EBAC - A Simple Admission Control Mechanism

Jens Milbrandt, Michael Menth, Simon Oechsner
Dept. of Distributed Systems, Inst. of Computer Science, University of Würzburg
Am Hubland, 97072 Würzburg, Germany
Phone +49 931 8886631, Fax +49 931 8886632
Email: {milbrandt,menth,oechsner}@informatik.uni-wuerzburg.de

I. Extended Abstract

Next Generation Networks (NGN) are expected to provide Quality of Service (QoS) to the customers. This can be achieved by bandwidth overprovisioning or by Admission Control (AC), which is the focus of this work. AC can be subdivided into parameter-based AC (PBAC) methods and measurement-based AC (MBAC) methods. PBAC methods limit the acceptable traffic by accounting effective bandwidths [1] of admitted flows. PBAC is often inefficient because the AC decisions are rather pessimistic and the traffic descriptors usually overestimate the actually sent rate to avoid packet loss and delay due to spacing or policing. In contrast, MBAC approaches [2], [3], [4], [5] measure the current network load and take an estimate of the current characteristics of the new flow and the admitted aggregate to perform the AC decision. Other approaches [6], [7] work on end-to-end (e2e) measurements. MBAC methods take advantage of network measurements and admit traffic as long as enough network capacity is available. The downside of the MBAC approach is its susceptibility to measurement accuracy and QoS attacks.

In this work, we propose the experience-based AC (EBAC) as a hybrid approach of PBAC and MBAC. It pertains on a single link but it can be easily extended to a network-wide scope. It relies on peak rate traffic descriptors, which may be significantly overestimated. The utilization of the reserved capacity gives an estimate for the peak-to-mean rate ratio and allows for the calculation of an overbooking factor. The idea is simple but safety margins are required to provide sufficient QoS and questions arise regarding its robustness against variable traffic streams. We elaborate a feasible EBAC concept and show the computation of the overbooking factor. Furthermore, we present some selected results of our simulation experiments concerning EBAC on a single link and describe the extention of the EBAC mechanism to an entire network which is work in progress and implemented as a software running in a real-world testbed at Siemens/Munich.

The idea of EBAC is briefly described as follows. An AC entity for a link l limits the access to its capacity $c(l)$ and records the admitted flows $\mathcal{F}(t)$ 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 for a peak rate $r(f_{new})$. If

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

holds, admission is granted and f_{new} joins $\mathcal{F}(t)$. Flows are removed from $\mathcal{F}(t)$ on termination. 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.

For the computation of the overbooking factor $\varphi(t)$, we define $R(t) = \sum_{f \in \mathcal{F}(t)} r(f)$ as the reserved bandwidth of all flows and $C(t)$ denotes their unknown aggregated mean rate. EBAC measures the utilized bandwidth $M(t)$ of the aggregate reservation $R(t)$ and a time statistic for the reservation utilization $U(t) = \frac{M(t)}{R(t)}$ is collected. $U_p(t)$ is the p_u-percentile of the empirical distribution of $U(t)$ and the reciprocal of this percentile is the overbooking factor $\varphi(t) = \frac{1}{U_p(t)}$.

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 $M_i = \frac{\Gamma_i}{\Delta_i}$ is determined by metering the traffic volume Γ_i sent during I_i.

Statistic Collection $P(t, U)$: The aggregate reservation $R(t)$ is known from the AC process. The utilization values $U(t)$ are sampled in constant time intervals and are stored as hits in bins for a time-dependent histogram $P(t, U)$. The time-dependent utilization quantile $U_p(t)$ can be derived from $P(t, U)$ by

$$U_p(t) \quad = \quad \min_u \{u : P(t, U \leq u) \geq p_u\}. \qquad (2)$$

To avoid an underestimation of $U_p(t)$ and an overestimation of $\varphi(t)$, enough statistical data must be collected before $\varphi(t)$ is calculated.

Statistics Aging: If traffic characteristics change over time, the EBAC utilization statistic must forget obsolete data to reflect the properties of the new traffic mix. Therefore, we devaluate the contents of the histogram bins regularly by a devaluation factor d.

Our method for evaluating the EBAC performance is the following. The objective of AC is to limit packet delay due to queueing and to avoid packet loss due to buffer overflow. If packet loss can be eliminated by sufficiently large buffers,

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

packet delay is the natural performance measure for the assessment of AC mechanisms. For a small traffic load and $C(t) << c(l)$, the really experienced delay can be very small even for too large overbooking factors like $\varphi(t) >> K(t)$. As the overbooking factor must still be safe if the link is highly utilized, the really experienced delay is not suitable for the validation of the EBAC. As it is impossible to scale up the given traffic realistically to simulate a link under heavy load, we scale down the link capacity such that it is just enough to meet the QoS requirements of the traffic. This equals a virtual server with capacity $C_v(t)$ and the observed waiting time is the virtual server delay W_v. We can estimate the mean rate of the traffic aggregate by $\frac{R(t)}{\varphi(t)}$ and we want to guarantee a maximum delay W_{max} with a probability p_W, i.e., $P(W \leq W_{max}) > p_W$. Therefore, we compute the required virtual server rate $C_v(t)$ based on a $N \cdot D/D/1$ queuing system with a mean arrival rate $\frac{R(t)}{\varphi(t)}$. Finally, we take the mean $E[W_v]$ of the virtual server delay W_v and primarily its 99%-quantile as performance measures.

The purpose of the EBAC performance study on a single link is manifold. The most important is the proof of the EBAC concept. The intrinsic idea of EBAC is the exploitation of the peak-to-mean rate ratio $K(t) = \frac{R(t)}{C(t)}$ of the traffic aggregate. In the following we show that the EBAC overbooking factor achieves this goal.

We performed simulations with different peak-to-mean rate ratios by using different peak rates for source traffic shaping. Figure 1 illustrates that EBAC reacts very well to traffic with different but constant peak-to-mean rate ratios. The average overbooking factor $E[\varphi]$ is approximately as large as the average peak-to-mean rate ratio $E[K]$ of the traffic aggregate. At the same time the virtual server delay W_v is well limited.

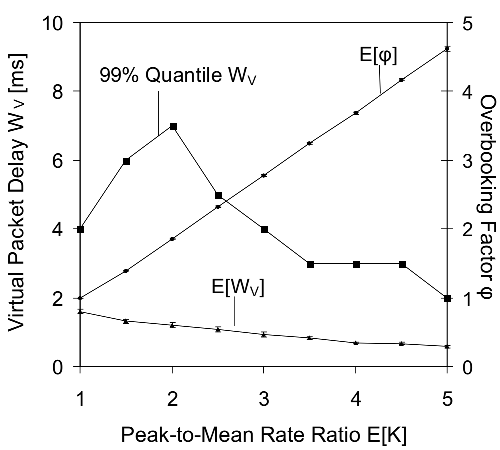

Fig. 1. Sensitivity of EBAC to the average peak-to-mean rate ratios.

Further simulation results allowed us to give recommendations for the EBAC parameters such as measurement interval length and reservation utilization percentile to obtain most reliable overbooking values. Others showed that the EBAC mechanism is robust against traffic variability in terms of packet size and inter-arrival time distribution as well as correlations thereof.

Our investigations of the EBAC mechanism on a single link are not completed yet. Further simulations shall reveal e.g. the impact of the economy of scale, i.e., rising link and aggregate bandwidths on the EBAC performance. As we limited our EBAC studies to steady-state simulation so far, we want to elaborate the EBAC behavior in transient phases in future experiments.

Among the work in progress, there is also an extention of the EBAC that allows for resource overbooking within an entire network. A first software protoype is already implemented and integrated in a large test network belonging to the project KING organized by Siemens/Munich. There the EBAC is part of the KING next generation network architecture [8] which uses border-to-border budget-based AC [9] where AC is performed by admitting flows to virtual border-to-border tunnels. By means of the EBAC mechanism, these tunnels can be overbooked such that the overal network efficiency is increased while QoS constraints are still met.

We proposed the experience-based AC (EBAC) as a new AC method. Flows signal their possibly overestimated peak rate demands to request a reservation and EBAC performs its AC decision taking an overbooking factor $\varphi(t)$ into account which is based on a time series of observed reservation utilizations. The overall EBAC objective is to increase resource efficiency while maintaining QoS guarantees in presence of overestimated peak rates and varying traffic rates. We briefly described the EBAC concept, explained our performance evaluation methodology and gave a first insight into our simulation results. Current and future work on EBAC concerns further single link simulations as well as the investigation of the EBAC as a network-wide overbooking mechanism.

References

[1] James Roberts, Ugo Mocci, and Jorma Virtamo, *Broadband Network Teletraffic - Final Report of Action COST 242*, Springer, Berlin, Heidelberg, 1996.

[2] R. Gibbens and F. Kelly, "Measurement-Based Connection Admission Control," in 15^{th} *International Teletraffic Congress*, June 1997.

[3] Matthias Grossglauser and David N. C. Tse, "A Framework for Robust Measurement-Based Admission Control," *IEEE/ACM Transactions on Networking*, vol. 7, no. 3, pp. 293–309, 1999.

[4] Lee Breslau, Sugih Jamin, and Scott Shenker, "Comments on the Performance of Measurement-Based Admission Control Algorithms," in *INFOCOM (3)*, 2000, pp. 1233–1242.

[5] Hans van den Berg and Michel Mandjes, "Admission Control in Integrated Networks: Overview and Evaluation," in *Proceedings 8th International Conference on Telecommunication Systems*, Nashville, US, 2000, pp. 132 – 151, Guter Ueberblick.

[6] Coskun Cetinkaya and Edward W. Knightly, "Egress Admission Control," in *INFOCOM (3)*, 2000, pp. 1471–1480.

[7] Ignacio Más and Gunnar Karlsson, "PBAC: Probe-Based Admission Control," in 2^{nd} *International Workshop on Quality of future Internet Services (QofIS2001)*, September 2001.

[8] C. Hoogendoorn, K. Schrodi, M. Huber, C. Winkler, and J. Charzinski, "Towards carrier-grade next generation networks," in *Proc. ICCT 2003*, Beijing, China, April 2003.

[9] Michael Menth, Stefan Kopf, Jens Milbrandt, and Joachim Charzinski, "Introduction to Budget Based Network Admission Control Methods," in 28^{th} *Annual IEEE Conference on Local Computer Networs (LCN2003)*, Bonn, Germany, Oct. 2003.