

マルチメディアコンテンツ転送向け 高性能 TCP/IP 通信ボードの構成と評価

長田 孝彦[†] 東海林 敏夫^{†,☆}
山下 博之^{†,☆☆} 塩川 鎮雄^{†,☆☆}

情報のマルチメディア化にともない、ネットワーク上を流れるデータ量は急激に増大している。100 Mbit/s 以上の高速回線が使用可能となってきたが、マルチメディア情報を扱うアプリケーションはその帯域を十分に利用できていない。それは、コンピュータ/端末における通信プロトコル処理に多大の時間を要するためである。我々は、従来のソフトウェア処理に代えてボード上のハードウェアで TCP/IP を高速処理する、ワークステーション拡張スロット装着用の ATM 通信ボードを開発した。同ボードでは、TCP/IP 系プロトコルの分析に基づいて得られた、TCP、IP ヘッダの編集、解析およびチェックサム生成、検証の各基本機能と、イベント、タイマおよびバッファといった通信プロトコル処理一般のリソース管理機能とを有する TCP/IP 処理 LSI を搭載する。この LSI とファームウェアとが、並列動作することにより、TCP/IP の高速処理を実現した。さらに、標準のバスインタフェースによる省スペース（1 スロット）実装、ファームウェア制御の採用によるサポート通信プロトコルの拡張性、コンフィグレーションの変更による既存アプリケーションへの影響（再コンパイル）回避、通信パス振り分け機能の挿入による既存 TCP/IP との共存、という特徴を有する。静止画像転送アプリケーションで 116 Mbit/s、ボードとしては 155 Mbit/s ATM のユーザ限界スループット（135 Mbit/s）を達成した。

Structure and Evaluation of High Performance TCP/IP Communications Boards for Multimedia Contents Transmission

TAKAHIKO NAGATA,[†] TOSHIO SHOJI,^{†,☆} HIROYUKI YAMASHITA^{†,☆☆}
and SHIZUO SHIOKAWA^{†,☆☆}

The size of data through communication networks is dramatically increasing with the progress of multimedia age. Multimedia applications can not fully use the bandwidth of high speed networks over 100 Mbit/s, because protocol processing gives heavy loads to the computers/terminals. The authors have developed TCP/IP/ATM communications boards for workstations that shift the protocol processing to hardware on the boards from software on the workstations. The board has a TCP/IP processing LSI that has basic functions which include TCP/IP header editing/analysis and checksum calculation, and has control functions for general protocol processing resources such as event, timer and buffer. High speed protocol processing has been realized by the TCP/IP processing LSI and firmware that run in parallel. Moreover, the board has following characteristics: small space (1 slot) implementation using standard bus interface, protocol extendibility by firmware control, avoidance of influence (recompile) on existing applications, and coexistence with resident TCP/IP. Using picture transmission application, the transmission performance (including application) was 116 Mbit/s, and that of the board itself reached user limit throughput (135 Mbit/s).

1. はじめに

近年のコンピュータネットワークの進展と伝達される情報のマルチメディア化にともない、ネットワーク上を流れる情報量は急激に増大している。また、これらの情報を転送するネットワークとしては、毎秒 100 メガビットを超える高速回線を使用できるようになってきた。しかしながら、現在のところこのような高速回線の帯域を十分に利用できていない。これは、各コンピュータ/端末において TCP/IP (Transmission Control Protocol/Internet Protocol) 等の通信プロ

[†] NTT 情報通信研究所

NTT Information and Communication Systems Laboratories

[☆] 現在, NTT アドバンステクノロジー株式会社情報ネットワークング事業部

Presently with Information Networking Division, NTT Advance Technology Corporation

^{☆☆} 現在, NTT 研究開発推進部

Presently with R&D Management Department, NTT

^{☆☆☆} 現在, NTT ソフトウェア株式会社ニュービジネス事業本部

Presently with New Business Sector, NTT Software Corporation

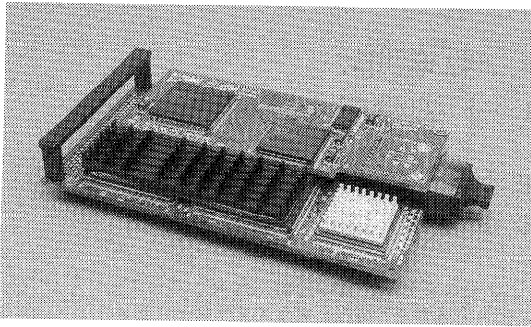


図1 TCP/IP/ATM 通信ボード
Fig.1 TCP/IP/ATM communications board.

トコルをソフトウェアで処理する従来の方法では、プロトコル処理に多大の時間を要し、データの転送効率が低下するためである¹⁾。そこで、アプリケーションが高速回線の帯域を十分に利用可能とするため、通信プロトコル処理の高速化が要求されている。

この要求に応える方法としては、高速プロトコルの新規開発、既存プロトコルの高性能実装が考えられる。前者の例としては、XTP (Xpress Transport Protocol)²⁾が高速プロトコルの標準候補として開発されているが、広く普及するには至っていない。また、後者の例としては、TCP/IP 処理のハードウェア化により高速化を図った例が報告されている³⁾。しかし、システム全体としては、既存の装置に容易に装着できる形態を有していない。既存の装置の変更を必要とせず、インターネット/イントラネット等で広く使用され、事実上の標準プロトコルとなっている TCP/IP の高速処理を可能にするシステムが望まれていた。

そこで今回、我々は、TCP/IP をボード上で高速処理する、ワークステーション拡張スロットに装着する通信ボードを開発した(図1)。この通信ボードは、ワークステーションにドライバプログラムをインストールするだけで使用可能であり、155 Mbit/s ATM 回線の帯域を有効に利用できる環境を容易に提供する。さらに、この通信ボードを使用することにより、TCP/IP の高速処理を可能にするだけでなく、TCP/IP プロトコルをボード上で処理するためワークステーションの CPU 負荷が大幅に軽減され、アプリケーションにより多くの CPU 処理能力を振り向けられ、システム全体の性能向上を図ることができる。

本論文では、開発した高性能 TCP/IP/ATM 通信ボードについて、2章で、TCP/IP 通信ボードに対する要求条件を述べ、3章で、同条件を満たすためのシステム設計について記述する。その後、4章で実装性能を評価し、5章で実用性に関して評価し、最後に、6

章でまとめる。

2. 通信ボードを含むシステムに対する要求条件

高性能 TCP/IP 通信ボードの設計にあたり、次の条件を満足する必要がある。

(1) 高速回線の全帯域のアプリケーションへの提供
今回開発した通信ボードは、155 Mbit/s ATM 回線向けである。ATM のヘッダ等のオーバーヘッドを除くと、アプリケーションが使用可能な帯域は、135 Mbit/s である。よって、TCP/IP の上位に位置するアプリケーションに対し、135 Mbit/s の帯域の提供が目標となる^{*}。

(2) 標準のバスインタフェースによる省スペース実装

広範なマルチメディアサービスに適用し、高性能のシステムを経済的に構築するため、既存システムに変更を強要することなく通信ボードを装着可能とし、さらに動画等のマルチメディア処理用ボードと共存できることが重要である。

(3) サポート通信プロトコルの拡張性

新たなマルチメディアサービスの要求に伴う通信プロトコルの進化等に対応するため、その機能の追加や変更が容易に行える構成とする必要がある。

(4) 既存アプリケーションの利用

通信ボードの容易な導入のため、FTP や Telnet 等既存 TCP/IP アプリケーションの変更および再コンパイルなしでの走行を保証することが重要である。

(5) 既存 TCP/IP との共存

システムの柔軟な構築を可能にするため、既存 TCP/IP を使用する Ethernet 等を経由する通信との同時動作の確保が必要である。

3. 通信ボードの設計

3.1 要求条件を満たす方策

前章で述べた各要求条件を満たすため、以下の設計を行った。

(1) ハードウェアにより TCP/IP 処理の高速化を実現した。ハードウェア化にあたり、TCP/IP 系プロ

^{*} アプリケーションが使用可能な帯域の算出方法：

1. SDH のオーバーヘッド (26 B/27 B)、2. セルヘッダのオーバーヘッド (48 B/53 B)、3. パケット化のオーバーヘッド (9140 B/9216 B) [ただし、ペイロード：9140 B, LLC/SNAP：8 B, AAL5 トレイラ：8 B, TCP ヘッダ：20 B, IP ヘッダ：20 B, セル分割フラグメント：20 B とした。]

以上を差し引くことにより、使用可能帯域は 134.5 Mbit/s となる。

トコルの分析に基づき、具体的な機能を LSI 化した。TCP/IP 処理 LSI は、TCP/IP 処理のアクセラレータとして動作し、定型的なアルゴリズムで実行される TCP および IP ヘッダの編集、解析およびチェックサム生成、検証の各基本機能を有する。このうち特に、チェックサム生成/検証は、TCP/IP 処理高速化のネックとなっていたものである¹⁾。また、通信プロトコル処理一般に共通して必要な実行頻度の高い処理として、イベント、タイマおよびバッファというボード全体の制御リソース管理機能を有する。

ボード上の CPU を用いて TCP/IP 処理を行うだけでも、ホスト装置上の処理との並列化により性能の向上が可能である。しかし、要求条件の達成には、高性能の CPU が必要となる。これは、コストおよび実装の面で問題があった。もちろん、CPU の性能-価格比の向上にともなってこの問題は低減するが、サポートすべきネットワークも高速化するため、将来にわたって TCP/IP 処理の LSI 化は必要であると判断した。

(2) SUN 互換ワークステーション拡張スロット装着用ボードとすることにより、既存システムの変更を行うことなく高速プロトコル処理エンジンの利用を可能とした。拡張スロット装着用であるため、ハードウェアを構成するうえでスペース上の制限が生じうるが、TCP/IP 処理 LSI を始め、周辺回路の ASIC 化により、1 スロットサイズで実現した。

(3) 要求条件 (1) を満足する範囲内で、プロトコル処理をファームウェアで行うこととした。すなわち、プロトコル処理のうち、高速化が要求される処理および定型的な処理を TCP/IP 処理 LSI で、可変要素の大きい処理をファームウェアで分担することにより、サポートプロトコルの追加や変更に対応可能とした。

(4) Solaris 環境において、ドライバの対外部および内部インタフェースを OS 標準の TLI⁴⁾ に統一するとともに、ネットワークコンフィグレーション用データベースファイル (/etc/netconfig) 中のドライバ指定を既存 TCP/IP ドライバから新規開発ドライバに書き換えることにより、既存アプリケーションの実行形式ファイルそのままの使用を可能とした。

(5) アプリケーションとドライバとの間に通信パスの動的な振り分けを行う機能モジュールを挿入することにより、既存 TCP/IP を使用する Ethernet 経由の通信との同時動作を可能とした。

3.2 通信ボードの構成

(1) 通信ボード全体

TCP/IP/ATM 通信ボードの構成を図 2 に示す。本

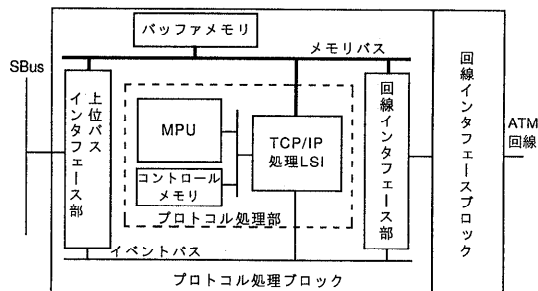


図 2 TCP/IP/ATM 通信ボードのブロック構成

Fig. 2 Block diagram of TCP/IP/ATM communications board.

通信ボードは、プロトコル処理ブロックと回線インタフェースブロックの大きく 2 つのブロックで構成した。1 スロットサイズ化のため、プロトコル処理ブロックはマザーボードで、回線インタフェースブロックはドーターボード上で実現した。本通信ボードは、33 MHz のクロックで動作する。

プロトコル処理ブロックは、メモリバスとイベントバスの 2 種類のバスを有する。いずれもデータは 32 bit 幅である。送受信データは、メモリバスを、制御信号は、イベントバスを経由する。バッファメモリ (4 MB) には、送受信ヘッダ、データおよびプロトコル処理用制御情報を格納する。コントロールメモリ (2 MB) には、MPU 上で走行するファームウェアを格納する。また、プロトコル処理ブロックは、プロトコル処理部、上位バスインタフェース部および回線インタフェース部から成る。

プロトコル処理部は、プロトコル処理用制御プログラム (ファームウェア) が走行する汎用 RISC プロセッサ (MPU)、および MPU のコプロセッサとして動作する TCP/IP 処理 LSI により構成している。なお、MPU として、SPARClike (MB86930: 33 MHz) を採用した。

上位インタフェース部および回線インタフェース部は、サポートする装置および回線のバリエーション展開に対応して切り出しが可能である。すなわち、今回開発した通信ボードは、ワークステーション用 SBus および 155 Mbit/s ATM 回線対応であるが、上位バスとして PCI バス等に、また、回線として ATM 以外に Fast Ethernet 等に、それぞれ容易に対応可能である。

回線インタフェースブロックは、汎用の SAR-LSI、FRAMER-LSI、PMD、コントロールメモリを有する。

(2) TCP/IP 処理 LSI

本 LSI のブロック構成および各ブロックの機能を、図 3

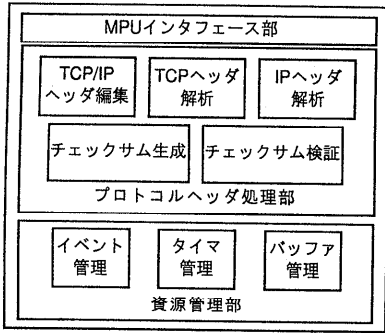


図3 TCP/IP処理LSIの機能ブロック構成

Fig. 3 Functional block diagram of TCP/IP processing LSI.

表1 TCP/IP処理LSIの機能概要

Table 1 Functions of TCP/IP processing LSI.

ブロック名	機能概要	
ヘッダ処理部	TCP/IPヘッダ編集	MPUからの情報に基づき、TCP、IP等のヘッダを編集する。
	チェックサム生成	チェックサム演算を行い、結果をヘッダに書き込む。
	チェックサム検証	チェックサム検証を行い、結果をイベント管理へ出力する。
	TCPヘッダ解析	TCPヘッダのフラグとパラメータ値のチェックを行う。
	IPヘッダ解析	IPヘッダのフラグとパラメータ値のチェックを行う。
資源管理部	イベント管理	イベントバス経由で受領したイベントを管理、MPUへイベントを通知する。
	タイマ管理	タイマ更新とタイムアウトチェックを行い、タイムアウトイベントをイベント管理へ出力する。
	バッファ管理	バッファメモリ上の送受信バッファブロックのアドレスを管理する。

および表1に示す。

MPUインタフェース部は、MPUに接続され、MPUの入出力制御および割り込み制御を実行する。

プロトコルヘッダ処理部は、プロトコルのヘッダに係わる処理を行う。そのほかに、イベントバスとメモリバスの競合制御も本LSIで行う。

特に多くの時間を費やしていることが明らかにされているチェックサム演算をTCP/IP処理LSIで行うことにより、ファームウェア処理に比べ数倍性能が向上している(図4)。

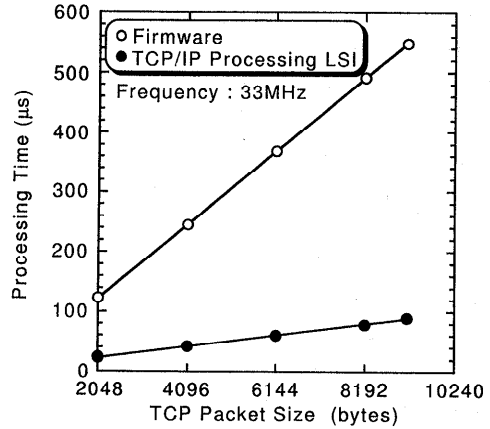


図4 TCPチェックサム演算処理時間の比較

Fig. 4 Comparison of TCP checksum processing times between TCP/IP processing LSI and firmware.

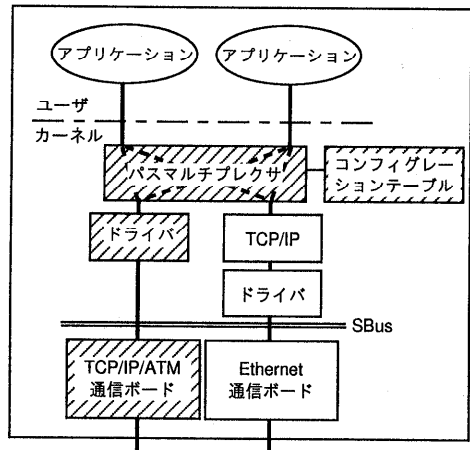


図5 ドライバの位置づけ

Fig. 5 Position of driver (system configuration including driver and board).

(3) ファームウェア

ファームウェアは、主に3種のタスク〔プロトコル処理(TCP/IPタスク)、上位バスインターフェース制御(BIFタスク)、回線インターフェース制御(LIFタスク)]により構成され、プロトコル処理に加え、ボード上の処理全体の制御も行う。これらは、それぞれ対応するイベントにより起動される。なお、TCP/IPのみでなく、UDP、ICMPもサポートしている。

(4) ドライバ

ドライバの位置づけを図5に示す。ドライバとアプリケーションとの間の機能モジュール(バスマルチプレクサ)が、コネクション確立時に設定するコンフィグレーションテーブルに基づき、TCP/IP/ATM通信

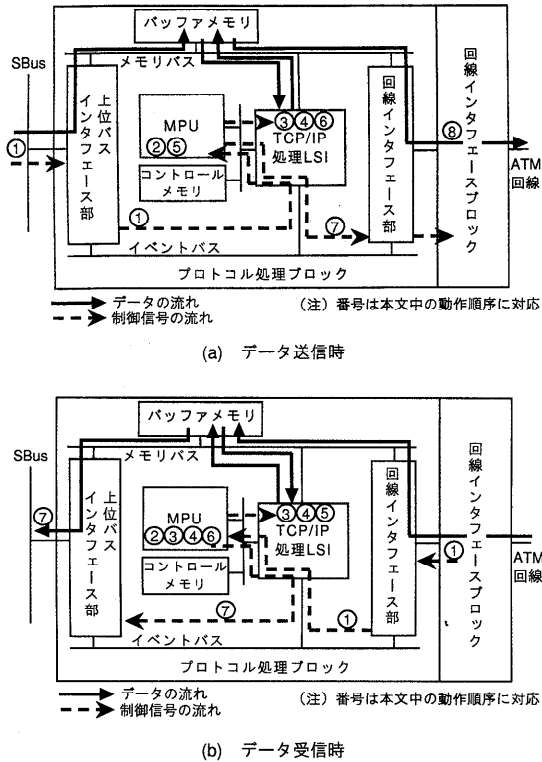


図6 通信ボードの動作。(a) データ送信時、(b) データ受信時
Fig. 6 Control and data flow in TCP/IP/ATM communications board: (a) send, (b) receive.

ボードを経由するパスと既存 TCP/IP および Ethernet を経由するパスとの動的な振り分けを行う。したがって、ネットワークコンフィグレーション用データベースファイル (/etc/netconfig) に設定するのは、このマルチプレクサとなる。いったん経由するパスが確定し、データ転送フェーズに入ると、パスマルチプレクサによるオーバーヘッドはない。

3.3 通信ボードの動作

(1) データ送信 (図 6(a))

- ①: ドライバはアプリケーションからのデータ送信要求を受け、通信ボード上のバッファ管理部よりバッファアドレスを取得し、送信データをバッファメモリ上のデータ領域に DMA により書き込む。書き込み終了後、データ送信要求イベントをイベント管理部へ発行し、MPU (ファームウェア) に通知する。〈ドライバ + BIF タスク〉
- ②: TCP 送信処理を開始する。利用可能なウィンドウを計算、再送タイマ設定、TCP ヘッダ情報を TCP/IP 処理 LSI に設定する。〈TCP/IP タスク〉
- ③: TCP/IP 処理 LSI のヘッダ編集 (TCP ヘッダ) 部・チェックサム生成部を起動する。〈TCP/IP タス

ク〉

- ④: TCP ヘッダ編集、チェックサム演算を行う。〈TCP/IP 処理 LSI〉
- ⑤: IP ヘッダ情報を TCP/IP 処理 LSI に設定する。〈TCP/IP タスク〉
- ⑥: TCP/IP 処理 LSI のヘッダ編集 (IP ヘッダ) 部・チェックサム生成部を起動する。〈TCP/IP タスク〉
- ⑦: 回線インタフェースブロックに対し回線インタフェース部への通知により、バッファメモリ上の所定エリアのデータの送信を指示する。〈LIF タスク〉
- ⑧: 回線インタフェースブロックは、バッファメモリ上のデータを DMA により取得し、回線へデータを送出する。

(2) データ受信 (図 6(b))

- ①: AAL5 フレームの処理終了後、回線インタフェース部があらかじめバッファ管理部から取得したバッファアドレスへ、受信データを DMA により書き込み、データ受信イベントで MPU (ファームウェア) に通知する。〈LIF タスク〉
- ②: 受信データから LLC/SNAP ヘッダを取り除き、IP データ受信イベントで MPU に通知する。〈LIF タスク〉
- ③: チェックサム検証部、ヘッダ解析部を起動し、IP ヘッダのチェックサム検証、ヘッダの解析を行う。〈TCP/IP タスク, TCP/IP 処理 LSI〉
- ④: チェックサム検証部を起動し、TCP のチェックサム検証を行う。〈TCP/IP タスク, TCP/IP 処理 LSI〉
- ⑤: チェックサム検証部は、検証結果を検証結果イベントで MPU に通知する。〈TCP/IP 処理 LSI〉
- ⑥: ヘッダ解析部を起動し、TCP ヘッダの解析を行う。〈TCP/IP タスク, TCP/IP 処理 LSI〉
- ⑦: MPU からの指示により、上位インタフェース部はバッファメモリ上の所定エリアのデータを DMA により装置側のメモリへ転送するとともに、ドライバに対しデータ受信割り込みを通知する。〈BIF タスク〉

4. 通信ボードの評価

4.1 性能

性能の評価は、本来の高速化の対象である TCP データ転送について行った。UDP については、チェックサム演算等に関して、程度は異なるものの、TCP と同様にスループット向上の効果が得られると考えられる。

4.1.1 通信ボード

(1) 評価環境

本通信ボードを 2 台のワークステーション (SPARC station20, 150 MHz) に装着して、光ファイバで直結

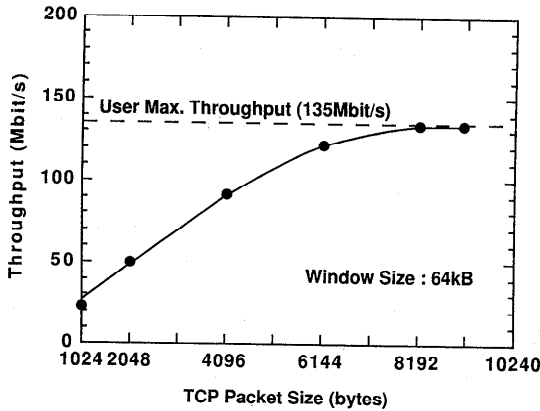


図7 通信ボードのスループット

Fig. 7 Throughput of TCP/IP/ATM communications board.

した。データの送受信は、ワークステーション上のメモリーメモリ間で行った。

(2) 静止画像転送性能

静止画像を転送するアプリケーションを使用してスループットを測定した。このアプリケーションは、一方から同じ静止画像を連続して送信し、他方で受信した静止画像を連続して表示するものである。測定結果は、116 Mbit/s [TCP パケット長 = 9140 B (最大セグメントサイズ)⁵⁾] であった。得られた結果は、ユーザ限界スループットの 85% であり、全帯域を使用していない。

なお、市販 ATM 通信ボード (FORE 社製 SBA200) を用いた場合のスループットは、60 Mbit/s であり、上記条件においては、本通信ボードを使用するシステムは約 2 倍の性能を有する。

(3) 通信ボード対向送受信性能

アプリケーションに依存しない通信ボードとしての能力の評価を目的に、ドライバレベルでボードを直接制御するプログラムを使用してスループットを測定した。

TCP パケット長とスループットとの関係を図 7 に示す。8192 B 以上の TCP パケット長においてほぼユーザ限界スループットを達成している。

4.1.2 ファームウェア

(1) 測定方法

割込みによる処理の中断がない TCP/IP 処理のトランザクションあたりのファームウェア性能を測定するため、1 パケットごとに送受信するプログラムを使用した。測定対象ファームウェアの範囲は、プロトコル処理タスク (TCP/IP タスク)、上位バスインタフェース制御タスク (BIF タスク)、回線インタフェース制御タスク (LIF タスク) とした。測定では、ファーム

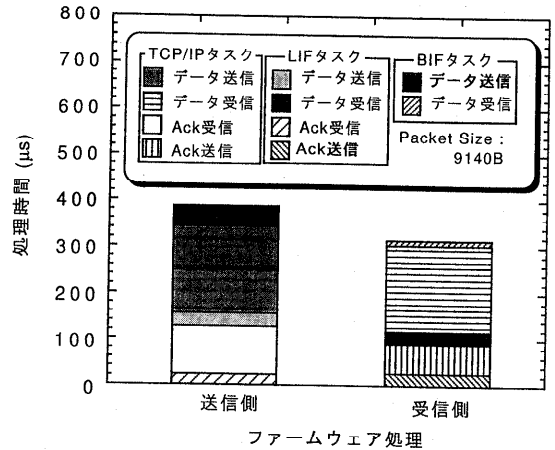


図8 送信側と受信側のファームウェア処理時間の比較

Fig. 8 Comparison of firmware processing times between transmission and reception.

ウェアにボード上の信号をオン/オフする命令を挿入し、その信号をロジックアナライザを用いてモニタすることにより、処理時間を計測した。

(2) 測定結果

送信側 (データ送信+ACK 受信) および受信側 (データ受信+ACK 送信) のファームウェア処理時間との関係を図 8 に示す。同図より、データ送信側の方がファームウェアの処理時間が受信側より長いこと、TCP/IP タスクが全ファームウェア処理時間の大半を占めていることが分かる。また、TCP パケット長とファームウェア処理性能との関係を図 9 に示す。データ長に依存する処理としない処理が存在することが分かる。TCP/IP タスクのデータ送受信処理が、データ長に依存している。このファームウェア処理性能から予測したスループットを図 10 の実線で示す。この値は、TCP パケット長を送信側ファームウェア総処理時間 (データ送信処理+ACK 受信処理) から導いた理想値である。送信側の総処理時間から算出するのは、送信側の総処理時間の方が受信側より大きく、スループットとしては送信側の処理能力に依存するためである。これより、TCP パケット長 6 kB 以上で 135 Mbit/s のユーザ限界スループットを得ることができると予測される。

4.1.3 考察

通信ボードの評価 (4.1.1 項) 結果から分かるように、本通信ボードとしては、155 Mbit/s ATM 回線のユーザ限界スループットを達成しているが、静止画像転送アプリケーションでは達していない。同じアプリケーションを 60 MHz の CPU を内蔵するワークステーションで動作させたとき、本通信ボード使用時は、

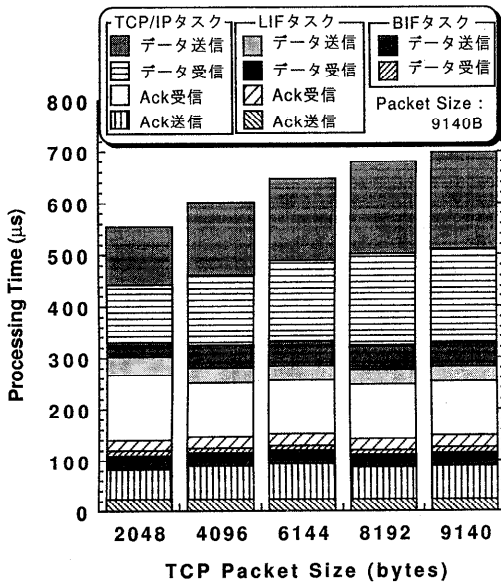


図9 ファームウェア処理時間の分析

Fig. 9 Analysis of firmware processing time.

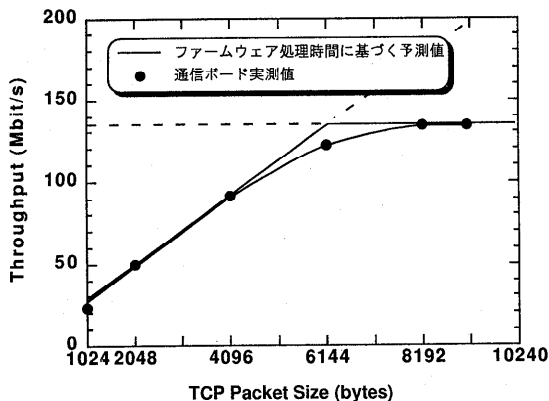


図10 ファームウェア処理時間に基づくスループット予測

Fig. 10 Throughput forecast from firmware performance.

90 Mbit/s, 市販ボード使用時は, 45 Mbit/s であり, 性能が低下した. 以上より, ユーザ限界スループットに達しないのは, アプリケーションである静止画像表示処理がマシン本体の CPU に対して大きな負担となっているためであると考えられる. しかしながら, 本通信ボード使用時の性能は, 市販ボードの使用時の性能の約 2 倍に達している. これらの性能は, 画像表示等のアプリケーション処理まで含んだ総合的な性能であることから, TCP/IP をボード上で高速処理する効果が大きいと分かる.

4.1.2 項の評価において, 6 kB 以下ではユーザ限界スループットに達していない原因を追究するため, ファー

ムウェアの処理について分析した.

まず, 4.1.1 項 (3) による通信ボードのスループット実測値と, 4.1.2 項によるファームウェアの処理時間実測結果から試算したスループット限界予測値とを比較する (図 10 参照). 両者はよく一致していることから, TCP パケット長 6 kB 以下でユーザ限界スループットに達しないのは, ファームウェアの処理性能に起因しているためと考えられる.

また, TCP/IP タスクのデータ送信処理において, データサイズによって処理時間が変化している (図 9). これは, IP のヘッダ編集処理が, TCP チェックサム演算処理終了後に開始されるためである (図 11 (a)). 一方, データ受信時には, TCP チェックサム検証の終了後に TCP 受信処理を行うシリアル処理となり (図 11 (b)), ファームウェア処理時間が TCP パケット長に依存している. すなわち, ファームウェアが idle 状態にある時間が生じる (ただし, idle 状態にあるファームウェアは, 他のコネクションの処理や ACK 送受信等の他の処理を行うこと可能である). 以上より, より短い TCP パケットの通信におけるスループット向上, さらなる高速回線への対応のためには, on-the-fly 処理等によるチェックサム処理の先行処理化やより並列処理性を高めること等, ファームウェアの高速処理化を図る必要がある.

4.2 実装

2 章で述べた各要求条件は満足したが, いくつかの問題が明らかになった. まず, 1 スロットスペースという高密度実装のため, 放熱対策を施す必要があった. また, 一部プロトコル処理をファームウェアで行っているため, 高速化の限界が明らかになった. ただしこれは, MPU の内部クロック周波数を上げることによりある程度向上できる. さらに, Socket インタフェースを使用する既存アプリケーションとの互換性に関し, 既存 OS 中の TCP/IP が TLI インタフェース規定に完全には準拠していないため, ドライバ/ファームウェアでもそのような処理をサポートする必要があった.

5. 実用性に関する考察

5.1 インテリジェント NIC の必要性

かつて Ethernet NIC に CPU を搭載し, 通信プロトコル処理を行うようにしたインテリジェント NIC は, ホスト/端末の CPU の高速化によって NIC 上の CPU の処理速度が端末よりも遅くなり実用化されなかった. しかしながら, CPU の高速化とともに処理量の大きいマルチメディアを扱うアプリケーションが現れている. あるいは, 電子メールやファクシミリを受

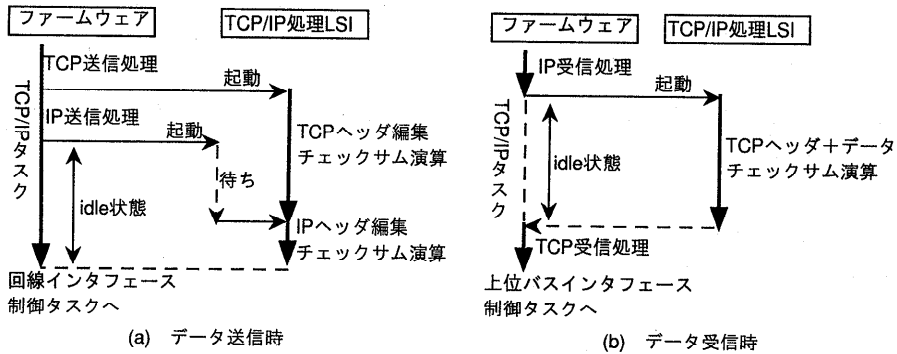


図 11 TCP チェックサム演算処理時のファームウェアと TCP/IP 処理 LSI の関係
 Fig. 11 Relation between firmware and TCP/IP processing LSI during TCP checksum: (a) send, (b) receive.

信, WWW サーバを自動的に巡回して新しい情報を収集する等, バックグラウンドで動作するアプリケーションが出現している。これは, 今後ますます増大する傾向にあると予想される。また, 利用できるネットワークも高速化されてきている。市販の ATM ボードの中には, プロトコルの一部 (チェックサム) をボード上で処理する動向もある。通信プロトコルを NIC 上で高速に処理するインテリジェント NIC の必要性は, 十分にあると考える。

5.2 長距離転送遅延問題への対策

ネットワークの帯域と遅延時間に対してウィンドウが不足すると, ネットワークに空き時間が生じ, 効率的なデータ転送ができなくなる。TCP/IP 長距離通信における問題は依然存在する。解決策の 1 つとして, 拡張ウィンドウ⁶⁾を用いる方法があるが, データが損失した際の再送によるオーバーヘッドが大きくなる可能性がある。そこで, 我々は, 解決方法の 1 つとして, 複数コネクションを利用する方法^{7),8)}を考えている。ネットワークの帯域と遅延時間から必要なウィンドウサイズを計算し, リンク数 (TCP コネクション) を決定し, ネットワークに空きが生じないようにする。ただし, 各リンクの最大ウィンドウサイズは, ネットワークの品質に応じて決定する。したがって, ネットワークの品質が悪い場合にも拡張ウィンドウを用いる方法に比べて, リンクごとのウィンドウサイズを小さくできるため, 再送による影響は少ない。このようにウィンドウサイズをリンク数倍にすることにより, 長距離通信時の問題を回避できると考える。さらに, ハードウェアによる通信プロトコル処理の高速化により, 1 つのアプリケーションあたりより多くのリンクを使用することが可能となり, 高性能を確保できると考える。

6. おわりに

高性能 TCP/IP/ATM 通信ボードの設計方針, システム構成および実装性能について述べた。データサイズ 8 kB 以上では, ユーザ限界スループットを達成しており, 本通信ボードの高性能性を実証した。IITTP, FTP 等を用いた画像データ等のマルチメディアコンテンツ転送は, 大きなデータサイズでの使い方が中心となると考えられ, 本通信ボードはそのようなアプリケーションに対して, 効率的な転送が期待でき有効である。本通信ボードの使用により, 高スループットが要求される新しいマルチメディアサービスを, 既存のワークステーションを用いても実現可能となる。また, 本通信ボードは, パソコンの PCI バスや Fast Ethernet 等にも容易に対応できる構成をしており, 優れた拡張性を有する。本通信ボードにおけるプロトコル高速処理技術は, ATM 等の高速回線を利用した, マルチメディア通信サービスの実現に大きな前進をもたらすものと考ええる。

今後, (1) 多重コネクション環境での評価, (2) 長距離通信での遅延対策を行う予定である。また, 広範なアプリケーションへの適用を目指し, バッファ間のデータコピーの影響など, 遅延時間に関する評価も実施したい。

謝辞 日頃ご指導いただく坂間保雄 NTT 情報通信研究所知的通信処理研究部長に感謝します。また, 長期にわたってご指導いただいた山下正秀 NTT マルチメディアシステム総合研究所研究企画部長に深謝します。さらに, 熱心に討論いただいた同僚に感謝します。

参 考 文 献

- 1) Clark, D.D., Jacobson, V., Romkey, J. and Salwen, H.: An Analysis of TCP Processing Overhead, *IEEE Communication Magazine*, Vol.27, No.6, pp.23-29 (1989).
- 2) XTP Forum: Xpress Transport Protocol Specification, XTP Revision 4.0, XTP Forum, CA (1995).
- 3) 丸山 充: ATM 高速通信ボードの構成と評価, 情報処理学会研究報告, 96-DPS-74, pp.103-108 (1996).
- 4) Transport Provider Interface Specification: UNIX International, OSI Special Interest Group, Revision 1.5 (1992).
- 5) Atkinson, R.: Default IP MTU for Use over ATM AAL5, RFC1626 (1994).
- 6) Jacobson, V., Braden, R. and Borman, D.: TCP Extensions for High Performance, RFC 1323 (1992).
- 7) Allman, M., Ostermann, S. and Kruse, H.: Data Transfer Efficiency over Satellite Circuits Using a Multi-Socket Extension to the File Transfer Protocol (FTP), *Proc. ACTS Results Conf.*, Network Section, NASA Lewis Research Center, Cleveland, OH (1995).
- 8) 三宅 優, 加藤聡彦: ウィンドウスケールオプションを実装した拡張 TCP の高速通信に関する一検討, 1997 年電子情報通信学会総合大会, B-7-158, p.287 (1997).

(平成 9 年 5 月 13 日受付)

(平成 9 年 11 月 5 日採録)



長田 孝彦

1970 年生。1993 年広島大学工学部第 2 類卒業。1995 年同大学院博士課程前期修了。同年, NTT 入社。現在, NTT 情報通信研究所勤務。プロトコル高速処理方式の研究

に従事。



東海林敏夫 (正会員)

1951 年生。1974 年北海道大学工学部電気工学科卒業。同年, 日本電信電話公社に入社。現在, NTT アドバンステクノロジー (株) 情報ネットワーク事業部担当課長。DIPS 通信制御処理装置ハードウェア, 高性能通信プロトコルアーキテクチャ, プロトコル高速処理方式の研究に従事。電子情報通信学会会員。



山下 博之 (正会員)

1957 年生。1979 年京都大学工学部情報工学科卒業。1981 年同大学院修士課程情報工学専攻修了。同年, 日本電信電話公社 (現 NTT) に入社。以来, DIPS 通信制御装置および通信系ファームウェアの研究実用化に従事。1991 年より, フォールトトレラントコンピューティング, ネットワークワイドのコンピューティングアーキテクチャ, 高速通信プロトコル処理などの研究に従事。現在, 研究開発推進部担当部長。電子情報通信学会, 応用物理学会, IEEE 各会員。



塩川 鎮雄 (正会員)

1944 年生。1968 年金沢大学電子工学科卒業。1973 年静岡大学大学院修士課程修了。同年, 日本電信電話公社に入社。現在, NTT ソフトウェア (株) ニュービジネス事業本部担当部長。DIPS 論理装置の研究実用化, その後, 情報ナビゲーションシステム, ユビキタスコンピューティングの研究に従事。電子通信処理学会会員。