Bridges, Spanning Trees & Switches

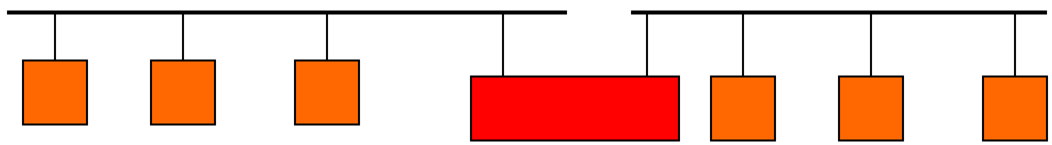
- Focus:
  - What to do when one shared LAN isn't big enough?

- Interconnecting LANs
  - Bridges and LAN switches
  - A preview of the Network layer

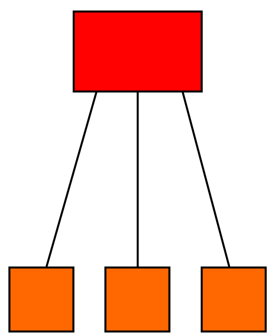
Application
Presentation
Session
Transport
Network
Data Link
Physical

Warning: confusing terminology / pictures

Original Ethernet (repeater)

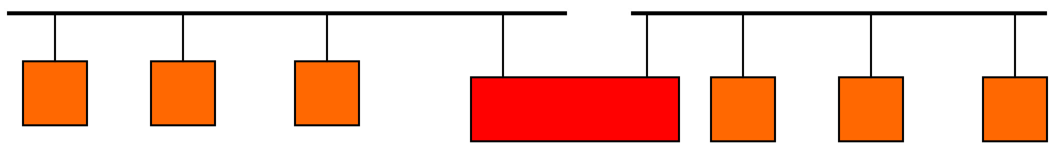


Modern Ethernet
(Hub)

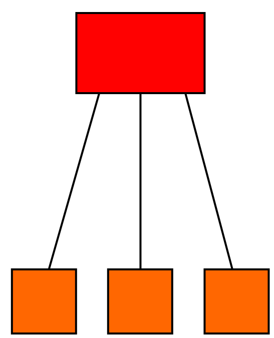


Not talking about these -- historical

Instead, we'll be talking about these



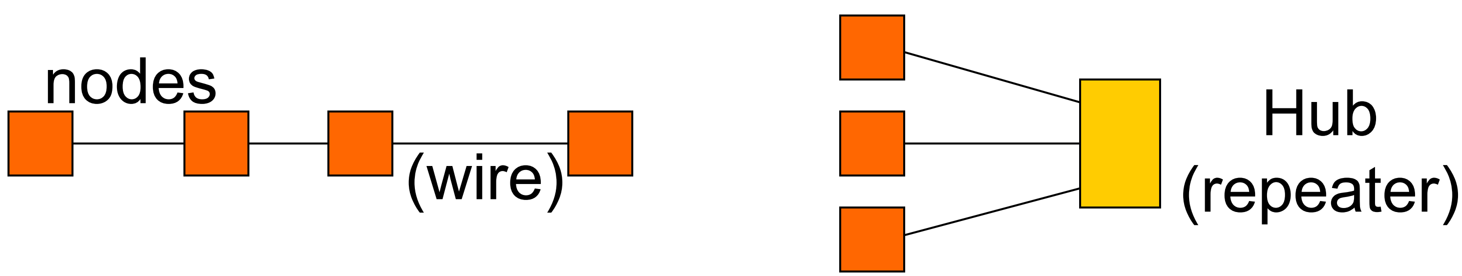
Bridges – well this is historical too …



Finally, a modern LAN switch – this is where we end up!

Limits of a LAN

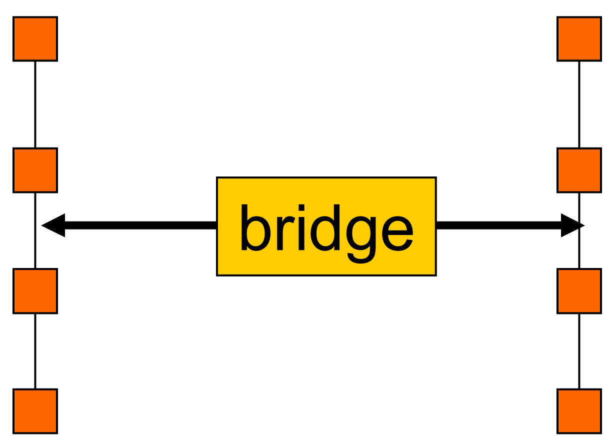
- One shared LAN can limit us in terms of:
  - Distance
  - Number of nodes
  - Performance



- How do we scale to a larger, faster network?
  - We must be able to interconnect LANs

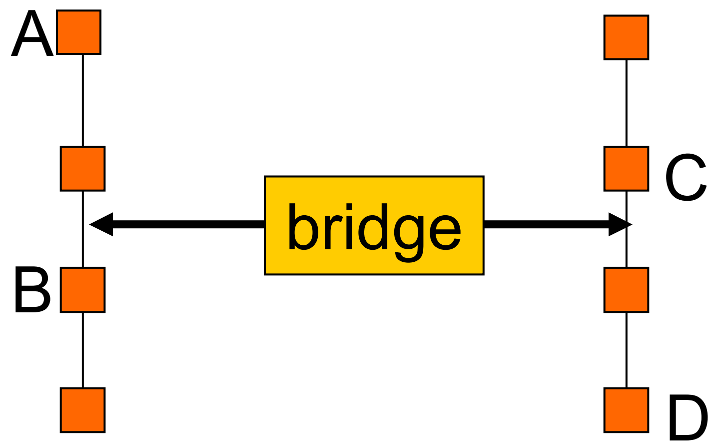
Bridges and Extended LANs

- "Transparently" interconnect LANs with bridge
  - Receive frames from each LAN and forward to the other
  - Each LAN is its own collision domain; bridge isn't a repeater
  - Could have many ports or join to a remote LAN



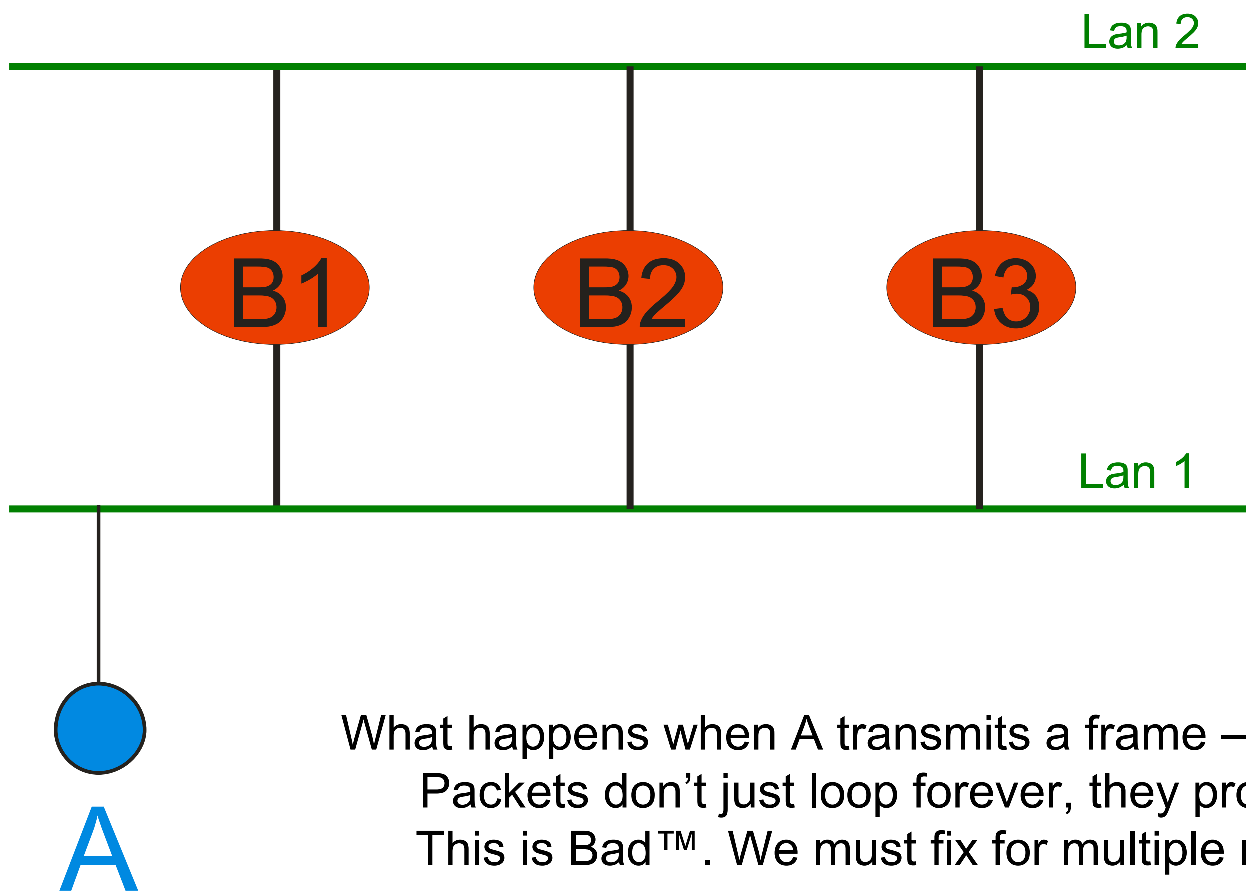
Backward Learning Algorithm

- To optimize overall performance:
  - Shouldn't forward A→B or C→D, should forward A→C and D→B



- How does the bridge know?
  - Learn who is where by observing <u>source</u> addresses and prune
  - Forward using destination address; age for robustness

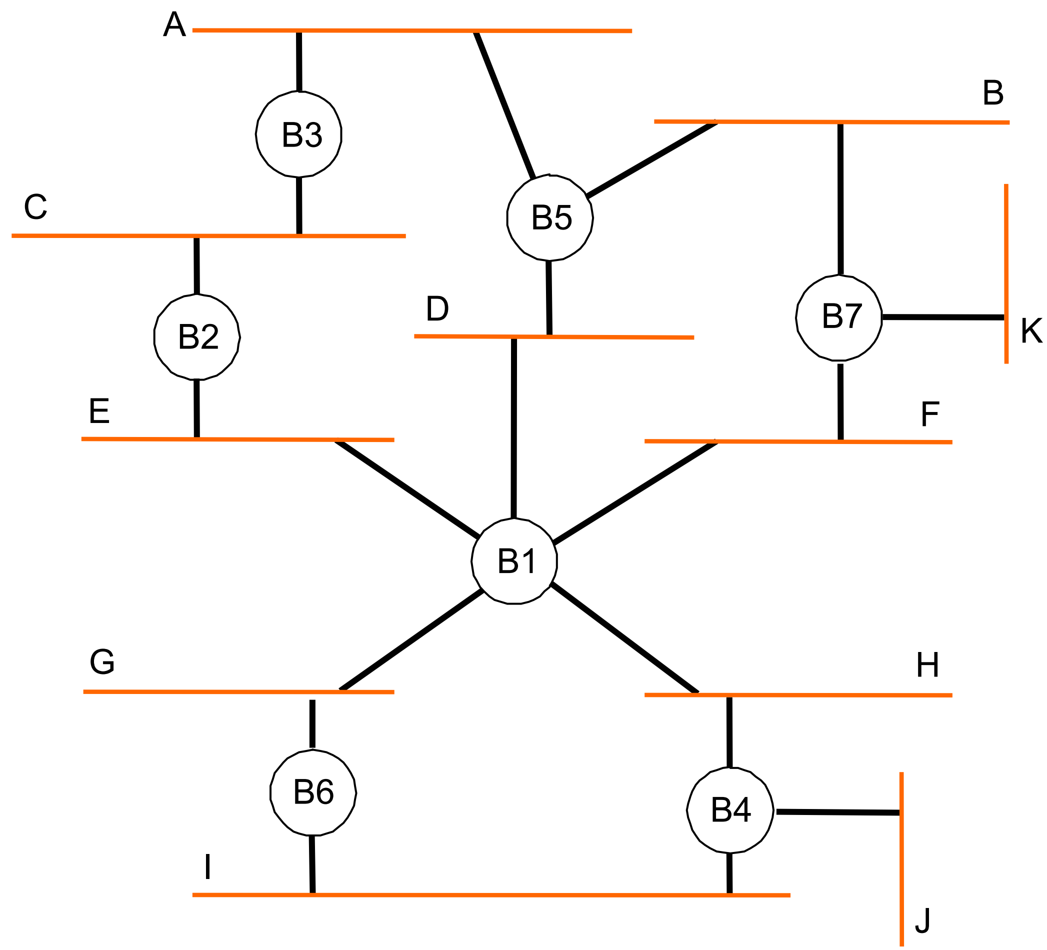
Is redundancy good or bad?



What happens when A transmits a frame – to anyone?
Packets don't just loop forever, they proliferate.
This is Bad™. We must fix for multiple reasons!

Large networks -- why stop at two?

- But to avoid loops we must forward only on select bridge ports!
- The Spanning Tree algorithm does this
- It is separate from backward learning



Radia Perlman says …

Algorhyme

I think that I shall never see
A graph more lovely than a tree.

A tree whose crucial property
Is loop-free connectivity.

A tree which must be sure to span
So packets can reach every LAN.

First the Root must be selected.
By ID it is elected.

Least cost paths from Root are traced.
In the tree these paths are placed.

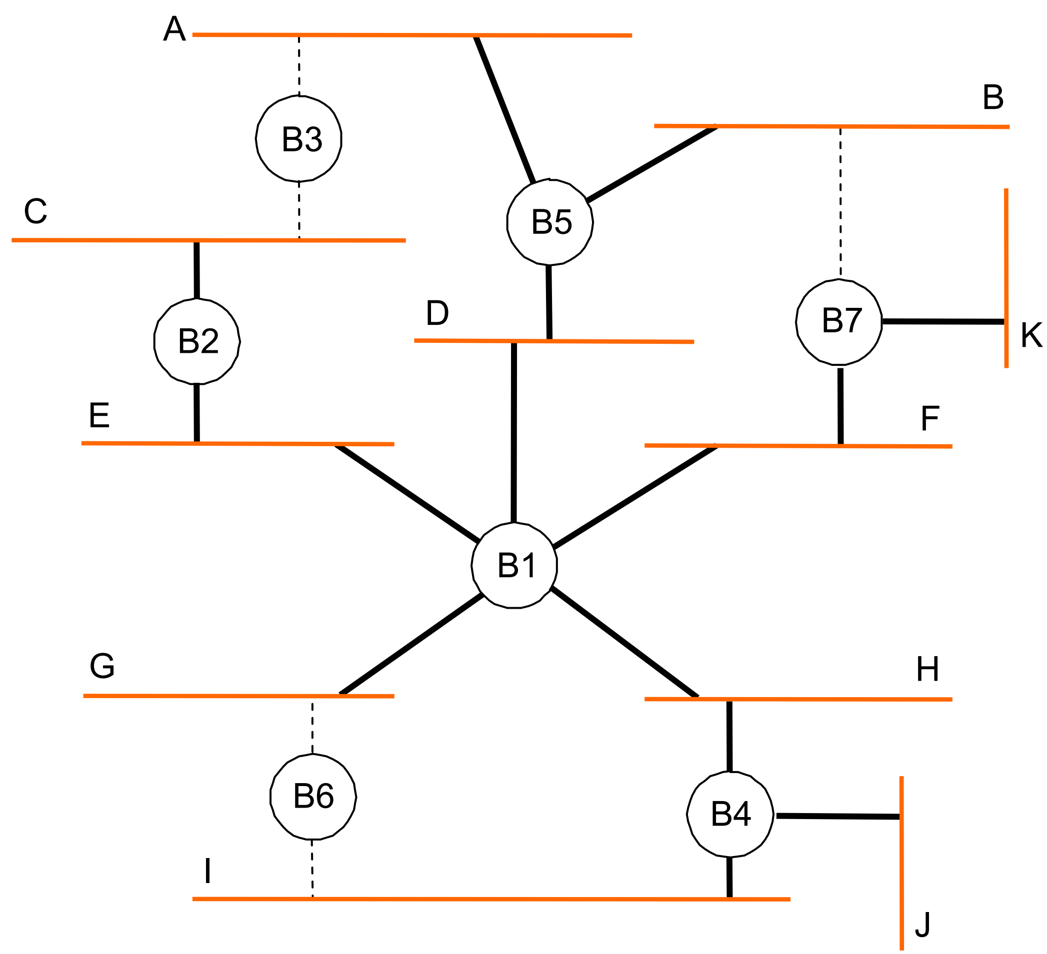
A mesh is made by folks like me
Then bridges find a spanning tree.

From: “An Algorithm for Distributed Computation of a Spanning Tree in an Extended LAN”,
R. Perlman, SIGCOMM 1985.

Spanning Tree Example

- Spanning tree uses select bridges so there are no cycles
  - Prune some ports
  - Only one tree

- Q: How do we find a spanning tree?
  - Automatically



Spanning Tree Algorithm

- Distributed algorithm to compute spanning tree
  - Robust against failures, needs no organization

- Outline:
  - Goal is to turn some bridge ports off
  1. Elect a root node of the tree (lowest address)
  2. Grow tree as shortest distances from the root (using lowest address to break distance ties)
  - All done by bridges sending periodic configuration messages over ports for which they are the "best" path
  - Then turn off ports that aren't on "best" paths

Algorithm continued

- Each bridge sends periodic messages to others containing:
  - Its address, address of the root bridge, and distance (in hops) to root
- Each bridge receives messages, updates "best" config.
  - Smaller root address is better, then shorter distance
  - To break ties, bridge with smaller address is better
- Initially, each bridge thinks it is the root
  - Sends configuration messages on all ports
- Later, bridges send only "best" configs
  - Add 1 to distance, send configs where still "best" (designated bridge)
  - Turn off forwarding on ports except those that send/receive "best"

General Design Principles

- All bridges to run the same algorithm
- All bridges start with zero information and operate in parallel
- Bridges send periodic messages about their own state
- State that hasn't been refreshed in a while is deleted (*soft state*)
- If we all have the same inputs and are running the same algorithm, we converge to a globally consistent state.

This is a common design pattern for network protocols that adapts to failures. Learn it. Live it. Love it.

Algorithm Example

- Message format: (root, dist to root, bridge)

- Sample messages sequences to and from B3:

1. B3 sends (B3, 0, B3) to B2 and B5
2. B3 receives (B2, 0, B2) and (B5, 0, B5) and accepts B2 as root

3. B3 sends (B2, 1, B3) to B5
4. B3 receives (B1, 1, B2) and (B1, 1, B5) and accepts B1 as root

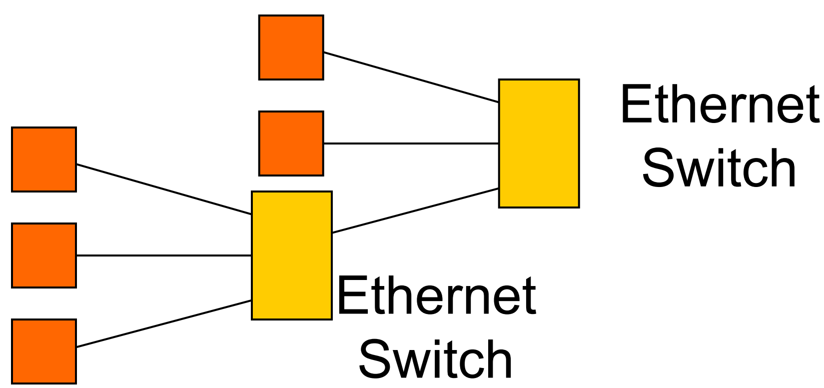
5. B3 wants to send (B1, 2, B2) but doesn’t as its nowhere “best”
6. B3 receives (B1, 1, B2) and (B1, 1, B5) again … stable
   – Data forwarding is turned off to the LAN A

Some other tricky details

- Configuration information is aged
  - If the root fails a new one will be elected
- Reconfiguration is damped
  - Adopt new spanning trees slowly to avoid temporary loops

LAN Switches – what is used today

- LAN switches are multi-port bridges
  - Modern, high performance form of bridged LANs
  - Frames are switched, not shared like hubs
  - Every host on a separate port, can combine switches



- Other important modern feature: VLANs = virtual LANs
  - Ability to split hosts on an extended switched LAN into groups
  - Groups appear on logically separate LANs -- helps with management.

Perspective on Bridges/Switches

- LAN switches form an effective small-scale network
  - Designed to work transparently with no host changes
  - Plug and play for real!

- But they don't scale well to large networks
  - Little control over forwarding paths
  - Size of bridge forwarding tables grows with number of hosts
  - Broadcast traffic flows freely over whole extended LAN
  - Spanning tree algorithm limits reconfiguration speed
  - Poor solution for connecting LANs of different kinds

- We can do better by changing hosts too. Next up: IP and routing.