

## IP マルチキャスト参加ホスト間の遅延差を軽減する

### パケット配信プロトコルの提案

永田 宏<sup>1)</sup>、伊藤 篤<sup>2)</sup>、飯作俊一<sup>1)</sup>、浅見 徹<sup>1)</sup>

(株)KDD 研究所<sup>1)</sup>

慶応大学 G-SEC<sup>2)</sup>

Abstract: 我々は IP マルチキャストにおける時間厳守型パケット配信、すなわちホスト間の遅延差を(出来る限り)解消するための新しい方法を研究している。今回は代表的な IP マルチキャストルーティングプロトコルである DVMRP に照準を絞り、マルチキャストグループに属するすべてのホストへのパケット同時配信方法を考案し、ソフトウェアデザインを行った。我々の方法は、マルチキャストルートルーターから遅延時間測定パケットを発信して各リーフルーターの遅延時間を測定し、それに基づいてマルチキャストデータグラムを各ホストに配信するタイミングを一致させようというものである。この方法は DVMRP だけでなく、PIM-DM にも有効であると考えられる。

## A New Protocol to Reduce Transmission Delay Differences

### Among Members of IP Multicast

Hiroshi Nagata<sup>1)</sup>, Atsushi Itoh<sup>2)</sup>, Shun-ichi Iisaku<sup>1)</sup>, and Tohru Asami<sup>1)</sup>

KDD R&D Laboratories Inc.<sup>1)</sup>

Keio G-Sec<sup>2)</sup>

Abstract: A new protocol to deliver multicast packets without delay differentials to all members participating a DVMRP multicast session is proposed. In our protocol, the maximum transmission delay of the multicast tree and the transmission delay of each leaf router belonging to the tree are measured and informed it to all the members. Each router delays the delivery of multicast packets to hosts over its sub-network according to the maximum delay and the delay at the router. As the result, multicast packets are delivered to all the hosts at the same time. The protocol is under implementation. It will be useful to realise fair packet delivery over IP multicasts.

## 1. はじめに

インターネットにおけるパケット伝送遅延問題を時間地図の概念で捉えなおし、遅延制御問題における主な課題を整理した。その中から「時間地図の形状制御」問題を本研究のテーマとした。具体的にはDVMRPマルチキャストツリーにおいて、全参加ホストに(ほとんど)同時にパケットを配信するためのプロトコルを考案した。

## 2. インターネットの時間地図

時間地図は、中心点から出発し目標地点に到達するまでに要する時間を尺度として作られる地図の一種である。図 1-a は東京駅を中心とし、JR の路線を使った場合の時間地図の一例である。各都市への実際の距離と到達時間は必ずしも一致しないため、時間地図の形状は通常の地図と異なっている。インターネットもパケットの伝送遅延時間を尺度とした時間地図によって表現することができる。図 1-b は我々の研究所から各ドメインまでの遅延時間をもとに描いた時間地図である。

時間地図の概念を用いると、インターネットにおける伝送遅延問題を次のように単純に整理することができる。すなわち

- (1) 時間地図の半径を縮小する。
- (2) 時間地図の形状の変動を減少させる。
- (3) 時間地図の形状そのものを制御する。

(1)はネットワークの高速化によって伝送遅延を出来る限り小さくしようという方向である。(2)は帯域の変動、すなわち jitter を抑えることにより、伝送遅延の変動を減少させようという方向である。実ネットワークでは jitter がつき物なので、インターネット時間地図は時々刻々形状を変化させている。Jitter を減らせば、時間地図の形を固定させることが出来る。(3)はより複雑で、特定のホストないしグループに対する伝送遅延時間を意図的に操作しようという方向である(図1-c)。特定

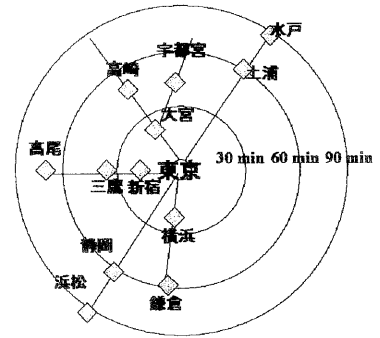


図 1-a. 東京駅を中心にして JR を利用したときの時間地図。

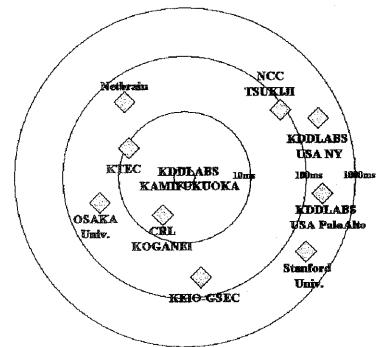


図 1-b. KDD 研究所(上福岡)を中心に、tracert を用いて作成した時間地図。

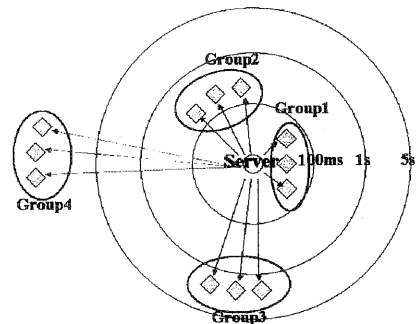


図 1-c. 遅延時間を意図的に操作することにより、時間地図の形状を制御する。

のホストやグループに対して決められた時間内にパケットを配信することが可能になれば、インターネット時間地図の形状を制御することができる。これら3つの課題が解決されれば、我々はインターネットをより自由に、かつ効率的に運用できることになるだろう。

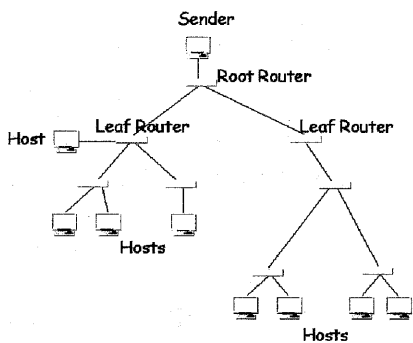


図 2-a. DVMRP のマルチキャストツリー

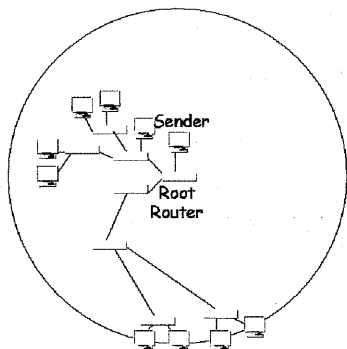


図 2-b. DVMRP マルチキャストツリーの時間地図

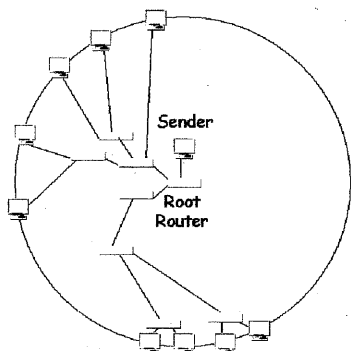


図 2-c. 遅延制御の目標

### 3. DVMRP マルチキャスト上のパケット同時配信

これら3つの問題のうち、我々が扱うのは3番目の問題である。本研究はその最初の取り組みである。特定のグループに対する遅延の意図的操作を行うのには、IP マルチキャストがもっとも適している<sup>1-4)</sup>。様々な IP マルチキャストルーティングプロトコルが存在するが、今回は DVMRP<sup>5-8)</sup>を選んだ。DVMRP は最短ルートの収束が遅く、スケーラビリティに対応できないなど様々な欠点を有するが、現在でも代表的な IP マルチキャストルーティングプロトコルとして使われている。しかも mroute のソースコードが公開されているという利点がある。

なお、今回の研究では、ネットワークは十分な帯域を有する LAN に限定し、jitter は十分に小さいという前提条件を設けた。

DVMRP では、図2-a のようなマルチキャストツリーが生成される。各ホストへのパケット配信遅延に基づいて時間地図を作ると、図2-b のようになる。すなわち一般的にルートルーターに近いホストほどパケットが早く到着し、遠いホストほど遅く到着する。本研究ではこれを図2-cのように、すべてのホストに同時にパケットを到着させる方法を検討した。これは時間地図の形状制御としてはもっとも単純な形であるが、実用価値は高いと考えられる。たとえばマルチキャストを用いて株式情報を配信し、各ホストがその情報に基づいて自動的に売買発注を行うようなビジネスソリューションでは、売買の公平性を保つうえで有効と考えられる。仮にパケットの到着遅延差が1秒程度であったとしても、その1秒間にホストコンピュータは株式情報を分析し、売買注文を発することができるだろう。ルートルーターにより近いホストが、より有利に取引を行うことができるわけである。こうした不公平を防ぐために、すべてのホストに対して同時にパケットを配信する仕組みが必要になる。

#### 4. 基本方針

この問題を解決するうえで重要なことは、DVMRPにおいては(その他のIPマルチキャストでも同様だが)、ルートルーターがマルチキャストツリーの大きさおよびツリーを構成するリーフルーターに関する情報を持っていないということである。そのためルートルーターは各リーフルーターまでのパケット配信遅延を予測することができない。同様に各リーフルーターは、自分がルートルーターからどれだけ離れているか(ホップ数がどれだけか)という情報を持っていないため、パケットの遅延時間を予測することができない。しかし各リーフルーターにおける遅延と、マルチキャストツリー全体の最大遅延が分かれば、図3に示す方法によって、各ホストへのパケット配信を同時に行うことが理論上は可能になる。

すなわち、マルチキャストデータグラムがルーター-Nに到着するまでに要する時間を $T(N)$ とし、最末端のルーター-Mに到着するのに要する時間を $T(Max)$ とする。また各リーフルーターからそれぞれのホストにデータグラムを配信するのに要する時間は十分に小さいものとする。この場合、Host-NとHost-Mに同時にデータグラムを到着させるためには、ルーター-Nにおいて $T(Max) - T(N)$ だけ、Host-Nへの配信を遅らせればよい。

そこで問題になるのが、 $T(N)$ と $T(Max)$ をどうやって知るかということになる。今回我々は、実際の遅延時間を測定することで、この問題を解決することにした。図4は我々が考えた問題解決のスキームである。まずルートルーターから遅延測定用のパケットを、マルチキャストツリーに沿って発信する。このパケットに対する返信を各ルーターが発信する。これによってリーフルーターは配下のルーターの遅延時間を知ることができる。次にルートルーターにおいて遅延テーブルを作成し、これをマルチキャストツリーに沿ってリーフルーターに配信する。遅延テーブルには、ツリーの最大遅延時間も記述しておく。各リーフルーターはこれらの数値を用いて、配下の

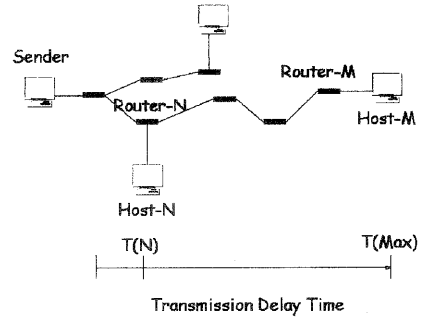


図3. すべてのホストに対して同時にデータグラムを配信するための方法。

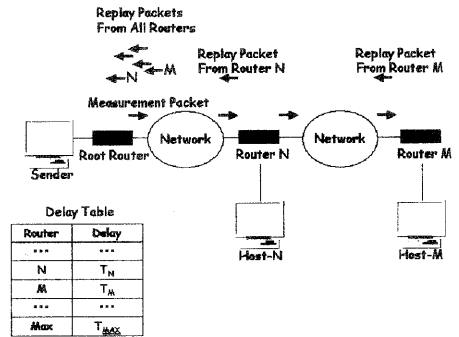


図4. 問題解決のためのスキーム。

ホストに対するデータグラム配信のタイミングを調節する。このようにして、すべての参加ホストに対して同時にデータグラム配信ができるようになる。

#### 5. 設計

このアイデアに基づいてソフトウェアの設計を行った(図5)。詳細は以下のとおりである。

最初に自分がルートルーターかリーフルーターかを、VDMRPルーティングテーブルを用いて判断する①。ルートルーターの場合には遅延測定用のパケットを発信し、返信パケットの到着を待つ③~⑥。測定パケットを発信するタイミングは、DVMRPがツリーを再構成した直後とする②。ま



かどうかは、ソフトウェアを実装してLAN上で動かし、確かめてみる必要がある。現在、その作業を進めているところである。

ソフトウェア設計の②の部分、すなわち測定パケットを発信するタイミングの取り方に関しては一考を要する。DVMRP ではリーフルーターが下流から Prune メッセージを受けてデータグラムの配信を停止しても、一定時間が経過すると Prune が無効になり、再びデータグラムの配信を開始する。このタイマーの値がルーターごとに異なっている場合、Prune 受信時間がまちまちになるので、全ルーターへの遅延時間測定を計測するタイミングが取りにくくなることが考えられる。この問題を回避するためには、遅延時間測定を完全に周期的に行うのという方法が考えられる。これに関しては今後、実験を通して検討していくつもりである。

十分に帯域が確保され、jitter が小さいネットワークでは、遅延時間そのものを測定するよりも、Hop 数を計測し、使用したほうがよいかもしれない。あるいはルートルーターから各リーフルーターまでの metric を使用するという方法も考えられる。これらの課題も今後、実験を通して検討する予定である。

我々のプロトコルが有効かどうかはまだ検証されていないが、仮にこのプロトコルで満足な結果が得られた場合には、同様のプロトコルを密集モードの PIM (PIM-DM) にも利用できるはずである。ただし散在モードの PIM (PIM-SM) では、この方法は使えないだろう。広域に利用できる Scalable な IP マルチキャストでの遅延時間制御に関しては、まったく別の発想に基づく方法を考案する必要があると考えられる。その問題に関しても、今後取り組んでいきたいと考えている。

#### 参考文献

1. R.Wittmann and M.Zitterbard, Multicast Communication, Morgan Kaufmann (1999)
2. T.A.Maufer 著、楠本博之訳、IP マルチキャスト入門、共立出版(2000)
3. D.Kosiur 著、苅田幸雄訳、マスタリング TCP/IP、オーム社(1999)

4. Kenneth M. Multicast Networking and Applications, Addison Wesley Longman, Inc. (1999)
5. Deering S. Multicast Routing in Internetworks and Extended LANs, Proceeding of ACM SIGCOMM 88. 55-64 (1988)
6. Deering S. & Cheriton D. Multicast routing in datagram internetworks and extended LANs, ACM Transaction on Computer Systems, 85-111 (May 1990)
7. RFC 1075, Distance Vector Multicast Routing Protocol, Waitzman D., Partridge C., and Deering S.E., November 1988 (Status: Experimental)
8. RFC 1112, Host extensions for IP multicasting, Deering S.E. August 1989 (Status: Standard)