LOCAL AREA NETWORKS: BUS and RING vs. COINCIDENT STAR

Donald C. Lindsay
Dynalogic Corporation
Ottawa, Canada

Summary

A local area networking scheme is proposed which is potentially more powerful than current approaches. A centralized star is described which could have equal reliability, could have simpler construction, and could support new services such as video teleconferencing.

Introduction

A large variety of Local Area Networks (LANs) have been introduced recently. Almost all are "buses" which use Collision Sense Multiple Access techniques in the spirit of Ethernet [1]. The most publicised alternative is the idea of "rings" [2,3]. The intention of this article is to introduce another reasonable alternative, the Coincident Star. Although a proper evaluation is difficult, arguments will be presented that this approach combines the VLSI and bandwidth advantages of the ring approach with the intrinsic reliability of a passive bus. It is also argued that the Coincident Star could be the least expensive of the three, and could support applications which are not currently feasible.

Constructing a Star

The essence of a star system is that all switching is done by a central element, and that all devices in the network will be attached to the central element by direct (and dedicated) communications links. However, we propose an untraditional central element, which uses a fully passive backplane, for the same reasons that Ethernet uses a fully passive bus.

The link boards which plug into the backplane are designed in the manner of telephone exchange boards, which can be inserted or extracted without perturbing the system. For example, a board's power and ground connections are made before signal path connections occur. Link boards would probably have on/off switches, and would be inserted "off". There would be dual power supplies, diode-or'd in the manner of Tandem systems, with on-board regulation. Again in the spirit of

Ethernet, it should be possible to design the link boards so that most faults would leave the board decoupled from the passive backplane.

A TTL implementation could operate a 32 bit wide backplane data bus at several megahertz. This means that the backplane bandwidth can easily be several hundred Mbits/second, and without difficulty can be increased past a gigabit/second. This utterly overwhelms the typical offered load of existing LANs. Therefore, we conclude that we do not need to find an arbitration scheme which copes with overload. If that ever occurs, it will be adequate to merely "forget" some of the bit stream. Ethernet uses acknowledgement packet protocols, with resend, to cope with damaged-packet and busy-receiver problems. The exact same technique will make a "cheap" arbitration scheme reliable.

The reasonable backplane protocols are of several different kinds. For example, we could use round-robin allocation. Each link board would assert an identifier onto an ID bus shortly after recognizing its own identifier. This resembles a token-passing protocol, and would have similar provisions for recovery should the chain break.

An important class of protocols are the synchronous ones, which require a clock. Since reliability forbids dependence on a single clock source, such protocols require either two clock sources, or else one per link board. This approach poses interesting technical problems such as coping when the clocks differ, and coping when a new clock is brought online. Regenerated clocking is a solution only as long as there is agreement about the source to use.

A less traditional approach would be to use contention on the arbitration bus. An imperfect analogy would be to think in terms of constructing a parallel Ethernet one metre long. Both open-collector drivers and tristate drivers permit a circuit which can assert an identifier, and check whether any other link board is "simultaneously" asserting. Note that the existence or non-existence of a collision can be determined by a trivial amount of digital circuitry in a few tens of nanoseconds.

The available circuit technology has an interesting consequence for colliding arbitration identifiers. Note that if both A and B are asserted, then "(A OR B)" appears on the backplane. If neither A nor B equal this value, then the data bus is not granted, and all players time out, as is done with Ethernet. However, if one of the identifiers equalled the or'd value, then that link board has "won" even in the face of collision. This means that an identifier has a priority which is linear in its count of zeroes. The major implication is that N contending link boards do not require N backplane connections. A secondary implication is that many collisions will not cause wasted data bus cycles. Another secondary implication is that there is zero cost to having "crash priority" messages, "background" hardware testing, and so on. We conclude that contention is both cheap and efficient, and therefore the preferred technique.

At some known time after arbitration, a winning link board would cycle the data bus. Since his identifier has already been broadcast, it is available to listeners. A destination identifier bus must also be cycled. The intended receiver is of course at liberty to ignore the data. This should only happen if the receiver is saturated by some other sender's packet. This data loss may be signalled down the backplane, or (data cycles being cheap) might merely be detected by host-to-host acknowledgement packet protocols.

It is important to note that an arbitration cycle and the previous data cycle can be overlapped. For example, receivers may latch data during a comparison, and one device delay later they will know if the latch contents are valid. (If not, they can simply re-latch during the next data bus cycle). It can be shown that 25 megahertz operation is possible with "S" TTL [6], and higher throughput is available with faster TTL logic families or ECL. Given the width of typical backplanes, the gigabit/second transfer rate mentioned above becomes a fairly modest goal.

The central switching component can have bandwidth which overwhelms the offered load. Therefore, there is no great point in having extensive buffering in the bus receive side of a link board. Assume a a ten megabit/second link, and a conservative 20 megahertz 32-bit data bus. It is highly probable that one bus cycle can be obtained in 32 bit times (64 bus cycles) of trying. Thus, the link board will only need a 32-bit serial-to-parallel converter, and a 32 bit latch. A conservative design might add a few words of FIFO, but even then the device count is very low. In fact, it is so low that one might reasonably create a deeper FIFO, and use multiple cycles of the data bus (per arbitration), simply to avoid having a wide backplane. This would become a valid tradeoff if the entire digital logic were implemented on a single chip which thereby became pin limited.

The electronics in a host interface is determined somewhat by the nature of the host. In particular, the host's I/O bandwidth determines whether the interface requires buffering. Even if it does, the digital logic is fairly easy, since addressing and contention take place at the central element, which is of course host independent.

For cost reasons, some link boards might support low bandwidth links. This does not impact the high bandwidth boards as long as packet buffering is placed into the slow board's bus receive logic.

Fourteen Points of Comparison

1. Ethernet-type LANs have a significant analog component, while rings and stars are almost entirely digital. This is significant because digital circuits are more testable and are more amenable to VLSI. The difference is caused by the multidrop nature of contention-controlled broadcast LANs. A given transmitter's signal must be receivable by all receivers on the cable, in spite of echoes, attenuation, and noise from idle transmitters. If N units are on the cable, then each transceiver must be designed conservatively enough so that all N(N-1) combinations will be reliable. Collision detection requires that an active transceiver be able to detect that it is not the only active one. Thus, the receiver part must be able to detect the weakest other transmitter during its own transmissions, and must be able to distinguish the other transmitter from its own transmitter's echoes. This can and has been done, but the resulting system has a large analog component at each node. In contrast, rings and stars only require point to point transmission, with strictly private links.

   Broadband LANs such as Wangnet fall in an intermediate category. Some analog problems are avoided, in particular by using components, such as wide-band amplifiers, which were developed for the cable TV industry. This approach is also too analog to profit from the potential of VLSI.

2. Because of the analog difficulties noted under point 1, Ethernet-type LANs must strain to reach 10 Mbit/second with a 200 node net. Rings and stars can be scaled to higher bandwidth links irrespective of net size. (The problem of scaling a Coincident Star's net size is addressed under point 8, below.)

3. Broadcast systems cannot make efficient use of higher bandwidth. The point is reached where the speed-of-light delay of a packet (its time of flight) becomes larger than the packet duration. In this situation, the system becomes an Aloha channel, with an intrinsic data capacity limit of about 18%

loaded [3]. No such limiting effect occurs in ring and star LANs.

In fairness, it should be noted that the limit only applies if the average packet size is "small" with respect to the bandwidth. A broadcast packet of 500 bits is efficient at 10 Mbits/second; it requires several times that to be efficent at 100 Mbits/second.

4. Ring and star LANs can easily use fiber optic technology, and in fact already have [2]. This very attractive and improving technology has already achieved transmission at five gigabits/second [7]. It cannot currently be used in a broadcast manner, and there is a very real possibility that it never will, in any sufficiently useful way.

5. Ethernet-type LANs have a problem relating to ground reference and power supply. It is important that a LAN not impose a uniform ground reference on all attached hosts. If it did, the network risks carrying large ground currents or creating ground loops. In order to obtain maximum transceiver performance, all present Ethernet designs seem to require direct coupling of an active component (e.g. the base of a transistor) to the cable, with consequent need for a power supply whose ground reference is the cable shield. To avoid adding a central, shared component, a per-node isolated power supply for the active part of the transceiver electronics seems to be a requirement of an Ethernet. Further, transient suppression (e.g. from lightning) requires that the cable be grounded at no more than one point. To enforce this requirement, and to maintain the ability to divide a long cable into sections for trouble shooting, the Ethernet specification [4] requires that there be no ground for the cable.

   Ring and star LANs, on the other hand, do not have a single cable. Further, the cables may be fiber optic, and if not may still achieve good bandwidth without placing active components beyond the ground isolation. Saltzer [3] reports that optical isolators and pulse transformers have both been used successfully to obtain this effect.

6. Ring LANs have a design difficulty relating to transmission rate. Not only must the repeaters agree on a clock rate, but that rate must also result in an integral number of bit times of delay when traversing the closed ring. The solutions to this (which do exist) must surely have costs, not only in engineering effort, but in testing and in analog complexity

(hence component count, price and reliability). Star LANs, by contrast, may have widely differing (even fluctuating) clock rates on their links.

7. The Coincident Star LAN applies contention at the level of words traveling across its backplane. This implies that the star will have an invariant packet delay as long as the receiver is not busy. By contrast, broadcast LANs have whole packets contending for a common medium. This results in high packet delay variance, which creates problems in applications such as voice and video transmission.

   An analysis of ring LANs is complicated by the existence of token passing protocols. It is suggested in [3] that such a protocol can be used to insure that the maximum contention time from N other nodes is no more than N packet times. This seems little better, and in fact both rings and broadcast LANs will have low variance only when lightly loaded.

   Existing implemented LANs are in fact operated with very light loading [5]. However, this loading is statistical, and applications such as voice and video transport are not attempted. The normal delay invariance of a Coincident Star is potentially a strong advantage and makes feasible some important applications.

8. When LANs are scaled to more and more nodes, they reach some limit. Broadcast LANs reach an analog limitation at perhaps 200 nodes. Ring LANs reach some high limit, which may be due to round trip delay, or which may be due to the fact that the offered load surpasses the bandwidth used in the links. Ring LANs may also be limited by cable costs (see further, point 14). Unfortunately, the Coincident Star is more likely to be limited by the size of a backplane - that is, more limited.

   When a LAN reaches a limit, two are used, and a gateway is established between them. In the case of broadcast LANs, this involves purchasing a computer and attaching it to both broadcast media. The delay through such a gateway will be one packet time, plus another contention time, plus software overhead, plus queuing delays.

   It should be possible to interconnect two Coincident Stars by running a simple link between them. The switching delay added by forwarding will be one word time, which is negligible. Of course, the boards at each end of the link must have certain features. They must be capable of recognizing multiple destination identifiers, or else must examine IDs for a certain prefix (in the manner of a telephone area code). They must also be capable of passing identifiers through, and hence are no longer protocol - ignorant. The added cost is therefore extra digital devices. If the board's digital logic were placed on a single VLSI chip, then this cost essentially vanishes, since nothing of great expense (such as pinout) has been added.

If the interLAN link is of high bandwidth compared to normal links, then it would be possible to time multiplex several packets at once between the nets. This is a little more difficult, due to the increased protocol requirements, but represents a perfectly reasonable attitude towards resources. Note that video teleconferencing in an interLAN environment would probably require either a multiplexed gateway link, or else multiple gateway links.

9. In previous LANs, any failed host interface could jam the network. The Coincident Star cannot be brought down in this way, since the central element can switch multiple simultaneous packets of unbounded length.

   Previous LANs prevented traffic overload by having a maximum packet size. With the Coincident Star, packet size is limited only by buffering considerations.

10. The Coincident Star does not appear to have any new problems with distributed initialization and distributed recovery. Reset, for example, can be done quite directly. It should pose simpler problems than those faced with rings. Token-passing protocols are not required to implement fairness and flow control.

11. With a Coincident Star, maintenance and trouble-shooting are centralized. Any failure of a link may be detected at the central unit. Any failed link board may be turned off, extracted, and replaced, all without perturbing the central switching.

    Ring LANs have a repeater at each node, and the whole net fails if any repeater fails. Worse, locating a failed repeater could require physically visiting each node. For this reason, a reliable ring LAN would probably divert each internode link so that it passes through a central point. At this "wire center", bypass switches or relays could be installed, so that ring troubleshooting and ring reconfiguration could be carried out quickly and perhaps automatically. The ring is now a star.

Broadcast LANs have a single cable, which (although passive) may fail or be attacked at any point. Trouble isolation may involve foot by foot inspection of the cable and visits to all nodes. For this reason, maintenance personnel may divide the cable into sections, which changes the quantity of effort (but not the nature of the effort).

12. Ring LANs depend on power being available at all nodes. Broadcast LANs depend on power being available at any nodes you want to have operating. The Coincident Star is the same, plus dependency on reliable power at the central unit. This unit has moderate power requirements, and could be supplied with (for example) battery backup.

13. Ring LANs must be disrupted when nodes are added or removed. Broadcast LANs and Coincident Star LANs may be edited freely.

14. Broadcast LANs lose performance as they are expanded to larger areas. However, it is easy to attach new nodes if the cable happens to pass near the desired locations.

    Ring LANs lose reliability as new nodes are added. New cable must always be added, although lower cost cables may be adequate, due to the point-to-point nature of the links (see comparison 1, above). As noted in comparison 11, rings may have starred wiring for reliability reasons. This tends to increase the amount of cable required. If a ring must cover a large area (such as a campus) then several wiring centers would probably be used.

    As noted in comparison 8, a Coincident Star LAN may be expanded easily by internetworking. A single Coincident Star is cabled in the manner of a starred ring.

Conclusions

It has been argued that a starred Local Area Network can be constructed so that it is cheap, simple and reliable. Construction techniques were suggested which are currently in use in ruggedized equipment, in telephone exchanges, and in Tandem computers. An unclocked arbitration scheme has been presented which combines efficiency with distributed control. Further, the scheme seems to have the minimum amount of analog circuitry, and appears to require an amount of digital circuitry which can be placed on a handful of chips. It does not appear to require difficult design, construction or maintenance. It is independent of the telecommunications technology employed, and can make good use of anticipated improvements in this area. It does not require all links to have the same bandwidth, which may mitigate the potentially high cable costs.

Because the Coincident Star is untried and unstandardized, it will not be desired by telephony companies, who have their own centralized switches. It may not be desired by the current suppliers of LANs. It should be desirable to the users of large mainframes, who already can overwhelm any available LAN. However, the most important feature of the Coincident Star LAN may be that it can support new services such as video teleconferencing. It is to be hoped that would-be suppliers of new and advanced services will consider the Coincident Star.

Acknowledgements

This paper is, in places, derivative of Saltzer [3]. The idea of a Coincident Star LAN is essentially an attempt to address the problems which that paper exposed.

References

[1] Metcalf,R. and Boggs,D. "Ethernet: Distributed Packet Switching for Local Computer Networks", CACM 19,7 p.395 (July 1976)

[2] Wilkes,M. and Wheeler,D. "The Cambridge Digital Communication Ring", Proc. Local Area Comm. Network Symposium, p.47, May 1979

[3] Saltzer,J., Clark,D., and Pogran,K. "Why a Ring?", Proc. Seventh Data Comm. Symposium, p.211, Oct. 1981 (SIGCOMM Computer Communications Review vol 11 , 4)

[4] "The Ethernet", Version 1.0, Digital Equipment Corp., Intel Corp., and Xerox Corp., Sept. 1980

[5] Needham, R., stated at Eighth Symposium on Operating System Principles

[6] Tait, J.B., unpublished designs

[7] Tell, R. and Eng, S., Applied Optics, 15 Nov 1981