A FEEDBACK QUEUE WITH OVERLOAD CONTROL

Bengt Wallström

Lund Institute of Technology, Sweden

ABSTRACT

A simple overload control scheme in a markovian single server system with feedback loops is considered. The server queue is protected by use of an external buffer to which new arrivals are fed when the former contains more than a predetermined number of customers.

Control performance is found to depend on several parameters, in particular
i) server load
ii) ratio of average feedback delay to average service time
iii) average number of delayed feedbacks per customer.

1. INTRODUCTION

The need for service protection against overload in telecommunication and computer systems is a question of increasing importance. This is so basically because modern high speed computers and communication links can be loaded near their maximum capacity before noticeable service degradation takes place.

Many overload control methods used in these systems, however, are rather intricate to analyse and there are reasons to look for simpler alternatives that can be more easily understood.

Let us consider an M/M/1 queue with two feedback loops, which may be entered by a customer each time he leaves the server (Fig 1).

Fig 1 The unprotected feedback queue

One loop, which is entered with probability \( \alpha \), feeds the customer without delay to the end of the queue. The other, being entered with probability \( \beta \), causes a delay of the customer before he joins the queue again. This delay we assume to be an exponentially distributed stochastic variable.

For service protection during heavy load periods we may use the simple arrangement shown in Fig 2. Here we have introduced an external queueing possibility and imposed certain restrictions concerning how to move customers from the external queue into the server queue.

Firstly we allow such transitions only at epochs where there are less than \( M+1 \) customers present in the server system (\( M = 0,1,... \)). Thus, with \( M = 0 \), customers may be moved from the external queue only when the server is idle. As \( M \to \infty \) we attain the system without service protection shown in Fig 1.

Secondly we assume that customers must be moved into the server queue in connection with certain events only, viz. at the epochs of new arrivals and at those epochs of service terminations after which the served customer either leaves the system or enters the delayed feedback loop. A reason for making these last assumptions is the gain of mathematical
simplicity, but in addition they can be recognized to model a case of interrupt mode operation.

This scheme was devised to serve as a simple model of the computer control of the Ericsson designed telephone exchange system AXE. The single server in this case is to model a central processor (CP) and the delayed feedback loop its interworking with a group of regional processors (RP). For simplicity, the RPs were modelled as an infinite server system with exponential service time distribution. The customers represent incoming calls to the exchange, each initiating a sequence of jobs in the CP and RPs. The external queue corresponds to a call buffer where new calls must wait for acceptance by the CP.

The model described above, however, is also of more general interest and we shall here attempt to find out how its performance depends on system parameters. These are

\[ \alpha, \beta \text{ and } M \text{ as defined above} \]

\[ \lambda, \text{ the arrival rate of the Poisson arrival process} \]

\[ 1/\mu, \text{ the mean of the exponential service time distribution} \]

\[ 1/\nu, \text{ the mean of the exponential feedback delay distribution.} \]

We define the state of the system at an arbitrary instant as the values taken by the three variables

\[ I, \text{ the number of customers waiting in the external queue} \]

\[ K, \text{ the number of customers in the server system, (waiting or being served)} \]

\[ L, \text{ the number of customers in the delayed feedback loop.} \]

It is easy to write down the equations determining the steady state probabilities

\[ P(i,k,l) = \Pr(I=i, K=k, L=l) \]

but these seem difficult to solve. Therefore we must confine the analysis to certain special cases. However, an important observation can be made immediately, viz. that the state probabilities

\[ P(m,l) = \Pr(I+K=m, L=l) \]

will obey the equations of the system without service protection, Fig 1. The solution of that system is known from queue net theory to be

\[ P(m,l) = \rho^m (1-\rho) \cdot \frac{1}{1!} \cdot e^{-\gamma} \]

with

\[ \rho = \frac{\lambda}{\mu(1-\alpha-\beta)} \]

\[ \gamma = \frac{\lambda \beta}{\nu(1-\alpha-\beta)} \]

In fact this result is valid more generally, e.g. for deterministic looping, provided only that \( \alpha \) and \( \beta \) are chosen so as to produce on average the correct numbers of loops. Denoting the latter by \( n_\alpha \) and \( n_\beta \) respectively, we have

\[ n_\alpha = \frac{\alpha}{1-\alpha-\beta} ; \quad n_\beta = \frac{\beta}{1-\alpha-\beta} ; \]

The expected number of times a customer visits the single server is

\[ n = n_\alpha + n_\beta + 1 = \frac{1}{1-\alpha-\beta} \]

Thus we conclude that for \( \rho < 1 \)

i) the total number of customers in the external queue and server system has the simple M/M/1-type distribution with mean \( \rho/(1-\rho) \) which is valid also for the number of customers in the unprotected server system of Fig 1.

ii) the number of customers in the delay loop has the Poisson distribution with mean \( \gamma \).

iii) the average time spent in the system by a customer is independent of \( M \) and, by Little's result, equal to

\[ T = \frac{1}{\lambda} \left( \frac{\rho}{1-\rho} + \gamma \right) . \]

In the derivation of state equations for special cases below it is useful to note the following

i) Our assumptions regarding acceptance of new customers imply that no more than one customer at a time can be moved from the external queue.

ii) The system can not move into a state where \( K \leq M \) and \( I \neq 0 \).

iii) A service completion followed by "\( \alpha \)-feedback" does never change the state of the system.

2. THE CASE \( \mu/\nu \to \infty \)

This case is approached when the average feedback delay is much greater than the average service time. As we shall see below, it is likely to represent the worst case of the server queue performance, but is still not too far away from reality. Clearly the number of customers in the delay loop will tend to infinity.
Also these customers will generate a Poisson-type feedback flow of rate $\lambda \beta/(1-\alpha-\beta)$ which is independent of the queues $I$ and $K$.

The state probabilities
\[ P(i,k) = \Pr(I=i, K=k) \]
must obey the following equations
\[ \frac{\lambda(1-\alpha)}{1-\alpha-\beta} \cdot P(0,0) = \mu(1-\alpha) \cdot P(0,1) \]
\[ \left[ \frac{\lambda(1-\alpha)}{1-\alpha-\beta} + \mu(1-\alpha) \right] P(0,k) = \frac{\lambda(1-\alpha)}{1-\alpha-\beta} \cdot P(0,k-1) + \mu(1-\alpha)P(0,k+1) \quad (k = 1, 2, \ldots, M) \]
\[ \left[ \frac{\lambda(1-\alpha)}{1-\alpha-\beta} + \mu(1-\alpha) \right] P(i,M+1) = \frac{\lambda(1-\alpha)}{1-\alpha-\beta} \cdot P(0,M) + \mu(1-\alpha)P(1,M+1) \]
\[ \quad + \mu(1-\alpha)P(0,M+2) \]
\[ \left[ \frac{\lambda(1-\alpha)}{1-\alpha-\beta} + \mu(1-\alpha) \right] P(i+1,M+1) = \frac{\lambda(1-\alpha)}{1-\alpha-\beta} \cdot P(i,M+1) + \mu(1-\alpha)P(i+1,M+1) \]
\[ \quad + \mu(1-\alpha)P(i,M+2) \quad (i = 1, 2, \ldots) \]
\[ \left[ \frac{\lambda(1-\alpha)}{1-\alpha-\beta} + \mu(1-\alpha) \right] P(i-1,M+1) = \frac{\lambda(1-\alpha)}{1-\alpha-\beta} \cdot P(i-1,M+1) \]
\[ \quad + \mu(1-\alpha)P(i-1,M+1) \]
\[ \quad + \mu(1-\alpha)P(i,k) \]
\[ \quad + \mu(1-\alpha)P(i,k+1) \quad (k = M+2, M+3, \ldots; i = 0, 1, \ldots) \]

Using the notations
\[ \rho = \frac{\lambda}{\mu(1-\alpha-\beta)} \]
\[ \tau = \frac{\rho \beta}{1-\alpha} \]
these equations can be rewritten as
\[ \rho \cdot P(0,0) = P(0,1) \quad (1) \]
\[ (\rho+1)P(0,k) = \rho P(0,k-1)+P(0,k+1) \quad (2) \]
\[ (\rho+1)P(0,M+1) = \rho P(0,M)+P(1,M+1)+P(0,M+2) \quad (3) \]
\[ (\rho+1)P(i,M+1) = (\rho-\tau)P(i-1,M+1)+P(i+1,M+1)+P(i,M+2) \quad (4) \]
\[ (\rho+1)P(i,k) = (\rho-\tau)P(i-1,k)+\tau P(i,k-1)+P(i,k+1) \quad (5) \]

From these equations we derive transforms of $P(i,k)$ and of the marginal distributions
\[ Q(i) = \sum_{k=0}^{\infty} P(i,k) \]
\[ R(k) = \sum_{i=0}^{\infty} P(i,k) \]
From (1) and (2) we obtain
\[ P(0,k) = \rho^k P(0,0); \quad k = 1, 2, \ldots, M+1 \]
and, since $I = 0$ whenever $k \leq M$,
\[ R(k) = \rho^k P(0,0); \quad k = 1, 2, \ldots, M \]
In equations (3), (4) and (5) we use the generating function
\[ \tilde{P}(y,z) = \sum_{i=0}^{\infty} \sum_{r=0}^{\infty} P(i,M+1+r)y^i z^r \]
and obtain
\[ \tilde{P}(y,z) = \frac{z(\frac{1}{y} - 1) P^{M+1}(0,0)-(\frac{z}{y} - 1) \tilde{P}(y,0)}{1-(\rho+1)z+((\rho-\tau)yz+\tau z^2} \quad (6) \]
In order to determine $P(0,0)$ first put $y = z$ to obtain
\[ \tilde{P}(z,z) = \frac{\rho^{M+1} P(0,0)}{1-\rho z} \]
Then calculate
\[ E(z^{I+K}) = \tilde{P}(z) = \sum_{k=0}^{M} P(0,k)z^k \tilde{P}(z,z) \]
Since $\tilde{P}(1) = 1$ we obtain
P(0,0) = 1 - \rho

and

\tilde{P}(z) = \frac{1 - \rho}{1 - \rho z}

in accordance with the more general result of the previous section.

Now put y = 1 in (6), which yields

\tilde{P}(1,z) = \frac{\tilde{P}(1,0)}{1 - \tau z}

Since

\tilde{P}(1,1) = 1 - \sum_{k=0}^{M} P(0,k) = \rho^{M+1}

we have \tilde{P}(1,0) = (1 - \rho) \rho^{M+1}

and

\tilde{P}(1,z) = \rho \rho^{M+1} \frac{1 - \tau}{1 - \tau z} (7)

Thus the marginal distribution of K can be written

R(k) = \rho^k (1 - \rho) \quad k = 0, 1, \ldots, M 

and

R(M+1+r) = \rho^{M+1} \tau^r (1 - \tau) \quad r = 0, 1, \ldots (8)

The two lower moments of R(k) are

E[K] = \frac{\rho}{1 - \rho} (1 - \rho) \rho^{M+1} + \frac{\tau}{1 - \tau} \rho^{M+1} (9)

E[K(K-1)] = \frac{2\rho^2}{(1 - \rho)^2} [1 - (M+1) \rho + M \rho^{M+1}] + \frac{2\tau (1 + M - M \tau)}{(1 - \tau)^2} \rho^{M+1} (10)

To obtain the marginal distribution of I, the external queue, we calculate from (6)

\tilde{P}(y,1) = \frac{\tilde{P}(y,0) - \rho \rho^{M+1} (1 - \rho)}{Y (\rho - \tau)}

The unknown function \tilde{P}(y,0) can be determined by using the condition that \tilde{P}(y,z) must be analytic in the region \(|y| < 1; |z| < 1\). Thus the numerator must vanish for any function z(y), in that region, that makes the denominator vanish. It follows that

\tilde{P}(y,0) = \rho \rho^{M+1} (1 - \rho) \frac{z_0 (1 - y)}{z_0 - Y} (11)

where \(z_0\) is the smallest root of the equation

\[ z^2 - \frac{\rho + 1 - (\rho - \tau) y}{\tau} \cdot z + \frac{1}{\tau} = 0 \] (12)

i.e.

\[ z_0 = \frac{1}{2\tau} [\rho + 1 - (\rho - \tau) y - \sqrt{(\rho + 1 - (\rho - \tau) y)^2 - 4}] \] (13)

Clearly there is no simple expression for the corresponding distribution, but moments can be calculated as.

\[ \text{E}(I) = [P'(y,1)]_{y=1} \]

\[ \text{E}(I(I-1)) = [P''(y,1)]_{y=1}, \text{etc.} \]

The required derivatives of \(z_0\) will be obtained most easily from (12). The first two moments are

\[ \text{E}(I) = \frac{\rho \rho^{M+1} (\rho - \tau)}{(1 - \rho)(1 - \tau)} \] (14)

\[ \text{E}(I(I-1)) = \frac{2\rho^2 \rho^{M+1} (1 - \rho)^2 (1 - \tau)}{(1 - \rho)^2 (1 - \tau)^3} \] (15)

3. THE CASE \(\mu/\nu \to 0\)

Here the average feedback delay is assumed negligible in comparison with the average service time. This case is of interest mainly because it seems to put a lower limit to the buffer space requirement of the server queue. Clearly the delay loop is almost always empty and just as in the previous case we need only to consider the state variables I and K. Note, however, that the two feedback loops are not quite equivalent, for an entrance to the delay loop can still cause acceptance of a new customer from the external queue. Also we find that K can not exceed M + 2, since there are no customers in the delay loop.

The state equations will read

\[ \lambda P(0,0) = \mu (1 - \alpha - \beta) P(0,1) \]

\[ [\lambda + \mu (1 - \alpha - \beta)] P(0,k) = \lambda P(0,k-1) \]

+ \mu (1 - \alpha - \beta) P(0,k+1)

\(k = 1, 2, \ldots, M\)
\[
[\lambda + \mu (1-\alpha-\delta)] P(0,M+1) = \lambda P(0,M) + \\
+ \mu (1-\alpha-\delta) P(1,M+1) + \\
+ \mu (1-\alpha-\delta) P(0,M+2)
\]

\[
[\lambda + \mu (1-\alpha)] P(1,M+1) = \lambda P(1-1,M+1) + \\
+ \mu (1-\alpha) P(i+1,M+1) + \\
+ \mu (1-\alpha) P(i,M+2)
\]

\((i = 1,2,\ldots)\)

\[
[\lambda + \mu (1-\alpha-\delta)] P(i,M+2) = \lambda P(i-1,M+2) + \\
+ \mu \delta P(i+1,M+1)
\]

\((i = 0,1,\ldots)\)

or, equivalently,

\[
\rho P(0,0) = P(0,1)
\]

\((\rho+1)P(0,k) = \rho P(0,k-1)+\rho P(0,k+1)
\]

\((\rho+1)P(0,M+1) = \rho P(0,M)+\rho P(1,M+1)+\rho P(0,M+2)
\]

\((\rho+1+\delta)P(i,M+1) = \rho P(i-1,M+1)+\rho P(i+1,M+1) + \\
+ \rho P(i,M+2)
\]

\((\rho+1)P(i,M+2) = \rho P(i-1,M+2)+\rho P(i+1,M+1) + \\
+ \rho P(i,M+2)
\]

where

\[
\rho = \frac{\lambda}{\mu (1-\alpha-\delta)}
\]

\[
\theta = \frac{\mu}{1-\alpha-\delta}
\]

Just as in the previous case we obtain from (16) and (17)

\[
P(0,k) = \rho^k P(0,0); \quad k = 1,2,\ldots,M+1
\]

and

\[
R(k) = \rho^k P(0,0); \quad k = 1,2,\ldots,M
\]

Using the generating function

\[
\tilde{Q}_k(y) = \sum_{i=0}^\infty P(i,k)y^i; \quad k = M+1,M+2
\]

in (18) – (20) yields the equations

\[
[(\rho+1+\delta-\rho y)-1]\tilde{Q}_{M+1}(y)-\tilde{Q}_{M+2}(y) = \\
- (\rho y-1)\rho P(0,M)
\]

From the relations

\[
\tilde{Q}_{M+1}(y) = \frac{1+\delta P(1-y)-\rho y}{[1+\delta P(1-y)](1-\rho y)}\rho P(0,M)
\]

\[
\tilde{Q}_{M+2}(y) = \frac{\rho y}{[1+\delta P(1-y)](1-\rho y)}\rho P(0,M)
\]

we can express the marginal distribution of \(K\) as

\[
R(K) = \rho^k (1-\rho) k = 0,1,\ldots,M
\]

\[
R(M+1) = (1-\rho) \rho^{M+1}
\]

\[
R(M+2) = \frac{\rho}{1-\rho} \rho^{M+2}
\]

The first two moments of this distribution are

\[
E(K) = \frac{\rho}{1-\rho} (1-\rho^{M+1}) + \rho^{M+1}
\]

\[
E(K(K-1)) = \frac{2\rho^2 [1-(1-M)\rho+\rho^M] M+1}{(1-\rho)^2}
\]

The marginal distribution of \(I, Q(i)\), has the generating function

\[
\tilde{Q}(y) = \sum_{k=0}^M P(0,k) \tilde{Q}_{M+1}(y) + \tilde{Q}_{M+2}(y)
\]

\[
\tilde{Q}(y) = \frac{1-\rho^M}{[1+\delta P(1-y)](1-\rho y)}\rho P(0,M)
\]
from which it is possible to obtain a series expansion and a recurrence relation for \( Q(i) \). Here we give the first two moments only

\[
E[I] = \frac{1 + \beta}{(1 + \beta)(1 - \rho)} \rho^{M+2}
\]

\[
E[I(I-1)] = \frac{2[1 + \beta(3 - \rho + \theta)]}{(1 + \theta)^2(1 - \rho)^2} \rho^{M+3}
\]

4. THE CASE \( \rho \geq 1 \)

In the case of persistent overload (\( \rho \geq 1 \)) the external queue tends to infinity, but the server system remains stable due to the control mechanism. Since there are always customers available in the external queue, a change of \( K \) from \( M+1 \) to \( M \) will be followed immediately by a change from \( M \) to \( M+1 \) again. Thus \( P[K \leq M] = 0 \), and it is convenient to define state probabilities as

\[
P(r,1) = P[K-M-1 = r, L = 1]
\]

\((r,1) = 0,1,...\)

The marginal distribution of \( L \) is a Poisson distribution with mean \( \mu \beta / \nu \), as the server is always busy and generates a Poisson arrival process of rate \( \mu \beta \) to the delay loop. For the joint distribution, \( P(r,l) \), the following state equations are obtained

\[
(u \beta + 1) \nu P(0,1) = u \beta [P(1,1-1) + P(0,1-1)] + u(1 - \alpha - \beta) P(1,1)
\]

\((l = 0,1,...)\)

\[
[u(1-\alpha)+1] \nu P(r,1) = (1 + \nu) P(r-1,1+1) + u \beta P(r+1,1)+1) + u(1-\alpha-\beta) P(r+1,1+)
\]

\((r = 1,2,...; l = 0,1,...)\)

We have so far not been able to solve these equations, not even to calculate moments of the server queue distribution. The two special cases when \( \mu / \nu \to \infty \) and when \( \mu / \nu \to 0 \), however, are simple enough. In fact the distribution of \( K \) will be obtained by putting \( \rho = 1 \) in the formula for \( R(k) \) derived in previous sections.

4.1 The case \( \rho \geq 1; \mu / \nu \to \infty \)

For \( \alpha + \beta < 1 \), \( \nu \in \nu \infty \), we obtain from (8), putting \( \rho = 1 \),

\[
R(k) = 0 ; k = 0,1,...,M
\]

\[
E[K-M-1] = \frac{\beta}{1-\alpha-\beta}
\]

\[
E[(K-M-1)(K-M-2)] = 2 \left( \frac{\beta}{1-\alpha-\beta} \right)^2
\]

4.2 The case \( \rho > 1; \mu / \nu \to 0 \)

Putting \( \rho = 1 \) in (27), we obtain

\[
R(k) = 0 ; k = 0,1,...,M
\]

\[
R(M+1) = 1 - \frac{\beta}{1-\alpha}
\]

\[
R(M+2) = \frac{\beta}{1-\alpha}
\]

and

\[
P[K-M-1] = \frac{\beta}{1-\alpha}
\]

\[
E[(K-M-1)(K-M-2)] = 0
\]

5. NUMERICAL STUDY

Below the system performance is illustrated by diagrams showing means \( m \) and standard deviations \( \sigma \) of the server system and external queues. These are calculated from formulae given in sections 2-4 as

\[
m = E[X]
\]

\[
\sigma^2 = E[X(X-1)] + m - m^2
\]

where \( X \) is the state variable under consideration.

Figures 3-6 show \( m \) and \( \sigma \) as functions of \( M \) for the server system and for the external queue in the two limiting cases studied in sections 2 and 3 above. The server load is \( \rho = 0.9 \).

Figures 3, 4 illustrate cases with an average \( n = 10 \) server visits per customer. The similar curves for \( n = 50 \) are shown in figures 5, 6. We have also varied \( \beta / (a + \beta) \) the proportion of delayed feedbacks.

Note that the case \( \beta = 0 \) could be used to model an SPC system without regional processors. In this case our limiting solutions for \( \mu / \nu \to 0, \infty \) coincide and yield the lowest possible values of the first two moments of the server queue. For \( \beta / (a + \beta) > 0, \alpha > 0 \), the model with infinite feedback delays \( (\mu / \nu \to \infty) \) produces strictly higher values of the first two moments than does the model with zero
feedback delays. It is natural, therefore, to conjecture that these cases will give lower and upper limits of the server queue moments as the ratio $\mu/\nu$ is varied from zero to infinity. As can be seen from figures 7-10 this conjecture was not contradicted by simulation results. The similar statement - with "upper" and "lower" reversed - seems to hold for the external queue.

Moreover it is seen that for $\mu/\nu = 0$ the moments are rather insensitive to variation of the ratio $\beta/(\alpha+\beta)$, while for $\mu/\nu = \infty$ they may vary a great deal, in particular when the average number of server visits per customer is high.

The mean and standard deviation curves of the server system indicate, as would be expected, that for maximum regulation effect $M = 0$ should be chosen.

The choice of $M$, however, does not seem very critical and may in practice be influenced by other factors. More important is the fact that service protection can be obtained even under conditions of persistent overload. It was found in section 4 that for $\rho > 1$ and $\alpha + \beta < 1$ the variable $K-M-1$ had a stationary distribution with finite moments in the two limiting cases $\mu/\nu \to 0, \infty$. In figures 9 and 10 we show examples of $m$ and $\sigma$ for these limiting distributions and, in addition, some intermediate approximate curves estimated by simulation. For good service protection the number of queueing places must be large enough to ensure that customers, once they have been accepted by the server, will be almost sure to get full service. Clearly the queue buffer requirement increases as the parameters $\mu/\nu, \nu$ and $\beta/(\alpha+\beta)$ are increased.

Before using our results for dimensioning purposes, however, the underlying assumptions should be carefully examined. For instance our delay loop model, which permits an unlimited number of customers to be present, is likely to overestimate the need for buffer memory in some applications.

6. CONCLUSION

The simple control method considered in this paper has the desirable property of keeping the server queue stable during overload periods. Still its efficiency can be said to vary in the sense that certain system parameters influence the response time and the buffer capacity needed to ensure proper service to customers accepted by the server.

The best efficiency in that meaning is obtained when $\beta = 0$, because for $\rho \geq 1$ we have then always $M+1$ customers in the server system and no extra queueing places are required. The apparently worst conditions occur when $\mu/\nu$ is large and $\beta$ is near unity. Intuitively this is so because in such cases there are many customers in the delay loop and these are not sensed by our control mechanism. Therefore it might seem natural to attempt to include the variable $L$ in the decision rule for acceptance of customers from the external queue. But if the server is to work at its maximum capacity it must accept a new customer immediately as it becomes idle, independently of the value of $L$. Thus under overload it seems impossible in this way to obtain better server queue performance than for $M = 0$ in our simple scheme, unless we allow the server occupancy to fall below unity.
Fig 3 Limits of state distribution averages as $\mu/\nu \to 0, \infty$ for $\rho = 0.9$; $\alpha + \beta = 0.9$ (n = 10)

Fig 4 Limits of state distribution standard deviations as $\mu/\nu \to 0, \infty$ for $\rho = 0.9$; $\alpha + \beta = 0.9$ (n = 10)

Fig 5 Limits of state distribution averages as $\mu/\nu \to 0, \infty$ for $\rho = 0.9$; $\alpha + \beta = 0.98$ (n = 50)

Fig 6 Limits of state distribution standard deviations as $\mu/\nu \to 0, \infty$ for $\rho = 0.9$; $\alpha + \beta = 0.98$ (n = 50)
Fig 7 Server system. State distribution averages. Simulation results (○) and limits as \( \frac{\mu}{\nu} \to 0, \infty \). \( \rho = 0.9; \alpha = \beta = 0.45 \)

Fig 8 Server system. State distribution standard deviations. Simulation results (○) and limits as \( \frac{\mu}{\nu} \to 0, \infty \). \( \rho = 0.9; \alpha = \beta = 0.45 \)

Fig 9 Server system. State distribution averages under persistent overload. Simulation results (○) and limits as \( \frac{\mu}{\nu} \to 0, \infty \). \( \alpha + \beta = 0.9 \).

Fig 10 Server system. State distribution standard deviations under persistent overload. Simulation results (○) and limits as \( \frac{\mu}{\nu} \to 0, \infty \). \( \alpha + \beta = 0.9 \).
Summary of Questions/Answers

Date: 10 June 1983
Session: 1.3
Paper: 4

Q.1 (Werner Bux)

You mention in the introduction that your study was motivated by an overload control scheme employed in a real telephone exchange. Have you made an attempt to validate your model through comparison with the behavior of the real system?

A.1 (Bengt Wallstrom)

The simple feedback model without overload control has been tested against a rather detailed simulation model of the AXE control system and found to be rather useful. The model with overload control has not so far been validated. It is in fact much simpler than the real AXE scheme.