University of Würzburg
Institute of Computer Science
Research Report Series

Experience-Based Admission Control with Type-Specific Overbooking

Jens Milbrandt, Michael Menth, Jan Junker

Report No. 382 August 2006

University of Wurzburg
Institute of Computer Science
Chair of Distributed Systems
Wurzburg, Germany
{milbrandt,menth,junker}@informatik.uni-wuerzburg.de

Experience-Based Admission Control with Type-Specific Overbooking

Jens Milbrandt, Michael Menth, Jan Junker
University of Wurzburg
Institute of Computer Science
Chair of Distributed Systems
Wurzburg, Germany
{milbrandt,menth,junker}@informatik.uni-wuerzburg.de

Abstract

Experience-based admission control (EBAC) is a hybrid approach combining the classical parameter-based and measurement-based admission control schemes. EBAC calculates an appropriate overbooking factor used to overbook link capacities with resource reservations in packet-based networks. This overbooking factor correlates with the average peak-to-mean rate ratio of all admitted traffic flows on the link. So far, a single overbooking factor is calculated for the entire traffic aggregate. In this paper, we propose type-specific EBAC which provides a compound overbooking factor considering different types of traffic that subsume flows with similar peak-to-mean rate ratios. The concept can be well implemented since it does not require type-specific traffic measurements. We give a proof of concept for this extension and compare it with the conventional EBAC approach. We show that EBAC with type-specific overbooking leads to better resource utilization under normal conditions and to faster response times for changing traffic mixes.

1 Introduction

Admission control (AC) may be used to ensure quality of service (QoS) in terms of packet loss and delay in packet-based communication networks. Many different approaches for AC exist and an overview can be found in [1]. In general, AC admits or rejects resource reservation requests and installs reservations for admitted flows. The packets of admitted flows are transported with high priority such that they get the desired QoS. Rejected flows are either blocked or their packets are handled only with lower priority.

Link admission control (LAC) methods protect a single link against traffic overload. They can be further subdivided into parameter-based AC (PBAC), measurement-based AC (MBAC), and derivatives thereof. PBAC methods [2–4] use traffic descriptors to

Parts of this work were funded by the Bundesministerium für Bildung und Forschung of the Federal Republic of Germany (Förderkennzeichen 01AK045) and Siemens AG, Munich, Germany. The authors alone are responsible for the content of the paper.

calculate a priori the expected bandwidth consumptions of admitted flows to get an estimate of the remaining free capacity which is required for future admission decisions. PBAC offers stringent QoS guarantees to data traffic that has been admitted to the network but it lacks scalability with regard to the signalling of resource reservations. In contrast, there are numerous measurement-based AC (MBAC) approaches which use real-time measurements to assess the remaining free capacity [5–13]. MBAC uses the available network resources very efficiently but relies on real-time traffic measurements and, therefore, it is susceptible to QoS violation.

Experience-based admission control (EBAC) is a hybrid solution [14]. It uses peak rate allocation based on traffic descriptors and calculates a factor to overbook a given link capacity. The calculation of this overbooking factor is based on the statistics of the utilization of past reservations that are obtained by measurements. Hence, EBAC does not require real-time measurements of the instantaneous traffic for admission decisions and is, therefore, substantially different from classical MBAC approaches and easier to implement. The major task of EBAC is the calculation of an appropriate overbooking factor for classical PBAC. This factor is obtained by measurements and correlates with the average peak-to-mean rate ratio (PMRR) of all admitted flows which only indicate their peak rate. In previous work, we have provided a proof of concept for EBAC [15]. We also investigated its robustness during sudden changes of the traffic properties to which all MBAC methods are susceptible [16]. So far, a single overbooking factor is calculated based on the traffic characteristics of the entire admitted traffic aggregate. This paper extends EBAC towards type-specific overbooking (TSOB) which provides a compound overbooking factor considering different types of traffic. The extension can be well implemented since it does not require type-specific measurements. We give a proof of concept for EBAC with TSOB and compare it with the conventional EBAC approach. We show that EBAC with TSOB leads to better resource utilization under normal traffic conditions and to faster response times in case of changing traffic mixes. Unlike conventional EBAC, the extension avoids congestion due to overreservation if the fraction of flows with low PMRR increases in the traffic mix, i.e. if the traffic intensity increases due to a changing composition of the admitted traffic mix.

All of the above sketched AC mechanisms apply for a single link but they can be extended on a link-by-link basis for a network-wide application. For the sake of clarity, we limit our performance study to a single link which can be done without loss of generality since the core concept of EBAC applies to any given single resource.

This paper is structured as follows. In Section 2, we briefly review the EBAC concept. Section 3 describes our simulation design and the applied traffic model and summarizes results from previous studies. Section 4 proposes the extension of EBAC towards type-specifc overbooking (TSOB). The simulation results in Section 5 show the superiority of EBAC with TSOB over conventional EBAC. Finally, Section 6 summarizes this work and points out further steps towards the application of type-specific overbooking in practice.

2 Experience-Based Admission Control (EBAC)

In this section, we briefly review the EBAC concept with emphasis on the EBAC memory which implements the experience based on which AC decisions are made.

An AC entity limits the access to a link l with capacity $c(l)$ and records all admitted flows $f \in \mathcal{F}(t)$ at any time t together with their requested peak rates $\{r(f) : f \in \mathcal{F}(t)\}$. When a new flow f_{new} arrives, it requests a reservation for its peak rate $r(f_{new})$. If

$$r(f_{new}) + \sum_{f \in \mathcal{F}(t)} r(f) \;\leq\; c(l) \cdot \varphi(t) \cdot \rho_{max} \tag{1}$$

holds, admission is granted and f_{new} joins $\mathcal{F}(t)$. If flows terminate, they are removed from $\mathcal{F}(t)$. The experience-based overbooking factor $\varphi(t)$ is calculated by statistical analysis and indicates how much more bandwidth than $c(l)$ can be safely allocated for reservations. The maximum link utilization threshold ρ_{max} limits the traffic admission such that the expected packet delay W exceeds an upper delay threshold W_{max} only with probability p_W. We calculate the threshold ρ_{max} based on the $N \cdot D/D/1 - \infty$ approach [17].

For the computation of the overbooking factor $\varphi(t)$, we define the reserved bandwidth of all flows as $R(t) = \sum_{f \in \mathcal{F}(t)} r(f)$. EBAC performs traffic measurements $M(t)$ on the link and collects a time statistic for the reservation utilization $U(t) = M(t)/R(t)$. The value $U_p(t)$ denotes the p_u-percentile of the empirical distribution of U and the reciprocal of this percentile is the overbooking factor $\varphi(t) = 1/U_p(t)$.

The EBAC system requires a set of functional components to calculate the overbooking factor $\varphi(t)$:

1. **Measurement Process for** $M(t)$ — To obtain $M(t)$, we use disjoint interval measurements such that for a time interval I_i with length Δ_i, the measured rate $M_i = \Gamma_i/\Delta_i$ is determined by metering the traffic volume Γ_i sent during I_i.

2. **Statistic Collection** $P(t,U)$ — For the values $R(t)$ and $M(t)$, a time statistic for the reservation utilization $U(t) = M(t)/R(t)$ is collected. The values $U(t)$ are sampled in constant time intervals and are stored as hits in bins for a time-dependent histogram $P(t,U)$. From this histogram, the time-dependent p_u-percentile $U_p(t)$ of the empirical distribution of U can be derived as
$$U_p(t) = \min_u \{u : P(t, U \leq u) \geq p_u\}. \tag{2}$$

3. **Statistic Aging Process for** $P(t,U)$ — If traffic characteristics change over time, the reservation utilization statistic must forget obsolete data to reflect the properties of the new traffic mix. Therefore, we record new samples of $U(t)$ by incrementing the corresponding histogram bins by one and devaluate the contents of all histogram bins in regular devaluation intervals I_d by a constant devaluation factor f_d. The devaluation process determines the memory of EBAC which is defined next.

4. **Memory of EBAC** — The histogram $P(t, U)$, i.e. the collection and the aging of statistical AC data, implements the memory of EBAC. This memory correlates successive flow admission decisions and consequently influences the adaptation of the overbooking factor $\varphi(t)$ to changing traffic conditions on the link. The statistic aging process, characterized by the devaluation interval I_d and the devaluation factor f_d, makes this memory forget about reservation utilizations in the past. The parameter pairs (I_d, f_d) yield typical half-life periods T_H after which collected values $U(t)$ have lost half of their importance in the histogram. Therefore, we have $\frac{1}{2} = f_d^{T_H/I_d}$ and define the EBAC memory based on its half-life period

$$T_H(I_d, f_d) = I_d \cdot \frac{-ln(2)}{ln(f_d)}. \tag{3}$$

3 EBAC Performance Simulation

In this section, we first present the simulation design of EBAC on a single link and the traffic model we used on the flow and packet scale level. Afterwards, we summarize recent EBAC simulation results from [15, 16].

3.1 Simulation Design

The design of our simulation is shown in Figure 1. Different types of traffic *source generators* produce flow requests that are admitted or rejected by the *admission control* entity. To make an admission decision, this entity takes the overbooking factor $\varphi(t)$ into account. In turn, it provides information regarding the reservations $R(t)$ to the *EBAC system* and yields flow blocking prababilities $p_b(t)$. For each admitted source, a *traffic generator* is instantiated to produce a packet flow that is shaped to its contractually defined peak rate. Traffic flows leaving the *traffic shapers* are then multiplexed on the buffered link with capacity $c(l)$. The link provides information regarding the measured traffic $M(t)$ to the EBAC system and yields packet delay probabilities $p_d(t)$ and packet loss probabilities $p_l(t)$. Another measure for the performance of EBAC is the overall response time T_R, i.e., the time-span required by the EBAC system to adapt the overbooking factor to a new traffic situation. The time T_R depends on the transient behavior of EBAC and is investigated in [16].

3.2 Traffic Model

In our simulations, the traffic controlled by EBAC is modelled on a flow scale level and a packet scale level. While the flow level controls the inter-arrival times of flow requests and the holding times of admitted flows, the packet level defines the inter-arrival times and the sizes of packets within a single flow.

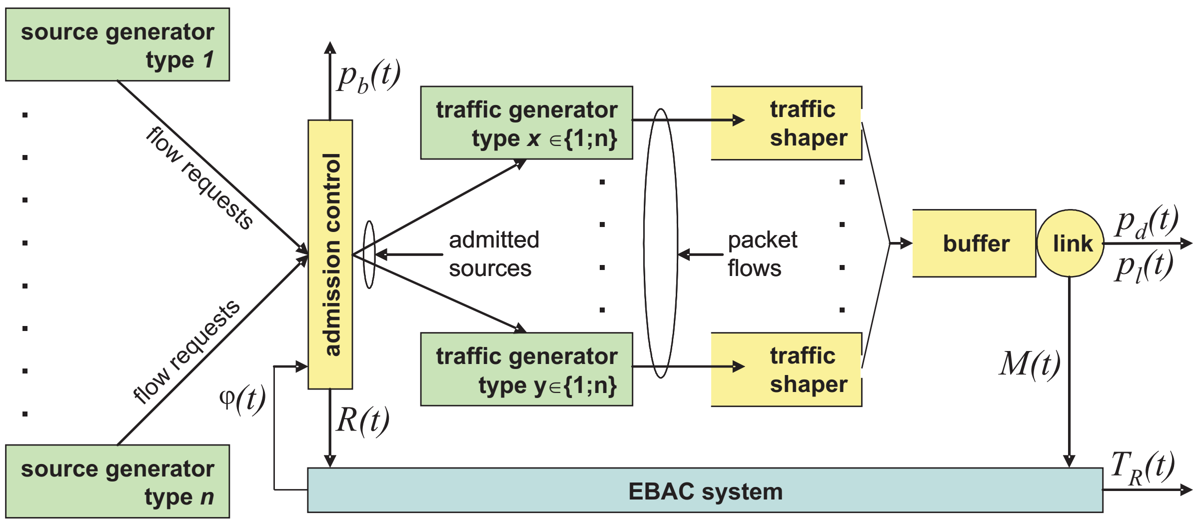


Figure 1: Simulation design for EBAC in steady and transient state.

3.2.1 Flow Level Model

On the flow level, we distinguish different traffic source types, each associated with a characteristic peak-to-mean rate ratio and corresponding to a source generator type in Figure 1. The inter-arrival time of flow requests and the holding time of admitted flows both follow a Poisson model [18], i.e., new flows arrive with rate λ_f and the duration of a flow is controlled by rate μ_f. The mean of the flow inter-arrival time is thus denoted by $1/\lambda_f$ and the holding time of a flow is exponentially distributed with a mean of $1/\mu_f$. Provided that no blocking occurs, the overall offered load $a_f = \lambda_f/\mu_f$ is the average number of simultaneously active flows measured in Erlang. To saturate an EBAC-controlled link with traffic, the load is set to $a_f \geq 1.0$. The latter assumption allows for an investigation of the EBAC performance under heavy traffic load such that some flow requests are rejected.

3.2.2 Packet Level Model

On the packet level, we abstract from the wide diversity of packet characteristics induced by the application of different transmission layer protocols. Since we are interested in the basic understanding of the behavior of EBAC, we abstain from real traffic patterns and define a flow of consecutive data packets simply by a packet size and a packet inter-arrival time distribution. Both contribute to the rate variability within a flow that is produced by a traffic generator in Figure 1. To keep things simple, we assume a fixed packet size per flow and use a Poisson arrival process to model a packet inter-arrival time ditribution with rate λ_p. We are aware of the fact that Poisson is not a suitable model to simulate Internet traffic on the packet level [19]. Therefore, we generate Poisson packet streams and subsequentially police the resulting flows with peak-rate traffic shapers (cf. Figure 1). The properties of the flows are significantly influenced by the configuration of these shapers.

In practice, applications know and signal the peak rates of their corresponding traffic flows. The type of an application can be determined, e.g., by a signalling protocol number. We use only this limited information in our simulations, i.e., the mean rates $c(f)$ of the flows are not known to the EBAC measurement process, they are just model parameters for the traffic generation. Therefore, we can control the rate of flow f by its peak-to-mean rate ratio $k = r(f)/c(f)$.

3.3 Simulation Studies of Conventional EBAC

3.3.1 EBAC Performance for Constant Traffic

The intrinsic idea of EBAC is the exploitation of the peak-to-mean rate ratio (PMRR) $K(t)$ of the traffic aggregate admitted to the link. In [15], we simulated EBAC on a single link with regard to its behavior in steady state, i.e., when the properties of the traffic aggregate were rather static. These simulations provided a first proof of concept for EBAC. We showed for different PMRRs that EBAC achieves a high degree of resource utilization through overbooking while packet loss and packet delay are well limited. The simulation results allowed us to give recommendations for the EBAC parameters such as measurement interval length and reservation utilization percentile to obtain appropriate overbooking factors $\varphi(t)$. They furthermore showed that the EBAC mechanism is robust against traffic variability in terms of packet size and inter-arrival time distribution as well as against correlations thereof.

3.3.2 EBAC in the Presence of Traffic Changes

In [16], we investigated the transient behavior of conventional EBAC after sudden traffic changes with special regard to the EBAC response time $T_R(t)$ and the QoS performance in terms of packet loss $p_l(t)$ and packet dealy $p_d(t)$ (cf. Figure 1) which are potentially compromised in case of traffic increases. As EBAC partly relies on traffic measurements, it is susceptible to changes of the traffic characteristics of admitted flows. There are certain influencing parameters coupled with this problem. One of them is the length of the EBAC memory which has been defined by its half-life period T_H. We tested the impact of the EBAC memory on a sudden decrease and increase of the traffic intensity expressed by changes of the PMRRs of the simulated traffic flows. We showed that, for a changing traffic intensity, the response time T_R required to adapt the overbooking factor to the new traffic situation depends linearly on the half-life period T_H. For decreasing traffic intensity, the QoS of the traffic is not at risk. For a suddenly increasing traffic intensity, however, it is compromised for a certain time span T_R^Q. The corresponding experiments used an unlimited link buffer and investigated the performance of EBAC under very extreme traffic conditions that correspond to a collaborative and simultaneous QoS attack by all traffic sources.

4 EBAC with Type-Specific Overbooking

In this section, we present type-specific overbooking (TSOB) as a concept extending EBAC. So far, we only consider the traffic characteristics of the entire aggregate of admitted traffic flows and calculate a single factor to overbook the link capacity. We now include additional information about the characteristics of individual traffic types and their share in the currently admitted traffic mix to calculate a compound type-specific overbooking factor. First, we describe the system extension and then we show how the compound overbooking factor for EBAC with TSOB can be estimated without type-specific traffic measurements. Finally, we present some simulation results showing the advantage of EBAC with TSOB over conventional EBAC.

4.1 EBAC System Extension

We assume that different applications produce traffic flows with typical peak-to-mean rate ratios (PMRRs) $K_i(t)$ which lead to different type-specific overbooking factors $\varphi_i(t)$. Parameter i then denotes a traffic type subsuming flows of different applications but with similar PMRRs K_i. The EBAC admission decision for a new flow f_i^{new} of type i is then extended to

$$r(f_i^{new}) \cdot U_{p,i}(t) + \sum_{f \in \mathcal{F}(t)} r(f) \cdot U_{p,type(f)}(t) \leq c(l) \cdot \rho_{max}. \tag{4}$$

In general, the aggregate $\mathcal{F}(t)$ is composed of flows of different traffic types i for which the PMRRs K_i remain rather constant over time. For admission, each flow is supposed to register at the AC entity with its peak rate and its traffic type. This yields type-specific reservations $R_i(t)$ for which $\sum_{i=0}^{n} R_i(t) = R(t)$ holds. On arrival of a new flow f_i^{new}, $R_i(t)$ is increased by the peak rate $r(f_i^{new})$ of the flow and it is decreased by the same rate when the flow terminates. The value $\alpha_i(t) = R_i(t)/R(t)$ then reflects the share of a traffic type i in the mix and the entire traffic composition consisting of n different traffic types is given by vector

$$\alpha(t) = \begin{pmatrix} \alpha_1(t) \\ \vdots \\ \alpha_n(t) \end{pmatrix}, \ \sum_{i=1}^{n} \alpha_i(t) = 1. \tag{5}$$

EBAC with TSOB uses the information about the PMRRs K_i and the time-dependent traffic composition $\alpha(t)$ to estimate type-specific reservation utilizations $U_i(t)$. The estimation of the reservation utilizations $U_i(t)$ is a rather complex task and described in Sec. 4.2. The values $U_i(t)$ are stored as hits in bins of separate histograms $P_i(t, U)$ which yield type-specific reservation utilization percentiles $U_{p,i}(t)$. We weight these percentiles by their corresponding shares $\alpha_i(t)$ and finally calculate the compound overbooking factor for EBAC with TSOB as

$$\varphi(t) = \frac{1}{\sum_i \alpha_i(t) \cdot U_{p,i}(t)}. \tag{6}$$

4.2 Estimation of Type-Specific Reservation Utilizations

A crucial issue for the performance of EBAC with TSOB is the estimation of the type-specific reservation utilizations $U_i(t)$. Making type-specific measurements $M_i(t)$ yields exact values for $U_i(t) = M_i(t)/R_i(t)$. For a reduced number of traffic classes, type-specific measurements seem feasible if we consider new network technologies such as differentiated services (DiffServ) [20] for traffic differentiation and multi protocol label switching (MPLS) [21] for the collection of traffic statistics. However, current routers mostly do not provide these type-specific traffic measurements and, therefore, we have to use the available parameters $M(t)$, $R(t)$, $R_i(t)$, and $\alpha(t)$ to estimate the $U_i(t)$. In the following, we develop two methods to obtain estimates for the type-specific reservation utilizations.

4.2.1 Estimation with Linear Equation Systems (LES)

The first simple method trys to calculate the type-specific reservation utilizations $U_i(t)$ exactly and uses the equation $U(t) = \sum_i \alpha_i(t) \cdot U_i(t)$ which leads to a linear equation system (LES) of the form

$$\begin{pmatrix} U(t_{j-n}) \\ \vdots \\ U(t_j) \end{pmatrix} = \begin{pmatrix} \alpha_1(t_{j-n}) \dots \alpha_n(t_{j-n}) \\ \vdots \qquad\qquad \vdots \\ \alpha_1(t_j) \quad \dots \quad \alpha_n(t_j) \end{pmatrix} \begin{pmatrix} U_1(t_j) \\ \vdots \\ U_n(t_j) \end{pmatrix} \tag{7}$$

where n is the number of traffic types and j denotes a time index. A unique solution of this LES requires probes of $U(t)$ and $\alpha(t)$ in the interval $[t_{j-n}, t_j]$ and n linearly independent matrix column vectors $(\alpha_i(t_{j-x}))_{1 \leq i \leq n}$. We calculate a new solution of the LES every time the vector $\alpha(t)$ changes significantly, i.e., $\exists k : \frac{|a_k(t_i) - a_k(t_{i-1})|}{a_k(t_i)} > \epsilon$. The problem of estimating the type-specific reservation utilization with the LES method is that the linear independence of the matrix columns in Equation (7) is not guaranteed at any time t_j when the traffic composition changes. In this case, a unique solution for the equation system does not exist and the values $U_i(t_j)$ cannot be included in the histogram $P_i(t, U)$. Then, we simply keep the resveration utilizations $U_i(t_{j-x})$ of the last feasible LES until a new linearly independent LES is found.

4.2.2 Estimation with Least Squares Approximation (LSA)

The second approach is more complex and estimates the type-specific reservation utilizations based on a least squares approximation (LSA, cf. e.g. [22]) of the values $U_i(t)$. For the ease of understanding, we illustrate this method, without loss of generality, with two different traffic types $i \in \{1, 2\}$. $U_1(t)$ and $U_2(t)$ denote their type-specific reservation utilizations . The global reservation utilization is then $U(t) = \alpha_1(t) \cdot U_1(t) + \alpha_2(t) \cdot U_2(t)$ and with $\alpha_1(t) + \alpha_2(t) = 1$, we get

$$U(t) = \alpha_1(t) \cdot (U_1(t) - U_2(t)) + U_2(t). \tag{8}$$

We substitute $a_j = U_1(t_j) - U_2(t_j)$ and $b_j = U_2(t_j)$ and obtain the least squares error for parameters $U_1(t)$ and $U_2(t)$ if we minimize the term

$$\mathcal{L} = \min_{a_m, b_m} \sum_{j=1}^{m} \left[U(t_j) - (\alpha_1(t_j) \cdot a_m + b_m)\right]^2. \tag{9}$$

The time index j thereby covers all values $U(t_j)$ and $\alpha(t_j)$ from the first ($j = 1$) to the last ($j = m$) probe ever determined by the EBAC system. We find the minimum of $\mathcal{L}$ where the first derivatives of Equation (9) yield zero, i.e., we set $\frac{\partial \mathcal{L}}{\partial a} \stackrel{!}{=} 0$ und $\frac{\partial \mathcal{L}}{\partial b} \stackrel{!}{=} 0$ and resolve these equations to parameters a_m and b_m which yields

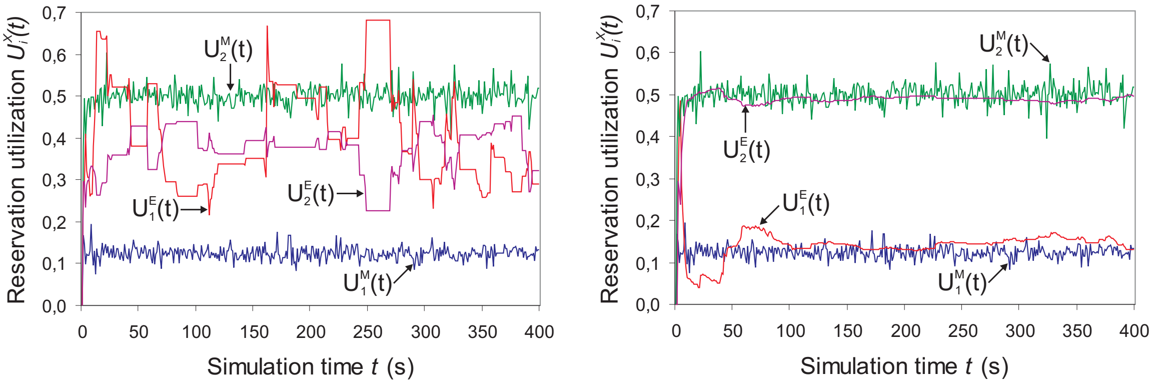
$$a_m = \frac{m \cdot \sum_j \alpha_1(t_j) U(t_j) - \sum_j \alpha_1(t_j) \cdot \sum_j U(t_j)}{m \cdot \sum_j \alpha_1(t_j)^2 - \left(\sum_j \alpha_1(t_j)\right)^2} \tag{10a}$$

$$b_m = \frac{\sum_j U(t_j) \cdot \sum_j \alpha_1(t_j)^2 - \sum_j \alpha_1(t_j) \cdot \sum_j \alpha_1(t_j) U(t_j)}{m \cdot \sum_j \alpha_1(t_j)^2 - \left(\sum_j \alpha_1(t_j)\right)^2} \tag{10b}$$

for $1 \leq j \leq m$. The sums in Equations (10a) and (10b) can be computed iteratively which helps to cope with the large set of instances observed over all times t_j. In addition, we apply the time exponentially weighted moving average (TEWMA) algorithm to these sums to blind out short-time fluctuations. Due to the lack of space, we omit any details of the TEWMA algorithm which is described in [23]. With the calculated parameters a_m and b_m, we finally obtain the type-specific reservation utilizations $U_1(t_m) = a_m + b_m$ and $U_2(t_m) = b_m$.

4.2.3 Comparison of Measured and Estimated Type-Specific Reservation Utilizations

We perform simulations with both methods estimating the type-specific reservation utilizations. Figure 2 shows a comparison of the measured type-specific reservation utilizations $U_i^M(t)$ and their corresponding estimates where $U_i^{LES}(t)$ is the estimate achieved with the LES method and $U_i^{LSA}(t)$ is the estimate obtained by the LSA method. The underlying simulation contains two traffic types $i \in \{1, 2\}$ of traffic. Type 1 has a PMRR $K_1 = 2$ and a mean share of $\alpha_1 = 0.2$ in the traffic mix. Type 2 has a PMRR $K_2 = 8$ and a mean share $\alpha_2 = 0.8$. All values K_i and α_i are averages. The type-specific reservation utilizations are determined every second. On the packet level, we have Poisson distributed inter-arrival times which lead to short-time fluctuations for the measured values $U_i^M(t)$. These fluctuations are clearly damped by the TEWMA algorithm used for the estimated values $U_i^{LES}(t)$ and $U_i^{LSA}(t)$. Obvioulsy, the simple LES method is not feasible for estimation since the resulting estimates deviate strongly from the exact measurements. In contrast, the LSA method provides good estimates for the corresponding measured values after some time. Hence, this approach enables EBAC with TSOB without type-specific traffic measurements.



(a) Estimation with Linear Equation Systems (b) Estimation with Least Squares Approximation

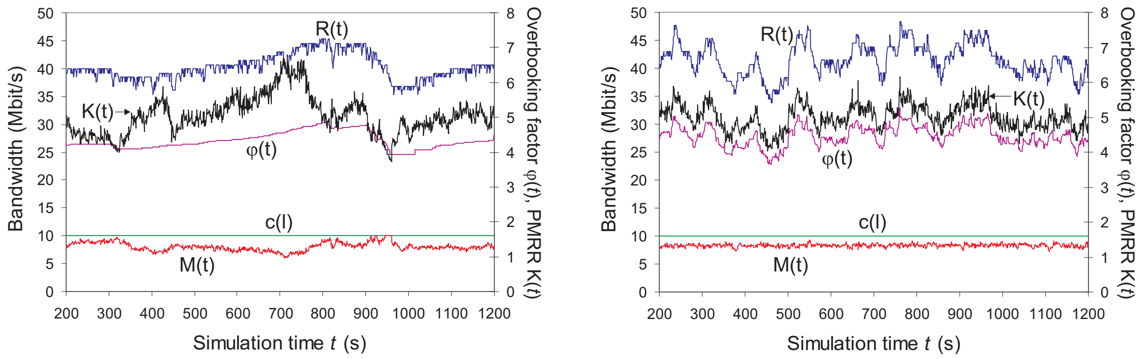
Figure 2: Comparison of measured and estimated type-specific reservation utilizations.

5 Performance Comparison of Conventional EBAC and EBAC with TSOB

To investigate EBAC with TSOB, we perform a number of simulations each associated with a different traffic situation. For all simulations, we use a link capacity $c(l) = 10$ Mbit/s and simulate with two traffic types $i \in \{1, 2\}$ with characteristic peak-to-mean rate ratios (PMRRs) $K_1 = 2$ and $K_2 = 8$. A flow f_i of any type i reserves bandwidth with a peak rate $r(f_i) = 768$ Kbit/s and has a mean holding time of $1/\mu_f = 90$ s. The mean interarrival time of flow requests is set to $1/\lambda_f = 750$ ms such that the link is saturated with traffic, i.e., some flow requests are rejected. For conventional EBAC we use the overbooking factor according to Section 2 and for EBAC with TSOB, we calculate it according to Equation (6). First, we investigate EBAC with TSOB for a rather constant traffic mix and then we study its behavior for a suddenly changing traffic composition $\alpha(t)$.

5.1 Simulation with Constant Traffic Mix

The first experiment simulates traffic with rather constant traffic shares $\alpha_i(t)$, i.e., the composition of the traffic mix remains constant except for statistical fluctuations. The results of a single simulation run are shown in Figure 3a for conventional EBAC and for EBAC with TSOB in Figure 3b. The mean shares of the traffic types in the mix are set to $\alpha_1 = 0.2$ and $\alpha_2 = 0.8$. We repeated this experiments 50 times to obtain reliable confidence intervals which proved to be very small. However, the illustration of a single simulation run shows more clearly the advantage of EBAC with TSOB over conventional EBAC. For EBAC with TSOB (cf. Figure 3b), the decreases of the PMRR $K(t)$ due to statistical fluctuations of $\alpha(t)$ lead to a significant decrease of the overbooking factor $\varphi(t)$. The increases of $K(t)$ due to changing $\alpha(t)$ lead to a significant increase of $\varphi(t)$. For conventional EBAC (cf. Figure 3a), these changes happen rather slow and, therefore,



(a) Conventional EBAC

(b) EBAC with TSOB

Figure 3: Overbooking performance of conventional EBAC vs. EBAC with TSOB for a constant traffic mix.

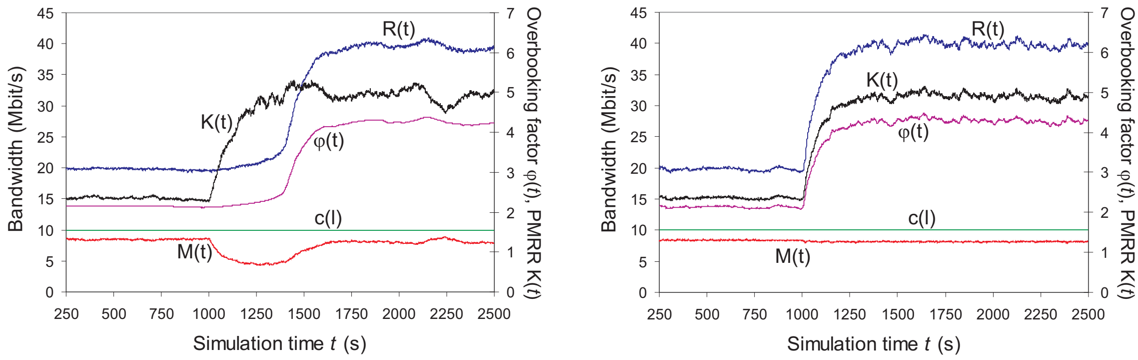
the QoS may be at risk or the link capacity is underutilized. In contrast, EBAC with TSOB adjusts its compound overbooking factor very quickly to the modifications of $\alpha(t)$ and, therefore, it is able to keep the measured rate $M(t)$ on a clearly more stable level than conventional EBAC. This, in turn, leads to a better QoS and to an improved resource utilization.

5.2 Simulation with Changing Traffic Mix

In the following two simulation experiments, we focus on the reaction of EBAC with TSOB after a decrease or an increase of the traffic intensity. We consider sudden changes of the traffic composition $\alpha(t)$ to have worst case scenarios and to obtain upper bounds on the EBAC response times.

5.2.1 Simulation with Decreasing Traffic Intensity

We investigate the change of the traffic intensity from a high to a low value. Figure 4 shows the average results over 50 simulation runs. We use the same two traffic types with their characteristic PMRRs as before. However, we start with mean traffic shares $\alpha_1 = 0.8$ and $\alpha_2 = 0.2$. At simulation time $t_0 = 1000$ s, the mean shares of both traffic types are swapped to $\alpha_1 = 0.2$ and $\alpha_2 = 0.8$ by changing the type-specific request arrival rates, i.e., the traffic intensity of the entire aggregate decreases due to a change in the traffic mix $\alpha(t)$. This leads to a sudden increase of the PMRR $K(t)$ which results in an immediate decrease of the measured traffic $M(t)$ for conventional EBAC (cf. Figure 4a). With observable delay, the conventional EBAC system adapts its overbooking factor $\varphi(t)$ as a result of the slowly decreasing p_u-percentile $U_p(t)$ in the histogram $P(t, U)$. From other simulations [16] we know that this delay strongly depends on the EBAC memory defined by the half-life period T_H in Equation (3). In contrast, EBAC with TSOB (cf. Figure 4b) increases its overbooking factor $\varphi(t)$ almost at once since the p_u-percentiles



(a) Conventional EBAC

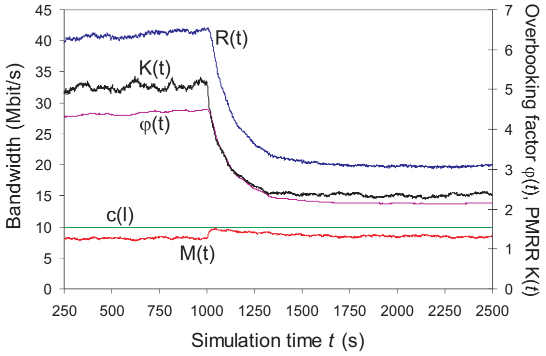
(b) EBAC with TSOB

Figure 4: Overbooking performance of conventional EBAC vs. EBAC with TSOB during a traffic intensity decrease.

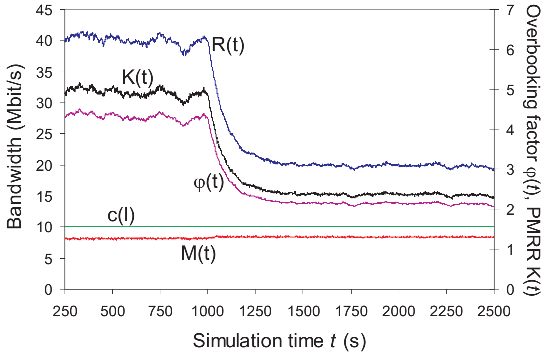
of the type-specific histograms $P_i(t, U)$ remain rather constant. As only the shares of the traffic types in the mix have changed, the compound $\varphi(t)$ is immediately adapted. As a consequence, the faster reaction of EBAC with TSOB leads to a higher and more stable mean link utilization.

5.2.2 Simulation with Increasing Traffic Intensity

Now, we change the traffic intensity from a low to a high value which leads to a decrease of the PMRR $K(t)$ of the traffic aggregate. The simulation results are shown in Figure 5. Using the same two traffic types as before, we start with mean traffic shares $\alpha_1 = 0.2$ and $\alpha_2 = 0.8$ and swap them at simulation time $t_0 = 1000$ s to $\alpha_1 = 0.8$ and $\alpha_2 = 0.2$ by changing the type-specific request arrival rates. This increases the traffic intensity of the aggregate due to a change in the traffic mix $\alpha(t)$. In this simulation experiment, the QoS is at risk because flows with low traffic intensity are successively replaced by flows with high intensity and, therefore, the load on the link is rising. Conventional EBAC (cf. Figure 5a) reacts again slower than EBAC with TSOB (cf. Figure 5b) although this time, their speed of adapting the overbooking factor differs less. From other simulations [16] we know that the response time of conventional EBAC is independent of the EBAC memory in case of a sudden traffic increase. Our simluation results show that conventional EBAC yields a slightly higher link utilization compared to EBAC with TSOB. However, this high utilization comes at the expense of a violation of QoS guarantees as the measured traffic $M(t)$ consumes the entire link capacity $c(l)$ for a short period of time (cf. Figure 5a). As a consequence, the packet delay probability $p_d = P(\text{Packet delay} \geq 50 \text{ ms})$ rises from $p_d = 0$ for EBAC with TSOB to a maximum of $p_d \approx 0.3$ for conventional EBAC. This obviously favours the extension of EBAC towards type-specific overbooking.



(a) Conventional EBAC



(b) EBAC with TSOB

Figure 5: Overbooking performance of conventional EBAC vs. EBAC with TSOB during a traffic intensity increase.

6 Conclusion

We reviewed the concept of experience-based admission control (EBAC) and summarized previous work regarding its robustness and adaptivity. EBAC overbooks the capacity of a single link with reservations according to the average peak-to-mean rate ratio of all admitted flows if the reservations are made based on signaled peak rates. The contribution of this paper is the extension of EBAC to use a compound type-specific overbooking factor for different traffic types subsuming flows with similar peak-to-mean rate ratios. The major challenge is the calculation of the type-specific reservation utilizations required for the compound overbooking factor. In general, the traffic cannot be measured type-specific and, as a consquence, the type-specific reservation utilizations cannot be obtained directly. Therefore, we proposed two mechanisms that derive the type-specific reservation utilizations depending on the reservation utilization of the entire traffic aggregate and the reserved rates of the type-specific aggregate shares. One simple approach is based on linear equation systems while the other one uses least squares approximation to calculate the type-specific reservation utilizations. Our simulation results revealed that only the second method is able to derive them with sufficiently high accuracy. We showed that EBAC with type-specific overbooking leads to larger resource utilization for stationary traffic mixes since more traffic can be savely admitted. We also simulated sudden changes of the traffic mix such that the share of flows with highly utilized reservations suddenly decreases or increases. If the share of these flows decreases, EBAC with type-specific overbooking (TSOB) adapts faster than conventional EBAC which leads to a significantly better resource utilization during the adaptation phase. If the share of these flows decreases, the advantage of EBAC with TSOB over conventional EBAC becomes even more obvious. While EBAC with TSOB can avoid overload situations, conventional EBAC has no appropriate means to prevent it.

This paper provided a proof of concept for EBAC with TSOB but many technical details must be clarified before it can be deployed in practice. E.g., a reliable network-wide measurement system needs to be installed, an appropriate number of different traffic types for type-specific overbooking must be found, and applications with similar peak-to-mean rate ratios have to be identified and classified. Certainly, these issues must be solved. However, we already had a prototype of the conventional EBAC running in a real network testbed which showed the ability to implement the concept.

References

[1] M. Menth, *Efficient Admission Control and Routing for Resilient Communication Networks*. PhD thesis, University of Würzburg, July 2004.

[2] M. Fidler and V. Sander, "A Parameter Based Admission Control for Differentiated Services Networks," *Computer Networks*, vol. 44, no. 4, pp. 463–479, 2004.

[3] J. L. Boudec, "Application of Network Calculus to Guaranteed Service Networks," *IEEE Transactions on Information Theory*, vol. 44, May 1998.

[4] J. Wroclawski, "RFC 2210: The Use of RSVP with IETF Integrated Services," September 1997.

[5] J. Qiu and E. Knightly, "Measurement-Based Admission Control with Aggregate Traffic Envelopes," *IEEE Transactions on Networking*, vol. 9, no. 2, pp. 199–210, 2001.

[6] I. Más and G. Karlsson, "PBAC: Probe-Based Admission Control," in 2^{nd} *International Workshop on Quality of future Internet Services (QofIS)*, September 2001.

[7] M. Mandjes and M. van Uitert, "Transient Analysis of Traffic Generated by Bursty Sources, and its Application to Measurement-Based Admission Control.," *Telecommunication Systems*, vol. 15, no. 3-4, pp. 295–321, 2000.

[8] L. Breslau, S. Jamin, and S. Shenker, "Comments on the Performance of Measurement-Based Admission Control Algorithms," in *IEEE Conference on Computer Communications (INFOCOM)*, pp. 1233–1242, March 2000.

[9] C. Cetinkaya and E. Knightly, "Egress Admission Control," in *IEEE Conference on Computer Communications (INFOCOM)*, pp. 1471–1480, March 2000.

[10] V. Elek, G. Karlsson, and R. Rönngren, "Admission Control Based on End-to-End Measurements," in *IEEE Conference on Computer Communications (INFOCOM)*, pp. 1233–1242, March 2000.

[11] M. Grossglauser and D. Tse, "A Framework for Robust Measurement-Based Admission Control," *IEEE Transactions on Networking*, vol. 7, no. 3, pp. 293–309, 1999.

[12] S. Jamin, S. Shenker, and P. Danzig, "Comparison of Measurement-Based Call Admission Control Algorithms for Controlled-Load Service," in *IEEE Conference on Computer Communications (INFOCOM)*, pp. 973–980, March 2000.

[13] R. Gibbens and F. Kelly, "Measurement-Based Connection Admission Control," in 15^{th} *International Teletraffic Congress (ITC)*, (Washington D. C., USA), June 1997.

[14] J. Milbrandt, M. Menth, and S. Oechsner, "EBAC - A Simple Admission Control Mechanism," in 12^{th} *IEEE International Conference on Network Protocols (ICNP)*, (Berlin, Germany), October 2004.

[15] M. Menth, J. Milbrandt, and S. Oechsner, "Experience-Based Admission Control," in 9^{th} *IEEE Symposium on Computers and Communications (ISCC)*, (Alexandria, Egypt), June 2004.

[16] J. Milbrandt, M. Menth, and J. Junker, "Performance of Experience-Based Admission Control in the Presence of Traffic Changes," in *IFIP-TC6 Networking Conference (NETWORKING)*, (Coimbra, Portugal), May 2006.

[17] J. Roberts, U. Mocci, and J. Virtamo, *Broadband Network Teletraffic - Final Report of Action COST 242*. Berlin, Heidelberg: Springer, 1996.

[18] A. M. Law and W. D. Kelton, *Simulation Modeling and Analysis*. Boston, USA: McGraw-Hill, 2000.

[19] V. Paxson and S. Floyd, "Wide-Area Traffic: The Failure of Poisson Modeling," *IEEE/ACM Transactions on Networking*, vol. 3, pp. 226–244, June 1995.

[20] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, and W. Weiss, "RFC2475: An Architecture for Differentiated Services," December 1998.

[21] E. Rosen, A. Viswanathan, and R. Callon, "RFC3031: Multiprotocol Label Switching Architecture," January 2001.

[22] A. Bjorck, *Numerical Methods for Least Squares Problems*. SIAM Society for Industrial & Applied Mathematics, 1996.

[23] R. Martin and M. Menth, "Improving the Timeliness of Rate Measurements," in 12^{th} *GI/ITG Conference on Measuring, Modelling and Evaluation of Computer and Communication Systems (MMB) together with 3rd Polish-German Teletraffic Symposium (PGTS)*, (Dresden, Germany), September 2004.