Ethernet

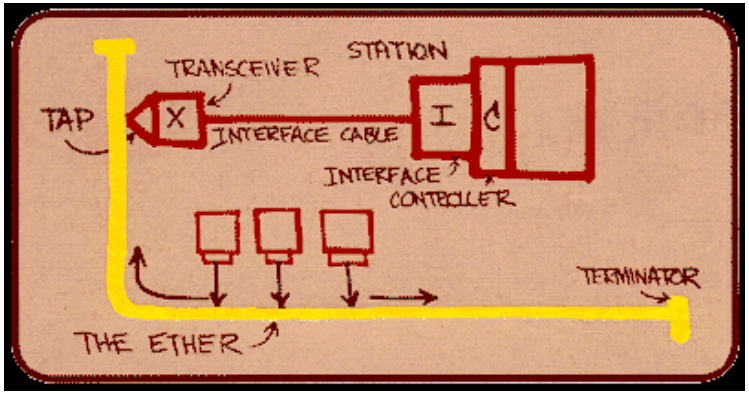
Fulvio Risso

Politecnico di Torino

Copyright notice

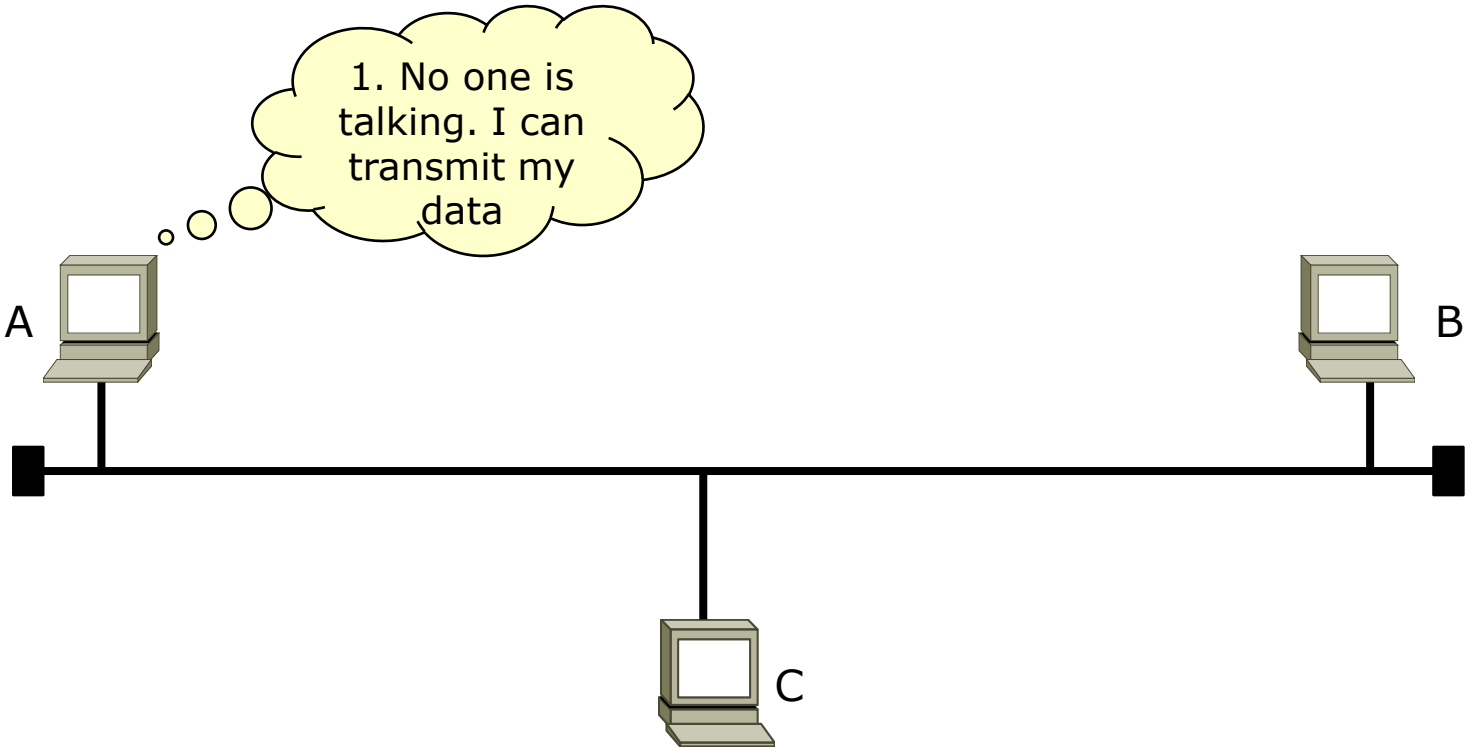
Ethernet in brief

- Bus (physical at the beginning, logical later)
- Carrier Sense, Multiple Access, Collision Detection (CSMA/CD)
- First prototype: 1973, 2.94Mbps
  - Xerox PARC
- First commercial specs: 1980, 10Mbps
  - Digital, Intel, Xerox (DIX), major participants in Ethernet design
  - Also known as Ethernet 2.0
- First successful LAN technology
  - The most common competitor (at the beginning) was Token Ring
- Very simple



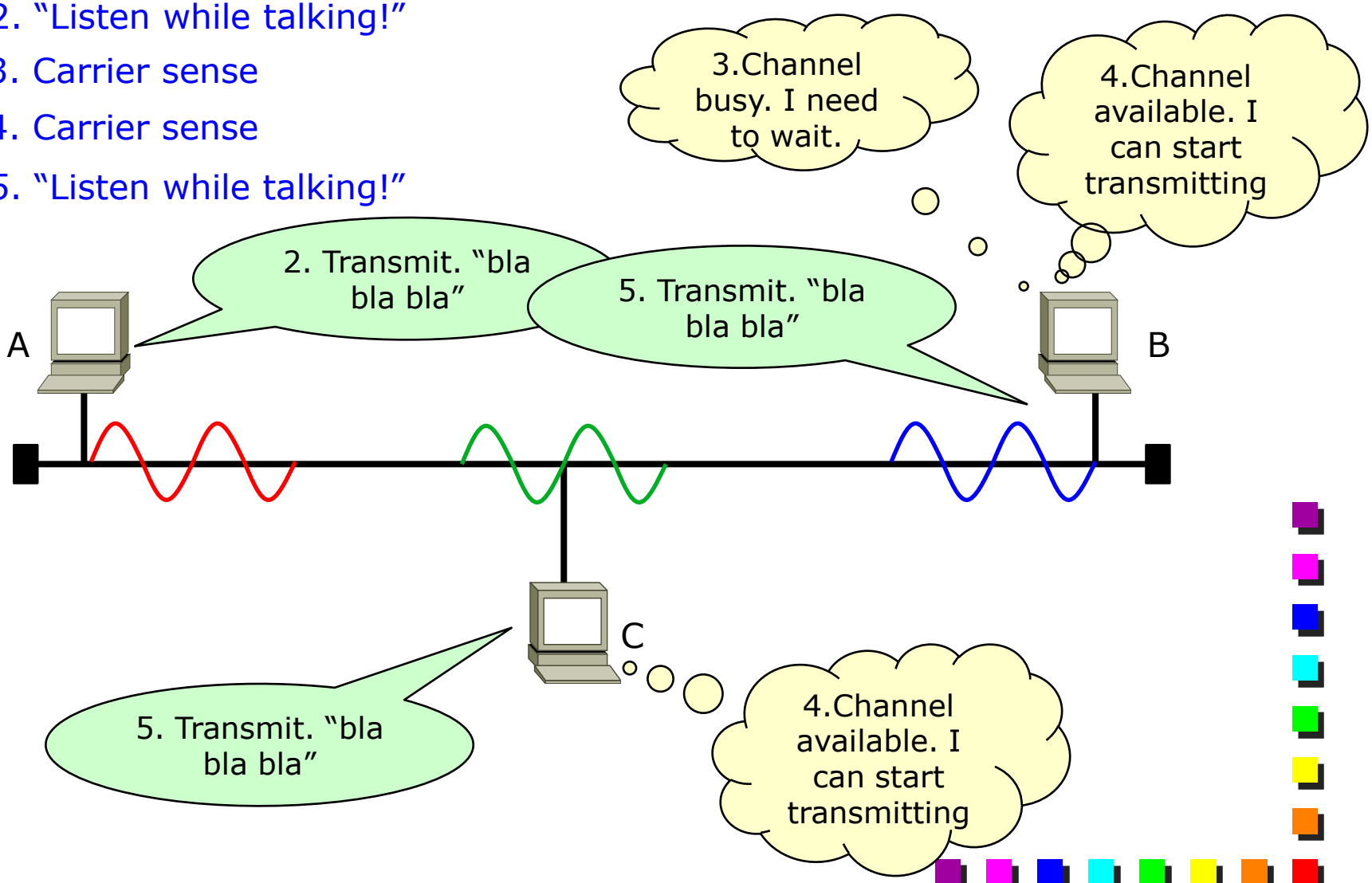
Ethernet principles (1)

1. Carrier sense



Ethernet principles (2)

2. “Listen while talking!”
3. Carrier sense
4. Carrier sense
5. “Listen while talking!”

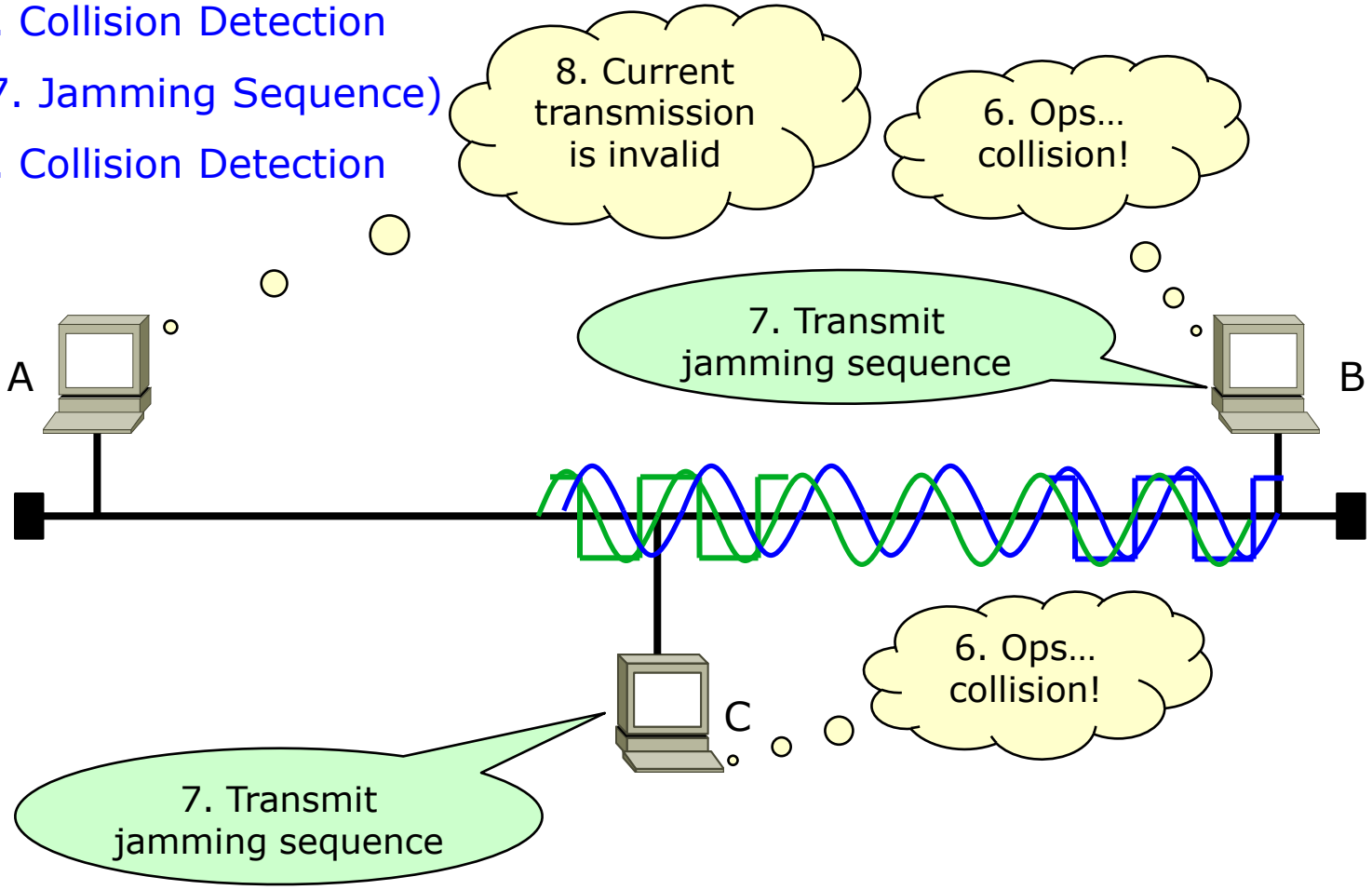


Ethernet principles (3)

6. Collision Detection

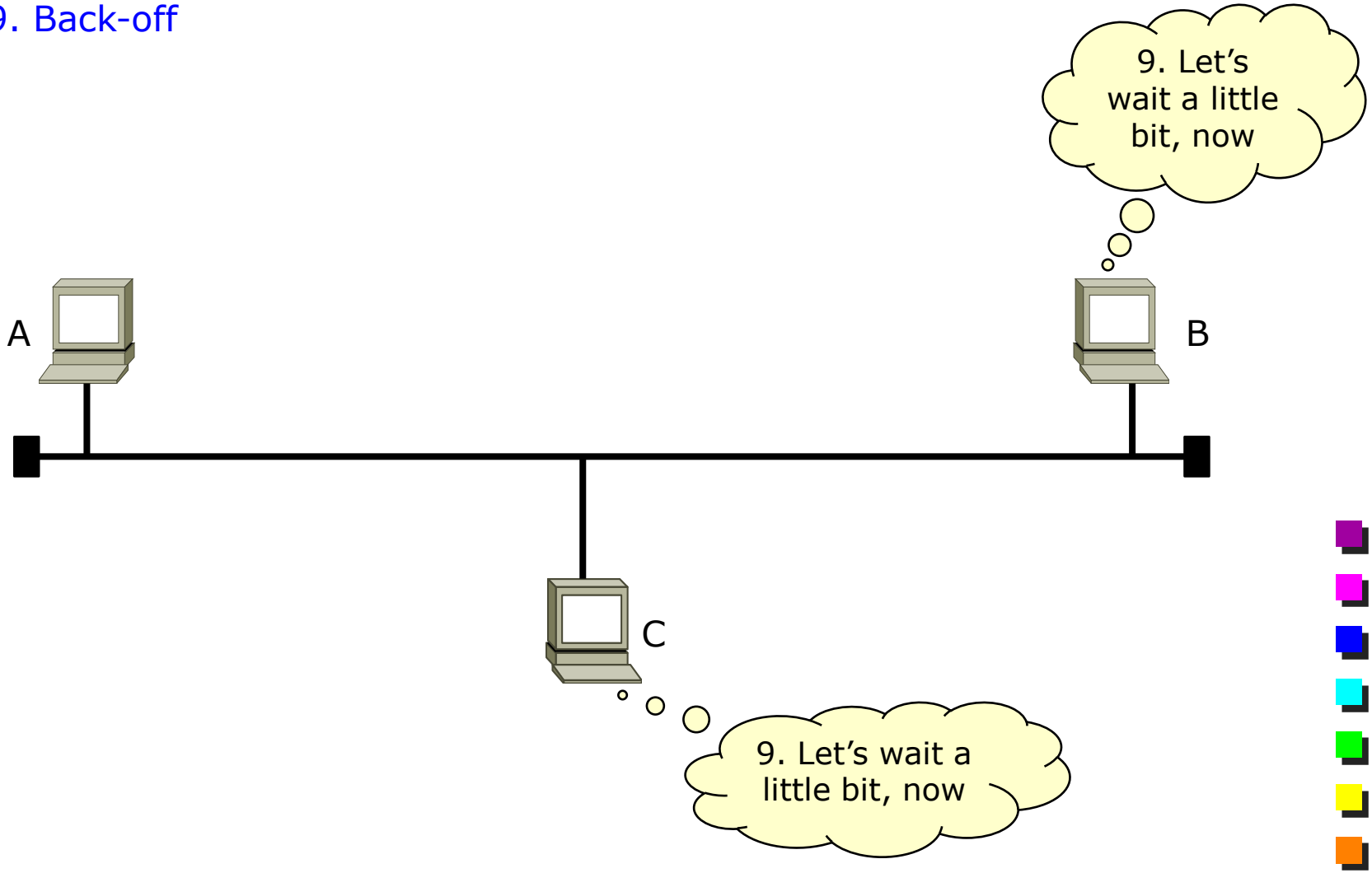
(7. Jamming Sequence)

8. Collision Detection



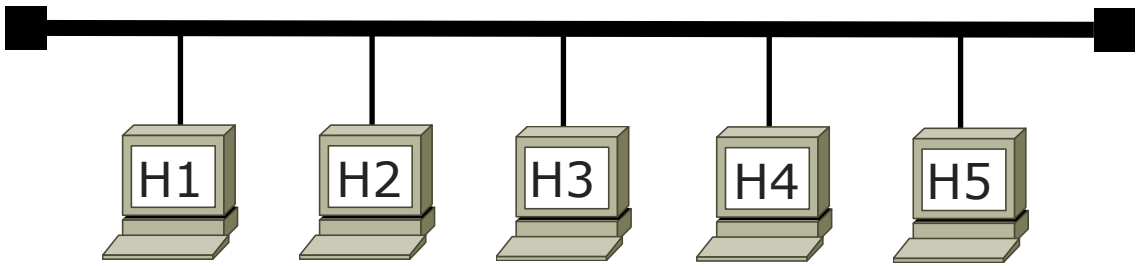
Ethernet principles (4)

9. Back-off



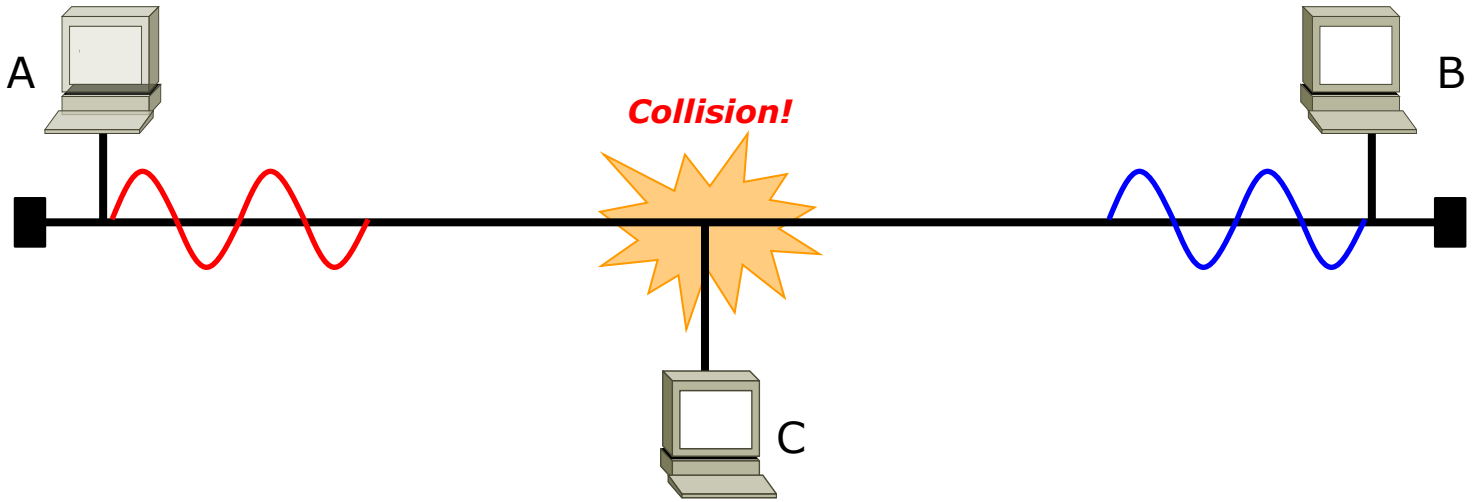
Ethernet principles (5)

- CSMA/CD
  - Carrier Sense: “Listen before talking”
  - Multiple Access: everyone can talk (potentially at the same time)
  - with Collision Detection: “Listen while talking”
- MAC Protocol
  - Defined for bus-based topologies
  - Non-deterministic
  - No upper limit for the waiting time



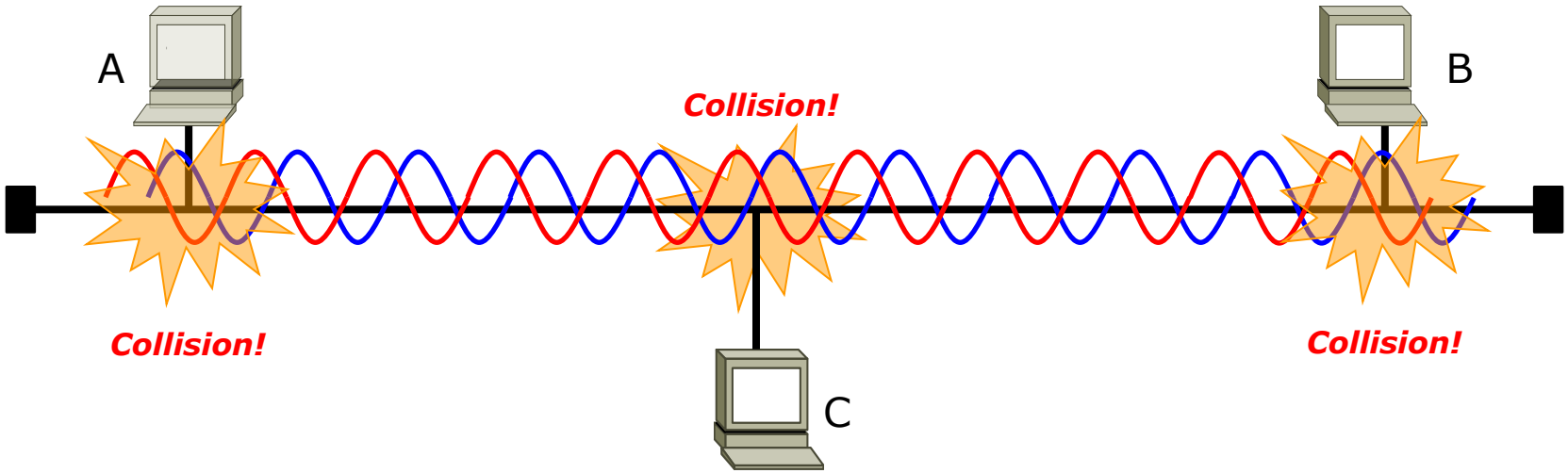
Collision Detection (1)

- Stations can talk at the same time
- In theory, the "Listen while talking" mechanism is able to detect the collision
- But… propagation speed is not infinite, therefore the collision may happen undetected
  - C detects the collision, while A and B do not

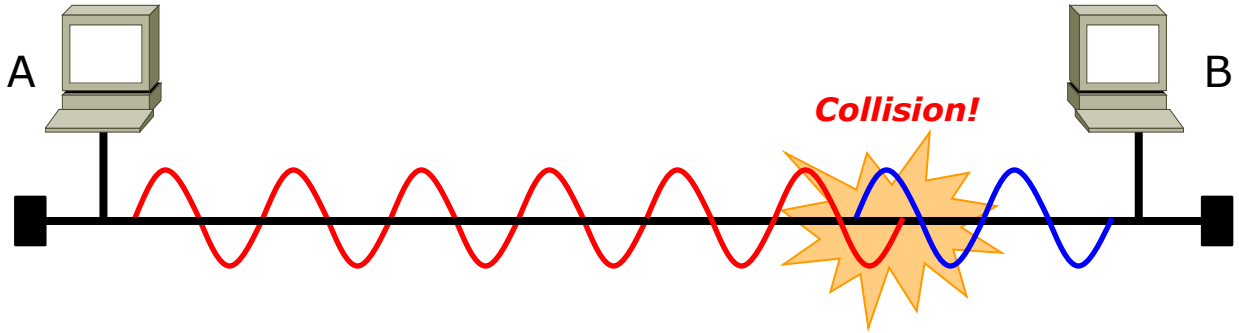


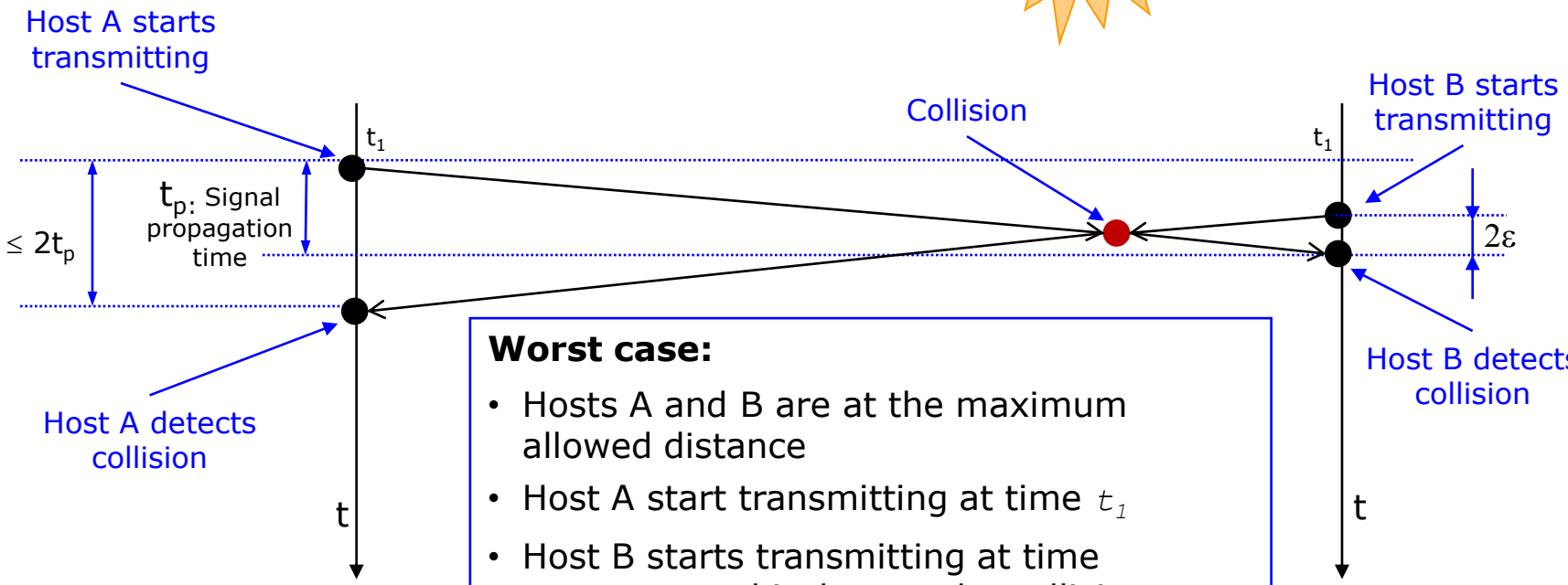
Collision Detection (2)

- In order to detect the collision
  - All talkers must still be active when the collision happens
    - This requires to define a “minimum talk duration”
  - The following entities are tied together
    - Max distance between stations
    - Signal propagation speed
    - Min duration of the talk



Collision Detection (3)





Worst case:

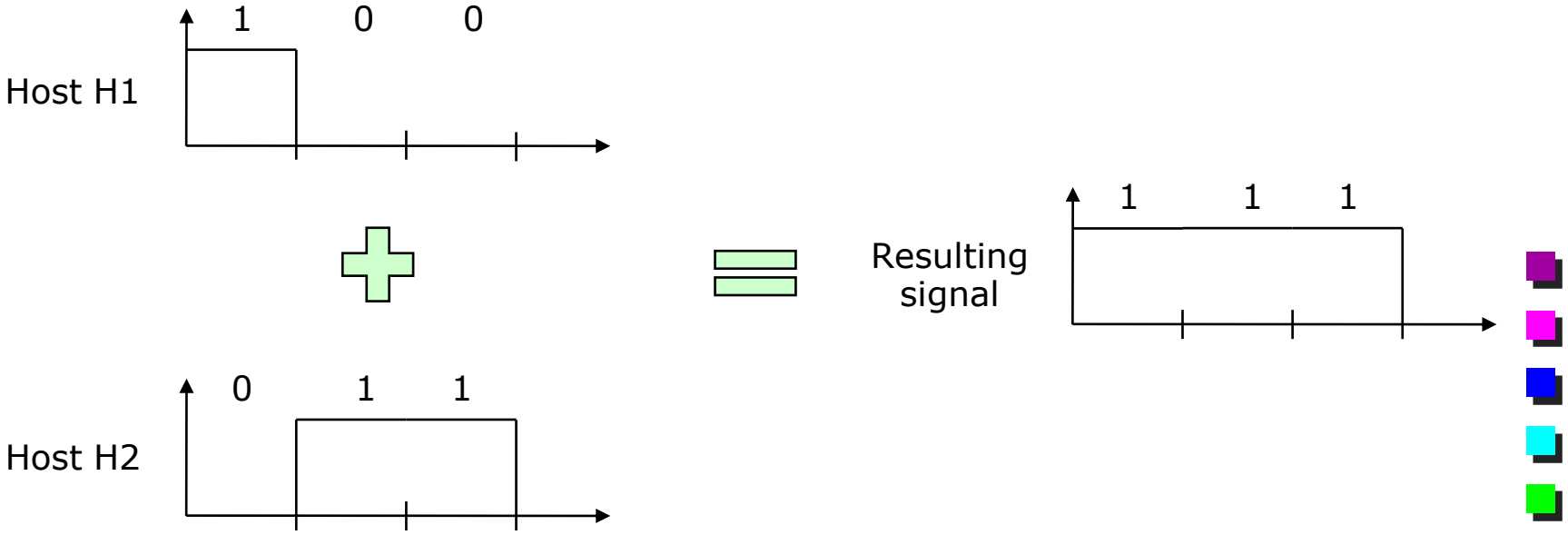
- Hosts A and B are at the maximum allowed distance
- Host A start transmitting at time t_1
- Host B starts transmitting at time $t_1 + t_p - \varepsilon$ and it detects the collision at time $t_1 + t_p$
- Host A detects the collision after $2\ t_p - 2\ \varepsilon$

Collision Detection in the real world (1)

- In the real world, collision is detected by
  - Coax cable: measurement of the average DC on link
  - Twisted pair, fiber
    - Transmitting stations: activity on both links (tx and rx)
      - Hubs propagate the signal to all ports except the one it comes from
    - Other stations: wrong CRC on the received frame
- A note
  - On coax networks, all the stations can report the correct number of collisions
  - On bidirectional networks, each station knows only its number of collisions
    - Other collisions appear as CRC errors

Collision Detection in the real world (2)

- Note: receiving stations may not always be able to detect the transmission by inspecting the signal
- This is the reason we need to rely on CRC



Jamming Sequence

- How can we guarantee that the CRC is invalid?
  - We transmit a particular Jamming Sequence that should maximize the probability to have a CRC error
- Jamming Sequence:
  - A 32 bit sequence, repeated until we reach the *min duration* of the talk
  - This guarantees a *transmission length* = *min frame length*, and no more
    - More efficient, since we no not take the channel for more that min_frame_length, irrespective of the actual size of the original frame
- Example
  - A station detect a collision after 100 bits transmitted
  - The jamming sequence is repeated (512-100)/32 times

A note about collision

- Collision is not a error, it is a feature!
- Collision is not a waste of bandwidth
  - Collision is the way Ethernet arbitrates the channel
    - Other technologies use tokens, special frames, etc
  - Instead, it is very efficient compared to other methods
    - If we do not have collision, the bandwidth used to arbitrate the channel is equal to zero (zero overhead)
    - On a token-based network, each transmission incurs in the overhead of the token
- Hence, although collision generates errors on frames, it is not an error itself

Back-Off (1)

- LANs are “reliable” communication medium, so we have to re-transmit the collided frames
- After a collision each station waits for a random time interval before re-trying the retransmission
  - We do not want both to retransmit at the same time (new collision)
  - In any case, after the random idle time, the Ethernet algorithm restarts from scratch
    - Wait that the channel is available and, if so, retransmit the frame
- The inactivity time is specified by a Truncated Binary Exponential Back-off algorithm

Back-Off (2)

- Parameters of the Truncated Binary Exponential Back-off algorithm:
  - τ = time required to transmit a 512 bits slot
  - n = number of collisions occurred on the current frame
- Algorithm
  - We choose a time value **r** randomly, according to the following inequality:

$$0 \leq r < 2^K \qquad k = \min(n, 10)$$

  - Between two consecutive transmissions, we have to wait at least:

$$T = r * \tau$$

  - Max 16 re-transmissions on the same frame
    - The transmission is then aborted (but please check the sanity of your network in that case)

Performances of the CSMA/CD

- Possible problems when load reaches 100%
  - Reasonable at 50% load
- Simple and distributed protocol
  - No intermediate devices, synchronization, etc
- No upper bound on delay
  - In theory not suitable for real-time communications
  - In practice…

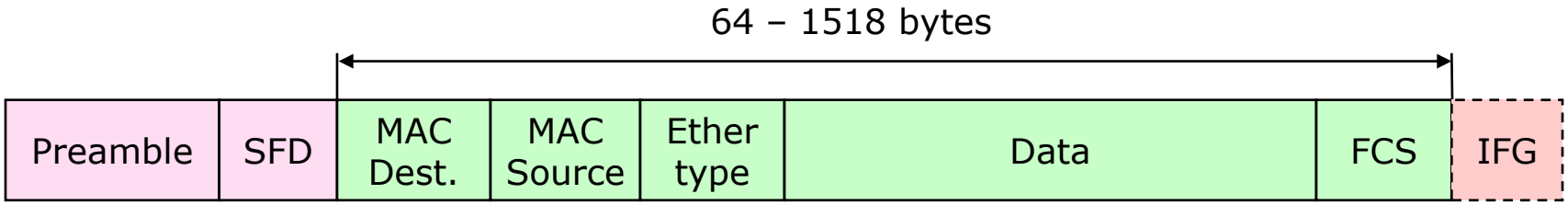
Ethernet in brief

- 10 Mbps maximum speed
- Multiple physical layers (coax, twister pair, fiber)
  - Multiple coding at physical layer
- Minimum frame: 64 bytes
  - In order to guarantee an adequate collision diameter
- Maximum frame: 1518 bytes
  - In order to guarantee an adequate statistical multiplexing over shared channels

Ethernet (DIX) frame format

- Frame format for the classical Ethernet (DIX) is very simple
  - MAC src/dst (addressing)
  - Ethertype (protocol demultiplexing)

- Actual channel occupancy include also Preamble, Start Frame Delimiter and Inter Frame Gap
  - 84 bytes (min frame)
  - 1538 bytes (max frame)



Ethernet DIX

Ethernet (DIX) frame format: delimiters

- Preamble
  - Up to 7 bytes for sync source/receiver
- Start of Frame Delimiter
  - Special byte for signaling the start of frame (invalid code at L1)
- Be careful: No end-of-frame delimiter
- Inter-frame gap: min silence between a frame and the following
  - 96 bit times
  - → 9.6us
  - Called “Inter-Frame Spacing” in IEEE 802.3
    - Not needed (“length” field), but retained for compatibility

Slot time: definition

- Time required to send a minimum Ethernet frame
- Expressed in “bit times”
  - I.e., number of bits
- Equal to 512 bit times on Ethernet
  - I.e., 51.2us
  - It does not include Preamble and Start Frame Delimiter

Collision Detection on Ethernet DIX (1)

- In case repeaters are present
  - Repeaters can shorten the Preamble
  - We should not consider the entire preamble as part of the minimum frame (for collision detection)
    - To be really sure, we should not consider the preamble at all
    - F_{min} may be considered 64 bytes + SFD – 1 bit (i.e., *519 bits*)
- However, the IEEE 802.3 considers the preamble, so we do
  - Please take repeaters into consideration if you work near to the limit
- IEEE 802.3 and repeaters
  - No more than 4 cascading repeaters are allowed in an Ethernet network
    - The preamble could become too short

Collision Detection on Ethernet DIX (2)

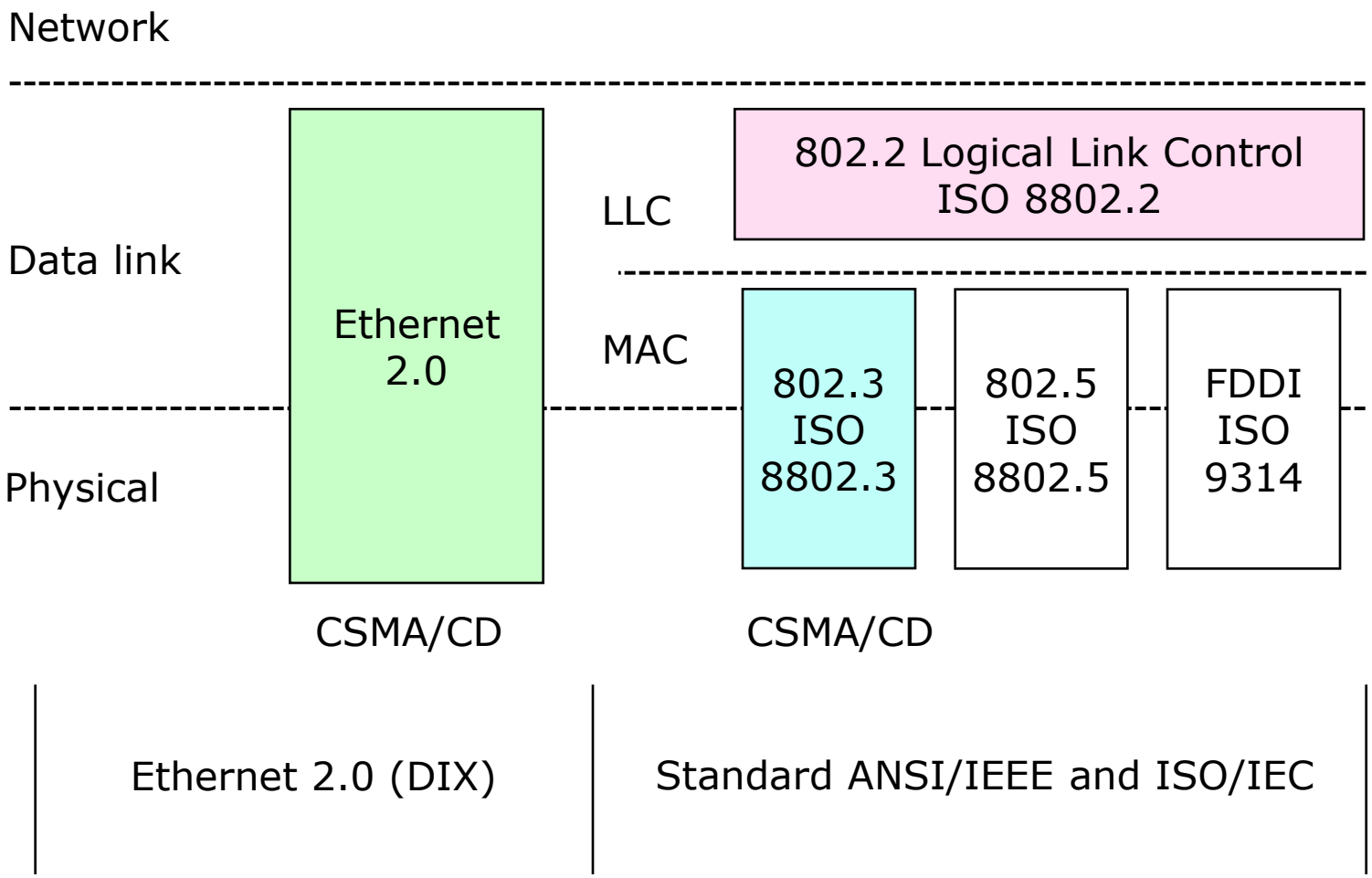
- Input data
  - Maximum distance: D_{max}
  - Signal propagation speed: S_{signal}
  - Min transmission size: F_{min}
  - Network bandwidth: B
- Some math
  - Propagation Time $t_p = D_{max} / S_{signal}$
  - Collision Window = Min Frame duration = $F_{min} / B = 2 * t_p$

$$D_{\max} = \frac{(F_{\min} - 1bit) \cdot S_{signal}}{2 \cdot B}$$

  - With B=10 Mbps, F_{min}=(64+7+1) bytes, S_{signal}=200.000 Km/s

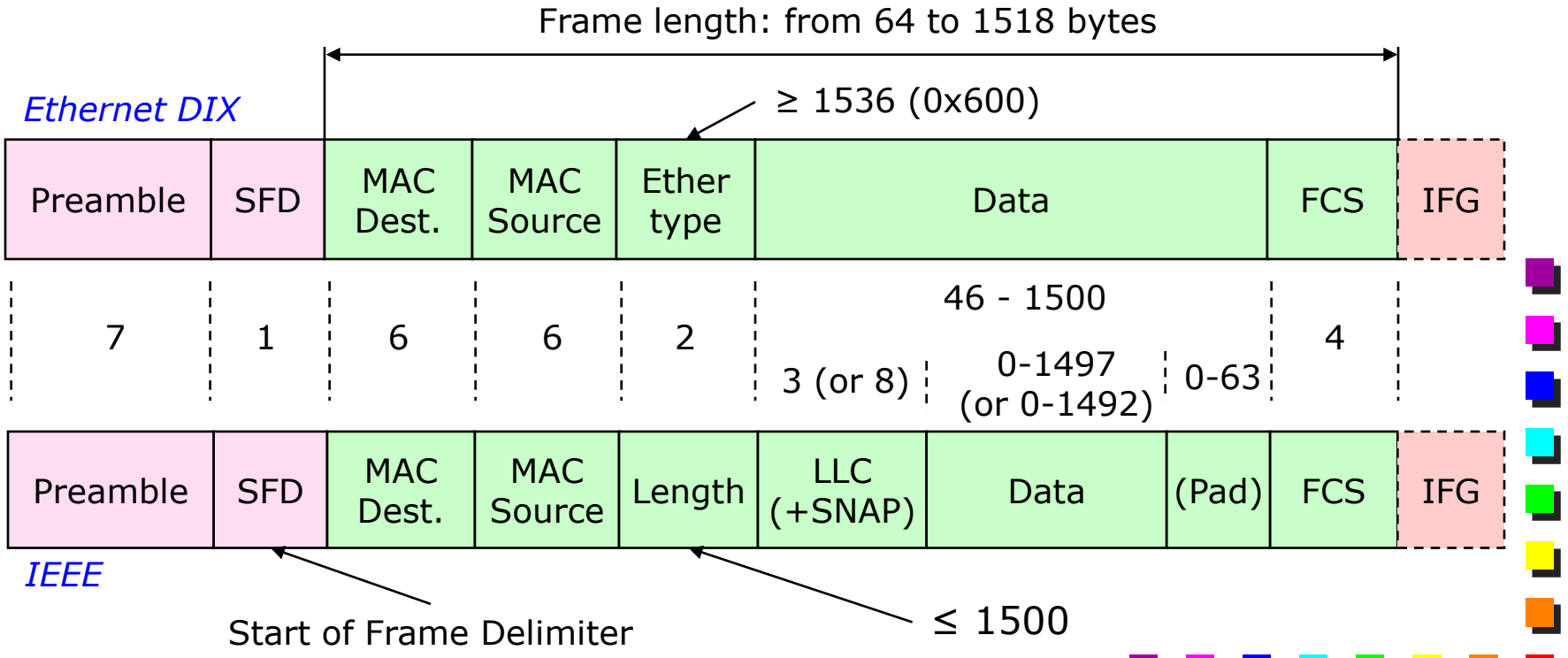
D_{max} = 5750 m

Ethernet v. 2.0 and IEEE 802.3 (1)



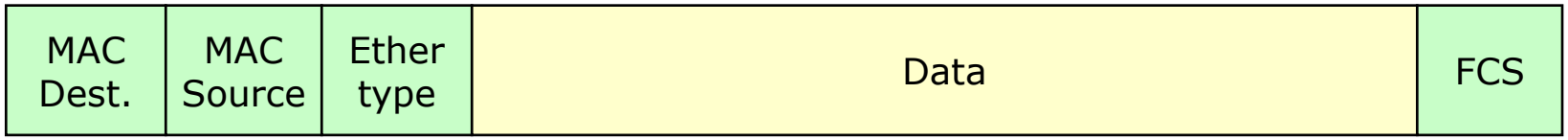
Ethernet v. 2.0 and IEEE 802.3 (2)

- Ethernet was specified and working before IEEE standards
  - Ethernet defined an "Ethertype" field for protocol demultiplexing
- IEEE had to specify a frame format compatible with the old Ethernet

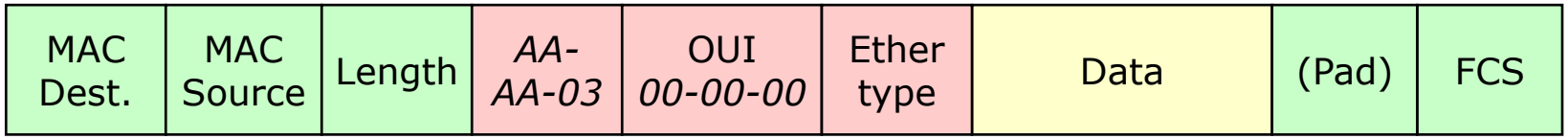


Ethernet v. 2.0 and IEEE 802.3 (3)

- Most protocols already working on the Ethernet retained the old format (version 2.0 or DIX)
  - E.g., IP has direct encapsulation in Ethernet
    - Why do we have to change something that is working fine?
    - Why do I have to add some extra complexity (btw, useless)?



Ethernet DIX



IEEE 802.3 + SNAP

Ethernet v. 2.0 and IEEE 802.3 (4)

- Some protocols that use direct encapsulation in Ethernet v2.0
  - IP: 0x800
  - ARP: 0x806
  - IPv6: 0x86DD
- Some (most IEEE-derived protocols such as 802.1d) use LLC
- Ethernet v. 2.0: missing “Padding” field
  - “Length” must be present in L3 protocols in order to detect the end of the packet
- IEEE 802.3: IFG kept for compatibility
  - Lo strictly needed because of the presence of the “length” field at MAC layer

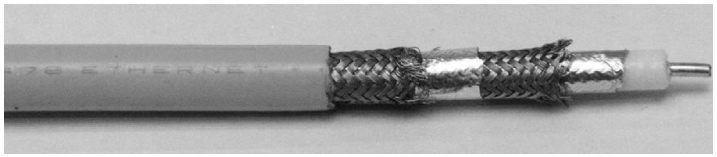
Ethernet physical layers

- 10 Mbps
  - 100ns bit time
- 802.3 defines different standards:
  - 10BaseT: twisted pairs (max 100 m)
- No longer in use:
  - 10Base5: thick coax (500 m)
  - 10Base2: thin coax (185 m)
  - FOIRL: fiber cable, asynchronous, for repeater-to-repeater connections (1000 m)
  - 10BaseFL: fiber cable, asynchronous, for stations or repeater-to-repeater connections (2000 m)
  - 10BaseFB: fiber cable, synchronous, for repeater-to-repeater connections (2000 m)
- Physical coding: usually Manchester
  - 10MHz

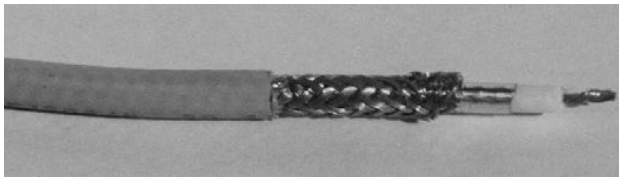
Physical layer: Coax cable

- No longer in use
  - Yellow cable (IBM)
  - “Thick” cable (RG213)
  - “Thin” cable (RG58)

RG213

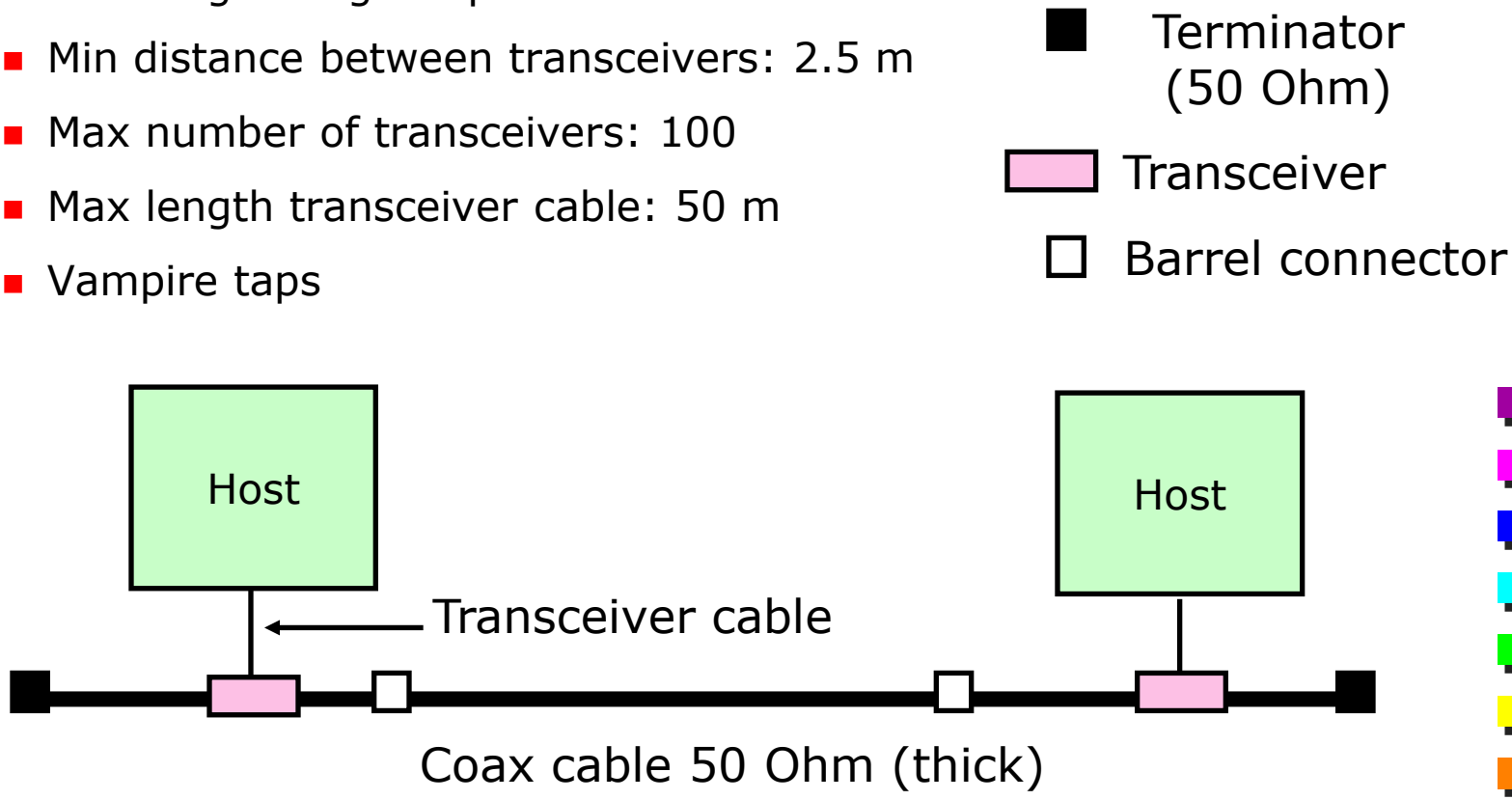


RG58



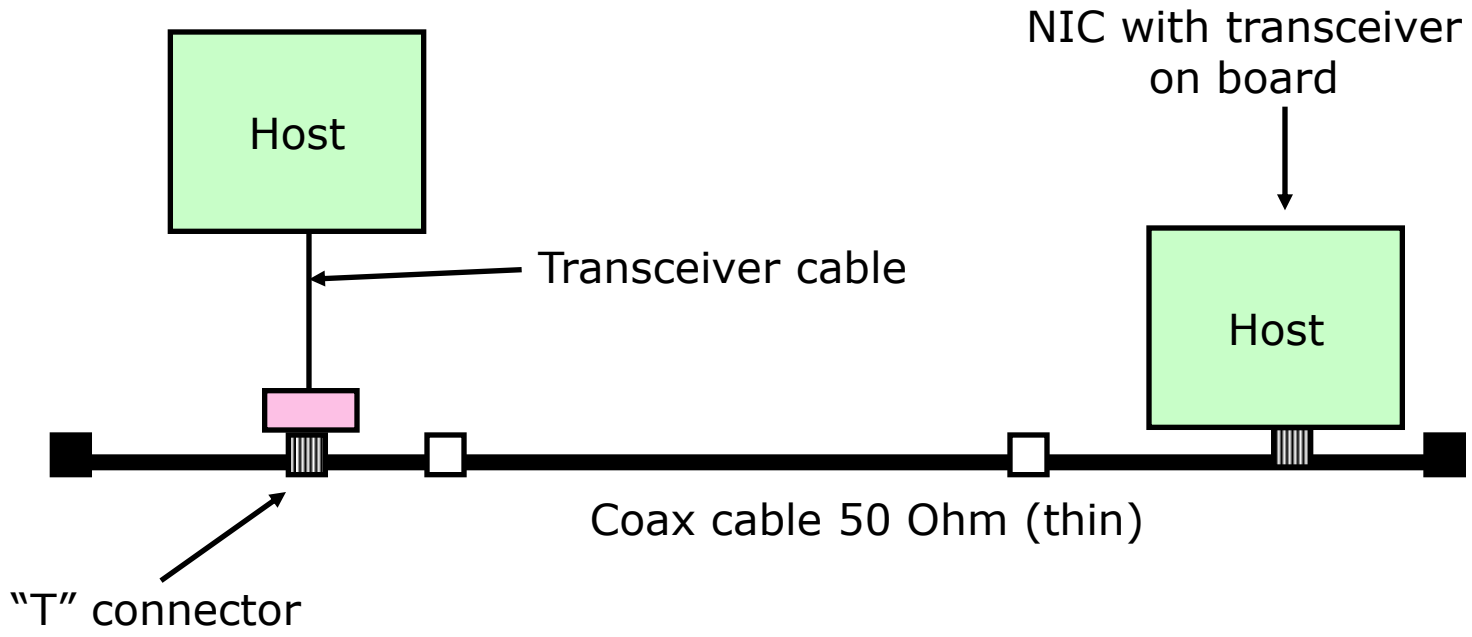
Physical layer: Thick coax cable

- Characteristics
  - Max length cable: 500 m
  - Max lenght single clip: 117 m
  - Min distance between transceivers: 2.5 m
  - Max number of transceivers: 100
  - Max length transceiver cable: 50 m
  - Vampire taps



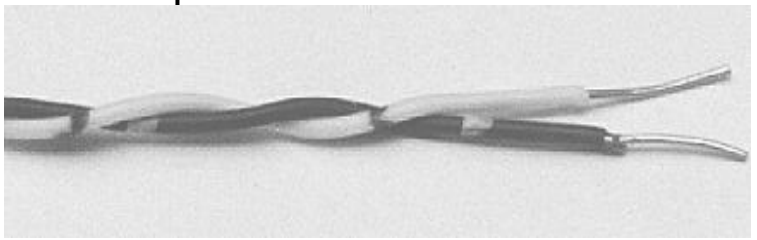
Physical layer: Thin coax cable

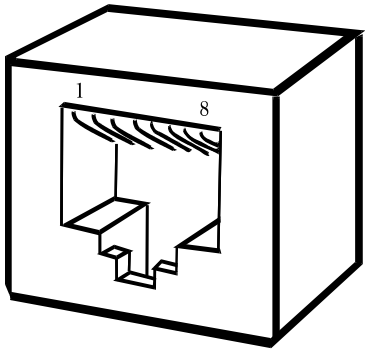
- Characteristics
  - Max length cable: 185 m
  - Max number of stations: 30
  - Min distance between stations: 0.5 m
  - Max length transceiver cable: 50 m



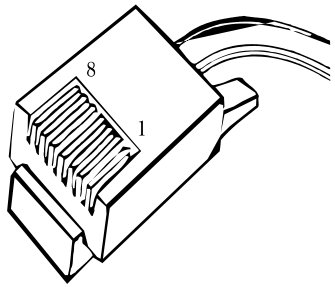
Twisted pair (1)

- UTP cable (min category 3)
- Max length: 100 m
  - Configuration details are more complicated





RJ45 wall socket

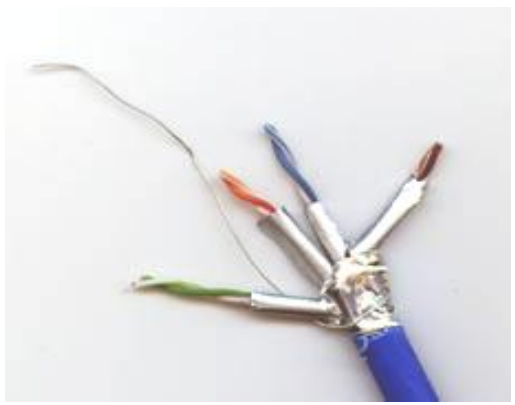


RJ45 connector

Twisted pair (2)

- Possible cables
  - UTP (Unshielded): not shielded
  - STP (Shielded): a single global shield
  - FTP (Foiled): global shield + a shield for each twisted pair
    - (also known as Fully Shielded Twisted Pair)



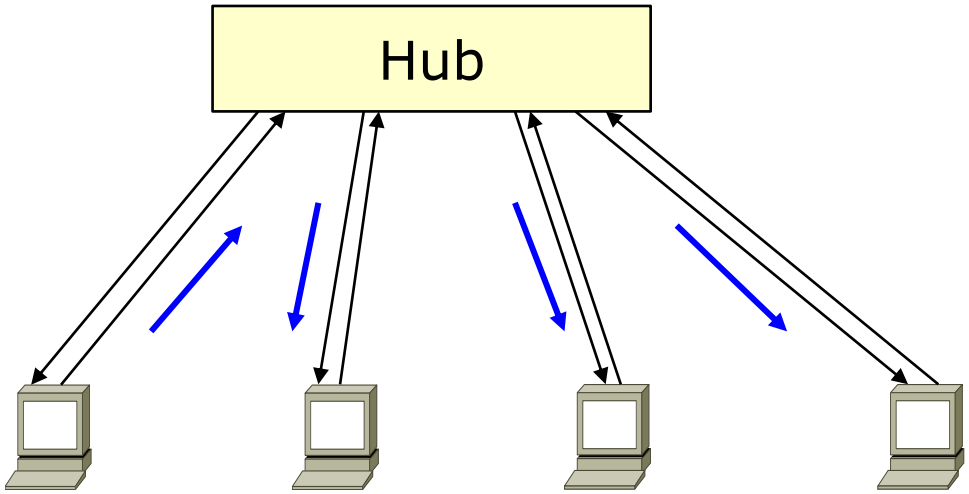



Fiber

- No sensitivity to electromagnetic fields
- Larger distances
- Higher costs
  - Cabling, crimping
- Less flexible
  - Fiber to the phone, in-field crimping

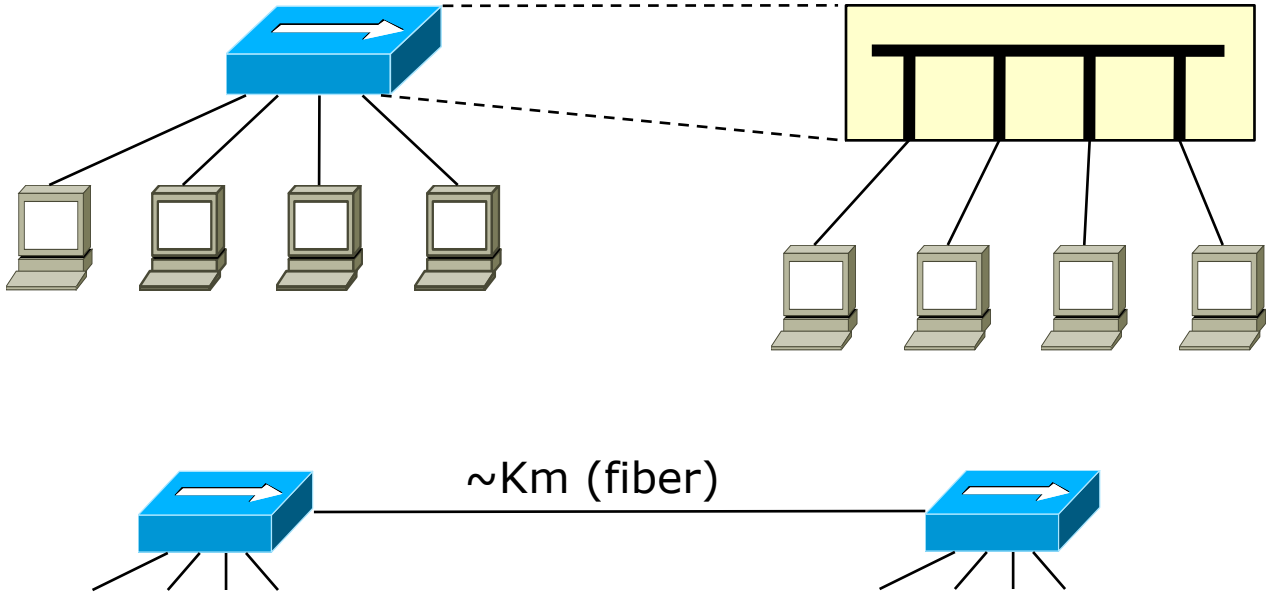
RX and TX on fiber and twisted pair cables

- Requires an intermediate device (hub)
  - Departure from the classical “bus” model
  - Hubs simulate the bus internally
- Two unidirectional cables
  - One TX only, one RX only
- Data repeated on all the other RX cables



Ethernet topology

- Rather limited
- ~200 m diameter
  - Due to twister pair cable limits
- Larger networks with repeaters and/or fiber (~ 3Km)



Conclusions

- Classical Ethernet networks (i.e., 10Mbps) no longer in use
  - Faster networks used
- However
  - Same principles
  - Same basic mechanisms
  - Some old choices still impact newer networks
- So… a detailed knowledge of the Ethernet is still needed