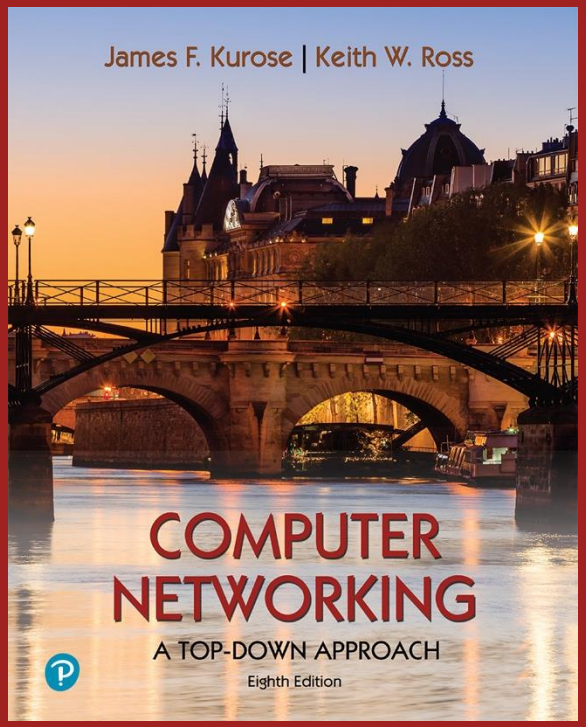
Communication Networks
W. Tavernier



Chapter 6
Data Link Layer

Computer Networking: A Top-Down Approach
8th Edition, 2020, Pearson,
James F. Kurose, Keith W. Ross

Link layer, LANs: outline

6.1 Introduction, services

6.2 Error detection, correction

6.3 Multiple access protocols

6.4 LANs

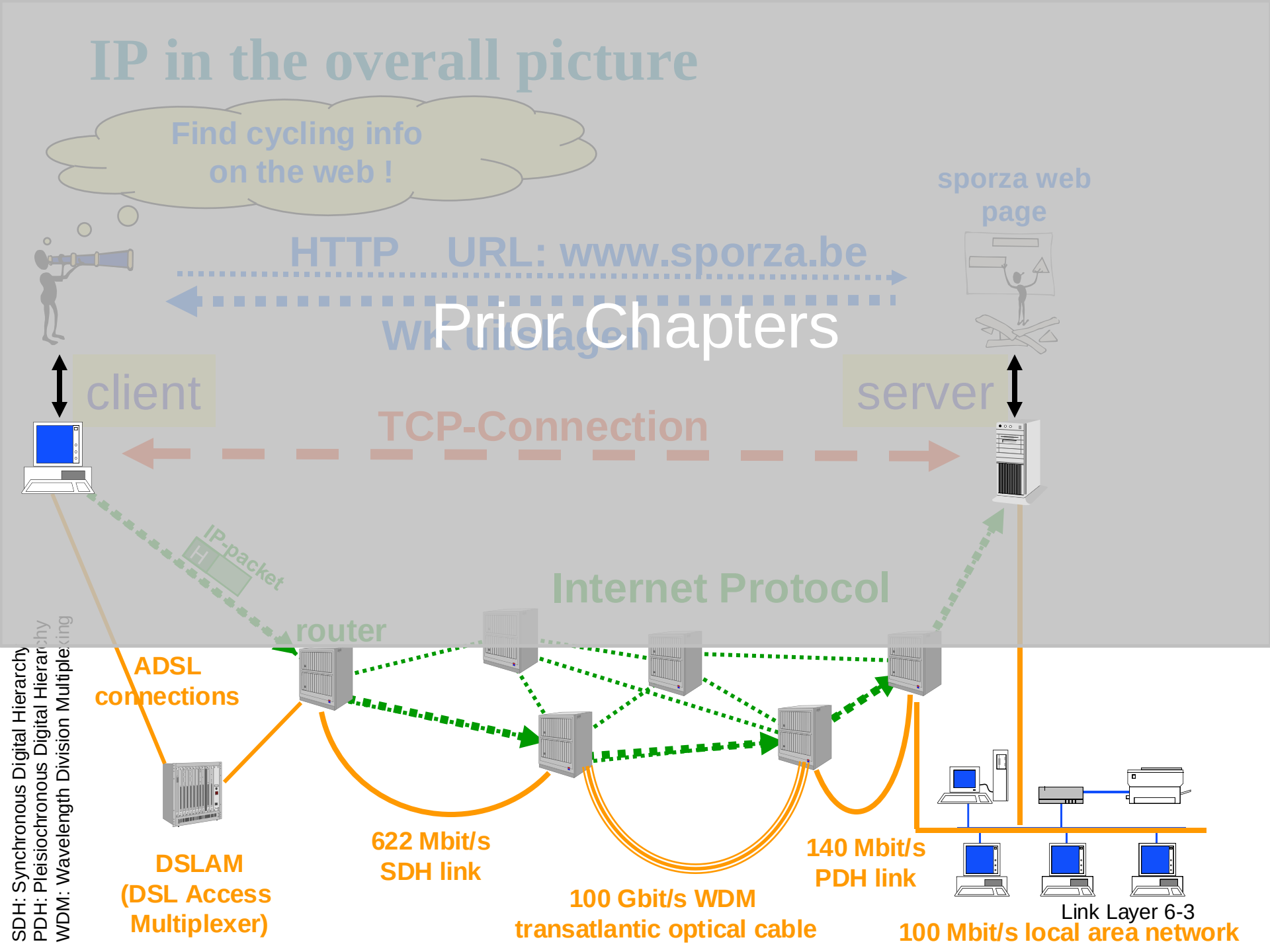
- addressing, ARP
- Ethernet
- switches
- VLANS

6.5 Link virtualization: MPLS

6.6 Data center networking

6.7 A day in the life of a web request

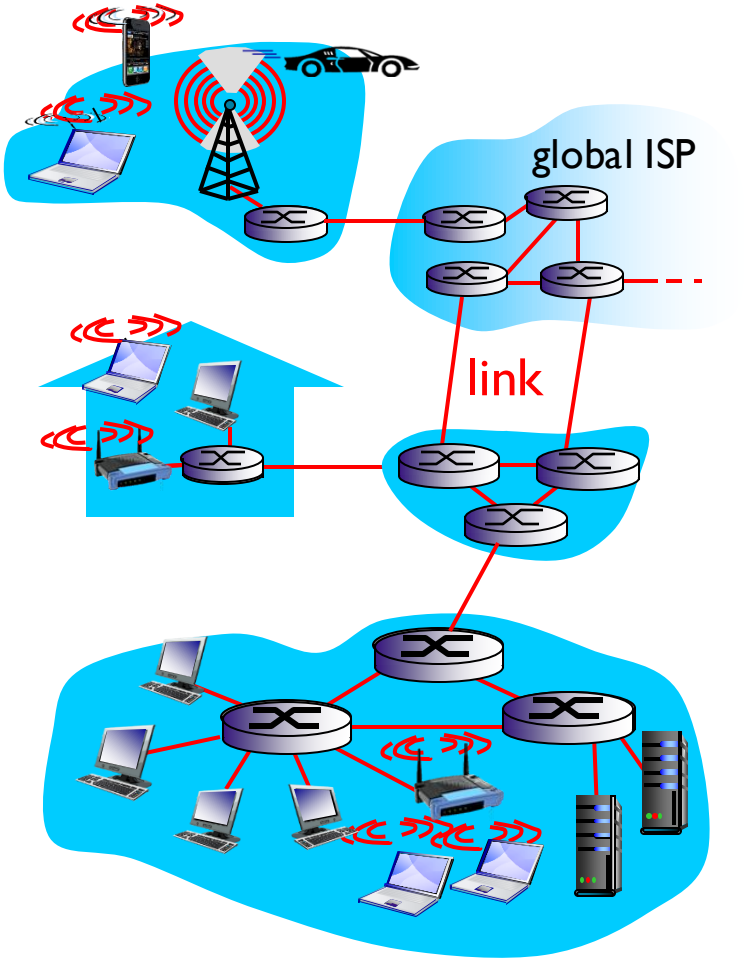
IP in the overall picture



Link Layer: Introduction

Terminology:

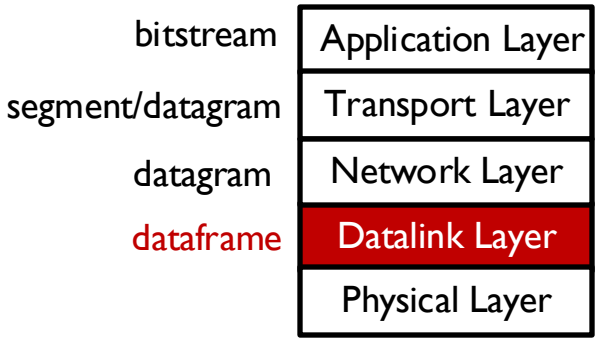
- hosts and routers are "***nodes***"
- communication channels that connect adjacent nodes along communication path are ***links***
  - wired links (point 2 point)
  - wireless links (multi-point)
  - LANs (multi-point)
- Layer 2-PDU is a ***frame***, encapsulates packet



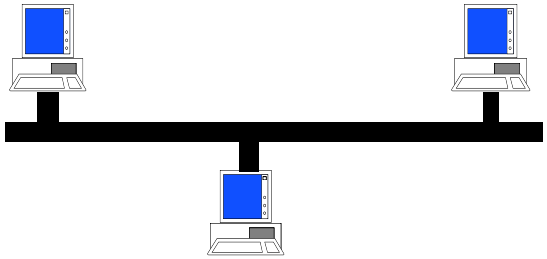
data-link layer (L2) has responsibility of transferring packets from one node to adjacent node over a link

Link Layer Services

- Framing:
  - en/de-capsulate packet into frame, adding header, trailer
  - header needs to indicate which is the upper-layer protocol (de/multiplexing)



- Medium access:
  - channel access if shared medium (MAC=Medium Access Protocol)
  - MAC addresses used to identify source and destination (different from IP address!)

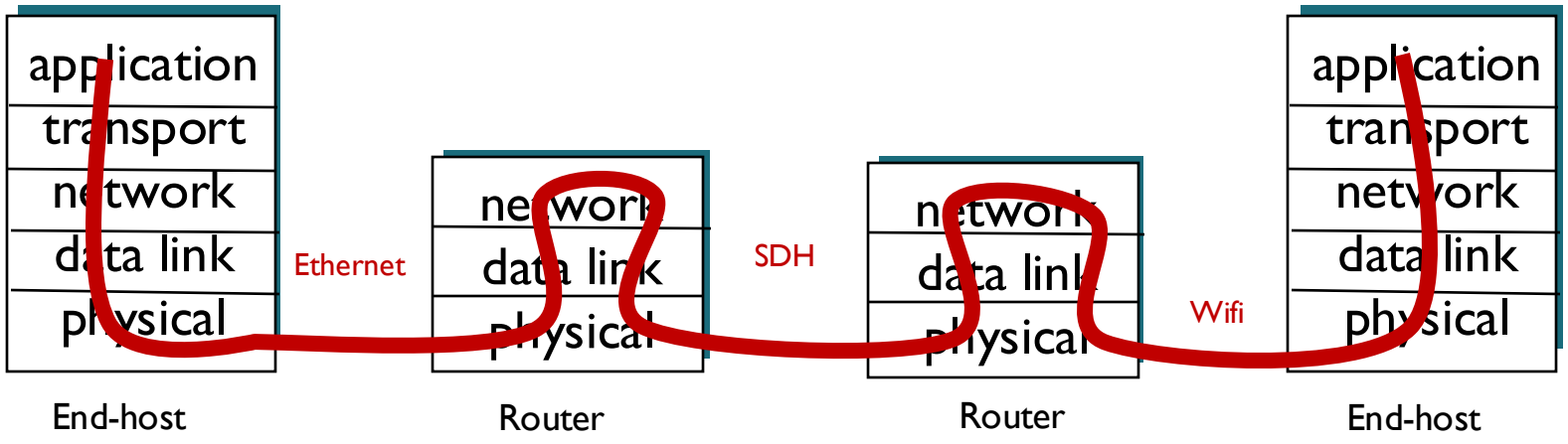


Link Layer Services (more)

- Reliable delivery between adjacent nodes (<> end-nodes in TCP)
  - seldom used on low bit error link (fiber, some twisted pair)
  - wireless links: high error rates
- Flow Control
  - pacing between adjacent sending and receiving nodes
- Error Detection & Correction (receiver side)
  - Detection: of errors caused by signal attenuation, noise -> signal to sender for retransmission
  - Correction: receiver identifies and corrects bit error(s) without resorting to retransmission
- Half-duplex and full-duplex
  - with half duplex, nodes at both ends of link can transmit, but not at same time

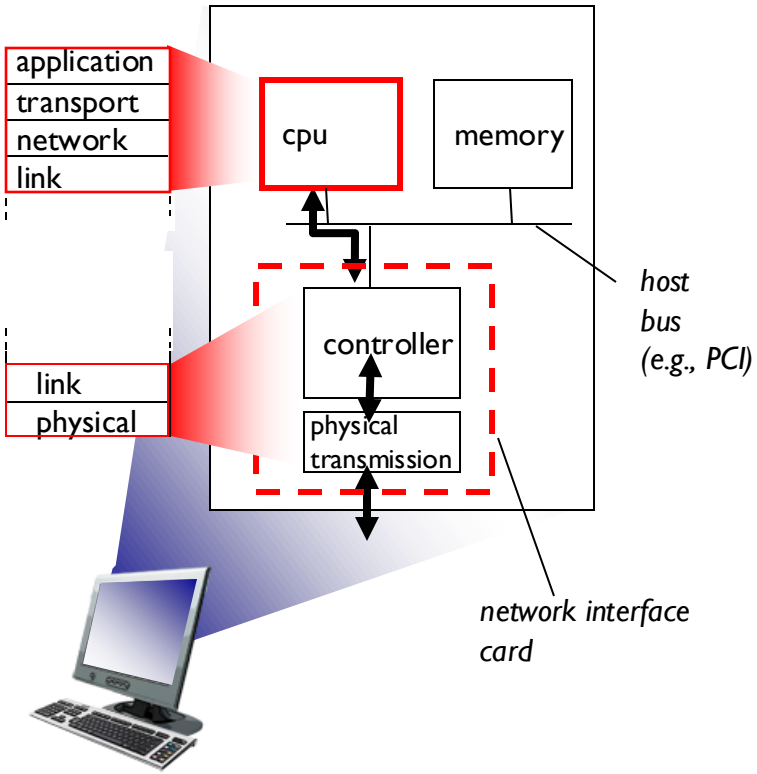
Link layer: context

- Data link layer connects two (or more) physically connected devices
  - Host-router
  - Router-router
  - Host-host
- Different links might use different data link protocols
  - IP layer is providing the internetworking
- Each datalink (L2) protocol provides different services
  - e.g., may or may not provide reliable data transfer over link

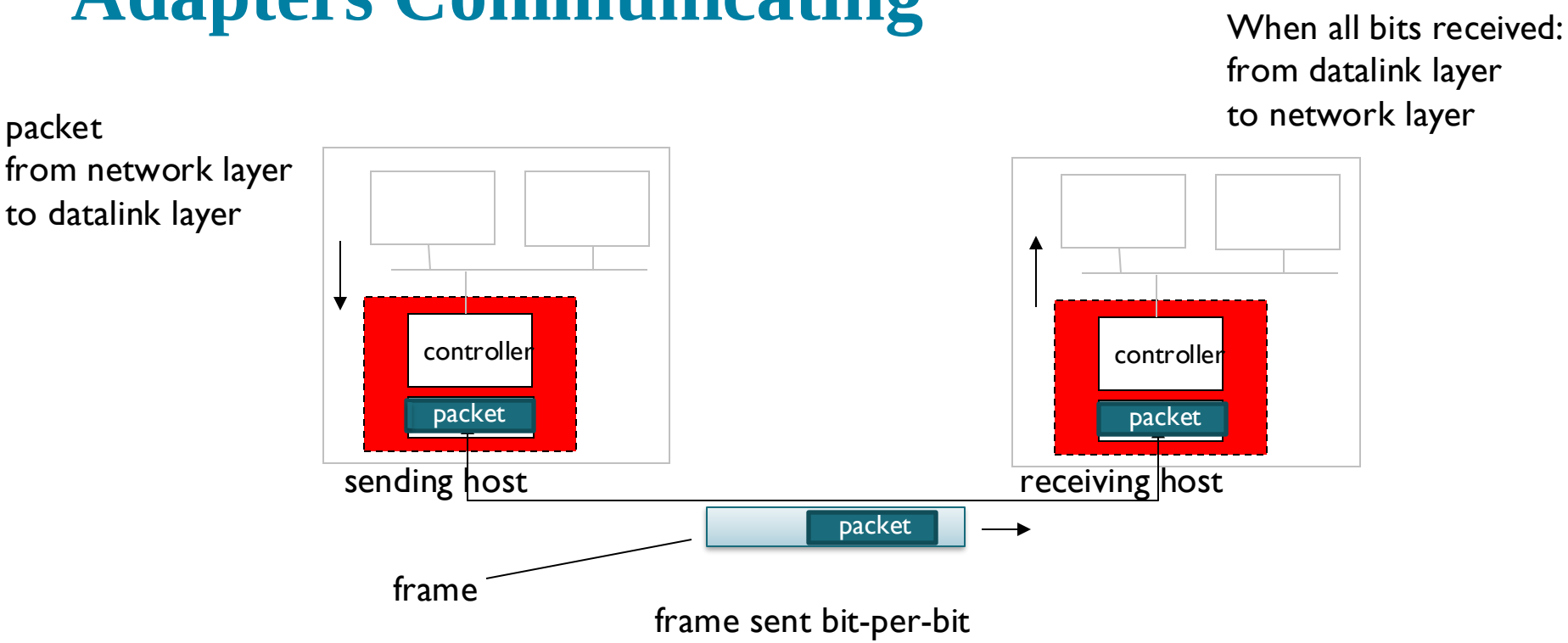


Where is the link layer implemented?

- in each and every *host + router*
- link layer implemented in "adapter" (*Network Interface Card* or *NIC*)
  - Ethernet card, 802.11 card, on board
  - implements link + physical layer
- attaches into host/router's system buses
- *combination of hardware, software, firmware*



Adapters Communicating

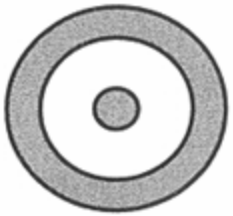


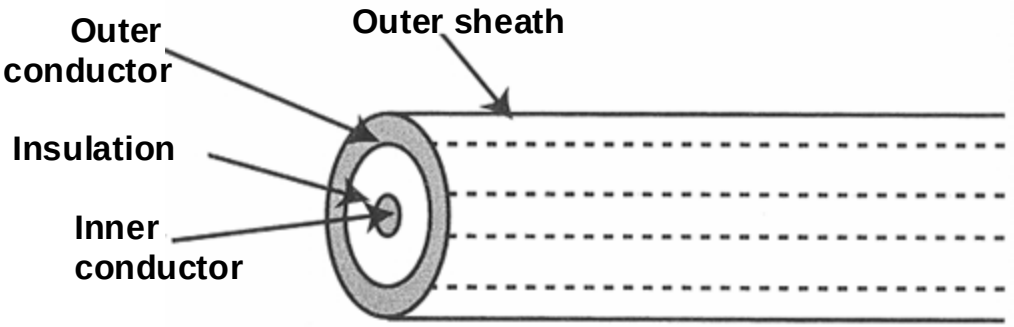
- sending side:
  - *encapsulates* packet in frame
  - *adds error checking bits*, reliable data transfer, flow control, etc.
- receiving side
  - *looks for errors*, reliable data transfer, flow control, etc
  - *extracts packet*, passes to upper layer at receiving side

Link Layer – Dependence on physical media

Twisted pair

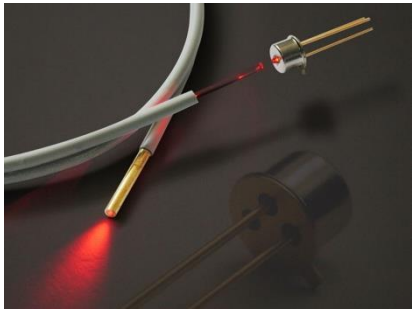
Coaxial cable

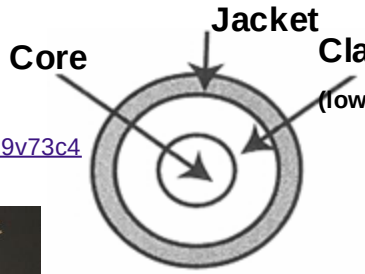


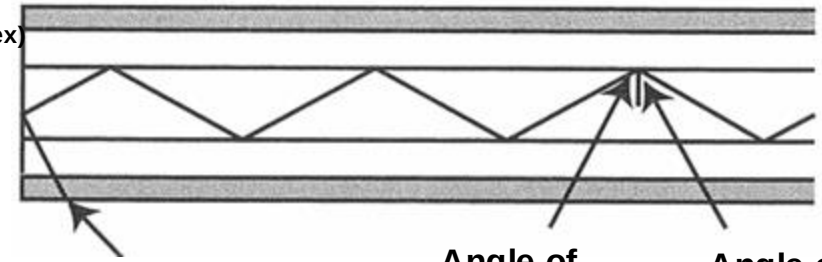


Optical fiber

https://www.youtube.com/watch?v=jZOg39v73c4







Light at less than critical angle is absorbed in jacket

Angle of incidence

Angle of reflection

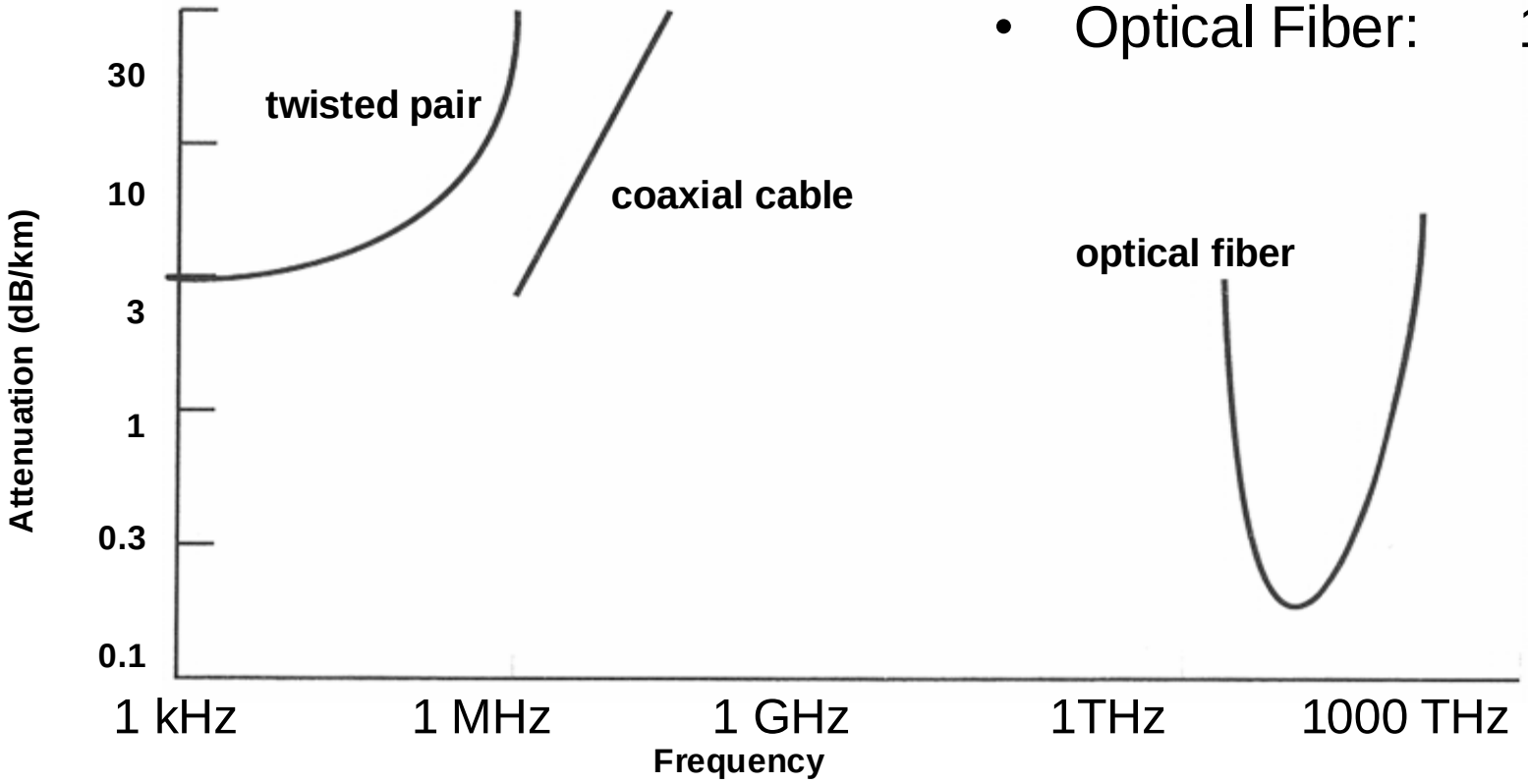
In netwerken zijn er drie belangrijke geleiders die gebruikt worden om informatie te verzen,den:
twisted pair – gebruikt om een huis te verbinden met het telefoonnetwork, maar ook in LAN netwerken wordt UTP (Universal Twisted Pair) vaak gebruikt.
coax kabel – gebruikt om een huis te verbinden met TV-distributie (CATV)
optische vezel – tot op heden vooral gebruikt binnen de core van ons internet of tussen telefooncentrales, maar ook steeds meer om huis te verbinden met ISP
Zonder geleider kan een signaal ook *door de lucht* verstuurd worden (bvb via 5G GSM).
Twisted pair *(TP)* draad bestaat uit twee geïsoleerde koperen draden die rond elkaar gedraaid zijn. Door die draaiing zijn de signalen op die draden minder gevoelig aan externe ruis. Meestal zijn een aantal twisted pairs gebundeld in een kabel.
Een ***coax kabel*** heeft een centrale kopergeleider, die omgeven wordt door een geleidend schild (koper, aluminium). Tussen de geleider en het schild wordt een isolator gebruikt. Het schild vormt een uiterst goede afscherming tegen externe ruis, het werkt immers als een kooi van Faraday. Tegelijk kan de bandbreedte op het medium verhoogd worden.
Een ***optische vezel*** of ***fiber*** bestaat uit een glasdraad, met een kern met een heel hoge weerkaatsing (refractieve index). Hierdoor wordt licht in de vezel geleid; door de interne reflectie ontsnapt quasi geen licht en is er dus ook quasi geen verzwakking (attenuatie) van het signaal.

Physical Layer – Attenuation of media

Attenuation in dB is 10 $\log_{10}$ (P_{in}/P_{out})
e.g. 10 dB = factor of 10 difference in power (W)

In real life:

- Twisted Pair: 10 MHz
- Coax Cable: 1 GHz
- Optical Fiber: 10 THz



Een signaal dat door een geleider wordt gestuurd (twisted pair, coax, fiber) zal zijn sterkte verliezen naarmate dit signaal verder moet gestuurd worden. Dit heet attenuatie en wordt uitgedrukt in dB/km.
Een waarde van 3 dB/km betekent bvb. dat de sterkte van het signaal gehalveerd wordt als het over een afstand van 1 km door de specifieke geleider verstuurd wordt. Als het over 3 km verstuurd dient te worden, vermindert de signaalsterkte dus met een factor 8.
De attenuatie die optreedt, is sterk afhankelijk van de gebruikte frequentie van het draagsignaal, zoals weergegeven in de figuur.
Die draagfrequentie bepaalt ook de maximum bitrate die kan behaald worden op de drager (theorema van Nyquist-Shannon). Hoe hoger de frequentie van het draagsignaal, hoe hoger de bitrate die kan verstuurd worden
We bemerken dat de attenuatie veel sterker is in een twisted pair dan in een coax kabel. Optische vezel heeft de laagste attenuatie voor signalen met een frequentie (of golflengte) die zich in het infrarode gebied bevinden.
Optische vezel heeft dus in se de grootste capaciteit om signalen te versturen.

Physical Layer – Medium Capacity Limit

What is the maximal bitrate we can transmit over a bandwidth limited channel (with noise)?

Shannon Capacity Limit:

$$C = B \log_2(1+SNR)$$

C = maximal capacity (in bits/s)
B = bandwidth of the transmission channel (in Hz)
SNR = signal power / noise power
$SNR_{dB} = 10 \log_{10}(SNR)$

E.g. classic telephony channel over twisted pair :
B = 4 kHz, SNR ≈ 1000 (30 dB)
$\Rightarrow C = 4000 \log_2(1001) \approx 4000 \times 10 = 40$ kbit/s

From the Shannon capacity limit, we can calculate the maximum bitrate C (bits/s) one can transport over a bandwidth limited channel with noise*.

$C = B \log_2(1+SNR)$

The bandwidth B is expressed in Hz. The noise is specified as a relative measure with respect to the signal strength : SNR or Signal to Noise Ratio = signal power/noise power. In many cases this is expressed in decibel (dB) : $SNR_{dB} = 10 \log_{10}$ (signal power/noise power). When the signal power is 1 and the noise intensity is 0.001, we have a SNR ratio of $1000=10^3$ or expressed in dB we have a SNR_{dB} of 30 dB (this is a typical value for telephone connections over twisted pair).

If we take an example of B=4 kHz and SNR_{dB}=30 dB, we obtain a maximum capacity of C = 40 kbit/s.

**Note that for a channel without noise the bitrate is in principle infinite : "you can send every second a signal with an amplitude specified with an infinite accuracy". For example at t=0 you send a sample with amplitude 1.384932829048321… (infinite number of digits) and at t=1 you send a next sample with amplitude 2.09473829103858392… and so on. This gives indeed an infinite transfer of information because you suppose that you can receive this signal with infinite accuracy (no noise is added). If there is a noise signal with amplitude e.g. 0.001, then the digits after 0.001 do not have any value anymore : due to the noise you do not know if the original signal was 1.384 or 1.385.*

Physical Layer – Encoding of information

encoding = representing information in a (digital) way such that it can be transmitted over a communication medium

Grade	Probability	Fixed-length Code	Variable-length Code
A	$1/3$	00	10
B	$1/2$	01	0
C	$1/12$	10	110
D	$1/12$	11	111

Fixed-length encoding for $BCBAAB$: 01 10 01 00 00 01 (12 bits)

sending 1000 grades always takes exactly 2000 bits

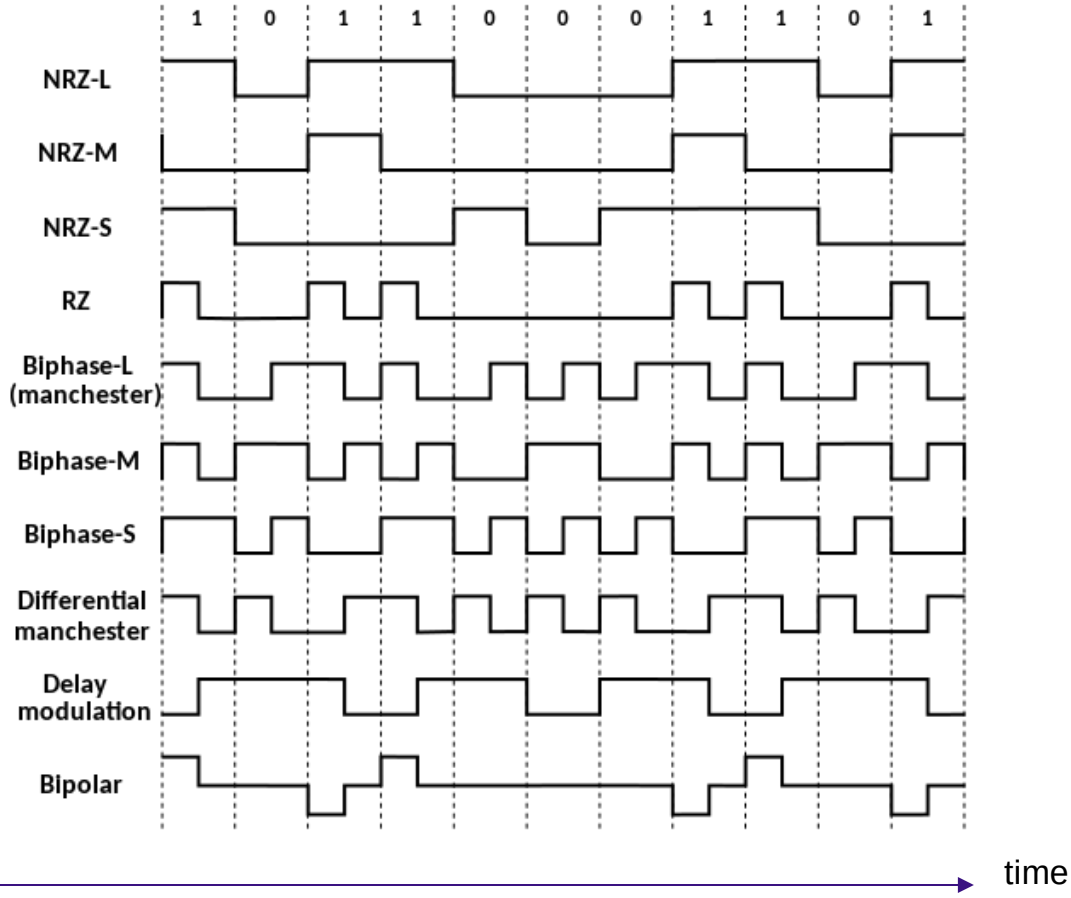
Variable-length encoding for $BCBAAB$: 0 110 0 10 10 0 (10 bits)

actual bit length depends on data

expected length = $1000[(\frac{1}{3})(2) + (\frac{1}{2})(1) + (\frac{1}{12})(3) + (\frac{1}{12})(3)] = 1000[1\frac{2}{3}] = 1666.7$ bits

Physical Layer – Encoding data into signals

Line code = pattern of voltage, current, or photons used to represent digital data as a signal



In telecommunication, a non-return-to-zero (NRZ) line code is a binary code in which ones are represented by one significant condition, usually a positive voltage, while zeros are represented by some other significant condition, usually a negative voltage, with no other neutral or rest condition.

An example of this is RS-232, where "one" is −12 V to −5 V and "zero" is +5 V to +12 V.

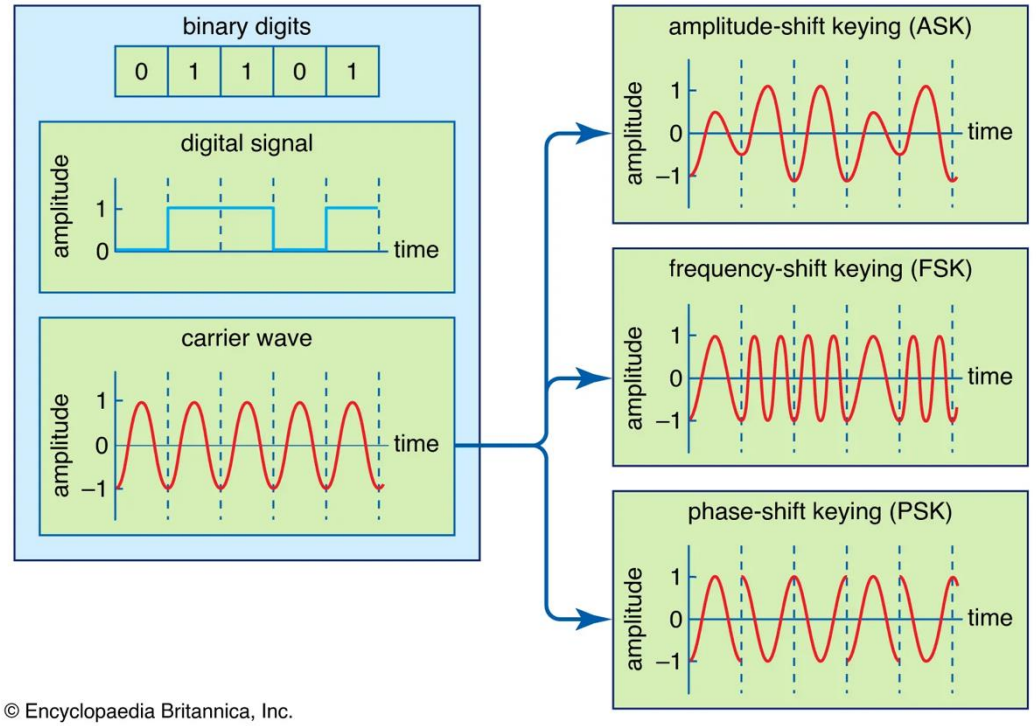
NRZMN = on-return-to-zero Mark serializer mapping {0: constant, 1: toggle}.
NRZS = Non-return-to-zero Space serializer mapping {0: toggle, 1: constant}.
…

Physical Layer – Signal Modulation

modulation = process of varying one or more properties of a periodic waveform (carrier wave) with a separate signal (digital signal)

Motivation:

- Decrease impact of **interference**
- Increased **range** (increasing power of carrier signal)
- Enable feasible **antenna sizes** (needs to be ¼ of amplitude of wave)
- More efficient **use of BW**: ability to multiplex (multiple data channels on 1 signal)



https://youtu.be/Iyzpt3bKTTI?si=YeRwR0IwNpaOY8qs

Link layer, LANs: outline

6.1 Introduction, services

6.2 Error detection, correction

6.3 Multiple access protocols

6.4 LANs

- addressing, ARP
- Ethernet
- switches
- VLANS

6.5 Link virtualization: MPLS

6.6 Data center networking

6.7 A day in the life of a web request

Multiple access links, protocols

two types of “links”:

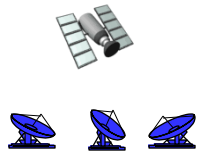
- point-to-point
  - point-to-point link between Ethernet switch, host
  - PPP for dial-up access

- broadcast (shared wire or medium)
  - old-school Ethernet
  - upstream HFC in cable-based access network
  - 802.11 wireless LAN, 4G/4G. satellite

shared wire (e.g., cabled Ethernet)

shared radio: 4G/5G

shared radio: WiFi



shared radio: satellite

Multiple access protocols

- single shared broadcast channel
- two or more simultaneous transmissions by nodes: interference
  - *collision* if node receives two or more signals at the same time

multiple access protocol

- distributed algorithm that determines how nodes share channel, i.e., determine when node can transmit
- communication about channel sharing must use channel itself!
  - no out-of-band channel for coordination

An IDEAL medium access control protocol

given: multiple access channel (MAC) of rate *R* bps

desiderata:

1. when one node wants to transmit, it can send at rate *R*.
2. when *M* nodes want to transmit, each can send at average rate *R/M*
3. fully decentralized:
   - no special node to coordinate transmissions
   - no synchronization of clocks, slots
4. simple

MAC protocols: taxonomy

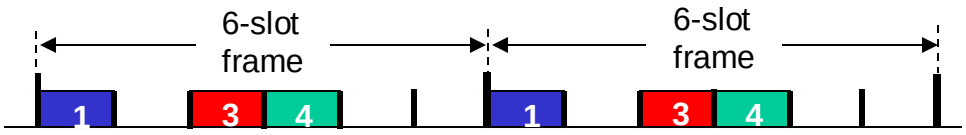
three broad classes:

- channel partitioning
  - divide channel into smaller “pieces” (time slots, frequency, code)
  - allocate piece to node for exclusive use
- random access
  - channel not divided, allow collisions
  - “recover” from collisions
- “taking turns”
  - nodes take turns, but nodes with more to send can take longer turns

Channel partitioning MAC: TDMA

TDMA: time division multiple access

- access to channel in “rounds”
- each station gets fixed length slot (length = packet transmission time) in each round
- unused slots go idle
- example: 6-station LAN, 1,3,4 have packets to send, slots 2,5,6 idle



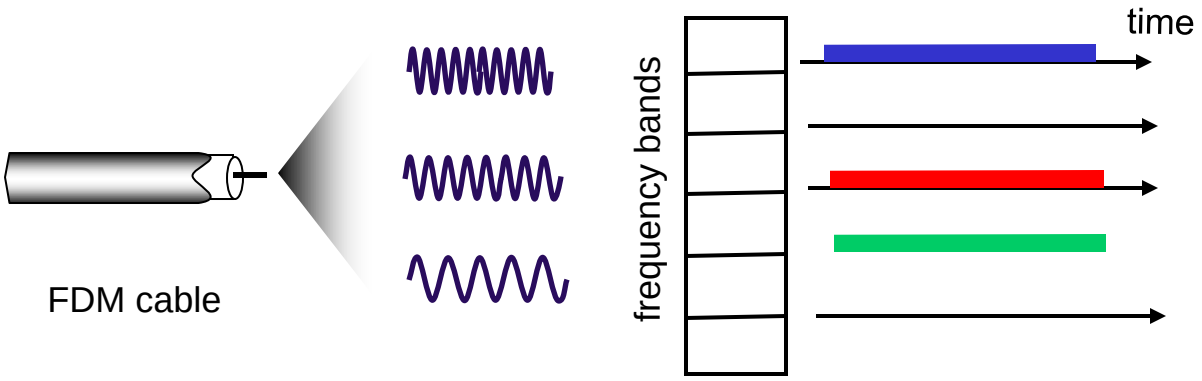
verdelen in tijdsloten, enlk station krijgt een vast slot toegewezen om tijdends elke ronde om te ververzenden

Channel partitioning MAC: FDMA

FDMA: frequency division multiple access

- channel spectrum divided into frequency bands
- each station assigned fixed frequency band
- unused transmission time in frequency bands go idle
- example: 6-station LAN, 1,3,4 have packet to send, frequency bands 2,5,6 idle

verdeelt het specturm van een kannaal in verschillende freqbanden,
elk station eigen freqband



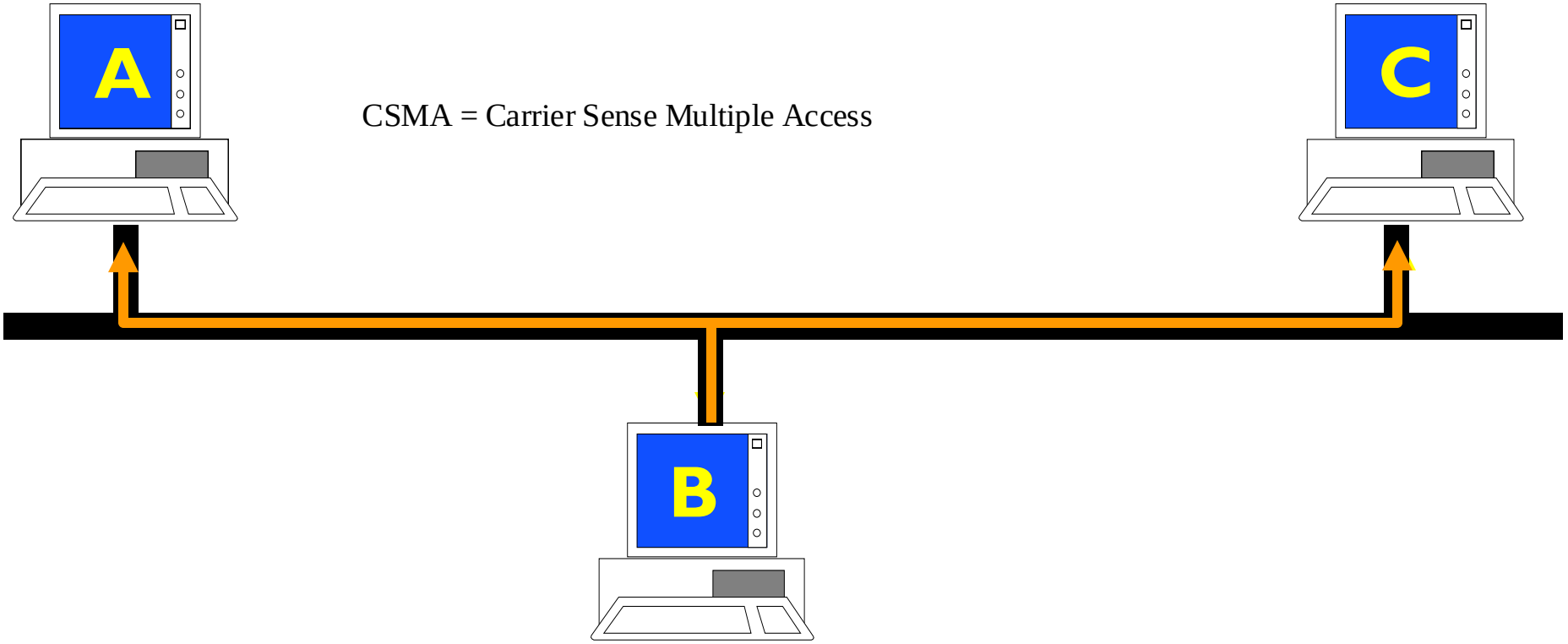
Random access MAC protocol

- when node has packet to send
  - transmit at full channel data rate R
  - no *a priori* coordination among nodes
- two or more transmitting nodes: "collision"

hier geen eerdere coordinatie met andere nodes

- random access protocol specifies:
  - how to detect collisions
  - how to recover from collisions (e.g., via delayed retransmissions)
- examples of random access MAC protocols:
  - ALOHA, slotted ALOHA
  - CSMA, CSMA/CD, CSMA/CA

MAC protocol (Carrier Sense Multiple Access)



CSMA = Carrier Sense Multiple Access

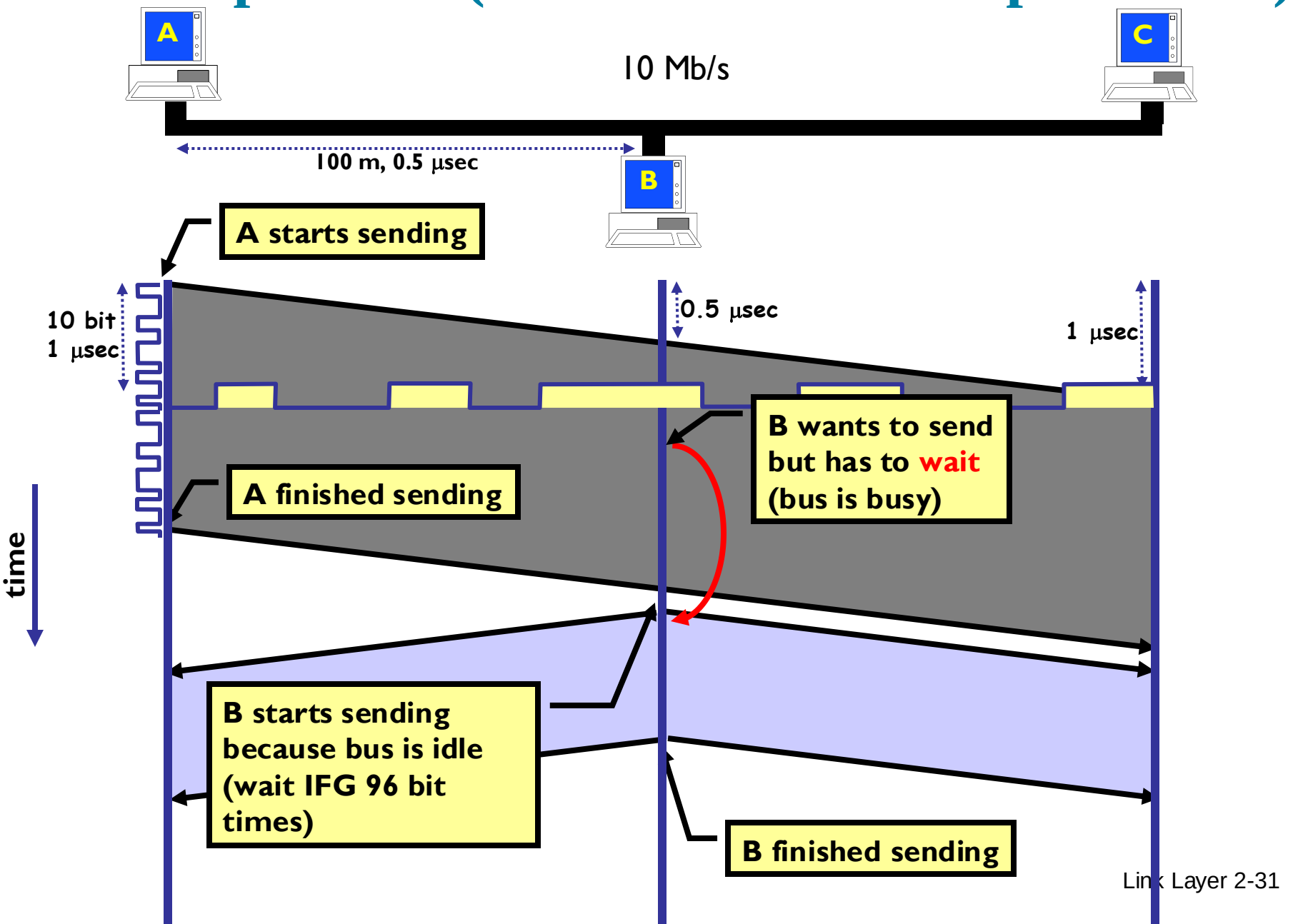
A has a frame to send : if bus idle ==> start sending

B has a frame to send : bus busy ==> wait sending

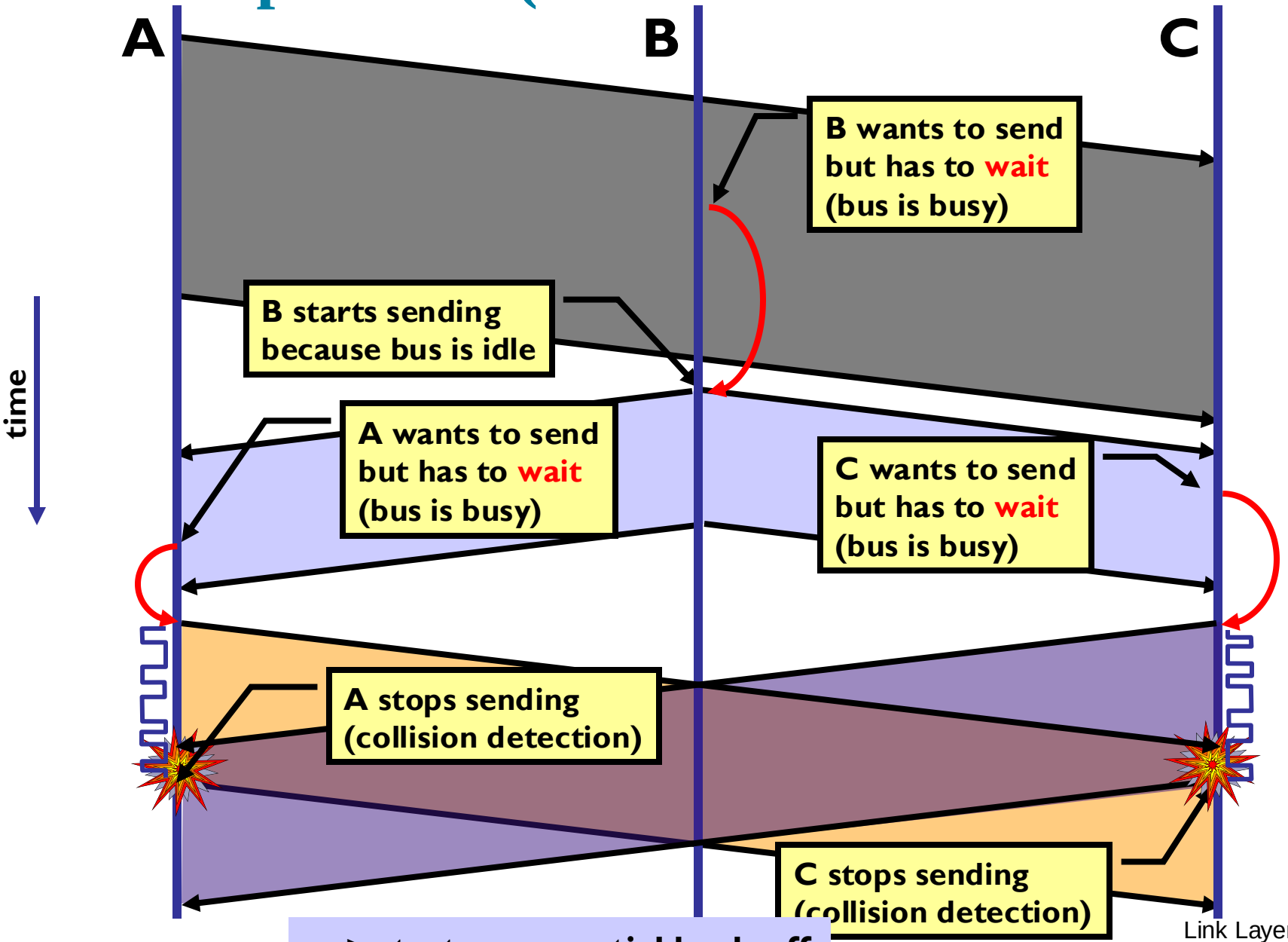
A finished sending : bus becomes idle

==> B starts sending

MAC protocol (Carrier Sense Multiple Access)



MAC protocol (CSMA/Collision Detection)



MAC protocol – Binary Exponential Backoff

COLLISION DETECTION ==>
Minimum frame length = 51.2 μsec

time slot (512 bit times)

2 slots: ½ chance for collision
4 slots: ¼ chance for collision
8 slots: 1/8 chance for collision
(indeed: if ons senders is using a slot, there are 7=8-1 other slots for the second sender where there is no collision ➔ 1/8 chance for collision)

two frames collide

two stations wait 0 or 1 time slot

two stations wait 0,1,2 or 3 time slots

two station wait 0 … 2^3-1 time slots

maximum of 1023 time slots

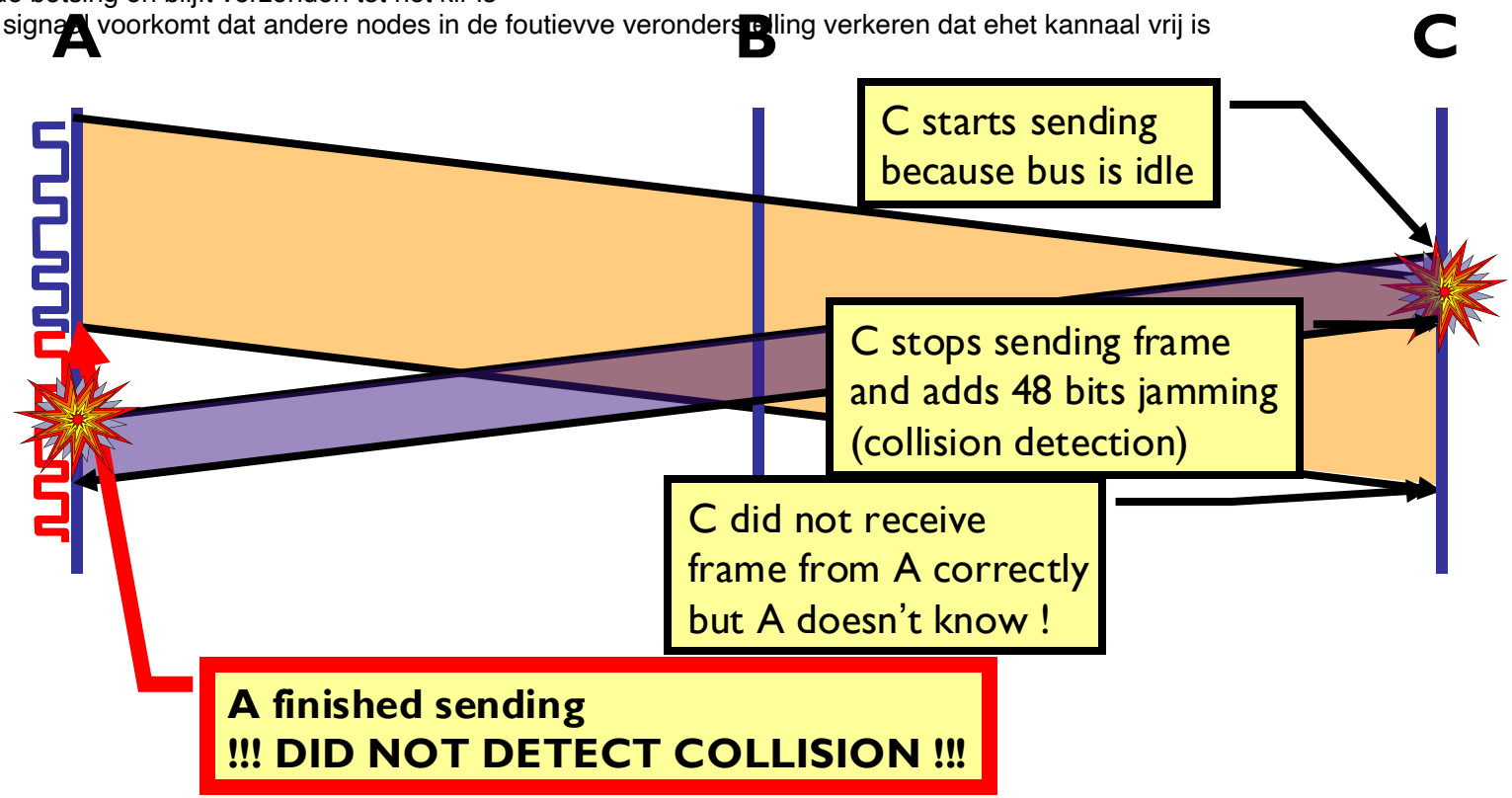
OK

node a en c beginnen op zelfde moment met verzenden,, detecteren elkaar niet het signaal tijd nodig heeft om te reizen

botsing: c detecteert de botsing en stopt met verzenden

a negeert de botsing en blijft verzenden tot het klr is
het jaming signaal voorkomt dat andere nodes in de foutievve veronderstelling verkeren dat ehet kannaal vrij is

MAC protocol (CSMA/Collision Detection)



Frame duration > RTT (=2 times propagation time) **: correct collision detection**
==> minimum frame length = 64 bytes or 512 bits
==> at 10 Mbit/s this gives a maximum of 51.2 μsec roundtrip delay
(5 μsec delay per km cable ==> max 5 km cable)

see notes

It is possible that a collision is not detected, as illustrated in the figure. Terminal A is sending a frame, but before it arrives at terminal C, this latter terminal also starts sending a frame (which is allowed because it is still sensing an idle bus). There will be shortly after a collision in C, and C stops sending (by finishing with a jamming signal). This signal (correct start of the frame + jamming signal due to collision) will propagate to A, but in the figure we observe that A finished its frame just before the signal from C arrives in A. As a result, A will not detect the collision of its frame and will suppose that it correctly arrived. In this way we observe that the collision detection is not working properly.

It is only when the frame emitted by A is longer than 2 times the propagation time between A and C, that the problem will not exist. In that case A will always detect a collision. As a result the exponential back-off algorithm will start (see later).

Note that the above mentioned problem will be solved by putting an upper limit to the length of an Ethernet network because there is a minimum frame length (64 bytes). The round-trip delay in the network should indeed be limited to 51.2 μsec (= duration of the minimum Ethernet frame = 64 bytes = 512 bits = 51.2 μsec at a 10 Mbit/s bitrate). This is the reason why one takes the minimum frame size equal to 64 bytes, in order to have reasonable cable lengths possible (maximum of 5 km). In practice the cable length will be further limited in order to obtain a low enough signal attenuation*.

*Note : a collision may be detected because the voltage on the cable is higher then the normally expected values. Indeed if one is receiving two signals (=collision), they will be added on the transmission line, resulting in a higher voltage then expected. This requires however that the signals are not to much attenuated along the line (limiting the length of the cable) : A strongly attenuated signal colliding with a normal signal will not be detected.

“Taking turns” MAC protocols

channel partitioning MAC protocols:

- share channel *efficiently* and *fairly* at high load
- inefficient at low load: delay in channel access, 1/N bandwidth allocated even if only 1 active node!

random access MAC protocols

- efficient at low load: single node can fully utilize channel
- high load: collision overhead

“taking turns” protocols

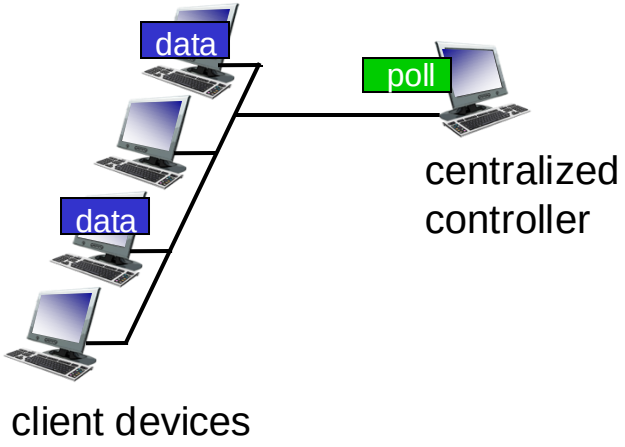
- look for best of both worlds!

“Taking turns” MAC protocols

The first drawback is that the protocol introduces a polling delay—the amount of time required to notify a node that it can transmit. If, for example, only one node is active, then the node will transmit at a rate less than *R* bps, as the master node must poll each of the inactive nodes in turn each time the active node has sent its maximum number of frames. The second drawback, which is potentially more serious, is that if the master node fails, the entire channel becomes inoperative. The Bluetooth protocol, which we will study in Section 6.3, is an example of a polling protocol.

polling:

- centralized controller “invites” other nodes to transmit in turn
- typically used with “dumb” devices
- concerns:
  - polling overhead
  - latency
  - single point of failure (master)
- Bluetooth uses polling



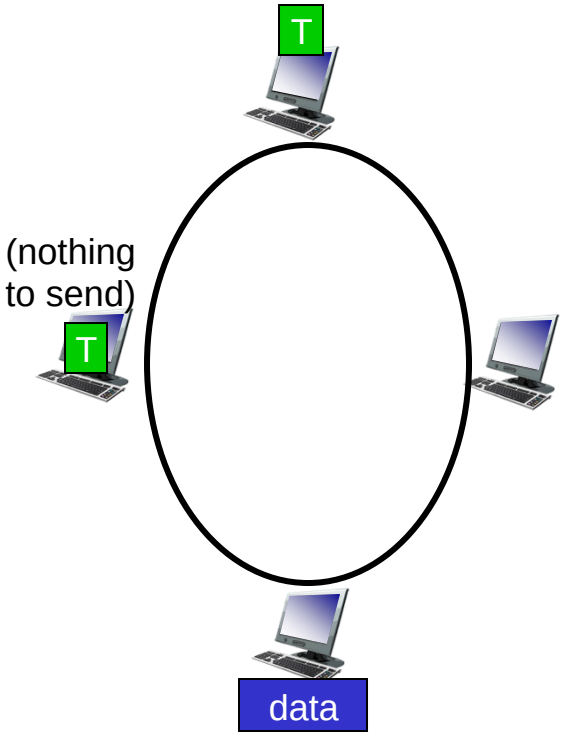
een controller nodigt uit andere nodes om de beurt om gegevens te verzender
- poling overhead: de controller moet elk apparaat periodiek controleren : extra tijd en middelen
latency: vertraging optreden voordat mag verzender
single poiont of failure: als de controller faalt kan hele netwerk n eit meer functioneren
bv bluetooth

“Taking turns” MAC protocols

node kan alleen gegevens verzenden als het het token in bezit heeft

token passing:

- control *token* message explicitly passed from one node to next, sequentially
  - transmit while holding token
- concerns:
  - token overhead
  - latency
  - single point of failure (token)



no master node

But it has its problems as well. For example, the failure of one node can crash the entire channel. Or if a node accidentally neglects to release the token, then some recovery procedure must be invoked to get the token back in circulation.

Link layer, LANs: outline

6.1 Introduction, services

6.2 Error detection, correction

6.3 Multiple access protocols

6.X Access network technologies (additional material!)

- **Introduction**
- ***DSL access technology**
- **Twisted pair-based access technology**
- **Fiber-based access technology**

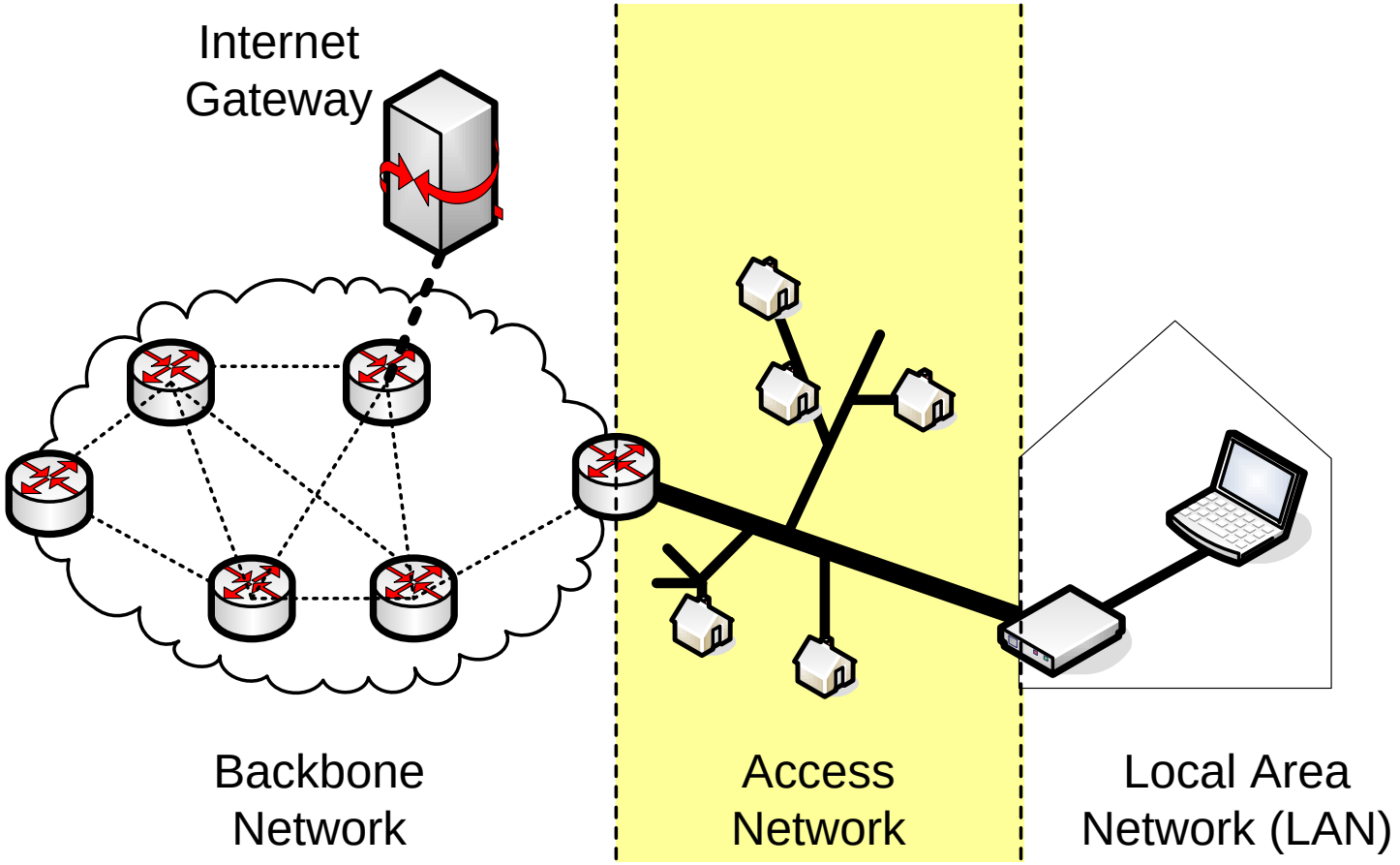
6.4 LANs

6.5 Link virtualization: MPLS

6.6 Data center networking

6.7 A day in the life of a web request

Generic network architecture



The (fixed) access network forms in important part of an end-to-end communication network. Although the boundaries of an access network are not always very clear, in general one may consider that part of the overall network that is connecting a customer terminal or customer premises network with the core network as an access network.

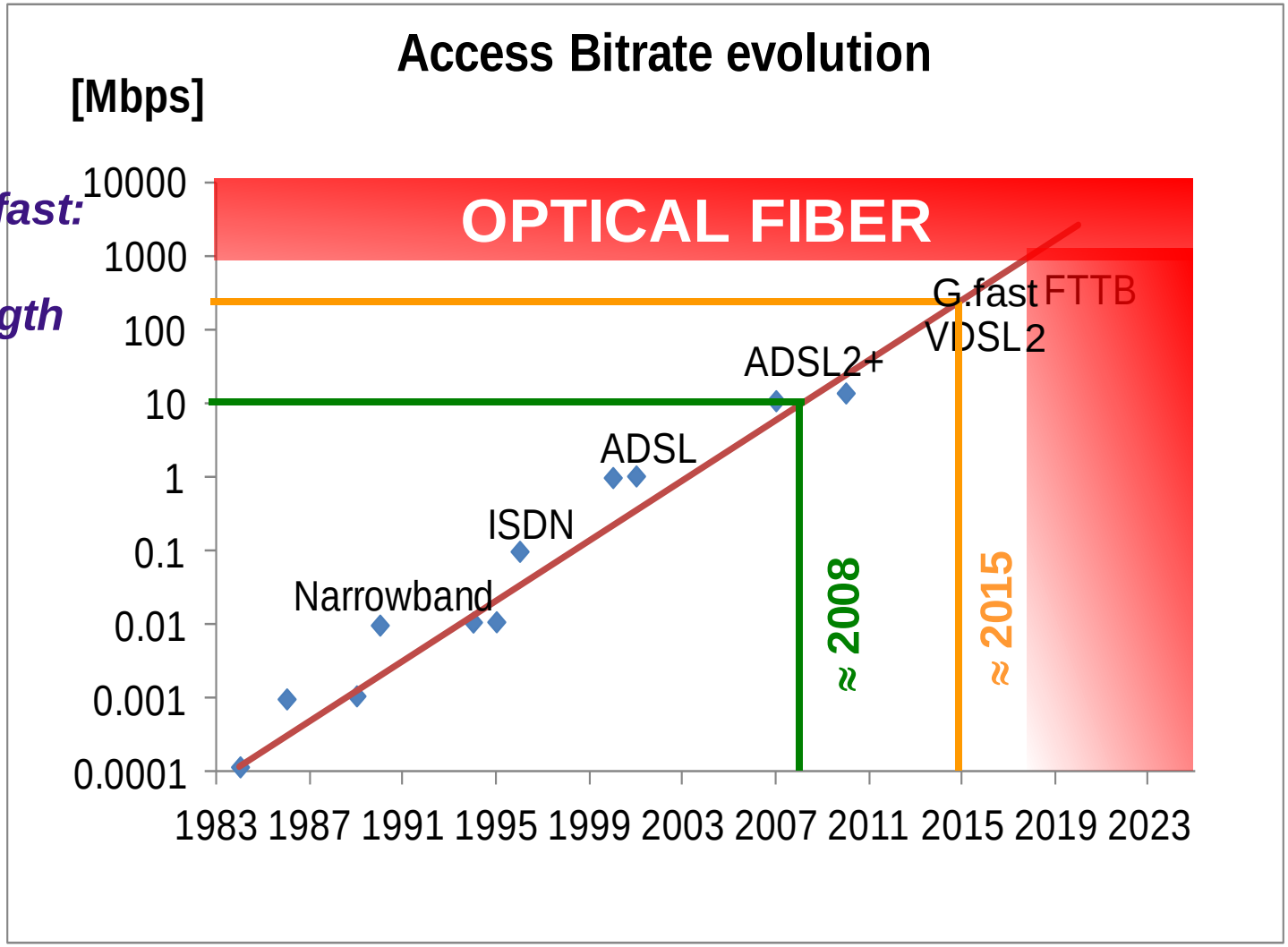
Different types of access networks may be distinguished based on the transmission medium used and on the services offered. In the past there was a clear separation between two types of fixed access networks :

- The twisted pair access was used for access to the telephone network (PSTN or Public Switched Telephone Network). The terminal is a telephone and the gateway is a telephone switch : local exchange (LEX).
- The coaxial cable was used for access to the CATV network (Community Antenna TeleVision). The terminal is a television and the gateway is a head end (where the TV-programs are captured and multiplexed on the coaxial cable).

There is a new type of fixed access network that is based on optical fiber. Originally, optical fiber was mostly used in backbone networks, for high bandwidth connections. Only in recent years, optical fiber has become important in access networks as well. Although the penetration rates are widely varying: in Belgium some first trials started in 2009 with Fiber To The Home (FTTH), while in Japan millions of households were already connected to fiber. One has of course also wireless access networks (4G/5G, WLAN, …), as discussed in other parts of the course.

Timeline wired access networks

DSL standard G.fast:
max. 1 Gbps
@ 100m loop length



The figure gives an idea when different technologies were introduced. This adoption matches very well with Nielsen's law of bandwidth evolution, stating that the user's connection speed roughly grows by 50 percent per year, or doubles every 21 months. Access speeds higher than 1 Gbit/s can only be delivered by an optical fiber towards the customer (i.e. fiber to the home).

- ADSL = Asymmetric Digital Subscriber Line, up to 12 Mbit/s downlink and 1.3 Mbit/s uplink.
- ADSL2+ = Asymmetric Digital Subscriber Line 2 plus, up to 24 Mbit/s downlink and 3.5 Mbit/s uplink.
- VDSL2 = Very-high-bit-rate Digital Subscriber Line 2 (VDSL2), sum of both directions up to 200 Mbit/s. In Belgium, Belgacom has rolled out VDSL2 with a coverage of 90.1% in Sept. 2014.
- G.fast = a DSL standard for local loops shorter than 250 m, sum of both directions from 250 Mbit/s (at 250 m) up to 1 Gbit/s (at 100m).
- FTTB (fiber to the building/basement) is a possible deployment scenario for G.fast, where the fiber node (terminating the optical fiber signal) is in the basement of a multi-dwelling unit and G.fast is used on the in-building telephone cabling.

Link layer, LANs: outline

6.1 Introduction, services

6.2 Error detection, correction

6.3 Multiple access protocols

6.X Access network technologies (additional material!)

- **Introduction**
- **Twisted pair-based access technology**
- **Cable-based access technology**
- **Fiber-based access technology**

6.4 LANs

6.5 Link virtualization: MPLS

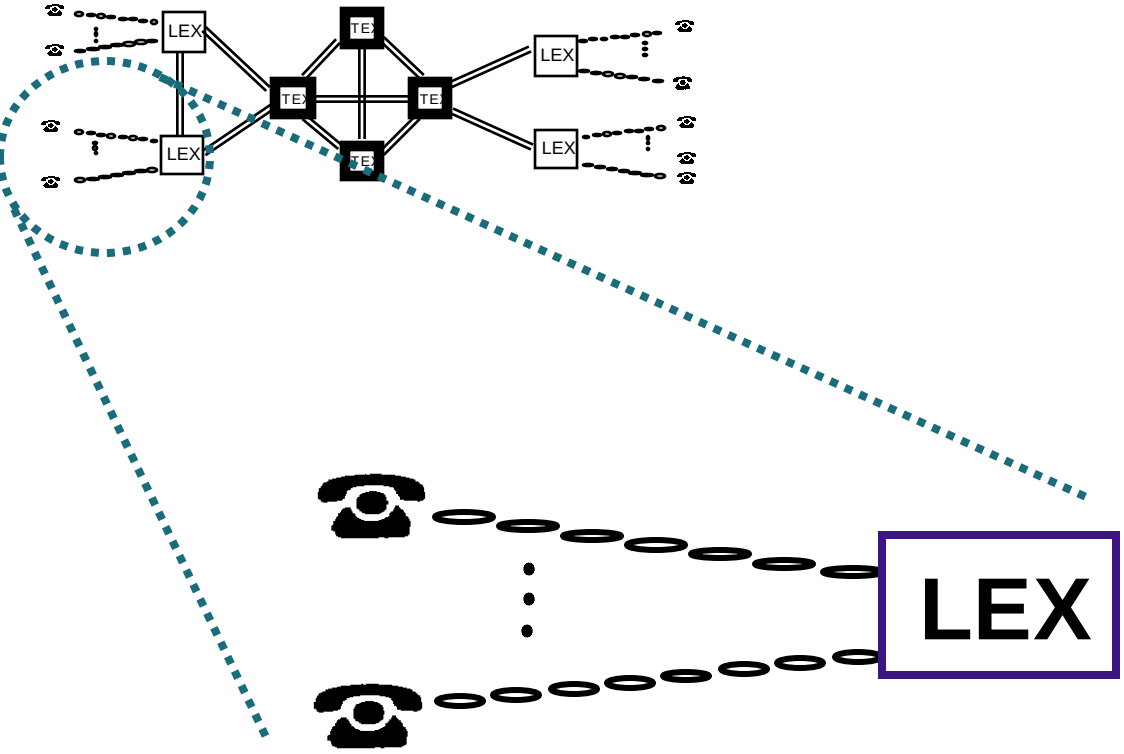
6.6 Data center networking

6.7 A day in the life of a web request

Twisted pair access: telephony

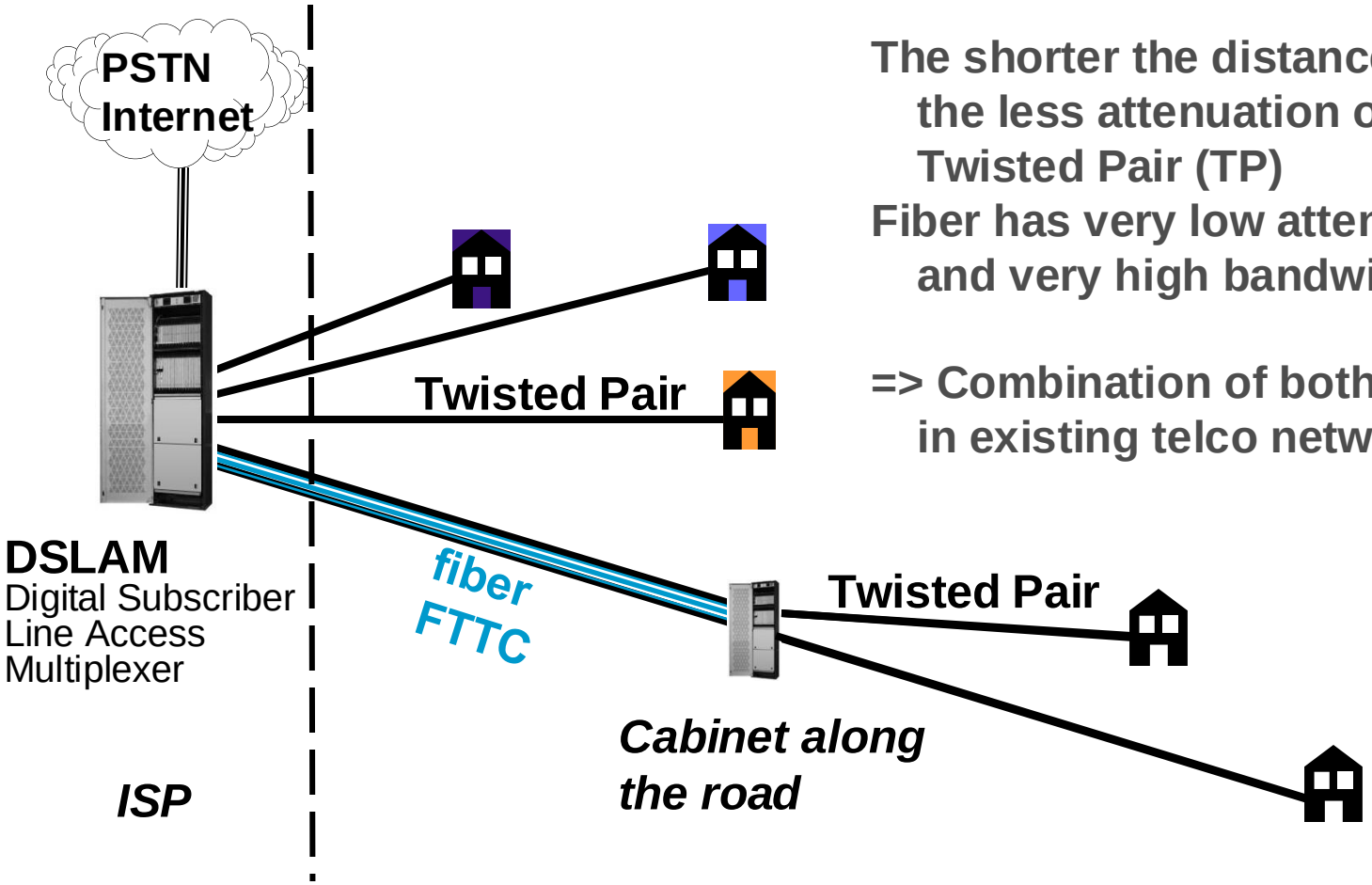
The telephone access network makes use of twisted pair to transmit the voice signal (0.3 - 3.4 kHz bandwidth) from the LEX (Local EXchange) [indicated as CO (central office in the optical access network)] to the user. The topology is a star with a typical radius between 1 and 10 km. The transmission is bidirectional point-to-point. In a typical access network one had over 10000 users connected to the same LEX and the penetration (number of users connected) is larger than 90%.
It is important to remember that about 80% of the network transmission cost is in the access network.

- **telephony**
- **twisted pair**
- **star (1-10 km)**
- **point-to-point**
- **bidirectional**
- **0.3 - 3.4 kHz**
- **>10000 users/LEX**
- **>90% penetration**



LEX = Local Exchange
TEX = Telephone Exchange

Twisted pair : Broadband



The shorter the distance,
the less attenuation on
Twisted Pair (TP)
Fiber has very low attenuation
and very high bandwidth

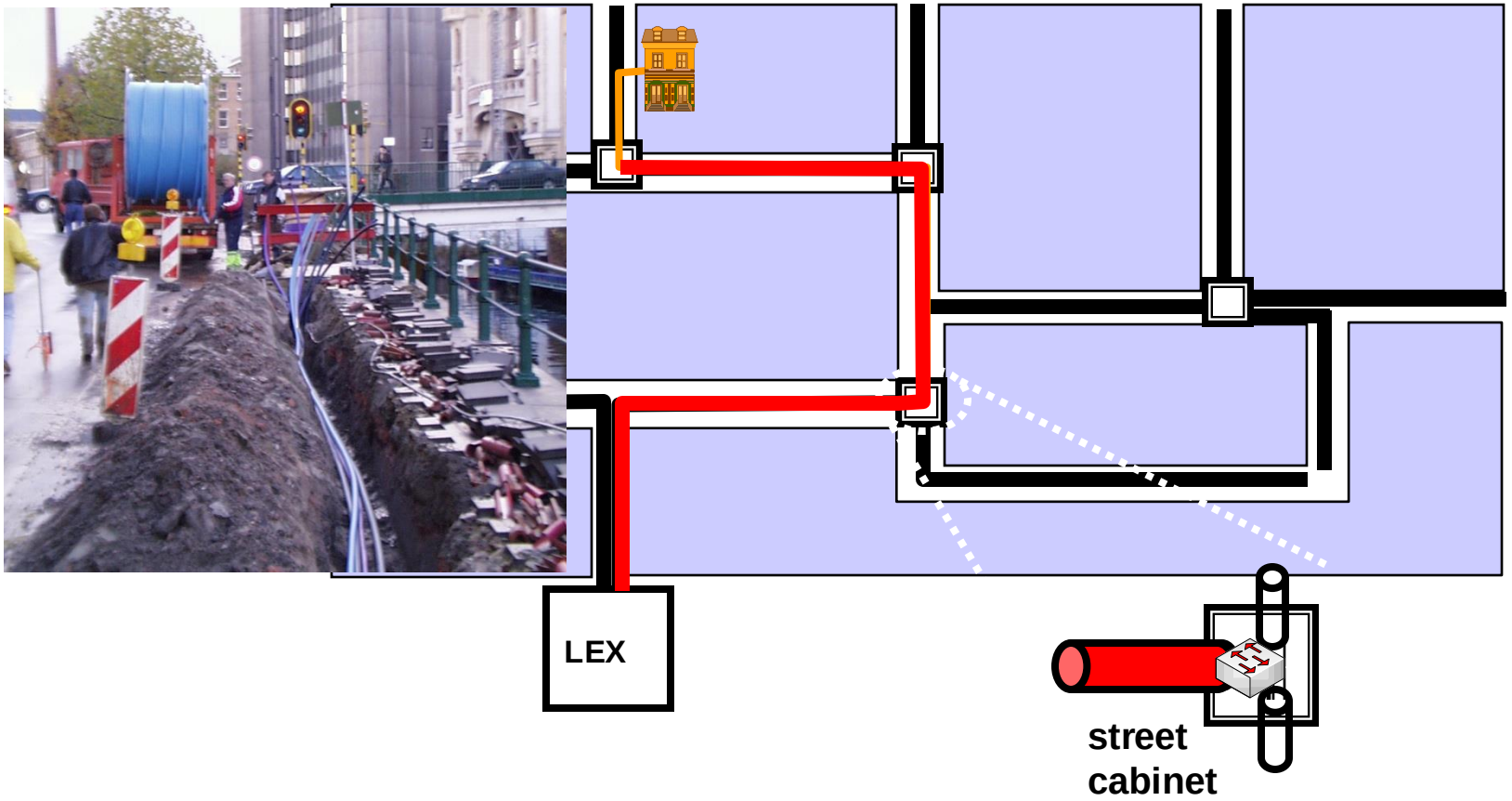
=> Combination of both
in existing telco networks

PSTN = Public Switched Telephone Network
FTTC = Fiber To The Cabinet

Twisted Pair (TP) was voor telefoonverbindingen uitermate geschikt: bij de lage frequenties was er weinig attenuatie.
Door de grote attenuatie die bij TP optreedt bij hogere frequenties, was het initieel enkel mogelijk om huizen die op korte afstand (< 3,5 km) van een centrale gelegen waren, te verbinden met een DSL modem (Digital Subscriber Line).
Bij de ISP wordt de aansluiting van de telefoonlijn doorverbonden met een DSLAM, een multiplexer die de DSL signalen van vele modems tegelijk kan verwerken. Deze is met het oude telefoonnet (PSTN) *en* het Internet verbonden.
De geleidelijke invoering van optische vezel als signaalgeleider biedt hier een oplossing. Optische vezel heeft immers quasi geen last van attenuatie.
De fibers worden geïnstalleerd tot op een zekere afstand in het access network, in de aansluitkasten van Proximus. De gebruiker bevindt zich zo slechts een “last mile” van een optische vezel capaciteit, ook al wonen ze vele kilometers van een centrale van de ISP. Voorwaarde is dat er conversie tussen elektrische en optische signalen kan gebeuren in de aansluitkast.
De “last mile” afstand wordt dus nog steeds bediend door TP: dit is immers het grootste deel van het geïnstalleerde network en zou uiterst duur zijn om te vervangen door vezel. Vaak wordt de waarde van deze aansluitingen op 80% van de waarde van het bedrijf geschat.
Proximus lanceert binnenkort voor *nieuwe* aansluitingen een combinatie van “telefoonkabel + wachtbuisje voor fiber”, met het oog op “fiber to the home”.
Afhankelijk van de bandbreedte die men wenst aan te bieden over de TP, kan men beslissen om de optische vezel te brengen tot een aansluitkast op enkele kilometers van de gebruiker verwijderd, of tot op enkele honderden meters.

Twisted pair access: broadband

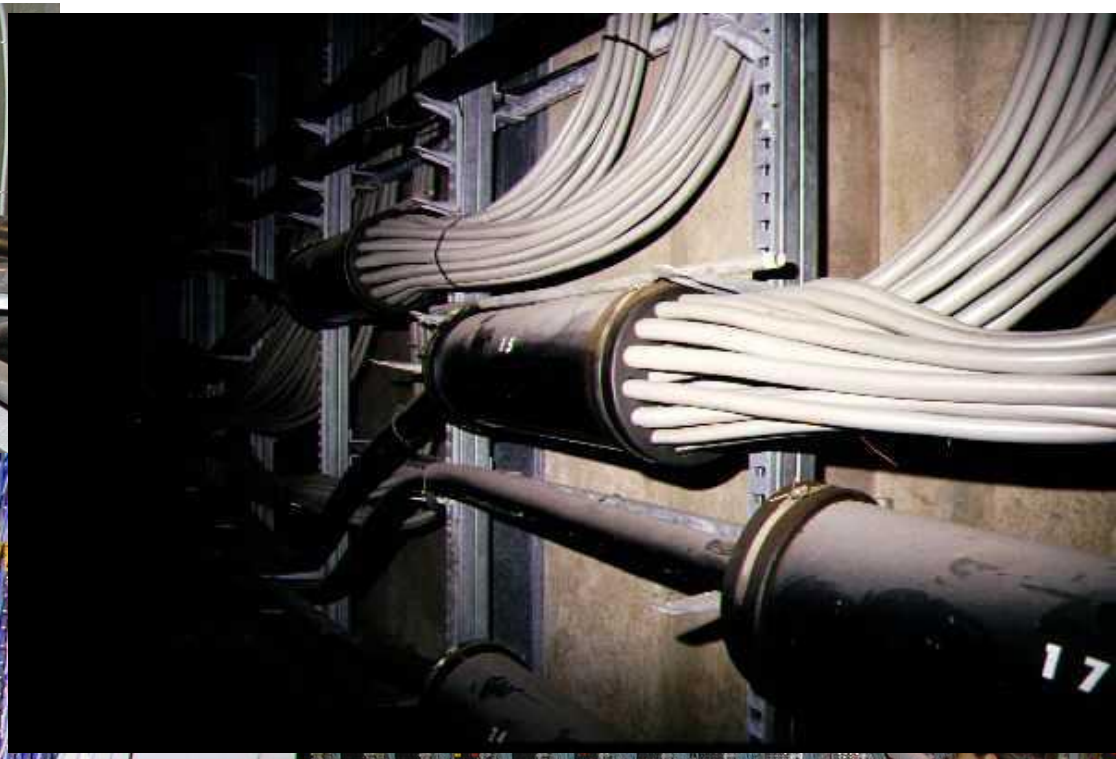
The figure above shows how the fiber is gradually introduced in the access network, getting closer to the user.

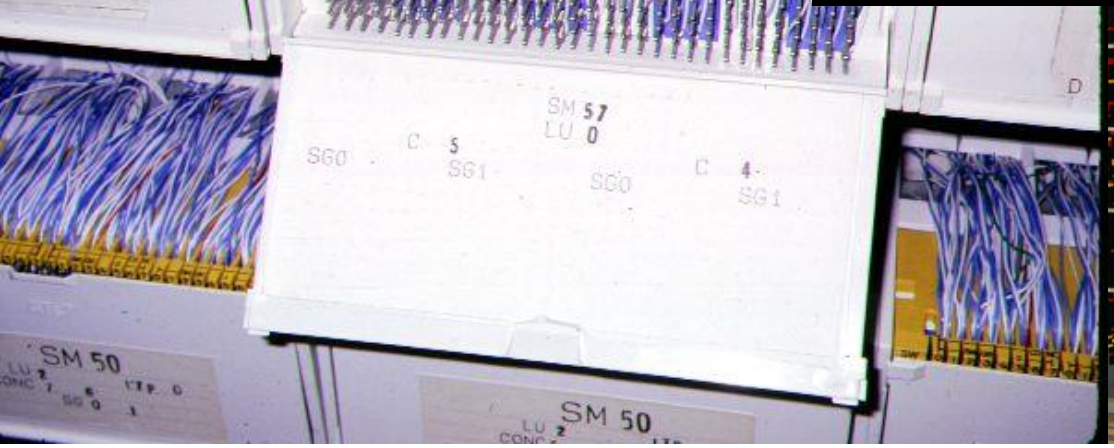


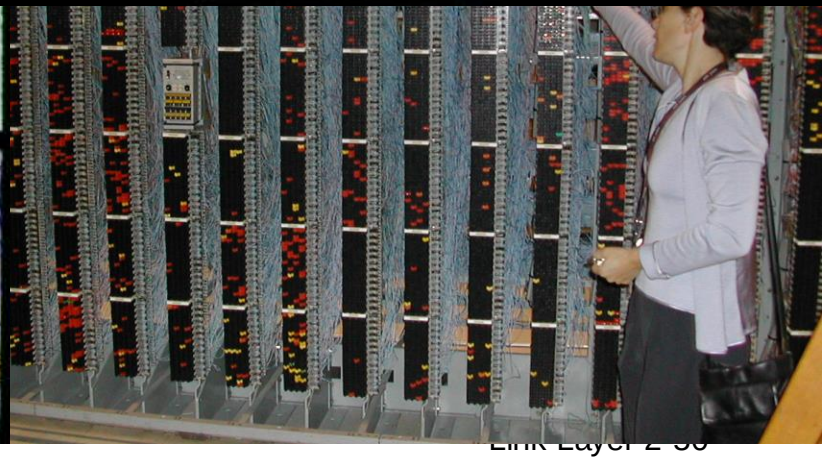
Introduction of Fiber

Some examples are given of the twisted pair cables used in the exchange buildings (top picture). Distribution frames (where the twisted pairs are terminated) are shown at the bottom.

Twisted pair access

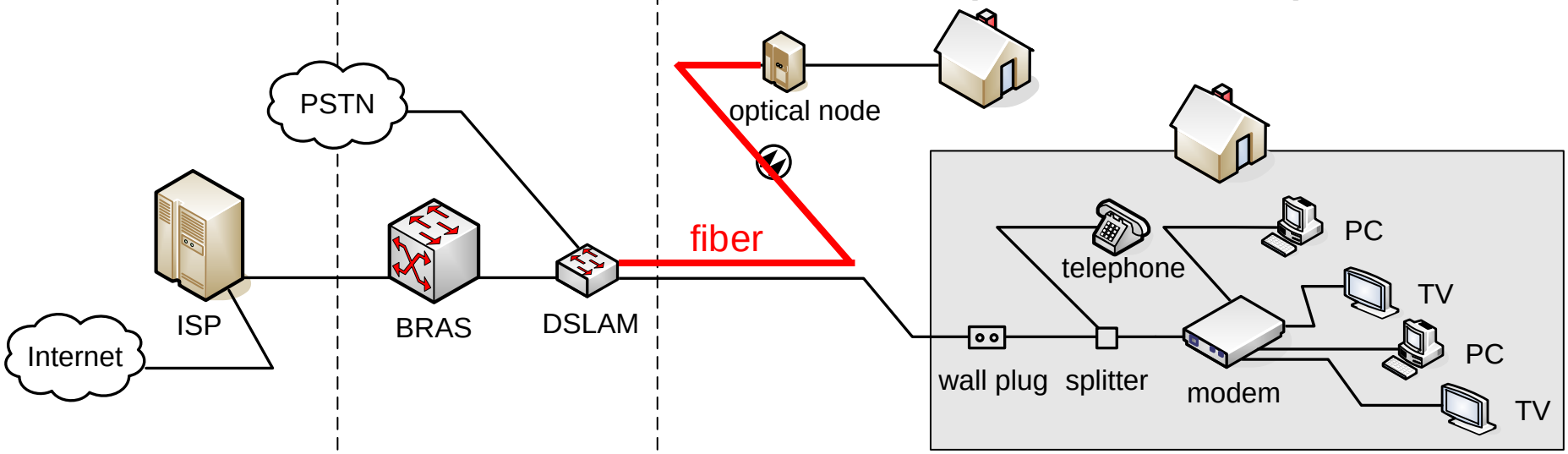






Twisted pair access: ADSL architecture

The figure shows the typical architecture of an ADSL based access network. At home, a computer or multiple computers are connected (via a switch or hub) towards an ADSL modem. A splitter is used to separate the telephone signal and the ADSL signal. A twisted pair will connect the home to the access multiplexer (DSLAM in central office). The DSLAM will connect to the access aggregation network using another L2 technology (e.g. Asynchronous Transfer Mode (ATM) or Ethernet). The L2 network is connected to a BRAS (Broadband Remote Access Server). The BRAS will connect to the public Internet or to other service providers.

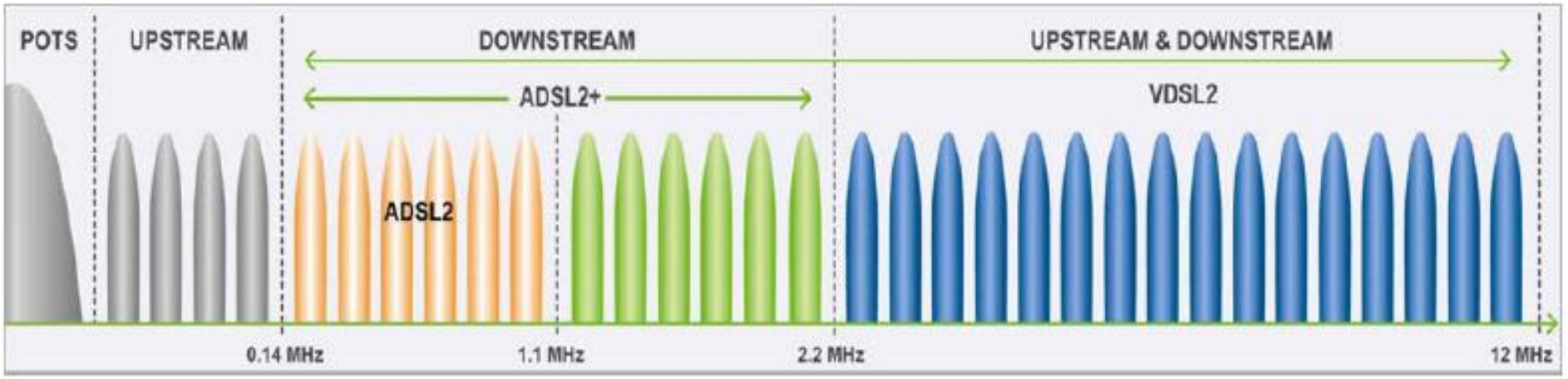


ADSL : Asymmetric Digital Subscriber Line
DSLAM : DSL Access Multiplexer
BRAS : Broadband Remote Access Server

Note: between DSLAM and BRAS: aggregation network

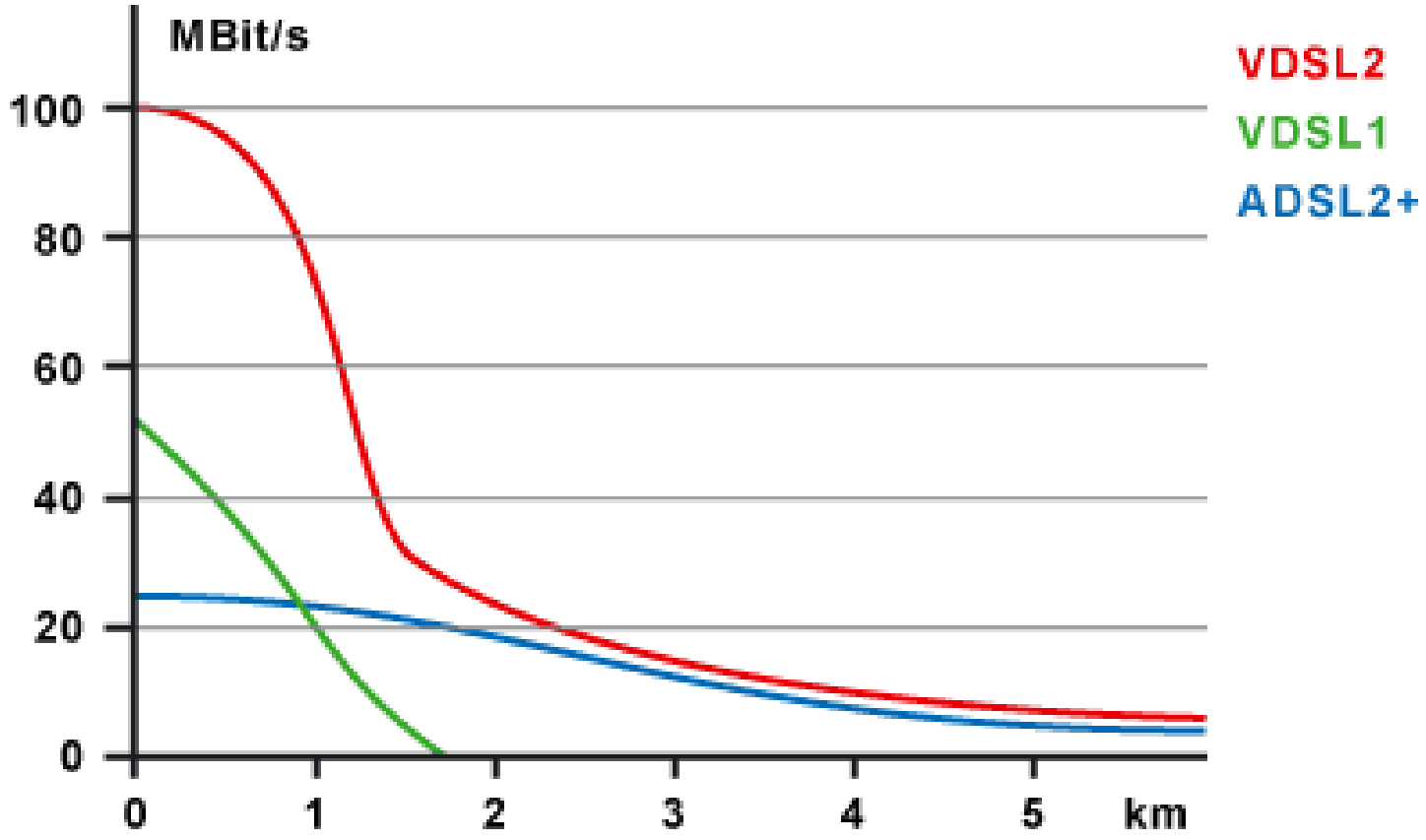
ADSL2+ / VDSL

Het verschil tussen ADSL2+ en de oudere technologieën (ADSL, ADSL2) zit vooral in de gebruikte bandbreedte: deze is verdubbeld van 1.1 MHz naar 2.2 MHz. Aangezien de verzwakking bij deze hogere frequenties groter is zal de afstand korter zijn vooraleer het signaal in de ruis verdwijnt. Om dus optimaal gebruik te kunnen maken van de grotere bandbreedte van ADSL2+ moet de afstand tot de aansluitkast korter zijn.



https://upload.wikimedia.org/wikipedia/commons/3/32/VDSL2_frequencies.png

ADSL / VDSL - attenuatie



Twisted pair : DSL versions

	Standard (ITU-T)	Bit rate [Mbps]		Max. TP length	Symmet-rical?	PSTN compatible?
		DS	US			
HDSL	G.991.1	2	2	7 km	Yes	-
SHDSL	G.991.2	4	4	3 km	Yes	-
ADSL	G.992.1	8	1	3.5 km	-	Yes
ADSL lite	G.992.2	1.5	0.5	5.5 km	-	Yes
RADSL	No ITU	8	1	5.5 km	-	Yes
ADSL2	G.992.3	12	1	3 km	-	Yes
RE-ADSL2	G.992.3 Annex L	12	1	7 km	-	Yes
ADSL2+	G.992.5	24	1	1.5 km	-	Yes
VDSL	G.993.1	6.5	1.2	1.5 km	-	Yes
		13	13	1 km	Yes	
		26	3.2		-	
		26	26	300 m	Yes	
		55	15		-	
VDSL2	G.993.2	100	100	< 300 m	Yes	Yes

DSL : Digital Subscriber Line

Een overview van de verschillende soorten DSL (Digital Subscriber Line) technologiën die bestaan, vind je in de tabel hierboven.
HDSL: High speed DSL
ADSL: Asymmetric DSL, ongelijke up & download bitrate
VDSL: Very high speed DSL
In België kennen ADSL en ADSL2 de grootste verspreiding, gezien de relatief goede downloadsnelheden (DS resp. 8 en 12 Mbps). De upload snelheid is typisch veel kleiner, maximum 1 Mbps. Gezien de afstand met TP tot de aansluitkast 3 km mag zijn, was dit de initiële investering die door Proximus gebeurde.
De bandbreedte laat VoIP en Digitale TV toe, maar als beide over een klassieke ADSL lijn worden gestuurd, schiet er niet veel bandbreedte meer over voor Internet toepassingen.
De opvolger VDSL is eveneens asymmetrisch, maar kent verschillende configuratiemogelijkheden: afhankelijk van de afstand die met TP moet overbrugd worden, kunnen andere up- & downloadbitrates ingesteld worden. Ook symmetrische bitrates kunnen ingesteld worden.
Deze hogere bandbreedte laat toe om naast gewoon Internet verkeer, ook zonder beperkingen Digitale TV en VoIP telefonie aan te bieden over de DSL lijn.
Bepaalde locaties in België kunnen bvb. niet aangesloten worden op Proximus TV, omdat de voorzieningen voor VDSL nog niet geïnstalleerd zijn.

Link layer, LANs: outline

6.1 Introduction, services

6.2 Error detection, correction

6.3 Multiple access protocols

6.X Access network technologies (additional material!)

- **Introduction**
- ***DSL access technology**
- **Cable-based access technology**
- **Fiber-based access technology**

6.4 LANs

6.5 Link virtualization: MPLS

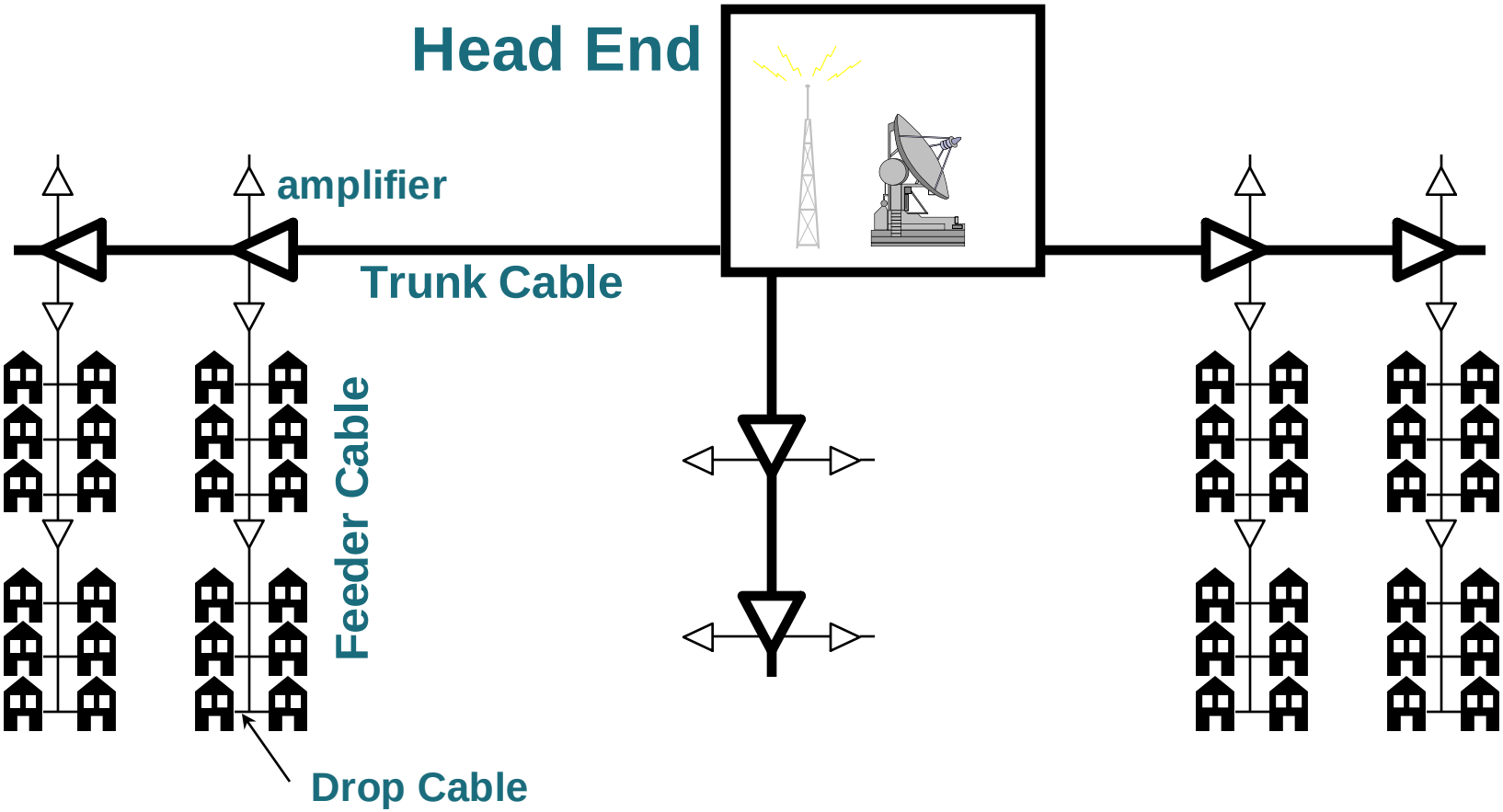
6.6 Data center networking

6.7 A day in the life of a web request

In de meeste huishoudens in Vlaanderen is niet enkel de twisted pair van de telefoon beschikbaar, maar eveneens een coax kabelaansluiting, die traditioneel voor TV distributie werd gebruikt.
Ook deze kabel kan als “last mile” gebruikt worden. De opbouw is gelijkaardig met het DSL gebaseerde access network. Ook hier worden bij de gebruikers thuis modems geplaatst, cable modems, die toelaten signalen te versturen op andere frequenties dan de TV distributie.
Bij de ISP moet een toestel deze signalen kunnen ontvangen; deze multiplexer wordt een CMTS of Cable Modem Termination System genoemd.
Het grootste verschil zit in de initiële opbouw van beide networken. Telefoonlijnen zijn eigenlijk privé verbindingen die lopen van thuis tot bij de ISP. De kabels lopen weliswaar samen in dezelfde kabelbundels, maar de verbindingen zijn individueel en bedoeld voor tweerichtingsverkeer.
Bij het kabeldistributie network is de opbouw een gedeeld medium: alle huizen in de straat zijn aangesloten op één en dezelfde coax kabel. Bovendien is het network gebouwd om data – TV signalen – te sturen naar het huis, maar niet om iets terug te sturen. Gedeeld - en eenrichtingsverkeer.
De aanpassingen die in dit network moeten gebeuren zijn dus verschillend dan bij het DSL network dat we hiervoor besproken hebben.

Topology

Key characteristics:
- Unidirectional
- Shared medium per branch



De figuur toont de opbouw van een traditioneel kabel network (CATV – Cable TV). In de typische “vertakte” topology wordt het (analoge) TV signaal opgevangen in een Head End. Van daar wordt het doorgestuurd naar de vele verbonden klanten thuis. Om de attenuatie en het splitsingsverlies op te vangen, zijn op regelmatige afstand versterkers geplaatst.
Het volledige systeem is unidirectioneel, omdat de versterkers unidirectioneel de analoge signalen versterken.
Het network bestaat uit 3 delen: trunk kabels (met trunk versterkers), feeder kabels in de straten (met feeder versterkers) en drop kabels die de klant tot in het huis verbinden.
Doordat alle kabels met elkaar verbonden zijn, is dit een gedeeld medium access network: alle bandbreedte moet gedeeld worden door alle klanten die verbonden zijn met een bepaalde “tak” van het network. Gezien bij TV distributie alle klanten dezelfde kanalen en dus dezelfde signalen ontvangen, is dit echter geen beperking.

CATV network : modification

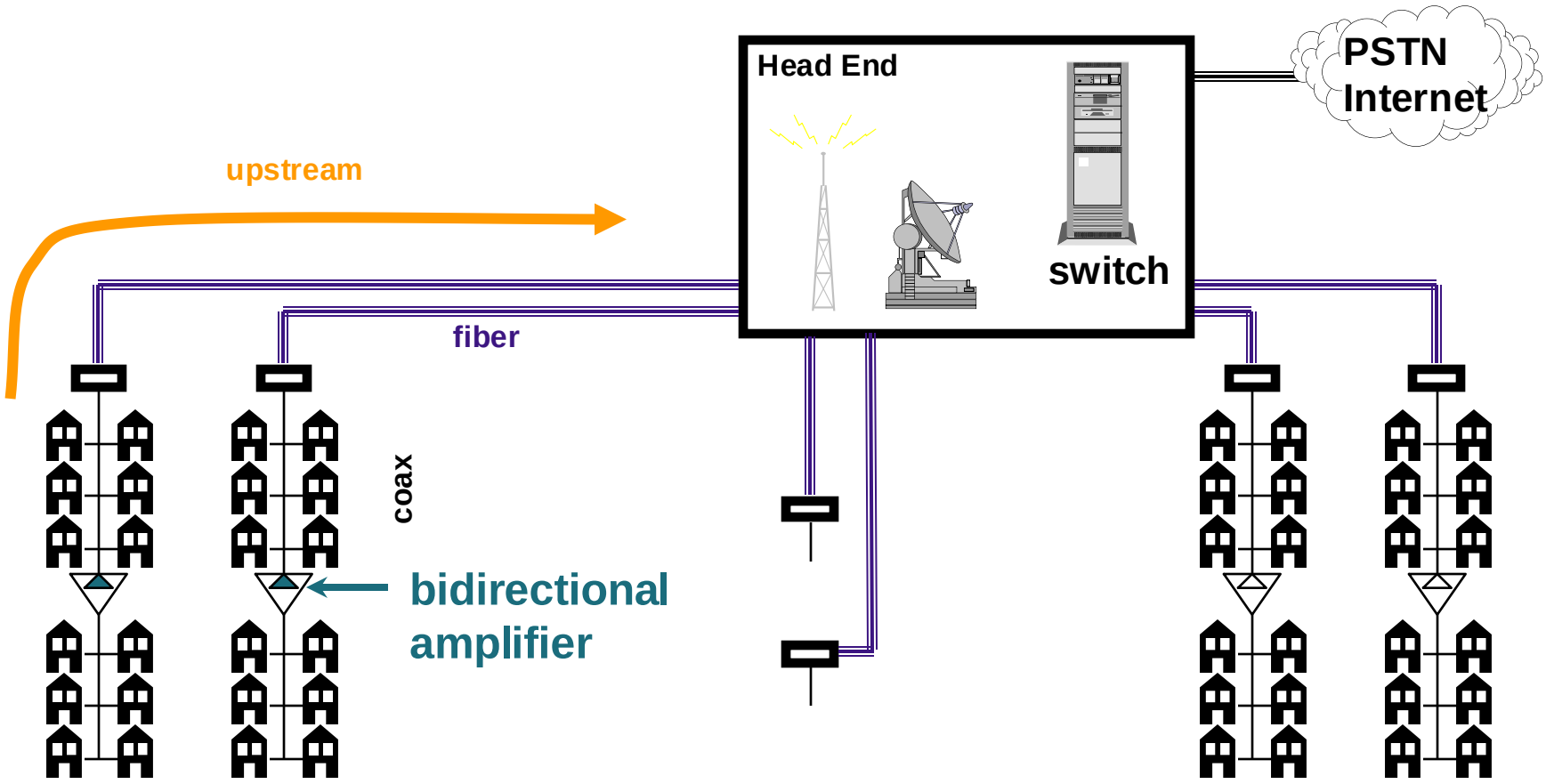
- CATV -> “ISP using CATV”
  - From: unidirectional, analog distribution
  - To: bidirectional, digital based IP network

- Modifications to the old cable network:
  - From analog −> digital signals
  - Bidirectional transmission
  - Reduce the distance of the coax connection
    - introduce optical fiber
  - Optimise Bandwidth usage
    - better modulation techniques on the existing cable network
  - Backbone network construction

De aanpassingen die in het coax network moeten gebeuren om ook Internet en andere interactieve multimudia mogelijk te maken, zijn van velerlei aard:
Overgang van analoge naar digitale signalen, maar met behoud van de analoge TV distributie
Bidirectionele transmissie mogelijk maken
De afstand die met coax wordt afgelegd reduceren door minder klanten op één tak aan te sluiten – klanten die hierdoor dus elk een groter deel van het gedeelde medium kunnen gebruiken
Bandbreedte gebruik verbeteren door geavanceerde modulatietechnieken
Het coax network niet enkel met een Head End maar ook met een backbone network verbinden (voor Internet en voor telefonie).
Belangrijk is dat deze aanpassingen steeds moeten gebeuren zonder de traditionele taak van analoge TV-distributie te onderbreken.
De technologische uitdagingen zijn dus velerlei, we bespreken de aanpassingen in het network op de volgende slides.
De head end moet eerst aangepast worden: deze locaties worden van ontvangststations van analoge signalen omgebouwd tot ISP points-of-presence die switching/routing mogelijkheden ondersteunen.

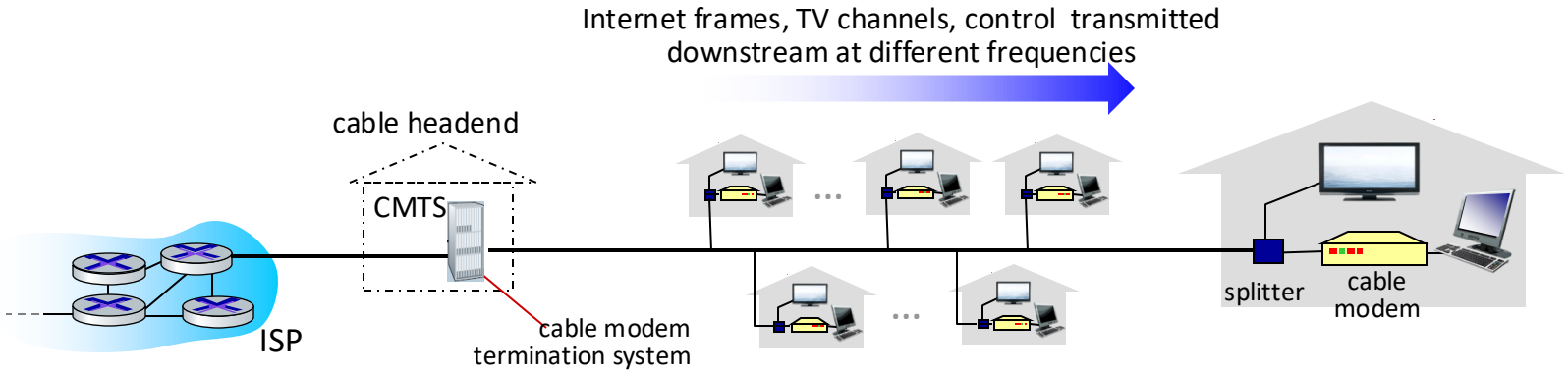
Evolution of coax access network: fiber

Hybrid Fiber Coax (HFC)



Eens een Head End aangesloten is op de backbone en dus op het Internet, kunnen verdere aanpassingen geactiveerd worden.
De trunk kabels van het traditionele CATV network worden vervangen door optische vezel verbindingen. Hierdoor kan enerzijds meer bandbreedte behaald worden, anderzijds wordt het network in kleinere branches opgesplitst. Per branch zijn gemiddeld 500 gebruikers aangesloten. Een andere grote investering is de installatie van nieuwe versterkers: deze moeten immers aangepast worden om bidirectioneel te gaan versterken en bovendien ook hogere frequenties te gaan versterken, aangezien de nieuwere modulatietechnieken in een ander deel van het spectrum worden toegepast.
Een verbeterd CATV network wordt een HFC of Hybrid Fiber Coax network genoemd.

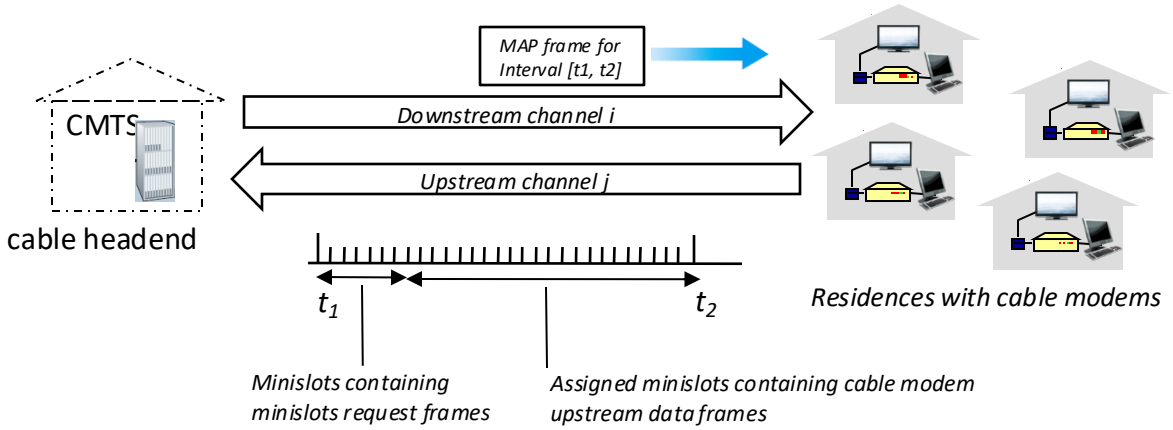
Cable access network: FDM, TDM + random access



- multiple downstream (broadcast) FDM channels: up to 1.6 Gbps/channel
  - single CMTS transmits into channels
- multiple upstream channels (up to 1 Gbps/channel)
  - multiple access: all users contend (random access) for certain upstream channel time slots; others assigned TDM

In de meeste huishoudens in Vlaanderen is niet enkel de twisted pair van de telefoon beschikbaar, maar eveneens een coax kabelaansluiting, die traditioneel voor TV distributie werd gebruikt.
Ook deze kabel kan als “last mile” gebruikt worden. De opbouw is gelijkaardig met het DSL gebaseerde access network. Ook hier worden bij de gebruikers thuis modems geplaatst, cable modems, die toelaten signalen te versturen op andere frequenties dan de TV distributie.
Bij de ISP moet een toestel deze signalen kunnen ontvangen; deze multiplexer wordt een CMTS of Cable Modem Termination System genoemd.
Het grootste verschil zit in de initiële opbouw van beide networken. Telefoonlijnen zijn eigenlijk privé verbindingen die lopen van thuis tot bij de ISP. De kabels lopen weliswaar samen in dezelfde kabelbundels, maar de verbindingen zijn individueel en bedoeld voor tweerichtingsverkeer.
Bij het kabeldistributie network is de opbouw een gedeeld medium: alle huizen in de straat zijn aangesloten op één en dezelfde coax kabel. Bovendien is het network gebouwd om data – TV signalen – te sturen naar het huis, maar niet om iets terug te sturen. Gedeeld - en eenrichtingsverkeer.
De aanpassingen die in dit network moeten gebeuren zijn dus verschillend dan bij het DSL network dat we hiervoor besproken hebben.

Cable access network: FDM, TDM + random access



DOCSIS: data over cable service interface specification

- FDM over upstream, downstream frequency channels
- TDM upstream: some slots assigned, some have contention
  - downstream MAP frame: assigns upstream slots
  - request for upstream slots (and data) transmitted random access (binary backoff) in selected slots

Het spectrum dat uitgestuurd wordt op coax kabels voor analoge TV distributie, is initieel bepaald om weinig attenuatie te kennen. Dit zijn de VHF en UHF band.
Het spectrum dat op de coax kabel kan verstuurd worden, kan verbreed worden, aangezien de afstand minder groot is geworden. Daarnaast wordt het spectrum opgesplitst in een band voor upstream (US) verkeer en een band voor downstream (DS, van network naar user zoals de originele distributie). De preciese frequenties hangen af van land tot land en van de regulator.
Typisch wordt de upstream gealloceerd in het laagste deel van de frequentieband, tussen 5MHz en 25-65MHz. De voormalige VHF-I band is hiervoor intussen opgeofferd. Deze upstream kan verschillende signalen kennen: naast IP pakketten ook digitale telefonie of signalen voor video-aanvragen (Video-on-Demand).
Het merendeel van het spectrum is echter voorbehouden voor downstream verkeer. De frequenties die gebruikt kunnen worden gaan tot 1 GHz. De hogere regionen van het spectrum zijn typisch voorbehouden voor nieuwe interactieve diensten. Een groot deel van het spectrum wordt nog steeds ingenomen door de UHF kanalen, die per kanaal 8 MHz “verbruiken”.
Bemerk dat de laatste jaren steeds meer analoge TV kanalen worden afgeschaft om meer bandbreedte toe te laten voor digitale diensten.
Als een filter wordt geïnstalleerd die enkel Internet toelaat (ToF of Telenet-only Filter), is dit opnieuw een eenvoudige RC filter die op de coax de signalen van de UHF band blokkeert.

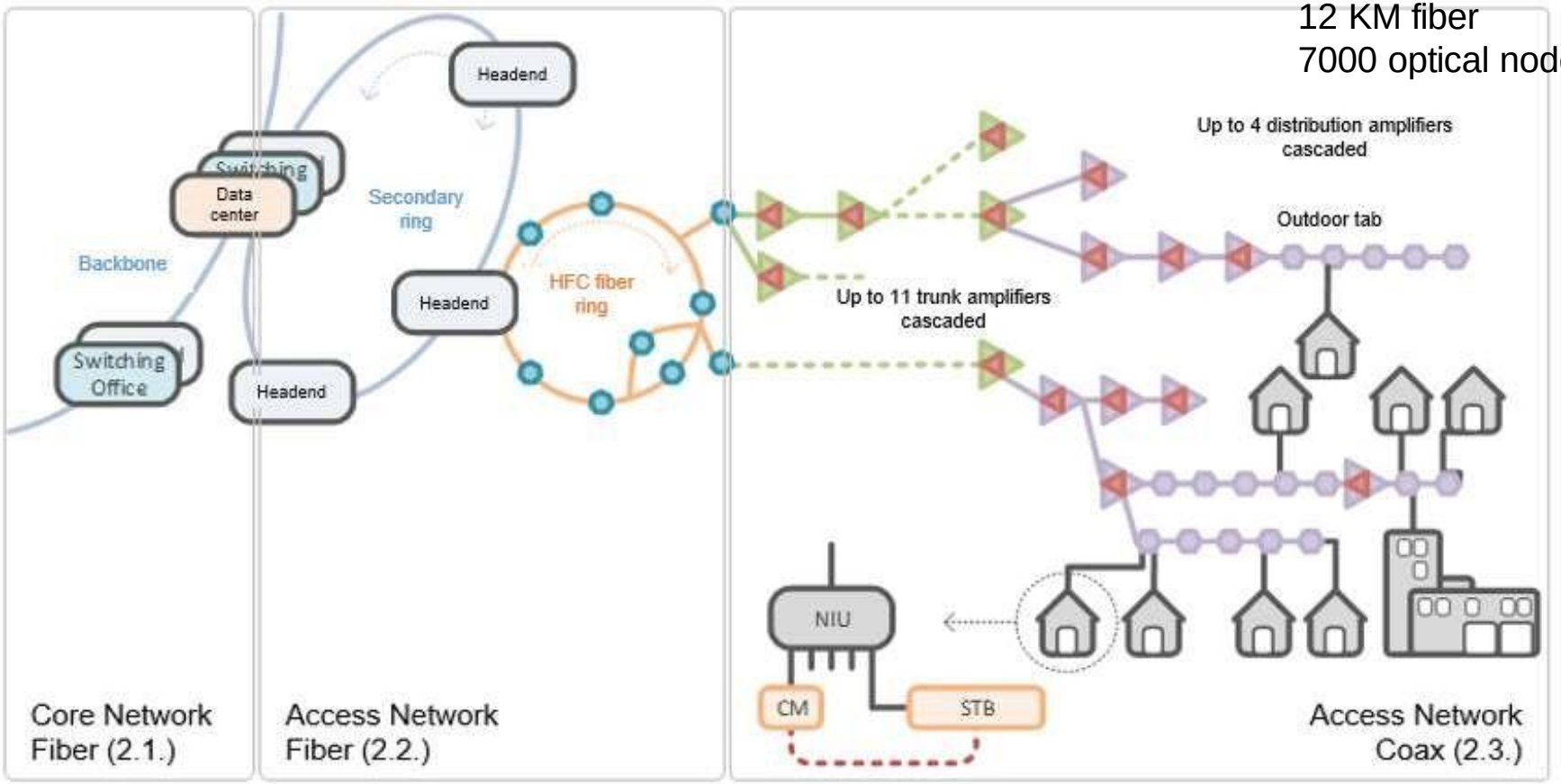
Euro-DOCSIS

- Type of cable modems: DOCSIS => Euro-Docsis
  => Data Over Cable Services Interface Specification, Europese versie
- Four versions:
  - Euro-DOCSIS 1.0: best-effort data service – only IP
  - Euro-DOCSIS 1.1: QoS + advanced security (Authentication)
  - Euro-DOCSIS 2.0: better physical layer in the upstream
    => symmetrical bandwidth possible
  - Euro-DOCSIS 3.0: o.a.
    - channel bonding : bandwidth extension
    - Better security
    - IP multicast, IPv6 support
  - Euro-DOCSIS 3.1:
    - standard 2013, first field tests 2016
    - gigabits speeds (10 down, 1 up)

cable modems werken allemaal op dezelfde standaard, genaamd DOCSIS (Data Over Cable System Interface Specification). Deze Amerikaanse standaard werd aangepast om te kunnen werken op Europese kabelnetworken, vandaar de naam Euro-Docsis.
De Docsis standaard specifieert zowel de fysische laag als de data-link laag. De fysische laag is anders gedefinieerd in Docsis, vergeleken met Euro-Docsis.
De data-link laag specifieert o.a. een Medium Access Control (MAC) protocol, dat de toegang regelt op het gedeelde medium voor de upstream toegang.
Versie 1.1 (april 1999, twee jaar na versie 1.0) bevat belangrijke verbeteringen wat betreft cryptografie en Quality of Service (QoS) ondersteuning die nodig is voor bijExample Voice over IP (VoIP).
Docsis 2.0 (januari 2002) bevat voornamelijk verbeteringen wat betreft de modulatiemethodes, die hogere bitrates mogelijk maken.
De huidige standaard, versie 3.0, kent onder meer hogere datasnelheden door het parallel gebruiken van downstreamkanalen (tot 340 Mb/s), genaamd channel bonding. Bovendien is er ondersteuning voor IP multicast en kan IPv6 gebruikt worden op dit type modems.

Telenet HFC network architecture

In 2023:
51 headends
54 000 KM coax
12 KM fiber
7000 optical nodes



Source: Vlaamse Media Regulator (rapport kostentoerekening ‘21-’23)

The active components for both Internet, fixed telephone and television (analogue and digital) are located in the headend locations. These headends are master facilities for receiving and transmitting signals for the processing and distribution of a cable system. Telenet currently possesses 51 headends, including the ones from the S F acquisition (Brussels and Rance in the region Mons). The length of the coax cables equals 54,432km, while the fiber network consists of 12,340km fiber cables. The HFC network is a mix of the owned network by Telenet and the network controlled and maintained by Fluvius (former Interkabel).

The headend facilities are connected to the secondary ring while supporting appliances are typically held in a data centre. Oneor multiple fiber rings flow out of a headend.

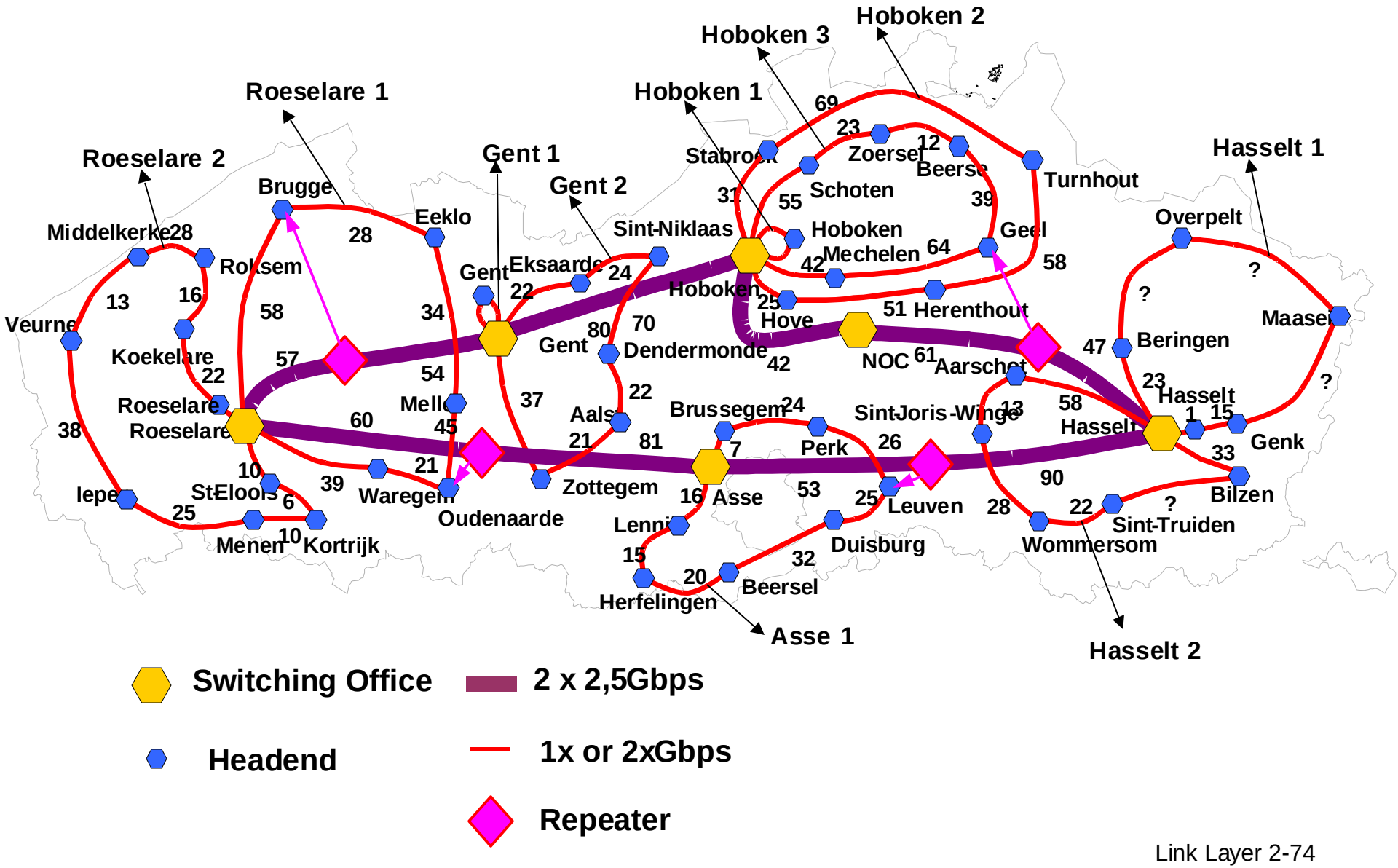
Optical nodes, located on these fiber loops, guarantee that signals can be translated from fiber into coax or from optical into electrical, in both down- and upstream directions. Telenet's network consists out of roughly 7000 of these optical nodes. Telenet Fixed access Engineering has designed the HFC fiber rings in a circular way to create redundancy to cover situations where a breach within the rings would occur.

The Coax segment further consists out of trunk and distribution sections. Passive components within the network, such as cables, splitters and outdoor taps will attenuate the strength of the signal, while active components or amplifiers provide a boost to the signal. Trunk amplifiers are connected to trunk coax cables which are typically causing less attenuation compared to distribution cables and can, therefore, cover a greater distance. The distribution section covers the further delivery of the signal to the customers.

The distribution part is foreseen with outdoor taps, these can be installed after or in-between amplifiers. These devices are used to connect clients to the network via drop cable lines. Drop cable lines run from an outdoor tap to the end customer or subscriber and receive signals from the HFC network.
At the home of a subscriber, the drop cable is connected to the Network Interface Unit (NIU). An NIU is a device that serves as a common interface for devices connected with coax within a customer's home or as a link to an outside network without losing signal strength. The NIU has an interactive port for a cable modem (CM) and downstream only ports for a Set-Top Box (STB) or TV.

Source: https://www.vlaamseregulatormedia.be/nl/file/14631/download?token=Ry8gTdTt

Example: Telenet network (2000)



Link layer, LANs: outline

6.1 Introduction, services

6.2 Error detection, correction

6.3 Multiple access protocols

6.X Access network technologies (additional material!)

- **Introduction**
- ***DSL access technology**
- **Cable-based access technology**
- **Fiber-based access technology**

6.4 LANs

6.5 Link virtualization: MPLS

6.6 Data center networking

6.7 A day in the life of a web request

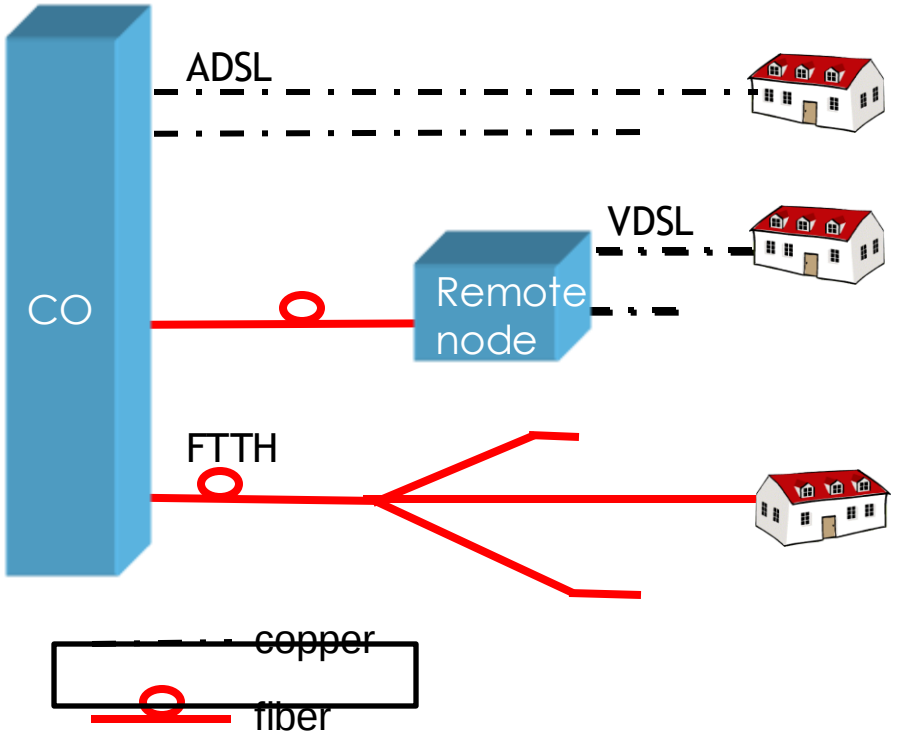
Introduction of fiber in access network

To increase the bitrate provided to customers, the optical fiber part of the network is expanding step by step, coming closer and closer to the user.
(CO = central office)

- Optical fiber in the access network
  - Copper network: shorter copper length, cf. VDSL
  - Cable network: smaller area per optical node

evolution

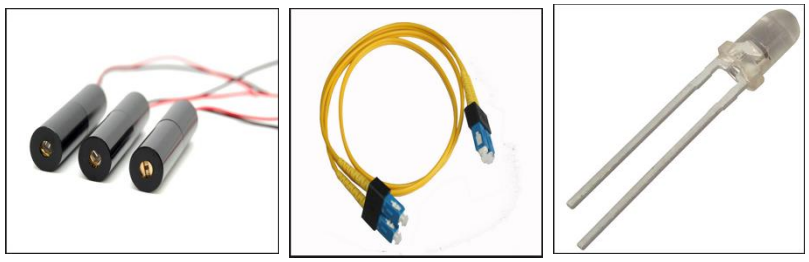
- FTTx: Fiber to the x
  - FTTB: Fiber to the Building
  - FTTH: Fiber to the Home

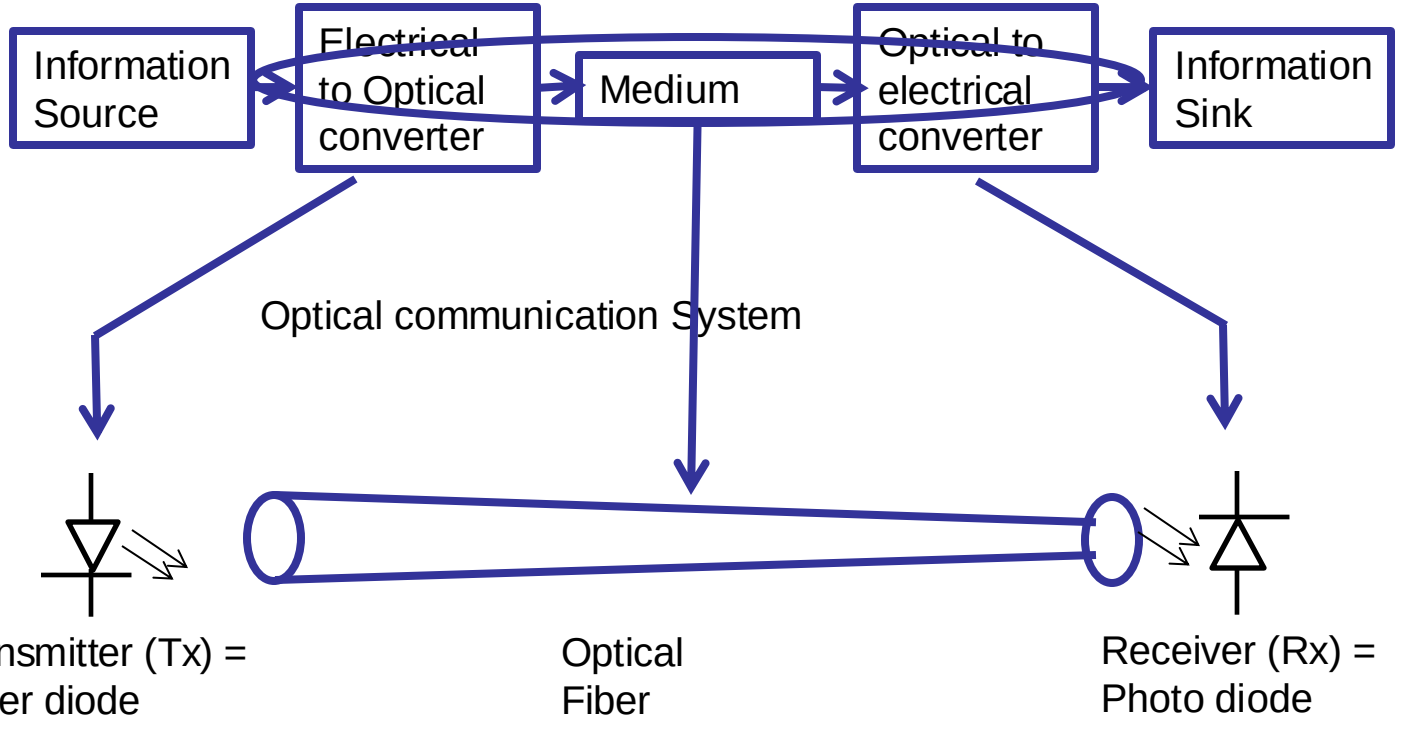


Optical communication

To allow transfer of information through an optical fiber, the electrical signal serves as input signal for a laser diode. The output signal of the laser is an optical signal, modulated according to this input signal. The optical signal is transmitted through an optical fiber and is captured by a photo diode on the other side. This photo diode produces an electrical output signal, modulated according to this incident optical signal. The electrical output signal can then be sent further to the destination.

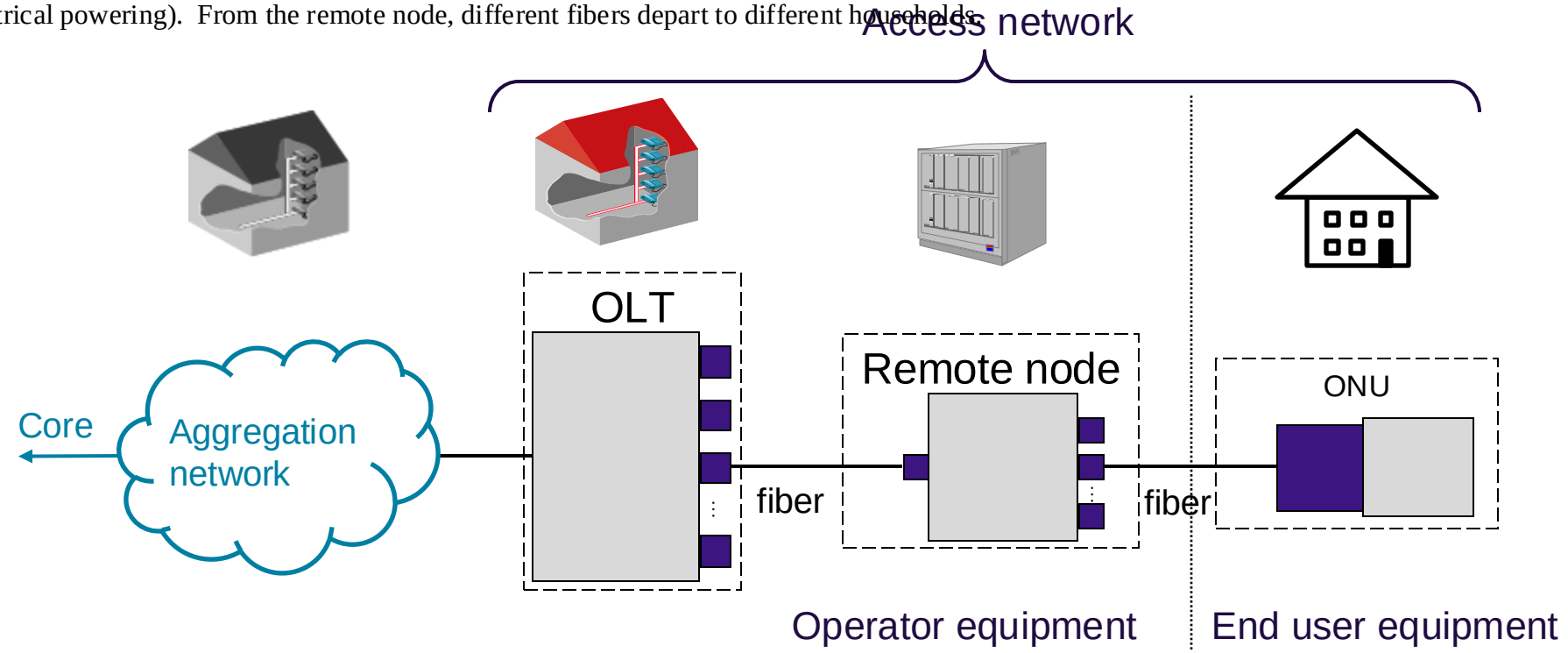
- *How do optical communication networks work?*





Optical fiber access – Fiber to the Home (FTTH)

For instance in the case of Fiber to the Home, an optical fiber is reaching each individual home. The typical design of such an access network is shown above. From the optical line terminal in the central office, different fibers depart, going to different intermediate locations, the 'remote nodes' (for instance located in a street cabinet). Such a remote node can be active (i.e. electrical powering is required) or passive (i.e. no electrical powering). From the remote node, different fibers depart to different households.



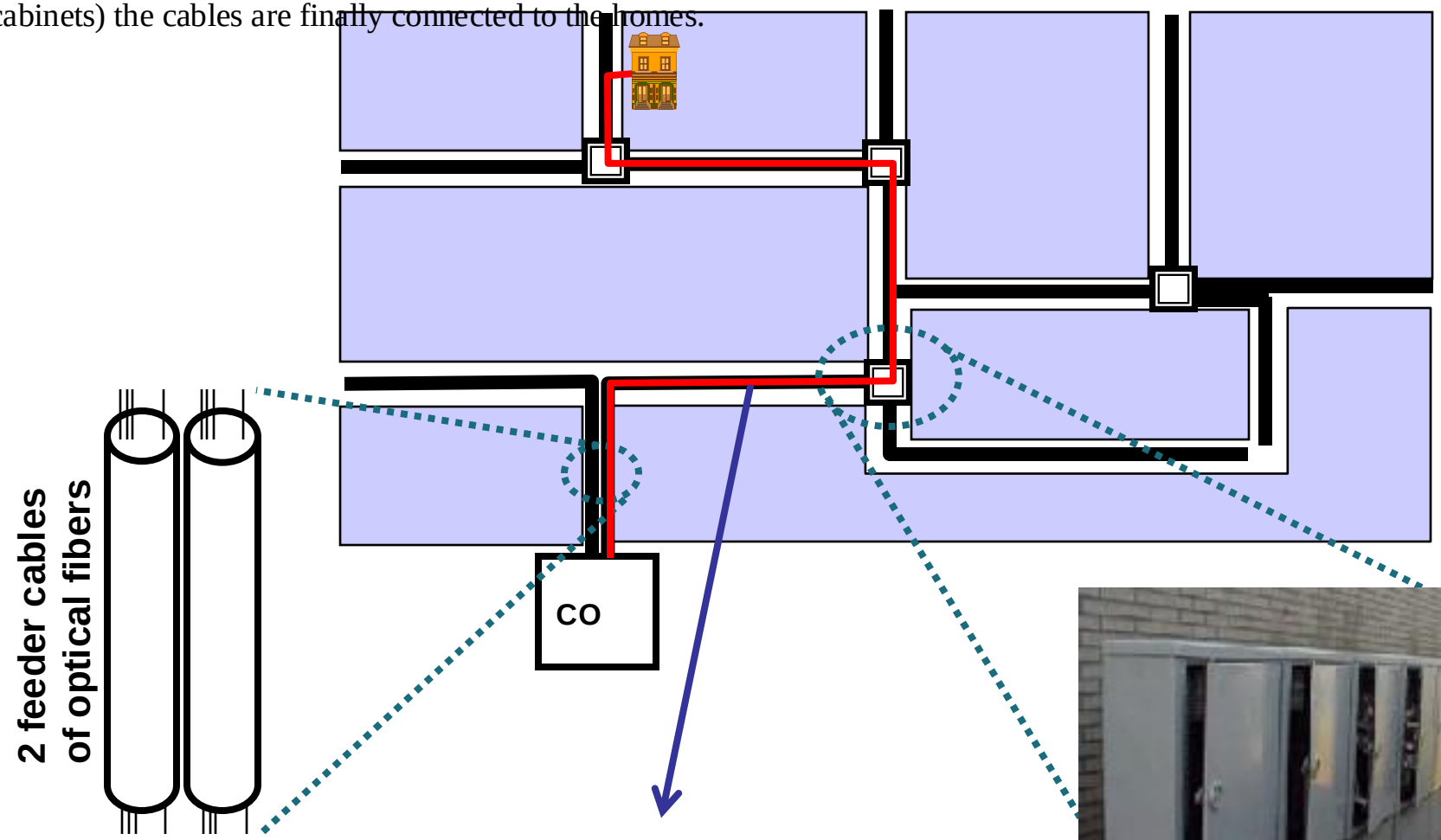
OLT = Optical Line Terminal / ONU = Optical Network Unit

Passive Optical Networks (PONs) ⇔ Active Optical Networks (AONs)

- PON: no active components between OLT and ONU
- AON: active switching / routing between OLT and ONU

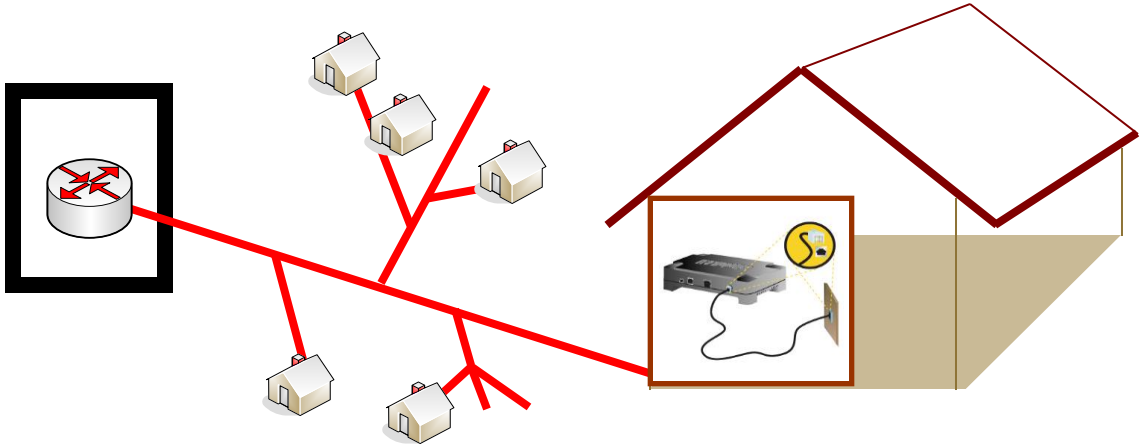
Optical fiber access: topology

The optical fiber access network makes use of cables with optical fibers, in a physical tree and branch structure. Very big cables leave the Central Office (CO) and are further split in smaller cables. In the distribution points (or street cabinets) the cables are finally connected to the homes.



- **Optical fiber**
- **Tree and branch**

Fiber to the Home - OLT

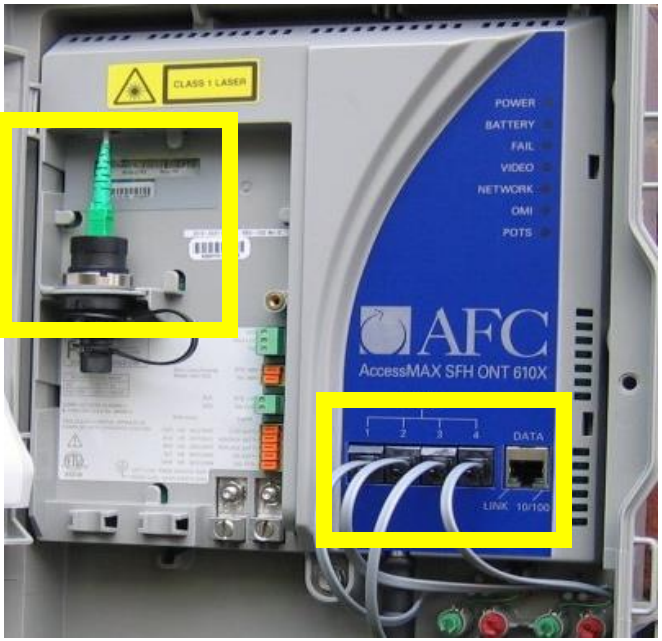


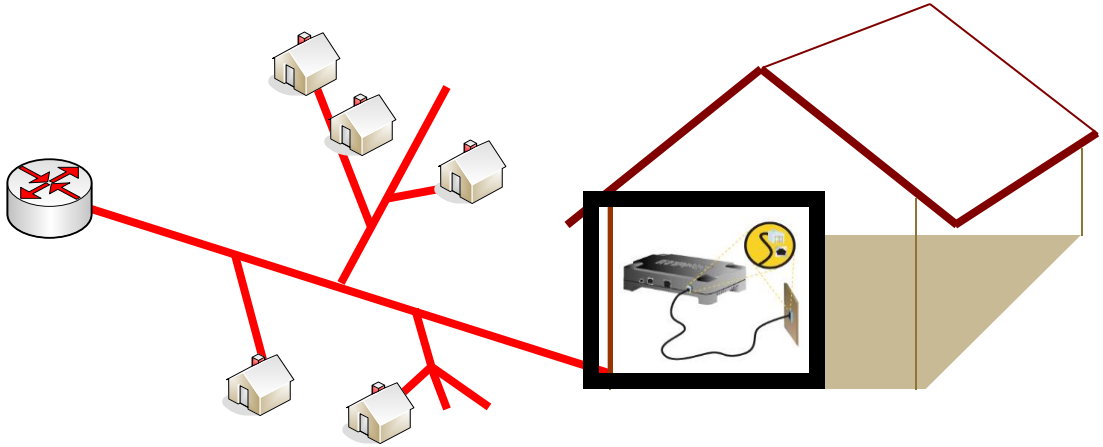
Network operator (central office (CO))
Optical Line Terminal (OLT)

The main building blocks of an FTTH are shown in the current and next slides.

In the central office, an optical line terminal is needed, to send the optical signal on the fiber(s) and to receive the incoming optical signals from the end users.

Fiber to the Home - ONU

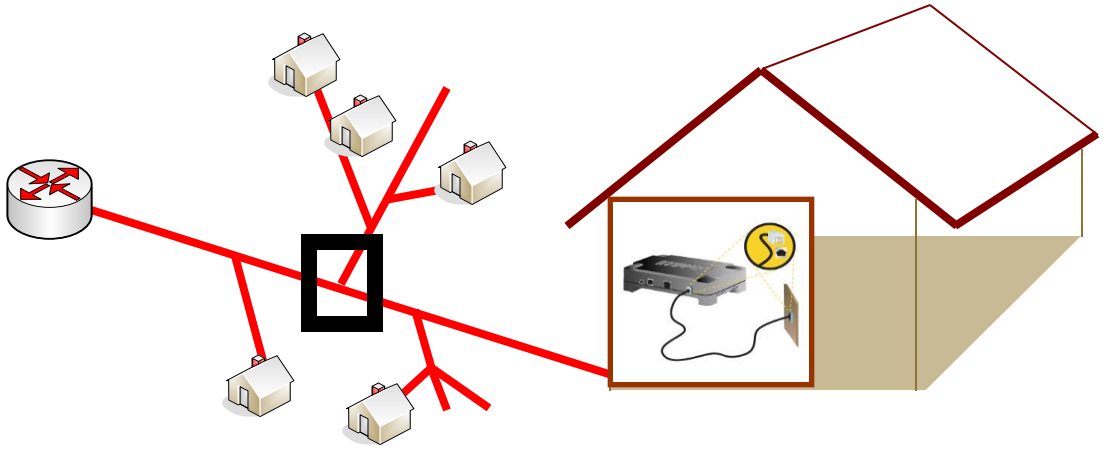




Customer Premises Equipment (CPE)
Optical Network Unit (ONU)

On the other side of the FTTH network, at the customers premises, an optical network unit in installed for termination of the optical signal.

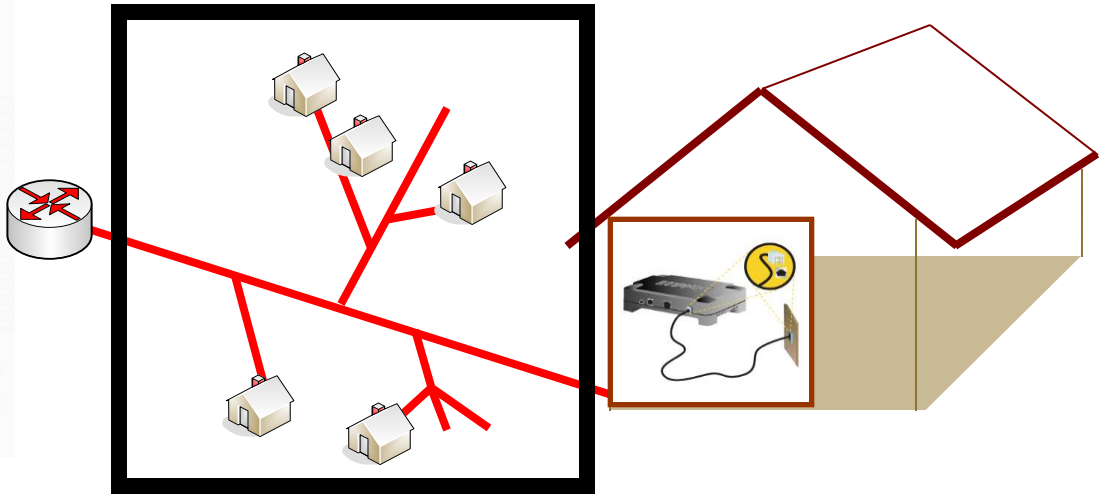
Fiber to the Home – remote node

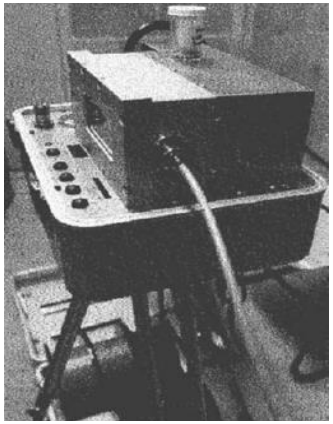
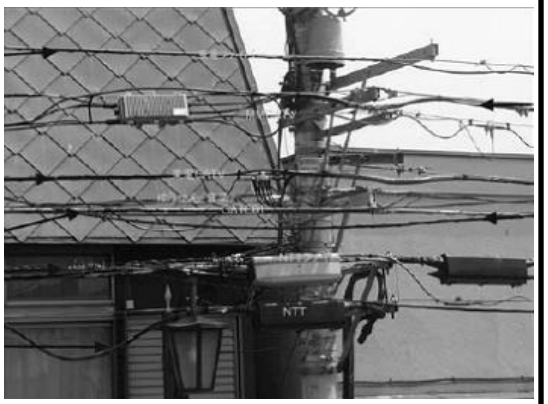


Network operator (street cabinet)
Remote node

At an intermediate location, in the remote node, some (active or passive) switching capability is provided.

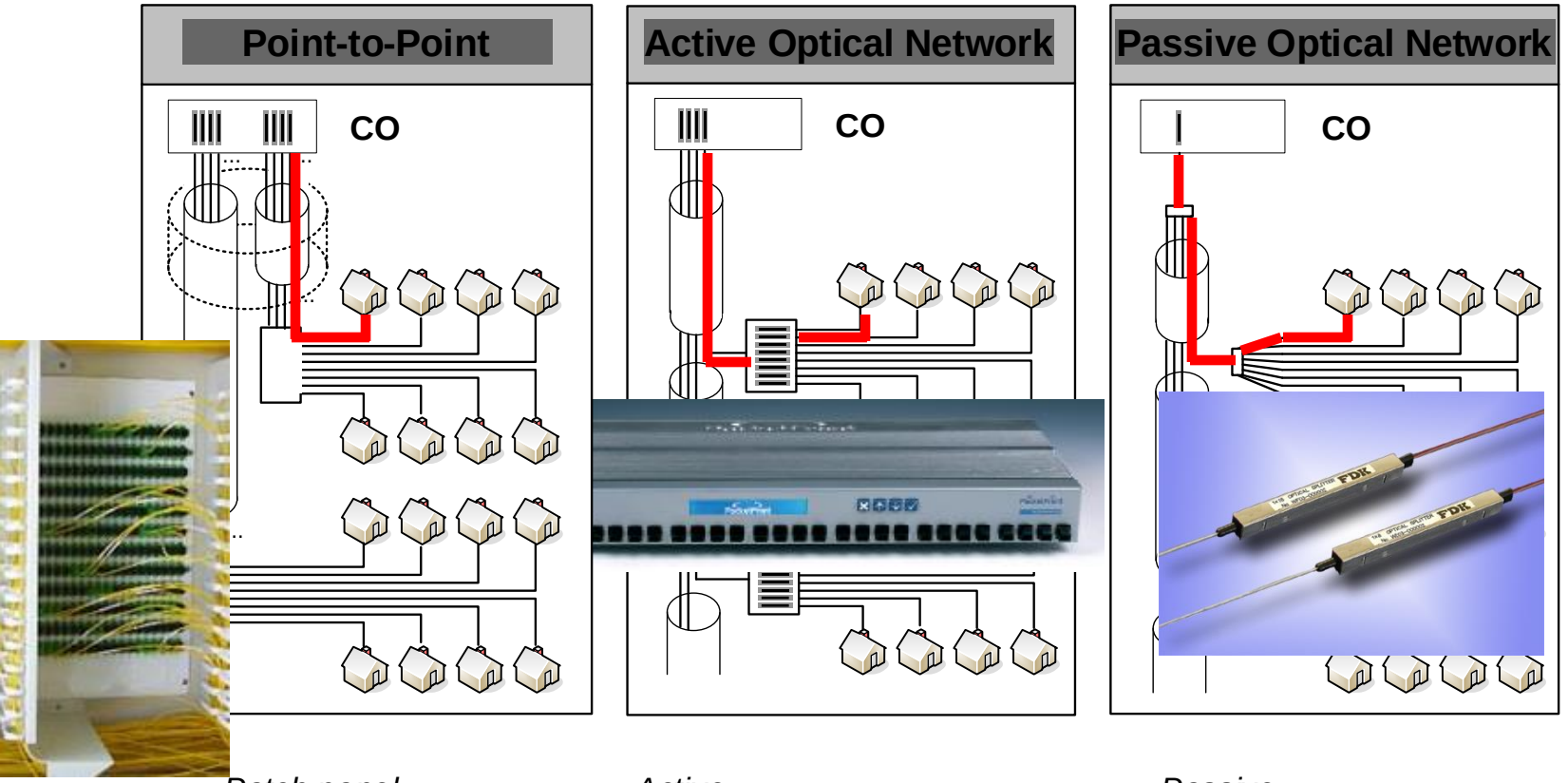
Fiber to the Home - outside



Digging	Blowing / pulling	Aboveground
		

The actual fiber connections typically represent the major investment to build an FTTH network.
In many cases, fiber cables are installed under the foot path, leading to substantial digging costs (e.g. 50 Euro/meter) and burden for the local residents.
If tubes were already installed before (e.g. at previous civil works), fibers can be blown or pulled through these tubes. This solution avoids digging open complete footpaths and hence leads to considerable cost savings.
Another possibility is to attach the fibers to poles or house facades. For instance in Japan this aboveground solution is frequently used. It is much cheaper than underground, but more vulnerable for cuts and less aesthetic.

Fiber to the Home (FTTH) architectures



Patch panel

Active (Ethernet) switch

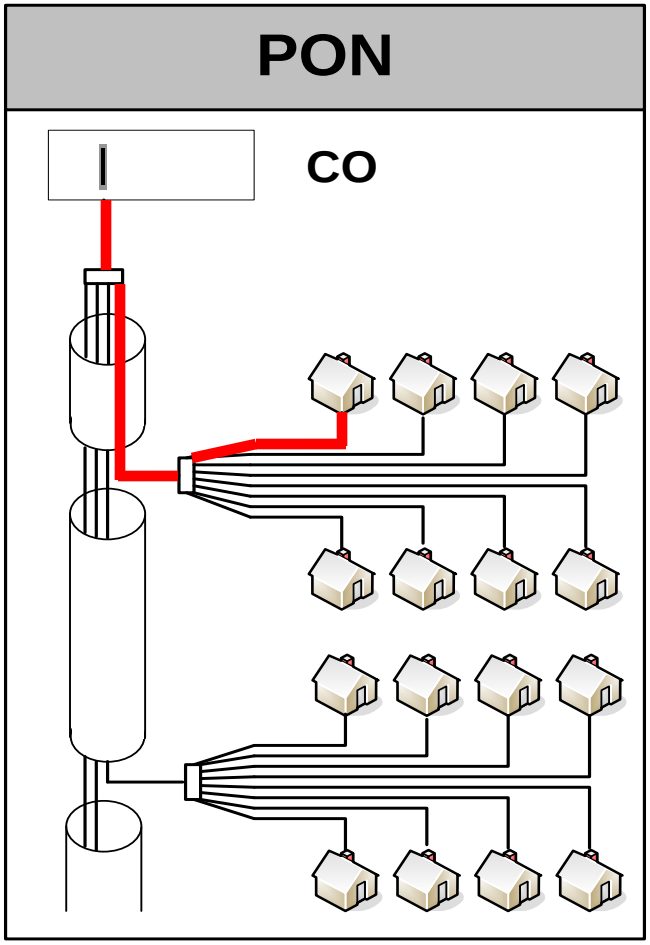
Passive (power) splitter

The figure illustrates 3 major options to provide Fiber To The Home (FTTH):

1. Point-to-Point: every home has a dedicated fiber to the central office (conceptually similar to twisted pair in telephony)
2. Active Optical Network (AON): Every home is connected with a fiber to a switch (e.g. located in a street cabinet). The street cabinet is connected to the central office with a single fiber.
3. Passive Optical Network (PON): The active switch in option 2. is replaced by a passive optical splitter. In the downstream direction the signal from the central office is distributed to all homes. In the upstream direction the signals from the home are combined in the passive 'splitter' (in fact 'combiner' in this direction) and sent to the central office (conceptually similar to CATV network).

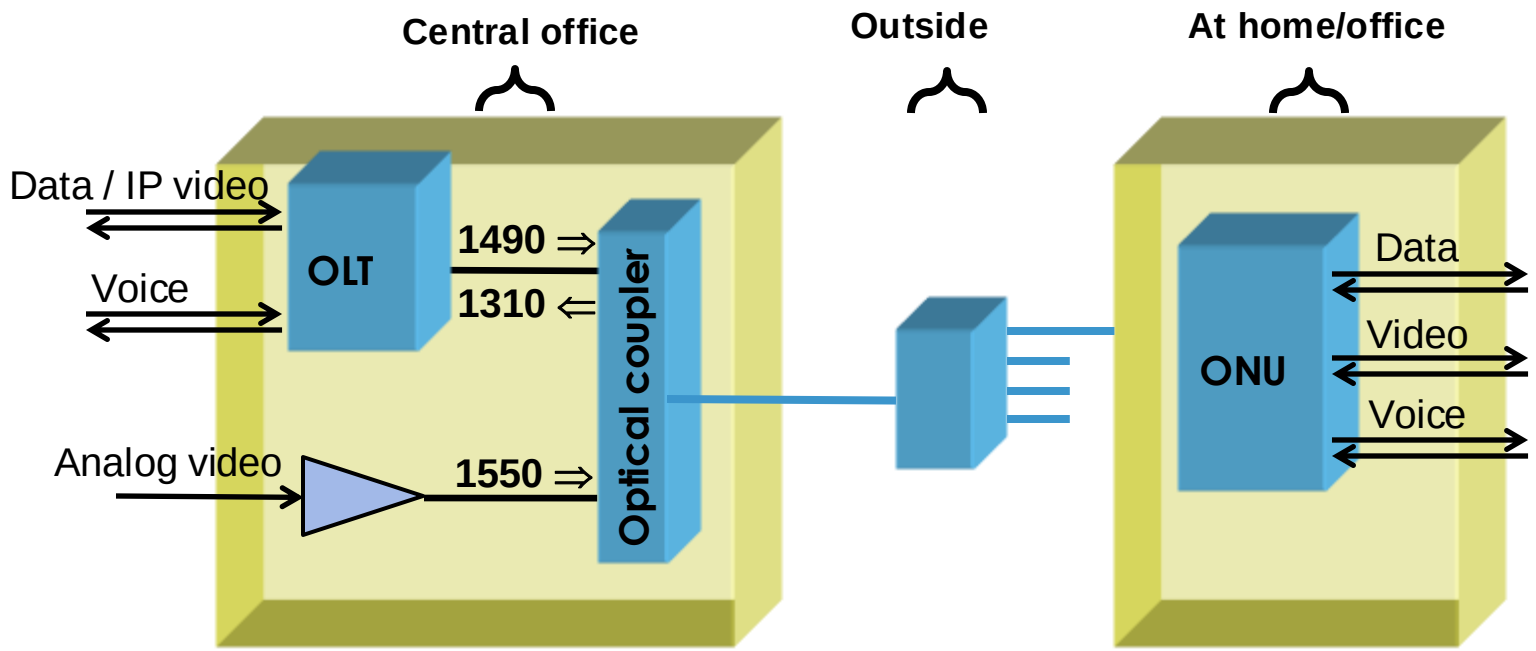
Passive Optical Networks - PON

- PON technology most widespread FTTH solution
- Currently deployed PONs are Time Division Multiplexing (TDM) based
  - Remote node = passive power splitter
  - Shared medium
  - Two (or three) wavelengths:
    - Upstream: 1310 nm
    - Downstream: 1490 nm
    - Analog Video (optional): 1550 nm
  - Bit rates (shared)
    - From 1 Gbps up to 10 Gbps and even 40 Gbps



Passive optical networks are the most widespread FTTH solution. A typical configuration is shown on the slide above (and the next slides).

Deployment example

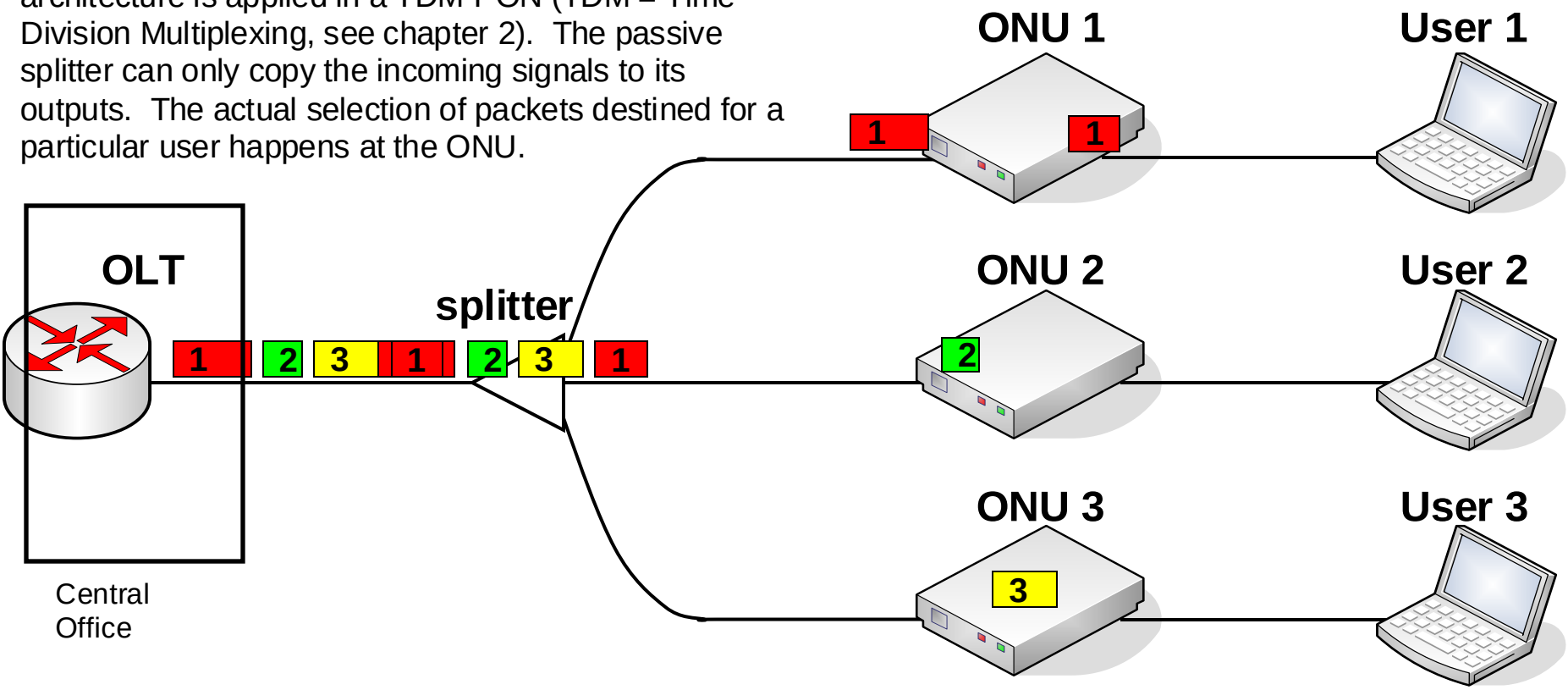


The top figure illustrates a typical architecture including the supported services and interfaces. Note that the PON is using CWDM (Coarse Wavelength Division Multiplexing) using the 1500 nm band for downstream traffic and the 1300 nm band for upstream traffic. An overlay network at 1550 nm is distributing CATV video signals. Note that only one fiber is needed between optical coupler and remote node, as the same fiber is used for both downstream and upstream traffic (at different wavelength bands).

The photo below illustrates the cables connected to poles (as used in Japan).

Downstream traffic TDM-PONs

In the downstream direction, a point-to-multipoint architecture is applied in a TDM-PON (TDM = Time Division Multiplexing, see chapter 2). The passive splitter can only copy the incoming signals to its outputs. The actual selection of packets destined for a particular user happens at the ONU.

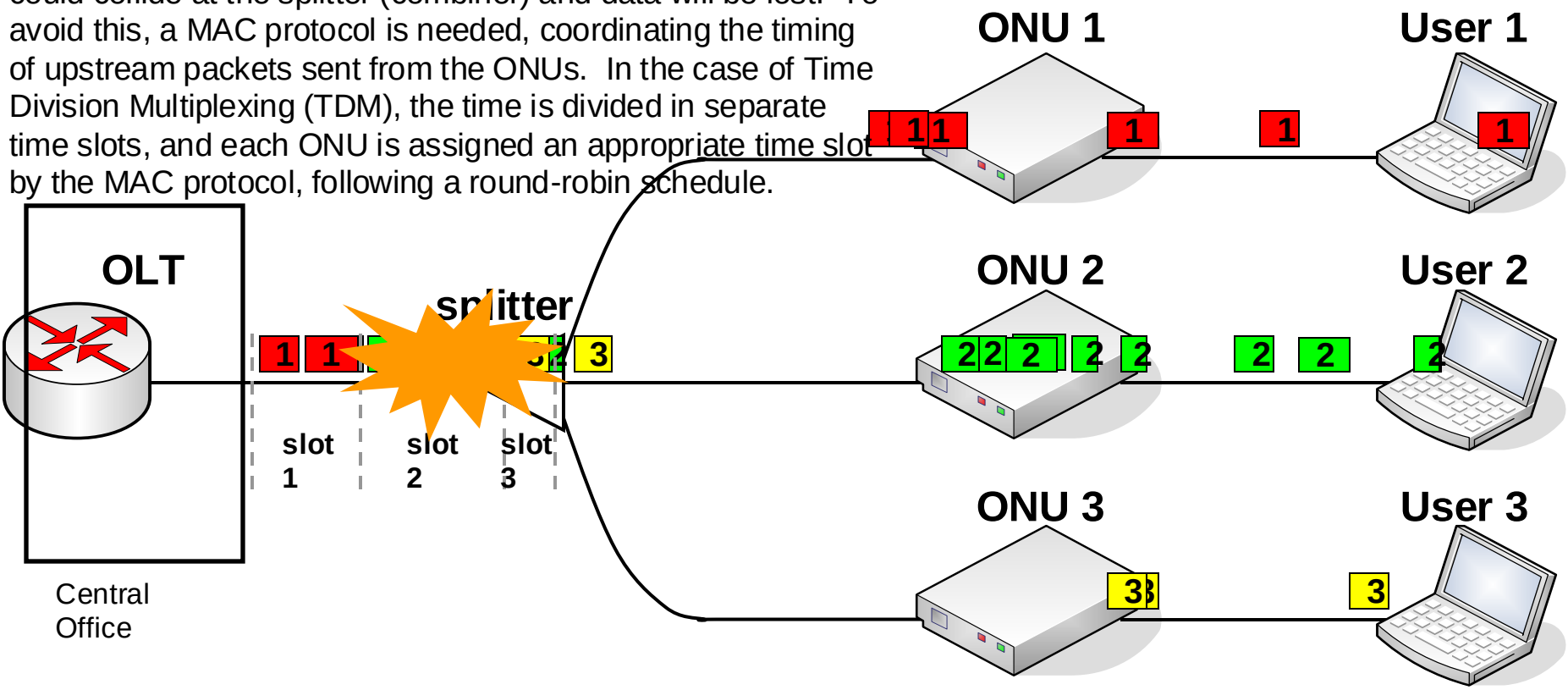


Downstream = point-to-multipoint

➔ Packets filtered at ONU

Upstream traffic TDM-PONs

In the upstream direction, the situation is more complicated. Without coordination, upstream packets from different ONUs could collide at the splitter (combiner) and data will be lost. To avoid this, a MAC protocol is needed, coordinating the timing of upstream packets sent from the ONUs. In the case of Time Division Multiplexing (TDM), the time is divided in separate time slots, and each ONU is assigned an appropriate time slot by the MAC protocol, following a round-robin schedule.

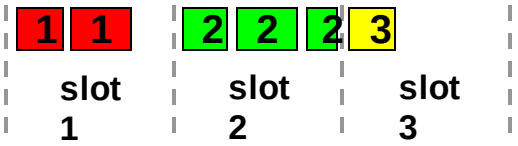


Upstream = multipoint-to-point

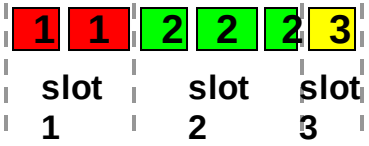
➔ Medium Access Control (MAC) protocol to avoid collisions

How to assign bandwidth to users?

- Fixed bandwidth allocation: assigns equal bandwidth to all users
  - + Simple
  - - Inefficient bandwidth usage, as it does not adapt to the user's load
  - - Increases network delay and resource under-utilization



- Dynamic bandwidth allocation (DBA): assigns bandwidth proportionally to load among the users
  - + Efficient bandwidth usage
  - + Helps in reducing network delays and resource under-utilization
  - - More complex



The most simple solution is to apply fixed length time slots. This results in a fixed bandwidth allocation between ONUs, but leads to inefficiencies.

Having time slots of variable length allows to adapt to the instantaneous load of users (ONUs) and hence improves on efficiency.

Different standards

Standard	Ethernet PtP	EPON	GPON	10G-EPON	XG-PON
Architecture	Point to Point	Point to Multi-point	Point to Multi-point	Point to Multi-point	Point to Multi-point
Downstream data rate	100 M 1 G	1.25 G	1.25 G 2.5 G	10.3 G	10 G
Upstream data rate	100 M 1 G	1.25 G	155 M 1.25 G 2.5 G	1.25 G 10.3 G	2.5 G 10 G
Coverage (without amplifiers)	10 km	10-20 km	20 km	10-20 km	20 km
Split ratio	not applic.	1:16	1:32, 1:64, 1:128	1:16,1:32	1:32, 1:64, 1:128

The table gives an overview of the most important technologies and related standards for passive optical access networks.

PtP = Point-to-point
EPON = Ethernet PON
GPON = Gigabit-capable PON
10G-EPON = 10 gigabit EPON
XG-PON = 10 Gigabit-capable PON

Link layer, LANs: outline

6.1 Introduction, services

6.2 Error detection, correction

6.3 Multiple access protocols

6.4 LANs

- **addressing, ARP**
- Ethernet
- switches
- VLANS

6.5 Link virtualization: MPLS

6.6 Data center networking

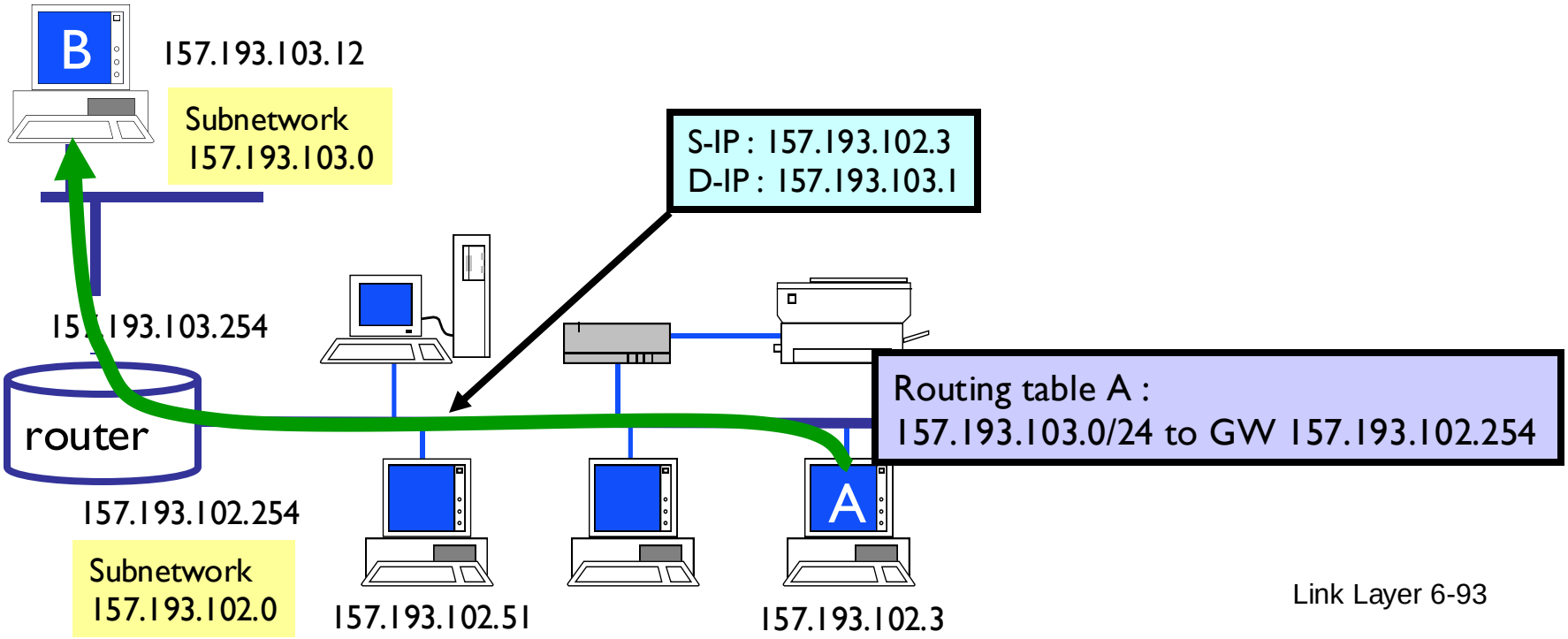
6.7 A day in the life of a web request

Interaction between L3 and L2 forwarding

Problem : 157.193.102.3 (A) wants to send IP packet to 157.193.103.12 (B)
Routing table in A : go to router (157.193.102.254 = gateway router)
Packet leaving A has destination address 157.193.103.12
(and not 157.193.102.254 !!)

Q: How does the packet get to the (gateway) router ?

A: Use of the layer 2 Ethernet address of the router !



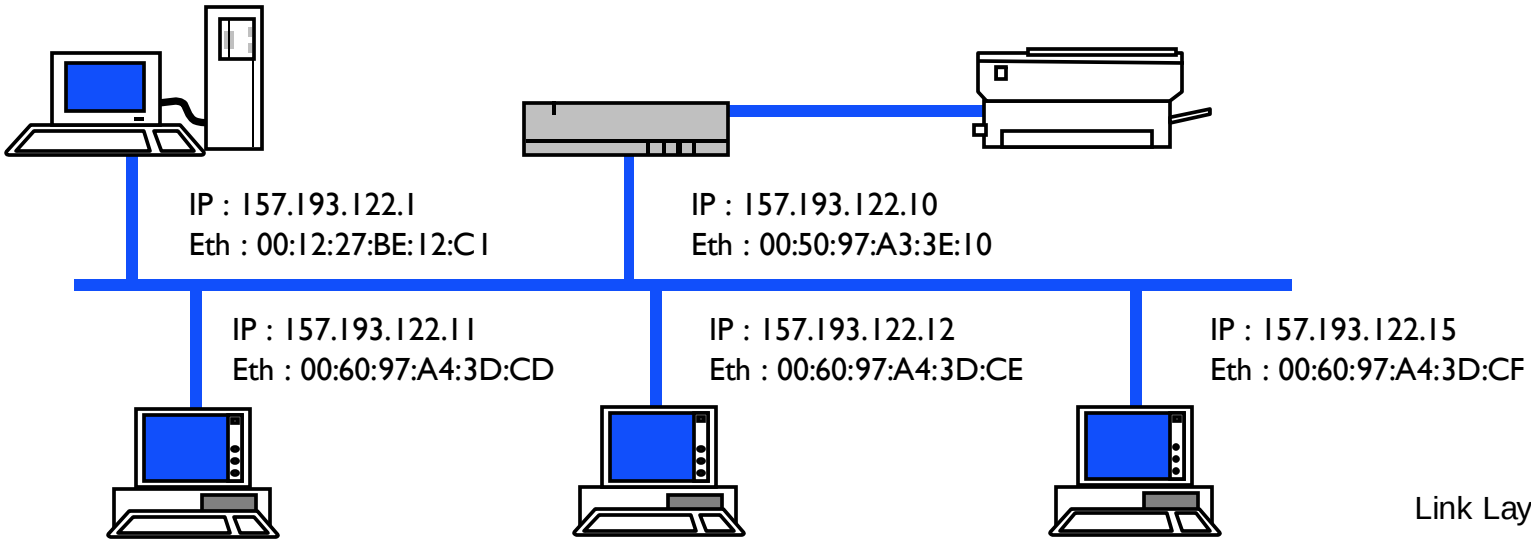
IP address vs. MAC address

32-bit IP address:

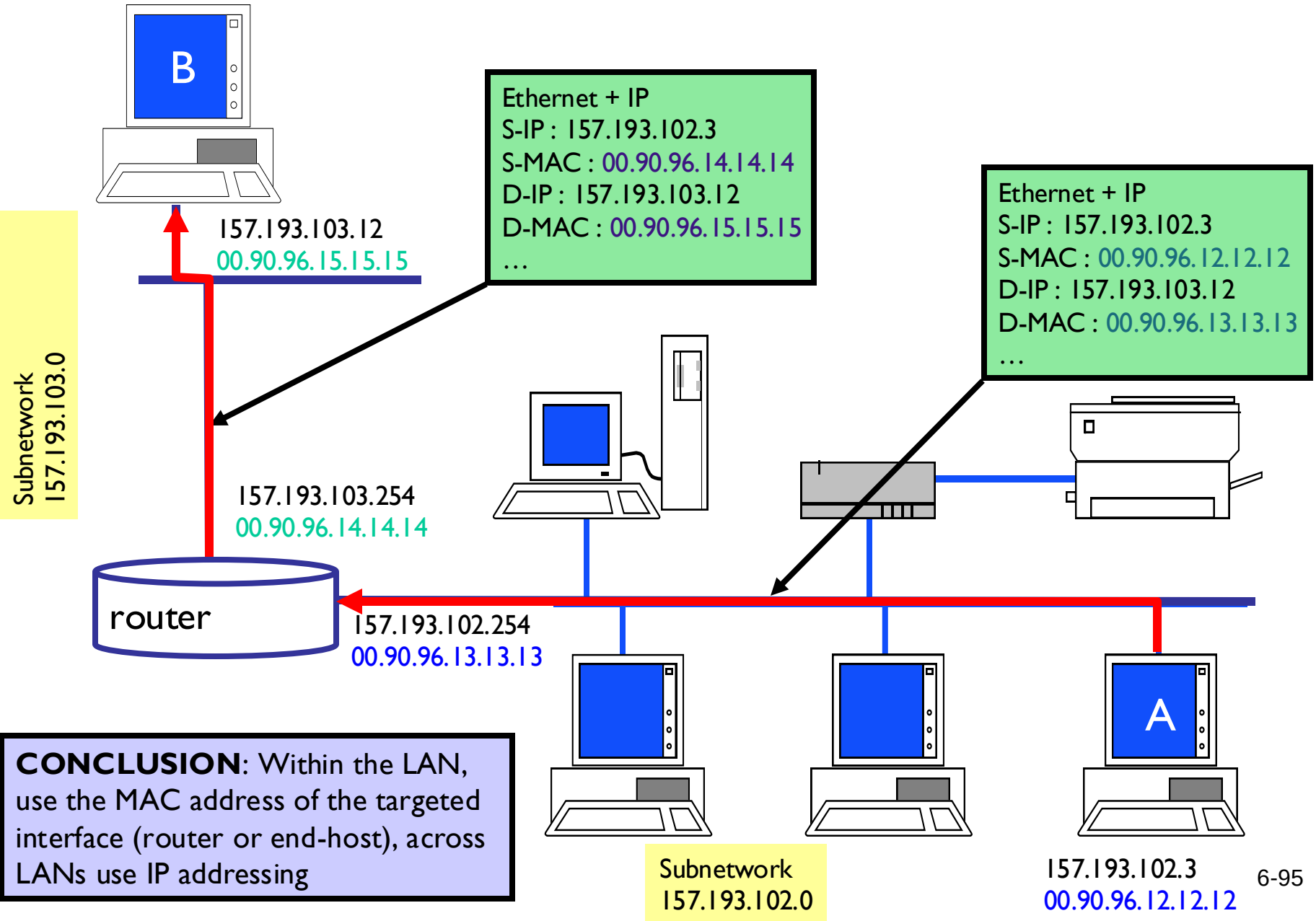
- network-layer address, e.g. 157.193.122.15
- to get packet to destination IP network (globally, across subnetworks)
- hierarchically structured (reflects in which subnet) -> not portable

48-bit MAC address (also physical or LAN or Ethernet address):

- link-layer address, e.g. 00:60:97:A4:3D:CF
- to get packet to other interface on same physically-connected, local network
- flat address structure, burned in the adapter ROM -> portable to other (sub-) network (MAC address space is allocated per vendor by IEEE)

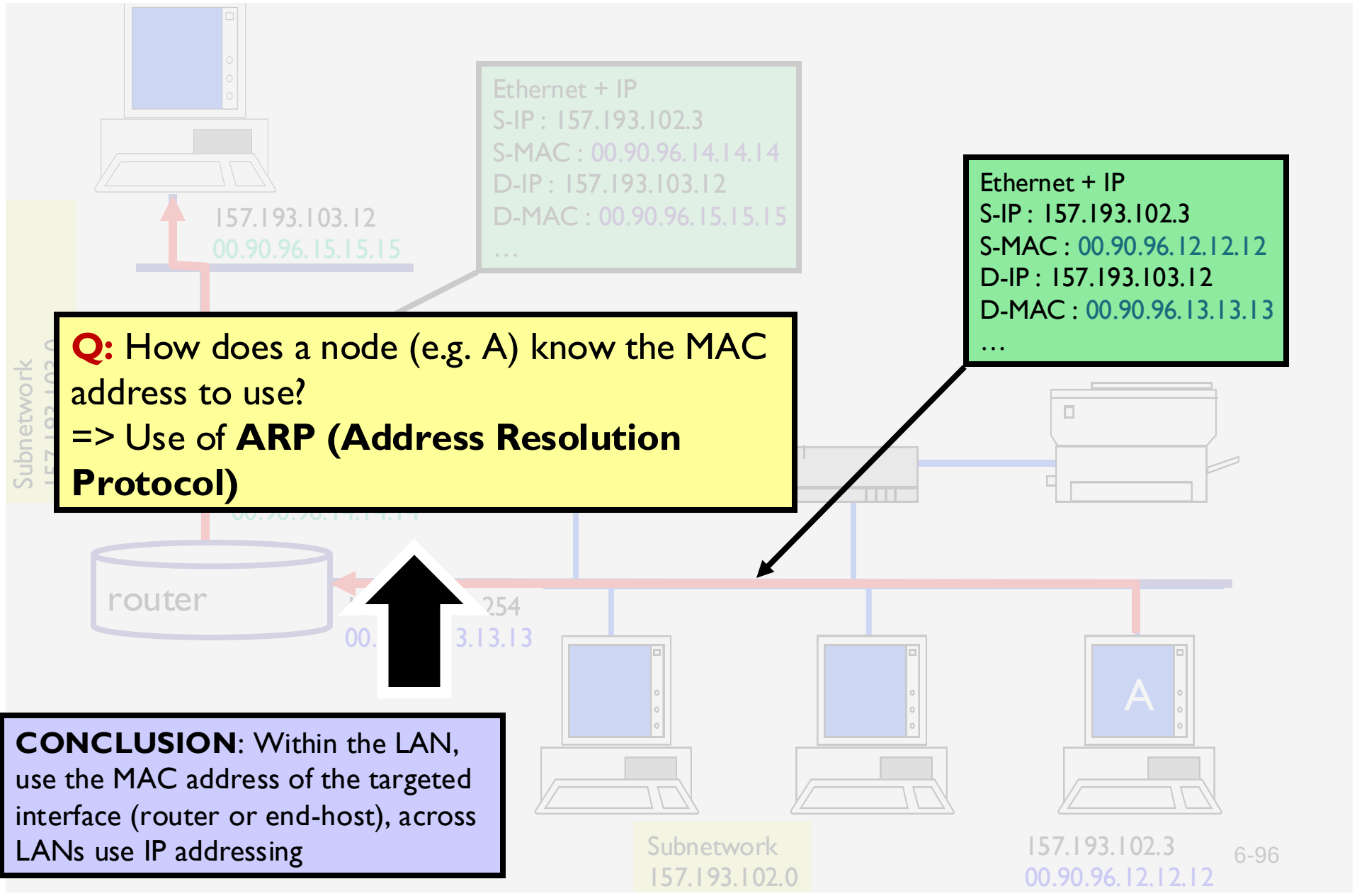


Interaction between L3 and L2 forwarding



CONCLUSION: Within the LAN, use the MAC address of the targeted interface (router or end-host), across LANs use IP addressing

Interaction between L3 and L2 forwarding

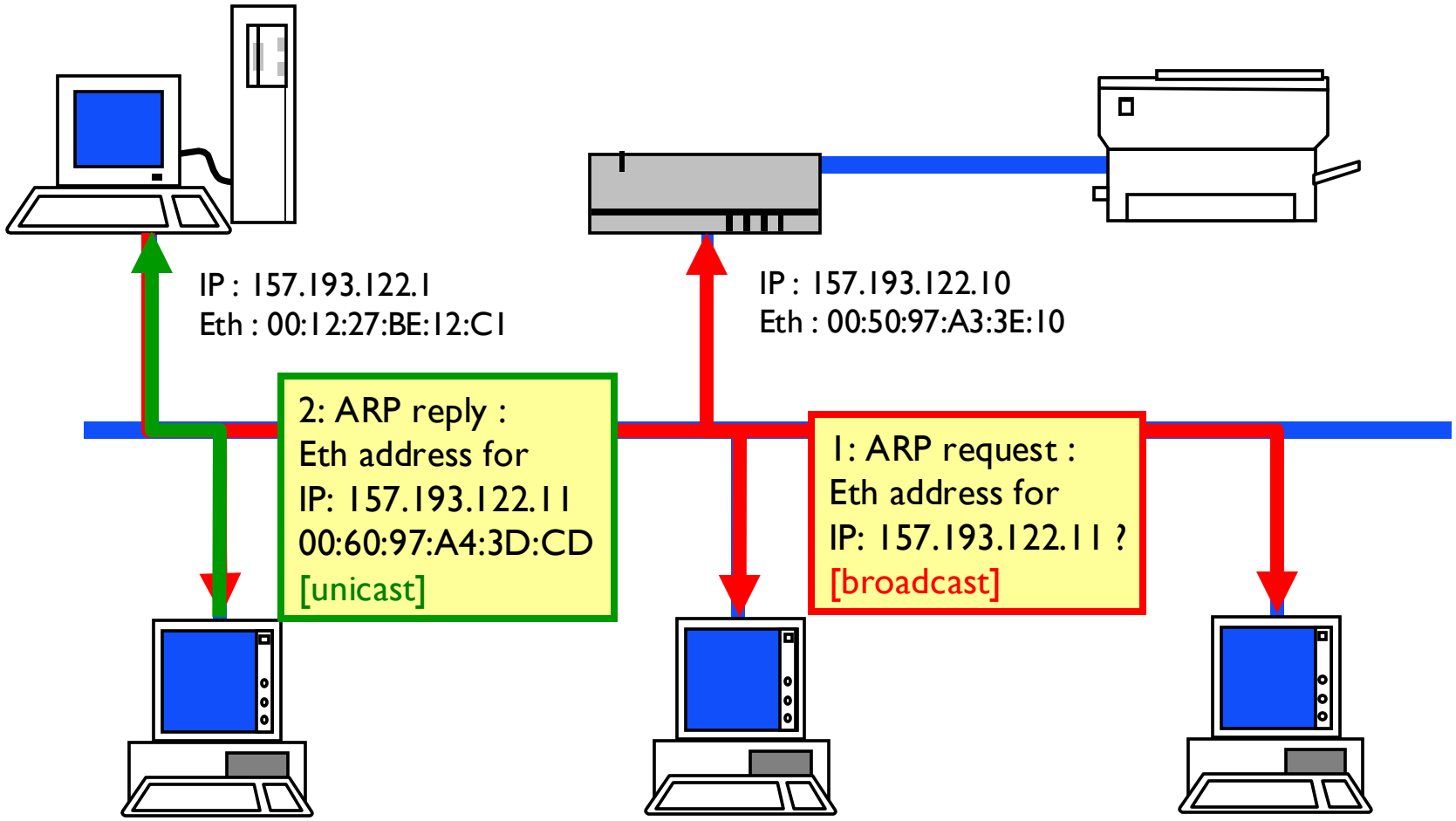


ARP protocol key characteristics

- ARP = *plug & play*, requires no manual administration
- Each node maintains a *local ARP cache/table* with the mapping between IP and MAC address
- Unknown MAC addresses can be *requested using a local LAN broadcast message*
  - Broadcast received by every node on the LAN
- ARP cache entries have a *soft state* (TTL)
  - Times out unless refreshed
  - ***Accept*** a refresh or not ? => chapter Security (ARP spoofing)

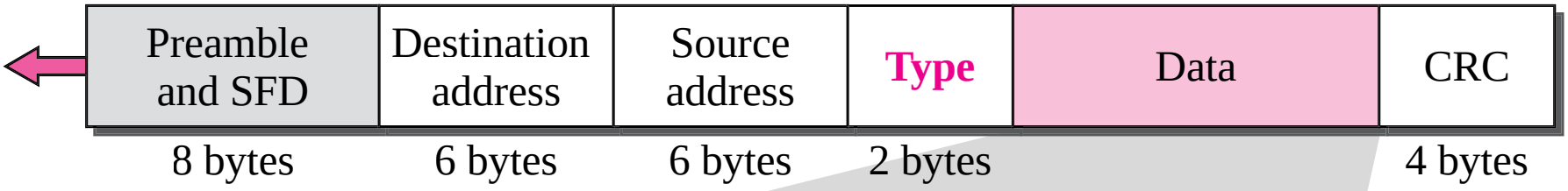
Address Resolution Protocol : ARP

ARP cache:
157.193.122.11 ➔ 00:60:97:A3:3D:CD (TTL = 5 min)



ARP message for Ethernet & IP

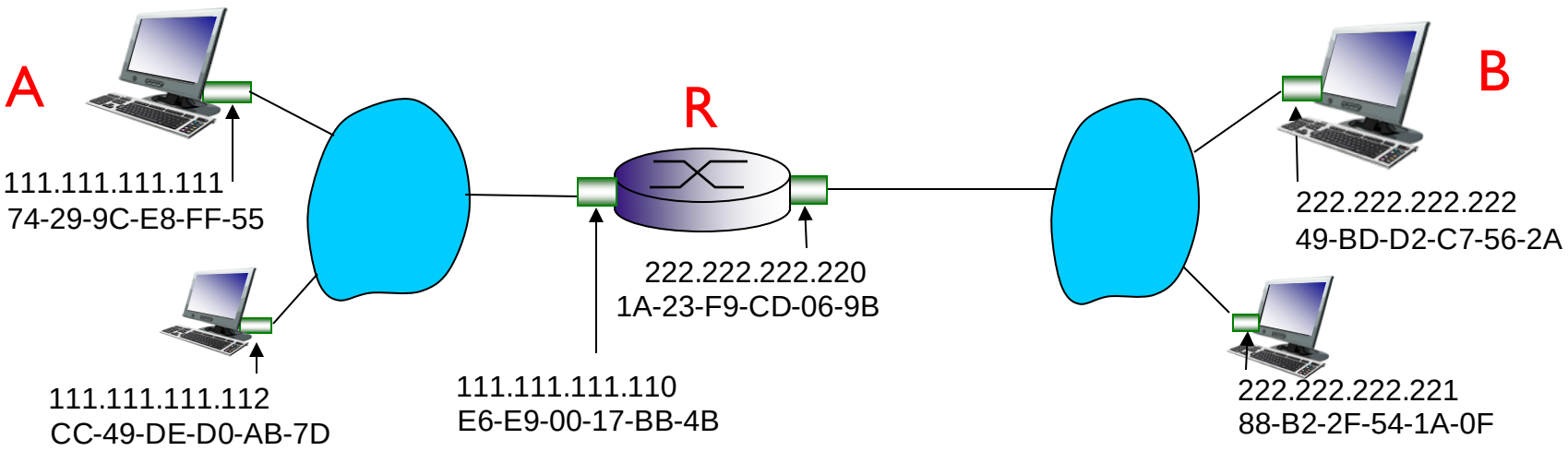
Type: 0x0806



Hardware Type		Protocol Type
Hardware length	Protocol length	Operation Request 1, Reply 2
Sender hardware address (For example, 6 bytes for Ethernet)		
Sender protocol address (For example, 4 bytes for IP)		
Target hardware address (For example, 6 bytes for Ethernet) (It is not filled in a request)		
Target protocol address (For example, 4 bytes for IP)		

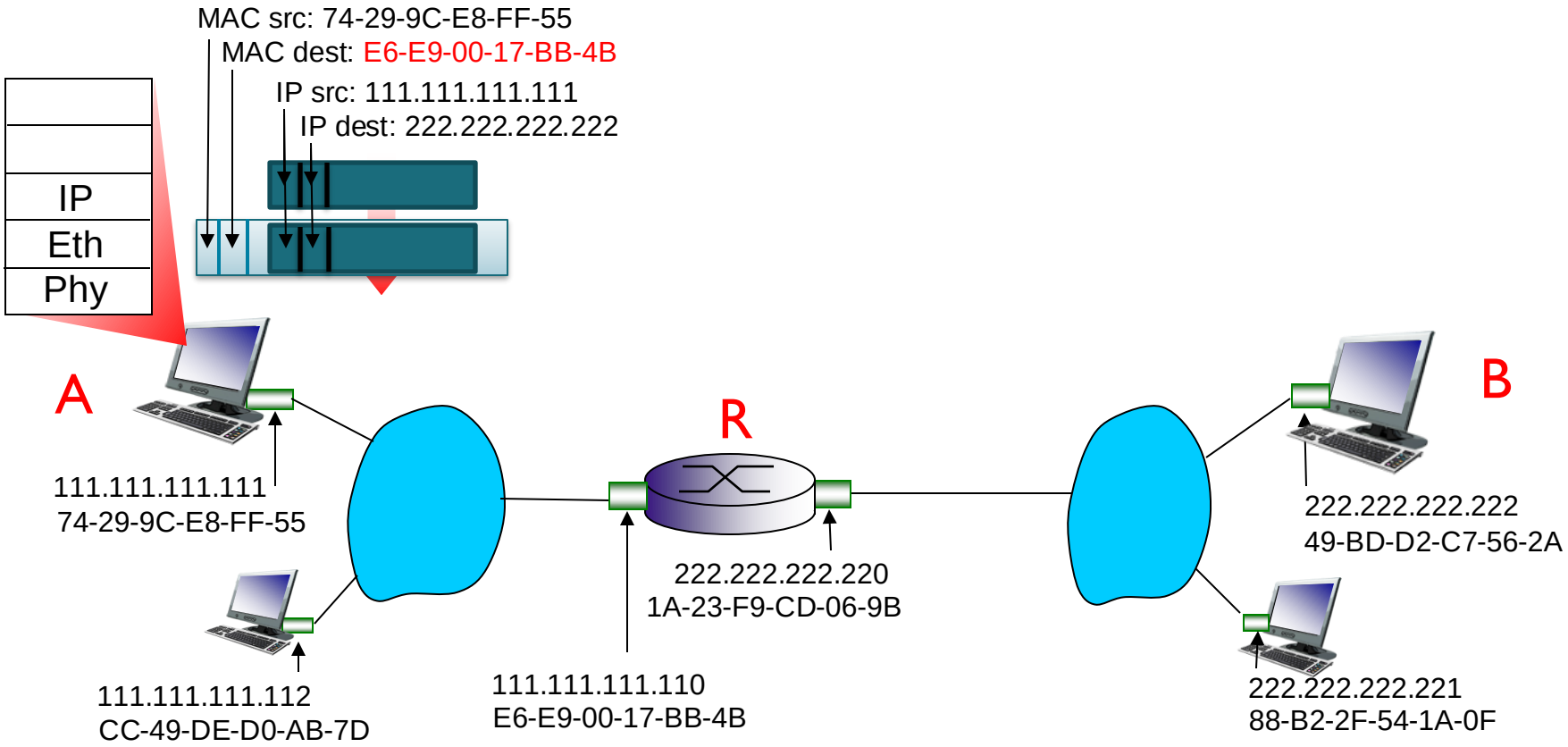
ARP: routing to another LAN

- walkthrough: send datagram from A to B via R
  - focus on addressing – at IP (datagram) and MAC layer (frame)
  - assume A knows B's IP address
  - assume A knows IP address of first hop router, R (how?)
  - assume A knows R's MAC address (how?)



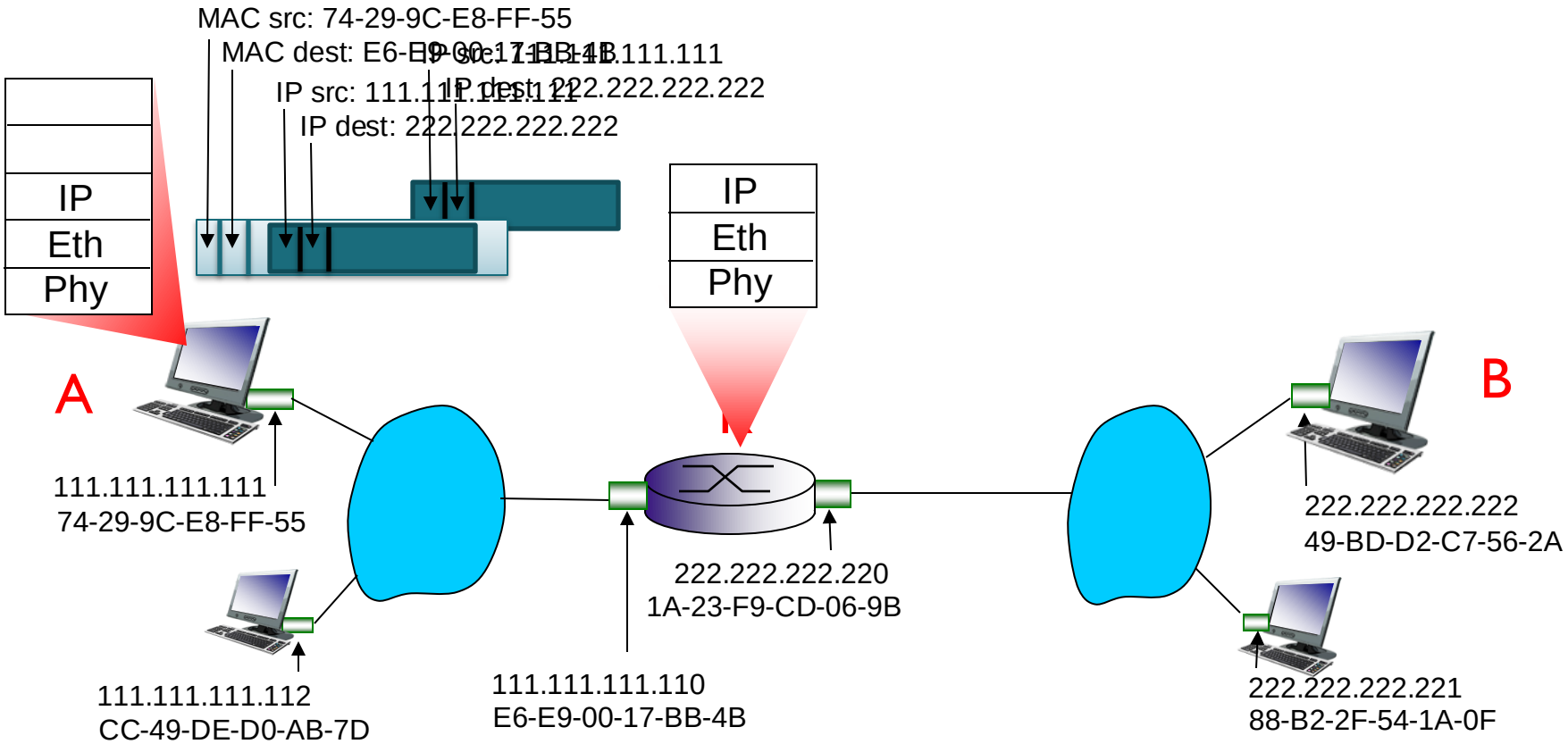
ARP: routing to another LAN

- A creates IP datagram with IP source A, destination B
- A creates link-layer frame with R's MAC address as destination address, frame contains A-to-B IP datagram



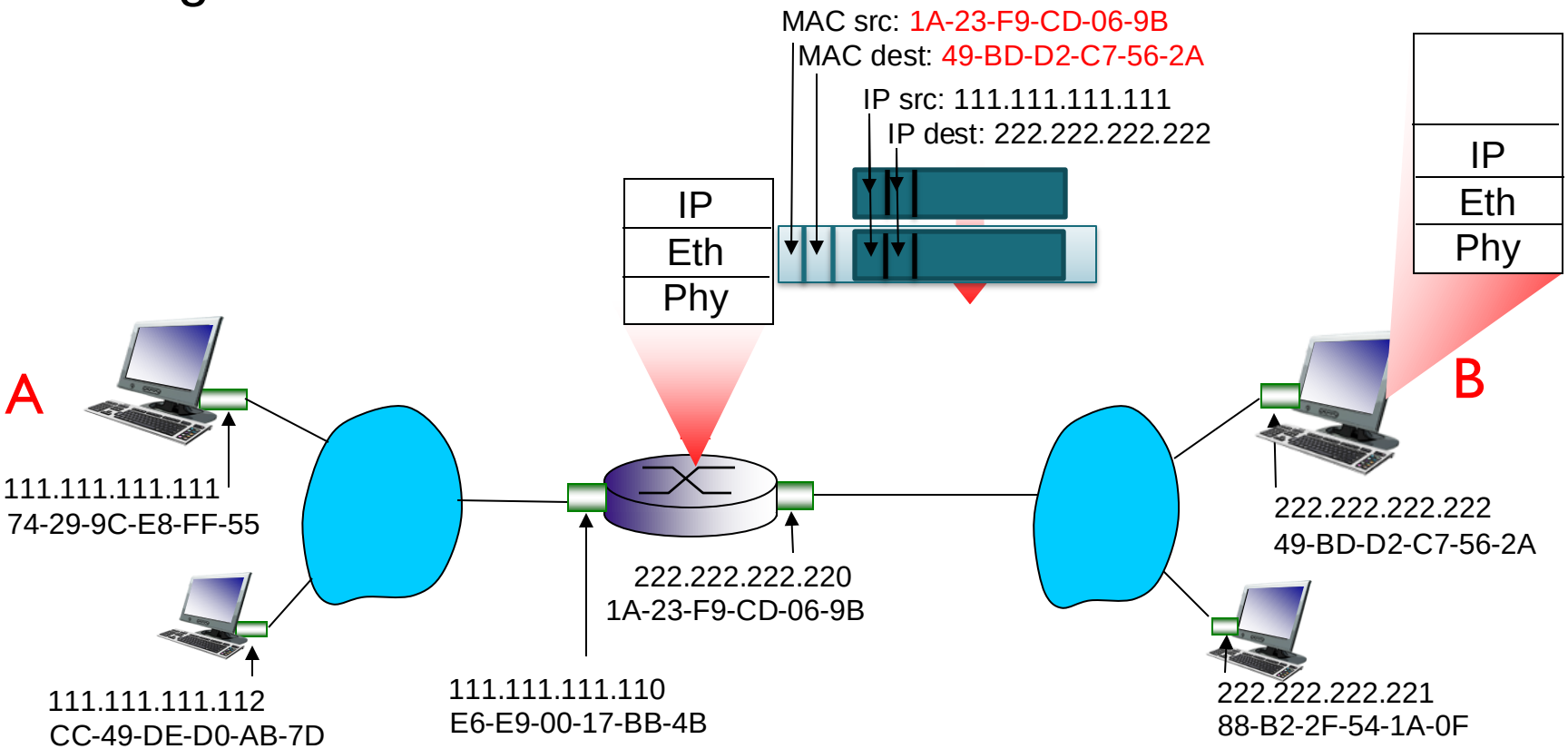
ARP: routing to another LAN

- frame sent from A to R
- frame received at R, datagram header removed, passed up to IP layer



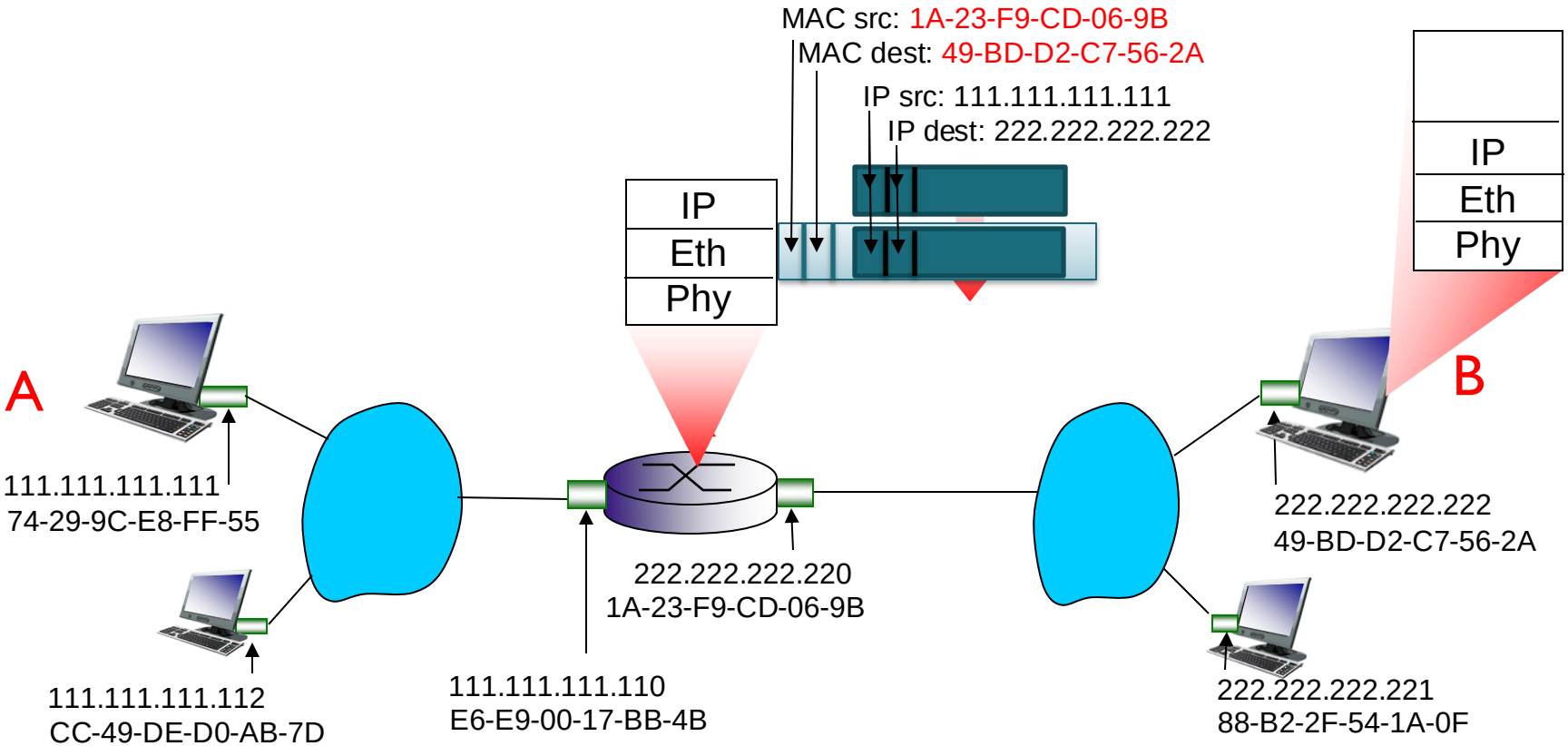
ARP: routing to another LAN

- R forwards datagram with IP source A, destination B
- R creates link-layer frame with B's MAC address as destination address, frame contains A-to-B IP datagram



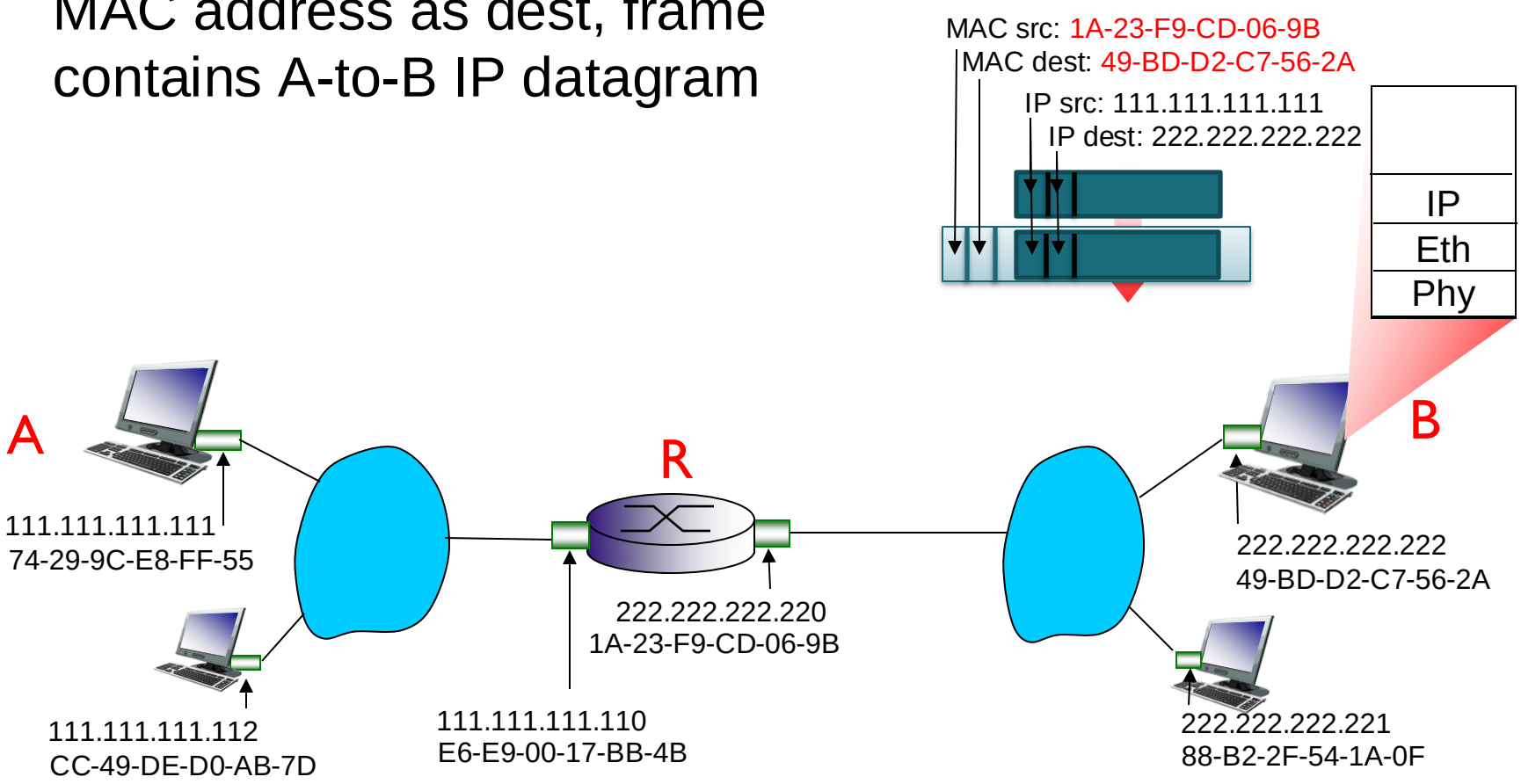
ARP: routing to another LAN

- R forwards datagram with IP source A, destination B
- R creates link-layer frame with B's MAC address as destination address, frame contains A-to-B IP datagram



ARP: routing to another LAN

- R forwards datagram with IP source A, destination B
- R creates link-layer frame with B's MAC address as dest, frame contains A-to-B IP datagram



Link layer, LANs: outline

6.1 Introduction, services

6.2 Error detection, correction

6.3 Multiple access protocols

6.4 LANs

- **addressing, ARP**
- **Ethernet**
- switches
- VLANS

6.5 Link virtualization: MPLS

6.6 Data center networking

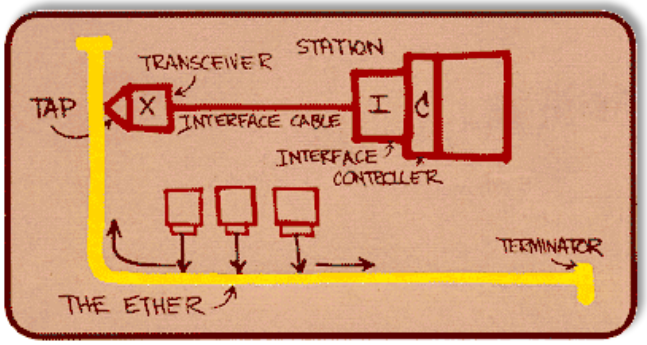
6.7 A day in the life of a web request

Ethernet

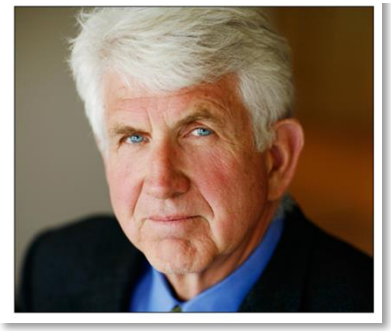
"dominant" wired LAN technology:

- first widely used LAN technology
- simpler, cheap
- kept up with speed race: 10 Mbps – 400 Gbps
- single chip, multiple speeds (e.g., Broadcom BCM5761)

Metcalfe's Ethernet sketch

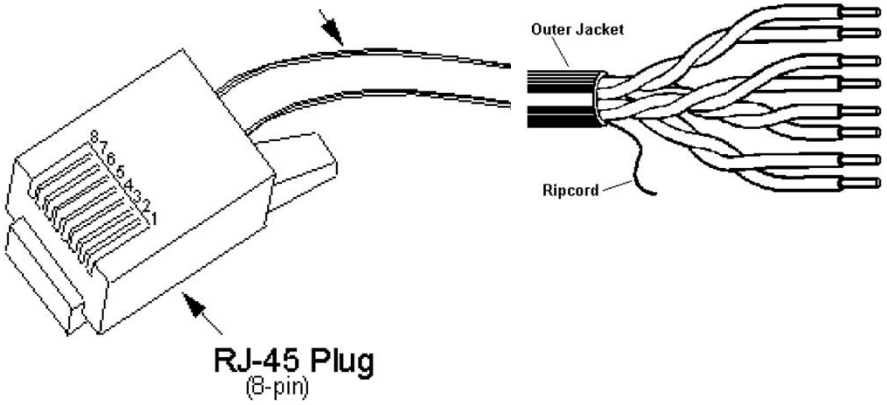


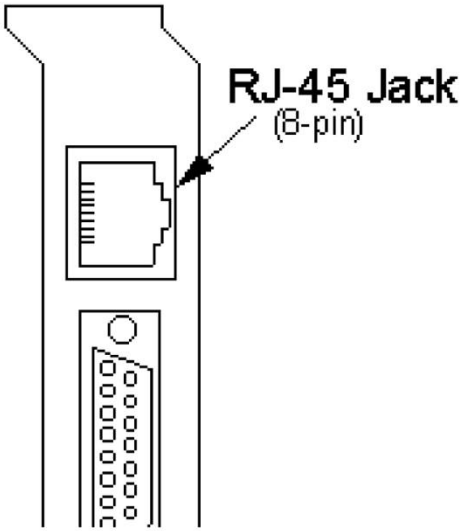
Bob Metcalfe: Ethernet co-inventor,
2022 ACM Turing Award recipient



https://www.uspto.gov/learning-and-resources/journeys-innovation/audio-stories/defying-doubters

Ethernet devices



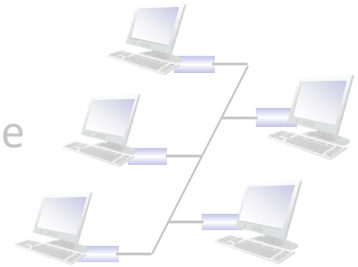


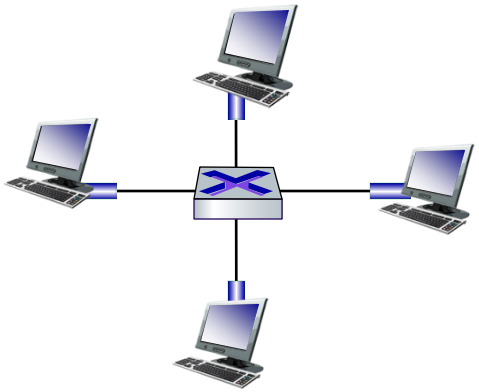
Ethernet physical topology

- bus: popular through mid 90s
  - physical-layer "dumb" repeater
  - bits coming in one link go out all other links at same rate
  - all nodes in same collision domain (can collide with each other)
  - no frame buffering
  - no CSMA/CD at hub: host NICs detect collisions

- switched: prevails today
  - active link-layer 2 *switch* in center
  - each "spoke" runs a (separate) Ethernet protocol (nodes do not collide with each other)

bus: coaxial cable

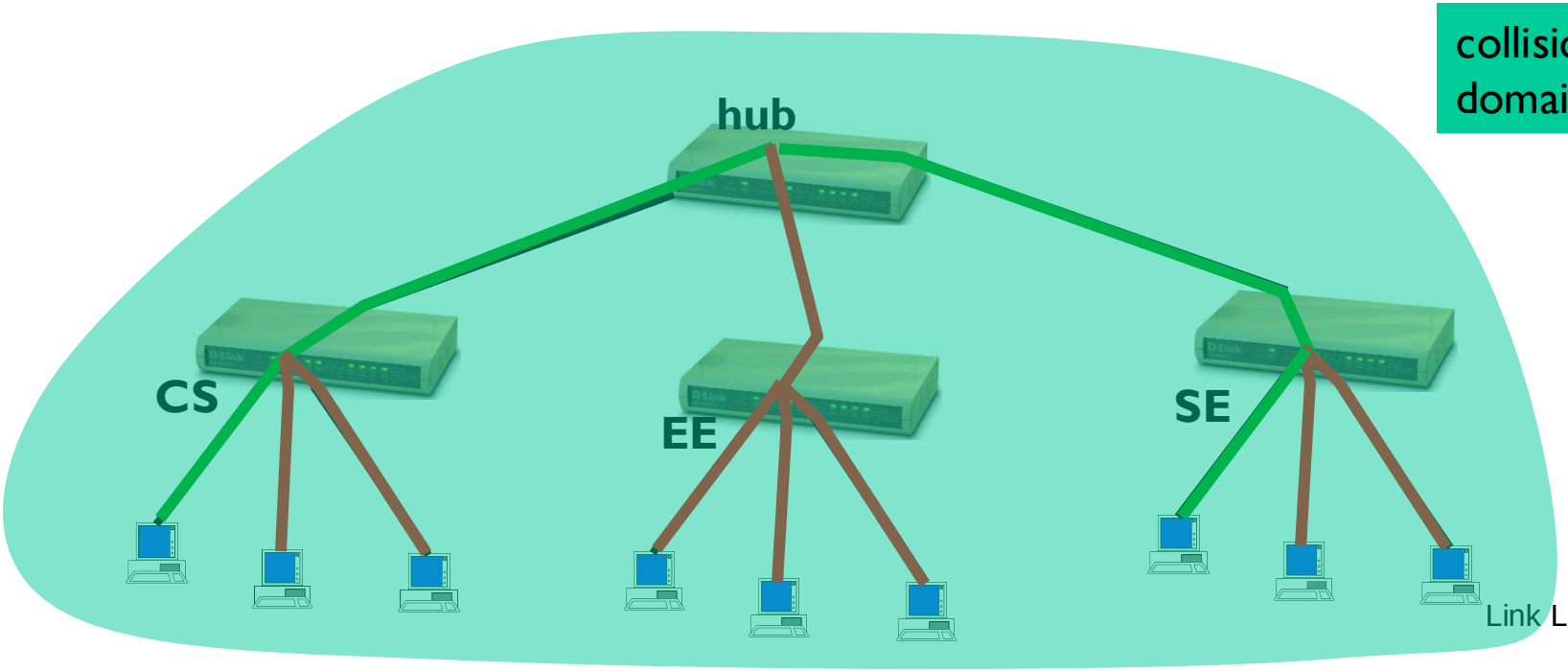




switched

Interconnecting with hubs

- Backbone hub interconnects LAN segments
- Extends max distance between nodes
- But individual segment collision domains become one large collision domain
  - if a node in CS and a node EE transmit at same time: collision
- Can't interconnect 10 Mb/s & 100 Mb/s at same time



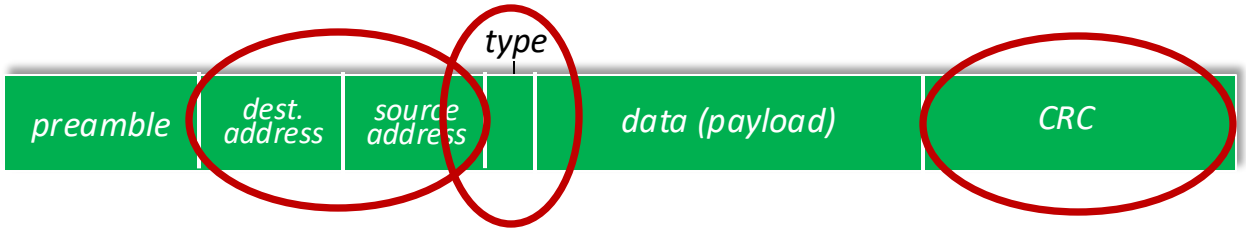
Ethernet frame structure

sending interface encapsulates IP datagram (or other network layer protocol packet) in Ethernet frame

preamble:

- used to synchronize receiver, sender clock rates
- 7 bytes of 10101010 followed by one byte of 10101011

Ethernet frame structure



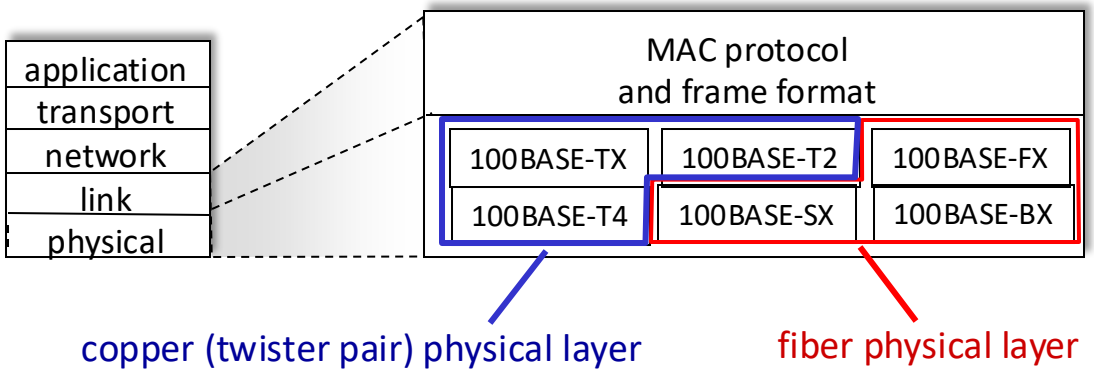
- addresses: 6 byte source, destination MAC addresses
  - if adapter receives frame with matching destination address, or with broadcast address (e.g., ARP packet), it passes data in frame to network layer protocol
  - otherwise, adapter discards frame
- type: indicates higher layer protocol
  - mostly IP but others possible, e.g., Novell IPX, AppleTalk
  - used to demultiplex up at receiver
- CRC: cyclic redundancy check at receiver
  - error detected: frame is dropped

Ethernet characteristics

- connectionless: no handshaking between sending and receiving NICs
- unreliable: receiving NIC doesn't send ACKs or NAKs to sending NIC
  - data in dropped frames recovered only if initial sender uses higher layer rdt (e.g., TCP), otherwise dropped data lost
- Ethernet's MAC protocol: unslotted CSMA/CD with binary backoff

802.3 Ethernet standards: link + physical layer

- *many* different Ethernet standards
  - common MAC protocol and frame format
  - different speeds: 2 Mbps, ... 100 Mbps, 1Gbps, 10 Gbps, 40 Gbps, 80 Gbps
    - different physical layer media: fiber, cable



Link layer, LANs: outline

6.1 Introduction, services

6.2 Error detection, correction

6.3 Multiple access protocols

6.4 LANs

- addressing, ARP
- **Ethernet**
- **switches**
- VLANS

6.5 Link virtualization: MPLS

6.6 Data center networking

6.7 A day in the life of a web request

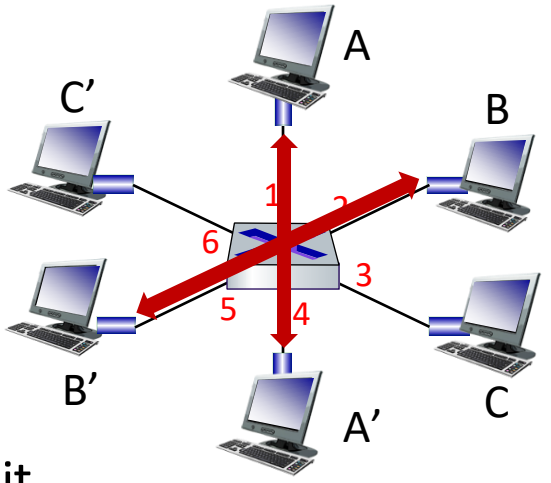
Ethernet Switch

- link-layer device: smarter than hubs, take active role
  - store, forward Ethernet frames
  - examine incoming frame's MAC address, selectively forward frame to one-or-more outgoing links when frame is to be forwarded on segment, uses CSMA/CD to access segment
- transparent
  - hosts are *unaware* of presence of switches (as if connected by a direct cable)
- plug-and-play, self-learning
  - switches do not need to be configured

A switch is sometimes also called "bridge"

Ethernet switch – simultaneous transmission

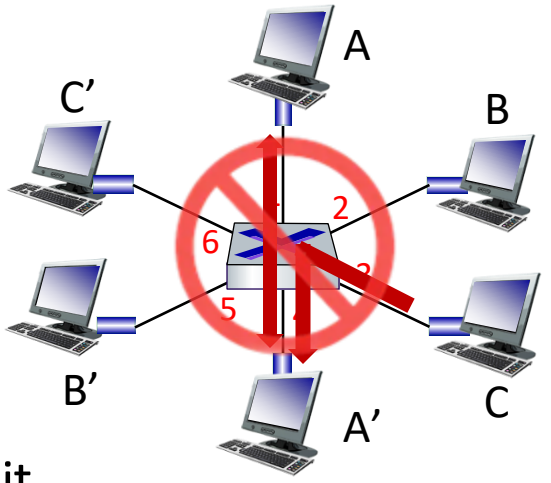
- hosts have dedicated, direct connection to switch
- switches buffer packets
- Ethernet protocol used on *each* incoming link, so:
  - no collisions; full duplex
  - each link is its own collision domain
- switching: A-to-A' and B-to-B' can transmit simultaneously, without collisions



switch with six interfaces (1,2,3,4,5,6)

Ethernet switch – simultaneous transmission

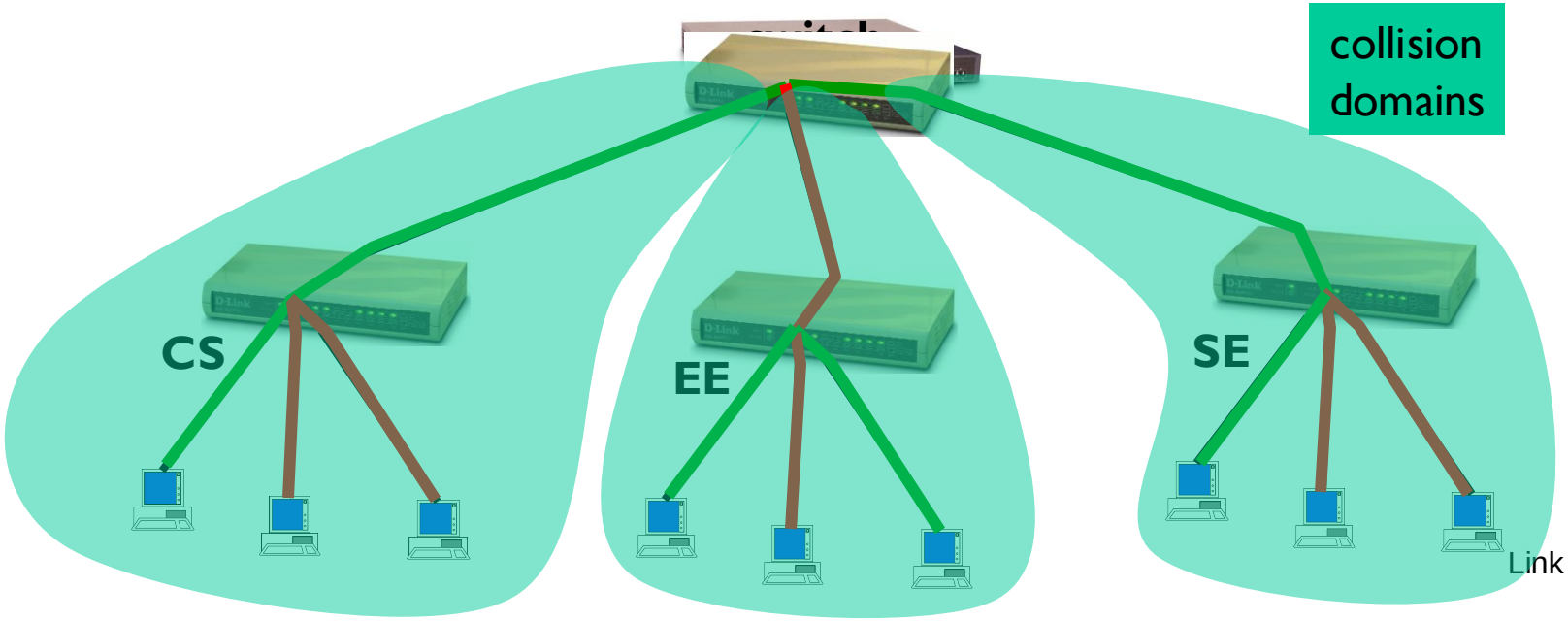
- hosts have dedicated, direct connection to switch
- switches buffer packets
- Ethernet protocol used on *each* incoming link, so:
  - no collisions; full duplex
  - each link is its own collision domain
- switching: A-to-A' and B-to-B' can transmit simultaneously, without collisions
  - but A-to-A' and C to A' can *not* happen simultaneously



switch with six interfaces (1,2,3,4,5,6)

Ethernet Switch - traffic isolation

- Switch installation breaks LAN into LAN segments
- Switch filters packets:
  - same-LAN-segment frames not usually forwarded onto other LAN segments
  - segments become separate collision domains



Ethernet switch - forwarding

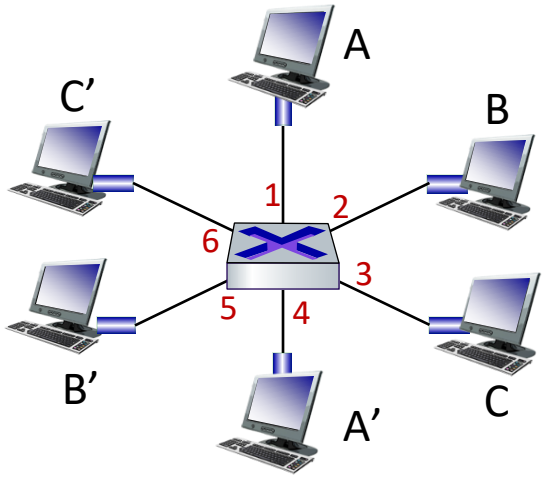
Q: how does switch know A' reachable via interface 4, B' reachable via interface 5?

A: each switch has a switch table, each entry:

- (MAC address of host, interface to reach host, time stamp)
- looks like a routing table!

Q: how are entries created, maintained in switch table?

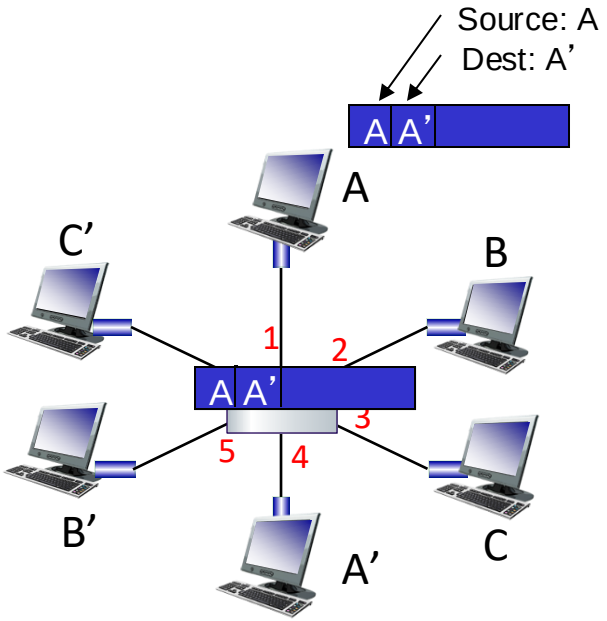
- something like a routing protocol?



Link Layer: 6-120

Ethernet switch - flooding

- switch acts as a *plug & play device* (no manual configuration)
- when booting (powered on), switch tables are empty
- default behaviour = …

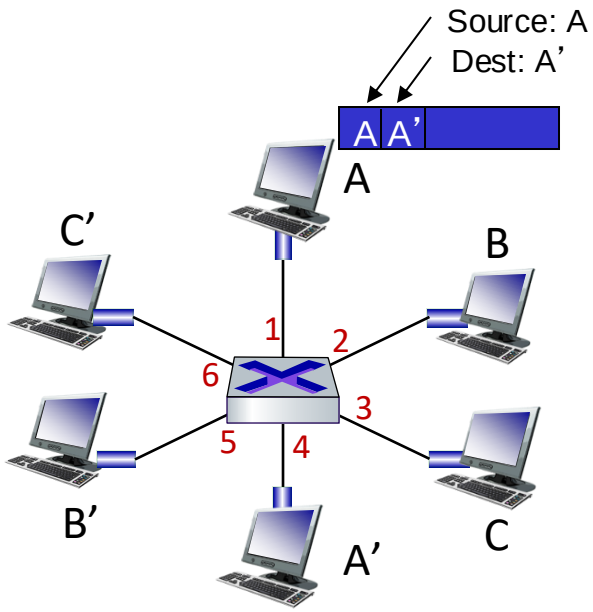


frame destination, A', location unknown: flood

Link Layer: 6-121

Ethernet switch – MAC learning

- switch *learns* which hosts can be reached through which interfaces
  - when frame received, switch "learns" location of sender: incoming LAN segment
  - records sender/location pair in switch table

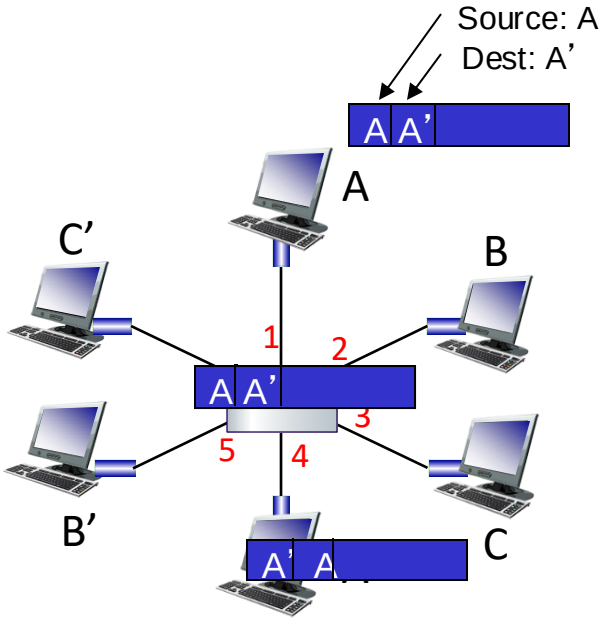


MAC addr	interface	TTL
A	*1*	*60*

Switch table (initially empty)

Ethernet switch - forwarding

- frame destination, A', location unknown: flood
- destination A location known: selectively send on just one link



MAC addr	interface	TTL
A	*1*	*60*
A'	*4*	*60*

switch table (initially empty)

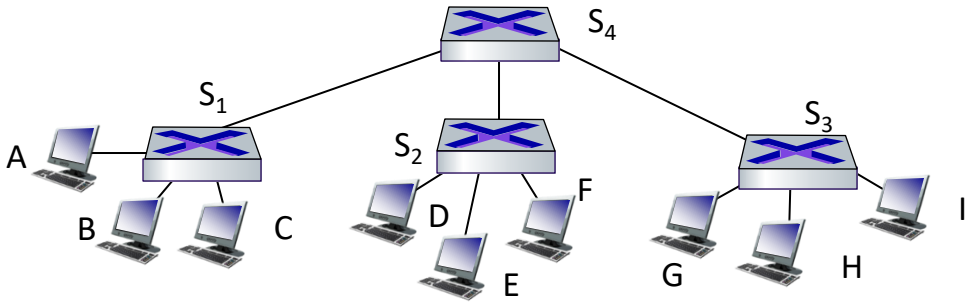
Ethernet switch – forwarding logic

when frame received at switch:

1. record incoming link, MAC address of sending host
2. index switch table using MAC destination address
3. if entry found for destination
   then {
   if destination on segment from which frame arrived
      then drop frame
       else forward frame on interface indicated by entry
    }
   else flood /* forward on all interfaces except arriving interface */

Ethernet switch – larger networks

self-learning switches can be connected together:

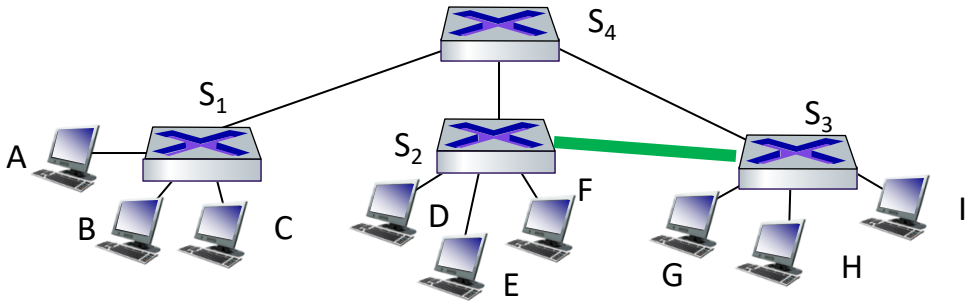


Q: sending from A to G - how does S_1 know to forward frame destined to G via S_4 and S_3?

- *A:* self learning! (works exactly the same as in single-switch case!)

Ethernet switch – larger networks

self-learning switches can be connected together, with an additional link …

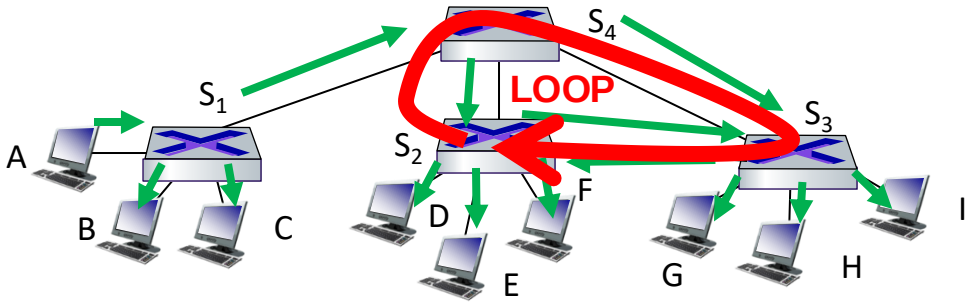


Q: sending from A to G:

what will happen the first time (all switches just powered on)?

Ethernet switch – larger networks

self-learning switches can be connected together, with an additional link …



Q: sending from A to G:

what will happen the first time (all switches just powered on)?

Ethernet Forwarding – STP

- IEEE 802.1D: *Spanning Tree Protocol (STP)*
- STP forms a spanning tree where *interfaces are blocked to avoid loops* in the network
- Switches communicate using 2 types of BPDU's (*Bridge Protocol Data Units*):
  - Configuration BPDU's (at start-up)
  - Topology Change Notification BPDU's (during operation)
- The spanning tree is built *automatically*

Ethernet Forwarding – STP

- **Configuration procedure:**
  - Step 1: all ports in blocking mode
  - Step 2: choose a root switch
  - Step 3: minimum spanning tree algorithm calculated in a distributed way using the Port Path Costs (cfr. Kruskal)
  - Step 4: ports will change to forwarding mode based on spanning tree
- **Q:** How to choose the root switch?
  - Based on Switch ID (lowest value)
  - Switch ID format:

Switch priority (2 bytes)	MAC address (6 bytes)

see notes

The collection of switches in a LAN can be considered a graph whose nodes are the switches and whose edges are the cables connecting the switches. To break loops in the LAN while maintaining access to all LAN segments, the switches collectively compute a spanning tree. The spanning tree is not necessarily a minimum cost spanning tree. A network administrator can reduce the cost of a spanning tree, if necessary, by altering some of the configuration parameters in such a way as to affect the choice of the root of the spanning tree.

The spanning tree that the switches compute using the Spanning Tree Protocol can be determined using the following rules.

Elect a root switch. The *root switch* of the spanning tree is the switch with the smallest switch ID. Each switch has a unique identifier (ID) and a configurable priority number; the switch ID contains both numbers. To compare two switch IDs, the priority is compared first. If two switches have equal priority, then the MAC addresses are compared. A network administrator can determine which switch is the root by configuring its priority to be higher (lower priority number) than any other switches on the LAN.

Determine the least cost paths to the root switch. The computed spanning tree has the property that messages from any connected device to the root switch traverse a least cost path, i.e., a path from the device to the root that has minimum cost among all paths from the device to the root. The cost of traversing a single network segment is configurable; the cost of traversing a path is the sum of the costs of the segments on the path. Different technologies have different default costs for network segments. Also, an administrator can configure the cost of traversing a particular network segment.

Ethernet Forwarding –STP example

router

RP: Root Port
DP: Designated Port
BP: Blocked Port

1 root

8

DP

RP

DP

DP

switch

7

RP

RP

5

DP

BP

DP

DP

DP

DP

RP

RP

DP

RP

BP

RP

BP

DP

DP

DP

BP

DP

BP

3

BP

2

6

DP

BP

4

hub

see notes

The property that messages always traverse least-cost paths to the root is guaranteed by the following two rules.

Least cost path from each switch. After the root switch has been chosen, each switch determines the cost of each possible path from itself to the root. From these, it picks the one with the smallest cost (the least-cost path). The port connecting to that path becomes the *root port* of the switch.

Least cost path from each network segment. The switches on a network segment collectively determine which switch has the least-cost path from the network segment to the root. The port connecting this switch to the network segment is then the *designated port* for the segment.

Disable all other root paths. Any active port that is not a root port or a designated port is a blocked port.

Modifications in case of ties. The above rules over-simplify the situation slightly, because it is possible that there are ties, for example, two or more ports on a single switch are attached to least-cost paths to the root or two or more switches on the same network segment have equal least-cost paths to the root. To break such ties:

Breaking ties for root ports. When multiple paths from a switch are least-cost paths, the chosen path uses the neighbor switch with the lower switch ID. The root port is thus the one connecting to the switch with the lowest switch ID.

Breaking ties for designated ports. When more than one switch on a segment leads to a least-cost path to the root, the switch with the lower switch ID is used to forward messages to the root. The port attaching that switch to the network segment is the *designated port* for the segment.

The final tie-breaker. In some cases, there may still be a tie, as when two switches are connected by multiple cables. In this case, multiple ports on a single switch are candidates for root port or designated port. In this case, the port with the lowest port number is used.

Ethernet Forwarding – xSTP

- Recovery times STP:
  - With standard timers = *30- 60 sec*
  - Acceptable in small LAN environment, but not for large Ethernet networks
- STP extension for *fast recovery* after failures:
  - IEEE 802.1w *Rapid Spanning Tree Protocol* (RSTP)
  - Recovery in order of *seconds*
- STP extension for VLAN (see later)
  - IEEE 802.1S Multiple Spanning Tree Protocol (MSTP) (multiple spanning trees possible using different VLAN's)

Link layer, LANs: outline

6.1 Introduction, services

6.2 Error detection, correction

6.3 Multiple access protocols

6.4 LANs

- addressing, ARP
- Ethernet
- switches
- **VLANS**

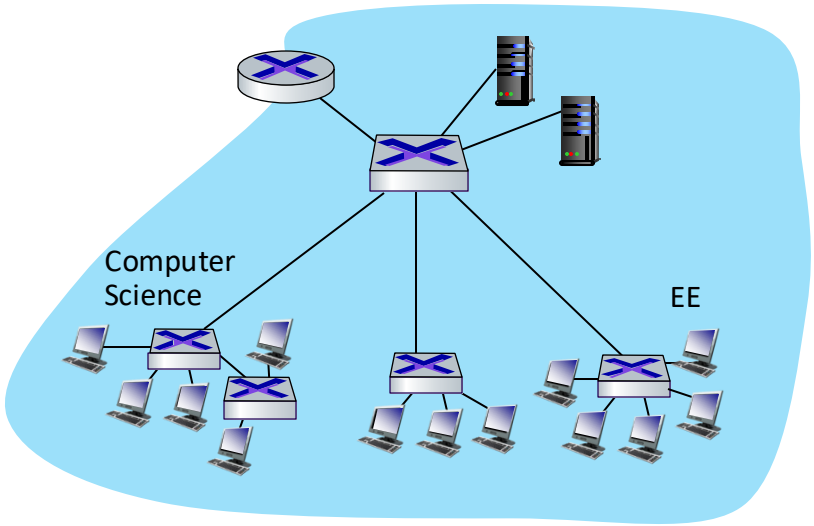
6.5 Link virtualization: MPLS

6.6 Data center networking

6.7 A day in the life of a web request

Virtual LANs - motivation

Q: what happens as LAN sizes scale, users change point of attachment?

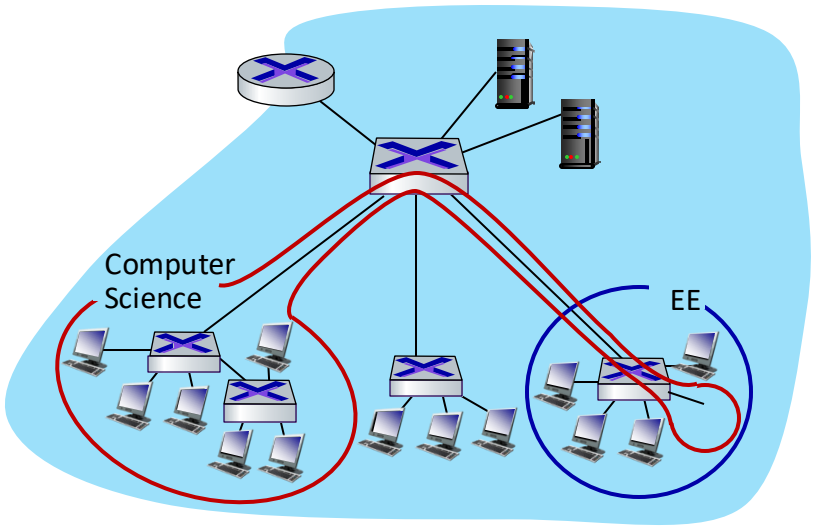


single broadcast domain:

- *scaling:* all layer-2 broadcast traffic (ARP, DHCP, unknown MAC) must cross entire LAN
- efficiency, security, privacy issues

Virtual LANs - motivation

Q: what happens as LAN sizes scale, users change point of attachment?



single broadcast domain:

- *scaling:* all layer-2 broadcast traffic (ARP, DHCP, unknown MAC) must cross entire LAN
- efficiency, security, privacy, efficiency issues

administrative issues:

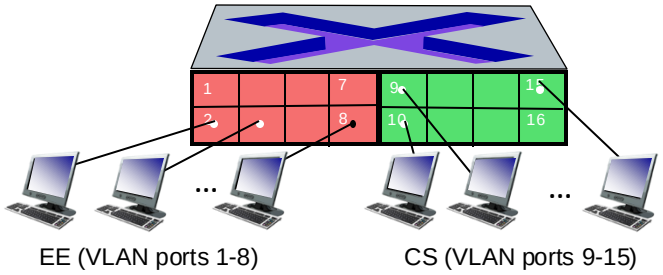
- CS user moves office to EE - *physically* attached to EE switch, but wants to remain *logically* attached to CS switch

Port-based VLANs

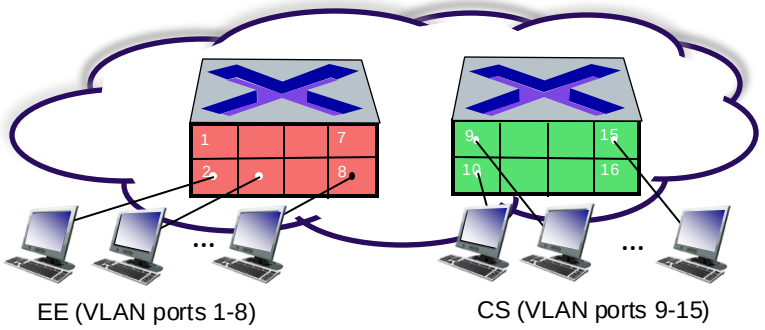
Virtual Local Area Network (VLAN)

switch(es) supporting VLAN capabilities can be configured to define multiple *virtual* LANS over single physical LAN infrastructure.

port-based VLAN: switch ports grouped (by switch management software) so that *single* physical switch ……

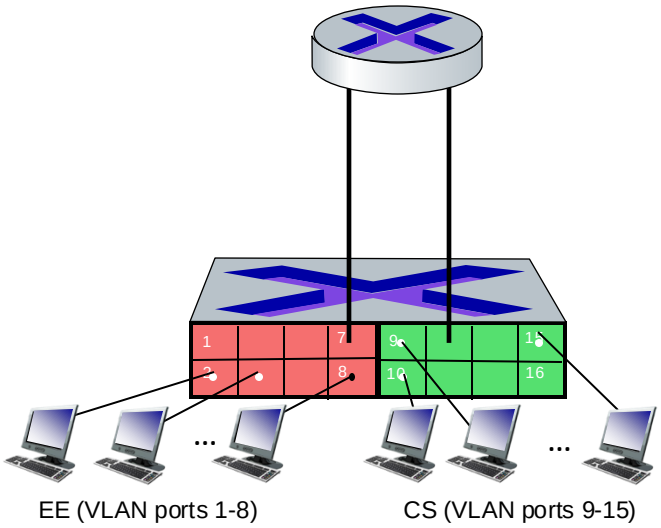


… operates as multiple virtual switches

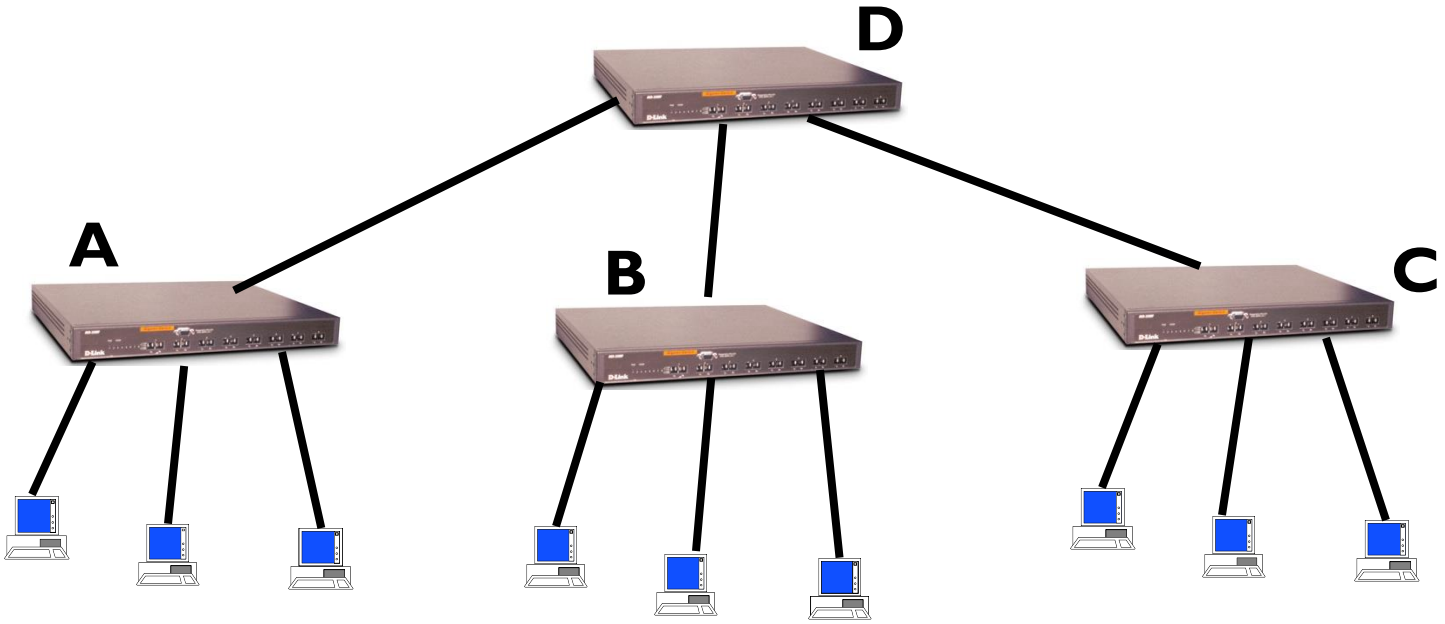


Port-based VLANs

- traffic isolation: frames to/from ports 1-8 can *only* reach ports 1-8
  - can also define VLAN based on MAC addresses of endpoints, rather than switch port
- dynamic membership: ports can be dynamically assigned among VLANs
- forwarding between VLANS: done via routing (just as with separate switches)
  - in practice vendors sell combined switches plus routers

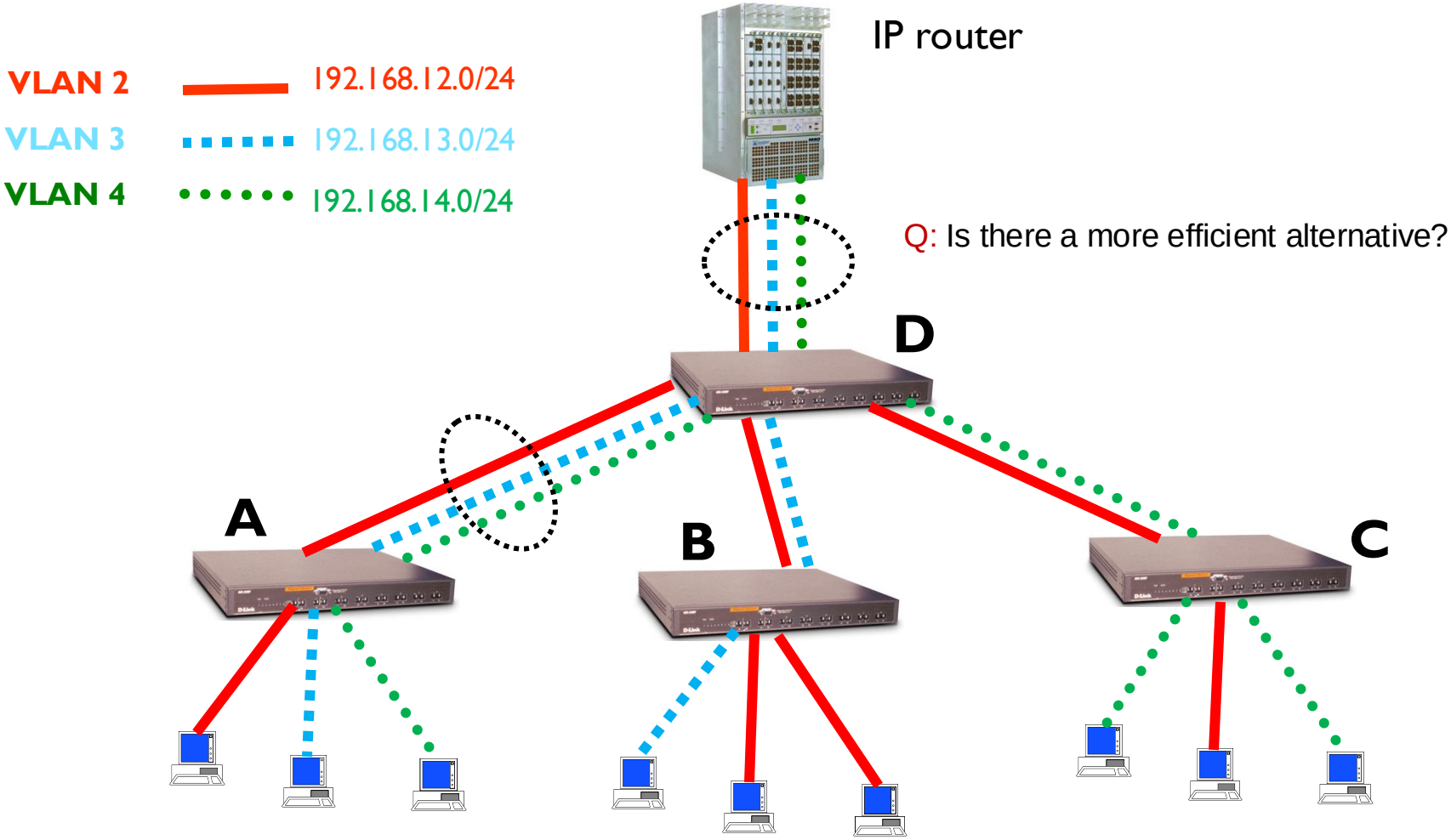


Port-based VLANs - example



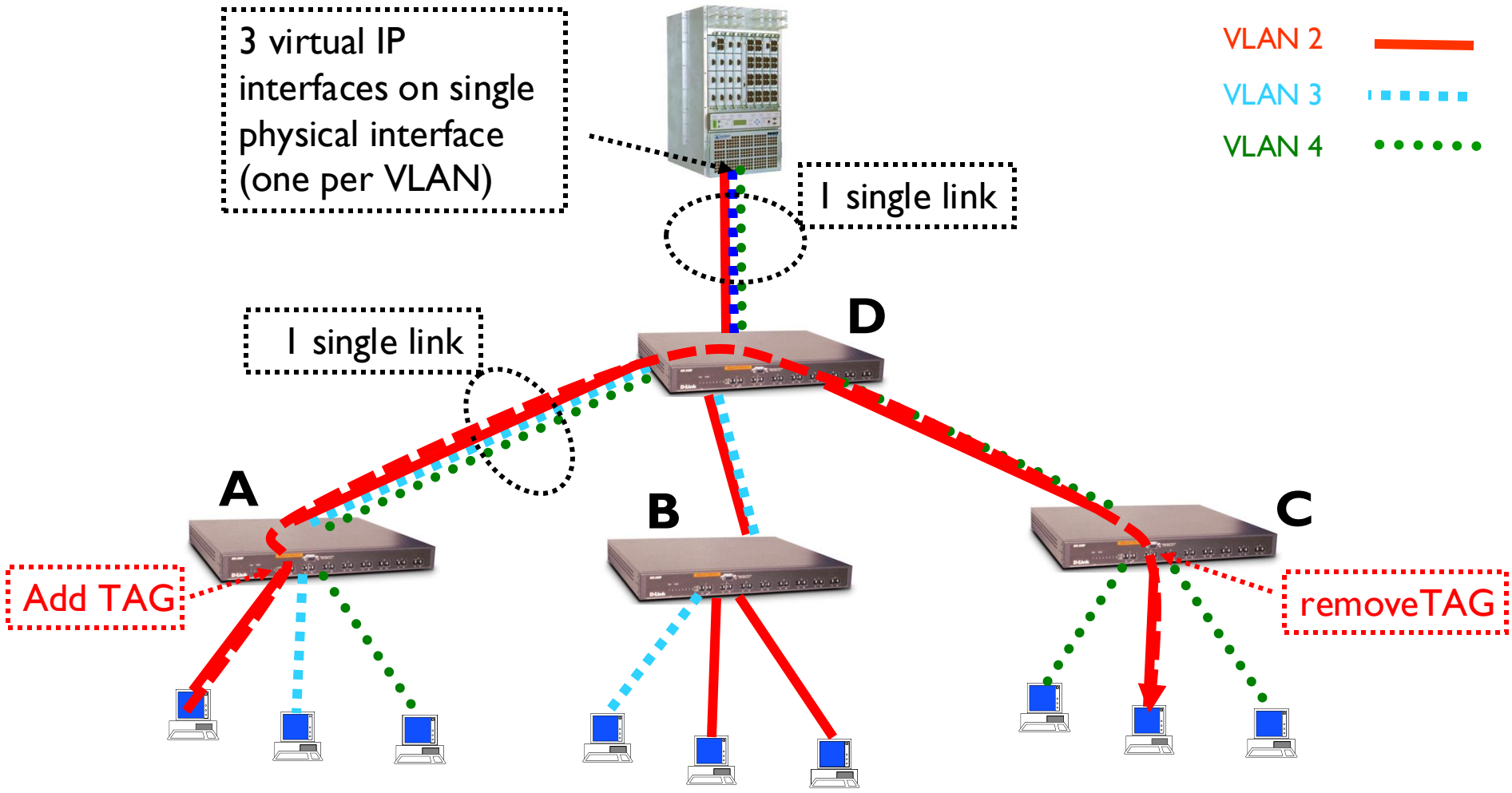
How to organize 9 hosts over 3 port-based VLANs ?

Port-based VLANs - example



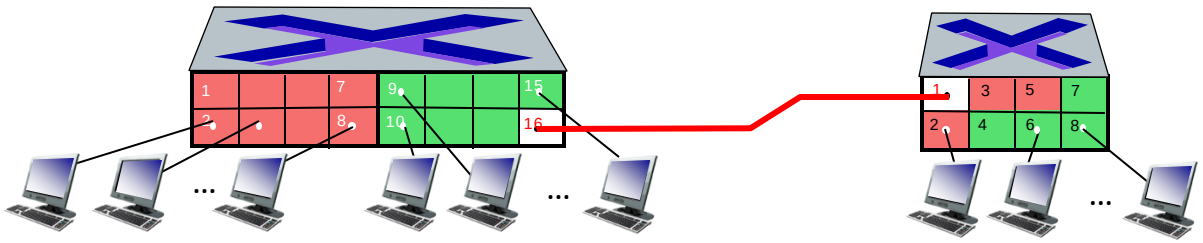
- Each VLAN = own subnet -> has own IP address block
- Multiple VLANs require separate ports -> multiple cables
- Forwarding between VLANs = forwarding between subnets -> via IP router

Tag-based VLANs



- Multiple VLAN's can use a single port (due to tagging) -> trunk port

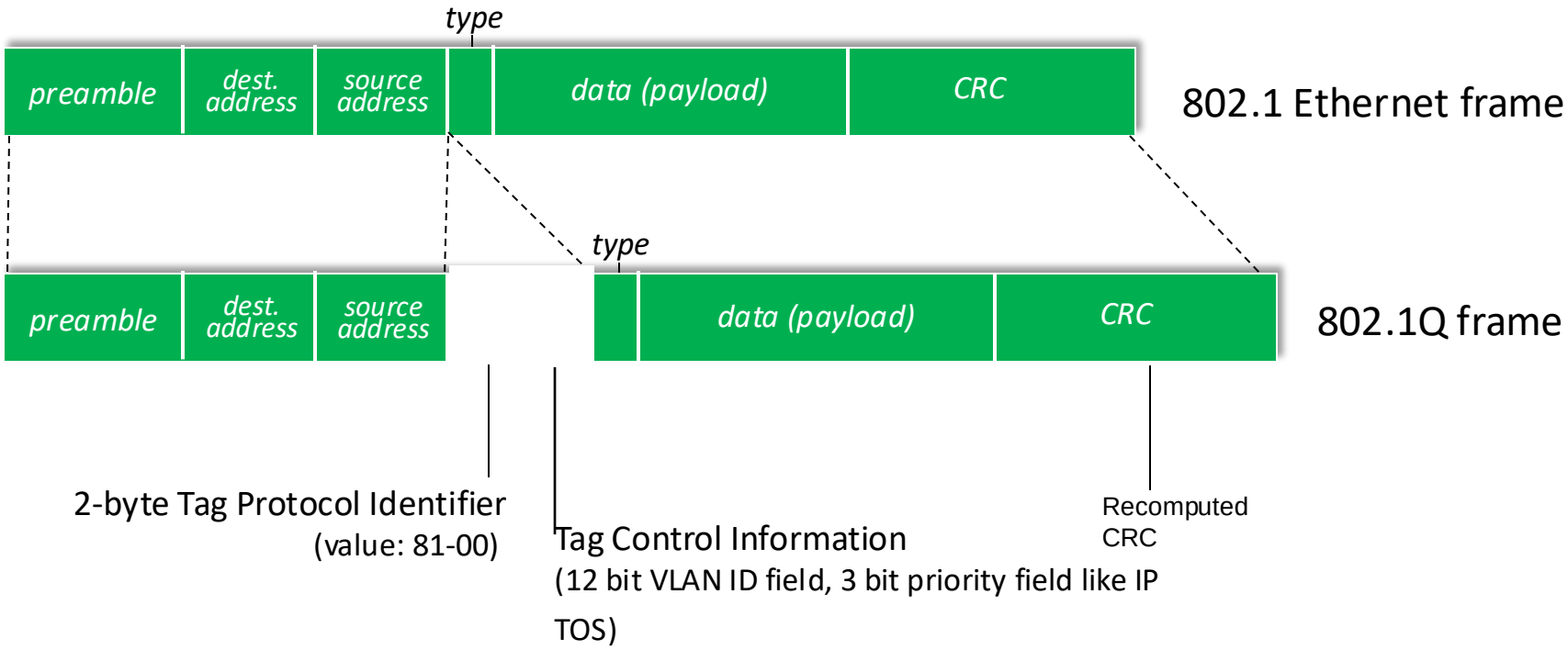
VLANs over multiple switches



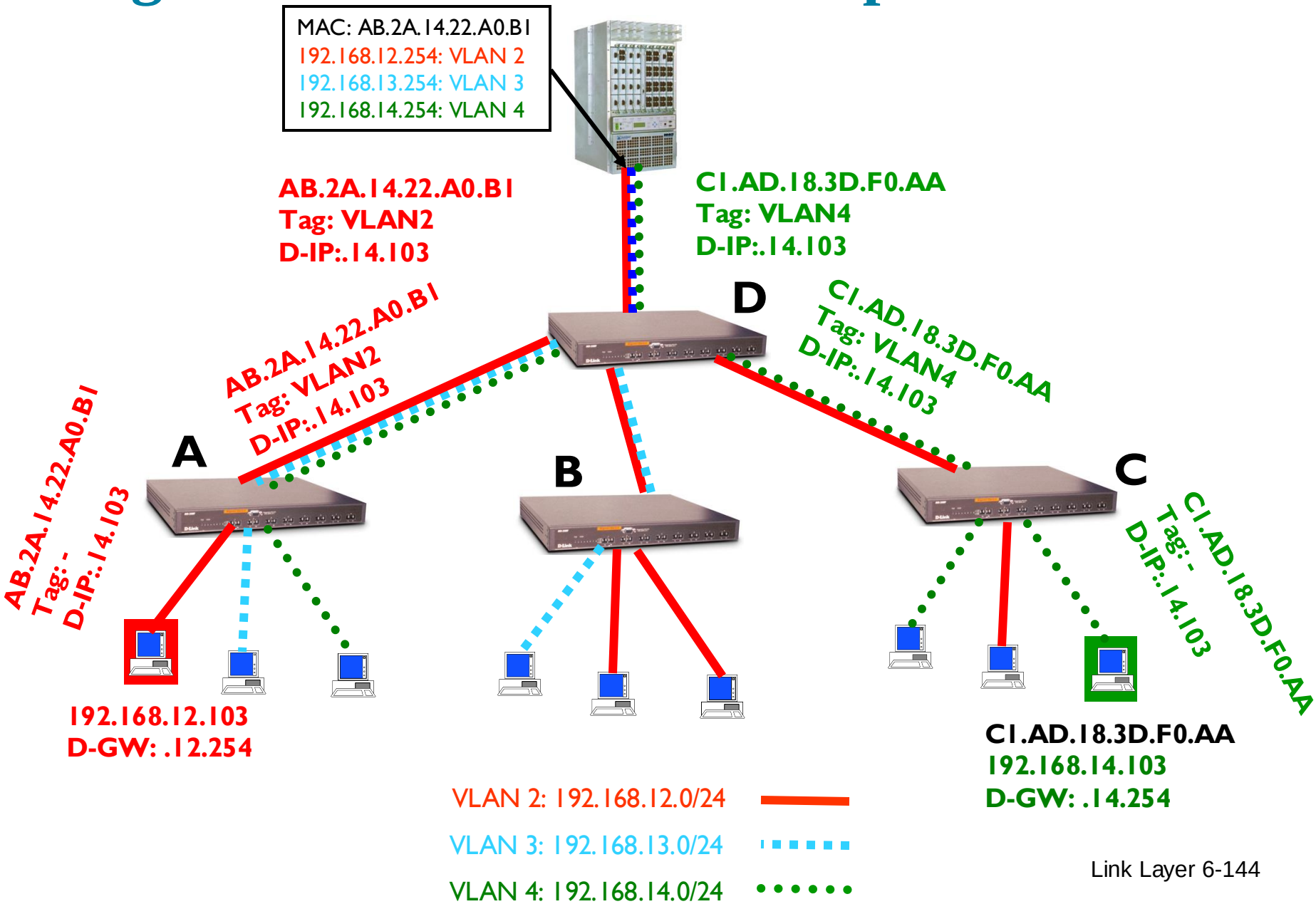
trunk port: carries frames between VLANS defined over multiple physical switches

- frames forwarded within VLAN between switches can't be vanilla 802.1 frames (must carry VLAN ID info)
- 802.1q protocol adds/removed additional header fields for frames forwarded between trunk ports

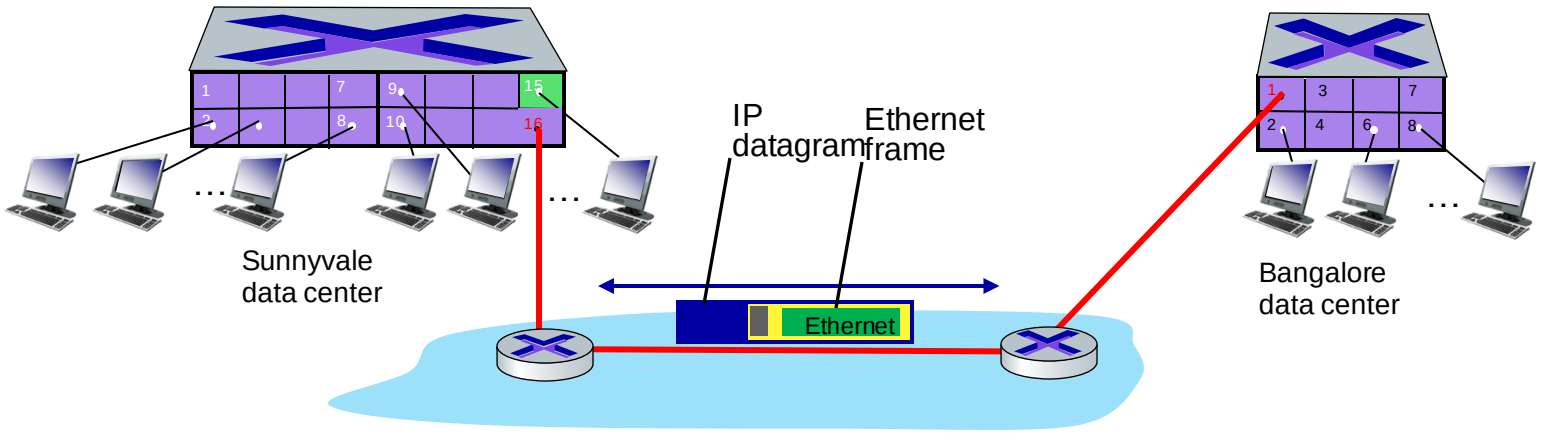
IEEE 802.1Q VLAN frame format



Tag-based VLAN – full example



Ethernet VPN with VXLAN =/= VLAN



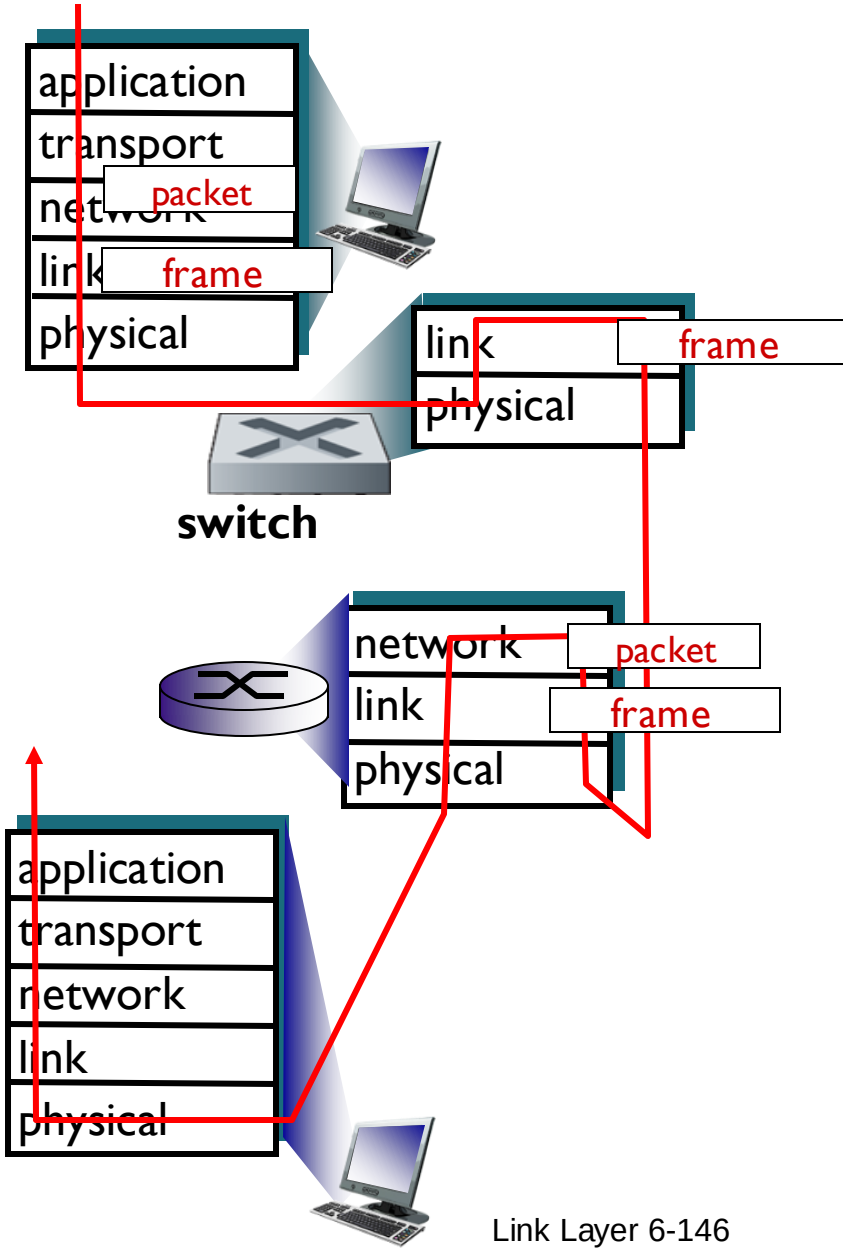
Layer-2 Ethernet switches *logically* connected to each other (e.g., using IP as an *underlay*)

- Ethernet frames carried *within* IP datagrams between sites
- "*tunneling* scheme to *overlay Layer 2 networks on top of Layer 3 networks* ... runs over the existing networking infrastructure and provides a means to "stretch" a Layer 2 network." [RFC 7348]

Routers & switches: similarities & diff's

- both are *store-and-forward*
  - ROUTERS: network-layer devices (examine network-layer headers)
  - SWITCHES: link-layer devices (examine link-layer headers)

- both have *forwarding tables*
  - ROUTERS: use longest-prefix matching in forwarding tables computed by routing protocols
  - SWITCHES: use MAC+VLAN matching in forwarding tables computed by MAC learning, otherwise flooding is used



Routers & switches: pro's & con's

Switches pro's & con's:

+ switch operation involve simpler (=cheap) packet processing

+ switch tables are self-learning (plug & play)

- topology confined to spanning tree

- flooding can come at a high BW cost (broadcast)

Routers pro's & con's:

+ arbitrary topologies can be supported, cycling is limited by TTL counters (and good routing protocols)

+ no flooding, forwarding only to longest matching entry in FWD table

- require IP address/forwarding/routing protocol configuration

- require higher packet processing

Bottomline: Switches do well in small (few hundred hosts) while routers used in large networks (thousands of hosts)

Link layer, LANs: outline

6.1 Introduction, services

6.2 Error detection, correction

6.3 Multiple access protocols

6.4 LANs

- addressing, ARP
- Ethernet
- switches
- VLANS

6.5 Link virtualization: MPLS

6.6 Data center networking

6.7 A day in the life of a web request

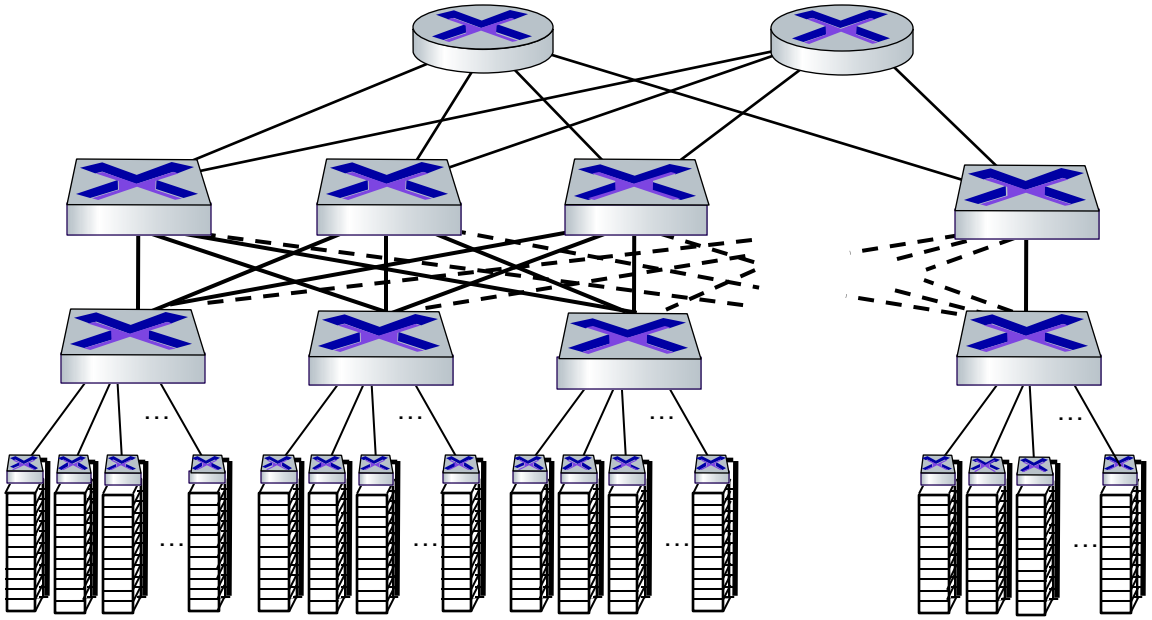
Data center networks

- 10's to 100's of thousands of servers, often closely coupled, in close proximity:
  - e-business (e.g. Amazon)
  - content-servers (e.g., YouTube, Akamai, Apple, Microsoft)
  - search engines, data mining (e.g., Google)

- Challenges:
  - multiple applications, each serving massive numbers of clients
  - managing/balancing load, avoiding processing, networking, data bottlenecks

Inside a 40-ft Microsoft container, Chicago data center

Data center network layers



Border routers
- connections outside datacenter

Tier-1 switches
- connecting to ~16 T-2s below

Tier-2 switches
- connecting to ~16 TORs below

Top of Rack (TOR) switch
- one per rack
- 100G-400G Ethernet to blades

Server racks
- 20- 40 server blades: hosts

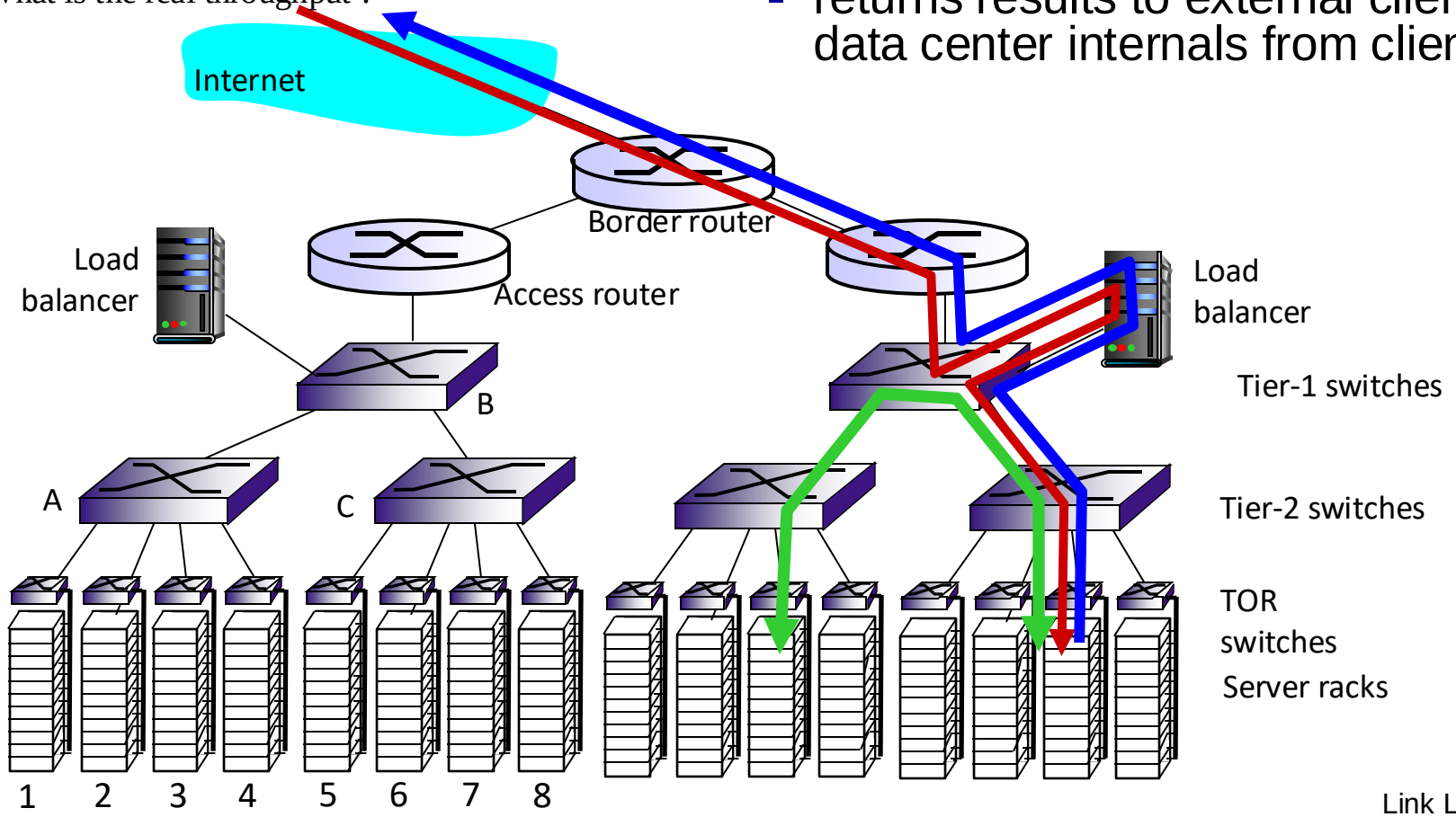
Data center networks – Fat Tree

Example:
- 1 Gbit/s between servers and TOR switches
- 10 Gbit/s between switches

Requested traffic pattern: 10 hosts from rack 1 each send 1 Gbit/s to 10 hosts in rack 5, similar between rack 2-6, 3-7 and 4-8.
What is the real throughput ?

Load Balancer: application-layer routing

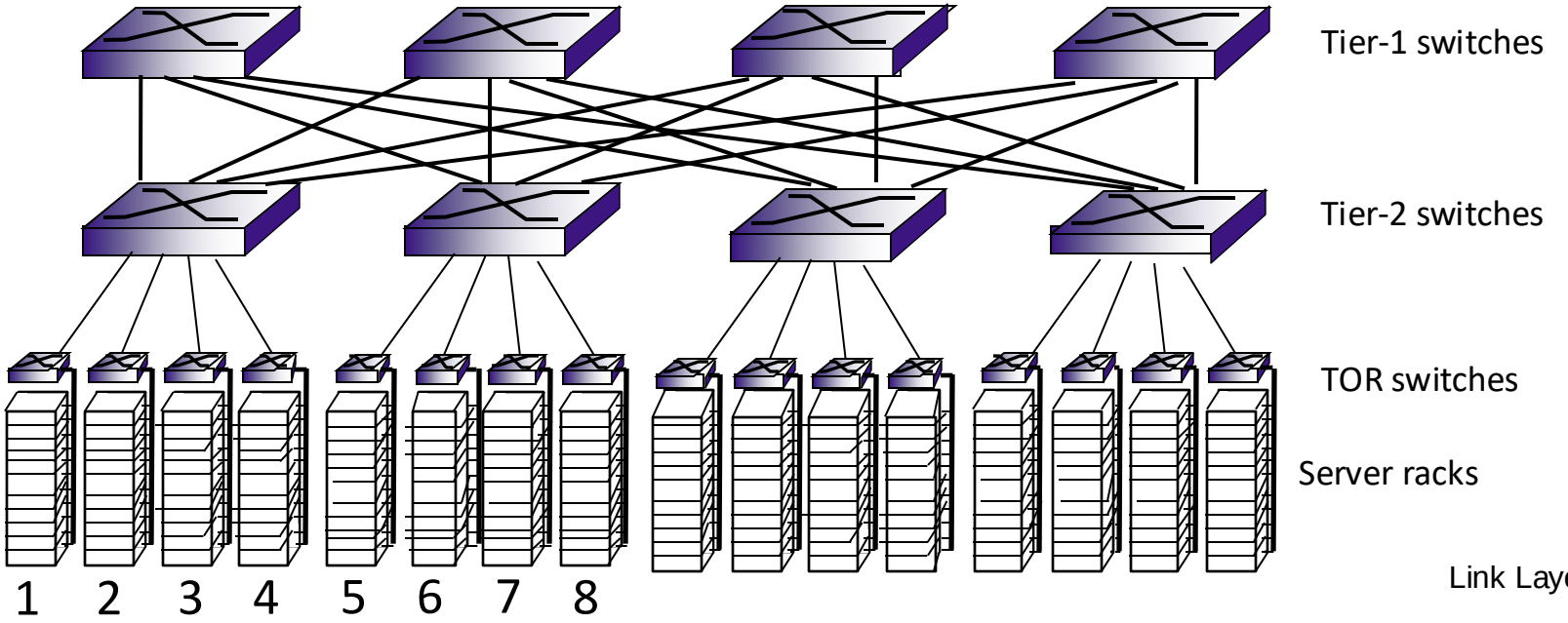
- receives external client requests
- directs workload within data center
- returns results to external client (hiding data center internals from client)



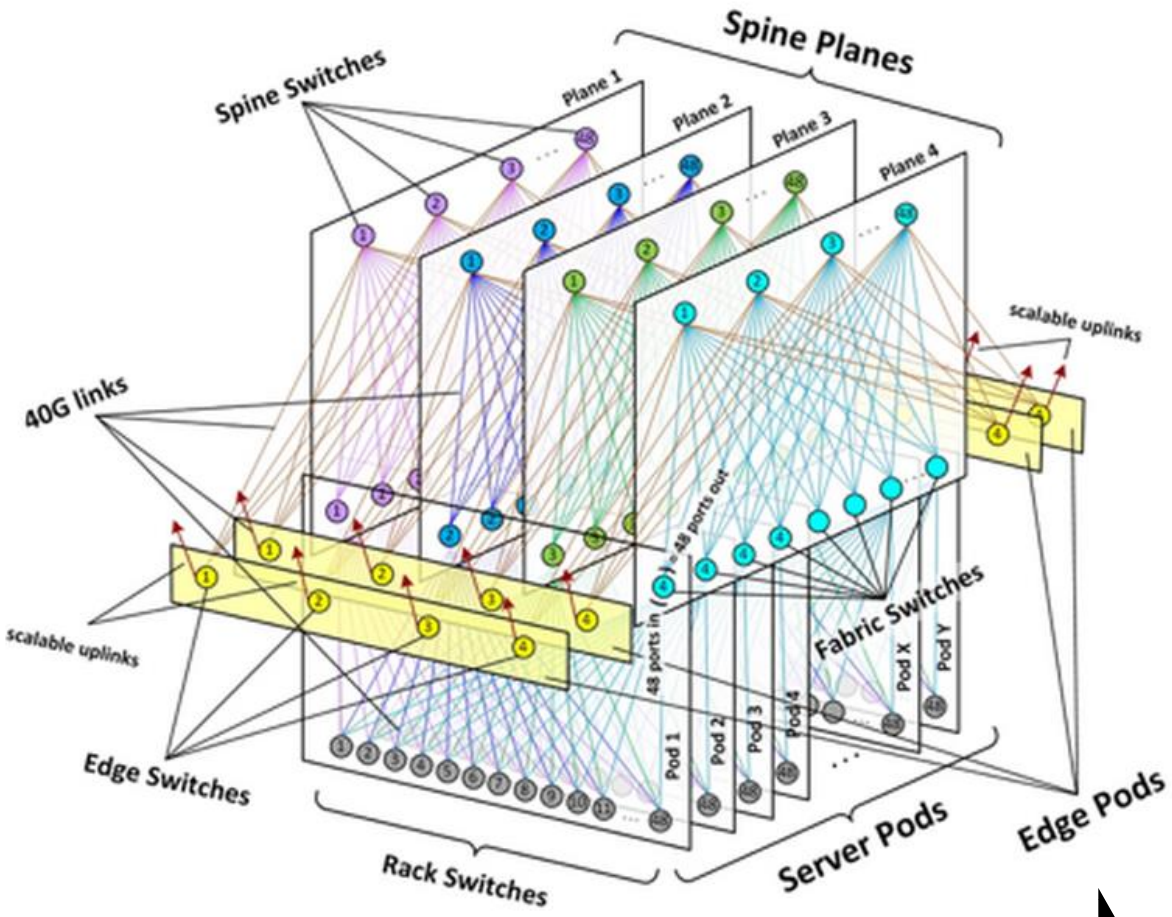
Data center networks – Beyond STP

- rich interconnection among switches, racks:
  - increased throughput between racks (multiple routing paths possible)
  - increased reliability via redundancy
- requires advanced L2/L3 switching & routing configuration
  - optimized Multiple Spanning Tree(s) per VLAN (MSTP)

Note: The routing is more complex then a simple spanning tree at L2 (out of scope of this course).



Data center networks – today (Facebook)



- Traditional DC topologies were limited by port density of cluster switches
- Dependency on limited set of vendors
- Topology and size of the top-down structure is unflexible in the long-term

Multi-dimensional structure

- Uniform Server pods
- Spine planes
- Edge pods
- Custom switch design
- Software-Defined Network (SDN) control

- Below the ToR, L2 switching is used
- Above the ToR to the edge, using Equal-Cost Multipath routing is used on IPv4 and IPv6 network