Local Area Networks: Ethernet

Dots. Toomas Tommingas
Chair of Real-Time Systems (RAS)

Objectives

- Understand shared medium access methods of Ethernet;
- Describe network aspects for an Ethernet network;

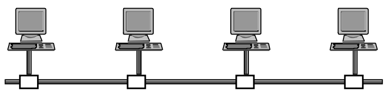
PART A: The CSMA/CD Access Method

PART B: Network Aspects

Part A: The CSMA/CD Access Method

Motivation for LANs

- goal: connect computers in same site (building, small campus)
- experience from host centric networks: bursty traffic
- basic idea: share a cable, no complex software in end system



- altenatives?
  - switch based LANs: connection oriented: ATM
  - switch based LANs: connectionless. Switched Ethernet

Access Method

- multiaccess communication = share a communication medium
  - radio channel, cellular networks, satellite links
  - machine bus
  - local area cable
- shared medium requires an **Access Method**
  - deterministic:
    - **Time Division Multiple Access (TDMA)**
    - **Token Passing (Token Ring, Token Bus, FDDI)**
  - non-deterministic
    - **Aloha**
    - **CSMA/CD**

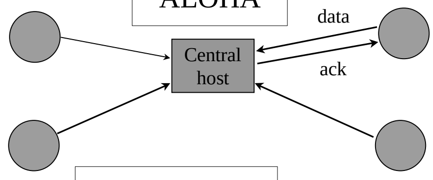
Access Method Topology

- Two topologies are used
  - **bus**:
    - all bits sent by one station are propagated to all stations
    - data die at end of bus
    - all stations see all frames
    - used by Ethernet, Token Bus
  - **ring:**
    - all bits are passed from one station to next station, then to next's neighbour, etc
    - bits eventually return to originating station which has to remove them
    - all stations see all frames
    - used by Token Ring and FDDI

Access Method Topology (cont.)

- **Access Method topology** = physical topology = topology used by bits
  - **logical topology** = topology used by the token in case of Token Bus (logical ring over a physical bus)
  - **cabling topology** = layout of cables = star in most cases (see later)

ALOHA



transmission

```
i = 1
while (i ≤ maxAttempts) do
        send packet
        listen for ack during one
RTD
        if ack received then leave
else resend
        increment I
end do
```


CSMA

- Improvement 1: Listen before you talk:
  "Carrier Sense Multiple Access"

```
i = 1
while (i ≤ maxAttempts) do
        listen until channel idle
        transmit immediately
        wait for acknowledgement or timeout
        if ack received then leave
        wait random time /* collision */
        increment i
end
```

- Improvement 2: Wait random time before retransmission

CSMA

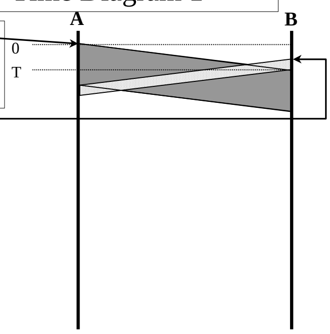
- improvement 3: detect collisions as soon as they occur : "Carrier Sense Multiple Access / Collision Detection"

```
i = 1
while (i ≤ maxAttempts) do
    transmit /* after interframe delay */ and listen
    wait until (end of transmission) or (collision detected)
    if collision detected
            then stop transmitting /* after 32 bits ("jam")*/
            else leave
 wait random time
 increment i
end
```

- improvement 4: acknowledgments replaced by CD
- This is Ethernet (≈ 802.3, the standard conformant version of Ethernet)

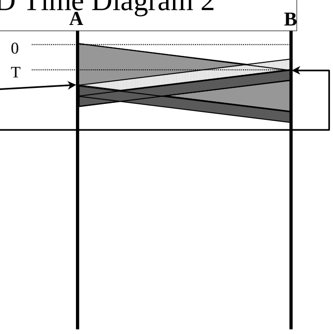
CSMA/CD Time Diagram 1

- A senses idle channel, starts transmitting
- shortly before T, B senses idle channel, starts transmitting



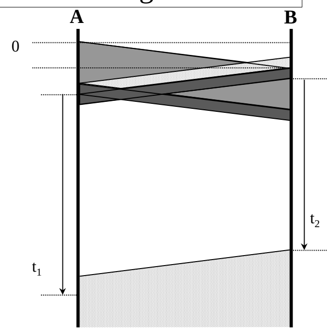
CSMA/CD Time Diagram 2

- A senses collision, continues to transmit 32 bits ("jam")
- B senses collision, continues to transmit 32 bits ("jam")



CSMA/CD Time Diagram 3

- A waits random time t1
- B waits random time t2
- B senses channel idle and transmits
- A senses channel busy and *defers* to B
- A now waits until channel is idle



CSMA/CD

- Stations detect *collisions* and *stop* transmitting.
- Re-attempt to transmit after a random time.
- Acknowledgements ⇨ absence of collision.
- The interframe delay ("gap") is 9.6 µs.
- If repeated collisions occur, then the time before retransmission increases exponentially.

Exponential Backoff

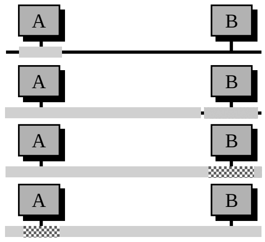
- **random time before re-transmission is given by:**

k = min (10, AttemptNb)
r = random (0, 2^k - 1) * slotTime

- **examples:**
  - ❍ first retransmission attempt:
    k = 1; r = 0 or r = slotTime
  - ❍ second retransmission attempt (if preceding one failed):
    k = 2; r = 0, 1, 2 or 3 * slotTime

Minimum Frame Size

- t = 0: A begins transmission
- t = 1- ε: B begins transmission
- t = 1 : B detects collision, stops transmitting
- t = 2- ε: A detects collision



SlotTime and Minimum Frame Size

- a minimum frame size equal to number of bits transmitted during one round trip is required to detect all collisions
- **`slotTime`** = number of bits transmitted by a source during the maximum round trip time for any ethernet network
  - ❑ slotTime = round trip time + jam time + safety margin
  - ❑ = 512 bits (corresponding to 51.2 µs)
    includes propagation time in repeaters + margin:
    4 repeaters + 5 segments + 2 stations = 2*21.2 µs + 2*1 µs = 44.4 µs
- o **rule**: in Ethernet, all frames must be as large as **`slotTime`**
- o **properties:**
  P1: all collisions are detected by sources
  P2: collided frames are shorter than **`slotTime`**

Proof: P1 see previous slide
P2 because collided frame are aborted by source at the latest after slotTime, including jam bits

CSMA/CD Performance

- Bound on throughput:

$$\theta \geq 1 / (1 + 3.1\ \alpha)$$

where $\alpha = \beta / L$ = 2 * propagation delay / transmission time
with L = frame size, β = bandwidth-delay product

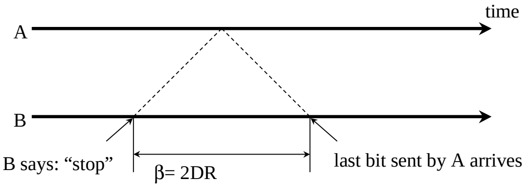
- Approximation:

$$\theta \approx 1 / (1 + 2.5\ \alpha)$$

- o Key for high utilization is: bandwidth delay product << frame size

Bandwidth Delay Product

- Interpretation of bandwidth-delay product



- large β means: delayed feedback

Bandwidth Delay Product

L [m]	v [m/s]	D[s]	R [bit/s]	ß=2*D*R [bit]	[Byte]	[KB]
200	2.3E+08	8.7E-07	10000000	17.39	2.17	0.00
1000	2.3E+08	4.3E-06	10000000	86.96	10.87	0.01
10000	2.3E+08	4.3E-05	10000000	869.57	108.70	0.11
36000000	2.3E+08	0.15652	10000000	3130434.78	391304.35	382.13
200	2.3E+08	8.7E-07	100000	0.17	0.02	0.00
1000	2.3E+08	4.3E-06	100000	0.87	0.11	0.00
10000	2.3E+08	4.3E-05	100000	8.70	1.09	0.00
36000000	2.3E+08	0.15652	100000	31304.35	3913.04	3.82
100	2.3E+08	4.3E-07	100000000	86.96	10.87	0.01
200	2.3E+08	8.7E-07	100000000	173.91	21.74	0.02
1000	2.3E+08	4.3E-06	100000000	869.57	108.70	0.11
10000	2.3E+08	4.3E-05	100000000	8695.65	1086.96	1.06
100	2.3E+08	4.3E-07	1E+09	869.57	108.70	0.11
200	2.3E+08	8.7E-07	1E+09	1739.13	217.39	0.21
1000	2.3E+08	4.3E-06	1E+09	8695.65	1086.96	1.06
10000	2.3E+08	4.3E-05	1E+09	86956.52	10869.57	10.61
36000000	2.3E+08	0.15652	1E+09	##########	#########	38213.32

CSMA/CD Performance

- We have:
  $x1 = R$
  $x2 = 1/\mu$ in average
  $E(x3 \mid \text{collision occured}) = 2R$; Prb (collision occured) = $1 - \exp(-R\mu)$
  $E(x3 \mid \text{successful transmission}) = T$; Prb (successful transmission) = $\exp(-R\mu)$

  The last formula is because collisions can occur only if an arrival occurs during the propagation time R, because of collision avoidance. The average cycle time is thus, for this worst case scenario:

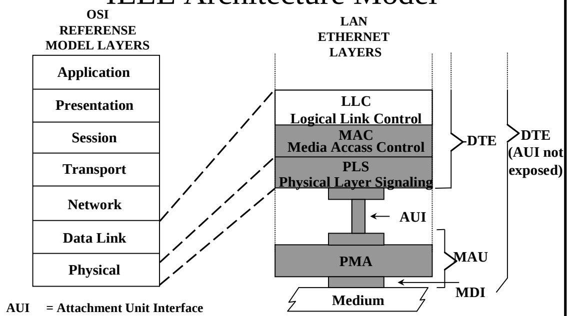
  $\tau = R + 1/\mu + 2R(1- \exp(-R\mu)) + T \exp(-R\mu)$

  and the corresponding utilization:

  θmax = average useful time per cycle / average cycle duration
  $= T \exp(-R\mu) / \tau$

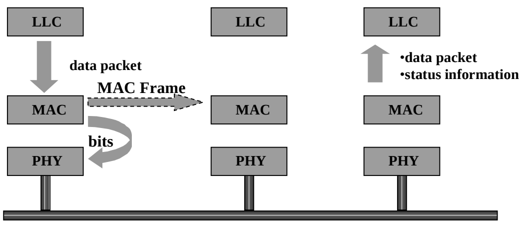
  computing the maximum of θmax with respect to $x = R\mu$ gives the formula (maximum obtained for $x = 0.43$). Note that $\alpha = 2R / T$.
- the approximation shown is based on simulations
- for a large network, β is close to 60 Bytes; for traffic with small frames (L = 64 bytes), the utilization is less than 30 %. For large frames (1500 Bytes), it is around 90%.

IEEE Architecture Model



AUI = Attachment Unit Interface
MAU = Medium Attachment Unit
MDI = Medium Dependent Interface
PMA = Physical Medium Attachment

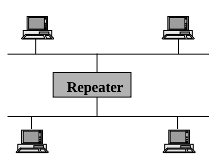
MAC Service



- **data packet = MAC service data unit (SDU)**
- **MAC frame = MAC protocol data unit (PDU)**

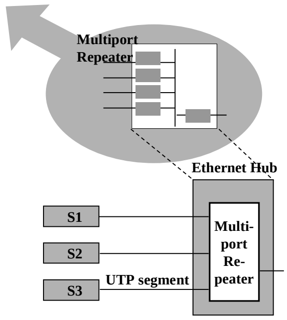
Repeaters

- Extend network beyond cable length limit
- function of a simple (2 port-) repeater:
  - repeat bits received on one port to other port
  - if collision sensed on one port, repeat random bits on other port
- One repeated network = one collision domain
- Even with repeaters, network is limited:
  - propagation time
  - 51.2us slotTime includes repeaters
  - at most 4 repeaters in one path
  - Repeaters perform physical layer functions only (bit repeaters)



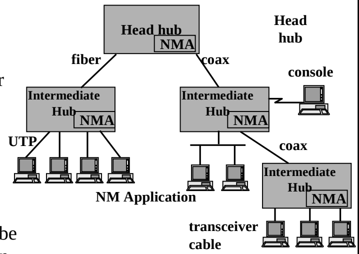
From Repeaters to Hubs

- **Multiport repeater:**
  - (*n* ports) logically equivalent to ***n*** simple repeaters connected to one **internal** Ethernet segment
  - Multi-port repeaters it is possible to use point-to-point segments (**Ethernet in the box**)
  - Value of point to point cabling ?
    - **ease of management**
    - **fault isolation**

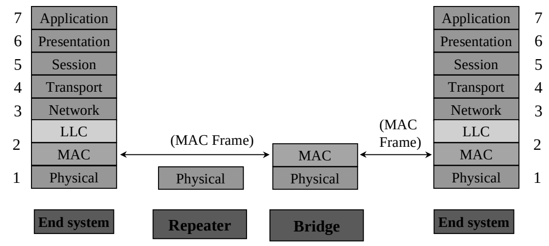


From Bus to Star and Tree

- Ethernet today = active concentrators allow star wiring
- UTP on point-to-point configurations only
- remote network management
- How many frames can be transmitted in parallel in this network?________



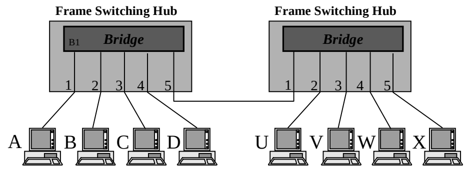
Repeaters and Bridges in OSI



- Bridges operate at layer 2
- Repeaters operate in layer 1
- Layer 3 intermediate systems - routers

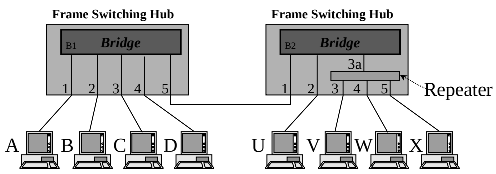
Switched Ethernet

- Switched Ethernet = **Bridge in the Box**
- Total Bandwidth is **not shared**: parallel frame transmission
- An Ethernet Switch is a **connectionless** data switch
- Ethernet used as a **point-to-point** mechanism!

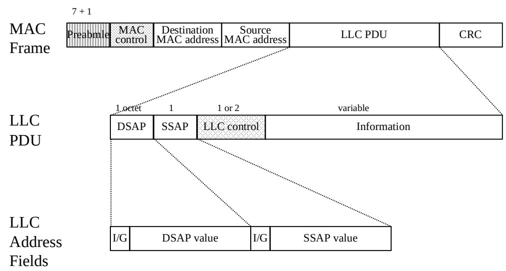


Today's Concentrators

- Concentrators (=hub) combine **bridging** (frame switching) and **port switching** (assign repeater ports to the same **collision domain**)
- NB! Broadcast!



MAC Frame Format

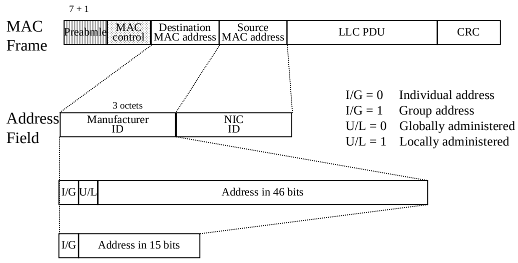


LLC - LSAP

LSAP - LLC Access point, SAP Source and Destination

VALUE	*Assignment*
00	TCP/IP (DoD)
10	NetWare
AA	SNAP (Subnetwork Access Protocol)
E0	NetWare
F0	NetBios
F5	LAN Network Manager Group

MAC Addresses



Token Ring (IEEE 802.5)

Frame

pa	SD	AC	FC	DA	SA	Data	FCS	ED	FS

Token

pa	SD	AC	ED

AC

PPP	T	M	RRR

PPP Priority bits; **T** Token bit
M monitor bit; **RRR** Reservation bits

ED

JK1JK1	I	E

J,**K** non-dada (CPFSK);
I intermiediate frame; **E** error bit

FS

A C r r	A C r r

A address recognition bit;
C Copy bit; **r r** - reserved

FDDI

Frame

pa	SD	FC	DA	SA	Data	FCS	ED	FS

Token

pa	SD	FC	FS

FC

C	L	FF	ZZZZ

C synchr./asynchr.; **L** 16/48 bit addressing
FF (LLC/MAC/..); **ZZZZ** control frame type

ED

JK1JK1	I	E

J,**K** non-dada (CPFSK);
I intermiediate frame; **E** error bit

FS

E A C r	E A C r

E error recognized; **A** address recognized;
C frame copied; **r** - reserved

FDDI MAC protocol

- Key differences compared to 802.5:
  - seize token by aborting the token transmission instead of flipping the T-bit;
  - early token release (already in 16 Mbps version)
- Frame Status (FS) field
  - each station checks any passing frames and sets E bit accordingly
  - MAC protocol does not attempt to retransmit the frame with E bit set. It is the responsibility of LLC or some higher-layer protocol.

FDDI Capacity Allocation

- The priority scheme of 802.5 not applicable due to “early token release”
- FDDI capacity-allocation scheme seeks to accommodate a mixture of stream and bursty traffic
- use of synchronous and asynchronous frames

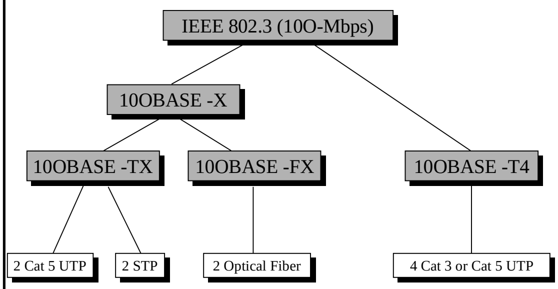
FDDI Capacity Allocation

☞TTRT - Target Token Rotation Time:

$$DMax + FMax + TokenTime + \Sigma\, Sa_i \le TTRT$$

Sa_i	= synchronous allocation for station *i*
DMax	= propagation delay for one complete cirquit of the ring
FMax	= time required to transmit a max length frame (4500 o)
TokenTime	= time required to transmit a token

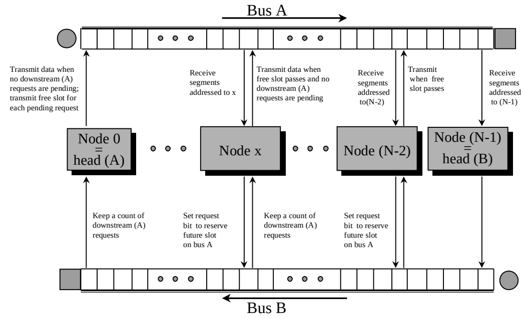
100 BASE-X



IEEE 802.3 Physical Layer

	100 BASE-TX		100 BASE-FX	100 BASE-T4
Transmission medium	2 pair, STP	2 pair, Cat 5 UTP	2 optical fibers	4 pair, Cat 3, 4 or 5 UTP
Signalling technique	4B5B, MLT-3	4B5B, MLT-3	4B5B, NRZI	8B6T, NRZ
Data rate	100 Mbps	100 Mbps	100 Mbps	100 Mbps
Maximum segment length	100 m	100 m	100 m	100 m
Network span	200 m	200 m	200 m	200 m

DQDB - Distributed Queue Dual Bus



DQDB (IEEE 802.6)