

Gigabit Ethernet NIC の性能評価

住元真司† 堀 敦史† 手塚宏史†
原田 浩† 高橋俊行† 石川 裕†

Gigabit Ethernet NIC(Network Interface Card) の TCP/IP 通信の性能評価について述べる。評価に用いた NIC は、Packet Engine 社 G-NIC II, Intel 社 PRO/1000 Gigabit Server Adapter, 3COM 社 3C985-SX, 及び、ODS/Essential 社 JackRabbit である。測定には Netperf を用い、Pentium II 400MHz 2CPU 搭載の PC 上で Windows NT 4.0 と Linux 2.1.132 上で測定した。

評価の結果、TCP/IP のバンド幅が最も性能が良いのは、G-NIC II(linux-2.1.132, 1 CPU) の場合で、2048 バイトメッセージ時に 49.5 MB/s であり、他社の NIC のバンド幅性能は 31~33 MB/s であった。ラウンドトリップ時間が最も良いのは、PRO/1000(Windows NT, 1CPU) の場合で、4 バイトメッセージ時に 187.1 μ s であり、他社 NIC は、270 μ s~410 μ s であった。

Performance Evaluation of Gigabit Ethernet NIC

SHINJI SUMIMOTO,† ATSUSHI HORI,† HIROSHI TEZUKA,†
HIROSHI HARADA,† TOSHIYUKI TAKAHASHI† and YUTAKA ISHIKAWA †

This paper presents the TCP/IP communication performance evaluation of Gigabit Ethernet NICs. Packet Engines G-NIC II, Intel PRO/1000 Gigabit Server Adapter, 3COM 3C985-SX and ODS/Essentials JackRabbit NIC are tested on Dual Pentium II 400MHz PCs which run Windows NT 4.0 and Linux 2.1.132 using the Netperf benchmark.

The performance results show that G-NIC II achieves 49.5 MB/s TCP/IP bandwidth on Linux using 1 CPU. The others are between 31 MB/s and 33 MB/s, and PRO/1000 achieves 187.1 μ s round trip latency on Windows NT using 1 CPU. The others are between 270 μ s to 410 μ s.

1. はじめに

近年、次世代標準 LAN を目指し登場した Gigabit Ethernet は、多くのベンダが機器販売を行ない、着々と広まりつつある。量産と価格競争により価格は下がって来ており、より対線を用いた 1000BASE-TX 対応の製品出荷が始まれば、更に価格が下がると予想される。Gigabit Ethernet を利用したコストパフォーマンスの高いクラスタシステム構築が可能となるだろう。

しかしながら、Gigabit Ethernet の性能 (特に NIC) について比較評価した文献は少ない。今後、コストパフォーマンスの高いクラスタシステム構築の指針となるように、Gigabit Ethernet NIC の通信性能を評価することが本稿の目的である。

本稿では、評価に用いた NIC の概要について述べた後、Netperf¹⁾ を用いた TCP/IP のバンド幅性能と

ラウンドトリップ性能評価について述べる。

2. Gigabit Ethernet NIC について

2.1 Gigabit Ethernet

Gigabit Ethernet は、Ethernet フレームを用い、双方向に 125MB/s のバンド幅を持つネットワークである。現在、製品化されているのは、主に光ファイバを用いた 1000BASE-SX, 1000BASE-LX である。より対線の 1000BASE-TX の標準化も最終段階にあり、標準化後、間もなく製品が出荷される見込みである。

2.2 評価 NIC の概要

今回の評価で用いた NIC の概要を表 1 に示す。すべて 1000BASE-SX の NIC である。

Packet Engine 社 G-NIC-II:

G-NIC II は、G-NIC に続く第 2 世代の NIC である。NIC 上に CPU と外付けメモリを持つが、33MHz PCI, 1CPU の G-NIC に比べ、66MHz PCI 対応、2CPU 構成である。(以降、G-NIC II)

3COM 社 3C985-SX:

Alteon 社 ACEnic の OEM であるが、現在手に

† 技術研究組合 新情報処理開発機構 つくば研究センター
並列分散システムソフトウェアつくば研究室
Parallel & Distributed System Software Laboratory
TRC, Real World Computing Partnership
<http://www.rwcp.or.jp>

表1 評価 NIC の概要

	G-NIC II	3C985-SX	PRO1000	J.Rabbit
PCI	32/64 bit 33-66MHz	32/64 bit 33MHz	32/64 bit 33MHz	32 bit 33MHz
CPU	2 CPU	1 CPU	不明	1 CPU
メモリ	512KB	1MB	不明	1MB
対応 OS				
Win NT	○	○	○	○
Linux	○	○		○
Solaris	○	○		○
Digital	○	○		○
IRIX	○	○		○

対応 OS は、主要 OS のみ、サードベンダ、OEM 先を含む
 J.Rabbit: JackRabbit
 Win NT: Windows NT, Digital: Digital UNIX,

はいる NIC より 1 つ前のハードウェアである。NIC 上に CPU と外付けメモリを持つ。他社の NIC と違い、MTU を 9000 バイトにする JUMBO Frame をサポートしている。現在の ACEnic は、66MHz PCI に対応し、2 CPU 構成となっている。(以降、3C985-SX)

Intel 社 PRO/1000 Gigabit Server Adapter:

他社の NIC と異なり、外付けメモリを持たない点の特徴。現状では、ドライバの OS 対応範囲が狭い。(以降、PRO/1000)

ODS/Essential 社 JackRabbit:

JackRabbit も、NIC 上に CPU と外付けメモリを持つ。新情報で開発を進めている GigaE PM⁽²⁾⁽³⁾ で使用。(以降、JackRabbit)

2.3 NIC の作りについて

Gigabit Ethernet NIC は、バスインターフェイス、Gigabit Ethernet MAC (Media Access Control)、及び、トランシーバ、光インターフェイスから構成されるが、主にフレームのバッファリング、TCP/IP 性能向上サポートと運用機能実現のため、NIC 上に CPU とメモリを持つものが多い。また、TCP/IP 性能向上をサポートする機構として、DMA 時のチェックサム計算機構を備えるものが多い。現状、NIC の作りは以下の 2 つに分類される。

- HIPPI コントローラに Gigabit Ethernet MAC を接続したもの: JackRabbit
- CPU, PCI DMA, Gigabit Ethernet MAC 回路などを内蔵した LSI を開発し使ったもの: G-NIC II, 3C-985-SX, PRO/1000 (CPU 搭載は不明)

3. TCP/IP 通信性能評価

本章では、4 種の Gigabit Ethernet NIC の TCP/IP 性能 (バンド幅性能とラウンドトリップ遅延) の性能評価について述べる。

3.1 測定環境

TCP/IP のバンド幅とラウンドトリップ遅延の測定

は、netperf 2.1 ベンチマーク¹⁾ を用いた。測定は、CPU 1 台、CPU 2 台の場合について一つのベンチマークプログラムを実行して測定した。バンド幅性能は TCP_STREAM オプション、ラウンドトリップ遅延測定は TCP_RR オプションで測定した。すべての測定では、nodelay オプションを用い、送受信のソケットバッファサイズは 128KB である。

以下に測定環境について述べる。

計算機: Pentium II 400MHz x 2, 440BX chipset, bus clock 100MHz, 256MB SDRAM

また、G-NIC II では、以下の計算機も利用した。

- PWS 600au: Alpha 21164 600MHz, 21174 chipset, bus clock 66MHz, 320MB SDRAM
- PC Cluster II: Pentium PRO 200MHz, 440FX chipset, bus clock 66MHz, 256MB FastPage DRAM

スイッチ: Extreme 社 Summit I

OS: Windows NT 4.0 Service Pack 3 (G-NIC I-I, PRO/1000, 3C985-SX), Linux 2.1.132 ac1 patch (G-NIC II, PRO/1000, JackRabbit).

ドライバ: Windows NT はベンダ製を使用した。又、Linux 2.1.132 は以下を利用した。

G-NIC II "hamachi.c:v0.07 8/17/98 Written by Donald Becker"⁴⁾

3C985-SX Linux-2.1.132 附属ドライバ "Written by Jes Sorensen, Jes.Sorensen@cern.ch"

JackRabbit Essential 社提供ドライバ, "Written by Donald Becker and Essential Communications", GigaE PM のドライバ

Windows NT 上の JackRabbit は測定不能であった。また、比較のために、Linux の性能測定では、Intel PRO/100 PCI Adapter (100BASE-TX) と、GigaE PM 上の TCP/IP の測定結果も示す。

3.2 Windows NT でのバンド幅測定結果

図 1 に 2CPU, 図 2 に 1CPU の時の Windows NT 上でのバンド幅測定結果を示す。結果より、2CPU の時では、1458 バイトメッセージ時に PRO/1000 の 35.2MB/s が、又、1CPU の時では、2900 バイト時に G-NIC II の 35.2MB/s が最大バンド幅であった。2CPU の時より 1CPU の時が全体的に良い値を示した。

3.3 Linux 2.1.132 でのバンド幅測定結果

図 3 に 2CPU, 図 4 に 1CPU の時の Linux 2.1.132 上でのバンド幅測定結果を示す。結果より、2CPU の時では、1280 バイトメッセージ時に G-NIC II の 30.9MB/s が、又、1CPU の時でも、2048 バイトメッセージ時に G-NIC II の 49.5MB/s が最大バンド幅であった。3C589-SX を除き、2CPU の時より 1CPU の時が全体的に良い値を示した。

G-NIC II では、2CPU の時のバンド幅は 1CPU の時の 62.5% であった (図 3, 図 4 の PE GNIC-II)。これは、Linux 2.1.132 の同じレベルの割り込みを複数

CPUに負荷分散する機構の影響が大きい。この機能をOFFにすることで41.0MB/sに改善した(図3, PE GNIC-II w/o INT DIST)。なお、他社のNICではこの影響は小さかった。

3.4 Windows NTでの遅延測定結果

図5に2CPU, 図6に1CPUの時のWindows NT上でのラウンドトリップ遅延測定結果を示す。2CPUの時では、4バイトメッセージ時にPRO/1000の199.6 μ sが、又、1CPUの時も、4バイトメッセージ時にPRO/1000の187.1 μ sが最小遅延であった。2CPUと1CPUとの差は12.5 μ sであった。他のNICのラウンドトリップ遅延は304 μ s以上と大きい。

3.5 Linux 2.1.132での遅延測定結果

図7に2CPU, 図8に1CPU時のlinux 2.1.132上でのラウンドトリップ遅延測定結果を示す。GigaE PMを除き、2CPUの時では、4バイトメッセージ時にG-NIC IIの295.0 μ sが、又、1CPUの時も、4バイトメッセージ時にG-NIC IIの270.3 μ sが最小遅延であった。2CPUと1CPUとの違いは24.7 μ sであった。G-NIC II, 3C589-SX, JackRabbitでは、メッセージサイズに関わらず、遅延一定の場合があった。

G-NIC IIの最小遅延(270.3 μ s)は、同じPentium II 400MHzを用いた100BASE-TXであるPRO/100(図中、100BASE-TX)の144.9 μ sに比べ86%遅い。また、JackRabbitの場合は、ベンダ供給のファームウェアによる最小の遅延が337 μ sに対して、同じJackRabbitを用いたGigaE PMでは、128.2 μ sを実現している。ファームウェア処理の違いと考えられる。

3.6 システムによる違い

最大バンド幅性能を出したG-NIC IIについて、CPUにPentium PRO, Alphaを用いたシステムで1CPUの時の通信性能を測定した。図9と図10に、バンド幅性能とラウンドトリップ遅延の測定結果を示す。結果より、バンド幅性能の差はCPUの処理能力とメモリ性能の差によるものといえる。また、遅延では、512バイト以下ではCPU能力の違いによる遅延への影響は小さかった。

3.7 Gigabit Ethernet NICの安定性について

Netperfベンチマークは、CPUとバスに負荷がかかるベンチマークのため、測定中、PRO/1000, G-NIC II, JackRabbitにおいて通信不能(インターフェイスダウン, システムハング)になるケースがあった。

3.8 相互接続性について

実験に用いたSummit 1では、試用したすべてのNICについて、接続上問題になることはなかった。

4. 考 察

4.1 Gigabit Ethernet NICの性能について

Pentium II 400MHz搭載のPCを用いることにより、TCP/IPで49.5MB/sのバンド幅性能が得られ

た。Gigabit Ethernetの持つバンド幅125MB/sの半分以下であったが、CPUとメモリ性能の向上、NICのハードウェア性能の向上により改善されつつある。他社NICは31~33 MB/sであった(図1-図4)。

また、PRO/1000とGigaE PM以外のNICは、ラウンドトリップで270 μ s~410 μ s(図5-図8)と、100BASE-TXの結果に比べて約86%以上大きい。クラスタシステム上で、遅延が問題になる並列プログラムで性能を出すのは難しい。GigaE PMでは100BASE-TXより速い結果が出ているので、改善を期待したい。

試験したNICの中では、Windows NTではPRO/1000, Linux 2.1.132では、G-NIC IIが最も良好な結果を示した。

4.2 NICの持つ遅延について

NIC本来の遅延を測定するため、ICMP ECHOを用いてラウンドトリップ遅延を測定した。測定は、Linux上でテストプログラムを動作させ、通信相手を変えて測定した。リファレンスはLinux上で最も遅延の小さいGigaE PMを用いた。図11に結果を示す。凡例の左側が測定プログラム実行ホストのNIC, NTはWindows NT, 及び表示無しはLinuxを示している。

結果より、3C589-SX(NT, Linux), G-NIC II(Linux)は、サイズに関わらずほぼ一定値を示した。一方、PRO/1000, GigaE PMは、サイズに比例した値となった。G-NIC II(Windows NT)については、Linux上での結果に反し、サイズに比例する結果となった。NICファームウェア処理の差と考えられる。

4.3 TCP/IPのオーバーヘッドについて

図11と図8のGigaE PM上の遅延結果より、TCP/IPのオーバーヘッドは、ラウンドトリップで28.1 μ sであった。この値は受信側処理(ユーザとsystem call)のコストを含む。遅延全体の21.9%であった。

5. ま と め

Gigabit Ethernet NICのTCP/IP通信性能評価について述べた。評価の結果、バンド幅は、G-NIC II上(linux-2.1.132, 1 CPU)で49.5MB/s, ラウンドトリップ遅延は、PRO/1000上(Windows NT, 1 CPU)で187.1 μ s, GigaE PMを含めると128.2 μ sの性能であった。他社NICの性能については、バンド幅性能は31~33 MB/s, 遅延は270 μ s~410 μ sであった。試験したNICの中では、Windows NTではPRO/1000, Linux 2.1.132ではG-NIC IIが最も良好な結果を示した。

バンド幅はGigabit Ethernetの持つバンド幅(125MB/s)の半分以下であったが、CPUとメモリシステム性能の向上、及び、NICのハードウェア性能の向上により改善されつつある。一方、遅延については、PRO/1000とGigaE PMを除き、100BASE-TXと比べても4バイトメッセージ時に約86%以上遅く、

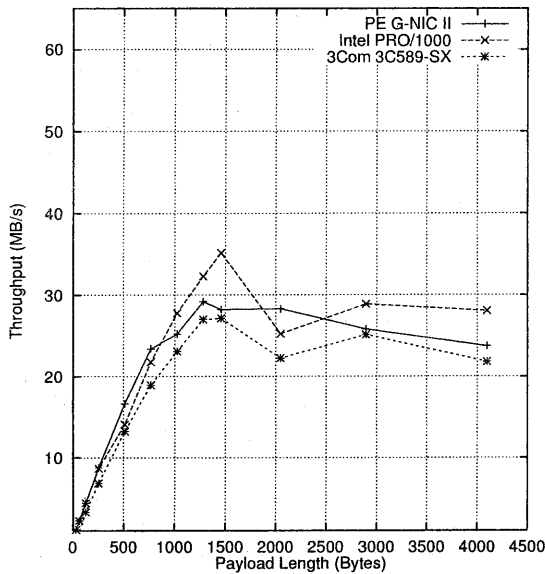


図1 Windows NT 2CPUでのTCP/IPバンド幅性能

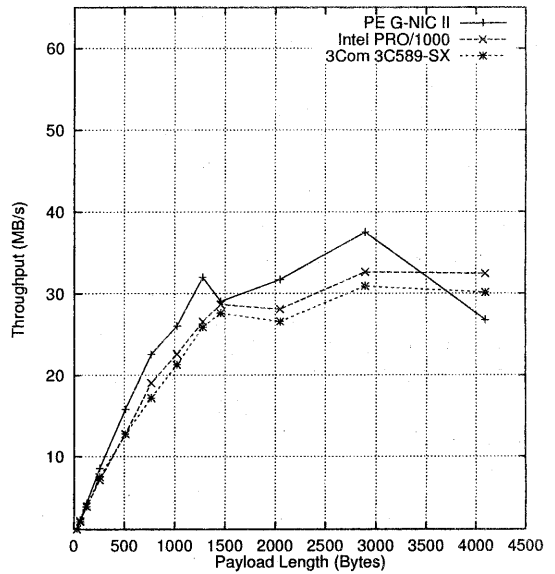


図2 Windows NT 1CPUでのTCP/IPバンド幅性能

改善が必要である。

謝辞 評価用NICを貸し出して頂いた、旭エレクトロニクス株式会社黒江様、ネクストコム株式会社陸田様、丸紅ソリューション株式会社宮垣様、東京リース株式会社毛利様に感謝します。

参考文献

- 1) <http://www.netperf.org/>.
- 2) 住元, 石川, 堀, 手塚, 原田, 高橋. Gigabit Ethernetを用いた高速通信ライブラリの設計. 情報処理学会研究報告 98-HPC-72, pp. 109-114, August 1998.
- 3) S. Sumimoto, H. Tezuka, A. Hori, H. Harada, T. Takahashi, and Y. Ishikawa. "High Performance Communication using a Gigabit Ethernet". Technical report, RWCP, 1998. TR-98003.
- 4) <http://www.tidalwave.net/becker/hamachi.html>.

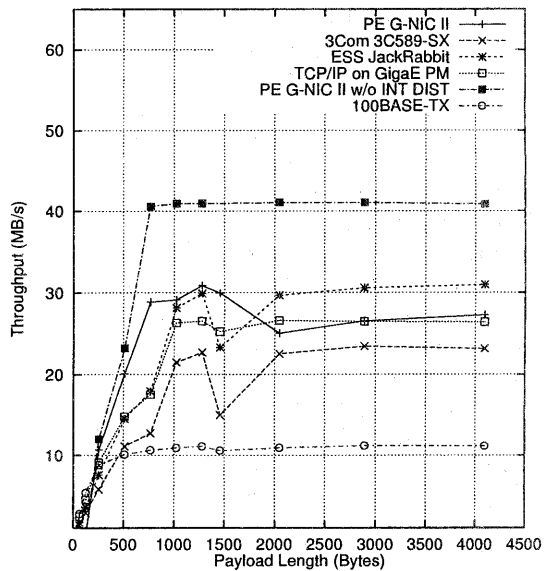


図3 linux 2.1.132 2CPUでのTCP/IPバンド幅性能

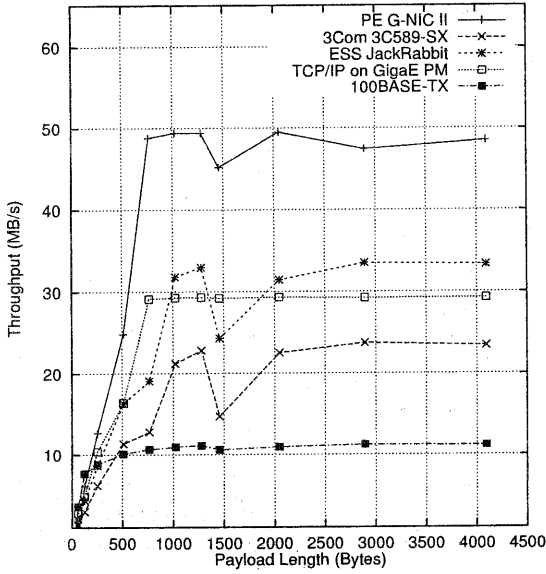


図4 linux 2.1.132 1CPUでのTCP/IPバンド幅性能

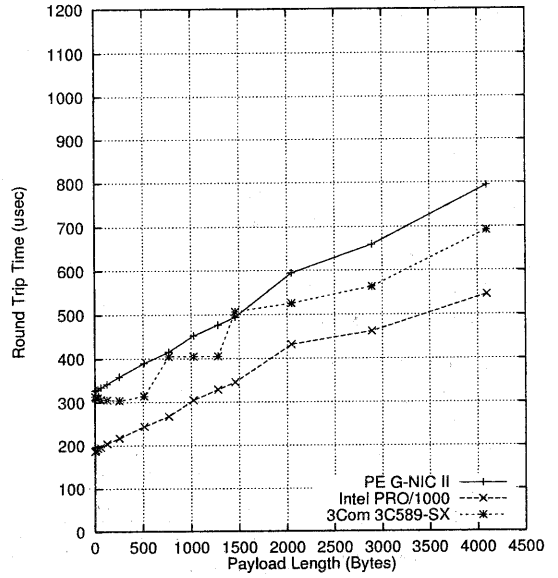


図6 Windows NT 1CPUでのTCP/IP遅延時間

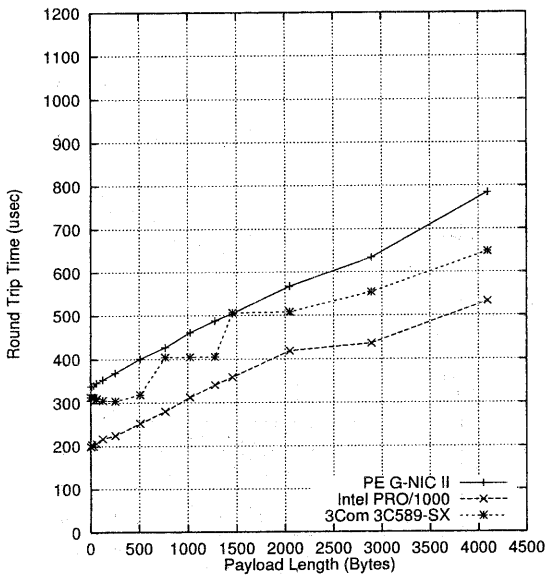


図5 Windows NT 2CPUでのTCP/IP遅延時間

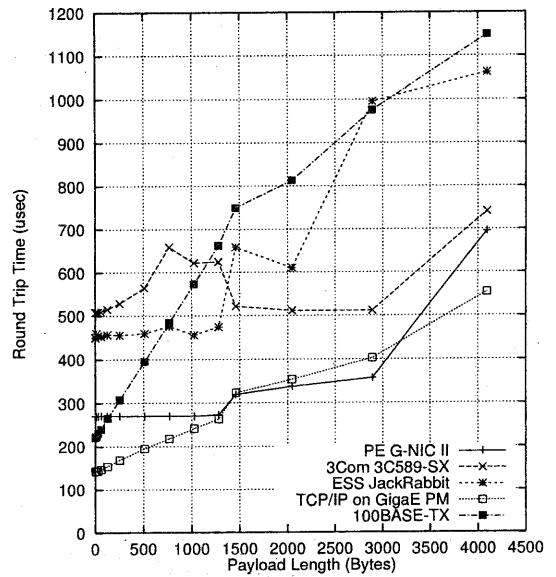


図7 linux 2.1.132 2CPUでのTCP/IP遅延時間

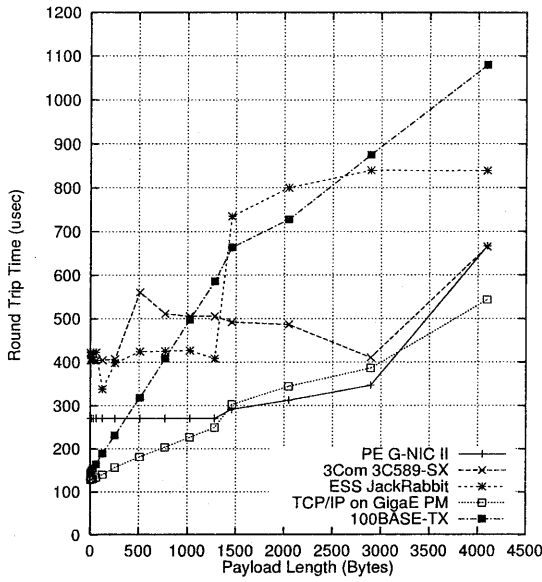


図 8 linux 2.1.132 1CPUでのTCP/IP遅延

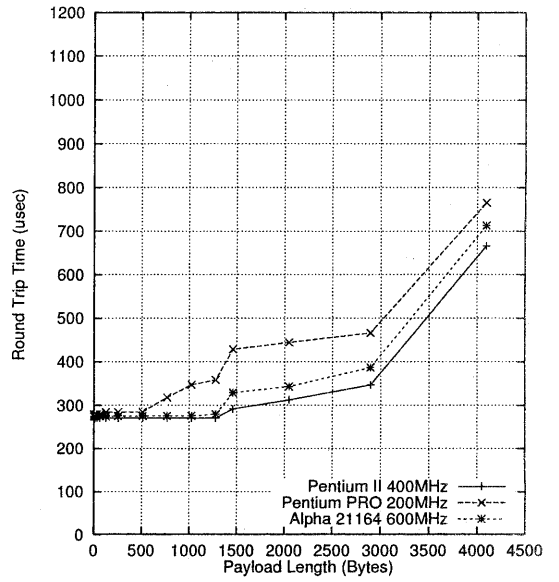


図 10 システムの違いによるTCP/IP遅延時間(G-NIC II)

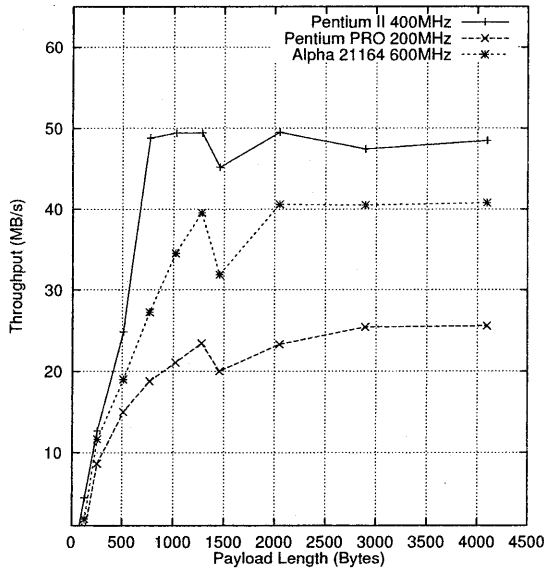


図 9 システムの違いによるTCP/IPバンド幅性能(G-NIC II)

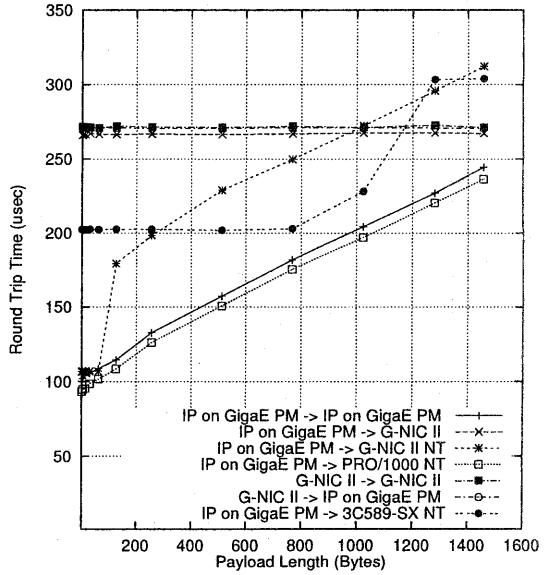


図 11 ICMP ECHOの遅延時間