

修士論文 1999年度 (平成11年度)

多段通信路ネットワークにおける信頼保証型マルチキャストデータ配送システムの設計と実装

慶應義塾大学大学院政策メディア研究科

氏名：西田 視磨

**多段通信路ネットワークにおける信頼保証型
マルチキャストデータ配送システムの設計と実装**

本論文では、従来のマルチキャストデータ配送システムの欠点を克服し、有効にマルチキャストを用いた信頼性のあるデータ配送を行う、Segmented Hierarchical Reliable Multicast (SHRM:セグメント化階層構造信頼性マルチキャストデータ配送システム)を設計、実装した。

従来のマルチキャストを用いた信頼保証型データ配送では、信頼性の不足、大規模性の不足、配送が効率的でない、といった欠点があり、実用的でない。本研究で提案する方式では、ネットワークをマルチキャスト可能なローカルセグメント (MLS: Multicastable Local Segment) に分割し、この MLS 内でデータ配送システムを提供する。データ配送の信頼性保証を行う責任範囲を限定することにより、効率的な信頼保証データ配送を実現した。また、各々自立した MLS が相互に協調することにより、ネットワーク上で規模性を維持しながら信頼性を保証するデータ配送を実現した。

本研究で提案するシステムでは、ネットワークを MLS に分割し、広帯域なものから狭帯域なものへと接続された MLS の階層構造が与えられた時に、最上位 MLS にデータソースを配置し、この上で有効な信頼保証型データ配送を行う。本システムでは、MLS 内でデータ配送の信頼性を保証するプロトコル、MLS 間を接続し協調させるプロトコル、データソースを最上位 MLS まで搬送するプロトコルを提供する。

本論文では、本研究の新しい手法を提案し実装した。本論文で提案する手法では、従来の信頼保証型マルチキャストデータ配送と比較して、効率よくデータが配送できる。

本論文で提案する手法を利用することにより、インターネット上で、デジタルコンテンツの配信などの場面で効率的に信頼性のあるデータ配送が行えるようになる。

キーワード

1. インターネット 2. マルチキャスト 3. 信頼性 4. データ配送 5. MLS

慶應義塾大学大学院政策・メディア研究科

西田視磨

SHRM: A Segmented Hierarchical Reliable Multicast

This paper proposes the "SHRM, Segmented Hierarchical Reliable Multicast", which is an efficient multicast data distribution architecture. This architecture overcame the weakness of conventional multicast data distribution systems.

Conventional reliable multicast techniques are useless because of scrumpy reliability, low scalability and lack of distribution efficiency. This paper suggests a new approach called "MLS : Multicastable Local Segment". SHRM technique separates wide area network to plural local segments, which have multicast capability, and guarantees transport reliability. SHRM guarantees reliability in limited the area, so that efficient reliable data distribution system is feasible. The collaboration of each MLS enabled scalable and reliable data distribution system on the network.

SHRM assumes plural MLSs compose hierarchy by descending order, and data source are put inn top-level MLS. This condition provides best efficiency in this mechanism.

This paper defines three protocols:

- 1) Protocol for Reliable multicast data transport in MLS internally
- 2) Protocol for harmonizing other external MLS
- 3) Protocol for carrying source data to top-level MLS

This paper discusses implementation of SHRM and present excellence of the system compared to conventional reliable multicast data distribution system. SHRM enables efficient and reliable data distribution on the Internet.

Key Word

1.Internet 2.Multicast 3.Reliable 4.Data Distribution 5. MLS

Keio University Graduate School of Media and Governance

Mikiyo Nishida

目次

第1章 序論	-----	1
1.1. 現在のインターネットを取りまく背景	-----	1
1.2. 現状での問題	-----	2
1.3. 本研究の目的	-----	2
1.4. 本論文の構成	-----	2
第2章 信頼保証型データ配送システムについての考察	-----	4
2.1. データ配送システム実現のための前提	-----	4
2.1.1. 配送データの巨大化	-----	4
2.1.2. 利用者数	-----	4
2.1.3. インターネット上の通信路とデータ配送システム	-----	5
2.2. データ配送における既存の枠組みの欠点	-----	5
2.2.1. データ配送に対する要求の種類	-----	5
2.2.2. ユニキャストとマルチキャスト	-----	6
2.3. データ配送システム実現上の問題点	-----	7
2.3.1. 実時間性と信頼性との関係	-----	7
2.3.2. 信頼保証手法と性能	-----	8
2.4. システム構築の方針	-----	10
第3章 関連研究	-----	12
3.1. Chang and Maxemcunk と RMP	-----	12
3.2. RMTP:A Reliable Multicast Transport Protocol	-----	13
3.3. SRM (Scalable Reliable Multicast)	-----	13
3.4. 既存の Reliable Multicast のまとめ	-----	14
第4章 解決手法の提案	-----	15
4.1. MLS 手法の導入	-----	15
4.1.1. MLS の定義	-----	15
4.1.2. MLS の設定	-----	15
4.2. MLS 分割による範囲限定型信頼保証	-----	16
4.3. MLS 手法を利用したデータ配送	-----	17
4.3.1. MLS 間でのデータの配送	-----	17

4.3.2. MLS 階層構造の形成	18
4.3.3. 配布データ配置	18
4.4. MLS 階層構造モデル	19
第5章 設計	20
5.1. システム設計の前提	20
5.2. データ構造	20
5.2.1. 配送単位	20
5.2.2. 配布データの有効期限	21
5.3. MLS 内データ配送	21
5.3.1. MLS の設定	21
5.3.2. データ配送	22
5.3.3. データ広報	22
5.3.4. データソースホストにおけるデータ配送	22
5.3.5. レシーバにおけるデータ受信	26
5.3.6. パケットのフォーマット	29
5.4. MLS 間データ配送	32
5.4.1. MLS の接続	32
5.4.2. MLS のデータ中継	33
5.5. 配布データ移送	33
第6章 実装	35
6.1. 実装環境	35
6.2. アプリケーション動作環境	35
6.3. アプリケーション動作	35
6.4. アプリケーション設定項目	36
6.4.1. アプリケーション動作の設定	36
6.4.2. 上位 MLS 接続インターフェースの設定	37
6.4.3. 下位 MLS 接続インターフェースの設定	37
6.5. アプリケーション構成	38
6.5.1. 配送データ管理モジュール	40
6.5.2. SHRM データ広報モジュール	41
6.5.3. SHRM データ配送モジュール	41
6.5.4. データ受信モジュール	42
6.5.5. SHRM 制御モジュール	43
6.5.6. SHRM データ中継モジュール	43
6.5.7. ネットワーク・アダプタ操作モジュール	43

6.5.8. ユーザーインターフェース・モジュール	_____	44
6.5.9. アプリケーション管理モジュール	_____	47
第7章 評価	_____	48
第8章 まとめと今後の課題	_____	50
8.1. 本論文のまとめ	_____	50
8.2. 今後の課題	_____	50
参考文献	_____	53

目次

図 2 - 1	ユニキャストを使用したデータ配布	7
図 2 - 2	マルチキャストを使用したデータ配送	8
図 2 - 3	FEC を用いた信頼性確保	9
図 2 - 4	ACK Explosion と ACK Aggregation	10
図 4 - 1	広域ネットワークと MLS	16
図 4 - 2	MLS 内部データ配送と MLS 間データ配送	17
図 4 - 3	MLS 階層構造	18
図 4 - 4	MLS データ配送システム・モデル	19
図 5 - 1	データとセル	21
図 5 - 2	配布データとデータ広報パケット	23
図 5 - 3	双方向モードでのデータ送出とセル番号	24
図 5 - 4	単方向モードでのデータ送出のセル番号	25
図 5 - 5	データソースホスト状態遷移図	26
図 5 - 6	セルテーブル構成図	26
図 5 - 7	双方向モードでのデータ受信とセル番号	28
図 5 - 8	単方向モードでのデータ受信とセル番号	28
図 5 - 9	レシーバ状態遷移図	29
図 5 - 10	広報パケットのフォーマット	30
図 5 - 11	配布パケットのフォーマット	31
図 5 - 12	制御パケットのフォーマット	32
図 5 - 13	配布データ配送	34
図 6 - 1	アプリケーション構成図	39
図 6 - 2	メインウィンドウ画面	44
図 6 - 3	アプリケーション設定画面ダイアログ	45
図 6 - 4	上位 MLS 接続インターフェース設定ダイアログ	45
図 6 - 5	上位 MLS 接続インターフェース設定ダイアログ	46
図 6 - 6	データソース追加ダイアログ	46

表目次

表 5 - 1	データソースホストの動作モード	_____	23
表 5 - 2	データソースホストの状態と意味	_____	23
表 5 - 3	レシーバでの状態と意味	_____	26
表 5 - 4	セルテーブルの持つ状態	_____	27
表 5 - 5	パケットの種類と用途	_____	29
表 5 - 6	広報パケットの各フィールド	_____	30
表 5 - 7	配布パケットの各フィールド	_____	31
表 5 - 8	制御パケットの各フィールド	_____	32
表 6 - 1	アプリケーション動作の設定項目	_____	36
表 6 - 2	上位 MLS 接続インターフェースの設定項目	_____	37
表 6 - 3	下位 MLS 接続インターフェースの設定項目	_____	38
表 6 - 4	配送データ情報	_____	40
表 6 - 5	コマンドルーティングを行うパケットの種類	_____	42
表 7 - 1	各 Reliable Multicast と本研究の比較	_____	49

第 1 章

序論

1.1. 現在のインターネットを取りまく背景

近年インターネットは、我々にとって必要不可欠な情報インフラストラクチャとして確固たる地位を占めつつある。特にネットワーク・アプリケーションの発達とデータリンク技術の進歩には矚目すべきものがある。

過去においては、インターネットを構成する通信路は帯域が狭く、インターネット上で使用されるアプリケーションも、文字情報を主体としたもので情報量は少なかった。しかし、World Wide Web の登場やマルチメディア技術の発展、同アプリケーションの浸透に伴ってインターネット上で交換されるデータの種類・量は級数的に増大してきている [WTM]。ネットワーク・アプリケーションが取り扱う情報の種類は、文字情報の他に、静止画像や音声、また動画、そしてこれらを組み合わせたもの等、広い範囲に及んでいる。動画・音声情報は、文字情報とは比較にならないデータ量を持ち、また実時間性を要求するもの、再生時の品質保証を要求するものなど、多岐にわたる情報伝達上の品質に対する要求が発生している。

インターネット利用者数は増大の一途をたどっており、インターネットを構成するデータリンクに目を向ければ、通信路、殊に基幹となるネットワークの通信路の帯域は格段に広域化している。しかし、一般の利用者の観点からインターネットを見た場合、個々人がインターネットへ接続する場合の通信路は多種多様であり、利用できる帯域は様々である。アナログ電話回線、ISDN 回線といった数十 Kbps の帯域のものから、ケーブルテレビ網、デジタル衛星放送網などの Mbps クラスの帯域を持つものもある。

一方で、インターネットを情報伝達媒体として利用したい、という要求が高まってきている。従来の新聞・雑誌などの印刷媒体や、テレビ・ラジオなどの放送媒体に代わって、インターネットを情報伝達の媒体として用いる、というものである。これは、一般家庭などでも容易にインターネットへの接続性を得ること出来るようになり、利用者が多くなることが期待できること、またデジタルデータを取り扱う技術、アプリケーションの発達によって、従来の情報伝達媒体よりも情報伝達にかかるコストが低減できることが期待されているためである。

1.2. 現状での問題

1.1. で述べてきたインターネットの現状から、以下のような問題点が挙げられている。

多数の利用者に対して同一の情報を効率よく配布したい、という要求は高いにも拘わらず、現状のインターネットの持つ通信路を効率よく使用し、同時多数の利用者に対して情報の配布を行う仕組みが提供されていない。

インターネット上で多数の利用者に対して同時に同一の情報を配布する手段として、Reliable Multicast が研究されているが、既存のものは信頼性・規模性・実装の容易さなどの点で、現実的でないものがほとんどである。動画や音声を取り扱うアプリケーションのうち、リアルタイム中継を行うものはいくつも提案され実用化されている。しかし、アーカイブの配送を行うためのシステムを実用化したものは少ない。

現在 IETF などで行われている Reliable Multicast の研究では、多くの研究者が「すべての目的に合致する Reliable Multicast の枠組みは存在しない」と、万能の Reliable Multicast が不可能であることに合意している。そして、特定の目的に合致する枠組みを複数用意し、状況に応じてそれを使い分けるように推奨している。IETF には、複数の手法が提案されているが、アーカイブの配送に特化して実用的なものは、まだ存在しない。

1.3. 本研究の目的

本研究では、上述してきたインターネットに内在する問題、要求から、Segmented Hierarchical Reliable Multicast (SHRM:セグメント化階層構造信頼性マルチキャストデータ配送システム) と呼ぶ新しいデータ配送システムを設計、実装する。

第一に、情報の信頼性を保証する。信頼性のあるデータ配送を実現することによって、配布する情報の品質が保証され、またデータの再利用も容易になる。第二に、利用者が接続されているネットワークの通信路の持つ帯域を有効に使用する。これによって、利用者が広帯域通信路で接続されているにも拘わらず、低速でしか配信されない、という問題点を解決する。第三に、既存のインターネットに適合可能な規模性を持つ。これによって、広い範囲の利用者が、本研究で提案するシステムを用いて効率よいデータ配送を行うことができる。

1.4. 本論文の構成

第2章では、本論文で論ずるデータ配送システムに対する要求と、問題点について述べ

る。第3章では、既存のデータ配送システムである Reliable Multicast について述べる。第4章では、本論文で提案する解決手法について述べる。第5章では、本論文で提案するデータ配送システムである、SHRM の設計について述べる。第6章では SHRM の実装について述べ、第7章ではシステムの評価について述べる。最後に、第8章ではまとめと今後の課題について述べる。

第2章

信頼保証型データ配送システムについての考察

前章で述べたとおり、大量のデータを信頼性を保証しつつ配送するシステムに対する要求は非常に高い。本章では、インターネット上で多数の利用者に対してデータを配布する際の前提とすべき条件及び問題点について述べ、これを解決するためのデータ配送システムが持つべき要件について考察する。

2.1. データ配送システム実現のための前提

本節では、インターネット上でのデータ配送を行う際に、前提として考慮すべき条件について述べる。

2.1.1. 配送データの巨大化

マルチメディア技術に発達によって、コンピュータやネットワーク上で、様々な種類のデジタル化されたデータが取り扱えるようになった。この中で、近年特に発達がめざましいものに、音声・動画像のデータが挙げられる。これらの種類のデータに特徴的なのは、従来の文字・静止画といったものと比較して、飛躍的にデータの大きさが大きくなっている点である。

2.1.2. 利用者数

インターネットの利用者が増加していることは、先に述べたとおりである。また近年では、家庭用コンピュータの低価格化・高機能化が進み、従来高速な処理能力が必要とされていたマルチメディアコンテンツの再生も、一般家庭などで容易に実現できるようになってきた。

コンピュータの利用者がこうしたデータを入手するための媒体として、インターネットが注目されている。現在、デジタルデータの配布は、CD-ROMなどを頒布する形態が一般的である。しかしCD-ROMなどの記録済み媒体と異なり、ネットワークを経由したデータの入手方法は、利用者が欲しい時に手に入れられる、という点で非常に自由度が高い。

こうした現状から、インターネットを経由したデジタルデータの配布に対する要求は非常に高い。また、インターネットを介したデータ配布という形態を持つ利用者数は、既存の雑誌などのマス・コミュニケーション・メディアや、音楽CDの頒布といったものと同数、あるいはそれ以上と推定される。

2.1.3. インターネット上の通信路とデータ配送システム

インターネットを構成する通信路はさまざまであり、ネットワークのトポロジーも一様ではない。データを利用したい利用者は、ネットワーク上の様々な場所に存在し、データの配布元からのネットワーク上の経路は、利用者によって異なる。これは、データの配布元から利用者までをつなぐネットワークの帯域が一様でない、といえる。

配布元から利用者までのデータ配送のための通信路の品質が一様でないことは、後述するデータ配布の効率と規模性に大きく関わる。

2.2. データ配送における既存の枠組みの欠点

本節では、TCP/IP プロトコル体系で一般的に使用されている通信方式と、本論文で論議するデータ配送システムとの関係について考察する。

2.2.1. データ配送に対する要求の種類

インターネット上では、様々な種類のデータが取り扱われる。データの特徴によって、その配送に対する要求も異なるが、個々のパケットの配送に要する時間とデータ全体の信頼性に着目したとき、大きくは以下の2つに分類される。

1つは、実時間性を重視するデータ配送である。実時間性を重視したデータ配送では、データの配布元から配布先までの到達時間が成る可く短くなることが要求される。この種のデータ配送の典型例には、映像・音声のインターネットを介した中継や、インターネット電話などのアプリケーションが挙げられる。

インターネット電話を例にとり考える。電話での会話において、データの欠落はそれほど重大な問題ではない。ある瞬間の音声の欠落したとしても、その前後から内容が推定

できる。それに対してデータの遅延は、重大な問題である。会話者の一方が発話してから他方に届くまでに時間差があると、会話が困難になることは想像に難くない。

こうした点から、リアルタイム配送はデータの信頼性の保証よりもデータの到達時間の保証が重要である。

もう一方は、信頼性を重視するデータ配送である。信頼性を重視したデータ配送では、配布元のデータがエラーなく配送先に複写されていなければならない。データの配送のほとんどは、この種の配送であるといえる。コンピュータの実行イメージや、事務に供する書類など、データがたとえ一部分にしる誤って配送されれば、そのデータは利用価値を失う。

インターネット上でのデータ配送を考える場合、配送システムがどちらの性質を要求しているのかを考慮しなければならない。

2.2.2. ユニキャストとマルチキャスト

本節では、2.2.1. で述べた遅延と信頼性とは異なった視点でデータ配送を考える。インターネット上でのデータ配送では、送信者と受信者が1対1で対応するユニキャストによる通信方式と、送信者1に対して複数の送信者が存在するマルチキャストによる通信方式、ブロードキャストによる通信方式が存在する。

インターネットで使用されている TCP/IP プロトコル体系では、ユニキャストでの通信では、信頼性を保証するトランスポートである TCP と、信頼性を保証しない UDP が、広く実装され、使用されている。マルチキャストでの通信には、ユニキャストでの TCP のような信頼性を保証するトランスポート層は一般的に実装、使用されていない。

これら既存の、ユニキャスト/マルチキャスト通信、トランスポート層プロトコルと本論文で論ずるデータ配送システムの関係について考える。

本論文で対象とするデータ配送システムは、2.1.2. で述べたように、同一のデータに対して多数の利用者が存在することが前提となる。この前提の上で、ユニキャストによる通信によって信頼性を保証するデータ配送を実現することを考える。

このときには、データの配布元とデータの利用者が、それぞれ個別にデータ配送を行うことになる。この概念図を、図2-1に示す。

図からも分かるとおり、ユニキャストを利用してデータを配送すると、データの配布元に近いネットワークにトラフィックが集中してしまう。このトラフィックは、同一のデータを配送しているにも拘わらず、受信者の分だけ発生してしまうため、受信者数に対して規模性を欠く。このため、多数の受信者に対するデータ配送にユニキャストを用いることは現実的でない。

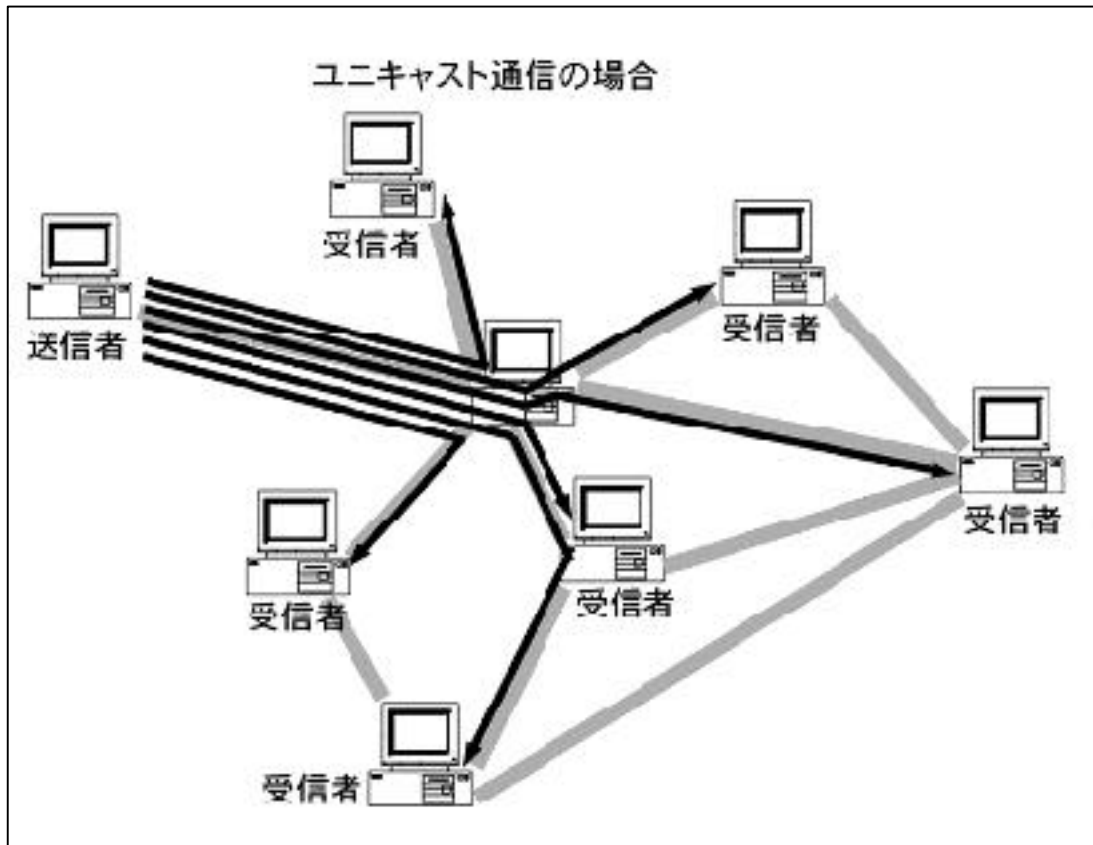


図 2 - 1 ユニキャストを使用したデータ配布

また、マルチキャストによる通信方法は、同一のデータグラムを重複なく、複数の受信者に対して送信できるため、ユニキャストで挙げたようなトラフィック集中の問題は発生しない。この概念図を、図 2 - 2 に示す。

しかし前述の通り、現在の TCP/IP プロトコル体系上では、一般的に実用に供されているマルチキャスト通信に対する信頼保証の仕組みが存在しない。

2 . 3 . データ配送システム実現上の問題点

本節では、本論文で論議するデータ配送システムを実現するときの問題となる点について考察する。

2 . 3 . 1 . 実時間性と信頼性との関係

2 . 2 . 1 . で、通信の種類を実時間性と信頼性に大別した。

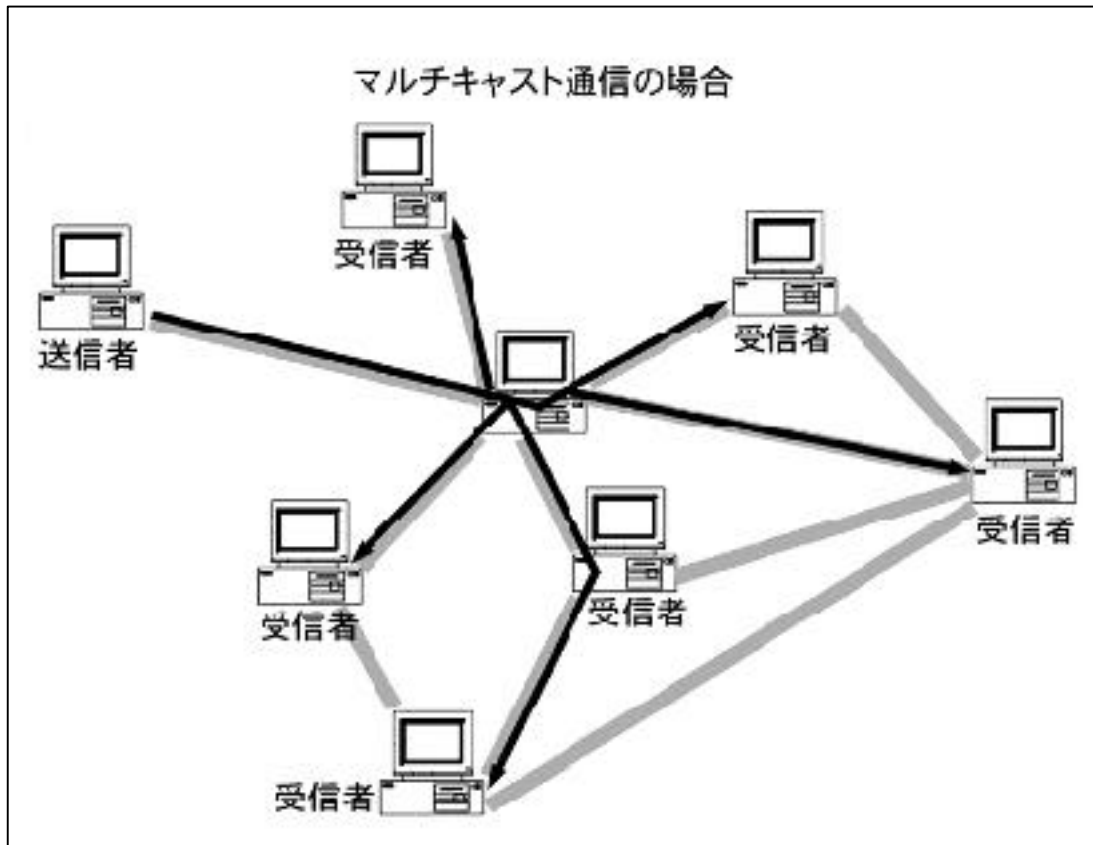


図2 - 2 マルチキャストを使用したデータ配送

実時間性と信頼性は、通信を実現する上で相反する性質である。信頼性を保証するためには、欠落やエラーの発生によって失われたデータグラムを、正しいデータへ訂正・回復する必要がある。これを行うためには、データグラムの欠落や誤りを検出し、正しいデータを保持しているホストやルータなどへ正しいデータを再び要求し、データグラムの再送を受ける必要がある。

データグラムの再送を行うと、当然再送にかかる時間がデータグラム送達の遅延となって立ち現れる。このため、実時間性を保とうとすれば信頼性を十分に保証することが出来なくなり、信頼性を保証しようとするれば実時間性を維持することが出来なくなる。

データ配送システムを考える上では、この二律背反を考慮に入れ、配送するデータの性質に応じて実時間性と信頼性のバランスを取る必要がある。

2.3.2. 信頼保証手法と性能

2.2.2. で述べたマルチキャスト通信を用いて、1対多での信頼性を保証する通信を行う場合、信頼性保証の手法には、現在大きく分けて3つの種類が提案されている。3つの方式には、それぞれ長所、短所があり、スループットや規模性、適応できるデータ配

送の性質が異なる。マルチキャスト通信を用いたデータ配送を行う場合、どのような手法で信頼性を保証するのかが問題になる。本節では、この信頼性を保証するための方式について考察する。

信頼性保証の方式の一つは、送信時にデータグラム中に前方誤り訂正符号(FEC: Forward Error Collection)を挿入し、受信時にFECに基づいて受信者が自力で誤りを検出、訂正するものである。この方式の概念図を、図2 - 3に示す。

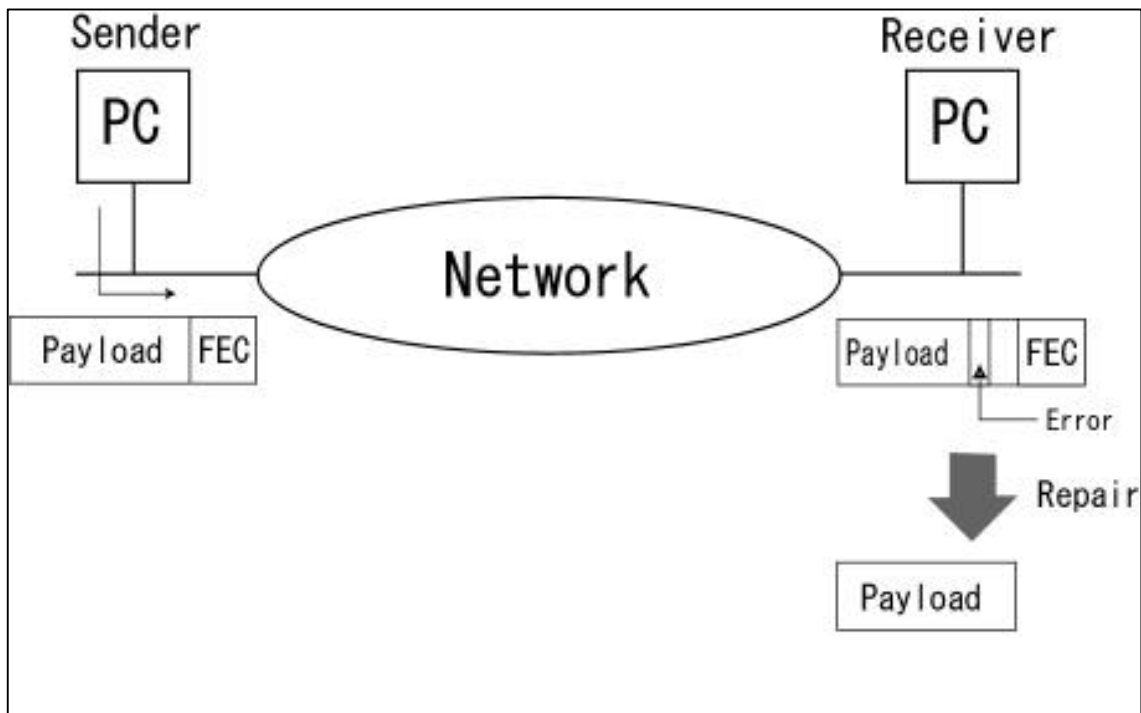


図2 - 3 FECを用いた信頼性確保

この方式では、データグラムの誤りを訂正するために再送を行わないため、再送機構によるスループットの低下や配送の遅延が発生しない、という利点がある。しかし、この方式は、データグラム中の誤りがFECで訂正できる範囲を超えていた場合や、データグラムそのものが配送されなかった場合に、受信側で正しいデータを再現できない。このため、この方式では完全に信頼性を保証することが不可能である。

信頼性保証の方式の二つ目は、送達確認を用いる手法である。この方式はユニキャストでのTCPと同様、送信者が送信したデータグラムに対して、受信者が受信した際に送信者に送達確認を送るものである。この方式では、多数の受信者がすべて送達確認を送信すると、送信者に対して送達確認が集中する、という問題(ACK Explosion)が発生するため、送達確認を集約する機構(ACK Aggregation)を併用するのが一般的である。この方式の概念図を、図2 - 4に示す。

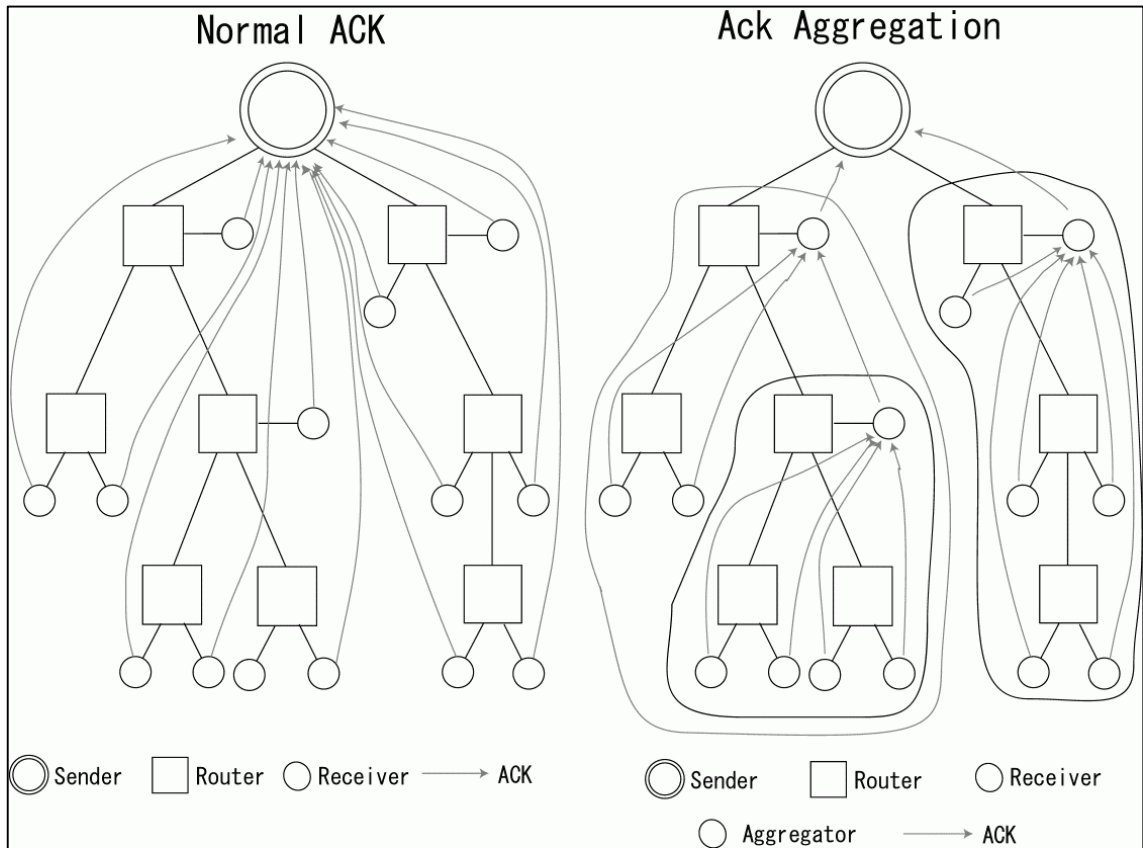


図2 - 4 ACK Explosion と ACK Aggregation

この方式では、受信者がデータグラムを確実に受信したことを確認できるため、データグラムの配信に完全な信頼性を保証できる。しかしこの方式では、発信者が多数の受信者の送達確認を管理する。このため、送達確認を行わないと次のデータグラムが送信できないので、受信に成功した受信者は、他の失敗した受信者が回復を行うまで待たされることになる。結果的に、データ配送システム全体としてのスループットが受信について能力の低い受信者に合わされることになり、効率的でない。また、個々の受信者の送達確認を個別に管理しては、送信者に対する負担が過大となるため、受信者の数に対する規模性を欠く。

2 . 4 . システム構築の方針

本節では、これまでに論議してきた事柄をふまえ、本論文で提案する配送システムが持つべき要件について述べる。

本システムは、データ配送について完全な信頼性を保証する必要がある。1 . 1 . で述べたように、本システムはアーカイブ配送に用いられることを前提としている。従って、

データの信頼性は必須のものである。

また本システムでは、受信者の接続する通信路の帯域に見合ったスループットを得られるような仕組みを提供する。3. で後述する既存の Reliable Multicast では、信頼性保証の機構のために、広帯域の通信路で接続された受信者が、十分に通信路の帯域を使用できない。本システムでは、この問題点を解決する。

信頼性保証・スループットの向上とともに、規模性も、配送システムの持つ重要な性能である。本システムは、現在のインターネットで成る可く広範に適用できるような、規模性を持つ必要がある。

第3章

関連研究

本章では、過去に提案されている Reliable Multicast の枠組みについて、本論文と関係の深いものを選び、それぞれの長所・短所について述べる。また、既存の信頼性マルチキャストについてのまとめと、信頼性マルチキャストの枠組みの開発について考慮すべき点について論ずる。

3.1. Chang and Maxemcunk と RMP

Reliable Multicast を使用したデータ配送の仕組みとして有名なものに、RMP (Reliable Multicast Protocol) [RMP]がある。RMP は、順序保証型の Reliable Multicast である。

RMP の動作の基本は、RMP 依然に提案されていた Chang and Maxemcunk プロトコル[Chang]を拡張したものである。Chang and Maxemcunk プロトコルは、最も初期に開発された Reliable Multicast プロトコルとして有名である。

このプロトコルの動作の概要は、以下の通りである。

このマルチキャスト配送に参加するすべてのホストは、論理的なリングを形成し、その内の一つのホストがトークン・サイトと呼ばれる、送達確認と再送について責任を持つホストとなる。リング中のソースがデータを送出すると、トークン・サイトはタイムスタンプを付加して送達確認をマルチキャストで送信する。またデータ受信に失敗したホストに対してユニキャストでデータを再送する。

トークン・サイトは、1つの送達確認毎にリングの次の参加者にトークンを渡す。

この方式の欠点は、以下に挙げる通りである。

参加者すべてが論理リングを構成し、またすべてがトークン・サイトになる。このため、ネットワークに接続された通信路の帯域が狭い参加者がトークン・サイトになった場合、そこがボトルネックとなり、全体のスループットが低下する。また、参加者に変更がある度に、論理リングを再構成しなければならない。このため、参加者の数に対する規模性が低い。

3 . 2 . RMTP:A Reliable Multicast Transport Protocol

RMTP(A Reliable Multicast Transport Protocol)[RMTP]は、送達確認を集約する方法で信頼保証型のマルチキャストデータ配送実現したものである。この方式で提案されている手法では、ネットワークの途上に Designated Receiver と呼ばれる、送達確認を集約する受信者を設置する。Designated Receiver は、マルチキャスト・ルーティングが木構造を持っている点を利用して、木構造の途上の結節点に位置し、そのサブツリー下から返送される送達確認を集約し、ツリーの上位に対して送達確認を代表する。これによって、2 . 3 . 1 . で述べた ACK Explosion の問題を解決している。また、Designated Receiver は受信データの一部をキャッシュし、サブツリー下からの再送要求に対して修復可能であれば、送信元に代わって修復パケットを送信できる。

この方式では、ネットワークを Designated Receiver 毎に分割して送達確認を管理している。この分割手法は、後述する本論文でのネットワーク分割手法に近い。この手法では、マルチキャスト受信者の参加者数に対する規模性が向上する。

しかし、この手法では、順序保証も行うため、低速でしか受信できないホストに影響を受けて、広帯域で受信できるホストでのスループットが十分に得られない。また、マルチキャスト・ルーティングを形成する木構造をベースにするため、これが原因でシステム全体としてのスループットは最も低帯域な通信路に影響される。

3 . 3 . SRM (Scalable Reliable Multicast)

SRM(Scalable Reliable Multicast)[SRM]は、上述の手法よりアプリケーションを考慮した Reliable Multicast Protocol である。この手法の特徴は、データの取り扱い単位を、アプリケーション毎に意味のあるデータのひとかたまりである ADU (Application Data Unit)としたことである。この手法は、多人数参加型のホワイトボード・アプリケーションのためのプロトコルとして設計された。従来の提案されている手法はすべてパケット単位であるのとは異なり、この手法では、ホワイトボードのアプリケーションについて意味のあるデータ(「線を引く」「塗りつぶす」と言ったオペレーション単位)を取り扱うデータの単位としている。パケットに対してセッション中に一意な番号であるシーケンス番号を付与するように、本手法では ADU に対して一意な名前を付加(Named ADU)する。データの単位に ADU を用いることによって、データの再送時に必ずしもデータの配布元から取得する必要がなくなり、欠落データ修復の自由度が増す。また Named ADU はアプリケーションにとって意味のある単位なので、順序の保証はアプリケーション側で行うことができ、システムの複雑さを低減できる。

しかし、この手法では、適応するアプリケーション毎に ADU を設計する必要があり、

アプリケーションに対する自由度が低い。従って、一般的なトランスポートの仕組みとして適用するのは難しい。

3 . 4 . 既存の **Reliable Multicast** のまとめ

Reliable Multicast の構築については、いくつかの提案がなされているが、ある一つの方法がすべての状況や、アプリケーションの要求に対して満足のできる性能を持つことは不可能である。IETF で合意された信頼性マルチキャスト・トランスポート及びアプリケーションのプロトコルの評価基準[RFC2537]でも述べられているように、信頼性マルチキャストを1つのプロトコルによってすべての要求に対して満足させることは不可能である。このため、信頼性マルチキャストの枠組みは、アプリケーションやネットワークの要求をよく考慮し、実現する範囲を限定し、その中で性能を十全に発揮できるものが開発されなければならない。信頼性マルチキャストを使用する際は、システムに対する要求を明らかにした上で、その要求に最も適する枠組みを選択して使用することが望まれる。

本章で取り上げた信頼性マルチキャスト通信の枠組みは、より一般的な枠組みを提供しようとした RMP や RMTP では、すべての要求に対して十分に応えることは出来ない。また、アプリケーションを考慮して提案された SRM では、前章までで論じてきた広域なアーカイブの配送に用いるには、欠点があり充分ではない。

本論文では、これら既存の提案をふまえた上で、それぞれの提案から良い点を抽出し、新たな手法を加味した、新しいデータ配送の仕組みを提案する。

第4章

解決手法の提案

本章では、前述してきた問題点を解決する新しいデータ配送システムである、Segmented Hierarchical Reliable Multicast(SHRM:セグメント化階層構造信頼性マルチキャストデータ配送システム)を提案する。

4.1.MLS 手法の導入

本節では、信頼性マルチキャストデータ配送に対して、新しい考え方である MLS の導入を提案する。MLS の定義と、ネットワーク上での MLS の実現方法について述べる。

4.1.1.MLS の定義

既存のマルチキャストを利用した信頼性保証型データ配送システムでは、それを広域なネットワーク上に適用した際の規模性が常に問題となる。この点を解決するため、本システムではマルチキャスト可能なローカルセグメント(MLS: Multicastable Local Segment)という考え方を導入する。MLS は、均等な品質でマルチキャスト・データグラムを配送可能な、限定的な範囲のネットワークとする。ネットワークを MLS 単位に分割することにより、広域ネットワークは、複数の MLS が相互に接続されたネットワークと見なすことができる。

広域ネットワークと MLS の関係を、図4 - 1に示す。

4.1.2.MLS の設定

MLS は、マルチキャスト・データグラムを配送可能なネットワークであれば、どのような範囲のものにも設定可能である。

しかし、前節で述べた品質は、データリンクの帯域にとどまらず、エラー発生割合、配送遅延も含む。この条件を満たすためには、実際には、MLS は同一のデータリンクによ

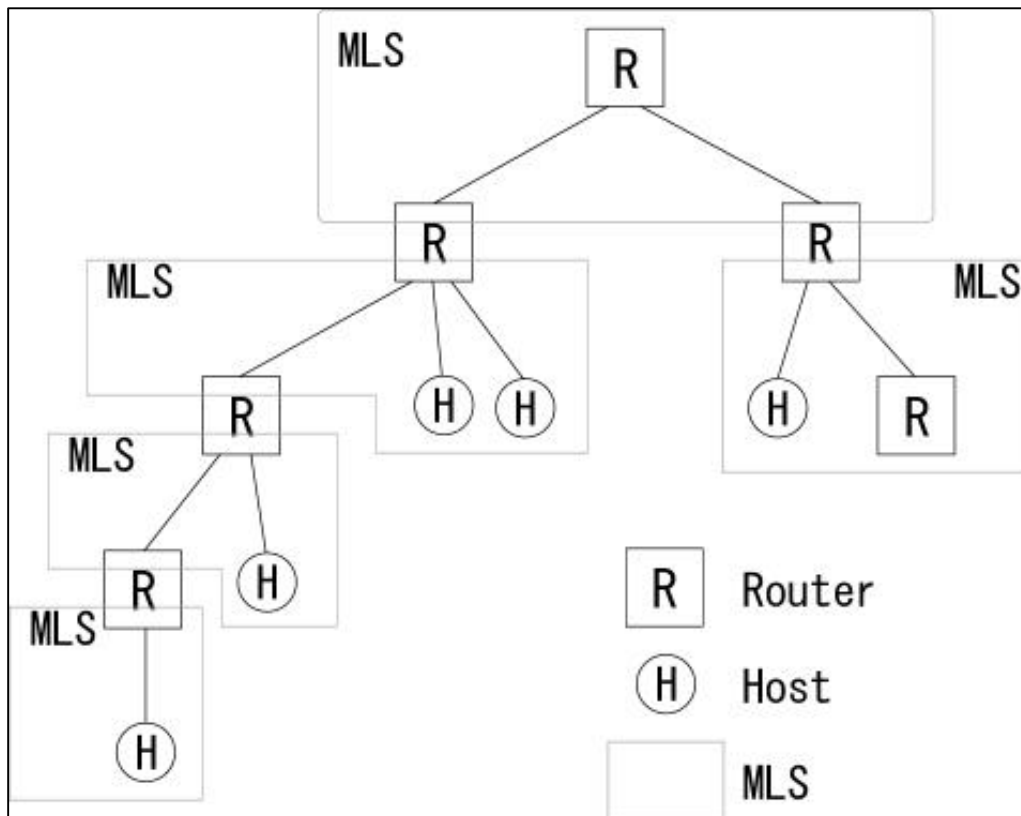


図 4 - 1 広域ネットワークとMLS

って区切られたネットワークとすることが現実的である。

現在のインターネットでは、インターネット上のすべてのルータがマルチキャスト経路制御を行えるとは限らない。現実的には、マルチキャスト経路制御を行わないルータの方が多い。このため、広域ネットワーク上でのマルチキャスト通信は、トンネリング技術を用いて実現されるのが一般的である。トンネリングは、配送すべき階層のデータグラムを、異なるデータリンク層またはそれより上位のプロトコルによってカプセル化し、配送するものである。このため、トンネリングを使用したデータグラムの配送では、データ配送の品質を推定することが難しい。よって MLS は、トンネリングを行う経路を含まないマルチキャスト配送の範囲とすることが望ましい。

4 . 2 . MLS 分割による範囲限定型信頼保証

MLS 単位に分割したネットワークにおいて、個々の MLS に注目する。

個々の MLS は、概ね均等な品質をもってマルチキャスト・データグラムの配送を行える点に着目し、1 つの MLS 単位の内部で完結されたデータ配送の信頼保証システムを構

築する。

ここで重要な点は、信頼保証を行う仕組みは個々の MLS 内部で完結し、特定の MLS の外部とは無関係に独立のものであるという点である。この信頼保証のシステムでは、従来の信頼保証のシステムに見られるデータリンクの帯域、遅延の有無といった品質の多様性を考慮する必要がない。従って、より簡易な仕組みを用いて信頼性を提供することが可能となる。

4.3. MLS 手法を利用したデータ配送

4.3.1. MLS 間でのデータの配送

個々の MLS 内でのデータ配送を行うだけでは、広域なネットワーク全体に跨ったデータ配送を行うことができない。広域ネットワーク全体の上で、データ配送を行うため、上述した MLS 同士でデータを配送する仕組みを提供する。

ここで重要な点は、MLS 間でデータを配送するシステムと、MLS 内部でデータを配送するシステムとは相互に独立であるという点である。MLS 間のデータ配送は、4.1.2. で述べた MLS 同士の接点となっている部分のみが処理を受け持つ。MLS の接点は、MLS 内部のデータ配送システムに対しては MLS 内部に存在する通常のホストと同様の振る舞いをする。

MLS 内部データ配送と、MLS 間データ配送の関係を、図 4 - 2 に示す。

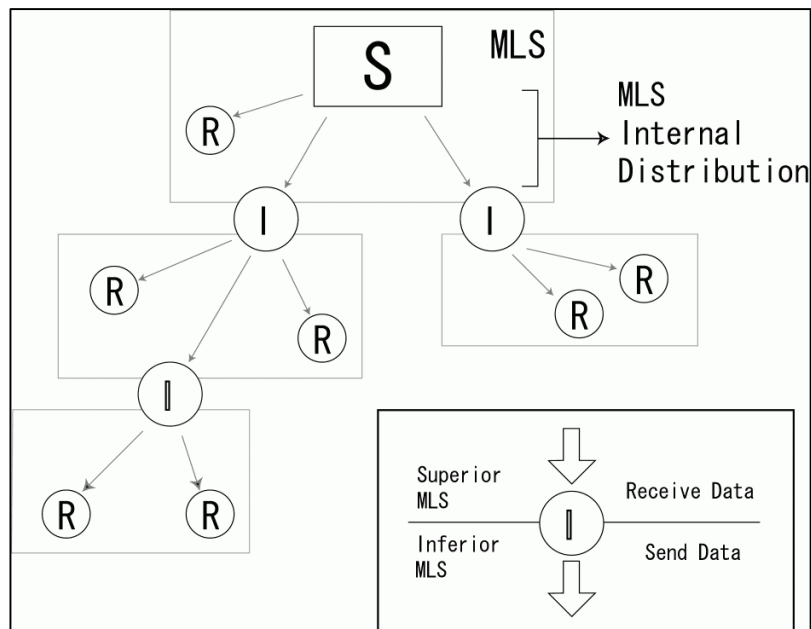


図 4 - 2 MLS 内部データ配送とMLS 間データ配送

4.3.2. MLS 階層構造の形成

ネットワークを MLS に分割した時、ネットワークは MLS の集合と見ることが出来る。本システムでは、この MLS の集合を階層構造として捉える。

MLS 階層構造は、原則として、最も帯域の広い MLS を最上位 MLS とし、その下に、順次帯域の狭い MLS が接続される構造を持つものとする。MLS 階層構造の概念図を、図 4 - 3 に示す。

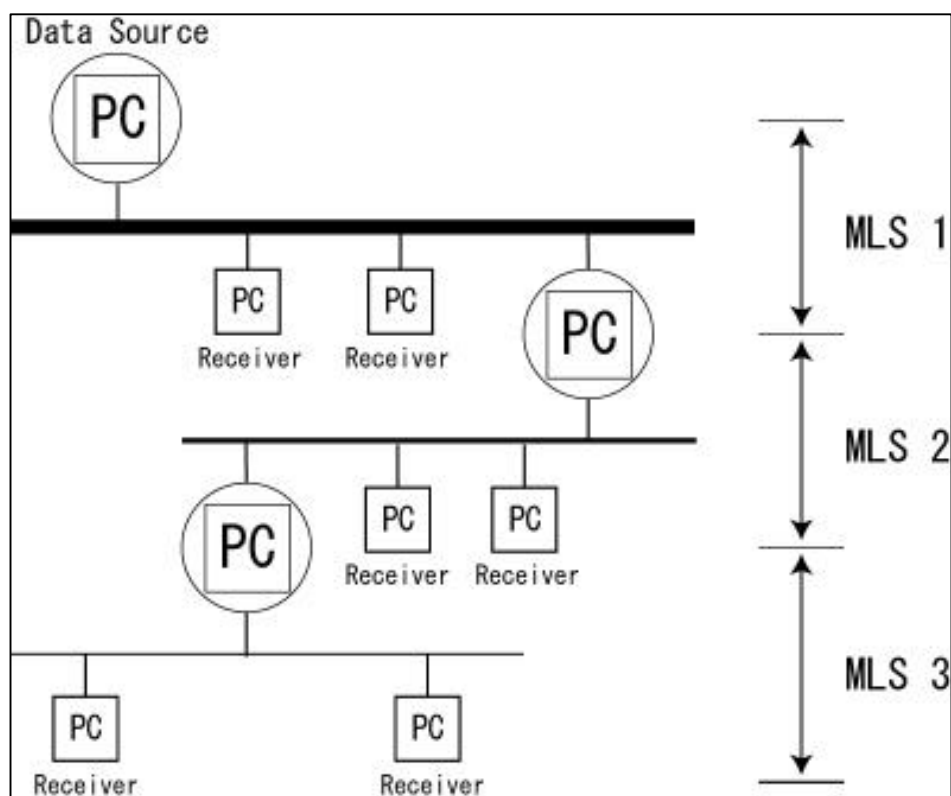


図 4 - 3 MLS 階層構造

MLS 階層構造を形成する必要性については、4.3.3.において配布データの配置場所についての問題とともに述べる。

4.3.3. 配布データ配置

配布データは、必ず 4.3.2. で述べた MLS 階層構造の最上位 MLS に置くものとする。

これは、配布データが必ず特定のMLSにあるとすれば、任意のMLSにおいてデータは必ず配布データが置かれているMLSに近い側のMLS接点から流入してくることになる。これにより、配布データの流通する方向は必ず一方向となるので、配送システムが配布データの所在位置を記憶しておく必要がなくなり、システムの複雑さを低下できる。

MLSの集合が4・3・2・で述べたような階層構造を形成しているとするなら、この中では配布データは最上位MLSに配置するのが最も効率が良い。

4.4. MLS 階層構造モデル

以上述べた、MLSを用いたデータ配送システムのモデルを、図4-4に示す。

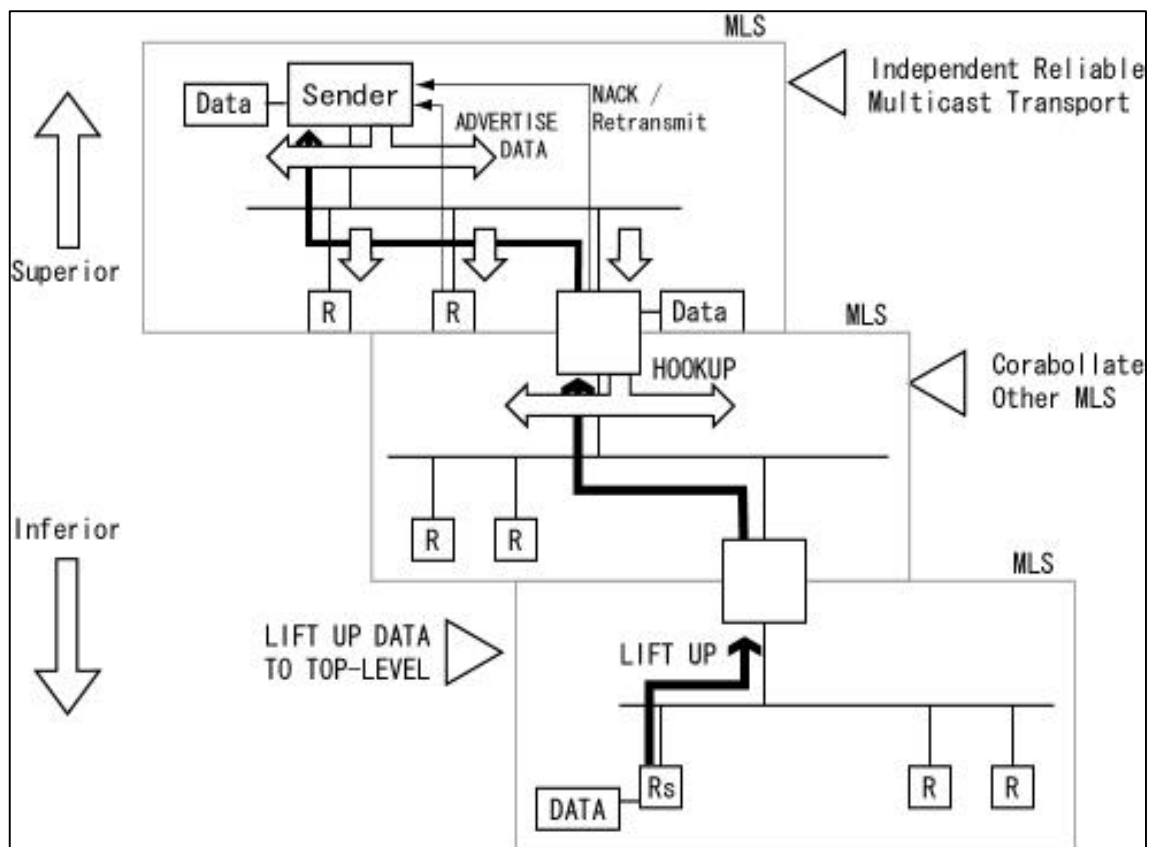


図4-4 MLS データ配送システム・モデル

第5章

設計

本章では、本論文で提案する SHRM の設計について述べる。

本システムは、MLS 内データ配送、MLS 間データ配送、配布データ配送に分かれる。

5.1 システム設計の前提

本システムは、以下の動作が行える環境を前提とする。

- ・IP マルチキャスト・データグラムが配送可能なネットワークであること

本システムは、MLS 内部でのデータ配送に、IP マルチキャストを使用する。従って、ネットワークは IP マルチキャストをサポートしている必要がある。

- ・マルチキャスト、ユニキャストで、UDP トランスポート層が使用可能であること

本システムでは、個々のパケットの配送に UDP を用いる。従って、ネットワークは UDP を配送可能な必要がある。

5.2 データ構造

5.2.1 配送単位

システムは配送を行う際、配送データを小さいサイズに分割して送信する。この小片化されたデータを、セルと呼ぶ。セルはヘッダとペイロードからなる。ペイロードのサイズは MLS 内部での MTU や転送効率を考慮し、自由に設定できる。セルには、配布データの先頭部分を格納しているセルから順次番号が付される。これをセル番号と呼ぶ。システムは、データの配送、信頼性保証の仕組みの動作の際に、セルを最小単位として取り扱う。これを、図5 - 1に示す。

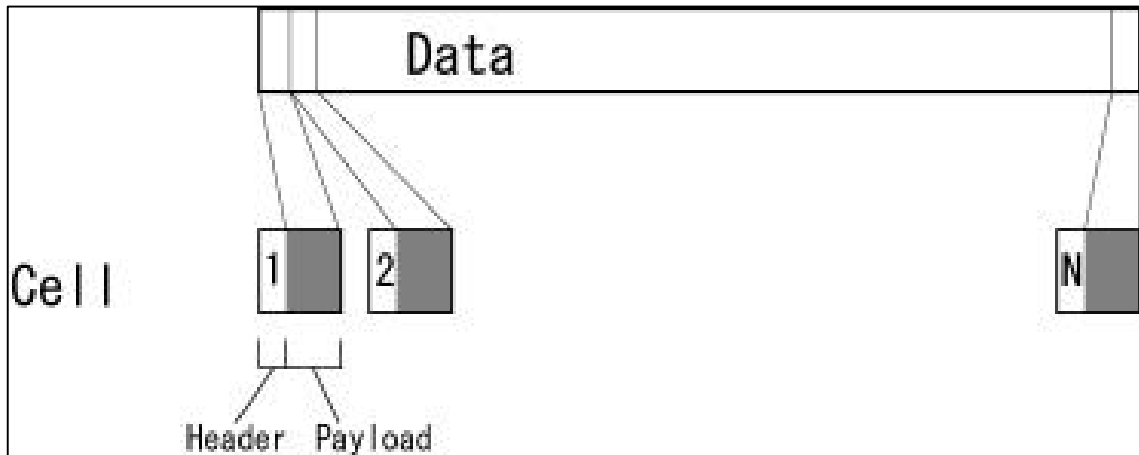


図5 - 1 データとセル

5 . 2 . 2 . 配布データの有効期限

配布データには、そのデータの有効期限が設定される。

有効期限を過ぎたデータは、配送しない。有効期限は、各 MLS 内で独自に設定し、MLS 内部でのみ有効であるものとする。

5 . 3 . MLS 内データ配送

MLS 内データ配送は、データ配送と、配送時の信頼性の保証、配送データに関わる情報の提供を行う。

5 . 3 . 1 . MLS の設定

4 . 1 . で述べたように、MLS は均等な品質でマルチキャスト・データグラムを配送可能な範囲のネットワークである。これを実現するため、特定の MLS の範囲を決められなければならない。

既存のマルチキャスト通信では、マルチキャスト・データグラムの配送範囲を限定するために、TTL を利用している。本システムでも、同様に TTL を利用する。

5.3.2. データ配送

本システムは、UDP を使用して実現される。

MLS 内データ配送では、ホストは配布元として配布データを保持するデータソースホストと、データを受信するレシーバに別れる。

データ配送は、配布すべきデータの広報と、配布データの配送に分類される。広報と配送は、独立したデータチャネル（UDP ポート番号）を使用して送信される。

データ広報は、MLS 内に独立の、広報に用いるポート番号を定める。このポート番号は永続的であり、データソースホスト、レシーバが共に共通に認知していなければならない。

またデータ配送は、動的に任意のポート番号を定める。データ配送に用いるポート番号は、データ広報の中に記述される。データ配送に用いるポート番号は、配送データ毎に独立であり、永続的ではない。

データソースホストは、MLS 内に定期的にデータ広報を行う。また、レシーバからの要求に従って、データ配送を行う。レシーバはデータを受信する際に、まずデータ広報を受信し、受信すべきデータの諸元を得た後、そのデータを受信する。

データソースホストにおけるデータ広報、データ配布の詳細、並びにレシーバにおけるデータ受信の詳細は、以降の節で述べる。

5.3.3. データ広報

データソースホストでは、配布すべきデータの情報をレシーバに対して通知する。このとき、配布すべきデータの情報は、予め与えられているものである。配布すべきデータの情報については後述する。

データ広報は、広報パケットを送出することによって行う。配布データと広報パケットの関係を図 5 - 2 に示す。

データソースホストは、配布すべきデータに対応する数のデータ広報パケットを生成し、定期的に MLS 内に対して送付する。

5.3.4. データソースホストにおけるデータ配送

データソースホストでは、レシーバからの要求に応じて、配布すべきデータを送付する。

データソースホストでは、2つの送付モードを持つ。各モードの名称と意味を、表 5 - 1 に示す。

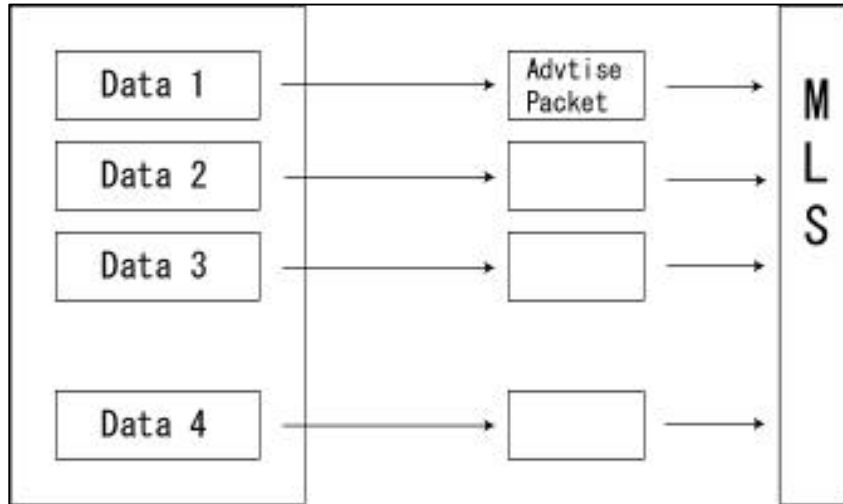


図 5 - 2 配布データとデータ広報パッケージ

動作モード	動作の種類
双方向モード(Bi-Directional)	データソースホストとレシーバが相互に通信可能な状態での動作
単方向モード(Uni-Directional)	データソースホストからレシーバへ向けての単一方のみ通信可能な状態での動作

表 5 - 1 データソースホストの動作モード

状態の種類	状態
初期状態(INITIALIZED)	レシーバが存在しない状態
常時送信(ALWAYSSEND)	常時データを送信する状態
レシーバ残数 1 時の送出(1RCV)	レシーバが 1 つだけ存在する状態
レシーバ残数 2 以上時の送出(2RCV)	レシーバが 2 つ以上存在する状態

表 5 - 2 データソースホストの状態と意味

データソースホストでは、4つの状態を持つ。各々の状態の名称と意味を、表 5 - 2 に示す。

5.3.1. で述べたように、データはセルに分割され、各々のセルには 0 から始まる順番番号が付与される。また、データソースホストでは、送出すべきデータに対応するレシーバ数を保持するカウンタ（レシーバカウンタ）を持つ。セルは、UDP を使用して配布される。このときの UDP のポート番号は、配布データ毎に動的に設定される。

双方向モードでのデータ配送

双方向モードでは、初期状態ではレシーバが1つも存在しないので、配布データの送出行わさない。レシーバから配布データの送出行要求を受信した場合、レシーバカウンタを1加算し、番号0のセルから送出行を始め、以降、最後のセルまで順次送出行を行う。

レシーバからの送出行要求がある度にレシーバカウンタを1加算する。すでにデータの送出行を行っている状態でレシーバからの要求を受けた場合には、そのときに送出行していたセルの番号を、最後に要求を受けた際に送出行を開始したセル番号（最終要求セル番号）として記憶する。

配布データを最後まで送出行した場合、レシーバカウンタを1減算する。レシーバカウンタが1より大きい場合は、再びセル番号0からデータの送出行を行う。レシーバカウンタが1の場合、最終要求セル番号の位置までデータを送出行し、レシーバカウンタを0に減じてデータの送出行を終了する。

データの送出行とデータソースホスト、セル番号の関係を、図5-3に示す。

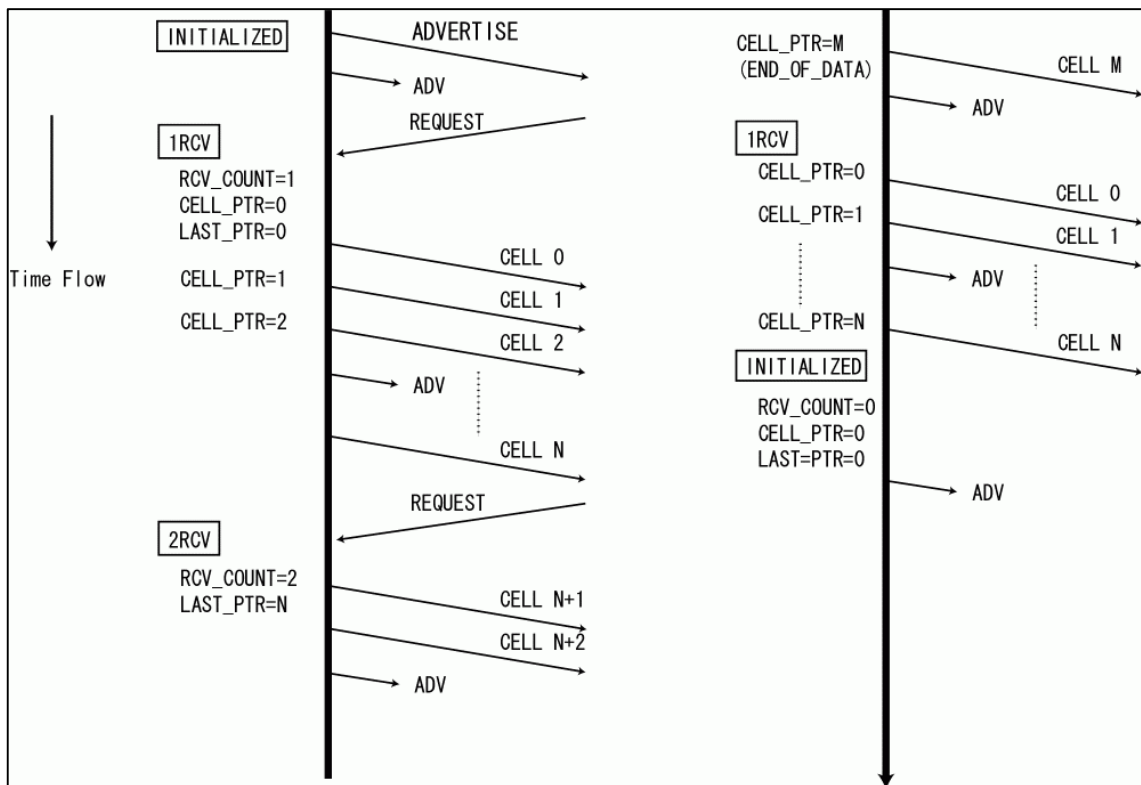


図5-3 双方向モードでのデータ送出行とセル番号

単方向モードでのデータ配送

単方向モードでは、レシーバが存在するか否かに関わりなく、常時配布データを送出する。

データは、セル番号0のセルから順次送出される。データの最後まで送出した場合、再びセル番号0に戻ってデータを送出する。

データの送出とレシーバ、セル番号の関係を、図5 - 4に示す。

データソースホストでの状態の遷移を、図5 - 5に示す。

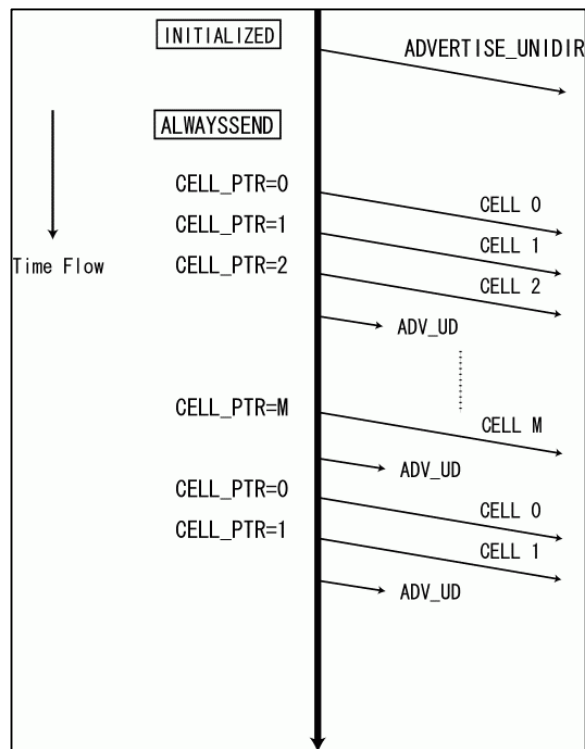


図5 - 4 単方向モードでのデータ送出とセル番号

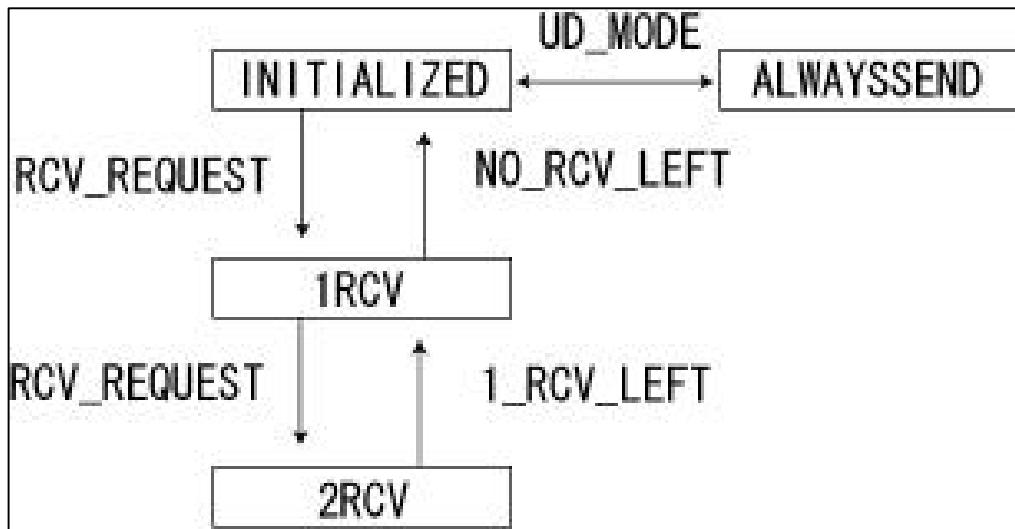


図 5 - 5 データソースホスト状態遷移図

5 . 3 . 5 . レシーバにおけるデータ受信

レシーバでは、4つの状態を持つ。各々の状態の名称と意味を、表 5 - 3 に示す。

状態の種類	状態
初期状態(INITIALIZED)	受信していない状態
巡回受信(CYCLE-RCV)	データを一方的に受信する
補正受信(CORRECT-RCV)	双方向モード時の欠落データ補完
巡回補正待ち受信(CYCLE-CORRECT-RCV)	単方向モード時の欠落データ補完

表 5 - 3 レシーバでの状態と意味

レシーバでは、受信すべき配布データの全セルの状態を保持する、セルテーブルを持つ。セルテーブルの構成を図 5 - 6 に示す。セルテーブルの持つセルの状態を、表 5 - 4 に示す。

CELL #	0	1	2	...	n-2	n-1	n
STATUS							

図 5 - 6 セルテーブル構成図

状態の種類	セルの状態
未受信(NOTRCVED)	受信していない
受信済(RCVED)	受信済み
欠落(DROPPED)	受信失敗

表5 - 4 セルテーブルの持つ状態

レシーバでは、まずデータ広報を受信し、データを受信するための諸元を得て、データ広報によって受信すべきデータを決定する。

双方向モード

データソースホストが双方向モードで動作している場合、データソースホストに対してデータの送出手を要求する。このとき、配布データの送出手がすでに行われていても、要求を送出手する。

レシーバは、まず巡回受信状態データソースホストの送出手するセルを一方向的に受信する。セルを受信し、セル番号に従って配布データを生成する。これを、セル番号が一巡するまで続ける。巡回受信状態データセルの受信に失敗した場合は、そのセルは欠落としてセルテーブルに記録する。

セルの受信が一巡したら、巡回受信状態を終了し、補正受信状態となる。セルテーブルを参照し、この中で欠落(DROPPED)としてマークされているセルの送出手を、データソースホストに対して要求する。

最終的に、すべてのセルを受信したら、受信を終了する。

このときのレシーバの各状態とセルテーブル、データの関係を図5 - 7に示す。

単方向モード

データソースホストが単方向モードで動作している場合、レシーバはまず巡回受信状態データソースホストの送出手するセルを一方向的に受信する。このとき、双方向モードとは異なり、データの送出手要求をデータソースホストに対して要求しない。

セルの受信が一巡したら、巡回受信状態を終了し、巡回補正待ち受信状態となる。セルテーブルを参照し、この中で欠落としてマークされているセルのみを受信する。セルテーブル中の欠落としてマークされているセルがなくなるまで続け、すべてのセルを受信したら終了する。

このときのレシーバの各状態とセルテーブル、データの関係を図5 - 8に示す。

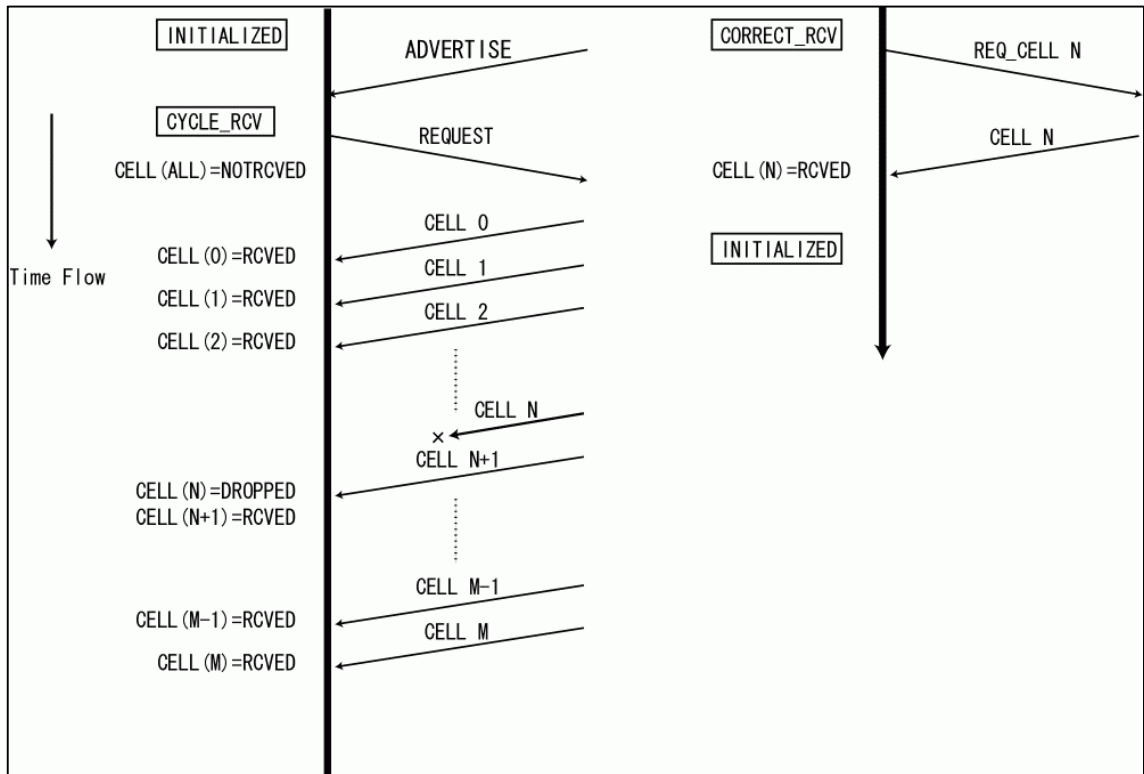


図5 - 7 双方向モードでのデータ受信とセル番号

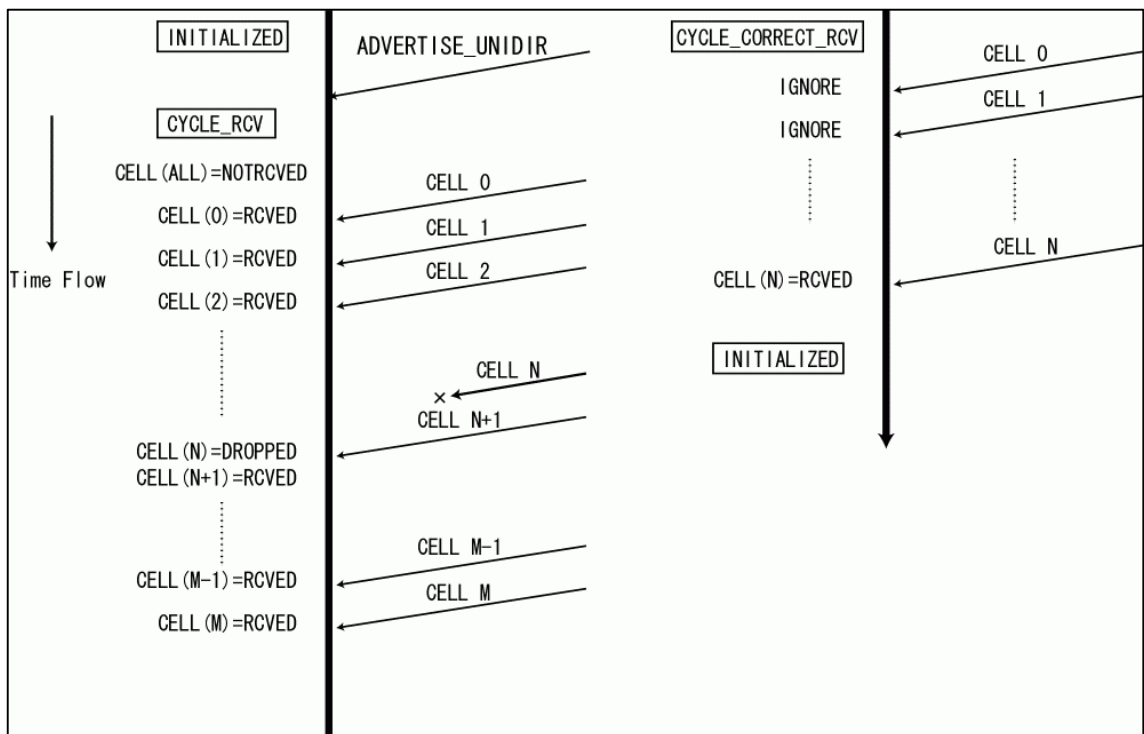


図5 - 8 単方向モードでのデータ受信とセル番号

レシーバの状態遷移図を図5 - 9 に示す。

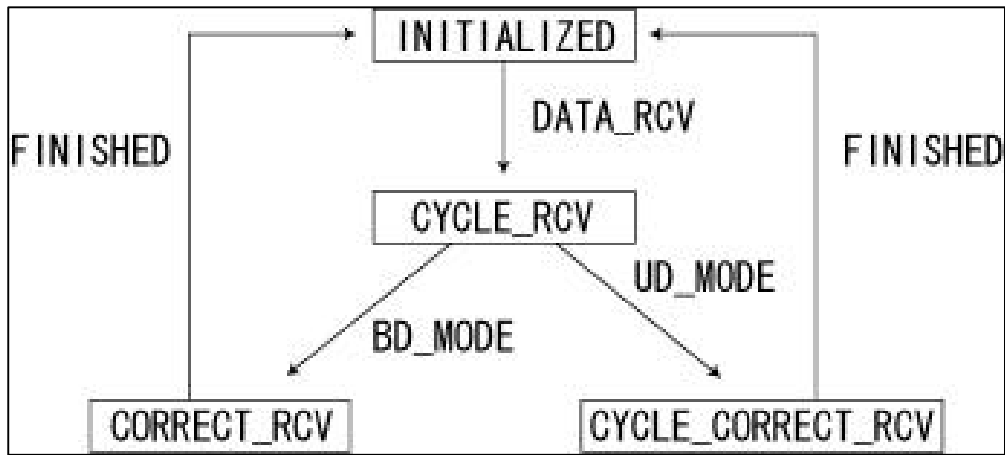


図5 - 9 レシーバ状態遷移図

5 . 3 . 6 . パケットのフォーマット

本節では、SHRM の用いるパケットの種類とそのフォーマットを示す。
 本システムでは、3つの種類のパケットを使用する。これを表5 - 5 に示す。

パケットの種類	使用用途	送受信の方向
広報パケット	データ広報	Data Source Receiver
配布パケット	セル配布	Data Source Receiver
制御パケット	データの要求	Receiver Data Source

表5 - 5 パケットの種類と用途

広報パケットは、データソースホストが配布データの情報を広報するために用いる。広報パケットには、表5 - 6 に挙げる情報を含む。広報パケットのフォーマットを、図5 - 10 に示す。

octet		
Version	TYPE	Packet length
Segment length		Control CH Port#
ADV Total		RSVD
ADV #		Data CH Port#
Data Archive Length		
Data ID Length	Data ID String	
Data ID String (var_len)		

図5 - 10 広報パケットのフォーマット

Field	Length(octet)	内容
Version	1	SHRM のバージョン (=0x01)
TYPE	1	パケット種別 0x01: ADVERTISE 0x02: ADVERTISE_UNIDIR
Packet length	2	パケット長
Cell Payload Length	2	セルのペイロード長
Control CH Port	2	制御ポート番号
ADV Total	3	広報中のデータ総数
RSVD	2	予約 (=0x0000)
ADV #	2	広報番号
Data CH Port	2	データ配布ポート番号
Data Archive Length	4	配布データのサイズ
Data ID Length	2	配布データ ID 文字列長
Data ID String	Variable	配布データ ID 文字列

表5 - 6 広報パケットの各フィールド

配布パケットは、データソースホストが各セルを送出するために用いる。配布パケットには、表5 - 7に挙げる情報を含む。配布パケットのフォーマットを、図5 - 11に示す。

Version	TYPE	Packet length
Cell Number		
Payload (var_len)		

図5 - 1 1 配布パケットのフォーマット

Field	Length(octet)	内容
Version	1	SHRM のバージョン (=0x01)
TYPE	1	パケット種別 0x03: DATA_START 0x04: DATA_CONT 0x05: DATA_EOT 0x06: DATA_SPECIFIED
Packet Length	2	パケット長
Cell Number	4	セル番号
Payload	variable	データ本体

表5 - 7 配布パケットの各フィールド

ここで、セルの持つペイロードのサイズについて考える。

配布データをセルに分割した際、個々のセルにはヘッダが含まれるので、配送すべきデータの総量は実際の配布データより増加する。従って、ペイロード長が大きいほど効率が高くなる。しかし、ネットワークの通信路はその種類によって最大 MTU が決まっており、それを越えた長さのデータグラムを取り扱うことは出来ない。

従って、セルのペイロード長は、以下の値を持つことが現実的であると考えられる。

ペイロード長 = MLS 内部の最大 MTU - IP ヘッダ長 - UDP ヘッダ長 - セルヘッダ長

制御パケットは、レシーバが欠落セルの送出手をデータソースホストに対して要求するために用いる。制御パケットには、表5 - 8 に挙げる情報を含む。広報パケットのフォーマットを、図5 - 1 2 に示す。

Version	TYPE	Packet length
Data ID Length		Data ID String
Data ID String (var_len)		
Cell Number (if need)		

図5 - 1 2 制御パケットのフォーマット

Field	Length(octet)	内容
Version	1	SHRM のバージョン (=0x01)
TYPE	1	パケット種別 0x07: REQ_DATA_SEND 0x08: REQ_SPECIFIED_CELL (0x09: ACK_RECEIVE)
Packet Length	2	パケット長
Data ID Length	2	対象となる配布データの Data ID 文字列長
Data ID String	Variable	対象となる配布データの Data ID 文字列
Cell Number	4	対象となるセルの番号

表5 - 8 制御パケットの各フィールド

5 . 4 . MLS 間データ配送

MLS 間データ配送は、上位 MLS と下位 MLS 間で、データを配送する。

5 . 4 . 1 . MLS の接続

上位 MLS と下位 MLS の接続は、MLS 間接続ホスト（中継ホスト）がこれを行う。

MLS 間接続ホストは、上位 MLS と下位 MLS とに同時に接続し、上位 MLS に対しては、MLS 内データ配送におけるレシーバとして振る舞う。下位 MLS に対してはデータソースホストとして振る舞う。

MLS 間接続ホストは、上位 MLS でのデータ広報、データ配送を下位 MLS に中継する。

5 . 4 . 2 .MLS のデータ中継

MLS 間接続ホストは、上位 MLS から受信したすべてのデータ広報を、下位 MLS に中継する。

データ広報を受信した MLS 間接続ホストは、広報に含まれる情報のうち、下位 MLS での MLS 内データ配送に必要な部分を適切なものに変換する。ポート番号、データ ID、セルのペイロード長の違いによるセル総数の違いなどがこれに当たる。

下位 MLS からデータ送出要求を受けた場合、これに該当する配布データを上位 MLS から受信する。上位 MLS から受信した配布データは MLS 間接続ホストに蓄積され、下位 MLS に対する配布データとして準備される。配布データは、下位 MLS での MLS 内データ配送に適するようにセル長などを変換され、下位 MLS に送出される。

また、上位 MLS から受信されたデータは、原則として MLS 間接続ホスト内で保持され、再度の送信要求に対して使用される。

5 . 5 . 配布データ移送

本節では、4 . 3 . 3 . で述べた、配布データを最上位 MLS へ配置する手法について述べる。

MLS 階層構造中の最上位 MLS でない MLS に属するホストが、データを配布する場合、このデータは、一旦最上位 MLS まで運ばれてから配布されなければならない。このため本システムでは、最上位 MLS への配布データ移送を行う。

4 . 3 . 2 . で述べたように、MLS は必ず一つの MLS 間接続ホストによって上位 MLS に接続されている。このときの MLS 間接続ホストは、その MLS 内でのデータソースホストである。

従って、ネットワーク上の任意の MLS から最上位 MLS へデータを配送するために、以下の機構を提供する。

配送データを持つホストは、自分の属する MLS で、配送データをデータソースホストに送信する。配送データを受け取ったデータソースホストは、自分が所属する MLS に上位 MLS に接続する MLS 間接続ホストが存在すれば、それに対して配送データを送信する。

これによって、ネットワーク中の任意のホストから最上位 MLS のデータソースホストに対して、配布データを配送することができる。この動作を、図 5 - 13 に示す。

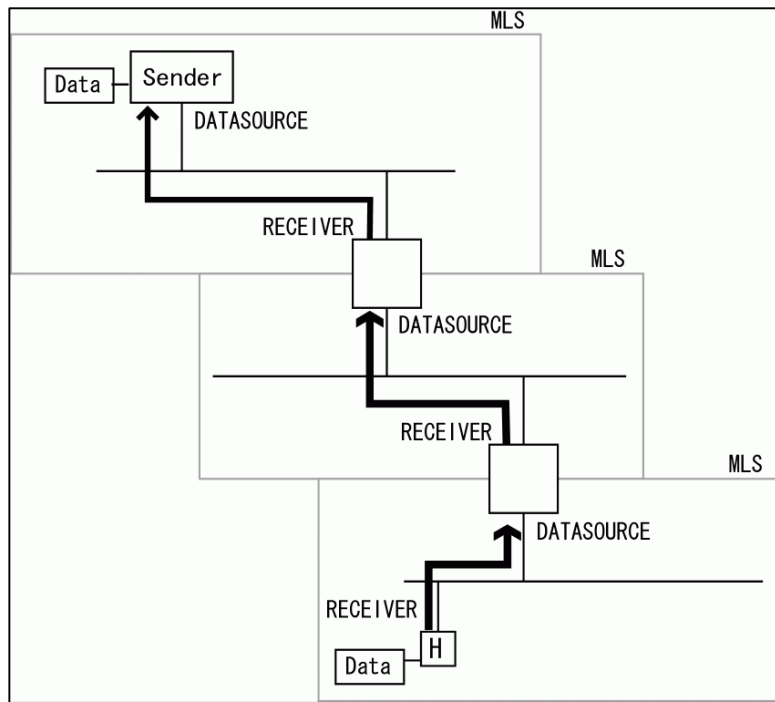


図 5 - 1 3 配布データ配送

第6章

実装

本研究では、本論文で提案する SHRM を実装した。本章では、SHRM の実装について述べる。

6.1. 実装環境

SHRM の実装は、Microsoft Windows98 オペレーティングシステム（以降 Windows98 と表記する）上で行った。プログラム言語には、Microsoft Visual C++ 6.0 を使用した。

実装は、Windows98 上のアプリケーションとして動作する。

6.2. アプリケーション動作環境

アプリケーションは、以下の環境で動作する。

CPU: Pentium 133MHz 以上

HDD: 10Mbyte 以上の空き容量

Memory: 32Mbyte 以上

ネットワークアダプタ（インターフェース）：1つ以上

OS: Microsoft Windows98 またはそれ以降のバージョンで、WindSock2.0 またはそれ以降をサポートしていること

6.3. アプリケーション動作

本アプリケーションは、単体の実行ファイル形式を持つ。

本アプリケーションは、以下の動作モードを持つ。

- ・データソースホスト・モード
- ・レシーバ・モード
- ・中継ホスト・モード

アプリケーション動作のための設定情報は、Windows98 上のレジストリに保持される。

6.4. アプリケーション設定項目

アプリケーションは、以下の設定項目を持つ。

- ・アプリケーション動作の設定
- ・上位 MLS 接続インターフェースの設定
- ・下位 MLS 接続インターフェースの設定

6.4.1. アプリケーション動作の設定

アプリケーション動作の設定値を、表 6 - 1 に示す。

設定	項目
ホスト動作モード	データソース / レシーバ / 中継
データ保存フォルダ	ディレクトリ
データ保存上限	サイズ
標準で使用する SHRM アドレス	マルチキャスト / ブロードキャスト / ユニキャスト
標準で使用する制御ポート	1025 ~ 65535
標準で使用する広報ポート	1025 ~ 65535
標準で使用する広報の送出間隔	1 ~ 3600
標準で設定する MLS 範囲 TTL	1 ~ 64

表 6 - 1 アプリケーション動作の設定項目

ホスト動作モード

ホストが SHRM でどの役割を行うかを決定する。データの受信のみを行う場合はレシーバ、最上 MLS にあってデータ配布を行う場合はデータソースホスト、2 つの MLS に接続しデータを中継する場合は中継ホストに設定する。

データ保存フォルダ

アプリケーションがデータソースホストの場合、下位 MLS から配布データを受信した時の一時的な格納場所として使用される。レシーバの場合、配布データを受信した際の保存場所として指定される。中継ホストの場合、上位 MLS から受信した、中継するデータの一時保存場所、また自身が受信したデータの保存場所として使用される。

データ保存上限

データ保存フォルダに格納されるデータの合計の上限を設定する。

標準で使用される SHRM アドレス / 制御ポート / 広報ポート / 広報の送出間隔 / MLS 範囲 TTL

各々、MLS 内データ配送時に標準の値として使用される設定値。

アプリケーション動作設定は、ユーザーインターフェース・モジュールを介して行う。

6.4.2. 上位MLS 接続インターフェースの設定

上位 MLS 接続インターフェースは、アプリケーションがレシーバまたは中継ホスト・モードで動作している時に有効である。アプリケーションがレシーバモードで動作している時は、MLS へ接続するインターフェースとして使用される。このインターフェースを介して MLS からデータを受信する。中継ホストモードの場合は、上位 MLS へ接続するインターフェースとして使用される。

上位 MLS 接続インターフェースの設定を、表 6 - 2 に示す。

設定	項目
使用アダプタ	ホストの持つアダプタから 1 つ選択
SHRM アドレス	上位 MLS で使用される SHRM のアドレス
広報受信ポート	上位 MLS で使用される広報のポート番号

表 6 - 2 上位 MLS 接続インターフェースの設定項目

上位 MLS 接続インターフェース設定は、ユーザーインターフェース・モジュールを介して行う。

6.4.3. 下位MLS 接続インターフェースの設定

下位 MLS 接続インターフェースは、アプリケーションがデータソースホストまたは中継ホスト・モードで動作している時に有効である。アプリケーションがデータソースホストで動作している時は、MLS へ接続するインターフェースとして使用される。このインターフェースを介して、MLS へデータを送信する。中継ホストモードの場合は、下位 MLS へ接続するインターフェースとして使用され、上位 MLS 接続インターフェースから受信したデータをこのインターフェースへ中継する。

下位 MLS 接続インターフェースの設定を、表 6 - 3 に示す。

設定	項目
使用アダプタ	ホストの持つアダプタから 1 つ選択
SHRM アドレス	下位 MLS で使用する SHRM のアドレス
広報送信ポート	下位 MLS で使用する広報のポート番号
広報送信間隔	広報を送信する間隔
MLS 範囲 TTL	下位 MLS の範囲を決定する TTL
通信路帯域	下位 MLS の通信路帯域

表 6 - 3 下位 MLS 接続インターフェースの設定項目

下位 MLS 接続インターフェース設定は、ユーザーインターフェース・モジュールを介して行う。

6 . 5 . アプリケーション構成

アプリケーションは、以下のモジュールによって構成される。

- ・ 配送データ管理モジュール
- ・ SHRM データ広報モジュール
- ・ SHRM データ配送モジュール
- ・ データ受信モジュール
- ・ SHRM 制御モジュール
- ・ SHRM データ中継モジュール
- ・ ネットワーク・アダプタ操作モジュール
- ・ ユーザーインターフェース・モジュール
- ・ アプリケーション管理モジュール

各モジュール間では、Windows98 のメッセージにより相互に命令を発行する。

各モジュール及び物理的なハードウェアの関係を、図 6 - 1 に示す。

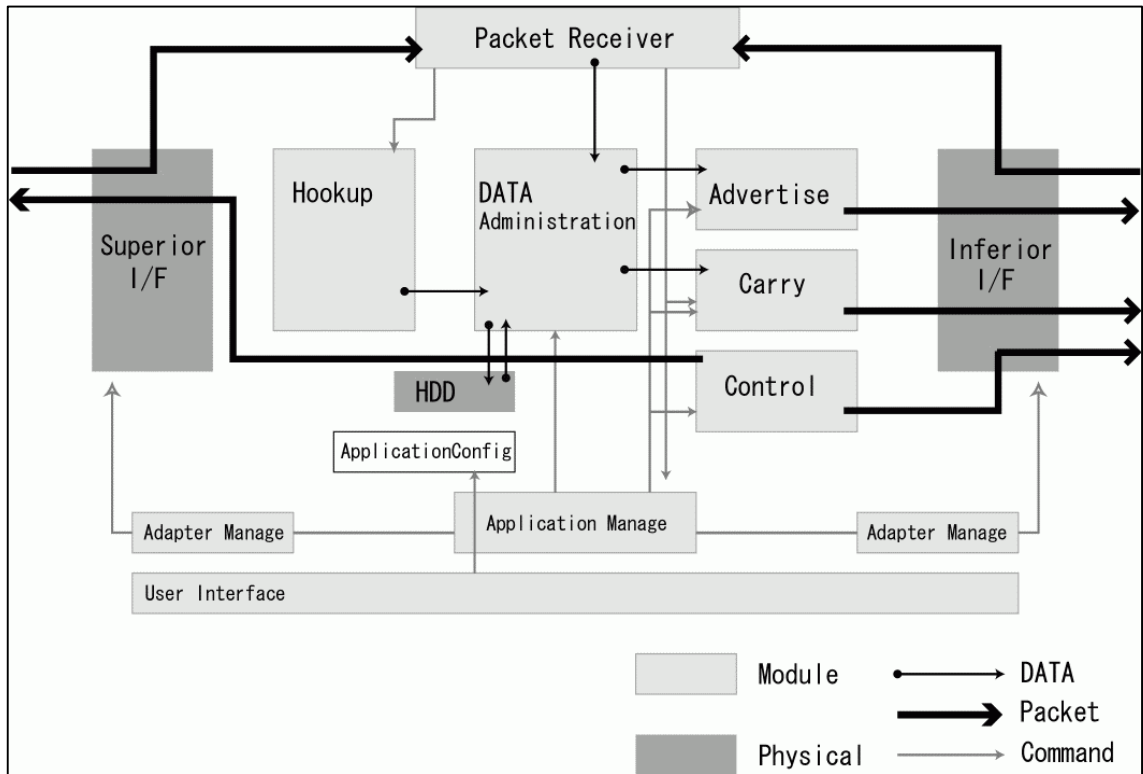


図6 - 1 アプリケーション構成図

6.5.1. 配送データ管理モジュール

配送データ管理モジュールでは、配送データの追加、広報パケットの生成、送出セルの順序の管理を行う。本モジュールは、アプリケーションがデータソースホスト・モードまたは中継ホスト・モードで動作している時に動作する。

配送データ管理モジュールでは、配送データ情報テーブルを持つ。配送データ情報テーブルは、各配送データに対して1組の配送データ情報を保持する。また、本モジュールは個々の配布データに基づいて、予め広報パケットを生成し、保持する。

配送データ情報は、表6 - 4に示す通りである。

情報種別	内容
チャンネル番号	データ配布に使用するポート番号
データ名	配布データの Data ID String
送信状態	停止 / 広報中 / 送信中
レシーバカウンタ	レシーバからの参照数
送信中セル番号	送信を行っているセルの番号
最終参照セル番号	最後に参照されたセルの番号
データ有効期限	配布データの有効期限

表6 - 4 配送データ情報

配布データが追加された場合、新しく配送データ情報を配送データ情報保持テーブルに追加する。データ名は、ユーザが手動で設定するか、または以下のルールによって名付けられる。

データ名 = 下位接続 MLS インターフェースの IP アドレス文字列 +
OS 内部時計の 32bit 値 + ファイル名

配布データが削除された場合は、該当する配送データ情報をテーブルから削除する。

送信状態は、停止、広報中、送信中の3つの値を取る。停止は広報・セルともに送信していないことを表し、広報中は、広報のみを送信しセルを送信していないことを表す。送信中はセルを送信していることを表す。

データ受信モジュールからデータ送信要求の受信を指示された場合は、レシーバカウンタを1加算する。

送信中セル番号は、送信を行うセルの番号を記録し、セルの送信が終了すると1加算される。レシーバカウンタが1の時は、送信中セル番号が最終参照セル番号に到達した時点でレシーバカウンタを0とし、送信状態を広報中とする。レシーバカウンタが2以上の場合は、送信中セル番号がデータの最終セルに到達したときにレシーバカウンタを1減算し、

送信中セル番号を 0 に設定する。

データ有効期限の過ぎた配布データは、自動的に配布データ情報保持テーブルから削除される。

6.5.2.SHRM データ広報モジュール

SHRM データ広報モジュールでは、配送データ管理モジュールの情報に応じて、下位 MLS 接続インターフェースを使用してデータ広報を送出する。本モジュールは、アプリケーション管理モジュールから、定期的に応報の送付を指示される。この間隔は、下位 MLS 接続インターフェースの設定による。

データ広報モジュールは、配布データ管理モジュールの持つ配布データ情報保持テーブル中で、送信状態が広報中または送付中となっている配布データを検索する。

送信状態が広報中または送付中の配布データについては、配布データ管理モジュールの持つ広報パケットを受け取り、下位 MLS 接続インターフェースに設定されたインターフェースから、UDP を使用して広報パケットを送出する。

6.5.3.SHRM データ配送モジュール

SHRM データ配送モジュールでは、配送データ管理モジュールの情報に応じて、下位 MLS 接続インターフェースを使用してデータ配布のためのセルを送出する。本モジュールは、アプリケーション管理モジュールからの指示によって駆動する。

データ配布は、通常のセル送付と、欠落セル補正配送に分かれる。

通常のセル送付では、データ配布モジュールは、配布データ管理モジュールの持つ配布データ情報保持テーブル中で、送信状態が送付中となっている配布データを検索する。

送信状態が送付中の配布データについては、配布データ情報保持テーブル中の送付中セル番号に対応するデータ部分からセルを生成し、下位 MLS 接続インターフェースに設定されたインターフェースから、UDP を使用してセルを送出する。セル送付後、配布データ管理モジュールにセル送付終了を通知する。

欠落セル補正配送は、データ受信モジュールから特定セルの送付を指示された場合に駆動する。指定された配送データ、指定セル番号から、セルを生成し、下位 MLS 接続インターフェースに設定されたインターフェースから、UDP を使用してセルを送出する。

6.5.4. データ受信モジュール

データ受信モジュールでは、インターフェースからのパケットの受信と受信パケットの解析、受信パケットの内容に応じたコマンドルーティングを行う。

受信パケットには、表 6 - 5 に挙げる種類がある。

パケット種別	パケットの内容
広報パケット	ADVERTISE/ADVERTISE_UNIDIR
配布パケット	DATA_START/DATA_CONT/DATA_EOT
	DATA_SPECIFIED
制御パケット	REQ_DATA_SEND
	REQ_SPECIFIED_CELL

表 6 - 5 コマンドルーティングを行うパケットの種類

コマンドルーティングは、アプリケーションの動作モードによって異なる。

データソースホスト

受信するパケットは、制御パケットの送信要求(REQ_DATA_SEND)、セル要求(REQ_SPECIFIED_CELL)である。

送信要求を受信した場合、配布データ管理モジュールに対して、配布データ送信要求を指示する。セル要求を受信した場合、SHRM データ配布モジュールに対して特定セルの配送を指示し、パラメータとしてセル要求に格納されている要求データ ID、要求セル番号を指示する。

レシーバ

受信するパケットは、広報パケット(ADVERTISE, ADVERTISE_UNIDIR)、配送パケットのセル巡回(DATA_START, DATA_CONT, DATA_EOT)、セル特定(DATA_SPECIFIED)である。

広報パケットを受信した場合、受信可能な配送データとして、受信した広報データの内容をアプリケーション管理モジュールに通知する。

配送パケットを受信した場合、受信したセルの番号に応じてデータを保存し、アプリケーション管理モジュールに対して受信したセルのセル番号を通知する。

中継ホスト

中継ホストでは、上位 MLS 接続インターフェースからは、レシーバと同様のパケットを受信する。また、下位 MLS 接続インターフェースからは、データソースホストと同様のパケットを受信する。各々の処理方法は、データソースホスト、レシーバと同様である。

6.5.5.SHRM 制御モジュール

SHRM 制御モジュールでは、制御パケットの送出を行う。

本モジュールは、アプリケーションがデータソースホストまたは中継ホスト・モードで動作している時に動作する。本モジュールは、アプリケーション管理モジュールからの指示によって駆動する。

データ送信要求を指示された場合、アプリケーション管理モジュールの指示するデータ ID に対する送信要求を行う制御パケットを生成し、上位 MLS 接続インターフェースから送信する。

欠落データ送信要求を指示された場合、アプリケーション管理モジュールの指示するデータ ID、セル番号に対する特定セルの送信要求を行う制御パケットを生成し、上位 MLS 接続インターフェースから送信する。

6.5.6.SHRM データ中継モジュール

SHRM データ中継モジュールでは、上位 MLS と下位 MLS の間の中継に伴う処理を行う。

データ中継では、上位 MLS からデータを受信し、中継ホストでデータを再構成した後、下位 MLS に対して広報する。

広報の中継では、広報番号の変換、データチャネルの変換、制御ポートの変換を行う。上位 MLS から受信した広報に基づいて、配布データ情報保持テーブルを生成する。配布データ情報保持テーブルの各配布データ情報に基づいて下位 MLS 用の広報パケットを生成、保持し、下位 MLS に対して広報を行う。

配布データの中継では、中継ホストでのデータを受信と、下位 MLS 用のセル生成を行う。まず上位 MLS から中継する配布データをすべて受信し、ディスク上に保存する。その後、配布データを下位 MLS で使用するセルに分割し、下位 MLS からのデータ送信要求に基づいてデータを配布する。

6.5.7.ネットワーク・アダプタ操作モジュール

ネットワークアダプタ操作モジュールでは、Windows98 を介して、ホストに接続されたネットワークインターフェース情報の取得と変更を行う。

本モジュールは、以下に挙げる機能を持つ。

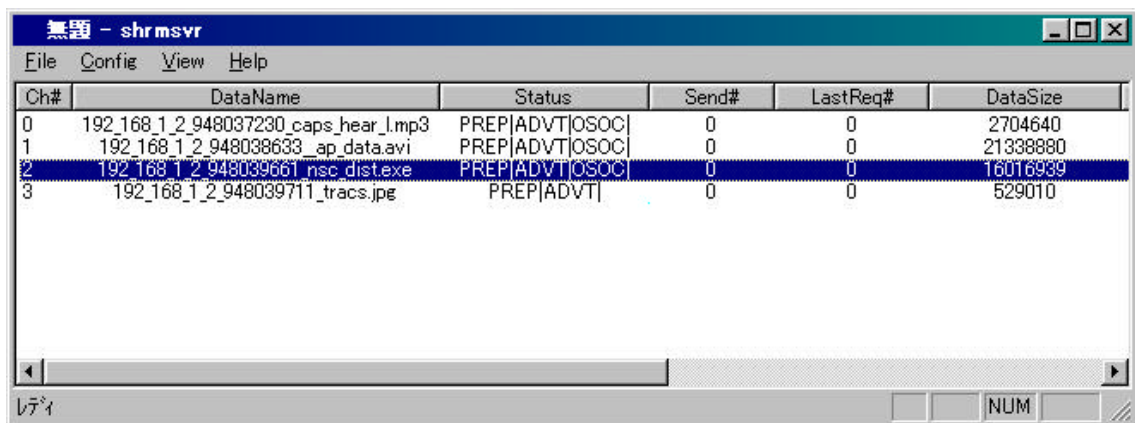
- ・ネットワーク・アダプタに付与されたアダプタ名称の取得
- ・IPアドレスの取得
- ・ネットマスクの取得
- ・マルチキャストグループへの参加・離脱、
- ・マルチキャスト送信パケットの TTL 設定

6.5.8. ユーザーインターフェース・モジュール

ユーザーインターフェース・モジュールは、ユーザに対してアプリケーションの設定、状態の確認を行うためのウィンドウを表示する。ウィンドウの表示は、Windows98 標準のコモンダイアログを使用する。

本モジュールは、以下のウィンドウを表示する。

- ・メインウィンドウ (図 6 - 2 参照)
- ・アプリケーション設定ダイアログ (図 6 - 3 参照)
- ・上位 MLS 接続インターフェース設定ダイアログ (図 6 - 4 参照)
- ・下位 MLS 接続インターフェース設定ダイアログ (図 6 - 5 参照)
- ・データソース追加ダイアログ (図 6 - 6 参照)



Ch#	DataName	Status	Send#	LastReq#	DataSize
0	192_168_1_2_948037230_caps_hear_lmp3	PREP ADVT OSOC	0	0	2704640
1	192_168_1_2_948038633_ap_data.avi	PREP ADVT OSOC	0	0	21338880
2	192_168_1_2_948039661_nsc_dist.exe	PREP ADVT OSOC	0	0	16016939
3	192_168_1_2_948039711_tracs.jpg	PREP ADVT	0	0	529010

図 6 - 2 メインウィンドウ画面

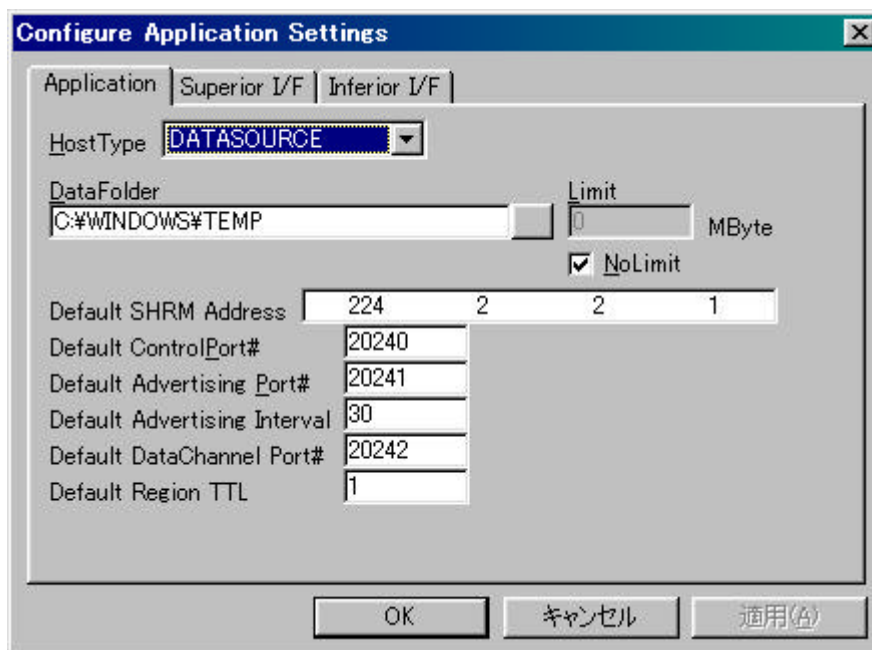


図 6 - 3 アプリケーション設定画面ダイアログ

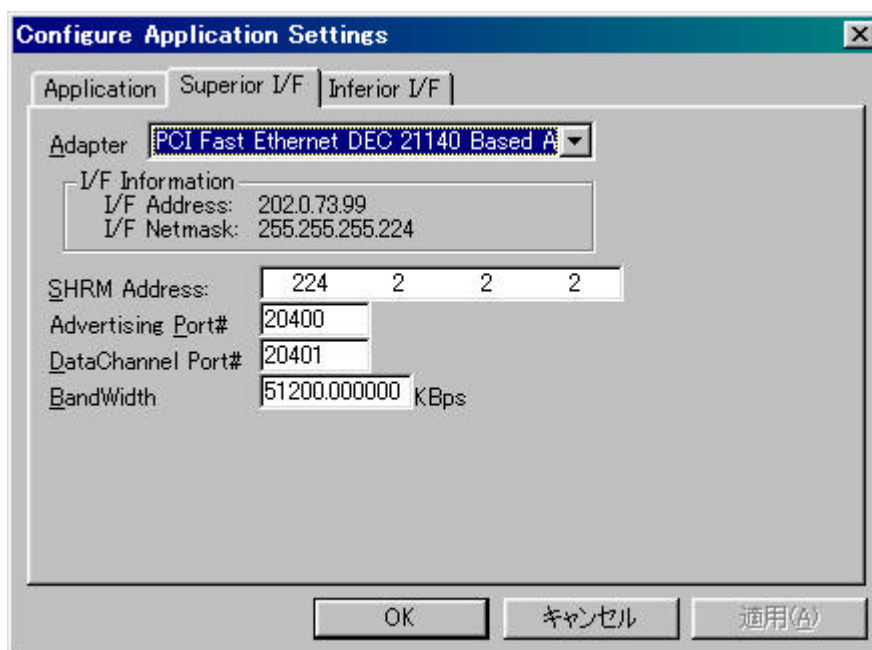


図 6 - 4 上位MLS 接続インターフェース設定ダイアログ

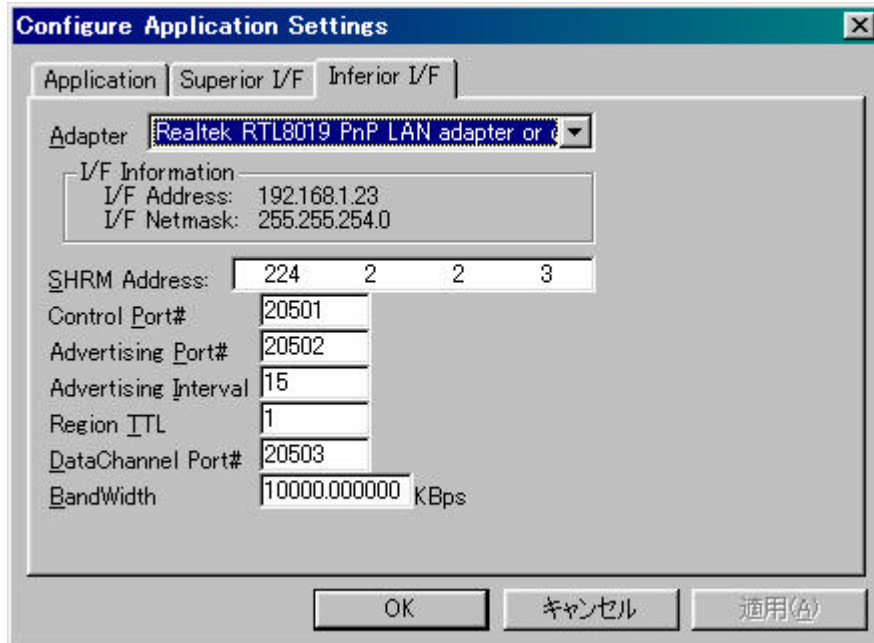


図 6 - 5 下位 MLS 接続インターフェース設定ダイアログ

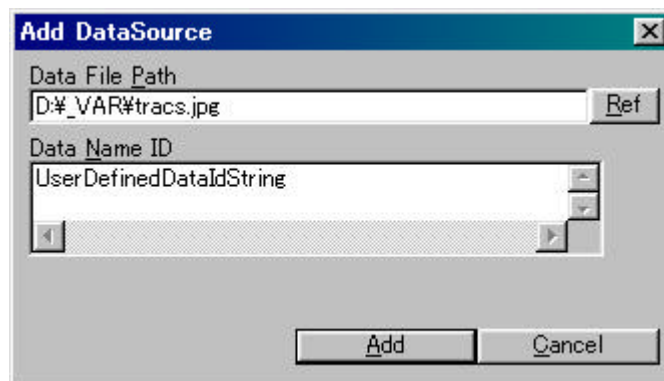


図 6 - 6 データソース追加ダイアログ

6 . 5 . 9 . アプリケーション管理モジュール

アプリケーション管理モジュールでは、各モジュールからの情報に基づいて、SHRM の各プロトコルを制御する。

本モジュールは、内部にタイマーを持つ。このタイマーを使用して、広報パケット送出間隔毎に SHRM 広報モジュールに広報の送出を指示する。また、単位時間内に送出されたパケットの総量を積算し、システム全体が単位時間あたりに送出するパケットの総量を制御する役割を持つ。

第7章

評価

本章では、これまで述べてきた SHRM の他の提案との定性評価について述べる。比較の対象として、以下のものを挙げる。

- RMTP (Reliable Multicast Transport Protocol)
- SRM (Scaleable Reliable Multicast)
- RMP (Reliable Multicast Protocol)
- TRAM (Tree-based Reliable Multicast)
- MFTP
- SHRM

比較は、以下の6点について行った。

- **信頼性保証の方法**

信頼性の保証を、どのような方法で行っているか

- **基本的なトポロジー**

信頼性保証を行う際に、基本としているトポロジーの型

- **フロー制御**

パケット流量の制御を、どのような方法で行っているか

- **途中参加の可否**

セッションに通信途中から参加できるか、また途中参加の場合、他の受信者と同様のデータが受信できるか

- **通信路の多様性への考慮**

狭帯域の通信路に広帯域の通信路上の通信が影響されないか

- **受信者数の規模**

受信者数の多寡に対する規模性

この比較の結果を、表7 - 1に挙げる。

Protocol	信頼性保証の方法	基本的なトポロジー	フロー制御	途中参加の可否	通信路の多様性への考慮	受信者数の規模
RMTP	ACK+ Aggregation	TREE	Cache	(一部のみ受信可)	×	大
MFTP	NACK	STAR	Rate		×	大
RMP	ACK	RING	-		×	小
TRAM	SACK	TREE	Window	(一部のみ受信可)	×	大
SHRM	NACK	TREE+ STAR	Rate (静的)			大

表7 - 1 各 Reliable Multicast と本研究の比較

複数の研究者が指摘しているように、すべての目的に使用できるような信頼性マルチキャストの実現は不可能である。このため、既存の信頼性マルチキャストはそれぞれ特徴を持って設計されており、本論文で提案する手法である SHRM においてもこれは同様である。

本論文で提案する手法では、フロー制御やツリー構造の形成、またパケットの到着順序保証を行わない点で他の手法と利用範囲が異なる。損失の多い狭帯域無線ネットワークなどの上での利用には適さず、ツリー構造を手動で形成する点で、自由度は高くない。また、到着順序が意味を持つアプリケーションのデータ配送には適さない。しかし、そのかわり従来のものにはない特徴として、通信路の多様性の考慮がある。他の手法を用いた場合、どうしても狭帯域で接続された受信者に起因する再送によって、広帯域通信路上の受信者が再送処理の間待たされる、という問題がある。SHRM では、この問題が解決され、広帯域ネットワーク上の受信者はより効率のよいデータ配送を行える。

データ配送が Reliable Multicast に対して必要とする要求をよく考慮し、SHRM がその要求に合致すれば SHRM を使用し、そうでない時は他の手法を用いることが、最も望ましいと言える。

第 8 章

まとめと今後の課題

8 . 1 . 本論文のまとめ

本論文では、現在のインターネット上でのアプリケーションで必要とされる、信頼性を持ったアーカイブ配送のための信頼性マルチキャストデータ配送システムについて提案した。

現在のインターネットとコンピュータ、アプリケーションの使用形態から、データ配送システムの必要性とそれが持つべき要件について述べ、またこれを実現する上での問題点について論じた。

既存のデータ配送システムである Reliable Multicast の提案を複数取り上げて考察し、Reliable Multicast を利用したデータ配送システムを開発する上での考え方について考察した。

本論文では、現在のデータ配送システムでの問題点を解決するために、マルチキャスト可能なローカルセグメント(MLS: Multicastable Local Segment)という手法を導入し、MLS を使用した広域なデータ配送システムを提案した。MLS 及び MLS によって構成される階層構造に基づいて、責任範囲を限定した信頼保証の仕組みと、MLS 相互の協調によって広域なネットワーク上で動作するデータ配送システムを提案した。本システムの設計について述べ、この設計に基づいてシステムを実装した。

本論文で提案したシステムが有効であることを示すため、本論文での実装についての評価を行った。これにより、本論文で提案する手法がインターネットを利用したアーカイブの配送について有効であることを示した。

本論文で提案する手法を用いることにより、広域なネットワーク上で効率のよいアーカイブの配送が実現できる。

8 . 2 . 今後の課題

本論文では、MLS 及び MLS 階層構造という新しい手法を提案したが、この手法には MLS 階層構造をどのように形成するか、という仕組みは含まれていない。SHRM は、MLS

が広帯域なものから狭帯域なものへと階層構造をなしている場合に、最も効率が良くなる。しかし、現実のネットワークでは、必ずしも回線の帯域の広いものから狭いものへと階層構造をなしているとは限らない。また、回線で使用できる帯域は、回線の混雑状況にも左右される。従って、SHRM を実際のネットワーク上に後広範に適用するために、SHRM が動作する基幹となるツリー構造は、現実のトポロジーや回線状態に即して動的に構成される必要がある。このため、MLS 階層構造をネットワークの状況に応じて動的に形成する仕組みを提供する必要がある。

また、本論文の実装では、通信路の持つ帯域に対する配布時の回線使用の割合を静的に設定している。この方法では、回線が輻輳した場合に、適切なスループットが得られず、また他の TCP による通信などを阻害するおそれがある。従って、配送システムに動的な輻輳制御の機構を導入する必要がある。Reliable Multicast における輻輳制御は、配送全体のスループットと深く関わっており、この点については更に研究の余地が大きい。

将来的には、これらの問題を解決し、インターネットでのアーカイブ配送について汎用なデータ配送システムを構築する予定である。

謝辞

本研究を進めるにあたり御指導を賜りました慶應義塾大学環境情報学部教授村井純博士並びに徳田英幸博士に深く感謝いたします。また、絶えず貴重な御助言と御指導を頂きました慶應義塾大学環境情報学部助教授楠本博之博士に、心より御礼を申し上げます。

本研究を進める上で終始惜しめない御助力を頂きました、慶應義塾大学政策メディア研究科関谷勇司氏、石橋啓一郎氏に深く感謝致します。また、本論文の執筆に当たりまして常に励ましと御協力を頂きました慶應義塾大学環境情報学部徳田・村井・楠本・中村研究室の諸兄、特に渡部陽仁氏、仲山昌宏氏に感謝致します。

参考文献

[WTM]前田亮, 植村俊亮, 「WWWにおけるマルチメディアファイルの容量分析」,第9回データ工学ワークショップ論文集, March 1998

[RMP] Whetten, B., Kaplan, S., Montgomery, T., "A High Performance Totally Ordered Multicast Protocol", submitted to INFOCOM'95

[Chang] Chang, J., and Maxemchuk, N., "Reliable Broadcast Protocols", ACM Transactions on Computer Systems, Vol.2, No.3, pp.251-275, Aug. 1984.

[RMTP] Jhon, C., Lin, Sanjoy Paul, "RMTP: A Reliable Multicast Transport Protocol", Proceeding of IEEE INFOCAM '96, March 1996, pp. 1414-1424

[SRM] Sally Floyd, Van Jacobson, Steven McCanne, "A Reliable Multicast Framework for Light-weight Sessions and Application Level Framing", ACM SIGCOMM'95, Aug. 7, 1995.