A New End-to-End Available Bandwidth Measurement Scheme Using Active Probing

Wayman TAN†, Marat ZHANIKEEV†, and Yoshiaki TANAKA†‡
† Global Information and Telecommunication Institute, Waseda University
1-3-10 Nishi-Waseda, Shinjuku-ku, Tokyo, 169-0051 Japan
‡ Advanced Research Institute for Science and Engineering, Waseda University
17 Kikuicho, Shinjuku-ku, Tokyo, 162-0044 Japan
E-mail: waymantan@akane.waseda.jp, maratishe@aoni.waseda.jp, ytanaka@waseda.jp

Abstract Available bandwidth of a network path is defined as the unused portion of the maximum throughput of the path. In this paper, a new available bandwidth measurement scheme using active probing is proposed. The proposed scheme is compared with *pathChirp*, and the result proves that the proposed scheme is more accurate than *pathChirp*.

Keywords end-to-end, probing, active measurement, available bandwidth

1. Introduction

A network path consists of a series of store-and-forward links that connect two hosts at each end of the path. There are two commonly used throughput-related metrics concerning the performance of the network path. They are the *capacity C* and the *available bandwidth A*.

In an individual link, the capacity C is the *transmission rate* or the *maximum throughput* of the link. Considering there are n hops (links) in the path and C_i is the capacity of the i-th link ($1 \leq i \leq n$), the capacity of the path is

$$C = \min (C_1, C_2, \dots, C_n) \tag{1}$$

The capacity of the path or *the end-to-end capacity* is defined as the maximum throughput the path can provide. It is equivalent to the capacity of the narrowest link along a path. Please note that the end-to-end capacity does not change once the path is fixed, provided no routing changes or multi-path forwarding occurs during measurement process.

Available bandwidth of a link is the unused portion of its capacity at a given moment. Suppose there are T bits traffic transmitted at the link during a time interval τ, then the available bandwidth of the link during the interval is:

$$A = C - \frac{T}{\tau} \tag{2}$$

where C is the capacity of the link. Similarly, if A_i denotes the available bandwidth of the i-th link, the end-to-end available bandwidth is:

$$A = \min (A_1, A_2, \dots, A_n) \tag{3}$$

Unlike end-to-end capacity, available bandwidth of a link or a path is defined along with a certain time interval. Depending on changes in traffic along the path, available bandwidth may be different for each interval.

The knowledge of end-to-end available bandwidth is very important and beneficial for various network applications such as multimedia streaming, server selection and optimal routing. However, the scale and complexity of the Internet and the dynamic nature of end-to-end available bandwidth makes its measurement a challenging task.

There are two categories of measurement techniques that estimate end-to-end available bandwidth: passive measurement [1] and active probing [2]-[5]. Passive measurement reads the trace files of existing traffic from routers along a path. While being potentially accurate and efficient, this procedure generally requires administrative accesses to the network devices inside the path. However, such access is rarely given to end users. On the other hand, active probing tools directly send probe traffic into a path from its ends. Such tools are free of privileged access requirement and therefore a pplicable to end users. This paper introduces a novel end-to-end available bandwidth measurement scheme using active probing that helps to solve the problem of very low reliability of end-to-end probing due to interference with cross-traffic.

2. Conventional Schemes and Proposed Scheme

2.1 Conventional Schemes

Various methods for end-to-end available bandwidth measurement using active probing have been proposed to the day. Generally, they can be split into two main approaches underlying the estimation procedure:

The *probe rate model* (PRM) uses the sending rate of the probe traffic (probing rate) to infer the end-to-end available bandwidth. If sending rate is lower than current available bandwidth along a path, probe packets experience no queuing delay. The arrival rate of probe packets at the receiver is roughly equal to the probing rate. By contrast, if the probe traffic is sent at a rate higher than the available bandwidth, probe packets will queue inside the path and thus will be delayed. In this case, the arrival rate is lower than the probing rate. The PRM techniques usually send the probe traffic from an initial low rate and gradually increase it. They search for *the turning point* at which the arrival rate becomes lower than the probing rate. Such turning point is believed to mirror the available bandwidth during the probing period. PRM techniques include *pathChirp* [2], *Pathload* [3], and *PTR* [4].

The *probe gap model* (PGM) compares the time gap of successive probe packets between the sender and the receiver to calculate the available bandwidth. Given two successive probe packets sent with a time gap Δ_{in}, they reach the receiver with a time gap Δ_{out} and are included in the same queue in a single bottleneck of the path. Then Δ_{out} should consist of two time segments: the transmission time of a single probe packet, and the time it takes to transmit all cross traffic that arrives between the two neighbouring probe packets. Therefore, the transmission time of the cross traffic is $\Delta_{out} - \Delta_{in}$. Consider the total transmission time of the probe traffic as well as the cross traffic is Δ_{out}. The rate of the cross traffic C_T is

$$C_T = C \times \frac{\Delta_{out} - \Delta_{in}}{\Delta_{out}} \quad , \tag{4}$$

where C is the end-to-end capacity. The available bandwidth is

$$A = C \times \left(1 - \frac{\Delta_{out} - \Delta_{in}}{\Delta_{out}}\right) \quad . \tag{5}$$

Eq.(5) is applied to a set of congested probe packets to get an average estimate. *IGI* [4] and *Spruce* [5] are examples of tools using PGM.

Both PRM and PGM tools generate a train of probe packets for a single probe so as to ensure robustness to of the bursty cross-traffic. However, they often produce different results. Since the PRM is based on a single turning point of the probe, it can readily respond to changes in cross-traffic. The main drawback of the PRM is its over-reacting, i.e., the estimate is either too high or too low depending on how abruptly cross-traffic changes. On the other hand, the PGM uses Eq.(5) in a set of congested probe packets to smoothen the estimate. It thus provides an averaged estimate but fails to follow changes promptly enough.

Both the PRM and PGM techniques hold certain assumptions: all routers along the path provide FIFO queuing; cross traffic follows a fluid model; its average rate changes slowly and keeps constant during a single probe. For the PGM, a single bottleneck link with the least capacity as well as the least available bandwidth along the path is expected. If these assumptions could not be satisfied, the performance of the techniques may be affected.

2.2 Overview of Proposed Scheme

The PRM and PGM techniques exhibit different properties. In fact, the properties of the two models are complementary. If the properties of both models can be combined in one tool, it offers a room for improvement on the accuracy of estimation. The proposed scheme is taking advantage of just this fact; it adopts both PRM and PGM methods. The proposed scheme generates two estimates with two methods and then uses the estimate of the PGM to filter that of the PRM.

In PRM family *pathChirp* is one of the

prominent members. It uses a *chirp* structure –packets are exponentially spaced within a single probe train to make the probing efficient [2]. The proposed scheme employs similar logic of *pathChirp* and uses PGM method to filter the result of *pathChirp*.

2.3 Structure of Probe Train

The structure of a probe train in the proposed scheme is shown in Figure 1. The probe train consists of N packets. In the first half of the train, the spaces (time gap) between successive packets decrease exponentially; the space of each packet pair is λ times larger than that of the next. λ is called a spread factor. While the sizes of all the packets of a train are the same, the spaces between each packet determine the per-packet probing rate. A probe train with exponentially spaced structure covers a wide range of per-packet probing rate values so it is efficient to discover the available bandwidth in various possible values. In the second half of the train, the spaces between successive packets increase exponentially with the same spread factor λ. The overall structure is symmetric.

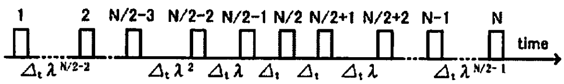

Fig.1 Probing structure of a probe train.

3. Detailed Implementation

3.1 Parameters and Procedure of the Scheme

The proposed scheme uses UDP packets for probing. There are some parameters of the scheme that need to be determined before sending the probe:

(1) The probe packet size P

(2) The length of a probe train N

(3) The initial probing rate E_0

(4) The exponential spread factor λ

P and N are user specified options while E_0 and λ are determined by P and N.

There are three steps in a single probe: initialization, the PRM and PGM calculation, and result filtering.

3.2 Initialization

During the initialization stage, the knowledge of the end-to-end capacity C is required. C is used by the PGM part of calculation. It also helps to calculate the exponential spread factor λ. C can be measured using existing techniques, for example *pathrate* [6]. Note that errors in the measurement of C will affect the accuracy of the available bandwidth estimate. However the end-to-end capacity measurement tools are fairly accurate according to our analysis and experience.

E_0 and λ can be set according to C. Since E_0 is the lowest rate within the whole probe train, it should be much smaller than C. A low E_0 ensures that the probe train covers a wide range of possible results of the available bandwidth. λ should be set so that the highest probing rate of the train is slightly higher than C. A probing rate higher than C guarantees self-congestion, which is necessary for all active probing techniques. Note that in a probe packet pair, the per-packet probing rate E, the time gap between two packets Δ_t and the packet size P have the following relationship:

$$E = \frac{P}{\Delta t} \quad . \tag{6}$$

After all the parameters are initialized, the probe train is transmitted.

3.3 PRM and PGM Calculation

The receiver receives the probe train and performs PRM and PGM calculations. The first half of the probe train is used with the PRM calculation and the second half is used with the PGM calculation.

For the first half of the probe train, a typical signature of one way delay difference can be seen as it is shown in Figure 2. The one way delay difference is the difference of one way delay between two successive probe packets.

The last slope of the one way delay difference will show a gradually increasing trend due to self-induced congestion. The last slope will not be terminating, i.e., it will not return to be zero.

The turning point of the probe train is the first packet of the last slope. As it is introduced before, the probing rate of the turning point is regarded as the available bandwidth during the probing period.

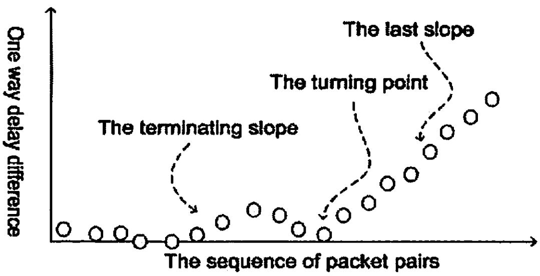

Fig.2 Typical signature of one way delay difference.

Theoretically, the one way delay difference of those probe packets before the turning point should be zero since the probing rate of them is lower than the available bandwidth, however, it is not always the truth. In Fig.2, some small slopes deviating from zero and terminating before the turning point can be seen. They are called *the terminating slopes*. They are caused by the burstiness of the cross traffic which makes the path temporarily congested.

When the bursty cross traffic comes, the instantaneous available bandwidth becomes lower than the probing rate, the probe packets are congested and the one way delay difference increases; after the bursty traffic leaves the available bandwidth returns higher than the probing rate and the one way delay difference drops back to zero. The per-packet probing rate of the probe packets on the inclining side of the terminating slopes reflects the instantaneous available bandwidth when the bursty cross traffic passes. Even though the assumption that the available bandwidth is constant during a single probe period is held by most active probing techniques, it is always better to take into account the burstiness of the cross traffic.

Each packet of the first half of the train has an estimated value E^k of the available bandwidth. When the packets are on the inclining side of the terminating slope, the estimated values are the probing rate of themselves; for all other packets of the first half the estimated values are the probing rate of the packet at the turning point. Taking a weighted average of all estimated values with the time gap at the sender Δ_{in} generates the estimation of the average available bandwidth for the PRM part.

$$D_{PRM} = \frac{\sum_{k=1}^{\frac{N}{2}-1} E^k \Delta_{in}^k}{\sum_{k=1}^{\frac{N}{2}-1} \Delta_{in}^k} \tag{7}$$

The second half of the probe train is used to filter the estimate of the first half later. The PGM method is applied. In the second half of the train, Eq.(5) is used in the set of probe packets from the first one of the second half to the packet with at the same probing rate of the packet at the turning point. Since the probing rate of the packet at the turning point is believed to match the available bandwidth, this set of probe packets should have a probing rate higher than the available bandwidth and thus induce congestion. Each pair of the successive probe packets can generates a per-packet estimate by using Eq.(5). Taking the average of the per-packet estimate yields the estimate of the second half.

$$D_{PGM} = \frac{\sum_{k=\frac{N}{2}}^{N-t-1} C \times \left(1 - \frac{\Delta_{out}^k - \Delta_{in}^k}{\Delta_{out}^k}\right)}{\frac{N}{2} - t}, \tag{8}$$

where C is the end-to-end capacity and t is the sequence number of the packet at the turning point.

3.4 Result Filtering

Before combining two estimates, the D_{PRM} and D_{PGM} need to be validated: since the available bandwidth is never higher than the end-to-end capacity C, D_{PRM} and D_{PGM} will be set as C if they are higher than C.

$$A_{PRM} = \begin{cases} D_{PRM}, & \text{if } D_{PRM} \le C \\ C, & \text{if } D_{PRM} > C \end{cases} \tag{9}$$

$$A_{PGM} = \begin{cases} D_{PGM}, & \text{if } D_{PGM} \le C \\ C, & \text{if } D_{PGM} > C \end{cases} \tag{10}$$

Finally the estimate of the first half is filtered by that of the second half with a coefficient α.

$$A = \alpha * A_{PRM} + (1-\alpha) * A_{PGM}, \tag{11}$$

where $0 \le \alpha \le 1$. The coefficient α is very important as it determines how much either part of the probe train affects the final estimate. The configuration of α will be analyzed in the next section.

Since the proposed scheme does not adapt the probing rate to the cross traffic, the probing rate is fixed after the initialization and it is possible to get an estimate of available bandwidth in a short time. However, the estimate result of a single probe may not be precise due to complexity of the Internet and the highly bursty nature of the cross traffic. The measurement should be smoothened by a number of probes. Usually, one should keep probing for a reasonable period of time and then calculate the average result of all probes. The result is the estimate of the average available bandwidth during the period.

4. Simulation

4.1 Simulation Conditions

The simulation is performed with OPNET Modeler. The background cross traffic includes popular Internet traffic such as HTTP, FTP, email and voice traffic. A three-hop network path is constructed and the middle hop is the bottleneck link. Some available bandwidth measurement tools perform well only in paths with high capacity, e.g., more than 100Mbps. In our simulation, the capacity of the bottleneck link is 10Mbps. The performance of the proposed scheme is verified in a low capacity environment. The proposed scheme is compared with *pathChirp*.

4.2 Simulation Results

The first test scenario is as follows: The size of probe packets is fixed to 1000 bytes. The length of each probe train is 30 packets. The time gap between the first successive packets is 8ms. The exponential spread factor is 1.3. The coefficient α is 0.5, which stands for a perfect middle-ground between PGM and PRM. The path is under medium utilization with the average utilization rate of 25%.

The result is shown in Figure 3. As can be confirmed from the figure, the proposed scheme generally outperforms *pathChirp* and provides an estimate close to the actual value.

In the next scenario, the parameters of the scheme remain unchanged and the utilization of the path is set to a higher average value of around 40%. The result is shown in Figure 4. Despite a few inaccuracies in form of spikes, the overall result is satisfactory.

When the utilization of the path is light, e.g., the average is lower than 15% and the capacity of the path is small (10Mbps in our test), both *pathChirp* and the proposed scheme cannot sense the change of the cross traffic. The result is shown in Figure 5. Under such condition, the coefficient α should be tuned well to bring the estimate closer to the average available bandwidth.

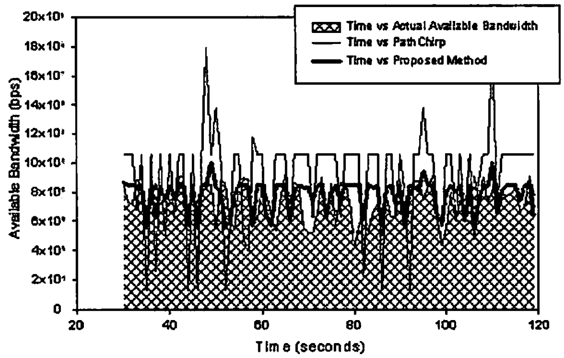

Fig.3 Comparison between the proposed scheme and *pathChirp* under medium utilization.

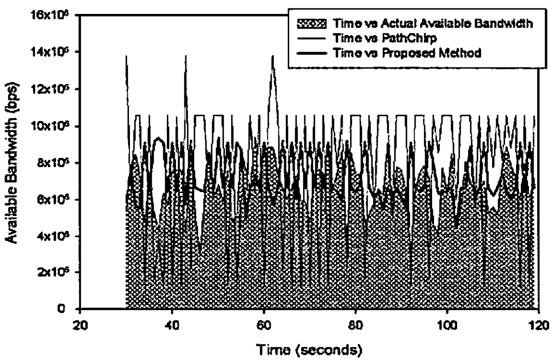

Fig.4 Comparison between the proposed scheme and *pathChirp* under heavy utilization.

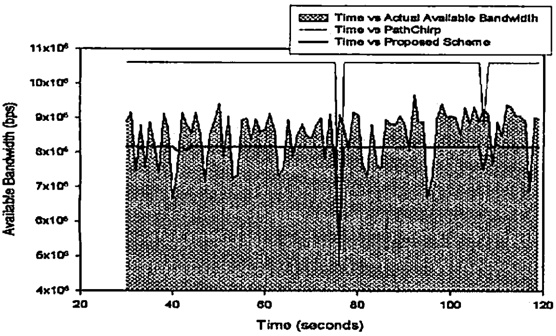

Fig.5 Comparison between the proposed scheme and *pathChirp* under light utilization.

4.3 Analysis of the Results

In the proposed scheme there are three user-specified parameters: the probe packet size, the probe train length and the coefficient α. For the first two parameters they are well studied in other tools [2]-[5]. They are optimized in our tests and not discussed within this paper. The coefficient α is unique to our proposed scheme and thus needs a detailed analysis.

To analyze α, it is configured with various values under different utilization. The results are shown in Figure 6. The error rate is the difference between the estimate and the actual available bandwidth value, in proportion to the total capacity. Note that during the measurement there are a number of probes and thus a number of samples of the error rates. The error rate of a single measurement is the average of all collected samples. In Fig.6, both the error rate and the coefficient value are displayed as ratio values.

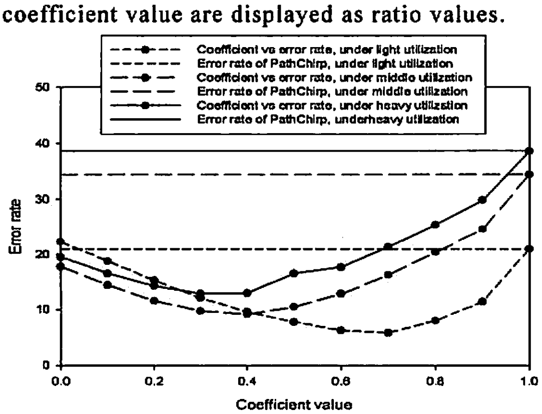

Fig.6 Error rate vs. coefficient value under different utilization environment.

From the numerical results, the coefficient value 0.4 generates the lowest error rate under the middle and heavy utilization. It is because the cross traffic may be highly fluctuating under such circumstances, the filter of the second half of the train needs to take more effect to balance the over-reacting of the first half. When the utilization is light, the coefficient value 0.7 offers the best performance. In this case, the estimate of the probe does not change too much, and there is not much over-reaction that needs to be filtered.

5. Conclusion

In this paper, an active probing scheme of end-to-end available bandwidth measurement is introduced. The scheme adopts both *probe rate model* and *probe gap model*. It takes advantage of the complementary properties of two models to improve the accuracy of the estimation. Simulation shows that the proposed scheme generally outperforms existing tools such as *pathChirp* and can provide an estimate close to the actual available bandwidth. The effect of different filtering weight in different circumstances is also studied.

References

[1] S. Seshan, M. Stemm, and R. H. Katz, "SPAND: Shared passive network performance discovery," Proc. 1st Usenix Symp. Internet Technologies Systems, Monterey, CA, pp.135-146, Dec. 1997.

[2] V. J. Ribeiro, R. H. Riedi, R. G. Baraniuk, J. Navratil, and L. Cottrell, "pathChirp: Efficient available bandwidth estimation for network paths," Proc. Passive and Active Measurement Workshop, no. SLAC-PUB-9732, March 2003.

[3] M. Jain and C. Dovrolis, "End-to-End available bandwidth: Measurement methodology, dynamics, and relation with TCP throughput," IEEE/ACM Trans. Networking, vol.11, no.4, pp.537-549, Aug. 2003.

[4] N. Hu and P. Steenkiste, "Evaluation and characterization of available bandwidth probing techniques," IEEE J. Select. Areas Commun., vol.21, no.6, pp.879-894, Aug. 2003.

[5] J. Strauss, D. Katabi, and F. Kaashoek, "A measurement study of available bandwidth estimation tools," Proc. 3rd ACM SIGCOMM Conf. Internet Measurement (IMC), pp.39-44, Oct. 2003.

[6] C. Dovrolis, P. Ramanathanm, and D. Moore, "What do packet dispersion techniques measure?" Proc. IEEE INFOCOM, pp.905-914, April 2001.