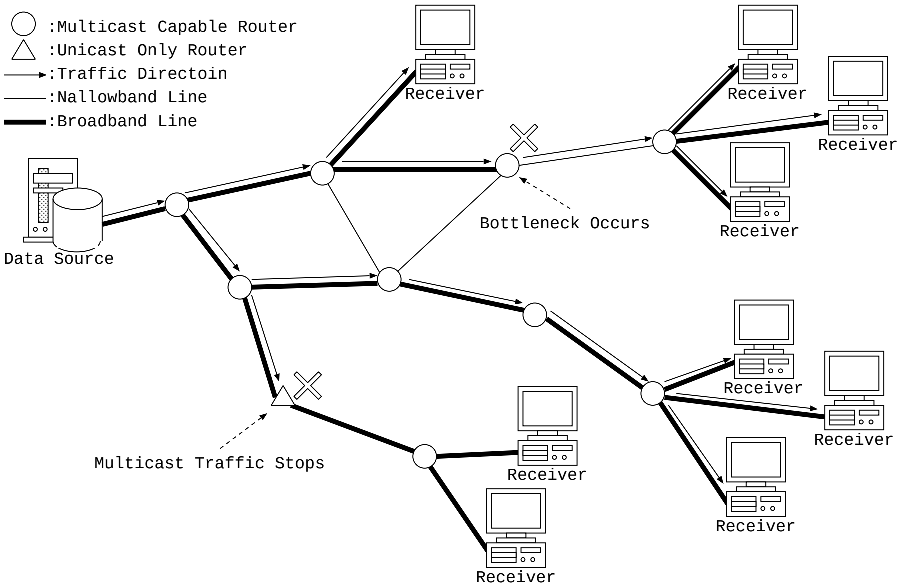

図 1.1: 既存のネットワークにおけるマルチキャストを用いる際の問題点

多数の受信者に対してマルチキャストを用いたデータ配信を行う際には、マルチキャストグループ管理やデータ配送などの技術が利用者数やネットワークの規模に対応する必要がある。しかし、ネットワークの規模が大きくなるにつれ、トラフィックが経由するルータの数が増加し、全体的なスループットが低下する可能性がある。

現在、衛星回線を片方向通信路として含む広域ネットワークを構築し、マルチキャストを用いたデータ配信が試みられている。衛星回線の持つ特徴とマルチキャスト通信は親和性が高く、１対多型の通信を行うには衛星回線を通信路として用いるのが有効である。しかし、UDLR と衛星回線を組み合わせたネットワークにおいて、多数の受信者を対象としてマルチキャストの通信を行う際には、受信者数の規模性に対応できない問題が生じる。

サービスが実用に耐える性能を持つためには、配信されるデータが受信者に到達する時点で定められた品質を保つ必要がある。サービスの形態によって要求される品質は異なり、配信するデータの即時性と完全性とに大別できる。これらの品質は二律排反の関係にあり、同時に満たすのは困難である。即時性を重視するデータ配信はいくつかの手法が提案され実現されている。しかし、多数の利用者が接続する不均質なネットワークにおいて、完全性を保証して効率的にデータ配信を行う手法はまだ実現されていない。

1.3 本研究の目的

本研究では、衛星回線を含むネットワークにおいて、マルチキャストを用いて多数の受信者を対象とするデータ配信を行う際に生じる問題を解決する。この問題の解決手法とし

て、衛星回線を片方向通信路として含むネットワークでの使用に特化したパケット集約機構を構築する。

第一に、既存のインターネットで動作するマルチキャストグループ管理機構を衛星回線を含む大規模なネットワークにおいても有効に動作させる。このとき、既存のIGMPにおけるパケットの送出抑制機構が衛星回線を含むネットワークの規模性に対応できず、完全には機能しない問題を解決する。

第二に、データの完全性を保証するマルチキャストデータ配信において、利用者数の規模に対応し、衛星回線の帯域資源を有効に利用する機構を構築する。多数の利用者が同時に同一パケットの欠損を検知した際に、重複するデータの再送要求や修復データ等が衛星回線の帯域を圧迫する問題を解決する。

1.4 本論文の構成

第2章では、衛星回線とマルチキャストを用いたデータ配送との親和性、衛星回線を片方向通信路として用いるための技術、データの信頼性を考慮した配信手法について述べる。第3章では、衛星回線を含む広域ネットワークにおいて、受信者数の規模性に起因する問題点について述べる。第4章では、第3章で述べる問題点を解決するパケット集約機構について述べる。第5章では、パケット集約機構の設計について、第6章では、実装についてそれぞれ述べる。第7章では、パケット集約機構の評価について述べる。第8章では、本研究のまとめと今後の課題について述べる。

第2章　衛星回線を用いたマルチキャストデータ配信についての考察

本章では、衛星回線を用いたマルチキャストデータ配信について論じる。

2.1　ユニキャストとマルチキャスト

インターネットにおける通信は、1対1型のユニキャスト、1対多型のマルチキャスト、およびブロードキャストに大別できる。

複数の受信者に同一データを配信する場合、ユニキャストによる通信では、個々の受信者に繰り返しデータを送信する。このため、データの送信者に直近のネットワークではトラフィックが集中し、データの送信を繰り返す分の帯域資源を消費する。これを図2.1に示す。

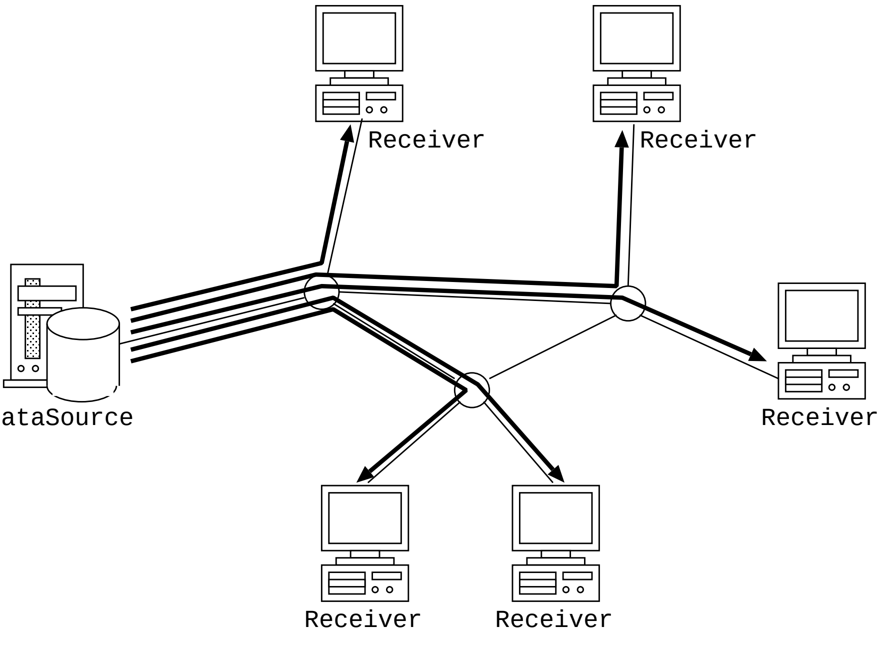

図2.1: ユニキャストによるデータ配信

マルチキャストによるデータ配信では、一度の送信で複数の受信者にデータを配送する。送信されたデータはマルチキャストルータを経由するごとに複製され、枝分れしながら末端に接続する受信者に到達する。このため、少ない帯域消費で多くの受信者にデータを配

送できる。これを図 2.2 に示す。

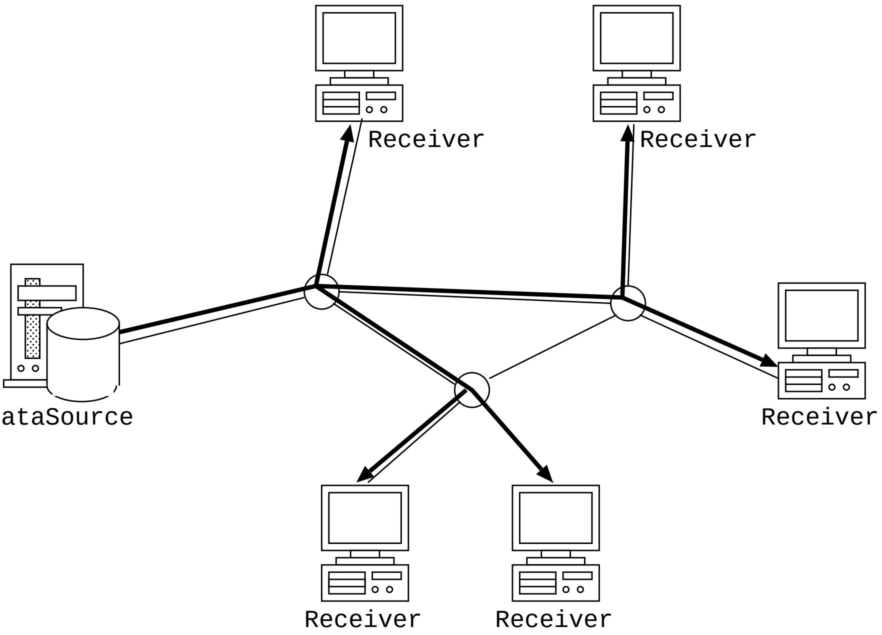

図 2.2: マルチキャストによるデータ配信

2.2　広域ネットワークにおけるマルチキャストの実現

マルチキャストは同一データを複数の利用者に配送する際に有効である。しかし、地上回線による広域ネットワークにおいて、マルチキャストを用いて多数の受信者に信頼性を考慮したデータ配信を行うには、中継ルータにかかる負荷の増大や、通信路の品質の差異をはじめとする制約が多く、現実的でない。

本研究では、衛星回線が持つ広域性、同報性などの特性がマルチキャストとの親和性が高いことに着目する。マルチキャストを行う通信路に衛星回線を用いることで、送信者の直近にあるマルチキャストルータとネットワーク的に離れた場所にいる受信者を直接接続できる。衛星回線はすべての受信者に概ね均質な通信路を提供し、単一の回線に複数の受信者が接続するため、効率的にデータを共有できる。これにより、中継ルータにかかる負荷の増大や、回線品質の差異によって生じる通信のボトルネックなど、地上回線による広域ネットワークで発生する問題の解決が容易になる。これを図 2.3 に示す。

2.2.1　インターネットにおけるマルチキャストグループ管理

マルチキャストでは、通信に参加するホスト群によって構成されるマルチキャストグループに対してパケットの配送を行う。ホストは、マルチキャストグループへ自由に参加し離脱できる。マルチキャストルータは、自分が接続するリンク上のホストがどのマルチキャストグループに参加しているかを管理する必要がある。

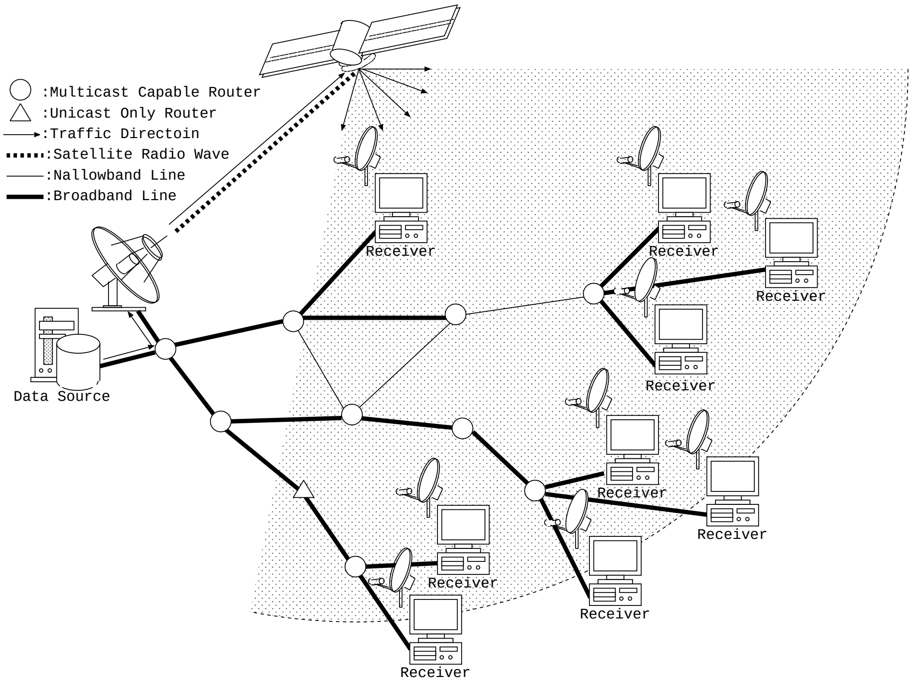

図 2.3: 衛星回線を用いた広域ネットワークにおけるマルチキャストの実現

本研究では、衛星回線を含むネットワークをインターネットの一部として扱う。このため、既存のマルチキャストグループ管理機構が、衛星回線を含むネットワーク上で正しく機能する必要がある。これを図2.4に示す。IGMP(Internet Group Management Protocol)[1]を用いて衛星回線におけるマルチキャストグループ管理を行う。

2.2.2 IGMPの概要

マルチキャストによって配送されるデータは、マルチキャストグループのIPアドレス宛に送信される。ホストがマルチキャストデータを受信するためには、目的のデータが配信されるマルチキャストグループに参加する必要がある。本節では、マルチキャストグループの管理プロトコルである、IGMPの概要を述べる。

マルチキャストグループへの参加と離脱

ホストは、マルチキャストグループに自由に参加し離脱できる。特定のマルチキャストグループに参加する際には、当該グループのマルチキャストグループアドレス宛にMembership Reportを送信しグループに参加する。マルチキャストグループから離脱する際には、Leave Messageを送信しグループから離脱する。

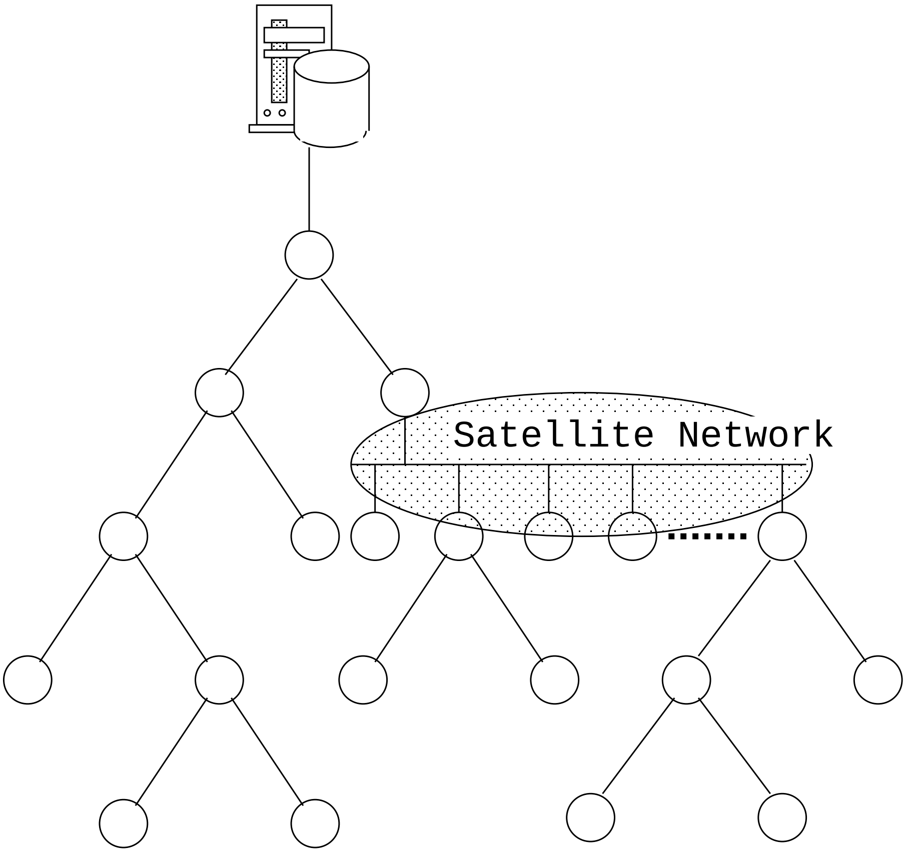

図 2.4: 衛星回線を含むネットワークの既存のマルチキャストツリーへの応用

マルチキャストルータによるグループの状態管理

マルチキャストルータは、IGMPを用いて自分が接続するリンク上のマルチキャストグループの情報を管理する。マルチキャストルータは、接続するリンクごとにマルチキャストグループのメンバ情報のリストを保持する。メンバ情報のリストは、グループごとに全メンバを記録するものではなく、参加するメンバがあるマルチキャストグループを記録するものである。

グループ参加メンバの有無の確認

マルチキャストルータは、参加するメンバがないマルチキャストグループへのトラフィックは転送しない。マルチキャストルータは、管理するマルチキャストグループについて、参加メンバの有無を確認するため、定期的にMembership Queryを送信する。Membership Queryを受信したホストは、Membership Reportをマルチキャストグループアドレス宛に送信する。マルチキャストルータがMembership Reportを受信すると、当該マルチキャストグループに参加メンバがあると判断する。

ホストによる**Membership Report**の送出遅延

Membership Queryを受信したホストは、グループごとに遅延タイマを設定し、Membership Reportを複数のホストが重複して送信するのを抑止する。タイマの値には、Membership Queryで指定されている最大応答時間内でランダムな値が設定される。設定されたタイマがタイムアウトした場合、Membership Reportをマルチキャストグループアドレス宛に送信する。

ホストで設定された遅延タイマが有効な間に他のホストからマルチキャストされたMembership Reportを受信した場合、遅延タイマを解除し、送出を待機していたMembership Reportを破棄する。遅延タイマを用いたMembership Reportの送出抑制は、多数のホストが同一セグメント上に接続している場合、重複したMembership Reportによるトラフィックの量を低減するのに有効である。

2.3 衛星回線を用いた広域ネットワークの構築

2.3.1 衛星回線の片方向通信路としての利用

衛星回線を用いる際、機器のコストや運用コストの面から、送信設備よりも受信設備の方が格段に利用しやすい。このため、衛星回線を少数の送受信可能な局に多数の受信のみ可能な局が接続する片方向通信路として用いるのが一般的である。

インターネットでは双方向の通信が可能な通信路を用いることを前提としている。このため、経路制御をはじめとする多くの既存のシステムが片方向通信路を含むネットワークにおいて完全には機能しない問題が発生する。

片方向通信路における経路制御機構(Unidirectional Link Routing: UDLR)[2]では、衛星回線と地上回線とを統合的に用いることによってこの問題を解決する。これにより、衛

星回線を含むネットワークに接続する各利用者間で双方向のブロードキャスト型の通信が可能なネットワークを構築し、既存のプロトコルがUDLを含むネットワークでも正常に機能する。

2.3.2 UDLRの概要

UDLRは、データリンクパケットのトンネリングとBroadcast Emulationにより、片方向通信路を仮想的な双方向のブロードキャスト型通信路として実現する。以下に、UDLRの仕組みについて述べる。

UDLRにおけるノード形態の定義

衛星回線に接続するノードは、回線への接続形態によって以下のに分類される。

- Receiver:衛星回線に対して受信のみの接続性を持つルータ、またはホスト
- Send-Only Feed:衛星回線に対して送信のみの接続性を持つルータ

本研究では、各ノードはUDLに接続するインターフェースと、地上回線をはじめとする双方向通信路(BDL: Bi-directional Link)に接続するインターフェースの両方を持つことを前提とする。また、本節で述べるUDLRについては、UDLネットワークに接続するノードを上述のSend-Only Feed(以降、Feedと記述)、およびReceiverとする。

ReceiverからFeedに向けたデータ伝送のプロセス

ReceiverはUDLに対して送信可能な接続性を持たないため、UDL上ではReceiverからFeedへ直接の通信ができない。一方、FeedはUDLに対して受信可能な接続性を持たないため、UDL上に流れるデータを受信できない。このため、ReceiverからFeedに向かうデータは、以下のプロセスを経て伝送される。

1. ReceiverがUDL側インターフェースから送信すべきデータリンクパケットを、BDL側インターフェースのIPアドレスがついたIPヘッダによりカプセル化
2. ReceiverのBDL側インターフェースからFeedのBDL側インターフェースにカプセル化したIPパケットを送信
3. FeedがBDL側インターフェースで受信したIPパケットを脱カプセル化し、元のデータリンクパケットをUDL側インターフェースで受信

IPデータグラムの中にデータリンクパケットをカプセル化する技術をトンネリングと呼ぶ。トンネリングを用いることで、UDL上のノード同士で双方向の通信が可能となる。これを図2.5と図2.6に示す。

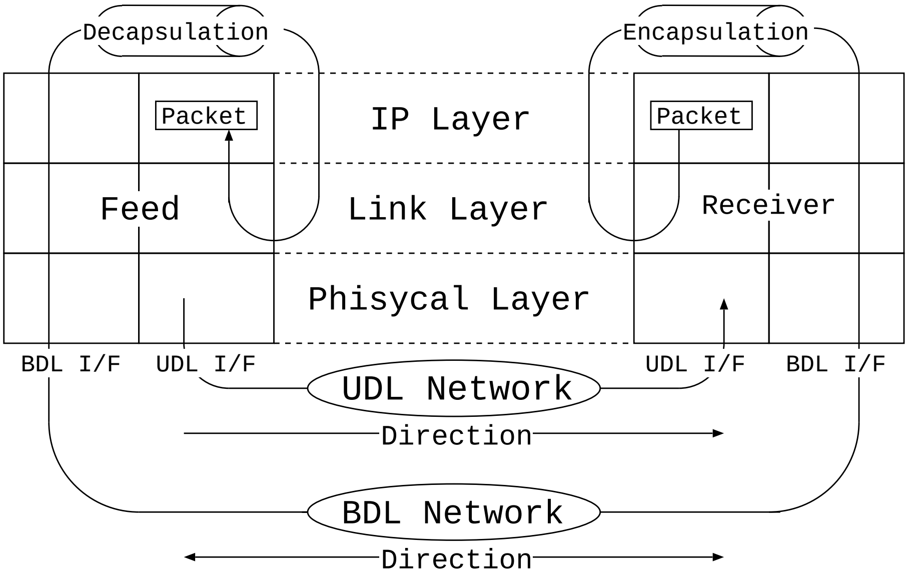

図 2.5: Feed と Receiver におけるパケットカプセル化の概念図

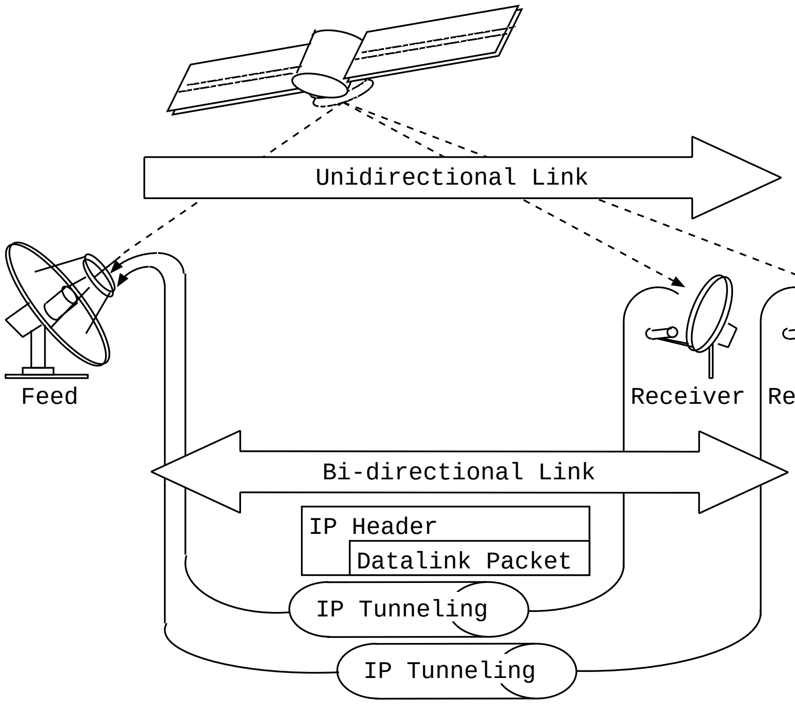

図 2.6: トンネリングを用いた通信の概念図

衛星回線を用いたブロードキャスト型の双方向通信路の仮想的な実現

あるReceiverからFeedや他のReceiverに向けて伝送された全てのユニキャストパケット、マルチキャストパケット、およびブロードキャストパケットは同様のプロセスを経た後、Feedが代行してUDLに転送する。衛星回線は同報型の通信路であるため、Feedが受信したパケットを衛星回線に転送すれば、すべてのReceiverが転送されたパケットを受信することになる。これを図2.7に示す。

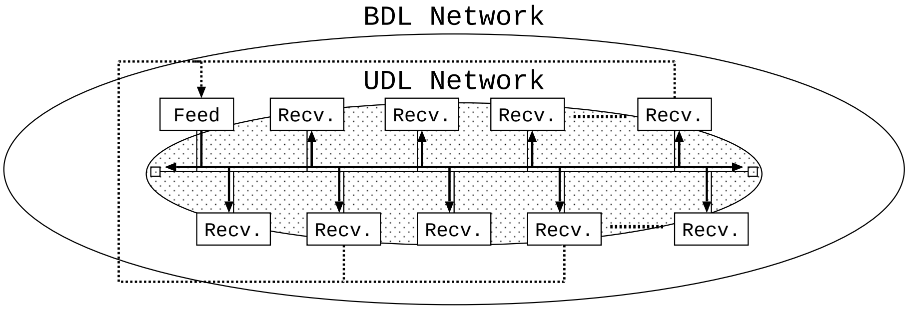

図2.7: Broadcast Emulationの概念図

2.4 信頼性保証型マルチキャストデータ配信

2.4.1 データ配信における信頼性の定義

データ配信における信頼性とは、受信者がデータを受け取る際に保たれているべき品質を指す。データの信頼性は即時性と完全性とに大別できる。前者は、送信者から受信者までなるべく短い時間でデータを配送することを指す。後者は、受信者に届くデータが送信者にあるデータの完全なコピーであることを指す。

ビデオ会議を例に考える。受信するデータに多少の欠損があることは、会話にさほど大きな影響を与えない。データに欠損がある場合、画像が乱れたり、音声が途切れたりするが、前後の会話から内容を推測できる。しかし、ネットワークの遅延が双方の間で異なる場合、会話を行うことが極めて困難である。従って、このようなサービスにはデータの即時性を重視した配送を行う必要がある。

また、ファイルのダウンロードを例に考える。ファイルが個々のデータに分割されて配信される場合、受信するデータの順序の如何は問題ではない。ファイルを構成する全てのデータが正しく受信されていれば、受信者側でそれらのデータをファイルとして再構成できる。しかし、受信するデータに欠損がある場合、当該ファイルがアプリケーションであれば実行できず、楽曲であれば音声が途切れてしまう。ダウンロードしたファイルが送信者のファイルと完全に同一でなければならない。従って、このようなサービスにはデータの完全性を重視した配送を行う必要がある。

2.4.2 信頼性保証の手法

信頼性を保証する手法の一つは、配送するデータに予め前方誤り訂正符号 (FEC: Forward Error Correction) を付加する手法である。この手法では、受信者が FEC にもとづいてデータの欠損を検出し、誤りを訂正する。また、データの再送を伴わないので、送信者にかかる負荷が少ない。しかし、データの欠損が FEC で修正できる範囲を超えている場合や、データそのものが受信者に到達しない場合がある。FEC を用いた手法では、配送するデータの完全な信頼性を保証できるわけではない。

信頼性を保証する二つめの手法は、TCP と同様に肯定的確認応答 (ACK) を用いてデータの再送を行う手法である。この手法では、受信者がデータを正しく受信する度に送信者に ACK を行い、送信者が ACK にもとづいてデータの配送状況を管理し、データの再送を行う。このため、受信者の数が多くなるにつれ、ACK が送信者に集中するという問題 (ACK Implosion) が発生するので、ACK の集約を併せて行うのが一般的である。これを図 2.8 に示す。しかし、正しくデータを受信できた受信者も他の受信者がデータの再送を受けるのを待つ必要があるため、全体的なスループットが低下する。また、受信者数が多大になった際に、送信者が配送状況を管理するための負荷が大きくなる。

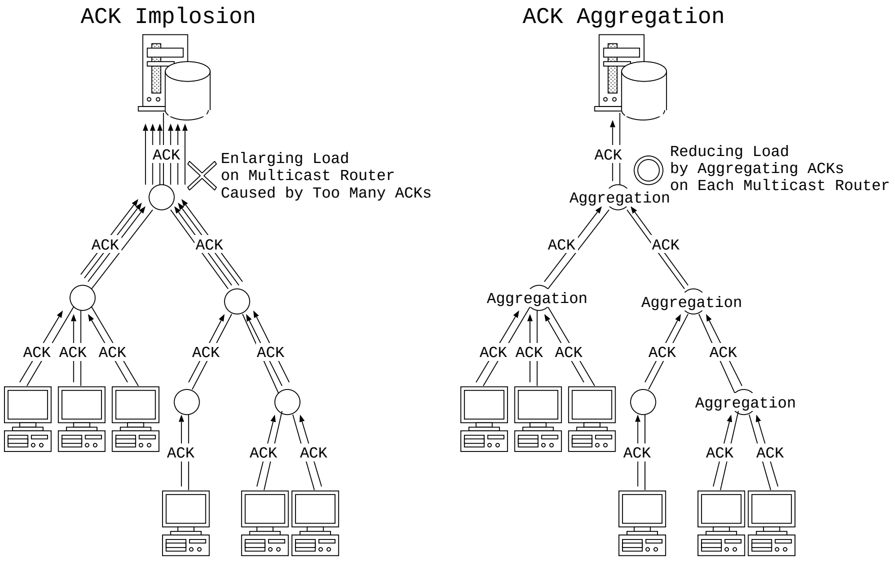

図 2.8: ACK Implosion の概念図

信頼性を保証する三つめの手法は、否定的確認応答 (NAK) を用いてデータの再送を行う手法である。この手法では、配信されたデータに欠損があった場合、受信者は送信者に NAK を送信し、送信者は受信者から送られてきた NAK にもとづいてデータを再送する。このため、データ配送時の誤りが少ないほど、NAK が行われる頻度が少なくなる。また、データの配送状況を受信者が管理するため、送信者にかかる負荷が少ない。これを図 2.9 に示す。

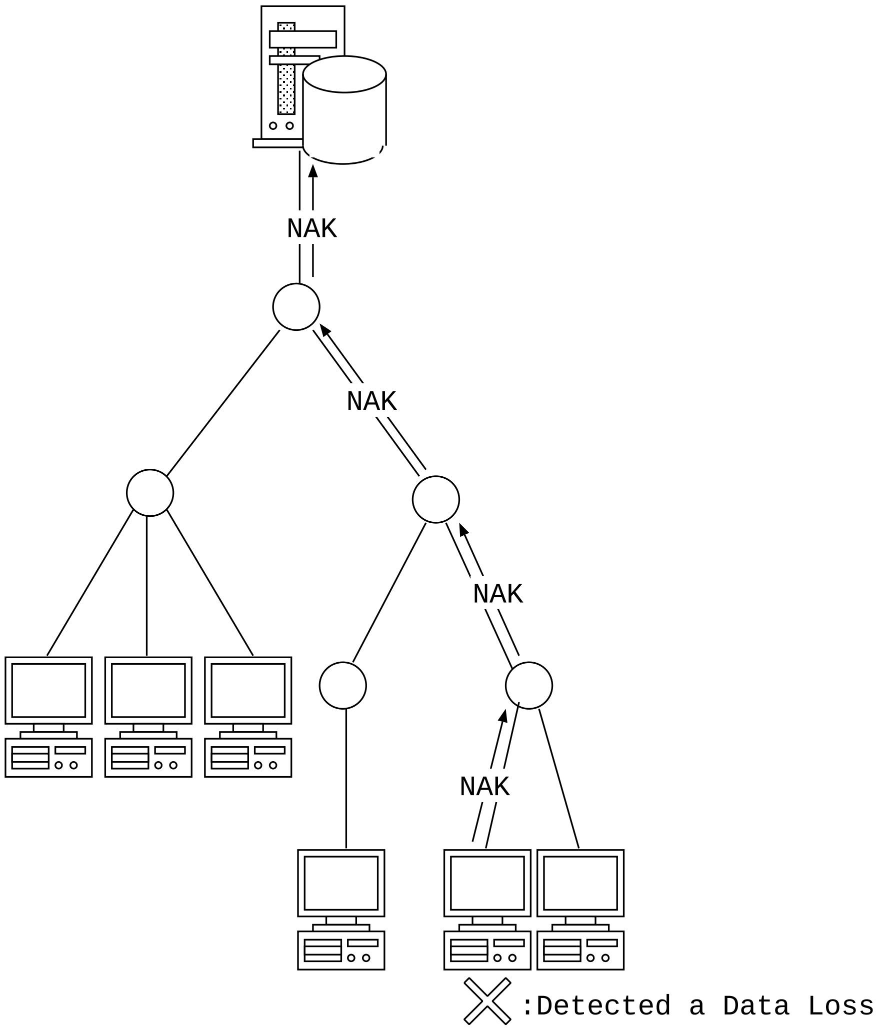

図 2.9: データに欠損がある場合のみ行われる NAK

2.4.3 衛星回線を含むネットワークにおいて有効な信頼性保証の手法

想定するネットワークの構成とその性質

衛星回線を含むネットワークにおいてデータ配信を行う際には、単一のデータ送信者に対して多数の受信者が、仮想的に実現される同一セグメント上のノードとして接続する。これを図 2.10 に示す。

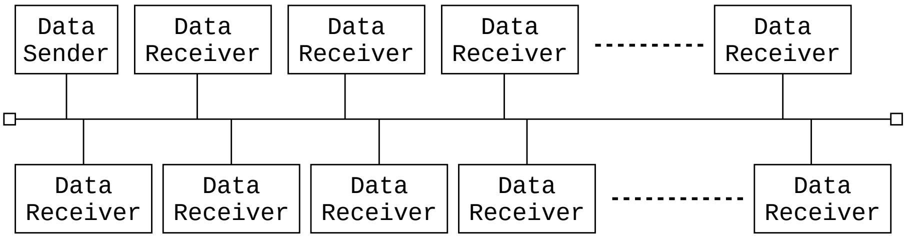

図 2.10: 仮想的に実現された同一セグメントの概念図

衛星回線は同報性を持つ通信路である。各受信者が接続する際の回線品質が概ね一定であるという衛星回線の特性により、データの送信者が欠損無くパケットを衛星回線に転送すれば、多くの受信者がこれを正しく受信すると想定できる。一方で、データが衛星回線に転送される時点で欠損が生じていた場合には、多くの受信者に欠損のあるデータを受信すると想定できる。

ACK と NAK の規模性による比較

データの信頼性を保証する手法のうち、FEC を用いる手法では、完全にはデータを修復できない場合がある。このため、データの再送を伴う ACK または NAK を用いた手法によって、データの完全性を保証する必要がある。

ACK を用いる手法では、受信者からの ACK によって、送信者が各受信者へのデータの配送状況を管理する。このため、データの送信者に対して多数の受信者が接続するネットワークでは、送信者にかかる負荷が増大し、規模性を欠く。

NAK を用いる手法では、ネットワークの規模性に対応するために、パケットの送出遅延タイマを用いたパケット送出の抑制機構を併用できる。各受信者がデータに欠損を検出してから NAK を送出するまでの間、遅延タイマを設定し NAK の送出を待機する。最初にタイムアウトした受信者は NAK をマルチキャストグループアドレス宛に送信する。同じマルチキャストグループに参加している受信者が、遅延タイマが有効な間に同一のデータに対する NAK を受信した場合は、送出を待機していた NAK を破棄する。これにより、重複した NAK パケットによるトラフィックを抑制できる。

衛星回線でデータの送信者と多数の受信者が直接結ばれるネットワーク構成では、ACK を用いるよりも、NAK を用いて信頼性を保証する方が規模性に優れている。

2.5 本章のまとめ

マルチキャストは1対多型の通信形態であり、衛星回線はその性質上、マルチキャストとの親和性が高い。UDLRを用いれば、受信のみ可能なホストが多数接続する広域ネットワークが構築できる。多数の受信者を対象とするデータ配信は、衛星回線とマルチキャストを組み合わせて行うのが有効である。

衛星回線を含むネットワークにおいて、多数の受信者を対象として信頼性を考慮したデータ配信を行う際には、データ受信の確認応答を行う手法を用いるのが現実的である。このとき、全体的なスループットを高い状態で維持するには、NAKを用いた信頼性保証の手法が有効である。

第3章　問題点

本章では、UDLR を動作させた環境において、衛星回線に特有の遅延と受信者数の規模性に起因して発生する、衛星回線の帯域資源が無駄に消費される問題について述べる。

3.1　UDLR が内包する問題点

衛星回線を片方向通信路として含むネットワークでは、送信ホスト、および各受信ホストは、UDLR を用いて仮想的に実現する双方向の通信が可能な同一セグメント上に接続する。UDLR では、Receiver から UDL に向けて送出されたパケットは、すべてトンネリングを用いて Feed に伝送される。Feed がトンネルインターフェースから受信したパケットは、Broadcast Emulation によって UDL に転送される。このため、受信者数が多大になった場合にも、すべての Receiver から送出されたパケットが UDL に転送されるため、衛星回線の帯域消費が増大する。UDLR では、Broadcase Emulation を行っているので、帯域消費の問題は避けられない。これを図 3.1 に示す。

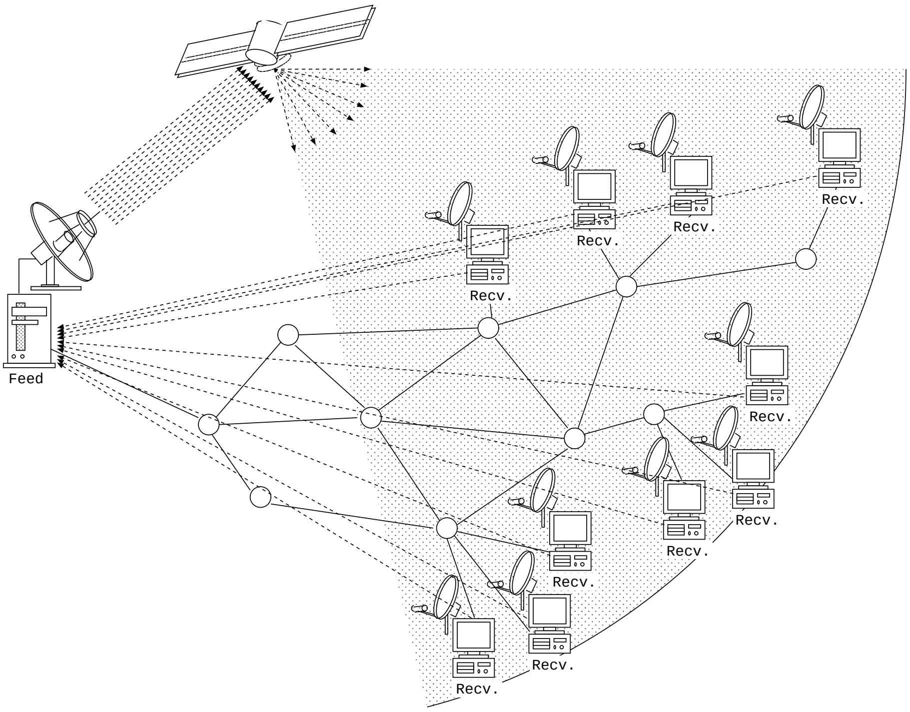

図 3.1: UDLR を用いることによる衛星回線の帯域消費

3.2 マルチキャストグループ管理における問題点

衛星回線を含むネットワークにおいてIGMPを用いた場合、前述のIGMPの送出遅延によるパケット抑制機構が完全には動作しない。衛星回線は高遅延な通信路である。マルチキャストルータからのMembership Queryを受信したホスト群のうち、最初にタイムアウトしたホストから送信されたMembership ReportはFeedを経由し、衛星回線を用いてネットワーク上の全ホストに到達する。よってこの遅延が大きい。

衛星回線には受信のみ可能なホストが多数接続すると想定されるので、既存のパケット抑制機構によって設定される遅延タイマの多くが、他のホストが送信したMembership Reportが到達する前にタイムアウトする。このため、重複したMembership Reportが多数送信され、UDLRのBroadcast Emulationにより、そのまま衛星回線に転送される。これを図5.1に示す。

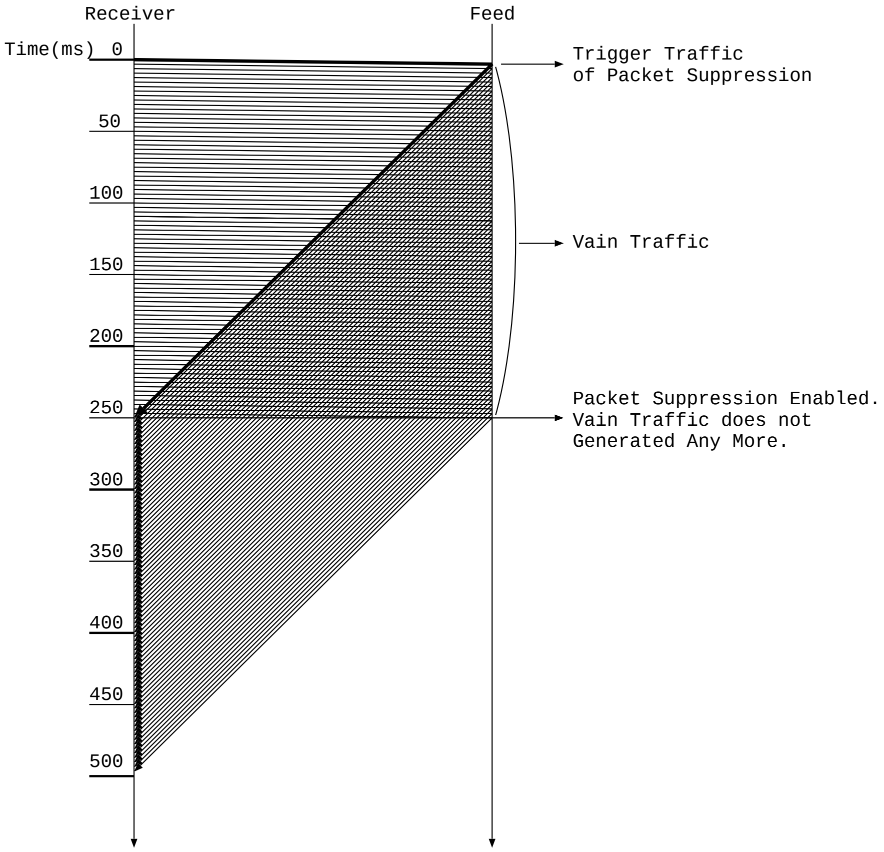

図3.2: 抑制機構が機能するまでに発生するトラフィック

衛星回線に転送される**IGMP**パケットによる帯域消費の試算

衛星回線に多数の受信者が接続するネットワークにおいて、各ReceiverからFeedを経由して衛星回線に転送されるIGMPパケットによる帯域消費を試算する。IGMPパケット

を含んでネットワーク上を伝送されるデータリンクパケットのサイズは、44Byte である。これを図 3.3 に示す。

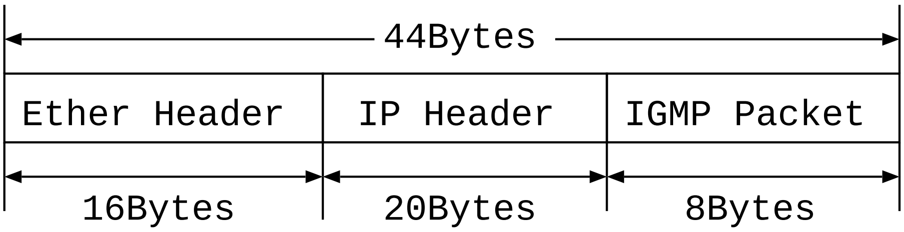

図 3.3: IGMP パケットを含むデータリンクパケットのサイズ

IGMP の仕様における最大応答時間の初期設定値は、100ms を 1 単位時間として 100 である。各ホストは、0 から最大応答時間までの単位時間をランダムな値でタイマを設定する。このとき、あるマルチキャストグループに参加しているホスト数を (n) とすると、1 単位時間あたりに送出される IGMP パケットによる帯域消費 (c) は、以下の式で表せる。

$$c(Byte) = \frac{n}{100} \times 44$$

他のホストから送出された Membership Report が到達するまで経過する単位時間 (t) あたりに、ある 1 つのマルチキャストグループに対して送出される Membership Report によって消費される帯域 (dc) を以下の式で表せる。

$$dc(Byte) = tc$$

参加するホスト数 (n) が 1000000 であるマルチキャストグループにおいて、衛星回線による 250ms の遅延を加味して経過する単位時間 (t) を 3 とするとき、3 単位時間あたりの帯域消費の合計 (dc) を以下のように試算できる。

$$\begin{aligned} dc(Byte) &= 3 \times \frac{1000000}{100} \times 44 \\ &= 1.32(MByte) \end{aligned}$$

参加するホスト数が多大である場合には、1 つのマルチキャストグループを管理するだけで、多くの帯域資源が消費される。複数のマルチキャストトラフィックを衛星回線を用いて配送する場合、それぞれのマルチキャストグループに参加するホスト数が多大であると、その規模性に対応できなくなる。

3.3 信頼性のあるデータ配信における問題点

低遅延な双方向通信路によって構成されたネットワークにおいて、多数の受信者を対象として信頼性のあるデータ配信を行う際に、NAK にもとづくデータの再送と遅延タイマによるパケットの送出抑制を組み合わせて行えば、パケットが送出される量が効率的に低減され、有効である。

しかし、衛星回線を含むネットワークにおいて同様のサービスを行う際には、各受信者から送出されたNAKパケットが他の受信者に伝送されるまでに衛星回線を経由する必要がある。この間の遅延によって、IGMPのパケット送出抑制機構と同様に、NAKのパケット送出抑制機構が正しく機能しない。その結果、多くの受信者から同一データに対するNAKが送出され、重複したNAKパケットが衛星回線に転送される。このとき、IGMPパケットによる帯域消費を試算した結果にあるように、NAKパケットによる衛星回線の帯域消費も大きいと考えられる。

3.4 本章のまとめ

UDLRでは、ネットワークの規模性の問題を議論の対象としていない。ネットワークの双方向性をBroadcast Emulationによって確保するため、Receiverからのトラフィックは、Feedによって衛星回線にそのまま転送される。受信者数が増大した際に、衛星回線の帯域消費が増大することは避けられない。

衛星回線が高遅延な通信路であるため、IGMPやNAKにおける遅延タイマを用いたパケットの送出抑制機構は完全には機能しない。このため、重複した内容のパケットが多数、衛星回線に転送される。必要以上に送出されるIGMPやNAKパケットが無駄なトラフィックとして衛星回線に転送される際の負荷が大きく、これを無視できない。

衛星回線が多くの受信者によって共有できる通信路であり、1対多型の通信方式であるマルチキャストとの親和性が高い。しかし、実用的なサービスを行うには、既存の機構では抑制しきれなかったトラフィックを如何に扱うかの問題を解決する必要がある。

第4章　解決手法の提案

衛星回線を片方向通信路として含んで構成される広域かつ大規模なネットワークにおいて、多数の受信者を対象にマルチキャストを用いてデータ配信を行う場合には、不要なパケットによるトラフィックが無視できない程増大し、帯域資源の無駄な消費が発生すると述べた。この問題を解決する手法として、衛星回線を含むネットワークでの使用に特化したパケット集約機構を提案する。

4.1　パケット集約機構の概要

パケット集約機構(Packet Aggregator: 以降、PAと記述)は、不必要なパケットが重複して衛星回線に転送されるのを防ぎ、通信路の帯域資源を有効に使用するための機構である。ある特定のプロトコル、またはホストに宛てられたパケットを捕捉し、定められたポリシに従って、それぞれ衛星回線に転送するかを選択する。これにより、衛星回線の帯域資源の無駄な消費を抑制する。これを図4.1に示す。

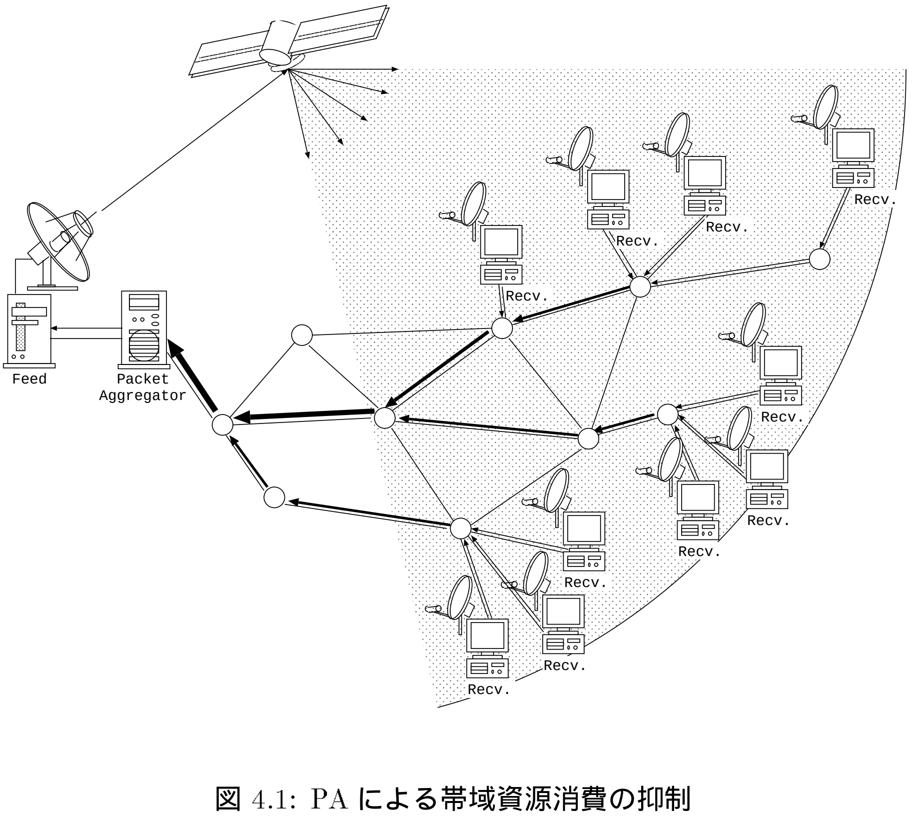

図4.1: PAによる帯域資源消費の抑制

衛星回線を含むネットワークでマルチキャストデータ配信を行う際に送出されるパケットのうち、IGMP パケットと NAK パケットに注目する。IGMP パケットと NAK パケットは、それぞれパケットの内容が重複する。これらのパケットを効率的に集約すれば、衛星回線の帯域資源を無駄な消費を少なくしてマルチキャストグループ管理、と信頼性を考慮したデータ配信が実現できる。両パケットに共通する特徴として、以下のような点が挙げられる。

- パケットの内容が重複している
- パケットがマルチキャストグループアドレスに向けて送出される
- 必ずしもすべてのホストがパケットを送出する必要があるわけではない
- 一度に多数のパケットが送出される
- パケット抑制機構が完全には機能しないために重複して送出されるパケットが帯域を無駄に消費する

PA は、各ホストから送出される IGMP パケットと NAK パケットの集約に焦点を当てる。重複した IGMP や NAK パケットが衛星回線に転送されると、それらが無駄なトラフィックとなって帯域資源を消費する。このとき、PA が、Receiver から UDL に向けて送出されたパケットのうち、IGMP と NAK パケットそれぞれについて、既に UDL に転送されたものと内容を比較し集約できる。これにより、IGMP と NAK パケットが衛星回線に転送される総量を必要最低限に抑制できる。

4.2 PA を用いたマルチキャストグループ管理

3 章に述べたように、既存の IGMP のパケット送出抑制機構は、衛星回線のような高遅延な通信路を用いて構成され、多数のホストが接続する大規模なネットワークでは完全には機能しない。このため、本来、マルチキャストグループ管理に必要かつ十分な IGMP パケットの他、これと重複する多数の IGMP パケットが衛星回線に転送される。この問題を解決するため、IGMP よるマルチキャストグループ管理に PA を適用する。

PA は、各 Receiver から送出された IGMP パケットを Feed に到達する前に捕捉し、宛先マルチキャストグループアドレスと IGMP のプロトコルタイプをもとに取捨選択する。これにより、不必要な IGMP パケットの転送を防ぎ、衛星回線の帯域資源の無駄な消費を防げる。これを図 4.2 に示す。

4.3 PA を用いた信頼性のあるデータ配信

一般的に、衛星回線によって接続する利用者は、衛星回線の持つ特性により、回線品質が概ね一様な通信路を共有する。このため、データが衛星回線に送信される時点で欠損が生じていれば、多くの受信者が同様の欠損を含むデータを受信する可能性が高い。受信したデータに欠損を検知した各ホストは、結果的に同一のデータに対して重複した NAK を送信することになる。

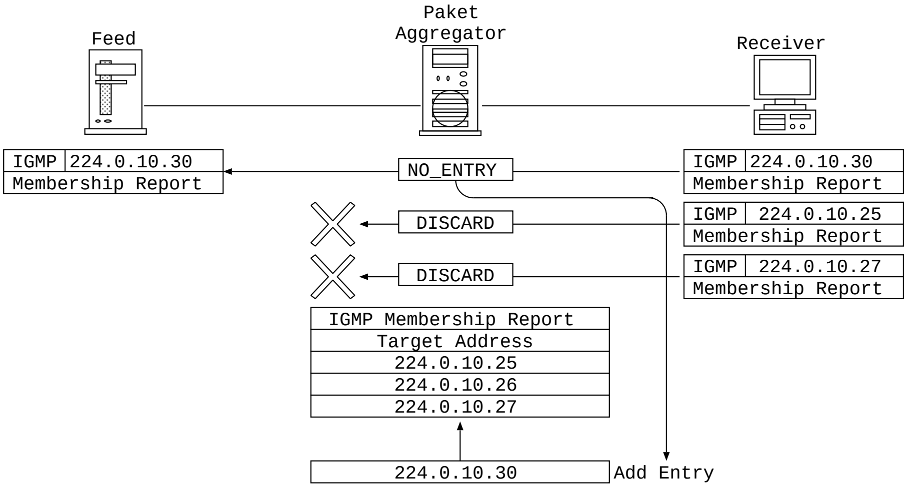

図 4.2: IGMP パケットの集約

衛星回線を用いたネットワークでは、あるホストから送出されたNAKパケットが他のホストに到達するまでの遅延が大きい。このため、NAKパケットの送出抑制に遅延タイマを用いても完全には機能せず、重複したNAKパケットが多数送出される。

このとき、PAを適用すれば、同一データの再送を要求するNAKパケットが、データの送信者に多数重複して到達するのを防げる。同時に、前述UDLRにおけるBroadcast Emulationによって重複パケットが衛星回線に転送されるのを防げる。これを図4.3に示す。

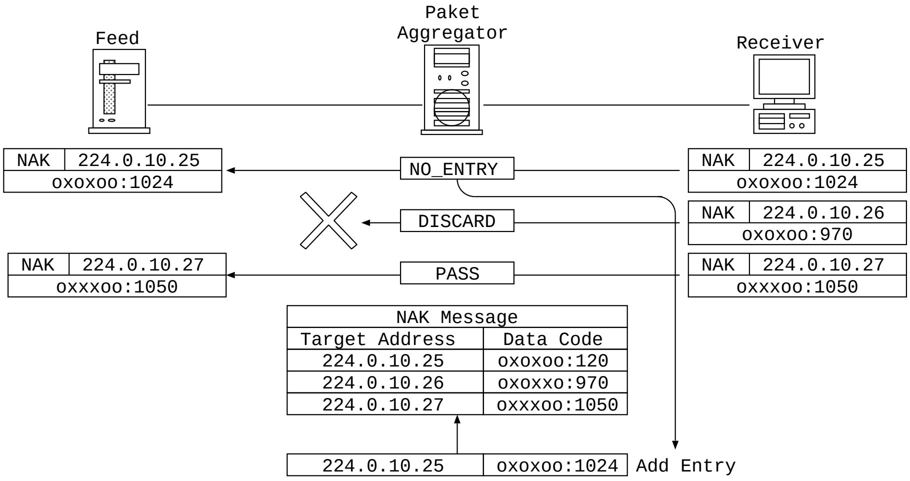

図 4.3: NAK パケットの集約

4.4　本章のまとめ

複数送出されるパケットの内容が同一であり、必ずしもすべてのホストが送出する必要がないパケットであれば、これらを一つに集約できる。このとき、IGMP や NAK パケットはその内容が重複し、必ずしもすべてのホストがパケットを送出する必要がない点に注目する。これらを効率的に集約すれば衛星回線に転送されるパケットの数量を低減できる。

これを実現する機構として、PA を提案した。PA は自分が Feed に転送したパケットの内容を記憶し、内容の重複するパケットを受信した際に、これを破棄する。これにより、各 Receiver から送出される重複パケットによって衛星回線の帯域消費が増大する問題を解決できる。

第5章　PAの設計

本章では、本論文で提案するPAの設計について述べる。

5.1　設計の方針

5.1.1　集約対象となるパケット

PAは、衛星回線を含むネットワークにおいて、遅延タイマを用いたパケットの送出抑制機構が完全には機能しないために発生する無駄なパケットを集約する。これにより、Feedに転送されるパケットの総量を抑制し、衛星回線の無駄な帯域消費を低減する。PAが集約の対象とするのは、IGMPパケットとNAKパケットである。

5.1.2　衛星回線に転送すべきパケット

各Receiverから送出されるパケットのうち、パケット送出抑制機構が機能するトリガーとなるものと、そうでないものがある。IGMPでは、マルチキャストルータからの問い合わせであるMembership Queryを受信すると、ランダムで設定されたタイマが直ちにタイムアウトするReceiverがある。当該Receiverから送出されるパケットが、他の受信者に到達することでパケット送出抑制機構が機能する。

Membership Queryを受信してから直ちに送出されたMembership Reportが他のReceiverに到達するまでの間に順次タイムアウトする多くのホストから、それぞれが無駄なパケットが送出される。これらのパケットのうち、パケットの送出抑制機構が機能するのに必要なパケットのみFeedによって衛星回線に転送されるのが望ましい。これを図5.1と図5.2に示す。

NAKパケットについてもIGMPと同様に、パケット送出抑制機構が機能するトリガーとして必要なパケットがFeedによって衛星回線に転送されるのが望ましい。

5.1.3　パケット集約方式

PAは、自分がFeedに転送したIGMPとNAKパケットの内容を、パケットの種類ごとに記録し、保持する。このための媒体をリストと呼び、パケットの内容をリストに格納するための媒体をエントリと呼ぶ。パケットの集約は、PAが保持するリストのエントリと受信したパケットの内容を比較し行う。パケットの集約は以下に示す処理を経る。

1. PAはReceiverからFeedに向けたパケットをすべて捕捉する

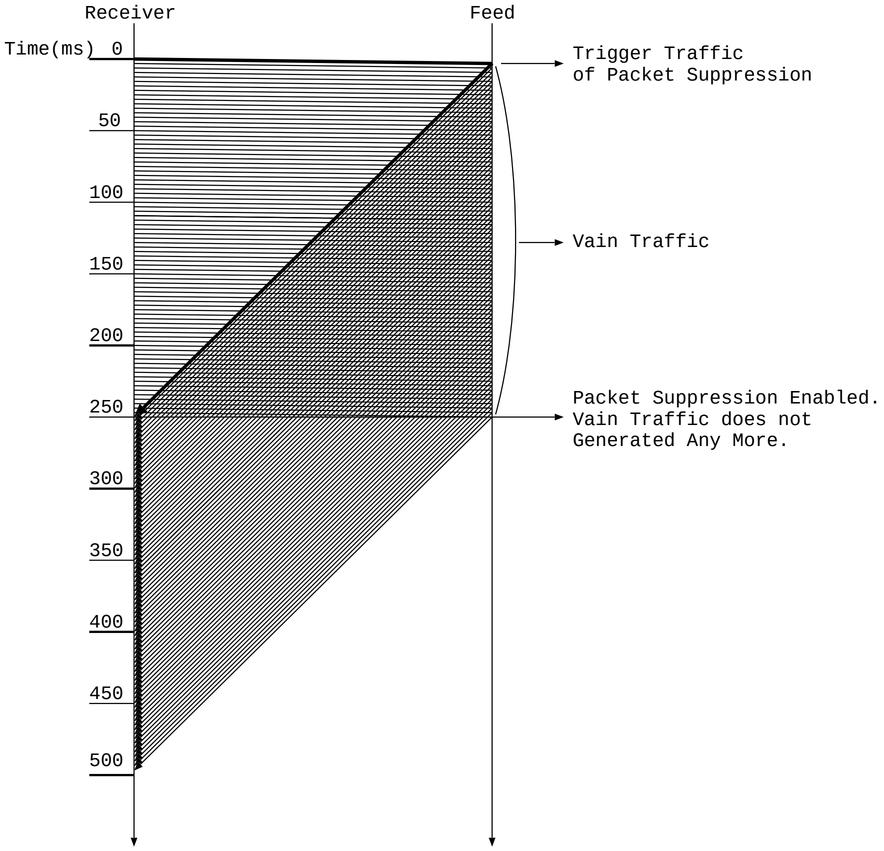

図 5.1: 抑制機構が機能するまでに発生するトラフィック

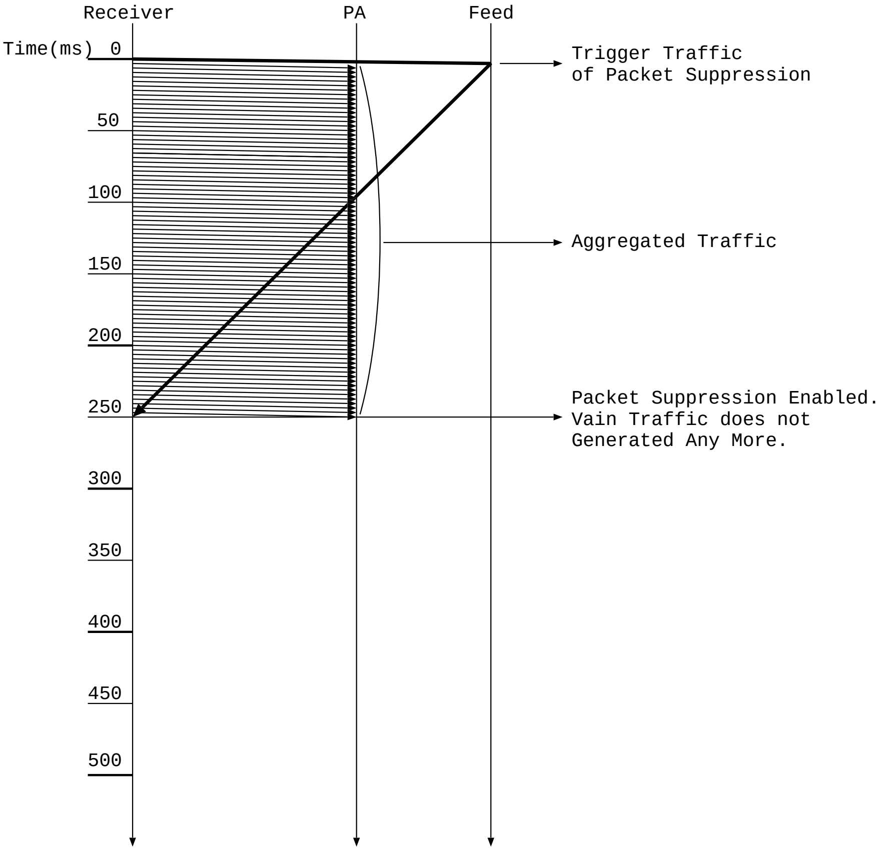

図 5.2: PA によって集約されるトラフィック

2. PA は、捕捉したパケットが集約の対象である場合、自分の持つリストを参照し、同一のパケットを既に Feed に転送したかを確認する
    - 自分が既に同一のパケットを Feed に転送していれば、受信したパケットを直ちに破棄する
    - PA が持つリストに無いパケットを受信した場合には、当該パケットを Feed に転送し、パケットの内容をリストに加える
3. PA は、保持するリストのエントリについて、一定期間が過ぎると、これを破棄する

5.1.4 パケットの集約ポリシの構成要素

PA がパケットを集約する際の仕様は、パケットの種類によって異なる。このため、パケットの種類ごとに集約の仕様を定める必要がある。これを集約ポリシと呼ぶ。集約ポリシに設定する内容を以下に示す。

- 集約するパケットの種類
- PA が持つパケットの種類ごとのリストの内容
- PA がリスト内にエントリを保持する時間

5.1.5 IGMP パケットの集約ポリシ

IGMP において、多数の Membership Report が集中して送出されるのは、Membership Query に対して、各受信者が返答を行う時である。

マルチキャストルータが Membership Query を送出する間隔、および Membership Report の最大応答時間は、ネットワーク管理者が一定の範囲内で設定する。IGMP パケットの集約ポリシは IGMP の仕様に合わせて設定できる必要がある。

5.1.6 NAK パケットの集約ポリシ

NAK を用いた信頼性保証の手法は、データ配信を行うアプリケーションの仕様に依存して、実際に用いられる NAK パケットの内容や送出条件などが異なる。このため、アプリケーションごとに集約ポリシを設定する必要がある。

5.2 PA を適用したネットワーク構成

PA は Feed に到達するパケットをその直前で捕捉する。集約の対象になっているパケットのうち、PA はパケット抑制機構が機能するトリガーとなるパケットを Feed に転送する。その後、Feed に転送したパケットが衛星回線を経由し、他の Receiver に到達するまでの間に順次到達する重複パケットを集約する。PA を含むネットワーク構成を図 5.3 に示す。

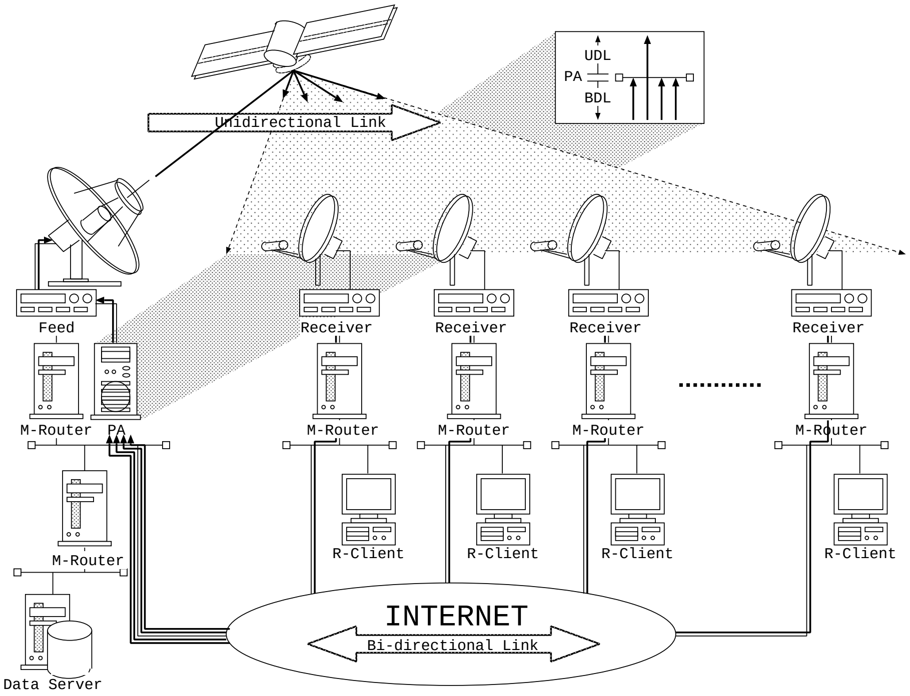

図 5.3: PA を適用したネットワークの構成

5.3 PA の構成

PA は以下に示すモジュール群によって構成される。

- パケット集約モジュール
  - パケットを捕捉し、当該パケットを転送するか破棄するかを判別し、実行する
- エントリ管理モジュール
  - パケットの集約ポリシに従って、エントリの生存時間を確認し、時間切れとなったエントリを削除する
- 集約ポリシ管理モジュール
  - ネットワーク管理者が設定する集約ポリシを読み込み、プロトコルごとに設定をリストに反映する

これを図 5.4 に示す。

5.4 パケット集約モジュール

5.4.1 モジュールの構成

パケット集約モジュールは、以下のルーチン群によって構成される。

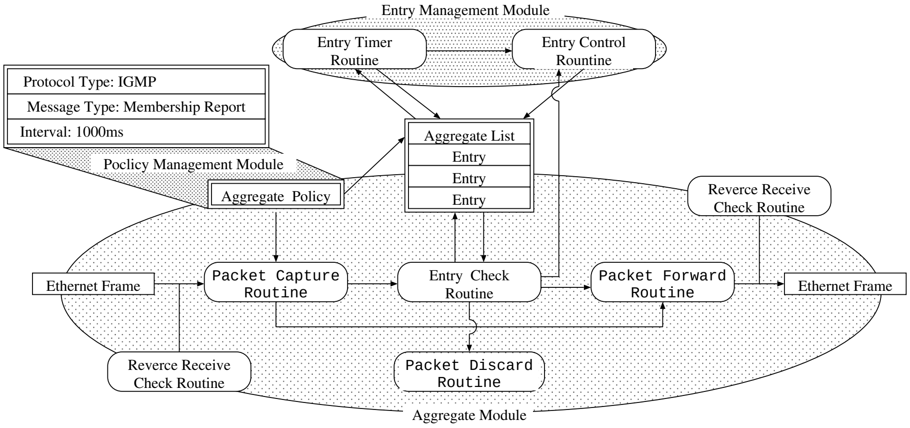

図 5.4: PA の構成

- パケット捕捉ルーチン
- リスト照合・判別ルーチン
- パケット転送ルーチン
- パケット破棄ルーチン

5.4.2 パケットの捕捉

パケット捕捉ルーチンは、BPF を用いて Receiver から Feed に向かうすべてのパケットを捕捉し、到達したパケットが集約の対象であるかを判別する。集約の対象であるパケットは、パケット照合・判別ルーチンに渡される。集約の対象外のパケットは、直ちに Feed に転送される。

5.4.3 パケットの照合と、転送または破棄の判別

パケット照合・判別ルーチンは、集約の条件となる各要素をパケット捕捉モジュールから渡されたパケットから抽出する。そして、パケットの各要素についてリスト内のエントリと照合する。以下に、IGMP パケットと NAK パケットを集約する際に、エントリと照合する要素を示す。

- IGMP パケットの集約に必要な要素
  - IGMP Protocol Type
  - Group Address
- NAK パケットの集約に必要な要素
  - Message Type

  - Data ID
  - Data Cell Sequence

パケットの各要素を照合した結果によって、当該パケットを Feed に転送するかを判別する。パケットを Feed に転送するかを判別する基準は、パケットの各要素を照合し、一致するエントリがリスト内にあるかである。

該当するエントリがリスト内にある場合は、パケットを破棄する。そうでない場合は、パケットの内容をリストのエントリとして新たに追加し、当該パケットを Feed に転送する。エントリの追加には、エントリ管理モジュールを呼ぶ。これを図 5.5 に示す。

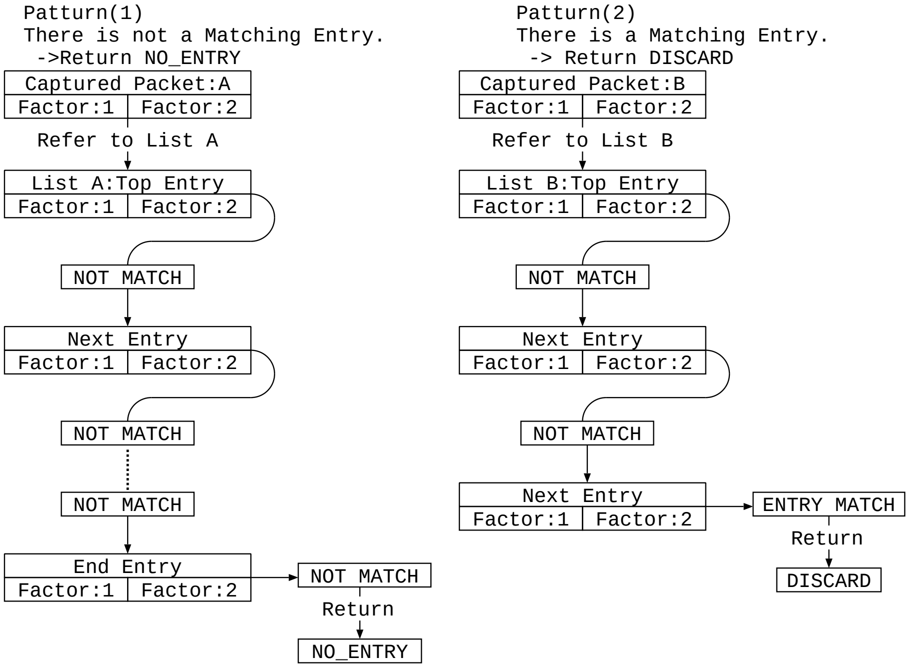

図 5.5: パケットの転送、および破棄の判別処理の概要

5.4.4 パケットの転送、および破棄

パケット転送ルーチンは、渡されたパケットを Feed に転送する。パケット破棄ルーチンは、渡されたパケットを破棄する。

5.5 エントリ管理モジュール

5.5.1 モジュールの構成

エントリ管理モジュールは以下のルーチン群によって構成される。

- 時刻管理ルーチン

- エントリ操作ルーチン

5.5.2 エントリの生存時間の管理

時刻管理ルーチンは、アグリゲートの対象となるパケットごとに設定されたポリシに従って、PA が保持するリスト内のエントリの生存時間を管理する。本ルーチンは、現在時刻を一定間隔で取得し、エントリが生成された際に付加されるタイムスタンプを参照する。設定された生存時間を過ぎたエントリは、リスト・エントリ操作ルーチンによって削除される。時刻管理ルーチンの概要を図 5.6 に示す。

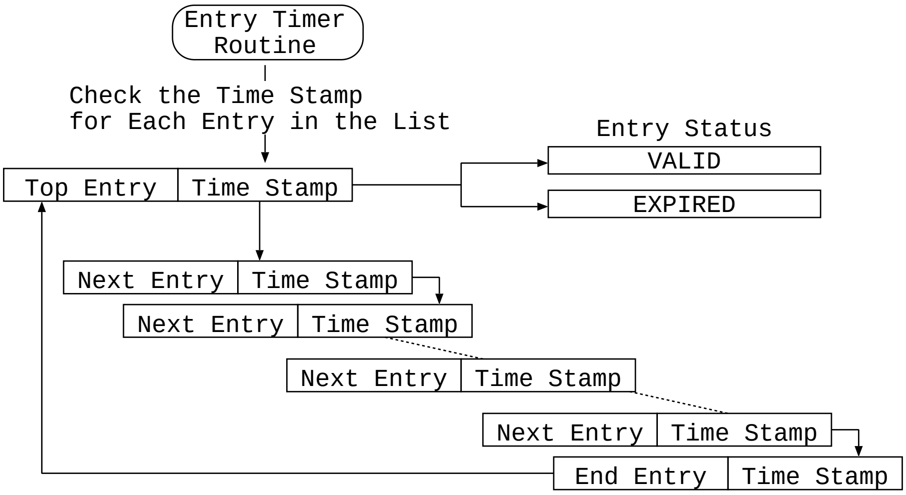

図 5.6: 時刻管理ルーチンの概要

5.5.3 エントリの操作

エントリ操作ルーチンは、エントリの追加と削除を行う。エントリの追加はパケット照合・判別ルーチンから呼ばれて行い、削除は時刻管理ルーチンから呼ばれて行う。これを図 5.7 に示す。

5.6 集約ポリシ管理モジュール

集約ポリシ管理モジュールは、ネットワーク管理者が集約ポリシを記述したファイルを PA の起動時に読み込み、リストの内容に反映する。

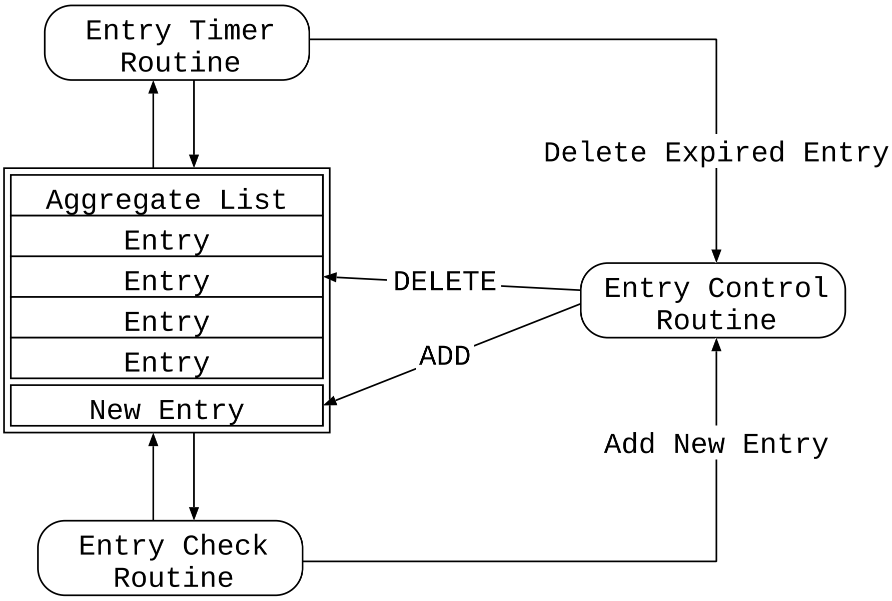

図 5.7: エントリ操作ルーチンの概念図

第6章　PAの実装

本章では、5章で述べた機構の実装について述べる。本研究では、IGMPパケットとNAKパケットの集約のうち、IGMPのMembership Reportの集約に焦点を絞り、この処理を行うモジュールを実装した。

6.1　実装環境

PAの実装環境を表6.1に示す。

表 6.1: PAの実装環境

Operation System	FreeBSD 4.4-RELEASE
CPU	Pentium III 1GHz
Memory	512MByte
Hard Disk	80GByte
gcc version	2.95.3

6.2　集約ポリシの設定

6.2.1　pa_policy構造体

pa_policy構造体は、プロトコルごとにポリシを格納する構造体である。集約ポリシの設定は固有のファイルから読み込む。PAは、起動時に設定ファイルの各項目をpa_policy構造体に格納し、これをもとにパケット集約モジュールとエントリ管理モジュール間で共有するリストを初期化する。

```
struct pa_policy {
    u_char   protocol_type;
    u_char   message_type;
    char     m_addr[INET_ADDRSTRLEN];  /* Multicast IP Address */
    char     u_addr[INET_ADDRSTRLEN];  /* Unicast IP Address */
    int      interval;
};
```

図 6.1: pa_policy構造体

6.2.2 IGMP パケットの集約ポリシの設定

IGMP パケットの集約ポリシの設定には、IGMP のプロトコルタイプ、パケットの種類、およびエントリの生存時間を指定する。本実装では、全てのマルチキャストグループに対して送出された Membership Report を集約するため、マルチキャストグループアドレスの項目には特定のアドレスを指定しない。IGMP パケットの送信元は考慮しないため、ユニキャスト IP アドレスの項目には特定のアドレスを指定しない。Membership Report の集約ポリシの設定ファイルを図 6.2 に示す。

```
[Policy-1]
       Protocol Type:IPPROTO_IGMP
        Message Type:IGMP_V2_MEMBERSHIP_REPORT
Multicast IP Address:LEARN
  Unicast IP Address:IGNORE
            Interval:10000
```

図 6.2: 集約ポリシ設定ファイルの例

集約ポリシの設定に固有の値を図 6.3 に示す。

```
#define LEARN      0x01    /* Learn Address in Process of Aggregation */
#define IGNORE     0x02    /* Ignore Specefied Address Field */
#define DEFAULT    250     /* Default Interval(ms)*/
```

図 6.3: 集約ポリシの設定に固有な値の定義

Membership Report の集約ポリシでは、マルチキャストグループアドレスには常に LEARN を、ユニキャスト IP アドレスには IGNORE を指定する。

6.3 集約するパケットの判別

捕捉したパケットが集約対象であるか否かを判別するには、捕捉したパケットに含まれる各ヘッダのプロトコルタイプを参照する。パケット判別ルーチンは、捕捉したデータリンクパケットに含まれる各ヘッダのプロトコルタイプを参照し、集約ポリシに設定されたプロトコルタイプと一致するかを調べる。これを図 6.4 に示す。

6.4 IGMP パケットの転送、または破棄の判別

IGMP パケットのリスト照合ルーチンは、集約対象のパケットを渡されると、宛先マルチキャストグループアドレスを、自分が保持するリストの各エントリに記録されているマルチキャストグループアドレスと照合する。このとき、既に Feed に転送したパケットのエントリとマルチキャストグループアドレスが一致した場合には、DISCARD を返す。該当するエントリが無い場合には、NO_ENTRY を返す。これを図 6.5 に示す。

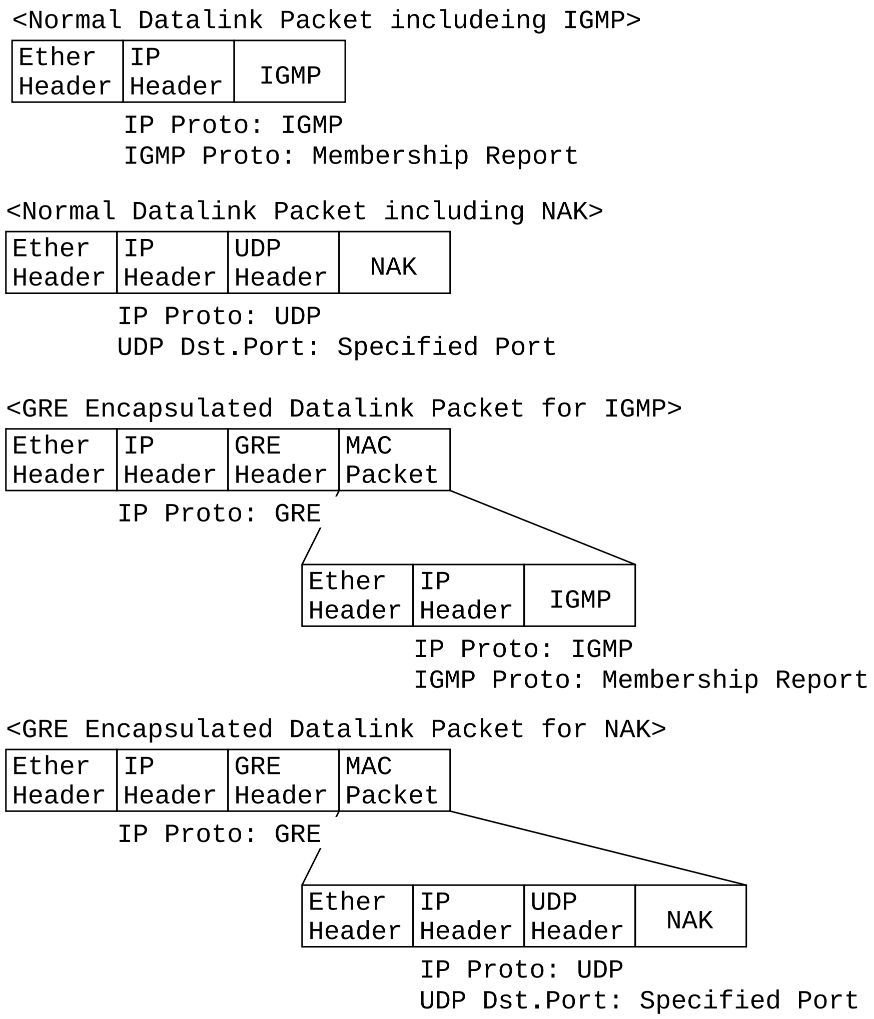

図 6.4: 捕捉したパケットにおける集約対象の判別

```
temp_entry = list->top;
while (temp_entry != list->end) {
    if ((strcmp(temp_entry->group_addr, addr)) == 0) {
        return(DISCARD)
    }
    temp_entry = temp_entry->next_entry;
}
return(NO_ENTRY);
```

図 6.5: IGMP のリスト判別ルーチンの一部

6.5 エントリの生存時間の判別

IGMPパケットの内容をリストに記録するエントリは、以下のように生存時間を確認する。リスト内のエントリについて、生存時間が過ぎているものは削除する。これを図6.6に示す。

```
if ((igmp_entry_timer(temp_entry, current_time)) == EXPIRED) {
    temp_entry = delete_igmp_entry(list, temp_entry);
    list->entry_count--;
}
```

図 6.6: IGMPのエントリ管理ルーチンの一部

6.6 BPFを用いることによるループ発生の回避

PAはBPFを用いてパケットを捕捉し転送を行う。このため、PAの送信ルーチンがパケットを送信すると同時に、BPFがこれを受信する。PAにおいてパケットのループが発生する。PAが転送したパケットを受信しても、これを直ちに破棄し、パケットのループを防ぐ必要がある。

PAでは、送信ルーチンは送信したパケットのMACアドレスを記録し、共有する。送信ルーチンはパケットを送出する際に、パケットの宛先MACアドレスをバッファし、送出フラグを立てる。捕捉ルーチンがパケットを受信した際に、フラグが立っていたら受信したパケットのMACアドレスを確認する。このとき、PAが送出したばかりのパケットであれば、これを破棄する。

第7章　評価

本章では、PA を評価ネットワークで動作させた際の評価について述べる。

7.1　評価項目

- PA を設置したネットワークにおける帯域の測定
- パケットの集約を行った際に PA が転送するパケット数の測定
- UDL の帯域消費の試算

7.2　評価環境

7.2.1　評価ネットワークの概要

PA の評価を行うため、以下の図 7.1 に示すネットワークを構築した。M-Router は 2 つのインターフェースを持ち、Network A と Network B とを接続する。M-Router ではマルチキャスト経路制御プロトコルとして DVMRP[3] を動作させる。本トポロジでは、M-Router は Feed を兼ねる。従って、M-Router に到達するパケットは、実ネットワーク上では UDL に転送されることになる。

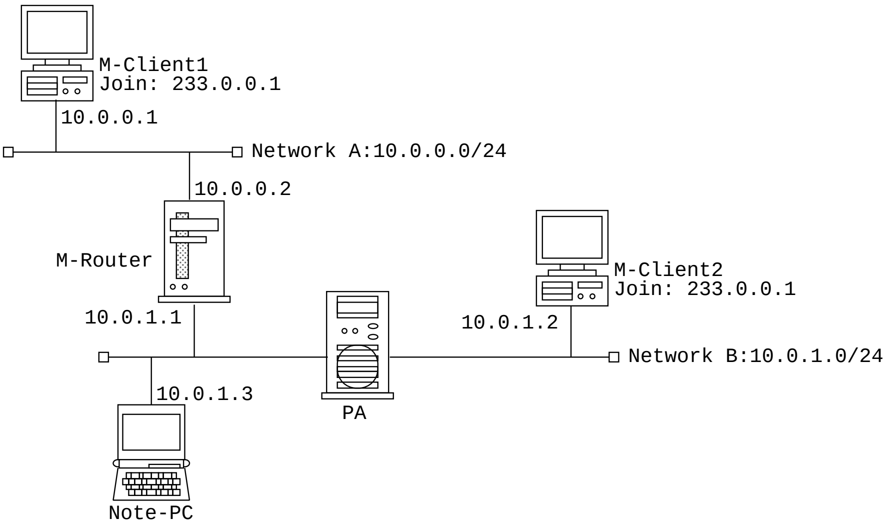

図 7.1: 評価ネットワーク

Netowrk A に接続する M-Client1 と Network B に接続する M-Client2 は、それぞれマルチキャストグループ 233.0.0.1 に参加する。Network A、B に接続する M-Router は、M-Client1 と M-Client2 からの Membership Report を受信する。このとき、M-Router の持つマルチキャストルーティングテーブルには、それぞれのリンクでメンバーが存在するマルチキャストグループが登録され、当該マルチキャストグループの状態管理を開始する。

7.2.2 評価に用いたマシンの仕様

評価に用いたマシンの仕様を以下の表 7.1 に示す。

表 7.1: 評価に使用した各マシンの仕様

	Operation System	CPU	Memory	NIC
PA	FreeBSD 4.4-RELEASE	Pentium III 1GHz	512MByte	10Mbps
M-Router	FreeBSD 4.4-RELEASE	Pentium III 1GHz	512MByte	10Mbps
M-Client	FreeBSD 4.4-RELEASE	Pentium III 1GHz	512MByte	10Mbps
Note-PC	FreeBSD 4.4-RELEASE	Pentium 233MHz	32MByte	10Mbps

7.3 PA によるネットワークのスループットへの影響の測定

ネットワークに PA を設置した場合と、そうでない場合のネットワークのスループットについて、FTP によるファイル転送を行い測定した。

ネットワークに PA を設置せず、M-Router と M-Client2 を直接接続した。3.5MByte のファイルを転送し、FTP クライアントが表示するスループットを記録した。その後、ネットワークに PA を設置し、パケットの集約を行わない状態で同様に測定した。FTP によるデータ転送は、PA の設置前後でそれぞれ 20 回ずつ行った。このときの測定結果を図 7.2 に示す。

PA を設置しない状態でのファイル転送を行った際のスループットの平均は 942.1Kbps であった。PA を設置し、集約を行わない状態でファイル転送を行った際のスループットの平均は 937.6Kbps であった。この結果、PA を設置した状態でのスループットは、設置しない状態でのスループットに比べおよそ 4Kbps ほど低下し、その割合はおよそ 0.5%であった。

7.4 PA が集約を行い、転送するパケット数の測定

7.4.1 パケット複製ツールによるトラフィックの発生

PA が転送するパケット数を測定するにあたり、同一セグメント上で重複するパケットによるトラフィックを大量に発生させるために、パケット複製ツール (Packet Reproducer: 以降 PR と記述する) を作成した。PR は起動時に複製するパケットの種類を指定する。

PA は BPF によって接続するリンクのすべてのパケットを捕捉し、パケットの内容を参

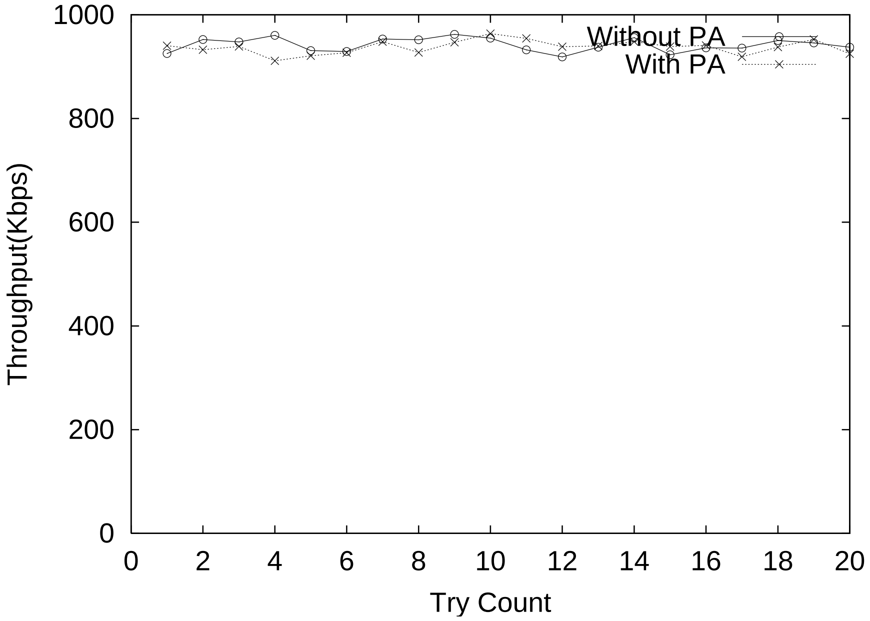

図 7.2: FTP によるスループットの測定結果

照しながら目的の種類のパケットが到達するまで待機する。目的の種類のパケットを捕捉すると、これをデータリンクパケットの状態で一定期間連続して BPF に書き込む。

本研究では、PR を Network B の M-Client2 に実装した。M-Router、または M-Client2 が Membership Report を送出した際に、当該パケットが多数複製される。これにより、Network B のセグメント上で、同一のマルチキャストグループアドレスに向けて多数のパケットが送出される状態を再現した。

7.4.2 PA が転送したパケット数の測定

パケットの集約ポリシの設定

PS における集約ポリシの設定を表 7.2 に示す。

表 7.2: PA の設定

PA の設定	
集約対象となるパケットの種類	IGMP_V2_MEMBERSHIP_REPORT
送信先マルチキャストグループアドレス	LEARN
送信元ユニキャスト IP アドレス	IGNORE
エントリの生存時間	1000ms

PR の設定

PR の設定を表 7.3 に示す。

表 7.3: PR の設定

PR の設定	
複製するパケットの種類	IGMP_V2_MEMBERSHIP_REPORT
パケットの送出期間	100ms

予測される測定結果

PA では、Network B にトラフィックが流れる時間より十分長くエントリの生存時間を設定した。このため、Network B に流れるトラフィックのうち、最初に PA に到達する Membership Report の内容を学習し、それ以降到達する同一内容の Membership Report はすべて PA によって破棄されなければならない。M-Router にはパケットが 1 つだけ転送されるはずである。

測定結果

PA でパケットの集約を行い、PA が転送したパケット数を測定した。測定は、10 回行った。その結果を以下の図 7.3 に示す。

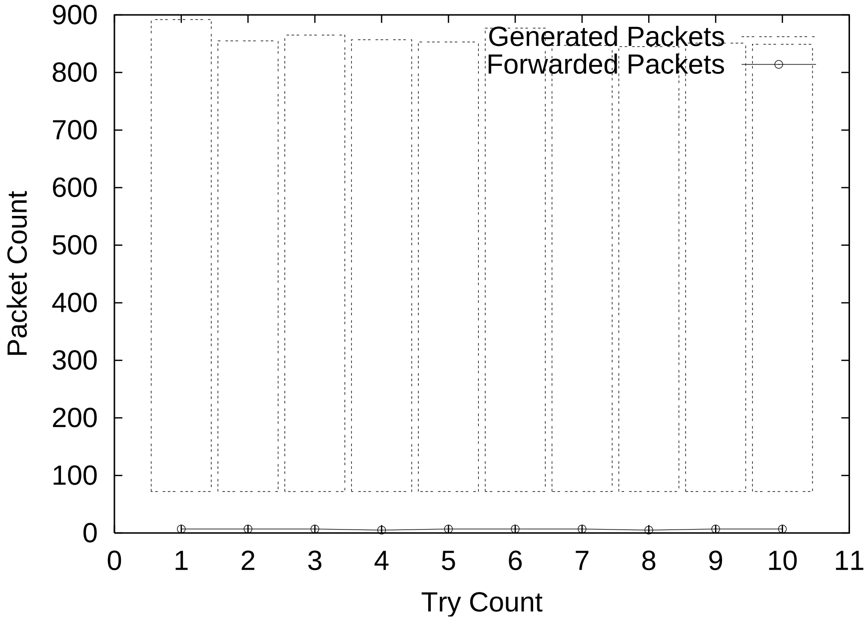

図 7.3: エントリの生存時間を 1000ms に設定した際の測定結果

測定結果から、PA が捕捉したパケット総数の平均は、およそ 859 であった。このうち、PA によって転送されたパケット総数の平均はおよそ 6 であった。

また、PR の設定には変更を加えず、PA の設定について、エントリの生存時間を 2000ms、10000ms にそれぞれ変更し、10 回ずつ PA によって転送されるパケット数の測定を行った。その結果、最初に行った測定結果とほぼ同様の結果を得た。これらの結果を図 7.4 と図 7.3 に示す。

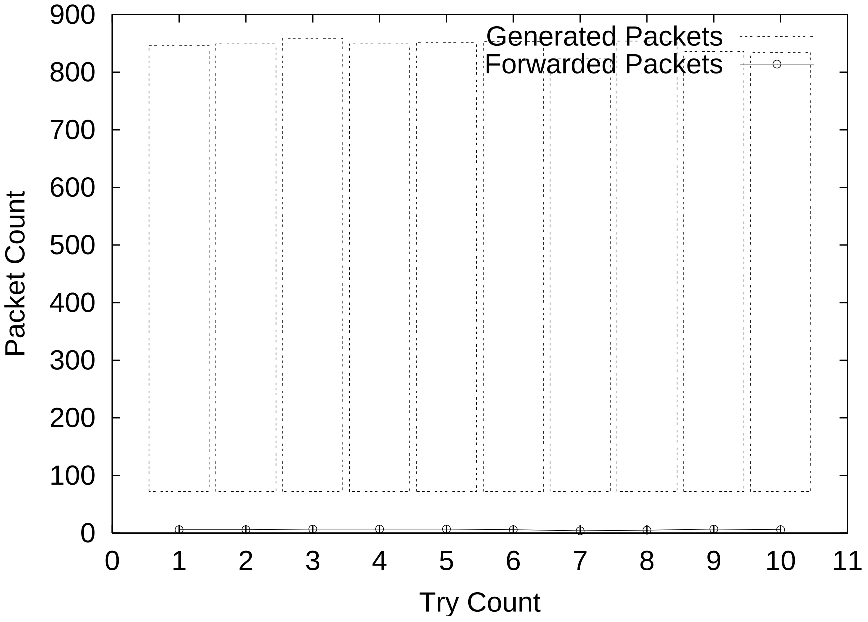

図 7.4: エントリの生存時間を 2000ms に設定した際の測定結果

限界性能の測定

PA の限界性能について、PR が送出する重複パケットの数を調節し測定した。その結果、0.1 秒間におよそ 3800 個のパケットを処理できた。0.1 秒間に集中するパケット数をおよそ 3800 個から 3900 個となるよう負荷を増大させせたとき、PA はコアダンプをともなう異常終了が発生した。

7.4.3 測定結果の考察

パケット数の測定結果の考察

PA の集約ポリシ、重複パケットが連続して送出されても、これらを処理するのに十分な長くエントリの生存時間を設定した。しかし、実際には複数のパケットが転送され、PA から重複したパケットが M-Router に到達した。

この理由について、M-Client2 と M-Router の直近に設置した Note-PC において、tcpdump を用いて検証した。その結果、 PA が学習していないマルチキャストグループドレ

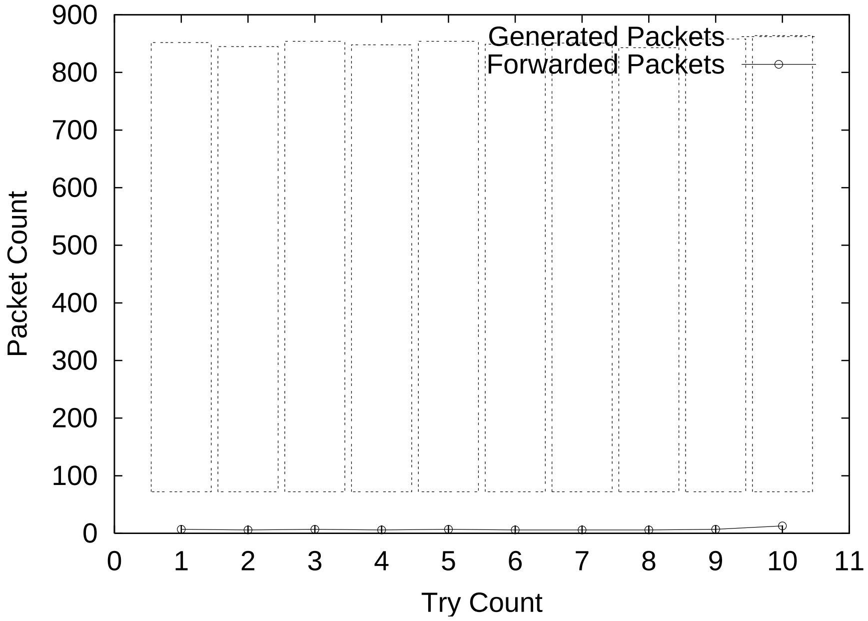

図 7.5: エントリの生存時間を 10000ms に設定した際の測定結果

ス宛の Membership Report が送出されており、これを転送したためだった。しかし、別のアドレスに転送されたパケットは 2 または 3 であり、依然として集約対象のパケットが転送されていた。

集約対象のパケットが重複して転送された原因として BPF を用いてパケットを捕捉し、転送している。パケットの集中によって PA に負荷がかかった際に、PA の各モジュールが正しく機能しなかったか、BPF のループ制御が正しく行われなかった可能性がある。

パケット数の測定結果にもとづく帯域消費の試算

1 単位時間あたりに同一のマルチキャストグループアドレス宛に送出された IGMP パケット数 (p) による帯域消費 (c) は以下の式によって表せる。

$$\begin{aligned} c(Byte) &= p \times 44 \\ &= 850 \times 44 \\ &= 37400 \\ &= 37.4(KByte) \end{aligned}$$

このとき、3 単位時間 (t) あたりに 1 つのマルチキャストグループアドレスに対して送出された Membership Report による帯域消費の合計 (dc1) は以下の式で表せる。

$$\begin{aligned} dc(Byte) &= tc \\ &= 3 \times 850 \times 44 \\ &= 112200 \\ &= 112.2(KByte) \end{aligned}$$

また、dc からこのマルチキャストグループに参加していると想定されるホスト数 (n) を算出すると、

$$
\begin{gathered}
dc(Byte) = 3 \times \frac{n}{100} \times 44 \\
3 \times \frac{n}{100} \times 44 = 112200 \\
n = 85000
\end{gathered}
$$

となる。1 単位時間に PA が M-Router に転送したパケットのうち、集約対象となっていたパケット (pa) が衛星回線の帯域消費 (dc2) は、

$$
\begin{aligned}
dc(Byte) &= 4 \times 44 \\
&= 176 \\
&= 0.2(KByte)
\end{aligned}
$$

であるため、実際に送出された Membership Report による帯域消費 (dc1) を 1%以下に集約できる。

限界性能の測定結果の考察

PA が IGMP の 1 単位時間あたりに処理できる最大パケット数を 3000 個とした場合、IGMP の最大応答時間を初期設定値の 100 単位時間とすると、PA が少なくとも 300000 ホストから送出される重複パケットを処理できる。また、最大応答遅延の値を設定上限である 255 とした場合、PA がおよそ 750000 ホストから送出される重複パケットを処理できることになる。PA が大規模ネットワークにおいても有効に機能する値を示したと言える。

7.5 評価のまとめ

PA を設置することによってネットワークのスループットに影響は少なく、大規模なネットワーク上に設置するに足る性能を持つと言える。しかし、大量のパケットが送出された際に、PA によって M-Router への重複したパケットの転送が認められたので、PA が完全に機能しているとは言えない。

重複するパケットが多数送出されるネットワークにおいて PA を機能させた場合、PA にパケットが集中してもその大部分を集約できた。PA を用いれば、重複するパケットが多数 UDL に転送されるため、UDL の帯域消費が増大するという問題は概ね解決できた。

第8章　まとめと今後の課題

8.1　まとめ

本研究では、衛星回線を片方向通信路として含むネットワークにおいて、多数の受信者を対象として信頼性のあるマルチキャストデータ配信を行う際に発生する、規模性の問題を解決した。

広域性、同報性を持つ衛星回線を通信路として含むネットワークにおいて1対多型の通信方式であるマルチキャストを用いれば、多数の利用者に対して効率的にデータを配信できる。衛星回線は少数の送信ホストに対して多数の受信ホストが接続する片方向通信路として用いるのが一般的である。UDLRを用いれば、片方向通信路である衛星回線を仮想的なブロードキャスト型の双方向通信路として用いられる。これにより、既存のインターネットにおいて通信の双方向性を前提として動作する各システムが、衛星回線を片方向通信路として用いる場合でも、正しく機能する。

しかし、衛星回線を含むネットワークでにおいて、多数の受信者を対象とする信頼性のあるマルチキャストデータ配信を行う際には、衛星回線に特有の遅延とネットワークの規模性に起因してUDLの帯域消費の増大する問題が生じる。UDLRではReceiverから衛星回線に向けて送出された全てのパケットが、Broadcast Emulationによって衛星回線にそのまま転送される。また、既存のマルチキャストグループ管理やデータの信頼性保証におけるパケットの送出抑制機構が機能するまでの遅延が大きいため、重複するIGMPやNAKパケットが多数送出される。このとき、衛星回線の帯域資源が無駄に消費される。

本研究では、上述の問題点を解決するため、UDLR環境での使用に特化したパケット集約機構(PA)の適用を提案した。PAは集中して到達する重複パケットのうち、Feedに転送すべきものと、PAにおいて破棄するものとを集約ポリシに従って選択する。

既存のパケット送出抑制機構が正しく機能するトリガーとなるパケットだけを直ちにFeedに転送し、パケット抑制機構の機能が不完全な間に順次送出されるパケットはFeedに転送せず、破棄する。重複パケットを有効に集約すれば、衛星回線の帯域消費を低減できる。この機構を実現するために、パケット集約機構を設計、実装した。

本研究では、集約が可能なパケットのうち、IGMPパケットの集約を行う機構を実装した。評価ネットワークにおいてPAの性能を評価した。その結果、単位時間内に送出されたパケット数に対する、PAがM-Routerに転送したパケット数の割合は、ネットワークの規模性に対応できる結果を示していた。衛星回線に多数の受信者が接続するネットワークにおいてPAを用いれば、衛星回線の帯域消費を十分に抑制できることを示した。

本研究により、多数の受信者が接続する衛星回線をインターネットの通信路の一部として用いた、大規模なマルチキャストデータ配信が実現できる。

8.2 今後の課題

本研究では、集約するパケットの種類をIGMPパケットに限定して実装した。今後、本研究において論じたNAKパケットを集約する機構の実装を継続する。また、全体的に処理能力を向上させるため、PAをカーネル内で動作させ、より複雑な集約ポリシが設定できるようにする。パケットの集約の汎用性について評価を行い、衛星回線をUDLとして含むネットワークにおいて、実用に耐える性能が得られるかを評価し、検証を行う予定である。

謝辞

本研究を進めるにあたり、御指導をいただきました、慶應義塾大学環境情報学部教授の徳田英幸博士、並びに村井純博士に深く感謝いたします。また、絶えず御助言と御指導をいただきました、同学部助教授の楠本博之博士、並びに中村修博士に深く感謝いたします。
　本研究の全般に渡り、終始惜しみない御助言と叱咤激励をしていただきました、慶應義塾大学政策メディア研究科の西田視磨氏、慶應義塾大学SFC研究所の泉山英孝氏に深く感謝いたします。また、本論文の執筆に際して、常に励ましと御協力をいただきました、慶應義塾大学政策メディア研究科の村上陽子氏、藤枝俊輔氏、小川浩司氏、環境情報学部の工藤信篤氏、熊木美世子氏をはじめとする、慶應義塾大学環境情報学部の徳田・村井・楠本・中村・南研究会の諸兄に深く感謝いたします。

参考文献

[1] W. Fenner. *Internet Group Management Protocol, Version 2*, November 1997. RFC 2236.

[2] E. Duros, W. Dabbous, H. Izumiyama, N. Fujii, and Y. Zhang. *A Link-Layer Tunneling Mechanism for Unidirectional Links*, March 2001. RFC 3077.

[3] D. Waitzman, C. Partridge, and S.E. Deering. *Distance Vector Multicast Routing Protocol*, November 1988. RFC 1075.