Allocation of End-to-end Delay Objectives for Networks Supporting SMDS

Frank Y.S. Lin

Bell Communications Research
444 Hoes Lane
Piscataway, New Jersey 08854
(908)699-4220
lin1@cc.bellcore.com

Abstract

In this paper, the end-to-end delay objective allocation problem for networks supporting Switched Multi-megabit Data Service (SMDS) is considered. Traditionally, for engineering tractability, end-to-end service objectives are allocated to network elements in such a way that, if the allocated service objective for each network element is satisfied then the end-to-end service objectives are satisfied. Such an objective allocation strategy is referred to as a *feasible* objective allocation strategy.

For networks supporting SMDS, the delay objectives state that 95% of the packets delivered from the origin Subscriber Network Interface (SNI) to the destination SNI should be within a given time threshold. From network monitoring point of view, this percentile type of delay objectives makes it complicated to compute feasible allocation strategies so that network elements instead of each origin-destination pair should be monitored. From network planning point of view, these end-to-end percentile-type delay objectives usually impose an excessively large number of nonconvex and complicated constraints (delay distributions are convolved, assuming delays are mutually independent).

The emphasis of this paper is three fold: (i) to propose an efficient and generic approach to replacing the set of end-to-end percentile-type delay constraints by a simpler set of network element utilization constraints (small size and convex), (ii) to investigate how this approach could be adopted in conjunction with a number of possible allocation schemes and (iii) to compare the relative effectiveness (in terms of the utilization thresholds determined) using $M/M/1$ and $M/D/1$ queueing models. The significance of this work is to provide a general and effective way to calculate engineering thresholds on the network element utilization factors when percentile types of end-to-end delay objectives are considered. This work lays a foundation to help the system planners/administrators monitor, service, expand and plan networks supporting SMDS.

1. INTRODUCTION

Switched Multi-megabit Data Service (SMDS) is a high-speed, connectionless, public, packet switching service that will extend Local Area Network (LAN)-like performance beyond the subscriber's premises, across a metropolitan or wide area. SMDS has been recognized as the first step towards the Broadband Integrated Services Digital Network (BISDN), and is commercially available. To support the planning and engineering functions of networks supporting SMDS (referred to as SMDS networks), it is important for the system planners/administrators to be able to efficiently and effectively consider the service performance objectives.

Since end-to-end performance is users' direct perception about the service quality, like many new services, e.g. Frame Relay Service (FRS), Asynchronous Transfer Mode (ATM) and Advanced Intelligent Network (AIN) services, performance objectives are specified on an end-to-end basis for the SMDS service. When network planning and engineering functions are performed, these end-to-end service objectives are typically difficult to handle principally due to the excessively large number of constraints they impose. (When the linkset sizing problem for SMDS network is considered[1], the number of original end-to-end delay constraints equals the total number of simple paths in the network.)

For networks supporting SMDS, the delay objectives state that 95% of the packets delivered from the origin Subscriber Network Interface (SNI) to the destination SNI should be within a given time threshold[2]. The percentile nature of the end-to-end delay performance objectives for SMDS networks makes it more difficult to handle the constraints, since each of the constraints is highly nonlinear (involving convolution when delays are assumed to be mutually independent) and nonconvex. The nonconvex property can be illustrated by the following example. Consider a network with only one network element. For a given time threshold, the probability that an arbitrary packet will not be delivered cross the network element within the time threshold, referred to as the *overdue* probability, shall increase with the utilization and approaches 1 asymptotically. The overdue probability is thus clearly not a convex function of the utilization.

To make the problem tractable, there is a need to circumvent the above two difficulties -- volume and nonconvexity of end-to-end delay constraints. One possible, and traditional, approach is to allocate the end-to-end performance objectives to each network element in such a way that, if the performance objective allocated to each network element is met then the end-to-end performance objectives are satisfied[3]. In addition, for monitoring purposes, the derived (surrogate) service objective for each network element is further converted into a threshold on the network element utilization.

The allocation approach has the following significance. First, it reduces the number of constraints greatly (from the number of simple paths to the number of links in the network). Second, this approach decouples the decision variables (utilization) in the constraint set (each constraint involves only one decision variable). Third, this approach makes the delay constraints linear. Fourth, if properly handled as will be discussed in the next section, the new set of constraints will define a feasible region which is a subset of the feasible region of the original constraints. In other words, any feasible solution to the surrogate problem is a feasible solution to the original problem.

This allocation approach is also commonly adopted when the network monitoring problem is considered. For a feasible allocation strategy, when the utilization of each network element is no greater than the allocated threshold, it can be sure that the end-to-end delay objectives be satisfied. Otherwise, the network planner/administrator may need to calculate/measure the end-to-end delay for every origin-destination (O-D) pair at the same time to evaluate the delay objectives.

To satisfy the feasibility property (a feasible solution to the surrogate problem is a feasible solution to the original problem) mentioned earlier, in this paper we propose a general approach to replacing the end-to-end percentile-type delay objectives by a set of link constraints in such a way that the end-to-end percentile-type delay objectives are satisfied under any routing assignment as long as the link constraints are satisfied. This approach is suitable for heterogeneous networks. How to implement this approach in a network planning/engineering problem to achieve the best result is also discussed.

Following the general approach, a number of possible allocation schemes are discussed. They include

1. longest delay control
2. complete decomposition
3. GI/G/1 bounding scheme
4. Markov inequality
5. Chebyshev inequality and
6. normal approximation.

Two allocation schemes for homogeneous networks are also discussed. They are

1. Convolution scheme and
2. Chernoff bounding scheme.

Each of the above schemes requires different input, e.g. interarrival time distribution, service time distribution, delay distribution, packet blocking probability and so on, and different degree of computational complexity. Those differences are discussed. In addition, the relative effectiveness in terms of the utilization determined in a homogeneous network is compared using $M/M/1$ and $M/D/1$ queueing models.

The significance of this work is to provide a general and effective way to calculate engineering thresholds on the network element utilization factors when percentile types of end-to-end delay objectives are considered. This work lays a foundation to help the system planners/administrators monitor, service, expand and plan networks supporting SMDS.

The remainder of this paper is organized as follows. In Section 2, a general approach to allocating the end-to-end percentile-type delay objectives to network elements in a heterogeneous environment is proposed. In Section 3, 6 schemes following the general approach proposed in Section 2 are discussed. In Section 4, 2 allocation schemes suitable in a homogeneous environment are discussed. In Section 5, two case studies are given to compare the schemes using $M/M/1$ and $M/D/1$ queueing models. Section 6 summarizes this paper.

2. GENERAL APPROACH

In this section, we propose a general approach to replacing the end-to-end percentile-type delay objective constraints by a simpler set of link constraints in such a way that, under any routing assignment, if the link constraints are satisfied then the end-to-end percentile-type delay objectives are satisfied. This approach is suitable for heterogeneous networks, and thus for homogeneous networks as special cases. In addition, this approach is applicable when end-to-end mean delay or loss rate constraints are considered. How to implement this approach in a network planning/engineering problem to achieve the best result is also discussed.

An SMDS backbone network is modeled as a directed graph $G(V,L)$ where delay elements (e.g., a trunk or a switch fabric) are represented by directed links and the junctions between delay elements are represented by nodes. The nodes (junctions) do not incur any delay. Let L be the set of links and V be the set of nodes in the graph (network).

Consider a planning/engineering problem for a heterogeneous SMDS network denoted by problem (P). The goal is to replace the original intractable and nonconvex end-to-end delay constraints into $|L|$ convex link constraints. More importantly, any solution satisfying the link constraints should satisfy the original end-to-end delay constraints. A procedure is proposed below.

The first step is to identify a set of performance indicators $m^1, m^2, ..., m^n$, e.g. the overdue probability and the time threshold, such that for a path consisting of links 1 to k the following condition

$$\mathop{O^j}_{l=1}^{k} m_l^j \le M^j \quad \forall j = 1,2,...,n \tag{1}$$

guarantees the overdue probability for this path be no greater than 5%, where O^j is an operator, e.g. $\sum$, for performance indicator m^j, m_l^j is the allocated type-j performance indicator to link l and M^j is a prespecified upper bound on the aggregate effect (through operator O^j) of type-j performance indicators.

Assume the longest-hop path involves K hops (or alternatively a hop constraint can be imposed). Then, let $m_l^j, \forall l \in L, j = ,1,2,...,n$, be the solution to the following equation

$$\mathop{O^j}_{i=1}^{K} m_l^j = M^j. \tag{2}$$

It is clear that with this assignment strategy, if all m_l^j's are satisfied then the end-to-end delay constraints are satisfied. Otherwise, given an $\{m_l^j\}_{l,j}$ assignment, we may need to identify those paths where the end-to-end delay constraints are violated and exclude them from the feasible set, which is usually intractable. The next step is for each link l to calculate the highest utilization where the allocated $\{m_l^j\}_j$ are satisfied.

This general approach can be generalized to allow non-even allocation of performance indicators, which can best be illustrated by the following example. If a particular network element l is much more expensive than the others and/or is obviously the bottleneck of the network, then we may want to increase its allowable utilization than the allocated value by the original approach. Then, we can increase m_l^j to be greater than the solution to Equation (2) and set m_i^j to be the solution to

$$O^j(m_i^j, m_i^j, \ldots, m_i^j, m_l^j, m_i^j, \ldots, m_i^j, m_i^j) = M^j$$

for every other link i.

Next, we describe an implementation procedure to achieve the best result when solving problem (P). Let problem (P_K) be problem (P) where the above general allocation approach is used to simplify the constraints and a hop constraint which requires each active path not involve more than K hops is added. Note that with this hop constraint, a feasible solution to problem (P_K) is still a feasible solution to problem (P). In a solution procedure to (P_K), a shortest-path algorithm is usually incorporated. The shortest-path problem seems to be the only part impacted by this additional hop constraint. However, the hop-constrained shortest-path problem is still solvable in polynomial time. The impact of K can be grossly described below. When K is increased, the allocated utilization for each link becomes smaller, which tends to increase the objective function value, but the routing flexibility is increased, which tends to decrease the objective function value. A series of problem (P_K)'s, where K may typically ranges from the network diameter to $|V|-1$, are then solved. The objective function value for each of the problems are recorded and the best is chosen.

3. ALLOCATION SCHEMES FOR HETEROGENEOUS NETWORKS

In this section, based upon the general allocation approach proposed in Section 2, we discuss a number of possible allocation schemes for heterogeneous networks. In Section 4, 2 allocation schemes suitable for homogeneous networks are discussed.

3.1 APPROACH 1: LONGEST DELAY CONTROL

One possible approach to controlling the network element loads so that the delay objectives will be met is to limit the longest packet delay on each network element. More precisely, a packet will either be transmitted strictly within the given delay threshold (by properly choosing the buffer size) or be dropped (due to buffer overflow). Note that for those packets discarded due to buffer overflow, the corresponding delays are infinite from the standpoint of an SMDS network. As such, the end-to-end packet loss probability should be no greater than 5%. However, a more stringent end-to-end packet loss performance objective is specified in [2], which requires that the end-to-end packet loss probability be no greater than 10^{-4}. Assume that the buffer overflow probability of queue i can be expressed as a function of the utilization factor ρ_i and the buffer size J_i (in packets), denoted by $B_i(\rho_i,J_i)$. Let t_i be the service time of the longest packet on server i. Let $\bar{t}$ be the end-to-end delay threshold. Consider a path comprising queues 1 to k. Then the problem is to find a feasible solution to the following system:

$$\begin{cases} \prod_{i=1}^{k} (1 - B_i(\rho_i,J_i)) \ge 1 - 10^{-4} \\ \sum_{i=1}^{k} J_i\, t_i \le \bar{t}. \end{cases}$$

To apply the general approach described in Section 2, for each link i we set

$$B_i(\rho_i,J_i) = 1 - (1 - 10^{-4})^{\frac{1}{K}}$$

$$J_i\, t_i = \frac{\bar{t}}{K}.$$

If $B_i(\rho_i,J_i)$ is a monotonically increasing function of ρ_i, ρ_i can be determined from the above system using standard line search techniques.

This approach is simple but may involve the following difficulties. First, a mechanism to control the buffer size may not be available for an SMDS Switch System (SS) through an Operations Support System (OSS). Second, as the number of hops increases, ρ_i may decrease fast due to the joint effect of decreased J_i and decreased B_i.

3.2 APPROACH 2: COMPLETE DECOMPOSITION

If the probability density function (pdf) of the delay on each network element is known and can be expressed as a function of the utilization factor, we may allocate the end-to-end delay objectives by properly allocating the overdue probability and the time threshold. Let T_i be the delay on network element i (a random variable). Let $f_{T_i}(t_i,\rho_i)$ be the pdf of T_i, where ρ_i is the utilization factor of network element i. Let $F_{T_i}(t_i,\rho_i)$ be the probability distribution function (PDF) of T_i. Let $F_{T_p}(t_p,R_p)$ be the PDF of the end-to-end delay T_p ($= \sum_{i \in h_p} T_i$) when path p is considered, where h_p is the set of links on path p and R_p is the vector of ρ_i $\forall i \in h_p$. The following lemma provides one possible way of allocating the end-to-end delay objectives.

Lemma 1: *If $F_{T_i}(b_i\, t,\rho_i) \ge 1 - a_i$, $b_i \ge 0$, $\sum_{i=1}^{k} b_i \le 1$, $\forall i = 1,2,...,k$, where T_i's are mutually independent, then $F_{T_p}(t,R_p) \ge \prod_{i=1}^{k}(1 - a_i)$.*

Proof: $F_{T_p}(t,R_p)$ is the integral of $f_{T_1}(t_1,\rho_1) \circledast f_{T_2}(t_2,\rho_2) \circledast \cdots \circledast f_{T_k}(t_k,\rho_k)$ over the subspace $S = \left\{(t_1,t_2,\ldots,t_k) \mid \sum_{i=1}^{k} t_i \le t\right\}$ where the symbol $\circledast$ is used to denote the convolution operator. This subspace contains $S' = \left\{(t_1,t_2,\ldots,t_k) \mid t_i \le b_i\, t, \forall i = 1,2,...,k\right\}$. Take the integral over subspace S' and then we have $F_{T_1}(b_1\, t,\rho_1) \times F_{T_2}(b_2\, t,\rho_2) \times \cdots \times F_{T_k}(b_k\, t,\rho_k)$. Use the preconditions and then the result follows. □

Following the general approach described in Section 2, for each link i we set

$$a_i = 1 - 0.95^{\frac{1}{K}}$$

$$b_i = \frac{1}{K}.$$

This approach has an advantage of simplicity. However, one potential drawback of this approach is that over-conservative decisions may be made, especially when K is large. A quality indicator of the utilization threshold determined is the space defined by S' divided by that defined by S. When the ratio is low, the determined utilization threshold tends to be low. As an example, when all b_i's are chosen to be $1/K$ (the ratio is maximized), the ratio becomes $K!/K^K$, which approaches zero as K goes to infinity.

3.3 APPROACH 3: GI/G/1 BOUNDING SCHEME

In this scheme, a $GI/G/1$ queueing model is considered and the required input is the interarrival time distribution and the service time distribution. This scheme is based upon a result due to Kingman[4]. Using another result also due to Kingman[5], a similar allocation scheme can be developed where the input is the first two moments of the interarrival time distribution and the service time distribution.

We briefly summarize the result in [4] below. Let $A_i^*(s)$ be the Laplace transform of the interarrival time distribution for link i. Let $B_i^*(s)$ be the Laplace transform of the service time distribution for link i. Let $C_i^*(s)$ be $A_i^*(-s) \times B_i^*(s)$. Let W_i be the waiting time on network element i. Let $F_{W_i}(t,\rho_i)$ be the probability distribution function of W_i. Then the result states

$$1 - F_{W_i}(t,\rho_i) \le e^{-s_0 t} \tag{3}$$

where s_0 is found from

$$s_0 = \sup\left\{ s > 0 : C_i^*(-s) \le 1 \right\}. \quad (4)$$

It is also worth mentioning that the average waiting time $\overline{W}_i$ is upper bounded by $1/s_0$. This result may be useful for other networks where the performance objective is the mean end-to-end delay.

From Equation (3), $F_{W_i}(t,\rho_i)$ is lower bounded by the PDF of an exponentially distributed random variable with mean s_0. In this allocation scheme, the performance indicator is s_0. For a path with K hops, let the allocated value of s_0 be the same, denoted by v. For simplicity and illustration purposes, we assume deterministic service times. Therefor, to consider the overdue probability, we can consider the total waiting time and adjust the time threshold accordingly (the original time threshold minus K times the service time).

Consider the following result.

Lemma 2: *If* $F_{T_1}(t,\rho_1) \le F_{T_2}(t,\rho_2)$ *and* $F_{T_3}(t,\rho_3) \le F_{T_4}(t,\rho_4)$ $\forall t$, *then* $F_{T_1+T_3}(t,\rho_1,\rho_3) \le F_{T_2+T_4}(t,\rho_2,\rho_4)$ $\forall t$.

Proof:

$$\begin{aligned} F_{T_1+T_3}(t,\rho_1,\rho_3) &= \int_0^t f_{T_1}(t-y,\rho_1) F_{T_3}(y,\rho_3)dy \\ &\le \int_0^t f_{T_1}(t-y,\rho_1) F_{T_4}(y,\rho_4)dy \\ &= \int_0^t f_{T_4}(t-y,\rho_4) F_{T_1}(y,\rho_1)dy \\ &\le \int_0^t f_{T_4}(t-y,\rho_4) F_{T_2}(y,\rho_2)dy \\ &= F_{T_2+T_4}(t,\rho_2,\rho_4). \quad \square \end{aligned}$$

This lemma immediately leads to the following proposition.

Proposition 1: *If* $F_{T_i}(t,\rho_i) \le F'_{T_i}(t,\rho_i)$ $\forall t \ge 0, i = 1,2,\ldots,k,$ *then* $F_{T_1}(t,\rho_1) \circledast f_{T_2}(t,\rho_2) \circledast f_{T_3}(t,\rho_3) \circledast \cdots \circledast f_{T_k}(t,\rho_k) \le F'_{T_1}(t,\rho_1) \circledast f'_{T_2}(t,\rho_2) \circledast f'_{T_3}(t,\rho_3) \circledast \cdots \circledast f'_{T_k}(t,\rho_k)$.

Then the PDF of the end-to-end delay is lower bounded by the PDF corresponding to the following Laplace transform

$$\frac{1}{s}\left[\frac{v}{s+v}\right]^K.$$

Calculating 1 minus the inverse Laplace transform of the above expression yields

$$e^{-vt}\left[1 + v\,t + \frac{v^2 t^2}{2!} + \cdots + \frac{v^{(K-1)}\,t^{(K-1)}}{(K-1)!}\right]. \quad (5)$$

The next step is to calculate v such that the above expression equals 5%. It can be verified that the above expression is a monotonically decreasing function of v. Therefore, standard line search techniques can be applied.

Once the value of s_0 (= v) for each link is calculated, we may apply Equation (4) to calculate the utilization threshold. Two examples are given below to demonstrate this step.

3.3.1 M/M/1

Consider an $M/M/1$ queue with mean arrival rate λ and mean service rate μ. We first determine the maximum s that satisfies $\frac{\lambda}{s+\lambda}\frac{\mu}{-s+\mu} \le 1$. It is easy to find that $s_0 = \mu - \lambda$. Therefore, the utilization threshold is set to $1 - s_0/\mu$.

3.3.2 M/D/1

The arrival process is Poisson with parameter λ. Let the service time for each packet be $t_{\max}$. We first determine the maximum s that satisfies $\frac{\lambda}{s+\lambda} e^{s\,t_{\max}} \le 1$ or equivalently $\frac{\rho}{s\,t_{\max}+\rho} \le e^{-s\,t_{\max}}$. It can be shown that at s_0 the equality hold and there are two roots (one is at 0) if the service rate ($1/t_{\max}$) is greater than the arrival rate. We can use standard line search schemes to find the second root.

3.4 APPROACH 4: MARKOV INEQUALITY

If the mean delay on each link is available and can be expressed as a function of the link utilization factor, the following allocation scheme based upon the Markov inequality is proposed. Consider path p. Let $\overline{T}_l(\rho_l)$ be the mean delay on link $l \in h_p$. Let t be the end-to-end time threshold. Then by the Markov inequality

$$1 - F_{T_p}(t,R_p) \le \frac{\sum_{l \in h_p} \overline{T}_l}{t}.$$

Setting $\sum_{l \in h_p} \overline{T}_l/t \le 5\%$, the problem is then reduced to allocating $0.05t$ to the links on path p. Then the maximum link utilization factors can be calculated through $\overline{T}_l(\rho_l)$.

Following the general approach proposed in Section 2, we set

$$\overline{T}_l(\rho_l) = \frac{0.05\,t}{K}.$$

Since $\overline{T}_l(\rho_l)$ is typically a monotonically increasing function of ρ_l, ρ_l can be calculated using standard line search techniques, if not analytically.

One comment on this scheme is given below. If $\overline{T}_l(\rho_l)$ is a convex function and the number of candidate paths is manageable, we may apply the formulation and solution approach in [6] to improve the effectiveness. More precisely, 0.05 t is not evenly preassigned to each link. Instead, for each path p we consider the following constraint

$$\sum_{l \in h_p} \overline{T}_l(\rho_l) \le 0.05\,t. \quad (6)$$

This alternative treatment is attributed to the convex property associated with Equation (1) in this scheme (shown in Equation (6)), which is unique among the schemes discussed in this section.

3.5 APPROACH 5: CHEBYSHEV INEQUALITY

If the mean and variance of delay on each link are available and they can both be expressed as functions of the link utilization factor, an allocation scheme based upon the Chebyshev inequality is possible. As will be shown shortly, this scheme has similarity to the normal approximation scheme to be discussed in the next subsection. However, one major difference between these two schemes is that the Chebyshev inequality scheme guarantees a feasible allocation strategy while the normal approximation scheme does not.

Let T be a random variable. The Chebyshev inequality makes use of the mean $\overline{T}$ and variance σ_T^2; it states that for any $t > 0$

$$P[\,|T - \overline{T}| \ge t] \le \frac{\sigma_T^2}{t^2}.$$

If (i) the mean delay $\overline{T}_l(\rho_l)$ and the variance of delay $\sigma_{T_l}^2(\rho_l)$ on link $l \in h_p$ are known, (ii) the delays are mutually independent, and (iii) the end-to-end time threshold is t, then by the Chebyshev inequality

$$\begin{aligned} 1 - F_{T_p}(t,R_p) &\le 1 - F_{T_p}(t,R_p) + F_{T_p}(t - 2\sum_{l\in h_p} \overline{T}_l(\rho_l),R_p) \\ &= P[\,|T_p - \sum_{l\in h_p} \overline{T}_l(\rho_l)| \ge t - \sum_{l\in h_p} \overline{T}_l(\rho_l)] \\ &\le \frac{\sum_{l\in h_p} \sigma_{T_l}^2(\rho_l)}{(t - \sum_{l\in h_p} \overline{T}_l(\rho_l))^2}. \end{aligned} \quad (7)$$

In this scheme, we use the mean and variance of delay as the performance indicators. Let the mean and the variance of delay allocated to each link be D and V, respectively. If the network is homogeneous, then for each link the mean and the variance have the same relationship. Therefore, the utilization threshold can be determined by solving

$$\frac{K\,V}{(t - K\,D)^2} = 0.05. \quad (8)$$

However, for heterogeneous networks, the relationship between the mean and the variance may not be the same for all links. Under this condition, a uniform/representative relationship must be described, e.g. $V = a\,D^2 + b\,D + c$ where a, b and c are constants. This relationship together with Equation (8) can be used to determine V and D. For each link, a utilization threshold can be determined by V and D, respectively. The smaller value is chosen.

3.6 APPROACH 6: NORMAL APPROXIMATION

If the first two moments of each link delay is known, then the standard normal approximation technique can be applied. Let $T_1, T_2, \ldots$ be independent identically distributed random variables having mean $\overline{T}$ and finite nonzero variance σ^2. Set $S_k = \sum_{i=1}^k T_i$. Then from the Central Limit Theorem

$$\lim_{k\to\infty} P\left[\frac{S_k - k\overline{T}}{\sigma\sqrt{k}} \le x\right] = \Phi(x), \qquad -\infty < x < \infty$$

where

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{(2\pi)^{1/2}} e^{-y^2/2}\,dy.$$

The Central Limit Theorem strongly suggests that for large k we can make the approximation

$$P\left[\frac{S_k - k\bar{T}}{\sigma\sqrt{k}} \le x\right] \approx \Phi(x), \qquad -\infty < x < \infty$$

or equivalently

$$P[S_k \le x] \approx \Phi\left[\frac{x - k\bar{T}}{\sigma\sqrt{k}}\right], \qquad -\infty < x < \infty. \tag{9}$$

Since we require that the overdue probability be no greater than 0.05, by a table look-up we have

$$\frac{x - k\bar{T}}{\sigma\sqrt{k}} \ge 1.645.$$

It is clear that this equation has the same structure as that of Equation (8). Therefore, as mentioned earlier, the procedure developed for the Chebyshev inequality scheme can be applied. We therefore omit the description of the procedure for the normal approximation scheme.

4. ALLOCATION SCHEMES FOR HOMOGENEOUS NETWORKS

In this section, 2 allocation schemes for homogeneous networks are discussed. Both schemes require the knowledge of the delay distribution for each link.

4.1 APPROACH 7: CONVOLUTION SCHEME

With full information about the link delay distribution and the assumption of mutual independence among link delays, this approach exactly calculate the end-to-end delay distribution (and thus the overdue probability) by convolution. By letting the utilization threshold be the same for each link, we consider a path with K hops and can express the overdue probability as a univariate monotonically increasing function of the utilization threshold. Numerical procedures for convolution (with high computational complexity) and line search can be applied to calculate the utilization threshold such that the overdue probability equals 0.05.

To establish the validity of this approach, we provide the following result.

Lemma 3: *If $F_{T_l}(t, \rho_l)$ is a monotonically decreasing function of ρ_l $\forall l \in \{1,2\}$ and $\rho_1' \ge \rho_1$, where T_1 and T_2 are independent, then $F_{T_2+T_1}(t, \rho_2, \rho_1) \le F_{T_2+T_1}(t, \rho_2, \rho_1')$ $\forall t$.*

Proof:

$$F_{T_2+T_1}(t, \rho_2, \rho_1) = \int_0^t f_{T_2}(t-y, \rho_2) F_{T_1}(y, \rho_1) dy$$

$$\le \int_0^t f_{T_2}(t-y, \rho_2) F_{T_1}(y, \rho_1') dy$$

$$= F_{T_2+T_1}(t, \rho_2, \rho_1'). \qquad \square$$

The monotonicity assumption made in Lemma 3 states that for a given time threshold the overdue probability increases as the utilization factor increases. For a typical queueing system, this assumption should be valid.

Proposition 2: *The exact overdue probability for path p is upper bounded by $1 - F_{T_p}(t_p, \rho, \rho, \ldots, \rho)$ where $\rho = \max_{l \in h_p} \rho_l$.*

Proof: Apply Lemma 3 $|h_p| - 1$ times. Then the result follows. $\square$

This proposition basically states that if any $\rho_l < \rho$, $l \in h_p$, then the end-to-end delay objective for path p is still satisfied.

4.2 APPROACH 6: CHERNOFF BOUNDING SCHEMES

A rather sophisticated means for bounding the tail of the sum of a *large* number of independent, identically distributed random variables is available in the form of the Chernoff bound. It involves an equality similar to the Markov and Chebyshev inequalities, but makes use of the entire distribution of the random variable itself. Again, line search techniques are required to calculate the utilization threshold. Due to the rather complicated form of the bounds, we do not show the inequality here. The interested reader is referred to Kleinrock[7] for details.

One major advantage of this approach is that legitimate upper bounds on the link utilization factors for a given time threshold are provided. Compared with the convolution scheme, the same amount of information is required, but lower computational complexity is involved. On the other hand, a disadvantage of this approach is that the Chernoff bound tends to be loose when the number of random variables is small.

5. COMPARISONS AMONG APPROACHES -- TWO CASE STUDIES

In this section, two comparisons among different schemes using $M/M/1$ and $M/D/1$ models, respectively, are made. Several points need to be emphasized regarding the comparisons. First, these comparisons should be deemed as an illustration of a number of theoretic results (e.g., the relative bound quality and the behavior of the bounds as the number of hops increases). Second, the criterion used in the comparisons is solely the bound quality. Third, to compare all schemes discussed, a homogeneous network is considered. Fourth, these comparisons certainly favor the convolution scheme because the convolution scheme calculates exact overdue probabilities and thus serves as a benchmark.

5.1 COMPARISON USING $M/M/1$ MODELS

In this set of performance test, $M/M/1$ queueing models are adopted. It is assumed that non-pipelining SSs are used. In other words, there is segmentation and reassembly operation performed in intermediate SSs. As such, for each network element the system time should be considered.

For the complete decomposition scheme, we want to find the minimum r such that $e^{-r\,t/k} \le a$ $(b_i = 1/k)$, where $(1-a)^k = 95\%$. Then, $e^{-r\,t/k} \le 1 - 0.95^{1/k}$. Consequently,

$$r\,t \ge -k \ln(1 - 0.95^{1/k}).$$

The minimum $r\,t$ (residual capacity and time threshold product) satisfying the above inequality is plotted in Figure 1.

For the Markov inequality scheme, $|h_p|/[(\mu - \lambda)\,t] \le 5\%$. Consequently, $(\mu - \lambda)\,t \ge 20\,|h_p|$.

For the Chebyshev inequality scheme,

$$\frac{\frac{|h_p|}{r^2}}{(t - \frac{|h_p|}{r})^2} \le 5\%.$$

where $r = \mu - \lambda$ is the residual capacity. After simple algebra, $r\,t \ge \sqrt{20\,|h_p|} + |h_p|$.

For the normal approximation scheme,

$$1 - F_{T_p}(t, R_p) \approx 1 - \Phi\left[\frac{t - |h_p|/r}{r\sqrt{|h_p|}}\right] \tag{10}$$

where r is again the residual capacity for each link on path p. Setting the RHS of (10) to 5% yields $r\,t = 1.645\sqrt{|h_p|} + |h_p|$

Applying the Chernoff bounding scheme, we solve the following equation

$$\exp\left[|h_p| - r\,t + |h_p| \ln\left(\frac{r\,t}{|h_p|}\right)\right] = 0.05.$$

When applying the convolution scheme, Expression (5) can be used to characterize the overdue probability where v in (5) is replaced by the residual capacity r. We then applied the bisecting search method (without using derivatives) to find the minimum $r\,t$.

For $M/M/1$ models, results by different schemes are summarized in Figure 1. The first observation from Figure 1 is that the curve corresponding to the complete decomposition scheme is convex in the observed range and tends to diverge from the exact curve (associated with the convolution scheme). Therefore, the complete decomposition scheme is not suggested when the number of hops is large. Second, bounds obtained by using the Markov inequality are loose and therefore only the first two points (for k equals 1 and 2) are plotted in Figure 1. One property which is not shown in the figure is that the curve associated with the Markov inequality scheme is linear. Third, the approaches using the Chebyshev inequality and the Chernoff bounds have comparable performance when the number of hops is small. Moreover, the Chebyshev inequality scheme degrades faster with the number of hops. Fourth, the curve corresponding to the normal approximation is close to the exact curve. In addition, the percentage error (from the true values) improves as the number of hops increases. However, the normal approximation does not guarantee that the residual capacities (or utilization factors) obtained will satisfy the end-to-end delay objective, as shown in this case. Last, the curve corresponding to the convolution scheme (true $M/M/1$) is concave.

5.2 COMPARISON USING $M/D/1$ MODELS

For $M/D/1$ queues, we assume that the link capacities are 45 Mbps and that the packet size is 83 Kbits. It is assumed that pipelining SSs are used. In other words, there is no segmentation and reassembly operation performed in intermediate SSs. Also because the length of a Level 2 Protocol Data Unit, L2_PDU, (53 octets) is small compared with a Level 3 Protocol Data Unit, L3_PDU, the packet (L3_PDU) transmission time will not be counted in the intermediate SSs. The delay threshold (after considering the deterministic service time) is assumed to be 15.625 msec.

For analyzing the packet loss probability of the longest delay control scheme, we consider a slotted $M/D/1/J$ system, where each time slot equals the packet (deterministic) service time and packets arriving during a slot cannot be transmitted until the beginning of the next slot. It is clear that this slotted system overestimates the packet loss probability, especially when the system load is low. A procedure described in [8] can be applied to calculate the packet loss probability of the slotted $M/D/1/J$ system.

For a number of the discussed schemed, the pdf of the delay on each $M/D/1$ queue must be known. This can be obtained by the following analysis. Let the number of packets in the system (in service and in the queue) be n. The probability mass function (pmf) of n is given below[9]

$$\begin{cases} p_0 = 1 - \rho \\ p_1 = (1-\rho)(e^{\rho} - 1) \\ p_n = (1-\rho) \sum_{k=1}^{n} (-1)^{n-k} e^{k\rho} \left[\frac{(k\rho)^{n-k}}{(n-k)!} + \frac{(k\rho)^{n-k-1}}{(n-k-1)!} \right] \quad (n \geq 2) \end{cases}$$

where p_n is the probability that there are n packets in the system and the second factor in p_n is ignored for $k = n$. When a new packet arrives, it will potentially experience two delays. One is the residual service time for the packet in service. The other is the waiting time for the packets that are ahead in the queue. It is easy to verify that the residual service time is uniformly distributed in $[0, t_{max}]$. Given the condition that n packets are in the system ($n-1$ packets are in the queue) upon the arrival of the new packet, then the waiting time w is uniformly distributed in $[(n-1)\, t_{max}, n\, t_{max}]$. Removing the condition on n (using p_n), the pdf of the waiting time can be obtained. To calculate the exact overdue probability, a numerical procedure was developed to conduct the convolution operation based upon the pdf obtained above.

For a number of the proposed approaches, the mean and variance of the waiting time on an $M/D/1$ queue are required. The mean and variance can both be obtained by the P-K formula. For the convenience of the reader the variance is given below

$$\frac{3x^2\rho^2 + 4x^2\rho(1-\rho)}{12(1-\rho)^2}$$

where x is the (mean) service time.

The comparison of different approaches using $M/D/1$ queues is shown in Figure 2. In general, Figure 2 shows the same relative performance among the schemes compared in Figure 1. However, a number of new observations are obtained from Figure 2. First, the $GI/G/1$ bounds are the closest among the legitimate ones except for the convolution scheme which serves as a benchmark. In addition, the quality of the $GI/G/1$ bounds tends to degrade (slightly) as the number of hops increases. Second, the longest delay control scheme is very conservative and the bound quality degrades fast with the number of hops.

From the above two comparisons we may draw the following conclusions. First, we may consider the longest delay control scheme, the complete decomposition scheme and the Markov inequality scheme as secondary candidates when the bound quality is the major concern. Second, although the normal approximation scheme provides the smallest absolute error bound, the allocation strategies determined by this scheme are not feasible for these two test cases. Third, the Chernoff bounding scheme performs consistently with the number of hops in terms of the bound difference. Fourth, the Chernoff bounding scheme and the Chebyshev inequality scheme are comparable, although the Chebyshev inequality scheme degrades faster.

6. SUMMARY

This paper deals with a very important problem in network planning and engineering: how to circumvent the difficulty of (i) the excessively large number of end-to-end performance objective constraints and (ii) the nonconvexity of each of these constraints particularly resulting from the percentile nature of the end-to-end delay constraints for the SMDS service. A general approach is to replace the set of end-to-end delay constraints by a simpler set of link utilization constraints in such a way that, if the link utilization constraints are satisfied then the end-to-end delay constraints are satisfied.

For end-to-end percentile-type delay constraints, to compute an allocation strategy satisfying the aforementioned feasibility criterion is nontrivial, especially for heterogeneous networks. We then propose a general allocation approach using the concept of allocating a set of performance indicators as an intermediate step. FOr each link, the highest utilization satisfying the allocated performance indicators can be calculated. We also address the issue of how to apply this general allocation approach in a network planning/engineering problem to achieve the best result.

Based upon the general allocation approach, a number of allocation schemes for heterogeneous and homogeneous networks, respectively, are discussed. Their relative effectiveness in terms of the utilization thresholds determined is compared for a homogeneous network using $M/M/1$ and $M/D/1$ queueing models.

From the case studies, it is indicated that a number of the discussed schemes are over-conservative and may be impractical. However, more studies should be performed in the future to confirm/reject this result when more information about traffic characteristics and available traffic measurements is available.

The general allocation approach proposed in this paper is applicable for many new services, such as the Frame Relay service, Asynchronous Transfer Mode and the Advanced Intelligent Network services, where end-to-end percentile-type performance objectives are considered. This work lays a foundation to help network planners/administrators perform various network planning and engineering functions for a wide class of services.

Acknowledgments

The author would like to thank Vincent Calcagno, Jack Kichura, Arthur Lin, Jonathan Lu, Marco Mostrel, Jonathan Wang, Ellen White, and Mehmet Ulema for valuable information and fruitful discussions.

REFERENCES

1. F.Y.S. Lin, *Link Set Sizing for Networks Supporting SMDS*, to be presented in the Second IEEE Network Management and Control Workshop, September 1993.
2. Bell Communications Research, *Generic System Requirements in Support of SMDS*, Bellcore technical reference, TR-TSV-000772, Issue 1, May 1991.
3. J.L. Wang, *An Integrated Methodology for Supporting Network Planning and Traffic Engineering with Considerations to SMDS Service*, Proc. IEEE Globecom, December 1991.
4. J.F.C. Kingman, *Inequalities in the Theory of Queues*, Journal of the Royal Statistical Society, Series B, 32, pp. 102-110, 1970.
5. J.F.C. Kingman, *On Queues in Heavy Traffic*, Journal of the Royal Statistical Society, Series B, 24, pp. 383-392, 1962.
6. F.Y.S. Lin and J.R. Yee, *A Real-time Distributed Routing and Admission Control Algorithm for ATM Networks*, Proc. IEEE Infocom'93.
7. L. Kleinrock, Queueing Systems, Vol. 1, New York: Wiley Interscience, 1975.
8. J.F. Hayes, Modeling and Analysis of Computer Communications Network, 1984.
9. D. Gross and C.M. Harris, Fundamentals of Queueing Theory, John Wiley & Sons, 1974.

Fig. 1. A comparison among different approaches using M/M/1 models

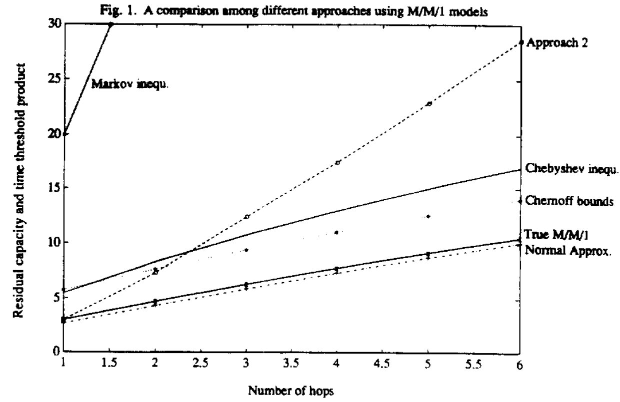

Fig. 2. A comparison among different approaches using M/D/1 models