****

ATM Forum Document Number: ATM Forum/95-1661

****

Title: More Strawvote Comments: TBE vs Queue sizes

****

Abstract:

We analyze the role of Source End System rule 6 in depth. We find that the queue lengths can grow considerably in some cases even if the value of TBE (a.k.a. CIF) is small.

****

Source:

Raj Jain, Sonia Fahmy, Shiv Kalyanaraman, Rohit Goyal, 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)

****

****

Major Comment:

Section 5.10.3.1 on page 43 of 95-0013R9 and Normative Annex E on page 72 in 95-0013R9:

Current text on page 43:

"Transient Buffer Exposure, TBE, is the negotiated number of cells that the network would like to limit the source to sending during idle startup period, before the first RM cell returns"

Replacement:

"Transient Buffer Exposure, TBE, is the negotiated number of cells that the network would like to limit the source to sending during initial startup period, before the first RM cell returns"

Current Text on page 72:

None

Suggested addition on page 72:

"TBE limits the queue length only during initial startup and cannot be relied upon during the close 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."

JUSTIFICATION FOR CHANGE:

Both changes proposed above relate to source end system rule 6. In TM4.0 (R9 page 45), source end system rule 6 states that:

"Before sending an in-rate forward RM-cell, and after adjusting ACR according to #5 above, if at least Crm in-rate forward RM-cells have been sent since the last backward RM-cell with BN=0 was received, then ACR shall be reduced by at least ACR*CDF, unless that reduction would result in a rate below MCR, in which case ACR shall be set to MCR"

Crm (a.k.a. Xrm) is computed as:

"Crm = ceiling of min (TBE/Nrm, PCR*FRTT/Nrm)"

We have analyzed Rule 6 in depth. In particular, we have examined the following cases:

1. Initial source startup
2. Infinite sources with VBR in background

We found that rule 6 does limit the queue length inside the network in case 1. In this case, there are no previous cells in the network from the same VC. In other cases, the queue lengths can grow arbitrarily, regardless of the value of TBE. We show several such cases. We present simulation results first and then present an analytical explaination of when and why this happens.

SIMULATION RESULTS:

The effect of rule 6 on initial source start up is well known and so here present only cases where queue lengths grow arbitrarily and rule 6 has no effect. Figures 1 through 3 included in part 2 (PostScript) of this contribution 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 [1,2]. 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.

```
   [S1]-\ +----------+                +--------+ /-[D1]
```

```
        \|           |                 |        |/
 [S2]---| Switch 1 |-----------------|Switch 2|---[D2]
        /|           |                 |        |\
[VS1]-/ +----------+                 +---------+ \-[VD1]

|1000 km|          |    1000 km     |        |1000 km|
```

The following parameter values are used:

PCR = 155.52 Mbps
MCR = 0 Mpbs
ICR = Min{155.52, TBE/FRTT}
RIF (AIR) = 1
Nrm = 32
Mrm = 2
RDF = 1/512
Crm = Min{TBE/Nrm, PCR*FRTT/Nrm}
TOF = 2
Trm = 100 ms
FRTT = 30 ms
TBE = {128, 512, 1024} (Three values)
CDF (XDF) = {0, 0.5} = {Without rule 6, With Rule 6}
TDF = 0 (Rule 5a disabled)
PNI = 1 (Rule 5b disabled)
TCR = 10 c/s

Notice that we disabled rule 5 since rule 5 is undergoing a change and there are many versions of it that are being considered. The baseline rule 5 as specified in AF-TM 95-0013R9 causes unnecessary oscillations on normal rise and confuses the issue.

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 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 1 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 2 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.

Figure 3 shows ABR rates and queue graphs for TBE of 1024 cells. The queue length with and without rule 6 is higher than that with TBE of 512.

ANALYTICAL EXPLANATION:

The reason for the inadequacy of Rule 6 in limiting the queue growth can be explained as follows. Assume that a certain source S is sending forward RM cells at an average rate of R cells per second (cps). The RM cells are turned by the destination and the backward RM cells are received by S at a different rate r cps. In this case, the inter-FRM cell time at the source is 1/R while the inter-BRM cell time at the source is 1/r. Source end system Rule

6 will trigger at S if inter-BRM time is much larger (more than Crm times larger) than the inter-FRM time. That is if:

1/r >= Crm*(1/R)

or

R >= Crm*r

Thus, Rule 6, as defined, triggers only when the inter-FRM interval is very small compared to inter-BRM time. 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 without any background traffic, r is zero and rule 6 triggers after TBE cells.

However, if the BRM flow is not totally broken and R<Crm*r, then the cells can accumulate in the network at the rate of (R-r)*Nrm and not trigger rule 6. In such cases, the queues can grow substantially.

The source rule 6 protects sources from cell loss in case of link failures. In these cases, if we refer to the inequality discussed above, the rate at which the source S receives feedback, r, becomes zero, since feedback cannot be received due to link failure. Hence, the inequality is satisfied, and rule 6 protects the source S from losing a large number of cells. But link failure is better detected at lower layers of the protocol stack, and it does not seem essential for this function to be performed at this level. It is worth considering whether the additional complexity introduced by source end system rule 6, and its associated parameters Crm and CDF, are well worth the functionality it provides.

MOTION 1:

Change "Idle Startup" to "Initial Startup" in the definition of TBE in Section 5.10.3.1 on page 43 of 95-0013R9

MOTION 2:

Add the following paragraph just before the "MCR reservation" paragraph of Normative Annex E on page 72 in 95-0013R9

"TBE limits the queue length only during initial startup and cannot be relied upon during the close 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."

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.

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

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