

Application Layer Multicast を用いた高性能マルチメディアコンテンツネットワークの構築

蘇 洲 早稲田大学基幹理工学部助教

1 はじめに

近年の高速回線の急速な普及に伴い、今までWEBブラウジングやメールといった比較的小さな情報量のやりとりから、映像配信や音楽配信といったマルチメディア通信や、テレビ電話、テレビ会議といった映像コミュニケーション通信のような大きな情報量のやりとりが増えている。この大きな情報量のやりとりを可能にする技術の一つとしてIP マルチキャストがある。しかしこのIP マルチキャストはルータを実装しなければならない、その実装困難さやマネジメントの困難さから大規模な普及には至っていない。そこでその代替技術として、ALM(Application Layer Multicast)が注目されている。(1-4)

ALMはマルチキャストの主要機能を各エンドホストに行わせることによって、仮想的なマルチキャストを実現させる技術である。IP マルチキャストではルータでパケットの複製をし送信していたのを、ALMではルータを用いず、ユーザレベルで各ホストがパケットを複製し送信させている。これにより、ALMは実装が容易であり、また柔軟なネットワークを構築することが可能である。しかし、ALMはマルチキャストのコア機能を信頼性が高いとは言えない各エンドホストに行わせるため、システム全体の信頼性がエンドホストに依存するという問題点を有する。(5-7)

ALMの既存の手法としてTree-based 方式とMesh-based 方式があるが、これらの手法は各ユーザの離脱によってシステムが分断しやすかったり、システムの規模が大きくなったときに通信処理の負荷が大きくなるといった問題点がある。また最近では、これらの問題点をカバーする手法としてTree-based 方式とMesh-based 方式を組み合わせた手法が考えられているが、そのシステムのルーティング方法や無駄なデータ通信によるシステム負荷の増大など、新たな問題が発生している(8)。大きな情報量の通信をする上で、マルチキャスト通信はサーバーの負荷を下げ、ネットワーク帯域の有効利用を図ることができる。それは大きな情報量をより速く、そして確実に届けることにつながる。しかし、現在までのマルチキャスト技術はまだ不完全であり、その研究、改善の余地は大いにあると思われる。本研究ではマルチキャスト技術の中でも特に柔軟性に富み、より効率よく大きなデータ情報を配信可能なALMに着目し、その中でもTree-based 方式とMesh-based 方式を組み合わせたTree-Mesh 混合型ハイブリッドALM システムの技術を改善、検討する。具体的には、従来のTree-Mesh 混合型ハイブリッドALM システムでの、ノードをツリーに構築する際の親ノードの選び方に新たなパラメータを追加したより有効性があると思われる親ノードの選び方と、グループ単位でのシステムマネジメント方法を追加した新たなTree-Mesh 混合型ハイブリッドALM システムを考察、検討し、システムの負荷を軽減させ、データ配信の効率性を向上させる手法を提案する。

2 一对多の応用通信手法

従来の通信手法に加えて、近年では従来の通信手法を応用した様々な通信手法が研究されている。以下にその代表例として

- P2P
- Application Layer Multicast

を挙げ、それぞれについて述べる。

2-1 P2P

P2P(Peer to Peer) とは、ユニキャストによって各ホスト(Peer) が不特定多数のホスト(Peer) と通信を行う技術である。P2P では、実ネットワークを意識せずに不特定多数と通信する。そのため、実ネットワーク上に仮想のネットワークを張って通信しているように見えることから、オーバーレイネットワークという総称で呼ばれている。

P2P ネットワークを構築するための主な形態としてPure P2P とHybrid P2P の二つがある。

Pure P2P

Pure P2P の大きな特徴として、純粋なピア群だけで構築されているP2P ネットワークであることが挙げられる。これは、P2P ネットワークに参加しているピアを管理するサーバを保有しないということである。Pure P2P では、参加するピアの管理をP2P ネットワークを構築するピアそのものが情報を共有し、分散して情報を保持、管理する形態をとっている。そのため、スケーラビリティについて非常に優れたモデルだと言える。また、ピア情報がP2P ネットワーク内に分散して保持されるため、ネットワーク自体の匿名性が非常に高いという特徴を持つ。しかしピア管理サーバを保有しないため、管理情報が分散してしまい、管理を行うこと自体が非常に難しいという問題点がある。

Hybrid P2P

Hybrid P2P は、Pure P2P とは逆にピア管理サーバを保有していることが特徴として挙げられる。ピア間の通信に関してはピア同士が相互自由に通信を行い、ピアの管理に関する処理のみをピア管理サーバが行う。Hybrid P2P は管理サーバを保有するため、Pure P2P と比較すると管理コストが発生するため、スケーラビリティについては劣っている。匿名性に関してもピア管理が行われていることから、Pure P2P ほど高くない。その一方でピア管理が行なわれているため、P2P ネットワーク内における情報検索、ピア制御などの処理については非常に優れたモデルであると言える。

2-2 Application Layer Multicast

マルチキャストはルータで送信パケットを複製し、ルータが受信者に対してパケットを送信することで通信を実現する技術である。そのパケット複製、及び送信機能を各ホストに分担させることでマルチキャストを実現する技術をALM(Application Layer Multicast)と言う。

ALMでは、各ホスト(ユーザ、エンドホスト)に本来マルチキャストでルータが行っていたパケット複製や送信制御を行わせ、情報をパケツリレーのような形で伝達している。ALMでは、仮想的なマルチキャストネットワーク、すなわちP2P ネットワークによるオーバーレイネットワークを構築することでマルチキャストを実現している。つまり、ALMはP2P ネットワークを利用したマルチキャストであるため、すべてユニキャストのみで実現することが可能となっている。

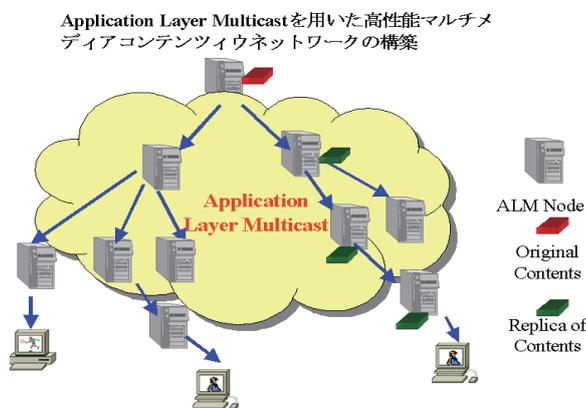


図1: Application Layer Multicast を用いた高性能マルチメディアコンテンツネットワークの構築

ALM とP2P、IP マルチキャスト(通常のルータやスイッチを利用したマルチキャスト)の相関関係は以下の通りである。ALMのネットワークでは、ユニキャストによって各ホストのデータの受信、データの利用データの転送を行わせることでマルチキャストを実現させているため、各ホストにおける処理がソフトウェアベースの処理となり、特殊な条件を必要としなく、実装が容易であるという利点を持っている。

しかし、ALMは完全にユーザ依存型のネットワーク構築であるという欠点がある。ALMは通常のマルチキャストと比較して実装や構築が容易である反面、不安定なネットワークを構築してしまいがちであるという問題点を有している。また、昨今のインターネットは主にウェブ向きに設計されており、P2P ソフトウェアの動作に適さない構造が各所に見られる。特にADSL などでの上り帯域幅の狭さは、ALMにとって非常に大きな問題となる。ALMでは、各ノードが受信したデータを他のノードに提供するので、受信のための下り帯域幅だけでなく、送信のための上り帯域幅も重要となってくる。上り帯域幅が広いほど、トラフィックをより大きく増幅できるのである。

3. 提案手法

3.1 ALMにおける安定なマルチメディアコンテンツ配信経路の構築

3.1.1 基本方針

本提案手法のALMシステムの構築の仕方について、まず各ノード間の距離(hops)をもとに、ノードをいくつかのグループに分ける。そしてそのグループのすべてのノードの中から、ノードの安定性、帯域幅、転送時間という3つのパラメータをもとにデータの送受信がグループの中で最も効率的に行えるノードを選び、そのノードをそれぞれそのグループのスーパーノードとする。スーパーノードを親ノードとし、グループのその他のノードとツリーを構築する。そして子ノード同士をメッシュで構築し、各グループでTree-Mesh 混合型ハイブリッドALM を作り上げる

3.1.2 スーパーノード選出アルゴリズム

従来手法では、スーパーノードを選ぶ際に、ノードがネットワークに滞在して自分の役割を全うすることを期待できる時間、つまり各ノードがどれだけ安定しているかのみを考慮してスーパーノードを選んでいる。提案手法ではそれに加え、各ノードの帯域幅の広さとデータ転送にかかる時間も考える。帯域幅も非常に重要な要素となるALMにおいては、安定しているだけでなく、帯域幅が十分にあるノードがスーパーノードになる必要があると考える。よって、提案手法ではこの3つの要素を総合的に考慮して、グループの中で最も最適なノードだと思われるものをスーパーノードに選出する。これにより、より安定したデータ配信が可能なALM システムを構築できると考える。

3.1.3 安定なツリーの構造

本提案手法では、まずノードの考えられうる状態を表1のように4パターンに分け、それらの状態においてのデータ転送効率を各ノードごとに比べるとという方式を採用。具体的には、あるノード*i*について、そのノードがネットワークに滞在して自分の役割を全うすることを期待できる時間、つまりノード*i*がどれだけ安定して機能できるかというノード*i*の安定性を a_i 、ノード*i*が保有する帯域幅を b_i とする。そしてこれらのパラメータにおいて、ノード*i*が安定していて、十分な帯域幅も保有している状態をパターン1とし、その際に、ノード*i*があるノードと接続してからデータを転送し終わるまでにかかる時間を $t_{i,1}$ とする。同様にして、ノード*i*が安定しているが、十分な帯域幅を持たない場合をパターン2とし、その際のデータ転送にかかる時間を $t_{i,2}$ 、ノード*i*が安定したノードであるとは言えないが、十分な帯域幅を保有している状態をパターン3とし、そのときのデータ転送にかかる時間を $t_{i,3}$ 、ノード*i*が安定したノードではなく、さらに十分な帯域幅を持たない場合をパターン4とし、そのときのデータ転送にかかる時間を $t_{i,4}$ とする。このようにあるノード*i*における考えられる状態を4パターンに分け、その一つ一つの値を出して比べるのではなく、この4パターンの各々でかかるデータ転送時間をすべて合計したノード*i*の一つの基準値とも言える値で、最もスーパーノードにふさわしいノードを選ぶ方針である。

表1:ノード*i*の各状態におけるデータ転送効率

ノード <i>i</i> の安定性	ノード <i>i</i> の帯域幅	ノード <i>i</i> のデータ転送時間
$a_i(\bigcirc)$	$b_i(\bigcirc)$	$t_{i,1}$
$a_i(\bigcirc)$	$1-b_i(\times)$	$t_{i,2}$
$1-a_i(\times)$	$b_i(\bigcirc)$	$t_{i,3}$
$1-a_i(\times)$	$1-b_i(\times)$	$t_{i,4}$

各ノードの考えられうる状態を4パターンに分けたが、その状態を簡潔にわかりやすくまとめるために、この4パターンの状態を以下のようなマトリックスで表す。

$$M = \begin{bmatrix} M_1 \\ M_2 \\ M_3 \\ M_4 \end{bmatrix} = \begin{bmatrix} a_i & b_i & t_{i,1} \\ a_i & 1-b_i & t_{i,2} \\ 1-a_i & b_i & t_{i,3} \\ 1-a_i & 1-b_i & t_{i,4} \end{bmatrix} \quad (1)$$

そしてここから、スーパーノードを決めるために、この4パターンの各々でかかるデータ転送時間をすべて合計したものを以下の式のように求め、これをノード*i*の一つの基準値とし、スーパーノードを選出する。

$$\begin{aligned}
\Delta &= \sum_i M \\
&= M_1 + M_2 + M_3 + M_4 \\
&= a_i \cdot b_i \cdot t_{i,1} + a_i \cdot (1 - b_i) \cdot t_{i,2} \\
&\quad + (1 - a_i) \cdot b_i \cdot t_{i,3} + (1 - a_i) \cdot (1 - b_i) \cdot t_{i,4}
\end{aligned} \tag{2}$$

3.2 データ転送負荷分散とコンテンツ管理

従来手法では、一つの大きなALM システムの中で、親ノードから子ノードへのデータ転送ルートだけを管理し、子ノード同士では従来のメッシュと同じようにデータを氾濫させ、データ転送先をランダムに選んで転送している。これでは、メッシュだけと比べたら多少無駄なデータ転送を制御できるものの、まだまだ無駄なデータ転送が多く、なによりデータ転送の際のシステムに与える負荷が集中し、負荷が大きくなりやすいという問題がある。提案手法ではノードをいくつかのグループに分け、各々のグループで小規模なALMを構築する。それにより、メッシュ間のデータ転送にかかる負荷を分散させるだけでなく、いくつものグループがお互いにカバーし合えるので、より安定したALM システムとなる。また、グループ内でのデータ転送ルート、グループ間でのデータ転送ルートをデータ転送効率の良さと安定したデータ転送が可能かどうかという条件に当てはめて管理する。このルーティング管理を明確にすることで、無駄なデータ転送を極力無くした転送効率の良いALM システムを構築できると考える。

4. シミュレーション

4.1 シミュレーションモデル

- ・ノードを20 個とし、それを1 グループ5 ノード、全部で4 つのグループを作るとする。
- ・システムにおけるビットレートは128bps とする。
- ・システム内の各ノード間の距離は、Power-Law 分布(9) に基づいてランダムに構成される。
- ・ノードの離脱率は乱数によってランダムに決定される。
- ・各ノードが保有する帯域幅はBPD(Bounded Pareto Distribution) 分布(9) によってランダムに決定される。
- ・送られるデータサイズは一定の乱数によってランダムに決められる。
- ・各コンテンツの人気度は、Zipf 分布(9) によってランダムに決定される。これにより、人気のあるコンテンツがより多くリクエストされるとする。
- ・ノードがコンテンツリクエストを行う時間間隔は、Poisson 分布(10) によりランダムに決められる。
- ・あるノードが離脱後、そのあるノードと再接続するためにかかる時間を10 秒とする。

4.2 シミュレーション結果

4.2.1 Average Traffic の比較

まず、ネットワークにあるコンテンツを1000 個とし、各コンテンツの人気度はそれほど変わらない状態で比較を行った(図2)。図2 からわかるように、提案手法はmTreebone よりも約19% トラフィックを減少できていることがわかる。このことより、提案手法は従来手法に比べネットワークの負荷を軽減できていると言える。次に、コンテンツの人気度はあまり変わらない状態で、ネットワークにあるコンテンツ数を5000 個にして比較を行った(図3)。図3 からわかるように、提案手法はmTreebone よりも約7% トラフィックを減少できていることがわかる。このことより、コンテンツ数が増えても、提案手法は従来手法よりも負荷を減らすことができ、その有効性を示せた。ただ、コンテンツ数が増えることによって、ネットワークにかかる負荷も大きくなり、ある一定のコンテンツ数までいくと従来手法とほとんど変わらなくなることが予想される。最後に、コンテンツの数は1000 個とし、各コンテンツの人気度にばらつきがでるようにして比較を行った(図4)。つまり、人気のあるコンテンツはとても人気があり、人気のないコンテンツも存在する状態である。図4 からわかるように、提案手法はmTreebone よりも約19% トラフィックを減少できていることがわかる。このことより、人気なコンテンツにアクセスが集中したとしても、提案手法は安定したコンテンツ配信が可能であると言える。

4.2.2 Average Delay の比較

ネットワークにあるコンテンツ数を1000 個、各コンテンツの人気度にそれほど変化のない状態でまずは比

較を行った(図5). 図5 からわかるように, 提案手法はmTreebone よりも約23% 遅延を減らせていることがわかる. このことより, 従来手法に比べ提案手法はデータの転送効率を上げることができたと言える. 次に, コンテンツの人気度はあまり変わらない状態で, ネットワークにあるコンテンツ数を5000 個にして比較を行った(図6). 図6 からわかるように, 提案手法はmTreebone よりも約6% 遅延を減少できていることがわかる.

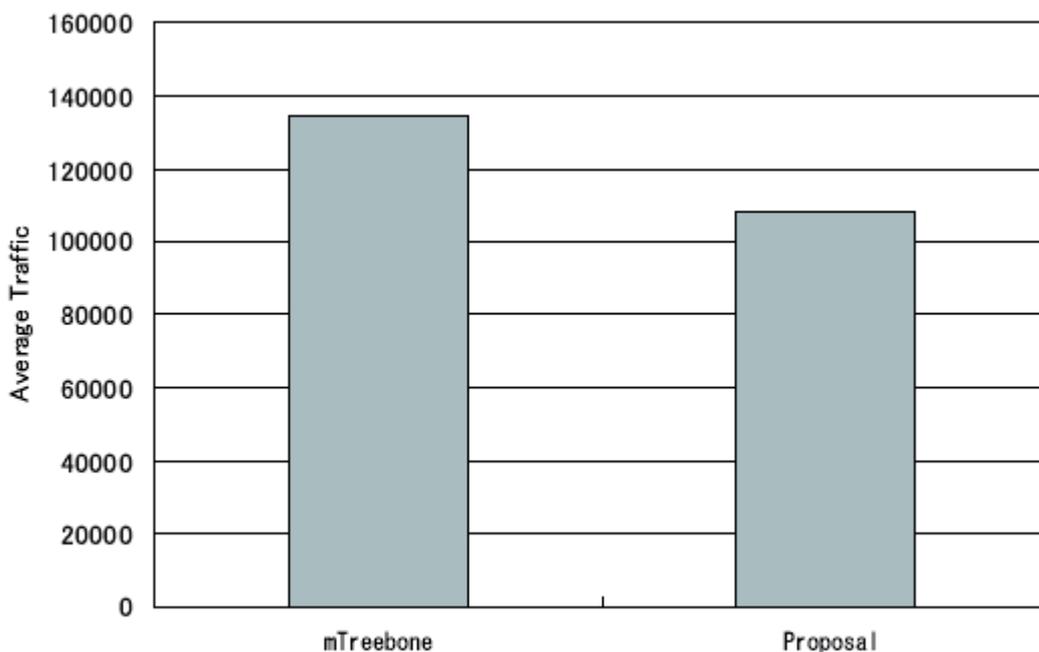


図2 : コンテンツ数1000, コンテンツ人気度平均時のAverage Traffic

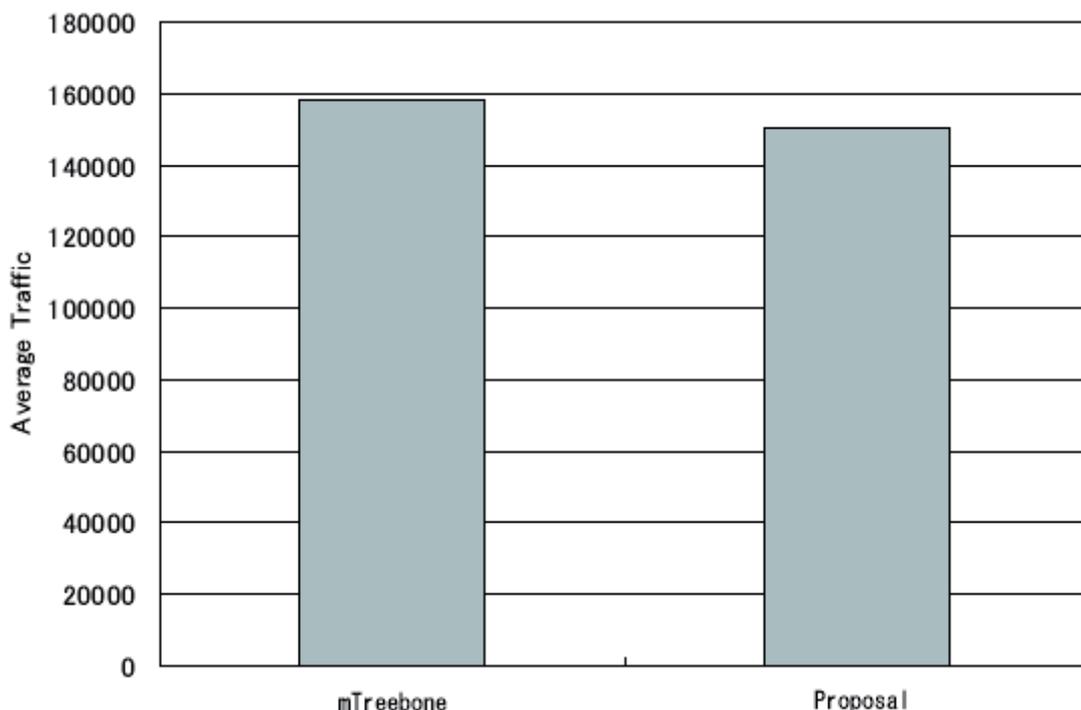


図3 : コンテンツ数5000, コンテンツ人気度平均時のAverage Traffic

コンテンツ数が増えることによって, ネットワークで起こる遅延も大きくなり, もっとコンテンツ数が増えれば従来手法とあまり変わらないデータ転送効率となることが予想されるが, コンテンツ数が多少増えても, 提案手法は従来手法よりも良い転送効率でデータを転送することができることを示せた. 最後にコンテンツ

の数は1000個で、各コンテンツの人気度にばらつきがでるようにして比較を行った(図7)。図7からわかるように、提案手法はmTreeboneよりも約23%遅延を減らせていることがわかる。このことより、人気なコンテンツにアクセスが集中したとしても、提案手法は従来手法に比べ効率の良いデータ転送ができると言える。

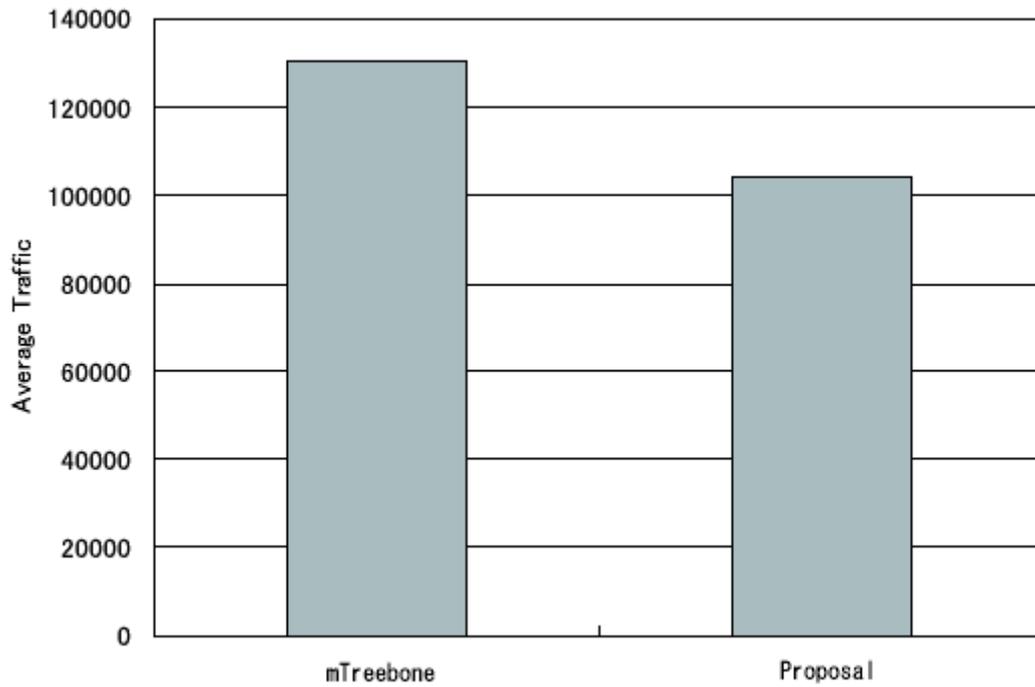


図4：コンテンツ数1000，コンテンツ人気度偏重時のAverage Traffic

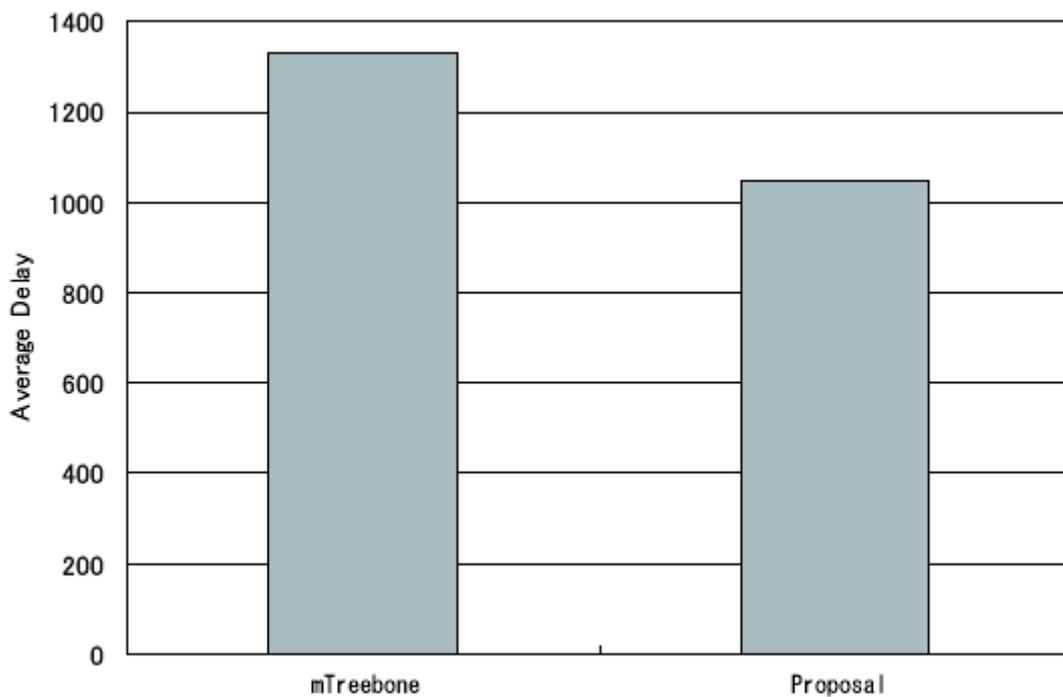


図5：コンテンツ数1000，コンテンツ人気度平均時のAverage Delay

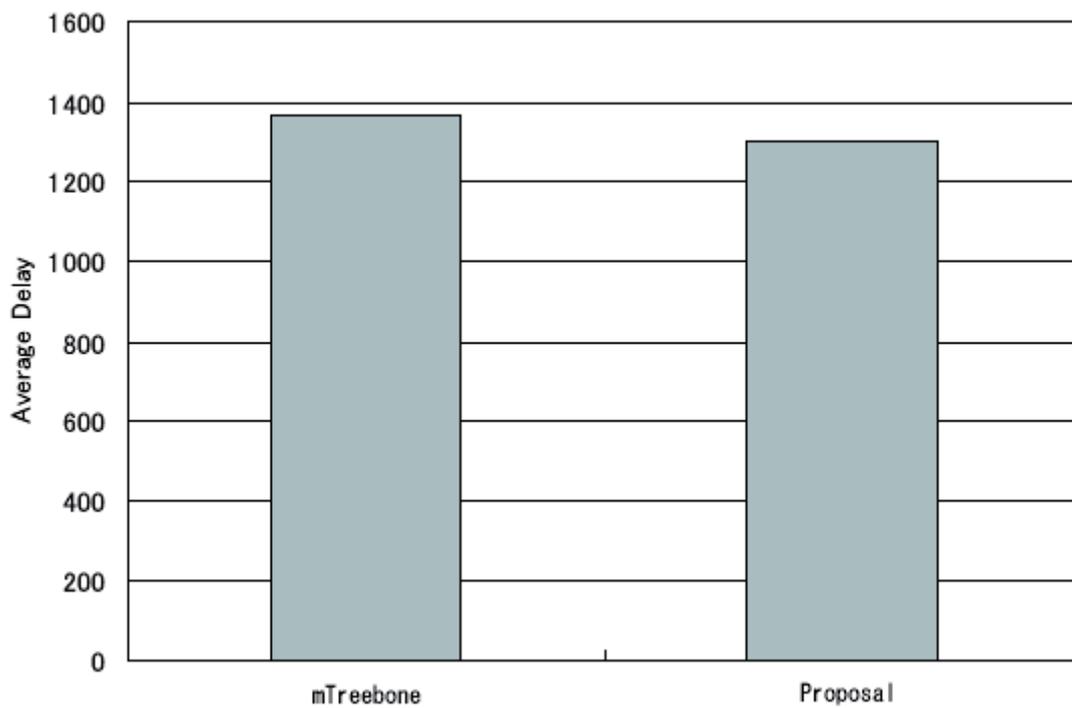


図 6 : コンテンツ数 5000, コンテンツ人気度平均時の Average Delay

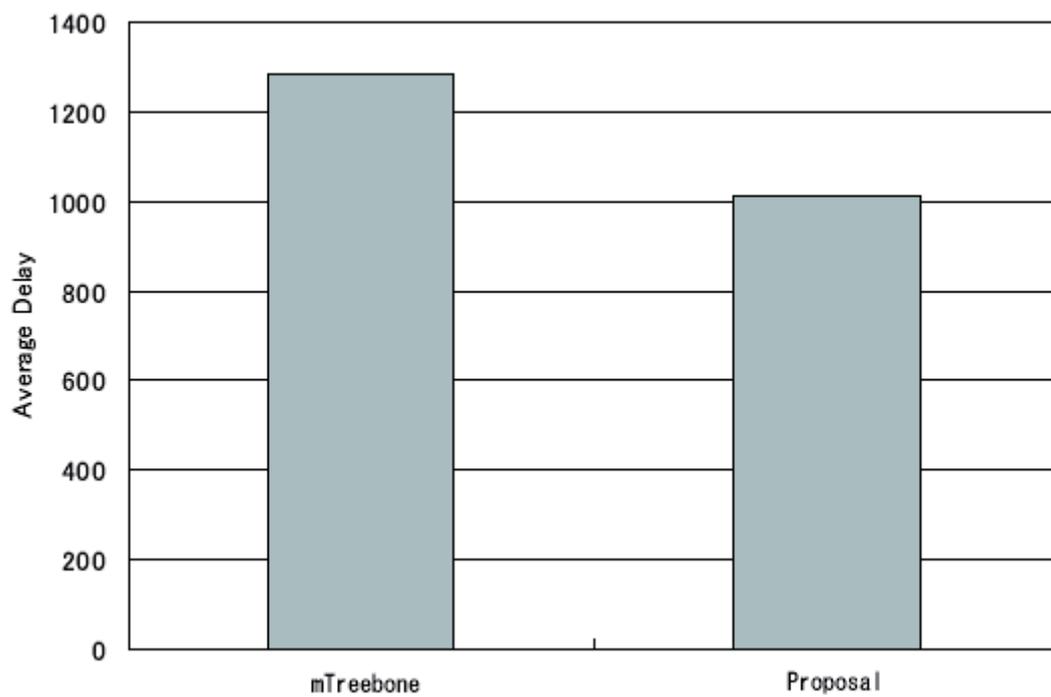


図 7 : コンテンツ数 1000, コンテンツ人気度偏重時の Average Delay

【参考文献】

- (1) IETF, “ RFC Pages, ”<http://www.ietf.org/rfc.html>

- (2) “インターネットの歴史と仕組み,”
<http://sta@.miyakyo-u.ac.jp/m-taira/Lecture/history-internet.html>
- (3) 小口敦司, “独立複数ツリー型分散ALMストリーミングシステムの提案,” 2006年度早稲田大学富永研究室修士論文
- (4) Cisco, “IPマルチキャストテクノロジーの概要,”
<http://www.cisco.com/japanese/warp/public/3/jp/index.shtml>
- (5) M.Hosseini, D.T.Ahmed, S.Shirmohammadi, N.D.Georganas, “A Survey of Application-Layer Multicast Protocols,” IEEE Communications Surveys, Tutorials, 2007
- (6) D.Li, J.Wu, Y.Cui, J.Liu, “QoS-Aware Streaming in Overlay Multicast Considering the Selfishness in Construction Action,” IEEE INFOCOM, May 2007
- (7) 首藤一幸 “アプリケーション層マルチキャスト:基本と応用,” UNIX magazine, October 2006
- (8) F.Wang, Y.Xiong, J.Liu, “mTreebone: A Hybrid Tree/Mesh Overlay for Application-Layer Live Video Multicast,” ICDCS, 2007
- (9) Lada A. Adamic, “Zipf, Power-laws, and Pareto - a ranking tutorial,” Information Dynamics Lab, HP Labs
- (10) “Poisson Distribution” <http://next1.cc.it-hiroshima.ac.jp/MULTIMEDIA/prob/node14>.

〈発表資料〉

題名	掲載誌・学会名等	発表年月
Efficient Construction in ALM with Assignment of Layered Degree and ALM-Bi-Cast	Proceedings of IEEE CCNC2009	2009, 1
ALMにおけるTree-Mesh混合型ハイブリッドALMシステムの検討	電子情報通信学会春季全国大会	2009, 3