OPERATIONS RESEARCH CENTER

working paper

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

A Two Priority M/G/1 Queue
with Feedback

By

J. Keilson and L.D. Servi

OR 177-88 March 1988

A Two Priority M/G/1 Queue with Feedback

by

J. Keilson
GTE Laboratories Incorporated and Mass. Inst. of Tech.

and

L. D. Servi
GTE Laboratories Incorporated

ABSTRACT

An M/G/1 queueing system with two priority classes of traffic and Bernoulli feedback for the low priority class is studied. For the low priority customers, the distributions of the ergodic time to completion, the first pass time, and the number in the system are found in the transform domain and the first moments are displayed. The distributed Poisson approximation introduced previously by the authors is then employed to analyze systems with a preempt-resume high priority clocked schedule and low priority traffic with feedback.

1. Introduction

The following system which we will refer to as the *feedback* system is studied: a) the system consists of two classes of traffic, C_1 and C_2, each having Poisson arrivals and i.i.d. service times; b) the C_1 customers have preempt-resume priority over C_2; c) upon completion of service of any C_2 task, that task feeds back with probability θ to the end of the C_2 queue to again await service. This system contains as special cases: i) the M/G/1 system, ii) a two class M/G/1 system with prempt-resume priority as considered, for example, by Keilson and Sumita [9], and iii) an M/G/1 system with feedback considered, for example, by Takács [11] and Disney [1],[2]. The elements of priority and feedback interact in the feedback system and compound the effort required for its analysis. Many of the formulae obtained, however, can be given a structurally simple and informative form.

The study was motivated by concern for the performance of a telecommunications switch with priority and feedback features. The distributed Poisson approximation introduced in Keilson and Servi [7] and the results obtained for our feedback system permit exact analysis of a system which is a fair approximation to one such telecommunications switch.

In Section 2, notation is first introduced. Next, using analysis of a system *without* feedback, the p.g.f. of the number of C_2 customers in the system in the *feedback* system is found as well as the mean time in the system for the C_2 customers. Using an approach related to that of Keilson and Kooharian [6] the equations for the state-space motion are found in Section 3 and solved in Section 4. In Section 5 this analysis is used to find the transform of the joint distribution of the duration of the first pass time in the system and the number in the system at the end of the first pass time. This leads to the Laplace transform of the first pass time and its expected value. In Section 6 the transform is found of the joint distribution of the sojourn time of the low priority customers and the number in the system at the epochs at which low priority customers leave the system. In Section 7 the busy period is examined. Finally, in Section 8 the analysis of Section 2 is applied to preempt-resume clocked schedules with feedback of low priority customers using the distributed Poisson approximation introduced analysis in [7]. The results are then validated numerically.

2.Notation. The mean time in system for low priority customers.

The following notation will be helpful for reference. Each row describes an M/G/1 system whose Poisson input stream is the class listed in the first column.

Class	arrival rate	service time	service time c.d.f	LST	utilization	ergodic waiting time
C_1	λ_1	T_1	$A_{T1}(x)$	$\alpha_{T1}(s)$	ρ_1	W_1
C_2	λ_2	T_2	$A_{T2}(x)$	$\alpha_{T2}(s)$	ρ_2	W_2
C_{2EFF}	λ_2	T_{2EFF}	$A_{T2EFF}(x)$	$\alpha_{T2EFF}(s)$	$\rho_{2EFF}=\rho_2(1-\theta)^{-1}$	W_{2EFF}
C_{12}	$\lambda_1+\lambda_2$	T_{12}	$A_{T12}(x)$	$\alpha_{T12}(s)$	$\rho_{12}=\rho_1+\rho_2$	W_{12}
C_S	$\lambda_1+\lambda_2$	T_S	$A_{TS}(x)$	$\alpha_{TS}(s)$	$\rho_S=\rho_1+\rho_2(1-\theta)^{-1}$	W_S

TABLE 2.1

The subscript 12 designates the priorityless system of C_1 and C_2 customers with no feedback . The corresponding service time transform is given by

$$\alpha_{T12}(s)=\frac{\lambda_1}{\lambda_1+\lambda_2}\alpha_{T1}(s)+\frac{\lambda_2}{\lambda_1+\lambda_2}\alpha_{T2}(s)\,. \tag{2.1}$$

The subscript 2EFF designates a class without feedback whose effective service time is a compound geometric mixture of multiple convolutions of $A_{T2}(x)$ with itself. The corresponding service time transform is given by

$$\alpha_{T2EFF}(s)=(1-\theta)\sum_{i=0}^{\infty}\theta^i\alpha^{i+1}{}_{T2}(s)\;=\;\frac{(1-\theta)\alpha_{T2}(s)}{1-\theta\alpha_{T2}(s)}\,. \tag{2.2}$$

This may be thought of as arising from the class of customers C_2 feeding back to the head of the line with probability θ.

The subscript S designates the priorityless superposition of the classes C_1 and $C_{2EFF.}$ The corresponding service time transform is given by

$$\alpha_{TS}(s)=\frac{\lambda_1}{\lambda_1+\lambda_2}\alpha_{T1}(s)+\frac{\lambda_2}{\lambda_1+\lambda_2}\alpha_{T2EFF}(s).$$

Equivalently, from (2.2),

$$\alpha_{TS}(s)=\frac{\lambda_1\alpha_{T1}(s)(1-\theta\alpha_{T2}(s))+\lambda_2(1-\theta)\alpha_{T2}(s)}{(\lambda_1+\lambda_2)(1-\theta\alpha_{T2}(s))}\,. \tag{2.3}$$

This should be contrasted with the feedback system of focal interest where the C_2 customers return to the back of the line with probability θ and C_1 has preempt-resume priority over C_2.

The following representative notation will also be employed.

$N(t)$	the number of C_2 customers in the system at time t.
T_{sys}	the ergodic time in system for C_2 customers in the feedback system.
θ	the probability of feedback for C_2 customers.
T_{BP1}	the busy period for C_1 customers in the absence of C_2.
$\sigma_{BP1}(s)$	$E[\exp(-sT_{BP1})]$, the busy period transform for C_1 customers.
$\phi_{PK12}(s)$	$E[\exp(-sW_{12})]$, the waiting time Pollaczek-Khinchin transform for the priorityless C_{12} system.
$\phi_{PKS}(s)$	$E[\exp(-sW_S)]$, the waiting time Pollaczek-Khinchin transform for the priorityless C_S system.

An upper case asterisk will always designate forward recurrence time, e.g.

$$\alpha^*_{T12}(s) = \frac{1 - \alpha_{T12}(s)}{E[T_{12}]s} \quad \text{and} \quad \alpha^*_{TS}(s) = \frac{1 - \alpha_{TS}(s)}{E[T_S]s} \tag{2.4}$$

With this notation

$$\phi_{PK12}(s) = \frac{1 - \rho_{12}}{1 - \rho_{12}\alpha^*_{T12}(s)} \quad \text{and} \quad \phi_{PKS}(s) = \frac{1 - \rho_S}{1 - \rho_S\alpha^*_{TS}(s)} \tag{2.5}$$

Case 1: The priority system with no feedback ($\theta = 0$)

The priority system with no feedback has been treated by Keilson and Sumita [9] and the distribution of the low priority customer delay is exhibited there. A much simpler derivation of this distribution is given next which sets the stage for the more difficult feedback case.

The arrival of a tagged C_2 customer in steady state finds an ergodic backlog distribution equal to that of a priorityless system of C_1 and C_2 traffic having arrival rates λ_1 and λ_2 and service times c.d.f. of $A_{T1}(x)$ and $A_{T2}(x)$. From TABLE 2.1 the L-S transform of this backlog is $\phi_{PK12}(s)$. All subsequent C_2 customers have no effect on the delay time of the tagged C_2 customer. However subsequent C_1 arrivals modify the delay time as interruptions of rate λ_1 with each interruption time equal to that of T_{BP1}, the C_1 busy period. It follows at once that the C_2

waiting time distribution has an Laplace transform given by $\phi_{PK12}(s + \lambda_1 - \lambda_1\sigma_{BP1}(s))$ [4], [5]. Using similiar reasoning the time in the system has a Laplace Transform

$$\phi_{T12}(s) = \phi_{PK12}(s + \lambda_1 - \lambda_1\sigma_{BP1}(s)) \ \alpha_{T2}(s + \lambda_1 - \lambda_1\sigma_{BP1}(s)) . \tag{2.6}$$

Case 2: The feedback system.

The process N(t) is insensitive to a certain variant on the underlying order of service discipline. This insensitivity enables one to employ Little's Law [10] to evaluate the expectation $E[T_{sys}]$ directly.

Theorem 2.1:

a) The system is stable when $\rho_S = \rho_1 + \rho_2 (1-\theta)^{-1} < 1$ and the idle state probability at ergodicity, $E(\infty)$, is given by $1 - \rho_S$.

b) The probability generating function of $N(\infty)$ is given by

$$h(u) = \phi_{PKS}(\zeta(u))\, \alpha_{T2EFF}(\zeta(u)) \tag{2.7}$$

where

$$\zeta(u) = \lambda_2(1-u) + \lambda_1[\, 1 - \sigma_{BP1}(\lambda_2(1-u))\,], \tag{2.8}$$

and $\alpha_{T2EFF}(s)$ and $\phi_{PKS}(s)$ are defined in (2.2) and (2.5).

c) $$E[N(\infty)] = \lambda_2 E[T_{sys}], \tag{2.9}$$

and

d) For a low priority customer, the mean time in the system is given by

$$E[T_{sys}] = \frac{(1-\theta)\lambda_1 E[T_1{}^2] + \lambda_2 E[T_2{}^2] + 2E[T_2](1-\rho_1-\rho_2)}{2(1-\rho_1)\{(1-\theta)(1-\rho_1)-\rho_2\}} . \tag{2.10}$$

Proof: The ergodic number of low priority customers in the system, $N(\infty)$, would be the same in distribution if the low priority customers that feed back were to feed back to the *front* of the line rather than to the *back* of the line. Hence, h(u) must also be the same under *either* schedule. System behavior under the front of the line schedule is equivalent to that for a two priority system in which the low priority customers have an arrival rate λ_2 and a service time T_{2EFF} having the

Laplace transform given in (2.2). Using the reasoning of (2.6) the time in system for the low priority customers of a two priority M/G/1 system with arrival rates λ_1 and λ_2 and service times T_1 and T_{2EFF} is given by

$$E[\exp(-sT_{sys})] = \phi_{PKS}(s + \lambda_1 - \lambda_1\sigma_{BP1}(s))\, \alpha_{T2EFF}(s + \lambda_1 - \lambda_1\sigma_{BP1}(s)) \tag{2.11}$$

where $\phi_{PKS}(s)$ and $\sigma_{BP1}(s)$ are defined in TABLE 2.1.

From the distributional form of Little's Law [8], provided for conveniencein Appendix A.1, the distribution of $N(\infty)$ and T_{sys} are related to each other according to

$$h(u) = E[u^{N(\infty)}] = E[\exp(-(\lambda_2(1-u))T_{sys})] \tag{2.12}$$

so (2.7) follows.

Equation (2.9) follows from either Little's Law [10] or from differentiation of (2.12). From (2.11)

$$E[T_{sys}] = \frac{E[W_S] + E[T_{2EFF}]}{1 - \rho_1} \tag{2.13}$$

where W_S , T_{2EFF} and ρ_1 are defined in TABLE 2.1. Equation (2.10) then follows from (2.13) and the Pollaczek - Khintchine formula,

$$E[W_S] = \frac{\rho_S}{1-\rho_S} \frac{E[T_S^2]}{2E[T_S]} \tag{2.14}$$

where T_S and ρ_S are defined in TABLE 2.1. (Alternatively (2.10) follows from (2.7) and the observation that $E[N(\infty)] = h'(1)$.)♦

Remark: Note that when $\lambda_1 = 0$ and $\theta = 0$, (2.10) coincides with $E[T_{sys}]$ for the ordinary M/G/1 queue. When $\lambda_1 = 0$, (2.10) agrees with $E[T_{sys}]$ for the M/G/1 queue with feedback as given by [11, equation (35)]. When $\theta = 0$, agreement is found with $E[T_{sys}]$ for the preempt-resume M/G/1 queue with two classes of traffic given in [9, equation (3.8f)].

In the case of no high priority customers, i.e., $\lambda_1 = 0$, equation (2.7) is equivalent to

$$h(u) = \frac{(1-u)(1-\theta-\rho_2)\,\alpha_{T2}(\lambda_2(1-u))}{\theta u \alpha_{T2}(\lambda_2(1-u)) + (1-\theta)\alpha_{T2}(\lambda_2(1-u)) - u}$$

so

$$h(u) - E(\infty) = \frac{-(1-\rho_2(1-\theta)^{-1})\, u\, (1-\alpha_{T2}(\lambda_2(1-u)))}{u - (\theta u+(1-\theta))\, \alpha_{T2}(\lambda_2(1-u))}$$

which is consistent with [11, equation (20)].

3. The motion of the system on its state space

To describe the feedback system, the following notation will be employed.

Let $N(t)$ be the number of low priority customers in the system.

Let $X(t)$ be the virtual time to availability of service for C_2 customers who are not in service at time t, i.e. let

$$X(t) = B_1(t) + R_2(t),$$

where $B_1(t)$ is the C_1 (high priority) backlog at time t and $R_2(t)$ is the residual service time of any C_2 customer who has been interrupted and is awaiting resumption of service.

The reader will note that $[N(t), X(t)]$ is *not* bivarate markov for the following reason. When $X(t)$ becomes zero at a C_2 service completion epoch, $N(t)$ decreases by one. On the other hand, when $X(t)$ becomes zero at a C_1 service completion epoch (when no C_2 customers having been served in the current busy period), $N(t)$ does not change. A third process $J(t)$ taking two values S and V, respectively, distinguishes between such system histories. Specifically:

a) $J(t) = V$ if the server is busy and, in the current busy period, no C_2 service has been provided;

b) $J(t) = S$ if the server is busy and some C_2 service has been provided during the current busy period.

Finally, let E be the idle state.

With this notation it is clear that $[N(t), X(t), J(t)]$ is a multivariate Markov process on the state space $\boldsymbol{N} = \boldsymbol{V} + \boldsymbol{S} + \boldsymbol{E}$ with

$\boldsymbol{S} = \{(n,x,S)\ ;\ 1 \le n < \infty, 0 < x < \infty\}$; $\boldsymbol{V} = \{(n,x,V); 0 \le n < \infty, 0 < x < \infty\}$; $\boldsymbol{E} = \{E\}$.

Let T_i be the service time of a type i customer having density $a_{Ti}(x)$, $i = 1,2$ Jumps on the state space N induced by arrivals, busy period terminations and low priority service completions are enumerated in TABLE 3.1.

Epoch	Changes from *V or E*	Changes from *S*
C_1 arrival	$(n,x,V) \dashrightarrow (n, x+T_1, V)$; $E \dashrightarrow (1, T_1, V)$	$(n,x,S) \dashrightarrow (n, x+T_1, S)$;
C_2 arrival	$(n,x,V) \dashrightarrow (n+1, x, V)$; $E \dashrightarrow (1, T_2, S)$	$(n,x,S) \dashrightarrow (n+1, x, S)$;
termination of the $J(t) = V$ mode	$(n,0,V) \dashrightarrow (n,T_2,S)$	
C_2 service completion with no feedback		$(n,0,S) \dashrightarrow (n-1,T_2,S)$; $n > 1$ $(1,0,S) \dashrightarrow \boldsymbol{E}$
C_2 service completion with feedback		$(n,0,S) \dashrightarrow (n,T_2,S)$

TABLE 3.1

When underlying service time distributions are absolutely continuous the distribution on the state space N can be described in terms of densities on $\boldsymbol{S}$ and $\boldsymbol{V}$. Let

$$f_{Vn}(x,t) = \frac{d}{dx}\text{Prob}[\, N(t) = n, X(t) \le x \,, J(t) = V] \,;$$

$$f_{Sn}(x,t) = \frac{d}{dx}\text{Prob}[\, N(t) = n, X(t) \le x \,, J(t) = S] \,;$$

$E(t) = \text{Prob}$ [System is in the idle state $\boldsymbol{E}$ at time t] .

The equations of motion on the state-space may be written down in the customary way (see, e.g., [6]).

One finds that

$$\frac{dE(t)}{dt} = -(\lambda_1+\lambda_2)\, E(t) + f_{V0}(0,t) + (1-\theta) f_{S1}(0,t) \tag{3.1}$$

$$\frac{\partial f_{V0}(x,t)}{\partial t} - \frac{\partial f_{V0}(x,t)}{\partial x}$$
$$= -\lambda_2 f_{V0}(x,t) + \lambda_1 E(t)\, a_{T1}(x) - \lambda_1 f_{V0}(x,t) + \lambda_1 f_{V0}(x,t) * a_{T1}(x), \tag{3.2}$$

$$\frac{\partial f_{Vn}(x,t)}{\partial t} - \frac{\partial f_{Vn}(x,t)}{\partial x}$$
$$= -\lambda_2 f_{Vn}(x,t) + \lambda_2 f_{Vn-1}(x,t) - \lambda_1 f_{Vn}(x,t) + \lambda_1 f_{Vn}(x,t) * a_{T1}(x), \quad n \geq 1, \tag{3.3}$$

$$\frac{\partial f_{Sn}(x,t)}{\partial t} - \frac{\partial f_{Sn}(x,t)}{\partial x}$$
$$= -\lambda_2 f_{Sn}(x,t) + \lambda_2 f_{Sn-1}(x,t) - \lambda_1 f_{Sn}(x,t) + \lambda_1 f_{Sn}(x,t) * a_{T1}(x)$$
$$+ f_{Vn}(0,t)\, a_{T2}(x) + (1-\theta) f_{Sn+1}(0,t) a_{T2}(x) + \theta f_{Sn}(0,t) a_{T2}(x), \quad n > 1 \tag{3.4}$$

$$\frac{\partial f_{S1}(x,t)}{\partial t} - \frac{\partial f_{S1}(x,t)}{\partial x}$$
$$= -\lambda_2 f_{S1}(x,t) + \lambda_2 E(t) a_{T2}(x) - \lambda_1 f_{S1}(x,t) + \lambda_1 f_{S1}(x,t) * a_{T1}(x)$$
$$+ f_{V1}(0,t) a_{T2}(x) + (1-\theta) f_{S2}(0,t) a_{T2}(x) + \theta f_{S1}(0,t) a_{T2}(x) \tag{3.5}$$

and

$$E(0) = 1. \tag{3.6}$$

Equation (3.1) for example states that the rate of change in E(t) is due to losses associated with the C_1 and C_2 arrivals and gains from busy period terminations. Equation (3.4) says that the time rate of change of $f_{Vn}(x,t)$ seen by an observer moving with velocity -1 consists of losses from {(n,x,S)} due to arrivals of C_2 customers, gains from {(n-1,x,S} due to C_2 arrivals, internal losses and gains on{(n,x,S} due to C_1 arrivals, gains from (n,0,V) for termination of busy periods initiated by a C_1 customer arrival, and service completion of C_2 customers. The other equations have a similar probabilistic interpretation.

4. Ergodic analysis

In the next theorem the transform of the joint distribution of $N(\infty)$, the number of low priority customers in the system at ergodicity, and of $X(\infty)$ are obtained from Equations (3.1)-(3.6).

Let

$$\phi_{V\infty}(u,x) = \sum_{n=0}^{\infty} f_{Vn}(x,\infty)\, u^n \,; \quad \Phi_{V\infty}(u,w) = L\,[\sum_{n=0}^{\infty} f_{Vn}(x,\infty)\, u^n]$$

$$\phi_{S\infty}(u,x) = \sum_{n=0}^{\infty} f_{Sn}(x,\infty)\, u^n \,; \quad \Phi_{S\infty}(u,w) = L\,[\sum_{n=0}^{\infty} f_{Sn}(x,\infty)\, u^n]$$

(where $L[\cdot]$ is the Laplace Transform operator) and let $\zeta(u)$ be the solution of the functional equation

$$\zeta(u) = \lambda_2[1-u] + \lambda_1\,[1 - \alpha_{T1}(\zeta(u))\,]\,. \tag{4.1}$$

This functional equation plays a key role in the solution. As shown in Appendix A.2, the functional equation (4.1) has the unique solution given in (2.8).

Theorem 4.1 (The Ergodic Distribution on the State Space)

The ergodic distribution of $[N(t), X(t), J(t)]$ on its state space is described by

$$\Phi_{S\infty}(u,w) = \frac{u\, E(\infty)\, \zeta(u)\, (\, \alpha_{T2}(w) - \alpha_{T2}(\zeta(u))\,)}{[\, w - \lambda_2(1-u) - \lambda_1(1-\alpha_{T1}(w))\,]\,[\, u - \alpha_{T2}(\zeta(u)) + \theta\alpha_{T2}(\zeta(u))(1-u)\,]} \tag{4.2}$$

$$\Phi_{V\infty}(u,w) = \lambda_1 E(\infty) \frac{\alpha_{T1}(\zeta(u)) - \alpha_{T1}(w)}{w - \lambda_2(1-u) - \lambda_1(1 - \alpha_{T1}(w))}\,, \tag{4.3}$$

and

$$E(\infty) = 1 - \rho_S = 1 - \rho_1 - \rho_2\,(1-\theta)^{-1}\,. \tag{4.4}$$

Proof: Equations (4.2)-(4.4) are obtained from the equations of motion, (3.1) - (3.6), by standard methods employing p.g.f.'s, Laplace transforms, regularity conditions and algebra. Details are provided in Appendix A.3. ♦

Remark: By definition $h(u) = \Phi_{S\infty}(u,0) + \Phi_{V\infty}(u,0) + E(\infty)$. Hence, from equations (4.2)-(4.4) and some algebra, equation (2.7) follows. Details are provided in Appendix A.4.

5. The first pass time

We define the first pass time of a C_2 customer to be the time from the arrival of that customer to the system until the completion of that customer's first service. The expected value of the first pass time at ergodicity is a measure of the performance of the system. Its distribution is also an important ingredient for the calculation of the total time in the system.

Theorem 5.1 (Ergodic First Pass Time and the Number in the System): The transform of the joint distribution of V_1^*, the duration of the first pass time, and of N_1^*, the number of C_2 customers in the system at the end of the first pass is given by

$$\begin{aligned}\vartheta_1(u,s) &= E[u^{N_1^*} e^{-sV_1}] \\ &= \alpha_{T2}(w)\,[\,\Phi_{V\infty}(z,w) + E(\infty)] + \Phi_{S\infty}(z,w)\,. \qquad (5.1)\end{aligned}$$

Here

$$z = z(u,s) = (1-\theta+\theta u)\,\alpha_{T2}(w(u,s)) \qquad (5.2)$$

and

$$w = w(u,s) = s + \lambda_2(1-u) + \lambda_1 - \lambda_1\,\sigma_{BP1}(s + \lambda_2\,(1-u)\,). \qquad (5.3)$$

Hence,

$$\vartheta_1(u,s) = \left[\frac{(1-(\rho_1+\rho_2))\,\alpha_{T2}(w)}{1-(\rho_1+\rho_2)\alpha^*_{T12}(w)+\theta(1-u)\alpha_{T2}(w)/w}\right]\left[E(\infty) + \frac{\theta\rho_2}{(1-\theta)}h(z)\frac{(u-z)}{wE[T_2]}\right] / (1-(\rho_1+\rho_2)) \qquad (5.4)$$

where $\alpha^*_{T12}(w)$ and $h(z)$ are defined in (2.4) and (2.6).

Proof: Equation (5.1) follows from a generalization of the argument of Theorem 2.1. Equation (5.4) follows from algebra involving (4.2), (4.3), and (5.1). Details are provided in Appendix A.5.

The Laplace transform of the distribution of the first pass time at ergodicity can be found by simply evaluating (5.4) at u=1. This transform is given in the next corollary:

Corollary 5.2.(Ergodic First Pass Time)
The Laplace transform of the ergodic first pass time distribution is given by

$$\begin{aligned}\Psi_1(s) &= \vartheta_1(1,s)\\ &= \phi_{PK12}(\tilde{w}(s))\,\alpha_{T2}(\tilde{w}(s))\;[p + (1-p)\,h(\alpha_{T2}(\tilde{w}(s)))\,\alpha^{*}_{T2}(\tilde{w}(s))\,]\end{aligned} \tag{5.5}$$

where

$$\tilde{w}(s) = w(1,s) = s+\lambda_1-\lambda_1\sigma_{BP1}(s),$$
$$p = E(\infty)\,/\,(1-\rho_1-\rho_2) = (1-\rho_1-\rho_2(1-\theta)^{-1})\,/\,(1-\rho_1-\rho_2),$$

and $\phi_{PK12}(s)$ is given in (2.5).

The mean is given by

$$\begin{aligned}E[W_1] &= -\Psi'_1(0)\\ &= \frac{(\,(1-\theta)(1-\rho_1)-\theta\rho_2\,)\{\,\lambda_2E[T_2^2]+(1-\theta)\lambda_1E[T_1^2]-2\rho_2E[T_2]\,\}}{2(1-\rho_1)^2\,(1-\theta)\;(\,(1-\theta)(1-\rho_1)-\rho_2\,)} + \frac{E[T_2](1-\theta)}{(1-\theta)(1-\rho_1)-\rho_2}\end{aligned} \tag{5.6}$$

Proof: Equations (5.5) and (5.6) follow from (5.4).♦

Corollary 5.3 (Joint Distribution under no feedback)

If $\theta = 0$

$$\vartheta_1(u,s) = \phi_{PK12}(w(u,s))\;\;\alpha_{T2}(w(u,s)) \tag{5.7}$$

where $w(u,s)$ and $\phi_{PK12}(s)$ are defined in (5.3) and (5.6).

Proof: Equation (5.7) follows from (5.4).♦

Remark: In the case of no feedback , i.e., $\theta = 0$, from (5.7) and (2.6)

$$\Psi_1(s) = \vartheta_1(1,s) = \phi_{T12}(s) \tag{5.8}$$

which is equivalent to equation (3.5) and (3.7) of [9]. If $\theta = 0$ and $\lambda_1 = 0$ this equation simplifies to $\Psi_1(s) = \phi_{PK1}(s)\,\alpha_{T2}(s)$ as expected.

Little's Law, [10], can be used to relate the average number of customers in a system to the average time in the system. The following corollary verifies the distributed form of this law for the case of $\theta = 0$.

Corollary 5.3 (A Distributional form of Little's Law):

If $\theta = 0$, then

$$\Psi_1(\lambda_2 - \lambda_2 u) = \vartheta_1(1,\lambda_2 - \lambda_2 u) = \vartheta_1(u,0) = h(u). \quad (5.9)$$

Hence

$$E[N(\infty)] = \lambda_2 E[T_{sys}] \quad (5.10)$$

and

$$Var[T_{sys}] = Var[\lambda_2 N(\infty)] + E[N(\infty)]. \quad (5.11)$$

Proof: If $\theta = 0$ then, from (2.8), (5.3) and (5.7),

$$\Psi_1(\lambda_2 - \lambda_2 u) = \vartheta_1(1,\lambda_2 - \lambda_2 u) = \vartheta_1(u,0) = \phi_{PK12}(\zeta(u)) \ \alpha_{T2}(\zeta(u)).$$

But if $\theta = 0$ then $\alpha_{T2EFF}(s) = \alpha_{T2}(s)$, $\phi_{PK12}(s) = \phi_{PKS}(s)$ so from (2.7) this equals h(u) so (5.9) follows. If $\theta = 0$ there is no feedback so the time in the system, T_{sys}, is equal to the first pass time, V_1^*. Therefore $\Psi_1'(0) = E[T_{sys}]$, $\Psi_1''(0) = E[T_{sys}^2]$, $h'(1) = E[N(\infty)]$, and $h''(1) = E[N(\infty)^2] - E[N(\infty)]$. Hence (5.10) and (5.11) follow from (5.9).♦

6. Total time in system

Theorem 6.1: Let $\chi(u,s)$ be the transform of the joint distribution of: a) the number of customers remaining in the system at departure and b) the total elapsed time from arrival to the system to departure from the system. Then $\chi(u,s)$ satisfies

$$\chi(u,s) = (1-\theta)\vartheta_1(u,s) + \theta\alpha_{T2}(w(u,s))\ \chi(z(u,s),w(u,s))) \quad (6.1)$$

where z(u,s), w(u,s) and $\vartheta_1(u,s)$ are defined in equations (5.2),(5.3) and (5.4) .

Proof: Using an argument similar to Theorem (5.1) one can show that $\vartheta_{i+1}(u,s) = \alpha_{T2}(w(u,s))\vartheta_i(z(u,s),w(u,s))$ where $\vartheta_i(u,s)$ is the joint transform of the duration of the first i passes and the number of C_2 customers in the system at the end of the ith pass. But $\chi(u,s) = \sum_{i=1}^{\infty} (1-\theta)\theta^{i-1}\vartheta_i(u,s)$ so equation (6.1) follows. Details are provided in Appendix A.6.♦

7. The busy period

The following theorem gives a generalization of the Takács busy period equation.

Theorem 7.1: **(The Busy period)**The busy period has a Laplace transform

$$\sigma_{BP}(s) = \frac{\lambda_1}{\lambda_1 + \lambda_2}\sigma^*_{BP1}(s) + \frac{\lambda_2}{\lambda_1 + \lambda_2}\sigma^*_{BP2}(s) \tag{7.1}$$

where

$$\sigma^*_{BP1}(s) = \alpha_{T1}[\ s + \lambda_1 - \lambda_1\sigma^*_{BP1}(s) + \lambda_2 - \lambda_2\sigma^*_{BP2}(s)\ ], \tag{7.2}$$

$$\sigma^*_{BP2}(s) = \alpha_{T2EFF}[\ s + \lambda_1 - \lambda_1\sigma^*_{BP1}(s) + \lambda_2 - \lambda_2\sigma^*_{BP2}(s)\ ], \tag{7.3}$$

$\alpha_{T1}(s)$ and $\alpha_{T2EFF}(s)$ are defined in TABLE 2.1,

and

$$E[\ T_{BP}] = \frac{\rho_1 + \rho_2(1-\theta)^{-1}}{1 - \rho_1 - \rho_2(1-\theta)^{-1}}\ \frac{1}{\lambda_1+\lambda_2} = \frac{E[T_S]}{1 - \rho_S} \tag{7.4}$$

Proof: See Appendix A.7.

Note that if λ_1 or λ_2 is set equal to zero then equations (7.2) and (7.3) reduce to the familiar Takács busy period equation.

8. Clocked schedules

Consider a system in which the priority tasks arrive with *deterministic* interarrival times $k\Delta$. Suppose further that: a) the service time of the priority tasks, T_1 is infinitely divisible, i.e., for every integer N, T_1 is distributed as the sum of N independent and identically distributed random variables; b) the low priority tasks arrive with an exponentially distributed interarrival time and have a service time with a general distribution; c) the low priority tasks feed back with probability θ.

The distributed Poisson approximation of [7] may be used to model the high priority traffic as a sequence of Poisson streams parametrized by the variable N with rate $\lambda_{1N} = N/\Delta$ and service time having Laplace Transform $\alpha_{1N}(s) = \alpha_{T1}{}^{1/N}(s)$.

As in the classical proof of de Finetti's Theorem [3], one has

$$\alpha_{T1}(s) = \lim_{N\to\infty} [\exp \{ -\lambda_{1N} \Delta [1 - \alpha_{1N}(s) \}], \tag{8.1}$$

i.e, the distribution of the high priority tasks arriving in Δ time units under the approximate Poisson model with arrival rates λ_{1N} and service times $\alpha_{1N}(s)$ is exactly $\alpha_{T1}(s)$. From (8.1) it can be shown that

$$\lim_{N\to\infty} \{ \lambda_{1N} \Delta \, E[T_{1N}] \} = E[T_1], \tag{8.2}$$

$$\lim_{N\to\infty} \{ \lambda_{1N} \Delta \, E[T_{1N}{}^2] \} = Var[T_1], \tag{8.3}$$

and

$$\rho_1 = \lim_{N\to\infty} \{ \lambda_{1N} E[T_{1N}] \} = E[T_1] / \Delta. \tag{8.4}$$

Equation (8.1) suggests that the clocked schedule with feedback could be accurately modeled as the limit, in N, of a system with two classes of Poisson traffic having arrival rates λ_{1N} and λ_2, service times with Laplace transforms $\alpha_{1N}(s)$ and $\alpha_2(s)$ and having a feedback parameter θ. In fact, in [7] this distributed Poisson approximation has been shown analytically as well as numerically to be highly accurate for the case of θ=0. The mean time to completion for the limiting system can now be readily obtained:

Theorem 8.1 (Mean Time to Completion of the Distributed Poisson Approximation to the Clocked Schedule with Feedback)

The mean time to completion is given by

$$E[T_{sys}] = \frac{(1-\theta)Var[T_1]/\Delta + \lambda_2 E[T_2{}^2] + 2E[T_2](1-\rho_1-\rho_2)}{2(1-\rho_1)\{(1-\theta)(1-\rho_1)-\rho_2\}}. \tag{8.5}$$

Proof: Equation (8.5) follows immediately from equations (2.10), (8.2), (8.3) and (8.4).♦

Results

To test the accuracy of (8.5) the following three examples were compared with simulations:

CASE 1: $\lambda_1 = 1$; T_1 has an Erlang-3 distribution with a mean of .2; $\lambda_2 = 1$; T_2 has an exponential distribution with a mean of .15; Four values of θ were examined: θ=0,.25,.5, and .75.

CASE 2: $\lambda_1 = 1$; T_1 has an Erlang-2 distribution with a mean of .2; $\lambda_2 = 5$; T_2 has an Erlang-2 distribution with a mean of .05; Five values of θ were examined: θ=0, .1, .2, .3, .4, and .5.

CASE 3: $\lambda_1 = 1$; T_1 has an exponential distribution with a mean of .2; $\lambda_2 = .1$; T_2 has an Erlang-5 distribution with a mean of .2; Eight values of θ were examined: θ=0, .25, .50, .75, .90, .95, .96, and .97.

In TABLE 8.1, for each case the value of $E[T_{sys}]$ is computed via simulations. The simulated value of $E[T_{sys}]$, $E[T_{sys}]_{sim}$, is given with a 95% confidence interval. In addition, the theoretical value of $E[T_{sys}]$ from (8.5), $E[T_{sys}]_{theory}$, is given along with the relative error,

$$\frac{|E[T_{sys}]_{sim} - E[T_{sys}]_{theory}|}{E[T_{sys}]_{sim}}.$$

CASE	θ	$E[T_{sys}]_{theory}$	$E[T_{sys}]_{sim}$	$\frac{\lvert E[T_{sys}]_{sim} - E[T_{sys}]_{theory}\rvert}{E[T_{sys}]_{sim}}$
1	.00	.244	.244± .090	0.0%
1	.25	.347	.352± .379	1.4%
1	.50	.617	.620± 1.060	0.5%
1	.75	3.042	3.066± 7.251	0.8%
2	.00	.107	.108± .251	0.9%
2	.10	.122	.122± .212	0.0%
2	.20	.144	.154± .095	6.5%
2	.30	.177	.187± .000	5.3%
2	.40	.233	.243± .095	4.2%
2	.50	.349	.357± .402	2.2%
3	.00	.284	.283± .232	0.4%
3	.25	.374	.368± .474	1.6%
3	.50	.554	.551± .943	0.5%
3	.75	1.135	1.128± 2.496	0.6%
3	.90	3.342	3.327± 8.425	0.5%
3	.95	9.962	9.962± 26.320	0.0%
3	.96	16.583	16.232± 43.390	2.2%
3	.97	49.688	50.171±136.89	1.0%

TABLE 8.1

APPENDIX

A. 1: Theorem (Keilson and Servi [8]):

Let an ergodic queueing system be such that for a given class C of customers,

a) arrivals are Poisson of rate λ_2

b) all arriving customers enter the system, and remain in the system until served

c) the customers leave the system one at a time in order of arrival

d) arriving customers do not affect the time in the system of previous customers

then $N(\infty)$ and T_{sys}, the ergodic number in the system and the time in the system, are related according to equation (2.12).

A.2 Proof that equation (2.8) is a solution to equation (4.1):

Let $\Omega(v) = v + \lambda_1(1 - \sigma_{BP1}(v))$ and $V(w) = w - \lambda_1(1 - \alpha_{T1}(w))$ where $\sigma_{BP1}(v)$ is the Laplace transform of a busy period of an M/G/1 queue with arrival rate λ_1 and service time Laplace transform $\alpha_{T1}(w)$. It can be easily shown that $\Omega(v)$ and $V(w)$ are inverses on the appropriate domain, i.e,

$$V(\Omega(v)) = \Omega(v) - \lambda_1\{1 - \alpha_{T1}(\Omega(v))\}$$
$$= v + \lambda_1 \{ 1 - \sigma_{BP1}(v) \} - \lambda_1\{1 - \alpha_{T1}(\Omega(v))\} = v$$

because the Takács equation is equivalent to $\sigma_{BP1}(v) = \alpha_{T1}(\Omega(v))$.

Hence, on the appropriate domain

$$w = \Omega(V(w))$$
$$= V(w) + \lambda_1 (1 - \sigma_{BP1}(V(w)))$$
$$= w - \lambda_1 (1 - \alpha_{T1}(w)) + \lambda_1 (1 - \sigma_{BP1}(w - \lambda_1 (1 - \alpha_{T1}(w))).$$

This implies that

$$\alpha_{T1}(w) = \sigma_{BP1}(w - \lambda_1 (1 - \alpha_{T1}(w))). \tag{A.1}$$

Evaluating (A.1) at $w = \zeta(u)$ gives

$$\alpha_{T1}(\zeta(u)) = \sigma_{BP1}(\zeta(u) - \lambda_1 [1 - \alpha_{T1}(\zeta(u))])$$

so, from (4.1),

$$\alpha_{T1}(\zeta(u)) = \sigma_{BP1}(\lambda_2[1-u]). \tag{A.2}$$

Therefore equation (2.8) follows from (4.1) and (A.2). ♦

A.3 Proof of Theorem 4.1:

From equations (3.3) and (3.4) the Laplace transform with respect to the variable x and the probability generating function with respect to the variable n is found and its limit is evaluated when t approaches ∞. Then

$$\phi_{V\infty}(u, 0) - w\, \Phi_{V\infty}(u,w)$$
$$= -\lambda_2(1-u)\Phi_{V\infty}(u,w) - \lambda_1\Phi_{V\infty}(u,w) + \lambda_1\Phi_{V\infty}(u,w)\alpha_{T1}(w) + \lambda_1 E(\infty)\alpha_{T1}(w),$$

so

$$\Phi_{V\infty}(u,w) = \frac{\phi_{V\infty}(u,0) - \lambda_1 E(\infty)\alpha_{T1}(w)}{w - \lambda_2(1-u) - \lambda_1(1 - \alpha_{T1}(w))} \tag{A.3}$$

From (4.1) the denominator of (A.3) vanishes at $w = \zeta(u)$. Regularity in w in the half-plane {w: Re(w) ≥ 0 } requires that the numerator also vanish there. One then has from equation (A.3)

$$\phi_{V\infty}(u, 0) = \lambda_1 E(\infty)\, \alpha_{T1}(\zeta(u)). \tag{A.4}$$

Equation (4.3) then follows from (A.3) and (A.4).

Also, from (3.1), $\frac{dE(\infty)}{dt} = 0$. Hence

$$(\lambda_1 + \lambda_2)E(\infty) = f_{V0}(0,\infty) + (1-\theta)f_{S1}(0,\infty)\,. \tag{A.5}$$

Equations (3.4) and (3.5) imply that

$$\begin{aligned} - L\{\frac{-\partial\phi_{S\infty}(u,x)}{\partial x}\} &= \phi_{S\infty}(u,0) - w\Phi_{S\infty}(u,w) \\ &= -\lambda_2(1-u)\Phi_{S\infty}(u,w) - \lambda_1(1 - \alpha_{T1}(w))\, \Phi_{S\infty}(u,w) \\ &\quad + [\, \phi_{V\infty}(u,0) - f_{V0}(0,\infty)\,]\, \alpha_{T2}(w) + \lambda_2 E(\infty)\alpha_{T2}(w)u \\ &\quad + (1-\theta)[\, \phi_{S\infty}(u,0)u^{-1} - f_{S1}(0,\infty)\,]\alpha_{T2}(w) + \theta\phi_{S\infty}(u,0)\alpha_{T2}(w). \end{aligned} \tag{A.6}$$

From (A.4), (A.5) and (A.6) one obtains,

$$\Phi_{S\infty}(u,w) = \frac{\phi_{S\infty}(u,0)\,[\, 1-\theta\alpha_{T2}(w)-(1-\theta)u^{-1}\alpha_{T2}(w)\,]}{w - \lambda_2(1-u) - \lambda_1(1-\alpha_{T1}(w))} + \frac{E(\infty)\alpha_{T2}(w)\,[\, \lambda_1(1- \sigma_{BP1}(\lambda_2(1-u))) +\lambda_2(1-u)\,]}{w - \lambda_2(1-u) - \lambda_1(1-\alpha_{T1}(w))}. \tag{A.7}$$

From (4.1) the denominator of (A.7) vanishes at $w = \zeta(u)$. Regularity in w in the half-plane $\{w:Re(w) \geq 0\}$ requires that the numerator also vanish there. One then has from equation (A.7)

$$0 = \phi_{S\infty}(u,0)\,[\,1 - \theta\alpha_{T2}(\zeta(u)) - (1-\theta)u^{-1}\alpha_{T2}(\zeta(u))\,] \\ + E(\infty)\alpha_{T2}(\zeta(u))\,[\,\lambda_1(\,1-\sigma_{BP1}(\lambda_2(1-u))) + \lambda_2(1-u)\,]$$

which, from (2.7), implies

$$\phi_{S\infty}(u,0) = \frac{-\,E(\infty)\alpha_{T2}(\zeta(u))\,\zeta(u)}{1 - \theta\alpha_{T2}(\zeta(u)) - (1-\theta)u^{-1}\alpha_{T2}(\zeta(u))}. \tag{A.8}$$

Equation (4.2) then follows from (2.8), (A.7) and (A.8). Equation (4.4) follows from equations (4.2) and (4.3) and the condition that $\Phi_{S\infty}(1,0) + \Phi_{V\infty}(1,0) + E(\infty) = 1$. ♦

A.4 Derivation of h(u) from equation (4.2)-(4.4):

From (4.2) and (4.3)

$$\Phi_{S\infty}(u,0) = \frac{-E(\infty)\,\zeta(u)\,(1 - \alpha_{T2}(\zeta(u))\,)}{\lambda_2(1-u)[\,1 - \theta\alpha_{T2}(\zeta(u)) - (1-\theta)u^{-1}\alpha_{T2}(\zeta(u))\,]}$$

and

$$\Phi_{V\infty}(u,0) + E(\infty) = E(\infty)\frac{\lambda_1(\,1-\alpha_{T1}(\zeta(u))\,) + \lambda_2(1-u)}{\lambda_2(1-u)}.$$

Then from (4.1)

$$\Phi_{V\infty}(u,0) + E(\infty) = \frac{E(\infty)\zeta(u)}{\lambda_2(1-u)}.$$

But

$$h(u) = \Phi_{S\infty}(u,0) + \Phi_{V\infty}(u,0) + E(\infty)$$

so, after some simplification,

$$h(u) = \frac{-E(\infty)\zeta(u)(1-\theta)\alpha_{T2}(\zeta(u))}{\lambda_2(\,u - \theta u\alpha_{T2}(\zeta(u)) - (1-\theta)\alpha_{T2}(\zeta(u))\,)} \tag{A.9}$$

$$= \frac{-E(\infty)\zeta(u)(1-\theta)\alpha_{T2}(\zeta(u))}{\lambda_2(1-\alpha_{T2}(\zeta(u))) - \lambda_2(1-u)\,(\,1-\theta\alpha_{T2}(\zeta(u))\,)},$$

which, from (4.1),

$$= \frac{-E(\infty)\zeta(u)(1-\theta)\alpha_{T2}(\zeta(u))}{\lambda_2(1-\alpha_{T2}(\zeta(u))) + [\lambda_1(1-\alpha_{T1}(\zeta(u))) - \zeta(u)]\,(1-\theta\alpha_{T2}(\zeta(u)))}$$

which, from (2.3),

$$= \frac{E(\infty)\zeta(u)}{\zeta(u) - (\lambda_1+\lambda_2)(1-\alpha_{TS}(\zeta(u)))} \quad \frac{(1-\theta)\alpha_{T2}(\zeta(u))}{1-\theta\alpha_{T2}(\zeta(u))}$$

$$= \frac{1-\rho_S}{1-\rho_S \dfrac{1-\alpha_{TS}(\zeta(u))}{\zeta(u)E[T_S]}} \quad \frac{(1-\theta)\alpha_{T2}(\zeta(u))}{1-\theta\alpha_{T2}(\zeta(u))}$$

equals (2.7).♦

A.5 Proof of Theorem 5.1:

The joint transform of V^*, the duration of the first pass, and N^*, the number of C_2 customers in the system at the end of the first wait conditioned on the value of (N,X,J) upon arrival, *if no future C_1 or C_2 customers arrive*, is

$$E[u^{N^*}e^{-sV^*} \mid N=n, X=x, J=V] = (1-\theta+\theta u)^n \alpha_{T2}{}^{n+1}(s)\ e^{-sx}. \tag{A.10}$$

The first term reflects the fact the each of the n C_2 customers will feed back with probability θ.

If J = S the right hand side of (A.10) is modified to reflect the fact that a C_2 customer is in service so

$$E[u^{N^*}e^{-sV^*} \mid N=n, X=x, J=S] = (1-\theta+\theta u)^n \alpha_{T2}{}^n(s)\ e^{-sx}, \tag{A.11}$$

and

$$E[u^{N^*}e^{-sV^*} \mid \text{system is idle i.e., } N(\infty)=0, X=0, J=E] = \alpha_{T2}(s). \tag{A.12}$$

Hence,

$$E[u^{N^*}e^{-sV^*}] = \sum_{n=0}^{\infty} \int_{s=0}^{\infty} (1-\theta+\theta u)^n \alpha_{T2}{}^{n+1}(s)\ e^{-sx}\ f_{Vn}(x,\infty)dx$$
$$+ \sum_{n=0}^{\infty} \int_{s=0}^{\infty} (1-\theta+\theta u)^n \alpha_{T2}{}^{n}(s)\ e^{-sx}\ f_{Sn}(x,\infty)dx + \alpha_{T2}(s)E(\infty). \tag{A.13}$$

Therefore,

$$E[u^{N^*}e^{-sV^*}] = \alpha_{T2}(s)\,\Phi_{V\infty}[(1-\theta+\theta u)\alpha_{T2}(s),s] + \Phi_{S\infty}[(1-\theta+\theta u)\alpha_{T2}(s),s] + \alpha_{T2}(s)E(\infty) \quad (A.14)$$

In the presence of future C_1 arrivals, the number of C_2 customers in the system at the end of the first wait, N^*_{P1}, is still N*. However, the C_1 arrivals modify V^* by interruptions of rate λ_1 with each interruption time is equal in distribution to that of T_{BP1}, the C_1 busy period. Hence, in the presence of future C_1 arrivals the joint wait transforms of, V^*_1, the time to completion of the C_2 customers and N^*_{P1} is [4],[5]

$$E[u^{N_{P1}^*}e^{-sV_1^*}] = E[u^{N^*}e^{-(s+\lambda_1-\lambda_1\sigma_{BP1}(s))V^*}] \qquad (A.15)$$

where $\sigma_{BP1}(s)$ is defined in TABLE 2.1.

In the presence of both future C_1 arrivals and future C_2 arrivals the time to completion C_2 customers, V_1, is still V^*_1. However, the number of C_2 customers in the system at the end of the first wait, N_{P1}, is a modification of N^*_{P1} which reflects the Poisson stream with rate λ_2 that arrives during V^*_1. Hence the joint distribution of N_{P1} and V_1 is

$$\vartheta_1(u,s) = E[u^{N_{P1}}e^{-sV_1}] = E[u^{N_{P1}^*}e^{-(s+\lambda_2-\lambda_2 u)V_1^*}]. \qquad (A.16)$$

Combining (A.14), (A.15) and (A.16) gives (5.1)-(5.3).

From (2.8) and (4.3),

$$\alpha_{T2}(w)\,[\Phi_{V\infty}(z,w) + E(\infty)] = \frac{\alpha_{T2}(w)E(\infty)(w-\zeta(z))}{w-\lambda_2(1-z)-\lambda_1(1-\alpha_{T1}(w))}. \qquad (A.17)$$

From (4.2) and (A.17),

$$\Phi_{S\infty}(z,w) + \alpha_{T2}(w)\,[\Phi_{V\infty}(z,w) + E(\infty)]$$

$$= \left[\frac{E(\infty)}{w-\lambda_2(1-z)-\lambda_1(1-\alpha_{T2}(w))}\right]$$

$$[\ \frac{z\,\zeta(z)\,(\alpha_{T2}(w)-\alpha_{T2}(\zeta(z)))}{z-\alpha_{T2}(\zeta(z))+\theta\alpha_{T2}(\zeta(z))(1-z)}+\alpha_{T2}(w)(w-\zeta(z))\ ].$$

From (5.2)

$$=[\ \frac{E(\infty)\alpha_{T2}(w)}{w-\lambda_2(1-(1-\theta+\theta u)\,\alpha_{T2}(w))-\lambda_1(1-\alpha_{T1}(w))}\ ]$$

$$[\ \frac{\zeta(z)\,(z-(1-\theta+\theta u)\alpha_{T2}(\zeta(z)))}{z-\alpha_{T2}(\zeta(z))+\theta\alpha_{T2}(\zeta(z))(1-z)}+w-\zeta(z)\ ],$$

which, from (2.1) and (2.4)

$$=[\ \frac{E(\infty)\alpha_{T2}(w)}{w-(\rho_1+\rho_2)w\alpha^*_{T12}(w)+\theta(1-u)\alpha_{T2}(w)}\ ][\ \frac{\theta\zeta(z)\,(z-u)\,\alpha_{T2}(\zeta(z))}{z-\alpha_{T2}(\zeta(z))+\theta\alpha_{T2}(\zeta(z))(1-z)}+w\ ]$$

which, from (A.9),

$$=[\ \frac{E(\infty)\alpha_{T2}(w)}{w-(\rho_1+\rho_2)w\alpha^*_{T12}(w)+\theta(1-u)\alpha_{T2}(w)}\ ][\ \frac{-h(z)\lambda_2\theta\,(z-u)}{E(\infty)(1-\theta)}+w\ ]$$

which simplifies to (5.4).♦

A.6 Proof of Theorem 6.1: Using the argument of Equation (A.10) one finds that if V^* is the joint distribution of the duration of the first i passes and N^* be the number of C_2 customers in the system at the end of the ith pass *if no other C_1 or C_2 customers arrive* then

$$E[u^{N*}e^{-sV*}\mid N_{P1}=n_1,\ V_1=v_1]\ =(1-\theta+\theta u)^{n_1}\,\alpha_{T2}{}^{n_1+1}(s)\ e^{-sv_1}\,. \tag{A.18}$$

In the presence of future C_1 arrivals, joint transform of the duration of the first i passes, V_i^*, and number of C_2 customers in the system at the end of the ith pass, N^*_{Pi}, is found using the argument of (A.15) to be

$$E[u^{N_{P2}{}^*}e^{-sV_2{}^*}]\ =E[\ u^{N^*}\ e^{-(s+\lambda_1-\lambda_1\sigma_{BP1}(s))V^*}\ ] \tag{A.19}$$

where $\sigma_{BP1}(s)$ is defined in TABLE 2.1.

In the presence of both future C_1 arrivals and future C_2 arrivals, the duration of the first i passes, V_i, number of C_2 customers in the system at the end of the ith pass, N_{Pi}, is found using the argument of (A.16), to be

$$\vartheta_2(u,s) = E[u^{N_{P2}} e^{-sV_2}] = E[u^{N_{P2}^*} e^{-(s+\lambda_2-\lambda_2 u)V_2^*}]. \tag{A.20}$$

Combining (A.18), (A.19) and (A.20) gives

$$\vartheta_2(u,s) = \alpha_{T2}(w(u,s))\vartheta_1(z(u,s),w(u,s))$$

where z(u,s) and w(u,s) are defined in equations (5.2) and (5.3).

Using a similar argument, for i = 1,2,...

$$\vartheta_{i+1}(u,s) = \alpha_{T2}(w(u,s))\vartheta_i(z(u,s),w(u,s)). \tag{A.21}$$

If we uncondition on the number of passes then we find that the joint distribution of the number of C_2 customers in the system and the total elapsed time is

$$\chi(u,s) = \sum_{i=1}^{\infty} (1-\theta)\theta^{i-1}\vartheta_i(u,s) \tag{A.22}$$

which equals

$$= (1-\theta)\vartheta_1(u,s) + \theta \sum_{i=1}^{\infty} (1-\theta)\theta^{i-1}\vartheta_{i+1}(u,s)$$

which, from (A.21),

$$= (1-\theta)\vartheta_1(u,s) + \theta\alpha_{T2}(w(u,s))\sum_{i=1}^{\infty} (1-\theta)\theta^{i-1}\vartheta_i(z(u,s),w(u,s))$$

which, from (A.16) equals (6.1)♦

A.7 Proof of Theorem 7.1:

If $\sigma_{BP1}(s)$ is the Laplace transform of an busy period of C_1 customers assuming no C_2 customer arrivals, i.e.,

$$\sigma_{BP1}(r) = \alpha_{T1}(r + \lambda_1 - \lambda_1\sigma_{BP1}(r)). \tag{A.23}$$

Define $\sigma^*_{BPi}(s)$ to be the Laplace transform of a busy period that begins with a C_i customers in the presence of both C_1 and C_2 arrivals. Then, $\sigma^*_{BP1}(s)$ corresponds to the duration of a T_{BP1} interrupted at rate λ_2 by a interruptions having a duration with Laplace Transform $\sigma^*_{BP2}(s)$. Hence,

$$\sigma^*_{BP1}(s) = \sigma_{BP1}(s + \lambda_2 - \lambda_2\sigma^*_{BP2}(s)). \tag{A.24}$$

The busy period of the C_2 customers is independent of the order of service. Therefore it will be assumed, as was done in Theorem 2.1, that the customers feed back to the front of the line rather than the back of the line and therefore have an effective service time given in equation (2.2).

If this service time is modified due to the interruptions of the Poisson stream of C_1 customers having busy periods with a Laplace transform $\sigma_{BP1}(r)$ then the modified service time has a Laplace transform,

$$\tilde{\alpha}_{T2EFF}(r) = \alpha_{T2EFF}(r + \lambda_1 - \lambda_1\sigma_{BP1}(r)) \tag{A.25}$$

and a busy period starting with a C_2 customer is given by

$$\sigma^*_{BP2}(s) = \tilde{\alpha}_{T2EFF}(s + \lambda_2 - \lambda_2\sigma^*_{BP2}(s)) \,. \tag{A.26}$$

From (A.23) and (A.24), setting $r = s + \lambda_2 - \lambda_2\sigma^*_{BP2}(s)$,
$\sigma^*_{BP1}(s) = \alpha_{T1}(s + \lambda_2 - \lambda_2\sigma^*_{BP2}(s) + \lambda_1 - \lambda_1\sigma_{BP1}(s + \lambda_2 - \lambda_2\sigma^*_{BP2}(s)))$
which from (A.24) implies (7.2).

From (A.25) and (A.26), setting $r = s + \lambda_2 - \lambda_2\sigma^*_{BP2}(s)$

$\sigma^*_{BP2}(s) = \alpha_{T2EFF}(s + \lambda_2 - \lambda_2\sigma^*_{BP2}(s) + \lambda_1 - \lambda_1\sigma_{BP1}(s + \lambda_2 - \lambda_2\sigma^*_{BP2}(s))$
which from (A.26) implies (7.3).

Equation (7.1) follows from the definition of $\sigma_{BP}(s)$, $\sigma^*_{BP1}(s)$, and $\sigma^*_{BP2}(s)$ and (7.4) from differentiations of (7.1), (7.2) and (7.3)♦

ACKNOWLEDGMENT: The authors are grateful to A. Khoo for the programming support which was used to generate TABLE 8.1.

REFERENCES

[1] Disney, R. L. 1984. "Stationary Queue-Length and Waiting-Time Distributions in Single-Server Feedback Queues". *Advances in Applied Probability*. **16**. 437-446.

[2] Disney, R. L. McNickle, D. C. and B. Simon. 1980. "The M/G/1 Queue with instanteous Bernoulli Feedback".*Naval Research Logistic Quarterly*. **27**. 635-644.

[3] Feller, W. "An Introduction to Probability Theory and Its Applications", Vol. 2, John Wiley and Sons, New York, 1971.

[4] Gaver, D. P. 1962. "A Waiting TIme with Interrupted Service, including Priorities", *Journal of the Royal Statistical Societty, Ser. B*. **24**. 698-720.

[5] Keilson, J. 1962. "Queues Subject to Service Interruptions". *The Annuals of Mathematical Statistics*. **33**. 1314-1322.

[6] Keilson, J. and A. Kooharian. 1960. "Time Dependent Queueing Processes". *Annals of Mathematical Statistics*. **31**,104-112.

[7] Keilson, J. and L. D. Servi, "A Distributed Poisson Approximation for Preempt-Resume Clocked Schedules", submitted to *IEEE Trans. on Communication.*

[8] Keilson, J. and L. D. Servi, "A Distributional Form of Little's Law", submitted to *Operations Research Letters.*

[9] Keilson, J. and U. Sumita, "Evaluation of the Total Time in System in a Prempt-Resume Priority Queue via a Modified Lindley Process, *Advances in Applied Probability,* Vol. 15, 1984, pp. 840-859.

[10] Little, J., "A Proof of the Theorem L = λW", *Operations Research, 9,* 383-387 (1961).

[11] Takács, L. "A Single-Server Queue with Feedback", *The Bell System Technical Journal,* March, 1963, pp. 505-519.