

## 論文の内容の要旨

論文題目 ネットワーク内処理によるパケット通信の高速・高信頼化に関する研究  
氏 名 長 谷 川 輝 之

第 1 章では、本研究の背景と概要を述べている。近年 IP パケット通信は急速に利用範囲を拡大し、WWW 等の従来のインターネットアクセスに加え、携帯電話でのデータ通信や家電製品間の相互接続、更には音声通話や放送配信への適用等、日常生活を支えるインフラ構成要素として必要不可欠な存在となっている。IP ネットワークのインフラ化に伴いノードの置換や実装変更は益々困難となっているが、ユーザが通信に期待するスループット・信頼性への要求は高まる一方である。本論文は、現在の IP ネットワークを構成する膨大な数のノードを一斉に変更することが困難であるとの認識に基づき、ネットワークを構築・運用するキャリアの視点で、ネットワーク内のボトルネック箇所に高度なパケット処理を行う付加装置を徐々に導入し、既設ノード実装を一切変更せず段階的にパケット通信の高速・高信頼化を実現する手法を述べその実用性を明らかにした。

第 2 章では、TCP/IP の概要説明と共に IP ネットワークの現状を俯瞰し、キャリアの視点から高速・高信頼化に向けた提案と課題抽出を行った。IP ネットワークはルータの機能を簡素化し個々の通信(フロー)に関与しないことでスケラビリティを保持し規模を拡大してきたが、今や膨大な数となったホストの存在自身が新機能展開を阻害している点を指摘し、手法検討に当たり以下の要求条件を挙げた。

- 既設ノードの実装を変更しない
- 一部導入時も効果が得られる
- 伝送媒体・ネットワーク構成に応じてフローを意識した制御が可能

これに対して、ネットワークへの設備設置やボトルネック把握が可能なキャリアの優位性を活かし、ネットワーク内に導入した付加装置により高位プロトコル処理を行う手法を提案し、実現に向けた制約条件と課題を以下のように整理した。

- ネットワーク内でフローを意識した処理を行うため、設置場所をフロー数が少ないネットワークエッジ近傍に限定する。
- 上記制約の中で、適用対象を将来に亘って主要なアプリケーションであることが見込まれる TCP 通信の高速化と IP マルチキャストの高信頼化に絞り具体的方式を確立する。
- 少数の装置からの段階的導入実現のため、汎用プラットフォームによるギガビットレート対応の実装技術を確立する。

第 3 章では、近年のキャリアへのギガビットイーサネット普及を踏まえ、付加装置実現のコア技術であるパケットモニタの高性能化検討を通じて、PC と NIC で構成された汎用プラットフォームで 1Gbps の処理が実現可能であることを示した。2001 年に開発した手法は、DMA 転送機能を持つ NIC を対象にソフトウェアのみでこれを実現している。主な特徴を以下に示す。

- NIC デバイスドライバとして動作する。DMA 転送先を工夫することで CPU の介在無しに複数

- 複数ヘッダ情報を UDP パケットにまとめ、自身又は他 PC へ転送する機能を持つ。本機能によりヘッダ抽出と解析を複数 PC で負荷分散できる。

性能評価の結果、1GHz×2 CPU の PC を抽出に用いた構成(他 PC へ転送)で、インターネットを想定した 384 バイトフレームでフルレート(約 300K パケット/秒)収集を約 30%の CPU 負荷で達成した。これに対して既存方式は、30-40%の収集率で CPU 負荷が 70-80%に達しており、大幅な性能向上が見られた。メモリコピーの省略やカーネル内処理による高速化は広く検討されている技術であるが、更に本研究では DMA 転送を工夫しメモリコピーを削減する等、パケットモニタに適切に応用し実現性を高めた。

第 4 章では「高位プロトコル処理を行う付加装置」の応用例として国際 LAN 間接続の高速化を検討した。1990 年代の国際網広帯域化に伴い、TCP のフロー・再送・輻輳制御がボトルネックとなりスループットが劣化する問題が顕在化した。本研究ではこれにいち早く着目し、国際網に接続する LAN エッジルータとして TCP までの処理を行う付加装置(TCP ゲートウェイ)を設置する手法を提案した。主な特徴を以下に示す。

- ゲートウェイ間に、伝送遅延に対して十分なウィンドウサイズを持ち広域網に適したトランスポートプロトコルを導入し、TCP と本プロトコル間の変換を行う。
- TCP コネクション確立・解放をホスト間でエンド・エンドに行う一方、送達確認やフロー制御を、ホスト・ゲートウェイ間、ゲートウェイ間、ゲートウェイ・ホスト間の 3 セグメントで独立、且つエンドホストからは透過的に提供する。

これにより既設ホストはゲートウェイの存在を意識せずにスループットが向上する。1995 年の初期実装は、数 100msec の往復遅延がある ATM 網を介した TCP スループットを数 100Kbps から最大 7.3Mbps に改善した。更に 1997 年の商用実装では、データリンク層実装による処理削減や選択再送機能導入による改善により、当時のワークステーション性能で 45Mbps(DS3)回線に対応した実用的な処理性能を実現した。1998 年頃から同様の商用製品も登場し現在は一般化した技術となったが、本研究はキャリアでの実用化に先鞭をつける取組みであると考ええる。

第 5 章では、衛星インターネットアクセスにおける TCP 通信の高速化について、加入者毎に付加装置(ゲートウェイ)を展開・運用することが困難であるとの認識に基づき、前出の 3 セグメント手法に代わる簡易な手法を検討した。具体的には、想定される通信の殆どがサーバからのダウンロード(下り)方向であることに着目し、衛星アクセス回線のキャリア側にのみ付加装置を設置し、複数加入者に対して下り方向のスループットを向上する手法を提案した。本手法は TCP コネクションを衛星区間とそれ以外に 2 分割し送達確認・フロー制御を行うため 2 セグメント手法と呼び、1998 年に初期実装を完了し 2001 年に改良を行った。主な特徴を以下に示す。

- 3 セグメント手法と同様にデータパケット(DT)を TCP 手順に従い予めサーバから引き出す。クライアントへ広告ウィンドウを超え投機的に DT を先送りし連続受信させる。
- DT 紛失時の先送り DT 大量破棄や帯域非対称性による上り回線輻輳を防ぐため、TCP-Vegas

- クライアントが継続的なウィンドウ閉塞状態であることを検知し、不要な先送りや付随する再送発生を防ぐ。

実装評価により、上記機能を実装した改良 TCP ゲートウェイが DT 紛失前にプロアクティブに送信量を調整し、スループットを 8 倍程度まで改善可能であることを確認した。本方式は、KDDI での実用化に加え地方自治体向け衛星ネットワークの標準仕様として全国各地に導入されている。また、ITU-R での標準化も完了している。対向側が既存 TCP のため性能・機能面で制約があるが、加入者毎の装置設置が不要であり導入・運用コストの面で実用的な手法である。

第 6 章では、上り・下り回線帯域が非対称な環境でのスループット向上のため、2 セグメント手法に基づく手法を提案した。非対称性は衛星回線以外にも 1xEV-DO や ADSL 等近年広く見られる。TCP は ACK 返送により DT 送信とは逆方向に帯域を消費するため、十分なスループットを得るには ACK 帯域の削減も必要となる。そこで、キャリア側に設置した付加装置(プロキシ)で加入者ホスト向けに最大パケットサイズを強制的に拡大した DT(オーバサイズ DT)を利用することで、ホストを変更すること無く ACK 数を削減する手法を提案した。主な特徴を以下に示す。

- 2 セグメント方式と同様にサーバ・加入者ホスト間の TCP コネクションを 2 分割し、ホストに送信する DT を蓄積する。
- 上記 DT を、ホストが想定する最大パケットサイズを超えるオーバサイズ DT に再構成して送信し、対応する ACK 数を削減する。
- 上記オーバサイズ DT は、下り方向の path MTU サイズに合わせ付加装置の IP 層で強制的に分割する。

実装評価の結果、検証した主要 OS 全ての正常動作を確認し、10MB の HTTP ファイルダウンロードで最大 3 倍のスループット向上効果を確認した。IP 分割による付加装置・ホストのオーバヘッド増加は殆ど無く、実用性のある手法であることも示した。

第 7 章では「高位プロトコル処理を行う付加装置」を用いた別の応用例として、今後急速な普及が見込まれるキャリアの IP マルチキャストネットワークを用いた放送配信の高信頼化手法を検討した。その主な特徴は以下の通りである。

- 付加装置をネットワーク両端のエッジルータ近傍に導入して放送配信を常時受信監視することで疎通断を高速に検出する。
- 付加装置は断検出時のみ 2 種類の予備マルチキャストを受信し、ルータのマルチキャスト転送テーブルと整合を取りつつホストへ代理配信する。
- 一方の予備マルチキャストには送信側で予め遅延を挿入する。これにより、付加装置で障害発生時点に遡って配信可能とする。
- キャリアネットワークの冗長構成を踏まえ、既存ルータの持つ機能のみで、障害点を回避した予備マルチキャスト木の構築が可能となる。

併せて、配信サーバのマルチキャスト送信タイミング平滑化のために、配信フォーマット非依存の送信制御手法を考案・実装した。実装システムによる HDTV 品質の映像配信実験により、提案手法がトラヒ

ックバーストを抑制しつつルータでのネットワーク障害検出・回復に伴う約 27,500 パケットの紛失を補償し、無瞬断でサービス提供できることを示した。関連研究動向としては IETF MPLS WG で現在標準化が進められている P2MP MPLS TE や Cisco 社の dual stream 技術等があるが、既設ノード実装変更の必要が無い点では提案手法が優れている。

第8章では、本研究の結論として以下を述べた。

- キャリアの視点から、ネットワーク内で高位プロトコル処理を行う付加装置を導入し、現状のIPネットワークで通信品質を段階的に改善する手法を提案した。
- 応用例と高性能実装技術を確立することでその実用性を明らかにした。

提案手法は、既存インフラを変更することなくネットワークに機能を段階的に追加可能という特長により、常に品質向上・差別化を求められるキャリアにおいて重要性は益々高まると考える。また、今後の課題として以下を提示した。

- ユーザエッジ機器への適用を想定した実装技術と応用例の開発
- 標準技術との効果的なインターワークや段階的な移行手法の確立