

プラットフォーム独立分散システム評価ツールによる 高速ネットワークの性能計測

山之上 卓, 戸田哲也, 望月雅光, 甲斐郷子, 中山 仁, 大西淑雅

九州工業大学情報科学センター

プラットフォーム独立分散システム評価ツールによって、ギガビットスイッチなどの高速ネットワーク機器の性能計測を行ったことについて述べる。ここで述べる分散システムの性能評価システムは、LAN やインターネットで動作する、遠隔操作、操作記録採取および操作再生が可能なアプリケーション群で構成されており、教育用分散システムで良く発生する一斉操作時の応答時間などを測定できる。アプリケーションの一つとして、TCP によるデータ転送時間を計測するものを加え、ネットワーク上で複数のデータ転送を同時に行い、その転送時間を計測することによって、高速ネットワーク機器の性能の計測を行うことができた。

Performance Measurement of High Speed Networks by a Platform independent Evaluation system for Distributed Systems

Takashi Yamanoue, Tetsuya Toda, Masamitsu Mochizuki, Kyoko Kai,
Hitoshi Nakayama, Yoshimasa Ohnishi

Information Science Center, Kyushu Institute of Technology

Performance measurement of high speed network equipment such like a Gigabit switch by a platform independent evaluation system for distributed systems is shown. This evaluation system includes remote controllable, record-able and replay-able applications. An application, which measures the bandwidth of the TCP connection between two nodes, is also included in this system. We have measured the performance of high speed network equipment by exchanging data between plural nodes of the network simultaneously by using this system.

1. はじめに

分散システムの構成要素であるネットワーク機器は分散システムの物理的および論理的な中心部分に位置しており、分散システムのなかでも重要なものの一つである。ネットワーク機器の性能そのものを測定するこ

とができれば、分散システムの構築や設計に役立てることができる。近年、ネットワーク機器の高速化が進み、カタログ値で 1Gbps の性能を持ったスイッチなども普及している。しかしながら、これらの機器が本当にそれだけの性能を発揮できるかどうかについて、従来は高価なネットワークテストなどを使って計測する必要があり、エンドユーザ側で調査

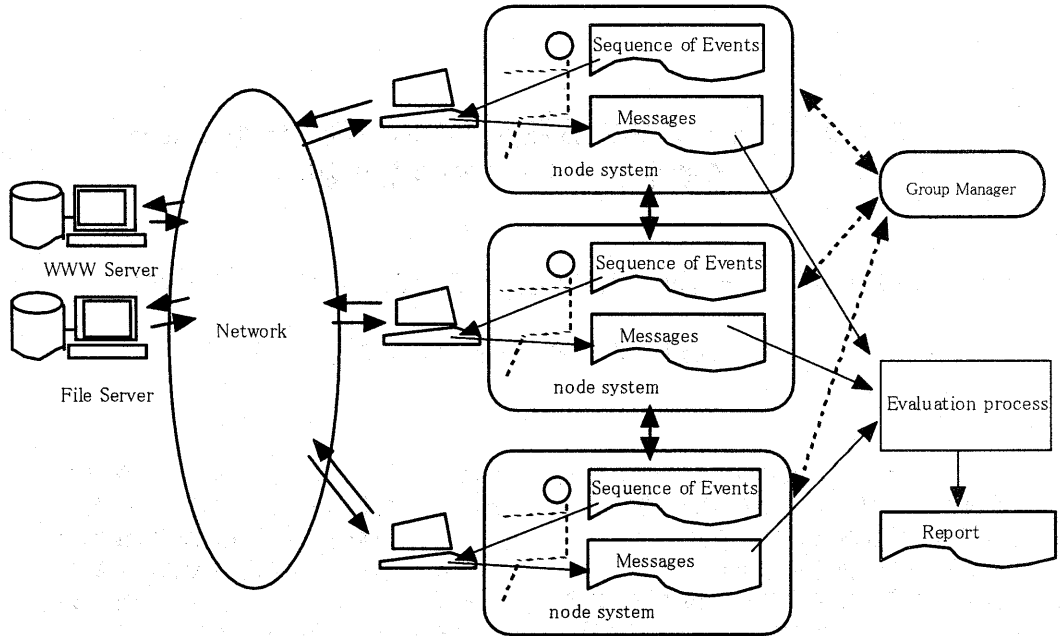


図1 分散システムの性能評価システムの概要

することは難しかった。

分散システムは、ここ数年のわずかな間で急激に普及しており、そのシステムを構成するネットワーク機器も技術的進歩により目まぐるしく変化している。このため、分散システムを導入し運用するための経験や知識の蓄積は、ユーザー側にも、メーカー側にも十分であるとは言いがたい。従って、分散システムの導入作業や管理運用作業には多大な労力を伴う場合が多い。また、システム運用時間になんらかの不安定動作を発見した場合、どの機器に原因があるか特定することは難しい。

これらの問題を解決する手段として、分散システムの性能評価基準を与えることが考えられる。この評価基準によりシステム全体の処理能力、コマンドやイベントに対する反応時間、MTBF、MTTFなどを自動的に公平に計測できれば、分散システムの比較も可能となり、分散システムの導入や管理コストの削減に役立つ。

我々は、遠隔操作、操作記録採取および操作再生が可能なプラットフォーム独立アプリ

ケーション群で構成された分散システムの評価システムを開発している[3][4]。このシステムは、マルチメディア作図システムやプログラミング環境などを含んでおり、分散システムの評価以外にも、教室内教育支援、遠隔講義、自習などに利用することができる。Javaのアプリケーションでシステムが実現されているため、ハードウェアやOSの種類にかかわらず、一般的な教育・オフィス用分散システムの多くで利用することができる。

本論文では本性能評価システムを使ってギガビットネットワークスイッチなどの高速ネットワーク機器の性能測定を行ったことについて述べる。

2. 分散システムの性能評価システム

本システムは分散システムを構成するコンピュータそれぞれの上で動作する「ノードシステム」と、ノードシステム間の接続を管理する「グループマネージャ」などによって構成される。図1に本システムの概要を示す。

2.1 ノードシステム

ノードシステムは、個々のユーザに対応しており、分散システム上のクライアントコンピュータや CPU サーバなどに分散配置され、ユーザの要求した処理を実行する。実行時の操作記録を採取し、これを再生することによって、ユーザが行った動作を自動的に繰り返すことができる。また、他のノードシステムとの間で操作記録を送受信することにより遠隔操作が可能である。

ノードシステムは「アプリケーション群」、
「コミュニケーションノード」、
「イベントレコーダ/プレイヤー」で構成されている。アプリケーション群はマルチメディア作図システム、プログラミング環境、ノード間データ転送速度測定ツールなどを持つ。コミュニケーションノードは、通信を行ったり、アプリケーションの起動を行ったりする。イベントレコーダ/プレイヤーは、通信記録や操作記録の採取および、これらの記録をアプリケーション群に伝えることによって操作の再生を行う。アプリケーションの操作はイベントの列であると考えられる。イベントは、その発生時刻と共に記録される。アプリケーション群が出力するメッセージも、時刻とともに記録される。

2.2 グループマネージャ

ノードシステム間は TCP で結合されて1つのグループを構成し、チャットや遠隔操作による一斉動作が行えるようになっている。

グループに新しいノードシステムを加えたり、グループからノードシステムが離れたときなどの処理を行うために、グループマネージャが分散システム上に配置されている。

ノード間の結合は、2分木になるよう階層的に行われる。このことによって数十台のノードを一斉に操作するときでも、手動では計測不可能な程度の時間内で、ほぼ同じに行うことができる[1]。

3. 高速ネットワーク機器の性能測定実験

本性能評価システムを使って、100Mbps

スイッチと 1Gbps スイッチの性能測定実験を行った。この実験は、測定対象となるスイッチにノードシステムを動作させるパソコンを複数接続し、この複数のノード間で一斉にデータ転送を行うことによって、スイッチの性能を測定するものである。

この測定には、ノード間データ転送速度測定ツールを使った。ノード間データ転送測定ツールは、他のノードのノード間測定ツールとの間と TCP 接続を行い、このノード対の間でデータを転送させて時間を計測することによって、データ転送速度を測定するものである。

本評価システムの遠隔操作機能により、複数のノード間データ転送測定ツールを一斉に操作させることによって、大量のデータを一斉にネットワークに流し、その転送速度を測定することができる。

ノードシステムを動作させるパソコンは、

- CPU: celeron400Mhz
- Memory: 192Mbytes
- Network Interface: 100Mbps
- OS: Turbo Linux
- JDK: Blackdown 1.1.7

を使用した[5]。ノード間の通信プロトコルは TCP を使用した。実験のまえにネットワークテストを利用して、本性能評価システムで計測されるデータ転送速度がほぼ正しいことを確認しておいた。

3.1. 100Mbps スイッチの性能測定

100Mbps ポートを 24 個持つスイッチにパソコンを 24 台接続して本性能評価システムを動作させ、このスイッチの性能測定を行った。図 2 にこの実験の概要を示し、表 1 と表 2 に実験結果を示す。

この実験では、24 のノードシステムを 2 台ずつ 12 対に分けて各ノードの対の間で、100Mbyte のデータを一斉に往復させて、データ転送時間を測定した。

その結果、12 のノード対の平均転送速度が 92 Mbps であり、このような一斉データ転送においても、ほぼ 100Mbps の速度が得られていることがわかる。ただし、50Mbps 未満の速度のノード対も 1 つあった。

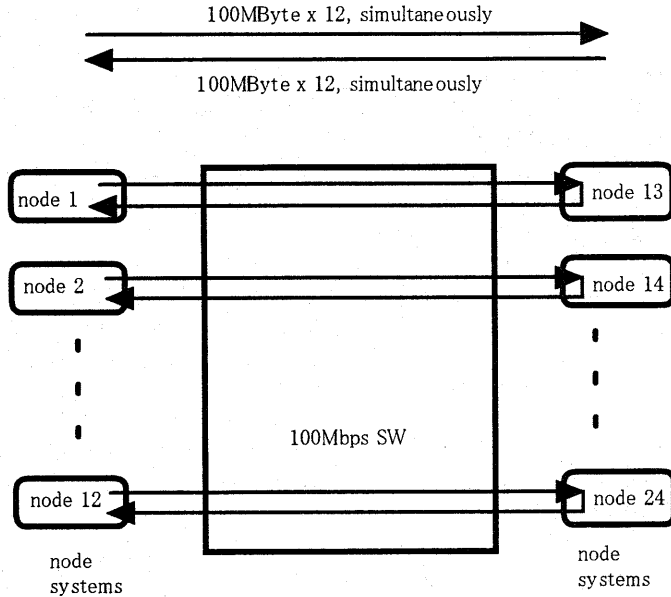


図2. 100Mbps スイッチ性能測定実験

表1. 100Mbps スイッチ性能測定実験結果

総合経過時間	34.64 秒
総合転送量	1.2Gbyte x 2 (往復)
総合通信容量	0.554Gbit/秒
ノード対の最大転送速度	111Mbps
ノード対の最小転送速度	46.4Mbps
ノード対の平均転送速度	92Mbps

表2. 図2の実験における
各ノード対の転送速度

速度	ノード対数
100Mbps 以上	5 対
90-100Mbps	4 対
80-90Mbps	0 対
70-80Mbps	1 対
60-70Mbps	1 対
50-60Mbps	0 対
40-50Mbps	1 対

3.2. 1Gbps スイッチの性能測定

1 Gbps スイッチの2つのポートに 1G-100Mbps スイッチを2台接続し、それぞれの 1G-100Mbps スイッチの 100Mbps ポートにパソコンを接続して本性能評価システムを動作させることによって、この 1G スイッチの性能測定を行った。図3にこの実験の概要を示し、表3に実験結果を示す。

この実験では、1Gbps スイッチをまたいだ 22 対のノードシステム間で 100Mbyte のデータを一斉に往復させて、データ転送時間

を測定した。

その結果、総合的なデータ転送速度として約 1Gbps が得られることを確認できた。

4. 関連研究

ネットワークの測定については、Netperf[6]、DBS(Distributed Benchmark System)[7]などのツールがある。これらのツールは TCP だけでなく UDP にも対応している。

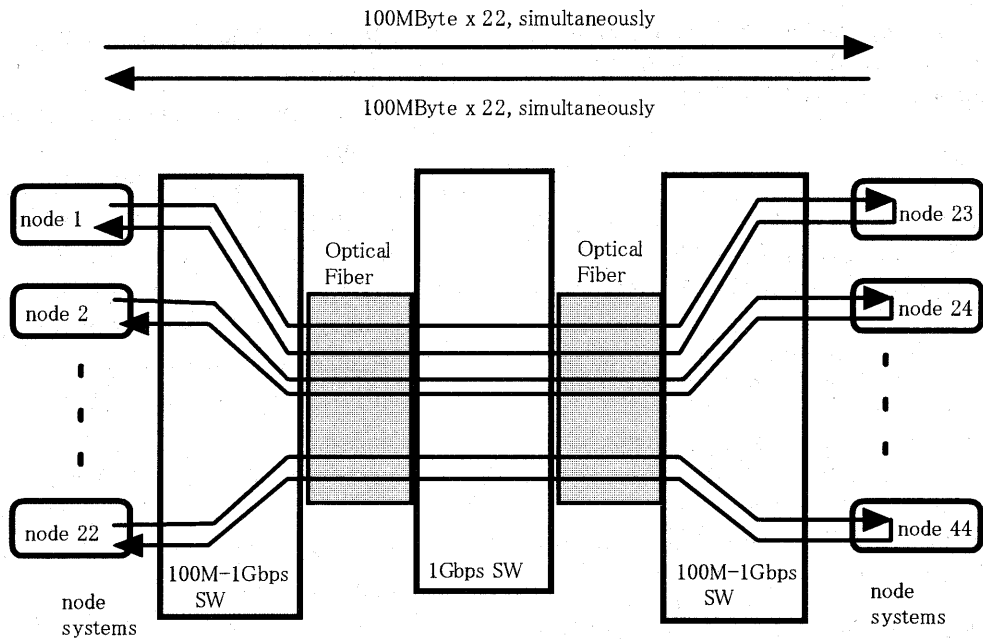


図 3. 1Gbps スイッチの性能測定実験

表 4. 1Gbps スイッチ性能測定実験結果

総合経過時間	33.04 秒
総合転送量	2.2Gbyte x 2 (往復)
総合通信容量	1.07Gbit/秒
ノード対の最大転送速度	75.2Mbps
ノード対の最小転送速度	49.7Mbps
ノード対の平均転送速度	67.7Mbps

Netperf は基本的には 1 対 1 の通信の性能測定を行うツールである。

DBS は複数のホスト間で同時に複数のデータ転送を行うことができると同時に、各々のデータ転送のスループットの時間変化も測定できる。また、TCP コントロールブロックの値を記録できるなどきめ細かな測定が可能である。しかしながら現時点で DBS を動作させるためには、UNIX 系の OS がインストールされていることが必要である。

これに対して本評価システムは現時点では UDP の測定を行えず、DBS ほどきめこ

まかな測定も行うことはできない。しかしながら、スクリプトに従って複数のホストを同時に制御することなどについては、DBS とほぼ同様な機能を持っている。また、UNIX 以外の OS でも本評価システムを動作させることができる。DBS はネットワーク性能測定専用ツールであるのに対して、本評価システムは分散システム全体の性能測定を行うことができる他、教育支援システムとしても利用可能である。

分散システムに負荷を与えてその性能を測定するツールとしては、本システムの他に、

Mercury Interactive 社の LoadRunner[1]があるが, LoadRunner は現在在のところで, Windows システムでしか利用できない.

5. おわりに

分散システムの評価システムを使って高速ネットワーク機器の性能測定を行ったことについて報告した. Java のアプリケーションでシステムを実現しているため, ハードウェアや OS の種類にかかわらず, 一般的な教育・オフィス用分散システムの多くで利用することができる.

本アプリケーション群は, マルチメディアデータを含むすべての操作履歴とその実行時メッセージを記録するため, メモリやファイルなどの大量の資源を消費する. また, Java でシステムを開発しても, OS や JDK のバージョンが異なると, 同じプログラムが異なる動作をする場合があるなどの問題もある. これらの原因により, 現状では安定性などに問題があり, 現在改良を行っている. また, 現時点ではセキュリティの機能を持っていない. 実際に利用するためにはセキュリティ機能を付加する必要もある.

今回は, ネットワーク上で行うことができる通信の, ごく一部の組み合わせの測定しか行っていない.

今後はもっと多くの組み合わせによって測定実験をおこない, より精度の高いネットワーク機器の評価を行っていく必要がある.

謝辞

本稿で述べた研究の一部は文部省科学研究費基盤研究一般(C)(2)09680401 の補助を受けた.

参考文献

[1] Takayuki Hirahara, Takashi Yamanoue, Hiroyuki Anzai and Itsujirou Arita, "SENDING AN IMAGE TO A LARGE NUMBER OF NODES IN SHORT TIME USING TCP", Proceedings of the ICME2000, IEEE International Conference on Multimedia and Expo, New York City,

- USA, July 30-Aug. 2, pp. 987-990, 2000.
- [2] "LoadRunner",
<http://www.mercury.co.jp/products/loadrunner/>
- [3] 山之上卓, 望月雅光, 甲斐郷子, "遠隔操作, 操作記録採取および操作再生が可能なプラットフォーム独立アプリケーション群", 情報処理学会分散システム/インターネット運用技術研究会報告情報処理学会研究報告, 15-10, pp.55-60, 1999.
- [4] 山之上卓, 望月雅光, 甲斐郷子, "分散システムとインターネットアプリケーションの性能評価システム", 情報処理学会分散システム/インターネット運用技術研究会報告情報処理学会研究報告, 20-3, pp.55-60, 1999.
- [5] 中山仁, 大西淑雅, 望月雅光, 山之上卓, 甲斐郷子, "Linux thin client を端末とする集合教育用計算機環境の構築", 情報処理学会分散システム/インターネット運用技術研究会報告, 2000-DSM-18, pp. 31-36, 2000.
- [6] Hewlett-Packard Company, "Netperf: A Network Performance Benchmark Revision 2.1", 1995.
- [7] Yukio Maruyama, Suguru Yamaguchi, "DBS: a powerful tool for TCP performance evaluations", Performance and Control of Network Systems, Proceedings of SPIE, Volume 3231, November 1997.