

インターネットにおける計測に基づいたボトルネック特性調査手法

Measurement Based Investigation into Characteristics of Network Bottlenecks in the Internet

的場 一峰

Kazumine MATOBA

大阪大学 大学院基礎工学研究科

Graduate School of Engineering Science,
Osaka University

阿多 信吾

Shingo ATA

大阪市立大学 工学部

Faculty of Engineering,
Osaka City University

村田 正幸

Masayuki MURATA

大阪大学 サイバーメディアセンター

Cybermedia Center,
Osaka University

1 はじめに

近年、トラフィック量が急速に増大しているインターネットにおいて、ユーザに安定した品質を提供するためには、将来必要となる帯域を正確に予測し、回線容量設計を速やかに行う必要がある。また、ユーザに適切なサービス品質を提供するためには、現在のトラフィックの状況を正確に把握し、最適なリンク帯域を導出する必要がある。そのためには、エンドホスト間の通信において発生するボトルネックを特定し、その改善方法について明らかにする必要がある。

本稿では、インターネット上のトラフィックの大部分を占めている TCP に着目し、ネットワークにおけるボトルネックの特定手法を提案する。さらに、ボトルネックとなる回線の利用率を求めることにより、最適な回線容量を設定する基準を提供する。

また、実際にトラフィック測定を行い、その結果に基づいたボトルネックの特定、および回線利用率の推定結果を示す。

2 ボトルネックの検知

ボトルネックとなる要因を調べるために、スループットをモデル化した式を利用する [1]。TCP のウィンドウサイズの期待値 $E[W]$ はパケットロス率をもとに算出できる。また、 $E[W]$ と受信側ホストのソケットバッファサイズである、ウィンドウサイズの最大値 W_{max} を比較することで、バッファサイズが不足しているかどうかを判断する。

また [1] によれば、TCP スループットはパケットロス率、ラウンドトリップ時間、受信側のウィンドウサイズの最大値、ACK の持つ到達確認の数、タイムアウト時のパケット再送間隔の 5 つのパラメータから推定することができる。ただ、このスループットの予測式は全てのボトルネックとなる要因を考慮にいれていないので、正確なスループットを予測できない場合がある。したがって、実測値との比較をするために TCP のスループットおよび物理的なリンクの帯域を計測する。

TCP のスループットを計測する際には、実際に通信をおこない、パケットの情報を tcpdump を用いて調べることによって観測をおこなう。また、リンク帯域を測定するために [2, 3, 4] で用いられている手法を利用する。

パケットロス率は tcpdump を使用し、送信したパケットに対する ACK パケットを調べることで算出する。タイムアウト時のパケットの再送間隔は、RTT の各計測ごとに再計算する [5]。

以上の計測結果を利用することで、ボトルネックとなる要因を判断することが可能になる。

3 ボトルネックの推定結果

ここでは実際に計測を行い、現在のインターネット上で TCP を利用した通信の特性を分析した結果を示す。

表 1 ~ 3 は、大阪大学のあるホストから計測を行った結果である。表の各値は、上から受信側のソケットバッファサイズ、RTT、パケットロス率、スループットの計測値、

表 1: 大阪市立大学 (bottleneck 1.47 Mbps)

buffer size	8 KB	32 KB	64 KB
RTT(ms)	4.03	17.4	21.7
packet loss(%)	0.000187	0.000946	0.00268
TCP[measure]	1.35 Mbps	1.36 Mbps	1.35 Mbps
TCP[estimate]	1.62 Mbps	2.57 Mbps	1.21 Mbps
buffer	shortage	enough	enough

表 2: ring.so-net.ne.jp (bottleneck 19.3 Mbps)

buffer size	8 KB	32 KB	64 KB
RTT(ms)	17.7	23.4	23.4
packet loss(%)	0.00007	0.00293	0.00406
TCP[measure]	2.13 Mbps	9.66 Mbps	15.6 Mbps
TCP[estimate]	2.62 Mbps	10.5 Mbps	20.8 Mbps
buffer	shortage	shortage	shortage

および推定値である。最後の欄では式 $E[W]$ を算出し、受信側のソケットバッファサイズが十分である (enough) かそうでない (shortage) かを判断した結果を示す。

表 2 の結果を見ると、RTT が小さいために受信側のバッファがボトルネックになっていることがわかる。これは、ソケットバッファサイズの欄が shortage であることと、ボトルネックであるバッファサイズを大きくすると、TCP のスループットも予測値にしたがって向上していることから判断できる。この状況は表 1、3 でバッファサイズを 8 KB にした時にも観測できる。

表 1 の結果では、受信側ホストのバッファサイズを 64 KB にするとパケットロスの総数と RTT がともに増加していることから、約 1.5 Mbps のリンクの帯域がボトルネックになっていることがわかる。

表 3 の結果において、pchar によるボトルネックリンク帯域の計測結果から、最大で約 30 Mbps のスループットが期待される。しかし実際の計測結果では、バッファサイズを増加させてもスループットは 4.8 Mbps 程度にとどまっている。また、リンク帯域がボトルネックの場合、表 1 と同様にパケットロスが増加するはずであるが、測定の結果、パケットロスに大きな変動は見られない。

これらのことを考慮に入れると、送信側ホストがパケットを送信する部分がボトルネックであると考えられ、具体的な要因として、以下の 2 つが考えられる。

- 送信側ホストから 1 ホップ目のリンクの帯域
このときにパケットの送信レートは 1 ホップ目のリンクの帯域以下になる。したがって、送信パケットが途中のルータでキューイングされないとき、パケットロス率や RTT は低い値になる。
- サーバによる帯域制限
ftp サーバ自体に帯域を制限する機構が働いている場合は、設定されているスループット以上のレートでパケットを送信することができないため、この場合もパケットロス率および RTT が低い値をとる。

上記の要因のどちらであるかを判断するためには、サーバ側から 1 ホップ目のリンクの帯域を求める必要がある。ただし、計測を行うホストと計測先のサーバとが同一のネッ

表 3: ftp.iij.ad.jp(bottleneck 29.6 Mbps)

buffer size	8 KB	64 KB	128 KB
RTT(ms)	20.8	22.9	20.7
packet loss(%)	0.000573	0.000288	0.000867
TCP[measure]	2.83 Mbps	4.33 Mbps	4.78 Mbps
TCP[estimate]	3.03 Mbps	35.3 Mbps	21.7 Mbps
buffer	shortage	enough	enough

トワーク内に存在し、ネットワーク内のリンク帯域やサーバにおける帯域制限の有無などの構成情報を知ることができる場合、これら 2 つの要因を容易に特定できる。

以上の議論より、以下の手続きを利用することでボトルネックとなる箇所を特定することができる。

1. RTT、パケットロス率、TCP スループットを計測し、ネットワーク特性に関する情報を収集する。
2. $E[W]$ を算出し、受信側のソケットバッファサイズが不足しているかどうかを判断する。
3. バッファサイズを変化させ、TCP スループットの変化を観測。

観測の結果、バッファサイズが十分でありパケットロス率が高い場合には、リンク帯域がボトルネックとなる。また、バッファサイズとスループットの変化に関連性が見られない場合、ボトルネックは送信側ホストがパケットを送信する部分と判断することができる。バッファが不足しており、バッファを変化させることでスループットが向上する場合は、ボトルネックが受信側のソケットバッファサイズにある。

ネットワークの帯域にボトルネックが存在する場合、回線増強によってボトルネックの解消を行わなくてはならない。以下では、この回線増強の指針の一つとなる回線利用率を求める手法の提案をおこなう。

4 ボトルネックリンクにおける回線利用率

4.1 回線利用率の導出

ボトルネックリンクの回線利用率を求めることのできる計測ツールとして、cprobe [6] がある。cprobe は ICMP の ECHO request パケットを連続して送信し、その応答パケットのパケット間隔を観測することで利用可能帯域 v (bps) を求める。この v の値と、ボトルネックリンクの帯域 b (bps) を利用し、ボトルネックリンクの回線利用率 u を $u = v/b$ で算出している。

cprobe では、計測トラヒックの影響を考慮に入れておらず、計測トラヒックが存在しない場合の回線利用率を導出できない。また、cprobe ではエンドホスト間の計測を行っているため、計測パス全体の情報が入手できず、ボトルネックとなるリンクに限った情報収集を行うことができない。このため、得られる測定結果が必ずしも正確ではない場合がある。

そこで、本稿ではより詳細な測定結果を用いることで、正確な回線利用率を求める手法を示す。ここでは、エンドホスト間のリンク n がボトルネックである場合の、リンク n の回線利用率を求める方法を示す。まず、ボトルネックリンクの直前のルータ $n-1$ 、ボトルネックリンク直後のルータ n のそれぞれに対して連続して計測パケットを送信する。送信パケットのサイズを s (bit)、ボトルネックリンクの帯域を b_n (bps)、普段ボトルネックリンクを通過しているトラヒックのレートを a (bps) とおく。このとき、到着するパケットの間隔がルータ $n-1$ との計測では l_{n-1} (s)、ルータ n との計測では l_n (s) であるとする。

ここで連続して送信した 2 つの計測パケットに着目する。最初のパケットがルータ n に到着したあと、 l_{n-1} (s) 後

表 4: 回線利用率計測結果 (bottleneck = 1.5 Mbps)

他のトラヒック	$l_{n-1} \rightarrow l_n$	回線利用率
なし	5.55 ms \rightarrow 5.56 ms	13.9%
0.5 Mbps	5.32 ms \rightarrow 6.62 ms	34.2%
1.0 Mbps	5.81ms \rightarrow 7.55 ms	55.1%

に 2 つ目のパケットがルータ n に到着することから、2 つのパケットの間には $l_{n-1}a$ (bit) のトラヒックが存在する。また、 l_{n-1}, l_n には、1 つ目の計測パケットを転送するために必要な時間が含まれているため、2 つのパケットの間隔 w_n は $w_n = l_n - s/b_n$ (s) となる。ボトルネックリンクでは常にトラヒックが途切れることなく通過していると考えると、 w_n (s) の区間に $l_{n-1}a$ のトラヒックが存在することから、ボトルネックリンクの帯域 b_n は $b_n = l_{n-1}a/w_n$ で表される。したがって、 a は $a = (l_n b_n - s)/l_{n-1}$ となり、これを用いてボトルネックリンクの回線利用率 u は

$$u = \frac{a}{b_n} = \frac{l_n b_n - s}{l_{n-1} b_n} \quad (1)$$

となる。

4.2 計測結果

表 4 は、大阪市立大学との間のボトルネックリンクに対して計測を行った結果である。計測時には、バックグラウンドに別のトラヒックを送出し、回線利用率の変化について調べる。

表の結果より、他のトラヒックが増加するにしたがい、回線利用率が上昇していることが分かる。特に、0.5 Mbps の場合には、理論上の回線利用率 (33.3%) に近い値が得られている。しかしながら一方で回線利用率の誤差は最大ボトルネックの 10% 程度存在している。このような誤差は、計測データのフィルタリング等により精度を向上させることができるが、その手法については今後の課題である。

5 まとめと今後の課題

本稿では、計測にもとづくネットワークのボトルネック検出手法および回線利用率の算出方法を提案した。パケットロス率、RTT、TCP スループット、計測ホストのソケットバッファサイズ、パケットロスのタイムアウトまでの時間を求めることで現在のネットワークの特性の調査を行った。さらに、回線利用率を算出する方法を提案し、それぞれの手法を適用した計測結果を示した。今後の課題としては、測定結果の精度の向上、回線利用率に基づいたネットワークの容量設計手法などが挙げられる。

参考文献

- [1] J. Padhye, V. Firoiu, D. Toesley, and J. Kurose, "Modeling TCP throughput: A simple model and its empirical validation," *Proceedings of ACM SIGCOMM'98*, pp. 303–314, September 1998.
- [2] V. Jacobson, "pathchar," <ftp://ftp.ee.lbl.gov/pathchar/>.
- [3] B. A. Mah, "pchar: A tool for measuring Internet path characteristics," <http://www.ca.sandia.gov/~bmah/Software/pchar>.
- [4] K. Matoba, S. Ata, and M. Murata, "Improving accuracy of bandwidth estimation for Internet links by statistical methods," *13th ITC Specialist Seminar*, pp. 29.1–29.10, September 2000.
- [5] W. R. Stevens, *TCP/IP Illustrated, Volume1: The Protocols*. Addison-Wesley, 1994.
- [6] R. L. Carter and M. E. Crovella, "Dynamic server selection using bandwidth probing in wide-area networks," *TR-96-007*, March 1996.