

CS 123: Lecture 14, Bridging

(Lecturer: Geoff Voelker

May 6, 2009

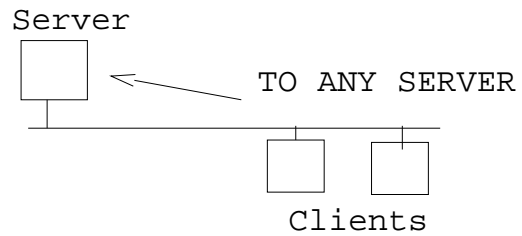
Local Area Networks Review

- Till last lecture, all the physical links we studied were a single wire with one sender and one receiver. A local area network (LAN) is like a shared wire with multiple senders and multiple receivers.
- Common LANs are Ethernet and token ring. In Ethernet, stations transmit when they want to; if two or more transmit at same time, there is a collision and they retransmit.
- Ethernets were successful but had limitations. Hence the need for bridging. But first, need to know about multicasting.

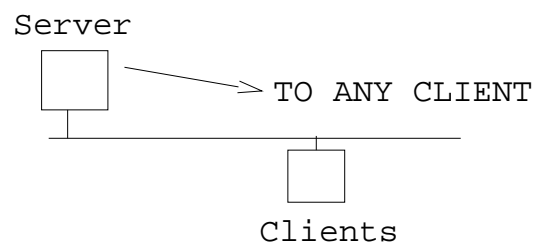
WHY MULTICAST : Autoconfiguration, Efficiency

High-order bit of DA = 1 for multicast

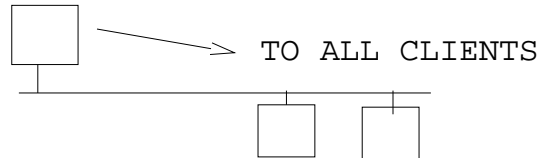
1) SOLICITATION



2) ADVERTISEMENT



3) FREE COPIES (a drug!)



Multicast Addresses

- 6 byte Ethernet Addresses assigned by 802 committee. Vendors buy a fixed 3 byte code. They can then assign remaining 3 bytes (for Dest Addresses) or 2 bytes (for 5-byte type fields). Can buy more codes.
- Multicast address denoted by Most Significant Bit. Get 2^{24} multicast and unicast address and 2^{16} types/block.
- Common multicast addresses (i.e., all IP endnodes, all IP routers) and type fields are standardized.
- Broadcast address is all 1's. Multicast better. Ethernet hardware should only pass up to software packets with DA = My address or a multicast that station listens to. Hashing or CAMs. See handout.

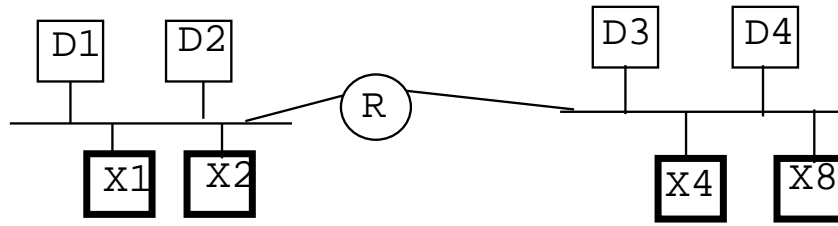
Bridging: an Exercise in Invention

- Problem Definition
- Initial Solution
- Refinements for Efficiency and Correctness.
- Generalization.
- Realization.

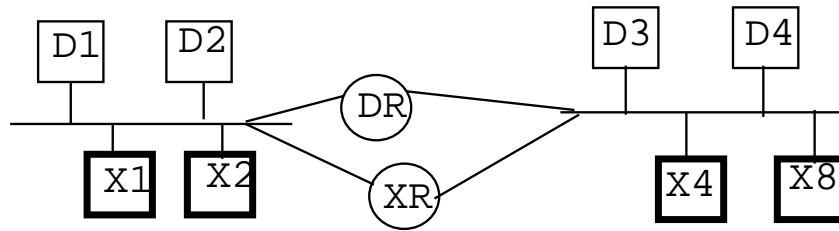
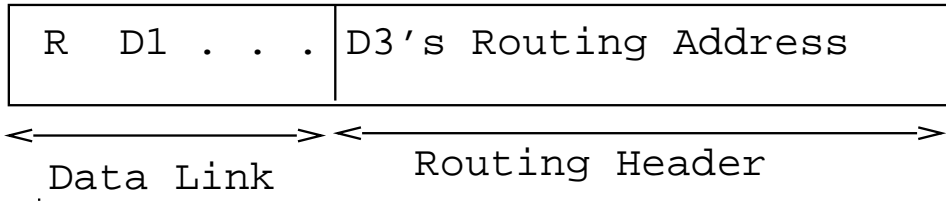
Problem Statement

Circa 1980 at Digital: Ethernets under attack

- Ethernet had limited distance (2.5 km) and stations (8000). Also perception that Ethernet collapsed at high loads. Token ring emerging.
- Question: how can we extend 2 Ethernets to make a larger Ethernet that has twice the distance, twice the bandwidth and twice the number of stations?
- Repeaters don't work as they repeat all bits everywhere. So bandwidth will not be twice. Routers work, but are expensive because we need different routers for each high-level routing protocol on the Ethernet (In next page figure, DR = DEC Router, XR = Xerox Router. D1, D2, D3, D4 are DEC stations. X1, X2, X3, X4 are Xerox).
- How can we extend LANs *transparently*: without end-stations knowing that they are on an extended LAN.



D1--> D3



Code

```
ReceiveFrame F on port X
  AddTable(F.Source, X)  (* learn source, refresh timer *)
  Y = Lookup (F.Dest) (* lookup destination *)
  If (Y = Nil) then (* unknown destination *)
    Forward Frame F on All Ports Y != X; (* flood*)
  Else if (Y != X) then
    Forward Frame F on Port Y
  Else (* received on same port as source *)
    Drop Frame F  (* filtering *)
  Endif
End;

TimerExpiry (E)  (* timer for entry E fires *)
  E.port = Nil;  (* reenable flooding *)
```

Generalizations

- Any LAN or Data Link that puts both source and destination addresses in frames (includes all 802 LANs)
- Any topology without cycles.
- More than two bridge ports.

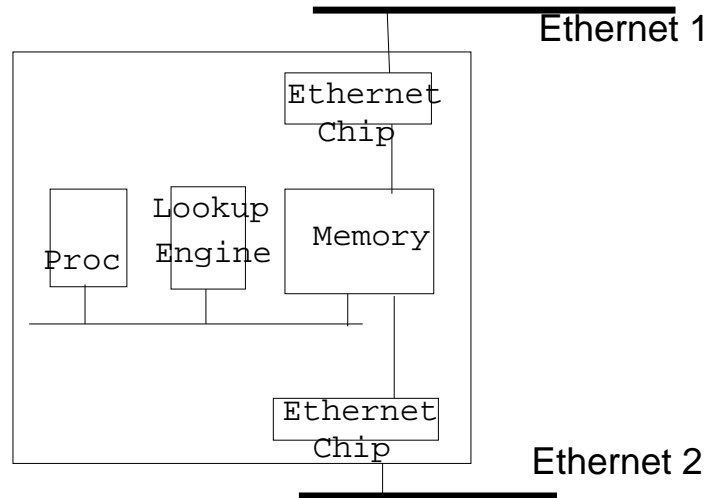
Bridge Terminology and Summary

- Transparency, Promiscuous Receive, Flooding, Filtering
- Main idea: Learn based on source addresses and forward based on destination addresses. Using flooding when there is no info, and timeout to handle stale learning information.

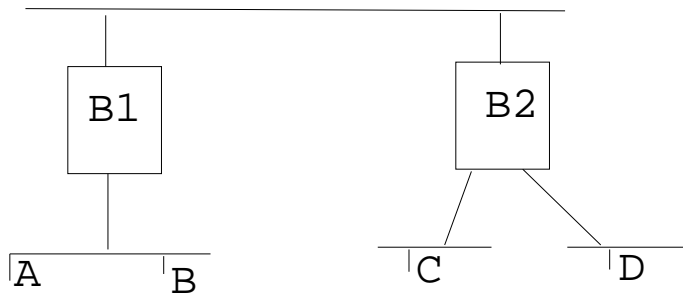
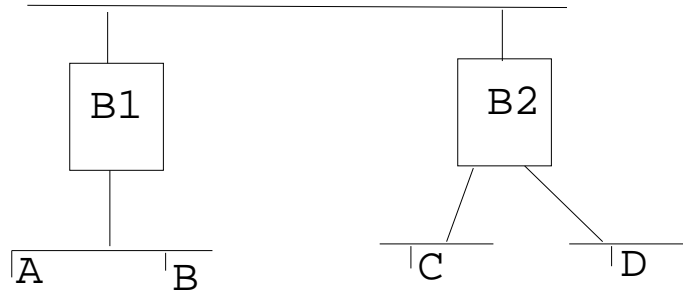
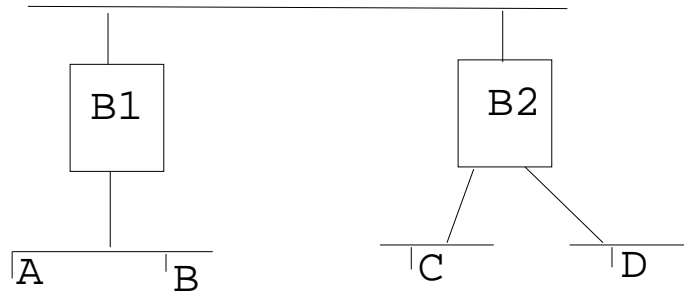
Realization

- Need much higher performance than router which handles all frames addressed to it.
- Need to decide whether to drop or forward frames in minimum interframe time on Ethernet, 51.2 usec on each port. Otherwise could drop some frames that need to be forwarded while examining others that have to be filtered.
- First DEC implementation by Mark Kempf in 1984 technology. Achieved forwarding in min frame time with low cost (1000 dollars).

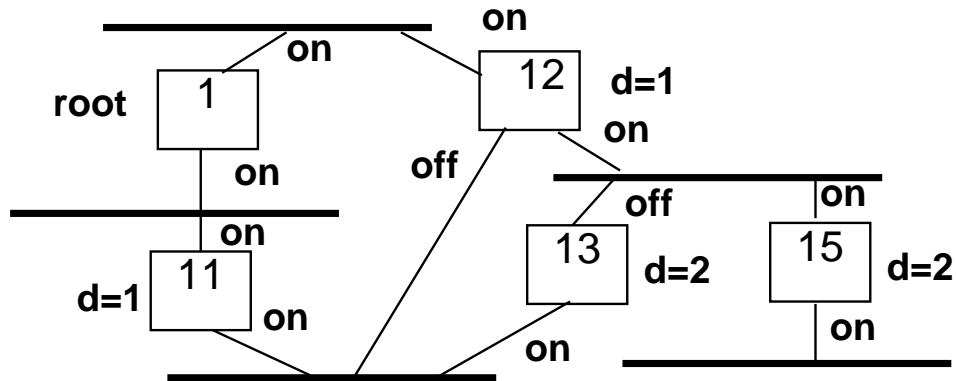
Bridge Implementation



- 1) 3 bus design to avoid congestion on processor bus
- 2) 4 ported memory design.
- 3) Hardware binary search lookup engine. Takes $\log(8000) = 13$ memory accesses of 100 nsec each = 1.3 usec
- 4) Processor stays in loop after a packet interrupt servicing as many packets as arrive to reduce context switching overhead.

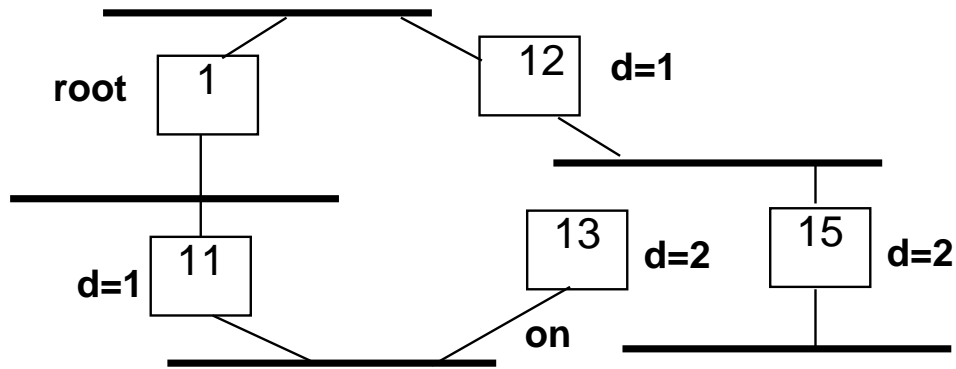


Centralized Algorithm



- Each bridge finds Min port — port through which it has shortest path to root (a.k.a. parent)
- Each bridge finds the ports for which this bridge is on the shortest path between root and corresponding LAN: Designated Ports.
- Each bridge turns ON Min port and all Designated Ports. ON, OFF are software states: always receive hello and management messages on all ports. Drop data packets to/from OFF port.

Final Tree



- To be a tree, each LAN must have a unique path to every other LAN.
- Algorithm guarantees that each LAN can get to root only through designated bridge for LAN and designated bridges have unique path through their parent LAN.

Distributed Algorithm

- Easy to calculate root and distance using a centralized algorithm. How to calculate if each bridge powers up knowing only its own ID and can only get info from received messages.
- Want to calculate Min ID node (root) and distance to Min ID node. Can keep an estimate of Root and Dist and send updates to neighbors and update own estimates based on neighbors distances.
- First try to find an update rule for root estimates based on received estimates from neighbors. Then find an update rule for distance, assuming your distance is right.