Corresponding author:
Debasis Mitra
Bell Labs, Lucent Technologies
Room 2C-382
600 Mountain Avenue
Murray Hill, NJ 07974
Tel: 908 582 4745
Fax: 908 582 3340
e-mail: mitra@lucent.com

TECHNIQUES FOR TRAFFIC ENGINEERING OF MULTISERVICE, MULTIPRIORITY NETWORKS

Debasis Mitra K. G. Ramakrishnan

Bell Laboratories, Lucent Technologies
Murray Hill, NJ 07974

Abstract

We present techniques for traffic engineering in QoS-supported data networks, and also illustrate the application of these techniques in a case study. For scalability we use multicommodity flow (MCF) solution techniques as primitives. The techniques address the design of topology and size of explicit routes in MPLS-supported IP networks and VPNs. The techniques are for network-wide optimization, subject to constraints on routing imposed by end-to-end QoS and other considerations. The notion of admissible route sets is used to differentiate real-time services, such as Internet telephony and video, from delay insensitive services, such as premium data. Different optimization techniques are given for Best Effort services. We also give an efficient and accurate design technique to handle priorities. Finally, a novel technique is presented for obtaining traffic engineering designs for stochastic traffic models from MCF-based designs, with only a small amount of incremental effort.

Area of Paper: Core network design and optimization

Key Words: Multicommodity Flow, Optimization, Quality of Service, Stochastic Traffic Demand, MPLS

TECHNIQUES FOR TRAFFIC ENGINEERING OF MULTISERVICE, MULTIPRIORITY NETWORKS

Debasis Mitra K. G. Ramakrishnan

Bell Laboratories, Lucent Technologies
Murray Hill, NJ 07974

Abstract

We present techniques for traffic engineering in QoS-supported data networks, and also illustrate the application of these techniques in a case study. For scalability we use multicommodity flow (MCF) solution techniques as primitives. The techniques address the design of topology and size of explicit routes in MPLS-supported IP networks and VPNs. The techniques are for network-wide optimization, subject to constraints on routing imposed by end-to-end QoS and other considerations. The notion of admissible route sets is used to differentiate real-time services, such as Internet telephony and video, from delay insensitive services, such as premium data. Different optimization techniques are given for Best Effort services. We also give an efficient and accurate design technique to handle priorities. Finally, a novel technique is presented for obtaining traffic engineering designs for stochastic traffic models from MCF-based designs, with only a small amount of incremental effort.

1 INTRODUCTION

This paper presents techniques, some of which are novel, for traffic engineering in emerging QoS-supported data networks. The case study serves as a medium for illustrating the techniques. An important property required of the techniques is scalability, so that networks ranging from tens to hundreds of nodes and many service classes may be handled. For this reason all the techniques described here use as primitives the techniques for solving MultiCommodity Flow (MCF), and, more generally, Linear Programming (LP) problems [1, 2]. These primitives are polynomial time algorithms, and the user may select from a large range of options, including off-the-shelf packages for moderate sized problems [3], interior point methods for large problems [4, 5], and algorithms based on combinatorial optimization and ϵ-approximation for even larger problems [6, 7, 8, 9]. Speed of response is important since we expect these traffic engineering tools to be used on-line for batches of demands, and on a semi-on-line basis to perform reconfigurations for larger batch sizes. The approach taken in this paper differs fundamentally from traffic engineering tools for capacity planning which, in the pursuit of global optimization and accuracy, deals with large stochastic models and are necessarily off-line. See [10] and references therein for examples.

In comparison, the approach presented in this paper has advantages in scalability and in speed of response to changes in traffic and network conditions. Stochastic variability is, of course, an important component of traffic engineering for data networks. Packet level variability is taken into account in our flow based approach primarily through the notion of effective bandwidth of flows. This important topic is revisited later in this section when we introduce a new, simple method for also taking flow-level variability into account.

A major incentive for the work here is the emergence of traffic engineering support in IP networks, for instance in networks supporting the MPLS protocol. The task of selecting the topology and size of MPLS "explicit routes" to carry QoS-homogeneous traffic is addressed here. MPLS supports multiple paths from source to destination since this offers the potential of better network-wide load balancing, and this point of view fits well with that of the present work. Also it has been observed that it is mutually advantageous to combine DiffServ and MPLS protocols by carrying ToS-homogeneous traffic on explicit routes [11]. We note that similar traffic engineering issues arise in the design of ATM and Frame Relay networks, and the techniques presented in this paper are applicable to these networks as well.

QoS-centric Virtual Private Networks (VPN) introduce additional complexities in traffic engineering. The number of policy and end-to-end constraints in the network proliferate when requirements on routing are introduced by individual VPNs, and the number of VPNs is large, as is expected to be the case. The main point is that the techniques described here for handling QoS constraints are naturally extendible to handle the requirements introduced by VPNs. The size of the problem is larger and the need for scalable techniques is even greater.

This paper develops and illustrates techniques for network-wide optimization subject to constraints on routing imposed by QoS and other considerations. A major motivation is the realization that the converged data networks will support Internet telephony as well as packetized video and other delay sensitive services, in addition to data in a variety of grades of services. To handle such end-to-end constraints we rely on the notion of *admissible route sets*, $\mathcal{R}(s,\sigma)$, which are specific to each QoS service class, s, and (source, destination) pair σ. For instance, real-time services, such as voice and video, may require routes with lengths not exceeding specified thresholds, on account of propagation delay, and also restrictions on the number of hops may apply, since each hop is associated with a switching node and consequent additional delay. In contrast, for delay insensitive service classes, such as premium data, the distance and hop constraints would be less stringent. Admissibility of a route will additionally depend on policy, nodal ability to handle certain services, etc., making the generation of route sets a substantial task in itself. We present techniques that handle routing constraints embedded in the admissible route sets. We also present a contrasting technique which is far more efficient in the absence of such constraints, as in the service class in our case study which we call "Best Effort".

Another important aspect of traffic engineering is priorities. In the case study only a single level of priority is considered; however, the approach naturally extends to multiple priority levels. In the case study priority access to bandwidth is given to the QoS service classes over Best Effort traffic. As a consequence the performance of the QoS service classes is oblivious to the presence of the Best Effort traffic. The design problem with priorities is formulated as a multicriterion optimization problem, where the first objective is the maximization of a weighted combination of QoS-assured carried bandwidth of the QoS service classes, and the second objective is to maximize a similar measure of the Best Effort traffic subject to the constraint that the measure for the QoS traffic is optimal. It is computationally burdensome to obtain an exact solution for this problem. For this reason we develop an efficient and accurate approximation based on a three-layered decomposition of the problem.

Our final major contribution is a technique for obtaining traffic engineering designs for stochastic models and environment by adapting MCF-based designs, with only a small amount of incremental effort. In the stochastic model that we consider the basic entities are flows or calls, which arrive as a stochastic point process and persist for randomly distributed time periods, and if accepted in the network, require an (effective) bandwidth which is characteristic of the burstiness and QoS requirements of their service class. In our technique the multicommodity flow rates, which are inputs to the MCF-based design problem, are inflated to compensate for the variabilities in the stochastic model.

The case study is for a small network with 8 nodes and 20 links, which is convenient because results for this network may be readily reproduced by the interested reader, and also prior results are available for comparison. All the results reported here have been obtained from the software package TALISMAN [12], which has been extended to automate the procedures described here and has been exercised for networks substantially larger than the one described here.

2 MODEL, PROBLEM FORMULATION AND PROCEDURES

We first describe the model of the services, traffic and routes in the network, then describe the problem formulation, and finally give the solution procedures. The results obtained by these procedures for the case study are reported in subsequent sections.

2.1 Model

We consider a connected network with N nodes and L directed links, where link ℓ has bandwidth or capacity C_ℓ. The service classes are classified into QoS and Best Effort. Typically, as in the present case study, there are several QoS service classes, each having characteristic specifications, such as on delay. Here we assume that there is a single Best Effort service class. The QoS group of service classes is denoted by $\mathcal{S}_{\mathrm{QoS}}$.

For each service class s we assume that there is a N by N matrix $\boldsymbol{T}_s$ of bandwidth demands from source node σ_1 to destination node σ_2, i.e., $\boldsymbol{T}_s = \{T_{s,\sigma}\}$. We will find it convenient to refer to (s, σ) as a stream, and to (s, r) as a service-route, where r is a route.

The objectives to be maximized are *network revenues*, which has two components, W_{QoS} and W_{BE} for the QoS and Best Effort service class groups, respectively. Since the former is given priority, comparisons between the two are not particularly meaningful; however, it is still useful to compare the measures from one design to another. We let

$$W_{\mathrm{QoS}} = \sum_{s \in \mathcal{S}_{\mathrm{QoS}}} \sum_{\sigma} \sum_{r \in \mathcal{R}(s,\sigma)} e_{sr} X_{sr} \tag{2.1}$$

where X_{sr} is the flow or carried bandwidth on service-route (s, r), and e_{sr} is the corresponding earnings per unit carried bandwidth. A special case of particular importance is $e_{sr} = e_{s\sigma}$, for all $r \in \mathcal{R}(s, \sigma)$. The earnings parameter may also be used to tune QoS. Also,

$$W_{\mathrm{BE}} = \sum_{\sigma} e_{\mathrm{BE},\sigma} F_\sigma \tag{2.2}$$

where F_σ is the total Best Effort flow or carried bandwidth for (source, destination) pair σ, and $e_{\mathrm{BE},\sigma}$ is the corresponding earning rate. In agreement with the previously stated assumptions on routing of Best Effort traffic, the earning rate does not depend on the route.

2.2 Formulation of the Traffic Engineering Problem

We give the formulation in three pieces, first for the QoS traffic in isolation, second for Best Effort traffic also in isolation, and finally the combined problem.

2.2.1 Admission Control/Routing for QoS Traffic

The problem is

$$W^*_{\mathrm{QoS}} = \max_{\{X_{sr}\}} W_{\mathrm{QoS}} \tag{2.3}$$

subject to

$$\left.\begin{array}{rcl} \sum_{r\in\mathcal{R}(s,\sigma)} X_{sr} & \leq & T_{s,\sigma} \\ X_{sr} & \geq & 0 \;\; \forall r\in\mathcal{R}(s,\sigma) \end{array}\right\} \qquad \forall s\in\mathcal{S}_{\mathrm{QoS}},\ \forall\sigma \qquad \begin{array}{r}(2.4)\\(2.5)\end{array}$$

$$\sum_{s\in\mathcal{S}_{\mathrm{QoS}}} \sum_{\sigma} \sum_{r\in\mathcal{R}(s,\sigma):\ell\in r} X_{sr} \leq C_\ell \qquad \forall\ell \tag{2.6}$$

Here X_{sr} is the flow in service-route (s,r) and the slack in (2.4) represents the loss, which may be ascribed to admission control, in stream (s,σ). The ratio of this loss to $T_{s\sigma}$ may be called the loss ratio, $L_{s\sigma}$.

2.2.2 Admission Control/Routing for Best Effort Traffic

The problem is

$$\max_{\{Y_{\sigma\ell}\}} W_{\mathrm{BE}} \tag{2.7}$$

subject to

$$0 \leq F_\sigma \leq T_{\mathrm{BE},\sigma} \qquad \forall\sigma \tag{2.8}$$

$$\sum_{\ell\in H(n)} Y_{\sigma\ell} - \sum_{\ell\in J(n)} Y_{\sigma\ell} = \left\{\begin{array}{cl} F_\sigma & \text{if } n=\sigma_2 \\ -F_\sigma & \text{if } n=\sigma_1 \\ 0 & \text{otherwise} \end{array}\right\} \qquad \begin{array}{l}\forall \text{ nodes } n,\\ \forall\sigma\end{array} \tag{2.9}$$

$$Y_{\sigma\ell} \geq 0 \qquad \forall\sigma,\forall\ell \tag{2.10}$$

$$\sum_{\sigma} Y_{\sigma\ell} \leq C_\ell \tag{2.11}$$

In (2.9),

$$H(n) = \text{ set of links directed into node } n, \tag{2.12}$$

and

$$J(n) = \text{ set of links directed out of node } n. \tag{2.13}$$

In the above link-based formulation, $Y_{\sigma\ell}$ represents the component of the total flow F_σ from source σ_1 to destination σ_2 that is carried on link ℓ [13]. Note that no part of the flow F_σ that is admitted into the network is dropped inside the network. As before, the slack $(T_{\mathrm{BE},\sigma} - F_\sigma)$ may be considered the loss due to admission control, and the loss ratio may be obtained by dividing it by $T_{\mathrm{BE},\sigma}$.

The solution to the above problem does not directly shed light on the utilized paths or routes, possibly more than one, from σ_1 to σ_2, and the share of F_σ on these routes. This is addressed separately in a procedure called *flow decomposition*, which involves processing the variables $\{Y_{\sigma\ell}\}$ and $\{F_\sigma\}$ from the solution to the problem in (2.7)–(2.13).

2.2.3 Combined Admission Control/Routing for QoS and Best Effort Traffic

The combined problem is

$$\max \; W_{\text{BE}} \tag{2.14}$$

subject to

$$W_{\text{QoS}} = W^*_{\text{QoS}}\,, \tag{2.15}$$

$$\sum_{s\in\mathcal{S}_{\text{QoS}}} \sum_{\sigma} \sum_{r\in\mathcal{R}(s,\sigma):\ell\in r} X_{sr} + \sum_{\sigma} Y_{\sigma\ell} \le C_\ell \quad \forall\ell \tag{2.16}$$

and (2.4), (2.5), (2.8)–(2.10). In (2.15) W^*_{QoS} is obtained from (2.3), i.e., by solving the problem for the QoS traffic in isolation, which has been formulated in Section 2.2.1. Hence, a tacit assumption in the above formulation is that the QoS problem is solved first, followed by the combined problem. Note that the variables $\{X_{sr}\}$, which are common to both problems, may have different values. That is, the only coupling between the two problems is expressed in (2.15).

The combined problem is substantially more complex than either of the problems for the QoS or Best Effort traffic. We give below an approach which decomposes the combined problem and reduces the complexity to the same order as the individual QoS and Best Effort problems, and yet is very effective in approximating the optimal solution. For our case study we compute and report on the exact solution to the combined problem as well as that obtained by the approximation based on the decomposition.

2.3 Multilayered Decomposition of the Combined Problem

In this approach we decompose the combined problem in (2.14)–(2.16) into three layered sub-problems. The first layer simply handles the route-based admission control/routing for the QoS classes, as described in Section 2.2.1, in isolation. The second layer, which requires the solution of a LP of the same order of complexity as the one in the first layer, undertakes resource conservation by minimizing the bandwidth-hops utilized by the QoS traffic, while satisfying the constraint that the earned revenue is optimal, i.e., $W_{\text{QoS}} = W^*_{\text{QoS}}$. An output of the procedure in layer 2 is the residual or unutilized bandwidth C'_ℓ on each link ℓ, which collectively define the residual network $\mathcal{N}'$. In layer 3 the Best Effort admission control/routing problem specified in Section 2.2.2 is solved with C_ℓ replaced by C'_ℓ in (2.11), i.e., the problem is solved on the residual network.

The resource conservation procedure in layer 2 is redundant if there is a unique admission control/routing solution from layer 1 for the QoS traffic. In fact, however, as a result of a phenomenon in typical practical designs, which is well rationalized by a scaling argument from stochastic loss networks but not developed here, the optimal solutions of layer 1 exist in a multidimensional subspace. In fact, in this paper we will provide extensive insights into the structure of the range of possible solutions from layer 1. To this end we employ the tactic of examining the solution of layer 1 which *maximizes* resource usage in bandwidth-hops by the QoS traffic, while also satisfying the constraint $W_{\text{QoS}} = W^*_{\text{QoS}}$. This solution and the solution from layer 2 are, in some sense, the corner points of the range.

A consequence of our multilayered approach is that the residual network $\mathcal{N}'$ on which the Best Effort traffic operates has the property that the sum of its link bandwidths over all links is maximum subject to the constraint $W_{\text{QoS}} = W^*_{\text{QoS}}$. Yet such a maximal residual network may have undesirable characteristics. For instance, it may isolate source nodes from which high earning Best Effort traffic originates or terminates. It is this aspect that the combined problem in Section 2.2.3 addresses and is overlooked in our decomposed solution procedure.

2.4 Route and Link Based Formulations

We draw attention to the differences in the formulations and the variables, $\{X_{sr}\}$ and $\{Y_{\sigma\ell}\}$, in the QoS and Best Effort problem formulations, respectively. The former is route-based, while the latter is link-based [2, 13]. The route-based formulation allows the routes for the different (source, destination) pairs and services to have distinct "policy" constraints. Such routing policy constraints are difficult to incorporate in the link based formulation. On the other hand, when the number of admissible routes is very large the complexity of the route based formulation is correspondingly large. Unless otherwise pruned by policy constraints, this is likely to be the case in richly connected, mesh-like networks where the number of candidate routes may be exponential in the number of nodes. The complexity of the link based formulation depends on the number of links and nodes, but is independent of the number of routes. The above considerations combine to make the route and link based formulations well-matched to the QoS and Best Effort problem layers, respectively.

We next draw attention to the possibility of a further reduction in complexity in the link based formulation of the Best Effort problem layer, at the cost of some loss of generality in the revenue definition, see (2.2). In this formulation only σ_2 would be used to index the flow $Y_{\sigma_2 \ell}$ for destination σ_2 on link ℓ. The reduction in the number of variables, and thereby complexity, is clear. The disadvantage is that in the expression (2.2) for revenue, the earnings rate e_{BE,σ_2} may now depend only on the destination. We do not consider such destination and link based formulation any further in this paper.

2.5 Flow Decomposition

We proceed to give a procedure that identifies the utilized routes, i.e., routes with positive Best Effort traffic flows, for each (source, destination) pair σ, from the link flows $\{Y_{\sigma,\ell}\}$ obtained from the layer 3 solution. The following procedure, which is for a particular σ, is repeated for all σ, starting in each case with the residual network $\mathcal{N}'$.

flow-decomposition (σ)

1. reduce graph of $\mathcal{N}'$ by deleting all links ℓ such that $Y_{\sigma\ell} = 0$, and denote it by $\mathcal{G}$
2. procedure trace-route (depth first search):
   trace a route r in $\mathcal{G}$ from source σ_1 to destination σ_2
   compute $X_{\mathrm{BE},r} = \min_{\ell \in r} Y_{\sigma,\ell}$

$$\begin{aligned} F_\sigma &\longleftarrow F_\sigma - X_{\mathrm{BE},r} \\ Y_{\sigma,\ell} &\longleftarrow Y_{\sigma,\ell} - X_{\mathrm{BE},r} \quad \text{for all } \ell \text{ in } r \end{aligned}$$

3. end if $F_\sigma = 0$
   else goto 1. ∎

We mention that an alternative equivalent to the above procedure requires solving a system of linear equations.

3 NETWORK FOR CASE STUDY

The infrastructure, on which the case study is based, is a fictitious network called Core NSF [12], which has eight nodes ($N = 8$), of which 10 pairs are directly connected, as shown in Figure 1. The network has 20 directed links ($L = 20$), one in each direction for each connected node pair. The typical bandwidth of a directed link is OC3 = 155 Mbps, with the exception of the links connecting Argonne ③ and Princeton ④, and also Houston ⑧ and Atlanta ⑦, which have bandwidths of 2 × OC3 = 310 Mbps. One measure of total resources in the network is 24 OC3-hops.

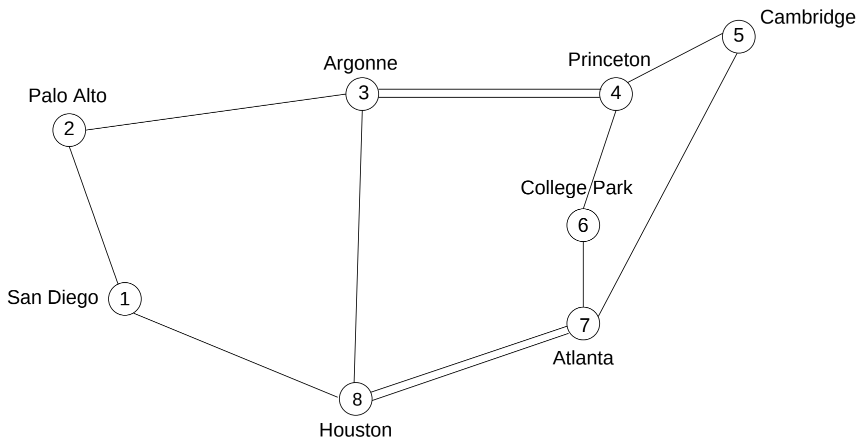

Figure 1: Infrastructure network

There are four service classes, of which three (voice, video and premium data, indexed by $s = 1, 2, 3$, respectively) are QoS, and the fourth is Best Effort. Classes 1 (voice) and 2 (video) are real-time, and have identical admissible route sets, namely, only routes with the minimum number of hops from source to destination are admissible. Class 3 is relatively delay-insensitive and its admissible routes have at most 4 hops, which is more than 3, the diameter of the network graph. There are a total of 68 routes in each of the admissible route sets for classes 1 and 2, and 160 routes in the set for class 3. As previously stated, there are no restrictions placed on the routes for the Best Effort traffic class.

The point-to-point bandwidth demand for the four services are incorporated in the matrices $\{\boldsymbol{T}_s\}$. For convenience we work with a single base matrix $\boldsymbol{T}$, in which the elements are in units of Mbps, and apply multipliers m_s to obtain $\boldsymbol{T}_s = m_s\boldsymbol{T}$. The multipliers are $m_s = 0.4$, 0.1, 0.25 and 0.25, for $s = 1, 2, 3$ and 4, respectively. The base matrix $\boldsymbol{T}$ is given in Table 1.

Since classes 1 and 2 have identical route sets and both belong to the QoS group of service classes, they have identical solutions in the flow optimization study, and may be merged. It is only in Section 5 where we undertake an evaluation of the flow-based design in a stochastic framework, that the two classes have different performances.

Since the sum of all entries in the base matrix $\boldsymbol{T}$ is 1408.8 Mbps, the total demands for the service classes are obtained by applying the appropriate multiplier: 563.5, 140.9, 352.2, 352.2 Mbps.

									out rates
	–	15.7	18.1	2.7	23.5	13.1	5.3	7.9	86.3
	18.3	–	70.1	8.4	90.4	39.2	15.6	23.5	265.5
	20.8	72.7	–	10.4	96.4	46.8	18.1	26.8	292.0
	2.7	8.4	7.9	–	10.6	5.3	2.7	2.7	40.3
	28.7	96.3	101.6	13.2	–	62.3	23.5	39.2	364.8
	13.1	36.4	41.6	5.3	54.9	–	10.4	15.6	177.3
	5.2	13.1	15.6	2.7	20.8	10.4	–	5.3	73.1
	7.9	20.8	26.0	2.7	31.2	15.6	5.3	–	109.5
in rates	96.7	263.4	280.9	45.4	327.8	192.7	80.9	121.0	1408.8

Table 1: Base matrix $\boldsymbol{T}$, in Mbps

4 RESULTS FROM THE CASE STUDY

The results of this section are in three parts: first, we give the results for the three QoS service classes; second, the exact and approximate solutions to the combined problem (Sec. 2.2.3) are presented, together with issues and results related to the resource conservation procedure in layer 2 of the decomposition; finally, qualitative features of the solution for the Best Effort service class are discussed.

4.1 Results for the QoS Service Classes

The results below are obtained after completion of layers 1 and 2. The earnings parameter $e_{sr} \equiv 1$, the default setting, so that the earned revenue is identical to the carried bandwidth. The foremost observation on the results is that all bandwidth demands are carried with QoS assurance, i.e., there is no loss. In other words, (2.4) holds with equality. The results are summarized in Table 2.

service class	voice (1)	video (2)	premium data (3)	total
bandwidth demand = carried QoS-assured bandwidth (Mbps)	563.5	140.9	352.2	1056.6

Table 2: Results for QoS service classes

Even though the network resources were not severely strained, as indicated by the no loss condition, the optimal routes that were obtained exhibited some notable departures from their shortest paths. For instance, the Cambridge ⑤ to Palo Alto ② traffic is routed as follows: the real-time (classes 1 and 2) streams are routed via Princeton ④ and Argonne ③, as this is the only admissible route for these streams, while the non-real-time (class 3) stream is split, with 78% on the above 3-hop shortest path and the remaining 22% on the longer 4-hop path via Atlanta ⑦, Houston ⑧ and San Diego ①. Another interesting case is traffic from College Park ⑥ to Cambridge. Class 1 and 3 streams are routed via Atlanta, and the class 2 stream is routed via Princeton. Both routes are 2-hop and shortest.

The following features of the residual network $\mathcal{N}'$ are helpful in understanding the routing dispositions for the QoS traffic. The Cambridge to Princeton link has zero capacity, even though the Princeton to Cambridge link has positive capacity. In contrast to the route splitting from Cambridge to Palo Alto, traffic of all QoS classes from Palo Alto to Cambridge are routed on the 3-hop shortest path via Princeton.

4.2 Combined Problem, Resource Conservation

The combined problem in Section 2.2.3 is solved exactly and also by the approximate three-layered technique described in Section 2.3. We select the earnings parameters $e_{\mathrm{BE},\sigma} \equiv 1.0$, see (2.2), which is the default setting for which the earned revenue W_{BE} is simply the total carried Best Effort bandwidth. The two values of W_{BE} that are obtained are almost identical, with the difference so small as to be plausibly due to numerical errors. The results are in Table 3. Note the heavy usage of the residual network, which together with the uniformity of the earnings parameter constitute conditions in which we expect the approximate solution to be close to the optimum.

total Best Effort bandwidth demand (Mbps):	352.2
W_{BE}, total Best Effort bandwidth carried (Mbps):	279.4
fraction of bandwidth lost (%):	20.5

Table 3: Results for the Combined Problem: exact and approximate solutions

The rest of this section is devoted to the role of layer 2, see Section 2.3, which undertakes conservation of resources utilized by the QoS traffic. As mentioned before, for comparative illustration we obtain the solution that maximizes resource usage, measured in bandwidth-hops, while carrying all QoS traffic demands, just as in the layer 2 solution. The results for these two solutions are in Table 4. Note, by using the result in Table 2, that the average number of hops in routes are about 2.0 and 2.5 in the two solutions, respectively.

Another view of the efficiency of resource conservation is obtained by examining the residual network $\mathcal{N}'$. Let $\mathrm{cap}(\mathcal{N}')$ denote the total capacity of $\mathcal{N}'$, which is obtained by

minimum resource usage by QoS service classes	= 2120 (Mbps-hops)
(layer 2 solution)	
maximum resource usage by QoS service classes	= 2668 (Mbps-hops)

Table 4: Resource usage by QoS traffic

cap($\mathcal{N}'$) if resource usage by QoS classes is minimized	= 10.3	OC3
cap($\mathcal{N}'$) if resource usage by QoS classes is maximized	= 2.85	OC3

Table 5: Residual network capacity

summing all its link capacities. The results for the two extreme solutions are presented in Table 5. For perspective recall that cap($\mathcal{N}$) = 24 OC3.

The effect on connectivity of $\mathcal{N}'$ is noteworthy too. With minimum resource usage by QoS classes only 3 links (Palo Alto $\leftrightarrow$ Argonne, Cambridge $\rightarrow$ Princeton) have zero capacity, while with maximum resource usage 12 links have zero capacity. In the former solution no node is completely isolated, while in the latter solution no Best Effort traffic can leave Palo Alto and Cambridge is isolated in both in and out directions.

We should point out that there is an important hidden cost to optimal resource usage by QoS classes. It has the effect of packing routes tightly, which is the opposite, in a certain sense, of load balancing. On the other hand, maximizing resource usage yields well balanced loads on routes. The effect of this dichotomy is not visible in the present flow-based framework, but shows up in the evaluation of designs in the stochastic framework, which is studied in the next section. Here we mention only that there is an important role for less aggressive resource conservation than is implemented in layer 2, for instance by introducing constraints that require link utilizations not to exceed certain levels, which may either be prespecified or obtained adaptively in the course of the design process.

4.3 Results for the Best Effort Class

We comment on the qualitative nature of the solutions obtained for Best Effort traffic routes from the flow decomposition procedure in Section 2.5. The procedure yields routes with positive flows having as many as 5 hops. Not too surprisingly, there are some extreme features in the optimal solution for the Best Effort class. For instance, all Best Effort traffic from Cambridge to Palo Alto and from Cambridge to College Park is lost, while the fraction of bandwidth from Palo Alto to Cambridge that is lost is 78%. On the other hand, Best Effort traffic due to all node pairs, with the exclusion of the aforementioned and all Best Effort traffic originating at Cambridge, is carried without any loss. Such extreme features may be considered unfair and therefore unacceptable. In this case the flexibility afforded by the earnings parameters $e_{\mathrm{BE},\sigma}$ may be exercised to obtain less extreme solutions. For instance, the default setting of the earnings parameters is obviously responsible for streams in which the source and destination are distant to suffer higher losses. The earnings parameters may be tuned to compensate for such phenomenon.

5 ADAPTATION OF FLOW-BASED DESIGNS FOR STOCHASTIC ENVIRONMENTS

In this section we will provide a framework for obtaining, with only a small amount of incremental effort, stochastic traffic engineering designs from flow-based designs, which have been the subject of this paper up to this point. There are two main ideas. First, flow rates, which are inputs to the flow-based design, are inflated to compensate for the variabilities in the stochastic environment. Second, the flow-based design on completion is mapped into the stochastic environment by holding invariant certain ratios of naturally corresponding traffic intensities. Also in this section we briefly recapitulate the analytic technique for evaluating performance in the stochastic environment, before proceeding to present performance results for the case study. Although the methodology clearly applies also to design for Best Effort traffic, we will only consider the QoS traffic in this section.

First let us provide some high level motivation for the need for a bridge to the stochastic environment. The "naive" approach is to apply the admission control/routing decisions from the flow-based design in which the traffic demands $\{T_{s,\sigma}\}$ are simply the corresponding mean values of processes in the stochastic model. The effectiveness of this approach may be judged by comparing the loss ratios $L_{s\sigma}$ from the flow-based designs to the loss probabilities, say also denoted by $L_{s\sigma}$, for streams (s, σ). In the case study, as previously noted, for all QoS service classes and streams, the loss ratios are zero. On the other hand, as we shall see in Section 5.5 and Table 6, the same design evaluated in the stochastic environment gives loss probabilities for certain streams as high as 0.137. This discrepancy may be considered unacceptably high. Moreover, it is clear that the reason that the naive approach gives overly optimistic performance predictions is that it is oblivious to the effects of variabilities in the stochastic model.

5.1 Flow-Based Design Outputs

We briefly review the outputs from the flow-based designs.

(a) X_{sr}, the "size" of the service-route (s, r).

(b) $(T_{s\sigma} - \sum_{r \in \mathcal{R}(s,\sigma)} X_{sr})$, the bandwidth of stream (s, σ) that is lost due to admission control.

5.2 Stochastic Traffic Model

The classical multiservice, multirate, connection-level, stochastic loss network model is considered here. See, for instance, Kelly [14], Ross [15] and Mitra, Morrison and Ramakrishnan [10, 12]. The notation here follows [10]. Here IP flows [16], connections and calls are terms used interchangeably. Each call of service type s requires (effective) bandwidth d_s. The notion of effective bandwidth [17, 18] encapsulates various packet-level issues, such as burstiness and QoS (delay, jitter, loss) at network elements. Requests for calls of stream (s, σ) arrive in a Poisson process with mean rate $\Lambda_{s\sigma}$. Admission control is administered in the form of a Bernoulli trial to determine if an arriving call of stream (s, σ) is offered to the network, and, if it is, to which route $r \in \mathcal{R}(s, \sigma)$ it is offered. If an arriving call is offered to a route, it may still be lost if there is insufficient bandwidth available on any link in the route. A practical interpretation of the associated loss probability is that it is the fraction of calls offered to the network for which the QoS requirements of its service class are not fully met. If the arriving new call is not lost, then the bandwidth on each link in its route is simultaneously held for the duration of the call. The holding period of a call of stream (s, σ) is random, generally distributed with mean $1/\mu_{s\sigma}$ and independent of earlier arrival and holding times. We let $\bar{\rho}_{s\sigma} = \Lambda_{s\sigma}/\mu_{s\sigma}$ denote the load or traffic intensity for the stream, and note that $d_s\bar{\rho}_{s\sigma}$ represents its mean bandwidth requirement.

We note in passing that in real heterogeneous networks the effective bandwidth of calls depends not only on the service type but also on the link, i.e., the representation should be $d_{s\ell}$. We handle this case but omit further discussion since it complicates the exposition unduly.

The asymptotic scaling that interests us is classical [19], and in it $d_s = O(1)$, while link capacities C_ℓ and loads $\bar{\rho}_{s\sigma}$ are both large.

5.3 Compensating for Variability

The basic mechanism for compensation that we propose is contained in the following expression:

$$T_{s\sigma}/d_s = \bar{\rho}_{s\sigma} + \alpha\sqrt{\bar{\rho}_{s\sigma}} \tag{5.1}$$

where $\alpha (\alpha \geq 0)$ is a parameter. For the stream (s, σ), the expression relates the bandwidth demand $T_{s\sigma}$ for the flow-based models to the parameters $(\bar{\rho}_{s\sigma}, d_s)$ of the stochastic model. The naive approach sets the *compensation factor* α to 0. The greater the variability in the point processes of the stochastic model, greater is the need for compensation and a larger value of α. For Poisson arrivals of calls and exponentially distributed holding times, and the network in critical loading, $\alpha \approx 0.5$ is appropriate, in our experience. The results presented below support this choice.

For $\alpha = O(1)$, $d_s = O(1)$ and both $\bar{\rho}_{s\sigma}$ and $T_{s\sigma}$ large, the typical case of interest, (5.1) may be inverted to yield the following to leading order of magnitude:

$$\bar{\rho}_{s\sigma} \sim T_{s\sigma}/d_s - \alpha\sqrt{T_{s\sigma}/d_s}\,. \tag{5.2}$$

To avoid negative values we qualify (5.2) so that it holds if $T_{s\sigma}/d_\sigma > \alpha^2$, and otherwise we let $\bar{\rho}_{s\sigma} = T_{s\sigma}/d_s$.

We should point out that in the literature on the asymptotic analysis of loss networks there is a long history of scalings related to the form in the right of (5.1). See, for instance, Jagerman [19].

5.4 Design Procedure

We now proceed to give the steps in our procedure for obtaining stochastic traffic engineering designs from flow-based designs. We assume here that the input parameters of the stochastic model, $\{\Lambda_{s\sigma}\}$, $\{\mu_{s\sigma}\}$, $\{d_s\}$, are given and the compensation factor α has been fixed. (Recall that $\bar{\rho}_{s\sigma} = \Lambda_{s\sigma}/\mu_{s\sigma}$.) The outputs of the design that are sought are the analogs of the quantities in Section 5.1, namely,

$\rho_{sr} = \lambda_{sr}/\mu_{s\sigma}$, the load (or traffic intensity) on the service-route (s, r), and

$(\bar{\rho}_{s\sigma} - \sum_{r \in \mathcal{R}(s,\sigma)} \rho_{sr})/\bar{\rho}_{s\sigma}$, the fraction of arriving calls of the stream (s, σ) that are not offered to the network due to admission control.

The procedure is as follows:

1. Obtain the traffic demands $\{T_{s\sigma}\}$ for the flow model from (5.1).

2. Obtain flow-based design, which gives $\{X_{sr}\}$.

3. Obtain $\{\rho_{sr}\}$ by utilizing the following relation:

$$\frac{\rho_{sr}}{\bar{\rho}_{s\sigma}} = \frac{X_{sr}}{T_{s\sigma}} \quad (r \in \mathcal{R}(s,\sigma)) \,. \tag{5.3}$$

A final step is the performance evaluation in the stochastic environment of the admission control/routing design thus obtained. Here we rely on the well-known approach [14, 15, 12] of fixed-point equations. This is an approximation technique based on the assumption that each link in the network acts to independently thin Poisson arrival processes, with the fixed point equations determining the thinning probabilities. A great deal is known about the nature of errors in this approximation, and also considerable experience has been accumulated in the use of the techniques. For typically large networks with reasonable mixing of flows, the technique has proven quite adequate in accuracy for design purposes. The performance results of interest are the loss probabilities L_{sr} for service-route (s, r).

5.5 Results from the Case Study

Ideally, for the case study, we would like to exhibit the procedure described above in Section 5.4. However, we have already obtained a flow-based design for input parameters $\{T_{s,\sigma}\}$, and have reported the results in Section 4. For this reason, we will retain this flow-based design as the benchmark, and compare its performance results, specifically loss ratios, to the corresponding results from the design for the stochastic model and evaluated for the stochastic environment. We will consider various values of the compensation factor, α; the effective bandwidths of individual calls $\{d_s\}$ are held fixed. Specifically, for each α, we follow the procedure below.

1. Compute $\bar{\rho}_{s\sigma}$ from $T_{s,\sigma}$ using (5.2).

2. Compute ρ_{sr} from (5.3) in which the right hand quantities are obtained from the benchmark flow-based design.

3. Evaluate performance, specifically loss probabilities $\{L_{sr}\}$, for the stochastic environment using fixed point equations for the design based on $\{\rho_{sr}\}$.

We will consider $\alpha = 0.0$, 0.5 and 1.0; recall that $\alpha = 0$ corresponds to the previously described "naive" approach. The effective bandwidth of individual calls of the three QoS classes are as follows.

$$\begin{array}{ll} \text{voice calls } (s=1), & d_1 = 16 \text{ Kbps} \\ \text{video calls } (s=2), & d_2 = 640 \text{ Kbps} \\ \text{premium data calls } (s=3), & d_3 = 384 \text{ Kbps} \end{array} \tag{5.4}$$

Now, recall that the performance results for the QoS service classes from the benchmark are simply stated: loss ratios on all service routes are zero. For the stochastic

compensation factor, α	$\alpha = 0$	$\alpha = 0.5$	$\alpha = 1.0$
total bandwidth demand (Mbps), $\sum_{s,\sigma} \rho_{s\sigma}$	1056.3	978.8	901.3
total carried QoS-assured bandwidth (Mbps), $\sum_{s,\sigma} \sum_{r \in \mathcal{R}(s,\sigma)} \rho_{sr} d_s (1 - L_{sr})$	1042.2	976.0	901.2
network-wide loss probability	0.0134	0.003	0.0001
service-routes with highest loss probabilities			
(video, Cambridge $\rightarrow$ Palo Alto)	.137	.035	.0017
(video, Palo Alto $\rightarrow$ Cambridge)	.122	.026	.0006
(video, Palo Alto $\rightarrow$ Argonne)	.073	.015	.0003
(video, Palo Alto $\rightarrow$ Princeton)	.073	.015	.0003
(video, Palo Alto $\rightarrow$ College Park)	.073	.015	.0003
(video, Cambridge $\rightarrow$ Princeton)	.071	.021	.0014
(video, Cambridge $\rightarrow$ Argonne)	.071	.021	.0014
(video, College Park $\rightarrow$ Palo Alto)	.071	.014	.0004
(video, Argonne $\rightarrow$ Palo Alto)	.071	.014	.0004
(video, Princeton $\rightarrow$ Palo Alto)	.071	.014	.0004

Table 6: Results from the stochastic model/design/performance evaluation

model/design/performance evaluation, small α has the advantage of requiring small compensation, but yields large loss probabilities, and therefore significant discrepancies between the two approaches. Hence, our goal is to obtain the smallest value of the compensation factor α such that the differences between the loss ratios and loss probabilities on service routes in the two approaches are all just within a user-specified error tolerance.

Our results are presented in Table 6. A notable feature is that the video service class is uniformly the most adversely affected in QoS. This phenomenon is expected and stems from the well-known feature of loss networks that calls with the most demanding bandwidth requirement are affected most. Notice too that $\alpha = 0$ and $\alpha = 1$ are less acceptable (for different reasons) than $\alpha = 0.5$.

6 CONCLUSIONS

We have presented and illustrated scalable techniques, which use multicommodity flow solution methods as primitives, for handling QoS requirements in the traffic engineering of data networks. Novel techniques have been presented for (a) adapting efficient MCF-based design methods for stochastic models, (b) incorporating priorities in the design by an efficient and accurate approximation based on a three-layered decomposition, and (c) matching the requirements of QoS and Best Effort traffic to the solution techniques, route and edge-based respectively, for their routing/admission control problems. Some topics for further investigation are as follows. There is great interest in evaluating various new and powerful combinatorial optimization and ϵ-approximation techniques for extremely large data networks, especially when the time requirements are severe. Also, the results presented in Section 4 suggest other formulations of the combined (priority) problem in which the incentive for packing the higher priority QoS traffic is not as great, thus allowing greater load balancing. In fact, such modifications are more conveniently made in layer 2 of our three layered decomposition. Similarly, it will be beneficial to investigate techniques for achieving greater fairness among (source, destination) pairs by tuning the earnings parameters, i.e., by increasing earnings for more widely separated terminal nodes. Yet another topic for further investigation is the comparative performance of sequential optimization where demands are routed one at a time, in contrast to the joint optimization of batches of demands as discussed here.

References

[1] D. G. Luenberger. *Linear and Nonlinear Programming.* Addison-Wesley, Reading, MA, 2nd edition, 1984.

[2] D. Bertsekas and R. Gallager. *Data Networks.* Prentice Hall, Englewood Cliffs, NJ, 2nd edition, 1992.

[3] B. A. Murtagh and M. A. Saunders. MINOS 5.1, User's Guide. Tech. Report SOL83-20R, Systems Optimization Laboratory, Dept. of Operations Research, Stanford University, 1983, revised 1987.

[4] I. Adler, N. K. Karmarkar, M. G. C. Resende, and G. Veiga. Data structures and programming techniques for the implementation of Karmarkar's algorithm. *ORSA J. Computing*, 1(2), Spring 1989.

[5] N. K. Karmarkar and K. G. Ramakrishnan. Computational results of an interior point algorithm for large scale linear programming. *Mathematical Programming*, 52:555–586, 1991.

[6] C. Stein. *Approximation Algorithms for Multicommodity Flow and Scheduling Problems.* PhD thesis, MIT, 1992.

[7] N. Garg and J. Konemann. Faster and simpler algorithms for multicommodity flow and other fractional packing problems. *Proc. IEEE Symp. on Foundations of Computer Science*, 1998.

[8] R. Hassin. Approximation schemes for the restricted shortest path problem. *Math. of Oper. Res.*, 17:36–42, Feb. 1992.

[9] A. Goel, D. Kataria, D. Logothetis, and K. G. Ramakrishnan. An efficient algorithm for constraint-based routing in data networks. To appear in *Proc. IEEE INFOCOM*, 2001.

[10] D. Mitra, J. A. Morrison, and K. G. Ramakrishnan. Virtual private networks: Joint resource allocation and routing design. *Proc. IEEE INFOCOM*, 1999.

[11] L. Wu, P. Cheval, and P. Vaananen. MPLS extensions for Differential Services. IETF draft-wu-mpls-diff-ext-00.txt, work in progress, Nov. 1998.

[12] D. Mitra, J. A. Morrison, and K. G. Ramakrishnan. ATM network design and optimization: A multirate loss network design. *IEEE/ACM Trans. Networking*, 4(4):531–543, 1996.

[13] S. Ahuja and J. Orlin. *Network Flows.* Addison-Wesley, Reading, MA, 1998.

[14] F. P. Kelly. Routing in circuit-switched networks: optimization, shadow prices and decentralization. *Adv. Appl. Prob.*, 20:112–144, 1988.

[15] K. W. Ross. *Multirate Loss Models for Broadband Telecommunications.* Springer, New York, 1995.

[16] K. Thompson, G. J. Miller, and R. Wilder. Wide-area Internet traffic patterns and characteristics. *IEEE Network*, Nov./Dec. 1997.

[17] *IEEE J. Selected Areas Communications*, 13(6), August 1995.

[18] M. Schwartz. *Broadband Integrated Networks.* Prentice Hall, New York, 1996.

[19] D. L. Jagerman. Some properties of the Erlang loss function. *Bell System Tech. J.*, 53:525–551, March 1974.