

## CHAPTER 7

### SOURCE RULE DESIGN FOR THE ABR SERVICE

In the earlier chapters of this dissertation we have examined switch scheme design in detail, including our proposals (OSU, ERICA and ERICA+). In this chapter, we will examine the design of source rules in the ATM Traffic Management framework. The source rules, as described in Chapter 2, determine the scheduling of data and bidirectional RM cells, the policies at the source when feedback is disrupted, or when the source does not use the allocated rate, and the response to binary and explicit rate feedback.

This dissertation work has helped design some of the source rules of the international standard (esp. SES Rules 5, 9, 11, and 13) and we shall examine the details later in this chapter. These rules can be broadly be considered as providing some open-loop functionality in ABR. SES Rules 5 and 13 deal with the problem of sources not using their allocated rates - the related policies are popularly called the Use-it-or-Lose-it policies. In a related work [70, 74], and in the introductory chapter 2, we consider parameter related issues in SES Rule 6. In Rule 9, we developed the “rescheduling” option to allow low rate sources to immediately use higher allocations. We helped develop SES Rule 11 which deals with the issue of low rate sources and out-of-rate

RM cells. In this chapter, we devote one section to the topic of Use-it-or-Lose-it policies, and another to the topic of low rate sources.

## 7.1 Use-It-or-Lose-It Policies

The ABR framework is predominantly closed-loop, i.e., sources normally change their rates in response to network feedback. Another form of control is open-loop control where sources change their rates independent of network feedback. Open-loop control can complement closed-loop control when the network delays are large compared to application traffic chunks, or when network feedback is temporarily disrupted. It is also useful to control applications which are bursty or source bottlenecked. Bursty application traffic alternates between active periods (application has data to send) and idle periods (application has no data to send). Source-bottlenecked applications cannot sustain a data rate as high as the network allocated rate. The ATM Forum debated on the issue of using open-loop control to reduce rate allocations of sources which do not use them. The proposed solutions, popularly known as the Use-It-or-Lose-It (UILI) policies, have had significant impact on the ABR service capabilities. In this section, we discuss and evaluate these policies, and their implications on the ABR service.

This section is subdivided as follows. Section 7.1.1 discusses the issues in the design of UILI policies. We then discuss early UILI proposals in section 7.1.2. We identify the problems with the early proposals in section 7.1.3 and present the final set of proposals which were debated in the ATM Forum in section 7.1.6. We then evaluate the performance of various alternatives in Section 7.1.12 and summarize the implications of UILI on ABR in section 7.1.13.

### 7.1.1 Issues in Use-It-or-Lose-It

When some VCs' present bursty or source-bottlenecked traffic, the network may experience underload even after rate allocation. It then allocates higher rates to all VCs without first taking back the unused allocations. As a result, the underloading sources retain their high allocations without using them. When these sources suddenly use their allocations, they overload the network. This problem is called "ACR Retention." A related problem is "ACR Promotion" where a source intentionally refrains from using its allocation aiming to get higher allocations in later cycles. The effect of ACR Retention/Promotion is shown in Figure 7.1. In the figure, before time  $t_0$  the source rate is much smaller than its ACR allocation. The ACR allocation remains constant. At time  $t_0$ , the source rate rises to ACR and the network queues correspondingly rise. These problems were first identified by Barnhart [7].

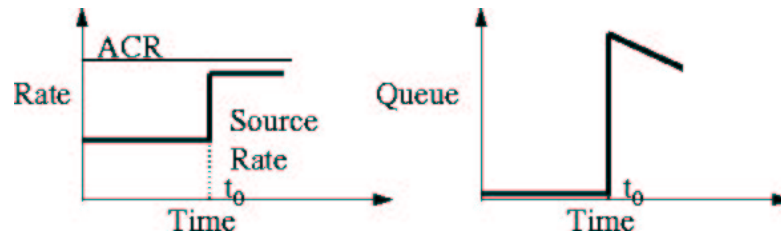


Figure 7.1: Effect of ACR Retention/Promotion

A solution to this problem is to detect an idle or source-bottlenecked source and reduce its rate allocation before it can overload the network. But this has an important side effect on bursty sources. If the rates are reduced after every idle period and the active periods are short, the aggregate throughput experienced by the source is

low. This tradeoff was discovered and studied carefully in the ATM Forum. The solutions proposed are popularly known as the Use-It-or-Lose-It (UILI) policies, referring to the fact that the source's ACR is reduced (lost) if it is not used.

### 7.1.2 Early UILI Proposals

The UILI function can be implemented at the SES (source-based) or at the switch (switch-based) or at both places. The early UILI proposals were all source-based. In these proposals, the test for ACR retention is done when an RM cell is being sent by the source. If ACR retention is detected, the source's ACR is immediately reduced using a rate reduction algorithm. Further, to prevent network feedback from overriding the ACR reduction, some proposals ignore the next feedback from the switch (if the feedback requests a rate increase). Over the February, April, May and June 1995 meetings of the ATM Forum, several UILI proposals were considered. The proposals differ in how the ACR retention is detected (additive or multiplicative metric), and in the algorithm used to reduce ACR.

In February 1995, Barnhart proposed a formula which reduced ACR as a function of the time since the last RM cell was sent or rate decrease was last done:

$$ACR_n = ACR_o(1 - T \times ACR_o/RDF)$$

$ACR_n$  is the new ACR and  $ACR_o$  is the old ACR. The time 'T' in the formula is the time which has transpired since the last *backward* RM cell was received or since the last ACR decrease. RDF is the rate decrease factor which is normally used to calculate the new rate for single-bit feedback. However, it is reused in the reduction formula to avoid choosing a new parameter. ACR retention is detected when the

source has sent out  $k$  RM cells ( $k$  is the Time out Factor (TOF) parameter) but does not hear from the network or has not decreased its rate during the same period.

In April 1995, several flaws with this proposal were corrected. Further, the ACR decrease function was found to be too aggressive and was changed to a harmonic function:

$$1/ACR_n = 1/ACR_o + T/RDF$$

The time ‘T’ in the function is now the time which has transpired since the last *forward* RM cell was sent. In the May and June 1995 meetings several other side effects were identified and corrected. For example, it was felt that the decrease function should not reduce the ACR below the negotiated Initial Cell Rate (ICR), because the source is allowed to start at that rate after an idle period. Kenney [83] observed that the harmonic ACR reduction formula was difficult to implement and proposed a linear reduction formula, which was similar to, but less aggressive than the February proposal:

$$ACR_n = ACR_o(1 - T \times TDF)$$

‘TDF’ is a new parameter called “Timeout Decrease Factor”. Incorporating these changes, the ABR SES (source) specification in August 1995 read as follows:

*“5. Before sending a forward in-rate RM-cell, if the time  $T$  that has elapsed since the last in-rate forward RM-cell was sent is greater than  $TOF \cdot N_{rm}$  cell intervals of  $(1/ACR)$ , and if  $ACR > ICR$ , then:*

*a) ACR shall be reduced by at least  $ACR * T * TDF$ , unless that reduction would result in a rate below ICR, in which case ACR shall be set to ICR, and TDF is equal to  $TDFR/RDF$  times the smallest power of 2 greater or equal to PCR,  $TDFR = \{ 0, 2^i, 2^j, 2^l \}$  (2 bits), where the values of the integers  $i, j,$  and  $l$  are to be determined*

*in the specification.*

*b) ACR shall not be increased upon reception of the next backward RM-cell.”*

The above UILI rule will also be interchangeably called “rule 5” henceforth, referring to the rule number in the ABR SES specification. The two parts are called “rule 5a” and “rule 5b” respectively.

### **7.1.3 Problems and Side Effects of Early Proposals**

In August 1995, Anna Charny et al [18] pointed out certain undesirable side effects in the above proposal. In particular, sources experience performance degradation in the transient phase when they increase from low ACR to high ACR. As a result, the links may be underutilized for a long period of time.

### **7.1.4 Worst Case Performance**

The worst case occurs when ICR is small and the source rises to a high rate from a very low rate, and when the backward RM cell (BRM) is received just before a forward RM cell (FRM) is sent. The BRM carries the network feedback and asks the source to increase its rate to a value greater than  $\text{TOF} \times (\text{old rate})$ . When the FRM is sent, the measured source rate  $S$  is close to the earlier low rate. This results in triggering UILI and the reduction of ACR by  $\text{ACR} \times T \times \text{TDF}$ . Now ACR is large and  $T$  is also large since it depends on the earlier low rate. Hence, ACR is reduced by a large amount upto ICR. Since ICR again is a small value, the cycle repeats when the BRM is received just before a FRM is sent. As a result, a source starting from a low ICR may never send at a rate higher than ICR.

### 7.1.5 Bursty and RPC Traffic Performance

Charny et al [18] also observed that bursty traffic having low ICR experienced a long-term performance degradation due to UILI resulting in large ACR fluctuations. Further, rule 5b prevents the increase of the source rate even though the network may have bandwidth available. In such bursty traffic configurations, it was found that rule 5a without rule 5b yielded better performance than both the parts together. However there was no way to selectively turn off rule 5b. Hence, it was decided to introduce a PNI (Prohibit No Increase) bit which when set turns off rule 5b selectively. Note that this also allows us to turn off rule 5 completely if TDF is also set to zero.

The performance degradation due to remote procedure call (RPC) ping-pong type traffic was independently observed by Bennet et al [11]. These authors pointed out that such applications may not want their rates to be decreased or reset to ICR after every idle period. They also suggested that UILI be performed by the switch and the source-based UILI be left optional.

We note that these side effects of rule 5 are not seen when the source is in the steady state (with source rate approximately equal to ACR) or in the transient phase when the source is decreasing from a high ACR to a low ACR. The main problem seemed to be due to the fact that the decrease function was proportional to T resulting in large ACR decreases after an ACR increase, leading to ramp-up delays.

Another problem which emerged was that some parameters like RDF and ICR were being used in multiple rules. Hence, choosing optimal values for these parameters became difficult due to their various side effects. These problems were addressed in the new set of proposals in December 1995 when the issue was voted upon to arrive at a final decision.

### 7.1.6 December 1995 Proposals

There were three main proposals in December 1995: the time-based proposal [8, 88, 90], our count-based proposal [56], and the switch-based proposal [102]. The time-based and the count-based proposals were later combined into one joint proposal. The ATM Forum voted between the switch-based proposal and the joint source-based proposal.

### 7.1.7 Unresolved UILI Issues

The following were the unresolved issues in UILI in December 1995. Essentially, a UILI proposal which works for both source-bottlenecked and bursty sources was desired.

- How to avoid UILI from affecting the normal rate increase (ramp up) of sources ?
- How long should the switch feedback be ignored after an ACR adjustment ?
- How to ensure good throughput and response time for bursty sources having small, medium and large active periods, when the idle periods are small, medium or large ?
- The floor of the August 1995 UILI ACR reduction function is ICR. If the source rate,  $S$ , is larger than ICR, the ACR may be reduced below the source rate down to ICR. We want a reduction function which does not decrease the ACR below the source's rate,  $S$ .
- "Headroom" measures how much the ACR is greater than the source rate,  $S$ , when it is declared as not an ACR retaining source. Should the headroom be



multiplicative ( $ACR \leq TOF \times S$ ) or additive ( $ACR \leq S + \text{headroom}$ ) ? Is a separate headroom parameter necessary (to avoid depending on ICR) ?

- Can UILI be done effectively in the switch ?
- Under what circumstances is UILI unnecessary or harmful ?

### 7.1.8 Count-Based UILI Proposal

The count-based UILI proposal [56] was made by us. It solved a large subset of the above problems and presented results of an extensive study on bursty traffic behavior.

#### Count vs Time

First, the count-based proposal removes the dependency of the ACR reduction function on the time factor,  $T$ , which is the time since the last FRM is sent. The reduction formula suggested is:

$$ACR = ACR - ACR \times TDF$$

The proposal is called “count-based” because a constant ACR decrease is achieved by triggering UILI  $n$  times. On the other hand, the time-based UILI decreases the ACR proportional to the time factor,  $T$ .

#### Multiplicative vs Additive Headroom

The count-based proposal uses an additive headroom for ACR detection ( $ACR \leq S + \text{headroom}$ ). Recall that if the ACR of the source is within the headroom, UILI is not triggered. The problem with multiplicative headroom ( $ACR \leq TOF \times S$ ) used in the August 1995 proposal is that depending upon the value of  $S$  it results in a large

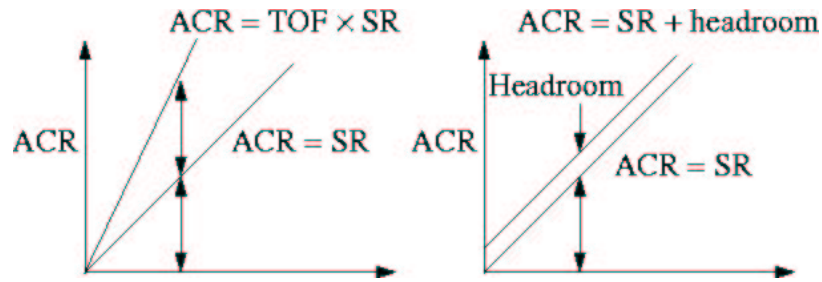


Figure 7.2: Multiplicative vs Additive Headroom

difference between ACR and source rate,  $S$ . A large difference ( $ACR - S$ ) results in large network queues when the source suddenly uses its ACR.

The additive headroom allows only a constant difference ( $ACR - S$ ) regardless of the source rate,  $S$ . The queue growth is hence bounded by a constant:  $(ACR - S) \times \text{Feedback Delay} \times \text{Number of Sources}$ . Hence, the additive headroom provides better network protection than the multiplicative headroom. The difference between the multiplicative and additive headroom is shown in Figure 7.2. Further, the latter is easier to implement since fewer multiply operations are required.

### Floor of the ACR Reduction Function

We also observed that the floor of the August 1995 UILI ACR reduction function is ICR and independent of the source rate,  $S$ . This is problematic because if  $S$  is larger than ICR, the ACR may be reduced below the source rate down to ICR. Therefore, we use a different floor function ( $S + \text{headroom}$ ) which ensures that the ACR does not decrease below  $S$  or the headroom. This floor function ensures that if the headroom equals the ICR, the ACR is guaranteed not to be decreased below ICR.

## Normal Rate Increase (Ramp Up)

The August 1995 proposal inhibited the ACR ramp up from a low rate because it triggered UILI immediately after the rate increase. Further, the amount of decrease could be large as explained in section 7.1.4.

Though our proposal does trigger UILI after ramp up from a low rate, it only reduces ACR by a step  $\Delta = \text{ACR} \times \text{TDF}$ . The next BRM cell brings the rate back to the ACR value before the decrease. If TDF is small, UILI is no longer triggered. For larger values of TDF, UILI may still be triggered multiple times. But, our new floor function ensures that the source rate consistently increases by at least the “headroom” value and eventually UILI is no longer triggered.

The count-based proposal also demonstrates a technique which avoids all oscillations due to normal rate increase. The UILI test is disabled *exactly once* after a normal rate increase. This allows the source rate to stabilize to the new (high) rate before the next UILI test, and thus UILI is not unnecessarily triggered. We use a bit called the PR5 (“Prohibit Rule 5”) bit which is enabled whenever there is a normal rate increase. The bit is cleared otherwise.

This technique also has one important side effect. Consider a source which is using its ACR allocation but suddenly becomes idle. Using the RM cells remaining in the network, the network may request a rate increase during the idle period. According to the above technique, the UILI test is disabled exactly once when the source becomes active again. Now observe that the first FRM cell opportunity after an idle period is the only opportunity for the source to reduce its ACR using UILI. This is because the memory of the prior idle period is lost when the next FRM is sent. As a result, UILI

is never triggered. However, the PR5 technique is not necessary and can be disabled if TDF is chosen to be small.

### Action on BRM

We observed that the ACR reduction function alone is not enough to ensure that that ACR retention is eliminated. For example, the August 1995 proposal requires that if the immediately next BRM feedback, after an UILI ACR reduction, requests a rate increase, and the PNI bit is not set, the BRM feedback is ignored. However, subsequent feedbacks may undo the ACR reduction and the problem of ACR retention still persists.

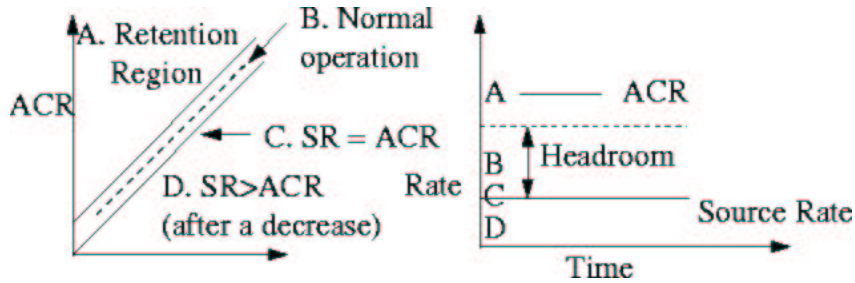


Figure 7.3: Regions of Operation

The count-based proposal ignores the BRM feedback as long as the source does not use its ACR allocation. The proposal uses the headroom area as a hysteresis zone in which network feedback to increase ACR is ignored. The proposal defines four regions of operation A, B, C, and D, as shown in Figure 7.3. Region A is called the ACR retention region. In this region,  $ACR > SR + Headroom$ , and UILI is triggered unless the PR5 bit (if used) is set. Region B is the headroom area. In this region,  $ACR \leq SR + Headroom$ , but  $ACR > SR$ . In this region BRM feedback requesting

Table 7.1: BRM Actions In The Different Regions Of Count-Based UILI

Region	Trigger UILI	Increase On BRM	Decrease On BRM
A	Yes unless PR5	No	Yes
B	No	No	Yes
C	No	Yes	Yes
D	No	Yes	Yes

increase is ignored. Region C has the source rate equal to ACR. Region D has source rate greater than ACR. Region D is touched briefly when the ACR decreases and the measured source rate is a mixture of the old and new ACRs. In regions C and D, the source obeys the feedback of the network to increase or decrease its ACR. In these regions, the source is not ACR retaining because its source rate is at least equal to its current ACR allocation. The actions in various regions are shown in Table 7.1. Note that there is no need for the PNI parameter, since UILI can be disabled by simply setting the parameter TDF to zero.

### Parameter Selection

The count-based proposal has two parameters: “headroom” and “TDF”. We recommended a separate “headroom” parameter is to avoid overloading the ICR parameter. This allows the ICR parameter to be set to a high value based on short-term congestion information. The headroom parameter can be set to a more conservative value. It controls how much the sources can lie about their rates at any time and determines how many cells the switch receives at once. However, as discussed in the

simulation results of bursty sources (Section 7.1.12), very small headroom is not desirable. A value of 10 Mbps is recommended. This allows LANE traffic to go at full Ethernet speed. Smaller values can be used for WANs.

The parameter TDF determines the speed of convergence to the desired UILI goals (region B in Figure 7.3). Hence, it determines the duration for which the network is susceptible to load due to sources suddenly using their ACRs. Larger values of TDF give faster convergence. However, a low value is preferred for bursty sources as discussed in Section 7.1.12, and TDF set to zero disables UILI. A value of 1/8 or 1/16 is recommended.

### **Pseudo Code For the Count-Based Proposal**

In the pseudo code for the count-based proposal given below, the variable 'ACR\_ok' indicates that the source has used its allocated ACR, and is allowed to increase its rate as directed by network feedback. The variable 'PR5' when set conveys the fact that the network has just directed an increase. 'SR' is a temporary variable and is not stored between successive execution of the code. Further, the proposal requested a separate parameter 'headroom' instead of using ICR in the UILI formula.

- **At FRM Send event:**

```
SR = Nrm/T;
```

```
ACR_ok = ((ACR ≤ SR) OR (TDF == 0.0));
```

```
IF (PR5 == FALSE)
```

```
    IF (ACR > SR + headroom)
```

```
        ACR = Max(SR + headroom, ACR × (1.0 - TDF));
```

```

        ENDIF
ELSE PR5 = FALSE;

• At BRM Receive event:
  IF (NI = 0 AND ACR_ok)
    IF (ACR < ER) PR5 = TRUE ELSE PR5 = FALSE;
    ACR = Min(ACR + AIR × PCR, PCR);
  ENDIF
  ACR = Min(ACR, ER);
  ACR = Max(ACR, MCR);

• Initialization
  ACR_ok = True;
  PR5 = False;

```

Note that the comparison ( $ACR \leq SR$ ) may always yield false due to the fact that cells may be scheduled only at certain fixed slots. There is typically a minimum granularity  $\Delta$  which dictates the cell scheduler at the source. To account for this scheduler, the comparison may be replaced by ( $ACR \leq SR + \Delta$ ).

### 7.1.9 Time-Based UILI Proposal

The time-based UILI proposal has a ACR reduction function which depends upon the time  $T$  since the last FRM was sent. While this aspect is similar to the August 1995 UILI proposal, the other changes suggested are:

1. The time-based proposal also independently observes the problem with using ICR as the floor of the reduction function (as discussed in Section 7.1.8). The proposal suggests two possible floor values:

a)  $ACR_{max} = \text{Max}(\text{ICR}, \text{TOF} \times \text{SR})$

b)  $ACR_{max} = \text{ICR} + \text{SR}$

2. IF (  $ACR > ACR_{max}$  )

$$ACR_{new} = \text{Max}( ACR \times (1 - T/Tc), ACR_{max} );$$

The recommended value for Tc is  $\text{Max}(\text{ADDF} \times \text{FRTT}, \text{TBE}/\text{PCR})$ , where ADDF has a default value of 2. FRTT is the Fixed Round Trip Time measured at connection setup.

The ACR reduction formula decreases ACR depending upon how long the idle period is compared to the round-trip time. A performance comparison of the count-based and the time-based alternatives is presented in Section 7.1.12.

### 7.1.10 Joint Source-Based UILI Proposal

The time-based and count-based camps agreed on a consensus, which we refer to as the “joint source-based proposal.” The proposal uses the count-based reduction function and a constant value for TDF. It uses the new floor of the reduction function and the additive headroom. However, ICR is used in the UILI function instead of the proposed “headroom” parameter. The hysteresis region (region B in Figure 7.3) suggested by the count-based proposal is not used. Rule 5b remains the same as the August 1995 proposal, and PR5 is not used since TDF is set to a small value (1/16), the count-based reduction formula is used.



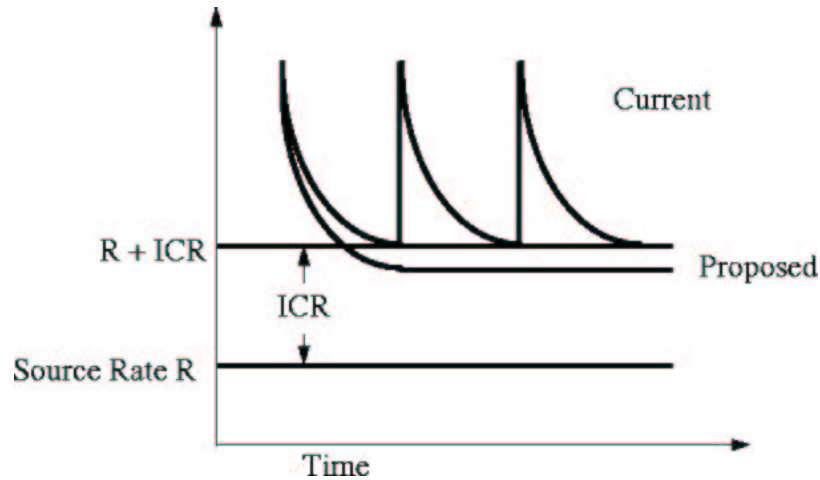


Figure 7.4: Joint Source-Based UILI Proposal vs Count-Based Proposal

The effect of removing the hysteresis region in the joint proposal is shown in Figure 7.4. In the joint proposal, the source will ignore one ER feedback after reducing the ACR to within the desired threshold. However, it may increase its rate-based upon ER feedback henceforth. The source thus re-enters the danger zone of ACR retention. In the count-based proposal, a source which reaches the desired operating zone ( $ACR \leq SR + ICR$ ), it remains in this region until the source actually uses its ACR allocation.

### 7.1.11 Switch-Based Proposal

AT&T [102] argued that the UILI function can be implemented in the switches on the following lines:

- Estimate rate of a connection and derive a smoothed average. This requires per-VC accounting at the switches.

- The switch maintains a local allocation for the VC based on the max-min fair allocation and the rate the VC claims to go at, i.e., its CCR.
- Use an “aging” function at the switch which allocates a rate to the VC based on the the ratio of the CCR and the actual rate-estimate. Basically, this function withdraws the allocations from ACR retaining sources.

A suggested aging function was  $(e^{\alpha u} - e^{\alpha \delta})$  where,  $u$  is the ratio of the expected rate and the actual rate, and,  $\alpha$  and  $\delta$  are parameters. The function has the property that the larger the difference between the CCR and the estimated actual rate, the greater the reduction factor. Essentially, the switch allocates conservatively to sources which it knows are not using their allocations.

A switch-based policy with no support from the source faces problems in handling sources which go idle because idle sources do not send RM cells. The switch may take away the allocation of an idle source after a timeout, but there is no way to convey this information to the idle source, since there are no RM cells from the source. Therefore, the switch-based UILI proposal suggests a simple timeout mechanism at the source which reduces the rate of the source to ICR after a timeout (parameter ATDF) of the order of 500 ms. Note that idle sources which become active before the timeout expires may still overload the network. The proposal does not implement UILI for such sources.

### 7.1.12 Simulation Results

In this section, we study the tradeoffs in the UILI design through simulation results. We look at both source-bottlenecked and bursty source configurations and present sample simulation results for the following five UILI alternatives:

1. No UILI
2. August 1995 UILI proposal
3. Baseline Rule 5 (enhanced August 1995) proposal, where the time-based reduction formula is replaced by the count-based formula, and an additive headroom (equal to ICR) is used in place of the multiplicative headroom.
4. The count-based UILI proposal
5. The time-based UILI proposal

A complete set of simulation results may be found in reference [81].

### **Source Bottlenecked Configuration**

The configuration is a network consisting of five ABR sources (Figure 7.5) going through two switches to corresponding destinations. All simulation results use ERICA switch algorithm. All links are 155 Mbps and 1000 km long. All VCs are bidirectional, that is, D1, D2, through D5 are also sending traffic to S1, S2 through S5. Some important ABR SES parameter values are given below. The values have been chosen to allow us to study UILI without the effect of other SES rules.

PCR = 155.52 Mbps, MCR = 0 Mbps, ICR = 155.52 Mbps, 1 Mbps  
RIF (AIR) = 1, Nrm = 32, Mrm = 2, RDF = 1/512  
Crm =  $\text{Min}\{\text{TBE}/\text{Nrm}, \text{PCR} \times \text{FRTT}/\text{Nrm}\}$   
TOF = 2, Trm = 100 ms, FRTT = 30 ms, TCR = 10 cells/sec  
TBE = 4096 (Rule 6 effectively disabled), CDF (XDF) = 0.5  
TDF =  $\{0, 0.125\} : \{0 \Rightarrow \text{No rule 5}, 0.125 \text{ for all versions of rule 5}\}$

$PNI = \{0, 1\} : \{1 \Rightarrow \text{No rule 5b}, 0 \Rightarrow \text{Rule 5b for August 1995 and Baseline UILI}\}$

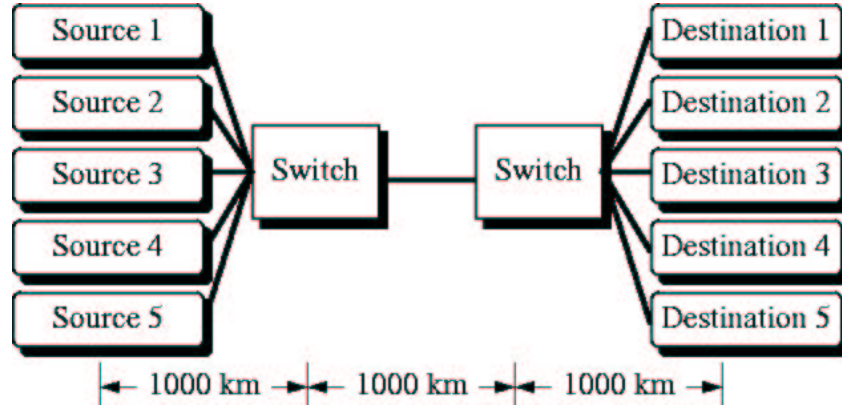
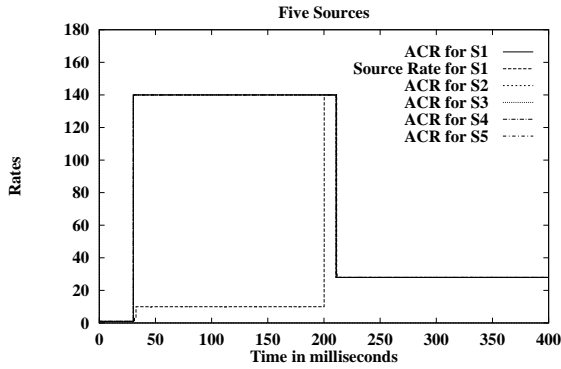


Figure 7.5: Five Sources Configuration

The simulation is run for 400 ms. For the first half of the simulation (200 ms), all the VCs are source-bottlenecked at 10 Mbps. After  $t=200$  ms, all sources are able to use their allocated rates.

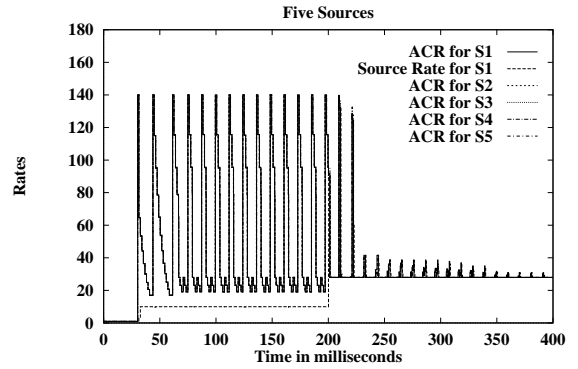
Figure 7.6 shows the ACR, and the actual source rates for the five UILI alternatives studied. There are six lines in each graph consisting of five ACR values and one actual source rate. Since all five sources are identical, the curves lie on the top of each other. With no UILI implemented (figure 7.6(a)) the ACR is initially much larger than the actual source rate. At 200 ms, the source rate jumps to the ACR and results in network overload. Figure 7.6(b) shows oscillatory behavior of the August 1995 proposal due to the wrong floor of the ACR reduction function. The Baseline UILI reaches the goal. However it oscillates between the goal and the network feedback. The count-based UILI converges quickly to the goal and does not have oscillations

2five-sources.r:option=8485:optionb=1:359:icr=1.0:time\_int=200.0:sw\_int=30:share=0.999:dis=1000:cif=4096:adf=0.5:tof=2.0:rtq=1.0:scr=10.0:trm=100000.0  
 adf=0.0:headroom=1.0:\_threshold=200000.0:maxsrcrate=10.0:HA=1.0:mib=20000:mib=20000:vbrate=124.41:rbv=120:a=1.15:b=1.05:qib=0.8 / Date:11/20/95  
 ICR: 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 / XRM: 128.00 128.00 128.00 128.00 128.00 128.00 128.00 128.00 128.00 / Graph: 2



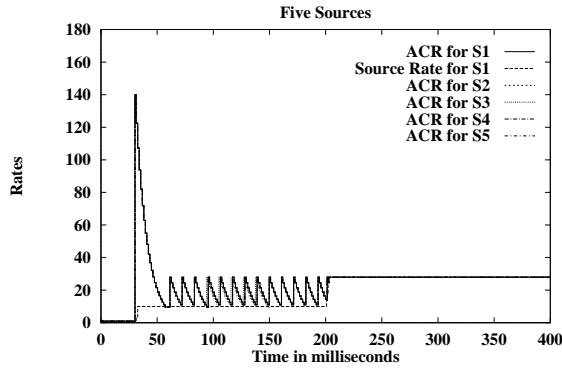
(a) No UILI

2five-sources.r:option=8485:optionb=79:icr=1.0:time\_int=200.0:sw\_int=30:share=0.999:dis=1000:cif=4096:adf=0.5:tof=2.0:rtq=1.0:scr=10.0:trm=100000.0  
 adf=0.125:headroom=1.0:\_threshold=200000.0:maxsrcrate=10.0:HA=1.0:mib=20000:mib=20000:vbrate=124.41:rbv=120:a=1.15:b=1.05:qib=0.8 / Date:11/20/95  
 ICR: 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 / XRM: 128.00 128.00 128.00 128.00 128.00 128.00 128.00 128.00 128.00 / Graph: 1



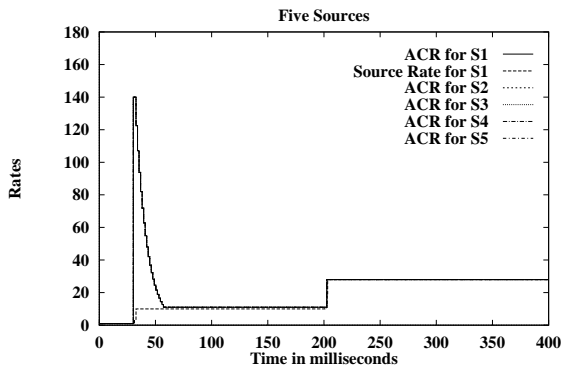
(b) Aug 1995 UILI

2five-sources.r:option=8485:optionb=71:icr=1.0:time\_int=200.0:sw\_int=30:share=0.999:dis=1000:cif=4096:adf=0.5:tof=2.0:rtq=1.0:scr=10.0:trm=100000.0  
 adf=0.125:headroom=1.0:\_threshold=200000.0:maxsrcrate=10.0:HA=1.0:mib=20000:mib=20000:vbrate=124.41:rbv=120:a=1.15:b=1.05:qib=0.8 / Date:11/20/95  
 ICR: 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 / XRM: 128.00 128.00 128.00 128.00 128.00 128.00 128.00 128.00 128.00 / Graph: 2



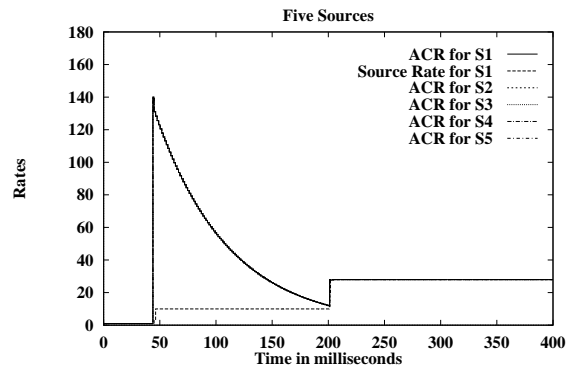
(c) Baseline UILI

2five-sources.r:option=8485:optionb=327:icr=1.0:time\_int=200.0:sw\_int=30:share=0.999:dis=1000:cif=4096:adf=0.5:tof=2.0:rtq=1.0:scr=10.0:trm=100000.0  
 adf=0.125:headroom=1.0:\_threshold=200000.0:maxsrcrate=10.0:HA=1.0:mib=20000:mib=20000:vbrate=124.41:rbv=120:a=1.15:b=1.05:qib=0.8 / Date:11/20/95  
 ICR: 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 / XRM: 128.00 128.00 128.00 128.00 128.00 128.00 128.00 128.00 128.00 / Graph: 6



(d) Count-Based UILI

2five-sources.r:option=8485:optionb=199:icr=1.0:time\_int=200.0:sw\_int=30:share=0.999:dis=1000:cif=512:adf=0.0:tof=2.0:rtq=1.0:scr=10.0:trm=100000.0  
 adf=0.125:headroom=1.0:\_threshold=200000.0:maxsrcrate=10.0:HA=1.0:mib=20000:mib=20000:vbrate=124.41:rbv=120:a=1.15:b=1.05:qib=0.8 / Date:12/01/95  
 ICR: 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 / XRM: 16.00 16.00 16.00 16.00 16.00 16.00 16.00 16.00 16.00 / Graph: 1



(e) Time-Based UILI

Figure 7.6: Five Source Configuration: Rates, Low ICR = 1.0 Mbps, Headroom = 1 Mbps, MaxSrcRate = 10 Mbps for 200 ms

after reaching the goal. The time-based UILI converges very slowly to the goal. Had the sources started using their ACR allocations earlier (than 200ms), it would have resulted in network overload.

## Bursty Sources

Recall that bursty sources have active periods when they send data at the allocated rate and idle periods when they do not have data to send. From the point of view of the bursty application, the following two measures are of interest (figure 7.7):

- *Burst response time* is the time taken to transmit the burst.
- *Effective throughput* is the average transmission rate of the burst.

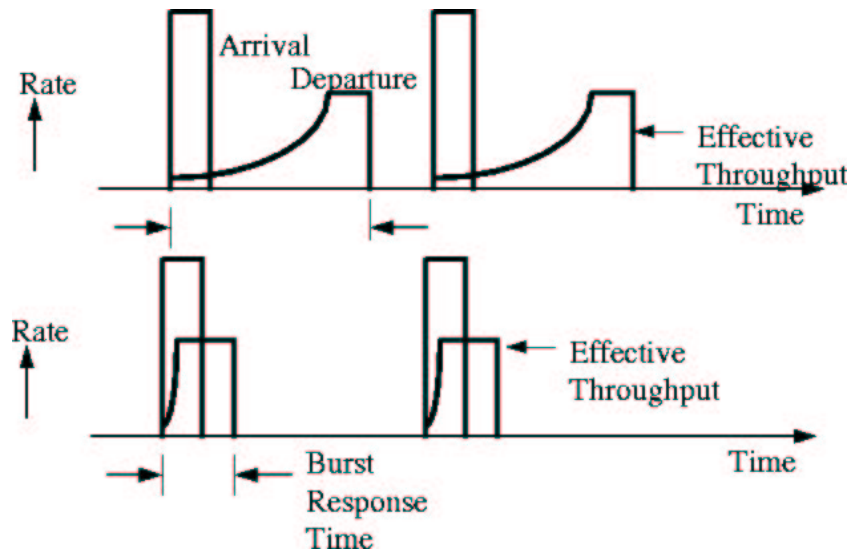


Figure 7.7: Burst Response Time vs Effective Throughput

Figure 7.7 shows the arrival and departure of a burst at an end system. The top part of the figure shows a burst which takes a long time to be transmitted, and the

bottom part shows one which is transmitted quickly. In the former case, the burst response time is short and effective throughput is higher, and vice versa for the latter case. Note that the effective throughput is related to the size of the burst and the burst response time.

Observe that the UILI goals conflict with the above bursty traffic performance goals. When UILI works, ACR is effectively reduced and a bursty source keeps restarting from low rates after every idle period. This results in a high burst response time which implies reduced performance. We study the effect of the UILI policy for different lengths of the active period: short (burst size is smaller than  $N_{rm}$ ), medium (burst time smaller than round trip time (RTT), but burst size larger than  $N_{rm}$ ) and large (burst time larger than RTT). Handling the network queues is usually not a problem for short or medium bursts. But it does become important when larger bursts active periods are used. The next section describes a model to generate short, medium and long bursts.

### **Closed-Loop Bursty Traffic Model**

We define a new “closed-loop” bursty traffic model as shown in Figure 7.8. The model consists of cycles of request-response traffic. In each cycle the source sends a set of requests and receives a set of responses from the destination. The next cycle begins after all the responses of the previous cycle have been received and an inter-cycle time has elapsed. There is a gap between successive requests called the inter-request time. The request contains a bunch of cells sent back-to-back by the application at rate PCR and the adapter controls the output rate to ACR.

The model as presented above may roughly represent World Wide Web traffic, transaction-oriented traffic, or client-server traffic. The model is “closed-loop” in the

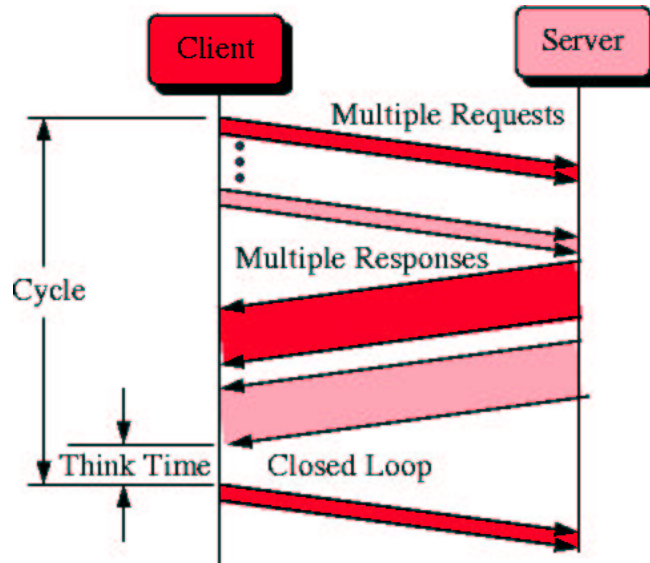


Figure 7.8: Closed-Loop Bursty Traffic Model

sense that the rate at which cycles (and hence requests) are generated depends upon the responsiveness of the network. If the network is congested the response take longer time to come back and the sources do not generate new requests until the previous ones have been responded to. In an “open-loop” traffic model like the packet-train model [64], bursts are generated at a fixed rate regardless of the congestion in the network.

Note that the time between two sets of requests (called a cycle time) is at least the sum of the time to transmit requests, the round-trip time and the inter-cycle time. Thus the idle time between two sets of requests is always greater than the round-trip time. All the RM cells from the previous set of requests return to the source before the new set of requests are sent. When a new burst starts there are no RM cells of the source in the network (ignoring second-order effects).



In our simulations, a cycle consists of one request from the client and one response from the server. We use a small response burst size (16 cells), and vary the request burst size.

### Single-Client Configuration and Parameter Values

The configuration we use is called the single-client configuration (Figure 7.9). It consists of a single client which communicates with the server, via a VC which traverses a bottleneck link. An infinite source is used in the background to ensure that the network is always loaded, and any sudden bursts of traffic manifest as queues. All the links run at 155 Mbps.

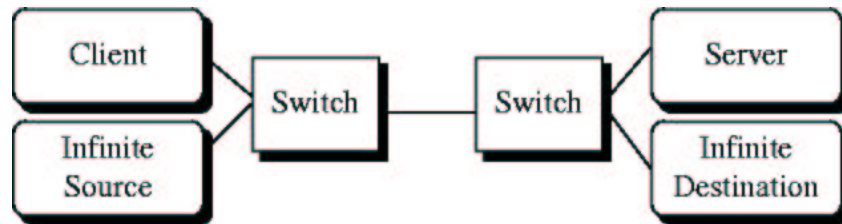


Figure 7.9: Client-Server Configuration With Infinite Source Background

The response size is kept constant at 16 cells. The request size can be 16, 256 or 8192 for small, medium or large bursts respectively. The inter-cycle time is chosen to be 1ms. All links are 500km long. The other source parameters are chosen to maximize ACR and disable the effects of other source rules:

ICR = 10 Mbps, TDF = 1/8, TCR = 10 cells/sec

TRM = 100 ms, TBE = 512, CDF = 0 to disable SES Rule 6.

The switch uses the ERICA algorithm to calculate rate feedback. The ERICA algorithm uses two key parameters: target utilization and averaging interval length.

The algorithm measures the load and number of active sources over successive averaging intervals and tries to achieve a link utilization equal to the target. The averaging intervals end either after the specified length or after a specified number of cells have been received, whichever happens first. In the simulations reported here, the target utilization is set at 90%, and the averaging interval length defaults to 100 ABR input cells or 1 ms, represented as the tuple (1 ms, 100 cells).

In the following sections, we pictorially describe the simulation results; a full set of graphs may be found in reference [81].

### Small Bursts

Small bursts are seen in LANE traffic. For example, the ethernet MTU, 1518 bytes is smaller than 32 (Nrm) cells. Since small bursts are smaller than Nrm cells, no RM cells are transmitted during certain bursts. As a result, no SES rules are triggered during these bursts. In other words, the entire burst is transmitted at one rate. However, when RM cells are finally transmitted, UILI is triggered which brings down the ACR to ICR. The source rate,  $S$ , is nearly zero due to the short burst time and long idle time. Hence,  $ICR + S$  is approximately equal to ICR.

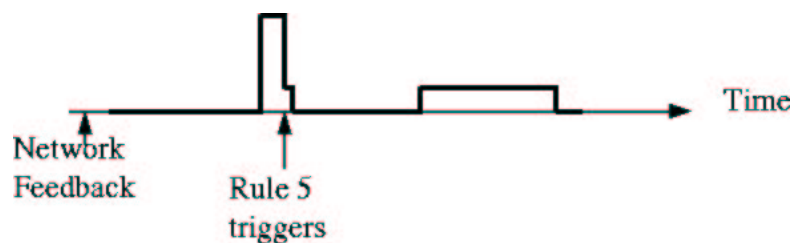


Figure 7.10: Effect of UILI on Small Bursts

Figure 7.10 shows the effect of ULLI on the source rate of small bursts. The network feedback first arrives when the source is idle, asking it to increase its ACR. The source uses its ACR to almost send the full burst. The first RM cell sent reduces its source rate back to ICR. The source rate goes back to zero when the source is idle. Now, the time-based and count-based proposals differ in the way they respond to subsequent network feedback.

In the time-based proposal, the feedback brought by the next RM cell is ignored because of rule 5b. Now there is no RM cell of the source in the network and at least two bursts are sent at ICR before the next RM cell is sent which results in an ACR increase. Note that the sending of this second RM cell does not decrease the ACR further because ACR is already at ICR. Therefore, on the average one out of every three bursts is sent at a higher rate.

In the count-based proposal, the rate-increase feedbacks are always ignored because the system is in region B (Figure 7.3). The ACR slowly reduces to ICR and then remains at ICR. Over the long term, all short bursts are sent out at ICR only. This can be improved by using a leaky bucket or GCRA [35] type burst tolerance mechanism where small bursts can be sent at link rate irrespective of ACR or ICR. Other alternatives include choosing a small TDF or a larger ICR. An ICR of 10 Mbps allows LANE traffic (the source of small bursts) to go through at full speed. On the other hand, since the burst is very short, there is not a significant time difference in transmitting the burst at ACR and transmitting it at ICR (assuming ICR is not very small). In such a case, the emphasis then shifts to supporting medium bursts and large bursts efficiently.

## Medium Bursts

Medium bursts are expected in ATM backbone traffic or in native mode ATM applications. Medium bursts contain more than  $N_{rm}$  (32) cells, but the active time is shorter than the round trip time. Though multiple RM cells are sent in a single burst, the network feedback for the burst arrives only after the burst has already been transmitted.

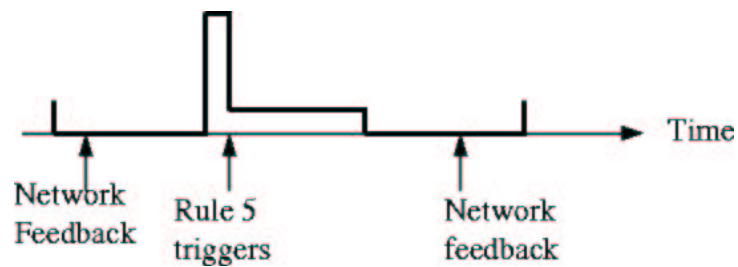


Figure 7.11: Effect of UILI on Medium Bursts

As shown in Figure 7.11, the UILI mechanism triggers once when the first RM cell is sent. In the time-based proposal, the amount of decrease is proportional to the idle time prior to the burst, while in the count-based UILI, the decrease is a constant amount. In the time-based proposal, if the idle time is large, almost the entire burst may be transmitted at ICR. Since, the count-based proposal sends the burst almost at  $ACR \times (1 - TDF)$ , it provides better burst response. Accordingly, simulation results in reference [81] show that the average source rate experienced by the bursts is higher for the count-based option (120 Mbps) compared to the time-based option (68 Mbps).

## Large Bursts

Large bursts are expected to be seen in backbone ATM links. Large bursts have a burst time larger than the round trip time. The network feedback returns to the source before the burst completes transmission.

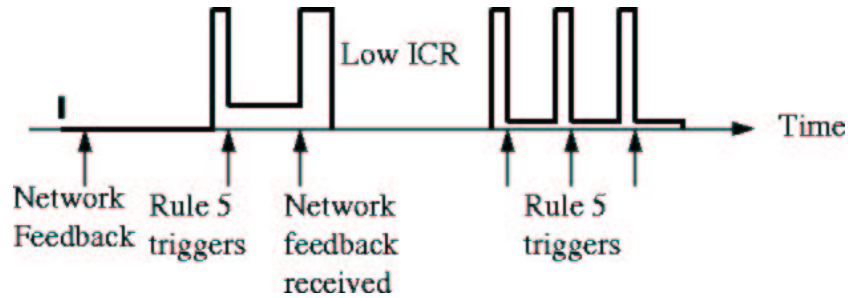


Figure 7.12: Effect of UILI on Large Bursts

Figure 7.12 shows the behavior of large bursts with the August 1995 proposal. When the burst starts, UILI triggers when the first RM cell is sent and brings the rate to ICR. Some part of the burst is transmitted at ICR. When network feedback is received, the ACR increases to the network directed value. If ICR is not very low there are no further oscillations and normal increase is not hampered. However if ICR is very low UILI is triggered after the ACR increase bringing the rate down to ACR again. The cycle is repeated and UILI triggers multiple times during the transmission of the burst resulting in low effective throughput and high burst response time.

The time-based UILI avoids the multiple triggering of UILI. It triggers once when the burst starts, and reduces the ACR proportional to the idle time. The count-based UILI also triggers once, and reduces the ACR by a constant value. Since the burst size is large, for large idle times ( $> RTT$ ), the network protection provided by the

count-based technique may be insufficient. However under such conditions a different SES rule (rule 6) can provide the required network protection.

### **7.1.13 Summary of UILI Alternatives and ATM Forum Decision**

The ATM Forum debated considerably over the UILI issue in December 1995 before putting the issue to vote. The summary of the arguments were the following:

The UILI policy can be implemented in switches or in NICs (sources) or both. The advantage of switch-only implementation is that NICs are simpler. The advantage of NIC implementation is that switches can be more aggressive in their bandwidth allocation without worrying about long-term implications of any one allocation. Without source-based UILI, the switches have to provision buffers to allow for overallocation of bandwidth.

Finally, the ATM Forum decided not to standardize an elaborate source-based UILI policy. A simple timeout is mandated for the source, where sources keep their rate allocations until a timeout (parameter ATDF, of the order of 500 ms) expires. After the timeout expires, ACR is reduced to ICR. The burden of implementing UILI is on the switches. However, NIC manufacturers can optionally implement a source-based UILI policy. The Informative appendix I.8 of the ATM Traffic Management 4.0 specification [35] briefly describes some source-based policies including the joint source-based proposal. The purpose of this paper has been to describe and evaluate the performance of various options.

## 7.2 Issues with Low Rate Sources

Normally, the RM cells sent by the source or turned around by the destination are counted in the source's rate (ACR). However, SES Rule 11 of the traffic management specification allows sources to send RM cells "out-of-rate" i.e., these RM cells are not counted towards the rate of the source. Out-of-rate RM cells are tagged with the CLP bit being set to one at the point of origination. The rate of such traffic is limited by the parameter TCR, which is 10 cells/s in all standard vectors. The text of the rule does not mandate the implementation of the out-of-rate RM cell policy.

We noted that if sources can set their ACRs to zero and switches can give a zero ACR feedback, then the out-of-rate mechanism is the only means to get out of the  $ACR = 0$  situation. Further, for unidirectional VCs the reverse direction has no data to send and may be initialized with a rate of zero. Under such conditions, the out of rate mechanism is required to either obtain a non-zero rate for sending BRM cells, or for sending the BRM cell itself as an out-of-rate cell. In general, the out-of-rate mechanism can be selectively used to improve transient performance.

Another related issue for low-rate sources is the source scheduling policy. Consider the event trace (from an actual simulation) shown in figure 7.13. Prior to the first event (at 42.2 ms), the bottleneck link capacity had been grabbed by VBR traffic, leading to  $ACR = 0$  allocation from the switch. At 42.2 ms, the out-of-rate RM policy triggers and an RM cell is sent, and the next opportunity to send is scheduled at 142.2 ms (due to the fact that ACR is currently zero, and the TCR value is 10 cells/s). Now new non-zero feedback is received from 43.2 ms (the switch uses the RM cells remaining in the network, including the out-of-rate RM cell). However, since the next send opportunity was scheduled at 142.2 ms, the rate allocations are ineffective

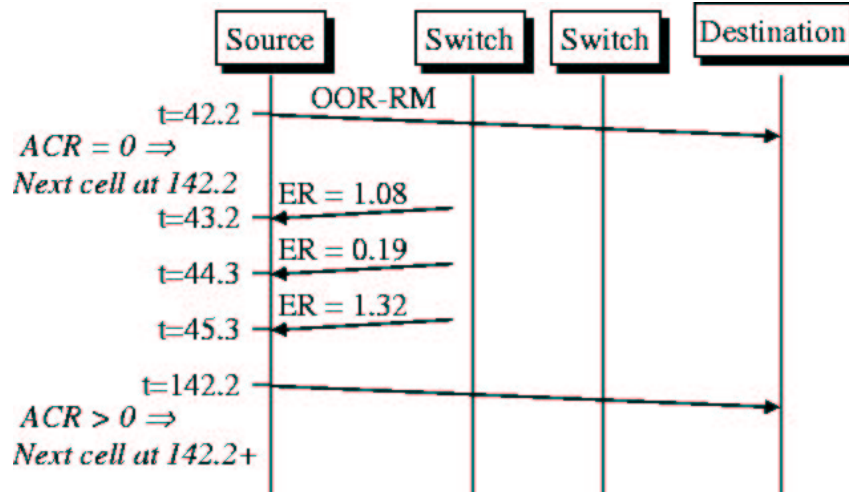


Figure 7.13: An event trace illustrating need for a rescheduling mechanism

till 142.2 ms, resulting in an unnecessary idle interval. In other words, the source once stopped is unable to use bandwidth for 100 ms even if it becomes available. The bandwidth is left unused not because there are no RM cells, but because the network feedback is ignored.

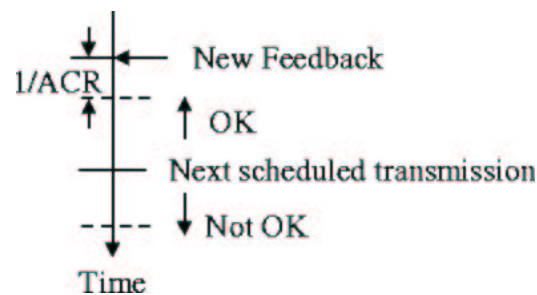


Figure 7.14: The Rescheduling Mechanism

To tackle this situation, we proposed that the sources may optionally “reschedule” their cell transmission opportunities based upon feedback. Specifically, the source



may reschedule upon receipt of new feedback if the next cell transmission opportunity calculated with the new ACR is earlier than the one currently scheduled. In terms of pseudo code:

```
IF( time_to_send > (now + 1/acr)
```

```
THEN time_to_send ← (now + 1/acr)
```

This mechanism is also illustrated in figure 7.14.

### **7.3 Summary of Source Rule Design Issues**

As explained in chapter 2, source and destination end system rules are important in complimenting the switch feedback calculation mechanisms. Specifically, the source rules provide “open-loop” control which is effective in cases when the source starts sending data after idle periods, and/or when the switch feedback to the sources is disrupted. Further, the sources have to consider the scheduling of RM and data cells, especially in the case of low rate traffic. This dissertation work has addressed the standardization aspects of Use-it-or-Lose-It policies (the issue arises when sources start sending data after idle periods), and that of low rate sources. Specifically, this work has helped design some of the source rules of the international standard (SES Rules 5, 9, 11, and 13).

## BIBLIOGRAPHY

- [1] Santosh P. Abraham and Anurag Kumar. Max-Min Fair Rate Control of ABR Connections with Nonzero MCRs. *IISc Technical Report*, 1997.
- [2] Yehuda Afek, Yishay Mansour, and Zvi Ostfeld. Convergence Complexity of Optimistic Rate Based Flow Control Algorithms. In *28th Annual Symposium on Theory of Computing (STOC)*, pages 89–98, 1996.
- [3] Yehuda Afek, Yishay Mansour, and Zvi Ostfeld. Phantom: A Simple and Effective Flow Control Scheme. In *Proceedings of the ACM SIGCOMM*, pages 169–182, August 1996.
- [4] Anthony Alles. ATM Internetworking. White paper, Cisco Systems, <http://www.cisco.com>, May 1995.
- [5] G.J. Armitage and K.M. Adams. ATM Adaptation Layer Packet Reassembly during Cell Loss. *IEEE Network Magazine*, September 1993.
- [6] Ambalavanar Arulambalam, Xiaoqiang Chen, and Nirwan Ansari. Allocating Fair Rates for Available Bit Rate Service in ATM Networks. *IEEE Communications Magazine*, 34(11):92–100, November 1996.
- [7] A.W.Barnhart. Changes Required to the Specification of Source Behavior. ATM Forum 95-0193, February 1995.
- [8] A.W.Barnhart. Evaluation and Proposed Solutions for Source Behavior # 5. ATM Forum 95-1614, December 1995.
- [9] A. W. Barnhart. Use of the Extended PRCA with Various Switch Mechanisms. ATM Forum 94-0898, 1994.
- [10] A. W. Barnhart. Example Switch Algorithm for TM Spec. ATM Forum 95-0195, February 1995.
- [11] J. Bennett, K. Fendick, K.K. Ramakrishnan, and F. Bonomi. RPC Behavior as it Relates to Source Behavior 5. ATM Forum 95-0568R1, May 1995.

- [12] J. Bennett and G. Tom Des Jardins. Comments on the July PRCA Rate Control Baseline. *ATM Forum 94-0682*, July 1994.
- [13] J. Beran, R. Sherman, M. Taqqu, and W. Willinger. Long-Range Dependence in Variable-Bit-Rate Video Traffic. *IEEE Transactions on Communications*, 43(2/3/4), February/March/April 1995.
- [14] U. Black. *ATM: Foundation for Broadband Networks*. Prentice Hall, New York, 1995.
- [15] P. E. Boyer and D. P. Tranchier. A reservation principle with applications to the atm traffic control. *Computer Networks and ISDN Systems*, 1992.
- [16] D. Cavendish, S. Mascolo, and M. Gerla. SP-EPRCA: an ATM Rate Based Congestion Control Scheme based on a Smith Predictor. Technical report, UCLA, 1997.
- [17] Y. Chang, N. Golmie, L. Benmohamed, and D. Siu. Simulation study of the new rate-based eprca traffic management mechanism. *ATM Forum 94-0809*, 1994.
- [18] A. Charny, G. Leeb, and M. Clarke. Some Observations on Source Behavior 5 of the Traffic Management Specification. *ATM Forum 95-0976R1*, August 1995.
- [19] Anna Charny. An Algorithm for Rate Allocation in a Cell-Switching Network with Feedback. Master's thesis, Massachusetts Institute of Technology, May 1994.
- [20] Anna Charny, David D. Clark, and Raj Jain. Congestion control with explicit rate indication. In *Proceedings of the IEEE International Communications Conference (ICC)*, June 1995.
- [21] D. Chiu and R. Jain. Analysis of the Increase/Decrease Algorithms for Congestion Avoidance in Computer Networks. *Journal of Computer Networks and ISDN Systems*, 1989.
- [22] Fabio M. Chiussi, Ye Xia, and Vijay P. Kumar. Dynamic max rate control algorithm for available bit rate service in atm networks. In *Proceedings of the IEEE GLOBECOM*, volume 3, pages 2108–2117, November 1996.
- [23] D.P.Heyman and T.V. Lakshman. What are the implications of Long-Range Dependence for VBR-Video Traffic Engineering ? *ACM/IEEE Transactions on Networking*, 4(3):101–113, June 1996.
- [24] Harry J.R. Dutton and Peter Lenhard. *Asynchronous Transfer Mode (ATM) Technical Overview*. Prentice Hall, New York, 2nd edition, 1995.

- [25] H. Eriksson. MBONE: the multicast backbone. *Communications of the ACM*, 37(8):54–60, August 1994.
- [26] J. Scott et al. Link by Link, Per VC Credit Based Flow Control. *ATM Forum 94-0168*, 1994.
- [27] L. Roberts et al. New pseudocode for explicit rate plus efci support. *ATM Forum 94-0974*, 1994.
- [28] M. Hluchyj et al. Closed-loop rate-based traffic management. *ATM Forum 94-0438R2*, 1994.
- [29] M. Hluchyj et al. Closed-Loop Rate-Based Traffic Management. *ATM Forum 94-0211R3*, April 1994.
- [30] S. Fahmy, R. Jain, S. Kalyanaraman, R. Goyal, and F. Lu. On source rules for abr service on atm networks with satellite links. In *Proceedings of First International Workshop on Satellite-based Information Services (WOSBIS)*, November 1996.
- [31] Chien Fang and Arthur Lin. A Simulation Study of ABR Robustness with Binary-Mode Switches: Part II. *ATM Forum 95-1328R1*, October 1995.
- [32] Chien Fang and Arthur Lin. On TCP Performance of UBR with EPD and UBR-EPD with a Fair Buffer Allocation Scheme. *ATM Forum 95-1645*, December 1995.
- [33] R. Fielding, J. Gettys, J. Mogul, H. Frystyk, and T. Berners-Lee. Hypertext Transfer Protocol – HTTP/1.1. Request For Comments, RFC 2068, January 1997.
- [34] ATM Forum. <http://www.atmforum.com>.
- [35] 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.
- [36] M. Garrett and W. Willinger. Analysis, modeling, and generation of self-similar vbr video traffic. In *Proceedings of the ACM SIGCOMM*, August 1994.
- [37] Matthew S. Goldman. Variable Bit Rate MPEG-2 over ATM: Definitions and Recommendations. *ATM Forum 96-1433*, October 1996.
- [38] Rohit Goyal, Raj Jain, Shiv Kalyanaraman, Sonia Fahmy, and Seong-Cheol Kim. Performance of TCP over UBR+. *ATM Forum 96-1269*, October 1996.

- [39] Rohit Goyal, Raj Jain, Shiv Kalyanaraman, Sonia Fahmy, Bobby Vandalore, Xiangrong Cai, and Seong-Cheol Kim. Selective Acknowledgements and UBR+ Drop Policies to Improve TCP/UBR Performance over Terrestrial and Satellite Networks. *ATM Forum 97-0423*, April 1997.
- [40] Rohit Goyal, Raj Jain, Shiv Kalyanaraman, Sonia Fahmy, Bobby Vandalore, Xiangrong Cai, and Seong-Cheol Kim. Selective Acknowledgements and UBR+ Drop Policies to Improve TCP/UBR Performance over Terrestrial and Satellite Networks. *ATM Forum 97-0423*, April 1997.
- [41] M. Grossglauser, S.Keshav, and D.Tse. RCBR: a simple and efficient service for multiple time-scale traffic. In *Proceedings of the ACM SIGCOMM*, August 1995.
- [42] S. Hrastar, H. Uzunalioglu, and W. Yen. Synchronization and de-jitter of mpeg-2 transport streams encapsulated in aal5/atm. In *Proceedings of the IEEE International Communications Conference (ICC)*, volume 3, pages 1411–1415, June 1996.
- [43] D. Hughes and P. Daley. More abr simulation results. *ATM Forum 94-0777*, 1994.
- [44] D. Hunt, Shirish Sathaye, and K. Brinkerhoff. The realities of flow control for abr service. *ATM Forum 94-0871*, 1994.
- [45] Van Jacobson. Congestion avoidance and control. In *Proceedings of the ACM SIGCOMM*, pages 314–329, August 1988.
- [46] J. Jaffe. Bottleneck Flow Control. *IEEE Transactions on Communications*, COM-29(7):954–962, 1980.
- [47] R. Jain. A delay-based approach for congestion avoidance in interconnected heterogeneous computer networks. *Computer Communications Review*, 19.
- [48] R. Jain. A timeout-based congestion control scheme for window flow-controlled networks. *IEEE Journal on Selected Areas in Communications*, 1986.
- [49] R. Jain. A comparison of hashing schemes for address lookup in computer networks. *IEEE Transactions on Communications*, 1992.
- [50] R. Jain. The eprca+ scheme. *ATM Forum 94-0988*, 1994.
- [51] R. Jain. The osu scheme for congestion avoidance using explicit rate indication. *ATM Forum 94-0883*, 1994.

- [52] R. Jain. Atm networking: Issues and challenges ahead. *Engineers Conference, InterOp+Network World*, 1995.
- [53] R. Jain. Congestion control and traffic management in atm networks: Recent advances and a survey. *Computer Networks and ISDN Systems*, 1995.
- [54] R. Jain, D. Chiu, and W. Hawe. A quantitative measure of fairness and discrimination for resource allocation in shared systems. *DEC TR-301*, 1984.
- [55] R. Jain, D. Chiu, and W. Hawe. A quantitative measure of fairness and discrimination for resource allocation in shared systems. *DEC TR-301*, 1984.
- [56] R. Jain, S. Kalyanaraman, R. Goyal, S. Fahmy, and F. Lu. A Fix for Source End System Rule 5. *ATM Forum 95-1660*, December 1995.
- [57] R. Jain, S. Kalyanaraman, R. Goyal, S. Fahmy, and F. Lu. Erica+: Extensions to the erica switch algorithm. *ATM Forum 95-1145R1*, 1995.
- [58] R. Jain, S. Kalyanaraman, and R. Viswanathan. Method and apparatus for congestion management in computer networks using explicit rate indication. *U. S. Patent application filed (S/N 307, 375)*, 1994.
- [59] R. Jain, S. Kalyanaraman, and R. Viswanathan. The transient performance: Eprca vs eprca++. *ATM Forum 94-1173*, 1994.
- [60] R. Jain, S. Kalyanaraman, R. Viswanathan, and R. Goyal. A sample switch algorithm. *ATM Forum 95-0178R1*, 1995.
- [61] R. Jain, K. Ramakrishnan, and D Chiu. Congestion avoidance scheme for computer networks. *U.S. Patent #5377322*, 1994.
- [62] R. Jain and K. K. Ramakrishnan. Congestion avoidance in computer networks with a connectionless network layer: Concepts, goals, and methodology. *Proc. IEEE Computer Networking Symposium*, 1988.
- [63] R. Jain, K. K. Ramakrishnan, and D. M. Chiu. Congestion Avoidance in Computer Networks with a Connectionless Network Layer. Technical Report DEC-TR-506, Digital Equipment Corporation, August 1987.
- [64] R. Jain and S. Routhier. Packet Trains - Measurement and a new model for computer network traffic. *IEEE Journal of Selected Areas in Communications*, 1986.
- [65] Raj Jain. Congestion Control in Computer Networks: Issues and Trends. *IEEE Network Magazine*, pages 24–30, May 1990.

- [66] Raj Jain. *The Art of Computer Systems Performance Analysis*. John Wiley & Sons, 1991.
- [67] Raj Jain. Myths about Congestion Management in High-speed Networks. *Internetworking: Research and Experience*, 3:101–113, 1992.
- [68] Raj Jain. ABR Service on ATM Networks: What is it? *Network World*, 1995.
- [69] Raj Jain. Congestion Control and Traffic Management in ATM Networks: Recent advances and a survey. *Computer Networks and ISDN Systems Journal*, October 1996.
- [70] Raj Jain, Sonia Fahmy, Shivkumar Kalyanaraman, Rohit Goyal, and Fang Lu. More Straw-Vote Comments: TBE vs Queue sizes. *ATM Forum 95-1661*, December 1995.
- [71] Raj Jain, Shiv Kalyanaraman, Rohit GOyal, and Sonia Fahmy. Source Behavior for ATM ABR Traffic Management: An Explanation. *IEEE Communications Magazine*, 34(11), November 1996.
- [72] Raj Jain, Shivkumar Kalyanaraman, Sonia Fahmy, and Fang Lu. Bursty ABR Sources. *ATM Forum 95-1345*, October 1995.
- [73] Raj Jain, Shivkumar Kalyanaraman, Sonia Fahmy, and Fang Lu. Out-of-Rate RM Cell Issues and Effect of Trm, TOF, and TCR. *ATM Forum 95-973R1*, August 1995.
- [74] Raj Jain, Shivkumar Kalyanaraman, Sonia Fahmy, and Fang Lu. Straw-Vote comments on TM 4.0 R8. *ATM Forum 95-1343*, October 1995.
- [75] Raj Jain, Shivkumar Kalyanaraman, Rohit Goyal, Ram Viswanathan, and Sonia Fahmy. Erica: Explicit rate indication for congestion avoidance in atm networks. U.S. Patent Application (S/N 08/683,871), July 1996.
- [76] Raj Jain, Shivkumar Kalyanaraman, and Ram Viswanathan. The osu scheme for congestion avoidance in atm networks: Lessons learnt and extensions. *Performance Evaluation Journal*, October 1997. to appear.
- [77] Raj Jain and Shivkumar Kalyanaraman Ram Viswanathan. ‘method and apparatus for congestion management in computer networks using explicit rate indication. U. S. Patent application (S/N 307,375), SepJuly 1994.
- [78] H. Tzeng K. Siu. Intelligent congestion control for abr service in atm networks. *Computer Communication Review*, 24(5):81–106, October 1995.

- [79] Lampros Kalampoukas, Anujan Varma, and K.K. Ramakrishnan. An efficient rate allocation algorithm for atm networks providing max-min fairness. In *6th IFIP International Conference on High Performance Networking (HPN)*, September 1995.
- [80] Shivkumar Kalyanaraman, Raj Jain, Sonia Fahmy, Rohit Goyal, and Jianping Jiang. Performance of TCP over ABR on ATM backbone and with various VBR traffic patterns. In *Proceedings of the IEEE International Communications Conference (ICC)*, June 1997.
- [81] Shivkumar Kalyanaraman, Raj Jain, Rohit Goyal, and Sonia Fahmy. A Survey of the Use-It-Or-Lose-It Policies for the ABR Service in ATM Networks. Technical Report OSU-CISRC-1/97-TR02, Dept of CIS, The Ohio State University, 1997.
- [82] D. Kataria. Comments on rate-based proposal. *ATM Forum 94-0384*, 1994.
- [83] J.B. Kenney. Problems and Suggested Solutions in Core Behavior. ATM Forum 95-0564R1, May 1995.
- [84] Bo-Kyoung Kim, Byung G. Kim, and Ilyoung Chong. Dynamic Averaging Interval Algorithm for ERICA ABR Control Scheme. ATM Forum 96-0062, February 1996.
- [85] H. T. Kung. Adaptive Credit Allocation for Flow-Controlled VCs. ATM Forum 94-0282, 1994.
- [86] H. T. Kung. Flow Controlled Virtual Connections Proposal for ATM Traffic Management. ATM Forum 94-0632R2, September 1994.
- [87] T.V. Lakshman, P.P. Mishra, and K.K. Ramakrishnan. Transporting compressed video over atm networks with explicit rate feedback control. In *Proceedings of the IEEE INFOCOM*, April 1997.
- [88] L.G.Roberts. Operation of Source Behavior # 5. ATM Forum 95-1641, December 1995.
- [89] Hongqing Li, Kai-Yeung Siu, Hong-Ti Tzeng, Chinatsu Ikeda, and Hiroshi Suzuki. Tcp over abr and ubr services in atm. In *Proceedings of IPCC'96*, March 1996.
- [90] S. Liu, M. Procanik, T. Chen, V.K. Samalam, and J. Ormond. An analysis of source rule # 5. ATM Forum 95-1545, December 1995.



- [91] B. Lyles and A. Lin. Definition and preliminary simulation of a rate-based congestion control mechanism with explicit feedback of bottleneck rates. *ATM Forum 94-0708*, 1994.
- [92] P. Newman. Traffic Management for ATM Local Area Networks. *IEEE Communications Magazine*, 1994.
- [93] P. Newman and G. Marshall. Becn congestion control. *ATM Forum 94-789R1*, 1993.
- [94] P. Newman and G. Marshall. Update on becn congestion control. *ATM Forum 94-855R1*, 1993.
- [95] Craig Partridge. *Gigabit Networking*. Addison-Wesley, Reading, MA, 1993.
- [96] Vern Paxson. Fast Approximation of Self-Similar Network Traffic. Technical Report LBL-36750, Lawrence Berkeley Labs, April 1995.
- [97] K. Ramakrishnan and R. Jain. A binary feedback scheme for congestion avoidance in computer networks with connectionless network layer. *ACM Transactions on Computers*, 1990.
- [98] K. K. Ramakrishnan, D. M. Chiu, and R. Jain. Congestion Avoidance in Computer Networks with a Connectionless Network Layer. Part IV: A Selective Binary Feedback Scheme for General Topologies. Technical report, Digital Equipment Corporation, 1987.
- [99] K. K. Ramakrishnan and P. Newman. Credit where credit is due. *ATM Forum 94-0916*, 1994.
- [100] K. K. Ramakrishnan and "Issues with Backward Explicit Congestion Notification based Congestion Control. Issues with backward explicit congestion notification based congestion control. *ATM Forum 94-0231*, 1993.
- [101] K. K. Ramakrishnan and J. Zavgren. Preliminary simulation results of hop-by-hop/vc flow control and early packet discard. *ATM Forum 94-0231*, 1994.
- [102] 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.
- [103] L. Roberts. The benefits of rate-based flow control for abr service. *ATM Forum 94-0796*, 1994.
- [104] L. Roberts. Enhanced prca (proportional rate-control algorithm). *ATM Forum 94-0735R1*, 1994.

- [105] L. Roberts. Rate-based algorithm for point to multipoint abr service. *ATM Forum 94-0772R1*, 1994.
- [106] Larry Roberts. Enhanced PRCA (Proportional Rate-Control Algorithm). *ATM Forum 94-0735R1*, August 1994.
- [107] A. Romanov. A performance enhancement for packetized abr and vbr+ data. *ATM Forum 94-0295*, 1994.
- [108] Allyn Romanov and Sally Floyd. Dynamics of TCP Traffic over ATM Networks. *IEEE Journal on Selected Areas in Communications*, May 1995.
- [109] W. Stallings. Isdn and broadband isdn with frame relay and atm. *ATM Forum 94-0888*, 1995.
- [110] Lucent Technologies. Atlanta chip set, microelectronics group news announcement, <http://www.lucent.com/micro/news/032497.html>.
- [111] Christos Tryfonas. MPEG-2 Transport over ATM Networks. Master's thesis, University of California at Santa Cruz, September 1996.
- [112] H. Tzeng and K. Siu. A class of proportional rate control schemes and simulation results. *ATM Forum 94-0888*, 1994.
- [113] H. Tzeng and K. Siu. Enhanced credit-based congestion notification (eccn) flow control for atm networks. *ATM Forum 94-0450*, 1994.
- [114] International Telecommunications Union. <http://www.itu.ch>.
- [115] Bobby Vandalore, Shiv Kalyanaraman, Raj Jain, Rohit Goyal, Sonia Fahmy, Xiangrong Cai, and Seong-Cheol Kim. Performance of Bursty World Wide Web (WWW) Sources over ABR. *ATM Forum 97-0425*, April 1997.
- [116] Bobby Vandalore, Shiv Kalyanaraman, Raj Jain, Rohit Goyal, Sonia Fahmy, and Pradeep Samudra. Worst case TCP behavior over ABR and buffer requirements. *ATM Forum 97-0617*, July 1997.
- [117] L. Wojnaroski. Baseline text for traffic management sub-working group. *ATM Forum 94-0394R4*, 1994.
- [118] Gary R. Wright and W. Richard Stevens. *TCP/IP Illustrated, Volume 2*. Addison-Wesley, Reading, MA, 1995.
- [119] Lixia Zhang, Scott Shenker, and D.D.Clark. Observations on the dynamics of a congestion control algorithm: The effects of two-way traffic. In *Proceedings of the ACM SIGCOMM*, August 1991.