

長距離 10Gbps 回線における TCP 中継器の評価

長谷川洋平[†] 地引昌弘[†]

近年、10Gbps 回線の普及も進み、長距離通信では TCP の性能がスループットを制限してしまう問題が顕在化してきている。これを解決するため、本研究では TCP 中継器を利用した長距離通信の高速化について取り組んでいる。本稿では 3320km の 10Gbps 回線を用いそのスループットの向上効果を評価する。今回の実験では、End-to-End の TCP に提案する TCP 中継器を組み合わせることで最大 13.8 倍にスループットを向上させた。また、TCP 中継器が 3320km の回線で 10Gbps のスループットを発揮することも確認した。

Evaluation of TCP Booster on long distance 10Gbps link

Yohei Hasegawa[†] and Masahiro Jibiki[†]

The rapid growth in link bandwidth has emphasized the throughput limitation of long distance TCP/IP data transfers. In this paper, we propose the TCP Booster (TCPB) to enhance TCP throughput via long distance links. With the TCPBs, TCP data transfer will be optimized for both terrestrial network and long distance link. We have evaluated the TCPB test bed system with 10Gbps 3320km link. We will show results that our test bed greatly improved TCP throughput. For example, TCP throughput using Windows XP hosts improved up to 13.8 times higher with the TCP Boosters. We also validated that our test bed boosted TCP throughput up to 10Gbps via 3320km cable.

1. はじめに

近年、インターネットの回線速度向上は著しく、10Gbps 回線の普及も進んでいる。端末においても既に 10Gbps のイーサネット回線が利用可能である。このように端末から長距離回線までの 10Gbps 通信環境が整いつつあり、長距離の大容量ファイル高速転送などの実現が期待されている。

一方、インターネットで用いられる TCP/IP は長距離通信においてスループットを発揮しにくいことが知られている。TCP のおよそのスループットは端末間の往復遅延 (RTT) に反比例し、また、同時にパケットロス率の平方根にも反比例し低下する[1]。このため、大きな RTT があり、かつパケットロスが発生する場合は、特にスループットが低下してしまう。

この問題を軽減させるため、いくつもの高速 TCP が提案されてきた[2,3,4,5]。これら高速 TCP は、TCP の送信レート制御を変更することで、通常の TCP よりも高いスループットの達成を目指すものである。しかし、これら高速 TCP では比較的高いスループットが得られるものの、一般的にアグレッシブな制御を実現しているため、パケットロス率の悪化を招く可能性もある。また、他の TCP トラヒックと競合した際などに、他 TCP の性能を阻害する可能性もある。このように多様な環境に適応し高い性能を発揮する TCP を実現することは難しい。

一方、ネットワーク内にて TCP を終端し中継することでスループットを向上させる方式なども研究されている[6,7]。文献[6]の研究は、TCP の制御方法を変更することに加え、ネットワーク内にて TCP のフィードバックループを分割することで、パケット再送による性能低下も抑え、より一層高いスループットの発揮を目指したものである。TCP 中継器を用いることで、TCP のスループット問題を軽減することが期待される。しかしながら、TCP は処理コストの高いプロトコルであることもあり、10Gbps の長距離回線において TCP 中継器を適用しスループットを向上させた例はこれまで報告されていない。

本稿では、これら問題を鑑み、ネットワークを特性毎に分解するよう TCP 中継器を利用する方法を提案し、3320km の 10Gbps 回線を用いて性能評価を行った結果を報告する。

以降、本稿は次のように構成される。2 章では長距離 TCP/IP 通信における性能問題と関連研究を説明する。3 章では、TCP 中継器による長距離通信のための通信方法を提案する。4 章では 3320km の通信回線を利用した TCP 中継器の性能評価結果を報告する。5 章では本稿を総括する。

[†] 日本電気株式会社
NEC Corporation

2. 長距離 TCP/IP 通信における性能問題

ネットワークの遅延とパケットロスに対して TCP のスループットが大幅に低下することは広く知られており、TCP のスループットを向上させる方法については長年にわたる研究が行われてきている。TCP のおよそのスループットは端末間の往復遅延 (RTT) に反比例し、また、同時にパケットロス率の平方根にも反比例し低下する。スループット B は簡単には次式のように表される[1].

$$B \approx \min\left(\frac{W}{d}, \frac{C}{d\sqrt{p}}\right)$$

ただし、 W は TCP の最大ウィンドウサイズ、 d は RTT、 p はパケットロス率、 C は定数である。

このスループットモデルの概形を図 1 に示す。 ($W=512\text{Kbyte}$, $C=\sqrt{3}/2$) このように、遅延とパケットロスが組み合わされた状況では、TCP のスループットはまったく発揮されない。近年の長距離回線では、FEC(前方誤り訂正符号)によってビットエラーが隠蔽されるため、パケットロスが発生しない長距離伝送を可能としている、しかし、この長距離リンクに接続されるネットワークでは一般に端末からのトラフィックが集約される過程でパケットロスが発生してしまい、結果として、端末間経路は、遅延とパケットロスの両方があるものになってしまう。

この問題を軽減させるため、いくつもの高速 TCP が提案されてきた。例えば、High Speed TCP[2], FAST TCP[3], CUBIC TCP[4], COMPOUND TCP[5]などがよく知られている。これら高速 TCP は、TCP の送信レート制御を変更することで、通常の TCP よりも高いスループットを達成することを目指すものである。

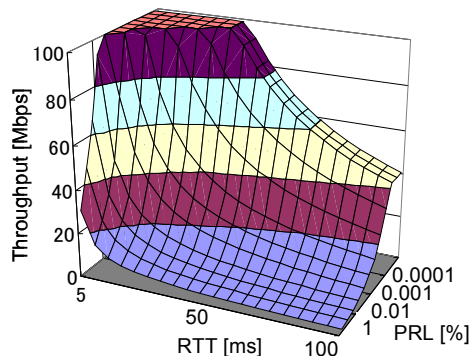


図 1 TCP のスループットと RTT, パケットロス率の関係

しかし、これら高速 TCP では比較的高いスループットが得られるものの、一般的にアグレッシブな制御を実現しているため、他の TCP トラフィックと競合した際などに、他 TCP の性能を阻害する可能性もある。例えば、CUBIC TCP はアグレッシブな制御により、ネットワークのパケットロス率が上昇し、他 TCP トラフィックのスループットが低下する可能性がある。また、逆に、他 TCP トラフィックとの競合を考慮しているため、理想的なスループットを発揮できない場合もある。このように多様な環境に適応し高い性能を発揮する TCP を実現することは難しい。

一方、ネットワーク内にて TCP を終端し中継することでスループットを向上させる方式なども研究されている[6,7,8]。もともと TCP 中継器は無線リンクなどビットエラーが発生するリンクが原因となるスループット低下を防ぐために用いられていたものである、これに対し、文献[6]の研究は、TCP の制御方法を変更することに加え、ネットワーク内にて TCP のフィードバックループを分割することで、パケット再送による性能低下も抑え、より一層高いスループットの発揮を目指したものである。TCP 中継器を用いることで、TCP のスループット問題を軽減することが期待される。しかしながら、TCP は処理コストの高いプロトコルであることもあり、10Gbps の長距離回線において TCP 中継器を適用しスループットを向上させた例はこれまで報告されていない。

3. 長距離回線のための TCP 中継器

TCP 中継器は、端末間の TCP コネクションをネットワーク内で分割し、データを転送していく。例えば、2つの TCP 中継器を用いることで、端末間のデータ転送は TCP1, TCP2, TCP3 の3つの TCP コネクションを介して行われることになる。図 3 には TCP 中継器を用いたパケットトランザクションの例を示す。この例では、コネクション開設を端末間で行った後、データ転送の際に TCP 中継処理を実施する例である。TCP 中継器は端末からは透過的に動作することができ、あたかも端末間で通信しているように利用できる。

本章では、これまで述べた問題を解決するため、長距離伝送回線をはさみ TCP 中継器を利用することを提案する。端末間の TCP1~3 はそれぞれ次のような部分ネットワーク区間を経由する。

TCP1, TCP3: パケットロスはあるものの遅延が小さい集線ネットワーク

TCP2: 遅延はあるもののパケットロスが発生しない長距離伝送ネットワーク

これによって、TCP1~3 が経験する RTT もしくはパケットロス率のどちらかが小さくなることになる。図 1 を参照すると、それぞれのスループットが指数的に向上することがわかる。これによって、端末間の通信スループットを大幅に向上させることができる。

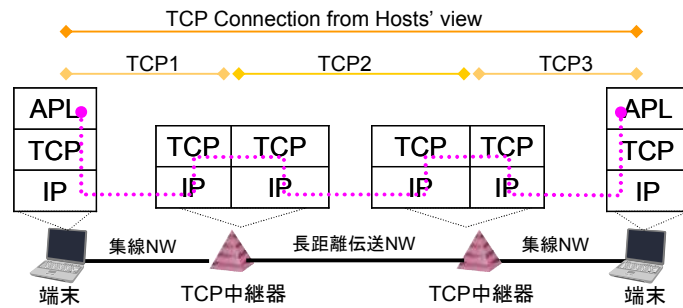


図 2 TCP 中継器を利用した通信の概要

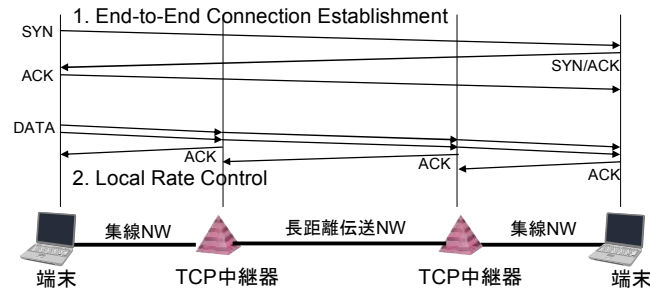


図 3 TCP 中継器によるパケットトランザクション

4. 性能評価

ここでは、3320km の長距離回線を用いた環境における TCP 中継器の性能評価について報告する。まず、評価環境を説明する。次に、提案する TCP 中継器を導入することで一般的な端末の TCP/IP 通信スループットがどれほど向上するのか、通信性能を評価する。その後、提案する TCP 中継器が 10Gbps のスループットを達成できるのか検証する。

4.1 評価環境

評価環境の概要を図 4 に示す。評価では、3320km もしくは 1740km の光ファイバケーブルを利用し、この回線終端装置(TE)をはさむように TCP 中継器を配置した。多数

の端末がこの長距離回線を利用することを想定し、集線過程におけるパケットロスを再現するためネットワークエミュレータを用意した。TCP 中継装置には NEC の Express サーバを用いた。TCP 中継器のスペック概要を表 1 に示す。また、長距離ケーブル、ネットワークエミュレータ、送信/受信端末の概要を表 2 に示す。

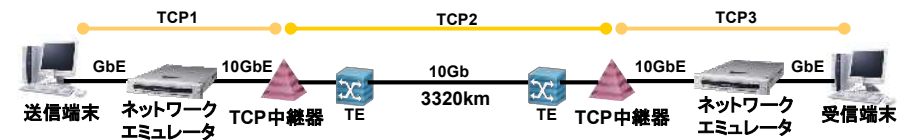


図 4 実験環境の概要

表 1 TCP 中継器のスペック

Machine	NEC Express server 5800/R120a-2
CPU	Intel XEON 5570 2.93 GHz × 2
RAM	12GB (DDR3-1066, 1GB × 12)
NIC	NEC 10GBASE-SR接続ボード N8103-123A (Chelsio S310E-SR)

表 2 実験環境のパラメタ

	項目	設定
長距離ケーブル	ケーブル長 (伝送遅延)	3320km (16.6ms)
	Bit Error	0 (Error free)
ネットワークエミュレータ	伝送遅延設定	1ms
	パケットロス率設定	0.1%
送信/受信端末	OS	Windows XP, Vista, Linux
	CPU	Intel Core2Duo(2.66GHz)
	NIC	1000BASE-T

4.2 通信性能評価

送信端末にそれぞれ Windows XP, Windows Vista, Linux を使用し, TCP 中継器を使用した場合と, 使用しない場合について, それぞれ端末間でスループットを計測した結果を表 3 に示す. 受信端末は Linux とした. また, 得られたスループットから DVD1 枚分 4.7GB のダウンロード推定時間を求め図 5 に示した.

Windows XP を用いた計測では, TCP 中継器により 13.8 倍のスループット向上効果が得られた. これは, DVD ダウンロード時間に換算すると, TCP 中継器を用いない場合では 43 分のところ, 3 分まで短縮されることになり, 体感速度の向上を大いに実感できるものである.

Windows Vista と Linux のを用いた場合には, TCP 中継器を利用すると 920Mbps 以上のスループットが計測された. 端末の回線インタフェースは 1Gbps のイーサネットであり, ヘッダ長のオーバーヘッドを考慮した最大スループットは 940Mbps ほどである. 実験では, これに近い性能を発揮できたことがわかる.

表 3 性能計測結果 (回線長 3320km)

端末 OS	TCP Window Size	TCP中継器 ON/OFF	Throughput [Mbps]	DVD Download Time [min.]	TCP中継器 ON/OFFの スループット比
Windows XP	64KByte	OFF	15.50	43.41	13.80
		ON	218.00	3.09	
Windows Vista	32MByte	OFF	136.00	4.95	6.78
		ON	922.00	0.73	
Linux	256KByte	OFF	115.00	5.85	8.13
		ON	935.00	0.72	

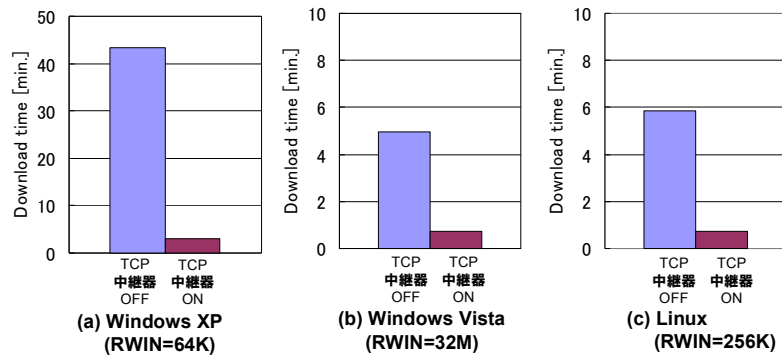


図 5 TCP 中継器の有無による DVD ダウンロード時間の違い

次に, 回線長の違いによるスループットの変化を評価する. ここでは, 1740km と 3320km の光ファイバをそれぞれ使用し, ネットワークエミュレータでは 5ms の遅延を付加した. つまり, 全ての RTT のうち 20ms は 2 台のネットワークエミュレータにより付加された遅延である. またネットワークエミュレータでは 0.1% のパケットロスを発生させた. 図 6, 図 7 は, 送信端末に Windows XP, Linux をそれぞれ用いた場合の結果を示す. 破線は TCP 中継器を用いない場合の推定スループットである.

どちらの実験結果も, TCP 中継器を利用した場合は, 回線長によらず, 短距離通信と同等のスループットを達成できたことがわかる. 図 6 の実験では, TCP 中継器を用いた場合のスループットは, 送信端末から TCP 中継器までの TCP (RTT=10ms) が発揮するスループットとほぼ等しいことがわかる. 図 7 では, 実測値のほうがスループットモデルよりも高い性能が得られており, Linux の TCP はわずかなパケットロスがあった場合でも性能が高いことを示している.

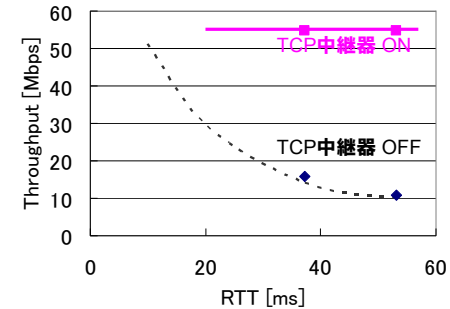


図 6 回線長を変化させた場合のスループット(Windows XP)

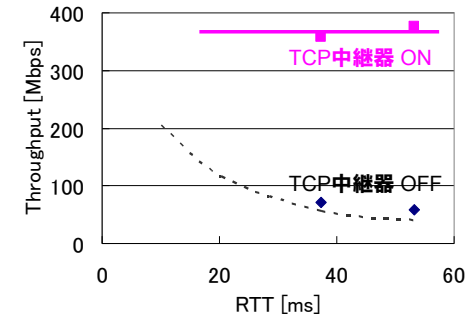


図 7 回線長を変化させた場合のスループット(Linux)

4.3 長距離 10Gbps 通信の検証

TCP 中継器が長距離回線を用いても 10Gbps の通信性能を発揮できることを検証するため、10Gbps 回線のみを用いた実験を行った。ここでは、ネットワークエミュレータ間で TCP トラフィックを発生させ、このスループットを計測している。

結果、TCP 中継器が 3320km 回線にて 10Gbps のスループットを達成できることを確認した。10Gbps を達成した結果の例を図 8 に示す。ここでは、TCP 中継器の送信ウィンドウサイズを 3.3MByte としている。コネクション数を変化させた。1 コネクション、2 コネクションの場合に 10Gbps が達成できていないのは、ウィンドウサイズが十分大きくなかったため、これを大きく設定すると、1 コネクションでも 10Gbps を達成することができたことも付け加えておく。

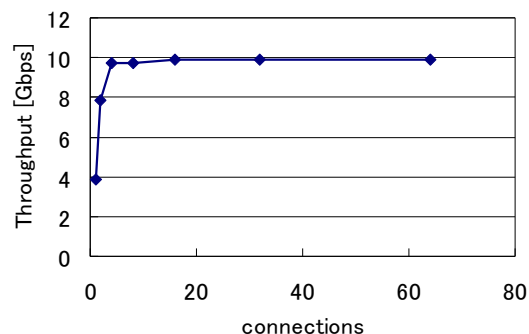


図 8 TCP 中継器を用いた 10Gbps スループットの例

5. おわりに

本稿では、ネットワークを特性毎に分解するよう TCP 中継器を利用する方法を提案し、3320km の 10Gbps 回線を用いて性能評価を行った結果を報告した。

実験では、一般的な端末間 TCP/IP 通信に提案する TCP 中継器を組み合わせることで最大 13.8 倍にスループットを向上させた。また、TCP 中継器が 3320km の回線で 10Gbps のスループットを発揮することも確認した。

今後は TCP 中継器を、多様な環境に対応させるべく検討を進める予定である。例えば、TCP 中継器を通過するが、IP フォワードされる背景トラフィックへの影響などについても検討する。

参考文献

- 1) J. Padhye, V. Firoiu, D. Towsley, J. Kurose, "Modeling TCP Throughput: A Simple Model and its Empirical Validation," Proc. ACM SIGCOMM, SEP 1998.
- 2) S. Floyd, "HighSpeed TCP for large congestion windows," RFC3649, IETF, DEC 2003.
- 3) C. Jin, D. Wai, and S. Low, "FAST TCP: Motivation, architecture, algorithms, performance," Proc. IEEE INFOCOM, vol. 4, pp.2490-2501, MAR 2004.
- 4) I. Rhee and L. Xu, "CUBIC: A new TCP-friendly high-speed TCP variant," Proc. PFLDNet, 2005.
- 5) K. Tan, J. Song, Q. Zhang and M. Sridharan, "A Compound TCP Approach for High-speed and Long Distance Networks," Proc. IEEE INFOCOM, APR 2006.
- 6) T. Murase, H. Shimonishi, and Y. Hasegawa, "TCP overlay network architecture," Proc. Comm. Conf IEICE'02, B-7-49, SEP, 2002.
- 7) Y. Liu, Y. Gu, H. Zhang, W. Gong, and D. Towsley, "Application Level Relay for High-bandwidth Data Transport," Proc. GridNets 2004, OCT 2004.
- 8) 長谷川洋平, 村瀬勉, "TCP 中継ノードの高速プロトコル処理方式と性能評価," 信学会ソサイエティ大会, SEP 2003.