Paper

Simple admission control procedure for QoS packet switched military networks

Damian Duda and Wojciech Burakowski

Abstract— Providing quality of service (QoS) into the networks based on the packet switched technologies, as ATM and IP, is currently the challenge for the military communication system designers. The main element for achieving QoS capabilities is to implement effective admission control (AC) function, which regulates the volume of submitted traffic to the network. The traditional approach for the AC is that it is invoked by each call requesting QoS. As a consequence, the call set-up latency is increasing and, in addition, the signaling traffic in the network is growing. This paper proposes a simple AC method that is based on the online traffic load measurements and assumes that the AC is involved only when the load exceeds a predefined threshold. As a consequence, for most of the connections the AC is not necessary to be executed and this causes lower set-up phase duration and limits the volume of signaling traffic. The numerical results showing effectiveness of the approach are included and compared with traditional AC performing.

Keywords— QoS, admission control, measurement based AC.

1. Introduction

The evolution scenario of military networks assumes that the packet switched networks, based on ATM/IP technologies, will substitute the existing circuits switched networks. Notice, that in the most of the NATO countries this process was finished or is currently running. Let us remark that packet networks designed for the military should provide quality of service (QoS) capabilities for the packet transmission level since some information needs to be transmitted as urgent or very urgent. So, we can not rather use the best effort service only as it is available in today's public Internet. The QoS requirements coming from the military applications are mainly referring to the timely context delivery and real time transmission capabilities for such applications as voice over Internet Protocol (VoIP), videoconference, radar data, etc.

One can distinguish between two methods for providing QoS capabilities into the packet switched networks. The more complex approach is to implement in the network additional QoS mechanisms working on different time scale, like traffic control mechanisms at the packet level (classifiers, schedulers, markers, etc.), traffic control mechanism at the call level – admission control (AC) as well as adequate QoS routing procedures. It is worth to mention that these mechanisms could be collected in the form of appropriate QoS architecture as differentiated services (DiffServ) architecture [1]. The second approach is to keep the network in the over-provisioned state. Anyway, such approach that can be effective in the public core IP-based networks is not appropriate to be applied in the military, especially on the tactical level, where we face with the radio or radio-relay links with limited capacity. As a consequence, the more attractive solution for providing QoS in the military IP-based networks is to follow the approach with engaging traffic control mechanisms, where the AC function plays a key role.

The traditional approach for the AC is that it is invoked by each call requesting QoS. As a consequence, the call set-up latency is increasing and, in addition, the signaling traffic in the network is growing. In order to overcome this problem, in this paper we propose and evaluate a new strategy for performing in an effective way AC function, targeted to limit invocation of the AC but keeping QoS capabilities. The discussed approach, named as AC with threshold (AC-T), is based on the online traffic load measurements. It assumes that the AC is invoked only when the traffic load exceeds a predefined threshold.

2. Proposed AC-T method

The strategy based on the invocation of AC each time when new call is submitted to the system is the most popular solution for achieving QoS guarantees in the network. Let us recall that the AC makes the decision about admission/rejection of new call on the basis of the traffic declarations and the current traffic load. Anyway, when decision is to accept, the network guarantees that the packet transfer characteristics will satisfy the QoS objectives as assumed for the AC. This strategy is typical for QoS networks.

The call request includes QoS needs coming from the user application, e.g., requested bandwidth, accepted packet losses, etc. The appropriate network control entities hold current status of the network resources and give the necessary information for making AC decision. The available resources are computed accordingly to the predefined QoS objectives. For example, QoS objectives for voice traffic are target values of maximum delay, maximum delay variation and packet loss ratio.

In the simplest case, when the provisioning of resources is based using static bandwidth allocation, the call request is accepted if the sum of currently used bandwidth and the bandwidth needed for new connection does not exceed the bandwidth allocated for QoS traffic. For instance, the AC algorithm may work accordingly to the formula:

$$Bw_{new} + \sum_{i=1}^{N} Bw_i \leq C\,, \tag{1}$$

where:

N – number of running connection,

Bw_{new} – bandwidth requested for a new connection,

Bw_j – bandwidth used by ith connection, $i = 1, \ldots, N$.

The acceptance decision is made only when formula (1) is satisfied. It should be noticed that this formula is executed each time when new call is submitted to the system and this causes that delay appears in setting the connection, even when the traffic is low.
The investigated in the paper approach assumes that for new calls we invoke the AC depending on the current traffic load. The proposed mechanism is based on the following principles (see Fig. 1):

- All requests are accepted when current load (or other specified but representative parameter) is below a given threshold, say thr_1.
- Acceptance of new call demands AC decision when load carried by the link exceeds the threshold thr_1.
- The invocation of AC is again stopped when the load decreases to a given threshold, say thr_2 ($thr_1 \geq thr_2$).

Similar mechanism as above was proposed in [2, 3], for handling TCP flows by implicit AC.

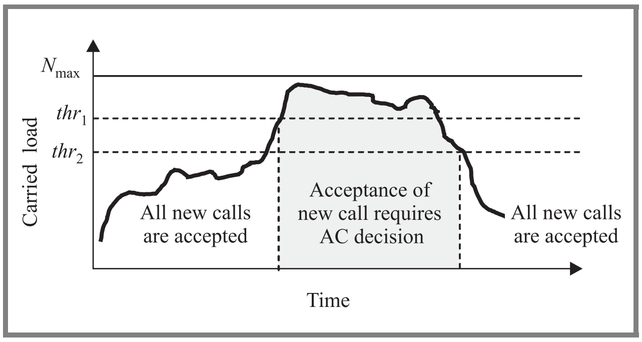


Fig. 1. Areas when the AC function is invoked.

In AC-T strategy, AC function is executed when current load – for simplicity, represented by the number of running connections N – exceeds predefined threshold $trsh_1$. In this case, a new connection is admitted only if:

$$(N+1) \leq N_{\max}\,, \tag{2}$$

where:

N – number of running connections,

$N_{\max}$ – maximum number of simultaneous connections allowed by the AC.

Assuming that each call requests the effective bandwidth R (homogenous traffic sources case) the $N_{\max}$ is calculated as:

$$N_{\max} = \frac{C}{R}\,, \tag{3}$$

where:

C – the link capacity.

3. Disadvantages of AC-T

We assume that in proposed approach the AC decision is based on information gathered from measurements. The usual problem of such approach is that accuracy of performing AC strictly depends on credibility of measurement process. In real networks the data for AC function are provided by measurement subsystem after some delay, always greater than zero. Let us name this measurement delay as D. It is likely that in time interval D, there may be a number of connections accepted violating link capacity Eq. (1) or a connection with large bandwidth request may be admitted, consuming or exceeding remaining capacity Eq. (2). So the probability of exceeding link limit $N_{\max}$ is non-zero:

$$\Pr\{N > N_{\max}\} \leq \varepsilon\,; \quad \varepsilon \geq 0\,. \tag{4}$$

The objective of AC-T performing is to keep as close to 0 as possible. The best quality is offered to a user if $\varepsilon = 0$. This state may be reached when there is no measurement delay, i.e., $D = 0$. In proposed approaches, where we assume that D is non-zero, the degradation of QoS is expected, as depicted in Fig. 2. The shadowed region corresponds to expected deterioration of running connections quality, e.g., increase of packet losses or packet transfer delay.

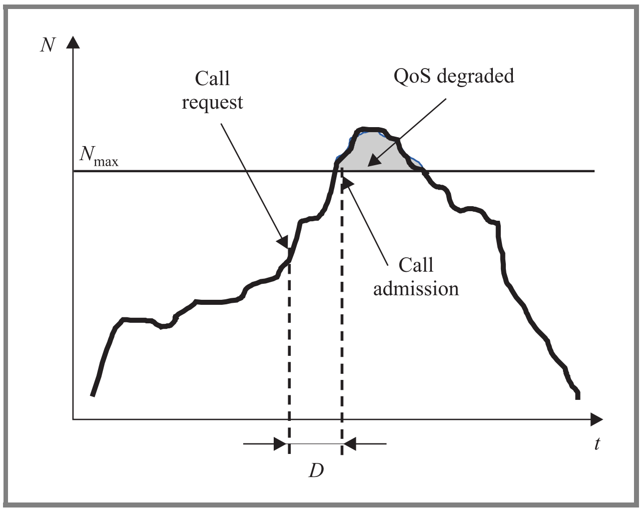


Fig. 2. The impact of measurement delay D on performance of AC function.

4. Simulation model

Let us consider a simple queuing system with single server of capacity C and infinite buffer B. The traffic submitted to this system may come from a number of users each of them requesting specified link bit rates. Assuming that each connection requests bit rate equal to its effective bandwidth R (homogenous case), we get simulation model with N_{max} available channels each of R bitrate as depicted in Fig. 3, where N_{max} is calculated, e.g., from Eq. (3).

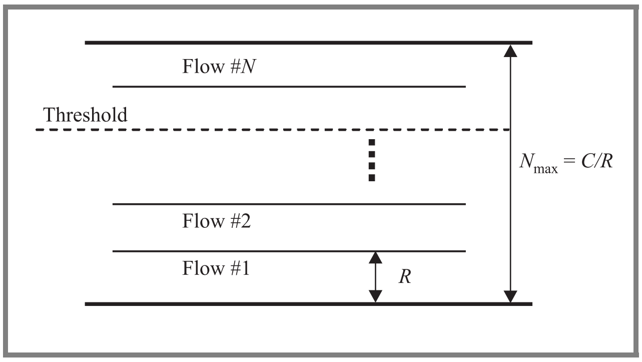


Fig. 3. Simulation model of the link with capacity of N_{max} connections.

The impact of measurements on AC decision were simulated by introducing the delay parameter D in the simulator. As pointed in the previous section, this parameter reflects the time after which the information about current link load is available for AC function.

5. Effectiveness criteria

Three parameters were taken into account for making an evaluation of the discussed approach for performing AC-T.

Accepted call ratio (*ACR*), defined as a ratio of calls accepted by AC to the total number of submitted calls:

$$ACR = \frac{accepted\_calls}{call\_attempts}. \tag{5}$$

Let us remark that the *ACR* is strictly related to the ratio of successfully completed calls, if we do not consider the number of unfinished calls at the end of the observation time interval.

Now, let us define the QoS experienced by the user as a percentage of connection time when QoS degradation is observed. Mostly, this happens when the number of running connections is greater than N_{max} and, as a consequence, the offered load is greater than the link capacity. Thus we introduce two next parameters for characterising the QoS degradation level: degradation time ratio and cumulative degradation time ratio.

Degradation time ratio (*DTR*) corresponds to the observation of a single connection duration time t_C and represents the part of t_C when we have more than N_{max} running connections in the system (jointly with the connection in question). In general, the N_{max} may be exceeded several times during connection time t_C (Fig. 4).

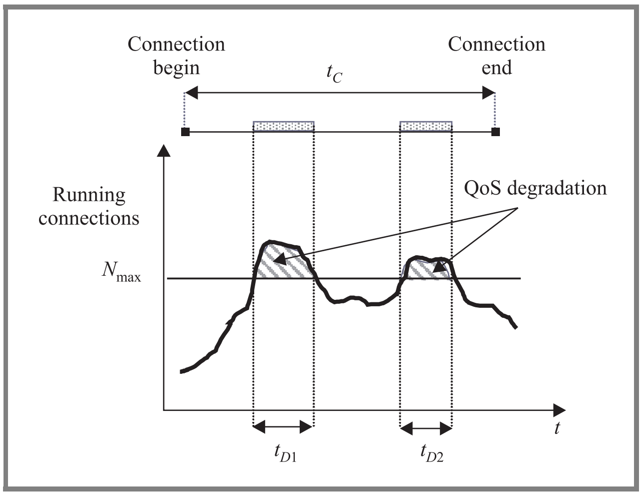


Fig. 4. Time intervals when QoS of the connection is degraded.

Thus, for the jth connection, the *DTR* is calculated as a sum of all time intervals when $N > N_{max}$, limited for jth connection duration time t_{Cj}:

$$DTR_j = \frac{\sum_k t_{Dk}}{t_{Cj}}; \quad k = 1, 2, \ldots \tag{6}$$

where:

t_{Dk} – kth time interval of jth connection when the number of admitted connections is greater than N_{max}.

Notice that if degradation time is equal or greater than connection duration time then *DTR* becomes 1.

Cumulative degradation time ratio (*CDTR*) is defined as the ratio of time when capacity limit N_{max} is exceeded to the total time of observation:

$$\frac{t_{N>N_{max}}}{T}, \tag{7}$$

where:

$t_{N>N_{max}}$ – time interval when in the system there are more than N_{max} running connections,

T – time of observation.

The same effectiveness metric as above was proposed in, e.g., [4].

6. Numerical results

In this section we present the exemplary numerical results showing the performance evaluation of the proposed strategy and compare it with the traditional admission control strategy, named AC. The simulation tests were carried under assumption of Poissonian calls arrival process with rate varied from 50 to 120 calls/s and exponential service time distribution normalised to 1. All calls required the same bandwidth, so the link capacity could be expressed by the maximum number of simultaneously running connections (see Eq. (3)). The capacity allocated for both of strategies was the same. The rest of parameters were set as follows:

– link capacity allocated $N_{\max}$ 100 connections,
– thresholds $trsh_1 = trsh_2 = trsh$,
– measurement delay $D = 0.1;\ 1;\ 5$ s.

At first, the comparison of accepted calls ratio versus load was performed. The numerical results of the test are shown in Figs. 5 and 6. The accepted calls ratio differs depending on the chosen strategy. For AC-T the resulting *ACR* factor is approximately the same as we can get from AC. The reason is that AC-T strategy uses accurate information about link load, when the link utilisation exceeds threshold *trsh*, here 0.8 of link capacity.

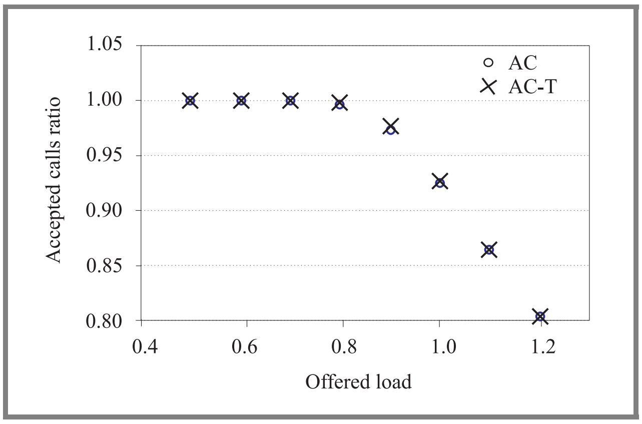


Fig. 5. Accepted calls ratio versus load for $trsh = 0.8$ and $D = 5$ s.

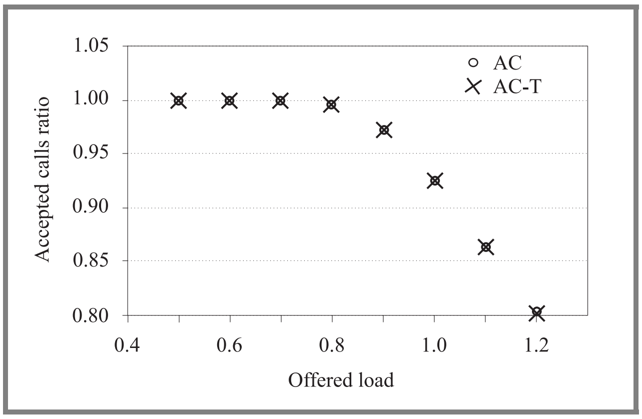


Fig. 6. Accepted calls ratio versus load for $trsh = 0.8$ and $D = 0.1$ s.

Figures 7 and 8 show the impact of threshold value on efficiency of proposed strategies. Notice that for AC-T the *ACR* is close to the received values for accurate AC strategy (see Fig. 7). For more accurate measurements, i.e., smaller *D*, the resulting *ACR* for AC-T is very close to the reference values (see Fig. 8). Comparing Figs. 7 and 8 we can see that there are more admitted connections when *D* increases but with less respect to available link capacity.

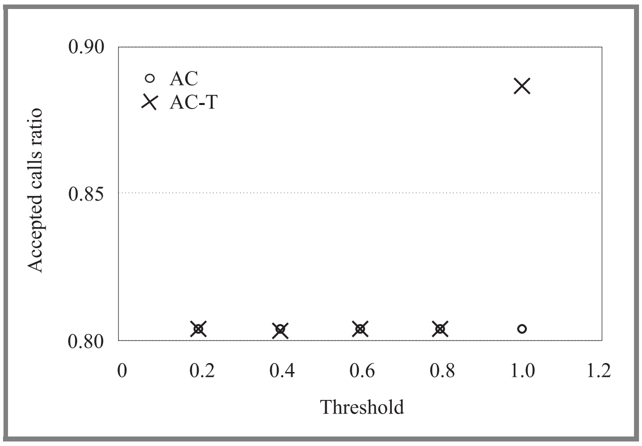


Fig. 7. Accepted calls ratio versus threshold for $\rho = 1.2$ and $D = 5$ s.

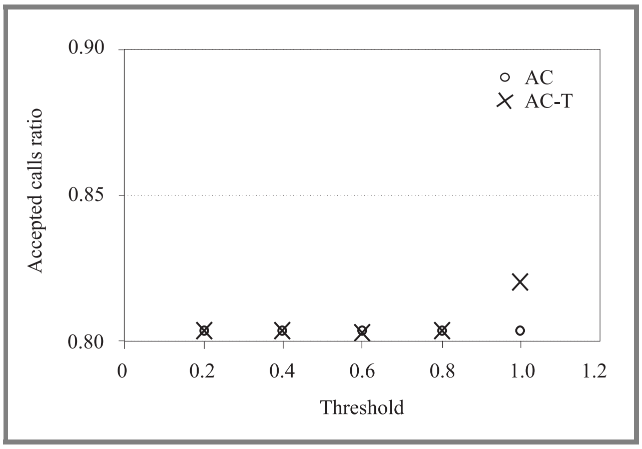


Fig. 8. Accepted calls ratio versus threshold for $\rho = 1.2$ and $D = 0.1$ s.

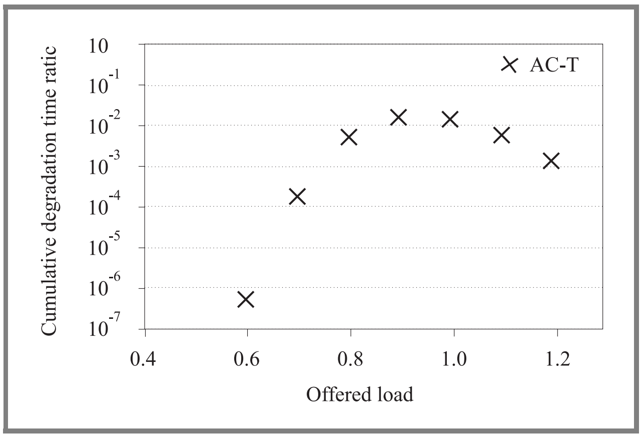


Fig. 9. Cumulative degradation time ratio versus load for $trsh = 0.8$ and $D = 5$ s (note: no events were observed for AC).

Next, the evaluation of *CDTR* was performed. As we could expect, *CDTR* parameter is affected by the accuracy of measurements. The simulation results for inaccurate measurements are depicted in Fig. 9. The resulting *CDTR* is relatively high, but the degradation ratio improves to 10^{-3} when link load exceeds threshold set to $trsh = 0.8$, i.e., the precise AC function is triggered. Note, that degradation concerns all running connections.
In the case of more accurate measurements, for $D = 0.1$ s, we observe negligible degradation level, that is less than 10^{-7} (not shown in the figure).

7. Conclusions

The paper presented and evaluated the approach to facilitate AC function in packet switched networks. It appears that we do not need to invoke the AC each time new call is submitted and we are able to provide QoS guarantees at the assumed level. The effectiveness of proposed AC-T algorithm was evaluated by simulations and compared with hypothetical, accurate AC algorithm, working according to M/M/N/N model.
Anyway, for effective implementation of the strategy AC-T we need to assure that the measurement delay is not too large. The sensitivity of proposed strategies on measurements accuracy can be seen as a trade-off to expected decrease of signaling overhead.
In further work we focus on developing the rule to assign proper threshold values, which seems as a way to manage the measurement issues.

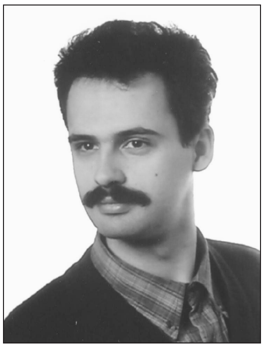
References

[1] S. Blake *et al.*, "An architecture for differentiated services", IETF, RFC 2475, Dec. 1998.

[2] S. Ben Fredj, S. Oueslati-Boulahia, and J. W. Roberts, "Measurement-based admission control for elastic traffic", in *Proc. Int. Teletraff. Congr. ITC-17*, Salvador, Brazil, 2001.

[3] N. Benameur, S. Oueslati, J. W. Roberts, "Experimental implementation of implicit admission control", Tech. Rep. 279TD(03)026, COST-279, 2003.

[4] R. Martin, M. Menth, and J. Charzinski, "Comparison of border-to border budget based network admission control and capacity overprovisioning", COST 279, 2003.



Damian Duda was born in Nysa, Poland, in 1973. He received M.Sc. degree in telecommunications from Military University of Technology, Faculty of Electronics, in 1998. He now works as an IT specialist for Military Communication Institute, Zegrze, Poland. His research is focused mainly on modeling, simulation and evaluation of traffic control mechanisms in packet switched QoS networks.
e-mail: d.duda@wil.waw.pl
Military Communication Institute
05-130 Zegrze, Poland

Wojciech Burakowski was born in Poland, in 1951. He received the M.Sc., Ph.D. and D.Sc. degrees in telecommunications from Warsaw University of Technology, in 1975, 1982 and 1992, respectively. He is now a Professor at the Institute of Telecommunications, Warsaw University of Technology. Since 1990, he has been involved in the European projects COST 224, COST 242, COST 257 and 279. His research interests include ATM and IP network design as well as traffic control mechanisms.
e-mail: wb@wil.waw.pl
Military Communication Institute
05-130 Zegrze, Poland