Path Delays in Communication Networks

IZHAK RUBIN*

Department of System Science
University of California
Los Angeles, California

ABSTRACT

A communication network is modelled by a weighted graph. The vertices of the graph represent stations with storage capabilities, while the edges of the graph represent communication channels (or other information processing media). Channel capacity weights are assigned to the edges of the network. The network is assumed to operate in a store-and-forward manner, so that when a channel is busy the messages directed into it are stored at the station, joining there a queue which is governed by a first-come first-served service discipline.

Assuming messages, with fixed length, to arrive at random at the network, following the statistics of a Poisson point process, we calculate the statistical characteristics of the message time-delays along a path in a communication network. We solve for the steady-state distributions of the message waiting-times along the path, for the distribution of the overall message delay-time, for the average memory size requirements at the stations, as well as for other statistical characteristics of the message flow and the queueing processes along a communication path.

1. Introduction. An information transmission (or processing) system which constitues of a network of channels and stations is called a communication network. Topologically, such a system is represented by a weighted graph $G = (V, \Gamma, W)$. The set of vertices V of the graph represent the stations, while the channels are generally represented by the set of edges Γ. Appropriate weights are assigned from the set W to the edges and vertices of the graph.

A large variety of information transmission (and processing) networks can be described by the above model. In a satellite communication system, the stations (vertices) of the network (graph) represent satellites, ground stations or air-borne stations; and these stations are interconnected by communication channels (edges). The weighting function associated with a satellite communication network assigns appropriate weights to the channels (like channel capacities, noise-characteristics, etc.) as well as weights to the stations (information processing capabilities, power

* This work was supported by the Office of Naval Research under Grant No. N00014-69-A-0200-4041.

APPLIED MATHEMATICS & OPTIMIZATION, Vol. 1, No. 3

limitations, etc.). A similar model is utilized to describe a telephone, telegraph, space-communication or a computer communication network. In the latter network, the stations represent the users' or the computers' processing units and communication channels inter-connect the various users and computing facilities. In certain situations, one may want to associate the communication channels with the vertices of the graph and the stations with its edges (as in cases where time delays in the network are evaluated and the main time delay involved is associated with information processing in the station, like an encoding or decoding procedure, rather than along the channel). Similar models are of considerable importance in many other areas, such as transportation, biology, management and operations research.

Communication networks are generally considered, as assumed here, to operate in a "store and forward" fashion. A message arriving at a station will be directed into the outgoing appropriate channel, following the system's routing policy, and be transmitted over this channel if it is free for transmission. If the latter channel is busy the message will be stored at the station, and join a queue of messages which is assumed to be governed by a first-come first-served service discipline.

The two major considerations when analyzing and synthesizing a communication network, are those of congestion and time-delays in the network and those of reliable transmission of information over the noisy channels. The latter considerations require one to develop appropriate channel and source coding procedures to decrease the overall probability of error in message transmission through the network and minimize a distortion measure, respectively. Such problems, associated with simple communication networks, have recently been considered by information theoretists, and are starting to draw considerable attention. See [5] for channel coding for broadcast channels (modelled by a network with one transmitting source and multiple receivers), and [6] for discussions concerning coding subject to a fidelity criterion for some simple networks under special noise characteristics.

When considering a communication network for which appropriate coding techniques have already been applied to combat noise interferences (for example, by conventional coding methods, as is the case for the existing computer and satellite communication networks), so that one can view the confronting network to be noiseless, the remaining only major consideration is that of congestion and time-delays. The latter is the subject of the present paper.

We assume that messages, of constant length, arrive at a station of a communication network at random times, governed by the statistics of a Poisson point process. Each channel (edge) in the network (graph) is assigned a capacity. Considering an arbitrary path in the network, which leads the messages from their origin to their destination through the network, we are interested to obtain the time delays experienced by a message over this path. Using results and methods from queueing theory, we derive the (steady-state) distribution functions for the message waiting times at the channels, and for the overall waiting and delay times, and obtain, as well, the average storage requirements at the various stations. Using our results, one can proceed, using time-delay considerations, with analysis and synthesis of communication networks, when non-simultaneous flows in paths are allowed between the various stations of the communication network.

Approximate (limiting) average time-delay expressions for communication networks have been obtained in [7], and applied to computer communication networks (see [8] and the references therein). In the latter analysis, message lengths are assumed to be exponentially distributed, so that Burke's theorem for an M/M/1 queueing system could be invoked to conclude that the message departure process at the first channel is (at steady-state) a Poisson point process, as that at the input. Then, to avoid statistical dependencies, an "independence assumption" is made (see [7] p. 50), which amounts to choosing the message length at random at each of the channels in the network. For large and (topologically) complexed networks, the latter results have been observed (by simulation) in [7] to be good approximations. In our situation when path delays are sought (and constant message lengths are assumed, as is the case, for example, in the present computer communication networks which utilize fixed length packets of messages), the latter approximation cannot be made, and we need to perform an exact analysis as presented by this paper. Related time-delay problems have been considered in association with queueing networks [9], queues in tandem [10] and a variety of computer processing systems. However, all the latter studies make the above-mentioned independence assumption (i.e., taking the service duration at any queueing channel statistically independent of that in any other channel), which cannot be made in our problem, since a message carrying a fixed contents of information is transferred through the network. For an approach to the analysis of queueing networks using the diffusion approximation, see [11]. For studies of networks with deterministic channel delays see [12], Chapter 9.

The Mathematical model of the communication network is presented in Section 2. Time delays over a single channel path follow directly by using results from queueing theory, and are given in Section 3. We solve then, in Section 4, for the waiting time distributions, as well as other statistical characteristics, for a two-channel path. The latter results are extended to an n-channel path in Section 5. In Section 6 we derive the distribution of the overall path delay, as well as conclude a capacity invariance property. The latter property asserts that the overall path delay is independent of the order of the various channel capacities, and allows one to subsequently obtain the average time-delay at each channel and the path delay distribution. After this paper has been written, we have noticed that the results of Section 6 have also been obtained, in an operations-research context, in [13]–[14] which study only *overall* waiting times in tandem queues with constant service times. In Section 7, we present a simple illustrating example.

2. Preliminaries. ***The Communication Path.*** A communication path, as shown in Fig. 1(a), is considered. The branch (v_i, v_{i+1}) between vertices v_i and v_{i+1}, $i = 1, 2, \ldots, n$, represents the i-th channel, whose capacity is equal to C_i [bits/sec]. Thus, each bit is delayed $1/C_i$ seconds by the i-th channel. The vertices v_i, $i = 1, 2, \ldots, n+1$, represent stations or buffer systems equipped with memory storage units or queueing facilities. No restrictions are imposed on the size of the storage unit. (It will be clear from the results of the analysis how to control the memory sizes.) Messages arrive, at the input station v_1, randomly in time according to a Poisson point process with intensity λ[messages/sec]. Each message (or packet of information) is assumed to be of a constant length of β[bits/message]. If the arriving message at v_1 finds channel 1 free, it is immediately transmitted to station v_2. The message transmission time-delay over channel 1 is clearly

$$a_1 = \beta/C_1 \quad \text{[sec/message]}.$$

The output of channel 1 is stored at v_2 until the remaining of the message arrives. The latter message is stored at v_2 until channel 2 is free and then it is transmitted to v_3. And so on until the message leaves the system at v_{n+1}. If the arriving message at v_1 finds channel 1 busy (i.e., another message is being transmitted over this channel at its time of arrival), it will be stored at v_1 and wait until channel 1 is free and then will be transmitted to v_2, and so on. It is assumed that a first-come first-served discipline is employed to serve the waiting messages at v_i, $i = 1, 2, \ldots, n$.

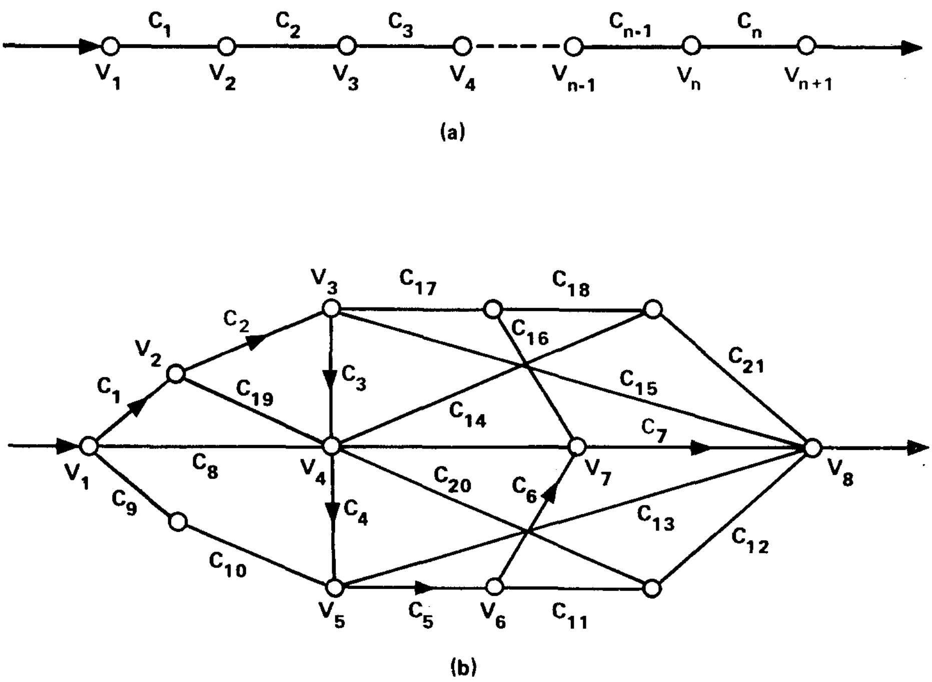


Figure 1. (a) A Communication Path.
(b) A Path in a Communication Network.

In this paper, we derive the steady-state distributions for the waiting times and storage requirements at the individual stations, and for the time-delay of a message over the n channel path.

We also note that β bits out of the storage of v_i, $i = 2, 3, \ldots, n$, can be associated with the facilities of channel $i - 1$ for the purpose of collecting the message departing from channel $i - 1$. The latter message is then immediately transferred to the remaining storage facility of v_i. A storage of β bits is required at v_{n+1} since a departure is declared only when all the message is transmitted.

The Mathematical Model. Each channel i with its storage facilities can be considered as a queueing system. For that purpose, consider a message to be transmitted over channel i, to be a customer which requires service from server i. The service time required by the customer (message) from the server (channel) i is equal to his transmission time over the channel and is clearly given by

$$a_i = \beta / C_i \tag{1}$$

seconds. Upon leaving the i-th queueing system, a customer enters the $(i+1)$-st queueing system, where he is served on a first-come first-served basis, and requires service time of a_{i+1} seconds. Customers (messages) arrive at v_1 according to a Poisson point process with intensity λ, and depart from the path at v_{n+1}. Our problem is thus equivalent to studying a system of queues in tandem, with constant service times. We will employ a queueing theoretical approach.

The following notations will be utilized throughout the paper. We let $X_t^{(i)}$ denote the number of messages stored at v_i or being transmitted from v_i (through channel i). Thus, $\{X_t^{(i)}, t \geq 0\}$ is the queueing process associated with channel i. Assume $X_0^{(i)} = 0, i = 1, \ldots, n$. The (random) instants of arrival of message at v_i are denoted by $\{t_n^{(i)}, n = 1, 2, 3, \ldots\}, t_0^{(i)} \triangleq 0$, and the departure stochastic point process from channel i is denoted as $\{r_n^{(i)}, n = 1, 2, \ldots\}$, where $t_n^{(i)}$ and $r_n^{(i)}$ denote, for v_i and channel i, the instants of the n-th arrival and departure, respectively. Clearly, for a communication path $t_n^{(i+1)} = r_n^{(i)}$. The waiting time at v_i of the n-th arriving message is denoted as $W_n^{(i)}$. The virtual waiting time $W_t^{(i)}$ is the waiting time to be assumed at v_i by a (hypothetical) message which arrives at t. Clearly, $W_n^{(i)} = W_{t_n^{(i)}}^{(i)}$.

Observing the evolution of the queueing process $\{X_t^{(i)}, t \geq 0\}$, we find that the process passes successively through idle and busy periods. We denote, for channel i, the sequence of idle periods (See Fig. 2) by $\{I_n^{(i)}, n = 0, 1, 2, \ldots\}$, and the sequence of busy periods by $\{B_n^{(i)}, n = 1, 2, \ldots\}$. Thus, for $\{X_t^{(i)}, t \geq 0\}$, $I_n^{(i)}$ is the duration of the $(n+1)$-st idle period (and during it $X_t^{(i)} = 0$), and $B_n^{(i)}$ is the duration of the n-th busy period (and during it $X_t^{(i)} > 0$). The number of messages transmitted through channel i during the n-th busy period (whose length is $B_n^{(i)}$) is denoted as $N_n^{(i)}$. The delay of the n-th message at channel i, $\gamma_n^{(i)}$, is defined as the total of its waiting time and transmission delay time at channel i. Thus, we have

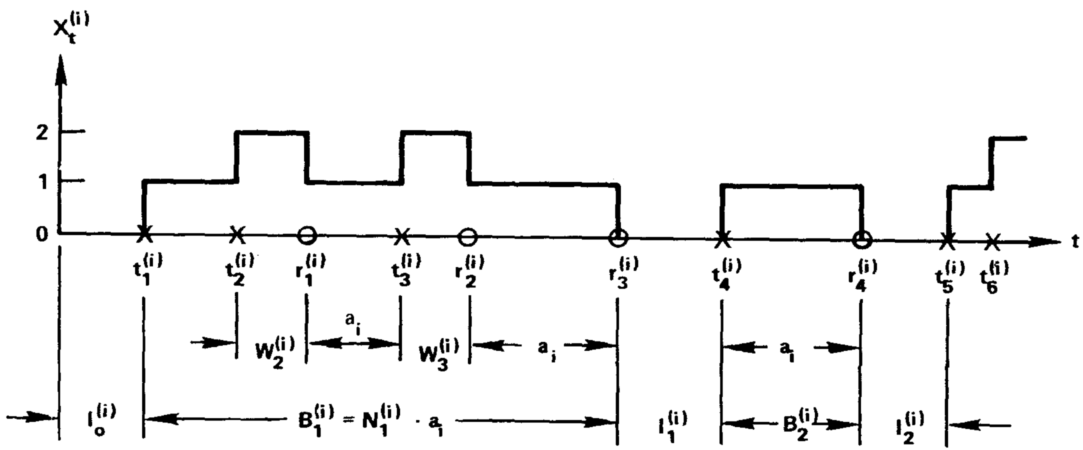
$$\gamma_n^{(i)} = W_n^{(i)} + a_i. \tag{2}$$

The overall time delay γ_m of the m-th arriving message through the n channel communication path is thus given by

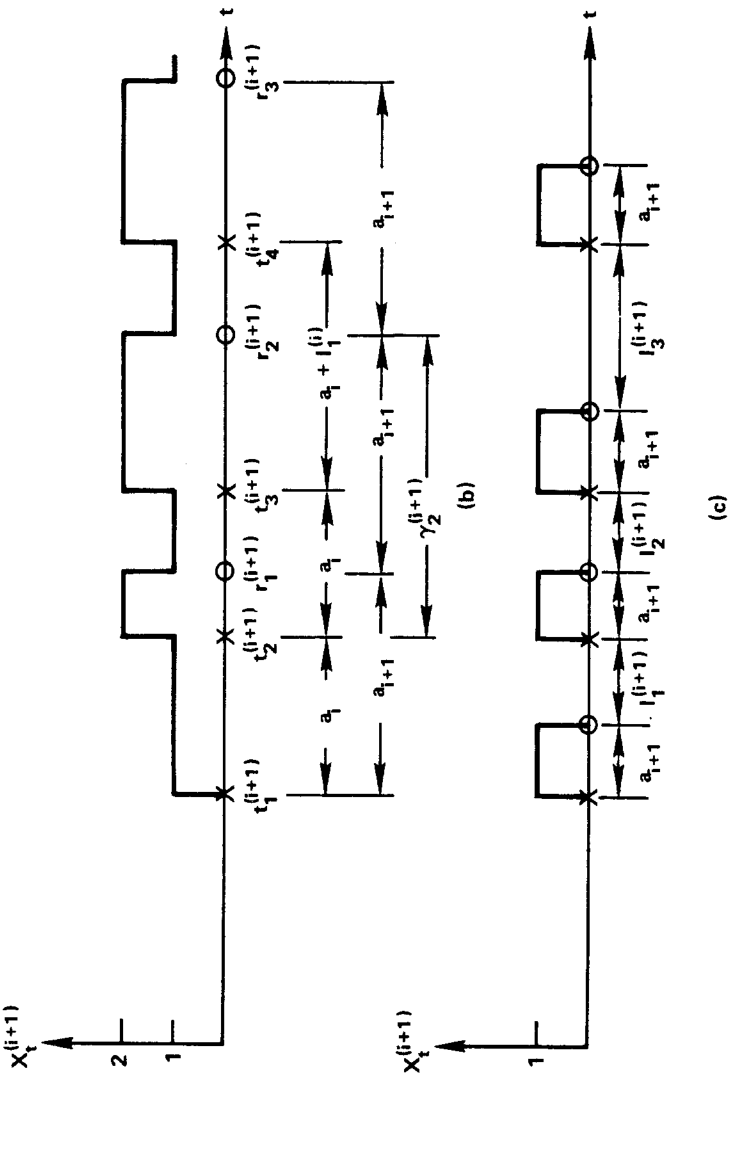
$$\gamma_m = \sum_{i=1}^{n} \gamma_m^{(i)}. \tag{3}$$

Thus, we are interested in obtaining the steady-state average message waiting times at the individual stations $\overline{W}^{(i)} = \lim_{m \to \infty} E\{W_m^{(i)}\}$, $i = 1, 2, \ldots, n$. Using these expressions we would calculate the average overall message delay time γ, $\gamma = \lim_{n \to \infty} E\{\gamma_n\}$, as well as the average required storage capability at v_i (which is related to $\lim_{t \to \infty} E\{X_t^{(i)}\}$). Moreover, we wish to derive the limiting distributions

Figure 2. Queueing Process Realizations
(a) Over the ith Channel.
(b) Over the $(i+1)$st Channel with $a_{i+1} > a_i$.
(c) Over the $(i+1)$st Channel with $a_{i+1} < a_i$.



(a)



$X_t^{(i+1)}$
2
1
t
$t_1^{(i+1)}$
$t_2^{(i+1)}$
$r_1^{(i+1)}$
$t_3^{(i+1)}$
$r_2^{(i+1)}$
$t_4^{(i+1)}$
$r_3^{(i+1)}$
a_i
a_i
$a_i + l_1^{(i)}$
a_{i+1}
a_{i+1}
a_{i+1}
$\gamma_2^{(i+1)}$
(b)
$X_t^{(i+1)}$
1
t
a_{i+1}
$l_1^{(i+1)}$
a_{i+1}
$l_2^{(i+1)}$
a_{i+1}
$l_3^{(i+1)}$
a_{i+1}
(c)

$\lim_{n\to\infty} P(W_n^{(i)} \le x)$ and $\lim_{n\to\infty} P(\gamma_n \le x)$ for the waiting-times and the overall delay respectively.

3. **The Single Channel Path.** Consider the case $n = 1$, so that the path includes only a single channel. In this case we are thus considering a queueing system into which customers (messages) arrive according to a Poisson process with intensity λ, and where each customer (message) requires a deterministic service of length a_1. Such a system is denoted in queueing theory as an M/D/1 system.* Results for this queueing system are readily available (see [2] pp. 32–38, [1] Chap. II.4). In particular, assuming the traffic intensity $\rho_1 = \lambda a_1 < 1$, so that $\{X_t^{(1)}, t \ge 0\}$ is ergodic, we have

$$\lim_{n\to\infty} P(W_n^{(1)} = 0) = \lim_{t\to\infty} P(W_t^{(1)} = 0) = 1 - \rho_1, \tag{4}$$

so that the traffic intensity ρ_1 expresses the limiting probability of channel 1 being busy. The limiting waiting time distribution is given by ([2] p. 35)

$$\begin{aligned} F_{W^{(1)}}(x) &= \lim_{n\to\infty} P\{W_n^{(1)} \le x\} \\ &= (1-\rho_1) \sum_{i=0}^{n} \exp[\lambda(x - ia_1)] \,.\, (i!)^{-1} \,.\, [-\lambda(x - ia_1)]^i, \end{aligned} \tag{5}$$

for $x = na_1 + t$, $n \ge 0$, $0 \le t < a_1$. In particular, the limiting average waiting time is then given by

$$W^{(1)} = \lim_{n\to\infty} E(W_n^{(1)}) = \frac{1}{2}\frac{\rho_1}{1-\rho_1} a_1, \tag{6}$$

so that the delay time is given as

$$\gamma^{(1)} = \lim_{n\to\infty} E(\gamma_n^{(1)}) = \overline{W}^{(1)} + a_1 = \frac{1-\rho_1/2}{1-\rho_1} a_1. \tag{7}$$

For the average queue size we have (see [1] p. 247, or just by Little's theorem [3] the average queue size is equal to $\lambda\gamma^{(1)}$),

$$\lim_{t\to\infty} E\{X_t^{(1)}\} = \frac{1-\rho_1/2}{1-\rho_1} \rho_1. \tag{8}$$

Since every message contains β bits, the average memory size required at v_1, to be denoted as $M^{(1)}$, is given by $\lim_{t\to\infty} E\{X_t^{(1)}\} \,.\, \beta$, so that we have

$$M^{(1)} = \frac{1-\rho_1/2}{1-\rho_1} \,.\, \beta\rho_1 \text{ bits}. \tag{9}$$

---

* The notation A/B/s is commonly used to designate a queueing system with s servers, for which the arrival renewal process has inter-arrival distribution A(·) and the customers' service times are i.i.d. random variables, statistically independent of the arrival process, governed by the service distribution B(·). The letters D, M and G are used to denote deterministic, exponential and general distributions, respectively. Thus, an M/D/1 queueing system is a 1-server system with Poisson arrivals and deterministic service times.

Of particular interest for the following analysis are the distributions of the idle and busy periods of $\{X_t^{(1)}, t \geq 0\}$. Since the arrival process is Poisson with intensity λ, the idle period is exponentially distributed with parameter λ. Thus, the distribution of any idle period $I_n^{(1)}$ is given by

$$P\{I_n^{(1)} \leq x\} = [1 - \exp(-\lambda x)]U(x), \tag{10}$$

where $U(x)$ is the unit step function, $U(x) = 1$ for $x > 0$ and $U(x) = 0$ for $x \leq 0$. The distribution of the number of customers served during a busy period, $N_m^{(1)}$, is obtained to be (see [2] p. 36)

$$P\{N_m^{(1)} = n\} = (n!)^{-1} (\lambda n a_1)^{n-1} \,.\, \exp(-\lambda n a_1), n \geq 1, \tag{11}$$

and the first two moments are obtained to be given as

$$E\{N_m^{(1)}\} = (1-\rho_1)^{-1} \tag{12a}$$

$$E\{[N_m^{(1)}]^2\} = (1-\rho_1)^{-3}. \tag{12b}$$

We note that $\{I_n^{(1)}, n \geq 0\}$ and $\{N_n^{(1)}, n \geq 1\}$ are stochastically independent sequences of i.i.d. random variables with distributions given by equations (10) and (11), respectively. Thus, the point process $\{{}_0t_n^{(1)}, n \geq 1\}$, where ${}_0t_n^{(1)} = \sum_{i=1}^{n} [I_i^{(1)} + N_i^{(1)} a_1]$, $n \geq 1$, ${}_0t_0^{(1)} = t_1^{(1)}$, is a renewal point process (i.e., ${}_0t_{n+1}^{(1)} - {}_0t_n^{(1)}$ are i.i.d. random-variables) and ${}_0t_n^{(1)}$ is the instant of arrival which initiates the $(n+1)$-st busy period.

4. The Two-Channel Path. We assume now that $n = 2$, so that a two-channel path is considered. The distribution and mean of the waiting time at v_1, $W_n^{(1)}$, are given by (5) and (6), respectively. The distribution of the waiting time at v_2, $W_n^{(2)}$, need now be derived. We observe that

$$t_n^{(2)} = r_n^{(1)}, n \geq 1, \tag{13}$$

so that the instant of n-th message arrival at v_2 is equal to the time of the n-th message departure from channel 1. The waiting time $W_n^{(2)}$ satisfies the recursive relationship.

$$W_{n+1}^{(2)} = [W_n^{(2)} + a_2 - T_{n+1}^{(2)}]^+, n \geq 1, \tag{14}$$

where $[X]^+ \triangleq \max(0, X)$, $W_1^{(2)} = 0$, and $T_{n+1}^{(2)} = t_{n+1}^{(2)} - t_n^{(2)}$ is the inter-arrival duration at v_2. However, by (13) we readily obtain

$$T_{n+1}^{(2)} = r_{n+1}^{(1)} - r_n^{(1)} = \begin{cases} a_1 & , \text{ if } W_{n+1}^{(1)} > 0 \\ a_1 + I_{(n+1)}^{(1)}, & \text{if } W_{n+1}^{(1)} = 0, \end{cases} \tag{15}$$

where $I_{(n+1)}^{(1)}$ denotes the duration of the idle period preceding $t_{n+1}^{(1)}$, for channel 1. Equation (15) follows the observation that during a busy period the inter-departure time is equal to the service time (which is a_1 for channel 1), while the inter-departure time between the last message in a busy period and the next message is given by the sum of the service time and the associated idle period.

From Eq. (15) one concludes that the arrival process at v_2 is not a renewal process, so that the queueing process $\{X_t^{(2)}, t \geq 0\}$ is not associated with a GI/D/1 queueing system. One also readily observes from Eqs. (14) and (15) that $T_{n+1}^{(2)}$ is not statistically independent of $W_n^{(2)}$ (since $W_n^{(2)}$ depends on $T_n^{(2)}$ which is not statistically independent of $T_{n+1}^{(2)}$), and subsequently one cannot employ here the available queueing theoretical techniques for solving for the limiting distribution of $W_n^{(2)}$ (using Lindley's Integral Equation method, see [1] Section II.6.3., or results from fluctuation theory and Spitzer's identity, see [1] Section I.6.6). To resolve these problems we develop here a new approach.

First we observe from (15) that $t_{n+1}^{(2)}-t_n^{(2)} \geq a_1$, so that if $a_2 \leq a_1$ the service time required by the messages arriving at v_2 is not larger than their inter-arrival times, and subsequently no waiting is required. We have thus observed the following property.

Proposition 1. For $a_2 \leq a_1$, we have

$$W_n^{(2)} = 0,$$

for each $n \geq 1$.

We thus consider henceforth only the case $a_2 > a_1$. We then obtain that the n-th message arrival at v_2 can find channel 2 free only if the n-th arrival at v_1 has found channel 1 free. In turn, if $\{W_n^{(1)} > 0\}$ then also $\{W_n^{(2)} > 0\}$. It is however possible that $W_n^{(1)} = 0$ while $W_n^{(2)} > 0$. This is indicated by the following lemma.

Lemma 1. For $a_2 > a_1$, event $\{W_n^{(2)} = 0\}$ can occur only if $\{W_n^{(1)} = 0\}$ has occurred.

Proof. By Eqs. (14) and (15), if $W_{n+1}^{(1)} > 0$ we have $W_{n+1}^{(2)} = [W_n^{(2)}+a_2-a_1]^+ = W_n^{(2)}+a_2-a_1 > 0$, since $W_n^{(2)} \geq 0$. Hence $\{W_n^{(1)} > 0\}$ implies $\{W_n^{(2)} > 0\}$. Subsequently, event $\{W_n^{(2)} = 0\}$ can occur only if $\{W_n^{(1)} = 0\}$. Q.E.D.

From Lemma 1 we conclude that the number of messages $N_n^{(2)}$ served (transmitted) during a busy-period at channel 2 is the sum of a certain number of variables $N_n^{(1)}$, each representing a number of messages served during a busy-period by channel 1, as expressed by the following.

Lemma 2. For $a_2 > a_1$, we have

$$N_n^{(2)} = \sum_{i=1}^{h_n^{(2)}} N_{k+1}^{(1)}, \tag{16}$$

for some integer $k \geq 0$, where $h_n^{(2)}$ is a random-variable, defined over the space of positive integers, which denotes the number of channel 1 busy-periods included in the n-th busy-period of channel 2. Both $\{N_n^{(2)}, n \geq 1\}$ and $\{h_n^{(2)}, n \geq 1\}$ are sequences of i.i.d. random-variables, and $h_1^{(2)}$ is given by

$$\begin{aligned} h_1^{(2)} &\triangleq \min \{n | n \geq 1, W^{(2)}_{\Sigma_{i=1}^{k} N_i^{(1)}+1} > 0, k=1,2,\ldots,n-1, W^{(2)}_{\Sigma_{i=1}^{n} N_i^{(1)}+1} = 0\} \\ &= \min \{n | n \geq 1, \sum_{i=1}^{k} [N_i^{(1)}(a_2-a_1)-I_i^{(1)}] > 0, k=1,2,\ldots,n-1, \\ &\qquad \sum_{i=1}^{n} [N_i^{(1)}(a_2-a_1)-I_i^{(1)}] \leq 0\}. \end{aligned} \tag{17}$$

Proof. By Lemma 1, $N_1^{(2)}$ must include all the messages served in $B_1^{(1)}$, whose number is $N_1^{(1)}$. If $W^{(2)}_{N_1^{(1)}+1} > 0$, by Lemma 1 the following $N_2^{(1)}$ messages are also served during $B_1^{(2)} = \sum_{i=1}^{h_1^{(1)}} N_i^{(2)}$, where $h_1^{(2)}$ is defined as expressed by the first line of Eq. (17). By Eqs. (14) and (15), it follows that for $k \leq h_1^{(2)}$ we have

$$W^{(2)}_{\sum_{i=1}^{k} N_i^{(1)}+1} = \left[\sum_{i=1}^{k}[N_i^{(1)}(a_2-a_1)-I_i^{(1)}]\right]^+, \tag{18}$$

which implies the second expression in Eq. (17). For $n \neq m$, $h_n^{(2)}$ and $h_m^{(2)}$ are determined by expressions similar to (17) involving different sets of variables $\{N_i^{(1)}\}$ and $\{I_i^{(1)}\}$. Since the latter are i.i.d. random-variables, we conclude that $\{h_n^{(2)}, n \geq 1\}$ is also a sequence of i.i.d. random-variables. Similarly for the sequence $\{N_n^{(2)}, n \geq 1\}$, which is determined by (16). Q.E.D.

We next define the random variable $\eta_n^{(i)}$ to be

$$\eta_n^{(i)} = n - \max\{m|0 \leq m \leq n,\ W_m^{(i)} = 0\}. \tag{19}$$

Thus, $\eta_n^{(i)}$ denotes the number of message arrivals at channel i, prior (and including) the n-th arrival and following the most recent arrival which has found the channel free. Denoting the instants of arrivals which has found channel i free by $\{{}_0t_n^{(i)}, n \geq 0\}$, ${}_0t_0^{(i)} = t_1^{(i)}$, we have that

$${}_0t_n^{(j)} = \sum_{i=1}^{n}[I_i^{(j)}+N_i^{(j)}a_j],\ n \geq 1. \tag{20}$$

In the interval $[{}_0t_n^{(i)}, {}_0t_{n+1}^{(i)})$, $N_{n+1}^{(i)}$ messages arrive at channel i. The variable $\eta_n^{(i)}$ is thus the backward-recurrent-number associated with the discrete-time renewal process whose sequence of intervals is $\{N_n^{(i)}, n > 1\}$. Considering now the two channel path, we can relate the waiting time of the n-th message at channel 2 to the waiting time of the n-$\eta_n^{(1)}$ message at channel 2, the latter being the recent message prior to the n-th message which had a zero waiting time at channel 1. The result is given by the following Lemma.

Lemma 3.

$$W_n^{(2)} = W^{(2)}_{n-\eta_n^{(1)}} + \eta_n^{(1)}[a_2-a_1]^+. \tag{21}$$

Proof. By Proposition 1, for $a_2 \leq a_1$, $W_n^{(2)} = 0\ \forall\, n$, so that Eq. (21) holds. We need thus consider only the case $a_2 > a_1$, for which $[a_2-a_1]^+ = a_2-a_1$. By Eqs. (14) and (15), proceeding recursively backward from $W_n^{(2)}$ to $W^{(2)}_{n-\eta_n^{(2)}} = 0$, we obtain

$$W_n^{(2)} = [\eta_n^{(2)}(a_2-a_1)-I^{(1)}(n-\eta_n^{(2)}, n)]^+, \tag{22}$$

where $I^{(i)}(m, n)$ denotes the duration of the idle period of channel i in the interval $[t_m^{(i)}, t_n^{(i)})$, $m < n$. Eq. (22) thus expresses the waiting time of the n-th message at channel 2 in terms of the backward recurrent number $\eta_n^{(2)}$ and the idle duration $I^{(1)}(n-\eta_n^{(2)}, n)$ at channel 1. Using Eq. (22), we have for $W^{(2)}_{n-\eta_n^{(1)}}$,

$$W^{(2)}_{n-\eta_n^{(1)}} = [\eta^{(2)}_{n-\eta_n^{(1)}}(a_2-a_1) - I^{(1)}(n-\eta_n^{(1)}-\eta^{(2)}_{n-\eta_n^{(1)}}, n-\eta_n^{(1)})]^+. \tag{23}$$

However, by Lemma 1 we conclude that

$$\eta^{(2)}_{n-\eta_n^{(1)}} = \eta_n^{(2)} - \eta_n^{(1)}, \tag{24}$$

since the event $\{W_m^{(2)} = 0\}$ cannot occur when $t_m^{(2)} \varepsilon\, (t^{(1)}_{n-\eta_n^{(1)}}, t_n^{(1)}]$ when $\eta_n^{(1)} \geq 1$. We also have

$$I^{(1)}(n-\eta_n^{(1)}-\eta^{(2)}_{n-\eta_n^{(1)}}, n-\eta_n^{(1)}) = I^{(1)}(n-\eta_n^{(2)}, n-\eta_n^{(1)}) = I^{(1)}(n-\eta_n^{(2)}, n), \tag{25}$$

where the first equality in (25) follows by using (24), and the second equality there results by observing that no channel 1 idle time can occur during $(t^{(1)}_{n-\eta_n^{(1)}}, t_n^{(1)}]$ so that $I^{(1)}(n-\eta_n^{(1)}, n) = 0$. We clearly let $I^{(1)}(n, n) \equiv 0$. Incorporating Eqs. (24) and (25) in (23) and using Eq. (22) we obtain

$$\begin{aligned} W_n^{(2)} - W^{(2)}_{n-\eta_n^{(1)}} = {} & [\eta_n^{(2)}(a_2-a_1) - I^{(1)}(n-\eta_n^{(2)}, n)]^+ \\ & - [(\eta_n^{(2)}-\eta_n^{(1)})(a_2-a_1) - I^{(1)}(n-\eta_n^{(2)}, n)]^+. \end{aligned} \tag{26}$$

Now if $W^{(2)}_{n-\eta_n^{(1)}} > 0$, then $W_n^{(2)} > 0$ and both terms on the RHS of Eq. (26) are positive. Subsequently, we obtain

$$W_n^{(2)} - W^{(2)}_{n-\eta_n^{(1)}} = \eta_n^{(1)}(a_2-a_1),$$

as claimed by Eq. (21). On the other hand, if $W^{(2)}_{n-\eta_n^{(1)}} = 0$, we obtain by Eq. (26), or Eq. (22), $W_n^{(2)} - W^{(2)}_{n-\eta_n^{(1)}} = [\eta_n^{(2)}(a_2-a_1) - I^{(1)}(n-\eta_n^{(2)}, n)]^+$. However, $W^{(2)}_{n-\eta_n^{(1)}} = 0$ implies that $\eta_n^{(2)} \leq \eta_n^{(1)}$. Since, by Lemma 1 we always have $\eta_n^{(2)} \geq \eta_n^{(1)}$, we conclude that

$$\eta_n^{(2)} = \eta_n^{(1)}, \tag{27}$$

when $W^{(2)}_{n-\eta_n^{(1)}} = 0$. Subsequently, we also obtain

$$I^{(1)}(n-\eta_n^{(2)}, n) = I^{(1)}(n-\eta_n^{(1)}, n) = 0. \tag{28}$$

Substituting Eqs. (27) and (28) in (22), we obtain that Eq. (21) holds also in the second case. Q.E.D.

Eq. (21) is of significant importance in our analysis, since it enables us to express $W_n^{(2)}$ in terms of $W^{(2)}_{n-\eta_n^{(1)}}$. We thus need to obtain first the distribution of the waiting times for the embedded sequence of arrivals at channel 2, which includes only message arrivals who had zero waiting time at channel 1. This sequence of waiting time is clearly $\{W^{(2)}_{n-\eta_n^{(1)}}\}$ and the corresponding embedded sequence of arrivals at channel 2 is readily observed to be $\{t^{(1)}_{n-\eta_n^{(1)}} + a_1, n \geq 1\}$.

Let the embedded waiting-time sequence $\{W^{(2)}_{n-\eta_n^{(1)}}\}$ be denoted as $\{W_n^{(2,1)}, n \geq 1\}$, $W_1^{(2,1)} = 0$, so that $W_n^{(2,1)}$ represents the waiting time at channel 2 for the n-th arriving message which had a zero waiting-time at channel 1. The following result will enable us to solve for the distribution of $W_n^{(2,1)}$ using standard results from the theory of M/G/1 queueing systems.

Proposition 2. For $a_2 > a_1$, the distribution functions for the waiting times $\{W_n^{(2,1)}, n \geq 1\}$ are identical to the distribution functions of the waiting times $\{W_n, n \geq 1\}$, $W_1 = 0$, arising in a M/G/1 queueing system for which the arrival process is a Poisson process with intensity λ, and the service times $\{X_n, n \geq 1\}$ are given by

$$X_n = N_n^{(1)}(a_2 - a_1). \tag{29}$$

Proof. For any M/G/1 queueing system, the waiting time sequence $\{W_n, n \geq 1\}$ satisfies the relationship

$$W_{n+1} = [W_n + X_n - (t_{n+1} - t_n)]^+, n \geq 1, \tag{30}$$

where $\{X_n, n \geq 1\}$ is the i.i.d. sequence of service times, and the interarrival durations $\{(t_{n+1} - t_n), n \geq 1\}$ are i.i.d. random-variables, exponentially distributed with parameter λ. The random-variables W_n, X_n, $(t_{n+1} - t_n)$, are all statistically interdependent $\forall n$. In the present case, using Eqs. (14) and (15) we obtain that the embedded waiting-time sequence $\{W_n^{(2,1)}, n \geq 1\}$ follows the relationship

$$W_{n+1}^{(2,1)} = [W_n^{(2,1)} + N_n^{(1)}(a_2 - a_1) - I_{(n+1)}^{(1)}]^+, n \geq 1. \tag{31}$$

The variable $I_{(n+1)}^{(1)}$ represents the duration of the recent idle-period of channel 1 prior to $t_{n+1}^{(1)}$, is exponentially distributed with mean λ^{-1}, since the arrival process at v_1 is Poisson with intensity λ, and is clearly statistically independent of $N_n^{(1)}$. The waiting time $W_n^{(2,1)}$ clearly depends only on the channel 1 variables $\{N_k^{(1)}, I_{(k+1)}^{(1)}, 1 \leq k \leq n-1\}$ and is therefore independent of $N_n^{(1)}$ and $I_{(n+1)}^{(1)}$. Consequently, if we let $X_n = N_n^{(1)}(a_2 - a_1)$ we conclude from Eqs. (30) and (31) that $W_n^{(2,1)}$ possesses the same distribution as W_n for the associated M/G/1 queueing system. Q.E.D.

We can now use known results from the theory of M/G/1 queueing systems to deduce the distribution of $W_n^{(2,1)}$. In particular, we are interested in steady-state results. However, it is known from M/G/1 theory that $N_n < \infty$ and $W_n < \infty$ w.p. 1 as $n \to \infty$ if the traffic intensity ρ is less than 1. Here, the traffic intensity $\rho^{(2,1)}$ for the equivalent M/G/1 system is (using Eq. (12a))

$$\rho^{(2,1)} = \lambda E(X_n) = \lambda E(N_n^{(1)})(a_2 - a_1) = (\rho_2 - \rho_1)/(1 - \rho_1). \tag{32}$$

We thus conclude the following (see also Lindley's theorem, [3] p. 41).

Corollary 1. If $\rho_2 < 1$ then the limiting distribution $\lim_{n\to\infty} P\{W_n^{(2,1)} \leq x\}$ exists and is independent of the distribution of $W_1^{(2,1)}$. If $\rho_2 \geq 1$, then $\lim_{n\to\infty} P\{W_n^{(2,1)} \leq x\} = 0$ for every x.

For an M/G/1 queueing system, the limiting waiting-time distribution $W(t) = \lim_{n\to\infty} P(W_n \leq t)$ and its Laplace transform $w(s) = \int_{0-}^{\infty} e^{-st} dW(t)$, for $\rho < 1$, are given by (see [1] p. 255)

$$w(s) = (1-\rho)\frac{\lambda^{-1}s}{\beta(s) - 1 + \lambda^{-1}s}, \ \mathrm{Re}\, s \geq 0, \tag{33a}$$

$$W(t) = (1-\rho)\sum_{n=0}^{\infty} \rho^n [\beta^{-1} \int_0^t \{1 - B(\tau)\} d\tau]^{n*}, \ t > 0, \tag{33b}$$

where $B(t) = P(X_n \leq t)$ and $\beta = E(X_n)$ are the distribution and mean of the service time, respectively, $\beta(s) = \int_0^\infty e^{-st} dB(t)$ and $[F(t)]^{n*}$ denotes the n-th convolution of $F(t)$ (and $[F(t)]^{0^*}$ is defined to be the unit step function). In particular, the moments of $W(t)$ can be obtained (see [1] p. 256). Of particular interest to us in the present analysis is the limiting mean waiting time. The latter is easily found from (33a) to be given by

$$\int_0^\infty t\, dW(t) = \frac{1}{2}\frac{\rho\beta}{1-\rho}\frac{\beta_2}{\beta^2}, \tag{34}$$

where $\beta_2 = E(X_n^2)$. Relation (34) is known as the Pollaczek-Khintchine formula. Using the latter relationships, the following results are deduced for the limiting distribution of the waiting time $W_n^{(2,1)}$, when equations (32), (12a) and (12b) are incorporated to express ρ, β and β^2, respectively.

Corollary 2. For $\rho_2 < 1$, the limiting waiting-time distribution

$$W^{(2,1)}(t) = \lim_{n\to\infty} P(W_n^{(2,1)} \leq t),$$

is given by (33b), and its transform by (33a), when we set $X_n = N_n^{(1)}(a_2-a_1)$. Hence, we have

$$W^{(2,1)}(t) = \frac{1-\rho_2}{1-\rho_1}\sum_{n=0}^{\infty}(\rho_2-\rho_1)^n\left\{\int_0^t P(N_n^{(1)} > [\tau/(a_2-a_1)])d\tau\right\}^{n*},\ t > 0, \tag{35}$$

where the distribution of $N_n^{(1)}$ is given by Eq. (11), and clearly $W^{(2,1)}(t) = 0$ for $t < 0$.

The limiting mean embedded waiting-time is given by

$$\overline{W}^{(2,1)} = \lim_{n\to\infty} E\{W_n^{(2,1)}\} = \frac{1}{2}\frac{1}{\lambda}\frac{(\rho_2-\rho_1)^2}{1-\rho_2}\frac{E\{(N_n^{(1)})^2\}}{E\{N_n^{(1)}\}}, \tag{36a}$$

and subsequently

$$\overline{W}^{(2,1)} = \frac{1}{2}\frac{1}{\lambda}\frac{(\rho_2-\rho_1)^2}{(1-\rho_1)^2(1-\rho_2)}. \tag{36b}$$

We return now to the non-embedded queueing process $\{X_t^{(2)}, t \geq 0\}$. The traffic intensity at channel 2, ρ_2, can be defined as the ratio between the average service time, a_2, and the average inter-arrival time. However, one readily shows that the latter quantity is given by, as $n \to \infty$,

$$E\{t_{n+1}^{(2)}-t_n^{(2)}\} = E\{r_{n+1}^{(1)}-r_n^{(1)}\} \to a_1\rho_1+(a_1+\lambda^{-1})(1-\rho_1) = \lambda^{-1}. \tag{37}$$

Hence, the traffic intensity ρ_2 is equal to $\lambda a_2 \triangleq \rho_2$. As for channel 1, $1-\rho_2$ is the limiting probability that channel 2 is free, as shown by the following.

Theorem 1. For $\rho_2 < 1$, we have

$$\lim_{n\to\infty} P(W_n^{(2)} = 0) = \lim_{t\to\infty} P(W_t^{(2)} = 0) = \lim_{t\to\infty} P(X_t^{(2)} = 0) = 1-\rho_2. \tag{38}$$

Proof. From M/G/1 theory (see [1] p. 255) and Proposition 2, we conclude that

$$\lim_{n\to\infty} P(W_n^{(2,1)} = 0) = 1-\rho^{(2,1)} = (1-\rho_2)/(1-\rho_1).$$

However, by definition of the embedded waiting sequence $\{W_n^{(2,1)}\}$ and Lemma 1 we have

$$\begin{aligned}\lim_{n\to\infty} P(W_n = 0) &= \lim_{n\to\infty} P(W_n^{(1)} = 0,\ W_n^{(2)} = 0)\\ &= \lim_{n\to\infty} P(W_n^{(1)} = 0)\, P(W_u^{(2)} = 0 \mid W_n^{(1)} = 0)\\ &= \lim_{n\to\infty} P(W_n^{(1)} = 0) P(W_n^{(2,1)} = 0) = (1-\rho_1)\,.\,(1-\rho_2)/(1-\rho_1)\\ &= 1-\rho_2,\end{aligned}$$

as claimed. We will show in next section that the idle periods in channel 2 are exponentially distributed. The latter property can readily be used to deduce that $\lim_{n\to\infty} P(W_n^{(2)} = 0) = \lim_{t\to\infty} P(W_t^{(2)} = 0)$. The latter equality can also be verified as follows. Applying the Key Renewal Theorem of Smith we obtain (see [1] pp. 290–291, and readily verifying the interarrival distribution to be non lattice), for $\rho_2 < 1$, that

$$\lim_{t\to\infty} P(W_t^{(2)} = 0) = \frac{E(I_n^{(2)})}{E(I_n^{(2)})+E(B_n^{(2)})}, \tag{39}$$

so that the limiting probability of the channel being idle at any time is given by the ratio of the average duration of the idle-period and the average duration of the busy-cycle (which is the overall duration of a busy-period and the following idle-period). We will show in the following analysis that $E(I_n^{(2)}) = \lambda^{-1}$ and $E(N_n^{(2)}) = (1-\rho_2)^{-1}$. Hence, $E(B_n^{(2)}) = E(N_n^{(2)})a_2 = a_2(1-\rho_2)^{-1}$. Substituting these means in (39) we obtain $\lim_{t\to\infty} P(W_t^{(2)} = 0) = 1-\rho_2$. Finally, we observe that $\{W_t^{(2)} = 0\}$ if and only if $\{X_t^{(2)} = 0\}$. Q.E.D.

The distributions of the random variables $N_n^{(2)}$, the number of messages transferred by channel 2 during a busy-period, and $h_n^{(2)}$, the number of channel 1 busy-periods included in a channel 2 busy-period, can also be derived using the statistics of the equivalent embedded queueing system. For that purpose, we show the following equivalence.

Lemma 4. For the equivalent M/G/1 queueing system with service times given by (29), let $B_n^{(2,1)}$ denote the duration of the n-th busy-period and $N_n^{(2,1)}$ represent the number of messages served during the n-th busy-period. Then, we have, $k \geq m$,

$$P\{N_n^{(2)} = k,\ h_n^{(2)} = m\} = P\{B_n^{(2,1)} = k(a_2-a_1),\ N_n^{(2,1)} = m\}. \tag{40}$$

Thus, $N_n^{(2)}$ and $h_n^{(2)}$ have the same distribution as $(a_2-a_1)^{-1}\, B_n^{(2,1)}$ and $N_n^{(2,1)}$, respectively.

Proof. The equivalent M/G/1 queueing system have been shown to describe the statistical behaviour of the channel 2 queueing process $\{X_t^{(2)},\ t \geq 0\}$ at the times

$\{_0t_n^{(1)}, n \geq 1\}$ corresponding to arrival-times of messages which had zero waiting-times at channel 1. Since $h_n^{(2)}$ is the number of the latter type of messages between two such messages which have zero waiting-time at channel 2, we conclude that $P(h_n^{(2)} = m) = P(N_n^{(2,1)} = m)$. Moreover, for the equivalent M/G/1 system we clearly have that $B_1^{(2,1)} = \sum_{i=1}^{N_1^{(2,1)}} X_i$. Hence, using Eq. (29) and $h_n^{(2)} \approx N_n^{(2,1)}$ (Where $X \approx Y$ denotes that X and Y have the same distribution), we obtain that $B_n^{(2,1)} \approx \sum_{i=1}^{h_n^{(2)}} N_i^{(1)}(a_2 - a_1)$. Since $N_n^{(2)} = \sum_{i=1}^{h_n^{(2)}} N_i^{(1)}$, we conclude that $B_n^{(2,1)} \approx N_i^{(1)}(a_2 - a_1)$, and that Eq. (40) holds. Q.E.D.

Using Lemma 4, and the known results for the distributions and moments of B_n and N_n for an M/G/1 queueing system (see, for example, [1] pp. 250–251), we obtain the distributions and moments of $N_n^{(2)}$ and $h_n^{(2)}$ as indicated by the following

Theorem 2. For $a_2 > a_1$, the joint distribution of $N_i^{(2)}$ and $h_i^{(2)}$, $i \geq 1$, is given by, $k \geq n$,

$$P\{N_i^{(2)} = k, h_i^{(2)} = n\} = (n!)^{-1}[(\rho_2 - \rho_1)k]^{n-1} \exp[-(\rho_2 - \rho_2)k] \cdot P(N_i^{(1)} = k)^{n*}, \tag{41}$$

where $P(N_i^{(1)} = k)$ is given by (11). For $\rho_2 < 1$, $P(N_i^{(2)} < \infty) = P(h_i^{(1)} < \infty) = 1$, and

$$E\{N_n^{(2)}\} = (1 - \rho_2)^{-1}, \tag{42a}$$

$$E\{[N_n^{(2)}]^2\} = (1 - \rho_2)^{-3}, \tag{42b}$$

$$E\{h_n^{(2)}\} = (1 - \rho_1)/(1 - \rho_2), \tag{42c}$$

$$E\{[h_n^{(2)}]^2\} = E(h_n^{(2)}) + 2E(h_n^{(2)}) \frac{\rho_2 - \rho_1}{1 - \rho_2}\left(1 + \frac{\rho_2 - \rho_1}{2(1 - \rho_1)(1 - \rho_2)}\right). \tag{42d}$$

Proof. Eq. (41) is obtained from M/G/1 theory as indicated above. The moments in (42) are readily obtained by utilizing the following expression for the joint characteristic function (see [1] p. 250 for its derivation),

$$\psi(r,s) = E\{r^{h_n(2)} e^{-sN_n(2)(a_2 - a_1)}\}, \ |r| \leq 1, \ \mathrm{Re}\ s \geq 0,$$

$$\psi(r,s) = r\beta\{s + \lambda - \lambda\psi(r,s)\}, \tag{43}$$

where

$$\beta\{s\} \triangleq E\{e^{-sN_n(1)(a_2 - a_1)}\}. \qquad \text{Q.E.D.}$$

We note that $N_1^{(2)} = \sum_{i=1}^{h_1^{(2)}} N_i^{(1)}$. Hence, by using Wald's Equation (which holds since, by (17), one deduces that $h_1^{(2)}$ is a stopping-time for $\{N_i^{(1)}, i \geq 1\}$) we conclude that $E(N_1^{(2)}) = E(h_1^{(2)})\, E(N_i^{(1)})$. However, by (12a), $E(N_i^{(1)}) = (1 - \rho_1)^{-1}$, and by a Renewal theorem (due to Erdös, Pollard and Feller, see [1] p. 103) we have (for the discrete-time renewal process with intervals $\{N_i^{(2)}, i \geq 1\}$),

$$\lim_{n \to \infty} P(W_n^{(2)} = 0) = [E(N_n^{(2)})]^{-1}. \tag{44}$$

Hence, using Eq. (38), we obtain that $E(N_1^{(2)}) = (1 - \rho_2)^{-1}$, and $E(h_1^{(2)}) = E(N_1^{(2)})/E(N_i^{(1)}) = (1 - \rho_1)/(1 - \rho_2)$, and thus verifying expressions (42a), (42c).

To obtain the limiting distribution of the waiting time $W_n^{(2)}$, we need to find the limiting distribution of $\eta_n^{(1)}$, defined by (19), and utilize relation (21). However, $\eta_n^{(i)}$ is clearly the backward-recurrence time for the discrete-time renewal point process $\{S_n^{(i)}, n > 0\}$, $S_0^{(i)} = 0$, where $S_{n+1}^{(i)} - S_n^{(i)} = N_{n+1}^{(i)}$, $n \geq 0$, or $S_n^{(i)} = \sum_{k=1}^{n} N_k^{(i)}$, $n \geq 1$ (i.e.; $\eta_n^{(i)}$ equals the duration between n and the recent event occurrence of $\{S_n^{(i)}\}$ prior to n). Hence, using renewal theory (see [1] pp. 113–114), we find the limiting distribution of $\eta_n^{(i)}$, when $E(N_n^{(i)}) < \infty$, to be given by

$$\eta^{(i)}(k) \triangleq \lim_{n\to\infty} P(\eta_n^{(i)} = k) = P(N_n^{(i)} > k)/E(N_n^{(i)}). \tag{45}$$

By eq. (45), the limiting mean of $\eta_n^{(i)}$ is obtained to be

$$\bar{\eta}^{(i)} \triangleq \lim_{n\to\infty} E(\eta_n^{(i)}) = E\{N_n^{(i)}(N_n^{(i)} - 1)\}/2E(N_n^{(i)}). \tag{46}$$

The limiting distribution and mean of $W_n^{(2)}$ subsequently follow.

Theorem 3. For $a_2 > a_1$, $\rho_2 < 1$, the limiting waiting-time distribution at channel 2, $W^{(2)}(t) = \lim_{n\to\infty} P\{W_n^{(2)} \leq t\}$, is given by

$$W^{(2)}(t) = (1-\rho_1)\sum_{k=0}^{\infty} W^{(2,1)}(t - k(a_2 - a_1))P(N_n^{(1)} > k), \tag{47}$$

where $W^{(2,1)}(t)$ is given by (35) and $P(N_n^{(1)} = k)$ by (11). The limiting average waiting time at channel 2, $\overline{W}^{(2)} = \lim_{n\to\infty} E(W_n^{(2)})$, is equal to

$$\overline{W}^{(2)} = \tfrac{1}{2}(a_2 - a_1)[(1-\rho_1)^{-1}(1-\rho_2)^{-1} - 1]. \tag{48}$$

Proof. Using relation (21), we obtain

$$\begin{aligned}
\lim_{n\to\infty} P(W_n^{(2)} \leq t) &= \lim_{n\to\infty} P\{W_{n-\eta_n^{(1)}}^{(2)} + \eta_n^{(1)}(a_2 - a_1) \leq t\} \\
&= \lim_{n\to\infty} \sum_{k=0}^{\infty} P\{W_{n-k}^{(2)} \leq t - k(a_2 - a_1) | \eta_n^{(1)} = k\} P\{\eta_n^{(1)} = k\} \\
&= \sum_{k=0}^{\infty} \lim_{n\to\infty} P\{W_n^{(2,1)} \leq t - k(a_2 - a_1)\} P\{\eta_n^{(1)} = k\} \\
&= \sum_{k=0}^{\infty} W^{(2,1)}(t - k(a_2 - a_1))\eta^{(1)}(k),
\end{aligned} \tag{49}$$

which yields (47) when (45) is incorporated to evaluate $\eta^{(1)}(k)$. The above interchange of limit and summation follows by the dominated convergence theorem, since

$$P\{W_{n-k}^{(2)} \leq t - k(a_2 - a_1), \eta_n^{(1)} = k\} \leq P(\eta_n^{(1)} = k) \leq P(N_i^{(1)} > k), \tag{50}$$

and

$$\sum_{k=0}^{\infty} P(N_i^{(1)} > k) = E\{N_i^{(1)}\} = (1-\rho_1)^{-1} < \infty. \tag{51}$$

The mean $\overline{W}^{(2)}$ is obtained from (21) to be given as

$$\overline{W}^{(2)} = \overline{W}^{(2,1)} + \bar{\eta}^{(1)}(a_2 - a_1). \tag{52}$$

Incorporating Eqs. (36b), (46) in (52), we obtain expression (48). Q.E.D.

The steady-state average delay time at channel 2, $\gamma^{(2)} = \overline{W}^{(2)}+a_2$, and the overall delay for the two-channel path, $\gamma = \gamma^{(1)}+\gamma^{(2)}$, can now be obtained using Eqs. (7), (48) and Proposition 1.

Corollary 3. For $\rho_1 < 1$, $\rho_2 < 1$, the limiting average delay at channel 2, $\gamma^{(2)}$, is equal to

$$\gamma^{(2)} = a_2+\tfrac{1}{2}[a_2-a_1]^+\{(1-\rho_1)^{-1}(1-\rho_2)^{-1}-1\}. \tag{53}$$

The overall delay for the two-channel path is given by

$$\gamma = \frac{1-\rho_1/2}{1-\rho_1}\,a_1+a_2+\tfrac{1}{2}[a_2-a_1]^+\{(1-\rho_1)^{-1}(1-\rho_2)^{-1}-1\}, \tag{54}$$

where $[x]^+ \triangleq \max(0,x)$.

The limiting average number of messages of channel i (waiting or being transmitted), $\bar{X}^{(i)} \triangleq \lim_{t\to\infty} E\{X_t^{(i)}\}$, can be evaluated by using Little's theorem (see [4]), which states that $\bar{X}^{(i)} = \lambda^{(i)}\gamma^{(i)}$. $\lambda^{(i)}$ denotes the limiting arrival rate at channel i (so that $[\lambda^{(i)}]^{-1} = E\{t_{n+1}^{(i)}-t_n^{(i)}\}$ as $n \to \infty$). By Eq. (37), we conclude that $\lambda^{(2)} = \lambda$, so that $\bar{X}^{(2)} = \lambda\gamma^{(2)}$. The average memory size $M^{(2)} = \beta\bar{X}^{(2)}$ is therefore

$$M^{(2)} = \beta\bar{X}^{(2)} = \lambda\beta\gamma^{(2)} = \beta\rho_2+\tfrac{1}{2}\beta[\rho_2-\rho_1]^+\{(1-\rho_1)^{-1}(1-\rho_2)^{-1}-1\} \quad \text{bits.} \tag{55}$$

The overall required average memory size $M = M^{(1)}+M^{(2)}$, is thus evaluated by incorporating Eqs. (9), (55). Hence, $M = \lambda\beta\gamma$ bits, where γ is given by (54).

5. Delays in an *n*-Channel Path. Consider an n-channel path, where $n \geq 2$. Due to the series structure of the path, we clearly have

$$t_n^{(i)} = r_n^{(i-1)},\ i \geq 2, \tag{56}$$

so that the interarrival time at channel i, $i \geq 2$, is given by

$$T_{n+1}^{(i)} = t_{n+1}^{(i)}-t_n^{(i)} = r_{n+1}^{(i-1)}-r_n^{(i-1)} = \begin{cases} a_{i-1} & \text{if} \quad W_{n+1}^{(i-1)} > 0 \\ a_{i-1}+I_{(n+1)}^{(i-1)}, & \text{if} \quad W_{n+1}^{(i-1)} = 0 \end{cases} \tag{57}$$

The waiting-time at channel i follows the relationship

$$W_{n+1}^{(i)} = [W_n^{(i)}+a_i-T_{n+1}^{(i)}]^+. \tag{58}$$

Hence, by (57)–(58), we conclude that

$$W_n^{(i)} = 0, \quad \text{if} \quad a_i \leq a_{i-1},\ i \geq 2. \tag{59}$$

However, the above expressions can be simplified as follows. We define a sequence of ladder indices $\{k_i,\ _i = 1, 2, \ldots, m\}$, so that

$$\begin{aligned} k_1 &= 1, \\ k_j &= \min\{i : k_{j-1} < i \leq n,\ a_i > a_{k_{j-1}}\}, \quad \text{for} \quad j \geq 2. \end{aligned} \tag{60}$$

Thus, $k_2 = i$ if $a_1 \geq a_j$, $2 \leq j \leq i-1$, and $a_1 < a_i$; so that k_2 is the index of the first element of $\{a_i,\ 1 \leq i \leq n\}$ which is larger than a_1. Similarly, k_j is the index of the first element of $\{a_i,\ 1 \leq i \leq n\}$, following $a_{k_{j-1}}$, which is larger than $a_{k_{j-1}}$.

For example, for the sequence $\{a_i\} = \{6,8,2,5,10\}$, we have $k_1 = 1$, $k_2 = 2$, $k_3 = 5$, $a_{k_1} = 6$, $a_{k_2} = 8$, $a_{k_3} = 10$ and $m = 3$. The number of ladder indices, m, is clearly defined by $\max_j k_j = k_m$, $1 \le m \le n$. We observe that $k_1 < k_2 < \ldots < k_m$ and $a_{k_1} < a_{k_2} < \ldots < a_{k_m}$. We will now show that a message will have a zero waiting time at any channel j, where j is not a ladder index of $\{a_i\}$. Thus, a message may have to wait only at channels whose index correspond to a ladder index of $\{a_i\}$. For that purpose, we first obtain the following property.

Lemma 5. For $i \ge 2$, we have

$$t_{n+1}^{(i)} - t_n^{(i)} = r_{n+1}^{(m_i)} - r_n^{(m_i)}, \tag{61}$$

where

$$m_i = \max\{k_j : 1 \le k_j \le m, k_j < i\}. \tag{62}$$

Proof. Consider an arbitrary channel i, $i \ge 2$. The integer m_i represents the largest ladder index smaller than i. Thus, if $m_i = i-1$, then Eq. (61) clearly holds, due to relation (56). If $m_i = i-2$, then $i-2$ is a ladder index and by (60) we observe that $a_{i-2} \ge a_{i-1}$. Consequently, by (59), $W_n^{(i-1)} = 0$ and $r_n^{(i-1)} = t_n^{(i-1)} + a_{i-1}$ so that $r_{n+1}^{(i-2)} - r_n^{(i-2)} = t_{n+1}^{(i-1)} - t_n^{(i-1)} = r_{n+1}^{(i-1)} - r_n^{(i-1)} = t_{n+1}^{(i)} - t_n^{(i)}$, and Eq. (61) holds. If $m_i = i-3$, then $i-3$ is a ladder index and $i-2$ and $i-1$ are not ladder indices. Hence, by (60), $a_{i-3} \ge a_{i-3}, a_{i-2} \ge a_{i-1}$. Subsequently, $W_n^{(i-2)} = 0$ and $r_n^{(i-2)} = t_n^{(i-2)} + a_{i-2}$, so that $r_{n+1}^{(i-3)} - r_n^{(i-3)} = t_{n+1}^{(i-2)} - t_n^{(i-2)} = r_{n+1}^{(i-2)} - r_n^{(i-2)} = t_{n+1}^{(i-1)} - t_n^{(i-1)}$. Hence, since $r_{n+1}^{(i-3)} - r_n^{(i-3)} \ge a_{i-3}$ and $a_{i-3} \ge a_{i-1}$, we conclude that $t_{n+1}^{(i-1)} - t_n^{(i-1)} \ge a_{i-3} \ge a_{i-1}$, and therefore using (58) we obtain that $W_n^{(i-1)} = 0$. Subsequently, $t_{n+1}^{(i-1)} - t_n^{(i-1)} = r_{n+1}^{(i-1)} - r_n^{(i-1)} = t_{n+1}^{(i)} - t_n^{(i)}$, so that Eq. (61) holds. Clearly, the proof extends to any arbitrary value $m_i = i-k$, $1 \le k \le i-1$. Q.E.D.

In particular, the following property follows readily from Lemma 5 and its proof, by observing that if $m_i = i-k$, $1 \le k \le i-1$, then $W_n^{(j)} = 0$ for $i-k+1 \le {}_j \le i-1$, for each $n \ge 1$.

Theorem 4. In an n-channel path, if i is not a ladder index for $\{a_i, 1 \le i \le n\}$ (i.e., i is not equal to any k_j, $1 \le j \le m$), then

$$W_n^{(i)} = 0,$$

for each $n \ge 1$, $i \ge 2$.

Thus, we have shown that for any channel i, $i \ge 2$, we have for each $n \ge 1$,

$$a_i \le a_{i-1} \quad \text{or} \quad a_{i-1} < a_i \le a_{m_i} \Rightarrow W_n^{(i)} = 0. \tag{63}$$

From Lemma 5 and Theorem 4 we conclude that in order to calculate the distributions of the waiting times at the channels, we need to consider just a "reduced" modified path. The modified path is generated from the original path by "short-circuiting" all the channels whose index does not correspond to a ladder index (to be called non-ladder channels in contrast to ladder channels). Thus, the modified path consists of a series connection of the ladder channels in order of increasing ladder indices. The first channel of the modified path is thus channel $k_1 = 1$, the second one is channel k_2, and the last one is channel k_m, $1 \le k_m \le n$. The corresponding transfer times for the ladder channels are $a_1 < a_{k_2} < a_{k_3} < \ldots < a_{k_m}$. Since the waiting time $W_n^{(i)}$ depends on the arrival process through the inter-arrival time $T_{n+1}^{(i)}$, as seen by (57)–(58), Lemma 5 implies that the waiting time $W_n^{(k_i)}$ will assume the same values in the original path and in the modified path. In the following analysis, we thus obtain the waiting-time distributions in a modified path. The delays in the original path are subsequently readily deduced.

Consider the m-channel modified path. Many of the properties for this path follow readily in the same manner as for the 2-channel path, and the reader is referred to the previous section for their detailed proofs. In particular, the properties indicated in the following proposition, follow as those in Lemma 1, Lemma 2 and Lemma 3.

Proposition 3. For $k_i = 1, 2, 3, \ldots, m$, we have:

$$1.\quad \{W_n^{(k_i)} = 0\} \Rightarrow \{W_n^{(k_{i-1})} = 0\}. \tag{64}$$

$$2.\quad N_1^{(k_i)} = \sum_{i=1}^{h_1(k_i)} N_i^{(k_{i-1})} \tag{65}$$

where

$$h_1^{(k_i)} = \min \{n | n \ge 1, \sum_{i=1}^{k} [N_i^{(k_{i-1})}(a_{k_i} - a_{k_{i-1}}) - I_i^{k_{i-1})}] > 0, k = 1, 2, \ldots, n-1,$$
$$\sum_{i=1}^{n} [N_i^{(k_{i-1})}(a_{k_i} - a_{k_{i-1}}) - I_i^{(k_{i-1})}] \le 0\}. \tag{66}$$

3. For any $n \ge 1$,

$$W_n^{(k_i)} = W_{n-\eta_n(k_{i-1})}^{(k_i)} + \eta_n^{(k_{i-1})}(a_{k_i} - a_{k_{i-1}}), \tag{67}$$

where $\eta_n^{(i)}$ is defined by (19).

It is interesting to observe that property (64), when used recursively, for each $k_i \ge 2$, implies that $\{W_n^{(k_i)} = 0\} \Rightarrow \{W_n^{(1)} = 0\}$, so that the following result can be stated.

Corollary 4. For any ladder index k_i, $2 \le k_i \le m$, and any $n \ge 1$, we have

$$\{W_n^{(k_i)} = 0\} \quad \text{only if} \quad \{W_n^{(1)} = 0\}. \tag{68}$$

Thus, Corollary 4 indicates that a message can have a zero waiting-time at any ladder channel $k_i \ge 2$, only if it had a zero-waiting time at channel 1, while entering the path. Clearly, at non-ladder channels all the messages have zero-waiting times.

The embedded waiting-time sequence $\{W_{n-\eta_n(k_{i-1})}^{(k_i)}\}$ is now denoted as $\{W_n^{(k_i,k_{i-1})}, n \ge 1\}$, $W_1^{(k_i,k_{i-1})} = 0$, $2 \le k_i \le m$. Thus $W_n^{(k_i,k_{i-1})}$ represents the waiting time at ladder-channel k_i for the n-th arriving message which had a zero waiting time at the preceding ladder-channel k_{i-1}. To solve for the distribution functions of $\{W_n^{(k_i,k_{i-1})}, n \ge 1\}$ we will obtain a result equivalent to that expressed by Proposition 2 for the two-channel path. Since the departure point processes from the channels are not Poisson processes, the utilization of a corresponding

M/G/1 queueing system as in Proposition 2, is not directly justified. However, the following result will enable us to resolve the problem.

Proposition 4. For any channel i, $1 \leq i \leq n$, the random sequences $\{N_n^{(i)}, n \geq 1\}$, $\{I_n^{(i)}, n \geq 1\}$, are statistically independent sequences of i.i.d. random-variables. For any channel i, $1 \leq i \leq n$, the idle-period $I_n^{(i)}$ is exponentially distributed with mean λ^{-1}, for any $n \geq 1$. Thus,

$$P\{I_m^{(i)} \leq \tau\} = [1-\exp(-\lambda\tau)]U(\tau), \tag{69}$$

for any m and i, $m \geq 1$, $1 \leq i \leq n$, where $U(\tau)$ is the unit-step function so that $U(\tau) = 0$ for $\tau \leq 0$ and $U(\tau) = 1$ for $\tau > 0$.

To prove Proposition 4, we need first to establish the following result, which yields a necessary and sufficient condition for $(W_{n+1}^{(k_i)} = 0)$ in terms of $W_n^{(k_i)}$ and an appropriate idle-period duration in channel k_{i-1}.

Lemma 6. For any ladder channel, k_i, $1 < k_i \leq m$, and any n, $n \geq 1$, we have

$$\{W_{n+1}^{(k_i)} = 0\} \Leftrightarrow \{W_n^{(k_i)} \leq I^{(k_{i-1})}(r_n^{(k_{i-1})}, t_{n+1}^{(k_{i-1})}) - (a_{k_i} - a_{k_{i-1}})\}, \tag{70}$$

where $I^{(j)}(r_n^{(j)}, t_{n+1}^{(j)})$ denotes the length of time the j-th channel is idle during $(r_n^{(j)}, t_{n+1}^{(j)})$.

Proof. By Eqs. (57) – (58) and Lemma 5, it follows that

$$\{W_{n+1}^{(k_i)} = 0\} \Leftrightarrow \{W_n^{(k_i)} + a_{k_i} - T_{n+1}^{(k_i)} \leq 0\}. \tag{71}$$

Assume that $W_{n+1}^{(k_i)} = 0$. Then, by Eq. (64), we also have that $W_{n+1}^{(k_{i-1})} = 0$, and subsequently by (57) and (61)

$$T_{n+1}^{(k_i)} = a_{k_{i-1}} + I^{(k_{i-1})}(r_n^{(k_{i-1})}, t_{n+1}^{(k_{i-1})}). \tag{72}$$

Substituting (72) into (71) yields the RHS of (70). Now assume that the RHS of (70) holds. Then, we conclude that

$$I^{(k_{i-1})}(r_n^{(k_{i-1})}, t_{n+1}^{(k_{i-1})}) \geq W_n^{(k_i)} + (a_{k_i} - a_{k_{i-1}}) \geq (a_{k_i} - a_{k_{i-1}}) > 0,$$

since $W_n^{(j)} \geq 0$ and $a_{k_i} > a_{k_{i-1}}$. However, the latter idle-time being positive clearly implies that $W_{n+1}^{(k_{i-1})} = 0$. Subsequently, using (57)-(58), it follows that

$$T_{n+1}^{(k_i)} = a_{k_{i-1}} + I^{(k_{i-1})}(r_n^{(k_{i-1})}, t_{n+1}^{(k_{i-1})}).$$

Hence, by incorporating the latter equality into the inequality expressed at the RHS of (70), we conclude that

$$W_n^{(k_i)} + a_{k_i} - T_{n+1}^{(k_i)} \leq 0,$$

and therefore, by (71), $W_{n+1}^{(k_i)} = 0$. Q.E.D.

Lemma 6 is now applied to prove Proposition 4.

Proof of Proposition 4. The properties indicated in the proposition clearly hold for channel 1. Using Lemma 5, relation (61), we deduce that if the proposition is shown to hold for ladder channels, it subsequently holds also for non-ladder channels. Hence, it is sufficient to prove the claims of the proposition for the modified path.

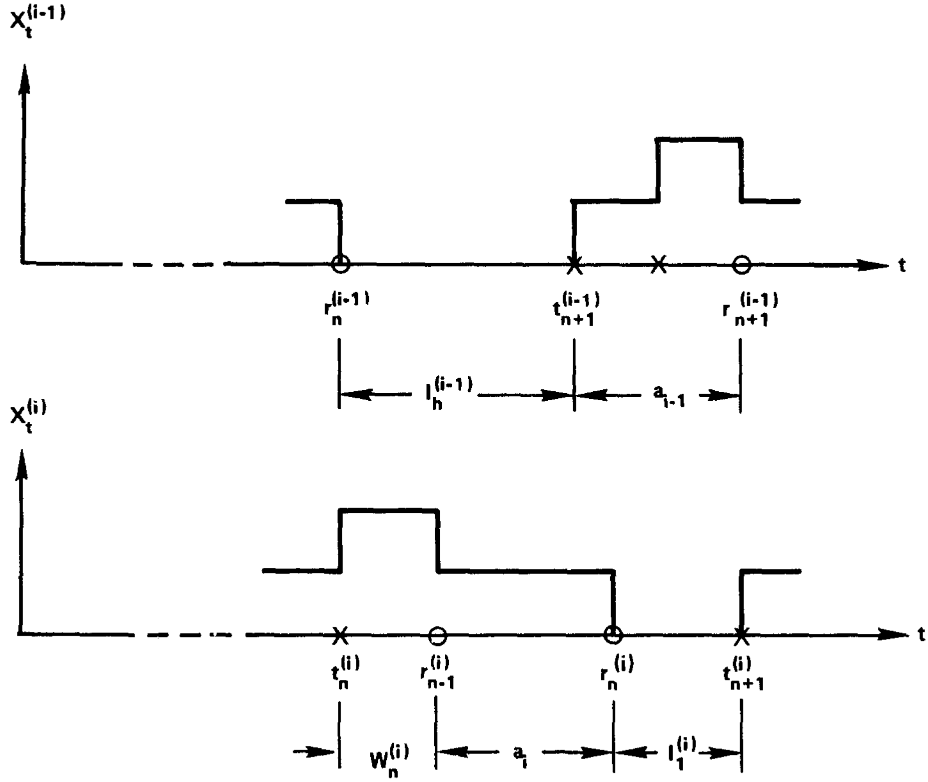


Figure 3. A Queueing Process Realization Over the ith and $(i-1)$st Channel, When $W_{n+1}^{(i)} = 0$, $a_i > a_{i-1}$.

Consider now two successive channels, k_{i-1} and k_i, $k_i \geq 2$, in the modified path. Assume $N_1^{(k_i)} = n$. Then $W_{n+1}^{(k_i)} = 0$. Hence, by (64), $W_{n+1}^{(k_i-1)} = 0$. We then have (see Fig. 3),

$$\begin{aligned} r_{n+1}^{(k_i-1)} - r_n^{(k_i-1)} &= (r_{n+1}^{(k_i-1)} - t_{n+1}^{(k_i-1)}) + (t_{n+1}^{(k_i-1)} - r_n^{(k_i-1)}) \\ &= I^{(k_i-1)}(r_n^{(k_i-1)}, t_{n+1}^{(k_i-1)}) + a_{k_{i-1}}, \end{aligned} \tag{73a}$$

where, clearly, $I^{(k_i-1)}(r_n^{(k_i-1)}, t_{n+1}^{(k_i-1)}) = I_{h_1(k_i)}^{(k_i-1)}$. Also,

$$t_{n+1}^{(k_i)} - t_n^{(k_i)} = (t_{n+1}^{(k_i)} - r_n^{(k_i)}) + (r_n^{(k_i)} - t_n^{(k_i)}) = I_1^{(k_i)} + W_n^{(k_i)} + a_{k_i}, \tag{73b}$$

since $(r_n^{(j)} - t_n^{(j)}) = W_n^{(j)} + a_j$, being the overall time the n-th message spends at the j-th channel, in waiting and being transmitted. Using relation (61), we can equate (73a) and (73b) to obtain

$$I_1^{(k_i)} = I_{h_1(k_i)}^{(k_i-1)} - (a_{k_i} - a_{k_{i-1}}) - W_n^{(k_i)}. \tag{74}$$

Therefore, using (74) and Lemma 6, Eq. (70), we obtain for $\tau > 0$,

$$\begin{aligned} &P\{I_1^{(k_i)} > \tau | N_1^{(k_i)} = n,\ h_1^{(k_i)} = h,\ W_n^{(k_i)} = w,\ W_{n+1}^{(k_i)} = 0\} \\ &= P\{I_h^{(k_i-1)} > \tau + w + (a_{k_i} - a_{k_{i-1}}) | I_h^{(k_i-1)} \geq w + (a_{k_i} - a_{k_{i-1}}), \\ &\qquad N_1^{(k_i)} = n, h_1^{(k_i)} = h,\ W_n^{(k_i)} = w\} \\ &= P\{I_h^{(k_i-1)}) > \tau + w + (a_{k_i} - a_{k_{i-1}}) | I_h^{(k_i-1)} \geq w + (a_{k_i} - a_{k_{i-1}})\}. \end{aligned} \tag{75}$$

The last equality in (75) follows by observing that the random variables $N_1^{(k_i)}$, $h_1^{(k_i)}$ and $W_n^{(k_i)}$ at channel k_i, are statistically independent of the queueing process at channel k_{i-1} for times $t > r_{N_1(k_i)}^{(k_{i-1})}$; and therefore, in particular, are independent of the duration $I_h^{(k_{i-1})}$. The latter observation is readily verified for k_2 (see the proof of Lemma 2), and subsequently recursively for any k_i, using Proposition 3. Equation (69) will now follow from (75) by induction on k_i. For $k_1 = 1$, the arrival process is Poisson, and therefore $P\{I_h^{(1)} > \tau\} = \exp(-\lambda\tau)U(\tau)$. Hence $P\{I_h^{(1)} > \tau+s | I_h^{(1)} > s\} = \exp(-\lambda\tau)$, $\tau > 0$. Equation (75) consequently implies that $P\{I_1^{(k_2)} > \tau\} = \exp(-\lambda\tau)$, $\tau > 0$, so that $I_1^{(k_2)}$ is exponentially distributed with mean λ^{-1}. Proceeding inductively, Eq. (75) thus yields result (69). In particular, we have obtained that the conditional distribution (75) equals to $\exp(-\lambda\tau)$, and thus verifying that $I_1^{(k_i)}$ is statistically independent of $N_1^{(k_i)}$. Combining the latter property with our previous observations, we readily conclude that the random sequences $\{N_n^{(i)}, n \geq 1\}$, $\{I_n^{(i)}, n \geq 1\}$ are statistically independent sequences of i.i.d. random-variables. Q.E.D.

Using Eqs. (57)–(58), we can write for the embedded waiting sequence at channel k_i, $k_i \geq 2$,

$$W_{n+1}^{(k_i,k_{i-1})} = [W_n^{(k_i,k_{i-1})} + I_{(n+1)}^{(k_{i-1})} + (a_{k_i} - a_{k_{i-1}})]^+. \tag{76}$$

By Proposition 4, the idle period $I_{(n+1)}^{(k_{i-1})}$ is exponentially distributed with mean λ^{-1} and is statistically independent of $W_n^{(k_i,k_{i-1})}$. Hence, proceeding as in the proof of Proposition 2, we deduce the following result.

Proposition 5. The distribution functions for the waiting times $\{W_n^{(k_i,k_{i-1})}, n \geq 1\}$ are identical to the distribution functions of the waiting times $\{W_n, n \geq 1\}$, $W_1 = 0$, arising in a M/G/1 queueing system for which the arrival process is a Poisson process with intensity λ, and the service times $\{X_n, n \geq 1\}$ are given by

$$X_n = N_n^{(k_{i-1})}(a_{k_i} - a_{k_{i-1}}). \tag{77}$$

Using the result stated by Proposition 5, we proceed to obtain the waiting-time and busy-period distributions at the channels of the communication path. The derivations of the latter distributions are identical to those presented for the two-channel path, when the modified path is considered. The results are thus summarized by the following theorem.

Theorem 5. For ladder channels k_i, $1 \leq k_i \leq m$, we have:

1. For $\rho_{k_i} = \lambda a_{k_i} \geq 1$, $\lim_{n\to\infty} P(W_n^{(k_i,k_{i-1})} \leq t) = 0$ for each t. For $\rho_{k_i} < 1$, the limiting embedded waiting time distribution exists and is given by

$$W^{(k_i,k_{i-1})}(t) = \lim_{n\to\infty} P(W_n^{(k_i,k_{i-1})} \leq t)$$

$$= \frac{1-\rho_{k_i}}{1-\rho_{k_{i-1}}} \sum_{n=0}^{\infty} (\rho_{k_i} - \rho_{k_{i-1}})^n \left\{\int_0^t P(N_n^{(k_{i-1})} > [\tau/(a_{k_i} - a_{k_{i-1}})] d\tau\right\}^{n*} \tag{78}$$

for $t > 0$, and $W^{(k_i,k_{i-1})}(t) = 0$ for $t < 0$.

2. For $n = 1, 2, \ldots,$

$$P(N_k^{(k_i)} = n) = (n!)^{-1}(n\rho_{k_i})^{n-1} \exp(-n\rho_{k_i}). \tag{79}$$

3. For $\rho_{k_i} < 1$, we have

$$\lim_{n\to\infty} P(W_n^{(k_i)} = 0) = \lim_{t\to\infty} P(W_t^{(k_i)} = 0) = \lim_{t\to\infty} P(X_t^{(k_i)} = 0) = 1-\rho_{k_i}. \tag{80}$$

4. For $k \geq n \geq 1$, we have

$$P\{N_j^{(k_i)} = k, h_j^{(k_i)} = n\} = (n!)^{-1}[(\rho_{k_i}-\rho_{k_{i-1}})k]^{n-1} \exp[-(\rho_{k_i}-\rho_{k_{i-1}})k] \, P(N_j^{(k_{i-1})} = k)^{n*}. \tag{81}$$

For $\rho_{k_i} < 1, P(N_j^{(k_i)} < \infty) = P(h_j^{(k_i)} < \infty) = 1,$ and

$$E(N_n^{(k_i)}) = (1-\rho_{k_i})^{-1}, \tag{82a}$$

$$E([N_n^{(k_i)}]^2) = (1-\rho_{k_i})^{-3}, \tag{82b}$$

$$E(h_n^{(k_i)}) = (1-\rho_{k_{i-1}})/(1-\rho_{k_i}) \tag{82c}$$

$$E([h_n^{(k_i)}]^2) = E(h_n^{(k_i)}) + 2E(h_n^{(k_i)}) \frac{\rho_{k_i}-\rho_{k_{i-1}}}{1-\rho_{k_i}} \left(1+\frac{\rho_{k_i}-\rho_{k_{i-1}}}{2(1-\rho_{k_i})(1-\rho_{k_{i-1}})}\right). \tag{82d}$$

5. For $\rho_{k_i} \geq 1, \lim_{n\to\infty} P(W_n^{(2)} \leq t) = 0$ for each t. For $\rho_{k_i} < 1$, the limiting waiting time distribution exists and is given by

$$\begin{aligned} W^{(k_i)}(t) &= \lim_{n\to\infty} P\{W_n^{(k_i)} \leq t\} \\ &= (1-\rho_{k_i}) \sum_{k=0}^{\infty} W^{(k_i,k_{i-1})}(t-k(a_{k_i}-a_{k_{i-1}}))P(N_n^{(k_{i-1})} > k), \end{aligned} \tag{83}$$

where $W^{(k_i,k_{i-1})}(t)$ is given by (78) and $P(N_n^{(k_{i-1})} = k)$ by (79). The limiting average waiting time at channel k_i, $\overline{W}^{(k_i)} = \lim_{n\to\infty} E(W_n^{(k_i)})$, is equal to

$$\begin{aligned} \overline{W}^{(k_i)} &= \tfrac{1}{2}(a_{k_i}-a_{k_{i-1}})[(1-\rho_{k_{i-1}})^{-1}(1-\rho_{k_i})^{-1}-1] \\ &= \frac{1}{2}\frac{\rho_{k_i}}{1-\rho_{k_i}} a_{k_i} - \frac{1}{2}\frac{\rho_{k_{i-1}}}{1-\rho_{k_{i-1}}} a_{k_{i-1}}. \end{aligned} \tag{84}$$

The overall average delay over an n-channel path and the memory-sizes are readily deduced from Eq. (84) and Little's Theorem to yield the following.

Corollary 5. The overall delay for the n-channel path is given by

$$\gamma = \sum_{i=1}^{m} \overline{W}^{(k_i)} + \sum_{1}^{n} a_i = \frac{1}{2}\frac{\rho_{k_m}}{1-\rho_{k_m}} a_{k_m} + \sum_{1}^{n} a_i. \tag{85}$$

The memory size for channel k_i, $1 \leq k_i \leq m$, is equal to

$$M^{(k_i)} = \lambda\beta\gamma^{(k_i)} = \lambda\beta(\overline{W}^{(k_i)}+a_{k_i}) \text{ bits}. \tag{86}$$

It is interesting to observe (see (85)) that the overall average waiting-time depends only on the largest transmission time a_{k_m}, and thus only on the value of the smallest capacity in the communication path. In the next section we will explain the latter property as well as derive the distribution function for the overall delay.

6. The Distribution of the Overall Path Delay, A Capacity Ordering Invariance Property.

Let

$$S_n^{(k)} \triangleq \sum_{i=1}^{k} W_n^{(i)},\ n \geq 1, k \geq 1, \tag{87}$$

denote the overall waiting time of the n-th message over channels $(1, 2, \ldots, k)$. To obtain the distribution of the m-th message delay γ_m, where

$$\gamma_m = S_m^{(n)} + \sum_{i=1}^{n} a_i, \tag{88}$$

we first obtain the distribution of $S_m^{(n)}$. For that purpose, we will utilize the following Lemma.

***Lemma** 7.* The random-variable $S_n^{(k)}$ satisfies the relationship, $k \geq 1, n \geq 1$,

$$S_{n+1}^{(k)} = [S_n^{(k)} + \max(a_1, a_2, \ldots, a_k) - T_{n+1}^{(1)}]^+. \tag{89}$$

Proof. We will prove (89) by induction on $k \geq 1$. First consider the case $k = 1$. Then $W_{n+1}^{(1)} = [W_n^{(1)} + a_1 - T_{n+1}^{(1)}]^+$, which is identical to (89). Assume now that (89) holds for $1 \leq k \leq m$, and show it to hold also for $k = m+1$. Let $a_j = \max(a_1, a_2, \ldots, a_m)$. If $a_{m+1} \leq a_j$, then $W_n^{(m+1)} = 0$ for each $n \geq 1$. Hence, $S_{n+1}^{(m+1)} = S_{n+1}^{(m)} + W_{n+1}^{(m+1)} = S_{n+1}^{(m)} = [S_n^{(m)} + a_j - T_{n+1}^{(1)}]^+ = [S_n^{(m+1)} + \max(a_1, \ldots, a_{m+1}) - T_{n+1}^{(1)}]^+$, by the induction hypothesis (and $S_n^{(m+1)} = S_n^{(m)}$ for each n). If now $a_{m+1} > a_j$, we consider two cases.

In the first case, we assume that $S_{n+1}^{(m)} > 0$. Then, by (64) we have $W_{n+1}^{(j)} > 0$ and subsequently $W_{n+1}^{(m+1)} > 0$. Hence, by (61) and (57), $T_{n+1}^{(m+1)} = a_j$. Consequently, using the induction hypothesis, we obtain

$$\begin{aligned} S_{n+1}^{(m+1)} &= S_{n+1}^{(m)} + W_{n+1}^{(m+1)} = [S_n^{(m)} + a_j - T_{n+1}^{(1)}] + [W_n^{(m+1)} + a_{m+1} - T_{n+1}^{(m+1)}] \\ &= [S_n^{(m)} + W_n^{(m+1)} + a_{m+1} - T_{n+1}^{(1)}] = [S_n^{(m+1)} + \max(a_1, \ldots, a_{m+1}) - T_{n+1}^{(1)}]^+, \end{aligned}$$

so that (89) holds for $k = m+1$.

Consider now the second case, where $a_{m+1} > a_j$ and $S_{n+1}^{(m)} = 0$. Clearly, $S_{n+1}^{(m)} = 0$ implies that $W_{n+1}^{(i)} = 0$ for each i, $1 \leq i \leq m$. Consequently, we obtain

$$T_{n+1}^{(k+1)} = T_{n+1}^{(k)} - W_n^{(k)}, \text{ for } 1 \leq k \leq m,$$

which when used recursively yields

$$T_{n+1}^{(m+1)} = T_{n+1}^{(1)} + \sum_{k=1}^{m} (T_{n+1}^{(k+1)} - T_{n+1}^{(k)}) = T_{n+1}^{(1)} - S_n^{(m)}.$$

We thus obtain

$$\begin{aligned} S_{n+1}^{(m+1)} = W_{n+1}^{(m+1)} &= [W_n^{(m+1)} + a_{m+1} - T_{n+1}^{(m+1)}]^+ = [W_n^{(m+1)} + a_{m+1} + S_n^{(m)} \\ &\quad - T_{n+1}^{(1)}]^+ = [S_n^{(m+1)} + \max(a_1, \ldots, a_{m+1}) - T_{n+1}^{(1)}]^+, \end{aligned}$$

which yields (89) for $k = m+1$. Q.E.D.

One readily observes in (89) that $T_{n+1}^{(1)}$ is statistically independent of $S_n^{(k)}$. Also recall that $T_{n+1}^{(1)}$ is the inter-arrival duration of a Poisson process with rate λ, and is therefore exponentially distributed with mean λ^{-1}. Hence, comparing (89) with the waiting-time recurrence relationship, $W_{n+1} = [W_n + a - T_{n+1}]^+$, for the waiting time W_n in an M/D/1 queueing system with service time a and arrival rate λ, we obtain the following result for the distribution of $\sum_{i=1}^{n} W_m^{(i)}$, using Eq. (5).

Theorem 6. The overall waiting time for the m-th message at an n-channel path, $S_m^{(n)} = \sum_{i=1}^{n} W_m^{(i)}$, has the same distribution as the waiting time W_m in an M/D/1 queueing system, with Poisson arrivals with rate λ and service time equal to $a_{k_m} = \max(a_1, a_2, \ldots, a_n)$. If $\rho_{k_m} < 1$, the limiting distribution of the overall waiting time exists and is given by

$$\begin{aligned} S(x) &= \lim_{m\to\infty} P(S_m^{(n)} \le x) = \lim_{m\to\infty} P\left(\sum_{i=1}^{n} W_m^{(i)} \le x\right) \\ &= (1-\rho_{k_m}) \sum_{i=0}^{n} \exp[\lambda(x - ia_{k_m})](i!)^{-1}[-\lambda(x - ia_{k_m})]^i, \end{aligned} \tag{90}$$

where $x = na_{k_m} + t$, $n \ge 0$, $0 \le t \le a_{k_m}$.|

In particular, the limiting average overall waiting time follows from (90) and (6) to be given by

$$\overline{W} = \lim_{m\to\infty} E\left(\sum_{i=1}^{n} W_m^{(i)}\right) = \frac{1}{2}\frac{\rho_{k_m}}{1-\rho_{k_m}} a_{k_m}. \tag{91}$$

The overall average delay γ is equal to $\overline{W} + \sum_{i=1}^{n} a_i$, so that expression (85) results.

We note that

$$W_n^{(k)} = S_n^{(k)} - S_n^{(k-1)}. \tag{92}$$

Hence, $E(W_n^{(k)}) = E(S_n^{(k)}) - E(S_n^{(k-1)})$, and in steady-state expression (84) for the average waiting-time results. One observes, however, that the distribution of $W_n^{(k)}$ cannot be directly calculated from (92) since $S_n^{(k)}$ and $S_n^{(k-1)}$ are not statistically independent. The limiting distribution of $W_n^{(k)}$ has been shown to be given by (83) (if k is a ladder channel, otherwise $W_n^{(k)} = 0$). We also note that the number of messages transferred between two zeros of $\{S_n^{(k_i)}, n \ge 1\}$ are equal to the variable $N_m^{(k_i)}$, for some $m \ge 1$. Hence, by Theorem 6, $N_m^{(k_i)}$ has the same distribution as N_m for an M/D/1 system with arrival rate λ and service time a_{k_i}. Subsequently, using Eq. (11), we have verified expression (79).

It is of particular interest to indicate that Lemma 7, Eq. (89), implies the following interesting property, which follows by observing that in (89) only $\max(a_1, \ldots, a_k)$ is utilized to evaluate the overall waiting time in the path.

Theorem 7. (Capacity Ordering Invariance Property).

The overall delay time over an n-channel path with capacities $(C_1, C_2, \ldots, C_n)$ is the same as that over an n-channel path with capacities $(C_{i_1}, C_{i_2}, \ldots, C_{i_n})$, where the latter sequence is an arbitrary ordering of $(C_1, \ldots, C_n)$. The overall waiting-time depends only on the minimal capacity, $\min(C_1, \ldots, C_n)$.

We notice that Theorem 7 implies the result of Theorem 6, since one may order the given service sequence $(a_1, \ldots, a_n)$ so that the resulting sequence has $a_{k_m} = \max(a_1, \ldots, a_n)$ as the first service time. For the latter case, we have $W_n^{(i)} = 0$ for each $n \ge 1$ and each i, $2 \le i \le n$, and $W_n^{(1)}$ being the waiting time for an M/D/1 system with service time a_{k_m}. The overall waiting-time distribution given by Theorem 6 subsequently follows.

7. An Example To illustrate the application of our results to communication networks, consider the network shown by Fig. 4 with channel capacities (in K bits/sec) as indicated there. Let the message length be $\beta = 1000$ bits/message (this is, for example, the packet length in the ARPA computer communication network). We wish to consider the overall average delays resulting when we transmit messages from v_1 to v_6 through a communication path, and obtain the path which yields the minimal overall delay.

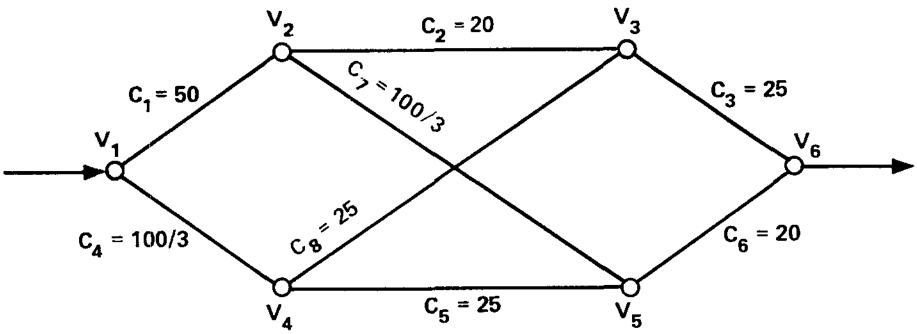


Figure 4. The Communication Network for the Example. Channel Capacities are in K bits/sec.

For the given network, there are four paths between v_1 and v_6, denoted by $\pi_1 = v_1v_4v_3v_6$, $\pi_2 = v_1v_2v_5v_6$, $\pi_3 = v_1v_2v_3v_6$, $\pi_4 = v_1v_4v_5v_6$. The overall transmission times $A \triangleq \sum a_i$, and the minimal capacity min C_i, are observed to be given for each path by:

$$A(\pi_1) = 0.11 \frac{\text{sec.}}{\text{mess.}}, \quad \min C_i(\pi_1) = 25 \frac{\text{K bits}}{\text{sec.}}$$

$$A(\pi_2) = 0.10 \frac{\text{sec.}}{\text{mess.}}, \quad \min C_i(\pi_2) = 20 \frac{\text{K bits}}{\text{sec.}}$$

$$A(\pi_3) = 0.11 \frac{\text{sec.}}{\text{mess.}}, \quad \min C_i(\pi_3) = 20 \frac{\text{K bits}}{\text{sec.}}$$

$$A(\pi_4) = 0.12 \frac{\text{sec.}}{\text{mess.}}, \quad \min C_i(\pi_3) = 20 \frac{\text{K bits}}{\text{sec.}}$$

The average delay $\gamma(\pi)$ is now readily calculated for each path using Eq. (85), for any arrival rate λ. To obtain the optimal path, we first observe that path π_2 yields always (for each λ) a smaller delay than paths π_3 and π_4. We need thus only compare π_1 and π_2.

By (85), we have

$$\gamma(\pi_1) = 0.11 + \frac{1}{2}\frac{0.04\lambda}{1-0.04\lambda}0.04,$$

$$\gamma(\pi_2) = 0.10 + \frac{1}{2}\frac{0.05\lambda}{1-0.05\lambda}0.05,$$

from which one readily concludes that:

$$\gamma(\pi_2) < \gamma(\pi_1) \text{ if } \lambda < \lambda_0$$

$$\gamma(\pi_1) < \gamma(\pi_2) \text{ if } \lambda_0 < \lambda < 25\,\frac{\text{mess.}}{\text{sec.}}$$

where $\lambda_0 \cong 9.33$ mess./sec. Hence, for incoming message rates less than λ_0 path π_2 will be utilized (since its overall transmission time is the shortest and λ is low enough so that the overall waiting time is not high), while path π_2 will be utilized for $\lambda > \lambda_0$. Clearly, $\gamma(\pi_2) = \infty$ for $\lambda \geq 20$ and $\gamma(\pi_1) = \infty$ for $\lambda \geq 25$. Other quantities for the communication paths are readily deduced as well, using our results.

8. Conclusions. We have solved for the steady-state distributions of the message waiting-times and overall delays along paths in a communication network. The average memory storage requirements at the stations have also been obtained. The following two points are readily observed.

1. Lemma 7, Eq. (89), holds for any incoming message point process. Hence, the distribution of the overall waiting time $S_m^{(n)}$ is deduced from (89), for any incoming message process (the message lengths are still assumed to be of fixed length). In particular, if the incoming messages follow the statistics of a renewal point process (i.e., $\{T_n^{(1)}, n \geq 1\}$ in (89) are i.i.d. random variables), results from GI/D/1 queueing system theory are used to obtain the limiting waiting time distribution and moments (see, for example, [1] Chapters II.5 and II.6).

2. The capacity assignment problem is readily solved. Thus, assume that the total capacity over the n-channel path is given, $\sum_{i=1}^{n} C_i = C$ (so that $C > n\lambda\beta$ bits/sec, to avoid infinite delays). We wish to find the values of the individual capacities, $i = 1, 2, \ldots, n$, so that the overall average delay γ is minimized. The delay γ is given by Eq. (85). Since the overall waiting-time depends only on the value of the minimal capacity, we wish to choose the largest possible value of the minimal capacity. Hence, the overall waiting time is minimized by choosing equal capacities, $C_i = C/n, i = 1, 2, \ldots, n$. The overall transmission time is $\sum_1^n a_i = \beta \sum_1^n C_i^{-1}$, which (by symmetry) is readily observed to be minimized as well by choosing equal channel capacities. Consequently, the average delay γ is minimized by choosing equal channel capacities along the communication path.

Further time-delay problems for communication networks are currently under investigation. The techniques and results presented here have been proved to be powerful tools for these studies. In particular, using time-delay and memory storage considerations, analysis and synthesis problems are being studied for more general message flows in communication networks.

References

[1] J. W. Cohen, *The Single Server Queue*, Amsterdam: North-Holland Publishing Company, 1969.

[2] N. U. Prabhu, *Queues and Inventories*, New York: Wiley, 1965.

[3] L. Takacs, *Introduction to the Theory of Queues*, New York: Oxford University Press, 1962.

[4] W. S. Jewell, "A Simple Proof of $L = \lambda W$", *Operations Research*, 15, pp. 1109–1116, 1967.

[5] T. M. Cover, "Broadcast Channels", *IEEE Trans. on Information Theory*, vol. 18, pp. 2–14, January, 1972.

[6] R. M. Gray and P. P. Bergmans, "Two Problems in Simultaneous Communications", *IEEE Transactions on Communications*, vol. 21, pp. 763–767, June, 1973.

[7] L. Kleinrock, *Communication Nets*, New York: McGraw-Hill, 1964.

[8] L. Kleinrock, "Scheduling, Queueing and Delays in Time-Shared Systems and Computer Networks", in *Computer—Communication Networks*, Ed. by N. Abramson and F. F. Kuo. Englewood Cliffs, New Jersey: Prentice-Hall, 1973.

[9] J. R. Jackson, "Networks of Waiting Lines", *Operations Research*, vol. 5, pp. 518–521, 1957.

[10] W. J. Gordon and G. F. Newell, "Closed Queueing Systems with Exponential Servers", *Operations Research*, vol. 15, pp. 254–265, 1967.

[11] H. Kobayashi, "Application of the Diffusion Approximation to Queueing Networks: Part I—Equilibrium Queue Distributions", *Proceedings of the ACM SIGME Symposium on Measurement and Evaluation* (Palo Alto), pp. 54–62, February, 1973.

[12] H. Frank and I. T. Frisch, *Communication, Transmission, and Transportation Networks*, Reading, Massachusetts: Addison-Wesley, 1971.

[13] H. D. Friedman, "Reduction Methods for Tandem Queueing Systems," *Operation Research*, Vol. 13, pp. 121–131, 1965.

[14] B. Avi-Itzhak, "A sequence of Service Stations with Arbitrary Input and Regular Service Times," *Management Science*, Vol. 11, No. 5, pp. 565–571, March 1965.

Received December 1973