The Optimal Buffer Allocation Problem

VENKAT ANANTHARAM, MEMBER, IEEE

Abstract —We have available a fixed number of buffer spaces to be allocated among the nodes of an open network of exponential servers with Bernoulli routing and Poisson arrivals. The goal is to optimize a performance criterion associated with the time to buffer overflow, such as maximizing its mean or maximizing the probability that it exceeds some value. We argue that for any such criterion the assignment should be done roughly in inverse proportion to the logarithms of the effective service rates at the nodes. Here, by effective service rate we mean the ratio of the service rate to the stationary arrival rate at the node in the network with infinite buffers.

I. Introduction

NETWORKS of queues are commonly used as models for the queueing processes taking place in communication networks, computer networks, and manufacturing systems; see [2] for an excellent overview connecting theory with practice in the context of communication networks. An important problem in these applications is how to allocate buffer spaces among the nodes of the network to avoid frequent buffer overflows. Indeed, much is known about the stationary behavior of networks of queues with infinite buffer space at each node, see e.g. [6] and [11]. For designs carried out on the basis of the stationary behavior, one would like to maximize the time the network spends in a regime where it is well approximated by its stationary model.

It is generally accepted that this problem is analytically intractable. In view of this, the problem of estimating the time to buffer overflow by simulation is currently being studied by several investigators. Simulation-based approaches to this problem include the large deviation theory ideas for fast simulation of Cottrell *et al.*, [3] and Parekh and Walrand [8], [12], and the perturbation analysis technique of Ho *et al.*, see [5]. For insight into the situation in light traffic, see Reiman [9]. Sheshkin [10], and Mitra and Tsoucas [7] also consider the problem of buffer allocation for tandem queues.

In this paper we use pathwise probabilistic arguments to justify a simple rule of thumb by which buffer allocation can be carried out. Our model for the underlying network is the skeleton of an open Jackson network. That is, we have J exponential servers with respective service rates $\mu_1, \cdots, \mu_J$, Bernoulli routing with $J+1 \times J+1$ routing matrix R, and exogenous Poisson arrivals of rate γ, with the usual independence assumptions. Recall that this means that customers in queue at node i are served in the order in which they arrive and require service for a time which is exponentially distributed with mean μ_i^{-1}, all service times being independent. The outside world is denoted by 0, and the arrival process is independent of the service times. A customer leaving node i, $0 \le i \le J$ is routed to j, $0 \le j \le J$ with probability r_{ij}, routing decisions being independent and independent of the arrival process and the service times. Further, R is an irreducible stochastic matrix.

By the optimal buffer allocation problem, we mean the problem of how to distribute in the best possible way a fixed number N of available buffer spaces among the nodes of the network. The goal is to optimize some performance criterion associated with the time to buffer overflow, such as its mean or the probability that it exceeds some value. We argue that for *any* such performance criterion the assignment should be done roughly in inverse proportion to the logarithms of the effective service rates.

II. Main Result

We assume that the network is stable, namely, that the solutions of the flow balance equations

$$\lambda_i = \gamma r_{oi} + \sum_{j=1}^{J} \lambda_j r_{ji}, \qquad 1 \le i \le J$$

satisfy

$$\lambda_i < \mu_i, \qquad 1 \le i \le J.$$

This is a natural requirement if the network is to operate for reasonably long periods of time without buffer overflow. Let $\eta_i \triangleq \mu_i / \lambda_i$ denote the effective service rate at node i.

We now introduce some notation. Throughout, σ denotes an infinite subset of the positive integers. We write $\lim_{N \underset{\sigma}{\to} \infty}$ for N going to ∞ along the subsequence σ. Given two functions $f(N)$ and $g(N)$ on the positive integers, we write $f(N) = o_\sigma(g(N))$ if $\lim_{N \underset{\sigma}{\to} \infty} f(N)/g(N) = 0$ and $f(N) = \omega_\sigma(g(N))$ if $\lim_{N \underset{\sigma}{\to} \infty} f(N)/g(N) = \infty$. Logarithms are to base 2.

A *buffer allocation scheme* is a way to assign, for each N, N_i buffers to node i, such that $N_i > 0$, $1 \le i \le J$, and $\sum_{i=1}^{J} N_i = N$, i.e., it is a function on the positive integers taking N to a positive J-tuple summing to N. Here in counting buffer spaces we include the space in service. Let T have the distribution of the time to buffer overflow

Manuscript received November 2, 1987; revised October 1988. This work was supported in part by National Science Foundation Grant NCR-8710840. This paper was presented at the 1988 Conference on Information Sciences and Systems, Princeton University, Princeton, NJ, March 1988.

The author is with the School of Electrical Engineering, 313 Phillips Hall, Cornell University, Ithaca, NY 14853, USA.

IEEE Log Number 8929030.

0018-9448/89/0700-0721$01.00

when the network is started empty. This depends on N and the associated buffer allocation, but the dependence will be suppressed from the notation. In the next section we prove the following theorem.

Theorem 1: For any buffer allocation scheme $(N_1,\cdots,N_J)$, $\sum_{i=1}^J N_i = N$, let $g(N) \triangleq \min_i \eta_i^{N_i}$. For any subsequence σ and any $\tau(N)$ such that $\tau(N) = o_\sigma(g(N))$, we have

$$\lim_{N \to_\sigma \infty} P(T \le \tau(N)) = 0. \tag{2.1}$$

On the other hand, for any $T(N)$ such that $T(N) = \omega_\sigma(Ng(N))$ we have

$$\lim_{N \to_\sigma \infty} P(T \le T(N)) = 1. \tag{2.2}$$

The theorem appears to justify the following heuristic buffer allocation rule. □

Rule of Thumb: In allocating N buffers to "maximize the time to buffer overflow," one should allocate roughly a fraction p_i of the N buffers to node i, where p_i is inversely proportional to $\log \eta_i$.

Justification: Let $c \triangleq p_i \log \eta_i$ (independent of i). For any buffer allocation scheme $(N_1,\cdots,N_J)$, $\sum_{i=1}^J N_i = N$, denote $Ng(N)/2^{cN}$ by $h^2(N)$. Suppose

$$\liminf_{N \to \infty} Ng(N)/2^{cN} = 0.$$

Then there is a subsequence σ such that $Ng(N)/2^{cN} = o_\sigma(1)$. Then also $h(N) = o_\sigma(1)$. If T and T^* have the distribution of the time to buffer overflow for this buffer allocation scheme and for the rule of thumb, respectively, then we have

$$\lim_{N \to_\sigma \infty} P(T \le 2^{cN}h(N)) = \lim_{N \to_\sigma \infty} P\left(T \le \frac{Ng(N)}{h(N)}\right) = 1 \tag{2.3}$$

by (2.2), whereas

$$\lim_{N \to_\sigma \infty} P(T^* \le 2^{cN}h(N)) = 0 \tag{2.4}$$

by (2.1). Clearly, the buffer allocation scheme is worse than the rule of thumb along σ. Modifying the scheme along σ by replacing it with the rule of thumb is an improvement in the following sense: if $\hat{T}$ has the distribution of the time to buffer overflow for the modified scheme, $t(N)$ is any function, and s any subsequence, then

$$\limsup_{\substack{N \to \infty \\ s}} P(\hat{T} \le t(N)) \le \limsup_{\substack{N \to \infty \\ s}} P(T \le t(N)). \tag{2.5}$$

Indeed, if σ and s are eventually disjoint, i.e., $\sigma \cap s$ is finite, then (2.5) is obviously true with equality, whereas, if $\sigma \cap s$ is infinite, it suffices to show that

$$\limsup_{\substack{N \to \infty \\ \sigma \cap s}} P(T^* \le t(N)) \le \limsup_{\substack{N \to \infty \\ \sigma \cap s}} P(T \le t(N)). \tag{2.6}$$

If $t(N) \ge 2^{cN}h(N)$ infinitely often along $\sigma \cap s$, the right side of (2.6) is 1 by (2.3), while if $t(N) \le 2^{cN}h(N)$ eventually along $\sigma \cap s$, the left side is 0 by (2.4). From this, (2.6) follows.

This suggests that any reasonable buffer allocation scheme must satisfy

$$\liminf_{N \to \infty} \frac{Ng(N)}{2^{cN}} > \epsilon$$

for some $\epsilon > 0$. It is a straightforward calculation to see that this implies that

$$|N_i - p_i N| < K \log N \tag{2.7}$$

for all sufficiently large N, for all $i = 1,\cdots,J$, where $K \triangleq \sum_{i=1}^J (\log \eta_i)^{-1}$. In this sense the rule of thumb is approximately optimal.

III. Proof

In this section we prove Theorem 1. Let $\mathbf{N}$ denote the network with service, routing and arrivals as in the preceding section but with infinite buffer space at each node. $\mathbf{N}$ is an open Jackson network which can be described by a Markov process $\boldsymbol{x} = (X(t), t \ge 0)$ with state space $X = \{(n_1,\cdots,n_J): n_i \ge 0, 1 \le i \le J\}$. Given N and $(N_1,\cdots,N_J)$ with $\sum_{i=1}^J N_i = N$, let $A = \{(n_1,\cdots,n_J) \in X:\ n_i > N_i$ for some $i = 1,\cdots,J\}$. With T defined as in the preceding section, it is easy to see that for any $t \ge 0$,

$$P(T > t) = P_0(X(s) \notin A, \quad \text{for all } 0 \le s \le t)$$

where the subscript 0 in P_0 denotes the state $(0,\cdots,0) \in X$. In this section we will, therefore, write T for $\inf\{t > 0: X(t) \in A\}$ and discuss the distribution $(P_0(T \le t),\ t \ge 0)$.

Given $\tau(N)$ as in the statement of the theorem, we first prove (1.1) by estimating $P_0(T \le \tau(N))$. Let $\alpha \triangleq P_0(X(t)$ visits A before returning to 0). It turns out it is enough to upper bound α.

We first lower bound ET using an ergodic argument. Let $t_0(t)$ and $t_A(t)$ denote, respectively, the time spent by $\boldsymbol{x}$ in state 0 and in A up to time t. From the ergodic theorem for Markov chains (see e.g. [4]) we have

$$\lim_{t \to \infty} \frac{t_A(t)}{t_0(t)} = \frac{\pi(A)}{\pi(0)} \tag{3.1}$$

where π denotes the stationary distribution of $\boldsymbol{x}$.

It is well-known (see e.g. [6], [11]) that π is given by

$$\pi(n_1,\cdots,n_J) = \prod_{i=1}^J \left(\frac{\lambda_i}{\mu_i}\right)^{n_i}\left(1 - \frac{\lambda_i}{\mu_i}\right). \tag{3.2}$$

From this a simple calculation gives that

$$\pi(A) = 1 - \prod_{i=1}^J \left(1 - \eta_i^{-(N_i+1)}\right) \le \frac{C}{g(N)} \tag{3.3}$$

where $g(N)$ is as in the statement of the theorem and C is a constant independent of N and $(N_1,\cdots,N_J)$. From (3.1)–(3.3) we have

$$\lim_{t \to \infty} \frac{t_A(t)}{t_0(t)} \le \frac{C}{g(N)} \tag{3.4}$$

where C is a constant independent of N and $(N_1,\cdots,N_J)$.

On the paths of the process we can construct independent identically distributed (i.i.d.) random variables $(T_k, k$

≥ 1) with the distribution of T and i.i.d. random variables $(S_k, k \geq 1)$ such that S_k has the distribution of the time taken to return to 0 starting from the hitting distribution of A. Formally,

$$T_k = \inf\left\{t > 0:\ X\left(t + \sum_1^{k-1} T_l + \sum_1^{k-1} S_l\right) \in A\right\}$$

$$S_{k-1} = \inf\left\{t > 0:\ X\left(t + \sum_1^{k-1} T_l + \sum_1^{k-2} S_l\right) = 0\right\}.$$

Since the state does not visit 0 on the intervals $[\Sigma_1^k T_l + \Sigma_1^{k-1} S_l, \Sigma_1^k T_l + \Sigma_1^k S_l)$, $k = 1, 2, \cdots$, a standard regenerative argument shows that

$$\lim_{t \to \infty} \frac{t_0(t)}{t} \leq \frac{ET}{ET + ES_1}. \tag{3.5}$$

Since the time spent in any state stochastically dominates an exponential random variable of rate $\gamma + \Sigma_{i=1}^J \mu_i$, we also have

$$\lim_{t \to \infty} \frac{t_A(t)}{t} \geq \frac{\left(\gamma + \sum_{i=1}^J \mu_i\right)^{-1}}{ES_1 + ET}. \tag{3.6}$$

From (3.4)–(3.6), it follows that

$$ET \geq Cg(N) \tag{3.7}$$

for a constant C, independent of N and $(N_1, \cdots, N_J)$.

We proceed to upper bound α. Let δ have the distribution of the time taken to return to 0, starting from 0 and conditioned on not visiting A. Let $\delta_1, \delta_2, \cdots$ be i.i.d. with the distribution of δ. Let Δ have the distribution of the time to hit A starting from 0 and conditioned on not returning to 0. Also, assume that Δ is independent of $(\delta_n, n \geq 1)$. Let ν be a geometric random variable independent of Δ and $(\delta_n, n \geq 0)$ with

$$P'(\nu = k) = \alpha(1-\alpha)^k, \qquad k = 0, 1, 2, \cdots. \tag{3.8}$$

Then it is easy to see that

$$T \overset{d}{=} \sum_{k=1}^{\nu} \delta_k + \Delta \tag{3.9}$$

where $\overset{d}{=}$ denotes equality in distribution. In particular,

$$ET = \frac{1-\alpha}{\alpha} E\delta + E\Delta = \frac{1}{\alpha}\left[(1-\alpha)E\delta + \alpha E\Delta\right].$$

Observe that $(1-\alpha)E\delta + \alpha E\Delta$ is the mean time taken to either return to 0 or visit A starting from 0. This time is stochastically dominated by the time to return to 0 starting from 0. It follows that

$$ET \leq \frac{C}{\alpha} \tag{3.10}$$

for a constant C independent of N and $(N_1, \cdots, N_J)$. From (3.7) and (3.10) we have

$$\alpha \leq \frac{C}{g(N)} \tag{3.11}$$

for a constant C independent of N and $(N_1, \cdots, N_J)$.

Now observe that δ stochastically dominates an exponential random variable of mean γ^{-1}. Indeed, if the network is empty, we have to wait at least that long for an arrival. From (3.8) and (3.9), and since an independent geometric sum of i.i.d. exponential random variables is exponential, it follows that T stochastically dominates an exponential random variable of mean $1 - \alpha/\alpha\gamma$. Hence

$$P(T \leq \tau(N)) \leq 1 - \exp\left(-\frac{\alpha\gamma\tau(N)}{1-\alpha}\right). \tag{3.12}$$

It is easy to see that for some $\epsilon > 0$ independent of N and $(N_1, \cdots, N_J)$ we have $\alpha < 1 - \epsilon$. Hence from (3.11) and (3.12) it follows that if $\tau(N) = o_\sigma(g(N))$, then

$$\lim_{N \underset{\sigma}{\to} \infty} P(T \leq \tau(N)) = 0,$$

completing the proof of (2.1). □

Given $T(N)$ as in the statement of the theorem, we proceed to prove (2.2). From (3.2), a calculation gives

$$\pi(A) \geq \frac{C}{g(N)} \tag{3.13}$$

for a constant C independent of N and $(N_1, \cdots, N_J)$. The time spent by x in A on an interval $[\Sigma_1^{k-1} T_l + \Sigma_1^{k-1} S_l, \Sigma_1^k T_l + \Sigma_1^k S_l)$, $k = 1, 2, \cdots$ is stochastically dominated by S_1. Thus we have

$$\lim_{t \to \infty} \frac{t_A(t)}{t} \leq \frac{ES_1}{ET + ES_1} \tag{3.14}$$

(cf. (3.6)). From the ergodic theorem we also have

$$\lim_{t \to \infty} \frac{t_A(t)}{t} = \pi(A). \tag{3.15}$$

From (3.13)–(3.15) it follows that

$$ET \leq Cg(N) ES_1 \tag{3.16}$$

for a constant C independent of N and $(N_1, \cdots, N_J)$. Applying the Markov inequality (see e.g. [1]) gives

$$P(T > T(N)) \leq C \frac{g(N)}{T(N)} ES_1 \tag{3.17}$$

for the same constant C.

It remains to estimate ES_1. This is done via the following lemma.

Lemma 1: Given $(m_1, \cdots, m_J)$ and $(m_1', \cdots, m_J')$ such that $m_i \leq m_i'$, $1 \leq i \leq J$, let $L = \Sigma_{i=1}^J m_i$ and $L + L' = \Sigma_{i=1}^J m_i'$, so that $L' \geq 0$.

Let $D = \{(n_1, \cdots, n_J): \Sigma n_i \leq L'\}$. Let S have the distribution of the time to empty when N is started with m_i customers in queue at node i, $1 \leq i \leq J$. Let S_D have the distribution of the time to hit D when N is started with m_i' customers in queue at node i, $1 \leq i \leq J$. Then

$$S_D \underset{s}{\leq} S$$

where $\underset{s}{\leq}$ denotes stochastic domination.

Proof: We prove the claim by coupling the paths in the two situations, using a coloring idea. The point behind this idea is that when a virtual service occurs at a node with a nonempty queue, we are free to decide which customer in the queue departs without affecting the process of the total number of customers at the nodes of the network. At time 0, start with m_i' customers in queue i. Color the leading m_i customers red and the remaining $m_i' - m_i$ customers blue. There is a total of L' blue customers initially. Arriving customers are colored red. Red customers always have precedence over blue customers, i.e., when a virtual service takes place at node i, a blue customer in queue at node i does not move unless there are no red customers in queue at node i. Let $(X_i(t), t \geq 0)$ denote the number of red customers in queue at node i and $(X_i'(t), t \geq 0)$ the total number of customers in queue at node i.

Let

$$\hat{S} = \inf\{t > 0: (X_1(t), \cdots, X_J(t)) = 0\}$$

and

$$\hat{S}_D = \inf\{t > 0: (X_1'(t), \cdots, X_J'(t)) \in D\}.$$

Then $\hat{S} \stackrel{d}{=} S$ and $\hat{S}_D \stackrel{d}{=} S_D$. At $\hat{S}$, all customers in the network are blue and their total number is at most L'. Hence

$$\hat{S}_D \leq \hat{S}$$

from which the claim follows.

Remark 1: The coupling idea of Lemma 1 gives a joint realization for the network started at $(m_1', \cdots, m_J')$ and the network started at $(m_1, \cdots, m_J)$ such that the total number of customers in the system for the first initial condition is always at most L' larger than the total number of customers in the system for the second initial condition. This result may be of independent interest.

Corollary 1: Let S^i, $1 \leq i \leq J$ have the distribution of the time to empty when N is started with all queues empty except for a single customer in queue at node i. Let S^* have the smallest distribution which stochastically dominates all the S^i, $1 \leq i \leq J$.

Suppose N is started with n_i customers in queue at node i. Let

$$L = \sum_{i=1}^{J} n_i.$$

Let S have the distribution of the time to empty. Then

$$S \underset{s}{\leq} \text{ sum of } L \text{ independent copies of } S^*.$$

Proof: We prove the claim by induction on L. Clearly, the claim is true for $L = 1$. Suppose the claim is true for all initial conditions where the total number of initial customers is at most $L-1$, and let N be started with n_i customers in queue at node i, $1 \leq i \leq J$, where $\sum_{i=1}^{J} n_i = L$. Let

$$D = \left\{(m_1, \cdots, m_J): \sum_{i=1}^{J} m_i \leq L-1\right\}.$$

Let j be any index such that $n_j > 0$. Let S_D denote the time of the first visit to D. By Lemma 1,

$$S_D \underset{s}{\leq} S^j \underset{s}{\leq} S^*$$

where the second inequality is by the definition of S^*.

At S_D the network is in a state in D. While the hitting distribution on D and S_D may be dependent, we know from the induction hypothesis that from any initial condition in D the time to empty is stochastically dominated by the sum of $L-1$ independent copies of S^*. The claim follows easily. □

Corollary 1 tells us that S_1 is stochastically dominated by the sum of N independent copies of a random variable with finite mean. Hence

$$ES_1 \leq CN \tag{3.18}$$

for a constant C independent of N and $(N_1, \cdots, N_J)$. From (3.17) and (3.18), we see that if $T(N) = \omega_\sigma(Ng(N))$, then

$$\lim_{N \to_\sigma \infty} P(T \leq T(N)) = 1$$

which proves (2.2). □

IV. Concluding Remarks

The rule of thumb suggests that at least for large N one could restrict attention to the buffer allocations satisfying (2.7). The cardinality of this set of allocations grows like $(\log N)^{J-1}$, in comparison to N^{J-1} for the set of all possible allocations. For a specific performance criterion, such as the mean time to buffer overflow, it may therefore be feasible in practice to determine the optimal allocation in the region determined by (2.7).

If the performance criterion is the mean time to buffer overflow, it is not necessary to study asymptotics to justify the rule of thumb. From (3.7), (3.16), and (3.18) we see that there are constants c_1 and c_2 such that

$$c_1 g(N) < ET < c_2 N g(N)$$

for *every* N. From this it follows easily that unless a buffer allocation $(N_1, \cdots, N_J)$ of N buffers satisfies

$$|N_i - p_i N| < K(\log N + \log c_2 c_1^{-1})$$

for every $i = 1, \cdots, J$, it has a smaller mean time to buffer overflow than the rule of thumb. The reader can easily write down explicit values for c_1 and c_2 from the proof in Section III.

References

[1] P. Billingsley, *Probability and Measure.* New York: Wiley Interscience, 1986.

[2] D. Bertsekas and R. Gallager, *Data Networks.* Englewood Cliffs, NJ: Prentice-Hall Inc., 1987.

[3] M. Cottrell, J. Fort and G. Malgouyres, "Large deviations and rare events in the study of stochastic algorithms," *IEEE Trans. Automat. Contr.*, vol. AC-28, no. 9, pp. 907–920, Sept. 1983.

[4] D. Freedman, *Markov Chains.* New York: Springer-Verlag, 1983.

[5] Y. C. Ho, M. A. Eyler and T. T. Chien, "A gradient technique for

general buffer storage design in a production line," *Int. J. Production Res.*, vol. 17, no. 6, pp. 557–580, 1979.

[6] F. Kelly, *Reversibility and Stochastic Networks*. New York: Wiley, 1979.

[7] D. Mitra and P. Tsoucas, "Convergence of relaxations for numerical solutions of stochastic problems," in *Computer Performance and Reliability*, G. Iazeolla, P. J. Courtois, and O. J. Boxma, Eds. New York: Elsevier Science (North-Holland), 1988, pp. 119–133.

[8] S. Parekh, "Rare events in networks," Ph.D. dissertation, Electrical Engineering and Computer Science Department, University of California, Berkeley, 1986.

[9] M. Reiman, "The optimal buffer allocation problem in light traffic," in *Proc. 26th IEEE Conf. Decision and Control*, Los Angeles, CA, Dec. 9–11, 1987, pp. 1409–1503.

[10] T. J. Sheskin, "Allocation of interstage storage along an automatic production line," *AIIE Trans.* vol. 8, no. 1, pp. 146–152, 1975.

[11] J. Walrand, *An Introduction to Queueing Networks*. Englewood Cliffs, NJ: Prentice-Hall, 1987.

[12] J. Walrand, "Quick simulations of networks of GI/G/1 queues," Plenary lecture presented at the 16th Conf. Stochastic Processes and Their Applications, Stanford University, Stanford, CA, Aug. 16–21, 1987.