

ネットワークテストベッドGNET-1のSNMPツール

佐藤博之^{A)} 児玉祐悦^{A)}
工藤知宏^{A)} 清水敏行^{B)}

^{A)} 産業技術総合研究所グリッド研究センター

^{B)} 株式会社シナジェテック

ネットワークテストベッドGNET-1は広域ネットワーク等の通信路上で発生する遅延やパケット損失などの現象をハードウェアレベルで発生させることができるネットワーク実験装置であり、また同時に通過パケットを詳細なレベルで解析できるネットワーク観測装置でもある。本報告ではこのGNET-1の制御ならびに観測データの取得をSNMPで行うための考察ならびに実装とその評価を行った。その結果リモートマシンからSNMPを利用して1ミリ秒間隔の高精度バンド幅情報を計測することが可能であることを確認した。

SNMP tool for Network Testbed GNET-1

HIROYUKI SATO^{A)} YUETSU KODAMA^{A)}
TOMOHIRO KUDOH^{A)} TOSHIYUKI SHIMIZU^{B)}

^{A)} Grid Technology Research Center, National Institute of Advanced Industrial Science and Technology

^{B)} Synergetech, Inc

GNET-1 is a powerful network testbed in order to measure network traffics, emulate networks, and develop new protocols. We have developed SNMP tools for this equipment. In this paper, detail discussion and evaluation about the implementation are presented. Our SNMP tool can get fine-grain bandwidth information those interval is one millisecond.

1. はじめに

近年、コンピューターネットワークの通信バンド幅は飛躍的な拡大を遂げ、国際間の広域網でGbpsを超える例もはや珍しくはなくなった。しかし、大陸間を跨ぐ様な長距離通信では遅延時間が大幅に伸びるため、現在の標準的なインターネット通信プロトコルではネットワーク帯域幅を有効に利用することができない。このような広帯域高遅延ネットワークを十二分に活用するため、各種ソフトウェアの開発が行われている。このようなソフトウェア開発には、他の通信の影響を受け動的に変化する実ネットワークではなく、遅延時間や帯域幅、輻輳等が再現性のある形で自由にコントロールできるネットワークエミュレーション環境を利用することは、極めて有用であると考えられる。

これらソフトウェアの改良や比較を行うためには詳細なネットワークの計測が必要^{[1][7]}であるが、標準的なネットワーク機器では輻輳のときなどに発生するトラフィックの変化を高精度な時間分解能で測定することはできない。

産総研で開発したGNET-1^{[1][6]}は、このような高精度な計測が可能で多様なネットワークエミュレーションが可能なネットワークテストベッドである。開発に際しては大規模FPGA (Field Programmable Gate Array) を使用することにより、必要な機能を自由に組み込める柔軟な仕様になっている。このGNET-1の高精度測定機構は、入出力バンド幅をミリ秒以下の間隔で観測・記録でき、GNET-1のコントロールを行うホストPCからUSBインターフェイスを経由して読み出すことができる。

従来このGNET-1のコントロールやバンド幅測定値の読み出しは、ユーザがホストPCへリモートログオンすることにより行なわれていた。しかしGNET-1は開発者だけでなく一般ユーザにも使われるようになりつつあり、もっと容易にコントロールやバンド幅の観測を行えるようにする必要がある。またGNET-1が複数設置され、それらから統計情報を収集することを考えると、個々のホストPCにログインして作業を行うのは効率が良くなく、可能ならば1台のマシンからの操作できることが望ましい。そこで我々はこれらの要求を満たすため、汎用的

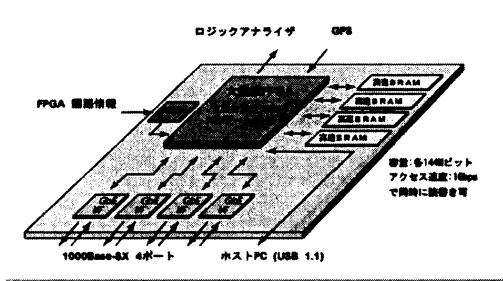


図1 GNET-1のブロック図

なネットワーク管理プロトコルSNMP (Simple Network Management Protocol) を使用しGNET-1を制御することを考えた。SNMPは多くのネットワーク機器に搭載されており、これによりGNET-1と他のネットワーク機器を同一のインターフェイスで扱うことが可能になる。

2. GNET-1でのSNMP

2.1. ネットワークテストベッドGNET-1の概要

GNET-1のブロック図を図1に示す。コアとなるFPGAにはXilinx社のXC2V6000を使用し、パケットは4つのGigabit Ethernet (GbE) のポートを経由して入出力され、自由に遅延や欠損を発生させることができる。また144Mbitsの高速SRAMが搭載されており、125Mbytes/sec以上で同時読み書きが可能である。GPSモジュールは高精度な時間情報を得るために使用される。このGNET-1の制御ならびに計測データの取得はUSB1.1を経由してホストPCからFPGA内のレジスタにアクセスすることによって行われる。

2.2. 標準的なSNMPとその動作

RFC1157^[2]で定義されたSNMPは、主としてネットワークに接続された機器類の監視や設定を外部のリモートマシンから行うためのプロトコルであり、管理される側である「SNMPエージェント」と、管理する側である「SNMPマネージャ」の2つで構成される。エージェントはマネージャからの要求(リクエスト)に応じ監視機器の情報を「MIB (Mnagement Information Base)」と呼ばれる管理情報ベースを参照して取得しそれを返答(レスポンス)する仕組みになっている(図2)。このMIBに格納されている情報は「オブジェクト(Object)」と呼ばれる単位で管理され、個々に「オブジェクトID(OID)」という識別子が割当てられている。MIBは標準なものが規定されているが、ベンダ等が独自に拡張することも可能である(拡張MIB)。

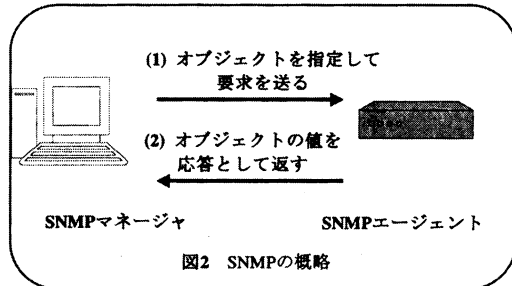


図2 SNMPの概略

現在SNMPはルータ等の多くのネットワーク機器に組み込まれ、その状態監視や統計情報等を収集するために利用されている。一般的なマネージャでは定期的にリクエストをエージェントへ送り、返ってきた値をマネージャ側で保存して、後から各種統計データを閲覧できる仕組みになっている。代表的なものとしてはWebで見ることができるMRTG^[3] (Multi Router Traffic Grapher) 等が挙げられる。

2.3. GNET-1でのSNMPの問題点

前節で述べたSNMPによる統計情報収集をGNET-1で行うと、特に高精度バンド幅測定において問題が発生する。通常SNMPではエージェントが返す統計情報には時刻情報は含まれておらず、MRTG等ではマネージャ側の時刻情報を用いている。これは通常のSNMPの用途では1秒程度以上の間隔で情報を収集することしか想定しておらず、マネージャ側の時刻を用いても実用上問題が生じないからである。これに対しGNET-1では数ミリ秒ごとに情報を収集することが可能であるため、正確な情報を得るにはエージェント側で統計情報に時刻を付加する仕組みが必要である。また、単位時間当たり生成される情報が多量になるため、1回の問い合わせで多数の時刻と統計情報の組を受け渡す仕組みも必要になる。

2.4. GNET-1でSNMPを利用する

このような問題を解決するため、本稿ではエージェントからマネージャにバンド幅情報と時刻情報の組を複数まとめて送る方式を試みた(図3)。本方式では、バンド幅情報のリクエストがマネージャから送られてきた場合、受信した時点から過去の複数個のデータを返送することとした。データの個数はあらかじめSNMPでマネージャ側から指定する。

ただし、リクエストのタイミングによっては同一のデータが重複してエージェントから返送されたり、送信できないデータが発生したりする。しかし、高精度バンド幅測定ではデータの欠損は許容しがたい問題である。そこで、本方式ではリクエストのイン

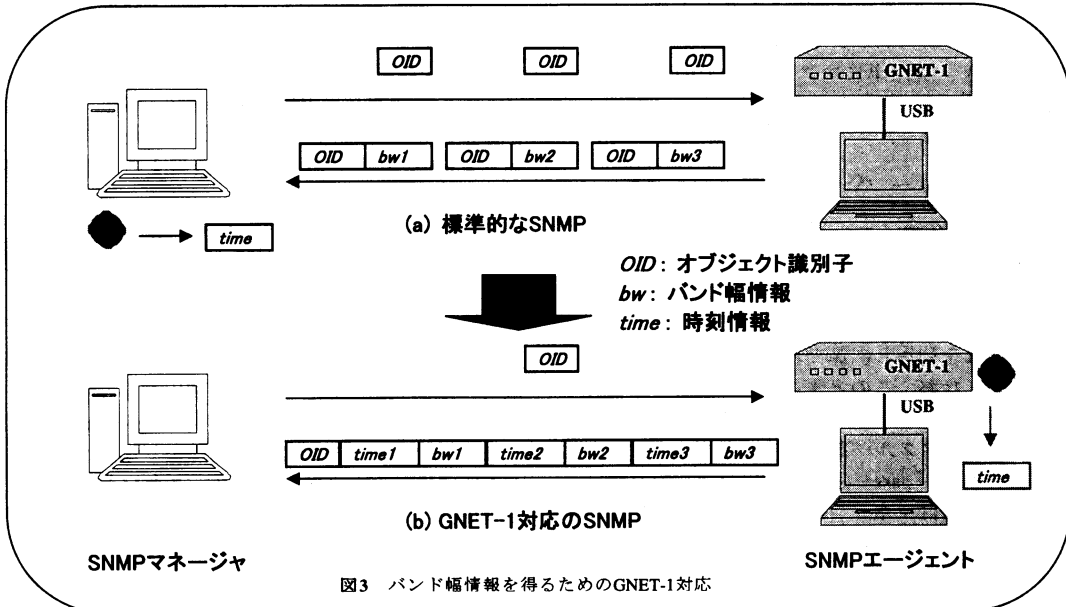


図3 バンド幅情報を得るためのGNET-1対応

ターバルよりも十分大きなデータ個数を指定することにより、データの重複送信は許容するが、欠損は発生しにくいような構成を採った。

本方式では、マネージャ側では重複受信したデータを除きながら記録しなければならないため処理負担が少々増加するが、もともとさまざまな統計処理をするものであることからこの程度の負荷の増加は特に問題はないものと考えられる。

3. GNET-1対応SNMPエージェントの実装

図4はここで実装されるプロセスとその間の関係を示している。マネージャからのSNMPリクエストはnet-snmppが受信しgn1dがGNET-1とのインターフェイスを受け持つ。また、バンド幅情報に関してはbwmdを経由する。以下にこれらについての詳細を述べる。

3.1. bwmd — バンド幅取得プログラム

最初にSNMPエージェントとは独立にGNET-1からバンド幅情報を取得するためのプログラムを作成した。このプログラムによりSNMPエージェントとなるホストPC内の共有メモリ上にはGNET-1からUSBを経由して読み込まれたデータが展開され、このデータには他のプロセスからIPC (Inter Process Communication) を利用してアクセスすることができる。1組のデータの大きさは16バイトで時刻とバンド幅がそれぞれに8バイトずつを占めている (double precision 形式)。また、使用される共有メ

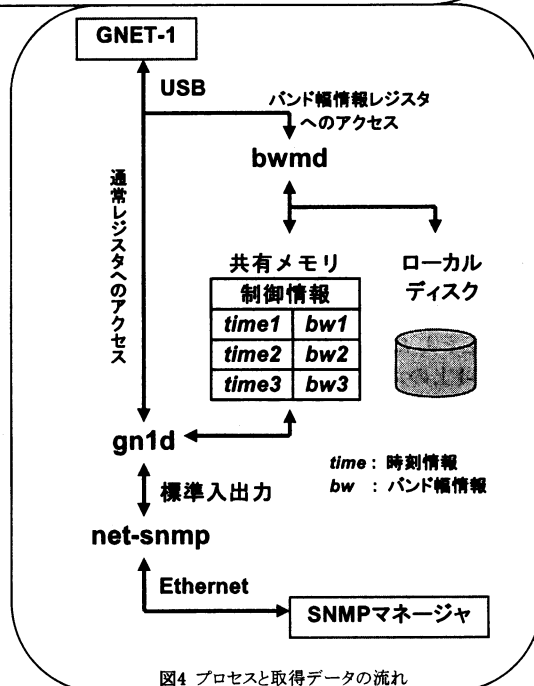


図4 プロセスと取得データの流れ

モリ領域はリングバッファの構成を採り、その大きさはプロセスの起動時にシステムの許容する範囲で任意に指定できるようになっている。なお、この共有メモリへのアクセスに際してはセマフォによる排他制御を行っている。

このようにGNET-1へのアクセスプロセスを

SNMPエージェントプロセスと分離した構成を採ることにより、エージェントが複雑化するのを避けることができ、また単にホストPC上でのみバンド幅の観測を行いたいときもこのプログラムを利用することが可能となった。^{*1}

3.2. net-snmp — ベースSNMPエージェント

SNMPインターフェイスとなるエージェントにはnet-snmp-5.0.8^[4]を使用した。このエージェントは指定したOIDを受信したときに特定のプログラムを起動できるパススルー制御と呼ばれる機能があり、本稿ではこれを利用し次節で述べるgn1dを呼び出すことにした。なお、このときコマンドやデータは標準入出力を経由してnet-snmpとやりとりされる。

またこのパススルー制御には起動プログラムを指定する方法に「pass」と「pass_persist」の2種類があり、これらはマネージャからの要求に指定したプログラムが返答した後に、プロセスが終了するか稼働し続けるのかが異なる。本稿ではホストPCがGNET-1と繋がっているUSBインターフェイスのオープンに時間が掛かり、プログラムの起動に時間を要することを考慮して後者のpass_persistを利用することにした。

3.3. gn1d — GNET-1のための拡張

GNET-1のコントロールは装置内のレジスタにアクセスすることによってなされる。本稿の実装においてはGNET-1上の全てのレジスタに対し個別のOIDを割り当て、リモートホストからはこのOIDにアクセスすることによってGNET-1のコントロールを可能とした(gn1d)。net-snmpはマネージャからGNET-1上のオブジェクトを示すOIDを受信するとパススルー制御によりこのgn1dプロセスを起動する。gn1dはnet-snmpから渡されたコマンドとOIDに基づきGNET-1内のレジスタへのアクセスを行い、結果を標準出力経由でnet-snmpへと返す。その後gn1dはデーモンとして常駐しnet-snmpからの再度の呼び出しを待つ。ただし一定時間呼び出しが無いと自動的に消滅する実装とした。以上の動作を行わせるためnet-snmpの設定ファイルに以下の1行を追加した。

```
pass_persist .1.3.6.1.4.1.18936.1 /usr/local/gnet1_snmpd/gn1d
```

ここで数字の部分はOIDを表し、それに続くパスは起動するプログラムを表している。「18936」は産総研に割り当てられたPEN (Private Enterprises Number)

であり、それに続く「1」はグリッド研究センターを意味している。この設定によりGNET-1のレジスタにSNMP経由でアクセスすることが可能となる。

一方、高精度バンド幅情報の取得に関してはgn1dが直接GNET-1にアクセスするのではなく、前節で述べたbwmdを利用しデータを取得する実装にした。bwmdをSNMPエージェントとは独立なプロセスとしてホストPC上に常駐させ、GNET-1からUSB経由でバンド幅情報を定期的に読み込み、共有メモリ上に展開する。gn1dが高精度バンド幅情報の取得を行うときは直接GNET-1にアクセスするのではなく、ホストPCの共有メモリにアクセスすることによって間接的に観測データを取得し、標準出力を経由させてベースのSNMPエージェントであるnet-snmpに渡す。

また、マネージャ側ではgn1dにより拡張されたオブジェクトにアクセスするため、それに対応する拡張MIBファイルを作成した。これにより、例えばGNET-1内部のクロックタイマーの情報を表示するには、net-snmpパッケージに付属しているsnmpgetを利用しマネージャ側から、

```
% snmpget -v 1 -c public1 192.168.0.2 %
AIST-SNMP-MIB::CLOCK.0
```

とすればよい(-vと-cはバージョンとコミュニティを指定し、IPアドレスはエージェントを指定している)。ここでCLOCKはこの拡張MIBによりGNET-1の内部クロックタイマーを示すレジスタに割り当てられており、AIST-SNMP-MIBはその拡張MIBの識別子である。また1台のホストPCに複数のGNET-1を接続できる仕様になっているため、最後にインターフェイス番号(この場合は0)を付加し区別している。

4. 性能評価

4.1. セットアップ

評価測定に使用したPCの諸元はマネージャ側とエージェント側で共通であり以下の通りである。

```
CPU : Intel Xeon 2.4GHz (FSB:533MHz)x2
Memory : DDR-SDRAM 2G
NIC : Intel PRO100
OS : RedHat8 (Linux kernel 2.4.21-smp)
```

^{*1}bwmdは0.5ミリ秒間隔までの連続バンド幅測定が可能である。

これら2台をネットワークに接続し双方のマシンにnet-snmpをインストールし、エージェント側のマシンをUSBケーブルでGNET-1と接続した。また、被観測用のテストストリームを流すためにGNET-1のGbEポートにGbEのNICを搭載した2台のPCをEthernetで接続し、内1台のラインにはビット反転によるパケットエラー遅延発生器^{★2}を挿入した(図5)。

4.2. SNMPによる高精度バンド幅情報の取得

リモートホストからSNMPによる高精度バンド幅情報が取得できることを確認するため、GNET-1を経由して接続されたPC間でiperfを使用したTCP通信を行い、その途中で人工的にパケットエラーを発生させ観測を試みた。

bwmdをGNET-1から1ミリ秒間隔のデータを取得するようにパラメータを指定して起動し、リモートホストから本稿で実装したsnmpエージェントら(net-snmp及びgn1d)を経由してデータ取得を行った。このときマネージャにはnet-snmpに付属していたサンプルプログラムを本実装に適合するように修正したものを使用した。リクエストを送信する時間間隔は10ミリ秒とし、1回で直前の20ミリ秒のデータを返すように指定した。

図6にこれによって実際にパケットエラーを発生

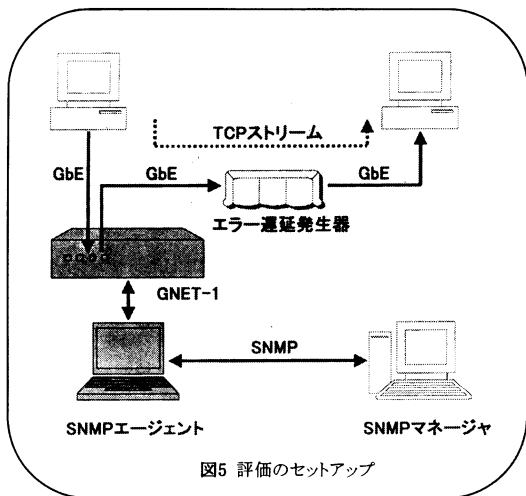


図5 評価のセットアップ

させたときに得られた時間とバンド幅をプロットしたものを示す。エラー発生時に一旦バンド幅が急激に減少しその後回復する様子をはっきりと見える。また図7はこの際のマネージャとエージェント間のパケットのやり取りをtcpdumpでモニタしたものである。マネージャ (IP:192.168.0.1) がリクエスト (GetRequest) を送信しエージェント (IP:192.168.0.2) がレスポンス (GetResponse) を行っていることがわかる。ペイロードを見ると時刻情報 (0.001秒おき) とバンド幅情報 (単位Mbps) が交互に20組送られていることがわかる。

次に遅延を入れ、GNET-1で測定間隔を100、10、1ミリ秒としたときの観測を行った。挿入した遅延時間は30ミリ秒で上記と同様にiperfを用いてTCPストリームを発生させた。ただし遅延に対応させるためオプションでウィンドウサイズを8メガバイトにした。結果を図8から10に示す。なおマネージャからエージェントへのリクエスト間隔はそれぞれ1000、100、10ミリ秒で、リクエストが到着した直前の20組のデータを返送するように指定した。つまり、直前の2000、200、20ミリ秒のデータが返送される。測定間隔を短くし粒度を上げていくことによって、TCPの詳細な振る舞いが観測できることがこの図からわかる。今までのSNMPによるバンド幅情報の収集方法ではこのような観測は不可能であり、新規プロトコルやソフトウェアの開発に大きく寄与できるものと考えられる。

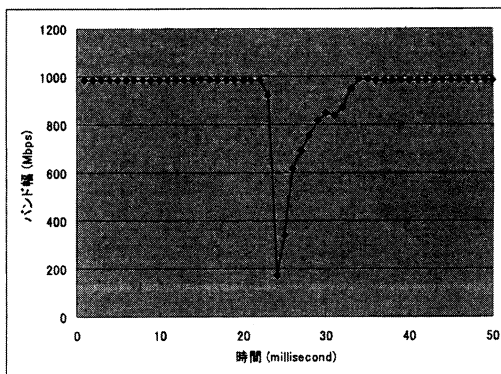


図6 エラー発生時の様子

```
11:22:38.996260 192.168.0.1.32775 > 192.168.0.2.snmp: GetRequest(33) E:18936.1.2.1.0.1 (DF)
11:22:38.996830 192.168.0.2.snmp > 192.168.0.1.32775: GetResponse(375) E:18936.1.2.1.0.1="0.022
984.048375 0.023 918.192692 0.024 170.488041 0.025 339.624791 0.026 617.912194 0.027 692.936295
0.028 752.736033 0.029 817.896793 0.030 845.048519 0.031 837.688381 0.032 872.592496 0.033
949.480031 0.034 985.489065 0.035 986.180273 0.036 984.518037 0.037 984.009863 0.038 984.389532
0.039 984.204638 0.040 984.938463 0.041 984.953873" (DF)
```

図7 SNMP通信のtcpdumpによるモニタ

★2 実体はもう1台のGNET-1でエラー遅延を発生させている。^[7]

5. おわりに

GNET-1はFPGAと高速メモリを搭載し多様な機能を実現することができるネットワークテストベッドである。本稿ではGNET-1をネットワーク機器として、その機能を十分に活用するために必要なSNMPインターフェイスの搭載を試み、GNET-1のコントロールや1ミリ秒毎の詳細なバンド幅情報取得がリモートホストから可能なことを示した。

今後はさらにプログラムとインターフェイスの改良を行い、GNET-1の機能が利用しやすいようにしていく予定である。

謝辞

本研究の一部は、新エネルギー・産業技術総合開発機構基盤技術研究促進事業（民間基盤技術研究支援制度）の一環として委託を受け実施している「大規模・高信頼サーバの研究」の成果である。

参考文献

- [1] GNET-1 : <http://www.gtrc.aist.go.jp/>
- [2] RFC1157 : <http://www.ietf.org/rfc.html>
- [3] MRTG : <http://people.ee.ethz.ch/~oetiker/webtools/mrtg/>
- [4] net-snmp : <http://net-snmp.sourceforge.net/>
- [5] 近藤秀樹他 : SNMPによるクラスタ性能計測手法の検討と評価, 情報処理学会研究報告 2003-HPC-93, pp. 19-24 (2003).
- [6] Y.Kodama at el: GNET-1: Gigabit Ethernet Network Testbed, to appear in *Cluster2004 Sept.20-23 2004, San Diego*.
- [7] O.Tatebe at el : The Second Trans-Pacific Grid Datafarm Testbed and Experiments for SC2003, *SAINT 2004 Wrokshop*

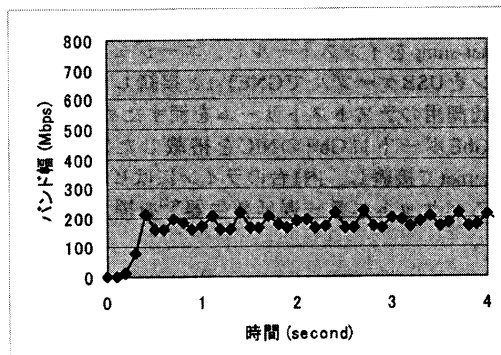


図8 測定間隔100ms

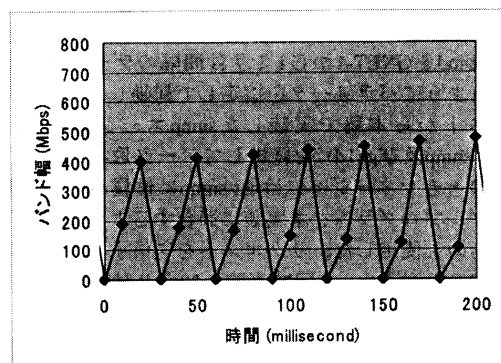


図9 測定間隔10ms

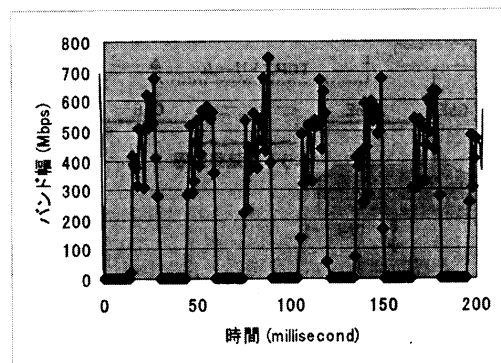


図10 測定間隔1ms