INTERNATIONAL COMPUTER SCIENCE INSTITUTE
1947 Center St. • Suite 600 • Berkeley, California 94704-1198 • (510) 643-9153 • FAX (510) 643-7684

Managing ABR Capacity in Reservation-based Slotted Networks

Roya Ulrich and Pieter Kritzinger*
{ulrich@icsi.berkeley.edu, psk@cs.uct.ac.za}

The Networks Group
International Computer Science Institute, and
The Computer Science Depatrment
University of Cape Town

TR-96-006
January 1996

Abstract

For slotted networks carrying full multi-media traffic to work successfully, it is essential that connection setup and management is done well under all traffic conditions. Major challenges remain with the current state of the technology, however, particularly on how one copes with traffic bursts. Existing reservation-based networks do not allow the user to dynamically adjust his bandwidth requirements on demand. In this paper we propose a new scheme, called the *reservoir scheme*, which allows dynamic and distributed resource allocation. The basic idea behind the scheme is to reserve bandwidth with a guaranteed bit rate for each virtual circuit. The user is allowed to decentrally allocate additional bandwidth from an Available Bit Rate (ABR) reservoir to satisfy dynamic changes of Variable Bit Rate (VBR) traffic. The duration and bandwidth of this dynamic access are negotiated in the call setup phase and do not require any renegotiation with the service provider so that this solution overcomes the rigidity of current static bandwidth reservation schemes. The additional management requirements are low compared to other dynamic bandwidth reservation schemes. We also describe an analytic model and simulation which we used to determine whether it would be practical to apply the proposed scheme in a slotted network.

*Pieter Kritzinger is in the Computer Science Depatrment, University of Cape Town, private Bag, Rondbosch 7700

1 Introduction

Slotted, multi-media networks are intended to carry traffic as diverse as voice (with a constant bit rate) and video (with a variable bit rate) in either a connection-oriented or connectionless mode. With such a wide range of traffic demands which may vary dynamically, connection management becomes a major challenge. In fact, despite the many proposals (e.g., [NS94, CLG94, LR94, IEE94]) there is, as yet, no confirmation of effective ways to manage the dynamic traffic load with its wide spectrum of service requirements.

The ITU recommendation I.362[ITU90] proposes four service classes which have to be supported by B-ISDN. Of these, connectionless (class D) services are not difficult to manage, but connection-oriented services assume that the service user and provider negotiate with one another in a call setup phase for a guaranteed service. In a connection request, the user has to characterise his own traffic behaviour in advance, by providing values for the various Quality of Service (QOS) parameters required by the network operator, who will then proceed to allocate available resources in either a *static* reservation or *dynamic* reservation mode.

In the case of static reservation, the service provider guarantees and reserves a certain transmission capacity exclusively for the user. In practice, this approach (particularly the peak rate allocation) is attractive for its simplicity. The problem, however, is that VBR traffic, by definition, has variable resource demands and the parameter values characterising the traffic in advance are, moreover, generally estimates only.

Figure 1 illustrates the problem with static bandwidth allocation for two fictive sources. If the bandwidth reservation is based on the average bit rate, data can be lost because of buffer overflow, illustrated by the left hand diagram. If the waiting time in the buffer at the sender side is too high, data can experience delays. Bandwidth reservation based on the peak rate means, in most cases, overallocation and leads, therefore, to wasted bandwidth as illustrated by the right hand diagram.

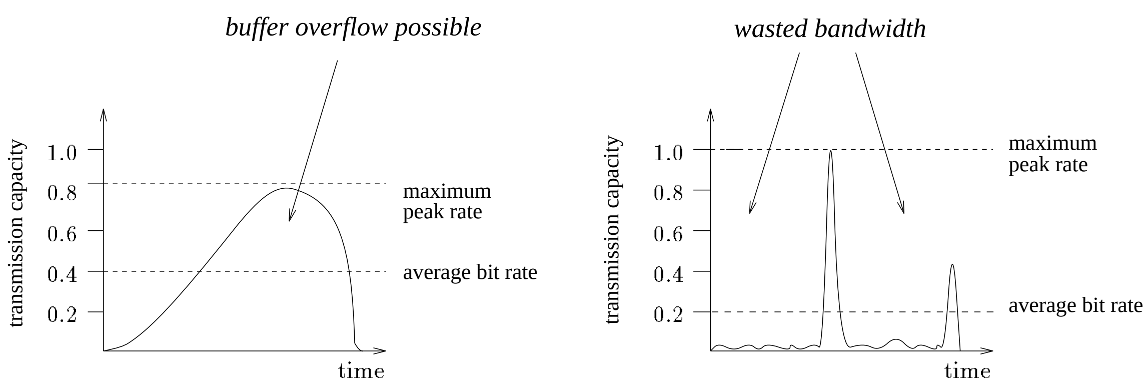

Figure 1: Typical variations in bandwidth requirements and static resource allocation

Several solutions to the problem of dynamic resource allocation have recently been proposed in the literature (e.g., [Sri93, NS94, CLG94]). SMITH AND NAHRSTEDT[NS94] propose a methodology for improved information exchange between the user and the provider by negotiation and renegotiation. They separate the duration of complex and time intensive applications into intervals or *eras* in which service requirements are constant. Renegotiation is required if changes in agreements benefit the application and the network. A new era then begins after the renegotiation. Using a set of QOS parameters, a "correctness condition" is formulated with which a new optimal allocation decision is possible.

CHONG *et al.*[CLG94] present a dynamic bandwidth allocation approach for VBR video traffic only. The authors propose on-line monitoring and prediction of scene changes in a video stream and separate the video traffic into a *low frequency domain* for slow time variation of consecutive scene changes, and a *high frequency domain* for strong autocorrelations between video frames. The authors show that the low frequency signal determines the bandwidth demand, and that for efficient resource usage, adaptive changes of the bandwidth allocation to the low frequency signal are required. Two bandwidth prediction schemes are examined: The first method is based on a relatively complex recursion. The second method uses a time delay neural network which has a lengthy training period (24 minutes CPU time on a SPARC-10 workstation). Other dynamic methods for ATM networks have been reported in the literature [LR94, IEE94].

In this paper, we propose a new scheme, called the *reservoir scheme*, which allows dynamic and distributed resource allocation for slotted networks. In this scheme, the user is allowed to decentrally allocate additional bandwidth to a virtual circuit in order to satisfy dynamic changes of VBR traffic. The duration and maximum capacity of this dynamic access are negotiated in the call setup phase and do not require any renegotiation with the service provider so that this solution overcomes the rigidity of current static bandwidth reservation schemes. The additional management requirements are low compared to other dynamic bandwidth reservation schemes. After describing the scheme in the next section, we outline an analytic model to determine the choice of optimum parameter values in Sec. 3. In Sec. 4 we describe a simulation testbed and the results of an experiment to evaluate the potential effectiveness of the scheme in a slotted network such as DQDB [IEE90].

2 Reservoir Scheme

The basic principle of the *reservoir scheme* is to reserve a minimum bandwidth for loss- and delay-sensitive[1] traffic with a guaranteed bit rate for each user. VBR traffic is, in addition, allowed to allocate bandwidth from a *reservoir* of capacity to accommodate dynamic resource requests decentrally. In other words, as illustrated in Fig. 2, in the reservoir scheme the available bandwidth is shared between the following different service classes:

[1]Called "sensitive traffic" in the remainder of the paper.

Bandwidth for non-sensitive traffic:

This fraction of the bandwidth is reserved for asynchronous, connectionless traffic. The size of the fraction may be reduced by sensitive traffic to a lower predefined, network specific boundary which is necessary for essential traffic such as that for network management.

Bandwidth for sensitive traffic:

This fraction is for the guaranteed bandwidth, connection-oriented services. Because of the bandwidth guarantee for each connection or virtual circuit, the size of this fraction is fixed for the duration of a connection.

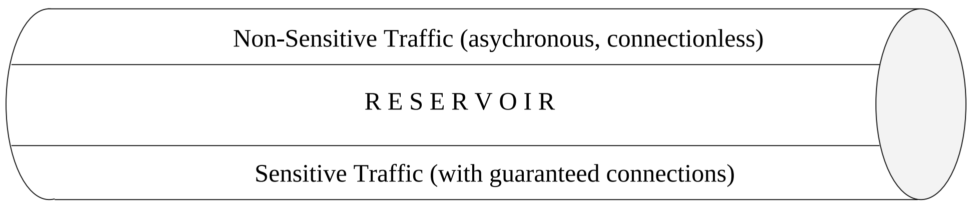

Figure 2: Bandwidth sharing in the reservoir scheme

The remainder of the transmission capacity is referred to as the *reservoir* including all free slots which are not used for asynchronous transmissions or which are not reserved for virtual circuits. Such slots are denoted as *gratis* slots. In a burst phase, gratis slots can be accessed decentrally and dynamically by the traffic class concerned. The mechanism for this is discussed in detail in the following sections.

2.1 Admission Control

Slotted networks require a bandwidth manager process which is the central administrative authority for the establishment and management of virtual circuits and the available bandwidth. The primary function of the bandwidth manager is to decide whether a request with certain QOS requirements can be satisfied. An obvious, necessary condition for accepting a request is that there be sufficient capacity available to allocate to the connection. In addition to bandwidth allocation, the bandwidth manager should check delay requirements of the connection. In the reservoir scheme, admission control depends upon whether

- there is currently sufficient capacity available in the reservoir to satisfy the predicted bandwidth requirements of the application, *also during a burst phase.*
- whether the required delay bounds can be guaranteed, and
- whether the required delay jitter limitation can be met.

A connection is refused if any one of the above conditions cannot be met. The request may then be repeated later.

In order to apply these conditions, the bandwidth manager has to translate the QOS parameters in Table 1 to the physical parameters of the network as listed in Table 2.

Parameter	Description
λ_{max}	peak bit rate
$\lambda_{reserved}$	guaranteed minimum bit rate
ϱ	reservation parameter
D_{max}	maximum tolerable slot delay
D_{jitter}	maximum tolerable slot jitter
L_{max}	maximum tolerable loss rate

Table 1: QOS parameters specified by the user

The reservation parameter ϱ is a new parameter introduced by the reservoir scheme. It determines the amount of the bandwidth to be reserved in addition to the average bit rate $\bar{\lambda}$ known to the user. I.e., from ϱ the user calculates the parameter $\lambda_{reserved}$ in Table 1 from an estimate of λ_{max} and $\bar{\lambda}$ as follows:

$$\lambda_{reserved} := \bar{\lambda} + \varrho \cdot (\lambda_{max} - \bar{\lambda}) \qquad \forall \varrho : \ 0 \leqq \varrho \leqq 1 \ \wedge \ \varrho \in \mathbb{R}$$

The value $\varrho = 0$ corresponds to a bandwidth reservation equal to the average bit rate; $\varrho = 1$ indicates a peak rate bandwidth reservation. The optimum choice of value ϱ is discussed in Sec. 4. Note that there is no need for the user to specify the mean bit rate $\bar{\lambda}$ in the reservoir scheme.

Parameter	Description
L	fraction reserved for network management
$S_{capacity}$	available bit rate in a slot payload (Eq. 3)
S_{number}	number of slots in a frame (Eq. 1)
$S_{propagation}$	slot propagation time
S_{max}	number of slots per frame corresponding to λ_{max} (Eq. 2)
$S_{max,i}$	S_{max} for existing $i-$th virtual channel
$S_{reserved}$	number of slots per frame corresponding to $\lambda_{reserved}$
$S_{distance}$	number of available slots between two reserved slots

Table 2: Physical network parameters

The QOS parameters specified by the user have to translated to values for the physical parameters of the network listed in Table 2. For instance, S_{number}, the number of slots per frame is computed from

$$S_{number} := \left\lceil \frac{C \cdot F \cdot \kappa}{S_{size} \cdot 8} \right\rceil \tag{1}$$

where C is the capacity of the network, F the duration of a frame, S_{size} the slot length and κ a factor which depends on the coding scheme used (e.g., 8/10 coding).

As far as the first question about satisfying the bandwidth requirements is concerned, the bandwidth manager has to prove whether the reservoir can cope with the total volume of the expected sensitive traffic if the peaks occur simultaneously. This requires the translation of λ_{max} into the network parameter S_{max} which gives the number of slots corresponding to the specified peak rate as follows:

$$S_{max} := \left\lceil \frac{\lambda_{max}}{S_{capacity}} \right\rceil \tag{2}$$

where $S_{capacity}$ gives the available transmission capacity of a slot and is calculated from

$$S_{capacity} := \frac{S_{size} \cdot 8}{F} \tag{3}$$

The bandwidth manager then compares the available capacity $(1 - L) \cdot S_{number}$ with the sum of the peak rates of all guaranteed service users, including the one currently being considered:

$$S_{max} + \sum_{i=1}^{N_{VC}} S_{max,i} \leq (1 - L) \cdot S_{number}$$

where N_{VC} is an integer value equal to the number of open virtual circuits. If this quantity is less than or equal to the available capacity, the new request can be allowed.

The second question about guaranteeing the required delay bounds, refers to the translation of the parameters D_{max} and $\lambda_{reserved}$ into the network parameters. The bandwidth manager checks whether the transmission delay between two reserved slots remains below the delay limitation required by the user. A violation of the delay condition leads to the rejection of the request. The criterion for this decision is

$$S_{distance} \cdot S_{propagation} < D_{max}$$

where $S_{distance}$ expresses the maximum distance between consecutive reserved slots, and $S_{propagation}$ represents the propagation time of a slot. The parameter $S_{distance}$ is calculated by taking into account the number of slots assigned to the virtual circuit (the minimum guaranteed bit rate) which is given by

$$S_{reserved} := \frac{\lambda_{reserved}}{S_{capacity}}$$

$S_{distance}$ can then determined by

$$S_{distance} := \left\lceil \frac{S_{number}}{S_{reserved}} \right\rceil$$

As far as the last question about satisfying the delay jitter limitations, the bandwidth manager checks to determine the location of the slots belonging to the new virtual circuit. This depends very much on the slot assignment strategy[2] used by the bandwidth manager to distribute the reserved slots in a framing period. Although bandwidth may be available, the request can be rejected if the jitter limitation cannot be maintained for two arbitrary consecutive slots of the virtual circuit. The problem is particularly crucial for the acceptance or rejection of isochronous service requests. When no jitter can be tolerated, a slot assignment in equal distances is required.

If all criteria can be satisfied, under the reservoir scheme the number S_{gratis} of *gratis* slots per frame is finally computed by the bandwidth manager from

$$S_{gratis} := S_{max} - S_{reserved}$$

The bandwidth manager then informs both parties in the connection of the number S_{gratis} of gratis slot accesses allowed per framing period. The gratis slot access is controlled decentrally at each end of the connection using a *gratis counter* as described next.

2.2 Connection Management

We distinguish between the *static* management of the $S_{reserved}$ slots under the reservoir scheme and the *dynamic* management of the permitted S_{gratis} slots.

2.2.1 Static Reservation

On acceptance of a request, the bandwidth manager determines the slot position within a frame and informs the slot generator about this and the corresponding virtual channel identifier (VCI). The communication partners do not need to be informed about the slot position. Channel access follows merely by comparing their own VCI with the VCI assignments of slots passing the station. This static slot assignment is demonstrated for an example case in Fig. 3, in which the virtual circuit is labelled with the identifier x and is located in the second last slot position in the frame.

2.2.2 Dynamic Access

Apart from these reserved slots, the user can also access S_{gratis} slots during a burst phase in the reservoir scheme. Whether a user is in a burst phase or not can be determined by considering the buffer utilization in the sender station, for example.

[2]Different assignment strategies are discussed in [Gus95].

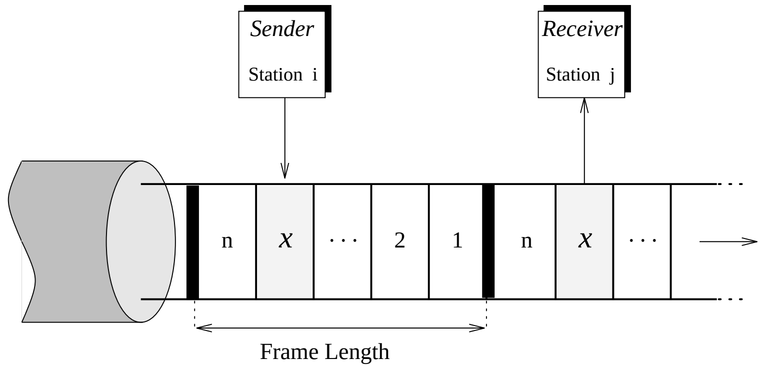

Figure 3: Typical static slot assignment for a virtual circuit

If the buffer length exceeds a predefined threshold, a burst phase can be declared and the user may access no more than the S_{gratis} slots allowed during each framing period. The latter is controlled using a separate *gratis slot counter* for each guaranteed service user of the station. Recognizing a frame start delimiter, the station has to set all gratis counters to the negotiated S_{gratis} value. The counter value is decremented with each gratis slot used. The station may continue to use gratis slots as long as the value of the corresponding gratis slot counter is greater than zero.

A typical example of this behaviour is illustrated in Fig. 4 which show a guaranteed service user in a burst phase. At the start of a framing period the sender sets the *gratis counter* to the negotiated value S_{gratis} (left hand diagram). This means that the next gratis slot can be used if the user is still in the burst phase. The counter value is decremented with each gratis slot used. Although the user is in a burst phase, a gratis slot access is not permitted if the counter value equals zero (right hand diagram).

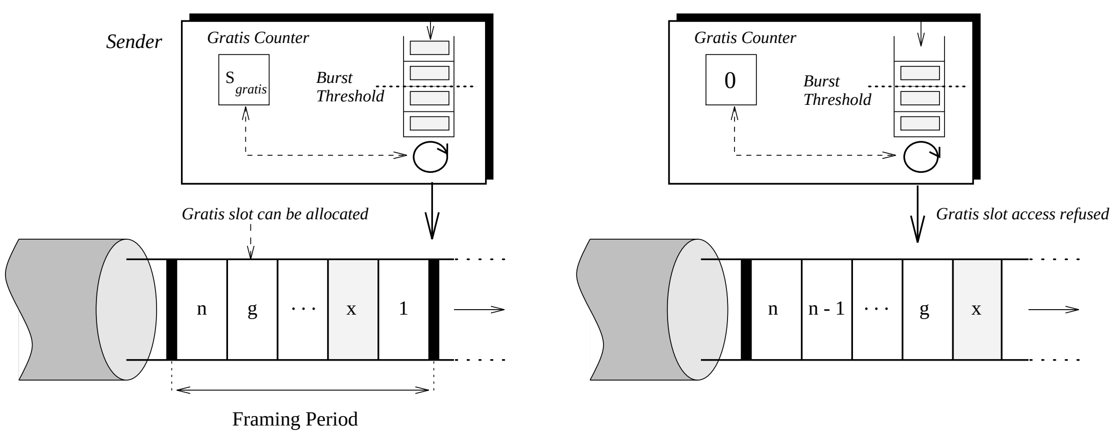

Figure 4: Dynamic slot assignment for a virtual circuit in burst phase

The gratis slot counter mechanism ensures fair behaviour of the guaranteed service users in the scheme. A single heavy user cannot allocate itself the entire available reservoir capacity. Once it has exceeded its negotiated number of gratis slots, it has to let gratis slots pass as illustrated in the right hand side of Fig. 4.[3]

Some conflict situations can occur by scheduling gratis access when the station is involved in multiple time sensitive applications and several service users fulfill the conditions for a gratis slot allocation. Various strategies can be implemented to solve this problem. One possibility is to give the access permission to the virtual circuit with the highest buffer utilization.

3 Analytic Performance Model

In order to determine whether the proposed reservoir scheme could work in practice, we simulated its operation in a slotted ring network. The scheme, however, expects a connection request to offer a value for the reservation parameter ϱ. It is far to expensive in terms of computing time to determine an optimum value of the latter by simulation and we therefore developed an analytic model to study the performance, as a function of ϱ, at a single station. We chose the number of slots waiting to be transmitted at a station as the measure of performance.

3.1 Analytic Model

The basic components of the analytic model are shown in Fig. 5. In that model the waiting line represents the buffer with a finite length of K slots at the sender station. Arrivals are assumed to occur at discrete-time instants, and the arrival process models the behaviour of the correlated traffic of a single VBR source as a double stochastic m-phase Markov Modulated Bernoulli Process with a phase dependent, binomially distributed batch arrival process X.

The server in that figure represents the physical medium, and has a dual character:

- A deterministic (D) (isochronous) service from reserved slots a distance $S_{distance}$ apart (see Sec. 2.2.1), and
- A service (GEO) due to the availability of gratis slots (see Sec. 2.2.2) where there is a probability P_{gratis} given by

$$P_{gratis} = P(T \leq t) = \sum_{t=1}^{S_{distance}} (1-\vartheta) \cdot \vartheta^{t-1},$$

  that a free gratis slot is available before the next reserved slot has passed the sender station. This probability depends on the network utilization $\vartheta \in [0,1]$, i.e., it is influenced by the behaviour of other network users. A 100% network utilization with $\vartheta = 1$ means that no free gratis slot is available, and that data can only be transmitted by the static reserved slots.

[3]A similar fairness mechanism exists for asynchronous services of the DQDB protocol in the Bandwidth Balancing Mechanism.

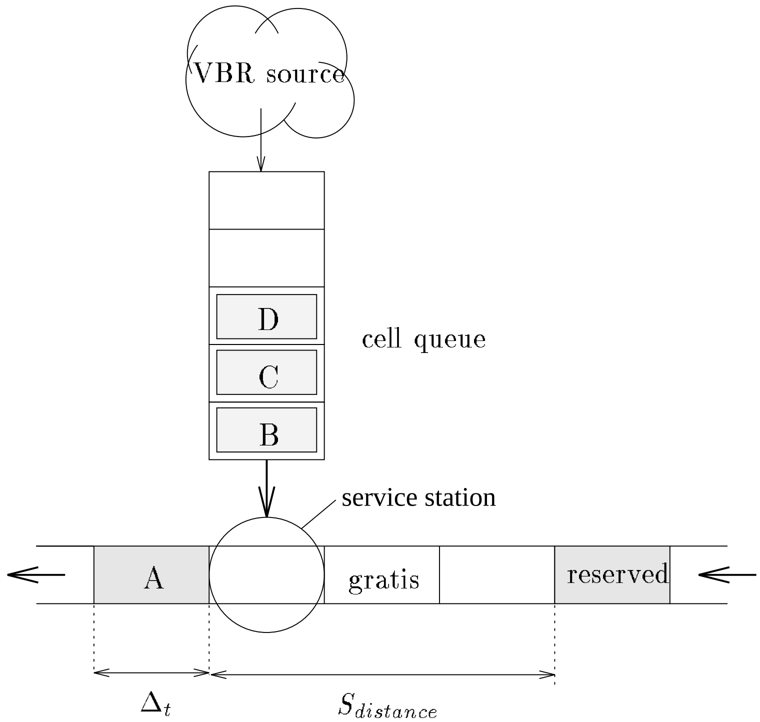

Figure 5: Queueing model of an individual service user

The distribution of the combined $(D + GEO)$ service is given by the discrete-time convolution of the deterministic and the geometric distributions. To analyse the $MMBP^{[X]}_{(m)}/(D + GEO)/1 - K$ queueing model, we used an iterative approach based upon a method first introduced by TRAN-GIA in [TG88] for analyzing $G/G/1 - K$ queueing systems. Unlike various standard continuous-time methods, this approach does not make use of the transition matrix and the corresponding solution of linear equation systems.

In this approach, which we call the *unfinished work approach* (UWA), the process is characterized by the stochastic description of the unfinished work in the system and the queue length distribution is computed directly from the probability mass function of the unfinished work. Other performance measures, such as loss probability, can then be easily computed from the queue length distribution. UWA allows one to efficiently compute the solution of queueing systems with MMBP phase-type arrival processes whose state spaces exceed one million states. The reader is referred to [UHK95] for a detailed description of the analytic model and its validation through simulation.

3.2 Results

We chose a bandwidth of 2.4 Gbps and a standard frame duration of 125 microseconds for the network parameter values used to obtain the analytic results. The parameter values for the load were, in turn, chosen as shown in Table 3 using the notation introduced in Table 1.

Parameter	Value
λ_{max}	40 Mbps
$\bar{\lambda}$	15 Mbps
L_{max}	10^{-9}
D_{max}	4 milliseconds
Burstiness factor	2.7
Burst size	0.5Mb

Table 3: Load parameter values

We first of all chose the network utilisation $\vartheta = 1.0$ to determine the behaviour of a station when no gratis slots are available. The are results given in Fig. 6 which illustrates the queue length distribution for different reservation parameters ϱ.

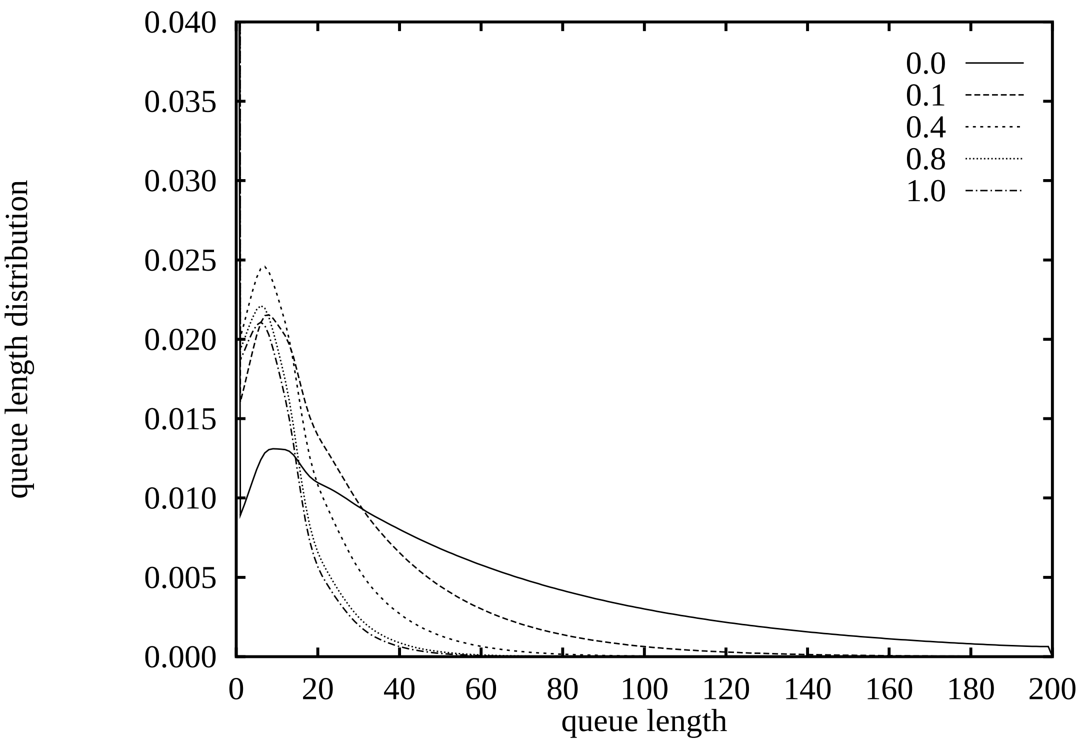

Figure 6: Queue length distribution at a station

Note from that figure that even a small ϱ value of 0.1 already leads to a rapid decrease in the queue length distribution in the figure; for reservation factors close to 1.0 (the peak rate reservation), the improvement is however less marked. Queue lengths greater than 140 have an almost zero probability for values of ϱ of 0.1 or more. As expected, the probability of shorter queue lengths improves with increased values of ϱ, but the improvement is again slight for reservation parameters higher than 0.4 and is not directly proportional to ϱ. The peaks at a queue length of seven in Fig. 6 is caused by binomially distributed batch arrivals.

The main advantage of the reservoir scheme is the fact that a station may access gratis slots dynamically during a burst phase. With $\vartheta = 1.0$ the results reported sofar did not take this into account since in that case there are no free slots. Recall that a station may claim gratis slots once the queue length at the station exceeds a predefined threshold. This threshold is assumed to be half of the maximum buffer size in the analysis. The results are shown in Fig. 7 for network utilizations of 100%, 99%, and 95% respectively. The reservation parameter ϱ was chosen to be 0.4 for reasons which should be apparent from the discussion above, and the maximum buffer size $K = 60$.

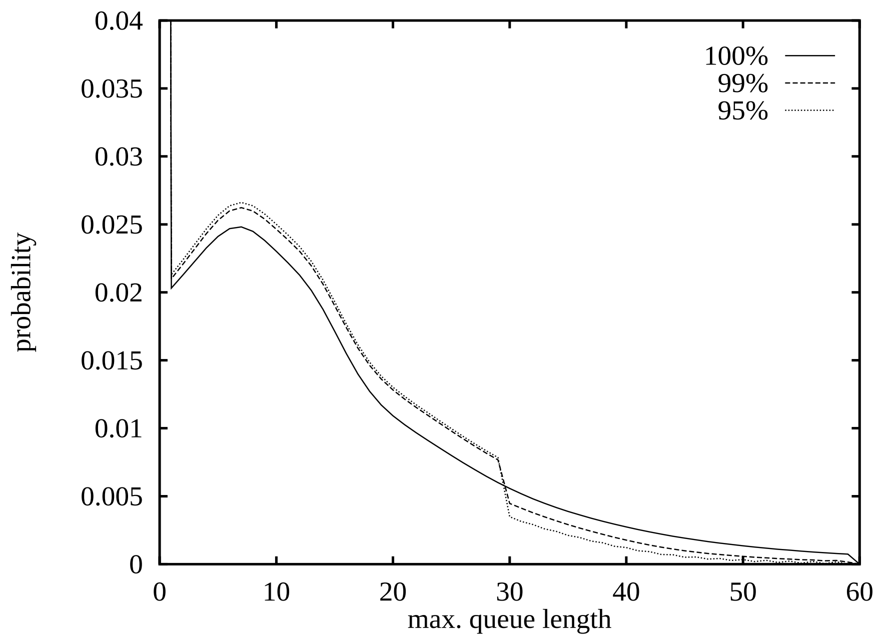

Figure 7: Queue length distribution with dynamic access

The effect of the dynamic access of gratis slot is clearly recognisable from the figure once the burst threshold of 30 slots is exceeded thus confirming that the reservoir scheme has a significant effect on the waiting time of slots at a station.

4 Simulation Model

In the previous section we modelled the behaviour of a single workstation in order to determine its performance under the reservoir scheme for different values of ϑ

for individual applications. In that model the load on the network is represented by the overall utilisation ϑ while, ideally, one should study the scheme under full traffic conditions in a slotted network. This is analytically intractable and we thus attempted a simulation of this scenario. Although it was possible to derive some results from the simulation it also proved impossible to simulate a typical heavy load scenario because there were simply too many events at such high speeds.

4.1 Simulation Testbed

The simulation testbed uses the Object-oriented Simulation of Slotted Communication ARchitectures (OSSCAR) class library [Ulr94], developed at the University of Erlangen-Nürnberg as a simulation tool for performance evaluation of slotted high speed networks. OSSCAR in turn, makes use of the Communication Network Class Library (CNCL) from the Technical University of Aachen [GJW93] which provides a powerful set of C++ classes supporting event-driven simulations.

A simulation process with OSSCAR is based on a technique [Her89] whereby a model of a complex system, the so-called *system model*, can be derived by mapping a *load model* to a *machine model*, where the

- load model represents the behaviour of individual data sources and determines a traffic scenario, and
- machine model represents the (hardware or software) components computing the load.

4.1.1 Machine Model

Using OSSCAR, the simulation testbed was developed for a single ring topology, easily configurable for different parameters, e. g. number of stations, distance between them, buffer size in each station, as well as the network transmission capacity as illustrated in Fig. 8. All stations in the test configuration of that figure could access the network after establishing a connection and any station could simultaneously be an active participant in several connections. Asynchronous VBR traffic, which also allocated the reservoir fraction dynamically, was simulated by the slot generator station. These asynchronous slots circulated in the ring and were removed in the slot generator station. The bandwidth management was performed by a separate process implemented in the slot generator station. For more technical details and further case studies the reader is referred to [Gus95].

The simulation test scenario presented here is based on a proposal[XVI91] from an ITU study group. The network connected 32 active stations which were arranged at equal distances on a ring. The total length of the medium was assumed to be 10 Km. The stations were numbered sequentially from 0 to 31 with the slot generation in station 0. The network provided a transmission capacity of 1.2 Gbps, where 10% of this capacity is reserved for high priority traffic such as network management operations. The slot size was chosen to be 53 bytes with a 48 byte payload.

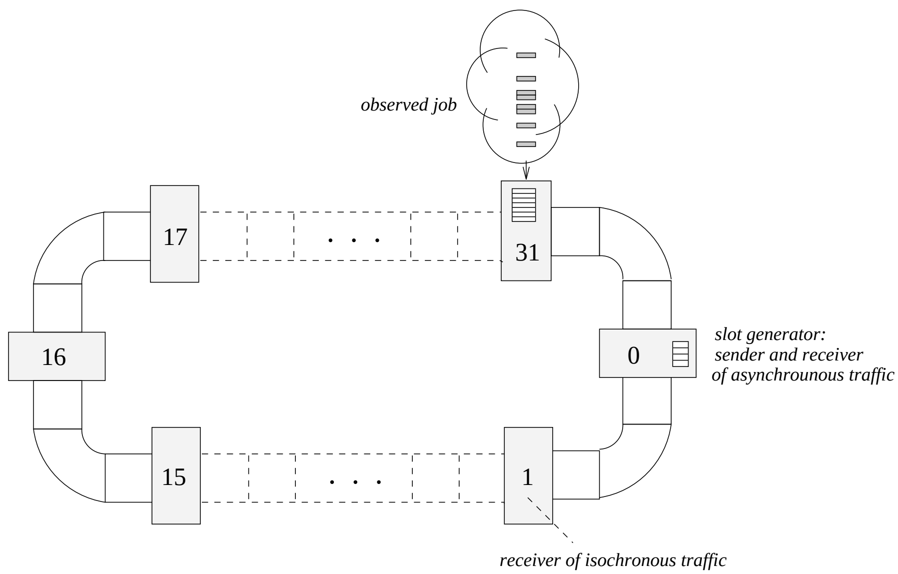

Figure 8: The network configuration used in the simulation

4.1.2 Load Generation

The other major part of OSSCAR is the load generation module. There are 13 load generators available representing the most common stochastic processes, e. g., *on-off*, *autoregressive*, and different *Markov Modulated* processes.

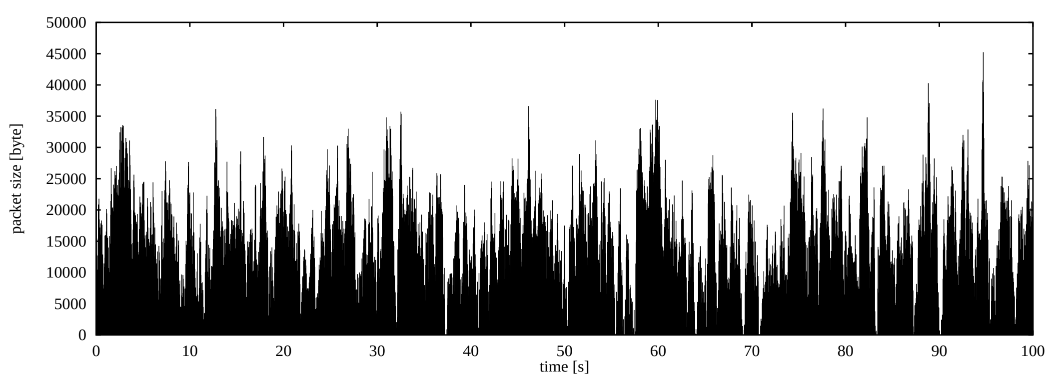

Figure 9: Output of load generator

A mix of VBR and CBR traffic with different bit rate requirements was used for the load on the network in the testbed. A total of 60 virtual circuits were established and were active during the whole simulation. The load scenario furthermore assumed that all nodes ranging from 2 to 31 communicate with node 1 behind the slot generator. We selected station 31 to monitor, just ahead of the slot generator. This represents a worst case scenario, in that if nodes before the monitor station go into a burst phase and consequently use reservoir space, the observed job will find a much reduced number of available gratis slots.

The monitor station is also the source of data for the application for which data are collected during the simulation. We chose the load generator illustrated in Fig. 9 with an average bit rate of 3.9 Mbps and a peak rate of 10.575 Mbps for this time sensitive application.

Using the analytical model described above we decided that in this experiment the quantitative requirements can be satisfied even by a small reservation parameter ϱ (a reservation close to the average bit rate) and small buffer sizes. Therefore, we chose a reservation parameter of $\varrho = 0.4$. The maximum buffer size K was set to 20 slots, and a gratis slot access was allowed as soon as the buffer contained more than 10 slots, i. e. the burst threshold was 50% of the maximum buffer size. The exact details of the experiment are described in the dissertation by Ulrich [Gus95].

4.2 Results

The simulation was run for 8000 framing periods, corresponding to an elapsed time of 1 second. This required about 3 hours and 50 minutes on a SPARC5 workstation of which around 45 minutes was required to generate the complete configured load scenario. The results in Fig 10 illustrate the probability distribution function of gratis slots per frame. Note that all the results were measured at the monitor station and that all measures are given with a confidence interval of 95%.

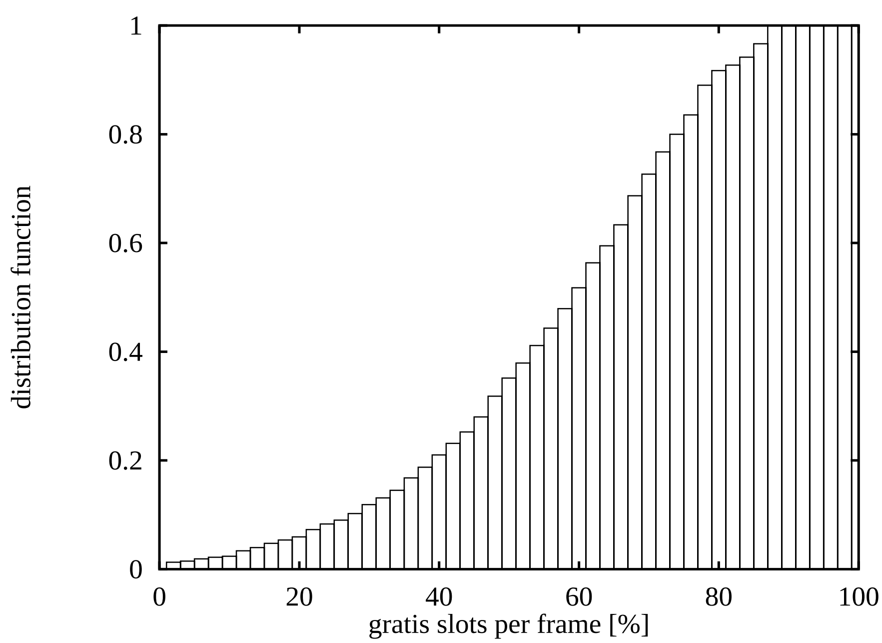

Figure 10: Probability distribution function of the number of gratis slots per frame at the observed station

Despite what appears to be a heavy traffic configuration with 60 connections and

60 asynchronous service users, the average network utilization was measured at 32% only (with a confidence interval of $\pm 4.40\%$). A drastic increase in the number of active load generators were required to obtain higher utilizations which increased the computing time to the point where it became unfeasible to do.

From Fig. 10 it appears that the probability distribution of available gratis slots has a binomial distribution, which confirms the assumptions made for the analytical model in the previous section. In any event, the mean value of the number of gratis slots per frame was 55% so that on the average more than half of the frame capacity was available, thus confirming that at fairly low network utilisation the principles of the reservoir scheme can be applied to utilise the available free bandwidth.

5 Conclusions

In this paper, we presented a new scheme, called the *reservoir scheme*, which allows dynamic and distributed resource allocation for slotted networks. Under the scheme, the user is allowed to decentrally allocate additional bandwidth to a virtual circuit in order to satisfy dynamic changes of VBR traffic. The duration and bandwidth access are negotiated in the call setup phase and do not require any subsequent renegotiation with the service provider and thus overcomes the rigidity of the static bandwidth reservation schemes. Fair use is ensured with a scheme similar to that used in the *Bandwidth Balancing Mechanism* for DQDB networks. The additional management required for the reservoir scheme is low compared to that required for other dynamic bandwidth reservation schemes.

That the proposal may work in practice was illustrated with an analytical model of individual workstations on a slotted network carrying correlated traffic with binomially distributed burst sizes. The network traffic is represented by the overall utilisation of the bus and not by the behaviour individual stations on the network. The latter scenario was modelled using a simulation testbed which proved again that sufficient free slots would be available to make the reservoir scheme work. Another result of this experiment was the discovery that it is just unfeasable to simulate heavy traffic conditions to the detail required to determine the feasibility or otherwise of the proposed connection management scheme.

References

[CLG94] S. Chong, S. Li, and J. Ghosh. Dynamic Bandwith Allocation for Efficient Transport of Real-Time VBR Video over ATM. In *Proceedings of the IEEE INFOCOM Conference*, pages 81–90, 1994.

[GJW93] C. Görg, M. Junius, and B. Walke. *Hand-on-Tools Praktikum: BOOST/CNCL*. RWTH Aachen, Oktober 1993. Version 0.14.

[Gus95] R. Gusella. *Reservoir-based Resource Management for Slotted High Speed Networks.* PhD thesis, Informatik VII, Universität Erlangen–Nürnberg, 1995.

[Her89] U. Herzog. Leistungbewertung und Modellierung für Parallelrechner. *Informationstechnik (it)*, 31(1):31–38, 1989.

[IEE90] IEEE 802.6: IEEE Standards for Local and Metropolitan Area Networks. *Distributed Queue Dual Bus (DQDB) Subnetwork of a Metropolitan Area Network (MAN)*, 1990.

[IEE94] IEEE Computer Society, IEEE Communications Society. *Networking for Global Communications*, volume 1 of *Proceedings of the IEEE INFOCOM Conference*, 1994.

[ITU90] ITU. *Recumendation I. 362: B-ISDN ATM Adaption Layer (AAL) Functional Description*, 1990.

[LR94] J. Labetoulle and J. W. Roberts, editors. *The Fundamental Role of Teletraffic in the Evolution of Telecommunications Networks*, volume 1a of *Proceedings of 14th International Teletraffic Congress, ITC 14*, 1994.

[NS94] K. Nahrstedt and J. M. Smith. Revision of QoS Guarantees at the Application/Network Interface. Technical report, University of Pennsylvania, Distributed Systems Laboratory, PA 19104-6389, 1994.

[Sri93] K. Sriram. Methodologies for Bandwidth Allocation, Transmission Scheduling, and Congestion Avoidance in Broadband ATM Networks. *Computer Networks and ISDN Systems*, 26(1):43–59, 1993.

[TG88] P. Tran-Gia. Zeitdiskrete Analyse verkehrstheoretischer Modelle in Rechner- und Kommunikationssystemen. 46. Bericht über verkehrstheoretische Arbeiten, Universität Stuttgart, 1988.

[UHK95] R. Ulrich, U. Herzog, and P.S. Kritzinger. Modelling buffer utilisation in cell-based networks. Technical Report /95, Informatik VII, Universität Erlangen–Nürnberg, 1995.

[Ulr94] R. Ulrich. Simulationsumgebung OSSCAR. Technical Report 19/94, Informatik VII, Universität Erlangen–Nürnberg, 1994.

[XVI91] CCITT Study Group XVIII. Temporary Document 34 (XVIII/8). Meeting Report of SWP 8/1 ATM, 11-28 June 1991.