Call Admission Control Schemes and ATM Network Topological Design

Shuzhi Z. Lo
Graduate Program of Operations Research

Brad A. Makrucki*
BellSouth Telecommunications

Griff L. Bilbro
Electrical and Computer Engineering

Salah E. Elmaghraby
Operations Research and Industrial Engineering

North Carolina State University, Raleigh, NC, 27695

Abstract

Call admission control criteria are not only important for call admission control itself, but also can be an important input to network topological design. In this paper, we show the difference in terms of network cost incurred by adopting different call admission control schemes in network topological design. We compare two call admission control schemes. Scheme 1 uses equivalent bandwidth as its call admission control criterion and Scheme 2 is based on modeling the volatility of call traffic using Reflected Brownian Motion. Though Scheme 2 increases the complexity of network topological design, it can give lower network costs. Our experimental results show that for the same traffic mix, the network cost can be as little as 10% and as much as 35% lower when Scheme 2 is used instead of Scheme 1. The differences between the pair of resulting networks suggests that network topological design can be used as one of the criteria for choosing the call admission control scheme.

Key Words: ATM network topological design, call admission control, integer programming, mixed integer programming

*Now with IBM, Research Triangle Park, NC

Abbreviations

ATM	Asynchronous Transfer Mode
CAC	Call Admission Control
CLR	Cell Loss Ratio
EBW	Equivalent Bandwidth
QOS	Quality Of Service
RBM	Reflected Brownian Motion

1 An Overview of ATM Networking Technology

Asynchronous Transfer Mode (ATM) networking is a new networking technology based on simple principles. It is presently being developed for use as the next generation of widely deployed telecommunications infrastructure. The technology was developed to accommodate the simultaneous transport of voice, video and data streams while assuring proper quality-of-service (QOS) to each connection.

ATM networks generally operate as follows: when a user needs to establish a connection, it places a "call" in a manner similar to what is done in today's telephone network. In ATM, though, the user specifies the type of connection required (for example, from various traffic classes) and the QOS required. If possible, the network establishes a connection from the source to the requested destination. After setup is completed the connection proceeds as in today's voice telecommunication networks. The network is a connection oriented network in the sense that there is a connection for each accepted call.

Since the network accepts or rejects the connection request and this establishes the efficiency of the network, the accept/reject decision is a critical one (since maximal use of network resources is a main objective of network operators). The accept/reject decision is based on whether or not the network would be able to provide the required QOS for the requesting connection while maintaining the QOSs already committed to for the existing connections. Call Admission Control (CAC) algorithms are executed by network switches as connection requests arrive and are responsible for achieving high network element utilization while assuring the requested QOSs. The algorithms are based on the switching node designs and on the nature of traffic in ATM networks.

ATM switching/transport is based on fixed length packets called “cells”. Data is submitted to the network in the form of cell streams (cells are submitted by the connection source once a connection has been established). Voice, video and data all use cell stream submission. ATM switches are network elements that switch cells from input ports/links to output ports/links as cells arrive. Since many arriving cells can be destined for the same output port, output port buffering is typically employed.

Whenever buffers of finite size are employed, there are probabilities of cell loss due to buffer overflow. One of the key QOS parameters is a connection’s cell loss ratio (CLR). It is this measure that is directly addressed by the CAC algorithms compared in this paper. Other measures such as delay and delay variance control are not directly considered here. Traffic sources, besides selecting the QOS that they require, supply a minimalistc descriptor of the traffic that they will submit during their connections. This descriptor (which has been standardized in telecommunications standards) is used by the CAC algorithms to determine the amount of resource that a connection’s traffic will require. In this paper, the traffic descriptor supplied by the source consists of the connection’s peak cell rate, its sustainable cell rate and its maximum burst length (at the peak rate).

An early method for resource use determination uses the simple concept of determining a connection’s “equivalent bandwidth” (EBW). An equivalent bandwidth for a connection can be added to other equivalent bandwidths for connections sharing a link and if the sum exceeds the link's speed (in bits/sec for example), then this would cause buffer overflows and the arriving call would be rejected. More recent work has pursued other directions for predicting resource use, and is overviewed in the next section and in the Appendix.

This paper focuses on the problem of sizing the switches and inter-switch link speeds. This is the "network topological design problem" in ATM. The solution to the design problem depends on the particular CAC scheme used in the optimization algorithm, and so is intimately tied to the nature of ATM switching. To separate the design problem from the CAC problem (which is dependent on the ATM technology specifics) would lead to generic solutions which are primarily illustrative; whereas our attempt has been to solve the ATM design problem by including as much detail as appropriate so that the resulting solution is as accurate as possible.

2 Call Admission Control Criteria

In ATM networks (or other high-speed networking technologies that are connection-oriented), the admission of a new call can be a complex process. Admitting a particular call may result in traffic overloads if that call transmits a large amount of traffic. On the other hand, if a call is rejected then network revenue is not generated. Therefore the goal of a CAC algorithm is to accept calls (and thereby traffic) if and only if they can be handled without violating QOS requirements of all existing traffic.

In ATM transport technology, data is submitted to the network in the form of data units called cells. All cells are the same size and a network connection's traffic contracts for a certain CLR to be provided by the network.

To address this problem at least two methods have been developed. One involves computing an "equivalent bandwidth" (EBW) that the requesting call's traffic will use [2, 4, 6, 5, 7, 10, 9]. More recently, another approach based on modeling the volatility of call traffic using Reflected Brownian Motion (RBM)[8] has been proposed.

In our study, we compare two CAC schemes in terms of network cost. Our results indicate that the selection of one from the many different schemes for CAC can be based on the resulting network cost, which can be deduced from the network topological design problem. We call the two CAC schemes, which we study here, Scheme 1 and Scheme 2.

Scheme 1 uses an EBW as the CAC criterion. Its three advantages for network topological design are

1. EBW corresponds directly to bandwidth requirement of a connection.

2. EBW is additive: we can sum up all the per-call EBWs on a link to determine the required capacity for that link.

3. EBW is the simplest CAC scheme for topological design.

Scheme 2 is based on modeling the volatility of call traffic using RBM [8] and allows the CLR for each connection calculated from the formulae derived from the

RBM model (see the Appendix for details). Though this scheme does not give a measure which corresponds to the bandwidth requirement of a connection directly, it offers the following two advantages:

1. It is more aggressive than EBW (which has been shown to be conservative [9]), in the sense that for the same link it permits more connections to be accepted by the network.

2. The RBM model parameters used in the CLR calculation are additive across connections. When a connection is rerouted, the value of the RBM model parameters of affected links can be easily modified. When a connection is added to (or removed from) a link, the values of the corresponding parameters for this connection are added to (or subtracted from) the values of the RBM model parameters for the link (see the Appendix for the additivity of the RBM parameters).

Both schemes are "additive" in some sense (EBW is additive and the RBM model parameters are additive). The reason that additivity is an advantage to network topological design is that it permits superposition of loads, which facilitates computation–addition is the simplest computation. Without additivity EBW or the RBM model parameters would have to be re-calculated when a connection is rerouted via the corresponding procedures, which are much more complicated and time-consuming.

3 ATM Network Topological Design

Given the traffic demand on a network and a cost structure of network components, *i.e.*, inter-switch optic fibers and switches, what is the most cost effective network topology? This question is answered by "network topological design".

However, traffic demand and cost structure of network components can not uniquely determine ATM network topology. In ATM networks the actual bandwidth requirement of a call depends not only on the call's traffic characteristics but also on the network's switch structure (due to the burstiness of traffic and the finite size of buffers). Therefore, the method of estimating the bandwidth requirement for a connection, *i.e.*, the CAC scheme, significantly affects network topological design and must be considered as an input to the design procedure.

Since network topological design has been shown to be NP-hard, we turn to heuristics for a solution. We will adopt the column generation based heuristic developed in [11] as our tool for network topological design.

3.1 Generic Model for ATM Network Topological Design

We define the following physical inputs:

Nodes The geographical sites where switches can be placed are called nodes. **Nodes** are typically a subset of the cities connected by the network.

Distance A distance matrix specifies the length of optic fiber necessary to connect two nodes. These distances are not assumed to be symmetric. They may not correspond to the shortest geometrical path between sites and they do not necessarily obey the triangle inequality . Distance determines the signal/data propagation delay.

Fiber Cost of the optic fiber, which is a three dimensional matrix, (node i, node j, capacity). This matrix gives the projected cost of laying fiber or leasing bandwidth of a specified capacity between two nodes during the planned lifetime of the network. This cost may include maintenance and associated expenses. Units of fiber cost can be in dollars or any other generic units as long as it is the same as the unit of switch cost.

Switch The set of switches from various vendors to be considered for deployment in the final network. The switch data is tabulated in a file and includes a record for each switch and for each switch option or configuration. The user describes each switch in an ASCII file with entries that specify:

- **price** The projected cost of installing, purchasing, or renting and maintaining this switch in a node during the lifetime of the network.
- **num_port** Number of ports. We presently have only exercised the code for cases in which the number of input ports equal to the number of output ports.

- **cap_port** Capacity of each port in bits per second. We have assumed it is the same for all the ports for the sake of simplicity, but it can differ among ports.

- **buffer** Buffer size of each output port, this is specified in cells and can also be port dependent.

Since there are many different switch structures provided by industry, the above description may not be sufficient to cover all of them. One of the difficulties in practical ATM network design is that vendors do not use readily comparable measures to specify the performance of their switches. However, the switch structure is not used as a direct input to the algorithm. It is only needed when we calculate the EBW in Scheme 1 and evaluate the CLR in Scheme 2. We developed this switch description for the CAC schemes that we have investigated. It is believed to be accurate enough for realistic network design.

Traffic Demand The traffic demand matrix is three dimensional and is indexed by origin, destination, and class. The third index, class, depends on the traffic descriptor for traffic demand and its QOS requirements. The traffic demand matrix is not symmetric in general. It is specified by an ASCII file containing three entries per record:

- **num_connection** Number of connections for each triple (origin, destination, class). The user may direct the optimizer to consider inserting switching substation by including a node with zero traffic originating from it and zero traffic terminating at it. When the number of connections is zero for all the traffic originating from and terminating at a node, that node is included only for the purpose of minimizing the cost of the final network. Industrial network designers require this feature for realistic applications.

- **traffic descriptor** The design optimizer uses the standards-defined traffic descriptors: a connection's peak cell rate, sustainable cell rate, and maximum burst length [1].

- **QOS** The design optimizer uses the standards-defined QOS parameters: the maximal acceptable CLR, and the maximal acceptable delay for each triple (origin, destination, class) [1].

Call Admission Control Scheme The method specified by the network designer to determine whether a call can be accepted or not.

Pathset(origin, destination, class) The set of all paths satisfying the end-to-end delay requirement. It is derived from the distance matrix and the maximal acceptable delay assuming that propagation delay dominates cross-network delays. (Propagation dominates only in high speed networks. In low speed networks, delay occurring at a node is traffic-dependent and can not be neglected. We can resolve this problem by associating a constant delay with a node and adding a penalty term to the objective function.)

A path is specified by a path-link vector.

$$\delta(o,d,k,path,i,j) = \begin{cases} 1, & \text{if } path \text{ for } (o,d,k) \text{ passes link } (i,j) \\ 0, & \text{otherwise} \end{cases}$$

Then the outputs will be the resulting network design which is specified by the following collection of decision variables.

- $x(o,d,k,path)$ The number of connections that will be routed on $path$ from **traffic**(o,d,k).
- $y(i,j,cap)$ The number of fibers of capacity(i,j,cap) placed on link (i,j).
- $z(i,s)$ The number of switches of s type placed on node i.

Since some CAC schemes do not give analytical formulae for the bandwidth requirement of a connection, we give a descriptive model of network topological design as follows:

Problem G:

minimize **cost of network**

subject to **traffic demand is satisfied**

where **cost of network** is determined by the sum of cost of all network components and **traffic demand is satisfied** is realized by verifying that every connection's QOS requirements are satisfied.

3.2 ATM Network Topological Design Under Scheme 1

In Scheme 1, we adopt EBW as the CAC scheme. In this scheme, the bandwidth requirement of a connection can be calculated explicitly and we have then an analytical model for network topological design.

Let ebw(o, d, k, s) be the corresponding EBW when **traffic**(o, d, k) passes through switch s. If we define

$$\text{link\_load}(i,j) = \sum_{(o,d,k)} \text{ebw}(o,d,k,s(i)) * \sum_{p} x(o,d,k,p) * \delta(o,d,k,p,i,j) \tag{1}$$

where $i \neq j$ and $s(i)$ is the switch in node i and

$$\text{link\_load}(i,i) = \sum_{(o,k)} \text{ebw}(o,i,k,s(i)) * \text{num\_connection}(o,i,k) \tag{2}$$

then we can model the problem as follows:

Problem 1:

$$\begin{aligned}
\text{minimize} \quad & \sum_{(i,j,cap)} \text{fiber_cost}(i,j,cap) * y(i,j,cap) + \sum_{(i,s)} \text{switch_cost}(i,s) * z(i,s) && (3)\\
\text{subject to} \quad & \sum_{cap} \text{capacity}(i,j,cap) * y(i,j,cap) \geq \text{link_load}(i,j), \forall i \neq j && (4)\\
& \sum_{s} \text{switch_capacity}(i,s) * z(i,s) \geq \sum_{j} \text{link_load}(i,j), \forall i && (5)\\
& \sum_{p} x(o,d,k,p) \geq \text{num_connection}(o,d,k), \forall (o,d,k) && (6)\\
& \sum_{s} z(i,s) \leq 1, \forall i && (7)
\end{aligned}$$

where $x(o, d, k, p)$, $y(i, j, cap)$ and $z(i, s)$ are integers.

The meaning of the notations in **Problem 1** is as follows:

- link_load(i, j) is the traffic, in terms of EBW, on link (i, j) and link_load(i, i) is the traffic, in terms of EBW, terminating at node i;
- fiber_cost(i, j, cap) is the cost during the planned lifetime of the network for laying a fiber with cap type capacity from node i to node j and capacity(i, j, cap) is the corresponding capacity;
- switch_cost(i, s) is the cost during the lifetime of the network for placing type s switch at node i and switch_capacity(i, s) is the corresponding capacity;
- num_connection(o, d, k) is the number of connections of traffic from node o, to node d, with class k QOS requirement.

The objective of **Problem 1** has two terms, total fiber cost and total switch cost. Constraint (4) is the link capacity constraint and correspondingly (5) is the node capacity constraint. Constraint (6) is the "flow conservation requirement" which forces every connection accommodated by the network. Since the objective is minimized, writing (6) as an inequality will not hurt in anyway but gives the flexibility to do analysis like Lagrangian relaxation. Constraint (7) limits at most one switch placed on a node.

This is a non-linear integer programming problem. The non-linearity is due to the dependence of EBW on switch structure. For a six node network with 11 classes of traffic, the number of integer variables is 390. One can see that solving a problem with realistic network size is not a trivial task.

We solve **Problem 1** via a column generation based heuristic developed by [11].

3.3 ATM Network Topological Design Under Scheme 2

In Scheme 2, we adopt RBM model to calculate CLR and hence do CAC. Since we do not have an analytical formula for connection bandwidth requirements in this case, we model the problem as follows:

Problem 2:

$$\text{minimize} \quad \sum_{(i,j,cap)} \text{fiber\_cost}(i,j,cap) * y(i,j,cap) + \sum_{(i,s)} \text{switch\_cost}(i,s) * z(i,s) \tag{8}$$

$$\text{subject to} \quad \sum_{p} x(o,d,k,p) \geq \text{num\_connection}(o,d,k), \forall(o,d,k) \tag{9}$$

$$\sum_{s} z(i,s) \leq 1, \forall i \tag{10}$$

$$\text{CLR is guaranteed for traffic}(o,d,k), \forall(o,d,k) \tag{11}$$

where $x(o,d,k,p)$, $y(i,j,cap)$, $z(i,s)$ are integers and CLR is calculated according to the Appendix.

Problem 2 is still a non-linear integer programming problem, but it is harder to solve due to the complexity of the CLR calculation. Whenever we reroute a connection, we have to recalculate the CLR of all the connections affected by the rerouting (actually, we only need to recalculate the CLRs of connections which pass the links which the rerouted connection is added to). **Problem 2** is also solved by the column generation based heuristic proposed by [11].

4 Experiments and Results

4.1 Input Data

We have developed 11 classes of traffic specifications. Many of them are derived from realistic traffic sources while others are used to test the CAC scheme which must operate for any traffic source characteristics.

Table 1–Traffic Descriptors of Traffic Demand

class	peak cell rate (Mbps)	sustainable cell rate (Mbps)	maximum burst length (cells)	delay (miles)
1	2.1	.032	150	3580
2	25.0	20.000	200	3600
3	50.0	1.000	50	3700
4	49.0	1.000	199	3800
5	10.0	.500	10	3900
6	10.0	.100	40	4000
7	40.0	.500	25	4100
8	5.0	4.000	500	4200
9	10.0	.100	400	4300
10	20.0	.100	100	4400
11	5.0	5.000	≥ 1	4500

Table 1 shows the 11 traffic class' characteristics. Since we presently consider propagation delay, the maximal delay for a connection can be specified in terms of distance. The CLR requirement for traffic is not given, it is left as a control variable for experiments.

We have 6 nodes. They are Dallas, Denver, Los Angels, San Francisco, Seattle and Washington DC. We use the direct-line distance [3] as our distance resulting in the matrix given by Table 2.

Table 2–Distance Matrix (miles)

	node 1	node 2	node 3	node 4	node 5	node 6
node 1	0	662	1240	1483	1681	1185
node 2	662	0	831	950	1021	1493
node 3	1240	831	0	347	958	2300
node 4	1483	950	347	0	678	2441
node 5	1681	1021	958	678	0	2328
node 6	1185	1493	2300	2441	2328	0

In this particular design case, we have only one fiber capacity, 120 Mbps. The cost of fiber for link (i, j) is computed from the formula:

$$fiber\_cost(i,j) = \begin{cases} a(i,j) + b(i,j) * k(i,j) & \text{if } k(i,j) > 0 \\ 0 & \text{otherwise} \end{cases}$$

where $a(i,j)$ is the initial cost for laying fiber on link (i,j) and $b(i,j)$ is the additional cost for laying one 120 Mbps fiber on link (i,j) and $k(i,j)$ is the number of fibers on link (i,j). The unit of cost is some generic unit and it is the same for switch cost. The $a(i,j)$ and $b(i,j)$ are given by Table 3, which happens to be symmetric in our experiment, $a(i,j) = a(j,i)$ and $b(i,j) = b(j,i)$. However they are not required to be symmetric.

Table 3–Costs for Fiber $(a(i,j), b(i,j))$

	node 1	node 2	node 3	node 4	node 5	node 6
node 1	(0,0)	(700,700)	(1300,130)	(1500,150)	(1700,170)	(1200,120)
node 2	(700,70)	(0,0)	(900,90)	(1000, 100)	(1100, 110)	(1500,150)
node 3	(1300,130)	(900,90)	(0,0)	(400,40)	(1000,100)	(2300,230)
node 4	(1500,150)	(1000,100)	(400,40)	(0,0)	(700,70)	(2500,250)
node 5	(1700,170)	(1100,110)	(1000,100)	(700,70)	(0,0)	(2400,240)
node 6	(1200,120)	(1500,150)	(2300,230)	(2500,250)	(2400,240)	(0,0)

Thus, for the link $(1,3)$ (i.e., Dallas–Los Angels) $a(1,3) = 1300$ and $b(1,3) = 130$.

For this design example, we have 5 types of switches. Each port capacity is 120 Mbps which roughly corresponds to an OC-3 and they all have output buffers of 100 cells. Switch costs are given by Table 4.

Table 4–Switch Capacity and Cost

type	number of ports	cost
1	4	200
2	8	360
3	16	450
4	32	700
5	64	1250

4.2 Optimization Results

The experimental setup is as follows: the total number of connections between an origin and a destination is selected randomly between 0 and 60. In addition, we compared three different traffic mixes. In **mix 1**, class 1 comprises 20% of the connections. Each of the other 10 classes comprise 8%. In **mix 2**, class 6 comprises 20% of the connections and each of the other 10 classes comprise 8%. Since both **mix 1** and **mix 2** are dominated by bursty traffic (see Table 4), to investigate a less bursty expected loading, we use another traffic mix, **mix 3**, in which traffic classes are uniformly distributed among all classes.

Since we have relatively small buffers (100 cells), according to results in [9], we can use peak cell rate as each connection's EBW.

The results of the simulated comparison of network design are presented in Table 5, which tabulates the ratio of network cost for Scheme 1 compared to Scheme 2.

Table 5–Ratio of the Cost of the Optimal Network

Using CAC Scheme 1 to Scheme 2

cell loss ratio	mix 1	mix 2	mix 3
1.0e-5	1.35	1.30	1.23
3.0e-8	1.26	1.26	1.12
1.0e-10	1.25	1.26	1.10

To illustrate the difference between CAC schemes 1 and 2, the resulting network topology under schemes 1 and 2 (for **mix 1** with CLR=1.0e-5) are presented in Figures 1 and 2 respectively.

The connectivity is not symmetric and is represented by double pointed arrows with fiber capacities given for each direction.

The capacity of each link is indicated by the number on the link so that 10 units of fiber (a unit is 120Mbps in this example) are necessary from 2 to 5 in Figure 1, but only 4 units are required in Figure 2. The type of switch at each node is given in the order of nodes. (In Figure 1, a switch of type 4 (see Table 3) must be assigned to node 1 to satisfy the QOS for the specified traffic demand. In Figure 2, only a switch of type 3 is necessary at node 1.)

5 Conclusions

From the experimental results, we can draw the following conclusions:

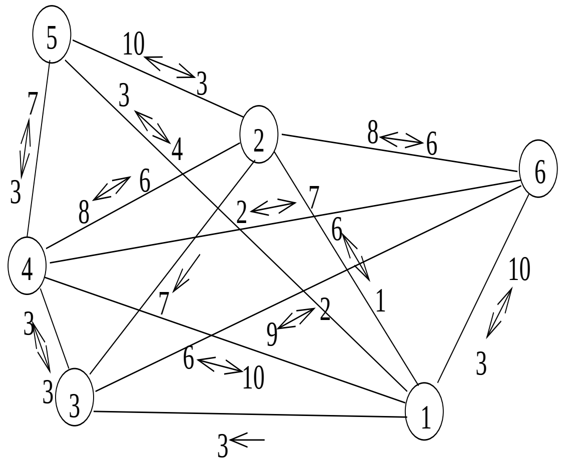

Figure 1   Result from Scheme 1 with Type 4,4,3,4,3,4 Switch in Node 1,2,3,4,5,6

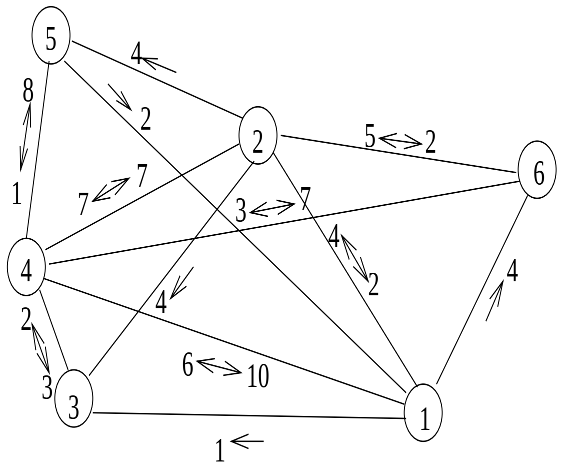

Figure 2   Result from Scheme 2 with Type 3,4,2,4,3,3 Switch in Node 1,2,3,4,5,6

- For network topological design,
  - CAC has to be considered in the network topological design process. When the network designer adopts different CAC schemes, the cost of the network can be dramatically different (as little as 10% and as much as 35% in our example).
  - Network topological design can be used not only for determining network topology, it can also be used as a tool to evaluate other aspects of network architecture, such as the CAC schemes in this study.
- For CAC scheme selection, we can see that network topological design can serve as one of the criteria for selecting CAC scheme. For example, in our experiment, Scheme 1 results in higher cost network topology in general which is consistent with the conservatism of EBW, and Scheme 2 is especially well suited to an expected traffic mix in which high bursty traffic dominates (**mix 1, mix 2**). Scheme 2 is favored compared to Scheme 1.

References

[1] CCITT. Recommendation I.371, 1992.

[2] A.I. Elwalid and D. Mitra. Effective bandwidth of general Markovian traffic sources and admission control of high speed networks. *IEEE/ACM Transactions on Networking*, 1:329–343, 1993.

[3] G.L. Fitzpatrick and M.J. Modlin. *Direct-Line Distances.* Scarecrow Press, 1986.

[4] R.J. Gibbens and P.J. Hunt. Effective bandwidth for the multi-type uas channel. *Queuing Systems*, 9:17–27, 1991.

[5] R. Guerin, H. Ahmadi, and M. Naghshineh. Equivalent capacity and its application to bandwidth allocation in high-speed networks. *IEEE Journal on Selected Areas in Communications*, 9:968–981, 1991.

[6] F.P. Kelly. Effective bandwidth at multi-class queues. *Queuing Systems*, 9:5–15, 1991.

[7] V.G. Kulkarni, L. Gun, and P.F. Chimento. Effective bandwidth vectors for multiclass traffic multiplexed in a partitioned buffer. *IEEE Journal on Selected Areas in Communications*, 13:1039–1047, 1995.

[8] B.A. Makrucki. A connection admission control for variable bit rate connections in ATM. *Performance Evaluation Special Issue on Traffic Control in ATM Networks.* to appear.

[9] K.M. Rege. Equivalent bandwidth and related admission criteria for ATM systems–a performance study. *International Journal of Communication Systems*, 7:181–197, 1994.

[10] K. Sriram. Methodologies for bandwidth allocation, transmission, scheduling, and congestion avoidance in broadband ATM networks. *Computer Networks and ISDN Systems*, 26:43–59, 1993.

[11] S. Zhang, G.L. Bilbro, and S.E. Elmaghraby. ATM network topological design: heuristic and lower bound. In *4th International Conference on Telecommunication Systems: Modelling and Analysis*, pages 275–284, Nashville, TN, March 1996. also available as Technical Report from Center of Advanced Computing and Communication, North Carolina State University, TR 95-15.

Appendix: The Reflected Brownian Motion (RBM) Model

A complete description of the RBM model of buffer dynamics is beyond the scope of this paper but is summarized in the following paragraphs.

A traffic source i specifies as its descriptor its *peak cell rate* $\hat{\lambda}_i$; its *sustainable cell rate* $\bar{\lambda}_i$; and its *maximum burst length* (at the peak rate) $\hat{B}_i$. The RBM model transforms these quantities into model parameters that indicate *traffic variability* and *steady-state behavior*. The variability parameter is termed α_i.

In the RBM model [8] network buffer lengths are modeled using Reflected Brownian Motion (RBM), from which cell loss ratios (and probabilities) are derived. Briefly, the model is constructed as follows: For each traffic source, model parameters are calculated, call them α_i and $\bar{\lambda}_i$ for source i (α_i is indicative of the "burstiness" of source i's traffic, $\bar{\lambda}_i$ is the source's sustainable cell/packet rate and is actually declared by the source at connection request). As a result, the RBM model parameters are additive across connections so that the RBM is characterized by the tuple $(\alpha = \sum_i \alpha_i, \beta = \sum_i \bar{\lambda}_i - \mu)$ where μ is the link rate at which a buffer is drained (for the buffer where the QOS is being calculated).

From the RBM model and parameters calculated above, each connection's QOS can be determined. When a new connection arrives and requests admission to the network, the network must check that it and all existing connections will receive their requested QOSs. This is done via detailed calculations based on the RBM model for buffer sizes.

The element of the model that must be reflected in the topological design work here is that the RBM model parameters are additive across connections and connection CLRs can be approximated with the following formulae [8]:

$$\alpha = \sum_i \alpha_i \tag{12}$$

$$\beta = \sum_i \bar{\lambda}_i - \mu \tag{13}$$

$$C = \text{buffer size} \tag{14}$$

$$\hat{B}_i = \text{burst length at the peak rate } \hat{\lambda}_i \text{ for connection } i \tag{15}$$

$\hat{B}_i$ is available from the connection's declared traffic descriptor. Then

$$CLR_i \approx \frac{\sum_{K=0}^{\hat{B}_i - 1}(1 - G^{\infty}(C-1, \alpha_t(i), \beta_t(i)))}{\hat{B}_i}, \text{ where } t = \frac{K}{\hat{\lambda}_i} \tag{16}$$

with

$$G^{\infty}(x, \alpha, \beta) = \frac{1 - e^{2\frac{\beta}{\alpha}x}}{1 - e^{2\frac{\beta}{\alpha}C}} \tag{17}$$

and $\alpha_t(i)$ and $\beta_t(i)$ are RBM parameters calculated at a sequence of times that correspond to cell arrivals from source i when it emits a burst of traffic. They are as follows:

$$\alpha_t(i) = \alpha - (1 - f(t))\alpha_i \tag{18}$$

$$\beta_t(i) = (1 - f(t))(\beta + \mu - \bar{\lambda}_i + \hat{\lambda}_i) + f(t)\sum_i \lambda_i - \mu \tag{19}$$

and

$$f(t) = e^{-\frac{1}{2}(\frac{\beta^2}{\alpha}+\alpha\frac{\pi^2}{C^2})t} \tag{20}$$

$f(t)$ is derived from transient analysis of RBM and represents the "response time" of the buffer in changing from a steady-state (represented by source i being in a background state of activity) to a system in which source i is transmitting a burst.

The procedure of calculating CLR can be summarized as follows:

Input:

1. *peak cell rate* $\hat{\lambda}_i$, *sustainable cell rate* $\bar{\lambda}_i$, and *maximum burst length* (at the peak rate) $\hat{B}_i$ for source i;
2. link capacity μ and buffer size C.

Calculate:

1. α_i for source i (see [8] for details);
2. α according to (12) and β according to (13);
3. $\alpha_t(i)$ according to (18) and $\beta_t(i)$ according to (19) for source i;
4. CLR_i according to (16).

Output: CLR_i

CLR_i is the approximation of the CLR for source i and it is used by Scheme 2 to examine the fulfillment of the CLR requirement for source i.