Link Dimensioning and LSP Optimization for MPLS Networks Supporting DiffServ EF and BE traffic classes * †

Kehang Wu and Douglas S. Reeves
Departments of Electrical and Computer Engineering and Computer Science
North Carolina State University
kwu@unity.ncsu.edu, reeves@eos.ncsu.edu

Capacity planning is indispensable for future Internet providing QoS. Accurate dimensioning is especially important when no per-flow signaling or control exists.

In this paper, we address the problem of link dimensioning and Lable Switching Path (LSP) optimization for MPLS networks supporting DiffServ EF and BE traffic classes. The problem is formulated as an optimization problem, where the goal is to minimize the nonlinear total link cost, subject to the performance constraints of both expedited forward (EF) and best effort (BE) traffic classes. The variables to be determined are the routing of LSPs carrying both EF and BE traffic, and the discrete capacities of the links.

We show that Lagrangean Relaxation and subgradient optimization methods can be used to effectively solve this difficult problem. Computational results show that the solution quality is verifiably good, while the running time remains reasonable on practical-sized networks. This is the first work on capacity planning for MPLS networks supporting multiple DiffServ service classes.

1. Introduction

Capacity planning is the process of designing and dimensioning networks to meet the expected demands on them. The nature of offering only best effort (BE) service makes capacity planning a straightforward matter[1] in the current Internet. With the popularization of e-commerce and new value-added services over IP, Quality of service (QoS), the ability of a network element to have some level of assurance that its traffic and service requirements can be satisfied, has become a must. Capacity planning will be an imperative part of IP network management to support various qualities of service.

Multiprotocol label switching (MPLS) [2] [3] and Differentiated services (DiffServ) [4] [5] are regarded as two key components for providing QoS in the Internet.

MPLS uses a short, fixed-length, locally significant label in the packet header to switch the packets. The initial label is chosen and inserted by the ingress node of a MPLS domain, based on the information in the IP header, associated QoS, or any other policies in effect. The intermediate nodes use the label as an index to find the next hop and the corresponding new label. A label distribution protocol (LDP) propagates label bindings among the nodes in order to establish and tear down the label switched path (LSP). The power of MPLS lies in the fact that the mappings between the packet flows and the LDPs are flexible, which enables IP network to be traffic engineered. Packets with the same source and destination addresses, which will inevitably follow the same path in the traditional IP network, can be assigned different labels

*This work is supported by DARPA and AFOSR (under grants F30602-99-1-0540 and F49620-99-1-0264).
†The authors wish to thank Dr. Cole Smith and Dr. Michal Pioro for their helpful comments and suggestions.

and subsequently be sent to separate LSPs. LSPs can be setup explicitly to optimize the resource utilization. The use of MPLS labels may also provide faster switching than the normal IP forwarding algorithm.

The essence of DiffServ is prioritization. In traditional IP networks, the DiffServ Code Point (DSCP) field in the headers of IP packets is marked at the edge of the network. Routers within the core of the network forward packets using different predefined per-hop behaviors (PHBs), according to their DSCP field. In MPLS networks, the DSCP field is not visible to the core LSRs. The label value and EXP field (3bit experimental field) are used instead to determine the PHB scheduling class (PSC) associated with packets [6]. Note that even with MPLS, the signaling and traffic control are still at the level of flow aggregation. Since there is no per-flow signaling or control, accurate dimensioning of the network is particularly important for achieving performance guarantees. To prepare for the deployment of DiffServ in MPLS network, it is necessary to study the capacity planning problem in the context of multiple class-of-service networks.

The IETF DiffServ working group has standardized two PHBs: Expedited Forwarding (EF) and Assured Forwarding (AF). The EF PHB [7] is defined as being such that the EF packets are guaranteed to receive service at or above a configured rate. The EF PHB can be used to build a low loss, low latency, low jitter, assured bandwidth , end-to-end service, through a DiffServ Domain. As has been discussed in [8], three expected major initial applications of QoS in the IP network are: 1) to distinguish "mission critical" or preferred customers; 2) to provide voice over IP service; 3) to enable services competitive with leased lines. It can be easily seen that the services based on the EF PHB are ideal for all three of these applications. Because of its great value, the EF PHB is very likely to be the first PHB to be put into action. The priority queue is widely considered to be the canonical way to implement the EF PHB, due to its ability to offer a tighter delay bound and smoother service over relatively short time scales [7]. AF PHBs are designed to realize different forwarding assurances, or dropping preferences, for IP packets. AF PHBs are considered useful to differentiate TCP traffic, where the performance is sensitive to the packet loss rate. However, simulations showed that the standard traffic control methods of routers, such as RED(Random Early Detection), do not satisfactorily differentiate between AF PHBs and best effort traffic [9]. Since our approach requires a precise performance model for optimization, we do not include the AF traffic classes in this paper, due to the lack of a consensus on the implementation of the AF PHB.

In this paper we address the problem of link dimensioning and path optimization for MPLS networks providing DiffServ EF and BE traffic classes. The problem is formulated as an optimization problem, where we jointly select the routes for edge to edge EF and BE user demand pairs, and assign a discrete capacity value for each link. The goal is to minimize the total link cost, subject to the performance constraints of both EF and BE classes. The non-bifurcated (i.e., single-path) routing model is used for the EF class as required by [7], so that the traffic from a single EF demand pair will follow the same LSP between the origin and the destination. Traffic in the BE class is allowed to be split across multiple LSPs. While the performance constraint of EF traffic is only represented by a bandwidth requirement, the performance constraint of the BE class is characterized by the average delay in each link. Queueing is modeled as M/G/1 strict priority queues. Our intention is to not only define the capacity planning problem for MPLS network supporting DiffServ and disclose feasible solutions, but also provide helpful insights for capacity planning of other QoS architectures.

Although there is no previous work specifically targeting the dimensioning and routing problems for MPLS networks supporting Diffserv, there are many papers addressing the issues of QoS routing in general. An extensive survey can be found in [10]. But because the routing and link dimensioning problems are closely related to each other, it is inappropriate to separate them. There are extensive works dealing with traffic engineering issues in MPLS network, such as [11] [12], or LSP setup and dimensioning problem [13]. However, in those works the link capacity is fixed and not subject to be optimized.

Papers where the routing and capacity assignment problems are treated simultaneously include [14] [15] [16][17] [18] [19] [20] [21]. Gerla and Kleinrock [14] presented heuristic methods based on the flow deviation algorithm [22]. Gavish and Neuman [15] formulated the problem as a non-linear integer programming problem, and proposed a Lagrangean relaxation based approach. [21] studied the network with elastic traffic and approximated the non-linear cost function to a piece-wise linear function. The networks studied in [14] [15] [21] only include one traffic class, though. Medhi and Tipper [19] proposed four approaches for reconfigurable ATM networks based on the Virtual Path concept. Even though ATM networks include multiple traffic classes, Medhi proposed a model that assumes the deterministic multiplexing of different virtual paths, which results in linear performance constraints.

We studied the problem of capacity planning for DiffServ networks in a previous work [23], where the MPLS protocol is not supported. Because of the absence of MPLS in [23], traffic demands with the same origin and destination will be constrained to follow the same path. [23] only considers the routing of EF demand pairs, while the routing of both EF and BE classes are optimized in this paper. A nonlinear cost function is assumed in this paper, while [23] uses a linear cost function.

The novel aspect of our capacity planning problem is the fact that two traffic classes, EF and BE, with independent behaviors and performance requirements, share the same capacity resource, which results in a complex non-linear performance constraint. In addition to the non-linear performance constraint, non-bifurcated routing and discrete link capacity constraints dramatically increase the degree of difficulty, and significantly limit the viable solution approaches.

The remainder of this paper is organized as follows. In Section 2, notation and detailed assumptions and models are presented. The problem definition is given in Section 3. Section 4 shows a Lagrangean relaxation of the original problem, and describes the subgradient procedure to solve the resulting dual problem. Section 5 presents some numerical results on the use of the method. The paper is concluded in Section 6.

2. Notation and Models

The following notation will be used throughout the paper.

K	set of (both EF and BE) Origin-Destination (O-D) pairs [3]
M_k	set of EF demands for O-D pair k, $k \in K$
L	set of links in the network
J_k	set of possible candidate LSP paths for the O-D pair k, $k \in K$
δ_j^l	link-path indicator; 1 if path j uses link l, $j \in J_k, k \in K, l \in L$, 0 otherwise
α_{km}^{ef}	average arrival rate of an EF traffic demand m, $m \in M_k$, $k \in K$
ρ_{km}^{ef}	requested bandwidth of an EF traffic demand m, $m \in M_k$, $k \in K$
x_{kmj}^{ef}	EF path routing variable; 1 if EF demand $m, m \in M_k$, $k \in K$ uses path $j \in J_k$, 0 otherwise.
η_l	total requested bandwidth of EF demand on link $l \in L$
β_l^{ef}	average arrival rate of total EF traffic demand on link $l \in L$
x_{kj}^{be}	BE path routing variable: the portion of BE demand k uses candidate path $j, j \in J_k$. x_{kj}^{be} can be any real value between 0 and 1
α_k^{be}	average arrival rate of a BE traffic demand, $k \in K$
γ_l	average arrival rate of extra BE traffic demand on link l, $l \in L$
β_l^{be}	average arrival rate of total BE traffic demand on link $l \in L$

d_l	average delay experienced by BE traffic on link $l \in L$
d_{lmax}	maximum value of d_l allowed for link $l \in L$
g_l	BE delay bound factor
T_l	index of available link types for link $l \in L$
u_{lt}	link type decision variable; 1 if link type t is used for link $l \in L$, 0 other wise.
ψ_{lt}	size of the capacity of link type $t, t \in T_l$
$\tilde{\psi}_l$	total capacity of link l, $l \in L$
C_{lt}	cost of the link type t, $t \in T_l$, in link l
$\tilde{C}_l$	total cost of link l, $l \in L$
$\tilde{y}$,$\widetilde{y^2}$	the first and second moment of packet size, (units: bits & bits^2)

Link based formulation is used in this paper. The network is defined by (L, K, J_k).

For an EF demand $m, m \in M_k$, we differentiate between the average arrival rate, α^{ef}_{km}, and the requested bandwidth, ρ^{ef}_{km}. ρ^{ef}_{km} is usually a value between the average arrival rate and the peak rate. It is noted in [7] that the packets of the EF traffic class belonging to the same flow should not be reordered. Consequently, traffic from the same EF demand can not be separated into different LSPs.

$$\beta^{ef}_l = \sum_{k \in K} \sum_{m \in M_k} \alpha^{ef}_{km} \sum_{j \in J_k} x^{ef}_{kmj} \delta^l_j \tag{1}$$

$$\eta_l \;=\; \Gamma(\{\rho^{ef}_{km}\}, \{\delta^l_j\}, \{x^{ef}_{kmj}\}) \;\leqslant\; \sum_{k \in K} \sum_{m \in M_k} \rho^{ef}_{km} \sum_{j \in J_k} x^{ef}_{kmj} \delta^l_j \tag{2}$$

where $\Gamma()$ is the EF demand multiplexing function (discussed further below). The inequality becomes equality only when there is no multiplexing gain.

For each O-D pair k , only one BE demand pair is defined. We allow an arbitrary portion of the BE demand to route through any candidate LSP. Therefore the aggregation of BE traffic would potentially improve the effectiveness of traffic engineering.

Because of the connectionless nature of IP traffic, it is unlikely that all the BE demand can be clearly mapped to specific O-D pairs. We introduce another variable γ_l, which is the average arrival rate of extra BE traffic in link l besides that from the BE demand pairs $\{k : k \in K\}$. Thus the total BE load (average arrival rate) on link l is:

$$\beta^{be}_l = \gamma_l + \sum_{k \in K} \alpha^{be}_k \sum_{j \in J_k} x^{be}_{kj} \delta^l_j \tag{3}$$

The capacity and the cost of link l, $\tilde{\psi}_l$ and $\tilde{C}_l$ respectively, are:

$$\tilde{\psi}_l = \sum_{t \in T} u_{lt} \psi_{lt}, \qquad \tilde{C}_l = \sum_{t \in T} u_{lt} C_{lt} \tag{4}$$

There is no linear relationship assumed between C_{lt} and ψ_{lt}, therefore $\tilde{C}_l$ is not necessarily a linear function of $\tilde{\psi}_l$.

There are many discussions about the original EF PHB [24] concerning the limits on EF utilization. Charny reported in [25] that the worst case delay jitter can be made arbitrarily large using a FIFO queue unless the utilization of EF traffic was limited to a factor smaller than $1/(H-1)$, where H is the number of hops in the longest path of the network. Other implementations of packet scheduling may improve the upper bound on the EF utilization. The revised EF PHB [7], RFC 3246, introduces an error term E_a for the treatment of the EF aggregate, which represents the allowed worst case deviation between the actual EF packet departure time and the ideal departure time of the same packet. It is not immediately clear whether this revision totally eliminates the constraint on the EF utilization, or simply allows a trade-off between the EF utilization and the delay jitter. In this paper, we assume that the projected EF user demand η_l is much less than the capacity of the link, so there is no concern about this limit on the EF utilization, and the exact form of the multiplexing function $\Gamma()$ does not have any impact on the final solutions.

How to specify the performance of BE traffic in the service level agreement (SLA) is still an active research topic. [26] suggests using the latency averaged over a large time scale as the primary criteria for the performance of BE traffic in IP network service level agreements (SLAs). We pick the average delay as the sole performance measurement for BE traffic in this paper. We evaluate the performance of BE traffic on a per-link basis (i.e., not end-to-end). The value $\frac{\tilde{y}}{\tilde{\psi}_l}$ stands for the average transmission delay of packets. We use $\frac{\tilde{y}}{\tilde{\psi}_l}$ as the basis for the delay bound. Let $d_{lmax} = g_l \frac{\tilde{y}}{\tilde{\psi}_l}$, where g_l is a parameter defined by the network designer. The larger the value of g_l, the more bandwidth is required for link l, therefore the lower the link utilization. We assume that the performance of BE traffic is satisfactory if $d_l \leqslant d_{lmax}$.

Every router is modeled as a M/G/1 system with Poisson packet arrivals and an arbitrary packet length distribution. While it has been suggested that the Internet traffic is long-range dependent [27] and thus bursty, a recent work [28] shows that the network traffic can be smooth and “Poisson-like”. [29] concludes, through both simulation and analytic study, that even though the traffic exhibits bursty behavior at certain time scales, the variance-mean relation is approximately linear over larger time scales, where the traffic can be treated as if it were smooth. Our choice of the Poisson arrival model is justified because we are more concerned about the average BE performance over a large time scale for capacity planning purposes.

From the average queueing delay formula of the priority queue [30], we obtain the performance constraint for BE traffic:

$$\frac{\tilde{y}}{\tilde{\psi}_l} + \frac{\widetilde{y^2}}{2\tilde{y}} \frac{\beta_l^{ef} + \beta_l^{be}}{(\tilde{\psi}_l - \beta_l^{ef})(\tilde{\psi}_l - \beta_l^{ef} - \beta_l^{be})} \leqslant g_l \frac{\tilde{y}}{\tilde{\psi}_l} \tag{5}$$

In order to have a meaningful solution for constraint (5), $\tilde{\psi}_l > \beta_l^{ef} + \beta_l^{be}$ is required.

With some rearrangement, (5) yields $\tilde{\psi}_l \geqslant f(\beta_l^{ef})$, where

$$f(\beta_l^{ef}) = \beta_l^{ef} + \frac{\beta_l^{be}}{2} + \frac{\widetilde{y^2}(\beta_l^{ef} + \beta_l^{be})}{4(\tilde{y})^2(g_l - 1)} + \frac{1}{2}\sqrt{(2\beta_l^{ef} + \beta_l^{be} + \frac{\widetilde{y^2}(\beta_l^{ef} + \beta_l^{be})}{2(\tilde{y})^2(g_l - 1)})^2 - 4\beta_l^{be}(\beta_l^{be} + \beta_l^{ef})}$$

3. Problem Formulation

The formal problem definition is presented below.

Given: $K, M_k, L, J_k, \alpha_{km}^{ef}, \rho_{km}^{ef}, \alpha_k^{be}, \gamma_l, \delta_j^l, T, \psi_t, C_{lt}$

Variable: $x_{kmj}^{ef}, x_{kj}^{be}, u_{lt}$

Goal: $\min \sum_{l \in L} \tilde{C}_l$

Subject to:

$$\tilde{\psi}_l \geqslant f(\beta_{ef}^l) \tag{6}$$

$$x_{kmj}^{ef} = 0/1, \sum_{j \in J_k} x_{kmj}^{ef} = 1 \tag{7}$$

$$\sum_{j \in J_k} x^{be}_{kj} = 1 \tag{8}$$

$$u_{lt} = 0/1, \sum_{t \in T_l} u_{lt} = 1 \tag{9}$$

Constraint (6) ensures the performance of BE traffic. (9) imposes a discrete constraint on the link capacities. (7) ensures that all traffic from one EF O-D pair will follow one single path.

Because C_l is a non-decreasing function of β_l^{ef} and β_l^{be}, this problem can be reformulated as:

$$\min \sum_{l \in L} \tilde{C}_l \tag{10}$$

Subject to (6) (7) (8) (9) and:

$$\beta_l^{ef} \geqslant \sum_{k \in K} \sum_{m \in M_k} \alpha^{ef}_{km} \sum_{j \in J_k} x^{ef}_{kmj} \delta^l_j \tag{11}$$

$$\beta_l^{be} \geqslant \gamma_l + \sum_{k \in K} \alpha^{be}_k \sum_{j \in J_k} x^{be}_{kj} \delta^l_j \tag{12}$$

We refer to the problem defined by (10, 6, 7, 8, 9, 11, 12) as problem (P) in the rest of this paper. As can be seen from the above problem formulation, problem (P) is a non-linear integer programming problem, which is very difficult to optimize in general.

4. Solution Method

4.1. Lagrangean Relaxation

Lagrangean Relaxation is a common technique for multicommodity flow problems [31]. It has been successfully applied to the capacity planning and routing problems [32] [15] [20] [18] [19]. We describe its use for our problem in this section.

Using Lagrangean Relaxation, relax (11) and (12), and we have the Lagrangean as:

$$\begin{aligned} L(x^{ef}_{kmj}, x^{be}_{kj}, u_{lt}, \lambda^{ef}_l, \lambda^{be}_l) &= \sum_l \tilde{C}_l - \sum_l \lambda^{ef}_l (\beta^{ef}_l - \sum_{k \in K} \sum_{m \in M_k} \alpha^{ef}_{km} \sum_{j \in J_k} x^{ef}_{kmj} \delta^l_j) \\ &\quad - \sum_l \lambda^{be}_l (\beta^{be}_l - \gamma_l - \sum_{k \in K} \alpha^{be}_k \sum_{j \in J_k} x^{be}_{kj} \delta^l_j) \end{aligned} \tag{13}$$

The Lagrangean dual problem (D) is then:

$$\max_{\lambda^{ef}_l, \lambda^{be}_l \geqslant 0} h(\lambda^{ef}_l, \lambda^{be}_l) \tag{14}$$

where:

$$h(\lambda^{ef}_l, \lambda^{be}_l) = \min_{x^{ef}_{kmj}, x^{be}_{kj}, u_{lt}} L(x^{ef}_{kmj}, x^{be}_{kj}, u_{lt}, \lambda^{ef}_l, \lambda^{be}_l) \tag{15}$$

Since $\beta^{ef}_l, \beta^{be}_l$ and $x^{ef}_{kmj}, x^{be}_{kj}$ are independent variables,

$$\begin{aligned} \min L &= \min \sum_l [\tilde{C}_l - \lambda^{ef}_l \beta^{ef}_l - \lambda^{be}_l (\beta^{be}_l - \gamma_l)] + \min \sum_{k \in K} \sum_{m \in M_k} \alpha^{ef}_{km} \sum_{j \in J_k} \sum_l \lambda^{ef}_l x^{ef}_{kmj} \delta^l_j \\ &\quad + \min \sum_{k \in K} \alpha^{be}_k \sum_{j \in J_k} \sum_l \lambda^{be}_l x^{be}_{kj} \delta^l_j \end{aligned} \tag{16}$$

$$\begin{aligned} &= \sum_l \min[\tilde{C}_l - \lambda^{ef}_l \beta^{ef}_l - \lambda^{be}_l (\beta^{be}_l - \gamma_l)] + \sum_{k \in K} \sum_{m \in M_k} \min(\alpha^{ef}_{km} \sum_{j \in J_k} \sum_l \lambda^{ef}_l x^{ef}_{kmj} \delta^l_j) \\ &\quad + \sum_{k \in K} \min(\alpha^{be}_k \sum_{j \in J_k} \sum_l \lambda^{be}_l x^{be}_{kj} \delta^l_j) \end{aligned} \tag{17}$$

4.2. Solving the Subproblems

Equation (17) shows that the problem (15) can be separated into the following three subproblems:

Subproblem (i):

$$\min[\tilde{C}_l - \lambda_l^{ef}\beta_l^{ef} - \lambda_l^{be}(\beta_l^{be} - \gamma_l)] \tag{18}$$

$$\text{Subject to: } f(\beta_l^{ef}, \beta_l^{be}, \tilde{\psi}_l) \leqslant d_l^{max} \tag{19}$$

Subproblem (i) can be solved by the gradient projection method [33].

Subproblem (ii):

$$\min(\alpha_{km}^{ef} \sum_{j \in J_k} \sum_l \lambda_l^{ef} x_{kmj}^{ef} \delta_j^l) \tag{20}$$

This is simply a shortest path problem where the cost of link l is set to λ_l^{ef}. The solution is to let $x_{kmj^*}^{ef} = 1$ for j^* satisfying:

$$P(j^*) = \min_j(P(j)) \tag{21}$$

where $P(j) = \sum_l \lambda_l^{ef} x_{kmj}^{ef} \delta_j^l$

Subproblem (iii):

$$\min(\alpha_k^{be} \sum_{j \in J_k} \sum_l \lambda_l^{be} x_{kj}^{be} \delta_j^l) \tag{22}$$

Similar to Subproblem (ii), the solution is to set $x_{kj^*}^{be}$ to 1 for j^* satisfying:

$$Q(j^*) = \min_j(Q(j)) \tag{23}$$

where $Q(j) = \sum_l \lambda_l^{be} x_{kj}^{be} \delta_j^l$

4.3. Subgradient Method

The subgradient method is used to update λ_l^{ef} and λ_l^{be}. Due to space limitations, the reader is referred to the standard reference [31], or to our earlier work [23], for a detailed description of the procedure and the choices of parameters.

At each iteration, the solution of $x_{kmj^*}^{ef}$ and $x_{kj^*}^{be}$ for the primal problem (P) can be obtained from the solution of subproblem (ii) and(iii). The link capacity $\tilde{\psi}_l$ can be computed according to (6). Consequently, the primary objective function can be derived. As the iteration proceeds, we store the best solution found so far for the primal problem (P). In this way, we are always able to obtain a feasible solution. The maximum number of iterations is set to 400 in the implementation [23][31]s.

The solution of the dual problem provides a lower bound for the primal problem. Therefore, the solution quality can be assessed by the duality gap, which is the difference between the solutions of problem (P) and problem (D). Note that because the duality gap is always no smaller than the actual difference between the obtained feasible solution and the optimal solution, it is a conservative estimate of the solution quality.

5. Computational Results

In this section, we present numerical results based on experimentation. The objective of our experiment is to evaluate the solution quality and running time of the algorithm. The program is implemented in C and the computation is performed on a Pentium IV 2.4GHz PC with 512M memory, running Redhat Linux 7.2.

The network topologies are generated using the Georgia Tech Internetwork Topology Models (GT-ITM) [34]. The locations of origins and destinations are randomly selected. For each O-D pair, 10 candidate paths are calculated using Yen's K-shortest path algorithm [35].

If not specified, EF and BE demand pairs are randomly generated with a uniform distribution from 0 Mbps to 10 Mbps, while the average BE traffic load of each link is also uniformly distributed from 30 Mbps to 100 Mbps. The number of EF demands for the same O-D pair is uniformly distributed from 1 to 10. The number of candidate link types for link l is uniformly distributed from 5 to 10. The capacities of link types are set to be multiples of 45Mbps, while the costs of the link types for link l are randomly generated in such a way that the cost goes higher and the unit cost per Mbps goes down as the link capacity increases. We use average packet size $\tilde{y} = 4396$ bits and second moment of packet size $\tilde{y^2} = 22790170$ bits^2 for all the test cases. They are calculated based on a traffic trace (AIX-1014985286-1) from the NLANR Passive Measurement and Analysis project [36].

In practice, the BE delay bound factor, g_l, should be carefully chosen to reflect the actual traffic pattern, the desired BE performance, and the expected link utilization. g_l is set to 2 for all links in our experiments. The average link utilization is about 60% when g_l equals 2.

The algorithm was tested on 8 different sizes of networks, ranging from 10 nodes to 1000 nodes. Some details of the network topologies are listed in Table 1. Note that the O-D demand number shown in Table 1 includes both EF and BE demands. To obtain confidence intervals, we generate 30 different topologies for each network size, with the same number of nodes, links, and O-D pairs.

$$\text{Duality Gap} = \left| \frac{s_p - s_d}{s_d} \right| \tag{24}$$

The solution quality is represented by the duality gap, which is the percentage difference between the solution of the primal problem and the dual problem. s_p and s_d are the solutions of primal problem and dual problem respectively. A value close to zero means the solution is very close to optimal.

Table 1 shows the running time and Duality Gap (where a value of 0 means optimal quality) of various network sizes, expressed in terms of the 95% confidence intervals. In all 240 test cases, the algorithm converges without difficulty. It is easy to see from the table that the Lagrangean

Table 1
Network topology information and experimental results

Node Number	*Link Number*	*O-D Demand Number*	*Duality Gap (%)*	*Running Time (sec)*
10	25	30	(0.15, 2.24)	(3.79, 5.20)
20	50	90	(0.01, 3.96)	(15.84, 22.68)
50	125	350	(0, 2.78)	(75.40, 98.38)
100	250	1000	(0.16, 3.51)	(207.55, 241.49)
200	500	3000	(0, 2.15)	(1430.79, 1935.30)
500	1250	12000	(0, 3.02)	(8895.36, 12237.93)
700	1750	20000	(0, 2.41)	(18470.24, 23918.52)
1000	2500	40000	(0, 2.93)	(34807.64, 44630.93)

Relaxation together with the subgradient method produces reasonable results as the duality gap is bounded by no more than 4%. Note the primal problem itself is approximated when reducing the size of candidate path set for all possible path set. But according to our experimental results, more than 99% of the time, the final solution is chosen among the 5 shortest candidate paths. Therefore, 10 candidate paths are considered adequate. Having more than 10 candidate paths will have minimal impact on the solution quality, while significantly increasing the running time. Given the large number of networks being tested, we are confident that the solution should have good quality for other sizes of networks.

Because capacity planning is usually performed on the time scale of weeks to months, the running time of the algorithm is not the most critical factor. But it is still desirable to know how the running time scales up with respect to the network size. The size of the largest network evaluated in this paper is representative of a large network, and is much larger than the test cases used in most work on capacity planning. It is fair to predict that the running time of the algorithm will stay reasonable for practical sized networks.

The other observation is that over 93% of the time, demands with the same O-D use only single candidate path. It means that when designing a least cost network, traffic engineering capability enabled by the MPLS protocol does not have a significant impact on the optimal result at the stage of network planning. This is understandable, since traffic engineering is most useful when the real traffic fluctuates in an operational network. Further investigation is required to have a better understanding of the effect.

6. Conclusions and Future Directions

In this paper, we addressed the problem of link dimensioning and routing for MPLS networks supporting DiffServ EF and BE traffic. We formulate the problem as an optimization problem, where the total link cost is minimized, subject to the performance constraints of both EF and BE classes. The performance guarantee of BE traffic results in nonlinear constraints. The variable here is the non-bifurcated (single-path) routing of EF demands, multipath routing of BE demands, and the discrete link capacities.

We presented a Lagrangean Relaxation-based method to effectively decompose the original problem. A subgradient method is used to find the optimal Lagrangean multiplier. We investigated experimentally the solution quality and running time of this approach. The results from our experiments indicate that our method produces solutions that are within a few percent of the optimal solution, while the running time stays reasonable for practical sized networks.

This paper presents a preliminary investigation of the capacity planning issue for MPLS networks supporting DiffServ. The novelty of the problem presented in this paper is that it involves two traffic classes, EF and BE, which have totally different forms of performance requirements. The problem formulation and solution approaches may be applied to other traffic classes and similar network architectures.

There are opportunities to extend this work in several directions. We are working on a method where an empirical performance model may be used, and thus AF traffic classes can be incorporated into the problem. We are also investigating the adaptation of this technique to other type of networks, when there are multiple classes of service.

REFERENCES

1. S. Keshav, An Engineering Approach to Computer Networking, Addison-Wesley, 1997.
2. D. A. et al., Requirement for traffic engineering over MPLS, IETF RFC 2702 .
3. E. Rosen, A. Viswanathan, R. Callon, Multiprotocol label switching architecture, IETF RFC 3031 .
4. S. Blake, et al., An Architecture for Differentiated Service, IETF RFC 2475 .

5. K. Nichols, V. Jacobson, L. Zhang, A Two-bit Differentiated Services Architecture for the Internet, IETF RFC 2638 .
6. L. Wu, B. David, etc, Multi-Protocol label switching MPLS support of differentiated services, IETF RFC 3270 .
7. B. David, A. Charny, et Al., An Expedited Forwarding PHB, IETF RFC 3246 .
8. B. E. Carpenter, K. Nichols, Differentiated services in the internet, Proceedings of the IEEE 90 (9) (2002) 1479–1494.
9. S. Ano, T. Hasegawa, N. Decre, Experimental TCP performance evaluation on diffserv AF PHBs over ATM SBR service, Telecommunication Systems 19 (3-4) (2002) 425–441.
10. S. Chen, K. Nahrstedt, An overview of quality-of-service routing for the next generation high-speed networks: Problems and solutions, IEEE Network Magazine 12 (1998) 64–79.
11. A. Elwalid, C. Jin, S. Low, I. Widjaja, MATE: MPLS adaptive traffic engineering (2001), in: IEEE Infocom '01, Vol. 3, IEEE, Anchorage, Alaska, 2001, pp. 1300–1309.
12. M. Kodialam, T. Lakshman, Minimum interference routing with applications to MPLS traffic engineering, in: IEEE Infocom '00, Vol. 2, IEEE, Tel Aviv, Israel, 2000, pp. 884 –893.
13. I. Akyildiz, T. Anjali, L. Chen, J. de Oliveira, C. Scoglio, A. Sciuto, J. Smith, G. Uhl, A new traffic engineering manager for DiffServ/MPLS networks: Design and implementation on an IP QoS testbed, Computer Communications 26 (2003) 388–403.
14. M. Gerla, L. Kleinrock, On the topological design of distributed computer networks, IEEE Transactions on Communications COM-25 (1) (1977) 48–60.
15. B. Gavish, I. Neuman, A System for Routing and Capacity Assignment in Computer Networks, IEEE Transaction on Communications 37 (4) (1989) 360–366.
16. H. Gersht, R. Weihmayer, Joint Optimization of Data Network Design and Facility Location, IEEE Journal on Selected Areas in Communications 8 (9) (1990) 149–152.
17. T. Ng, D. Hoang, Joint Optimization of Capacity and Flow Assignment in a Packet-switched Communication Network, IEEE Transaction on Communications COM-35 (2) (1987) 202–209.
18. D. Medhi, Multi-Hour, Multi-Traffic Class Network Design for Virtual Path-based Dynamically Reconfigurable Wide-Area ATM Networks, IEEE/ACM Transactions on Networking 3 (6) (1995) 809–818.
19. D. Medhi, D. Tipper, Some Approaches to Solving a Multihour Broadband Network Capacity Design Problem with Single-path Routing, Telecommunication Systems 13 (2-4) (2000) 269–291.
20. A. Amiri, A System for the Design of Packet-Switched Communication Networks with Economic Tradeoffs, Computer Communications 21 (18) (1998) 1670–1680.
21. G. Malicsko, G. Fodor, M. Pioro, Link capacity dimensioning and path optimization for networks supporting elastic services, in: IEEE ICC 2002, Vol. 4, 2002, pp. 2304–2311.
22. L. Fratta, M. Gerla, L. Kleinrock, The Flow Deviation Method: An Approach to Store-and-Forward Communication Network Design, Networks 3 (2) (1973) 97–133.
23. K. Wu, D. S. Reeves, Capacity planning of DiffServ networks with Best-Effort and Expedited Forwarding traffic, in: To appear in IEEE 2003 International Conference on Communications, Anchorage, Alaska, 2003.
24. V.Jacobson, K.Nichols, K.Poduri, An Expedited Forwarding PHB, IETF RFC 2598 .
25. A. Charny, J.-Y. L. Boudec, Delay Bounds in Network with Aggregate Scheduling, in: First International Workshop on Quality of Future Internet Service, Berlin, Germany, 2000, pp. 1–13.
26. J. Martin, A. Nilsson, On service level agree ments for IP networks, in: IEEE Infocom 02', IEEE, New York, 2002.
27. V. Paxson, S. Floyd, Wide-area traffic: The failure of poisson modeling, in: ACM SIGCOMM '94, London, 1994.
28. R. Morris, D. Lin, Variance of aggregated web traffic, in: IEEE Infocom '00, IEEE, Tel Aviv, Israel, 2000, pp. 360–366.
29. X. Tian, J. Wu, C. Ji, A unified framework for understanding network traffic using independent wavelet models, in: IEEE Infocom '02, Vol. 1, IEEE, 2002, pp. 446–454.
30. L. Kleinrock, Queueing Systems, Volume II: Computer Application, Wiley Interscience, 1976.
31. R. Ahuja, T. Magnanti, J. Orlin, Network Flows, Theory, Algorithms and Applications, Prentice Hall, New Jersey, 1993.
32. B. Gavish, I. Neuman, Capacity and Flow Assignment in Large Computer Networks, in: IEEE Infocom '86, IEEE, 1986, pp. 275–284.
33. D. Bertsekas, R. Gallager, Data Networks, Prentice Hall, Englewood Cliffs, NJ, 1987.
34. E. Zegura, K. Calvert, M. J. Donahoo, A Quantitative Comparison of Graph-based Models for Internet Topology, IEEE/ACM Transactions on Networks 5 (6) (1997) 770–783.
35. J. Yen, Finding the K Shortest Loopless Paths in a Network, Management Science 17 (11) (1971) 712–716.
36. NLANR passive measurement and analysis project, http://pma.nlanr.net (2001).