

A Mathematical Model for IP over ATM

Irena Atov and Richard J. Harris

Royal Melbourne Institute of Technology,
GPO Box 2476V Melbourne, Vic. 3001, Australia,
{Irena, Richard}@catt.rmit.edu.au
<http://www.catt.rmit.edu.au/>

Abstract. We consider IP over ATM networking scenario, widely deployed today, where Asynchronous Transfer Mode (ATM) is used as a backbone network to provide high-speed transport for the Internet Protocol (IP) traffic. Telecommunication network carriers and Internet Service Providers (ISPs), that have deployed ATM as their backbone networks, need to model and characterize the IP traffic in order to plan and manage these networks to meet specified performance measures demanded by their customers. This paper addresses the problem of modelling IP traffic that is being transported over links of an ATM network. It provides mathematical models that can be used to take measurement data and translate it into a form that is suitable for various planning and management functions performed by network carriers.

1 Introduction

We consider an IP over ATM networking scenario (Fig. 1), where ATM is used as a backbone network to provide high-speed transport for the IP internetwork traffic. The IP over ATM networking scenario involves setting up a mesh of permanent virtual circuits (PVCs) between Internet Gateway Routers (IGRs) around an ATM cloud, and the Next Hop Resolution Protocol (NHRP) achieves a similar result with switched virtual circuits (SVCs). In such a scenario, the Internet traffic is first aggregated by IGRs before being sent on an ATM backbone. Typically, in this case each ATM virtual connection carries the traffic corresponding to a potentially large number of IP connections. We address the problem of modelling the aggregate traffic of multiple IP connections as it enters the ATM backbone into a form that is suitable for the dimensioning process of the ATM network.

ATM networks use the notion of *equivalent bandwidth* to transform the multi-level traffic problem associated with the ATM networks into a multi-slot circuit-switched problem, which enables ATM networks to be dimensioned in a fashion similar to telephony networks [1] [2] [3]. The Guérin et al [4] method for equivalent bandwidth assumes that the cell generation of each traffic type in the ATM network is a fluid flow on-off process, characterized by the parameters: mean cell rate, peak cell rate, and mean burst period. The ATM Forum has defined the same set of parameters as source traffic descriptors for the ATM Variable Bit

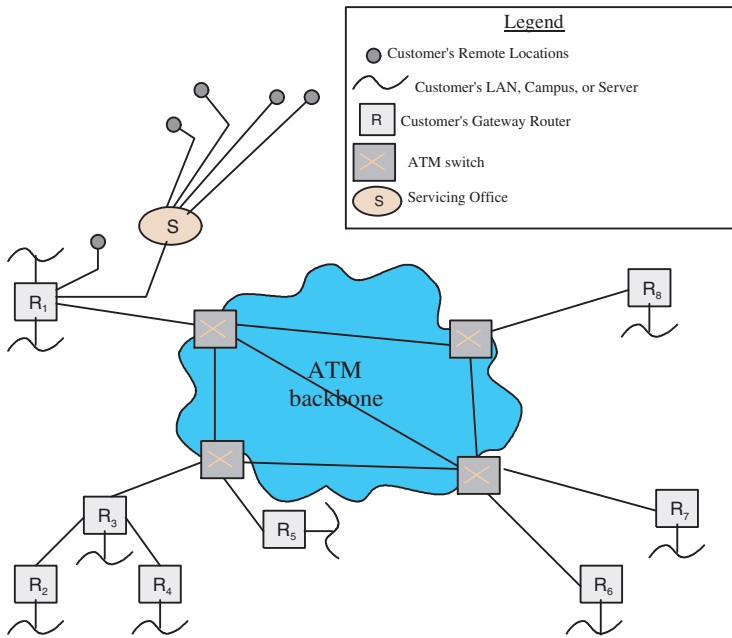


Fig. 1. IP Over ATM Networking Scenario

Rate (VBR) service [8]. For ATM networks, VBR is one of the effective solutions to accommodate IP traffic and in order to guarantee high quality service, it is important to select appropriate VBR parameter values. In addition, these three parameters are used for traffic policing at the ingress of the ATM network [8].

Our contribution can be summarized as follows. First, we model the traffic from multiple IP connections that is being transported over links of ATM network as an aggregate traffic stream characterized by the on-off traffic parameters: mean cell rate, peak cell rate, and mean burst period. The motivation is clearly to obtain a traffic model that is consistent with that used in [4] for determining an equivalent bandwidth. Then, we discuss approaches for incorporating our model into the ATM dimensioning procedures.

For the modelling, we analyze separately Transmission Control Protocol (TCP)-based and User Datagram Protocol (UDP)-based IP traffic, as they possess different characteristics that have to be accounted for when dimensioning the ATM backbone. The modelling of the TCP-based traffic involves characterization of the retransmissions of TCP/IP packets as a result of transmission errors inherent to ATM networks. When large TCP/IP packets are segmented to ATM cells in order to adapt to ATM transportation, a single cell lost at a buffer or corrupted in some way, will result in retransmission of the whole packet. Thus most of the cells comprising the packet are being retransmitted even though they were error free [5]. This repeated attempt aspect of the transmission of TCP traffic produces an overhead, which has to be taken into account when dimensioning

the ATM backbone. Furthermore, the modelling of both TCP-based and UDP-based traffic has to take account of the additional overhead associated with the protocols at Layer 2 and Layer 3 (i.e., the ATM Layer, and AAL5 Layer).

This paper is organized as follows: In Section 2 we introduce input parameters, notation and assumptions used for the modelling. Section 3 presents the IP over ATM model that can be used to map parameters of the input IP traffic mix to three principal parameters of the Guérin et al model [4], viz: mean, peak and burst period. Section 4 discusses approaches for incorporating the IP over ATM model into the ATM dimensioning procedures. Section 5 provides verification of the results from simulation studies. A summary and proposed directions for future research are given in Section 6.

2 Input Parameters and Notation

The IP over ATM model takes as input the traffic estimates or forecasted point-to-point traffic (i.e., IGR-to-IGR) that are categorized by: protocol type (TCP or UDP), packet sizes and packet arrival rates (packets per second). The model takes as input, the IP packet sizes measured as cells instead of bytes, in order to account for the protocol overheads applicable at Layer 2 and Layer 3. IP packets are variable in length and must be encapsulated and segmented to fit the fixed cell size requirements of ATM for transportation. The total data encapsulation overheads at the AAL5 and the ATM layer consist of: LLC+SNAP header with a length of 8 bytes, AAL5-trailer with a length of 8 bytes, PAD field with a length of (0-47) bytes and 5 byte ATM headers. The conversion of IP packet size from bytes (L_{byte}) to cells (L_{cell}) is easily obtained from $L_{cell} = \lceil \frac{L_{byte} + X}{48} \rceil$, where X denotes the protocol overhead, which is a sum of the LLC+SNAP header and the AAL5 trailer.

Notation used in the mathematical analysis is as follows:

N - IP packet length in cells

N_{MAX} - Maximum length of IP packet in cells

IP_{N_TCP} - Total number of TCP packets with length N cells per second

$P_{IP}(N)$ - IP packet error probability with length N cells

p - Cell error probability

p_{bit} - Bit error probability of ATM network

$P_{TCP}\{length = N\}$ - Probability that an IP packet has length N cells in the fresh TCP flow

$P'_{TCP}\{length = N\}$ - Probability that an IP packet has length N cells in the total TCP flow

F_{TCP} - Fresh TCP flow mean rate (i.e., offered traffic) [cell/s]

R_{TCP} - Repeated TCP flow mean rate [cell/s]

F'_{TCP} - Total TCP flow mean rate (i.e., carried traffic) [cell/s]

F_{UDP} - Fresh UDP flow mean rate [cell/s]

F'_{UDP} - Total UDP flow mean rate [cell/s]

m_{TCP} - Mean bit rate of the total TCP flow [bit/s]

$\sigma_{m_{TCP}}'^2$ - Variance of the bit rate of the total TCP flow [bit²/s²]
 IP_{N_UDP} - Total number of UDP packets with length N cells per second
 $P_{UDP}\{length = N\}$ - Probability that an IP packet has length N cells in the fresh UDP flow
 $P'_{UDP}\{length = N\}$ - Probability that an IP packet has length N cells in the total UDP flow
 m'_{UDP} - Mean bit rate of the total UDP flow [bit/s]
 $\sigma_{m_{UDP}}'^2$ - Variance of the bit rate of the total UDP flow [bit²/s²].
 m'_{IP} - Mean bit rate of the total IP flow [bit/s]
 $\sigma_{m_{IP}}'^2$ - Variance of the bit rate of the total IP flow [bit²/s²].
 R'_{peakIP} - Peak bit rate of the total IP flow [bit/s]
 b'_{IP} - Burst Period of the total IP flow [cells]

Assumptions used in the mathematical analysis:

1. $P_{TCP}\{length = N\}$ is known. It is derived from the input data.
2. $P_{UDP}\{length = N\}$ is known. It is derived from the input data.
3. p_{bit} is known.

It follows that:

$P_{IP}(N)$ can be easily derived from a Binomial distribution. An IP packet of length N cells will be repeated if at least one cell (out of N) is corrupt:

$$P_{IP}(N) = \sum_{i=1}^N \binom{N}{i} p^i (1-p)^{N-i} = 1 - (1-p)^N$$

Cell error probability p is derived from p_{bit} from a Binomial distribution, as well. One cell will be repeated if at least one bit out of 424 bits is corrupt (1 cell = 424 bits).

3 IP over ATM Model

3.1 TCP Model for the Mean Rate

In order to characterize the effects of the transmission errors on the packet arrival rates of the TCP traffic we make use of a simple model used in ITU-T documentation to analyze repeated attempt traffic in circuit switched networks. First, we briefly review the simple model and then we extend this model, in a straightforward manner, to the case of TCP traffic being transported across an ATM network.

A Simple Model for Repeated Attempts. Consider the following simplified picture of traffic that is presented to a telecommunication network (Fig.2). On the left of the diagram, fresh calls (first attempt traffic) are presented to the network. Due to the fact that the network has finite resources, some of these

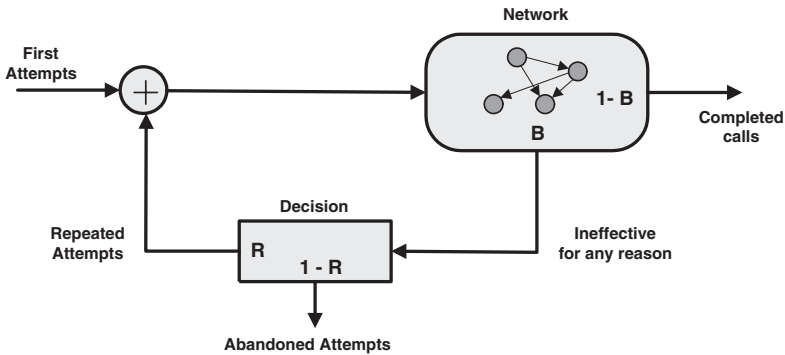


Fig. 2. Simple Model for Repeated Attempt Traffic

calls will be blocked with probability B . Conversely, the probability of call completion is given by $(1 - B)$. The ineffective calls can be abandoned or repeated. Calls are repeated with probability R and abandoned with probability $(1 - R)$. The repeated calls are added to the incoming fresh calls and represented to the network (at the '+' sign). By performing a simple analysis of this system and by defining the first attempt traffic using the symbol F and the total call attempts by T , one gets:

$$T = \frac{F}{(1 - RB)} \quad (1)$$

It should be noted that the model assumes that calls continue to repeat indefinitely until they are answered. Furthermore, the model is essentially deterministic in nature and contains no advanced stochastic modelling assumptions. Despite its simplicity, this model provides an excellent insight into the macro-operation of a circuit-switched network and the effect of repeated attempts on network performance.

Extension to TCP Traffic Carried on ATM Networks. In order to extend the model to the case of TCP traffic where packets are of variable length, we assume that a “call” is a flow consisting of TCP packets of constant length. We now give a detailed description of the new model that is based on flows that consist of TCP packets with constant length.

We define the fresh TCP flow F_{TCP} as a sum of N_{MAX} different fresh flows:

$$F_{\text{TCP}} = F_{\text{TCP}}(1) + F_{\text{TCP}}(2) + F_{\text{TCP}}(3) + \cdots + F_{\text{TCP}}(N_{\text{MAX}}) \quad (2)$$

where $F_{\text{TCP}}(N)$ represents the fresh flow consisting of IP packets of length N cells ($N = 1, 2, 3, \dots, N_{\text{MAX}}$):

$$F_{\text{TCP}} = \sum_{N=1}^{N_{\text{MAX}}} F_{\text{TCP}}(N) = \sum_{N=1}^{N_{\text{MAX}}} N \cdot IP_{N_TCP} \quad (3)$$

As a result of the above consideration, we can apply the method of the original model to obtain a formula for F'_{TCP} . Hence, as a general case we need to analyse only the case of *one* fresh flow $F_{\text{TCP}}(N)$.

The repeated flow R is simply given by the sum of N_{MAX} repeated flows caused by N_{MAX} different fresh flows: $R = R(1) + R(2) + \dots + R(N_{\text{MAX}})$, where $R(N)$ is the repeated flow caused by lost IP packets from the fresh flow $F_{\text{TCP}}(N)$, $R(N) = P_{\text{IP}}(N) \cdot F_{\text{TCP}}(N)$. From $F'_{\text{TCP}} = F_{\text{TCP}}(N) + R(N)$ it follows:

$$F'_{\text{TCP}} = \sum_{N=1}^{N_{\text{MAX}}} F'_{\text{TCP}}(N) \quad (4)$$

Here we demonstrate a way to derive F'_{TCP} using the original model for repeated attempts as the basis for the approach (we omit the subscript TCP for clarity):

$$1\text{-st step: } F'_1(N) = F(N) = N \cdot IP_{N_TCP}$$

$$R_1(N) = P_{\text{IP}}(N)F'_1(N) = P_{\text{IP}}(N)F(N)$$

$$2\text{-nd step: } F'_2(N) = F(N) + R_1(N) = F(N)(1 + P_{\text{IP}}(N))$$

$$R_2(N) = P_{\text{IP}}(N)F'_2(N) = F(N)(P_{\text{IP}}(N) + P_{\text{IP}}^2(N))$$

⋮
⋮
⋮

$$n\text{-th step: } F'_n(N) = F(N) + R_{n-1}(N) = F(N)(1 + P_{\text{IP}}(N) + \dots + P_{\text{IP}}^{n-1}(N))$$

$$R_n(N) = P_{\text{IP}}(N)F'_n(N) = F(N)(P_{\text{IP}}(N) + \dots + P_{\text{IP}}^n(N))$$

For $n \rightarrow \infty$, $F'_n(N)$ tends to:

$$F'_{\text{TCP}}(N) = \frac{F_{\text{TCP}}(N)}{1 - P_{\text{IP}}(N)} \quad (5)$$

Using (4), for the mean cell rate of the total TCP flow one gets:

$$F'_{\text{TCP}} = \sum_{N=1}^{N_{\text{MAX}}} \frac{F_{\text{TCP}}(N)}{1 - P_{\text{IP}}(N)} = \sum_{N=1}^{N_{\text{MAX}}} \frac{N \cdot IP_{N_TCP}}{1 - P_{\text{IP}}(N)} \quad (6)$$

Accordingly, the mean bit rate of the total TCP flow is:

$$m'_{\text{TCP}} = 424 \cdot F'_{\text{TCP}} \quad (7)$$

3.2 UDP Model for the Mean Rate

In a similar way, the fresh UDP flow F_{UDP} is modelled as a sum of N_{MAX} flows that consist of IP packets with constant length:

$$F_{\text{UDP}} = \sum_{N=1}^{N_{\text{MAX}}} F_{\text{UDP}}(N) = \sum_{N=1}^{N_{\text{MAX}}} N \cdot IP_{N_UDP} \quad (8)$$

The carried traffic consisting of IP packets of length N cells, $F'_{\text{UDP}}(N)$, is simply the product of the offered flow $F_{\text{UDP}}(N)$ and the probability of the traffic being successfully transmitted across the network. Thus, for $F'_{\text{UDP}}(N)$ we have:

$$F'_{\text{UDP}}(N) = N \cdot IP_{N_UDP}(1 - P_{\text{IP}}(N)) \quad (9)$$

Consequently, for the mean cell rate of the total UDP flow one gets:

$$F'_{\text{UDP}} = \sum_{N=1}^{N_{\text{MAX}}} N \cdot IP_{N_UDP}(1 - P_{\text{IP}}(N)) \quad (10)$$

The mean bit rate of the total UDP flow is simply: $m'_{\text{UDP}} = 424 \cdot F'_{\text{UDP}}$.

Finally, for the mean bit rate of the aggregate IP traffic carried over ATM network we have:

$$m'_{\text{IP}} = m'_{\text{TCP}} + m'_{\text{UDP}} \quad (11)$$

3.3 Model for the Peak Rate

Since, our interest is in mapping the parameters of the aggregate IP traffic into appropriate on-off traffic parameters, we can compute the peak rate from the following on-off relation: $\sigma'^2_{m_{\text{IP}}} = m'_{\text{IP}}(R'_{\text{peak}_{\text{IP}}} - m'_{\text{IP}})$. Thus, it remains only to derive the variance of the bit rate of the aggregate IP traffic, $\sigma'^2_{m_{\text{IP}}}$, which on the other hand can be expressed as a sum of the variances of the TCP and UDP flows as follows:

$$\sigma'^2_{m_{\text{IP}}} = \sigma'^2_{m_{\text{TCP}}} + \sigma'^2_{m_{\text{UDP}}} \quad (12)$$

From the distribution of IP packet length of the fresh flow, one can derive the probabilities that an IP packet has length N cells in the total TCP flow and UDP flow respectively, in the following way:

$$P'_{\text{TCP}}\{\text{length} = N\} = \frac{IP'_{N_TCP}}{\sum_{N=1}^{N_{\text{MAX}}} IP'_{N_TCP}} = \frac{\frac{IP_{N_TCP}}{1 - P_{\text{IP}}(N)}}{\sum_{N=1}^{N_{\text{MAX}}} \frac{IP_{N_TCP}}{1 - P_{\text{IP}}(N)}} \quad (13)$$

$$P'_{\text{UDP}}\{\text{length} = N\} = \frac{IP'_{N_UDP}}{\sum_{N=1}^{N_{\text{MAX}}} IP'_{N_UDP}} = \frac{IP_{N_UDP}(1 - P_{\text{IP}}(N))}{\sum_{N=1}^{N_{\text{MAX}}} IP_{N_UDP}(1 - P_{\text{IP}}(N))} \quad (14)$$

Then, for the variance of the IP packet length for the total TCP flow $\sigma'^2_{l_{\text{TCP}}}$ - we have:

$$\sigma'^2_{l_{\text{TCP}}} = \sum_{N=1}^{N_{\text{MAX}}} (N - \overline{N'})^2 \cdot P'_{\text{TCP}}\{\text{length} = N\} \quad (15)$$

Where $\overline{N'}$ is the mean IP packet length of the total TCP flow. From (13) for the number of IP packets with length N cells in the total TCP flow we have:

$$IP'_{N_TCP} = P'_{\text{TCP}}\{\text{length} = N\} \sum_{N=1}^{N_{\text{MAX}}} IP'_{N_TCP} = sP'_{\text{TCP}}\{\text{length} = N\} \quad (16)$$

The sum $\sum_{N=1}^{N_{\text{MAX}}} IP'_{N_TCP}$ is constant, and thus we designate it using the constant s . Then, by substituting (16) in (7) for the mean bit rate one gets:

$$m'_{TCP} = 424 \sum_{N=1}^{N_{\text{MAX}}} N \cdot IP'_{N_TCP} = S \cdot \overline{N'} \quad (17)$$

where, for simplicity, we use $S = 424 \cdot s$. Having a relation between the mean bit rate and the mean IP packet length of the total TCP flow, one can expect that the same relation holds between the bit rate, ν_b , and the IP packet length:

$$\nu_b = S \cdot N \quad (18)$$

The variance of the bit rate for the total TCP flow - $\sigma'^2_{m_{TCP}}$ can be derived from:

$$\sigma'^2_{m_{TCP}} = \sum_{N=1}^{N_{\text{MAX}}} (\nu_b - m'_{TCP})^2 \cdot P\{\nu_b = S \cdot N\} \quad (19)$$

By applying (17) and (18) in (19) and by assuming that the probability an IP packet has length N cells is equal to the probability that the bit rate is $S \cdot N$, $P\{\nu_b = S \cdot N\}$, for the variance of the bit rate for the total TCP flow we have:

$$\sigma'^2_{m_{TCP}} = S^2 \cdot \sigma'^2_{l_{TCP}} \quad (20)$$

The validity of this can be shown under the assumption that a time scale is divided into time slots with constant length T and that only one IP packet enters the system in each time slot. Consequently, the corresponding flow is one packet per unit of time or $s = \frac{1}{T}$ and the corresponding bit rate is $\nu_b = 424 \cdot N \cdot s = S \cdot N$.

The variance of the bit rate of the total UDP flow can be obtained in the same way as for the TCP flow, except in the above analysis one should apply (14) instead of (13).

Finally, in order to obtain an unbiased estimator for the variance of the bit rate of the total flow - $\hat{\sigma}'^2_{m_{IP}}$, (which is important because we estimate the variance directly from sample measurements) we have to multiply the estimated variance of bit rate, $\sigma'^2_{m_{IP}}$, with a factor to remove the bias in this calculation. In our case, the bias factor is $\frac{C}{C-1}$ where C is the number of observations, which, in our case, is the number of total IP packets entering the system per second, $\sum_{N=1}^{N_{\text{MAX}}} (IP_{N_TCP} + IP_{N_UDP})$. Since, we know the variance of the bit rate of the total flow (i.e its unbiased estimator) and the mean bit rate of the total flow (11), we can easily derive the peak bit rate of the total flow as:

$$R'_{peakIP} = \frac{\hat{\sigma}'^2_{m_{IP}}}{m'_{IP}} + m'_{IP} \quad (21)$$

3.4 Model for the Burst Period

One approach to determine the burst period (e.g., maximum number of cells that can be transmitted at peak rate) is similar to the one used to dimension

buffers inside IP routers. In the case of best-effort traffic, traditional rule is to dimension the buffer according to the bandwidth-delay (or *rtt*) product. In our case, where the aggregate IP traffic uses VBR service, the maximum burst size in cells is obtained as a product of the mean cell rate and the round-trip time, *rtt* :

$$b'_{IP} = \frac{m'_{IP}}{424} \cdot rtt \quad (22)$$

Some measurement studies have reported that such estimation works correctly for large aggregates of IP traffic [10].

In addition, to the above approach we are currently considering another approach for modelling the burst period, which is based on analysis of the User Parameter Control (UPC) functions at the ingress of ATM. Namely, we consider scenario where source is policed by two leaky buckets operating on different time-scales: one leaky bucket polices the peak rate, R'_{peakIP} , with a tolerance τ_p , and the other polices the mean rate, m'_{IP} , with a tolerance τ_m . These tolerances, τ_p and τ_m , translate as bucket sizes of the leaky bucket policers for the peak and the mean rate, respectively. Under the assumption that the source do not violate the negotiated peak rate, the maximum burst of cells that can pass through the mean rate policer is [8]:

$$b'_{IP} = \frac{\tau_m}{1 - \frac{m'_{IP}}{R'_{peakIP}}} \quad (23)$$

In order to obtain an estimate of the tolerance τ_m , we study G/D/1/N - delay loss system, which is an exact model for the cell loss probability of the leaky bucket mechanism (if violating cells are discarded). One approximate solution for the cell loss probability of such system with on-off source model, which is sufficiently accurate in most cases, has been reported in [9]. We want to determine the cell loss probability as a function of the queue capacity (bucket size), for a given traffic load and given leaky bucket rate. The leaky rate is fixed by introducing an overdimensioning factor C ($C > 1$), as a product of this factor and the source mean rate. The relationship between the cell loss probability and the bucket size has two distinct parts: a cell-scale component which decreases exponentially (linearly on a log scale) with the increase of the bucket size and the burst component where the slope of the curve changes dramatically as we reach certain bucket size. The intercept between these two components define solution for the required bucket size and tolerance τ_m . However, we do not present any results here as this is still an ongoing research.

4 Incorporating the Model into the ATM Dimensioning

Our model defines the aggregate IP traffic between two OD pairs (e.g. IGR-to-IGR) as a single IP traffic stream¹ and provides its characterization in terms

¹ Note that the model can be extended to incorporate QoS and thus define the aggregate IP traffic as multiple IP traffic streams. For that, it is necessary for the input traffic parameters to be categorized by QoS as well.

of the three parameters: mean rate, peak rate, and burst period. From these parameters one can calculate directly the equivalent bandwidth required for this traffic stream and thus model the aggregate IP traffic as an “IP call” with a capacity requirement equal to the equivalent bandwidth.

The equivalent bandwidth computed for the “IP call” can now be used in two different ways. The first approach is to model this aggregate IP traffic as call arriving at some average arrival rate (Poisson distributed) and persisting for an average time period (Negative exponentially distributed); it requires the nominal equivalent bandwidth for the duration of the call. This approach enables these “IP calls” to be treated in a “unified” fashion along with other ATM calls that can be characterized by their mean arrival rate, service time and equivalent bandwidth. When dimensioning ATM networks using this approach, a call loss probability is specified and the capacity determined using their mean arrival and service rates in conjunction with their equivalent bandwidths. The second approach recognizes the fact that IP traffic coming from a LAN or WAN is likely to be “always on” and that we simply need to allocate capacity based on the equivalent bandwidth computed for our single “IP call” emanating from the LAN or WAN. Adopting the first approach in this case would lead to gross over-dimensioning of the capacity requirements. This means that if such traffic is present, one should carefully separate the “always on” traffic from the traffic that can be characterized by an arrival and service rate.

Finally, we consider another approach that can be taken for modeling the aggregate IP traffic stream to fit the “unified” ATM dimensioning model. We can transform our VBR “IP call” of capacity equivalent to its equivalent bandwidth into equivalent CBR “IP call” of different characteristics. According to the equivalent burst approximation [6], a general traffic stream which is the superposition of a large number of independent traffic streams can be replaced by an equivalent process with simpler characteristics. The equivalent process is chosen to have the same values for the following parameters: (1) Mean cell rate, m (2) Variance of instantaneous rate, σ^2 (3) Asymptotic variance of the number of cell arrivals in a long time interval, ν . This equivalent traffic process has a Poissonian arrival process λ , for independent but equally distributed bursts with peak rate R , and bursts durations exponentially distributed with mean μ . The fitting relations are given by:

$$\lambda = \frac{2m^2}{\nu} \quad R = \frac{\sigma^2}{m} \quad \mu = \frac{\nu}{2\sigma^2} \quad (24)$$

Thus, by fitting the mean, the instantaneous variance, and the asymptotic variance of the aggregate IP traffic into this equivalent process, we get as a result a CBR “IP call” of capacity R , and traffic intensity $\lambda\mu$. Accordingly, it only remains the asymptotic variance of the aggregate IP traffic to be derived. The expression for the asymptotic variance of the number of offered cells in a long time interval, follows from the application of the formulae for asymptotic variance of a cumulative regenerative process [7].

5 Simulation Results

Purpose built real-time simulator was used to analyse the analytical model. The simulator is a very simple implementation of the IP over ATM model. However, it tries to test the model with more realistic data flows. The simulator tests the effects of the transmission errors on the packet arrival rates. To be able to easily understand these effects and their behaviour, the system is kept very simple by limiting it to one ATM link.

Traffic from individual IP connections is modelled as sessions which arrive according Poisson with given mean and negative exponential durations with mean values selected from a given range. The sessions are generated as TCP or UDP according to a specified probability. The packets from UDP sessions are generated deterministically, whereas the generation of packets from TCP sessions is governed by a protocol based on a TCP sliding window protocol, but has not been implemented in full detail. First a full data of window is sent (the window is set to a default value) and every other packet is sent after an arrival of ACK from the receiver. The packets are being retransmitted if an ACK has not been received within a fixed time interval. The packets from a session are of fixed length, which are selected from a given range.

The simulator gives results for the mean rate and the peak rate of the carried IP traffic on ATM link as a function of BER, as well as, for the distribution of IP packet lengths of the offered traffic, which is needed for the input of the analytical model. The results in Table 1 are obtained for input traffic of mean arrival rate equal to 40 sessions per second, out of which 90 % is TCP and 10 % UDP. The analytical model for the mean rate gives values within the range of (0.39 %, 7.25 %) from the results obtained from the simulator. The analytical model for peak rate gives results within the range of (1.43 %, 9.23 %) from the results obtained from the simulator. Discrepancies bigger than 0.55 % for the mean rate and bigger than 1.76 % for the peak rate were recorded for $BER > 10^{-6}$.

Table 1. Mean Rate and Peak Rate as a function of BER

BER	Mean (an) [Mbps]	Mean (sim) [Mbps]	an/sim %	Peak (an) [Mbps]	Peak (sim) [Mbps]	an/sim %
1E-10	1.264	1.259	0.39	1.835	1.809	1.43
1E-9	1.264	1.259	0.39	1.835	1.809	1.43
1E-8	1.264	1.259	0.39	1.835	1.809	1.43
1E-7	1.265	1.260	0.40	1.836	1.810	1.46
5E-7	1.269	1.263	0.47	1.842	1.812	1.52
1E-6	1.276	1.269	0.55	1.849	1.817	1.76
5E-6	1.310	1.287	1.78	1.900	1.846	2.89
1E-5	1.359	1.310	3.74	1.972	1.893	4.15
5E-5	1.907	1.830	4.21	2.684	2.252	6.50
1E-4	3.210	2.993	7.25	4.100	3.740	9.23

6 Conclusions

In this paper we have described a mathematical model that characterises the aggregate traffic of multiple TCP/IP connections as it enters the ATM backbone into a form that is suitable for the dimensioning processes of the ATM network. Specifically, we modelled the aggregate IP traffic that is being carried over ATM link as an “aggregate IP call” characterized by the three principal parameters of the Guérin et al model, viz [4]: mean rate, peak rate, and mean burst period. This enabled us to calculate the effective bandwidth of the “aggregate IP call” and to use it subsequently for the dimensioning of the ATM networks.

To validate the modelling, we have developed a simulation model. The simulation results have shown that our mathematical model is very accurate in demonstrating the effects of increasing cell loss probabilities (i.e., bit error rates) on the performance of the IP traffic over ATM networks. The success of the method shows that we can translate the IP traffic measurement data into their equivalent cell-level parameters for direct application into the ATM dimensioning procedures.

Future work will involve the development of tools to implement the procedures described in this paper for optimal design of ATM networks.

References

1. Berry, L.T.M., Harris, R.J., Puah, L.K.: Methods of Trunk Dimensioning in a Multiservice Network. In Proceedings of GLOBECOM'98. (1998) 282–287
2. Kaufman, J.S.: Blocking in a Shared Resource Environment. *IEEE Transactions on Communications*. **29** (1981) 1474–1481
3. Roberts, J.W.: A Service System with Heterogeneous User Requirements - Application to Multi-Service Telecommunications Systems. In Proceedings of Performance of Data Communication Systems and their Applications, G. Pujolle (ed.). (1981) 423–431
4. Guérin, R., Ahmadi, H., Naghshineh, M.: Equivalent Capacity and its Application to Bandwidth Allocation in High-Speed Networks. *IEEE Journal on Selected Areas In Communications*. **9** (1991) 968–981
5. Hassan, M., Breen, J.: Performance Issues for TCP/IP over ATM. 7th International Network Planning Symposium - Planning Networks and Services for the Information Age. (1996) 575–580
6. Lindberger, K.: Analytical Methods for the Traffical Problems with Statistical Multiplexing in ATM Networks. In Proceedings of the 13th International Teletraffic Congress. (1991) 807–813
7. Smith, W.L.: Renewal Theory and its Ramifications. *Journal of Royal Statistical Society B*. **20** (1958) 243–302
8. ATM Forum. ATM Traffic Management Specification Version 4.1. (1999)
9. Butto, M., Cavallero, E., Toniatti, A.: Effectiveness of the “Leaky Bucket” Policing Mechanism in ATM Networks. *Journal on Selected Areas in Communications*. **9** (1991) 335–342
10. Bonaventure, O.: PhD. Integration of ATM Under TCP/IP to Provide Services with Minimum Guaranteed Bandwidth. Université de Liège (1998)