Computer Networks
B.Tech, Computer Engineering,
Delhi Technological university
Module -2.5
Ethernet and switching
Instructor: Divya Sethia

1

Objective

•Ethernet 802.3
  - Cabling
  - Frame Format
•Switching

2

Ethernet

- The Ethernet MAC Sublayer Protocol
- The Binary Exponential Backoff Algorithm

Repeaters

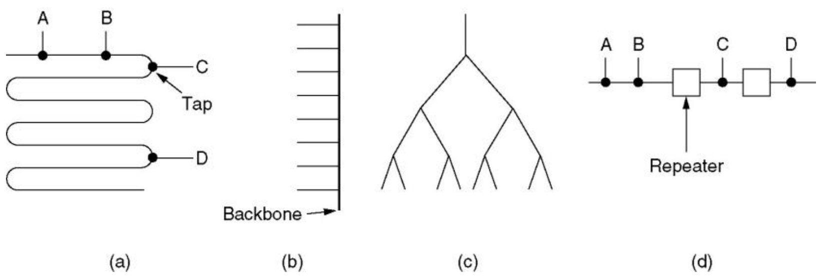
•Allow larger networks, multiple cables can be connected by **repeaters**

•Receives, amplifies (regenerates), and retransmits signals in both directions

•Multiple cable segments and multiple repeaters

•No two transceivers may be more than 2.5 km apart and no path between any two transceivers may traverse more than four repeaters

Ethernet Cabling



Cable topologies. (a) Linear, (b) Spine, (c) Tree, (d) Segmented.

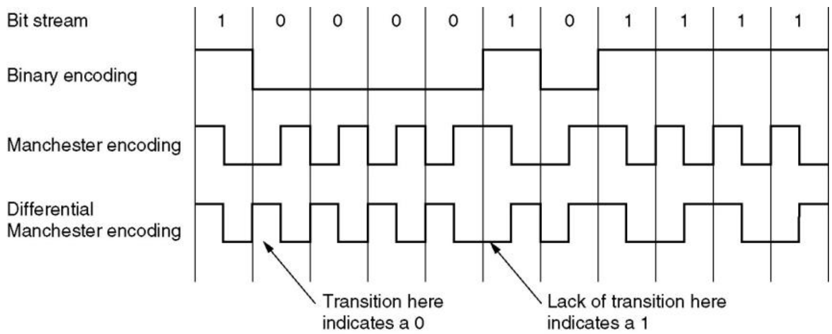
Ethernet

a) IEEE standard 802.3 series.
b) Developed at Xerox PARC in early 1970's
c) Versions:
   - Traditional ("classic"): 10 Mb/s
   - Fast: 100 Mb/s
   - Gigabit: 1Gb/s

d) Original intention was that the Ethernet variants would be compatible for physical layer signalling, but this has not happened.

Divyashikha Sethia (DTU)

Ethernet Encoding



(a) Binary encoding, (b) Manchester encoding,
(c) Differential Manchester encoding.

Divyashikha Sethia (DTU)

Ethernet MAC Sublayer Protocol

Frame formats. IEEE 802.3.

Preamble	SOF	Destination address	Source address	Length	Data	Pad	Check-sum

Preamble: 7 bytes of alternating pattern of one and zeros (10101010s)
- to allow Receiver to synchronize with incoming transmission

•**Start of Frame (SOF):** 1 byte ends with two consecutive 1 bits(10101011 is the start flag in Ethernet IT helps to synchronize the frame reception portions of all stations on the LAN. Also known as **SFD Start Frame Delimiter.**

Note: There is no provision for ACK/NAK
This must be implemented at a higher layer ⇒ Ethernet is an unreliable medium

Ethernet MAC Sublayer Protocol

Frame formats. IEEE 802.3.

Destination address (DA): It consists of 6 bytes. The DA field identifies which station(s) should receive the frame.

Source addresses (SA): It consists of 6 bytes. The SA field identifies the sending station.

I/G = 0 in case of individual address
= 1 in case of group address
U/L = 0 in case of global administered address
= 1 in case of local administered address
Unicast: Defines a single destination
Broadcast: FFFFFFFFFFFF each station on the network receive and accept frames
Multicast: A group address defines multiple recipient

Hardware Addresses

Hardware addresses must be unique on a LAN
Example: **`00-0A-E4-09-85-2B`**

How can those address be assigned and who is responsible for uniqueness?

- Static
  - Hardware manufacturer assigns permanent address to each interface. Manufacturer must ensure every interface has a unique address
- Dynamic
  - Address can be set by end user, either through switches or jumpers on the interface or through software
  - System administrators must coordinate to avoid conflict
- Automatic
  - Interface automatically assigns hardware address each time it is powered up
  - Automatic scheme must be reliable to prevent conflicts

Broadcasting

•Some applications want to broadcast messages to all stations on the LAN
•Shared communication channel can make broadcast efficient - message is delivered to all stations
•Special broadcast address used to identify broadcast messages, which are captured by all stations

Ethernet MAC Sublayer Protocol

Frame formats. IEEE 802.3.

•**Length:** Indicates the number of bytes of data that follows this field.

•**Data**: It is a sequence of n bytes of any value, where n is less than or equal to 1500

• **Frame check sequence (FCS):** It consists of 4 bytes. This sequence contains a 32-bit cyclic redundancy check (CRC) value, which is created by the sending MAC and is recalculated by the receiving MAC to check for damaged frames. The FCS is generated over the DA, SA, Length/Type, and Data fields.

Maximum Data limit

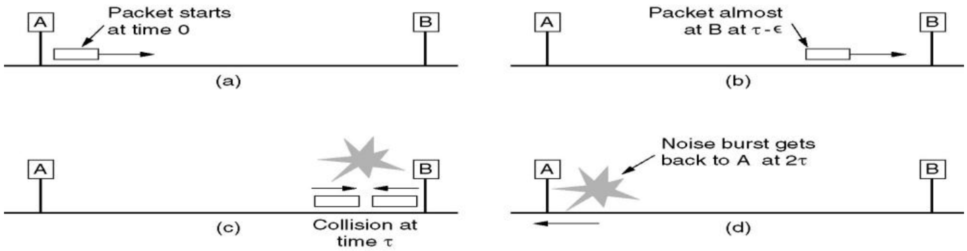
•Maximum limit: 1500 bytes

•This limit was chosen somewhat arbitrarily at the time the DIX standard was cast in stone, mostly based on the fact that a transceiver needs enough RAM to hold an entire frame and RAM was expensive in 1978. A larger upper limit would have meant more RAM, hence a more expensive transceiver.

Minimum Ethernet Data limit

•On collision transreceiver truncates current frame – stray bits of frames on the cable
•Need to distinguish 0 byte data frames from garbage – Hence valid frame of at least 64 bytes from destination to checksum both inclusive.
•Use Pad field to round frame to 64 byte
•

Minimum Ethernet Data limit



Collision detection can take as long as 2τ.

•Station tries to transmit short frame, and collision occurs, but the transmission completes before the noise burst gets back at 2τ.
•The sender will incorrectly conclude that the frame was successfully sent.
•Prevention: frames must take more than 2τ to send so that the transmission is still taking place when the noise burst gets back to the sender.
• 10-Mbps LAN with a maximum length of 2500 meters and four repeaters worst case round-trip time: 50 μsec
•Time for a bit at 10 Mbps : 100 nsec
•=> 500 bits min frame size rounded to 512 bits = 64 bytes

Binary Exponential Backoff Algorithm

•After a collision, time is divided into discrete slots whose length is equal to the worst-case round-trip propagation time on the ether (2τ). To accommodate the longest path allowed by Ethernet, the slot time has been set to 512 bit times

•first collision, each station waits either 0 or 1 slot times before trying again. If two stations collide and each one picks the same random number, they will collide again

• second collision, each one picks either 0, 1, 2, or 3 at random and waits that number of slot Times

•general, after *i collisions, a random number between 0 and 2i - 1 is chosen, and that* number of slots is skipped.

Variants of Ethernet

a) 10Mb/s (“10base”):
- Manchester encoding for physical layer
- 10Base5: (IEEE 802.3)
  - thick co-axial cable bus
  - maximum length: 500 metres
  - minimum ~2.5 metres between stations
- 10Base2: (IEEE 802.3a)
  - thin co-axial cable bus
  - maximum length: 185 metres
  - minimum 0.45 metres between stations
- 10BaseT: (IEEE 802.3i)
  - UTP cable with central hub, star topology
  - maximum length 100 metres

Variants of Ethernet

a) 100Mb/s (“100base”, “Fast Ethernet”) (IEEE 802.3u)
- 100BaseTX: Uses 2 wires in UTP cable
  - maximum length: 100 metres
  - Encoding: 4B/5B plus MLT-3
  - MLT-3: multi-level line transmission, 3-level
    - cycles through states in order -, 0, +, 0, -, etc.
    - 0 data bit: stay at current state
    - 1 data bit: move to next state in cycle
- 100BaseFX: Uses optical fibre
  - maximum length: 136 metres
  - Encoding: 4B/5B plus NRZI
- 100BaseT4:
  - UTP cable with central hub, star topology
  - maximum length 100 metres
  - Encoding: 8B/6T

Variants of Ethernet

a) 1000Mbs (“1000base”, “Gigabit ethernet”) (IEEE 802.3u)
   - 1000BaseT: (IEEE 802.3ab)
     - Uses 4 pairs of UTP cable
     - maximum length: 100 metres
     - Encoding: complex scheme called 4D-PAM5
   - 1000BaseLX: (IEEE 802.3z)
     - Uses long-wave optical fibre
     - maximum length: 550 to 5000 metres, depending on fibre
     - Encoding: 8B/10B + NRZ
   - 1000BaseSX:(IEEE 802.3z)
     - Uses short-wave optical fibre
     - maximum length: 220 to 550 metres, depending on fibre
     - Encoding: 8B/10B + NRZ

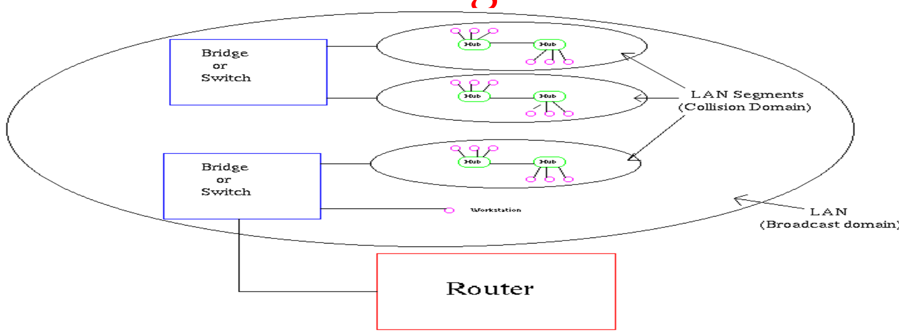
LAN

- Is a network of computers located within the same area
- Defined as a single **broadcast domain** - user broadcasts information on his/her LAN, the broadcast will be received by every other user on the LAN

- Broadcasts prevented to leave LAN using router

- Bridges and switches will not forward collisions, but will allow broadcasts (to every user in the network) and multicasts (to a pre-specified group of users) to pass through

LAN segment



- Workstations, hubs, and repeaters together form a LAN segment. A LAN segment is also known as a **collision domain** since collisions remain within the segment
- Area within which broadcasts and multicasts are confined is called a **broadcast domain** or LAN.
- LAN can consist of one or more LAN segments
- Everyone on a LAN must be located in the same area

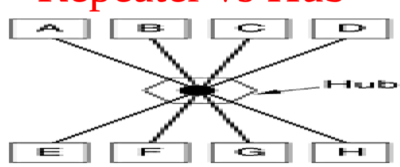
Interconnecting LANs (1)

a) Why not one big LAN?
   - Administrative separation (e.g. separate domain for SITE versus university-wide)
   - Limited amount of supportable traffic: on a single LAN, all stations must share bandwidth
   - Limited length: Ethernet has maximum cable length in IEEE802.3 standard
   - Limited number of stations: token passing delays at each station in IEEE802.4, IEEE802.5
   - Reliability: containment of potential faults
   - Security: differing levels of security access, confidentiality of data

Interconnecting LANs (2)

a) Bridges versus Repeaters
- Repeater:
  - Copies (regenerates) bits between LAN segments.
  - Physical layer interconnection of LANs.
  - Transparent to LAN functionality
- Bridge:
  - Receives and stores/forwards (when appropriate) packets between LANs (usually using same technology for physical and data link layers)
  - Has several layers of protocol stack: physical, data link, possibly network layers.
  - Acts as a filter.

Repeater vs Hub



• A hub has a number of input lines that it joins electrically. Frames arriving on any of the lines are sent out on all the others. If two frames arrive at the same time, they will collide, just as on a coaxial cable.

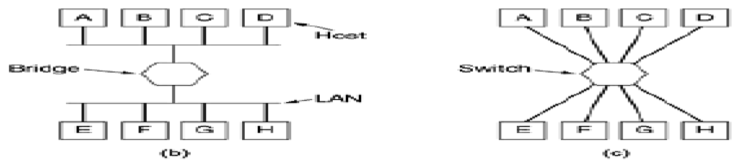
•Entire hub forms a single collision domain. All the lines coming into a hub must operate at the same speed.

•Hubs differ from repeaters :
- do not (usually) amplify the incoming signals
- designed to hold multiple line cards each with multiple inputs (multi port)

• Like repeaters, hubs do not examine the 802 addressesor use them in any way.

Bridge vs Switch

•Similarity : both route on frame addresses
•switch is most often used to connect individual computers



(a) host *A* wants to send a frame to host *B, the bridge gets the frame but just discards it*
(b) switch must actively forward the frame from *A to B because there* is no other way for the frame to get there

Switch
1. Connects multiple computers
Each line card provides buffer space

for frames arriving on its ports. Since
Each port is its own collision domain,
switches never lose frames to collisions.
If frames come in faster than they can be retransmitted,
the switch may run out of buffer space
and have to start discarding frames

Bridge
1. Connects LAN segments
2. Each LAN segment has a collision domain

Switch types

•**Store and forward switch:**
 - Store the entire frame and then forward based on the header information
 - If frames come faster than they can be retransmitted , buffers may be overfull and frames are discarded

• **Cut through switch**
 - switches start forwarding frames as soon as the destination header field has come in, but before the rest of the frame has arrived
 - done in hardware

Interconnection LANS(3)

a) Bridges versus Routers

– Routers:

•At routers the frame header and trailer are stripped off and the packet located in the frame's payload field is passed to the routing software.
•Routing software uses the packet header to choose an output line.
•For an IP packet, the packet header will contain a 32-bit (IPv4) or 128-bit (IPv6) address, but not a 48-bit 802 address.
• Des not see the frame addresses and does not even know whether the packet came in on a LAN or a point-to-point line

Interconnection Lans

Application layer	Application gateway
Transport layer	Transport gateway
Network layer	Router
Data link layer	Bridge, switch
Physical layer	Repeater, hub

Devices in different layers

Bridges

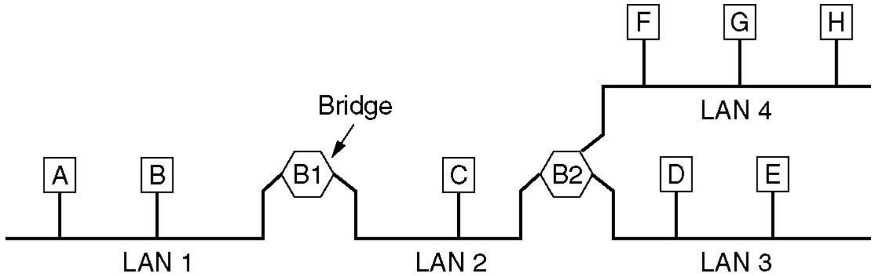
a) Properties:
   - No modification to frames
   - Need buffer space for anticipated traffic
   - Needs to know about addresses and routing
     - Specifically, which hardware addresses are on which LAN.

Types of Bridge Routing

a) Fixed:
   - Manually create table of addresses for each LAN connected to bridge

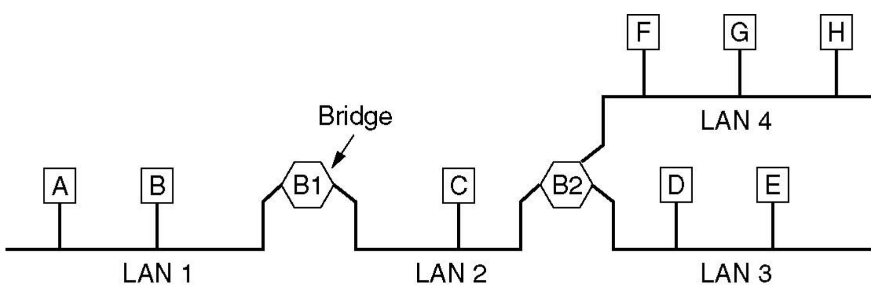
b) Adaptive / Learning (Transparent Bridge with Backward learning)
   - Bridge eventually determines how to route by examining source of sent messages
   - Use a lookup table that will match a destination address with a LAN segment

Local Internetworking



Bridge B1 sees:
Frame from C coming on LAN2 – maps C to LAN2 in B1 Hash Table
Subsequent frames sent from LAN1 for C will be forwarded to LAN2
Frames sent from LAN2 to C will be discarded

Local Internetworking



- Transparent bridge operates in promiscuous mode, accepting every frame transmitted on all the LANs to which it is attached.

- A frame arriving at bridge B1 on LAN 1 destined for A can be discarded immediately, because it is already on the correct LAN

- Frame arriving on LAN 1 for C or F must be forwarded.

Bridge Hash Table / MAC Table

•Destination address in a big (hash) table inside the bridge.

•Lists each possible destination and which output line (LAN) it belongs on.

•Hash table empty initially

•Flooding: every incoming frame for an unknown destination is output on all the LANs to which the bridge is connected except the one it arrived on

•Transparent bridge learn destination segments by **Backward Learning**

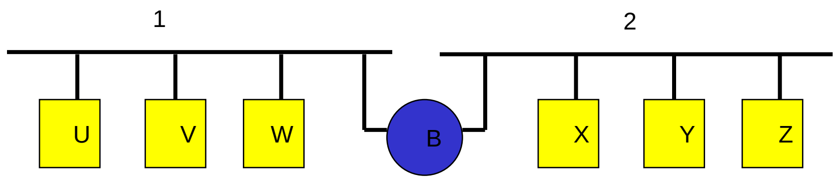
•Whenever a packet reaches a bridge it looks at the source address and the Lan segment the frame arrived. Enters in the Hash table the source MAC address with Lan segment information

•Hash Table is periodically updated and old entries are deleted

Adaptive Bridge Routing

a) When the bridge boots, the lookup table is initially empty
b) When a frame arrives at the bridge from a LAN segment, the bridge reads the source and destination addresses. Then:
   - If the destination cannot be found in the lookup table, or is the broadcast address, the frame is stored, and re-transmitted on all other LAN segments.
   - If the destination is found in the lookup table:
     - If it is on the same segment from which the frame arrived, no further action is taken.
     - If on a different segment, the frame is re-transmitted on the destination's segment only.
   - The frame on which the segment arrived is now known to be the segment for the sender's hardware address, and the lookup table updated.

Adaptive Routing Example



Event	Lookup table	Sent to:
Boot	empty	
U sends to V	(U,1)	2
V sends to U	(U,1), (V,1)	Not forwarded
Z broadcasts	(U,1), (V,1), (Z,2)	1
Y sends to V	(U,1), (V,1), (Z,2), (Y,2)	1
Y sends to X	(U,1), (V,1), (Z,2), (Y,2)	1
X sends to W	(U,1), (V,1), (Z,2), (Y,2), (X,2)	1
W sends to Z	(U,1), (V,1), (Z,2), (Y,2), (X,2), (W,1)	2
U sends to W	(U,1), (V,1), (Z,2), (Y,2), (X,2), (W,1)	Not forwarded

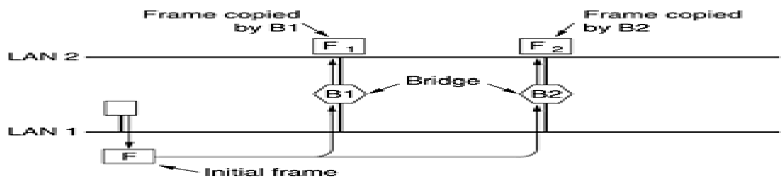
Notes on Adaptive Routing

a) Algorithm shown here depends on only one path between any two nodes
   - If topology includes alternate paths (i.e. not a tree structure), then a minimum spanning tree algorithm (IEEE 802.1) should be used.
     - Requires that bridges have identifiers, a cost is assigned to each bridge port, and that they communicate with each other on startup.

b) Periodically, entries in lookup table should expire, to allow for updates in topology

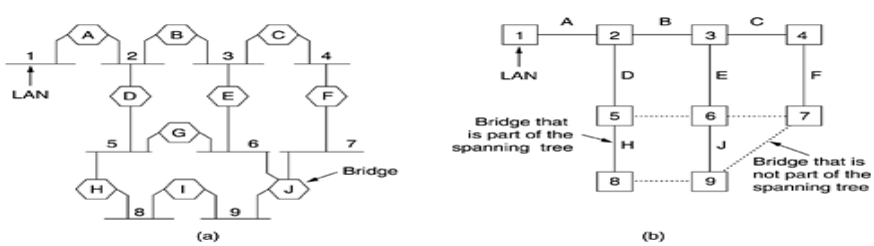
Spanning Trees

For reliability: sites use two or more bridges in parallel between pairs of LANs



•Problem: frame, F, with unknown destination is flooded
Each bridge, uses flooding for unknown destinations, here means copying it to LAN 2.
B1 sees F2, with an unknown destination, which it copies to LAN 1, generating F3 (not shown).
B2 copies F1 to LAN 1 generating F4 (not shown).
B1now forwards F4 and B2 copies F3. This cycle goes on forever.

Spanning Trees

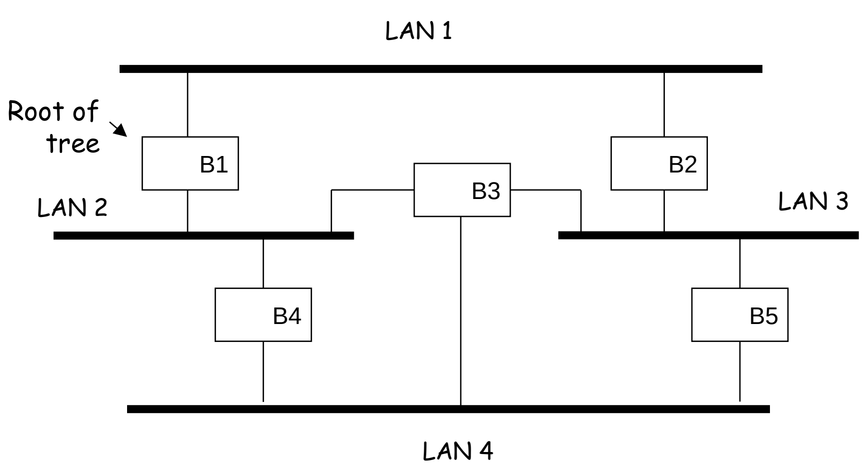


(a) nine LANs interconnected by ten bridges. This configuration can be abstracted into a graph with the LANs as the nodes. An arc connects any two LANs that are connected by a bridge.
(b) graph can be reduced to a spanning tree by dropping the arcs shown as dotted lines
- exactly one path from every LAN to every other LAN

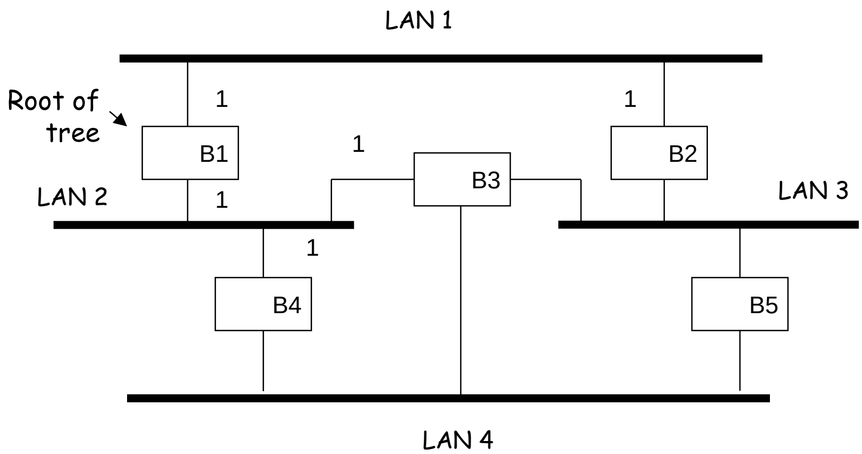
Spanning Trees (3)

- Bridges have to choose one bridge to be the root of the tree.
- Each bridge broadcast its unique serial number. The bridge with the lowest serial number becomes the root.
- Tree of shortest paths from the root to every bridge and LAN is constructed. This tree is the spanning tree.
- If a bridge or LAN fails, a new one is computed.
- Unique path is established from every LAN to the root and thus to every other LAN.
- Tree spans all the LANs, not all the bridges are necessarily present in the tree (to prevent loops).
- Even after the spanning tree has been established, the algorithm continues to run during normal operation in order to automatically detect topology changes and update the tree

Network that is not a tree



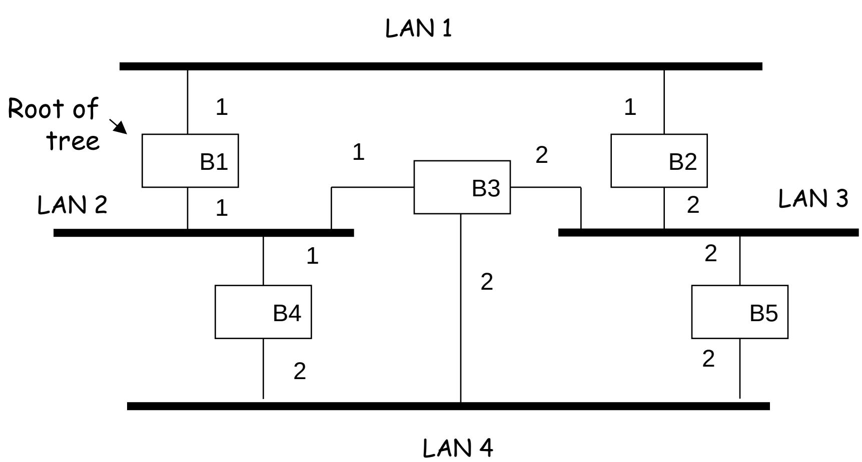
Hop counts



41

Divyashikha Sethia (DTU)

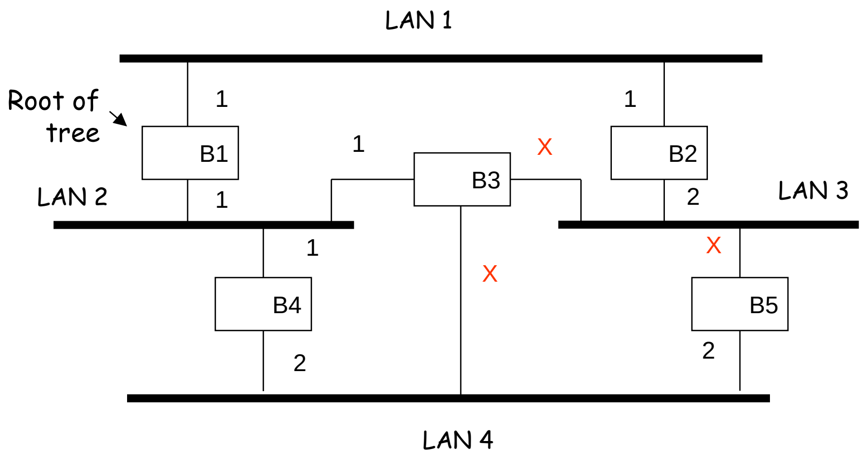
Hop counts



42

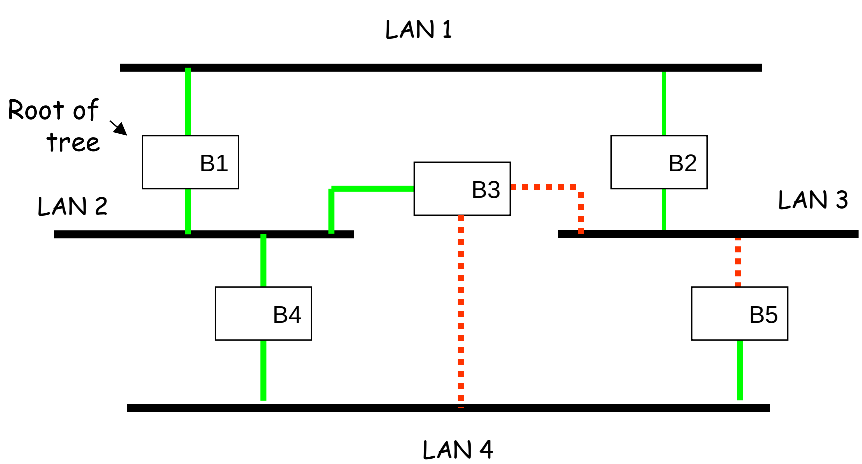
Divyashikha Sethia (DTU)

Hop counts



Discard edges that produce cycles

After minimum spanning tree



Bridge will not forward messages incoming from these ports.

Animations

- Animations: http://www.visualland.net
-
- LAN Flooding
- Unicasting
- Broadcasting

THANKS

Divyashikha Sethia (DTU)

References

Tanenbaum chapter 4