

高速環境におけるTCPの実装評価と考察

1 V-5

近藤栄一 渥美幸雄
(株) 超高速ネットワーク・コンピュータ技術研究所

1. はじめに

近年、回線の高速化、マシンの性能向上に対して、実効スループットの向上が求められている。最も多く用いられているプロトコルであるTCPについても対応が求められており、昨年、RFC2018として採番されたTCP Selective Acknowledgement Options (以下SACKと記す)も、その1つである。

本稿では、BSD/OS下での高速通信環境を作成し、SACKを含むTCPの基本特性を評価したので、その結果を報告する。

2. 評価システム

本評価システムの作成にあたり、TCPの動作が明らかにされている[1]、BSD/OSをベースとした。高速な通信I/Fは155M/ATMとし、対応するドライバを移植した[2]。また、高速対応TCPとしてSACKをサポートするよう改造した[3]。これらにより、構築された評価システム構成を図1に示す。

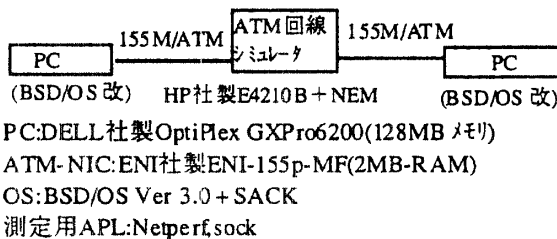


図1 評価システム構成

(1) 高速I/F

エンドシステムのLAN-I/Fとして、コストパフォーマンスの高い、高速なEthernetが注目されている。これに対し、高速WAN環境におけるI/Fとしては、ATMが有望とされている。ここでは、後者を想定し、155M/ATMを用いることとした。RFC1577 (Classical IP and ARP over ATM) によれば、IPパケットをAAL5でカプセル化する場合、IEEE 802.2 LLC/SNAPで定義される識別子を挿入することになっているが、本ATM-NICドライバではAAL5に直接IPパケットを設定するものとなっている。

TCP evaluation in high speed network

Eiichi KONDOH, Yukio ATSUMI

Ultra-high Speed Network and Computer Technology Laboratories

(2) SACK

従来のTCPでは、ヘッダの構成要素であるAcknowledgement Numberにより送達確認を行い、パケット紛失が発生した場合、原則として以降のパケットすべてを送信するようになっている。これに対し、SACK-TCPでは部分的な送達確認を行うことで、確認された部分を除いた再送を行うことが可能となり、結果として帯域を有効に利用することができる。SACK実装に関し、RFC2018では既存のアルゴリズムは保存されるべきと記されており、本改造でも再送及び輻輳制御に関するTCP-Renoアルゴリズムはそのまま継承している。

本環境にてスループット測定を行ったところ、最大134Mbpsが得られており、エンドシステムの性能として、ボトルネックにはならないと言える。

3. 評価結果

(1) TCPセグメントサイズとスループット

TCPセグメントサイズが小さい場合、パケット単位の処理が多くなり、スループットが上がらない。図2は回線遅延時間無、セルロス無の環境にてTCPセグメントサイズを変化させたときのスループットを示したものである。この結果から、5KB以上のセグメントサイズなら本環境におけるスループットには、ほぼ影響がないことが分かる。また、BSD/OSにおけるTCPセグメントサイズは、ネットワークI/FのMTUサイズに合わせて設定されるが、RFC1577で記述されている、ATM-MTUサイズのデフォルトの最大値、9180Bも本環境条件では妥当な値である。

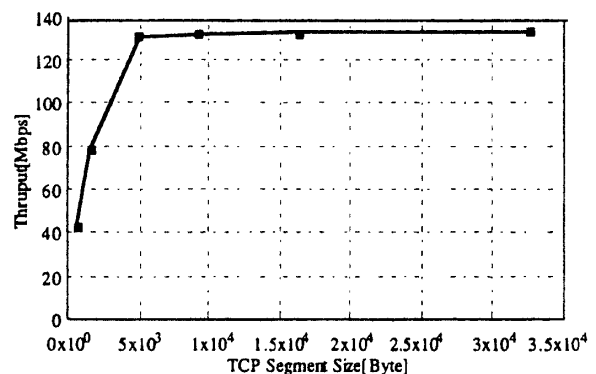


図2 TCPセグメントサイズとスループット

(2) ソケットバッファサイズ

TCPでは送達確認を行うため、回線遅延時間の影響が避けられない。従って、高遅延環境であるWANでの高スループット化を想定した場合、多くのバッファを確保し、ウィンドウサイズを大きく設定する必要がある。TCPヘッダにおけるウィンドウサイズは64kBまでであるが、BSD/OSではウィンドウスケールオプション[4]がサポートされており、本評価システムでは、246kBまで設定可能である。図3は、バッファサイズ毎に遅延時間とスループットを測定したもので、バッファサイズが大きい程、回線遅延の影響によるスループット低下が少ないことが分かる。

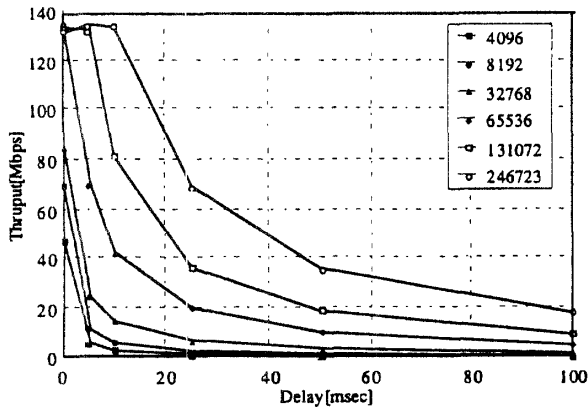
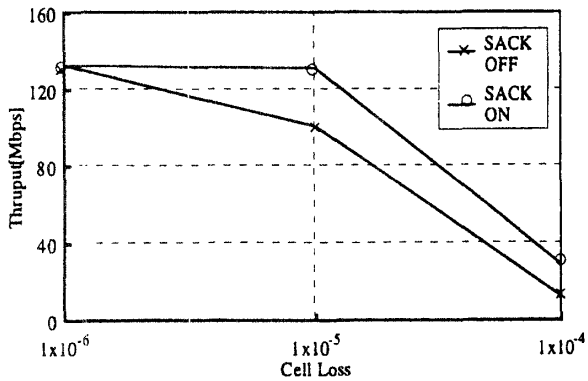


図3 バッファサイズ毎の往復遅延時間とスループット

(3) SACKの効果

セルロスが発生させた場合のSACK有/無のスループットを比較したものが図4である。

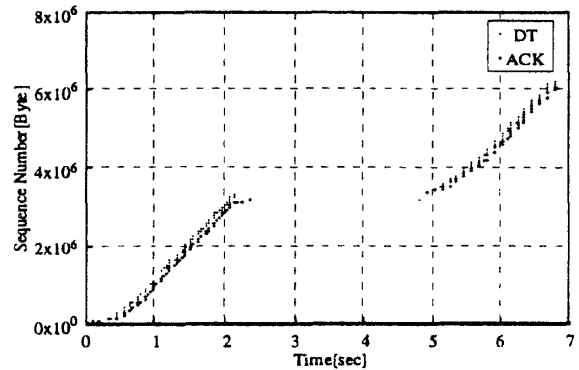


測定条件：セルロス発生パタン：指数分布，ATM-MTUサイズ：9180Byte
 バッファサイズ：246kB，回線遅延：なし，測定時間：20sec

図4 セルロス発生時のスループット

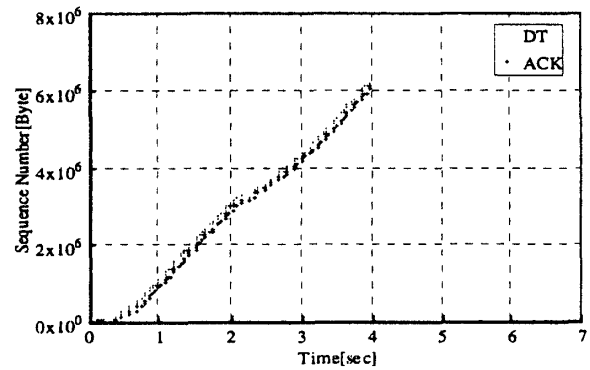
この結果から、(i)セルロス $1E-5$ 以下なら、SACK有の場合はスループット低下がほとんどない。(ii) $1E-5$ で約30%、 $1E-4$ で約100%の改善が得られる。ことがわかる。性能差が大きいのは、複数パケット紛失事象が多いためと考えられるので、具体的なシーケンスの違いを明らかにする。1パケット

の間隔で2パケットのみ紛失したケースのDTとACKのSequence Numberを示したものが、図5(SACK無)、図6(SACK有)である。SACK無のケースでは2パケット目の紛失に対するFastRetransmitが行われずにタイムアウトによる再送を行うのに対し、SACK有のケースでは、2パケット目の再送後も、タイムアウトを待たずに送信を再開していることが分かる。



測定条件：往復回線遅延：100msec，受信側にて廃棄フル使用
 他は、図4と同様

図5 2パケット紛失時のシーケンス(SACK無)



測定条件：図5と同様

図6 2パケット紛失時のシーケンス(SACK有)

4. おわりに

今回、高速環境におけるTCPに関する評価システムを作成し、その基本特性を評価した。

今後、SACK-TCPを中心とする詳細な動作を明らかにし、高速環境におけるTCPの高スループット化に向けた検討を行ってゆく予定である。

[1] Gary R. Wright, W. Richard Stevens,
 TCP/IP Illustrated, Volume 2 "The Implementation"
 [2] Internet上に公開されている、以下をベースとした。
<http://www.cerc.wustl.edu/pub/chuck/bsdnetm/wucs.html>
 [3] Internet上に公開されている、以下をベースとした。
<ftp://daedalus.cs.berkeley.edu/pub/tcpsock/>
 [4] RFC1323: "TCP Extensions for High Performance"