Reservoir-based ABR Services in ATM Networks

Wolfgang Frohberg[1] and Roya Ulrich
{wfrohber, ulrich}@icsi.berkeley.edu
The Networks Group
International Computer Science Institute
1947 Center Street., Berkeley, CA 94704
TR-96-024
July 1996

Abstract

ATM technology tends to be the major networking technology for the Broadband-ISDN. Motivated by the growing amount of Internet traffic, which will be carried over ATM networks, we extend the *reservoir-based resource management* proposed in [ULR95] to ATM networks, where it can be used to provide an available bit rate (ABR) bearer service. ABR is connection-oriented and performs a variable bit rate data transport without timing constraints. The *reservoir-based ABR scheme* (ResABR) proposed in this paper assigns bandwidth on demand of the sources, taking into account the network utilization at the request time. The basic idea of ResABR is to divide virtual paths into two logical parts. One part contains the bandwidth necessary to guarantee a minimum bandwidth for all connections. Another part of each VP provides a reservoir of extra bandwidth, which can be used by one or more of the ResABR connections for a short time to send bursts. The advantages of the ResABR scheme are: resource management actions are necessary only when a burst occurs; no extra storage of cells inside the network is necessary; the scheme is robust and it provides less computational effort; it is fair between sources.

---

1. Wolfgang Frohberg is with the ALCATEL Telecom Research Division - Location Stuttgart, D-70430 Stuttgart, Germany

1 Introduction

ATM technology tends to be the major networking technology for the Broadband-ISDN. It has the capability to meet Quality of Service and performance requirements of different application domains by using one single technology in the network.

ATM equipment vendors and network operators have started to design and to develop basic transport capabilities in ATM networks. As proposed in [ITU90] these bearer services are based on four traffic classes, with different requirements concerning timing behavior, bit rate, and transmission mode, as shown in table 2.

TABLE 1. Traffic Classification

	Class A	**Class B**	**Class C**	**Class D**
Timing	required		not required	
Bit Rate	constant	variable		
Transmission Mode	connection-oriented			connectionless

Up to now there have been bearer services covering classes A, B and C under discussion:

- Constant Bit Rate (CBR) for class A;
- Variable Bit Rate (VBR) for class B;
- Available Bit Rate (ABR) and Unspecified Bit Rate (UBR) for class C.

ABR is a connection-oriented service, which performs a variable bit rate data transport that does not have timing constraints. It guaranties a minimum of bandwidth to connections and uses transmission capacity temporarily available in the network to extend the minimum bandwidth.

The ABR principle assumes that the network is the active part. That means the network notifies ABR service users about the available resources. The ABR mechanism assumes, that traffic sources demand to send a certain amount of data as quickly as possible. However, there are a lot of applications which do not have this demand. One example is image browsing where a human browses through an image database. The real-time requirements of such applications are not too strong. Additionally, the sending end (the image database) does not send continuous data, especially when the human is viewing the image and deciding what to do next. Such applications require transmitting a burst of data from time to time (bulk sources). With increasing interest in the world wide web, such applications build the major part of Internet traffic.

Hence, it is not very efficient to use a transport service in which the network always tells all sources about the resources they can get regardless of their current needs.

In [ULR95], a scheme is developed which enables available capacity to be shared dynamically depending on traffic changes without any negotiation. The proposed so called reservoir-scheme allows users to access excess bandwidth (or a reservoir range) during burst phases. The reservoir scheme is basically developed for slotted shared medium networks.

Motivated by the growing amount of Internet traffic, which will be carried over ATM networks, we extend the reservoir-scheme to switched networks like ATM networks. Additionally, the extension requires the consideration of asynchronous and non-framed multiplexing of ATM networks. With these extensions, a dynamic and robust ABR bearer service can be provided which is also easy to implement.

The reservoir-based ABR scheme (ResABR) proposed in this paper assigns bandwidth to connections *on demand of the sources* taking into account the network utilization at the request time. Although the ResABR scheme can handle other traffic sources, it has special advantages with bulk sources.

The rest of the paper is organized as follows: First, we describe the ABR service as it is in the quasi standardization process of the ATM Forum (section 2). Following, the basic ideas of the reservoir-based resource management for slotted high speed networks are summarized (section 3). We then give a description of the reservoir-based ABR scheme (section 4). We finally end up with conclusions and some remarks on future work (section 5).

2 Available Bit Rate Service

ABR is a connection-oriented service, which performs a variable bit rate data transport without timing constraints. The ABR service requires connection set up and release. Once a connection has been established, the network guarantees a minimum of bandwidth to this connection. The network uses transmission capacity temporarily available to provide additional bandwidth. The available capacity is shared among connections. This can be done rate-based or credit-based.

The basic idea behind both principles is that the network is the active part. The network tells the traffic sources how to behave. The ABR mechanism assumes that there is the demand of the traffic source to send a certain amount of data as quickly as possible.

ATM Forum and ITU-T specify ABR service as a rate based service, which controls the bandwidth of connections directly [BFG95], [ITU96]. According to these specifications, feedback from network switches gives the end systems the information necessary to respond to changes in the available bandwidth, by modifying their transmission rates.

Establishment of an ABR connection requires a call setup. During the call setup, the values for a set of ABR-specific parameters ar e negotiated with the network. These parameters are:

- a minimum cell rate (MCR) as the lower bound for the source rate,
- a peak cell rate (PCR) as the upper bound for the source rate,
- an initial value for the allowed cell rate (ACR),
- an increment to increase ACR, and
- an increment to decrease ACR.

Some of these parameters are requested by the source and possibly modified by the network, e.g. MCR and PCR, while others are directly chosen by the network, e. g. ACR and its increments.

Once call setup is completed and a virtual channel is established, the source begins with the cell transmission. The rate at which a source is allowed to send cells is denoted by ACR which is always restricted between MCR and PCR. The ACR value may vary during a connection depending on the changes of available network resources.

The transmission of user cells begins with sending a resource management cell (RMC). The source rate is controlled by the return of these RMCs, which contain an update of ACR after they returned from the network. The source will periodically send RMCs, typically after a certain amount of data cells.

RMCs contain the following information:

- the current cell rate (CCR) at which the source is sending at the moment,
- the explicit cell rate (ECR) at which the source wishes to send (usually its PCR), and
- a congestion indicator (CI) which indicates a congestion in the intermediate switches along the virtual channel.

The source places its ACR in the CCR field, and fills in the ECR field. Then the RMC travels forward through the network, thus providing the switches in its path with the information in its fields. This information is used for bandwidth sharing among connections. Switches may decide at this time to reduce the value of ECR if they are unable to maintain this rate for the connection. They also may set CI to 1 if they detect congestion.

When the RMC arrives at the destination its direction changes and it returns to the source. As the RMC travels backward through the network, switches may update ECR, too. Switches should modify ECR on either the forward or backward journeys of RMCs, but not on both.

When the RMC arrives back at the source, the source resets its ACR, based on the information carried by the RMC. If CI is not set, the source increases its ACR by the fixed incr ement determined at the call setup, towards (or up to) the ECR value returned, but never exceeding PCR. If CI is set, the source must decrease its ACR by an amount which is also determined at the call setup. If ACR is greater than the returned ECR, the source must decrease its ACR to the returned ECR, although never below MCR.

Note that the available bandwidth for a certain connection depends on the load situation in the network segments which are involved in this connection. Whenever a connection is established or released, that effects more or less a large part of the network. Major effects are on connections sharing the same network links; minor effects are on connections which share links with effected connections.

The mechanisms applied in the switches to reassign available bandwidth (resulting in an ECR change) must be carefully chosen to be stable, efficient, fast and fair to all connections.

3 Reservoir -based Resource Management

In [ULR95] a new scheme for resource management is proposed, called the *reservoir scheme*, which allows dynamic and distributed resource allocation in slotted high speed networks.

The basic idea behind the reservoir scheme is to reserve bandwidth with a guaranteed bit rate for each virtual channel (VC). The user is allowed to distributed allocate additional transmission capacity from an excess bandwidth or reservoir range 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.

In the reservoir scheme the available bandwidth is shared between the following different service classes:

- Bandwidth for non-sensitive traffic: this fraction of the bandwidth is r eserved for asynchronous, connectionless traffic.
- Bandwidth for sensitive traffic: this fraction is for the guaranteed bandwidth, connection-oriented services with guarantied bandwidth. Because of the bandwidth guarantee for each connection or VC, the size of this fraction is fixed for the duration of a connection.

The remainder of the transmission capacity is referred to as the *reservoir* including all free slots which are not used for connectionless transmissions or which are not reserved for VCs. Such slots are denoted as *gratis slots*. In a burst phase, gratis slots can be accessed decentralized and dynamically by the traffic class concerned.

Slotted networks require having a central timing source which is responsible for the synchronization of slot flow on the medium. Slots are continuously generated in constant-length time periods, called framing periods or frame in short, by a slot generator station. The resource management is performed centrally and the slot generator station acts ideally as bandwidth manager. The bandwidth manager (BM) process is the central administrative authority for the establishment and management of virtual circuits and the available bandwidth. The primary function of the (BM) 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 BM should check delay requirements of the connection. 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.

On acceptance of a request, the BM 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.

Apart from these reserved slots, the user can also access 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. If the buffer length exceeds a predefined threshold, a burst phase can be declared and the user may send a number of 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. Once it has exceeded its negotiated number of gratis slots, it has to let gratis slots pass. The gratis slot counter mechanism ensures fair behavior of the guaranteed service users in the scheme. A single heavy user cannot allocate itself the entire available reservoir capacity.

In that way, the reservoir scheme overcomes the rigidity of the static bandwidth reservation based on average or peak bit rate. The additional management required is low compared to the dynamic bandwidth reservation scheme which react adaptive to the changes in the traffic as well as to altering network capabilities.

4 Reservoir ABR scheme (ResABR)

The reservoir scheme has been basically developed for slotted shared-medium networks. ATM networks will be mostly wide spread meshed networks consisting of a large number of nodes and links. Figure 1 illustrates a part of such a network consisting of five A TM switches *A* to *E* and three connections already established. The sketch also contains the physical links, which are necessary to carry the traffic of the existing connections. They ar e illustrated by solid lines. All explanations of the ResABR scheme will be based on this example network.

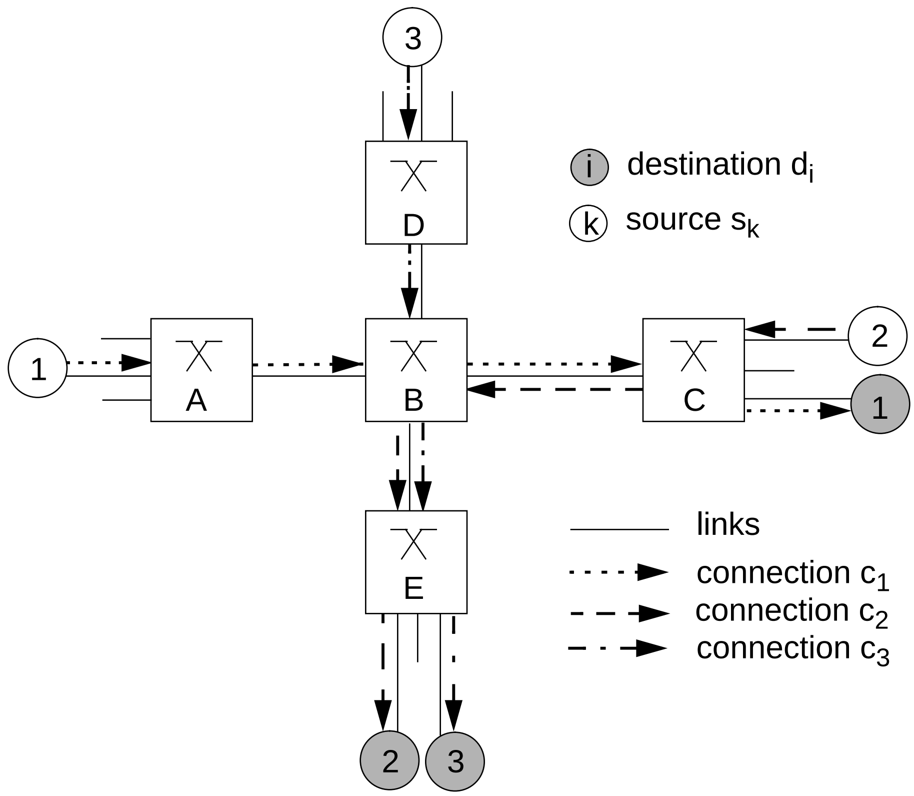

FIGURE 1. An Example Network

Different connections go along different routes through the network; links are shared by different connections. For example c_1 goes from s_1 via the switches **A**, **B** and **C** to d_1, c_2 between s_2 and d_2 passes the nodes **C**, **B** and **E**. c_3 shares network resources on link **BE** with c_2.

Bandwidth access on any link by a certain connection has to take into account other connections passing the same link. The ResABR scheme is able to handle this with a minimum management effort based on the burst demands of connections.

As mentioned in the previous section, the original reservoir scheme is based on frame based transmission protocols. As ATM networks do not provide any protocol elements for time synchronization, and cell multiplexing is done asynchronously, we extend the reservoir scheme with mechanisms to cell access in burst situations in switched ATM networks. We call these extensions reservoir based ABR (ResABR). For the realization of ResABR services, we assume the network consists of a logical virtual paths (VP) subnetwork including VCs for each ResABR connection. Again, consider the network presented in Figure 1, we have now a VP subnetwork as shown in figure 2.

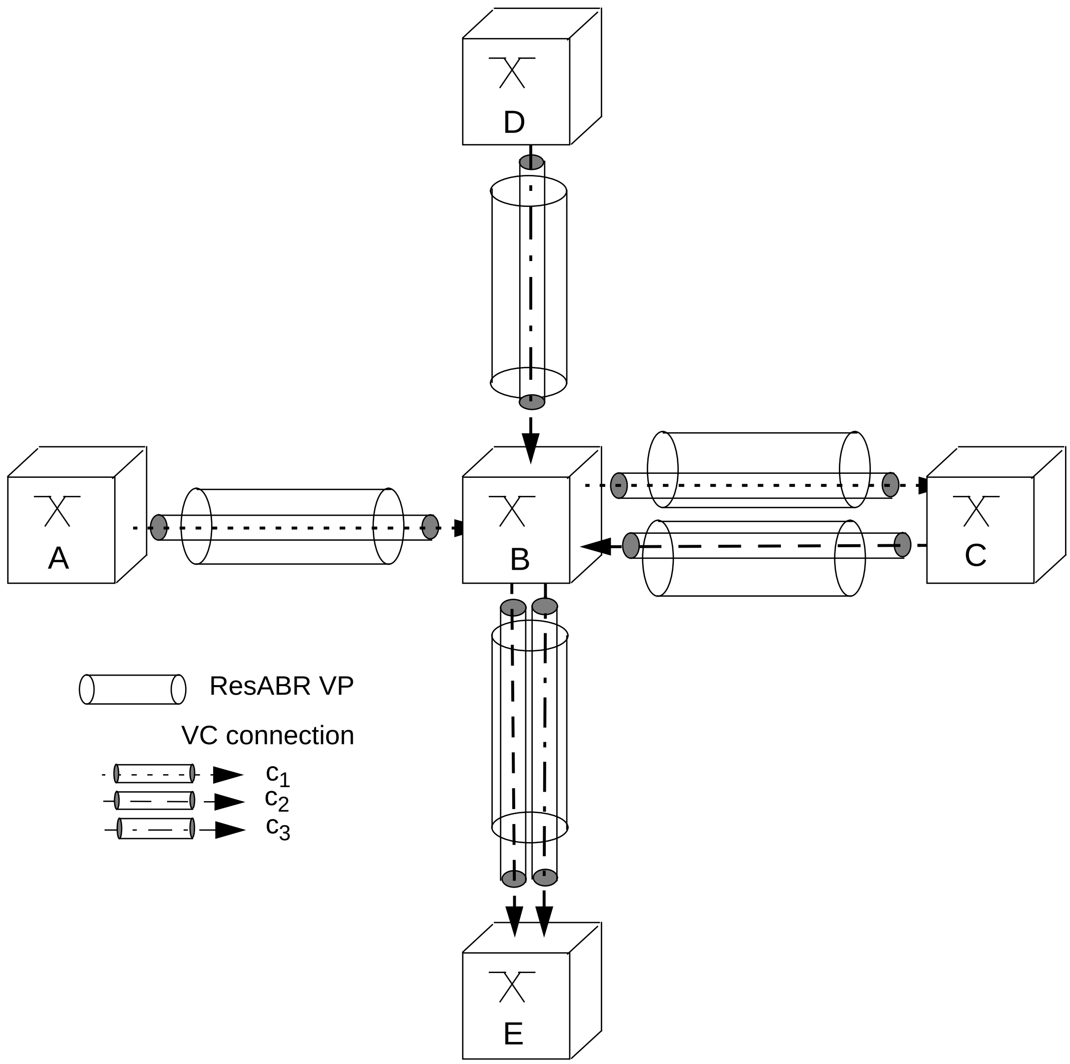

FIGURE 2. ResABR VP Subnetwork

Between the nodes there are directed physical connections, as shown from node **A** to node **B**. These nodes also contain VPs not necessarily belonging to the ResABR logical VP network. The VPs used for this example are referenced by their beginning and target node, e.g. the VP named **BC** is the VP carrying ResABR traffic from node **B** to node **C**. Note that the ResABR traffic from node **C** to node **B** is carried in another VP, named **CB**. As both physical connections and VPs are directed, **BC** and **CB** belong to different physical connections. The transmission rate assigned to each of the VPs in the network might be changed by the network management but should be stable during a reservoir usage. ResABR traffic is only carried in ResABR VPs.

The basic idea of ResABR is to divide its VPs into two logical parts. One part contains the bandwidth necessary for guaranteeing the MCR of all connections. Another part of each VP provides a reservoir of extra bandwidth, which can be used by one or more of the ResABR connections for a short time to send bursts.

The parameters used to control ResABR connections are

- the minimum cell rate λ_{min} which is guarantied to the connection for the whole connection time,
- a maximum cell rate λ_{max} which is the upper bound for any cell rate during a burst,
- an actual cell rate λ,
- a maximum burst duration θ,
- a time-out τ_1 for waiting between unsuccessful burst requests, and
- a time-out τ_2 for waiting before sending the next burst request after a successful burst.

As the VP bandwidth is limited, the reservoir part of the VP bandwidth depends on the number of existing connections and their λ_{min}. Figure 3 explains the composition of the ResABR VP. In the lower part the λ_{min} of some connections is shown, while the upper part contains the reservoir.

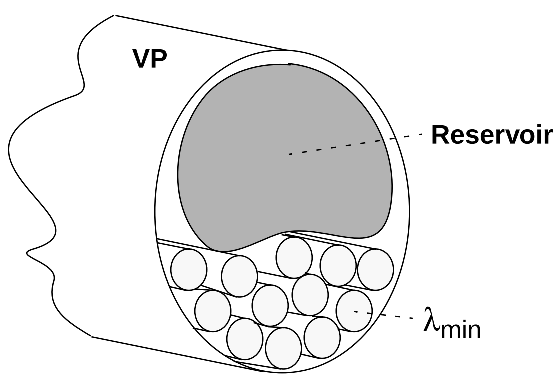

FIGURE 3. Composition of a ResABR VP

To each ResABR VP there is attached a bandwidth manager. It deals with the bandwidth partitioning between the portion for all λ_{min}s and for the reservoir in the call setup phase but mainly it deals with the reservoir assignment to different connections when they want to send bursts. So we call it Reservoir Bandwidth Manager (RBM).

In the following section the ResABR scheme is described in two parts as it deals with two levels of controlling the concerned traffic:

- the establishment of a connection (connection admission control), and
- the actions performed when a sender station wants to send a burst (dynamic access).

4.1 Admission Control

Connection admission control is initiated by the source. We describe the set up of one unidirectional connection, although there are usually two such connections involved in a call. ResABR applies for a single unidirectional connection either on admission control level or on dynamic access level.

λ_{min} depends on the application. It's the cell rate necessary to operate the application without congestion when no bursts occur. Typically λ_{min} is between 0 and λ_{max}. In the case of image transmission λ_{min} represents the rate needed for an alive signal in-between image transmission. It can be very low. If just one VP along the route cannot provide λ_{min}, the connection request will be rejected. Reasons for not being able to provide the minimum cell rate are less bandwidth resources in at least one of the VPs which is used by the connection.

λ_{max} is the rate with which the source intends to send after it got the allowance to send a burst. This parameter can be negotiated with the network. The upper limit of λ_{max} is the bandwidth available in the reservoir. If the negotiation of λ_{max} is not successful, the connection request will be rejected.

The burst duration θ is a parameter either given by the network or negotiated between network and source. Although θ is dependent on the application, it cannot be determined by the source only. The network has to ensure fairness between all connections, as the basic idea is to give the reservoir bandwidth or at least a portion of it to one connection for the burst duration.

4.2 Dynamic Access

Dynamic access on the reservoir range permits the user to allocate additional capacity in a traffic burst phase. Sending a burst pr e-supposes a burst request. A burst request can be indicated by reaching a certain buffer limit at the source, for instance. To request a burst, the source generates a burst reservation cell (BRC) containing the cell rate λ with which the source wishes to send. λ must not exceed λ_{max} as negotiated at call setup.

The BRC travels trough the network along the route of the corresponding connection. It passes all RBMs of the VPs of the connection. Each intermediate RBM tests if it can provide the required λ from its reservoir. If it cannot provide any bandwidth, it sets λ to zero and sends the BRC back to the sender on the same route which passes all RBMs the cell has already passed when carrying the burst request. If one RBM can only provide a bandwidth less than λ, it lowers λ to the value of available capacity. Finally, the RBM reserves the cell rate λ for the connection and sends the BRC to the next RBM on the route of the connection. The last RBM on the route of the connection puts the BRC on the way back. While travelling back, the BRC contains the λ with which the connection may send. At least this cell rate is assigned in all RBMs. For the connection it may happen, that some RBMs have reserved a cell rate to a specific burst request, which is higher than the actual λ given back to the source, because λ could have been limited after passing this RBM. For that reason on the way back, all reservations of reservoir bandwidth assigned to the connection are set to the actual λ. Figure 4 contains the way of two BRCs belonging to connections c_2 and c_3 in the example network shown in figure 1.

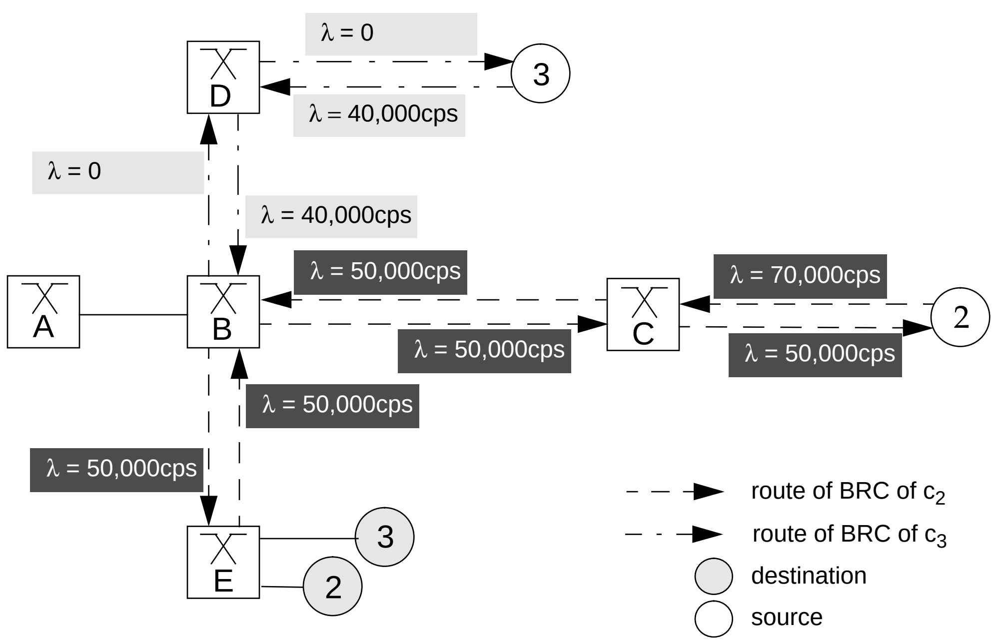

FIGURE 4. BRC Travel Paths

In Figure 4, c_2 requests a burst with a cell rate of 70,000 cells per second (cps). In node **C** this request cannot be satisfied due to limited bandwidth on the link **CB**. λ is reduced to 50,000cps by switch **C**. This can be provided by all other parts of the connection, too. So the BRC returns a λ of 50,000cps. After that, c_3 requests a burst with a cell rate of 40,000cps. There is no more reservoir on the VP between **B** and **E**, the reservoir is used by c_2 . For that reason, node **B** returns the BRC of c_3 with $\lambda = 0$.

After returning the BRC, the source knows whether it is allowed to send a burst and at which rate it can send. Now it can start sending the burst without any other conditions. It has to take care not to supersede its λ. The source is supposed to send the burst only for a limited time θ which is the maximum burst duration negotiated during call setup. If the returned λ is zero (no burst allowed at the moment), the source has to wait a time specified in the parameter τ_1 before generating the next BRC, for the same burst.

After the time period θ, the source stops sending burst cells. The burst is followed by a burst release cell (BEC). The BEC takes the same route as all other cells of the burst and notifies the various RBMs to release bandwidth assigned to the most recent burst. The source returns to its minimum sending rate λ_{min} and has to wait a time τ_2 before it is allowed to request another burst.

The choice of the parameter θ plays a crucial role in the ResABR scheme. The burst duration θ allows one source to use the reservoir or parts of it for a certain time. However, for the same time other sources are prevented from using the same resources. If the value of θ is small, sources do more often have a chance to request a burst. But sources also more often *will* request a burst due to the fact that they are unable to send all the data they wanted to send during θ. Each burst which is not completed because θ is over leads to a request for one (or more) new bursts. Additionally, each burst request generates extra traffic in the form of BRCs. Although the amount of this traffic is small compared to the payload traffic, BRCs have to be processed in each RBM. If the value of θ is large, a source can probably send the whole data connected to the burst, but other sources are prevented from sending their bursts for a long time. So it is difficult to provide fairness among the sources. Parameters τ_1 and τ_2 play an important role for fairness in that case, too. Figure 5 shows different burst lengths compared to different θ values.

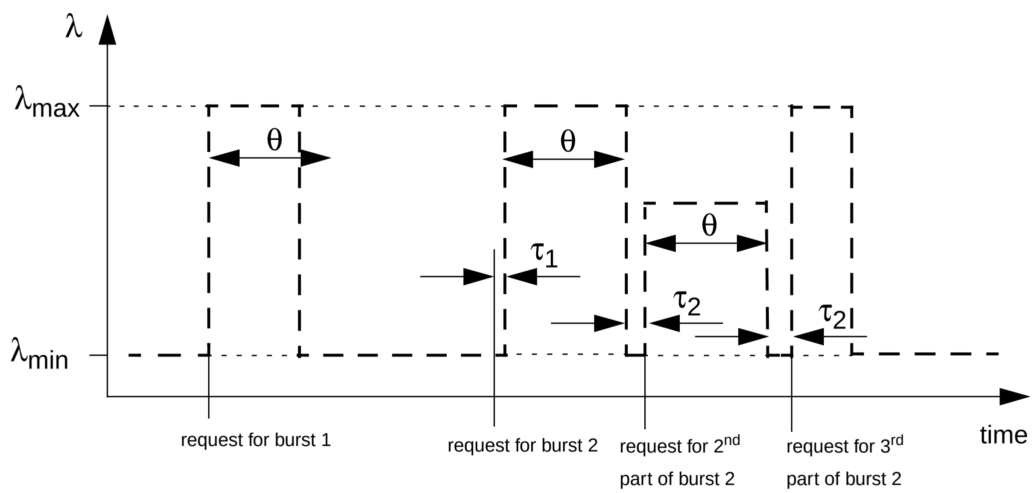

FIGURE 5. Bursts from a Data Source and their Transmission with ResABR

The first burst in Figure 5 has been accepted with the first request. It has been sent with λ_{max} and it lasted shorter than θ. The second burst has been accepted with only the second request. The amount of data, which the source wanted to send, was more than during one θ could be sent. So another request after $\theta + \tau_2$ was generated. The request for this second part has been accepted immediately, but λ

was limited to a value smaller than λ_{max}. A third burst request was necessary to complete the transmission of the data the source intended to send.

Another question is how to measure θ. There are two possibilities: A certain amount of time or a certain amount of data. In the first case, θ is independent from λ. In the second case, the burst duration in terms of time becomes longer, when a source is limited $\lambda < \lambda_{max}$. It is suggested to measure θ in terms of time and to use a value near the common length of bursts, which the applications usually generate. Nevertheless further investigation is needed to study the influence of θ on fairness and performance. Note that applications must be able to hold back their bursts for some burst duration periods in the worst case.

Both time-outs τ_1 and τ_2 play an important role for the fairness of the ResABR scheme. First of all τ_1 should be longer than one round trip of a BRC. Otherwise it will happen that a BRC finds r esources locked by forward traveling BRCs of other connections, although the locked resources are not really needed for some burst. τ_1 has to be adjusted in a way that there are not too many BRCs traveling in the network at the same time. A deadlock behavior could be the result. The difference between τ_1 and τ_2 gives some credit to sources, which did not send a burst in the last time. So, τ_2 must be larger than τ_1. Assumed that the burst duration equals the common natural burst length of the applications, it is suggested to fix τ_2 also equal to the common natural burst length of the applications, so at least one other source gets a chance to send its burst. But, extensive study is necessary to describe the influence of the time-out parameters on the systems behavior.

5 Conclusions and Future Work

In this paper we described a scheme, ResABR, based on the reservoir scheme proposed in [ULR95]. The ResABR scheme can be the basic scheme for connection admission and dynamic bandwidth allocation for a best effort service in ATM networks.

Compared to the rate-based ABR bearer service [BFG95],[ITU96], ResABR provides the following advantages:

- ABR bandwidth allocation is done on demand by the source, no rate-based bandwidth renegotiation is necessary, whenever the network status has changed. Resource management actions are necessary only, when a burst occurs;
- there is no extra storage of cells inside the network, acknowledged bursts will find the bandwidth they need. The scheme is independent fr om large buffer requirements concerning the used hardware;
- the scheme is robust and it provides less computational effort;
- it is fair among active.

ResABR uses a reservoir range to provide extra bandwidth to connections which need to transmit a burst of data, the reservoir itself can be dynamic, as well. So a high flexibility is given for the provider of ATM bearer services.

ResABR is especially suitable for Internet traffic (e. g. image traffic), because of the bulk character of such traffic. Bulk traffic is characterized by long periods of low cell rates interrupted by bursts. Connection-oriented file transfer might be performed using a bearer service based on ResABR, but the scheme cannot use all of its advantages in this case.

Further studies are necessary concerning the performance of the scheme regarding the following parameters:

- burst length and inter-arrival times of bursts. Both are dependent on the application. There must be found a common characterization of the applications for which ResABR is most efficient.
- the scheme immanent burst duration θ, time-out τ_1 for waiting between unsuccessful burst requests, and time-out τ_2 for waiting before sending the next burst request after a successful burst. There must be compared the meaning of the burst duration as a time interval and certain amount of data.

The optimal set of scheme immanent parameters depends on the applications and the characterization of their traffic. Further investigations will point to optimal parameter sets. Performance evaluation will focus to find the sensitivity of the performance of the scheme when the source characteristics are changing.

References

[BFG95] F. Bonomi and K. Fendick and N. Giroux. *The Available Bit Rate Service*. The ATM Forum newsletter, Vol. 3, No. 4, October 1995

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

[ITU96] ITU. *Recommendation I.371: Traffic Control and Congestion Control in B-ISDN*, May 1996

[ULR95] Roya Ulrich. *Reservoir-Based Resource Management for Slotted High Speed Networks*. Dissertation, Universitaet Erlangen-Nuernberg, Arbeitsberichte des IMMD, Vol. 28, No. 8, 1995