ABR Engineering: Roles and Guidelines for Setting ABR Parameters[1]

Sonia Fahmy[2], Raj Jain, Rohit Goyal and Bobby Vandalore
Department of Computer and Information Science
The Ohio State University
2015 Neil Avenue, Columbus, OH 43210-1277, USA
New Address: Raj Jain, Washington University in Saint Louis,
jain@cse.wustl.edu, http://www.cse.wustl.edu/~jain

ABSTRACT

This paper discusses the role of parameters negotiated and used for flow control in the ATM ABR service, and gives guidelines for setting each parameter. The effect of the speed of the links on the path from the source to the destination, and the round trip time of the connection, on the ABR parameter values is discussed. We also give simulation results to illustrate the effect of various parameter values on the performance of the ABR connection, in terms of both throughput and queue lengths.

Keywords: ATM networks, available bit rate, ABR service, ABR end system, ABR parameters

1 Introduction

ATM networks currently offer five service categories: constant bit rate (CBR), real-time variable bit rate (rt-VBR), non-real time variable bit rate (nrt-VBR), available bit rate (ABR), and unspecified bit rate (UBR). The ABR and UBR service categories are specifically designed for data traffic[3]. The ABR service provides better service for data traffic than UBR by frequently indicating to the sources the rate at which they should transmit. ABR can thus provide minimum rate guarantees and low cell loss to ABR sources.

The ABR source end system is allowed to send data at a given rate called Allowed Cell Rate (ACR), which ranges between a negotiated Peak Cell Rate (PCR) and Minimum Cell Rate (MCR). Immediately after establishing a connection, ACR is set to an Initial Cell Rate (ICR), which is also negotiated with the network. The source sends a Resource Management (RM) cell every Nrm−1 data cells (default Nrm value is 32), and the destination end system turns the RM cells around. As seen in figure 1, the RM cells traveling from the source to the destination are called forward RM cells (FRMs), while the RM cells traveling from the destination back to the source are called backward RM cells (BRMs). The RM cells collect the network feedback and return to the source,

---

[1] Submitted to "Journal of Computer Networks." This paper and all our papers and contributions are available through http://www.cis.ohio-state.edu/˜jain/

[2] Corresponding author

[3] The guaranteed frame rate (GFR) service is currently being finalized, and can also be used for data traffic.

which adjusts its allowed cell rate according to that feedback.

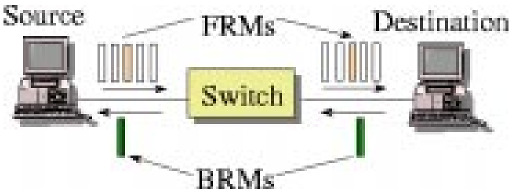


Figure 1: Forward and backward RM cells

Most resource management cells generated by the sources are counted as part of the source load in the sense that the total rate of data and RM cells should not exceed the ACR of the source. Such RM cells are called "in-rate" RM cells. Under exceptional circumstances, switches, destinations, or sources can generate extra RM cells. These "out-of-rate" RM cells are not counted in the ACR of the source and are distinguished by having their cell loss priority (CLP) bit set, meaning that the network will carry them only if there is enough bandwidth and discard them if congested.

There are three ways for the network switches to indicate their feedback to the sources. First, each cell header contains a bit called Explicit Forward Congestion Indication (EFCI), which can be set by a congested switch. Second, RM cells have two bits in their payload, called the Congestion Indication (CI) bit and the No Increase (NI) bit, that can be set by congested switches (switches using CI and NI are called relative rate marking (RRM) switches). The destination also sets the CI bit in the RM cells it turns around if the EFCI bit is marked in data cells. Third, the RM cells have a field called explicit rate (ER) that can be reduced by congested switches to any desired value. When sources receive the returning RM cells, they adjust their ACR according to certain rules. A complete explanation of the source and destination rules is presented in [9].

This paper examines how to set ABR parameters, and explains the factors which influence their values. The remainder of the paper is organized as follows. We begin by summarizing the ABR parameters, specifying the the units, range, default value and negotiation process for each. Then we examine each set of related parameters in consecutive sections, describing the role and effect of each, and giving some guidelines on how to set their values. We conclude by summarizing our findings.

2 ABR Parameters

At the time of connection setup, ABR sources negotiate several parameters with the network. A complete list of parameters used in the ABR mechanism is presented in table 1. The relevant parameters are further explained as they occur in the ensuing discussion.

3 Rate Upper and Lower Bounds: PCR and MCR

3.1 Role

The peak cell rate (PCR) and the minimum cell rate (MCR) are used in **source rule 1**. The rule states that source should always transmit at a rate equal to or below its computed ACR, which cannot exceed PCR and need not go below MCR, i.e.,

$$\text{MCR} \leq \text{ACR} \leq \text{PCR}$$

$$\text{Source Rate} \leq \text{ACR}$$

PCR is the maximum value at which a source can transmit. It must be negotiated down and has no default value. MCR is the minimum value that a source need not reduce its rate beyond. It is negotiated down to the minimum acceptable MCR, MCRmin, only if MCRmin is signaled.

3.2 Values

The sources can initially set PCR to the maximum possible value if they are willing to pay for it (for example, they can set it according to the capacity of the application or host, or the link bandwidth of the link from the host to the next node). Of course, billing and pricing considerations play an important role here, and sources may select a lower PCR value if they are not willing to pay for a large one.

MCR can be set according to the user requirements (e.g., video applications require some minimum rate guarantee), and the billing and pricing policy as well. Unless the traffic is high priority, MCR can be set to zero, to make the service a best effort one. Most applications, especially TCP/IP applications, however, work better with an MCR greater than zero, to prevent timeouts.

Table 1: List of ABR Parameters

Label	**Expansion**	**Units and Range**	**Default Value**	**Signaled?**
PCR	Peak Cell Rate	cells/second from 0 to 16M	–	down
MCR	Minimum Cell Rate	cells/second from 0 to 16M	0	down to MCRmin
ACR	Allowed Cell Rate	cells/second from 0 to 16M	–	no
ICR	Initial Cell Rate	cells/second from 0 to 16M	PCR	down
TCR	Tagged Cell Rate	constant	10 cells/s	no
Nrm	Number of cells between FRM cells	power of 2 from 2 to 256	32	optional
Mrm	Controls bandwidth allocation between FRM, BRM and data cells	constant	2	no
Trm	Upper Bound on Inter-FRM Time	milliseconds, 100 × power of 2 from -7 to 0	100 ms	optional
RIF	Rate Increase Factor	power of 2 from 1/32768 to 1	1	down
RDF	Rate Decrease Factor	power of 2 from 1/32768 to 1	1/32768	up, or down by $\leq$ RIF decrease factor
ADTF	ACR Decrease Time Factor	seconds, from 0.01 to 10.23 seconds in steps of 10 ms	0.5 s	optionally down
TBE	Transient Buffer Exposure	cells from 0 to 16,777,215	16,777,215	down
CRM	Missing RM-cell Count	integer of unspecified size	$\lceil \frac{TBE}{Nrm} \rceil$	computed
CDF	Cutoff Decrease Factor	zero or a power of 2 from 1/64 to 1	1/16	optionally up
FRTT	Fixed Round-Trip Time	microseconds from 0 to 16.7 seconds	–	accumulated

Observe that charging considerations may limit PCR to be a multiple of MCR, i.e.,

$$PCR = k \times MCR$$

where:

$$2 \leq k \leq 10$$

This makes it easier for traffic to be shaped.

The switches can reduce the PCR and MCR according the connection admission control (CAC) algorithm. One possible simple policy is to ensure that the following is satisfied:

$$\Sigma PCR_{CBR} + \Sigma SCR_{VBR} + \Sigma MCR_{ABR}(\text{including the new connection}) \leq \text{link bandwidth}$$

Hence, the MCR of the new connection can be computed as:

$$MCR_i \leq min(\text{User-requested:}MCR_i, \text{link bandwidth} - \Sigma PCR_{CBR} - \Sigma SCR_{VBR} - \Sigma_{j \neq i} MCR_{ABRj})$$

If the signaled MCR is less than the minimum acceptable MCR, i.e., $MCR_i < MCRmin_i$, the connection is rejected.

As for the PCR of the ABR connection, it is only limited by the bandwidth of the links on the path from the source to the destination.

$$PCR_i = min(PCR_i, \forall j, j \in \text{links on path from source to destination, minimum (link bandwidth)}_j)$$

Thus, PCR and MCR are dependent on the bottleneck link bandwidth, but not on the round trip time (RTT) of the connection.

Parameter	**Speed?**	**RTT?**	**Source initializes according to**	**Switch/Dest. modifies according to**
PCR	increases	no effect	link bandwidth or host/application capacity and pricing	bottleneck link bandwidth
MCR	increases	no effect	application requirements (e.g, video) and pricing	connection admission control (available resources)

4 Control of Frequency of RM Cells: Nrm, Mrm and Trm

4.1 Role

The three parameters Nrm, Mrm and Trm control the frequency of generation of resource management cells at the source. They are used in **source rule 3**. At any instant, sources have three kinds of cells to send: data cells, forward RM cells, and backward RM cells (corresponding to the reverse flow). The relative priority of these three kinds of cells is different at different transmission opportunities.

The sources are required to send an FRM after every Nrm cells. But if the source rate is low, the time between RM cells will be large and network feedback will be delayed. To overcome this problem, a source should send an FRM cell if more than Trm milliseconds have elapsed since the last FRM was sent. This introduces another problem for low rate sources. In some cases, at every transmission opportunity, the source may find that it has exceeded Trm and needs to send an FRM cell. In this case, no data cells will be transmitted. To overcome this problem, an additional condition was added that there must be at least two (Mrm) other cells between FRMs.

A waiting BRM has priority over waiting data, given that no BRM has been sent since the last FRM. Of course, if there are no data cells to send, waiting BRMs may be sent. The second and third part of source rule 3 ensure that BRMs are not unnecessarily delayed and that all available bandwidth is not used up by the RM cells. Figure 2 illustrates the scheduling of FRMs, BRMs and data cells.

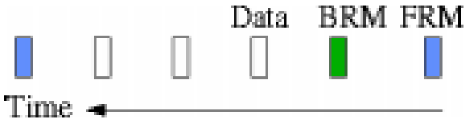


Figure 2: Scheduling of forward RM, backward RM, and data cells

4.2 Values

Mrm is constant at 2, and is not negotiated at connection setup. We next discuss the setting of Nrm and Trm.

4.2.1 Nrm

The specifications [5] select a default value of 32 for Nrm to ensure that the control overhead does not exceed approximately 6% (the value with window-based flow control). During normal operation, $\frac{1}{32}^{nd}$ or 3% of all cells are FRM cells. Another 3% of cells are BRM cells resulting in a total overhead of 6% [7]. Nrm is independent of link speed and round trip time, since it is simply a ratio.

In practice, the choice of Nrm affects the responsiveness of the control and the computational overhead at the end systems and switches. For a connection running at 155 Mbps, the inter-RM cell time is 86.4 μs while it is 8.60 ms for the same connection running at 1.55 Mbps. The inter-RM interval determines the responsiveness of the system. Sources, destinations, and switches may wish to increase Nrm if their processing power is limited, or if they wish to minimize the rate variations of the ABR connection, or increase the data cell frequency. They may wish to decrease Nrm if fast rate changes are desirable, and responsiveness to network feedback is advantageous.

At high data rates, a small RM cell interval can result in high frequency rate variations caused by the ABR feedback. *If traffic such as video is being transported over ABR, the rate variations must be minimized to reduce variations in the quality of service.* Users require a constant quality of service in a real-time application such as real-time video. One way of reducing the ABR rate changes is to send RM cells less frequently, i.e., set Nrm to a large value, instead of 32. Sending RM cells at end of each video frame is one possible option.

In the experiments shown next, we vary Nrm and examine the allowed cell rates at the sources, as well as the queue lengths at the switches, the link utilizations and the throughput at the destinations. Since the Nrm value must be a power of two that is allowed to range between 2 and 256 (according to the current specifications [5]), we have conducted experiments with all the allowed Nrm values (2, 4, 8, 16, 32, 64, 128 and 256). However, we only show the simulation results for Nrm = 8, 32 and 256 here. The reason why we have selected these values is that values smaller than 8 incur a very high control cell overhead and are not very realistic. 32 is the default value, and 256 is the maximum allowed value. In our simulations, all links are 155.52 Mbps links. The initial cell rate (ICR) of all sources is set to 150 Mbps, while the remaining ABR parameters are set to their default values as specified in the specifications. In particular, note that the value of the rate increase factor (RIF) parameter is set to 1/16. The ERICA [13] scheme is used in this study, with switch averaging interval set to a fixed time of 5 ms, and target utilization set to 90% of the link capacity. The configuration simulated consists of two ABR sources: source 1 sends data at its ACR throughout the simulation, while source 2 is a transient source that comes on at 100 ms and sends data for about 100 ms. All link lengths are 1000 km. The main aim of this configuration is to test how the responsiveness of the system is affected by the Nrm value, both during rate increases and rate decreases.

ABR performance for the two source transient configuration is shown in figures 3 through 5 for Nrm = 8, 32, and 256 respectively. The figures show the ACRs of the two sources, the queue length for switch 1, and the link utilization at the bottleneck link. In all cases, source 1 ACR quickly comes down to its target value of about 140 Mbps. When source 2 starts to send data, the ACRs of both sources are brought down to 70 Mbps. When source 2 stops sending data, the ACR for source 1 comes back up to 140 Mbps. There is a difference in the rate of increase of ACR for the three Nrm values. Since RIF is set to 1/16, the ACR comes up in steps on the receipt of every BRM cell. With Nrm = 8, the source receives BRMs more frequently than with Nrm = 256. As a result, the ACR for source 1 first reaches 140 Mbps faster for Nrm = 8. The overhead with small Nrm values is quite high, however. This can be clearly observed by measuring the throughput at the application layer at the destinations (these plots are not shown here). Another interesting observation is that for smaller Nrm values, source 1 does not start rising as fast as with larger Nrm values because the high RM cell overhead causes the data of the second source to take a longer time to be transmitted, and hence the two sources must share the bottleneck link for a longer time.

Table 2 shows the variation of inter-RM cell time with link speed, and with Nrm value. The source is assumed to be sending at link rate for the values shown in the table. A general heuristic is to use Nrm of 32 at speeds below OC-3 and to use Nrm of 256 for OC-3 and higher speeds.

Parameter	**Speed?**	**RTT?**	**Source initializes according to**	**Switch/Dest. modifies according to**
Nrm	maybe should increase with speed	no effect	processing speed and application type (real-time should increase it)	switch scheme and switch speed

Table 2: Inter-RM cell time for different speeds and Nrm

Total ABR Capacity	DS0 64 kbps	T1 1.5 Mbps	OC-3 155 Mbps	OC-24 1.2 Gbps
Nrm = 8	0.5 s	24 ms	24 μs	3 μs
Nrm = 32	2.3 s	96 ms	96 μs	12 μs
Nrm = 256	18.4 s	768 ms	768 μs	96 μs

4.2.2 Trm

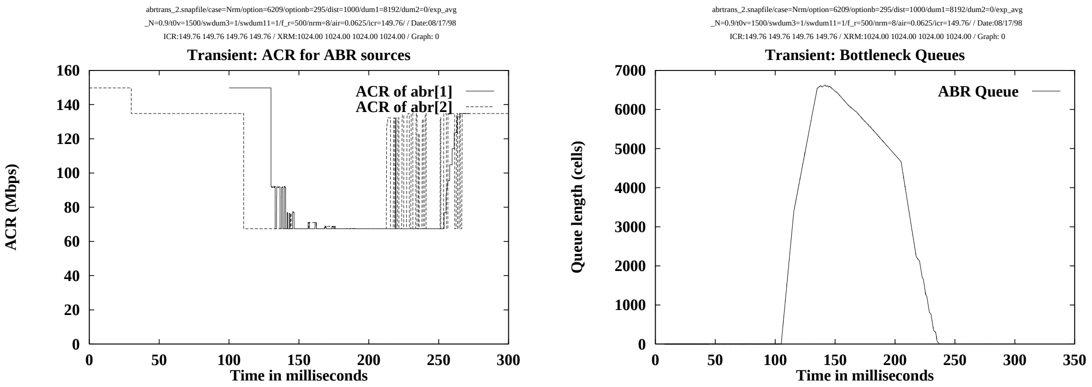
The Trm parameter is used with low rate sources, when Trm milliseconds are compared to the time elapsed since the last in-rate FRM cell was sent. Sources may get a low ACR due to high amplitude VBR traffic sharing the same resources as the ABR connection, a large number of ABR sources, or low bottleneck link speeds (T1 links). Trm allows low rate sources to sense the network state more frequently than normal. When more bandwidth suddenly becomes available, the network may not be able to allocate the source more bandwidth until it sees an RM cell from the source.

Smaller Trm values result in shorter time between RM cells, leading to faster transient response (rise from low rate to high rate). Small Trm values, however, cause high overhead with low rate sources. The choice of Trm depends on the link speed. For example, at a rate of 155 Mbps, the inter-cell time is 2.7 μs, while at a rate of 1.5 Mbps, the inter-cell time is 270 μs, and at a rate of 2.4 Gbps, the inter-cell time is 0.42 ns. Thus, a Trm value of 100 ms seems more appropriate for 1.5$-$155 Mbps than with higher (2.4 Gbps+) speeds, where a Trm of 100 ms is too long to wait before sending an FRM cell to sense the state of the network. Trm should be reduced in such cases. The switches or destination can compare Trm to the inter cell time calculated as the reciprocal of the negotiated PCR (which may indicate the bottleneck link bandwidth). A good value for Trm (based on heuristics) would be:

$$Trm = \frac{1}{PCR} \times c$$

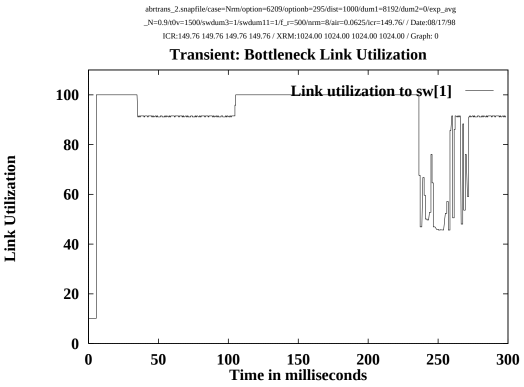
One choice of c can be $\frac{1{,}000{,}000}{27}$. This is based upon the intuition that 100 ms was observed to be suitable for OC-3 links (2.7 microsecond = 0.0027 millisecond inter-cell time).

Trm is independent of the round trip time, i.e., whether the connection is local to a LAN, crosses a WAN, or traverses a satellite link of hundreds of milliseconds delay. This is because Trm is compared to the time since the last in-rate FRM cell was sent, so it is independent of the time the



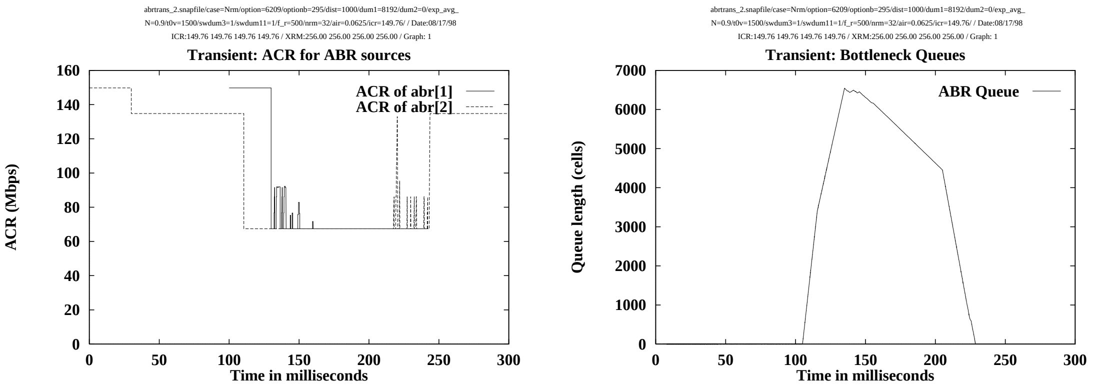
(a) Allowed Cell Rate

(b) Queue Length



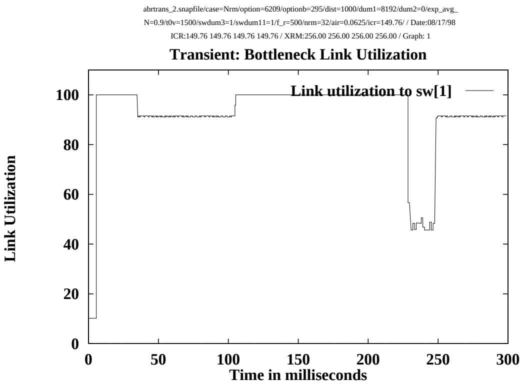
(c) Link Utilization

Figure 3: Results for a WAN transient configuration, Nrm = 8



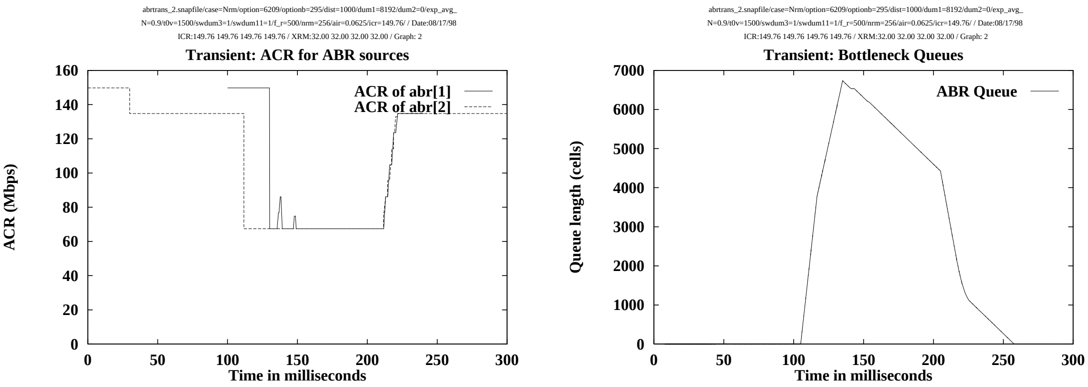
(a) Allowed Cell Rate

(b) Queue Length



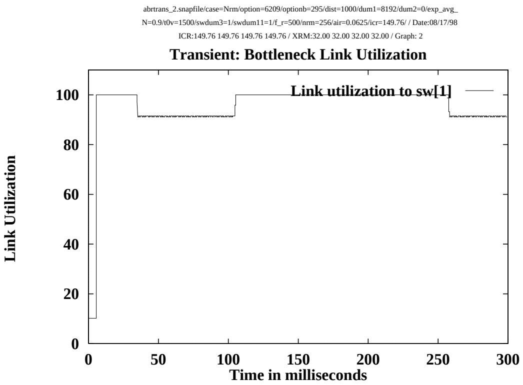
(c) Link Utilization

Figure 4: Results for a WAN transient configuration, Nrm = 32



(a) Allowed Cell Rate

(b) Queue Length



(c) Link Utilization

Figure 5: Results for a WAN transient configuration, Nrm = 256

RM cell reached the destination, or the time the RM cells returns back to the source.

We have performed an experiment with various Trm values (1, 10 and 100 ms). We assume VBR background traffic always has a higher priority. We also assume a simple on/off VBR model where VBR is on for 20 ms and off for 20 ms. When VBR is on, it sends at a rate of 138 Mbps. Simulation results (see figure 6) show that in this case, capacity may be unused for a long time for large Trm values (100 ms), when VBR goes away and capacity for ABR becomes available. Lower Trm (1 ms and 10 ms) results in more frequent RM cells, and hence a faster response [10]. *This is especially important for small or zero minimum cell rate.*

Parameter	**Speed?**	**RTT?**	**Source initializes according to**	**Switch/Dest. modifies according to**
Trm	decreases	no effect	processing speed and application type	switches can reduce Trm for a high PCR, or increase it for low switch speed

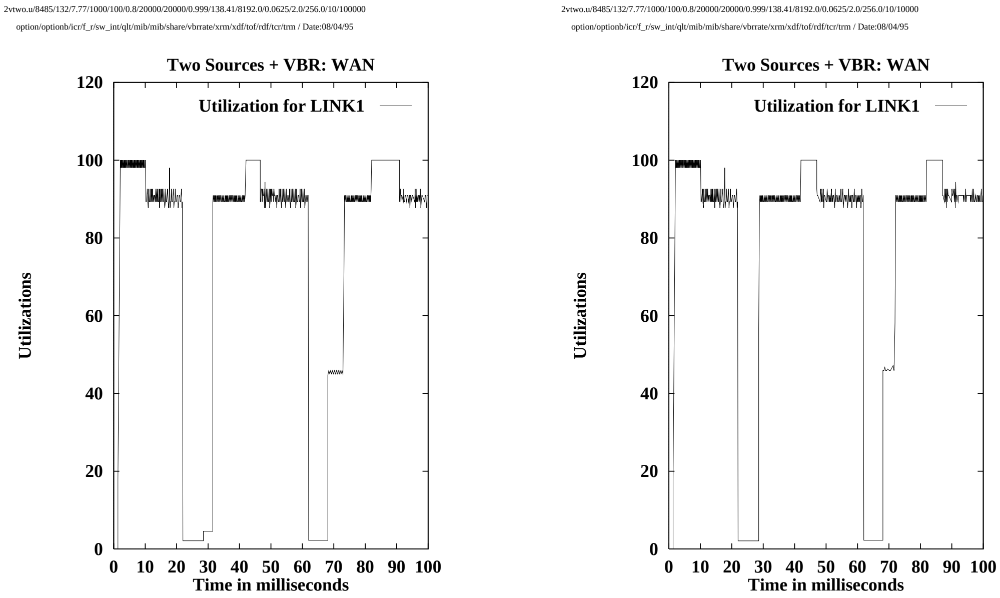
5 Rate Increase and Decrease Factors: RIF and RDF

5.1 Role

The rate increase factor (RIF) and rate decrease factor (RDF) are used in **source rules 8 and 9**. Source rules 8 and 9 describe how the source reacts to network feedback. The feedback consists of the explicit rate (ER), congestion indication bit (CI), and no increase bit (NI). A source does not simply change its ACR to the new ER due to the following reasons:

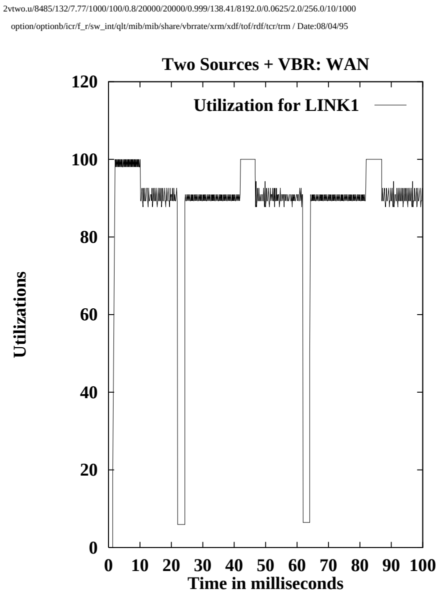
1. If the new ER is very high compared to current ACR, switching to the new ER may cause sudden overload in the network. Therefore, the amount of increase is limited. The rate increase factor (RIF) parameter determines the maximum allowed increase in any one step. The source cannot increase its ACR by more than RIF × PCR.

2. If there are any EFCI or relative rate marking (RRM) switches in the path, they do not change the ER field. Instead, they set EFCI bits in the cell headers, or CI and NI bits in RM cells. The destination monitors EFCI bits in data cells, and returns the last seen EFCI



(a) Trm = 100 ms

(b) Trm = 10 ms



(c) Trm = 1 ms

Figure 6: Link utilization results for two sources and VBR in a WAN

bit in the CI field of a BRM. A CI of 1 means that the network is congested and that the source should reduce its rate. The decrease is determined by the rate decrease factor (RDF) parameter. Unlike the increase, which is additive, the decrease is multiplicative in the sense that:

$$\text{ACR} \leftarrow \text{ACR} \times (1 - \text{RDF})$$

3. The no-increase (NI) bit handles mild congestion by allowing a switch to specify an ER, but instruct the source not to increase its rate if ACR is already below the specified ER.

The actions corresponding to the various values of CI and NI bits are as follows:

NI	**CI**	**Action**
0	0	$\text{ACR} \leftarrow \min(\text{ER}, \text{ACR} + \text{RIF} \times \text{PCR}, \text{PCR})$
0	1	$\text{ACR} \leftarrow \min(\text{ER}, \text{ACR} - \text{ACR} \times \text{RDF})$
1	0	$\text{ACR} \leftarrow \min(\text{ER}, \text{ACR})$
1	1	$\text{ACR} \leftarrow \min(\text{ER}, \text{ACR} - \text{ACR} \times \text{RDF})$

Once the ACR is updated, the subsequent cells sent from the source conform to the new ACR value. However, if the earlier ACR was very low, it is possible that the very next cell is scheduled a long time in the future. In such a situation, it is advantageous to "reschedule" the next cell, so that the source can take advantage of the high ACR allocation immediately [10].

5.2 Values

RIF and RDF play an important role when the connection passes through EFCI or RRM switches. In addition, some ER schemes work better with conservative RIF values, while others, such as ERICA [13] are insensitive to the RIF value, and work well with an RIF of 1.

5.2.1 RIF

The rate increase factor determines the maximum increase when a BRM cell indicating underload is received. If the RIF is set to a fraction less than one, the maximum increase at each step is limited to RIF × the peak cell rate for the VC. Setting RIF to small values is a more conservative strategy that controls queue growth and oscillations, especially during transient periods. It, however, may slow down the response of the system when capacity suddenly becomes available, leading to underutilization.

If there are no EFCI switches in a network, setting RIF to 1 allows ACRs to increase as fast as the network directs it (through the ER field). This allows the available bandwidth to be used quickly. For EFCI networks, or a combination of ER and EFCI networks, RIF should be set conservatively to avoid unnecessary oscillations [14]. Thus, sources can initially set RIF to large values, even 1, according to the application requirements. During connection setup, any switch which does not implement an explicit rate scheme or implements a scheme which requires a conservative RIF (such as EPRCA) must reduce RIF to a conservative value such as 1/16 or less. RIF can be set to more conservative (smaller) values for high speeds (as indicated by PCR) and long round trip times (long delay links) to avoid congestion loss. The fixed part of the round-trip time (FRTT) is accumulated during connection setup. This is the minimum delay along the path and does not include any queueing delay [11].

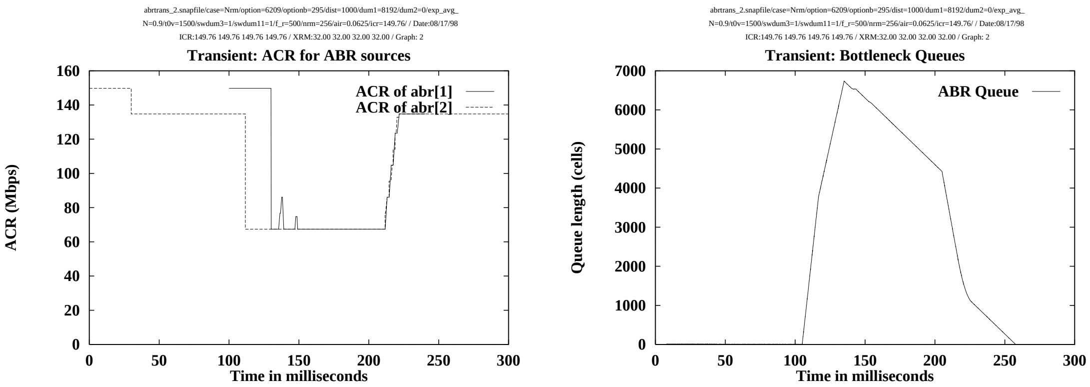
Figures 7 and 8 compare the performance of a transient configuration (same as the configuration used in the Nrm experiments) with RIF set to 1/16 (the default value) and RIF set to 1. The basic ERICA scheme [13] is used in these simulations. Nrm is set to 256 to slow down the feedback rate, in order to emphasize the effect of RIF. All other parameters are the same as with the Nrm experiments. It is clear from figure 7 that an RIF value of 1/16 results in a step increase of the rate of the non-transient source when the transient source stops transmission. With RIF set to 1 (figure 8), the rate of the non-transient source increases to the full rate as soon as the inactivity of the transient source is detected.

Parameter	**Speed?**	**RTT?**	**Source initializes according to**	**Switch/Dest. modifies according to**
RIF	no, but may be decreased	no, but may be decreased	application requirements	EFCI and RRM switches, and ER switches sensitive to RIF should reduce it depending on FRTT, PCR and scheme

5.2.2 RDF

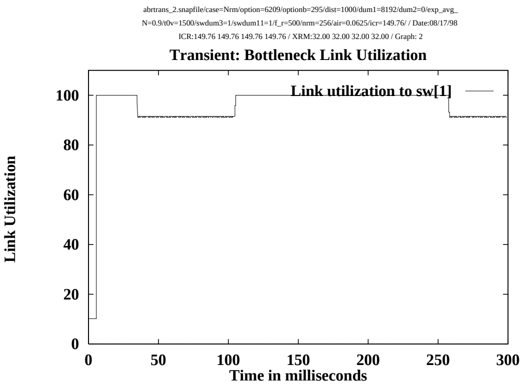
When the network is congested (the CI bit is set), the source reduces its rate as follows:

$$\text{ACR} \leftarrow \text{ACR} \times (1 - \text{RDF})$$



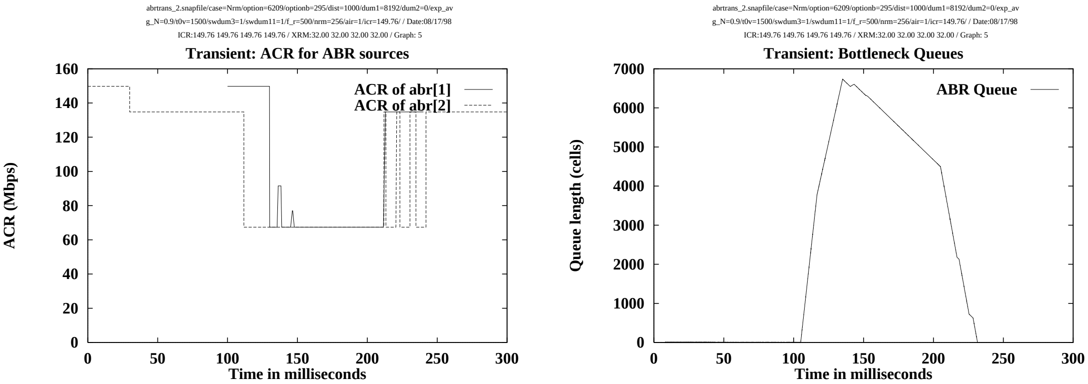
(a) Allowed Cell Rate

(b) Queue Length



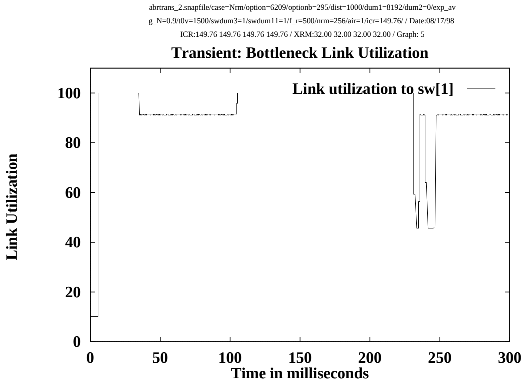
(c) Link Utilization

Figure 7: Results for a WAN transient configuration (Nrm = 256) RIF = 1/16



(a) Allowed Cell Rate

(b) Queue Length



(c) Link Utilization

Figure 8: Results for a WAN transient configuration (Nrm = 256) RIF = 1

Thus the RDF parameter determines how fast the rate is reduced in case of congestion. This multiplicative decrease only occurs if the CI bit is set, either by the switches, or by the destination when the EFCI bits of data cells are set.

The source should initially set RDF to a moderate value. Switches should reduce RDF dependent on the schemes they use for setting EFCI bits or CI bits. *Explicit rate switches need not modify RDF*. The RDF parameter should be set more conservatively (to smaller values) for higher speeds and longer round trip times to avoid a large amount of cell loss during congestion. Switches can examine the round trip time (in the FRTT field) and bottleneck link speed (as indicated by the PCR), and reduce RDF accordingly. If the switch or destination detects a large FRTT or large PCR (indicating a high bottleneck link speed), then RDF should be reduced.

Parameter	**Speed?**	**RTT?**	**Source initializes according to**	**Switch/Dest. modifies according to**
RDF	no, but may be decreased	no, but may be decreased	application requirements	EFCI and RRM switches should reduce dependent on FRTT, PCR and scheme

6 Rate Reduction under Abnormal Conditions and Startup after Idle Periods: TBE, CRM, CDF, ICR and ADTF

Since CRM and CDF are used with source rule 6, they are both discussed together. Both CRM and ICR are computed using the TBE parameter, so TBE and ICR are also discussed here, as well as ADTF that is used in conjunction with ICR.

6.1 Role

We first discuss the rule 6 parameters, and then we discuss the rule 2 and 5 ones.

6.1.1 TBE, CRM and CDF

The three parameters transient buffer exposure (TBE), missing RM cell count (CRM), and cutoff decrease factor (CDF) are used in **source rule 6**. This rule deals with the following scenario: if a network link fails, or becomes highly congested, RM cells are blocked and the source does not receive feedback. To protect the network from continuous in-flow of traffic under such circumstances, the sources are required to reduce their rate if the network feedback is not received in a timely manner.

In steady state, a source should receive one BRM for every FRM sent. The sources keep a count of the RM cells sent, and if no backward RM cells are received for a long time, the sources reduce their rate by a factor of "Cutoff Decrease Factor (CDF)." The "long time" is defined as the time to send CRM forward RM cells at the current rate. When rule 6 triggers once, the condition is satisfied for all successive FRM cells until a BRM is received. Thus, this rule results in a fast exponential decrease of ACR.

CRM is computed from another parameter called transient buffer exposure (TBE) which is negotiated at connection setup. TBE determines the maximum number of cells that may suddenly appear at the switch during the first round trip before the closed-loop phase of the control takes effect. During this time, the source will have sent TBE/Nrm RM cells. Hence,

$$\mathrm{CRM} = \lceil \frac{\mathrm{TBE}}{\mathrm{Nrm}} \rceil$$

6.1.2 ICR and ADTF

At the beginning of a connection, sources start at the initial cell rate (ICR) as specified in **source rule 2**. During the first round trip, a source may send as many as ICR × FRTT cells into the network. Since this number is negotiated separately as TBE, the following relationship exists between ICR and TBE:

$$ICR \times FRTT \leq TBE$$

or:

$$ICR \leq \frac{TBE}{FRTT}$$

The sources are required to use the ICR value computed above if it is less than the ICR negotiated with the network:

$$\text{ICR used by the source} = \min(\text{ICR negotiated with the network}, \frac{TBE}{FRTT})$$

According to **source rule 5**, the rate allowed to a source is valid only for approximately ADTF seconds. If a source does not transmit any RM cells for this duration, it cannot use its previously allocated ACR, particularly if the ACR is high. The source should re-sense the network state by sending an RM cell and decreasing its rate to the initial cell rate (ICR) negotiated at connection setup. If the source ACR is already below ICR, it should not increase to ICR. The timeout interval is set to the ACR Decrease Time Factor (ADTF) parameter, whose default value is 500 ms.

Rule 5 is intended to solve the problem of *ACR retention*, when a source retains a rate allocated to it under light loads, and uses it when the network is highly loaded, causing congestion. Several solutions to this problem (called *use it or lose it* (UILI) solutions) were proposed [19, 12]. The ATM Forum standardized a policy that reduces ACR to ICR when the timeout ADTF expires. Vendors are free to implement additional proprietary restraints at the source or at the switch.

6.2 Values

We first discuss the value of TBE and the two parameters which depend on it (CRM and ICR). Then, we discuss CDF, and finally ADTF.

6.2.1 TBE, CRM, and ICR

As previously mentioned, TBE determines the exposure of the switch to sudden traffic transients. It determines the the number of cells that may be received at the switch during initial start up (or after any long idle period of time). TBE is specified in cells while CRM is specified in RM cells. Since there is one RM cell per Nrm cells, the relationship between CRM and TBE is as follows:

$$\text{CRM} \leftarrow \lceil \text{TBE}/\text{Nrm} \rceil$$

In negotiating TBE, the switches have to consider their buffer availability. As the name indicates, the switch may be suddenly exposed to TBE cells during the first round trip (and also after long idle periods). For small buffers, TBE should be small and vice versa. On the other hand, TBE should also be large enough to prevent unnecessary triggering of rule 6 on long delay paths or with very high speeds. Thus TBE is highly affected by speed and round trip time (the delay bandwidth product of the connection).

TBE can thus be set to:

$$PCR \times FRTT + \Sigma_i, i \in \{\text{switches on path}\}, \text{buffer sizes}$$

to account for the speed, link delays, and buffer sizes.

Effect of speed and round trip time on TBE and CRM:

For long-delay links, such as satellite links, our simulation results revealed that source rule 6 can unnecessarily trigger and cause oscillations during start up and after idle periods, unless TBE is large enough. This can degrade the throughput considerably. Figure 9 shows the configuration used to illustrate the problem. All the links are OC-3 links operating at a rate of 155.52 Mbps. The link connecting the two switches is a satellite link, while the links connecting the switches to the end systems are each 1 km long. The one-way propagation delay of the satellite link is 275 ms, while the propagation delay of each LAN link is 5 microseconds. The traffic is bidirectional, and the sources are persistent. The ERICA [13] algorithm is used with target utilization 90%. The ABR source parameter values are as follows: PCR = 155.52 Mbps, MCR = 0 Mbps, ICR = 0.9×PCR = 140 Mbps, Nrm = 32, RIF = 1, CDF = 1/16, and CRM = 32, 256, 1024, 4096, 6144, 8192.

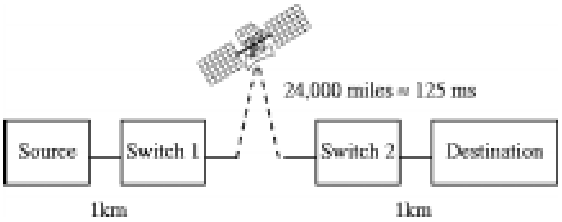
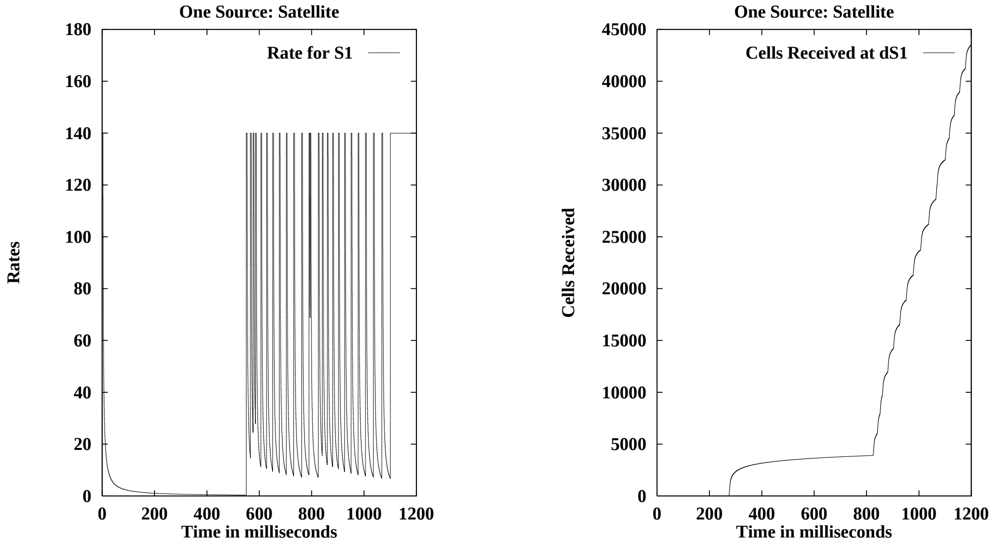


Figure 9: One source configuration

Figure 10 illustrates the performance of the system with CRM set to 32 (the default value before August 1995). Figure 10(a) shows the allowed cell rate of the source over 1200 ms, and figure 10(b) shows the number of cells received at the destination during the same period of time [3]. As seen in figure 10(a), the initial rate is 140 Mbps (90% of 155 Mbps). After sending 32 RM cells (or CRM×Nrm = 32×32 = 1024 cells), rule 6 triggers and the rate rapidly drops. The first feedback is received from the network after around 550 ms (275 ms×2), because the one-way delay of the satellite link is 275 ms. The network asks the source to go up to 140 Mbps. The source increases its rate but rule 6 triggers again. The rule triggers again because the time between returning RM cells is large (they were sent at a low rate). This phenomenon of increase and decrease repeats resulting in high-frequency oscillations between very low rates and very high rates. The rapid rate drops occur due to the triggering of source rule 6, while the rate increases occur because the network feedback is consistently at 140 Mbps (90% of 155 Mbps).



(a) Allowed Cell Rate (ACR)

(b) Cells Received at Destination

Figure 10: Simulation results for a one source configuration. CRM = 32.

Figure 10(b) shows the number of cells received at the destination. From this figure, it is possible to compute instantaneous throughput by computing the slope of the curve. It is also possible to compute average throughput over any interval by dividing the cells received (increase in the y-value) during that interval by the period of time (x-value) of the interval. The average throughput during the interval from 275 ms to 825 ms is 32 Mbps and that during the interval from 825 ms to 1200 ms is 45 Mbps. During the first 550 ms, the source is mostly sending at a very low rate until the first feedback is received after about 550 ms. The effect of the receipt of feedback can be observed at the destination after 550+275=825 ms. After the first feedback is received, the rate oscillations result in reduced throughput. The results do not significantly vary for different values of CDF. The low throughput values in figure 10(b) are a result of the unnecessary triggering of source rule 6 for small CRM values. Rule 6 limits the number of cells that can be in flight during start up periods.

For full throughput, we need to set the value of TBE such that the number of cells in flight can be as large as those required to fill the path both ways. This number is equal to the round trip time (FRTT) multiplied PCR. Hence, the number of RM cells in flight (CRM) should be (1/Nrm)th of this value:

$$\text{CRM} \geq \frac{\text{FRTT} \times \text{PCR}}{\text{Nrm}}$$

For 155 Mbps links, CRM should be greater than or equal to 6144 (550 ms×365 cells per ms/32 cells). For 622 Mbps links, CRM should be greater than or equal to 24576 (6144×4). For two 622 Mbps satellite hops, CRM should be greater than or equal to 49152 (24576×2). For n 622 Mbps satellite hops, CRM should be greater than or equal to 24576×n. Since the size of the TBE parameter is 24 bits and Nrm is normally 32, a 24-bit TBE allows a 19-bit CRM, which is sufficient for most situations (long delay links and high speeds).

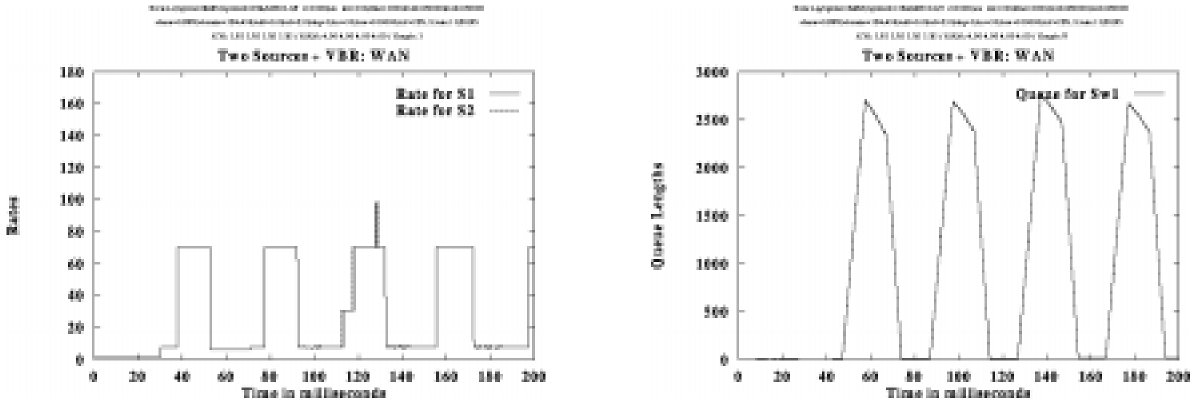
Effect of TBE on queue sizes:

It has been incorrectly believed that cell loss could be avoided by simply negotiating a TBE value below the number of available buffers in the switches. We show in [8] that it is possible to construct workloads where queue sizes could be unreasonably high even when TBE is very small. TBE limits the queue length only during initial startup and after idle periods when there are no previous cells in the network from the same VC. In this case, the queue length can be given by the following equation:

$$\text{Queue length} = (\text{number of sources} - 1) \times min(TBE, \text{burst size})$$

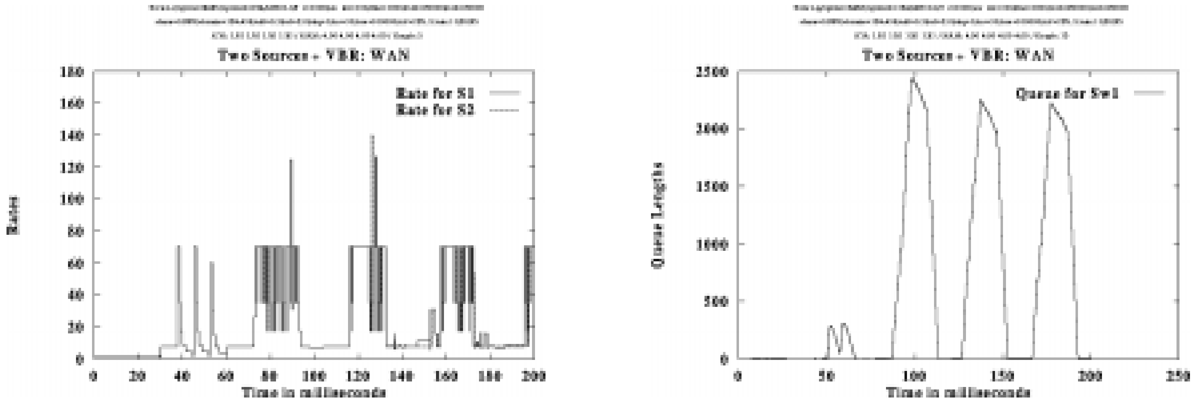
TBE cannot be relied upon during the closed-loop operation phase of a connection. During this latter phase, the contribution of a VC to the queue at a switch can be more than its TBE. The buffer usage at a switch can be more than the sum of TBEs allocated to active VCs. In steady state, rule 6 rarely triggers and is overridden by subsequent explicit feedbacks. Since the reverse flow is not stopped completely, the forward flow continues and keeps filling the queues. TBE does not significantly affect the maximum queue length.

Figures 11 and 12 show ACR and queue lengths for a network consisting of two ABR and one VBR sources going through two switches to corresponding destination. All simulation results use ERICA switch algorithm [13]. All links are 155 Mbps and 1000 km long. All VCs are bidirectional, that is, D1, D2, VD1 are also sending traffic to S1, S2 and VS1. The following parameter values are used: PCR = 155.52 Mbps, MCR = 0 Mbps, ICR = min155.52, TBE/FRTT, RIF = 1, Nrm = 32, RDF = 1/512, CRM = TBE/Nrm, Trm = 100 ms, FRTT = 30 ms, TBE = {128, 512, 1024} (three values), CDF = {0, 0.5} = {Without rule 6, With Rule 6}. The VBR source generates a square waveform of 20 ms on and 20 ms off. During on period, its amplitude is 80% of the link rate. During off period, the amplitude is zero. The first VBR pulse starts at t=2 ms. Thus, it is on



(a) Transmitted Cell Rate: Without Rule 6

(b) Queue Length: Without Rule 6



(c) Transmitted Cell Rate: With Rule 6

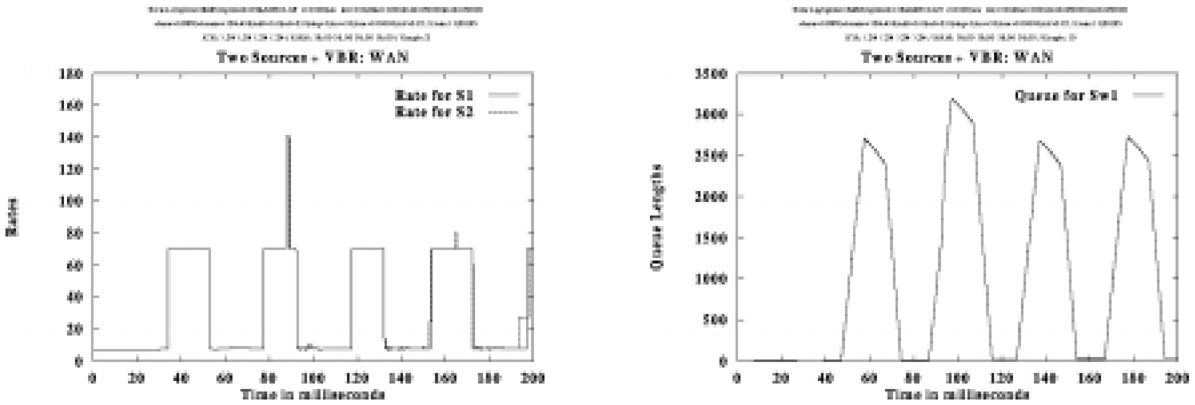
(d) Queue Length: With Rule 6

Figure 11: Two Sources and VBR on a WAN, TBE = 128 cells

from 2 to 22 ms and off from 22 to 42 ms and so on. The target utilization is 90%. The scheduler gives preference to VBR and so there are no VBR queues.

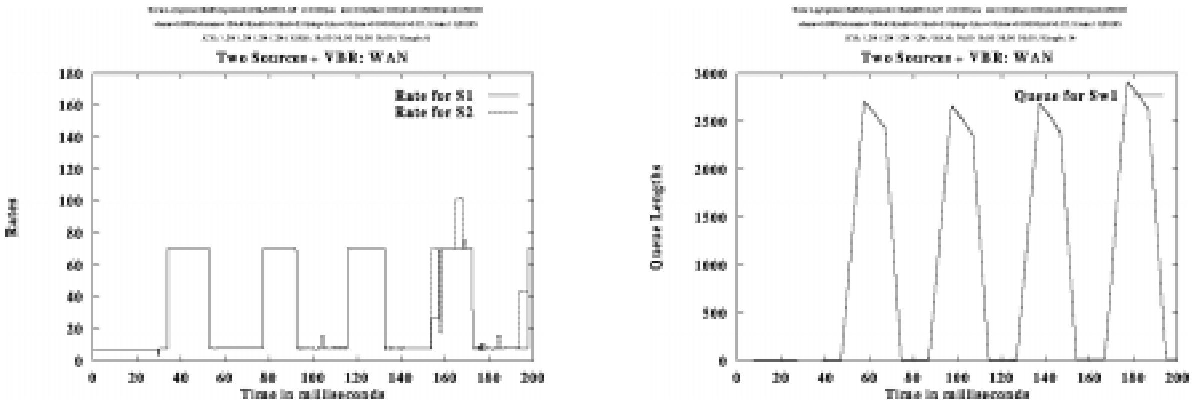
Figure 11 shows the ABR rates and queue graphs for TBE of 128 cells. With just two sources the queue length (without rule 6) is of the order of 2500 cells. The situation does not change significantly with rule 6. Rule six does trigger during initial start up, but is not triggered at all once the flow is set up. Figure 12 shows ABR rates and queue graphs for TBE of 512 cells. Once again with or without rule 6 the queue length is 2500 to 3000 cells. This queue length is more than that with TBE of 128 but there is no simple relationship between TBE and queue length.

The reason for the inadequacy of rule 6 in limiting the queue growth can be explained as follows (figure 13). Assume that a certain source S is sending forward RM cells at an average rate of R



(a) Transmitted Cell Rate: Without Rule 6

(b) Queue Length: Without Rule 6



(c) Transmitted Cell Rate: With Rule 6

(d) Queue Length: With Rule 6

Figure 12: Two Sources and VBR on a WAN, TBE = 512 cells

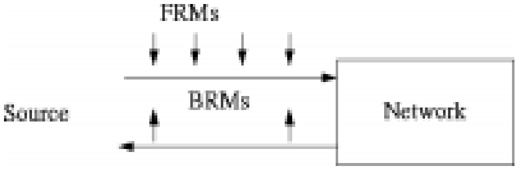


Figure 13: Source rule 6 does not trigger if BRM flow is maintained

cells per second (cps). The RM cells are turned around by the destination and the backward RM cells are received by S at a different rate r cps. In this case, the inter-forward-RM cell time at the source is $1/R$ while the inter-backward-RM cell time at the source is $1/r$. Source end system Rule 6 will trigger at S if the inter-backward-RM time is much larger (more than CRM times larger) than the inter-forward-RM time [8]. That is, if:

$$1/r \geq \text{CRM} \times (1/R)$$

or:

$$R \geq \text{CRM} \times r$$

In the case of initial startup, r is zero and so after TBE cells, rule 6 triggers and protects the sources. Similarly, in the case of a bursty source, r is zero and rule 6 triggers after TBE cells. However, if the BRM flow is not totally stopped and $R < CRM \times r$, then the cells can accumulate in the network at the rate of $(R - r) \times Nrm$ and not trigger rule 6. In such cases, the queues can grow substantially. The maximum queue length is a function of PCR, the target utilization, and the VBR amplitude, multiplied by the feedback delay [8].

Parameter	**Speed?**	**RTT?**	**Source initializes according to**	**Switch/Dest. modifies according to**
TBE	increases	increases	application type, pricing and host capacity	buffering and resources, and PCR and FRTT

ICR:

ICR should be set by the source as desired according to pricing and the application type. For TCP/IP applications and lower link speeds, ICR should be close to the peak cell rate (PCR).

Switches should reduce their ICR to reflect their availability of buffers, as well as the bandwidth available for the connection. ICR is related to the availability of resources as computed during connection setup, and should correspond to the anticipated ACR for the connection at that time. Finally, the source takes the minimum of that ICR and $\frac{TBE}{FRTT}$ to correspond to the rate at which the source should initially send for the first round trip or after idle periods, before feedback is received. ICR depends on the bottleneck link speed and the round trip time.

Parameter	**Speed?**	**RTT?**	**Source initializes according to**	**Switch/Dest. modifies according to**
ICR	increases	source takes minimum of signaled ICR and $\frac{TBE}{FRTT}$	pricing, host capacity and application	buffering and resources, PCR and FRTT

6.2.2 CDF

When source rule 6 is triggered, the source reduces its rate by a factor of CDF, but not below the minimum cell rate. That is,

$$\text{ACR} \leftarrow max(\text{MCR}, \text{ACR} - \text{ACR} \times \text{CDF})$$

where the value of CDF can be zero (for no rate decrease), or it can be a power of two that ranges from 1/64 to 1.

This means that after CRM RM cells are sent (or CRM×Nrm total cells are sent), and no backward RM cell is received:

$$\text{ACR} = \text{ACR}_{initial} \times (1 - \text{CDF})$$

Note that if rule 6 is triggered once, it usually triggers on sending successive forward RM cells (as long as no backward RM cells are being received).

Thus, after CRM+1 RM cells (or (CRM+1)×Nrm cells) are sent:

$$\text{ACR} = \text{ACR}_{initial} \times (1 - \text{CDF})^2$$

After CRM+k RM cells (or (CRM+k)×Nrm cells) are sent:

$$\text{ACR} = \text{ACR}_{initial} \times (1 - \text{CDF})^{k+1}$$

Such repeated rate reductions result in an exponential rate drop when source rule 6 triggers, as long as no feedback is being received as shown in figure 14.

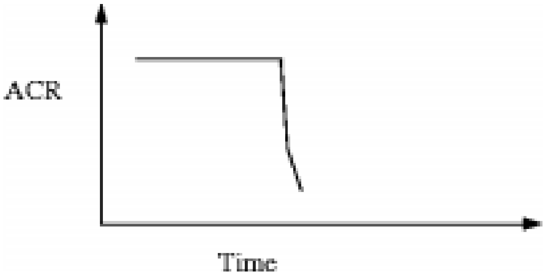


Figure 14: Rule 6 results in a sudden drop of rate

The smaller CDF value, the more rapid the rate decrease when rule 6 is triggered. It is possible to disable source rule 6, by setting CDF to zero. This may be desirable if TBE cannot be set to a reasonable value, or if TBE must be set to a low value to decrease ICR, but rule 6 should not be triggered unnecessarily. However, disabling rule 6 in this manner risks a large amount of cell loss in case of link failures or congestion collapse. CDF may be set to smaller values for high speeds and long RTTs to avoid big losses. It is set according to the application type, confidence in TBE value, confidence in links, and availability of resources.

Parameter	**Speed?**	**RTT?**	**Source initializes according to**	**Switch/Dest. modifies according to**
CDF	may be smaller for high speeds	may be smaller for long RTTs	application type, confidence in TBE value	confidence in TBE value, confidence in links, availability of resources

6.2.3 ADTF

As previously mentioned, the purpose of the ADTF timeout is to avoid the ACR retention problem that may cause congestion. ACR retention can cause sudden queue growth of:

$$(ACR - \text{source rate}) \times \text{feedback delay} \times (\text{number of sources} - 1)$$

VCs that disable rule 5 (e.g., by setting ICR=PCR) can be vulnerable to sudden arrivals. The default value of 500 ms was selected to correspond to the timer granularity used with most TCP/IP implementations using slow start.

ADTF especially affects bursty traffic. ADTF is independent of the bottleneck link speed of the connection since traffic is smoothed in the ATM network. ADTF should be larger than the round trip time:

$$ADTF > RTT$$

to prevent unnecessary rate reductions for long round trip times. Sources can set ADTF according to the application traffic characteristics (the expected burstiness of the traffic). Switches can reduce ADTF if they have little available resources.

Parameter	**Speed?**	**RTT?**	**Source initializes according to**	**Switch/Dest. modifies according to**
ADTF	no effect	no effect (may increase)	application traffic characteristics	if little available resources, reduce ADTF

7 Upper Bound on Out of Rate RM Cells: TCR

Although the tagged cell rate (TCR) is not signaled, we include a brief discussion here on its role, and the significance of the value chosen for it.

7.1 Role

As stated in **source rule 11**, the out-of-rate FRM cells generated by sources are limited to to a rate below the tagged cell rate (TCR) parameter, which has a default value of 10 cells per second.

7.2 Values

Although higher TCR values improve transient response with zero or very low ACRs, since feedback is more frequent, increased TCR does increase the RM cell overhead in such cases. Rescheduling becomes important in cases where ACR is very low and the new ACR will allow cells to be scheduled earlier than their previously scheduled time [10]. There are currently no guidelines on how to space out-of-rate RM cells.

It seems like TCR should depend on the bottleneck link speed, and perhaps a ratio, such as Nrm, should be used. 10 cells per second may be too low for very high speeds, e.g., 2.4 Gbps+. Perhaps it would be better to state, for example, that no more than$x$%, say 2.7×10^{-5}% of the link bandwidth should be used for out-of-rate RM cells. The value 2.7×10^{-5}% is based on the intuition that 10 cells per second seems to be a good value for OC-3 links (10 cells per second out of 365 cells per millisecond).

8 Summary

Table 3 summarizes the discussion in this paper (the conclusions from individual sections are just repeated for faster reference). For each of the parameters, the table indicates what the value the source end system sets for the parameter, how switches and destinations negotiate the parameter, how the parameter is affected by link speeds, and how it is affected by the round trip time of the connection.

References

[1] Flavio Bonomi and Kerry W. Fendick. Rate-based flow control framework for the available bit rate ATM service. *IEEE Network Magazine*, 9(2):25–39, March/April 1995.

[2] Thomas M. Chen, Steve S. Liu, and Vijay K. Samalam. The available bit rate service for data in ATM networks. *IEEE Communications Magazine*, 34(5):12, May 1996.

[3] Sonia Fahmy, Raj Jain, Shivkumar Kalyanaraman, Rohit Goyal, and Fang Lu. On source rules for ABR service on ATM networks with satellite links. In *Proceedings of the First International Workshop on Satellite-based Information Systems*, pages 108–115, Rye, New York, November 1996.

[4] Kerry W. Fendick. Evolution of controls for the available bit rate service. *IEEE Communications Magazine*, 34(11):35–39, November 1996.

[5] The ATM Forum. The ATM forum traffic management specification version 4.0. ftp://ftp.atmforum.com/pub/approved-specs/af-tm-0056.000.ps, April 1996.

[6] G. Hasegawa, H. Ohsaki, M. Murata, and H. Miyahara. Performance evalulation and parameter tuning of TCP over ABR service in ATM networks. *IEICE transactions on communications*, E79-B(5):668–683, May 1996.

Table 3: Summary of Parameter Value Recommendations

Parameter	**Speed?**	**RTT?**	**Source initializes according to**	**Switch/Dest. modifies according to**
PCR	increases	no effect	link bandwidth or host/application capacity and pricing	bottleneck link bandwidth
MCR	increases	no effect	application requirements (e.g, video) and pricing	connection admission control (available resources)
ICR	increases	source takes minimum of signaled ICR and $\frac{TBE}{FRTT}$	pricing, host capacity and application	buffering and resources, PCR and FRTT
Nrm	maybe should increase with speed	no effect	processing speed and application type (real-time should increase it)	switch scheme and switch speed
Trm	decreases	no effect	processing speed and application type	switches can reduce Trm for a high PCR, or increase it for low switch speed
RIF	no, but may be decreased	no, but may be decreased	application requirements	EFCI and RRM switches and ER switches sensitive to RIF should reduce it depending on FRTT, PCR and scheme
RDF	no, but may be decreased	no, but may be decreased	application requirements	EFCI and RRM switches should reduce dependent on FRTT, PCR and scheme
ADTF	no effect	no effect (may increase)	application traffic characteristics	if little available resources, reduce ADTF
TBE	increases	increases	application type, pricing and host capacity	buffering and resources, and PCR and FRTT
CDF	may be smaller for high speeds	may be smaller for long RTTs	application type, confidence in TBE value	confidence in TBE value, confidence in links, availability of resources

[7] R. Jain. Congestion control and traffic management in ATM networks: Recent advances and a survey. *Computer Networks and ISDN Systems*, 28(13):1723–1738, November 1996.

[8] Raj Jain, Sonia Fahmy, Shivkumar Kalyanaraman, Rohit Goyal, Fang Lu, and Saragur Srinidhi. More strawvote comments: TBE vs queue sizes. ATM Forum/95-1661, December 1995.

[9] Raj Jain, Shivkumar Kalyanaraman, Sonia Fahmy, Rohit Goyal, and S. Kim. Source behavior for ATM ABR traffic management: An explanation. *IEEE Communications Magazine*, 34(11):50–57, November 1996.

[10] Raj Jain, Shivkumar Kalyanaraman, Sonia Fahmy, and Fang Lu. Out-of-rate RM cell issues and effect of Trm, TOF, and TCR. ATM Forum/95-0973R1, August 1995.

[11] Raj Jain, Shivkumar Kalyanaraman, Sonia Fahmy, and Fang Lu. Straw-vote comments on TM 4.0 R8. ATM Forum/95-1343, October 1995.

[12] Raj Jain, Shivkumar Kalyanaraman, Rohit Goyal, Sonia Fahmy, and Fang Lu. A fix for source end system rule 5. ATM Forum/95-1660, December 1995.

[13] Raj Jain, Shivkumar Kalyanaraman, Rohit Goyal, Sonia Fahmy, and Ram Viswanathan. ERICA switch algorithm: A complete description. ATM Forum/96-1172, August 1996.

[14] A. Koike, H. Kitazume, H. Saito, and M. Ishizuka. On end system behavior for explicit forward congestion indication of ABR service and its performance. *IEICE transactions on communications*, E79-B(4):605–610, April 1996.

[15] Ram Krishnan. Rate based control schemes for ABR traffic - design principles and performance comparison. In *Proceedings of the IEEE global telecommunications conference (GLOBECOM)*, volume 2, pages 1231–1235, November 1996.

[16] D. Lee, K. K. Ramakrishnan, W. M. Moh, and A. U. Shankar. Protocol specification using parameterized communicating extended finite state machines- a case study of the ATM ABR rate control scheme. In *Proceedings of ICNP*, pages 208–217, October 1996.

[17] D. Lee, K. K. Ramakrishnan, W. M. Moh, and A. U. Shankar. Performance and correctness of the ATM ABR rate control scheme. In *Proceedings of the IEEE INFOCOM*, volume 2, pages 785–794, March 1997.

[18] H. Ohsaki, M. Murata, and H. Miyahara. Robustness of rate-based congestion control algorithm for ABR service class in ATM networks. In *Proceedings of the IEEE global telecommunications conference (GLOBECOM)*, volume 2, pages 1097–1101, November 1996.

[19] K. K. Ramakrishnan, P. P. Mishra, and K. W. Fendick. Examination of alternative mechanisms for use-it-or-lose-it. ATM Forum/95-1599, December 1995.

[20] H. Saito, K. Kawashima, H. Kitazume, A. Koike, M. Ishizuka, and A. Abe. Performance issues in public ABR service. *IEEE Communications Magazine*, 34(11):40–48, November 1996.

[21] Nanying Yin. Analysis of a rate-based traffic management mechanism for ABR service. In *Proceedings of the IEEE global telecommunications conference (GLOBECOM)*, volume 2, pages 1076–1082, November 1995.

AUTHOR'S NOTE:

Some of the graphs have some text in a very small font on top. This is just used to mark the parameter values used in the simlation, and will not appear in the final version of the paper.