UNE IMP

EHE ★ QMIPS ★ CWI

INR UER

ZAR UTO

Quantitative Modeling in Parallel Systems

Response Time Distributions in Packet-switched Banyan Networks

Peter G. Harrison Afonso de C. Pinto
Department of Computing
Imperial College
180 Queen's Gate, London SW7 2BZ

D W2.T1.11

22 September 1994

Distribution Level: All partners

Approved by: Peter G. Harrison

Abstract An approximation for the distribution of time delays experienced by tasks in finite-buffered multistage interconnection networks is presented. We focus on asynchronous packet-switched banyan networks which are important in ATM telecommunication systems as well as in parallel computer architectures although our method is applicable to any feed-forward network. The networks we consider are organised in a finite number of stages through each of which a task passes successively in its transmission. The method presented considers successive pairs of switches in a task's path, the queue (output) joined on arrival at the second having length dependent on that which existed previously on arrival at the first. The transient period between arrivals at successive nodes is modelled in two ways: by solving the appropriate Kolmogorov equations and by an approximation which considers the net change in the queue length at the second. The networks may be non-homogeneous in that different servers may have different rates and the arrival processes need not be identical. Numerical results are obtained for a range of network parameters and both analytical methods are compared with simulation results.

The QMIPS Project The QMIPS project is a collaborative research project supported by the CEC as ESPRIT-BRA project no 7269. It is being carried out by the following organisations: CWI (Amsterdam), EHEI (University of Paris V), Imperial College (London), INRIA (Sophia-Antipolis), University of Erlangen, University of Newcastle, University of Torino and University of Zaragoza.

1 Introduction

Recently, much attention has been given to the problem of determining response time distributions in queueing networks [2, 3, 9, 11, 12, 16, 18]. However, apart from a very few studies (e.g., [3, 9]), there has not been much research on the application of such results to multistage interconnection networks (MINs). Task-based performance measures of MINs, time delays in particular, have been determined, in general, in terms of their mean values, typically using Little's law together with certain resource-based measures such as mean queue-length and throughput of the subsystem under consideration. However, the mean value alone is frequently insufficient since it is possible to have a situation where the frequency in which the observed time delay falls outside a desirable interval is quite high (i.e., high variability), whereas the mean value itself is within an acceptable range.

In this paper, an approximation method for obtaining the response time of a tagged task in an aynchronous, finite-buffered, packet-switched banyan network is presented. The approach adopted follows closely that employed in [6] where each node's queue-length distribution is determined in terms of the distributions of the corresponding next stage nodes (i.e., those that are connected to it). The approximation is characterised by the fact that the method considers successive pairs of servers in a task's path, the queue joined on arrival at the second server having length dependent on that which existed previously on arrival at the first. A similar method, the Paired Center Approximation (PCA), applied to closed queueing networks with servers with negative exponential service time distributions was proposed by Harrison [4].

Although the analysis is tailored to MINs, the approach can be applied to more general open networks with no feedback and finite capacity nodes, once each node's steady-state queue length distribution is known. These results are obtained from an approximation algorithm for the queue length probability distribution at each switch of the network for the case where every switch (including those located at the first stage) are assumed to be finite [15]. This results in two algorithms which are the fruition of the preliminary work presented by the authors in [14].

MINs occur commonly in large scale ATM networks to support B-ISDN communication. Banyan MINs provide full connectivity between a set of sources and a set of destinations and in a B-ISDN environment, support several different types of traffic concurrently, e.g. data, voice and video. Consequently models must be able to represent various traffic characteristics such as burstiness which corresponds to (at least) video which has to be batched. Burstiness is often well represented by the generalised exponential (GE) interarrival time distribution which we use in our algorithm. However, ATM works on the basis of fixed length packets which are often analysed using discrete-time models. Nevertheless, a whole network will not operate with lock-step synchronisation and a single node transmitting fixed length packets can be considered as simply having deterministic service times. However, the effect of burstiness together with blocking means that the *effective* service times will be random and so continuous time methods become appropriate.

This paper is organized as follows. **Section 2** describes the model assumptions and the approach. The algorithms are given in **Section 3** and **Section 4** presents numerical results and validation with respect to simulation. Conclusions are given in **Section 5**.

2 Model Specification and Notation

We consider a regular buffered banyan network consisting of M stages and L levels in a packet-switching environment. The basic component is the two-input, two-output switching element (2×2 crossbar switch). For an $N \times N$ network (i.e., N inputs and N outputs), with N a power of 2, it follows that $M = \log_2 N$ and $L = N/2$. An 8×8 banyan network is illustrated in Figure 1. We shall call switch–(i, j) the switching element located at the ith level of the jth stage of the network. The upper input pin of each switch is labeled "input–0" and the lower input pin, "input–1", and similarly for output pins.

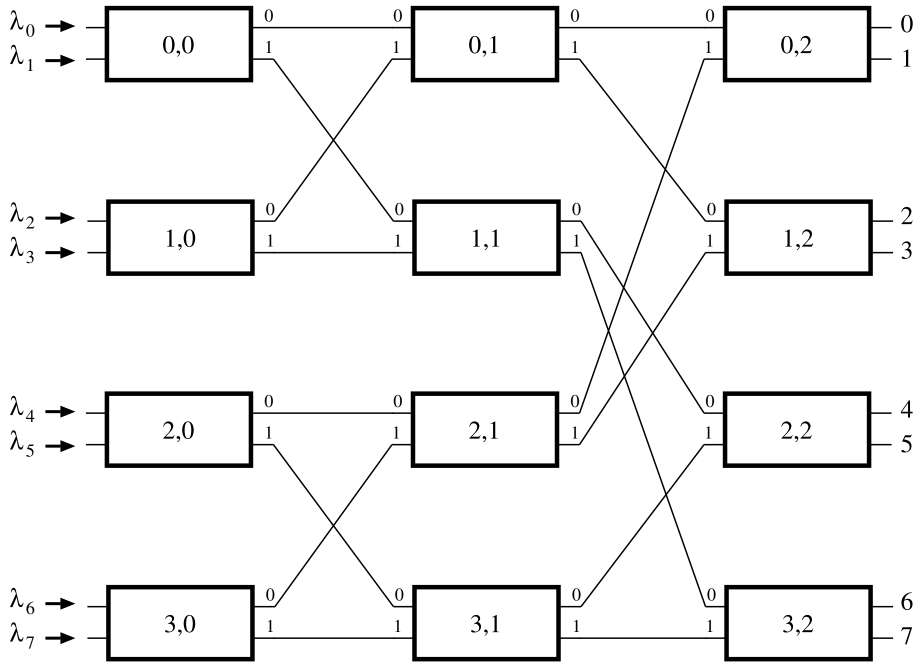


Figure 1: 8×8 Banyan Network.

Let $\mathbf{R} = \{r_{ij}\}_{N \times N}$ be the routing probability matrix where r_{ij} is the probability that a task originating at input pin i (first stage) has output pin j (last stage) as its destination, where i, j are numbered from 0 at the top as above. Tasks are assumed to arrive at the switches located in the first stage as a Poisson process with arrival rate λ_i at input pin i. Let then, $\boldsymbol{\lambda} = [\lambda_0, \lambda_1, \ldots, \lambda_{N-1}]$ be the external arrival rate vector. It is assumed that each switch has a single finite buffer with the exception of the first stage switches which are assumed to have a single buffer with infinite size. Let $\boldsymbol{\mathcal{C}}$ be the switch-capacity matrix where the (i,j)th element is given by $\mathcal{C}_{ij}$, the capacity, or buffer size of switch–(i,j), $0 \leq i \leq L-1$, $0 \leq j \leq M-1$.

Incoming tasks join a switch's single buffer and each corresponding output picks the next task requesting its service. The service at each output is assumed to be FCFS. In our model each output pin can have any service time distribution. However, for analytical and computational tractability, we use generalised exponential (GE) distributions to approximate general distributions with known first two moments. This type of approximation has been found to be a robust and versatile tool for the analysis of general queueing systems, particularly when bursty traffic is present, [8]. Let $S_{ijx}(t)$ be the service time distribution at output–x of switch–(i,j), $x = 0, 1$, $0 \leq i \leq L-1$, $0 \leq j \leq M-1$. The distribution $S_{ijx}(t)$ is defined in terms of the parameters μ_{ijx} and θ_{ijx} as follows:

$$S_{ijx}(t) = 1 - \theta_{ijx} e^{-\mu_{ijx} t},$$

with LST $S^*_{ijx}(s)$ given by:

$$\begin{aligned} S^*_{ijx}(s) &= (1 - \theta_{ijx}) + \theta_{ijx} \frac{\mu_{ijx}}{s + \mu_{ijx}} \\ & \qquad 0 \leq i \leq L-1, \;\; 0 \leq j \leq M-1, \;\; x = 0, 1. \end{aligned} \tag{1}$$

We define $\mathbf{S}$ as the 'service-time matrix' whose (i, j)th element, s_{ij}, is given by the two ordered-pairs $[(\mu_{ij0}, \theta_{ij0}), (\mu_{ij1}, \theta_{ij1})]$ which give the parameters of the service time distributions of outputs 0 and 1 of switch–(i, j), respectively.

Every task that arrives at one of the inputs of the first stage, goes to its output destination (last stage) passing through one switch at every stage of the network in a tandem fashion. Tasks are prevented from entering a full switch according to *Blocking After Service* (BAS), i.e., service is suspended at the output from which the task is trying to leave.

The approximation algorithms presented in **Section 3** produce as result the time delay distribution of a tagged task arriving at one of the N external inputs of a MIN and passing along a path of M switches determined according to the task's destination output. The time delay of a task is defined to be the time elapsed between its arrival instant at one of the external inputs of the network (stage–0) and its departure from one the network's external outputs (stage–$(M-1)$). Note that the total number of paths in a $N \times N$ banyan network is N^2 (one path for each pair external input-external output).

Without loss of generality, for notational conciseness, we assume that the queues in the path taken by the tagged task are numbered $0, 1, \ldots, M-1$. We call "conjugate queue (output)" a queue (output) which is not in the tagged task's path but is in a switch whose other queue (output) belongs to the said path.

We label by q_k the queue in the tagged task's path located in stage–k of the network, and $\overline{q}_k$ its conjugate queue, $0 \leq k \leq M-1$. Such labels are used in Figure 2 which shows a schematic representation of a path. The shadowed part of Figure 2 indicates the actual path taken by the tagged task. Although similar to a tandem network, the path can not be analysed by standard

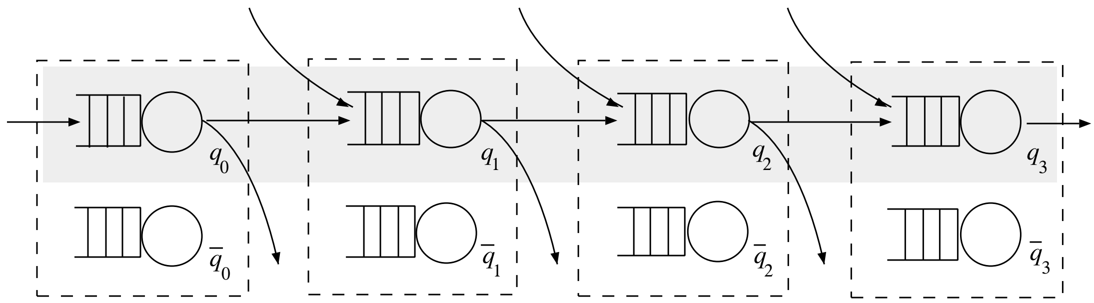


Figure 2: Tagged task's path.

modelling techniques used for this type of network. This is due to the possibility of external arrivals at each node originating from sources other than the previous queue in the path (with the exception of the first node), possible departures from intermediate nodes of tasks that are ahead of the tagged task (thus reducing its overall delay) and the fact that the state of each conjugate queue does affect the arrival rates at its corresponding node (since each switch has finite capacity). These issues are discussed in detail in the next section. Note that although all queues in the tagged task's path depicted in the representation shown in Figure 2 are located at the upper output of their respective switches, that does not necessarily correspond to their actual positions in the network.

Let $\mathcal{S}_i$ be the random variable for the sojourn time of the tagged task at queue q_i, $i = 0, 1, \ldots, M-1$. Defining $T_0 = 0$ and $T_i = \mathcal{S}_0 + \mathcal{S}_1 + \ldots + \mathcal{S}_{i-1}$, for $1 \leq i \leq M$, it follows that $\{T_i \mid 0 \leq i \leq M-1\}$ is the set of arrival instants of the tagged task at queues $q_0, q_1, \ldots, q_{M-1}$, and T_M is the random variable for the instant the tagged task leaves queue q_{M-1} (and consequently the network).

We define the queue length's process as $\{Q_i(t) \mid 0 \leq i \leq M-1, t \geq 0\}$, assumed to be in steady-state initially, where $Q_i(t)$ is the length of queue q_i at time t. Similarly, $\{\overline{Q}_i(t) \mid 0 \leq$

$i \leq M-1, t \geq 0\}$ is defined as the conjugate queue length's process, where $\overline{Q}_i(t)$ is the length of queue $\overline{q}_i$ at time t. Then, it follows that $Q_i(T_i)$ is the queue length faced by the tagged task on arrival at q_i (including the tagged task). Let $\mathbf{P}\{Q_i = n, \overline{Q}_i = \overline{n}\}$ be the probability that, in equilibrium, queue q_i has length n and queue $\overline{q}_i$ has length $\overline{n}$, where $Q_i = Q_i(\tau)$ and $\overline{Q}_i = \overline{Q}_i(\tau)$, $0 \leq i \leq M-1$, for some $\tau \geq 0$. Then $\mathbf{P}\{Q_i = n\}$ and $\mathbf{P}\{\overline{Q}_i = \overline{n}\}$ are the marginal probabilities that in equilibrium, queue q_i has length n and queue $\overline{q}_i$ has length $\overline{n}$, respectively, for $0 \leq i \leq M-1$. Results for the distributions $\mathbf{P}\{Q_i = n, \overline{Q}_i = \overline{n}\}$, $0 \leq i \leq M-1$, are those obtained in [6].

In what follows, we analyse the joint probability distribution of $(T_i, Q_i(T_i))$, for $0 \leq i \leq M-1$. We define the distributions $D_i(t,n)$, for $0 \leq i \leq M-1$, as:

$$D_i(t,n) = \mathbf{P}\{T_i \leq t,\ Q_i(T_i) = n\}, \quad 1 \leq n \leq \mathcal{C}_i, \quad t \geq 0. \tag{2}$$

Note that the minimum value that n can assume is 1 since it includes the tagged task itself. From the above definition, and from the fact that $T_0 = 0$, we have:

$$D_0(t,n) = \mathbf{P}\{T_0 \leq t,\ Q_0(T_0) = n\} = \mathbf{P}\{\tilde{Q}_0 = n\}, \tag{3}$$

where

$$\mathbf{P}\{\tilde{Q}_i = n\} = \frac{\mathbf{P}\{Q_i = n\}}{1 - \mathbf{P}\{Q_i = 0\}}, \quad 1 \leq n \leq \mathcal{C}_i, \quad 0 \leq i \leq M-1, \tag{4}$$

which in terms of the joint probability distibution of Q_i and $\overline{Q}_i$ is qiven by:

$$\mathbf{P}\{\tilde{Q}_i = n\} = \frac{\sum_{k=0}^{\mathcal{C}_i - n} \mathbf{P}\{Q_i = n, \overline{Q}_i = k\}}{1 - \sum_{k=0}^{\mathcal{C}_i} \mathbf{P}\{Q_i = 0, \overline{Q}_i = k\}}, \ \ 1 \leq n \leq \mathcal{C}_i, \ \ 0 \leq i \leq M-1,$$

i.e., $\mathbf{P}\{\tilde{Q}_i = n\}$ is the equilibrium probability that queue q_i has length n, conditional on a nomempty queue. Thus, Equation 3 states that $D_0(t,n)$ is equal to the equilibrium probability that upon arrival the tagged task finds n tasks (itself included) already in q_0.

In the next section two approximate algorithms for the time delay distribution of a tagged task traversing a buffered MIN along a given path are derived as recurrence formulas.

3 Two Approximate Algorithms for Time Delay Distributions

In this paper, the expression "effective service time" is used in the same context as in [6], i.e., to indicate that the blocking effect is taken into consideration by adding the eventual blocking time to the actual service time. Let $B_i(t)$ and $\overline{B}_i(t)$ be the effective service time distributions at q_i and $\overline{q}_i$, respectively, for $0 \leq i \leq M-1$.

The total delay experienced by the tagged task at each node in its path has at most three components: (i) the tagged task's effective service time; (ii) the remaining effective service time of the task at the head of the queue (if any) at the instant the tagged task arrives at the considered node; and (iii) the sum of all "full" effective service times of the tasks (if any) positioned immediately after the task at the head of the queue and ahead of the tagged task. Obviously, the total delay will always include (i). If, upon arrival, the tagged task finds exactly one task already in the queue (not considering the tagged task itself), the total delay will also include (ii). However, if it finds more than one task then the total delay will include both (ii) and (iii). Now, let $\mathcal{F}^*_{\mu,\theta,r}(s)$ be the LST of $\mathcal{F}_{\mu,\theta,r}(t)$, the distribution of the tagged task's sojourn time at each queue in its path, given that the tagged task finds r tasks already in the queue upon arrival (itself included) where the parameters of the GE approximation of the effective service time distribution are μ and θ. It then follows that $\mathcal{F}^*_{\mu,\theta,r}(s)$ is given by:

$$\mathcal{F}^*_{\mu,\theta,r}(s) = \begin{cases} [(1-\theta) + \dfrac{\theta\mu}{\mu+s}]^{r-1}(\dfrac{\mu}{\mu+s}), & r > 1, \\ (1-\theta) + \dfrac{\theta\mu}{\mu+s}, & r = 1. \end{cases}$$

According to the argument presented in the previous paragraph, for $r = 1$, the tagged task's sojourn time is simply its effective service time. For $r > 1$, the sojourn time is given by the sum of a remaining effective service time (therefore the $\frac{\mu}{\mu+s}$ term), and $r-1$ effective service times (the tagged task's included). Anti-transforming $\mathcal{F}^*_{\mu,\theta,r}(s)$ we obtain $f_{\mu,\theta,r}(t)$, the probability density function of $\mathcal{F}_{\mu,\theta,r}(t)$, given bellow:

$$f_{\mu,\theta,r}(t) = \begin{cases} \mu e^{-\mu t} \displaystyle\sum_{k=0}^{r-1} \binom{r-1}{k} (1-\theta)^k \frac{(\theta\mu t)^{r-1-k}}{(r-1-k)!}, & r > 1, \\ \theta\mu e^{-\mu t}, & r = 1. \end{cases}$$

We now derive a recurrence formula for the distribution $D_i(t,n) = \mathbf{P}\{T_i \le t,\ Q_i(T_i) = n\}$, $0 < i \le M-1$. We initially apply the Law of Total Probability to Equation 2 which results in:

$$\begin{aligned} D_i(t,n) \;=\; & \sum_{\mathbf{r}_i \in \mathcal{D}_i} \mathbf{P}\{T_i \le t,\ Q_i(T_i) = n,\ Q_j(T_j) = r_j\ (0 \le j \le i-1), \\ & Q_i(T_{i-1}) = r_i,\ \overline{Q}_i(T_{i-1}) = \overline{r}_i\}, \end{aligned} \tag{5}$$

where $\mathcal{D}_i = \{\mathbf{r}_i = (r_0, r_1, r_2, \ldots, r_{i-1}, r_i, \overline{r}_i) \,|\, 1 \le r_j \le \mathcal{C}_j\ (0 \le j \le i-1),\ 0 \le r_i \le \mathcal{C}_i,\ 0 \le \overline{r}_i \le \mathcal{C}_i - r_i\}$. Note that the dimension of vector $\mathbf{r}_i$ depends on the value of i. For example, the vectors belonging to the set $\mathcal{D}_1$ have only three elements (i.e., they are vectors of the type $\mathbf{r}_1 = (r_0, r_1, \overline{r}_1)$) while the vectors belonging to $\mathcal{D}_{M-1}$ have $M+1$ elements (i.e., they are vectors of the type $\mathbf{r}_{M-1} = (r_0, r_1, \ldots, r_{M-2}, r_{M-1}, \overline{r}_{M-1})$).

Equation 5 can be re-written as:

$$\begin{aligned} D_i(t,n) \;=\; & \sum_{\mathbf{r}_i \in \mathcal{D}_i} \mathbf{P}\{Q_i(T_i) = n \,|\, T_i \le t,\ Q_j(T_j) = r_j\ (0 \le j \le i-1), \\ & Q_i(T_{i-1}) = r_i,\ \overline{Q}_i(T_{i-1}) = \overline{r}_i\} \\ & \mathbf{P}\{T_i \le t,\ Q_j(T_j) = r_j\ (0 \le j \le i-1),\ Q_i(T_{i-1}) = r_i, \\ & \overline{Q}_i(T_{i-1}) = \overline{r}_{i-1}\}. \end{aligned} \tag{6}$$

By replacing the random variable T_i with $T_{i-1} + \mathcal{S}_{i-1}$ (see T_i's definition) in Equation 6, we require the convolution of the distributions for T_{i-1} and $\mathcal{S}_{i-1}$. Conditioning $\mathcal{S}_{i-1} \le v$ on the remaining events, as well as switching the summation with the integral (which appears due to the convolution) we obtain:

$$\begin{aligned} D_i(t,n) \;=\; & \sum_{\mathbf{r}_i \in \mathcal{D}_i} \int_0^t \mathbf{P}\{Q_i(T_i) = n \,|\, T_{i-1} \le t - v,\ \mathcal{S}_{i-1} = u, \\ & Q_j(T_j) = r_j\ (0 \le j \le i-1),\ Q_i(T_{i-1}) = r_i,\ \overline{Q}_i(T_{i-1}) = \overline{r}_i\} \\ & \mathbf{P}\{T_{i-1} \le t - v,\ Q_j(T_j) = r_j\ (0 \le j \le i-1), \\ & Q_i(T_{i-1}) = r_i,\ \overline{Q}_i(T_{i-1}) = \overline{r}_i\} \\ & d\mathbf{P}\{\mathcal{S}_{i-1} \le v \,|\, T_{i-1} \le t - u,\ Q_j(T_j) = r_j\ (0 \le j \le i-1), \\ & Q_i(T_{i-1}) = r_i,\ \overline{Q}_i(T_{i-1}) = \overline{r}_i\}. \end{aligned} \tag{7}$$

Now, from the Markov property, and from the fact that, under FCFS service discipline, $\mathcal{S}_{i-1}$ depends only on $Q_{i-1}(T_{i-1})$, $Q_i(T_{i-1})$ and $\overline{Q}_i(T_{i-1})$, as well as making the approximating assumption that $Q_i(T_i)$ is independent of $Q_j(T_j)$, $(0 \leq j < i-1)$ (i.e., assuming that $Q_i(T_i)$ depends on $Q_{i-1}(T_{i-1})$ but not on $Q_{i-2}(T_{i-2}), Q_{i-3}(T_{i-3}), \ldots, Q_0(T_0)$), Equation 7 reduces to:

$$
\begin{aligned}
D_i(t,n) \;=\; & \sum_{\mathbf{r}_i \in \mathcal{D}_i} \int_0^t \mathbf{P}\{Q_i(T_i) = n \,|\, \mathcal{S}_{i-1} = v,\ Q_{i-1}(T_{i-i}) = r_{i-1}, \\
& Q_i(T_{i-1}) = r_i,\ \overline{Q}_i(T_{i-1}) = \overline{r}_i\} \\
& \mathbf{P}\{T_{i-1} \leq t - v,\ Q_{i-1}(T_{i-i}) = r_{i-1},\ Q_i(T_{i-1}) = r_i,\ \overline{Q}_i(T_{i-1}) = \overline{r}_i\} \\
& d\mathbf{P}\{\mathcal{S}_{i-1} \leq v \,|\, Q_{i-1}(T_{i-i}) = r_{i-1}, Q_i(T_{i-1}) = r_i,\ \overline{Q}_i(T_{i-1}) = \overline{r}_i\}. \qquad (8)
\end{aligned}
$$

From Equation 8 it is clear that in order to calculate $D_i(t,n)$, it is first necessary to evaluate the following three probabilities:

1. $\mathbf{P}\{Q_i(T_i) = n \,|\, \mathcal{S}_{i-1} = v,\ Q_{i-1}(T_{i-i}) = r_{i-1},\ Q_i(T_{i-1}) = r_i,\ \overline{Q}_i(T_{i-1}) = \overline{r}_i\}$.
2. $\mathbf{P}\{T_{i-1} \leq t - v,\ Q_{i-1}(T_{i-i}) = r_{i-1},\ Q_i(T_{i-1}) = r_i,\ \overline{Q}_i(T_{i-1}) = \overline{r}_i\}$.
3. $\mathbf{P}\{\mathcal{S}_{i-1} \leq v \,|\, Q_{i-1}(T_{i-i}) = r_{i-1},\ Q_i(T_{i-1}) = r_i,\ \overline{Q}_i(T_{i-1}) = \overline{r}_i\}$.

We initially obtain the probability, $\mathbf{P}\{T_{i-1} \leq t - v,\ Q_{i-1}(T_{i-i}) = r_{i-1},\ Q_i(T_{i-1}) = r_i,\ \overline{Q}_i(T_{i-1}) = \overline{r}_i\}$ given above as follows:

$$
\begin{aligned}
& \mathbf{P}\{T_{i-1} \leq t - v,\ Q_{i-1}(T_{i-i}) = r_{i-1},\ Q_i(T_{i-1}) = r_i,\ \overline{Q}_i(T_{i-1}) = \overline{r}_i\} = \\
& \quad \mathbf{P}\{Q_i(T_{i-i}) = r_i,\ \overline{Q}_i(T_{i-1}) = \overline{r}_i \,|\, T_{i-1} \leq t - v,\ Q_{i-1}(T_{i-i}) = r_{i-1}\} \cdot \\
& \quad \mathbf{P}\{T_{i-1} \leq t - v,\ Q_{i-1}(T_{i-i}) = r_{i-1}\}. \qquad (9)
\end{aligned}
$$

As mentioned previously, the network is assumed to be initially in equilibrium. Therefore, the queue length distributions at queues q_i and $\overline{q}_i$ at any time prior to the arrival of the tagged task to q_i (i.e., any time prior to T_i), do not depend on the queue length of any other queue located before q_i on the tagged task's path. It then follows that

$$
\begin{aligned}
& \mathbf{P}\{Q_i(T_{i-i}) = r_i,\ \overline{Q}_i(T_{i-1}) = \overline{r}_i \,|\, T_{i-1} \leq t - v,\ Q_{i-1}(T_{i-i}) = r_{i-1}\} = \\
& \quad \mathbf{P}\{Q_i(T_{i-i}) = r_i,\ \overline{Q}_i(T_{i-1}) = \overline{r}_i \,|\, T_{i-1} \leq t - v\}.
\end{aligned}
$$

The fact that queues q_i and $\overline{q}_i$ are assumed to be in equilibrium at any time t, $t < T_i$, also implies that the queue length distributions at queues q_i and $\overline{q}_i$ do not depend on t itself, for $t < T_i$ (and consequently do not depend on T_{i-1}). Thus,

$$
\begin{aligned}
& \mathbf{P}\{Q_i(T_{i-i}) = r_i,\ \overline{Q}_i(T_{i-1}) = \overline{r}_i \,|\, T_{i-1} \leq t - v\} = \\
& \quad \mathbf{P}\{Q_i = r_i,\ \overline{Q}_i = \overline{r}_i\}.
\end{aligned}
$$

Substituting the above results into Equation 9 and recalling the definition of $D_i(t,n)$ we obtain:

$$
\begin{aligned}
& \mathbf{P}\{T_{i-1} \leq t - v,\ Q_{i-1}(T_{i-i}) = r_{i-1},\ Q_i(T_{i-1}) = r_i,\ \overline{Q}_i(T_{i-1}) = \overline{r}_i\} = \\
& \quad \mathbf{P}\{Q_i = r_i,\ \overline{Q}_i = \overline{r}_i\}\ D_{i-1}(t - v, r_{i-1}). \qquad (10)
\end{aligned}
$$

The third probability, $\mathbf{P}\{\mathcal{S}_{i-1} \leq v \,|\, Q_{i-1}(T_{i-i}) = r_{i-1},\ Q_i(T_{i-1}) = r_i,\ \overline{Q}_i(T_{i-1}) = \overline{r}_i\}$, can be approximated by $\mathbf{P}\{\mathcal{S}_{i-1} \leq v \,|\, Q_{i-1}(T_{i-i}) = r_{i-1}\}$ by considering the service time distribution at queue q_{i-1} as its effective service time distribution, which takes into account the effect of the queue length distributions at queues q_i and $\overline{q}_i$ on the tagged task's sojourn interval at queue q_{i-1}. Therefore,

$$
\mathbf{P}\{\mathcal{S}_{i-1} \leq v \,|\, Q_{i-1}(T_{i-i}) = r_{i-1}\} = \mathcal{F}_{b_{i-1},\beta_{i-1},r_{i-1}}(v), \qquad (11)
$$

where (b_{i-1}, β_{i-1}) are the parameters of the GE approximation of the effective service time distribution at queue q_{i-1}.

Finally, we must obtain an expression for the conditional probability $\mathbf{P}\{Q_i(T_i) = n \,|\, \mathcal{S}_{i-1} = v,\ Q_{i-1}(T_{i-i}) = r_{i-1},\ Q_i(T_{i-1}) = r_i,\ \overline{Q}_i(T_{i-1}) = \overline{r}_i\}$, i.e., the probability that upon arrival, the tagged task finds n tasks (itself included) at queue q_i given that: (i) at time T_{i-i}, the number of tasks in the previous queue in its path, i.e., q_{i-1}, is r_{i-1}; (ii) at time T_{i-i}, the number of tasks in $\overline{q}_{i-1}$ is $\overline{r}_{i-1}$; and (iii) the transient sojourn interval of the tagged task at queue q_{i-1} is v (i.e., $T_i - T_{i-i} = v$). Figure 3 illustrates the state of the pair of switches considered (in terms of queue length at each relevant queue) at times T_i and T_{i-1}. The probability $\mathbf{P}\{Q_i(T_i) = n \,|\, \mathcal{S}_{i-1} = v,\ Q_{i-1}(T_{i-i}) = r_{i-1},\ Q_i(T_{i-1}) = r_i,\ \overline{Q}_i(T_{i-1}) = \overline{r}_i\}$ is the probability that the subsystem composed of the two switches, is in the state depicted by the left part of Figure 3 at time T_{i-1}, and goes to the state shown on the right part of that figure at T_i (for all $\overline{n}$, $0 \leq \overline{n} \leq \mathcal{C}_i - n$). The shaded tasks in Figure 3 indicate the position of the tagged task at times T_i and T_{i-1}, respectively.

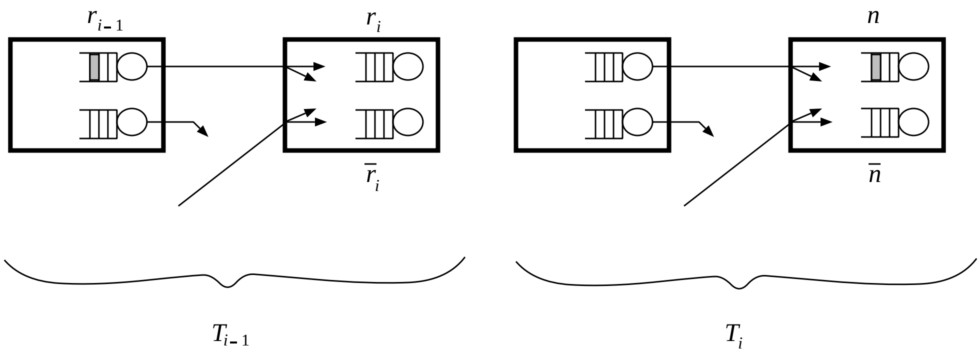


Figure 3: Two consecutive switches in the tagged task's path.

In the next section we derive the Kolmogorov forward differential-difference equations for a simplified version of the subsystem shown twice in Figure 3 (for times T_i and T_{i-1}). The solution of the said equations, given appropriate boundary conditions, produces the time-dependent result (in terms of v) for the probability $\mathbf{P}\{Q_i(T_i) = n \,|\, \mathcal{S}_{i-1} = v,\ Q_{i-1}(T_{i-1}) = r_{i-1},\ Q_i(T_{i-1}) = r_i,\ \overline{Q}_i(T_{i-1}) = \overline{r}_i\}$. Before we proceed though, we introduce yet another term which will be used in the remainder of this paper. We call switch–i, $0 \leq i \leq M-1$, the switch where queues q_i and $\overline{q}_i$ are located, with q_i, as mentioned earlier, being a queue in the tagged task's path, and $\overline{q}_i$, its conjugate.

3.1 Transient Analysis: Kolmogorov Equations

In order to analyse the subsystem composed of queues q_{i-1}, q_i and $\overline{q}_i$ during the transient period $[T_{i-1}, T_i)$, we use a simplified version of this subsystem as shown in Figure 4. The simplication is basically the replacement of queue q_{i-1} by an external source of arrivals (to switch–i) with the properties that it generates only a finite number of arrivals (in fact, a total of r_{i-1} arrivals, the number of tasks in queue q_{i-1} at time T_{i-1}), and that the last one of those arrivals occurs at time T_i. This "finite external source" is represented by a square on the top-left corner of Figure 4. The interarrival times of the arrival process originating from this source are distributed according to a generalised exponential distribution with parameters equal to those of the GE approximation of the effective service time distribution at queue q_{i-1} (i.e., b_{i-1} and β_{i-1}) obtained in [6]. Except for the last task in this finite external source (which is the tagged task as we are assuming a FCFS service discipline in all outputs) all tasks can go to either q_i or $\overline{q}_i$ with probabilities u and

$\overline{u}$, respectively, with $u + \overline{u} = 1$.

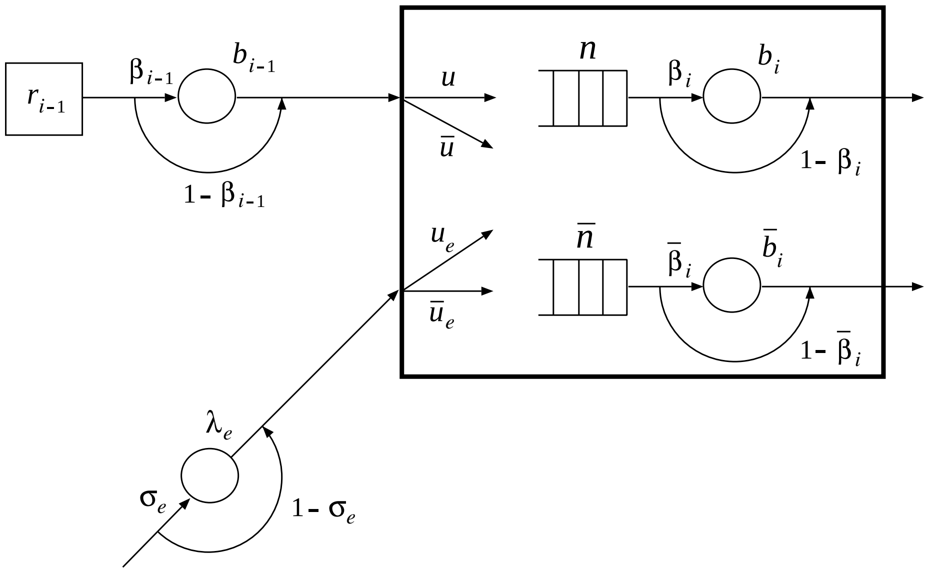


Figure 4: Transient Analysis.

The interarrival times of the external arrival process at the other input of switch–i are also distributed according to a generalised exponential distribution, in this case with parameters equal to those of the GE approximation of the effective arrival time distribution at that particular input which were obtained in [6], denoted here by (λ_e, σ_e) (see Figure 4). All tasks originating from this external source can go to either q_i or $\overline{q}_i$ with probabilities u_e and $\overline{u}_e$, respectively, with $u_e + \overline{u}_e = 1$. The probabilities u, $\overline{u}$, u_e and $\overline{u}_e$ can be calculated from the effective arrival rate at each input pin of switch–i, from the routing probability matrix $\mathbf{R}$ and the network's topology.

As in this transient analysis, switch–i is considered in isolation, the parameters of the GE approximations for the service time distributions at queues q_i and $\overline{q}_i$ are the parameters of the *effective* service times at those queues which were also obtained in [6], denoted here by (b_i, β_i) and $(\overline{b}_i, \overline{\beta}_i)$, respectively. By considering effective service times instead of actual service times we are taking into account the eventual blocking effect caused by the switches connected to switch–i which are located at the next stage of the network.

We call the "arrival phase" of an arriving process whose interarrival times are distributed according to a GE distribution, the exponential stage of that GE. Using this terminology, it can be said that a task enters the arrival phase of, say, the external source shown in Figure 4 with probability σ_e, and skips it (thus originating a bulk arrival) with probability $1 - \sigma_e$.

Now, considering the system shown in Figure 4, let the probability $p_i(n, \overline{n}, r, t)$ be defined as:

$$\begin{aligned} p_i(n, \overline{n}, r, t) \;\; = \;\; & \mathbf{P}\{Q_i(t) = n,\ \overline{Q}_i(t) = \overline{n},\ Q_{i-1}(t) = r\}, \\ & t \in [T_{i-1}, T_i], \\ & 0 \leq n \leq \mathcal{C}_i,\ \ 0 \leq \overline{n} \leq \mathcal{C}_i - n,\ \ 0 \leq r \leq r_{i-1}, \\ & 0 \leq i \leq M-1. \end{aligned} \qquad (12)$$

The time-dependent balance equations for the underlying Markov process are:

$$p_i(n, \overline{n}, r, t + \Delta t) =$$

$$p_i(n, \overline{n}, r, t)\{1 - [I_{n+\overline{n}<\mathcal{C}_i}\,(b_{i-1} \sum_{k=1}^{\min(r, \mathcal{C}_i - n - \overline{n})} (1 - \beta_{i-1})^{k-1} \beta_{i-1}^{I_{k<r}} + \lambda_e \sigma_e \sum_{k=1}^{\mathcal{C}_i - n - \overline{n}} (1 - \sigma_e)^{k-1})$$

$$
\begin{aligned}
&\qquad\qquad + I_{n>0}\, b_i \sum_{k=1}^{n} (1-\beta_i)^{k-1} \beta_i^{I_{k<n}} \\
&\qquad\qquad + I_{\overline{n}>0}\, \overline{b}_i \sum_{k=1}^{\overline{n}} (1-\overline{\beta}_i)^{k-1} \overline{\beta}_i^{I_{k<\overline{n}}}] \} \, \Delta t \\
&+ b_{i-1}\, u\, \beta_{i-1}^{I_{r>0}} \beta_i\, (1-\beta_{i-1})^{n-1} \cdot \\
&\qquad \sum_{d=0}^{\mathcal{C}_{i-1}-r-n} p_i(0,\overline{n},r+n+d,t)(1-\beta_{i-1})^d (1-\beta_i)^d\, I_{r+n+d \le \mathcal{C}_{i-1}}\, \Delta t \\
&+ p_i(0,\overline{n},r,t) \left(\frac{\lambda_e u_e \sigma_e \beta_i (1-\sigma_e)^{n-1}}{\sigma_e + \beta_i - \sigma_e \beta_i} \right) \Delta t \\
&+ b_{i-1}\, \overline{u}\, \beta_{i-1} \overline{\beta}_i\, (1-\beta_{i-1})^{\overline{n}-1} \cdot \\
&\qquad \sum_{d=0}^{\mathcal{C}_{i-1}-r-1-\overline{n}} p_i(n,0,r+\overline{n}+d,t)(1-\beta_{i-1})^d (1-\overline{\beta}_i)^d\, I_{r+\overline{n}+d \le \mathcal{C}_{i-1}}\, \Delta t \\
&+ p_i(n,0,r,t) \left(\frac{\lambda_e \overline{u}_e \sigma_e \overline{\beta}_i (1-\sigma_e)^{\overline{n}-1}}{\sigma_e + \overline{\beta}_i - \sigma_e \overline{\beta}_i} \right) \Delta t \\
&+ I_{n>0} \sum_{k=1}^{n-1} [\, p_i(n-k,\overline{n},r+k,t)\, I_{r+k \le \mathcal{C}_{i-1}}\, b_{i-1}\beta_{i-1}(1-\beta_{i-1})^{k-1} u \\
&\qquad\qquad + p_i(n-k,\overline{n},r,t)\, \lambda_e \sigma_e (1-\sigma_e)^{k-1} u_e] \, \Delta t \\
&+ I_{\overline{n}>0} \sum_{k=1}^{\overline{n}-1} [\, p_i(n,\overline{n}-k,r+k,t)\, I_{r+k \le \mathcal{C}_{i-1}}\, b_{i-1}\beta_{i-1}(1-\beta_{i-1})^{k-1} \overline{u} \\
&\qquad\qquad + p_i(n,\overline{n}-k,r,t)\, \lambda_e \sigma_e (1-\sigma_e)^{k-1} \overline{u}_e] \, \Delta t \\
&+ I_{n+\overline{n}<\mathcal{C}_i} \sum_{k=1}^{\mathcal{C}_i-n-\overline{n}} [\, p_i(n+k,\overline{n},r,t) b_i\, \beta_i\, (1-\beta_i)^{k-1} \\
&\qquad\qquad + p_i(n,\overline{n}+k,r,t) \overline{b}_i\, \overline{\beta}_i\, (1-\overline{\beta}_i)^{k-1}] \, \Delta t \\
&+ o(\Delta t), \\
&\qquad\qquad 0 \le r \le r_{i-1}, \;\; 0 \le n \le \mathcal{C}_i, \;\; 0 \le \overline{n} \le \mathcal{C}_i - n,
\end{aligned}
\tag{13}
$$

where I_x is the "indicator" function defined by:

$$
I_x = \begin{cases} 1 & \text{if } x \text{ is True,} \\ 0 & \text{otherwise.} \end{cases}
$$

We make the assumption that all tasks in an arriving batch (from either external source) go to the same queue in switch–i (i.e., they all go to either q_i or $\overline{q}_i$). In other words, tasks arriving in the same batch do not go to different queues. Therefore, any batch including the tagged task must go to q_i. The indicator function I_x is included in the above equations in order to suppress invalid transitions from states with negative valued queue lengths.

Now, the first term of the right-hand side of Equation 13 gives the probability that the system is in state $(n,\overline{n},r)$ at time t and that during a time interval Δt (i.e., from instant t to $t+\Delta t$) it does not leave state $(n,\overline{n},r)$. This is, of course, given by $1 - \mathbf{P}\{$there is "at least" one transition out of state $(n,\overline{n},r)$ during $\Delta t\}$. This probability is composed of four terms. The first one gives the probability that an arriving batch from the finite external source which henceforth we shall refer to as "the finite source," has size k. Note that k can assume values ranging from 1 to $\min(r, \mathcal{C}_i - n - \overline{n})$, where r is the size of the finite source at time t. This is due to the fact that the maximum number of tasks that can arrive from the finite source is either the remaining available positions in switch–i at time t, $\mathcal{C}_i - n - \overline{n}$, or the size of the finite source at time t, r, whichever is the smallest. Note also that the size of the finite source decreases from r_{i-1} at time T_{i-1} to 0 at time T_i since, during this interval, all tasks up to the tagged task (included) do join switch–i and, as it is assumed that the service discipline at each queue is FCFS, any task that joins q_{i-1} during such interval is placed behind the tagged task and therefore has no effect on the present transient analysis. The second term gives the probability that an arriving batch from the other external source which henceforth we shall refer to as simply "the external source," has size k. In this case, the maximum value that k can assume is the remaining available positions in switch–i at time t, $\mathcal{C}_i - n - \overline{n}$, since we make (in our approximation) no restrictions on the number of arrivals originated from the external source. The third and fourth terms account for the probability that there is a service completion at q_i and $\overline{q}_i$, respectively, and also, in each case, there is the subsequent decision of some of the remaining tasks located at each queue to leave the corresponding queue without being served (with probabilities $(1-\beta_i)$ and $(1-\overline{\beta}_i)$, respectively), resulting in bulk departures. Note that the third term is multiplied by $\beta_i^{I_{k<n}}$. This is to take into account two possible situations. The first is the case where k tasks ($k < n$) depart from q_i leaving behind the remaining $n-k$ tasks. This occurs with probability β_i since the first task to remain in q_i enters service with that probability. The second situation is the case where all tasks in q_i do leave in a bulk departure ($k = n$) which implies that there is no remaining task in q_i to either skip service (with probability $1 - \beta_i$) or start service (with probability β_i). Therefore, in this case, the third term is simply multiplied by 1 (since $I_{k<n} = 0$). For a similar reason, the fourth term is multiplied by $\overline{\beta}_i^{I_{k<\overline{n}}}$.

The second term of the right-hand side of Equation 13 gives the probability that the system is in state $(0, \overline{n}, r+n+d)$ at time t (for $d \leq \mathcal{C}_{i-1} - r - n$) and that during a time interval Δt it goes to state $(n, \overline{n}, r)$. This can occur if, during Δt a number d of tasks arrive from the finite source at an empty q_i and immediately skip service, followed by a task which enters service at q_i and $n-1$ further tasks which queue. The corresponding transition rate is

$$b_{i-1}\, u\, \beta_{i-1}^{I_{r>0}}\, \beta_i\, (1-\beta_{i-1})^{n-1}(1-\beta_{i-1})^d(1-\beta_i)^d\, I_{r+n+d\leq\mathcal{C}_{i-1}}.$$

The reason for having $\beta_{i-1}^{I_{r>0}}$ in the above transition is that, for $r > 0$, one task must enter the arrival phase in order to limit the size of the batch to $n + d$. Therefore, for this purpose, a task enters the arrival phase of the finite source (i.e., it starts being served in queue q_{i-1}) with probability β_{i-1}. For $r = 0$, all remaining tasks in the finite source at time t do join q_i which implies that there is no task left there to enter the arrival phase.

The third term is similar to the second with the difference that, in this case, the batch arrives from the external source (and not the finite one) which implies a transition from state $(0, \overline{n}, r)$ to state $(n, \overline{n}, r)$. Again, this occurs if, during Δt a number d of tasks arrive from the external source at an empty q_i and immediately skip service, followed by a task which enters service at q_i and $n-1$ further tasks which queue. In this case d can go up to infinity since, as mentioned earlier, we make (in our approximation) no restrictions on the number of arrivals originated from the external source. The corresponding transition rate is

$$\lambda_e u_e \sigma_e \beta_i \sum_{d=0}^{\infty}(1-\sigma_e)^{d+n-1}(1-\beta_i)^d = \frac{\lambda_e u_e \sigma_e \beta_i (1-\sigma_e)^{n-1}}{\sigma_e + \beta_i - \sigma_e\beta_i}.$$

The fourth term of the right-hand side of Equation 13 gives the probability that the system goes from state $(n, 0, r+\overline{n}+d)$ to state $(n, \overline{n}, r)$ in Δt units of time. The transition rate is obtained similarly to the second term explained above, with the difference that, in this case, the arriving tasks enter $\overline{q}_i$, as opposed to q_i. Note that here we have β_{i-1} instead of $\beta_{i-1}^{I_{r>0}}$. This is due to the fact that the last task in the finite source is the tagged task which obviously does not enter $\overline{q}_i$ (since $\overline{q}_i$ does not belong to the tagged task's path). Therefore the batch size must be smaller than or equal to $r-1$ (note the difference between the upper limits of the summations which are part of the transitions associated with the second and fourth terms of the right-hand side of Equation 13).

The fifth term gives the probability that the system goes from state $(n,0,r)$ to state $(n,\overline{n},r)$ in Δt units of time. The transition rate is obtained similarly to the third term explained above, with the difference that, as in the case detailed in the previous paragraph, the arriving tasks enter $\overline{q}_i$, as opposed to q_i.

The sixth term of the right-hand side of Equation 13 gives the probability that the system goes to state $(n,\overline{n},r)$ in Δt units of time, from a state where the number of tasks at q_i and $\overline{q}_i$, at time t, are $n-k$ and $\overline{n}$, respectively, for $1 \leq k \leq n-1$. This type of transition occurs if, during Δt, a batch of size k arrives from either the finite or the external source. If it arrives from the finite source then we have a transition from state $(n-k,\overline{n},r+k)$ to state $(n,\overline{n},r)$. If, on the other hand, the batch of size k arrives from the external source, then we have a transition from state $(n-k,\overline{n},r)$ to state $(n,\overline{n},r)$. The seventh term of Equation 13 is derived similarly since it gives the probability that the system goes to state $(n,\overline{n},r)$ in Δt units of time from a state where the number of tasks at q_i and $\overline{q}_i$, at time t, are n and $\overline{n}-k$, respectively, for $1 \leq k \leq n-1$.

The eight term of the right-hand side of Equation 13 gives the probabilities that the system goes from states $(n+k,\overline{n},r)$ and $(n,\overline{n}+k,r)$ to state $(n,\overline{n},r)$ in Δt units of time. In these cases, there must a service completion at respectively, q_i and $\overline{q}_i$, followed by $k-1$ further tasks which skip service with the next one entering service.

Now, rearranging the equations 13 and dividing by Δt, the resulting Kolmogorov forward equations for the subsystem considered are, in the limit $\Delta t \to 0$,

$$
\begin{aligned}
&\frac{\partial p_i(n,\overline{n},r,t)}{\partial t} = \\
&\quad -p_i(n,\overline{n},r,t)\,[I_{n+\overline{n}<\mathcal{C}_i}\,(b_{i-1}\sum_{k=1}^{\min(r,\mathcal{C}_i-n-\overline{n})}(1-\beta_{i-1})^{k-1}\beta_{i-1}^{I_{k<r}} + \lambda_e\sigma_e\sum_{k=1}^{\mathcal{C}_i-n-\overline{n}}(1-\sigma_e)^{k-1}) \\
&\qquad\qquad + I_{n>0}\, b_i\sum_{k=1}^{n}(1-\beta_i)^{k-1}\beta_i^{I_{k<n}} \\
&\qquad\qquad + I_{\overline{n}>0}\,\overline{b}_i\sum_{k=1}^{\overline{n}}(1-\overline{\beta}_i)^{k-1}\overline{\beta}_i^{I_{k<\overline{n}}}] \\
&\quad + b_{i-1}\,u\,\beta_{i-1}^{I_{r>0}}\beta_i\,(1-\beta_{i-1})^{n-1}\cdot \\
&\qquad\qquad \sum_{d=0}^{\mathcal{C}_{i-1}-r-n} p_i(0,\overline{n},r+n+d,t)(1-\beta_{i-1})^d(1-\beta_i)^d\, I_{r+n+d\leq\mathcal{C}_{i-1}} \\
&\quad + p_i(0,\overline{n},r,t)\cdot\frac{\lambda_e u_e\sigma_e\beta_i(1-\sigma_e)^{n-1}}{\sigma_e+\beta_i-\sigma_e\beta_i} \\
&\quad + b_{i-1}\,\overline{u}\,\beta_{i-1}\overline{\beta}_i\,(1-\beta_{i-1})^{\overline{n}-1}\cdot \\
&\qquad\qquad \sum_{d=0}^{\mathcal{C}_{i-1}-r-1-\overline{n}} p_i(n,0,r+\overline{n}+d,t)(1-\beta_{i-1})^d(1-\overline{\beta}_i)^d\, I_{r+\overline{n}+d\leq\mathcal{C}_{i-1}} \\
&\quad + p_i(n,0,r,t)\cdot\frac{\lambda_e\overline{u}_e\sigma_e\overline{\beta}_i(1-\sigma_e)^{\overline{n}-1}}{\sigma_e+\overline{\beta}_i-\sigma_e\overline{\beta}_i} \\
&\quad + I_{n>0}\sum_{k=1}^{n-1}[\,p_i(n-k,\overline{n},r+k,t)\,I_{r+k\leq\mathcal{C}_{i-1}}\,b_{i-1}\beta_{i-1}(1-\beta_{i-1})^{k-1}u \\
&\qquad\qquad + p_i(n-k,\overline{n},r,t)\,\lambda_e\sigma_e(1-\sigma_e)^{k-1}u_e]
\end{aligned}
$$

$$
\begin{aligned}
&+ I_{\overline{n}>0} \sum_{k=1}^{\overline{n}-1} [\, p_i(n, \overline{n}-k, r+k, t)\, I_{r+k \le \mathcal{C}_{i-1}}\, b_{i-1}\beta_{i-1}(1-\beta_{i-1})^{k-1}\overline{u} \\
&\qquad\qquad + p_i(n, \overline{n}-k, r, t)\, \lambda_e \sigma_e (1-\sigma_e)^{k-1}\overline{u}_e] \\
&+ I_{n+\overline{n}<\mathcal{C}_i} \sum_{k=1}^{\mathcal{C}_i-n-\overline{n}} [\, p_i(n+k, \overline{n}, r, t) b_i\, \beta_i\, (1-\beta_i)^{k-1} \\
&\qquad\qquad\quad + p_i(n, \overline{n}+k, r, t)\overline{b}_i\, \overline{\beta}_i\, (1-\overline{\beta}_i)^{k-1}], \\
&\qquad\qquad\qquad 0 \le r \le r_{i-1},\ \ 0 \le n \le \mathcal{C}_i,\ \ 0 \le \overline{n} \le \mathcal{C}_i - n. \qquad (14)
\end{aligned}
$$

The boundary conditions for the above linear system of differential equations are:

$$
p_i(x, \overline{x}, y, 0) = \begin{cases} 1, & x = r_i \ \ \& \ \ \overline{x} = \overline{r}_i \ \ \& \ \ y = r_{i-1} \\ 0, & \text{otherwise} \end{cases}
$$

for $0 \le x + \overline{x} \le \mathcal{C}_i$, and $0 \le y \le r_{i-1}$. The normalised form of the solution of Equations 14 is obtained by dividing by $\sum_{x=0}^{\mathcal{C}_i} \sum_{\overline{x}=0}^{\mathcal{C}_i-x} \sum_{y=0}^{r_{i-1}} p_i(x, \overline{x}, y, t)$. This solution, together with the boundary conditions given above, produces the time-dependent result for the probability $\mathbf{P}\{Q_i(T_i) = n, \overline{Q}_i(T_i) = \overline{n} \,|\, \mathcal{S}_{i-1} = v, Q_{i-1}(T_{i-1}) = r_{i-1}, Q_i(T_{i-1}) = r_i, \overline{Q}_i(T_{i-1}) = \overline{r}_i\}$ as indicated below:

$$
\begin{aligned}
\mathbf{P}\{Q_i(T_i) = n, \overline{Q}_i(T_i) = \overline{n} \,|\, \mathcal{S}_{i-1} = v,\ Q_{i-1}(T_{i-1}) = r_{i-1},& \\
Q_i(T_{i-1}) = r_i,\ \overline{Q}_i(T_{i-1}) = \overline{r}_i\} \ =\ & p_i(n, \overline{n}, 0, v), \\
& 0 \le n \le \mathcal{C}_i,\ \ 0 \le \overline{n} \le \mathcal{C}_i - n.
\end{aligned}
$$

Note that the conditional events $(Q_i(T_{i-1}) = r_i)$ and $(\overline{Q}_i(T_{i-1}) = \overline{r}_i)$ are considered as boundary conditions whereas the event $(Q_{i-1}(T_{i-1}) = r_{i-1})$ is taken into account as the capacity of the finite source shown in Figure 4.

We can finally obtain the marginal probability distribution $\mathbf{P}\{Q_i(T_i) = n \,|\, \mathcal{S}_{i-1} = v,\ Q_{i-1}(T_{i-1}) = r_{i-1},\ Q_i(T_{i-1}) = r_i,\ \overline{Q}_i(T_{i-1}) = \overline{r}_i\}$ as follows:

$$
\begin{aligned}
\mathbf{P}\{Q_i(T_i) = n \,|\, \mathcal{S}_{i-1} = v,\ Q_{i-1}(T_{i-1}) = r_{i-1},& \\
Q_i(T_{i-1}) = r_i,\ \overline{Q}_i(T_{i-1}) = \overline{r}_i\} \ =\ & \sum_{\overline{n}=0}^{\mathcal{C}_i-n} p_i(n, \overline{n}, 0, v), \qquad 0 < n < \mathcal{C}_i.
\end{aligned}
$$

The numerical solution of Equations 14 is obtained from built-in routines in *Mathematica v2.1*. Although these routines are very efficient and extremely fast (even for quite large state spaces) there is still a need for a faster algorithm since in order to obtain the end-to-end delay of a tagged task, it is necessary to solve a similar system of equations to those given above for each node in the tagged task's path.

In view of this fact, we propose an alternative approximate algorithm which, although being cruder than the method presented here in terms of assumptions made, can be more easily implemented by computer.

3.2 Transient Analysis: An Alternative Approach

In this section we present an alternative (approximate) approach for obtaining the probability $\mathbf{P}\{Q_i(T_i) = n, \,|\, \mathcal{S}_{i-1} = v,\ Q_{i-1}(T_{i-1}) = r_{i-1},\ Q_i(T_{i-1}) = r_i,\ \overline{Q}_i(T_{i-1}) = \overline{r}_i\}$. For notational convenience and clarity, let $p_i(n|v, r_{i-1}, r_i, \overline{r}_i)$ denote such probability.

We initially define the following random variables:

Z_v : number of "external" arrivals at the ith node of the tagged task's path during the time interval $[T_{i-1}, T_i)$ of duration v.

X_v : number of departures from the queue in the tagged task's path located at the ith stage of the network, q_i, during the time interval $[T_{i-1}, T_i)$ of duration v.

$\overline{X}_v$: number of departures from $\overline{q}_i$, the conjugate queue of q_i, during the time interval $[T_{i-1}, T_i)$ of duration v.

We argue that the probability $p_i(n|v, r_{i-1}, r_i, \overline{r}_i)$ can be evaluated by the following approximation:

$$p_i(n|v, r_{i-1}, r_i, \overline{r}_i) = \sum_{\overline{n}=0}^{\mathcal{C}_i - n} \mathbf{P}\{Z_v - (X_v + \overline{X}_v) = n + \overline{n} - r_{i-1} - r_i - \overline{r}_i\}. \quad (15)$$

Equation 15 states that $p_i(n|v, r_{i-1}, r_i, \overline{r}_i)$ is the probability that, during the time interval $[T_{i-1}, T_i)$, the difference between the number of external arrivals to the switch located at the ith node of the tagged task's path (switch–i) and the total number of departures from that switch (which includes departures from both outputs) is equal to the (given) difference between the number of tasks at queues q_i and $\overline{q}_i$ at time T_i, and the number of tasks at queues q_{i-1}, q_i, and $\overline{q}_i$ at time T_{i-1}. Since the service discipline is assumed to be FCFS, by time T_i (the tagged task's arrival instant at queue q_i) all r_{i-1} tasks that were at queue q_{i-1} at time T_{i-1} (including the tagged task) must have entered switch–i. Note that the right-hand side of Equation 15 is summed over all possible values of $\overline{n}$, the queue length at $\overline{q}_i$ at time T_i. This is due to the fact that $p_i(n|v, r_{i-1}, r_i, \overline{r}_i)$ does not depend on the number of tasks found in $\overline{q}_i$ at time T_i.

Except for the tagged task whose next destination is determined by its path, all other $r_{i-1}-1$ tasks can go to either q_i or $\overline{q}_i$ with probabilities u and $\overline{u}$ (as defined in **Section 3.1**), respectively. Supposing that k tasks, $0 \leq k \leq r_{i-1} - 1$, do go to queue $\overline{q}_i$, we define the function $z^{\min}(\overline{n}, k)$ as follows:

$$z^{\min}(\overline{n}, k) = \max(0, n - r_{i-1} + k - r_i) + \max(0, \overline{n} - \overline{r}_i - k), \quad (16)$$

which gives the minimum value the random variable Z_v can assume. The above expression simply states that, given that at time T_i, there are n tasks in q_i, and at time T_{i-1}, there were r_{i-1} tasks at queue q_{i-1} and r_i tasks at queue q_i, and during the interval $[T_{i-1}, T_i)$ a total of $r_{i-1} - k$ tasks arrive at q_i from q_{i-1}, then there must be a minimum number of external arrivals given by the difference between n and $r_{i-1} - k + r_i$ (if positive, otherwise the minimum number of external arrivals is zero). The same reasoning applies to queue $\overline{q}_i$ resulting in the second term of the right-hand side of Equation 16.

Still supposing that k tasks go from queue q_{i-1} to queue $\overline{q}_i$ during the time interval $[T_{i-1}, T_i)$, we define the function $x^{\min}(k)$ as follows:

$$x^{\min}(k) = \max(0, r_{i-1} - k + r_i - n), \quad (17)$$

which gives the minimum value the random variable X_v can assume. That is, given that at time T_i, there are n tasks in q_i, and at time T_{i-1}, there were r_{i-1} tasks at queue q_{i-1} and r_i tasks at queue q_i, and during the interval $[T_{i-1}, T_i)$ a total of $r_{i-1} - k$ tasks do arrive at q_i from q_{i-1}, then there must be a minimum number of departures from q_i given by the difference between $r_{i-1} - k + r_i$ and n (if positive, otherwise the minimum number of departures from q_i is zero). Similarly, we define $\overline{x}^{\min}(\overline{n}, k)$ below, as the minimum number of departures from $\overline{q}_i$ during the time interval $[T_{i-1}, T_i)$ (i.e., the minimum value the random variable $\overline{X}_v$ can assume):

$$\overline{x}^{\min}(\overline{n}, k) = \max(0, k + \overline{r}_i - \overline{n}). \quad (18)$$

Now, supposing that the number of external arrivals during the transient period $[T_{i-1}, T_i)$ is y, $z^{\min}(\overline{n}, k) \leq y < \infty$, the total number of departures during this period from both q_i and q_{i-1} is given by:

$$l(y, \overline{n}) = y + r_{i-1} + r_i + \overline{r}_i - n - \overline{n}, \quad (19)$$

which is simply the difference between the sum of the number of tasks in switch–i at time T_{i-1} (i.e., $r_i + \overline{r}_i$) and the number of tasks that arrive at that switch during the interval $[T_{i-1}, T_i]$ (i.e., $y + r_{i-1}$), and the the number of tasks found at time T_i in switch–i (i.e., $n + \overline{n}$). Note that the interval is closed on the right since it includes the arrival of the tagged task. Also, note that the random variables Z_v, X_v, $\overline{X}_v$, and the functions $x^{\min}(k)$, $\overline{x}^{\min}(\overline{n}, k)$, $z^{\min}(\overline{n}, k)$, and $l(y, \overline{n})$ are all dependent on the values of n, r_{i-1}, r_i, and $\overline{r}_i$. We just do not make such dependency explicit for notational simplicity.

For any two discrete-valued random variables, A and B, the probability that the difference between them is equal to a constant c, $\mathbf{P}\{A - B = c\}$, can be obtained as follows:

$$\begin{aligned} \mathbf{P}\{A - B = c\} &= \sum_i \mathbf{P}\{A = a_i, B = a_i - c\} \\ &= \sum_i \mathbf{P}\{B = a_i - c \,|\, A = a_i\}\, \mathbf{P}\{A = a_i\}. \end{aligned} \tag{20}$$

Applying this result to Equation 15 we have:

$$\begin{aligned} & p_i(n | v, r_{i-1}, r_i, \overline{r}_i) = \\ & \quad \sum_{\overline{n}=0}^{\mathcal{C}_i - n} \sum_{k=0}^{r_{i-1}-1} \binom{r_{i-1} - 1}{k} \overline{u}^k\, u^{r_{i-1}-1-k} \sum_{y=z^{\min}(\overline{n},k)}^{\infty} \mathbf{P}\{Z_v = y\} \cdot \\ & \quad \mathbf{P}\{X_v + \overline{X}_v = l(y, \overline{n}) \,|\, Z_v = y\}. \end{aligned} \tag{21}$$

As mentioned earlier, the minimum values that the (independent) random variables X_v and $\overline{X}_v$ can assume are, for a given k, $x^{\min}(k)$ and $\overline{x}^{\min}(\overline{n}, k)$, respectively. Therefore the convolution $\mathbf{P}\{X_v + \overline{X}_v = l(y, \overline{n}) \mid Z_v = y\}$ is in fact a "truncated" convolution, given by:

$$\mathbf{P}\{X_v + \overline{X}_v = l(y, \overline{n}) \mid Z_v = y\} = \sum_{w=\overline{x}^{\min}(\overline{n},k)}^{l(y,\overline{n}) - x^{\min}(k)} \mathbf{P}\{X_v = l(y, \overline{n}) - w\} \mathbf{P}\{\overline{X}_v = w\}.$$

Note that the condition $(Z_v = y)$ appears implicitly in the function $l(y, \overline{n})$ on the right-hand side of the above equation. Applying this result to Equation 21 we have:

$$\begin{aligned} & p_i(n | v, r_{i-1}, r_i, \overline{r}_i) = \\ & \quad \sum_{\overline{n}=0}^{\mathcal{C}_i - n} \sum_{k=0}^{r_{i-1}-1} \binom{r_{i-1} - 1}{k} \overline{u}^k\, u^{r_{i-1}-1-k} \sum_{y=z^{\min}(\overline{n},k)}^{\infty} \mathbf{P}\{Z_v = y\} \cdot \\ & \quad \sum_{w=\overline{x}^{\min}(\overline{n},k)}^{l(y,\overline{n}) - x^{\min}(k)} \mathbf{P}\{X_v = l(y, \overline{n}) - w\} \mathbf{P}\{\overline{X}_v = w\} \end{aligned} \tag{22}$$

The random variable Z_v, for a given v, is one in a family of random variables given by the counting process $\{Z_t, t \geq 0\}$. This counting process is associated with a generalised exponential distribution with parameters (λ_e, σ_e) which is the distribution of the inter-arrival times of the arrival process originated from the external source shown in Figure 4. The counting processes $\{X_t, t \geq 0\}$ and $\{\overline{X}_t, t \geq 0\}$ are also associated with GE distributions. $\{X_t, t \geq 0\}$ is related to the distribution of the effective service time at q_i whose parameters are (b_i, β_i), whereas $\{\overline{X}_t, t \geq 0\}$ is related to the distribution of the effective service time at $\overline{q}_i$ with parameters $(\overline{b}_i, \overline{\beta}_i)$. In order to evaluate the right-hand side of Equation 22 we must obtain the distributions of Z_t, X_t, and $\overline{X}_t$, $t \geq 0$.

Let $\{N_t, t \geq 0\}$ be a Poisson process with rate λ. This process can also be viewed as a sequence of non-negative random variables with identical exponential distribution of rate λ.

N_t is said to be the counting process associated with the point process determined by this sequence of random variables. Now, let $\{\mathcal{N}_t, t \geq 0\}$ be the counting process associated with a GE distribution of parameters (λ, θ). This is simply a generalisation of the Poisson process which allows simultaneous occurrences. The probability that, at a given point, the number of simultaneous occurrences (or batch size) is k is given by $\theta(1-\theta)^{k-1}$. It then follows that the probability that the total number of occurrences in a counting process associated with a GE distributions of parameters (λ, θ) is k up to time t (i.e., $\mathbf{P}\{\mathcal{N}_t = k\}$), is given by:

$$\begin{aligned} \mathbf{P}\{\mathcal{N}_t = k\} &= \sum_{i=1}^{k} \mathbf{P}\{N_t = i\} \binom{k-1}{i-1} \theta^i (1-\theta)^{k-i} \\ &= e^{-\lambda t} \sum_{i=1}^{k} \frac{(\lambda t)^i}{i!} \binom{k-1}{i-1} \theta^i (1-\theta)^{k-i}, \qquad k \geq 1. \end{aligned}$$

The above expression simply states that the probability that $\mathcal{N}_t$ equals k is given by all possible combinations ranging from 1 point with k simultaneous occurrences to k points with only one occurrence each. The probability that $\mathcal{N}_t$ equals 0 can be obtained as follows as a check:

$$\begin{aligned} \mathbf{P}\{\mathcal{N}_t = 0\} &= 1 - e^{-\lambda t} \sum_{k=1}^{\infty} \sum_{i=1}^{k} \frac{(\lambda t)^i}{i!} \binom{k-1}{i-1} \theta^i (1-\theta)^{k-i} \\ &= 1 - e^{-\lambda t} (e^{\lambda t} - 1) \\ &= e^{-\lambda t} = \mathbf{P}\{N_t = 0\}. \end{aligned}$$

For notational conciseness we define the function $F_t(x_1, x_2, x_3)$ as:

$$F_t(x_1, x_2, x_3) = e^{-x_2 t} \, [I_{x_1=0} + I_{x_1>0} \sum_{i=1}^{x_1} \frac{(x_2 t)^i}{i!} \binom{x_1 - 1}{i-1} x_3^i (1-x_3)^{x_1 - i}], \tag{23}$$

which gives the probability that the number of events in a counting process associated with a GE distribution of parameters (x_2, x_3) is x_1. We can finally obtain an expression for $p_i(n|v, r_{i-1}, r_i, \overline{r}_i)$ by substituting Definition 23 into Equation 22, which results in:

$$\begin{aligned} & p_i(n|v, r_{i-1}, r_i, \overline{r}_i) = \\ & \sum_{\overline{n}=0}^{\mathcal{C}_i - n} \sum_{k=0}^{r_{i-1}-1} \binom{r_{i-1} - 1}{k} \overline{u}^k \, u^{r_{i-1}-1-k} \sum_{y=z^{\min}(\overline{n},k)}^{\infty} F_v(y, \lambda_e, \sigma_e) \cdot \\ & \sum_{w=\overline{x}^{\min}(\overline{n},k)}^{l(y,\overline{n})-x^{\min}(k)} F_v(l(y,\overline{n}) - w, b_i, \beta_i) \, F_v(w, \overline{b}_i, \overline{\beta}_i) \end{aligned} \tag{24}$$

As remarked earlier in this section, the above expression is quite simple and can be easily implemented by computer, thus avoiding the use of heavy symbolic manipulation which is required to solve Equations 14.

3.3 The Approximation Algorithm

We can now present the complete approximation method for obtaining the end-to-end delay of a tagged task traversing a buffered multistage interconnection network.

As mentioned in **Section 2**, the queues in the path taken by the tagged task are numbered $0, 1, \ldots, M-1$. Let $D_M(t)$ be defined as $D_M(t) = \mathbf{P}\{T_M \leq t\}$, where T_M is the instant the tagged task leaves the last queue in its path. As the tagged task enters the network at time $T_0 = 0$, the random variable T_M also gives the tagged task's end-to-end delay. Therefore, $D_M(t)$

is the probability distribution of the time elapsed between the tagged task's arrival instant at one of the external inputs of the network and its departure from one of the network's external outputs. This distribution can be obtained as follows:

$$D_M(t) = \sum_{n=1}^{\mathcal{C}_{M-1}} D_{M-1}(\cdot, n) * \mathcal{F}_{\mu_{M-1},\theta_{M-1},n}, \tag{25}$$

where $*$ indicates the convolution operation on distributions.

The term $D_{M-1}(t,n)$, for $t \geq 0$, gives the joint distribution of: (i) the time spent in the network by the tagged task up to its arrival at queue q_{M-1}, and (ii) the queue length found at q_{M-1} by the tagged task at that instant. $\mathcal{F}_{\mu_{M-1},\theta_{M-1},n}(t)$ gives the distribution of the tagged task's sojourn time at q_{M-1}, given that the tagged task finds n tasks already in q_{M-1} upon arrival (itself included) and the parameters of the GE distribution of the actual service time are μ_{M-1} and θ_{M-1}. The actual service time is used in the last queue of the tagged task's path as opposed to the effective service time (used in all previous queues) due to the fact that there is no blocking in the last stage of network and, therefore, there is no need to take its effect into account.

Considering initially the case where the transient analysis is approached via the solution of the Kolmogorov forward differential-difference equations (see **Section 3.1**), the distributions $D_i(t,n)$, $1 \leq i \leq M-1$, $t \geq 0$, $1 \leq n \leq \mathcal{C}_i$, are given by:

$$\begin{aligned} D_i(t,n) &= \sum_{\mathbf{r}_i \in \mathcal{D}_i} \mathbf{P}\{Q_i = r_i, \overline{Q}_i = \overline{r}_i\} \cdot \\ & \sum_{\overline{n}=0}^{\mathcal{C}_i - n} \int_0^t p_i(n, \overline{n}, 0, v) D_{i-1}(t-v, r_{i-1}) f_{b_{i-1},\beta_{i-1},r_{i-1}}(v) dv, \end{aligned} \tag{26}$$

which is simply Equation 8 rewritten after the appropriate substitutions for the individual probabilities are made.

For the case where the transient analysis is approached via the alternative method proposed in **Section 3.2**, the distributions $D_i(t,n)$, $1 \leq i \leq M-1$, $t \geq 0$, $1 \leq n \leq \mathcal{C}_i$, are obtained as follows:

$$\begin{aligned} D_i(t,n) &= \sum_{\mathbf{r}_i \in \mathcal{D}_i} \mathbf{P}\{Q_i = r_i, \overline{Q}_i = \overline{r}_i\} \cdot \\ & \int_0^t p_i(n|v, r_{i-1}, r_i, \overline{r}_i) D_{i-1}(t-v, r_{i-1}) f_{b_{i-1},\beta_{i-1},r_{i-1}}(v) dv. \end{aligned} \tag{27}$$

Note that the difference between Equations 26 and 27 obviously just lies on the terms referring to the transient sojourn times.

Finally, for $i = 0$, the distribution $D_0(t,n)$, $t \geq 0$, $1 \leq n \leq \mathcal{C}_i$, is given by Equation 3 which we reproduce here:

$$D_0(t,n) = \mathbf{P}\{T_0 \leq t,\ Q_0(T_0) = n\} = \mathbf{P}\{\tilde{Q}_0 = n\}.$$

4 Numerical Results

In this section we give some numerical results of the analytical models proposed in this paper and compare these with simulation. Results are presented for 4 different sets of network parameters (see Figures 5 – 8). For each one of these sets the end-to-end delay probability distribution is given for two different paths. We consider an 8×8 Banyan network with the same topology as that shown in Figure 1. For Figures 5 – 6, each output pin has a GE-distributed service time with parameters $(0.5, 0.5)$. These values were chosen (partly) so that the average service time at each output equals 1. The external arrival rate at each input pin is $\lambda_i = 0.1, \ \ i = 0, \ldots, 7$, so that the network total arrival rate equals 0.8. All tasks are assumed to be stochastically identical. The top output (hot-spot) is selected with the same probability, r_{i0}, by all tasks and all other outputs are selected with equal probability. For Figure 5 the routing probabilities r_{i0}, $0 \le i \le 7$, are equal to 0.23, while the remaining probabilities, r_{ij}, for $1 \le i \le 7$ and $0 \le j \le 7$, are all equal to 0.11. For Figure 6, the routing probabilities r_{i0}, $0 \le i \le 7$, are equal to 0.86, while the remaining probabilities, r_{ij}, for $1 \le i \le 7$ and $0 \le j \le 7$, are all equal to 0.02. This is therefore an extreme hot-spot but the predictions retain reasonable accuracy. The effective arrival rate at each input pin of any stage is obtained from an iterative procedure based on the model presented in [15] whereby the true value of the effective arrival rates at stage–0 inputs $\bar{\boldsymbol{\lambda}}$ are approximated by successive iterations (the rates used as initial conditions of the said procedure are the actual external arrival rates $\boldsymbol{\lambda}$).

Figure 7 gives results for a non-uniform traffic pattern characterized by the routing probability matrix **R'**, given below:

$$\mathbf{R'} = \begin{bmatrix} 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.2 & 0.2 \\ 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.2 & 0.2 \\ 0.1 & 0.1 & 0.1 & 0.1 & 0.26 & 0.26 & 0.04 & 0.04 \\ 0.1 & 0.1 & 0.1 & 0.1 & 0.26 & 0.26 & 0.04 & 0.04 \\ 0.125 & 0.125 & 0.125 & 0.125 & 0.125 & 0.125 & 0.125 & 0.125 \\ 0.125 & 0.125 & 0.125 & 0.125 & 0.125 & 0.125 & 0.125 & 0.125 \\ 0.125 & 0.125 & 0.125 & 0.125 & 0.125 & 0.125 & 0.125 & 0.125 \\ 0.125 & 0.125 & 0.125 & 0.125 & 0.125 & 0.125 & 0.125 & 0.125 \end{bmatrix}$$

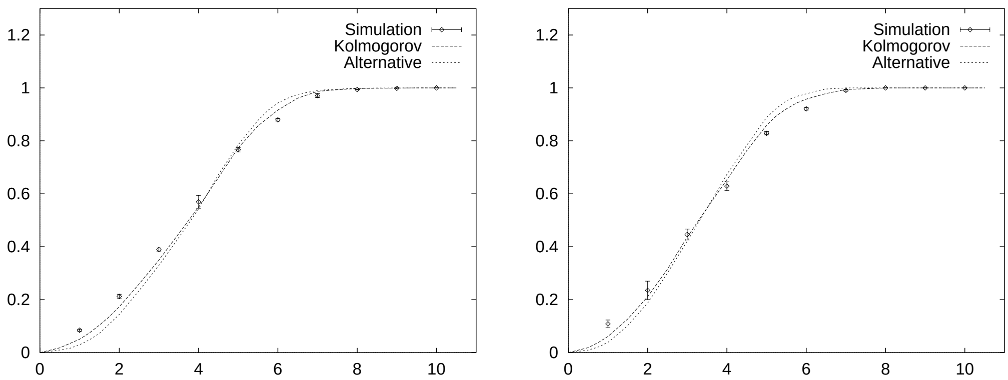
These values where chosen so that the traffic flows arriving on distinct inputs at a particular internal switch–$(1, 1)$ have routing bias towards distinct switch outputs. Considering the topology of the network shown in Figure 1, the traffic flow arriving at input–0 of switch–$(1, 1)$ is given by $\bar{\lambda}_0(r_{04} + r_{05} + r_{06} + r_{07}) + \bar{\lambda}_1(r_{14} + r_{15} + r_{16} + r_{17})$, where $\bar{\lambda}_0(r_{04} + r_{05}) + \bar{\lambda}_1(r_{14} + r_{15})$ is the traffic flow from input–0 to output–0, and $\bar{\lambda}_0(r_{06} + r_{07}) + \bar{\lambda}_1(r_{16} + r_{17})$ is the traffic flow from input–0 to output–1. According to the routing matrix **R'**, given above, these traffic flows are then respectively $0.2(\bar{\lambda}_0 + \bar{\lambda}_1)$ and $0.4(\bar{\lambda}_0 + \bar{\lambda}_1)$. Analogously, the traffic flow arriving at input–1 of switch–$(1, 1)$ is given by $\bar{\lambda}_2(r_{24} + r_{25} + r_{26} + r_{27}) + \bar{\lambda}_3(r_{34} + r_{35} + r_{36} + r_{37})$, where $\bar{\lambda}_2(r_{24} + r_{25}) + \bar{\lambda}_3(r_{34} + r_{35})$ is the traffic flow from input–1 to output–0, and $\bar{\lambda}_2(r_{26} + r_{27}) + \bar{\lambda}_3(r_{36} + r_{37})$ is the traffic flow from input–1 to output–1. Again, according to **R'**, we have a rate $0.52(\bar{\lambda}_2 + \bar{\lambda}_3)$ from 1 to 0 and $0.08(\bar{\lambda}_2 + \bar{\lambda}_3)$, from 1 to 1. For the case of Figure 7 the external arrival rates are given by $\boldsymbol{\lambda} = [0.02,\ 0.02,\ 0.2,\ 0.2,\ 0.02,\ 0.02,\ 0.2,\ 0.2]$. Again, looking at switch–$(1, 1)$, it can be noticed that there is much more traffic arriving at input–0, 0.24 in this case, than at input–1, which is 0.024. The simulation indicates that the analytical model performs reasonably well even for these cases where there is the combined effect of unbalanced traffic and routing bias.

In Figures 7 – 8, the network outputs are slower than all internal switches, which in practice is often the case. To produce such effect, each output pin has a GE-distributed service time with parameters $(0.5, 0.7)$ resulting in an average service time of 1.4 as opposed to 1 for all the internal switches. Figure 8 gives results for heavy traffic. The external arrival rate at each input is 0.5 resulting in a total arrival rate of 4.

In general there are three levels of approximation:

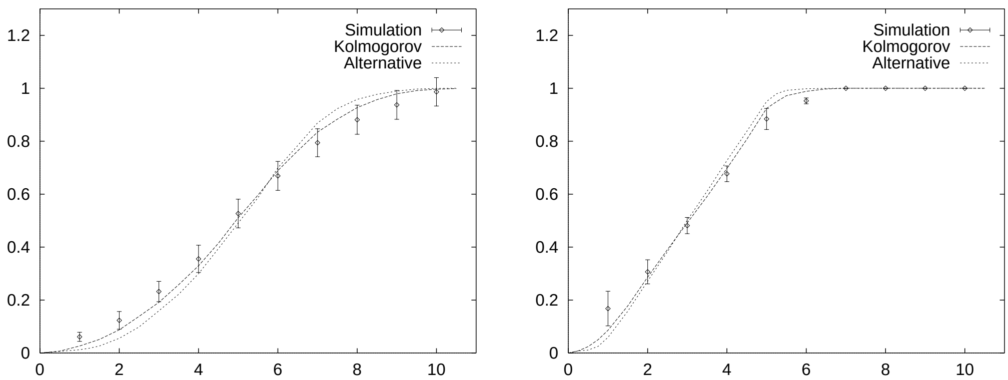
(i) In the PCA method's indepedence assumption there is only dependence between adjacent switches in a path.

(ii) The equilibrium probability distribution of queue *lengths* (see [15]) which is used in the PCA method.



(i) (ii)

Figure 5: (i) Path–(0,0) Time Delay Distributionr; (ii) Path–(4,7) Time Delay Distribution.



(i) (ii)

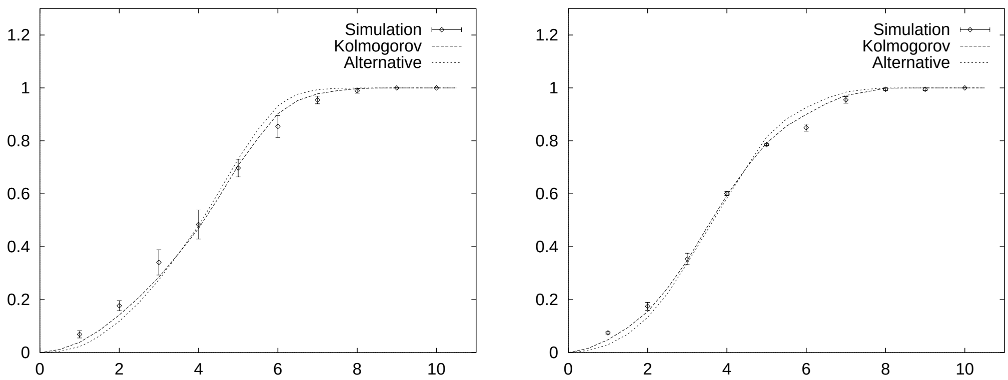
Figure 6: (i) Path–(0,0) Time Delay Distribution; (ii) Path–(4,7) Time Delay Distribution.

(iii) The approximations involved in the transient analysis.

The effect of (iii) can be assessed by comparision with the first model (**Section 3.1**). The effect of (ii) could be checked by using more accurate estimates of the equilibrium queue length probability distributions obtained by simulation.

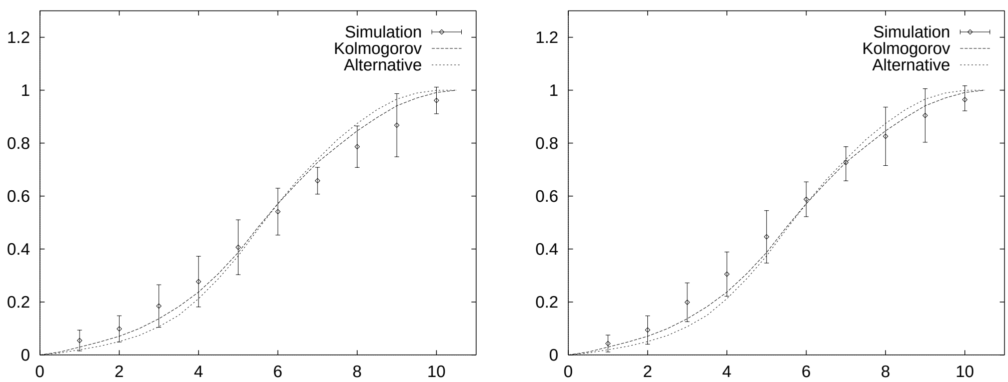
5 Conclusion

An approximation method for obtaining the response time of a tagged task in an aynchronous, finite-buffered, packet-switched banyan network has been presented. The approach taken considers successive pairs of servers in a task's path; the queue joined on arrival at the second server having length dependent on that which existed previously on arrival at the first. Results are obtained using two methods for analysing the transient period between arrivals at successive nodes. The first solves the corresponding Kolmogorov equations. The second is a more efficient approximation which assumes that the probability distribution, during the transient time, of the difference between the number of external arrivals to the switch concerned and the total number of departures from it (which includes departures from both outputs) is given in terms



(i) (ii)

Figure 7: (i) Path–(0,5) Time Delay Distribution; (ii) Path–(4,0) Time Delay Distribution.



(i) (ii)

Figure 8: (i) Path–(0,0) Time Delay Distribution; (ii) Path–(4,7) Time Delay Distribution.

of the number of tasks in the sub-system being analysed at both ends of the considered transient interval. The main advantages of this method over the direct solution of the Komolgorov equations are basically its conciseness and ease of implementation, as well as the fact that it requires much less processing time, especially for systems with larger state space (i.e., switches with bigger buffer capacities). On the other hand, the lesser cause and effect basis of the method may be regarded as the major factor behind its poorer performance in terms of accuracy.

References

[1] A. E. Conway and D. E. O'Brien, "Estimating Response Time Distributions in Queueing Networks," In *Proc. of the IFIP WG 7.3 International Conference on the Performance of Distributed Systems and Integrated Communication Networks*, Kyoto, Japan, pp. 225–244, 1992.

[2] H. Daduna, "Passage Times for Overtake-free Paths in Gordon-Newell Networks," *Adv. Appl. Prob.*, **14**, 1982, pp. 672–686.

[3] A. J. Field and P. G. Harrison, "Transmission Times in Unbuffered Full Crossbars with Cyclic Arbitration," In *Proc. ICPP' 92*.

[4] P. G. Harrison, "An Enhanced Approximation by Pair-Wise Analysis of Servers for Time Delay Distributions in Queueing Networks," *IEEE Trans. Comput.*, **C–35**, no 1, Jan. 1986, pp. 54–61.

[5] P. G. Harrison and A. de C. Pinto, "Blocking in Asynchronous, Buffered Banyan Networks," In *Proc. of the IFIP WG 7.3 International Conference on the Performance of Distributed Systems and Integrated Communication Networks*, Kyoto, Japan, ed. T. Hasegawa, H. Takagi and Y. Takahashi, North-Holland, pp. 169–188, 1992.

[6] P. G. Harrison and A. de C. Pinto, "An Approximate Analysis of Asynchronous, Packet-switched Buffered Banyan Networks with Blocking," *Performance Evaluation*, **19**, 1994, pp. 223–258.

[7] P. G. Harrison and N. M. Patel, *Performance Modelling of Communication Networks and Computer Architectures*, Addison-Wesley, 1992.

[8] D. Kouvatsos and N. Xenios, "MEM for Arbitrary Queueing Networks with Multiple General Servers and Repetitive-service Blocking," *Performance Evaluation*, **10**, pp. 169–195, Sep. 1989.

[9] C. P. Kruskal, M. Snir and A. Weiss, "The Distribution of Waiting Times in Clocked Multisage Interconnection Network," *IEEE Trans. Comput.*, **37**, no 11, pp. 1337–1352, Nov. 1988.

[10] C. Langaris and B. Conolly, "On the Waiting Time of a Two-stage Queueing System with Blocking," *European Journal of Operational Research*, **19**, 1985, pp. 222–232.

[11] J. McKenna, "Asymptotic Expansions of the Sojourn Time Distribution Functions of Jobs in Closed, Product-form Queuing Networks," *Journal of the ACM*, **34**, no 4, 1987, pp. 985–1003.

[12] B. Melamed and M. Yadin, "Numerical Computation of Sojourn-time Distributions in Queueing Networks," *Journal of the ACM*, **31**, no 4, 1984, pp. 839–854.

[13] H. Perros, "A Bibliography of Papers on Queueing Networks with Finite Capacity Queues," *Performance Evaluation*, **10**, 1989, pp. 255–260.

[14] A. de C. Pinto, "Time Delay Distribution in Buffered Multistage Interconnection Networks," In *Proc. of the 1st UK Workshop on Performance Modelling and Evaluation of ATM Networks*, ed. D. D. Kouvatsos, 1993.

[15] A. de C. Pinto, "Models of Finite-buffered Packet-switched Multistage Interconnection Networks," PhD thesis, Dept. of Computing, Imperial College, University of London, in preparation.

[16] K. E. E. Raatikainen, "Approximating Response Time Distributions," *Performance Evaluation Review*, **17**, no 1, 1989, pp. 190–199.

[17] H. J. Siegel, *Interconnection Networks for Large-Scale Parallel Processing: Theory and Case Studies*, 2nd ed., McGraw-Hill Publishing Co., 1990.

[18] J. Walrand, *An Introduction to Queueing Networks*, Prentice-Hall, 1988.

[19] H. Yoon, K. Y. Lee and M. T. Liu, “Performance Analysis of Multibuffered Packet-Switching Networks in Multiprocessor Systems,” *IEEE Trans. Comput.*, **39**, no 3, pp. 319–327, Mar. 1990.