#190

ANALYSIS OF SHARED STORAGE IN A COMPUTER NETWORK ENVIRONMENT*

Farouk Kamoun and Leonard Kleinrock
Computer Science Department
University of California, Los Angeles 90024

Abstract

We consider various schemes for sharing a pool of buffers among a set of communication channels in a communication network environment. Four sharing schemes are examined and the results of the analysis are presented and displayed in a fashion which permits one to establish the tradeoffs among blocking probability, utilization, throughput and delay.

1. INTRODUCTION AND MODEL

Queueing models for computer networks often assume infinite storage at the switching nodes. Such an assumption is questionable. As a result, a storage constraint must be introduced in realistic network models. This we do for a single node in this paper.

In Store-and-Forward (S/F) computer nets, the outgoing channels of a node share a certain number of buffers (S/F buffers). If no feedback is considered, (i.e., no retransmission of rejected messages), the S/F function of a node may be modelled as a set of M/M/1 queueing systems which share a finite waiting room, under some scheme.

The purpose of this paper is to analyze and compare a few existing and/or intuitive storage sharing schemes. The first (and simplest) is the Complete Partitioning (CP) scheme where actually no sharing is provided, but where the entire finite storage (waiting room) is permanently partitioned among the (say) R servers. At the other extreme is the second scheme, Complete Sharing (CS), which is such that an arriving customer is accepted if any storage space is available, independent of the server to which it is directed. CS succeeds in achieving a better performance than CP (smaller probability of blocking) under normal traffic conditions and for fairly balanced input systems. However, for highly asymmetrical input rates (λ_i i=1,..., R) and equal service rates, CS tends to heavily favor servers with higher input rates, even though they may be close to saturation (input rate close to service rate). The failure to recognize servers at or near saturation results in most of the space being occupied by customers waiting for those servers, at the detriment of the others. Moreover, even with perfectly balanced arrival rates (i.e., $\lambda_i=\lambda$ i=1, ..., R), under overload conditions, CS fails (where CP succeeds) in securing a full utilization of all the R servers.

The above considerations intuitively indicate that contention for space must be limited in some way.

In order to avoid the possible utilization of the entire space by any particular output channel, we impose a limit on the number of buffers to be allocated at any time, to any server. This idea is incorporated in our third scheme: Sharing with Maximum Queues (SMXQ). Of course, the sum of those maxima must be greater than the total space if some sharing is to be provided. The SMXQ, however, does not guarantee a full utilization of the servers under heavy traffic conditions. This deficiency motivates the fourth scheme: Sharing with a Minimum Allocation (SMA) scheme. With SMA, a minimum number of buffers is always reserved for each server and, in addition, a common pool of buffers is to be shared among all the servers, with no further constraints on the queue size.

Rich and Schwartz [RICH 75] studied a scheme very similar to SMA except that the entire common storage is dynamically allocated to one server at a time. Drukey [DRUK 75] analyzed the CS scheme with the assumption that all the ρ_i's are equal; for the general case of different ρ_i's he restricts his study to 2 channels. Moreover, problems of this sort are frequently encountered in telephony and are referred to as "graded" systems [SYSK 60]. Their main interest, however, is in sharing (extra) lines as opposed to storage.

In this paper, we intend to characterize the four schemes under steady state conditions; namely, we derive expressions for the probabilities of blocking, the average time in system and the throughput. A comparison of the sharing schemes is also provided. The key to the analysis lies in the fact that, in steady state, the joint probability distribution obeys the well known product form solution for networks of queues [JACK 57], [BASK 75], [KLEI 76] (and the bibliographies therein.)

2. ANALYSIS

2.1 GENERAL SOLUTION

We consider R M/M/1 queueing systems which share a

*This research was supported by the Advanced Research Projects Agency of the Department of Defense under Contract DAHC-73-C-0368.

finite waiting room of size B under one of the above schemes. Queueing system $i(i=1, \ldots, R)$ is characterized by a Poisson input stream at a rate λ_i and an exponential service time of mean $1/\mu C_i$. Customers to be served by server i are referred to as type or class i customers. Arriving customers not admitted to the queue (because of the sharing scheme) depart without service. Accepted (non-rejected) customers are served on a First-Come-First-Serve basis.

The sharing of space introduces dependencies among the R queueing systems. The entire system is a birth-death process [KLEI 76], whose state can be simply described by the vector $\underline{n}=(n_1, \ldots, n_R)$, where n_i is the number of type-i customers. The basic equation which describes the behavior of the system of queues in steady state obeys the well known product form solution for a network of queues; i.e.,

$$\begin{cases} P(\underline{n}) = C_x \rho_1^{n_1} \cdots \rho_R^{n_R} & \text{for } \underline{n} \in F_x \\ \quad = 0 & \text{otherwise} \end{cases} \qquad [1]$$

where $\rho_i = \lambda_i/\mu C_i$.

The subscript x indicates the scheme referred to, i.e., $x \in \{a,b,c,d\}$ where a stands for CP, b for CS, c for SMXQ and d for SMA. F_x represents the set of possible system states.

In what follows, we first characterize C_x for the 4 schemes, then, from the joint probability distribution, we obtain the probability of blocking, the throughput and average delay. We restrict our study here, to the particular case of equal ρ_i's $\rho_i=\rho)$ with some constraints imposed on all types of customers. The more general case of different ρ_i's and further details about this special case may be found in [KAMO 76].

The above considerations lead to the same probability of blocking PB_x, for all types of customers and consequently to the same utilization of the servers $\rho(1-PB_x)$. If we also choose equal λ_i's and μ_i's, then all customers, independent of their types, will incur the same average delay t_x. We also notice that C_x is simply the probability of an empty system $C_x = P(\underline{0})$, and that it can be computed by summing all the probabilities to one, i.e.,

$$C_x^{-1} = \sum_{\underline{n} \in F_x} \rho_1^{n_1} \cdots \rho_R^{n_R} \qquad [2]$$

2.2 CP SCHEME

CP is a degenerate case where actually all the R queueing systems are independent. The basic equations describing the behavior of any of the queues are well known (see for example [KLEI 76]). Furthermore each of those systems is equivalent to CS with only one type of customer. Thus, we have the characterization of CP and of CS. However, we note that

$$F_a = \{\underline{n} \mid o \leq n_i \leq B_o, \ i = 1, \ldots, R\}$$

where B_o is the number of buffers reserved for each type of customer.

2.3 CS SCHEME

We now combine all the individual buffers into a global pool of size B, $(B=RB_o)$. Empty space is allocated FCFS regardless of the type of arriving customer. The above considerations (Eqs. (1) and (2)) lead to

$$F_b = \left\{ \underline{n} \mid o \leq n_i \leq B, \ i = 1, \ldots, R \ ; \ \sum_{i=1}^{R} n_i \leq B \right\}$$

$$C_b^{-1} = \sum_{K=0}^{B} G(K) \qquad [3]$$

where

$$G(K) = \sum_{\Sigma n_i = K} \rho^K = \binom{K+R-1}{R-1} \rho^K \qquad [4]$$

Since we have Poisson arrivals, there the probability of blocking, PB_b, is simply the probability of having B customers in the system; i.e., $PB_b=C_b G(B)$. The average delay, T, may be determined by first finding the average number of customers of a given type and, then by applying Little's result. This leads to

$$T = \frac{1/\mu C}{1-\rho} \; \frac{\sum_{K=0}^{B-1} G(K) - \binom{B+R-1}{R} \rho^B}{\sum_{K=0}^{B-1} G(K)}$$

Of interest are the 2 cases when $\rho=1$ and $\rho=\infty$

$$\rho = 1 : \qquad PB_b = \frac{R}{R+B} = \frac{1}{1+Bo}$$

In other words, it is exactly the same as for a single M/M/1 queue with Bo buffers. This means that for $\rho=1$, CS and CP lead to the same probability of blocking. This fact will be illustrated in the figures below.

$$\rho \to \infty : \qquad PB_b \to 1 \quad \text{and} \quad \rho(1-PB_b) \to B/(B+R-1)$$

The above shows that infinite input rates do not lead to full utilization of the servers, except for R=1 which is the case of CP. Hence, as opposed to CP, a non-degenerate CS (i.e., R>1), behaves worse than CP for $\rho\to\infty$, and in fact, as we will see below, it behaves worse than CP for all $\rho>1$ and better for $\rho<1$ (see Figs. 1,2). This concludes the analysis of CS.

2.4 SMXQ SCHEME

Like CS, SMXQ allows the sharing of a pool of B buffers with a further constraint imposed on the number of buffers to be allocated to any server, at any time. Let b be the maximum number of buffers that can be used by any type of customer; the set of feasible states becomes

$$F_c = \left\{ \underline{n} \;\middle|\; 0 \leq \sum_{i=1}^{R} n_i \leq B, \ 0 \leq n_i \leq b \quad i = 1, \ldots, R \right\}$$

We find that

$$C_c^{-1} = \sum_{K=0}^{B} Q(K)$$

where

$$Q(k) = \sum_{\substack{\sum_{i=1}^{R} n_i=K,\ 0\leq n_i \leq b}} \rho_1^{n_1} \ldots \rho_R^{n_R} = \sum \rho^{\Sigma n_i}$$

Note that the difference between Q(K) and G(K) comes from the added constraint, $n_i \leq b_i$ which greatly complicates the evaluation of C_c^{-1}.

However, if we assume that each queue is allowed to occupy more than half of the entire space, i.e., $b \geq B/_2$, then we obtain the following simple expression for C_c^{-1}

$$C_c^{-1} = \sum_{K=0}^{B} G(K) - R\rho^{b+1} \sum_{K=0}^{B-b-1} G(K)$$

The probability of blocking of any type of customer, say i, is equal to the probability that the entire space is full or that n_i is equal to its maximum b; this leads to

$$PB_c = C_c \rho^B \left[\binom{B + R - 1}{R - 1} - R \binom{B - b + R - 2}{R - 1} \right] + C_c \rho^b \sum_{K=0}^{B-b-1} \binom{K + R - 2}{R - 2} \rho^K$$

The expression for the average delay T is complicated and is not shown here (see [KAMO 76]). We note the following.

1. if b=B then SMXQ is equivalent to CS
2. if $b=B_0$ then SMXQ is equivalent to CP
3. if R=2 then SMXQ is equivalent to SMA with a minimum allocation per queue equal to B-b.

This fact will be taken advantage of in the comparison of the 4 schemes (below).

4. If $\rho \to \infty$ then $\rho(1-PB_c)$, the utilization of any server does not reach one except for R=1 and R=2. Hence for R>2, (and $\rho \to \infty$) SMXQ still does not provide a full utilization of the servers. Our next scheme is motivated by this deficiency.

2.5 SMA SCHEME

Similar to CS, SMA allows the sharing of a pool of B buffers, of which a buffers are permanently allocated to type-i customers, i=1, ..., R. As a result, the set of feasible states becomes

$$F_d = \left\{ \underset{\sim}{n} \;\middle|\; \sum_{i=1}^{R} \sup\{0, n_i - a\} \leq B,\ 0 \leq n_i \leq B+a \quad i=1,..,R \right\}$$

The evaluation C_d is based on the partitioning of the set F_d into subsets corresponding to known subsets of types of customers which exceed their minimum allocations. This leads to

$$C_d^{-1} = \sum_{p=0}^{R} \binom{R}{p} \left(\frac{1-\rho^a}{1-\rho} \right)^{R-p} \rho^{pa} \sum_{K=0}^{B} \binom{K+p-1}{p-1} \rho^K$$

The probability of blocking of any type of customer, say r, is equal to the probability that the shared buffers (B) are full and that $n_r \geq a_r$

$$PB_d = C_d \sum_{p=1}^{R} \binom{R-1}{p-1} \left(\frac{1-\rho^a}{1-\rho} \right)^{R-p} \rho^{pa} \binom{B+p-1}{p-1} \rho^B$$

From the above, we note that if $a \neq 0$ (i.e., non-degenerate SMA) then: $\rho \to \infty \to \rho(1-PB_d)=1$. Hence, as expected, SMA will always secure full utilization under heavy load conditions. The expression of the average delay T is complicated and is not shown here (see [KAMO 76]).

3. COMPARISON OF THE SHARING SCHEMES

We first compare CP and CS, and then the 4 schemes.

3.1 COMPARISON OF CP AND CS

Fig. 1 illustrates the behavior of the probability of blocking PB, with respect to ρ and for a set of values of R, (R=1,...,4). R=1 corresponds to CP, R=2,3,4 correspond to the merging of 2,3,4 single queues. Note that all the curves (with ame B_0) meet at $\rho=1$ where $PB=1/(1+B_0)$. Note also that for $0 \leq \rho \leq 1$, CS leads to a smaller PB, hence a better performance than CP. The improvement is quite considerable for small values of B_0 and increases with R. However, for $\rho \geq 1$, CP shows a slightly better performance (smaller PB) than CS, namely for small values of B_0.

Fig. 2 shows the respective channel utilizations $\rho(1-PB)$. Note the loss in limiting throughput ($\rho \to \infty$) with CS for the smaller value of B_0.

The average delay curves are not shown here; they show, for each set (B_0=4,9),similar behavior as in Fig. 5. In general, the average delay increases as more buffers are provided, i.e., as R increases.

3.2 COMPARISON OF THE FOUR SCHEMES

We assume R=2 and B=6. The maximum queue length b is chosen to satisfy the relation $B/2 \leq b \leq B$, i.e., b=3,4,5,6. From our previous considerations we know that b=3 leads to CP, b=4 and b=5 lead to non-degenerate SMXQ, SMA, and b=6 leads to CS.

Figs. 3,4 and 5 show respectively the probability of blocking PB, the channel utilization $\rho(1-PB)$, the normalized average message delay μCT, obtained with the four schemes. With respect to blocking and utilization, the optimal b (i.e., the optimal scheme) is a function of ρ. We note that for small values of ρ, b=6 (i.e., CS) is optimal; as ρ increases b=5 then b=4 (i.e., SMXQ, SMA) become optimal, and finally, for a larger ρ, b=3 (i.e., CP) becomes optimal. With respect to the average delay, it is an increasing function of b.

We conclude that no one scheme is always optimal; one should select a scheme to fit the particular operational environment. This study shows sharing with some restrictions on the contention of space is certainly more advantageous than no-sharing, especially when little storage is available.

REFERENCES

[BASK 75] F. Baskett, K.M. Chandy, R.R. Muntz and F.G. Palacios, "Open, Closed and Mixed Networks of Queues with Different Classes of Customers," Journal Assoc. Comput. Mach., April 1975.

[DRUK 75] D.L. Drukey, "Finite Buffers for Purists," TRW Systems Group Report No. 75.6400-10-97.

[JACK 57] J.R. Jackson, "Networks of Waiting Lines," Operations Research, 5,518-521, 1975.

[KAMO 76] F. Kamoun, Design Considerations for Large Computer Communication Networks, Ph.D. Dissertation Comp. Sci. Dept., University of Calif. L.A., 1976.

[KLEI 76] L. Kleinrock, Queueing Systems, Vol. II: Computer Applications, Wiley, New York 1976.

[RICH 75] M.A. Rich, "Buffer Sharing in Computer-Communication Network Nodes," Ph.D. Dissertation, Polytechnic Institute of New York, 1975.

[SYSK 60] R. Syski, "Introduction to Congestion in Telephone Systems," Oliver & Boyd, London, 1960.

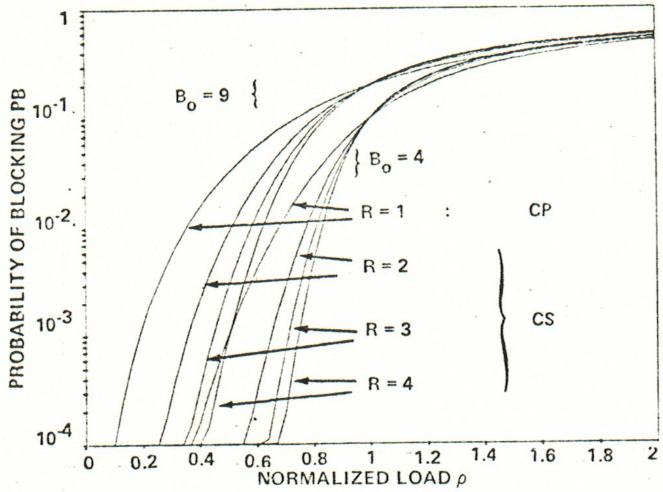

Fig. 1 Comparison of CP and CS: Blocking

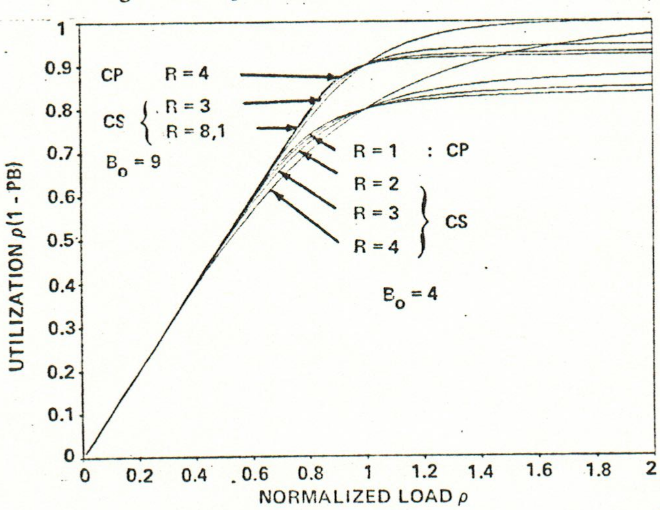

Fig. 2 Comparison of CP and CS: Utilization

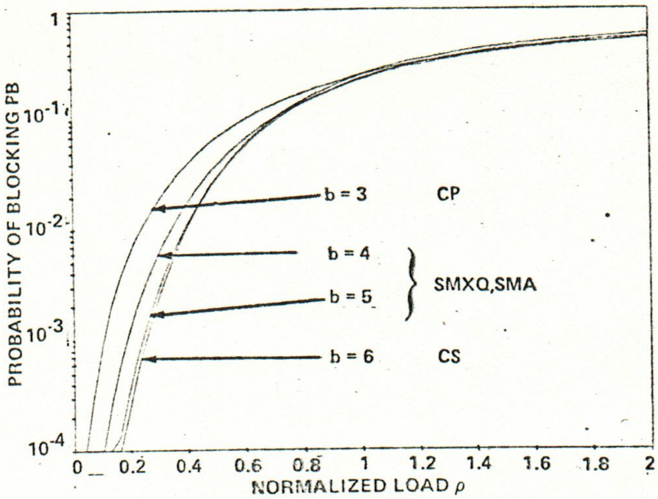

Fig. 3 Comparison of 4 Schemes: Blocking

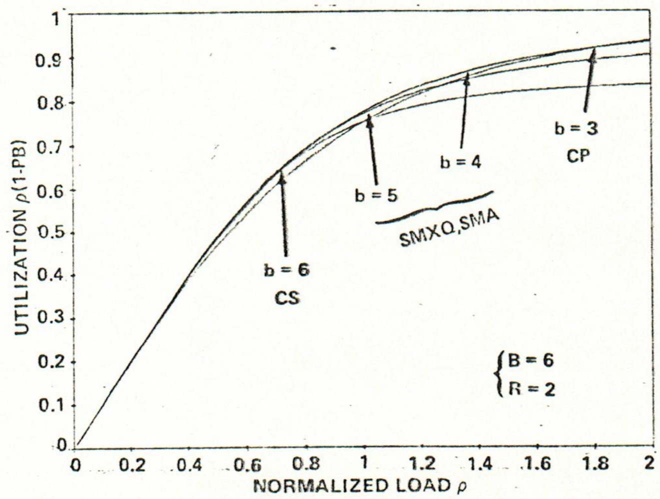

Fig. 4 Comparison of 4 Schemes: Utilization

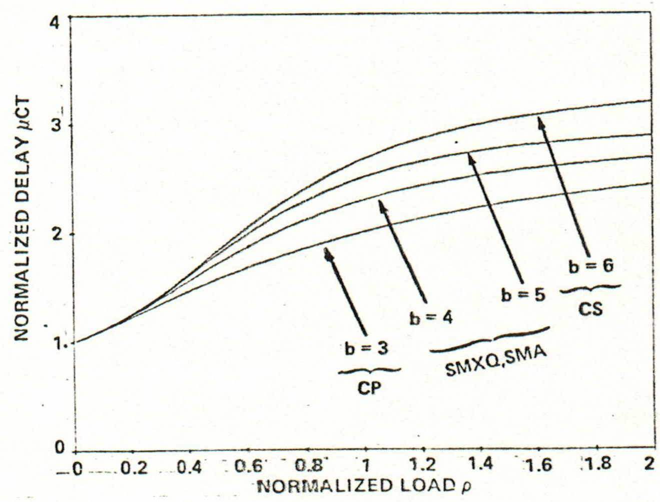

Fig. 5 Comparison of 4 Schemes: Delay

BIOGRAPHIES

Farouk Kamoun was born in Sfax, Tunisia on October 20, 1946. He received the Engineering Degree from Ecole Superieure d'Electricité, Paris, France, in 1970 and the M.S. degree in computer science from UCLA in 1972. He is now a graduate student in the Computer Science Department at UCLA.

Leonard Kleinrock, for biography, see p. of these Proceedings.