A Resequencing Model for High Speed Networks

Ka-Cheong Leung and Victor O. K. Li
Department of Electrical and Electronic Engineering
The University of Hong Kong
Pokfulam Road, Hong Kong, China

Abstract—**In this paper, we propose a framework to study the resequencing mechanism in high speed networks. This framework allows us to estimate the packet resequencing delay, the total packet delay, and the resequencing buffer occupancy distributions when data traffic is dispersed on multiple disjoint paths. In contrast to most of the existing work, the estimation of the end-to-end path delay distribution is decoupled from the queueing model for resequencing. This leads to a simple yet general model, which can be used with other measurement-based tools for estimating the end-to-end path delay distribution to find an optimal split of traffic. We consider a multiple-node $M/M/1$ tandem network as a path model. When end-to-end path delays are Gaussian distributed, our results show that the packet resequencing delay, the total packet delay, and the resequencing buffer occupancy drop when the traffic is spread over a larger number of homogeneous paths, although the network performance improvement quickly saturates when the number of paths used increases. We find that the number of paths used in multipath routing should be small, say up to three. Besides, an optimal split of traffic occurs at paths with equal loads.**

I. Introduction

The convergence of the computer, communications, entertainment, and consumer electronics industry is driving an explosive growth in multimedia applications [16]. Recent studies show that multimedia traffic exhibits variability or correlation on various time scales [3]. Such long-range dependence property has a considerable unpleasant impact on queueing performance, and is a dominant characteristic for a number of packet traffic engineering problems [7].

Multipath routing has recently been found to be an effective method to alleviate the effects of such traffic correlation [9]. Precisely, *multipath routing* or *spatial traffic dispersion* [4], [9], [18] is a load balancing technique in which the total load from a source to a destination is spatially distributed over several paths. However, packets which travel along different paths may arrive out of order. Those packets arriving out of order may have to be stored in a buffer, called a resequencing buffer, until they can be delivered to the end process in the proper order. This is called *resequencing*, which is an important issue associated with communications using multiple paths.

Multimedia applications generally have stringent quality of service requirements. Multimedia data generated by these applications must arrive correctly to the receiving end within a specified period of time in order to be useful. In addition, those data that do not satisfy such delay constraint have to be confined below a certain percentage. To see whether and when multipath routing can have a better performance than single path routing, models are required to characterize the resequencing mechanism.

Existing work can be grouped into two major categories. The first category consists of work that characterizes the disordering network[1] as a queueing system with several servers sharing a single queue [5], [6], [20]. In [20], the source node together with the set of outgoing links have been modeled as an $M/M/m$ queue with servers of different rates. The resequencing delay distribution and the average resequencing delay were derived. The model was then used to compare the performance of channel level and virtual circuit resequencing. In [5], the author extended the problem addressed in [20] by allowing possibly different average message lengths for messages from different virtual circuits. An $M/M/\infty$ model was assumed in which the average service rates depended only on the virtual circuits. Only average values on resequencing delay and buffer length have been deduced. In [6], distributions of resequencing delay and total delay have been evaluated for three different queueing models with homogeneous servers, namely, the $G/M/m$ model, the $G/M/\infty$ model, and the $M/H_K/\infty$ model.

Under this category of work, the queue length distribution can be computed as in [13]. The resequencing delay distribution is then calculated by conditioning on the number of other customers (messages) being served when a tagged customer goes into service. The major weakness of this approach is that it allows at most one customer on any one path. Customers are always transmitted on the next idle path while it is impractical to obtain such signaling information in high speed networks because of terribly large bandwidth delay product.

In the second category of work, the disordering network is described as a queueing system with several servers and queues, where each server has its own dedicated queue [8], [11], [12], [15], [19]. In this case, customers can be transmitted on any paths without waiting for them to become idle. In addition, any orderly dispersion sequences can easily be incorporated into resequencing models. In [11], two parallel paths have been modeled as two heterogeneous $M/M/1$ queues. The mean resequencing delay of a tagged customer was derived by conditioning on the number of customers in the other queue being circumvented when the tagged customer has just finished its service. Besides, the impact of the interfering traffic on the resequencing delay has also been studied. This work has been extended in [8] in that a difference in the fixed delays on the two paths is inserted to illustrate the effect of fixed delays on the optimal splitting probability. The effect of deterministic, instead of probabilistic, routing sequence to queueing performance has been considered in [15].

Regardless of the fact that most of the existing work assumed Markovian arrival processes, the queueing analysis would become much more difficult or computationally infeasible when a path consists of multiple hops. Hence, there is a desire to decouple the estimation of the end-to-end path delay distributions from the calculation of resequencing delay distributions. Furthermore, as far as we know, there is no analytical model available to estimate the resequencing buffer occupancy distribution. This distribution is useful for network administrators who must estimate the size of the resequencing buffer so as to satisfy the quality of service constraints.

The objective of this work is to propose a framework to study the resequencing mechanism in high speed networks. This framework allows us to estimate the packet resequencing delay, the total packet delay, and the resequencing buffer occupancy distributions under an orderly dispersion of traffic on multiple disjoint paths. We will investigate the effectiveness of multipath routing by examining three basic questions:

- Does multipath routing improve the system performance? If so, when?
- What is the optimal split of traffic to achieve the best performance?
- What is the cost of employing multipath routing?

This paper is organized as follows. Section II gives a traffic model for the proposed framework. Section III presents an analytical model to compute the packet resequencing delay, the total delay, and the buffer occupancy distributions. Section IV develops an end-to-end path delay model for the framework. Section V examines the analytical results derived from the framework and studies the effectiveness of multipath routing. Section VI concludes our work.

II. Traffic Model

Our traffic model consists of a disordering network connecting the source to the destination, and flows of packets. The disordering network consists of a set of N disjoint paths, namely path 1, path 2, ..., path N, connecting the source to the destination, such that customers may arrive at the destination in a different order as they are sent. We assume the source has an unlimited supply of packets, which are delivered to paths according to a deterministic routing sequence. Let $\mathcal{R}(i)$ be the routing sequence function of the i^{th} packet. If $\mathcal{R}(i) = j$, the i^{th} packet will be routed on path j. Denote $s_{j,i}$, $d_{j,i}$, and $a_{j,i}$ respectively be the sending time of the first bit of the i^{th} packet from the source, the packet end-to-end path delay, and the arrival time of the last bit of the packet at the destination, which is delivered on path j. The relationship among these three quantities can be formulated as:

$$a_{j,i} = s_{j,i} + d_{j,i}$$

Given any two packets m and n, they arrive out of order only when packet m is sent before packet n at the source but arrives later at the destination. This means that $s_{\mathcal{R}(m),m} < s_{\mathcal{R}(n),n}$ and $a_{\mathcal{R}(m),m} > a_{\mathcal{R}(n),n}$. The condition can be simplified as:

$$d_{\mathcal{R}(m),m} - d_{\mathcal{R}(n),n} - i_{(\mathcal{R}(m),m),(\mathcal{R}(n),n)} > 0$$

[1] Disordering network is a network that allows customers to leave in a different order as they arrive [2].

0-7803-5284-X/99/$10.00 © 1999 IEEE.

where $i_{(\mathcal{R}(m),m),(\mathcal{R}(n),n)}$ is the inter-departure time between packets m and n at the source. Therefore, $i_{(\mathcal{R}(m),m),(\mathcal{R}(n),n)} = s_{\mathcal{R}(n),n} - s_{\mathcal{R}(m),m} > \frac{l_m}{c_{\mathcal{R}(m)}}$, where l_m is the length of packet m in bits and c_j is the bandwidth of path j in bits per time unit.

Besides, these two packets may arrive out of order only when they are sent on two different paths, i.e. $\mathcal{R}(m) \neq \mathcal{R}(n)$. This is true by the in-order channel assumption, which states that if an arbitrary packet m is transmitted from the source before another packet n on the same path connecting the source to the destination, packet m will arrive at the destination before packet n.

III. Resequencing Model

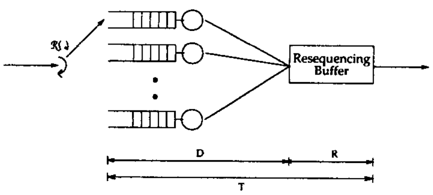


Fig. 1. The resequencing model.

Figure 1 illustrates the resequencing model. It consists of N queues in parallel where each queue j, $1 \leq j \leq N$, which corresponds to a transmission path, has a sojourn time distribution $D_j(\cdot)$. The system times are assumed to be mutually independent. The service discipline at each queue is first come first served (FCFS) so that the in-order channel assumption can be preserved. Customers or packets arrive into the system with a *general* inter-arrival time distribution. Each customer is routed deterministically according to the routing sequence function $\mathcal{R}(\cdot)$. The relative load to queue i, p_i, is defined as the portion of dispersed traffic to be routed to queue i, where $\sum_{i=1}^{N} p_i = 1$. To achieve the best system performance, customers are routed under a cyclical fashion so that the arrival instants of any two successive customers to each queue is as uniformly distributed as possible. Algorithms for constructing such routing sequences can be found in [1]. Upon service completion, they join the resequencing buffer to await the arrivals of all those customers that have entered the system before it. The total delay of each customer i in the system is defined as the sum of its sojourn time and resequencing delay. That is,

$$t_{\mathcal{R}(i),i} = d_{\mathcal{R}(i),i} + r_{\mathcal{R}(i),i}$$

The discussion will proceed as follows. Section III-A discusses how to determine the resequencing delay and total delay distributions. An estimation of buffer occupancy distribution is presented in Section III-B.

A. Determination of Resequencing Delay and Total Delay Distributions

Because of the in-order channel assumption, the delay analysis can be greatly simplified to merely consider $N-1$ packets by applying the following lemma.

Lemma 1: To determine the resequencing delay of an arbitrary packet, it is sufficient to consider the last packet to be transmitted before it on every path other than the one taken by it.

Proof: This has been proved in [19] for the case when the routing sequencing function is round robin. We are going to extend it to any general multipath routing discipline. The in-order channel assumption guarantees that the last packet transmitted before the arbitrary packet on a path other than the one taken by it arrives at the destination last among all packets transmitted before it on that path. The resequencing delay of the arbitrary packet is determined by the maximum time it waits for a packet that has been transmitted before it and received after it. Thus, it is sufficient to consider the last packet to be transmitted before it on every path other than the one taken by it. □

Define the random variables D_i, R_i, and T_i be the end-to-end path delay, the resequencing delay, and the total delay incurred for a packet to be transmitted on path i. Thus,

$$T_i = D_i + R_i$$

Define another random variable $I_{i,j|k}$ be the inter-arrival time to the system between a packet transmitted on path i and another packet transmitted on path j, $j \neq i$, given that the tagged packet is on path k, where k can either be i or j. The computation of the probability distribution of $I_{i,j|k}$, $P_{I_{i,j|k}}(\cdot)$, can be illustrated with the use of a two-path example. The time diagram of the packet arrival process for the example is shown in Figure 2. Consider the periodic routing sequence such that two consecutive packets are transmitted on path 1, following by a packet transmitted on path 2, and so on. The inter-arrival time between two successive packets to the system is exactly one time unit. $P_{I_{i,j|k}}(\cdot)$ can be calculated as:

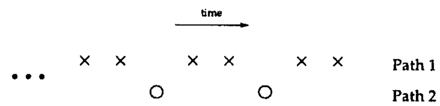


Fig. 2. The time diagram for two-path example.

$$P_{I_{1,2|1}}(d) = P_{I_{2,1|1}}(d) = \begin{cases} 0.5 & \text{if } d = 1,2; \\ 0 & \text{otherwise.} \end{cases}$$

$$P_{I_{1,2|2}}(d) = P_{I_{2,1|2}}(d) = \begin{cases} 1 & \text{if } d = 1; \\ 0 & \text{otherwise.} \end{cases}$$

Let $R_{i|j}$ be a random variable denoting the waiting time of a packet transmitted on path i, given that it has to wait until a packet transmitted on path j arrives. R_i is the maximum $R_{i|j}$, $j = 1, 2, ..., i-1, i+1, ..., N$. The cumulative distribution function of R_i can be computed as:

$$\begin{aligned} F_{R_i}(t) &= \text{Prob}(\max(R_{i|1}, R_{i|2}, \cdots, R_{i|i-1}, R_{i|i+1}, \cdots, R_{i|N}) \leq t) \\ &= \prod_{\substack{j=1 \\ j \neq i}}^{N} \text{Prob}([D_j - D_i - I_{j,i|i}]^+ \leq t) \\ &= [\prod_{\substack{j=1 \\ j \neq i}}^{N} \int_{-\infty}^{t} f_{D_j}(\varphi) \otimes f_{D_i}(-\varphi) \otimes f_{I_{j,i|i}}(-\varphi)\, d\varphi] \cdot u(t) \end{aligned}$$

where $f_X(t)$ is the density function of any random variable X, and $u(t)$ is the unit step function.

By the law of total probability [10], the cumulative distribution function of the packet resequencing delay is given by:

$$F_R(t) = \sum_{i=1}^{N} p_i\, F_{R_i}(t)$$

Since the total delay of a packet is the sum of its end-to-end path delay and resequencing delay, the cumulative distribution function of total delay for a packet can be calculated as:

$$F_T(t) = \sum_{i=1}^{N} p_i\, F_{T_i}(t) = \sum_{i=1}^{N} p_i \cdot [\int_{-\infty}^{t} f_{D_i}(\varphi) \otimes f_{R_i}(\varphi)\, d\varphi]$$

B. Estimation of Buffer Occupancy Distribution

The following lemma guides the evaluation of the resequencing buffer occupancy.

Lemma 2: The resequencing buffer occupancy seen by an arbitrary packet is the total number of packets, each of which has arrived at the destination before it and waiting for the arrival of a packet (which may or may not be the arbitrary packet) which has been transmitted before it.

Proof: This lemma can be arrived at directly from the definition of resequencing. □

Suppose a packet transmitted on path i has to wait in the resequencing buffer until a packet transmitted on path j arrives. We call path j the *bottleneck path* of the packet transmitted on path i. If such a bottleneck path does not exist, $j = i$. The resequencing buffer occupancy distribution, $P_B(\cdot)$, can be derived by following these steps:

1. Derive the probability distribution of the resequencing buffer occupancy on the condition that those awaiting packets have been transmitted on path j after the tagged packet transmitted on path i, $P_{A_a(j|i)}(\cdot)$
2. Deduce the probability distribution of the resequencing buffer occupancy on the condition that those awaiting packets have been transmitted on any paths other than path j before the tagged packet and the bottleneck path is path j, $P_{A_b(j|i)}(\cdot)$.
3. Determine the unconditional resequencing buffer occupancy probability distribution, $P_B(\cdot)$.

B.1 Evaluation of $P_{A_a(j|i)}(k)$

Suppose the random variable G_j is the inter-arrival time of two consecutive packets transmitted on path j. Let the random variable G_j^k be the sum of k independent, identically distributed random variables, each of which is statistically identical to the random variable G_j. If $A_a(j|i) = k$, this means that the first k packets transmitted on path j later than an arbitrary tagged packet transmitted on path i arrive at the destination before it, and that the $(k+1)^{\text{th}}$ packet arrives after it. Thus, the probability mass function of the resequencing buffer occupancy on the condition that those awaiting packets have been transmitted on path j after the tagged packet can be written as:

$$P_{A_a(j|i)}(k) = \begin{cases} \text{Prob}(D_i - D_j - I_{i,j|i} - G_j^k < 0) \\ \cdot[1 - \sum_{h=0}^{k-1} P_{A_a(j|i)}(h)] & \text{if } j \neq i, k = 0, 1, 2, \cdots; \\ 1 & \text{if } j = i, k = 0; \\ 0 & \text{otherwise.} \end{cases}$$

where $\text{Prob}(D_i - D_j - I_{i,j|i} - G_j^k < 0) = \int_{-\infty}^{0} f_{D_i}(\varphi) \otimes f_{D_j}(-\varphi) \otimes f_{I_{i,j|i}}(-\varphi) \otimes f_{G_j}^{(k)}(-\varphi)\, d\varphi$, and $f_{G_j}^{(k)}(t)$ denotes the k-folded convolution of $f_{G_j}(t)$.

B.2 Evaluation of $P_{A_b(j|i)}(k)$

Let random variable $M(q|i)$ denote the number of packets transmitted on path q earlier than an arbitrary tagged packet transmitted on path i and which arrive at the destination after it. The probability mass function of $M(q|i)$ can be given by:

$$P_{M(q|i)}(k) = \begin{cases} \text{Prob}(D_q - D_i - I_{q,i|i} - G_q^k < 0) \\ \cdot[1 - \sum_{h=0}^{k-1} P_{M(q|i)}(h)] & \text{if } q \neq i, k = 0, 1, 2, \cdots; \\ 1 & \text{if } q = i, k = 0; \\ 0 & \text{otherwise.} \end{cases}$$

Given that $M(q|i) \leq u$, the conditional probability mass function of $M(q|i)$ can be determined as:

$$P_{M(q|i)}(k|u) = \frac{P_{M(q|i)}(k)}{\sum_{h=0}^{u} P_{M(q|i)}(h)}$$

If the tagged packet sees n packets transmitted on path i waiting in the resequencing buffer, it can be inferred that the maximum number of packets transmitted on path q waiting in the buffer is:

$$w(n|q,i) = \lceil \frac{(n+1) \cdot p_q}{p_i} \rceil$$

If the tagged packet sees n packets transmitted on path i waiting in the resequencing buffer and the bottleneck path is path j, $j \neq i$, the maximum number of packets waiting in the buffer can be calculated as:

$$Y(j,i) = \sum_{\substack{q=1 \\ q \neq j}}^{N} w(n|q,i) - \frac{I_{i,j|j}}{Z}$$

where $G_i > I_{i,j|j}$ and Z is the random variable denoting the inter-arrival time between any two successive packets to the system.

Given that the tagged packet sees n packets transmitted on path i waiting in the resequencing buffer, the probability mass function of the number of awaiting packets transmitted on any paths before the tagged packet and the bottleneck path is path j can be written as:

$$P_{A_b(j|i)}(k|n) = P_{Y(j,i)}(k) \bigotimes_{\substack{q=1 \\ q \neq i,j}}^{N} [z_{q,j|i}\, P_{M(q|i)}(-k|w(n-1|q,i)) + z_{j,q|i}\, P_{M(q|i)}(-k|w(n|q,i))]$$

where $z_{q,j|i} = \text{Prob}(I_{q,i|q} > I_{j,i|j} \mid I_{q,i|q} \neq I_{j,i|j})$.

Denote $P_m(j|i)$ as the probability that path j is the bottleneck path of the tagged packet transmitted on path i. It can be computed as:

$$P_m(j|i) = \begin{cases} \text{Prob}(D_j - D_i - I_{j,i|i} > 0) \cdot [1 - \sum_{\substack{q=1 \\ q \neq i,j}}^{N} z_{q,j|i} P_m(q|i)] & \text{if } j \neq i; \\ 1 - \sum_{\substack{q=1 \\ q \neq i}}^{N} P_m(q|i) & \text{otherwise.} \end{cases}$$

The probability that the tagged packet sees n packets transmitted on path i waiting in the resequencing buffer can be found as:

$$P_o(n|i) = \prod_{\substack{q=1 \\ q \neq i}}^{N} \{ \sum_{h=0}^{w(n|q,i)} P_{M(q|i)}(h) - [w(n|q,i) - \frac{(n+1) \cdot p_q}{p_i}] \cdot P_{M(q|i)}(w(n|q,i)) \} - \sum_{u=0}^{n-1} P_o(u|i) - P_m(i|i)$$

Finally, the probability mass function of the number of awaiting packets transmitted on any paths before the tagged packet and the bottleneck path is path j can be approximated as:

$$P_{A_b(j|i)}(k) \approx \frac{\sum_{n=0}^{\infty} P_o(n|i)\, P_{A_b(j|i)}(k|n)}{1 - P_m(i|i)}$$

since we have used the average value of $P_o(n|i)$, instead of the corresponding value, conditioned on the bottleneck path being path j. We believe this approximation will not cause any inaccuracy.

B.3 Evaluation of $P_B(k)$

Denote random variable $A(j|i)$ as the total number of packets waiting in the resequencing buffer and seen by an arrival of a packet from path i and the bottleneck path is path j. By using the results in Sections III-B.1 and III-B.2, the probability mass function of $A(j|i)$ can be calculated as:

$$P_{A(j|i)}(k) = \begin{cases} \bigotimes_{\substack{h=1 \\ h \neq i}}^{N} P_{A_a(h|i)}(k) & \text{if } j = i; \\ P_{A_b(j|i)}(k) \bigotimes_{\substack{h=1 \\ h \neq i,j}}^{N} P_{A_a(h|i)}(k) & \text{otherwise.} \end{cases}$$

By applying the law of total probability [10], the probability mass function of the resequencing buffer occupancy seen by a packet from path i can be written as:

$$P_{B_i}(k) = \sum_{j=1}^{N} P_m(j|i)\, P_{A(j|i)}(k)$$

Further application of the law of total probability results in the resequencing buffer occupancy probability mass function:

$$P_B(k) = \sum_{i=1}^{N} p_i\, P_{B_i}(k)$$

where k is any non-negative integer, and $\sum_{k=0}^{\infty} P_B(k) = 1$.

The packet loss probability can also be estimated as follows. Let $\mathcal{B}$ be the size of the resequencing buffer, which is in terms of packets. From the system point of view, it is more realistic to assume that packets are all fixed size. However, the loss probability of any packet from path i can be approximated as:

$$P_{L_i} \approx \text{Prob}(B_i \geq \mathcal{B}) = 1 - \sum_{k=0}^{\mathcal{B}-1} P_{B_i}(k)$$

since we have neglected the effect on the queueing performance due to lost packets because of overflow at the resequencing buffer.

Using the law of total probability, the packet loss probability can be computed as:

$$P_L = \sum_{i=1}^{N} p_i\, P_{L_i} \approx 1 - \sum_{i=1}^{N} \sum_{k=0}^{\mathcal{B}-1} p_i\, P_{B_i}(k)$$

IV. End-to-End Path Delay Model

Generally, a path consists of several hops. Traffic measurements on the ARPANET[2] indicated that two sites were separated by 5.32 hops, on the average. Moreover, the distribution of round-trip delay appeared to be roughly normally distributed [14]. Modern high speed networks are much more complex in terms of domain architectures and connectivity. Thus, the mean path length between any two sites in a high speed network tends to be higher than the above value.

The central limit theorem [10] strongly suggests that the end-to-end path delay, which is the sum of a large number of independent hop delays, is approximately normally distributed. The mean and the variance of the end-to-end path delay provide sufficient information to generate an approximate distribution, which can then be utilized to compute the resequencing delay, the total delay, and the buffer occupancy distributions. This approach, which was used to solve the end-to-end percentile-type delay objective allocation problem for networks supporting Switched Multi-megabit Data Service (SMDS), has been shown to provide the best approximation to the reference values [17].

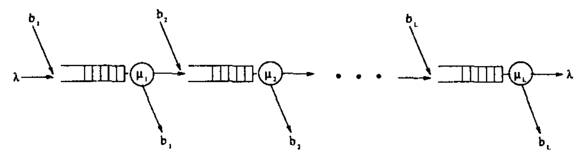


Fig. 3. The L-hop path example.

To complete our end-to-end path delay model, an L-hop path is modeled as a multiple-node $M/M/1$ tandem network as illustrated in Figure 3. L $M/M/1$ queues are connected in tandem. The i^{th} queue receives input from two traffic sources: the tagged dispersed traffic of rate λ, and the interfering or background traffic of rate b_i. The service rate of the i^{th} server is μ_i. Denote $\bar{D}$ and σ_D^2 as the mean and the variance of the end-to-end path delay. It can be shown [13] that:

$$\bar{D} = \sum_{i=1}^{L} \frac{1}{\mu_i(1-\rho_i)} \qquad \sigma_D^2 = \sum_{i=1}^{L} \frac{1}{\mu_i^2(1-\rho_i)^2}$$

where the utilization of the i^{th} server, $\rho_i = \frac{\lambda + b_i}{\mu_i}$.

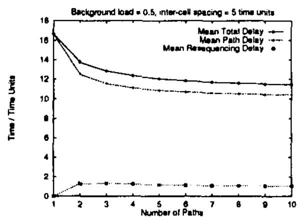
V. Analytical Investigation

This section discusses the numerical results based on the analytical expressions obtained in Sections III and IV. With the help of some numerical examples, we can illustrate the effectiveness of multipath routing by answering the three basic questions posed in Section I. Without loss of generality, packets, called cells, are all fixed size. The inter-sending time between any two consecutive cells from the source, denoted as the inter-cell spacing time, is a constant. A path is modeled as a five-node $M/M/1$ tandem network. Each server serves a packet with an average service time of one time unit. Its end-to-end path delay is assumed to be Gaussian or normally distributed with its mean and variance computed from the path model exhibited in Section IV.

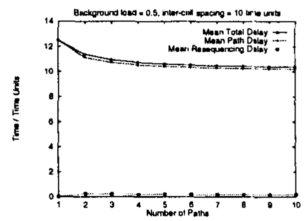
The results are provided in two sets. The first set studies the effectiveness of multipath routing with different number of homogeneous paths used, where the relative load to each path is the same. It provides the mean total delays, the mean path delays, the mean resequencing delays, the mean resequencing buffer occupancies, and the bounds on the complementary functions for the resequencing buffer occupancies, for each of the 40 cases (two background loads, two inter-cell spacing times, and ten path configurations). It also provides the resequencing buffer distributions for three cases (three path configurations). The second set studies the effectiveness of multipath routing using two heterogeneous paths, with possibly different relative loads according to a given dispersion ratio. The study includes the mean total delays, the mean path delays, the mean resequencing delays, the mean resequencing buffer occupancies, and the bounds on the complementary functions for the resequencing buffer occupancies, for each of the eight cases (eight dispersion ratios).

We examine the first set of results. Figure 4 shows the mean total delays, the mean path delays, and the mean resequencing delays. Two different background loads, namely link utilizations of 0.5 and 0.75, and two different inter-cell spacing, namely five and ten time units, are chosen. The mean total delay and the mean path delay drop as the number of paths used increases. Clearly, the decrease

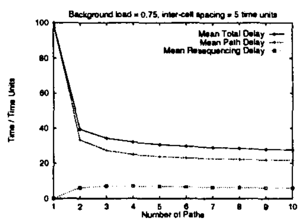
[2] ARPANET stands for the U.S. Department of Defense (DoD) Advanced Research Projects Agency (ARPA) Network. It began as an experimental packet-switched network and provided the groundwork for the development of the modern Internet.



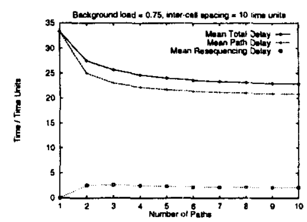
(a) Cell delay for background load of 0.5 and inter-cell spacing of 5 time units.



(b) Cell delay for background load of 0.5 and inter-cell spacing of 10 time units.



(c) Cell delay for background load of 0.75 and inter-cell spacing of 5 time units.



(d) Cell delay for background load of 0.75 and inter-cell spacing of 10 time units.

Fig. 4. Cell delay plots for various settings.

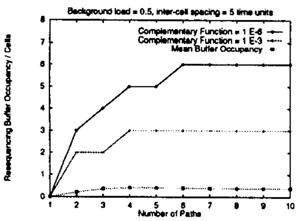
in the mean path delay comes from the decrease in utilization of each path. However, the improvement flattens with further increases in the number of dispersed paths. Surprisingly, the mean resequencing increases slightly and then flattens (or falls slightly) as the number of paths used increases. This results from the fact that when the number of dispersed paths is sufficiently large (say three), the fluctuation of delays between different paths is offset by the reduction in variances of path delays. Furthermore, the performance improvement is more significant when inter-cell spacing is smaller, as the load from the dispersed traffic to each path is higher. Thus, multipath routing is effective in performance improvement when the dispersed traffic load and the network load are both high.

Figure 5 exhibits the mean resequencing buffer occupancies, and two bounds of their complementary distributions, namely at 10^{-3} and 10^{-6}. They represent, to some extent, the system cost of multipath routing, because network administrators need to allocate sufficient buffer resources before a multipath connection can be established. From the figure, the mean resequencing buffer occupancy flattens as the number of dispersed paths increases, but this may not be the case when a cell loss bound is considered. This means that in order to provide the cell loss quality of service guarantee, possibly a much larger buffer size has to be allotted for resequencing when a larger number of dispersed paths is used. Besides, the probability distribution of the resequencing buffer occupancy tends to have a heavier tail for the cases when a larger number of paths is used, as shown in Figure 6. Thus, this argument does not favor using a large number of paths, say more than three, in multipath routing.

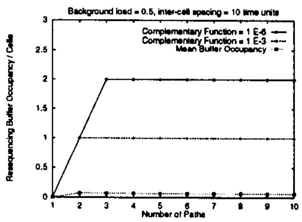
The second set of results demonstrates how to adjust the load distribution to each dispersed path to further improve the performance. Consider there are two heterogeneous paths, with background traffic of different loads. Define R as the dispersion ratio such that cells are transmitted on these two paths in a ratio of $R : 1$. Our result in Figure 7 shows that the resequencing delay and the resequencing buffer occupancy attain their minima when the total traffic load along these paths are the same. In other words, an optimal split of traffic is to ensure that the total load of each dispersed path is more or less the same, which is intuitively satisfying.

VI. Conclusions

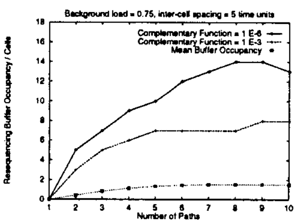
In this paper, we have proposed a framework to study the resequencing mechanism in high speed networks. This framework allows us to estimate the packet resequencing delay, the total packet delay, and the resequencing buffer occupancy distributions when traffic are dispersed on multiple disjoint paths.



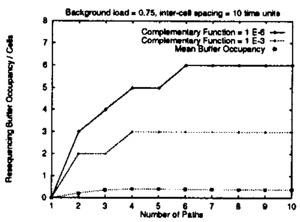
(a) Resequencing buffer occupancy for background load of 0.5 and inter-cell spacing of 5 time units.



(b) Resequencing buffer occupancy for background load of 0.5 and inter-cell spacing of 10 time units.



(c) Resequencing buffer occupancy for background load of 0.75 and inter-cell spacing of 5 time units.



(d) Resequencing buffer occupancy for background load of 0.75 and inter-cell spacing of 10 time units.

Fig. 5. Resequencing buffer occupancy plots for various settings.

The traffic model has been constructed in a flexible manner so that any multipath routing mechanisms can be modeled easily. The resequencing model has been devised to allow us to compute all necessary performance metrics for resequencing. The end-to-end path delay model has been built to allow the queueing model for resequencing to be decoupled from that for a path. This leads to a simple yet general model, which can be used with other measurement-based tools for estimating end-to-end path delay distributions to find an optimal split of traffic.

To illustrate our proposed framework, we have considered a multiple-node $M/M/1$ tandem network as a path model. When end-to-end path delays are Gaussian distributed, our results show that the packet resequencing delay, the total packet delay, and the resequencing buffer occupancy drop when the traffic is spread over a larger number of homogeneous paths, although the network performance improvement quickly saturates when the number of paths used increases. Multipath routing is effective in using a small number of paths, say up to three. Besides, an optimal split of traffic occurs at paths with equal loads.

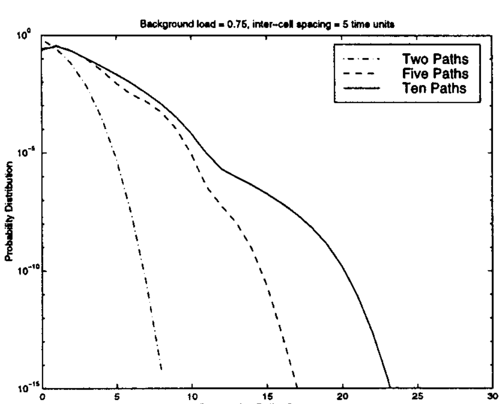
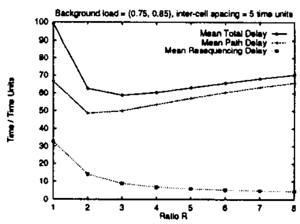
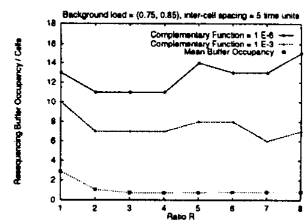


Fig. 6. Resequencing buffer occupancy plots for background load of 0.75 and inter-cell spacing of 5 time units.



(a) Cell delay plot.



(b) Resequencing buffer occupancy plot.

Fig. 7. Performance plots against different traffic splits.

REFERENCES

[1] Y. Arian and Y. Levy. Algorithms for Generalized Round Robin Routing. *Operations Research Letters*, Vol. 12, No. 5, pp. 313-319, November 1992.

[2] F. Baccelli and A. M. Makowski. Queueing Models for Systems with Synchronization Constraints. *Proceedings of the IEEE*, Vol. 77, No. 1, pp. 138-161, January 1989.

[3] J. Beran, R. Sherman, M. S. Taqqu, and W. Willinger. Long-Range Dependence in Variable-Bit-Rate Video Traffic. *IEEE Transactions on Communications*, Vol. 43, No. 2-4, pp. 1566-1579, February-April 1995.

[4] S. N. Chiou and V. O. K. Li. Diversity Transmissions in a Communication Network with Unreliable Components. *Proceedings of IEEE ICC '87*, Vol. 2, pp. 968-973, Seattle, WA, USA, 7-10 June 1987.

[5] S. Chowdhury. On Message Resequencing in Computer Networks. *Proceedings of the Eighth Annual International Phoenix Conference on Computers and Communications*, pp. 231-235, Scottsdale, AZ, USA, 22-24 March 1989.

[6] S. Chowdhury. Distribution of the Total Delay of Packets in Virtual Circuits. *Proceedings of IEEE INFOCOM '91*, Vol. 2, pp. 0911-0918, Bal Harbour, FL, USA, 7-11 April 1991.

[7] A. Erramilli, O. Narayan, and W. Willinger. Experimental Queueing Analysis with Long-Range Dependent Packet Traffic. *IEEE/ACM Transactions on Networking*, Vol. 4, No. 2, pp. 209-223, April 1996.

[8] N. Gogate and S. P. Panwar. On a Resequencing Model for High Speed Networks. *Proceedings of IEEE INFOCOM '94*, Vol. 1, pp. 40-47, Toronto, Ont., Canada, 12-16 June 1994.

[9] E. Gustafsson. Traffic Dispersion in ATM Networks. Ph.D. Dissertation, Royal Institute of Technology, TRITA-IT R 97:03, Kista, Sweden, June 1997.

[10] W. W. Hines and D. C. Montgomery. *Probability and Statistics in Engineering and Management Science*. Third Edition. John Wiley & Sons, 1990.

[11] I. Iliadis and Y.-C. Lien. Resequencing in Distributed Systems with Multiple Classes. *Proceedings of IEEE INFOCOM '88*, pp. 0881-0888, New Orleans, LA, USA, 27-31 March 1988.

[12] A. Jean-Marie and L. Gün. Parallel Queues with Resequencing. *Journal of the Association for Computing Machinery*, Vol. 40, No. 5, pp. 1188-1208, November 1993.

[13] L. Kleinrock. *Queueing Systems (Volume I: Theory)*. John Wiley & Sons, 1975.

[14] L. Kleinrock. *Queueing Systems (Volume II: Computer Applications)*. John Wiley & Sons, 1976.

[15] R. Krishnan and J. A. Silvester. Choice of Allocation Granularity in Multipath Source Routing Schemes. *Proceedings of IEEE INFOCOM '93*, Vol. 1, pp. 322-329, San Francisco, CA, USA, 28 March - 1 April 1993.

[16] V. O. K. Li and W. Liao. Distributed Multimedia Systems. *Proceedings of the IEEE*, Vol. 85, No. 7, pp. 1063-1108, July 1997.

[17] F. Y. S. Lin. Allocation of End-to-End Delay Objectives for Networks Supporting SMDS. *Proceedings of IEEE GLOBECOM '93*, Vol. 3, pp. 1346-1350, Houston, TX, USA, 29 November - 2 December 1993.

[18] N. F. Maxemchuk. Dispersity Routing. *Proceedings of IEEE ICC '75*, pp. 41-10 - 41-13, San Francisco, CA, USA, June 1975.

[19] N. T. Plotkin and P. P. Varaiya. Performance Analysis of Parallel ATM Connections for Gigabit Speed Applications. *Proceedings of IEEE INFOCOM '93*, Vol. 3, pp. 1186-1193, San Francisco, CA, USA, 28 March - 1 April 1993.

[20] T.-S. P. Yum and T.-Y. Ngai. Resequencing of Messages in Communication Networks. *IEEE Transactions on Communications*, Vol. COM-34, No. 2, pp. 143-149, February 1986.