Networking Technologies and Applications

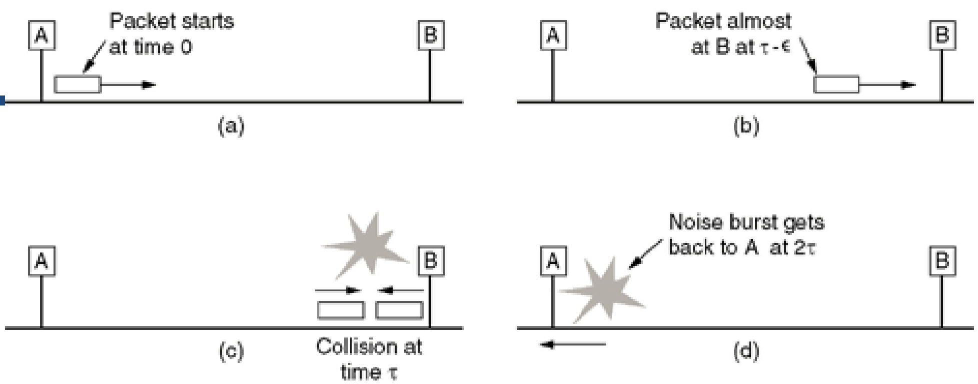
Rolland Vida
BME TMIT

October 3, 2019

Ethernet = CSMA/CD

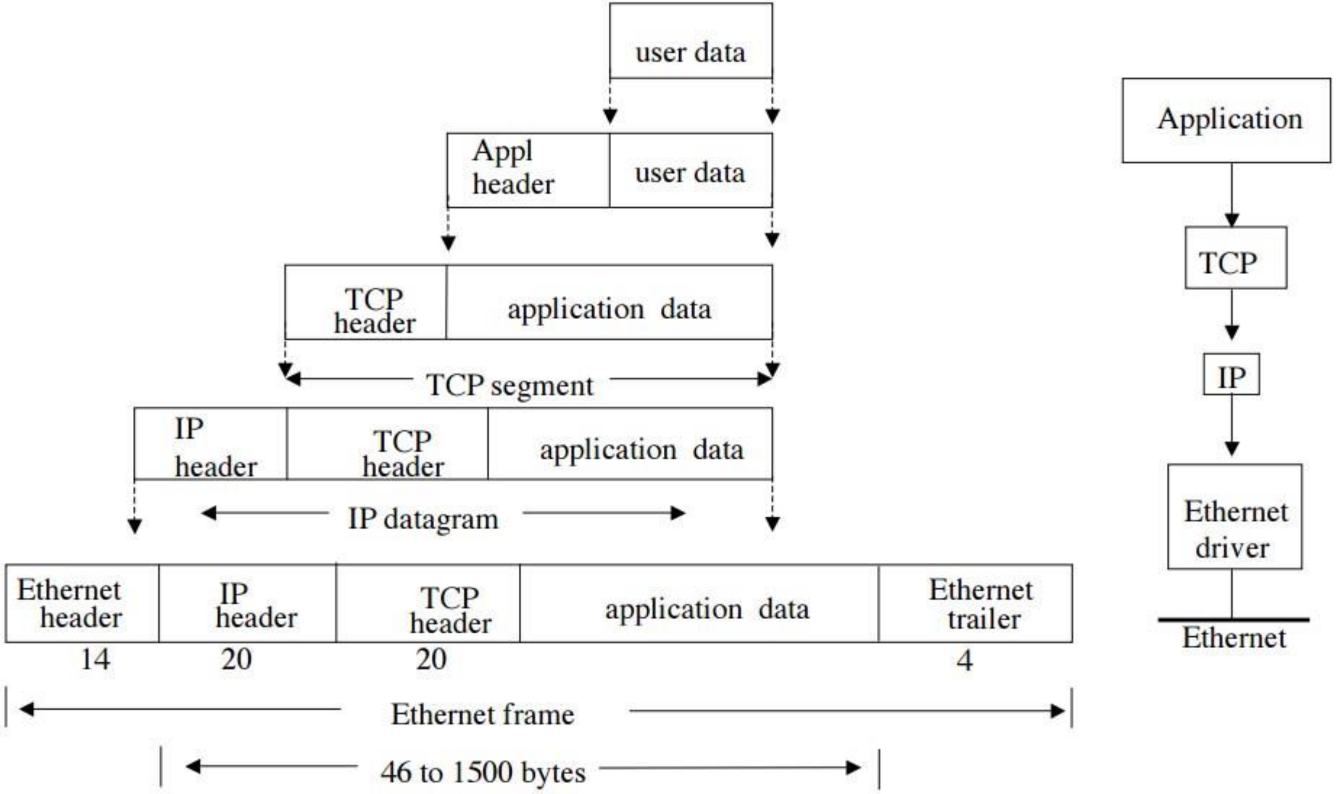
- Before transmission, hosts on the same Ethernet cable first listen to the channel (**CSMA – Carrier Sense Multiple Access**)
  - If busy, they wait for the transmission to end
  - If free, start sending
    - Not immediately, but after a **slot time**
      - If there is a signal on the channel, it leaves time for it to be received
    - Slot time = maximum round-trip time on the cable
      - For 10 Mb/s Ethernet it is 51,2 µs, for 100 Mb/s is 5,12 µs
- Two stations might think in parallel that the channel is free
  - Both start sending, a collision occurs
  - If collision, they detect it (**CD – Collision Detection**), and send a jam signal to ensure that others detect the collision as well

Ethernet = CSMA/CD

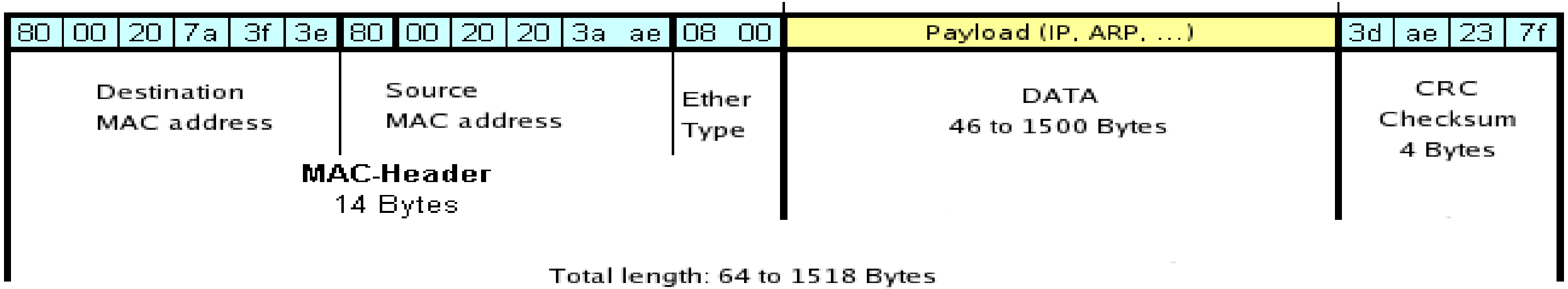


- Wait for a random time interval and retry afterwards
  - Set a timer to a random value from the $[0, 1, .., 2^{m-1}]$ x t_{av} interval, where t_{av} is the default waiting time (51,2 µs), m = min (10,n) and n is the number of collisions.
  - After each collision the maximum waiting time is doubled, until reaching an upper limit - **truncated binary exponential backoff**
- CSMA/CD not possible on Alohanet
  - Two users on the remote islands could not hear each other

Encapsulation



Ethernet Frame



- The DATA field is at most **1500 bytes** (**MTU – Maximum Transmission Unit**)
  - – If frame too large, it occupies the channel for long time
  - – Higher possibility of an error, you will need to resend large frames
- Minimum length **46 bytes** (minimal frame size = 64 bytes = 512 bits)
  - If the frame too small, collision detection cannot be used
    - – Transmission is terminated very fast, before the first bit reaching the end of the cable
    - – 51,2 µs round trip time / 0,1 µs bit time = 512 bit
  - Even if there's a collision, the sender is not informed about it
  - – Packets that are too small are filled with **padding data** (bits with no utility)

Carrier Extension

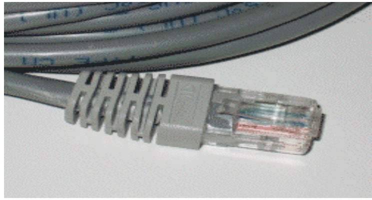
- If the speed of the network increases...
  - Either increase the minimum frame size...
  - Or decrease the maximum cable length
    - On a 2500 m cable, for a 1 Gb/s speed, the minimum frame size is 6400 bytes
    - If the minimum size is 640 bytes, the cable can be only 250 m long
  - Very annoying restrictions on a Gigabit network

- **Carrier Extension**
  - The sender puts the useless bits after the CRC field
  - The receiver cuts it of, not included in the CRC
    - Still a serious waste of capacity
- **Frame Bursting**
  - During a single transmission, several consecutive frames transmitted
  - Increases efficiency considerably

First Ethernet versions

- 10Base5 – thick Ethernet
  - Coaxial cable, 10 Mb/s, 500 m long segments

- 10Base2 – thin Ethernet
  - Coaxial cable, 200 m segments



- 10Base-T
  - Twisted pair, star topology around a hub, 100 m segments

- 10Base-F
  - Optical fiber, 2km long segments

Fast Ethernet

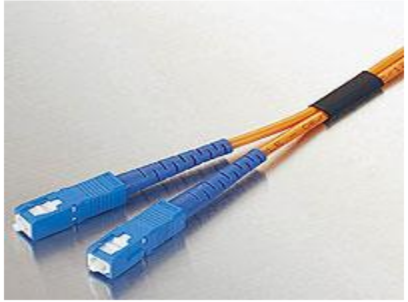
- 100Base-TX
  - 2 twisted pairs
  - One for the upstream, one for the downstream, 100 Mb/s duplex speed
- 100Base-T4
  - 4 twisted pairs
  - One for the upstream, one for the downstream data, the other two can be used as needed
  - Maximum 100 m long segments
- 100Base-FX
  - Multi-mode fiber in both directions
  - 100 Mb/s duplex speed
  - Maximum 2 km between the hub and the stations

Gigabit Ethernet

- IEEE 802.3z (1998), 802.3ab (1999)
- Only point-to-point setups
  - No shared segments, as in traditional 10 Mb/s Ethernet
- Two operation modes:
  - Duplex – traffic in both directions in the same time
    - A central **switch** links the stations on the periphery
    - All the connections are buffered
      - Any station can send data at any time
      - No need to sense the channel, no contention
      - No need for CSMA/CD, not really Ethernet anymore
  - Half-duplex
    - Stations are connected to a simple hub
    - No buffering, collisions are possible

Gigabit Ethernet

- Different versions
  - 1000Base-SX
    - Multi-mode fiber
    - Maximum 550 m long segments
  - 1000Base-LX
    - Single- or multi-mode fiber
    - Maximum 5000 m long segments
  - 1000Base-T
    - 4 pairs of Cat. 5 UTP cables
    - Maximum 100 m long segments
- IEEE 802.3ae – 10 Gb/s Ethernet (2002)
  - Only on optical cables
- IEEE 802.3ba – 40Gb/s and 100 Gb/s Ethernet (2010)
  - Lucent Technologies Bell Labs experimental results
  - Standard adopted in June 2010



Hub

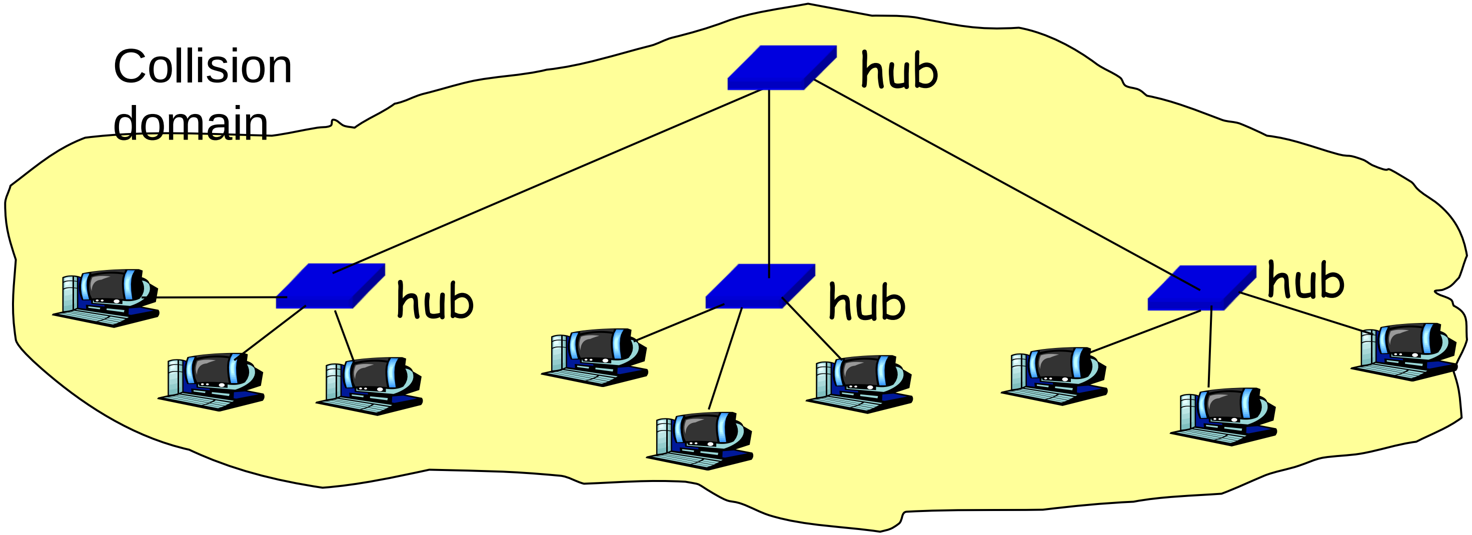
- Physical layer repeater device
  - Repeats the packet on bit level
  - An incoming packet is forwarded immediately on all the other interfaces
    - Everyone receives all the packets
- If many simultaneous transmissions - collision
  - The „collision domain” is not changed
- Usually a hierarchical, tree-like hub topology

Hub – advantages and drawbacks

- Each station can collide with any other station on the hub
  - Lowers the efficiency of the network
  - Lowers scalability
  - Anyone can see anyone's traffic
- Different Ethernet versions cannot be joined in the same network
  - If one 10Mbps station in the network, the entire network switches back to 10 Mbps operation mode

Hub

- Not efficient to build a large network using only hubs
  - One large collision domain

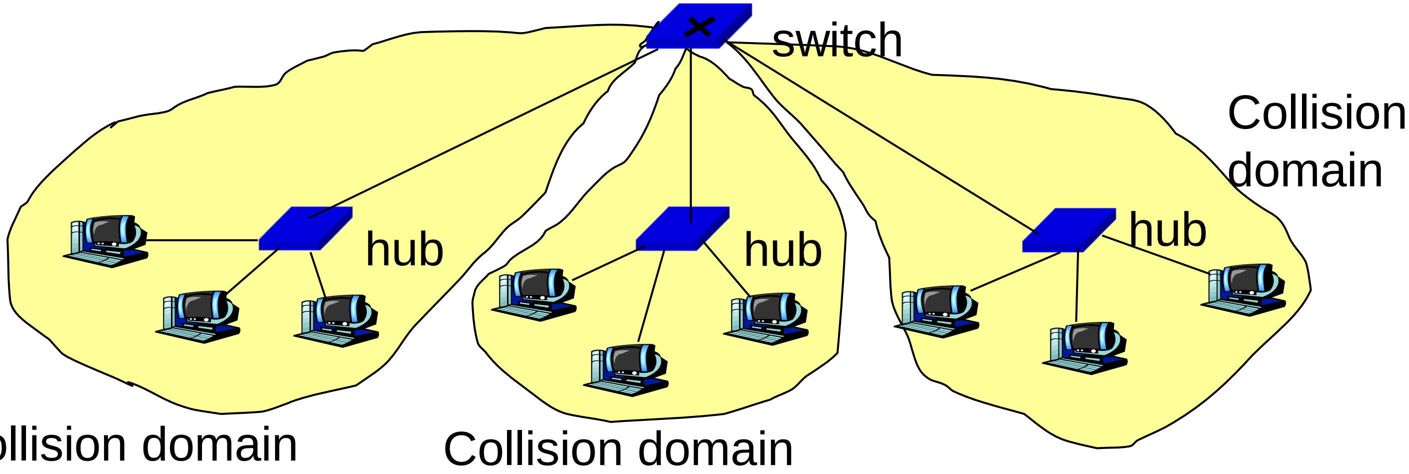


Switch (bridge)

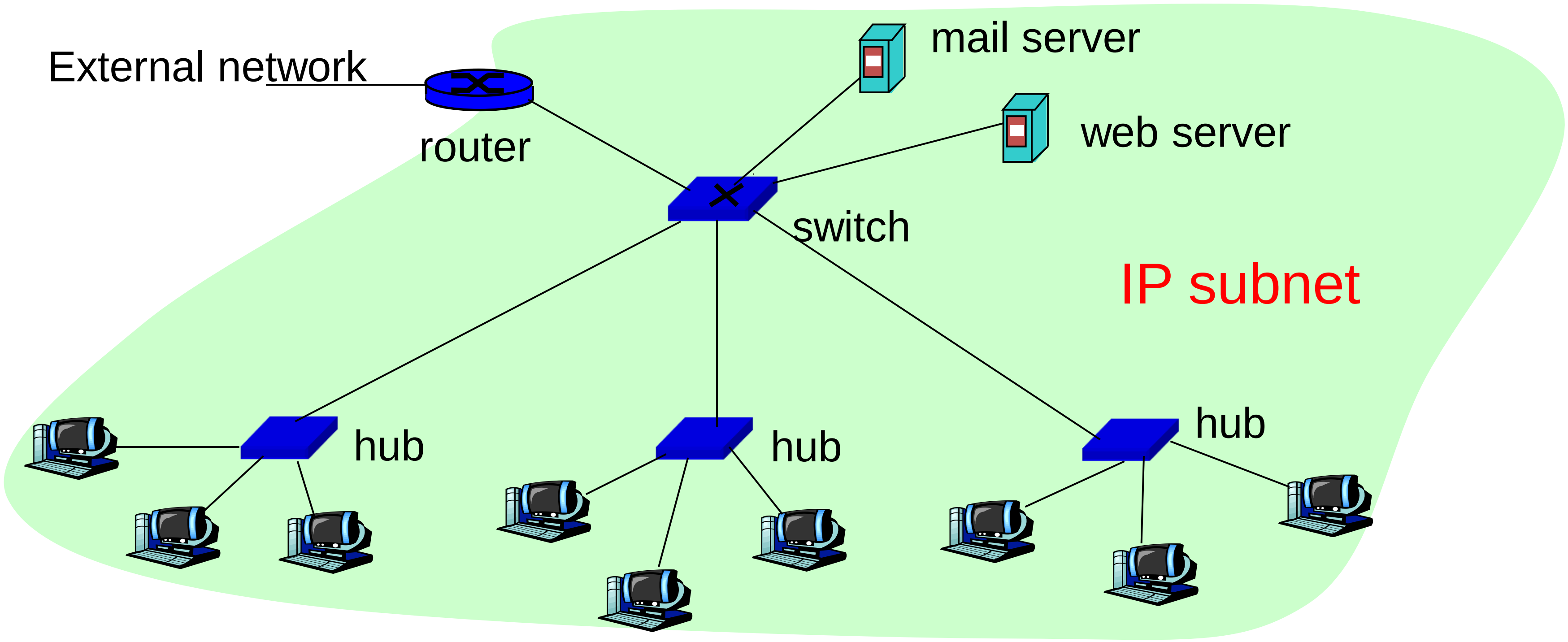
- Link Layer device
  - Checks the MAC header, and forwards selectively
    - switch table: (MAC address, interface, timer)
      - Built from the received packets
      - If one address is unknown, the packet is forwarded to all the interfaces
  - Separates the collision domains
    - Buffers the packets
    - Forwards them only to the appropriate segments

Switch

- Advantages:
  - Higher scalability
  - More efficient, more secure
  - Buffering and switching tables makes the connection of different Ethernet versions possible inside the same network

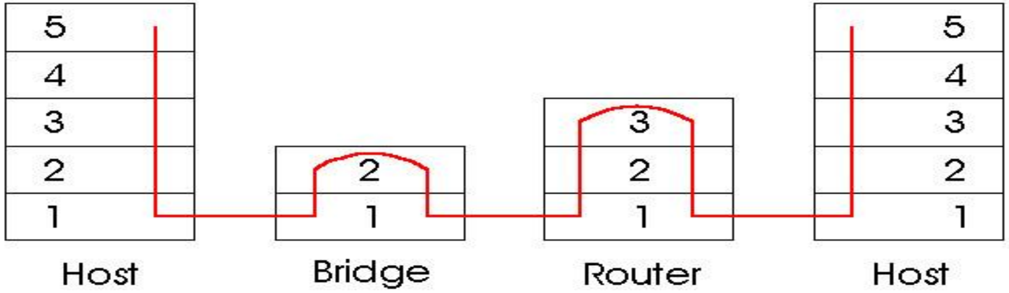


Corporate network



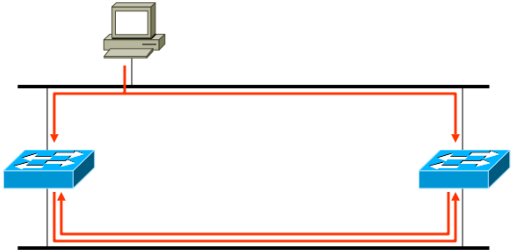
Switch (bridge) vs. router

- Intelligent store-and-forward devices
- Router
  - In the network layer (L3), based on IP addresses
  - Stores routing tables, uses routing protocols
- Switch
  - In the data link layer (L2), based on MAC addresses
  - Stores switching tables, uses address learning algorithms

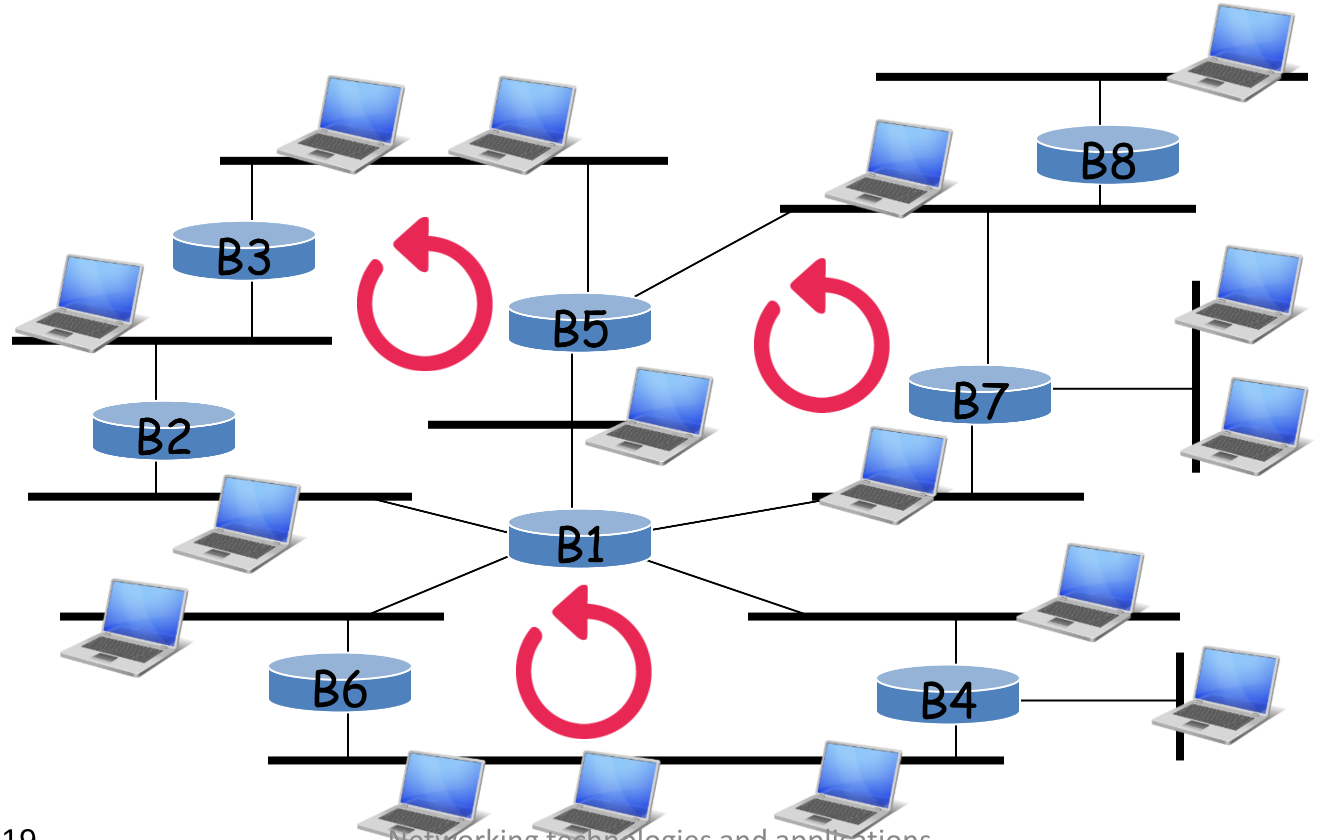


STP

- Spanning Tree Protocol
  - Part of the IEEE 802.1D standard
    - Radia Perlman (MIT, DEC)
  - Loop-free trees on a bridged LAN
    - No TTL in Ethernet (Time To Live)
      - In case of a loop, packets travel indefinitely in the network
    - Need for redundancy
      - In case of an error, there should be an alternative path



Example topology

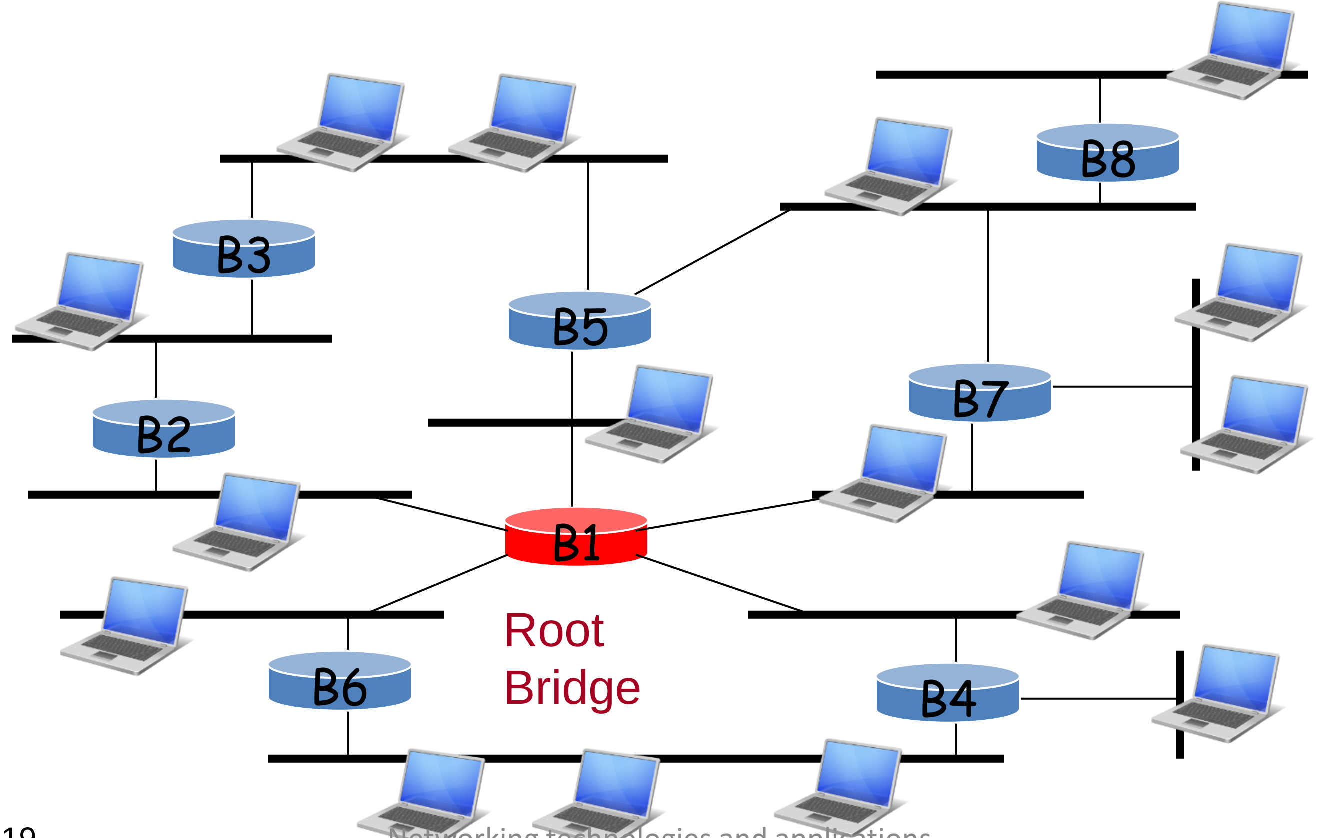


STP operation

- Choosing the root bridge
  - Each bridge has a MAC address and a configurable priority number
    - BID – Bridge Identification (64 bits)

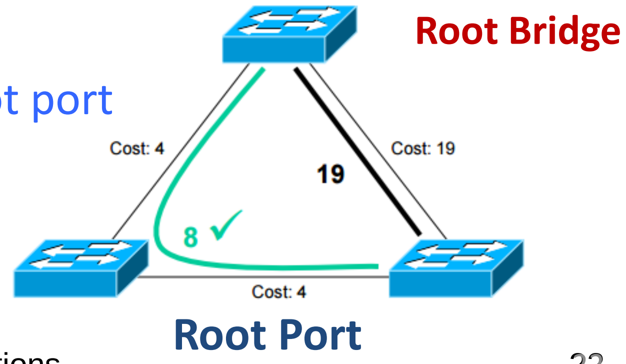
  - The bridge with the lowest priority will be the root
    - In case of equal priorities, the lowest MAC address wins
    - There will be a secondary (backup) root as well
  - Totally automatic, but if the network manager wants a specific device to be the root, it sets a low priority number

Choosing the root bridge

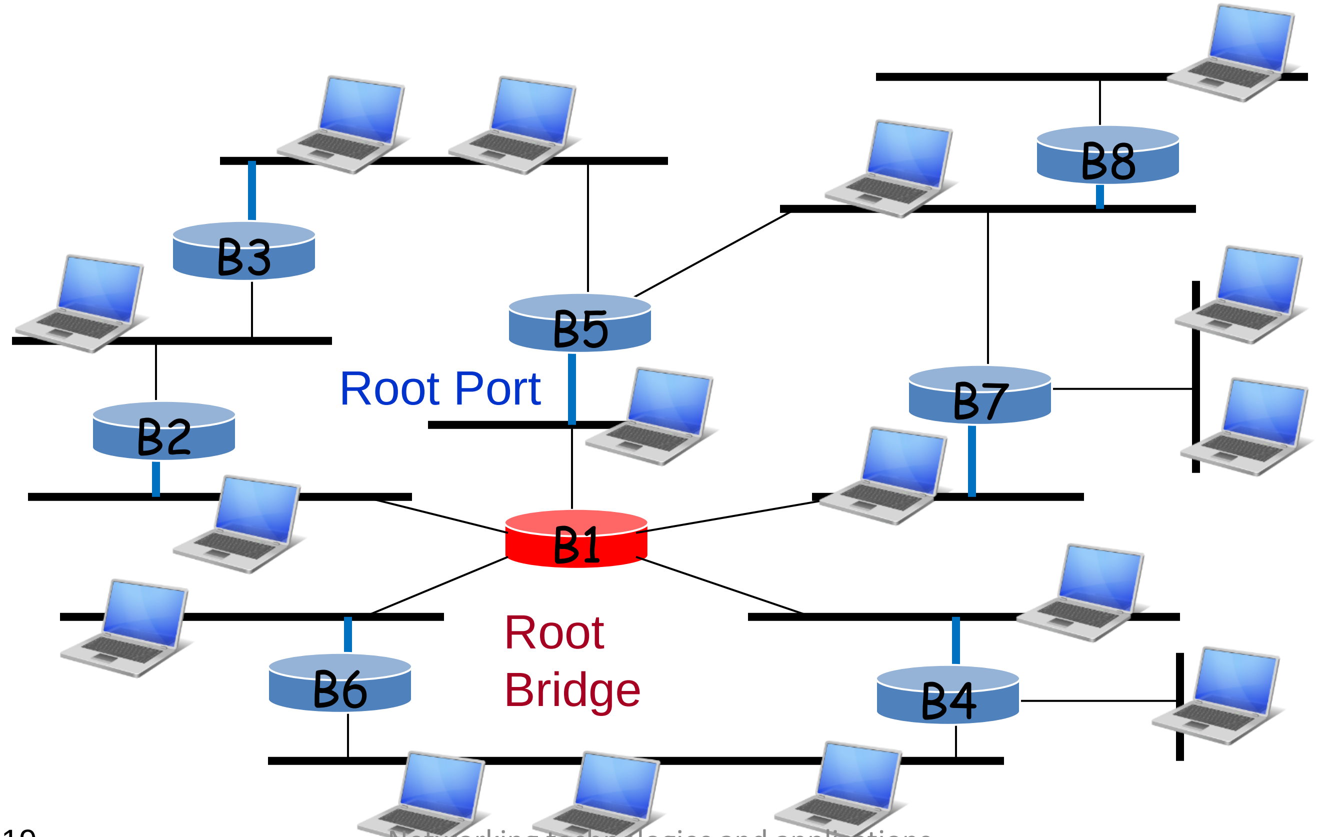


STP operation

- Finding the „cheapest" path to the root bridge
  - BPDU – Bridge Protocol Data Units
    - Sent periodically (2s) among the bridges
  - A bridge calculates the cost of all the possible paths to the root bridge
    - Each port has a *Port Cost*
      - Administrative value, e.g., inversely proportional with the bandwidth
  - Chooses the least-cost path
    - The port belonging to that path will be the root port
    - If several paths with the same cost, the lower Port ID wins



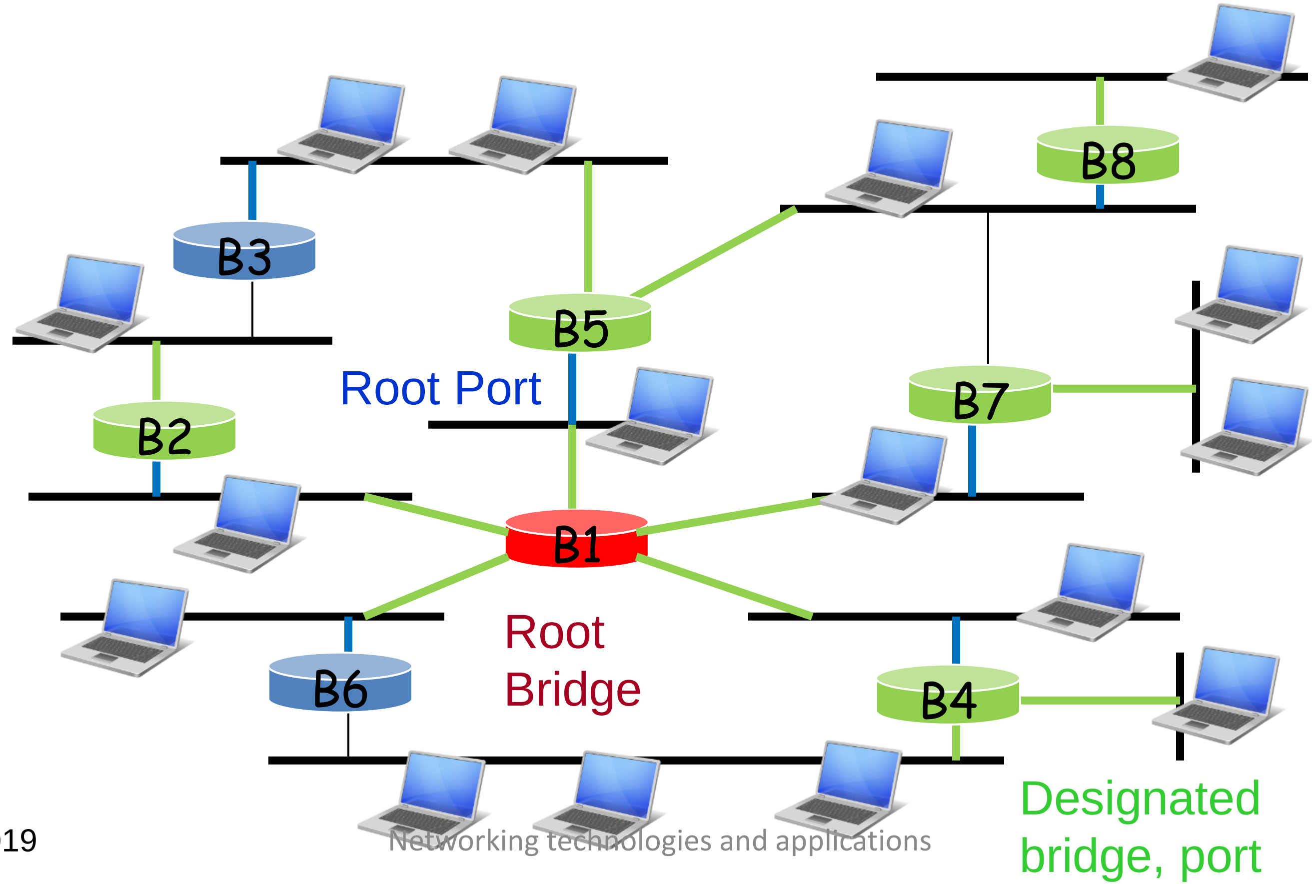
Choosing the root port



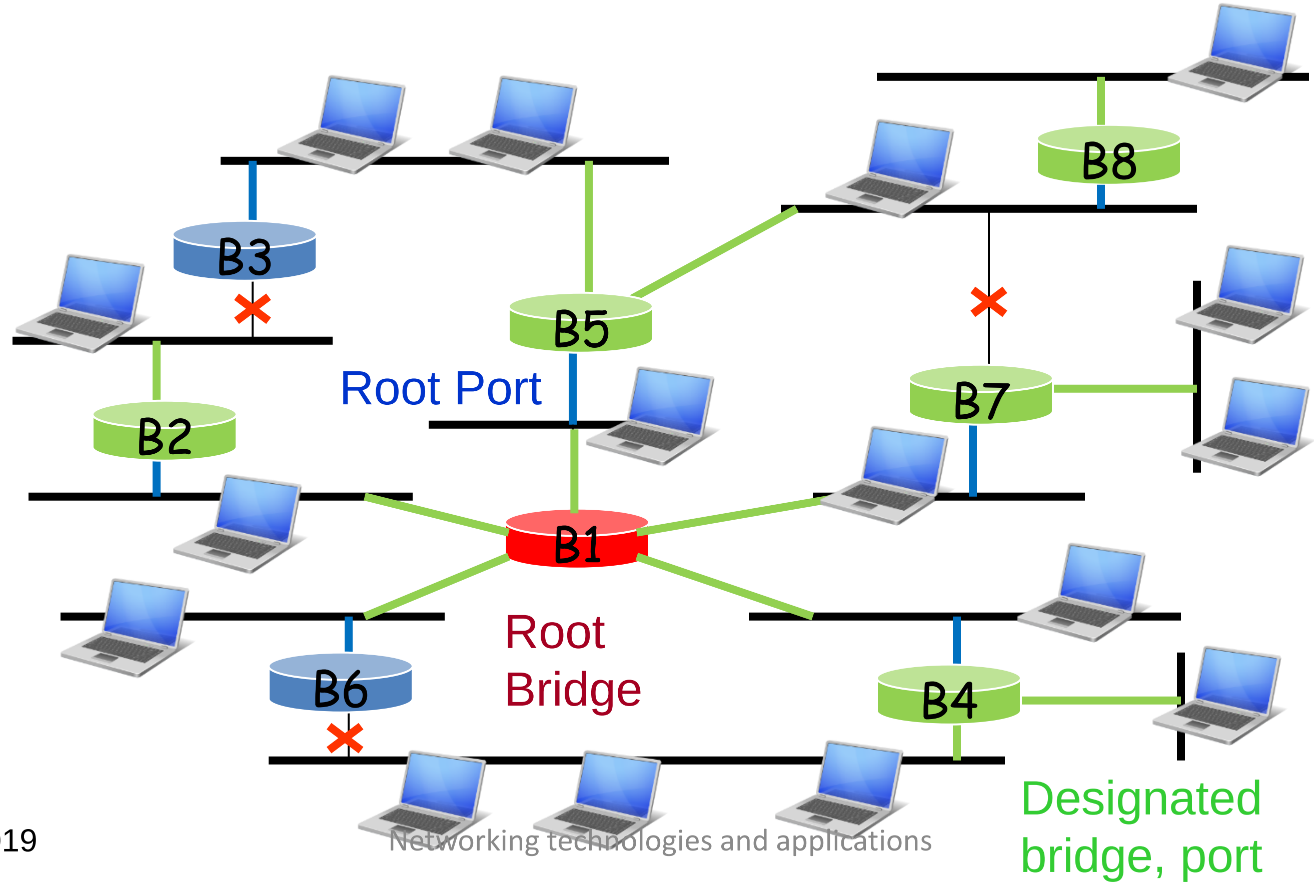
STP operation

- Finding the „cheapest” paths to the root bridge for each LAN segment
  - The bridges calculate together, for each LAN segment, which is the bridge that belongs to the least-cost path towards the root bridge
    - *Designated bridge, designated port*
  - The designated and root ports are switched to *forwarding state*
  - On all the other ports traffic is blocked
    - Only BPDUs pass
- After building the tree, addresses are learned
  - 15 seconds learning time

Choosing the Designated bridge/port



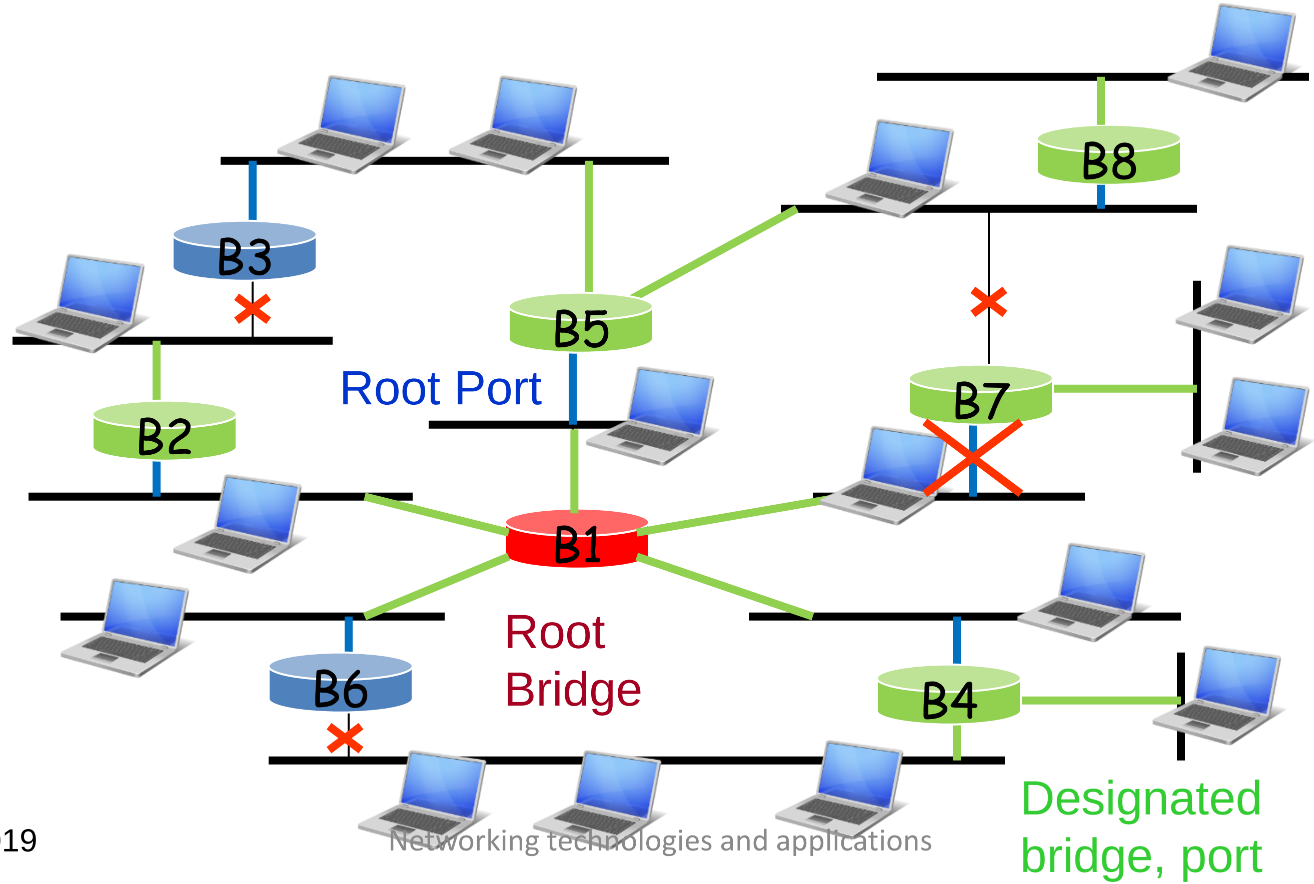
Port blocking



Handling errors

- BPDUs sent periodically
- Two BPDUs missed means an error
  - The bridges recalculate the topology
  - If there is a blocked port, they will use it
- New topology built in 15 sec
- Then, MAC addresses are learned again
  - In 30 secs the network is operational again

Handling errors



Handling errors