Integrating Effective-Bandwidth-Based QoS Routing and Best Effort Routing

Stephen L. Spitler and Daniel C. Lee
University of Southern California
Department of Electrical Engineering
Los Angeles, CA, USA
stephen.l.spitler@boeing.com, dclee@usc.edu

Abstract—A methodology is presented for integrating effective-bandwidth-based routing for QoS-sensitive traffic and datagram routing of the best-effort traffic. To prevent excessive delays of best-effort traffic in a network domain, we develop (1) a constraint, stated in the form of a residual link bandwidth, and (2) a cost function for application to routing of QoS connections. Link-based and path-based problem formulations and algorithms are presented. For the case that a cost quantization condition holds, we develop an efficient implementation of a link-based routing strategy that first minimizes a QoS cost, then secondarily minimizes a best-effort cost. The performance of this approach is further enhanced by explicitly accounting for the difference between the effective bandwidth and the average bandwidth of traffic. Simulation results illustrate the application of our BE-friendly method to an algorithm for path routing with restoration.

Keywords— Effective bandwidth; quality of service; path restoration; dynamic routing; constraint-based routing; MPLS; best-effort traffic

I. INTRODUCTION

In traditional IP routing each packet typically seeks the shortest path to its destination through the connectionless, hop-by-hop routing; the idea of connection-based end-to-end routing, e.g., Multiprotocol Label Switching (MPLS) [1], has been subsequently added. An advantage of the connection-based routing would be its ability to facilitate the quality of service (QoS) through traffic regulation and engineering for the traffic carried through the connection. In this paper, we address how to integrate the QoS routing and the best-effort (BE) routing in the same IP administrative domain; e.g., an internet service provider (ISP).

In order to emphasize the applicability of the methodology, we consider an internet domain supporting MPLS and use MPLS terminology.   However, the methodology can be generally applied to implementations other than MPLS, and what is described by the term, label switched path (LSP), in this paper can be regarded as any path and bandwidth provision along that path for a micro-flow or a set of micro-flows with a certain QoS requirement. We consider an internet network domain with a heterogeneous traffic mix of various classes using both MPLS and the destination-based, hop-by-hop IP routing. Specifically, we assume that LSPs are provided for traffic trunks, i.e., micro-flows or aggregates of micro-flows that share a common class of service (CoS). We assume that the traffic trunks routed require certain QoS guarantees. The network domain also carries the BE (best-effort) class, which does not require any QoS guarantees. Although a service provider may also set up LSPs to carry BE traffic that requires no QoS guarantees, in this paper we use the term, BE traffic, to refer to traffic that is forwarded via traditional destination-based, hop-by-hop routing tables. Furthermore, we assume that, although BE traffic can take advantage of unused bandwidth that is allocated to LSPs for QoS trunks, it uses it on a strictly lower priority basis.

Referring to Fig. 1, suppose that a QoS traffic trunk, t, and BE traffic both share the minimum-hop path from ingress label switched router (LSR) s to egress LSR d, i.e., path s-u-d. Further, suppose that additional micro-flows are being added to trunk t so that t becomes larger, i.e., it uses more bandwidth. Consequently, BE traffic gets crowded out and may begin to suffer excessive delays as bandwidth at links (s, u) and (u, d) is preempted to support t. Now suppose that the LSP routing algorithm sets up another path from s to d to support a new traffic trunk. If paths are selected on the basis of hop count, then s-v-w-d and s-x-y-d are equally attractive choices. However, if the routing algorithm arbitrarily selects s-x-y-d, then the new trunk will impact BE traffic that is flowing from y to d while selection of s-v-w-d would have left that BE traffic undisturbed.
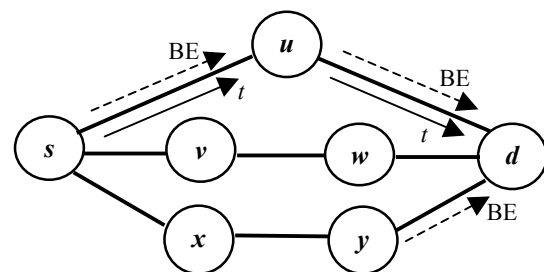

Figure 1. Sharing of a minimum-hop path by traffic trunk t and BE traffic.

This paper presents a methodology to make QoS routing more BE-friendly. We constrain QoS routing so as to prevent it from causing excessive BE traffic delays but mitigate the undesirable consequences for QoS routing, e.g., increased blocking probability. We do this by making use of what we call 'excess effective bandwidth,' the margin by which a QoS trunk's effective bandwidth exceeds the average bandwidth that it actually consumes. We take excess effective bandwidth into account when we apply our BE traffic delay constraint in determining which paths are feasible for QoS routing. Also, we define a cost quantization condition and, for the case that it applies, exploit it to develop a computationally efficient implementation of a two-stage optimization strategy for QoS routing. This strategy first selects feasible paths that minimize a QoS cost. As an example, this QoS cost may be the path's overall reservation of network bandwidth. However, the optimization method presented in this paper will work with a general class of QoS cost functions. Then, in the case that there are multiple paths that are tied with respect to this QoS cost, the second stage of the strategy is to select from these tied paths one that minimizes a measure of BE traffic delay. Finally, we demonstrate the practicality of our methodology by integrating it into a well-known algorithm for routing LSPs with path restoration.

Section II introduces our notations and definitions. Section III develops constraints and costs to account for BE traffic delay when routing LSPs. Our BE-friendly two-stage optimization strategy for QoS routing, along with an efficient implementation of it, are introduced in Section IV. Section V shows how our methodology can be applied to LSP routing with path restoration and includes simulation results. Finally, Sections VI and VII mention related work and conclude the paper, respectively.

II. PRELIMINARIES

The MPLS network domain is represented by the directed graph, $G = (V, E)$, where V is the set of vertices or nodes (symbolizing LSRs) and E is the set of directed links or edges. Each directed link, e, has bandwidth capacity C_e. We also refer to links by their end-nodes, e.g., (i, j) to represent the link directed from node i to node j. Each traffic trunk, t, is described by its ingress node, s_t, egress node, d_t, class of service, c_t, supporting LSP, p_t, and average transmission rate, b_t. The average transmission rate, b_t, of t is the sum of the average transmission rates of all the micro-flows in trunk t. An LSP or path, p, is defined as an ordered subset, $\{e_1, e_2, \ldots, e_{h(p)}\}$, of E where $h(p)$ is the number of hops along p and the order indicates the sequence of links along p from ingress node s to egress node d. We also refer informally to paths by their node sequences ordered from s to d, e.g., s-u-d.

We also associate with a trunk its effective bandwidth which we assume to account for the trunk's QoS requirements (e.g., constraints on packet delay, jitter, or loss) [2] and to represent the amount of bandwidth that is reserved for the trunk if it is routed. Although there are different definitions of effective bandwidth [3-9], in general, it is a value intermediate between the mean and peak flow rates of a connection or

trunk. It indicates the amount of bandwidth that must be reserved for a trunk in order to satisfy its QoS requirements. A trunk's effective bandwidth may be defined in terms of link operating points as well as the trunk's traffic characteristics [8, 9]. Therefore, we use α_e^t to refer to the effective bandwidth of trunk t at link e.

Let τ represent a traffic trunk that is requesting an LSP. Denoting τ's descriptors using τ in place of t, we assume that an LSP request provides the network admission control and routing decision maker with τ's transmission characteristics and CoS which imply b_τ. In general, to specify α_e^τ, knowledge of link e's operating point is also necessary.

For each directed link, $e \in E$, we define

$$R_e^{eff} \equiv C_e - \sum_{\{t | e \in p_t\}} \alpha_e^t, \qquad (1)$$

and refer to R_e^{eff} as the residual bandwidth of e expressed in terms of effective bandwidth. One of the necessary conditions for link e to be feasible for assignment to an LSP for trunk τ is that e must satisfy the residual effective bandwidth constraint,

$$\alpha_e^\tau \leq R_e^{eff}. \qquad (2)$$

A candidate path, p, under consideration for assignment to τ, is feasible for τ if and only if it is composed entirely of feasible links.

Let $Cost_e^\tau$ be a measure of the cost to the network of using link e in an LSP for trunk τ, exclusive of any effects on BE traffic. We refer to $Cost_e^\tau$ as the 'base cost' of e and also define

$$Cost_p^\tau \equiv \sum_{e \in p} Cost_e^\tau. \qquad (3)$$

as the corresponding base cost of selecting path p to support τ. We can leave the link base costs as general functions to some extent. However, we incorporate the residual effective bandwidth constraint into the cost functions by penalizing the $Cost_e^\tau$ to be infinite for any e where (2) is not satisfied. A simple example of a link base cost function is a hop count cost, implemented by defining

$$Cost_e^\tau \equiv \begin{cases} 1, & \text{if } \alpha_e^\tau \leq R_e^{eff} \\ \infty, & \text{otherwise.} \end{cases} \qquad (4)$$

Another link base cost function, which is aimed at minimizing a path's overall reservation of network bandwidth, is the bandwidth reservation cost, implemented by defining

$$Cost_e^\tau \equiv \begin{cases} \alpha_e^\tau, & \text{if } \alpha_e^\tau \leq R_e^{eff} \\ \infty, & \text{otherwise.} \end{cases} \qquad (5)$$

We limit the scope here to single path routing for each trunk. Although multipath routing may provide the potential to improve network resource utilization, it increases the complexity of the signaling process for LSP setup, increases the likelihood of out-of-order packet arrivals, and can eliminate the additivity property of effective bandwidths [10].

III. BE TRAFFIC CONSIDERATIONS

If BE traffic is not considered when routing LSPs for QoS traffic trunks, then admission and routing for an LSP request may consist of searching for a path that minimizes $Cost_p^\tau$ of (3). We next describe LSP routing constraints and costs to account for the effects of this routing on BE traffic.

A. BE Traffic Available Bandwidth

We let

$$C_e^{BE} \equiv C_e - \sum_{\{t | e \in p_t\}} b_t, \qquad (6)$$

and refer to C_e^{BE} as the average bandwidth that is available for serving BE traffic at link e. Note that, by the notion of effective bandwidth, a trunk's average bandwidth consumption does not exceed its effective bandwidth, i.e.,

$$\alpha_e^t \geq b_t, \quad \forall e \in p_t, \forall t. \qquad (7)$$

Then, by (1), (6), and (7), it follows that

$$C_e^{BE} \geq R_e^{eff}, \quad \forall e. \qquad (8)$$

Therefore, C_e^{BE}, the average bandwidth available for BE traffic, may include some bandwidth that is allocated to but, on average, is not consumed by LSPs. Here, we assume that the routers provide packet scheduling that enables BE traffic to take advantage of this excess effective bandwidth without interfering with the QoS trunks. For example, under such scheduling disciplines as weighted fair queueing (WFQ) [11] and deficit round robin (DRR) [12], low priority traffic can be queued separately from higher priority classes and serviced whenever the higher priority queues are empty. Such disciplines also provide the means of distributing to QoS trunks their allocated bandwidths.

B. BE Traffic Delay Metric

Our approach is to define a metric that is indicative of average BE traffic delay and to use this metric to define a BE traffic delay constraint that we apply in determining which paths are feasible for QoS routing. Also, in the case that there are multiple paths that minimize the base cost, $Cost_p^\tau$, we then select from them one that minimizes a BE cost that is associated with the BE traffic delay metric. In this paper, we demonstrate how to utilize two example BE traffic delay metrics, one based on an M/M/1 queueing model approximation and the other, described in Section III.E, based

on a G/G/1 queueing delay bound. Other delay metrics can be employed in a similar manner.

For the M/M/1 delay metric, we follow [13, Ch. 5], defining D_{BE} to approximate the average time that it takes for a BE packet to traverse the network from its source to destination. Let

$$D_{BE} \equiv \frac{1}{\gamma} \sum_{e \in E} M_e \qquad (9)$$

where

$$M_e \equiv \frac{F_e}{C_e^{BE} - F_e} + d_e F_e, \qquad (10)$$

F_e is the average BE traffic load at link e, d_e is the processing and propagation delay for e, and γ is the average total rate of BE traffic entering the network. Referring to (10), by Little's theorem, M_e represents the average amount of BE data that is either queued at or being transmitted across link e. Then $\sum_{e \in E} M_e$ in (9) represents the average total amount of BE data in the network and, again by Little's theorem, D_{BE} represents the average time a BE packet spends in the network. We assume that the parameters, γ, $\{d_e\}$, and $\{F_e\}$, are constant over an interval of interest and are approximately known. For example, F_e might represent the expected BE traffic load on e during a given busy hour of a weekday.

C. BE Traffic Residual Bandwidth Constraint

Implicit in (10), the definition of M_e, is the assumption that

$$C_e^{BE} > F_e, \quad \forall e \in E. \qquad (11)$$

To prevent excessive BE traffic delays, we impose the further constraint that the delay cost be limited, i.e., that

$$D_{BE} \leq D_{max} \qquad (12)$$

for some value, D_{max}. Moreover, from a practical point of view, we would like to guarantee a certain BE performance at each link. Therefore, we require that

$$M_e \leq \frac{\gamma D_{max}}{|E|}, \quad \forall e \in E, \qquad (13)$$

where we have selected the limit on the right-hand side of (13) such that (12) is satisfied if (13) is satisfied at each link. Then (10) and (13) imply that

$$C_e^{BE} \geq \frac{F_e |E|}{\gamma D_{max} - d_e F_e |E|} + F_e, \quad \forall e \in E. \qquad (14)$$

Suppose that, to support a new traffic trunk, τ, we are considering a candidate path that includes directed link e whose current available bandwidth for BE traffic is C_e^{BE}. If e

does end up supporting τ, then, by (6), its available bandwidth for BE traffic will become $C_e^{BE} - b_\tau$. Then the BE traffic delay cost bound will still be satisfied only if (14) holds with $C_e^{BE} - b_\tau$ substituted for C_e^{BE}, i.e.,

$$C_e^{BE} - b_\tau \geq \frac{F_e|E|}{\gamma D_{max} - d_e F_e|E|} + F_e. \tag{15}$$

Inequality (15) can be expressed as

$$b_\tau \leq R_e^{ave}, \tag{16}$$

where

$$R_e^{ave} \equiv C_e^{BE} - F_e - \frac{F_e|E|}{\gamma D_{max} - d_e F_e|E|}. \tag{17}$$

Like R_e^{eff}, R_e^{ave} represents a residual bandwidth capacity of directed link e. Whereas R_e^{eff} limits the effective bandwidth, α_e^τ, of a new trunk that link e can support, R_e^{ave} limits the average transmission rate, b_τ, of a new trunk that e can support. Therefore, to be feasible for supporting new trunk τ, directed link e must satisfy both residual bandwidth constraints, (2) and (16).

D. BE Traffic Cost of a Candidate Path

Again suppose that we select a path, p_τ, to support a new traffic trunk, τ. Then, for each $e \in p_\tau$, e's available bandwidth for BE traffic will change from C_e^{BE} to the new value, $C_e^{BE} - b_\tau$. Consequently, by (10), the new value of M_e for each $e \in p_\tau$ will be

$$\frac{F_e}{C_e^{BE} - b_\tau - F_e} + d_e F_e,$$

and, letting $1_{\{H\}}$ represent the indicator function, i.e.,

$$1_{\{H\}} = \begin{cases} 1, & \text{hypothesis } H \text{ is true,} \\ 0, & \text{otherwise,} \end{cases}$$

by (9), the new value of D_{BE} with p_τ selected will be

$$D_{BE} = \frac{1}{\gamma} \sum_{e \in E} \left(\frac{F_e}{C_e^{BE} - F_e - 1_{\{e \in p_\tau\}} b_\tau} + d_e F_e \right)$$

$$= \frac{1}{\gamma} \sum_{e \in p_\tau} \left(\frac{F_e}{C_e^{BE} - F_e - b_\tau} - \frac{F_e}{C_e^{BE} - F_e} \right)$$

$$+ \frac{1}{\gamma} \sum_{e \in E} \left(\frac{F_e}{C_e^{BE} - F_e} + d_e F_e \right).$$

Note that the last summation above is independent of path p_τ. Consequently, we define the BE traffic delay cost associated with a candidate path, p, under consideration for serving τ as

$$Cost_p^{BE} \equiv \sum_{e \in p} Cost_e^{BE}, \tag{18}$$

where

$$Cost_e^{BE} \equiv \begin{cases} \dfrac{F_e b_\tau}{\gamma \left(C_e^{BE} - b_t - F_e\right)\left(C_e^{BE} - F_e\right)}, & \text{if } b_\tau \leq R_e^{ave} \\ \infty, & \text{otherwise.} \end{cases} \tag{19}$$

$Cost_e^{BE}$, which we refer to as the BE traffic delay cost associated with using link e to support trunk τ, is the increase in the average BE delay that results if link e is used in an LSP to support trunk τ. Again, we define this cost to be infinite if its corresponding residual bandwidth constraint, (16) in this case, is not satisfied.

E. An Alternative BE Cost Function

A simple alternative to the approximation of using an M/M/1 model to represent BE traffic delay is to make use of a well-known (e.g., [13, Ch. 3]) upper bound on the time spent waiting in a G/G/1 queue. Letting W represent the average time that a customer waits in a G/G/1 queue, we have that $W \leq W^{ub}$ where

$$W^{ub} \equiv \frac{\lambda\left(\sigma_a^2 + \sigma_b^2\right)}{2\left(1 - \lambda/\mu\right)},$$

σ_a^2 is the variance of the customer interarrival time, and σ_b^2 is the variance of the customer service time. Now we define the cost metric,

$$W_{BE} \equiv \frac{1}{\gamma} \sum_{e \in E} F_e W_e \tag{20}$$

where

$$W_e \equiv \frac{F_e\left(\sigma_a^2 + \sigma_b^2\right)}{2\left(1 - F_e/C_e^{BE}\right)}, \tag{21}$$

we have we substituted F_e for λ and C_e^{BE} for μ, and σ_a^2 and σ_b^2 can be now used to characterize the arriving and departing traffic of the link, respectively. Letting W_{max} be the maximum allowable value of W_{BE}, we can ensure that $W_{BE} \leq W_{max}$ by requiring that

$$F_e W_e \leq \frac{\gamma W_{max}}{|E|}, \quad \forall e \in E, \tag{22}$$

which, along with (21), implies that

$$C_e^{BE} \geq \frac{2\gamma W_{max} F_e}{2\gamma W_{max} - F_e^2 |E| (\sigma_a^2 + \sigma_b^2)}, \ \forall \, e \in E. \quad (23)$$

In order for link e to support new traffic trunk τ with average transmission rate b_τ, (23) must hold with $C_e^{BE} - b_\tau$ substituted for C_e^{BE}. This condition is expressed by (16) where, in this case,

$$R_e^{ave} \equiv C_e^{BE} - \frac{2\gamma W_{max} F_e}{2\gamma W_{max} - F_e^2 |E| (\sigma_a^2 + \sigma_b^2)}. \quad (24)$$

We next define a BE traffic cost of a candidate path in terms of the G/G/1 bound. If path p_τ is selected to support new traffic trunk τ with average transmission rate b_τ, then, by (20) and (21) with $C_e^{BE} - b_\tau$ substituted for C_e^{BE}, the new value of W_{BE} will be

$$W_{BE} = \frac{F_e^2 (\sigma_a^2 + \sigma_b^2)}{2\gamma} \sum_{e \in p_\tau} \left(\frac{C_e^{BE} - b_\tau}{C_e^{BE} - b_\tau - F_e} - \frac{C_e^{BE}}{C_e^{BE} - F_e} \right)$$
$$+ \frac{F_e^2 (\sigma_a^2 + \sigma_b^2)}{2\gamma} \sum_{e \in E} \frac{C_e^{BE}}{C_e^{BE} - F_e}.$$

We take the path dependent part of W_{BE} as the BE traffic delay cost of candidate path, p, i.e., using the G/G/1 bound, $Cost_p^{BE}$ is given by (18) with

$$Cost_e^{BE} \equiv \begin{cases} \dfrac{F_e^3 (\sigma_a^2 + \sigma_b^2) b_\tau}{2\gamma (C_e^{BE} - b_t - F_e)(C_e^{BE} - F_e)}, & \text{if } b_\tau \leq R_e^{ave} \\ \infty, & \text{otherwise.} \end{cases} \quad (25)$$

IV. BE-FRIENDLY LSP ROUTING

We next describe our routing optimization strategy and an efficient method for implementing it. Then we present our BE-friendly routing with problem formulations and algorithms.

A. Optimization Strategy

An arriving LSP request for new trunk τ is admitted if there exists a feasible path from s to d, i.e., one composed entirely of links that satisfy residual bandwidth constraints (2) and (16). For the general case in which there are multiple feasible paths, we apply a two-stage optimization strategy (which is conceptually similar to one in [14]) for path selection. In the first stage, we identify all feasible paths that achieve a finite minimum value of the base cost, $Cost_p^\tau$, of (3). Then we apply the second stage, which is to select from all such paths one that achieves a finite minimum value of $Cost_p^{BE}$ of (18). In the case that such a path exists, we assign it to the LSP to support new trunk τ.

This optimization strategy can be implemented directly. However, if a certain quantization condition is satisfied, a more computationally efficient implementation is possible. Suppose that the base cost of any directed link e can be expressed as a nonnegative integer multiple of a positive real, i.e., for some quantum, $q > 0$,

$$Cost_e^\tau \in \{nq \mid n \in Z^+\} \cup \{\infty\}. \quad (26)$$

Condition (26) clearly holds for the simple example of (4) for $q = 1$ and $n = 1$. More generally, by selecting some q sufficiently small, it may be possible to approximate base cost by a quantized version of $Cost_e^\tau$ without having any appreciable effect on the resulting routing process.

Whenever (26) holds, we can use the following method to indirectly implement our two-stage optimization. Consider a path cost that is a weighted sum of base and BE costs, i.e., define

$$Cost_p^{WS} \equiv Cost_p^\tau + w_{BE} Cost_p^{BE}, \quad (27)$$

where w_{BE} is a positive weighting coefficient for the BE cost (with unity weighting of $Cost_p^\tau$). A path that minimizes $Cost_p^{WS}$ does not necessarily minimize $Cost_p^\tau$. However, if we select a value for w_{BE} that is sufficiently small, then we can insure that any path that achieves a finite minimum value of $Cost_p^{WS}$ also minimizes $Cost_p^\tau$. Specifically, we have the following

Proposition. Suppose that condition (26) holds for $q > 0$ and that $w_{BE} > 0$ is such that

$$w_{BE} Cost_p^{BE} < q, \text{ for any feasible candidate path, } p. \quad (28)$$

If feasible path p_1 is such that

$$Cost_{p_1}^{WS} = \min_{\{p \text{ feasible}\}} Cost_p^{WS}, \quad (29)$$

then it follows that

$$Cost_{p_1}^\tau = \min_{\{p \text{ feasible}\}} Cost_p^\tau.$$

Proof: Suppose that (29) holds and that there exists feasible path p_2 such that

$$Cost_{p_2}^\tau < Cost_{p_1}^\tau.$$

Condition (26) then implies that

$$Cost_{p_2}^\tau \leq Cost_{p_1}^\tau - q. \quad (30)$$

Then

$$Cost_{p_2}^{WS} \;=\; Cost_{p_2}^{\tau} + w_{BE}\, Cost_{p_2}^{BE} \qquad \text{(by (27))}$$

$$\le\; Cost_{p_1}^{\tau} - q + w_{BE}\, Cost_{p_2}^{BE} \qquad \text{(by (30))}$$

$$<\; Cost_{p_1}^{\tau} \qquad \text{(by (28))}$$

$$\le\; Cost_{p_1}^{WS}, \qquad \text{(by (27))},$$

contradicting (29). ∎

Furthermore, since w_{BE} is positive, if path p is a minimum-base-cost path (i.e., it minimizes $Cost_p^{\tau}$) and p minimizes (27), then p has the minimum BE cost among all minimum base cost paths.

Now we comment on the values of w_{BE} that satisfy condition (28). It is the BE residual bandwidth constraint, (16), that allows us to select w_{BE} such that condition (28) is satisfied. That is, (16) forces the BE traffic delay metric, and therefore $Cost_p^{BE}$, to be upper bounded. Defining for any candidate path, p, the variables,

$$x_e \;\equiv\; 1_{\{e \in p\}}, \quad \forall e \in E, \qquad (31)$$

where $1_{\{H\}}$ is again the indicator function for hypothesis H, by (18), condition (28) is equivalent to

$$w_{BE} \sum_{e \in E} x_e\, Cost_e^{BE} \;<\; q. \qquad (32)$$

For the case that the BE residual bandwidth is given by (17) (M/M/1 approximation), it follows that, for any feasible link, e,

$$C_e^{BE} - F_e - b_\tau \;\ge\; \frac{F_e|E|}{\gamma D_{max} - d_e F_e |E|}, \qquad (33)$$

and

$$C_e^{BE} - F_e \;\ge\; b_\tau. \qquad (34)$$

Therefore, the BE cost of using feasible link e is upper bounded as

$$Cost_e^{BE} \;\le\; \frac{F_e b_\tau}{\gamma \left(C_e^{BE} - b_t - F_e \right)\left(C_e^{BE} - F_e \right)} \qquad \text{(by (19))}$$

$$\le\; \frac{D_{max}}{|E|} - \frac{d_e F_e}{\gamma} \qquad \text{(by (33) and (34))}$$

$$\le\; \frac{D_{max}}{|E|}, \quad \forall e \in E. \qquad (35)$$

Then, by selecting

$$w_{BE} \;\equiv\; \frac{\delta q}{D_{max}} \qquad (36)$$

for some $\delta \in (0,\,1)$, we are assured that the BE cost contribution to $Cost_p^{WS}$ for any feasible path, p, is less than q because

$$w_{BE} \sum_{e \in E} x_e\, Cost_e^{BE} \;\le\; w_{BE}\, |E|\, \max_{e \in E} Cost_e^{BE}$$

$$\le\; w_{BE}\, D_{max} \quad \text{(by (35))}$$

$$<\; q \qquad \text{(by (36)).} \qquad (37)$$

Alternatively, for the case that the BE residual bandwidth is given by (24) (G/G/1 bound), a similar argument shows that condition (28) is satisfied if w_{BE} is specified as

$$w_{BE} \;\equiv\; \delta q\, \frac{2\gamma}{2\gamma W_{max} - F_e^{\,2}|E|(\sigma_a^2 + \sigma_b^2)} \qquad (38)$$

for some $\delta \in (0,\,1)$.

B.   LSP Routing Problem Formulation

Assuming that conditions (26) and (28) hold, the two-stage optimization for routing an LSP to support a new trunk can be described as the selection of $\{x_e\}$ that minimizes

$$Cost_p^{WS} \;=\; \sum_{e \in E} x_e \left(Cost_e^{\tau} + w_{BE}\, Cost_e^{BE} \right) \qquad (39)$$

subject to

$$\sum_{j \in V} x_{(i,j)} - \sum_{j \in V} x_{(j,i)} \;=\; 0, \quad i \ne s, d \qquad (40)$$

$$\sum_{j \in V} x_{(s,j)} - \sum_{j \in V} x_{(j,s)} \;=\; 1, \qquad (41)$$

$$\sum_{j \in V} x_{(d,j)} - \sum_{j \in V} x_{(j,d)} \;=\; -1, \qquad (42)$$

$$x_e \in \{0,\,1\}, \quad \forall e \in E, \qquad (43)$$

where x_e is defined by (31), s and d are the ingress and egress nodes, respectively, $Cost_e^{BE}$ is given by (19), and the base cost is defined so that $Cost_e^{\tau} = \infty$ for any link e with $\alpha_e^{\tau} > R_e^{eff}$. The flow conservation constraints, (40) for all vertices in V other than s and d, (41) for s, and (42) for d, force $\{x_e\}$ to describe a single path from s to d.

There may be other constraints placed on this path selection problem, e.g., administrative or service class constraints. An example of an administrative constraint is a restriction on what links can be used in the path. If only links in $E' \subseteq E$ are admissible, then we substitute E' for E in (39)-(43). An example of a service class constraint is a maximum

hop count, $H(c_\tau)$, associated with the CoS, c_τ, of the trunk requesting an LSP [14-16]. This constraint can be accommodated by adding the inequality,

$$\sum_{e \in E} x_e \leq H(c_\tau), \tag{44}$$

to problem formulation (39)-(43).

If there is no feasible $\{x_e\}$ for the problem formulation, then the LSP request is rejected. Otherwise, an optimal path to support new trunk τ is constructed as the ordered set,

$$p_\tau \equiv \{(v_n, v_{n+1}) \mid n = 1, 2, \ldots, \sum_e x_e^*\}, \tag{45}$$

where $\{x_e^*\}$ is an optimal solution to (39)-(43), $v_1 = s$, and v_{n+1} is the unique vertex such that $x_{(v_n, v_{n+1})}^* = 1$.

C. LSP Routing Algorithm

Although problem formulation (39)-(43) is a 0-1 integer programming problem, it has a special structure that lends itself to solution by Dijkstra's algorithm. By defining as our link metric,

$$Cost_e^{WS} \equiv Cost_e^{\tau} + w_{BE} \, Cost_e^{BE}, \tag{46}$$

we have that a path from s to d with shortest possible length under metric $Cost_e^{WS}$ also minimizes $Cost_p^{WS}$ of (39).

The algorithm, WS-OPT, shown in Fig. 2, computes metric $Cost_e^{WS}$ for each link and then applies Dijkstra's algorithm to search for an optimal path to assign to the LSP for trunk τ. In this figure, we adopt the pseudocode conventions of [17], omitting **end** statements and indicating block structures solely by indentation.

WS-OPT($V, E, s, d, \alpha_e^\tau, b_\tau, \{R_e^{eff}\}, \{R_e^{ave}\}, \{F_e\}, D_{max}, \{d_e\}, w_{BE}$)

1   **for** each $e \in E$
2   **do**   compute $Cost_e^\tau$, e.g., by (4) or (5)
              ($Cost_e^\tau \leftarrow \infty$ if $\alpha_e^\tau > R_e^{eff}$)
3   compute $Cost_e^{BE}$, e.g., by (19) or (25)
              ($Cost_e^{BE} \leftarrow \infty$ if $b_\tau > R_e^{ave}$)
4           $Cost_e^{WS} \leftarrow Cost_e^\tau + w_{BE} \, Cost_e^{BE}$
5   $p_\tau \leftarrow$ Dijkstra's algorithm for (V, E), s, d, and $\{Cost_e^{WS}\}$
6   **if** no such p_τ exists
7   **then** reject LSP request
8   **else** admit LSP request, assign path p_τ

Figure 2. WS-OPT to implement weighted sum optimization.

A link set satisfying administrative constraints, $E' \subseteq E$, can be substituted for E in WS-OPT. If there is a maximum hop count service class constraint, $H(c_\tau)$, it can be accommodated with a slight modification of Dijkstra's algorithm. Consider its main **while** loop, which continues iterating while the priority queue of vertices to be processed remains nonempty (e.g., [17]). If we limit the number of loop iterations to be equal to

$H(c_\tau)$, then the algorithm will discover only those paths that do not exceed $H(c_\tau)$ hops.

Regarding a time bound for WS-OPT, lines 1-4 take time $O(|E|)$ and Dijkstra's algorithm can be run in time $O(|E| + |V| \log |V|)$ if its vertex priority queue is implemented as a Fibonacci heap [17]. Therefore, WS-OPT takes time $O(|E| + |V| \log |V|)$. By comparison, a direct implementation of the two-stage optimization would require an exhaustive search for all minimum-base-cost paths, followed by selection of a minimum-base-cost path with minimum BE cost. Thus, if conditions (26) and (28) hold, then indirect implementation of the two-stage optimization by WS-OPT significantly reduces computation complexity since it only applies Dijkstra's algorithm once.

D. Path-based LSP Routing

The problem formulation and algorithm presented in the previous two subsections are for a link-based implementation of BE-friendly LSP routing. This approach assumes knowledge of the network topology for routing LSP requests. An alternative approach is path-based LSP routing. In this case, upon arrival of an LSP request, a path is selected from a fixed pre-determined set of candidate paths for the ingress-egress pair and service class. Path-based routing may be practical in a very sparsely connected network, or when administrative or service class constraints have the effect of winnowing down the candidate paths to a manageable number.

Representing the candidate path set by P and letting

$$z_p \equiv 1_{\{\text{path } p \text{ is selected}\}}, \quad \forall p \in P, \tag{47}$$

we apply the first stage the two-stage optimization for path-based routing by selecting $\{z_p\}$ to minimize

$$\sum_{p \in P} z_p \, Cost_p^\tau \tag{48}$$

subject to

$$\sum_{p \in P} z_p = 1, \tag{49}$$

$$z_p \in \{0, 1\}, \quad \forall p \in P, \tag{50}$$

where constraint (49) forces selection of a single path. If (48) is infinite for all possible selections of $\{z_p\}$, then we reject the LSP request. Otherwise, we let $\{z_p^*\}$ represent an optimal solution to (48)-(50) and implement the second optimization stage by selecting a new set, $\{z_p\}$, that minimizes

$$\sum_{p \in P} z_p \, Cost_p^{BE} \tag{51}$$

subject to (49), (50), and

$$\sum_{p \in P} z_p \, Cost_p^\tau = \sum_{p \in P} z_p^* \, Cost_p^\tau. \tag{52}$$

The 0-1 integer program of (48)-(52) is readily solved by a straightforward algorithm that (i) computes the base cost, $Cost_p^{\tau}$, of each path in P, (ii) forms set $P*$ containing all paths in P that have finite minimum base cost, (iii) computes the BE cost, $Cost_p^{BE}$, of each path in $P*$, and (iv) selects a path in $P*$ that has finite minimum BE cost (if it exists). This algorithm takes time $O(|P|\,|E|)$.

V.   BE-FRIENDLY ROUTING WITH PATH RESTORATION

A.   Application to PI-ER Routing

In this section, we illustrate that our BE-friendly QoS-routing can be combined with restorable routing. Much recent work has been done on restorable QoS routing, e.g., [18-20]. Reference [20] describes an algorithm called partial information with exact reservations (PI-ER) which dynamically routes LSPs with 1:1 path protection. The term, dynamic routing, refers to routing LSP requests that arrive one-by-one; 1:1 path protection refers to supporting each active path with a link-disjoint backup path that is used only if the active path fails. In protecting against single link failures, the PI-ER algorithm minimizes the backup bandwidth reservations by sharing bandwidth among multiple backup paths. For example, let us suppose that there are only two active LSPs in the network. The algorithm may use a particular link, e_b, in the backup path for each of two trunks. If these trunks have mutually link-disjoint active paths, then the algorithm reserves an amount of backup bandwidth at e_b equal to the bandwidth requirement of the trunk that requires the most bandwidth. This reservation is enough for link e_b to support either trunk along its backup path if a single link failure affects the trunk's active path. On the other hand, if the two trunks' active paths share a common link, e_a, then the algorithm reserves enough backup bandwidth at link e_b to support the sum of the trunks' bandwidth requirements. This reservation protects the trunks in the event that link e_a fails. The 'partial information' used by the PI-ER algorithm consists of the following for each link: the total bandwidth reservation for active paths, the total bandwidth reservation for backup paths, and the residual bandwidth. The link's residual bandwidth is that part of its bandwidth (capacity) that is not reserved for active or backup paths. 'Exact reservations' refers to how the bandwidth reservations for a trunk's backup path, once its active and backup paths have been routed, are exactly what the trunk requires. (However, when actually routing a trunk, the PI-ER algorithm makes use of upper bounds on the backup bandwidth reservations that the trunk will require.) More details are referred to [19,20].

We apply our link-based, two-stage optimization approach to constrain PI-ER routing of active paths to be BE-friendly. We do not impose BE-friendly routing of the backup paths as they are used only in the event of a link failure. Given an LSP request for a new QoS traffic trunk, τ, the routing optimization, framed as a 0-1 integer programming problem as in [19, 20], is to minimize

$$\Sigma_{e\in E}\, a_e\, x_e + \Sigma_{e\in E}\, c_e\, y_e, \tag{53}$$

subject to (40)-(43), analogous constraints for $\{y_e\}$, and

$$x_e + y_e \le 1, \forall e \in E, \tag{54}$$

where x_e (y_e) is one if link e lies in the active (backup) path and zero otherwise, and coefficients a_e and c_e, for link e, both depend on the trunk's bandwidth requirement which we take to be its effective bandwidth. Both a_e and c_e are calculated based on a variable, M, which is iterated by the PI-ER algorithm throughout the set of values of the net bandwidth reservation for active paths at each link. Coefficient c_e is also a function of G_e, which represents the current bandwidth reservation at e for supporting backup paths. In our BE-friendly version of the PI-ER algorithm, we make the simplifying assumption (as in [19, 20]) that the bandwidth requirement of each trunk does not depend on link operating points. That is, we assume that

$$\alpha_e{}^t = \alpha^t, \forall e \in p_t, \forall t.$$

We also suppose that M, α^{τ}, α^t $\forall t$, and G_e are all quantized with quantum, q. In this case, for BE-friendly active path routing with application of residual bandwidth constraints (2) and (16), we define coefficient a_e as

$$a_e \equiv \begin{cases} \alpha^{\tau} + w_{BE}Cost_e^{BE}, & \text{if } \sum_{\{t|e\in p_t\}}\alpha^t \le M \\ & \text{and (2) and (16) hold} \\ \infty, & \text{otherwise.} \end{cases} \tag{55}$$

For example, $Cost_e^{BE}$ is defined by (19) and w_{BE} is defined by (36) for the M/M/1 approximation or (38) for the G/G/1 bound. As in [19, 20],

$$c_e \equiv \begin{cases} 0, & \text{if } M+\alpha^{\tau} \le G_e \\ M+\alpha^{\tau}-G_e, & \text{if } M+\alpha^{\tau} > G_e \\ & \text{and } M+\alpha^{\tau}-G_e \le R_e^{eff} \\ \infty, & \text{otherwise.} \end{cases}$$

Here, and in (2) as it is used in (55), the residual effective bandwidth is

$$R_e^{eff} \equiv C_e - \sum_{\{t|e\in p_t\}}\alpha^t - G_e, \forall e \in E. \tag{56}$$

The dual-based solution algorithm of [19] is then applied.[1]

---

[1] In applying this algorithm, we truncate to multiples of q each of the summands that are summed to form the reduced cost, r_{ij}, of [19]. This technical detail prevents BE cost from influencing the dual variables, $\{\sigma_{ij}\}$, of [19]. Then BE cost does not influence the selection of backup paths.

B. Simulation of BE-Friendly LSP Routing

To illustrate operation of a BE-friendly PI-ER algorithm, we use a 15-node test network configuration [19, 20] shown in Fig. 3 where, in this case, we let the bandwidth capacity of each directed link be 100 Mb/s. (The undirected links shown represent links that are directed in both directions.) Figures 4 and 5 depict results from multiple simulations of a scenario with Poisson arrivals of QoS LSP requests and exponentially distributed LSP holding times. In these figures, the blocking probability (BP) faced by an arriving LSP request is plotted versus the average BE traffic load, F_e, which we take to be common to all links. The parameter, ρ, is the offered load in Erlangs (E), i.e., the ratio of the LSP request arrival rate to the departure rate of each LSP. In these simulations, we approximate steady state conditions by running each simulation long enough for 30,000 LSP requests to arrive. The ingress-egress pair of each LSP request is randomly chosen with uniform probability for each pair. We assume that the effective bandwidth for each LSP request is uniformly distributed over {1, 2, …, 15} Mb/s. In computing coefficient a_e by (55), we use R_e^{ave} as defined by (17) (M/M/1 approximation) in constraint (16). Other simulation details are that we set $q = 1$ Mb/s, $\delta = \frac{1}{2}$, $D_{max} = 1$ s, $d_e = 0 \; \forall \; e$, and assume that BE effort packets traverse an average of 3 hops through the network so that $\gamma = (1/3) \, \Sigma_e \, F_e$.
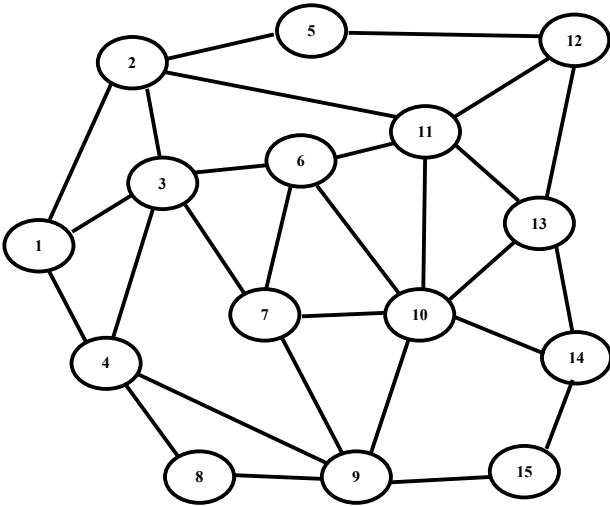

Figure 3. A 15-node test network configuration [19, 20].

As shown in Fig. 4, which is for the case of $\alpha^t = b_t \; \forall e \in p_t, \; \forall t$, there is not much of a penalty in terms of BP to implementing BE-friendly routing (i.e., specifying coefficient a_e by (55)) as F_e increases up to a certain point. For example, for $\rho = 40$ E, the BP is well under 1% for all BE traffic loads up to 50 Mb/s. Beyond that point, the BP deteriorates rapidly. Whereas Fig. 4 is for $\alpha^t = b_t \; \forall e \in p_t, \; \forall t$. Fig. 5 shows the corresponding results for $\alpha^t = 2b_t \; \forall e \in p_t, \; \forall t$. In this case, the BP of LSPs is essentially constant over an even wider range of BE traffic loads. For example, for $\rho = 40$ E, the BP is well under 1% for all BE traffic loads up to 75 Mb/s (as compared to 50 Mb/s in Fig. 4). Therefore, the Fig. 5 results suggest

that, by explicitly accounting for excess effective bandwidth, we can significantly reduce the BP penalty of implementing BE-friendly QoS routing.
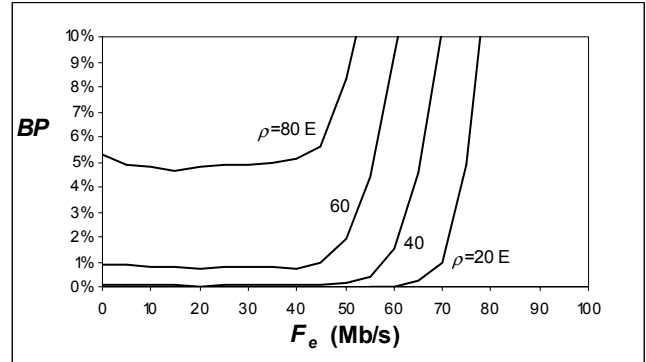

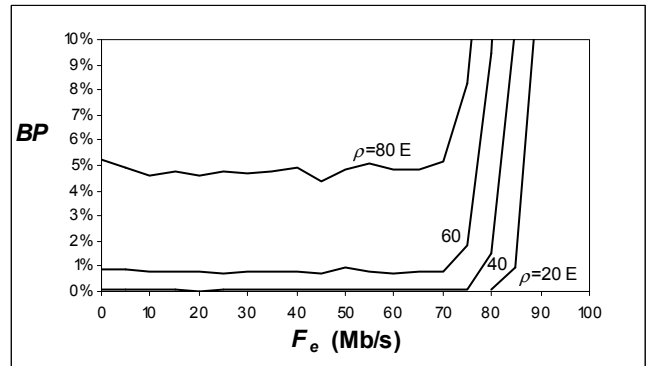Figure 4. *BP* versus F_e for $\alpha^t = b_t \; \forall e \in p_t, \forall t$.

Figure 5. *BP* versus F_e for $\alpha^t = 2b_t \; \forall e \in p_t, \forall t$.

VI. RELATED WORK

There are some works that have addressed the consequences of QoS routing on BE traffic. The virtual residual bandwidth (VRB) concept [21-23] is to adjust each link's actual residual bandwidth (bandwidth unassigned to QoS connections) to reflect the BE traffic load on the link. A link's VRB is made lower (higher) than its residual bandwidth if the BE traffic load on the link is heavy (light). QoS connections can then be routed via shortest paths using link costs that are functions of the VRB, e.g., the reciprocal of the VRB [21, 22]. A related virtual cost technique is also used in [23]. Another approach, framed in a setting with connection-less, hop-by-hop QoS routing, is the enhanced bandwidth-inversion shortest path algorithm [24]. It uses a link cost that is the reciprocal of the residual bandwidth scaled so as to penalize longer paths. Finally, [14] presents a problem formulation for routing batches of QoS and BE demands that maximizes revenues earned by QoS demands and, secondarily, maximizes revenues earned by BE demands.

The methodology that we present in this paper is directed toward connection-oriented QoS routing that uses any link metric. We propose a constraint on QoS routing to preserve a minimum level of performance for low priority BE traffic. This constraint, which we state in terms of a (second) residual link bandwidth, is to keep the average BE delay from exceeding some maximum acceptable limit. We mitigate the impact of this constraint on QoS routing by explicitly accounting for 'excess effective bandwidth.' Also, for the case that a quantization condition applies, we propose an efficient implementation of a two-stage optimization strategy that minimizes the cost of QoS routing and, secondarily, minimizes BE delay.

VII. Conclusions

We have proposed, for integration into effective-bandwidth-based QoS routing algorithms, mechanisms to quantitatively account for the impact that any given path selection has on low priority BE traffic. With this methodology, the degree to which path routing is restricted by BE-friendly constraints can be traded off, via a few parameters, with BE traffic delays. Our simulation results for a simple test network configuration suggest that degradation of QoS routing performance (e.g., increased LSP blocking probability) due to a BE-friendly constraint need not be prohibitive, especially when this constraint accounts for excess effective bandwidth being utilized by BE traffic. Our two-stage optimization strategy selects from among otherwise equal-cost paths (that acknowledge a residual bandwidth constraint due to BE traffic), one that has the least impact on BE delay. We have shown that, assuming quantization of QoS link costs, the two-stage optimization strategy can be implemented very efficiently, i.e., with the complexity of Dijkstra's algorithm.

References

[1] E. Rosen, A. Viswanathan, R. Callon, "Multiprotocol Label Switching Architecture," *IETF, RFC 3031*, January 2001.

[2] K. Kar, M. Kodialam, T. V. Lakshman, "Minimum interference routing of bandwidth guaranteed tunnels with MPLS traffic engineering applications," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 12, December 2000.

[3] G. de Veciana, G. Kesidis, J. Walrand, "Resource management in wide-area ATM networks using effective bandwidths," *IEEE JSAC*, vol. 13, pp. 1081-90, 1995.

[4] W. Whitt, "Tail probability with statistical multiplexing and effective bandwidths in multi-class queues," *Telecommun. Syst.*, vol. 2, pp. 71-107, 1993.

[5] A. I. Elwalid, D. Mitra, "Effective bandwidth of general Markovian traffic sources and admission control of high speed networks," *IEEE/ACM Transactions on Networking*, vol. 1, no. 3, pp. 329-343, June 1993.

[6] F. P. Kelly, "Notes on effective bandwidths," in *Stochastic Networks*, F. P. Kelly, S. Zachary, I. Ziedins, Eds., Oxford: Clarendon Press, pp. 141-68, 1996.

[7] C. S. Chang, J. A. Thomas, "Effective bandwidths in high-speed digital networks," *IEEE JSAC*, vol. 13, pp. 1091-1100, 1995.

[8] F. P. Kelly, "Effective bandwidths at multi-class queues," *Queueing Systems*, vol. 9, no. 1, pp. 5-16, 1991.

[9] C. Courcoubetis, V.A. Siris, G. Stamoulis, "Application of the many sources asymptotic and effective bandwidths for traffic engineering," *Telecommunications Systems*, vol. 12, pp. 167-191, 1999.

[10] C. Courcoubetis, V. A. Siris, "Managing and pricing service level agreements for differentiated services," *Proceedings of IWQoS '99*, June 1999.

[11] A. Demers, S. Keshav, S. Shenker, "Analysis and Simulation of a Fair-queueing Algorithm," *Proc. ACM SIGCOMM '89* (September 1989): 1–12.

[12] M. Shreedhar, G. Varghese, "Efficient Fair Queueing Using Deficit Round Robin," *Proc. ACM SIGCOMM '95*, vol. 25, no. 4 (October 1995):231–242.

[13] D. Bertsekas, R. Gallager, *Data Networks*, Englewood Cliffs, NJ: Prentice-Hall; 1992.

[14] D. Mitra, K. G. Ramakrishnan, "A case study of multiservice, multipriority traffic engineering design for data networks," *Proceedings of Globecom '99*, December 1999.

[15] M. K. Girish, B. Zhou, J. Q. Hu, "Formulation of the traffic engineering problems in MPLS based IP networks," *Proceedings of ISCC 2000*, July 2000.

[16] Y. Seok, Y. Lee, Y. Choi, C. Kim, "Dynamic constrained multipath routing for MPLS networks," *IEEE ICCCN 2001*, October 2001.

[17] T. H. Cormen, C. E. Leiserson, R. L. Rivest, *Introduction to algorithms*, 2nd ed., Cambridge, Mass.: MIT Press; Boston : McGraw-Hill; 2001.

[18] C. Huang, V. Sharma, "Building reliable MPLS networks using a path protection mechanism," *IEEE Communications Magazine*, vol. 40, no. 3, March 2002.

[19] M. Kodialam, T.V. Lakshman, "Dynamic routing of bandwidth guaranteed tunnels with restoration," *Proceedings of IEEE INFOCOM 2000*, vol. 2.

[20] M. Kodialam, T.V. Lakshman, "Restorable dynamic quality of service routing," *IEEE Communications Magazine*, vol. 40, no. 6, June 2002.

[21] Q. Ma, P. Steenkiste, "Supporting dynamic inter-class resource sharing: a multi-class QoS routing algorithm," *IEEE INFOCOM '99*, vol. 2, pp. 649-660, 1999.

[22] H. Kochkar, T. Ikenaga, Y. Oie, "QoS routing algorithm based on multiclasses traffic load," *IEEE GLOBECOM '01*, vol. 4, pp. 2193–2198, 2001.

[23] Y.-W. Chen, R.-H. Hwang, "QoS routing algorithms for multiple traffic classes," *IEEE ICC 2002*, vol. 4, pp. 2217-2221, 2002.

[24] J. Wang, K. Nahrstedt, "Hop-by-hop routing algorithms for premium-class traffic in diffserv networks," *IEEE INFOCOM 2002*, vol. 2, pp. 705–714, 2002.