Minimal Backlogging Method for Estimation of Available Bandwidth

Seung Yeob Nam, Sunggon Kim, and Dan Keun Sung
Department of Electrical Engineering and Computer Science, KAIST,
Daejeon, 305-701, KOREA

***Abstract*—This paper is concerned with estimating the available bandwidth of a network path. We first develop a theory to estimate the available bandwidth of a queueing system. In order to estimate the available bandwidth, we propose a probing method called a *minimal backlogging method*, and a statistic based on the service rate of minimally backlogging probing traffic. We show that the available bandwidth of a queueing system can be estimated by the statistic if probing packets are sent to the queueing system by the minimal backlogging method. For a network path consisting of multiple hops, we extend the approach for a single server by introducing a simplified path model. Since the proposed mechanism can estimate the available bandwidth quickly and track it adaptively, a reasonable range of available bandwidth for a short time interval can be obtained using the mean and variance of the estimated available bandwidth. The performance of the proposed available bandwidth estimation mechanism is evaluated by simulation in a multiple hop network topology.**

***Index Terms*— Available bandwidth, probing, measurement, available bandwidth estimation, minimal backlogging**

I. Introduction

The reliable estimation of available bandwidth for a path is very important for high utilization of network resources as well as QoS guarantee for real-time flows. If the *available bandwidth* (AB) for a specific network path is known to a traffic source node, the source node can avoid paths in congestion in advance [1] or the information about AB can be used for capacity provisioning, network troubleshooting, and traffic engineering (TE) in IP or MPLS networks [2], [3]. Thus, monitoring of AB is very important to exploit network resource efficiently.

For a network path $\mathcal{P}$ between a node pair consisting of H serially connected links, AB C_a for the path in a given time interval is usually defined as

$$C_a = \min_{1 \le i \le H} C_i(1 - u_i),$$

where C_i and u_i be the link rate and the utilization of the i-th link in the given time interval, respectively. The link with the least unused bandwidth of C_a is referred to as *tight link* and the link with the minimum link rate is referred to as *bottleneck link*.

Several methods have been proposed to estimate the AB for a path. The first attempt to measure available bandwidth was C-probe [4]. The C-probe is to estimate the available bandwidth from the dispersion of trains of eight packets. They assumed that the dispersion of long packet trains is inversely proportional to the available bandwidth. However, Dovrolis et al.[5] showed that this is not true. Melander et al.[6] proposed a TOPP probing method which is an extension to the packet pair probing technique. They estimated the available bandwidth and the capacity of the link with the smallest link rate from the relation between the input and output rates of different packet pairs. TOPP is computationally intensive to implement. Jain and Dovrolis[7] proposed a tool called *pathload*. Pathload is to estimate the range of available bandwidth iteratively, not the value of available bandwidth and has a rather long convergence time. Hu et al.[8] proposed the initial gap increasing (IGI) method and the packet transmission rate (PTR) method.

In order to estimate the available bandwidth quickly and reliably by overcoming the drawbacks of existing schemes, we propose a new available bandwidth estimation mechanism. Basically, we assume that every node serves packets in a First-Come-First-Served (FCFS) manner. The proposed available bandwidth estimation mechanism is based on two key components. The first one is a minimal backlogging concept and the second one is a simplified path model. We estimate available bandwidth by monitoring probing packets sent according to the minimal backlogging method, which is explained in detail in the next section. We first develop an available bandwidth estimation theory for a single server and extend the approach to a network path. We introduce a simplified path model in order to simplify the estimation problem for a multiple hop path. We focus on tracking the dynamically varying AB for a relatively short time period to finally obtain a reasonable range of the AB.

The rest of this paper is organized as follows. In Section II, we introduce the minimal backlogging method. We propose a statistic based on the service rate of minimally backlogging probing traffic to estimate available bandwidth of a queueing system. In Section III, we introduce a simplified path model for multiple hop paths. We extend the approach for a single server to multiple hop paths by exploiting the simplified path model. In Section IV, The validity of the proposed statistic is verified by simulation for a single server and the performance of the proposed available bandwidth estimation mechanism is evaluated by simulation in a multiple hop network topology. Finally, conclusions are presented in Section V.

II. Estimation of Available Bandwidth for a Single Server

Before taking into account the AB estimation problem for multiple hop routes, we introduce some concepts and theory for a single server. We consider a queueing system with an FCFS service policy. The service rate is C, and the arrival rate of packets except probing packets is λ. Suppose that the service time of a packet is given by the packet size divided by the service rate C of the system. Let L' denote the average length of the packets except probing packets. Then, for the queueing system, available bandwidth C_a is defined as

$$C_a = C(1-\rho),$$

where $\rho = \lambda L'/C$ is the traffic load to the system. If the parameters C, λ and L' representing a queueing system are unknown, this system is said to be *unidentified* in this paper. We assume that $\rho < 1$ for the stability of the system. For the estimation of the available bandwidth of an unidentified queueing system, we propose a probing scheme as follows:

Definition 1: Suppose that we send probing packets into a queueing system so that there exists one and only one probing packet in the system. This probing method is called a *minimal backlogging method*.

If we send a probing packet into a queueing system just at the departure time of the previous probing packet, then there exists one and only one probing packet in the system. We now propose a statistic in order to estimate the available bandwidth of a queueing system.

Definition 2: The available service $\hat{Y}_{[s,t]}$ for a queueing system is the amount of probing packets served in interval $[s,t]$ when probing packets are sent to the queueing system according to the minimal backlogging method.

Before we investigate the characteristics of the available service analytically, we briefly explain why the term of *available service* is used for $\hat{Y}_{[s,t]}$. In case that the minimal backlogging method is not used, an *idle period*, i.e. a time interval when the server is not busy, can exist if the load of non-probing packets is less than 1. In case that the probing packets are sent to the queueing system according to the minimal backlogging method, there always exists at least one probing packet in the queueing system, and thus, there is no idle period during the probing time. If there is no non-probing packet in the system, probing packets will be served continuously until a new non-probing packet arrives. Thus, we can know that the amount of probing packets served in a given time interval will be at least the maximum amount of service that the server can additionally support while serving all arriving non-probing packets according to an FCFS policy.

Theorem 1: Let $\hat{Y}_t$ be the available service for a $G/G/1$ queueing system. The size of each probing packet is fixed to a constant of L. Then, for $0 < q < \infty$,

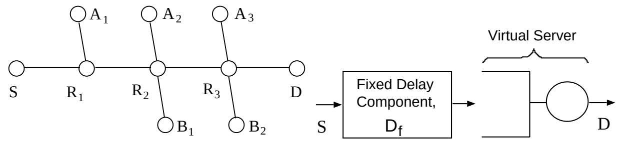
$$\lim_{t\to\infty} E\left[\left|\frac{\hat{Y}_t}{t} - C(1-\rho)\right|^q\right] = 0.$$

Proof: The proof is given in [9]. ■

Thus, the service rate of minimally backlogging probing traffic, $\hat{Y}_t/t$, can be used as an estimator of the AB for a $G/G/1$ queueing system and this statistic is used to estimate the available bandwidth for a multiple hop path in the next section.

III. Estimation of available bandwidth for a multiple hop path

The AB estimation mechanism for a single server developed in the previous section can not be directly applied to AB estimation for a network path between a specific node pair because a network path usually consists of multiple hops. Before considering AB estimation for a multiple hop path, we introduce a simplified path model. We consider a single tight link along a multiple hop path because multiple tight links are not likely to occur frequently in real networks due to variation of the AB at each link. However, the proposed mechanism can be applied to multiple tight link environments.



(a) Network topology (b) Simplified path model

Fig. 1. An end-to-end network path model

If we assume that the delay variation at other links except the tight link is negligible, then the summation of delays at those links except the tight link is constant and we denote the summation as D_f. Then, we can obtain a simplified path model consisting of a fixed delay component (D_f) and a virtual server $\mathcal{S}$ for the tight link as shown in Fig. 1. Suppose that a probing packet p arrives at the path at time a_p and departs from the path at time d_p. Then, the packet p arrives at the virtual server $\mathcal{S}$ at time $a_p^s = a_p + D_f$.

If N probing packets are sent to the virtual server by the minimal backlogging method, then, by Theorem 1, AB for the virtual server in the interval $[a_1^s, d_N]$ can be estimated by $NL/(d_N - a_1^s)$, where $a_1^s = a_1 + D_f$ and D_f is the fixed delay for the current probing period. However, it is not possible to send probing packets according to the minimal backlogging method due to the fixed delay component D_f. Instead, we attempt to emulate the minimal backlogging method by sending bursts of probing packets at an adaptive rate. We send a fleet of N probing packets at a time and the time interval of $[a_1, d_N]$ is called a *probing period*. Then, AB for the path is estimated as follows.

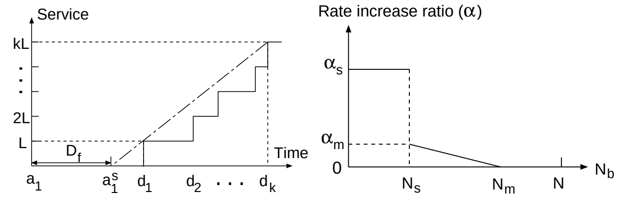
Since the inter-probing-packet spacing is fixed during a probing period by the corresponding probing rate in real applications, several busy periods of probing packets may exist during a probing period. Consider the i-th busy period containing k continuously backlogged probing packets. Probing packets arriving during the busy period are indexed from 1 to k. Fig. 2(a) illustrates a sample service curve for the busy period showing the amount of probing packets served for $[a_1, t]$. The *Measured Probing Rate* (*MPR*) for the i-th busy period is defined as:

$$MPR(i) = \frac{kL}{d_k - a_1^s} = \frac{kL}{(d_k - a_1) - D_f}.$$

where D_f is estimated in the longest busy period of the previous probing period by

$$\tilde{D}_f = (d_1 - a_1) - \frac{d_k - d_1}{k-1}.$$

while satisfying $a_p^s \le d_p$ for any probing packet p. The *MPR* for the longest busy period during a probing period is used to reliably estimate AB.



(a) Example of a service curve

(b) Rate increase ratio curve

Fig. 2. Probing rate adaptation scheme

We now consider a probing rate adaptation scheme. Let $N_b(i)$ be the number of probing packets belonging to the longest busy period in a probing period. If N probing packets are sent to the virtual server according to the minimal backlogging method, then there will be only a single busy period containing N probing packets during a probing period, and thus, $N_b = N$. However, since the inter-packet-spacing is fixed during a probing period, even if probing packets are sent at the rate which is the average rate of minimally backlogging probing packets, N_b may be less than N. We attempt to maintain N_b within a reasonable range by an adaptive probing scheme. A small value of N_b is due to a lower probing rate than for minimal backlogging and a large value of N_b is due to a higher rate. If N_b is in the reasonable range, we may assume that the minimal backlogging occurs. Thus, *MPR* is a reliable estimate of the AB. Let $(N_s, N_m]$ be the reasonable range of N_b. Fig. 2(b) shows the proposed probing rate adaptation scheme, which is explained as follows:

Case 1: If $N_b > N_m$, then *MPR* is considered to be larger than the AB due to a higher probing rate than for minimal backlogging, and the next input rate is set to *MPR*. The AB is estimated by *MPR* since *MPR* quickly approaches to the AB.

We can explain the reason for the use of *MPR* as the next probing rate by the following example. For an FCFS server with a link rate of C and an AB of C_a, if the probing packets arrive at a rate of r $(\geq C_a)$, they are served at a rate of

$$m(r) = \frac{r}{C - C_a + r}\, C.$$

In case that we adjust the $(n+1)$-th probing rate by $r_{n+1} = m(r_n)$, we can easily show that if $r_1 \geq C_a$ and $C_a > 0$, then $\lim_{n\to\infty} r_n = C_a$, that is, *MPR* converges to the AB.

Case 2: If $N_b \leq N_s$, *MPR* for this short busy period may be inaccurate because the minimal backlogging condition is not satisfied. Thus, the current AB is estimated by the AB at the last probing period for $N_b > N_s$. If $N_b \leq N_s$ consecutively i times since the last probing period with $N_b > N_s$, then the next input rate is set to AB $\cdot\, (1 + \alpha_s)^i$. α_s determines the tracking speed of the proposed algorithm when the probing rate is lower than the AB. When the current probing rate is lower than the AB, if α_s is large, then *MPR* quickly approaches to the AB, but large values of α_s may cause temporary ripples.

Case 3: If $N_s < N_b \leq N_m$, then *MPR* is a reliable estimate of the AB. However, it is necessary to maintain the probing rate slightly higher than AB in order to obtain a reliable value of *MPR*. Thus, the next input rate is increased to $MPR \cdot (1 + \alpha(N_b))$, where $\alpha(N_b) = \alpha_m(N_m - N_b)/(N_m - N_s)$, and α_m is the maximum rate increase ratio in the medium busy period range. If the value of *MPR* is close to that of AB, then the next probing rate is higher than AB by a ratio of $\alpha(N_b)$. For a given value of N_b, if α_m or N_m increases, $\alpha(N_b)$ also increases.

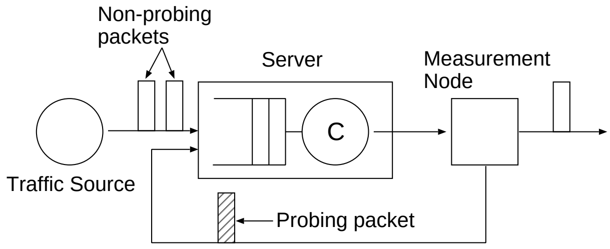


Fig. 3. A measurement setup for estimation of the available bandwidth of a queueing system

As explained above, the proposed probing scheme attempts to send probing packets at a slightly higher rate than the AB. Thus, the load offered to the tight link may slightly exceed one during a probing period. In order to prevent degradation of the throughput of data traffic at the tight link due to overload, consecutive probing periods are separated by at least one probing period length of $d_N - a_1$. Then, the average load offered by the probing traffic is approximately equal to or lower than half of the AB in a longer time interval than the duration of one probing period, and thus, the tight link is not overloaded in a long time scale.

IV. Numerical Results

First, we investigate the validity of the theory developed in Section II for a single server. Next, we evaluate the performance of the AB estimation mechanism developed in Section III for a multiple hop path.

A. Single Server Case

In Section II, we showed that it is possible to estimate the AB of a queueing system by measuring probing packets if the queueing system is probed by the minimal backlogging method for an infinite time duration. However, it is not possible to probe a queueing system for an infinite time duration. Thus, we evaluate the accuracy of the proposed statistic numerically in case of probing a queueing system during a finite time duration.

Fig. 3 shows a simulation setup for estimation of the available bandwidth of a queueing system. The measurement node directly connected to the queueing system sends probing packets to the queueing system by the minimal backlogging method, i.e., the node sends a new probing packet upon arrival of the previous probing packet and calculates the value of the statistic. The measurement node bypasses every non-probing packet.

The traffic source generates two types of non-probing packet traffic patterns: Poisson and self-similar traffic. The traffic patterns of today's IP networks have been known to exhibit self-similarity and long-range dependence [11]. Neither of them can be modeled using conventional Markovian models. Thus, we use a multi-fractal model [12] to generate self-similar traffic. The Hurst parameter is 0.8. The sizes of both probing and

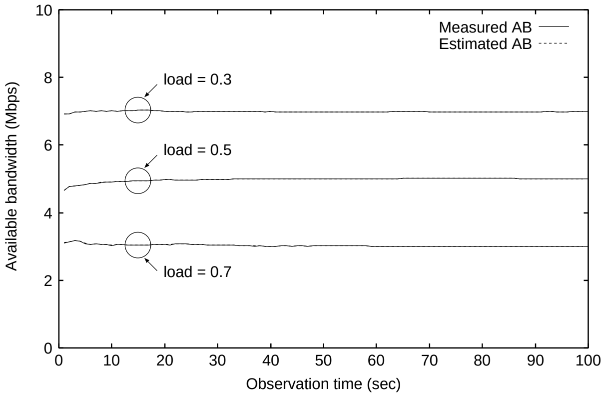


Fig. 4. Comparison of the estimated AB and the measured AB under a Poisson traffic load

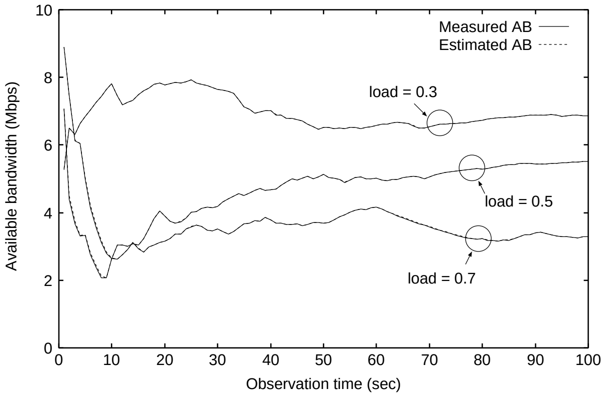


Fig. 5. Comparison of the estimated AB and the measured AB under a self-similar traffic load

non-probing packets are fixed to 4000 bits. The service rate (C) of the queueing system is 10 Mbps.

Fig. 4 compares the estimated AB with the measured AB under a Poisson traffic load. We can observe that the estimation results agree very well with the measured AB for all traffic loads. In addition, the estimation results are accurate even when the observation time is short.

Fig. 5 compares the estimated AB with the measured AB under a self-similar traffic load. The sigma/mean ratio of self-similar traffic is 0.68, 0.54, and 0.39 for the loads (ρ) of 0.3, 0.5, and 0.7, respectively. First, we can observe that the traffic is even burstier than the case of Poisson traffic. Thus, it takes longer time for the average rate of the traffic to converge to $C\rho$ for all traffic loads compared with the case of Poisson traffic. However, the estimated AB agrees very well with the measured AB regardless of the observation time duration for various input loads as shown in Fig. 5. Due to the burstiness and the long-range dependence of the traffic, it takes longer time for the estimated AB to converge to $C(1-\rho)$ compared with the case of Poisson traffic.

B. Multiple Hop Path Case

We compare the performance of the AB estimation mechanism for a multiple hop path developed in Section III with that of pathload [7] through OPNET simulation. A multiple hop topology is illustrated in Fig. 1(a). Each node is modeled as an output queued router with a FIFO queue. We estimate AB for the path $S-R_1-R_2-R_3-D$. Every link except R_2-R_3 has a link rate of 20 Mbps and a propagation delay of 5 ms. Link R_2-R_3 with a link rate of 10 Mbps is the bottleneck link. The sizes of both probing packets and data packets are 4000 bits. For the proposed mechanism, the number of probing packets sent in one probing period (N) is 100. The values of the rate adaptation related parameters are set to $N_m = 0.95 \times N = 95, N_s = 0.30 \times N = 30, \alpha_m = 0.10,$ and $\alpha_s = 1.0$. For the pathload [7], the user-specified resolution of AB ω is set to 0.2 Mbps and the grey-region resolution χ is 0.3 Mbps.

Two types of traffic patterns are used for non-probing packet sequence: constant bit rate (CBR) and self-similar traffic generated using a multi-fractal model [12]. The Hurst parameter is 0.8 and the sigma/mean ratio of a flow is approximately 0.5. The mean rate of each flow is 4 Mbps except a flow which is sent from A_2 to B_2 and has a rate of 2 Mbps. During a simulation time of 200 seconds, 4 flows with a lifetime of 70 seconds are sent on route $A_1-R_1-R_2-B_1$ sequentially at an interval of 10 seconds from time 0. 4 flows with a lifetime of 70 seconds are sent on route A_3-R_3-D sequentially at an interval of 10 seconds from time 100. Thus, link R_1-R_2 is a tight link in the interval $[20, 80]$. Link R_2-R_3 is a tight link in the intervals of $[0, 30]$, $[70, 130]$, and $[170, 200]$. Link R_3-D is a tight link in the interval $[120, 180]$. Thus, two tight links exist in the intervals of $[20, 30]$, $[70, 80]$, $[120, 130]$, and $[170, 180]$.

Fig. 6 compares the AB of the proposed mechanism with that of the pathload under a CBR traffic load. The pathload iteratively estimates the range $[R^{min}, R^{max}]$ of AB. The trace of $(R^{min} + R^{max})/2$ is plotted in Fig. 6 and the range of $[R^{min}, R^{max}]$ is also shown at the instant of termination. The pathload is restarted just after it terminates. We can observe that it takes about 8 seconds for the pathload to terminate. The pathload sometimes yields a significant error in the estimation of AB, especially at time 153.8 as shown in Fig. 6. However, the proposed mechanism closely tracks the AB even if AB changes abruptly, there exist two tight links or the tight link is different from the bottleneck link. The error observed in the intervals of $[30, 70]$ and $[130, 170]$ is due to the fact that the proposed probing scheme tries to maintain the probing rate slightly higher than the AB to obtain a reliable value of *MPR*. If N_m or α_m is decreased, this error can also be decreased, but more ripples may occur due to unreliable values of *MPR*.

The reason why the proposed scheme can estimate the AB faster than the pathload can be explained as follows. The pathload changes the probing rate using a binary search to find the AB, while our scheme tries to find and track the AB continuously by adapting the probing rate based on the previous estimation value of the AB and the observed value of N_b.

Fig. 7 compares the mean (μ) and the standard deviation (σ) of the AB estimated by the proposed mechanism with those of the measured AB under a self-similar traffic load. The range of $[\mu-\sigma, \mu+\sigma]$ is plotted based on the measurement at an interval of 10 seconds. We can observe that the mean of the measured AB lies within σ from μ of the estimated AB for every estimation time.

Fig. 8 compares the AB of the proposed mechanism with that of the pathload under the same traffic trace as Fig. 7. The curve

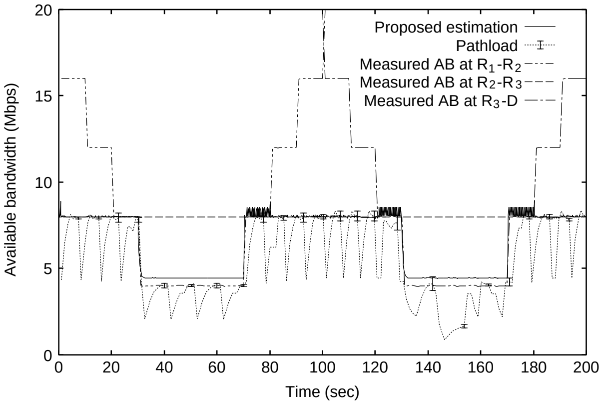


Fig. 6. Comparison of available bandwidth estimated by the proposed mechanism and by the pathload under a CBR traffic load

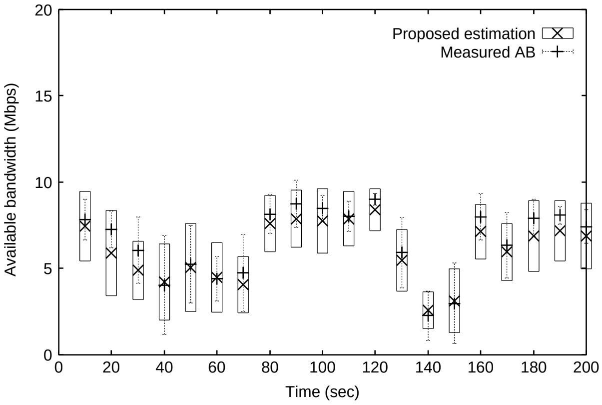


Fig. 7. Available bandwidth estimated by the proposed mechanism under a self-similar traffic load

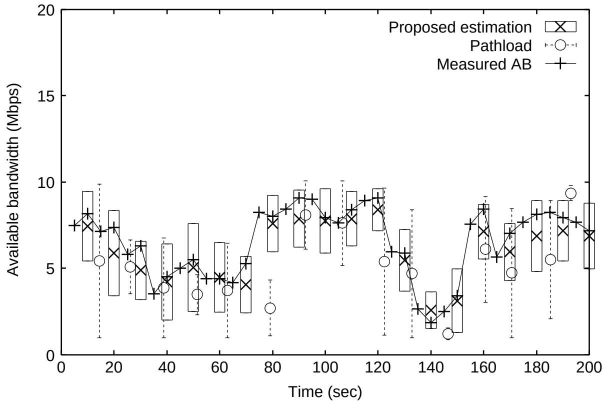


Fig. 8. Comparison of available bandwidth estimated by the proposed mechanism and by the pathload under a self-similar traffic load

for the measured AB is the value of measured available bandwidth averaged for every 5 seconds. The curve for the proposed estimation mechanism is the same as that in Fig. 7. We can observe some problems of pathload from the curve for the pathload. First, the convergence time increases for bursty traffic. We can observe that the average convergence time is longer than 10 seconds in this case. Second, pathload frequently fail to give a converged range of AB when the traffic load changes dynamically. Third, even the estimation range of AB sometimes deviates from the average value of the measured available bandwidth, especially at time 51, 79, 147, and 193 seconds. Thus, it is difficult to obtain a reasonable range of available bandwidth for a short period of 10 seconds by the pathload because of a long convergence time. On the other hand, the proposed mechanism gives a reasonable range of the available bandwidth even when the traffic load significantly changes.

V. Conclusions

A new mechanism for estimation of available bandwidth is proposed in this paper. The proposed mechanism is based on two key components: the minimal backlogging concept and a simplified path model. We first developed an available bandwidth estimation theory for a single server. We proposed a statistic based on the service rate of minimally backlogging probing traffic and showed that the statistic can be used to estimate the available bandwidth of a queueing system if probing packets are sent to the queueing system according to the minimal backlogging method. For a network path consisting of multiple hops, we extended the approach for a single server by introducing a simplified path model and proposing a probing rate adaptation scheme. Simulation results show that the estimation results based on the proposed statistic agree well with the measured available bandwidth in case of a single server even for a finite probing time. In a multiple hop topology, it is observed that the proposed mechanism tracks the available bandwidth rather accurately even when the available bandwidth changes abruptly. Thus, the proposed mechanism can be used to obtain a reasonable range of dynamic available bandwidth for a network path in a short time interval.

References

[1] N. Benameur, S. B. Fredj, S. Oueslati-Boulahia, J. W. Roberts, "Quality of service and flow level admission control in the Internet," *Computer Networks*, vol. 40, pp. 57-71, 2002.

[2] A. Bosco, R. Mameli, E. Manconi, F. Ubaldi, "Edge distributed admission control," *IEEE Communications Letters*, vol. 7, no. 2, pp. 88-90, Feb. 2003.

[3] J. Boyle *et al*, "Applicability statement for traffic engineering with MPLS," IETF RFC 3346, 2002.

[4] R. L. Carter and M. E. Crovella, "Measuring bottleneck link speed in packet-switched networks," *Performance Evaluation*, vol. 27-28, pp. 297-318, 1996.

[5] C. Dovrolis, P. Ramanathan, and D. Moore, "What do packet dispersion techniques measure?," *Proc. IEEE Infocom'01*, pp. 905-914, Apr. 2001.

[6] B. Melander, M. Bjorkman, and P. Gunningberg, "A New End-to-End Probing and Analysis Method for Estimating Bandwidth Bottlenecks," *Proc. IEEE Globecom'00*, pp. 415-421, Nov. 2000.

[7] M. Jain and C. Dovrolis, "Pathload: a measurement tool for end-to-end available bandwidth," *Proc. Passive and Active Measurements (PAM) Workshop*, Mar. 2002.

[8] 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.

[9] Seung Yeob Nam, Sunggon Kim, and Dan Keun Sung, "Estimation of Available Bandwidth for an Unidentified Queueing System," submitted.

[10] A. G. Pakes, "Some Conditions for Ergodicity and Recurrence of Markov Chains," *Oper. Res.*, vol. 17, pp. 1058-1061, 1969.

[11] V. Paxson and S. Floyd, "Wide-area traffic: the failure of Poisson modeling," *IEEE/ACM Trans. Networking*, vol. 3, pp. 226-244, 1995.

[12] R. H. Riedi, M. S. Course, V. J. Ribeiro, and R. G. Baranuik, "A Multifractal Wavelet Model with Application to Network Traffic," *IEEE Trans. Inform. Theory*, vol. 45, pp. 992-1018, 1999.