Hysteresis-based congestion control at the B-ISDN access

Johannes Wissing

University of Hagen
Institute of Communications Systems, Bergischer Ring 100, D-5800 Hagen 1,
Federal Republic of Germany

ABSTRACT

Multiplexing of bursty sources and refined congestion control strategies are still subject of numerous research activities. Broadband applications with very high peak-to-mean bitrate ratio and long silence periods like still picture video gave rise to different ideas of rate control at the B-ISDN network access. Contributions on Input Rate Control for source coded traffic as well as on Server Rate Control within a LAN/ATM Interworking Unit have recently been presented. This paper addresses a congestion avoidance strategy at the network access regarding the aggregated traffic of bursty sources. Depending on the number of active sources as well as on certain defined congestion levels the cell rate at the network access is controlled. The proposed analytical approach is based on the model of uniform and continuous arrival and service. The selected underlying Markov chain contains "split" states in order to handle the congestion correlation. The proposed model is extended to an adaptive Non-Markov system where the buffer filling level is evaluated using a switching hysteresis. This type of congestion measurement turns out to be very useful for an adaptive rate control mechanism that guarantees a certain Quality of Service while still achieving a good statistical gain. The analytical approach is confirmed by results of a computer simulation that is extended to the more complex case of adaptive rate control.

1. INTRODUCTION

The interconnection of LANs and MANs has been identified to be a challenge for the ATM-based B-ISDN concerning bandwidth allocation for high speed computer communications. Virtual Channels (VCs) and Virtual Paths (VPs) are going to be used to interconnect customer premises networks, either employing VC/VP switching entities or supported by Connectionless Service Functions (CLSF)[1,2]. The traffic to be carried is expected to be "bursty", and it is recognized[3] that typical characteristics can be approximated by means of a superposition of high speed ON-OFF sources, possibly with long silence periods. File transfer as well as dialogue oriented workstation applications or still picture video are frequently cited examples, where the burstiness defined by the peak-to-mean bitrate ratio is high, say about 20. Within this domain a plain bandwidth allocation per call, applying any determination of a constant "effective bitrate" will waste a lot of transmission capacity to guarantee an acceptable grade of service, i. e. maximum limits for cell loss probability and transfer delay.

The challenge of creating appropriate protocols for dynamic bandwidth allocation according to actual requirements gave rise to several proposals. A procedure called Fast Reservation Protocol (FRP)[4] has been developed with detailed specifications. The protocol consists of an allocation mechanism between the FRP control units at the network access and the involved network nodes. A bandwidth demand message is only acknowledged if the bottleneck node is able to carry the increased cell rate. Another message going around is used to confirm or draw back the preceding demand. A different approach has been proposed by Gerla[5] et. al. The concept called Bandwidth Advertising supposes a mechanism to inform the network access entities (e.g. customer gateways) about the free available bandwidth on the VP. This additional capacity is used for the CLP=1 traffic thus exploiting bandwidth that would otherwise be wasted. Another concept of dynamic bandwidth allocation[6] is based on the measurement of the entire traffic flow that is buffered at the network access and then forwarded according to the actually calculated and announced cell rate.

The subject of the present paper is to present a model of refined server rate adaptation and its extension to a congestion avoidance strategy that is based on the evaluation of the access buffer filling level and a defined switching hysteresis. Thus, the enhanced procedure of bandwidth allocation per burst is supplemented by a concept of bandwidth allocation per congestion.

The proposed model employing a three-dimensional Markov chain is extended to an adaptive Non-Markov model. The idea is to switch over to a slightly different variant of the original model, if the access buffer is getting congested. The change between the two variants of the Markov model is controlled by a switching hysteresis, defined by two thresholds. The combination of both, the proposed server rate adaptation (per burst) and the adaptive mechanism (per congestion) turns out to be a proper strategy for both resource utilisation and congestion avoidance.

The paper is organized as follows. In section 2 various models of server rate adaptations are discussed. A description of the analytical approach based on the model of uniform and continuous arrival and service is given in section 3. The theory is essentially confirmed by the results of a computer simulation that is moreover extended to the Non-Markov case of adaptive congestion control (section 4). Finally, concluding remarks are given in section 5.

2. DESCRIPTION OF THE PROPOSED MODELS

Two-dimensional Markov chains have been employed among other things to model source coded traffic[7] or to adapt a server rate[6]. The behaviour of the sources is described by means of a one-dimensional Markov chain and its extension to a two-dimensional system, i.e. parallel Markov chains, where each chain represents a value of another system variable. In the present context this system variable is equal to the server rate at which ATM cells are delivered to the ATM network (Figure 1). The server rate is controlled by the source activity process and by the network via bandwidth allocation acknowledgement messages. The source traffic is assumed to be packet-oriented and a segmentation into ATM cells is supposed. This functionality is included in the computer simulation to assess the loss probability due to packetization, for a packet is lost if it cannot entirely be buffered. The affected connection may be marked to eliminate subsequent packets in order to alleviate the access buffer congestion. It may be better to limit packet loss to few connections, since a single lost packet may cause higher end-to-end protocols to retransmit a multiple of the lost data. In the following this kind of congestion control is not applied and the effect of retransmission due to packet loss is neglected. The extension to the Non-Markov model for hysteresis-based congestion control depending on the buffer filling level will be presented later in this section. Let us focus on several Markov models to point out some properties concerning bandwidth allocation and congestion avoidance strategies.

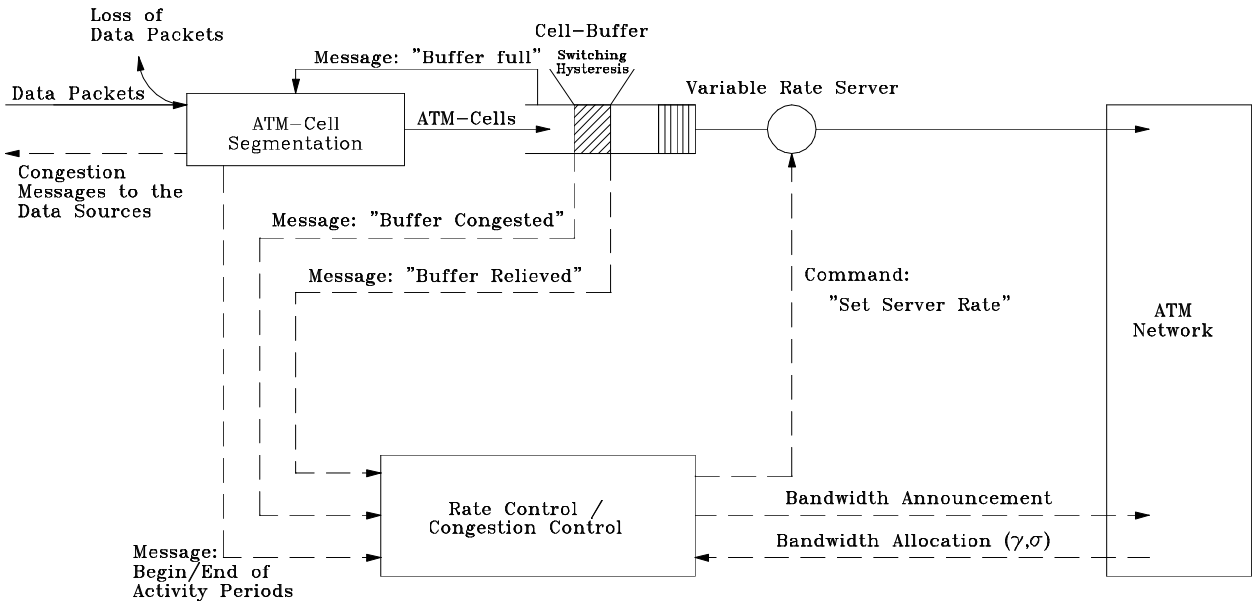


Figure 1 : Server rate control at the network access

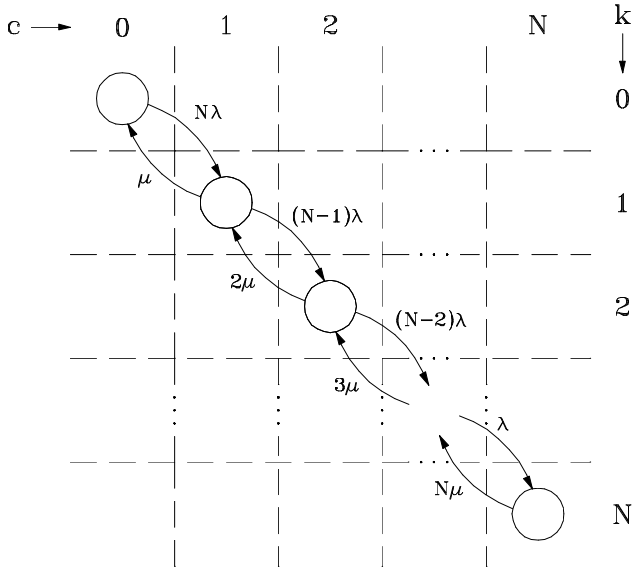


Figure 2: Markov chain without Server rate control

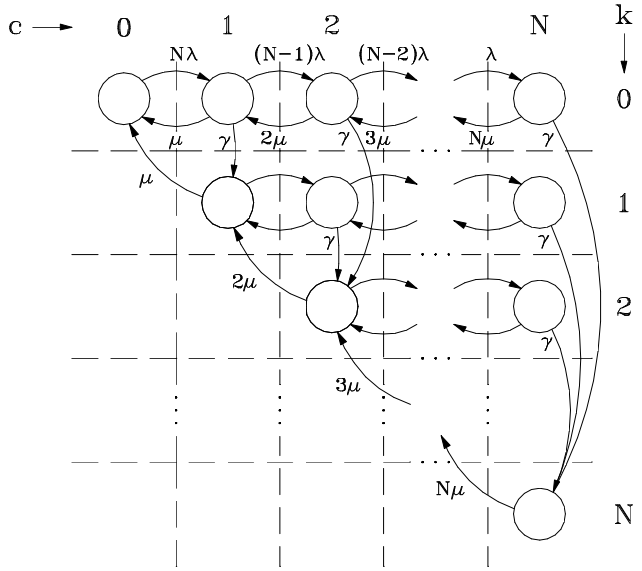


Figure 3: Two-dimensional Markov chain

At first a trivial case that does not include a server rate control is given in Figure 2. The columns (*c=0...N*) stand for the number *c* of active sources and each row (*k=0...N*) represents a rate at which ATM cells are delivered to the network. If there is any change in the number of active sources, the cell rate is adapted immediately. Assuming that a coincident reaction is nearly achievable this scheme does not employ any buffer space except for supporting the segmentation process. There is no influence on the cell stream flowing into the ATM network. From the network point of view the bandwidth reservation principle will exclusively be based on the statistical behaviour of the sources and this might be very ineffective, if both the burstiness and peak-to-link ratio is high. The opposite scheme is given in Figure 3 where any increase in the number *c* of active sources causes a bandwidth request while the actual server rate is kept and additional cells are stored in a buffer. The bandwidth request is confirmed within an exponentially distributed response time with mean $1/\gamma$ and according to the requirement the server rate is adapted. A decrement in the number of active sources will cause a bandwidth release message and the server rate is decreased immediately. This system with $(N+1)(N+2)/2$ states has been investigated by Schödl et al.[6] and is based on the renewal property of the network acknowledgement events. Network congestion limitation is supported using negative response messages, modelled as late network acknowledgements. If a network congestion situation lasts, several late acknowledgements will occur and cause an overflow at the access buffer, thus preventing the network from serious overload. This model is characterized by a reservation strategy that strictly buffers additional cells, if the number of active sources increases.

In the following the proposed models are based on a concept, that enables a rate increment in spite of a missing acknowledgement. At first some properties of Model A (Figure 4) will be described that is later on used as a reference model. As in the previous one an increment in the number of active sources causes a bandwidth request message that is positively answered. If there is another source arrival within the waiting time, another reservation message is sent behind the first one indicating that the server rate has been increased by one additional source in spite of the missing acknowledgement. This message will mostly cross the acknowledgement message thus avoiding unnecessary waiting times. Of course we have introduced uncertainty for the network to manage overload situations. A certain amount of bandwidth can permanently be reserved to absorb a rapid increase in the number of active sources. On the other hand the probability of the process moving along the upper diagonal line in Figure 4 is very low, if the reaction to a bandwidth request is on the average much faster than a change in the number of active sources. This holds true for the selected set of parameters described below, where the process is mainly determined by the probability of being in one of the lower numbered states. For example state number 1 of Model A is characterized by $\gamma >> \mu + (N-1)\lambda$. The dominance of the γ-transitions is a necessity to successfully perform the dynamic bandwidth reservation strategy. For the analysis and the computer simulation the parameter values of μ, λ and γ have been adopted from the cited approach[6] (Figure 3). The aggregated traffic is approximated by a superposition of N independent and identically distributed burst-silence sources.

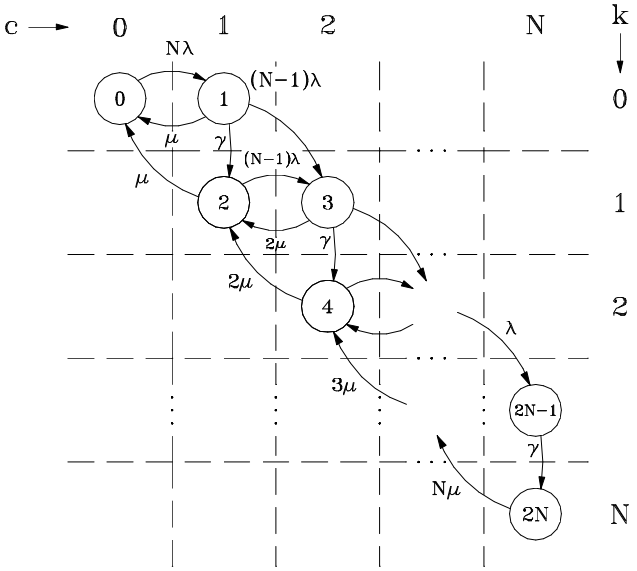


Figure 4: Model A

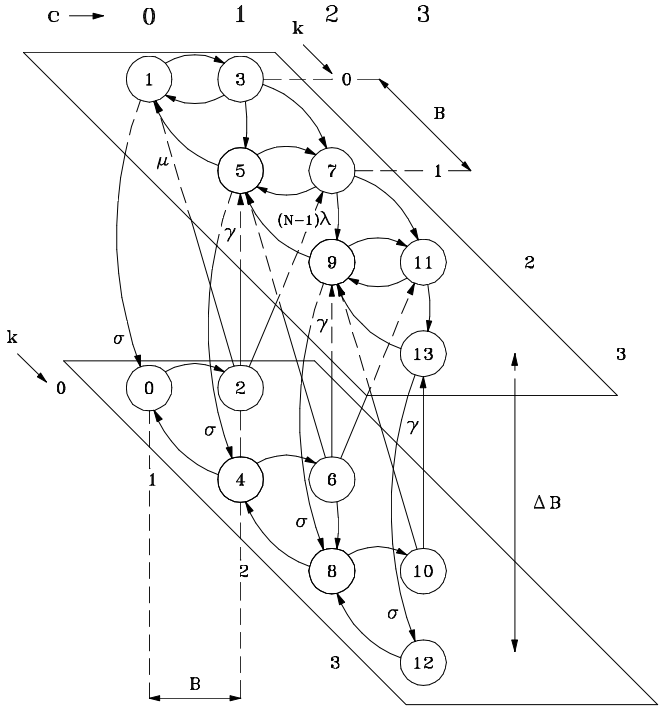


Figure 5: Model B

The burst periods with mean $1/\mu$ as well as the silence phases with mean $1/\lambda$ are assumed to be exponentially distributed. During an activity period (burst) a source emits *B* packets/s whereas no packets are being sent within a silence period . The actual packet rate of the superimposed sources is given by the binomial distribution with mean *NaB*, where $a = \lambda / \lambda + \mu$, the probability of the source being active, also denoted as the source activity factor. To guarantee the stability of the queuing system the mean server rate must be higher than the mean aggregated bandwidth requirements of the sources. The process being in one of the upper diagonal states of Figure 4 causes an increment of the buffer filling level or even packet loss. The change to a diagonal state is accompanied by a server rate adaptation that guarantees the decrement of the buffer. A straightforward way of overdimensioning bandwidth can be performed using a bandwidth offset B_{stat}, that is permanently allocated and simply added to the server rate independent of the systems state. Additionally, a linear approach according to *kfB*, $f > 1$ can be used to increase the server rate.

A refined concept of bandwidth allocation is shown in Figure 5. For the sake of simplicity the diagram is confined to the case $N = 3$. Model B is an extension of Model A by splitting the states creating two layers. Each state of the upper layer is characterized by an increased server rate with a difference ΔB compared to the related state of the lower layer. The process being in an upper diagonal state of the lower layer will change to the upper layer for a certain amount of time to decrease the buffer filling level. The change back to the lower layer is mainly determined by an event that is supposed to be supported by the network in order to advertise congestion information. This information is included in the bandwidth acknowledgement messages and is internally treated as a self-created event with the renewal property. Depending on the network congestion level there is a different impact on the duration of staying within the upper layer. If the network congestion level is low, more ATM cells will be delivered thus improving bandwidth utilisation. On the other hand, if at least one of the affected network nodes is getting congested, the duration within the upper layer is shortened increasing the access buffer filling level, whereas the node congestion will be alleviated. The network congestion may be advertised using a range e. g. from 0 to 1. This information is scaled according to the transition rates of the actual state including σ. Thus the duration within the upper layer is mainly determined by the network response. The gain in bandwidth utilisation is proven by the results given in section 4. Additionally, this scheme can be extended to support a congestion avoidance strategy within the network. The impact of the latter is currently examined and is beyond the scope of this paper.

The next step, a further extension of Model B, involves a hysteresis-based congestion control for the access buffer. If the buffer filling level exceeds an upper threshold, a congestion message is sent to the network indicating that the system parameter σ has been changed to a lower value, thus increasing the mean allocated server rate. Notice, that even in this situation the emptying of the buffer will not be enforced but is still controlled by the network. Hence Model C is characterized by the switching between two variants of Model B. The switching back to the former value of σ and the issue of a congestion release message occurs, if the buffer filling level is again below a defined lower threshold. The introduction of a switching hysteresis is important to minimize the number of congestion messages and to avoid rapid and unnecessary changes between the to variants of Model B. The concept of congestion messages that has been introduced implies another category of bandwidth management. Bandwidth reservation is separated into three domains: per call, per burst and per congestion. The announcement of a congestion situation does not necessarily cause a bandwidth renegotiation but can be handled similar to a burst level bandwidth allocation. The impact on other cell streams within the network may be controlled by a reservation mechanism preferring VPs having announced an access buffer congestion.

3. ANALYSIS OF THE PROPOSED MODELS

The components q_{ij} of the transition rate matrix **Q** of the Markov chain for Model A and Model B can be found as shown in Table 1. Let i be a state number as it is given in Figure 4 and Figure 5, respectively. The column number c denotes the number of active sources.

Definition of the q_{ij} for Model A:

for i from 0 to $2N$
for c from 1 to N

if	$i=2N-2c-1$	then	$q_{i,i+2} := c\lambda$
if	$i=2N-2c$	then	$q_{i,i+1} := c\lambda$
if	$i=2c-1$	then	$q_{i,i-1} := c\mu$
if	$i=2c$	then	$q_{i,i-2} := c\mu$
if	$i=2c-1$	then	$q_{i,i+1} := \gamma$

Definition of the q_{ij} for Model B:

for i from 0 to $4N+1$
for c from 1 to N

if	$i=4N-4c-2$	then	$q_{i,i+5} := c\lambda$
if	$i=4N-4c-1$	then	$q_{i,i+4} := c\lambda$
if	$i=4N-4c$ or $i=4N-4c+1$	then	$q_{i,i+2} := c\lambda$
if	$i=4c-2$	then	$q_{i,i-1} := c\mu$
if	$i=4c-1$	then	$q_{i,i-2} := c\mu$
if	$i=4c..$or$..i=4c+1$	then	$q_{i,i-4} := c\mu$
if	$i=4c-1$	then	$q_{i,i+2} := \gamma$
if	$i=4c+1$	then	$q_{i,i-1} := \sigma$

Additionally, there is a transition for $c=0$: $q_{1,0} := \sigma$.

Table 1

Of course, to complete **Q**, the missing transitions are represented by zeros and diagonal elements are finally composed:

$$q_{ii} = -\sum_{\substack{j=0 \\ j\neq i}}^{2N} q_{ij} \text{ and } q_{ii} = -\sum_{\substack{j=0 \\ j\neq i}}^{4N+1} q_{ij}, \text{ respectively.} \quad (1)$$

Let $\mathbf{P}=(p_0, p_1, \ldots, p_l)$ be the vector of probabilities for the $l+1$ states, where $l = 2N$ (Model A) and $l = 4N+1$ (Model B).

The equilibrium probabilities are given by

$$\mathbf{PQ} = \mathbf{0} = (0,0,\ldots,0) \tag{2}$$

and

$$\sum_{i=0}^{l} p_i = 1 \quad . \tag{3}$$

The queuing analysis is based on the model of uniform and continuous arrival and service, also known as fluid flow model. This means, that the packet nature of input and output streams is neglected. Input rate and output rate as well as the buffer filling level and its capacity L are represented by continuous quantities.

Let **D** be a diagonal matrix

$$\mathbf{D}=\mathrm{diag}(d_1, d_2, \ldots, d_l), \tag{4}$$

$$d_i = R_{in_i} - R_{out_i} \tag{5}$$

where R_{in_i} = input rate of state i and R_{out_i} = output rate of state i. (6)

The d_i can be found as shown in Table 2:

Definition of the d_i for Model A:
for i from 0 to $2N$
for c from 0 to N
for k from 0 to N
if (i=2c-1 or i=2c) and (i=2k+1 or i=2k) then $d_i := cB - kfB - B_{stat}$
Definition of the d_i for Model B
for i from 0 to $4N$+1
for c from 0 to N
for k from 0 to N
if (i = 4c-2 or i=4c) and (i=4k+2 or i=4k) then $d_i := cB - kfB - B_{stat}$
if (i=4c-1 or i=4c+1) and (i=4k+3 or i=4k+1) then $d_i := cB - kfB - B_{stat} - \Delta B$

Table 2

It is useful to define the ratio

$$r_m = \frac{\textit{Mean allocated bandwidth}}{\textit{Mean bandwidth requirement}} = \frac{\sum_{i=0}^{l} p_i R_{out_i}}{\sum_{i=0}^{l} p_i R_{in_i}}. \tag{7}$$

As mentioned before the overdimensioning ($r_m > 1$) is a necessity for the stability of the queuing system.

For the regarded superposition of N independent and identically distributed burst silence sources the expression for the denominator in (7) is simply equal to NaB.

Now we are interested in the joint probability $F_i(x)$ of the process being in state *i* while the queue length is less than or equal *x*. The analysis is based on the fundamental work of Anick, Mitra and Sondhi[8] and on the extension to the finite buffer case given by Tucker[9].

$$\mathbf{F}(x) = (F_0(x), F_1(x), \text{K} \ , F_l(x))^T \tag{8}$$

$$\mathbf{DF'}(x) = \mathbf{Q^T F}(x) \tag{9}$$

$$F_i(x) = \sum_{k=0}^{l} a_k \left\{ \mathbf{\Phi}_k \right\}_i e^{z_k x}, \text{ for } 0 \le x \le L, \tag{10}$$

where

z_k: k-th eigenvalue of $\mathbf{D^{-1}Q^T}$

$\left\{ \mathbf{\Phi}_k \right\}_i$: *i*-th component of the right eigenvector $\mathbf{\Phi_k}$ associated to z_k.

and the a_k are constants to be determined from boundary conditions[9]:

The buffer cannot be empty, if the process is in any buffer-increasing state:

$$F_i(0) = 0, \text{ for } d_i > 0 \tag{11}$$

The buffer cannot remain full, if the process is in any buffer-decreasing state:

$$F_i(L^-) = p_i, \text{ for } d_i < 0 \tag{12}$$

The loss probability of state *i* is given by $p_i - F_i(L^-)$ and $\sum_{i=0}^{l} d_i (p_i - F_i(L^-))$ is the amount of loss per unit of time, i.e. the loss rate.

The overall loss probability can be defined as the ratio of loss rate and input rate

$$p_L = \frac{\sum_{i=0}^{l} d_i (p_i - F_i(L))}{NaB}. \tag{13}$$

In the following the Quality of Service is expressed in terms of Packet Loss Probability and the factor of overdimensioning given by r_m. The considered domain is characterized by $r_m = 1.1$, that is priority was given to the enhanced resource utilisation. On the average only 10% of the required bandwidth is additionally booked.

4. RESULTS

As mentioned before the effect of a packetized data stream has been neglected applying the fluid flow analysis. For the computer simulation the choice of a packet length corresponding to a payload of 57 ATM cells has been adopted[6]. This is based on the assumption of a fixed packet length used within a customer premises network that delivers packets to the ATM network. An ATM cell is assumed to carry a payload of 44 Bytes. During an activity period a single source emits data resulting in a peak bitrate of 10.032 Mbit/s. This leads to a packet rate *B* of 500 packets/s. The mean burst length is set to 100 packets and together with a burstiness of *b*=20 we get $\mu = 1/0.2\text{s}$ and $\lambda = 1/3.8\text{s}$. The mean response time of the

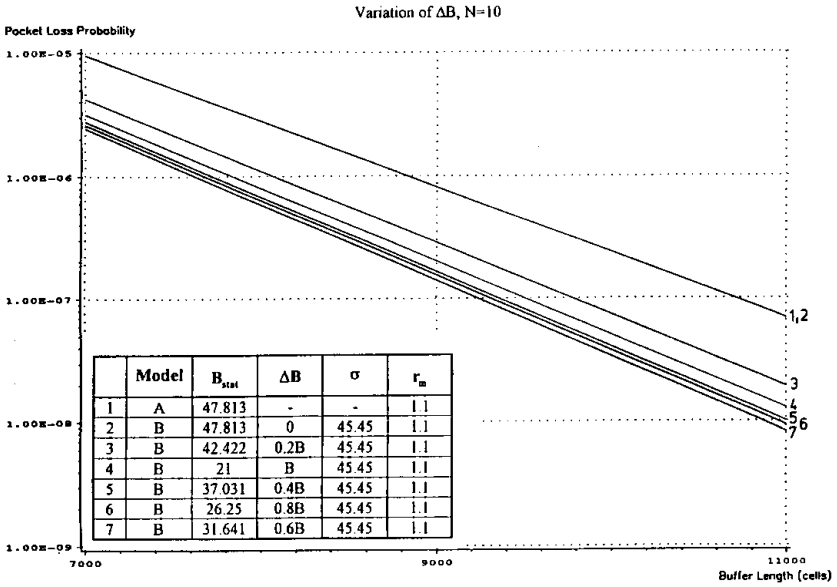
bandwidth acknowledgement messages is set to $1/\gamma = 22\text{ms}$, hence one additional active source will on the average cause an increment of the buffer filling level of 11 packets or 627 ATM cells, if the impact of $\mu-$ and $\lambda-$ transitions is neglected and $B_{stat} = 0$. The mean duration of staying within the lower layer is approximated by $1/\gamma$. The temporarily accumulated amount of ATM cells has to be worked off within the upper layer of Model C. An assessment of this amount is given in (14) for $f = 1$ neglecting the impact of $\mu-$ and $\lambda-$ transitions :

$$(B - B_{stat})\frac{1}{\gamma} = (\Delta B + B_{stat})\frac{1}{\sigma}. \tag{14}$$

The left hand side of (14) denotes the assessment of the mean buffer increment (in packets) while the process is waiting for the bandwidth acknowledgement message, e.g. in state number 2 of Figure 5. The right hand side is an approximation of the mean buffer decrement (e.g. in state number 5, where $\sigma >> \mu + (N-1)\lambda$) assuming that transitions within the upper layer are rare. To outline the impact of ΔB on the Packet Loss Probability analytical results are given in Figure 6 for a constant $\sigma = \gamma = 45.45/\text{s}$. The values for ΔB have been selected within a range between $\Delta B = 0$, where Model B is identical to Model A, and $\Delta B = B$. The static part of the allocated bandwidth B_{stat} is adapted to obtain a constant $r_m = 1.1$. A significant improvement of the Packet Loss Probability is obtained by Model B compared to Model A, if ΔB is appropriately selected. To clarify this, an interpolation for an extended range of ΔB and a Buffer Length of 9000 cells is shown in Figure 7. Obviously the assessment of equation (14) is quite accurate and, as expected, the impact of the transitions within the layers can be neglected.

In the following (14) has been applied with $\sigma = \gamma$ to assess ΔB:

$$\Delta B = B - 2B_{stat} \tag{15}$$



	Model	B_{stat}	ΔB	σ	r_m
1	A	47.813	-	-	1.1
2	B	47.813	0	45.45	1.1
3	B	42.422	0.2B	45.45	1.1
4	B	21	B	45.45	1.1
5	B	37.031	0.4B	45.45	1.1
6	B	26.25	0.8B	45.45	1.1
7	B	31.641	0.6B	45.45	1.1

Figure 6

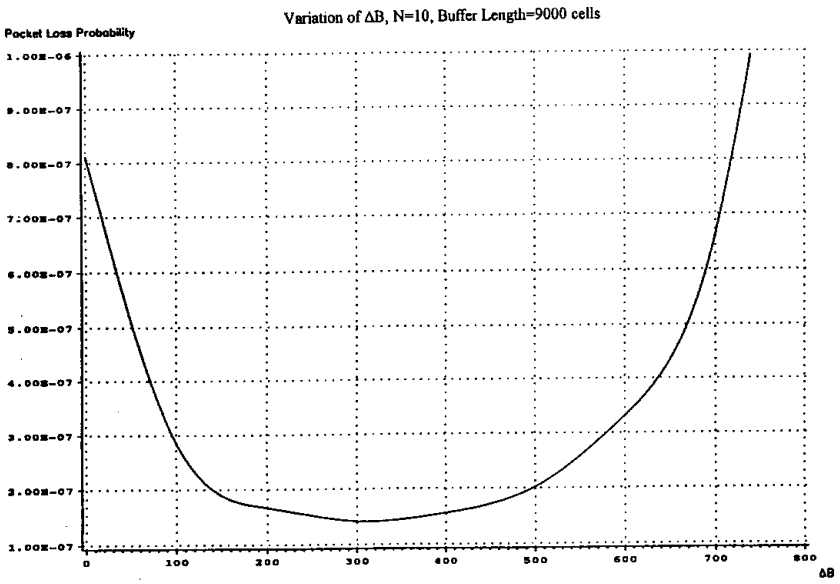
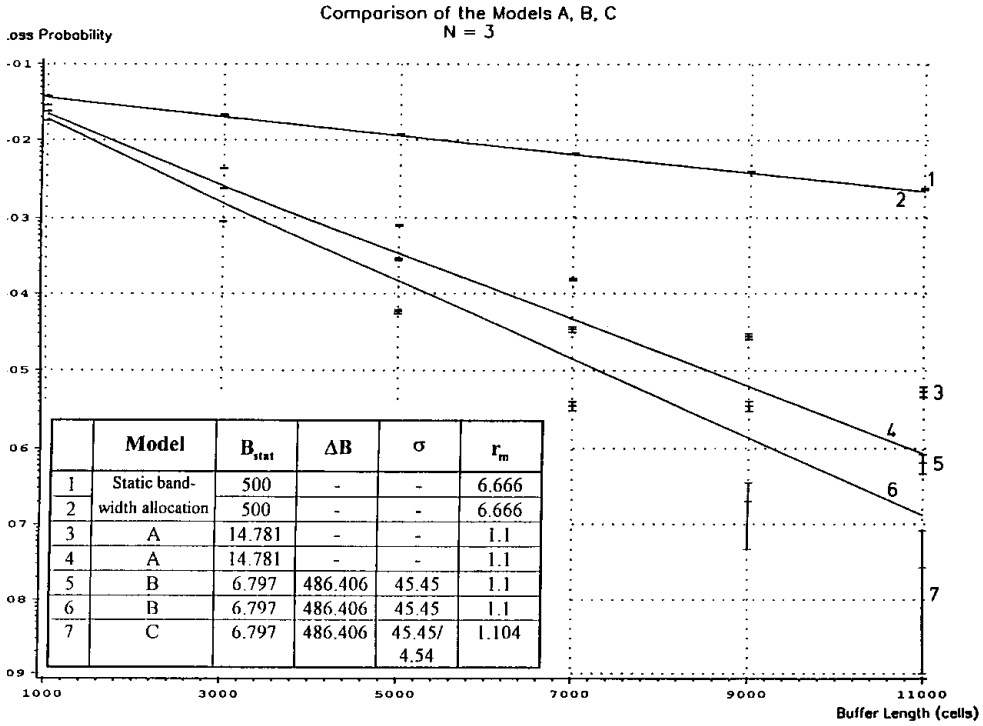


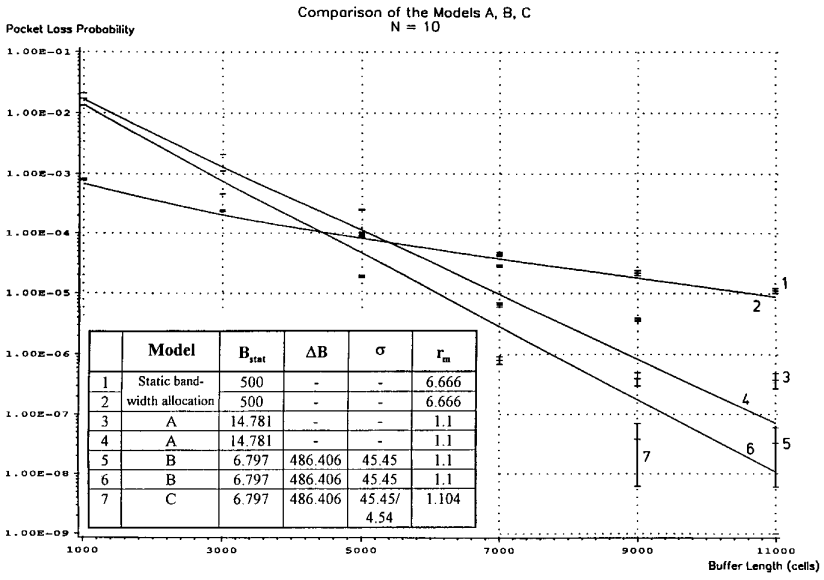
Figure 7

A comparison of the models is shown in Figure 8 and Figure 9 for *N*=3 and *N*=10, respectively. The solid lines (2, 4, 6) represent the calculation of the fluid flow analysis. The results of the computer simulations are given by the related confidence intervals (1, 3, 5). Model C (7) is included, but simulated only. The diagrams clearly demonstrate the underestimation of (2, 4, 6), a well known property of the fluid flow analysis. The comparison of the results for *N*=3 and *N*=10 shows a difference traceable to the effect of traffic mixing. The gain for *N*=10 is very well marked by the static bandwidth allocation (1, 2), where $B_{stat} = 500$ packets/s. This example has been included for comparison purposes. The mean allocated bandwidth was set to $r_m = 6.666$ in order to reach an acceptable Packet Loss Probability. This value is about 6 times higher than the selection for the other models. However, the simple dynamic bandwidth allocation mechanism of Model A (3, 4) turns out to be more efficient for higher values of the Buffer Length. As mentioned before, B_{stat} and ΔB (for Model B) are adapted according to $r_m = 1.1$. The gain of the more complex Model B (5, 6) becomes evident, the Packet Loss Probability is about ten times less compared to Model A. Another factor of a 10 times less Packet Loss Probability is achieved using Model C, where σ is switched between 45.45/s and 4.54/s depending on the buffer filling level. The lower value of σ is applied, if the buffer is getting congested. The switching hysteresis was set to 20% of the Buffer Length and was placed in the middle of the buffer. The measured value of $r_m = 1.104$ shows that the temporary change to $\sigma = 4.54$/s does not significantly affect the mean allocated bandwidth. However, Model C saves about 17% to 20% (for *N*=10) of the Buffer Length compared to Model B achieving the same Packet Loss Probability.



	Model	B_{stat}	ΔB	σ	r_m
1	Static band-	500	-	-	6.666
2	width allocation	500	-	-	6.666
3	A	14.781	-	-	1.1
4	A	14.781	-	-	1.1
5	B	6.797	486.406	45.45	1.1
6	B	6.797	486.406	45.45	1.1
7	C	6.797	486.406	45.45/ 4.54	1.104

Figure 8



	Model	B_{stat}	ΔB	σ	r_m
1	Static band-	500	-	-	6.666
2	width allocation	500	-	-	6.666
3	A	14.781	-	-	1.1
4	A	14.781	-	-	1.1
5	B	6.797	486.406	45.45	1.1
6	B	6.797	486.406	45.45	1.1
7	C	6.797	486.406	45.45/ 4.54	1.104

Figure 9

5. CONCLUSION

A concept of dynamic bandwidth allocation at the B-ISDN network access has been proposed. The model (Model C) is characterized by two variants of a three-dimensional Markov chain (Model B). The switching between the two variants is controlled by the access buffer filling level and its evaluation is based on a switching hysteresis defined by two thresholds. A change between the two variants of Model B causes a bandwidth announcement message or bandwidth release message, a mechanism called bandwidth allocation per congestion. Within the network the messages are handled similar to the bandwidth allocation per burst. Additionally, for this kind of bandwidth allocation a priority scheme may be applied. Both, the analysis and simulation results shows an improvement of the access buffer Packet Loss Probability compared to a straightforward reference model of a dynamic bandwidth allocation (Model A). This improvement is given by a factor of about 10 for Model B and a factor of about 100 for Model C. The improvement achieved by Model C has been pointed out in terms of a decrement of the Packet Loss Probability. The drawback is due to the increased expense for the implementation of a switching hysteresis at the access buffer. Additionally the bandwidth acknowledgement messages have to be supplemented by the network nodes in order to advertise network congestion that is evaluated at the network access. In the present paper the interaction of the access buffer control functions and network congestion management functions has not been considered. The impact of the proposed scheme on the network utilisation will separately be published.

Summarizing the results it has been shown that the advantage of both, the gain achieved by Model B and the additional improvement employing Model C enables a dynamic bandwidth allocation with a very low factor r_m of overdimensioning. The concept of bandwidth allocation per congestion offers a strategy to avoid congestion situations within the network.

6. ACKNOWLEDGEMENTS

The author would like to thank Mr. Hagemann for his interest in this work, the useful discussions and suggestions concerning the application of the employed mathematical software tool. Thanks also to Mrs. Welzel for performing the time-consuming parameter adaptations to obtain a constant r_m and for the graphical presentations.

7. REFERENCES

1. CCITT COM XVIII Recommendation I.211 B_ISDN service aspects, Geneva, 1990

2. CCITT COM XVIII Recommendation I.327 B_ISDN functional architecture, Geneva, 1990

3. A. Baiocchi, N. Melazzi, M. Listani, A. Roveri, R. Winkler,

"Loss performance analysis of an ATM Multiplexer loaded with high speed ON-OFF Sources", IEEE Journal on Selected Areas in Communications, Vol. 9, No. 3, pp. 388-393, April 1991

4. P.E. Boyer, D.P. Tranchier,

"A reservation principle with applications to the ATM traffic control", Computer Networks and ISDN Systems 24 (1992) pp. 321-334

5. M. Gerla, T.-Y. Tai, J.A.S. Monteiro, G. Gallassi,

"Interconnecting LANs and MANs to ATM", Proc. 16th Conference on Local Computer Networks, Minneapolis, Minnesota, USA, October 1991

6. W. Schödl, U. Briem, H. Krämer, T. Theimer,

"A Bandwidth Allocation Mechanism for LAN/MAN Interworking with an ATM Network", 5th IEEE Workshop on Metropolitan Area Networks, Taormina, Italy, May 1992

7. N. Yin, M.G. Hluchyj,

"A Dynamic Rate Control Mechanism for Source Coded Traffic in a Fast Packet Network", IEEE Journal on Selected Areas in Communications, Vol. 9, No. 7, pp. 1003-1012, September 1991

8. D. Anick, D. Mitra, M. M. Sondhi ,

"Stochastic Theory of a Data-handling System with Multiple Sources, AT&T Bell System Technical Journal, Vol. 61, No. 8, October 1982

9. R. C. F. Tucker,

"Accurate Method for Analysis of a Packet-Speech Multiplexer with Limited Delay", IEEE Transactions on Communications, Vol. 36, No. 4, pp. 479-483, April 1988