V. Carl Hamacher
Gerald S. Shedler

Collision-Free Local Area Bus Network Performance Analysis*

This paper deals with port access control for local area computer communication bus networks. Emphasizing properties of the algorithms and delay-throughput performance, we focus on two collision-free access control schemes recently proposed by Eswaran, Hamacher, and Shedler. We also provide a comparison of these schemes to other available bus access techniques. The performance analysis is based on representation of the bus network as a closed queueing system with nonpreemptive priority service.

1. Introduction

There is a great deal of current interest in the development of local area networks of small to medium size computers as alternatives to shared mainframes. The trend has been noted by Liebowitz [1, 2]. Variable-length program and data file transfers among the computers on such a network may require communication path bandwidths of a few megabits per second to provide acceptable system response. Bit-serial transmission facilities such as coaxial cables, twisted pairs of wires, or optical fibers are used for the geographically limited environments of business offices, hospitals, or manufacturing plant sites.

Decentralized local area networks are usually configured in a ring or bus topology as described by Clark *et al.* [3] and by Chlamtac *et al.* [4]. This paper focuses on port access control for the bus topology. In particular, we describe two access control schemes (denoted A1 and A2) recently proposed by Eswaran, Hamacher, and Shedler [5]. The paper emphasizes properties of the algorithms and delay-throughput performance of the schemes. We also provide some comparisons to other available bus access techniques [6–10].

2. Bus access control

Let N be the number of ports (numbered $1, 2, \cdots, N$) on the bus network, as shown in Fig. 1. Message packet traffic on the passive bilateral bus is transmitted/received by port J at bus tap $\mathbf{B}(J)$. In addition to the bus, a one-way logic *control wire* also links the ports. Associated with each port J is a flip-flop, $\mathbf{S}(J)$, called the *send flip-flop*. The signal $\mathbf{P}(J)$, called the OR-*signal*, tapped at the control wire input to port J is the inclusive OR of the send flip-flops of all ports to the left of port J.

The bus and control wire may be of the order of a kilometer in length. Because of this, care must be taken in considering how propagation delays affect access control. Our notation for propagation delays is as follows: T denotes end-to-end bus propagation delay, and $R(J)$ denotes delay from port J to port N on the control wire. For technical reasons, we actually must take T to be the end-to-end bus propagation delay plus a small (fixed) quantity. Signal propagation delay along the control wire includes gate delays, and we assume that propagation delay along the control wire is larger than along the bus.

The control scheme A1 given below, implemented in the port interface logic of each port, achieves collision-free communication among ports. Control scheme A2 also achieves collision-free communication, and in addition provides a bounded, guaranteed time to transmission for each port. Thus, control scheme A2 ensures that a packet which becomes available to a port for transmission will be transmitted and that transmission will begin within a bounded amount of time. The proofs given in [5] of the

*A portion of this paper has been presented at the National Telecommunications Conference, Houston, Texas, November 30–December 4, 1980.

properties of control schemes A1 and A2 require no assumptions regarding the mechanism by which packets arrive and become available to the individual ports for transmission. In particular, any port may have a next packet available for transmission immediately after the end of transmission of a current packet.

Specification of distributed control scheme A1 is in terms of an algorithm for an individual port J. We assume that packets for transmission by port J, which arrive while an execution of the algorithm by port J is in progress, queue externally. Upon completion of this execution of the algorithm, one of any such packets immediately becomes available to port J for transmission and the next execution of the algorithm begins.

Algorithm A1

- Set **S**(J) to 1.
- Wait for a time interval $R(J) + T$.
- Wait until the bus is observed (by port J) to be idle AND **P**(J) = 0; then begin transmission of the packet, simultaneously resetting **S**(J) to 0.

We shall see in Section 3 that, although control scheme A1 provides collision-free communication among the ports of a bus network, in general it does not guarantee transmission access for all ports. The second control scheme, A2, does provide a bounded, guaranteed time to transmission for all ports. Control scheme A2 is obtained from scheme A1 by adding an initial wait for the bus to be idle throughout a time interval of length $2T$.

Algorithm A2

- Wait until the bus is observed (by port J) to be idle throughout a time interval of length $2T$.
- Set **S**(J) to 1.
- Wait for a time interval $R(J) + T$.
- Wait until the bus is observed (by port J) to be idle AND **P**(J) = 0; then begin transmission of the packet, simultaneously resetting **S**(J) to 0.

If it is desirable for all ports to execute exactly the same algorithm, then $R(J)$ can be replaced by $R(1)$ in either A1 or A2 without altering the essential properties of control.

To illustrate the manner in which events occur when several ports execute Algorithm A2 asynchronously with respect to each other, we use time line diagrams as in Fig. 2. This figure pertains to a network with five ports that are uniformly spaced along the bus. Each horizontal line is a time axis (or time line) for displaying events at a particular port. Time increases to the right from an assumed 0 origin, and the unit of time is chosen to be T. The vertical axis represents distance along the bus,

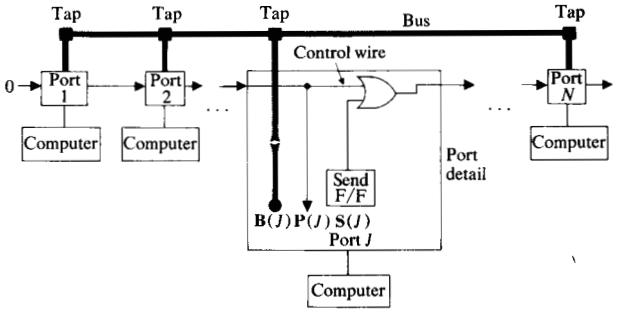

Figure 1 Bus network and ports.

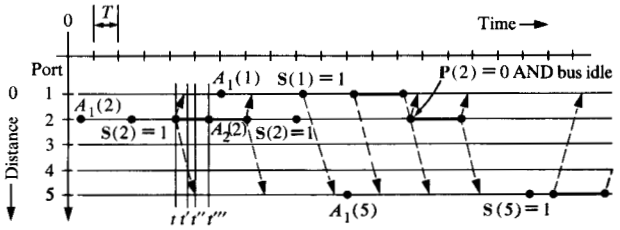

Figure 2 Time line diagram for five-port network.

directed from left to right. For this five-port network, the distance between each of the ports is 1/4 of the bus length. The slanted, dashed arrows which intersect the port time lines connect the times at which the spatially distributed ports observe the occurrence of a particular event (*e.g.*, start of transmission, setting of a send flip-flop, etc.) at some remote port. The slopes of these arrows are consistent with the horizontal time scale and vertical distance scale. A dashed arrow emanates from an event, and its effect propagates towards the head of the arrow. For example, in Fig. 2, the leftmost downward arrow indicates that the start (at time t) of transmission by port 2 of its first packet is observed by port 4 at time t' and by port 5 at time t''. While port 2 is transmitting its first packet, a second packet arrives at time t'''. This packet becomes available to port 2 at the end of transmission of its first packet, *i.e.*, at time $A_2(2)$. [We denote by $A_m(J)$ the time at which the mth packet becomes available to port J for transmission.] Immediately after the completion of transmission of its first packet, port 2 starts the execution of Algorithm A2 again, setting **S**(2) to 1 after observing the bus to be idle throughout an interval of length $2T$. After waiting for a time interval $R(2) + T$, port 2 observes **P**(2) = 1. Therefore, port 2 cannot begin transmission of its second packet, but must wait until it observes that port 1 has ended transmission and that **P**(2) equals 0. Port 2 begins transmission of its second packet at the time labeled "**P**(2) = 0 AND bus idle."

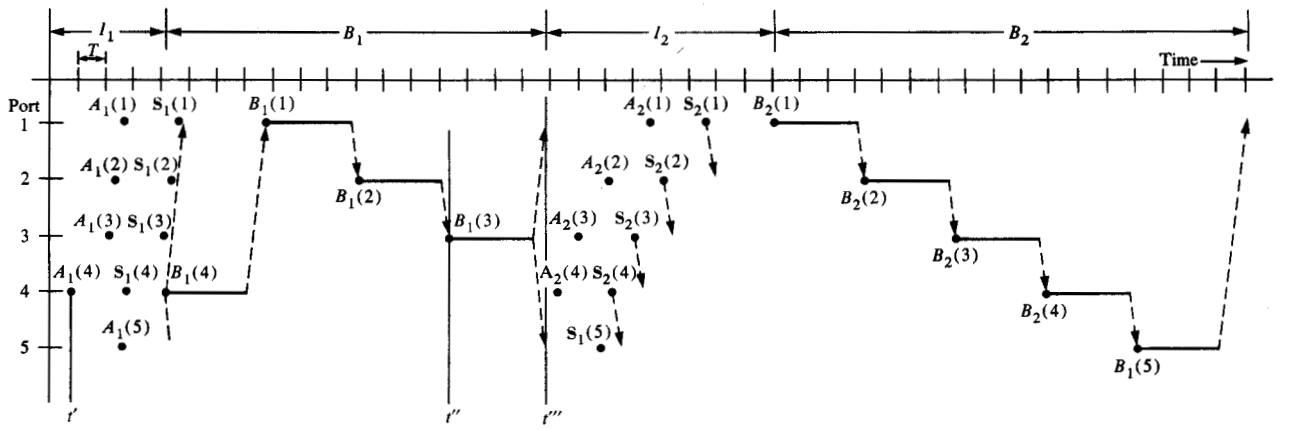

Figure 3 Bus cycles for five-port network.

3. Properties of control schemes A1 and A2

In this section we indicate several properties of control schemes A1 and A2. Formal arguments used to establish these properties are in [5]. The strength of the control schemes lies in these properties and the simplicity of the schemes.

- *Collision avoidance*

We show that both A1 and A2 are collision-free by assuming that transmission from two ports (I and J) collide, and then deducing a contradiction. Without loss of generality, suppose that $I < J$. Denote by $B_m(I)$ [resp. $B_n(J)$] the time at which port I [resp. port J] begins transmission of its mth [resp. nth] packet, and assume that these transmissions collide. There are two cases: either $B_m(I) > B_n(J) + T$ or $B_m(I) \leq B_n(J) + T$. If $B_m(I)$ occurs after $B_n(J) + T$, port I would have observed the bus to be busy at time $B_m(I)$ and therefore would not have begun transmission. In the case that $B_m(I) \leq B_n(J) + T$, port J would have observed $\mathbf{P}(J) = 1$ at time $B_n(J)$. This is because port I must have set $\mathbf{S}(I)$ to 1 for transmission of its mth packet no later than time $t = B_m(I) - \{R(I) + T\}$. Therefore, port J would not have begun transmission. Thus we have

Proposition 1 A bus network with distributed control A1 or A2 is collision-free.

- *Guaranteed time to transmission under A2*

Although control scheme A1 is collision-free, in general it does not provide a guaranteed time to transmission for all ports. It is easy to see this by considering a three-port network in which a next packet is available to each port immediately after it has finished transmitting a current packet. In this case, port 3 never observes the bus to be idle and $\mathbf{P}(3)$ to be 0 simultaneously, and thus never transmits. Ports 1 and 2 will alternate use of the bus.

To illustrate that the initial wait for the bus to be idle throughout a time interval of length $2T$ in control scheme A2 ensures transmission by all ports, again consider the bus network with three ports. Assume that port 2 is equidistant from ports 1 and 3 and that all ports have set their flip-flops. Suppose port 1 transmits first. When transmission of the initial packet by port 1 ends at, say, time t, port 2 has set $\mathbf{S}(2)$ and has waited for a time interval of length $R(2) + T$, and port 3 has set $\mathbf{S}(3)$ and has waited for a time interval of length $R(3) + T$. Then at time $t + T/2$, port 2 observes the bus to be idle and $\mathbf{P}(2) = 0$. Therefore, it begins to transmit. Note that port 1 observes the bus to be idle throughout the time interval $[t, t + T]$; but then it observes the transmission by port 2, and consequently cannot yet set $\mathbf{S}(1)$ to 1 preparatory to transmitting its second packet. Port 3 observes the bus to be idle and $\mathbf{P}(3) = 0$ at time $T/2$ following the end of transmission by port 2 and begins transmission. When each of the ports observes the bus to be idle (for a time interval of length $2T$) after port 3 ends transmission, it sets its send flip-flop, and subsequent packet transmissions from ports 1, 2, and 3 proceed cyclically. This discussion assumes packet transmission time is longer than $R(1) + T$, which is not restrictive in practice.

The analysis needed to obtain the best (least upper) bound for the guaranteed time to transmission is somewhat tedious. In order to suggest how long transmission of an available packet can be delayed, we consider a specific five-port network and suppose that the time to transmit any packet is P, where $P > R(1) + T$.

Figure 3 shows a series of events which give rise to a worst-case time to transmission for port 5. We assume that at time $A_1(5)$ a first packet becomes available to port 5 for transmission. Although port 5 observes the bus to be idle at this time, port 4 begins transmission of a packet at time $B_1(4)$, and the transmission is observed by port 5 just before $A_1(5) + 2T$, so that port 5 is unable to set S(5) to 1. Successively, ports 4, 1, 2, and 3 transmit packets. At time $t + 2T = S_1(5)$, port 5 has observed the bus to be idle throughout a time interval of length $2T$ and now sets S(5). It then waits for a time interval of length $R(5) + T = T$, at which time it observes $P(5) = 1$, because port 4 has set S(4) to 1. Port 4 cannot begin transmission after setting S(4) and waiting for a time interval of length $R(4) + T$ [since it observes $P(4) = 1$ because port 3 has set S(3), etc.]. Eventually, ports 1, 2, 3, and 4 each transmit a packet before port 5 observes the bus to be idle and $P(5) = 0$. Finally, port 5 begins transmission. Note that with this series of events, all ports other than port 5 transmit two packets after time $A_1(5)$ before port 5 begins transmission of its packet.

The example of Fig. 3 generalizes easily to the N-port case. A detailed analysis of the general situation establishes

Proposition 2 If the time to transmit a packet is P, the guaranteed time to transmission for port J in a bus network with control scheme A2 is bounded above by

$$(J + 6)T + T(1, J) + R(J) + (N + J - 1)P + \sum_{i=2}^{J} R(i),$$

where $T(1, J)$ is the propagation delay along the bus from port 1 to port J.

If NT and $NR(1)$ are very small relative to P (which would be the case in many practical situations), then the above bound is approximately NP for port 1 and $(2N - 1)P$ for port N.

● *Delay-throughput performance under A1*

The fact that control scheme A1 does not provide a guaranteed time to transmission leads us to a consideration of the actual delay-throughput characteristics of A1. In an earlier paper [11], we modeled a bus network with control scheme A1 as a closed queueing system with nonpreemptive priority service. Our analysis of the bus network emphasizes throughput and delay experienced by individual ports.

Definition of the model

The closed queueing system (see Fig. 4) provides service to N stochastically nonidentical *jobs* (ports) labeled $1, 2, \cdots, N$. The queueing system comprises $N + 1$ single server *service centers* (denoted $0, 1, \cdots, N$) which can provide service concurrently. Center 0 (bus) provides exponentially distributed nonpreemptive service (packet transmission) of duration W_0 to each of the N jobs. Center J ($J = 1, 2, \cdots, N$) provides service only to job J. Upon completion of service to job J at center 0, with probability p_J ($0 \leq p_J < 1$) job J joins the queue at center 0, and with probability $1 - p_J$ moves to center J where (with no queueing delay) service begins. Service at center 0 is according to a fixed priority ordering of jobs, lower labeled jobs having higher priority. With respect to this priority scheme, a job joining a nonempty queue at center 0 upon completion of service at center 0 does not contend for the next center 0 service. Exponentially distributed service of duration W_J at center J ($J > 0$) represents a time delay associated with the availability to port J of a next packet for transmission. The interpretation of the routing from center 0 is that with probability p_J, upon completion of packet transmission service, a next packet immediately becomes available to port J for transmission.

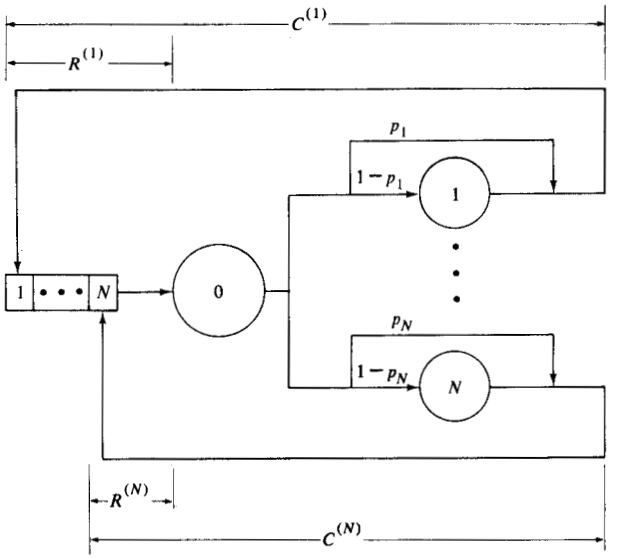

Figure 4 Queueing model for N-port network. Center 0 provides nonpreemptive packet transmission service. Service at center 0 is according to priority ordering of ports. Service at each of centers 1 through N represents a time delay at its associated port.

Note that although the model does not explicitly incorporate the quantities $R(J) + T$, the priority rule effectively preserves the order of port access to the bus in the network. There is, however, a situation in which the bus access assignment in the model is different from that in the network. Suppose that upon completion of a packet transmission, two or more ports have packets available.

According to the priority rule in the model, the leftmost port, L, among them gains access to the bus. In the network, however, port L may not gain access to the bus if it has waited for a time interval less than $R(L) + T$. This occurs infrequently because the time intervals $R(J) + T$ are very much smaller than other time intervals (*e.g.*, packet transmission times) in the network.

We make the following *probabilistic assumptions*:

1. Successive service times at center 0 [resp. $1, 2, \cdots, N$] form a sequence of independent random variables, identically and exponentially distributed as W_0 [resp. $W_1, \cdots, W_N$];
2. Service times at the centers are independent of the routing of jobs from center 0;
3. The sequences in (1) are mutually independent.

Specification of response times

Denote the increasing sequence of service completion times (irrespective of center identity) by $\{\tau_n : n \geq 0\}$. For $J = 1, 2, \cdots, N$ and $t \geq 0$, let

$$Z_J(t) = \begin{cases} 2 & \text{if job } J \text{ is in service at center 0,} \\ 1 & \text{if job } J \text{ is queued for service at center 0,} \\ 0 & \text{if job } J \text{ is in service at center } J, \end{cases}$$

at time t and for $n \geq 0$ set

$$Z_n = (Z_1(\tau_n), \cdots, Z_N(\tau_n)).$$

It is easy to check that the process $\{(Z_n, \tau_n) : n \geq 0\}$ is an irreducible Markov renewal process with finite state space, E. This means (*cf.*, *e.g.*, Çinlar [12], p. 313) that $\{(Z_n, \tau_n) : n \geq 0\}$ satisfies

$$P\{Z_{n+1} = z, \tau_{n+1} - \tau_n \leq t \mid Z_0, \cdots, Z_n; \tau_0, \cdots, \tau_n\} = P\{Z_{n+1} = z, \tau_{n+1} - \tau_n \leq t \mid Z_n\}$$

with probability one for all $n \geq 0$, $z \in E$, and $t \geq 0$. Moreover, the discrete time Markov chain $\{Z_n : n \geq 0\}$ is irreducible.

It follows that the continuous time process

$$\mathbf{X} = \{X(t) : t \geq 0\}$$

defined by

$$X(t) = Z_n \quad \text{if } \tau_n \leq t < \tau_{n+1}$$

is an irreducible semi-Markov process with state space E. The holding times in $\mathbf{X}$ are exponentially distributed (with parameters depending only on the current state), and the process is a continuous time Markov chain when $p_J = 0$, $J = 1, 2, \cdots, N$. If $p_J > 0$ for at least one J, however, jumps in $\mathbf{X}$ from a state to the same state can occur. For example, jumps from $i = (2, 0, \cdots, 0)$ to i occur if $p_1 > 0$.

We denote the one-step transition matrix of the embedded (discrete time) Markov chain $\{X(\tau_n) : n \geq 0\}$ by $\mathbf{R} = (r_{ij})$, and let $\mathbf{q}$ be the vector of (rate) parameters for the exponentially distributed unconditional holding times in $\mathbf{X}$.

The response times for job J are specified in terms of four subsets of the set E : $A_1^{(J)}$, $A_2^{(J)}$, $B_1^{(J)}$, and $B_2^{(J)}$, $J = 1, 2, \cdots, N$. The sets $A_1^{(J)}$ and $A_2^{(J)}$ [resp. $B_1^{(J)}$ and $B_2^{(J)}$] determine when to start [resp. terminate] the clock measuring response times for job J. These subsets of E are

$$A_1^{(J)} = \{(x_1, \cdots, x_N) \in E : x_J \neq 1\},$$

$$A_2^{(J)} = \{(x_1, \cdots, x_N) \in E : x_J \neq 0\},$$

$$B_1^{(J)} = \{(x_1, \cdots, x_N) \in E : x_1 = \cdots = x_N = 0 \text{ or } x_J \neq 0\},$$

and

$$B_2^{(J)} = \{(x_1, \cdots, x_N) \in E : x_J = 2\}.$$

For $n \geq 1$, denote the start [resp. termination] time of the nth response time for job J by $S_{n-1}^{(J)}$ [resp. $T_n^{(J)}$]. Setting $T_0^{(J)} = 0$ for $J = 1, 2, \cdots, N$,

$$S_n^{(J)} = \inf\{\tau_k > T_n^{(J)} : X(\tau_k) \in A_2^{(J)}, X(\tau_{k-1}) \in A_1^{(J)}\}, n \geq 0$$

and

$$T_n^{(J)} = \inf\{\tau_k \geq S_{n-1}^{(J)} : X(\tau_k) \in B_2^{(J)}, X(\tau_{k-1}) \in B_1^{(J)}\}, n \geq 1.$$

Then the nth response time for job J is

$$R_n^{(J)} = T_n^{(J)} - S_{n-1}^{(J)}.$$

For $t \geq 0$ let $L(t)$ denote the last state visited by the semi-Markov process $\mathbf{X}$ before jumping to $X(t)$, and set

$$V(t) = (L(t), X(t)).$$

The process $\mathbf{V} = \{V(t) : t \geq 0\}$ has a finite state space, F, which consists of all pairs (i, j) of states of $\mathbf{X}$ for which a one-step transition from i to j can occur with positive probability. For job J, define two subsets $S^{(J)}$ and $T^{(J)}$ of F according to

$$S^{(J)} = \{(i, j) \in F : i \in A_1^{(J)}, \quad j \in A_2^{(J)}\}$$

and

$$T^{(J)} = \{(i, j) \in F : i \in B_1^{(J)}, \quad j \in B_2^{(J)}\}. \tag{1}$$

The entrances of $\mathbf{V}$ to $S^{(J)}$ [resp. $T^{(J)}$] correspond to the starts [resp. terminations] of response times for job J.

A key observation is that the process

$$\{(X(S_n^{(J)}), R_{n+1}^{(J)}) : n \geq 0\}$$

is a regenerative process (Smith [13]) in discrete time. (Heuristically, a stochastic process is regenerative if there exists a sequence of random time points at which

the process probabilistically restarts.) The regenerative property guarantees (see Miller [14]) that the sequence $\{R_n^{(J)} : n \geq 1\}$ converges in distribution to a random variable $R^{(J)}$, the limiting response time for job J.

Since $\mathbf{X}$ is an irreducible, finite state semi-Markov process, it is a regenerative process in continuous time. The regenerative structure ensures that the "steady state" of the process is determined (as a ratio of expected values) by the behavior of the process in a cycle, *i.e.*, between any two successive regeneration points. Our analysis is based on the selection of a particular sequence of regeneration points (returns to a fixed state, j_0) for $\mathbf{X}$. Entrances of the process $\mathbf{X}$ to state j_0 correspond to the starts of response times for a particular job, J_0, with a lower priority job, K_0, in service at center 0, and no other jobs in queue at center 0. For each quantity of interest, we establish a ratio formula in terms of cycles defined by the returns to state j_0. We then apply computational results of the kind developed by Hordijk, Iglehart, and Schassberger [15] for discrete time and continuous time Markov chains.

Analysis of the bus network model can be based on other sequences of regeneration points. It is, however, computationally advantageous to use cycles defined by state j_0 since only the jump matrix $\mathbf{R}$ and vector $\mathbf{q}$ of the underlying semi-Markov process $\mathbf{X}$ (rather than the corresponding quantities for the process $\mathbf{V}$) are needed. Moreover, in order to compute delay characteristics for all ports, only a single matrix inversion is required.

Analysis for response times

Select J_0 and K_0 with $1 \leq J_0 < K_0 \leq N$. Now let $j_0 = (x_1', \cdots, x_N') \in A_2^{(J_0)}$ with $x'_{J_0} = 1$, $x'_{K_0} = 2$, and $x_i' = 0$ for $i \neq J_0$, K_0, and take $X(0) = j_0$. Because of the nonpreemptive priority service discipline at center 0, $X(\tau_k) = j_0$ implies that $X(\tau_{k-1}) = i_0$, where $i_0 = (x_1 \cdots, x_N)$ with $x_{K_0} = 2$ and $x_i = 0$ for $i \neq K_0$. Thus, successive entrances of the process $\mathbf{X}$ to state j_0 correspond to starts of response times for job J_0 such that job K_0 is in service at center 0, and job i is in service at center i, $i \neq J_0$, K_0. Set $\beta_0 = 0$ and denote the time of the kth entrance of $\mathbf{X}$ to state j_0 by β_k, $k \geq 1$. Also define $\alpha_k = \beta_{k-1}$, $k \geq 1$.

Let $\{V_k : k \geq 0\}$ be the embedded jump chain of $\mathbf{V}$ and for convenience designate state $(i_0, j_0) \in F$ as state 0. Denote by $\{\gamma_k : k \geq 1\}$ the lengths in discrete time units of the successive 0-cycles (successive returns to the fixed state 0) for $\{V_k : k \geq 0\}$. (These correspond to the successive entrances of $\{X(\tau_n) : n \geq 0\}$ to state j_0 from i_0.) Now fix J. The number of response times for job J in the first 0-cycle of $\mathbf{V}$ is

$$N_1^{(J)} = \sum_{k=0}^{\gamma_1 - 1} 1_{\{V_k \in S^{(J)}\}}, \tag{2}$$

where $1_{\{V_k \in S^{(J)}\}} = 1$ if $V_k \in S^{(J)}$ and 0 otherwise, and the sum of the response times in the first 0-cycle is

$$Y_1^{(J)} = \sum_{k=1}^{N_1^{(J)}} R_k^{(J)}. \tag{3}$$

Denote the analogous quantities in the mth 0-cycle by $N_m^{(J)}$ and $Y_m^{(J)}$, $m \geq 1$. Since $\mathbf{X}$ is a regenerative process, $\mathbf{V}$ is also. It follows that the pairs of random variables $\{(N_k^{(J)}, Y_k^{(J)}) : k \geq 1\}$ are independent and identically distributed (i.i.d.). Standard arguments (*cf.* Iglehart and Shedler [16], Appendix 2) establish the ratio formula

$$E\{R^{(J)}\} = E\{Y_1^{(J)}\}/E\{N_2^{(J)}\}. \tag{4}$$

We now show how to calculate the quantities on the righthand side of Eq. (4).

Recall that $\mathbf{R}$ is the one-step transition matrix of the embedded Markov chain $\{X(\tau_n) : n \geq 0\}$ and denote by ${}_0\mathbf{R}$ the matrix obtained by setting the j_0 column of $\mathbf{R}$ equal to 0. (We assume a fixed enumeration of the states of $\mathbf{X}$ and that the j_0 column corresponds to state j_0.) We consider vectors to be column vectors, view a real-valued (measurable) function such as f having domain E in this way, and denote it by $\mathbf{f}$. In addition, $\circ$ denotes the Hadamard product of vectors; *i.e.*, for vectors $\mathbf{u} = (u_1, u_2, \cdots, u_k)$ and $\mathbf{v} = (v_1, v_2, \cdots, v_k)$, the symbol $\mathbf{u} \circ \mathbf{v}$ denotes the vector $(u_1v_1, u_2v_2, \cdots, u_kv_k)$. The component of the vector $\mathbf{u}$ corresponding to state j is denoted by $[\mathbf{u}]_j$.

Let f be a real-valued function with domain F. An argument analogous to that used to obtain Theorem (3.1) of [15] for discrete time Markov chains shows that

$$E\left\{\sum_{k=0}^{\delta_1 - 1} f(X(\tau_k), X(\tau_{k+1}))\right\} = [(\mathbf{I} - {}_0\mathbf{R})^{-1}\mathbf{g}]_{j_0}, \tag{5}$$

where $\mathbf{I}$ is the identity matrix and for $k \in E$,

$$g(k) = \sum_{m \in E} f(k, m) r_{km}.$$

We use this result to calculate the quantity $E\{N_1^{(J)}\}$. Take f to be the indicator function of the set $S^{(J)}$: for $(x_1, \cdots, x_N, x_1', \cdots, x_N') \in F$, $f(x_1, \cdots, x_N, x_1', \cdots, x_N')$ equals 1 if $(x_1, \cdots, x_N) \in A_1^{(J)}$ and $(x_1', \cdots, x_N') \in A_2^{(J)}$, and equals 0 otherwise. Then

$$N_1^{(J)} = \sum_{k=0}^{\gamma_1 - 1} f(X(\tau_k), X(\tau_{k+1}))$$

and Eq. (5) gives

$$E\{N_1^{(J)}\} = [(\mathbf{I} - {}_0\mathbf{R})^{-1}\mathbf{g}]_{j_0}, \tag{6}$$

where for $k \in E$,

$$g(k) = \begin{cases} \sum_{m \in A_2^{(J)}} r_{km} & \text{if } k \in A_1^{(J)}, \\ 0 & \text{otherwise.} \end{cases}$$

Similarly, for a real-valued f having domain E and $Y_1(f)$ defined by

$$Y_1(f) = \int_0^{\beta_1} f(X(s))ds,$$

it can be shown (*cf.* Theorem (3.10) of [15] for continuous time Markov chains) that

$$E\{Y_1(f)\} = [(\mathbf{I} - {}_0\mathbf{R})^{-1}(\mathbf{f} \circ \mathbf{q}^{-1})]_{j_0}, \tag{7}$$

We use this result to calculate the quantity $E\{Y_1^{(J)}\}$. For $(x_1, \cdots, x_N) \in E$ take f to be the function defined by

$$f(x_1, \cdots, x_N) = \begin{cases} 1 & \text{if } x_J = 1, \\ 0 & \text{otherwise,} \end{cases}$$

and observe that

$$\int_0^{\beta_1} f(X(s))ds = \sum_{k=1}^{N_1^{(J)}} R_k^{(J)}.$$

It then follows directly from Eq. (7) that

$$E\{Y_1^{(J)}\} = [(\mathbf{I} - {}_0\mathbf{R})^{-1}(\mathbf{f} \circ \mathbf{q}^{-1})]_{j_0}. \tag{8}$$

Combining Eqs. (4), (6), and (8), we obtain $E\{R^{(J)}\}$.

In much the same way, $P\{R^{(J)} = 0\}$ can be calculated. Observe that $R_n^{(J)} = 0$ if and only if $V(S_n^{(J)}) \in D^{(J)}$, where

$$D_1^{(J)} = \{(x_1, \cdots, x_N) \in E : x_J \neq 1 \text{ and for } k \neq J, \quad x_k = 0\},$$

$$D_2^{(J)} = \{(x_1, \cdots, x_N) \in E : x_J = 2 \text{ and for } k \neq J, \quad x_k = 0\},$$

and

$$D^{(J)} = \{(x_1, \cdots, x_N, x_1', \cdots, x_N') \in F : (x_1, \cdots, x_N) \in D_1^{(J)}, (x_1', \cdots, x_N') \in D_2^{(J)}\}.$$

It is easy to show (using the fact that $\{V_k : k \geq 0\}$ is an irreducible, finite-state Markov chain) that the pairs of random variables $\{(M_k^{(J)}, N_k^{(J)}) : k \geq 1\}$ are i.i.d., where $E\{N_1^{(J)}\}$ is given by Eq. (6) and

$$M_1^{(J)} = \sum_{k=0}^{\gamma_1 - 1} 1_{\{V_k \in D^{(J)}\}}.$$

Moreover,

$$P\{R^{(J)} = 0\} = E\{M_1^{(J)}\}/E\{N_1^{(J)}\}. \tag{9}$$

It follows directly from Eq. (5) that

$$E\{M_1^{(J)}\} = [(\mathbf{I} - {}_0\mathbf{R})^{-1}\mathbf{h}]_{j_0}, \tag{10}$$

where for $k \in E$,

$$h(k) = \begin{cases} \sum_{m \in D_2^{(J)}} r_{km} & \text{if } k \in D_1^{(J)}, \\ 0 & \text{otherwise.} \end{cases}$$

Analysis for expected queue length and throughput

Since $\mathbf{X}$ is a regenerative process,

$$X(t) \Rightarrow X$$

as $t \to \infty$, where $\Rightarrow$ denotes convergence in distribution. (The random variable X is the "steady state" of the regenerative process.) Define a function c having domain E and range $\{0, 1, \cdots, N\}$ according to

$$c(x_1, \cdots, x_N) = \sum_{j=1}^{N} 1_{\{x_j \neq 0\}}$$

for $(x_1, \cdots, x_N) \in E$. For $t \geq 0$, $c(X(t))$ is the number of jobs waiting or in service at center 0 at time t, and Q_0, the "steady state" expected queue length at center 0, is the quantity $E\{c(X)\}$.

Properties of regenerative processes (*cf.* Crane and Iglehart [17]) ensure that the pairs of random variables $\{(Y_k(c), \alpha_k) : k \geq 1\}$ are i.i.d. and that

$$Q_0 = E\{Y_1(c)\}/E\{\alpha_1\}. \tag{11}$$

Equation (7) implies that

$$E\{Y_1(c)\} = [(\mathbf{I} - {}_0\mathbf{R})^{-1}(\mathbf{c} \circ \mathbf{q}^{-1}]_{j_0} \tag{12}$$

and

$$E\{\alpha_1\} = [(\mathbf{I} - {}_0\mathbf{R})^{-1}(\mathbf{1} \circ \mathbf{q}^{-1})]_{j_0}, \tag{13}$$

where 1 is the function identically equal to one.

We define the "steady state" throughput, U, of the bus to be the limiting probability that the bus is busy; *i.e.*, U is the quantity $E\{b(X)\}$, where

$$b(x_1, \cdots, x_N) = \begin{cases} 0 & \text{if } x_1 = \cdots = x_N = 0, \\ 1 & \text{otherwise.} \end{cases}$$

[For $t \geq 0$, $b(X(t))$ equals 1 if there is a job in service at center 0 at time t, and equals 0 otherwise.] Since $\mathbf{X}$ is a regenerative process, the pairs of random variables $\{Y_k(b), \alpha_k) : k \geq 1\}$ are i.i.d. and

$$U = E\{Y_1(b)\}/E\{\alpha_1\}. \tag{14}$$

It follows directly from Eq. (7) that

$$E\{Y_1(b)\} = [(\mathbf{I} - {}_0\mathbf{R})^{-1}(\mathbf{b} \circ \mathbf{q}^{-1})]_{j_0}, \tag{15}$$

and $E\{\alpha_1\}$ is given by Eq. (13).

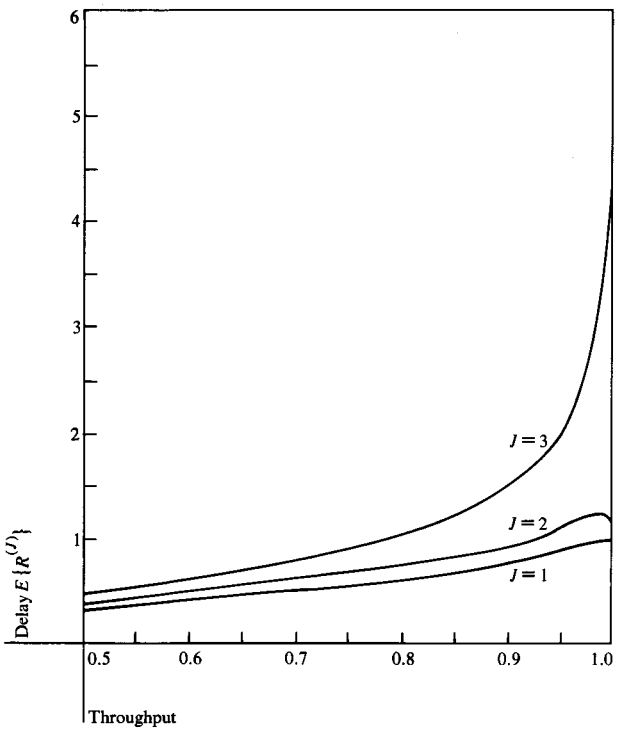

Figure 5 Delay-throughput characteristics for three-port network.

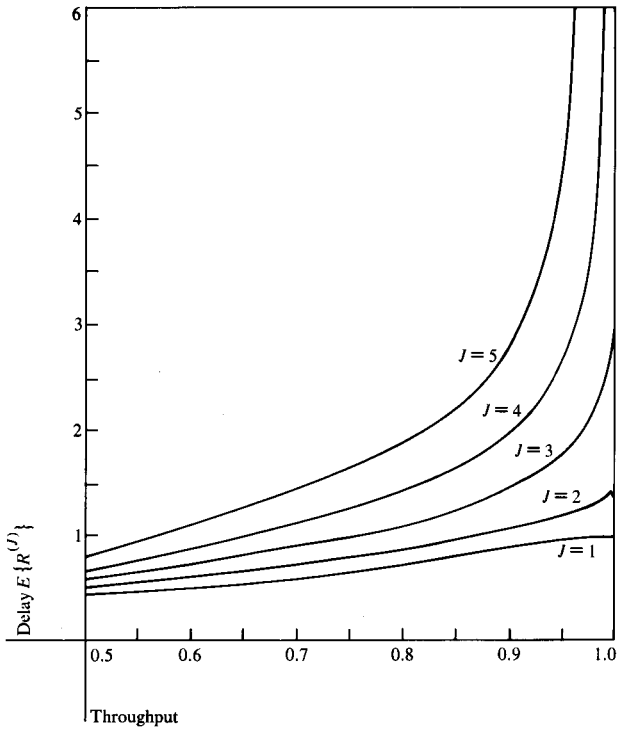

Figure 6 Delay-throughput characteristics for five-port network.

Figure 5 shows typical results for a three-port network. The abscissa is throughput, and the ordinate is delay, measured by $E\{R^{(J)}\}$ in normalized units of $E\{W_0\} = 1$. To generate these particular curves, calculations were made for various values of $E\{W_1\} = E\{W_2\} = E\{W_3\}$, with $p_1 = p_2 = p_3 = 0$. For example, throughput is 0.47 when $E\{W_1\} = E\{W_2\} = E\{W_3\} = 5.0$; throughput is 0.79 when $E\{W_1\} = E\{W_2\} = E\{W_3\} = 2.0$.

Consideration of the delay-throughput curves for the three-port network leads to the following conclusions. As throughput increases above 0.8, $E\{R^{(3)}\}$ begins to increase very rapidly, while $E\{R^{(1)}\}$ and $E\{R^{(2)}\}$ remain near 1. In the limiting case of throughput equal to 1, corresponding to $E\{W_1\} = E\{W_2\} = E\{W_3\} = 0$, it is easy to argue directly (*cf.* the example in Section 3) that $E\{R^{(1)}\} = E\{R^{(2)}\} = 1$ and port 3 does not gain access to the bus ($E\{R^{(3)}\} = \infty$). In this situation, transmissions by ports 1 and 2 alternate. Note that $E\{R^{(2)}\}$ actually attains values larger than 1 when throughput is close to, but less than, 1; see [11] for a discussion of this unintuitive phenomenon.

In an N-port network operating under control scheme A1, ports 1 and 2 experience response times that are qualitatively similar to the response times they experience in a three-port network, including the limiting case (throughput = 1) in which transmissions by the two ports alternate. Ports 3 through N behave like port 3 of the three-port case, with $E\{R^{(3)}\} < E\{R^{(4)}\} < \cdots < E\{R^{(N)}\}$ over the full range of throughput achievable when $E\{W_1\} = E\{W_2\} = \cdots = E\{W_N\}$. For throughput equal to 1, all ports $J \geq 3$ are denied access to the bus. Figure 6 shows results for a five-port network. Calculations were made for various values of $E\{W_1\} = E\{W_2\} = \cdots = E\{W_5\}$ with $p_1 = p_2 = \cdots = p_5 = 0$.

Note that our definition of throughput pertains to the data bus being busy. This means that throughput can be 1 when ports 1 and 2 alternate use of the bus, keeping it busy all of the time. The control wire is also part of the total communication facility; and since it does not carry packets, it is clear that the total facility is never fully utilized for data transmission in the conventional communication system sense.

Simulation for response times

The analysis of the previous section yields an assessment of the performance of control scheme A1 in terms of the expected response times for individual ports. It is also of interest (in particular for comparison with the guaranteed transmission time provided by control scheme A2) to study the variability of port response times. This can be

done by discrete event simulation of the model, *e.g.*, in terms of percentiles of the limiting response time distributions.

The stochastic setting for response times we have developed provides a basis for simulation of the bus network model (*cf.* Iglehart and Shedler [16], Section 10). Let f be a real-valued (measurable) function with domain $[0, +\infty)$. We assume throughout that $P\{R^{(J)} \in D(f)\} = 0$, where $D(f)$ is the set of discontinuities of the function f. Then the goal of the simulation is estimation of

$$r^{(J)}(f) = E\{f(R^{(J)})\}.$$

For example, to estimate the percentile $P\{R^{(J)} \leq x\}$ (x fixed), take $f(t) = 1_{[0,x]}(t)$, where $1_{[0,x]}(t)$ equals 1 if $t \leq x$ and equals 0 otherwise. Recall that $R^{(J)}$ is the limiting response time for job J.

Although the regenerative method of Crane and Iglehart [17] cannot be applied directly, point estimates and confidence intervals for $r^{(J)}(f)$ can be obtained from a single simulation run according to the following procedure.

Algorithm R: response time simulation

1. Select J_0 and K_0 with $1 \leq J_0 \leq K_0 \leq N$. Now let $j_0 = (x_1, \cdots, x_N)$ with $x_{J_0} = 1$, $x_{K_0} = 2$, and $x_i = 0$ for $i \neq J_0, K_0$. Begin the simulation with $X(0) = j_0$.
2. Carry out the simulation of **X** for a fixed number, n, of cycles (having random length) defined by the successive entrances of **X** to the state j_0.
3. In each cycle measure all the response times for job J.
4. For $k \geq 1$, denote the number of response times for job J observed in the kth cycle by $N_k^{(J)}$, and compute the sum $Y_k^{(J)}(f)$ of the quantities $f(R_m^{(J)})$ for response times $R_m^{(J)}$ in the kth cycle.
5. Take as a point estimate (based on n cycles) for $r^{(J)}(f)$ the quantity

   $$\hat{r}_n^{(J)}(f) = \bar{Y}_n^{(J)}(f)/\bar{N}_n^{(J)},$$

   where $\bar{Y}_n^{(J)}(f)$ and $\bar{N}_n^{(J)}$ are sample means over the cycles.
6. Take as a $100(1 - 2\gamma)\%$ confidence interval (based on n cycles) for $r^{(J)}(f)$ the interval

   $$\hat{r}_n^{(J)}(f) \pm z_{1-\gamma} s_n/(\bar{N}_n^{(J)} n^{\frac{1}{2}}).$$

   Here $z_{1-\gamma} = \Phi^{-1}(1 - \gamma)$, where $\Phi(\cdot)$ is the distribution function of a standardized (mean zero, variance one) normal random variable. The quantity s_n is

   $$s_n = [s_{11} - 2\hat{r}_n^{(J)}(f)s_{12} + \{\hat{r}_n^{(J)}(f)\}^2 s_{22}]^{\frac{1}{2}},$$

   where s_{11}, s_{22}, and s_{12} are the usual unbiased estimates for var $\{Y_k^{(J)}(f)\}$, var $\{N_k^{(J)}\}$, and cov $\{Y_k^{(J)}(f), N_k^{(J)}\}$, respectively.

This estimation procedure rests on the observation that the pairs of random variables $\{N_k^{(J)}, Y_k^{(J)}(f)): k \geq 1\}$ are i.i.d. Moreover, provided that $E\{|f(R^{(J)})|\} < \infty$,

$$E\{f(R^{(J)})\} = E\{Y_1^{(J)}(f)\}/E\{N_1^{(J)}\}.$$

Confidence intervals for $r^{(J)}(f)$ are based on the central limit theorem

$$n^{\frac{1}{2}}[\hat{r}_n^{(J)}(f) - r^{(J)}(f)]/[\sigma/E\{N_1^{(J)}\}] \Rightarrow N(0, 1),$$

where σ^2 is the variance of $Y_1^{(J)}(f) - r^{(J)}(f)\, N_1^{(J)}$ and $N(0, 1)$ is a standardized normal random variable.

4. Comparisons to other control schemes

In this section we briefly discuss five access control schemes for local area bus networks that have been developed recently by others. All of the schemes assume bit-serial transmission in the megabit/second range on a passive bilateral bus. We denote the end-to-end bus propagation delay by T, and when there is fixed or maximum packet transmission time, we denote it by P.

- *Ethernet*

The Xerox Ethernet system [6] allows collisions, detects them, and adjusts retry times randomly. In more detail, a port that begins transmitting a packet after it observes the line to be idle can detect whether or not some other transmission begins to interfere with its transmitted signal. If a collision is detected, the port stops its transmission, and after a random waiting period, attempts retransmission. The parameters of the probability distributions that are used to determine the waiting periods in the individual ports are adjusted if more than one retry is necessary. The stochastic nature of the retry waiting periods and the dynamic changes to the distributions themselves are intended to achieve a reduction of collisions, especially following the end of transmission of a packet. The Ethernet control is thus asynchronous and distributed, as are A1 and A2, but it is not collision-free. Also, it is possible, but unlikely, for a port to be blocked indefinitely from transmitting a packet without collision. The control is efficient in the sense that the collision-retry strategy wastes only a small amount of usable line transmission capacity when the line is lightly loaded. Control schemes A1 and A2 also waste small amounts of usable line transmission capacity during the various waiting intervals.

- *HXDP*

In the Honeywell HXDP system [7], access control is distributed and requires more hardware than Ethernet access control; but HXDP is collision-free and provides bounded, guaranteed time to transmission. A coded glob-

al clock signal is used to step each port synchronously through the line access control algorithm. This signal is transmitted on the broadcast bus itself by using a special line-signal sequence that cannot be confused with packet data signals. Access control is built around a 256-bit vector stored in each port. The global clock signal steps each of the ports through its respective vector, one step for every termination of a line usage interval. There is exactly one port with a 1 at any vector address; this signifies that the port may use the line during the next usage interval. During this interval, the distinguished port transmits a packet if one is available to it, and then transmits the clock signal. If there is no packet available, the port immediately transmits the clock signal, thus effecting transfer of access control to the next port. The number of 1's in a port's vector determines its fraction of bus usage intervals during a complete sweep through the vector. Using this scheme with N ports, an N-bit access vector in each port, and a maximum packet transmission time of P, then $(N-1)P$ is the (bounded) guaranteed time to transmission.

- *BRAM*

BRAM [8] is actually a family of four related decentralized access protocols. Two of these are collision-free, and one of them, called fair BRAM, is described here. All ports must monitor the bus continuously. Whenever a message is observed on the bus, the number of the transmitting port (contained in the message) is noted in each of the other ports. Assume that a number of ports have packets available for transmission and are deferring to an ongoing transmission from port J. Each waiting port I computes the value $H(I, J) = (I - J + N) \bmod N$ and waits $TH(I, J)$ time units after it observes the end of transmission from port J. If the bus, as observed by port I, has not become busy after $TH(I, J)$ time units, then port I transmits. Since the $H(K, J)$ values are distinct for all K $(1 \leq K \leq N)$, there are no collisions. If we assume that port J does not contend for permission to transmit immediately after it has used the bus, there is bounded guaranteed time to transmission for all ports. In particular, assuming $NT << P$, the bound is approximately $(N-1)P$.

- *Spaniol proposal*

An interesting variation on the Ethernet access protocol has been proposed by Spaniol [9]. A slotted Ethernet is developed in which all ports must use a fixed-length packet slot $P >> T$. Collisions may occur if more than one port attempts to transmit in an open slot. If a collision occurs, it is detectable in the initial portion of the slot, and the remaining (major) part of the slot is used in a time-division multiplexed mode to schedule subsequent slot allocations for the colliding packets. Each of the N ports has a predetermined time position in this scheduling interval in which it signifies that it is a participant in the collision. The slots immediately following the collision slot are then claimed by these ports using a simple priority rule, and no collisions can occur until each of these ports has transmitted its packet. The slot immediately following the last of these is open and any port may now attempt to transmit. Either no port transmits, exactly one port tries and is successful, or two or more ports attempt to transmit and a collision occurs. In the latter case, the time division multiplexed arbitration referred to above is used. Note that a particular packet collides at most once, and no port is delayed longer than $2NP$ time units in achieving a successful transmission. The bound is actually dependent on the priority scheme and ranges from $(N+1)P$ to $2NP$.

- *Mark proposal*

Mark [10] has studied the use of a separate control wire for access control synchronization, and has adapted a collision-free access technique originated by Rothauser and Wild [18] to the two-path bus environment. The control path operates as a bilateral bus with bit-time intervals that are longer than the bus propagation delay T. Access is determined by bit-serial port address comparisons on the control path using address of length $\lceil \log_2 N \rceil$ bits. The port with the highest numbered address wins. There is no lockout of lower address ports because all ports voluntarily do not contend for subsequent data slots after they have transmitted until the control path goes idle. When the control path goes idle, the end of the current cycle of serving all active ports with a data slot has been reached; and all ports can again contend for access. Address bit reversal can be used on successive cycles of operation if it is desired to remove the effects of priority that are induced by address values. There is a bounded guaranteed time to transmission with or without the address reversal action. The bound on transmission access is $(N-1)P$ if address reversal is used; and, depending on port address values, the bound ranges from $(N-1)P$ to $2(N-1)P$ if address reversal is not used. Because of the flexibility in assigning port addresses, a multiple priority request system can be implemented as discussed in [19].

The first two of the above five schemes for bus access control have been implemented, and a version of the Mark proposal is currently being implemented [19]. To our knowledge, BRAM and the Spaniol proposal have not been implemented.

Concluding remarks

The control schemes A1 and A2 described in this paper are distributed, have no global clock signaling, and are

collision-free. However, in addition to the single, shared-bus communication path, they require a separate logic control wire to propagate a one-way logic signal from one end of the bus to the other. In this last respect, our schemes are in the same class as that of Mark [10, 19]. Our use of the control wire path assists in the implementation of collision-free operation, but at an expense that is potentially less than the bit-vector approach of HXDP [7] or the address/priority manipulations required by BRAM [8], Spaniol [9], and Mark [10, 19].

It should be noted that the way in which we use the control wire bears some structural resemblance to the decentralized daisy chain techniques of conventional digital bus access control methods as discussed by Thurber *et al.* [20]. However, a closer examination shows that our open-loop use of a logic control wire is much simpler than the closed-loop daisy chain. Indeed, as Vranesic [21] has recently discussed, there are some subtle timing problems involved with implementing the closed-loop chain. None of these problems exist in our situation.

Acknowledgment

The authors are greatful to Suresh Jasrasaria for some of the programmed calculations used to obtain the curves displayed in Figs. 5 and 6.

References

1. B. H. Liebowitz, "Multiple Processor Minicomputer Systems—Part 1: Design Concepts," *Computer Design*, 88–95 (October 1978).
2. B. H. Liebowitz, "Multiple Processor Minicomputer Systems—Part 2: Implementation," *Computer Design*, 121–131 (November 1978).
3. D. D. Clark, K. T. Pogran, and D. P. Reed, "An Introduction to Local Area Networks," *Proc. IEEE* **66**, 1497–1517 (November 1978).
4. I. Chlamtac, W. R. Franta, P. C. Patton, and W. Wells, "Performance Issues in Back-End Storage Networks," *Computer* **13**, 18–31 (February 1980).
5. K. P. Eswaran, V. C. Hamacher, and G. S. Shedler, "Asynchronous Collision-Free Distributed Control for Local Bus Networks," *Research Report RJ2482*, IBM Research Division, San Jose, CA, 1979. To appear in *IEEE Trans. Software Eng.* **SE-7** (November 1981).
6. R. M. Metcalfe and D. R. Boggs, "Ethernet: Distributed Packet Switching for Local Computer Networks," *Commun. ACM* **19**, 395–404 (1976).
7. E. D. Jensen, "The Honeywell Experimental Distributed Processor—An Overview," *Computer* **11**, 28–38 (January 1978).
8. I. Chlamtac, W. R. Franta, and K. D. Levin, "BRAM: The Broadcast Recognizing Access Method," *IEEE Trans. Commun.* **COM-27**, 1183–1190 (August 1979).
9. O. Spaniol, "Modelling of Local Computer Networks," *Computer Networks* **3**, 315–326 (November 1979).
10. J. W. Mark, "Distributed Scheduling Conflict-Free Multiple Access for Local Area Communication Networks," *IEEE Trans. Commun.* **COM-28**, 1968–1976 (December 1980).
11. V. C. Hamacher and G. S. Shedler, "Performance of a Collision-Free Local Bus Network Having Asynchronous Distributed Control," *Proceedings of the 7th International Symposium on Computer Architecture*, IEEE Publ. No. 80CH1494-4C, La Baule, France, 1980, pp. 80–87.
12. E. Çinlar, *Stochastic Processes*, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1975.
13. W. L. Smith, "Renewal Theory and its Ramifications," *J. Royal Statist. Soc. B* **20**, 243–302 (1956).
14. D. R. Miller, "Existence of Limits in Regenerative Processes," *Ann. Math Statist.* **43**, 1263–1282 (1972).
15. A. Hordijk, D. L. Iglehart, and R. Schassberger, "Discrete Time Methods for Simulating Continuous Time Markov Chains," *Adv. in Appl. Prob.* **8**, 772–788 (1976).
16. D. L. Iglehart and G. S. Shedler, *Regenerative Simulation of Response Times in Networks of Queues, Lecture Notes in Control and Information Sciences*, Vol. 26, Springer-Verlag, Berlin, 1980.
17. M. A. Crane and D. L. Iglehart, "Simulating Stable Stochastic Systems: III, Regenerative Processes and Discrete Event Simulation," *Oper. Res.* **23**, 33–45 (1975).
18. E. H. Rothauser and D. Wild, "MLMA: A Collision-Free Multiaccess Method," *Research Report RZ802*, IBM Research Division, Zurich, Switzerland, 1976.
19. T. D. Todd and J. W. Mark, "Waterloo Experimental Local Network (WELNET) Physical Level Design," *Conference Record 1980 National Telecommunications Conference*, IEEE Publication No. 80CH11539-6, December 1980, pp. 41.4.1–41.4.5.
20. K. J. Thurber, E. D. Jensen, L. A. Jack, L. L. Kinney, P. C. Patton, and L. C. Anderson, "A Schematic Approach to the Design of Digital Bussing Structures," *Proc. Fall Joint Computer Conf.* (AFIPS) **41**, AFIPS Press, Montvale, NJ, 1972, pp. 719–740.
21. Z. G. Vranesic, "Multivalued Signalling in Daisy Chain Bus Control," *Proceedings of the 9th International Symposium on Multiple-Valued Logic*, IEEE Publication No. 79CH1408-4C, Bath, England, 1979, pp. 14–18.

Received April 3, 1981; revised June 15, 1981

V. Carl Hamacher is located at the Departments of Electrical Engineering and Computer Science, University of Toronto, Toronto, Ontario, Canada. Gerald S. Shedler is located at the IBM Research Division Laboratory, 5600 Cottle Road, San Jose, California 95193.