
ATM Forum Document Number: ATM Forum/95-1660

Title: A Fix for Source End System Rule 5

Abstract:

In the October 1995 meeting of the ATM Forum, it was found that source end system rule 5 does not handle ACR retention problem correctly. It causes undesirable oscillations in many cases and significantly reduces the performance of the bursty sources. In this contribution, we analyze the solution that was proposed in October 1995 meeting and finally present a better solution. While the case of small bursts needs to be tackled, the rule works well for other simple configurations.

Source:

Raj Jain, Shiv Kalyanaraman, Rohit Goyal, Sonia Fahmy, and Fang Lu
The Ohio State University

Raj Jain is now at Washington University in Saint Louis, jain@cse.wustl.edu <http://www.cse.wustl.edu/~jain/>

Saragur Srinidhi
NASA Lewis Research Center and Sterling Software
Scientific Systems Division
21000 Brookpark Road, MS 141-1
Cleveland, OH 44135
Phone: 216-433-8987, Fax: 216-433-8000

The presentation of this contribution at ATM Forum is sponsored by NASA.

Date: December 9-15, 1995

Distribution: ATM Forum Technical Working Group Members
(Traffic Management)

Notice: This contribution has been prepared to assist the ATM Forum. It is offered to the Forum as a basis for discussion and is not a binding proposal on the part of any of the contributing organizations. The statements are subject to change in form and content after further study. Specifically, the contributors reserve the right to add to, amend or modify the statements contained herein.

Source Rule 5 was originally intended to avoid ACR Retention, a phenomenon where the source sends at a rate lesser than the assigned ACR but may suddenly use its ACR causing queues in the network. The solution is for the sources to automatically reduce ACR to to be close to their actual source rates.

In the October 95 meeting, one of the suggestions for Rule 5 (by Barnhart, Jain et al) was to remove the dependency of the Rule 5 decrease on T, the time since the last FRM was sent. That is, the formula

$$ACR = ACR - ACR * T * TDF$$

was to be replaced by:

$$ACR = ACR - ACR * TDF$$

and the multiplicative Tof in the triggering condition was to be replaced by an additive factor, ICR. Hence, instead of

$$ACR_ok = (ACR > ICR) \text{ .and. } (T \leq Tof * Nrm / ACR)$$

we have

$$ACR_ok = (ACR > ICR) \text{ .and. } (T \leq Nrm / (ACR - ICR))$$

We experimented with the new proposal and found that it works much better than the original Rule 5. However, it resulted in continuous oscillations. Once the source reaches the safety region after a series of Rule 5 drops i.e. $ACR < \text{Source Rate} + ICR$, an ER feedback can immediately bring the ACR to a high value. This can happen even with Rule 5b which ignores ER just once. So, there is oscillation between ER and desired value of ACR. We solved this problem by continually applying rule 5b unless the source does use all of its allocation.

Second minor problem was that there were unnecessary oscillations during normal rise from a low ACR to a high ACR even in the absence of ACR Retention. This is because the Source Rate assumes a transient value (which is somewhere between the old ACR and the new ACR) during normal rise and this is confused to be ACR Retention. This can be easily solved by introducing a boolean variable called PR5 that prohibits rule 5 for one RM interval after a normal rise.

Third minor problem was that the following two statements in the october proposal

$$\begin{aligned} ACR &= ACR - ACR * TDF; \\ ACR &= \text{Max}(ACR, \text{headroom}) \end{aligned}$$

can reduce ACR below source rate if ACR is initially above headroom (ICR in october proposal) but TDF is large so that ACR is reduced considerably. We solve this by combining the two lines and replacing them by:

$$ACR = \text{Max}(ACR * (1 - TDF), SR + \text{Headroom})$$

Finally, we noticed that in the proposal ICR also plays the role of "headroom." Here, headroom is defined as the maximum allowed difference between the ACR and the actual source rate Nrm/T . If the difference $ACR - Nrm/T$ exceeds the headroom, rule 5 triggers. The considerations for setting this headroom optimally are very different from those for setting "initial cell rate" and so the pseudocode proposed here uses two different parameters for these two different roles. This discussed further later in this contribution under parameter setting guidelines.

Before reading the following pseudocode, please realize that source rate SR is a temporary variable and is not stored between successive executions of the code.

```
/* ----- Sending FRM : New Rule 5a ----- */
SR = Nrm/T;
ACR_ok = ((ACR <= SR) || (TDF == 0.0));
IF (PR5 == FALSE)
    IF (ACR > SR + headroom)
        ACR = Max(SR + headroom, ACR*(1.0 - TDF));
    ENDIF
ELSE
    PR5 = FALSE;

/* ----- Receiving BRM : New Rule 5b ----- */
IF (NI = 0 AND ACR_ok)
    IF (ACR < ER) PR5 = TRUE;
    ACR = Min(ACR + AIR * PCR, PCR);
ENDIF
ACR = Min(ACR, ER);
ACR = Max(ACR, MCR);

/* ----- initialization ----- */
ACR_ok = TRUE;
PR5 = FALSE;
```

Explanatory Notes:

1. The ACR decrease formula does not allow a decrease below SR + headroom. Thus, sources are allowed to keep ACR anywhere between SR and SR+Headroom.
2. When the source is ready to rise, it has to use its current ACR before rising further.
3. ACR_ok is not used in the Rule 5 test. The test depends only upon the ACR, SourceRate and the headroom. ACR_ok is set to true when the source rate is at least ACR. Increase of ACR (ACR < ER) is not allowed when ACR_ok is false.
4. The PR5 (Prohibit Rule 5) bit forces Rule 5 to be ignored once when there is a normal increase.
5. The NI = 0 in the condition takes care of the EFCI based switches. For normal decrease or during ACR Retention, there is no need to add AIR since we are not going to increase ACR using this value of ER.

Basically, we divide the operation into four regions:

- A. ACR above source rate+headroom
- B. ACR between Source rate and source rate+headroom
- C. ACR is equal to source rate
- D. ACR is less than the source rate

Region A is where there is the problem of "ACR Retention." This is the triggering condition for Rule 5. As long as the source remains in this region, rule 5 continually triggers.

Region B is the normal region of operation. Rule 5 does not trigger. Once a system enters this region, the ACR generally remains fixed. Explicit increase feedbacks from the network are ignored until the source uses up its past allocation and hence moves to line C. This prevents oscillations.

Region C is actually a straight line. Region D happens generally during normal decrease. Only when the source uses up the past allocation (that is it is operating in Region C or D), it is allowed to use explicit increase feedbacks from the network.

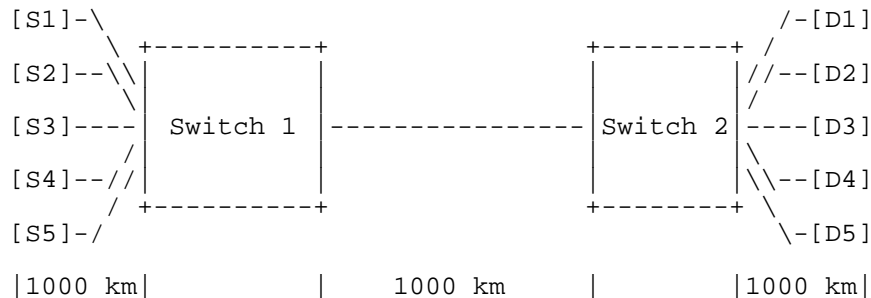
Region B acts as a hysteresis zone to prevent continuous oscillations. In this zone, ACR is held constant. The source is required to ignore explicit increases (ACR_ok remains false).

SIMULATION RESULTS:

We present simulation results for the following four alternatives for various configurations:

1. No rule 5
2. Baseline Rule 5 as in AF-TM95-0013R9
3. Oct Rule 5 as tentatively proposed by Barnhart and Jain, et al in October meeting
4. New rule 5 as proposed here

Figures 1 through 4 included in part 2 (PostScript) of this contribution show ACR, actual source rates and queue lengths for a network consisting of five ABR sources going through two switches to corresponding destinations. All simulation results use ERICA switch algorithm [1,2]. 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.



The following parameter values are used:

```

PCR = 155.52 Mbps
MCR = 0 Mbps
ICR = {155.52 Mbps, 1 Mbps} (Two values)
RIF (AIR) = 1
Nrm = 32
Mrm = 2
RDF = 1/512
Crm = Min{TBE/Nrm, PCR*FRRT/Nrm}
TOF = 2
Trm = 100 ms
  
```

FRTT = 30 ms
TBE = 4096 (Rule 6 is effectively disabled)
CDF (XDF) = 0.5
TDF = {0, 0.125} = {0 => No rule 5, 0.125 for all versions of rule 5}
PNI = {0, 1} = {1 => No rule 5b, 0=>Rule 5b for baseline and oct rule 5}
TCR = 10 c/s

The simulation is run for 400 ms. For the first one-half (200 ms), the VCs are bottlenecked (probably because of other VCs from that source) at 10 Mbps. The system is underloaded and asks all sources to increase. The ER for all sources is high. At t=200 ms, the situation changes. All sources are able to use their allocated ER. In all these graphs we have set XRM to 128 which ensures that rule 6 is not triggered. This avoids the oscillations and reductions caused by rule 6, which are the subject of another contribution.

Figure 1 shows ACRs and Source rates for the four alternatives studied. There are six lines in each graph consisting of five ACR values and one source rates. Since all five sources are identical, the curves lie on the top of each other. Notice that with rule 5 and baseline rule 5, the ACR can be high (compared to source rate). With Oct rule 5, the ACR comes down to source rate plus headroom. Here headroom value of 1 Mbps was used. However, after ignoring one explicit feedback from the network (rule 5b), it jumps back to the rate sent by the network and thus oscillates between approximately 30 Mbps and 11 Mbps. The new rule 5 does not have these oscillations.

Figure 2 shows queue lengths for the four alternatives depicted in Figure 1. Notice that time t=200 ms, the sources start using their high ACR and the queue length suddenly grows with no rule 5 and baseline rule 5. The oct rule 5 and the current rule 5 do not have this problem. The ACR and source rates are close to each other and the queue never goes over 5.

Figure 3 show the same information (ACR and source rates) as Figure 1 for ICR of 1 Mbps. Notice that baseline rule 5 causes oscillations. This problem is well known and was pointed out in [3,4,5,6]. October rule 5 oscillates between the goal and the network feedback. The new rule 5 does not have this problem.

Figure 4 show the queue lengths for the low ICR cases. Notice that no rule 5 and baseline rule 5 again result in large queues (even though ICR is low). Oct rule 5 and new rule 5 limit the queue length to 5.

In our ATM forum presentation, we will present an extended set of simulations with several parameter values and traffic patterns.

We have found that the proposal solves the ACR retention problem for all cases except small bursts. With small bursts, the averaged source rate is rather small and hence ACR comes down to headroom. There are three solutions:

1. Use a GCRA type of burst tolerance mechanism to let them go at PCR even when ACR is low
2. Use a small TDF 1/64
3. Ask for a larger headroom and possibly limit CIF (TBE). [To be studied]

PARAMETER SELECTION GUIDELINES

1. Initial Cell Rate:

As proposed here, ICR is used only during the VC setup and before the return of first RM cell. It does not have any longer term implication. Switches can set ICR to any value depending upon their current traffic. Subsequent feedback bring the rates to whatever is optimal later. For example, ICR = 155 Mbps may be acceptable in many networks.

2. Headroom:

The headroom controls "how much the sources can lie." Most switches will want this to be small. The system operates acceptably even if headroom is set to 0+ (very slightly above zero). The headroom is a commitment for the entire duration of the VC, which can be a long interval.

For many switch algorithms, the headroom will determine the maximum queue size after the ACR retention phase is over. Therefore, most switches will want to be conservative in setting headroom.

For small bursts, the headroom also controls the average throughput over a long interval. Such sources can benefit from a higher headroom value.

Based on these tradeoffs, we recommend a default headroom value of 10 Mbps. This will allow Ethernet frames (1518 bytes = 32 cells or less) to always go at their expected Ethernet rate.

Another flexibility with a separate headroom parameter is that it can be used to enable/disable rule 5. If headroom is 155 Mbps, rule 5 is effectively disabled. There is no need for PNI. However, we found that operation with high headroom can lead to large queues. Most network administrators will want to keep this new rule 5 enabled.

3. TDF:

This parameter controls the time to converge from "ACR Retention" to "No ACR Retention" phase. Larger TDF gives faster convergence. For infinite sources, we found that TDF=1/2 gives the best performance. However, for bursty sources, the average source rate is small and rule 5 causes undesirable decreases and so smaller value of TDFs give better throughput. We recommend using a default value of 1/8.

A FURTHER SIMPLIFICATION OF THE PROPOSED RULE 5:

We have experimented with many different variations of the proposal. Some were more complex while others were simpler. Simplifications, in general, lead to sacrificing performance in some cases. One such simplification, with only slight reduction in performance, is present below.

In this simplification, we remove the lines using the PR5 bit. The simplified pseudo code is given below.

```
/* ----- Sending FRM : New Rule 5a without PR5 ----- */
SR = Nrm/T;
ACR_ok = ((ACR <= SR) || (TDF == 0.0));

IF (ACR > SR + headroom)
    ACR = Max(SR + headroom, ACR*(1.0 - TDF));
ENDIF
```

```
/* ----- Receiving BRM : New Rule 5b ----- */
```

```
IF (NI = 0 AND ACR_ok)
    ACR = Min(ACR + AIR * PCR, PCR);
ENDIF
```

```
ACR = Min(ACR, ER);
ACR = Max(ACR, MCR);
```

```
/* ----- initially ----- */
ACR_ok = TRUE;
```

This code causes oscillations during normal rise. The size of the oscillations depends upon the TDF parameter. Larger values of TDF, cause larger oscillations. Fortunately, unlike the original rule 5, these oscillations die out fast. Recall that the original rule 5 (as in 95-0013R8 and before) caused persistent oscillations for small values of ICR. Thus, even the simplified rule 5 is better than the original rule 5.

References:

- [1] Raj Jain, Shiv Kalyanaraman, Ram Viswanathan, and Rohit Goyal, "A Sample Switch Algorithm," ATM Forum/95-0178R1, February 1995.
- [2] Raj Jain, Shiv Kalyanaraman, Rohit Goyal, Sonia Fahmy, and Fang Lu, "ERICA+: Extensions to the ERICA Switch Algorithm," ATM Forum/95-1346, October 1995.
- [3] Anna Charny, G. Leeb, and M. Clarke, "Some Observations on Source Behavior 5 of the Traffic Management Specification," AF-TM 95-0976R1, August 1995.
- [4] Raj Jain, Shiv Kalyanaraman, Sonia Fahmy, and Fang Lu "Out-of-Rate RM Cell Issues and Effect of Trm, TOF, and TCR," AF-TM/95-973R1, August 1995.
- [5] Raj Jain, Shiv Kalyanaraman, Sonia Fahmy, and Fang Lu "Straw-Vote Comments on TM 4.0 R8," AF-TM/95-1343, October 1995.
- [6] Raj Jain, Shiv Kalyanaraman, Sonia Fahmy, and Fang Lu "Bursty ABR Sources," AF-TM/95-1345, October 1995.

All our past ATM forum contributions and presentations can be obtained on-line:

<http://www.cse.wustl.edu/~jain/>