

x-kernel 上へ実装した XTP プログラムの性能評価

1 T-7

三宅 優 加藤 聡彦 鈴木 健二

国際電信電話株式会社 研究所

1. はじめに

高速ネットワーク上での高機能・高性能な通信プロトコルとして XTP^[1]が提案されている。筆者等は、FDDI 用の通信ボードや市販のワークステーション上で動作させるために、x-kernel をプロトコル実装のプラットフォームとして採用し、その上に XTP プログラムの実装を行ってきた^[2]。本稿では、国際間等の伝送遅延が大きなネットワークでの利用も考慮にいれ、遅延発生装置を使用して、新たに作成した XTP プログラムの性能評価を行ない、既存の TCP との比較を行なったので、以下に報告する。

2. x-kernel 上への XTP の実装概要

XTP の実装は、x-kernel 内に存在する複数のプロトコルオブジェクトの1つとして行なわれている。各プロトコルオブジェクトは、x-kernel のコンパイル時に設定したプロトコルグラフによって関連付けられる。筆者等の実装においては、XTP の下位プロトコルに FDDI とイーサネットを想定し、LLC および ETH(イーサネットプロトコル) 上で動作可能である。XTP とこれらの下位プロトコルオブジェクトの間には、ネットワークアドレスにしたがって適切なインタフェースを選択する機能を持つ、VNET(Virtual Network Protocol) と呼ばれるプロトコルオブジェクトが存在する。

3. 評価方法

3.1 評価システムの構成

性能評価においては、XTP プログラム、x-kernel 上へ実装された TCP、ワークステーションの OS へ実装された TCP の3つを比較した。このため、XTP プログラムとしてワークステーション上へ実装されたものを用いることとした。また、高速 LAN であることと、遅延発生装置の便宜を考慮して、性能評価で使用するネットワークとして ATM LAN を用いた。図 1 に、今回の評価に使用したシステムの構成を示す。評価は、評価対象となるプロトコルそれぞれに対して、2 台のワークステーション間の通信スループットを計測することにより行なった。評価では、SunOS 4.1.2 が動作する SPARCserver 670MP および SPARCstation IPC を使用し、この2つを ATM LAN (140Mbps) で接続した。評価において、伝送遅延の影響を調べるために、ATM 交換機の DS-3 (45Mbps) のインタフェースに遅延発生

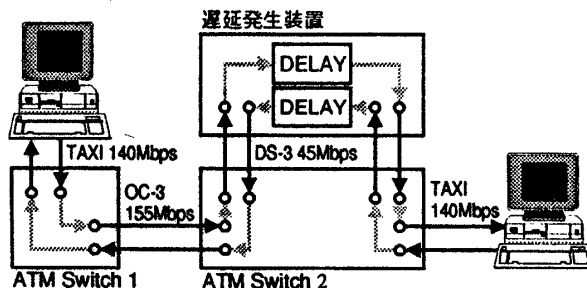


図 1: 評価に使用したシステムの構成

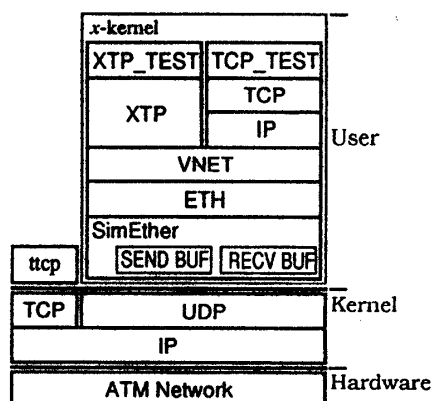


図 2: 性能評価用のプログラム構成

装置を接続し、ワークステーション間の送受信データが遅延発生装置を経由するように PVC を設定した。

3.2 ソフトウェア構成

図 2 に、評価に使用したプログラム構成を示す。x-kernel は、SunOS 上でプロセスとして動作し、プロトコルの処理を行う。ネットワークインタフェースとして ETH(イーサネットプロトコル) が存在し、デバイスドライバに相当する SimEther(Simulated Ethernet Driver Protocol) が、SunOS の UDP を利用してデータの送受信を行なう。ただし、イーサネット本来の MTU(Maximum Transmission Unit) サイズでは、データの送受信のためのシステムコール部分が全体のスループットのボトルネックになってしまうため、ATM LAN の MTU に適応させ ETH の MTU サイズを 9000 バイトに変更した。また、SimEther 内で確保される送受信バッファもこれに伴って増加させた。なお、データの送受信の際には、この送受信バッファでデータのコピーが行われている。

3.3 評価用上位プロトコルプログラム

x-kernel に実装されているプロトコルを評価するため

“Evaluation of XTP program on x-kernel”
Yutaka MIYAKE, Toshihiko KATO and Kenji SUZUKI
KDD R & D Laboratories

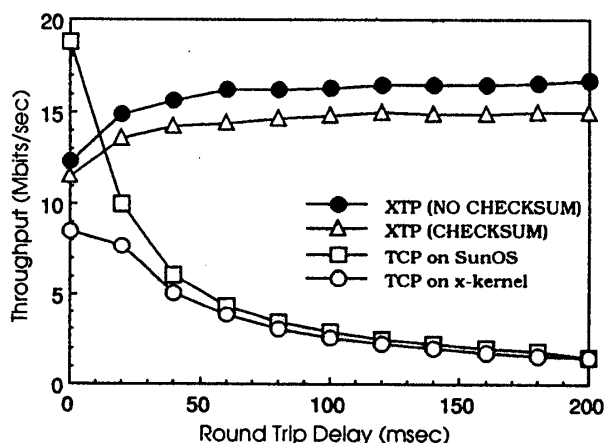


図 3: 往復遅延時間とスループットの関係

に、XTP_TEST および TCP_TEST プロトコルを作成した(図 2 参照)。これらのプロトコルプログラムは、対象となる各プロトコルの送受信スループットを計測するために、2つのワークステーション間で片方向にメモリ上のデータを連続して転送する。

SunOS 内に実装されている TCP の評価には、スループット計測のために一般的に用いられている `ttcp` を使用した。

3.4 測定方法

測定は、`x-kernel` 上に実装した XTP において、チェックサムを使用するモードと使用しないモードの 2 種類、`x-kernel` 上に実装された TCP、そして、SunOS に実装されている TCP の計 4 種類について、遅延発生装置を使用して挿入する往復遅延時間を 0 から 200msec の間で 20msec づつ変化させて行なった。スループット性能に大きな影響を与える送受信バッファサイズは、TCP の場合、SunOS の最大値である 53K バイト、XTP の場合、1M バイトとした。

4. 結果と考察

図 3 に、スループット測定によって得られた結果を示す。

4.1 スループット性能の比較

今回評価を行なった XTP は、

- (1) UNIX 上の 1 プロセスである `x-kernel` 内で動作している。
- (2) 送受信バッファの最大値が 50K バイト程度の UDP のソケットインタフェースを使用している。
- (3) プロセス内でネットワークインタフェースをエミュレートするためにコピーを行なっている。

等の理由により、カーネル内に実装されているプロトコルに比べ、高スループットを得るには不利な点が多い。そのため、最高スループットが 17Mbits/sec 程度と、SunOS に実装されている TCP の最高スループットである約 19Mbits/sec を下回る値しか得られなかった。しかし、同じ条件である `x-kernel` 上の TCP との比較で

は、最高スループットの値が、約 2 倍程度高くなっている。この差は、主に各プロトコルの誤り回復機構等の違いや、`x-kernel` 上の TCP が 4.3BSD のソースを移植しているのに対し、XTP は最初から `x-kernel` を対象に作成されたため、XTP の方が効率良く動作する等により生じたと考えられる。

4.2 往復遅延時間の増加に伴うスループットの変化

`ttcp` を使用して得られたスループットの最大値は、約 19Mbits/sec を得られているが、遅延が増加するに従い、スループットが急激に減少している。`x-kernel` 上に実装されている TCP の場合も、同様の傾向が見られる。

このスループットの減少は、TCP のウィンドウサイズの最大値が 64K バイトであるためである。SunOS の場合は、BSD の実装を継承しているため、最大値は 53K バイトとなる。通信路に十分な帯域が確保されていても、ウィンドウサイズが小さい場合には、往復遅延時間が大きくなるほどスループットが低下することになり、これが結果として現れたことになる。ウィンドウサイズの拡大は、RFC 1323 による TCP オプションで実現可能であるが、今回評価した TCP には実装されていない。

XTP では、64K バイト以上のウィンドウサイズを確保できる。今回の測定では、受信バッファとして 1M バイトを確保しており、この値に連動して最大ウィンドウサイズも 1M バイトとなる。ウィンドウサイズが大きすぎると、輻輳などによりパケットロスが発生する可能性も高くなる。しかし、結果より、このような比較的大きな値に対しても、本実装に導入した輻輳回避機構^[2]によりフロー制御が正常に動作し、遅延が増加してもスループットが低下せずに動作する事が確認された。

4.3 XTP のスループット変化

XTP において、挿入した往復遅延時間が 0msec の場合には、他の遅延時間の場合と比べてスループットが減少している。これは、プロトコル内で使用する往復遅延時間の計測回数^[2]が、往復遅延時間が短いほど多くなり、パケットの送受信が増加するためである。この部分のアルゴリズムの改良が必要である。

5. まとめ

本稿では、`x-kernel` 上に実装した XTP プログラムの評価を行ない、その結果について述べた。本ソフトウェアは、TCP と異なり、遅延の増加に影響なく、高い通信スループットを維持できる。また、遅延の小さい場合でも同じ条件の TCP と比較において、スループット向上を達成した。最後に日頃御指導頂く KDD 研究所 浦野所長、眞家次長に感謝します。

参考文献

- [1] "XTP Protocol Definition Revision 3.6," *Protocol Engines Inc.*, PEI 92-10, 11 January 1992.
- [2] 三宅、加藤、鈴木、"高速通信プロトコル XTP の `x-kernel` 上への実装," 第 49 回情報大全, 6C-3, September 1994.