S. S. Lam

Queuing Networks with Population Size Constraints

Abstract: The class of queuing networks with multiple routing subchains is extended to include mechanisms of state-dependent lost arrivals and triggered arrivals. A sufficient condition is found, involving the loss and trigger functions, for the equilibrium network state probability distribution to have the product form; the known class of queuing networks with a product form solution is thus enlarged. Such queuing networks are useful models for systems with various population size constraints. Potential applications to modeling computer communication systems with storage and flow control constraints are indicated.

Introduction

Networks of queues are important models of multiprogramming computer systems and computer-communication networks. Jackson [1] studied queuing networks with exponential servers and showed that the equilibrium network state probability distribution has a *product form.* The class of queuing networks with a product form solution has recently been extended by Baskett, Chandy, Muntz and Palacios [2] to include different classes of customers and general service time distributions for certain service disciplines. In their model, customers travel through a network of four types of service stations and may change class membership while making a transition from one service station to another. The underlying Markov chain of station and class transitions is assumed to be decomposable into K ergodic subchains (to be referred to as *routing subchains*). They considered networks with open, closed and mixed routing subchains. Either one of two kinds of population size constraints can be defined: a) the number of customers in each routing subchain is constrained (unconstrained) independently of other routing subchain population sizes, or b) the total number of customers in the network is constrained (unconstrained).

Let S denote the network state. In state S, we define $n_k(S)$ to be the customer population size in the kth routing subchain and $N(S)$ to be the total number of customers in the network. Jackson [1] first introduced mechanisms of lost and triggered arrivals as a function of $N(S)$. In this paper [3, 4], we consider mechanisms of lost and triggered arrivals as a function of the vector $\mathbf{n}(S) = (n_1(S), n_2(S), \cdots, n_K(S))$, which will be referred to as the *population vector* in state S. Depending upon $\mathbf{n}(S)$, arrivals to the network from external Poisson streams may be lost (lost arrivals); also, upon the departure of a customer from the network, a new customer may be injected instantaneously into the network (triggered arrivals). Such state-dependent lost and triggered arrivals give rise to much more general population size constraints than a) and b) above. To motivate consideration of such more general constraints, we give the following examples of models of computer communication systems with population size constraints which are beyond the scope of the class of queuing networks studied by Baskett et al. [2]. We shall later come back to these examples and show that they can be represented as special cases of queuing networks with state-dependent lost and triggered arrivals.

Example 1

Consider a queuing network model of a single store-and-forward node within a packet switching network [5, 6]. Packets received by the node are routed to one of K output channels. Let $n_k(S)$ denote the number of packets at channel k in state S of the node model. Suppose the node has a total of B buffers for the storage of packets being forwarded and the following buffer management scheme is adopted [7, 8]. The pool of B buffers is shared by all output channels. However, the number of buffers that can be allocated to channel k cannot exceed B_k where $B_k < B$. Under this scheme, feasible states of the node model must satisfy the constraints

$$\sum_{k=1}^{K} n_k(S) \le B$$

and

$$n_k(S) \le B_k, \qquad k = 1, 2, \cdots, K.$$

Example 2

A closed queuing network model has often been used to model time-sharing computer systems with a fixed degree of multiprogramming [9]. However, in an actual system with a fixed number of units, B, of main memory, the degree of multiprogramming fluctuates depending on the storage requirements of the jobs residing in main memory. Suppose that K classes of jobs can be distinguished such that a job in the kth class requires m_k, $1 \leq k \leq K$, units of main memory. Let $n_k(S)$ denote the number of class k jobs residing in main memory in state S of the system model. In this case, feasible states of the system model must satisfy the constraint

$$\sum_{k=1}^{K} m_k n_k(S) \leq N.$$

Note that the degree of multiprogramming is equal to $\Sigma_{k=1}^{K} n_k(S)$.

- *Outline of this paper*

An outline of the balance of this paper is given here. In the next section, the class of queuing networks considered in [2] is defined. This class of networks will be denoted by $\mathcal{N}$. Mechanisms of state-dependent lost and triggered arrivals and the associated loss and trigger functions which define the state dependency are next introduced. The class of networks in $\mathcal{N}$ with the addition of the loss and trigger mechanisms will be denoted by $\mathcal{N}^*$. A theorem is presented which gives a sufficient condition, involving the loss and trigger functions, for the equilibrium network state probability distribution to have the product form. The known class of queuing networks with a product form solution is thus enlarged. It is then shown that evaluation of the normalization constant for this new class of networks can be accomplished by considering a single queuing network with multiple closed subchains. Numerical solution for the normalization constant can thus utilize the convolutional method of Reiser and Kobayashi [10]. Finally, potential applications to modeling computer communication systems are discussed.

Proof of the theorem is based upon the technique of local balance equations described by Chandy [11]. In Appendix A, the notation needed to describe the network state and the balance equations is defined. It is shown that the complete set of local balance equations can be classified into four kinds and represented by four general equations. Based upon the notation and definitions introduced in Appendix A, a proof of the theorem is given in Appendix B.

Definition of queuing networks in $\mathcal{N}$

There is a finite number M of service stations and a finite number R of different classes of customers. A customer who completes service at station i in class r will next require service at station j in class s with a fixed probability $p_{ir;js}$. The routing matrix $\mathbf{P} = [p_{ir;js}]$ can be considered as defining a Markov chain with states indexed by the pairs (i, r). This Markov chain is assumed to be decomposable into K ergodic subchains (routing subchains). Let $E_1, E_2, \cdots, E_K$ denote the sets of (i, r) in these subchains. If n_{ir} denotes the number of class r customers at station i in network state S, then by our earlier definitions

$$n_k(S) = \sum_{(i,r)\in E_k} n_{ir}, \qquad k = 1, 2, \cdots, K,$$

and

$$N(S) = \sum_{k=1}^{K} n_k(S).$$

Four types of service stations are considered:

1. First-come-first-served service discipline, one or more servers (FCFS);
2. Processor-sharing service discipline, a single server (PS);
3. No queuing, arbitrarily many servers (IS);
4. Last-come-first-served preemptive resume service discipline, a single server (LCFS).

In an FCFS service station, all customers have the same exponential service time distribution with a service rate $\mu(j)$ dependent on the number j of customers at the service station. In PS, IS or LCFS service stations, each class of customers may have its own general service time distribution which has a rational Laplace transform. For the sake of clarity, several other forms of state-dependent service rates [2] are not considered for the moment.

Customers arrive from external sources according to one of two possible types of state-dependent arrival processes:

1. Customers arrive from a single Poisson stream with state-dependent rate $\lambda(N(S))$. An arrival enters station i for service in class r with a fixed probability q_{ir};
$$\sum_{i=1}^{M} \sum_{r=1}^{R} q_{ir} = 1.$$
2. Corresponding to the K routing subchains, there are K Poisson arrival streams with state-dependent rates $\lambda_k(n_k(S))$, $1 \leq k \leq K$. An arrival from the kth Poisson stream enters station i for service in class r with a fixed probability q_{ir};
$$\sum_{(i,r)\in E_k} q_{ir} = 1 \qquad \text{for all } k.$$

State-dependent lost and triggered arrivals

Without any loss of generality we shall assume that for a Type 1 arrival process $\lambda(m)$ is positive for $m \geq 0$ and for a Type 2 arrival process $\lambda_k(m)$ is positive for $m \geq 0$ and $k = 1, 2, \cdots, K$. Also, the service rate of any service station is positive whenever there is a customer waiting for service. Define

$$\mathbf{y} = (y_1, y_2, \cdots, y_K)$$

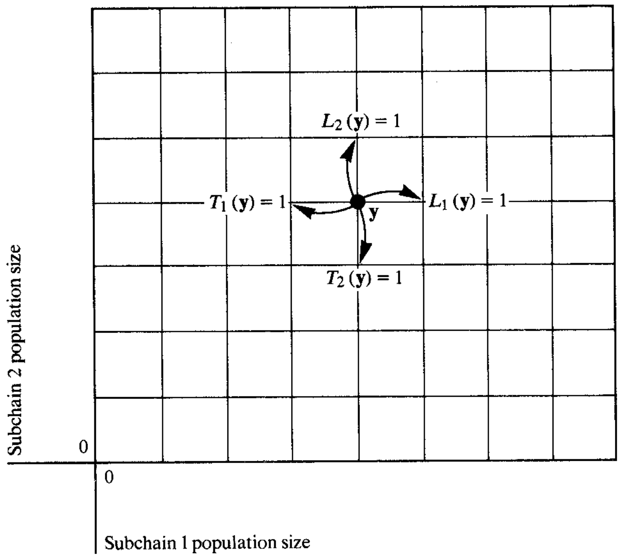


Figure 1 A graphical representation of the loss and trigger functions.

and let Y be the set of K-tuples $\{\mathbf{y} | y_k$ nonnegative integers$\}$.

Two sets of functions $L_k(\cdot)$ and $T_k(\cdot)$, $1 \le k \le K$, are now introduced which map the space Y into the set $\{0, 1\}$. Here $L_k(\cdot)$ is defined to be the *loss function* for the kth routing subchain such that in network state S, an external arrival from the kth Poisson stream is accepted if $L_k(\mathbf{n}(S)) = 1$; it is lost if $L_k(\mathbf{n}(S)) = 0$. Here $T_k(\cdot)$ is defined to be the *trigger function* for the kth routing subchain such that in network state S, the departure of a customer from the kth routing subchain triggers the immediate injection of an external arrival into the kth subchain if $T_k(\mathbf{n}(S)) = 0$; there is no triggered arrival if $T_k(\mathbf{n}(S)) = 1$. A class r customer completing service at station i leaves the network with probability

$$1 - \sum_{(j,s)\in E_k} p_{ir;js},$$

where $(i, r)\in E_k$. However, if $T_k(\mathbf{n}(S)) = 0$, a new customer is instantaneously injected into the network and it joins station j in class s with probability q_{js}, $(j, s)\in E_k$.

Specification of the loss and trigger functions completes the definition of a queuing network in $\mathcal{N}^*$.

State space decomposition and representation

It is known that if the Laplace transform of a given service time distribution is a rational function, the distribution can be represented by a series of exponential stages [2, 12]. Consequently, by introducing an appropriate (discrete) state space, the queuing network behavior can be characterized by a birth-death process (a continuous time Markov chain not to be confused with the Markov chain defined by the routing matrix $\mathbf{P}$). The equilibrium probability distribution $P(S)$, if it exists, is determined by a system of linear equations, also known as *balance equations*, namely

$$P(S)[\text{rate of flow out of } S] = \sum_{S'\in\mathcal{S}} P(S')[\text{rate of flow from } S' \text{ to } S] \quad \text{for all } S\in\mathcal{S}, \qquad (1)$$

where $\mathcal{S}$ is the set of feasible states for the network under consideration, determined by the specific service disciplines, service time distributions, and routing probabilities as well as by the loss and trigger functions.

Now $\mathcal{S}$ can be decomposed into disjoint subsets of feasible states in the following manner. Let

$$\mathcal{S}(\mathbf{y}) \triangleq \{S\in\mathcal{S} | \mathbf{n}(S) = \mathbf{y}\}.$$

Thus,

$$\mathcal{S} = \bigcup_{\mathbf{y}\in Y} \mathcal{S}(\mathbf{y}).$$

Two kinds of state transitions can be distinguished. The first kind is between adjacent states within the same subset $\mathcal{S}(\mathbf{y})$. The second kind is between a state in $\mathcal{S}(\mathbf{y})$ and an adjacent state in a neighboring subset $\mathcal{S}(\mathbf{y} + \mathbf{l}_k)$ or $\mathcal{S}(\mathbf{y} - \mathbf{l}_k)$ where $\mathbf{l}_k$ is defined to be the unit K-vector with its kth component equal to 1 and all other components equal to zero. The latter kind of transition corresponds to arrivals into or departures from the network and is controlled by the loss and trigger functions.

In Fig. 1, a graphical representation of the loss and trigger functions is shown for a network with two routing subchains ($K = 2$). The subset $\mathcal{S}(\mathbf{y})$ of feasible states is represented by a dot at point $\mathbf{y}$; $L_k(\mathbf{y}) = 1$ is represented by an arrow from $\mathbf{y}$ to $\mathbf{y} + \mathbf{l}_k$. The arrow signifies a positive rate of flow from the states $\mathcal{S}(\mathbf{y})$ into the states $\mathcal{S}(\mathbf{y} + \mathbf{l}_k)$. The absence of an arrow implies that the function has the value zero at $\mathbf{y}$. Similarly, $T_k(\mathbf{y}) = 1$ is represented by an arrow from $\mathbf{y}$ to $\mathbf{y} - \mathbf{l}_k$. The absence of an arrow implies that $T_k(\mathbf{y}) = 0$.

By using the above graphical representation, networks in $\mathcal{N}$ with open, closed and mixed routing subchains [2] are illustrated in Figs. 2(a), 2(b) and 2(c) for $K = 2$.

Jackson [1] first introduced state-dependent lost and triggered arrivals as a function of the total number $N(S)$ of customers in an exponential-server network. His ideas can be extended in a straightforward manner to define networks in $\mathcal{N}$ with the total population size $N(S)$ constrained between an upper bound and a lower bound [Fig. 2(d) if a Type 1 arrival process is assumed] or with individual routing subchain population sizes constrained between upper and lower bounds [Fig. 2(e) if a Type 2 arrival process is assumed]. All net-

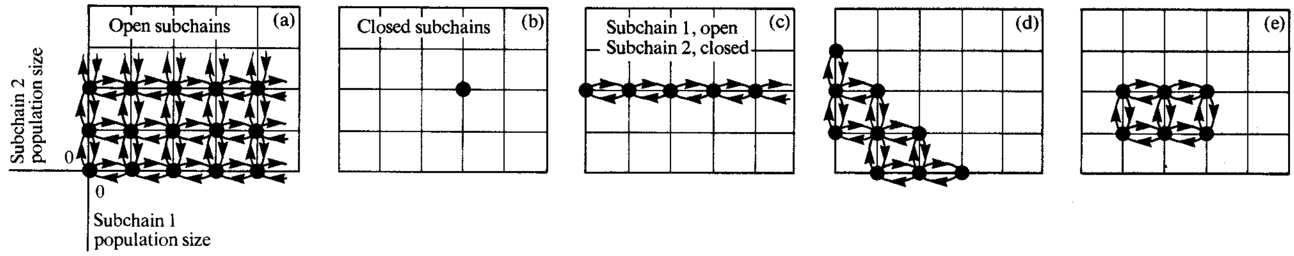


Figure 2 Queuing networks in $\mathcal{N}$ with different population size constraints.

works shown in Fig. 2 are thus within $\mathcal{N}$, which is known to possess a product form solution [2].

Introduction of state-dependent lost and triggered arrivals as a function of $\mathbf{n}(S)$ defines the extended class of networks $\mathcal{N}^*$ with arbitrary population size constraints. As an example, one such network with $K = 2$ is shown in Fig. 3. Since arbitrary loss and trigger functions are allowed, it is possible for the set of feasible states to be decomposable into two or more closed sets [13]. In Fig. 3, three closed sets are illustrated, namely

$$\mathcal{S}(V) = \bigcup_{\mathbf{y}\in V} \mathcal{S}(\mathbf{y});$$

$$\mathcal{S}(V') = \bigcup_{\mathbf{y}\in V'} \mathcal{S}(\mathbf{y});$$

$$\mathcal{S}(V'') = \bigcup_{\mathbf{y}\in V''} \mathcal{S}(\mathbf{y}),$$

together with some transient states. We are primarily interested in the equilibrium probability distribution, assuming a known initial network state. Thus, without any loss of interesting generality we can consider an irreducible Markov chain defined on a single closed set. We shall let $\mathcal{S}(V)$ denote the set of feasible states of the irreducible Markov chain where V is the corresponding set of feasible population vectors.

A sufficient condition

Theorem For a queuing network in $\mathcal{N}^*$, if the Markov chain defined on the set $\mathcal{S}(V)$ of feasible states is ergodic and if the loss and trigger functions satisfy the condition

(A) For each $k = 1, 2, \cdots, K$, $T_k(\mathbf{y}) = 1$ if and only if $L_k(\mathbf{y} - \mathbf{1}_k) = 1$ for all pairs of $\mathbf{y}$ and $\mathbf{y} - \mathbf{1}_k$ in V,

then the equilibrium network state probability distribution is given by the product form

$$P(S = (X_1, X_2, \cdots, X_M)) = C\, d(S) \prod_{i=1}^{M} f_i(X_i), \quad S \in \mathcal{S}(V), \tag{2}$$

where X_i represents the conditions at service station i in network state S, the functions $d(\cdot)$ and $f_i(\cdot)$ are given in Appendix A, and the normalization constant C is given by

$$\frac{1}{C} = \sum_{\mathbf{y}\in V} \sum_{S\in\mathcal{S}(\mathbf{y})} P(S). \tag{3}$$

A proof of the theorem is given in Appendix B. With our graphical representation of loss and trigger functions, condition (A) in the above theorem means that neighboring points in V must either be "doubly connected" or not connected at all. For instance, in Fig. 3, condition (A) is satisfied for $\mathcal{S}(V)$. It is not satisfied for $\mathcal{S}(V')$.

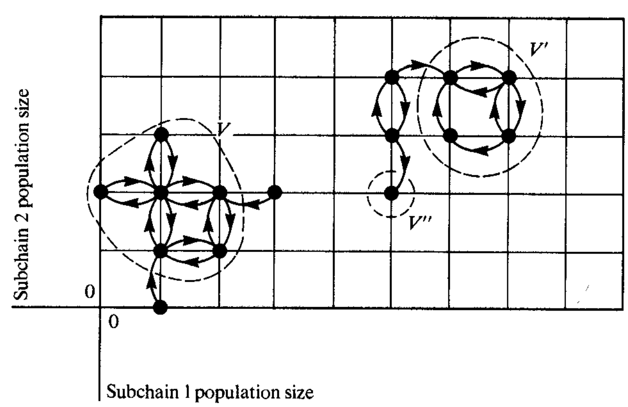
- *State-dependent service rates*

The product form given in Appendix A may be slightly modified to apply to networks in $\mathcal{N}$ with several forms of state-dependent service rates as shown by Baskett et al. [2]. It can be easily shown that the above theorem remains valid for such networks.

- *Evaluation of the normalization constant*

The normalization constant C in Eq. (3) may be computed by first evaluating the sums

Figure 3 A queuing network in $\mathcal{N}^*$ with general loss and trigger functions.



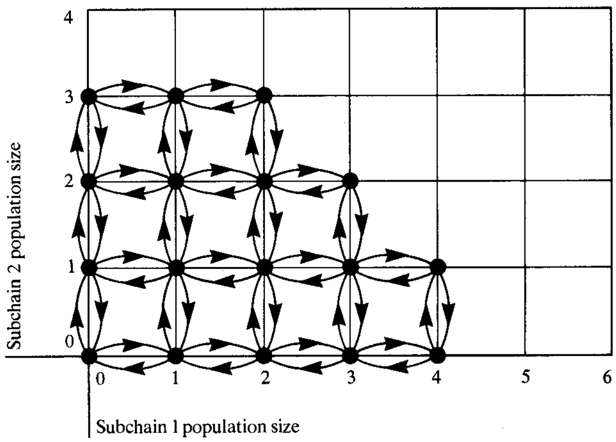


Figure 4 Loss and trigger functions for a store-and-forward node model.

$$\frac{1}{C(\mathbf{y})} \triangleq \sum_{S \in \mathscr{S}(\mathbf{y})} \prod_{i=1}^{M} f_i(X_i), \qquad \mathbf{y} \in V.$$

Note that $C(\mathbf{y})$ is by definition the normalization constant of a network with closed subchains and can be determined by the convolution method of Reiser and Kobayashi [10]. A very useful observation here is that in the process of evaluating $C(\mathbf{y})$, all $C(\mathbf{y}')$ where $y_k' < y_k$ for any $k = 1, 2, \cdots, K$, are also determined. Define

$$y_k^* \triangleq \max_{\mathbf{y} \in V} y_k$$

and

$$\mathbf{y}^* \triangleq (y_1^*, y_2^*, \cdots, y_K^*).$$

If $C(\mathbf{y}^*)$ is computed by the convolution method, then C can be evaluated from

$$\frac{1}{C} = \sum_{\mathbf{y} \in V} \frac{d(S)}{C(\mathbf{y})}$$

where $C(\mathbf{y})$, $\mathbf{y} \in V$, have been obtained in the computation for $C(\mathbf{y}^*)$.

Applications

• *Models for store-and-forward nodes*

Example 1 discussed in the Introduction is now reexamined as a queuing network with state-dependent lost arrivals. Specifically, suppose that $K = 2$, $B = 5$, $B_1 = 4$ and $B_2 = 3$, and the set of feasible states is given by Fig. 4. Notice that condition (A) is satisfied for a product form solution.

The particular buffer management scheme considered in Example 1 and other buffer management schemes have been analyzed using a queuing network model [7, 8]. Similar schemes have been implemented in practice [14]. Store-and-forward nodes employing such schemes can be modeled as queuing networks with population size constraints. Applicability of the product form as a solution can be verified by checking condition (A).

• *Modeling of network flow control*

In a store-and-forward network, flow control algorithms are needed to prevent the network from being overwhelmed by input sources as well as to prevent a single user or a single user group from hoarding resources of the network to the detriment of others. A two-level control scheme has also been analyzed using a queuing network model [15]. At the first level a limit is placed on the total number of messages in the network. At the second level, disjoint groups of source-destination pairs are defined and separate limits are placed on the number of messages belonging to each group. Note that in the context of routing subchains, such population size constraints are equivalent to those of Example 1. Applicability of the product form as a solution for the queuing network model is a simple consequence of the above theorem.

• *Models for multiprogramming computer systems*

Multiprogramming computer systems discussed earlier in Example 2 may be modeled by queuing networks in $\mathscr{N}^*$. For instance, if in Example 2, $K = 2$, $m_1 = 1$, $m_2 = 2$ and $N = 5$, the set of feasible states in Fig. 5 represents a system with a varying degree of multiprogramming which ranges from 2 to 5. In this model, it is assumed that there is always at least one job in each of the two classes waiting to be swapped into main memory as a triggered arrival. In an actual system, these jobs must merely be previously lost arrivals. Jobs are swapped into main memory on two kinds of occasions, namely, a) at the arrival time of a new job if its storage requirement is less than the available main memory space at that time (otherwise, it is "lost"); and b) at the departure time of a job if the departure gives rise to more than two units of available main memory space (an arrival is triggered). We see from Fig. 5 that condition (A) is satisfied for a product form solution.

Conclusion

In this paper, we recognized that the traditional formulation of open, closed and mixed networks of queues is inadequate for modeling systems with various population size constraints. Mechanisms of state-dependent lost and triggered arrivals were introduced to model such constraints. A sufficient condition was found for the equilibrium network state probability distribution to have the product form. Thus, the known class of queuing networks with a product form solution has been extended. Potential applications to modeling computer communication systems with storage and flow control constraints were discussed.

Acknowledgments
The author thanks L. S. Woo of the IBM Thomas J. Watson Research Center for many helpful discussions. He is also indebted to the anonymous referees for their constructive comments.

Appendix A

In this appendix, the notation needed to describe the network state S and the balance equations is defined. The product form of $P(S)$ is given. It is shown that the set of local balance equations can be differentiated into four kinds so that the entire set can be represented by four general equations.

• *Network state description*
A service time distribution having a rational Laplace transform may be represented by a series of exponential stages [2, 12]. Consider the service time distribution of a class r customer at station i. Let u_{ir} be the total number of exponential stages. The time spent in the lth stage has a negative exponential distribution with mean $1/\mu_{irl}$. The probability that a customer completes service and leaves after the lth stage is $1 - a_{irl}$, where a_{irl} is the probability that the customer goes on to the $(l+1)$th stage. (Note that $a_{iru_{ir}} = 0$.) Define

$$A_{ir1} \triangleq 1$$

and

$$A_{irl} \triangleq \prod_{j=1}^{l-1} a_{irj}, \qquad 2 \leq l \leq u_{ir}.$$

The network state S is represented by a vector $(X_1, X_2, \cdots, X_M)$ where X_i represents the conditions prevailing at station i. Let n_i denote the number of customers at station i. If station i is FCFS, then $X_i = (x_{i1}, x_{i2}, \cdots, x_{in_i})$, where x_{ij} is the class of the customer who is jth in FCFS order. If station i is PS or IS, then $X_i = (v_{i1}, v_{i2}, \cdots, v_{iR})$, where v_{ir} is a vector $(m_{ir1}, m_{ir2}, \cdots, m_{iru_{ir}})$ where m_{irl} is the number of class r customers in the lth exponential stage of service. Finally, if station i is LCFS, then $X_i = ((r_1, l_1), (r_2, l_2), \cdots, (r_{n_i}, l_{n_i}))$ where r_j is the class and l_j is the stage of service of the jth customer in LCFS order.

We next define the following set of linear equations for each routing subchain:

$$e_{js} = \sum_{(i,r)\in E_k} e_{ir} p_{ir;js} + q_{js}, \qquad (j, s)\in E_k, \tag{A1}$$

where e_{ir} may be interpreted as the aggregate arrival rate of class r customers to station i. If the network is closed with respect to E_k, then $q_{js} = 0$ for any $(j, s)\in E_k$ in Eq. (A1). In this case, the e_{ir} are determined to within a multiplicative constant for each subchain.

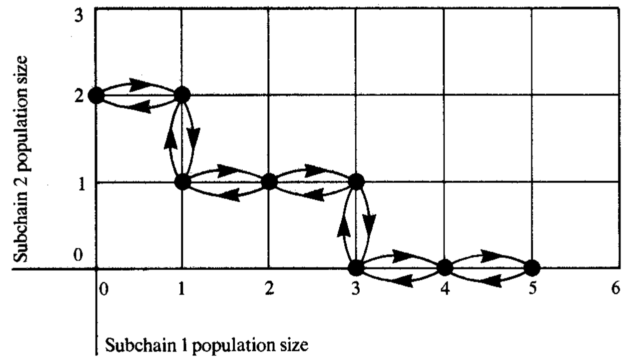


Figure 5 Loss and trigger functions for a multiprogramming system model.

• *The product form*
The product form of the equilibrium network state probability distribution is given by [2]

$$P(S = (X_1, X_2, \cdots, X_M)) = C\, d(S) \prod_{i=1}^{M} f_i(X_i),$$

where 1) C is a normalization constant; 2) f_i is a function that depends on the type of station i:

$$f_i(X_i) = \prod_{j=1}^{n_i} \left\{ e_{ix_{ij}} \frac{1}{\mu_i(j)} \right\} \qquad \text{if station } i \text{ is FCFS,}$$

$$f_i(X_i) = n_i! \prod_{r=1}^{R} \prod_{l=1}^{u_{ir}} \left\{ \left[\frac{e_{ir} A_{irl}}{\mu_{irl}} \right]^{m_{irl}} \frac{1}{m_{irl}!} \right\} \qquad \text{if station } i \text{ is PS,}$$

$$f_i(X_i) = \prod_{r=1}^{R} \prod_{l=1}^{u_{ir}} \left\{ \left[\frac{e_{ir} A_{irl}}{\mu_{irl}} \right]^{m_{irl}} \frac{1}{m_{irl}!} \right\} \qquad \text{if station } i \text{ is IS,}$$

$$f_i(X_i) = \prod_{j=1}^{n_i} \left\{ e_{ir_j} A_{ir_j l_j} \frac{1}{\mu_{ir_j l_j}} \right\} \qquad \text{if station } i \text{ is LCFS;}$$

and 3) if the network is closed allowing no external arrivals and departures, then $d(S) = 1$; otherwise,

$$d(S) = \prod_{m=0}^{N(S)-1} \lambda(m) \qquad \text{for a Type 1 arrival process,}$$

or

$$d(S) = \prod_{k=1}^{K} \prod_{m=0}^{n_k(S)-1} \lambda_k(m) \quad \text{for a Type 2 arrival process.}$$

• *A classification of local balance equations*
Recall that $P(S)$ must satisfy the set of balance equations in Eq. (1). It was observed by Chandy [11] that each such balance equation, which he termed a *global* balance equation, can be partitioned into a number of *local* balance equations. A solution that satisfies each of the local balance equations must also satisfy the sum of the local balance equations, namely, the global balance equation.

In order to write down the balance equations, additional notation is needed.

Let $S \triangleq (X_1, X_2, \cdots, X_M)$ and $\tilde{S} \triangleq (\tilde{X}_1, \tilde{X}_2, \cdots, \tilde{X}_M)$ be two adjacent network states. The following notation is defined:

- The state $\tilde{S} = S[X_i^{+r}]$ is the same as S except for $\tilde{X}_i$ which is $(r, x_{i1}, x_{i2}, \cdots, x_{in_i})$ where station i is FCFS.
- The state $\tilde{S} = S[X_j^{-s}]$ is the same as S except for $\tilde{X}_j$ which is $(x_{j1}, x_{j2}, \cdots, x_{j,n_j-1})$ where station j is FCFS and $s = x_{jn_j}$.
- The state $\tilde{S} = S[X_{il}^{+r}]$ is the same as S except for
  - i. $\tilde{m}_{irl}$ which is equal to $m_{irl} + 1$ where station i is PS or IS;
  - ii. $\tilde{X}_i$ which is $((r, l), (r_1, l_1), (r_2, l_2), \cdots, (r_{n_i}, l_{n_i}))$ where station i is LCFS.
- The state $\tilde{S} = S[X_{jl}^{-s}]$ is the same as S except for
  - i. $\tilde{m}_{jsl}$ which is equal to $m_{jsl} - 1$ where station j is PS or IS;
  - ii. $\tilde{X}_j$ which is $((r_2, l_2), (r_3, l_3), \cdots, (r_{n_i}, l_{n_i}))$ where station j is LCFS and $(s, l) = (r_1, l_1)$.
- A combination of the above notations (for example, $S[X_{il}^{+r}, X_j^{+s}]$) will have the same combined interpretation. For the special case $\tilde{S} = S[X_{j,l-1}^{+s}, X_{jl}^{-s}]$, where station j is LCFS and $(s, l) \triangleq (r_1, l_1)$, $\tilde{S}$ is the same as S except for $\tilde{X}_j$ which is $((r_1, l_1 - 1), (r_2, l_2), \cdots, (r_{n_j}, l_{n_j}))$.
- The rate of flow $\sigma_r(X_{il})$ due to a class r customer leaving the lth stage of service at station i is defined to be

  i. $\dfrac{m_{irl}\mu_{irl}}{n_i}$ if station i is PS;

  ii. $m_{irl}\mu_{irl}$ if station i is IS;

  iii. μ_{irl} if station i is LCFS.

  Similarly, the rate of flow $\sigma_r(X_{il}^{+r})$ is defined to be

  i. $\dfrac{(m_{irl} + 1)\mu_{irl}}{n_i + 1}$ if station i is PS;

  ii. $(m_{irl} + 1)\mu_{irl}$ if station i is IS;

  iii. μ_{irl} if station i is LCFS.

Let I_1 be the index set of FCFS service stations and I_2 be the index set of PS, IS and LCFS service stations. The set of local balance equations can be grouped into the following four categories.

1) *Local balance equations of the first kind* equate the rate of flow out of a feasible state S due to a customer leaving a FCFS station to the rate of flow into S due to a customer entering the same station. Consider, for instance, a network in $\mathcal{N}$. Local balance equations of the first kind for $S \in \mathcal{S}$ and $j \in I_1$ are given by

$$P(S)\mu_j(n_j) = \sum_{r=1}^{R} \Big\{ \sum_{i \in I_1} P(S[X_i^{+r}, X_j^{-s}])\mu_i(n_i + 1)p_{ir;js} + \sum_{i \in I_2} \sum_{l=1}^{u_{ir}} [P(S[X_{il}^{+r}, X_j^{-s}])\sigma_r(X_{il}^{+3}) \times (1 - a_{irl})p_{ir;js}] \Big\} + \tilde{\lambda} q_{js} P(S[X_j^{-s}]), \tag{A2}$$

where $s = x_{jn_j}$ in state S,

$$\tilde{\lambda} = \begin{cases} \lambda(N(S) - 1) & \text{for a Type 1 arrival process;} \\ \lambda_k(n_k(S) - 1) & \text{for a Type 2 arrival process, and } (j, s) \in E_k. \end{cases}$$

Note that if n_j in state S is zero, both sides of the above local balance equation are equal to zero.

2) *Local balance equations of the second kind* equate the rate of flow out of a feasible state S due to a class s customer leaving the first stage of service at a PS, IS, or LCFS station to the rate of flow into S due to a class s customer entering the first stage of service at the same station. Consider, for instance, a network in $\mathcal{N}$. Local balance equations of the second kind for $S \in \mathcal{S}$ and $j \in I_2$ are given by

$$P(S)\sigma_s(X_{j1}) = \sum_{r=1}^{R} \Big\{ \sum_{i \in I_1} P(S[X_i^{+r}, X_{j1}^{-s}])\mu_i(n_i + 1)p_{ir;js} + \sum_{i \in I_2} \sum_{l=1}^{u_{ir}} [P(S[X_{il}^{+r}, X_{j1}^{-s}])\sigma_r(X_{il}^{-r}) \times (1 - a_{irl})p_{ir;js}] \Big\} + \tilde{\lambda} q_{js} P(S[X_{j1}^{-s}]), \tag{A3}$$

where $\tilde{\lambda}$ is as defined above.

3) *Local balance equations of the third kind* equate the rate of flow out of a feasible state S due to a class s customer leaving the lth stage of service at a PS, IS or LCFS station to the rate of flow into S due to a class s customer entering the lth stage of service at the same station, where $l \geq 2$. Consider, for instance, a network in $\mathcal{N}$. Local balance equations of the third kind for $S \in \mathcal{S}$, $j \in I_2$ and $2 \leq l \leq u_{js}$ are given by

$$P(S)\sigma_s(X_{jl}) = P(S[X_{j,l-1}^{+s}, X_{j1}^{-s}])\sigma_s(X_{j,l-1}^{+s})a_{js,l-1}. \tag{A4}$$

4) *Local balance equations of the fourth kind* equal the rate of flow out of a feasible state S due to the arrival of a customer from an external source into a routing subchain to the rate of flow into S due to a customer departing from the same routing subchain. Consider, for instance, a network in $\mathcal{N}$. Local balance equations of the fourth kind for $S \in \mathcal{S}$ and $k = 1, 2, \cdots, K$ are given by

$$P(S)\hat{\lambda} \sum_{(i,r)\in E_k} d_{ir} = \sum_{\substack{(i,r)\in E_k\\ i\in I_1}} \Big[P(S[X_i^{+r}])\mu_i(n_i+1) \times \Big(1 - \sum_{(j,s)\in E_k} p_{ir;js}\Big)\Big] + \sum_{\substack{(i,r)\in E_k\\ i\in I_2}} \Big[\Big(1 - \sum_{(j,s)\in E_k} p_{ir;js}\Big) \times \sum_{l=1}^{u_{ir}} P(S[X_{il}^{+r}])\sigma_r(X_{il}^{+r})(1-a_{irl})\Big], \quad \text{(A5)}$$

where

$$\hat{\lambda} = \begin{cases} \lambda(N(S)) & \text{for a Type 1 arrival process;} \\ \lambda_k(n_k(S)) & \text{for a Type 2 arrival process.} \end{cases}$$

Appendix B

• *Proof of theorem*

The theorem is proved by showing that the product form given in Appendix A satisfies the four kinds of local balance equations for all feasible states in $\mathscr{S}(V)$.

Consider a feasible network state $S \in \mathscr{S}(\mathbf{y})$ where $\mathbf{y}$ is a feasible population vector in V.

We first note that local balance equations of the third kind are not affected by the loss and trigger functions. They are still given by Eq. (A4) and are thus satisfied by the product form as a solution.

Consider local balance equations of the first and second kinds and focus upon class s customers at station j, $(j,s)\in E_k$. With loss and trigger functions which satisfy condition (A), there are three possible cases:

Case 1 $\mathbf{y} - \mathbf{1}_k$ is not in V and $y_k > 0$, which imply that $T_k(\mathbf{y}) = 0$. (If $y_k = 0$, then the local balance equations given below reduce to the trivial case of $0 = 0$.)

Local balance equations of the first kind are given by

$$P(S)\mu_j(n_j) = \sum_{\substack{i\in I_1\\ (i,r)\in E_k}} \Big\{ P(S[X_i^{+r}, X_j^{-s}])\,\mu_i(n_i+1) \times \Big[p_{ir;js} + \Big(1 - \sum_{(m,t)\in E_k} p_{ir;mt}\Big)q_{js}\Big]\Big\} + \sum_{\substack{i\in I_2\\ (i,r)\in E_k}} \Big\{\Big[p_{ir;js} + \Big(1 - \sum_{(m,t)\in E_k} p_{ir;mt}\Big)q_{js}\Big] \times \sum_{l=1}^{u_{ir}} P(S[X_{il}^{+r}, X_j^{-s}])\sigma_r(X_{il}^{+r})(1-a_{irl})\Big\}, \quad \text{(B1)}$$

where $s = x_{jn_j}$ in state S. Local balance equations of the second kind are given by

$$P(S)\sigma_s(X_{j1}) = \sum_{\substack{i\in I_1\\ (i,r)\in E_k}} \Big\{ P(S[X_i^{+r}, X_{j1}^{-s}])\mu_i(n_i+1) \times \Big[p_{ir;js} + \Big(1 - \sum_{(m,t)\in E_k} p_{ir;mt}\Big)q_{js}\Big]\Big\} + \sum_{\substack{i\in I_2\\ (i,r)\in E_k}} \Big\{\Big[p_{ir;js} + \Big(1 - \sum_{(m,t)\in E_k} p_{ir;mt}\Big)q_{js}\Big] \times \sum_{l=1}^{u_{ir}} P(S[X_{il}^{+r}, X_{j1}^{-s}])\sigma_r(X_{il}^{+r})(1-a_{irl})\Big\}. \quad \text{(B2)}$$

After substituting the product form in Appendix A for $P(S)$, $P(S[X_i^{+r}, X_j^{-s}])$, $P(S[X_{il}^{+r}, X_j^{-s}])$, $P(S[X_i^{+r}, X_{j1}^{-s}])$ and $P(S[X_{il}^{+r}, X_{j1}^{-s}])$ into the above equations, dividing throughout by $P(S)$ and rearranging, it is sufficient to show

$$e_{js} = \sum_{(i,r)\in E_k} e_{ir}\Big[p_{ir;js} + \Big(1 - \sum_{(m,t)\in E_k} p_{ir;mt}\Big)q_{js}\Big]. \quad \text{(B3)}$$

We proceed by considering the right-hand side (RHS) of Eq. (B3).

$$\begin{aligned} \text{RHS} &= \sum_{(i,r)\in E_k} e_{ir}p_{ir;js} + q_{js}\Big[\sum_{(i,r)\in E_k} e_{ir} - \sum_{(i,r)\in E_k}\sum_{(m,t)\in E_k} e_{ir}p_{ir;mt}\Big] \\ &= \sum_{(i,r)\in E_k} e_{ir}p_{ir;js} + q_{js}\sum_{(m,t)\in E_k}\Big[e_{mt} - \sum_{(i,r)\in E_k} e_{ir}p_{ir;mt}\Big] \\ &= \sum_{(i,r)\in E_k} e_{ir}p_{ir;js} + q_{js}\sum_{(m,t)\in E_k} q_{mt} \\ &= \sum_{(i,r)\in E_k} e_{ir}p_{ir;js} + q_{js} \\ &= e_{js} = \text{left-hand side of Eq. (B3).} \end{aligned}$$

In the above derivation we have made use of Eq. (A1) which defines the $\{e_{ir}\}$.

Case 2 $\mathbf{y} - \mathbf{1}_k$ is in V, $T_k(\mathbf{y}) = 0$ and $L_k(\mathbf{y} - \mathbf{1}_k) = 0$. In this case, local balance equations of the first and second kinds are again given by Eqs. (B1) and (B2), respectively, and are thus each satisfied by the product form as a solution.

Case 3 $\mathbf{y} - \mathbf{1}_k$ is in V, $T_k(\mathbf{y}) = 1$ and $L_k(\mathbf{y} - \mathbf{1}_k) = 1$. In this case, local balance equations of the first and second kinds are given by Eqs. (A2) and (A3), respectively, and are thus satisfied by the product form as a solution.

Finally, consider local balance equations of the fourth kind and focus upon the kth subchain. There are again three possible cases:

Case 1 $\mathbf{y} + \mathbf{1}_k$ is not in V which implies that $L_k(\mathbf{y}) = 0$. The local balance equation for the kth subchain reduces to the trivial case of $0 = 0$.

Case 2 $\mathbf{y} + \mathbf{1}_k$ is in V, $L_k(\mathbf{y}) = 1$ and $T_k(\mathbf{y} + \mathbf{1}_k) = 1$. In this case, local balance equations of the fourth kind are given by Eq. (A5) and are thus satisfied by the product form as a solution.

Case 3 $\mathbf{y} + \mathbf{1}_k$ is in V, $L_k(\mathbf{y}) = 0$ and $T_k(\mathbf{y} + \mathbf{1}_k) = 0$. The local balance equation for the kth subchain reduces to the trivial case of $0 = 0$.

References and note

1. J. R. Jackson, "Jobshop-Like Queueing Systems," *Manage. Sci.* **10**, 131 (1963).
2. F. Baskett, K. M. Chandy, R. R. Muntz, and F. G. Palacios, "Open, Closed and Mixed Networks of Queues with Different Classes of Customers," *J. ACM* **22**, 248 (1975).
3. This paper is a new version of the report [4] with extensive revisions. In particular, the main theorem has been strengthened to cover both Type 1 and Type 2 arrival processes.
4. S. S. Lam, "Queueing Networks with Lost and Triggered Arrivals," *Research Report RC 5540*, IBM T. J. Watson Research Center, Yorktown Heights, New York, July 1975.
5. S. S. Lam, "Store-and-Forward Buffer Requirements in a Packet Switching Network," *IEEE Trans. Commun.* **COM-24**, 394 (1976).
6. P. J. Schweitzer and S. S. Lam, "Buffer Overflow in a Store-and-Forward Network Node," *IBM J. Res. Develop.* **20**, 542 (1976).
7. M. Irland, "Queueing Analysis of a Buffer Allocation Scheme for a Packet Switch," *Conference Record*, National Telecommunications Conference, New Orleans, December 1975, p. 24.
8. F. Kamoun and L. Kleinrock, "Analysis of Shared Storage in a Computer Network," *Proceedings of the Ninth Hawaii International Conference on System Sciences,* Honolulu, January 1976, p. 89.
9. J. P. Buzen, "Queueing Network Models of Multiprogramming," Ph.D. Thesis, Harvard University, Cambridge, MA, May 1971.
10. M. Reiser and H. Kobayashi, "Queuing Networks with Multiple Closed Chains: Theory and Computational Algorithms," *IBM J. Res. Develop.* **19,** 283 (1975).
11. K. M. Chandy, "The Analysis and Solutions for General Queueing Networks," *Proceedings of the Sixth Annual Princeton Conference on Information Sciences and Systems*, Princeton University, NJ, March 1972, p. 224.
12. D. R. Cox, "A Use of Complex Probabilities in the Theory of Stochastic Processes," *Proc. Cambridge Phil. Soc.* **51,** 313 (1955).
13. W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. I, John Wiley and Sons, Inc., New York, 1950.
14. J. M. McQuillan, et al., "Improvements in the Design and Performance of the ARPA Network," *AFIPS Conference Proceedings*, (Fall Joint Computer Conference, Anaheim, CA) **41,** 741 (1972).
15. J. Wong and M. Unsoy, "Analysis of Flow Control in Switched Data Networks," to be presented at IFIP Congress, Toronto, August 1977.

Received September 23, 1976; revised March 10, 1977

The author is located at the IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598.