Characterization of Cell Traffic Generated by an ATM Source

Julián ANDRADE  Wojciech BURAKOWSKI  Manuel VILLEN-ALTAMIRANO

Telefónica I+D, Madrid, Spain

In the study of the teletraffic aspects of ATM networks, one of the first problems which must be solved is to find a proper description of the traffic generated by the sources, especially those which have a variable bit rate. This paper proposes a set of three parameters which characterize the impact on network performance of the cell traffic generated by an ATM source. The proposed set has been obtained following a systematic testing methodology which assures its validity for a large variety of source types.

1 Introduction

The B-ISDN networks based on ATM allow a flexible integration of a large variety of services, with very different traffic characteristics. The main traffic characteristic to define the impact of a service on network performance is its bandwidth requirement. This bandwidth requirement may be constant along the call, as in CBR services, or variable to better meet user needs, as in the case of VBR services. It is easy to describe the bandwidth requirement in the CBR services since only one parameter, the cell rate, is sufficient. In the case of VBR sources, this description is quite more complex.

In this case, it is necessary to describe the changing bandwidth requirement, i.e., the cell generation process within a call, in a way which allows to determine the appropriate network dimensioning. An inaccurate description, e.g. by only the peak or the mean cell rate, would lead to network overdimensioning, and it would not be possible to obtain the desired multiplexing gain.

However, a detailed description of the cell generation process by means of a comprehensive set of parameters is not feasible either. Those parameters would have to be forecasted and used for dimensioning, declared by the user, evaluated by the CAC and surveyed by the policing function. Thus, the cell generation process of a call has to be described by a reduced set of parameters.

We consider that the identification of those key parameters is a crucial point in the solution of other traffic related problems which are presently under investigation. Let us examine some of these problems:

• Development of analytical methods: The analytical methods apply to simplified source models, with assumptions as, e.g., geometrically distributed number of cells in each state, a condition which probably will not be met by most of the actual sources. How can one know which theoretical model will give the same results as an actual source? We think that by identifying the key parameters of the actual source and using an equivalent model with the same values for them.

• Investigations on CAC function: The CAC function has to accept or reject a call based on its parameters. It is clear that only if they really represent the impact of the call on the network, can the CAC function be efficient.

• Investigations on policing function: This function has to detect violation of the declared parameters and to take the suitable actions. But, which are the suitable actions if a parameter with insignificant impact on the network performance is violated? Obviously, the design of an appropriate policing function requires an identification of the parameters which impact on the network, before solving the problem of how to police them.

• Design of audio and video codecs: The knowledge of these parameters is important to choose a proper design solution, i.e., a solution which optimizes the balance between fidelity of the codec and its impact on the network.

Several papers (see [1]-[5]) have studied the performance of a system substituting the real sources by equivalent ones which are analytically tractable. The variables defining the equivalent sources are obtained by matching certain parameters of the arrival process in the original and equivalent sources. The results obtained with both types of sources are compared, proving that the parameters chosen are representative for the cases studied.

This paper follows a systematic testing methodology to look for a set of parameters which define the impact of the sources on network performance for a large variety of sources. As a result of this approach, a set of
three parameters is proposed. Section 2 describes the methodology followed, section 3 presents the proposed set of parameters and section 4 summarizes the study performed, showing the tests results.

2 Methodology followed in the study

The influence of source traffic characteristics on the performance of an ATM multiplexer is taken as an indicator of their impact on network performance.

The ATM multiplexer is fed by a finite number of independent and homogeneous (generating cells according to the same process) sources. A burst/cell scale simulation model, with a constant number of calls along time has thus been built. The model consists of a queue with one server, which can only start a service in discrete instants separated by a constant time equal to the service time of a cell (i.e. to the duration of a slot). A general source model which allows to take a large variety of cell arrival processes into consideration has been used.

The performance of such a system is defined by the queue length distribution at cell arrivals. Since the service time is constant and the queue discipline is FIFO, this distribution is equal to the waiting time distribution measured in number of slots.

The number of waiting places has been considered fixed to a high value (1500) to avoid a new variable in the study. Nevertheless, the cell loss probability for smaller numbers of waiting places can be approximately estimated from the queue length distribution at arrivals.

The model uses the RESTART technique for accelerating rare event simulations described in a companion paper [6], which has allowed to obtain very confident simulation results for low probability values.

The validity of a tested set of parameters has been checked by simulating homogeneous sources with an arbitrary cell generation pattern and comparing the results with those obtained for other sources with another cell generation pattern but tested parameters with the same values. If the results are similar, the experiment is repeated for a variety of source patterns and parameter values to confirm the validity of the tested set.

An homogeneous source model has been used since it allows a better analysis of the results, and simplifies the search of the correct parameters; the satisfactory set of parameters found with homogeneous source models will have to be confirmed in the heterogeneous source case.

3 Proposed set of parameters

After many trials of different sets (the most relevant ones described in section 4), the following set of parameters, which have given satisfactory results in the tests made, are proposed:

- The mean cell rate:

\[
    m = \frac{E[N(t)]}{t}
\]

where \(N(t)\) is the number of arriving cells in a certain period of duration \(t\) and \(E[N(t)]\) is the expected value of \(N(t)\) when the beginning of the period is arbitrarily taken.

- The variance of the number of arriving cells during the mean interarrival time:

\[
    V(1/m) = E[N(1/m) - 1]^2
\]

- The ratio between the variance and the mean of the number of arrivals in a period \(t\), for the value of \(t\) in which this ratio is maximum:

\[
    \max_{t} \frac{V(t)}{M(t)} = \max_{t} \frac{E[N(t) - mt]^2}{mt}
\]

An equivalent source which follows a certain predefined pattern can be obtained for any arbitrary source by making the values of the three parameters equal in both sources. The pattern chosen for the equivalent source is ON/OFF with two states: the active state or burst has a geometrically distributed number of cells with constant interarrival time, and the silence state has a geometrically distributed number of slots. The equivalent sources obtained with this pattern, which is mathematically tractable, can be used for evaluations and dimensioning instead of the original source.

4 Summary of the study and checking of the parameters

The simulation results presented in this section correspond to eight types of sources, called here original sources, and their equivalent ones. As several sets of parameters have been tried out before choosing the proposed set, several equivalent sources have been obtained for each original one.

The original sources are numbered from 1 to 8. The equivalent ones have a name of the type XYZ, where \(X\) is the number of its original source, \(Y\), equal to \(R\) (rate) or \(V\) (variance), reflects the second parameter which has been matched, and \(Z\), equal to a number or to \(M\) (maximum), reflects the third parameter. Tables 1 and 2 describe the original and equivalent sources, respectively. The state diagrams refered in these tables are shown in Figure 1.

In all the cases studied, 80 homogeneous sources feed the ATM multiplexer, each offering a load of 0.01 Erl (\(m = 0.01\)), providing a total link load of 0.8 Erl.

Testing will start with three parameters commonly used in the literature and, according to the test results, they will be modified until those proposed in this paper are reached.
State type diagram | Active states | Silence state (duration) |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1 I A det(20)</td>
<td>det(10)</td>
<td>geo(1800)</td>
</tr>
<tr>
<td>2 I A dist(8.45)</td>
<td>det(10)</td>
<td>geo(760.5)</td>
</tr>
<tr>
<td>3 I A det(60)</td>
<td>det(10)</td>
<td>3 x geo(1800)</td>
</tr>
<tr>
<td>4 I A 4 x geo(12.9)</td>
<td>det(10)</td>
<td>4 x geo(1161)</td>
</tr>
<tr>
<td>5 I A 2 x geo(21.2)</td>
<td>det(10)</td>
<td>2 x geo(1908)</td>
</tr>
<tr>
<td>6 II A1 det(33)</td>
<td>det(22)</td>
<td>geo(2727)</td>
</tr>
<tr>
<td>7 III A1 det(33)</td>
<td>det(22)</td>
<td>geo(495.8)</td>
</tr>
<tr>
<td>8 I A1 det(63)</td>
<td>seq.</td>
<td>geo(5454)</td>
</tr>
</tbody>
</table>

\[ \text{det}(n) = \text{deterministic distribution with mean } n \]
\[ \text{geo}(n) = \text{geometric distribution with mean } n \]

\[ K \times \text{geo}(n) = \text{convolution of } K \text{ geometric distributions each one with mean } n \]

\[ x = \begin{cases} 
\text{geo}(1.5 \, n) & \text{with probability 0.5} \\
\text{geo}(0.5 \, n) & \text{with probability 0.5} 
\end{cases} \]

\[ \text{seq} = 33 \text{ interarrivals of 22 slots and 30 of 4 slots alternated in a fixed order: 22, 22, 4, 22, 4, etc...} \]

Time measured in slots.

**Table 1:** Description of the original sources.

<table>
<thead>
<tr>
<th>Model</th>
<th>( n )</th>
<th>( T )</th>
<th>( b )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 1R_\infty, 2R_\infty, 3R_\infty )</td>
<td>10.5</td>
<td>10</td>
<td>945</td>
</tr>
<tr>
<td>( 1RM, 2RM )</td>
<td>10.5</td>
<td>10</td>
<td>945</td>
</tr>
<tr>
<td>( 3R1120 )</td>
<td>31.8</td>
<td>10</td>
<td>2862</td>
</tr>
<tr>
<td>( 3R6000 )</td>
<td>15.2</td>
<td>10</td>
<td>1368</td>
</tr>
<tr>
<td>( 3RM, 4RM, 5RM )</td>
<td>24.1</td>
<td>10</td>
<td>2169</td>
</tr>
<tr>
<td>( 6RM )</td>
<td>14.1</td>
<td>7</td>
<td>1311</td>
</tr>
<tr>
<td>( 7RM )</td>
<td>6.2</td>
<td>7</td>
<td>576.6</td>
</tr>
<tr>
<td>( 7VM )</td>
<td>8.6</td>
<td>22</td>
<td>670.8</td>
</tr>
<tr>
<td>( 8RM )</td>
<td>27.8</td>
<td>7</td>
<td>2585</td>
</tr>
<tr>
<td>( 8VM )</td>
<td>31.7</td>
<td>13</td>
<td>2758</td>
</tr>
</tbody>
</table>

\( n \) : mean number of cells in the burst
\( T \) : interarrival time inside the burst
\( b \) : mean silence length in slots

**Table 2:** Description of the equivalent sources.

### 4.1 Long bursts with the same rate

Sources with long bursts (longer than the mean cell interarrival time \( 1/m \)) and the same constant cell rate in all their bursts were tested first (see cases 1, 2 and 3 of Table 1). Their equivalent sources were obtained adjusting these parameters:

- \( m \) : mean cell rate;
- \( R \) : cell rate during the burst \( (R = 1/T) \) where \( T \) is the interarrival time within the burst;
- \( V(\infty)/M(\infty) \) : limit of the ratio variance to mean of the number of cell arrivals in a period \( t \), when \( t \) tends to infinity.

The three sources have the same values for the three parameters, leading to a same equivalent source \( 1R_\infty = 2R_\infty = 3R_\infty \).

Figure 2a shows the waiting time distributions obtained from the simulations run for each type of source. It can be seen that, while similar results are obtained for the cases 1, 2 and \( 1R_\infty \), they are very different from those of case 3.

By observing Figure 3a, which provides the ratio variance to mean of the number of cell arrivals in a period \( t \), it can be deduced that these three parameters may be good to describe sources for which the maximum value of this ratio is at infinity, but not for sources which, as in the case of source 3, have this maximum for another value of \( t \). For these sources, the third parameter, \( V(\infty)/M(\infty) \), has to be changed.

Thus, new trials have been undertaken by replacing this parameter with the ratio variance to mean for other period lengths. Period lengths of 1120 and 6000 slots have been tried, obtaining the equivalent sources to the original source 3, \( 3R1120 \) and \( 3R6000 \) respectively.

Figure 3b shows how the ratios variance to mean have been matched for those period lengths. The obtained waiting times (see Figure 2b) are greater in the source \( 3R1120 \) and smaller in \( 3R6000 \). The main problem is that the result obtained is very different depending on the value of \( t \) taken to match the ratio \( V/M \) and, as was reported in [7], it is difficult to find criteria (if these criteria exist) to choose the appropriate value of \( t \). Thus, the parameter \( V/M \) for a certain value of \( t \) was rejected.

The next trial was to take as the third parameter, the maximum of ratio \( V/M \). Based on such parameter, the equivalent model \( 3RM \) was obtained which (see Figure 3b) has the same maximum of \( V/M \) as source 3 (although for a different value of \( t \)). Figure 2b shows that the results obtained for sources 3 and \( 3RM \) are fairly similar.
Fig. 2: Waiting time distributions of cases 1, 2, 3, 4, 5, 6 and equivalents

Fig. 3: V/M of cases 1, 2, 3, 4, 5, 6 and equivalents
This suggests the idea that the results depend on the maximum of $V/M$, but not (or, at least, not too much) on the value of $t$ for which this maximum occurs. This idea has been confirmed with several cases, e.g., Figure 2c shows the results obtained for cases 3, 4, 5 and 3RM(=4RM=5RM), which (see Figure 3c) reach the same maximum value of $V/M$ for different values of $t$ (1120, 1280, 1770 and infinity, respectively). The results obtained are quite similar in the four cases. The waiting times tend to be a little higher when the maximum of $V/M$ occurs for a higher value of $t$. The differences do not seem important. Anyway, as all the equivalent sources with geometric/geometric pattern have the maximum of $V/M$ for $t = 8$, their use for dimensioning lead to safe side solutions.

4.2 Long bursts with different rates between bursts

The definition of the second parameter considered in section 4.1, -the cell rate during the burst-, needs to be extended to be applied to original sources with bursts of different cell rates. An extension of this parameter is the equivalent burst cell rate, $R_{eq}$, defined as:

$$R_{eq} = \frac{\sum_{i} p_i n_i R_i}{\sum_{i} p_i n_i}$$

where $p_i$ is the proportion of bursts with $n_i$ cells and cell rate $R_i$ (This definition applies when the bursts have constant cell rate). Note that if all the bursts have the same cell rate, $R$, as occurred in section 4.1, then $R_{eq} = R$. Figure 2d shows the waiting time distribution for an original source of type 6 and for its equivalent one 6RM obtained by making equal:

- $m$, the mean cell rate;
- $R_{eq}$, the equivalent burst cell rate;
- $\max[V(t)/M(t)]$, the maximum of $V/M$.

The results are similar for both sources, arriving to the conclusion that these three parameters are able to define a source of these characteristics.

4.3 Long/short bursts with different rates between/within the bursts

Cases 7 and 8 have been obtained by rearrangement of the cells of case 6 in such a way that:

- The intearrival times of cells are the same (4 and 22), as in case 6 and in the same proportion, leading to a same value, 1/7, of $R_{eq}^*$;
- In case 7, there are ten short bursts, with only 3 cells;
- In case 8, the cell rate is continuously changing within the burst.

4a)  
4b)  
5a)  
5b)  

I : 95% Confidence interval
Fig. 4: Waiting time distribution of cases 7, 8 and equivalents
Fig. 5: $V/M$ of cases 7, 8 and equivalents

* The definitions of $R_{eq}$ have been applied to case 8 by considering its burst as a set of 63 consecutive bursts of rates 1/4 or 1/22. The same has been done applying the definition of instantaneous cell rate.
The comparison of results of cases 7 and 8 with their equivalent ones 7RM and 8RM, shown in Figures 4a and 4b, says to us that these parameters are not appropriate for this type of sources. The differences observed in the distribution for small values of \( t \) seem to indicate that the second parameter, the equivalent burst cell rate, is the responsible of this disagreement. Let us see how to change this parameter.

Assume that the instantaneous cell rate in all the instants of a burst is equal to the burst cell rate. Under this assumption \( R_{eq} \) is equal to the ratio of the second to the first moment of the instantaneous cell rate.

In case of long bursts with constant cell rate, the instantaneous cell rates are similar to the average cell rates in short periods, let us say in periods of duration \( 1/m \). Thus \( R_{eq} \) and

\[
R_{eq}' = \frac{E[N(1/m)/(1/m)]^2}{E[N(1/m)/(1/m)]}
\]

have similar values, e.g., in case 5, \( R_{eq} = 0.10 \) and \( R_{eq}' = 0.1092 \) and in case 6, \( R_{eq} = 0.14 \) and \( R_{eq}' = 0.11 \). On the contrary, in cases 7 and 8, with short bursts or bursts with changing cell rate, \( R_{eq} \) and \( R_{eq}' \) are quite different: \( R_{eq} = 0.14 \) and \( R_{eq}' = 0.04 \) in case 7, and \( R_{eq} = 0.14 \) and \( R_{eq}' = 0.07 \) in case 8. Thus, as the results show that:

- when \( R_{eq} \) is similar to \( R_{eq}' \) (cases 1 to 6), \( R_{eq} \) is a good parameter,
- when \( R_{eq} \) is very different from \( R_{eq}' \) (cases 7 and 8), \( R_{eq} \) is not a correct parameter,

it can be deduced that a better parameter could be \( R_{eq}' \) instead of \( R_{eq} \). Since:

\[ R_{eq}' = m[V(1/m) + 1] \]

\( V(1/m) \), the variance of the number of cell arrivals in a period of duration \( 1/m \), is taken as parameter.

We now have arrived to the set of three parameters stated in section 4. Based on this set, the equivalent sources 7VM and 8VM have been obtained. Figures 4a and 4b show that the problems of the RM equivalent sources for small values of \( t \) are solved using the VM ones.

If we observe the left part of the \( V/M \) curves in figures 5a and 5b, we see that 7VM and 8VM approximate to 7 and 8 better than 7RM and 8RM. However, in cases 1 to 6, in which there were long constant cell rate bursts, the left part of the \( V/M \) curves was quite approximated to the RM sources (see Figures 3a to 3d). Thus, it seems that the key points to define a source, apart form its mean cell rate, are the left part of the \( V/M \) curve of and, as was stated in section 4.1, the maximum value of such curve.

5 Conclusions

A set of three parameters which characterizes the impact on network performance of the cell traffic generated by an ATM source has been proposed. The number of parameters has been fixed to three, since a lower number would not allow to characterize a source with a minimum of accuracy, and a higher number, although providing more accuracy, would be impracticable to be used in network operation.

The proposed set has been obtained following a systematic testing methodology which assures its validity for a large variety of source types. Nevertheless, new tests need to be undertaken to confirm its validity with other source types, loads or number of sources per link, as well as in a heterogeneous source environment.

Once these pending tests have confirmed these parameters or led to new ones, a study has to be made on how to measure and police them or, in general, how to use them in network operation; these operation requirements may force some changes of the parameters, but these changes must be done bearing in mind that the basic requirement of the parameters is that they are able to define the impact of the services on network performance.

Thus, a large work is pending on this basic problem of cell traffic characterization. The authors desire that this contribution encourage others to investigate on it.

References


