

GIOSプロジェクトにおける 高速マルチキャストネットワークの構築

長橋 賢吾¹ 神明 達哉² 江崎 浩³ 加藤 朗⁴ 村井 純⁵

¹ 慶應義塾大学院政策・メディア研究科 ² 株式会社 東芝 研究開発センター 通信プラットフォームラボラトリー
^{3,4} 東京大学情報基盤センター ⁵ 慶應義塾大学環境情報学部

1 はじめに

GIOSプロジェクトでは、大学(コアサイト)間での授業交換にIPv6マルチキャストを使用している。本稿では、こうした広帯域を使用するアプリケーションにおいてどのようにマルチキャストネットワークを構築し、運用すべきかということについて実際に起きた問題を踏まえながら述べる。

2 ネットワーク構成

ネットワーク構築にあたって以下の要求事項を考慮してネットワークを構築した。

高品質なネットワーク 10~17Mのデジタルビデオデータを流すには、帯域にボトルネックがあると著しく品質が低下する。

コアサイトからのフィードバック 各サイトはマルチキャストで送信する一方で、一つのコアサイトから質疑応答などのフィードバックを受けるためユニキャストでの受信が必要となる。

高品質なネットワークに関しては、物理的には、ギガビットイーサネット、ATM-PVCで接続し、物理層でボトルネックが生じることはないように設計した。

経路制御に関しては、ユニキャストルーティングに関しては、スタティックルートをそれぞれのルータで設定し、マルチキャストルーティングにはPIM-Sparsemode(PIM-SM)[1]を用い、図1のように、WIDE-KDD大手町NOCをRendezvous Point(RP)としたShared Treeを形成するように設計した。

ルータおよびホストはすべてPentiumIIIクラスのCPU、メモリは128MBのPCから構成され、ネットワークインターフェイスとして、ファーストイーサネットおよびOC-3 ATMカードを使用し、このATMカード上に3本のPVCを設定した。OSはFreeBSD 3.4-RELEASE上に、IPv6に対応させるためKAMEキットを用いた。

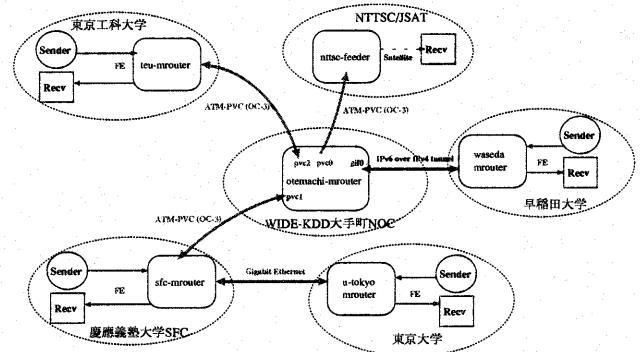


図1: ネットワーク構成図

3 運用上の問題点

3.1 問題点

上記のネットワーク環境において発生した問題としてRPによるパケットロス进行を挙げることができる。これは複数のソースが高いレートでデジタルビデオデータをマルチキャストで送信する際における現象であり、これによって授業の品質が著しく低下してしまう。

3.2 問題の検証

上記の問題を検証するため、各ルータでマルチキャストパケットに関するトラフィック測定をおこなった。実施環境は、以下の通りである。

- 慶應義塾大学SFCのsenderは、マルチキャストグループG1にストリームを送信
- 早稲田大学のsenderは、G2にストリームを送信
- 東京工科大学のreceiverは、G1にjoin
- 東京大学のreceiverは、G2にjoin
- G1、G2ともにRPは、otemachi-mrouterである

二つのマルチキャストデータを流している状況において、G1、G2は送信レート(16.8Mbps,11.9Mbps,10.7Mbps)を変化させながら1秒おきに30秒間、各インターフェイス(I/F)のパケット入出力数をnetstatコマンドを用いて計測し、平均化した。平均化する理由として、デジタルビデオのトラフィックパターンにはむらがあり、平均化することによって、そのむらを均等にするためである。G1に関する結果を表1、G2に関する結果を表2に示す。

表1、表2の結果によりトラフィックが重複するRPでパケットロスがおきていることがわかる。このパケットロスの原因として挙げられるのが、まず一つにデジタル

Constructing High-Speed Multicast Network in the GIOS Project

¹ Keio University, Graduate School of Media and Governance

² Communication Platform Lab. Corporate R&D Center, Toshiba Corp.

^{3,4} Information Technology Center, The University of Tokyo

⁵ Keio University Faculty of Information Environment

表 1: G1 に関する各 router の input/output packet 数/秒)

サイト	sfc	otemachi		teu
I/F	pvc0-out	pvc1-in	pvc2-out	pvc0-in
16.8Mbps	2945	2643	2432	2423
11.9Mbps	2243	2126	1934	1913
10.7Mbps	1971	1968	1893	1739

表 2: G2 に関する各 router の input/output packet 数/秒)

サイト	waseda	otemachi		u-tokyo
I/F	gif0-out	gif0-in	pvc1-out	ti0-in
16.8Mbps	2982	2833	2712	2704
11.9Mbps	2312	2296	2183	2113
10.7Mbps	1940	1928	1878	1894

ビデオが UDP パケットを輻輳制御することなしに送っていることが挙げられる。次に挙げられる原因として RP における処理のオーバーヘッドである。RP では、First Hop Router から RP カプセル化して送信されるデジタルビデオデータが含んだ PIM-Register パケットを処理し、Reverse Path Forwarding(RPF) を計算し、自身のマルチキャストルーティングテーブルを参照し、データを Outgoing Interface(OIF) に送信する。複数かつ広帯域のバルクデータでは、これらのパケットを出力する際にオーバーヘッドが増大していると考えられる。

3.3 問題解決へのアプローチ

前節でパケットロスの原因として、輻輳制御をしていないことと RP における処理のオーバーヘッドについて指摘したが、本稿では後者の問題について取り扱う。この問題を解決する一つの手段として、RP の分散がある。これは各サイトのマルチキャストルータが各サイトのグループアドレスを広告する方法である。そして、このアプローチをもとに、測定をした。条件はほぼ前節と同じであるが、異なる点として、G1、G2 の RP は”otemachi-mrouter”であったが、G1 の RP は’sfc-mrouter’、G2 の RP は、”waseda-mrouter” に設定し、同様の方法でトラフィックを測定した。

表 3、表 4 から明らかなように、このトポロジーでは RP を分散させることにより、RP を集中させる方法では”otemachi-mrouter”で G1、G2 におけるパケットロスが 10 パーセントを越えているのに対して、RP を分散される方法では、パケットロスはごくわずかである。これは従来は、一つの RP が、二つのグループに関する Register パケット処理し、RPF を計算し、OIF にマルチキャストデータを送信していたが、今回は RP を 2 つに分散させたために、上記の処理を一つの RP が一つのグループに

表 3: G1 に関する各 router の input/output packet 数/秒)

サイト	sfc	otemachi		teu
I/F	pvc0-out	pvc1-in	pvc2-out	pvc0-in
16.8Mbps	2881	2882	2879	2879
11.9Mbps	2234	2236	2232	2228
10.7Mbps	1983	1968	1963	1939

表 4: G2 に関する各 router の input/output packet 数/秒)

サイト	waseda	otemachi		u-tokyo
I/F	gif0-out	gif0-in	pvc1-out	ti0-in
16.8Mbps	2985	2983	2992	2994
11.9Mbps	2338	2346	2334	2332
10.7Mbps	1940	1928	1913	1904

関してパケットを処理するのみになったからである。

4 結論および今後の課題

高速ネットワークにおいてマルチキャスト、とくに Shared Tree を形成するマルチキャストルーティングプロトコルでは、RP におけるトラフィックの重複がアプリケーションの品質に致命的な影響を与えるといえる。本稿では、RP を複数に分散させることによってそれを避け、広帯域のマルチキャストネットワークを構築した。今後の課題としては、このマルチキャストネットワークをさらに踏み込んで評価する、PIM-SM で SharedTree と Shortest Path Tree(SPT) とを形成した場合にどの程度バルクトラフィックの伝送に影響をあたえるのかの評価、そして今回のような PC ベースでのマルチキャストパケット転送の限界がどの程度であるか、あるいはマルチキャストパケットをルータで複製することによって生じるジッタについても計測する必要がある。

また、本稿ではソース、レシーバとも少ない環境であったが、将来的にはレシーバは当然増加してくるはずである。そうした状況において、いかにスケーラビリティを確保するかということについても今後の課題である。

参考文献

- [1] D. Estrin, D. Farinacci, A. Helmy, D. Thaler, S. Deering, M. Handley, V. Jacobson, C. Liu, P. Sharma, L. Wei, "Protocol Independent Multicast-Sparse Mode (PIM-SM): Protocol", rfc2362, 1998