Ethernet

The LAN Killer

“Ethernet works in practice but not in theory.”

Robert Metcalfe

History (1)

- Late 1960s: **Aloha protocol University of Hawaii**
- Late 1972: **Robert Metcalfe developed first Ethernet system based on CSMA/CD**
  - **Xerox Palo Alto Research Center (PARC)**
  - **Exponental Backoff Algorithm was key to success (compared with Aloha)**
  - **2.94 Mbit/s**

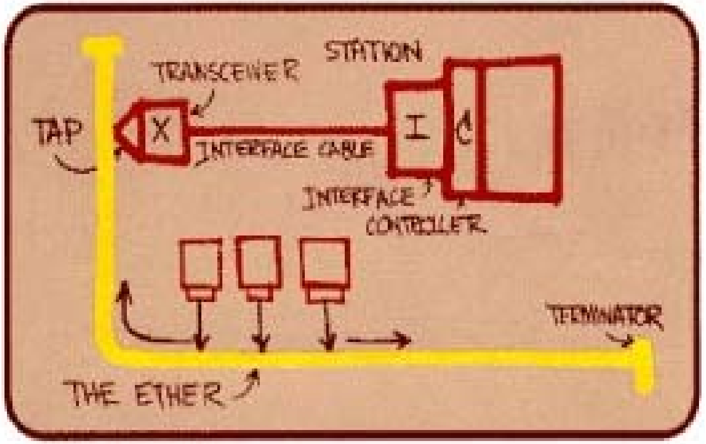
Original Ethernet Frame

Sync	Destination Address	Source Address	Data	CRC
1	8	8	about 4000 bits	16

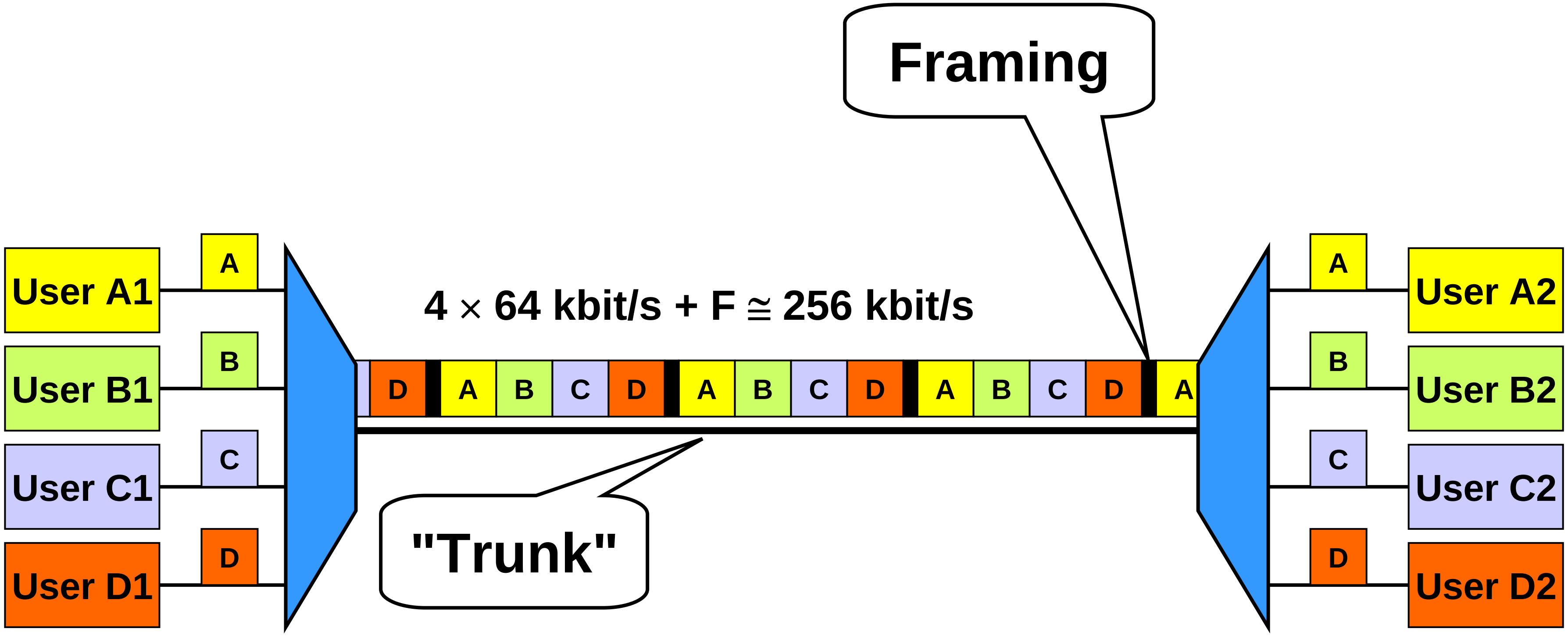
History (2)

- 1976: **Robert Metcalfe released the famous paper: "Ethernet: Distributed Packet Switching for Local Computer Networks"**

Original sketch

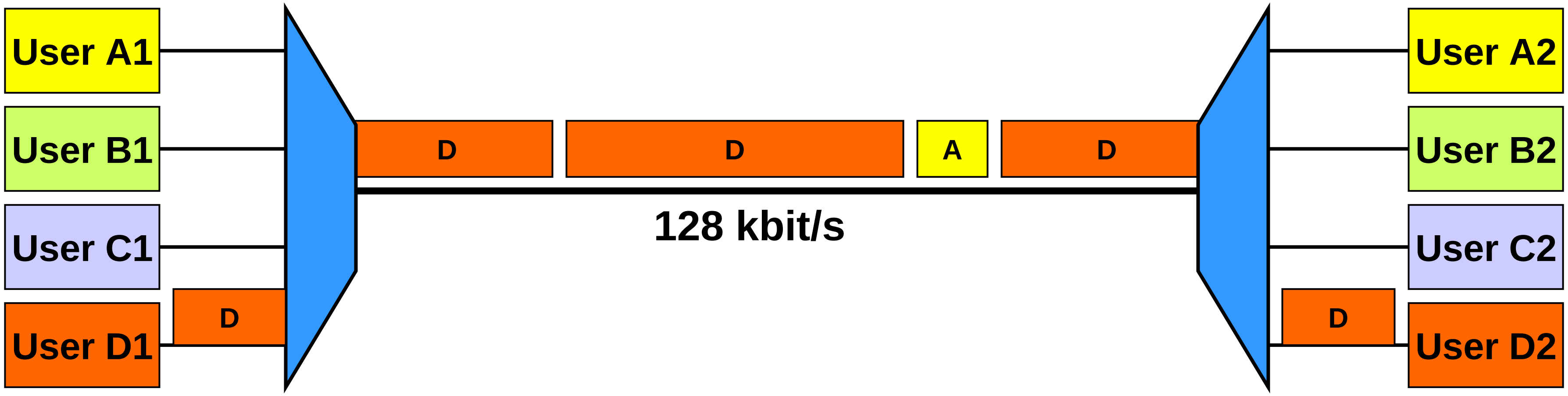


Deterministic (synchronous) TDM



Trunk speed = Number of slots × User access rate
Each user gets a constant timeslot of the trunk

Statistical (asynchronous) TDM

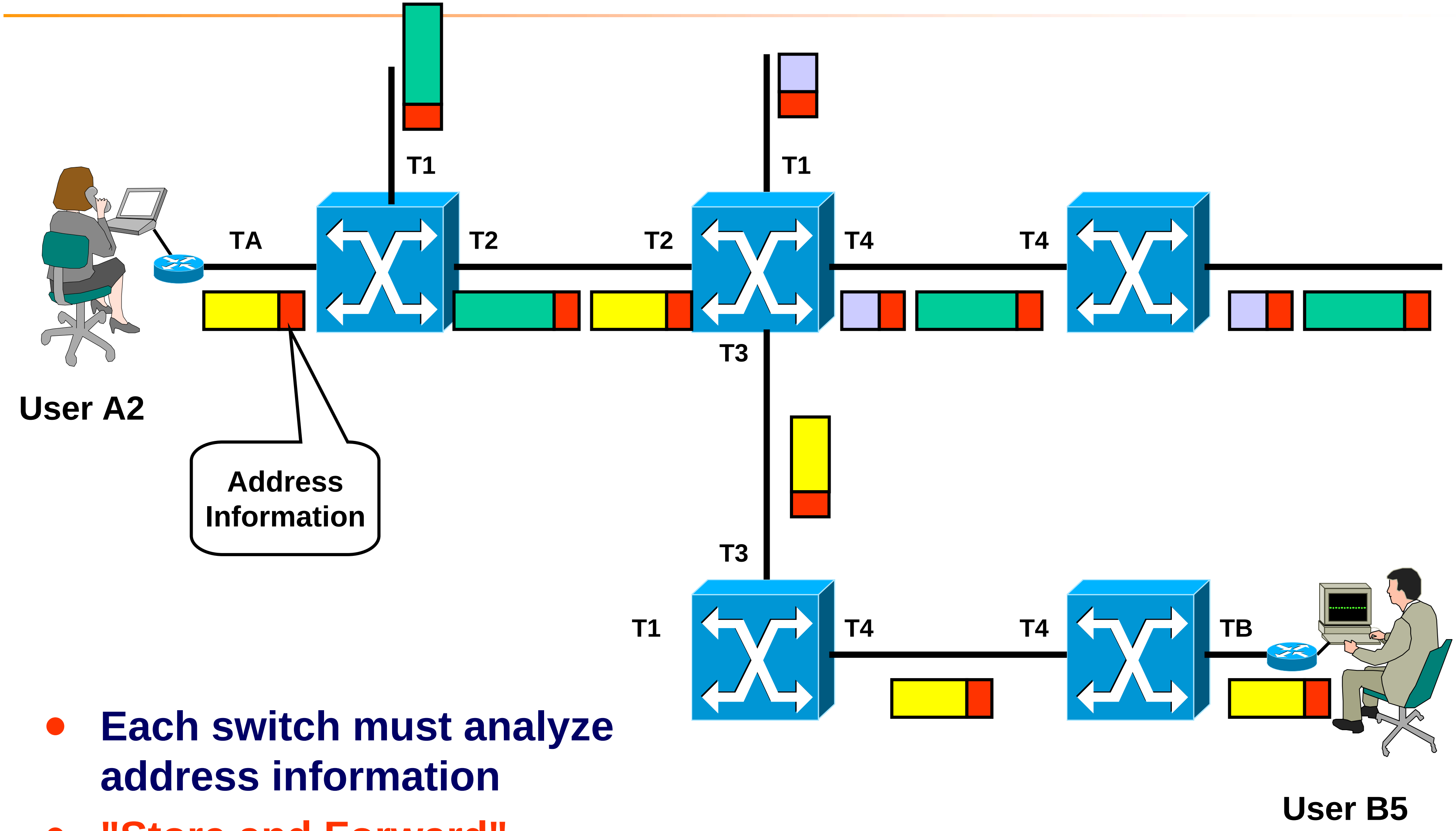


- If other users are silent, one (or a few) users can fully utilize their access rate

Networking Techniques

- **Synchronous or asynchronous time division multiplexing principles used in a network environment**
  - **<u>Circuit switching</u> based on synchronous TDM**
  - **<u>Packet switching</u> based on asynchronous (statistical) TDM**

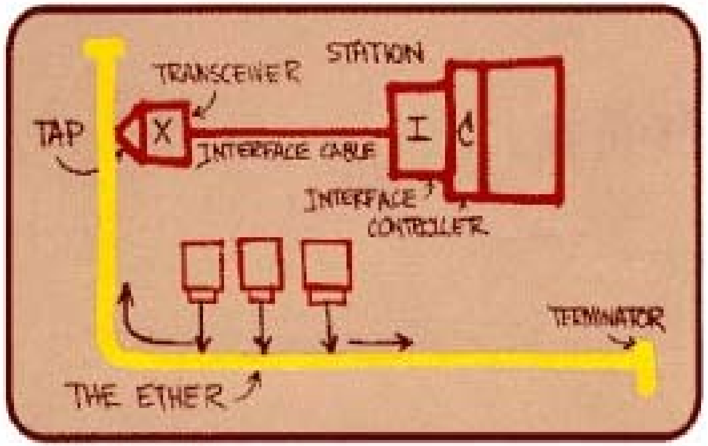
Packet Switching



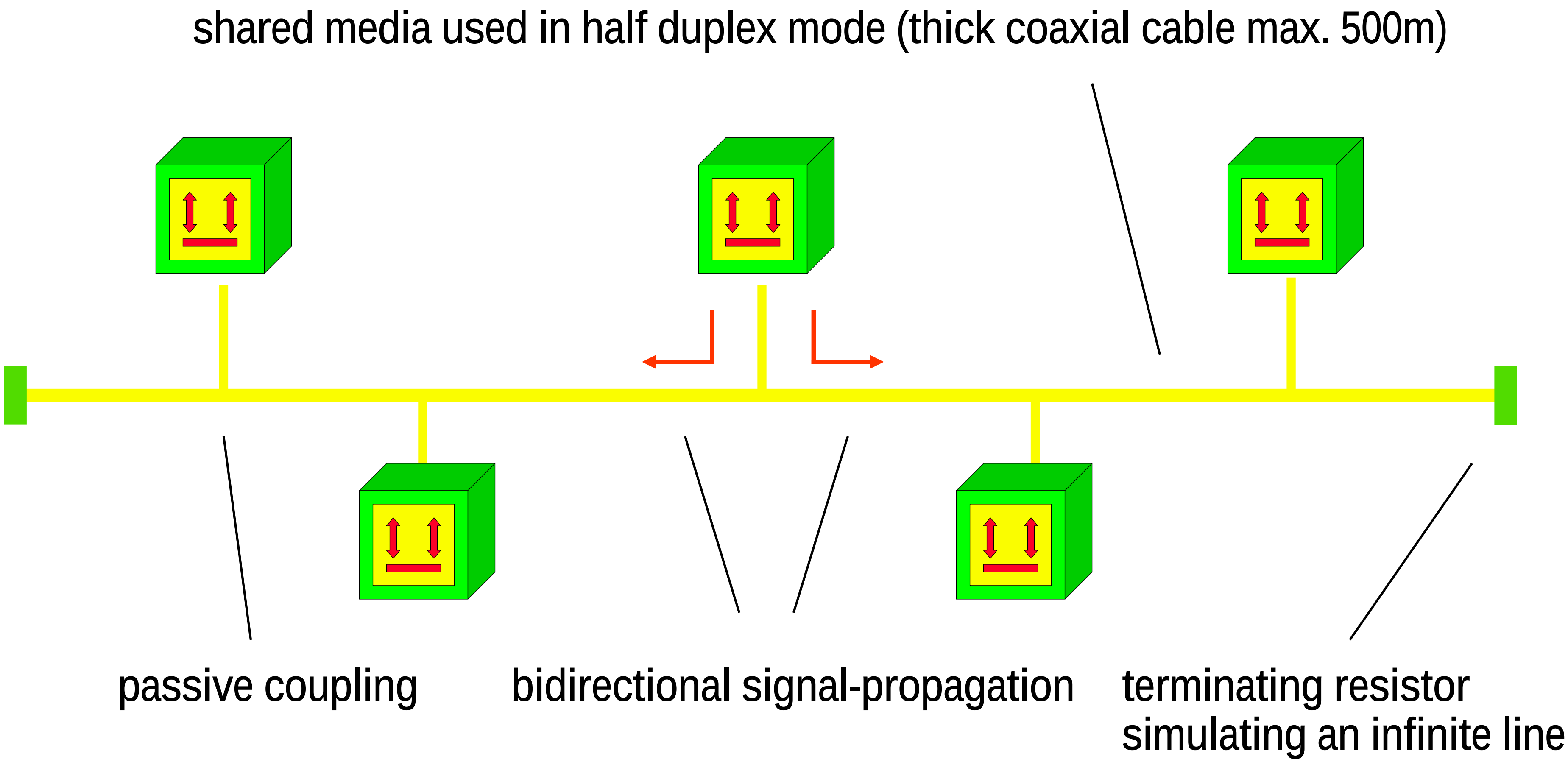
- Each switch must analyze address information
- "Store and Forward"

- **"Ethernet: Distributed Packet Switching for Local Computer Networks"**

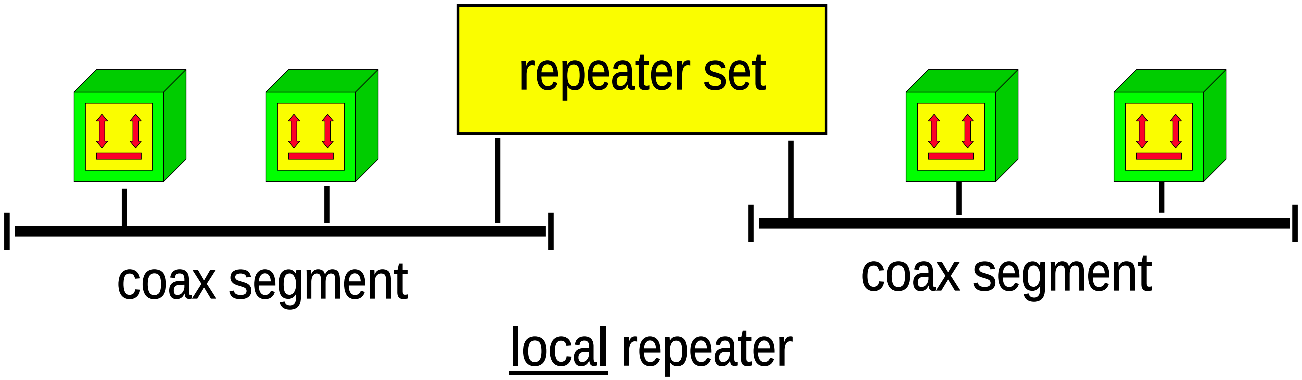
Original sketch

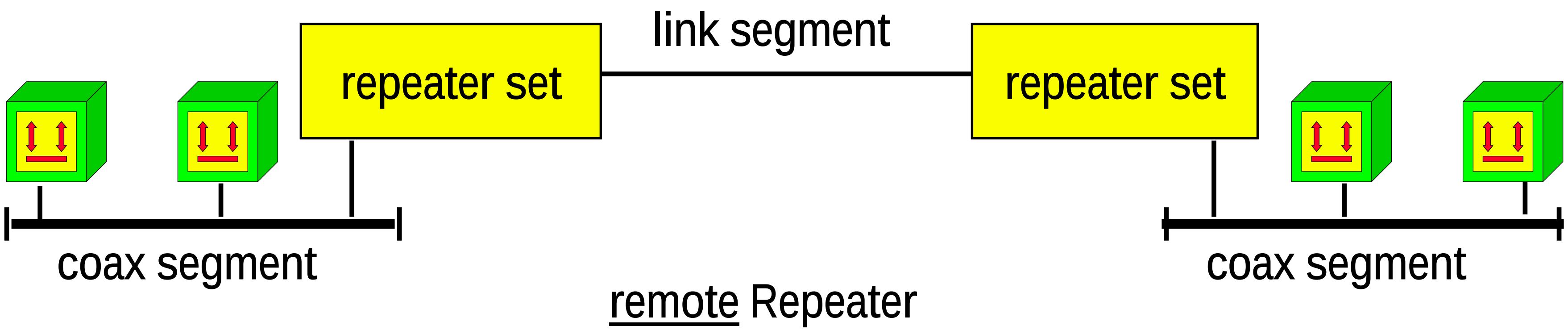


Basic Idea of Ethernet Bus System



Local / Remote Repeater





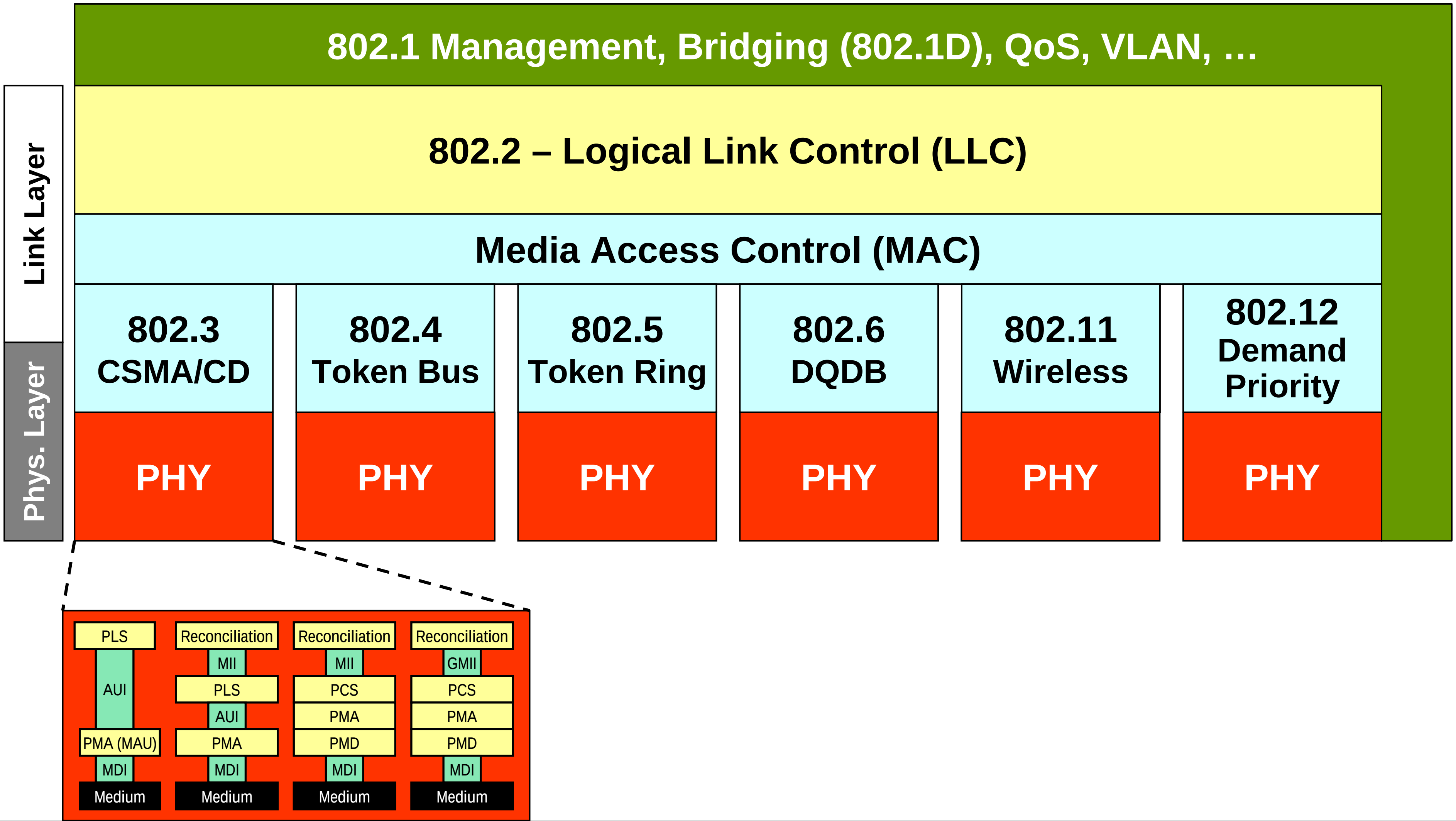
History (3)

- 1978: **Patent for Ethernet-Repeater**
- 1980: **DEC, Intel, Xerox (DIX) published the 10 Mbit/s Ethernet standard**
  - **"Ethernet II" was latest release (DIX V2.0)**
- Feb 1980: **IEEE founded workgroup 802**
- 1985: **The LAN standard IEEE 802.3 had been released**

IEEE 802

- **LAN Standardization is done**
  - **by IEEE (Institute of Electrical and Electronics Engineers)**
  - **OSI Layer 1 and 2 are suffieient for communication between two LAN stations**
- **But OSI Data Link Layer (Layer 2)**
  - **was originally designed for point-to-point line**
  - **but LAN = multipoint line, shared media**
- **Therefore OSI Layer 2 must be split into two sublayers**
  - **Logical Link Control**
  - **Media Access Control**

IEEE 802 Layer Model



Tasks of LAN Layers

- Layer 1
  - physical layer (PHY) specifies actual transmission technique
  - provides
    - electrical/optical and mechanical interface
    - encoding
    - bit synchronisation
  - consists of
    - MAU (Medium Attachment Unit)
    - AUI (Attachment Unit Interface)
    - PLS (Physical Layer Signalling)

Tasks of LAN Layers

- **Layer 2**
  - **MAC (Media Access Control) takes care for medium access algorithms, framing, addressing and error detection**
    - **avoid collisions**
    - **grant fairness**
    - **handle priority frames**
  - **LLC (Logical Link Control) provides original services of data link layer**
    - **connection-oriented services**
    - **connection-less service**
    - **SAPs (Service Access Points) for the higher layers**

The IEEE Working Groups

- 802.1 Higher Layer LAN Protocols
- 802.2 Logical Link Control
- 802.3 Ethernet
- 802.4 Token Bus
- 802.5 Token Ring
- 802.6 Metropolitan Area Network
- 802.7 Broadband TAG
- 802.8 Fiber Optic TAG
- 802.9 Isochronous LAN
- 802.10 Security
- 802.11 Wireless LAN
- 802.12 Demand Priority
- 802.13 Not Used
- 802.14 Cable Modem
- 802.15 Wireless Personal Area Network
- 802.16 Broadband Wireless Access
- 802.17 Resilient Packet Ring

Superstition?

IEEE 802.3/Ethernet

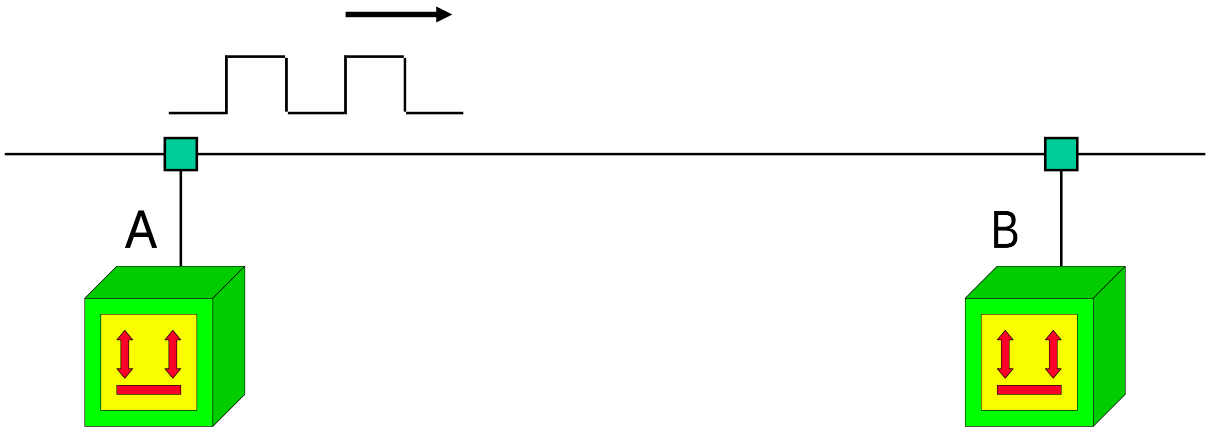
- Since 1984 the IEEE also maintains the DIX Ethernet standard
- Both frame types are supported by "Ethernet NICs"
  - Network Interface Cards

CSMA/CD

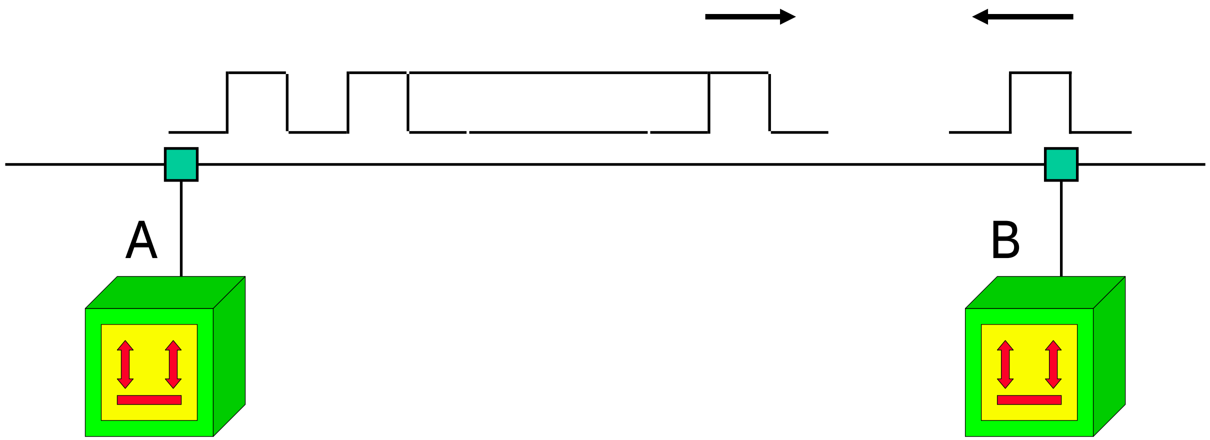
- **Carrier Sense Multiple Access Collision Detection**
  - **Improvement of ALOHA**
  - **"Listen before talk" plus**
  - **"Listen while talk"**
- **Fast and low-overhead way to resolve any simultaneous transmissions**

1) Listen if a station is currently sending
2) If wire is empty, send frame
3) Listen during sending if collision occurs
4) Upon collision stop sending
5) Wait a random time before retry

Collision Window / Slot Time 1

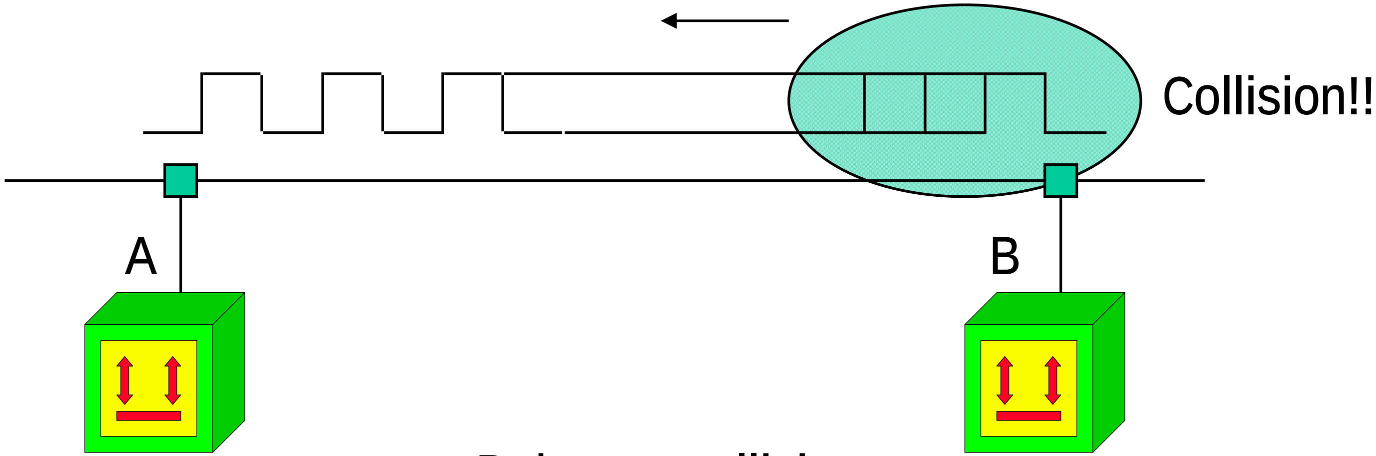


t = x ... A starts transmission

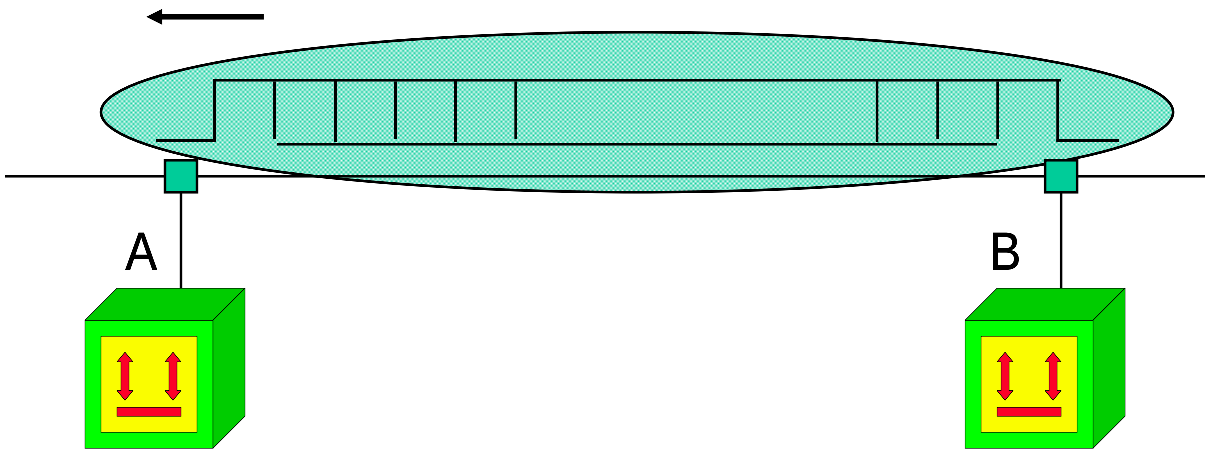


t = x + dt ... B starts transmission

Collision Window / Slot Time 2



t = x+ tp ... B detects collision
tp ... signal propagation time from A to B



t = x + 2tp ... A detects collision

Slot Time

- Minimum frame length has to be defined in order to safely detect collisions
- Each frame sent must stay on wire for a RTT duration – at least
- This duration is called "slot time" and has been standardized to be 512 bit-times
  - 51,2 µs for 10 Mbit/s

Slot Time Consequences

- So minimum frame length is 512 bits (64 bytes)
- With signal speed of 0.6c the RTT of 512 bit times allows a network diameter of
  - 2500 meters with 10 Mbit/s
  - 250 meters with 100 Mbit/s
  - 25 meters with 1000 Mbit/s (!)

NOTE:
Only valid on shared media (!)

Exponential Backoff (1)

- Most important idea of Ethernet !
- Provides maximal utilization of bandwidth
  - After collision, set basic delay = 512 x slot time
  - Total delay = basic delay * rand
  - 0 <= rand < 2^k
    - k = min (number of transm. attempts, 10)
- Allows channel utilization

Exponential Backoff (2)

- **After 16 successive collisions**
  - **Frame is discarded**
  - **Error message to higher layer**
  - **Next frame is processed, if any**
- **Truncated Backoff (k<=10)**
  - **1024 potential "slots" for a station**
  - **Thus maximum 1024 stations allowed on half-duplex Ethernet**

Channel Capture

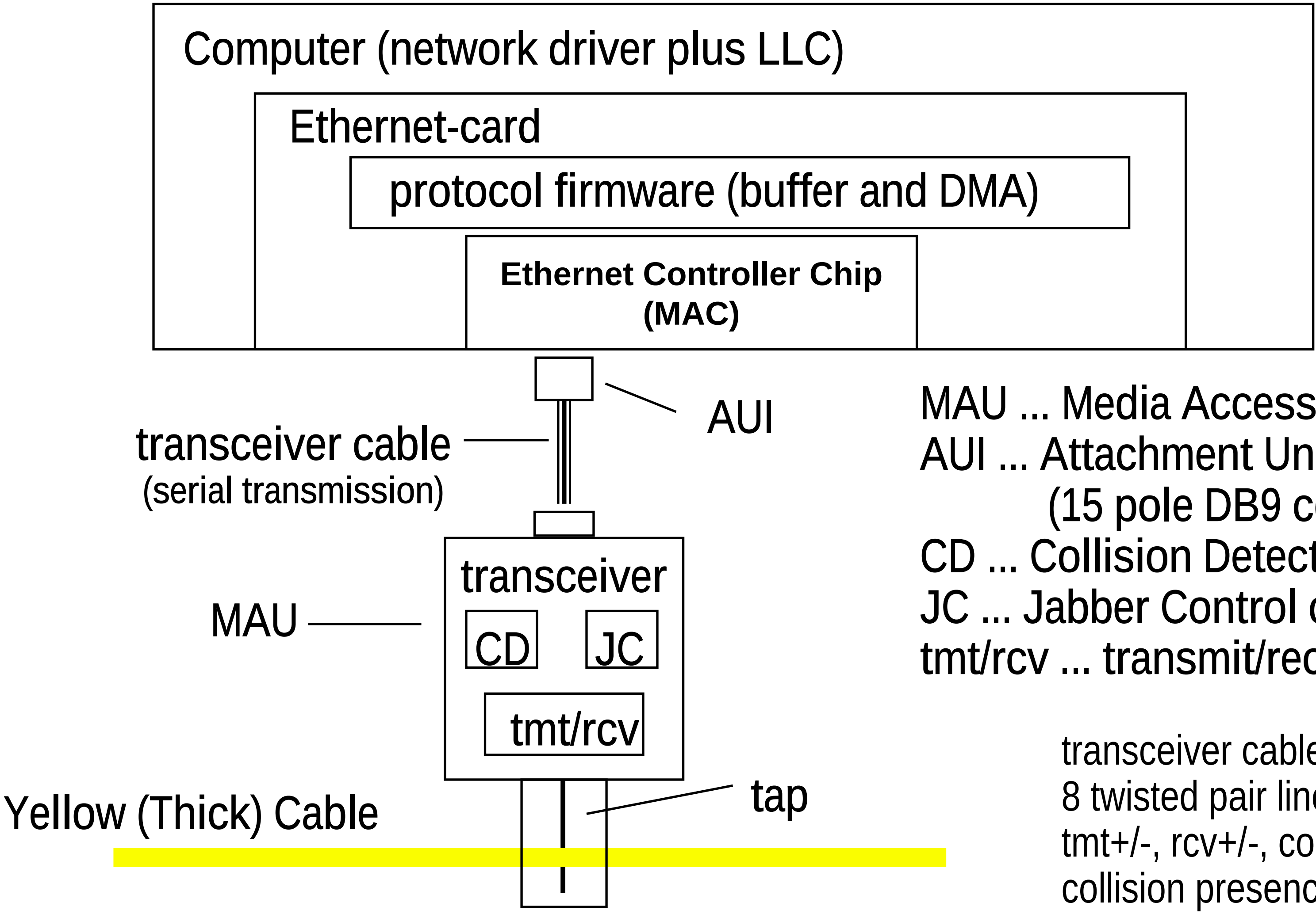
- **Short-term unfairness on very high network loads**
- **Stations with lower collision counter tend to continue winning**
- **10 times harder to occur on 100 Mbit/s Ethernet**
- **Rare phenomena, so no solution against it**

But would I choose Ethernet for mission-critical realtime applications…?

Collision Detection

- **10Base2, 10Base5**
  - **Manchester with –40 mA DC level**
  - **"high" = 0 mA, "low" = –80 mA**
- **10BaseT**
  - **Manchester with no DC offset**
  - **Collisions are detected by Hub who sends a "Jam" signal back**
  - **Similarily at 100BaseT and 1000BaseT**

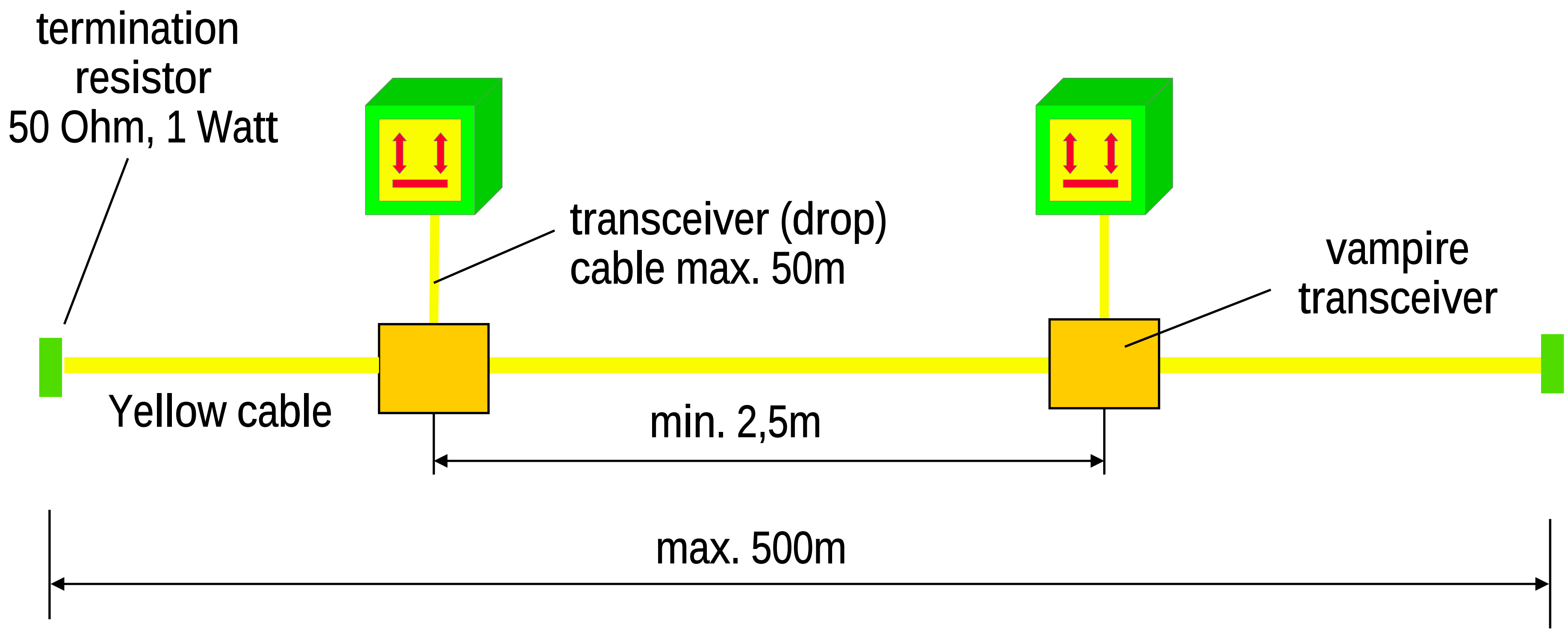
AUI-Connection with 10Base5 Transceiver



MAU ... Media Access Unit
AUI ... Attachment Unit Interface
(15 pole DB9 connector)
CD ... Collision Detecting circuits
JC ... Jabber Control circuits
tmt/rcv ... transmit/receive circuits

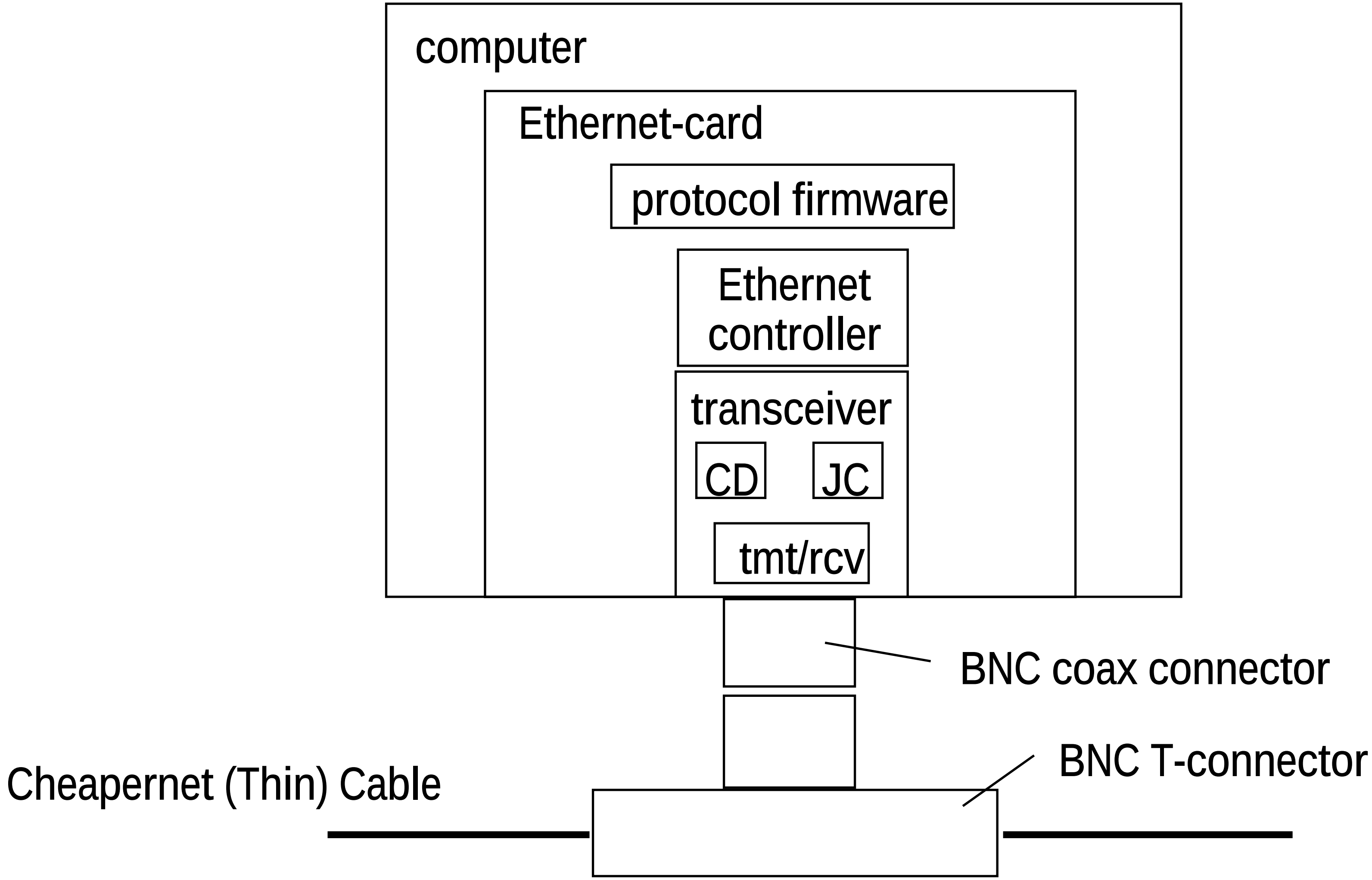
transceiver cable:
8 twisted pair lines for
tmt+/-, rcv+/-, control +/-
collision presence +/-,
3 lines for power, earth, shield

10Base5 Parameter

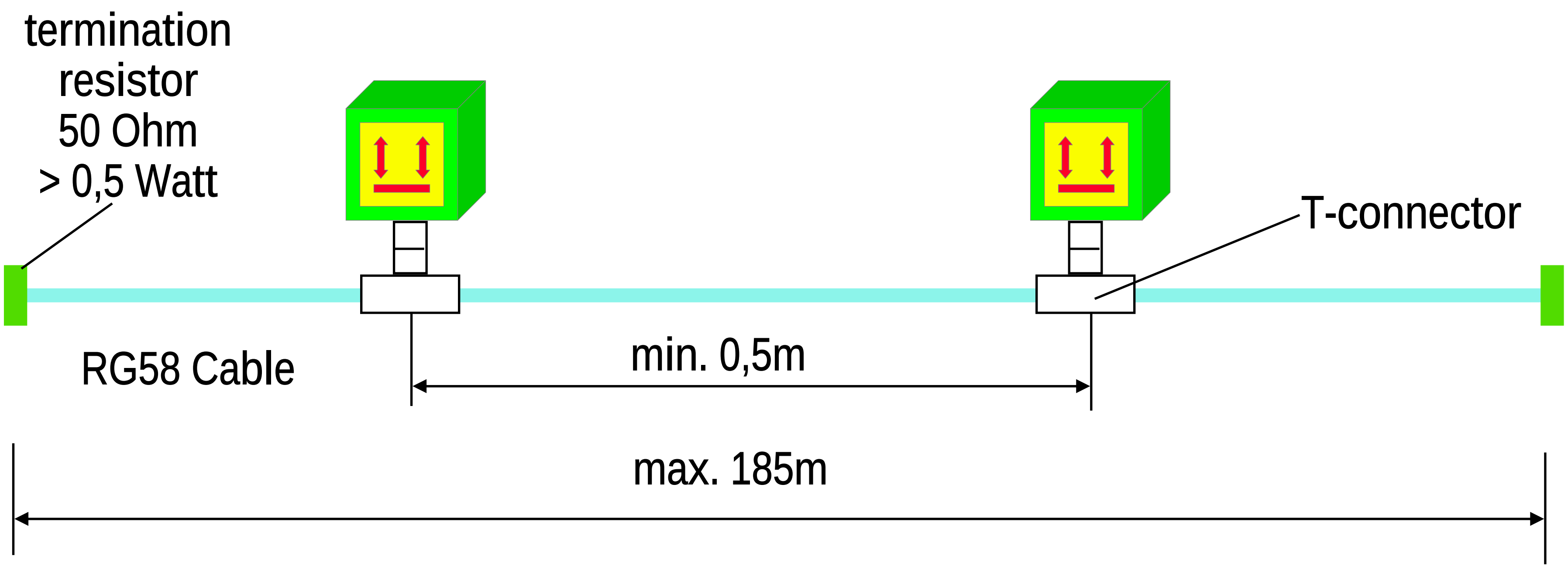


- maximal number of stations: 100
- attachable only at marked points
- cable splitting via coax couplers
- individual cable parts have a length of 23,4m or 70,2m or 117,5m (wave minimum on standing waves due to inhomogeneous media)
- smallest bending radius: 254mm

Integrated Transceiver for 10Base2

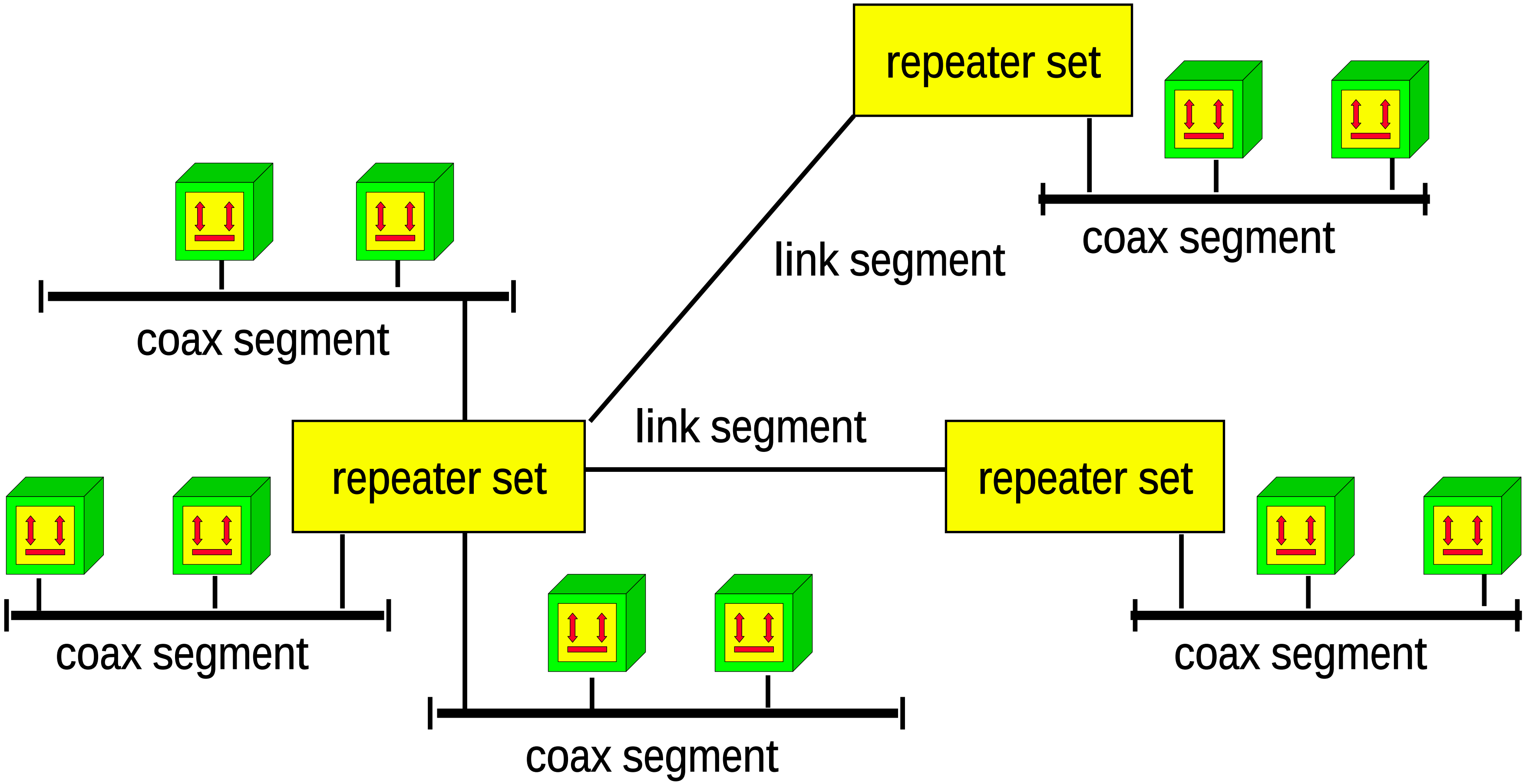


10Base2 Parameter



- maximal number of stations: 30
- attachable at any points
- smallest bending radius: 50 mm

Multiport Repeater - One Collision Domain



Link Segments for Repeater Interconnection

- link segment
  - first implementation for repeater interconnection only
  - point-to point connection
    - only two devices are connected by a physical cable
  - several types were defined
    - fibre based
    - copper based
  - FOIRL (Fiber Optic Inter Repeater Link)
    - maximal length 1000m
    - first FO specification
    - repeater – repeater
  - 10BaseFL (Fiber)
    - asynchronous
    - maximal length 2000m
    - repeater - repeater, end system - multiport repeater
  - 10BaseFB (Fiber)
    - synchronous (idle signals during communication pauses)
    - maximal length 2000m
    - for repeater - repeater links only

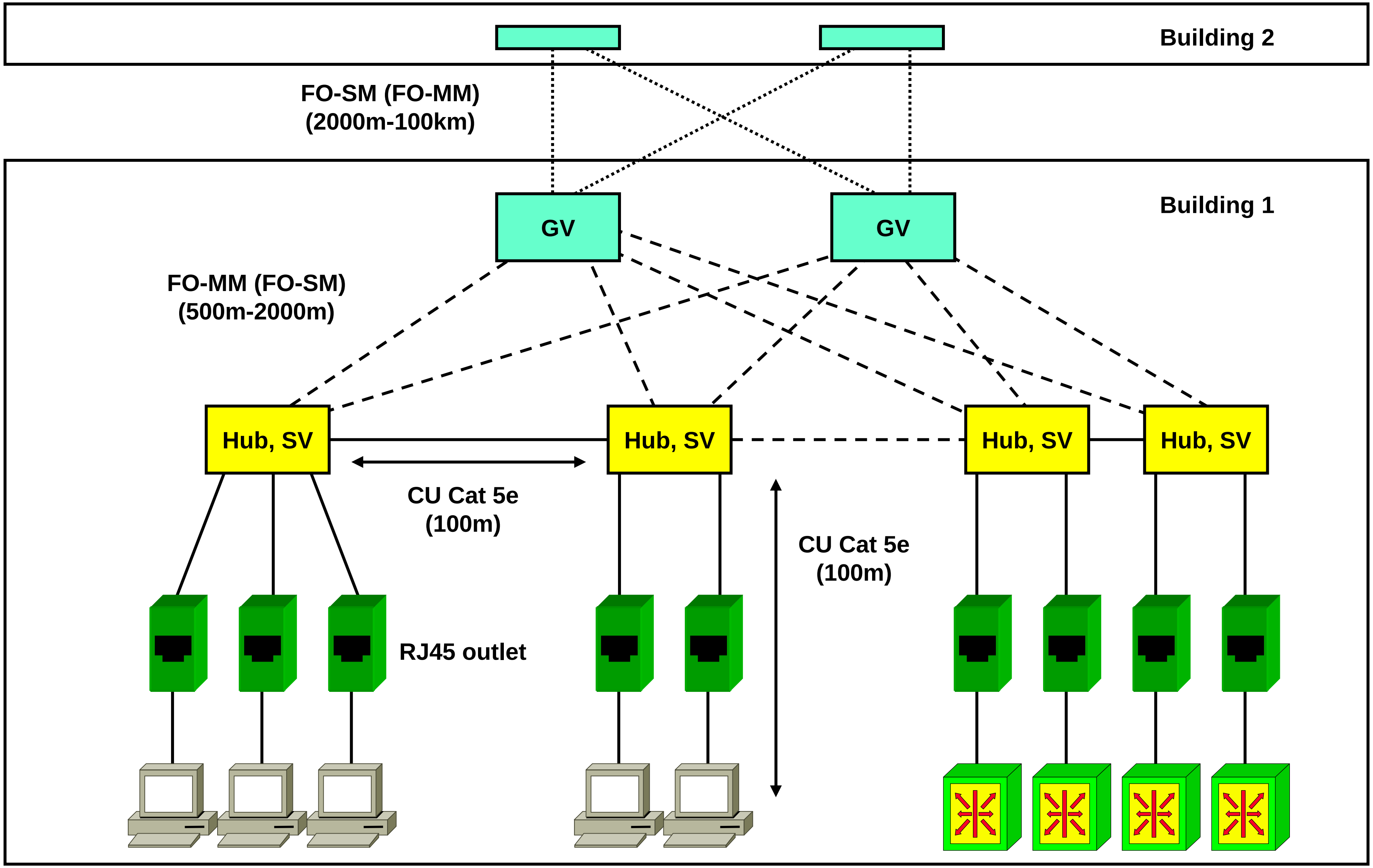
Link Segments for End Systems

- link segment
  - was later also defined for connection of a network station (end system) to a multiport repeater
    - using a dedicated point-to-point line
  - 10BaseT (unshielded twisted pair)
    - maximal length 100m
    - 2 lines Tmt+-, 2 lines Rcv+-, RJ45 connector
  - Reason for that:
    - Ethernet was originally based on coax cabling and bus topology
    - later an international standard for structured cabling of buildings was defined
      - star wired to a central point(s)
      - based on twisted pair cabling
    - that excellently fits to Token ring cabling
    - Ethernet had been adapted to that in order to survive

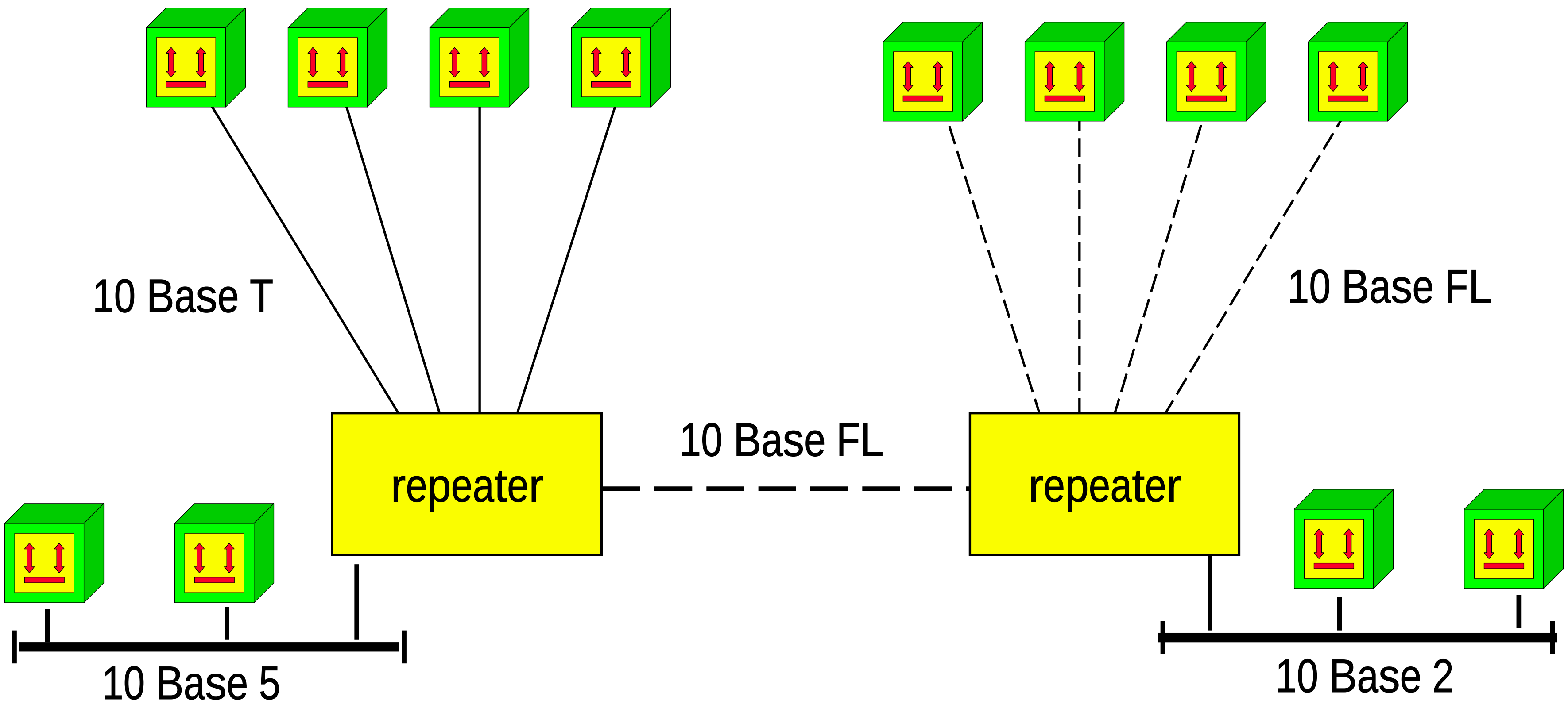
Structured Cabling (LAN)

- **Physical Wiring**
  - Should follow the principle of structured cabling
  - Primary
    - End system to first “Hub” (Repeater or nowadays a L2 Switch)
      - “Stockwerkverteiler”
    - CU-UTP, Category 5e or better
    - FO for extreme conditions only
  - Secondary
    - Hubs to central functions
      - “Gebäudeverteiler”
    - FO-MM (FO-SM)
  - Tertiary
    - Interconnections of buildings
    - FO-MM (FO-SM)

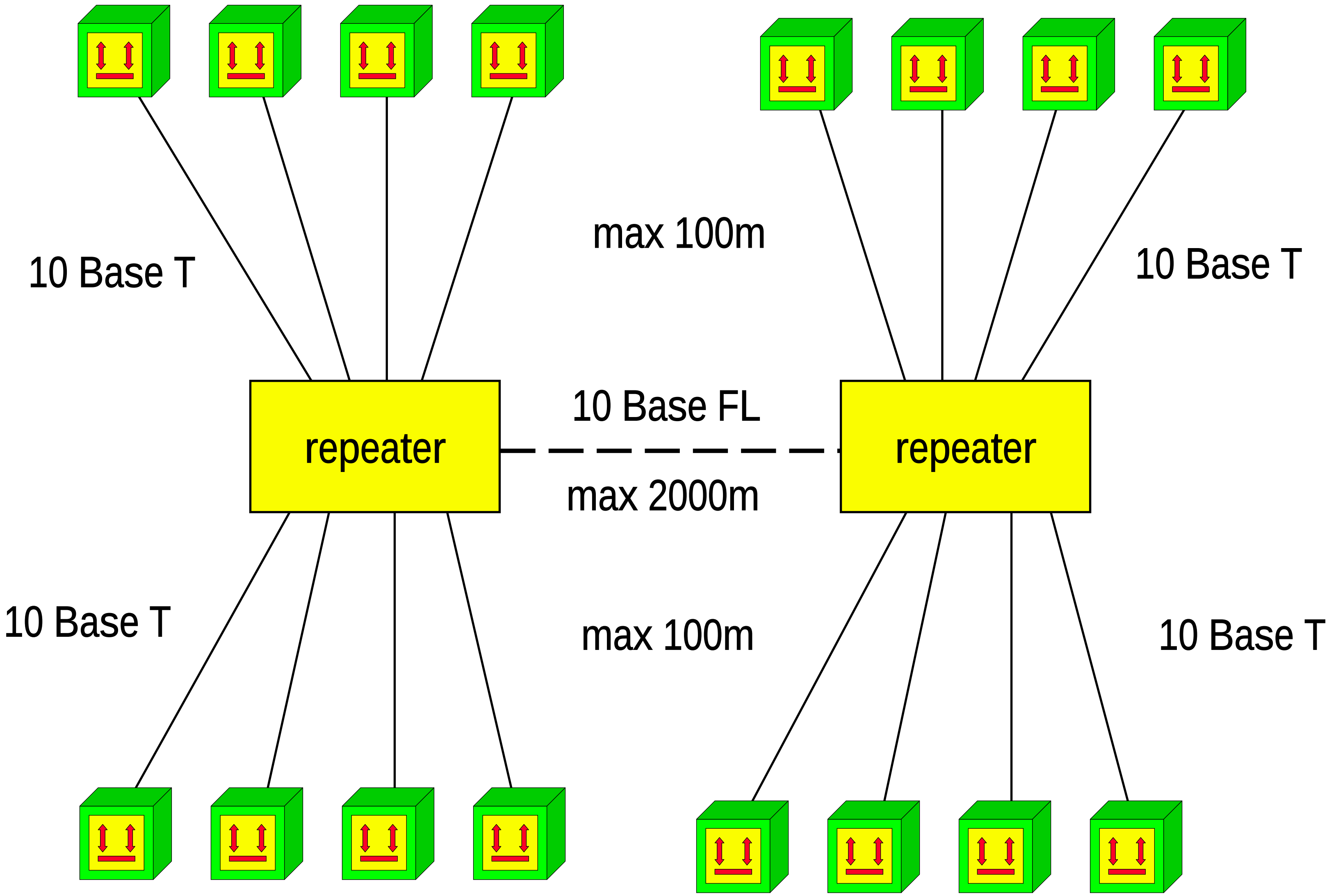
Structured Cabling (LAN)



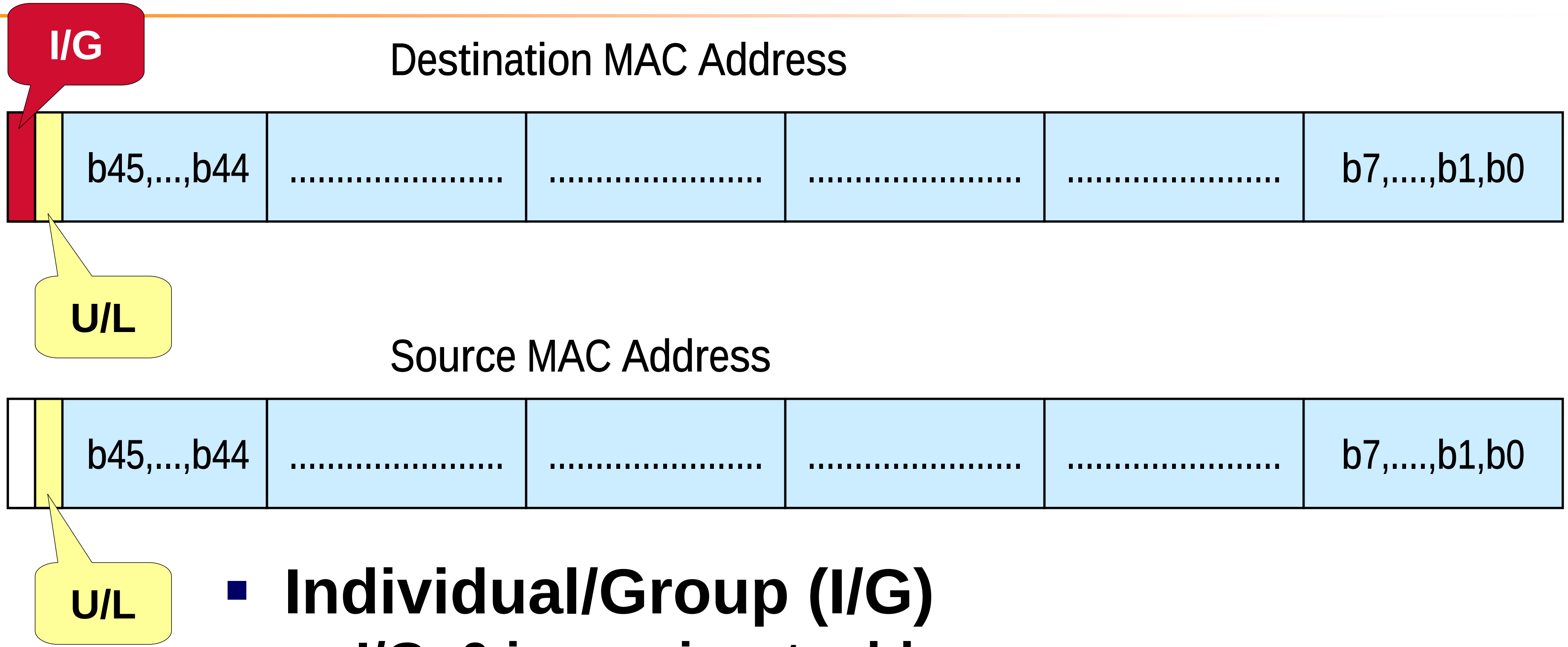
Multiport - Repeater



Multiport Repeater as „Hub“



6 Byte MAC Addresses



- **Individual/Group (I/G)**
  - **I/G=0 is a unicast address**
  - **I/G=1 is a group (broadcast) address**
- **Universal/Local (U/L)**
  - **U/L=0 is a global, IEEE administered address**
  - **U/L=1 is a local administered address**

Storage Format of 802.3 MAC-Address

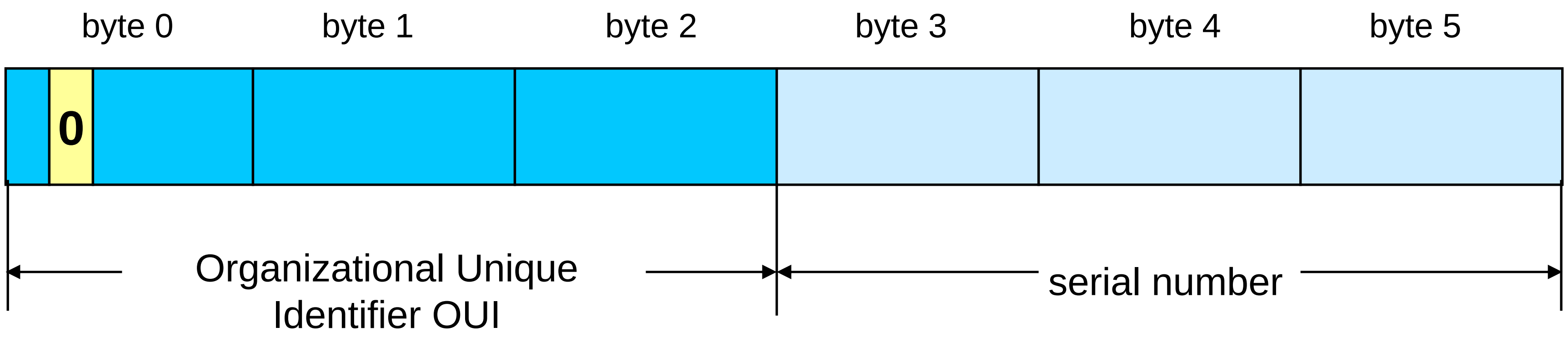
- **basic rule:**
  - I/G bit must be the first bit on the medium, so the transmitted address must have the following format:

I/G	U/L	b45, ..., b40	b39, ..., b32	...	b15, ..., b8	b7, ..., b1, b0
0	1	00 1000	0000 0000	....	0000 0000	1100 0001

  - <u>802.3 sends the least significant bit of each byte at first</u>
  - so 802.3 must store each byte in memory in reverse order:
    - also called "Canonical" Format

b40, ..., b45	U/L	I/G	b32, ..., b39	...	b8, ..., b15	b0, b1, ..., b7
0001 00	1	0	0000 0000	....	0000 0000	1000 0011

MAC Address Structure



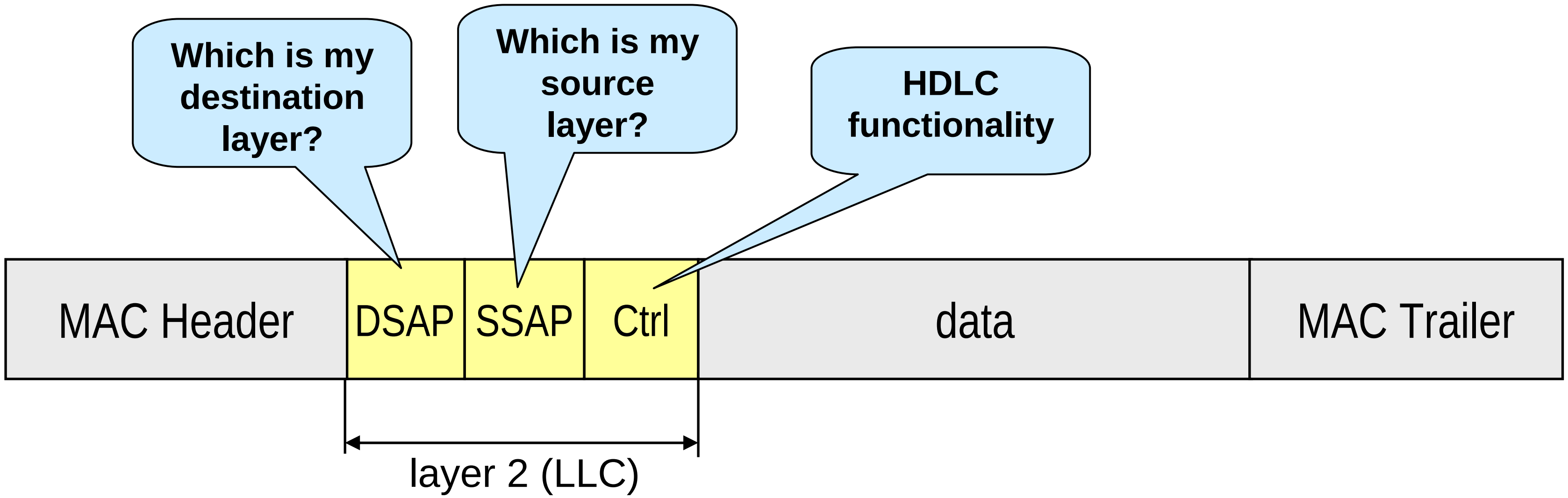
- **Each vendor of networking component can apply for an unique vendor code**
- **Administered by IEEE**

Ethernet Frames

- **Due to different development branches, there are two different frame types**
  - **IEEE type: consists of MAC and LLC**
  - **DIX type: consists of a Type field**
- **Why using both?**
  - **Different applications have been defined for either IEEE or DIX**

IEEE 802.2 (LLC)

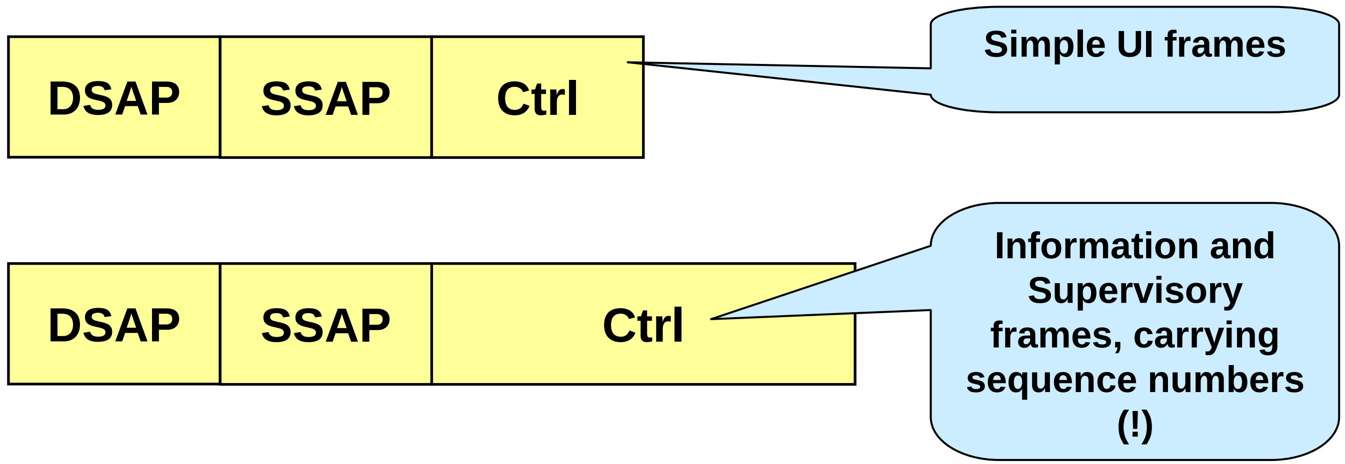
- **Every IEEE LAN/MAN protocol carries the Logical Link Control header**
  - **DSAP (Destination Service Access Point),**
  - **SSAP (Source Service Access Point)**
  - **Control Field = HDLC heritage**



Basic frame format of every IEEE protocol

LLC Details

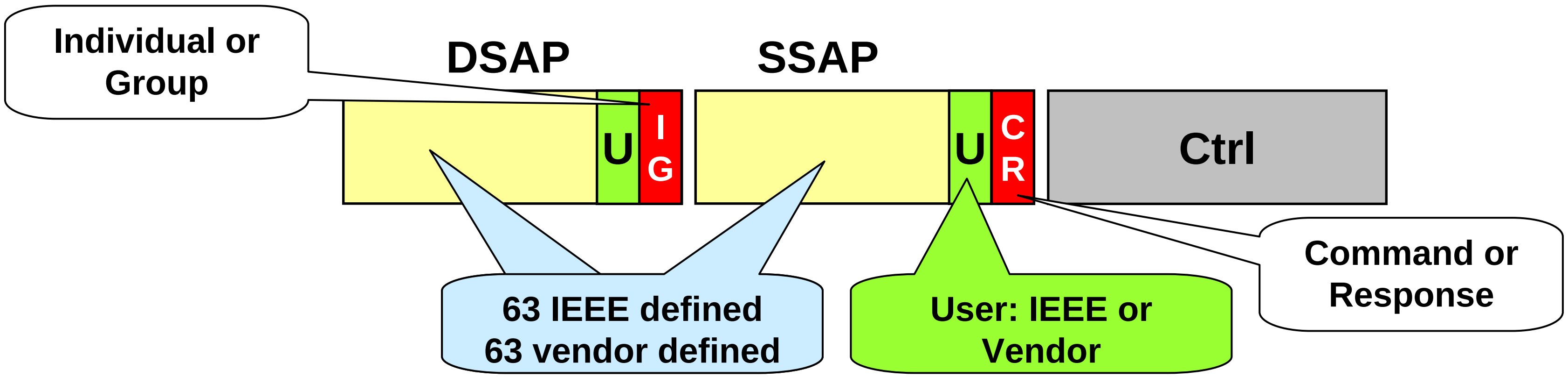
- **According sophisticated HDLC functionalities, 4 LLC classes defined**
  - **Class 1 is most important (UI, no ACKs)**



Either 1 or 2 bytes for control field

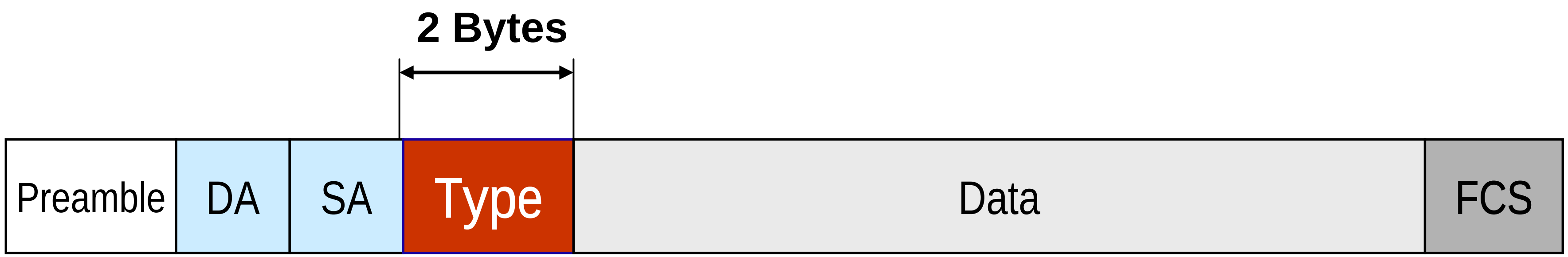
SAP Identifiers

- 128 possible values for protocol identifiers
- Examples:
  - 0x42 … Spanning Tree Protocol 802.1d
  - 0xAA… SNAP
  - 0xE0… Novell
  - 0xF0… NetBios



DIX Type field

- 2-bytes Type field to identify payload (protocols carried)
  - Most important: IP type 0x800
- No length field

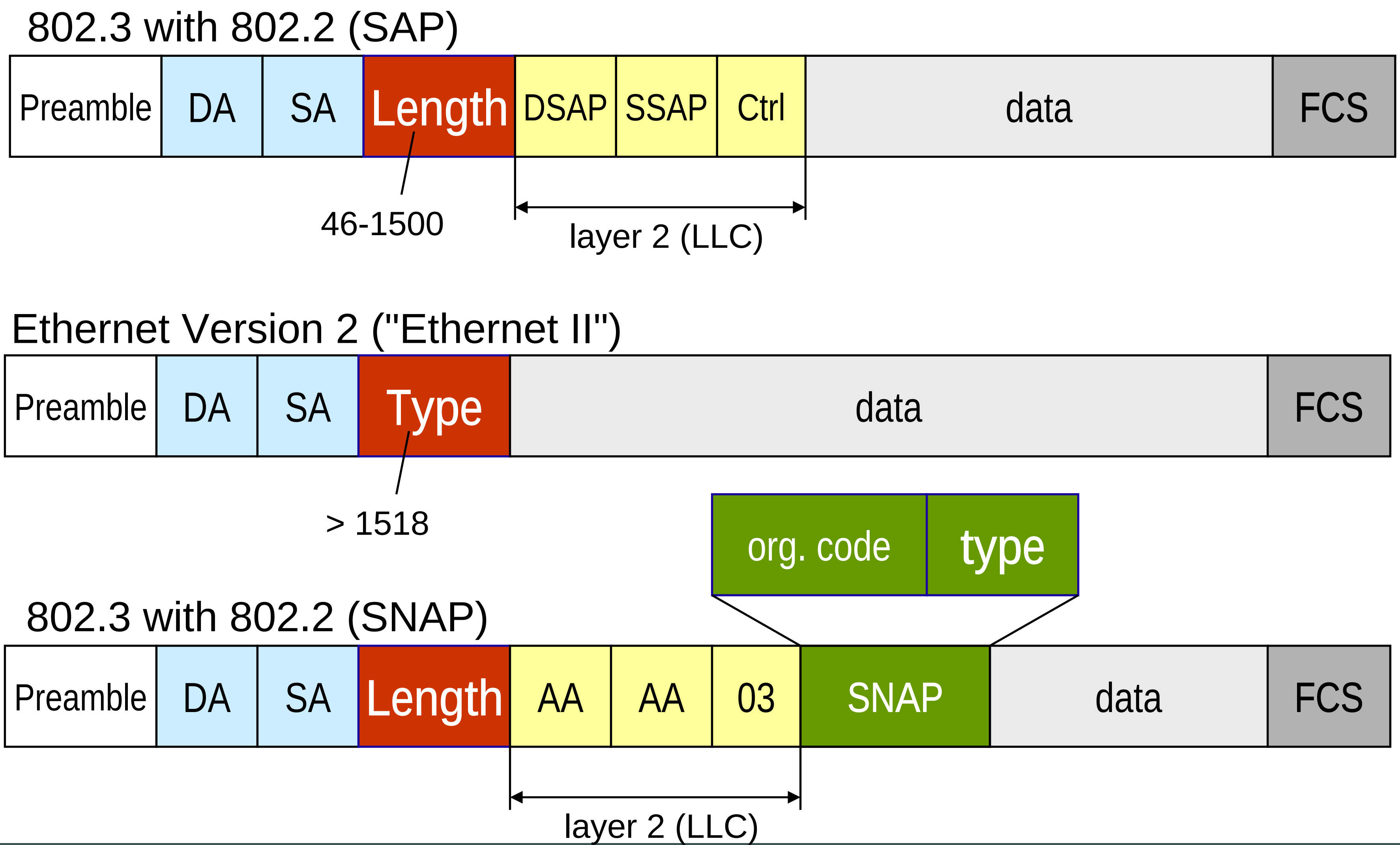


"THE" Ethernet Frame

SNAP

- **Demand for carrying type-field in 802.4, 802.5, 802.6, ... also !**
- **Subnetwork Access Protocol (SNAP) header introduced**
  - **If DSAP=SSAP=0xAA and Ctrl=0x03 then a 5 byte SNAP header follows**
  - **Containing 3 bytes organizational code plus 2 byte DIX type field**

Frame Types Summary



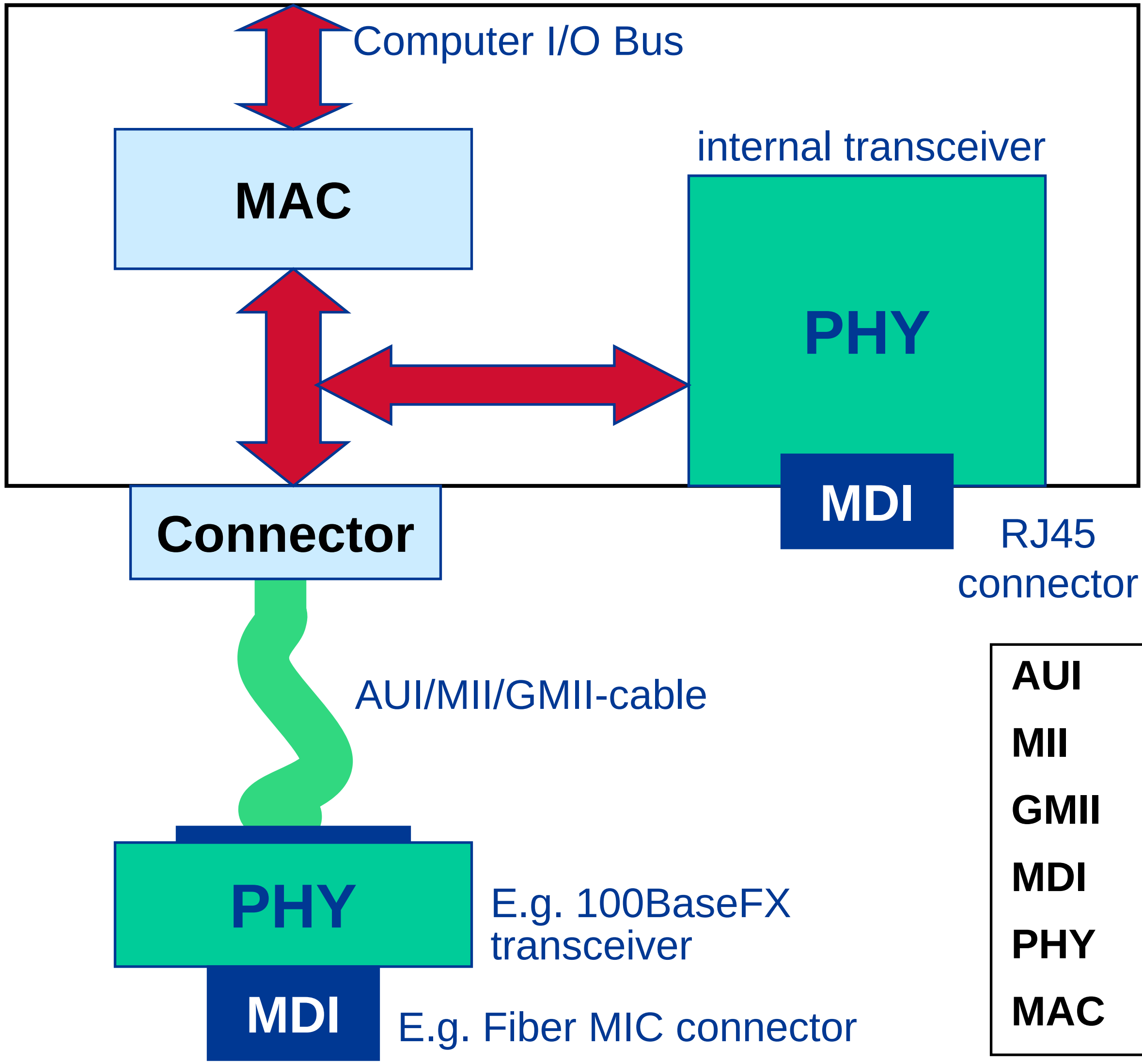
PHY Variants

- 10Base2 (10 Mbit/s, 200 meters)
- 10Base5 (500 meters)
- 10BaseT (star-like cabling, hub needed)
- 10BaseF (fiber)
- 10Broad36 (broadband cable)
- 100BaseT
- 1000BaseT
- 1000BaseX

Twisted Pair Cabling

- Category X cables
  - Cat 3 (Voice grade)
  - Cat 4
  - Cat 5
  - Cat 5e (1000BaseT, unshielded)
  - Cat 6
  - Cat 7
- Category depends on twisting cycles per length unit, isolation, and shielding

Typical NIC Design



AUI	**Attachment Unit Interface**
MII	**Media Independent Interface**
GMII	**Gigabit MII**
MDI	**Medium Dependent Interface**
PHY	**Physical Layer Device**
MAC	**Media Access Control Unit**

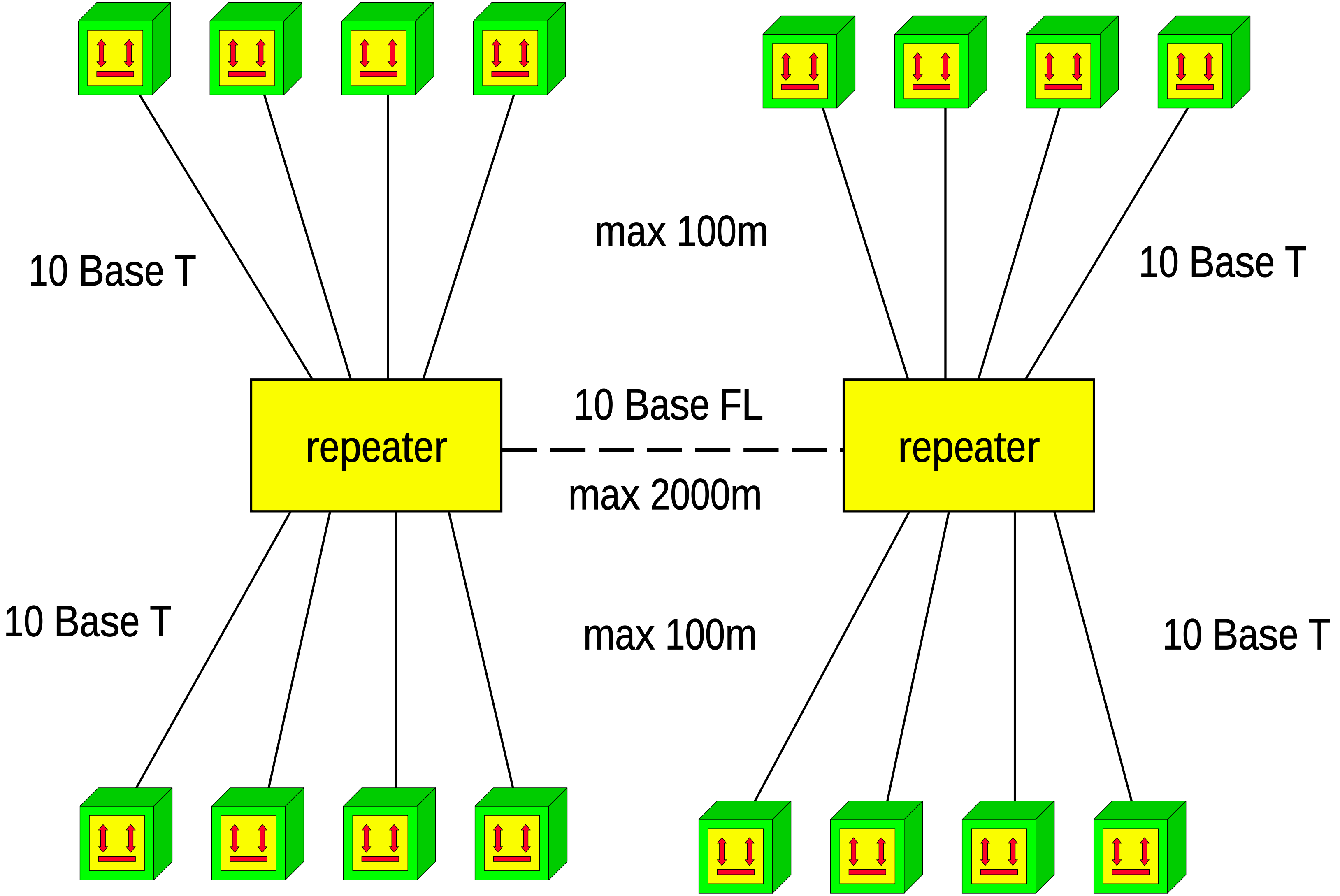
Summary

- Successful because simple
- Two frames: DIX (Ethernet2) and IEEE (802.3)
- Shared medium has consequences
  - Collisions → Slot time → Network diameter
  - Unpredictable, bad for realtime
- Increased data rate until today → 10 GE already available (!)

Quiz

- <u>What is a hub?</u>
  <u>List typical properties</u>:
  - Half/full-duplex?
  - Different data rates?
  - Collision behavior?
- What is the canonical addressing format?
- What is a jam signal?
- What is 802.3u and 803.3z ?
- What is a runt? What is the opposite?

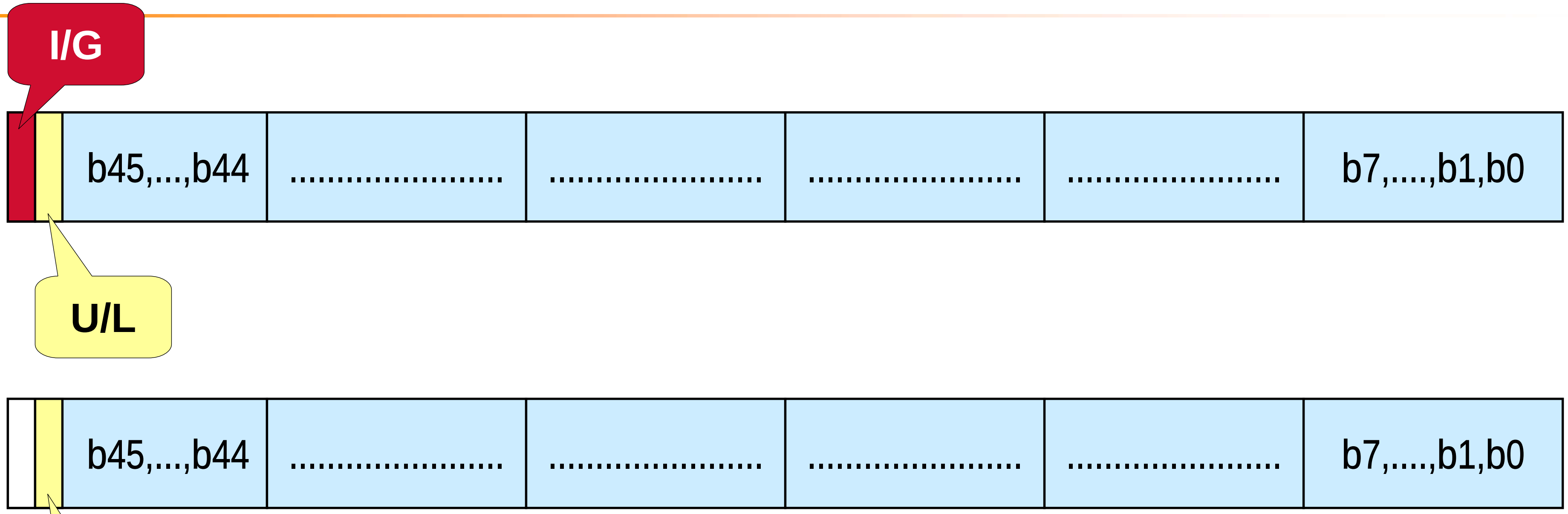
Multiport Repeater as „Hub“



Quiz

- What is a hub?
  List typical properties:
  - Half/full-duplex?
  - Different data rates?
  - Collision behavior?
- What is the canonical addressing format?
- What is a jam signal?
- What is 802.3u and 803.3z ?
- What is a runt? What is the opposite?

6 Byte MAC Addresses



- **Individual/Group (I/G)**
  - **I/G=0 is a unicast address**
  - **I/G=1 is a group (broadcast) address**
- **Universal/Local (U/L)**
  - **U/L=0 is a global, IEEE administered address**
  - **U/L=1 is a local administered address**

Storage Format of 802.3 MAC-Address

- **basic rule:**
  - I/G bit must be the first bit on the medium, so the transmitted address must have the following format:

I/G	U/L	b45, ..., b40	b39, ..., b32	...	b15, ..., b8	b7, ..., b1, b0
0	1	00 1000	0000 0000	....	0000 0000	1100 0001

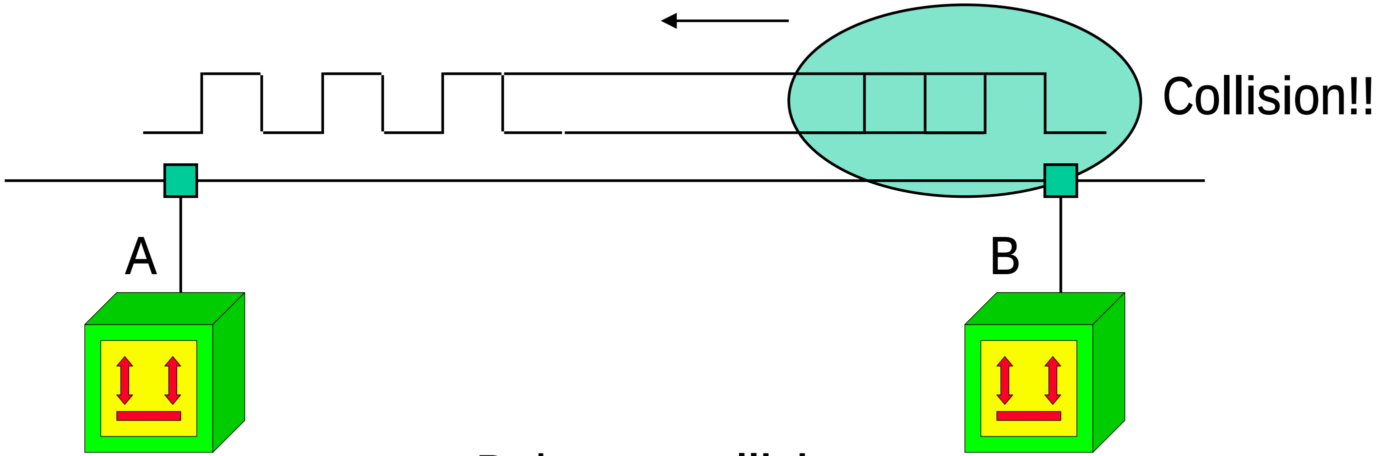
  - <u>802.3 sends the least significant bit of each byte at first</u>
  - so 802.3 must store each byte in memory in reverse order:
    - also called “Canonical” Format

b40, ..., b45	U/L	I/G	b32, ..., b39	...	b8, ..., b15	b0, b1, ..., b7
0001 00	1	0	0000 0000	....	0000 0000	1000 0011

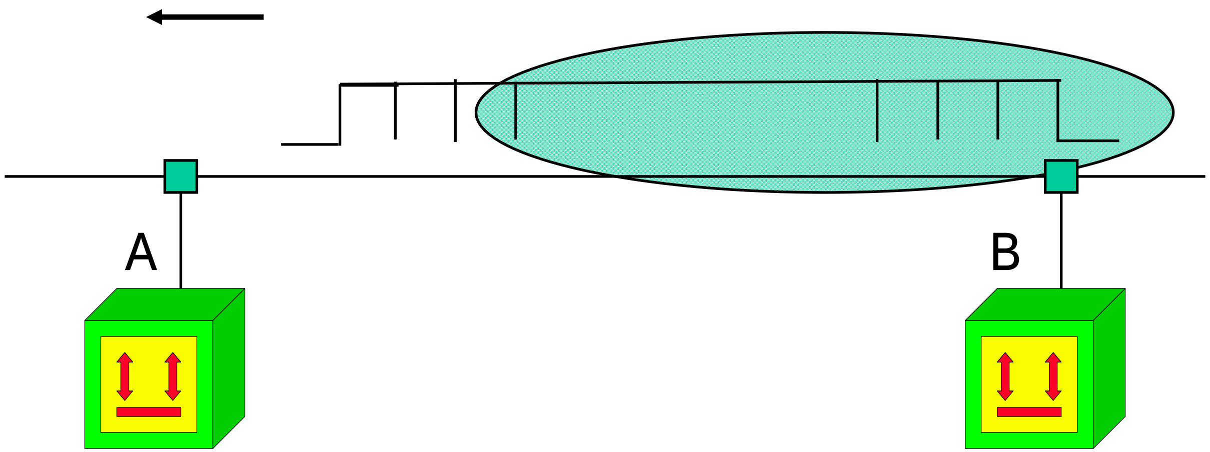
Quiz

- What is a hub?
  List typical properties:
  - Half/full-duplex?
  - Different data rates?
  - Collision behavior?
- What is the canonical addressing format?
- <u>What is a jam signal?</u>
- What is 802.3u and 803.3z ?
- What is a runt? What is the opposite?

Small Collision extended by JAM



t = x+ tp ... B detects collision
tp ... signal propagation time from A to B



t = x + 2tp ... A detects collision

Quiz

- **What is a hub? List typical properties:**
  - **Half/full-duplex?**
  - **Different data rates?**
  - **Collision behavior?**
- **What is the canonical addressing format?**
- **What is a jam signal?**
- **What is 802.3u and 803.3z ?**
- **What is a runt? What is the opposite?**

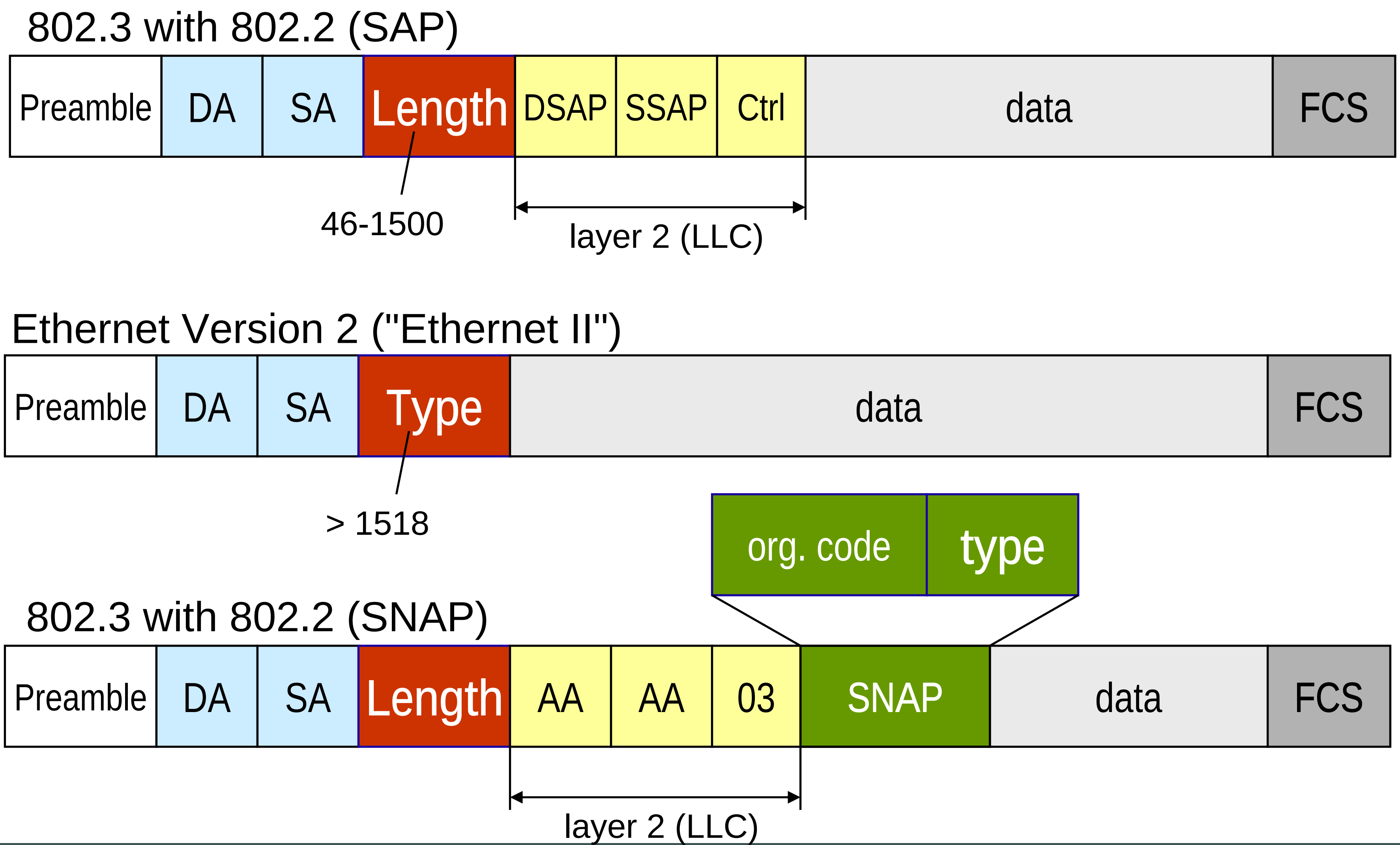
PHY Variants

- **10Base2 (10 Mbit/s, 200 meters)**
- **10Base5 (500 meters)**
- **10BaseT (star-like cabling, hub needed)**
- **10BaseF (fiber)**
- **10Broad36 (broadband cable)**
- **100BaseT**
  - **100BaseTX = 802.3u (integrated in 802.3-2008)**
- **1000BaseT**
  - **1000BaseT = 802.3ab (integrated in 802.3-2008)**
- **1000BaseX**
  - **1000BaseX = 802.3z (integrated in 802.3-2008)**

Quiz

- What is a hub?
  List typical properties:
  - Half/full-duplex?
  - Different data rates?
  - Collision behavior?
- What is the canonical addressing format?
- What is a jam signal?
- What is 802.3u and 803.3z ?
- <u>What is a runt? What is the opposite</u>?

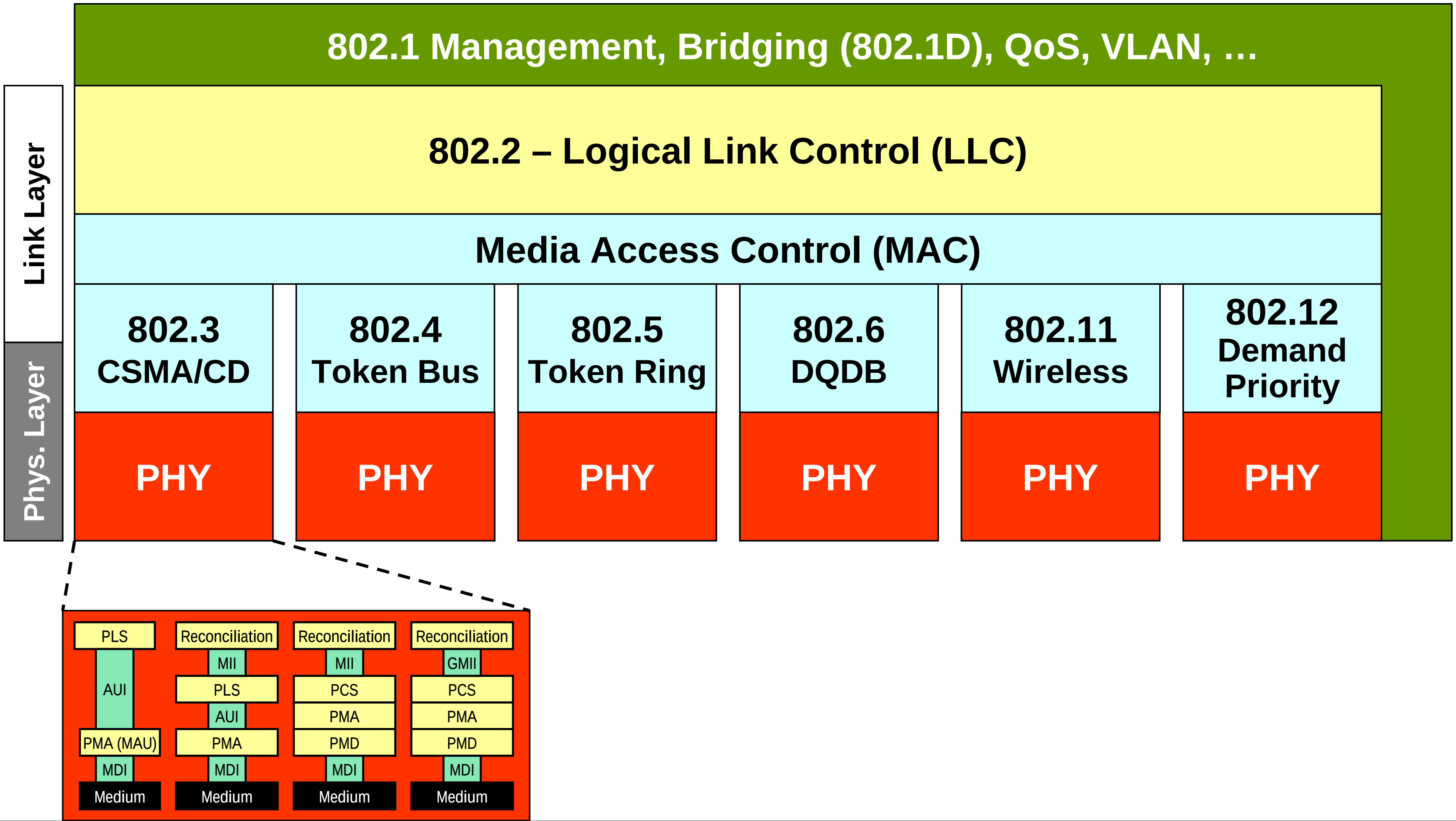
Frame Types Summary



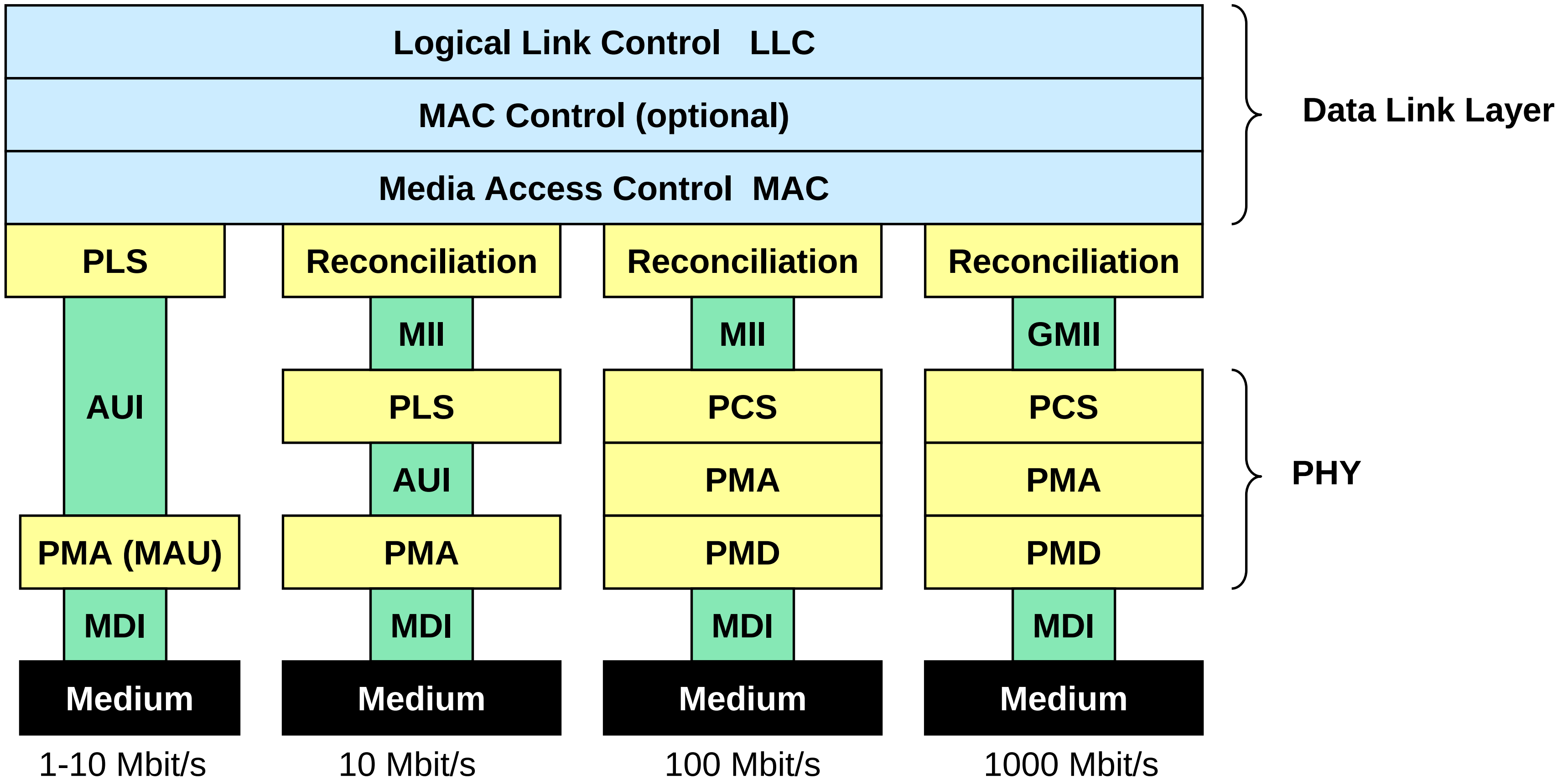
Quiz

- Explain NIC Design:
  - PHY, AUI, MII, GMII, MDI
    - AUI -> serial cable, 50m max. between NIC and 10Mps Ethernet transceiver
    - MII -> parallel interface (4-bit) between MAC controller and 100Mbps Ethernet transceiver
    - GMII -> parallel interface (8-bit) between MAC controller and 1000Mbps Ethernet transceiver

IEEE 802 Layer Model

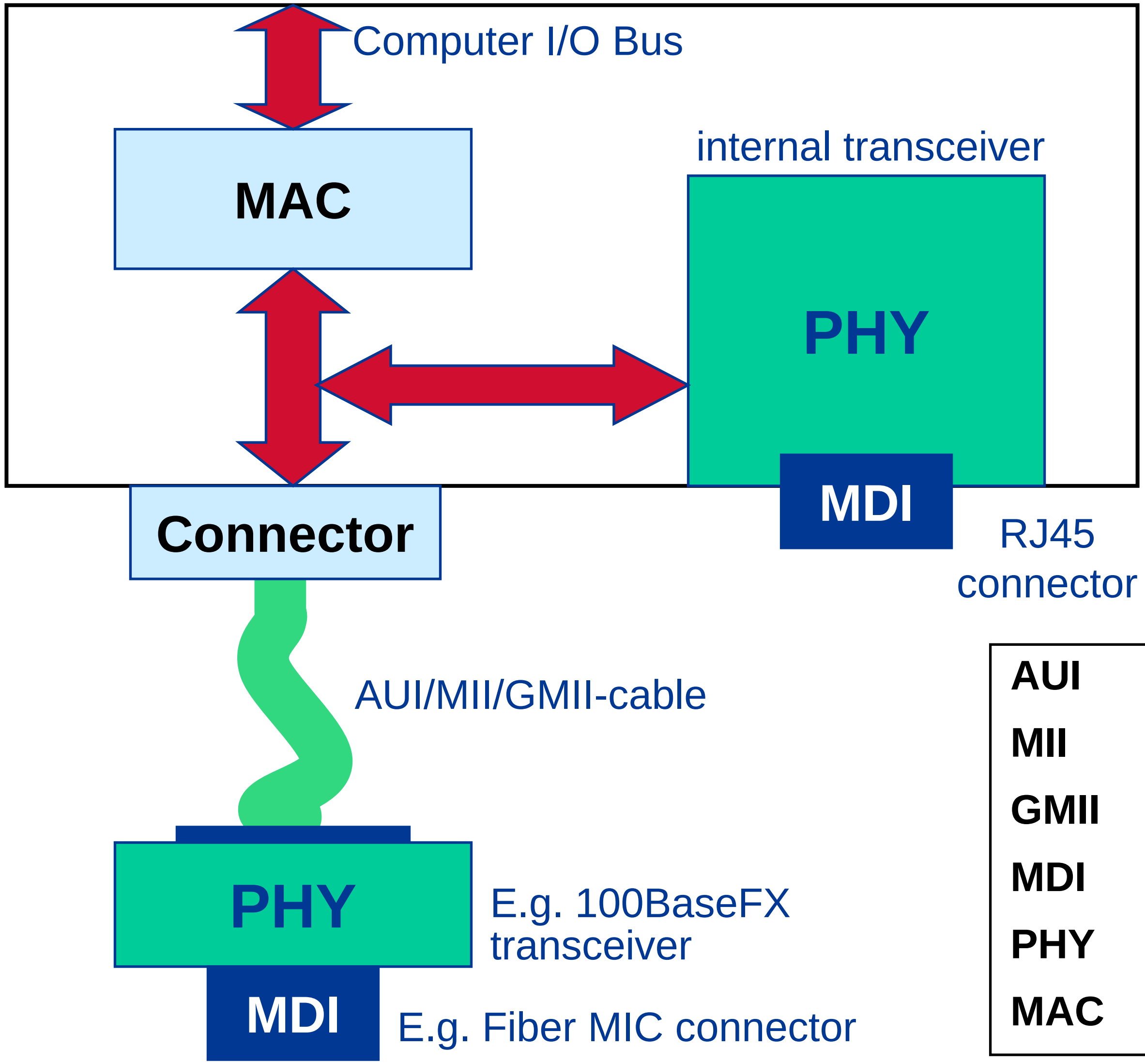


Several Physical Media Supported



AUI Attachment Unit Interface, **PLS** Physical Layer Signaling, **MDI** Medium Dependent Interface **PCS** Physical Coding Sublayer, **MII** Media Independent Interface, **GMII** Gigabit Media Independent Interface, **PMA** Physical Medium Attachment, **MAU** Medium Attachment Unit, **PMD** Physical Medium Dependent

Typical NIC Design



AUI	**Attachment Unit Interface**
MII	**Media Independent Interface**
GMII	**Gigabit MII**
MDI	**Medium Dependent Interface**
PHY	**Physical Layer Device**
MAC	**Media Access Control Unit**