EVALUATION OF TWO STATISTICAL CAC METHODS FOR VARIABLE BIT RATE TRAFFIC SOURCES

Matej Kavacký — Ivan Baroňák *

The aim of this paper is to present a complex comparison of two methods of connection admission control in ATM network. In the introduction the value of traffic management is briefly presented. In the next part, selected CAC methods, which have been compared, are presented. The following part is dealing with comparison of selected CAC methods on the basis of simulations in MATLAB for different sources of variable bit rate (VBR) traffic.

Keywords: ATM, traffic management, connection admission control methods

1 INTRODUCTION

ATM (Asynchronous Transfer Mode) is a powerful and flexible switching and multiplexing broadband technology that allows concurrent transmission of voice, video, pictures and data. It is based on fast cell switching with a constant length of 53 bytes. The ability to define the provided Quality of Service (QoS) is one of the main advantages of ATM technology. This capability is implemented on the highest level in comparison with other broadband networks. Hence, the customer can choose the required QoS level and pay only for what he or she really demands. Traffic control mechanism — traffic management — should be implemented in the network to secure QoS.

The ATM technology offers several traffic control mechanisms [1,2]. Traffic control can be based either on preventive or on reactive functions. Preventive traffic control is mainly used in ATM networks and it utilizes functions as eg Connection Admission Control (CAC), Traffic Shaping and Network/User Parameter Control (NPC/UPC). A large number of various CAC methods have been offered but none of them is universal for all traffic types.

Considering the fact that ATM is a connection oriented mode, it is possible to secure QoS for connections on the level agreed in advance. The required QoS level is arranged in the connection set-up through traffic contract. The quality of Service means provision and observation of required traffic parameters (cell losses, cell delay, etc) in the course of connection. The traffic contract states duties for both parties — network contracting to support the traffic at the arranged level and user’s agreement not to exceed the arranged performance boundaries.

Connection Admission Control CAC is a preventive function of traffic control that determines whether new incoming connection to the network will be accepted or rejected. The decision process is based on various traffic parameters, such as the current traffic load, the values of characteristic parameters (e.g., mean and peak rates), availability of network resources and required QoS of existing and new connections. Many CAC methods and their modifications have been offered. Some of them are based on mathematical models, statistics and theory of probability. Another large group of methods is based on on-line traffic measurements and on the analysis of the buffer states in the network nodes [3]. Methods based on artificial intelligence eg fuzzy systems [4] and artificial neural networks are also developed and extensively studied.

2 COMPARISON OF SELECTED CAC METHODS

2.1 Selected CAC Methods

Following the performed analysis in the field of methods of connection admission control verification and complex comparison of two selected CAC methods were chosen. The selection of methods was made according to these criteria — option of comparison on the basis of identical input parameters, affordable computing fastidiousness and simple implementation of the methods in applied solutions. The following methods are the most suitable according to our criteria — Method of effective bandwidth [5] and Diffusion approximation based method [6], [7]. As the main parameter for the efficiency comparison of these two methods the maximum number of accepted connections was chosen and the following dependences were studied:

– influence of the buffer size on the number of accepted connections,
– statistical multiplexing gain,
– influence of the required cell loss on the number of accepted connections,
– admission regions for two types of traffic. During the observation of individual methods efficiency, an ATM

* Department of Telecommunications, Faculty of Electrical Engineering and Information Technology Bratislava, Ilkovičova 3, 812 19, Bratislava, Slovakia; matej.kavacky@stuba.sk, ivan.baronak@stuba.sk

ISSN 1335-3632 © 2008 FEI STU
switch with a \( N \) input line was included, as well as one output line with a capacity of 155.520 Mbit/s (STM-1) and buffer store for the input line with the end capacity \( B \). During detection of the influence of the buffer size and the required cell loss one class of connections was included, i.e., all incoming connections have identical traffic parameters (peak cell rate, average cell rate, and average cell clusters size). During the observation of admission regions two classes of connections were sent to the ATM switch. Analogous to the previous case connections belonging to one class have the same traffic parameters. Simulations were made in MATLAB 6.5 simulation environment.

2.1.1 Effective Bandwidth Based Method (Equivalent Capacity Method)

The method based on the effective bandwidth is also called the equivalent capacity method. It is a frequently used and popular method. Resolution about the acceptance of the group of connections from different sources depends on the amount of capacities of individual sources equivalent with the total line capacity.

The equivalent capacity can be defined with the help of the following thought. Each source is simulated as an interrupted fluid process (IFP). If \( R \) is the peak source rate, \( r \) is the time when the source is active and \( b \) is the average time of active source period, then IFP source is fully described by vector \((R, r, b)\). Let the source strike the waiting queue with the end capacity \( K \), then from the distribution of the queue length the rate of traffic \( c \) can be calculated which matches the given cell loss \( \varepsilon \). This traffic rate shapes the equivalent source capacity, let us say effective transmission area, and we can calculate it by solving this formula:

\[
\varepsilon = \beta \exp \left\{ -\frac{K(c - rR)}{b(1 - r)(R - c)e} \right\},
\]

where

\[
\beta = \frac{(c - cR) + \varepsilon r(R - c)}{(1 - r)e}. \tag{2}
\]

For the calculation of the equivalent capacity \( c \) the simplification is used, when \( \beta = 1 \) (usually \( \beta < 1 \)). Then for the equivalent capacity of one source we get

\[
c = \frac{a - K + \sqrt{(a - K)^2 + 4K\alpha r}}{2a} R, \tag{3}
\]

where

\[
a = b(1 - r)R \ln \frac{1}{\varepsilon}. \tag{4}
\]

For \( N \) traffic sources the equivalent capacity is also defined as the traffic rate, which secures the required cell loss \( \varepsilon \). Its calculation is more difficult, therefore the approximation is used

\[
c = \min \left\{ \rho + a'\sigma, \sum_{i=1}^{N} c_i \right\}, \tag{5}
\]

where \( c_i \)-represents equivalent capacity of \( i \) source, calculated following the ratio for one source, \( \rho \)-aggregate average bit rate, \( \sigma r = \sum_{i=1}^{N} \rho_i \), when \( \rho_i \) is the average bit rate of \( i \) source, \( \sigma \)-the amount of variants of bit rates of individual sources \( \sigma_i^2 = \rho_i(R_i - \rho_i) \).

This approximation for equivalent capacity of \( N \) sources only applies to the following cases:
- \( N \) multiplexed sources suffice with the equivalent capacity, which is the summary of equivalent capacities of individual sources.
- Stationary bit rate of \( N \) sources has approximately normal distribution with the average value \( \rho \) and variance \( \sigma^2 \).

2.1.2 Diffusion Approximation Based Method

This CAC method is based on the assessment of the statistical bandwidth, on the basis of ratios for the model of diffusion approximation. Two models for ATM switch are used in order to improve the exactness of this method:
- model for ATM switch with the finite buffer store size \( \text{FB} \) (Finite Buffer),
- model for ATM switch with the buffer store of infinite size \( \text{IB} \) (Infinite Buffer).

In the model with the infinite buffer store the probability of cell loss is calculated with the help of the overflow probability, which is the general formulation of overflow probability of the actual buffer store capacity \( K \) in a system with an infinite buffer store size. The cell loss probability expressed from these two models is

\[
L_{\text{FB}} = \frac{1}{\sqrt{2\pi}} \exp \left\{ \frac{2K}{\alpha} (\lambda - C) \right\} \exp \left\{ -\frac{(\lambda - C)^2}{2\sigma^2} \right\}, \tag{6}
\]

\[
L_{\text{IB}} = \sigma L_{\text{FB}}. \tag{7}
\]

For \( N \) sources of interrupted fluid processes (IFP) characterized by three parameters — peak transmission rate \( R_i \) period when the source is active \( r_i \) and average duration of active source period \( b_i \) for each \( i \) source — we gain the total variance \( \sigma^2 \) and total average value \( \lambda \) of transmission rate of consolidated traffic as

\[
\sigma^2 = \sum_{i=1}^{N} \sigma_i^2 \tag{8}
\]

\[
\lambda = \sum_{i=1}^{N} \lambda_i, \tag{9}
\]

where

\[
\sigma_i^2 = \lambda_i(R_i - \lambda_i), \tag{10}
\]

\[
\lambda_i = r_i R_i. \tag{11}
\]
The immediate variance of the cell access process $\alpha$ can be calculated according to the following reference:

$$\alpha = \sum_{i=1}^{N} \lambda_i \left( \frac{1 - (1 - \beta_i T_i)^2}{(\beta_i T_i + \gamma_i T_i)^2} \right),$$

where $T_i = 1/R_i$, $\beta_i = 1/b_i$ represents the average value of active period duration, $1/\gamma_i$ is the average value of passive period duration of $i$ source transmission.

The statistical bandwidth is defined as the bandwidth necessary for multiplexed connections to secure the cell loss probability less than the required cell loss $\varepsilon$. For the two aforementioned models we get references for the statistical bandwidth:

$$C_{FB} = \lambda - \delta + \sqrt{\delta^2 - 2\sigma^2\omega_1},$$
$$C_{IB} = \lambda - \delta + \sqrt{\delta^2 - 2\sigma^2\omega_2},$$

where

$$\delta = \frac{2K}{\alpha}\frac{\sigma^2}{\omega_1 = \ln(\varepsilon \sqrt{2\pi})},$$
$$\omega_2 = \ln(\varepsilon \sqrt{2\pi}) - \ln(\sigma).$$

As the assessment for the worst status of the traffic, the definite value of statistic bandwidth $C_{DF}$ is defined as the minimum from the two previous bandwidths

$$C_{DF} = \max(C_{FB}, C_{IB}).$$

### 2.1.3 Traffic Sources

As traffic sources those sources were chosen which match the real video traffic sources [8] and interactive web application [9,10] with their parameters. The traffic sources were characterized by two-state ON-OFF sources [11–14] with parameters presented in Tab. 1. Source number 1 represents the characteristics of Jurassic park, source number 2 represents the characteristics of the Total Recall and source number 3 corresponds to an interactive web application.

### 2 SIMULATION RESULTS

#### 2.1 Buffer Size \( vs \) Admitted Connections

In the following simulations during the comparison of selected CAC methods those cases were studied, in which the buffer size in ATM switch ranged from 2 to 64 kB (approximately from 40 to 1240 ATM cells), which is a small or medium buffer size [15]. The dependences of the maximum number of accepted connections on the buffer size of the switch have been studied for three different sizes of required cell losses.

Figures 2 and 4 show the simulated dependences of the number of accepted connections on the ATM switch buffer size, expressed in ATM cells. The size of ATM switch has been changed with the step 5 cells. With the method of the effective bandwidth the buffer size has been changed from 5 to 1500 cells. With the diffusion approximation based method the dependence on the number of accepted connections is shown only for sizes up to 500, 250 and 150 ATM because further expansion of the memory would not have any influence on the number of accepted connections, forasmuch as the upper bound was reached using the smaller memory size.

The obtained results for all three types of traffic sources show that assessments of the diffusion method (Figs. 2a, 3a and 4a) in almost all cases of the studied ATM switch buffer size are approaching the upper bound of the number of accepted connections $N_{\max}$, let us say they are reaching it. The most significant influence of the buffer store is in the field of very small values, approximately up to 30 ATM cells. But this influence is decreasing with the growing factor of the source activity. Using the least activity factor (source no. 1) differences in the number of accepted connections in the diffusion approximation based method with small memory values are ranging in tens of connections, during which time using the source with the highest activity factor (source no. 3) these differences are only in drives of connections. With all three sources it is obvious that the influence of the required cell loss on the assessment of the diffusion approximation based method is very small, during which time with the growing buffer size this influence is completely minimized. From the results it is seen that the diffusion approximation based method enables the best usage of the transmission range, because its assessments are approaching the upper bound of maximum number of accepted connections. But the problem comes into being in the case when one of the connections exceeds the agreed cell transmission rate. There might not be a transmission range back-up of sufficient extent out the transmission and the overloading may occur easily.

![Fig. 1. ON-OFF traffic source](image-url)
Fig. 2. left – Maximum numbers of accepted connections using the method of diffusion approximation in dependence on the ATM switch buffer size for source No 1 and three different required values of CLR parameter, right – maximum numbers of accepted connections using the method of effective bandwidth in dependence on the ATM switch buffer size for source No 1 and three different required values of CLR parameter.

Fig. 3. left – Maximum numbers of accepted connections using the method of diffusion approximation in dependence on the ATM switch buffer size for the source No 2 and three different required values of CLR parameter, right – maximum numbers of accepted connections using the method of effective bandwidth in dependence on the ATM switch buffer size for the source No 2 and three different required values of CLR parameter.

Fig. 4. left – Maximum numbers of accepted connections using the method of diffusion approximation in dependence on the ATM switch buffer size for source No 3 and three different required values of CLR parameter, right – Maximum numbers of accepted connections using the method of effective bandwidth in dependence on the ATM switch buffer size for source No 3 and three different required values of CLR parameter.
A completely different situation occurs using the method of effective bandwidth. In diagrams in Figs. 2b, 3b and 4b one can see the strong dependences of the number of accepted connections on the buffer size and on the required cell loss. With the smallest buffer size the assessments about the method of effective bandwidth are approaching the lower-bound of the number of accepted connections $N_{\text{min}}$. With the growing buffer size the number of accepted connections is gradually increasing, though it is not approaching assessments of the diffusion methods in the studied bounds. We can see a considerable lower number of accepted connections in comparison with the diffusion method. Accordingly we can see that more strict requirements on the service quality (the cell loss in this case) mean a markedly smaller number of accepted connections in the whole studied range of the buffer size. In comparison with the diffusion approximation based method the method of effective bandwidth offers a markedly worse transmission area utilization, especially at lower values of ATM switch buffer store. On the other hand it ensures the agreed service quality to individual connections because in the case when some of connections exceed the agreed parameters, there is a sufficient back-up of the transmission range to manage the unexpected increase of the transmitting rate of one of connections.

As it is seen in the previous diagrams, using the method of effective bandwidth, the upper bound of the number of accepted connections in the studied extend of buffer size (up to 64 kB, ie 1240 ATM cells) has not been reached. The plots in Figs. 5a and 5b show the dependence of the number of accepted connections for the method of effective bandwidth also for greater values of the buffer size, approximately up to 512 kB (10000 cells). In the first studied case (Fig. 5a) the traffic source no. 1 from Tab. 1 was used, in the second case (Fig. 5b) the traffic source no. 3 was used.

From the previous two diagrams it is clear that the number of accepted connections using the method of effective bandwidth is approaching the upper bound only at the buffer size over 2000 ATM cells in the case of source no. 1, let us say over 5000 cells in the case of the source no. 3. Also it is seen that the stricter the requirements on the service quality are, in this case the cell loss, the later occurs the upper mentioned approach to the upper bound of the number of accepted connections. At source no. 3 we can mention that the increase of the number of the accepted connections is less sharp than in the case of source no. 1 with the lower activity factor. The influence of the required cell loss on the number of accepted connections vanishes only at great values of the buffer size. The gained results of simulations confirm the assumptions that the method of effective bandwidth is more conservative in the acceptance of new connections than the diffusion approximation based method, but on the other hand it offers more security of the agreed service quality.

2.2 Statistical Multiplexing Gain

The effective CAC method is supposed to secure a certain gain of statistical multiplexing. If $N_{\text{MIN}}$ is the number of $i$ class accepted connections using the method of transmission area allocation on the basis of the peak cell rate, ie $N_{\text{MIN}} = \lceil C/\text{PCR} \rceil$, and $N_{\text{MAX}}$ is the number of $i$ class accepted connections using the average rate, ie $N_{\text{MAX}} = \lceil C/\text{SCR} \rceil$, then the statistical multiplexing gain for CAC method is defined as the ratio of the number of accepted connections using the given method $N_i$ to the number of accepted connections using the peak rate $N_{\text{MIN}}$. The CAC method is effective only if

$$\frac{N_i}{N_{\text{MIN}}} > 1.$$  \hfill (18)

The gains of statistical multiplexing for the method of diffusion approximation and individual traffic sources for the buffer size $B = 16$ kB are shown in Tab. 2.
From the results for the diffusion method it is obvious that the traffic parameters have the ultimate influence on the statistical multiplexing. The required cell loss and ATM switch buffer size have only a nominal influence on the amount of the statistical multiplexing gain. These two parameters lower the gain at small loss values \((10^{-9})\) and at the same time at the small buffer store (2 kB). It is seen that the biggest statistical multiplexing gain was reached with the source of the smallest activity factor, ie with the smallest relationship between the average and the peak cell transmission rate.

In Table 3 there are the values of the statistical multiplexing gain for the method of effective bandwidth for the buffer size \(B = 16\) kB shown. From the results the most noticeable are the great differences between the studied cases for individual traffic sources. This showed the stronger dependence on the used buffer size, as well as on the required cell loss, in comparison with the diffusion method. With all three sources it has been discovered that the statistical multiplexing gain grows with increasing buffer size and with decreasing requirements on the cell loss. In any of the studied cases the values of this gain did not reach the values discovered for the method of diffusion approximation. But here we can notice that the greatest values of the statistical multiplexing gain were reached with the source of the lowest activity factor. As it is clear, the stronger is the relationship between the total capacity of the output line and the average source rate, the bigger statistical multiplexing gain is reached.

### 2.3 Cell Loss Ratio vs Accepted Connections

The cell loss ratio (CLR) is a parameter which determines the reliability of the source. It is one of the main QoS parameters determining the service quality. Some services, as for instance voice transmission, are able to tolerate a relatively high cell loss (about \(10^{-3}\)). Though, the majority of current high-speed services require much smaller losses. The typically required cell loss ranges from \(10^{-6}\) to \(10^{-9}\). In the following simulations the dependence of the number of accepted connections has been studied, using the CAC methods in dependence on the required cell loss. The number of accepted connections has been studied for the required cell loss at intervals from \(10^{-9}\) to \(10^{-3}\). Simulations were performed also for different buffer sizes (20, 310, 620 and 1240 cells). Simulation results are graphically shown in Figs. 6 and 7.

Diagrams in Figs. 6a and 6b show the gained simulation results for the method of diffusion approximation. With all studied traffic sources only the a weak dependence of the number of accepted connections on the required cell loss was discovered. In all cases the assessments of the diffusion method are identical with the upper limit \(N_{\text{max}}\), or they are approximating it with diminishing requirements on the service quality. There are dependences only for 2 or 3 small buffer sizes shown in the diagrams because with the bigger buffer sizes it was
found that in the whole range of the required cell loss the number of the accepted connections was identical with the upper limit of the number of accepted connections $N_{max}$. In all diagrams the regress of the number of accepted connections for the smallest values of the cell loss ratio is seen.

In the diagrams there is a determined dependence of the number of accepted connections on the required cell loss for the method of effective bandwidth shown in Figs. 7a and 7b. In contrast to the diffusion method, a much stronger dependence of assessments of the method of effective bandwidth on the required cell loss is visible. At the smallest buffer size (40 cells) the number of accepted connections in all three cases is just above the lower-bound of the number of accepted connections $N_{min}$ for the given source type. In this case the sensitivity of the number of accepted connections to the required cell loss is minimal, only drives of connections in the whole range of studied cell loss. For the buffer sizes 310 and 620 ATM cells (16 and 32 kB) it was found that all sources had the sharpest dependence on the required cell loss. With these two buffer store values the assessment of the effective bandwidth method is lowered for tens of connections. With the biggest buffer store (1240 cells) the regress of the number of accepted connections is slower than with the average buffer sizes in all studied cases.

2.4 Admission Regions

In the previous simulations it was counted with one traffic source type inputting the ATM switch. In the next part the behaviour of both selected CAC has been studied for the case when the input traffic consists of several traffic source types. Two particular source types with different traffic parameters have been studied in simulations. Sources no. 1 and 3 from Table 1 were chosen for the research. In these simulations admission regions have
been studied. They are combinations of both types of connections, which the given method is able to accept. Simulation results are graphically shown in Figs 8 and 9. The line shows the maximal combinations of connections which can be accepted. All combinations lying below this line can be accepted by this method, and any combination lying above this line will not be accepted.

With the very small buffer store (40 cells) it was found that for all the studied required cell losses there was a significant gap in the admission regions of accepted connections for both methods used. The admission region for the method of effective bandwidth in this case is very small compared to the result of the diffusion method. For the required $CLR = 10^{-3}$ it is visible, that with the increasing buffer size the admission region for the method of effective bandwidth is increasing very fast, whereas with the diffusion method it changes only minimally, for drives of connections. With the biggest buffer size (1240 cells) the admission regions for both methods are approaching each other in this case.

With increasing requirements on the cell loss ($CLR = 10^{-9}$) it was found that admission regions for the diffusion method were almost the same for all the studied buffer sizes. Admission regions for the method of effective bandwidth are increasing with the increasing buffer store, but with the increasing requirements on the cell loss this growth is markedly lower.

2.5 Discussion of Results

The diffusion approximation based method shows the competence to utilize the available transmission area on the transmission line. At the same time it did not show a sharp dependence on the service quality parameters, which have been studied in this paper. Disadvantage of the excellent utilization of the transmission area by the diffusion method is the bigger tendency to overload the network in the case when one of the connections suddenly increases its transmission cell rate over the agreed peak rate, which can cause the failure of the agreed service quality of some connections. On the contrary, the method of the effective bandwidth shows a sharp dependence on these parameters in the whole studied range of selected service quality parameters. The stricter the service quality requirements, the fewer connections accepted. The advantage of such a method is the considerably bigger security of the service quality provision, in comparison with the previous case, when the transmission area was utilized to 100%. The disadvantage is the lower efficiency of the transmission area utilization for the strict requirements on the service quality.

The method of effective bandwidth is recommended to be used in the case, when there are customers who demand high security of the agreed service quality and less efficiency of the transmission area utilization is allowed, offset by a higher price for the customer. The usage of the diffusion approximation based method is suitable in the case when we want the transmission bandwidth to be utilized to the nines, hence to locate on the transmission line as many connections as possible even with the bigger requirements on the service quality (cell loss, buffer size). There is a bigger probability of possible overloading in this case that can be offset with eg a lower service price for the customer compared to the previous case.

3 CONCLUSION

ATM technology came into existence as a reaction on the necessity to have the broadband network in the telecommunication environment which is able to transmit numbers of different new services and applications, from voice to image and multimedia services. Each of these services has different traffic characteristics and requires a different service quality. This requested the implementation of the traffic management mechanism. In comparison
with alternative networks (synchronous networks, packet networks) the ATM network has this traffic management sophisticated on the highest level, with the current optimal and efficient network resource utilization. Despite all these advantages of the sophisticated traffic management the price and complexity of the ATM technology were the reasons, why it did not have as big success among competitive technologies, as for instance IP or MPLS.

Mainly the preventive mechanism of connection admission control is used in ATM networks. Until now a large number of methods of connection admission control (CAC) have been proposed, which differ in their principle, availability of application and exactness of the assessment.

Acknowledgement

This work is part of research activities conducted at the Slovak Technical University Bratislava, Faculty of Electrical Engineering and Information Technology, Department of Telecommunications, within the scope of the project VEGA No. 1/3118/06 “Traffic in convergent telecommunication systems and networks”.

References


Received 24 January 2008

Ivan Baroňák, was born in Žilina, Slovakia, on July 1955. He received the electronic engineering degree from the Slovak Technical University Bratislava in 1980. Since 1981 he has been a lecturer at Department of Telecommunications STU Bratislava. In 1992 he submitted PhD work from the field of Terminal telephone equipment. In 1995 he became an associate professor for the subject applied information. Nowadays he works as an associate professor in Department of Telecommunications of FEI STU in Bratislava. Scientifically, professionally and pedagogically he focuses on problems of digital switching systems, ATM, Telecommunication management (TMN), Next Generation Networks, problem of optimal modeling of private telecommunication networks and services.

Matej Kavacký was born in Nitra, Slovakia, in 1979. He received the engineering degree in telecommunications in 2004 from Faculty of Electrical Engineering and Information Technology of Slovak University of Technology (FEI STU) Bratislava. Since 2006, after finished his PhD study he has been working as researcher at the Department of Telecommunications FEI STU and his scientific research is focused on QoS in broadband networks.