Capacity Planning of DiffServ Networks with Best-Effort and Expedited Forwarding Traffic *

Kehang Wu and Douglas S. Reeves

Departments of Electrical and Computer Engineering and Computer Science

North Carolina State University

Abstract

For networks providing a specific level of service guarantees, capacity planning is an imperative part of network management. Accurate dimensioning is especially important in DiffServ networks, where no per-flow signaling or control exists.

In this paper, we address the problem of capacity planning for DiffServ networks with only Expedited Forwarding (EF) and best effort (BE) traffic classes. The problem is formulated as an optimization problem, where the total link cost is minimized, subject to the performance constraints of both EF and BE classes. The edge to edge EF demand pairs and the BE demands on each link are given. The variables to be determined are the non-bifurcated routing of EF traffic, and the discrete link capacities.

We show that Lagrangean Relaxation and subgradient optimization methods can be used to effectively solve the problem. Computational results show that the solution quality is verifiably good while the running time remains reasonable on practical-sized networks. This represents the first work for capacity planning of multi-class IP networks with non-linear performance constraints and discrete link capacity constraints.

Keyword: DiffServ, Capacity Planning, Lagrangean Relaxation

1 Introduction

Capacity planning is the process of designing and dimensioning networks to meet the expected demands of users. If networks need to stay ahead of the growth of user demand while still being able to provide a satisfactory service, capacity planning is indispensable. Recent years have witnessed the spectacular growth of the Internet traffic. The nature of offering only best effort (BE) service has made capacity planning a straightforward matter (Keshav 1997).

Quality of service (QoS) is the ability of a network element to have some level of assurance that its traffic and service requirements can be satisfied. With the popularizing of e-commerce and new value-added services over IP, like Voice over IP, QoS has become a must. Capacity planning will be an imperative part of IP network management to support various qualities of service.

Differentiated Services (DiffServ) is regarded as one of the key components for providing QoS in the Internet (Blake *et al*. 1998) (Nichols *et al*. 1999). The essence of DiffServ is prioritization. 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. Since there is no signaling or per-flow control, accurate dimensioning of the network is particularly important for achieving performance guarantees. To prepare for the deployment of DiffServ, 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 (David *et al*. 2002) 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 (Carpenter & Nichols 2002), three expected major initial applications of QoS 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 line. It can be easily seen that the services based on the EF PHB are ideal for all those three 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 (David *et al*. 2002).

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 (Ano *et al*. 2002). The lack of a consensus on the implementation of the AF PHB prevents us from having a definite and precise model for performance evaluation.

Operations Research techniques traditionally were applied to a system only after it had already been in place and well modeled. DiffServ is still under active research. New models and revisions of existing standards are actively being proposed. Performance issues related to the traffic aggregation and the interaction between multiple classes are not fully understood. Those factors impose challenges on the formulation of capacity planning problems. Given the urgency for the need of capacity planning in DiffServ network on the one hand, and the difficulty and complexity of the problem on the other hand, we have to resort to simplifications and approximations.

In this paper, we address the problem of capacity planning for DiffServ networks with only EF and best effort traffic classes. We will study the AF PHB and possibly other PHBs in the future when the situation is improved. The problem is formulated as an optimization problem, where we jointly select the route for each EF user demand pair, and assign a discrete capacity value for each link to minimize the total link cost, subject to the performance constraints of both EF and BE classes. The non-bifurcated routing model is used for the EF class, where the traffic from a single EF demand pair will follow the same path between the origin and the destination. 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 the DiffServ networks 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 issues of DiffServ networks, there is a lot of work dealing with design and planning issues of communication networks, especially for connection-oriented networks [see (Medhi & Tipper 2000) and references therein]. Traditionally, network design has been focused on optimizing either the network cost or performance by tuning the network topology, link capacity, and routing strategies; see, for example, (Gerla & Kleinrock 1977) (Gavish & Neuman 1986) (Gerla *et al*. 1989) (O.Gerstel & S.Zaks 1996) (Mitra *et al*. 1996) (Medhi 1995) (Girish *et al*. 2000). More recently, network survivability has attracted attention as well (M. Grotschel & Stoer 1995) (Crochat & Boudec 1998) (Modiano & Narula-Tam 2001).

The literature focusing on the routing problem, where link capacities are given, is abundant (Gavish & Hantler 1983) (Tcha & Maruyama 1985) (Pirkul & Amiri 1994). But because the routing and link dimensioning problems are closely related to each other, it is not generally appropriate to separate them. Papers where the routing and capacity assignment problems are treated simultaneously include (Gerla & Kleinrock 1977) (Gavish & Neuman 1989) (Gersht & Weihmayer 1990)(Ng & Hoang 1987) (Medhi 1995) (Medhi & Tipper 2000) (Amiri 1998). Gerla and Kleinrock (Gerla & Kleinrock 1977) presented heuristic methods based on the flow deviation algorithm (Fratta *et al*. 1973). Gavish and Neuman (Gavish & Neuman 1989) formulated the problem as a non-linear integer programming problem, and proposed a Lagrangean relaxation based approach. The networks studied in (Gerla & Kleinrock 1977)(Gavish & Neuman 1989) only include one traffic class, though. Medhi and Tipper (Medhi & Tipper 2000) proposed four approaches for reconfigurable ATM networks, based on the Virtual Path concept. Even though ATM includes multiple traffic classes, Medhi proposed a model that assumes the deterministic multiplexing of different virtual paths, which results in linear performance constraints. The novel aspect of our DiffServ network 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 Notations and Models

The following notation will be used throughout the paper:

K	set of EF traffic demand node pairs
L	set of links in the network
J	set of possible candidate paths for EF demand pairs
x_{kj}	path routing variable; 1 if demand pair $k \in K$ uses path $j \in J$, 0 otherwise. $\forall k \in K$, $\sum_j x_{kj} = 1$
δ_j^l	link-path indicator; 1 if path j uses link $l \in L$, 0 otherwise
α_k	average arrival rate of an EF traffic demand of node pair $k \in K$
ρ_k	requested bandwidth of an EF traffic demand pair $k \in K$
β_{ef}^l	average arrival rate of total EF traffic demand on link $l \in L$
η_l	total requested bandwidth of EF demand on link $l \in L$
β_{be}^l	average BE traffic load on link $l \in L$

d_l average delay experienced by BE traffic on link $l \in L$

u_l number of units of capacity needed on link $l \in L$. u_l is integer

γ size of a unit of capacity

C_l cost of a unit of capacity on link $l \in L$

$\tilde{y}$, $\widetilde{y^2}$ the first and second moment of packet size, (units: bits & bits^2)

In the DiffServ capacity planning problem, we are given a network, which is defined by a set of links L, candidate paths J, and link-path indicators $\{\delta_j^l\}$. We are also given the EF and BE user demands. Since the purpose of EF PHB is to build a leased line type of service, the projected EF user demands are defined in the form of *origin-destination (O-D) pairs* (i.e., the origin and destination of each traffic source is given)[1]. For BE user demands, the traffic volume is much larger and much more variable. It is therefore more difficult and less likely to have a clear characterization of BE traffic for each possible origin and destination. Accordingly, in this paper we take the approach of defining BE user demands for each *link*, which is far easier to measure and predict in practice.

For an EF demand pair k, we differentiate between the average arrival rate, α_k, and the requested bandwidth, ρ_k. ρ_k is usually a value between the average arrival rate and the peak rate. BE demand on link l is characterized by β_{be}^l, the average arrival rate. User demand estimation is out of the scope of this paper. We would like to point out, however, that even if demand estimation is inaccurate, a high-quality solution method for the capacity planning problem, given the estimated demands, is still very desirable.

The goal of the capacity planning problem is to find the minimum cost network that satisfies the projected user demands of both EF and BE user classes. The variables here are the path routing variable, $\{x_{kj}\}$, which shows whether O-D pair k will use the candidate path j, and the discrete link capacity, $\{u_l\}$. The non-bifurcated routing model is used for EF class, where the traffic from a single EF demand pair will follow the same path between the origin and the destination. According to (David *et al*. 2002), the non-bifurcated routing assumption is necessary to ensure no packets are delivered out of order. We assume here that the cost is a linear function of the link capacity.

Let $\Gamma()$ be the EF demand multiplexing function. An expression for the EF traffic demand on link l is

$$\begin{aligned} \eta_l &= \Gamma(\{\rho_k\}, \{\delta_j^l\}, \{x_{kj}\}) \\ &\leqslant \sum_{k \in K} \rho_k \sum_{j \in J} \delta_j^l x_{kj}. \end{aligned} \tag{1}$$

The inequality becomes equality only when there is no multiplexing gain.

There are many discussions about the original EF PHB (V.Jacobson *et al*. 1999) concerning the limits on the EF utilization. Charny reported in (Charny & Boudec 2000) 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 (David *et al*. 2002), 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. Therefore

$$\eta_l << u_l\gamma,\ \forall l \in L. \tag{2}$$

Assumption (2) ensures that the bandwidth requirement for the EF class, as stated in the EF PHB standard (David *et al*. 2002), is complied with. The EF PHB standard also mentions that the delay will be bounded once sufficient bandwidth is given. We do not specifically consider the delay constraint for the EF traffic in this paper.

How to specify the performance of BE traffic in the service level agreement (SLA) is still an active research topic. (Martin & Nilsson 2002) 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. As the BE demands are given link by link, we evaluate the performance of BE traffic on a link basis as well. The value $\frac{\tilde{y}}{u_l\gamma}$ stands for the average transmission delay of packets. We use $\frac{\tilde{y}}{u_l\gamma}$ as the basis for the delay bound. Let $d_{lmax} = g_l\frac{\tilde{y}}{u_l\gamma}$, where g_l is a parameter defined by the network designer. The value of g_l should of course be greater than 1. A g_l value of 1 means there is only transmission delay but no queueing delay, which implies infinite link capacity with the assumption of Poisson arrival. We assume that the performance of BE traffic is satisfactory if:

$$d_l \leqslant d_{lmax}. \tag{3}$$

We further assume a strict non-preemptive priority queue is used in the routers to support the EF class. A priority queue is considered as the canonical example of an implementation of EF

(David *et al*. 2002). All EF packets share a single FIFO queue which has the highest priority. All other packets are sent to the second priority queue. Note that the priority queue gives the most preferable treatment to the EF packets possible among all schedulers, and it will also have the greatest negative impact on the performance of BE traffic. Given the same performance constraint of BE traffic, the choice of priority queue implies a larger link capacity requirement.

Also assume that 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 (Paxson & Floyd 1994) and thus bursty, a recent work (Morris & Lin 2000) shows that the network traffic can be smooth and "Poisson-like". (Tian *et al*. 2002) 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.

According to the average queueing delay formula of the priority queue (Kleinrock 1976), we have the performance of BE traffic:

$$d_l = \frac{\tilde{y}}{u_l\gamma} + \frac{\widetilde{y^2}}{2\tilde{y}} \frac{\beta_{ef}^l + \beta_{be}^l}{(u_l\gamma - \beta_{ef}^l)(u_l\gamma - \beta_{ef}^l - \beta_{be}^l)} \tag{4}$$

$$\text{where} \quad \beta_{ef}^l = \sum_{k\in K} \alpha_k \sum_{j\in J} \delta_j^l x_{kj}. \tag{5}$$

Note that to have a meaningful solution for constraint (4),

$$u_l\gamma > \beta_{ef}^l + \beta_{be}^l. \tag{6}$$

With some simplification, (4) and (3) yields:

$$u_l\gamma \geqslant f(\beta_{ef}^l) \tag{7}$$

where

$$\begin{aligned} f(\beta_{ef}^l) &= \beta_{ef}^l + \frac{\beta_{be}^l}{2} + \frac{\widetilde{y^2}(\beta_{ef}^l + \beta_{be}^l)}{4(\tilde{y})^2(g_l-1)} \\ &+\frac{1}{2}\sqrt{(2\beta_{ef}^l + \beta_{be}^l + \frac{\widetilde{y^2}(\beta_{ef}^l + \beta_{be}^l)}{2(\tilde{y})^2(g_l-1)})^2 - 4\beta_{be}^l(\beta_{be}^l + \beta_{ef}^l)}. \end{aligned} \tag{8}$$

Let $\theta = \frac{\widetilde{y^2}}{2(\tilde{y})^2(g_l-1)}$. We have:

$$\frac{df(\beta_{ef}^l)}{d\beta_{ef}^l} = 1 + \frac{\theta}{2} + \frac{1}{4}\frac{2(\theta^2+4\theta)\beta_{ef}^l + (2\theta^2+6\theta)\beta_{be}^l}{\sqrt{(\theta^2+4\theta)(\beta_{ef}^l)^2 + (2\theta^2+6\theta)\beta_{be}^l\beta_{ef}^l + (\theta+1)^2(\beta_{be}^l)^2}} > 0 \tag{9}$$

and

$$\frac{d^2f(\beta_{ef}^l)}{d(\beta_{ef}^l)^2} = \frac{2\theta(\beta_{be}^l)^2}{[(\theta^2+4\theta)(\beta_{ef}^l)^2 + (2\theta^2+6\theta)\beta_{be}^l\beta_{ef}^l + (\theta+1)^2(\beta_{be}^l)^2]^{\frac{3}{2}}} > 0. \tag{10}$$

Equation (9) shows that $f(\beta_{ef}^l)$ is an increasing function, while (10) means that $f(\beta_{ef}^l)$ is a convex function.

3 Problem Formulation

The formal problem definition is presented below. Our goal is to minimize the total costs.

$$\min(\sum_{l\in L} c_l u_l \gamma) \tag{11}$$

subject to:

$$u_l\gamma \geqslant f(\beta_{ef}^l),\ \forall l \in L \tag{12}$$

$$\text{where } \beta_{ef}^{l} = \sum_{k \in K} \alpha_k \sum_{j \in J} \delta_j^l x_{kj} \tag{13}$$

$$u_l \geqslant 0 \text{ and integer}, \ \forall l \in L \tag{14}$$

$$\sum_{j \in J} x_{kj} = 1, \ \forall k \in K \tag{15}$$

$$x_{kj} = 0/1, \ \forall j \in J, \ k \in K. \tag{16}$$

Constraint (12) ensures the performance of EF and BE traffic. (14) imposes a discrete constraint on the link capacities. (15) and (16) ensure that all traffic from one EF O-D pair will follow one single path.

The fact that the function $f(\beta_{ef}^l)$ is increasing in β_{ef}^l enables us to replace the constraint (13) with (17):

$$\beta_{ef}^{l} \geqslant \sum_{k \in K} \alpha_k \sum_{j \in J} \delta_j^l x_{kj} \tag{17}$$

The resulting equivalent problem is:

$$\min(\sum_{l \in L} c_l u_l \gamma) \tag{18}$$

subject to:

$$(12), \ (17), \ (14), \ (15), \text{ and } (16).$$

We refer to the problem defined by (18,17,12,14,15,16) 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 difficult in general. Other than the non-linear constraint (12), problem (P) is identical to the multiple choice multiconstrained knapsack problem (Sinha & Zoltners 1979), which is known to be NP-complete. The non-linear constraint (12) leads to a more difficult problem.

4 Lagrangean Relaxation

Lagrangean Relaxation is a common technique for multicommodity flow problems (Ahuja *et al*. 1993). It has been successfully applied to the capacity planning and routing problems (Gavish & Neuman 1986) (Gavish & Neuman 1989) (Amiri 1998) (Medhi 1995) (Medhi & Tipper 2000). We describe its use for our problem in this section.

Note that there are three constraints that make the problem (P) difficult:

- non-linear BE performance constraint (12),
- non-bifurcated routing constraint (16),
- and the discrete link capacity constraint (14).

Relaxing either one of those above constraints still results in either an integer programming or a nonlinear integer programming problem, which makes most heuristics inefficient. The novel approach shown in this section relaxes the equation (17), which in turn produces simple separable subproblems. This approach is also shown to yield good solution quality, according to a large number of test cases (section 5).

Let $\lambda = (\lambda_l)$ be the dual multiplier associated with the constraint (17). Then the Lagrangean can be expressed as

$$\begin{aligned} L(x,u,\lambda) &= \sum_{l\in L} c_l u_l \gamma + \sum_{l\in L} \lambda_l (\beta_{ef}^l - \sum_{k\in K} \alpha_k \sum_{j\in J} \delta_j^l x_{kj}) \\ &= \sum_{l\in L} (c_l u_l \gamma + \lambda_l \beta_{ef}^l) + \sum_{k\in K} \sum_{l\in L} -\alpha_k \lambda_l \sum_{j\in J} \delta_j^l x_{kj}. \end{aligned} \tag{19}$$

The Lagrangean dual problem (D) is then:

$$\max_{\lambda \leqslant 0} h(\lambda) \tag{20}$$

$$\text{where } h(\lambda) = \min_{x,u} L(x,u,\lambda) \tag{21}$$

subject to the constraint:

$$u_l \gamma \geqslant f(\beta_{ef}^l),\ \forall l \in L \tag{22}$$

$$u_l \geqslant 0 \text{ and integer},\ \forall l \in L \tag{23}$$

$$\sum_{j \in J} x_{kj} = 1,\ \forall k \in K \tag{24}$$

$$x_{kj} = 0/1,\ \forall j \in J,\ k \in K. \tag{25}$$

4.1 Solving the Lagrangean Dual Problem (D)

For a given λ, the Lagrangean is separable in x and u. (21) is reduced to solving two independent subproblems,

$$\min_{x,u} L(x,u,\lambda) = \min_{u} L_1(u,\lambda) + \min_{x} L_2(x,\lambda). \tag{26}$$

Subproblem (D1):

$$\min_{u} L_1(u,\lambda) = \min_{u}\{\sum_{l \in L}(c_l u_l \gamma + \lambda_l \beta_{ef}^l)\} \tag{27}$$

subject to the constraint:

$$u_l \gamma \geqslant f(\beta_{ef}^l),\ \forall l \in L \tag{28}$$

$$u_l \geqslant 0 \text{ and integer},\ \forall l \in L. \tag{29}$$

Subproblem (D2):

$$\min_x L_1(x,\lambda) = \min_x (\sum_{k\in K}\sum_{l\in L} -\alpha_k \lambda_l \sum_{j\in J}\delta_j^l x_{kj}) \tag{30}$$

subject to the constraint:

$$\sum_{j\in J} x_{kj} = 1,\ \forall k \in K \tag{31}$$

$$x_{kj} = 0/1,\ \forall j \in J,\ k \in K. \tag{32}$$

Subproblem (D1) can be separated into L problems, one for each link. The problem for each link l can be expressed as:

$$\min_u (c_l u_l \gamma + \lambda_l \beta_{ef}^l) \tag{33}$$

subject to the constraint:

$$u_l \gamma \geqslant f(\beta_{ef}^l) \tag{34}$$

$$u_l \geqslant 0 \text{ and integer.} \tag{35}$$

Because $f(\beta_{ef}^l)$ is increasing in β_{ef}^l and $\lambda_l \leqslant 0$, the problem (33) can be simply rewritten as:

$$\min_u (c_l u_l \gamma + \lambda_l f^{-1}(u_l \gamma)),\ u_l \geqslant 0 \text{ and integer.} \tag{36}$$

The value of u_l^*, which satisfies the equation (37) below, can be obtained numerically through any one-dimensional optimization method, such as Newton's method (Avriel 1976).

$$\frac{dz(u_l)}{du_l} = 0, u_l \geqslant 0 \tag{37}$$

where $z(u_l) = c_l u_l \gamma + \lambda_l f^{-1}(u_l \gamma)$.

Let u_{l1}^* and u_{l2}^* be the two non-negative integers closest to u_l^*. We know that the solution of u_l to the problem (36) is either u_{l1}^* or u_{l2}^*. Among them, the one that minimizes (36) is picked as the final solution.

Subproblem (D2) can be separated into K subproblems, one for each O-D pair. The problem for each k is:

$$\min_x (\sum_{l \in L} -\alpha_k \lambda_l \sum_{j \in J} \delta_j^l x_{kj}) \tag{38}$$

subject to the constraint:

$$\sum_{j \in J} x_{kj} = 1 \tag{39}$$

$$x_{kj} = 0/1, \ \forall j \in J. \tag{40}$$

The solution of subproblem (D2) is then easily obtained by setting $x_{kj^*} = 1$ for j^* satisfying:

$$P(j^*) = \min_j (P(j)) \tag{41}$$

where $P(j) = \sum_{l \in L} -\alpha_k \lambda_l \delta_j^l x_{kj}$.

4.2 Subgradient Optimization Procedures

We use the subgradient method to search for optimal multipliers λ. (For the detailed description of the subgradient method and choice of parameters, please refer to (Ahuja *et al*. 1993).

For a given initial λ, once we solve the problem (D), a dual subgradient is computed as follows:

$$\omega_l = \beta_{ef}^l - \sum_{k \in K} \alpha_k \sum_{j \in J} \delta_j^l x_{kj}, \forall l \in L. \tag{42}$$

The subsequent values of the Lagrangean multipliers are updated:

$$\lambda_l \leftarrow \min(0, \lambda_l + t\omega_l), \forall l \in L \tag{43}$$

where the step size, t, is defined by:

$$t = \phi \frac{h^*(u) - h(u)}{||\omega||^2} \tag{44}$$

where $h^*(u)$ is the value of the best feasible solution found so far, and ϕ is a scalar between 0 and 2. ϕ is set to 2 initially in our study and is halved if the solution does not improve in 10 iterations.

At each iteration, the solution of $\{x_l\}$ for the primal problem (P) can be generated from the solution of subproblem (D2). The value of $\{u_l\}$ can be computed according to (12). Consequently, the primary objective function can be obtained. 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. We set the maximum iterations to 400 in the implementation.

The benefit of Lagrangean optimization procedures is that 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 computational work 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) (Zegura *et al*. 1997). Link cost is set to be proportional to its length.

The locations of originations and destinations are randomly selected. For each O-D pair, 10 candidate paths are calculated using Yen's K-shortest path algorithm (Yen 1971). The length of each link is set to C_l, the cost of a unit of capacity. 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 increas the running time.

If not specified, EF demands 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 link unit γ is set to be 45 Mbps. We use $\tilde{y} = 4396$ (bits) and $\tilde{y^2} = 22790170$ (bits^2) for all the test cases. These are calculated based on a traffic trace (AIX-1014985286-1) from the NLANR Passive Measurement and Analysis project (NLANR 2001).

The BE delay bound factor, g_l, is set to 2 for all links in our experiments. In practice, g_l should be carefully chosen to reflect the desired BE performance and link utilization. Figure (1) shows the value of $f(\beta_{ef}^l)$ with respect to β_{ef}^l, for g_l=1.5, 2, 4, and 8. As can be seen from the figure, the link utilization varies significantly as g_l changes.

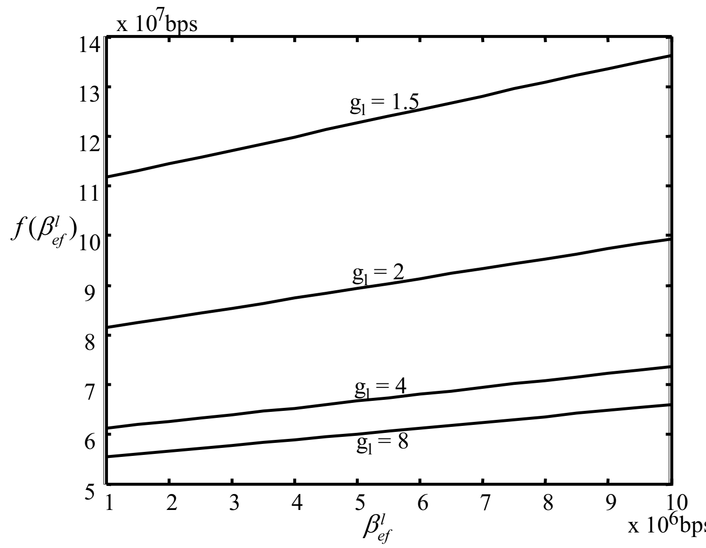

Figure 1: $f(\beta^l_{ef})$ vs. β^l_{ef}, ($\beta^l_{be} = 5 \times 10^7$ bps)

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. To obtain confidence intervals, we generate 30 different topologies for each network size, with the same number of nodes, links, and O-D pairs.

The duality gap is used to evaluate the solution quality. The duality gap is represented by the percentage difference between the solution of the primal problem and the dual problem. It is an upper bound on the actual gap between solution of the primal problem and the optimum solution.

$$\text{Duality Gap} = \left| \frac{s_p - s_d}{s_d} \right| \tag{45}$$

where s_p and s_d are the solutions of primal problem and dual problem respectively.

Table 1: Network topology information and experimental results

Number of Nodes	*Number of Links*	*Number of O-D Pairs*	*Duality Gap (%)*	*Running Time (sec)*
10	25	30	(0.25-5.24)	(0.99-1.16)
20	50	90	(0-5.04)	(4.79-5.60)
50	125	350	(0.13-4.13)	(23.88-27.59)
100	250	1000	(0.16-5.49)	(57.08-68.14)
200	500	3000	(0-4.51)	(491.90-562.88)
500	1250	12000	(0-4.12)	(5000.45-6360.11)
700	1750	20000	(0-3.28)	(10367.61-13217.79)
1000	2500	40000	(0-3.91)	(27905.69-35644.40)

Table 1 shows the running time and solution quality of various network sizes with 95% confidence intervals. In all 240 test cases, the algorithm converges without difficulty. It is easy to see from the table that the Lagrangean Relaxation together with the subgradient method produces reasonable results as the duality gap is bounded by no more than 6%. 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. As can be seen from the Figure (2), while the running time goes up more quickly than the number of nodes, it increases approximately linearly with respect to the number of O-D pairs in our test cases. This is understandable; since $K >> L$, when the network size increases, the dominant portion of the running time is spent on

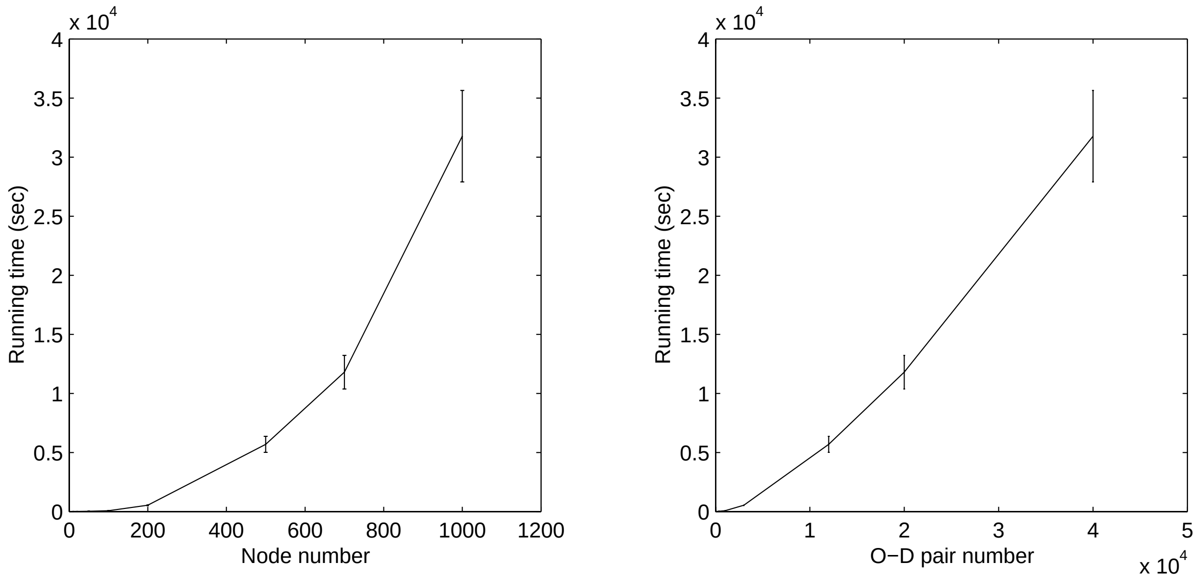

Figure 2: Running Time

the subproblem (D2), which in turn has K subproblems. The running time of each one of those K subproblems is insensitive to the size of the network, due to the fixed number of candidate paths and the reuse of the link cost throughout all iterations. 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.

6 Conclusions and Future Directions

Services based on the EF PHB are likely to be deployed in the near future. In this paper, we addressed the problem of link dimensioning and routing for DiffServ networks with only 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 routing of EF demands and the discrete link capacities.

We presented a Lagrangean Relaxation-based method to effective 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, and that the running time scales linearly with the number of EF O-D demand pairs.

This paper presents a preliminary investigation of the capacity planning issue for DiffServ networks. 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 is opportunity to extend this work in several directions. We will incorporate AF traffic into the model when there is a consensus on the correct implementation of AF service. We are also investigating the adaptation of this technique to MPLS traffic engineering, when there are multiple classes of service possible.

Acknowledgement and Notes

* 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.

1. Note that what we are studying in this paper is the core network. Therefore, when we speak of O-D pairs, we are referring to the traffic between pairs of edge routers, not between individual hosts.

References

Ahuja, R., Magnanti, T. & Orlin, J. (1993), *Network Flows, Theory, Algorithms and Applications*, Prentice Hall, New Jersey.

Amiri, A. (1998), 'A System for the Design of Packet-Switched Communication Networks with Economic Tradeoffs', *Computer Communications* **21**(18), 1670–1680.

Ano, S., Hasegawa, T. & Decre, N. (2002), 'Experimental TCP performance evaluation on diffserv AF PHBs over ATM SBR service', *Telecommunication Systems* **19**(3-4), 425–441.

Avriel, M. (1976), *Nonlinear Programming*, Prentice Hall.

Blake, S. & *et al*. (1998), 'An Architecture for Differentiated Service, IETF RFC 2475'.

Carpenter, B. E. & Nichols, K. (2002), 'Differentiated services in the internet', *Proceedings of the IEEE* **90**(9), 1479–1494.

Charny, A. & Boudec, J.-Y. L. (2000), 'Delay Bounds in Network with Aggregate Scheduling', in: *Proc. of the First International Workshop on Quality of Future Internet Service*, Berlin, Germany, pp. 1–13.

Crochat, O. & Boudec, J. L. (1998), 'Design Protection for WDM Optical Networks', *IEEE Journal on Selected Areas in Communications* **16**(7), 1158–1165.

David, B., Charny, A. & *et al*. (2002), 'An Expedited Forwarding PHB, IETF RFC 3246'.

Fratta, L., Gerla, M. & Kleinrock, L. (1973), 'The Flow Deviation Method: An Approach to Store-and-Forward Communication Network Design', *Networks* **3**(2), 97–133.

Gavish, B. & Hantler, S. (1983), 'An Algorithm for Optimal Route Selection in SNA Networks', *IEEE Transactions on Communications* **COM-31**(10), 1154–1161.

Gavish, B. & Neuman, I. (1986), 'Capacity and Flow Assignment in Large Computer Networks', in: *Proc. of IEEE Infocom '86*, IEEE, pp. 275–284.

Gavish, B. & Neuman, I. (1989), 'A System for Routing and Capacity Assignment in Computer Networks', *IEEE Transaction on Communications* **37**(4), 360–366.

Gerla, M. & Kleinrock, L. (1977), 'On the topological design of distributed computer networks', *IEEE Transactions on Communications* **COM-25**(1), 48–60.

Gerla, M., Monteiro, J. & Pazos, R. (1989), 'Topology Design and Bandwidth Allocation in ATM Nets', *IEEE Journal on Selected Areas in Communications* **7**(8), 1253–1262.

Gersht, H. & Weihmayer, R. (1990), 'Joint Optimization of Data Network Design and Facility Location', *IEEE Journal on Selected Areas in Communications* **8**(9), 149–152.

Girish, M., Zhou, B. & Hu, J. (2000), 'Formulation of the Traffic Engineering Problems in MPLS Based IP Networks', in: *Proc. of the Fifth IEEE Symposium on Computers and Communications*, pp. 214–219.

Keshav, S. (1997), *An Engineering Approach to Computer Networking*, Addison-Wesley.

Kleinrock, L. (1976), *Queueing Systems, Volume II: Computer Application*, Wiley Interscience.

M. Grotschel, C. M. & Stoer, M. (1995), *Network Models, Handbooks in Operations Research and Management Science*, Vol. 7, North-Holland, Amsterdam, chapter 10, pp. 617–672.

Martin, J. & Nilsson, A. (2002), ‘On service level agree ments for IP networks’, in: *Proc. of IEEE Infocom ’02*, IEEE, New York.

Medhi, D. (1995), ‘Multi-Hour, Multi-Traffic Class Network Design for Virtual Path-based Dynamically Reconfigurable Wide-Area ATM Networks’, *IEEE/ACM Transactions on Networking* **3**(6), 809–818.

Medhi, D. & Tipper, D. (2000), ‘Some Approaches to Solving a Multihour Broadband Network Capacity Design Problem with Single-path Routing’, *Telecommunication Systems* **13**(2-4), 269–291.

Mitra, D., Morrison, J. & Ramakrishnan, K. (1996), ‘ATM Network Design and Optimization: A Multi-rate Loss Network Framework’, *IEEE/ACM Transactions on Networking* **4**(4), 531–543.

Modiano, E. & Narula-Tam, A. (2001), ‘Survivable Routing of Logical Topologies in WDM Networks’, in: *Proc. of IEEE Infocom ’01*, Vol. 1, pp. 348–357.

Morris, R. & Lin, D. (2000), ‘Variance of aggregated web traffic’, in: *Proc. of IEEE Infocom ’00*, IEEE, Tel Aviv, Israel, pp. 360–366.

Ng, T. & Hoang, D. (1987), 'Joint Optimization of Capacity and Flow Assignment in a Packet-switched Communication Network', *IEEE Transaction on Communications* **COM-35**(2), 202–209.

Nichols, K., Jacobson, V. & Zhang, L. (1999), 'A Two-bit Differentiated Services Architecture for the Internet, IETF RFC 2638'.

NLANR (2001), 'NLANR passive measurement and analysis project, http://pma.nlanr.net'.

O.Gerstel, I. & S.Zaks (1996), 'The Layout of Virtual Paths in ATM Networks', *IEEE/ACM Transactions on Networking* **4**(6), 873–884.

Paxson, V. & Floyd, S. (1994), 'Wide-area traffic: The failure of poisson modeling', in: *Proc. of ACM SIGCOMM '94*, London.

Pirkul, H. & Amiri, A. (1994), 'Routing in Packet-switched Communication Networks', *Computer Communications* **17**(5), 307–316.

Sinha, P. & Zoltners, A. (1979), 'The Multiple Choice Knapsack Problem', *Operation Research* **27**(3), 503–515.

Tcha, D. & Maruyama, K. (1985), 'On the Selection of Primary Paths for a Communication Network', *Computer Networks and ISDN Systems* **9**(4), 257–265.

Tian, X., Wu, J. & Ji, C. (2002), 'A unified framework for understanding network traffic using independent wavelet models', in: *Proc. of IEEE Infocom '02*, IEEE, New York.

V.Jacobson, K.Nichols & K.Poduri (1999), 'An Expedited Forwarding PHB, IETF RFC 2598'.

Yen, J. (1971), ‘Finding the K Shortest Loopless Paths in a Network’, *Management Science* **17**(11), 712–716.

Zegura, E., Calvert, K. & Donahoo, M. J. (1997), ‘A Quantitative Comparison of Graph-based Models for Internet Topology’, *IEEE/ACM Transactions on Networks* **5**(6), 770–783.