melecon'87
MEDITERRANEAN ELECTROTECHNICAL CONFERENCE
34th CONGRESS ON ELECTRONICS
JOINT CONFERENCE

FINAL PROGRAM

ROME, ITALY - PALAZZO DEI CONGRESSI
March 24, 25, 26, 1987
BANDWIDTH AND BUFFER ALLOCATION FOR MULTIRATE SIGNALS

by
A.S. Drigas, E.M. Protonotarios, G.I. Stassinopoulos and E.D. Sykas

Department of Electrical Engineering
National Technical University of Athens
42 Patission Street
106 82 Athens, Greece

Abstract

Resource allocation aspects are considered in cases of random contention for M identical resources from K statistically different customer types. In particular attention is focused on data, voice and video traffic within the ISDN framework. The features of this application vary significantly form the assumption taken in the literature when examining resource allocation schemes in a similar framework.

The general problem is to determine the optimal policy for accepting or rejecting a call when the type of the requesting customer is known as well as the state vector, with components the numbers of customers of each type that are in service.

The objective of this study is to develop analytic models and computational algorithms for the determination of the optimal state subset for slotted time systems with call traffic modeled as stationary independent arrival processes and with deterministic service time. The parameters optimized are the ones generally accepted, as throughput, utilization and blocking of the system.

Introduction

Problems of buffer and bandwidth allocation among several types of customers are considered in this paper. It is a usual situation in a computer communication network that a limited number of resources are shared among several communities of customers. Some examples of such situations are the following: a) A number of customer types try to get access to a host. There is an upper limit to the number of virtual circuits that can terminate to the host at same instance, b) k types of jobs are looking forward to be served by a limited number of processors in a multiprocessor machine or in a computer network [2], c) several types of customers contend to set up virtual circuits through a limited bandwidth channel [8], d) in some cases a memory of limited size is shared among some packet communities within a computer communication network [1],[4],[5],[6]. Especially in an ISDN framework, communities of data, voice and video packets share a common memory before they get served by any server.

Recently, many articles have appeared in the literature considering such problems [1],[7],[8]. Usually the objective of such articles is to determine the optimal policy for a specific system or to choose the better policy among some of them, and to develop computational methods for doing such a thing. As a policy usually is meant the operation for accepting or rejecting a call when the type of requesting customer is known as well as the system state characterized by the allocation policy. The state is a vector with components the number of customers of each type that are in service.

In this article we consider discrete time systems. A limited number of resources are shared among packet communities. We have developed an algorithm for the determination of the system with best performance characteristics among some simple and usually appearing in the literature policies. We have developed the model for a discrete time system instead of continuous one as usually appears in the literature. The arrival rates are supposed to be geometrically distributed with different bit rates among different types of customers.

The mathematical model

Consider a server with a common waiting area with a total accommodation of N storage places, Figure 1.

![Fig. 1. Server with common waiting room.](image)

This buffer is shared between two types of customers. Each customer from one type is going to be served with a first-come first-served policy. The time required to serve a customer from any type is constant. If we are concentrating in a packet switching network, the constant service time comes from the constant packet length, and last exactly a time slot. So in each time slot only one job can be served. Also in each time slot one customer of type 1 can arrive with probability \( \lambda_1 \) or not arrive with probability \( (1-\lambda_1) \). The same holds for the customers of type 2. One arrives in this slot with probability \( \lambda_2 \) and doesn't arrive with probability \( (1-\lambda_2) \). This is a rational consideration and isn't made for simplification of the problem. That comes from the fact that the inputs to the considered system are outputs of similar systems, that can serve one customer in each time slot. So if there are customers to that other system they are served one each time slot and arrive to our system. So the probability \( \lambda_1 \) under which one customer arrives express the probability for any customer to exist to that other system.

When a job of any type requests service at its arrival may be accepted or rejected according to the policy and the number of each customer type that are already in the waiting room. We denote \( Q_1(t)=n_1 \) and \( Q_2(t)=n_2 \) the number of each customer type that are waiting for service of the beginning of the \( t \) slot. And that is what we call a policy now. A policy is the decision to accept or reject a call when \( Q_1(t) \), \( Q_2(t) \) is known as well as the identity of the requesting customer. Also with no loss of generality we can assume that the arrivals at each time slot take place exactly at the end of the slot. At the beginning of each time slot the server chooses by chance one of the waiting customers and serves him during the slot. We can observe that when the considered system is at a state, any one type customer can be served and leave the system so that we can understand that any policy must be a coordinate convex set of admissible states as originally stated by Aein [3].

Calculation of the state transition probabilities

Considering the state \((n_1,n_2)\) lying in the interior of the state space in \( \mathbb{R}^2 \), we have seven possible transitions into it. These are depicted in Fig. 2 and...
are:

(i) Arrival of two customers, one of each type with service completion of one customer from either type (transitions from \((n_1, n_2-1)\)).

(ii) Arrival of one customer of any type, with service completion of a customer of the same type (transition from \((n_1, n_2)\)).

(iii) Arrival of one customer of either type, with service completion of a customer of the other type (transition from \((n_1-1, n_2+1)\) or \((n_1+1, n_2-1)\)).

(iv) No new arrivals, with service completion of a customer from either type (transition from \((n_1, n_2+1)\) or \((n_1+1, n_2)\)).

For each policy now we have special boundary equations. Since our policy is effective by reflected on the remaining part of the boundary, we place special emphasis on the system description there.

Thus, for the complete sharing policy we obtain:

If \(n_1+n_2 = N\) (this case holds also for limited complete sharing policy)

\[
Pr(n_1, n_2) = (n_1/(n_1+n_2))\lambda_1\lambda_2 Pr(n_1, n_2-1) + (n_2/(n_1+n_2))\lambda_1\lambda_2 Pr(n_1-1, n_2) + \sum_{m=0}^{k} k!/(m! (k-m)!) \text{transitions if } m \text{ customer from any type arrive. So we have a total of:}
\]

\[
Pr(n_1, N) = \lambda_1(1-\lambda_2) Pr(N, 0) + (1/(n_1+1))(1-\lambda_1)(1-\lambda_2)
\]

and for \(n_2 = N\), we take

\[
Pr(0, N) = (1-\lambda_1)\lambda_2 Pr(0, N) + (1/(n_2+1))(1-\lambda_1)(1-\lambda_2)
\]

For complete partitioning and limited complete sharing policies there are 4 common boundary equations.

\[
Pr(N_1, 0) = \lambda_1(1-\lambda_2) Pr(N_1, 0) + (1/N_1)(1-\lambda_1)(1-\lambda_2)
\]

\[
Pr(N_1, 1) = (1-\lambda_1)\lambda_2 Pr(N_1, 0) + (1/(N_1+1))(1-\lambda_1)(1-\lambda_2)
\]

\[
Pr(0, N_2) = (1-\lambda_1)\lambda_2 Pr(0, N_2) + (1/(N_2+1))(1-\lambda_1)(1-\lambda_2)
\]

\[
Pr(1, N_2) = (1-\lambda_1)\lambda_2 Pr(N_1, N_2) + (1/(1+1))(1-\lambda_1)(1-\lambda_2)
\]
When for complete partitioning policy \(n_1 < n_1 N_2\) or for limited complete sharing policy \(n_2 = N_2, n_1 + n_2 < N\), we have:

\[
Pr(n_1, n_2) = \frac{(n_1 + n_2 - 1)}{(n_1 + n_2 - 1)} \lambda_1 \lambda_2 Pr(n_1, n_2 - 1) + \frac{(n_2/ (n_1 + n_2 - 1)) \lambda_1 \lambda_2 Pr(n_1, n_2) + (n_1/(n_1 + n_2)) \lambda_1 (1-\lambda_2) Pr(n_1, n_2) + ((n_2 + 1)/(n_1 + n_2)) (1-\lambda_1) \lambda_2 Pr(n_1, n_2) + ((n_1 + 1)/(n_1 + n_2)) (1-\lambda_1) \lambda_2 Pr(n_1 + 1, n_2) + ((n_1 + 1)/(n_1 + n_2)) (1-\lambda_1) \lambda_2 Pr(n_1 + 1, n_2)}{(1-\lambda_2) Pr(n_1 + 1, n_2)}
\]

For limited complete sharing policy there are two more boundary equations.

For \(n_1 = N_1, n_1 + n_2 < N\), we have

\[
Pr(n_1, n_2) = \frac{(n_1 + n_2 - 1)}{(n_1 + n_2 - 1)} \lambda_1 \lambda_2 Pr(n_1, n_2 - 1) + \frac{(n_2/ (n_1 + n_2 - 1)) \lambda_1 \lambda_2 Pr(n_1, n_2) + (n_1/(n_1 + n_2)) \lambda_1 (1-\lambda_2) Pr(n_1, n_2) + ((n_2 + 1)/(n_1 + n_2)) \lambda_1 (1-\lambda_2) Pr(n_1 - 1, n_2) + ((n_1 + 1)/(n_1 + n_2)) \lambda_1 \lambda_2 Pr(n_1 - 1, n_2 - 1)}{(1-\lambda_2) Pr(n_1, n_2)}
\]

For \(n_2 = N_2, n_1 + n_2 < N\), we take

\[
Pr(n_1, n_2) = \frac{(n_1 + n_2 - 1)}{(n_1 + n_2 - 1)} \lambda_1 \lambda_2 Pr(n_1, n_2 - 1) + \frac{(n_2/ (n_1 + n_2 - 1)) \lambda_1 \lambda_2 Pr(n_1, n_2) + (n_1/(n_1 + n_2)) \lambda_1 (1-\lambda_2) Pr(n_1, n_2) + ((n_2 + 1)/(n_1 + n_2)) \lambda_1 (1-\lambda_2) Pr(n_1 - 1, n_2) + ((n_1 + 1)/(n_1 + n_2)) \lambda_1 \lambda_2 Pr(n_1 - 1, n_2 - 1)}{(1-\lambda_2) Pr(n_1, n_2 - 1)}
\]

The method

Let \(P_n = [P_{n0}, \ldots, P_{nj}, \ldots, P_{nm}]^T\), where \(T\) means transposed and \(m\) depends on the policy. From the construction of the general equation independently of the policy we are following, we can perceive that every \(P_k\) can be expressed as follows:

\[
P_1 = A_1 P_1 + B_0 P_0
\]

\[
P_k = A_k P_k + B_{k-1} P_{k-1} + C_k P_{k-2}
\]

so that

\[
P_1 = (I-A_1)^{-1} B_0 P_0
\]

\[
P_k = (I-A_k)^{-1} (B_{k-1} P_{k-1} + C_k P_{k-2})
\]

then we have

\[
P_1 = D_1 P_0
\]

\[
P_k = D_k P_0
\]

where

\[
D_1 = (I-A_1)^{-1} B_0
\]

\[
D_k = (I-A_k)^{-1} (B_{k-1} D_{k-1} + C_k D_{k-2})
\]

where

\[
A_k, B_k, C_k, D_k\]

are arrays with appropriate dimensions

From the construction of the state diagram Fig. 3 we perceive that writing down the equations for states (0, 0), (0, 1), ..., and independently from the policy we are following, we can express \(P(1,0), P(1,1), \ldots\), in other words \(P_1\) as a function of \(P_0\), and we can find the terms of the \(A_1\) and \(B_0\) arrays. Also from equations for states (1, 0), (1, 1), ..., we can express \(P_2\) as a function of \(P_1\) and \(P_0\), and also we can determine the terms for \(A_2, B_1, C_0\) arrays. So we can determine every \(A_k, B_k, C_k\) array in other words every \(D_k\) and so that every \(P_k\), writing down the transition equations for every system's state except the states that are the horizontal projections of (0, 0), (0, 1), ..., (0, N) states, to the boundary of the policy.

\[\]

\[\]

\[\]

Fig. 3. State diagram for any policy

As a result we can express \(N\) probabilities of the policy's boundary (these we mentioned earlier) as a function of \(P_0\). Writing down the transition equations for these \(N\) probabilities we take an equation system with \(N\) equations and \(N\) unknowns. The \(N^{th}\) equation is linearly dependent from the residual \(N-1\) equations. So that we substitute it with the normalization equation

\[
\sum_{i,j} P_{ij} = 1
\]

\[
(i,j) \in D
\]

Where \(D\) is the set of admissible states according to the policy. So that we have a linear system with \(N\) equations and \(N\) unknowns which can be solved.

An example elucidating the method

Consider a buffer in a packet switching network which can contain a maximum number of 4 packets. Suppose that the assumptions we made in the mathematical model paragraph holds in hear, with \(\lambda_1 = 0,3\) and \(\lambda_2 = 0,2\).

For complete sharing policy we have the transition diagram as in Fig. 4.

\[\]

\[\]

\[\]

Fig. 4. Transition state diagram for complete sharing policy.

According to the method we have

\[

P_0 = [P_{00}, P_{01}, P_{02}, P_{03}, P_{04}]^T
\]

\[
P_k = [P_{00}, \ldots, P_{4-k}]^T
\]

\[
P_1 = D_1 P_0\]

where according to the method:
\[
\begin{align*}
P_k &= D_k P_0 \\
D_1 &= \begin{bmatrix} 0.785 & -1 & 0 & 0 & 0 \\ -0.392 & 3.571 & -2 & 0 & 0 \\ 0.147 & -1.339 & 5.357 & -3 & 0 \\ -0.049 & 0.446 & -1.785 & 7.142 & -3 \end{bmatrix}
\end{align*}
\]

writing the equation \( \sum \sum P_{ij} = 1 \), and transition equation for states \((0,4), (1,3), (4,0), (3,1)\) we have

\[
\begin{align*}
4.456x+0.012w-0.002z &= 1 \\
-0.002x+0.015w-0.062z+0.250w-m &= 0 \\
-0.127x+0.991w-3.141z+7.463w-4m &= 0 \\
3.764x+17.773w+16.368z-6.71w+m &= 0 \\
-4.644z+27.173w-35.767z+20.386w-4m &= 0
\end{align*}
\]

This equation system solved gives:

\[
P_0 = \begin{bmatrix} 0.2242 & 0.0721 & 0.0296 & 0.0070 & 0.0005 \end{bmatrix}^T
\]

and so that

\[
P_1 = \begin{bmatrix} 0.1036 & 0.1110 & 0.0746 & 0.0162 \end{bmatrix}^T
\]

\[
P_2 = \begin{bmatrix} 0.0539 & 0.0990 & 0.1622 \end{bmatrix}^T
\]

\[
P_3 = \begin{bmatrix} 0.0161 & 0.0297 \end{bmatrix}^T
\]

\[
P_4 = \begin{bmatrix} 0.0028 \end{bmatrix}
\]

For limited complete sharing policy we have

\[
\begin{align*}
P_1 &= D_1 P_0 \\
D_1 &= \begin{bmatrix} 0.785 & 1 & 0 & 0 & 0 \\ -0.392 & 3.571 & -2 & 0 & 0 \\ 0.147 & -1.339 & 5.357 & -3 & 0 \\ -0.049 & 0.446 & -1.785 & 7.142 \end{bmatrix}
\end{align*}
\]

\[
\text{Fig. 5. Transition state diagram for limited complete sharing policy.}
\]

\[
\begin{align*}
P_1 &= D_1 P_0 \\
D_1 &= \begin{bmatrix} 1.26 & -3.57 & 1 & 0 \\ -1.596 & 10.427 & -10.714 & 3 \\ 1.186 & -8.482 & 22.066 & -21.428 \end{bmatrix}
\end{align*}
\]

\[
D_3 = \begin{bmatrix} 2.324 & -9.083 & -5.356 & -1 \\ -4.784 & 29.129 & -38.758 & 21.422 \end{bmatrix}
\]

Following some steps as before we find:

\[
-0.119x+21.099w-18.478z+7.136w = 1 \\
-0.127x+0.991w-3.141z+7.463w = 0 \\
2.562x-11.815y+10.353z-3.999w = 0 \\
-4.004x+24.221w-33.180z+19.387y = 0
\]

This equation system solved gives:

\[
P_0 = \begin{bmatrix} 0.2242 & 0.0723 & 0.0296 & 0.0067 \end{bmatrix}^T
\]

so that

\[
P_1 = \begin{bmatrix} 0.1036 & 0.1110 & 0.0746 & 0.0162 \end{bmatrix}^T
\]

\[
P_2 = \begin{bmatrix} 0.0539 & 0.0990 & 0.1622 \end{bmatrix}^T
\]

\[
P_3 = \begin{bmatrix} 0.0161 & 0.0297 \end{bmatrix}^T
\]

For complete partitioning policy we have

\[
P_k = D_k P_0 \text{ with}
\]

\[
\text{Fig. 6. Transition state diagram for complete partitioning policy.}
\]

\[
D_2 = \begin{bmatrix} 1.26 & -3.57 & 1 & 0 \\ -1.596 & 10.427 & -10.714 & 3 \\ 1.186 & -8.482 & 22.066 & -21.428 \end{bmatrix}
\]

\[
D_3 = \begin{bmatrix} 2.324 & -9.083 & -5.356 & -1 \\ -4.784 & 29.129 & -38.758 & 21.422 \end{bmatrix}
\]

Following the developed method we have:

\[
2.212x+1.282y+14.028z = 1 \\
1.303x-5.080y+2.993z = 0 \\
0.874x-6.687y+15.916z = 0
\]

Solving this equation system we take

\[
P_0 = \begin{bmatrix} 0.2796 & 0.0831 & 0.0195 \end{bmatrix}^T
\]

so that

\[
P_1 = \begin{bmatrix} 0.1363 & 0.1481 & 0.0342 \end{bmatrix}^T
\]

\[
P_2 = \begin{bmatrix} 0.0751 & 0.2113 & 0.0110 \end{bmatrix}^T
\]

We are now in the position to calculate the per-
formance criteria for these policies in this example. The idleness for every policy is the probability for the system to be without any customers. So that idleness for complete sharing is \( P_{100} = 0.2242 \), idleness for limited complete sharing is \( P_{200} = 0.2242 \) and idleness for complete partitioning is \( P_{300} = 0.2796 \) so that utilization for every system is \( 1-P_{00} \). Now we will calculate rejection probabilities for each policy. For complete sharing policy we have:

\[
\text{rejection 1} = (P_{04}+P_{13}+P_{22}+P_{31}+P_{40}).\lambda_1 = 0.0624
\]

\[
\text{rejection 2} = (P_{04}+P_{13}+P_{22}+P_{31}+P_{40}).\lambda_2 = 0.0164
\]

for type 1 and 2 packets respectively.

For limited complete sharing policy, holds:

\[
\text{rejection 1} = (P_{30}+P_{13}+P_{22}+P_{31}).\lambda_1 = 0.0672
\]

\[
\text{rejection 2} = (P_{03}+P_{13}+P_{22}+P_{31}).\lambda_2 = 0.0429
\]

And at last for complete partitioning policy we take:

\[
\text{rejection 1} = (P_{22}+P_{30}+P_{20}).\lambda_1 = 0.0892
\]

\[
\text{rejection 2} = (P_{02}+P_{12}+P_{22}).\lambda_2 = 0.0129
\]

**Bandwidth allocation**

Consider a limited bandwidth shared between two classes of customers. Consider that call set up packets and call clear packets from the same customer type arrive under the same probability, \( \lambda \) for the type one customers and \( \lambda \) for type two customers. Consider also that the two classes are distinguished and from the different needs in bandwidth space. So that the transition state diagram becomes as in Fig. 7 (if for example every call from type one customers need 3 times more bandwidth space),

![Bandwidth transition state diagram](image)

Fig. 7. Bandwidth transition state diagram.

We can perceive from the special construction of the state diagram that we can implement the developed method and for this kind of problems. So we can find every \( P_k \) in terms of \( P_0 \) simply determining every \( D_k \). As a conclusion we can tell that we developed a simple computational method for the determination of the performance behaviour for every coordinate convex policy.