

研究種目：基盤研究 C

研究期間：2007～2009

課題番号：19500055

研究課題名（和文）

インラインネットワーク計測技術に基づく新たなトランスポートサービスの創出

研究課題名（英文）

Researches on transport-layer services based on inline network measurement technologies

研究代表者

長谷川 剛 (HASEGAWA GO)

大阪大学・サイバーメディアセンター・准教授

研究者番号：00294009

研究成果の概要（和文）：

TCP によるデータ転送の最中にそのコネクションが通過するパスの物理帯域・利用可能帯域がわかることで、さまざまな応用が可能となる。たとえば計測結果を上位アプリケーションに伝えることで、P2P ネットワークにおけるピア選択や、データグリッドにおけるダウンロードサーバの選択などである。本研究ではパスの帯域情報をデータ転送制御に用いる TCP Symbiosis 方式の提案・評価を行った。

研究成果の概要（英文）：

I construct a robust, selfadaptive, scalable, and robust congestion control mechanism for TCP named as TCP Symbiosis. The largest difference of our proposed mechanism from existing TCP variants is that I borrowed algorithms from biophysics to update the window size: the logistic growth model and the Lotka-Volterra competition model. I treat the competitive relationship among species in biophysics models as the bandwidth sharing among competing TCP connections in the Internet. I then enhance the congestion control mechanism to deal with the environmental changes, by adding the noise term to Lotka-Volterra competition model equation and adaptively setting control parameters. I investigate the effectiveness of the proposed mechanisms by both simulation and implementation experiments and confirm that TCP Symbiosis can provide high performance regardless of the network bandwidth delay product, and survive against measurement errors and sudden environmental changes.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2007年度	1,000,000	300,000	1,300,000
2008年度	1,600,000	480,000	2,080,000
2009年度	900,000	270,000	1,170,000
年度			
年度			
総計	3,500,000	1,050,000	4,550,000

研究分野：総合領域

科研費の分科・細目：情報学・計算機システム・ネットワーク

キーワード：ネットワーク計測、利用可能帯域、インライン計測、TCP、公平性、数理生態学

## 1. 研究開始当初の背景

現在のインターネットにおける主要なトランスポートプロトコルであるTCP (Transmission Control Protocol)は、近年ますますその重要性を増している。従来 UDP (User Datagram Protocol)が用いられてきた動画像・音声ストリーミングトラフィックが、近年ではファイアウォールやNAT (Network Address Translation)の配下に存在するユーザへの安定的なサービス提供を目的として、TCPを用いるケースが増加している (Gya0, Skypeなど)。また、ファイル交換などのP2Pネットワークサービスが増加し、従来受信側になることが多かったエンド端末が、送信側となる頻度が増している。

さらに、オーバレイネットワークの登場によって、エンドホストによるトラフィックのルーティングや、ネットワーク内部にレイヤ4やレイヤ7の終端装置を設置し、トラフィック制御をするサービスが登場しつつある。このことは、TCPの終端がネットワーク内部に置かれることを意味するため、ネットワーク・アプリケーション性能が、TCP性能へ依存する度合いがますます大きくなっている。

TCPは本来、エンドホスト間に存在するネットワークをブラックボックスと捉えて輻輳制御・エラー回復制御などを行うが、近年のネットワークの多様化・大規模化・高速化などによって、その手法には限界が見られつつある。たとえば、1Gbpsを超えるような回線においてTCPがその帯域を十分に使えないことはよく知られた問題であり、本申請者を含めた多くの研究者によってその改善が試みられている。その結果、次期Windows (Vista)には従来用いられたTCP Renoではなく、Compound TCPと呼ばれる方式が採用される予定であり、また現在のLinuxにおいてはCUBICと呼ばれるTCPの改良方式が採用されている。これらの新しい方式や、特に高速・高遅延環境で有力とされるHighSpeed TCPやFAST TCPなどに共通するのは、従来のTCP Renoがネットワーク輻輳のシグナルとしてパケット廃棄の発生のみを用いていたのに対して、RTT (Round Trip Time)や、ウィンドウサイズそのものの値の変化を輻輳のシグナルとして用いる点にある。すなわち、従来のTCP Renoの欠点を補い、高い性能を得るためには、エンドホスト間パスのネットワーク資源に関する情報を、計測によって得ることが重要となる。

## 2. 研究の目的

本研究では、BTC と呼ばれる予想 TCP スループットをインライン計測によって獲得し、それを利用した TCP 輻輳制御方式を提案する。BTC は TCP コネクションを用いて長時間データ転送を行えば自然と得られるが、それを転送初期の段階で推測を行うことができれば、さまざまな状況において利用可能帯域の予測よりも有用であることは明白である。たとえば上述したピア選択やダウンロードサーバ選択、並列 TCP 手法における TCP コネクション数の決定などが挙げられる。また、TCP の輻輳制御手法そのものにとっても、BTC 予測が行えるとすると、極めて効率的な、輻輳制御手法の設計が可能になると考えられる。

すなわち、本研究の目的は、インラインネットワーク計測と、その結果を利用するさまざまな輻輳制御手法が確立されることによって、現在のインターネットの基本コンセプトである KISS (Keep it simple, stupid) を崩すことなく、高度なトランスポートサービスを展開することである。

## 3. 研究の方法

まず、BTC を TCP コネクション内でインライン計測によって推測するためのアルゴリズムの提案を行う。できるだけ早期に推測値を得ることを目標とする。われわれがこれまでに提案した利用可能帯域・物理帯域のインライン計測手法においては、計測結果に統計的信頼性を与えている。これは、これまでに他研究者によって提案されている数多くのネットワーク計測ツール (インライン計測ではない) には見られない特長である。したがって、BTC 計測アルゴリズムにおいても、計測結果に統計的な信頼性を示す指標を同時に提示することを目標とする。

さらに、インライン計測によって取得可能なネットワークパスの利用可能帯域および物理帯域を利用した、トランスポートサービスの提案を行う。TCP コネクションのスループットは一般的に RTT (ラウンドトリップ時間) やパケット廃棄率に大きく影響を受けるが、インライン計測によって得られるネットワークの空き帯域に関する情報を基に、ウィンドウサイズを増加させる幅を動的に変化させることによって、どんな環境でも一定のスループットを確保することを可能とする。ネットワークの空き帯域を監視することによって、過剰な制御によってネットワークが破綻することも防止する。

#### 4. 研究成果

提案した TCP Symbiosis のウィンドウサイズ制御には数理生態学において生物の個体数の変化を表すモデルを適用したアルゴリズムを用いている。一般に、環境中の生物の個体数が多くなるにつれてその増殖速度は大きくなる。また、自然界では環境や資源について様々な制約があるために、個体数の上限となる環境容量が存在し、個体数が環境容量に近づくにつれて、種内での競争などの過密効果により増殖率が次第に低下する。ロジスティック増殖モデルは、このような限られた領域の中で生息している 1 種の生物における種内の競争による影響を考慮した個体数の時間経過に伴う変化を表すモデルである。一方、ロトカ・ヴォルテラ競争モデルは、ロジスティック増殖モデルを拡張し、種内の競争による影響だけでなく 2 種以上の種間の競争による影響を考慮したモデルである。

TCP Symbiosis はインラインネットワーク計測によってネットワークパスの物理帯域および利用可能帯域の情報を取得するが、帯域情報を取得できるまでは従来の TCP Reno と同じアルゴリズムで動作する。帯域情報を取得すると、ACK パケットを受信する度に 1 つ前の ACK を受信した時刻を 0 とすることによりウィンドウサイズを求める。また、パケット廃棄やタイムアウトの発生時においても TCP Reno と同じアルゴリズムを用いる。すなわち、パケット廃棄が発生した場合にはウィンドウサイズを半減させ、タイムアウトが発生した場合には Slow-Start フェーズを開始する。

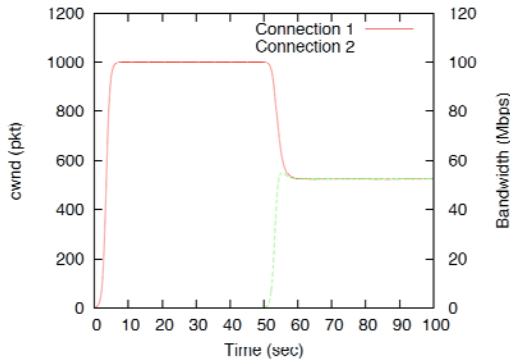


図 1: TCP Symbiosis の基本動作

TCP Symbiosis の基本的な動作例を図 1 に示す。本図はボトルネックリンク帯域を共有する 2 本の TCP Symbiosis コネクションのウィンドウサイズの変化をシミュレーションした結果である。図 1 より、TCP Symbiosis コネクションのウィンドウサイズがパケット廃棄を発生させることなくすばやく収束し、2 本のコネクションが帯域を公平に共有していることがわかる。

本研究ではさらに、TCP Symbiosis がネットワーク環境変動から受ける影響について考察し、変動に対する耐性向上を行うための改良を行った。以下にその概要を示す。

一般に、利用可能帯域の計測値には誤差が含まれる。一方、物理帯域の変動は利用可能帯域に比べて非常に小さいため、計測値の誤差は無視することができると考えられる。これより、TCP Symbiosis は、時刻  $t$  における利用可能帯域の計測値を  $A'(t)$ 、物理帯域を  $K$  とすると次のようになる。

$$\frac{d}{dt}w(t) = \epsilon \left( 1 - \frac{w(t) + \gamma(K - A'(t))\tau}{K\tau} \right) w(t)$$

利用可能帯域の計測にはパケット送信・受信間隔を用いたアルゴリズムが用いられるため、計測ノイズの大きさは計測する帯域の大きさに影響を受けると考えることができる。そこで、上式における  $(K - A'(t))$  の項は、ノイズの割合を  $n(t)$  として、時刻  $t$  における真の利用可能帯域の大きさ  $A(t)$ 、ノイズの割合を  $n(t)$  を用いて次のように表されると仮定する。

$$K - A'(t) = (1 + n(t))(K - A(t))$$

これより、上式は下記のように書き替えられる。

$$\gamma'(t) = (1 + n(t))\gamma,$$

$$\frac{d}{dt}w(t) = \epsilon \left( 1 - \frac{w(t) + \gamma'(t)(K - A(t))\tau}{K\tau} \right) w(t)$$

ロトカ・ヴォルテラ競争モデルにおいて種の個体数が 0 でない数で安定するためには、 $0 < \gamma < 1$  を満たす必要がある。 $\gamma > 1$  のように条件を満たさない場合には、絶滅にいたる種もある。これは、TCP においては一部の TCP コネクションが非常に低いスループットになる状況に相当する。

式(10)のように  $\gamma$  がノイズによって揺れる場合には  $\gamma$  の値が 1 を超えることがあるため、TCP Symbiosis の性能が低下する。このような状況を避けるためには、 $\gamma$  を安全側 (1 より十分に小さい値) に設定することが必要となる。しかしながら、式(5)よりウィンドウサイズの収束値  $w^*$  は  $w^* = \{\gamma A + (1 - \gamma)K\} \tau$  となるため、 $\gamma$  が小さい場合には収束状態においてボトルネックリンクにおけるキュー長が増加する。

シミュレーション結果を用いて上述した  $\gamma$  の特性を示す。ここでは、計測ノイズの影響を調べるために利用可能帯域の計測結果にホワイトノイズを加え、その大きさを100秒ごとに2倍にする。 $\gamma = 0.95$  および  $\gamma = 0.5$  とした場合のシミュレーション結果をそれぞれ図2と図3に示す。

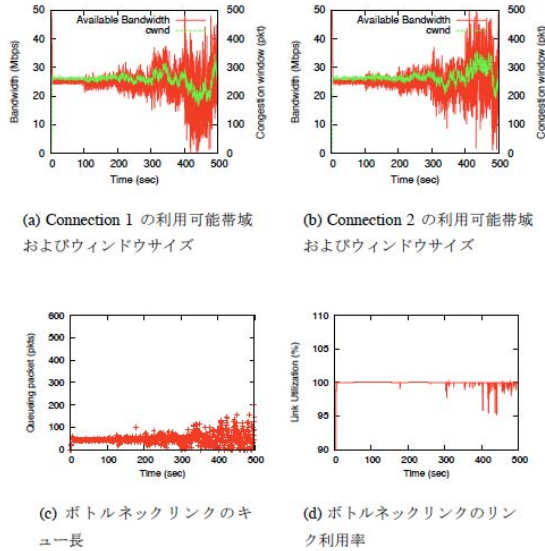


図2:  $\gamma = 0.95$  の場合の評価結果

図2より、 $\gamma$  が1に近いとき、つまり  $\gamma = 0.95$  のとき、2本の接続のウィンドウサイズはノイズが大きくなるにつれて不安定になることがわかる。これは利用可能帯域の計測結果のノイズが大きいと、式(10)における  $\gamma'(t)$  が頻繁に1より大きくなるためである。このことにより、2本の接続のウィンドウサイズが不公平に陥り、キュー長が振動することにより、リンク利用率が低下している。一方、 $\gamma$  が1に比べて非常に小さい場合(図3)、2つの輻輳ウィンドウサイズはノイズの大きさに関係なく安定しており、リンク利用率は100%を維持している。しかし、図3(c)が示すように、ボトルネックリンクのキュー長が  $\gamma = 0.95$  の場合に比べて大きいところがある。

このことは、ボトルネックリンクバッファにおけるキューイング遅延を増加させ、バックグラウンドトラフィックが急激に増加した場合におけるバッファオーバーフローの確率を増加させる。これらの結果より、ノイズが含まれる環境においては  $\gamma$  を固定値にすると良い性能が得られないことがわかる。

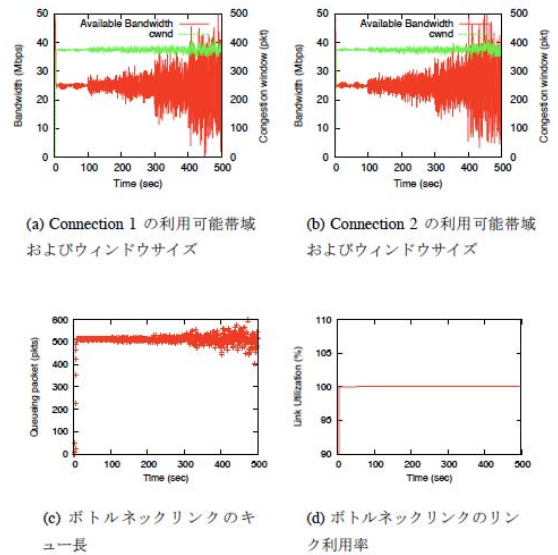
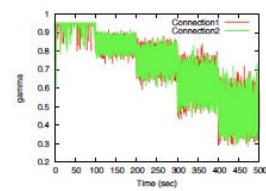
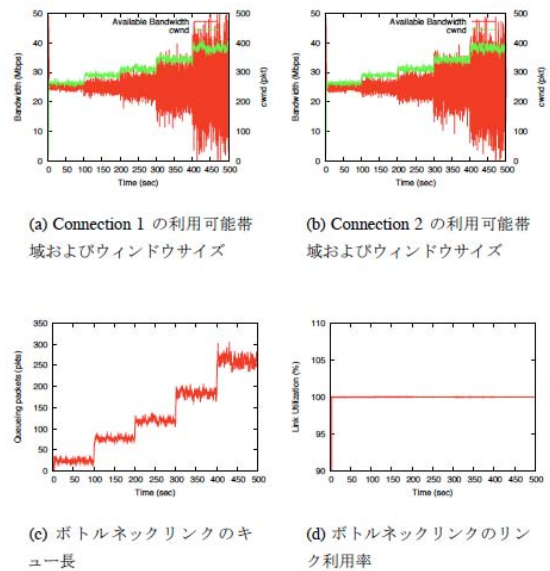


図3:  $\gamma = 0.5$  の場合の評価結果



(e)  $\gamma$  の変化

図4: 提案手法のノイズ耐性強化結果

上述の問題を解決するために、利用可能帯域の計測結果に応じて  $\gamma$  を動的に設定する方法を提案する。具体的には、計測した利用可能帯域の変動の大きさを基にして  $\gamma$  の値を決定する。上式における  $n(t)$  の項は  $(K - A'(t))$  に含まれるノイズの割合を意味するため、 $(K - A'(t))$  の計測値の変動係数  $C_v$  で置き換えることができる。これより、

上式を利用して、時刻  $t$  における  $\gamma$  の値、 $\gamma(t)$  を次のように動的に決定することができる。

$$\gamma(t) = \frac{\gamma_t}{1 + I C_v}$$

ここで、 $\gamma_t$  とは利用可能帯域の誤差がないとした場合に理想とする  $\gamma$  の値であり、 $\gamma(t)$  の目標値とする。また、 $I$  は  $\gamma$  の設定の際にどの程度ノイズの大きさを考慮するかを決定するパラメータである。

図 4 は図 2, 3 と同じ環境でシミュレーションを行った結果である。シミュレーションにおいては、 $g=0.125$ 、 $h=0.25$ 、 $\gamma_t=0.95$ 、および  $I=4$  としている。図 4(e) は  $\gamma(t)$  の変化を示している。図 4(a)、4(b) より、2 つのコネクションのウィンドウサイズがノイズの大きさに関係なく安定していることがわかる。また、図 4(d) より、ボトルネックリンクにおけるリンク利用率が 100 % を維持していることがわかる。さらに、ボトルネックリンクのキュー長もノイズの大きさによって適切に調整されており、ノイズが小さい場合には図 4(c) に比べて短くなっている。

これらの結果から、提案方式 TCP Symbiosis のノイズ耐性の強化が確認された。

#### 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 3 件)

1. Go Hasegawa, Kana Yamanegi, and Masayuki Murata, TCP Congestion Control Mechanisms for Achieving Predictable Throughput using Inline Network Measurement, IEICE Transactions on Communications, 査読有, Vol. E91-B, No. 12, pp. 3945-3955, Dec. 2008.
2. Cao Le Thanh Man, Go Hasegawa and Masayuki Murata, Monitoring overlay path bandwidth using an inline measurement technique, IARIA International Journal on Advances in Systems and Measurements, 査読有, vol 1, no 1, pp. 50-60, 2008.
3. Cao Le Thanh Man, Go Hasegawa and Masayuki Murata, Inline Bandwidth Measurement Techniques for Gigabit Networks, International Journal of Internet Protocol Technology (IJIPT), 査読有, Vol. 3, No. 2, pp. 81-94, 2008.

[学会発表] (計 5 件)

1. Mizuho Kodama, Go Hasegawa, and Masayuki Murata, Bandwidth-based congestion control for TCP: measurement noise-aware parameter settings and self-induced oscillation, in Proceedings of the 2nd International Workshop on the Network of the Future (Future-Network'09), Dresden(Germany), 18<sup>th</sup> June 2009.
2. Mizuho Kodama, Go Hasegawa, and Masayuki Murata, Implementation experiments of TCP Symbiosis: bio-inspired mechanisms for Internet congestion control, in Proceedings of CQR 2008, Arizona(USA), 30<sup>th</sup> April 2008.
3. Go Hasegawa, Kana Yamanegi, and Masayuki Murata, A Transport-layer Approach for Achieving Predictable Throughput for Internet Applications, in Proceedings of ICN 2008, Cancun(Mexico), 15<sup>th</sup> April 2008.
4. Tomoaki Tsugawa, Cao Le Thanh Man, Go Hasegawa and Masayuki Murata, Inline bandwidth measurements: Implementation difficulties and their solutions, in Proceedings of the 5th Workshop on End-to-End Monitoring (E2EMON 2007), Munchen(Germany), 21st May 2007.
5. Cao Le Thanh Man, Go Hasegawa and Masayuki Murata, Inferring Available Bandwidth of Overlay Network Paths Based on Inline Network Measurement, in Proceedings of ICIMP 2007, Sun Jose(USA), 2<sup>nd</sup> July 2007.

[その他]

<http://www.anarg.jp/imtcp/> にて、提案手法のソースコードを公開している。

#### 6. 研究組織

(1) 研究代表者

長谷川 剛 (HASEGAWA GO)  
大阪大学・サイバーメディアセンター・  
准教授  
研究者番号：00294009