

高遅延回線における TCP 速度性能劣化の改善

柳生理子* 田中 功一*

*三菱電機(株)情報技術総合研究所

衛星回線などを介したTCP通信における伝送遅延による速度性能劣化の改善を目的とし、高遅延回線前のルータによる送達確認応答(ACK)の代理応答を用いたアルゴリズムを開発した。シミュレーションによる性能評価の結果、本アルゴリズムを適用する事により、回線速度の80%以上のバンド幅の利用が可能となる事が分かった。

IMPROVEMENT IN THE LACK OF TCP PERFORMANCE USING LARGE DELAY BANDWIDTH PRODUCTS

Riko Yagiu * Kouichi Tanaka*

*Mitsubishi Electric Corp. Information Technology R & D Center

We developed an algorithm which gives a solution for a problem that the performance of TCP using large delay-bandwidth products is so poor by delay of ACKs. Proposed Algorithm is for a router that is placed just before large-delay communication link. The router acts as the receiver and send back ACKs to the sender. Our performance simulation shows that the algorithm provides more than 80% bandwidth utilization.

1 背景

近年、大容量のデータに対し、高速かつ確実な通信への要求が高まってきている。これに伴い衛星回線上でTCPを用いた大容量のデータに対する高速・高信頼のサービスが、始まってきている。しかし、衛星回線は高速かつ大容量の通信網である長所をもつ一方で、物理的伝送距離の長さゆえの伝送遅延がある。この遅延は、TCPの輻輳制御機構との間に、以下に示す二つの相互作用をもたらす。

・衛星回線の伝送遅延の影響により、TCP送信側においてセグメントが最終宛先に到達したかどうかを判別するまでに、長い時間を要し、これはセグメントの再

送¹を促す。

・TCPの輻輳制御機構は、「セグメントの再送の頻発は、回線の輻輳の信号である」との解釈に基づき、輻輳を緩和させるために送信量を抑制する

この結果、ネットワークが輻輳していないにもかかわらず、TCPは不必要にその転送速度を抑制し、幾つかのTCP輻輳制御アルゴリズムおよびtelnetなどの対話型のアプリケーションに対し、致命的な影響を及ぼす。このため、衛星回線など遅延の大きな回線を介したTCP通信における速度性能の劣化の改善が求められている。

¹ 衛星回線ではエラー率の高さが更にこれを助長。

2 衛星遅延による速度性能の劣化への取り組み

2.1 通信衛星利用共有データ配信システム

上述の問題に対し、我々は図1に示した構造をもつ通信衛星利用共有データ配信システム (SDDS: Satellite Data Distribution System) の開発を行ってきた[1][2]。SDDSは、アプリケーション層に実装されたS/Wであり、データ送信は衛星回線上でUDPを、送達確認は地上回線上で独自アプリケーションを用いている。

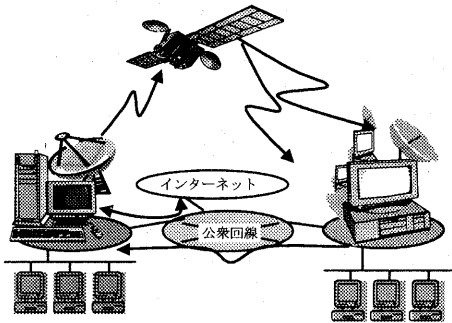


図1 SDDSの構成

SDDSは、TCPと同等の高い信頼性とIP透過性の確保、および速度性能の劣化の改善を可能とする一方で、一般的なTCP上のアプリケーションおよび双方向の衛星回線を用いた配信の性能劣化に対しては適用できないという課題があった。

2.2 TCP Spoofing アルゴリズムの開発

上述の課題のために、我々は、通信系路上にSpoofingと呼ばれる特殊な機能を備えたルータを導入するアプローチを採り、このSpoofing機能を実現するS/W(TCP Spoofing S/W)のアルゴリズムの研究開発を行った[3][4]。

図2は、TCP Spoofing アルゴリズムを用いた仕組みと、その特徴および働きを

示したものである。

Spoofingとは、図に示すように衛星区間の手前に設置されたルータにおいて、パケットを転送すると共に、受信側TCPにおいて転送パケットを受信した場合に作成する送達確認応答(ACK)と同等のもの(擬似ACK)を作成し、送信側TCPに対し、返送(ACKの代理送信)する機能[5][6]である。このACKの代理送信機能により、送信元のTCPは宛先にまでセグメントが到達したとSpoof(偽証)され、送信Windowを移動し連続的なデータ送信が可能となる。この手法は通信クライアントのS/Wへの変更が必要ないことから、ユーザから処理を隠蔽し、一般的なTCP上のアプリケーションを用いた大容量データに対する高速通信を可能とするものである。

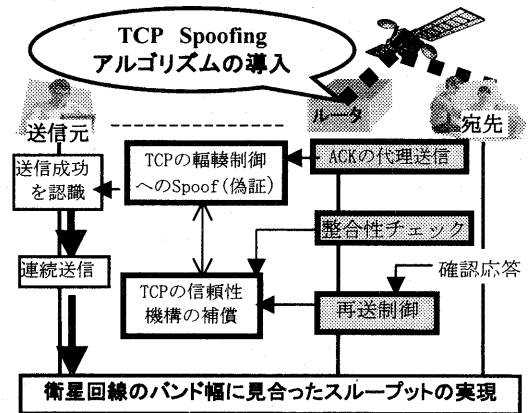


図2 TCP Spoofing アルゴリズム

2.2.1 アルゴリズムの特徴

TCP Spoofing S/W、以下に示す2つの機能を持つ。

- ① TCP上のアプリケーションに対するスループット低下の改善

送達確認の代理送信:

- ・宛先におけるACKと同等のセグメントを作成し、送信元に対し返送

② TCP 通信における信頼性の保障

通信結果の矛盾の回避:

- ・ 送達確認を代理送信したデータの保持
- ・ 受信確認および再送制御送信元 - 受信側間における配信結果の整合性チェック

このように、TCP Spoofing アルゴリズムは、TCP の輻輳制御の緩和によりスループットの減少の改善を促し、またこれにより TCP の信頼性の劣化を起ささないために、再送制御を行うものである。そこで、次に、TCP の輻輳制御とスループットについて述べる。

2.2.2 TCP の輻輳制御とスループット

通信開始時、ネットワークおよびは通信先の状況は不明であるため、TCP は 3-way-hand shake という互いの状況を確認しあうコネクションの確立を行う。

コネクション確立後は、ネットワークの輻輳状況を確認めながらスループットを上げることがを目的とし、輻輳制御 Window (cwnd) を用いて、与えられた Window サイズ (sssthresh) に達するまでの間、ACK の受信により送信量を増やす輻輳制御アルゴリズム (スロースタートと輻輳回避) を用いる。スロースタートでは、1ACK 受信毎にセグメントサイズ (segsz) 分、輻輳回避では 1ACK 受信毎に、 $segsz * segsz / cwnd$ 分、cwnd を増加させる。従って常に送信 Window がスライド可能な状況ではない場合、スループットは、ACK 受信毎の cwnd の増加量、送信 Window 内における送信済みデータ量および確認済みデータ量によって定まる。

エラーが発生し、再送タイムアウト (RTO) または Duplicate ACK のいずれかを検知した場合、TCP は再送を行う。この時まず、

現在の Window サイズ²の 1/2 のサイズに sssthresh を設定する。タイムアウトによる輻輳の検知の場合には、さらに cwnd を 1 セグメントに設定する。即ち、タイムアウトの発生イベントを受けた場合は、スロースタートが働き、duplicate ACK の受信をした場合は、輻輳回避アルゴリズムが働く。

2.2.3 エラー時の TCP の動き

以上のことから、TCP Spoofing アルゴリズムを導入した環境において、エラーが発生した場合、次のことがいえる。

- ・ 送信元ルータ間におけるエラーが発生した場合、検知方法の違いは、スループットに大きく影響し、特に cwnd が小さい区間においてはさらに顕著に表れる。
- ・ 一方、ルータ転送後、即ち衛星区間および受信側 TCP におけるエラーが発生した場合、ルータの再送タイムアウト値は、衛星区間の RTT より大きいという条件を満たす範囲において、ルータの受信バッファのサイズに応じて決定するものとした。従って、衛星区間におけるエラーの発生による速度への影響は、エラーの検知方法ではなく、ルータおよび受信側ホストの受信バッファ (Window) の大きさに、より多く依存する。

3 シミュレーションによる性能評価

今回、アルゴリズムのボトルネックとなる部分の発見と検証を目的とし、数値計算による速度性能を中心としたシミュレーションおよびそれに基づく性能評価を行うこととした。

² $\text{Max}(\text{Min}(\text{cwnd}, \text{受信側から告知されたウィンドウ}), 2 \text{セグメント})$

3.1 シミュレーションの方針

先に述べた TCP の輻輳制御を考慮し、以下を主方針としシミュレーションを行った。

- ACK 受信ごとに表 1 に示す評価データを累積。
- 速度性能に関しては、接続確立後、最初のデータセグメント送信から接続切断のためのパケット送信直前までの区間における、送信側 TCP の平均スループットを算出。
- 上記に関連し、データ送信開始時、送信元ルータ間におけるエラーによる Duplicate ACK 検知時、送信元ルータ間におけるエラーによる RT0 発生時を中心に検証。
- 比較対象として、無遅延回線上での通常 TCP 通信時および TO 発生時のスループットを設定。
- 衛星区間におけるエラー発生時のルータのバッファの制約条件は本シミュレーションの検証対象外。

データ項目	定義
ACK 受信数	送信元が受信した ACK 或いは擬似 ACK 数
cwnd	送信元の ACK 受信後の cwnd のセグメント数
送信総データ	送信元が最初の ACK 受信後から該当 ACK 受信後、送信可能なデータを送信終了した時点までの、送信した総バイト数
経過時間	通信確立後、最初のデータ送信から該当 ACK 受信後、送信可能なデータを送信終了した時点までの、経過時間

表 1

3.2 シミュレーション方法

前述の検証対象および比較対象の、11 セグメント目のエラーによる Duplicate

ACK 検知時、RTO 発生時について、接続確立後のデータ通信開始区間（開始時スロースタート区間）、TCP Spoofing アルゴリズム使用区間に分けて計算を行い、表 1 の記述データを算出した。

3.2.1 各区間の算出式

以下に送信元ルータ間の遅延を α 、衛星区間の遅延を β 、データセグメントの連続送信の際の 1 セグメント辺りの送信時間を γ 、ヘッダのみのセグメント (SYN や ACK などの情報) の連続送信の際の 1 セグメント当たりの送信時間を ε とした場合の、各区間の算出式を示す。

① 開始時スロースタート区間

この区間においては、最初の送信データに対する ACK が戻ってくる前に、タイムアウトが発生し、データ (1) の再送を行う。従って、 $\varepsilon + T0$ 時間を要する。

② TCP Spoofing 区間

ACK (実際には Spoofing S/W 作成の擬似 ACK) を一つ受け取り、Window を一つスライドさせ、データを増加 Window+1 セグメント分送信する。通信確立後からの所要時間は、 $2\alpha + 2\beta + \varepsilon$ となる。この区間においては、エラーがない限り、連続送信が行われるため、一つデータセグメントを送信し、ACK を受け取るまで γ 時間を要する。

3.2.2 想定環境

次に、前述の式に対し想定環境 (データを当てはめ、データの算出を行った。今回は通信ホスト間の MTU を 1500Byte、通信区間における回線の速度は全て 1 Mbps、衛星回線の RTT を 540msec、衛星以外の区間における RTT (遅延および処理) を 0、送信元 TCP の RTO を 500msec、通信ホストの最大 Window サイズは 64KByte、TCP のスロースタート開始時の Window サイズが 1 との条

件を仮定した。

4 シミュレーション結果

前述のシミュレーションにより、算出したデータをグラフ化し、評価を行った。

4.1 ACK 受信数と cwnd

図 3 は、データ送信量と cwnd(セグメント数)の推移の関連を表すグラフである。送信

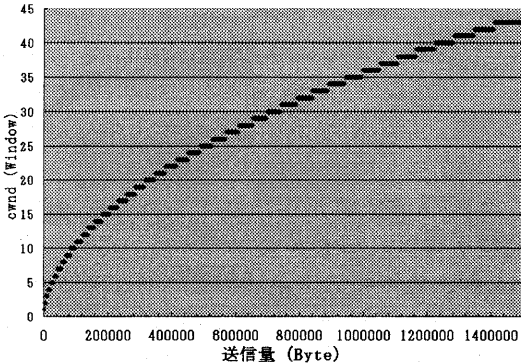


図 3 データ送信量と cwnd

はセグメント単位で行われるため、階段状の緩やかなカーブを描き、最大サイズ 43 ($44 > 65535 \div 1500 = 43.69 > 43$) に達するまで増加する。これ以降、タイムアウト発生によるスロースタートアルゴリズムの起動がない限り、送信はこの Window サイズで行われる。

4.2 正常時のスループット

図 4 は、エラーが発生しなかった場合の、

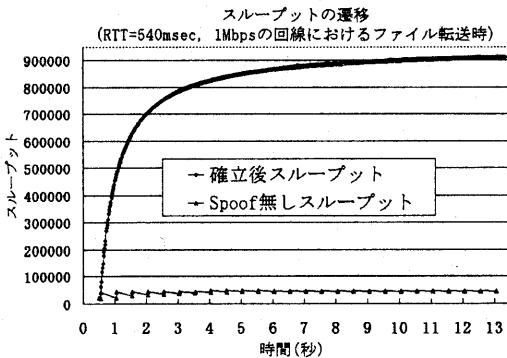


図 4

衛星回線上の TCP 通信において、TCP Spoofing を適用した場合と、しなかった場合を比較したグラフである。グラフより、1M ファイル転送時に、適用しなかった場合のスループットが 5%未満であるのに対し、適用した場合は 80%以上をのスループットを可能とすることが分かる。

4.3 エラー発生時のスループット

図 5 は、以下の場合におけるスループットの遷移を表すものである。

- ・ TCP Spoofing の適用を行い、エラーが発生しなかった場合 (正常時)
- ・ TCP Spoofing の適用を行い、エラーが発生し送信元が Duplicate ACK の検出により再送を行った場合 (Duplicate ACK 検出時)
- ・ TCP Spoofing の適用を行い、エラーが発生し送信元のタイムアウトにより再送を行った場合 (TO 発生時)
- ・ 遅延のない環境で通常の (TCP Spoofing アルゴリズムを用いない) TCP のファイル転送を行った際にエラーが発生し送信元のタイムアウトにより再送を行った場合 (遅延無し TCP の TO 発生時)

グラフより、正常時および Duplicate ACK

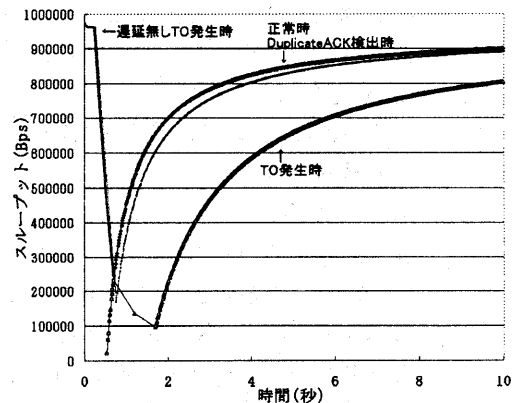


図 5 エラー検知方法によるスループットの比較

検出時と比べ TCP Spoofing アルゴリズム使用時にタイムアウトが発生した場合、最

大約 50%性能値が下がり、また回復も遅くなること分かる。

5 まとめと今後の課題

我々は、伝送遅延の大きな回線における TCP/IP 通信の高速化を目的とした TCP Spoofing アルゴリズムの研究を行っている。今回、数値計算によるシミュレーションと性能の評価を行った。性能評価の結果として以下のことが得られた。

- 通信開始時の輻輳制御 Window の増加率は通常の TCP 通信よりも緩やかであるが、それ自体は速度性能に対し大きく影響するものではない。エラーの発生がない場合は、1Mbps の回線上で 1MByte のファイル転送時には、通常の TCP 通信の 15 倍以上、遅延のない環境の 8 割以上のスループットを実現。
- 送信側とルータの間におけるエラー発生時も、送信側 TCP のタイムアウトが発生しない場合は、正常時とほぼ同等の性能を実現。
- 送信側 TCP において Duplicate ACK の受信により再送が行われた場合に比べ、タイムアウト発生により再送が行われた場合は、大きく速度性能が下がる。従って、タイムアウト発生時のアルゴリズムの改善が今後の課題となる。
- 過剰な速度での衛星回線への送信は、受信先のバッファオーバーフローを招き、衛星回線上での再送の多発による速度性能の低下をもたらす危険がある。従って、衛星回線における輻輳制御の改善とあわせて、バッファの管理方式の改善が今後の課題である。

これらの結果を踏まえ、今後、タイムア

ウト発生時のアルゴリズム、下り方向の速度性能および輻輳制御の改善を行ってゆく予定である。

Reference

- [1] 秋山他：衛星利用共有型情報配信システムの実装と評価、情処学会マルチメディア通信と分散処理研究会 No. 91, 1999
- [2] 柳生他：衛星利用共有型情報配信システムの設計と実装、情処学会マルチメディア通信と分散処理研究会 No. 96, 2000
- [3] 柳生他：伝送遅延の大きな回線における TCP を用いた通信に関する一考察、情処学会第 61 回全国大会 2H-02, 2000
- [4] 柳生他：高伝送遅延回線における速度性能劣化の改善、電子情報通信学会 2001 年 総合大会, B-8-14
- [5] Y. Zhang. "Satellite Communications in the Global Internet: Issues, Pitfalls and Potential". Proceeding of INET'97
- [6] T. R. Henderson. "Transport Protocol for Internet-Compatible Satellite Networks". IEEE Journal of Selected Areas of Communication, vol. 17, no. 2, pp. 345-359, February 99.