

ユビキタスインターネットのための高位レイヤスイッチング技術の研究開発 (0211036)

Researches on upper-layer switching technology for the ubiquitous Internet

村田正幸 大阪大学大学院情報科学研究科

Masayuki Murata, Graduate School of Information Science and Technology, Osaka University

長谷川剛† 牧一之進†† 村瀬勉††† 下西英之††† 蔵杉俊康††† 小林正好†††

長谷川洋平††† 山崎康広†††

Go Hasegawa †, Ichinochin Maki † †, Tutomu Murase † † †, Hideyuki Shimonishi † † †,

Toshiyasu Kurasugi † † †, Masayoshi Kobayashi † † †, Yohei Hasegawa † † †,

and Yasuhiro Yamasaki † † †

†大阪大学サイバーメディアセンター ††大阪大学大学院情報科学研究科

†††NECシステムプラットフォーム研究所

† Cybermedia Center, Osaka University,

† † Graduate School of Information Science and Technology, Osaka University,

† † † System Platforms Research Laboratories, NEC

研究期間 平成 14 年度～平成 16 年度

概要

本研究課題では、ネットワーク内部に第 4 層以上のプロトコル処理機能を有するオーバーレイ論理ネットワークの構築を実現した。そのために、上位レイヤプロトコル (TCP 等) をネットワーク内部のルータで高速に、かつスケーラブルに処理する技術の研究開発を行った。われわれの開発したシステムにおいてはユビキタス環境における多様なエンド間データ通信の品質を現状に比して数倍向上させることができ、その結果、高い付加価値を有するルータを実現し、わが国のネットワークノード技術の発展に寄与することが出来た。

Abstract

In this project, we establish new network architecture for the future high-speed and ubiquitous Internet environment. It is an upper-layer switching network; the transport-layer overlay network on the layer-3 legacy IP network. The objectives of our project includes: providing high-speed and scalable layer-4 protocol processing architecture at the network router, mathematical analysis and evaluation of the layer-4 overlay networks. The experiment evaluation results show that our TCP proxy node can handle TCP connections at wire speed of 1Gbps link, and obtain larger than 100% throughput gain in the data transmission on the public Internet between Tokyo and Osaka.

はじめに

本研究課題においては、ネットワーク内部に第 4 層以上のプロトコル処理機能を有する論理ネットワークの構築を実現することを目的とした。提案するネットワークアーキテクチャは、第 3 層プロトコルである IP ネットワーク上に、第 4 層プロトコルである TCP を用いてオーバーレイネットワークを構築するものである。そのために、TCP コネクションをネットワーク内部のルータで高速に、かつスケーラブルに処理する技術の研究開発し、それを用いた第 4 層におけるオーバーレイネットワーク (TCP オーバレイネットワーク: 図 1 参照) の構築、評価、実験評価を行った。特に実証実験では、実装した TCP プロキシが 1Gbps の回線速度を使いされる処理速度を持つことを確認した。また、東京-大阪間の公衆ネットワーク回線を用いた実験から、データ転送速度が 3 倍程度に向上することを確認した。以下に詳細な研究項目に関する研究内容を説明する。

TCP プロキシを用いたコネクション分割・中継手法の確立

本研究課題で提案する TCP オーバレイネットワーク構築のためには、ネットワーク内のノードにおいて TCP コネクションの終端処理を行い、データ・ACK パケットの中継転送を行う。図 2 に提案する高位レイヤ (第 4 層) スwitching アーキテクチャを示す。本研究課題ではまず、このコネクション分割・中継を可能とするための、中継ノード (TCP プロキシ) における SYN・SYN/ACK パケットの処理方法、パケットバッファリング手法などについて検討を行った。その結果、送受信端末で特別な設定を必要とせず、かつネットワーク内で柔軟にコネクション分割の可否判断や分割・中継処理を行う手法を確立した。

TCP コネクション分割・中継転送手法の数学的解析手法による性能解析

TCP オーバレイネットワークにおいては、通常エンドホスト間に設定される TCP コネクションを、ネットワーク内のノード (TCP プロキシ) で終端することによって TCP コネクションを複数に分割し、分割されたコネクションごとにパケッ

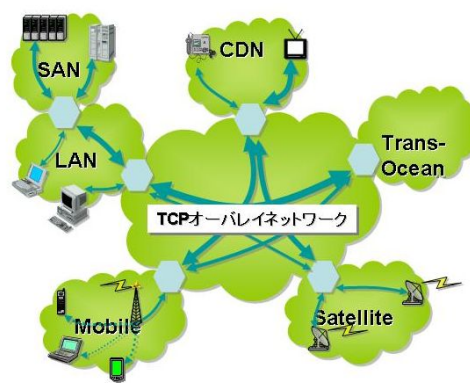


図1. TCPオーバレイネットワーク

トを中継しながら転送を行う。しかし、既存システムの変更を最小限にとどめるために、TCPの輻射制御アルゴリズムを各中継ノードにおいて独立に動作させる場合、それらが互いに干渉し、その結果期待するほどのスループットが得られないことが明らかになった。そこで本研究課題においては、この問題を考慮したエンドホスト間のスループット解析手法を示し、その妥当性をシミュレーションとの比較により検証した。



図2. 高位レイヤスイッチングアーキテクチャ

解析結果を用いることで、特にネットワーク環境が均一な場合に、TCP オーバレイネットワークによって期待されるスループット向上の度合いが、最大で55%削減されることが明らかとなった。また、そのスループット低下を防止する手法として、プロキシノードにおけるバッファリング処理の改善手法を提案し、その効果を確認した。また、効果的なTCPコネクション分割機構の導入シナリオについても議論を行った。

TCP オーバレイネットワークアーキテクチャを用いたさまざまなデータ転送手法の提案

TCP オーバレイネットワークにおいてはTCPコネクションが分割されるため、プロキシノード間のTCPコネクションに特殊なTCPを用いることで、性能向上を図ることができる。本研究課題においては、1つの手法としてマルチパス方式を提案した。これは、TCP オーバレイネットワークにおいて分割されたTCPコネクションのうち、TCPプロキシ間のコネクションを複数本設定してデータ転送を行うことにより、スループット、耐障害性の向上を図る手法である。提案手法をシミュレーションおよび試作機を用いた実験を通じて評価した結果、従来のIP層におけるマルチパス方式に比べて、2倍以上のスループット向上が可能であることがわかった。

また、プロキシ間のTCPコネクションに、HSTCPなどの高速TCPと呼ばれる手法を適用することで、プロキシ間のネットワークが多様な帯域遅延を持ち、従来のTCPでは十分なスループットが得られないような環境においても、高いスループットを得ることができると確認した。これらさまざまなデータ転送手法は、エンド端末である送受信端末を変更することなく実現可能であることが大きな特長である。

TCPプロキシノード・TCPプロキシ機構の実ネットワークにおける実装実験評価

提案手法が実ネットワーク環境においても有効であることを示すために、本研究課題においては、TCPプロキシ機構を実コンピュータ上に実装し、その性能評価を実験ネットワーク、および実ネットワークを用いた実証実験を通じて行った。実装においては、処理速度を考慮し、カーネルシステム内でコネクション中継処理を行う方式を提案した。提案手法は、アプリケーションプログラムの介在なしにTCPプロキシ処理を行うため、ネットワークノード、端末において問題となるメモリコピー処理のオーバーヘッドを大きく削減することが可能である。また、パケットのチェックサム処理をできるだけ軽減する手法も提案した。その結果、提案手法は回線の遅延時間にかかわらず、最大約1.4Gbpsの速度でTCP中継

処理を行うことが可能であることが明らかとなった。本研究ではさらに、実装を行ったTCPプロキシ機構を用いて、その上に我々が提案する高速かつTCP-Renoとの親和性の高いTCP-Adaptive Reno、およびgentle HSTCPの実装を行った。また、このTCPプロキシを実験室環境および、東京-大阪間のインターネットを利用した実ネットワークで評価した。その結果、本中継器を用いることで、実験ネットワーク環境では最大3倍程度、東京-大阪間の公衆インターネット環境では2倍以上のスループットが得られた。また、本提案機構によって中継されたTCPコネクションは、中継を行わない既存TCPコネクションと公平に帯域を共有することも強調すべき点である。なお、本実験において実装したTCPプロキシ機構の一部に関しては、NECが1-2年後を目処に製品化を目指

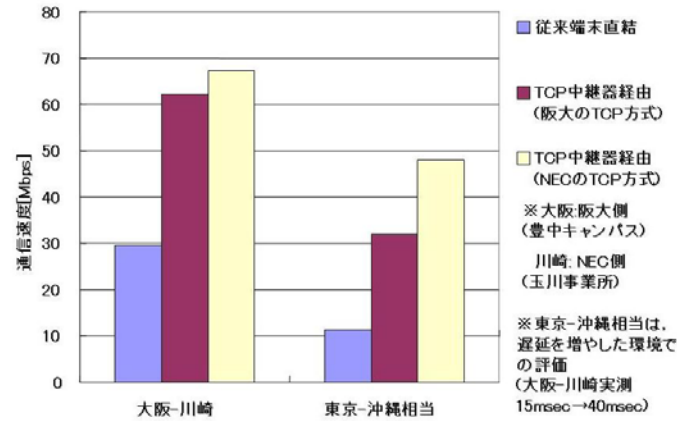


図3. 公衆インターネット回線を用いた実験結果

しており、実装に必要な細部の検討を開始している。

申請特許リスト

- [1] 長谷川洋平、下西 英之、山崎 康広、村瀬 勉、「セッション中継装置、セッション中継方法」、日本、2003年09月05日
- [2] 長谷川洋平、村瀬 勉、「通信装置およびその通信方法ならびにプログラム」、日本、2003年10月22日
- [3] 山崎康広、「セッション中継装置、セッション中継方法及びセッション中継プログラム」、日本、2003年12月03日他1申請

報道発表リスト

- [1] “光回線・ADSL 性能引き出し速度最大10倍 阪大とNEC、家庭向け新技術”、日本経済新聞 2005年4月15日付朝刊1面