An Optimal Station Allocation Policy for Tree Local Area Networks

Debashis Saha¹ and Amitava Mukherjee²

¹ Dept. of Computer Science & Engg, Jadavpur University, Calcutta, India ² Price Waterhouse Associates, Salt Lake, Calcutta, India

This paper reports on the simulation results of a heuristic solution to the station allocation problem in a tree topology Local Area Network (LAN). A local network is a data communication network where communication remains confined within a moderate sized area, such as a plant site, an office building or a university campus. Tree LANs with collision avoidance switches and multiple broadcast facility have, recently, become popular due to their suitability for high speed light wave communications. Given a tree LAN with fanout F and given the total number of stations N to be connected, a combinatorial optimization problem arises regarding how to allocate the stations to the leaf nodes so that the total system availability (a network performance criteria) is maximized. This is known as the optimal station assignment problem. In this paper, it is formulated as a non-linear optimization problem which can be solved by the Lagrangean relaxation and the subgradient optimization techniques. A simple heuristic is developed based on these techniques. The simulation studies show that the proposed heuristic is relatively fast operating only in a subspace of the complete solution space.

Keywords: LAN, tree topology, multiple broadcast protocol, concurrent transmissions, Lagrangean relaxation and subgradient heuristic.

1. Introduction

A Local Area Network (LAN) is a data communication network, where communication remains confined within a moderate sized area and which supports high data rates (> 10 Mbps) over a communication medium which is shared by all network participants [1],[13],[15]. Conventional LAN topologies are ring and bus [1],[13]. Today, most of the operational LANs belong to either ring (e.g.,Cambridge ring) or bus (e.g.,Ethernet) topology. Ring LANs use token ring or slotted ring or message insertion protocol, and bus LANs use CSMA/CD or token bus protocol. Reliability is a major concern with both the bus and ring topologies, which use either redundant paths or fail-soft bypassing to cope with the problem of node failures [1].

Sometime ago, a tree topology was proposed for high speed light wave LANs by various groups [2]-[5], as a modification of the broadcast star network [6]. Tree LANs (TLANs) with collision avoidance (CA) switches [3] have several advantages [2]-[5], [10]-[12]. In contrast to a ring or a bus network, a recursively defined tree-based network [7]-[12] is adaptable to various purposes and applications within a local environment. For instance, a tree network is most suitable for a client-server architecture where server resides at the root and the clients at the leaves [10] [11]. However, in this paper, we consider communication between leaf nodes only i.e., the root node is just another switching point. All the clients and servers reside at the leaves. This is the original architecture for which TLANs were proposed [2]-[5]. In addition, a tree network provides a high degree of fail-softness, high bandwidth utilization, lowdelay characteristic, simple access protocol and modularity (flexibility) in design. One major weakness of the tree topology is that the communication between different nodes and links is very much interdependant. So, from the reliability point of view, a tree, being a minimally connected graph [8], is highly susceptible to a single node or link failure which may even totally block the operation of the whole tree. To get rid of this problem, several reliable tree architectures for TLANs have been proposed in the literature [7], [10]. They are characterized

by their high degree of fault tolerance and their capability of maintaining concurrent transmissions [5] even in the event of multiple failures.

Recently, in order to satisfy the increasing demand of wider bandwidth communication at high speed, much attention has been devoted to new LAN topology, such as the tree, where concurrent transmissions are possible. Extension of this type of tree LANs (TLANs) has been described in literature both for interconnecting user stations in large geographical areas and for carrying different types of traffic (e.g., voice, video, fac-simile, real time data, high resolution graphics etc.). To this end, TLANs have been conceived utilizing fiber optics as transmission media, for instance, in Hubnet [10] and in Alberonet [12]. The station assignment problem, however, is not addressed by the researchers in these works [2]-[12], excepting the Reference [8]. In this paper the problem has been formulated as a nonlinear optimization problem and it has been solved using Lagrangean relaxation followed by a subgradient optimization heuristics. No recent literature, excepting our previous work [19], has addressed the problem from

the non-linear combinatorial optimization point of view.

The objective of this paper is to investigate optimal partitioning of a tree-based network, also known as a Lookahead network [4], as shown in Fig. 1. Given the total number of stations to be connected by a TLAN, it is required to find out the optimum number of levels vis-á-vis fanout (i.e., arity) in the tree such that some performance constraint of the network is optimized. This is important for practical implementations where the laying of fibers is restricted.

2. Network Topology and Transmission Protocol

The structure underlying a TLAN is modeled by a rooted tree [14] (also known as an oriented tree) whose vertices represent communication switches and edges represent communication links. Formally, a rooted tree T is defined as a connected undirected graph $\langle V, E \rangle$, with a set V of m vertices and a set E of (m - 1) edges so that one element r in V is specified as its root.



Fig. 1. Hierarchical Layout for Lookhead Network

The choice of r induces direction on the edges of the free tree T. From now onwards, we use the tree instead of a rooted tree when no ambiguity arises. Given two vertices v_1 and v_2 , v_1 is an ancestor of v_2 (equivalently, v_2 is a descendant of v_1), iff v_1 is a vertex on the unique path from v_2 to r; moreover, if v_1 is the first vertex on such a path, then it is called the father of v_2 (equivalently, v_2 is called a child of v_1). The number of descendants of a vertex v is called the fanout of the vertex V. If all the vertices in a tree have an identical fanout, then the tree is called uniform (or symmetric); otherwise, the tree is called nonuniform (or, nonsymmetric). In a uniform tree, the common fanout also defines the airty of the tree. The depth of a vertex ν is the number of vertices on the path from such a vertex to r and it is denoted by d(v). The number $L = \max\{d(v), v \in V\}$ is called the level of the tree. A uniform F-ary tree is a tree in which every vertex at a depth smaller than Lhas F children, whereas all vertices at depth Lare leaves.

In a tree network, the network components are identified as different elements of V and E. We assume, in our network, that stations, switches and communication channels are represented by leaves, branch vertices and edges, respec-The channels are bi-directional, and tively. the switches are Collision Avoidance Multiple Broadcast (CAMB) switches [2]-[5], [12]. The stations are connected to leaves at depth L, and there are P stations per leaf. Station and switch protocols for the Collision Avoidance Multiple Broadcast (CAMB) tree LANs have been given in details in [5] and [12]. The edges considered are all full-duplex channels so that, in general, a switch is connected by a single full-duplex line to its parent switch and also by a set of fullduplex lines to its children. It is to be noted that the root switch has children connections only. A packet sent by a station climbs up the tree until it reaches its proper ancestor and then gets broadcast down by the ancestor to the subtree below it. The proper ancestor of a packet is the switch which roots the minimal subtree containing both the leaves which are connected to the source station and the destination station of the packet.

Except for the root switch, the architecture of a CAMB switch in a TLAN consists of an Uplink Selector (US), an Address Decoder (AD) and a Downlink Broadcaster (DB)[5]. The root switch does not have an AD for obvious reasons. The US randomly selects one of the contending children uplinks and disables the rest, thus avoiding the collision among the incoming packets. The packet, arriving on the selected uplink, is forwarded by US to the AD which finds out from the packet header whether the switch is the proper ancestor of the packet. If yes, then AD passes the packet to DB, provided the latter is free; if DB is busy, AD discards the packet because switches have no buffer to store packets. If no, then AD simply passes the packet to its uplink. DB gives higher priority [3] to the packets obtained from its parent downlink in comparison with those obtained from its own AD, and so DB preempts the broadcast of an AD packet by a packet from its parent switch. This ensures that a packet, once broadcast by its proper ancestor, does not get blocked in the midway.

Optimization of TLAN

A tree or Lookahead Network [4] offers many advantages suitable for a LAN environment. However, before a practical and efficient system can be constructed, optimization of key LAN parameters needs to be investigated. One such parameter is the availability of a network. Failure (or down time or non-availability), N, is defined as the steady state probability that a component in the system has failed, and availability, A, is defined as the steady state probability that a component in the system is operational. That is, N+A=1, or A=(1-N) and N=(1-A). (1)

As the number of leaf nodes (or stations) approaches an arbitrary large value M, and, consequently, increases the number of levels, L, as shown in Fig. 2, the optimum values of fanout F and partition P will offer a maximally available network. This is known as fanout and partition analysis [8]. It is to be noted that the value of partition P is the number of resources connected in series [8] to a single leaf as shown in Fig. 2. The availability of the network, as the number of nodes (M) increases, need to be described for the analysis of the optimum conditions. So, before starting the fanout and partition analysis,



Fig. 2. Tree Network

a mathematical model for the availability of a tree network system is developed.

For a Lookahead Network, as shown in Fig. 2, four different failures are possible [8]. They are as follows: i) Facility failure, N_f , defined as the steady state probability that a node intended to be accessed by another facility has failed, ii) Non-catastrophic failure, N_{nc} , defined as the steady state probability of failure of a node which has successfully bypassed itself from the network (i.e., by transmitting a logical "1" into the propagate line and disabling the generate line); iii) Catastrophic failures, N_c , defined as the steady state probability of failure of a node which has successfully bypassed itself from the network (i.e., random transmission a logical "1" and "0" into both the propagate line and the generate lines) and may affect the whole access mechanism, thereby disrupting the normal operation of the network, and iv) Switching failures, N_s , defined as the steady state probability of failure of switching components inside the distribution panel [2].

Assuming that the probability of failures N_f , N_{nc} , N_c and N_s is statistically independent of each other, from the Total Law of Probability [8], derive an expression for the availability of

the complete Lookahead System, A_{sys} , which can be derived as:

$$A_{sys} = A_f A_{nc} (A_c)^P [(A_s)^F]^{(2L-1)}$$
(2), where

$$A_f = 1 - N_f,$$

$$A_{nc} = 1 - N_{nc},$$

$$A_c = 1 - N_c,$$

and

$$A_s = 1 - N_s.$$

The heuristic considered in this paper optimizes A_{sys} . The values of A's normally range between 0.8 and 1.0.

4. Definition of the Optimization Problem and the Solution Technique

The following notations are used to describe the problem:

M = Total number of nodes

F = Fanout of a uniform tree

P =Size of partitions

 A_{sys} = Total system availability

 $\vec{\lambda} = \text{Lagrangean multiplier vector}$

G = Lagrangean

L = Number of levels

A = Availability

N =Non-availability (i.e., failure) = 1 - A

 $\vec{i} = (i_1, i_2, \dots, i_K) = K$ th level tree identified by a vector of ancestors, $K \in [0, L-1]$

 $C_K(\vec{\iota}_{K+1}) = L$ th level tree which includes (K-1)th level sub-trees

 $n_K(\vec{\imath}_{K+1}) =$ the size of $C_K(\vec{\imath}_{K+1}), K \in [0, L-1]$

 n_L = the size of the L-th level subtree

In this section, we first present the optimization model for a general non- symmetric tree having non-identical fanouts for vertices. Since this model is too hard to solve analytically [17], we next reduce the model for a symmetric case (i.e., uniform tree) and present a solution procedure thereof.

A kth-level cluster, C_K , is recursively in terms defined as a set of (k - 1)th-level clusters. C_K corresponds to a node at level k in a tree. According to the Dewey notation [18], a O kth level cluster is represented by a vector of predecessors, $\vec{i}_{K+1} = (i_m, i_{m+1}, \dots, i_{K+1})$ and $C_K(\vec{\iota}_{K+1})$ identifies C_K . The index i_m indicates the (m-1)st-level cluster to which C_K belongs; i_{m-1} indicates the (m-2)nd level clusters and so on. The size of a kth level cluster, C_K , is defined as the number of (k - 1)st-level clusters which are included in C_K . The size of $C_K(\vec{\iota}_{K+1})$ is denoted $n_K(\vec{\iota}_{K+1})$ and we define that $\vec{n}_K = \vec{n}_K(\vec{i}_{K+1})$, where \vec{i}_{K+1} is the vector of sizes of all the kth level clusters. It is also assumed that $\vec{n} = (\vec{n}_1, \vec{n}_2, \dots, \vec{n}_m)$ is a size vector.

With the above discussion, we write the nonlinear optimization problem for a non-symmetric network as follows:

Maximize: $A_{sys} = A_f A_{nc} (A_c)^P [(A_s)^F]^{(2L-1)}$ (3)

Subject to:

$$M = P \sum i_{L} = 1^{n_{L}} \sum_{i_{2}=1}^{n_{L-1}(\vec{i}_{L})} \sum_{i_{2}=1}^{n_{2}(\vec{i}_{3})} \sum_{i_{1}=1}^{n_{1}(\vec{i}_{2})} n_{0}(\vec{i}_{1})$$
(4)

$$L \ge 2 \tag{5}$$

$$F \ge 2 \tag{6}$$

$$A_{sys} < 1 \tag{7}$$

As mentioned earlier, the above maximization problem is difficult to solve due to the constraint equation (4). So it would be useful to simplify the model for a uniform tree which is normally acceptable in TLAN designs.

Given a uniform tree architecture with fanout F, levels L and partitions P, M number of stations are to be connected to the tree as its leaves. The problem is how to partition the stations (i.e., to cluster the nodes) so that the total availability (or reliability) of the network is maximized. We assume a symmetric network having identical links and switches, and further assume that all stations are at the same level of the tree. Since there are L levels and the tree has a uniform fanout of F, the number of switches at level L is F^L . If P stations are connected to each such switch, the total number of stations becomes PF^L which must be equal to M, i.e., $M = P(F)^{L}$. The problem can then be posed as follows.

Maximize: $A_{sys} = A_f A_{nc} (A_c)^P [(A_s)^F]^{(2L-1)}$

Subject to:
$$M = PF^L$$
 (8)

$$L \ge 2 \tag{9}$$

$$F \ge 2 \tag{10}$$

$$A_{sys} < 1 \tag{11}$$

The constraint equations (8), (9) and (10) are self-explanatory. It is to be noted that O the equation (8) is a special case of the equation (4) because when $n_0 = F$ and $n_1(\vec{i}_1) = n_2(\vec{i}_2) = \dots = n_{L-1}(\vec{i}_{L-1}) = n_L = L$, then

$$M = P \sum_{i_L=1}^{L} \sum_{i_{L-1}=1}^{L} \sum_{i_1=1}^{L} F = P \prod_{i=1}^{L} F = P F^{L}.$$

We write the Lagrangean relaxation G of equation (3), subject to equations (8) and (9), as follows:

$$G(\lambda, P) = A_f A_{nc} (A_c)^P [(A_s)^F]^{(2L-1)} + \lambda_1 (-M + PF^L) + \lambda_2 (F-2)$$

where, λ_1 and λ_2 are Lagrangean multipliers.

Differentiating G partially with respect to P and equating the differential to 0, we obtain

$$\partial G / \partial P = A_f A_{nc} (A_c)^P (\ln A_c) [(A_s)^F]^{(2L-1)} + \lambda_1 F^L = 0$$

or,

$$(A_c)^P = -\lambda_1 F^L / A_f A_{nc} (\ln A_c) [(A_s)^F]^{(2L-1)}$$

Finally,

$$P = \ln(-\lambda_1) - \ln(A_f A_{nc} (\ln A_c) [(A_s)^F]^{(2L-1)} / F^L) / \log A.$$
(12)

We have solved equation (12) for P by a subgradient optimization technique. In this technique, we consider that λ^* be an optimal solution of Lagrangean relaxation. The subgradient optimization algorithm derives lower bounds on the optimal primal function G. In this technique, the gradient method is used by replacing the gradients with subgradients. If an initial multiplier 10 is given, the following expression generates a sequence of multipliers:

$$\lambda^{n+1} = \lambda^n + t_n((PF^L)^n - M)$$

P is obtained from an optimal solution to Lagrangean relaxation *G* and t_n , a positive scalar, is stepsize. The expression of t_n is:

$$t_n = \delta_n (A_{sys} - G) / \{ (PF^L - M)^2 \}^{1/2}$$

where δ_n , a scalar, satisfies $0 \le \delta_n \le 2$. Initially, δ_n is taken as 2 and then it is halved when *G* does not improve in a given number of consecutive iterations. The subgradient algorithm terminates either if it reaches one of the following conditions:

(i) the gap between $A_{sys} - G$ is within a specified limit or,

(ii)after 300 iterations

The flow-chart of the algorithm is given in the following Fig. 3.





5. Results

We have tested our method on several example networks of varying sizes between 50 to 1000. A typical LAN usually does not contain more than 1000 stations. The results are summarized in Table I and Table II. The availability parameters are assumed to be $A_f = 1.0$, $A_{nc} = 1.0$, $A_c = 1.0$ and $A_s = .95$.

As the number of stations grows, both L and F tend to grow to make P optimum. But, for a fixed M, if P decreases, then the total system availability decreases. This is expected because a low value of P means more switching nodes. So, the chance of failure then increases.

For a fixed number of stations, total system availability decreases with the increase in either fanout or number of levels. Increasing the number of levels L increases the depth of a tree LAN, and this, in turn, causes a corresponding increase in the round trip delay. So, to minimize delay, L should be as low as possible. Again, in order to make a switch design less complex, the fanout F should not be too high. However, lower values of F and L make the value of P to be quite high (of the order of tens) which may not always be acceptable.

Figure 4 shows the plot of A_{sys} versus M, for F = 3, L = 3 and optimum values of P, in each class of M. It is evident from Figure 4 that the total system availability increases as we expand TLAN from 50 stations to 300

Μ	F	L	Р	A
50	3	2	5.77	0.80
		3	1.90	0.40
100	3	3	3.74	0.66
200	3	3	$7.42(P_{opt})$	0.94
		4	2.49	0.48
300	3	4	3.73	0.48
500	4	3	7.81	0.75
		4	2.08	0.29
1000	4	4	3.96	0.34

Table I

stations, while maintaining an optimum partitioning scheme. The maximum availability is reached somewhere near M = 200 when $P \cong 8$.

Fig. 5 also shows A_{sys} against M. While Fig. 4 is drawn for fixed values of F and L, Fig. 5 is drawn for optimum values of all P, F and L. In this graph, Asys normally decreases as we extend a TLAN by incorporating more and more stations.

Since for large networks the system availability becomes poor, a better approach will be to partition the network into smaller clusters and optimize the station allocation policy for each cluster. This will be a hierarchical design of large TLANs the global optimization of which is the topic of our future work.





Μ	L	F	Fopt	Р	Popt	Asys
50	2	3	4	5.77	$3.21 \approx 4$	0.8
		4		3.21		0.73
		5		2.01		0.56
100	2	4	4	6.31	$6.31 \approx 7$	0.98
		5		4.01		0.67
200	3	3	3	7.42	$7.42 \approx 8$	0.94
		4		3.13		0.58
300	3	4	4	4.69	$4.69 \approx 5$	0.56
		5		2.42		0.34
500	4	3	3	6.17	$6.17 \approx 7$	0.61
		4		2.08		0.29
1000	4	4	4	3.96	$3.96 \approx 4$	0.34

Table II

6. Conclusion

LANs provide shared access to resources in a fashion which allows efficient use of the technological trend of the decade. There is a number of topologies, access methods, media, hardware designs, and software designs being developed to achieve the full promise of LANs. In this paper, we have concentrated on the most recent emerging topology for high speed LANs, namely the tree, and their station allocation problem.

In near future, LANs will play an important role in B-ISDN [1]. However, most of the present LAN architectures are not designed with an eye to B-ISDN, and, hence, they may find it difficult to accommodate themselves in B-ISDN. Tree LANs are free from this apprehension. Station

Fig	5
rig.	5.

allocation to tree LANs can be formulated as a non-linear optimization problem, as shown in this paper. A straightforward heuristics, based on the subgradient optimization technique, is developed to solve the problem. It works well, even for large tree networks with about a thousand stations. The system availability also remains within acceptable limits (Asys>0.5) for small number of levels and lower fanout ranges. A partitioning technique is currently being formulated for bigger networks with more than a thousand stations in order to make the system as much as possible.

References

- [1] W. STALLINGS, Data and Computer Communication, New York:Mcmillan, 1985.
- [2] G. J. LIPOVSKI, A. GOYAL AND M. MALEK, "Lookahead Networks", AFIPS Proc. Nat. Comput. Conf., Vol. 51, pp. 153–165, 1982.
- [3] T. SUDA, Y. YEMINI AND M. SCHWARTZ, "Tree networks with collision avoidance switches", Proc. IEEE Infocom'84, 1984.
- [4] A. GOYAL, "Lookahead Network", IEEE Trans. Commun., Vol. COM-33, pp. 1160–1169, 1985.
- [5] T. SUDA AND S. MORRIS, "Tree LANs with collision avoidance: Station and Switch Protocols", Comput. Networks & ISDN Sys., Vol. 17(2), pp. 101–110, 1989.
- [6] A. E. KAMAL, "Star Local Area Networks: A Performance study", IEEE Trans. Comput., Vol. C–36, No. 4, April 1987, pp. 483–499.
- [7] K. Y. SRINIVASAN AND A. K. SOOD, "Design of highly reliable tree architectures", Computers and Elect. Engn., Vol. 14, No. 1/2, pp. 43–52.

- [8] P. D. STIGALL AND S. AZIMI, Optimization of topologically tree-based local area network, Computers & Elect. Engng Vol 14 No 1/2 pp 1–9, 1988.
- [9] C. S. RAGHABENDRA, A. AVIZIENIS AND M. ERCE-GOVAC, "Fault-tolerance in binary tree architectures", IEEE Trans. Comput., Vol. C-33, pp. 568-572,June 1984.
- [10] E. S. LEE AND P. I. P. BOULTON, "The principles and performance of hubnet: A 50 Mbit/s glass fibre LAN", IEEE J. SAC, Vol. 1, no. 5, pp. 711–720, Nov. 1983.
- [11] E. S. LEE AND P. I. P. BOULTON, "Bus, ring, star and tree local area networks", ACM Comp. Commun. Rev., vol. 13, no. 3, pp. 19–24, July 1983.
- [12] Y. YEMINI, T. SUDA AND M. SCHWARTZ, "Protocol architecture of a tree network with collision avoidance switches", Proc. ICC'84, pp. 423–427, 1984.
- [13] C. D. TSAO, "A local area network architecture overview", IEEE Comm. Mag., vol. 22, no. 7, July 1984.
- [14] E. HOROWITZ AND A. ZORAT, "The binary trees as an interconnection network: Applications to multiprocessor systems and VLSI", IEEE Tran. Comput., vol. C-30, no. 4, pp. 247–253, April 1981.
- [15] D. D. CLARK, K. T. POGRAN AND D. P. REED, "An introduction to local area networks", Proc. IEEE, vol. 66, pp. 1497–1517, 1987.
- [16] D. SAHA AND P. DHAR, "Fault tolerant architectures for tree local area networks", Computer Communication, vol. 14, no. 4, May 1991.
- [17] D. SAHA AND A. MUKHERJEE, "A study on the design and optimization of tree LANs", Technical Report, Dept CSE, JU, India, 1994.
- [18] L. KLIENROCK AND F. KAMOUN, "Optimal clustering structures for hierarchical topological design of large computer networks", Networks, vol. 10, pp. 221–248, 1980.
- [19] D. SAHA AND A. MUKHERJEE, "A simulation study on the optimization of station allocation in tree local area networks", Proceedings of International Symposium on New Transmission and Switching Technologies, Sep 1996, China.

Received: April, 1996 Accepted: April, 1997

Contact address: Debashis Saha Dept. of Computer Science & Engg Jadavpur University Calcutta 700 032 Amitava Mukherjee Price Waterhouse Associates Salt Lake Calcutta 700 091 India DEBASHIS SAHA received his B.H. degree in Electronics & Telecommunication Engineering from the Jadavpur University, Calcutta, India in 1986, the M. Tech and Ph.D. degrees in Electronics & Electrical Communication

Engineering from the Indian Institute of Technology, Kharagpur, India in 1988 and 1996, respectively. Since July 1992, he is an Assistant Professor in the Department of Computer Science & Engineering at the Jadavpur University, Calcutta, India. His research interests include topological design of computer networks, local area networks, formal models of protocols, internetworking, protocol conversion and validation, parallel algorithms and high-speed networks. He is the author of over 35 technical papers and one book. He is a member of IEEE, IEEE Computer Society, and is an active executive member of the Calcutta Chapter of the IEEE Computer Society.

AMITAVA MUKHERJEE received his Ph.D. degree in Computer Science from the Jadavpur University, Calcutta, India. He was in the Department of Electronics and Telecommunication Engineering at Jadavpur University, Calcutta, India from 1983 to 1995. Since June 1995, he is a Senior Manager in Price Waterhouse Associates, Calcutta, India. His research interests are in the area of high speed networks, topological design of computer networks, combinatorial optimization and distributed systems. His interests also include the mathematical modeling and its applications in the fields of societal engineering and international relations. He is author of over 40 technical papers, one monograph and one book. He is a member of IEEE, IEEE Communication Society.