

# IP ネットワークの利用可能帯域幅推定方式

4 F-6

菊池 慎司 青木 武司 安達 基光 勝山 恒男  
(株)富士通研究所 ネットワークコンピューティング研究部

## 1 はじめに

近年、ネットワークが大規模・複雑化しており、ネットワーク性能の把握が困難になってきている。その一方で、IP ネットワークにおける品質保証、広域負荷分散などにおいて、ネットワークの性能を把握したいという要求が高まってきている。しかし、従来のネットワーク計測技術は、Treno[1] のように正確な計測が可能でもネットワーク資源を浪費したり、Bprobe[2] のようにリソース消費は少ないが精度は著しく低いなどの問題がある。

本稿では、測定精度とリソース消費のトレードオフを解決する IP ネットワーク経路の利用可能帯域幅推定方式 NEPRI を提案する。

## 2 利用可能帯域幅推定方法

Treno などのように、ネットワークの経路の利用可能帯域幅を正確に計測するツールは、実際に多量の計測用パケットを経路に送出し、(転送できたデータ量(bit))/計測期間(sec)を計算して、利用可能帯域幅を求めるものが多かった。しかし、この方式では他のネットワーク利用者の利用可能帯域を計測用パケットが圧迫してしまう。この問題を回避するために、パケット流量が経路の最大転送速度を越えるとパケットのラウンドトリップタイム(RTT)が急激に上昇するという性質[3]に着目し、少量の計測用パケットのRTTの挙動を観察し、RTTの上昇が顕著になるときの計測用パケットの転送速度を調べることにより、利用可能帯域幅を推定する手法を考案した。

### 2.1 経路のルータモデルと帯域幅推定法

IP ネットワークの経路におけるパケットの遅延の挙動を、図1のようなルータモデルで考える[4]。ここで、ルータの最大パケット通過性能を $\mu$  pps、ルータへの計測用パケットの流入量を $P$  pps、計測用パケット以外のパケットの流入量を $I$  ppsとする。また、計測用パケットの数を $n$ 個とし、それぞれのRTTを $RTT_1, RTT_2, \dots, RTT_n$ とする。なお、経路の利用可能帯域幅は $\mu - I$ とする。

ルータへのパケット流入量がルータの最大通過性能を越えない( $P + I \leq \mu$ )とき、パケットはルータのキューに蓄積されないので、計測用パケットのRTTは全てほぼ等しくなる( $RTT_i \approx RTT_{i+1}$ )。

一方、ルータへのパケット流入量がルータの最大通過性能を越える( $P + I > \mu$ )ときには、計測用パケッ

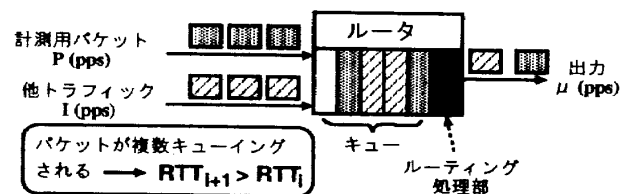


図1: 経路のルータモデル

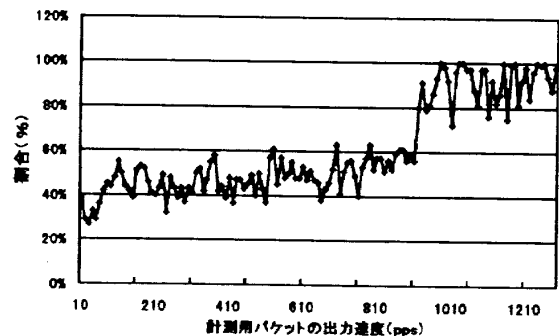


図2:  $RTT_{i+1} > RTT_i$ となる割合の例

トはキューに順次蓄積されていくので、 $i+1$ 番目の計測用パケットの処理時間は $i$ 番目のパケットの処理時間より長くなる( $RTT_{i+1} > RTT_i$ )。このとき、計測用パケットの転送速度 $P$ は経路の利用可能帯域幅の上限を越えており、現在この経路の利用可能帯域幅は $P$ pps以下であると推測できる。

このように、計測用パケットのRTTが $RTT_{i+1} > RTT_i$ となるかどうかを調べることにより、計測用パケットの速度 $P$ が経路の利用可能帯域幅 $\mu - I$ を越えているかを判定することができる。 $P$ を変化させながらこの調査を繰り返せば、 $P \approx \mu - I$ となる速度 $P$ を見つけられることになる。

ある経路に送出速度を変化させながら計測用パケットを送出し、 $RTT_{i+1} > RTT_i$ となる割合を調査した例を図2に示す。この図では、割合は900pps位までは50%前後だが、それ以降になると急激に上昇し、100%に近づく。よって、この経路の利用可能帯域幅は900pps程度であると推測できる。

### 2.2 帯域推定法 NEPRI

上記の考え方を基にした「遅延増加割合に基づくネットワーク計測法(Network Estimation by Per-

centage of RTT's Increasing : NEPRI)」のアルゴリズムを以下に示す。

- (1) 利用可能帯域幅を調べたい経路に対して、少量の RTT 測定用 ICMP echo パケット (大きさは  $E$  bit) を時間間隔  $\delta$  秒で転送する。
- (2) 測定用パケットの RTT が  $RTT_{i+1} > RTT_i$  となるような割合 ( $Q$ ) を調べる。
- (3)  $Q$  が閾値  $R$  より大きい ( $Q > R$ ) ならば、その経路の最大性能が出ている、つまり利用可能帯域幅  $BW_N$  は、 $BW_N \leq E/\delta$  であると判断する。
- (4) TCP/IP ではウィンドウサイズ分のパケットを送出したら、ACK パケットが戻るまで次のパケットは送出できない。よって、TCP/IP での利用可能帯域幅  $BW_T$  は  $BW_T \leq E/2\delta$  となる。

### 3 NEPRI の妥当性の検証

利用可能帯域幅推定方式 NEPRI の妥当性の検証のために、NEPRI と Treno との計測結果と比較した。Treno は標準的な TCP 性能計測ツールであり、UDP パケットを使って TCP の挙動をシミュレートすることにより、経路の TCP 性能を正確に計測する。

#### 3.1 実験内容

広域ネットワーク内の 301 本の経路に対して、以下の処理を行なった。

- (1) 経路の終端の PC から、Treno により経路の TCP 性能 ( $BW_T$  pps) を 20 秒間計測する。
- (2) 転送速度を 10pps から 1300pps まで増加させながら、PC から経路のもう一方の終端ノードに対して計測用 ICMP echo パケット (大きさ 1440byte) を 11 個ずつ出力。これを 10 回繰り返す。
- (3) この増加の途中で、 $RTT_{i+1} > RTT_i$  となる割合が 2 回連続して閾値 (今回は 0.8 採用) を越えた時点での転送速度 ( $BW_N$  pps) を記録する。

#### 3.2 実験結果および考察

Treno の計測結果  $BW_T$  と NEPRI の計測結果  $BW_N$  の関係を図 3 に示す。この図からは、 $BW_T \approx 0.6BW_N$  という関係が読み取れる。 $BW_T$  (経路の TCP 性能) が  $BW_N$  の約半分になっていることは、TCP は ACK が返ってくるまで次のパケットは送出できないという制約 (前節のアルゴリズム (4) 参照) から考えてもほぼ妥当な結果である。また、計測結果全体の 70% において  $BW_T$  と  $0.6 * BW_N$  の差が  $\pm 60$ pps 以下であった。しかし、 $(BW_T - 0.6 * BW_N) \leq -60$ pps となるものも 20% 存在した。このような場合、計測パケットが経路上で廃棄されているケースが多かった。これは、経路のキューのサイズが小さいために、RTT の増加が検出される前にパケット廃棄による TCP の輻輳回避が行なわれていることを表しており、それが

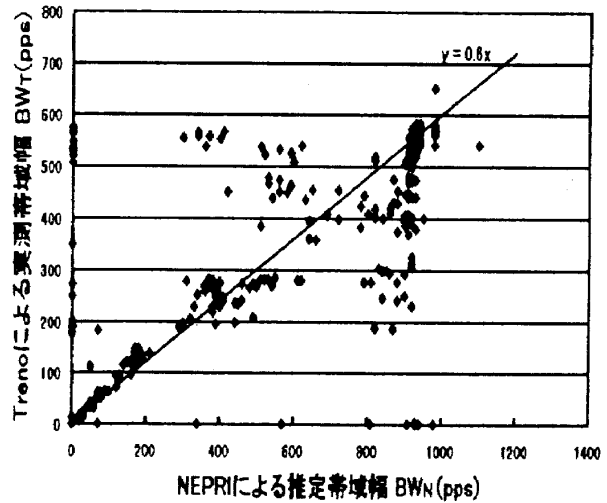


図 3: Treno と NEPRI の計測結果の比較

原因で TCP 性能が NEPRI で測定した性能を下回っていることが考えられる。

また、今回は計測用パケットの転送速度を 10pps から 1300pps まで単調増加させたため、多量の計測パケットを必要としたが、二分法などにより転送速度を操作すれば、計測パケット数を Treno の 1/10 以下に減らすことができる。

### 4 まとめ

本稿では、計測用 ICMP echo パケットの RTT の挙動に着目したネットワーク経路の利用可能帯域幅推定方式 NEPRI を提案した。そして、Treno との出力結果を比較し、70% 以上の割合で誤差  $\pm 60$ pps 以下で計測可能であることを確認した。また、Treno の 1/10 以下の計測パケット数で帯域幅が計測可能なることを示した。

今後は、計測パケットが経路上で廃棄された場合に発生する誤差を修正する手法について検討する必要がある。

### 参考文献

- [1] M. Matthiis and J. Mahdavi, "Diagnosis Internet Congestion with a Transport Layer Performance Tool," Proceedings of INET98, Montreal, June 1998.
- [2] R. Carter and M. Crovella, "Measuring Bottleneck Link Speed in Packet-Switch Networks," Tech. Report Bu-CS-98-006, Computer Science Department, Boston University, Mar. 1998.
- [3] S. B. Moon, J. Kurose, P. Skelly, D. Towsley, "Correlation of Packet Delay and Loss in the Internet", Technical Report UM-CS-1998-011, University of Massachusetts, 1998.
- [4] J.-C. Bolot, "End-to-End Packet Delay and Loss Behavior in the Internet," Proceedings of SIGCOMM '93, pp. 289-298, September 1993.