#145

10th International Symposium
on Fault-Tolerant Computing
October, 1980, Kyoto, Japan

A HIGHLY RELIABLE, DISTRIBUTED LOOP NETWORK ARCHITECTURE+

A. Grnarov*, L. Kleinrock and M. Gerla
Computer Science Department
University of California
Los Angeles, California 90024, USA

ABSTRACT

Loop network architectures offer attractive advantages over mesh networks when the host computers are limited in power and memory resources, are located in proximity one to the other, and require high performance in the data transfer (i.e., high reliability, high bandwith and low delay).

In this paper, we propose a new distributed loop architecture in which each loop interface is connected to four neighbor interfaces (daisy-chain) and each link can carry traffic in alternate directions. We show that the daisy-chain architecture is significantly more reliable than previously proposed loop networks. This property is based on the fact that the failure of one or more non-neighbor interfaces does not disconnect the loop. Furthermore, four links must fail before an interface becomes disconnected from the loop.

The paper includes an analysis of network throughput, delay and link utilization as a function of link failures, node failures and offered rates. The comparison of these results with previously published results proves the advantages of the daisy-chain network over existing loop architectures.

I INTRODUCTION

Distributed processing has become increasingly popular in recent years, mainly because of the advancement in computer network technology and the falling cost of hardware, particularly of microprocessors. Intrinsic advantages of distributed processing include high performance due to parallel operation, modular growth, fault resilience and load leveling.

A distributed system, in which several computers cooperatively share a set of resources, must be equipped with a communications subsystem. There are three topologically different approaches: the store and forward mesh; the common bus or highway, and; the loop communications system[1]. The store and forward mesh is usually associated with high reliability, long distances and a sophisticated communications protocol which leads to complex software design. Compared with the mesh, the highway is easier to implement due to the lack of any routing problems, but suffers from problems of reliability and, for longer distances, of performance.

The loop network offers attractive advantages over mesh and highway communications systems when the host computers are limited in power and memory resources, are located in proximity one to the other and require high performance in the control and data transfer (i.e., high reliability, high bandwith and low delay). A loop may, in principle, be considered as a cycle formed by point-to-point links. Instead of store and forward, the mode employed in most loop networks is check and forward. In this mode, instead of being completely buffered at each station before transmission, the messages are subject to a delay of only a few bit times.

There exists a large variety of different loop communications systems. Concerning the control, loops can be divided in two classes: loops with centralized control and loops with distributed control.

The first class includes the Farmer and Newhall loop[2], the Pierce loop[3], the Fraser loop[4], the Cambridge University loop[5], the National Security Agency loop[6], the "IBM 2790"[7], the Weller loop[8], the Serial Camac loop[9] and the Star-ring system[10].

To the second class belong the Farber loop[11], the Liu loop[12–14], and the Karsruhe loop[15].

From the reliability point of view, loops with distributed control are preferable since a failure of the loop supervisor will destroy the loop when centralized control is used.

Concerning link and interface failure, the following considerations about loop reliability can be made:

A simple, unidirectional loop system is very vulnerable in that any failure in the communications line or in a loop interface destroys the functioning of the loop. To achieve maximum reliability of the system, loop interfaces must themselves be as reliable as possible. A common practice is to separate out the receiver and transmitter into a small module and to power it from the line so that local power failure does not affect the whole loop. Another technique is to have a relay which is powered locally so that loss of power causes the relay to isolate the interface. Link failure may be handled in similar ways.

A general way of handling line failures is by use of a standby loop connected in parallel to the main loop, with each loop interface being connected to both loops. In case of the failure of one segment in one loop, the segment in the back-up loop is used.

Zafiropulo [16] discusses the use of a double loop in the case where there is a loop supervisor controlling the reconfiguration of the loop. Three techniques are defined for the case of double loop: the "bypass" technique, the "self-heal" technique, and a hybrid of the two.

In considering the performance of a particular technique, the values of three parameters are important. These are the terminal

---

+This research was supported in part by the Office of Naval Research under contract N00014-79-C-0866, and in part by the Advanced Research Projects Agency of the Department of Defense under contract MDA 903-77-0272.

*Currently on leave from the Electrical Engineering Department, University of Skopje, Yugoslavia.

pair reliability, the offered load and the length of the loop in terms of total propagation delay. The last two parameters are usually evaluated for fault-free mode of operation. It is important however to evaluate these parameters also during fault mode, namely, as a function of loop reliability. By comparing throughput and delay parameters for non-faulty and faulty mode of operation, one can obtain a better appraisal of overall loop fault-tolerance.

In the paper, we present a loop scheme which meets the following requirements:

- distributed control
- high terminal pair reliability relatively insensitive to node pair distance
- resilience to one or more non-neighbor interface failures
- no standby links
- high throughput
- low delay
- very good performance in degraded (fault) mode of operation

Section II presents a description of the proposed loop scheme. The loop topology is a daisy-chain in which, as a difference from the Farber and Lommis loop, a) all links are active; and, b) the links form one loop in one direction (basic loop) and one or two loops in the opposite direction (backward loops).

Recently Liu[13,14] proposed a Distributed Double-Loop Computer Network (DDLCN), which is designed as a fault-tolerant distributed system, and claimed that the DDLCN has better reliability as well as better communications performance than all other loops. In section III the reliability of the proposed loop is analyzed and is shown to be superior to that of the DDLCN loop.

Section IV and V analyze the communications performance of the proposed loop. From the analysis it can be concluded that the proposed loop has also better communications performance than DDLCN in both fault-free and fault mode of operation.

II DESCRIPTION OF THE PROPOSED LOOP TOPOLOGY

The proposed loop topology for a 10-node system is shown in Figure 1. It can be seen that the reliability is improved by the use of node skipping links known as a daisy-chained loop. It should be noted that all of the links are active, so that all loop elements are fully utilized. In fault-free mode of operation, the links are unidirectional and form one forward (basic) loop and two backward loops.

In the case of interface failure, the loop is still connected through the neighbors of the failed node. The loop remains connected also in the case of two or more non-neighbor failures. In the case of two or more neighbor interface failures, the loop becomes disconnected, but communications between the remaining interfaces are still maintained, even if additional single or multiple non-neighbor interface failures occur.

As the loop transmission mechanism we use the delay register insertion technique. This mechanism combines the best features of the mechanisms used in Newhall[2] and Pierce[3] loops. The delay register insertion technique permits concurrent generation and direct transmission of arbitrary-length messages onto the loop with completely distributed control. The analysis and simulation of this mechanism shows the superior performance of DLCN over the Newhall and Pierce loops.

The functional organization of a loop interface is shown in Figure 2. A loop interface must accept one stream of locally generated messages and two streams of incoming relayed messages. Conflicts of the simultaneous arrivals of messages from the three streams are resolved in the same way as in Wolf[14], i.e. by delaying incoming relayed messages in variable-length shift registers located in the loop interface. To achieve better reliability, a loop interface is split into two identical modules which have separate control and separate line driver/receiver (D/R). Both modules share transmitter (T) and receiver (R) for communications with the host (H) connected to this interface.

A message format is shown in Figure 3. The message format differs from the DLCN message format only in the Direction Bit (D). This bit is set by the origin node to inform its neighbor interface on the basic loop whether to relay the message in the same (forward) direction (D=1) or to relay it onto the backward loop (D=0). The neighbor interface checks also the destination address; if the destination address matches with the interface address, the message is removed from the loop and routed to the user device connected to the interface. If the address does not match, the interface relays the message according to the direction bit. If the message is coming from the backward link, only the destination address needs to be checked. If the address matches with the interface address, the message is routed to the interface Host; if the address matches with the address of the neighbor interface on the basic loop the message is rerouted onto the basic loop. When the matching is not satisfied, the message is simply relayed onto the backward loop.

A link is a twisted-pair channel connected to a line driver/receiver. Using a hardware interface as for DDLCN, which has a tri-state control device associated with each of the four links connected to it, an interface (node) remains connected to the loop even in the case of a three link failure. Some of the cases of link failures are shown in Figure 4. Direction of links in fault-free mode of operation is shown in Figure 4a. In the case when both outgoing links fail (links 1 and 3), the line driver/receiver for link 4 becomes a driver so that link 4 can carry traffic in the outgoing direction. Three link failure situation is shown in Figure 4c; the link number 2 is operating in "time-sharing" mode, i.e. it alternates between R and D function, and carries traffic in alternate directions. The "time-sharing" mode of operation is possible since in this case, link number 2 is carrying only local traffic to or from the Host connected to this interface .

III RELIABILITY ANALYSIS

It is customary, as an aid to reliability, to use double parallel loops with one loop used for standby operation. In this way the reliability will be improved at the expense of cost effectiveness. Namely, the standby loop does not improve throughput and delay performance during fault-free operation.

Recently the DDLCN topology was proposed consisting of two active loops (Figure 5). One loop is dedicated to message transmissions in the clockwise direction, and the other to the counter clockwise direction. The main motivation for designing the DDLCN was to protect the loop from disconnection in the case of link failures. In the case of an interface failure, the loop is disconnected, but still the communication between other interfaces is possible. It should also be noted that a failure of two or more non-neighbor interfaces makes the communication between some sets of interfaces impossible.

In order to compare the terminal pair reliabilities of the proposed loop and the DDLCN, we computed a lower bound on the terminal reliability for the proposed loop, and we carried out the exact computation for the reliability of DDLCN. The approximate terminal reliability was calculated using only three paths between interface pair. The resulting terminal reliabilities for interface pairs with distance $d = N/2$ links on the basic loop, without considering the possibility of changing a link direction, are plotted in Figure 6. The terminal reliability values are represented as a function of the number of interfaces, with link (P_L) and interface (P_N) reliability as parameters. It can be seen from the diagrams in Figure 6 that the proposed loop topology is much more reliable than the DDLCN's topology.

Using the algorithm given in Grnarov[17], we then calculated for a 10 node loop and for $P_N = P_L = 0.95$ the exact terminal pair reliabilities as a function of the hop distance between interfaces on the basic loop. The obtained results, presented in Figure 7, show that the daisy-chain loop has much better terminal reliablity. Furthermore, reliability is relatively insensitive to node distance. Specifically, the increase in node distance (on the basic loop) from 1 to 5 results in a decrease of terminal pair reliability of only 1.94%. For the DDLCN, this decrease is equal to 11.1%. The better terminal reliability of the daisy-chain loop is the result of the larger number of simple paths between terminal pairs, as compared with the DDLCN. For example, in the daisy-chain loop, there are 11, 10 and 8 simple paths between node pairs at distance 2, 3 and 4 respectively, as compared to only two paths in the DDLCN.

IV MAXIMUM AND AVERAGE NUMBER OF HOPS

In the previous section we showed that the proposed loop topology change results in significant increase of terminal pair reliability. The terminal pair reliability improvement, obtained by changing the loop topology rather than increasing loop elements reliability, also leads to an improvement of throughput and delay performance, as expected.

Figure 8 shows the hop number for the proposed and the DDLCN loops as a function of the number of interfaces. The average number of hops AVERAGE is defined as

$$\text{AVERAGE} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{N-1}\sum_{j=1}^{N} d_{ij}\right) \qquad (1)$$

while the maximum number of hops MAXHOP is defined as:

$$\text{MAXHOP} = \max d_{ij} \qquad \text{(over all (i,j) pairs)}$$

where d_{ij} is minimal number of hops between interfaces i and j, $d_{ii} = 0$ and N is the number of interfaces.

For a 10 node loop topology, the DDLCN MAXHOP and AVERAGE values are 25% and 19% higher than the corresponding daisy-chain values. In the case of N=30 these values are 36% and 38% higher. The numbers show that the improved topology will also result in a smaller delay.

Daisy chain topology shows better hop performance than the DDLCN after failure also. Figure 9a shows the impact of a link failure on the number of hops for messages in a 10 interface daisy-chain loop (Figure 1), evaluated out of each loop interface. It can be seen that a failure of the link number 1 results in an increase of 50% of MAXHOP and 41% of AVERAGE for the interface number 1. This link failure has impact on two other interfaces as well, as shown in Figure 9a. The increase of MAXHOP is 25% for both these interfaces, while the increase of AVERAGE is 7% for interface number 9 and 20% for interface number 10.

A link failure in a backward loop (link number 11) has impact on MAXHOP only for interface number 1 (an increase of 25%). The increase of AVERAGE is 33%, 20% and 5% for interface numbers 1,3 and 5 respectively.

The increase of overall AVERAGE, as defined in (1), is 7.1% in the case of link 1 failure and 5.7% in the case of link 11 failure.

A failure of link number 1 in the DDLCN has impact on four interfaces. This failure results in an increase of MAXHOP of 80%, 20%, 40% and 60% for interfaces 1,8,9 and 10 respectively. The increase of AVERAGE for this node respectively is 80%, 8%, 22% and 48%, while the overall AVERAGE increase is 16%.

Figure 10 shows the impact of an interface failure (interface number 1) on the hop number. For the daisy-chain MAXHOP shows an increase of 25% for interfaces 3, 9 and 10, and because there is no traffic to interface 1, a decrease of 25% for interface number 6. The maximum increase of AVERAGE is 28% (for interfaces number 3 and 10) while the increase of overall AVERAGE is 6.5%.

In the DDLCN loop, an interface failure has impact on 7 interfaces and results in maximal increase of MAXHOP and AVERAGE of 60% and 62% respectively. The overall AVERAGE increases by 20%.

From the previous results it can be concluded that the proposed loop topology has much better fault-mode performance than DDLCN in the case of link failures as well as of interface failures.

V THROUGHPUT CONSIDERATIONS

In this section, we show that the daisy-chain topology also offers greater throughput than the DDLCN topology. We make the following assumptions[18]: (1) Poisson message generation with total aggregate rate γ, (2) homogeneous network and (3) service rate of a transmitter equal to μ. We compute the loop maximal traffic in fault-free and fault modes of operation, in a way similar to Wolf's[14] relating each link utilization to the total message arrival rate and setting the maximal utilization to one.

For a 10-node daisy-chain loop, the maximal total rate is

$$\gamma_{SAT} = 8.18\ \mu$$

Under the same assumptions, the DDLCN throughput is:

$$\gamma_{SAT} = 7.2\ \mu$$

The previous results show that in fault-free mode of operation, the daisy-chain loop can carry 13.6% more traffic than the DDLCN.

Figure 11 shows link utilization in the case of a link failure (link number 1) for the 10 node loop. For the daisy-chain loop, we also specify the depth at which the failure information is propagated into the loop. For the case in Figure 11a. , we assume that only interfaces number 3,9 and 10 (besides 1 and 2) are informed of the failure. It can be seen that, even under this rather conservative assumption, links in the daisy chain loop are better utilized than in DDLCN leading thus to higher throughput. Namely, $\gamma_{SAT} = 5.455\mu$ for the daisy-chain loop and $\gamma_{SAT} = 3.6\mu$ for the DDLCN, which means that the daisy-chain loop can carry 51.2% more traffic in the case of a link failure. By propagating the link failure information also to interfaces number 4 and 8 (in the daisy-chain), the throughput performance becomes 62% better than for the DDLCN. Notice that in the DDLCN we assume that all interfaces are informed of the failure.

Figure 11b. shows link utilization in the case of an interface failure. It can be seen that the daisy-chain loop offers 33.3% greater throughput than the DDLCN in this case.

From the above it can be concluded that the daisy chain loop has consistently better fault-tolerant performance than the DDLCN.

VI CONCLUDING REMARKS

In the paper, a new loop network configuration is proposed. The loop combines the advantages of the daisy-chain topology and the delay register insertion technique.

It is shown that, for the same interface and link reliability, the proposed loop topology has better terminal pair reliability than DDLCN. As an example, a daisy-chain loop with probabilities $P_L = P_N = 0.95$ has the same terminal pair reliability as a DDLCN with perfect links. In the case $P_N = P_L = 0.95$ and 30 interface loop, the terminal pair reliability of diametrically opposed interfaces in daisy-chain topology is two times greater than in the DDLCN loop. Also, the terminal pair reliability of the daisy-chain loop shows relative insensitivity with respect to node pair distance. In addition to improving the loop terminal reliability, the topological change also improves the overall loop performance. For example, in fault-free mode of operation, the 30 interface DDLCN loop has an average number of hops 36% larger than the daisy-chain loop. Also, the 10 interface daisy-chain loop, in fault-free mode of operation has throughput 13.6% greater than the DDLCN.

The proposed loop has better fault-tolerant performance as well. For example, in a 10 node loop, a link failure increases the average number of hops by 16% for the DDLCN loop and by 7.1% for the daisy-chain loop. In the case of an interface failure, these numbers are 20% and 6.5% respectively. Also, in the case of a link failure, a daisy-chain loop, in which only three nodes are notified of the failure, can carry 51% more traffic than a DDLCN loop in which all nodes are informed of the failure. This value is 33.3% in the case of an interface failure.

Based on the above results, we conclude that the proposed daisy-chain topology provides a substantially better overall performance than the DDLCN topology, at the cost of a 50% increase in loop cable length.

REFERENCES

[1] B. Penney, A. Baghadi, "Survey of Computer Communications Loop Networks: Part 1", *Computer Communications* , Vol. 2 no. 4, August 1979, pp. 165-180.

[2] D. Farmer, E. Newhall, "An Experimental Distributed Switching System to handle Bursty Computer Traffic" *Proc. ACM Symposium on Problems in the Optimization of Data Comms.*, October 1969, pp. 1-33.

[3] J. Pierce, "Network for Block Switches of Data" *Bell Syst. Tech. J.*, Vol 51, No. 6, July/August 1972, pp. 1133-1145.

[4] A. Fraser, "Spider - A Data Communications Experiment" *Computing Science Tech. Rep.*, No. 23, Bell Laboratories, Murray Hill, USA, 1974.

[5] A. Hopper, "Data Ring at Computer Laboratory," *University of Cambridge Conference on Local Area Networking,* NBS Special Publication 500-31, Washington, USA, April 1978.

[6] T. Hassing et al., "A Loop Network for General Purpose Data Communications in a Heterogeneous World Data Network: Analysis and Design," *Third Data Communications Symposium,* St. Petersburg, USA, November 1973.

[7] R. Hippert, "IBM 2790 Digital Transmission Loop" *IBM J. Res. Dev.*, November 1970, pp. 662-667.

[8] D. Weller, "A Loop Communication System for I/O to a small Multi-User Computer" *Proc. 5th Annual IEEE Conf. on Hardware, Software, Firmware and Tradeoffs,* Boston, USA, September 1971, pp. 49-50.

[9] IEEE Standard Serial Highway Interface System (CAMAC) IEEE Standard 595-1976, ESONE Standard EUR 6100.

[10] J. Potvin, "The Star Ring System of Loosely Coupled Digital Devices", *Technical Report* CSRG-7 Computer System Research Group, University of Toronto, 1971.

[11] M. Hopwood, D. Loomis, L. Rowe, "The Design of a Distributed Computing System", *Tech. Rep.*, no. 25, University of California, Irvine, USA, January 1973.

[12] Y. Oh, M. Liu, "Interface Design for Distributed Control Loop Networks" *Proc. 1977 National Telecommunications Conf.*, Los Angeles, USA, December 1977, pp. 31:4(1)-31:4(6).

[13] J. Wolf, M. Liu, B. Weide, D. Tsay, "Design of a Distributed Fault-Tolerant Loop Network" *The 9th Annual International Symposium of Fault-Tolerant Computing,* Madison, Wisconsin, USA, June 1979, pp. 17-24.

[14] J. Wolf, B. Weide, M. Liu, "Analysis and Simulation of the Distributed Double-Loop Computer Networks", *Proc. Computer Networking Symposium,* NBS, Gaithersburg, USA, December 1979, pp. 82-89.

[15] D. Heger, "Comparison of Two Methods of Communication Control in Distributed Process Computer Systems With Regard to Statistical Characteristics" *Internal Rep.* Institute for Inforamations verarbeitung in Technik und Biologie, Karlsruhe, Germany, October 1977.

[16] P. Zafiropulo, "Performance Evaluation of Reliability Improvement" Techniques for Single loop Communications Systems", *IEEE Trans. Commun.*, Vol. 22, No. 6, June 1974, pp. 742-751.

[17] A. Grnarov, L. Kleinrock, M. Gerla, "A New Algorithm for Network Reliability Computation", *Proc. Computer Networking Symposium, NBS,* Gaithersburg, USA, December 1979, pp. 17-20.

[18] L, Kleinrock, *Queueing Systems, Vol. 2: Computer Applications,* Wiley Interscience, New York, 1976.

Aksenti Grnarov is Associate Professor of Computer Science at the University of Skopje, Yugoslavia. He received his Ph.D. degree from the University of Belgrade. Before joining the University of Skopje, he was associated with the Institute for Automation and Telecommunications, "Mihailo Pupin". In 1978 he came to the Computer Science Department at UCLA as a Fulbright fellow and was involved in research on on-line arithmetic and on reliability issues of computer communication networks. He is a member of IEEE and ACM.

Leonard Kleinrock is Professor of Computer Science at the University of California at Los Angeles. He received his Ph.D. from MIT. Dr. Kleinrock is the author of three major books: Stochastic Nets: Stochastic Message Flow and Delay (1964), Queueing Systems, Vol I: Theory (1975), and Queueing Systems, Vol II: Computer Applications (1976), and has published over 100 articles. He is a Guggenheim Fellow, an IEEE Fellow, and a member of the National Academy of Engineering. At UCLA, he directs a group in advanced teleprocessing systems and computer networks.

Mario Gerla earned a graduate degree in Electrical Engineering from the Politecnico di Milano, Italy, in 1966, and a Ph.D. in Computer Science from the University of California at Los Angeles in 1973. From 1973 to 1976 he was with Network Analysis Corporation, New York, where, as Manager of Computer Networking, he was involved in the planning, design and performance evaluation of several data networks for both government and industry. Since 1977 he has been on the faculty of the Computer Science Department, where he is currently involved in research on distributed processing systems and networks. Dr. Gerla is a member of the IEEE.

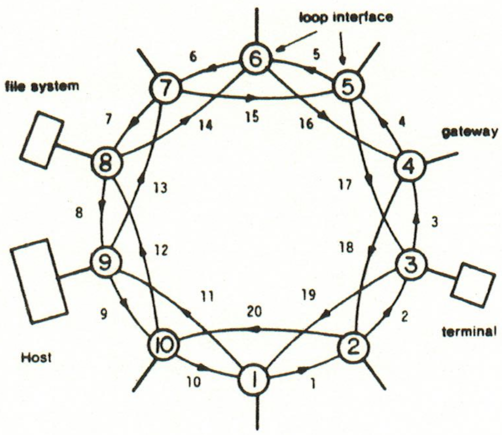

Figure 1. A 10 Node Daisy Chain Loop Network

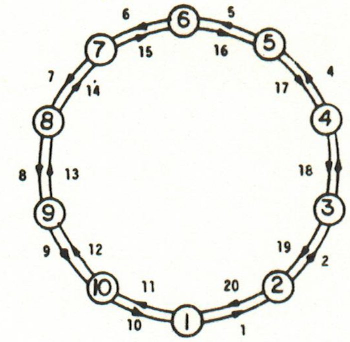

Figure 5. A 10 Node Distributed Double - Loop Computer Network (DDLCN)

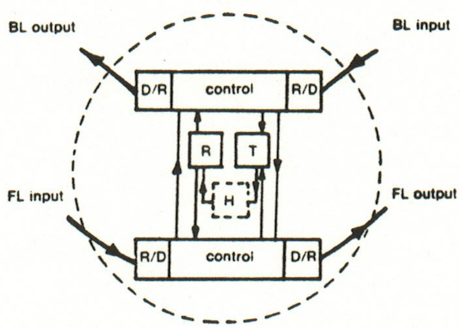

Figure 2. Functional Organization of a Loop Interface

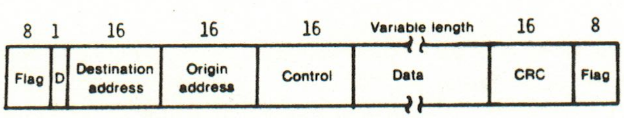

Figure 3. A Message Format

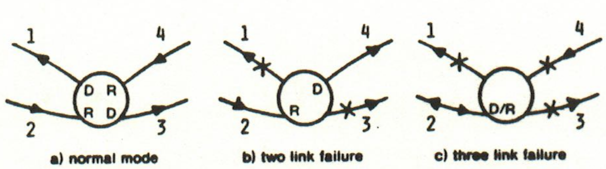

Figure 4. Increased Reliability by Carrying Traffic in Alternate Directions

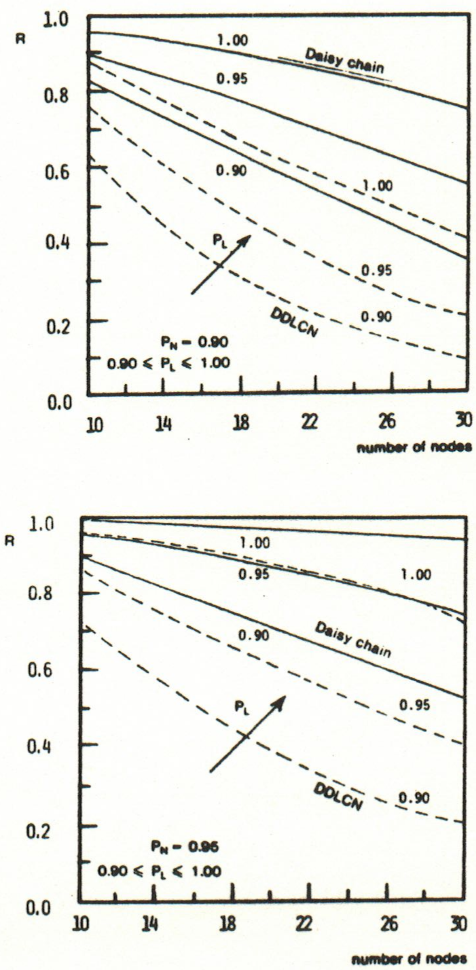

Figure 6. Terminal Pair Reliability Between Nodes on Distance N/2 for the Proposed (solid line) and DDLCN (broken line) Loops

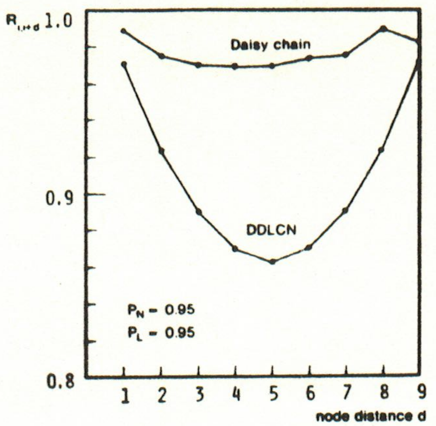

Figure 7. Terminal Pair Reliability as a Function of Node Distance

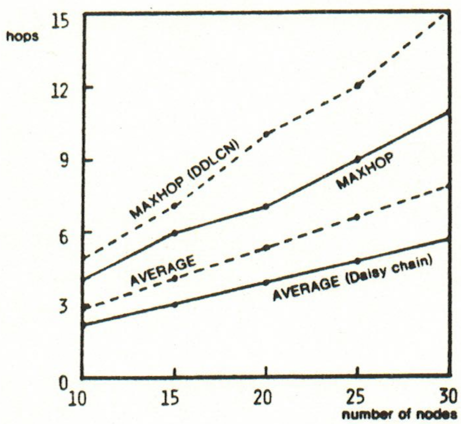

Figure 8. Maximal and Average Number of Hops for Daisy Chain

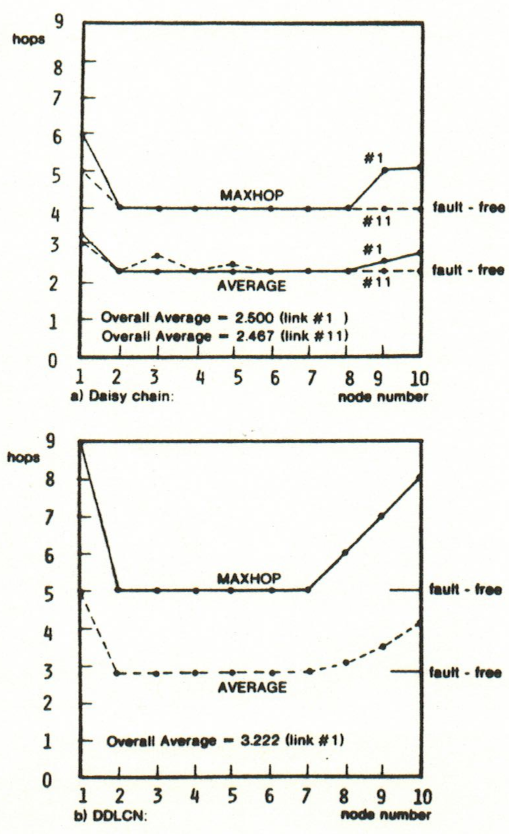

Figure 9. Impact of a Link Failure on the Hop Number

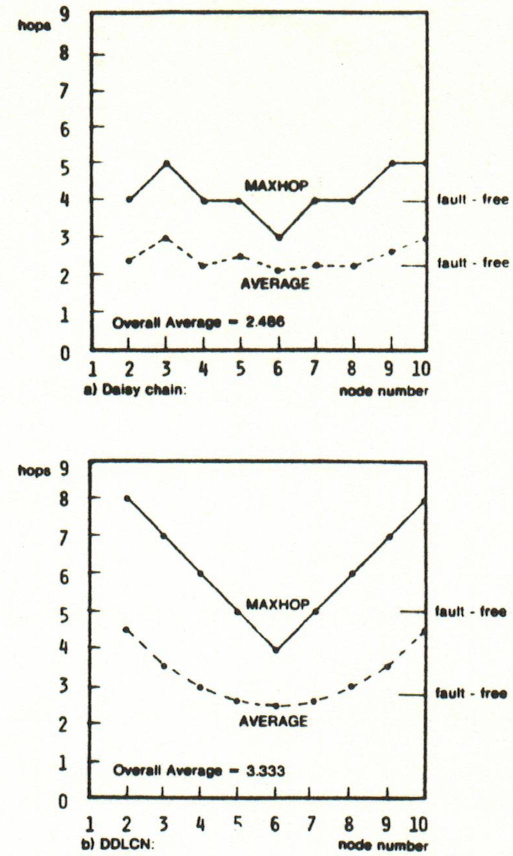

Figure 10. Impact of a Node Failure (node #1) on the Hop Number

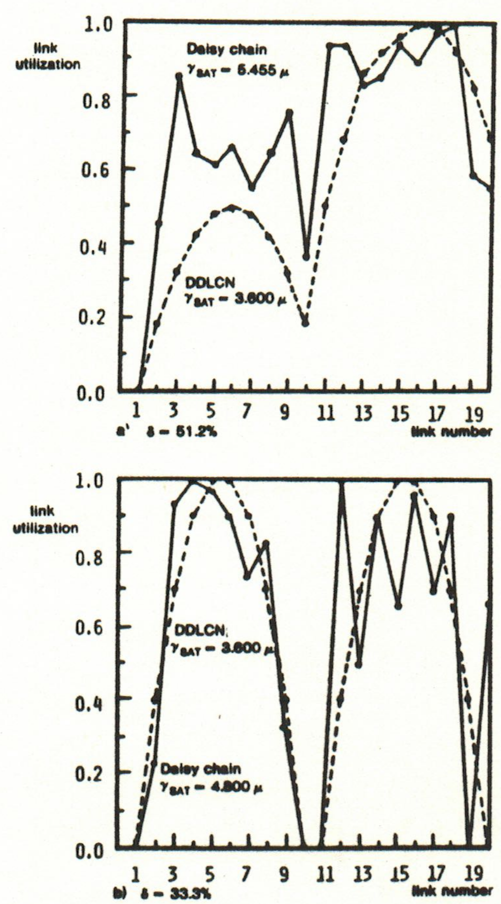

Figure 11. Link Utilization in the Case of Link Number 1 (a) and Node Number 1 (b) Failure