

[特別講演] インターネットにおける トランスポート層アーキテクチャに関する研究

長谷川 剛[†]

[†] 大阪大学 サイバーメディアセンター 〒560-0043 大阪府豊中市待兼山町 1-32

E-mail: hasegawa@cmc.osaka-u.ac.jp

あらまし Transmission Control Protocol (TCP) はインターネットにおいて最も頻繁に用いられるトランスポート層プロトコルであり、近年増加している P2P トラヒックや動画像ストリーミングなどのアプリケーションのほとんどが利用している。そのため、TCP の性能、特に輻輳制御機構がネットワーク性能に与えるインパクトは極めて大きい。本報告では、まず TCP の輻輳制御機構の概要をまとめ、これまで指摘されてきた問題点を整理し、それに対する解決方法の概略を紹介する。さらに、筆者の研究グループにおける近年の研究例である、TCP の挙動変更によって無線 LAN 経由のデータ転送を省電力化する省電力 TCP 手法に関して述べる。

キーワード Transmission Control Protocol (TCP), 輻輳制御, プロトコル階層構造, トランスポート層

[Special Talk] Research histories and trends on transport-layer protocols in the Internet

Go HASEGAWA[†]

[†] Cybermedia Center, Osaka University 1-32, Machikaneyama-cho, Toyonaka, Osaka, 560-0043 Japan

E-mail: hasegawa@cmc.osaka-u.ac.jp

Abstract Transmission Control Protocol (TCP) is the most important transport-layer protocols in the current Internet. Its performance, especially the performance of congestion control mechanisms, plays an important role since most of current Internet applications, including Web, P2P file sharing, VoD streaming services, utilize TCP. In this report, we first summarize congestion control mechanisms of TCP, and present recent researches on its problems. We also introduce our latest research topic on energy-efficient TCP, which tries to decrease power consumption of data transmission over Wireless LAN by modifying TCP's data transmission behaviors.

Key words Transmission Control Protocol (TCP), Congestion control, Protocol layering structure, Transport layer

1. はじめに

Transmission Control Protocol (TCP) [1] は現在のインターネットにおけるトランスポート層プロトコルとして最も多くのネットワークアプリケーションが利用しており、TCP トラヒックは現在のインターネットトラヒックの大部分を占めている [2]。また、インターネットは様々な種類のネットワークを取り込むことによって大規模化し、指数関数的な拡大を続けている [3]。その結果、TCP が誕生した 1970 年代当初には想定することができなかったネットワーク環境が発生している。

TCP において最も重要な機能はネットワーク輻輳を回避・検知・解消する輻輳制御機構 [4] である。多様化するネットワーク環境において安定的に通信を行なうことができることは、現在の TCP の最も重要な性質であるが、ネットワーク環境の変

化にともない、様々な問題が指摘され、解決方法が模索され続けている。

そこで本報告においては、まず、TCP の輻輳制御機構の概要をまとめ、これまで指摘されてきた問題点を整理し、それに対する解決方法の概略を紹介する。合わせて、筆者の研究グループにおける研究例についても紹介する。さらに、筆者の研究グループにおける近年の研究例である、TCP の挙動の変更によって無線 LAN 経由のデータ転送を省電力化する、省電力 TCP 手法に関する研究内容を述べる。

以下、2. 章において、TCP の輻輳制御機構の概要、およびその問題点について触れ、我々の研究グループにおける研究例を紹介する。3. 章においては、省電力 TCP 方式に関する研究内容の概略を紹介する。最後に 4. 章で本稿のまとめと今後の課題について述べる。

2. TCPの輻輳制御機構

2.1 TCPの輻輳制御機構の概略

TCPはウィンドウフロー型輻輳制御機構を用いて輻輳制御を行う。すなわち、送受信端末のTCPがウィンドウサイズと呼ばれる、確認応答無しに送信することができるデータ量を管理し、ネットワークに送出するデータ量を調節する。特に、送信側端末が持つ輻輳ウィンドウサイズは、ネットワークの輻輳状況に応じて増減されるため、インターネットにおける輻輳制御の重要な役割を果たしている。以下では、代表的なTCPのバージョンであるTCP Renoの送信側端末における、輻輳ウィンドウサイズの更新方法の概略を述べる。

TCP Renoの輻輳制御方式は、スロースタートフェーズおよび輻輳回避フェーズと呼ばれる2つのフェーズから構成され、それぞれにおいて輻輳ウィンドウサイズ(cwnd)の増加速度が異なる。スロースタートフェーズにおいては、1つのACKパケットを受信するごとに輻輳ウィンドウサイズを1パケット増加させる。一方、輻輳回避フェーズにおいては、1つのACKパケットを受信するごとに輻輳ウィンドウサイズをその逆数分だけ増加させる。すなわち、TCP Renoの輻輳ウィンドウサイズを w_{reno} とすると、その更新アルゴリズムは以下のように表すことができる。

$$w_{\text{reno}} \leftarrow \begin{cases} w_{\text{reno}} + 1 & (w_{\text{reno}} < s_{\text{reno}}) \\ w_{\text{reno}} + \frac{1}{w_{\text{reno}}} & (w_{\text{reno}} \geq s_{\text{reno}}) \end{cases} \quad (1)$$

ここで、 s_{reno} は、TCP Renoがスロースタートフェーズから輻輳回避フェーズに移行する時のしきい値 $ssthresh$ である。送信側端末においてACKパケットを受信する度に上式が用いられることによって、 $w_{\text{reno}} < s_{\text{reno}}$ の場合にはRTTごとにcwndが2倍になり、 $w_{\text{reno}} \geq s_{\text{reno}}$ の場合にはRTTごとにcwndが1だけ増加する。

一方、パケット廃棄を検出した場合には、次式のように輻輳ウィンドウサイズを減少させる。パケット廃棄検出が重複ACKによって行われたか、タイムアウトによって行われたかによって挙動が異なる。

$$w_{\text{reno}} \leftarrow \begin{cases} w_{\text{reno}}/2 & (\text{重複 ACK}) \\ 1 & (\text{タイムアウト}) \end{cases} \quad (2)$$

すなわち、TCP Renoはパケット廃棄を検出するまで輻輳ウィンドウサイズを増加させ続け、パケット廃棄をきっかけに減少させる。これは、TCP Renoがパケット廃棄の発生をネットワーク輻輳の指標と見なしていることに起因する。本論文ではこのようにパケット廃棄をネットワーク輻輳の指標として用いる手法をloss-based手法と呼ぶ。

2.2 TCP輻輳制御機構の問題点

インターネットのプロトコルは階層構造を持ち、おおよそ5層(物理層、データリンク層、ネットワーク層、トランスポート層、アプリケーション層)に区別されることが一般的である。それぞれの階層では、下位層から提供される機能を基に、様々な機能を上位層に対して提供する。また、本稿において着目しているトランスポート層プロトコルは、一般的にはネットワーク内では動作せず、エンド端末においてのみ動作する。これに

より、インターネットを構成するネットワークはネットワーク層プロトコルであるInternet Protocol (IP)のパケットを転送することに集中することができる。

プロトコル階層構造は、階層毎の入れ替えや改良が容易である点において優れている。特にインターネットは、砂時計モデルと呼ばれる構造を持つ。これは、ネットワーク層プロトコルであるIPを中心とし、より上位層のプロトコルはIPに対応することでネットワークサービスをインターネット上で展開することが可能となり、より下位層のプロトコルはIPに対応することで、そのプロトコルによって実現されるネットワークをインターネットに接続することができるものである。

一方、プロトコル階層構造は各階層の実装を他の階層から隠蔽するため、機能の重複や予期せぬインタラクションによって性能向上が阻まれることがある。以下ではその例をいくつか挙げ、我々の研究グループにおける研究例を紹介する。

2.2.1 輻輳制御機構の重複

Asynchronous Transfer Mode (ATM)のような、下位層においてトラヒック制御を行うようなプロトコルが存在する場合、上位層のTCP輻輳制御機構と機能が重複し、いわゆる屋上屋を架す状態となり、性能低下を招く。同様の問題はSSL-VPNなどで見られるTCP over TCP問題でも確認することができる。これは、エンド端末間パスの途中にアプリケーション層プロトコルによるVPNが構築されている場合、VPN区間のTCP輻輳制御ループと、エンド端末間のTCP輻輳制御ループが重複して存在することによって、意図しない性能低下を招く問題である。また、TCPより上位層のプロトコルとの相互作用の問題として、ビデオストリーミングなどのアプリケーション層によるレート制御などが、TCPの輻輳制御機構によるネットワークスループットの変化に対応できない問題なども挙げられる。

これらの問題は、TCPの輻輳制御機構、あるいは他の階層の制御を改変することによって、階層間の相互作用を軽減することによって解決されることが一般的である。我々の研究グループでも過去に、TCP over ATMネットワーク[5-7]、TCPを用いたビデオストリーミング[8,9]などにおいてそのような手法による解決を試みた。また、ある階層で実現されている機能を、他の階層へ取り込むことによって、相互作用を軽減することも有効である。我々の研究グループにおいては、従来アプリケーション層プログラムによって行われてきたエンド端末間パスの帯域計測を、TCPが行うインライン計測手法の検討を行った[10-12]。

2.2.2 異なるバージョンのTCPの混在

TCPの実装はバージョンによって大きく異なる上、同じバージョンの実装であってもOSによって細かい挙動が異なることが多い。このことは、それらのTCPコネクションがボトルネックを共有した際に、スループットの不公平を招く。さらに、TCPのような輻輳制御を行わないUser Datagram Protocol (UDP)とTCPの間でも深刻な不公平が発生することが指摘されている。これらの問題は、ネットワーク内において第4層以上の制御を行わないインターネットにおいては本質的に避けることができない。

これらの問題は、一方の(主に後発の)プロトコルの輻輳制御機構を改変し、従来から存在するプロトコルの存在を何らかの方法で検知し制御を変化させるなどの手法によって解決され

てきた。我々の研究グループにおいても過去に、TCP Reno と TCP Vegas [13] 間の公平性や HighSpeed TCP [14] と TCP Reno 間の公平性などに着目し、輻輳制御機構の改変による問題の解決を試みた [15,16]。

2.2.3 様々なネットワーク環境への対応

TCP の特性の 1 つとして、どのようなネットワーク環境においても安定的な通信を行なうことができる、というロバスト性が挙げられるが、このことは、個々のネットワーク環境における性能の最適化の観点では劣る場合があるということを意味する。これは、上述したようにインターネットがさまざまなネットワークを IP というプロトコルで接続しているという性質を鑑みると、止むを得ない性質であると考えられる。しかし、特に近年の光ファイバ技術や無線ネットワーク技術によるアクセスネットワーク環境の劇的な進歩にともない、そのような環境における TCP の性能が着目されることが多くなり、様々な問題が指摘されつつある。例えば、ネットワーク輻輳に加えてリンクエラーによりパケット廃棄が発生、および変動する無線ネットワーク [17–22] や、端末の移動によりエンド間の経路が通信中に変化し、スループット低下が発生するモバイル環境 [23] などが挙げられる。これらを始めとするさまざまな問題を解決するために、これまでに多くの TCP に対する改善が行われてきた (例えば [24,25])。我々の研究グループにおいても、受信側 TCP の ACK パケット生成方法の改変によるスループット低下の防止 [26,27] や、ACK パケットの大量廃棄に対して早期に輻輳制御を実行することによる無線 LAN における上下フロー間の公平性改善 [28] などに取り組んできた。

また、ネットワークの高速化により、TCP コネクションが利用できるネットワークの帯域遅延積 (リンク帯域とエンドホスト間の伝播遅延時間の積) が飛躍的に増大している。例えば、ラウンドトリップ時間 (RTT) が約 130 msec となる太平洋を狭んだ 2 台のエンドホスト間の最低帯域が 100 Mbps から 1 Gbps 程度である、という環境も一般に利用可能となりつつある [29]。このような高速・高遅延ネットワーク環境において、現在多くの OS の TCP 実装が基本としている TCP Reno を用いると、その輻輳制御方式の特徴が原因となって、リンク帯域を十分使うことができない、という問題が指摘されている。これは、TCP Reno が旧来の低速ネットワークを想定して設計されていること、またインターネットユーザがよりスループットなどの性能に敏感になっていることなどに起因していると考えられる。

この問題に対しては、新たな TCP 輻輳制御機構によって高速・高遅延環境における高スループットを実現する試みが多数行われている。例えば、従来の TCP Reno よりも平均的な輻輳ウィンドウサイズを高く保つためのウィンドウ更新アルゴリズム [14,30]、パケット廃棄だけではなく、ラウンドトリップ時間の変動に応じて輻輳ウィンドウサイズを更新する delay-based 手法 [31]、あるいはそれらの複合によって、高スループットと従来 TCP との公平性を維持する手法 [32–34] などが挙げられる。また、通常の TCP コネクションを複数本用いることによってアプリケーションスループットを向上させる並列 TCP 手法が実装された例も存在する [35]。我々の研究グループにおいても、従来 TCP との公平性を改善する HighSpeed TCP の改良 [16]、帯域計測に基づく輻輳制御機構 [36] に関する研究に取り組んだ。近年においても、大規模データセンタにおける TCP の性能

劣化問題 [37] などが指摘され、解決手法 [38] が提案されており、新たなネットワーク環境における TCP の性能向上の課題は今後も発生し続けると考えられる。

3. 無線 LAN 環境における省電力 TCP データ転送の消費電力解析

本章では、筆者らの研究グループにおいて取り組んでいる、無線 LAN を介した TCP データ転送の省電力化 [39–41] に関する研究の概略を示し、消費電力解析について述べる。なお、本章の内容は [41] を基にしている。

無線ネットワーク技術の発展にともない、無線機能を有した小型の端末 (無線端末) を利用したインターネットアクセスが一般的になってきている。無線端末は通常バッテリー駆動であるため、駆動時間の長期化という観点から消費電力の削減は重要な課題である。IEEE 802.11 無線 LAN においては、無線通信が消費する電力が全体の 10% から 50% を占めることが報告されており、無線通信の消費電力を削減することが機器全体の消費電力を削減するうえで重要である。

我々は [39] において、TCP の動作を考慮した消費電力モデルを提案した。本稿では、[39] において提案した消費電力モデルを拡張し、複数のデータセグメントをまとめてバースト転送した場合の消費電力モデルを構築する。数値解析により、バースト転送を利用することによって省電力効果が大きくなることを明らかにする。このようなバースト転送は TCP の Delayed ACK オプションを用いて実現することが可能である。

3.1 TCP の動作を考慮した消費電力モデル

[39] においては、MAC レベルの消費電力モデルと TCP レベルの消費電力モデルを組合せることによって、無線 LAN において TCP データ転送を行った場合の消費電力がモデル化している。TCP レベルのモデルにおいては、TCP データ転送をスロースタートフェーズとそれ以外の定常フェーズに分け、さらに定常フェーズを TD 期間と TO 期間に分けて消費電力を解析している。なお、TD 期間は TCP の輻輳回避フェーズにおいてパケット廃棄が発生してから次に発生するまでの期間を表し、TO 期間はタイムアウトが発生している期間を表す。

本稿では、理想的なスリープをした場合の消費電力モデルを拡張することにより、バースト転送を利用してスリープした場合の消費電力モデルを構築する。具体的には、TD 期間において、図 1 に示すようなバースト転送を適用し、TD 期間における消費電力 J^{TD} の構成要素である、スリープ状態への遷移回数およびスリープ時間の期待値 ($E[N_{td}^s]$ および $E[T_{td}^s]$) を導出する。なお、無線 LAN においてはフレームが損失せず、有線ネットワーク区間においてデータセグメントが廃棄されると仮定して、消費電力モデルを構築する。

3.1.1 バースト転送を利用してスリープした場合における TCP データ転送の消費電力モデル

TD 期間において、 m 個のデータセグメントをバースト転送する場合に送受信されるパケットシーケンスを図 1 に示す。通常 TCP の輻輳回避フェーズには、1 RTT 間に図 1(a) のようにパケットが送受信される。このとき、図 1(b) のように、データセグメントを m 個まとめてバースト転送するとパケットが送受信され、一つあたりのパケット間隔 (packet interval) が大きく

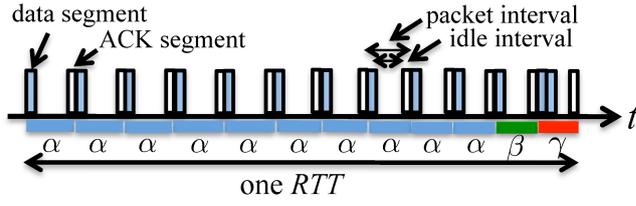
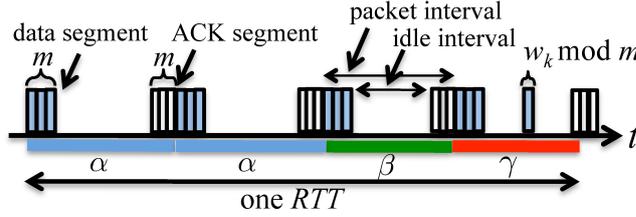
(a) $m = 1$ (b) $m = 3$

図1 TD 期間におけるパケットシーケンス

なる。このように複数個のデータセグメントをまとめてバースト転送することにより、省電力効果の向上とパケット送受信タイミングの推定が容易になることが期待できる。

k 番目のラウンドにおけるウィンドウサイズを w_k とすると、バースト転送をすることで w_k 個分のデータセグメントが m 個ずつ送信される。 w_k 個のデータセグメントが m で割り切れない場合は、その端数の $(w_k \bmod m)$ 個のデータセグメントがまとめて送信される。なお、 $m = 1$ はバースト転送を利用しない場合に相当する。パケット送受信が行われてない区間をアイドル間隔 (idle interval) とすると、スリープするためにはアクティブ状態からスリープ状態への遷移時間 (T^{as}) とスリープ状態からアクティブ状態への遷移時間 (T^{sa}) が必要となる。ここで、パケット間隔を図1のように α 区間、 β 区間および γ 区間の3つの区間に分類し、アイドル間隔においてスリープ可能である条件を求める。 γ 区間は、 k 番目のラウンドにおける輻輳ウィンドウにおいて、最後の $(m + (w_k \bmod m))$ 個のデータセグメントが送信されるパケット間隔を表し、 β 区間は γ 区間の一つ前のパケット間隔を表す。それ以外のパケット間隔はすべて α 区間とする。

以上の議論に基づき、各区間においてスリープ可能であるときの条件を求める。 k 番目のラウンドにおいては、その3つの区間の数の合計は $\lceil \frac{w_k-1}{m} \rceil$ である。したがって、各区間の時間はそれぞれ $\frac{RTT}{\lceil \frac{w_k-1}{m} \rceil}$ となる。 α 区間においては、 m 個のデータセグメントと m 個の ACK セグメントが送受信されるため、このときのアイドル間隔は $\frac{RTT}{\lceil \frac{w_k-1}{m} \rceil} - mT^t - mT^r$ となる。以上から、 α 区間でスリープ可能となるときに1となるステップ関数 $U_\alpha(k)$ を導入すると、 $U_\alpha(k)$ は以下ようになる。

$$U_\alpha(k) = \begin{cases} 1 & \text{if } \frac{RTT}{\lceil \frac{w_k-1}{m} \rceil} - mT^t - mT^r > T^{as} + T^{sa} \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

同様に、 β 区間および γ 区間においてスリープ可能である場合に1となるようなステップ関数をそれぞれ $U_\beta(k)$ および $U_\gamma(k)$ とすると、それらは次のように導出できる。

$$U_\beta(k) = \begin{cases} 1 & \text{if } \frac{RTT}{\lceil \frac{w_k-1}{m} \rceil} - mT^t \\ & - (w_{k-1} - m(\lceil \frac{w_k-1}{m} \rceil - 1))T^r > T^{as} + T^{sa} \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

$$U_\gamma(k) = \begin{cases} 1 & \text{if } m > 1 \text{ and} \\ & \frac{\frac{RTT}{\lceil \frac{w_k-1}{m} \rceil} - (w_k - m(\lceil \frac{w_k-1}{m} \rceil - 1))T^r}{\lceil \frac{w_k}{m} \rceil - \lceil \frac{w_{k-1}}{m} \rceil + 1} \\ & > T^{as} + T^{sa} \\ 1 & \text{if } m = 1 \text{ and} \\ & \frac{RTT}{w_{k-1}} - (w_k - w_{k-1} + 1)T^t - T^r > T^{as} + T^{sa} \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

以降では、(3)式、(4)式および(5)式を利用して、バースト転送を利用してスリープした場合における $E[N_{td}^s]$ および $E[T_{td}^s]$ を求める。

まず、 $E[N_{td}^s]$ を導出する。 k 番目のラウンドにおけるスリープ状態への遷移回数を N_k^s とし、 r_{td} を TD 期間においてパケット廃棄が発生したときのラウンド番号とすると、 $E[N_{td}^s]$ は $E[N_{td}^s] = \sum_{k=1}^{r_{td}+1} N_k^s$ となる。なお、 $r_{td} = E[W]/2$ である。

次に N_k^s を決定する。 $\lceil \frac{w_k-1}{m} \rceil = 1$ であるときは γ 区間が1つだけであり、 γ 区間においてスリープできるならば $(\lceil \frac{w_k}{m} \rceil - \lceil \frac{w_{k-1}}{m} \rceil + 1)$ 回スリープ状態へ遷移できる。 $\lceil \frac{w_k-1}{m} \rceil > 1$ である場合は、 $(\lceil \frac{w_k-1}{m} \rceil - 2)$ 個の α 区間と、 β 区間と γ 区間がそれぞれ1つずつ存在する。したがって、 N_k^s は次のように計算できる。

$$N_k^s = \begin{cases} U_\gamma(k) (\lceil \frac{w_k}{m} \rceil - \lceil \frac{w_{k-1}}{m} \rceil + 1) & \text{if } \lceil \frac{w_k-1}{m} \rceil = 1 \\ U_\alpha(k) (\lceil \frac{w_k-1}{m} \rceil - 2) + U_\beta(k) \\ & + U_\gamma(k) (\lceil \frac{w_k}{m} \rceil - \lceil \frac{w_{k-1}}{m} \rceil + 1) & \text{otherwise} \end{cases} \quad (6)$$

なお、 $m = 1$ である場合は、バースト転送を利用しない場合であるため、 N_k^s は以下のように計算できる。

$$N_k^s = \begin{cases} U_\alpha(k)w_{k-1} & \text{if } \lceil \frac{w_k-1}{m} \rceil = 1 \\ (w_{k-1} - 1)U_\alpha(k) + U_\gamma(k) & \text{otherwise} \end{cases} \quad (7)$$

また、 r_{td} 番目のラウンドにおいてスリープが可能であれば、 $(r_{td} + 1)$ 番目のラウンドにおいてもスリープが可能である。このとき、スリープ状態には $\lceil E[W]/m \rceil$ 回遷移できる。したがって、 $N_{r_{td}+1}^s$ は以下のように計算できる。

$$N_{r_{td}+1}^s = U_\alpha(r_{td}) \lceil E[W]/m \rceil \quad (8)$$

次に、 $E[T_{td}^s]$ を求める。 k 番目のラウンドにおけるスリープ時間を T_k^s とすると、 N_{td}^s と同様に、 $E[T_{td}^s] = \sum_{k=1}^{r_{td}+1} T_k^s$ と書くことができる。

次に T_k^s を決定する。 $\lceil \frac{w_k-1}{m} \rceil = 1$ である場合は、 γ 区間が1

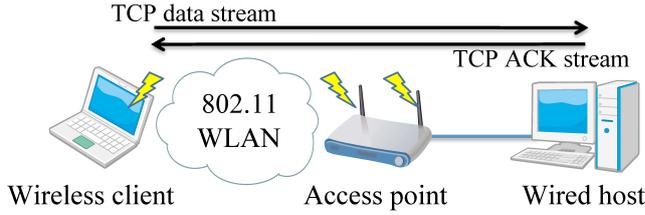


図2 無線 LAN 環境

表1 無線 LAN パラメータ

項目	値	項目	値
データレート	54 Mbps	PLCP プリアンプル	16 μ s
Slot time	9 μ s	MAC ヘッダ	24 バイト
SIFS	16 μ s	LLC ヘッダ	8 バイト
DIFS	34 μ s	CW_{min}	15

つだけであり、 γ 区間でスリープ可能であればスリープ可能である。 $\lceil \frac{w_{k-1}}{m} \rceil > 1$ である場合は、 $(\lceil w_{k-1} \rceil - 2)$ 個の α 区間および β 区間と γ 区間がそれぞれ 1 つずつ存在する。このとき、それぞれの区間におけるスリープ時間の合計が k 番目のラウンドにおけるスリープ時間となる。したがって、 T_k^s は以下のように計算できる。

$$T_k^s = \begin{cases} U_\gamma(k) \left(\frac{RTT}{\lceil w_{k-1}/m \rceil} - (w_k - m(\lceil \frac{w_{k-1}}{m} \rceil - 1))T^t - T^r \right) \\ \quad - N_k^s(T^{as} + T^{sa}) & \text{if } \lceil \frac{w_{k-1}}{m} \rceil = 1 \\ U_\alpha(k) \left(\frac{RTT}{\lceil w_{k-1}/m \rceil} - 2 \right) \left(\frac{RTT}{\lceil w_{k-1}/m \rceil} - mT^t - mT^r \right) \\ \quad + U_\beta(k) \left(\frac{RTT}{\lceil w_{k-1}/m \rceil} - mT^t - (w_{k-1} - m(\lceil \frac{w_{k-1}}{m} \rceil - 1))T^r \right) \\ \quad + U_\gamma(k) \left(\frac{RTT}{\lceil w_{k-1}/m \rceil} - (w_k - m(\lceil \frac{w_{k-1}}{m} \rceil - 1))T^t - T^r \right) \\ \quad - N_k^s(T^{as} + T^{sa}) & \text{otherwise} \end{cases} \quad (9)$$

また、 $(r_{td} + 1)$ 番目のラウンドにおいては、 $\frac{2E[W]-1}{E[W]}RTT$ の間に $E[W]$ 個のデータセグメントが送信され、 $E[W]$ 個の ACK セグメントが受信される。したがって、 $T_{r_{td}+1}^s$ は以下のように計算できる。

$$T_{r_{td}+1}^s = U_\alpha(r_{td}) \left(\frac{2E[W]-1}{E[W]}RTT - \frac{E[W]}{2}T^t - E[W]T^r \right) - N_{r_{td}+1}^s(T^{as} + T^{sa}) \quad (10)$$

3.2 数値例

3.2.1 パラメータ設定

図2のIEEE 802.11a 無線 LAN において、TCP を利用して無線端末から 100 MB のファイルを有線ホストへ転送するときの消費電力を評価する。また、図2のアクセスポイントと有線ホストの間の有線ネットワーク間でデータセグメントが確率によって廃棄される。無線 LAN 規格には IEEE 802.11a を想定し、無線 LAN に関するパラメータは表1のものを利用した。Atheros AR5004 チップのデータシート [42] より、送信、受信、アイドルおよびスリープ時の消費電力はそれぞれ 1.4 W、0.9 W、0.8 W および 0.016 W とした。また、アクティブ状態からスリープ状態へ遷移するときの消費電力はアイドル時と同様

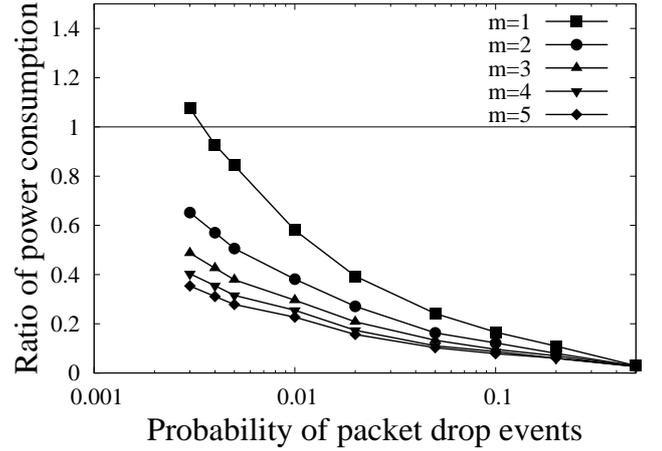


図3 消費電力比 (RTT: 100 ms)

とし、スリープ状態からアクティブ状態へ遷移するときの消費電力は送信時と同様とした。 T^{as} および T^{sa} はそれぞれ 1 ms とした。TCP データセグメントサイズは 1500 バイトとし、TCP ACK セグメントサイズは 40 バイトとした。

3.2.2 解析結果

スリープをしなかった場合の消費電力に対するバースト転送を利用してスリープした場合の消費電力の比を図3に示す。図3の x 軸は有線ネットワーク区間におけるパケット廃棄イベントの発生確率を表し、バースト転送の際にまとめて送信するパケット数 m を変化させた場合の結果である。なお、 $m = 1$ である場合はバースト転送を利用せずにスリープした場合に相当する。

図3より、パケット廃棄イベントの発生確率が大きいほど省電力効果が大きいことがわかる。また、バースト転送においてまとめるパケット数 m が大きいほど省電力効果が大きいことがわかる。例えば、パケット廃棄イベントの発生確率が 0.01 であるとき、 $m = 1$ である場合は消費電力が 60%、 $m = 2$ である場合は消費電力が 40% まで削減できている。これは、 m によってスリープ時間はほぼ変わらないものの、 m が大きいほどスリープ状態への遷移回数が少なくなり、状態遷移にかかる消費電力が削減できたためである。

3.3 今後の課題

今後の課題としては、バースト転送方式を IEEE802.11PSM などの既存の省電力プロトコル上で動作させた場合の挙動や消費電力解析を行いたい。また、複数 TCP コネクションや複数端末が存在する状況への拡張も重要である。さらに、消費電力モデルをマルチ MAC やマルチパス環境を想定したモデルに拡張したい。

4. おわりに

本報告においては、TCP の輻輳制御機構の概要をまとめ、これまで指摘されてきた問題点を整理し、それに対する解決方法の概略を紹介し、筆者の研究グループにおける研究例を述べた。さらに、筆者の研究グループにおける近年の研究例である、TCP の挙動の変更によって無線 LAN 経由のデータ転送を省電力化する、省電力 TCP 手法に関する研究内容を述べた。

謝 辞

本報告の一部は、独立行政法人 情報通信研究機構委託研究「新世代ネットワークを支えるネットワーク仮想化基盤技術の研究開発(課題ウ)」の支援による。

文 献

- [1] J. B. Postel, "Transmission control protocol," *RFC 793*, Sept. 1981.
- [2] M. Fomenkov, K. Keys, D. Moore, and K. Claffy, "Longitudinal study of Internet traffic in 1998-2003," in *Proceedings of Winter International Symposium on Information and Communication Technologies (WISICT 2004)*, Jan. 2004.
- [3] Hobbes' Internet timeline 10. available at <http://www.zakon.org/robert/internet/timeline/>.
- [4] V. Jacobson, "Congestion avoidance and control," in *Proceedings of ACM SIGCOMM '88*, pp. 314-329, Aug. 1988.
- [5] G. Hasegawa, H. Ohsaki, M. Murata, and H. Miyahara, "Performance evaluation and parameter tuning of TCP over ABR service in ATM networks," *IEICE Transactions on Communications*, vol. E79-B, pp. 668-683, May 1996.
- [6] G. Hasegawa, H. Ohsaki, M. Murata, and H. Miyahara, "Performance improvement of TCP over EFCI-based ABR service class by tuning of congestion control parameters," *IEICE Transactions on Communications*, vol. V80-B, pp. 1444-1453, Oct. 1997.
- [7] 長谷川剛, 大崎博之, 村田正幸, 宮原 秀夫, "誤再送を考慮した TCP のふくそう制御方式/エラー回復方式の改善," *電子情報通信学会論文誌*, vol. J82-B, pp. 2074-2084, Nov. 1999.
- [8] 長谷川剛, 尾池健二, 天野勝博, 村田正幸, "インラインネットワーク計測技術のビデオ会議アプリケーションへの応用," *電子情報通信学会技術研究報告* (2010-25), vol. 110, pp. 13-18, July 2010.
- [9] H. Hisamatsu, G. Hasegawa, and M. Murata, "Network friendly transmission control for progressive download over tcp," to appear in *Journal of Communications, SI on Multimedia Streaming*, 2012.
- [10] L. T. M. Cao, G. Hasegawa, and M. Murata, "ImTCP: TCP with an inline measurement mechanism for available bandwidth," *Computer Communications Journal special issue of Monitoring and Measurements of IP Networks*, vol. 29, pp. 1614-1626, June 2006.
- [11] L. T. M. Cao, G. Hasegawa, and M. Murata, "A simultaneous inline measurement mechanism for capacity and available bandwidth of end-to-end network path," *IEICE Transactions on Communications*, vol. E89-B, pp. 2469-2479, Sept. 2006.
- [12] L. T. M. Cao, G. Hasegawa, and M. Murata, "ICIM: An inline network measurement mechanism for highspeed networks," in *Proceedings of NOMS 2006 E2EMON Workshop 2006*, Apr. 2006.
- [13] L. S. Brakmo, S. W. O'Malley, and L. L. Peterson, "TCP Vegas: New techniques for congestion detection and avoidance," in *Proceedings of ACM SIGCOMM '94*, pp. 24-35, Oct. 1994.
- [14] S. Floyd, "HighSpeed TCP for large congestion windows," *Request for Comments 3649 (Experimental)*, Dec. 2003.
- [15] G. Hasegawa, K. Kurata, and M. Murata, "Analysis and improvement of fairness between TCP Reno and Vegas for deployment of TCP Vegas to the Internet," in *Proceedings of IEEE ICNP 2000*, Nov. 2000.
- [16] Z. Zhang, G. Hasegawa, and M. Murata, "Performance analysis and improvement of HighSpeed TCP with TailDrop/RED routers," *IEICE Transactions on Communications*, vol. E88-B, pp. 2495-2507, June 2005.
- [17] F. Lefevre and G. Vivier, "Understanding TCP's behavior over wireless links," in *Proceedings of Symposium on Communications and Vehicular Technology (SCVT 2000)*, pp. 123-130, Oct. 2000.
- [18] V. Tsaoussidis and I. Matta, "Open issues on TCP for mobile computing," *Wireless Communications and Mobile Computing*, vol. 2, pp. 3-20, Feb. 2002.
- [19] E. S. Chang and R. Taborek, "Recommendation of 10e-13 bit error rate for 10 gigabit ethernet," in *Proceedings of IEEE 802.3 high speed study group July 1999 plenary week meeting*, Jul. 1999.
- [20] IEEE, "IEEE standard for local and metropolitan area networks: overview and architecture." IEEE Std 802-2001, Dec. 2001.
- [21] C. H. Nam, S. C. Liew, and C. P. Fu, "An experimental study of ARQ protocol in 802.11b wireless LAN," in *Proceedings of Wireless Personal Multimedia Communications (WPMC 2002)*, Oct. 2002.
- [22] 3GPP, "Services and service capabilities." Technical Specification TS 22.105 v6.2.0 (2003-6), Jun. 2003.
- [23] T. Goff, J. Moronski, D.S. Phatak, and V. Gupta, "FreezeTCP: A true end-to-end TCP enhancement mechanism for mobile environments," in *Proceedings of IEEE INFOCOM 2000*, Mar. 2000.
- [24] V. Jacobson and R. Braden, "TCP extensions for long-delay paths," *Request for Comments 1072*, Oct. 1988.
- [25] S. Floyd and T. Henderson, "The NewReno modification to TCP's fast recovery algorithm," *Request for Comments 2582*, Apr. 1999.
- [26] G. Hasegawa, M. Nakata, and H. Nakano, "Receiver-based ACK splitting mechanism for TCP over wired/wireless heterogeneous networks," *IEICE Transactions on Communications*, vol. E90-B, pp. 1132-1141, May 2007.
- [27] G. Hasegawa, M. Nakata, and H. Nakano, "Modeling TCP throughput over wired/wireless heterogeneous networks for receiver-based ACK splitting mechanism," *IEICE Transactions on Communications*, vol. E90-B, pp. 1682-1691, July 2007.
- [28] M. Hashimoto, G. Hasegawa, and M. Murata, "A transport-layer solution for alleviating TCP unfairness in a wireless LAN environment," *IEICE Transactions on Communications*, vol. E94-B, pp. 765-776, Mar. 2011.
- [29] C. Marcondes, A. Persson, M. Sanadidi, M. Gerla, H. Shimonishi, T. Hama, and T. Murase, "Inline path characteristic estimation to improve TCP performance in high bandwidth-delay networks," in *Proceedings of PFLDnet 2006*, Feb. 2006.
- [30] T. Kelly, "Scalable TCP: Improving performance in highspeed wide area networks," *ACM SIGCOMM Computer Communication Review*, vol. 32, Apr. 2003.
- [31] C. Jin, D. X. Wei, and S. H. Low, "FAST TCP: motivation, architecture, algorithms, performance," in *Proceedings of IEEE INFOCOM 2004*, Mar. 2004.
- [32] K. T. J. Song, Q. Zhang, and M. Sridharan, "Compound TCP: A scalable and TCP-friendly congestion control for high-speed networks," in *Proceedings of PFLDnet 2006*, Feb. 2006.
- [33] H. Shimonishi, T. Hama, and T. Murase, "TCP-Adaptive Reno: Improving efficiency-friendliness tradeoffs of TCP congestion control algorithm," in *Proceedings of PFLDnet 2006*, Feb. 2006.
- [34] I. Rhee and L. Xu, "CUBIC: A new TCP-friendly high-speed TCP variant," in *Proceedings of PFLDnet 2005*, Feb. 2005.
- [35] W. Allcock, "GridFTP: Protocol extensions to FTP for the Grid," Available at: <http://www.ggf.org/documents/GFD.20.pdf>, Apr. 2003.
- [36] G. Hasegawa and M. Murata, "TCP symbiosis: congestion control mechanisms of TCP based on Lotka-Volterra competition model," in *Proceedings of Inter-Perf 2006*, Oct. 2006.
- [37] R. Griffith, Y. Chen, J. Liu, A. Joseph, and R. Katz, "Understanding TCP incast throughput collapse in datacenter networks," in *Proceedings of WREN Workshop 2009*, Aug. 2009.
- [38] M. Alizadehzy, A. Greenberg, D. A. Maltz, J. Padhyey, P. Pately, B. Prabhakar, S. Senguptay, and M. Sridharan, "Data center TCP (DCTCP)," in *Proceedings of SIGCOMM 2010*, Aug. 2010.
- [39] M. Hashimoto, G. Hasegawa, and M. Murata, "Modeling and analysis of power consumption in TCP data transmission over a wireless LAN environment," in *Proceedings of ICC 2011 4th International Workshop on Green Communications*, June 2011.
- [40] M. Hashimoto, G. Hasegawa, and M. Murata, "Energy efficiency analysis of TCP with burst transmission over a wireless LAN," in *Proceedings of IEEE ISCT 2011*, Oct. 2011.
- [41] 橋本匡史, 長谷川剛, 村田正幸, "無線 LAN 環境におけるバースト転送を利用した TCP フローの消費電力解析," *電子情報通信学会総合大会 BS-5-7*, Mar. 2011.
- [42] Wistron NeWeb Corp., "CM9: WLAN 802.11 a/b/g mini-PCI Module." available at <http://site.microcom.us/CM9.pdf>.