# The Ethernet Evolution

From 10Mbit/s to 10Gigabit/s Ethernet Technology

From Bridging to L2 Ethernet Switching and VLANs

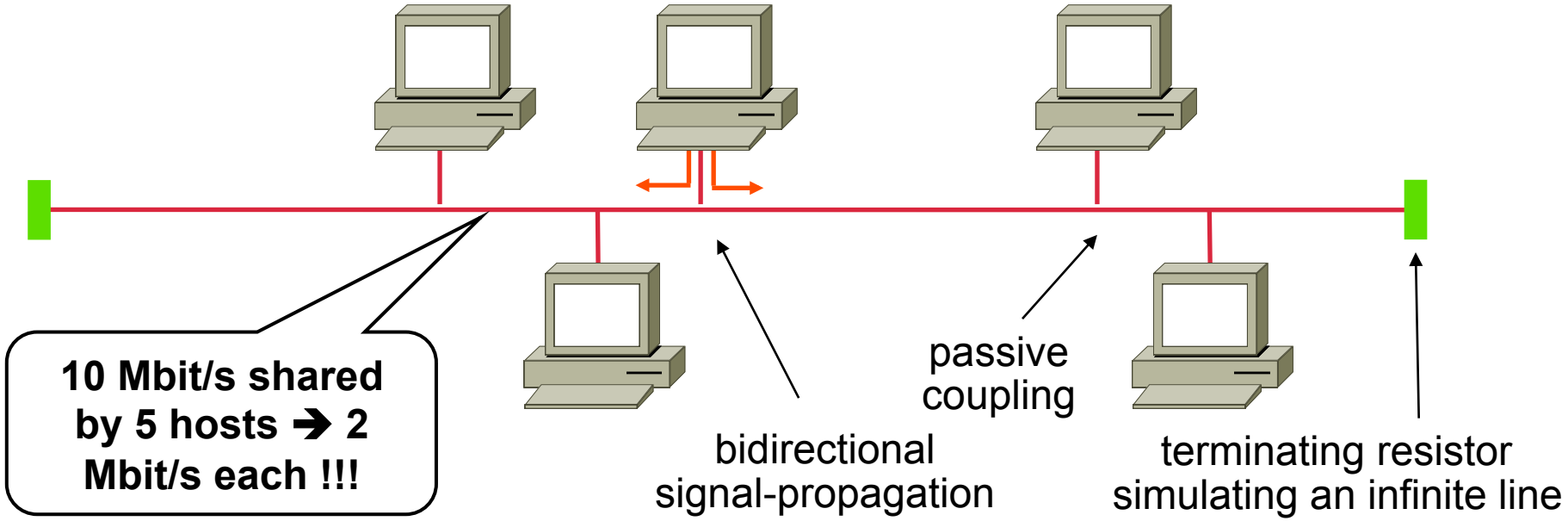
From LAN to WAN Transmission Technique

# Agenda

- **Ethernet Evolution**
- **VLAN**
- **High Speed Ethernet**
  - Introduction
  - Fast Ethernet
  - Gigabit Ethernet
  - 10 Gigabit Ethernet

# History: Initial Idea

- Bus topology based on coax cable (two wires)
  - ➔ half duplex transmission, natural broadcast behavior
- No active network elements (no store and forward) ➔ low latency
- Shared media needs media access control (MAC)
  - ➔ CSMA/CD (Carrier Sense Multiple Access / Collision Detection)
  - ➔ MAC addresses (OSI L2 address)
- One single collision domain and one broadcast domain



*"Use common sense in routing cable. Avoid wrapping coax around sources of strong electric or magnetic fields. Do not wrap the cable around fluorescent light ballasts or cyclotrons, for example."*

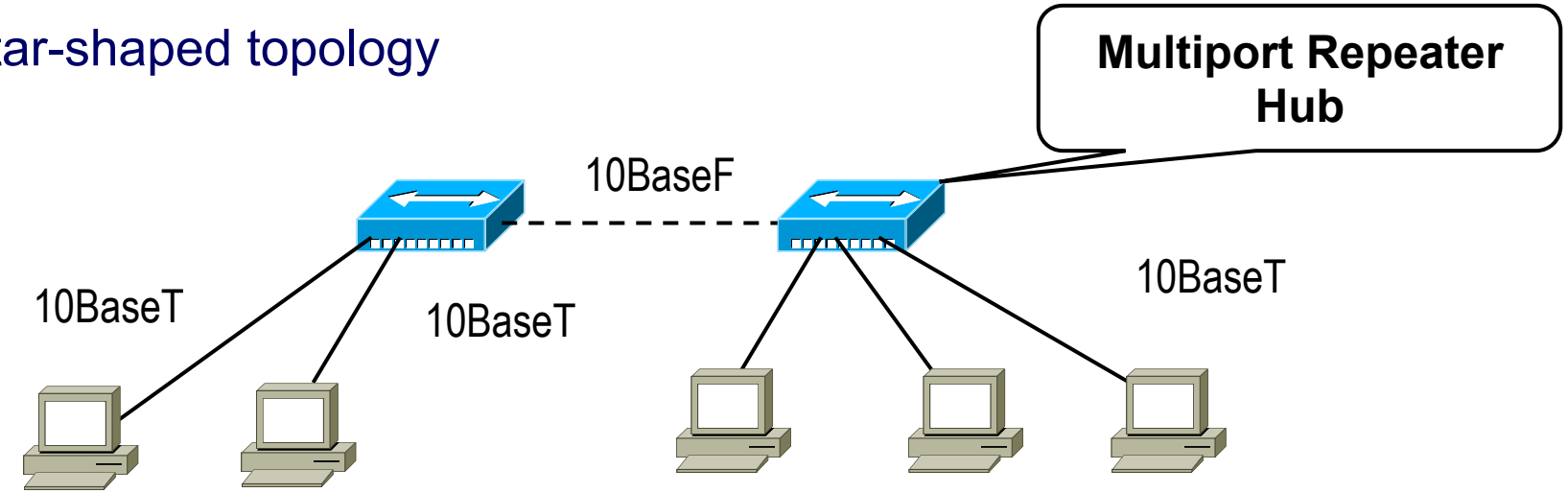
**Ethernet Headstart Product Information and Installation Guide, Bell Technologies, pg. 11**

# History: Multiport Repeaters

- Demand for structured cabling
  - 10BaseT (Cat3, Cat4, ...)
  - (voice-grade twisted-pair)
- Multiport repeater ("Hub") created
  - "CSMA/CD in a box“
  - The 180 Degree Turn (“Star instead Bus”)
- Still half duplex communication
- Still one collision domain and broadcast domain
- Star-shaped topology

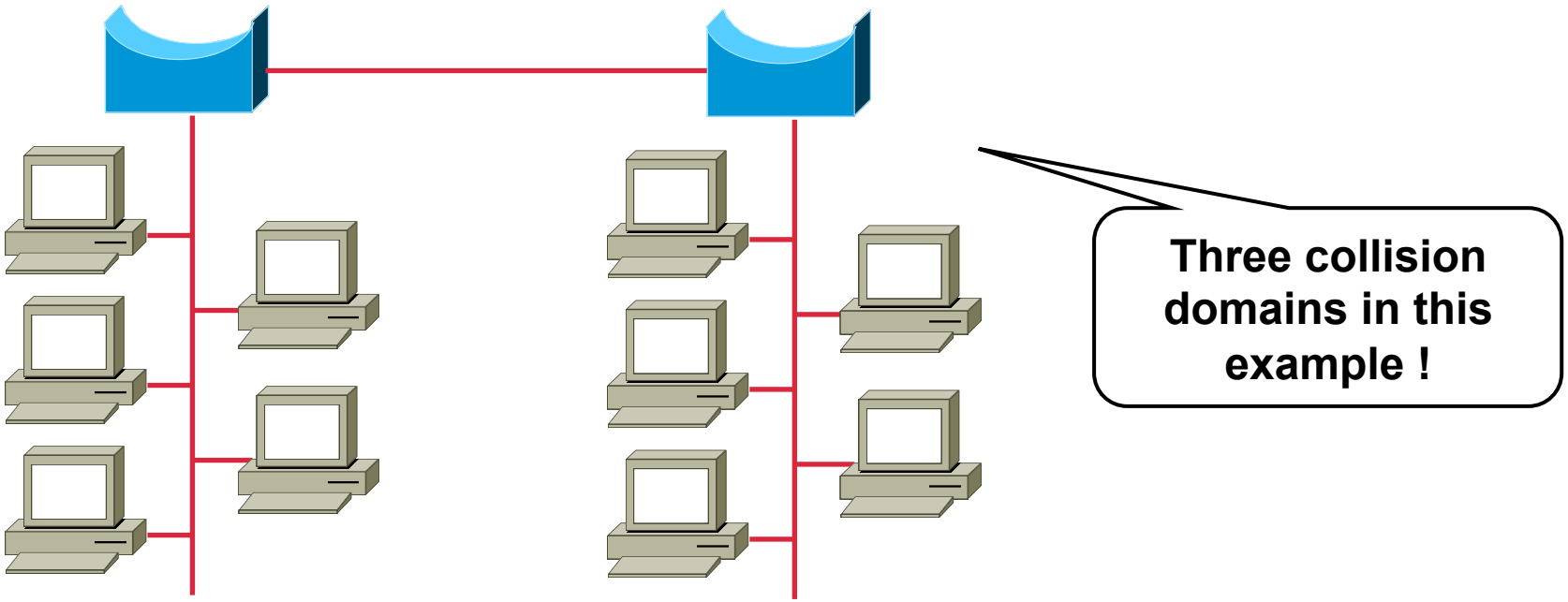
——— represents four CU wires
2 for Tmt, 2 for Rcv
(e.g. 10BaseT)

- - - - represents two FO wires
e.g. 10BaseF



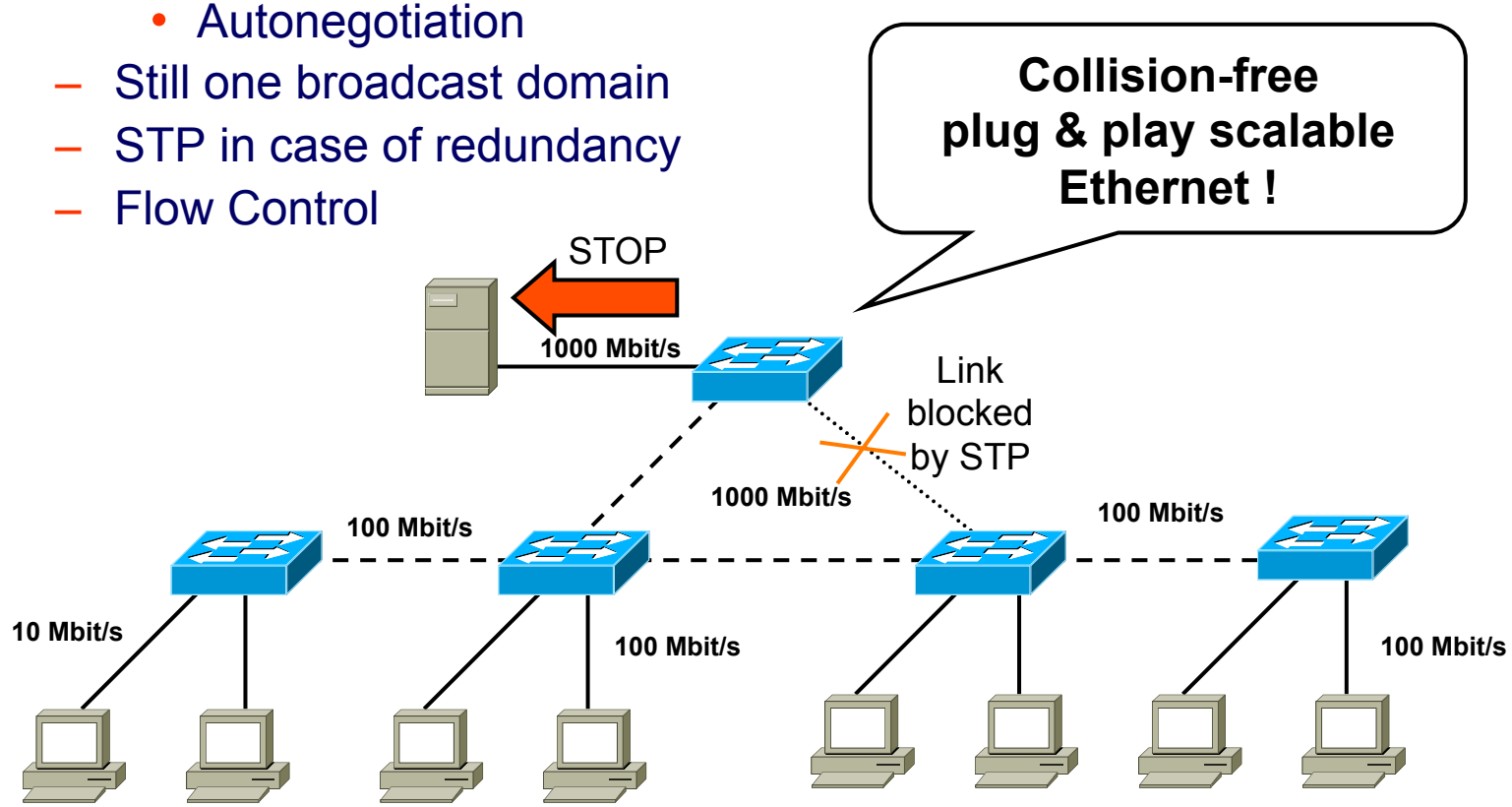
# History: Bridges

- Store and forwarding according destination MAC address
- Separated collision domains
- Improved network performance
- Still one broadcast domain



# History: Switches

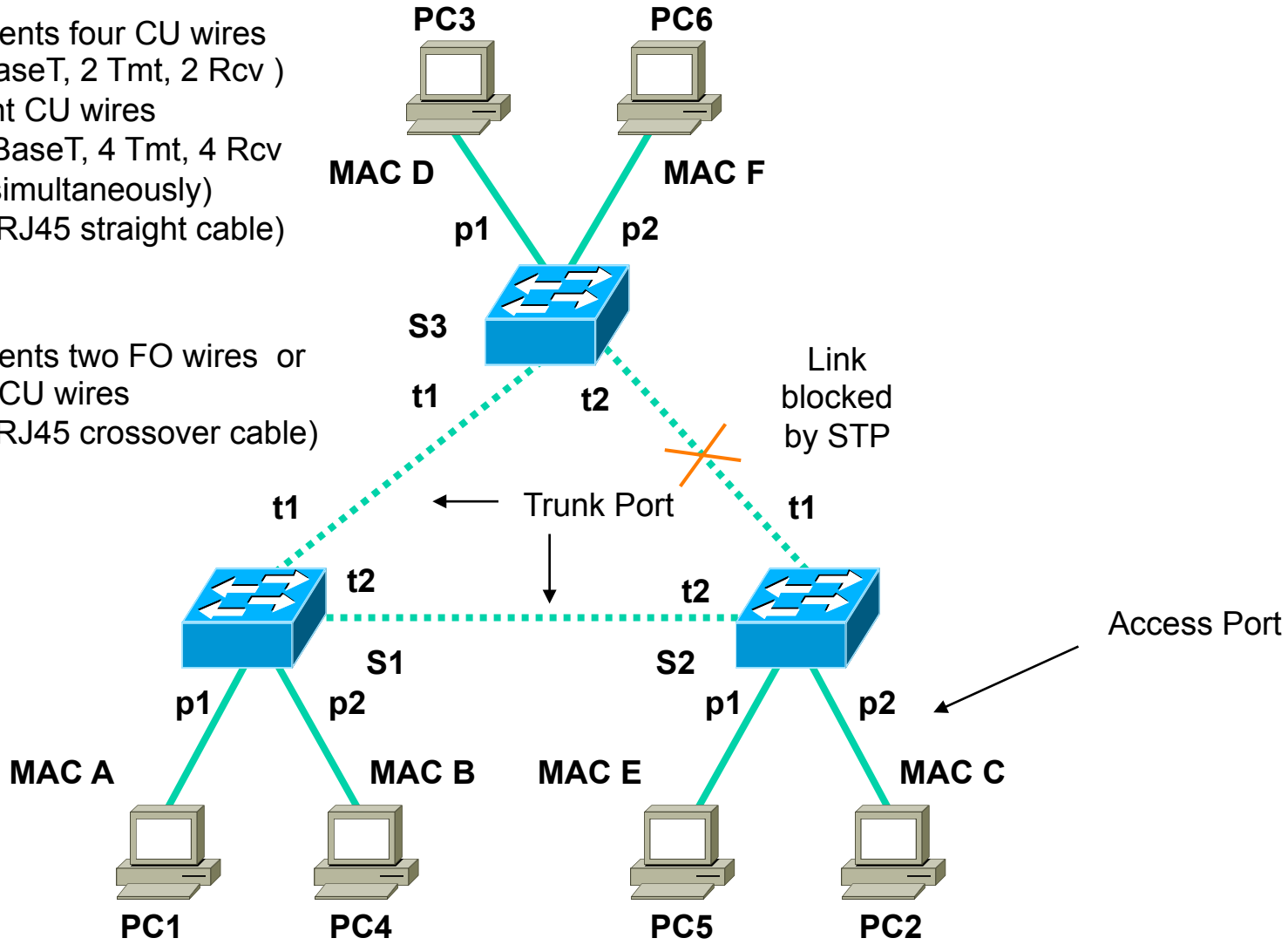
- Switch = Multiport Bridges with HW acceleration
- Full duplex ➔ Collision-free Ethernet ➔ No CSMA/CD necessary anymore
  - No collision domains anymore
- Different data rates at the same time supported
  - Autonegotiation
- Still one broadcast domain
- STP in case of redundancy
- Flow Control



# Ethernet Switching Topology
## (MAC Address Table - Empty)

represents four CU wires
(100BaseT, 2 Tmt, 2 Rcv )
or eight CU wires
(1000BaseT, 4 Tmt, 4 Rcv
used simultaneously)
(Rj45-RJ45 straight cable)

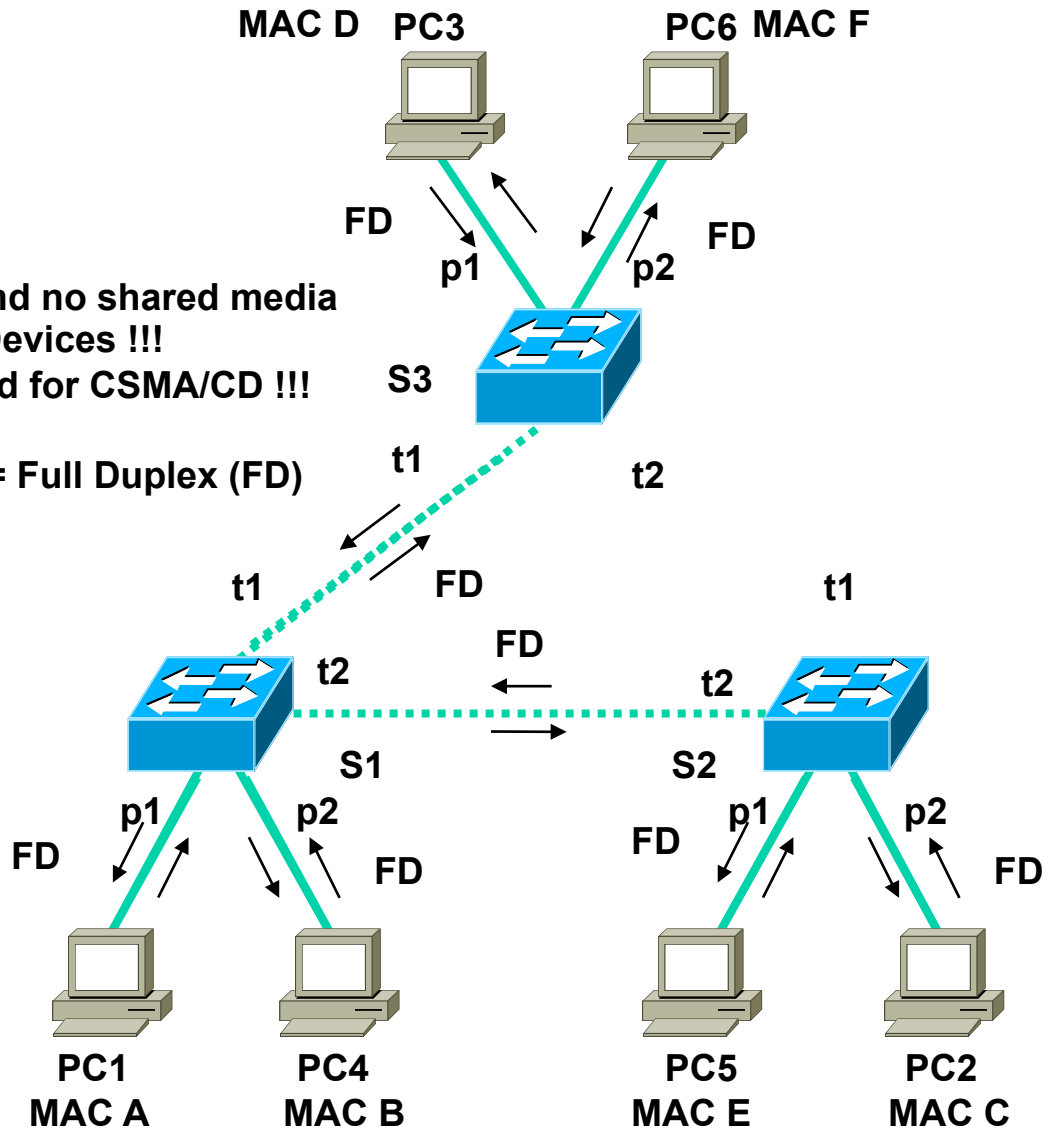
represents two FO wires or
4 or 8 CU wires
(Rj45-RJ45 crossover cable)



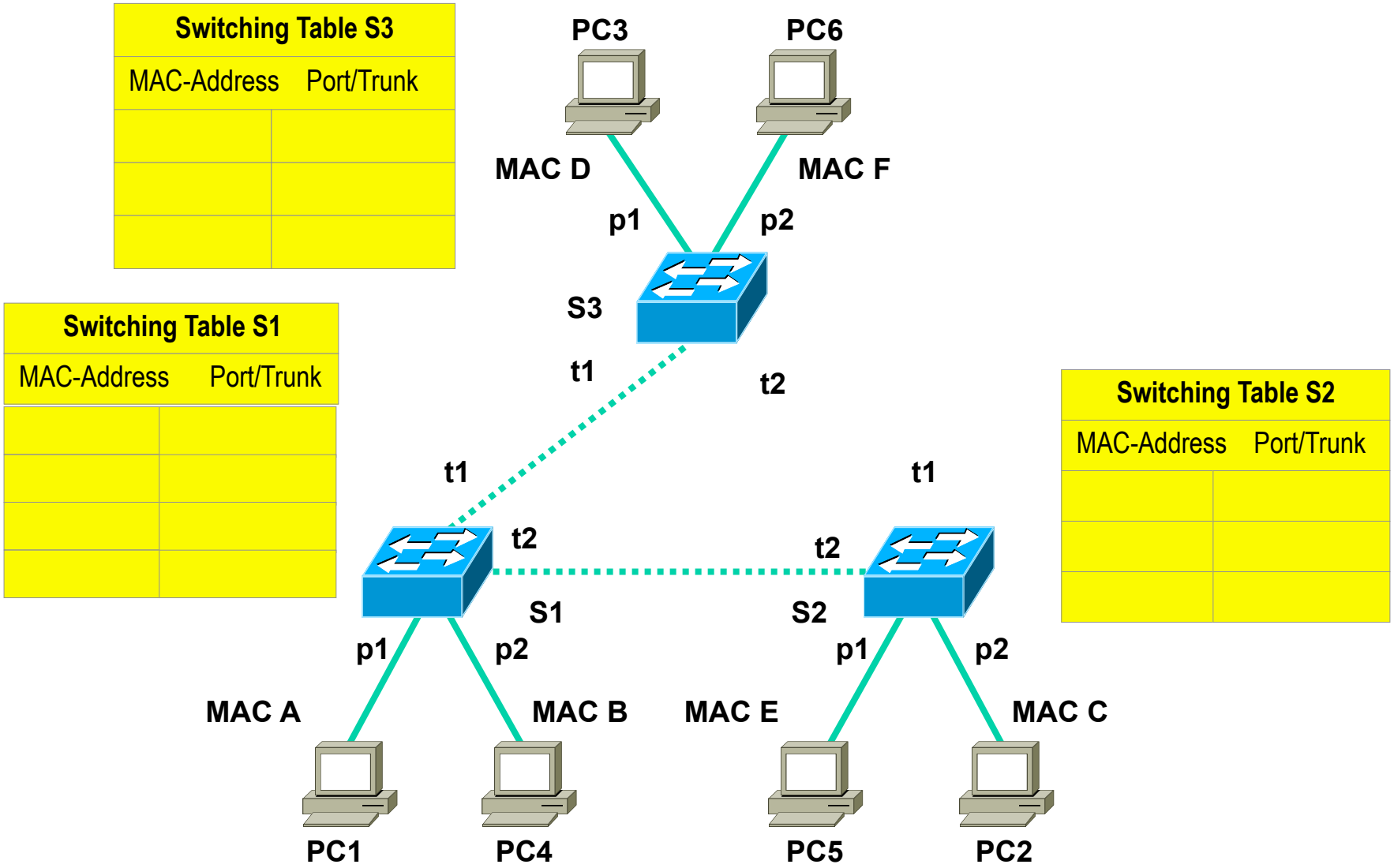
# Ethernet Switching – Full Duplex (FD)
## (Point-to-Point Links and FD Everywhere)

Only PTP links and no shared media for more than 2 Devices !!!
Therefore no need for CSMA/CD !!!

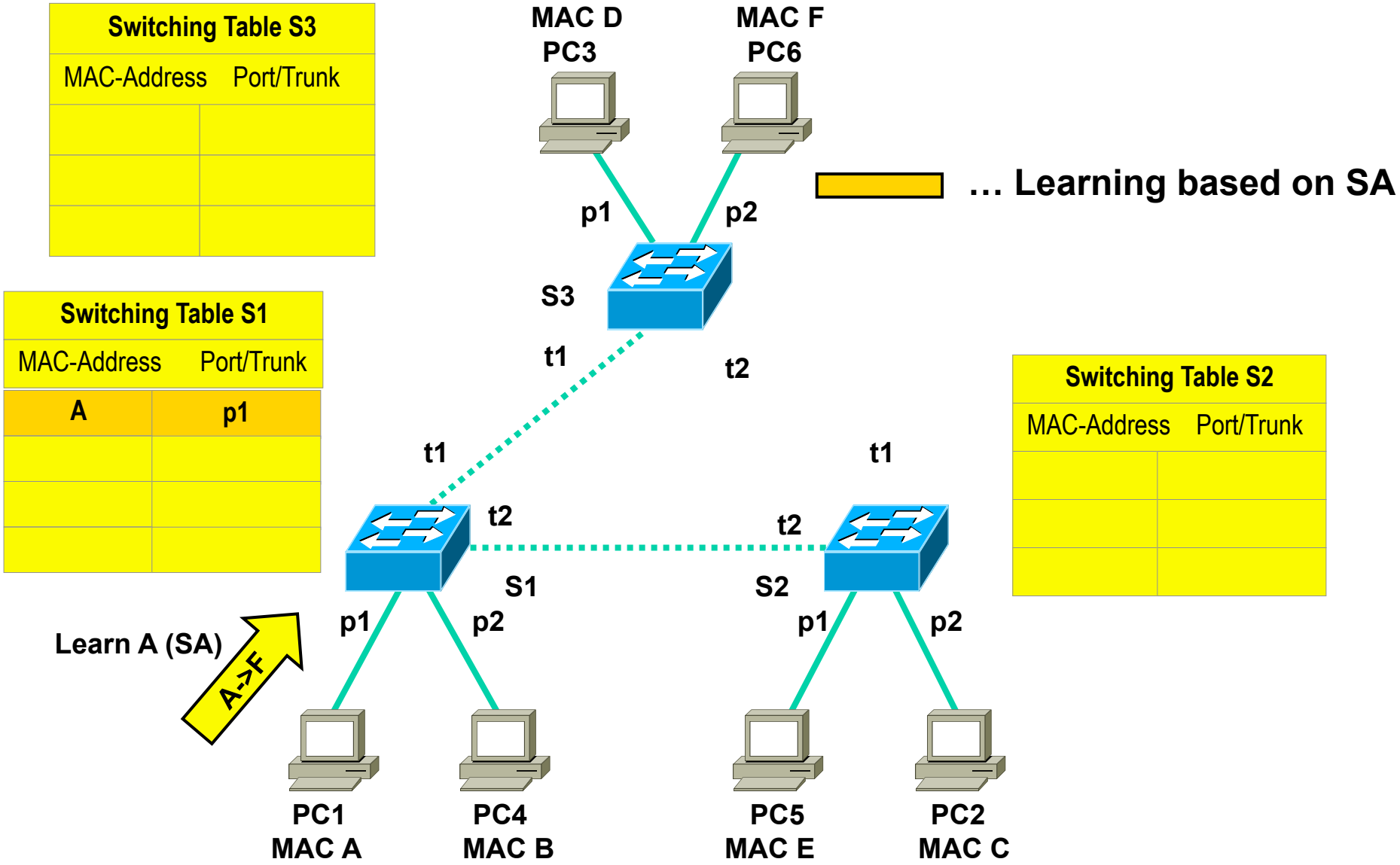
CSMA/CD OFF == Full Duplex (FD)



# Ethernet Switch Table - Power On
## (MAC Address Table - Empty)



# Table Building for Ethernet Frame MAC-A to MAC-F 1

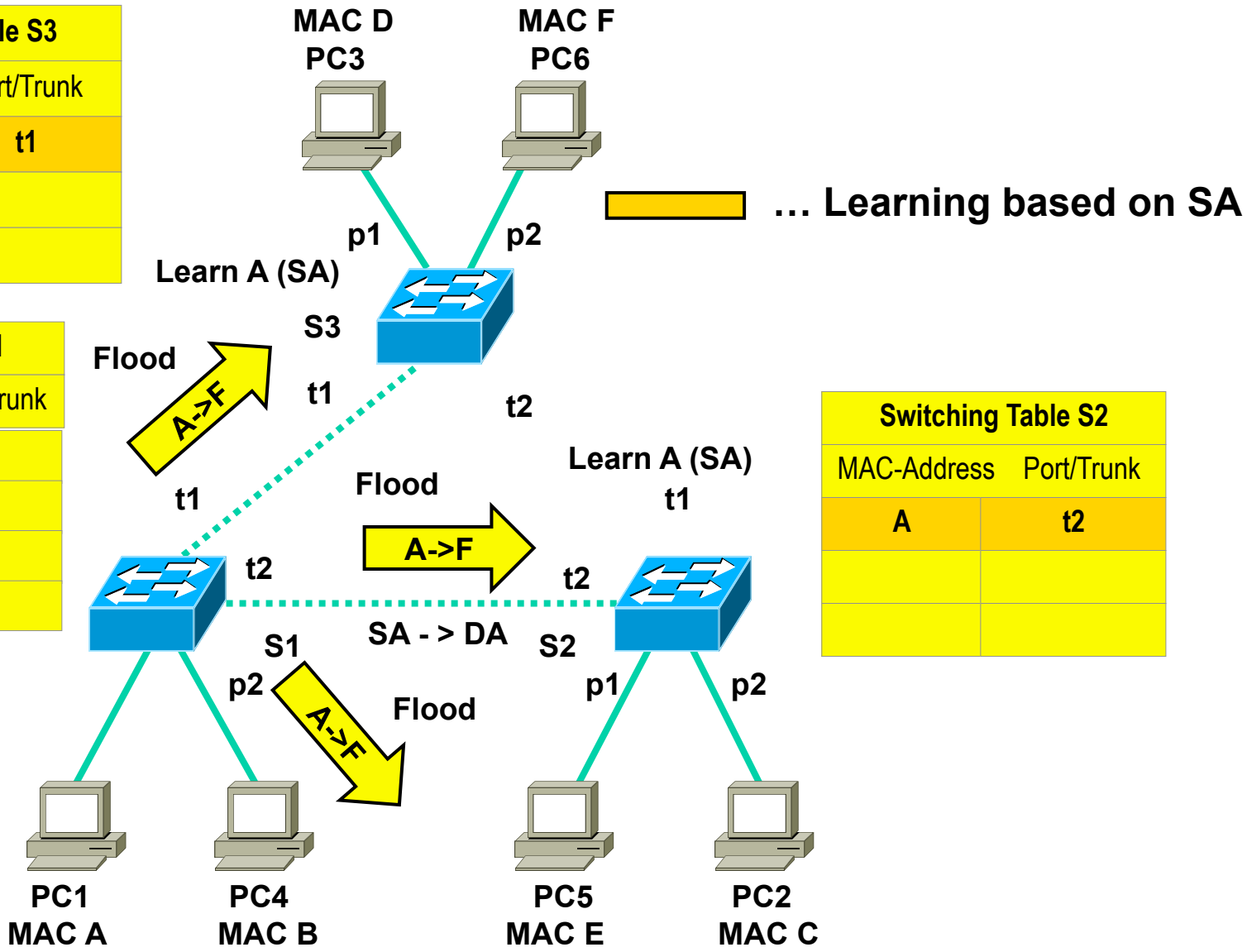


# Table Building for Ethernet Frame MAC-A to MAC-F 2

| Switching Table S3 | |
|---|---|
| MAC-Address | Port/Trunk |
| A | t1 |
| | |
| | |

| Switching Table S1 | |
|---|---|
| MAC-Address | Port/Trunk |
| A | p1 |
| | |
| | |
| | |

| Switching Table S2 | |
|---|---|
| MAC-Address | Port/Trunk |
| A | t2 |
| | |
| | |

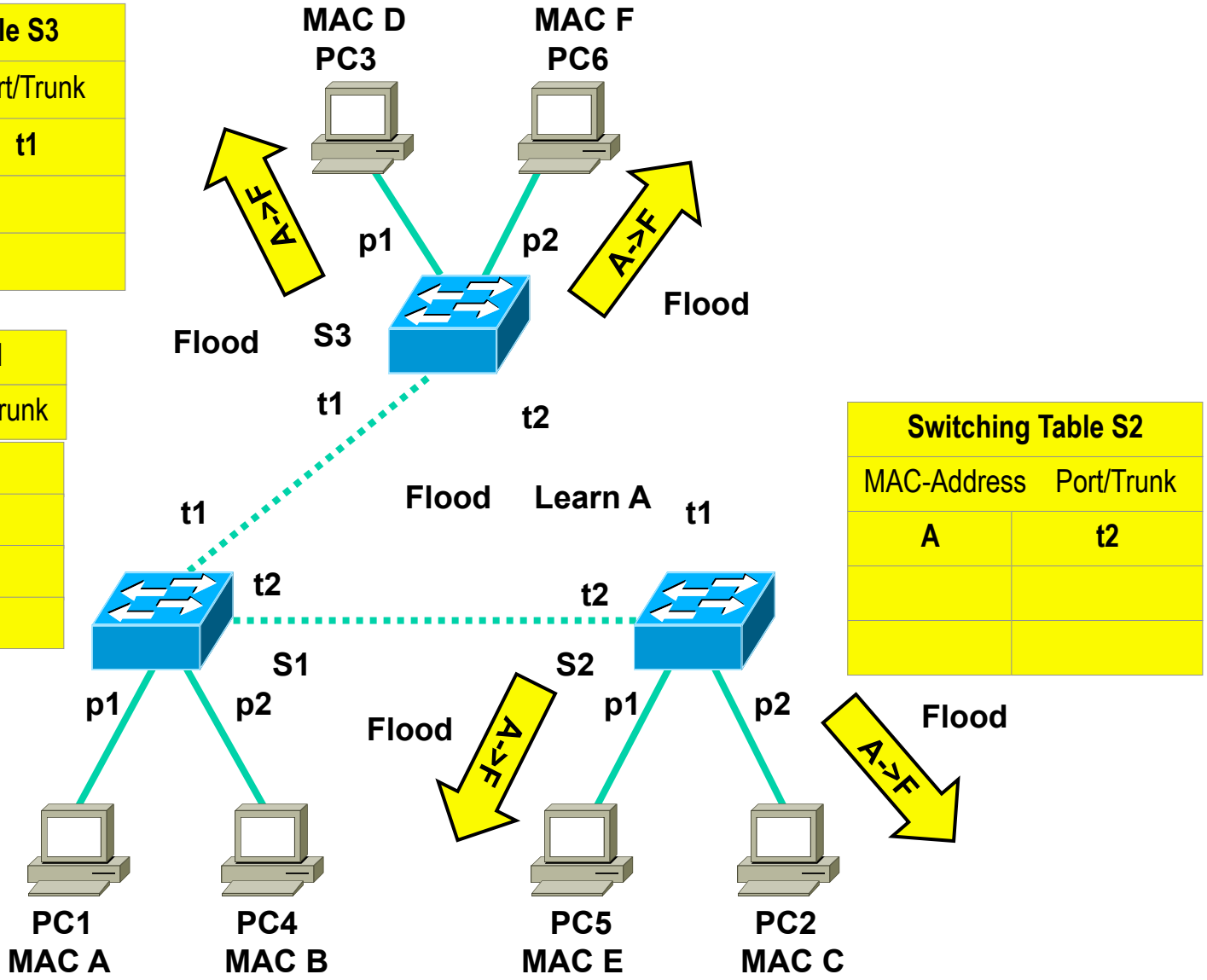


# Table Building for Ethernet Frame MAC-A to MAC-F 3

| Switching Table S3 | |
|---|---|
| MAC-Address | Port/Trunk |
| A | t1 |
| | |
| | |

| Switching Table S1 | |
|---|---|
| MAC-Address | Port/Trunk |
| A | p1 |
| | |
| | |
| | |

| Switching Table S2 | |
|---|---|
| MAC-Address | Port/Trunk |
| A | t2 |
| | |
| | |



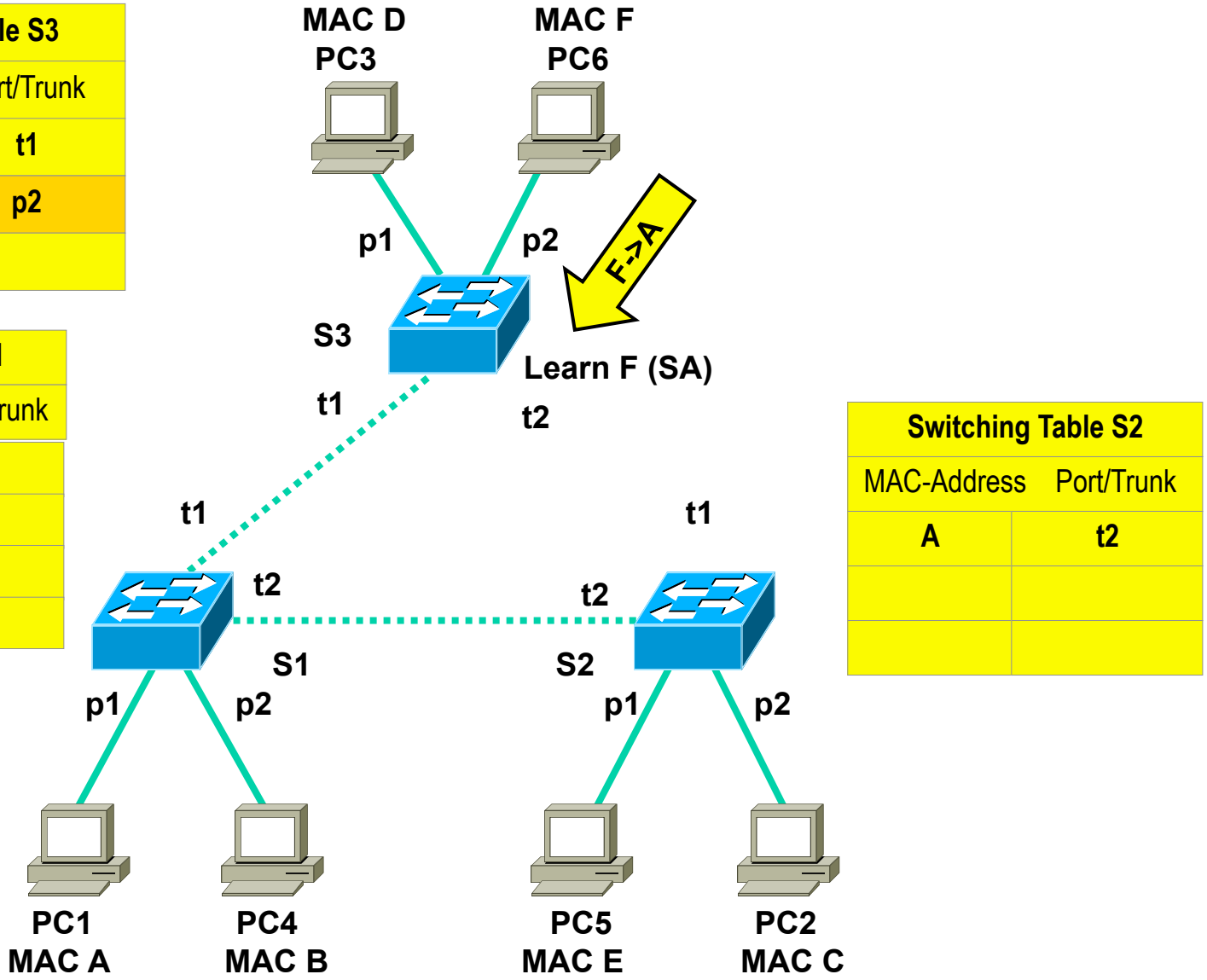
# Table Building / Table Usage for Ethernet Frame MAC-F to MAC-A

1

| Switching Table S3 | |
|---|---|
| MAC-Address | Port/Trunk |
| A | t1 |
| F | p2 |
| | |

| Switching Table S1 | |
|---|---|
| MAC-Address | Port/Trunk |
| A | p1 |
| | |
| | |
| | |

| Switching Table S2 | |
|---|---|
| MAC-Address | Port/Trunk |
| A | t2 |
| | |
| | |

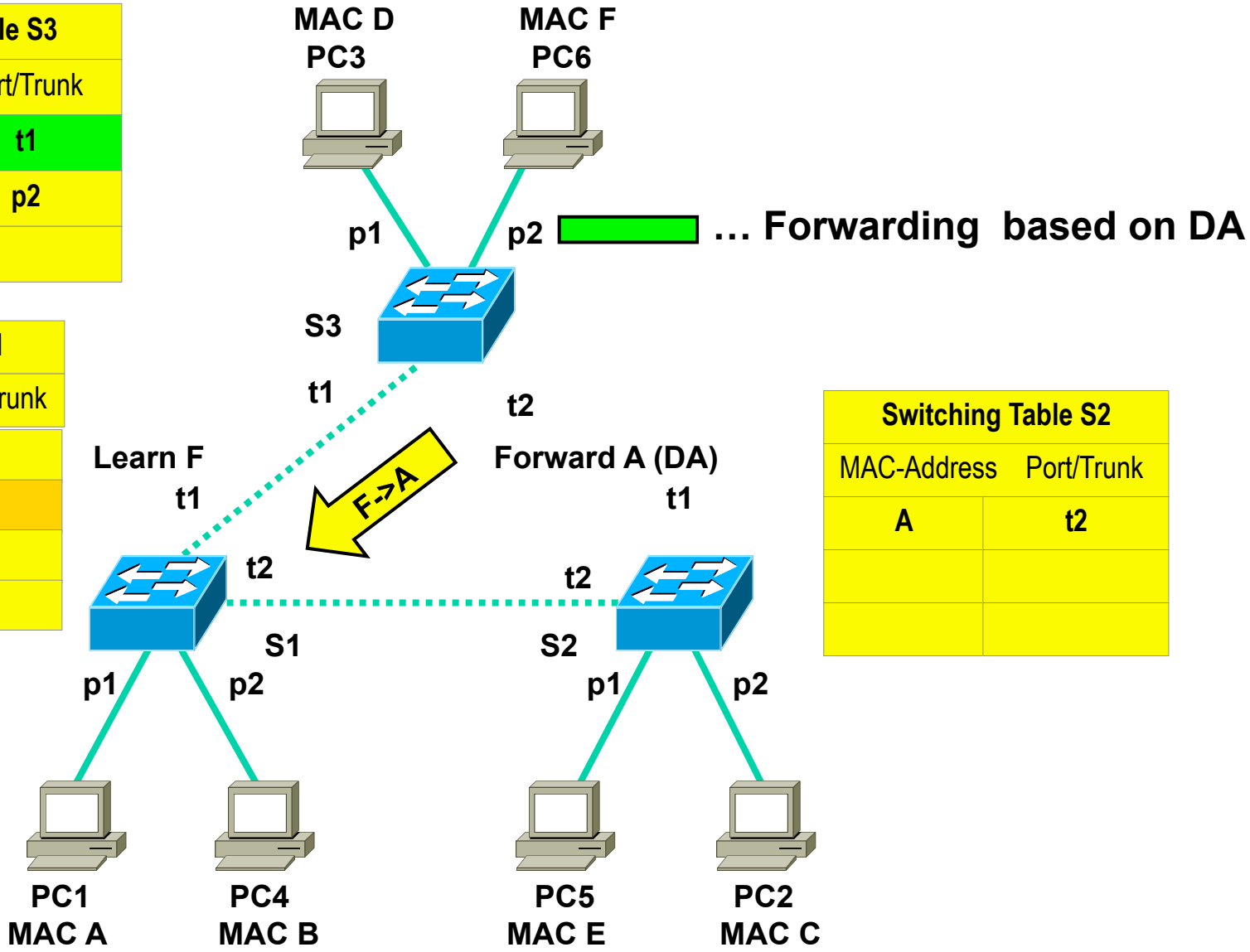


# Table Building / Table Usage (Forwarding Decision) for Ethernet Frame MAC-F to MAC-A 2

| Switching Table S3 | |
|---|---|
| MAC-Address | Port/Trunk |
| A | t1 |
| F | p2 |
| | |

| Switching Table S1 | |
|---|---|
| MAC-Address | Port/Trunk |
| A | p1 |
| F | t1 |
| | |
| | |

| Switching Table S2 | |
|---|---|
| MAC-Address | Port/Trunk |
| A | t2 |
| | |
| | |



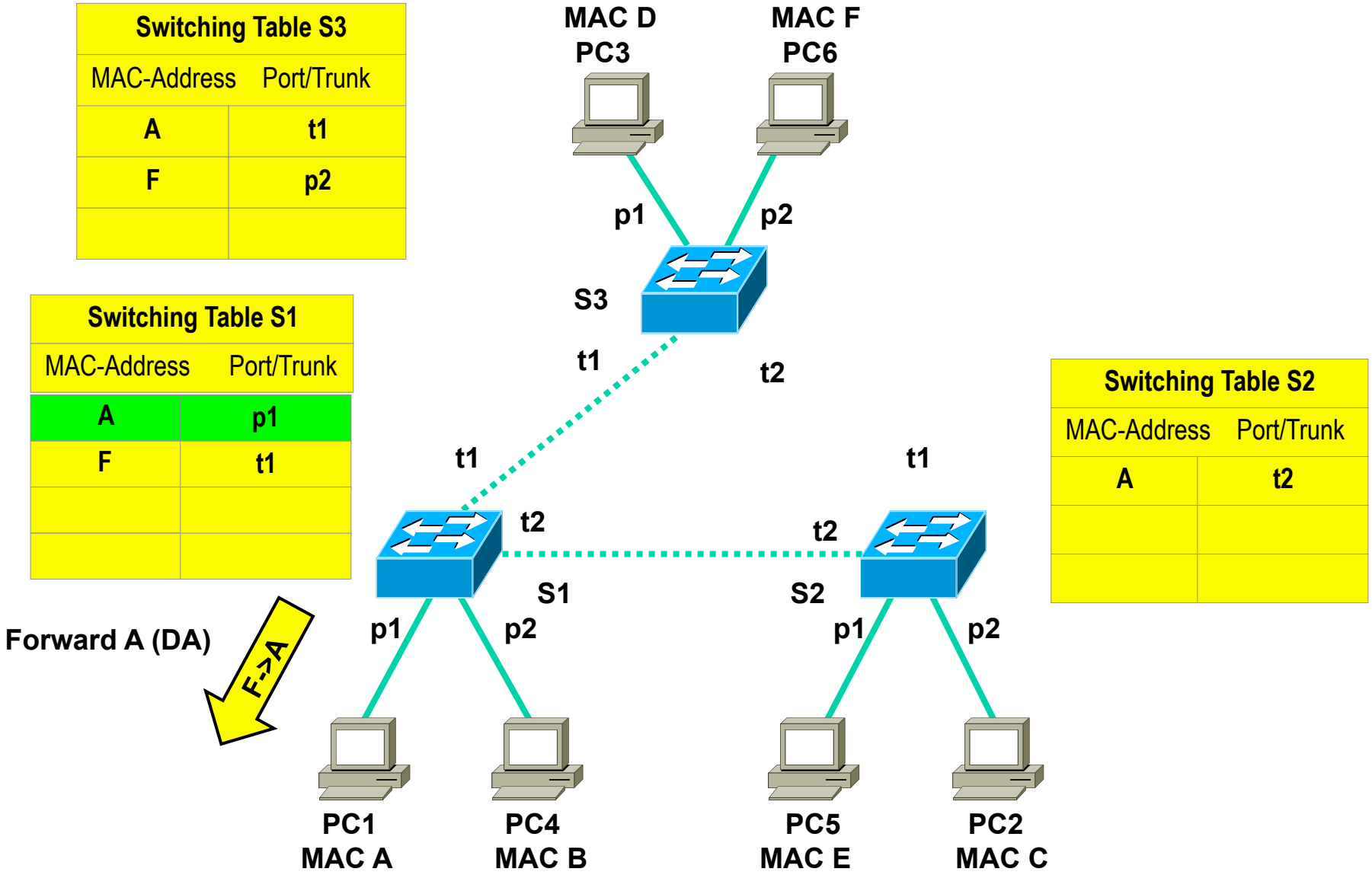
# Table Building / Table Usage (Forwarding Decision) for Ethernet Frame MAC-F to MAC-A



| Switching Table S3 | |
|---|---|
| MAC-Address | Port/Trunk |
| A | t1 |
| F | p2 |
| | |

| Switching Table S1 | |
|---|---|
| MAC-Address | Port/Trunk |
| A | p1 |
| F | t1 |
| | |
| | |

| Switching Table S2 | |
|---|---|
| MAC-Address | Port/Trunk |
| A | t2 |
| | |
| | |

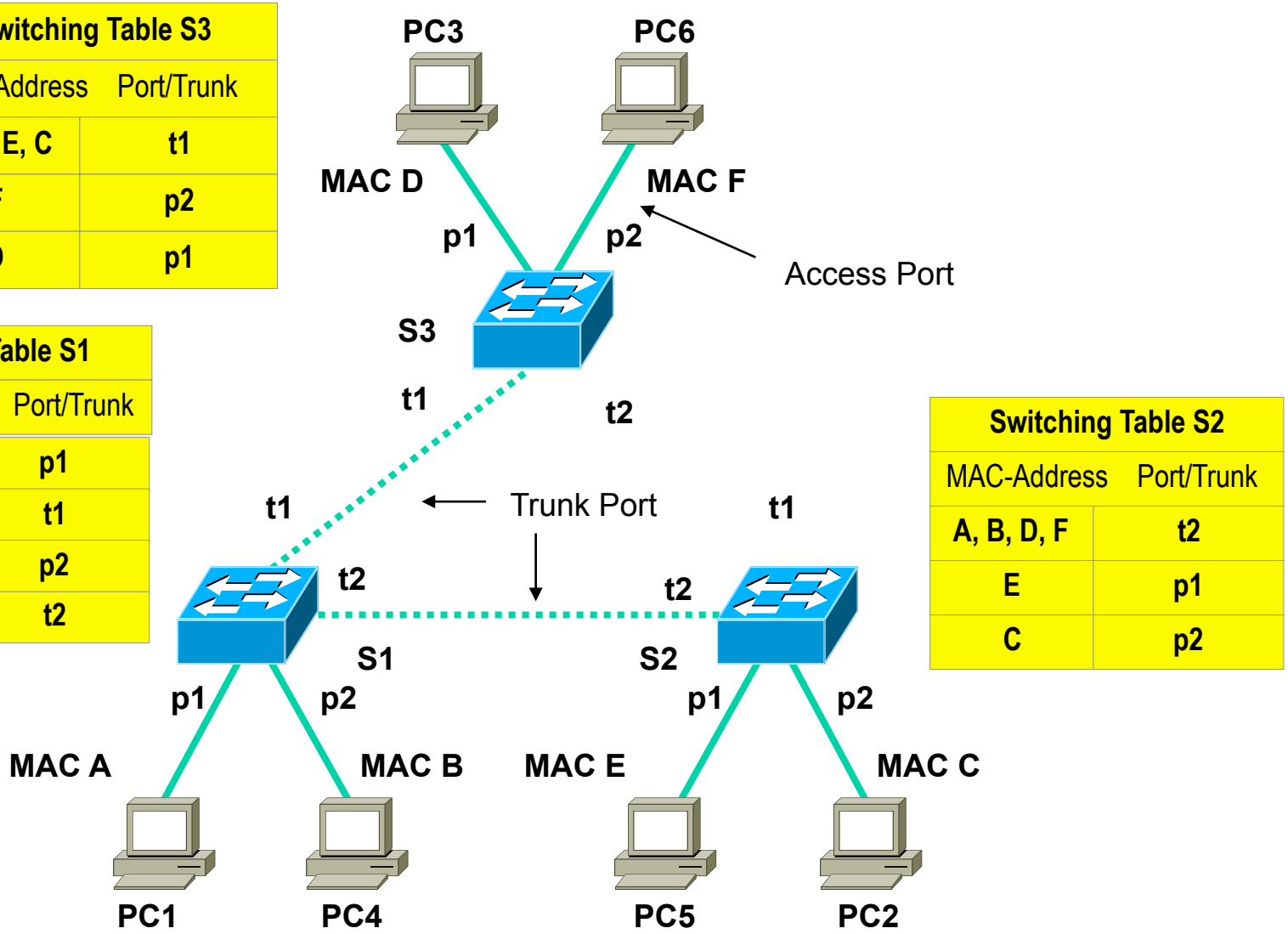


# Ethernet Switch Table – Final State
## (All MAC addresses learned)

| Switching Table S3 | |
|---|---|
| MAC-Address | Port/Trunk |
| A, B, E, C | t1 |
| F | p2 |
| D | p1 |

| Switching Table S1 | |
|---|---|
| MAC-Address | Port/Trunk |
| A | p1 |
| F, D | t1 |
| B | p2 |
| E, C | t2 |

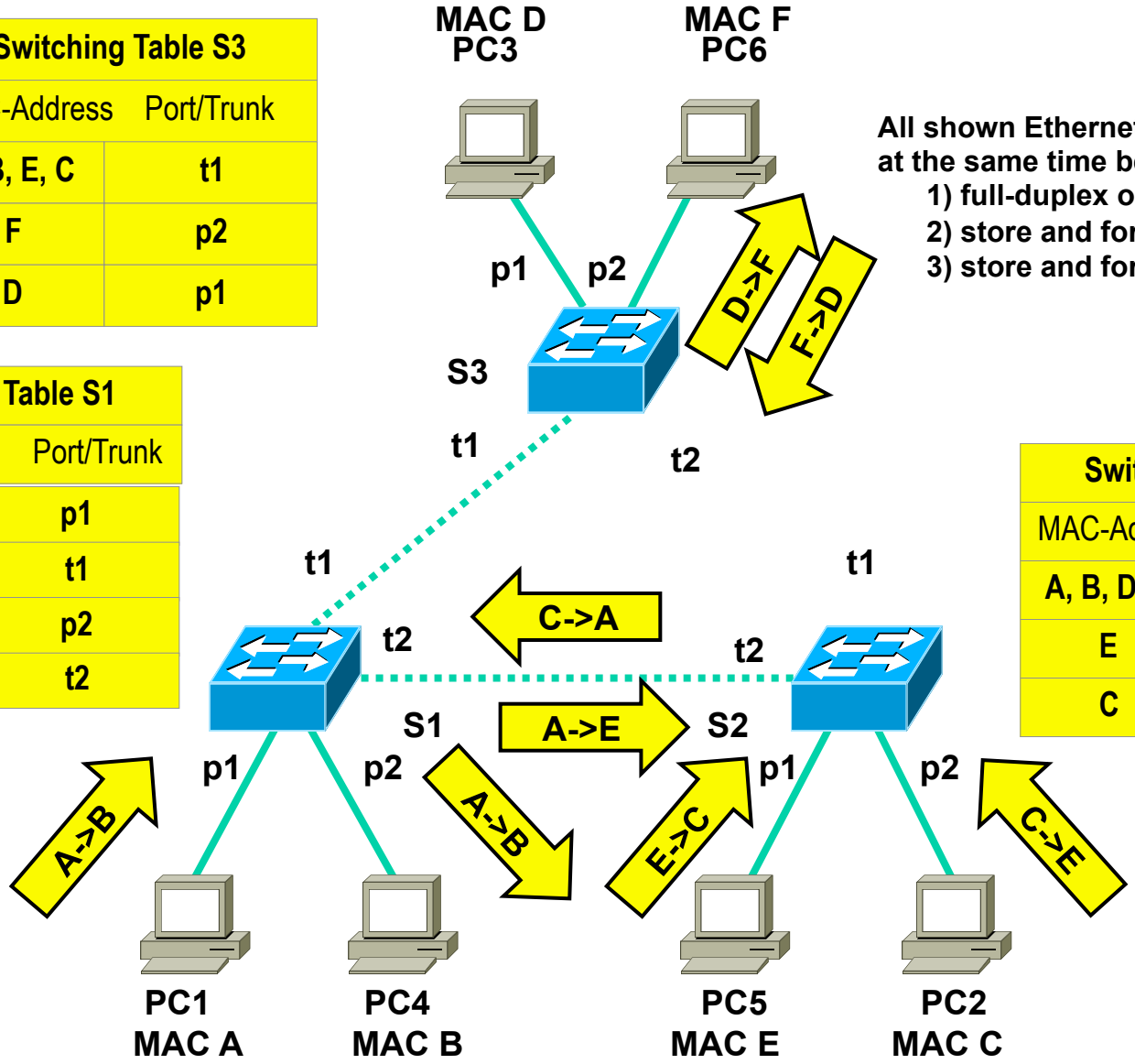
| Switching Table S2 | |
|---|---|
| MAC-Address | Port/Trunk |
| A, B, D, F | t2 |
| E | p1 |
| C | p2 |



# Ethernet Switching – Decoupling (Improved Performance <-> Collision Domains)

| Switching Table S3 | |
|---|---|
| MAC-Address | Port/Trunk |
| A, B, E, C | t1 |
| F | p2 |
| D | p1 |

| Switching Table S1 | |
|---|---|
| MAC-Address | Port/Trunk |
| A | p1 |
| F, D | t1 |
| B | p2 |
| E, C | t2 |



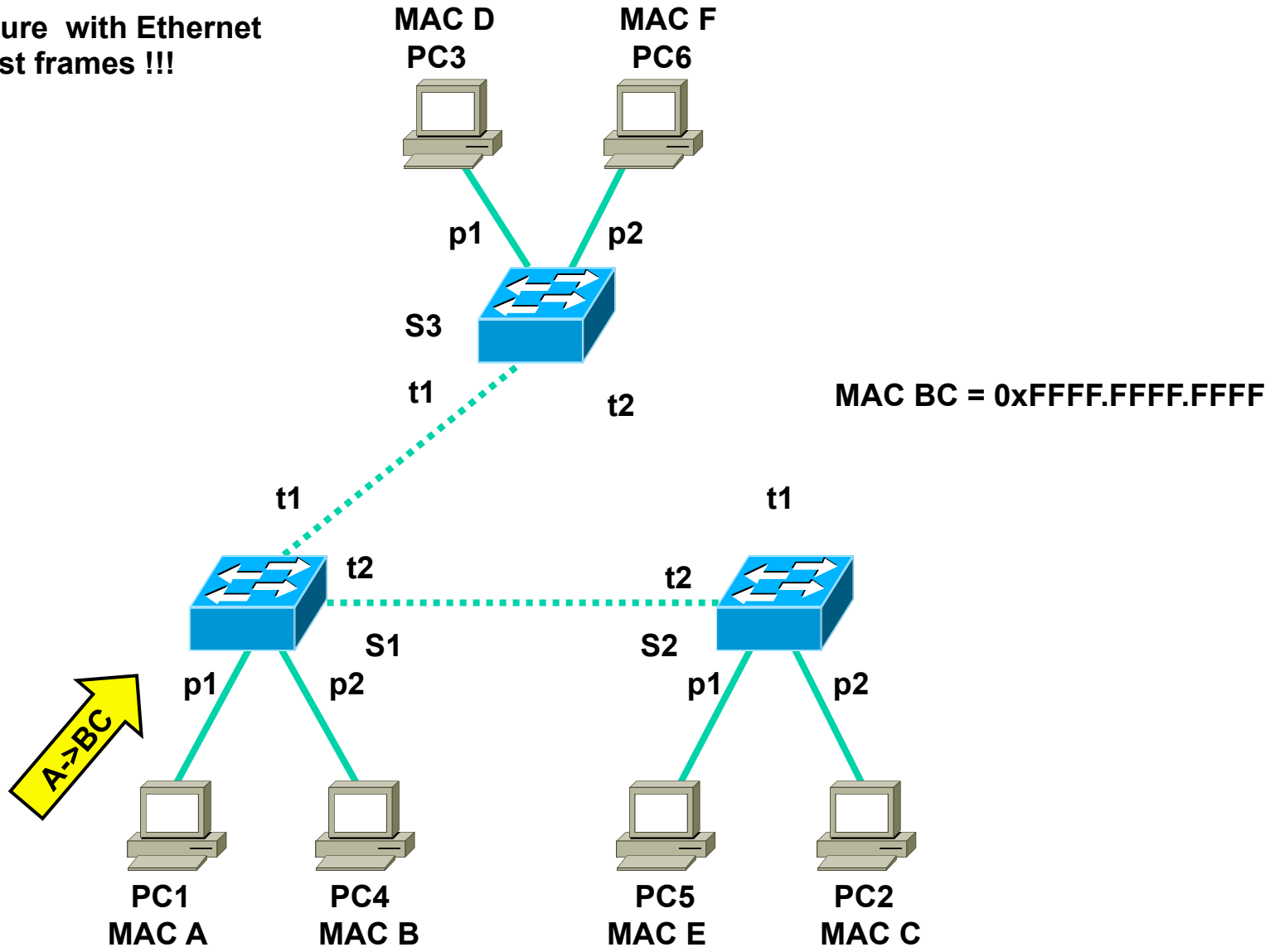

All shown Ethernet frames are forwarded at the same time because of:

1) full-duplex on every link
2) store and forward at switch
3) store and forward at NIC of PC

| Switching Table S2 | |
|---|---|
| MAC-Address | Port/Trunk |
| A, B, D, F | t2 |
| E | p1 |
| C | p2 |

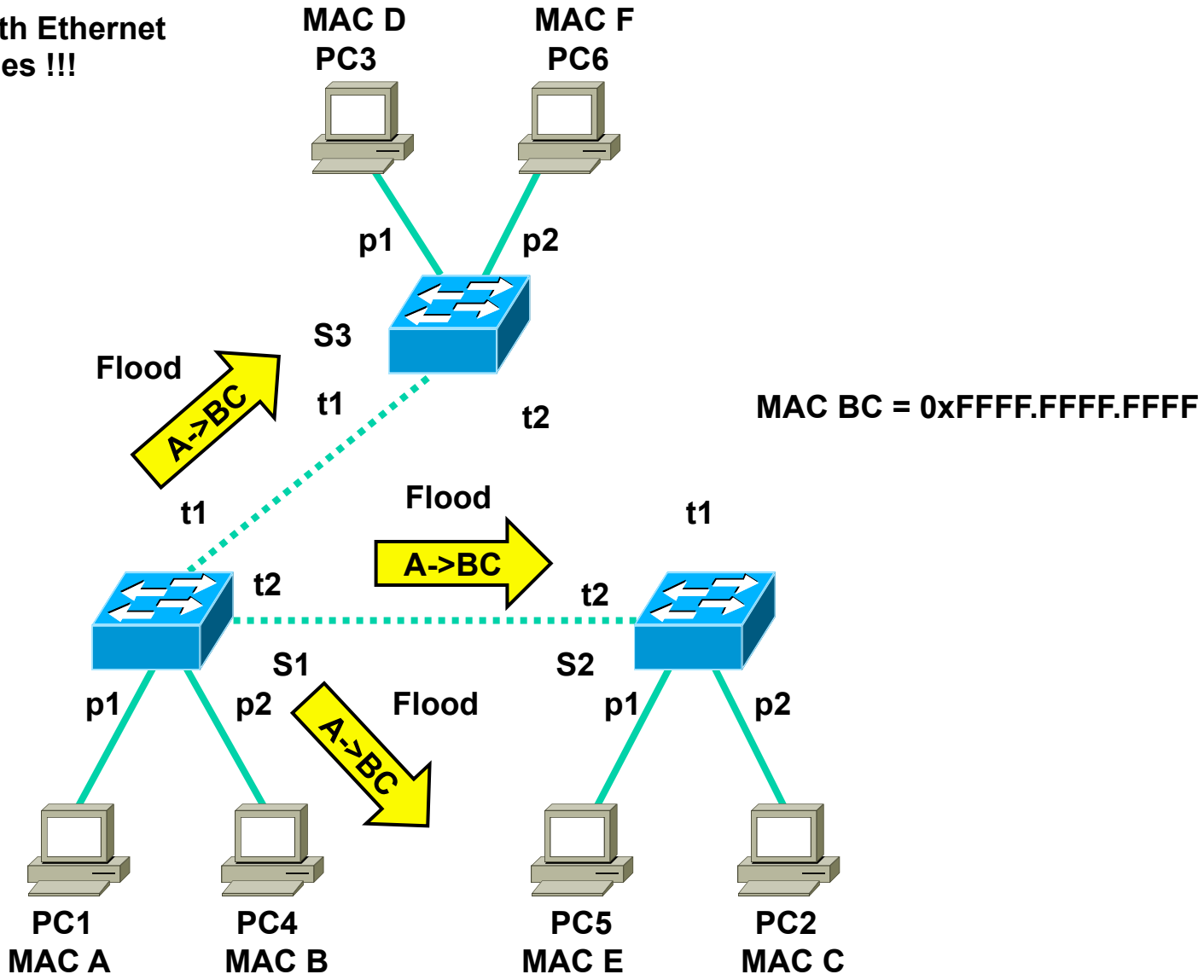
# Ethernet Broadcast (BC) 1

Same procedure with Ethernet multicast frames !!!



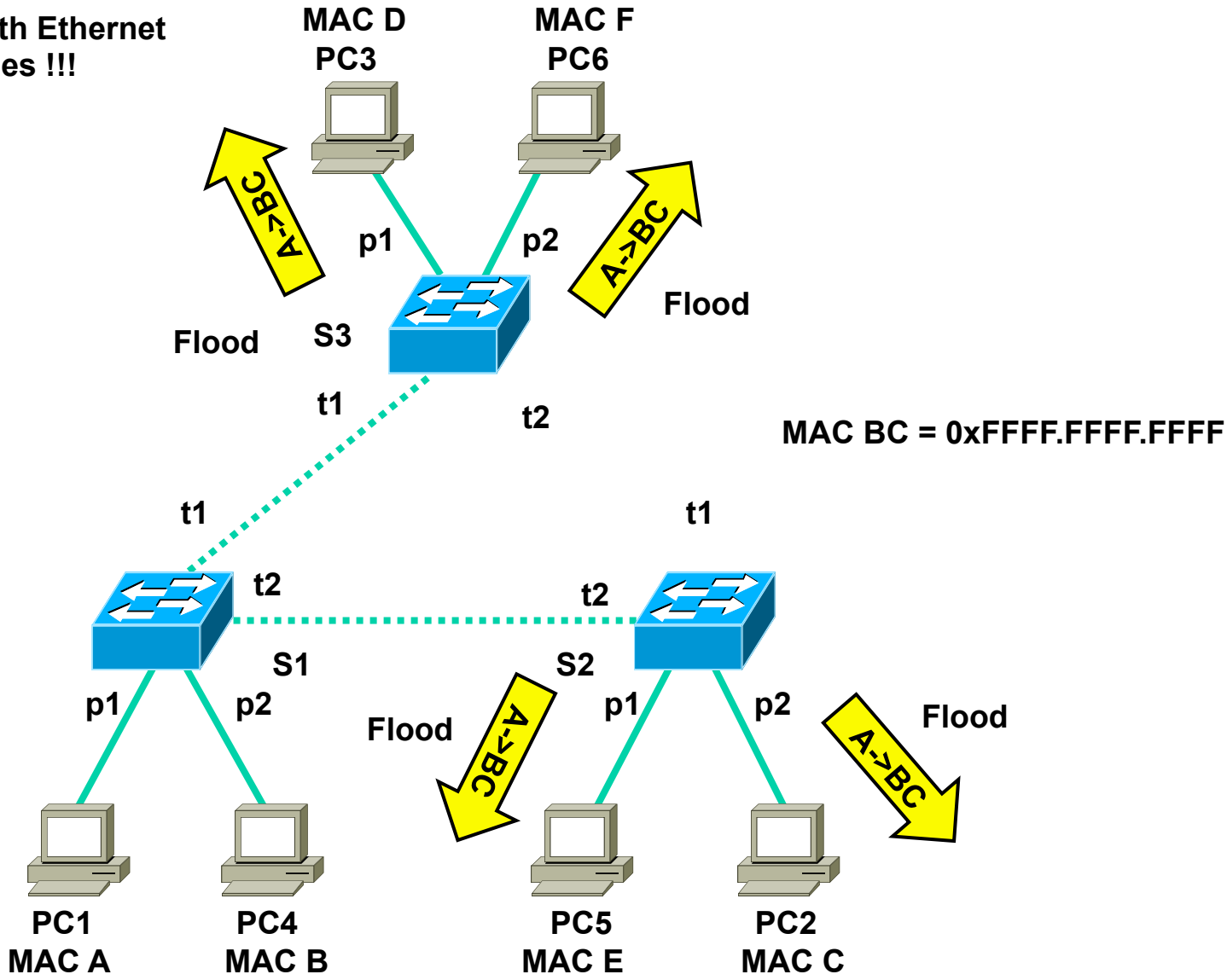
# Ethernet Broadcast (BC) 2

Same procedure with Ethernet multicast frames !!!

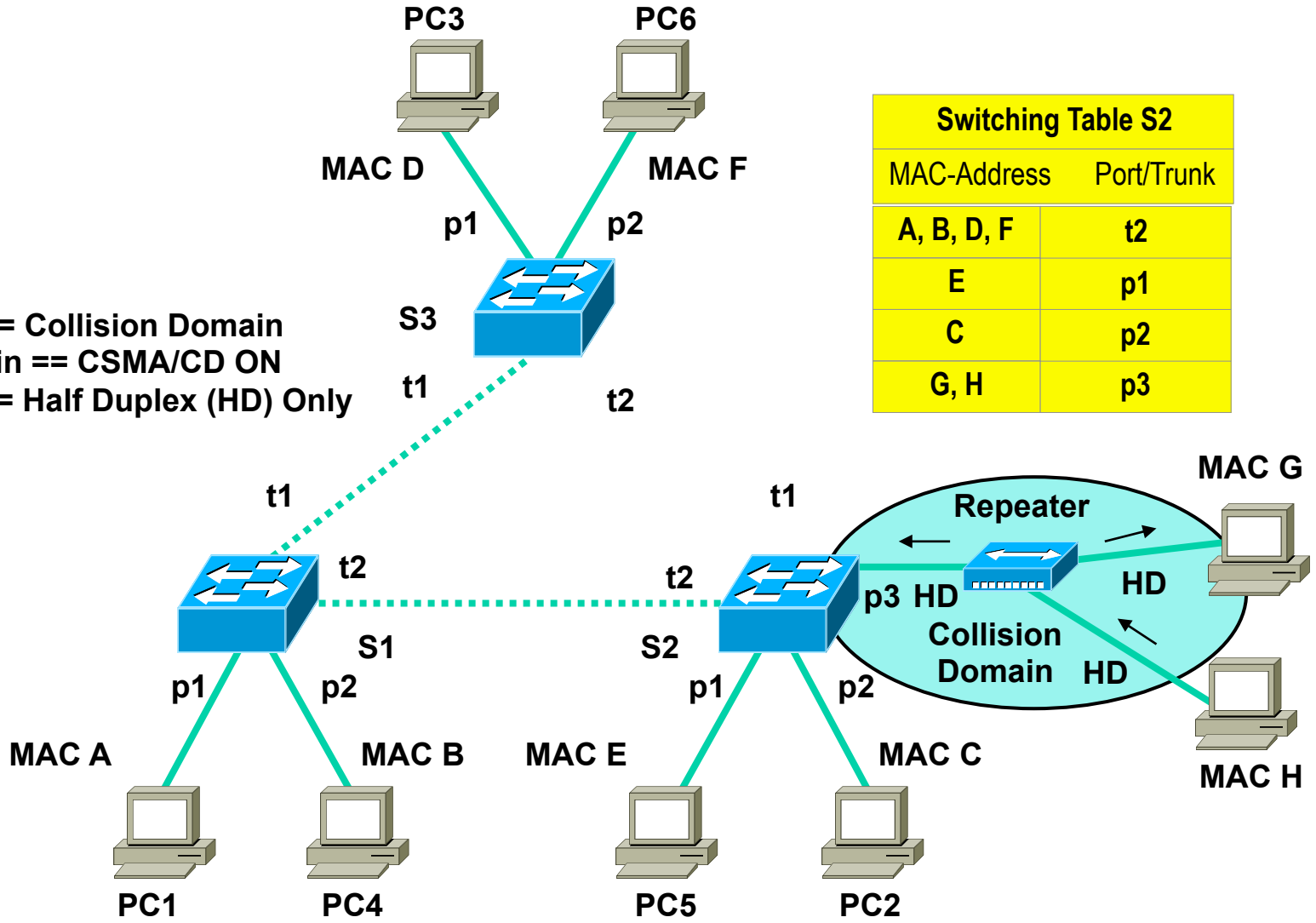


# Ethernet Broadcast (BC) 3

Same procedure with Ethernet multicast frames !!!



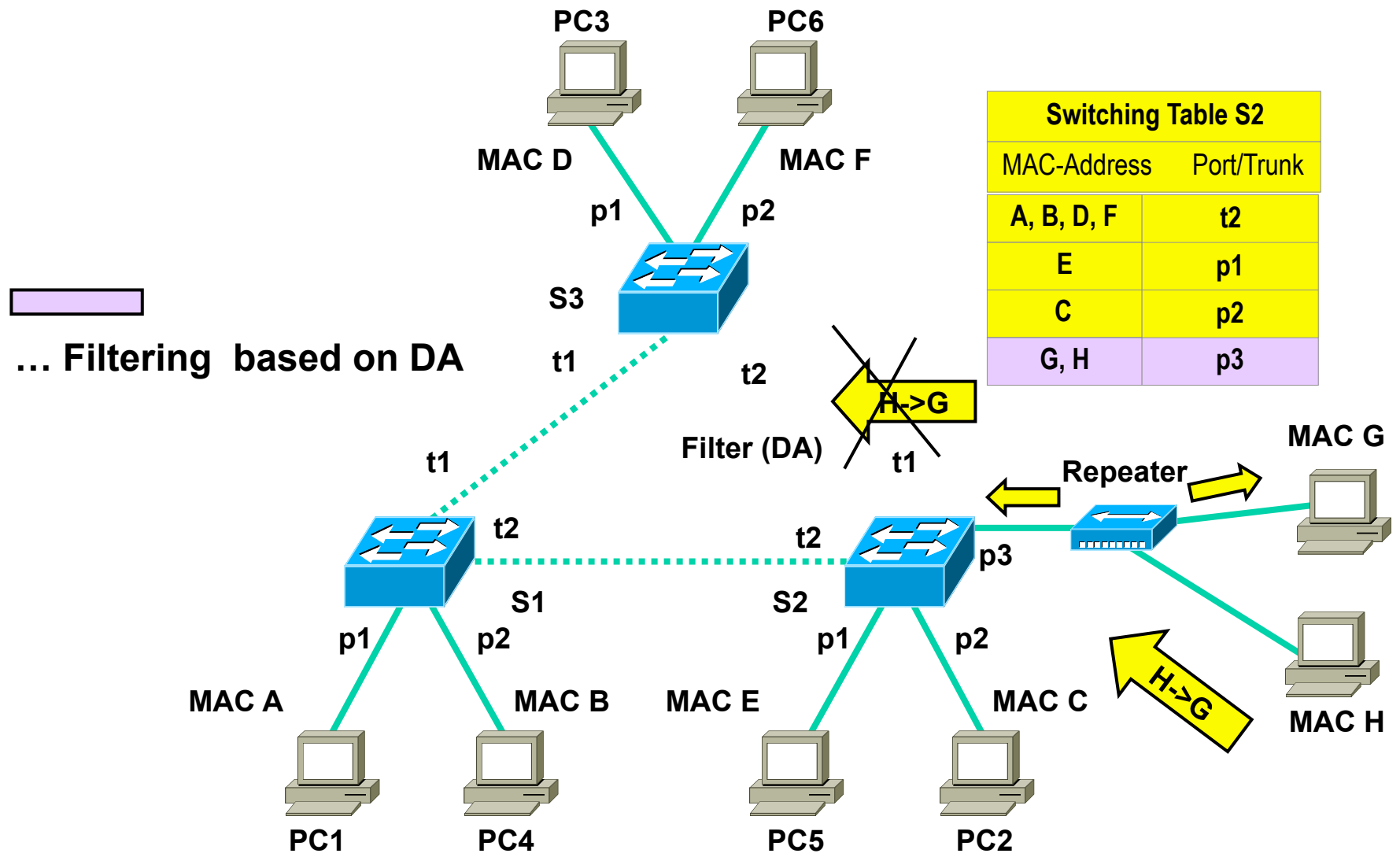
# Ethernet Switching – Repeater (Hub)
## (On One Link -> Shared Media – Half Duplex (HD))




Shared Media == Collision Domain
Collision Domain == CSMA/CD ON
CSMA/CD ON == Half Duplex (HD) Only

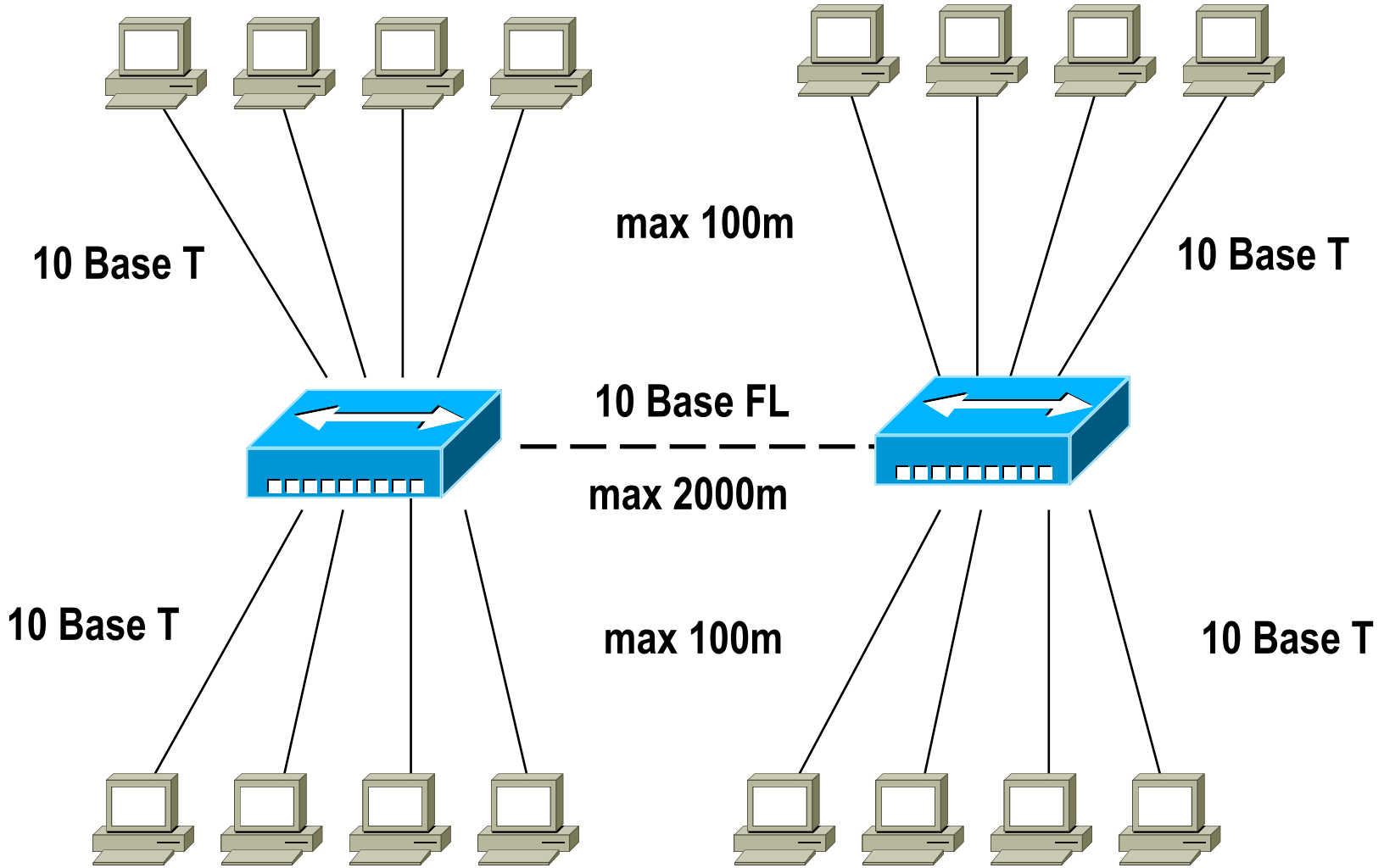
| Switching Table S2 | |
|---|---|
| MAC-Address | Port/Trunk |
| A, B, D, F | t2 |
| E | p1 |
| C | p2 |
| G, H | p3 |

# Table Usage (Filtering Decision) for Ethernet Frame MAC-H to MAC-G




| Switching Table S2 | |
|---|---|
| MAC-Address | Port/Trunk |
| A, B, D, F | t2 |
| E | p1 |
| C | p2 |
| G, H | p3 |

# Ethernet with Repeater: Network Sniffing? Yes -> Ethernet Card -> Promiscuous Mode

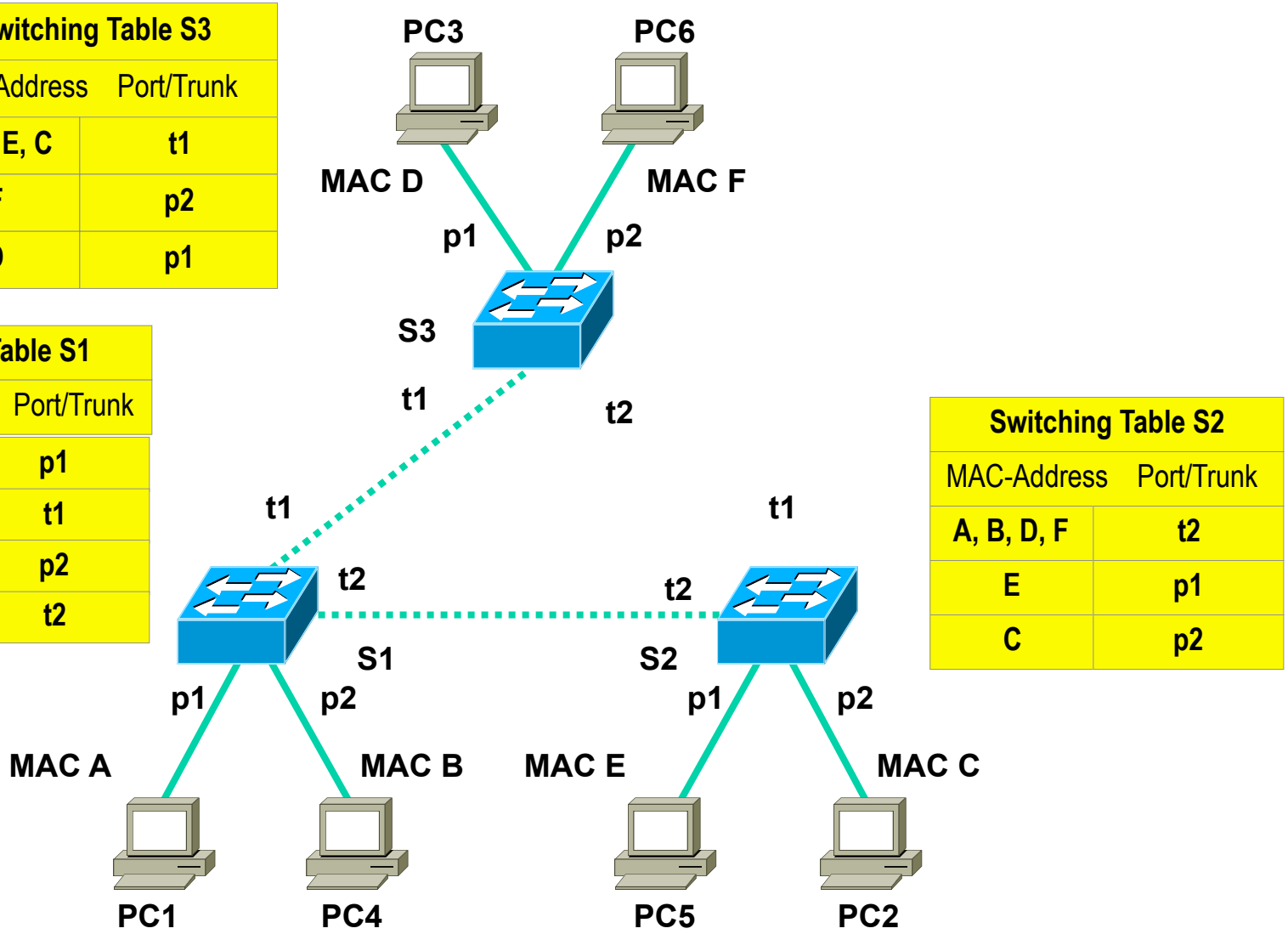


# Ethernet with Switches: Network Sniffing? Not So Easy -> Because of Inherent Filtering

| Switching Table S3 | |
|---|---|
| MAC-Address | Port/Trunk |
| A, B, E, C | t1 |
| F | p2 |
| D | p1 |

| Switching Table S1 | |
|---|---|
| MAC-Address | Port/Trunk |
| A | p1 |
| F, D | t1 |
| B | p2 |
| E, C | t2 |

| Switching Table S2 | |
|---|---|
| MAC-Address | Port/Trunk |
| A, B, D, F | t2 |
| E | p1 |
| C | p2 |



# Modern Switching Features

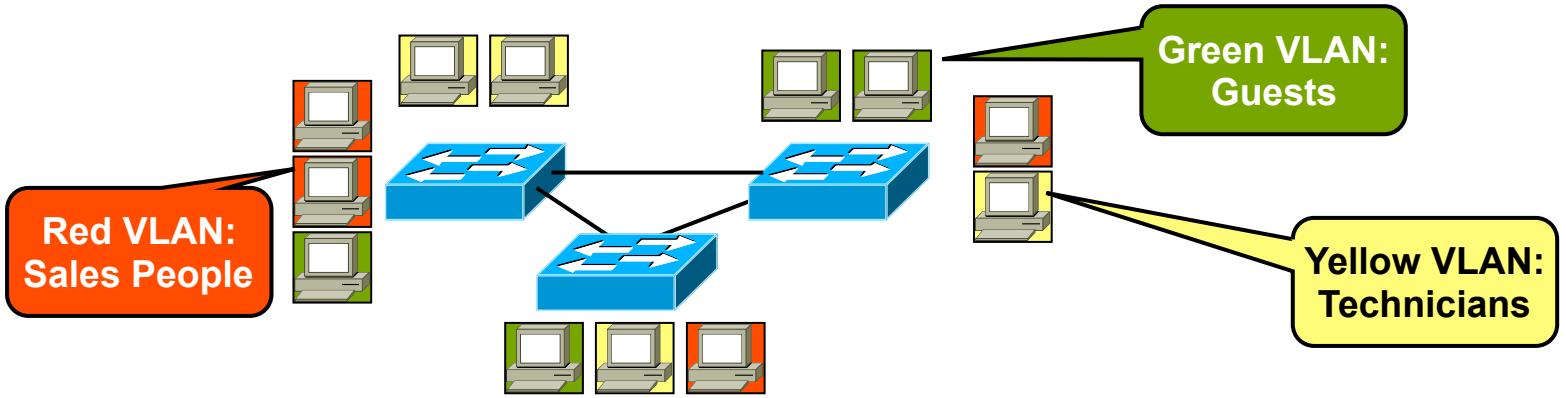
- **Different data rates supported simultaneously**
  - 10, 100, 1000, 10000 Mbit/s depending on switch
- **Full duplex operation**
- **QoS (802.1p)**
  - Queuing mechanisms
  - Flow control
- **Security features**
  - Restricted static mappings (DA associated with source port)
  - Port secure (Limited number of predefined users per port)
- **Different forwarding**
  - Store & Forward
  - Cut-through
  - Fragment-Free
- **VLAN support (tagging, trunking, 802.1Q)**
- **Spanning Tree (RSTP, MSTP, PVST+)**
- **SPAN (for monitoring traffic)**

# Agenda

- **Ethernet Evolution**
- **VLAN**
- **High Speed Ethernet**
  - Introduction
  - Fast Ethernet
  - Gigabit Ethernet
  - 10 Gigabit Ethernet

# Virtual LANs

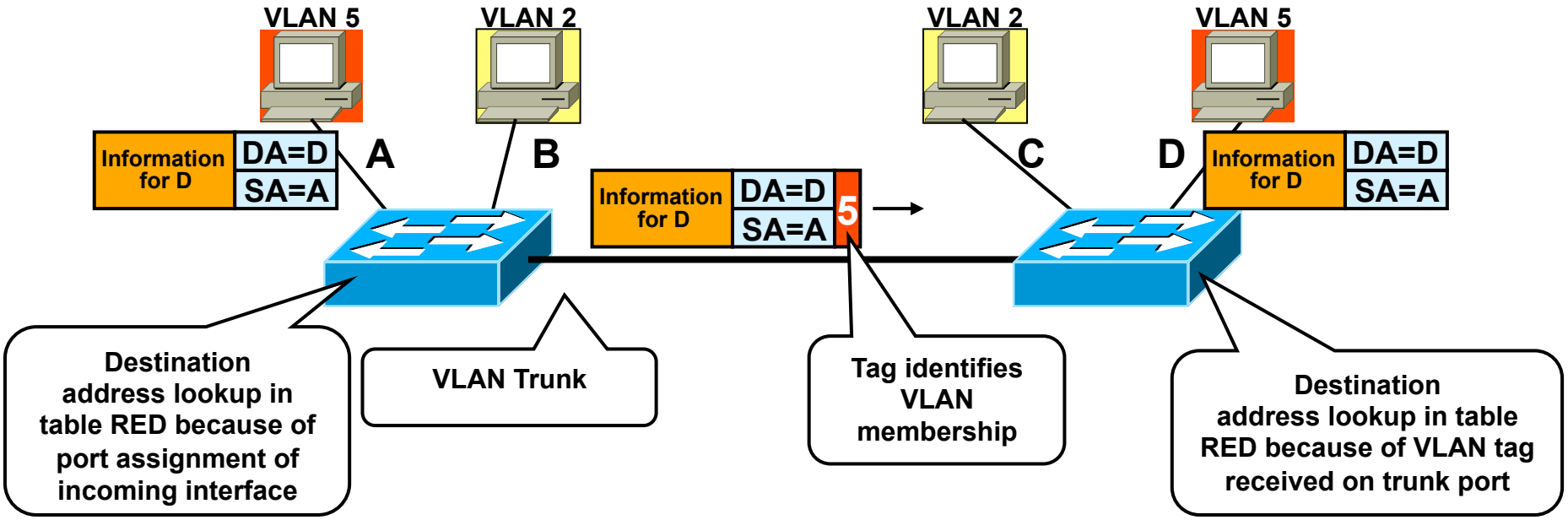
- **Separate LAN into multiple broadcast domains**
  - No global broadcasts anymore
  - For security reasons
- **Assign users to "VLANs“**
- **Base Idea:**
  - Multiplexing of several LANs over the same infrastructure (Ethernet switches and connection between switches)



# Host to VLAN Assignment

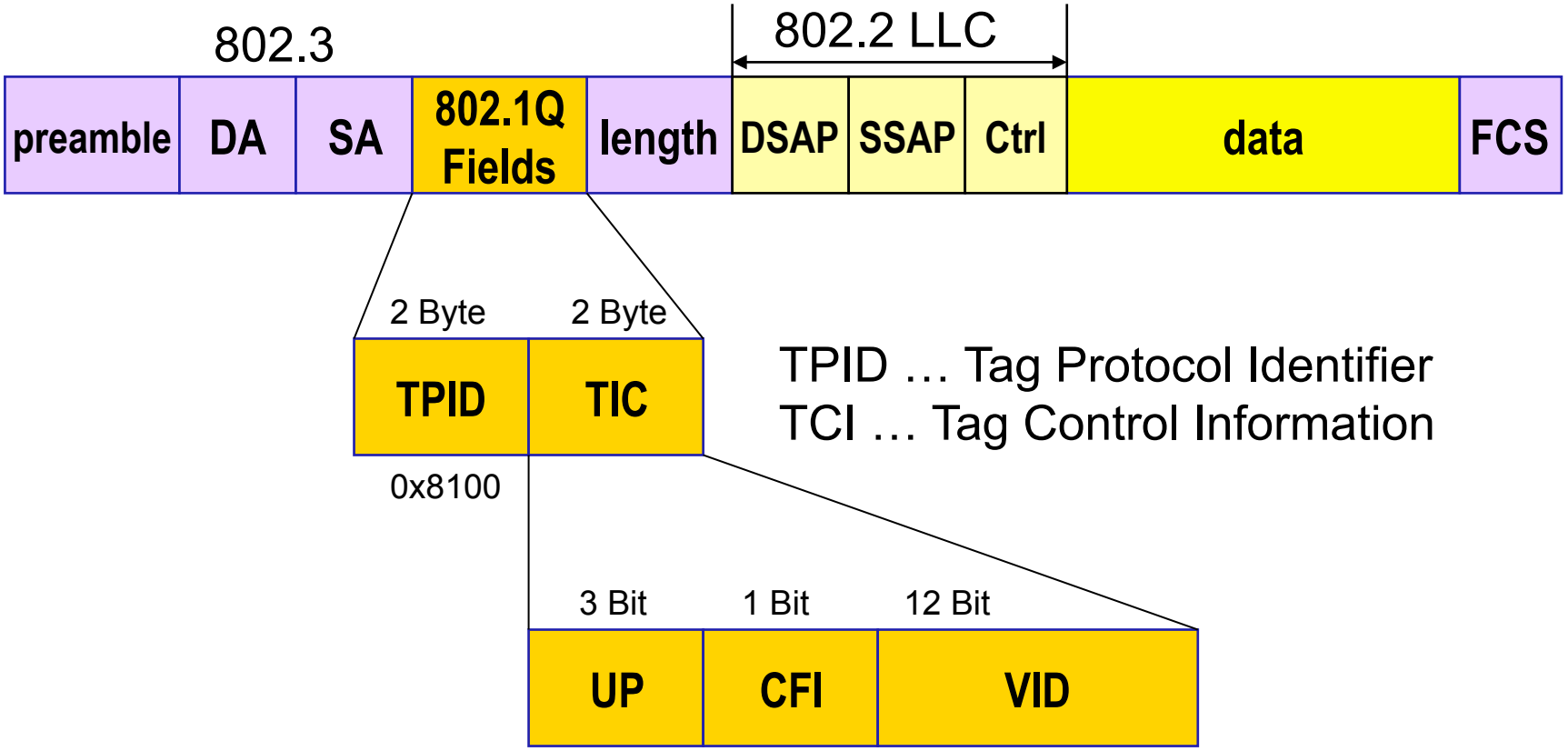
- **Different solutions**
  - Port based assignment
  - Source address assignment
  - Protocol based
  - Complex rule based
  - 802.1X based on the credentials of a user / machine provided by EAP authentication
- **Bridges are interconnected via VLAN trunks**
  - IEEE 802.1Q (former 802.1s)
  - ISL (Cisco)
  - IEEE 802.10 (pre 802.1Q temporary solution, outdated)

# VLAN Trunking with Port Assignment




- **Packets across the VLAN trunk are tagged**
  - Either using 802.1Q or ISL tag
  - So next bridge is able to constrain frame to same VLAN as the source
- **Inter-VLAN communication not possible**
  - Only IP router can forward inter-VLAN traffic

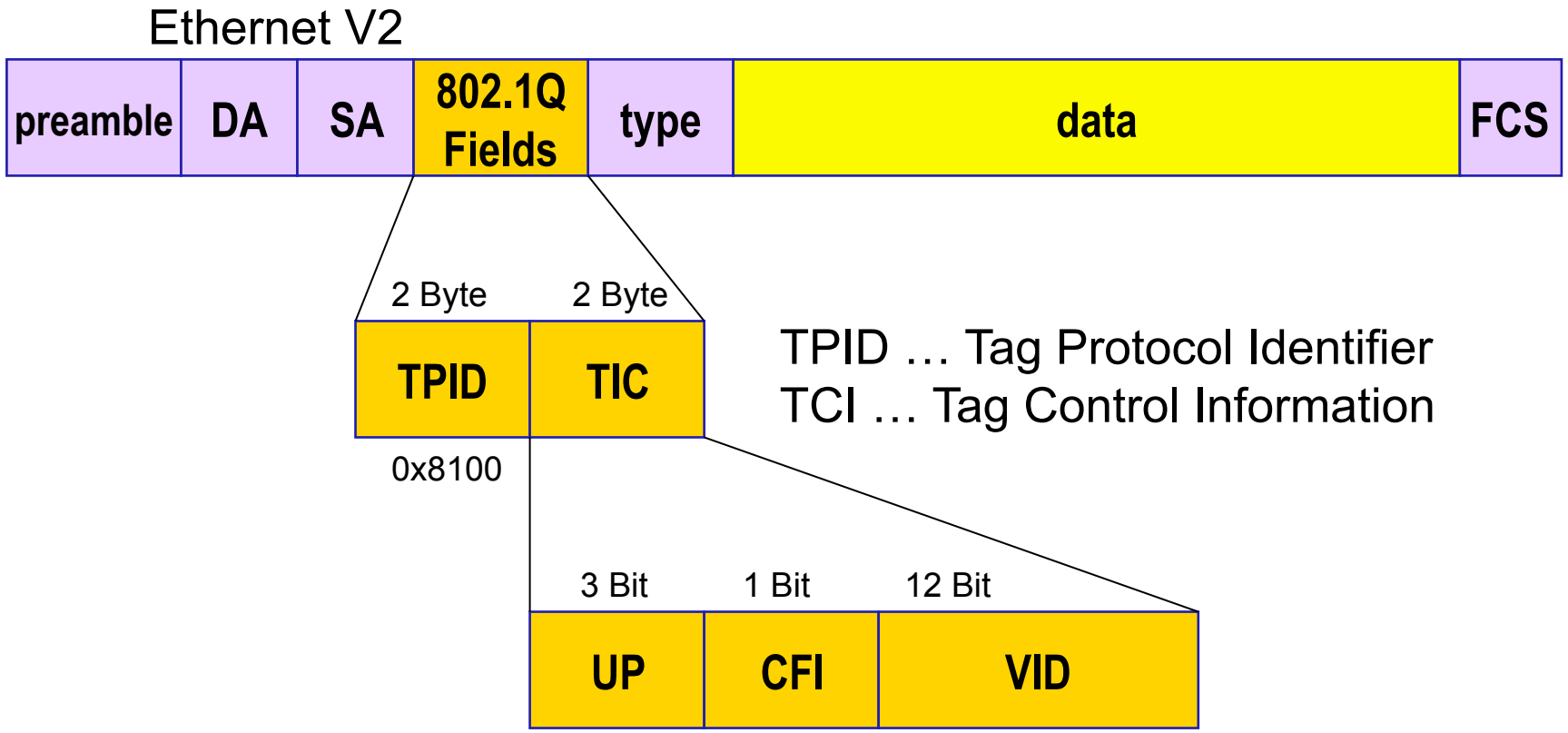
# 802.1Q VLAN Tagging – LLC (1)




note: With tagging Ethernets maximal frame length = 1522, minimal frame length = 68

UP … User Priority for L2 QoS = COS
CFI … Canonical Format Identifier
VID … VLAN Identifier

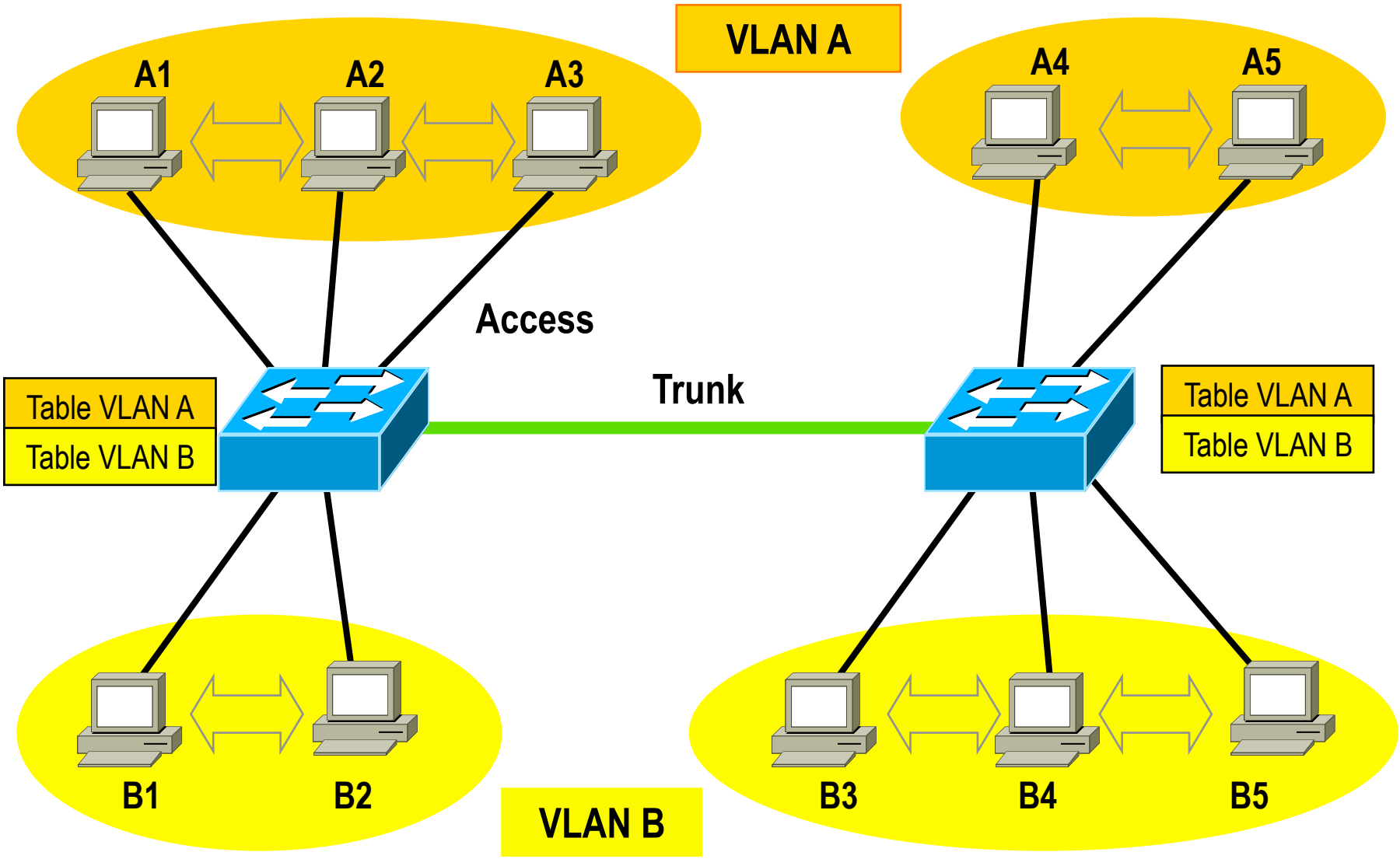
# 802.1Q VLAN Tagging – Ev2 (2)



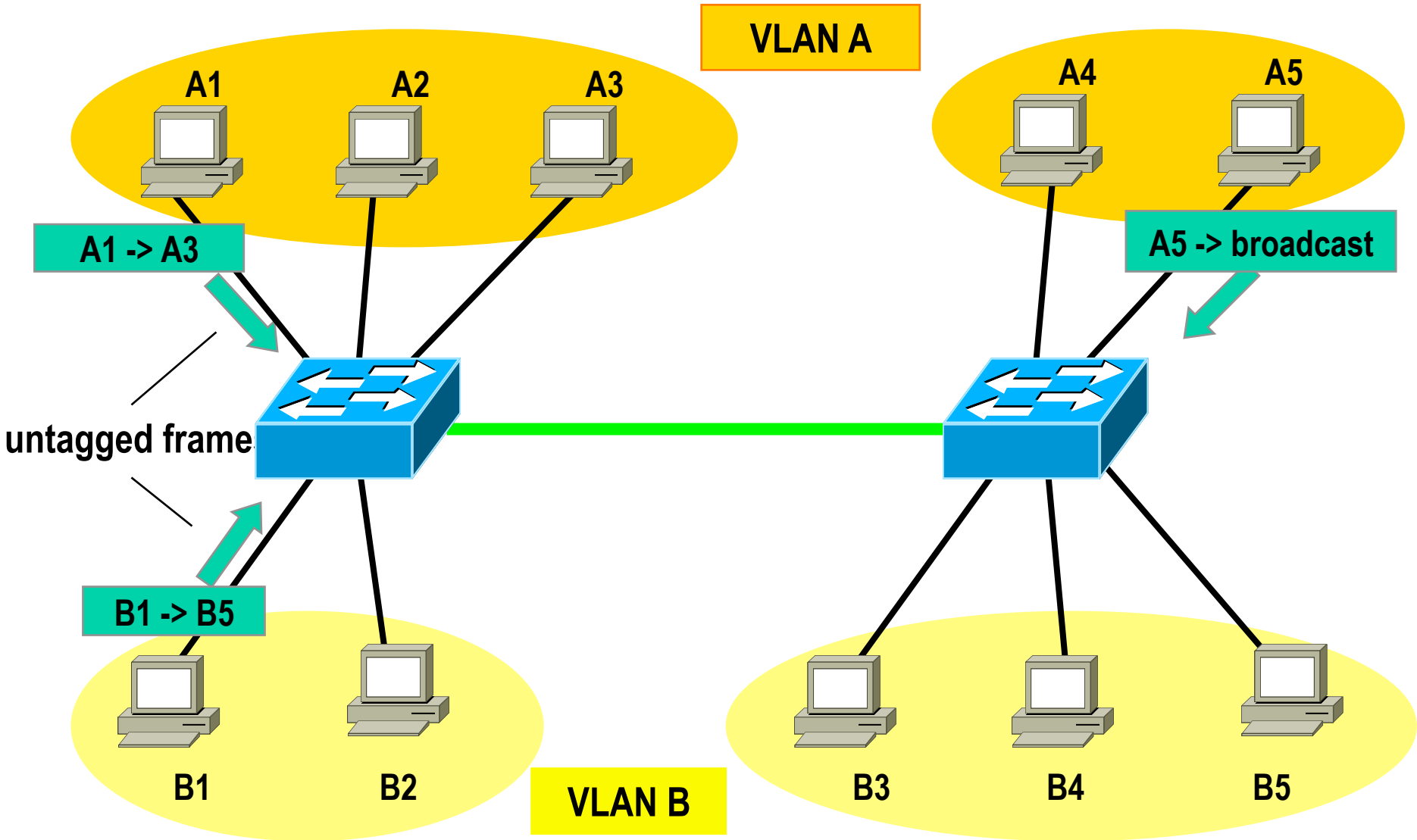

note: With tagging Ethernets maximal frame length = 1522, minimal frame length = 68

UP … User Priority for L2 QoS = COS
CFI … Canonical Format Identifier
VID … VLAN Identifier

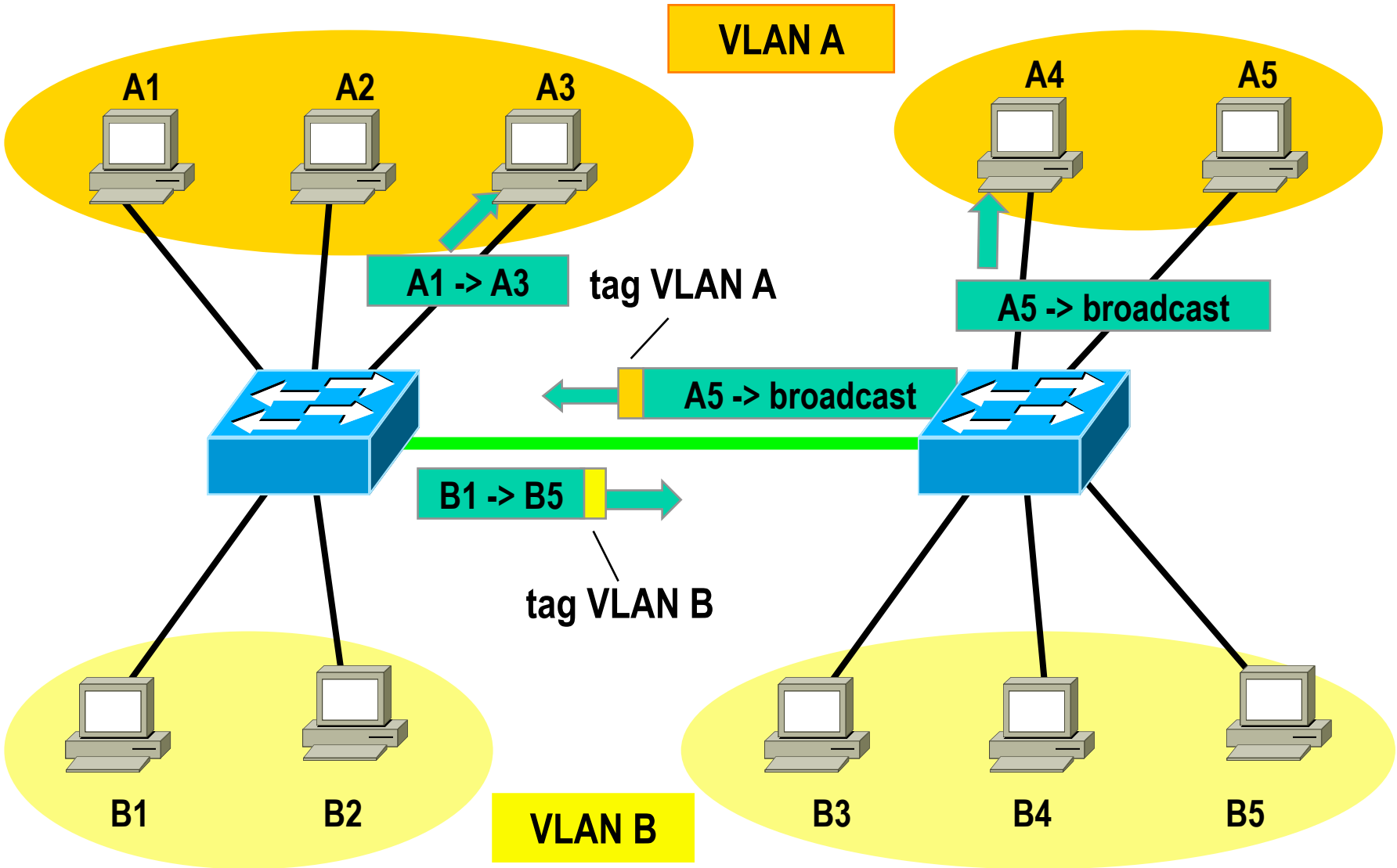
# VLAN Trunking Example



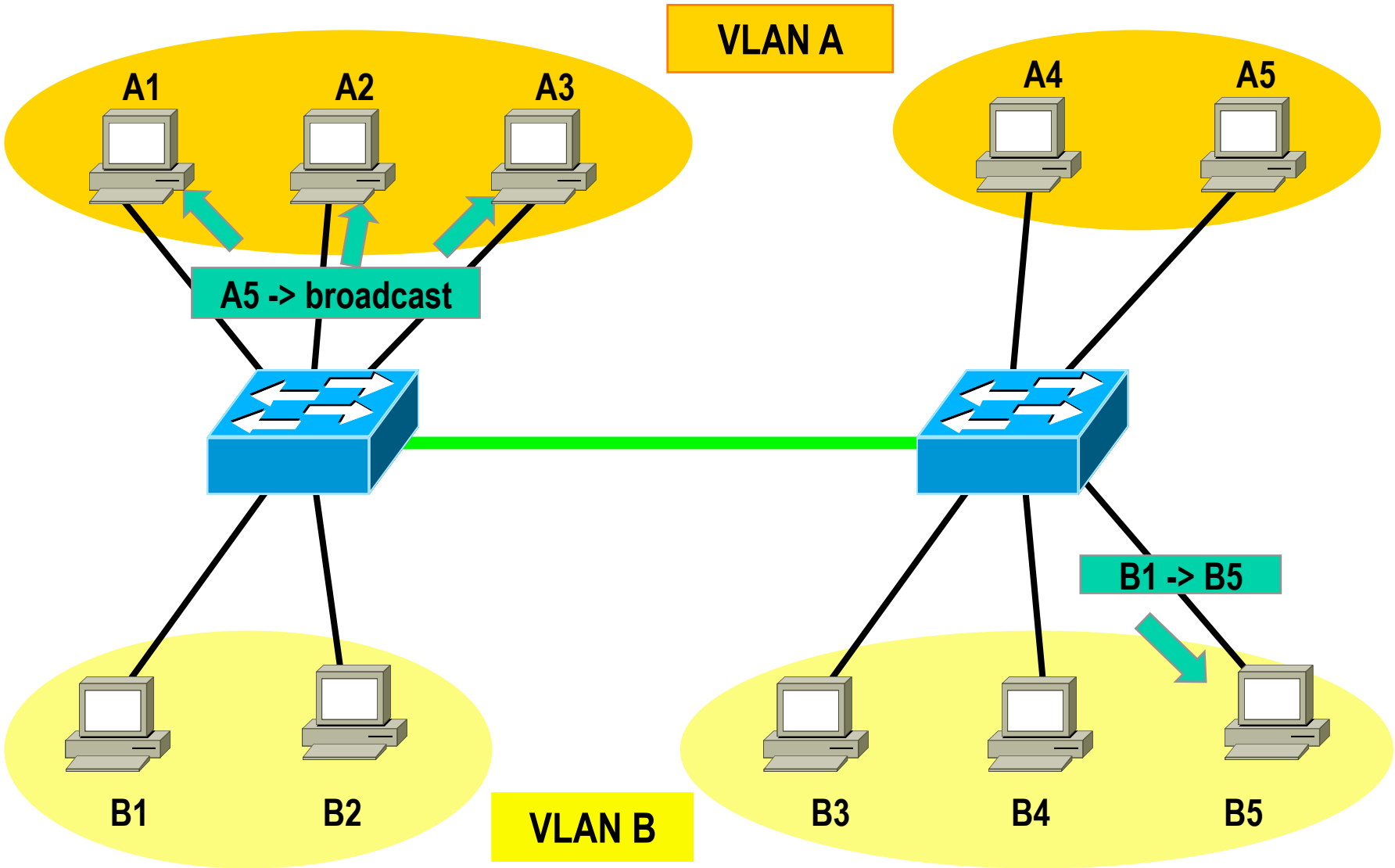
# VLAN in Operation (1)



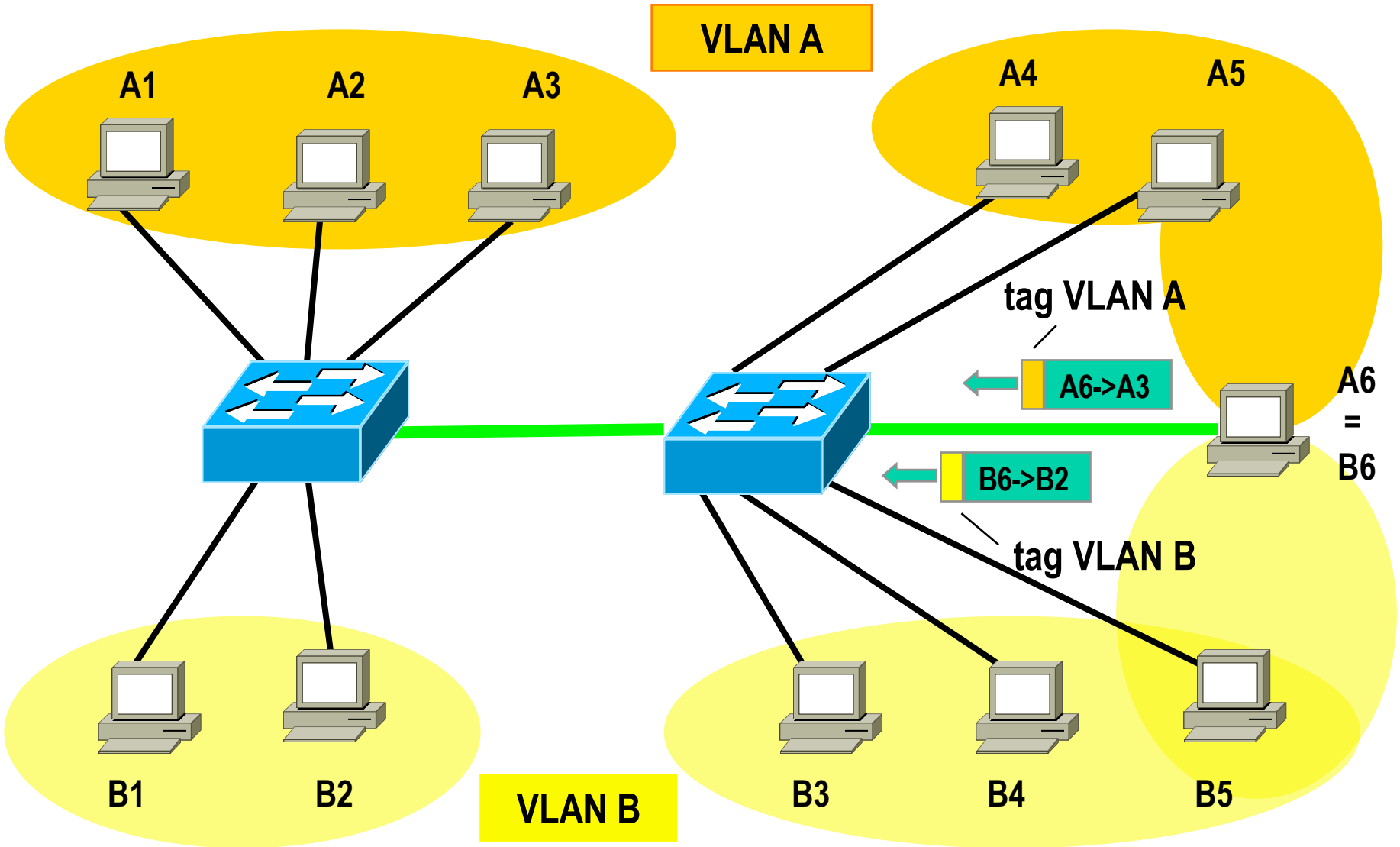
# VLAN in Operation (2)



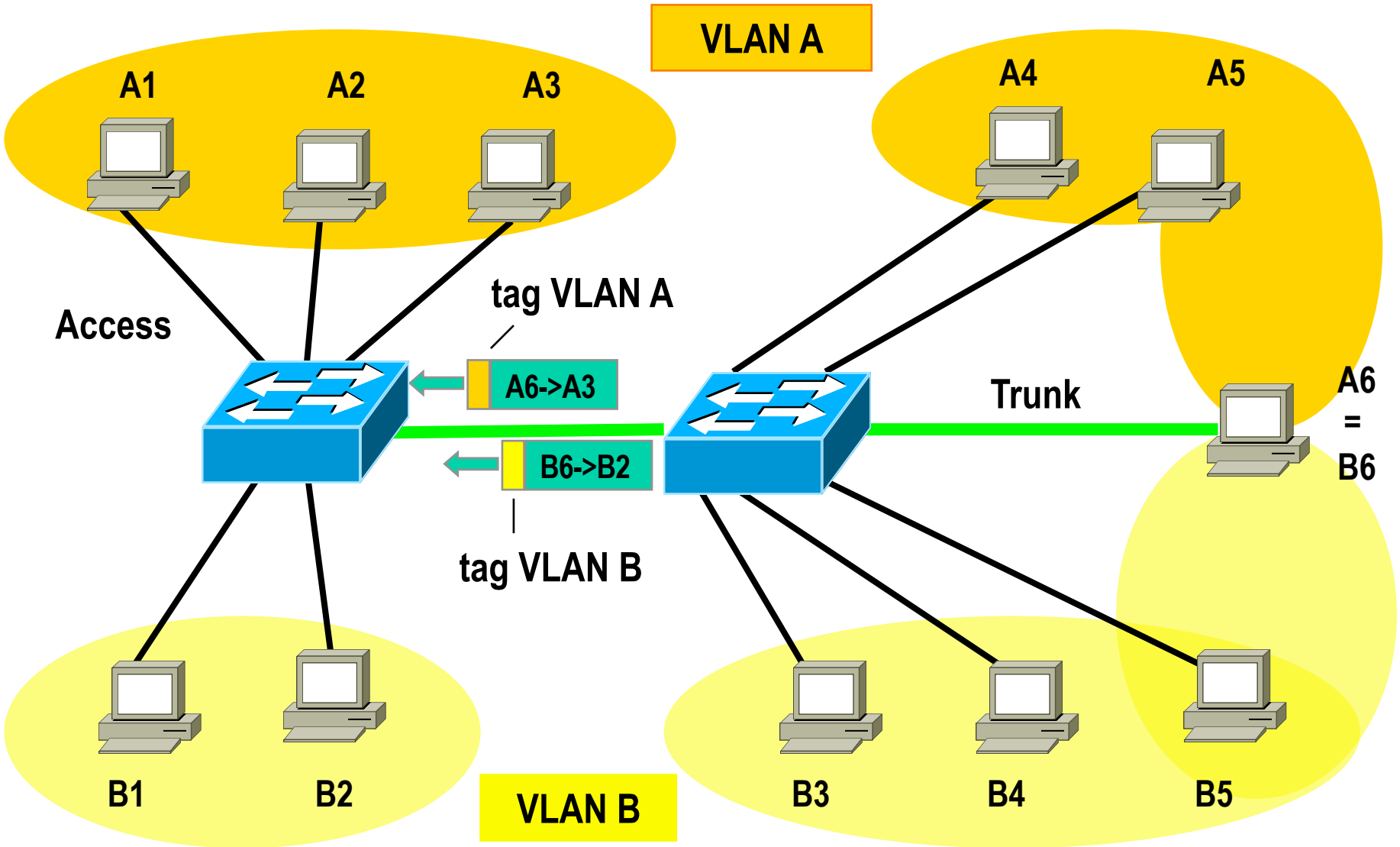
# VLAN in Operation (3)



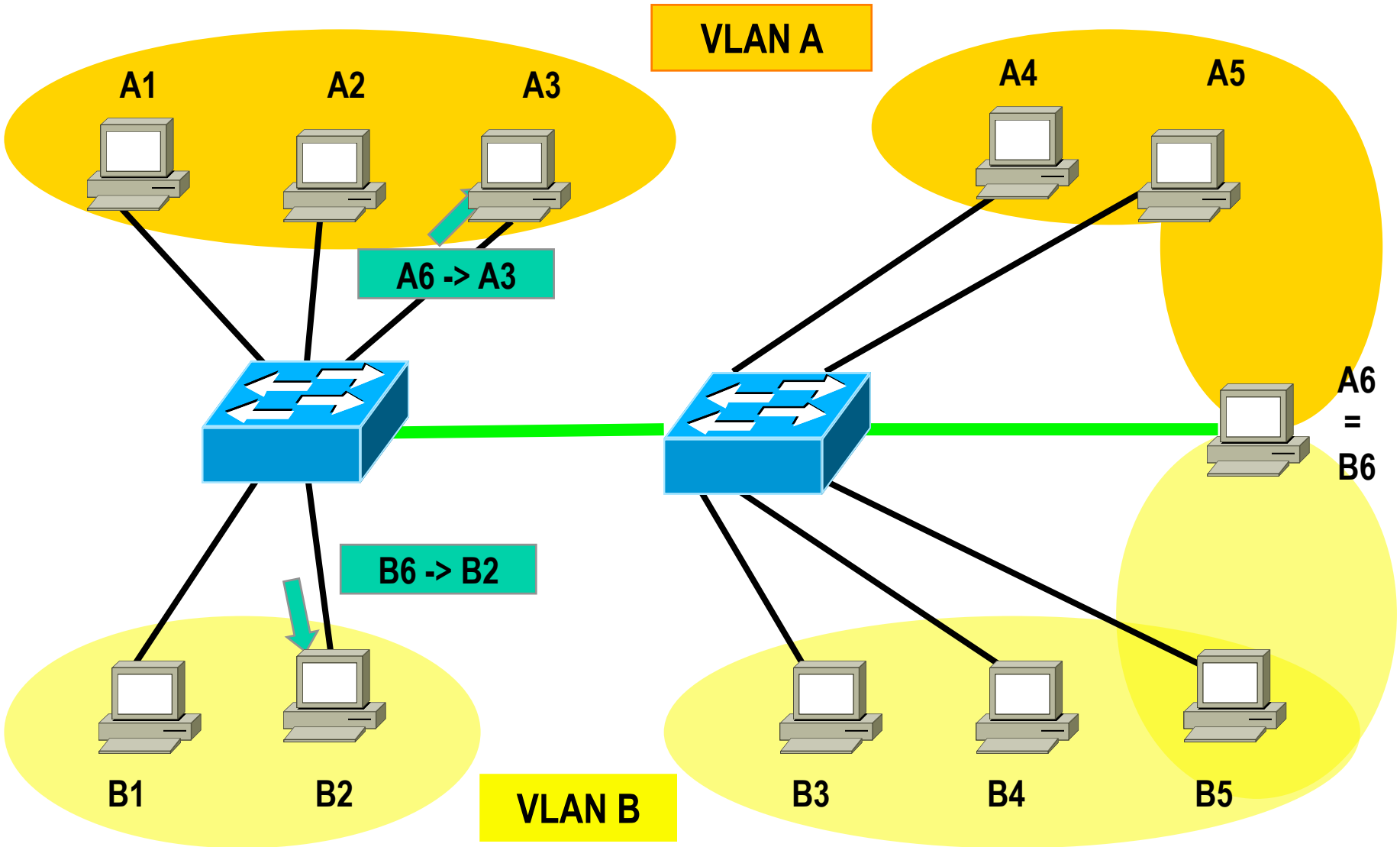
# Multihomed VLAN (1)



# Multihomed VLAN (2)

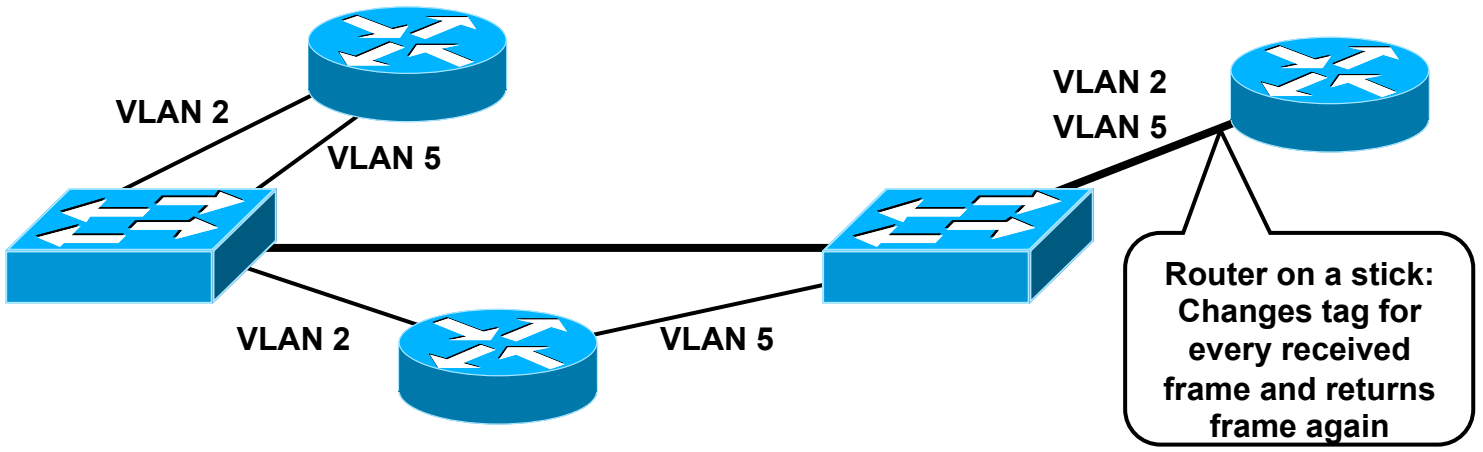


# Multihomed VLAN (3)

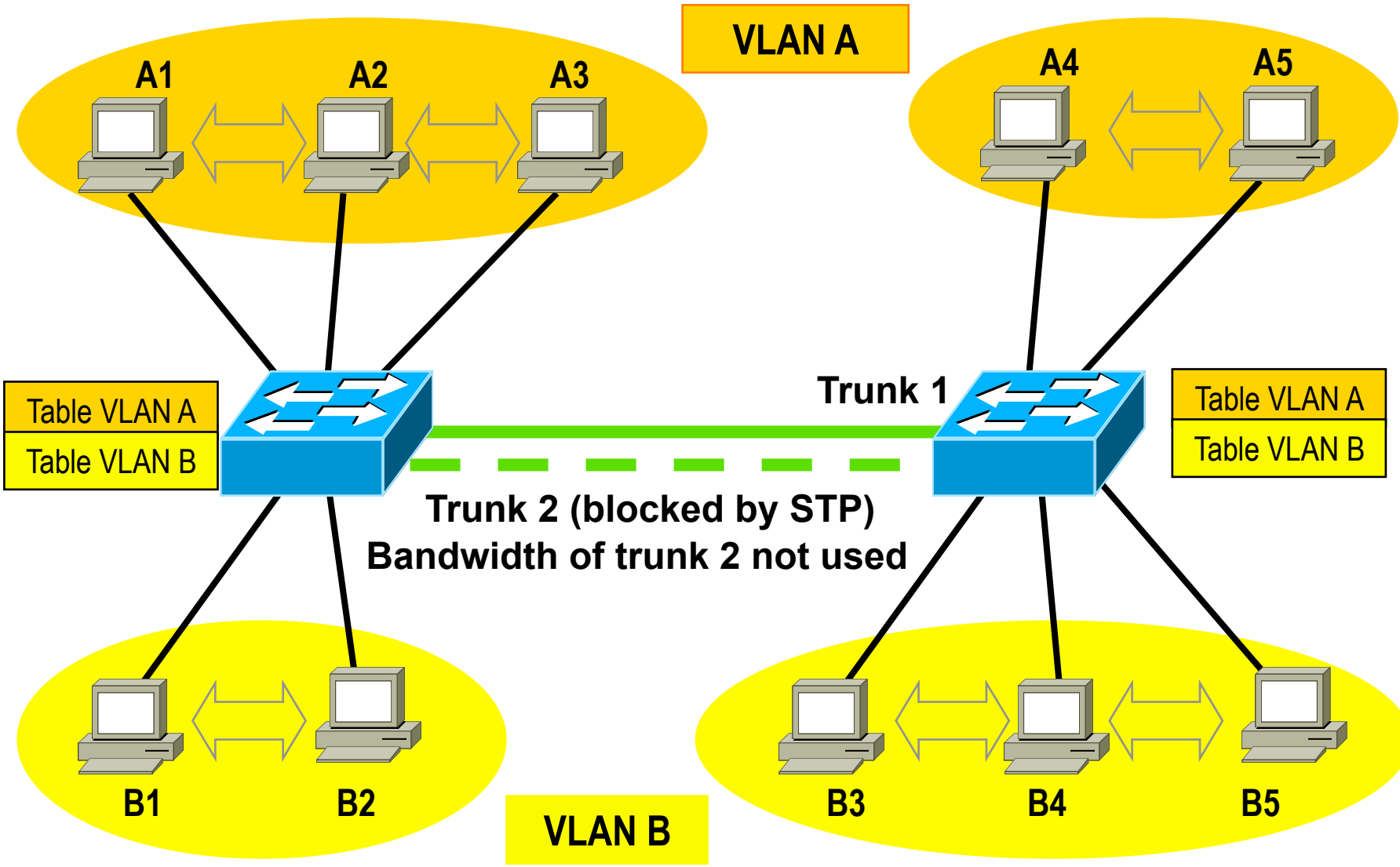


# Inter-VLAN Traffic

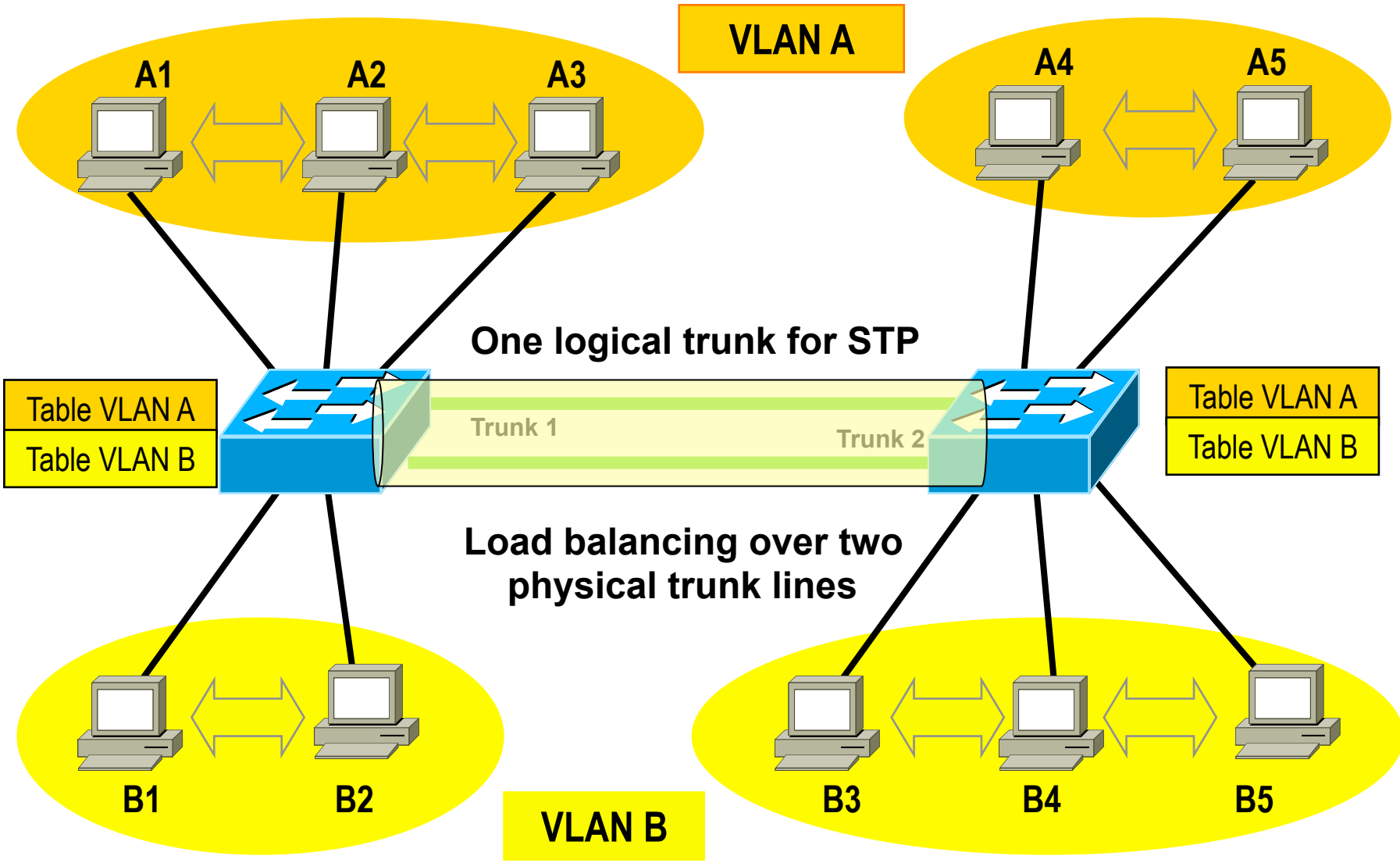
- **Router can forward inter-VLAN traffic**
  - Terminates Ethernet links
  - Requirement: Each VLAN in other IP subnet !
- **Two possibilities**
  - Router is member of every VLAN with one link each
  - Router attached on VLAN trunk port ("Router on a stick")



# Trunking without LACP / FEC / GEC



# Trunking with LACP / FEC / GEC



# Agenda

- **Ethernet Evolution**
- **VLAN**
- **High Speed Ethernet**
  - Introduction
  - Fast Ethernet
  - Gigabit Ethernet
  - 10 Gigabit Ethernet

# Ethernet Switching

- **Ethernet switches can connect end systems with 10 Mbit/s, 100 Mbit/s or 1000 Mbit/s for example**
  - Clients may using 100 Mbit/s and server may use 1000 Mbit/s using a full duplex, point-to-point connection with switch.
  - Note: multiport repeater is not able to do this!
  - Ethernet frame has not changed!
- **It is still connectionless packet switching on L2 based**
  - Asynchronous TDM principle, buffers
  - Flow control would be great
    - Modern switches can avoid congestion can by supporting a new MAC control frame (so called pause command)

# IEEE 802.3 (2008)

- **The latest version specifies**
  - Operation for 10 Mbit/s, 100 Mbit/s, Gigabit/s and 10Gigabit/sEthernet over copper and fiber
  - Full duplex Ethernet
  - Auto-negotiation
  - Flow control
- **It is still backward compatible to the old times of Ethernet**
  - CSMA/CD (half-duplex) operation in 100 and 1000 Mbit/s Ethernets with multiport repeater possible
  - Frame bursting or carrier extension for ensuring slot-time demands in 1000 Mbit/s Ethernet
  - 10Gigabit/s Ethernet is full duplex only
    - CSMA/CD has died!!!
  - Ethernet frame is identical across all speeds

# What About Gigabit Hubs?

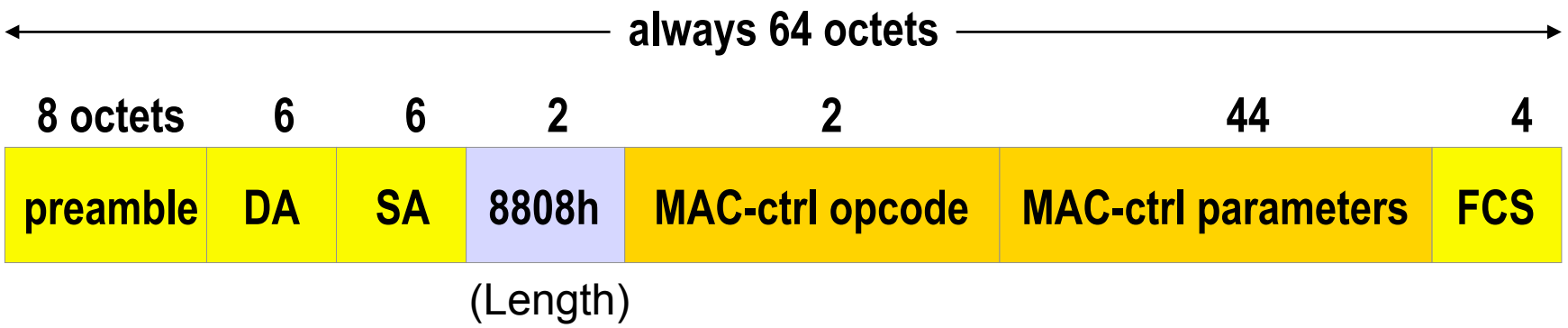
- **Would limit network diameter to 20-25 meters (Gigabit Ethernet)**
- **Solutions**
  - Frame Bursting
  - Carrier Extension
- **No GE-Hubs available on the market today ➔ forget it!**
- **No CSMA/CD defined for 10GE (!)**

# Flow Control

- **Speed-requirements for switches are very high**
  - Especially in full duplex operation also powerful switches can not avoid buffer overflow
  - Earlier, high traffic caused collisions and CSMA/CD interrupted the transmission in these situations, now high traffic is normal
- **L4 flow control (e.g. TCP) between end-systems is not efficient enough for a LAN**
  - switches should be involved to avoid buffer overflow
- **Therefore a MAC based (L2) flow control is specified**
  - MAC control frame with the Pause command

# MAC Control Frame

- **Identified among other frames by setting length field = 8808 hex**




MAC ctrl opcode ..... defines function of control frame

MAC- trl parameters .... control parameter data; always filled up to 44 bytes, by using zero bytes if necessary

- **Currently only the "pause" function is available (opcode 0x0001)**

# The Pause Command 1

- **On receiving the pause command**
  - Station stops sending normal frames for a given time which is specified in the MAC-control parameter field
- **This pause time is a multiple of the slot time**
  - 4096 bit-times when using Gigabit Ethernet or 512 bit-times with conventional 802.3
- **Paused station waits**
  - Until pause time expires or an additional MAC-control frame arrives with pause time = 0
  - Note: paused stations are still allowed to send MAC-control-frames (to avoid blocking of LAN)

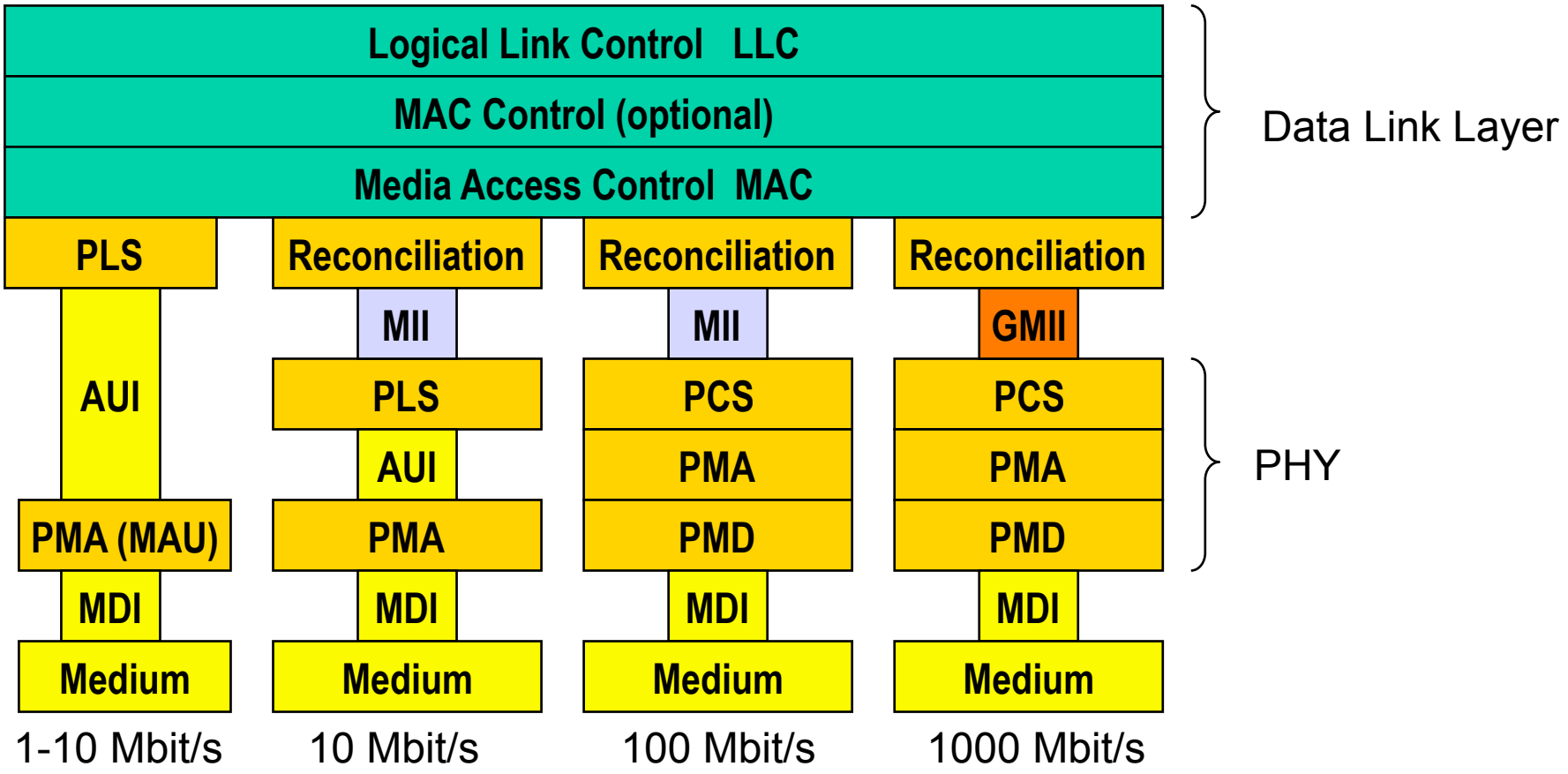
# The Pause Command 2

- **Destination address is either**
  - Address of destination station or
  - Broadcast address or
  - Special multicast address 01-80-C2-00-00-01
- **The special multicast address prevents bridges to transfer associated pause-frames to not concerned network segments**
  - Hence flow-control (with pause commands) affects only the own segment

# Demand for Higher Speed

- **Higher data rates need more sophisticated coding**
  - 10 Mbit/s Ethernet: Manchester coding
  - Fast Ethernet (100 Mbit/s): 4B/5B block code
  - Gigabit Ethernet 1000 Mbit/s): 8B/10B block code
- **New implementations should be backwards-compatible**
  - Old physical layer signaling interface (PLS), represented by AUI, was not suitable for new coding technologies
- **AUI has been replaced**
  - MII (Media Independent Interface) for Fast Ethernet
  - GMII for Gigabit Ethernet

# New Physical Sublayers




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

# PHY Sublayers (1)

- **Physical Layer Signaling (PLS) serves as abstraction layer between MAC and PHY**
- **PLS provides**
  - Data encoding/decoding (Manchester)
  - Translation between MAC and PHY
  - Attachment Unit Interface (AUI) to connect with PMA
- **Several new coding techniques demands for a Media Independent Interface (MII)**
- **Today coding is done through an media-dependent Physical Coding Sublayer (PCS) below the MII**

# PHY Sublayers (2)

- **PLS has been replaced with the Reconciliation sublayer**
  - Reconciliation layer transforms old MAC PLS-primitives into MII control signals
- **MII / GMII serves as an interface between MAC and PHY**
  - Hides coding issues from the MAC layer
  - MII: often a mechanical connector for a wire; GMII is an interface specification between MAC-chip and PHY-chip upon a circuit board
  - One independent specification for all physical media
  - Supports several data rates (10/100/1000 Mbits/s)
  - MII:4 bit / GMII: 8 bit parallel transmission channels to the physical layer

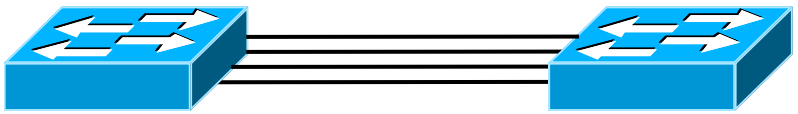
# PHY Sublayers (3)

- **Physical Coding Sublayer (PCS)**
  - Encapsulates MAC-frame between special PCS delimiters
  - E.g. 4B/5B or 8B/10B encoding
  - Appends idle symbols
- **Physical Medium Attachment (PMA)**
  - Interface between PCS and PMD
  - (de) serializes data for PMD (PCS)
- **Physical Medium Dependent (PMD)**
  - Serial transmission of the code groups
  - Specification of the various connectors (MDI)

# Today:

- **No collisions ➔ no distance limitations !**
- **Gigabit Ethernet becomes WAN technology !**
  - Over 100 km link span already
- **Combine several links to "Etherchannels"**
  - Link Aggregation Control Protocol (LACP, IEEE 802.3ad)
  - Cisco proprietary: Port Aggregation Protocol (PAgP), FEC, GEC
- **Trend moves towards layer 3 switching**
  - High amount of today's traffic goes beyond the border of the LAN
  - Routing stop broadcast domains, enable load balancing and decrease network traffic
  - "**Route if you can bridge if you must**"


**1 Gbit/s or even 10 Gbit/s long reach connection !!!**



**Ether Channel**

**Note: Spanning Tree regards this as one logical link!**
**=> Load balancing!**

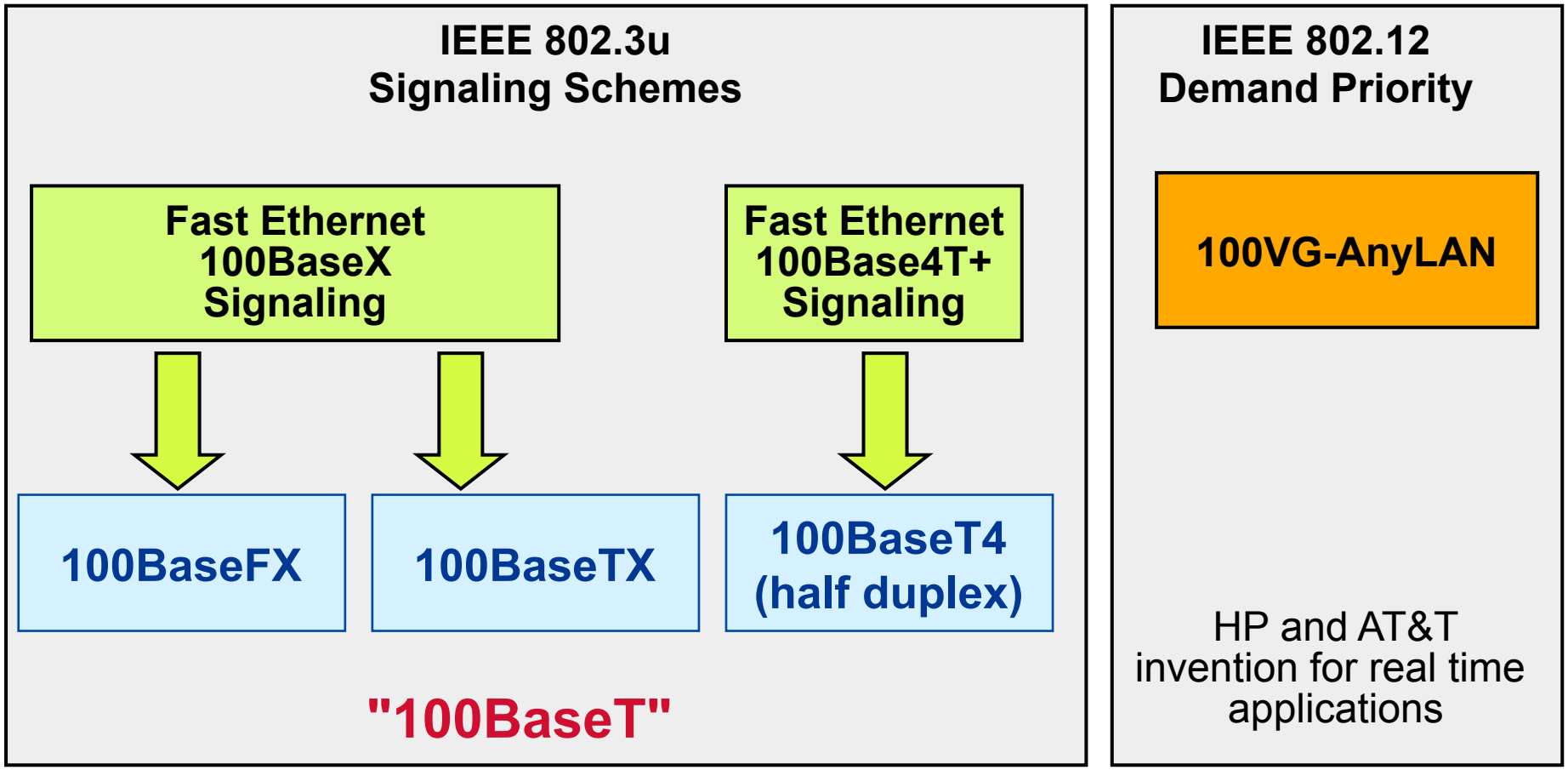
# Agenda

- **Ethernet Evolution**
- **VLAN**
- **High Speed Ethernet**
  - Introduction
  - Fast Ethernet
  - Gigabit Ethernet
  - 10 Gigabit Ethernet

# 100 Mbit/s Ethernet

- **Access method disagreement split 100 Mbit/s LAN development into two branches:**
  - Fast Ethernet - IEEE-802.3u (today 802.3-2008)
  - 100VG-AnyLAN - IEEE-802.12 (disappeared)
- **Fast Ethernet was designed as 100 Mbit/s and backwards-compatible 10Mbit/s Ethernet**
  - CSMA/CD but also
  - Full-duplex connections (collision free)
- **Network diameter based on collision window requirement (512 bit times)**
  - Reduced by factor 10
  - e.g. 250m compared with 2500m at 10 Mbit/s

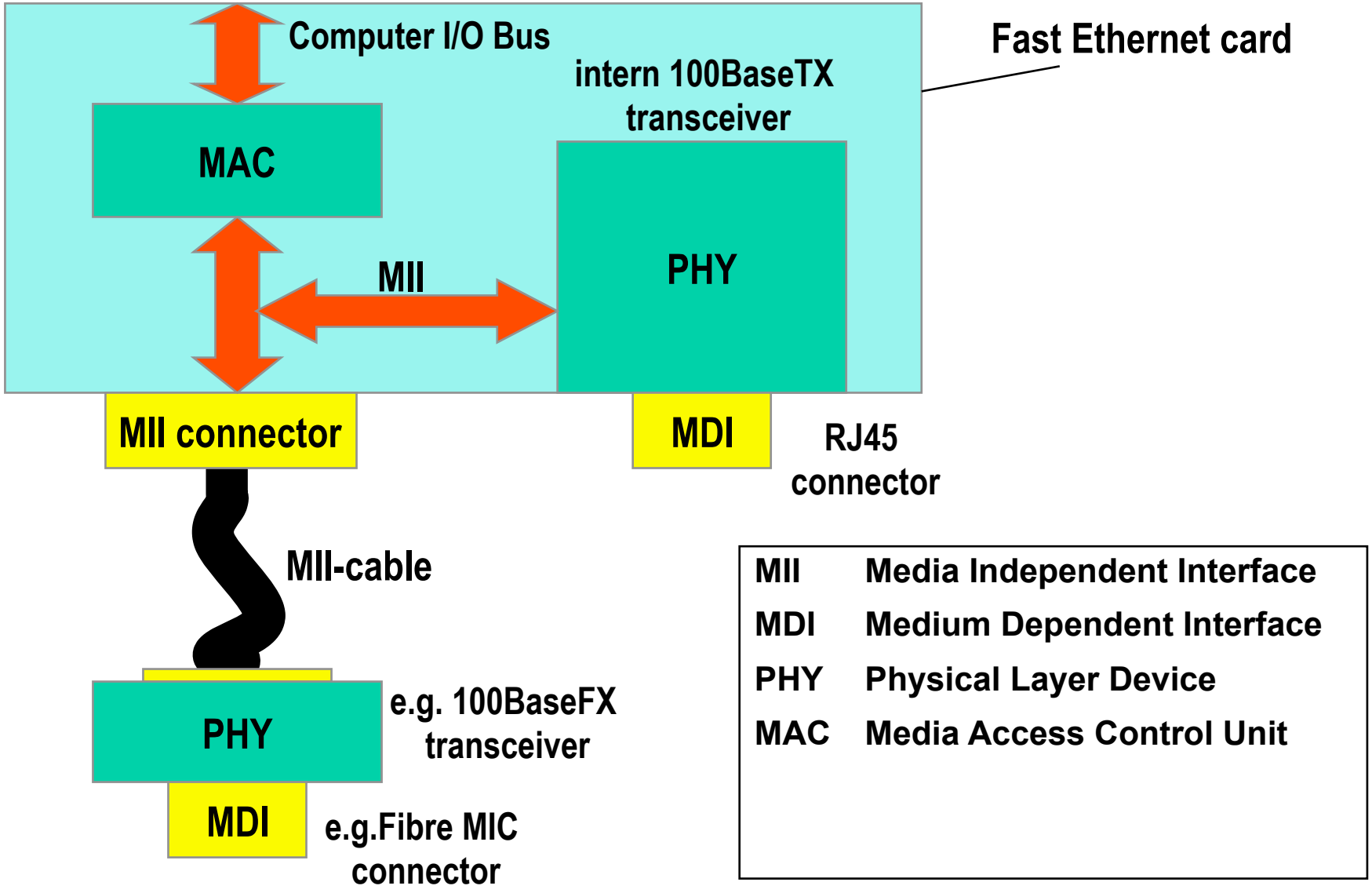
# 100 Mbit Ethernet Overview



# Fast Ethernet

- **AUI has been replaced with the Media Independent Interface (MII)**
  - New coding (4B/5B, 8B/6T, PAM 5x5) and bandwidth constrains demand for a redesigned abstraction layer
- **MII defines a generic 100BaseT interface**
  - Allows utilization of a 100BaseTX, 100BaseFX, 100BaseT4 or a 100BaseT2 transceiver
    - On-board or cable-connector with
    - 20 shielded, symmetrically twisted wire pairs -> 40 poles
    - One additional main-shield
    - 68 Ohm impedance; 2.5 ns maximal delay
    - 50 cm maximal length

# Typical Fashion

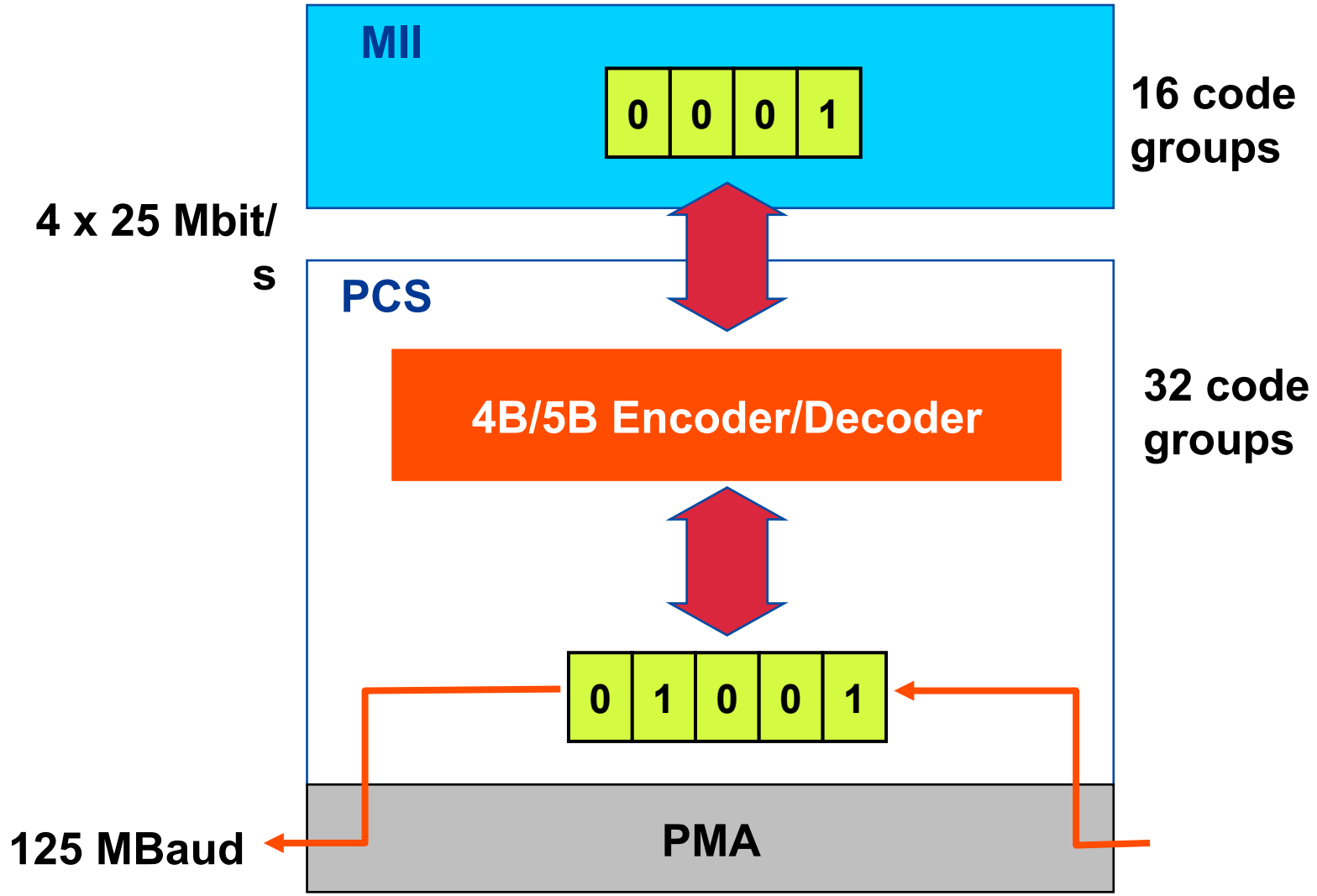



| | |
|---|---|
| MII | Media Independent Interface |
| MDI | Medium Dependent Interface |
| PHY | Physical Layer Device |
| MAC | Media Access Control Unit |

# 100BaseX Signaling – 4B/5BCoding

- **Each 4-bit group is encoded by a 5 bit run-length limited "code-group”**
  - Code groups lean upon FDDI-4B/5B codes
  - Some additional code groups are used for signaling purposes; remaining code groups are violation symbols -> easy error detection
  - Groups determinate maximal number of transmitted zeros or ones in a row -> easy clock synchronization
    - Keeps DC component below 10%
- **Code groups are transmitted using NRZI-encoding**
  - Code efficiency: 4/5 = 100/125 = 80% (Manchestercode only 50 %)

# 4B/5B Coding


MII
0 0 0 1
16 code groups
4 x 25 Mbit/s
PCS
4B/5B Encoder/Decoder
32 code groups
0 1 0 0 1
125 MBaud
PMA

# Code Group Table

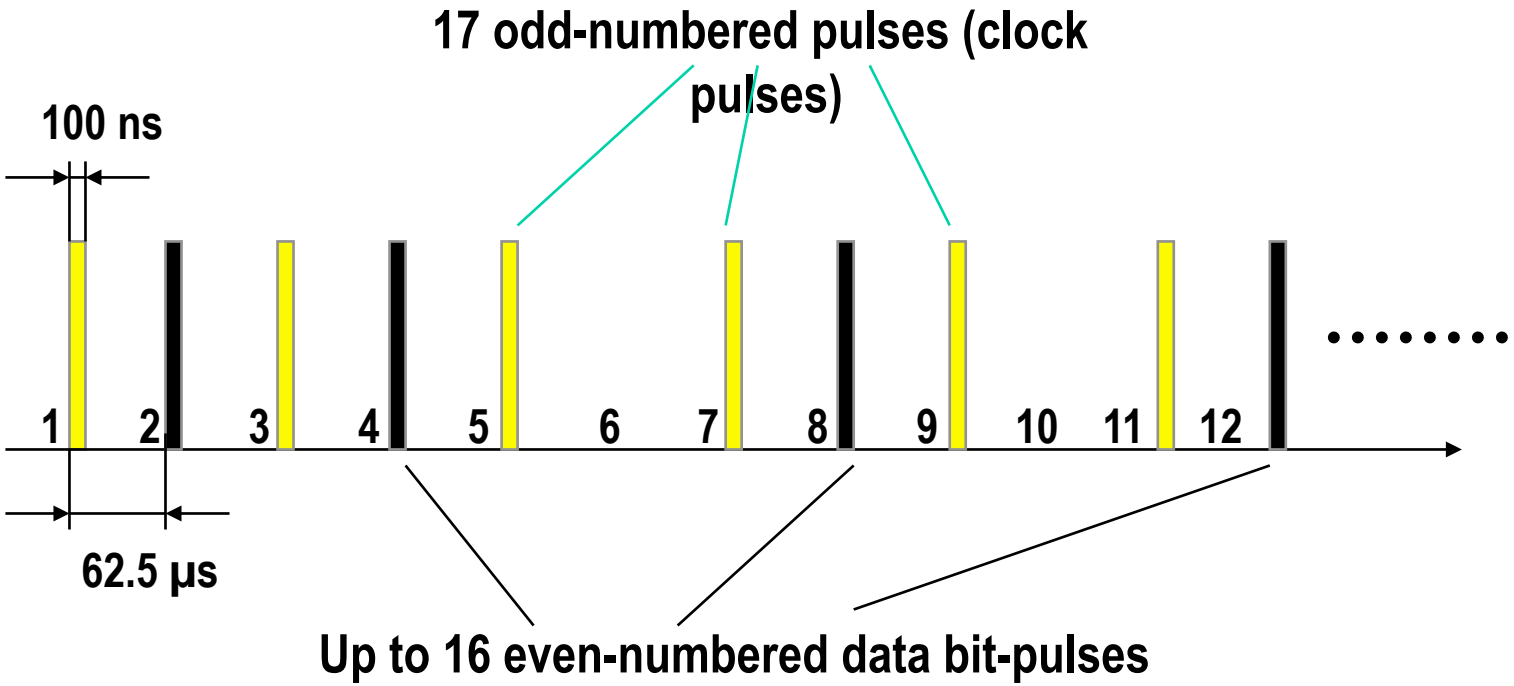
| | PCS code-group | name | MII group | |
|---|---|---|---|---|
| DATA | 11110 | 0 | 0000 | |
| | 01001 | 1 | 0001 | |
| | 10100 | 2 | 0010 | |
| | 10101 | 3 | 0011 | |
| | 01010 | 4 | 0100 | |
| | 01011 | 5 | 0101 | |
| | 01110 | 6 | 0110 | |
| | 01111 | 7 | 0111 | |
| | 10010 | 8 | 1000 | |
| | 10011 | 9 | 1001 | |
| | 10110 | A | 1010 | |
| | 10111 | B | 1011 | |
| | 11010 | C | 1100 | |
| | 11011 | D | 1101 | |
| | 11100 | E | 1110 | |
| | 11101 | F | 1111 | |
| Control | 11111 | I | undefined | Idle pattern between streams |
| | 11000 | J | 0101 | Start of Stream Delimiter (1st part) |
| | 10001 | K | 0101 | Start of Stream Delimiter (2nd part) |
| | 01101 | T | undefined | End of Stream Delimiter (1st part) |
| | 00111 | R | undefined | End of Stream Delimiter (2nd part) |
| | 00100 | H | undefined | signals receiving errors |

Remaining code groups are not valid (triggers error detection)

# Autonegotiation (1)

- **Enables each two Ethernet devices to exchange information about their capabilities**
  - Signal rate, CSMA/CD, half- or full-duplex
- **Modern Ethernet NICs send bursts of so called**
  - Fast-Link-Pulses (FLP) for autonegotiation signaling
  - Each FLP burst represents a 16 bit word
- **FLP**
  - Consists of 17-33 so called Normal-Link-Pulses (NLPs)
  - NLP are used for testing link-integrity
    - NLP technique is used in 10BaseT to check the link state (green LED)
    - 10 Mbit/s LAN devices send every 16.8 ms a 100ns lasting NLP, no signal on the wire means disconnected

# FLP Burst Coding




**= 1 1 0 1 0 1 ....**

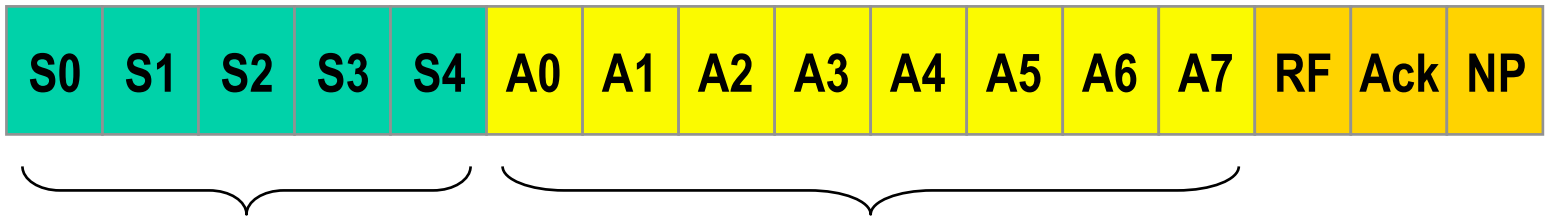
# Autonegotiation (2)

- **FLP-bursts are only sent on connection-establishments**
- **100BaseT stations recognizes 10 Mbit/s stations by receiving a single NLP only**
- **Two 100BaseT stations analyze their FLP-bursts and investigate their largest common set of features**
- **Last frames are sent 3 times -> other station responds with acknowledge-bit set**
- **Negotiated messages are sent 6-8 times**
  - FLP- session stops here

# Autonegotiation (3)

- **The first FLP-burst contains the base-link codeword**
- **By setting the NP bit a sender can transmit several "next-pages"**
  - Next-pages contain additional information about the vendor, device-type and other technical data
- **Two kinds of next-pages**
  - Message-pages (predefined codewords)
  - Unformatted-pages (vendor-defined codewords)
- **After reaching the last acknowledgement of this FLP-session, the negotiated link-codeword is sent 6-8 times**

# Base Page




provides selection of up to 32 different message types; currently only 2 selector codes available:

10000....IEEE 802.3

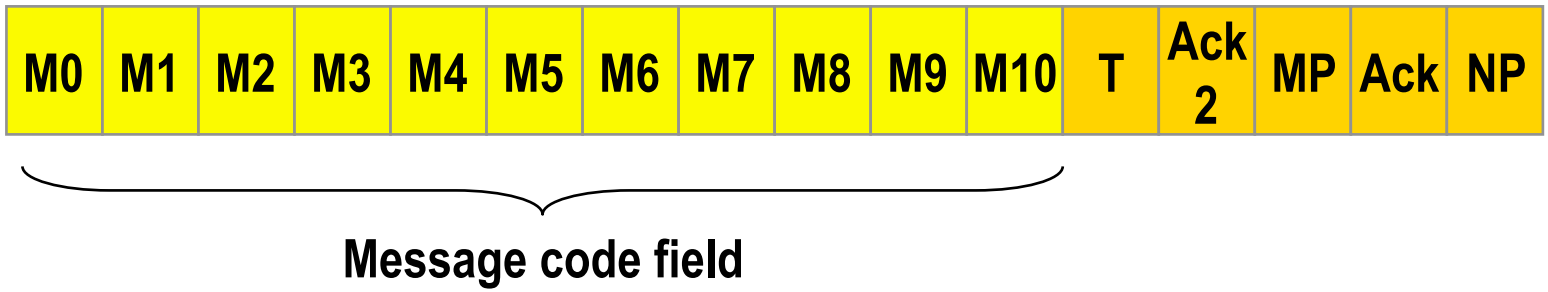
01000....IEEE 802.9

(ISLAN-16T)

(ISO-Ethernet)

| Bit | Technology |
|---|---|
| A0 | 10BaseT |
| A1 | 10BaseT-full duplex |
| A2 | 100BaseTx |
| A3 | 100BaseTx-full duplex |
| A4 | 100BaseT4 |
| A5 | Pause operation for full duplex links |
| A6 | reserved |
| A7 | reserved |

# Next-Pages Codeword

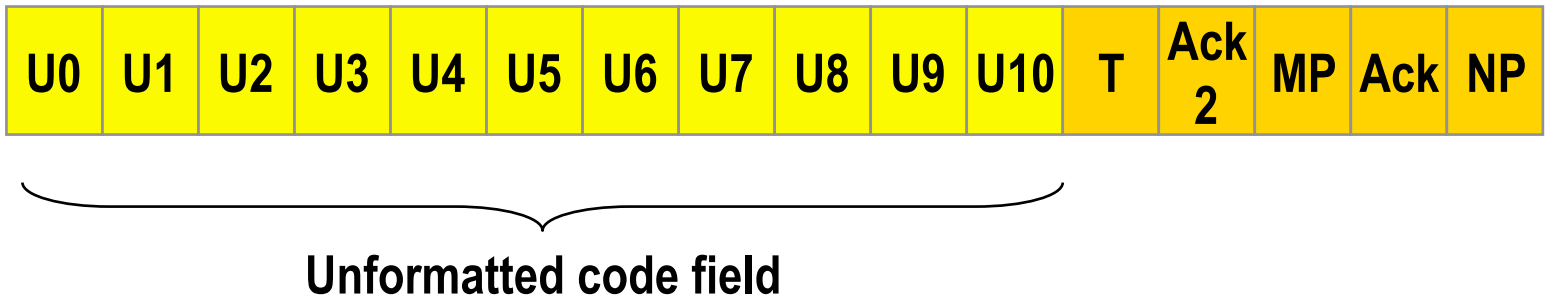



**Examples:**

**10000000000 ....null message, station has no further information to send**

**01000000000 ....next page contains technology ability information**

**10100000000 ....next 4 pages contain Organizationally Unique Identifier (OUI) information**



# Signaling Types (1)

- **Three signaling types :**
  - 100BaseX:
    - refers to either the 100BaseTX or 100BaseFX specification
  - 100BaseT4+
- **100BaseX**
  - combines the CSMA/CD MAC with the FDDI Physical Medium Dependent layer (PMD)
  - 4B/5B code
  - allows  full duplex operation on link

# Signaling Types (2)

- **100BaseT4+**
  - allows half duplex operation only
  - 8B6T code
  - Uses 4 pairs of wires; one pair for collision detection, three pair for data transmission
  - One unidirectional pair is used for sending only and two bi-directional pairs for both sending and receiving
  - Same pinout as 10BaseT specification
  - Transmit on pin 1 and 2, receive on 3 and 6; bi-directional on 4 and 5; bi-directional on 7 and 8

# 100BaseTX and 100BaseFX

- **100BaseTX:**
  - 125 MBaud symbol rate, full duplex, binary encoding
  - 2 pair Cat 5 unshielded twisted pair (UTP) or 2 pair STP or type 1 STP
  - RJ45 connector; same pinout as in 10BaseT (transmit on 1 and 2, receive on 3 and 6)
- **100BaseFX:**
  - 125 MBaud symbol rate, full duplex, binary encoding
  - Two-strand (transmit and receive) 50/125 or 62.5/125-µm multimode fiber-optic cable
  - SC connector, straight-tip (ST) connector, or media independent connector (MIC)

# 100BaseT4

- **100BaseT4:**
  - 25 MBaud, half duplex, ternary encoding
  - Cat3 or better, needs all 4 pairs installed
  - 200 m maximal network diameter
  - maximal 2 hubs

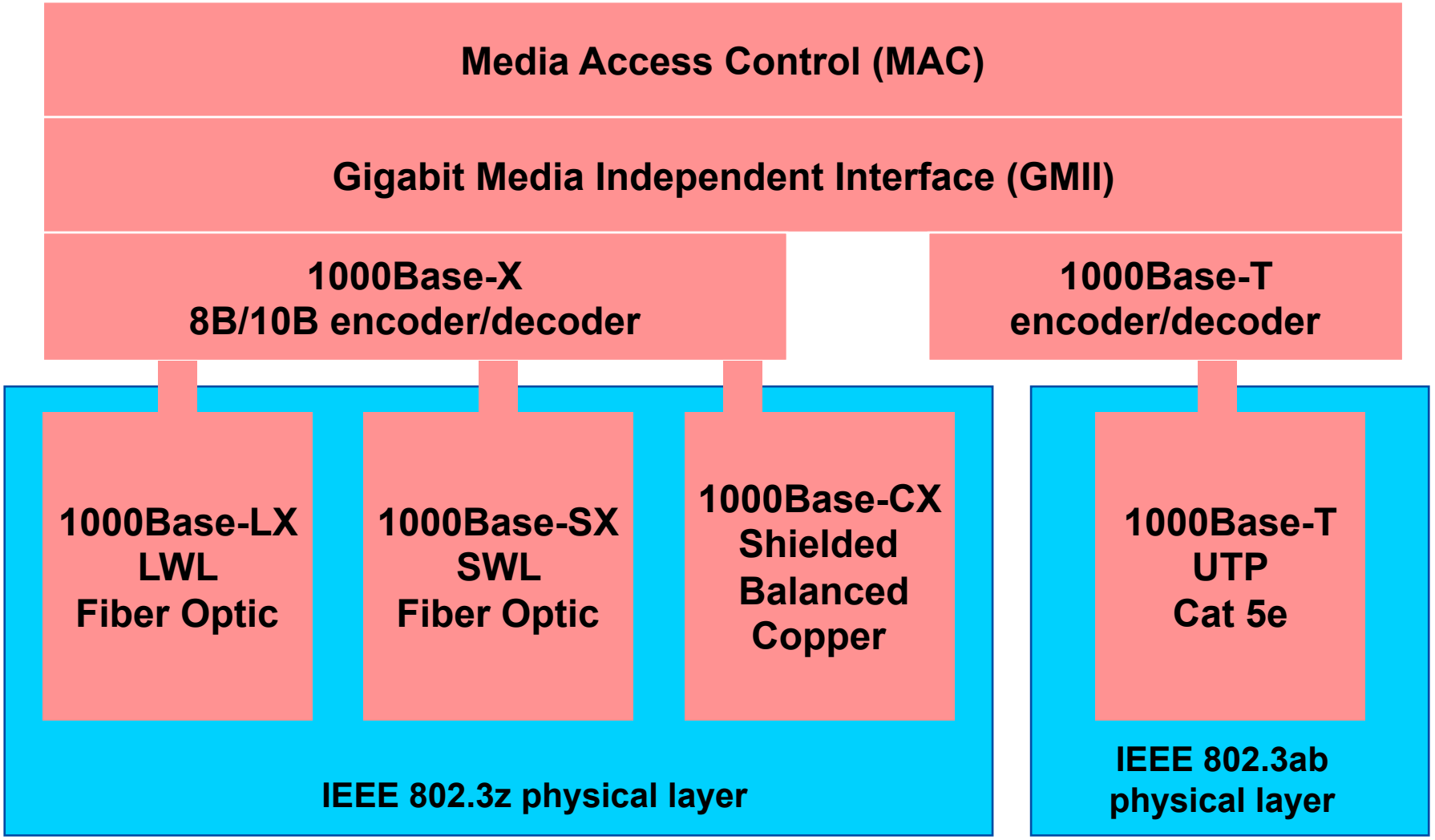
# Agenda

- **Ethernet Evolution**
- **VLAN**
- **High Speed Ethernet**
  - Introduction
  - Fast Ethernet
  - Gigabit Ethernet
  - 10 Gigabit Ethernet

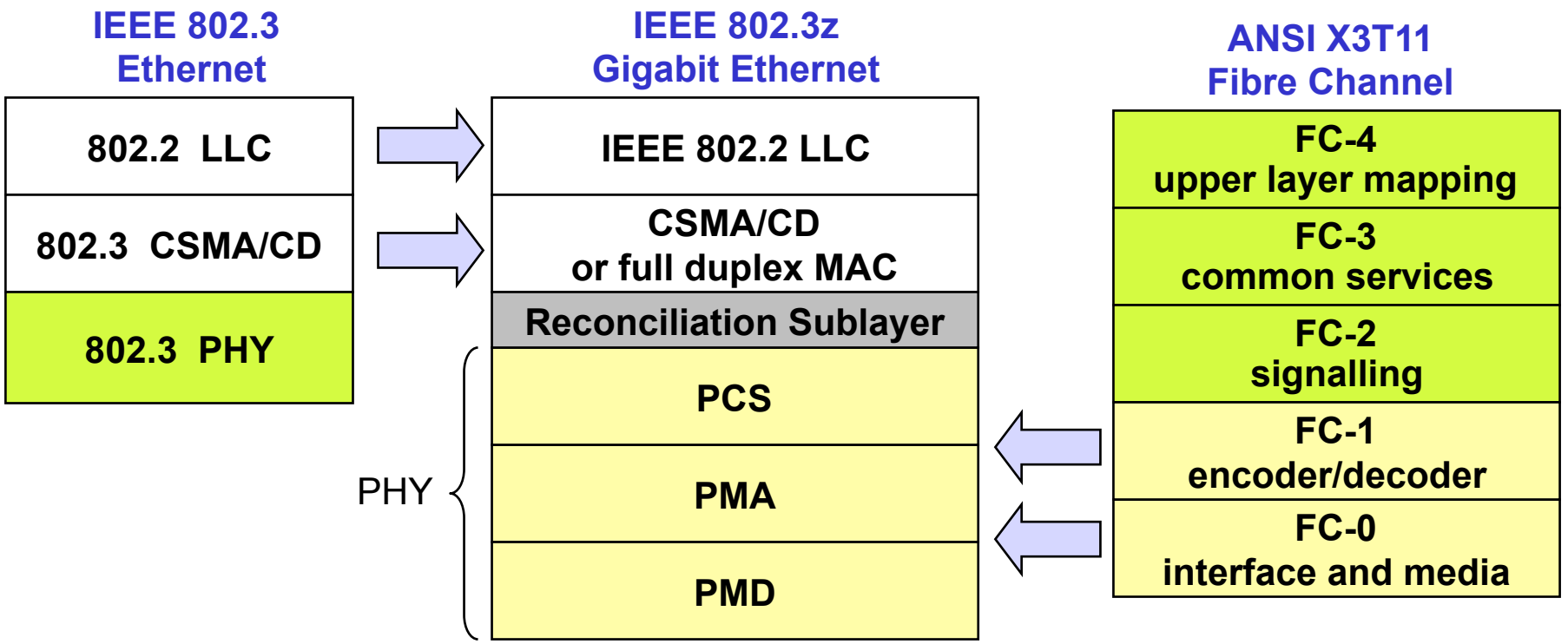
# Gigabit-Ethernet: IEEE-802.3z / IEEE802.3ab

- **Easy integration into existing 802.3 LAN configurations**
  - Because of backward compatibility
  - Access methods: CSMA/CD or full duplex
  - Autonegotiation and flow control
  - IEEE-802.3z/802.3ab are now part of IEEE 802.3-2008
- **Backbone technology**
  - GE link as WAN transmission technique
  - Reaches 70 km length using fibre optics

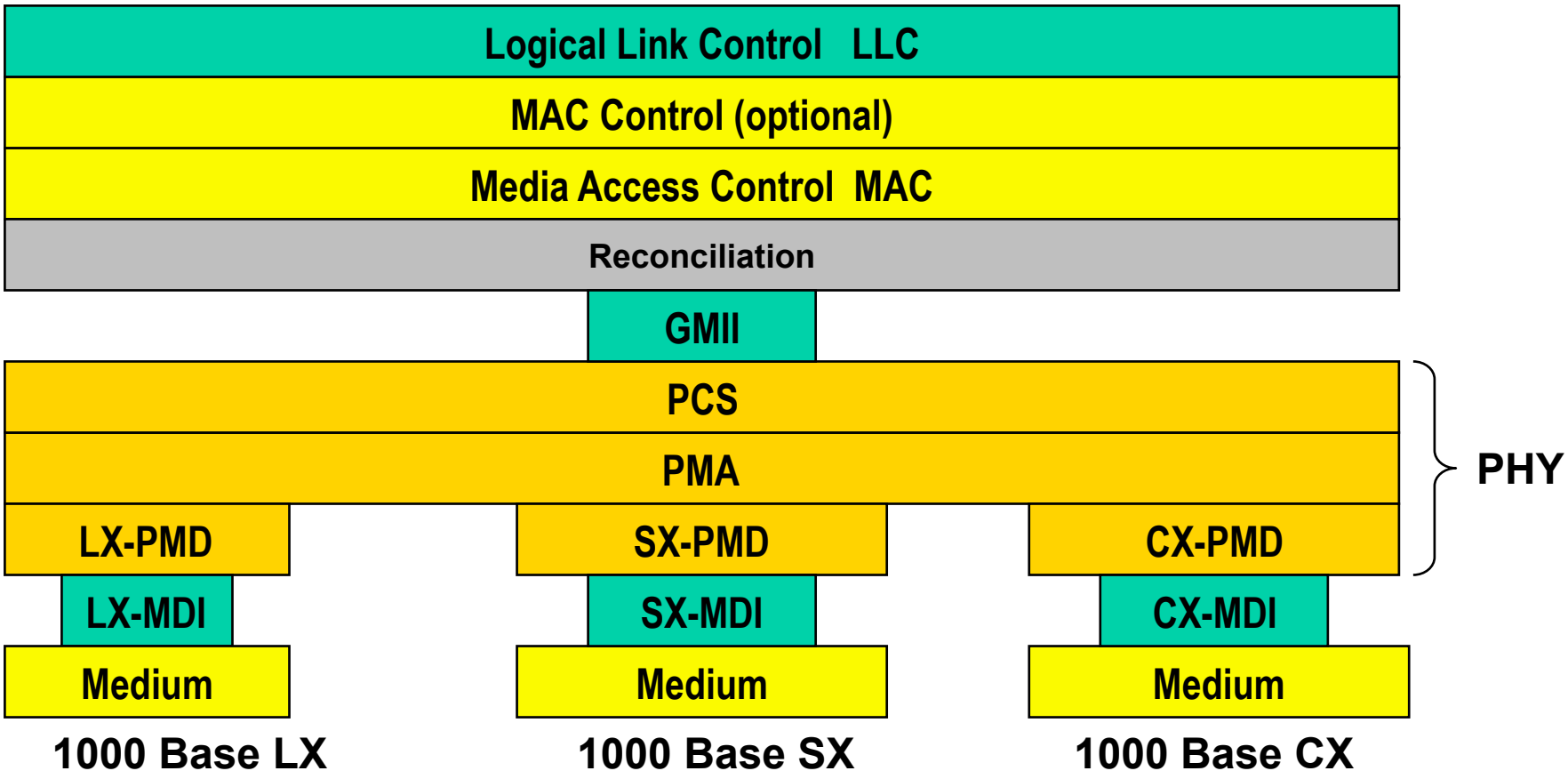
# Gigabit Ethernet



# GE Signaling



# Physical Sublayers

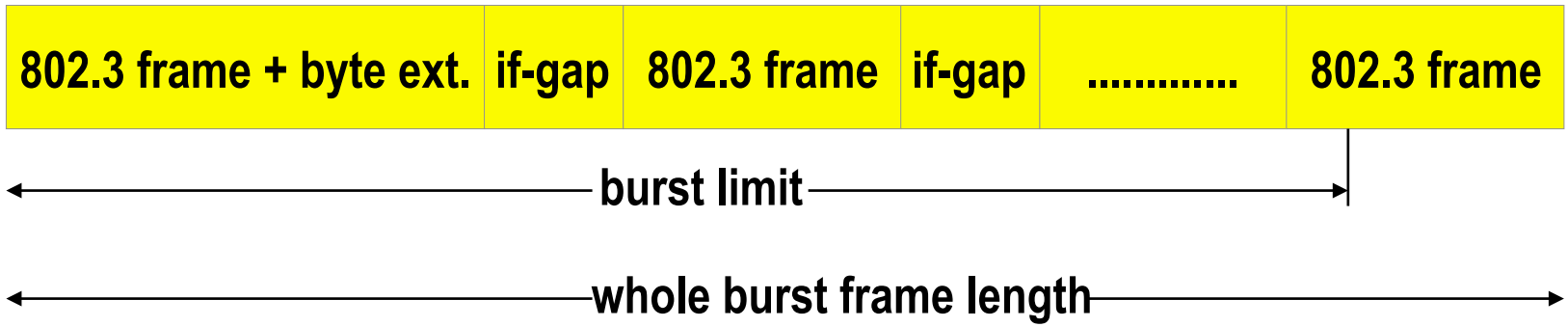


# CSMA/CD Restrictions (Half Duplex Mode)

- **The conventional collision detection mechanism CSMA/CD**
  - Limits the net diameter to 20m!
- **Solutions to increase the maximal net expansion:**
  - Carrier Extension:
    - Extension bytes appended to (and removed from) the Ethernet frame by the physical layer
    - Frame exists a longer period of time on the medium
  - Frame Bursting:
    - To minimize the extension bytes overhead, station may chain several frames together and transmit them at once ("burst").
  - With both methods the minimal frame length is increased from 512 to 4096 bits

# Frame Bursting

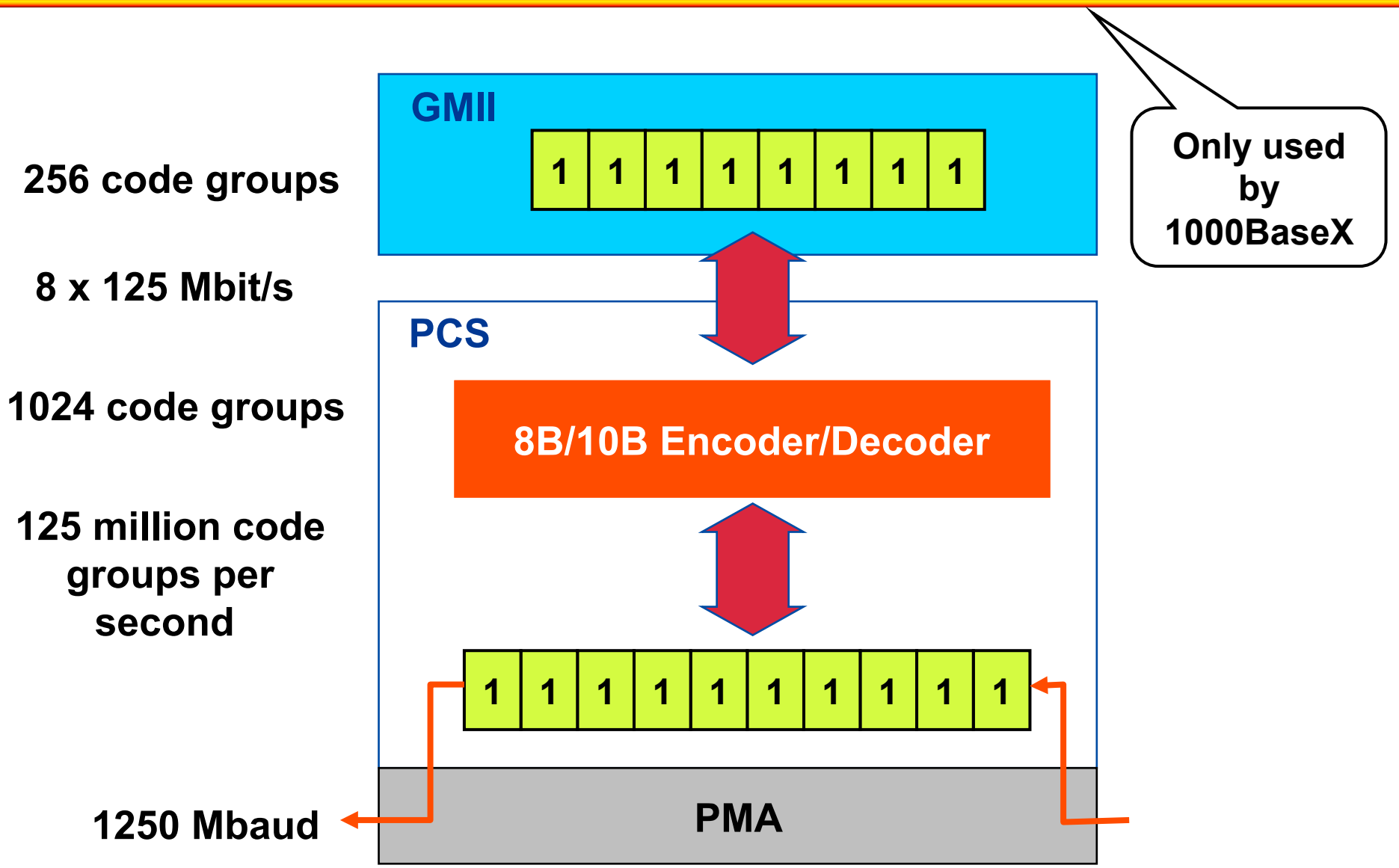
- **Station may chain frames up to 8192 bytes (=burst limit)**
  - Also may finish the transmission of the last frame even beyond the burst limit
- **So the whole burst frame length must not exceed 8192+1518 bytes**
  - Incl. interframe gap of 0.096 µs = 12 bytes



# 1000BaseX Coding

- **8B/10B block encoding: each 8-bit group encoded by a 10 bit "code-group" (symbol)**
  - Half of the code-group space is used for data transfer
  - Some code groups are used for signaling purposes
  - Remaining code groups are violation symbols
    - -> Easy error detection
  - Groups determine the maximal number of transmitted zeros or ones in a 10 bit symbol
    - -> Easy clock signal detection (bit synchronization)
  - No baselinewander (DC balanced)
    - lacking DC balance would result in data-dependent heating of lasers which increases the error rate
  - Code efficiency: 8/10 = 1000/1250 = 80%

# 8B/10B Coding (1)



# 8B/10B Coding (2)

- **Each GMII 8 bit group (data) can be represented by an associated pair of 10 bit code groups**
  - Each pair has exactly 10 ones and 10 zeros in sum
- **Sender toggles Running Disparity flag (RD) to remember which code group to be sent for the next data-octet**
- **Hence, only non-symmetric code groups need a compensating code group**
  - symmetric code groups already have equal number of ones and zeros

# 8B/10B Coding (3)

- **Code groups which are not registered in the code-table are considered as code-violation**
  - These code groups are selected to enable detection of line errors with high probability
- **256 data and 12 control code-group-pairs are defined**
- **Control-code-groups are used independently or in combination with data-code-groups**
  - Autonegotiation is performed by using such control-code-groups
  - Therefore no kind of NLP or FLP bursts for that purpose

# 8B/10B Coding (4)

- **Control-code-groups are classified by "ordered sets" after their usage:**
  - Configuration C  for autonegotiation
  - Idle I  used between packets
  - Encapsulation:
    - R for separating burst frames
    - S as start of packet delimiter
    - T as end of packet delimiter
    - V for error propagation

# 1000BaseX

- **Two different wavelengths supported**
- **Full duplex only**
  - 1000Base-SX: short wave, 850 nm MMF, up to 550 m length
  - 1000Base-LX: long wave, 1300 nm MMF or SMF, up to 5 km length
- **1000Base-CX:**
  - Twinax Cable (high quality 150 Ohm balanced shielded copper cable)
  - About 25 m distance limit, DB-9 or the newer HSSDC connector

# 1000BaseT

- **1000Base-T defined by 802.3ab task force**
  - UTP uses all 4 line pairs simultaneously for duplex transmission!
    - Using echo-cancelling: receiver subtracts own signal
  - 5 level PAM coding
    - 4 levels encode 2 bits + extra level used for Forward Error Correction (FEC)
  - Signal rate: 4 x 125 Mbaud = 4 x 250Mbit/s data rate
    - Cat. 5 links, max 100 m; all 4pairs, cable must conform to the requirements of ANSI/TIA/EIA-568-A
  - Only 1 CSMA/CD repeater allowed in a collision domain
    - note: collision domains should be avoided

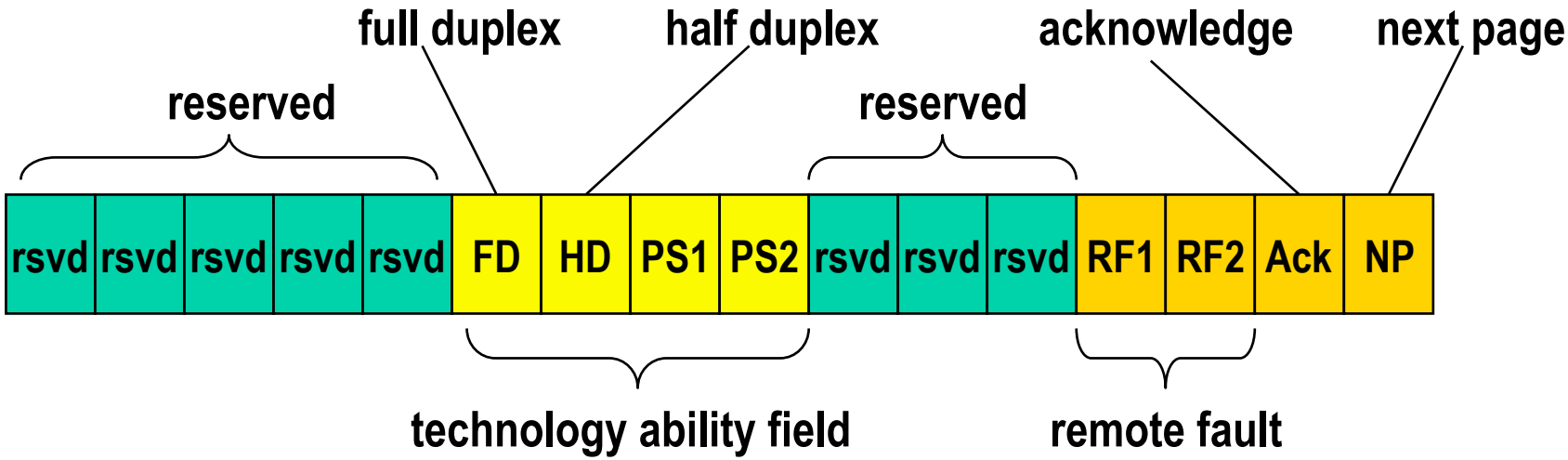
# Autonegotiation

- **Both 1000Base-X and 1000Base-T provide autonegotiation functions to determinate the**
  - Access mode (full duplex - half duplex)
  - Flow control mode
- **Additionally 1000Base-T can resolve the data rate**
  - Backward-compatibility with 10 Mbit/s and 100 Mbit/s
  - Also using FLP-burst sessions

# 1000BaseX Autonegotiation

- **1000Base-X autonegotiation uses normal (1000Base-X) signaling !**
  - "Ordered sets" of the 8B/10B code groups
  - No fast link pulses !
    - Autonegotiation had never been specified for traditional fiber-based Ethernet
    - So there is no need for backwards-compatibility
- **1000Base-X does not negotiate the data rate !**
    - Only gigabit speeds possible
- **1000Base-X autonegotiation resolves**
  - Half-duplex versus full-duplex operation
  - Flow control

# Base-Page

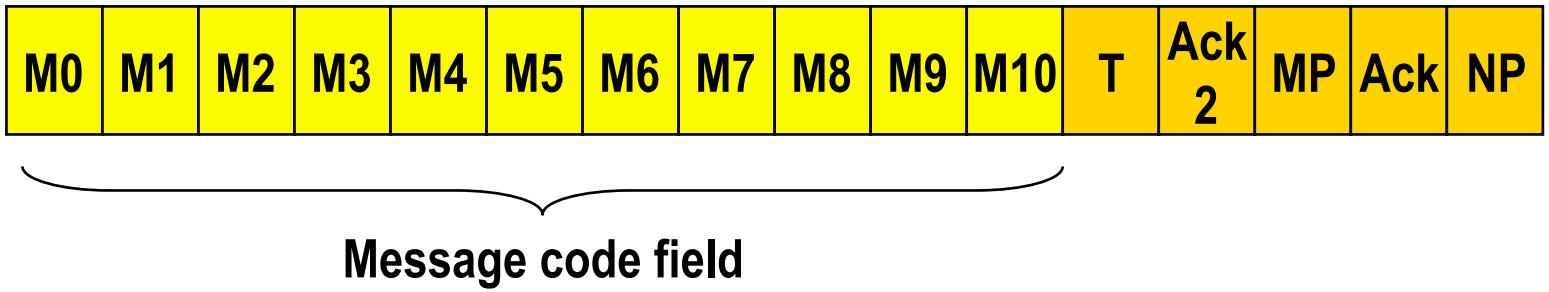



| PS1 | PS2 | description |
|---|---|---|
| 0 | 0 | no pause |
| 0 | 1 | asymmetrical pause |
| 1 | 0 | symmetrical pause |
| 1 | 1 | symmetrical and asymmetrical pause |

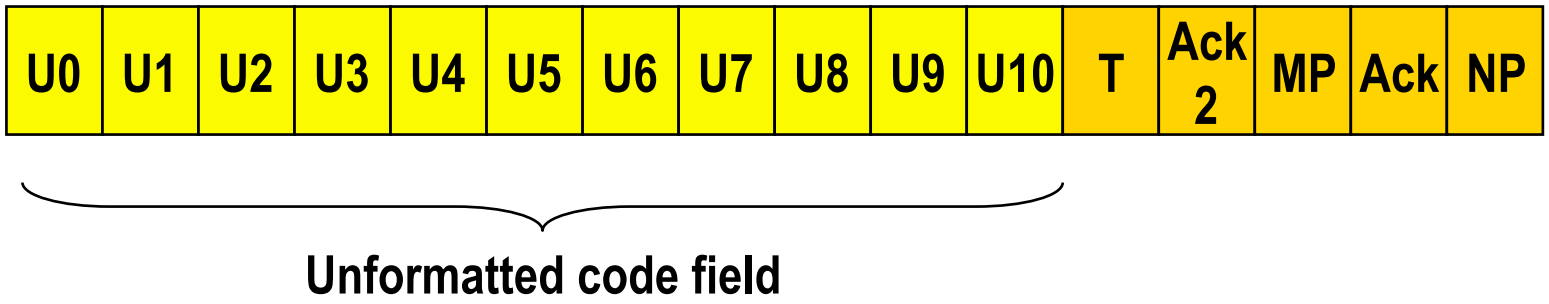
| RF1 | RF2 | description |
|---|---|---|
| 0 | 0 | no error |
| 0 | 1 | offline |
| 1 | 0 | connection error |
| 1 | 1 | autonegotiation error (no common capabilities) |

# Next-Pages

## Normal message-page (predefined codes)




## Vendor specific page (non predefined codes)



# 1000BaseT Autonegotiation

- Autonegotiation is only triggered when the station is powered on
- At first the stations expects Gigabit-Ethernet negotiation packets (replies)
- If none of them can be received, the 100Base-T fast link pulse technique is tried
- At last the station tries to detect 10Base-T stations using normal link pulses

# Agenda

- **Ethernet Evolution**
- **VLAN**
- **High Speed Ethernet**
  - Introduction
  - Fast Ethernet
  - Gigabit Ethernet
  - 10 Gigabit Ethernet

# 10 Gigabit Ethernet (IEEE 802.3ae)

- **Preserves Ethernet framing**
- **Maintains the minimum and maximum frame size of the 802.3 standard**
- **Supports only full-duplex operation**
  - CSMA/CD protocol was dropped
- **Focus on defining the physical layer**
  - Four new optical interfaces (PMD)
    - To operate at various distances on both single-mode and multi-mode fibers
  - Two families of physical layer specifications (PHY) for LAN and WAN support
  - Properties of the PHY defined in corresponding PCS
    - Encoding and decoding functions

# PMDs

- **10GBASE-L**
  - SM-fiber, 1300nm band, maximum distance 10km
- **10GBASE-E**
  - SM-fiber, 1550nm band, maximum distance 40km
- **10GBASE-S**
  - MM-fiber, 850nm band, maximum distance 26 – 82m
  - With laser-optimized MM up to 300m
- **10GBASE-LX4**
  - For SM- and MM-fiber, 1300nm
  - Array of four lasers each transmitting 3,125 Gbit/s and four receivers arranged in WDM (Wavelength-Division Multiplexing) fashion
  - Maximum distance 300m for legacy FDDI-grade MM-fiber
  - Maximum distance 10km for SM-fiber

# WAN PHY / LAN PHY and their PCS

- **LAN-PHY**
  - 10GBASE-X
  - 10GBASE-R
    - 64B/66B coding running at 10,3125 Gbit/s
- **WAN-PHY**
  - 10GBASE-W
    - 64B/66B encoded payload into SONET concatenated STS192c frame running at 9,953 Gbit/s
    - Adaptation of 10Gbit/s to run over traditional SDH links

# IEEE 802.3ae PMDs, PHYs, PCSs

| | | PCS | | |
|---|---|---|---|---|
| PMD | 10GBASE-E | 10GBASE-ER | | 10GBASE-EW |
| | 10GBASE-L | 10GBASE-LR | | 10GBASE-LW |
| | 10GBASE-S | 10GBASE-SR | | 10GBASE-SW |
| | 10GBASE-L4 | | 10GBASE-LX4 | |
| | | LAN PHY | | WAN PHY |

# 10 Gigabit Ethernet over Copper

- **IEEE 802.3ak defined in 2004**
  - 10GBASE-CX4
  - Four pairs of twin-axial copper wiring with IBX4 connector
  - Maximum distance of 15m
- **IEEE 802.3an defined in 2006**
  - 10GBASE-T
  - CAT6 UTP cabling with maximum distance of 55m to 100m
  - CAT7 cabling with maximum distance of 100m
- **Nowadays 802.3ab, 802.3ak, 802.3an are covered by the IEEE 802.3-2008 document**