# Exploring the Queueing Behaviour of ATM Switches

Nils Bjorkman, Alexander Latour-Henner, and Aziz Miah
Telia Research Laboratories

Simon Crosby and Ian Leslie
Cambridge University Computer Laboratory

Matthew Davey, Raymond Russell, and Fergal Toomey
Dublin Institute for Advanced Studies.

March 1996

**Abstract**

In this paper we describe methods which we have used to verify queueing models of two ATM switches, namely the Fairisle ATM switch developed at Cambridge University Computer Laboratory and an ATM switch installed in the experimental test-bed at Telia Research Laboratories. Our approach relies on methods from large deviation theory, which we believe provides an appropriate scale on which agreement may be expected between a simple model and an actual ATM switch. We found that the single-server queue provides an adequate model, at the large deviation scale, of the Fairisle switch; and preliminary results with the Telia switch lead us to a similar conclusion. The model was evaluated using both analytic techniques and simulation. We emphasize that the model is not exact, indicating that very complex models may be required in order to obtain results more detailed than those afforded by large deviation theory. This calls into question the common practice of using detailed models of network traffic in conjunction with simplified models of the network itself.

## 1 Introduction

The queueing behaviour of traffic passing through a switch has an important impact on the quality of service it receives. Once a queueing model of the switch has been formulated, the behaviour of the traffic and thus the quality of service achieved can be analysed using a variety of techniques. It is desirable to have a simple switch model; however, practical switches generally have complex designs and therefore the formulation of a model is not a trivial task. Queueing may occur at several points within a switch, but only a small number of these queues may actually make a significant contribution to delays. The problem therefore is to determine which features of a switch have significance for its queueing properties and which can safely be neglected.

In this paper we describe an experimental procedure which we have used to analyse two ATM switches, the Fairisle switch developed at Cambridge University Computer Laboratory and a switch installed in the ATM test-bed at Telia Research Laboratories. The objective of our work was to develop switch models which could be used as the basis of a large deviation analysis of their performance. Therefore, we were particularly interested in testing the validity of the models at the large deviation scale. (The relevance of large deviation theory to queueing systems is explained in section 2). Our approach was empirical: we loaded the test switch with known (stochastic) traffic patterns and then compared its behaviour with simple queueing models. Comparison was made both with analytical and simulation results. We found that a simple single-server queue provides an adequate model of the Fairisle switch under the conditions we explored. This model is *not* exact: it is accurate at the scale of large deviations but does not reproduce every aspect of the switch's behaviour. It seems likely therefore that detailed results on queueing behaviour obtained using this simple model, along with a complex traffic model, would in fact be invalid. Preliminary results available for the Telia switch indicate that it too may be modelled as a single-server queue, at the large deviation scale.

Fairisle is an experimental ATM platform built at the University of Cambridge Computer Laboratory to permit research into performance, architectural, and management issues in ATM networks [1]. It is based on a network architecture composed of ATM switches which are interconnected to form a general topology ATM LAN. The Fairisle switch is based on a simple, $4\times4$ input buffered crossbar. Switches of degree greater than 4 are constructed by using multiple $4\times4$ crossbars in a 2-stage delta interconnection network.

The hardware design of the Fairisle network components permits a high degree of experimental flexibility owing to the use of field programmable gate arrays which implement a number of features to support experimental work, including high resolution clocks and hardware time stamping of cells. Fairisle can be configured as either an input or output buffered switch. Each port is equipped with 2K cell buffers, 4 MB of RAM and an ARM RISC processor, which implements all policy-specific traffic control functions, using a flexible control plane which exports RPC-style interfaces [2, 3] for signalling and management. This has enabled us to implement a wide range of measurement facilities to accurately monitor network performance.

The switch at Telia is an 16 port output-buffered switch operating at a line rate of 155 Mb/s. The test-bed incorporates high resolution measurement tools allowing the experimenter to record traffic traces and cell-delay distributions, as well as a synthetic traffic generator capable of producing a wide variety of CBR and VBR traffic streams for test purposes. For a description of some of the tools in use at Telia see [4].

## 2 Methodology

Large deviation theory predicts that, for a wide range of stochastic traffic models, the queueing properties of the traffic can be deduced from its large deviation rate function. For example, if traffic is passed through a single-server queue served at constant rate $s$, then the resulting steady-state queue length distribution will have exponential asymptotics of the form,

$$\mathrm{P}\{Q > q\} \sim \exp(-\delta q), \tag{1}$$

with a decay parameter $\delta$ which can be calculated from the rate function $l(\cdot)$ of the model:

$$\delta = \min_{c>0} l(c+s)/c. \tag{2}$$

Similar results are known to hold for a variety of other simple queueing systems: see, for example, [5, 6]; a general introduction to this area may be found in [7].

A model of a switch may thus be evaluated by comparing the switch's actual performance under known load conditions with what we would expect from large deviation theory. In our experiments with the Fairisle and Telia switches we used the simple discrete time 'on-off' model to generate traffic. This is a two-state Markov model, which generates back-to-back cells when in the active state and produces empty slots in the silent state. The lengths of successive active and silent periods are independent and geometrically distributed. The transition matrix of the model is thus,

$$T = \begin{bmatrix} 1-a & a \\ d & 1-d \end{bmatrix}, \tag{3}$$

where $a$ is the probability of transitions from the silent state to the active state, and $d$ is the probability of transitions in the reverse sense. The mean activity of the source is then $a/(a+d)$. As a measure of source burstiness we use the parameter $\gamma = 1-a-d$, which in steady-state is equal to the correlation between the activity in successive time slots. $\gamma$ takes values between $+1$ and $-1$, representing respectively strong positive or strong negative correlations. When $\gamma = 0$ the model produces Bernoulli traffic. The large deviation properties of this simple traffic model can readily be calculated, as follows.

Let $A_n$ represent the number of cells generated in an interval of length $n$. The rate function $l(\cdot)$ of the traffic process is the Fenchel-Legendre transform of the scaled cumulant generating function (CGF) $\lambda(\cdot)$:

$$l(x) = \max_{\theta}\{\theta x - \lambda(\theta)\}, \tag{4}$$

where $\lambda(\theta)$ is defined by,

$$\lambda(\theta) = \lim_{n\to\infty} \frac{1}{n} \ln \mathrm{E} \exp(\theta A_n). \tag{5}$$

The rate function and the CGF contain the same information. For example, the decay parameter $\delta$ of the queue length distribution in a single-server queue may be calculated either from $l(\cdot)$ using equation (2), or alternatively from $\lambda(\cdot)$ using the formula,

$$\delta = \max\{\theta > 0 : \lambda(\theta) < s\theta\}. \tag{6}$$

Thus $\delta$ is the ordinate of the intercept of the graph of $\lambda(\cdot)$ and a line of slope $s$ through the origin. Let $\hat{T}(\theta)$ be the *twisted transition matrix* associated with $T$:

$$\hat{T}(\theta) = \begin{bmatrix} 1-a & ae^{\theta} \\ d & (1-d)e^{\theta} \end{bmatrix}, \tag{7}$$

and let $\phi(\theta)$ be the spectral radius of $\hat{T}(\theta)$. Then it can be shown that $\lambda(\theta) = \log\phi(\theta)$ (see [8], theorem 3.1.2). This formula for the CGF can easily be evaluated numerically.

As an alternative to calculations with models the large deviation properties of a traffic process can also be estimated directly from traffic measurements, using an approach described in [9]. For example, a simple estimator of the CGF is provided by

$$\hat{\lambda}_T(\theta) = \frac{1}{T} \log \frac{1}{N} \sum_{i=1}^{N} \exp(\theta A_T^{(i)}), \tag{8}$$

where $A_T^{(1)}, A_T^{(2)}, \ldots$ are the numbers of cells emitted in successive time intervals of length $T$. Estimates of the CGF can be used to compare the large deviation properties of different traffic samples, or to predict their queueing properties. Various issues involved in estimating the CGF are discussed in [9, 10, 11].

As well as looking at the predictions of large deviation theory we have also compared the behaviour of the two switches with a simple simulation model. In simulations as in experiments the performance parameters are of course 'random' objects which depend on the seed used to initialise the traffic model. Therefore experiments must be repeated many times with different, independent, initialisation conditions in order to build up a picture of the distribution of the parameters of interest.

# 3 Results

The experiments with the Fairisle switch were conducted at Cambridge Computer Laboratory. The switch was configured as an output-buffered switch with FIFO queueing. Traffic traces from the two-state Markov model were prepared from a simple traffic generator and then passed through the switch in the absence of any other traffic. Load conditions were varied using a feature of Fairisle which allows the input and output link clock rates to be set independently of each other; thus the transmission rate can be slowed down relative to the arrival rate. Another feature of Fairisle allows the empirical queue length distribution during an experiment to be recorded, and then compared with analytical and simulation predictions. A trace of the departure process was also recorded.

The Telia switch was tested at Telia Research Laboratories in Stockholm. Two-state Markov traffic was produced using the Synthetic Traffic Generator, a traffic simulator implemented in hardware which is described in [12]. Load conditions were varied by multiplexing two-state and CBR traffic of fixed intensity into the same output port. A trace of the departure process associated with the two-state stream was recorded, along with the end-to-end cell-delay distribution.

In both series of experiments the parameters of the traffic model were set to $a = 1/16$, $d = 3/16$. This gives a mean source activity of 0.25 and burstiness parameter $\gamma$ equal to 0.75. Using the Fairisle switch experiments were performed at loads of 65%, 75%, 85%, and 95% by adjusting the link clock rates as described in the previous section. For each load, ten experiments with a run length of $10^6$ cells each were carried out using different seeds for the traffic generator. In the case of the Telia switch the mean intensity of the two-state traffic was set to 25% of the output link rate. In successive experiments this was multiplexed with CBR traffic of intensity 30%, 35%, 45%, and 50% of the link rate, giving loads of 36%, 38%, 45%, and 50%.

The empirical queue length distributions from the Fairisle switch were compared with distributions obtained by passing the same traffic through a software implementation of a single-server queue with constant service rate. The load in the simulated queue was adjusted to be equal to that obtained in the switch during the experiment. The two sets of distributions are not identical; however they both exhibit asymptotically exponential decay, with decay parameters clustered around the calculated theoretical value (see figure 1). The spread of the decay parameters from each set is approximately the same. Furthermore, they agree sample path by sample path; in other words, for the same initialisation conditions, the value of $\delta$ obtained from the switch and that obtained by simulation were approximately the same. Thus from a large deviation perspective the queue length distribution in the switch can be regarded as identical with that obtained in a single-server queue with constant service rate.

Similar results were found for the departure process from Fairisle. Figure 2 shows a section of the departure trace from the switch and that from a simulated single-server queue for one choice of load (65%) and seed. The traces are very similar, but not entirely the same. Of particular interest to us is the question of whether they are distinguishable on a large deviation scale. To determine this, we compared the measured scaled cumulant generating functions of each trace. Figure 3 shows the measured CGFs of the two traces depicted in figure 2. The measured CGF is of course a statistical object which varies between different samples of the same process. We found that the difference between the two CGFs in the figure is much smaller than the 'natural' variation of the measured CGF for this traffic model; hence they are statistically indistinguishable. We conclude that the observed differences between the two departure traces are neglible at the large deviation scale.

The measurements made with the Telia switch again indicate the sort of large deviation behaviour expected from a single-server queue. The cell-delay distribution is expected to have an exponential tail of the form,

$$\mathrm{P}\{D > d\} \sim \exp(-\delta s), \tag{9}$$

where $\delta$ is the decay parameter found from equation (2) and $s$ is the service rate. The delay distributions obtained from the switch are indeed found to exhibit the exponential decay prescribed by (9). A preliminary comparison with simulation indicates that the departure process is also similar to that obtained from single-server queue.

# 4 Conclusions

The results of our experiments show that the single-server queue with constant service rate provides, at the large deviation scale, an appropriate model of the output-buffered Fairisle switch with FIFO queueing. That the results from the model were not entirely identical with the results from the switch is presumably due to the complexity of the switch hardware. Nevertheless, large deviation predictions for the queue length distribution and for the departure process were accurate. Results available at this stage for the Telia switch indicate that the single-sever queue is a valid model at the large deviation scale in this case also.

We believe that the approach we have used here will be useful for developing models of more complex systems. Given the complexity of modern switches it is unlikely that any simple model will produce exact agreement with experiment. However, large deviation theory provides an appropriate scale on which agreement may be expected.

# References


[1] R. Black, I. Leslie, and D. McAuley. Experiences of building an ATM switch for the Local Area. In *SIGCOMM*, volume 24(4), September 1994.

[2] R. Black and S. Crosby. Experience and results from the implementation of an ATM socket family. In *Proceedings USENIX Winter Technical Conference*, San Francisco, January 1994.

[3] U. of Cambridge Systems Research Group. ATM document collection. Technical note, University of Cambridge Computer Laboratory, Pembroke Street, Cambridge, CB2 3QG, UK, Mar. 1994. available as: http://www.cl.cam.ac.uk/Research/SRG/bluebook.html.

[4] N. Bjorkman and A. Latour-Henner. High time-resolution cell loss measurements of mixed cbr and vbr traffic. In *IFIP Conference on Broadband Communications '94*, Paris, March 1994.

[5] N. G. Duffield and N. O. Connell. Large deviations and overflow probabilities for the general single server queue, with applications. To appear in Proc. Cam. Phil. Soc.; available as: ftp:stp01.stp.dias.ie/DAPG/dapg9330.ps.

[6] N. O'Connell. Large deviations for queue lengths at a multibuffered resource. Submitted to Appl. Prob.; available as: ftp:stp01.stp.dias.ie/DAPG/dapg9434.ps.

[7] A. Schwartz and A. Weiss. *Large Deviations for Perfomance Analysis.* Chapman and Hall, 1995.

[8] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications.* Jones and Bartlett Publishers, 1993.

[9] N. G. Duffield, J. T. Lewis, N. O'Connell, R. Russell, and F. Toomey. Entropy of ATM traffic streams. *IEEE Journal on Selected Areas in Communications, Special issue on advances in the fundamentals of networking – part 1*, 13(6), August 1995.

[10] S. Crosby, I. Leslie, J. T. Lewis, N. O. Connell, R. Russell, and F. Toomey. Bypassing modelling: an investigation of entropy as a traffic descriptor on the fairisle ATM network. In *Proceedings of the 12th UK Teletraffic Symposium*, Windsor, March 1995.

[11] S. Crosby, I. Leslie, M. Huggard, J. T. Lewis, F. Toomey, and C. Walsh. Bypassing modelling: further investigations of entropy as a traffic descriptor in the fairisle ATM network. In *Proceedings of the First IFIP Workshop on ATM Traffic Management*, Paris, December 1995.

[12] B. Helvik, O. Melteig, and L. Morland. The synthezised traffic generator; objectives, design, and capabilities. In *Proceedings of IBCN&S*, Copenhagen, 1993.

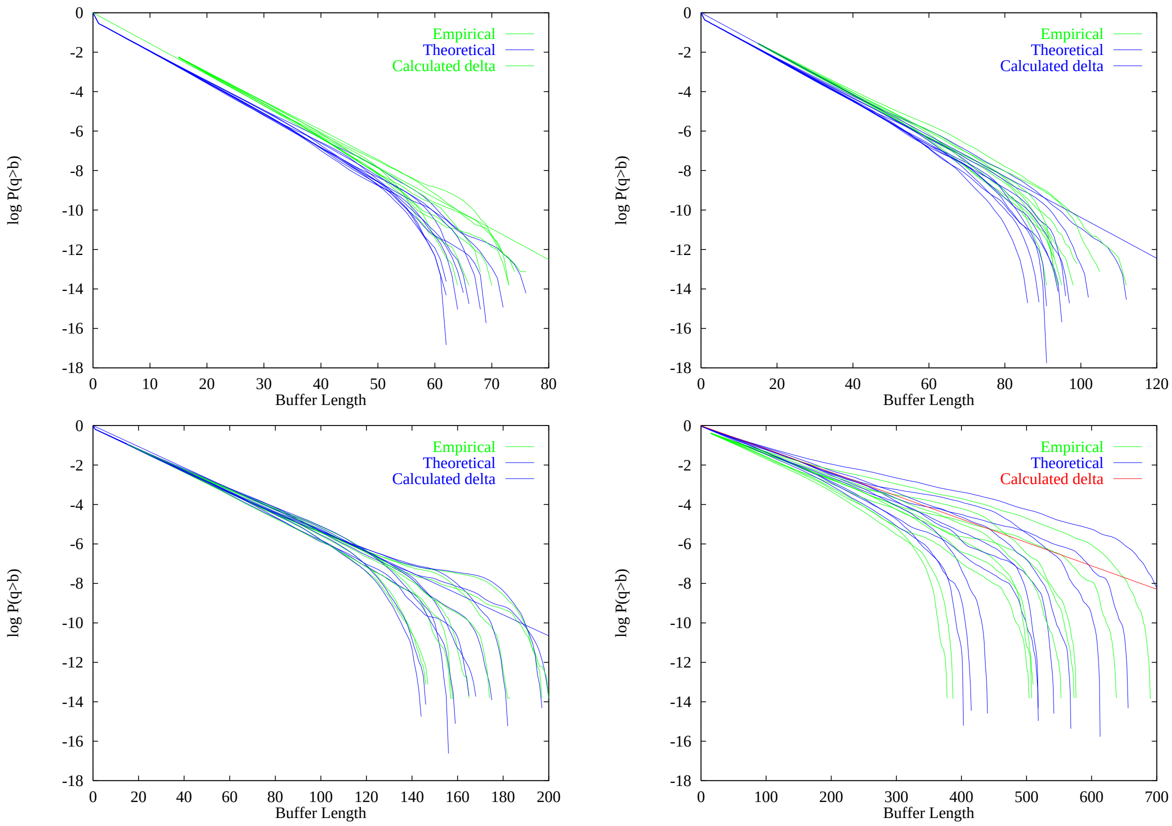


Figure 1: Empirical queue length distributions from the Fairisle switch and from simulation of a single-server queue, at loads of 65%, 75%, 85%, and 95%.

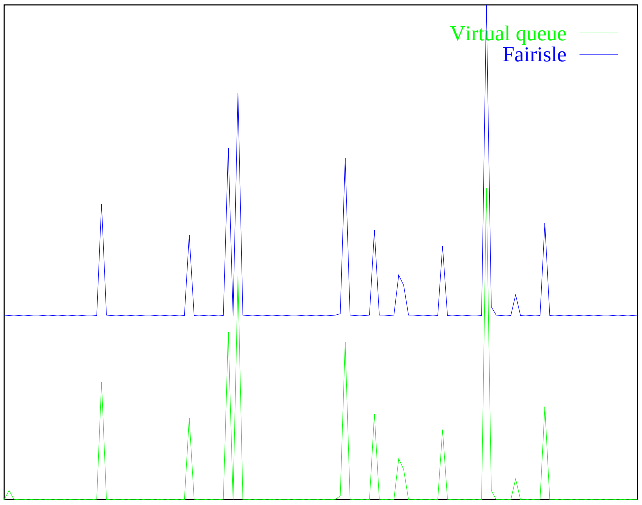


Figure 2: Departures from the Fairisle switch and from simulation (viewed at the burst level in cells per unit time vs. time)

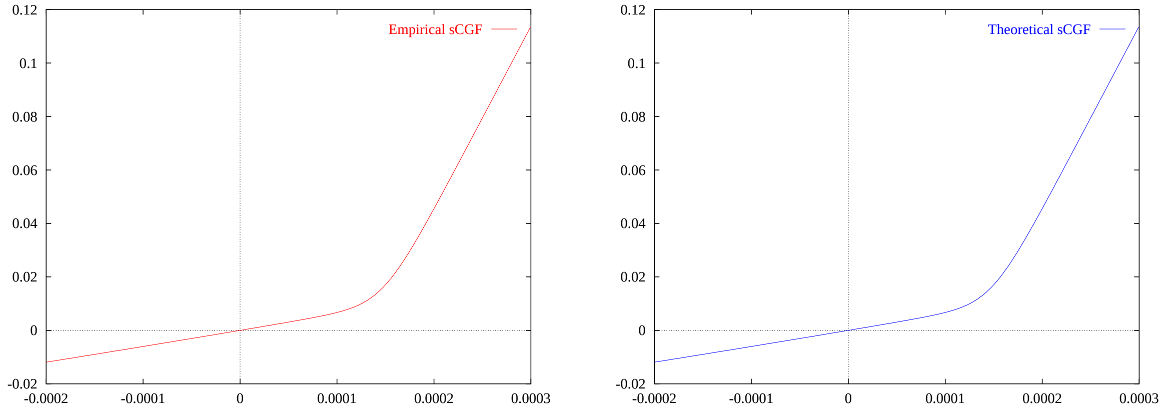


Figure 3: Measured CGFs of the Fairisle and simulation departure traces.