End-to-End Quality of Service Constrained Routing and Admission Control for MPLS Networks

Desire Oulai, Steven Chamberland, and Samuel Pierre

Abstract: **Multiprotocol label switching (MPLS) networks require dynamic flow admission control to guarantee end-to-end quality of service (QoS) for each Internet protocol (IP) traffic flow. In this paper, we propose to tackle the joint routing and admission control problem for the IP traffic flows in MPLS networks without rerouting already admitted flows. We propose two mathematical programming models for this problem. The first model includes end-to-end delay constraints and the second one, end-to-end packet loss constraints. These end-to-end QoS constraints are imposed not only for the new traffic flow, but also for all already admitted flows in the network. The objective function of both models is to minimize the end-to-end delay for the new flow. Numerical results show that considering end-to-end delay (or packet loss) constraints for all flows has a small impact on the flow blocking rate. Moreover, we reduces significantly the mean end-to-end delay (or the mean packet loss rate) and the proposed approach is able to make its decision within 250 msec.**

Index Terms: **Admission control mechanism, delay constraints, end-to-end QoS constraints, mathematical programming models, multiprotocol label switching (MPLS) networks, packet loss constraints, quality of service (QoS).**

I. INTRODUCTION

Multiprotocol label switching (MPLS) networks are typically designed to offer end-to-end quality of service (QoS) for Internet protocol (IP) traffic flows. Since QoS is important, for instance, for interactive voice and video applications, dynamic flow admission control is a central mechanism to accept or reject a new flow based on the QoS level requested and the available resources in the network. If there is no mechanism, the network could admit a new flow that overloads one or more links and then downgrade the QoS of several flows. Traffic routing is also an important mechanism in that context to find a path to the destination router while respecting QoS constraints. The admission control could use the routing results, but it may be independent from routing. For instance, even if there exists a feasible path, the admission control could reject a flow based on policies. However, most of the admission control mechanisms are routing-based.

Typically, QoS constraints are twofold: link constraints or path constraints. A link constraint is applicable locally to a link whereas path constraints are end-to-end. Examples of path constraints are end-to-end delay, packet loss, jitter, and bandwidth. Those constraints could also be considered locally to a link.

The admission control could be centralized [1], [2] or distributed [3], [4]. In the former case, a server gathers information on the network state to make the decision. In the latter, the edge routers keep total or partial information on the network state and the decision is made by the ingress node. A centralized server may imply a longer setup delay, but the distributed approach is subject to non-optimal decisions. Moreover, the admission control could be reservation based [1] or measurement based [5].

Even though many authors have worked in this area, most of the proposed mechanisms want to satisfy end-to-end delay or packet loss constraints for the new request without considering those constraints explicitly for the already admitted flows. To our knowledge, no solution has been proposed to satisfy the end-to-end delay constraints nor the end-to-end packet loss constraints of all flows without rerouting one or more already established flows. In this paper, the term "rerouting" a flow means finding a new path for this flow and using it for all packets forming this flow (i.e., no "flow splitting" is allowed). We want to avoid rerouting because it may be service affecting.

In this paper, we propose two models for the joint routing and admission control problem for the IP traffic flows in MPLS networks. The term joint routing and admission control means that the decision to admitted of not a new flow is taken simultaneously with the routing procedure, i.e., finding a set of label switched paths (LSPs) for the new flow to reach the destination. The first model includes end-to-end delay constraints and the second one, end-to-end packet loss constraints. These constraints are imposed not only for the new traffic flow, but also for all already admitted flows. The objective function is to minimize the end-to-end delay for the new flow.

The rest of the paper is organized as follows. Section II presents a literature review related to routing and admission control mechanisms. Section III presents preliminaries essential for understanding the proposed models. Section IV presents the model and numerical results for the joint routing and admission control (JRAC) with end-to-end delay constraints. In this paper, the models are solved to optimality. In fact, before spending time and efforts to develop heuristic algorithms, we first want to verify the value of the proposed models by solving them to optimality. Section V presents the model and numerical results for the JRAC with packet loss constraints. Finally, conclusion remarks are presented in Section VI.

II. RELATED WORKS

Routing-based admission control such as constraint-based routing (e.g., with delay or packet loss constraints) is an important area of research. For instance, Kodialam and Lakshman [6] proposed the minimum interference routing algorithm (MIRA).

Manuscript received August 30, 2007; approved for publication by Young Han Kim, Division III Editor, June 19, 2008.

The completion of this research was made possible thanks to Bell Canada's support through its Bell University Laboratories R&D program.

The authors are with the Department of Computer and Software Engineering, École Polytechnique de Montréal, Montréal, Canada, email: {desire.oulai, steven.chamberland, samuel.pierre}@polymtl.ca

The objective is to route the flow request over a path which minimizes the interference with possible future requests. Bagula *et al.* [7] introduced the least interference optimization algorithm (LIOA). LIOA calculates a cost for each link based on the number of connections through the link and the remaining capacity and then solved a shortest path. LIOA is less time consuming than MIRA and provides less blocking. Capone *et al.* [1] propose a virtual flow deviation method which allows to split and balance the flows among several paths. Widyono [2] presents an optimal centralized algorithm to find the least cost delay-constrained path. The algorithm (called constrained Bellman-Ford algorithm) performs a breadth-first search to find the optimal path. It is important to mention that end-to-end packet loss constraints are multiplicative constraints that can be transformed into additive constraints. Therefore, the algorithms for delay-constrained routing problems could be applied to packet loss constrained routing problems.

A more difficult problem is the multiconstrained admission control that considers simultaneously several constraints like delay, jitter, packet loss and bandwidth. The solutions proposed by Cui *et al.* [8] and Yuan [9] aimed at precomputing "optimal" constrained shortest paths. However, with dynamic traffic flows, the paths may not be optimal for the future requests. Jaffe [10] suggests to linearly combine the weights related to each constraint in order to obtain a composite weight for every link. The shortest paths are found using this composite weight. Another approach is the fallback algorithm described in [11]. The principle is to sequentially compute the shortest paths with regard to one QoS measure while hoping that it will satisfy all the constraints.

The admission control can also be done at link (or node) level where each link (or node) has a QoS threshold that cannot be exceeded (e.g., see Cui *et al.* [12], Nordstrom and Dziong [13] and Spitler and Lee [14]). The complexity of the problem is then reduced but the end-to-end QoS may not be fulfilled.

When doing the admission control for a new request, the service provider has to meet the QoS requirements of the already admitted flows. Indeed, most of the QoS measures such as delay and packet loss are related to the volume of traffic passing on the links. Therefore, the impact of accepting the new flow request has to be evaluated. In that sense, Khan *et al.* [15] introduce a utility model for optimal routing and admission control. Upon a request, a revenue function is maximized while observing every session with a QoS guarantee. k shortest paths are computed for each connection and the one that minimizes the revenue function is chosen. The algorithm then solve a global optimization problem for each request, which is time-consuming. To get optimal solution of each problem, the authors allow to reroute the flows, which can be service affecting. Ali *et al.* [3] propose an approach to reserve bandwidth for the already admitted flows to respect delay constraints. Their approach is based on the work by Paresh *et al.* [16]. The main drawback is that only link admission control is considered and this may not be enough for end-to-end delay objectives.

As mentioned before, no authors have considered the joint routing and admission control problem (with end-to-end delay and/or packet loss constrained) without rerouting while guaranteeing end-to-end QoS for all flows in the network.

III. PRELIMINARIES

A. The Notation

The following notation is used throughout the paper.

Sets: Let N the set of nodes (routers), M, the set of unidirectional links, L, the set of unidirectional LSPs, P_h, the set of paths joining all origin-destination pairs composed by at most h LSPs and finally, let T be the set of already admitted flows (where the flow $t \in T$ starts at node $O(t) \in N$, terminates at node $D(t) \in N$ and has a traffic request of α^t (in bps), a maximum delay limit of β^t (in sec) and a maximum end-to-end packet loss rate limit of ϕ^t).

Constants: Let y_{ij}^t be a 0-1 constant such that $y_{ij}^t = 1$ if and only if the flow $t \in T$ passes on the link $(i,j) \in M$ and z_{ij}^{ab} a 0-1 constant such that $z_{ij}^{ab} = 1$ if and only if the LSP $(a,b) \in L$ passes on the link $(i,j) \in M$.

Delay and packet loss functions: Let $d_{ij}(f_{ij})$ be the delay on the link (i,j), $p_{ij}(f_{ij})$, the packet loss rate on the link (i,j), and $r_{ij}(f_{ij})$, the packet transmit rate on the link (i,j), that is, $r_{ij}(f_{ij}) = 1 - p_{ij}(f_{ij})$. In this paper, the $M/M/1/k$ queuing model is used [17]. As a result,

$$d_{ij}(f_{ij}) = \frac{\rho(1 + k\rho^{k+1} - (k+1)\rho^k)}{\lambda(1-\rho)(1-\rho^k)} + a_{ij} \quad (1)$$

$$p_{ij}(f_{ij}) = 1 - r_{ij}(f_{ij}) = \frac{\rho^k(1-\rho)}{1-\rho^{k+1}} \quad (2)$$

$$\lambda = \frac{f_{ij}}{\ell} \quad (3)$$

$$\rho = \frac{f_{ij}}{c_{ij}} \quad (4)$$

where

- λ, the mean arrival rate (in packet/sec) on the link (i,j);
- ρ, the average utilization of the link (i,j);
- k, the buffer size (in packets);
- ℓ, the mean packet length (in bits);
- c_{ij}, the capacity (in bps) of the link (i,j);
- f_{ij}, the traffic (in bps) on the link (i,j) and finally,
- a_{ij}, the propagation delay on the link (i,j) plus the processing delay at node $i \in N$ (in sec).

Since this model is easy to analyze and considering that any delay and packet loss models can be used with the proposed JRAC framework, we have selected this model. Indeed, the delay and packet loss parameters are only input parameters of the proposed JRAC framework. In a real network, the delay and the packet loss on a link can be measured by the routers and then, only the delta parameters have to be estimated using, for instance, stochastic models.

Variables: Let x_{ab} be a 0-1 variable such that $x_{ab} = 1$ if and only if the new flow passes on the LSP (a,b) and y_{ij}, a 0-1 variable such that $y_{ij} = 1$ if and only if the new flow passes on the link (i,j).

B. Problem Formulation and Preprocessing

The joint routing and admission control problem proposed in this paper consists in finding a path for the new flow from its origin node o to its destination node d and having a traffic request of α and with an end-to-end delay limit of β or an end-to-end packet loss ratio limit of ϕ while considering already admitted flows in the network. If such a path does not exist, the new flow is blocked.

We consider a logical topology already laid. This topology is formed by LSPs set up between the edge nodes (routers). In this paper, there is one LSP between each pair of edges nodes and an LSP is seen as one hop in the logical topology. The LSPs are routed using Dijkstra's shortest path algorithm. The LSPs' bandwidth is adjusted upon acceptance of new flows. A protocol like RSVP [18] could be used to adjust bandwidth for the LSPs. Since a flow can use multiple LSPs, the IP header is inspected at an LSP termination to figure out if the current nodes is the destination or if the packet has to be forward through another LSP.

To formulate the mathematical models, preprocessing is necessary. Note that if the new flow does not pass on the link (i, j), the traffic on that link will be

$$F_{ij} = \sum_{t\in T} \alpha^t y_{ij}^t \tag{5}$$

and the delay on the link (i, j) will be $D_{ij} = d_{ij}(F_{ij})$, the packet loss rate $P_{ij} = p_{ij}(F_{ij})$ and the packet transmit rate $R_{ij} = 1 - P_{ij}$. Otherwise, the traffic on that link will be

$$\overline{F}_{ij} = \sum_{t\in T} \alpha^t y_{ij}^t + \alpha \tag{6}$$

and the delay on the link (i, j) will be $\overline{D}_{ij} = d_{ij}(\overline{F}_{ij}) = D_{ij} + \Delta D_{ij}$, the packet loss rate $\overline{P}_{ij} = p_{ij}(\overline{F}_{ij})$ and the packet transmit rate $\overline{R}_{ij} = 1 - \overline{P}_{ij}$.

Similarly, if the new flow does not pass on the LSP (a, b), the end-to-end delay, the transmit rate and the packet loss rate are respectively given by the following equations

$$D_{ab} = \sum_{(i,j)\in M} D_{ij} z_{ij}^{ab} \tag{7}$$

$$R_{ab} = \prod_{(i,j)\in M: z_{ij}^{ab}=1} R_{ij} \tag{8}$$

$$P_{ab} = 1 - R_{ab}. \tag{9}$$

Otherwise, if the new flow passes on the LSP (a, b),

$$\overline{D}_{ab} = \sum_{(i,j)\in M} \overline{D}_{ij} z_{ij}^{ab} = D_{ab} + \Delta D_{ab} \tag{10}$$

$$\overline{R}_{ab} = \prod_{(i,j)\in M: z_{ij}^{ab}=1} \overline{R}_{ij} \tag{11}$$

$$\overline{P}_{ab} = 1 - \overline{R}_{ab}. \tag{12}$$

IV. DELAY CONSTRAINED ADMISSION CONTROL

A. The Model

The mathematical model for the joint routing and admission control problem in MPLS networks without flow rerouting and with end-to-end delay constraints, denoted JRAC-D (joint routing and admission control with delay constraints), can now be given

JRAC-D:

$$\min_{\{x_{ab}:(a,b)\in L\}} \sum_{(a,b)\in L} \overline{D}_{ab} x_{ab} \tag{13}$$

subject to

$$y_{ij} = \sum_{(a,b)\in L} z_{ij}^{ab} x_{ab}, \quad \forall (i,j) \in M \tag{14}$$

$$0 \le y_{ij} \le 1, \quad \forall (i,j) \in M \tag{15}$$

$$\sum_{(a,b)\in L} x_{ab} \le h \tag{16}$$

$$\sum_{(i,j)\in M} (D_{ij} + y_{ij} \Delta D_{ij}) y_{ij}^t \le \beta^t, \quad \forall t \in T \tag{17}$$

$$\sum_{(a,b)\in L} \overline{D}_{ab} x_{ab} \le \beta \tag{18}$$

$$\sum_{b:(a,b)\in L} x_{ab} - \sum_{b:(b,a)\in L} x_{ba} = \begin{cases} 1 & \text{if } a = o \\ -1 & \text{if } a = d \\ 0 & \text{otherwise} \end{cases} \quad \forall a \in N \tag{19}$$

$$x_{ab} \in \{0, 1\} \quad \forall (a,b) \in L. \tag{20}$$

The objective function (13) of JRAC-D is to minimize the end-to-end delay for the new flow. Constraints (14) force the variable y_{ij} to be equal to the number of LSPs used by the new flow passing on the link (i, j) and constraints (15) impose this number be less than or equal to one, i.e., the new flow is allowed to pass on a link at most once. Note that the $y_{i,j}$ variables can be removed from the model (i.e., constraints (14) and (15) can be combined in a single expression and the $y_{i,j}$ variables in (17) can be replaced with the summation term in (14)). However, the model is more readable with those variables Constraint (16) limits the number of LSPs in the path used by the new flow to be at most h. This limit is applied to facilitate the admission control process. It allows us to use a path-based approach. Rather than scanning every flow during the admission control process, we just scan the possible paths composed with $1, 2, \cdots, h$ LSPs. Constraints (17) impose each already admitted flow in the network to respect the end-to-end delay limit and constraint (18), the new flow to respect the end-to-end delay limit. Constraints (17) and (18) are important because they force all flows to respect the QoS requirements. Constraints (19) are the flow conservation constraints and, finally, constraints (20) are integrality constraints.

JRAC-D is NP-hard (transformation from the shortest weight-constrained path problem [19]). However, since the number of integer variables is small, JRAC-D can be solved to optimality for real-size instances of the problem within a small amount of computational time.

B. The Reduction Algorithm

Since this number of constraints (17) is equal to $|T|$ (i.e., the number of already admitted flows), it could be computationally expensive to consider all these constraints. However, these constraints can be easily reduced as follows. First note that only traffic flows that share a link with the new one will be affected.

A path-based approach is considered. As mentioned before, a path is a set of LSPs (i.e., one or more $(a,b) \in L$) used from an origin node to reach a destination node. Let D_p^{lim} be the end-to-end delay limit for a given path p, i.e., the minimum of the β^t (i.e., the maximum end-to-end delay limit) of each flow t using the path p from its origin to its destination. If there is no flow using the path p, $D_p^{lim} = \infty$ (i.e., set to a very large value). In this paper, to reduce the number of paths, we limit the maximum number of LSPs composing a path to two (i.e., $h = 2$). Therefore, we maintain a table of all LSPs and all chains of two LSPs not forming a cycle. For each path, D_p^{lim} is computed.

We define $\overline{D}_p$ the delay of the path p if the new flow passes on that path, i.e., the sum of the $\overline{D}_{ab}$ for all LSPs (a,b) forming the path p. If $\overline{D}_p$ is less than or equal to D_p^{lim}, we do not need to consider the end-to-end delay constraints for all flows using that path. Indeed, constraints (17) could be rewritten as follows

$$\sum_{(a,b)\in L\cap p} \sum_{(i,j)\in M} (D_{ij} + y_{ij}\Delta D_{ij}) z_{ij}^{ab} \le D_p^{lim} \quad \forall p \in P_h : \overline{D}_p \ge D_p^{lim}. \tag{21}$$

The reduction algorithm is now given.

Reduction Algorithm

Step 1: (Initialization)

1.1 For all $(i,j) \in M$, set $\overline{F}_{ij} := \sum_{t\in T} \alpha^t y_{ij}^t + \alpha$ and $\overline{D}_{ij} := d_{ij}(\overline{F}_{ij})$.

1.2 For all $(a,b) \in L$, set $\overline{D}_{ab} := \sum_{(i,j)\in M : z_{ij}^{ab}=1} \overline{D}_{ij}$.

1.3 For all $p \in P_h$, set $D_p^{lim} := \infty$ and $\overline{D}_p := \sum_{(a,b)\in L:(a,b)\in p} \overline{D}_{ab}$.

Step 2: (Model generation)

2.1 Generate the model JRCA-D without constraints (17).

2.2 For each path $p \in P_h$ do
If $\overline{D}_p \ge D_p^{lim}$, add the following constraint in the model

$$\sum_{(a,b)\in L\cap p} \sum_{(i,j)\in M} (D_{ij} + y_{ij}\Delta D_{ij})\, z_{ij}^{ab} \le D_p^{lim}.$$

Step 3: Solve the model. (In this paper, the CPLEX Mixed Integer Optimizer 9.0 [20] is used for this step.)

C. Numerical Results

In this section, we evaluate the performance of the proposed routing-based admission control mechanism. All algorithms were programmed in the C language on a Linux workstation with 8 GB of RAM and a 2.4 GHz processor. For solving the model JRAC-D, the CPLEX Mixed Integer Optimizer 9.0 (see [20] for more information about CPLEX) is used. Note that the algorithm used by the CPLEX is the branch-and-bound algorithm. The default settings of CPLEX are used.

We perform tests on the well-known MIRA network, presented in Fig. 1, that have been used for simulations in several papers (see, for instance, [6] and [7]). In this figure, the dark lines are 48 Mbps links and the light lines are 12 Mbps links. The nodes not connected to dark lines are the edges nodes.

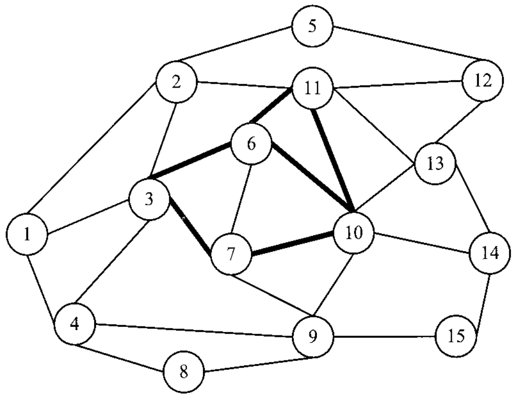

Fig. 1. The MIRA network.

Table 1. Features of the test networks.

$\lvert N\rvert$	$\lvert T\rvert$	Number of edge nodes	Number of links
5	400	3	10
10	800	6	15
20	1600	12	30
30	2400	18	45
40	3200	24	60
50	4000	30	75
60	4800	36	90
70	5600	42	105
80	6400	48	120

In addition to the MIRA network, random networks are used to assess the proposed approach for large-size instances of the problem. Those networks are generated as follows: we first built an Hamiltonian cycle (for two-connectivity) and additional links are randomly added. The capacity of each link is set to 100 Mbps and 60% of the nodes are chosen to be edge nodes. The features of the random networks are presented in Table 1.

The traffic flows are randomly generated. For each request, we randomly choose a pair of origin-destination edge nodes. For the MIRA network, we generate sets of 5000, 6000, 7000, and 8000 requests. The bandwidth of each flow is randomly taken (with equal probability) from the set {10, 20, 30, 40} kbps. For the random networks, each flow has a bandwidth of 1 Mbps.

For the tests, the end-to-end delay limit of each request is randomly taken from the set {50, 100, 150, 200, 250, 300} in msec and, finally, the packet length is set to 1500 bytes and the buffer size to 432 kbytes (i.e., $k = 288$ packets) or 1200 kbytes (i.e., $k = 800$ packets).

In this section, JRAC-D is compared to three other algorithms.

- LIOA [7]: This algorithm computes the least interference path. The link cost on the link (i,j) is calculated using the formula

$$u_{ij} = I_{ij}^{\omega} / S_{ij}^{1-\omega}$$

where I_{ij} is the number of flows carried on link (i,j) and S_{ij} is the remaining reservable bandwidth on link (i,j). Simulations proved that $\omega = 0.5$ provide better results re-

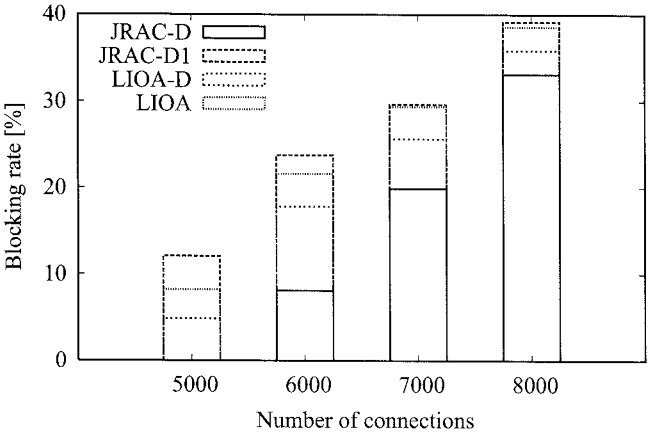

Fig. 2. Flow blocking rate for $k = 288$ (MIRA network).

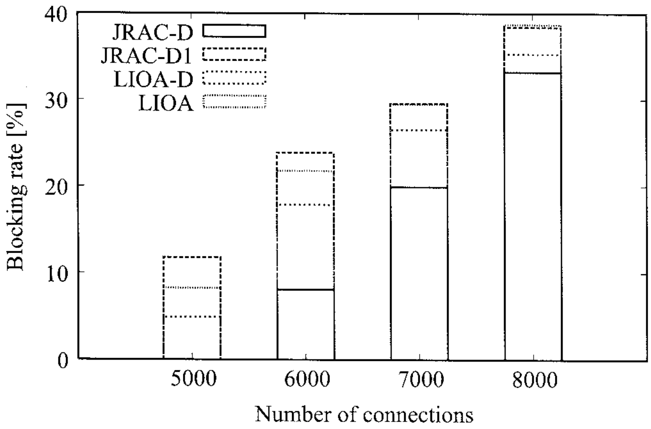

Fig. 3. Flow blocking rate for $k = 800$ (MIRA network).

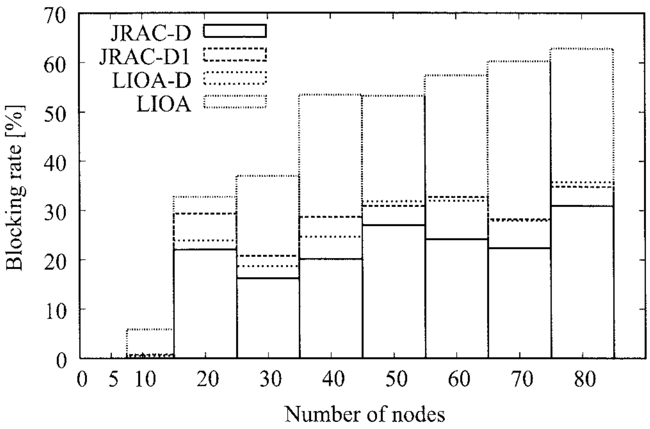

Fig. 4. Flow blocking rate for $k = 288$ (random networks).

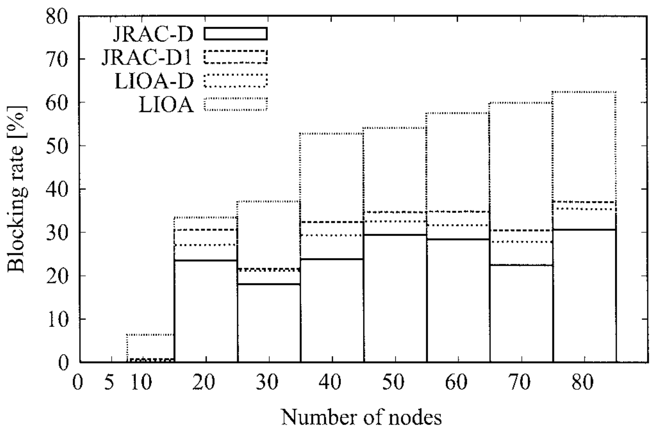

Fig. 5. Flow blocking rate for $k = 800$ (random networks).

garding blocking rate. LIOA, also consider an end-to-end delay constraint for the new flow;

- LIOA-D: LIOA with additional end-to-end delay constraints for all flows;
- JRAC-D1: JRAC-D with $h = 1$.

The metrics of interest to evaluate the performance of our mechanism are the flow blocking rate, the mean end-to-end delay, the ratio of constraints (17) violated (i.e., the proportion of flows exceeding their delay limit), and finally, the CPU execution time.

Figs. 2 and 3 present the flow blocking rate for the MIRA network and Figs. 4 and 5 for the random networks. In each scenario, JRAC-D provides the lowest blocking rate with a maximum of 33%. This is due to the limit on the number of LSPs per path and to the objective function chosen, i.e, minimizing the end-to-end delay for the new flow. Indeed, minimizing end-to-end delay reduce the impact of the new flow on the admission of future requests. LIOA-D, is the second best algorithm regarding blocking rate. The introduction of end-to-end delay constraints for all flows contribute to limit the delay on each link which is good to accept future requests.

For the MIRA network, LIOA provides better results than JRAC-D1, but with a difference less than 4%. We observe that this difference decreases when the number of connections increases. For the random networks, JRAC-D1 performs better than LIOA. This can be explained by the fact that JRCA-D imposes a limit on the number of LSPs per path and that its objective function is to minimize the end-to-end delay.

Figs. 6 and 7 present the end-to-end delay for the MIRA network and Figs. 8 and 9 for the random networks. We observe that for all scenarios and algorithms except for LIOA, the mean end-to-end delay is less than 25 msec. For the LIOA algorithm, the peak delay is 230 msec. This may result from the fact that LIOA allows longer paths. Moreover, delays are higher when using $k = 800$, because there is more waiting time in the buffers.

Figs. 10 and 11 present the mean CPU execution time for the proposed admission control mechanism. We illustrate the worst cases, i.e., when $k = 800$. These figures show that this approach can be used for real-size networks and the decision time (to admit or not a new flow) is less than 245 msec (including the preprocessing and the model resolution). For the MIRA network, the total time is under 3.5 msec. CPLEX takes less than 50 msec to solve the model. Note that for $k = 288$, the total time is under 200 msec.

To conclude this part, we can say that JRAC-D provides good results. Our approach allows us to guarantee the end-to-end delay limits for all flows without rerouting. Moreover, JRAC-D provides less blocking rate than other methods while running in a reasonable amount of time. We currently work on efficient heuristics to decrease the CPU execution time to find quasi-optimal solutions.

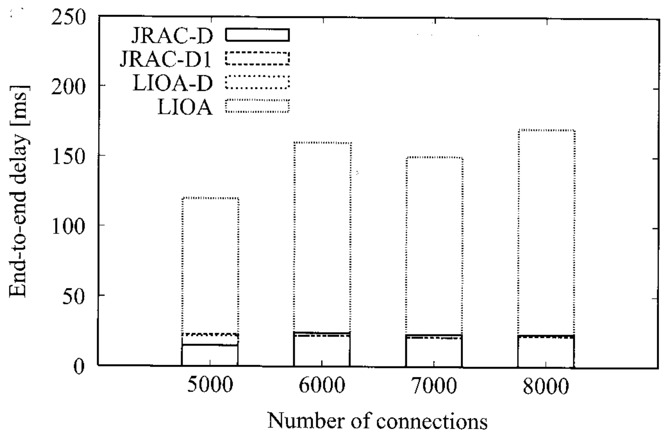

Fig. 6. End-to-end delay for $k = 288$ (MIRA network).

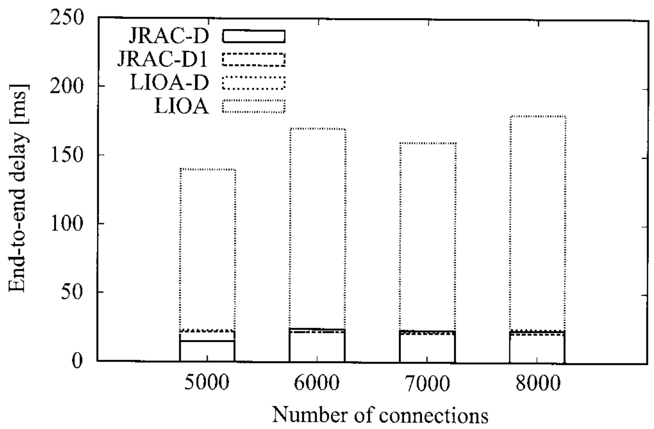

Fig. 7. End-to-end delay for $k = 800$ (MIRA network).

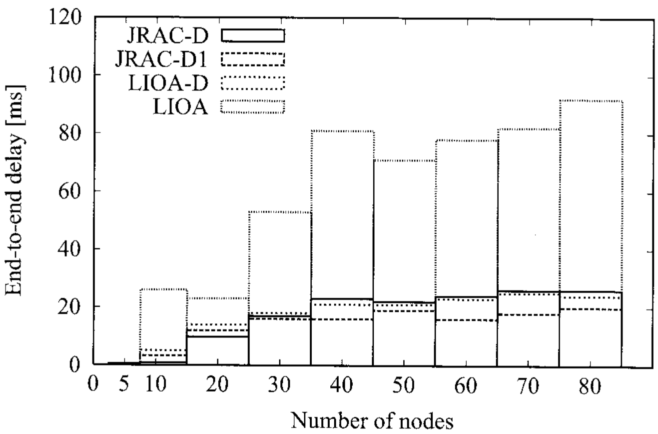

Fig. 8. End-to-end delay for $k = 288$ (random networks).

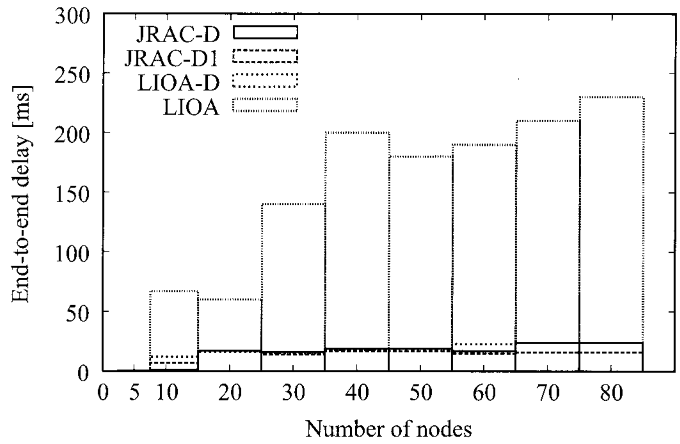

Fig. 9. End-to-end delay for $k = 800$ (random networks).

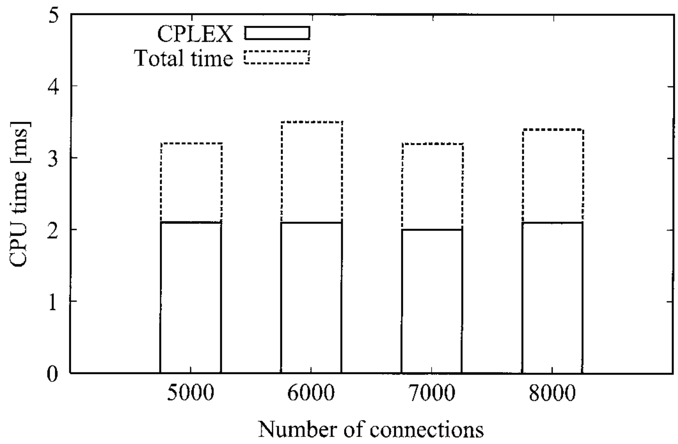

Fig. 10. CPU execution time for JRAC-D ($k = 800$, MIRA network).

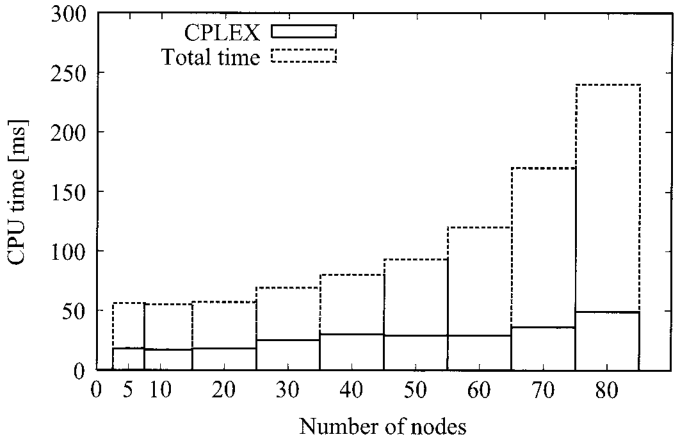

Fig. 11. CPU execution time for JRAC-D ($k = 800$, random networks).

V. PACKET LOSS CONSTRAINED ADMISSION CONTROL

A. The Model

The mathematical model for the joint routing and admission control problem in MPLS networks without flow rerouting and with end-to-end packet loss constraints, denoted JRAC-P (joint routing and admission control with packet loss constraints), can now be given

JRAC-P:

$$\min_{\{x_{ab}:(a,b)\in L\}} \sum_{(a,b)\in L} \overline{D}_{ab} x_{ab} \tag{22}$$

subject to (14)–(16), (19), (20), and

$$\sum_{(i,j)\in M} (\ln\ R_{ij} + y_{ij} \ln \frac{\overline{R}_{ij}}{R_{ij}}) y_{ij}^t \geq \ln(1-\phi^t) \quad \forall t \in T \tag{23}$$

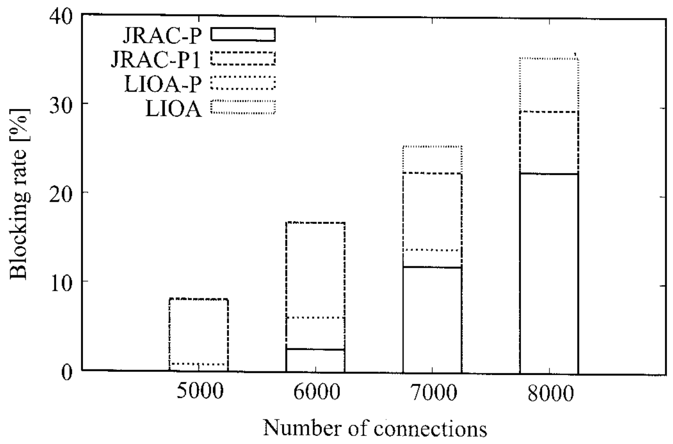

Fig. 12. Flow blocking rate for $k = 288$ (MIRA network).

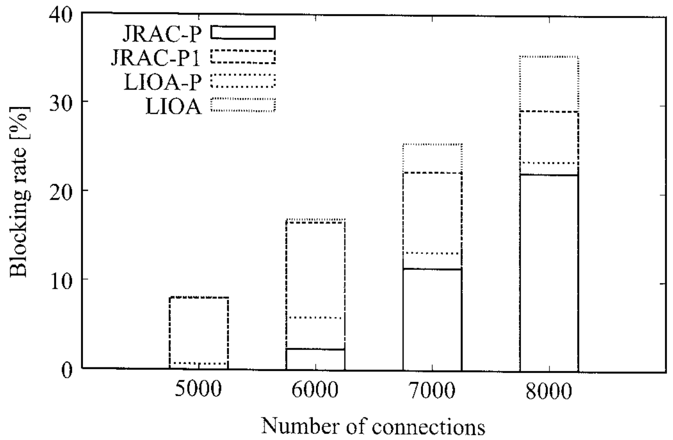

Fig. 13. Flow blocking rate for $k = 800$ (MIRA network).

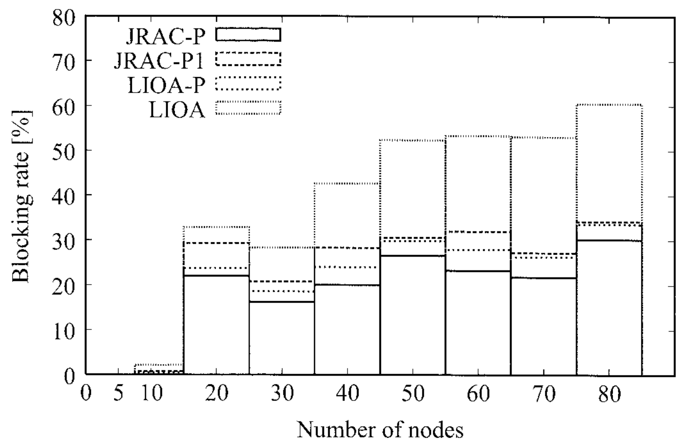

Fig. 14. Flow blocking rate for $k = 288$ (random networks).

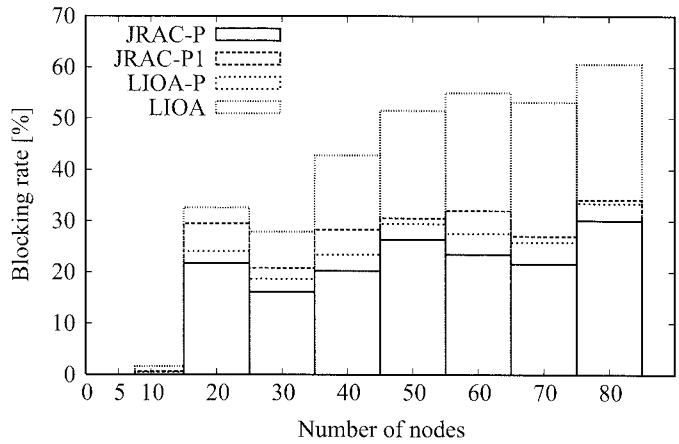

Fig. 15. Flow blocking rate for $k = 800$ (random networks).

$$\sum_{(a,b)\in L} x_{ab} \ln \overline{R}_{ab} \geq \ln(1-\phi). \qquad (24)$$

The objective function (22) of JRAC-P is to minimize the end-to-end delay for the new flow. This objective function is used because delay is an important QoS parameter for most of the applications. Constraints (23) and (24) impose all flows to respect the packet loss requirements. These linear constraints are obtained by logarithmic transformations. To enumerate these constraints, the preprocessing should verify that $\phi < 1$, $\phi^t < 1$ for all $t \in T$, $\underline{R}_{ij} > 0$ for all $(i,j) \in M$, $\overline{R}_{ij} > 0$ for all $(i,j) \in M$ and $\overline{R}_{ab} > 0$ for all $(a,b) \in L$.

As for JRAC-D, JRAC-P is NP-hard. However, we will demonstrate that it requires a small amount of computational time to solve real-size instances of the problem to optimality.

For the same reasons and with the same approach described in Section IV-B, we can reduce the number of constraints (23).

B. Numerical Results

For the tests, the end-to-end packet loss limit of each flow is randomly selected from the set $\{0.01, 0.02, 0.03, 0.04, 0.05\}$. The other test parameters are the same as those presented in Section IV-C. Here, JRAC-P1 is similar to JRAC-P but with a maximum of one LSP per path and LIOA-P is the same as LIOA with end-to-end packet loss constraints for all flows.

Figs. 12–15 present the flow blocking rate for the test networks. Note that, as expected for the MIRA network, the blocking rate increases with the load. In all scenarios, JRAC-P provides the lowest blocking rate with a maximum of 30%. The second best algorithm is LIOA-P with a maximum of 4% more blocking than JRAC-P.

Figs. 16–19 present the mean end-to-end packet loss. As expected, since the constraints (23) assure the end-to-end packet loss constraints to be respected, we obtain better results with those constraints. Whenever our new constraints are applied, the mean end-to-end packet loss is under 0.5%.

Finally, Figs. 20 and 21 show the mean CPU execution time for the proposed admission control mechanism. These figures assess that a reasonable amount of computational time is necessary for the proposed approach for real-size networks. The total decision time is less than 250 msec. We currently work on efficient heuristics to decrease the CPU execution time to find quasi-optimal solutions.

VI. CONCLUSIONS

In this paper, we have proposed two models for the joint routing and admission control problem for the traffic flows in MPLS networks. The first model includes end-to-end delay constraints and the second one, end-to-end packet loss constraints. These

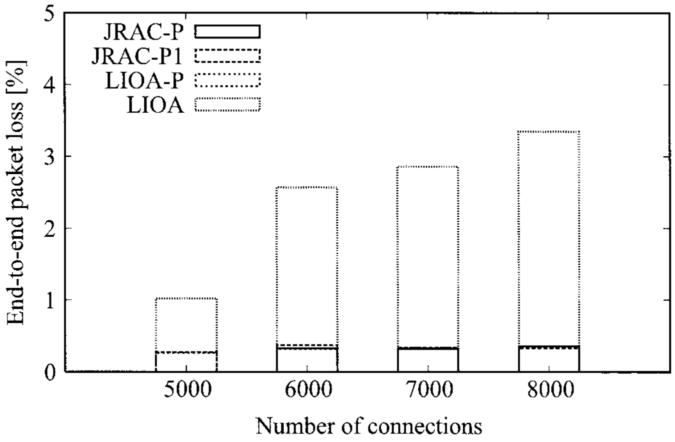

Fig. 16. End-to-end packet loss for $k = 288$ (MIRA network).

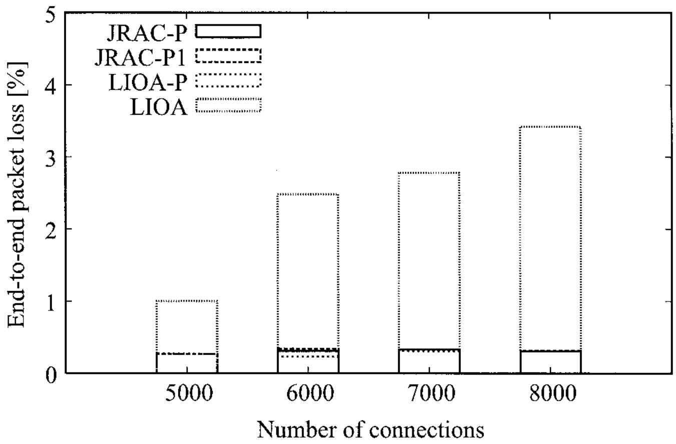

Fig. 17. End-to-end packet loss for $k = 800$ (MIRA network).

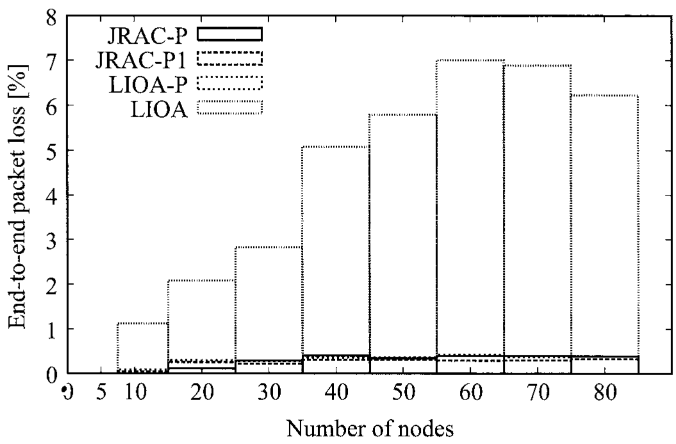

Fig. 18. End-to-end packet loss for $k = 288$ (random networks).

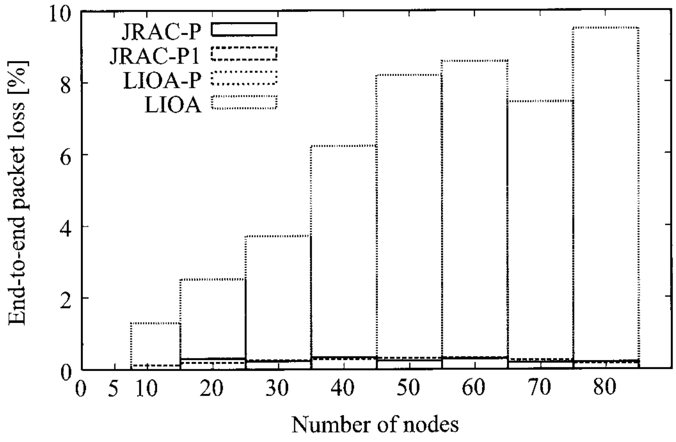

Fig. 19. End-to-end packet loss for $k = 800$ (random networks).

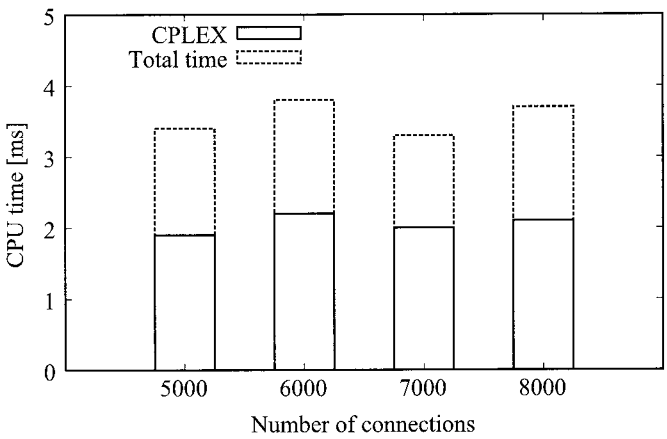

Fig. 20. CPU execution time for JRAC-P ($k = 800$) (MIRA network).

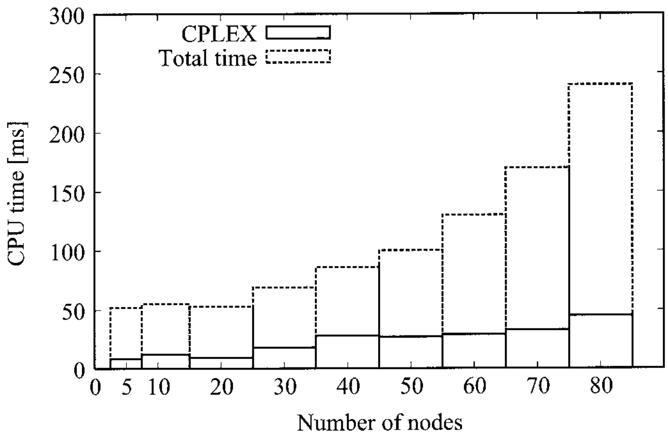

Fig. 21. CPU execution time for JRAC-P ($k = 800$) (random networks).

end-to-end QoS constraints are imposed not only for the new traffic flow, but also for all already admitted flows in the network. The objective function is to minimize the end-to-end delay on the path used by the new flow. These models are solved exactly (after preprocessing) within the admission control mechanism.

The numerical results show that considering end-to-end delay (or packet loss constraints) for all flows while limiting the number of LSPs per path, permit to offer less flow blocking rate. Moreover, we reduced significantly the mean end-to-end delay (or the mean packet loss rate) and the proposed approach is able to make the decision to admit or not a new flow in the network within 250 msec.

There are several avenues of research open at this point. First, we want to consider other objective functions like revenue function. We currently work on efficient heuristics to find good (quasi-optimal) solutions rapidly. It will also be interesting to test our approach on multiconstrained models.

REFERENCES

[1] A. Capone, L. Fratta, and F. Martignon, "Dynamic routing of bandwidth guaranteed connections in MPLS networks," *Int. J. Wireless and Optical Commun.*, vol. 1, no. 1, pp. 75–86, May 2003.

[2] R. Widyono, "The design and evaluation of routing algorithms for real-time channels," University of California at Berkeley, Tech. Rep., ICSI TR-94-024, June 1994.

[3] N. A. Ali, H. T. Mouftah, and S. Gazor, "Online distributed statistical-delay MBAC with QoS guarantees for VPLS connections," *Int. Conference on Telecom.*, June 2005, vol. 2, pp. 383–390.

[4] D. S. Reeves and H. S. Salama, "A distributed algorithm for delay-constrained unicast routing," *IEEE/ACM Trans. Netw.*, vol. 8, no. 2, pp. 239–250, Apr. 2000.

[5] H. Uzunalioglu, D. J. Houck, and Y. T. Wang, "Call admission control for voice over IP," *Int. J. Commun. Systems*, vol. 19, no. 4, pp. 363–380, Apr. 2006.

[6] M. S. Kodialam and T. V. Lakshman, "Minimum interference routing with applications to MPLS traffic engineering," in *Proc. IEEE INFOCOM*, Mar. 2000, vol. 2, pp. 884–893.

[7] A. B. Bagula, M. Botha, and A. E. Krzesinski, "Online traffic engineering: The least interference optimization algorithm," in *Proc. IEEE ICC*, June 2004, vol. 2, pp. 1232–1236.

[8] Y. Cui, K. Xu, and J. Wu, "Precomputation for multiconstrained QoS routing in high-speed networks," in *Proc. IEEE INFOCOM*, Apr. 2003, vol. 2, pp. 1414–1424.

[9] X. Yuan, "Heuristic algorithms for multiconstrained quality-of-service routing," *IEEE/ACM Trans. Netw.*, vol. 10, no. 2, pp. 244–256, Apr. 2002.

[10] J. M. Jaffe, "Algorithms for finding paths with multiple constraints," *Netw.*, vol. 14, no. 1, pp. 95–116, Apr. 1983.

[11] F. Kuipers, P. Van Mieghem, T. Korkmaz, and M. Krunz, "An overview of constraint-based path selection algorithms for QoS routing," *IEEE Commun. Mag.*, vol. 40, no. 12, pp. 50–55, Dec. 2002.

[12] Y. Cui, K. Xu, and J. Wu, "Multiconstrained end-to-end admission control in core-stateless networks," *Comput. Science*, vol. 3090, pp. 420–429, Feb. 2004.

[13] E. Nordstrom and Z. Dziong, "CAC and routing for multi-service networks with blocked wide-band calls delayed, part I: Exact link MDP framework," *Euro. Trans. Telecom.*, vol. 17, no. 1, pp. 21–36, Jan. 2006.

[14] S. L. Spitler and D. C. Lee, "Optimal call admission control under packet and call level QoS constraints and effect of buffering," *Int. J. Commun. Systems*, vol. 16, no. 7, pp. 647–662, May 2003.

[15] S. Khan, K. F. Li, E. G. Manning, R. Watson, and G. C. Shoja, "Optimal quality of service routing and admission control using the utility model," *Future Generation Comput. Systems*, vol. 19, no. 7, pp. 1063–1073, Oct. 2003.

[16] A. K. Parekh and R. Gallager, "A generalized processor sharing approach to flow control in integrated services networks: The multiple node case," *IEEE/ACM Trans. Netw.*, vol. 2, no. 2, pp. 137–150, Apr. 1994.

[17] D. Bertsekas and R. Gallager, *Data networks*. 2nd ed., Prentice-Hall, 1992.

[18] R. Braden, L. Zhang, S. Berson, S. Herzog, and S. Jamin, "Resource reservation protocol (RSVP) - Version 1 functional specification," IETF RFC 2205, Sept. 1997.

[19] M. R. Garey and D. S. Johnson, *Computers and intractability: A guide to the theory of NP-completeness*. W. H. Freeman, Co. New York, 1979.

[20] ILOG, Inc., "Using the CPLEX callable library and CPLEX mixed integer library," ILOG, Inc., 2005.

Desire Oulai received the Computer Engineering degree in 2002, the master degree in Computer Engineering in 2003 and the Ph.D. degree in Computer Engineering in 2007 all from the École Polytechnique de Montréal. He is currently a Postdoctoral Researcher at Ericsson Research Canada and at the École Polytechnique de Montréal. His research interests include flow admission control and routing in data networks, wireless networks, mathematical modeling, and meta-heuristics.

Steven Chamberland received the Electrical Engineering degree from the École Polytechnique de Montréal, Canada, in 1994 and the Ph.D. degree in Operations Research in 1998 also from the École Polytechnique. From 1998 to 1999, he was with the network architecture group of Bell Canada and from 1999 to 2001, he was a Professor at the École de technologie supérieure. Since 2001, he has been an Assistant, Associate and now a Full Professor of Computer Engineering at the École Polytechnique de Montréal where he is Director of the Bell Canada Research Chair in Network Architecture. His research interests include telecommunication network planning and design and network architecture. He is a Member of IEEE and ACM. He also serves on the Editorial Board of *Computers and Operations Research* and *Information Systems and Operational Research* (INFOR).

Samuel Pierre received the B.Eng. degree in Civil Engineering in 1981 from the École Polytechnique de Montréal, the B.Sc. and M.Sc. degrees in mathematics and computer science in 1984 and 1985, respectively, from the Université du Québec, the M.Sc. degree in economics in 1987 from the Université de Montréal and, the Ph.D. degree in Electrical Engineering in 1991 from the École Polytechnique de Montréal. He is currently a Professor of Computer Engineering at École Polytechnique de Montréal where he is Director of the Mobile Computing and Networking Research Laboratory (LARIM) and NSERC/Ericsson Industrial Research Chair in Next-generation Mobile Networking Systems. His research interests include wireless networks, mobile computing, performance evaluation, artificial intelligence, and electronic learning. He is a Fellow of Engineering Institute of Canada, Senior Member of IEEE, a Member of ACM and IEEE Communications Society. He is a Fellow of Engineering Institute of Canada, Senior Member of IEEE, a Member of ACM and IEEE Communications Society.