Latency-Optimizing File Splitting for Transmission over a Large Multi-Hop Network

Raj Rao Nadakuditi & Mingyan Liu
Dept. of Electrical Engineering & Computer Science
University of Michigan, Ann Arbor, Michigan 48104
Email: {rajnrao}, {mingyan}@umich.edu

***Abstract*—We consider the problem of latency-optimizing file splitting for transmission over a large multi-hop network. We utilize recent results from random matrix theory to analytically express the latency distribution as a function of the relevant network parameters and the file size.**

The resulting minimax problem can be solved using standard nonlinear programming techniques for a variety of latency metrics. We illustrate the reduction in latency obtained using our technique, relative to some commonly-used heuristics, in the large-file-large-multi-hop network setting and highlight the role of random matrix theory in realizing the resulting performance gains.

I. Introduction

We study the problem of transmitting a large file over paths of potentially many hops, and seek optimal ways of splitting the file into a large number of packets over multiple paths, each with different operating parameters over its hops, to minimize the end-to-end delay. The form of delay we consider consists primarily of random queueing delay and transmission delay at each intermediate hops.

Compared to queueing analysis of tandem of queues, see e.g., [1] on reversible queueing networks, and [2], [3] on non-reversible queueing systems, a key distinguishing feature of our system is that in our tandem of queues there is no arrival process. All packets are assumed to be available at the beginning of the transmission. In this sense ours is not a steady state analysis. In addition, we do not assume a discrete time structure as is done in [2], [3]. We do assume that service times at each hop are independent and exponentially distributed. In this paper we focus on the setting where the queues are initially all empty; the results can be extended to the setting where the queues are assumed to be in equilibrium without much difficulty by utilizing the results in [12].

The above assumptions make our model more suited for applications like streaming and file transfer, rather than the delivery of randomly arriving packets. We assume that the file is divided into equal-length packets. In light of this, the independent and exponential service time assumption is justified by noting that at each hop the node may be serving other applications/files, thus the effective service time perceived by one particular stream/file is random and may be independent from one node to another. This also applies to a wireless scenario where the transmission times of equal-length packets vary depending on the level of congestion and interference in the environment due to backoff and contention resolution schemes used at the MAC layer.

Since our system never reaches steady state, our analysis does not involve the derivation of queue occupancy distribution which is commonly done in queueing analysis. Instead, we use recent results from random matrix theory [12] to analytically express the latency distribution as a function of the relevant network parameters and the file size in the large-file-size, large-network asymptotic regime.

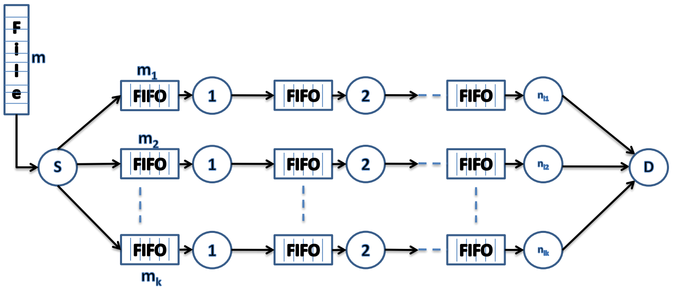


Fig. 1. The problem of splitting a large file for transmission over a large many-path network.

II. Problem: Optimal file split for fast delivery

We consider the following file transfer problem. Suppose we have a file segmented into M equal-length packets; M can potentially be very large. (The size of each packet is arbitrary, and can be normalized according to the nominal rate of transmission.) As depicted in Figure 1, we wish to transfer this file from a source node S to a destination node D. There are multiple routes we can potentially choose from for the transfer, we assume for the time being that these routes are mutually exclusive, i.e., they do not share common nodes. We ask the general question of how to decide what portion of the file to allocate to each route so as to minimize a certain delay measure.

Formally, we will denote the set of k routes between S and D by $\mathcal{R} = \{R_1, R_2, \cdots, R_k\}$. A route R_i contains an ordered list of l_i *relay* nodes, in addition to starting with S and ending with D, *i.e.*, $R_i = \{S, n_1^{(i)}, n_2^{(i)}, \cdots, n_{l_i}^{(i)}, D\}$. The service rates of nodes $n_j^i, \forall i = 1, \cdots, k, j = 1, \cdots, l_i$, denoted by $\mu_j^{(i)}$ are assumed to be known. Let $\lambda_j^{(i)} = 1/\mu_j^{(i)}$ be the corresponding (mean) service times. We assume that the packet service times at each node is exponentially distributed with mean $\lambda_j^{(i)}$ with independent service times at every queue and every packet.

A. Case 1: mean delay

Consider the first case where we are only concerned with the overall end-to-end delay of the task, *i.e.*, we would like to minimize the time from the start of the transfer till the last packet reaches the destination. Note that in this case packets do not need to be delivered in sequence, so the last packet just mentioned refers to the last packet to reach the destination, not necessarily the last packet in the sequence of M. In general, when multiple routes are used packets will be delivered out of order.

Specifically, suppose we split the set M into k sets, each denoted as M_i, $i = 1, \cdots, k$, and send the set M_i onto the ith route. With a slight abuse of notation, we will use M (as well as M_i) to denote both the set and the size of the set.

Denote by $t_i(M_i)$ the time of arrival of the last packet in the set M_i over route i. Note that $t_i(M_i)$ is a function of all service rates along route i, $\mu_1^{(i)}, \cdots, \mu_{l_i}^{(i)}$ or equivalently the mean service times $\{\lambda_j^{(i)} = 1/\mu_j^{(i)}\}$. It is however *not* a function of the positions of the packets within the set M_i, as these packets are assumed to be available at the beginning.

The optimization problem we would like to solve is the following and will be referred to as problem (P1):

$$\text{(P1)} \qquad \min_{M_1 \cup \cdots \cup M_k = M} E\left[\max\{t_1(M_1), \cdots, t_k(M_k)\}\right] \qquad (1)$$

As formulated above, (P1) is a constrained integer programming problem. Let $p_i = M_i/M$ be the proportion assigned to link i. Then the continuous relaxation of (P1), referred to as (P1a), yields the constrained optimization problem:

$$\text{(P1a)} \qquad \min_{p_1 + \cdots + p_k = 1} \max E\left[\{t_1(p_1), \cdots, t_k(p_k)\}\right]. \qquad (2)$$

B. Case 2: elastic delay deadlines

Suppose the file being transferred is being played back as packets are received. This will result in, possibly hard, deadline constraints imposed on individual packets. Assume the file transfer starts at time 0, and denote by $\bar{t}_i, i = 1, \cdots, M$ the desired time at which packet i should be received at node d. For instance, $\bar{t}_i$ could be the time at which packet i is scheduled to be played.

Suppose we use a function $f(t, \bar{t})$ to measure how much the deadline is missed, or the penalty assessed for missing it, when the deadline is $\bar{t}$ while the actual arrival time is t. This could be a simple difference, i.e., $f(t, \bar{t}) = [t - \bar{t}]^+$, or a uniform penalty $f(t, \bar{t}) = c \cdot I_{\{t > \bar{t}\}}$, or something more complicated.

The same delay measure $f()$ applies to vectors of times, i..e, $f(\mathbf{t}, \bar{\mathbf{t}})$, where $\mathbf{t} = \{t_1, t_2, \cdots, t_k\}$ denotes the arrival times of the last packets from each of the k routes and $\bar{\mathbf{t}} = \{\bar{t}_1, \cdots, \bar{t}_k\}$ are the respective deadlines.

Under such a setting, we would like to solve the following problem, referred to as problem (P2):

$$\text{(P2)} \qquad \min_{M_1 \cup \cdots \cup M_k = M} E\left[f(\mathbf{t}, \bar{\mathbf{t}})\right] \qquad (3)$$

As before with $p_i = M_i/M$, we have the relaxed continuous optimization problem:

$$\text{(P2a)} \qquad \min_{p_1 + \cdots + p_k = 1} \max E\left[\{f(t_1(p_1), \bar{t}_1), \cdots, f(t_k(p_k), \bar{t}_k)\}\right], \qquad (4)$$

where $f(t, \bar{t})$ measures how much the deadline was missed by the *last* packet. One can also perform the optimization with respect to minimizing the probability that the latency is greater than a certain deadline. We note that (P2) and (P2a) can handle, in principle at least, individual deadlines for each packet but, as given, we are only considering the "elastic" deadline for the last packet.

III. Pertinent result from random matrix theory

The relevant result from random matrix theory is stated next.

Theorem 3.1:

$$\frac{t_i(p_i) - p_i M \mu_i(p_i)}{\sqrt[3]{p_i M}\, \sigma_i(p_i)} \xrightarrow{D} \mathrm{TW}_2, \qquad (5)$$

where $\xrightarrow{D}$ denotes *almost sure* convergence and TW_2 is the type-2 Tracy-Widom distribution [4], [10] with $\mu_i(p_i)$ and $\sigma_i(p_i)$ given by:

$$\mu_i(p_i) = \frac{1}{y_i}\left(1 + \frac{1}{p_i M}\sum_{j=1}^{n_{l_i}^{(i)}} \frac{\lambda_j^{(i)} y_i}{1 - \lambda_j^{(i)} y_i}\right) \qquad (6)$$

$$\sigma_i(p_i) = \frac{1}{y_i^3}\left(1 + \frac{1}{p_i M}\sum_{j=1}^{n_{l_i}^{(i)}} \left(\frac{\lambda_j^{(i)} y_i}{1 - \lambda_j^{(i)} y_i}\right)^3\right), \qquad (7)$$

and y_i is the unique solution in $[0, 1/\max(\lambda_1^{(i)}, \ldots, \lambda_{l_i}^{(i)})]$ of the equation:

$$\sum_{j=1}^{l_i}\left(\frac{\lambda_j^{(i)} y_i}{1 - \lambda_j^{(i)} y_i}\right)^2 = p_i M \qquad (8)$$

Proof: In [5] the authors utilize Johansson's results [9] to show the equivalence between the latency distribution, assuming initially empty queues, and the last passage percolation problem. This theorem establishes the large-filesize, large network (or double asymptotic) analog of the Glynn-Whitt results in [8]. The general expressions for μ_i and σ_i were obtained in [7]. The extension of these results to the setting where the queues are in equilibrium appears in [12]. ■

Theorem 3.1 provides the desired analytical characterization for the latency distribution as a function of the service rates (or mean service times) via the parameter p_i. These expressions can be solved numerically and plugged in to solve the minimax problems (P1a) and (P2a).

We note that $\max E[\{t_1(p_1), \cdots, t_k(p_k)\}]$ in (P1a) is a strictly quasi-convex function (this is readily proved by noting that each $t_i(p_i)$ is a strictly increasing function in p_i and the maximum of quasi-convex functions is quasi-convex [6].) Consequently, (P1a) and (P2a) can be interpreted as the minimization of a strictly quasi-convex function on a convex set (*i.e.* $\sum_i p_i = 1$) so that the local minimum is a global minimum (proof of feasibility and existence of the unique minimum is relatively straightforward) and standard numerical optimization techniques can be used to find this optimum [6].

IV. Numerical Experiments

In what follows, we solve (P1a) by plugging in the analytical results from Theorem 3.1 and by noting that the TW_2 distribution has a mean of -1.7711 and a variance of 0.813194 [10].

Consider a simple setting where we have $k = 2$ routes and a file of $M = 200$ packets is to be transmitted with minimum mean delay. Routes 1 and 2 contain N_1 and N_2 hops with identically distributed service times with mean $\lambda_1 = 1$ and $\lambda_2 = 2$, respectively. The optimal allocation for Route 1, denoted by p_{opt} is computed by solving (P1a) using the analytical expressions for the latency as a function of the service times as described in Section III.

We will contrast this optimum with two simple heuristics (referred to as H1 and H2, respectively). The first one simply assigns packets to a route proportional to the end-to-end delay on that route. The resulting allocation to Route 1 is denoted by $p_{h,1}$ and is given by:

$$p_{h,1} = \frac{N_2\lambda_2}{N_1\lambda_1 + N_2\lambda_2} \ .$$

The second heuristic tries to balance the delay on the two routes, using only the average statistics (this would be a precise calculation if all services times are deterministic). The resulting allocation to Route 1 is denoted by $p_{h,2}$ and is obtained as follows. The total delay on Route 1 is

$$N_1\lambda_1 + (p_{h,2}M - 1)\lambda_1 \ ,$$

because it takes $N_1\lambda_1$ for the first packet to arrive, and an additional $(p_{h,2}M - 1)\lambda_1$ for the last packet to arrive assuming perfect pipelining. We can then equate the total delay on both routes and solve $p_{h,2}$:

$$N_1\lambda_1 + (p_{h,2}M - 1)\lambda_1 = N_2\lambda_2 + ((1 - p_{h,2})M - 1)\lambda_2 \ ,$$

which yields

$$p_{h,2} = \frac{(N_2 - 1)\lambda_2 - (N_1 - 1)\lambda_1 + M\lambda_2}{M(\lambda_1 + \lambda_2)} \ .$$

If the RHS of the above equation is negative – this happens if the condition $(N_1 - 1)\lambda_1 > (N_2 - 1)\lambda_2 + M\lambda_2$ holds – then the delay on the two routes cannot be balanced, and this heuristic would simply allocate all packets to Route 2. Similarly, if the RHS of the above equation exceeds 1 – this happens if the condition $(N_2 - 1)\lambda_2 - (N_1 - 1)\lambda_1 + M\lambda_2 > M(\lambda_1 + \lambda_2)$ holds – then the heuristic would allocate all packets to Route 1. Figure 2 shows the improvement in latency obtained using the new algorithm for different values of N_1 and N_2.

In this setting, the (asymptotic) latency for each route can be written explicitly. Let $p_1 = p$ so that $p_2 = 1 - p$. Then, we have that:

$$t_1(p) = \lambda_1(\sqrt{N_1} + \sqrt{pM})^2 + o(M^{1/4}),$$

and

$$t_2(p) = \lambda_2(\sqrt{N_2} + \sqrt{(1-p)M})^2 + o(M^{1/4}),$$

so that

$$p_{opt} = \max\{t_1(p), t_2(p)\}$$

subject to $p \in [0, 1]$. Let $c_1 = N_1/M$, $c_2 = N_2/M$ and $\beta = \lambda_2/\lambda_1$. After neglecting the $o(M^{1/4})$, the above optimization problem is equivalent to:

$$p_{opt} = \max\{\widetilde{t}_1(p), \widetilde{t}_2(p)\}$$

where

$$\widetilde{t}_1(p) = (\sqrt{c_1} + \sqrt{p})^2,$$

and

$$\widetilde{t}_2(p) = \beta(\sqrt{N_2} + \sqrt{(1-p)})^2.$$

Figure 3 shows the optimal allocation as a function of β and c in the setting where $c_1 = c_2 = c$. When $c \longrightarrow \infty$, $p_1 = p_{opt} = I_{\{\lambda_1 < \lambda_2\}}$. This corresponds to the setting where the $M \ll N_1, N_2$ so that the optimal allocation is to send all (the relatively few) packets on the route with the smaller mean service time. When $c \longrightarrow 0$, then $p_1 = p_{opt} = \lambda_2/(\lambda_1 + \lambda_2)$. This corresponds to the setting where $M \gg N_1, N_2$ so that a proportional throughput-maximizing allocation is also the optimum. Our algorithm yields an answer between these two extrema in the setting where $M = O(N_1), O(N_2)$ so that the filesize is on the order of the number of hops. Figure 4 shows the gains obtained in a more heterogenous setting.

V. The impact of $o(n)$ bottleneck hops

An important consequence of Theorem 3.1, via its connection with random matrix theory [12], [11] is that it allows us to examine the impact of a few bottleneck paths on the latency and consequently the optimal allocation.

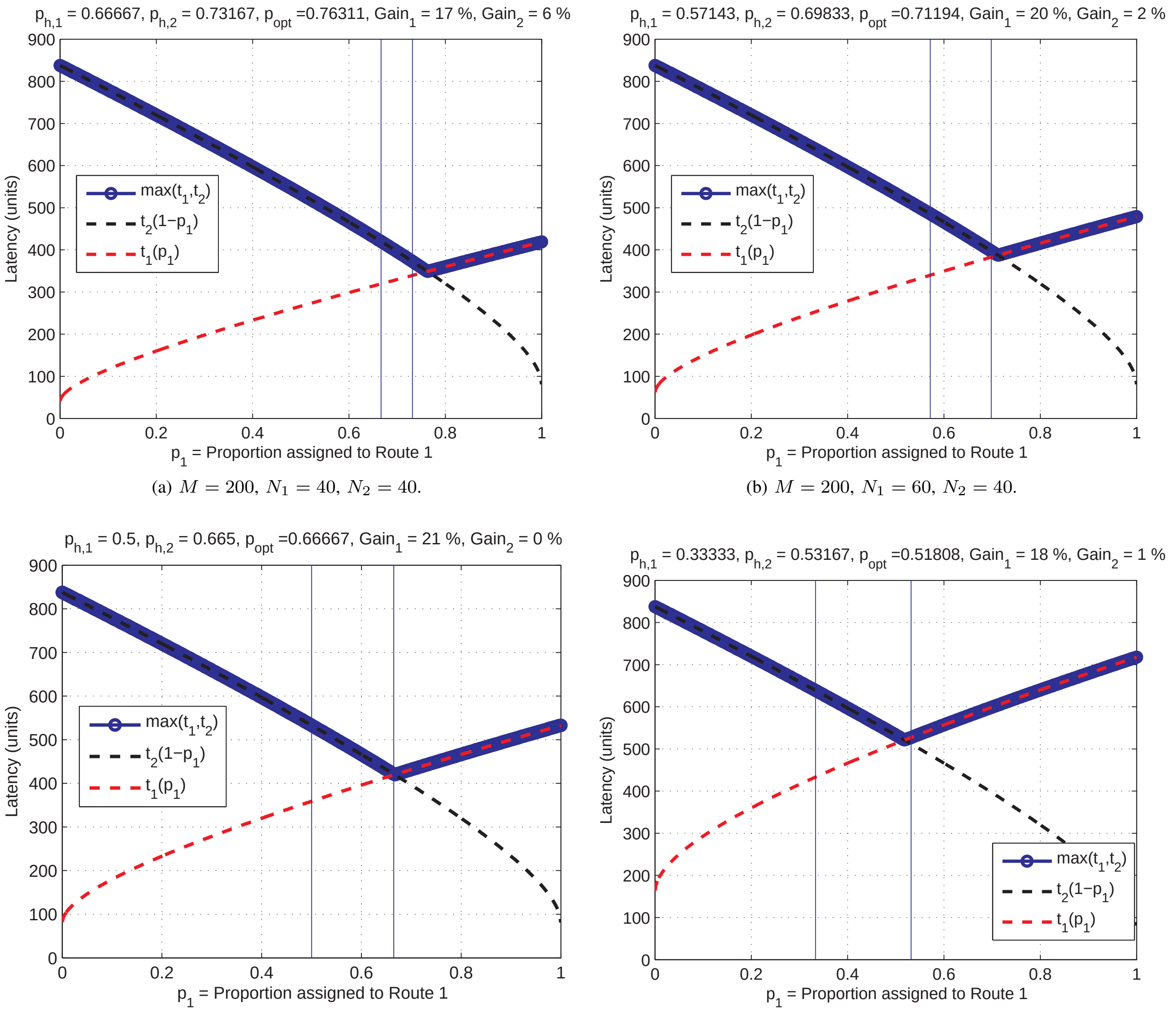
Theorem 5.1: Consider mean service times $\lambda_1, \ldots, \lambda_n$ with one bottleneck having $\lambda_n > \max\{\lambda_1, \ldots, \lambda_{n-1}\}$. Then as $n, M \to \infty$ we have that:

$$\frac{t(M,n)}{n} \xrightarrow{\text{a.s.}} \begin{cases} \dfrac{\lambda_n}{n}\left(1 + \dfrac{1}{M}\sum_{i=1}^{n-1}\dfrac{\lambda_i}{\lambda_n - \lambda_i}\right) & \text{if } \lambda_n > \tau \\ \mu(n/M) \text{ in (6)} & \text{otherwise.} \end{cases} \tag{9}$$

where τ is the solution of the equation:

$$\frac{n}{M} = \frac{1}{n}\sum_{i=1}^{n-1}\left(\frac{\lambda_i}{\tau - \lambda_i}\right)^2. \tag{10}$$

Proof: The above result exploits the connection between the latency analysis problem, the last passage percolation problem and the largest eigenvalue of complex Wishart distributed sample covariance matrices established in [5].



(a) $M = 200$, $N_1 = 40$, $N_2 = 40$.

(b) $M = 200$, $N_1 = 60$, $N_2 = 40$.

(c) $M = 200$, $N_1 = 80$, $N_2 = 40$.

(d) $M = 200$, $N_1 = 160$, $N_2 = 40$.

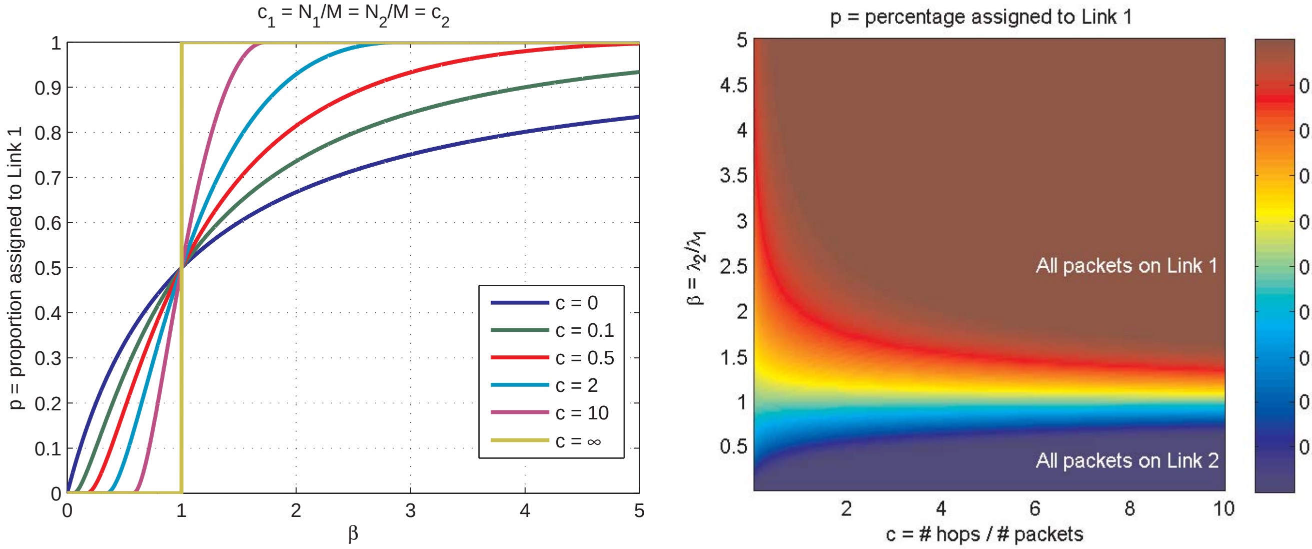
Fig. 2. Here we have $k = 2$ routes as in Figure 1 with $M = 200$ packets to be transmitted. Routes 1 and 2 contain N_1 and N_2 hops with identically distributed service times with mean $\lambda_1 = 1$ and $\lambda_2 = 2$, respectively. The optimal allocation for Route 1, denoted by p_{opt} is computed by solving (P1a) using the analytical expressions for the latency as a function of the service times as described in Section III. The optimum is contrasted with the two heuristics H1 and H2, respectively.

The explicit condition for the phase transition threshold appears in [11] where it is shown that the results hold for $o(n)$ bottlenecks as well. The distributional results appear in [7]. ■

The above result highlights the possible sub-optimality of any heuristic for (P1a) that exclusively utilizes service time statistics of a single hop without taking into account the file size and the service time parameters of the other hops. A similar result holds for the distribution as well and consequently for (P2a)- we omit the details here.

VI. Discussion and Extension

There are a number of extensions and relaxations of assumption we would like to pursue. The first is to extend this framework to the case of non-disjoint routes. In this case as different streams of packets share the same node/resources, the delay analysis on different routes becomes coupled. The analysis will also have to take into account the scheduling (or resource sharing) policy adopted by a node in transmitting packets from different streams. If this problem turns out to be tractable, then this delay analysis could be a powerful tool in finding the best set of routes.



(a) Optimal allocation to link 1.

(b) Phase diagram.

Fig. 3. We are in the same setting as in Figure 2. We let $c_1 = N_1/M$, $c_2 = N_2/M$ and $\beta = \lambda_2/\lambda_1$ and compute the optimal allocation for Route 1 for various values of β in the setting where $c_1 = c_2 = c$. When $c \longrightarrow 0$, $p_1 \longrightarrow \beta/(1+\beta)$; when $c \longrightarrow \infty$, $p_1 = I_{\{\beta<1\}}$.

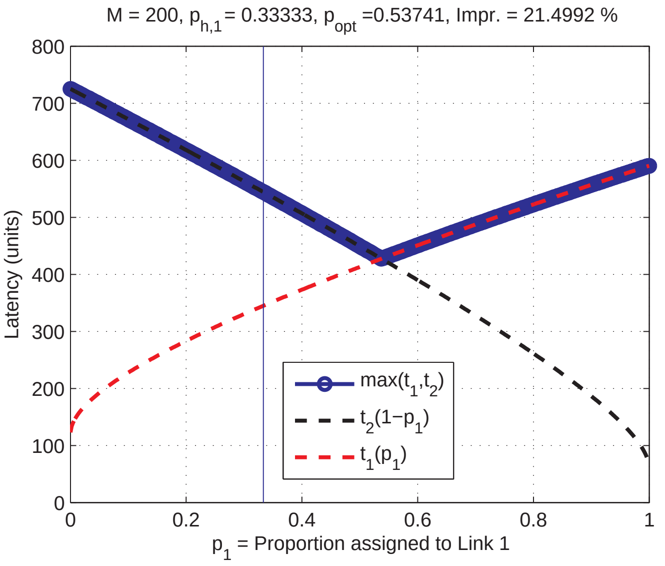


Fig. 4. Here we have $k = 2$ routes as in Figure 1 with $M = 200$ packets to be transmitted. Routes 1 and 2 contain $N_1 = 160$ and $N_2 = 40$ hops with service times given by the vectors $\lambda_1 = [1, ..., 1, 0.5, \cdots, 0.5]$ and $\lambda_2 = [2, ..., 2, 1, \cdots, 1]$, respectively. The optimal allocation for Route 1, denoted by p_{opt} is computed by solving (P1a) using the analytical expressions for the latency as a function of the service times as described in Section III. The optimum is contrasted with the total-route-delay based allocation $p_{h,1} = N_2\lambda_2/(N_1\lambda_1 + N_2\lambda_2)$.

A second extension is to relax the assumption that service times are independent across different nodes. In practice there are a number of reasons why service times may not be independent, even if all routes are disjoint and all packets are of unit size. A prime example is a wireless scenario, where different node within the same interference domain compete for channel access, and thus their services times are all coupled. In general interference relationship can be very complicated and sometimes form a chain: e.g., A may interfere with B, who interferes with C, but A and C do not interfere, etc. This relationship extends to the streams as well, i.e., stream 1 may interfere with stream 2, which interferes with stream 3, but 1 and 3 do not interfere, and so on. Also note that two streams can have multiple points along their path where nodes interfere. In such cases the delay analysis may need to take into account not only the topology (interference relationship among routes) but also the MAC layer channel access and transmission scheduling policies.

Interesting questions to be asked include (1) whether it's possible to obtain delay as a function of these scheduling policies, and (2) when the file size is large and the routes are long, to what extend do these scheduling policies really matter.

References

[1] D. P. Bertsekas and R. Gallager, *Data Networks*, Prentice-Hall, Inc., New Jersey, 1992.

[2] J. A. Morrison, "Two discrete-time queues in tandem," *IEEE Transactions on Communications*, vol. COM-27, no 3, March 1979.

[3] M. Neely, "Exact queueing analysis of discrete time tandems with arbitrary arrival processes," *ICC* 2004.

[4] J. Baik, P. Deift, and K. Johansson, "On the distribution of the length of the longest increasing subsequence of random permutations," *J. of Amer. Math. Soc.* 12(4), pp. 1119–1178, 1999.

[5] J. Baik, G. Ben Arous, S. Péché, "Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices," *Annals of Probability*, pp. 1643-1697, 2005.

[6] S. P. Boyd and L. Vandenberghe, *Convex optimization*, 2004.

[7] N. El Karoui, "Tracy-Widom limit for the largest eigenvalue of a large class of complex sample covariance matrices," *Annals of Probability* 35(2):663-714, 2007.

[8] P. W. Glynn and W. Whitt, “Departures from many queues in series,” *The Annals of Applied Probability*, 1(4):546-572, 1991.

[9] K. Johansson, “Shape fluctuations and random matrices,” *Comm. in Math. Phys.*, 209(2):437-476, 2000.

[10] I. M. Johnstone, “On the distribution of the largest eigenvalue in principal components analysis,” *Annals of Statistics*, pages 295–327, 2001.

[11] R. R. Nadakuditi, J. W. Silverstein, “Fundamental limit of sample generalized eigenvalue based detection of signals in noise using relatively few signal-bearing and noise-only samples,” *J. Sel. Topics in Signal Proc.*, `http://arxiv.org/abs/0902.4250`, 2009.

[12] R. R. Nadakuditi, J. Baik, “End-to-end transient delay analysis in large tandem networks”, In preparation.