

# Chapitre 6

## Link Layer and Local Area Networks

Laurent Schumacher (UNamur)

Dernière mise-à-jour : 16 janvier 2019

Materials used with permission from Pearson Education

© 1996-2012 J.F Kurose and K.W. Ross, All Rights Reserved

# Outline

## Data link layer services

Error-detection and error-correction techniques

Multiple access protocols

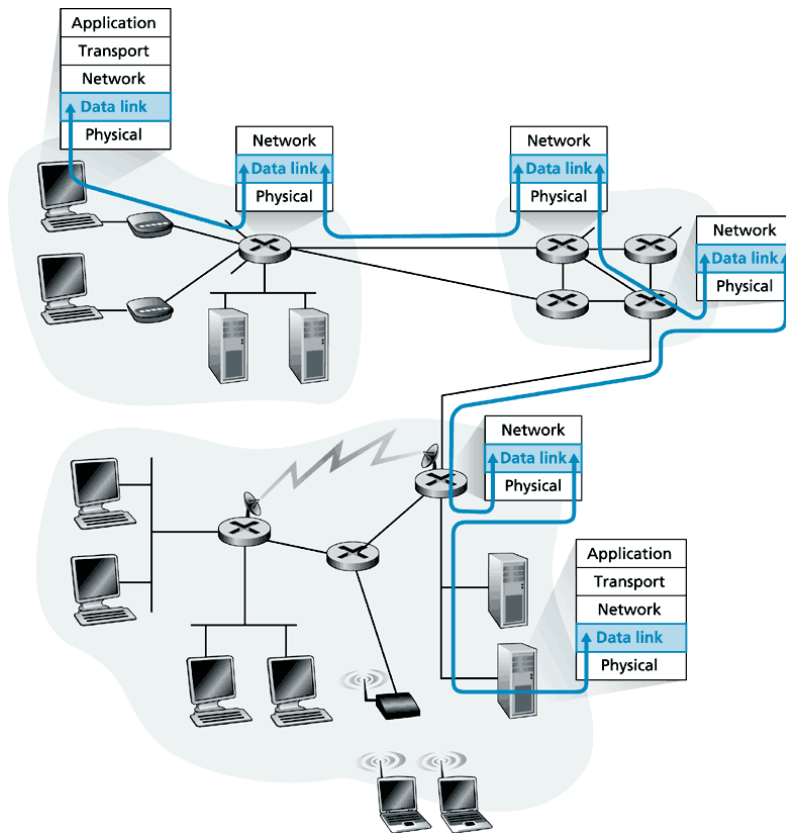
LAN addresses – Neighbour discovery (IPv6), ARP (IPv4)

Ethernet

Hubs, bridges, switches and routers

# Data Link Layer Services

## Introduction



- Hosts, routers, bridges and switches are *nodes*
- Communication channels that connect adjacent nodes are *links*
- 2-PDU is a *frame*, encapsulates datagram
- Data-link layer has responsibility of transferring network-layer datagram from one node to an adjacent node over a link

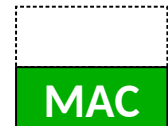
# Data Link Layer Services

## Context

- Datagram transferred by different link protocols over different links:
  - e.g., EPON on first link, Ethernet / 802.3 on intermediate links, WiFi / 802.11 on last link
- Each link protocol provides different services (reliability, error check, flow control, etc.)
- Transportation analogy : trip from Princeton to Lausanne
  - Car: Princeton to JFK
  - Plane: JFK to Geneva
  - Train: Geneva to Lausanne
  - Tourist = datagram
  - Transport segment = communication link
  - Transportation mode (car, plane, train) = link layer protocol
  - Travel agent = routing algorithm

# Data Link Layer Services Protocol

- Logical Link Control protocol (LLC)
  - Non architecture specific
  - Framing
  - Encapsulate network-layer datagram within frame
  - Physical address ( $\neq$  IP network address) used in header to identify source and destination
- Medium Access protocol (MAC)
  - Specific to physical layer (PHY)
  - Link access
    - No/simple protocol in point-to-point links
    - Access to shared medium in broadcast links
  - Reliable delivery
    - ACK + retransmissions (Go-Back-N, Selective Repeat, etc.)
    - Seldom used on low bit error link (fiber, some twisted pair)
    - Wireless links
      - High error rates
      - Enables to react locally instead of E2E



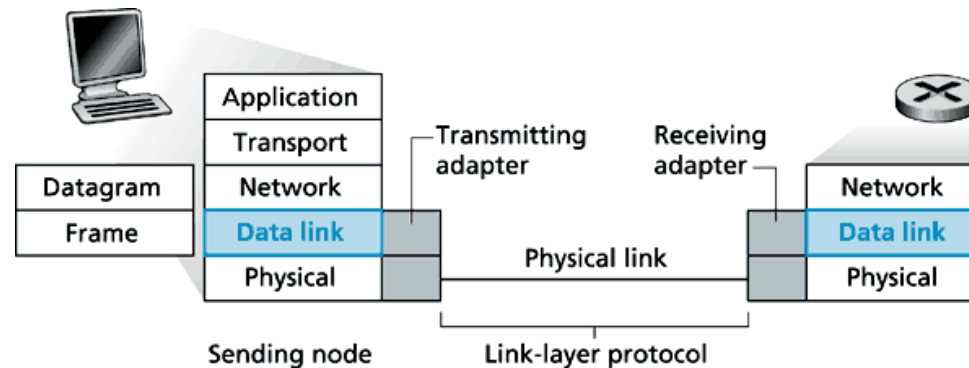
# Data Link Layer Services

## Services

- Reliability
  - Error *detection*
    - Errors caused by attenuation, electromagnetic ambient noise
    - Receiver detects presence of errors
    - Requests retransmission or drops frame
    - Usually more sophisticated in L2 than in L3/L4
    - Implemented in hardware
  - Error *correction*: receiver identifies and corrects bit error(s) without resorting to retransmission
- Flow control between adjacent nodes
- Full- or half-duplex: ability to send and receive simultaneously
- Strong parallels with transport-layer services, but not E2E. Only node-to-node.

# Data Link Layer Services

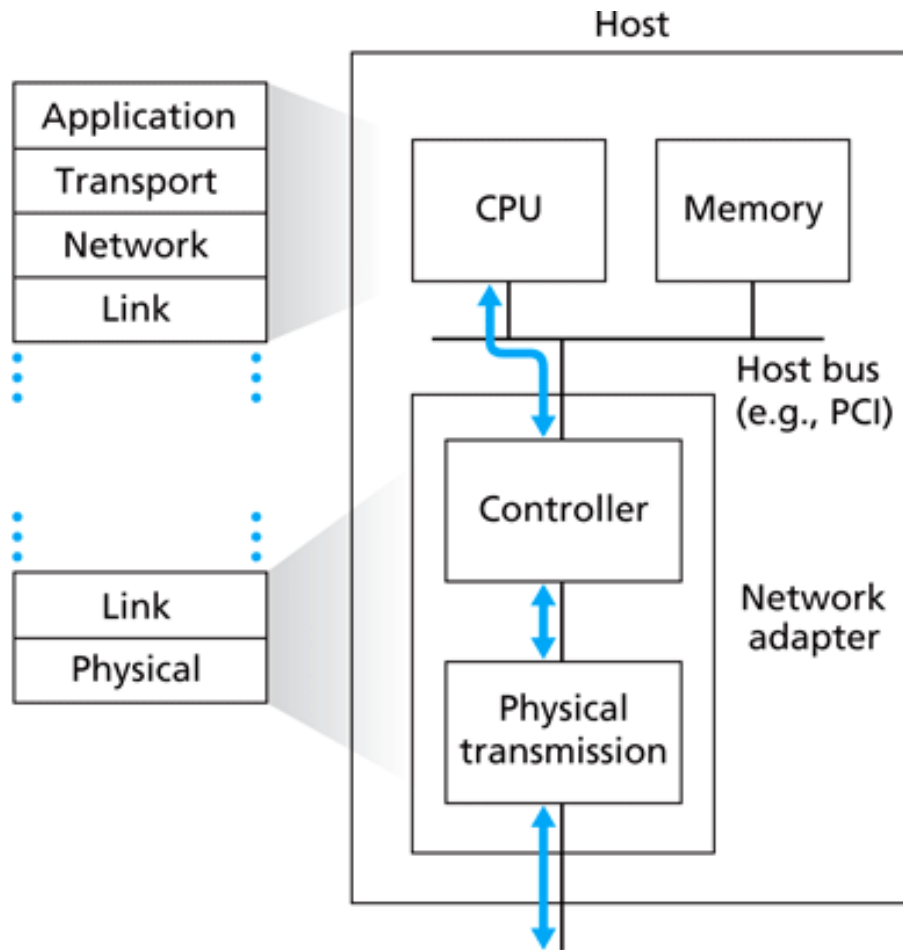
## Adapters



- Link layer implemented in “adapter”
- Source
  - Encapsulates datagram in a frame
  - Adds error checking bits, numbering, flow control, etc.
- Destination
  - Looks for errors, sequence numbers, flow control, etc.
  - Extracts network-layer datagram
- Adapter is semi-autonomous
  - Can discard erroneous frame without notifying upper layers
  - Still under control of the node

# Data Link Layer Services

## Network Interface Card (NIC)



- High-level features (Frame addressing, interrupts to/from NIC controller) in software (CPU)
- Low-level features (Framing, medium access, flow control, error detection) in hardware (Intel 8254x, Atheros AR5006, etc.)



# Outline

Data link layer services

Error-detection and error-correction techniques

Multiple access protocols

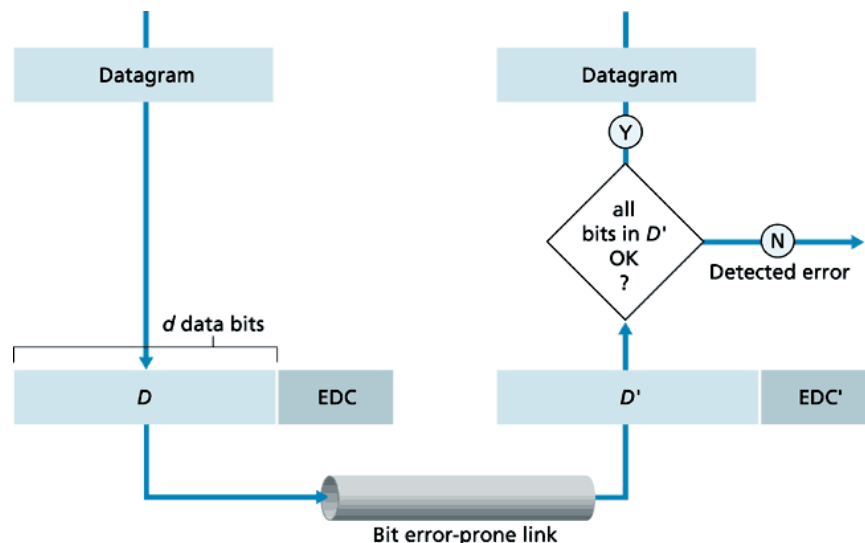
LAN addresses – Neighbour discovery (IPv6), ARP (IPv4)

Ethernet

Hubs, bridges, switches and routers

# Error Detection and Correction Background

- Bit-level error detection/correction



Assume

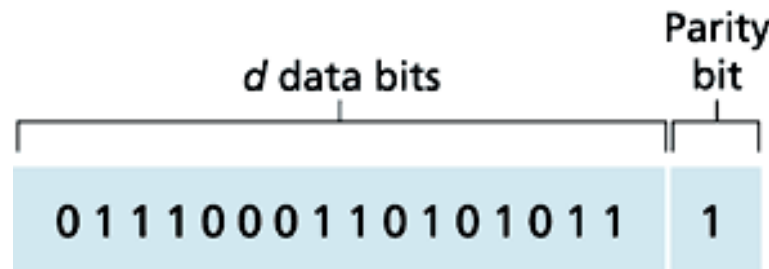
- $D$  = Data protected by error checking, may cover header fields
- EDC = Error Detection and Correction bits (redundancy)

- Goal: minimise non detection of bit errors
- The larger the EDC, the lower the risk of undetected errors
- Techniques: Parity check, checksum and Cyclic Redundancy Check (CRC)

# Error Detection and Correction

## Parity Check (1/2)

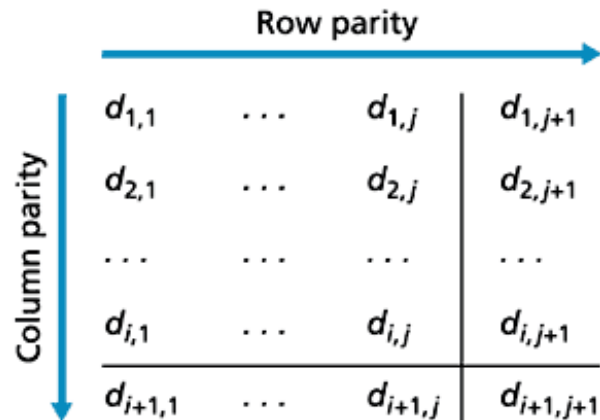
- Simplest form of error detection
- Single bit parity
- Enables to detect odd number of bit errors



- Limitation: errors usually occur in bursts
- Probability to detect bit error with single-bit parity = 50% under burst conditions

# Error Detection and Correction

## Parity Check (2/2)



No errors

1	0	1	0	1	1
1	1	1	1	0	0
0	1	1	1	0	1
0	0	1	0	1	0

Correctable  
single-bit error

1	0	1	0	1	1
1	0	1	1	0	0
0	1	1	1	0	1
0	0	1	0	1	0

Parity error

Parity error

- Two-dimensional parity
- Longitudinal/Vertical Redundancy Checking (LRC/VRC)
- Forward Error Correction (FEC) → able to detect and correct bit errors

# Error Detection and Correction

## Checksumming methods

- Goal: detect “errors” (e.g. flipped bits) in transmitted segment
- Sender
  - Treat segment contents as sequence of 16-bit integers
  - Checksum: addition (1’s complement sum) of segment contents
  - Sender puts checksum value into UDP checksum field
- Receiver
  - Compute checksum of received segment
  - Check if computed checksum equals checksum field value
  - NO - Error detected
  - YES - No error detected. Sure?

# Error Detection and Correction

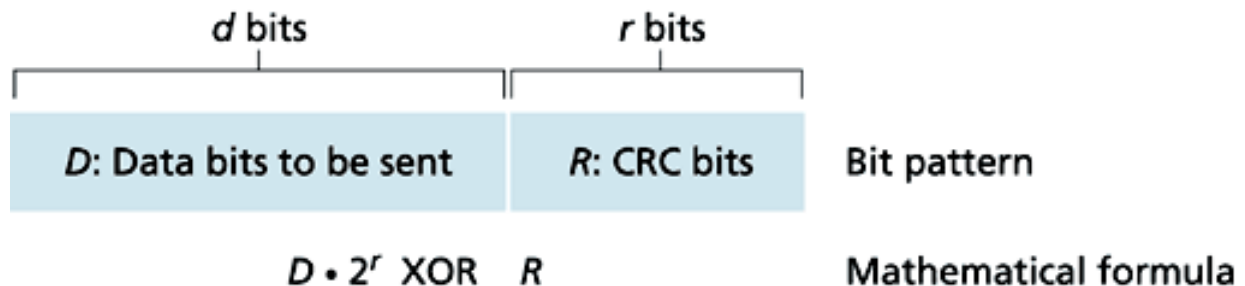
## Checksumming methods – Discussion

- Pro's
  - Little packet overhead (16 bits in TCP/UDP)
  - Simple and fast
  - Easily implemented in software (transport-layer method, located at edge)
- Con
  - Weak protection

# Error Detection and Correction

## Cyclic Redundancy Check (CRC)

- Assume  $d$ -bit binary data word  $D$
- Choose  $(r+1)$ -bit pattern (generator)  $G$
- For a given  $D$ , choose the  $r$ -bit additional CRC  $R$  such that  $(d+r)$ -bit word  $\langle D, R \rangle$  exactly divisible by  $G$  modulo 2



- Receiver knows  $G$ , divides  $\langle D, R \rangle$  by  $G$
- If non-zero remainder: error detected
- CRC can detect error bursts less than  $(r+1)$  bits

# Error Detection and Correction

## Cyclic Redundancy Check (CRC)

- How does the sender compute the CRC?

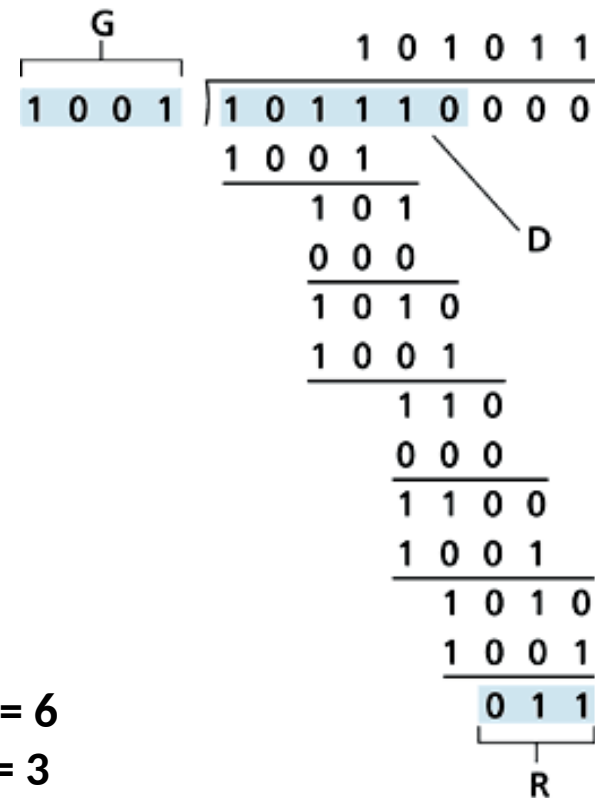
$$D \bullet 2^r \text{ XOR } R = nG$$

$$D \bullet 2^r = nG \text{ XOR } R$$

$$R = \text{remainder} \frac{D \bullet 2^r}{G}$$

A	B	A XOR B
0	0	0
0	1	1
1	0	1
1	1	0

$d = 6$   
 $r = 3$





# Outline

Data link layer services

Error-detection and error-correction techniques

Multiple access protocols

LAN addresses – Neighbour discovery (IPv6), ARP (IPv4)

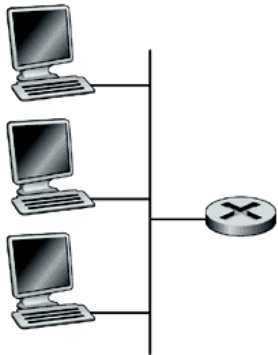
Ethernet

Hubs, bridges, switches and routers

# Multiple Access Protocols

## Introduction

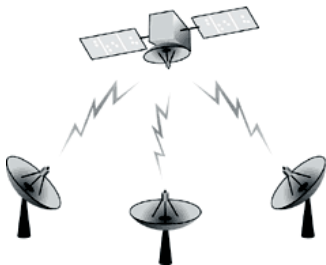
Shared wire  
(for example, Ethernet)



Shared wireless  
(for example, Wifi)



Satellite



Cocktail party



## Two types of “links”

### 1. Point-to-point

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

### 2. Broadcast (shared wire or medium)

- Traditional Ethernet
  - Upstream HFC
  - 802.11 wireless LAN
- Multiple access problem

# Multiple Access Protocols

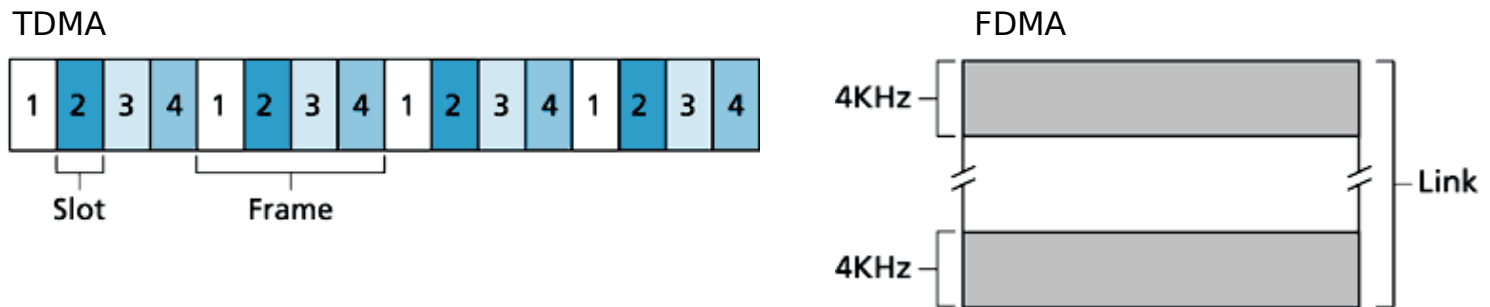
## Solution – Criteria

- On broadcast link, more than two nodes can transmit at the same time → collision and jamming of transmitted frames
- Multiple access protocols coordinate transmissions
- Ideally, for a broadcast channel of rate  $R$ 
  - When one node wants to transmit, it can send at rate  $R$
  - When  $M$  nodes want to transmit, each can send at average rate  $R/M$
  - Fully decentralised protocol (no SPOF – Single Point Of Failure)
  - Simple and inexpensive to implement
- Three categories of MAC protocols
  1. Channel partitioning
  2. Random access
  3. Taking-turns

# Multiple Access Protocols

## Channel partitionning


- Divide channel into smaller “pieces” (time slots, frequency, code)
- Allocate single piece to node for exclusive use



- Fair: each node gets an average rate of  $R/M$  bps
- Drawbacks
  - Single active user can not use empty slots ( $R/M$  bps is upper bound on rate)
  - Single active user must wait its turn or fit into bandwidth

# Multiple Access Protocols

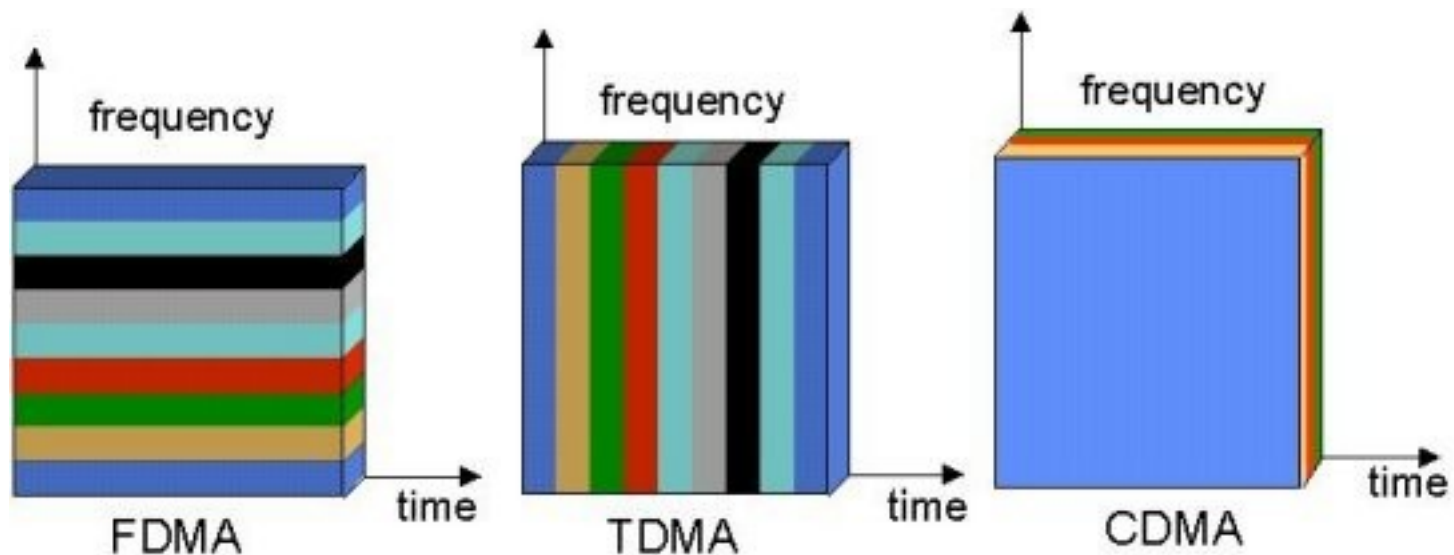
## Channel Partitionning – CDMA

- Code Division Multiple Access (CDMA)
- Unique “code” assigned to each user; i.e., code set partitioning
- Used mostly in wireless broadcast channels
- All users share same frequency, transmit simultaneously, but each user has own “chipping” sequence (i.e., code) to encode data
- Encoding = (original data) x (chipping sequence)
- Decoding = inner-product of encoded signal and chipping sequence
- Multiple users can “coexist” with minimal interference (if codes are “orthogonal”) 

# Multiple Access Protocols

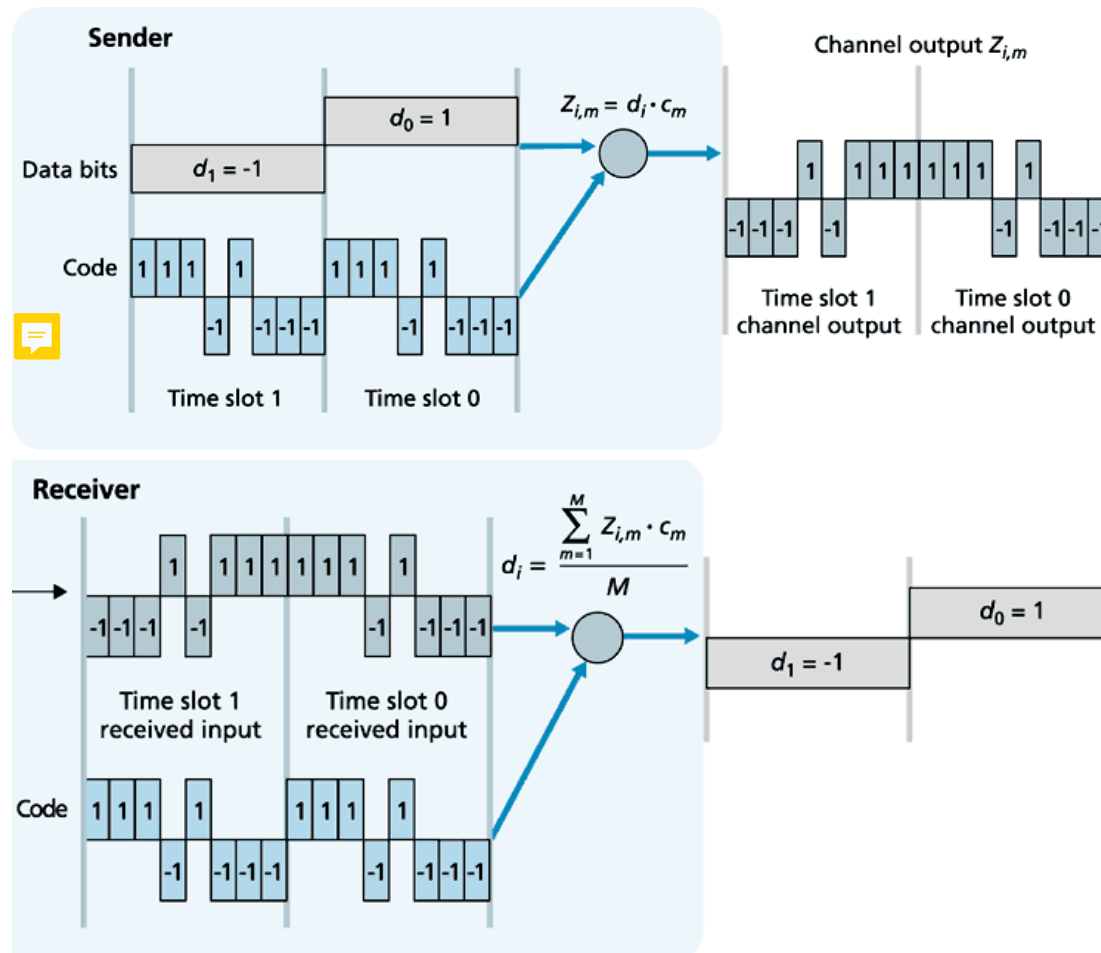
## Channel Partitioning - CDMA

- All users share same frequency, transmit simultaneously, but each user has own “chipping” sequence (i.e., code) to encode data



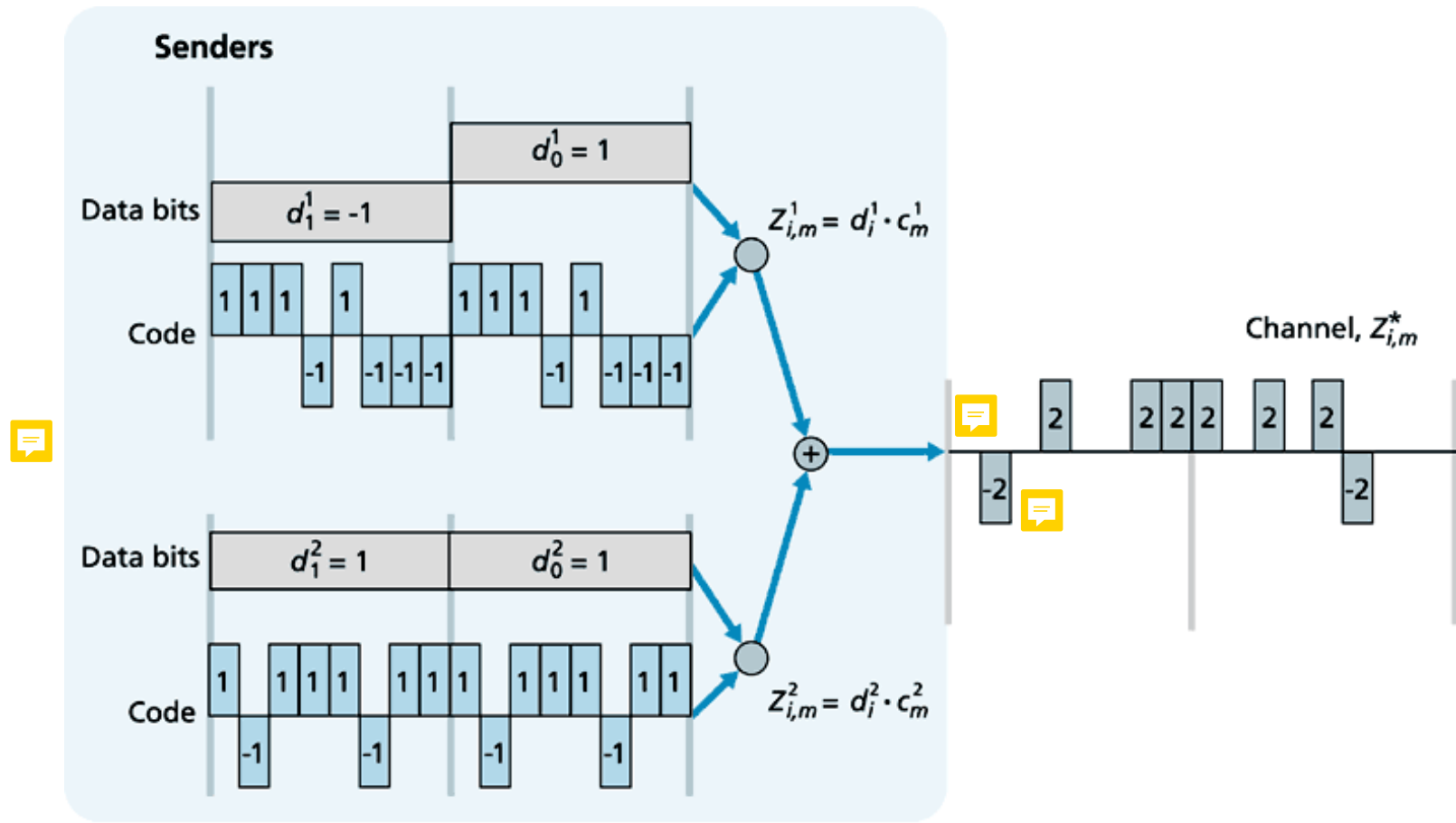
# Multiple Access Protocols

## Channel Partitionning – CDMA, single user



# Multiple Access Protocols

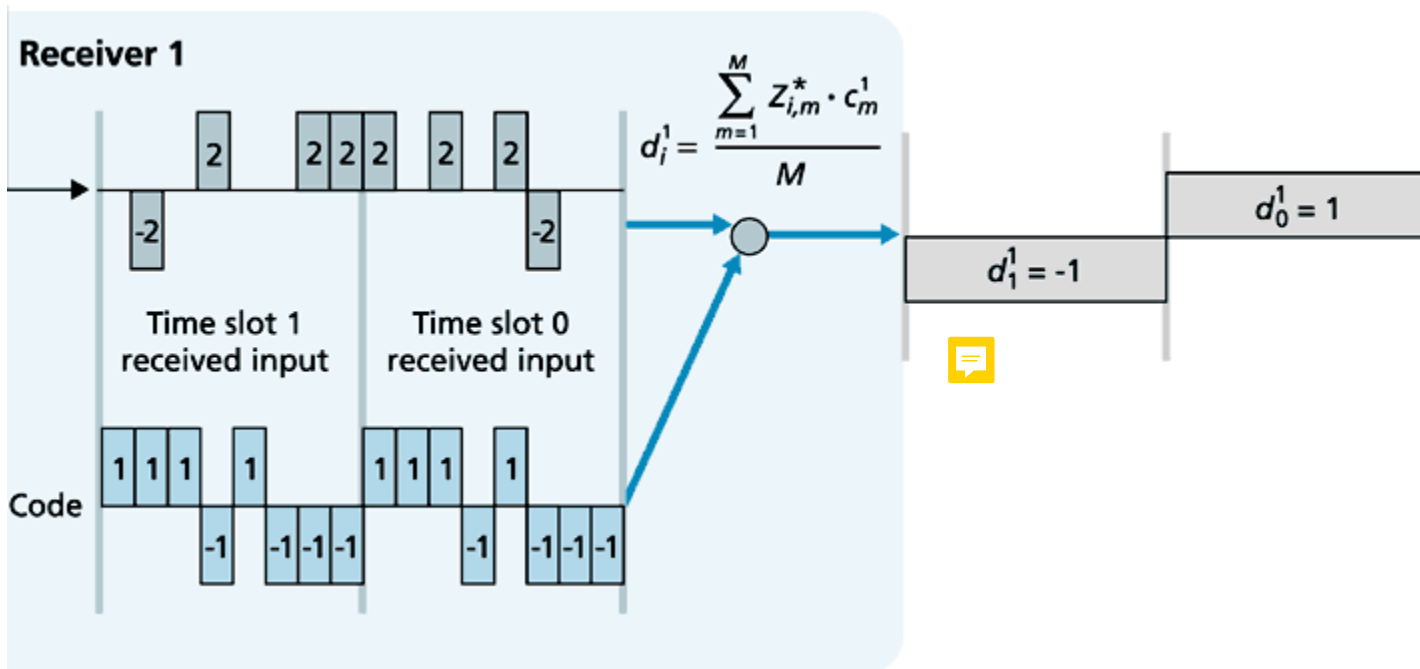
## Channel Partitionning – CDMA, interference (1/2)





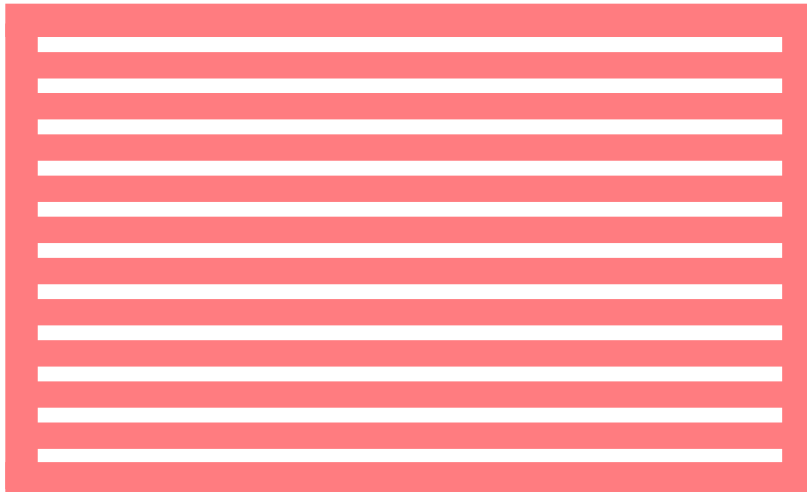
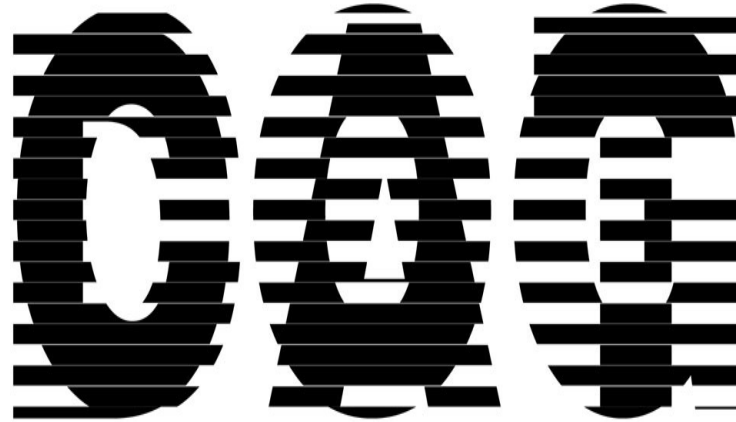
# Multiple Access Protocols

## Channel Partitionning – CDMA, interference (2/2)



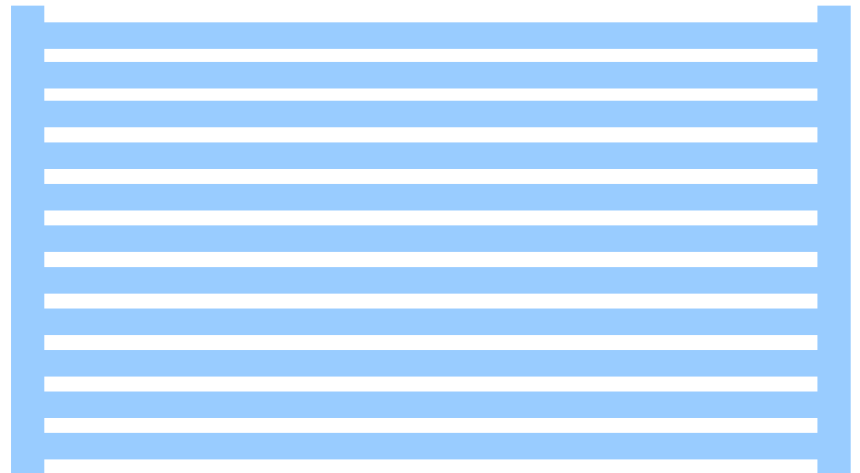
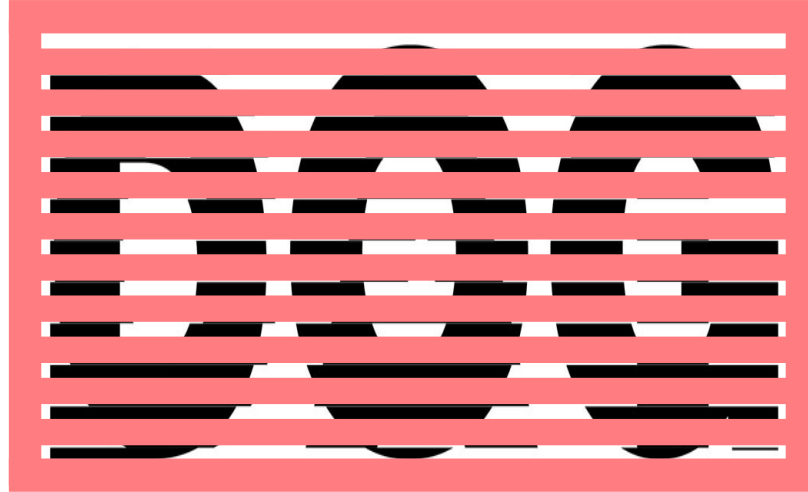
# Multiple Access Protocols

## Channel Partitioning – CDMA, demonstration (1/3)



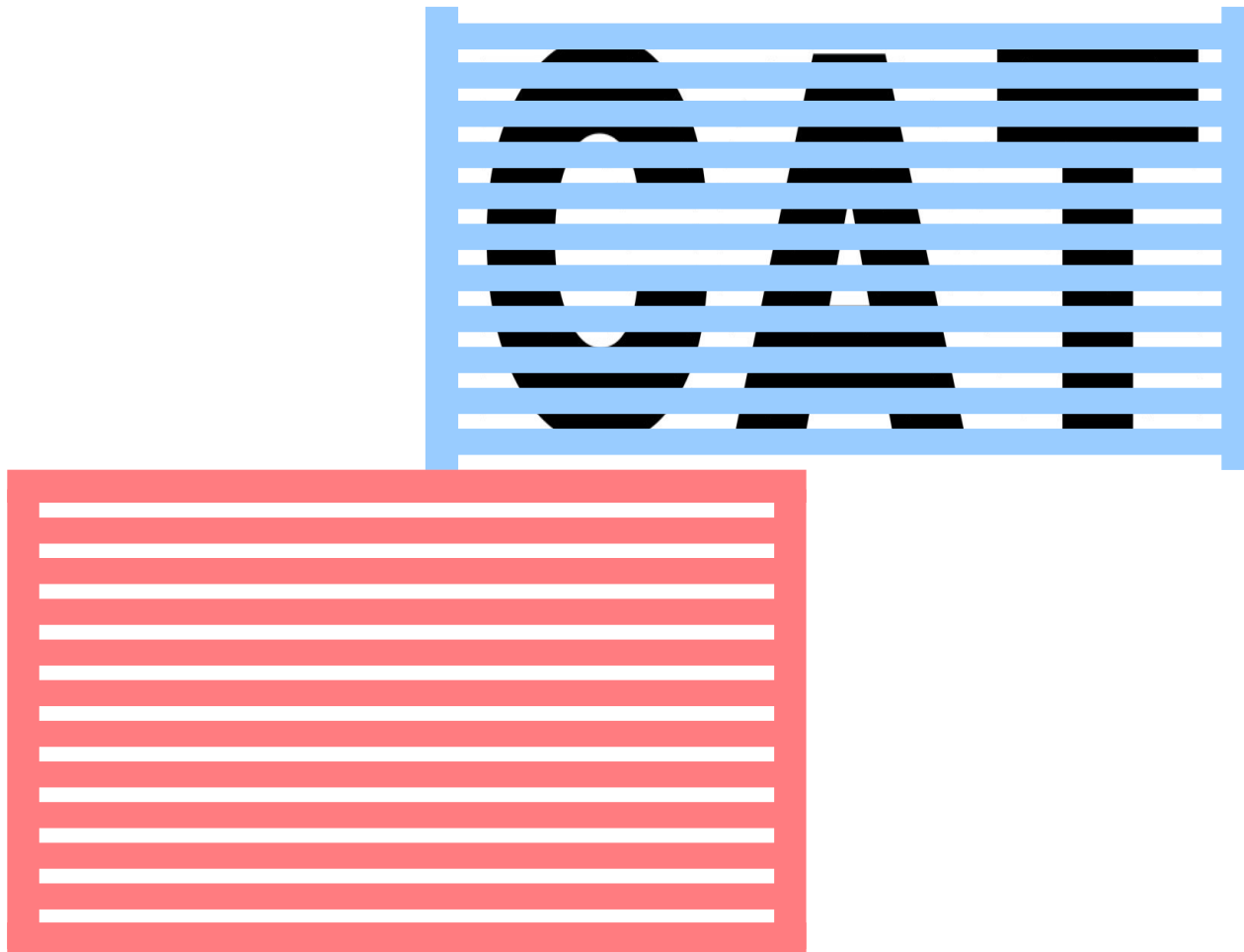
# Multiple Access Protocols

## Channel Partitioning – CDMA, demonstration (2/3)



# Multiple Access Protocols

## Channel Partitioning – CDMA, demonstration (3/3)



# Multiple Access Protocols

## Random Access Protocols

- 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”
- Random access MAC protocols specify
  - 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

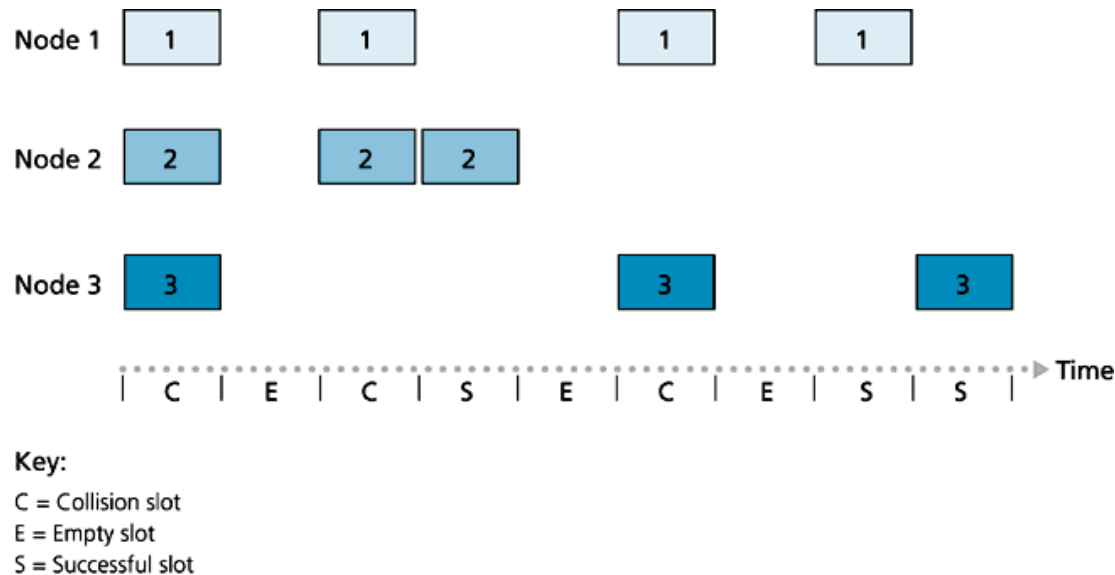
# Multiple Access Protocols

## Random Access Protocols – Slotted ALOHA

- Assumptions
  - All frames have same size ( $L$  bits)
  - Time is divided into equal size slots, corresponding to the time to transmit one frame ( $L/R$  seconds)
  - Nodes start to transmit frames only at beginning of slots
  - Nodes are synchronised
  - If two or more nodes transmit in slot, all nodes detect collision
- Operation
  - When node obtains fresh frame, it transmits in next slot
  - If no collision, node can send new frame in next slot
  - If collision, node retransmits frame in each subsequent slot with probability  $p$  until success

# Multiple Access Protocols

## Random Access Protocols – Slotted ALOHA



Advantages	Drawbacks
<ul style="list-style-type: none"><li>Simple</li><li>Highly decentralised: each node detects collision and decide when to transmit</li><li>Single active node can continuously transmit at full rate</li></ul>	<ul style="list-style-type: none"><li>Collisions result in wasted slots</li><li>Idle slots when nodes refrain from transmitting</li><li>Synchronisation required</li></ul>

# Multiple Access Protocols

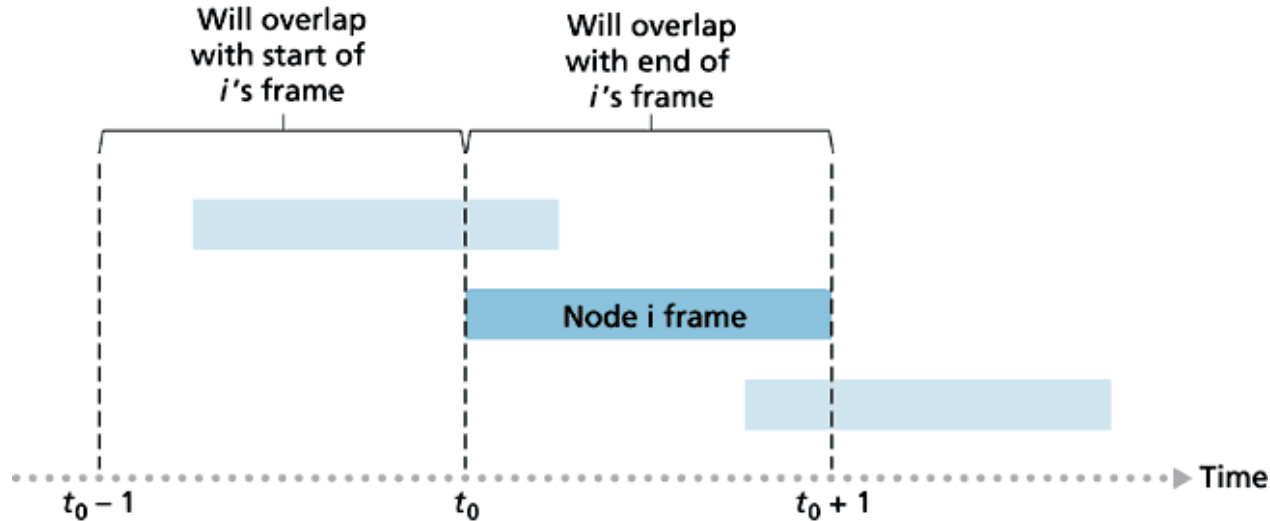
## Random Access Protocols – Slotted ALOHA

- Efficiency is the long-run fraction of successful slots when there are many nodes, each with many frames to send
- Suppose  $N$  nodes with many frames to send, each transmits in slot with probability  $p$
- Probability of success for a given node =  $p(1-p)^{N-1}$
- Probability of success for an arbitrary node =  $Np(1-p)^{N-1}$
- For MAX efficiency with  $N$  active nodes, find  $p^*$  that maximises  $Np^*(1-p^*)^{N-1} \rightarrow p^* = 1/N$
- For large number of nodes, take limit of  $Np^*(1-p^*)^{N-1}$  as  $N \rightarrow \infty$ , gives MAX efficiency =  $1/e = 0.37$
- At best, channel used for useful transmissions 37% of time!



# Multiple Access Protocols

## Random Access Protocols – Pure ALOHA



- Unslotted Aloha
  - Simpler : when frame first arrives, transmit immediately
  - No synchronization
- Collision probability increases → Efficiency  $\leq 18 \%$
- Frame sent at  $t_0$  collides with other frames sent in  $[t_0 - 1, t_0 + 1]$

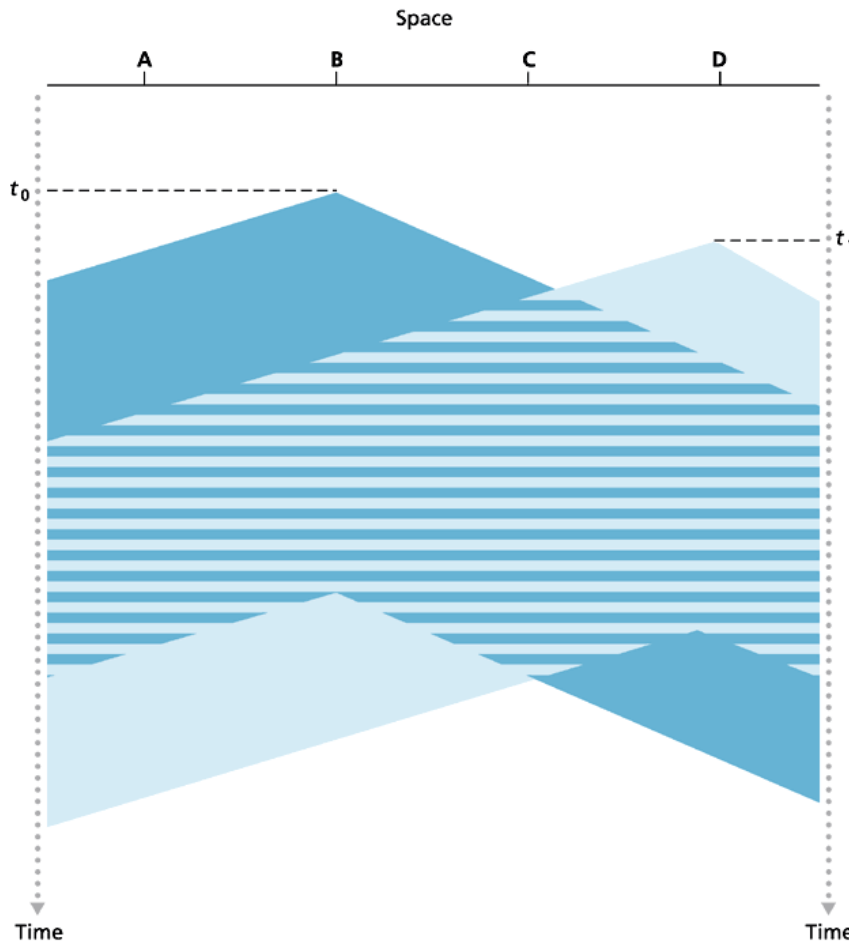
# Multiple Access Protocols

## Random Access Protocols – CSMA

- Carrier Sense Multiple Access (CSMA)
- Contrary to ALOHA, listen before transmit
- If channel sensed idle, transmit entire frame
- If channel sensed busy, defer transmission
- Human analogy
  - ALOHA → boorish partygoer
  - CSMA → conversation

# Multiple Access Protocols

## Random Access Protocols – Collisions in CSMA



- Collisions can still occur: propagation delay means two nodes may not hear each other's transmission
- Distance and propagation delay determine collision probability
- In case of collision, entire packet transmission time is wasted

# Multiple Access Protocols

## Random Access Protocols – CSMA/CD

- Carrier Sense Multiple Access/Collision Detection (CSMA/CD)
- Carrier sensing, deferral as in CSMA
- Collision detection
  - Collisions detected within short time
  - Colliding transmissions aborted, reducing channel wastage
  - Easy in wired LANs: measure signal strengths, compare transmitted, received signals
  - Difficult in wireless LANs: receiver shut off while transmitting
- Human analogy: polite conversation



# Multiple Access Protocols

## Random Access Protocols – CSMA/CD Algorithm (1/2)

- Adapter gets datagram from and creates frame
- If adapter senses channel **idle**, it starts to transmit frame.
- If adapter senses channel **busy**, waits until channel idle and then transmits.
- If adapter transmits entire frame without detecting another transmission, the adapter is **done** with frame.

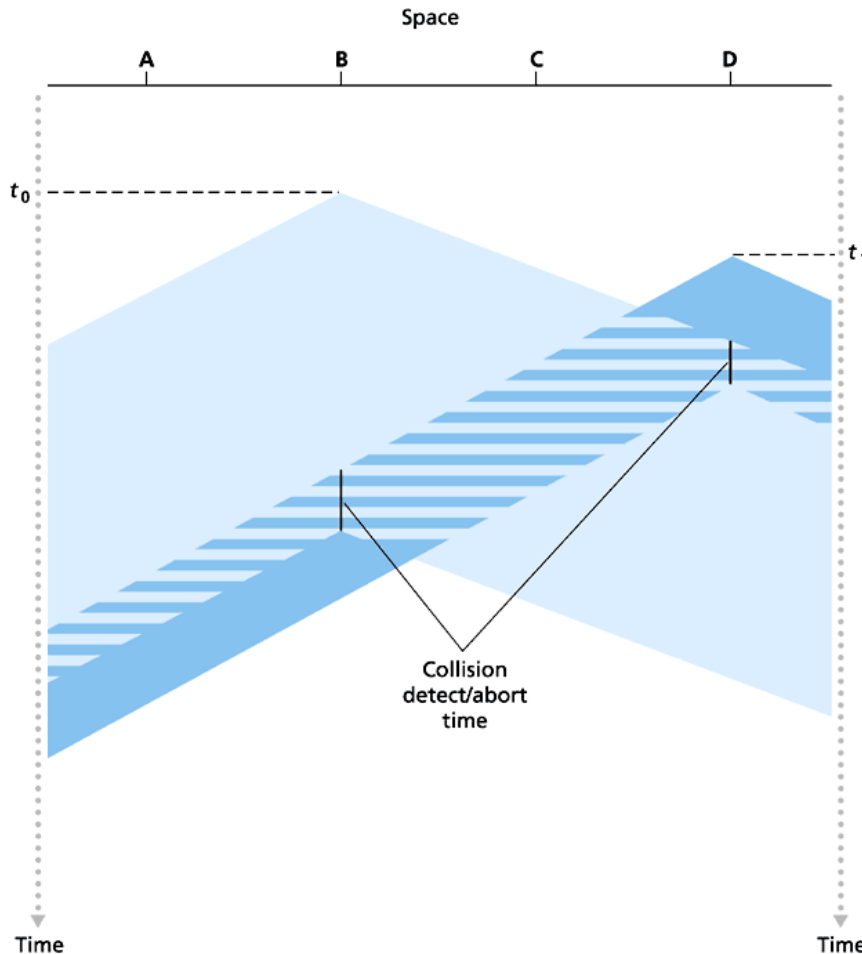
# Multiple Access Protocols

## Random Access Protocols – CSMA/CD Algorithm (2/2)

- If adapter detects another transmission while transmitting, aborts and sends jam signal (48 bits long in Ethernet) to make sure that all other transmitters are aware of **collision**
- After aborting, adapter enters exponential backoff
  - After the  $n$ th collision, adapter chooses  $K$  at random from  $\{0, 1, 2, \dots, 2^m - 1\}$  with  $m = \min(n, 10)$   
  - Adapter waits  $K * 512$  bit times (one bit =  $0.01 \mu\text{s}$  on 100 Mbps Ethernet) and returns to Step 2
  - $K$  grows as adapter experiences more and more collisions, hinting that there are many active adapters

# Multiple Access Protocols

## Random Access Protocols – Collisions in CSMA/CD



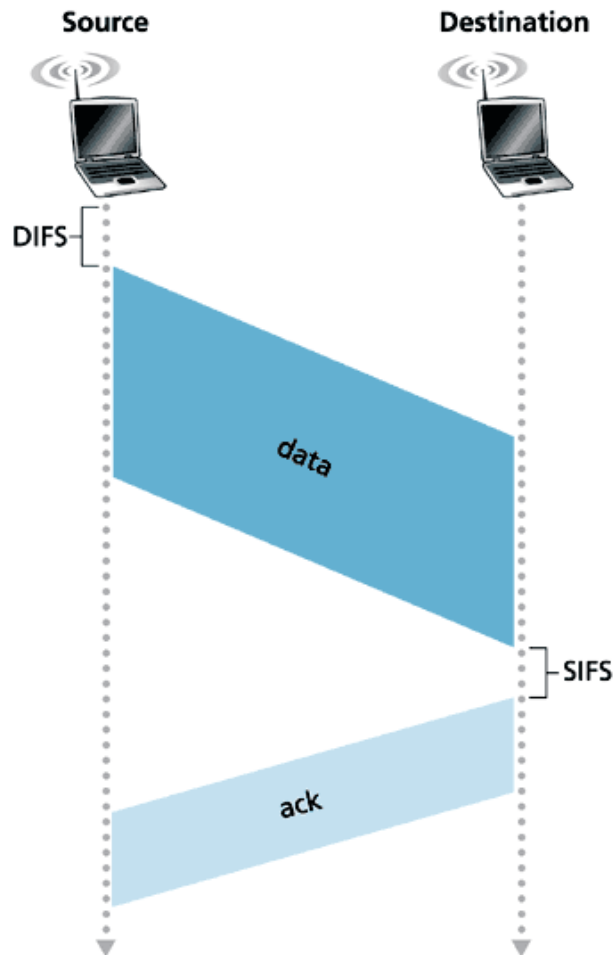
- Collisions detected within short time
- Colliding transmissions aborted, reducing channel wastage



CSMA/CD Applet

# Multiple Access Protocols

## Random Access Protocols – CSMA/CA

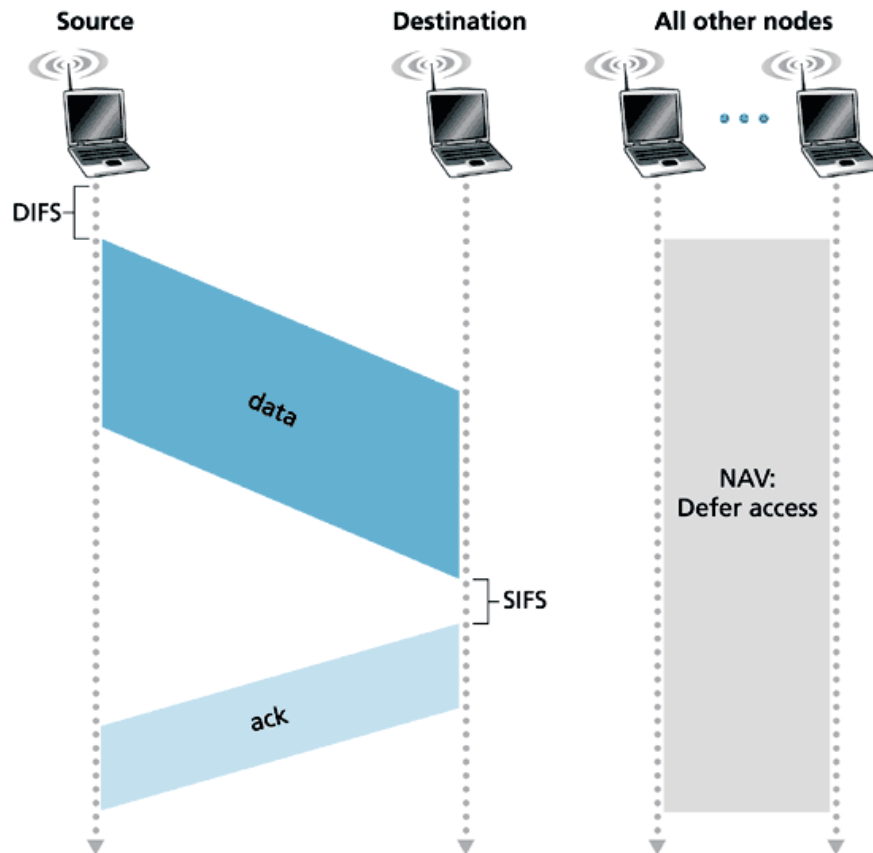


- Sender
  - If sense channel **idle** for DIFS (Distributed Inter Frame Space) then transmit entire frame
  - If sense channel **busy** then backoff
  - Backoff (DIFS + random time doubled whenever channel sensed **busy**)
  - No **collision** detection
- Receiver
  - If received OK, return ACK after SIFS (Short Inter Frame Spacing)
  - ACK required because no **collision** detection



# Multiple Access Protocols

## Random Access Protocols – Collision Avoidance

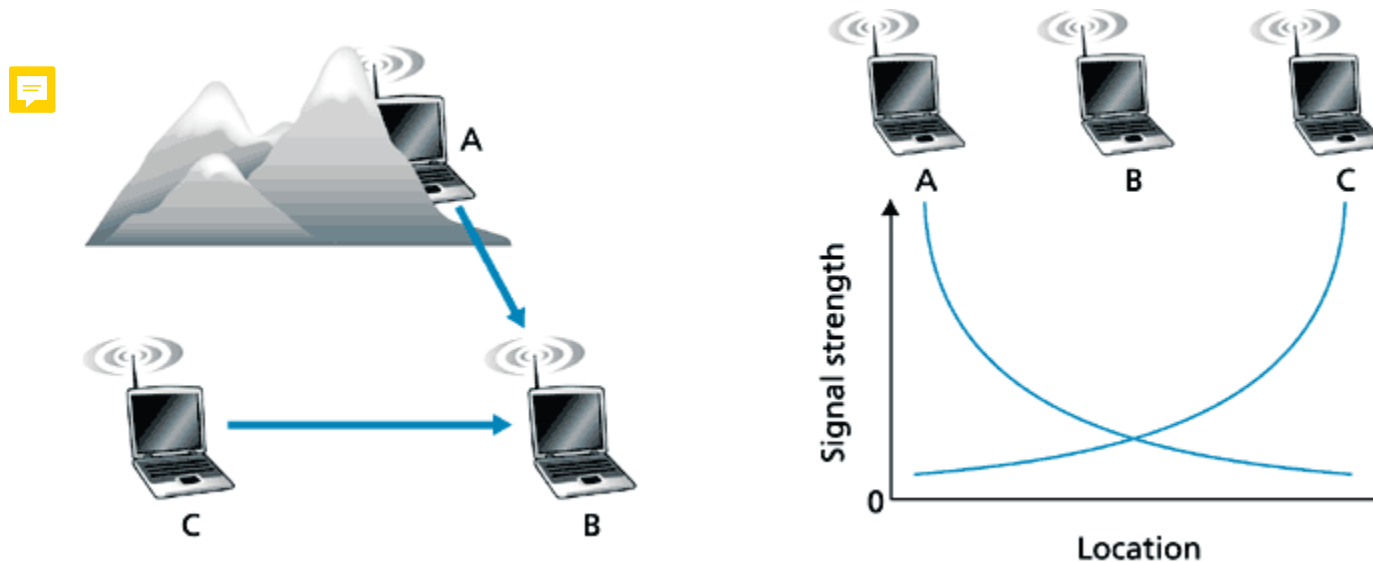


- Idea: avoid collisions instead of detecting and recovering
- Sender indicates duration of its transmission
- Nodes thus know the Network Allocation Vector (NAV), e.g. the amount of time by which they should defer their own transmission

# Multiple Access Protocols

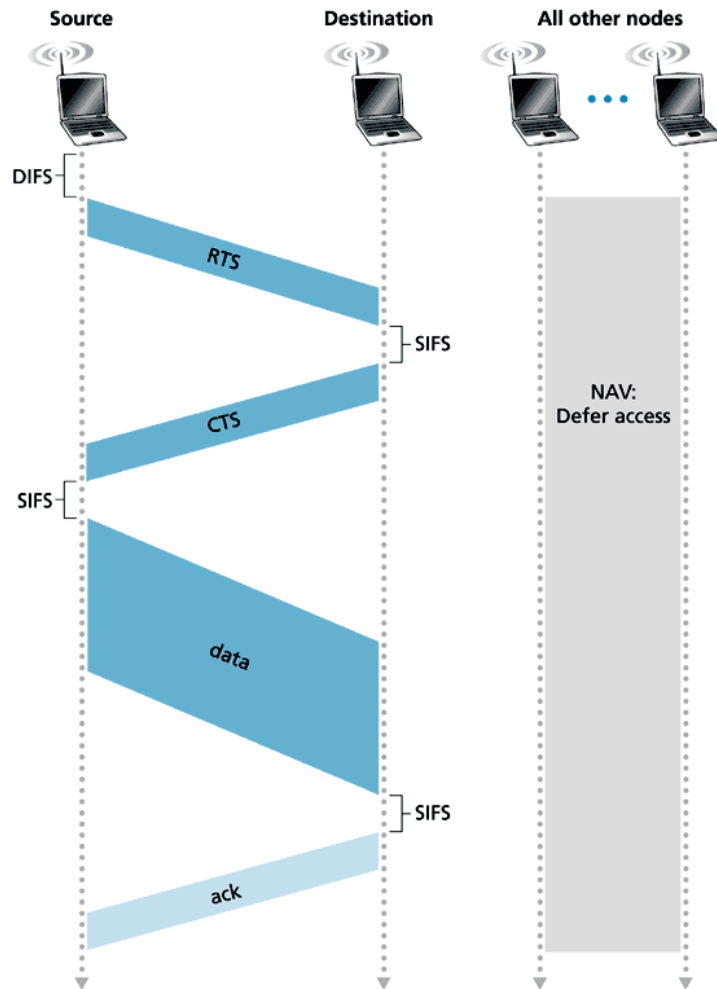
## Random Access Protocols – Hidden Terminal Problem

- CSMA/CA mostly popular for wireless deployments
- Very costly to simultaneously transmit and receive (unless Full Duplex radio) → no collision detection
- Hidden terminal problem → Neither collision detection nor collision avoidance work



# Multiple Access Protocols

## Random Access Protocols – RTS/CTS



- Reservation mechanism
  - Sender transmits short Request To Send (RTS) packet indicating duration of transmission
  - Receiver replies with short Clear To Send (CTS) packet
  - (Possibly hidden) nodes are notified, do not transmit for NAV
- Since RTS and CTS are short, impact of collision negligible

CSMA/CA  
Applet  
Without  
Hidden  
Terminals



CSMA/CA  
Applet  
With  
Hidden  
Terminals



# Multiple Access Protocols

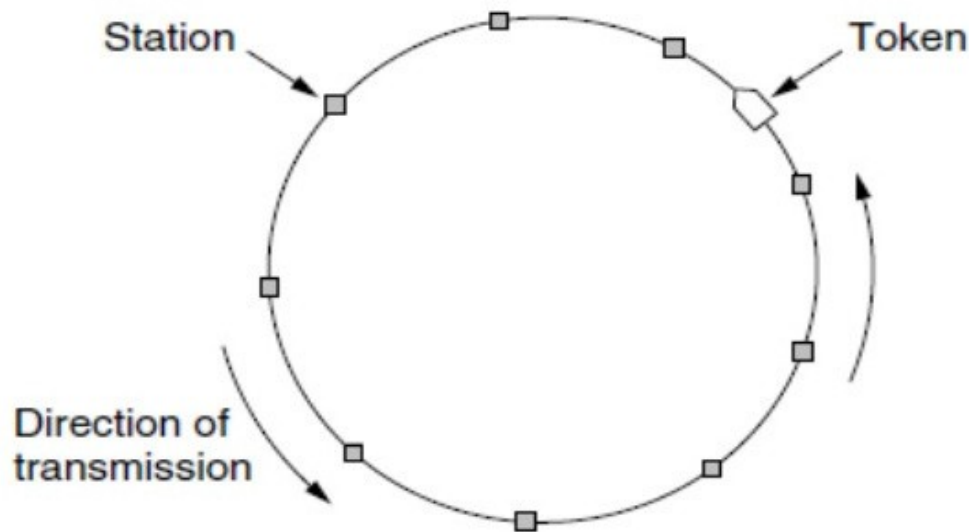
Random Access Protocols – « Taking Turns » protocols

		MAC Protocols	
		Channel partitionning	Random Access
Load	Low	Inefficient: delay in channel access, $1/N$ bandwidth allocated even if only one active node	Efficient: single node can fully utilise channel
	High	Channel share efficiently and fairly	Collision overhead

- Taking Turns protocols : looking for the best out of both worlds

# Multiple Access Protocols

## « Taking Turns » Protocols



- Polling: a master node “invites” slave nodes to transmit in turn
- Token passing: control token passed from one node to next sequentially
- Current: Master-Slave /Token Passing (MS/TP) building automation networks (RS 485)
- Legacy : token bus (IEEE 802.4), token ring (IEEE 802.5), FDDI

Pro's	Con's
No collision No empty slot	Latency SPOF (master or token) Overhead (master election, token generation)

# Outline

Data link layer services

Error-detection and error-correction techniques

Multiple access protocols

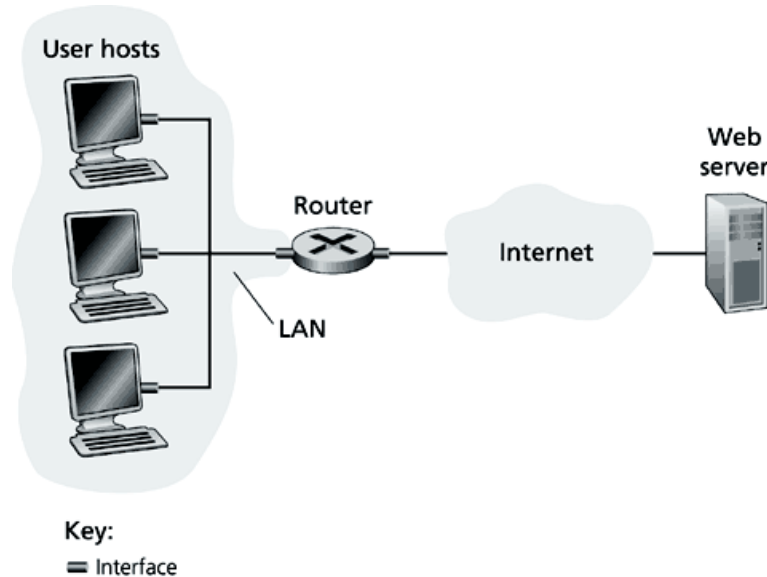
LAN addresses – Neighbour discovery (IPv6), ARP (IPv4)

Ethernet

Hubs, bridges, switches and routers

# LAN Addresses

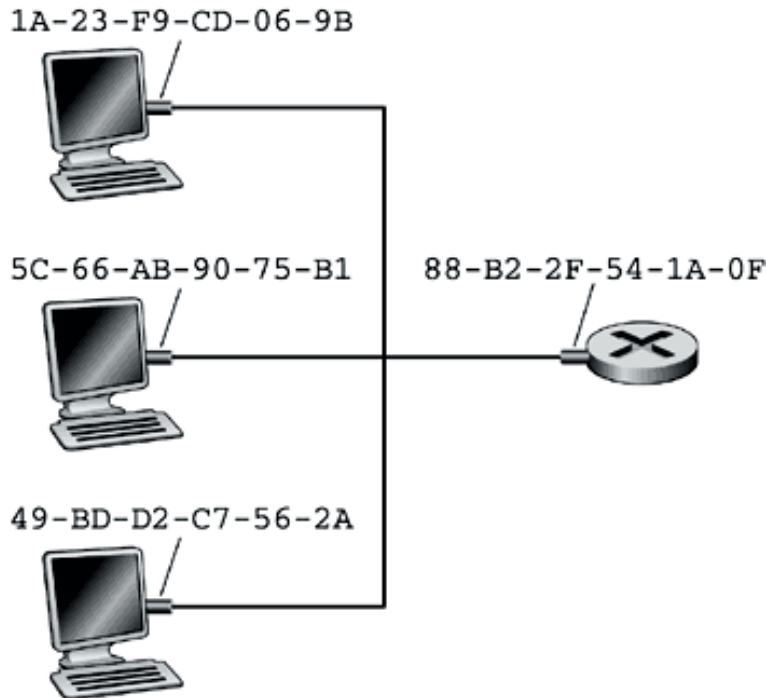
## Introduction



- Network-layer address
  - **128-bit IPv6** address (formerly **32-bit IPV4**)
  - Used to get datagram to destination IP network
- LAN (or MAC or physical or Ethernet) address
  - 48-bit MAC address burned in the adapter ROM → 281.5e12 addresses
  - Used to get datagrams from one interface to another physically-connected interface (same IP network)

# LAN Addresses

## MAC vs. IP addresses



- MAC address is flat  
→ portability
  - Each adapter on LAN has unique address
  - MAC address allocation administered by IEEE
  - Manufacturer buys chunks of  $2^{24}$  MAC addresses → 16.7e6 addresses
- IP address is hierarchical  
→ no portability

- Analogy

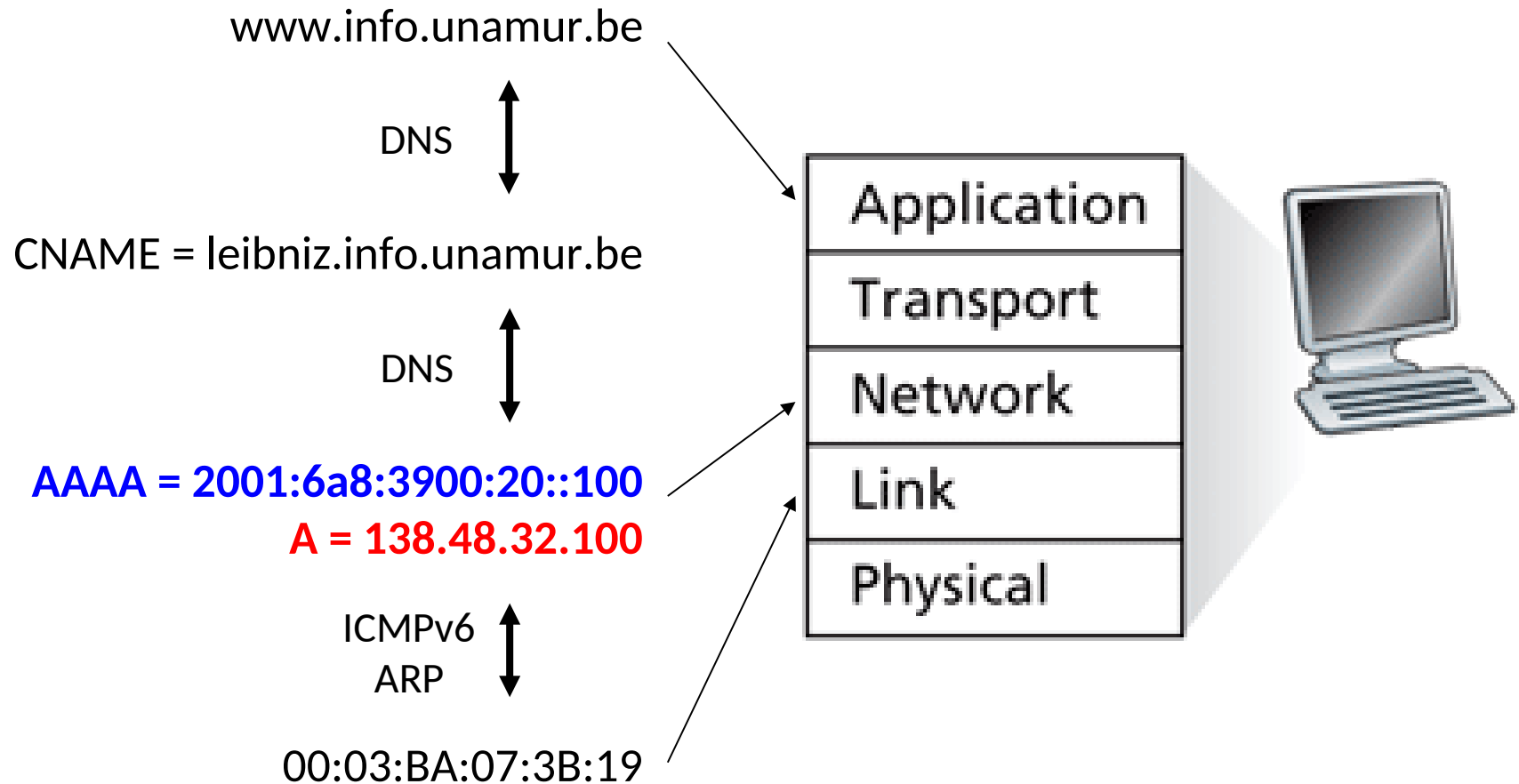
- MAC address = ID number (permanent)
  - IP address = postal address (variable)





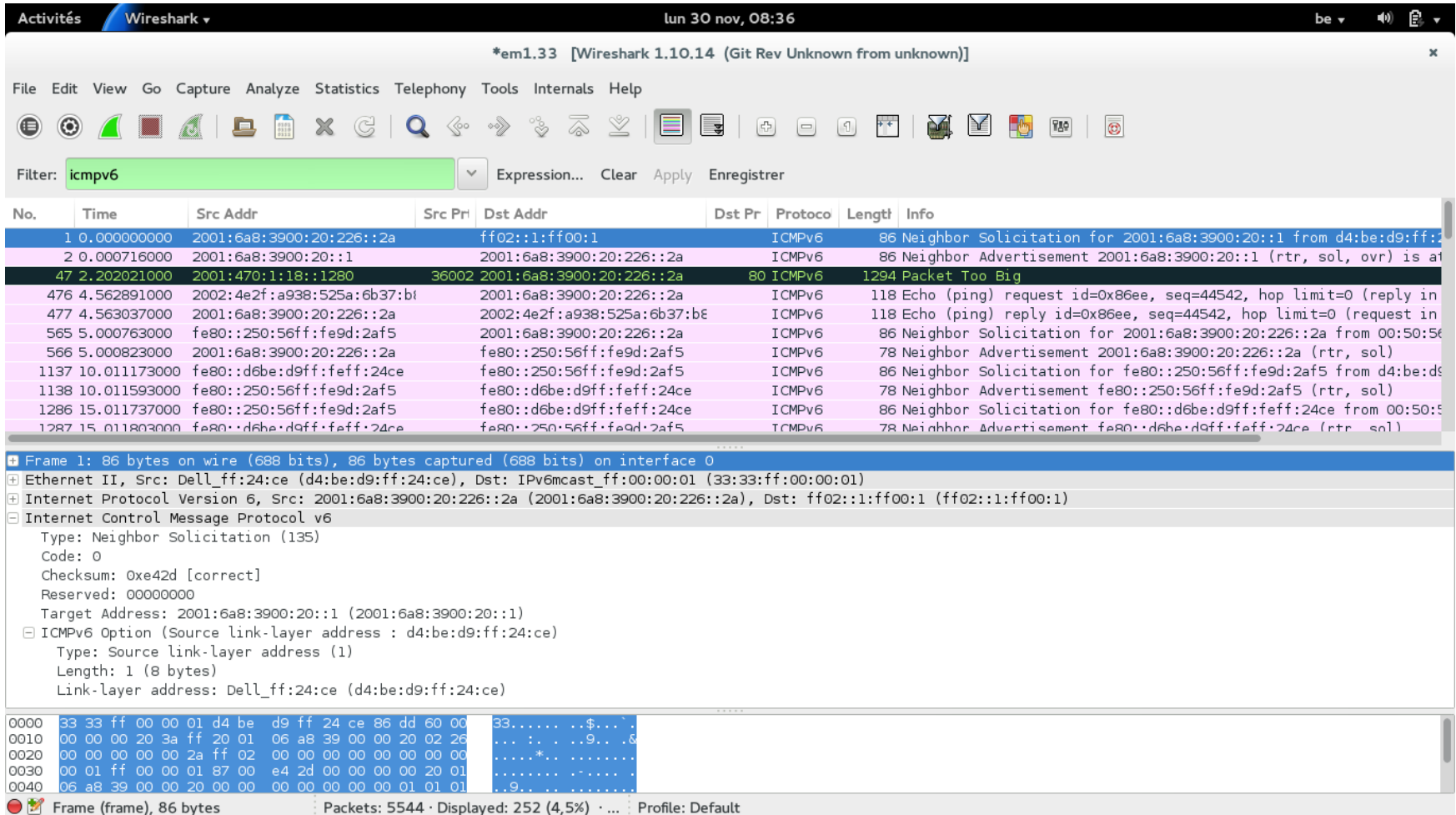
# LAN Addresses

## Addressing schemes



# LAN Addresses

## IPv6 Address Resolution = Neighbour Discovery



The screenshot shows a Wireshark capture of ICMPv6 traffic. The filter is set to 'icmpv6'. The packet list shows several Neighbor Solicitation (NS) and Neighbor Advertisement (NA) messages. The packet details pane shows the structure of the first packet (Frame 1):

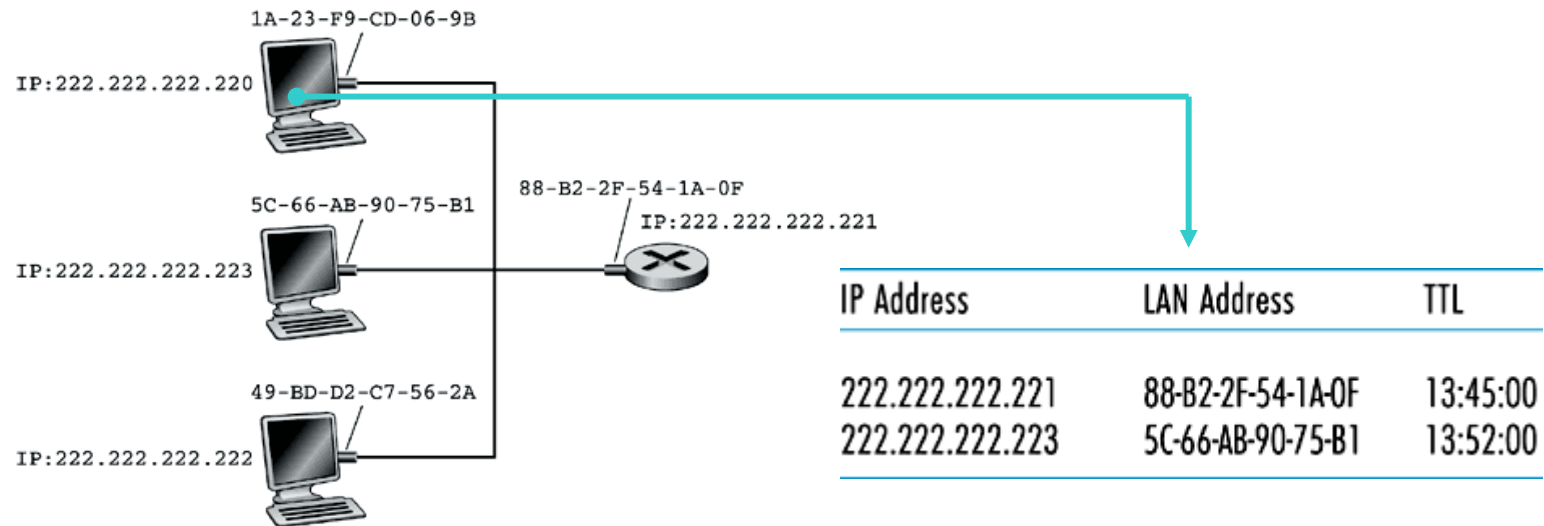
- Frame 1: 86 bytes on wire (688 bits), 86 bytes captured (688 bits) on interface 0
- Ethernet II, Src: Dell\_ff:24:ce (d4:be:d9:ff:24:ce), Dst: IPv6mcast\_ff:00:00:01 (33:33:ff:00:00:01)
- Internet Protocol Version 6, Src: 2001:6a8:3900:20:226::2a (2001:6a8:3900:20:226::2a), Dst: ff02::1:ff00:1 (ff02::1:ff00:1)
- Internet Control Message Protocol v6
  - Type: Neighbor Solicitation (135)
  - Code: 0
  - Checksum: 0xe42d [correct]
  - Reserved: 00000000
  - Target Address: 2001:6a8:3900:20:226::1 (2001:6a8:3900:20:226::1)
  - ICMPv6 Option (Source link-layer address : d4:be:d9:ff:24:ce)
    - Type: Source link-layer address (1)
    - Length: 1 (8 bytes)
    - Link-layer address: Dell\_ff:24:ce (d4:be:d9:ff:24:ce)

The packet bytes pane shows the raw data of the packet, including the Ethernet II header, IPv6 header, and ICMPv6 payload.

- Linux: `$ ip -6 neigh show`

# LAN Addresses


## IPv4 Address Resolution = ARP



- Each IP node on a LAN has an ARP table
- ARP Table contains IPv4/MAC address mappings
- Record format **<IP address; MAC address; TTL>**
- TTL is typically 20 minutes
- Soft state information: time out unless refreshed
- Linux: **\$ arp -a**

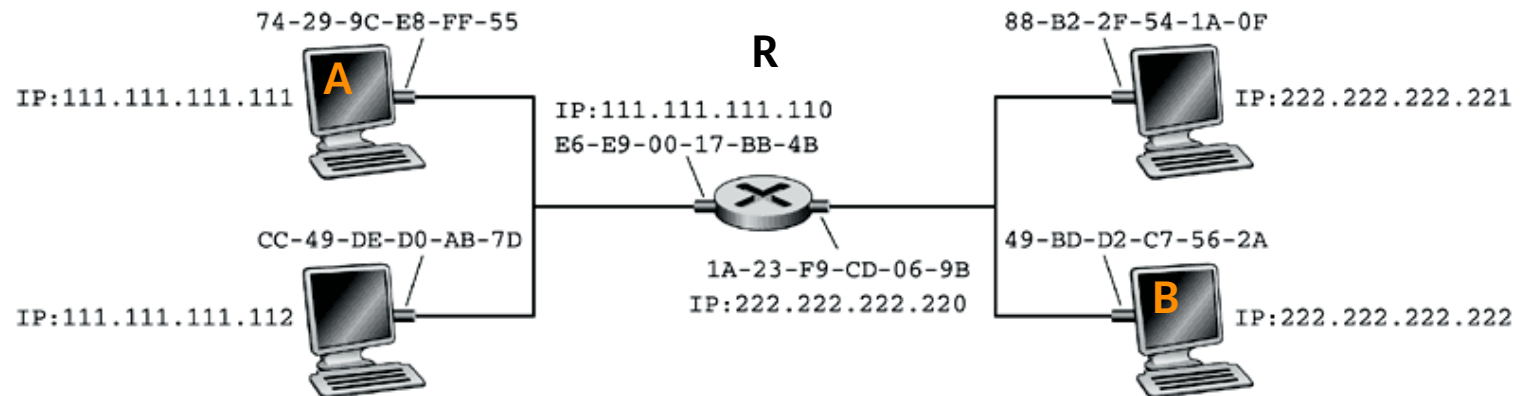
# LAN Addresses

## Address Resolution

- Assume 
  - A wants to send datagram to *B*
  - A knows *B*'s **IPv6** / **IPv4** address.
  - *B*'s MAC address is not in A's **Neighbours** / **ARP** table
- Query
  - A **multi-**/ **broad**-casts ICMPv6 / ARP query packet, containing *B*'s IP address
  - All machines on LAN receive query, including *B*
- Reply
  - *B* replies to A with its (*B*'s) MAC address
  - Frame sent to A's MAC address (unicast)
  - A caches IP-to-MAC mapping in its **Neighbours** / **ARP** table until TTL elapses
- “Plug-and-play”: nodes create their mapping tables without external intervention

# LAN Addresses

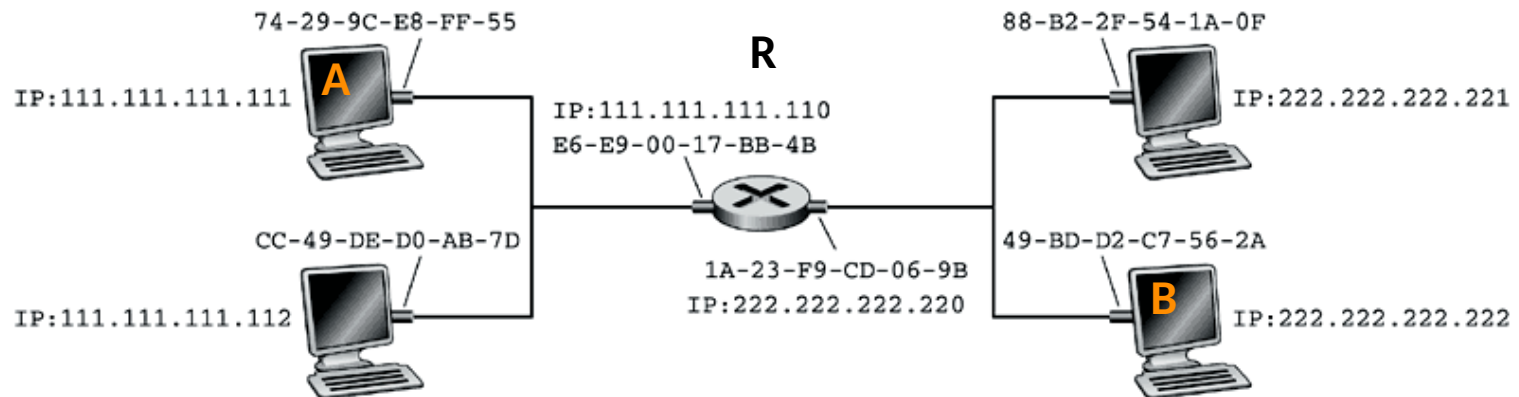
## Sensing a datagram to a node off the LAN (1/2)



- Send datagram from A to B via R
- Assume
  - A knows B's IPv4 address
  - Two ARP tables in router R, one for each IP network (LAN)
- A creates datagram with source A, destination B. Actually, should send to R
- A uses ARP to get R's MAC address for 111.111.111.110

# LAN Addresses

## Sensing a datagram to a node off the LAN (2/2)



- A creates link-layer frame with R's MAC address as destination, sends it to R's data link layer
- Frame contains A-to-B IPv4 datagram
- R removes IPv4 datagram from Ethernet frame, sees its destination is B
- R uses ARP to get B's MAC address
- R creates new frame containing A-to-B IPv4 datagram and sends it to B

# Outline

Data link layer services

Error-detection and error-correction techniques

Multiple access protocols

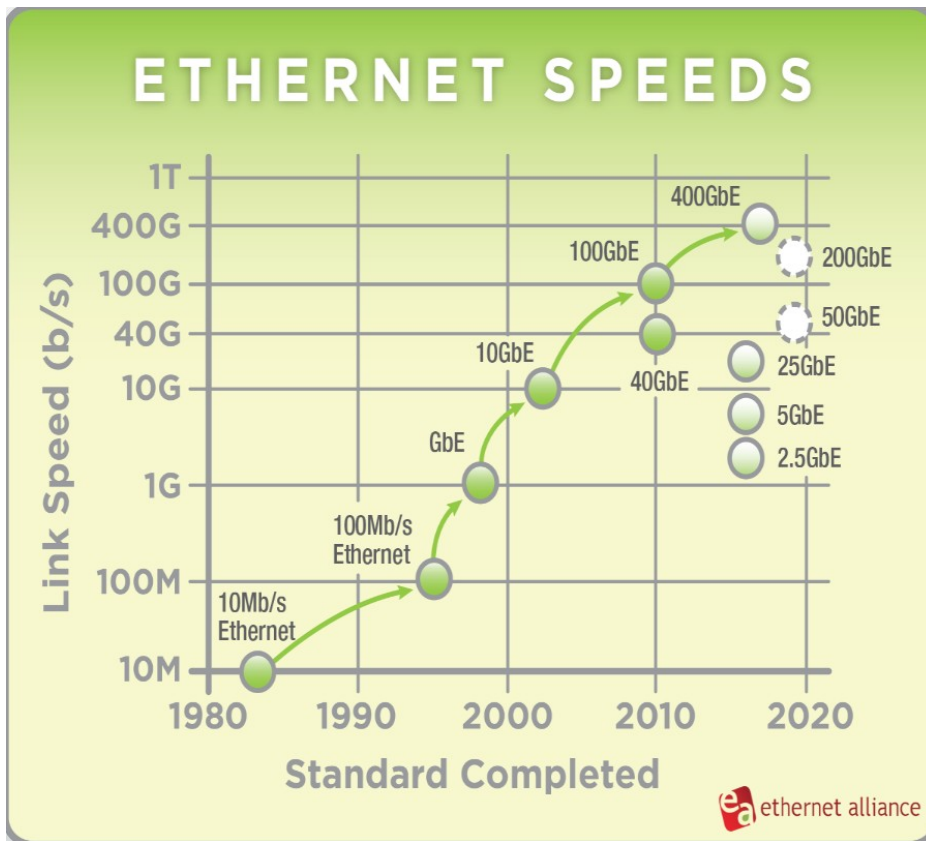
LAN addresses – Neighbour discovery (IPv6), ARP (IPv4)

Ethernet

Hubs, bridges, switches and routers

# Ethernet

## Introduction

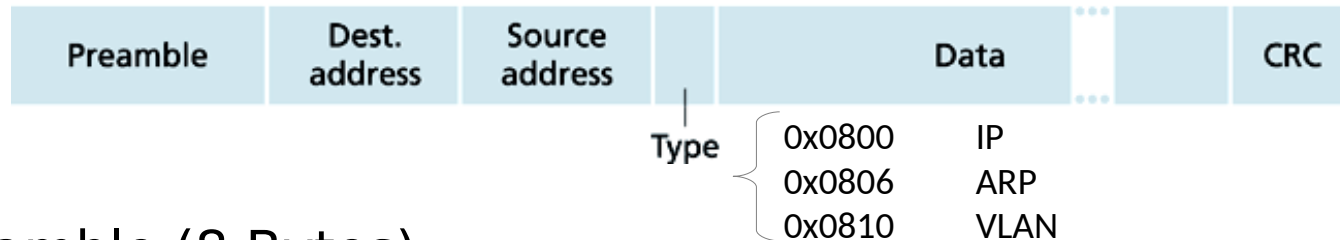


- Dominant LAN technology
- Used to be challenged by token ring, FDDI, ATM
- Success story
  - Large-scale deployment → cheap hardware
  - Simpler design (full decentralisation) and cheaper than token LANs and ATM
  - Kept up with speed race: from  $10^1$  Mbps up to  $10^5$  Mbps
- Versatile
  - Bus or star topology
  - Over coax, twisted pair or fiber optics



# Ethernet

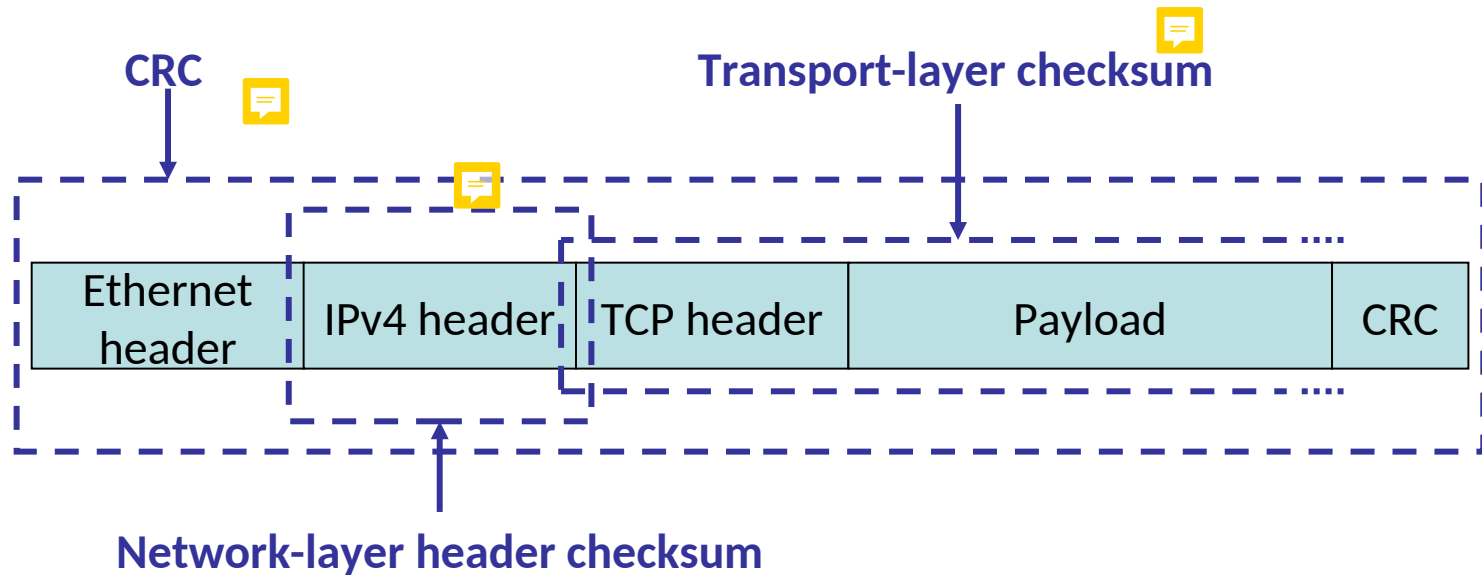
## Frame Structure (1/2)



- Preamble (8 Bytes)
  - Used to synchronise clock rates
  - 7 bytes with pattern 10101010 followed by 1 byte with pattern 10101011
- Addresses (6 Bytes each)
  - If adapter receives frame with matching destination address, or with broadcast address, it passes data in frame to network-layer protocol
  - Otherwise, adapter discards frame
- Type (2 Bytes) : indicates the network layer protocol
- Data (46 to 1500 Bytes): sending adapter encapsulates network layer protocol PDU in Ethernet frame
- CRC (4 Bytes): if error detected at receiver, frame is dropped

# Ethernet

## Frame Structure (2/2)



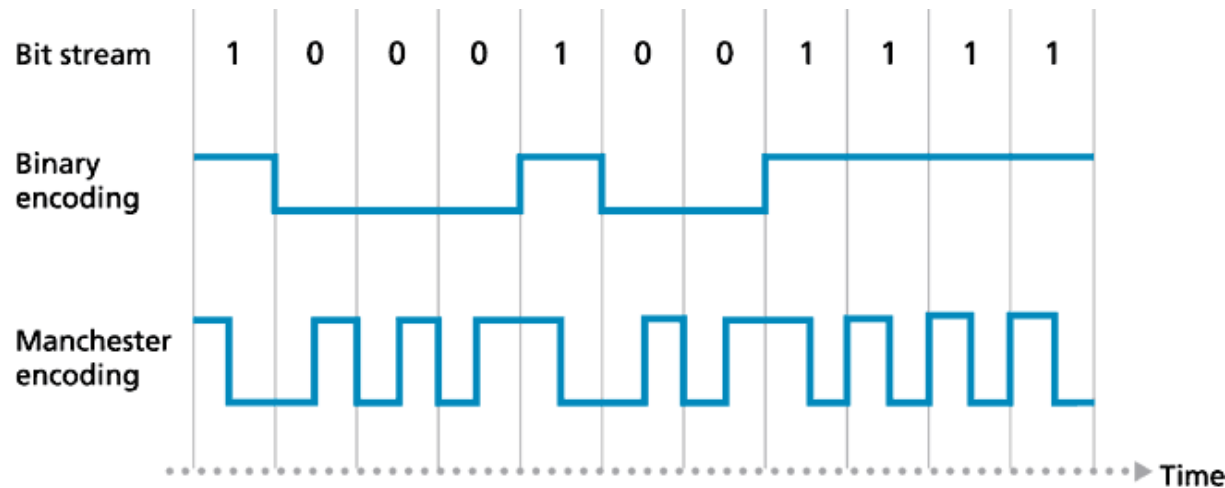
# Ethernet

## Service description

- Connectionless: no handshaking between sending and receiving adapter
- Unreliable
  - If CRC check succeeds, frame payload passed to network layer, but no ACK sent to source
  - If CRC check fails, frame discarded but no NAK sent to source
  - Stream of datagrams passed to network layer can have gaps
  - Gaps will be filled if application is using TCP
  - Otherwise, application will see the gaps

# Ethernet

## Baseband transmission – Manchester encoding



- Baseband: digital signal directly sent on medium as electric signal taking a limited set of values ( $\neq$  xDSL, cable modem)
- Each bit has a transition
- Allows clocks in nodes to synchronise to each other
- No need for a centralised, global clock
- Actually a PHY-layer issue

# Ethernet

## MAC protocol

- CSMA/CD
- No slots
- Adapter doesn't transmit if it senses that some other adapter is transmitting (CSMA)
- Transmitting adapter aborts when it senses that another adapter is transmitting (CD)
- Before attempting a retransmission, adapter waits a random period of time, typically small with respect to frame duration

# Ethernet

## MAC protocol – Efficiency

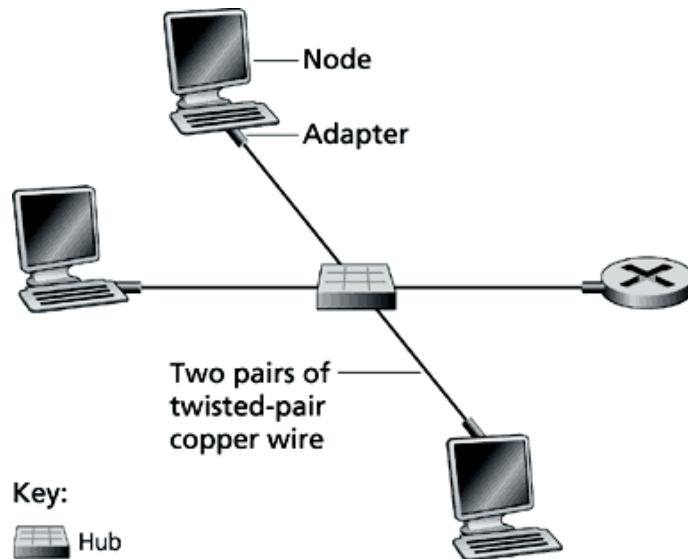
- Assume
  - $t_{prop}$  = MAX propagation time between 2 nodes in LAN
  - $t_{trans}$  = time to transmit MAX-size frame

$$\text{Efficiency} = \frac{1}{1 + 5 \left( \frac{t_{prop}}{t_{trans}} \right)}$$

- Efficiency goes to 1 as
  - $t_{prop}$  goes to 0 (close nodes → collision quickly detected)
  - $t_{trans}$  goes to  $\infty$
- Much better than ALOHA, but still decentralised, simple, and cheap

# Ethernet

## Fast Ethernet – 10BaseT, 100BaseT



- 10/100 → 10/100 Mbps rate
- T → Twisted Pair
- Nodes connect to a hub: “star topology”
- MAX 100m distance between nodes and hub
- Manchester for 10BaseT, 4B5B for 100 Base T

- Hubs are essentially physical-layer repeaters
  - Bits coming in one link go out all other links
  - No frame buffering
  - No CSMA/CD at hub: adapters detect collisions
  - Provides net management functionalities (auditing, automatic disconnection of malfunctioning adapters)

# Ethernet

## Gigabit Ethernet

- Use standard Ethernet frame format
- Backward compatible with 10BaseT and 100BaseT
- Supports both point-to-point links and shared broadcast channels
  - Point-to-point links : full-duplex at 1 Gbps using switches
  - Shared mode
    - CSMA/CD is used with hubs, called here “Buffered Distributors”
    - Short distances between nodes to be efficient
- IEEE standards
  - 802.3z = 1 Gbps
  - 802.3ae = 10 Gbps



# Ethernet

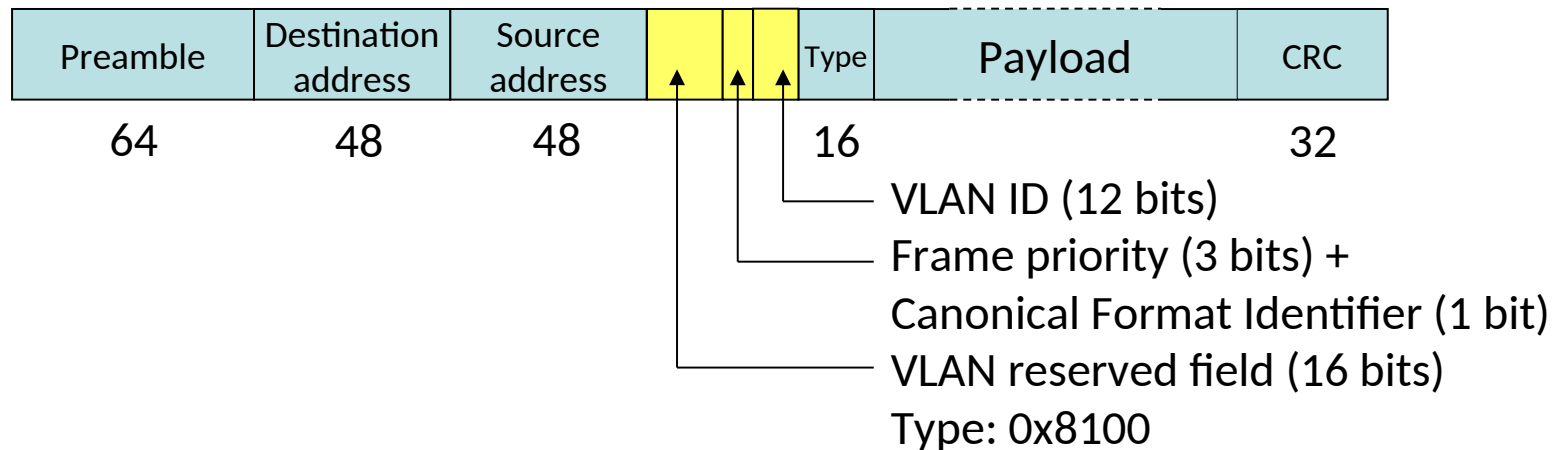
## IEEE 802.1Q – Virtual LANs

- Purposes
  - Restrict access to network resources without regard to physical topology
  - Reduce the size of the collision domain in a large single Ethernet segment
  - Interconnect distant LANs (Metro-Ethernet)
- Configuration
  - Switch port based: port ↔ VLAN membership
  - Layer 2 based: MAC address ↔ VLAN membership
  - IP subnet based: IP address ↔ VLAN membership
  - Higher protocol based

# Ethernet

## IEEE 802.1Q – Virtual LANs

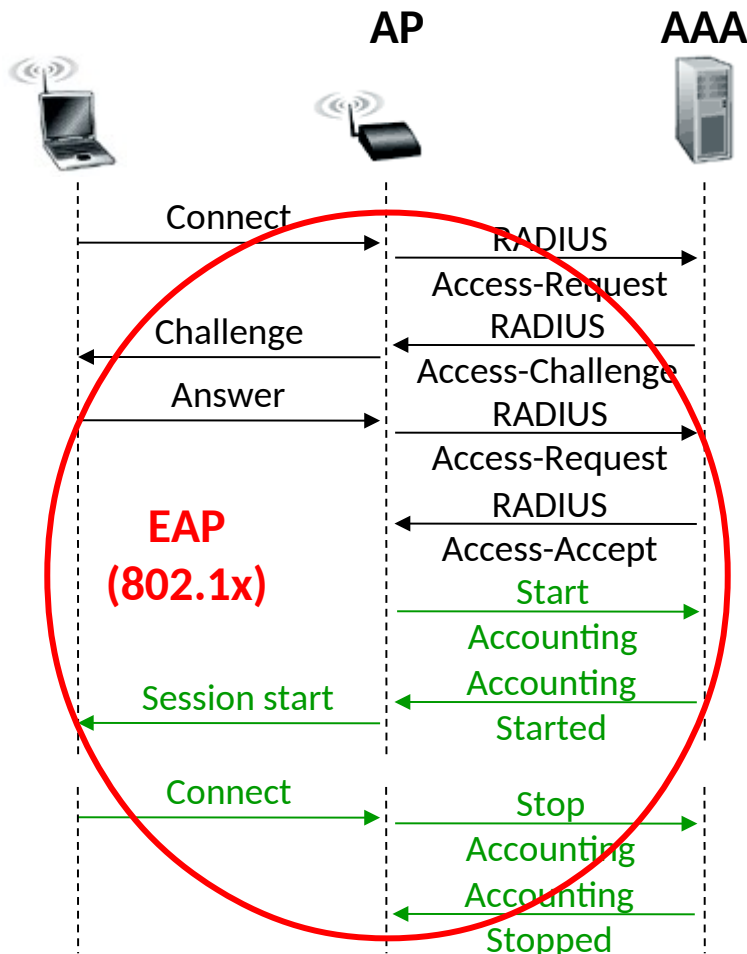
- Establishment
  - Frame-tagging



- Frame-filtering
- Scalability issue:  $2^{12} = 4,096$  simultaneous VLANs

# Ethernet

## IEEE 802.1X – Access Control



- Remote Authentication Dial In User Service (RADIUS) in Wi-Fi context
- Instead of requiring every AP to maintain a list of authorised users, RADIUS Access-Requests are forwarded to AAA Server
- Access-Request attributes: user's name, user's password, port ID, etc.
- AAA server checks user's authentication and autorisation
- RADIUS part of Extensible Authentication Protocol (EAP) in IEEE 802.1x standard

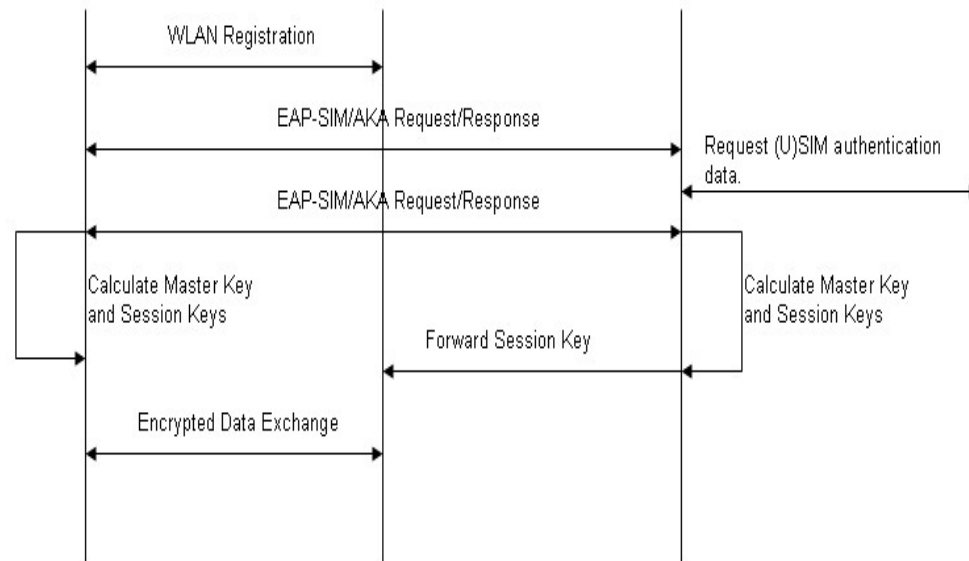
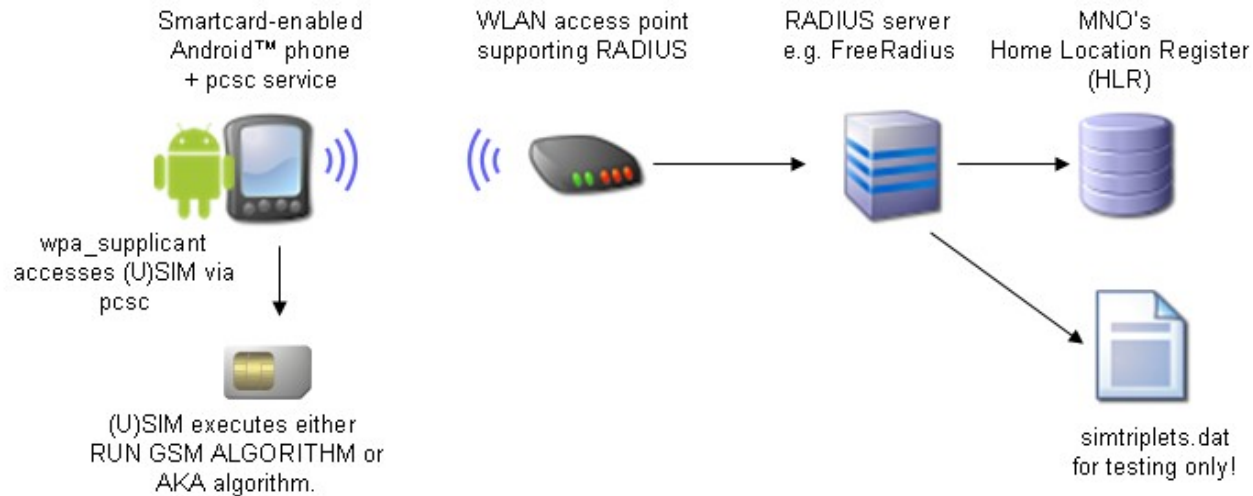
# Ethernet

## RADIUS vs. DIAMETER

- RADIUS not scalable
  - 8-bit ID → 256 pending requests
  - No provision for congestion control
- Diameter planned as RADIUS successor (RFC 3588)
  - 32-bit ID →  $4.3 \cdot 10^9$  pending requests
  - Congestion control
  - Based on TCP or SCTP
  - Other features
    - P2P, any node can issue request
    - More secure thanks to the use of IPSec and TLS

# Ethernet

## EAP-SIM – Proximus Smart WiFi (launched October 2015)



# Outline

Data link layer services

Error-detection and error-correction techniques

Multiple access protocols

LAN addresses – Neighbour discovery (IPv6), ARP (IPv4)

Ethernet

Hubs, bridges, switches and routers

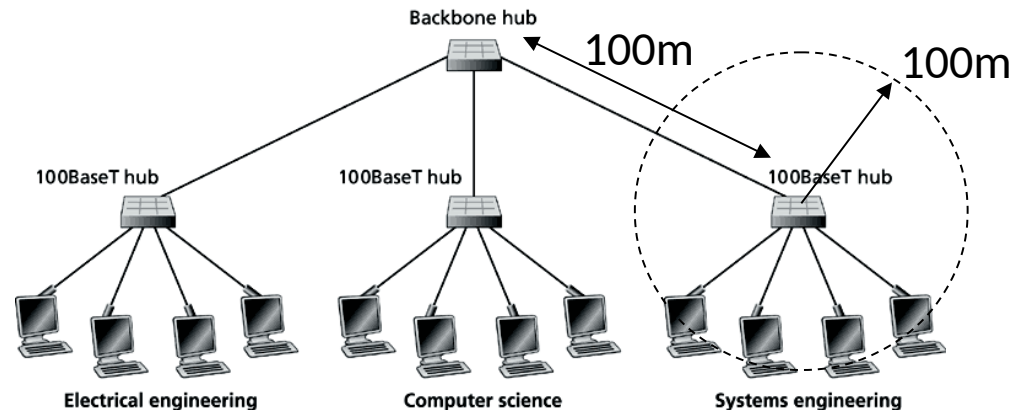
# Hubs, bridges, switches and routers

## Interconnecting LANs

- Usually several departments within a given company/institution/university/school
- Each department has its own LAN
- How to interconnect these LANs?
- Four options
  1. Hubs
  2. Bridges
  3. Switches
  4. Routers

# Hubs, bridges, switches and routers

## Interconnecting with hubs



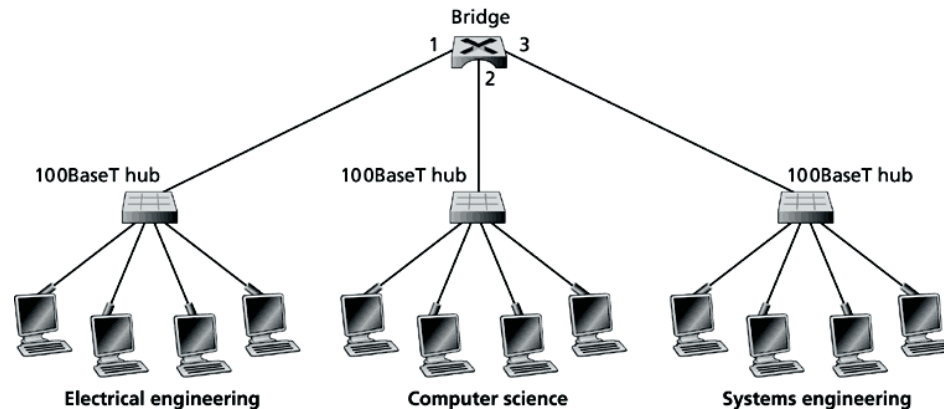
- Hubs are physical-layer devices (L1)
- Backbone hub interconnects LAN segments
- Extends reach (but still MAX 100m between nodes)
- Drawbacks
  - Individual collision domains become a single one
  - Three 100Mbps LAN become a single 100 Mbps network
  - Can not interconnect different technologies
  - Reach extended but still limited





# Hubs, bridges, switches and routers

## Interconnecting with bridges (legacy)

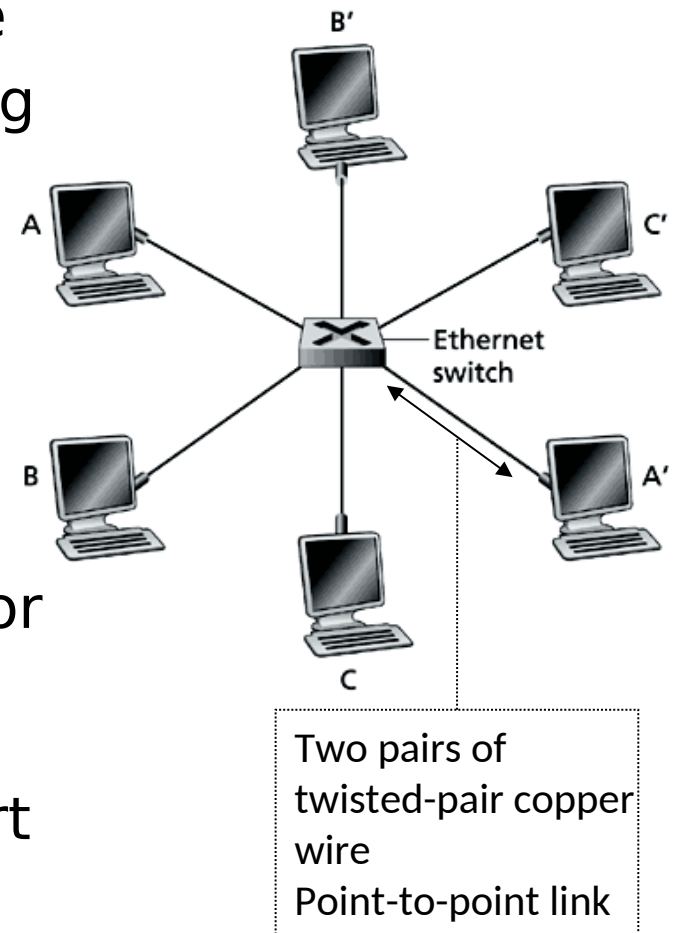


- Bridges are link-layer devices (L2)
  - Store and forward Ethernet frames
  - Examine frame header and selectively forward frame based on MAC destination address → Preserve distinct collision domains
  - When frame is to be forwarded on segment, use CSMA/CD to access segment
- Transparent: hosts are unaware of presence of bridges
- Plug-and-play, self-learning: no configuration required

# Hubs, bridges, switches and routers

## Interconnecting with switches

- Essentially a multi-interface bridge
- Layer 2 (frame) forwarding, filtering using LAN addresses
- Full-duplex: A-to-A' and B-to-B' simultaneously, without collisions
- Direct, dedicated host-switch connections (upstream + downstream)
- Ethernet without carrier sensing nor collision detection
- Cut-through switching: frame forwarded from input to output port without awaiting for assembly of entire frame → reduced latency



# Hubs, bridges, switches and routers

## Self-learning bridges/switches

- Bridges and switches have mapping tables
- Record in table
  - (LAN Address, Interface, Time Stamp)
  - Stale entries in table dropped
  - TTL can be 60 min

```
cr1-32.net.fundp.ac.be (138.48.32.1) at b8:af:67:63:6d:2d [eno1]  
monitor.info.fundp.ac.be (138.48.33.180) at 54:54:00:1c:3d:57 [eno2]  
dci1.info.fundp.ac.be (138.48.32.5) at b8:af:67:d5:57:a2 [eno3]
```

- Devices learn which host can be reached through which interface
- When a frame is received from a given sender, the device
  - Learns sender's details (LAN address and interface)
  - Records sender's details in table

# Hubs, bridges, switches and routers

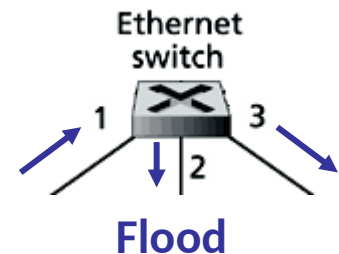
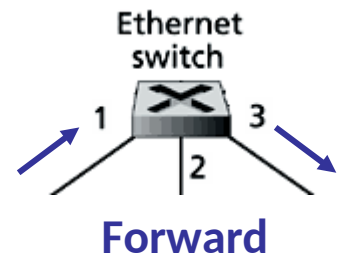
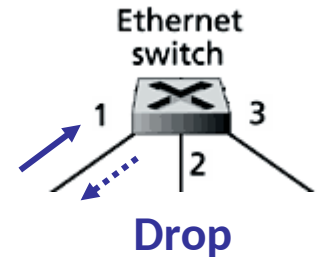
## Filtering/forwarding

When bridge/switch receives a frame  
index table using MAC destination address  
if entry found for destination  
then {

if destination on segment  
from which frame arrived  
then **drop** the frame  
else **forward** the frame  
on interface indicated

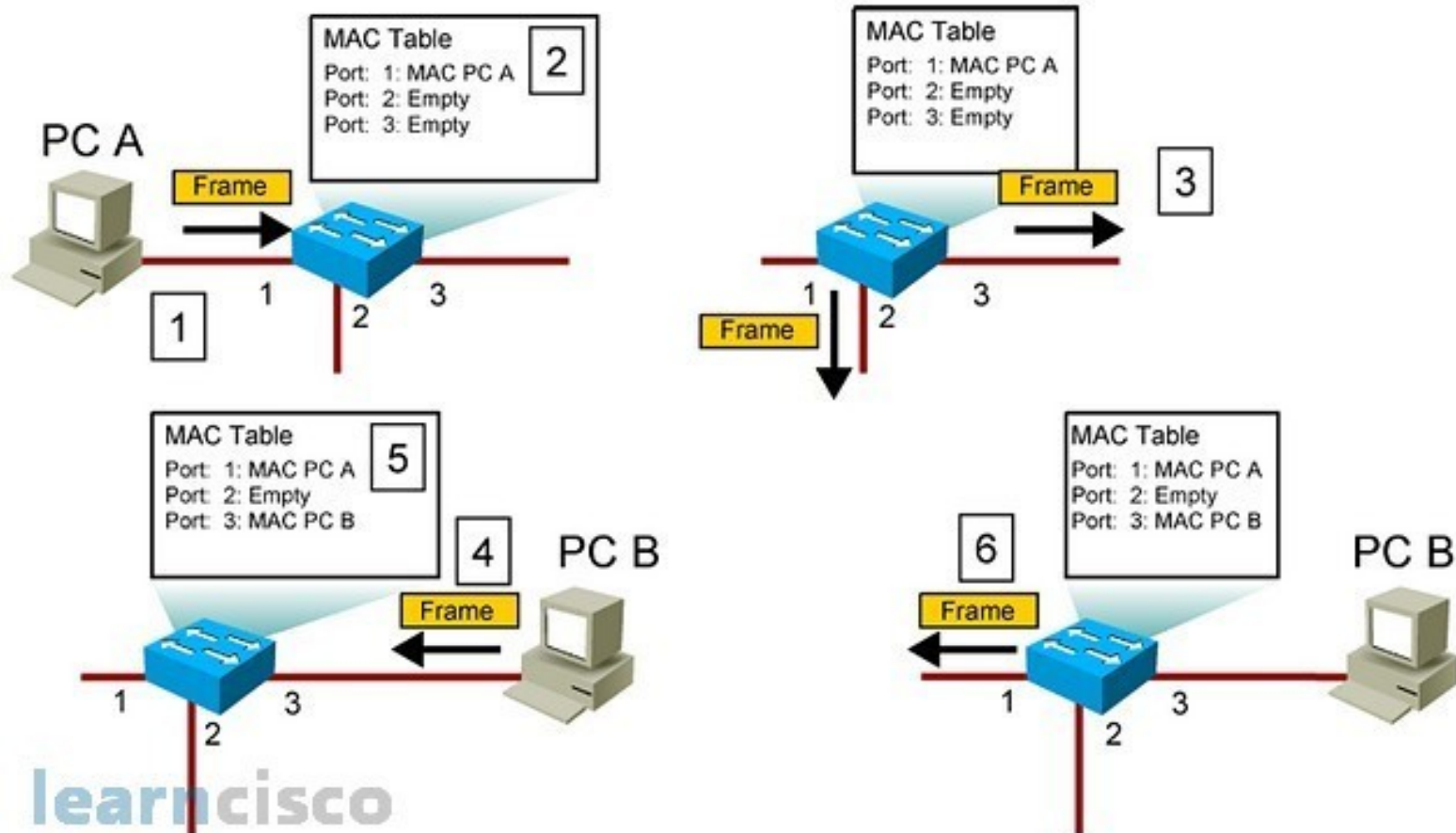
}

else **flood** all other interfaces



# Hubs, bridges, switches and routers

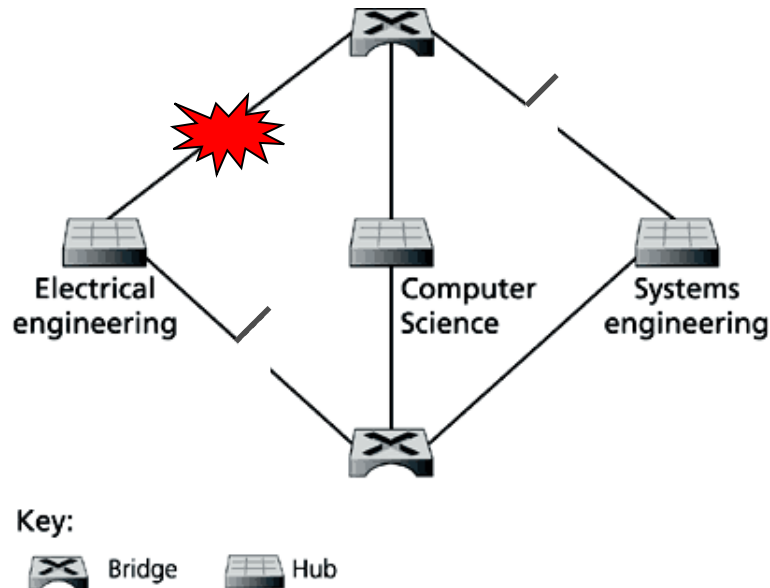
## Bridge/switch at work



Source: learncisco.net, modified Nov. 21, 2016

# Hubs, bridges, switches and routers

## Spanning Tree



- For increased reliability, desirable to have redundant, alternative paths from source to destination
- Issue: with multiple paths, cycles result
- Bridges may multiply and forward frame forever
- Solution: organise bridges in a spanning tree (subset of original topology without loops) by disabling subset of interfaces

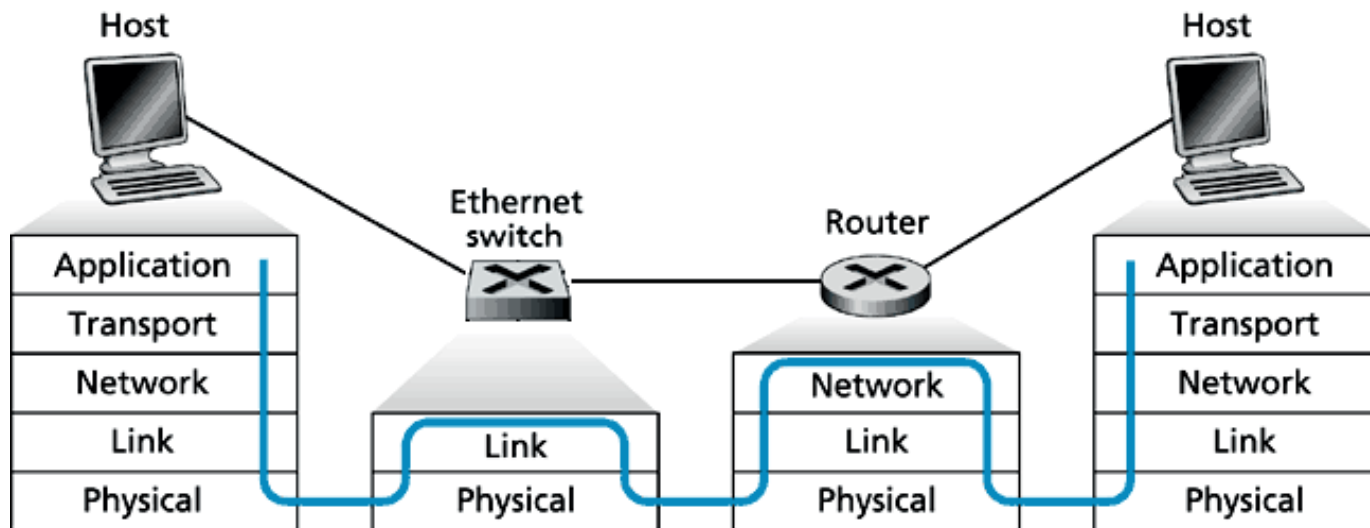
# Hubs, bridges, switches and routers

## L1 (Hub) vs. L2 (switch) interconnection

- Switches isolate collision domains resulting in higher total MAX throughput
- “Plug-and-play” (transparent): no configuration necessary
- Can connect different Ethernet types
- Limitless number of nodes and geographical coverage
- Could span the entire globe, but with HUGE ARP tables!

# Hubs, bridges, switches and routers

## Switches vs. routers



Switches	Routers
Link-layer devices (L2)	Network-layer devices (L3)
Maintain switch tables Implement filtering, learning and spanning tree algorithms	Maintain forwarding tables Implement routing algorithms



# Hubs, bridges, switches and routers

## Switches pro's and con's

Switches are self learning (“plug-and-play”)

Operation is simpler, requiring less packet processing (only up to layer 2)

All traffic confined to spanning tree, even when alternative path is available

No protection against broadcast storms

# Hubs, bridges, switches and routers

## Routers pro's and con's

Arbitrary topologies can be supported

Efficient: no spanning tree, best path can be selected

Robust

- Provide protection against broadcast storms
- Cycling is limited by TTL (and routing protocols)

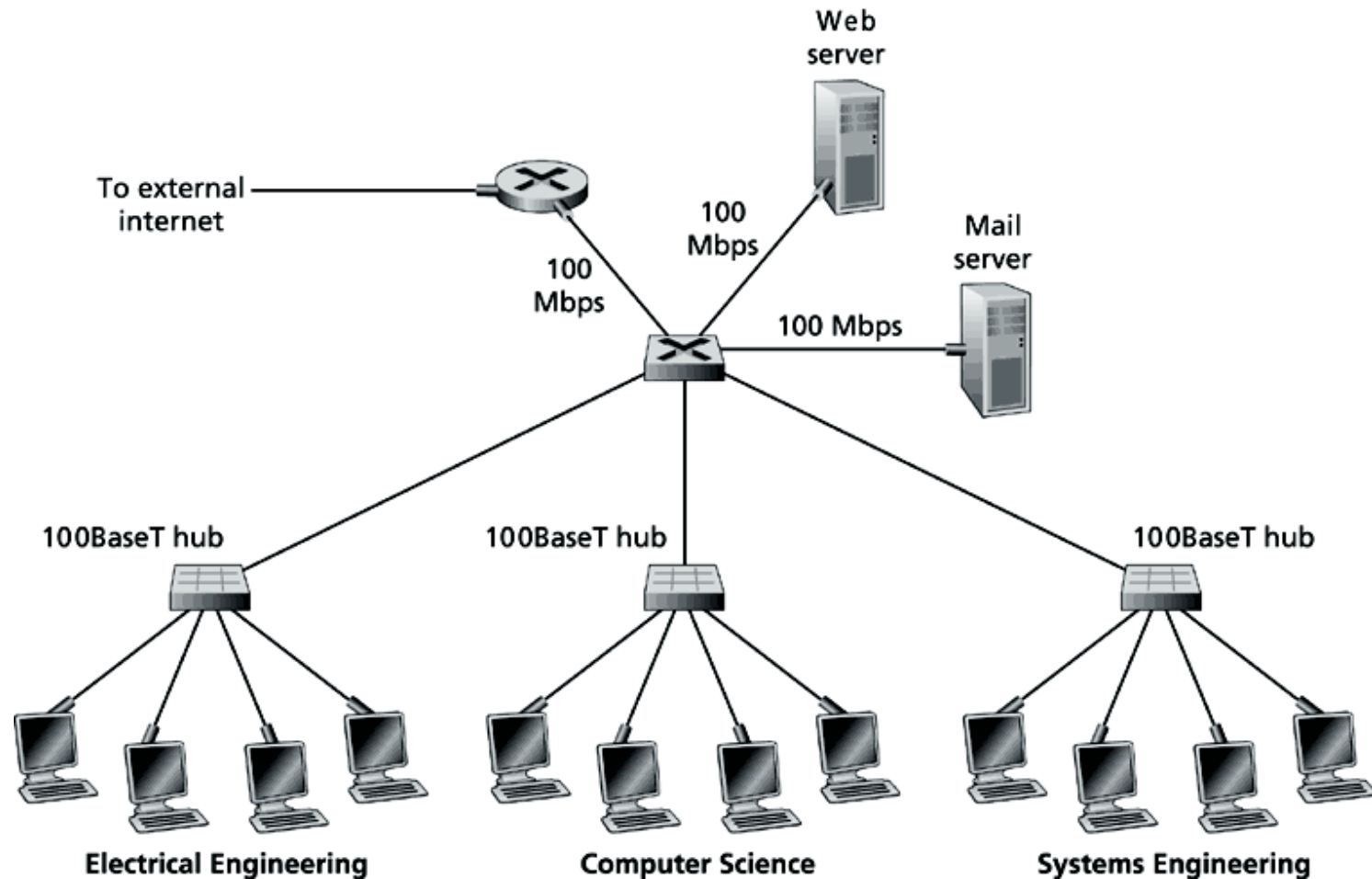
Require higher packet processing

Require address configuration (not “plug-and-play”)

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

# Hubs, bridges, switches and routers

## Typical LAN



# Hubs, bridges, switches and routers

## Comparison

		Traffic isolation	Plug-and-play	Optimal routing	Cut-through
Router	L3	Yes	No	Yes	No
Switch	L2	Yes	Yes	No	Yes
Bridge		Yes	Yes	No	No
Hub	L1	No	Yes	No	Yes

# Summary

- Principles behind data link layer services
  - Error detection, correction
  - Sharing a broadcast channel – Multiple access
    - Partitionning
    - Random Access : ALOHA, CSMA-CD/CA
    - Link layer addressing – IPv6 Neighbour Discovery, ARP
- Link layer technologies: Ethernet, hubs, bridges, switches, IEEE 802.11 LANs
- Journey down the protocol stack now over!

# Review questions

- Suppose two nodes start to transmit at the same time a packet of length  $L$  over a broadcast channel of rate  $R$ . Denote the propagation delay between the two nodes as  $t_{prop}$ . Will there be a collision if  $t_{prop} < L / R$ ? Why or why not?
- Would token-ring be efficient if LAN had a very large perimeter? Why?
- How big are the LAN, IPv4 and IPv6 address spaces?
- Suppose a 100 Mbps adapter sends into a channel an infinite stream of 1's using Manchester encoding. How many transitions per second has the signal emerging from the adapter?