Discrete Applied Mathematics 37/38 (1992) 45-64
North-Holland

Design and analysis of fault-tolerant multibus interconnection networks

Pietro Camarda

Dipartimento di Elettrotecnica ed Elettronica, Politecnico di Bari, 70125 Bari, Italy

Mario Gerla

Computer Science Department, UCLA, Los Angeles, CA 90024, USA

Received 22 june 1989
Revised 27 July 1990

Abstract

Camarda, P. and M. Gerla, Design and analysis of fault-tolerant multibus interconnection networks, Discrete Applied Mathematics 37/38 (1992) 45–64.

In this paper a new class of fault-tolerant multibus interconnection networks is presented and analyzed. Efficiency and fault tolerance have been the driving forces in the design of these structures. The most common types of faults have been explicitly considered and in particular the jabbering problem has been adequately resolved. The analysis covers the evaluation of capacity, throughput and average delay and it includes faults of one or more channels. The system is shown to be very efficient and to be able to adequately support channel and station faults.

1. Introduction

Several challenging problems arise in studying and designing high speed interconnection networks. Several thousands of processing elements can be integrated in a typical multiprocessor environment, generating a variety of traffic (data, video, etc.) with strict requirements in terms of bandwidth and permissible delay [18]. An overall bandwidth in the order of the gigabit is often the minimum requisite to support integrated traffic demands. Next to delay and bandwidth requirements, fault tolerance and reliability are the most important criteria in the design of these networks [6]. Appropriate topologies and protocols must be chosen to satisfy all mentioned constraints.

In this paper we propose and analyze a class of high speed fiber optic interconnection networks which show high efficiency and at the same time excellent fault-tolerance characteristics. These networks consist of several parallel buses to which stations are multiply connected. The proposed solution presents several attractive properties as discussed below.

1.1. Fault-tolerant properties

Two types of faults must be considered:

- Channel faults: This case may occur when a channel, for any reason (maintenance, fault, etc.) cannot be used. In this case, parallel channels which connect other pairs of stations may improve fault tolerance.
- Station faults: In networks with passive taps, as is the case here, a station fault does not affect the rest of the network except for the case when the fault consists of the so-called "jabbering mode" [19]. In this case a station gets stuck on transmit and therefore keeps the bus busy transmitting garbage and preventing other stations from access. It can happen that the jabbering station can jam one or more of its ports. To be conservative, we will assume that a station will jam all of its ports. If every station was connected to all channels, then a jabbering station may possibly jam all the channels rendering the network completely unavailable. Therefore, the multichannel arrangement can impose fault tolerance only if the station is connected to a subset of the channels.

The remedies for the two types of faults are evidently in contrast. With channel fault it is necessary to increase the number of ports at each station and to connect it to as many channels as possible. At the very least every station should be connected to all channels. In this case the network is still connected if at least one channel is working. On the other hand, station faults require that every station be connected to the minimum number of channels to minimize interference in case of jabbering. The connection scheme must find a compromise among these contrasting factors.

1.2. Bandwidth and delay efficiency

In addition to fault tolerance, the multibus network imposes bandwidth and delay performance. We have chosen, for their intrinsic robustness, networks implemented with fiber optic unidirectional bus systems (UBS). Each bus composing the multibus network can have three basic topologies. In a double bus topology (Fig. 1(a)) all stations can transmit and receive from both buses. Fasnet, tokenless and DQDB are some of the protocols studied for this topology [14,15]. A single folded bus is shown in Fig. 1(b), for which efficient protocols have been studied [1,2]. In a double folded topology (Fig. 1(c)), an outbound channel is used for sensing and transmitting, while the reception is assured by an inbound channel. Expressnet is based on this latter topology [22].

Most of the protocols studied for UBS use efficient access schemes called "attempt and defer" [20]. In these schemes, a generic backlogged station, i.e., a station with a ready packet, as soon as it detects the end of the carrier, starts the transmission

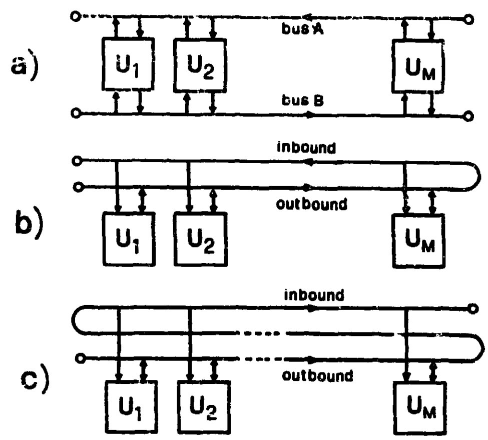

Fig. 1. Topologies for fiber optic LANs.

of its own packet, aborting it if the transmission of an upstream station is detected on the bus. Thus, only the most upstream station completes its transmission, while all other backlogged stations are interrupted and must repeat the process. In this way, a train of packets is formed on the bus. The most upstream backlogged station starts a new train at the end of the previous one with a mechanism which varies from protocol to protocol. In Fig. 2, a typical behaviour of the system is shown. The interpacket time t_d is of the order of the station reaction time (generally a few bit times), while the intercycle time ($\beta\tau$), is a multiple of the end-to-end propagation delay (τ). The particular protocol used determines the value of β (for most protocols, $1<\beta<3$).

For this class of protocols, the channel utilization or throughput, evaluated by the ratio between the average time spent in packets transmission in a cycle and the average cycle time, has the following expression

$$S=\frac{E[N]T}{E[N](T+t_d)+\beta\tau} \tag{1}$$

or equivalently

$$S=\frac{E[N]}{E[N]x+y} \tag{2}$$

where $x=1+t_d/T$ and $y=\beta\tau/T$. $E[N]$ is the average number of stations that transmit in a cycle. T is the packet transmission time. The channel capacity, i.e.,

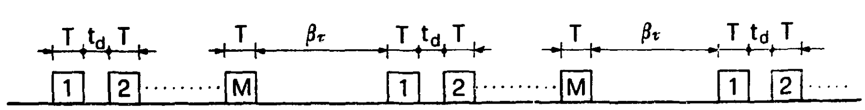

Fig. 2. Transmission cycles.

the maximum channel utilization, obtained when all stations transmit in cycle, is given by

$$C=\frac{M_c}{M_c x+y} \tag{3}$$

where M_c is the number of users connected to the bus.

The previous well-known results refer to a single channel network. In Section 4 they are extended by evaluating the average delay and considering the multichannel case with the particular connection scheme reported in the following section.

2. Connection scheme

The connection scheme, i.e., the selection of channels to which the various stations must be connected, must follow the criteria outlined in the previous section. In addition, the network, in order to minimize the delay, should provide direct (i.e., single hop) paths between all pairs of stations. Let us define:

- K: total number of parallel channels,
- P: number of LAN interfaces per station,
- M: total number of stations connected to the system,
- K_c: minimum interstation connectivity, i.e., the required minimum number of channels connecting two generic stations.

In normal conditions, i.e., when all channels are available, the connection scheme must allow any pair of stations to communicate directly on the same channel. The multi hop store and forward transmission mode should be required only during an emergency either to transmit normal communication packets, or to locate the jabbering station. To satisfy the one hop constraint the following inequality must be verified:

$$K_c \geq 1.$$

Furthermore we notice

$$K_c \leq P \leq K. \tag{4}$$

Let us subdivide the M stations in groups of $M_g=\binom{K}{P}$ elements, and, let us suppose, for simplicity, that M is a multiple of M_g, i.e., $M=M_g n$ where n is an integer. Let us consider only a particular group of M_g stations (the process can be repeated for all other groups). The value of the binomial coefficient $\binom{K}{P}$ represents the number of distinct subsets of P channels where K is the total number of available channels. Now the number of stations of a considered group corresponds exactly to the value of the binomial coefficient, and each of these stations has P ports; it follows that all stations of the group can be connected to a distinct subset of channels. An example will clarify the concept. Let us suppose that $K=6$, $P=4$. It follows that $M_g=15$. The connection scheme, considering a group of 15 stations, is shown in Fig. 3, the black dots indicating the connections. The need for direct communication

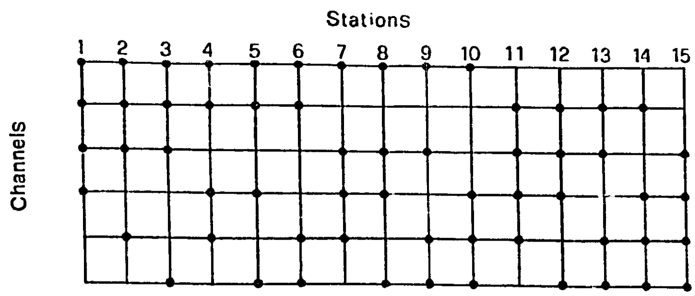

Fig. 3. Connection scheme ($K=6$, $P=4$).

between any pair of stations places some constraints on the values of P and K. In particular, $2P-K$ represents the actual minimum interstation connectivity and from our previous positions the following equation must hold

$$2P-K \geq K_c,$$

i.e.,

$$P \geq \frac{K_c+K}{2}. \tag{5}$$

Next, by considering equations (4) and (5) we have

$$\frac{K_c+K}{2} \leq P \leq K. \tag{6}$$

Given K and K_c, equation (5) permits us to choose P so that single connectivity is satisfied. When choosing the value of P, it must be remembered that for the network to survive in the case of jabbering, a station cannot be connected to all channels ($K=P$). Moreover, it may be useful to choose the interstation connectivity K_c to be at least equal to 2 ($K_c \geq 2$). In this case, a channel disconnection from the network (for maintenance reasons, for instance) does not preclude the stations communicating directly. Thus, only a worsening (rather an interruption) of the service is produced. The final selection of P must consider, besides the previous considerations, performance optimization of the system as well as cost minimization.

The proposed connection scheme is well balanced; in fact, every channel supports the same number of stations (M_c), where

$$M_c = n\binom{K-1}{P-1}, \tag{7}$$

where n indicates the number of groups while the second factor considers, for a single group, the number of stations connected to a generic channel. In fact, it counts all possible ways to connect $P-1$ ports to $K-1$ channels given that the Pth port is connected to the considered Kth channel.

The proposed connection scheme allows direct communication between all possible pairs of stations. However, in the case of faults some pairs of stations may no longer be directly connected. In this case store and forward is needed to assure communication. Some faults may even disconnect a subset of stations from the network. Let us indicate by U the number of unavailable channels. When $U \leq K_c - 1$ any pair of stations can communicate directly. For $K_c \leq U < P$ the communication is still possible but some pairs require store and forward. This means that a bridge must monitor the network, identify the fault conditions and act in the appropriate way. In the last case, when $U \geq P$, some stations are completely disconnected from the network. In particular, for any value of U, the number of station pairs of a single group which require store and forward is given by [9]

$$F_u = \sum_{L=K_c}^{U_u} Z_L \left[\frac{G}{2} + \sum_{i=L+1}^{U_u} \binom{U}{i} \binom{K-U-(P-L)}{P-i} \right] \tag{8}$$

where U_u is defined as $\min[U, P-1]$. By considering all n groups, the total number of pairs of stations requiring store and forward is given by

$$n_u = n^2 F_u. \tag{9}$$

See the Appendix for the derivation of (8), (9) and for a definition of Z_L and G. In (8) the second summation must be considered only for $i \leq U_u$ and, as usual, $\binom{x}{y}$ is considered 0 for $x < y$.

Interconnection networks similar to those presented in this paper have been considered in [6], where, unlike the connection scheme proposed in this paper, all pairs of stations share the same number of channels. Here we prefer to preserve the simplicity of the proposed interconnection scheme by considering that it may be useful to have some stations that interact more strongly than others. Moreover, subsets of this connection scheme, as shown in Fig. 4, have the symmetric property mentioned above and could be eventually used.

3. Communication protocols

It is quite natural, in a multiprocessor environment, to require a broadcast medium when a node needs to interact with all other nodes of the network. Alternatively,

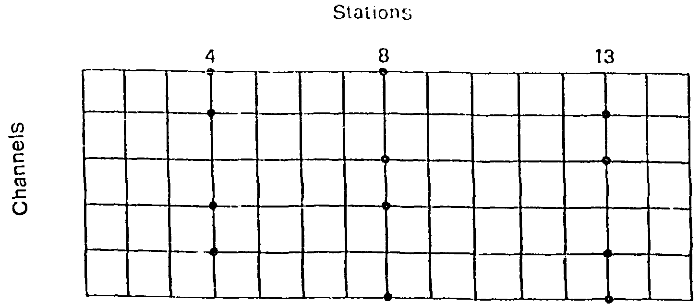

Fig. 4. Minimum number of bridges ($K=6$, $P=4$).

a node may interact only with a subset of other nodes or even with another single node. The need to address a subset of nodes arises in strongly coupled tasks, where it is important to interact with the cooperating tasks in the smallest time, so it is necessary for the subset to be physically close and it must be possible for them to be addressed like a single entity. The proposed multichannel network fully satisfies the above outlined requirements; in fact, a single transmission on a channel reaches all stations connected to that channel. We have to transmit usually on more than one channel for broadcast communications.

The proposed connection scheme permits, in normal operation, any pair of stations to communicate directly. In emergency mode ($U \geq K_c$), however, some pairs may require some store and forward hopping to connect. To this end a subset of stations can be dedicated to carry out the function of bridges. To permit the system to survive down to two available channels the bridges must ensure the connection between any pair of channels. As an example, for the system of Fig. 3, the three bridges considered in Fig. 4 permit the interconnection between any pair of channels. Thus, even with four unavailable channels all stations but one (i.e., the one which only connects to the failed channels) can still communicate with each other using one of the three bridges. In general, a subset of stations, dedicated to bridge functions, must be selected so as to permit the interconnection between any pair of channels. The number of connections between channels is $\binom{K}{2}$ which represents the number of combinations of K channels in groups of two. Obviously, by assigning each connection to a different bridge, the maximum number of bridges (B_M) is

$$B_M = \binom{K}{2}. \tag{10}$$

By considering that a bridge has P ports, it can interconnect $\binom{P}{2}$ channels so that the minimum number of bridges results to be

$$B_m = \left\lceil \frac{\binom{K}{2}}{\binom{P}{2}} \right\rceil \tag{11}$$

where $\lceil X \rceil$ indicates the lowest integer greater or equal to X. Typically an intermediate number of bridges between B_M and B_m will be used. As an example, Tables 1 and 2 show a possible association of the various links to the bridges in the case of maximum and minimum number of bridges respectively.

In general, there could be multiple paths (through different bridges) between the same pair of channels. This may lead to duplicate packets delivered to the same destination. This can be avoided by using a "spanning tree algorithm" to organize the bridges as a spanning tree, thus precluding loops in the interconnection topology [17]. Address learning and filtering is also required in the bridges in order to forward the packet to the proper channel [7].

Table 1. Association of links to 15 bridges

Stations	Links
1	1-2
2	1-3
3	1-6
4	1-4
5	2-4
6	1-5
7	3-5
8	3-4
9	5-6
10	4-5
11	2-3
12	3-6
13	2-5
14	2-6
15	4-6

3.1. Station-to-station communication

A generic station can communicate directly with any other station as long as $U<K_c$. A simple algorithm, which exploits the symmetric assignment of station to buses, permits a station to identify the channels on which the target station is connected. Among those, the channel on which to communicate can be randomly chosen.

When $U \geq K_c$, for some pairs of stations, store and forwarding through bridges is required to ensure the communication. In this case, if the sender station does not share a common bus with the target station, the channel on which to transmit can be chosen randomly. A bridge will automatically retransmit it on a channel connected to the target.

3.2. Broadcast communication

Broadcasting in the network without store and forward is always possible as long as $U<K_c$. Considering that the first channel on which the last station of a group is connected is the channel number $P-K_c+1$ and noting that all previous stations of the group are connected to previous channels, it follows that a generic station must transmit on $P-K_c+1$ channels to ensure broadcast.

Table 2. Association of links to three bridges

Stations	Links
4	1-2, 1-4, 1-5, 4-5, 2-4
8	1-3, 1-6. 3-4, 3-6, 4-6
13	2-3, 2-5, 2-6, 3-5, 5-6

When $U \geq K_c$, to ensure broadcast communication it is necessary to do store and forward on bridges. Every bridge being aware of the network status retransmits the received message on the channels to which the sender is not connected. It must be noticed that the maximum number of different channels on which the message is transmitted is limited to $P - K_c + 1$.

4. System performance

Throughput and delay analysis, the most important performance measures of a communication network, are reported, for a single channel system, in [10] and adapted to multichannel networks in [8]. The analysis considered here is based on the nongated sequential service discipline (NGSS) [4]. In this discipline the station, receiving the implicit token, transmits at most one packet conflict free. The successive transmission is allowed only after receiving again the implicit token in the successive cycle. It is assumed that the packet generation rate for each station is a Poisson process with constant generation rate λ (packets/s), and thus average interarrival time $1/\lambda$. Only one buffer is assumed for each station. No packet can be generated when the buffer is full. With the above hypotheses, a simple model is used to derive system performance in both normal and degraded mode.

The single channel throughput can be evaluated by the expression (2), where [10]

$$E[N] = \sum_{n=1}^{M_c} n p_n \tag{12}$$

and p_n can be evaluated as follows

$$p_n = p_0 \binom{M_c}{n} \prod_{j=0}^{n-1} [e^{\Lambda(jx+y)} - 1] \tag{13}$$

where $\Lambda = \lambda T$ and p_0 is a normalizing constant evaluated by

$$\sum_{n=0}^{M_c} p_n = 1.$$

The maximum channel throughput is then given by (3).

Let us define the average delay d_i for station i as the interval from the generation of a packet and the completion of its transmission. From the model stated before, the average packet delay for a generic station is simply the difference between expected interdeparture time and expected interarrival time [10]. It is easy to see from Fig. 5 that

$$d_i = t_3 - t_1 = \frac{T}{S_i} - \frac{1}{\lambda} \tag{14}$$

where t_3 and t_0 represent the departure time of two consecutive packets while t_1 is the arrival time of the packet whose transmission starts at t_2 ending at t_3. S_i is the

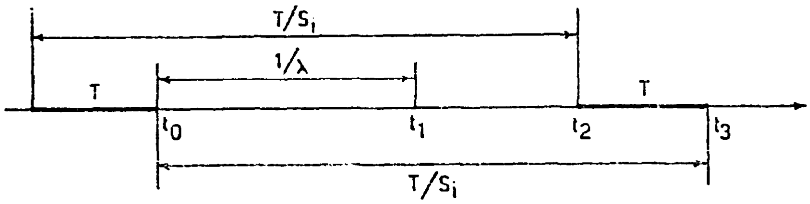

Fig. 5. Interdeparture and interarrival times.

throughput of station i and the quantity T/S_i represents the average interdeparture time for a generic station i (in seconds). Then average channel delay is

$$D=\sum_{i=1}^{M_c}\frac{S_i}{S}d_i=\frac{M_cT}{S}-\frac{1}{\lambda} \tag{15}$$

normalizing to the packet transmission time we obtain

$$D_0=\frac{M_c}{S}-\frac{1}{\Lambda}. \tag{16}$$

Let us now consider a multichannel system with K parellel broadcast channels to which the stations are connected as shown in Fig. 3. It is usually assumed that the interfaces can receive packets simultaneously from several buses [4,23]. For transmission, several alternatives, depending on buffer size and transmission parallelism, can be considered. In this paper we assume a single separate buffer for each station interface, thus allowing simultaneous independent transmissions on all buffers. With this policy a station chooses a priori the channel on which to transmit the packet and keeps that position until its transmission. This procedure corresponds to the random choice (RC) policy previously studied for multichannel random access protocols [3].

Let us define:

- W_j [bit/s] bandwidth of channel j,
- B [bit] packet length,
- $T_j=B/W_j$ [s] packet transmission time on channel j,
- $a_j=\tau/T_j$ normalized end-to-end propagation delay,
- W [bit/s] total bandwidth of the system, i.e., $W=\sum_{j=1}^{K}W_j$.

It follows that

$$a_j=a\frac{T}{T_j} \tag{17}$$

that is

$$\frac{a}{a_j}=\frac{T_j}{T}=\frac{W}{W_j}. \tag{18}$$

The offered traffic to a generic channel depends on the protocol used, namely, broadcast or station-to-station communication. For broadcast communications, by considering U $(0 \le U \le K-1)$ unavailable channels and with the further hypothesis that the transmit channels are randomly chosen among those available, it is possible to show that a Poisson arrival process is offered at every channel. In fact, recalling that a superposition and a random decomposition of a Poisson process is still a Poisson process, and that the transmit channels are $C_B = \min(P-K_c+1, K-U)$, it is possible to show that the generation rate of the process offered to a generic working channel is given by

$$\lambda_j = \frac{\lambda M_u C_B}{K-U} \tag{19}$$

where

$$M_u = n\left[\binom{K}{P} - \binom{U}{P}\right] \tag{20}$$

is the number of stations which are still connected to the system when U channels are not available. The average traffic offered by each station connected to a generic channel has the generation rate

$$\lambda_{ij} = \frac{\lambda_j}{M_c}. \tag{21}$$

For station-to-station communications only one channel is randomly selected among those connecting the source to the sink. In the case when the two stations are not directly connected, the packet is transmitted, by the source station, on a random channel. A bridge will retransmit it on another channel appropriately selected. In this way, the traffic is spread out uniformly among all available channels. The offered traffic to a generic channel is a Poisson process with a generation rate given by

$$\lambda_j = \frac{\lambda M_u C_s}{K-U} \tag{22}$$

where C_s is the average number of used buses in case of U unavailable channels. We have

$$C_s = 1 + \frac{n_u}{\binom{M}{2}} \tag{23}$$

where n_u represents the number of station pairs which need store and forward in order to communicate. The generation rate of the traffic offered by station i to a generic channel j is

$$\lambda_{ij} = \frac{\lambda_j}{M_c}. \tag{24}$$

The last case of interest, i.e., communication among a subset of stations, consists of transmitting on one or more channels to reach the intended subset. The performance of this case is comprised between the broadcast and the station-to-station case, and it will not be considered explicitly in the rest of the paper.

For the evaluation of system throughput it is evident from the definition that

$$S_j^c = \sum_{i=1}^{M} S_{ij} \tag{25}$$

where S_j^c is the throughput of channel j and S_{ij} is the throughput of station i on channel j. Let us recall that S_j^c/T_j represents the average number of packets transmitted on channel j during a second. By supposing that there are U $(0 \le U \le K-1)$ unavailable channels, the average number of packets transmitted by all available channels in a second is given by

$$\sum_{j=1}^{K-U} \frac{S_j^c}{T_j}. \tag{26}$$

For broadcast communication, the same packet is repeated on C_B channels. In this case, the system throughput can be evaluated by

$$\frac{S_u}{T} = \frac{1}{C_B} \sum_{j=1}^{K-U} \frac{S_j^c}{T_j}, \tag{27}$$

i.e.,

$$S_u = \frac{1}{C_B} \sum_{j=1}^{K-U} \frac{T}{T_j} S_j^c = \frac{1}{C_B} \sum_{j=1}^{K-U} \frac{W_j}{W} S_j^c \tag{28}$$

where S_u represents the system throughput.

In the case of station-to-station communication, a packet is transmitted only on one channel as long as the two communicating stations are connected. In the case of channel faults, however, some pairs of stations are not directly connected needing thus to use two channels for their communication. The system throughput is given by

$$S_u = \frac{1}{C_s} \sum_{j=1}^{K-U} \frac{W_j}{W} S_j^c. \tag{29}$$

The average channel delay can be evaluated by

$$D_j = \sum_{i=1}^{M} \frac{S_{ij}}{S_j^c} d_{ij} \tag{30}$$

here d_{ij} represents the average delay of station i transmitting on channel j. The average system delay can be expressed, averaging over all channels, as

$$D = \sum_{j=1}^{K-U} \frac{1}{S_u} \frac{W_j}{W} \frac{S_j^c}{C_P} D_j \tag{31}$$

where C_P assumes the value of C_B in the case of broadcast, while it assumes the value of C_s in the case of station-to-station communication. As can be easily verified from (28), $(W_j S_j^c)/(W C_P)$ represents the throughput rate of channel j. Substituting (30) in (31) we get, using (18)

$$D = \frac{1}{C_P} \sum_{j=1}^{K-U} \sum_{i=1}^{M} \frac{W_j}{W} \frac{S_{ij}}{S_u} d_{ij}. \tag{32}$$

By considering that

$$d_{ij}=\frac{T_j}{S_{ij}}-\frac{1}{\lambda_{ij}}, \tag{33}$$

we have, after a little algebra

$$D=\frac{1}{C_P S_u}\sum_{j=1}^{K-U}\sum_{i=1}^{M}\frac{W_j}{W}\left[T_j-\frac{S_{ij}}{\lambda_{ij}}\right]. \tag{34}$$

The previous model assumes dishomogeneous bandwidth subdivision among all channels. However, it has been shown in [21,13] that, for random access and round robin multichannel systems, the global system capacity is maximized with a uniform bandwidth subdivision. In this case, it results that

$$W_1=W_2=\cdots=W_K.$$

Then (18) becomes

$$\frac{a}{a_j}=\frac{T_j}{T}=\frac{W}{W_j}=K$$

and we have

$$S_1^c=S_2^c=\cdots=S_K^c=S^c.$$

Moreover, it results

$$\lambda_1 T_1=\lambda_2 T_2=\cdots=\lambda_K T_K=\Lambda_c.$$

The equations (19) and (22) can be written as

$$\lambda_j T_j=\frac{\lambda T_j M_u C_P}{K-U}, \tag{35}$$

i.e.,

$$\Lambda_c=\frac{K\Lambda M_u C_P}{K-U}. \tag{36}$$

The equations (28), (29) become

$$S_u=\frac{1}{C_P}\,\frac{K-U}{K}S^c. \tag{37}$$

The equation (34), normalizing to T and considering (37), becomes

$$D_0=M_u K\left[\frac{1}{S^c}-\frac{1}{\Lambda_c}\right]. \tag{38}$$

By considering a specific protocol we obtain the values of x and y permitting us to evaluate S_c from (2) and then S_u and D_0 from (37) and (38) for both station-to-station and broadcast, provided that the equation (36) is used to evaluate the generation rate of the traffic offered to a channel.

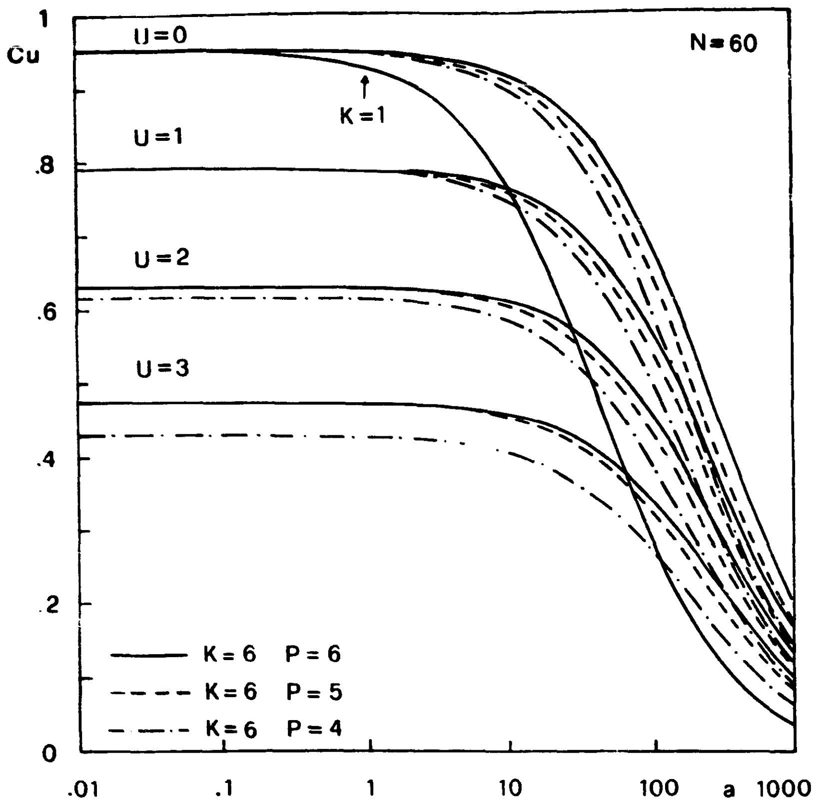

Fig. 6. Capacity vs normalized propagation delay (station-to-station).

In a similar way it can be shown that the system capacity has the following expression

$$C_u = \frac{1}{C_P} \frac{K-U}{K} C \tag{39}$$

where the channel capacity C can be evaluated by (3).

5. Numerical results

The previous model, valid for any round robin protocol, is applied to evaluate the characteristics of expressnet when this protocol is utilized with the studied multichannel topology. In this case, x and y can be approximated by [14]

$$x = 1 + \frac{t_p}{T_j} + \frac{t_r}{T_j},$$

$$y = 1.5a_j + 2\frac{t_r}{T_j},$$

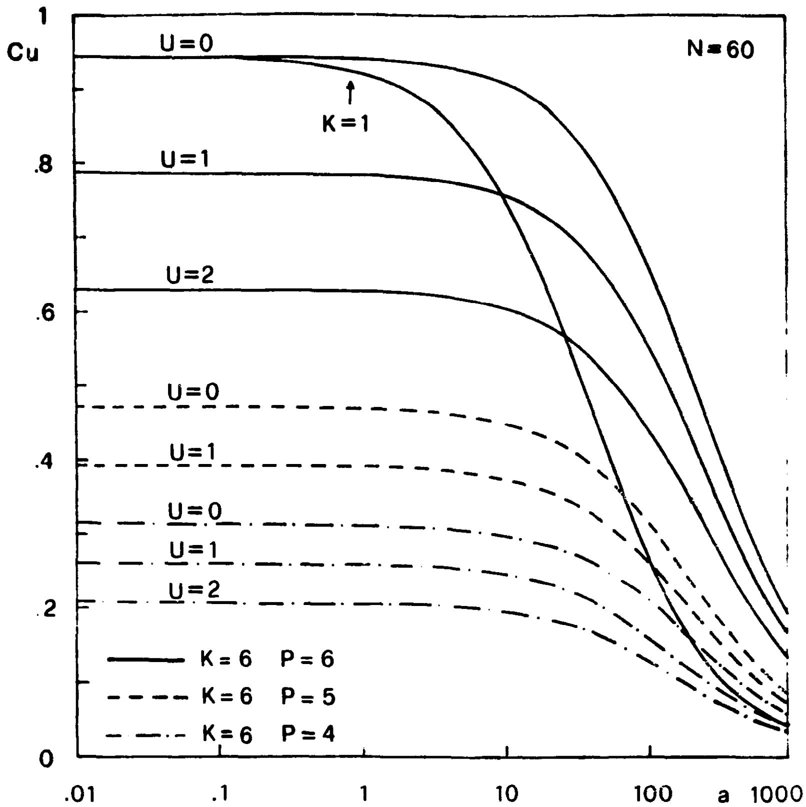

Fig. 7. Capacity vs normalized propagation delay (broadcast).

where t_p and t_r are preamble transmission time and station reaction time respectively. To allow a direct comparison, the considered topologies have the same global data rate. Whenever possible, a comparison is also made between a multichannel and a single channel network. The single channel network has a data rate equal to the sum of multichannel topology data rates. The station-to-station case always behaves better than broadcast. Moreover, in the case of station-to-station the capacity is almost independent from the value of *P*. Figures 6 and 7 report system capacity, as a function of normalized propagation delay for various number of unavailable channels (U) in the case of station-to-station and broadcast communication. Figures 8 and 9 report system throughput, as a function of the offered traffic for various connection schemes considering again station-to-station and broadcast protocols where t_p/T_j and t_r/T_j are respectively 0.05 and 0.01. Normalized propagation delay versus throughput is reported in Figs. 10 and 11. The analysis of results shows that the proposed multichannel system has a throughput considerably better than single channel network in the case of station-to-station communication. In the case of broadcasting, however, the throughput of multichannel network is, in several cases, worse than the throughput of an equivalent single channel network. This behaviour can be explained by considering that in this case the same packet is

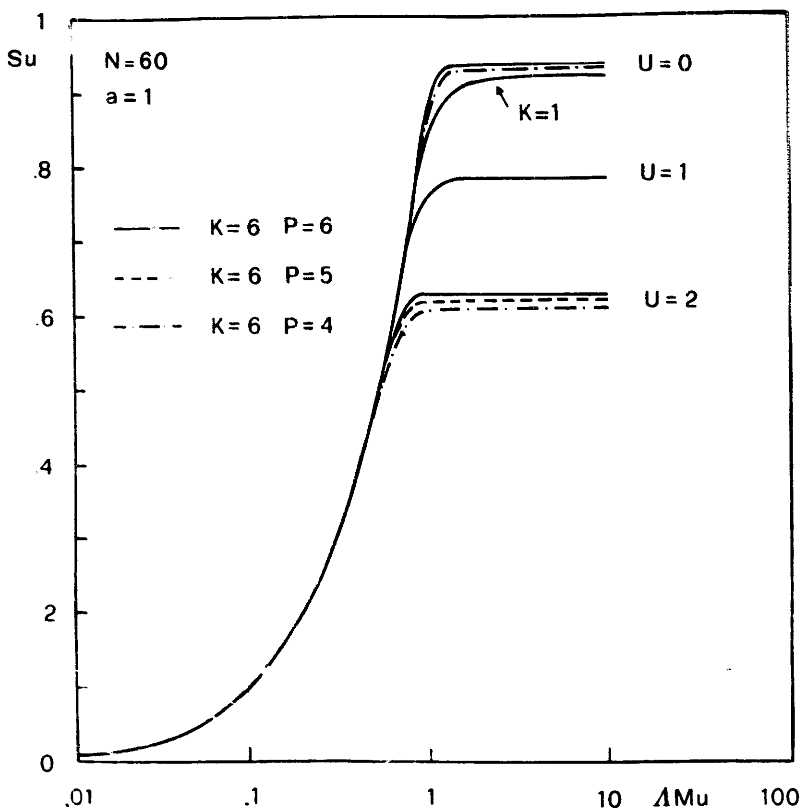

Fig. 8. Throughput vs offered traffic (station-to-station).

repeated on several channels to ensure broadcasting. For an equivalent single channel network the delay results lower than multichannel cases for a large range of system parameters, in accordance with analytical results [12]. It should be considered, however, that the single channel network, used as a comparison, does not provide any fault tolerance. Furthermore, for high data rate channels (e.g., 1 Gb/s), the equivalence single channel network can be very costly (or even impossible) to implement considering electronic component speed limitation.

6. Conclusions

The problem of fault tolerance in interconnection networks is attracting increasing interest, especially in environments where a service interruption is not acceptable. The multibus scheme proposed in this paper offers an efficient solution to the problem. In particular, the proposed connection scheme eliminates the jabbering problem and allows a gradual performance degradation of the system. The cost of such a system can be kept quite low considering that with a multibus system lower speed (and therefore lower cost) technology can be used instead of a high

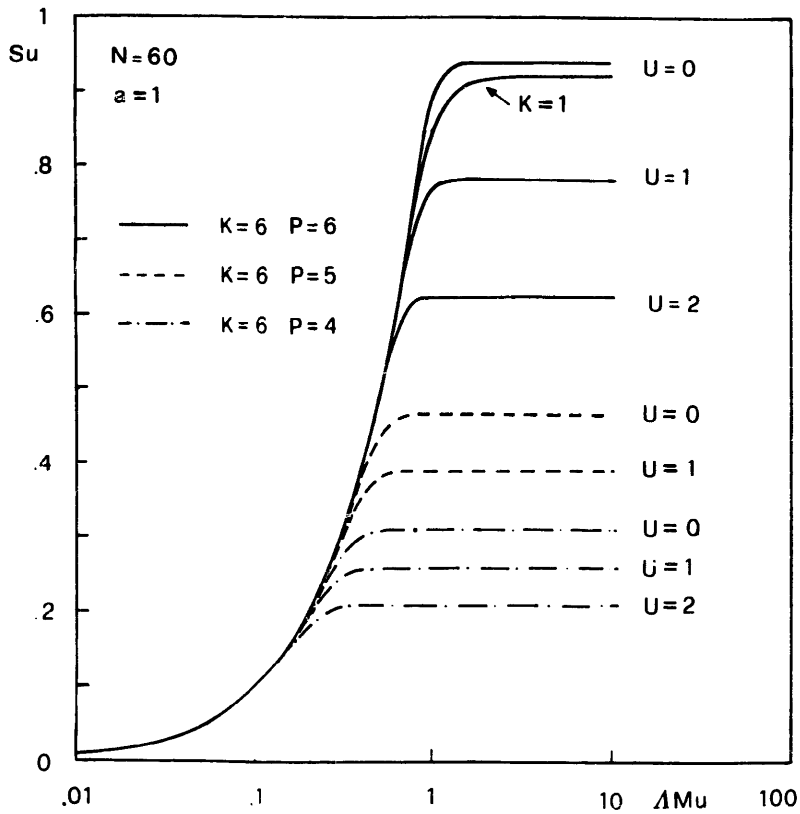

Fig. 9. Throughput vs offered traffic (broadcast).

speed, high cost technology that becomes necessary in an equivalent single channel network.

Appendix

To show the correctness of (8) we notice that the first summation considers the number of all possible unavailable channels to which a generic station can be connected. The values of L up to K_c are not considered because they do not generate station pairs needing store and forward. Z_L indicates the number of stations which are connected to the same number of unavailable channels (same L). The value of Z_L is given by

$$Z_L = \binom{U}{L}\binom{K-U}{P-L}. \tag{A1}$$

The first factor considers all possible ways to connect L ports to U unavailable channels, while the second factor corresponds to the number of ways to connect the remaining ports of the considered station to working channels. For each station

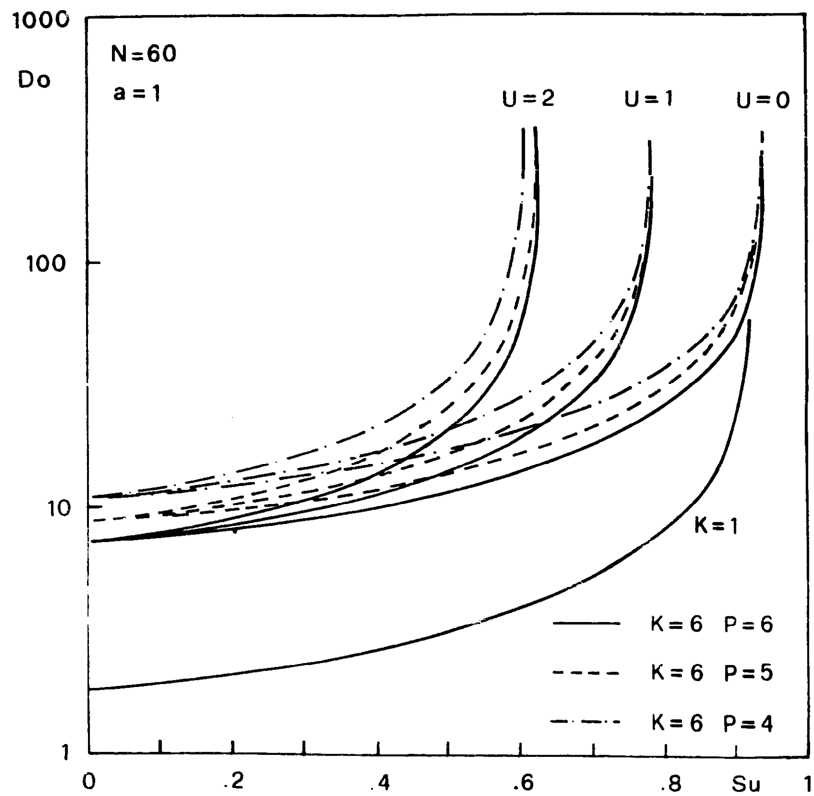

Fig. 10. Delay vs throughput (station-to-station).

considered in Z_L (target station), let G be the number of remaining stations with the same value of L which cannot communicate with this target station. We have

$$G=\binom{K-U-(P-L)}{P-L}\binom{U}{L}. \tag{A2}$$

The second binomial coefficient has the same meaning as in (A1), while the first factor corresponds to all possible ways to connect all usable interfaces to the working channels minus those occupied by the target station. Considering that a generic pair (i,j) is equivalent to the pair (j,i) it follows that the pairs which cannot communicate directly and which have the same value of L are $GZ_L/2$. The second summation in (8) considers the number of stations which cannot communicate with the target station and are connected to a number of unavailable channels greater than L. This has been obtained following the same considerations used in (A2).

The previous derivations consider only one group of M_g stations. By considering all n groups, we find that the total number of station pairs requiring store and forward in order to communicate (n_u) is given by

$$n_u=n^2F_u. \tag{A3}$$

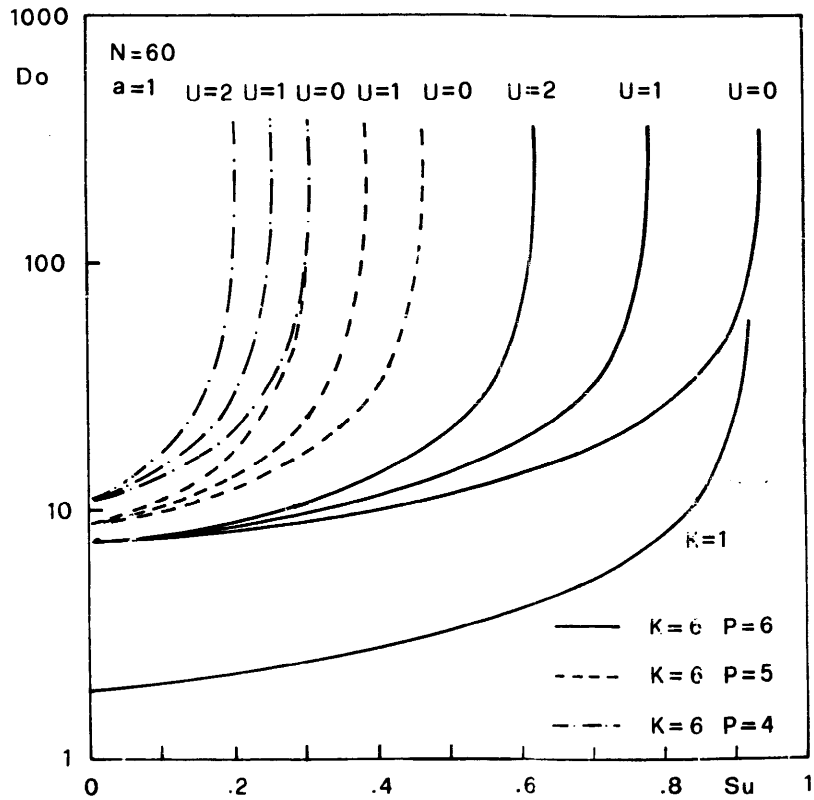

Fig. 11. Delay vs throughput (broadcast).

In fact, each station which cannot communicate with another one in a particular group, obviously cannot communicate with the corresponding stations in all other groups either. Considering that the same situation is repeated in all groups we get the previous expression. A more detailed analysis of the previous aspects is presented in [9].

References

[1] M. Ajmone Marsan and G. Albertengo, Integrated voice and data network, Comput. Comm. 5 (1982) 119-127.

[2] M. Ajmone Marsan and P. Camarda, A multiple access protocol for a local area network using two synchronous unidirectional buses, in: IASTED International Conference Telec. and Control, Halkidiki (1984) 63-66.

[3] M. Ajmone Marsan, P. Camarda and D. Roffinella, Throughput and delay characteristics of multichannel CSMA/CD protocols, in: Proceedings Globecom 83, San Diego, CA (1983) 1147-1151.

[4] M. Ajmone Marsan and D. Roffinella, Multichannel local area network protocols, IEEE J. Select. Areas Comm. 1 (1983) 885-897.

[5] A. Albanese et al., Overview of Bellcore METROCORE network, in: Workshop on High Speed LAN's, Liege (1988).

[6] B.E. Aupperle and J.F. Meyer, Fault-tolerant BIBD networks, IEEE Trans. Comput. Sci. 18 (1988) 306-311.
[7] F. Backes, Transparent bridges for interconnection of IEEE 802 LANs, IEEE Network 2 (1988) 5-9.
[8] P. Camarda, B. Castagnolo and G. Leaci, Optical fiber multichannel local area networks, in: IEEE International Conference on Communications, Philadelphia, PA (1988) 1514-1518.
[9] P. Camarda and M. Gerla, Fault-tolerance in multichannel local area networks, in: IPCCC '89, Phoenix, AZ (1989) 133-137.
[10] M. Fine and F.A. Tobagi, Performance of unidirectional broadcast local area networks: Expressnet and fasnet, IEEE J. Select. Areas Comm. 1 (1983) 913-926.
[11] M. Fine and F.A. Tobagi, Demand assignment multiple access schemes in broadcast bus local area networks, IEEE Trans. Comput. 33 (1984) 1130-1159.
[12] L. Kleinrock, Queueing Systems - Vol. 2: Computer Applications (Wiley, New York, 1976).
[13] G. Leaci, Studio delle prestazioni e sintesi di protocolli ad accesso ordinato per reti multicanale, Electrical Engineering Degree Thesis, University of Bari, Bari (1987) (in Italian).
[14] J.O. Limb and C. Flores, Description of fasnet, a unidirectional local area communications networks, Bell System Tech. J. 61 (1982) 1413-1440.
[15] R.M. Newman and J.L. Hullet, Distributed queueing: A fast and efficient packet access protocol for QPSX, in: Proceedings 8th ICCC '86, Munich (1986) 294-299.
[16] P. Rodriguez, L. Fratta and M. Gerla, Tokenless protocols for fiber optic local area networks, IEEE J. Select. Areas Comm. 3 (1985) 928-940.
[17] M. Soha and R. Perlman, Comparison of two LAN bridge approaches, IEEE Network 2 (1988) 37-43.
[18] H.S. Stone, High Performance Computer Architecture (Addison-Wesley, Reading, MA, 1987).
[19] R.R. Talbott, Feed forward rings, Fiber and Integrated Optics 7 (1988) 275-298.
[20] F.A. Tobagi, F. Borgonovo and L. Fratta, Express-net: A high performance integrated services local area network, IEEE J. Select. Areas Comm. 1 (1983) 898-913.
[21] T.D. Todd, An instantaneous bandwidth reconfigurable access control protocol, in: IEEE INFOCOM '85, Washington, DC (1985) 226-232.
[22] C. Tseng and B. Chen, D-net, new scheme for high data rate optical local area networks, IEEE J. Select. Areas Comm. 1 (1983) 493-499.
[23] Q. Yang, D. Ghosal and L.N. Bhuyan, Performance analysis of multiple token ring and multiple slotted ring networks, in: IEEE Computer Networking Symposium (1986) 79-86.