# The Danger of Data Traffic Models

31 April 1995

Sam Manthorpe*

**Abstract**

Emerging ATM networks are being dimensioned to carry the data traffic seen on today's LANs and WANs. The modelling of such traffic is therefore not only of interest from the point of view of advancing networking science, but one could even go as far as describing it is a "hot topic". This paper gives a resume of work that has been done and that is underway (a large amount of it in the COST projects 224 and 242) into finding data traffic models and applying them to network dimensioning problems. It then highlights the danger of modelling this traffic at the ATM layer and then using an "open-loop" approach to tackling network dimensioning problems. Examples from network analysis, simulation and experimentation with ATM networks carrying TCP/IP traffic are given to demonstrate the possible gross inaccuracies that may occur.

## 1. Introduction

### 1.1. Background

Data communication is currently the *raison-d'être* of ATM; one only has to look at the ATM Forum as it frenetically beats ATM into a marketable condition. This is not surprising in light of the explosion of Internet traffic in the last three years [PAX94], resulting mainly from the bandwidth-hungry applications that constitute the World Wide Web. Suddenly the Internet *needs* broadband technology and quickly. ATM is there (almost) to satisfy this need. For some time now work has been underway to try and characterise this data traffic, in particular by Leland and his team from Bellcore labs [LEL91]. With data networks moving in to the broadband era, however, a new importance has been placed on the modelling of this traffic, for the purposes of network dimensioning.

Leland and Wilson have shown that data traffic is bursty over a wide range of timescales [LEL91]. This makes it very difficult to find simple models that accurately reflect the important characteristics of the traffic and at the same time can be applied to solving network dimensioning problems. Nevertheless, significant progress has been made in this area, not in the least by the COST 242 project. The self-similarity of the data traffic samples from an operational LAN at Bellcore was first noticed by Leland et al. and in [NOR93], Norros proposed a 3-parameter Gaussian model for the traffic, discussed the relevance of its parameters and gave an approximate solution for the occupancy distribution of a queue with self-similar input. His paper gives a deep yet clear insight into the nature of an ensemble of data traffic with self-similar property.

Recently Robert [ROB95b] and Andersen et al [AAL95] have proposed modelling the self-similar data traffic with Markovian traffic models, demonstrating that such models are capable of modelling self-similar behaviour over several timescales. Work in this direction continues.

In this paper we will discuss the modelling of data traffic and point out the danger of modelling samples of traffic and then applying these models for network dimensioning purposes. The traditional telecommunications approach to network dimensioning was to observe real network traffic, find a model that represented its important features (these "important features" depend on what the model will subsequently be used for), fit its parameters to measured traffic from the real world and, finally, apply the model to a network dimensioning problem. This is shown in figure 1 below. This method has been used by many authors [AAL92,MAN94,MUR94,YOK94] for the dimensioning of ATM data networks. Here we borrow a signal processing term and call this the "open-loop" approach to modelling, since there is no feedback[1].

* Laboratoire de Réseaux de Communication, EPFL, CH-1015 Lausanne, Switzerland.
email: manthorp@tcomhp20.epfl.ch, web: http://lrcsuns.epfl.ch

[1]Not to be confused with the terms open-loop and closed-loop for queueing networks.

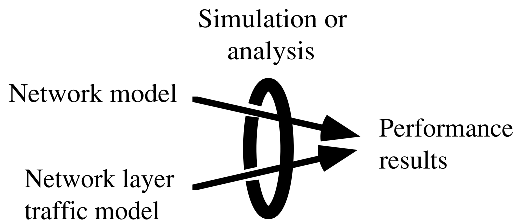



Figure 1 The traditional approach to network dimensioning

This approach, however, when used in a data communication scenario, may be grossly inaccurate unless the traffic control mechanisms at higher layers in the protocol stack are taken into account [MAN95]. The traffic measured at Bellcore *is* very bursty, since the measurements were made on a shared medium LAN with low average utilisation, thus allowing some long data bursts to do as they please in the network. However, one must be wary of applying models of such traffic to dimensioning problems where the underlying network is different from that of the original traffic sample. The reason for this is that the traffic originally observed *is heavily influenced by the topology and dimensions of the network in which it travels.*

As a simple example, let's consider the traffic that passes between two workstations, A and B, when A is sending data to B as shown in figure 2. Station A sends packets to station B, which sends corresponding acknowledgement messages back to A. Let's assume that the traffic between the two is regulated by means of a window-based flow control mechanism that effectively limits the number of unacknowledged packets in the network at any time. The diagram below shows the traffic on the network that will be observed for two different scenarios: (a) negligible propagation delay (LAN scenario), and (b) significant propagation delay (wide area scenario).

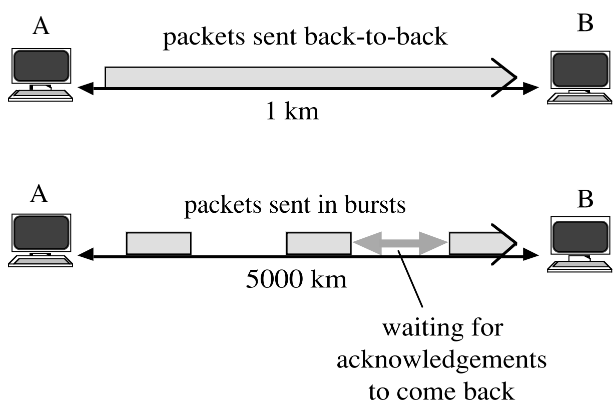



Figure 2 - Example of how the network influences the traffic patterns in it.

In the local area network, the low propagation delays mean that the acknowledgements from station B reach station A before station A has been able to transmit enough packets to fill its flow control window. This means that traffic will be transmitted by station A at a high rate, with packets being sent almost back-to-back. However, in the long-distance case, station A will transmit packets until its window is full, at which time it will wait for acknowledgements to come back from station B. Due to the long propagation delay, the window will fill before the acknowledgement for the first packet arrives at station A, resulting in a series of bursts being transmitted by station A.

This is one example of how the underlying network has a strong influence on the behaviour of the traffic in the network. In real networks many other phenomena may be observed. Packet loss in the network has important consequences on the behaviour of congestion control mechanisms, which will typically throttle the sending station when loss is detected. Variation in the network delay affects the traffic behaviour in systems where loss is detected through timouts, whose values are calculated from on-line round trip time measurements. Such a system is the Transport Control Protocol (TCP) [COM91], which is very widely used in today's data networks [CAC91] and whose performance is examined in section 2 and 3.

As another example, let us imagine that we have a "perfect traffic model", that models exactly the behaviour of a certain sample of data traffic, such as the Bellcore traces. We then use this model to dimension a buffer, based upon a certain cell loss ratio requirement. We may find that this buffer needs to be rather large in order to have low loss probability and high bitrate utilisation. However, a large buffer will introduce large delays which will change, rather significantly, the behaviour of the traffic. The original model is then no longer valid.

Figure 3 shows the influences between different layers of the protocol stack for TCP over ATM. Descriptions of the functionality and dynamics of TCP may be found in [COM91] and [JAC88].

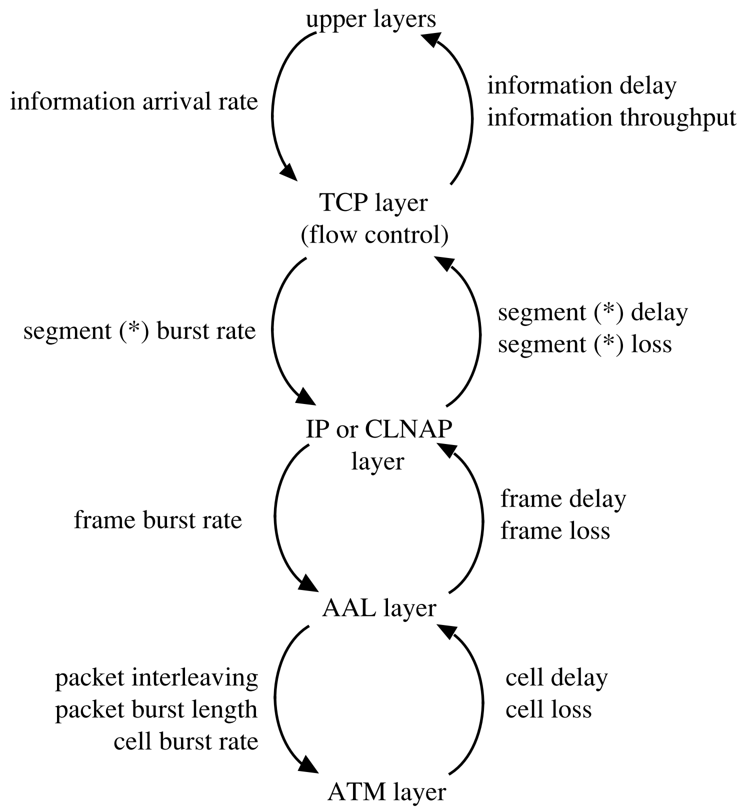



Figure 3 Interaction between different protocol layers

As can be seen from the above diagram, the mutual influence between different protocol layers is complex. Some authors have considered multilayered models and a good survey of this work may be found in [CON91]. Analysis techniques generally rely on the Newton substitution technique, whereby a whole network is replaced by a single queue that approximately represents its behaviour. Murata and Takagi have modelled the MAC and transport layers together in [MUR88], using a mean value analysis approach. Their model, however, does not take into account the strong influence of the transport layer congestion control mechanism on network behaviour. Fdida et al. [FDI90] also use mean value analysis to solve a network of semaphore queues that model end-to-end flow control, but again, there is no model for the congestion control algorithm. In both approaches, the only coupling between layers is through traffic flow and losses are not modelled.

In [NOR95] Norros introduced the concept of "free traffic", which is the traffic that would be seen if the network resources were unlimited. The aim of the free traffic concept is to try and isolate the inherent nature of the traffic from any influence that flow and congestion control mechanisms may have on the traffic. Certainly it seems intuitively correct that the self-similarity property observed in the Bellcore data is a fundamental property of this type of traffic and has little to do with the particular type of flow control mechanisms used at the transport layer or at any other layers. Norros asserts that the Bellcore traffic patterns, which were sampled in 1989, can effectively be considered as an example of free traffic, since the mean utilisation of the network was at that time low and network resources could be considered to be unlimited[2]. This seems reasonable and should be true for any shared-medium LAN that has a low operation utilisation (e.g. <25%).

In ATM networks, however, high bitrate utilisation is demanded and the network resources are anything but unlimited. Indeed, for the proposed Available Bitrate (ABR) service, the available bitrate and delay experienced

---

[2]Care should be taken here, since while this may be true for the local traffic samples, it is not so for the external traffic samples, where more significant propogation delays may be seen.

may undergo large fluctuations in very short periods of time. Data traffic in such a scenario will most likely look quite different. Whether it maintains its self-similar property remains an open question. If one uses the argument that the self-similar nature of data traffic is a direct result of human influence, an argument that sounds quite nice, then it would seem likely that the self-similarity will remain. This is an area for further study.

### 1.2. Performance Parameters

Network dimensioning in the telecommunications world is usually anchored to certain Quality of Service (QOS) parameters, typically for packet loss and packet delay. Of course, what the end user is really interested in is the quality of service that he/she perceives and not the ratio of packets dropped in the network. Luckily, there is usually a fairly simple relationship between the two; a lost packet for a telephony service results in a clicking noise, too many of which will irritate the user. For data traffic, there is also a relationship between data link and application layer performance, but it is, unfortunately, not as simple as is often assumed.

The main difference, again, arises from the presence of congestion control algorithms, located at the intermediate layers. Congestion control algorithms that adjust transmission rates according to network conditions and implement retransmission for lost packets effectively translate loss in network layer and below into delay observed at the application layer and above.

Let us take the following example. We wish to dimension a shaping buffer so that the probability of packet loss in the buffer is $\sigma$. We make a model of typical data traffic, for example the Bellcore traffic, and do some queueing analysis to find out what service rate we need for the buffer, which turns out to be $\mu$. However, when we run real traffic through the shaper with service rate $\mu$, we notice that the loss ratio is very different to $\sigma$. The reason for this is that some congestion control algorithm (e.g. TCP) is adjusting the transmission rate to find its optimal operational loss ratio. This will be demonstrated through experimentation later in the paper. Note that this optimal loss ratio is not necessarily zero, since performance may be improved by transmitting packets even when there is a risk that they may be lost[3]. The above phenomena is illustrated in sections 2 and 3.

A more useful parameter for network dimensioning would be the throughput as observed from above the congestion control algorithms. Such an approach is usually adopted for TCP simulations over various network technologies. The disadvantage is that results become specific to the implementation of the flow control mechanisms used. If one is dimensioning the data link layer, it is rather in-elegant to have to rely on information on the upper layers in order to be able to do it. Moreover, in a situation where QOS guarantees are made for the data link layer (such as a network interconnection service, for example), it is not possible to make any assumptions on or requirements of the upper layers.

### 1.3. Organisation of paper

Examples of the inaccuracies that may result from the open-loop approach to dimensioning are given in section 2, which contains already published dimensioning results for a MAN to WAN shaping buffer [MAN95]. It is shown that if the Bellcore traffic samples are force-fed into the buffer[4], an extremely large buffer capacity is required which would significantly alter the delay observed in the network and hence the behaviour of the traffic. In section 3 results from experiments with TCP over ATM are presented that show the inaccuracies of the open-loop modelling approach at the ATM layer alone. In the open-loop approach, bandwidth or buffer capacity would be allocated so as to ensure a certain loss probability at a buffer. The experimental results show that it is the actually the TCP congestion control algorithm that *really* controls the loss rate in the buffers. Section 4 concludes with a discussion of the issues raised.

---

[3] This point is relevant to the ABR service currently being specified by the ATM Forum, since they place a lot of importance on services that offer "zero loss".

[4] One can consider the traffic sample as a kind of perfect model within a limited timeframe, since it reflects the measured traffic 100% accurately.

## 2. Dimensioning approaches

This section describes some of the approaches to buffer dimensioning for data traffic that have been used, with a comparison of their accuracies. Results are taken from [MAN94] and [MAN95], where the goal was to dimension the shaping buffer in a DQDB to ATM interworking unit. Four approaches are considered:

(i) Poissonian packet arrival model, using the packet loss probability as the dimensioning criteria. If we assume instantaneous arrivals, which is the worst-case for this model, then the system can be modelled as an M/M/1 [AAL92,NOR93].

(ii) Simulation of shaper using Bellcore's sampled traffic, using the packet loss probability as the dimensioning criteria [MAN94]. The method is analogous to the "perfect model" described in the introduction, since it completely models the sampled traffic. More details may be found in appendix B.

(iii) Simulation of TCP through shaper, using the packet loss probability as the dimensioning criteria. See [MAN94] for more details.

(iv) Simulation of TCP through shaper, using the TCP goodput[5] as the dimensioning criteria. TCP goodput[6] is much more closely related to the quality of service as perceived by the human user than is the cell loss ratio and was adopted as the primary parameter for judging the performance of the IWU in this work. Descriptions of the functionality and dynamics of TCP may be found in [COM91] and [JAC88]. Full details of the simulations may be found in [MAN95] and in appendix A.

Graph 1 below shows the results obtained using the different dimensioning approaches. The vertical axis shows the utilisation of the shaper, which may either be the utilisation at the ATM layer, or the normalised TCP goodput, which we define to be the TCP goodput divided by the shaper rate at the ATM layer[7]. The curly brackets in the legend box indicate the layer at which the utilisation was measured. For those approaches where the utilisation is measured at the ATM layer, the dimensioning criteria is that the probability of packet loss in the buffer be $10^{-4}$. For the TCP case utilisation axis shows the maximum TCP goodput that was possible for that shaper size.

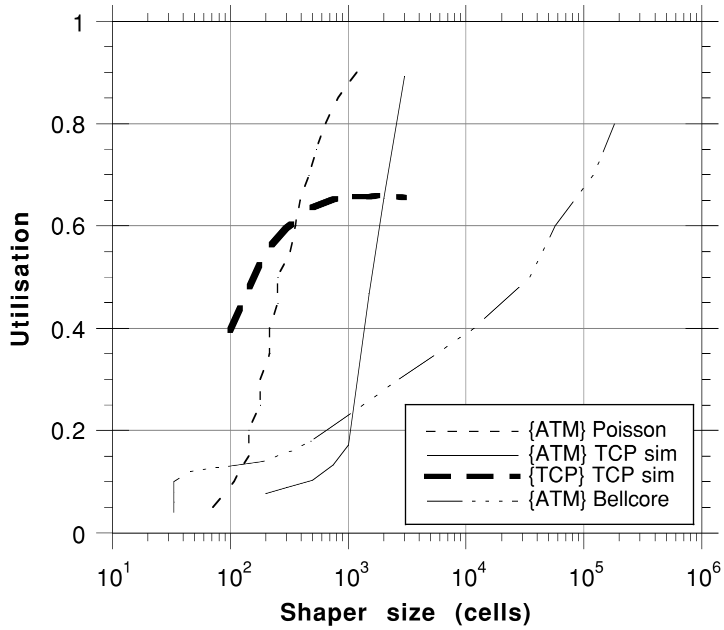
The most striking feature of this graph is the wide diversity of its curves. The Poissonian traffic model, with dimensioning performed based upon the packet loss probability, shows that a shaper size of around 1000 cells is required. We notice that for the same shaper size, good TCP performance is achieved. If we were to use the packet loss probability as the dimensioning criteria and were to use a TCP simulation to analyse the network, we would need a larger buffer. However, the most striking curve is that for the simulations with the Bellcore traffic samples which show that the buffers required to have a packet loss probability of 10-4 and good utilisation are <u>enormous</u>. The traffic samples represent the above mention "perfect model", yet such large buffers would introduce very large delays in the network which would, in turn, change the behaviour of the traffic that we originally wanted to model.

---

[5] We define the goodput to be the amount of bits per unit time communicated between the transport layers at the sending and receiving stations. The goodput is the amount of useful information transfered and therefore excludes padding at the ATM layer, retransmissions and protocol header/tail overhead. However, our simulator employs a packet based approach at the transport layer, rather than the octet based approach used in the real algorithm and since the packet size is constant, padding is introduced in the packets, which is included in the goodput measurements. It does include, however, the padding introduced at the TCP layer due to the fixed size packets.

[6] Rate at which useful information successfully reaches the destination station (does not therefore include lost packets or incorrectly re-transmitted packets) in our case, averaged over the duration of a file transfer.

[7] It is important to specify the layer when talking about bitrates, so as to take into account the overheads introduced at each layer.



Graph 1 Comparison of dimensioning approaches for the shaping buffer

## 3. Experimentation

The COMBINE project has also tackled the problem of buffer dimensioning and bandwidth allocation for the design and configuration of their interworking units. A model for the IWU, which is designed to perform routeing between a DQDB network and an ATM network, is shown in figure 4. Input from the DQDB network is on the left-hand side and output to the ATM network is on the right-hand side and must pass through the Usage Parameter Control (UPC) [I371]. The shaping buffer is required in order to achieve good utilisation of the Virtual Path Connections (VPCs) on the ATM network in the presence of the UPC.

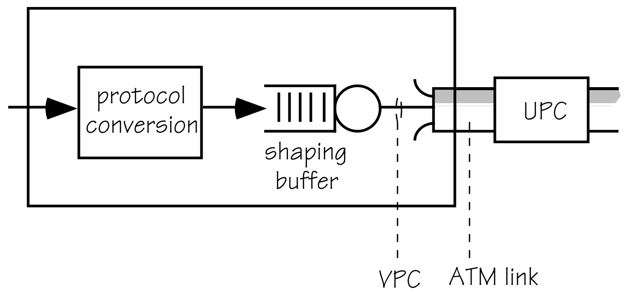



Figure 4 Interworking Unit (IWU)

To dimension the size of the cell-to-slot mode shaper, Alto et al. adopted a simple Poissonian packet arrival model and analysed it as a fluid buffer [AAL92,NOR93]. The argument for using a Poissonian model was that burstiness at timescales higher than that of a packet arrival are a problem to be tacked by flow control at higher layers. The traffic was therefore bursty at the cell layer, since packets arrived in bursts of cells, but above the packet layer was memoryless. This dimensioning resulted in the cell-to-slot interworking unit, to be equipped with a shaper of size 2000 cells.

Experiments were performed to measure the TCP performance through this IWU and so assess the end-to-end sensitivity to cell loss in the shaping buffer. The experiment described below aimed to assess the performance of the COMBINE IWU for TCP traffic. The basic experiment scenario is shown in figure 5, where two Sparc architecture workstations, Tiger and Mulle, communicate via a heterogeneous network consisting of Ethernet and ATM. The hardware used is as follows:

- Two Sparc 10 architecture workstations, connected to two Ethernets.
- Alcatel Ethernet_LIM. This is an interworking unit that acts as a bridge between Ethernet and ATM. It had a maximum throughput of 5 Mbits/s.
- Fore ASX200 switch.
- HP Broadband test equipment
- COMBINE cell-to-slot IWU

The ATM Pilot part of the network was included to introduce a more realistic propagation delay. A multiple loopback configuration with the Jutland Telephone Company (JTAS) was set up, with the traffic being sent to Jutland and back 7 times. This introduced a delay of approximately 16ms. The influence of traffic correlation in the MOU switches is ignored, since each VPC was allocated 5Mbits/s, so the traffic is too feeble to introduce any serious correlation effects.

Note that the traffic only goes through the ATM Pilot network in one direction. This was because in the LIMs used, no output shaping was present and Ethernet frames were reassembled and then transmitted in a burst at the link rate. It was therefore necessary to pass the traffic through the IWU before the pilot network so that it could be shaped to avoid Usage Parameter Control (UPC) cell loss. With the added requirement of not mixing the data traffic and its acknowledgements in the same shaping buffer, only one of the LIM to LIM traffic streams could be passed through the IWU and therefore only one could be passed through the pilot network. The importance of shaping, at least to avoid such network-unfriendly, high-bitrate bursts from message mode equipment was highlighted in the COMBINE project and was again evident in this practical case.

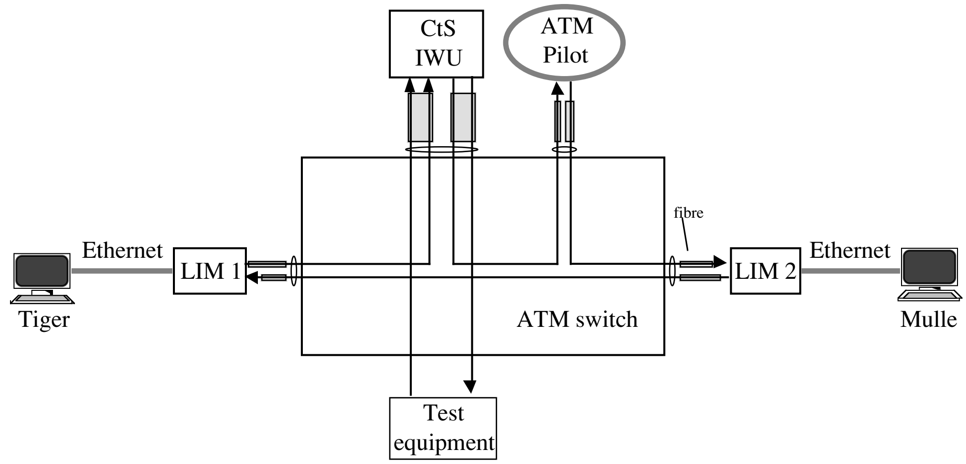



Figure 5  General experiment configuration

3.1. Measuring TCP performance

The File Transfer Protocol (FTP) application was used to test the TCP protocol's performance. After a file transfer, FTP gives some rudimentary statistics on the performance of the transfer, specifically the mean goodput and the time taken to transmit the file. The precision of this information is quite low (due to the low precision of internal timers), but was sufficient to be able to draw meaningful conclusions. A test file was set up on the local disk of workstation Tiger, from where it was then transferred to the local disk of Mulle and the time taken for the transfer and mean transfer rate measured. Cell loss in the network was measured using the traffic flow statistics of the FORE switch, which gives the number of cells going into and the number of cells coming out of the IWU on a per VCC basis.

3.2. Problems with controlling cell loss

One of the aims of this experiment was to show the influence of cell loss on the TCP performance. However, there was one main problem with this objective, in that the only way of introducing loss was by overflowing the shaping buffer, using background traffic. This made it difficult to isolate the effects of cell loss and those of delay, since a shaper overflow will always correspond to significant delays in the shaper (approximately 170ms for a shaper rate of 5Mbits/s). At the time the experiments were performed, no other means were available for introducing loss.

The problem was tackled by running two experiments; one control and one test. The control experiment measured the delay and loss in the shaper while varying the shaper rate, with no background load. The results show the influence of the delay on the TCP goodput. Note that in this case, the TCP congestion control algorithm has almost complete control over the queueing behaviour in the shaper and is able to ensure a low loss rate for a wide range of shaper service rates. The second experiment had a constant shaper rate of 5Mbits/s, but the background rate was adjusted to introduce loss in the shaper. It was possible to introduce loss in the shaper using the background traffic, since the TCP congestion control algorithm has less influence on the behaviour of the queue and has less power to control the loss rate. Here the TCP goodput is influenced both by the delay and the loss in the shaper. A comparison of the results from the two experiments allows one to more clearly see the influence of the cell loss on the performance observed by the user.

3.3. External factors

Although the network itself is empty (apart from a small amount of NFS traffic), the TCP behaviour was also influenced by other processes running in the station. It is not expected that this has a significant influence on the results, especially bearing in mind that each experiment was performed at least 5 times, often up to 18 times.

3.4. Results

*3.4.1. Varying the shaper rate*

In order to observe the performance of TCP over the network, file transfers were performed for different shaper rates, using the configuration shown in figure 6. As mentioned in the introduction, one of the problems of the configuration used in the experiment, where loss is introduce by overflowing the IWU shaper is that significant delays automatically accompany cell loss. In order to attempt to assess the performance of the IWU, the TCP goodput was measured without background traffic, but with a range of shaper rates. In a system where there is just one TCP communication through the network, the congestion control algorithm will have a very strong influence on the behaviour of the network and should thus easily be able to keep the cell loss rate at a low level.[8]. The delay in the network, however, will strongly depend on the shaper rate and we can thus observe the influence of the shaper delay on the TCP performance.

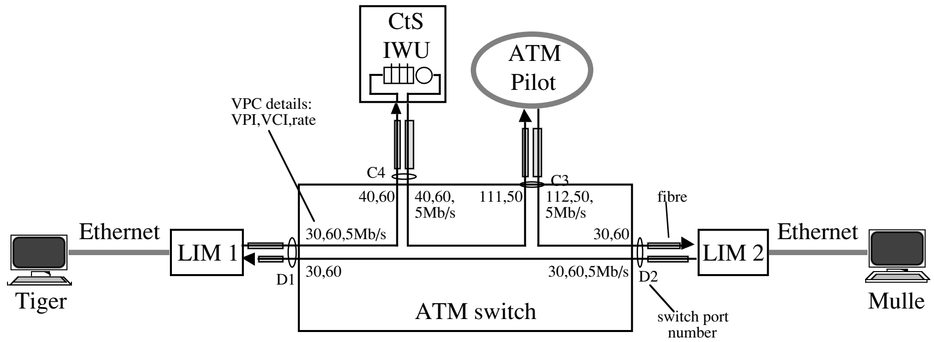
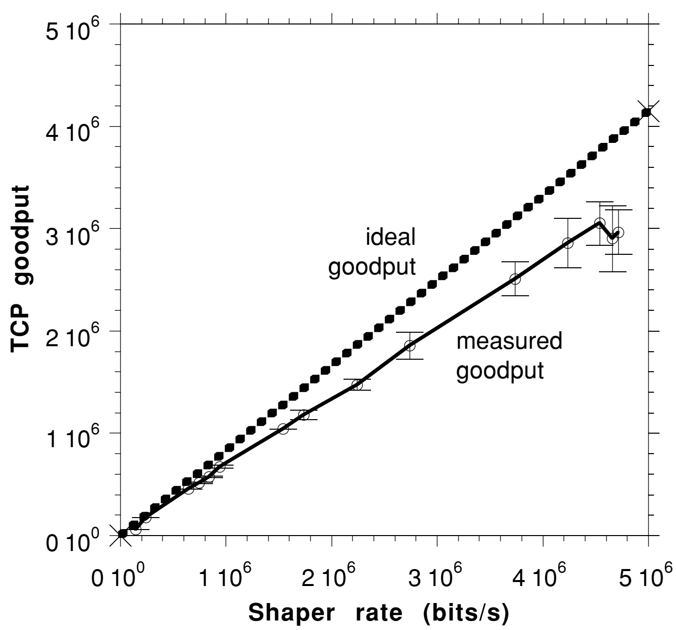



Figure 6 Configuration for varying shaper rates

For each set of parameters, the experiment was performed several times in order to be able to construct confidence intervals. Graph 2 shows the results for the TCP goodput as a function of the IWU shaper bandwidth, together with bars showing the 95% confidence intervals. The "ideal goodput" line shows the maximum goodput possible, taking into account the protocol overheads from TCP and all the layers below it.




Graph 2 TCP goodput measured for varying shaper rates, with 95% confidence intervals.

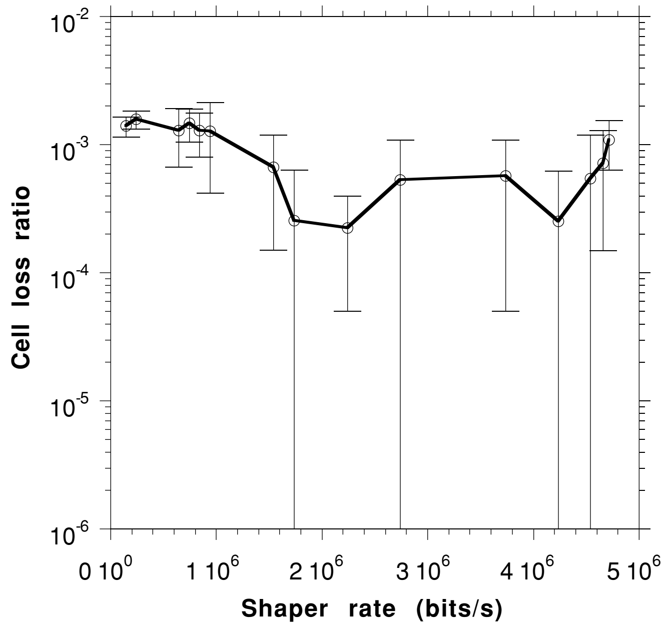
As can be seen, TCP adapts well to the bottleneck bandwidth introduced by the shaper, achieving a goodput close to the ideal for all shaper rates. This adds weight to the argument that ATM layer dimensioning should be able to cope with burstiness up until the packet layer timescale, but that longer term effects are to be dealt with by transport layer traffic control mechanisms. For shaper rates greater than 4.5 Mbits/s, the goodput appears to be levelling out. This is due to the maximum throughput rate of the LIMs, which is specified as being 5Mbits/s.

Graph 3 shows the cell loss ratios through the IWU, measured for different shaper rates, also with 95% confidence interval bars. Not only does the CLR vary in a way seemingly independent of the shaper rate, but some points have

---

[8]The phenomena of the bottleneck link speed being indicated by the rate of acknowledgement arrivals at the source is described by Jacobsen in [JAC88]. In this scenario, the TCP on Tiger will send packets at a rate more or less equal to the shaper rate, whatever it may be, thus ensuring low cell loss.

very large confidence intervals. It is not sure that the cell loss observed here results from shaper overflow, since if this were the case there should be some simple, observable relationship between the cell loss ratio and the shaper rate. The cell loss behaviour in this scenario is probably due to some other factor (behaviour of the operating system, timeout inaccuracies) and is not considered relevant for the results. What is important here is that the CLR does not seem to depend on the shaper rate, but is controlled by the TCP congestion control algorithm. This differs from the open-loop school of thinking that assumes a strong relationship between service rate and loss rate, ignoring any higher layer influence.




Graph 3 Cell loss ratio for different shaper sizes, with 95% confidence intervals

*3.4.2. With background load*

In order to be able to better control the cell loss in the shaper, it was then loaded with background traffic. In this scenario, the TCP congestion algorithm no longer has complete control over the network behaviour, since it cannot influence the arrival pattern of the background traffic. Its ability to control the cell loss is thus reduced and we were able to better study its behaviour when a significant proportion of its traffic is lost in the shaper. The experimental configuration used is shown in figure 7

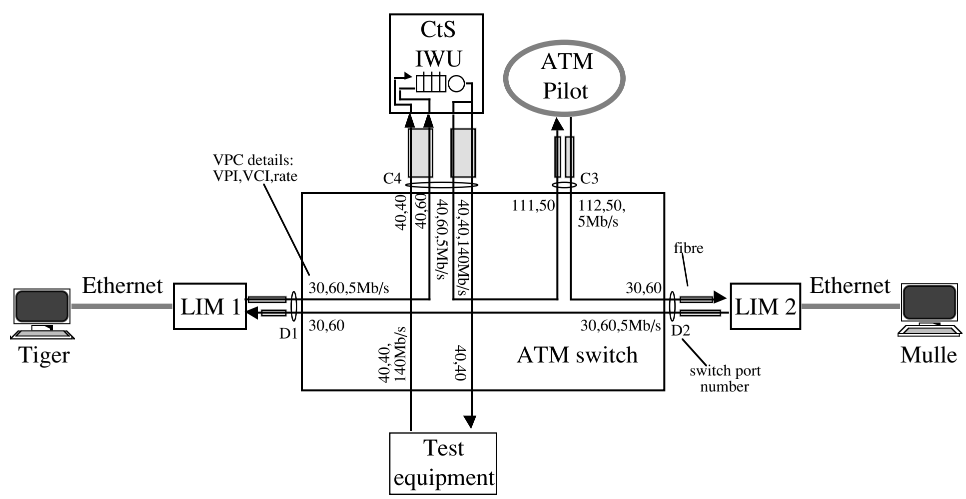
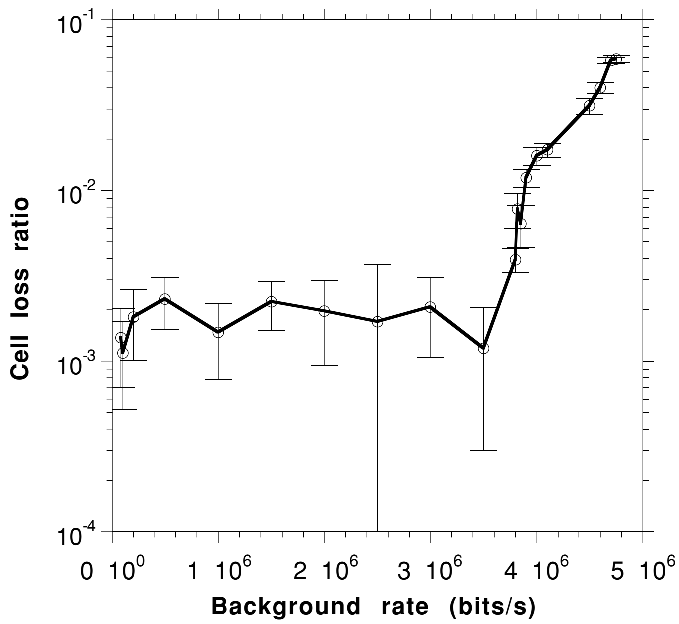


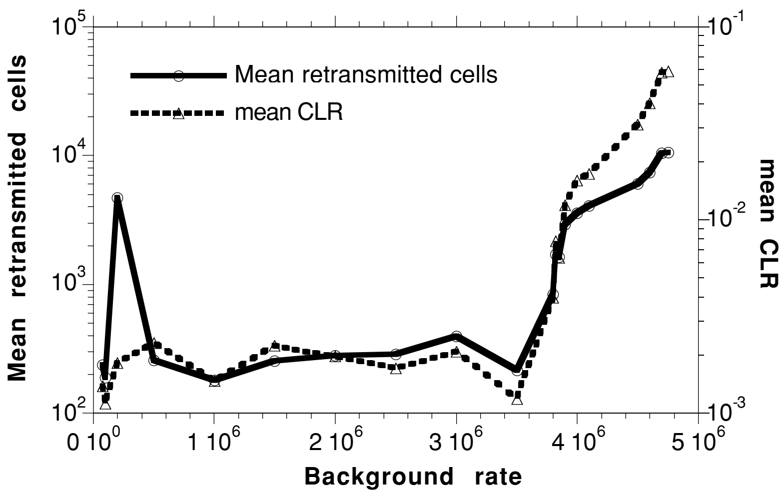

Figure 7 Experimental configuration with background traffic

The VPC leaving the shaper was loaded with a certain amount of background traffic, the rest of the bandwidth being available for the TCP traffic. We will call this traffic the available bitrate. It is calculated by taking the mean background rate and using it, together with the cell loss rate, to estimate the mean rate of background traffic that will actually be queued in the shaper and not lost at its entrance. This bitrate is then subtracted from the VPC rate to get the estimated available bitrate (see below). Graph 4 shows the cell loss experienced by the TCP traffic. The curve can be divided into two parts: for background loads less than 3.5 Mbits/s, the TCP congestion control algorithm is able to maintain the cell loss ratio in the order of $10^{-3}$. For background loads greater than 3.5Mbits/s, the cell loss ratio starts to increase significantly, up to two orders of magnitude greater than for the low background load case. For loads of up to 80%, the background load had little influence on the loss rate, due to TCP's congestion avoiding efforts. This is an important result since it shows that in most cases it is the TCP algorithm that controls the loss rate, rather than availability of buffer space and bandwidth.




Graph 4 Cell loss ratios for varying background rate, with 95% confidence intervals.

Graph 5 shows the relationship between the number of retransmissions (in units of cells) and the cell loss ratio for different background rates. The close correlation between the two curves shows that the retransmission policy of TCP is accurate in this case and there are not too many unnecessary retransmissions. The exception to this is for a background load of around 0.25 Mbits/s, which remains unexplained. Notice that at some points, the cell loss ratio is higher than the retransmission ratio; this is possible since the packet loss ratio will always be less than or equal to the cell loss ratio, multiple cell losses occurring in the same packet resulting in only one packet being lost.

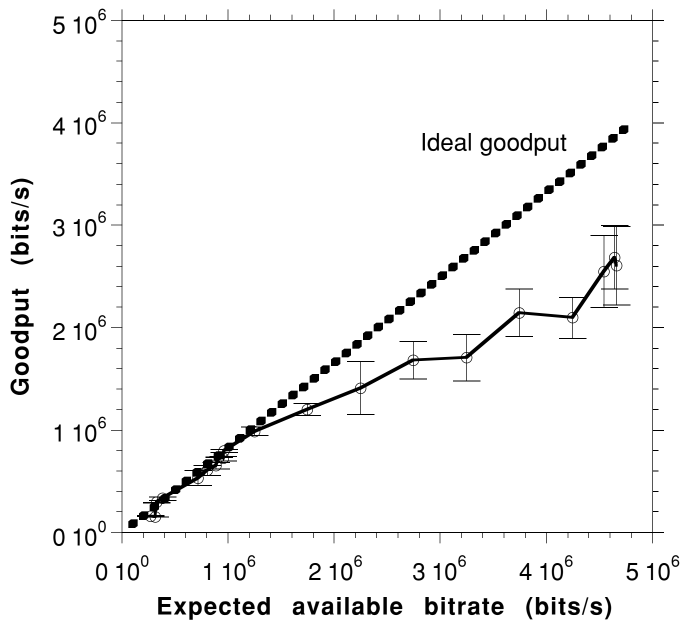



Graph 5 Retransmission overhead and cell loss ratio for background load experiment.

In order to compensate for the lost background traffic, the goodput was re-plotted against the "expected available bitrate", *a*, which is:

$$a = s - (1-\gamma)b \tag{1}$$

Where s is the shaper rate, $\gamma$ is the loss ratio and b is the background traffic rate. The relationship between the expected available bitrate and the TCP goodput is shown in graph 6.

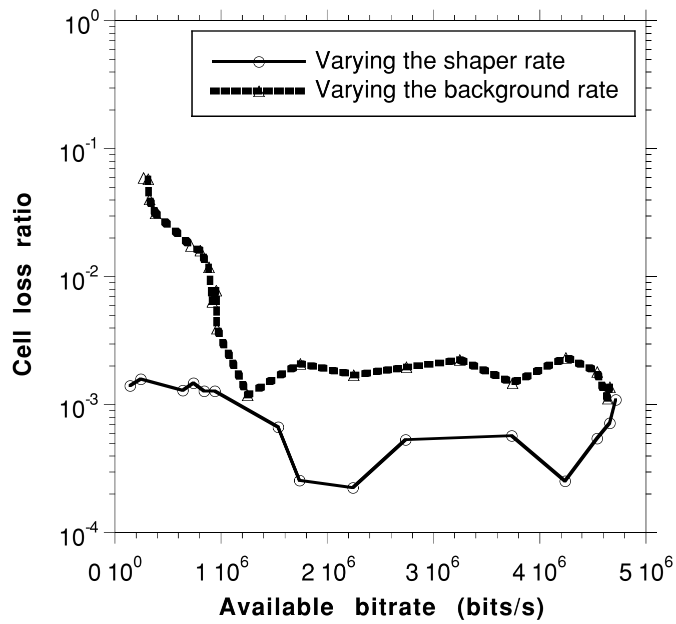



Graph 6 TCP goodput and ideal goodput for different available bitrates, taking into account the reduction of background bandwidth resulting from shaper cell loss.

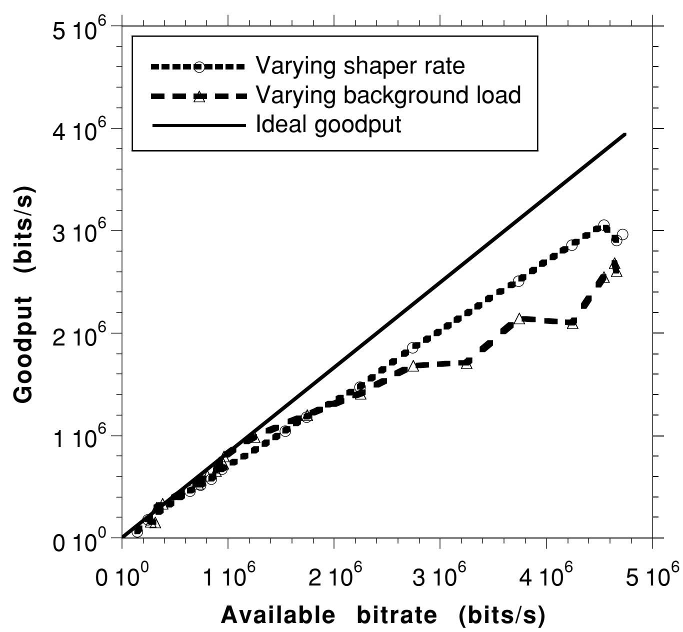
Here we see that although the goodput is impaired by loss and delay in the shaper, even under high load conditions, the throughput is quite acceptable.

### *3.4.3. Influence of cell loss on TCP goodput*

Results have shown that reducing the available bitrate by reducing the shaper rate has no noticeable influence on the cell loss ratio. Reducing the available bitrate by increasing the background rate, on the other hand, does introduce loss, since the TCP congestion control algorithm no longer has complete control over the behaviour of the queue. Comparing these two sets of results allows us to observe more objectively the influence of cell loss on the TCP goodput, since the influence of the shaper delay on the goodput will be the same in the two sets of results. Graph 7 shows the effect of varying the shaper rate and the background rate on the cell loss ratio for the two experiments.




Graph 7 Effect of varying the shaper rate and background rate on the cell loss ratio




Graph 8 Comparison of throughputs by varying the shaper rate and the background rate

Graph 8 compares the relationships between the goodput and the available bandwidth for variations in the shaper capacity and in the background load. The goodput is less for the background traffic experiments than for the shaper rate experiments, noticeably for high available bitrates (i.e. low background loads). However, we see that even in the presence of an 8% cell loss ratio in the background load case, the difference between that curve and the curve obtained from adjusting the shaper rate is not that large and the goodput achieved is still 60% of the ideal goodput. Again, we have an example of how reality diifers from the closed-loop school of thinking's assumption that low loss is required in order to get good performance at the higher layers.

## 4. Conclusions

This paper has highlighted the dangers of modelling data traffic and subsequently using this model for network dimensioning at the data link layer alone. I have described the "open-loop" approach to modelling, where a traffic model is developed and then applied to a network model, using either simulation or analysis, to obtain some performance results for the network under study. No matter how accurate the traffic model, the open-loop approach to performance analysis is fundamentally inaccurate for data traffic scenarios, due to the presence of flow and congestion control mechanisms at different layers in the protocol stack.

Examples of open-loop and closed-loop dimensioning approaches have been given and significant difference in results shown. Through experimentation, several open-loop school of thinking assumptions have been shown to be inaccurate for data traffic. More specifically, it was shown that it is the TCP congestion control algorithm that had most control over the loss rate in the shaping buffer of the COMBINE IWU and not the service rate of the shaper. It was also shown that the assumption that transport layer performance is very sensitive to loss at the data link layer is not entirely true.

Norros has introduced the concept of "free-traffic", which describes the traffic that would be observed if the underlying network had infinite capacity. This is a useful notion, since free-traffic models could be used together with so-called "closed-loop" network models that model the complex traffic flow in the network, to provide a more accurate performance analysis approach. A possible scheme is illustrated in figure 8, which shows the 7 layers of the protocol stack and suggests how each layer could be modelled. On the left hand side of the diagram, suggestions are given for the phenomena that have the largest influence of the traffic behaviour at each layer of the stack.

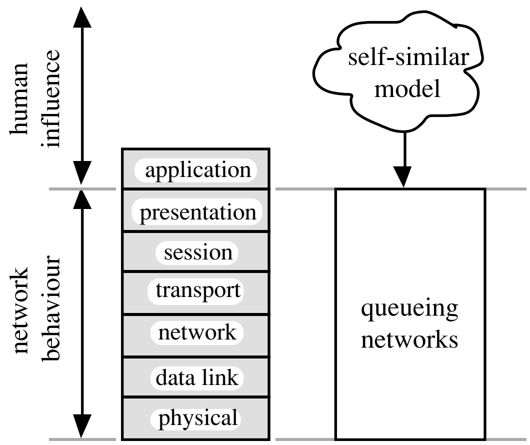



Figure 8 The primary influence on the behaviour of different protocol layers and the suggested approach to modelling them.

Here it is proposed that the self-similar nature of the data traffic is mainly a results of the human influence on the traffic. Whether the relationship between the lower six layers of the stack results in a self-similar behaviour of the traffic is not known and is an area for further study.

**5. Acknowledgements**

Many thanks to all members of the COMBINE project, especially Jorma Virtamo, Ilkka Norros, Bo Larsen and Max Perl who have helped very much with performing the experiments and interpreting the results. This work has been supported by the Swiss Office Fédéral de l'Education et de la Science and by the Swiss PTT Telecom.

## 6. Appendices

A. Simulation model

*A.1. Network model*

To study the behaviour of TCP/IP traffic in a broadband heterogeneous network, a simulation model consisting of a DQDB MAN and ATM was adopted. The model network is shown in figure A.1. We consider the case where data traffic is generated on a DQDB MAN and routed to an Interworking Unit (IWU), which transmits the traffic on the ATM network. The DQDB network is modelled as a process sharing queueing system [AAL92], which models the segment interleaving resulting from the MAC layer. Once at the IWU, the traffic is routed, via the CLON, to the destination network. The IWU works in cell-to-slot mode; that is, it does not reassemble frames, but converts segments directly into ATM cells. Before being sent on the ATM network, the traffic stream is shaped, and the size of the shaping buffer is one of the key parameters in this study. The CLS(s) and the destination network were not modelled in detail, but were represented by a constant delay. This was justified since the delay in the CLS is assumed constant and since the buffers in ATM switches are generally small, the deviation from the mean delay will be small. The simplicity of the model topology allows us to isolate the influence of the shaping buffer on the TCP traffic, as well as reducing simulation times. Exact values for the model values that are not variable in these studies are presented in section A.4.

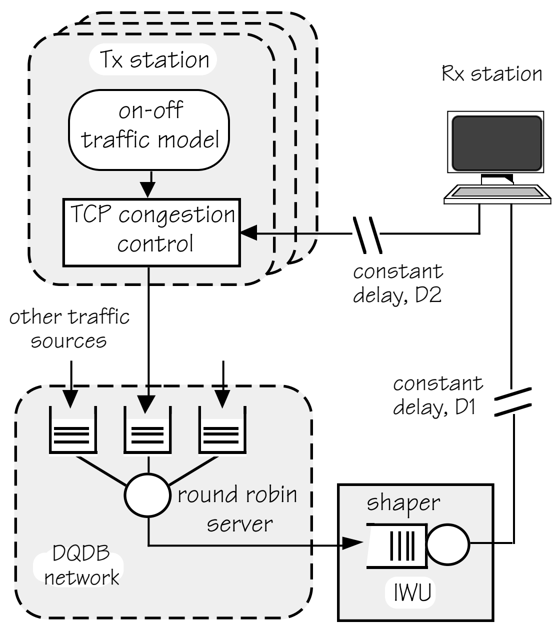
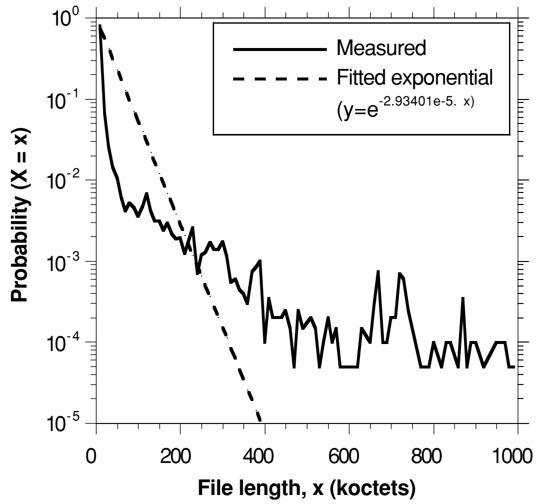



Figure A.1 Model network.

*A.2. Traffic sources*

In these studies we adopt an on-off source to model file transfer applications such as FTP, which generate long bursts of data, and as such are representative of the traffic that is particularly unfriendly to the ATM network. To find the parameters of the on-off model used, a statistical study was made of the filelengths of all the files on an EPFL file-server of 7 Goctets in size. Graph A.1 shows the probability distribution of the file lengths. We approximated this distribution with a geometrical distribution, also shown in the graph. The silence periods between file transmissions were also assumed to be geometrically distributed, with the mean silence length being chosen to give the mean traffic rate desired.

The TCP congestion and flow control algorithm was simulated at the transport layer. The implementation adopted was that described by Van Jacobsen in [JAC88] with a window based mechanism, the window size being controlled by the slow-start and congestion avoidance algorithms together. We assumed that all the receiving stations are sufficiently powerful that they never have to slow the sources down using window advertising. For full details, the reader is referred to [COM91] and [JAC88].




Graph A.1 Distribution of file lengths on an EPFL server

Note that the parameters of the traffic model specify the traffic offered to the TCP layer. However, the throughput of the TCP layer will usually be less than the rate at which traffic is generated during the transmission of a file and so the time taken for the file to be processed by the transport layer will be greater than the time taken to send the file to the transport layer. The file transmission is not considered to be finished until the last acknowledgement for the very last packet has been received and only at this point is the exponentially distributed time until the start of the next file calculated. This elongation of file transmission time means that the rate at which a source generates traffic is less than or equal to the rate specified by its parameters, the difference being a direct result of the TCPs throttling of the traffic throughput. This is illustrated in figure A.2 below. The upper bar illustrates the time that it would take to transmit the file if there was no transport layer traffic control and the lower bar illustrates the time that it actually takes for the file to successfully arrive at the destination.

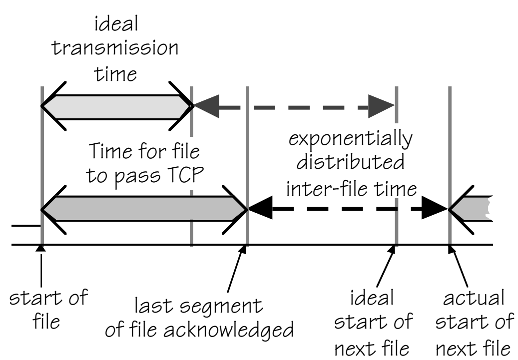



Figure A.2 Elongation of file transmission time due to TCP layer in the simulation model

In this paper we adopt the goodput at the TCP layer as the relevant performance parameter from the point of view of user-perceived quality of service. We define the goodput to be the amount of bits per unit time communicated between the transport layers at the sending and receiving stations. The aforementioned file elongation results in a reduction of TCP goodput and is influenced by delay and loss at the ATM layer.

*A.3. Simulator*

The simulations were run using software developed in the RACE 2032 COMBINE project, based upon the ATLAS-II library of C functions [RUM91] for network simulation. Simulation model times varied between 5000 and 50000 seconds, depending on the time required to achieve sufficiently small confidence intervals for the packet loss probability in the shaper.

### *A.4. Simulation parameters*

| **Parameter** | **Value** |
|---|---|
| DQDB bus bitrate | 100 Mbits/s |
| DQDB MAC access queues | 1000 packets (effectively infinite) |
| Peak rate during TCP connection | 10 Mbits/s |
| Maximum TCP window size | 75 koctets |
| TCP segment sizes | 1500 octets |
| Maximum TCP timeout value | 10 seconds |
| $\beta$ used for TCP | 2 (unitless) |
| Exponential timeout backoff rate, $\gamma$ | 2 (unitless) |
| Mean file length | 34083 octets |

## B. Simulation of Bellcore traffic

Simulations were performed using the Bellcore traffic samples. The traffic represents the external traffic that leaves the Ethernet and as such is a good example of the traffic that would enter the SIWU considered in this paper. The data has the following statistics:

| | |
|---|---|
| Smallest inter arrival time | 16 ms |
| Mean inter packet time | 122.7 ms |
| Mean packet length | 3.64 ATM cells |
| Mean bitrate | 9.137 kbits/s |

On a 10 Mbit/s Ethernet, this traffic represents a load of approximately 1%. When examining the behaviour of the shaping buffer in the IWU, we do not model the DQDB network itself, since we ignore the MAN background traffic and the Ethernet MAC delay is already incorporated in the sampled data. The simulation model is shown below in figure B.1.

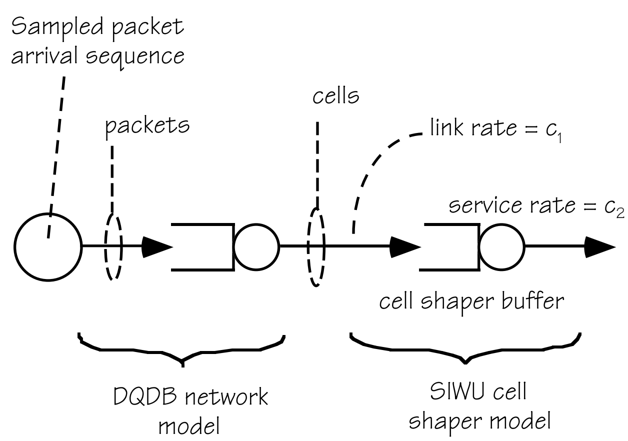



Figure B.1 Queueing model used for simulation of sampled traffic

Of the packets in this sample, 99.5% are IP packets. The TCP/IP protocol can generate very bursty traffic; the lengths of the burst coming from one source depend on: the rate at which data can be created by the application, the rate at which data can be processed by the receiving application and the size used for the flow control window.

## References

[AAL92] S.Aalto, I.Norros, J.Virtamo, K.Kvols, S.Manthorpe, "Performance aspects of streaming and message modes of interworking", Proceedings of the Interworking 92 conference, Bern, November 18-20 1992.

[ALL95] A.T.Andersen, A.Jensen, B.F.Nielsen, "Modelling of apparently self-similar packet arrival processes with Markovian Arrival Processes (MAP)", COST242 technical document TD(95)05, Cambridge meeting, January 17, 1995.

[CAC91] R.Cáceres, "Multiplexing Traffic at the Entrance to Wide-Area Networks", Ph.D. Dissertation, Report No. UCB/CSD 92/717, UCLA, December 92.

[COM91] D.E.Comer, "Internetworking with TCP/IP - Volume 1; Principles, Protocols and Architecture", second edition, Prentice-Hall International 1991, ISBN 0-13-474321-0.

[CON91] A.E.Conway, "A Perspective on the Analytical Performance Evaluation of Multilayered Communication Protocol Architectures", Journal on Selected Areas in Communications, Vol.9, No. 1, January 1991, pp4-14.

[I371| ITU-T Recommendation I.371.

[JAC88] V. Jacobsen, M.Karels, "Congestion Avoidance and Control", Proceedings of SIGCOM '88, Stanford, Aug 1988, ACM.

[LEL91] W.Leland, D.Wilson, "High Time-Resolution Measurement and Analysis of LAN Traffic: Implications for LAN Interconnection", Proceedings of Infocom 91.

[MAN94] S.Manthorpe, X. Garcia, "Shaper Dimensioning for MAN to ATM based WAN", Proceedings of the Interop 94 Engineers' Conference, Las Vegas, May 1994.

[MAN95] S.Manthorpe, X.Garcia, "TCP Performance Over ATM Based LAN Interconnection Services", Proceedings of the Interop 95 Engineers' Conference, Las Vegas, March 1995.

[MUR94] M.Murata, H.Miyahara, "LAN Internetworking through Broadband ISDN", IEICE Trans. Commun. Vol. E77-B, No.3, March 1994, pp294-305.

[NOR93] I.Norros (editor), K.Kvols (workpackage leader), "IWU Design Choices Based on Performance Analysis", Deliverable 4, R2032 COMBINE project, 30th June 1993, R2032/KTA/WP1/DS/P/004/b1.

[NOR94] I.Norros, "On the use of fractional Brownian motion in the theory of connectionless networks", COST242 technical contribution TD94-33, September 1994.

[PAX94] V.Paxson, "Growth Trends in Wide-Area TCP Connections", IEEE Network Magazie, July/August 1994, pp 8-17.

[RAO95] S.Rao, M.Rumekasten, M.Frank, (editors), "Performance Results from the Heterogeneous Broadband Testbed with DQDB-ATM Interworking Units and CL-Server", R2032 COMBINE deliverable 13, R2032/KTA/WP4/DS/P/013/1.

[ROB95b] S.Robert, "Can self-similar traffic be modelled by Markovian processes", Contribution to COST242 meeting in Stockholm, 15 May 1995.

[ROB95] S.Robert, C. van den Branden Lambrecht, "An SSMP Traffic Model Based on Network Measurements", 1995 IEEE International Phoenix Conference on Computers and Communications, March 28-31, 1995, Phoenix.

[RUM91] M. Rümekasten, A. Calcagno, "ATLAS Reference Manual", release 5, University of Paderborn, Dept. Mathematics and Computer Science, December 1991.

[YOK94] T.Yokotani, T.Shikama, "Bandwidth Allocation for Connectionless Service in Private Networks Based on ATM Technology", IEICE Trans. Commun., Vol.E77-B, No.3, March 1994, 386-395.