23

On the accelerated simulation of VBR virtual channel multiplexing in a single-server FIFO buffer

M. J. Tunnicliffe, D. J. Parish,
Department of Electronic and Electrical Engineering,
Loughborough University, Ashby Road, Loughborough,
Leicestershire, United Kingdom, LE11 3TU.
Tel.: (01509) 228117, Fax.: (01509) 222854.
e-mail: M.J.Tunnicliffe2@Lboro.ac.uk
WWW: http://info.Lboro.ac.uk/departments/el/research/hsn/index.html

Abstract

While direct *cell-level* simulation accurately predicts congestion in cell-switched networks, excessive run-times are often required to obtain significant results. Methods of *Accelerated* simulation have therefore been developed, examples of which include the *cell rate* technique (which represents the discrete cell-streams as continuous fluids) and the *histogram* method (which merges the multiplexed streams into an aggregate cell-rate histogram and performs independent statistical analysis on each bin). The current work applies both these techniques to a simple ATM multiplexer and explores their respective advantages and drawbacks. While the cell-rate method provides accurate predictions under a rapidly varying bit rate, the histogram method is more successful under quasi-static conditions. This suggests the possibility of a hybrid cell-rate/histogram model which is accurate at both extremes.

Keywords

ATM networks, simulation techniques, statistical analysis.

1 INTRODUCTION

The recent proliferation of cell-switched communication networks has led to increasingly complex problems in their design, evaluation and management. Such problems, many of which arise from congestion as virtual channels are multiplexed, lead to cell-losses and transmission-time *jitter*, the latter being particularly harmful to real-time services such as video.

D. D. Kouvatsos (ed.), *ATM Networks*

Various computer-aided techniques have been devised for the analysis of networking problems (see Kurose and Mouftah 1988 and Frost et al. 1988). The most direct approach is *cell-level simulation,* in which network components are directly represented within the software, and cell arrivals and transmissions are mimicked by pseudorandom sequences. Such simulations are highly processor-intensive, and enormous run-times are often required to simulate relatively short periods of operation. For example, the Orwell simulator recently developed at Loughborough University (Parish et al., 1994) can take several hours to simulate one minute of real-time, and since *acceptable* loss rates are of the order of 10^{-9} (i.e. 1 lost cell in 10^9), several weeks may be required to obtain statistically significant characterization.

For this reason, numerous workers have investigated *accelerated* simulation techniques, which allow run-times to be reduced without major loss of accuracy. One example is *variance reduction* which manipulates the statistical properties of a cell-level model in order to reduce the stochastic variability of its output, thus shortening the run-time needed to obtain statistical significance (Frost et al, 1988). However, the current paper concentrates on the following recently-published techniques:

- The *cell rate* method, developed by Pitts et al. (1994 a,b) at Queen Mary & Westfield College, London, represents the various discrete cell-streams applied to the input buffer of an ATM multiplexer as continuous fluids, whose flow-rates are modulated by the bulk-traffic characteristics. This method has been shown to produce accurate cell-loss predictions in the *burst-scale* where the aggregate cell-rate exceeds the channel capacity.
- The *histogram* method, introduced by Skelly et al. (1993) at the University of Columbia, NY, converts the incoming cell-streams of an ATM multiplexer into arrival-rate histograms and convolves them together to form an aggregate histogram. Statistical queueing analysis is applied separately to each histogram bin, and the results are then combined as a weighted sum. It is a fundamental assumption of this model that the system is in statistical quasi-equilibrium, and it is therefore unsuitable for rapidly varying bit-rates.

The current paper applies variants both these techniques to a simple two-channel multiplexer. The predictions are compared with the results of a stochastic cell-level simulator and their respective accuracies and run-times are contrasted.

2 CELL-LEVEL SIMULATION

Before the accuracies of any accelerated simulation techniques could be tested, it was first necessary to establish a cell-level simulator against which their predictions could be compared. Figure 1(a) shows a schematic diagram of the ATM multiplexer modelled in the software (which was written in Turbo-C and ran upon a 486-based desktop microcomputer). Time was quantised into *cycles,* during each of which up to one cell could arrive on each input channel and up to one cell could be read by the server. The latter operated in a *geometric* mode, in which there was a constant probability (μ) per cycle of a cell being read.

Each of the two buffer inputs could be fed with any user-defined data-stream. If both channels generated a cell within the same cycle (i.e. a *batch* arrival) both were simultaneously loaded onto the buffer in a randomly selected order (i.e. each cell had equal probability of getting first place in the buffer).

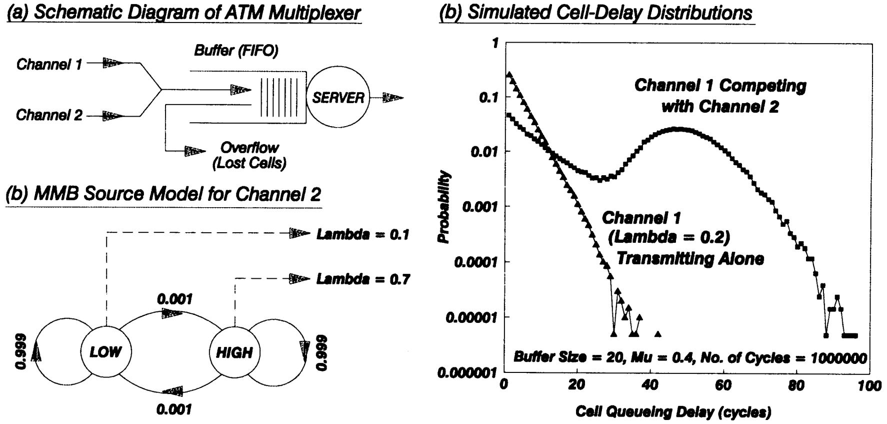

Figure 1: Cell-level simulation of channel interaction in an ATM multiplexer.

Figure 1(c) shows some typical results obtained from the simulator using a buffer-length (N) of 20 and a geometric service-probability of 0.4 per time-slot. Firstly, an unmodulated Bernoulli stream (arrival probability λ=0.2) was sent through the buffer on Channel 1 with Channel 2 inactive, and the cell-delay distribution was recorded. The experiment was then repeated with an additional 2-state Markov-modulated Bernoulli stream applied to Channel 2 (Figure 1(b)), and the subsequent deterioration of transmission quality (i.e. increased cell-delay) is clearly visible in the results (Figure 1(c)). The upper mode in the cell-delay distribution clearly represents the *burst* component, where the aggregate cell arrival rate exceeds the server capacity and the buffer becomes normally full.

3 HISTOGRAM SIMULATION

The analysis presented in this section assumes that the buffer is in statistical equilibrium, and hence that the equilibrium probabilities Π_0 .... Π_N remain constant with time. (Π_n is the probability that the buffer contains n cells).

Statistical Queueing Analysis

If only a single Bernoulli stream is applied then the buffer can be modelled as a discrete-time Geo/Geo/1/N queue, the solution of which is a matter of simple textbook theory. For $0 \leq n < N$, the equilibrium probabilities are given by:

$$\Pi_n = \frac{(1-\gamma)\cdot\gamma^n}{1 - \frac{\lambda}{\mu}\cdot\gamma^N} \quad where \quad \gamma = \frac{\lambda\,(1-\mu)}{\mu\,(1-\lambda)} \tag{1}$$

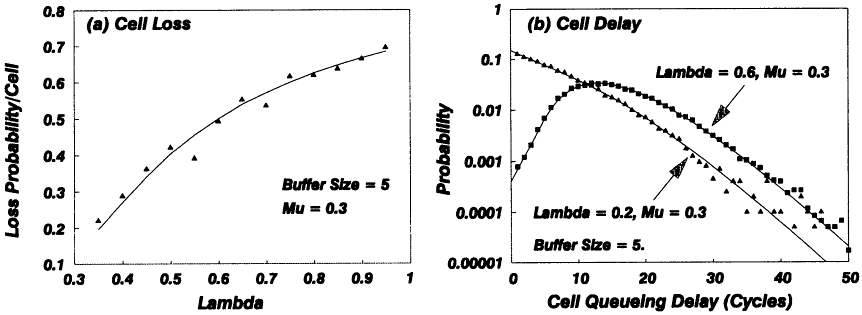

Figure 2: Simulated behaviour of Geo/Geo/1/N buffer compared with analytical model predictions. (Discrete points indicate simulations, solid lines indicate the analytical model.)

while for $n = N$:

$$\Pi_N = \frac{(1-\gamma)(1-\lambda)\cdot\gamma^N}{1 - \frac{\lambda}{\mu}\cdot\gamma^N} \tag{2}$$

If an arriving cell finds n ($<N$) cells ahead of it, then it remains in the buffer until the latter has been read $(n+1)$ times. Hence the probability that queueing delay is equal to k cycles is given by

$$P(k) = \sum_{n=0}^{N-1} \Pi_n \, \mu^{n+1} (1-\mu)^{k-(n+1)} \binom{k-1}{n} \tag{3}$$

Substituting Eqn.(1) for Π_n and simplifying yields

$$P(k) = \Pi_0 \, \mu (1-\mu)^{k-1} \left[(1-\lambda)^{1-k} - \sum_{i=N}^{k-1} \left[\frac{\lambda}{1-\lambda}\right]^i \binom{k-1}{i} \right] \tag{4}$$

Since the final term in (assumed zero for $k<N+1$) is a truncated binominal series, it may be replaced by the *incomplete beta function* $I_\lambda(N, k-N)$. The expression may now be re-written:

$$P(k) = \Pi_0 \, \mu \left[\frac{1-\mu}{1-\lambda}\right]^{k-1} \cdot \left[1 - I_\lambda(N, k-N)\right] \tag{5}$$

If an incoming cell finds N cells already in the buffer then the latter is full and the new cell must therefore be lost. Hence the loss probability is equal to Π_N and may therefore be computed using Equation 2. Figure 2 compares the analytical cell-loss and cell-delay

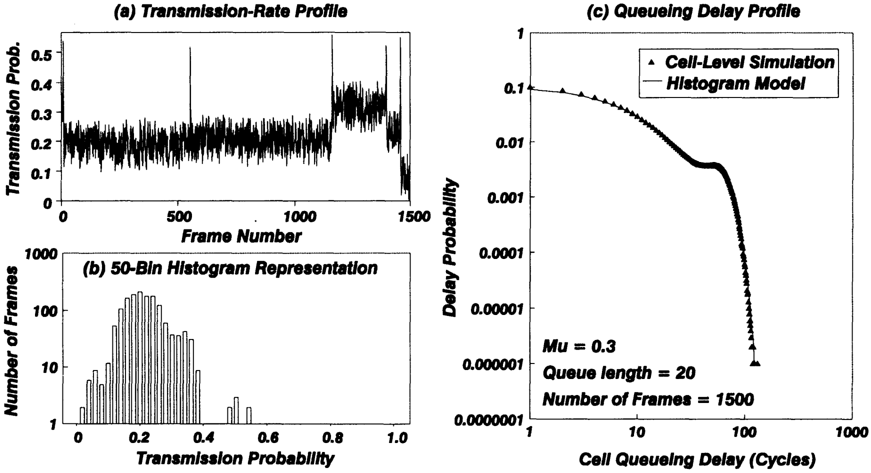

Figure 3: Example of single-channel histogram simulation.

characteristics with the results of cell-level simulations. (Numerical values for the incomplete beta function I_λ were computed using the algorithm supplied in *Numerical Recipes in C* by Press et al. (1988).)

Single Channel Histogram Simulation

Figure 3(a) shows an example of the simulated variable bit rate (VBR) video profiles used in this study. (The video simulation was based upon the output of an experimental VBR codec during the compression of "head-and-shoulders" image sequences. The occasional high cell-rate excursions correspond to scene-changes within the sequence, while the smaller variations indicate activity within individual scenes.) The duration of each video frame was 28 276 cycles, over which the arrival probability λ remained constant. (This period was sufficiently long for the assumption of statistical equilibrium to be approximately valid).

Figure 3(b) shows the same video profile expressed as a 50-bin arrival-probability histogram. Independent statistical queueing analysis was performed upon each bin (for a buffer size of N=20 and service rate λ=0.3), after which the results were weighted according to their relative frequencies, and finally summed to obtain the overall delay and loss characteristics. Figure 3(c) shows the resulting queue-delay distribution (solid lines) compared with a cell-level simulation of the same scenario (discrete points). The cell-loss ratio was computed as 2.56% by the histogram model, compared to 2.14% predicted by the cell-level simulation.

Since the distributions in Figure 3(c) are practically indistinguishable, the histogram curve clearly provides a highly accurate approximation of the simulated data. As the histogram results were obtained in approximately 1% of the cell-level run-time (i.e. 111 seconds compared to 10 494 for the cell-level simulation), the experiment illustrates the potential value of the histogram method as a means of reducing simulation run-time.

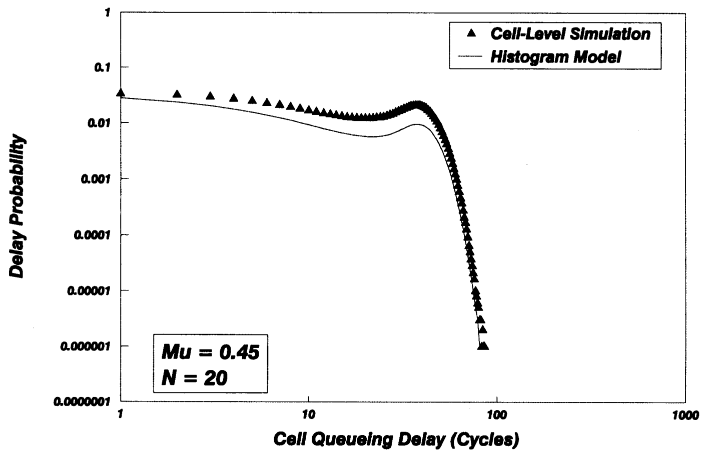

Figure 4: Typical results of two-channel histogram simulation.

Two-Channel Histogram Simulation

Since the primary focus of this paper is the interaction of competing virtual-channels in a common buffer, the above theory must be extended to cover the effects of multiple Bernoulli inputs. If the cell arrival probabilities for channels 1 and 2 are λ_1 and λ_2 respectively and p_n is the probability of n arrivals per cycle, then:

$$p_0 = (1 - \lambda_1)(1 - \lambda_2) = 1 - (\lambda_1 + \lambda_2) + \lambda_1 \lambda_2 \tag{6}$$

$$p_1 = \lambda_1 (1 - \lambda_2) + \lambda_2 (1 - \lambda_1) = \lambda_1 + \lambda_2 - 2 \cdot \lambda_1 \lambda_2 \tag{7}$$

$$p_2 = \lambda_1 \lambda_2 \tag{8}$$

The resultant arrival stream is an example of a *batch* Bernoulli process (maximum batch size = 2) whose effects upon discrete-time queueing have been analytically studied by Dafermos et al. (1971) and more recently by Hashida et al. (1991). However, the current analysis employs the following simplifying assumption: If λ_1 and λ_2 are both $\ll 1$ then $\lambda_1\lambda_2$ becomes negligible and the aggregate stream approximates to a standard Bernoulli process with arrival probability $(\lambda_1 + \lambda_2)$.

Figure 4 shows some typical histogram and cell-level results for the interaction of two independent VBR streams in a common buffer. The latter were initially converted into individual arrival-rate histograms, which were then convolved together to form the aggregate histogram. Although a wide divergence exists in some parts of the graph, the general agreement in the shapes of the curves illustrates the value of the technique. In view of the accuracy of the one-channel case, these errors are almost certainly associated with the

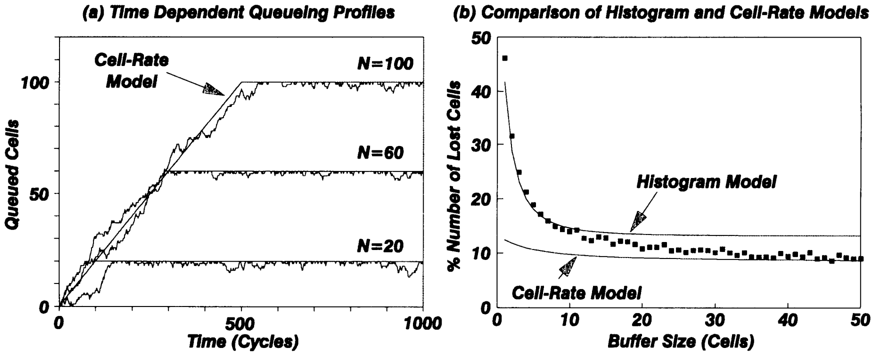

Figure 5: Illustration of cell-rate model, and comparison with the predictions of cell-level and histogram simulations.

aggregation approximation (Section 3) and/or the histogram convolution technique which is, strictly speaking, applicable only to stationary stochastic processes. Extension of the time-domain would therefore be expected to improve prediction accuracy.

4 CELL-RATE METHOD

The cell-rate simulation technique described below is a simplified version of that published by Pitts et al. (1994 a,b). Its main function is to show how the primary properties of this algorithm differ from those of the histogram method described above.

Basically, the cell-rate model ignores the discrete nature of the cell streams, and represents them as continuous fluids modulated by a *burst traffic* profile. The latter is composed of constant cell-rate *bursts*, punctuated by discontinuous cell-rate changes known as *events*. Within each burst, when the buffer is neither full nor empty, the number *n(t)* of cells in the queue varies with time *t* (cycles) according to the equation

$$n(t) = n_0 + \left[\sum_{r=1}^{k} \lambda_r - \mu \right] (t - t_0) ; \quad 0 < n(t) < N. \tag{9}$$

where t_0 is the time at which the burst began, n_0 is the number of queued cells when $t=t_0$ and k is the number of multiplexed streams. This transient phase ends when the queue becomes full or empty, and n remains equal to N or 0 until the end of the burst. Figure 5(a) shows these *transient* and *steady state* phases compared with the corresponding cell-level simulations for an initially empty Geo/Geo/1/N queue. When the buffer is full and the aggregate cell-rate exceeds the server capacity, losses occur at a rate of $(\Sigma\lambda - \mu)$ cells per cycle, and are distributed between Channels 1 and 2 according to the ratio $\lambda_1:\lambda_2$.

Unlike the Pitts et al. model (which handles network events concurrently) the cell-rate simulator program tracks buffer occupancy from burst-to-burst throughout the range of the

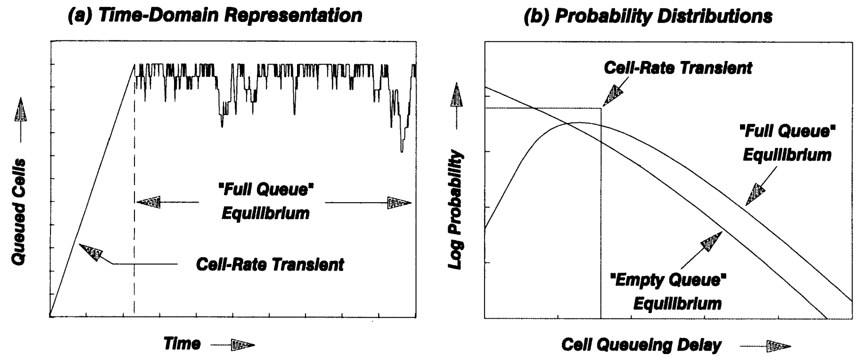

Figure 6: Illustration of proposed hybrid cell-rate/histogram model.

simulation, recording the number of lost cells. Figure 5(b) shows some typical cell-loss characteristics compared with the results of the cell-level and histogram models. During a high arrival-rate burst when $\Sigma\lambda/\mu > 1$ and cell-losses become significant, the length of the expanding queue is constrained by the buffer capacity and the transient phase can be expected to be of the same order as the buffer-filling time, i.e. $t_{\mathrm{tran}} \sim N / (\Sigma\lambda - \mu)$. Hence the observed increase in accuracy with increasing N. However, when N is small and $t_{burst} \gg t_{tran}$, transient phenomena can be entirely neglected and the queue assumed to be in equilibrium throughout the simulation. Hence when N becomes small, the histogram model provides the best predictions.

5 CONCLUSIONS AND FUTURE WORK

This paper has presented some early results from an ongoing study of computer-aided communication-network modelling. The initial stage of the work involved the design and testing of a cell-level simulator for a single-server FIFO buffer with a geometric read-time distribution, fed by two independent cell-streams. (This system could provide a module in a full network simulator). The same system was also modelled using the *cell-rate* technique and a *histogram* model based upon statistical queueing theory. Although both these models produced results within significantly shorter run-times than the cell-level simulator, they were found to be accurate only within certain regions of parameter-space. We now consider the possibility of a *hybrid* model, combining the respective virtues of these two algorithms.

Such an algorithm would be required to model the stochastic nature of both the transient and steady-state conditions of operation. Although several transient models are available for the unbounded, continuous-time M/M/1/∞ queue (a computationally efficient formula has recently been developed at Bradford University (Bundey, 1995)), the finite capacity of the ATM buffer presents severe theoretical difficulties. One possible solution is illustrated in Figure 6(a): The cell-rate model is applied during all transient phases of operation, while the statistical equilibrium model is used during periods of statistical equilibrium (i.e. when the value of n predicted by the cell-rate model is either 0 on N). The cell delay distribution in a transient phases might be represented to some degree of accuracy by a rectangular function of width μN and height $1/\mu N$ (Figure 6(b)).

It should be noted that the results represent only the most preliminary findings of an ongoing study of network modelling, and are not intended to form a definitive treatment of the subject. Investigations have so far been confined to a single network component, consisting of a single queue and a single server, under relatively simple traffic-loading conditions (although some of the bulk statistics *were* based upon a realistic VBR video-source model). The model must ultimately be extended to cover a network of many such interconnected units under more generalized traffic, which may include such complicating effects as correlation in the cell-generation process (Skelly, 1994). The validity of the resulting model must then be checked by comparing its predictions against the operational statistics of an actual hardware network under realistic traffic-loading conditions.

6 ACKNOWLEDGEMENTS

The authors wish to thank all members of the High Speed Networks research group at Loughborough University for their suggestions and technical assistance. The work was funded by an EPSRC-ROPA grant.

7 REFERENCES

Bunday,B. (1995), Dept. of Mathematics, University of Bradford, private communication.

Dafermos,S.C., Neuts,M.F. (1971) A Single Server Queue in Discrete Time, *Cahiers du Centre D'étude de Rechherche Opérationnelle*, **19**, 23-40.

Frost,V.S., Wood Larue,W, Shanmugan,K.S. (1988) Efficient Techniques for the Simulation of Computer Communications Networks, *IEEE J-SAC*, **6**, 146-57.

Hashida,O, Takahashi,Y, Shimogawa,S (1991) Switched Batch Bernoulli Process (SBBP) and the Discrete-Time SBBP/G/1 Queue with Application to Statistical Multiplexer Performance, *IEEE J-SAC*, **9**, 394-401.

Kurose,J.F., Mouftah,H.T. (1988) Computer-Aided Modelling, Analysis, and Design of Communication Networks, *IEEE J-SAC*, **6**, 130-45.

Parish,D.J., Rogers,C, Nche,C, Ruiz,I (1994) Modelling the Orwell Network Access Protocol on a Slotted Ring, in *Computer and Telecommunication Systems Performance Engineering* (eds. M.E.Woodward, S.Datta, S.Szumko), Pentech Press, London, 108-13.

Pitts,J.M., Cuthbert,L.G., Bocci,M., Scharf,E.M. (1994a) Cell Rate Modelling: An Accelerated Simulation Technique for ATM Networks, ibid., 94-107.

Pitts,J.M., Cuthbert,L.G., Bocci,M, Scharf,E.M. (1994b) An Accelerated Simulation Technique for Modelling Burst-Scale Queueing Behaviour in ATM, *Teletraffic Congress, 14th. International Conference*, **1**, 777-86.

Press,W.H, Teukolsky,S.A., Vetterling,W.T., Flannery,B.P. (1988) Numerical Recipes in C, Cambridge University Press.

Skelly,P, Schwartz,M, Dixit,S (1993) A Histogram-Based Model for Video Traffic Behaviour in an ATM Multiplexer, *IEEE/ACM Trans. Networking.*, **1**(4), 446-59.

8 BIOGRAPHIES

Martin J. Tunnicliffe holds B.Eng. and Ph.D. degrees from the Universities of Bradford and Loughborough respectively. His early research was in the field of semiconductor reliability, but he has more recently been involved in the monitoring and analysis of communication networks. He is currently employed as a contract researcher in the High Speed Networks Group at Loughborough University.

David J. Parish holds B.Sc. and Ph.D. degrees from the University of Liverpool. He has worked as a Scientific Officer at the UKAEA Culham Laboratory and as a Demonstrator in the Electrical Engineering Department at Liverpool University. From 1983 he has held the position of Lecturer and later Senior Lecturer in the Department of Electronic and Electrical Engineering at Loughborough University. His research interests concern the management, operation, monitoring and application of High Performance Networks. Specifically, he leads Loughborough's input to the BT funded research programme into the management of high speed networks using SuperJanet.