Integration of Explicit Effective-Bandwidth-Based QoS Routing With Best-Effort Routing

Stephen L. Spitler, *Member, IEEE*, and Daniel C. Lee, *Senior Member, IEEE*

***Abstract*—This paper presents a methodology for protecting low-priority best-effort (BE) traffic in a network domain that provides both virtual-circuit routing with bandwidth reservation for QoS traffic and datagram routing for BE traffic. When a QoS virtual circuit is established, bandwidths amounting to the traffic's effective bandwidths are reserved along the links. We formulate a new QoS-virtual-circuit admission control and routing policy that sustains a minimum level of BE performance. In response to a QoS connection request, the policy executes a two-stage optimization. The first stage seeks a minimum-net-effective-bandwidth reservation path that satisfies a BE protecting constraint; the second stage is a tie-breaking rule, selecting from tied paths one that least disturbs BE traffic. Our novel policy implementation efficiently executes both optimization stages simultaneously by a single run of Dijkstra's algorithm. According to simulation results, within a practical operating range, the consideration that our proposed policy gives to the BE service does not increase the blocking probability of a QoS connection request.**

***Index Terms*—Best-effort (BE) traffic, constraint-based routing, dynamic routing, effective bandwidth, quality of service (QoS).**

I. Introduction

IN traditional best-effort (BE) Internet Protocol (IP) routing, each packet typically seeks a shortest path to its destination through connectionless hop-by-hop routing (datagram routing). Explicit virtual-circuit (connection-oriented) routing capability, such as Multiprotocol Label Switching (MPLS) [1], and signaling protocols (e.g., RSVP-TE [2] or CR-LDP [3]) have been subsequently added. These capabilities facilitate implementation of traffic engineering (TE) and routing with quality-of-service (QoS) guarantees. In this paper, we consider a network administrative domain, e.g., an Internet service provider (ISP) network, that supports both virtual-circuit routing with bandwidth reservation for high-priority QoS traffic, and connectionless, hop-by-hop routing for low-priority BE traffic. As noted in [4] and [5], in a network domain supporting both QoS routing and BE routing, the BE performance can suffer if QoS routing is performed without regard for the low-priority BE traffic; that is, routing of QoS connections through links that are heavily utilized by BE flows can cause or aggravate BE traffic congestion. (We assume strict priority of QoS traffic over the BE traffic in the queueing discipline.) For example, referring to Fig. 1, routing of a new QoS connection from node s to node d through the minimum-hop path, $s - u - d$, although efficient without regard to the BE traffic, may congest the BE traffic flowing through links (s, u) and (u, d). Although the BE traffic could adapt its hop-by-hop routing to traffic conditions in order to bypass congested links, such adaptation raises the possibility of traffic oscillation. Under the Open Shortest Path First (OSPF) protocol, the shortest paths that BE packets follow may be determined solely on the basis of link bandwidths [6], with no consideration of traffic loads.

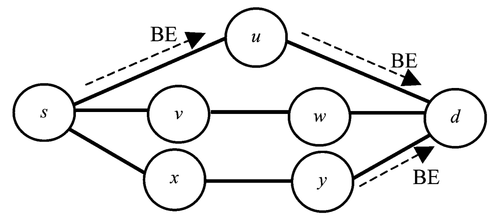

Fig. 1. Example network with heavy BE traffic on some links marked by dotted arrows.

This paper presents a practical methodology for sustaining a minimum level of BE performance in a network domain that also supports QoS routing for higher-priority traffic. This capability has relevance because BE traffic may remain a large part of a network's total traffic even after QoS routing is deployed. To emphasize the applicability of our methods, we will often use MPLS terminology in describing the QoS traffic and the connection-oriented QoS routing. However, the applicability of our methods is not limited to MPLS—our methods can be applied to any network that practices both datagram routing and virtual-circuit routing. What is described by the term, label switched path (LSP), in this paper can be regarded as any path (virtual circuit) and bandwidth provision along that path for supporting a microflow or set of microflows with a certain QoS requirement regarding, e.g., packet delay, jitter, or loss.

Our approach to protecting the BE traffic is to design a combined QoS connection admission control and routing policy that considers the impact of the QoS flows on BE traffic. We do not alter the BE routing. In our approach, an additional constraint is placed on the path (LSP) selection that occurs in response to a connection request for QoS traffic. This new QoS routing constraint is aimed at assuring a minimum performance level for the BE service in each link. In response to a QoS connection request, the network selects a path that satisfies the constraint. If no QoS path satisfying the constraints exists, then the connection request is denied. (In this paper, we limit the scope to a

Manuscript received December 13, 2004; revised June 21, 2006, and January 10, 2007; first published March 12, 2008; last published August 15, 2008 (projected); approved by IEEE/ACM Transactions on Networking Editor Z.-L. Zhang. This work was presented in part at IEEE INFOCOM 2003.

S. L. Spitler was with the Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089 USA. He is now with Boeing Company, El Segundo, CA 90245 USA.

D. C. Lee is with the School of Engineering Science, Simon Fraser University, Burnaby, BC, V5A 1S6, Canada (e-mail: dchlee@sfu.ca).

Digital Object Identifier 10.1109/TNET.2007.911425

greedy admission policy.) We then have the goals of quantitatively designing the BE protecting constraint, and finding a QoS path selection algorithm that satisfies this constraint without significantly degrading the resulting service of the QoS traffic. In particular, it is undesirable to increase the QoS call blocking probability, i.e., the probability that a request for a new QoS connection is denied.

With these goals in mind, we formulate a new two-stage network optimization problem [7] for selecting a path in response to a QoS connection request. The objective of the first-stage optimization is to find a path that consumes the minimum total effective bandwidth (the summation of bandwidths to be reserved over all the links in the path)—the rationale is to let the network retain available bandwidth for future QoS connection requests as much as possible. Included in this first-stage optimization problem is our constraint for protecting BE traffic. A key ingredient in our formulation is what we call excess effective bandwidth. We assume here that a QoS connection's effective bandwidth takes into account its QoS requirements [8] and represents the amount of bandwidth that should be reserved for the connection. Excess effective bandwidth then refers to the margin by which a QoS connection's effective bandwidth exceeds its average bandwidth usage. There are queueing disciplines, e.g., weighted fair queueing (WFQ) [9] and deficit round robin (DRR) [10], that enable BE traffic to take advantage of excess effective bandwidth when that bandwidth is not used by QoS traffic. Accounting for this efficiency enables us to design our BE protecting constraint to allow more bandwidth to be reserved by QoS connections, while still providing bandwidth for BE traffic.

The second stage of our two-stage optimization for routing QoS connections functions as a tie-breaking rule. Given a connection request, if there are multiple candidate paths that have the same total effective bandwidth and satisfy the BE protecting constraint, the second-stage optimization selects from among them one that is most favorable to the performance of the BE service. For example, referring again to Fig. 1, suppose that candidate paths s-v-w-d and s-x-y-d are tied for optimality in the first-stage optimization for routing a new QoS connection from node s to node d. Then the second-stage optimization selects path s-v-w-d because this selection least disturbs BE traffic. (The two-stage optimization is similar in concept to that in [11] for batch routing.)

Note that the QoS routing and admission control decisions are dynamically made in response to QoS connection requests, and are made based on different instances of this two-stage optimization because the available bandwidths for links are different at the times of different connection requests. An important issue to be addressed in this paper is the blocking probability associated with this routing and admission control policy, which employs the two-stage optimization described above. An interesting simulation result to be presented in this paper is that, within a specific operating range, the consideration that our proposed policy gives to the BE service does not increase the blocking probability of a QoS connection request. As long as the policy is designed to protect BE traffic loads that do not exceed a particular threshold, the blocking probability is essentially the same with our BE protecting constraint imposed as it is without this QoS routing constraint. On the other hand, above this threshold, the blocking probability of a QoS connection request steeply increases. Section V will elaborate on this phase-transition phenomenon.

With regard to the solution of the two-stage optimization, a direct execution of this optimization can be inefficient as we show, so we demonstrate a novel way to add the second stage of optimization for a very small additional computational cost. Subject to mild conditions, it can be folded into the same single run of a shortest path algorithm such as Dijkstra's algorithm that performs the first-stage QoS routing optimization.

Section II presents preliminary notation and definitions and introduces the problems of admission control and routing of QoS connections. Section III develops a constraint and cost function to account for BE traffic performance when routing LSPs for QoS connections in a network domain that also serves BE traffic through connectionless routing. In Section IV, we present an efficient computation of our two-stage QoS routing strategy and present a numerical example illustrating the behavior of the QoS call blocking probability. Section V provides a theoretical explanation of the phase transition exhibited by the QoS call blocking probability.

II. Preliminaries

A network domain is represented by the directed graph $G = (V, E)$, where V is the set of vertices or nodes (symbolizing routers or switches) and E is the set of directed links. Each directed link e has bandwidth 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 QoS connection supports a single micro-flow or aggregate of micro-flows that share the same forwarding equivalence class (FEC) [1]. An existing connection c is described by its ingress node s_c, egress node d_c, routing path p_c, average flow rate b_c, and its effective bandwidth $\alpha_{c,e}$ at each link e in the path. Path p_c is defined as an ordered subset $\{e_1, e_2, \ldots, e_{h(p_c)}\}$ of E, where $h(p_c)$ is the number of hops along p_c and the order indicates the sequence of links along p_c from ingress node s_c to egress node d_c. Although there are different definitions of effective bandwidth (see [12]–[22] and references therein), in general, it is a value intermediate between a connection's average and peak flow rates. It indicates the amount of bandwidth that must be reserved for a connection in order to satisfy its QoS requirements. A connection's effective bandwidth may depend upon link operating points as well as the connection's traffic characteristics [15], [22], so our notation $\alpha_{c,e}$ has subscript e to denote the effective bandwidth at each link $e \in p_c$. We let χ represent a connection request. Denoting χ 's descriptors using χ in place of c, we assume that in response to connection request χ, admission control and routing is performed for χ as based on s_χ, d_χ, b_χ, and $\{\alpha_{\chi,e} \mid e \in E\}$.

Letting

$$A_e^{QoS} \equiv \sum_{\{c | e \in p_c\}} \alpha_{c,e} \tag{1}$$

which represents the amount of bandwidth that is reserved for the existing QoS connections at directed link $e \in E$ at the time of request χ 's arrival, we define

$$R_e^{\text{eff}} \equiv C_e - A_e^{QoS} \tag{2}$$

which we refer to as the residual effective bandwidth of link e. One of the necessary conditions for link e to be feasible for assignment to an LSP to support connection request χ is that the residual effective bandwidth constraint

$$\alpha_{\chi,e} \leq R_e^{\text{eff}} \tag{3}$$

must be satisfied at link e. A candidate path p is feasible for supporting connection request χ if and only if it is composed entirely of feasible links.

We define the function $J_\chi^{(1)}(e)$ such that, for feasible link e, $J_\chi^{(1)}(e)$ is a measure of the cost to the network of reserving a sufficient amount of bandwidth along link e for requested connection χ in order to guarantee the QoS of its traffic. We refer to $J_\chi^{(1)}(e)$ as the base cost of link e and define

$$J_\chi^{(1)}(p) \equiv \sum_{e \in p} J_\chi^{(1)}(e) \tag{4}$$

as the corresponding base cost of selecting path p to support χ. We can leave the link base costs as general functions to some extent so that the routing and call admission control based on this cost function can be tailored to different operational environments. However, we incorporate the feasibility constraints into the cost functions by penalizing $J_\chi^{(1)}(e)$ to be infinite for any link e that is not feasible. In this paper, we use the net-effective-bandwidth reservation cost

$$J_\chi^{(1)}(e) \equiv \begin{cases} \alpha_{\chi,e}, & \text{if link } e \text{ is feasible} \\ \infty, & \text{otherwise} \end{cases} \tag{5}$$

where link feasibility is defined as satisfaction of constraint (3). With this cost definition, the minimization of $J_\chi^{(1)}(p) \equiv \Sigma_{e \in p} J_\chi^{(1)}(e)$ would[1] be to find a route that reserves the least amount of bandwidth for QoS connection χ without any regard to the performance of the best-effort (BE) traffic sharing the network resources. The two-stage optimization problem to be formulated in Section III for the routing and admission control of the QoS connections will have another constraint, which is designed to sustain the performance of the BE traffic.

We limit the scope here to single-path routing for each QoS connection. Although multipath routing may provide the potential to improve network load-balancing [11], [23], it increases the complexity of the signaling process for LSP setup and increases the likelihood of out-of-order packet arrivals.

III. BE Traffic Considerations

This section develops an LSP routing constraint and cost function that account for the effects of QoS routing on BE traffic.

[1]Other useful definitions of $J_\chi^{(1)}(e)$ for feasible link e include the minimum-hop path cost, 1, and the reciprocal of residual effective bandwidth [5], $1/R_e^{\text{eff}}$.

A. BE Traffic Available Bandwidth

Letting

$$B_e^{QoS} \equiv \sum_{\{c | e \in p_c\}} b_c \tag{6}$$

which represents the average bandwidth usage of QoS connections that are in progress at directed link $e \in E$, we define

$$C_e^{\text{BE}} \equiv C_e - B_e^{\text{QoS}} \tag{7}$$

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 connection's average bandwidth consumption does not exceed its effective bandwidth, i.e.,

$$\varepsilon_{c,e} \equiv \alpha_{c,e} - b_c \geq 0, \quad \forall\, e \in p_c, \ \forall\, c \tag{8}$$

where $\varepsilon_{c,e}$ is what we refer to as the excess effective bandwidth of connection c at link e. Then, by (1), (2), and (6)–(8), it follows that

$$C_e^{\text{BE}} \geq R_e^{\text{eff}}, \quad \forall\, e. \tag{9}$$

When strict inequality holds in (9), C_e^{BE}, the average bandwidth available for BE traffic, includes some bandwidth that is allocated to but, on average, is not consumed by the QoS connections.

In this paper, we assume that the routers or switches in the network domain implement a queueing discipline that enables BE traffic to take advantage of this excess effective bandwidth without interfering with the QoS connections, that is, at any moment at which QoS connections are not fully utilizing all the bandwidth reserved for them, the BE traffic can immediately grab this otherwise unused bandwidth. This capability can be provided by such queueing disciplines as weighted fair queueing (WFQ) [9] and deficit round robin (DRR) [10], which allow different traffic classes to be queued separately.

B. BE Traffic Performance Metric

Our approach is to set a minimally protected BE bandwidth, which we denote as F_e, through link e and to define a metric that is indicative of BE traffic performance. We use this metric to define a BE traffic performance 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, we then select from them one that least penalizes the BE traffic performance. We use a very general BE traffic performance metric of the form

$$D_{\text{BE}} \equiv \sum_{e \in E} g_e \left(C_e^{\text{BE}} \right) \tag{10}$$

where g_e, in (10) for each link, e, is nonnegative, monotonically nonincreasing, and right-continuous on $[F_e, \infty)$. By the nature of the connection admission control of the QoS traffic, which will be presented later, C_e^{BE}, the average bandwidth available for BE traffic, is maintained such that

$$C_e^{\text{BE}} \geq F_e. \tag{11}$$

We note that parameter F_e and function $g_e\left(C_e^{\mathrm{BE}}\right)$ are to be selected or designed by the network service provider. Appendix B briefly discusses the network service provider's (operator's) choices of parameter F_e and performance function $g_e\left(C_e^{\mathrm{BE}}\right)$. The essential characteristic of (10) is that g_e is nonincreasing with its variable C_e^{BE} and non-decreasing with parameter F_e. An example of the BE performance metric would be an M/M/1 approximating function

$$g_e\left(C_e^{\mathrm{BE}}\right)=\frac{F_e}{\gamma\left(C_e^{\mathrm{BE}}-F_e\right)} \tag{12}$$

where γ is another parameter to be set by the service provider in accordance with its network planning. (Setting γ and F_e to the average BE traffic entering the network and the average BE traffic through link e, respectively, if these averages are available to the network planner, will make (10), with (12) plugged in, the M/M/1 approximation of a BE packet's delay in the network as in [24, Ch.5].)

C. BE Protecting Residual Bandwidth Constraint

To protect BE performance, from BE performance metric (10), we can impose the following constraint:

$$D_{\mathrm{BE}} \leq \overline{D} \tag{13}$$

for some positive value, $\overline{D}$. However, this alone can result in unfair treatment of BE flows going through different paths. Therefore, we require that C_e^{BE}, the average bandwidth available for BE traffic, be such that

$$g_e\left(C_e^{\mathrm{BE}}\right) \leq \overline{g}_e, \quad \forall\, e \in E \tag{14}$$

where the set of positive values $\{\overline{g}_e\}$ is selected such that

$$\sum_{e\in E} \overline{g}_e \leq \overline{D}. \tag{15}$$

We next define

$$\underline{C}_e^{\mathrm{BE}} \equiv \inf\{x \mid x \geq F_e, g_e(x) \leq \overline{g}_e\}, \quad \forall\, e \in E \tag{16}$$

and

$$\Delta_e \equiv \underline{C}_e^{\mathrm{BE}} - F_e \geq_{(\text{by (16)})} 0, \quad \forall\, e \in E \tag{17}$$

and note that, by the right-continuity of g_e, we have

$$g_e\left(\underline{C}_e^{\mathrm{BE}}\right) \leq \overline{g}_e, \quad \forall\, e \in E. \tag{18}$$

Therefore, since g_e is nonincreasing on $[F_e, \infty)$, we can express constraint (14), the link-wise BE performance bound, as

$$C_e^{\mathrm{BE}} \geq \underline{C}_e^{\mathrm{BE}} =_{(\text{by (17)})} F_e + \Delta_e, \quad \forall\, e \in E. \tag{19}$$

Now, suppose that, to support connection request χ, we are considering a candidate path that includes directed link e whose current available bandwidth for BE traffic is C_e^{BE}. If link e does end up supporting χ, then, by (6) and (7), its available bandwidth for BE traffic will become $C_e^{\mathrm{BE}} - b_\chi$. Then, constraint (14) will still be satisfied only if (19) holds with $C_e^{\mathrm{BE}} - b_\chi$ substituted for C_e^{BE}. This condition can be expressed as

$$b_\chi \leq R_e^{\mathrm{ave}}, \tag{20}$$

where

$$R_e^{\mathrm{ave}} \equiv C_e^{\mathrm{BE}} - \underline{C}_e^{\mathrm{BE}} = C_e^{\mathrm{BE}} - F_e - \Delta_e. \tag{21}$$

Like R_e^{eff}, R_e^{ave} represents a residual bandwidth of directed link e. Whereas R_e^{eff} limits [via (3)] the effective bandwidth of a new connection that link e can support, R_e^{ave} limits [via (20)] the average transmission rate of a new connection that link e can support. Therefore, for directed link e to be feasible for supporting connection request χ, both residual bandwidth constraints (3) and (20) must be satisfied at link e.

D. BE Traffic Cost of a Candidate Path

Suppose that we are considering a candidate feasible path p to support connection request χ. If path p is selected, then, for each $e \in p$, link e's available bandwidth for BE traffic will change from C_e^{BE} to the new value, $C_e^{\mathrm{BE}} - b_\chi$. Consequently, by (10), the increase in BE performance metric D_{BE} due to using the feasible link, $e \in p$, will be

$$g_e\left(C_e^{\mathrm{BE}} - b_\chi\right) - g_e\left(C_e^{\mathrm{BE}}\right)$$

and the corresponding increase in D_{BE} due to using feasible path p will be $\sum_{e\in p}\left[g_e\left(C_e^{\mathrm{BE}} - b_\chi\right) - g_e\left(C_e^{\mathrm{BE}}\right)\right]$. We use these increases in BE traffic metric D_{BE} as the basis for defining the link and path costs for the second stage of the routing optimization. These represent the costs to BE traffic of using a given link or path in a new LSP in response to a QoS connection request. What we refer to as the BE cost of using link e to support connection request χ is then given by

$$J_\chi^{(2)}(e) \equiv \begin{cases} g_e\left(C_e^{\mathrm{BE}} - b_\chi\right) - g_e\left(C_e^{\mathrm{BE}}\right), \\ \quad \text{if } e \text{ satisfies (3) and (20)} \\ \infty, \\ \quad \text{otherwise.} \end{cases} \tag{22}$$

The corresponding BE path cost of candidate path p, under consideration for serving χ, is

$$J_\chi^{(2)}(p) \equiv \sum_{e\in p} J_\chi^{(2)}(e). \tag{23}$$

IV. BE-Friendly LSP Routing

We next describe our routing optimization strategy and an efficient method for executing it. Then we present our BE-friendly routing for QoS traffic with a problem formulation, algorithm, and numerical example.

A. Optimization Strategy

Arriving QoS connection request χ is admitted if there exists a feasible path from s_χ to d_χ, i.e., one composed entirely of links that satisfy residual bandwidth constraints (3) and (20).

For the general case in which there are multiple feasible candidate paths, we apply a two-stage optimization strategy (which is conceptually similar to one in [11]) for path selection. In a direct execution of this strategy, we identify in the first stage all feasible paths that achieve a finite minimum value of the base path cost, $J_\chi^{(1)}(p)$, of (4). Then, we apply the second stage, which is to select from all such paths one that achieves a finite minimum value of $J_\chi^{(2)}(p)$ of (23). In the case that such a path exists, we use it to support connection request χ.

Although this two-stage optimization strategy can be implemented directly as described above, such an implementation can be inefficient. The reason is that, in the worst case, the number of feasible, minimum-base-cost paths from s_χ to d_χ can be exponential with $|V|$, the number of nodes; Appendix C provides a simple constructive proof of this fact. Thus, an algorithm that enumerates all feasible minimum-base-cost paths has a worst case running time that is exponential in $|V|$. However, we next show that an efficient computation of the two-stage optimization is possible if certain conditions are met.

Given QoS connection request χ, let P_{feas} denote the set of feasible candidate paths from s_χ to d_χ. Then, P_{feas} is constrained such that each $p \in P_{\text{feas}}$ is a path composed entirely of links that satisfy constraints (3) and (20). (Further constraints can also be imposed on P_{feas} by a network administration—e.g., by requiring that the hop count of each path in P_{feas} not exceed a maximum value, H.) Suppose that there exists a positive value q, with the following property: for any feasible candidate path, if its base cost is not the minimum value, then it is larger than the minimum by at least q —i.e.,

$$\begin{aligned} \forall\, p \in P_{\text{feas}},\ \text{if } J_\chi^{(1)}(p) &> \min_{p' \in P_{\text{feas}}} \left\{ J_\chi^{(1)}(p') \right\} \\ \text{then } J_\chi^{(1)}(p) &\geq \min_{p' \in P_{\text{feas}}} \left\{ J_\chi^{(1)}(p') \right\} + q. \end{aligned} \tag{24}$$

A simple sufficient condition for satisfying (24) is to limit the range of $J_\chi^{(1)}(e)$ to a countable set of values $\{J_1, J_2, \ldots\}$, such that

$$\inf_{i,j \in \{1,2,\ldots\}, i \neq j} \{|J_i - J_j|\} \geq q. \tag{25}$$

For example, (25) is satisfied if $J_\chi^{(1)}(e)$ is quantized with quantum q, i.e.,

$$J_\chi^{(1)}(e) \in \{nq \mid n \in Z^+\} \cup \{\infty\}, \quad \forall\, e \in E. \tag{26}$$

Then, the base path cost $J_\chi^{(1)}(p)$ is similarly quantized. A simple example satisfying (26) is the minimum-hop link cost $J_\chi^{(1)}(e) = 1$ for each feasible link e. More generally, if a sufficiently small value of q is selected, then the base cost may be approximated by a quantized version of $J_\chi^{(1)}(e)$ without having any appreciable effect on the resulting routing process.

Whenever (24) holds, we can use the following method to efficiently perform our two-stage optimization. Consider a path cost that is a weighted sum of base and BE costs, i.e., define

$$J_\chi^{\text{ws}}(p) \equiv J_\chi^{(1)}(p) + wJ_\chi^{(2)}(p) \tag{27}$$

where w is a positive weighting coefficient for the BE cost (with unity weighting of the base cost). A path that minimizes $J_\chi^{\text{ws}}(p)$ does not necessarily minimize the base cost, $J_\chi^{(1)}(p)$. However, as shown in the following proposition, if we can select a value for w that is sufficiently small so that

$$wJ_\chi^{(2)}(p) < q, \quad \forall\, p \in P_{\text{feas}} \tag{28}$$

then we can insure that any path that achieves a finite minimum value of $J_\chi^{\text{ws}}(p)$ also minimizes $J_\chi^{(1)}(p)$. Furthermore, such a path has the minimum BE cost among all feasible minimum-base-cost paths. These results enable us to use a single run of a shortest path algorithm to efficiently implement both the first-stage and second-stage optimizations, as we will describe shortly.

Proposition 1: Given connection request χ, suppose that q and w are such that conditions (24) and (28) hold. If feasible path p_w is such that

$$J_\chi^{\text{ws}}(p_w) = \min_{p \in P_{\text{feas}}} \left\{ J_\chi^{\text{ws}}(p) \right\} \tag{29}$$

then it follows that

$$J_\chi^{(1)}(p_w) = \min_{p \in P_{\text{feas}}} \left\{ J_\chi^{(1)}(p) \right\} \tag{30}$$

and that

$$\begin{aligned} &J_\chi^{(2)}(p_w) \leq J_\chi^{(2)}(p) \text{ for every } p \in P_{\text{feas}} \text{ such that} \\ &J_\chi^{(1)}(p) = \min_{p' \in P_{\text{feas}}} \left\{ J_\chi^{(1)}(p') \right\}. \end{aligned} \tag{31}$$

Proof: Let p_1 be an arbitrary path in P_{feas} that achieves the following minimum:

$$J_\chi^{(1)}(p_1) = \min_{p \in P_{\text{feas}}} \left\{ J_\chi^{(1)}(p) \right\}. \tag{32}$$

For a proof of (30) by contradiction, suppose that

$$J_\chi^{(1)}(p_w) > \min_{p \in P_{\text{feas}}} \left\{ J_\chi^{(1)}(p) \right\}. \tag{33}$$

Then

$$\begin{aligned} J_\chi^{\text{ws}}(p_w) &\geq_{(\text{by }(27))} J_\chi^{(1)}(p_w) \\ &\geq_{(\text{by }(24)\text{ and }(33))} \min_{p \in P_{\text{feas}}} \left\{ J_\chi^{(1)}(p) \right\} + q \\ &=_{(\text{by }(32))} J_\chi^{(1)}(p_1) + q \\ &>_{(\text{by }(28))} J_\chi^{(1)}(p_1) + wJ_\chi^{(2)}(p_1) \\ &\geq_{(\text{by }(27))} \min_{p \in P_{\text{feas}}} \left\{ J_\chi^{\text{ws}}(p) \right\}, \end{aligned}$$

which contradicts (29). Therefore, (30) is established. For (31), again let $p_1 \in P_{\text{feas}}$ be such that (32) holds. Then

$$\begin{aligned} J_\chi^{(2)}(p_1) &=_{(\text{by }(27)\text{ and }(32))} \frac{1}{w} \\ &\quad \times \left[J_\chi^{\text{ws}}(p_1) - \min_{p \in P_{\text{feas}}} \left\{ J_\chi^{(1)}(p) \right\} \right] \\ &\geq_{(\text{by }(29))} \frac{1}{w} \left[J_\chi^{\text{ws}}(p_w) - \min_{p \in P_{\text{feas}}} \left\{ J_\chi^{(1)}(p) \right\} \right] \end{aligned}$$

$$=_{(\text{by }(30))} \frac{1}{w}\left[J_\chi^{\text{ws}}(p_w) - J_\chi^{(1)}(p_w)\right]$$
$$=_{(\text{by }(27))} J_\chi^{(2)}(p_w).$$

∎

With regard to the availability of the sufficiently small value of w, residual bandwidth constraint (20) guarantees the existence of w that satisfies condition (28). We now verify this claim. Constraint (20) with (21) implies that

$$C_e^{\text{BE}} - b_\chi \geq F_e + \Delta_e = \underline{C}_e^{\text{BE}}, \quad \forall e \in p \tag{34}$$

which, in turn, from (18) and monotonicity of g_e implies that

$$g_e\left(C_e^{\text{BE}} - b_\chi\right) \leq \overline{g}_e, \quad e \in p. \tag{35}$$

Therefore, $J_\chi^2(p)$ is bounded as

$$\begin{aligned} J_\chi^{(2)}(p) &=_{(\text{by }(22)\text{ and }(23))} \sum_{e\in p}\left[g_e\left(C_e^{\text{BE}} - b_\chi\right) - g_e\left(C_e^{\text{BE}}\right)\right] \\ &\leq \sum_{e\in p} g_e\left(C_e^{\text{BE}} - b_\chi\right) \\ &\leq_{(\text{by }(35))} \sum_{e\in p}\overline{g}_e \\ &\leq \sum_{e\in E}\overline{g}_e. \end{aligned} \tag{36}$$

Then, by selecting

$$w = \frac{\delta q}{\sum_{e\in E}\overline{g}_e} \tag{37}$$

for some $\delta \in (0,1)$, we are assured of satisfying condition (28) because for $p \in P_{\text{feas}}$

$$wJ_\chi^{(2)}(p) \leq_{(\text{by }(36))} w\sum_{e\in E}\overline{g}_e <_{(\text{by }(37))} q. \tag{38}$$

B. LSP Routing Problem Formulation

Given QoS connection request χ, for each $e \in E$, let

$$y_e \equiv \begin{cases} 1, & \text{if an LSP to support } \chi \text{ includes link } e \\ 0, & \text{otherwise.} \end{cases} \tag{39}$$

Assuming that conditions (24) and (28) hold, the two-stage optimization for routing an LSP to support connection request χ can be described as the selection of $\{y_e\}$ that minimizes

$$J_\chi^{\text{ws}}(p) =_{(\text{by }(4),\,(23),\text{ and }(27))} \sum_{e\in E} y_e\left[J_\chi^{(1)}(e) + wJ_\chi^{(2)}(e)\right] \tag{40}$$

subject to

$$\sum_{j\in V} y_{(i,j)} - \sum_{k\in V} y_{(k,i)} = 0 \quad \forall\, i \neq s_\chi, d_\chi \tag{41}$$

$$\sum_{j\in V} y_{(s_\chi,j)} - \sum_{j\in V} y_{(j,s_\chi)} = 1 \tag{42}$$

$$\sum_{j\in V} y_{(d_\chi,j)} - \sum_{j\in V} y_{(j,d_\chi)} = -1, \tag{43}$$

$$y_e \in \{0,1\}, \quad \forall\, e \in E. \tag{44}$$

The flow conservation constraints (41) for all vertices in V other than s_χ and d_χ, (42) for s_χ, and (43) for d_χ force $\{y_e\}$ to describe a single path from s_χ to d_χ.

There may be other constraints placed on this path-selection problem. An example 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 (40)–(44). Another example is a maximum hop count H as may be associated with the FEC of the QoS connection request [11], [25], [26]. This constraint can be accommodated by adding the inequality

$$\sum_{e\in E} y_e \leq H \tag{45}$$

to problem formulation (40)–(44).

If there is no feasible $\{y_e\}$ for the problem formulation, then the connection request is denied. Otherwise, an optimal path p^* to support connection request χ is constructed as the ordered set

$$p^* \equiv \left\{(v_n, v_{n+1}) \mid n = 1, 2, \ldots, \sum_{e\in E} y_e^*\right\} \tag{46}$$

where $\{y_e^*\}$ is an optimal solution to (40)–(44), $v_1 = s_\chi$, and v_{n+1} is the unique vertex such that $y^*_{(v_n,v_{n+1})} = 1$.

C. LSP Routing Algorithm

Although problem formulation (40)–(44) is a 0–1 integer programming problem, it has a special structure that lends itself to solution by a shortest path algorithm such as the Dijkstra's or Bellman–Ford (BF) algorithm. By defining as our link metric

$$J_\chi^{\text{ws}}(e) \equiv J_\chi^{(1)}(e) + wJ_\chi^{(2)}(e) \tag{47}$$

we have by (4), (23), and (27) that a path from s_χ to d_χ with shortest possible length under metric $J_\chi^{\text{ws}}(e)$ also minimizes $J_\chi^{\text{ws}}(p)$ of (27). Thus, if conditions (24) and (28) hold, then, by Proposition 1, a shortest path algorithm run from source s_χ using link metric $J_\chi^{\text{ws}}(e)$ will seek a minimum-base-cost path to each other node that, among such paths, minimizes the BE cost. If such an optimal path exists from s_χ to d_χ with a finite base-cost defined in (22) and (23), then it is assigned to support connection request χ; otherwise, the connection request is denied.

Regarding constraints on path selection, if only links in $E' \subseteq E$ are admissible for routing, then the shortest path algorithm can be run on graph (V, E'). If there is a maximum-hop count constraint H, then the BF algorithm is convenient for path selection because at its h th iteration, it identifies optimal paths not exceeding h hops from the source to the other nodes [27]. Thus, H iterations of BF will find an optimal path from s_χ to d_χ that does not exceed H hops, if such a path exists.

A shortest path algorithm can execute our two-stage optimization strategy very efficiently. Dijkstra's algorithm can be coded to run in time $O(|E| + |V|\log|V|)$ and the BF algorithm runs in time $O(|V||E|)$ [28]. By comparison, a direct implementation of the two-stage optimization would have a worst case running time that was exponential in $|V|$.

D. Numerical Example

We present results from a simulation of BE-friendly QoS routing and admission control using a "sample US nationwide topology" [29] with 24 nodes and 86 directed links. For the cases simulated, QoS connection request arrivals are modeled as a Poisson process and connection holding times are exponentially distributed. We let ρ represent the QoS offered load in Erlangs (E), i.e., the product of the QoS connection request arrival rate and the mean connection holding time. Steady-state conditions are approximated by running each simulation long enough for 250,000 connection requests to arrive. The ingress–egress pair of each connection request is randomly chosen from all possible combinations with equal likelihood. For simplicity, we let the effective bandwidth of arriving connection request χ be link-invariant and denote it by α_χ. In our numerical study, the capacity of each link is set to 160 units, and the bandwidth request α_χ takes values of 0.1, 0.15, 0.6, 1, 2.5, 5, and 10 units. The ratio of these different bandwidth requests' arrival rates is set to be 50: 20:10:10:4:2:1, which is described in [29] as a "typical bandwidth distribution in a practical network." We perform simulations with a ratio of effective bandwidth to average bandwidth usage

$$r_\chi \equiv \alpha_\chi / b_\chi \tag{48}$$

being 1, with r_χ uniformly distributed over [1, 2], and uniformly distributed over [1.5, 2.5].[2]

We apply the BE link cost function of (22) with g_e as given by (12), which is based on an M/M/1 model of packet queueing. Referring to (12), for simplicity, we let the minimally protected BE bandwidth F_e have the same value F at every link and vary F in the range $[0, C]$. Regarding γ in (12), let h be the average number of hops that BE packets traverse from ingress to egress. Then, we set

$$\gamma = F\,|E|/h \tag{49}$$

which would be the average BE traffic entering the network if the average BE traffic through each link were equal to F. We let $h = 3$, which is reasonable in view of comments in [31] and [32] about the average hop count over Internet backbone links. For bounding BE traffic performance, we set

$$\overline{g}_e = \overline{D}/|E|, \quad \forall\, e \in E \tag{50}$$

which satisfies (15) with equality. In accordance with the M/M/1 model (12), $\overline{D}$ can be interpreted as an upper bound on the BE packet delay in the network, and we let $\overline{D} = 0.2$ s.[3] Parameter

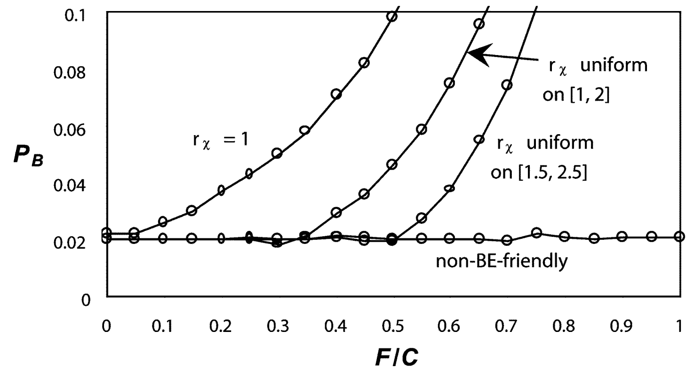

Fig. 2. QoS call blocking probability P_B versus F/C for $\rho = 7000E$.

Δ_e, which appears in definition (21) for R_e^{ave}, is calculated per (17) where $\underline{C}_e^{\text{BE}}$ is, by (12), (16), and (50), such that

$$\frac{F}{\gamma\left(\underline{C}_e^{\text{BE}} - F\right)} = \frac{\overline{D}}{|E|}$$

$$\underline{C}_e^{\text{BE}} = \frac{F(|E| + \gamma\overline{D})}{\gamma\overline{D}}. \tag{51}$$

Then, in accordance with (17) and (51), for our simulation we use

$$\Delta_e = \frac{F\,|E|}{\gamma\overline{D}} \equiv \Delta, \quad \forall\, e \in E. \tag{52}$$

Fig. 2 shows the QoS call blocking probability P_B plotted versus F/C, where C is the bandwidth capacity for each of the links. A QoS offered load, ρ, of 7000 E (Erlang) was used for Fig. 2. The plot labeled "$r_\chi = 1$" is for the case of zero excess effective bandwidth, i.e., $\varepsilon_\chi \equiv \alpha_\chi - b_\chi = 0$ (by (48) for $r_\chi = 1$) for each arriving QoS connection request, χ. In this case, P_B increases quickly with F/C. In contrast, the plots labeled "r_χ uniform on [1, 2]" and "r_χ uniform on [1.5, 2.5]" exhibit distinct phase transitions. P_B is relatively constant with F/C up to a certain point, beyond which it increases. Below this threshold, P_B is approximately the same as for the "non-BE-friendly" plot, which is for the case of no constraint (20) and no BE-friendly second-stage optimization.

Fig. 3 repeats the case of r_χ uniform on [1.5, 2.5] for the indicated values of QoS offered load ρ. In Fig. 3, P_B is plotted on a logarithmic scale in order to illustrate that P_B's phase transition occurs at approximately the same threshold value of F/C regardless of ρ.

Simulation results were also obtained for cases in which ingress–egress pairs for arriving LSP requests were nonuniformly distributed. We randomly selected from node set V, a subset containing eight nodes. For an LSP request, each node in this subset was selected as the ingress node with the same probability, which was ten times the probability common to all nodes outside the subset. After ingress node selection, the egress node was similarly selected from the remaining nodes, i.e., nodes in the favored subset were selected with 10 times the probability of nodes outside this subset. Fig. 4 shows plots for eight random selections of this subset for the case of r_χ uniform on [1.5, 2.5] and $\rho = 10000E$. The threshold value of

[2][30] presents Internet WAN traffic traces for several links that indicate ratios of effective bandwidth to average bandwidth between 1 and 2.6.

[3]In comparison, typical Internet backbone routers provide enough buffering for queueing delays to grow to as much as 0.5 s at a link (e.g.,[33]).

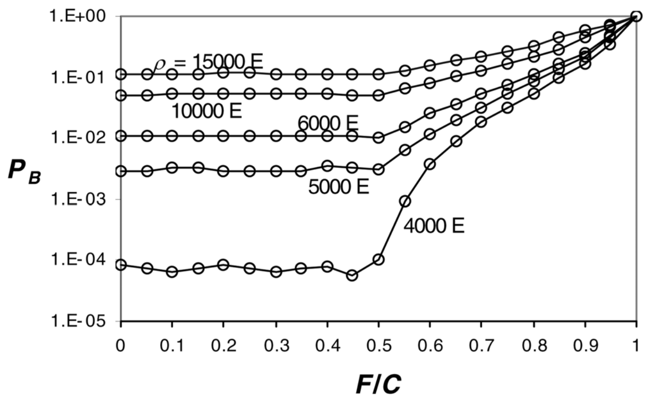

Fig. 3. P_B versus F/C with r_χ uniform on [1.5, 2.5].

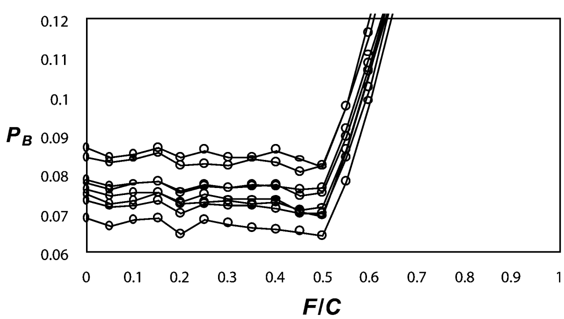

Fig. 4. P_B versus F/C for eight simulations for $\rho = 10000E$ with nonuniformly distributed QoS ingress–egress pairs.

F/C at which P_B's phase transition occurs is again relatively invariant for the various cases simulated.

Finally, we mention that we also performed the simulations summarized by Figs. 2–4 for a 15-node topology [34] and observed similar features in those results.

V. Theoretical Basis for QoS Call Blocking Phase Transition

Here, we look for a theoretical basis for the phase transition phenomena we observe in our simulation results. For insight into the phase transition exhibited by P_B, consider the two residual bandwidth constraints (3) and (20) that are imposed when attempting to route QoS connection request χ. Generalizing (48), we define

$$r_{\chi,e} \equiv \alpha_{\chi,e}/b_\chi \tag{53}$$

and express (3) and (20) collectively as

$$\alpha_{\chi,e} \le \min\left\{R_e^{\mathrm{eff}}, r_{\chi,e}\, R_e^{\mathrm{ave}}\right\}, \quad \forall\, e \in E. \tag{54}$$

Inequality (20), the constraint to protect BE traffic, contributes the term $r_{\chi,e}R_e^{\mathrm{ave}}$ in (54). Therefore, (20) does not actually constrain QoS routing unless

$$r_{\chi,e}R_e^{\mathrm{ave}} < R_e^{\mathrm{eff}}. \tag{55}$$

To simplify analysis, let us remove any randomness in $r_{\chi,e}$, i.e., suppose that $r_{\chi,e} = \overline{r}$, a constant, for every arriving connection request and link e and that $r_{c,e} \equiv \alpha_{c,e}/b_c = \overline{r}$ for any connection c in progress and link e. Then, with (1) and (6), it follows that

$$A_e^{QoS} = \overline{r}B_e^{\mathrm{QoS}}. \tag{56}$$

Then, (2), (7), (21), and (56) imply that (55) is equivalent to

$$\begin{aligned} \overline{r}(C_e - B_e^{\mathrm{QoS}} - F_e - \Delta_e) &< C_e - \overline{r}B_e^{\mathrm{QoS}} \\ F_e &> \overline{F}_e \equiv C_e(\overline{r}-1)/\overline{r} - \Delta_e. \end{aligned} \tag{57}$$

Therefore, if the minimally protected BE bandwidth F_e does not exceed the threshold value $\overline{F}_e$ of (57), then the BE-friendly constraint (20) for link e does not actually constrain QoS routing because the right-hand side of (54) is R_e^{eff} for this case.

In the simulations of Section IV-D, we assume that $C_e = C$, $F_e = F$, and $\Delta_e = \Delta$ for every link $e \in E$. With these simplifying assumptions, if

$$F \le \overline{F} \equiv C(\overline{r}-1)/\overline{r} - \Delta, \tag{58}$$

then (20) does not constrain QoS routing at any link and thus does not increase P_B. For this case, P_B exhibits a phase transition with increasing F if $\overline{F}$ is positive, i.e., if

$$\Delta < C(\overline{r}-1)/\overline{r}.$$

Consider the example of Section IV-D and suppose that C, which was given as 160 units, is 160 Mb/s. We were using an M/M/1 model of packet queueing (12), and the packet delay bound $\overline{D} = 0.2$ s. Accordingly, by (49) and (52) with $h = 3$, we have $\Delta = 15$ packets/s. Assuming an average of 3200 bits per packet, we can express Δ as 0.048 Mb/s. If we set $\overline{r}$ to the mean of r_χ for the case of r_χ uniform on [1.5, 2.5], i.e., $\overline{r} = 2$, then by (58), $\overline{F}/C = 1/2 - \Delta/C \cong 0.5$. Referring again to Figs. 2–4, we note that, for each case of r_χ uniformly distributed on [1.5, 2.5], P_B exhibits a relatively crisp phase transition located approximately at $F/C = 0.5$, despite the randomness of r_χ.

VI. Related Work

The works in [4], [5], and [35]–[37] have addressed integration of QoS and best-effort routing. Each presents a link cost for shortest path routing of QoS connections that gives consideration to BE traffic in some manner. Notable in this work is the virtual residual bandwidth (VRB) concept of [5], which is to adjust each link's residual bandwidth to reflect BE traffic loads. These adjustments are based on max-min fair share rates (e.g., [24, Ch. 6]), which are also used to route the BE traffic. The link cost in this case is defined as the reciprocal of VRB. Reference [35] presents an alternative definition of VRB that does not require computation of max-min fair share rates. Also related is a virtual link cost described in [36]. [4] presents a probabilistic approach to link costs that considers the uncertainty in a node's knowledge of residual bandwidths due to changing network conditions and the limited update rate of a link-state protocol. Another routing algorithm that gives consideration to BE traffic, though in a different setting with connectionless, hop-by-hop QoS routing, is the enhanced bandwidth-inversion shortest path algorithm of [37].

Our methods are basically compatible with those described in [4], [5], [35], [36]. The link costs of these papers can be used with our BE-friendly QoS routing constraint in the first optimization stage of QoS routing. Our second-stage optimization, the tie-breaking rule, is also feasible with, e.g., quantized versions of these papers' link costs. Also, our methods do not preclude the particular BE routing techniques described in [4], [5], [36]. However, as in [35], we do not prescribe the BE routing. While [4], [5] include the notion of setting aside bandwidth for BE traffic, our QoS routing constraint provides bandwidth for BE traffic more efficiently by accounting for BE exploitation of excess effective bandwidth. This efficiency in our constraint may yield a range of bandwidth that can be provided for BE traffic without significantly affecting the QoS call blocking probability. Furthermore, as described in Section III, our QoS routing constraint provides a new feature in that it sustains a specific minimum level of service for the BE traffic at each link.

VII. Conclusion and Future Work

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 QoS path routing is restricted by BE-friendly constraints can be traded off, via a few parameters, with BE traffic performance. Our simulation results suggest that the degradation of QoS routing performance (e.g., increased call 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 the BE-protecting constraint, one that has the least impact on BE performance. We have shown that, subject to mild conditions (e.g., quantization of QoS link costs), the two-stage optimization strategy can be implemented very efficiently, i.e., with the complexity of Dijkstra's algorithm.

Noteworthy in our simulation results are the phase transitions exhibited by plots of the QoS call blocking probability versus the minimally protected BE bandwidth. Below a threshold value of the minimally protected BE bandwidth, the blocking probability is not significantly affected by imposing our BE-protecting constraint. We demonstrated that this threshold results from the way our BE-protecting constraint accounts for BE exploitation of excess effective bandwidth.

In this paper, we have considered among QoS performance metrics only the QoS call blocking probability. We took the approach of representing packet-level QoS metrics, such as the QoS packet overflow probability and QoS latency, by the effective bandwidth bestowed. The relation between the packet-level QoS and the given effective bandwidth can be further studied separately. Also left as future work is exploring and examining particular cost functions (e.g., g_e, $J_\chi^{(1)}$, and $J_\chi^{(2)}$ in this paper) to suit the network service provider's particular interests. Function g_e used in (22) is a generalization of the average packet delay and can accommodate a wide class of average delay expressions that are constructed from empirical data (e.g., expression (7) of [38] for self-similar traffic). In Appendix D, we discuss using utility-based BE performance measure for $J_\chi^{(2)}$.

We are currently scrutinizing the idea of applying the BE-protecting QoS routing of the present paper to reliable routing [34], [39] with path protection.

Finally, we have limited our scope of admission control to the greedy policy, i.e., a policy that admits a QoS connection request if there exists a feasible path to support it. A nongreedy policy, considering QoS call requests to arrive in the future [40], may sometimes deny a connection request even though a feasible path exists, or forego the use of a particular feasible link when routing a connection (e.g., [8]). Theoretically, the optimal policy that allows for nongreedy ones should yield a lower call blocking probability for the QoS calls. Our future work will further address the challenging problem of optimizing combined admission control and routing problem formulations that allow nongreedy solutions.

Appendix A
Glossary of Notation

A_e^{QoS}	Bandwidth reserved for QoS connections at link e, (1).
$\alpha_{c,e}$, $\alpha_{\chi,e}$	Effective bandwidths of QoS connection c and request χ, respectively, at link e.
B_e^{QoS}	Average bandwidth usage of QoS connections at link e, (6).
b_c, b_χ	Average bandwidth usage of QoS connection c and request χ, respectively.
C_e	Bandwidth capacity of link e.
C_e^{BE}	Available bandwidth at link e for BE traffic, (7).
$\underline{C}_e^{\mathrm{BE}}$	A lower limit placed on C_e^{BE}, (16) and (19).
c, χ	Indices for QoS connections and requests, respectively.
D_{BE}	BE traffic performance metric, (10).
$\overline{D}$	Upper bound placed on D_{BE}.
d_c, d_χ	Egress nodes of QoS connection c and request χ, respectively.
Δ_e	Amount by which $\underline{C}^{\mathrm{BE}}$ exceeds F_e, (17).
δ	Positive fraction used in defining weighting coefficient w, (37).
E	Set of directed links in network, indexed by e, f.
$\varepsilon_{c,e}$	Excess effective bandwidth of QoS connection c at link e, (8).
F_e	Minimally protected BE bandwidth at link e.
$\overline{F}_e$	Threshold value of F_e at which a phase transition of P_B occurs, (57).
(i,j)	Link in E directed from node i to node j.
G	Graph (E,V).

$g_e(C_e^{\mathrm{BE}})$	Nonincreasing function of C_e^{BE} used for bounding BE performance at link e, e.g., (12).
$\overline{g}_e$	Upper bound placed on $g_e(C_e^{\mathrm{BE}})$, (14).
γ	Average total rate of BE traffic entering the network, (49).
H	Maximum hop count.
h	Average number of hops that BE packets traverse from ingress to egress.
$h(p_c)$	Number of hops of path p_c.
$J_\chi^{(1)}(e)$, $J_\chi^{(1)}(p)$	Base costs of using link e and path p, respectively, to support QoS connection request χ, (5) and (4), respectively.
$J_\chi^{(2)}(e)$, $J_\chi^{(2)}(p)$	BE costs of using link e and path p, respectively, to support QoS connection request χ, (22) and (23), respectively.
$J_\chi^{\mathrm{ws}}(p)$	Weighted sum of the base and BE costs of path p, (27).
P_B	QoS call blocking probability.
P_{feas}	Set of feasible candidate paths for a QoS connection request.
p, p^*	Candidate path and optimal path, respectively, for supporting a QoS connection request, (46) for p^*.
p_c	Path for connection c.
q	Quantum of base cost, (26).
R_e^{ave}	Residual bandwidth of link e in terms of QoS connection average transmission rates, (21).
R_e^{eff}	Residual effective bandwidth of link e, (2).
$r_{\chi,e}$	Connection request χ's ratio of effective bandwidth to average bandwidth usage at link e, (53).
$\overline{r}$	Constant value of $r_{\chi,e}$, Section V.
ρ	QoS offered load in Erlangs.
s_c, s_χ	Ingress nodes of QoS connection c and request χ, respectively.
V	Set of network nodes (indices include i, j, k).
w	Weighting factor used in defining weighted sum cost, (27), (37).
y_e, y_e^*	Indicators of whether link e is used in a path, and in an optimal path, respectively, to support a new QoS connection, (39).
Z^+	Set of positive integers (indexed by n).

APPENDIX B
CHOICE OF F_e AND FUNCTION $g_e\left(C_e^{\mathrm{BE}}\right)$

With the scheme proposed in this paper, link e maintains at least bandwidth F_e for the BE traffic all of the time. This parameter is to be set by the network service provider. An exemplary value would be the long-term average load of BE traffic that has to go through link e. In order to clarify the meaning of this "long-term average load," we consider the following example. Let us view the message generation by the application layer as a random (discrete-event) arrival process. Denote by λ_{sd} the rate at which application programs at node s generate messages destined to node d. Denote by random variable X_{sd} the size (in bits) of a message generated at node s and destined for node d. Then the long-term average load to transport data from source s to destination d is $\lambda_{sd}E(X_{sd})$ (bps). We note that this quantity does not depend on QoS traffic. Suppose that the BE routing protocol determines one path, p_{sd}, for each $s-d$ pair. (This is a realistic assumption because service providers often use the number of hops, or sum of inverses of link bandwidths, as the distance measure for their shortest-distance routing. In fact, adaptive routing is likely to cause oscillation.) Then, the average load of link e is

$$\sum_{s,d}\sum_{\{p_{sd}\mid e\in p_{sd}\}}\lambda_{sd}E(X_{sd}). \tag{59}$$

If value (59) may be assessed by the service provider as a part of its capacity planning, it is a good exemplary value for using as parameter F_e. Then (10), with $g_e\left(C_e^{\mathrm{BE}}\right)$ given by (12), has the form of an expression for the average time a BE packet spends traversing the network domain from ingress to egress as based on using an M/M/1 model of queueing at each link [24, Ch. 5]. Analogous delay metrics can be readily constructed based on the Pollaczek-Khinchin (P-K) formula for M/G/1 queues, or a G/G/1 waiting time bound (e.g., [24]).

In any choice of F_e and $g_e\left(C_e^{\mathrm{BE}}\right)$, at least bandwidth F_e is protected for BE traffic at link e. Our study indicates that there may be little adverse impact on QoS call blocking if F_e can be selected to not exceed a threshold value (e.g., (58)). We envision that the network service providers can adjust their choice of F_e and $g_e\left(C_e^{\mathrm{BE}}\right)$ on the basis of their policy and experiences.

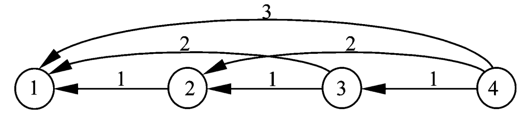

Fig. 5. Example network for $N=4$.

APPENDIX C
GRAPHS WITH AN EXPONENTIAL NUMBER OF SHORTEST PATHS

Consider a directed graph Γ_N with N nodes designated by integers $1,2,\ldots,N$, to be constructed in the following way. For each node j such that $j\in\{2,3,\ldots,N\}$, let there be one directed edge from j to i for each $i\in\{1,2,\ldots,j-1\}$ and let its length be $j-i$. Fig. 5 illustrates the graph for $N=4$; the lengths of each edge are as indicated. In this graph, there is a path from node n to node m if and only if $m<n$. For the case $m<n$, every path from node n to node m is a shortest path because all paths from node n to node m have the same length. Specifically, for $m<n$, all paths from node n to node m are of length $n-m$. Let $K(n)$ denote the number of shortest paths

from node n to node 1. We now prove by induction that in graph Γ_N for each N

$$K(n) = 2^{n-2}, \text{ for each } n \in \{2, 3, \ldots, N\}. \tag{60}$$

We first note that (60) is true for $N = 2$, i.e., there is one path from node 2 to node 1. Suppose, as an induction hypothesis that (60) holds for an integer N. Then, we are to show that (60) also holds for $N + 1$. In graph Γ_{N+1} node n does not have a path to node $k > n$. Therefore, $K(n) = 2^{n-2}$, for $n = 2, 3, \ldots, N$, even in graph Γ_{N+1}. With regard to $K(N + 1)$, consider paths from node $N + 1$ to node 1 that take node $j \in \{1, 2, \ldots, N\}$ as the first hop. There is one such path for $j = 1$. For $j \in \{2, 3, \ldots, N\}$, a path with its first hop from node $N+1$ to node j then follows one of the paths from node j to node 1, of which there are 2^{j-2}. Therefore

$$K(N+1) = 1 + \sum_{j=2}^{N} 2^{j-2} = 2^{N+1-2}$$

proving that (60) holds with $N+1$ substituted for N. With (60) holding for $N = 2$, we have by induction that (60) holds for all $N \in \{2, 3, \ldots\}$.

Therefore, the number of shortest paths can grow exponentially with the number of nodes in the graph. (If we replace all the directed edges of Γ_N with bi-directional edges with the same length, $j - i$, for both directions between each pair of nodes, i and j, such that $i < j$, the number of shortest paths from node N to node 1 in this new graph is no less than $K(N)$.)

Appendix D
BE Utility Criteria

The performance criterion for the BE traffic specified in Section III is based on the performance function $g_e\left(C_e^{\mathrm{BE}}\right)$ in individual links, as shown in (22) and (23). This function is a generalization, based on the fundamentals of queueing theory, of the average packet delay in the link. The assumptions about this function are quite general, so the results of this paper will be valid for an extremely wide class of empirically derived performance functions. In this section, we consider another class of BE performance criteria, which is not based on average delays but rather on the utility functions of BE sessions. It is known [41] that the sessions' equilibrium flow rates resulting from TCP/AQM protocols constitute a solution to the utility maximization problem, which in the context of the present paper is expressed as

$$\begin{aligned} \max_{x \in X} \quad & \sum_s U(x_s) \\ \text{subject to} \quad & \sum_{s \in S(e)} x_s \le C_e^{\mathrm{BE}}, \quad \forall\, e \end{aligned} \tag{61}$$

for a nondecreasing concave utility function U, where x_s here denotes the flow rate of BE session s, $S(e)$ denotes the set of BE sessions that go through link e, and X is the set of vectors whose components, say x_s, are all nonnegative. Note from [41] that a utility function U is associated with a specific TCP/AQM protocol. Let us denote a solution to this maximization by x_s^* (optimal flow of session s for each s, given the link bandwidth capacity constraints). Then, the total utility for BE traffic at this equilibrium is $\sum_s U(x_s^*)$. Suppose path p is selected for an incoming QoS connection χ, which takes away bandwidth b_χ along the path. Let us denote by x_s^p the flow rate of the new equilibrium, which is the solution to the following maximization:

$$\begin{aligned} \max_{x \in X} \quad & \sum_s U(x_s) \\ \text{subject to} \quad & \sum_{s \in S(e)} x_s \le C_e^{\mathrm{BE}} - b_\chi, \quad \forall\, e \in p \\ & \sum_{s \in S(e)} x_s \le C_e^{\mathrm{BE}}, \quad \forall\, e \notin p. \end{aligned} \tag{62}$$

Then, for the selection of the path for an incoming connection χ, we can consider using $\sum_s U(x_s^*) - \sum_s U(x_s^p)$ as function $J_\chi^{(2)}(p)$ in place of (23). Obviously, $\sum_s U(x_s)$ is bounded above due to the limited bandwidth in each link. This implies that theoretically the optimization strategy in Section IV-A can be used. Namely, in place of (40), we can use

$$J_\chi^{\mathrm{ws}}(p) = \sum_{e \in E} y_e J_\chi^{(1)}(e) + w\left[\sum_s U(x_s^*) - \sum_s U(x_s^p)\right]. \tag{63}$$

(Note that variables y_e in constraint set (41)–(44) form a path p.) A practical difficulty of this optimization is that the network operator would have difficulty in keeping track of all BE sessions. We now consider an interesting heuristic method.

According to the duality theory, the overall utility at equilibrium prior to accepting incoming connection is identical to the result of the following optimization problem, which is the dual of (61):

$$\min_{\mu \ge 0} f(\mu) \tag{64}$$

where

$$f(\mu) \equiv \max_{x \in X}\left[\sum_s U(x_s) - \sum_e \mu_e \left(\sum_{s \in S(e)} x_s - C_e^{\mathrm{BE}}\right)\right]. \tag{65}$$

With a similar procedure, we can establish that the optimization dual of (62) is

$$\min_{\mu \ge 0}\left[f(\mu) - \sum_{e \in p} \mu_e b_\chi\right]. \tag{66}$$

Denote

$$\mu* \equiv \operatorname*{argmin}_{\mu \ge 0} f(\mu), \quad \mu^p \equiv \operatorname*{argmin}_{\mu \ge 0}\left[f(\mu) - \sum_{e \in p} \mu_e b_\chi\right].$$

Then, the decrease of BE utility at equilibrium as the result of giving path p to the incoming QoS connection χ is

$$f(\mu*) - \left[f(\mu^p) - \sum_{e \in p} \mu_e^p b_\chi\right] \tag{67}$$

which is bounded below by

$$f(\mu*) - \left[f(\mu*) - \sum_{e \in p} \mu_e^* b_\chi \right] = \sum_{e \in p} \mu_e^* b_\chi \tag{68}$$

and bounded above as

$$f(\mu*) - f(\mu^p) + \sum_{e \in p} \mu_e^p b_\chi \leq \sum_{e \in p} \mu_e^p b_\chi. \tag{69}$$

(Bounds (68) and (69) follow from the definitions of μ^p and $\mu*$.) According to [41], variable μ_e is the congestion measure of link e, and the dynamics of a TCP/AQM protocol can be viewed as primal-dual optimization iterations. The value of μ_e indicates the congestion level of link e and, in a TCP/AQM algorithm, is effectually communicated to the source of the flow that goes through link e. Therefore, the value of μ_e can be viewed as a part of link state information, which can be monitored by the network operator. According to [41], μ_e^* indicates the congestion level of link e at the equilibrium prior to adding a path for an incoming QoS connection χ. In accordance with (68), $\sum_{e \in p} \mu_e^* b_\chi$ is a lower bound of the utility decrease resulting from giving path p to the incoming connection χ. Therefore, for the BE utility criterion, $\sum_{e \in p} \mu_e b_\chi$ is a good heuristic candidate for an indication of the impact on BE traffic of giving path p to the incoming connection χ. Note that the network operator can compute $\sum_{e \in p} \mu_e b_\chi$ from link states $\{\mu_e\}$, without knowledge of individual sessions s. To implement this approach, we can use

$$J_\chi^{(2)}(e) \equiv \begin{cases} \mu_e b_\chi, & \text{if } e \text{ satisfies (3) and } b_\chi \leq C_e^{\mathrm{BE}} - F_e' \\ \infty, & \text{otherwise} \end{cases}$$

in place of (22). (F_e' is the bandwidth to be maintained for BE traffic in link e all the time and is to be determined by the network operator with queue stability in mind.) This choice of $J_\chi^{(2)}(e)$ by (23) yields $J_\chi^{(2)}(p) = \sum_{e \in p} \mu_e b_\chi$ for a feasible path p. Then, the boundedness of utility (and thus boundedness of utility decrease) makes the optimization strategy in Section IV-A valid for this choice of $J_\chi^{(2)}(e)$. Moreover, the structure of this cost function enables us to collapse the two-stage optimization into a single run of the shortest-path problem, as presented in Section IV-C.

References

[1] E. Rosen, A. Viswanathan, and R. Callon, Multiprotocol label switching architecture IETF, RFC 3031, Jan. 2001.

[2] D. Awduche *et al.*, RSVP-TE: Extensions to RSVP for LSP tunnels IETF, RFC 3209, Dec. 2001.

[3] B. Jamoussi *et al.*, Constraint-based LSP setup using LDP IETF, RFC 3212, Jan. 2002.

[4] K. Nahrstedt and S. Chen, "Coexistence of QoS and best effort flows-routing and scheduling analysis," in *Proc. 10th IEEE Tyrrhenian Int. Workshop Digit. Commun.: Multimedia Commun.*, Ischia, Italy, Sep. 1998.

[5] Q. Ma and P. Steenkiste, "Supporting dynamic inter-class resource sharing: A multi-class QoS routing algorithm," in *Proc. IEEE INFOCOM*, 1999, pp. 649–660.

[6] OSPF Design Guide [Online]. Available: http://www.cisco.com/warp/public/104/1.html

[7] A. Ephremides and S. Verdu, "Control and optimization methods in communication network problems," *IEEE Trans. Automat. Contr.*, vol. 34, no. 9, pp. 930–942, Sep. 1989.

[8] K. Kar, M. Kodialam, and T. V. Lakshman, "Minimum interference routing of bandwidth guaranteed tunnels with MPLS traffic engineering applications," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 12, pp. 2566–2579, Dec. 2000.

[9] A. Demers, S. Keshav, and S. Shenker, "Analysis and simulation of a fair-queueing algorithm," in *Proc. ACM SIGCOMM*, Sep. 1989, pp. 1–12.

[10] M. Shreedhar and G. Varghese, "Efficient fair queueing using deficit round robin," in *Proc. ACM SIGCOMM*, Oct. 1995, pp. 231–242.

[11] D. Mitra and K. G. Ramakrishnan, "A case study of multiservice, multipriority traffic engineering design for data networks," in *Proc. IEEE GLOBECOM*, Dec. 1999, pp. 1077–1083.

[12] G. d. Veciana, G. Kesidis, and J. Walrand, "Resource management in wide-area ATM networks using effective bandwidths," *IEEE J. Sel. Areas Commun.*, vol. 13, no. 6, pp. 1081–90, Jun. 1995.

[13] G. Kesidis, J. Walrand, and C. Chang, "Effective bandwidths for multi-class Markov fluids and other ATM sources," *IEEE/ACM Trans. Netw.*, vol. 1, no. 4, pp. 424–428, Aug. 1993.

[14] F. P. Kelly, "Notes on effective bandwidths," in *Stochastic Networks*, F. P. Kelly, S. Zachary, and I. Ziedins, Eds. Oxford: Clarendon Press, 1996, pp. 141–168.

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

[16] E. W. Knightly and N. B. Shroff, "Admission control for statistical QoS: Theory and practice," *IEEE Network*, vol. 13, pp. 20–29, 1999.

[17] C. Li, A. Burchard, and J. Liebeherr, "A network calculus with effective bandwidth," Univ. of Virginia, 2003 [Online]. Available: http://www.cs.virginia.edu/\simtechrep/CS-2003-20.pdf

[18] D. N. C. Tse, R. G. Gallager, and J. N. Tsitsiklis, "Statistical multiplexing of multiple time-scale Markov streams," *IEEE J. Sel. Areas Commun.*, vol. 13, no. 8, pp. 1028–1038, Aug. 1995.

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

[20] A. I. Elwalid and D. Mitra, "Effective bandwidth of general Markovian traffic sources and admission control of high speed networks," *IEEE/ACM Trans. Netw.*, vol. 1, no. 3, pp. 329–343, Jun. 1993.

[21] C. S. Chang and J. A. Thomas, "Effective bandwidths in high-speed digital networks," *IEEE J. Sel. Areas Commun.*, vol. 13, no. 8, pp. 1091–1100, Aug. 1995.

[22] C. Courcoubetis, V. A. Siris, and G. Stamoulis, "Application of the many sources asymptotic and effective bandwidths for traffic engineering," *Telecommun. Syst.*, vol. 12, pp. 167–191, 1999.

[23] A. Elwalid, C. Jin, S. Low, and I. Widjaja, "MATE: MPLS adaptive traffic engineering," in *Proc. IEEE INFOCOM*, Apr. 2001, pp. 1300–1309.

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

[25] M. K. Girish, B. Zhou, and J. Q. Hu, "Formulation of the traffic engineering problems in MPLS based IP networks," in *Proc. IEEE ISCC*, Jul. 2000, pp. 214–219.

[26] Y. Seok, Y. Lee, Y. Choi, and C. Kim, "Dynamic constrained multipath routing for MPLS networks," in *Proc. IEEE ICCCN*, Oct. 2001.

[27] R. Guerin and A. Orda, "Computing shortest paths for any number of hops," *IEEE/ACM Trans. Netw.*, vol. 10, no. 6, pp. 613–620, Oct. 2002.

[28] T. H. Cormen, C. E. Leiserson, and R. L. Rivest, *Introduction to Algorithms*, 2nd ed. Cambridge,, MA: MIT Press/McGraw-Hill, 2001.

[29] S. Rai, O. Deshpande, C. Ou, and B. Mukherjee, "Reliable multi-path provisioning for high-capacity optical backbone mesh networks," in *Proc. IEEE ICC*, May 2005, pp. 1741–1745.

[30] V. Siris, Large Deviation Techniques for Traffic Engineering [Online]. Available: http://www.ics.forth.gr/netgroup/msa/

[31] K. G. Coffman and A. Odlyzko, "The size and growth rate of the internet," *First Monday*, vol. 3, no. 10, Oct. 1998.

[32] K. G. Coffman and A. M. Odlyzko, I. Kaminow and T. Li, Eds., "Growth of the internet," in *Proc. SPIE Optical Fiber Telecommun. IV B: Syst. Impairments*, 2002, pp. 17–56.

[33] A. Kantawala and J. Turner, "Queue management for short-lived TCP flows in backbone routers," in *Proc. IEEE GLOBECOM*, 2002, pp. 2380–2384.

[34] M. Kodialam and T. V. Lakshman, "Dynamic routing of restorable bandwidth-guaranteed tunnels using aggregated network resource usage information," *IEEE/ACM Trans. Netw.*, vol. 11, no. 3, pp. 399–410, Jun. 2003.

[35] H. Kochkar, T. Ikenaga, and Y. Oie, "QoS routing algorithm based on multiclasses traffic load," in *Proc. IEEE GLOBECOM*, 2001, pp. 2193–2198.

[36] Y.-W. Chen and R.-H. Hwang, "QoS routing algorithms for multiple traffic classes," in *Proc. IEEE ICC*, 2002, pp. 2217–2221.

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

[38] Y. G. Kim, A. Shiravi, and P. S. Min, "Congestion prediction of self-similar network through parameter estimation," in *Proc. 10th IEEE/IFIP Netw.Oper. Manag. Symp.*, Apr. 2006, pp. 1–4.

[39] C. Huang and V. Sharma, "Building reliable MPLS networks using a path protection mechanism," *IEEE Commun. Mag.*, vol. 40, pp. 156–162, Mar. 2002.

[40] X. Lin and N. B. Shroff, "Simplification of network dynamics in large systems," *IEEE/ACM Trans. Netw.*, vol. 13, no. 4, pp. 813–826, Aug. 2005.

[41] S. H. Low, "A duality model of TCP and queue management algorithms," *IEEE/ACM Trans. Netw.*, vol. 11, no. 4, pp. 525–536, Aug. 2003.

Stephen L. Spitler received the B.S. degrees in electrical and mechanical engineering from the University of California at Davis in 1981 and the M.S. degree in mechanical engineering from Stanford University, Stanford, CA, in 1982. He is currently working toward the Ph.D. degree in electrical engineering at the University of Southern California, Los Angeles.

He has worked in industry on control systems for inertial navigation applications, first at the Northrop Grumman Corporation from 1983 to 1999, and presently as an Engineer/Scientist at the Boeing Company. His research interests are in theoretical aspects of computer and communication networks, including optimal routing, connection admission control, and various issues arising in integrating multiple mechanisms for quality of service.

Daniel C. Lee received the B.S. degree in electrical engineering (with honors) and the B.S. degree in mathematics from the University of Maryland at College Park, both in 1985, and the M.S. and Ph.D. degrees from the Massachusetts Institute of Technology, Cambridge, in 1987 and 1992, respectively, both in electrical engineering and computer science.

From 1993 to 1998, he devoted his research to the systems engineering of networks and communication systems at the U.S. Naval Research Laboratory (NRL), Washington, DC. His active academic career began in 1998, when he joined the faculty of the Electrical Engineering Department, University of Southern California, Los Angeles, as an Assistant Professor. He is currently an Associate Professor at Simon Fraser University, Burnaby, BC, Canada. His main research interests have been quality of service and resource allocation issues in communication systems and networks. Applications of his research include wireless communications and networking, sensor networks, optical networks, and internet multimedia.