An Approximate Model of Processor Communication Rings under Heavy Load

E. G. Coffman, Jr., *L. Flatto* and *E. N. Gilbert*

Bell Laboratories, Lucent Technologies
Murray Hill, New Jersey 07974

A. G. Greenberg

AT&T Labs – Research
Murray Hill, New Jersey 07974

ABSTRACT

A communication ring of N cells rotates unidirectionally in discrete steps, carrying messages (packets) among N processors at fixed locations around the ring. Each cell holds only one packet. A packet's destination is chosen so that packets are delivered after being on the ring for i.i.d. random numbers of steps. When a cell makes a delivery, the receiving processor uses the same cell to send a new packet. Thus, at every step every cell is moving a packet.

The packet delivery process in this model gives a useful approximation to the corresponding process in rings where stable, but typically very long queues of waiting packets form at the processors (see [1, 2]). The distribution of times between successive deliveries is derived for a general distribution of packet transit times on the ring. A limit law shows that, for large N, the former distribution is approximately exponential when packet destinations are chosen uniformly at random from among the processors.

1. Introduction

A number N of processors is served by a ring of N cells carrying packets of information among the processors, with no cell carrying more than one packet at a time. The ring resembles a circular conveyor of cells rotating clockwise in discrete steps past fixed processors. Cells and processors are labeled by integers $0, 1, \ldots, N-1$, with cells numbered counterclockwise and processors clockwise. If cell i is now at processor j then t steps later cell $i + t \pmod N$ will be at processor j and cell i will be at processor $j + t \pmod N$. When a packet completes its time on the ring, its cell discharges the packet and accepts a new packet in return, all in the same step. Figure 1 shows an example.

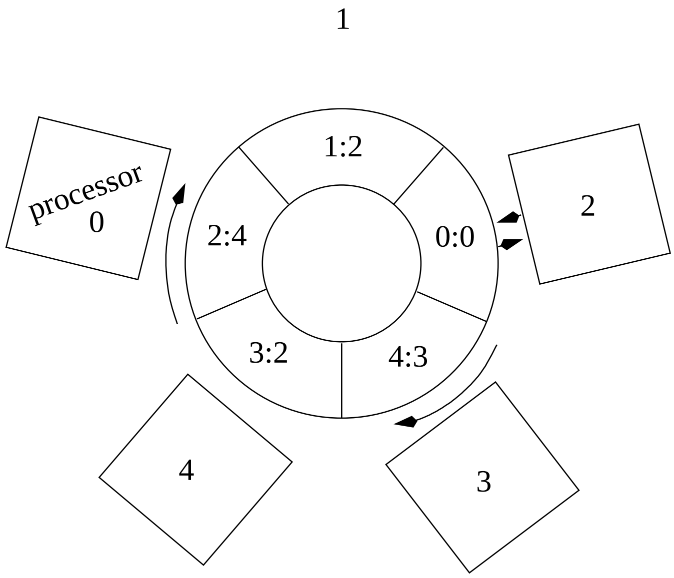

Figure 1: A ring of 5 cells. The first number inside each cell is the label of that cell while the second is the number of steps that remain before its packet will leave. Thus processor 2 is receiving a packet from cell 0 and giving cell 0 a new packet. None of the cells 1, 2, 3, 4 now have packets for processor 2; processor 2 will have to wait at least 3 more steps to load its next packet on the ring (it will be exactly 3 steps only if, after 2 steps, processor 1 loads a packet that leaves the ring after 1 step). Processor 1, however, will surely load packets after 2 and 3 steps.

The times that packets spend on the ring are assumed to be independent samples of a random variable X with given probabilities $g_k = P(X = k)$, $k \geq 1$. Note that the support of $\{g_k\}$ can be any subset of the positive integers; a packet may go around the ring one or more times before getting off at its destination. Of interest is the time that a processor must wait between its opportunities to send a new packet.

We emphasize that a processor always has a new packet ready to put into a cell making a delivery; at every time step, every cell moves a packet from one processor to the next.

Thus, queueing theory plays no role in the model defined here. On the other hand, our model approximates the packet delivery process in the queueing model of [1, 2] under heavy loading. In the latter model, which has the same ring structure, a stable queue of waiting packets forms at each processor, X is geometric with mean N/μ, and new packets arrive at each processor in a Bernoulli process with mean λ/N per step. For fixed λ and μ, with $\lambda < \mu$ for stability, the mean waiting time in a queue was found to be $o(N)$ for large N. That bound was improved to $O(1)$ in [2]. See [1] for many other references to the literature on processor rings.

Returning to our model, note that the cell at any given processor will return there only after complete cycles of the ring, i.e., after multiples of N steps. Section 2 finds the distribution of the number C of cycles between successive deliveries to a given processor by a given cell. Section 3 uses that result to derive the distribution of the time between successive deliveries at a given processor, taking all cells into account, and to get numerical results. Section 3 also derives the distribution of a processor's wait W for service, starting from a random time instead of a delivery. Section 4 proves limit theorems for the distributions of C and W for large N when X is uniformly distributed on $\{1, \ldots, N-1\}$.

2. Single Cell

This section studies the ring from the point of view of a single cell, say cell 0. Let ξ_n denote the processor to which cell 0 delivers its n^{th} packet. Then ξ_n is a Markov chain with states $0, \ldots, N-1$, the integers modulo N. In terms of the transit-time probabilities g_k, the transition probabilities of $\{\xi_n\}$ are

$$P(\xi_n = (\xi_{n-1} + j) \bmod N | \xi_{n-1}) = g_j + g_{N+j} + \cdots.$$

Let K be the set of integers k for which $g_k > 0$, and let D be the set of differences $|k - k'| > 0$ with k, k' in K. *We assume hereafter that the greatest common divisor of the integers in the set* $\{N\} \cup D$ *equals 1.* This condition is both necessary and sufficient for the irreducibility and aperiodicity of the chain $\{\xi_n\}$ (see, e.g., Spitzer [6]). Further, $\{\xi_n\}$ will then have the limiting uniform distribution on $\{0, \ldots, N-1\}$.

Now consider a particular processor, say 0, and the times at which cell 0 delivers packets to processor 0. Since the state of the Markov chain is known to be 0 at these times, the numbers of (complete) cycles between these successive delivery times are independent samples of a random variable C. The mean of C is easy to find; one need only observe that, since $1/E[X]$ is the rate at which cell 0 delivers packets to processors, both $1/(NE[X])$ and $1/(NE[C])$ give the rate at which cell 0 delivers to processor 0. This proves

Theorem 1.

$$E[C] = E[X].$$

The probability $r_c = P(C = c)$ will now be derived. Let the Markov chain start in state $\xi_0 = 0$ so that r_c is the probability of a first return to state 0 at time cN. The connection between r_c and g_k can be made in terms of their generating functions $R(z) = \sum_{c=1}^{\infty} r_c z^c$ and $G(z) = \sum_{k=1}^{\infty} g_k z^k$.

Theorem 2. *For* $|z| < 1$,

$$R(z^N) = 1 - \frac{1}{\frac{1}{N}\sum_{m=0}^{N-1}[1 - G(\zeta_m z)]^{-1}},$$

where $\zeta_m = e^{2\pi im/N}$, $0 \leq m \leq N-1$.

Proof: The n^{th} delivery by cell 0 occurs at a time that is a sum of n independent samples from the probability distribution g_k (by definition, the 0^{th} delivery occurs at time 0). Then $G^n(z)$ is the generating function for the probability distribution of the time of the n^{th} delivery to some processor. Observe that $|G(z)| < 1$ for $|z| < 1$ so a sum on n gives

$$U(z) = \frac{1}{1 - G(z)} \tag{1}$$

for the generating function of the probability u_j that cell 0 makes some delivery at time j. This delivery will be to processor 0 if j is a multiple of N. Then $v_c = u_{cN}$ is the probability of a delivery (not necessarily the first) to processor 0 after c complete cycles. A generating function $V(z)$ for v_c follows from (1). Indeed, $V(z^N)$ is obtained by deleting from $U(z)$ all terms with index not divisible by N. The deletion from (1) can be done formally by using the following identity for sums of powers of roots of unity:

$$\begin{aligned} \frac{1}{N}\sum_{m=0}^{N-1} \zeta_m^j &= 1 \quad \text{if } N|j \\ &= 0 \quad \text{otherwise.} \end{aligned}$$

The result is

$$V(z^N) = \frac{1}{N}\sum_{m=0}^{N-1} U(\zeta_m z). \tag{2}$$

The probability v_c is the probability that some unspecified number of independent first passages from state 0 back to itself is completed on cycle c. Exactly as in the derivation of (1), one now has

$$V(z) = \frac{1}{1 - R(z)}. \tag{3}$$

Solving (3) for $R(z^N)$ and using (2) proves the theorem. □

In computations, the expression for $R(z)$ in terms of roots of unity is less useful than recurrences. The desired recurrences can be derived directly or, just as easily, from (1) and (3). Thus (1), rewritten as $U(z) = 1 + G(z)U(z)$, implies the following recurrence for u_j:

$$u_0 = 1, \qquad u_j = \sum_{k=1}^{j} u_{j-k} g_k, \quad j \geq 1, \tag{4}$$

The solution of (4) also gives $v_c = u_{cN}$. Then (3), rewritten $R(z) = V(z) - 1 - R(z)(V(z) - 1)$, gives a recurrence for r_c:

$$r_1 = v_1 = u_N, \qquad r_c = v_c - \sum_{k=1}^{c-1} r_{c-k} v_k, \quad c \geq 2. \tag{5}$$

3. Waiting Times

This section considers the sequence of times at which a particular processor, say 0, receives deliveries from all cells. The number S of steps between successive deliveries to processor 0 is a random variable with the distribution $q_t = P(S = t)$. The mean of S is simply

$$E[S] = E[X], \tag{6}$$

since the rate $N/E[S]$ at which packets arrive to the ring must equal the rate $N/E[X]$ of departures from the ring. To find q_t, suppose cell 0 delivers a packet to processor 0 at time 0. Write $t = cN + a$ for integers a, c with $0 \leq a \leq N-1$. Then cell a is the cell at processor 0 at time t. Further, $S = t$ if cell a delivers at time t and no cell has delivered to processor 0 at times $1, \ldots, t-1$. Since the N cells operate independently of one another, q_t will be a product of N factors; each factor is a probability that a particular cell obeys the conditions required for $S = t$.

The factor for cell 0 will contain probabilities r_c found in Section 2, and the probabilities $R_c = 1 - r_1 - \cdots - r_c$ that cell 0 does not deliver to processor 0 on cycle $1, \ldots, c$. Other cells have different factors because they do not start out delivering a packet. The formulas below for these other cells apply to the stationary regime. In this regime, when a cell other than 0 makes its first visit to processor 0, it will contain a packet with a random remaining transit time having the residual-life distribution

$$y_0 = \frac{1}{E[X]}, \qquad y_m = \frac{1 - g_1 - \cdots - g_m}{E[X]}, \quad m > 0. \tag{7}$$

The factors for a cell other than 0 contain the probabilities r_c^* and R_c^*, where r_c^* is the probability that such a cell makes its first delivery to processor 0 exactly c cycles after its initial visit, and where R_c^* is the probability that such a cell does not deliver to processor 0 on cycle $0, 1, \ldots, c$, i.e.,

$$R_c^* = 1 - r_0^* - \cdots - r_c^*.$$

A means for calculating the r_c^* will be given after the proof of the following theorem.

Theorem 3. *Suppose cell* 0 *releases a packet at processor* 0 *at time* 0*, and that the remaining cells at their first subsequent visit to processor* 0 *carry packets with the residual-life distribution of transit times. Write* $t = cN + a$ *for integers* $c \geq 0$ *and* a, $0 \leq a \leq N-1$. *Then*

$$q_t = \begin{cases} \frac{1}{E[X]}\left(1 - \frac{1}{E[X]}\right)^{a-1} & c = 0, a > 0 \\ r_c (R_{c-1}^*)^{N-1} & c > 0, a = 0 \\ r_c^* R_c (R_c^*)^{a-1} (R_{c-1}^*)^{N-1-a} & c, a > 0. \end{cases} \tag{8}$$

Proof: Consider the three cases shown in (8):
(i) $c = 0,\ a > 0$. In this case, no cell has visited processor 0 more than once in $[0, t]$. At the arrival of the first a cells, each independently makes a release with probability $y_0 = 1/E[X]$, so one obtains the geometric probabilities shown for this case.
(ii) $c > 0,\ a = 0$. In this case, cell 0 has made its next release on its c^{th} return to processor 0 (with probability r_c), and the remaining cells have made no release in their next $c-1$ visits to processor 0 (each of these $N-1$ independent events having probability R_{c-1}^*). The equation for the second case follows.
(iii) $c, a > 0$. Under this assumption, cell a has released its first packet on its $(c+1)^{\text{st}}$ visit to processor 0 (an event of probability r_c^*); cell 0 released no packet on any of its first $c+1$ returns to processor 0 (an event of probability R_c); cells $1, \ldots, a-1$ have visited processor 0 $c+1$ times, each time without releasing a packet (probability R_c^* of each such event); and cells $a+1, \ldots, N-1$ have visited processor 0 only c times, each time without releasing a packet (probability R_{c-1}^* for each event). The third case in (8) follows. □

Now consider the calculation of the r_c^*. When a cell with remaining transit time m visits processor 0, it will deliver packets after times $j \geq m$ with probability u_{j-m}. For cells other than cell 0, m is sampled from the distribution (7). For those cells, the probability of a delivery j steps after the initial visit at 0 is

$$u_j^* = \sum_{m \geq 0} y_m u_{j-m}. \tag{9}$$

The probability of a delivery c complete cycles after the initial visit to processor 0 is $v_c^* = u_{cN}^*$. The probability that this delivery is the first one that the cell makes to processor 0 is r_c^*, as needed in (8). With $R(z)$, $V^*(z)$, and $R^*(z)$, the generating functions for r_c, v_c^*, and r_c^*, an identity

$$V^*(z) = R^*(z)[1 + R(z) + R^2(z) + \cdots] = \frac{R^*(z)}{1 - R(z)} \tag{10}$$

follows because each delivery to processor 0 is either a first delivery or a first delivery followed by some unspecified number of first deliveries spaced integer numbers of cycles apart. Solving (10) for $R^*(z)$ and expanding in powers of z, one finds a recurrence for r_c^*,

$$r_1^* = v_1^*, \qquad r_c^* = v_c^* - \sum_{j=1}^{c-1} r_{c-j} v_j^*, \quad c \geq 2. \tag{11}$$

Another delay of interest to processor 0 is the waiting time W to its next chance to discharge a packet, starting from a random initial time instead of from a delivery time. If the initial time happens to coincide with a delivery, processor 0 discharges a packet immediately and the waiting time is taken to be 0. The distribution $\{P(W = t)\}$ can be computed as the residual-life distribution of $\{q_t\}$, i.e., $P(W = t) = [1 - \sum_{i=1}^{t} q_i]/E[S]$. The awkward calculations of this approach are avoided in the proof below.

Theorem 4. *For $t = cN + a$ with $c = \lfloor t/N \rfloor$ and $0 \leq a \leq N - 1$,*

$$P(W > cN + a) = [1 - \theta_{c+1}]^{a+1}[1 - \theta_c]^{N-a-1}, \tag{12}$$

where $\theta_0 = 0$ and

$$\theta_n = [R_0 + \cdots + R_{n-1}]/E[X], \qquad n \geq 1. \tag{13}$$

Proof: The tail probability $P(W > t)$ is the product of N probabilities $w^{(j)}(t)$ that the first delivery to processor 0 from cell j happens after step t. Deliveries to processor 0 from cell j form a renewal process having a distribution $\{r_c\}$ for the number c of cycles between renewals (deliveries). The probability that cell j first delivers to processor 0 on its c^{th} visit is $R_{c-1}/E[X]$, where $R_0 = 1$ and $R_c = 1 - r_1 - \cdots - r_c$, $c \geq 1$, as before. If cell j visits processor 0 c times in $[0, t]$, then

$$w^{(j)}(t) = 1 - [R_0 + \cdots + R_{c-1}]/E[X] = 1 - \theta_c.$$

As in Theorem 3, cells $j = 0, \ldots, a$ make $c + 1$ visits in time $t = cN + a$ and the remaining $N - a - 1$ cells make c visits. The desired probability, $\prod_{j=0}^{N-1} w^{(j)}(t)$, is (12). $\square$.

Perhaps the most interesting special case has transit times with the uniform distribution $g_k = 1/(N - 1)$, $k = 1, \ldots, N - 1$. Table 1 gives numerical values of q_t, calculated from (8) for this case.

Table 1: Distribution of times between deliveries at one processor. Transit times are uniform.

t/N	q_t $N=5$	10	25	geometric approx ($N=25$)
0.2	.4000	.1280	.0573	.0573
0.4	.2400	.1024	.0378	.0378
0.6	.1440	.0655	.0249	.0249
0.8	.0864	.0419	.0164	.0164
1.0	.04670	.0236	.00937	.01081
1.2	.03535	.0181	.00726	.00713
1.4	.02028	.0114	.00476	.00470
1.6	.01163	.00719	.00312	.00310
1.8	.00667	.00453	.00205	.00204
2.0	.00378	.00280	.00131	.00134
2.2	.00223	.00182	.884e-3	.886e-3
2.4	.00127	.00114	.579e-3	.584e-3
2.6	.725e-3	.718e-3	.379e-3	.385e-3
2.8	.413e-3	.451e-3	.249e-3	.254e-3
3.0	.236e-3	.284e-3	.163e-3	.167e-3
3.4	.763e-4	.112e-3	.699e-4	.727e-4
3.8	.248e-4	.443e-4	.300e-4	.316e-4
4.2	.805e-5	.175e-4	.129e-4	.137e-4
4.6	.262e-5	.693e-5	.553e-5	.596e-5

The tabulated variable t/N is the time to the next delivery in units of complete cycles. For $N = 25$, Table 1 also compares q_t with the geometric distribution $(2/N)(1-2/N)^{t-1}$. Since $E[X] = N/2$, this geometric distribution is exactly q_t in (8) when $t < N$. It is a good approximation for the first few cycles and becomes increasingly accurate as $N \to \infty$, as the next section will show.

4. Limit Laws

This section derives limit laws as $N \to \infty$ for rings with X distributed uniformly on $\{1, \ldots, N-1\}$. In units of the mean transit time $N/2$, C and W become exponentially distributed, a fact anticipated by Table 1, where q_t appears to become exponential.

Similar results hold for geometrically-distributed transit times, with mean $E[X] = AN$ for constant A. In this case, a processor just sees deliveries arriving independently at rate $1/(AN)$. Then $P(\frac{S}{AN} > x)$, $P(\frac{C}{AN} > x)$, and $P(\frac{W}{AN} > x)$ are all $[1 - \frac{1}{AN}]^{ANx} \to e^{-x}$, as $N \to \infty$. But the geometric case has limited practical interest because it allows a packet to travel more than once around the ring.

Theorem 5. *For X uniform on $\{1, \ldots, N-1\}$,*

$$\lim_{N\to\infty} P\left(\frac{C}{N/2} > x\right) = e^{-x}, \quad x \geq 0.$$

so on substituting into (14), one can conclude

$$\lim_{N\to\infty} E\left[e^{\frac{tC}{N/2}}\right] = \frac{1}{1-t},\ t<0.$$

But $1/(1-t)$ is the Laplace transform of the exponential distribution $1-e^{-x},\ x \geq 0$, so the Lévy continuity theorem [5, p. 191] completes the proof of the limit law for C. □

In general, an exponential approximation $P(\frac{W}{E[X]} > x) \approx e^{-x}$ to the waiting time distribution can be expected whenever N is large and each individual cell serves a given processor only rarely. Feller [4, pp. 355–356] outlined an argument for a similar result, but some care is needed to justify the approximations. The next theorem supplies these details when X is uniformly distributed.

Theorem 6. *For X uniformly distributed on* $\{1, \ldots, N-1\}$,

$$\lim_{N\to\infty} P\left(\frac{W}{N/2} > x\right) = e^{-x}, \quad x \geq 0.$$

Proof: The argument examines $P(W > t)$ with $t = cN + a = xN/2$, and hence $c = \lfloor \frac{x}{2} \rfloor$. Let $\alpha = \frac{x}{2} - \lfloor \frac{x}{2} \rfloor$, so that $\alpha = \frac{a}{N}$. Since $1 - R_c = r_1 + \cdots + r_c = P(C \leq c)$, Theorem 5 shows that $R_c \to 1$ for any fixed c. From (13),

$$\theta_c = \frac{2}{N}[c + o(1)] = \frac{2}{N}\left[\left\lfloor \frac{x}{2} \right\rfloor + o(1)\right]$$

and (12) becomes, as $N \to \infty$,

$$\begin{aligned} P\left(\frac{W}{N/2} > x\right) &\sim \left[1 - \frac{2}{N}\left(\left\lfloor \frac{x}{2} \right\rfloor + 1\right)\right]^{\alpha N+1} \left[1 - \frac{2}{N}\left\lfloor \frac{x}{2} \right\rfloor\right]^{(1-\alpha)N-1} \\ &\sim e^{-2\alpha(\lfloor x/2 \rfloor + 1)} e^{-2(1-\alpha)\lfloor x/2 \rfloor} \\ &\sim e^{-2(\lfloor x/2 \rfloor + \alpha)} = e^{-x}. \end{aligned}$$

□

References

[1] E. G. Coffman, Jr., E. N. Gilbert, A. G. Greenberg, F. T. Leighton, Philippe Robert, and A. L. Stolyar. Queues served by a rotating ring. *Comm. Statist.–Stochastic Models*, **11**(3), 371–394 (1995).

[2] E. G. Coffman, Jr., Nabil Kahale, and F. T. Leighton. Processor ring communication: A tight asymptotic bound on packet waiting times. *SIAM J. Comput.* (to appear).

[3] W. Feller. *An Introduction to Probability Theory and Its Applications, Vol. I*, 3rd Edition. Wiley & Sons, New York, 1968.

[4] W. Feller. *An Introduction to Probability Theory and Its Applications, Vol. II.* Wiley & Sons, New York, 1966.

[5] M. Loève. *Probability Theory.* Van Nostrand, New York , Third Edition, 1963.

[6] F. Spitzer. *Principles of Random Walk.* Van Nostrand, New York, 1964.