

機関番号：14401

研究種目：若手研究(A)

研究期間：2007～2010

課題番号：19680003

研究課題名(和文) テラビット級トランスポート層通信プロトコルの研究開発

研究課題名(英文) Research and Development of Transport Protocol for Tera-bit Networks

研究代表者

大崎 博之(OHSAKI HIROYUKI)

大阪大学・大学院情報科学研究科・准教授

研究者番号：00294166

研究成果の概要(和文)：

本研究では、理論的・体系的なアプローチにより、テラビット級トランスポート層通信プロトコルの実現を目指し、(1) 流体近似法を用いたトランスポート層通信プロトコルの解析、(2) 流体シミュレーションを用いたトランスポート層通信プロトコルの評価、(3) テラビットネットワークに適した輻輳制御フレームワークの解明、の 3 つの研究課題に取り組んだ。テラビットネットワークに適した通信プロトコルを解析・設計するとともに、大規模ネットワークへのための流体シミュレータ技術を確立した。

研究成果の概要(英文)：

This research project aimed at realization of terabit-scale transport protocol using theoretical and systematic approaches. Specifically, this research project focused on the three challenges: (1) establishment of fluid-based analysis techniques for transport protocols, (2) performance evaluation of transport protocols using fluid models, and (3) comparative study on congestion control frameworks for terabit-scale networks. Limitations of conventional transport protocols have been analyzed and clarified, and a novel terabit-scale transport protocol and a large-scale network simulator have been designed.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2007年度	5,500,000	1,650,000	7,150,000
2008年度	5,000,000	1,500,000	6,500,000
2009年度	4,400,000	1,320,000	5,720,000
2010年度	4,400,000	1,320,000	5,720,000
年度			
総計	19,300,000	5,790,000	25,090,000

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：インターネット高度化、ネットワーク、通信工学、通信プロトコル、超高速通信

## 1. 研究開始当初の背景

近年、半導体技術や情報処理技術の急速な発展により、情報の多様化・大容量化が急速に進んでいる。このような情報爆発時代に対応するためには、それを支える超高速通信技術が不可欠である。

近年、OSI 参照モデルにおける下位レイヤの研究開発が進んでいる。例えば、物理層における通信技術として、光ファイバの伝送技術の高度化が進んでいる。これにより、単一の光ファイバで数 Tbps～数十 Tbps 程度の情報量の転送が可能になりつつある。また、ネ

ットワーク層の通信技術として、光パケット交換技術が開発され、2006年の時点で160 Gbpsの情報量の転送が実現されていた。

しかし、テラビット級の通信を真に実現するためには、物理層からネットワーク層までの伝送技術だけでは不十分であり、トランスポート層の通信技術におけるブレイクスルーが不可欠である。高速トランスポート層通信技術の研究が活発に行われているが、2006年の時点で約7.2 Gbps程度の転送速度しか実現されていない。特に、従来の高速トランスポート層通信技術の研究の大半は、単一の大容量ファイル転送のみを対象としており、通信速度の異なる膨大な数の人間や計算機が相互に大量の情報交換を行うような通信にはまったく対応できない。

テラビットネットワークでは、ネットワークの帯域遅延積が膨大となり、ギガビットネットワークと比較して制御の困難さが飛躍的に増大する。従来のトランスポート層通信プロトコルのような、経験に頼ったアドホックな設計手法はすでに限界であり、理論的・体系的な設計手法が強く求められていた。

## 2. 研究の目的

そこで本研究では、理論的・体系的なアプローチにより、テラビット級トランスポート層通信プロトコルの実現を目指し、以下のような研究課題に取り組む。

まず、テラビットネットワーク上での、既存のトランスポート層通信プロトコルの限界を定量的に明らかにする。本研究では、コアネットワークの回線速度として1 Tbps～100 Tbps、端末のインターフェース速度として10 Gbps～1 Tbps程度を想定する。ルータが多段接続された複雑なタンデムネットワークにおいて、異なる種類のトランスポート層通信プロトコルが混在した環境を対象とする。

本研究では、このようなテラビットネットワークにおける、既存の高速・広帯域向けトランスポート層通信プロトコルの性能限界を定量的に示す。具体的には、高速・広帯域向けトランスポート層通信プロトコルとして、近年、ギガビットネットワーク上での有効性が示されているXCPを対象とする。テラビットネットワークにおける、このような通信プロトコルの定常特性（スループット、パケット棄却率、転送遅延、公平性）および過渡特性（安定性、収束特性）を明らかにする。特に、膨大な数の人間や計算機が相互に大量の情報交換を行う時の特性を明らかにする。次に、テラビットネットワークにおける、3種類の輻輳制御フレームワーク（エンド-エンド型、ルータ支援型、ミドルウェア支援型）それぞれの性能限界を理論的に示す。これにより、テラビットネットワークに適した輻輳

制御フレームワークを明らかにする。

## 3. 研究の方法

### - 流体近似法を用いたトランスポート層通信プロトコルの解析

まず、流体近似法により、テラビットネットワーク上での既存のトランスポート層通信プロトコルの性能限界を明らかにする。本研究では、高速・広帯域向けトランスポート層通信プロトコルとしてXCPを対象とし、流体近似モデルを構築する。

さらに、テラビットネットワークにおけるルータの流体近似モデルを構築する。ルータの種類として、現在広く用いられているDropTailルータおよび近年活発に研究が行われているアクティブキュー管理ルータ（RED、GRED）を対象とする。

その後、トランスポート層通信プロトコルとルータの流体近似モデルを相互に接続し、ネットワーク全体の流体近似モデルを構築する。これにより、ルータが多段接続された複雑なタンデムネットワークにおいて、異なる複数のトランスポート層通信プロトコルが混在した環境の解析を可能とする。

数値シミュレーションおよび制御理論を応用した特性解析により、テラビットネットワーク上での、既存のトランスポート層通信プロトコルの特性を明らかにする。これにより、テラビットネットワークにおける、通信プロトコルの定常特性（スループット、パケット棄却率、転送遅延、公平性）および過渡特性（安定性、収束特性）を明らかにする。

### - 流体シミュレーションを用いたトランスポート層通信プロトコルの評価

次に、高速・広帯域向けトランスポート層通信プロトコルを流体シミュレータに実装する。流体シミュレーションにより、テラビットネットワーク上での、既存のトランスポート層通信プロトコルの特性を解明する。

流体シミュレーションでは、特に、大規模ネットワークにおける、トランスポート層通信プロトコルの非線型な挙動に着目する。これにより、複雑なネットワークにおける、トランスポート層通信プロトコルの非線型な挙動が、テラビットネットワーク全体にどのような影響を与えるかを明らかにする。本研究では、大規模（100,000ノード以上）のネットワークをシミュレートすることを計画している。

### - テラビットネットワークに適した輻輳制御フレームワークの解明

次に、テラビットネットワークにおける、3種類の輻輳制御フレームワーク（エンド-エンド型、ルータ支援型、ミドルウェア支援型）それぞれの性能限界を理論的に示す。

エンド-エンド型の輻輳制御フレームワークとは、エンドホストがネットワークの輻輳状況を検出する手法である。既存の TCP Reno など、エンド-エンド型の輻輳制御フレームワークは広く用いられている。エンド-エンド型の輻輳制御フレームワークは、ネットワークの下位レイヤに依存しないという利点がある。しかし、エンド-エンド型の輻輳制御フレームワークは、ギガビット程度の低速なネットワークへの適用が限界と考えられる。

ルータ支援型の輻輳制御フレームワークは、ルータがエンドホストに明示的にフィードバック情報を通知する手法である。古くは、ATM における ABR サービスクラスや IP ネットワークの ICMP Source Quench/ECN、最近では XCP や RCP など、さまざまな研究が行われている。ルータ支援型の輻輳制御フレームワークは、テラビット級のトランスポート層通信を実現するための解の一つと考えられるが、これまではギガビットネットワーク程度の低速なネットワークを対象とした研究しか行われていない。

本研究では、流体近似法を用いた数学的解析および流体シミュレーションにより、どの輻輳制御フレームワークが、どのようなテラビットネットワークにおけるトランスポート層通信プロトコルのフレームワークとして適しているかを解明する。

#### 4. 研究成果

- 流体近似法を用いたトランスポート層通信プロトコルの解析

広帯域・広域ネットワークにおける TCP の問題を解決するためのトランスポート層プロトコルとして、XCP (eXplicit Control Protocol) が提案されている。XCP は、明示的なフィードバック情報を、ルータからエンドホストに通知することにより、効率的な輻輳制御を目指している。これまで、さまざまなシミュレーション実験により XCP の性能評価が行われているが、一般的なネットワーク環境における XCP の安定性は明らかにされていない。そこで本研究では、伝搬遅延の異なる複数の XCP フローが混在するネットワークにおける XCP の安定性を解析した。その結果、(1) 各 XCP フローの伝搬遅延が異なる場合、すべての XCP フローの伝搬遅延が等しい場合よりも、XCP の制御がより安定すること、(2) ただし、各 XCP フローの伝搬遅延のばらつきが非常に大きい場合、逆

に XCP の制御は不安定となること、(3) XCP ルータの出力リンクの帯域は XCP の安定性に影響を与えないこと、などを明らかにした。さらに、本研究では、スケールフリーネットワーク (BA (Barabasi Albert) ツリー) における TCP フローのスループット・ラウンドトリップ時間・パケット棄却率の分布を導出した。スケールフリーネットワークでは、ノードとノードの接続確率はノードの次数と独立ではなく、解析的な取り扱いが容易ではない。Fekete らは、ツリー構造のスケールフリーネットワーク (BA ツリー) にネットワークトポロジを限定することによって、スケールフリーネットワーク上での TCP フローの平均スループットを導出している。ネットワークのスケールフリー構造と、TCP フローのエンド-エンド性能との関係を分析し、さまざまなネットワーク設計や制御へ応用することを考えると、TCP フローの平均的な特性だけではなく、それらの分布を知ることも重要となる。そこで本研究では、Fekete らの解析を拡張することにより、スケールフリーネットワークにおける TCP フローのスループット・ラウンドトリップ時間・パケット棄却率の分布を導出した。Fekete らの解析と同様に、スケールフリーネットワークとして BA ツリーを対象とした。TCP フローのダイナミクスは、流体近似モデルによって記述した。

- 流体シミュレーションを用いたトランスポート層通信プロトコルの評価

本研究では、大規模ネットワークの性能評価のための、フローレベルシミュレータ FSIM (Fluid-based SIMulator) を提案するとともに、実装した FSIM を用いてその有効性を検証した。フローレベルシミュレータ FSIM は、従来のフローレベルシミュレータよりも高精度かつ高速なシミュレーションが可能であるという特徴を持つ。シミュレーション精度の向上のため、より高精度な流体近似モデルを採用した。また、シミュレーション速度の向上のため、フローレベルシミュレーションで用いられる、微分方程式の数値計算アルゴリズムを高速化した。さらに、フローレベルシミュレータ FSIM は、既存のネットワーク性能評価ツールとの高い親和性を実現する。本研究では、実装したフローレベルシミュレータ FSIM を用いた実験を行い、FSIM のシミュレーション速度を評価した。その結果、フローレベルシミュレータ FSIM は、従来のフローレベルシミュレータよりも倍以上高速なシミュレーションが可能であることを示した。

さらに本研究では、現在広く用いられている TCP や、ルータ支援型の輻輳制御である XCP、

輻輳制御を行わない UDP のような、異種のフローが混在するネットワークに対応したフローレベルシミュレータ FSIM+ (Fluid-based networkSIMulator plus) を設計し、実装した。近年、パケットレベルシミュレータよりも圧倒的に高速なシミュレーションが可能な、フローレベルシミュレータが注目を浴びている。ただし、既存のフローレベルシミュレータは、データ転送を連続的に行うような TCP フロー (持続的 TCP フロー) のシミュレーションにしか対応していない。そこで本研究では、非持続的な TCP フローの流体近似モデル、XCP フローの流体近似モデル、UDP フローの流体近似モデルを利用することにより、異種のフローが混在するネットワークに対応したフローレベルシミュレータ FSIM+ を構築した。さらに、実装したフローレベルシミュレータ FSIM+ を用いた実験により、その有効性を定量的に評価した。その結果、フローレベルシミュレータ FSIM+ は、パケットレベルシミュレータ ns2 と比較して、(1) シミュレーション結果の誤差は 10% 以下であること、(2) シミュレーション速度は 10 倍以上高速であること、(3) 消費メモリ量は 10% 以下であること、などを示した。

- テラビットネットワークに適した輻輳制御フレームワークの解明

本研究では、ルータ支援型の輻輳制御フレームワークの一種であるトランスポート層通信プロトコル XCP が、ネットワークの変動に対してロバストでないことを明らかにするとともに、その解決策を提案した。XCP は、ルータからの明示的なフィードバックを利用して輻輳制御を行なうトランスポート層通信プロトコルである。これまでに、さまざまな XCP の性能評価が行なわれてきた。しかし、トラフィック変動に対する XCP のロバスト性に着目した研究はこれまでで行なわれていなかった。そこで本研究では、まず、シミュレーション実験により、(1) XCP トラフィックの変動が発生するとボトルネックリンクの利用率が低下してしまう、(2) TCP 以外の非 XCP トラフィックと XCP トラフィックが混在する環境では XCP の制御が安定しなくなる、といった問題が存在することを示した。さらに、トラフィック変動に対する XCP のロバスト性を向上させる XCP-IR (XCP with Increased Robustness) を提案した。シミュレーション実験により、XCP-IR がトラフィック変動に対して高いロバスト性を持つことを示した。

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 8 件)

- ①Y. Sakumoto, H. Ohsaki, and M. Imase, "Improving robustness of XCP (eXplicit Control Protocol) for dynamic traffic," IEICE Transactions on Communications, vol. E93-B, pp. 3013-3022, Nov. 2010.
- ②Y. Sakumoto, H. Ohsaki, and M. Imase, "Stability analysis of XCP (eXplicit Control Protocol) with heterogeneous flows," IEICE Transactions on Communications, vol. E92-B, pp. 3174-3182, Oct. 2009.
- ③H. Ohsaki, K. Sugiyama, and M. Imase, "Congestion propagation among routers with TCP flows," the International Journal of Computer Networks & Communications (IJCNC), vol. 1, pp. 112-128, July 2009.
- ④ S. Hasegawa, Y. Sakumoto, M. Wakabayashi, H. Ohsaki, and M. Imase, "Delay performance analysis on ad-hoc delay tolerant broadcast network," IEICE Transactions on Communications (Special Section on Ad Hoc and Mesh Networking for Next Generation Access Systems), pp. 728-736, Mar. 2009.
- ⑤ T. Ito, H. Ohsaki, and M. Imase, "GridFTP-APT: Automatic parallelism tuning mechanism for GridFTP in long-fat networks," IEICE Transactions on Communications, vol. E91-B, pp. 3925-3936, Dec. 2008.
- ⑥T. Ito, H. Ohsaki, and M. Imase, "On parameter tuning of data transfer protocol GridFTP for wide-area networks," International Journal of Computational Science and Engineering (IJCSSE), vol. 2, pp. 177-183, Oct. 2008.
- ⑦T. Ito, H. Ohsaki, and M. Imase, "On automatic parameter configuration mechanism for data transfer protocol GridFTP," International Transactions on Systems Science and Applications, vol. 3, pp. 220-226, Oct. 2007.

[学会発表] (計 33 件)

- ①Y. Sakumoto, H. Ohsaki, and M. Imase, "Fluid-based analysis of TCP flows in a scale-free network," to be presented at the IEEE/IPSJ International Symposium on Applications and the Internet (SAINT 2011), July 2011.
- ②Y. Sakumoto, T. Nishioka, H. Ohsaki, and M. Imase, "On the effectiveness of stable numerical solution for flow-level network simulation," to be presented at the NECSI

Eighth International Conference on Complex Systems (ICCS 2011), June 2011.

③Y. Sakumoto, H. Ohsaki, and M. Imase, "On the effectiveness of Thorup's shortest path algorithm for large-scale network simulation," in Proceedings of the First Workshop on High Speed Network and Computing Environments for Scientific Applications (HSNCE 2010), pp. 339-342, July 2010.

④T. Nishioka, Y. Sakumoto, H. Ohsaki, and M. Imase, "Design and implementation of flow-level simulator for a network with heterogeneous flows," in Proceedings of the 9th Annual International Symposium on Applications and the Internet (SAINT 2009), pp. 78-84, July 2009.

⑤F. Inoue, H. Ohsaki, Y. Nomoto, and M. Imase, "On maximizing iSCSI throughput using multiple connections with automatic parallelism tuning," in Proceedings of the 5th IEEE International Workshop on Storage Network Architecture and Parallel I/Os (SNAPI 2008), pp. 11-16, Sept. 2008.

⑥S. Hasegawa, H. Furuya, H. Ohsaki, and M. Imase, "Study on DTN (Delay Tolerant Network) based network applied to construction," in Proceedings of the 11th Symposium on Construction Robotics in Japan (11th SCR) (in Japanese), Sept. 2008.

⑦H. Ohsaki and M. Imase, "On the effect of scale-free structure of network topology on TCP performance," in Proceedings of IEEE Communications Society Communications Quality and Reliability Workshop (CQR 2008) (CD-ROM), May 2008.

⑧F. Inoue, T. Ito, H. Ohsaki, and M. Imase, "Implementation and evaluation of GridFTP automatic parallelism tuning mechanism for long-fat networks," in Proceedings of the 7th Asia-Pacific Symposium on Information and Telecommunication Technologies (APSITT 2008), pp. 232-237, Dec. 2008.

⑨Y. Sakumoto, H. Ohsaki, and M. Imase, "Increasing robustness of XCP (eXplicit Control Protocol) for dynamic traffic," in Proceedings of 50th IEEE Global Telecommunications Conference (GLOBECOM 2007), pp. 2025-2030, Nov. 2007.

⑩Y. Sakumoto, R. Asai, H. Ohsaki, and M. Imase, "Design and implementation of flow-level simulator for performance evaluation of large scale networks," in Proceedings of 15th Annual Meeting of the

IEEE International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS) 2007, pp. 166-172, Oct. 2007.

⑪Y. Sakumoto, H. Ohsaki, and M. Imase, "On XCP stability in a heterogeneous network," in Proceedings of 12th IEEE Symposium on Computers and Communications (ISCC'07), pp. 531-537, July 2007.

[産業財産権]

○出願状況 (計1件)

名称: iSCSI セッションの TCP コネクション数制御方法、iSCSI ホスト装置、および iSCSI イニシエータの構成 プログラム

発明者: 野本 義弘、大崎 博之、今瀬 真、井上 史斗

権利者: 大阪大学、NTT

種類: 特願番号: 2008-37692

出願年月日: 平成 20 年 2 月 19 日

国内外の別: 国内

6. 研究組織

(1) 研究代表者

大崎 博之 (OHSAKI HIROYUKI)

研究者番号: 00294166