

オーバレイネットワークを用いたマルチホーム手法の地域における応用

菊池 豊†

† 高知工科大学 総合研究所

〒782-0003 高知県香美市土佐山田町宮ノ口185 高知工科大学連携研究センター 306B

E-mail: tyu@kikuken.org

あらまし オーバレイネットワークによるマルチホーム手法の地域における応用を提案する。オーバレイネットワーク手法は宇多らの提唱する多重経路型マルチホーム手法を用いている。まず、これを地域IP網に応用することで、地域ISPの柔軟な運用を可能にすることを示す。また、複数経路による冗長なストリーム伝送により、地域放送事業に十分なインターネット上での映像伝送が可能になることを示す。

キーワード マルチホーム、オーバレイネットワーク、地域情報化

Applications of the Multihoming Technology using Overlay Networks in Regional Areas

Yutaka KIKUCHI†

† 306B, Research Institute, Kochi University of Technology

Miyanokuchi 185 Tosayamada-cho, Kami-city, Kochi 781-0003 Japan

E-mail: tyu@kikuken.org

Abstract We propose applications of multihoming method on overlay network especially in regional areas. We have employed an overlay architecture proposed by Uda et.al. Firstly, we present it allows easier management of ISPs that use Flet's access networks provided by NTT east and west as the ISP's access line. Secondly, we present bi-cast on the overlay network architecture makes video streams on the Internet much reliable for CATV.

Key words multihome, overlay network, regional ICT

共 著 者

佐古 正浩 岸本 建也 柴田 祐輔 栗田 博生
STNet 愛媛 CATV

尾田 成樹 上野 圭 大平 真

NTT 西日本一四国

松本 浩明 山本 幸生 片岡 幸人 西内 一馬
高知システムズ シティネット

増岡 博文 松嶋 聡 濱崎 哲一 宇多 仁
西南地域 ソフトバンク 富士通 北陸先端大

ネットワーク テレコム 四国システムズ

Coauthor

SAKO, Masahiro KISHIMOTO, Ken'ya
STNet

ODA, Naruki UENO, Kei

OHIRA, Makoto

NTT-West Shikoku

SHIBATA, Yusuke KURITA, Hiroataka
Ehime CATV

MATSUMOTO, Hiroaki YAMAMOTO, Kosei
Kochi Systems

KATAOKA, Yukihito NISHIUCHI, Kazuma
Citynet

MASUOKA, Hirofumi MATSUSHIMA, Satoru
Swan TV Softbank Telecom

HAMASAKI, Norikazu UDA, Satoshi
Fujitsu Shikoku Systems Japan Advance Institute
of Science and Technology

1. ま え が き

BGP ver.4 [1] を用いる現状のインターネットの経路制御では、トラフィックエンジニアリング (TE) を行うことが困難である。

- AS 間に複数の経路がある場合の経路選択、特に流入するトラフィックを運用者が制御しにくい
- 経路断による再収束に時間がかかる
- フローごとでの経路選択ができない
- マスク長の長い経路の増加がルータの能力を圧迫している

これは AS 単位での大きな経路制御だけで問題になるのではなく、個人が複数の ISP に接続する際にも、同様の課題が発生する。宇多らはこの課題に対して「多重ルーティング型マルチホームアーキテクチャ」を提案している [2] [3] [4]。

我々は、宇多らの手法を応用することで、個人のインターネット接続だけでなく、地域のインターネットの課題を解決できると考えた。以下では、まず、背景となる技術について述べ、その後で本研究プロジェクトの基本的な考え方と、これまでに得られた知見について述べる。

1.1 BGP4 による経路制御

インターネットの経路制御技術は、統一した経路制御ポリシーを持つ自律システム (AS) 内の制御と AS 間の制御の大きく 2 つの階層に分かれている。前者を行うプロトコルを一般に IGP と後者を行うプロトコルを EGP と呼ぶ。インターネットでは EGP として事実上 BGP ver.4 のみが用いられている。

BGP ルータは自 AS から目的地 AS までに経由する AS の列 AS-path を保持しており、互いにこれを交換することでインターネット全体の AS-path を保持する。IP データグラムを受け取った場合は宛先 IP アドレスから宛先 AS を決定し、宛先 AS が自 AS でない場合は AS-path の最も短い経路を検索し隣接 BGP ルータに転送する。

運用者は、自 AS から出て行くデータグラムに対しては、local preference と呼ばれるパラメータにより転送先の隣接 BGP ルータを指定することができる。また自 AS に入ってくるデータグラムに対しては、prepend と呼ばれる手法で特定経路を避けるように誘導することができる。これが BGP による基本的なトラフィック制御 (TE) である。

Prepend は、自 AS の経路情報を隣接ルータに広告する際に、AS-path 長を長く見せることで、他 AS のルータに当該経路を避けるように間接的に通知する手法である。このため、自 AS に流入するトラフィックについては、必ずしも運用者の意図した通りに制御出来るものではない。典型的な例として、複数の隣接 AS を持つ場合に、トラフィックの種類に応じて適切なリンクへの配分が困難であることがある。

1.2 多重経路型マルチホーム技術

ここでは宇多らの提案である「多重ルーティング型マルチホームアーキテクチャ」について簡単に解説する。以降、メカニズム上のキーワードを用いて、この手法を単に DR-UR と記述する。

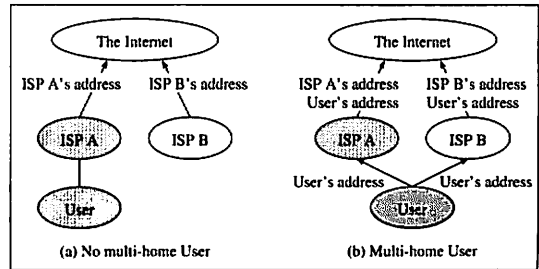


図1 マルチホーム

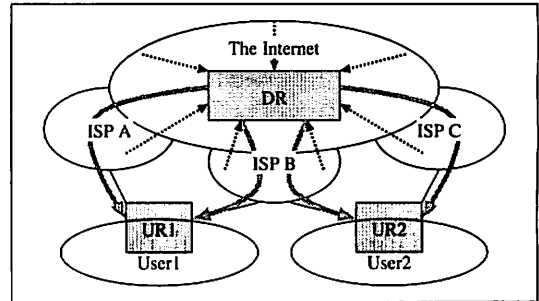


図2 DR-UR 基本アーキテクチャ

ある個人がインターネットに ISP A 経由で既に接続しているとする (図 1(a))。これをシングルホームと呼ぶ。このユーザが ISP B 経由でも接続したとする (図 1(b))。これをマルチホームと呼ぶ。

この場合、ユーザは ISP A と ISP B のリンクを何らかの形で有効に利用したいと考えるだろう。例えば、一方のリンクが故障などで切断した場合にユーザには認識されないような方法で他方のリンクに切り替わることや、ストリームや VoIP を一方に、WWW やファイル転送は他方にしたいこともあるだろう。従来のインターネットではアドレス空間の配布の問題や経路制御の問題により、このような要望は十分には満たしがたい。

宇多らの提案する DR-UR 手法は、ユーザのインターネットトラフィックを交換するために、以下を行う (図 2)。

- インターネット空間に DR と呼ばれるルータを、ユーザ側には UR と呼ばれるルータを用意する。
- ユーザに ISP A から ISP B からも独立なアドレスを与え、そのアドレスを含む CIDR ブロックを DR より広告する。
- DR-UR 間に ISP A と ISP B とを経由するトンネルを準備し、ユーザに出入りするトラフィックはトンネルを経由させる。
- DR-UR は UR 運用者のユーザや DR 運用者のポリシーに従い、通過するトラフィックをどちらかのリンクに割り当てる。

これにより、ユーザの要望を満たすことができる。

なお DR は複数設置することができるので、一般には、冗長性を持たせたり適度に短い経路を選択出来るようにするのが望ましい (図 3)。

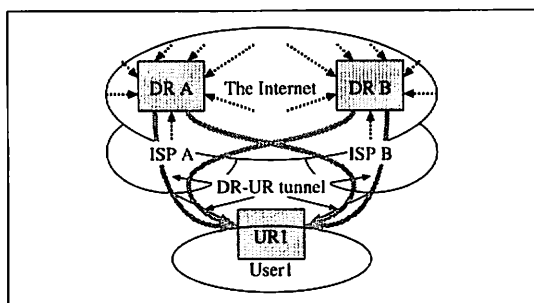


图 3 冗长 DR 構成

1.3 プロジェクト概要

地域の ISP はインターネット経路制御網上の葉にある場合が多く、主体的に経路制御を行うような環境が与えられていないことが多い。著者らは DR-UR の手法を、元々の目的にあった個人利用だけでなく、地域におけるインターネット環境へ応用可能であると考え、以下を目的として研究開発活動を行っている [5]。

- 地域における情報通信インフラを有効に活用する
- トラフィック制御を適切に行える
- インターネット全体に特別な仕掛けや負担をかけない
- 標準的なスキルのオペレータによる運用が可能

本プロジェクトのメンバーは四国内の地域 ISP を中心に構成している。プロジェクト名は Traffic Engineering for Regional Communities を略して “TEReCo” と称している。

TEReCo プロジェクトは、技術的な内容により、基本技術、IPv6、高品質、地域 IP 網応用、の 4 つのワーキンググループ (WG) を持つ。

このうち基本技術 WG では、PI アドレスを持たない PA アドレスのみの ISP がマルチホームを行う手法を提案し実験室レベルでの検証を行った [6]。

本稿では、第 4. 節で高品質 WG で行った複数ストリームによる高信頼性・高品質伝送について、第 2. 節で地域 IP 網応用 WG で行ったマルチホームの応用手法について述べる。

1.4 実験ネットワーク

地域において様々なネットワークポロジでの実験が可能になるように、実験用ネットワークを構築してある。

まず、各組織は L2 リンクで高知 IX に接続している。高知県の参加組織は、L2 リンクとして主に高知県が設置運営する高知県新情報ハイウェイの MPLS L2VPN を用いている。愛媛県にある参加組織は、高知県新情報ハイウェイと JGN II、さらに商用の L2 リンクサービスを用いて接続している。

この他、いくつかの参加組織は NTT 西日本の地域 IP 網（いわゆるフレッツサービス）に加入し、さらに相互に他の参加組織の運用する商用 ISP に加入している。地域 IP 網は 2 セッションの PPPoE を張れる契約であり、同時に 2 つの ISP への L3 リンクを持つことが出来る。

高知 IX は、菊池の経営する有限会社ナインレイヤーズが設置運用する地域 IX であり、L2 の地域内ピアリングサービスと

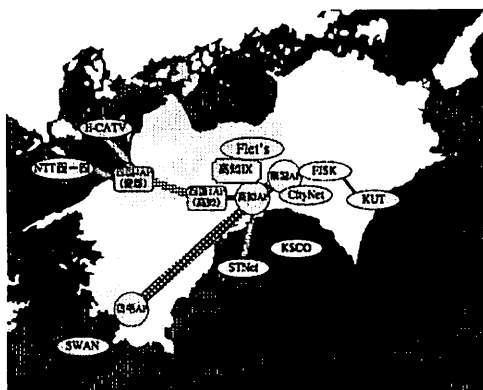


図 4 実験ネットワーク四国内 L2 構成図

L3のインターネットトランジットサービスを提供している。トランジット ISP を2つ有している。

実験用に高知 IX のトランジット ISP の iDC 内（東京と大阪に各 1 台）にハウジングスペースを借り、DR を設置した。その他、高知 IX 自体と、一部の参加組織に DR を設置した。また、全参加組織に UR を設置してある。

DR および UR は、ラックマウント型の 1U サーバに NetBSD を搭載し、その上に商用ライセンスの DR-UR の実装をインストールした。サーバには HP 製 DL320G2 を採用した。事前の調査で NetBSD (実験開始当時は ver. 1.6.2) との相性の良かったこと、遠隔から IP でコンソール KVM を操作することを可能にするオプションがあったことが採用理由である。

遠隔操作に専用の Ethernet ポートが用意されており、東京と大阪の iDC に設置したサーバにはこれにもグローバル IP を割り当てた。これにより、実験中にインバンドでの制御が不能になった際にもアウトバンドでサーバの回復が可能となった。実験中は経路情報も触ることなどより、実際に頻繁に不通となる。このような際にも東京や大阪の iDC に出向かなくて良く、実験の作業効率改善に貢献した。

2. ISP での応用

宇多らが DR-UR 手法を考案した元々の動機は、個人宅等のネットワークをマルチホーム接続することにあった。このため、DR-UR の提案自体は、/29 程度のエンドユーザのネットワークを UR 経由で接続するものである。

一方で ISP のネットワークまで含めて考えると、ネットワークのどこで UR を用いるかで、4 つのバリエーションを考えることができる。

- ユーザネットワーク（従来提案）
- ユーザ PC 単体
- ISP ネットワーク全体
- ISP の一部のネットワーク

まず、ISP のネットワーク全体を UR 配下に置く場合には、BGP4 による経路制御とは独立にトラフィック制御を行うことができる。またトランジットリンクの障害時に BGP4 の遅い切

り替え時間ではなく、DR-UR 間の死活監視による高速（パラメータの設定に依存する。デフォルト 300ms）な切り替えが可能となる。

また、マルチホームが可能になると言う他にも、トランジット ISP が提供する IP アドレスからの非依存性や、エンドホストの IP が移動先環境でも変化しないという性質を積極的に利用することができる。

例えば ISP は、地域 IP 網に収容されているユーザや、サーバを持たない個人や SOHO 等のユーザへは、PPPoE や DHCP により動的に IP アドレスを割り当てることが多い。このような場合で固定の IP アドレスを割り当てたい状況が発生すると、アドレス管理上の例外になったり、後日アドレスの付け替えが必要な場合にユーザの協力が必要になるなどの、運用上の障害となる可能性がある。DR-UR を用いて、このようなユーザ空間に対して、接続用リンク用の動的な IP アドレスの他に、ユーザが利用する固定の IP アドレスを供給することができる。

さらに、全ての動的 IP アドレスを割り当てるユーザの環境に UR を用いることで、接続を提供するアクセス網でトポロジー変更や運営上の理由による IP アドレス変更が生じた場合でも、ユーザの手を煩わせずに新しい環境に移行することが可能になる。

このような構造を実験室で仮想的に構築して、正しく動作することを確認している。

この他、地域 IP 網経由のローミング接続への応用がある。ISP の提供するインターネットサービスは、アクセス網のトラフィックを一旦自組織で集めてからインターネットと接続する場合の他、NTT 地域会社の地域 IP 網を用いて他通信事業者へのローミングによりインターネット接続をする場合があり、後者の場合は ISP 自体がユーザのトラフィックを制御することが出来ない。DR-UR を用いると、ユーザのトラフィックを ISP の設置した DR に一旦収容することができるので、ある程度の TE が可能となる。

3. モバイル通信への応用

DR-UR を用いると、エンドホストの IP アドレスが移動先環境でも変化しないという性質を積極的に利用することができる。

3.1 マルチホームとモバイル通信

従来は DR-UR はトランジット ISP 非依存のマルチホーム技術として考案された。一方で、MIP や NEMO と言ったモバイル技術においては、トンネリングを利用して移動先 IP 環境に依存しない通信を可能にする手法がとられる。このため結果として両者は強い類似性を持つ。DR-UR 手法は、ほぼそのままモバイル通信技術として利用することが出来る。実際、本実験で利用している DR-UR の実装での最近の版では DR-UR 間通信に MIP のプロトコルを援用している。

単体の PC に UR を実装することで、PC の接続リンクが変わっても、同一の IP アドレスで通信し続けることが可能になる。例えば、一定のアドレスブロック間だけの IP 到達性を許可し、かつ個々の IP アドレス自体を PC 個体と対応させるよう

なセキュリティ対策を施した場合を考える。このような場合に、通常は移動先で与えられる IP アドレスは PC 個体とは無関係であり、地理的に制限された範囲でしか利用ができなくなる。

DR-UR を利用すると、PC 単体に割り当てた固有 IP アドレスを移動先のリンクアドレスと独立に利用することが可能になる。この場合、PC 固有アドレスとリンクアドレスとのマッピングを行う際に悪意のあるホストが固有の IP アドレスを詐称できる余地がある。現在利用している実装では、ホストが移動して UR が新たに DR との接続を要求する際に、DR が UR の認証を行うことでこのリスクを回避している。

3.2 検証

実ネットワークにおいて、DR-UR を用いたモバイル環境が実現できることを確認した。

今回用いている DR-UR の実装は NetBSD を用いている。このため、ユーザが OS として直接 NetBSD を利用する場合には、当該 OS に UR を実装することで難なく実現できる。より広い範囲のユーザを対象としての検証をするため、ユーザが Windows XP を使うような環境で VMware を用いて実現してみた。

VMware のホスト OS は Windows XP Professional を用意し、ゲスト OS として NetBSD を搭載した。モバイル側の PC と外との通信は、PC の I/F を VMware の Bridge モードで行う。これで UR を動作させる NetBSD の仮想 I/F が外部リンクからみた宛先アドレスとなる。UR からユーザ側のネットワークは、VMware の Host Only モードで行う。これで UR とユーザの利用する OS とのネットワークが構成される（図 5）。

以上のメカニズムで、ユーザの利用する OS は、必ず UR およびトンネリング先の DR を経由してインターネットと IP を交換するようになる。ユーザ OS の IP アドレスは、あらかじめ与えてある固定の IP アドレスのままでインターネットとの接続性を維持する。リンクの IP アドレス（この場合は Bridge モードで接続した NetBSD の I/F 側のアドレス）が移動により変更されても、UR がその都度 DR とのトンネルを再構成し、Windows XP 側の IP アドレスは固定のままで通信できる。

実験では無線 LAN によるリンクを用いて、移動によりアクセスポイントが変更されても、通信が維持できることを確認した。具体的には WWW ブラウザで特定の WWW サーバをアクセスすることで確認をしている。

4. 高品質ストリーム伝送

本節では、映像伝送等の高い品質を要求される伝送への、DR-UR 手法の応用について述べる。

インターネット上での映像伝送は、サーバからユーザへの番組配信のみならず、事前には専用リンクを持たない局間での伝送や、取材現場から局へのライブ映像の伝送など、広い応用が期待できる。

一方で、テレビ放送に用いる伝送路には高い信頼性が要求され、インターネットでのいわゆるベストエフォートでは全く不十分である。放送事業者の意見をいくつかの場所で集めたとくろ、パケットロスはなくなくしてはならず、かつ障害によ

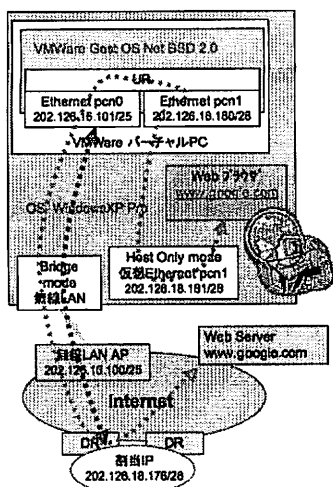


図 5 モバイル通信

る経路の切り替わりの不通時間が 50ms でも長すぎるかもしれないということであった。これは SDH/Sonet といった伝送路をそのまま用いてようやく実現できるレベルであり、それでも不十分かもしれない。

また、動画の符号化の観点で考えると、遅延の少ない DV 等の伝送の場合では通信路のジッタの影響があり、フレーム間圧縮をする MPEG 等の伝送の場合では 1 パケット落ちが多くのフレームのデコードに影響を与える。パケット落ち対策には FEC が多く用いられるものの、散発的なパケットロスに有効であっても、経路切り替えに伴うようなバースト落ちには向いていない。

さらに CATV 用に市販されている MPEG2 エンコーダを試験したところ、隣接するパケットの順序逆転が直ちにパケットロスと扱われたり、重複するパケットがあるとデコードが正常にできずに画面が激しく乱れるような製品があった。これは 1 本の L2 リンク上でエンコーダとデコーダを接続するような状況で用いることが暗黙に想定されているようであり、パケットの損失・重複・順序逆転が起こるインターネットでは利用できない。

今回用いた DR-UR の実装は独自の死活監視機能を持っており、デフォルトでは 100ms にプローブパケットを交換し、3 連続で損失が起るとリンク断と判断して経路を切り替えることができる。すなわち最悪 300ms でバックアップリンクへ切り替えることができる。このパラメータは変更可能で、より短い間隔での死活監視も可能ではある。これは、異なる AS 間での経路切り替え時間としては、最悪 90s 程度かかる BGP4 よりはるかに高速である。しかしながら、この死活監視機能を用いても映像伝送には十分とは言えない。

我々は以上のような状況から、映像伝送にインターネットを用いる際に、1 本のストリームに高い信頼性を与えることで実現することは極めて難しいと判断した。このため、複数のストリームでの冗長な伝送を行うことで、実現する手法にアプロー

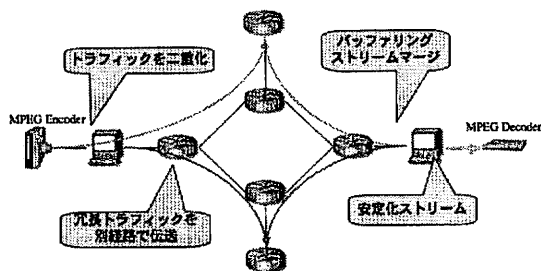


図 6 Bi-cast ネットワーク実験

チした。

4.1 Bi-cast による映像伝送

複数ストリームによる冗長な伝送のバリエーションとしては、どのような冗長性を取るか、どれぐらいのストリームを用いるかがある。今回は、完全に同じペイロードを持った 2 本のストリームによる、bi-cast による冗長伝送での実験を行った。

およそ以下の想定による模擬的なネットワークを実験室で準備した。伝送の送信側と受信側は異なる AS のネットワークに収容され、そのネットワークはそれぞれ UR を持つものとする。AS 間は直接のトランジットやピアの関係を持たず、それぞれ 2 本のリンクによりインターネットに接続している。インターネット側には複数の DR がありどの UR ともトンネルを構成している (図 6)。

映像の送信側には、サンプルビデオ映像を再生する DVD プレーヤの出力を、MPEG2 エンコーダで 1 本の RTP/UDP/IP ストリームとする。FreeBSD 上に構成したプログラムにより、このストリームを複製する。一方のストリームはそのまま出力され、他方は IP ヘッダの宛先ポート番号のみを変更して出力する。送出側の UR では経路制御ポリシーとして、ポート番号で送出 DR を振り分けるようにする。これによる 2 本のストリームは個別の経路を通るように制御できる。

どちらのストリームも DR 経由で受信側の UR に届き、宛先 IP アドレスは同じであるため 1 つのデコーダに向かう。ストリームがデコーダに届く前の PC に、ストリームのマージを行うプログラムを仕掛けてある。

これは受信バッファキューを持っており、通過しようとする RTP パケットを奪い RTP のシーケンス番号に従ってキューに蓄積し、OS のタイマによりデキューして宛先に送信する。このとき、複数のストリームにパケット損・ジッタ・順序逆転があっても、デキューされるまでに当該パケットが到着していれば、欠落なく正しい順序と RTP ヘッダが示すタイミングでキューからパケットが出て行くようになっている。

この実験で、ネットワークシミュレータを用いて、トラフィックに強制的に外乱を起こした例を示す。一方のストリームはパケット損失 10% で揺らぎの最大値が 40ms (図 7)、他方のストリームはパケット損失 1% で揺らぎの最大値が 70ms (図 8) である。これがマージされてパケット損失 0% で揺らぎの最大値が 5ms に改善された (図 9)。

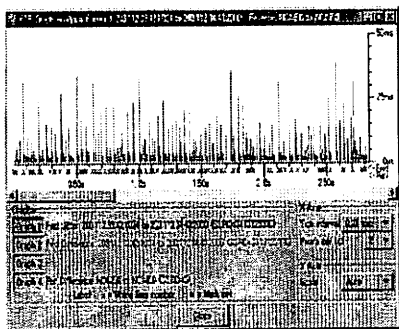


図 7 パケット損失が大きく揺らぎが大きいストリーム

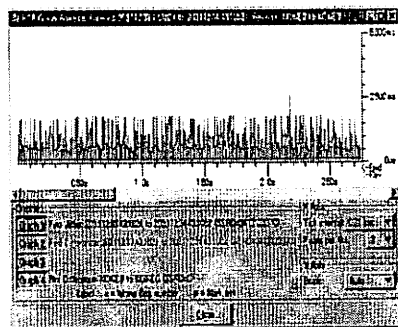


図 9 ストリームをマージした結果

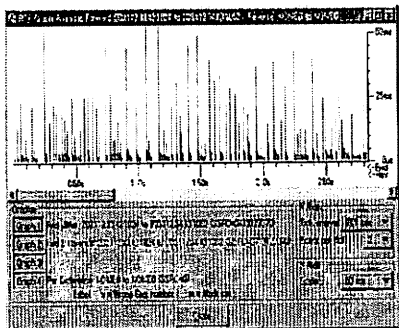


図 8 パケット損失が小さく揺らぎが大きいストリーム

5. ま と め

本研究では、トンネリングによるマルチホーム・モバイル技術を用いて、インターネットにおけるトラフィックエンジニアリングや高品質伝送の実現を目指している。インターネットトポロジーの葉に近い部分に位置している地域 ISP や、映像伝送においてインターネットによる恩恵を受けていない地域 CATV 等で応用を期待できると考えている。

実用性を持たせるための実験が不十分であり、今後は実験室レベルでの測定を増やすことや実証実験レベルでの経験が必要であると考えます。

現在の計画では、地域 IP 網によるアクセス線上で bi-cast を利用して安定な伝送品質を得る実験や、IPv4 のネットワーク上に IPv6 トラフィックを流す実験、トンネリングに用いている GRE のシーケンス番号からトンネルの品質を測定する手法

の提案等を行っていく予定である。

謝 辞

本研究を行うにあたり、DR-UR の提案者である東京大学江崎浩教授、株式会社インテック・ネットコアの中川郁夫様、永見健一様、小柏伸夫様、高知県企画振興部情報基盤課のみなさま、高知県新情報ハイウェイ運用担当 NTT コムウェアのみなさまには多くのご協力を賜りました。ここに記して感謝します。

また本研究では実験プラットフォームの一部に JGN II を利用しています (JGN2-A16056)。このたび JGN II 利用促進賞を賜りました。このような形での応援は大変な励みになります。JGN II 関係者のみなさまに御礼申し上げます。

本研究は総務省 SCOPE (戦略的情報通信研究開発推進制度 研究主体育成型研究開発 地域情報通信技術振興研究開発 受付番号 042309002) による助成を受けています。

文 献

- [1] Y. Rekhter and T. Li: "RFC 1771: A Border Gateway Protocol 4 (BGP-4)", IETF (1995).
- [2] 宇多, 小柏, 永見, 近藤, 中川, 篠田, 江崎: "多重ルーティング型マルチホームアーキテクチャの提案", 電子情報通信学会論文誌, J87-B, 10, pp. 1564-1573 (2004).
- [3] S. UDA, N. OGASHIWA, K. NAGAMI, K. KONDO, I. NAKAGAWA, Y. SHINODA and H. ESAKI: "Design and implementation of overlaying multi-homing architecture", Proceedings of SAINT2004 Workshops, IEEE Computer Society Press, pp. 165-172 (2004).
- [4] K. NAGAMI, S. UDA, N. OGASHIWA, R. WAKIKAWA, H. ESAKI and H. OHNISHI: "Multi-homing for small scale fixed network using mobile IP and NEMO", IETF Internet Draft: draft-nagami-mip6-nemo-multihome-fixed-network-03.txt (2005).
- [5] 菊池他: "多重ルーティング型マルチホームの地域 ISP における応用", 情報処理学会研究報告 2005-DSM-39, pp. 19-24 (2005).
- [6] 福家, 菊池: "多重経路ルーティング技術を用いた PA アドレスを持つ ISP のマルチホームに関する研究", 情報処理学会研究報告 2006-DSM-40, pp. 133-138 (2006).