

トラヒック測定技術：NEPRI

Traffic Measurement Technique: NEPRI

あらまし

近年、IPネットワークにおいて通信品質の保証や、負荷分散などの動的制御を行いたいという要求が高まってきており、そのためには、ネットワークの性能を的確に把握することが重要になってきている。しかし、今日のIPネットワークは、多種多様なネットワーク機器を相互接続してボトムアップに構築しているため、性能を把握することはなかなか困難である。

本稿では、従来技術の問題点である測定時の負荷増大および測定精度の低下について述べるとともに、富士通で開発を進めているネットワーク経路のボトルネックの帯域幅をプロービングによって効率良く測定できる新しい手法“NEPRI”を紹介する。本手法は、従来手法であるRTT(Round Trip Time)の値を求めるのではなく、プロービングパケット間のRTTの差が下限となるボトルネックの帯域幅を算出することによって、輻輳したネットワークでも少量のパケットを送信するだけで従来手法と同程度の精度で帯域幅を測定できる。

Abstract

Recently, there have been strong demands for better communication quality in the IP network and dynamic control of load distribution. To answer these demands, precise measurement of network performance has become vital. However, because today's IP network has been constructed bottom-up by interconnecting various network devices, such measurement is very difficult. This paper describes the increase in measurement load, the deterioration of measurement accuracy, and the problems associated with conventional technologies. Also, this paper introduces a new technique that Fujitsu is developing called“NEPRI.” NEPRI enables efficient measurement by probing the bottleneck bandwidth in the network path. This technique does not calculate the Round Trip Time(RTT) which is the conventional way, but calculates the available bandwidth of the bottleneck from the correlation between the RTTs of probing packets. This calculation enables the bandwidth to be measured as accurately as with the conventional technique simply by transmitting a small number of packets over a congested network.



勝山恒男(かつやま つねお)

(株)富士通研究所IPサーバプロジェクト部 所属
現在、IPネットワークのミドルウェア関連の研究開発に従事。



安達基光(あだち もとみつ)

(株)富士通研究所IPサーバプロジェクト部 所属
現在、IPネットワークの管理、制御技術に関連したミドルウェアの研究開発に従事。

ま え が き

IPネットワークの成功と発展は、多種多様のネットワーク機器から成るネットワークを相互接続させながら、段階的に規模を拡大していくボトムアップ型のネットワークの構築方法によるところが大きい。しかし、その反面、事前のトラフィック予測に基づいて設計された電話網とは違い、いわゆるベストエフォート型^(注1)のネットワークであり、ネットワーク管理者であっても、ネットワークの性能やサービス品質を正確に把握することはきわめて難しくなっている。ところが、ネットワークの情報インフラとしての重要性が増すにつれて、このようなベストエフォート型のサービス提供では不十分になり、何らかの品質保証を実現することが必要になってきている。つまり、ネットワークは、サービス機能を提供するだけではなく、性能面での監視・制御をしなければならなくなっている。一方、最近ではポリシー制御型ネットワーク^{(1),(2)}のように、動的にネットワークを制御する枠組みが提唱されており、これからはネットワークの稼働状態を逐次把握しながら、ネットワークの動的制御や運用管理を行っていくことが重要であり、そのための基本要素技術として測定技術の果たす役割は大きくなっている。

従 来 技 術

IPネットワークの計測技術とは、図-1に示すようにリンクの帯域、サーバの負荷状態などの稼働状態を把握し、逐次制御やネットワーク設計などを実現するための技術である。計測技術を体系的にまとめたものを以下に述べ

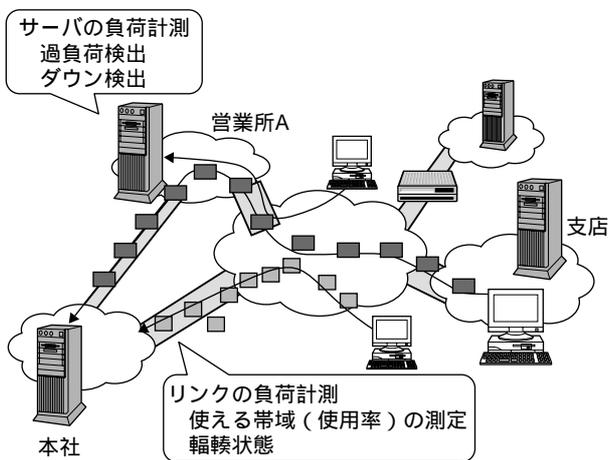


図-1 IPネットワークの負荷計測
Fig.1-IP network performance measurement.

(注1) 利用状況によって通信性能が変化する。

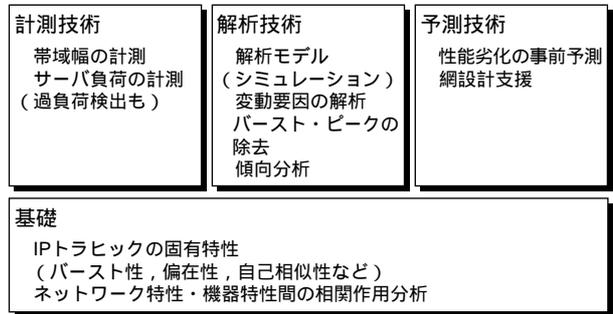


図-2 IPトラフィック計測技術・予測技術の体系
Fig.2-Outline of IP traffic measurement and prediction technology.

る(図-2)。

- (1) 実際の稼働状態を測定する技術
- (2) そのデータをもとに要因分析や傾向分析などを行う解析技術
- (3) 将来のトラフィック動向やネットワーク設計支援を実現する予測技術
- (4) IPパケットの基本的な振舞いや特性を分析する基礎技術

以下、本稿においては、とくにリンクの状態を把握する計測技術について詳しく述べる。

従来のIPネットワークの性能測定の方法としては、アクティブ型の計測技術とパッシブ型の計測技術に大きく分けられる。パッシブ型計測と言われるものは、ルータやブリッジなどのネットワーク機器が持つ管理情報ベース(MIB: Management Information Base)をSNMP(Simple Network Management Protocol)やRMON(Remote Monitoring)によって取得して、ネットワーク機器の性能を測定する方法が取られている。しかし、これらの性能測定では、個々のネットワーク機器の性能は分かっても、ネットワーク全体の性能を総合的に把握することは難しい。そこで、最近では、着目しているネットワーク部や経路の性能を直接測定する方法として、プローピングによるアクティブ型の性能測定が研究されている。これはプローピングパケットをネットワークに送信し、そのレスポンス時間などを測定することにより、ネットワーク性能を実測する方法であり、調べたい経路の性能をよりの確に測定できる。

従来のアクティブ型計測技術

アクティブ型の性能測定とは、プローピングパケットをネットワークに送信してそのレスポンスを測定することで、経路のホップ数^(注2)、パケット遅延、ボトルネックの

(注2) 経路中に存在するルータの数。

帯域幅や利用可能帯域幅、パケット廃棄率などのネットワークの経路の性能を実測する技術である。古典的な方法は、ICMP(Internet Control Message Protocol)エコーパケットを利用したping^(注3)であり、経路上のホストの到達可能性とパケットのラウンドトリップタイム(以下、RTT)を測定できる。またtracerouteは、TTL(Time To Live)を一つずつ増やしながらUDP(User Datagram Protocol)エコーパケットを送信することによって、経路上のルータのIPアドレスを一つずつ調べ、経路上に存在する任意のルータまでのRTT、廃棄率を測定する。さらに、最近のpathchar⁽³⁾と言われるツールは、いろいろなサイズのUDPパケットを組み合わせてプローピングを行い、RTTの最小値を求め、パケットサイズに対するRTT最小値の増加傾向から各ホップのリンクの帯域幅を測定しており、tracerouteの測定精度を大きく向上させている。

また、Treno⁽⁴⁾はtracerouteのようにTTLを増やしながらUDPエコーパケットを送信し、ICMP TTL超過パケットを受信することによってリンク性能を測定するが、TCP(Transmission Control Protocol)のウィンドウ制御と同様^{ふくそう}の輻輳回避アルゴリズムを行いながらパケットを送信する。したがって、TCP/IPによる通信を行った場合の性能(スループット、廃棄率、ウィンドウサイズ、RTT)を測定することができる。しかし、Trenoやpathcharは測定精度の向上のために、多数のパケットを送信し、ネットワークに過大な負荷をかけるため、実用には制限がある。

着眼点

前節で述べた従来技術の問題点は、測定精度の改善のために、プローピングの回数を増やす方法を取っていることにある。pathcharでは、プローピング回数を増やすことにより、RTTの最小値を正確に求めることができ、リンクの帯域幅の測定精度を向上できる。また、TrenoはTCP/IPによる通信を行う時間を長くすればするほど、ウィンドウ制御のサイクルが定常状態になり、スループットの測定精度が良くなる。しかし、ほかのトラヒックがプローピングパケットの間に混在するなどの外乱が入った場合は、精度が著しく低下するといった問題点がある。プローピングは元来、ほかのトラヒックの影響などネットワークの外乱要因による測定誤差を含むものであり、精度を改善するためにはプローピングの回数を増やしてノイズを除去せざるを得ない。しかし、これでは、測定精度は上がっても測定にかかるオーバーヘッドが大きくなり、精度とコストのトレードオフの壁にぶつ

かってしまう。測定精度の改善のために、プローピングの数を増やしていくという方法には必ずと限界があり、本質的に誤差の影響が避けられない。これは、本来ばらつきが多いプローピングパケットの個々のレスポンスからノイズを除去しようとするのは困難であるために、多数のプローピングを行って統計的に処理をして、誤差をフィルタリングした結果である。

これに対して、個々のプローピングを独立に評価するのではなく、プローピング間に見られる相関に着目すれば、個々のプローピングに誤差が含まれていても、その誤差の影響を受けずに、測定精度を上げることができるものと考え、ボトルネックにプローピングパケットがキューイングされたときに現れるプローピング間の相関を検出し、プローピングの数を増やすことなしに性能を計測する技術を考案した。

新計測技術

上記ポイントに着目した新しい計測技術NEPRI(NETwork Performance measurement by Round-trip-time Increase)では、RTTにばらつきがあるような、実運用中のネットワークでも、外乱の影響を受けずに、精度の良い測定が可能になり、精度とコストのトレードオフの問題を根本的に改善することができる。

基本原理

プローピング間の相関は、プローピングパケットがボトルネックに同時にキューイングされたときに特異的に現れる。それは、プローピングパケットが同時にキューイングされない場合、各プローピングは独立した事象になるが、プローピングパケットが同時にキューイングされた場合、前のプローピングパケットが送出されてからでない、後のプローピングパケットは送出されないため、必然的に前後のプローピング間で相関が現れるからである。以下、本技術の基本原則、特徴を示す。

(1) 相転移をマクロに観察

経路に計測用パケットを一定速度でプローピングし、そのRTTを測定する。図-3のようにRTTの時系列 $i = 1, 2, \dots$ を xy 平面にプロットすると、速いプローピングのときや、ほかのトラヒックが多いときは、「相転移」^(注4)が現れ、下限線が検出される。これはボトルネックのルータに計測用パケットが二つ以上同時にキューイングされた状態を示すものである。ボトルネックのリンク速度は、この下限線の切片の値から算出することができ、

(注3) ホスト間でエコーパケットを往復させ、それにかかる時間を計測する。

(注4) パケットの遅延の振舞いが変化すること。

他方、利用可能帯域幅は、相転移が現れ始めたときのプロービング速度から求めることができる。

(2) 輻輳やホップ数によらない測定が可能

本手法は経路性能に固有な相転移を検出することで性能測定を行うため、外乱の影響を受けにくい。したがって、経路が輻輳していたり、ホップ数が大きかったりする場合でも、測定精度が劣化しない。社内の運用ネットワークを用い、その経路上の10 Mbpsのリンクに20~80%の負荷を与え、利用可能帯域幅を計測した。本手法による計測結果(表-1)は以下のとおりまとめられ、本手法ではネットワークへの負荷を抑えつつ、従来技術と同程度の精度を得られることが確認できた。

- ・計測ばらつき誤差：10~20%以内
(0.75 Mbps以内、リンクの負荷状態による)

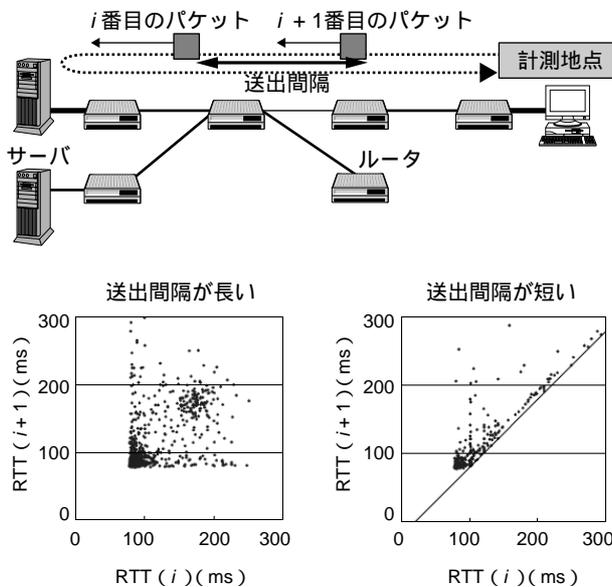


図-3 プロービング速度を変えたときのRTTの相関関係
Fig.3-Relation between RTTs when probing speed is changed.

表-1 富士通社内の運用ネットワークでの計測結果

利用可能帯域	8 Mbps (80%空き)	6 Mbps (60%空き)	4 Mbps (40%空き)
計測誤差	10%以内	15%以内	20%以内
与えた負荷	1.2 Mbps (15%負荷)	1.1 Mbps (18%負荷)	0.79 Mbps (20%負荷)
計測時間	2.9秒	3.3秒	4.8秒

表-2 従来方式との比較

計測ツール	計測対象	計測手法	リソース消費	精度
Treno	TCP スループット	擬似TCP通信	多量のデータを長時間送受信	実測値
Pathchar	最小物理帯域	最小遅延	精度向上のために繰り返しプロービングが必要	ほかのトラヒックによる遅延のばらつきの影響大
NEPRI	利用可能帯域	相転移検出	少数のパケットでも検出が容易	遅延のばらつきの影響小

- ・ネットワークへの負荷：求める値の20%以下
- ・計測に要する時間：3~5秒以内
- ・そのほか：ホップ数の制限なし

従来方式との比較

NEPRIとpathcharを比較する。pathcharでは、ばらつきの大きいRTTをそのまま使うのではなく、パケットサイズを変えて多数のプロービングを行い、RTTの最小値を割り出し、帯域幅の計算に用いている。RTTが最小になるのはルータのキューにパケットがキューイングされていないときである。最小RTTを検出するためには多数のプロービングを行う必要があること、非常に輻輳したネットワークでは最小RTTがうまく検出されないため、測定誤差を減らすのが困難になることがこの手法の問題点である。つまり、実運用中のネットワークでは利用に制限がある。それに対して、NEPRIでは、最小RTTを用いるのではなく、プロービング速度を変化させたときに、RTTの差が下限を与えるようになる時点を検出しているだけである。ほかのトラヒックがルータに混在している場合でも、RTTの差は下限を与えるので、NEPRIは輻輳したネットワークでも計測が可能である点、pathcharより有利である。

一般に従来方式は、RTTの誤差を減らすため、多数のプロービングパケットのRTTの統計量(最小値や平均値)を用いている。本手法は、プロービングパケット間のRTTの差が正になる割合の不連続な変化を検出し、その時点でのプロービング速度から帯域幅を求めているので、RTTのばらつきがあっても測定精度に影響しないことが大きな特徴となる(表-2)。

む す び

IPネットワークの経路のボトルネックの帯域幅を測定する新しい手法NEPRIを開発した。

実運用中のネットワークを用いて、外乱の影響を受けずに一定の精度で経路の利用可能帯域幅を測定できることを確認した。従来手法では、遅延の値そのものを使って利用可能帯域幅を推定する技術が多かったが、それらは、ほかのトラヒックの影響による遅延のばらつきによって誤差が発生するため、トラヒックの多い経路では

計測精度が劣化する。しかし、NEPRIでは遅延のばらつきによる影響を受けない現象を検出しているため、ほかのトラヒックの影響で計測精度が劣化することはない。社内の運用ネットワークを用いて、20～80%負荷でも測定精度が劣化しないことを確認した。

本稿で述べた計測技術は、ネットワーク管理者が運用設計する上で欠かせないものである。逐次、動的にネットワーク性能を把握する技術は、wwwサーバのように複製をネットワーク内に設置してアクセスを分散させる負荷分散サービスや、ネットワークの稼働状況に応じてアクセスを制御するポリシー制御型ネットワークと呼ばれる新しいネットワーク運用サービスなどにおいても、最適なネットワーク活用の実現に大いに貢献するものと考えている。

参考文献

- (1) 河井保博ほか：ポリシー・ネットへのアプローチ．日経インターネットテクノロジー，No.027，p.84-105(1999)
- (2) Policy Framework Working Group：IETF.
<http://www.ietf.org/html.charters/policy-charter.html>
- (3) V. Jacobson：pathchar - A Tool to Infer Characteristics of Internet Paths. MSRI，1997.
<ftp://ftp.ee.lbl.gov/pathchar/msri-talk.pdf>
- (4) M. Mathis et al.：Diagnosing Internet Congestion with a Transport Layer Performance Tool. 1996.6.
<http://www.psc.edu/mathis/htmlpapers/inet96.treno.html>

