Measurement-Based Admission Control at Edge Routers

Seung Yeob Nam, *Member, IEEE,* Sunggon Kim, and Dan Keun Sung, *Senior Member, IEEE*

***Abstract*— It is very important to allocate and manage resources for multimedia traffic flows with real-time performance requirements in order to guarantee quality-of-service (QoS). In this paper, we develop a scalable architecture and an algorithm for admission control of real-time flows. Since individual management of each traffic flow on each transit router can cause a fundamental scalability problem in both data and control planes, we consider that each flow is classified at the ingress router and data traffic is aggregated according to the class inside the core network as in a DiffServ framework. In our approach, admission decision is made for each flow at the edge (ingress) routers, but it is scalable because per-flow states are not maintained and the admission algorithm is simple. In the proposed admission control scheme, an admissible bandwidth, which is defined as the maximum rate of a flow that can be accommodated additionally while satisfying the delay performance requirements for both existing and new flows, is calculated based on the available bandwidth measured by edge routers. The admissible bandwidth is a threshold for admission control, and thus, it is very important to accurately estimate the admissible bandwidth. The performance of the proposed algorithm is evaluated by taking a set of simulation experiments using bursty traffic flows.**

Keywords: **Admission control; measurement; available bandwidth; admissible bandwidth; Quality of Service (QoS)**

I. Introduction

Although the capacity of core networks has increased tremendously due to advanced optical transmission equipments and high-speed routers/ethernet switches, quality-of-service (QoS) is not well guaranteed in the current IP networks. Integrated Services (IntServ) [1] is one of the approaches proposed to address this problem. While IntServ is capable of providing QoS within a domain, it is not scalable since every router is required to manage per-flow information. On the other hand, DiffServ [2] scales well since core routers treat not per-flow information, but only class-level traffic aggregate. There are two types of approaches for supporting QoS under DiffServ framework: reactive and preventive approaches. In the reactive approaches, QoS is supported by adaptively changing the source traffic load based on the network status [3, 4]. Resource is usually not reserved, but this reactive approach may not be directly applicable to the applications which do not change the traffic rate adaptively. Admission control is a typical preventive approach. The traffic rate does not need to be adjusted adaptively in this case and we focus on this preventive approach in this paper. There have been many efforts to provide statistical or soft QoS by incorporating a new admission control scheme in the framework of DiffServ [5–16]. However, the statistical QoS is still not well guaranteed [17].

There are two important goals of admission control algorithms. The first one is to guarantee the contracted QoS for real-time flows, and the other one is to achieve high network utilization. We propose a new admission control scheme to achieve these goals. We consider delay as a QoS target because real-time flows are more sensitive to delay than loss. If admission-controlled traffic is treated in the same way as best-effort traffic that is not subject to admission control at each router, then the required QoS may not be guaranteed due to the uncontrolled traffic rate of the best-effort traffic. Thus, we assume that there are two classes: a high priority class that is subject to admission control and a low priority class that is not subject to admission control. Core routers are assumed to use the strict priority policy for scheduling of different classes and the per-class scheduling can be implemented in the DiffServ framework.

In our proposed admission control scheme, each ingress router manages admissible bandwidth, which is a threshold for admission control, for each relevant egress router. Admission decision is made for each flow by comparing the peak rate of the flow with the admissible bandwidth. We derive a simple equation for admissible bandwidth considering the delay QoS based on the available bandwidth, which is estimated by the egress router through monitoring probing packets. The contribution of our approach can be summarized as follows.

First, our scheme statistically guarantees the delay bound of admission-controlled traffic for moderate delay bound values while maintaining high resource utilization. Conservative resource (e.g. bandwidth) allocation may guarantee delay bound, but high resource utilization can not be achieved. Achieving these two goals is still a very challenging issue in admission control area. Second, since the admissible bandwidth is calculated in advance and the admission decision is done by simple comparison, our scheme can perform admission control even for the requests arriving at the rate of up to the link rate. In addition, both edge and core routers need not manage any per-flow state. Thus, our scheme is scalable in terms of both the number of flow requests and the number of flows.

The rest of this paper is organized as follows. In Section II,

This work was supported by the Korea Research Foundation Grant funded by Korea government (KRF-2004-214-D00377).

S. Y. Nam was with the CNR Lab, Korea Advanced Institute of Science and Technology (KAIST) and the CyLab, Carnegie Mellon University, and is with the Department of Information & Communication Engineering, Yeungnam University, Gyeongsan 712-749, Korea.

S. Kim is with the Department of Information Statistics, Gyeongsang National University, Jinju 660-701, Korea.

D. K. Sung is with the Division of Electrical Engineering, School of Electrical Engineering and Computer Science, KAIST, Daejon 305-701, Korea.

we discuss related works. In Section III, we propose a concept of admissible bandwidth and a scalable architecture for admission control. In Section IV, we explain a probing method called minimal-backlogging method and a simplified path model as preliminaries. In Section V, we derive an estimator of the admissible bandwidth, propose an admission control algorithm, and discuss the scalability issues of the proposed scheme. In Section VI, we evaluate the performance of the proposed admission control scheme by simulation. Finally, conclusions are presented in Section VII.

II. Related Works

Admission control algorithms for internet flows can be classified into two categories. The first one is a traffic-model-based approach [18–21] and the second one is a measurement-based approach [5–16]. In the traffic-model-based approach input traffic is usually mathematically modeled and admission is determined based on the model. The accuracy of model-based approaches depends on the reliability of the assumed source models. There were some approaches calculating the effective bandwidth for a fluid input model or leaky-bucket regulated input traffic [18, 19]. However, these approaches do not consider a long-range dependence property which is an important characteristic of the current internet traffic [22–24]. It is possible to define effective bandwidth for the fractional Brownian input traffic which has self-similarity and long range dependence [20, 21, 25]. However, even the concept of effective bandwidth based on large deviation theory is not fully compatible with the realistic internet traffic according to [26]. In addition, if we calculate the effective bandwidth just based on the parameters of long-range dependent traffic considering some QoS such as loss probability, the utilization of the bandwidth can be very low due to huge rate fluctuation.

However, if we monitor the network status periodically, we can increase the bandwidth utilization by capturing the dynamic network status and allocating the resource accordingly. Measurement-based admission control algorithms (MBACs) can achieve a much higher utilization than traffic-model-based algorithms while providing somewhat relaxed QoS [5]. We can classify the MBAC schemes into two categories depending on the location of admission decision. First, admission decision is made at ingress end hosts. The end host probes the network by sending probe packets at the data rate it wants to reserve and recording the resulting level of packet losses (or ECN congestion marks [27]). The host then admits a flow only if the loss (or marking) percentage is below some threshold value. This kind of admission control is called as *endpoint admission control* [6, 7]. Endpoint admission control requires no explicit support from routers; routers keep no per-flow states and do not process reservation requests, and routers drop or mark packets in a normal manner. Thus, the endpoint admission control avoids the scalability problem of per-flow state management at each router. However, probing inherently involves a rather long set-up delay, on the order of seconds. In addition, probing overhead can cause a non-negligible problem especially when the network utilization is high. For example, when premium-class flash crowds are accessing to a specific server concurrently, it may incur denial-of-service (DoS) situation (also referred to as *thrashing* in [6]), where almost no flow is accepted by the low measured performance due to the overwhelming probing traffic. Thus, endpoint admission control has a scalability problem in terms of the number of flow requests. On the contrary, the proposed scheme is not subject to this scalability problem since admission decision is made promptly upon request arrival without further probing.

Second, admission decision is made at network nodes. Several measurement-based admission control algorithms belonging to this type have been proposed [5, 8–16] and our scheme also belongs to this category. Since it is difficult to predict future behavior accurately with traffic measurements, MBAC can lead to occasional violation of the contracted QoS. It is reported that the admission control algorithms of [5, 8–16] can not meet statistical QoS targets in terms of loss ratio [17]. Each of these algorithms makes the admission decision on a link-by-link basis. Thus, these algorithms require the cooperation of intermediate nodes in the admission control process. However, Cetinkaya *et al.*'s *Egress Admission Control* scheme [28] and our scheme are not subject to this coordination constraint because both of them work on an end-to-end basis. Cetinkaya *et al.*'s scheme achieves scalability by making admission control decisions only at egress routers without maintaining per-flow states. However, this admission control algorithm is not feasible when the load of existing flows between a specific ingress/egress router pair is very low or zero, since it is difficult to obtain a reliable service envelope [28] for a given interval of length T in this case. Our scheme works no matter how low the offered load is on a given path since the path is actively probed with probing packets. The proposed scheme is compared with Cetinkaya *et al.*'s scheme in more detail in Section VI.

III. System Architecture

Consider an autonomous system as depicted in Fig. 1. Routers A, E, F, G, and I are edge routers, and B, C, D, and H are core routers. Routers which provide interface to access networks are edge routers, and core routers do not operate as an interface. In the proposed architectural solution, an ingress router manages admissible bandwidth for the path to each relevant egress router. For example, Edge Router A manages admissible bandwidths for Egress Nodes E, F, G, and I, individually. Traffic arrivals at ingress routers of DiffServ domain are differentiated by the given QoS requirements. All arriving traffic with the same QoS requirements is treated as the same class.

Admissible bandwidth is managed separately according to the classes. Admissible bandwidth between a specific ingress/egress node pair is defined considering the level of services that can be provided. In this paper, we consider only delay bound violation probability as a QoS requirement. Let R'_j denote the admissible bandwidth for the j-th class between Ingress Router A and Egress Router E. Let d_j and ε_j be the delay bound and the threshold for the delay violation probability, respectively. $D_j(0)$ is a random variable representing the current end-to-end delay, and $D_j(R)$ is a random variable representing the end-to-end delay which the total traffic of

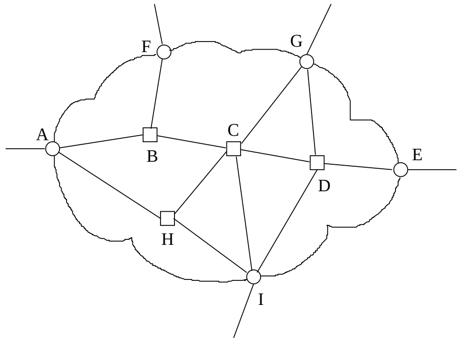

Fig. 1. Reference network model

class j experiences after admitting a flow with a rate of R. Then, the admissible bandwidth R'_j is defined by:

$$R'_j = \max\{R : P(D_j(R) > d_j) \ \leq \ \varepsilon_j\}. \qquad (1)$$

Thus, R'_j is the maximum available bandwidth that can be supported additionally satisfying the delay constraint.

In order to support QoS for a new flow while guaranteeing the contracted QoS for the existing flows, a negotiation is needed between the network and a new end-point application. The network determines whether to admit a new flow or not according to an admission control policy/algorithm assuming that the user complies with the contract. The characteristics of the new flow should be included in the contract because the network can not determine whether the required QoS will be satisfied or not if it does not know how much traffic will be offered by the new flow. Thus, we assume that the contract is made just based on the peak rate r_p of a flow. Peak rate r_p is the only traffic parameter used in our admission algorithm, and we assume that each flow is policed so that the instantaneous traffic rate can be maintained less than or equal to the peak rate r_p.

If the request from a new flow, which is destined to Router E and has a peak rate of r_p, arrives at Edge Router A, then Router A can accept the flow as the j-th class if the following condition is satisfied:

$$r_p < R'_j. \qquad (2)$$

Then, the delay constraint can be satisfied for both the existing and the new traffic. Since the proposed admission control algorithm is simple and ingress routers determine whether it accepts the new flow or not, admission control can be performed very quickly for real-time flows.

In this scheme, ingress routers need not calculate the admissible bandwidth whenever a new flow arrives. An ingress router sends probing packets to relevant egress routers to monitor the condition of each path, especially the available bandwidth for the path and calculates the admissible bandwidth R'_j for each ingress/egress node pair in advance.

IV. Preliminaries

Before the admission control scheme is proposed, we need to introduce an important concept of minimal backlogging [29], because this concept plays an important role in the proposed admission control scheme.

Calculation of the admissible bandwidth considering the delay QoS is the key problem in the proposed admission control scheme. We need to distinguish available bandwidth from admissible bandwidth reflecting QoS. For example, we consider a queueing system with a First-Come-First-Served (FCFS) service policy. C and λ denote the service rate in bits per second and the arrival rate of data packets in packets per second, respectively. Let L' denote the average length of the packets. Then, for the queueing system, available bandwidth C_a is defined as

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

where $\rho = \lambda L'/C$. This available bandwidth is the maximum spare service rate that the server can provide while maintaining stability of the system. In case of accepting a new flow with a rate of C_a, the desired QoS is usually not satisfied. Thus, the admissible bandwidth reflecting QoS is usually lower than the available bandwidth. However, we need to know the available bandwidth in order to obtain the admissible bandwidth. In [29], we proposed a probing scheme to estimate the available bandwidth of a single server. We briefly introduce the probing scheme and the available bandwidth estimation mechanism.

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

If we send a new probing packet into a queueing system just at the departure time of the previous probing packet, then there exists only one probing packet in the system. In order to introduce an estimator for available bandwidth, we define available service as follows:

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

Suppose that the size of probing packets is fixed to L. Then, we obtain in [29, Theorem 8] that for a $G/G/1$ queueing system,

$$\lim_{t\to\infty} E\left[\left|\frac{\tilde{Y}_{s,t}}{t-s} - C(1-\rho)\right|^q\right] = 0, \quad 0 < q < \infty. \qquad (3)$$

Thus, the service rate of probing traffic is equal to the available bandwidth of the queueing system probed by the minimal-backlogging method for an infinite duration, which implies that the service rate of minimally backlogging probing traffic can be used as an estimator of the available bandwidth.

V. Admission Control Scheme

As described in the previous sections, calculation of the admissible bandwidth is a crucial part of the proposed admission control scheme. If the calculated value is larger than the real available capacity, then delay QoS may not be guaranteed due to excessive amount of input traffic. On the other hand, if the calculated value is smaller than the real capacity, the utilization of the network resource decreases. In order to evaluate the admissible bandwidth between a specific ingress/egress router pair, we derive a relation that predicts the delay distribution if a new flow with rate R is accepted. If the new delay distribution can be predicted, then the admissible bandwidth can be calculated from (1). We also investigate a method to estimate the available bandwidth for a path between a given ingress/egress node pair by sending probing packets. We state a simple admission control scheme and discuss the complexity and scalability issues of the proposed scheme.

A. Model

We assume that there are only two classes of flows in the core network. The first is the premium class in which all flows abide by their peak rate constraints and have delay QoS requirements. This is the only class that is subject to admission control. The second is the best-effort class. Intermediate routers are assumed to give a strict priority to the premium class in managing two classes so that the delay of the premium class traffic is not affected by the best-effort traffic. Traffic is served according to the first-come-first-service (FCFS) policy in the same class.

We model a network path from a specific ingress router to an egress router as a simple path which is a concatenation of a fixed delay component (D_f) and a virtual server $\mathcal{S}$ as in [30]. In this model, the end-to-end delay of a packet D_e is decomposed as

$$D_e = D_f + D,$$

where D is the delay experienced by the packet at the virtual server. 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 arrives at $\mathcal{S}$ at time $a_p^s = a_p + D_f$. When the packet arrives at the destination node, it departs from both the path and the virtual server. Let $X_{u,\,v}$ denote the amount of traffic arriving at $\mathcal{S}$ in the time interval $[u, v]$. The traffic departed from $\mathcal{S}$ during the time interval $[u, v]$ is considered to be served by $\mathcal{S}$ during that time interval and is denoted by $Y_{u,\,v}$.

We assume that there is no backlogged traffic in the virtual server $\mathcal{S}$ at time 0. Then, the amount of the backlogged traffic of $\mathcal{S}$ at time t is given by

$$Q_t = \sup_{0\le s\le t} \{X_{s,\,t} - Y_{s,\,t}\}.$$

Let D_t denote the virtual delay which a virtual bit may experience if it arrives at $\mathcal{S}$ at time t. Since $\mathcal{S}$ is assumed to be empty at time 0, D_t is expressed as

$$D_t = \min\left\{\eta : \eta \ge 0 \,\text{and}\, X_{0,t} \le Y_{0,t+\eta}\right\}. \tag{4}$$

Then, D_t is the elapsed time from time t to the first instance at which the departure amount is greater than or equal to the arrival amount.

Now, suppose that a new flow is admitted and it starts sending traffic at a rate of R from time τ. We consider only a constant rate flow for the new flow because we want to evaluate the maximum admissible bandwidth R'_j of (1). Let $X^e_{u,v}$ and $X^n_{u,v}$ be the amount of traffic arriving at $\mathcal{S}$ from the existing and the new flows, respectively, in the time interval $[u, v]$. Then, the aggregate arriving traffic $X_{u,v}$ is the sum of $X^e_{u,v}$ and $X^n_{u,v}$. $Y^e_{u,v}$ denotes the amount of traffic, from the existing flow, served by $\mathcal{S}$ during the time interval $[u, v]$ and $Y^n_{u,v}$ denotes the amount of served traffic from the new flow in the same interval.

B. Evaluation of Admissible Bandwidth

In this subsection, we propose how to evaluate the admissible bandwidth when we know the available bandwidth. The amount of input traffic to a network path can be treated as being continuous in high speed communication networks. We assume $X_{u,\,v}(X^e_{u,\,v}, X^n_{u,\,v})$ and $Y_{u,\,v}(Y^e_{u,\,v}, Y^n_{u,\,v})$ to be continuous in this subsection.

Let D^n_t be the virtual delay of the new flow at time t. Since there is no priority between the existing flow and the new flow, the server treats the two traffic streams from the existing and new flows as if they come from the same flow. This implies that there is no difference in virtual delay at a given time no matter whether the virtual bit is of new flow or not. Thus, it follows:

Proposition 1: Suppose that a new flow starts at time $\tau \ge 0$. Then,

$$D^n_t = D_t,\ t > \tau.$$

For the virtual server with the arriving traffic amounts of $X^e_{u,v}$ and $X^n_{u,v}$ and the service amounts of $Y^e_{u,v}$ and $Y^n_{u,v}$, if we focus only on the arrival and service traffic of the new flow, we can know that a virtual bit arriving at time t from the new flow can be served just after the traffic arriving from the new flow during the interval $[0, t]$, $X^n_{0,t}$, is served completely under the assumption that $X^n_{0,s}$ $(s \ge 0)$ is increasing. Thus, D^n_t can be interchangeably expressed as $D^n_t = \min\{s : s \ge 0,\ X^n_{0,t} \le Y^n_{0,t+s}\}$.

Suppose that $Y^n_{\tau,t}$ can be obtained for any $t \ge \tau$. Since the new flow has a constant rate R, it is possible from the identity $D^n_t = \min\{s : s \ge 0,\ X^n_{0,t} \le Y^n_{0,t+s}\}$ to evaluate D^n_t. Then, the above proposition gives the value of D_t. However, $Y^n_{0,t}$ can be measured only when the new flow is really offered.

To estimate D_t before the new flow is offered into the network, we consider a queueing system with FCFS service policy, where $X^n_{\tau,t}$ is the input process and $\tilde{Y}^n_{\tau,\,t}$ is the possible amount of traffic served by the server, i.e. available service, during $[\tau, t)$. Let $\tilde{Q}^n_t$ denote the amount of backlogged traffic of the new flow in the queueing system at time t. Then, the virtual delay $\tilde{D}^n_t$ in this system is expressed as

$$\tilde{D}^n_t = \min\{\eta : \eta \ge 0,\ X^n_{b(t),\,t} \le \tilde{Y}^n_{b(t),\,t+\eta}\}, \tag{5}$$

where $b(t)$ is given by

$$b(t) = \sup\{s : \tau \le s \le t, \tilde{Q}^n_s = 0\}. \tag{6}$$

Note that $b(t)$ is the start time of the current backlogged period if $\tilde{Q}^n_t > 0$.

$\tilde{D}^n_t$ is defined in a different way from the definition of D_t, (4) because $\tilde{Y}^n_{s,\,t}$ can be larger than $X^n_{s,\,t}$ when the amount of arriving traffic $X^n_{s,\,t}$ is rather small. We can obtain the following relation between D^n_t and $\tilde{D}^n_t$.

Proposition 2: Suppose that traffic from a new flow is offered into a queueing system from time $\tau \ge 0$, then we have

$$D^n_t \le \tilde{D}^n_t, \quad \text{for } t \ge \tau.$$

Proof) The proof is given in [31, Proposition 5.2]. □

By Propositions 1 and 2, we can use $\tilde{D}_t^n$ as an upper bound of D_t. Although $\tilde{D}_t^n$ does not depend on $Y_{0,t}^n$, $b(t)$ in (5) is difficult to treat. Thus, we derive an alternative expression for the distribution of $\tilde{D}_t^n$ which avoids the use of $b(t)$.

Theorem 3: Suppose that a new flow starts at time $\tau \geq 0$. Then, we have that for $t \geq \tau$ and $s > 0$,

$$\tilde{D}_t^n > s \quad \text{if and only if} \quad \tilde{Q}_t^n > \tilde{Y}_{t,t+s}^n.$$

Proof) If $\tilde{Q}_t^n = 0$, then $b(t) = t$. It follows from (5) that $\tilde{D}_t^n = 0$. Thus, the two events in the above relation do not occur for any nonnegative s. Suppose $\tilde{Q}_t^n > 0$. Since the the queue is backlogged during $(b(t), t]$, we have that

$$\tilde{Q}_t^n = X_{b(t),t}^n - \tilde{Y}_{b(t),t}^n.$$

Thus, the event that $\tilde{Q}_t^n > \tilde{Y}_{t,t+s}^n$ is equivalent to that $X_{b(t),t}^n > \tilde{Y}_{b(t),t+s}^n$. By (5), the latter is equivalent to that $\tilde{D}_t^n > s$. □

Let d_0 be the delay bound which needs to be guaranteed for the class subject to admission control. Suppose that the system is stationary. In other words, there exists a random variable D^* such that

$$D^* = \lim_{t\to\infty} D_t.$$

Then, the delay D of a packet arriving at $\mathcal{S}$ can be estimated by D^*. The delay bound violation probability $P(D_e > d_0)$ can be expressed as:

$$\begin{aligned} P(D_e > d_0) &= P(D + D_f > d_0) \\ &= P(D^* > d_0'), \end{aligned} \tag{7}$$

where $d_0' = d_0 - D_f$. If $d_0' < 0$, i.e. $D_f > d_0$, then $Pr(D_e > d_0) = 1$ by (7) because D is non-negative. Since the delay bound can not be guaranteed in this case, no more traffic can be admitted, i.e. the admissible bandwidth R_j' is zero. Thus, we assume that $d_0' = d_0 - D_f > 0$ hereafter. Using Propositions 1 and 2, we can obtain an upper bound of delay violation probability as follows:

$$P(D^* > d_0') \leq P(\lim_{t\to\infty} \tilde{D}_t^n > d_0').$$

Applying Theorem 3 to the above inequality, we have

$$P(D^* > d_0') \leq P(\lim_{t\to\infty} \tilde{Q}_t^n > \lim_{t\to\infty} \tilde{Y}_{t,t+d_0'}^n).$$

It follows from (7) that

$$P(D_e > d_0) \leq P(\lim_{t\to\infty} \tilde{Q}_t^n > \lim_{t\to\infty} \tilde{Y}_{t,t+d_0'}^n). \tag{8}$$

In order to express the right hand side of the above inequality in a more explicit form, we assume that $\{\tilde{Y}_{\tau,t}^n, t \geq \tau\}$ has independent increments and the increments have a Gaussian distribution. More specifically,

$$\tilde{Y}_{\tau,t}^n = a(t-\tau) + \sigma B_{t-\tau}, \tag{9}$$

where $\{B_t, t \geq 0\}$ is a standard Gaussian process with independent increments and $B_0 = 0$. Then, a is the mean of $\tilde{Y}_{t,t+1}^n$, and σ^2 is the variance of $\tilde{Y}_{t,t+1}^n$. Then, $\tilde{Q}_t^n$ can be expressed as

$$\begin{aligned} \tilde{Q}_t^n &= \sup_{\tau \leq s \leq t} \{X_{s,t}^n - \tilde{Y}_{s,t}^n\} \\ &= \sup_{\tau \leq s \leq t} \{-\sigma B_{t-s} - (a-R)(t-s)\}. \end{aligned}$$

If we define $\tilde{Q}^n$ as $\tilde{Q}^n = \lim_{t\to\infty} \tilde{Q}_t^n$, then $\tilde{Q}^n$ has the following distribution [32, p.361]:

$$P(\tilde{Q}^n > x) = e^{-\mu x},$$

where $\mu = 2(a-R)/\sigma^2$.

Real-time applications or services will require a small value of d_0, usually less than 1 second. According to Recommendation G.114 [33] of Telecommunication standardization sector of International Telecommunication Union (ITU-T), one-way transmission time of up to 150 msec is acceptable for most user applications. Thus, $d_0' = d_0 - D_f$ is likely to be much smaller than 1. For such small values of d_0', the sigma-mean ratio of $\tilde{Y}_{t,t+d_0'}^n$ is too high, which means the non-negligible possibility of a negative amount of served traffic. Since the negative service amount is not realistic, we treat $\tilde{Y}_{t,t+d_0'}^n$ as if it were a constant of its mean ad_0' to obtain the following approximation:

$$\begin{aligned} P(\lim_{t\to\infty} \tilde{Q}_t^n > \lim_{t\to\infty} \tilde{Y}_{t,t+d_0'}^n) &\cong P(\tilde{Q}^n > ad_0') \\ &= e^{-\mu a d_0'}. \end{aligned}$$

Then, from the above equation and (8), we can obtain the following upper bound of delay bound violation probability:

$$P(D_e > d_0) \leq \exp\left(-\frac{2(a-R)a(d_0 - D_f)}{\sigma^2}\right) \tag{10}$$

Let $g(R)$ denote the right hand term of (10). If we evaluate R^* by

$$R^* = \max\{R : g(R) \leq \varepsilon\}, \tag{11}$$

then R^* becomes a lower bound of the admissible bandwidth for the class between the selected ingress/egress node pair. The explicit form of R^* can be obtained from (10) and (11) as

$$R^* = a + \frac{\log(\varepsilon)\sigma^2}{2(d_0 - D_f)a}. \tag{12}$$

This lower bound R^* is used to estimate the admissible bandwidth. From the above equation, we can obtain some insights about the behavior of the admissible bandwidth. First, we can observe that R^* increases as the average of the available bandwidth a increases because the first term on the right hand side of (12) is dominant when a increases. Second, if the variance of the available bandwidth σ^2 increases, R^* decreases because $\log(\varepsilon)$ is negative for $\varepsilon < 1$. Thus, the second term accommodates the burstiness of traffic by decreasing R^* for a large variance. Third, as the constraint is becoming more strict, that is, as the value of ε decreases, R^* also decreases. This is natural because in order to satisfy a more rigorous requirement, less traffic has to be admitted. Fourth, as the delay bound increases, R^* increases. This is also reasonable if we consider the limiting case that d_0 goes to infinity. Thus, we can know the behavior of the admissible bandwidth through the explicit form of R^*, (12), and the calculation complexity of R^* is very low since the value of R^* can be evaluated directly from the simple equation of (12) if the mean and variance of the available service are obtained through measurements.

C. Estimation of Available Service

In this subsection, we describe how to estimate the parameters a and σ of the available service $\tilde{Y}^n_{\tau,t}$ in (9) by using probing packets. We can obtain the value of $\tilde{Y}^n_{\tau,t}$ if we can provide the minimally backlogging probing traffic exactly. However, this is not possible in real networks. Instead, we send the probing packets by the scheme described in [30], which enable the probing packets to be offered to the virtual server of the network path satisfying the minimal backlogging condition approximately.

The probing scheme is window-based. Let T denote the duration of one window. Since the minimal backlogging condition is not satisfied exactly during probing, several busy periods of probing packets may exist during a window. Let a^s_p and d_p be the times when a probing packet p arrives at the virtual server $\mathcal{S}$ corresponding to the specific path and p departs from the server, respectively. For each probing packet p, the virtual server is considered to be backlogged from the arrival time a^s_p to the departure time d_p. Then, the virtual server is continuously backlogged for k probing packet transmissions from the j-th probing packet in the interval $[a^s_j, d_{j+k-1}]$ if

$$d_{j+m} \ \geq \ a^s_{j+m+1}, \quad \text{for all } 0 \leq m \leq k-2, \tag{13}$$

for $k \geq 2$.

Eqn. (3) means that the service rate of minimally backlogging probing traffic $\tilde{Y}^n_{s,t}/(t-s)$ is an asymptotically unbiased estimator of the available bandwidth. Thus, if there are k probing packets in the longest busy period of the current window, we approximate the available bandwidth $\tilde{Y}^n_{t,t+1}$ for an interval of 1 second in the current window as

$$\hat{R} = \frac{kL}{d_k - a^s_1}, \text{[1]}$$

where L is the size of the probing packets. We use the longest busy period for reliable estimation. We let R_i denote the value of $\hat{R}$ for the i-th previous window from the current window, and $i = 0$ corresponds to the current window. Then, a is estimated as

$$a' = \frac{1}{M_a} \sum_{m=0}^{M_a-1} R_m,$$

where M_a is the number of most recent windows considered for estimation of the mean of $\tilde{Y}^n_{t,t+1}$ and an estimator for σ^2

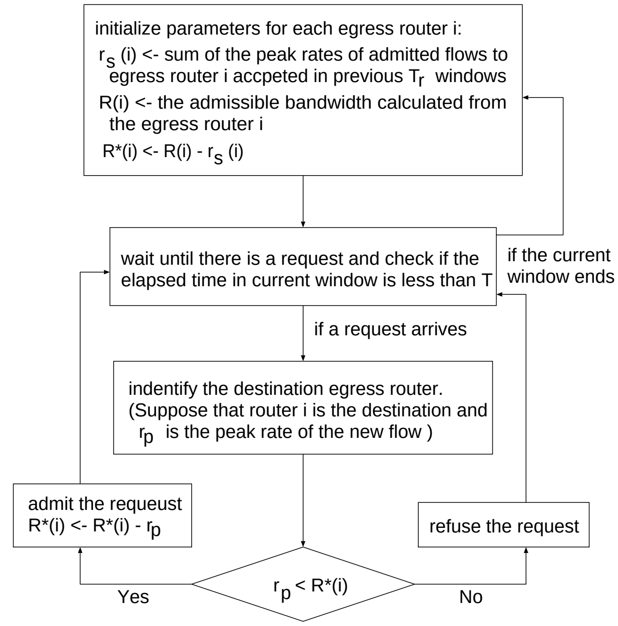

Fig. 2. Admission control algorithm for an egress router i

is given as

$$\sigma'^2 = \gamma \cdot \frac{1}{M_v - 1} \sum_{m=0}^{M_v-1} (R_m - \bar{R})^2,$$

where M_v is the number of most recent windows considered for estimation of the variance of $\tilde{Y}^n_{t,t+1}$, $\bar{R} = \sum_{m=0}^{M_v-1} R_m/M_v$, and γ is the variance multiplication factor (VMF). VMF is a factor used to compensate the error in estimation of the variance of $\tilde{Y}^n_{t,t+1}$. We will investigate the effect of VMF through simulation in the next section.

The scheme proposed in [30] estimates the fixed delay component D_f as well as the available bandwidth. In order to estimate a and σ of $\tilde{Y}^n_{t,t+1}$, other available bandwidth estimation schemes [35–37] can also be used. However, D_f may need to be estimated separately when other schemes are used.

D. Admission Control Algorithm

Let's consider an admission control algorithm for a specific ingress/egress router pair. The egress router calculates the lower bound of the admissible bandwidth R^* using (12) once every T seconds and sends it back to the ingress router.

Then, the ingress router performs admission control according to the algorithm described in Fig. 2. If the ingress router has not given admission to any flow in the previous window, the ingress router admits the request of a new flow with a peak rate of r_p if the following condition is satisfied:

$$r_p < R^* - r_s, \tag{14}$$

where r_s is the sum of the peak rates of the flows admitted in the current window before the current request.

In fact, more flows need to be considered for the r_s term in the above inequality. If we consider the time interval between the end of the probing period for the previous window and

[1] When we estimate the available bandwidth $\tilde{Y}^n_{t,t+1}$ for an interval $[t, t+1]$, if the duration of the longest busy period $d_k - a^s_1$ is too short compared with the interval length, one second, then the behavior of cross traffic out of the range of probing period may not be reflected, which causes an error in available bandwidth estimation. As the probing duration increases, the estimation accuracy improves according to [34]. However, since the increased probing duration implies high overhead of probing traffic, a proper probing duration needs to be determined considering the tradeoff between the estimation error and the probing traffic overhead. This issue will be investigated further in the future.

the end of the probing period for the current window, then the most recent value of R^* is the one obtained in the previous window because the probing for the current window is not finished yet. In this case, the rates of the flows admitted after the probing period in the previous window are not reflected in the calculation of R^*, and thus, the summation of the rates of these flows of the previous window also needs to be included in r_s.

In addition, if there are some flows that start to send packets after an initial idle period, the effective rates of these flows may not be detected and reflected on R^* in the next probing period. In this case, for conservative estimation of the admissible bandwidth, the number of windows considered for r_s, T_r may need to be increased.

E. Complexity and Scalability Issues

The admissible bandwidth is calculated according to the simple equation of (12) and the admission decision is made by just comparing the peak rate of the requesting flow with the admissible bandwidth according to (14). In addition, the admissible bandwidth is not calculated on demand, but it is calculated periodically in an interval of at least one second. Thus, the proposed scheme has a low complexity and can perform per-flow admission control even at a high request arrival rate through high speed links.

We now investigate scalability issues of the proposed admission control scheme. Our scheme does not require per-flow state management or processing at the core routers except the class-level scheduling. The class-level scheduling, especially priority scheduling, can be implemented in the framework of DiffServ. Since even the edge routers do not manage per-flow states, our scheme is scalable in terms of the number of flows.

In our scheme, an ingress router needs to manage admissible bandwidth for the relevant egress routers. Although there are enough memory space for many egress routers, the number of monitored egress routers may be limited due to the probing overhead. Let M and Z denote the total number of egress routers which can interact with a given ingress router, and the number of egress routers which can be probed in a given time interval T, respectively. Then, since Z is usually smaller than M, we consider the following approach. We give priority to the egress routers which interacted with the ingress router more recently in managing the admissible bandwidth information and monitor the admissible bandwidth for the most recently active Z egress routers. These recently active Z routers are managed in the *probing list*. In order to check the activity of the egress routers in the probing list, an individual clock is allocated for each managed egress router. If we increase every clock at a constant rate and reset the clock individually whenever there is a new request for the corresponding egress router, then each clock measures the inactivity time of each egress router in terms of flow requests. If a new active egress router appears, then the egress router which has been inactive for the longest period can be deleted from the probing list yielding to the new egress router. Although an egress router is evicted from the probing list, the admissible bandwidth for that router can be retained in the memory depending on the memory management policy. If a new request arrives destined to the egress router for which no admissible bandwidth information is retained, then probing is done consecutively for a few times in order to estimate the admissible bandwidth based on the mean and standard deviation of the available bandwidth. Afterwards, the corresponding egress router is registered in the probing list and the admissible bandwidth is updated periodically.

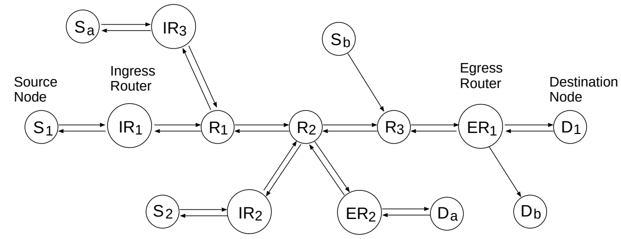

Fig. 3. Network topology for simulation

VI. Numerical Results

In this section, we evaluate the performance of the proposed admission control scheme in terms of delay QoS, i.e. delay violation probability, and utilization through OPNET simulation. We consider a network topology as shown in Fig. 3 for simulation. Nodes IR_1, IR_2 and IR_3 are ingress routers and Nodes ER_1 and ER_2 are egress routers. Nodes R_1, R_2, and R_3 are core routers. Nodes S_1, S_2, S_a, and S_b are source nodes where data traffic is generated. Nodes S_1 and S_2 generate only real-time flows that are subject to admission control, Node S_a generates both high and low priority traffic and S_b generates only background traffic that is not subject to admission control. Flows from source nodes S_1 and S_2 are always destined to the destination node D_1. Traffic streams generated at Nodes S_a and S_b are directed to Nodes D_a and D_b, respectively. We first focus on the scenario where only ingress routers IR_1 and IR_2 are admitting premium class traffic and later consider the situation where the ingress router IR_3 also admits high priority class traffic. The proposed admission control scheme does not guarantee the end-to-end delay QoS from Source Nodes S_1 or S_2 to Destination Node D_1, but guarantees the delay QoS from the ingress routers (IR_i, $i = 1, 2, 3$) to the egress routers (ER_j, $j = 1, 2$).

Each node is modeled as an output queued router with a strict-priority (SP) scheduling policy. Premium class traffic that is allowed by admission control is given a strictly higher priority than best-effort traffic that is not subject to admission control. Each link has a link rate of 10 Mbps and a propagation delay of 5 msec. The sizes of all probing packets and data packets are fixed to 4000 bits. The duration of one time window T is 1 sec, and the number of probing packets sent per time window, N is 100 packets. Thus, the average probing traffic rate is 400 kbps. The values of the probing-related parameters are set: $N_m = 0.70 \times N = 70, N_s = 0.05 \times N = 5, \alpha_m = 0.1,$ and $\alpha_s = 0.6$, where N_s and N_m are two thresholds which determines short and long busy periods of probing packets, respectively, α_s is the probing rate multiplication factor used for short busy periods and α_m is the maximum rate multiplication factor for medium busy periods [30].

We consider three types of traffic patterns for the flows that are subject to admission control: exponential on-off traffic, Pareto on-off traffic, and peak rate-policed self-similar traffic. The lifetime of each flow is exponentially distributed with the same mean of 200 seconds. The first one is simple on-off traffic whose *on* and *off* period lengths are exponentially distributed. The average lengths of both *on* and *off* periods are 0.5 second. No traffic is generated during *off* periods and packets are generated at the peak rate of r during *on* periods. Thus, the average rate is $r/2$. The peak rate of each flow is fixed to 512 kbps. The flow inter-arrival time is exponentially distributed with an average of 1 second.

The second one is also on-off traffic, but the lengths of *on* and *off* periods have a Pareto distribution. If X has a Pareto distribution with the shape parameter α and the scale parameter β, then X has a density function $f(x)$ and a distribution function $F(x)$ of

$$f(x) = \frac{\alpha\beta^{\alpha}}{x^{\alpha+1}}, \quad F(x) = 1 - \left(\frac{\beta}{x}\right)^{\alpha}, \quad \text{for } x \geq \beta.$$

If the shape parameter α is less than 2, X has an infinite variance. If α is less than 1, X has infinite mean and variance. We fix the value of α to 1.9 for both *on* and *off* periods, and the values of β's are 0.17 for both *on* and *off* periods so that the mean *on* and *off* lengths are 0.36 second as in [28] according to $E[X] = \beta\alpha/(\alpha-1)$ $(\alpha > 1)$. The peak rate of each flow is either 256 or 512 kbps. The average inter-arrival time of Pareto flows is also 1 second.

The third type of input traffic is self-similar traffic. But, the traffic is reshaped so that the instantaneous rate can not exceed the pre-specified peak rate. We use a multi-fractal model [38] to generate self-similar traffic patterns. The Hurst parameter of each flow is 0.8. The peak rate is fixed to 512 Kbps and the average rate is 256 Kbps. The average flow inter-arrival time is 1 second.

We consider 150 msec as the maximum allowable end-to-end delay according to ITU-T Recommendation G.114 [33]. In case of 3GPP, an end-to-end delay of 150 msec is preferred and transfer delay is defined as a 95-percentile of the distribution of delay for all delivered data packets [39, 40]. We set the threshold for the delay bound violation probability (ε) to 0.01 rather conservatively.

First, we investigate the effect of T_r, the number of windows (except the current window) considered for calculation of r_s in (14). T_r is referred to as complement window because it is used to complement admissible bandwidth by covering flows that are not considered in the calculation of the last admissible bandwidth R^*. Fig. 4 shows the delay violation probability for various values of complement window T_r under exponential on-off traffic loads, that is, in this case only exponential on-off flows are admitted by IR_1 and IR_2. Each probability value is obtained from 20 simulations with different seeds in the random number generator. A zero value of T_r implies that only the current window is considered to calculate r_s. If $T_r = i, \ i \geq 1$, then T_r includes up to the i-th previous window from the current window. The delay bound d_0 is set to 150 msec for both Ingress Routers 1 and 2. We can observe that when $T_r = 0$, the delay performance requirement is significantly violated. The measured delay bound violation probability is near 0.1, i.e. much higher than the target value of 0.01. This is because the rates of flows that are accepted after the last probing in the previous window are not reflected in the calculation of the admissible bandwidth. However, as the value of T_r increases to 1, the measured delay violation probability decreases significantly since the rates of flows that are accepted after the last probing in the previous window are now reflected in the calculation of the admissible bandwidth through the term of r_s. We observe that the measured delay violation probability decreases as the value T_r increases. However, since every accepted flow starts from busy period in the simulations, the T_r values need not be larger than 1. Thus, the value of T_r is fixed to 1 hereafter. We need to find a different reason for the violated delay QoS. There is no significant difference between the performance of the traffic streams from IR_1 and from IR_2.

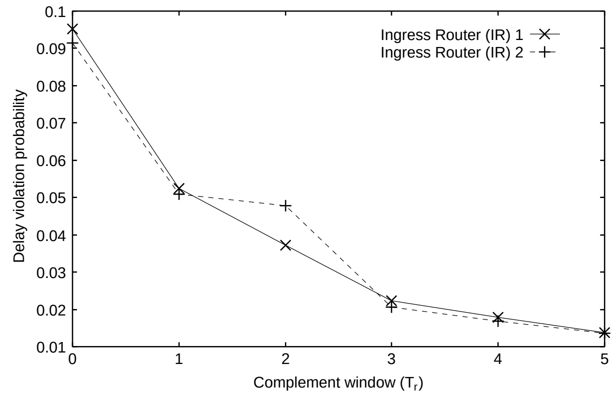

Fig. 4. Delay violation probabilities for various values of complement window T_r under exponential on-off traffic loads

The following simulation result shows that the violation of delay QoS occurs due to the estimation error of the probing mechanism. Fig. 5 compares the measured available bandwidth and the available bandwidth estimated by the probing mechanism proposed in [30] for the path between IR_1 and ER_1. The egress router obtains the measured available bandwidth for the path from the available bandwidth measured at each intermediate node every time window. Fig. 5 also shows the admissible bandwidth calculated by (12). We can observe that the available bandwidth estimated by the probing scheme is closely tracking the measured available bandwidth. The calculated admissible bandwidth is much lower than the available bandwidth. This can be easily explained with the following example. Let us consider a case where the bandwidth resource is allocated according to the peak rate, which is referred to as a *peak rate allocation scheme* in this paper. If flows with a peak rate of 512 kbps and an average lifetime of 200 seconds arrive at an average interval of 1 second from IR_1, then the link $\mathrm{R}_3 - \mathrm{ER}_1$ just before the egress router ER_1 will be occupied by 19 flows in about 19 seconds. If flows also arrive from IR_2, then it would take less time to fill up the link $\mathrm{R}_3 - \mathrm{ER}_1$ in case of peak rate allocation. In that

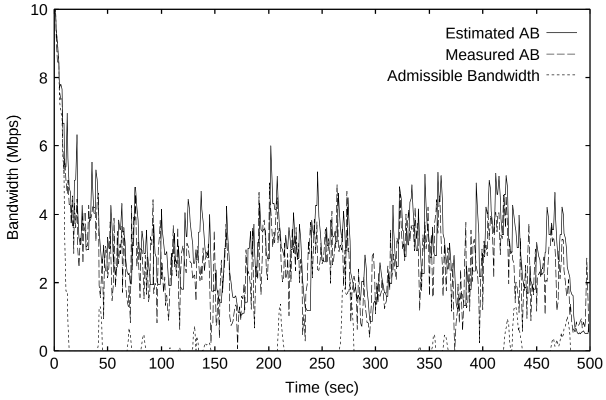

Fig. 5. Measured/Estimated available bandwidths and admissible bandwidth over time

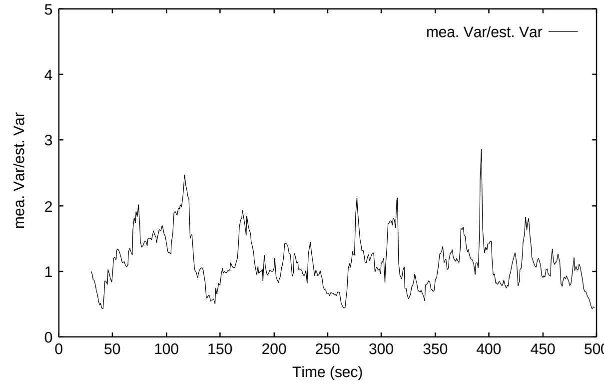

Fig. 6. Ratio of the measured variance to the estimated variance of available bandwidth under exponential on-off traffic loads

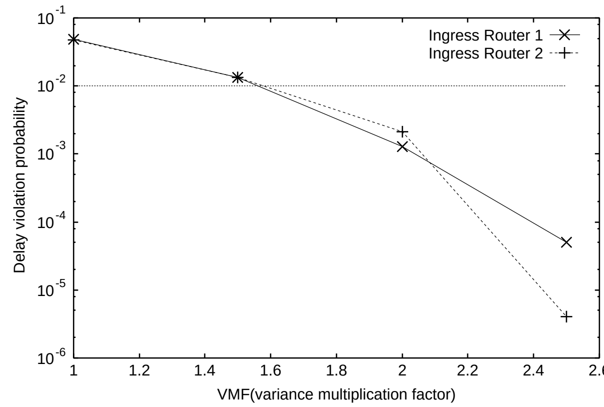

Fig. 7. Delay violation probabilities for various VMF values under exponential on-off traffic loads

case, no more flows can be admitted from either IR_1 or IR_2 until at least one of existing flows ends. Since the lifetime of each flow is much longer than the flow inter-arrival time, the any available resource, i.e. bandwidth, is likely to be taken quickly by frequently arriving flows. Thus, the admissible bandwidth will remain low compared with the peak rate of flows during full simulation time. In case of measurement-based admission control, the situation is a little better, but the admissible bandwidth should be low enough to guarantee the required delay QoS. Although the estimation by the probing mechanism seems rather accurate in Fig. 5, we can find the difference between the measurement and the estimation of available bandwidth from the following figure.

Fig. 6 shows the ratio of the measured variance to the estimated variance over time when admission control is performed under exponential on-off traffic loads. The measured variance of available bandwidth is obtained from 30 most recent windows ($M_v = 30$). The estimated variance is also obtained from 30 windows. The number of windows for variance estimation M_v needs to be sufficiently large for reliable estimation of variance. However, we set the number of windows for estimation of the average of available bandwidth M_a to 10 in order to follow the changing value of available bandwidth quickly. As we can observe in the figure, the ratio of the measured variance to the estimated variance is in the range of $[0.5, 3.0]$ for most of the simulation time. If we underestimate the variance σ^2, then the admissible bandwidth will be overestimated by (12). In that case, delay QoS may not be guaranteed due to excess traffic. Thus, variance multiplication factor VMF γ may be needed in order to complement this error in estimation of the variance of the available bandwidth and guarantee delay QoS for real-time flows.

Fig. 7 shows the delay violation probability for various values of VMF γ when exponential on-off traffic loads are offered. The delay bound d_0 is 150 msec and the delay violation probability threshold ε is 0.01. Simulation is performed for 500 seconds. When $\gamma = 1.5$, the measured delay violation probability is approximately 0.013 for both IR_1 and IR_2. Thus, the delay QoS is slightly violated for $\gamma = 1.5$. The delay QoS is well satisfied for $\gamma \geq 2.0$. From this figure, we find that VMF is effective in guaranteeing the delay QoS since VMF can complement the variance estimation error. However, a large value of VMF may lead to a low utilization as shown in the following figure.

Fig. 8 shows the utilization of the link between Egress Router ER_1 and Destination Node D_1 for various values of VMF γ in the same environment as Fig. 7. The utilization is measured during the last half of the simulation time. We observe that the utilization decreases as the value of γ increases. VMF is required in order to complement the variance estimation error of the probing scheme. However, if a too large value of VMF is used, then the admissible bandwidth will be underestimated by (12) due to the overestimated variance. Since a low admissible bandwidth can admit a smaller number of flows, a large value of γ tends to decrease the utilization as shown in Fig. 8. Thus, from Figs. 7 and 8 we find that there is a trade-off between the delay bound guarantee and utilization, and VMF is a tuning parameter which can adjust this trade-off. Hereafter, we investigate the effect of VMF on the performance, delay QoS and utilization, in more detail through a diverse set of simulations.

Fig. 9 compares the utilization of the link between ER_1 and D_1 for $\gamma = 1.0$ with that for $\gamma = 1.5$ when delay bound (d_0) has various values from 50 msec to 150 msec under Pareto on-off traffic loads. We use two traffic rates, 256 Kbps and 512 Kbps, for the peak rate of each flow while maintaining *on* and *off* period lengths at 360 msec in order to check the effect of

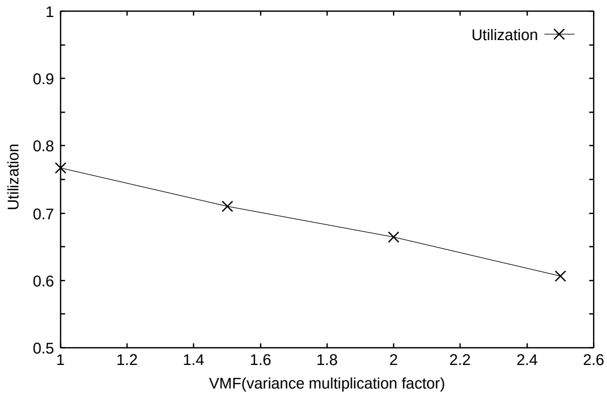

Fig. 8. Measured utilization of link $\mathrm{ER}_1 - \mathrm{D}_1$ for various VMF values under exponential on-off traffic loads

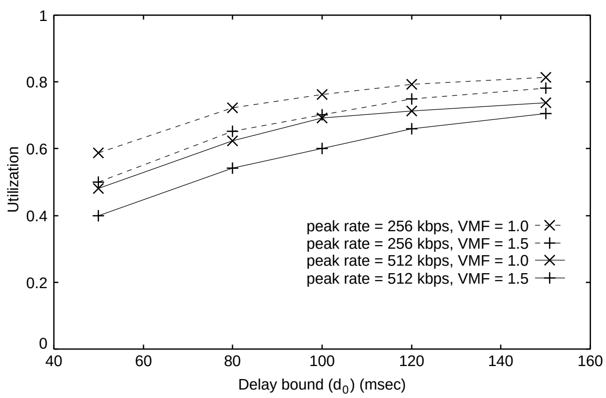

Fig. 9. Measured utilization of link $\mathrm{ER}_1 - \mathrm{D}_1$ for various delay bounds under Pareto on-off traffic loads

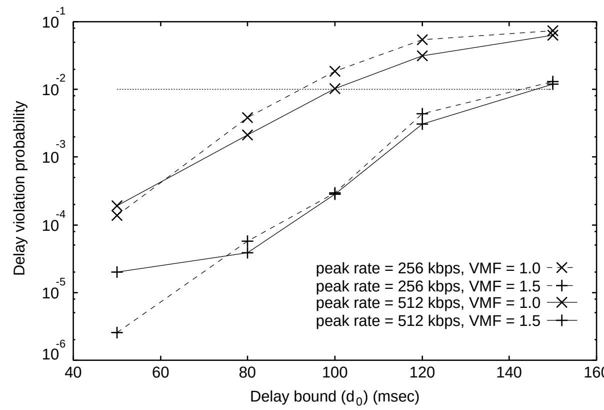

Fig. 10. Delay violation probabilities for various delay bounds under Pareto on-off traffic loads

flow's peak rate on the performance. We can observe that the utilization is higher for $\gamma = 1.0$ than for $\gamma = 1.5$ by the same reason explained for Fig. 8. We also observe that when the peak rate of each flow is lower, the utilization is higher. This phenomenon can be explained intuitively with the following simple example. Let us assume that the admissible bandwidth between a given ingress/egress router pair is 15 Mbps in the current time interval. If the peak rate of every flow is 16 Mbps, then no flow can be admitted during the current time interval. However, if the peak rate of each flow is 8 Mbps, then 1 flow can be admitted while leaving 7 Mbps of bandwidth still unused during the current interval. If the peak rate of each flow is 1 Mbps, then 15 flows can be admitted up to the available capacity. Thus, a smaller peak rate can contribute to a high utilization.

Fig. 10 shows the delay violation probability obtained by the proposed admission control scheme in the same environment as Fig. 9. Since the delay violation probability for IR_2 is very similar to that for IR_1, the results for only IR_1 are shown in Fig. 10. We can observe that the delay QoS is satisfied for almost all values of d_0 when $\gamma = 1.5$ since the variance σ^2 is rather conservatively estimated. On the other hand, the delay QoS is satisfied for $d_0 \leq 0.08$ when $\gamma = 1.0$. Thus, the VMF value of 1.5 is better for guaranteeing the delay performance. We can observe that the measured delay violation probability increases as the delay bound d_0 increases. This tendency can be explained as follows. From (12) we can expect that the admissible bandwidth will increase if the delay bound d_0 is increased abruptly because the second term of the right hand side is negative. We can also expect this behavior intuitively if we consider the case of very large value of d_0. Thus, if more traffic flows are admitted as d_0 increases, it is likely that the rate of the aggregate traffic fluctuates more frequently with a bigger amplitude. Then, the difference between the variance of real available bandwidth and that of the estimated available bandwidth, especially the ratio of the two variances, may increase. Since underestimation of variance of the available bandwidth implies overestimation of the admissible bandwidth by (12), delay violation probability increases due to more excess traffic as d_0 increases.

We observe that the delay violation probability is slightly higher when the peak rate of each flow is lower from Fig. 10. The reason can be explained with the following example. Let us consider an $M/M/1$ queueing system, where the service rate is μ' and the arrival rate is λ'. Then, the packet sojourn time W in the queueing system is exponentially distributed with parameter $(\mu' - \lambda')$ [41] and the expected sojourn time is

$$E[W] = \frac{1}{\mu' - \lambda'} = \frac{1}{\mu'}\frac{1}{1-\rho'},$$

where $\rho' = \lambda'/\mu'$ is the offered to the system. We assume that μ' is fixed considering constancy of link rates. From the above relation, we know that the packet delay increases drastically as the offered load ρ' approach 1. It means if some flow is admitted at high loads, then the delay increment is likely to be higher than the case where the same flow is admitted at low loads. From Fig. 9 we know that the utilization is higher when the peak rate is lower. Since the probing scheme may have an error in estimation of the variance of the available bandwidth, if a similar number of flows are admitted over the capacity due to variance estimation error, the system in a high utilization is likely to be affected more than the relatively low utilized system yielding a higher delay violation probability.

We now test the proposed admission control scheme for various input traffic loads. We evaluate the performance of the proposed scheme for exponential on-off traffic loads, Pareto

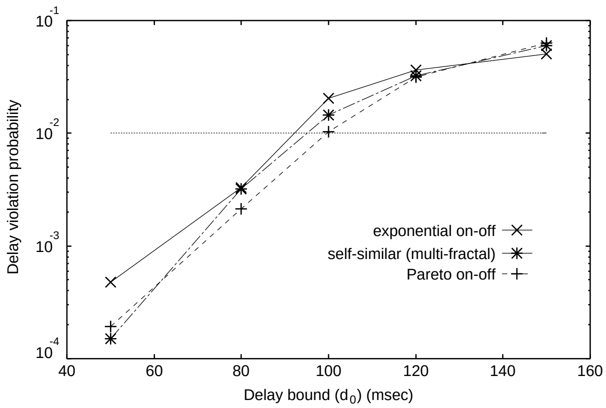

Fig. 11. Delay violation probabilities for various delay bounds under exponential on-off, Pareto on-off, and self-similar traffic loads

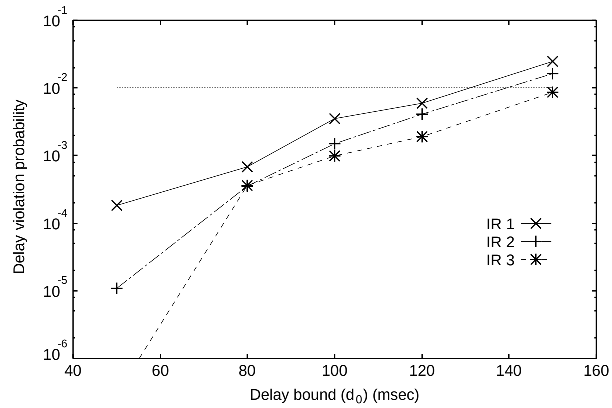

Fig. 12. Delay violation probabilities in the presence of high priority cross traffic (exponential on-off traffic) (VMF = 1.5)

on-off traffic loads, and self-similar traffic loads, separately, and Fig. 11 compares the test results. The peak rate and the average rate are the same as 512 Kbps and 256 Kbps, respectively, for all traffic patterns. The value of γ is fixed to 1. We observe that a similar performance is obtained although the detailed traffic pattern is different. Thus, we find that the detailed traffic pattern is not a crucial factor in determining the delay violation probability if the peak rate and the average rate are the same.

Thus far, we considered the case where there is no cross traffic, i.e., only ingress routers IR_1 and IR_2 admitted high priority traffic destined to ER_1. We now consider the case where there is high priority cross traffic. In this scenario, Ingress Router IR_3 admits high priority flows destined to D_a (via ER_2) from the source node S_a. Then, the traffic from IR_3 to ER_2 interfere with the high priority traffic from IR_1 to ER_1 at the link $R_1 - R_2$. Exponential on-off traffic patterns are used for high priority traffic flows that are subject to admission control. The value of γ is set to 1.5 in this case. Fig. 12 shows the delay violation probabilities obtained for the flows from IR_1, IR_2, and IR_3, respectively. We find that the traffic from IR_1 has a higher delay violation probability than the traffic from other Ingress Routers. The performance of IR_1 and IR_2 was very similar when there was no high priority cross traffic. The reason why the performance of IR_1 is worse than other Ingress Routers is as follows. Only the traffic from IR_1 interferes with two traffic streams from IR_2 and IR_3 while other traffic flows from IR_2 and IR_3 interfere with only one traffic stream from IR_1. Due to this higher interference, IR_1 may experience higher delay violation probability. However, even for IR_1 the delay bound is statistically guaranteed for the delay bounds of up to 120 msec when $\gamma = 1.5$.

We also evaluated the proposed admission control scheme in the presence of low priority cross traffic by sending low priority traffic from Nodes S_a and S_b to Nodes D_a and D_b, respectively. However, we found that the delay violation probability and the utilization were not much affected by the low priority cross traffic because the interference of low priority cross traffic on the service of premium class traffic is negligible due to the priority scheduling at core routers.

We now investigate the effect of measurement time window T and compare the proposed admission control scheme with Egress Admission Control (EAC) scheme [28]. Thus far, the value of T has been fixed to 1 second. Fig. 13 first shows the measured delay violation probability for various values of T and also shows the performance of EAC. No cross traffic is offered and the values of γ is 1 for our scheme. We can observe that the delay performance requirements are well satisfied as the the value of T increases in our scheme. Especially, when $T = 5.0$, the delay QoS is satisfied for all delay bounds. During one window of the length T, bandwidth resources are reserved according to the peak rate of a flow. Therefore, we can expect the bandwidth resource will be allocated more conservatively as T increases. As a consequence of conservative resource allocation, the delay QoS is well satisfied for large values of T. EAC guarantees the every delay bound well as shown in the figure. Fig. 14 shows the measured utilization of the link $\mathrm{ER}_1 - \mathrm{D}_1$ for various values of T and also shows the performance of EAC. Although the delay bound is well guaranteed as T increases in Fig. 13, the utilization decreases as the value of T increases in Fig. 14. Thus, the value of T needs to be selected considering the tradeoff between the delay QoS and the resource utilization.

On the other hand, EAC exhibits very low utilization as shown in Fig. 14. We find that the service envelope of EAC may estimate the service capacity of a given path very conservatively. As an example, let us consider a server with a fixed service rate C and we assume that we can measure the arrival and departure time of each arriving packet. If one packet with the packet size of l arrives, then the traffic amount of l is served during an interval of l/C. From the service envelope, we can estimate the service rate as $l/(l/C) = C$. Thus, when there is no fixed delay component, the service capacity may be estimated reasonably. As a more realistic case, let us assume that the packet sent by the sender can arrive at the server after a non-zero fixed delay d_f. Since EAC neither detects a fixed delay component nor removes the fixed delay component from the end-to-end delay, even though packets are arriving at an interval of $d_f + l/C$ the service envelope can be constructed. In this case, the service rate estimated from the service envelope becomes $l/(d_f + l/C)$ when there is no cross traffic. We need to note that the estimated service rate

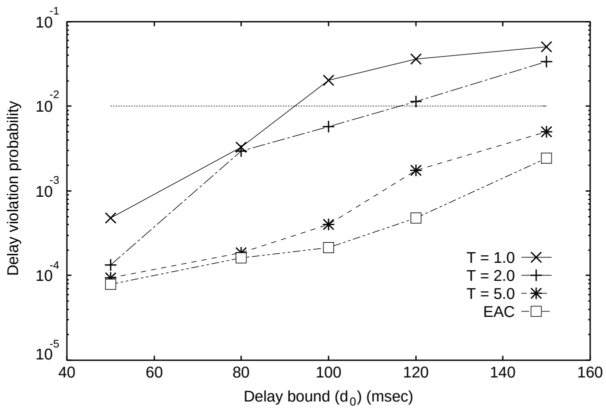

Fig. 13. Effect of T on the delay violation probability and performance comparison with egress admission control (EAC) scheme

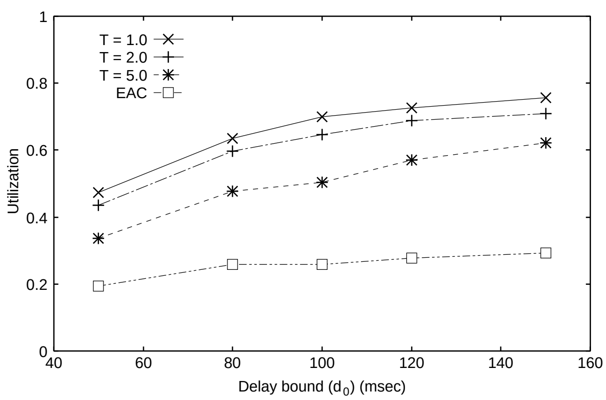

Fig. 14. Utilizations of link $\mathrm{ER}_1 - \mathrm{D}_1$ for various measurement window T and comparison with EAC

can become very small regardless of cross traffic when d_f is much larger than the packet service time l/C. Since the service capacity is significantly underestimated in our scenario by this reason, the utilization is low for EAC.

Thus far, we have fixed the link rate of each link to 10 Mbps. We now investigate the performance of the proposed admission control scheme for different link rates with a mix of traffic patterns. Fig. 15 compares the delay violation probabilities for two different link rates, 50 Mbps and 10 Mbps. The measurement window length T is fixed to 1 second. In this scenario, high priority cross traffic is applied from S_a to D_a and all the three types of traffic patterns (exponential on-off, Pareto on-off, self-similar) are used for the high priority traffic flows. The average flow inter-arrival time is 1 second for each traffic pattern at each Ingress Router $\mathrm{IR_i}(i = 1, 2, 3)$. Since IR_1 has the highest delay violation probability among ingress routers, Fig. 15 shows the delay violation probability for only IR_1. We can observe that the delay violation probability is lower for higher link rates. This is because the resources are used more conservatively in case of a link rate of 50 Mbps. Under the same admission request arrival patterns, the net amounts of available and admissible bandwidths are usually higher for a link rate of 50 Mbps than for 10 Mbps. Consequently, the effective arrival rate of accepted flows for a time window of the length T is higher for a link rate of

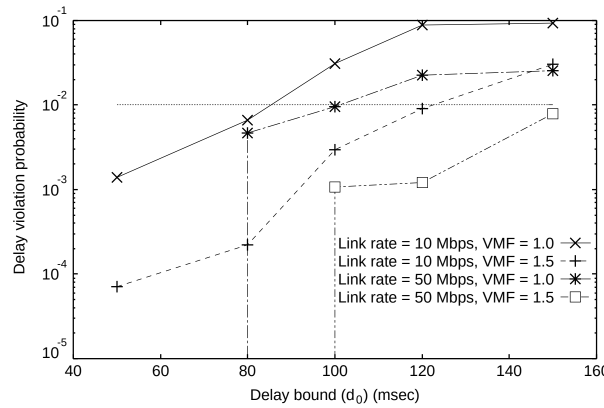

Fig. 15. Comparison of the delay violation probabilities for different link rates with a mix of traffic patterns (exponential on-off, Pareto on-off, self-similar traffic) in the presence of high priority cross traffic

50 Mbps than for 10 Mbps. Since bandwidth resources are allocated conservatively according to the peak rate of each flow during a window, a higher acceptance rate during a window implies more frequent conservative allocation of resources. Due to this conservative allocation of resources, the delay QoS is well satisfied for a link rate of 50 Mbps. When the variance estimation error is compensated with the VMF of 1.5, the guaranteed delay bound increases from 80 msec to 120 msec for the link rate of 10 Mbps and it also increases from 100 msec to 150 msec for the link rate of 50 Mbps. Thus, VMF is helpful to guarantee the statistical delay bound for a rather large delay bound. Through many simulations, we find that no VMF is needed for the delay bound of up to 80 msec, the VMF of 1.5 can guarantee the delay bound of up to 120 msec. According to (12), the admissible bandwidth R^* is determined by the bound on the variable component of delay $(d_0 - D_f)$, but not by just the delay bound d_0. In the current simulation setting, the fixed delay component D_f is only around 15-20 msec, i.e. three or four times the propagation delay (5 msec). However, in real networks the value of D_f is likely to be larger than 20 msec due to the larger number of hops and some processing delay component at the end hosts. Thus, if the fixed delay component is over 50 msec, then VMF of 1.5 can guarantee the delay bound of up to 150 msec.

Fig. 16 compares the utilizations of two links $\mathrm{ER}_1 - \mathrm{D}_1$ and $\mathrm{ER}_2 - \mathrm{D_a}$ for two link rates, 50 Mbps and 10 Mbps. The environment is the same as the case of Fig. 15. We only show the utilization for the VMF of 1.0 since we explained the difference in utilization caused by VMF at Fig. 8. We can observe that the utilization of $\mathrm{ER}_2 - \mathrm{D_a}$ is lower than that of $\mathrm{ER}_1 - \mathrm{D}_1$ when the link rate is the same. The link $\mathrm{ER}_1 - \mathrm{D}_1$ receives traffic from two Ingress Routers IR_1 and IR_2, but the link $\mathrm{ER}_2 - \mathrm{D_a}$ receives traffic from only one Ingress Router IR_3. Since the link $R_1 - R_2$ is shared by both IR_1 and IR_3, $R_1 - R_2$ is the bottleneck link on the path between IR_3 and ER_2. Thus, if half of the link rate of $R_1 - R_2$ is used by IR_1, the utilization of $\mathrm{ER}_2 - \mathrm{D_a}$ can not exceed half the link rate of $R_1 - R_2$. On the other hand, when the traffic rate from IR_1 to D_1 is low, IR_2 can send more traffic to D_1. Thus, the utilization of $\mathrm{ER}_1 - \mathrm{D}_1$ is usually higher than that of $\mathrm{ER}_2 - \mathrm{D_a}$.

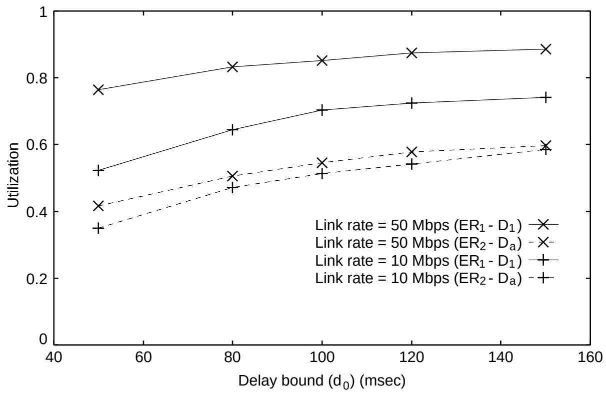

Fig. 16. Comparison of the utilizations of the links $\mathrm{ER}_1-\mathrm{D}_1$ and $\mathrm{ER}_2-\mathrm{D}_\mathrm{a}$ for different link rates with a mix of traffic patterns (exponential on-off, Pareto on-off, self-similar traffic) in the presence of high priority cross traffic

We also observe that the utilization increases as the link rate increases. We can find a reason for the improved utilization from (12). If $\sigma = 0$, then the admissible bandwidth R^* will be equal to the available bandwidth a. If we admit flows up to the rate of R^* according to the admission control algorithm of Subsection V-D, then the utilization of the tight link which has the minimum available bandwidth on the given path will be 1, since $a = C - \lambda$, where C is the link rate of the tight link and λ is the arrival rate of cross traffic at the tight link in a given time interval. If bandwidth resources are allocated up to R^* when $\sigma \neq 0$, the the utilization u can be expressed as

$$u = \frac{\lambda + R^*}{C} = 1 + \frac{\log(\varepsilon)\sigma^2}{2(d_0 - D_f)aC}. \tag{15}$$

We need to note that the second term on the right hand side of the above equation is negative due to $\log(\varepsilon)$. If we assume that the available bandwidth a is proportional to C approximately, then the second term is proportional to σ^2/a^2. The sigma/mean ratio, σ/a of the available bandwidth tends to decrease as the link rate C increases and more traffic flows are multiplexed. Thus, a decrease in the absolute value of the second term yields higher utilization as the link rate C increases. These days link speeds are high on the order of Gbps. At these high link speeds, our scheme can be effective in achieving a high utilization by (15).

VII. Conclusions

In this paper, we proposed a new admission control scheme. In the proposed scheme, admission decision is made for each flow at the ingress routers, but it is scalable because per-flow states are not managed and the admission algorithm is simple. An ingress router manages the admissible bandwidth, which is a threshold for admission control, for each relevant egress router. Since the admissible bandwidth is calculated considering the delay QoS, it is possible to guarantee the delay performance by the proposed admission control scheme. We derive an expression for a lower bound of the admissible bandwidth. We use this lower bound as an estimate for the admissible bandwidth. Since the bound is explicitly expressed in terms of delay bound (d_0), threshold for the delay violation probability (ε), fixed delay component (D_f), and the mean a and variance σ^2 of the available bandwidth, we can understand the effect of each factor on the admissible bandwidth intuitively. Using the probing scheme developed in [30], we can estimate the available bandwidth and can obtain the mean and variance from the history of the available bandwidth. In case that the probing scheme can not accurately track the available bandwidth due to too frequent and large-scale changes, a variance multiplication factor (VMF) can be used in order to compensate the variance (σ^2) estimation error.

Through simulations, we investigated the effect of VMF γ, measurement time window T, and link rate C on the performance of the proposed admission control scheme. As the delay bound (d_0) increases, more traffic flows are admitted and the variance estimation error tends to increase collaterally. Thus, larger values of VMF are required for large delay bounds. Through many simulations, we find that VMF is not needed when $d_0 - D_f$ is less than 60 msec and the VMF value of 1.5 can statistically guarantee the delay bound when $d_0 - D_f$ is less than 100 msec, where D_f is the fixed delay component of the target path. If D_f is 50 msec, then the VMF of 1.5 is enough for d_0 of up to 150 msec. As the measurement window T increases, the delay QoS is well satisfied, but the utilization is lowered since bandwidth resources are conservatively allocated during a time window of T. Thus, too large values of T are not good in terms of utilization. Since large values of T lead to conservative allocation of bandwidth resources, a proper value of T can yield a high utilization while guaranteeing delay QoS for a wide range of delay bound. Finally, we find that the proposed admission control scheme yields higher utilization at high link speeds by (15). Since the link speeds are high on the order of Gbps these days, our scheme can be effective in achieving a high utilization while guaranteeing the delay bound in real networks.

Since the proposed admission control scheme satisfies the delay performance requirements for most delay bounds and yields a rather high utilization without cooperation of core routers and per-flow state management, the proposed admission control algorithm and architecture can be a scalable solution for the delay QoS guarantee problem in IP networks.

References

[1] R. Branden, D. Clark, and S. Shenker, "Integrated services in the Internet architecture: an overview," IETF RFC 1633, June 1994.

[2] S. Blake, D. Black, M. Carlson, E. Davis, Z. Wang, and W. Weiss, "An architecture for differentiated services," IETF RFC 2475, Dec. 1998.

[3] A. Bose, M. E. Gendy, and K. G. Shin, "Sapphire: Statistical characterization and model-based adaptation of networked applications," *IEEE Trans. Parall. Distrib. Syst.*, vol. 17, no. 12, pp. 1512-1525, Dec. 2006.

[4] E. Lochin, L. Dairaine, and G. Jourjon, "gTFRC, a TCP friendly QoS-aware rate control for DiffServ assured service," *Telecommunication Systems*, vol. 33, no. 1-3, pp. 3-21, Dec. 2006.

[5] S. Jamin, P. B. Danzig, S. J. Shenker, and L. Zhang, "A measurement-based admission control algorithm for integrated services packet networks," *IEEE/ACM Trans. Netw.*, vol. 5, pp. 56-70, Feb. 1997.

[6] L. Breslau, E. W. Knightly, S. Shenker, I. Stoica, H. Zhang, "Endpoint admission control: architectural issues and performance," in *Proc. ACM SIGCOMM*, pp. 57-69, Aug. 2000.

[7] F. P. Kelly, P. B. Key, and S. Zachary, "Distributed admission control," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 12, pp. 2617-2628, Dec. 2000.

[8] S. Crosby *et al*, “Statistical properties of a near-optimal measurement-based CAC algorithm,” in *Proc. IEEE ATM’97*, pp. 103-112, May 1997.

[9] Z. Dziong, M. Juda, and L. Mason, “A framework for bandwidth management in ATM networks - aggregate equivalent bandwidth estimation approach,” *IEEE/ACM Trans. Netw.*, vol. 5, pp. 134-147, Feb. 1997.

[10] S. Floyd, “Comments on measurement-based admission control for controlled-load services,” Technical report, Lawrence Berkeley Laboratory, July 1997.

[11] R. Gibbens, and F. Kelly, “Measurement-based connection admission control,” *15th International Teletraffic Congress*, Jun. 1997.

[12] R. J. Gibbens, F. P. Kelly, and P. B. Key, “A decision-theoretic approach to call admission control in ATM networks,” *IEEE J. Sel. Areas Commun.*, vol. 13, no. 6, pp. 1101-1114, Aug. 1995.

[13] M. Grossglauser, and D. Tse, “A framework for robust measurement-based admission control,” *Computer Communications Review*, vol. 24, pp. 237-248, Oct. 1997.

[14] M. Grossglauser, and D. Tse, “A time-scale decomposition approach to measurement-based admission control,” in *Proc. IEEE INFOCOM*, vol. 3, pp. 1539-1547, Mar. 1999.

[15] M. Grossglauser, D. Tse, J. Kurose, and D. Towsley, “A new algorithm for measurement-based admission control in integrated services packet networks,” in *Proc. Fifth International Workshop on Protocols for High-Speed Networks*, Antipolis, France, Oct. 1996.

[16] S. Jamin, S. Shenker, and P. Danzig, “Comparison of measurement-based admission control algorithms for controlled-load service,” in *Proc. IEEE INFOCOM*, pp. 973-980, Apr. 1997.

[17] L. Breslau, S. Jamin, “Comments on the performance of measurement-based admission control algorithms,” in *Proc. IEEE INFOCOM*, vol. 3, pp. 1233-1242, Mar. 2000.

[18] R. Guerin, H. Ahmadi, and M. Naghshineh, “Equivalent capacity and its application to bandwidth allocation in high-speed networks,” *IEEE J. Sel. Areas Commun.*, vol. 9, no. 7, pp. 968-981, Sep. 1991.

[19] A. Elwalid, D. Mitra, R. H. Wentworth, “A new approach for allocating buffers and bandwidth to heterogeneous, regulated traffic in an ATM node,” *IEEE J. Sel. Areas Commun.*, vol. 13, no. 6, pp. 1115-1127, Aug. 1995.

[20] F. P. Kelly, “Note on effective bandwidth,” in *Stochastic Networks: Theory and Applications*. ser. Royal Statistical Society Lecture Notes, F. P. Kelly, S. Zachary, and I. B. Ziedins, Eds, London, U.K.: Oxford Univ Press, 1996, vol. 4, pp. 141-168.

[21] S. Kim, S. Y. Nam, and D. K. Sung, “Effective bandwidth for a single server queueing system with fractional Brownian input,” *Perform. Eval.*, vol. 61, no. 2-3, pp. 203-223, Jul. 2005.

[22] J. Beran, R. Sherman, M. S. Taqqu, and W. Willinger, “Long-range dependence in variable bit rate video traffic,” *IEEE Trans. Commun.*, vol. 43, no. 2/3/4, pp. 1566-1579, Feb./Mar./Apr. 1995.

[23] M. W. Garrett and W. Willinger, “Analysis, modeling and generation of self-similar VBR video traffic,” in *Proc. ACM SIGCOMM*, London, England, pp. 269-280, Aug. 1994.

[24] S. Kim, J. Y. Lee, and D. K. Sung, “A shifted gamma distribution model for long-range dependent Internet traffic,” *IEEE Commun. Lett.*, vol. 7, no. 3, pp. 124-126, Mar. 2003.

[25] N. G. Duffield and N. O’Connell, “Large deviations and overflow probabilities for the general single-server queue, with applications,” *Math. Proc. Camb. Philos. Soc.*, vol. 118, pp. 363-374, 1995.

[26] J. R. Gallardo, D. Makrakis, and M. Angulo, “Dynamic resource management considering the real behavior of aggregate traffic,” *IEEE Transactions on Multimedia*, vol. 3, no. 2, pp. 177-185, June 2001.

[27] K. Ramakrishnan and S. Floyd, “A proposal to add explicit congestion notification (ECN) to IP,” IETF RFC 2481, 1999.

[28] C. Cetinkaya, V. Kanodia, and E. W. Knightly, “Scalable services via egress admission control,” *IEEE Transactions on Multimedia*, vol. 3, no. 1, pp. 69-81, Mar. 2001.

[29] S. Y. Nam, S. Kim, and D. K. Sung, “Estimation of Available Bandwidth for an Unidentified Queueing System,” Technical report KAIST-CNR-06-01, KAIST, 2006.

[30] S. Y. Nam, S. Kim, J. Kim, and D. K. Sung, “Probing-Based Estimation of End-to-End Available Bandwidth,” *IEEE Commun. Lett.*, vol. 8, no. 6, pp. 400-402, June 2004.

[31] S. Y. Nam, “ Available bandwidth estimation and measurement-based admission control in IP networks,” *Ph.D. thesis*, Dept. EECS, KAIST, Korea, 2004.

[32] S. Karlin and H. M. Taylor, *A first course in stochastic processes, 2nd ed.*, Academic Press, 1975.

[33] One-way transmission time, ITU-T Recommendation G.114, May 2000.

[34] A. Shriram and J. Kaur, “Empirical study of the impact of sampling timescales and strategies on measurement of available bandwidth,” in *Proc. Passive and Active Measurement (PAM) Conference*, Mar. 2006.

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

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

[37] N. Hu and P. Steenkiste, “Evaluation and Characterization of Available Bandwidth Probing Techniques,” *IEEE J. Sel. Areas Commun.*, vol. 21, no. 6, pp. 879-894, Aug. 2003.

[38] R. H. Riedi, M. S. Crouse, V. J. Ribeiro, and R. G. Baraniuk, “A multifractal wavelet model with application to network traffic,” *IEEE Tran. Information Theory*, vol. 45, no. 3, pp. 992-1018, Apr. 1999.

[39] 3GPP Technical Specification 22.105, Services and service capabilities.

[40] 3GPP Technical Specification 23.107, Quality of Service (QoS) concept and architecture.

[41] R. W. Wolff, *Stochastic Modeling and the Theory of Queues*, Prentice Hall, 1989.

Seung Yeob Nam (S’97-A’04-M’05) received the BS, MS, and PhD degrees in electrical engineering from the Korea Advanced Institute of Science and Technology (KAIST), Daejon, Korea, in 1997, 1999, and 2004, respectively. From 2004 to July 2006, he was a post doc. fellow at CyLab in Carnegie Mellon University. Between Aug. 2006 and Feb. 2007, he was a post doc. fellow in the Dept. of EECS, KAIST. In Mar. 2007, he joined the Department of Information & Communication Engineering, Yeungnam University, Gyeongsan, Korea as a faculty member. His research interests include network monitoring, traffic engineering, network security, high-speed switching systems, wireless networks, etc. He received the Best Paper Award from the APCC 2000 Conference and Bronze prize from 2004 Samsung Humantech paper contest.

Sunggon Kim received B.S., M.S., and Ph. D. degrees in Mathematics from POSTECH(Pohang University of Science and Technology), Pohang, Republic of Korea, in 1996, 1998 and 2001, respectively. From 2001 to 2004, he was a post doc. fellow and a research professor at Department of Electrical Engineering and Computer Science, KAIST, Deajeon. In Dec. 2004, he joined the Department of Information Statistics, Gyeongsang National University, Jinju as an Assistant Professor. His research interests are including stochastic processes, queueing theory, and its application on communication networks. He received the Young Statistician Award from the Korean Statistical Society in 2003 and Silver prize from 2003 Samsung Humantech paper contest, and Bronze prize at 2004.

Dan Keun Sung (S’80-M’86-SM’00) received the B.S. degree in Electronics Engineering from Seoul National University in 1975 and the M.S. and Ph.D. degrees in Electrical & Computer Engineering from the University of Texas at Austin, in 1982 and 1986, respectively. In 1986 he joined the faculty of KAIST where he is currently Professor at the department of Electrical Engineering and Computer Science. He was Director of the Satellite Technology Research Center (SaTReC) of KAIST from 1996 to 1999. He is Editor of IEEE Communication Magazine. He is also Division Editor of the Journal of Communications and Networks. His research interests include mobile communication systems & networks, high speed networks, next generation IP based networks, traffic control in wireless & wireline networks, signaling networks, intelligent networks, performance & reliability of communication systems, and microsatellites. He received National Order of Merits, Dongbaek Medal in 1992, Research Achievement Award in 1997, MoMuc Paper Award in 1997, Academic Excellent Award in 2000, Best Paper Award in APCC2000, and This Month’s Scientist Award by MOST and KOSEF in 2004. He is a Senior Member of IEEE and a member of the National Academy of Engineering of Korea.