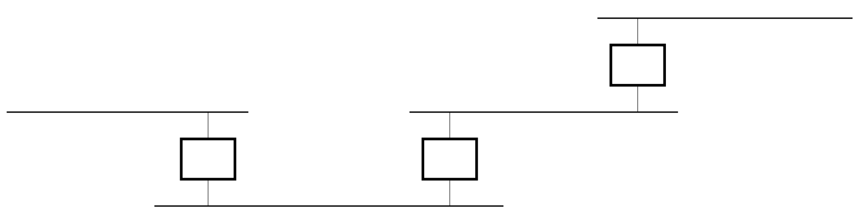
# Bridges and Extended LANs

## Overview

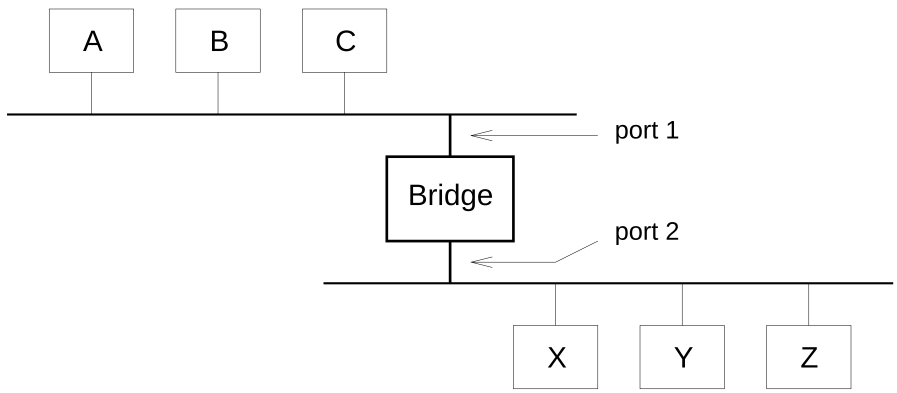
- LANs have physical limitations (e.g., 1500m Ethernet)
- Connect two or more LANs with a *bridge*
  - accept and forward strategy
  - level 2 connection (does not add packet header)



- Collection of LANs connected by bridges called an *extended LAN*

# Learning Bridges

- Do not forward when unnecessary



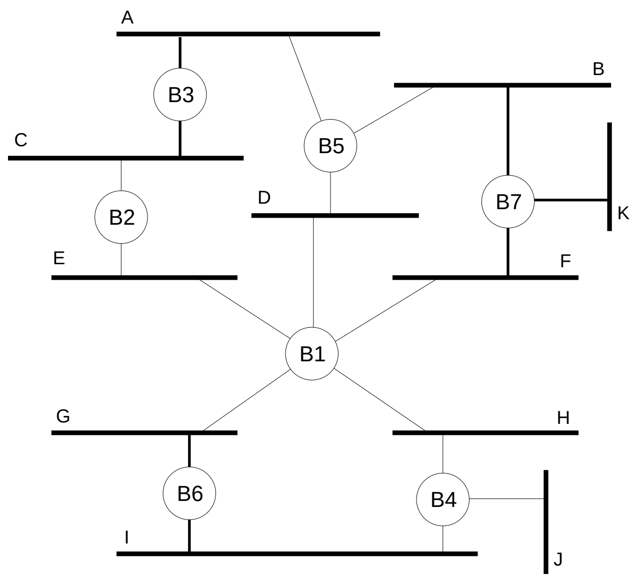

- Maintain forwarding table

| Host | Port |
|---|---|
| A | 1 |
| B | 1 |
| C | 1 |
| X | 2 |
| Y | 2 |
| Z | 2 |

- Learn table entries based on source address
- Table is an optimization; need not be complete
- Always forward broadcast frames

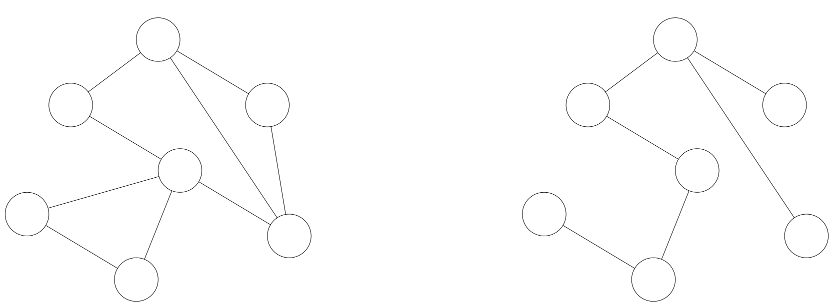
# Spanning Tree Algorithm

Extended LANs sometimes have loops



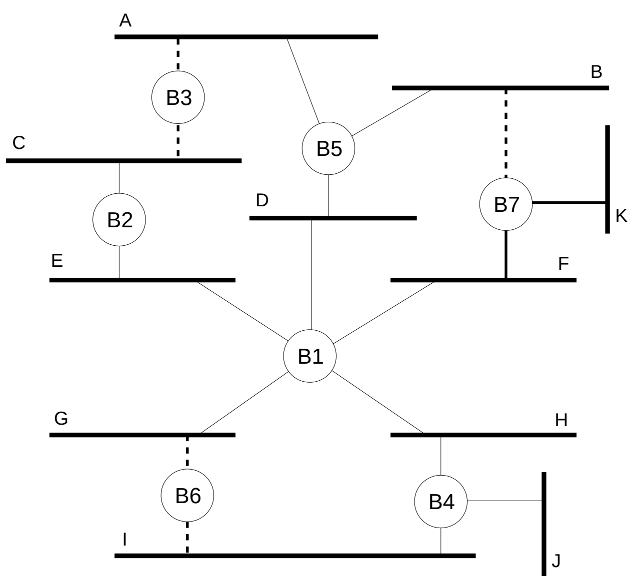

Bridges run a distributed spanning tree algorithm

- Select which bridges actively forward frames
- Developed by Radia Perlman at DEC
- Now IEEE 802.1 specification



# Algorithm Overview

- Each bridge has a unique id (e.g., B1, B2, B3)
- Select bridge with smallest id as root
- Select bridge on each LAN that is closest to the root as that LAN's designated bridge (use id to break ties)
- Each bridge forwards frames over each LAN for which it is the designated bridge

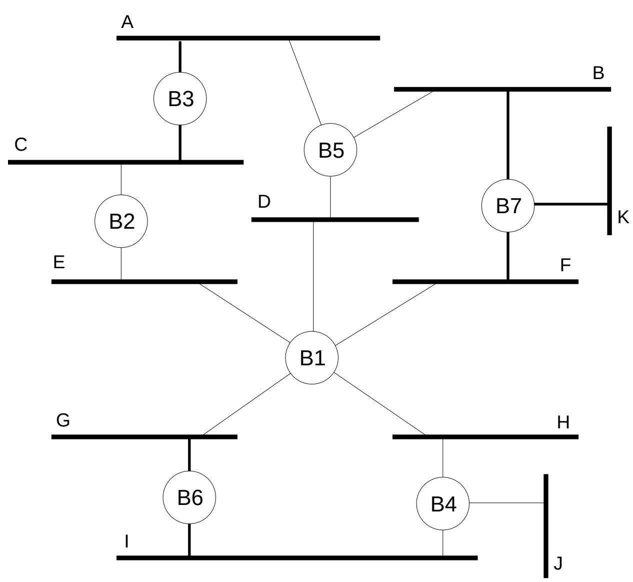


# Algorithm Detail

- Bridges exchange configuration messages
  - id for bridge sending the message
  - id for what the sending bridge believes to be root bridge
  - distance (hops) from sending bridge to root bridge
- Each bridge records current best configuration message for each port
- Initially, each bridge believes it is the root
- When learn not root, stop generating configuration message
  - in steady state, only root generates configuration messages
- When learn not designated bridge, stop forwarding configuration messages
  - in steady state, only designated bridges forward configuration messages
- Root bridge continues to send configuration messages periodically
- If any given bridge does not receive configuration message after a period of time, starts generating configuration messages claiming to be to be the root

# Broadcast and Multicast

- Forward all broadcast/multicast frames (current practice)
- Learn when no group members downstream
- Accomplished by having each member of group G send a frame to bridge multicast address with G in source field



# Limitations of Bridges

- Do not scale
  - spanning tree algorithm does not scale
  - broadcast does not scale
- Do not accommodate heterogeneity
- Caution: beware of transparency

# Route Propagation

Idea: Impose a second hierarchy on the network that limits what routers talk to each other. (The first hierarchy is the address hierarchy that governs how packets are forwarded.)

- Autonomous System (AS)
  - corresponds to an administrative domain
  - examples: University, company, backbone network
  - assign each AS a 16-bit number
- Two-level route propagation hierarchy
  - interior gateway protocol (each AS selects its own)
  - exterior gateway protocol (Internet-wide standard)

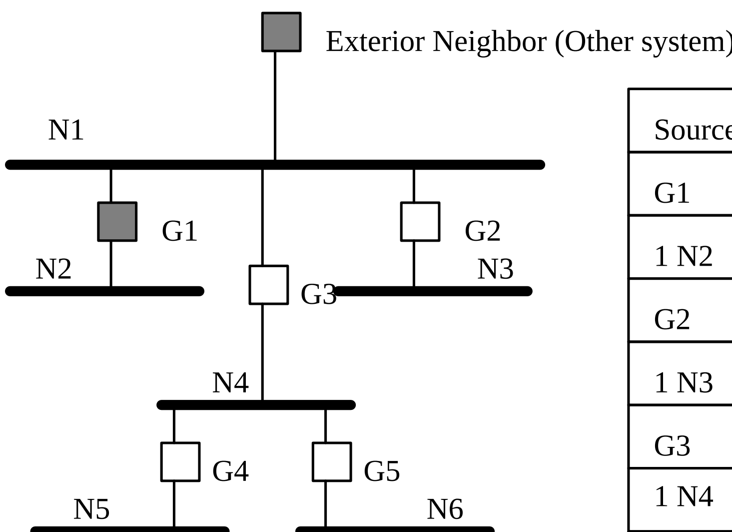
# Popular Interior Gateway Protocols

- RIP: Route Information Protocol
  - developed for XNS
  - distributed with Unix
  - distance-vector algorithm
  - based on hop-count
- OSPF: Open Shortest Path First
  - recent Internet standard
  - uses link-state algorithm
  - supports load balancing
  - supports authentication

# EGP: Exterior Gateway Protocol

- Overview
  - designed for tree-structured Internet
  - concerned with *reachability*, not optimal routes
- Protocol messages
  - neighbor acquisition: one router requests that another be its peer; peers exchange reachability information
  - neighbor reachability: one router periodically tests to see if the other router is still reachable; exchange HELLO/ACK messages; uses a k-out-of-n rule
  - routing updates: peers periodically exchange their routing tables (distance-vector)

# EGP Example




| Source Net: N1 |
|---|
| G1 |
| 1 N2 |
| G2 |
| 1 N3 |
| G3 |
| 1 N4 |
| 2 N5 |
| 2 N6 |

# BGP-4: Border Gateway Protocol

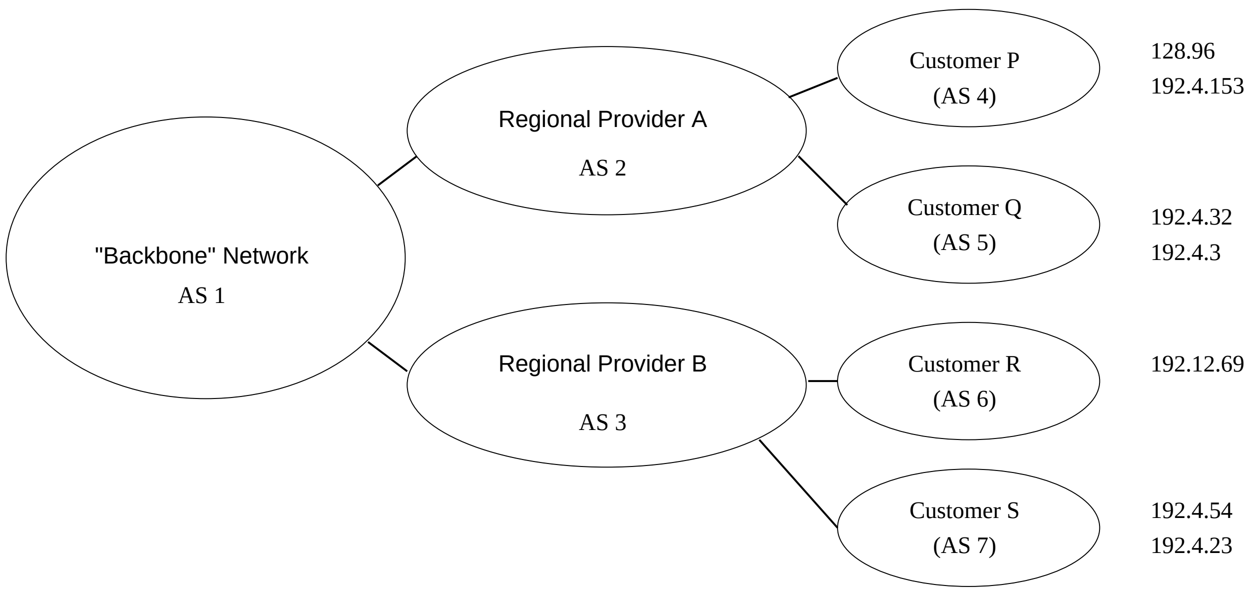
Assumes the Internet is an arbitrarily interconnected set of AS's. Define*local traffic* as traffic that originates at or terminates on nodes within an AS, and *transit traffic* as traffic that passes through an AS, we can classify AS's into three types:

- Stub AS: an AS that has only a single connection to one other AS; such an AS will only carry local traffic.
- Multihomed AS: an AS that has connections to more than one other AS, but refuses to carry transit traffic.
- Transit AS: an AS that has connections to more than one other AS, and is designed to carry both transit and local traffic.

Each AS has:

- One or more border routers
- One BGP *speaker* that advertises:
  – local networks
  – other reachable networks (transit AS only)
  – gives *path* information

# BGP Example




- Speaker for AS 2 advertises reachability to P and Q

  Network 128.96, 192.4.153, 192.4.32, and 192.4.3, can be reached directly from AS 2.

- Speaker for backbone network then advertises

  Networks 128.96, 192.4.153, 192.4.32, and 192.4.3 can be reached along the path ⟨AS 1, AS 2⟩.

- Speaker can also cancel previously advertised paths