ECE 463
Introduction to Computer Networks

Lecture: LAN Interconnects

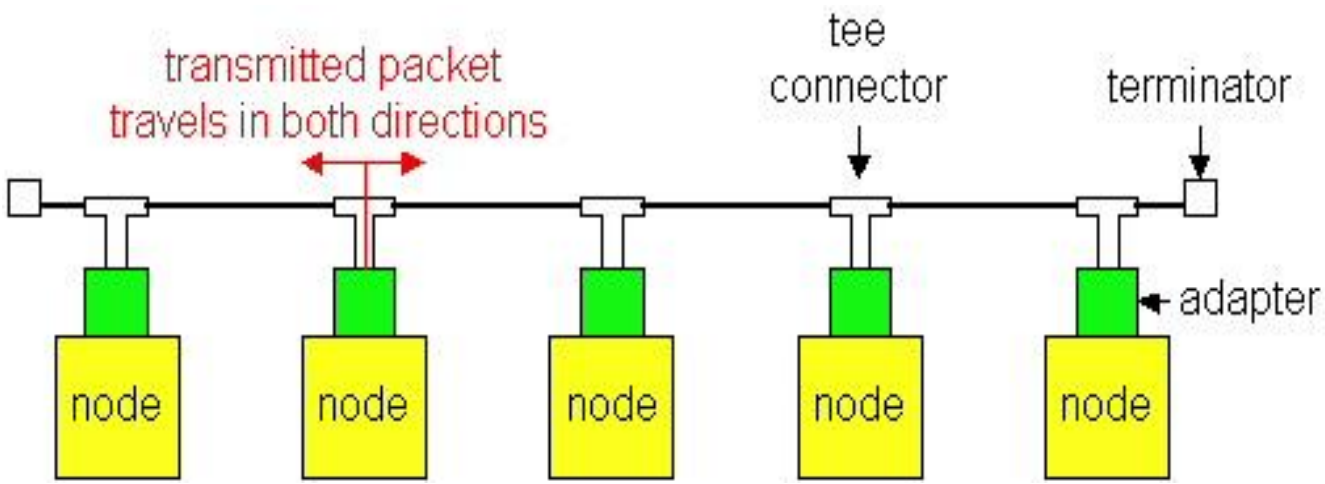
Sanjay Rao

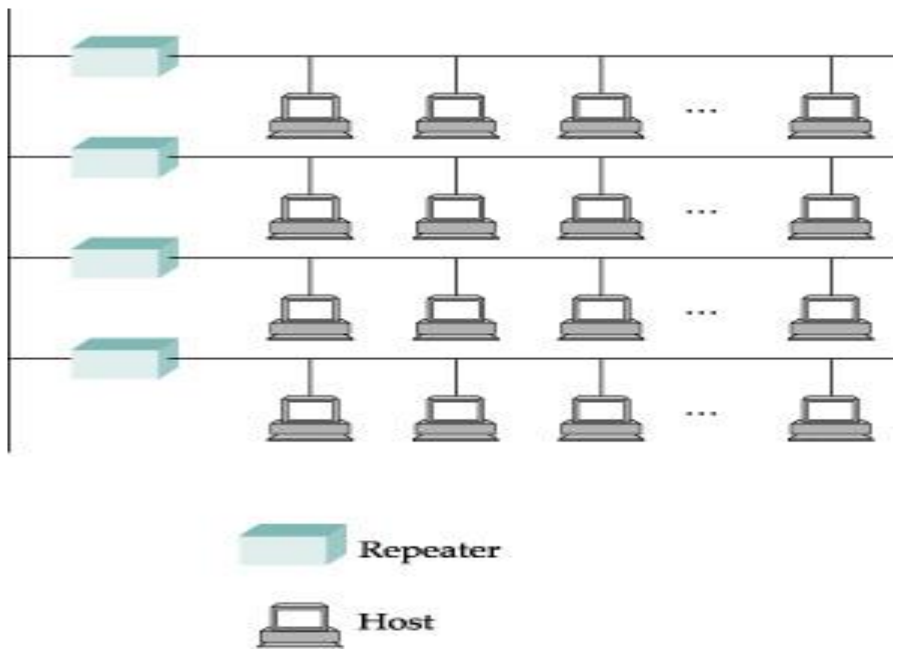
Ethernet

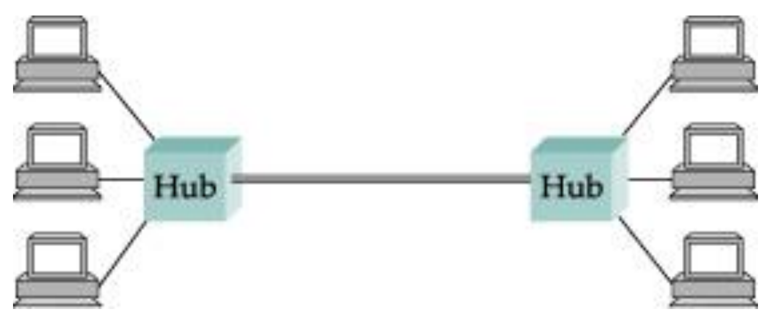
- Chief Advantages:
  - Easy to manage.
  - Cheap
- Concerns:
  - Does not work well under high load
  - Cannot work with too many nodes in system
  - Limitation on the physical distance between the nodes
- How do we get entire Purdue campus connected?

Ethernet Interconnects

- Repeaters/Hubs
  - Dumb physical layer device that forwards digital signals
  - Devices on either side in the same collision domain
- Bridges/Switches:
  - More intelligent devices that forward data only to hosts needing them.

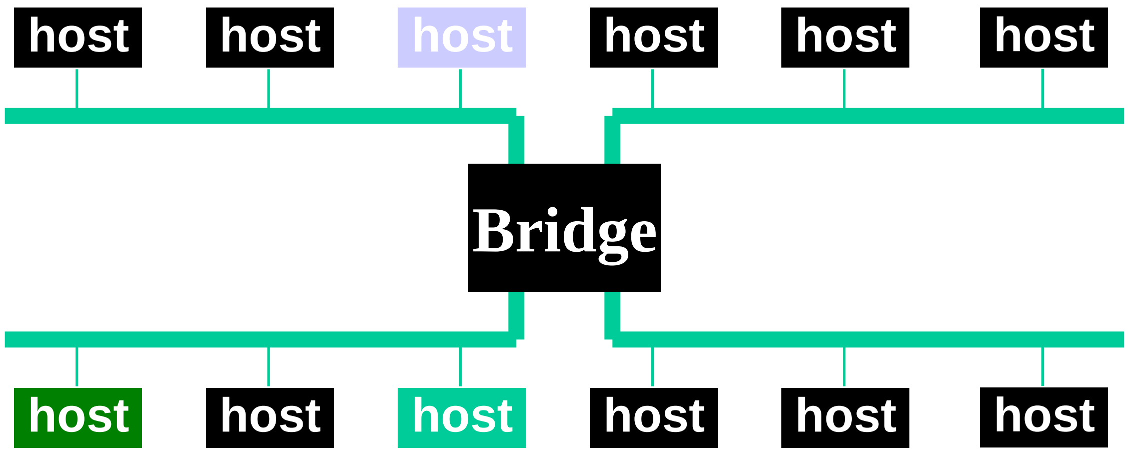

transmitted packet
travels in both directions
tee
connector
terminator
adapter
node
node
node
node
node


Repeater
Host


Hub
Hub

Building Larger LANs: Bridges

- Bridges connect multiple IEEE 802 LANs at layer 2.
  - Only forward packets to the right port
  - Reduce collision domain compared with single LAN



Transparent Bridges

- Overall design goal: **Complete transparency**
  - "Plug-and-play"
  - Self-configuring without hardware or software changes
  - Bridges should not impact operation of existing LANs

- Three parts to transparent bridges:

  (1) Forwarding of Frames

  (2) Learning of Addresses

  (3) Spanning Tree Algorithm

Frame Forwarding

- Each bridge maintains a **forwarding database** with entries

  `< MAC address, port, age>`

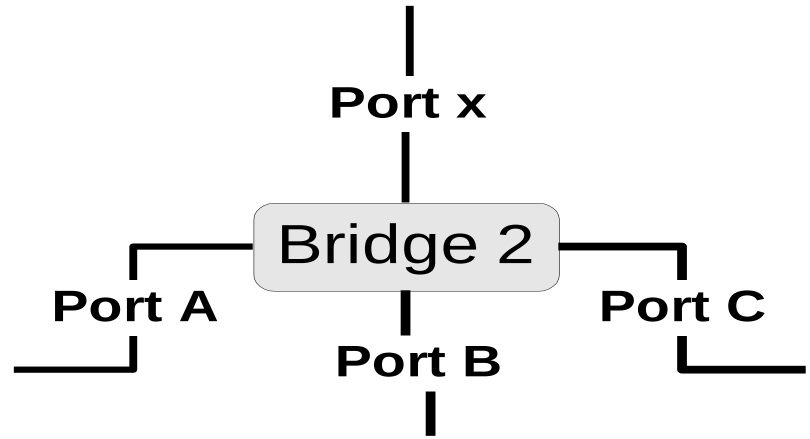
`MAC address:`	host name or group address
`port:`	port number of bridge
`age:`	aging time of entry

with interpretation:

- a machine with **`MAC address`** lies in direction of the **`port`** number from the bridge. The entry is **`age`** time units old.

Frame Forwarding 2

- Assume a MAC frame arrives on port x.



Search if MAC address of destination is listed for ports A, B, or C.

Found?

Forward the frame on the appropriate port

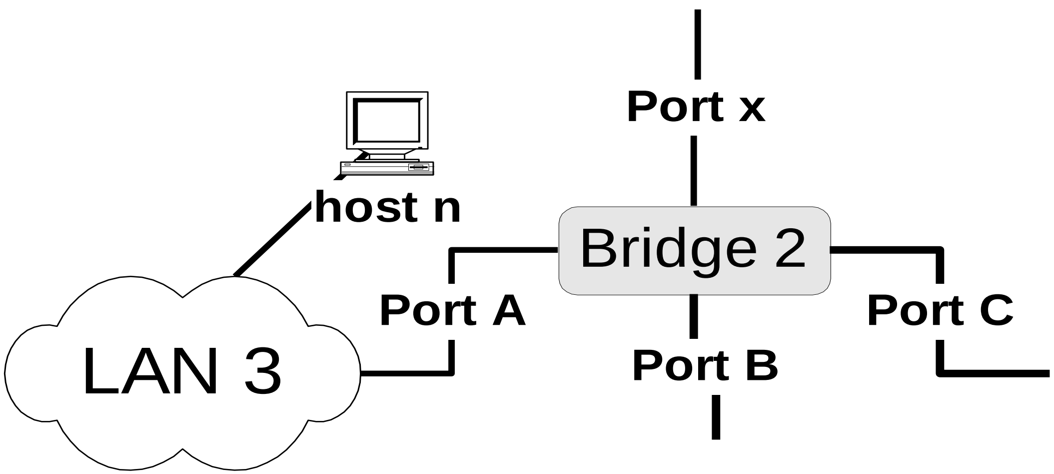
Not found ?

Flood the frame, i.e., send the frame on all ports except port x.

Address Learning

- In principle, the forwarding database could be set statically (=static routing)
- In the 802.1 bridge, the process is made automatic with a simple heuristic:

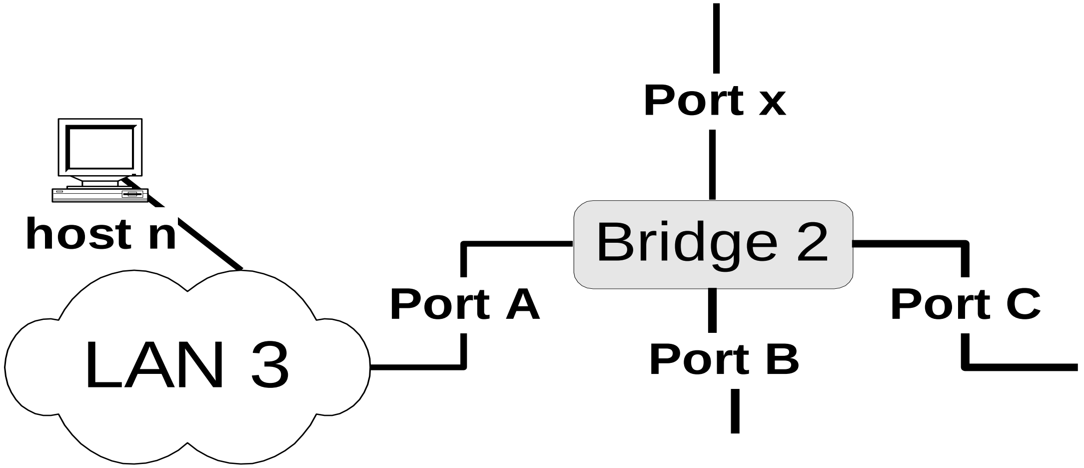
The source field of a frame that arrives on a port tells which hosts are reachable from this port.



Address Learning 2

Algorithm:

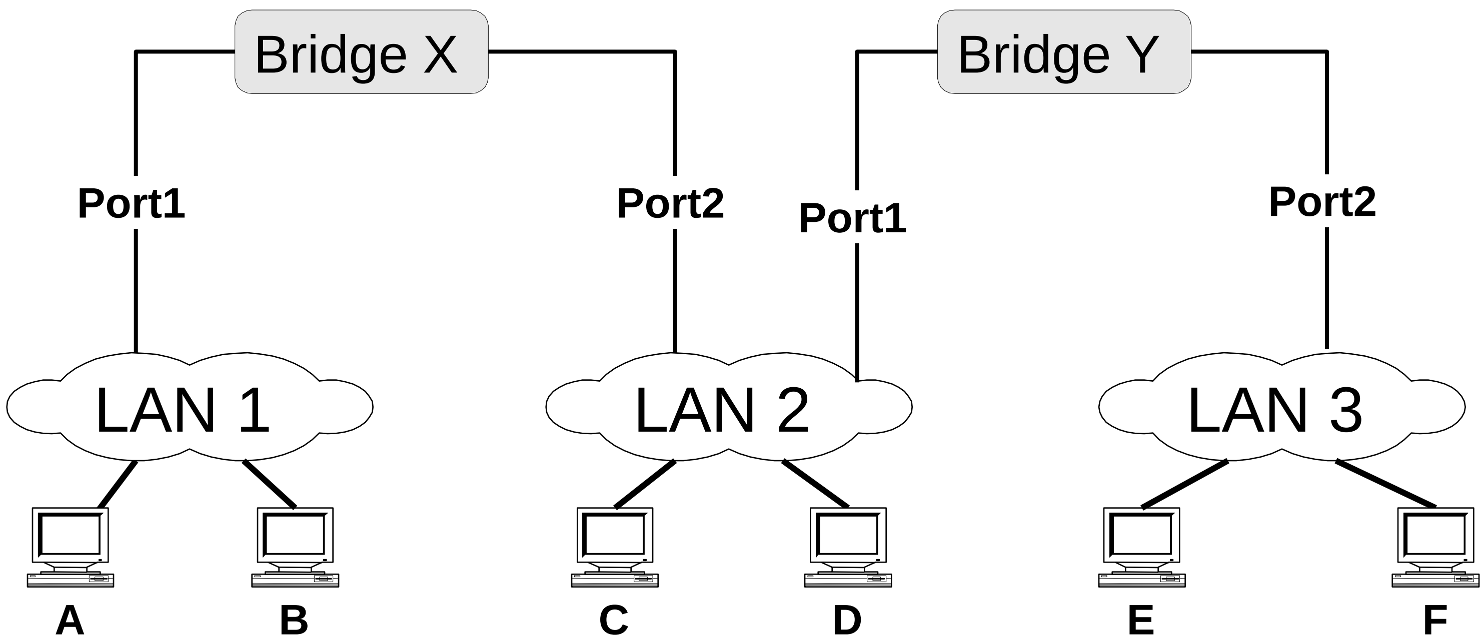
- For each frame received, the source stores the source field in the forwarding database together with the port where the frame was received.
- All entries are deleted after some time



Example

•Consider the following packets:
<Src=A, Dest=F>,   <Src=C, Dest=A>, <Src=E, Dest=C>

•What have the bridges learned?



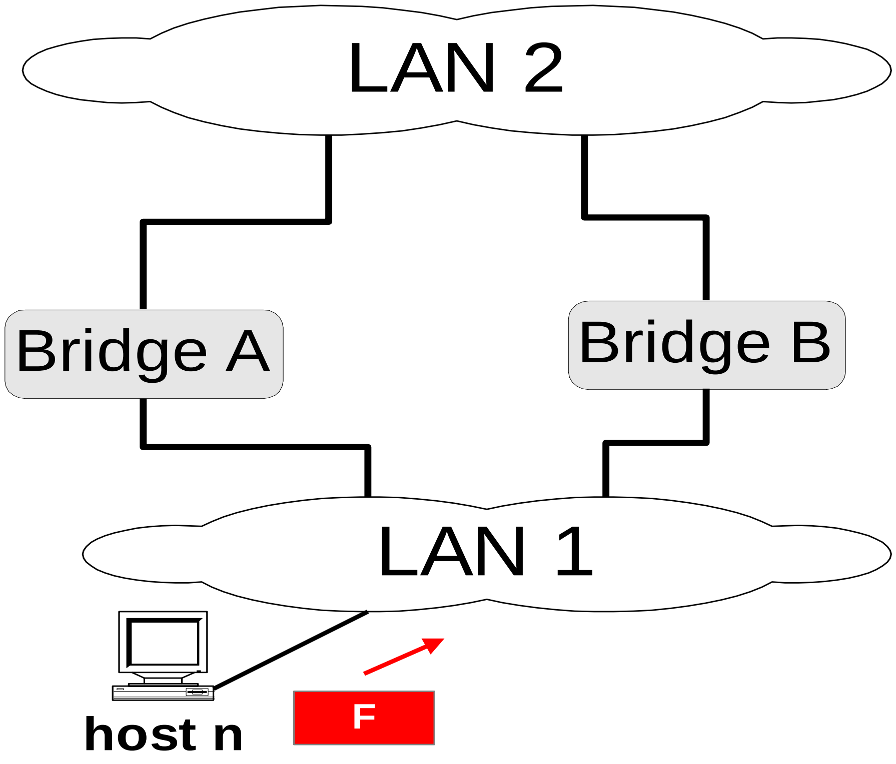
Issue

- Consider the two LANs that are connected by two bridges.
- Assume *host n* is transmitting a frame F with unknown destination.

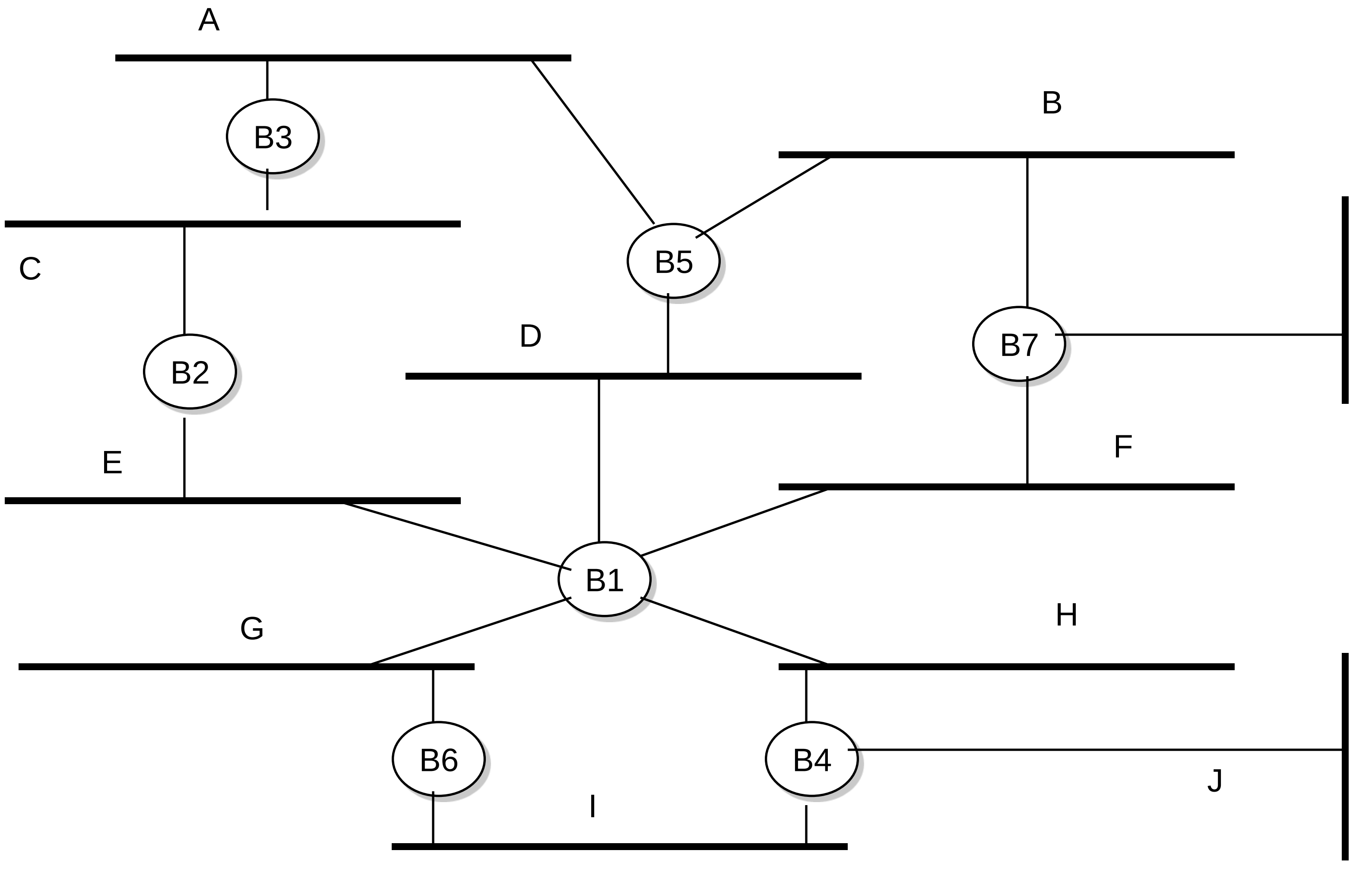
What is happening?

- Bridges A and B flood the frame to LAN 2.
- Bridge B sees F on LAN 2 (with unknown destination), and copies the frame back to LAN 1
- Bridge A does the same.
- The copying continues

Where's the problem? What's the solution ?



A Bridged Network



Spanning Trees

- The solution to the loop problem is to not have loops in the topology
- IEEE 802.1 has an algorithm that builds and maintains a spanning tree in a dynamic environment.
- Bridges exchange messages to configure the bridge (Configuration Bridge Protocol Data Unit, Configuration BPDUs) to build the tree.

What do the BPDUs do?

With the help of the BPDUs, bridges can:

- Elect a single bridge as the **root bridge**.
- Calculate the distance of the shortest path to the root bridge
- Each LAN can determine a **designated bridge**, which is the bridge closest to the root. The designated bridge will forward packets towards the root bridge.
- Each bridge can determine a **root port**, the port that gives the best path to the root.
- Select ports to be included in the spanning tree.

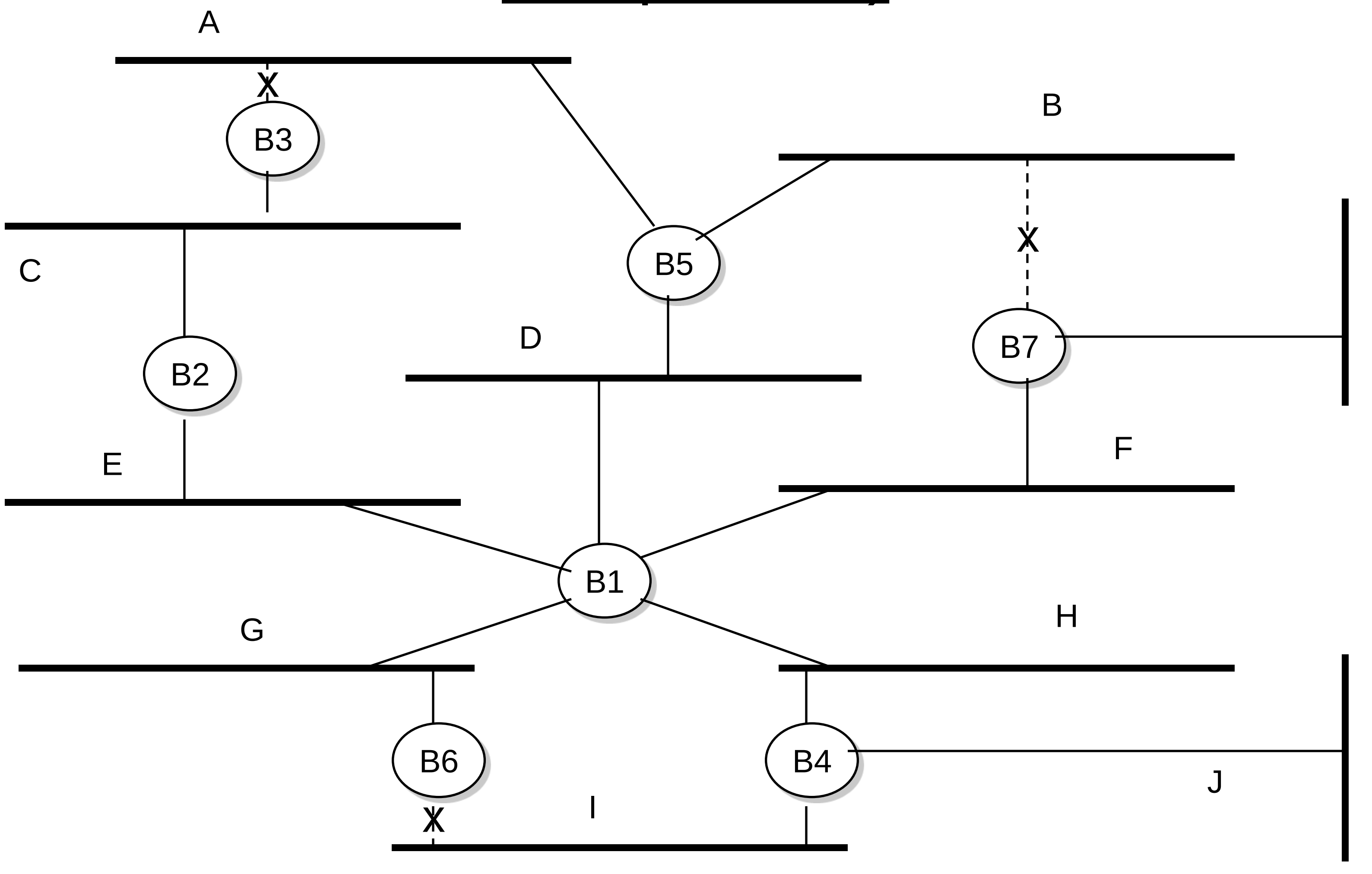
Concepts

- Each bridge has a unique identifier: Bridge ID
- Each port within a bridge has a unique identifier (port ID).

- **Root Bridge:** The bridge with the lowest identifier is the root of the spanning tree.
- **Root Port**: Each bridge has a root port which identifies the next hop from a bridge to the root.
- **Root Path Cost**: For each bridge, the cost of the min-cost path to the root, measured in #Hops to the root.
- **Designated Bridge**
  - Bridge on LAN that provides minimal cost path to root for this LAN
  - if two bridges have the same cost, select one with lower identifier
- **Designated Port:**
  - If a bridge is the designated bridge for a LAN, the appropriate port is the designated port.

- Bridge disables all ports which are neither root ports, nor designated ports.

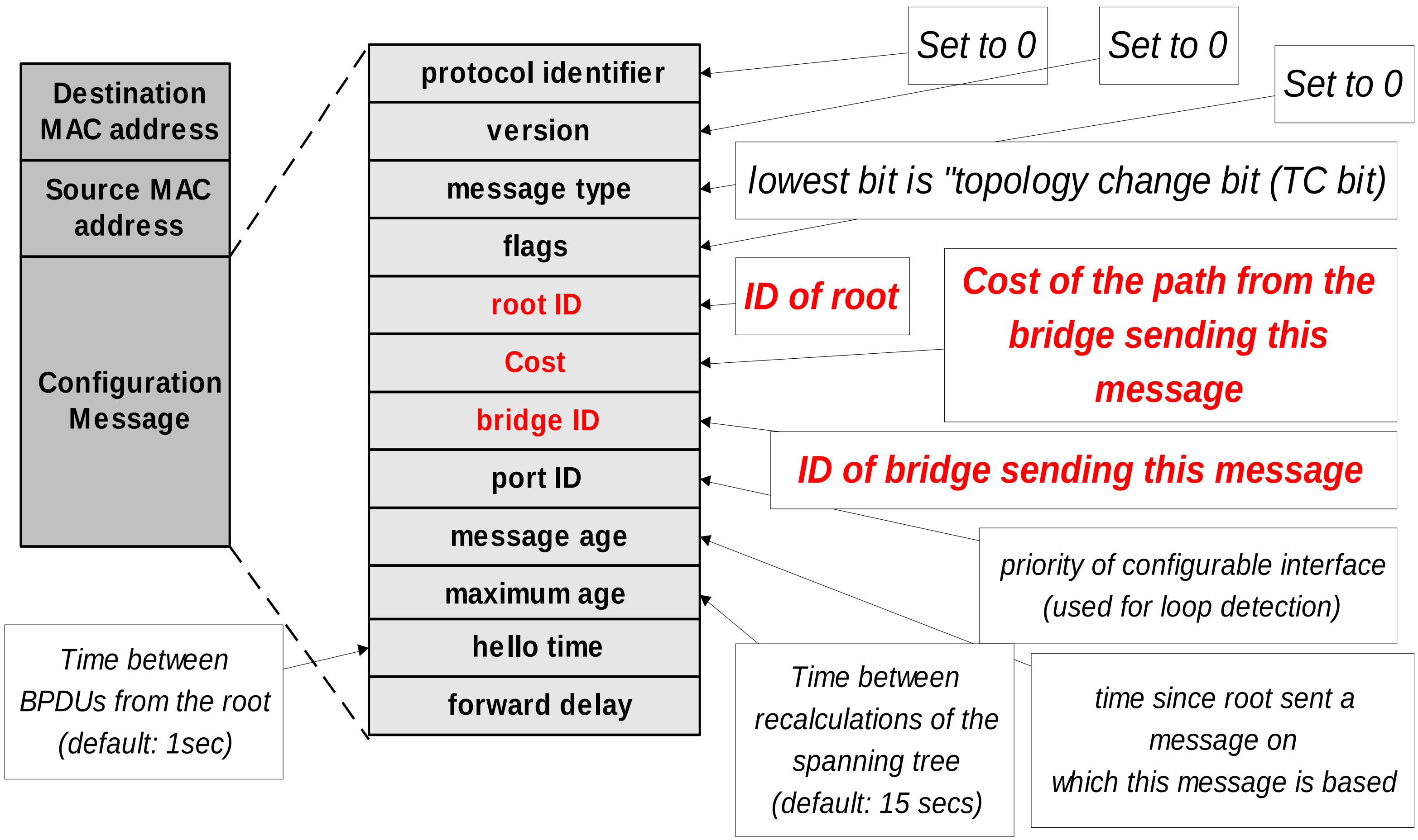
A Bridged Network (End of Spanning Tree Computation)



Centralized Vs. Distributed

- Easy to see how this algorithm works if we have the full picture.
- But in reality, each bridge has very limited information
- Need for Distributed algorithms
  - Bridges exchange messages with each other
  - Messages enable them to learn information needed to make the right decisions.
- Topology may not be correct initially, eventually gets to the right one.

Configuration BPDUs



Steps of Spanning Tree Algorithm

1. Determine the root bridge

2. Determine the root port on all other bridges

3. Determine the designated port on each LAN

- Each bridge is sending out BPDUs that contain the following information:

root ID	cost	bridge ID/port ID

root bridge (what the sender thinks it is)
root path cost for sending bridge
Identifies sending bridge

Initialization and Operation

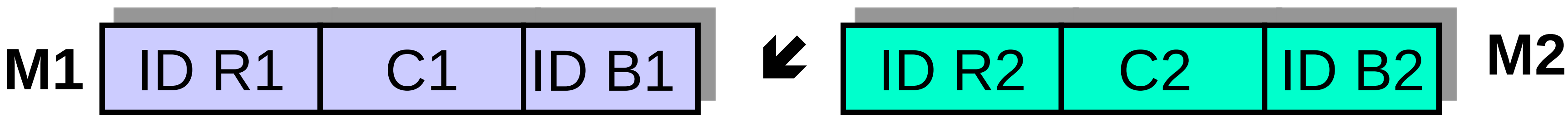
- Initially, all bridges assume they are the root bridge.
- Each bridge B sends BPDUs of this form on all its ports:

B	0	B

- Upon receiving BPDU on a port
  - Bridge checks if received BPDU is “better” than best recorded for that port (initially its own BPDU is best).
  - If “better”, it discards old information for that port, and saves new information, adding 1 to the distance field.

Ordering of Messages

- We can order BPDU messages with the following ordering relation "⬋" (let's call it "lower cost" or "better"):



If (R1 < R2)

M1 ⬋ M2

elseif ((R1 == R2) and (C1 < C2))

M1 ⬋ M2

elseif ((R1 == R2) and (C1 == C2) and (B1 < B2))

M1 ⬋ M2

else

M2 ⬋ M1

Determining Root Bridge, Root Port

- If bridge receives message from lower bridge id:
  - It knows it is not the root.
  - It updates its belief of who the root is, say R.
- Bridge B determines the Root Path Cost (Cost) as follows:
  - *If B = R :* Cost = 0.
  - *If B ☼ R:* Cost = {Smallest Cost in any of BPDUs that were received from R} + 1
- **B's root port** is the port from which B received the lowest cost path to R (in terms of relation "⬋").
- Knowing R and Cost, B can generate its BPDU (but will not necessarily send it out):

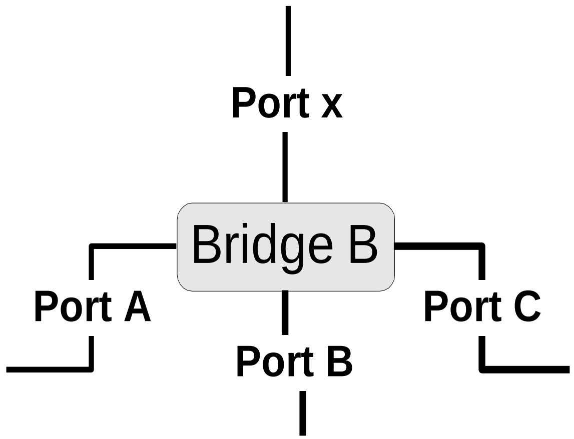
R	Cost	B

Calculate the Root Path Cost Determine the Root Port

- At this time: B has generated its BPDU

R	Cost	B

- B will send this BPDU on one of its ports, say **port x**, only if its BPDU is lower (via relation "⬋") than any BPDU that B received from port x.
- In this case, B also assumes that it is the **designated bridge** for the LAN to which the port connects.



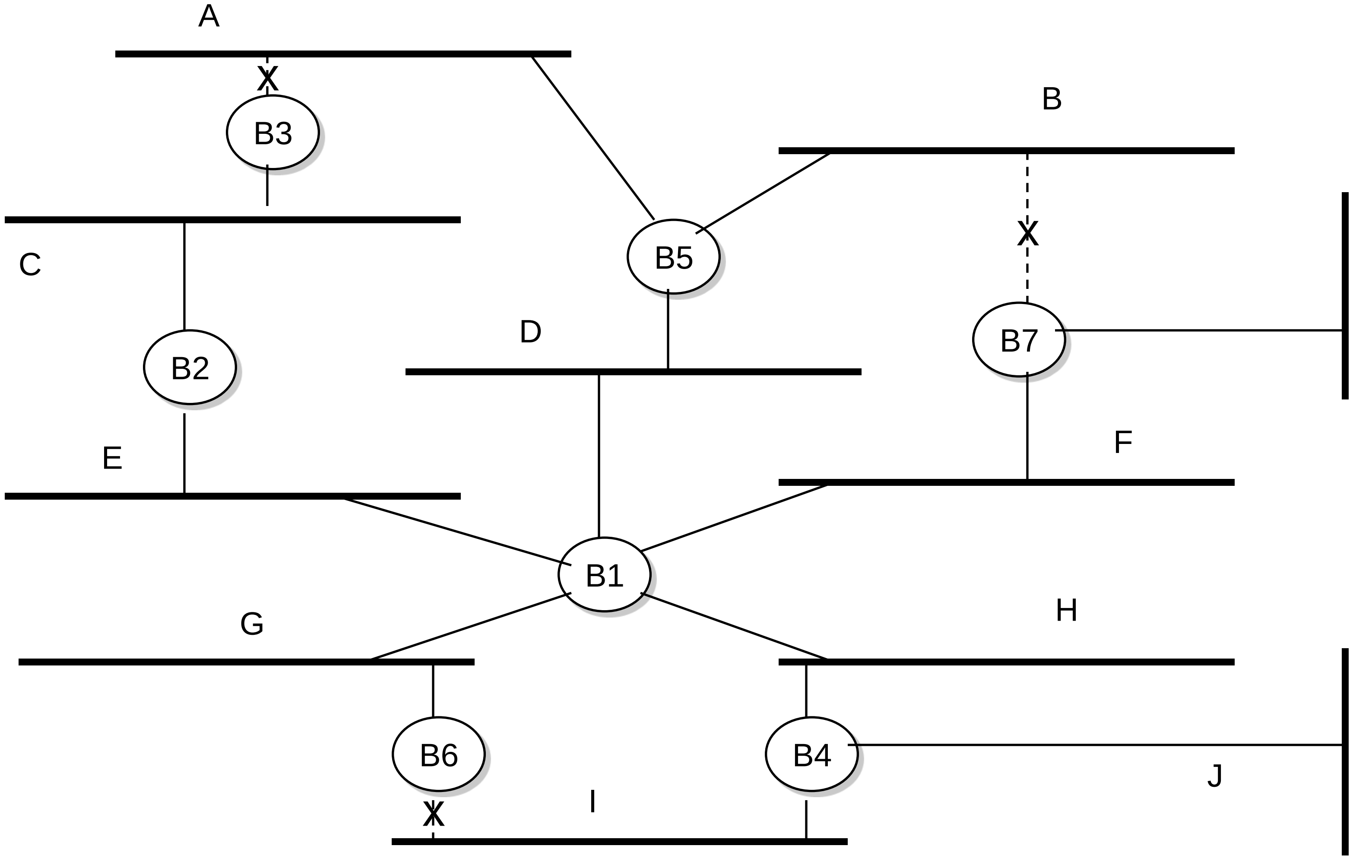
Selecting the Ports for the Spanning Tree

- At this time: Bridge B has calculated the root, the root path cost, and the designated bridge for each LAN.
- Now **B can decide which ports are in the spanning tree:**
  - B's root port is part of the spanning tree
  - All ports for which B is the designated bridge are part of the spanning tree.
- B's ports that are in the spanning tree will forward packets **(=forwarding state)**
- B's ports that are not in the spanning tree will not forward packets **(=blocking state)**

Example

- B3 rcvs (B2,0,B2). What happens?
- Makes B2 root. Send what to B5?
  - Sends (B2,1,B3) to B5
- Meanwhile B2 accepts B1 as root
  - Sends (B1,1,B2) to B3
- B3 updates its root to: B1
- Meanwhile B5 accepts B1 as root
  - Sends (B1,1,B5) to B3
- B3: no change to value of root

A Bridged Network (End of Spanning Tree Computation)



Ethernet Switches

- Bridges make it possible to increase LAN capacity.
  - Packets are no longer broadcasted - they are only forwarded on selected links
  - Adds a switching flavor to the broadcast LAN
- Ethernet switch is a special case of a bridge: each bridge port is connected to a single host.
  - Simplifies the protocol and hardware used (only two stations on the link) – no longer full CSMA/CD

Can the Internet be One Big Switched Ethernet?

- Inefficient
  - Too much flooding
- Explosion of forwarding table
  - Need to have one entry for every Ethernet address in the world!
- Poor performance
  - Tree topology does not have good load balancing properties
  - Hot spots
- Etc…