Distinguished Speaker:
Dr. James S. Meditch

IC^{3}N

Proceedings of the
Third International Conference on
Computer Communications and Networks

September 11-14, 1994
San Francisco, California, USA

Sponsored by

USC Communication Sciences Institute

IEEE Communications Society

Elsevier Science

Conference Chair: T. Suda
Technical Chair: J.W. Wong
Tech. Vice Chair: E.K. Park
Publication Chair: W. Liu (editor)

Third International Conference on Computer Communications and Networks (IC^3N)

Sponsored by USC Communication Sciences Institute
In cooperation with IEEE Communication Society
With additional support by Elsevier Science

KEYNOTE ADDRESS

Professor James S. Meditch
University of Washington
"The National Information Superhighway: Fact, Fiction or Folly?"

Conference Chair
Prof. T. Suda
Detp. of Info. Computer Science
Univ. of California at Irvine
Irvine, CA 92717

Technical Chairs
Prof. J.W. Wong (Chair)
University of Waterloo
Waterloo, Ontario N2L 3G
Canada

Dr. E.K. Park (Vice Chair)
United States Naval Academy
Annapolis, MD 21402-5002

Steering Committee
- V.O.K. Li (Chair)
- J.F. Chang
- B.G. Lee
- Y.C. Lee
- H. Okada
- G.S. Poo
- R. Pickholtz
- I. Rubin

TECHNICAL PROGRAM COMMITTEE

J.W. Wong	University of Waterloo (Chair)
E.K. Park	United States Naval Academy (Vice Chair)
M.H. Ammar	Georgia Institute of Technology
M. Chen	IBM T.J. Watson Research Center
D. Dykeman	IBM Zurich Research Laboratory
N.D. Georganas	University of Ottawa
J.J. Harms	University of Alberta
H.S. Hassanein	Kuwait University
I. Khan	SRI International
J. Kim	California State University at Northridge
V.O.K. Li	University of Southern California
J.W. Mark	University of Waterloo
G. Neufeld	University of British Columbia
B. Mukherjee	University of California at Davis
T. Suda	University of California at Irvine
P. Tran-Gia	University of Wuerzburg
P. Yegani	IBM Networking Systems

Local Arrangement Chair
Dr. I. Khan
SRI International

Registration Chair
Dr. K. Makki
Univ. of Nevada

Publicity Chair
Dr. N. Pissinou
Univ. S.W. Louisiana

Publication Chair
Dr. W. Liu
U.M.B.C.

Performance of Hop-by-Hop Distributed Routing and Resource Allocation in an ATM Network

R. Bolla, F. Davoli, M. Marchese

Department of Communications, Computer and Systems Science (DIST)
University of Genoa
Via Opera Pia 11/A, I-16145 Genova, Italy

Abstract.

A node-by-node admission control and routing strategy is described, which is based on the subdivision of traffic into a number of service classes, characterized by different performance requirements. At each network node, for all outgoing links, link capacity partitions are periodically assigned to the service classes, as the result of an optimization problem over a fixed time interval, and local access control rules compute the maximum number of connections of each class that a link can accept, within the assigned capacity. Incoming call connection requests are forwarded in hop-by-hop fashion. Each node traversed first checks the presence of the resources needed to accept a new connection and guarantee all Quality of Service (QoS) requirements, by using the local access control rule; then, it chooses the next node along the path on the basis of a distributed routing strategy, which minimizes a cost function accounting for local instantaneous information, as well as for aggregate information that is passed along periodically among adjacent nodes. Simulation results are presented, which also show a comparison of the proposed routing strategy with two other solutions, based on centralized and completely decentralized information, respectively.

Keywords: ATM Traffic Control, Quality of Service, Admission Control, Dynamic Routing

1. Introduction

One of the main problems that arise from the statistical multiplexing nature of the Asynchronous Transfer Mode (ATM) technique is that of guaranteeing Quality of Service (QoS) requirements in the presence of traffic flows that exhibit different statistical and performance characteristics. In this respect, a larger amount of control has to be exerted at the network boundaries (and within the network itself) than that required in other transfer modes (e.g., STM), where the allocation of bandwidth is somehow more "structured".

This specific aspect of ATM has given rise to a large amount of investigations in bandwidth allocation ([1-3], among others), admission and congestion control [4, 13], and (though to a lesser extent) routing [14]. Among the possible approaches, control architectures that impose a certain structure on the allocation of the resources have been considered, where typically traffic is subdivided into classes, which are homogeneous in terms of statistical or performance characteristics [11, 15-17, 18-20]. This often allows the decomposition of a very complex overall control task, which is in general characterized by very different time scales and requirements, according to the level where the system dynamics is considered (e.g., cell and call level), into smaller and somehow independent problems. For instance, an essential decoupling between cell and call level is achieved in [17] through the concept of schedulable region, whereas a hierarchical decomposition has been used by the authors in previous works [18-20], by adopting an approach already introduced in the context of TDM systems [21-22].

In this paper, we further investigate the integrated admission control and routing strategy previously introduced in [20], which is based on the above mentioned philosophy. More specifically, at each ATM network node, several traffic classes share the outgoing links. Each traffic class is characterized by statistical parameters (like peak and average bandwidth), as well as by QoS requirements at the cell level (in terms of cell loss probability and cell delay). On each link, each traffic class is dedicated a separate call admission controller, which applies a "local" fixed strategy, designed to maintain the required QoS, given the buffer space and bandwidth (percentage of cells) assigned to the class on the link. The bandwidth shares are periodically recomputed on-line by a bandwidth allocation controller that plays the role of a coordinator in a hierarchical dynamic control scheme.

In this context, in order to check if all the nodes to be traversed from source to destination have enough resources to serve an incoming connection, while still guaranteeing QoS to the connections in progress, and possibly optimize the usage of network resources, we have defined a combined routing and admission control scheme. The routing algorithm is implemented in hop-by-hop fashion at call set-up time: a call request is forwarded from node to node, each time dynamically choosing the next hop; upon traversing a node, after checking the availability of the resources for admission,

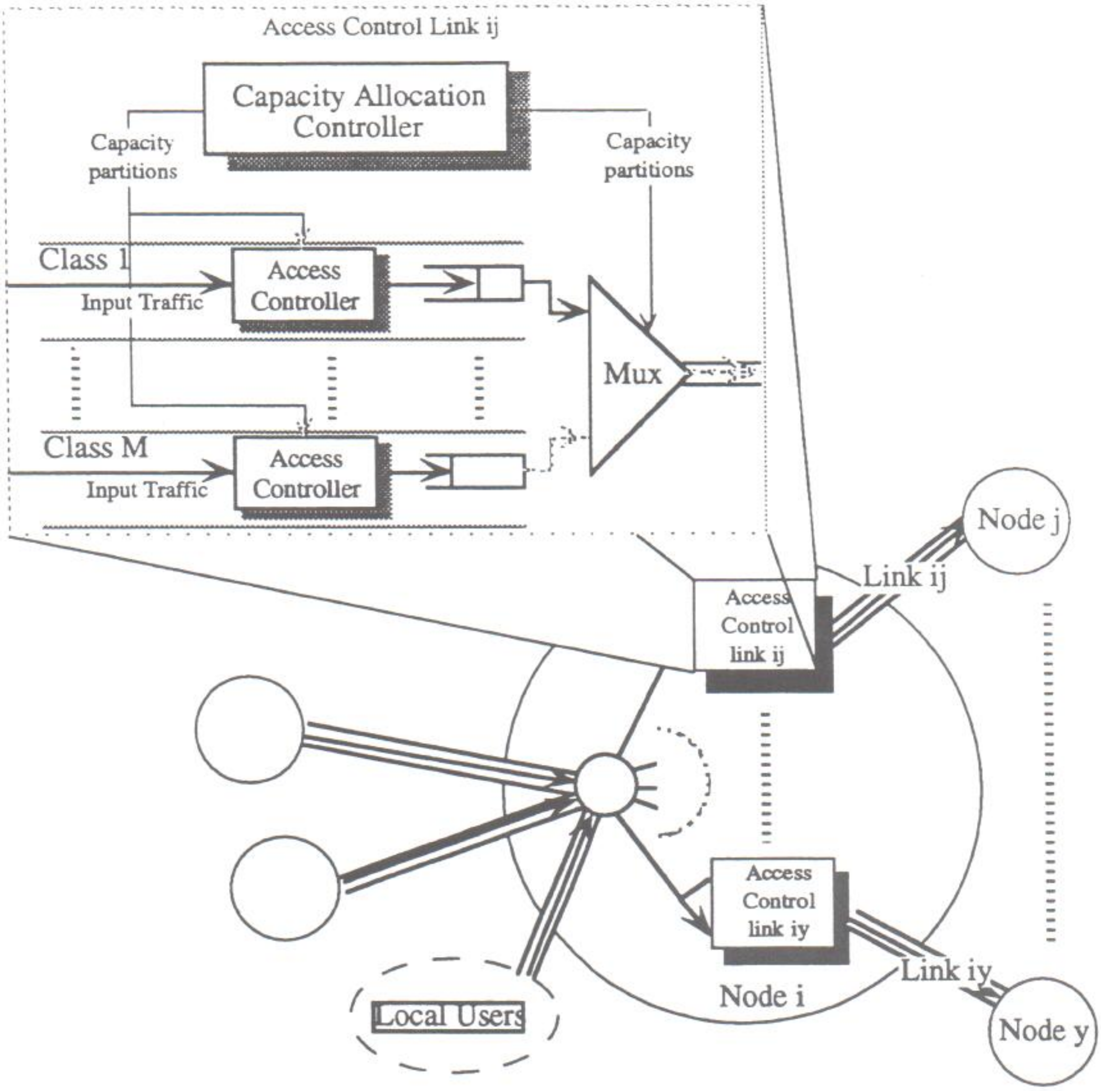

Figure 1 Structure of a node and overall control architecture of a link.

an outgoing link is chosen according to the routing strategy, the request is forwarded to the next node on the path and the same action is repeated; if the request is refused or the destination node has been reached, the event is notified to the source node. Globally, a connection is accepted in the network and the corresponding Virtual Circuit (VC) is created if every node up to the destination accepts the new connection request.

Routing decisions are taken on the basis of a distributed strategy, whereby every node decides the output link which the connection request packet must be sent to. The output link is chosen by minimizing a cost function composed by two terms: a "local" and a "global" one. The local term takes into account the situation of the node by using a simple function, whose value depends on the number of connections that every link could accept at the moment the decision is made. The global term holds aggregate information about the situation of the network, and it is updated periodically by every node and passed along to its neighbours.

The paper is organized as follows. In the next Section, we briefly recall the structure of the ATM node and the admission control and bandwidth allocation strategies that have been introduced and analyzed in [19]. The overall admission control and routing scheme is described in Section 3. Section 4 reports and discusses several simulation results; in particular, the present routing strategy is compared with two "extreme" situations, namely, centralized shortest path and decentralized hot-potato routing. Section 5 contains the conclusions and directions for further research.

2. Admission control and bandwidth allocation

We suppose the traffic on the network to be divided into M classes, each one characterized by statistical parameters like peak and average transmission rate, as well as by QoS requirements, like cell loss probability and cell delay.

The general structure of a node is depicted in Fig. 1, where incoming links from other nodes, a local traffic input link and outgoing links (over which the incoming traffic is partitioned) are shown. A resource allocation scheme and a local control access rule of the type presented in [19] are implemented for every outgoing link, and their structure, also shown in the figure, will be briefly summarized in this Section.

Each traffic class is assigned a separate buffer, whose output is statistically multiplexed on the outgoing link by a scheduler, which substantially divides the global channel capacity C_T among the classes, according to proportions that last over a certain time period. Connection requests, which can come from the users directly connected to the node or from other nodes (by means of special call request packets) are also processed on a class basis. Two controls are exerted on the system, by using a two level hierarchical control scheme, one acting on the scheduler and the other on the admission of the connections.

We pose our problem in discrete time, where the duration of a cell transmission corresponds to that of a slot and represents the time unit. At the higher level of the hierarchical control scheme, a bandwidth allocation controller periodically reassigns capacity partitions to every class. The scheduler receives the values of the partitions and must assure that every buffer is assigned a percentage of slots equal to the ratio between the total capacity and the capacity assigned to its class. The new capacity partitions $V_m^{(h)}$, $h = 1, \ldots, M$, are computed by the controller at discrete time instants $m = 0, K, 2K, \ldots$ (K is the length of the intervention period in slots), based on the minimization of a cost function that takes into account the overall cell loss probability.

At the lower level, M access controllers decide about the acceptance of a connection request, independently for each class. By assuming the quasi-stationarity of the connections activity process, the packet loss rate and the rate of cells exceeding a delay requirement for the connections of one class over an outgoing link can be computed [19]. Such assumption can be considered acceptable whenever the buffer dynamics is characterized by a cell-scale behaviour (which happens if the buffer dimension $Q^{(h)}$ is much smaller than the average burst length; this is assumed throughout). The acceptance decisions are taken on the basis of the two above mentioned quantities, and therefore depend on the capacity currently allocated to the class, on the current number of connections in progress, and on the statistical and performance characteristics of the specific traffic. As regards the latter, class h traffic is supposed to be made up by bursty connections with identical and independent statistical characteristics. We indicate with $b^{(h)}$, $B^{(h)}$, and $P^{(h)}$ the burstiness, the average burst length, and the peak bit rate, respectively, of the h-th class. Connected sources are supposed to be of the on-off type, and are modeled simply as a two state (idle and active, respectively) Markov chain. When a connection is in the active state, it can generate a cell per slot with a certain probability (equal to $P^{(h)}/C_T$) , and it does not generate cells when it is in the idle state.

After every access controller receives its capacity assignment from the bandwidth allocation controller, it computes the maximum number of connections of the h-th class that can be supported on link ij as

$$N_{ij}^{(h)}(m) = \min\left\{N_{ij,L}^{(h)}(m); N_{ij,D}^{(h)}(m)\right\} \qquad (1)$$

where $N_{ij,L}^{(h)}(m)$ is the maximum number of connections on the link capable of maintaining the cell loss rate below a given upper bound $\varepsilon^{(h)}$, and $N_{ij,D}^{(h)}(m)$ is the maximum number of connections computed by imposing a similar limit on cell delay, namely, that the probability of exceeding a delay of $D^{(h)}$ slots be less than an upper bound $\delta^{(h)}$ (the details of this computation can be found in [19]). Summing up, the local acceptance rule is: a new connection of the h-th class arriving at time slot k, $m+\Delta \le k \le m+\Delta+K-1$ (where Δ indicates the number of slots required for the computations), can be accepted on link ij if

$$N_{ij,A}^{(h)}(k) + 1 \le N_{ij}^{(h)}(m) \qquad (2)$$

where $N_{ij,A}^{(h)}(k)$ is the number of connections of the h-th class in progress (i.e., previously accepted on the link and not terminated) at time slot k.

The aim of the allocation controller is that of balancing the bandwidth distribution among the classes, by setting the values $V_m^{(h)}$ at the beginning of a new K-slot interval, on the basis of a performance measure, which should be capable of reflecting changes in the offered load of all traffic classes. We intend the offered load of a traffic class in a certain period of time as the total number of service requests, including both accepted and blocked ones. In this respect, we have chosen to use a cost function which reflects the steady-state value of the cell loss rate for the h-th class, due to the connections present in the system, as well as the additional loss that would have been incurred if all calls presented in the previous interval had been accepted. An equality constraint (the sum of the assigned capacity must be equal to the total capacity of the link) and a set of inequality constraints, which assure service quality for the connections already in progress, must be taken into account in the minimization procedure. The latter can be treated as a mathematical programming problem, which is solved by using a gradient projection method.

It is worth noting that all computations involved, i.e., the calculation of the capacity partitions and of $N_{ij,L}^{(h)}(m)$, $N_{ij,D}^{(h)}(m)$, are performed only at the beginning of a decision interval, whose duration should be quite long compared to the slot, because the cost function reflects the dynamics of the connection requests, which take place on a much longer time scale. Thus, if the number of time slots Δ to perform the computations after the reallocation instants represents a small fraction of K, we can avoid the use of special approximations to speed up the computations. On the other hand, the acceptance rule is very simple and then it can react very fast, during the K-slot interval, to the connection requests.

3. The DLCP routing scheme

Let us suppose, for simplicity, that a route has already been found; in this case a scheme to verify the availability of this route is the node-by-node control suggested in [13]: a special call request packet is forwarded from node to node, which is just a messenger asking whether a connection can be accepted or not. The packet runs through every node of the route, undergoing the access control rule at each hop. Each node must take into account the resources it has allocated; if the call is accepted, the messenger runs along the route, otherwise, the packet stops, the call is rejected and a message is sent back, through the same route, so the intermediate nodes can free the allocated resources.

Node-by-node control and routing can be joined and managed together. In our approach the best route is not chosen beforehand and then verified, as we have supposed above for simplicity, but, at each node, we choose the "best" outgoing channel among the available ones by means of a cost function associated with each link. The cost function should take into account the link's local traffic and the traffic associated with the subsequent hops to the destination, and it should be considered only for those links that are not congested (in the sense defined by the admission control rule).

A bit stream in the messenger packet is used to remember the nodes traversed along the route. At each node, a table look-up is performed to find out congested channels, the link with the lowest cost function among the non-congested ones is chosen, and the packet is sent along that route. If every channel is congested at an intermediate node, the connection is rejected and a message is sent back through the same route, so resources can be released; otherwise, if the messenger reaches its destination, the connection is accepted, a VC is established, and the path that the cells of that connection will follow is fixed.

At instant $\bar{k}$ (in slots), a generic node i chooses the link to which to forward a call request packet generated by a class h connection request, by minimizing (over all successor nodes j) the quantity

$$c_{ij}^{(h)} = c_{ij,L}^{(h)}(\bar{k}) + c_j^{(h)}(\bar{s}) \tag{3}$$

where $c_{ij,L}^{(h)}(\bar{k})$ is a local cost related to link ij and $c_j^{(h)}(\bar{s})$ is an aggregate cost referring to the traffic conditions of node j and its successor at instant $\bar{s} < \bar{k}$ (in slots). $c_{ij,L}^{(h)}(\bar{k})$ should weigh the local congestion of link ij. A possible expression, which was adopted in the simulations in [20] and in the ones reported here, can be

$$c_{ij,L}^{(h)}(\bar{k}) =$$

$$= \begin{cases} \dfrac{1}{N_{ij}^{(h)}(\bar{m}) - N_{ij,A}^{(h)}(\bar{k})} & \text{if } N_{ij}^{(h)}(\bar{m}) > N_{ij,A}^{(h)}(\bar{k}) \\ Z & \text{if } N_{ij}^{(h)}(\bar{m}) = N_{ij,A}^{(h)}(\bar{k}) \end{cases} \tag{4}$$

where $\bar{m} \le \bar{k} < \bar{m}+K$, i.e., $\bar{m}$ is the reallocation instant when the access control rule parameters, active at the time of the request packet's arrival, have been recomputed. By using this type of function, we have that the link cost increases with decreasing available space on the link, expressed in number of acceptable connections. When the link is saturated, the cost value is Z, which should be high enough to ensure that no saturated link will be chosen if non-congested links are available.

The cost referred to a generic node i is composed by two terms

$$c_i^{(h)}(\bar{s}) = \alpha_i c_{i,L}^{(h)}(\bar{s}) + \beta_i c_{i,A}^{(h)}(\bar{s}) \tag{5}$$

where α_i and β_i are two weighting coefficients. $c_{i,L}^{(h)}(\bar{s})$ represents the average situation of the node with respect to its congestion state, and $c_{i,A}^{(h)}(\bar{s})$ is an aggregate information on the average congestion of its successor nodes. More specifically, we have defined

$$c_{i,L}^{(h)}(\bar{s}) = \frac{1}{L_i} \sum_{j \in Succ(i)} c_{ij,L}^{(h)}(\bar{s}) \tag{6}$$

$$c_{i,A}^{(h)}(\bar{s}) = \frac{1}{L_i} \sum_{j \in Succ(i)} c_j^{(h)}(\bar{s}) \tag{7}$$

being Succ(i) the set of nodes that are successors of the node i. As can be seen, the values related to the successor nodes are referred to the instants $\bar{s}$, where $\bar{s}$=T, 2T, ..., with T equal to a fixed number of slots. This means that each node i sends its costs $c_i^{(h)}(\bar{s})$, h=1,...,M, to its predecessors every T slots and then, after receiving the costs from its successors, recomputes its new aggregate information on the congestion of the network. The passage of the parameters is "one step", in the sense that the cost sent back is computed before receiving the new values from the successor nodes.

It is worth noting that the proposed strategies are distributed, based on a mix of local real time (dynamic) and overall delayed aggregate information, and do not require the presence of a real time supervisory controller, which would be questionable in a wide area network. Due to these characteristics, we term the overall routing scheme Distributed Least Congested Path (DLCP).

4. Simulation results

In this Section, we report the results of several simulations that have been performed on a simple six-node test network, in order to obtain some indication on the performance of the proposed routing scheme, and to compare it with other possible solutions. The following data has been used:

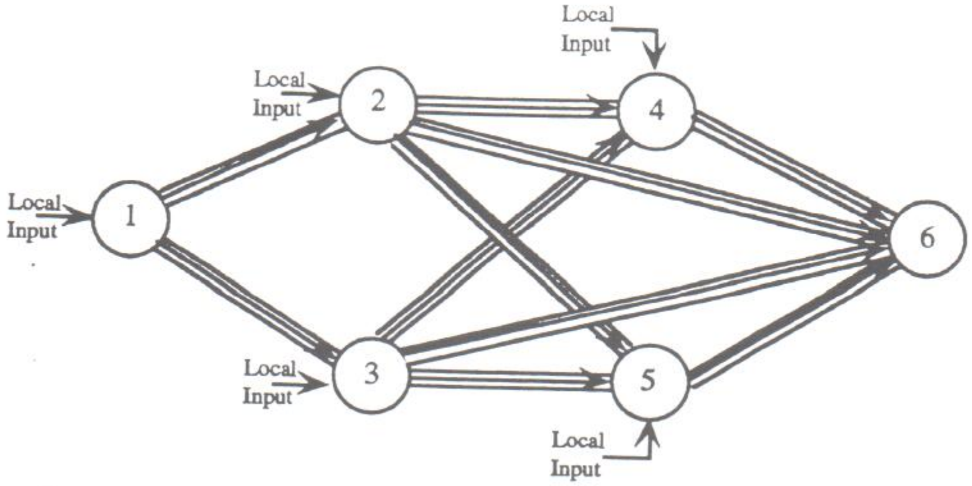

Figure 2 Topology of the test network.

$C_T = 150$ Mbit/s; $M = 3$; $K = 2\cdot 10^7$ cells

T_S = slot duration = $2.83 \cdot 10^{-6}$ s (53 bytes/cell)

$P^{(1)} = 384$ kbit/s; $P^{(2)} = 1$ Mbit/s; $P^{(3)} = 10$ Mbit/s

$b^{(1)} = 2$; $b^{(2)} = 5$; $b^{(3)} = 10$

$B^{(1)} = 100$; $B^{(2)} = 1000$; $B^{(3)} = 1000$ cells (average burst length)

$1/\mu^{(1)} = 15$ s; $1/\mu^{(2)} = 8$ s; $1/\mu^{(3)} = 10$ s (average connection duration)

$\varepsilon^{(1)} = \varepsilon^{(2)} = \varepsilon^{(3)} = 1\cdot 10^{-4}$

$\delta^{(1)} = \delta^{(2)} = \delta^{(3)} = 1\cdot 10^{-3}$

$D^{(1)} = 10$; $D^{(2)} = 70$; $D^{(3)} = 1000$ slots

$N_a^{(1)}=200$; $N_a^{(2)}=200$; $N_a^{(3)}= 40$ Erlangs

(average traffic intensities; the call arrival processes follow independent Poisson distributions)

$Q^{(1)} = 11$; $Q^{(2)} = 12$; $Q^{(3)} = 12$ cells

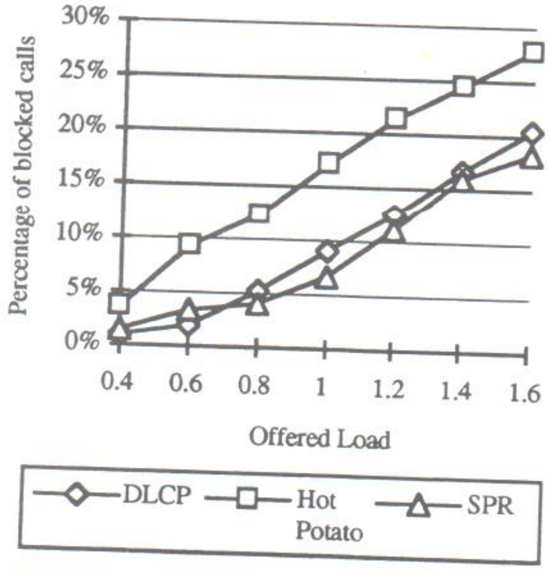

Figure 3. Total percentage of blocked calls versus the offered load.

We refer to the traffic flow generated by the above data as an offered load 1; an offered load "x" corresponds to the same data, except for the traffic intensities $N_a^{(h)}$, h=1, 2, 3, which are multiplied by x.

The topology of the network that has been used in the simulations is shown in Fig. 2. and is composed of six nodes, only one of which (node 6) is a destination node.

The behaviour of the access control and bandwidth allocation procedure at a node was tested extensively in [19], and some initial simulation results on the performance of the routing scheme are reported in [20]. Our aim in the present work is to further assess the performance of the overall technique, and to compare it with two other solutions, which may be regarded as two extreme situations with respect to the use of information for routing purposes, namely, a centralized shortest path and a totally decentralized "hot-potato" strategy.

All the simulations have a duration of 988.8 s, corresponding to 12 reallocation intervals.

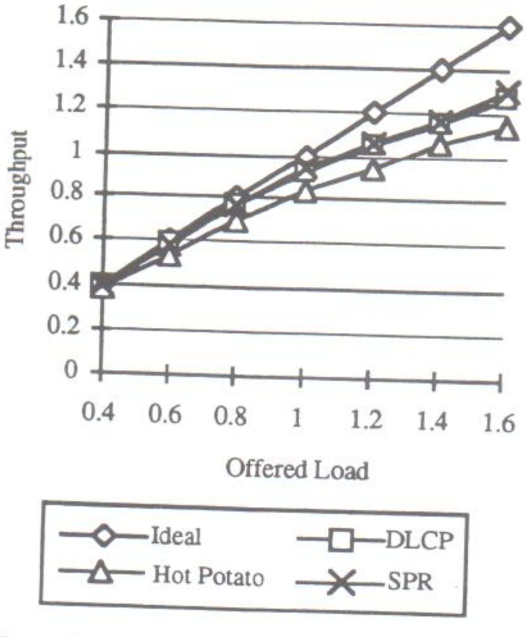

Figure 4. Total network throughput versus the offered load.

Fig. 3 shows the percentage of blocked connections for the DLCP routing scheme versus the offered load, by choosing the parameter values $\alpha_i=\alpha=0.7$, $\forall i$, and T=K/100. The results obtained are compared with a centralized Shortest Path Routing (SPR) strategy, where the cost of each link is the same as in (4), and with a local Hot Potato strategy, which is considered as a possible lower bound on performance. The percentage of blocked calls for DLCP is quite close to that of SPR.

The effect can be better appreciated from Fig. 4, where the normalized network throughput is plotted as a function of the offered load, showing an increase close to the ideal throughput characteristic up to the value 1.0 and a behaviour very similar to the SPR throughput.

Fig. 5 shows the percentage of blocked connections versus the weighting coefficient α; the DLCP routing with T=K/100 is compared with the same routing strategy for T=K; the graph shows the growing importance of the aggregate cost when the value T decreases.

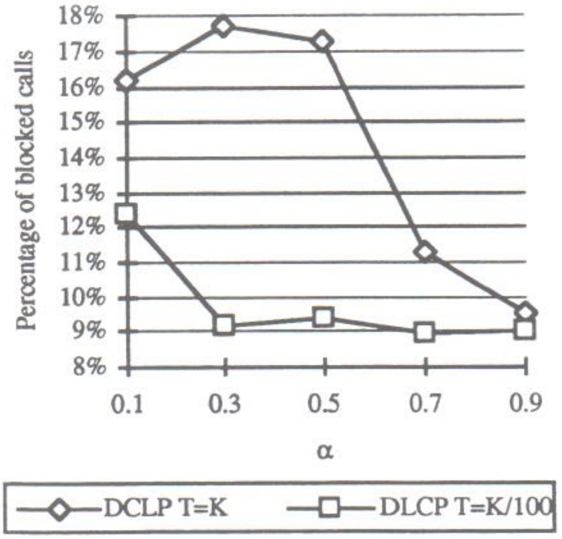

Figure 5. Total percentage of blocked calls versus the weighting coefficient α.

The percentage of blocked calls for the DLCP routing with $\alpha=0.7$ and T=K/100 is compared with the DLCP strategy with $\alpha=0.5$ and T=K in Fig. 6. The noticeable difference stresses the significance of a right choice both of the weighting coefficient α and of the updating interval T. The same meaning has the graph in Fig. 7, where the network throughput for the DLCP routing with $\alpha=0.7$, T=K/100 and $\alpha=0.5$, T=K and the ideal throughput are plotted versus the offered load.

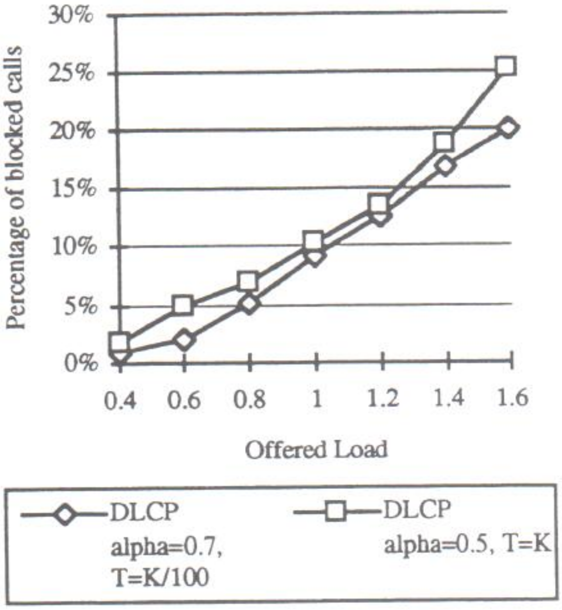

Figure 6. Total percentage of blocked calls versus the offered load.

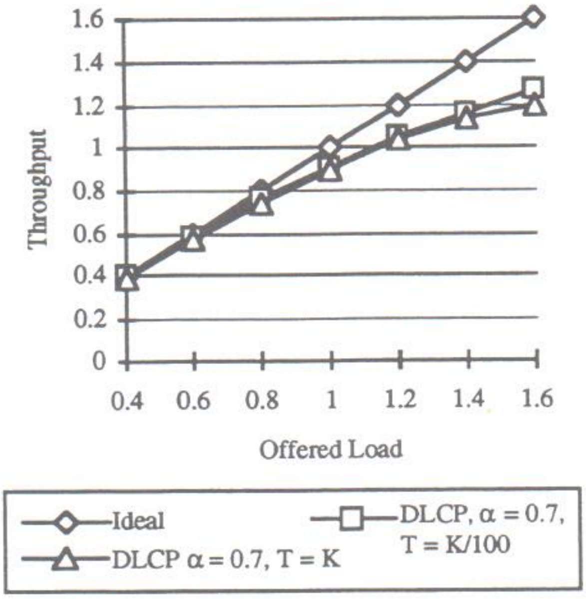

Figure 7. Total network throughput versus the offered load.

5. Conclusions

A global control architecture for access control, bandwidth allocation and routing has been considered, in an ATM network environment. The traffic is organized into service classes, characterized by specific performance requirements, and bandwidth partitions are dynamically allocated among them by controllers assigned specifically to each link. Access control and routing are performed separately for each class on a hop-by-hop basis: the former is exerted by local controllers for each link, whereas the latter stems from a distributed procedure based on local (real time) as well as on aggregated (delayed) information. The routing strategy has been explicitly defined in the paper. Simulation results have been reported and compared with other routing strategies like local Hot Potato and centralized Shortest Path Routing. The results highlight a low call rejection rate as an overall effect of the control structure over a large range of network load values.

Acknowledgments

This work was partially supported by the Italian Ministry of University and Scientific and Technological Research (MURST), within the coordinated program "Integration and Management of Multimedia Services in ATM Networks".

References

[1] J.Y.Hui, "Resource allocation for broadband networks", *IEEE J. Select. Areas Commun.*, vol. 6, n. 9, pp. 1598-1608, Dec. 1988.

[2] G.M.Woodruff, R.Kositpaiboon, "Multimedia traffic management principles for guaranteed ATM network performance", *IEEE J. Select. Areas Commun.*, vol. 8, n. 3, April 1990.

[3] S.Erfani, M.Malek, "Dynamic access capacity management in a multiservice packet-mode environment", *Proc. Workshop on Network Control and Management*, New York, 1990, pp. 277-285.

[4] R.Guérin, H.Ahmadi, M.Naghshineh, "Equivalent capacity and its application to bandwidth allocation in high speed networks", *IEEE J. Select. Areas Commun.*, vol. 9, n. 7, pp. 968-981, Sept. 1991.

[5] F.P.Kelly, "Effective bandwidths at multi-class queues", *Queueing Systems*, vol. 9, pp. 5-15, 1991.

[6] R.J.Gibbens, P.J.Hunt, "Effective bandwidths for the multi-type UAS channel", *Queueing Systems*, vol. 9, pp. 17-27, 1991.

[7] T.Kamitake, T.Suda, "Evaluation of an admission control scheme for an ATM network considering fluctuations in cell loss rate", *Proc. Globecom '89*, Dallas, TX, Nov. 1989, pp. 1774-1780.

[8] A.Lombardo, S.Palazzo, D.Panno, "Admission control over mixed traffic in ATM networks", *Internat. J. of Digital and Analog Commun. Systems*, vol. 3, Aug. 1990.

[9] Z.Dziong, J.Choquette, K.Q.Liao, L.Mason, "Admission control and routing in ATM networks", *Comput. Networks and ISDN Syst.*, vol. 20, pp. 189-196, Dec. 1990.

[10] C.Rasmussen, J.H.Sørensen, K.S.Kvols, S.B.Jacobsen, "Source independent call acceptance procedures in ATM networks", *IEEE J. Select. Areas Commun.*, vol. 9, n. 4, pp. 351-358, April 1991.

[11] Y.Takagi, S.Hino, T.Takahashi, "Priority assignment control of ATM line buffers with multiple QOS classes", *IEEE J. Select. Areas Commun.*, vol. 9, pp. 1078-1092, Sept. 1991.

[12] T.Murase, H.Suzuki, S.Sato, T.Takeuchi, "A call admission control scheme for ATM networks using a simple quality estimate", *IEEE J. Select. Areas Commun.*, vol. 9, pp. 1461-1470, Dec. 1991.

[13] D.Hong, T.Suda, "Congestion control and prevention in ATM networks", *IEEE Network Mag.*, vol. 5, no. 4, pp. 10-16, July 1991.

[14] S.Gupta, K.W.Ross, M.El Zarki, "Routing in virtual path based ATM networks", *Proc. Globecom'92*, Orlando, FL, Dec. 1992, pp. 571-575.

[15] A.A.Lazar, A.Temple, R.Gidron, "An architecture for integrated networks that guarantees quality of service", *Internat. J. of Digital and Analog Commun. Systems*, vol. 3, pp. 229-238, July 1990.

[16] J.M.Hyman, A.A.Lazar, G.Pacifici, "Real time scheduling with quality of service constraints", *IEEE J. Select. Areas Commun.*, vol. 9, pp. 1052-1063, Sept. 1991.

[17] J.M.Hyman, A.A.Lazar, G.Pacifici, "A separation principle between scheduling and admission control for broadband switching", *IEEE J. Select. Areas Commun.*, vol. 11, May 1993.

[18] R.Bolla, F.Davoli, A.Lombardo, S.Palazzo, D.Panno, "Adaptive bandwidth allocation by hierarchical control of multiple ATM traffic classes", *Proc. Infocom '92*, Florence, Italy, May 1992, pp. 30-38.

[19] R.Bolla, F.Danovaro, F.Davoli, M.Marchese "An integrated dynamic resource allocation scheme for ATM networks", *Proc. Infocom '93*, San Francisco, CA, March 1993, vol. 3, pp. 1289-1297.

[20] R.Bolla, F.Davoli, M.Marchese, "A distributed routing and access control scheme for ATM networks", *Proc. IEEE Internat. Conf. on Commun., ICC '94*, New Orleans, LA, May 1994, vol. 1, pp. 44-50.

[21] M.Aicardi, R.Bolla, F.Davoli, R.Minciardi, "A parametric optimization approach to admission control and bandwidth assignment in hybrid TDM networks", *Internat. J. of Digital and Analog Communication Systems*, vol. 6, pp. 15-27, Jan. 1993.

[22] R.Bolla, F.Davoli, "Dynamic hierarchical control of resource allocation in an integrated services broadband network", *Computer Networks and ISDN Systems*, vol. 25, no. 10, pp. 1079-1087, May 1993.