INF3190 - Data Communication

Data Link Layer (cntd)

Carsten Griwodz

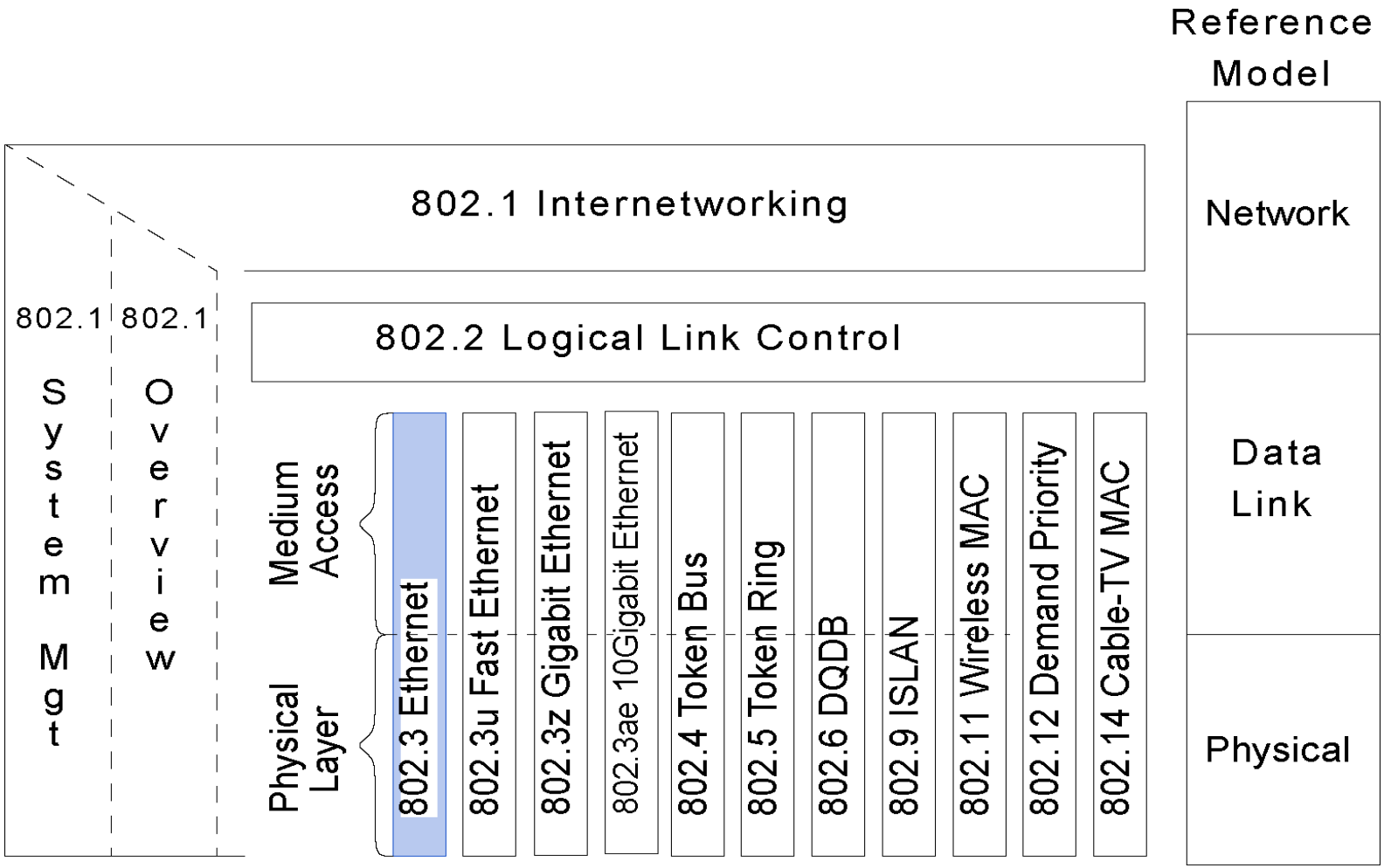
Email: griff@ifi.uio.no

MAC sublayer
Random access protocols

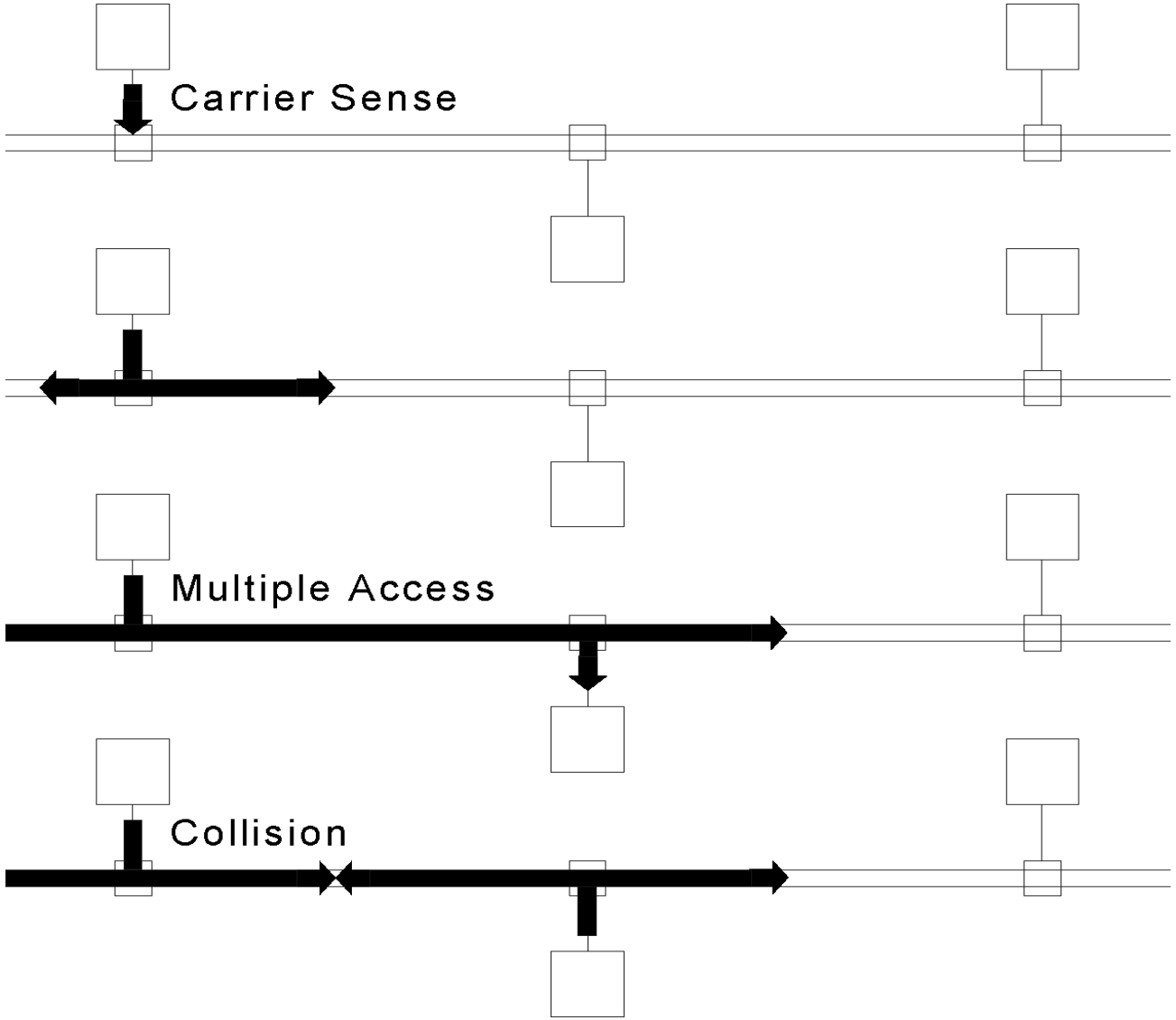
Comparing ALOHA, CSMA.., CSMA CD

		channel is checked (regarding decision to send, not with regard to collision)			behavior in case of desire to send and if one of the following states has been determined			Time slot
		before	during	after	busy	available	collision	
ALOHA	pure			X	sender does not know these conditions		re-transmit after random time interval	
CSMA	nonpersist	X		(X)	re-check channel only after random time interval	sends immediately	wait random time interval then re-check channel and send (if possible) (depending on algorithm "available/ busy")	
	1 persist.	X		(X)	Continuous wait until channel is Available			
	p persist.	X		(X)	initially: continuous wait until chnl/slot available	sends with probability p, waits with probability 1-p (for next slot, then re-checks status)		X
CSMA/CD		X	X		depending on procedure, (see above) 1-persistent is e.g. Ethernet		Terminates sending immediately, waits random time	

802.3: Protocol Family



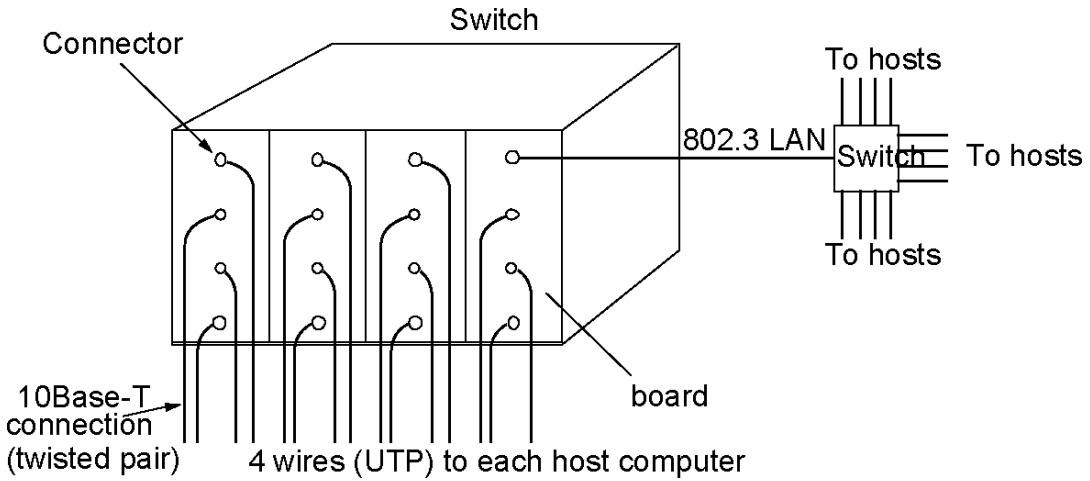
IEEE 802.3: CSMA / CD



Switched 802.3 LANs

Increasing the throughput of 802.3 versions

Switch as relaying center



- station sends frame
- switch tries to locate receiver
  - remember (cache) port of stations that have been **senders** before
  - if unknown, send to all

Collision domain

- the stations that can affect each other through collisions
  - when receiver is known: senders addressing same receiver at same time
  - when receiver is unknown: all stations

802.3: Properties

+ most widely spread
+ stations connect without shutting down the network
+ practically no waiting period during low workload

- analog components for collision recognition
- minimum frame size (64 bytes)
- not deterministic (no maximum waiting period)
- no prioritizing
- when load increases, collisions also increase

- → poor throughput at high load

Ethernet variants

Standardizing Ethernet

802.2 Logical Link Control

802.3 Contention Bus Standard 10base 5 (Thick Net)

- ~~802.3a Contention Bus Standard 10base 2 (Thin Net)~~
- ~~802.3i Twisted-Pair Standard 10base T~~
- ~~802.3j Contention Bus Standard for Fiber Optics 10base F~~
- 802.3u 100-Mb/s Contention Bus Standard 100base T
- 802.3x Full-Duplex Ethernet
- 802.3z Gigabit Ethernet
- 802.3ab Gigabit Ethernet over Category 5 UTP
- 802.3ae 10 Gigabit Ethernet over fiber
- 802.3av 10 Gigabit Ethernet over Passive Optical Network (EPON)
- 802.3bm 100G/40G Ethernet for optical fiber
- ...

IEEE 802.3u: Fast Ethernet

- History
  - High-Speed LAN compatible with existing Ethernet
  - 1992:
    - IEEE sets objective to improve existing systems
  - 1995:
    - 802.3u passed as an addendum to 802.3
    - (alternative solution containing new technology in 802.12)

- Principle
  - retain all procedures, format, protocols
  - bit duration
    - reduced from100 ns to 10 ns

- Properties: CSMA/CD at 100 Mbps
  - cost efficient extension of 802.3
  - very limited network extension
    - sender has to be able to recognize collision during simultaneous sending
      - network extension must not exceed the size of the min. frame
      - frame at least 64 byte, i.e. 5 ms at 100 Mbps per bit
    - i.e. extension only a few 100 meters "collision domain diameter" = 412 m
      - (instead of 3000m)
  - many collisions (lower utilization)

IEEE 802.3u: Fast Ethernet

- Basics
  - actually 10Base-T (Unshielded Twisted Pair)
  - ***Hub*** on L2

- Medium

Name	Cable	Max. segment	Advantages
100Base-T4	**Twisted pair**	**100m**	**Uses category 3UTP**
100Base-TX	**Twisted pair**	**100m**	**Full duplex at 100Mbps (5UTP)**
100Base-F	**Fiber optics**	**2000m**	**Full duplex at 100Mbps**

- 100Base-F (fiber optics):
  - maximum segment length of 2000 m too long for collision recognition
    → may be used only in context with buffered hub ports
    - collisions not possible

- usually improved procedure required
  - for 100 Mbps and more
  - to transmit data in real time

IEEE 802.3z: Gigabit Ethernet

Desirable principle

- if 100% compatible
  - retain all procedures, formats, protocols
  - bit duration reduced from 100 ns over 10 ns to 1 ns
- but, then
  - maximum extension would also be
    - 1/100 of the 10 Mbit/s Ethernet,
  - i. e. (depending on the type of cable) approx. 30 m

IEEE 802.3z: Gigabit Ethernet

Principle for

point-to-point links

- full duplex mode
- interconnected by switch function
- with 1 Gbps in both directions
- no change of packet size

→ i.e. no need for further details

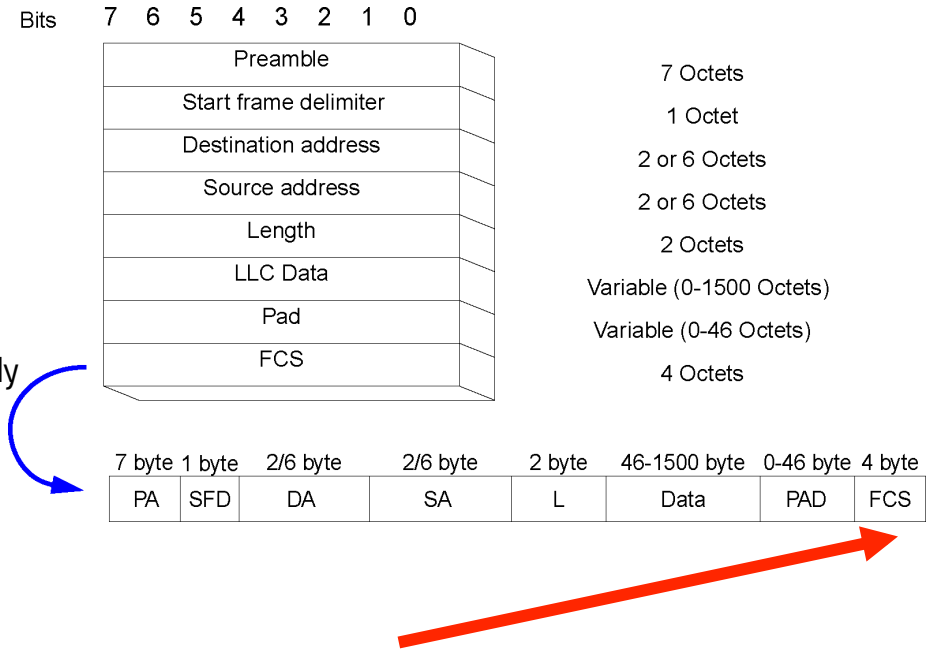
shared broadcast mode

- half duplex mode
- CSMA/CD
- interconnected by hub function
- tradeoff between distance and efficiency

→ i.e. see the following details

IEEE 802.3z: Gigabit Ethernet: Shared Broadcast Mode

- Principle:
  - maintain (as far as possible)
    - CSMA-CD with 64 byte minimum length
  - introducing two features
    - carrier extension
    - frame bursting



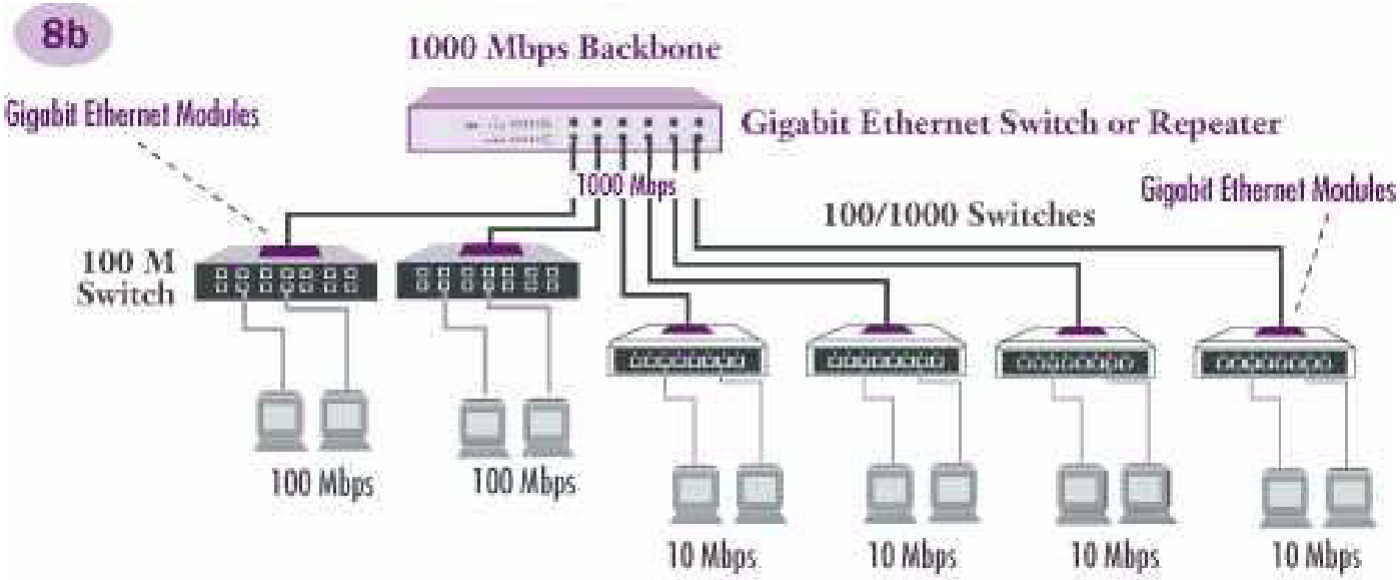
- Carrier extension
  - from 512 bit (64 byte) length, previously
  - to 512 byte length
  - i. e. by attaching a new extension field
    - following the FCS field (Frame Check Sum)
    - to achieve the length of 512 byte
  - Doing:
    - added by sending hardware and
    - removed by receiving hardware
    - software doesn't notice this
  - low efficiency
    - transmit 46 byte user data using 512 byte: 9%

- Frame bursting
  - allow sender to transmit CONCATENATED SEQUENCE OF MULTIPLE FRAMES in single transmission
    - needs frames waiting for transmission
    - better efficiency

IEEE 802.3z: Gigabit Ethernet: Shared Broadcast Mode

Maximum extension of a segment (i.e. of a Collision Domain)

- 5 UTP 100 m
- coax 25 m
- multimode fiber 550 m
- single mode fiber 5 km



IEEE 802.3ae: 10Gbit Ethernet

History
- 1999: IEEE 802.3ae task force founded
- 2002: approval as a standard

Objectives
- to preserve 802.3 frame format
  - incl. minimal and maximal frame sizes
- to support full duplex operation only
  → no CSMA/CD required

Type of media used
- works over optical fiber only, no UTP or coax

Supported distances:
- 850nm: 300 m
- 1310nm: 10 km
- 1550nm: 40 km

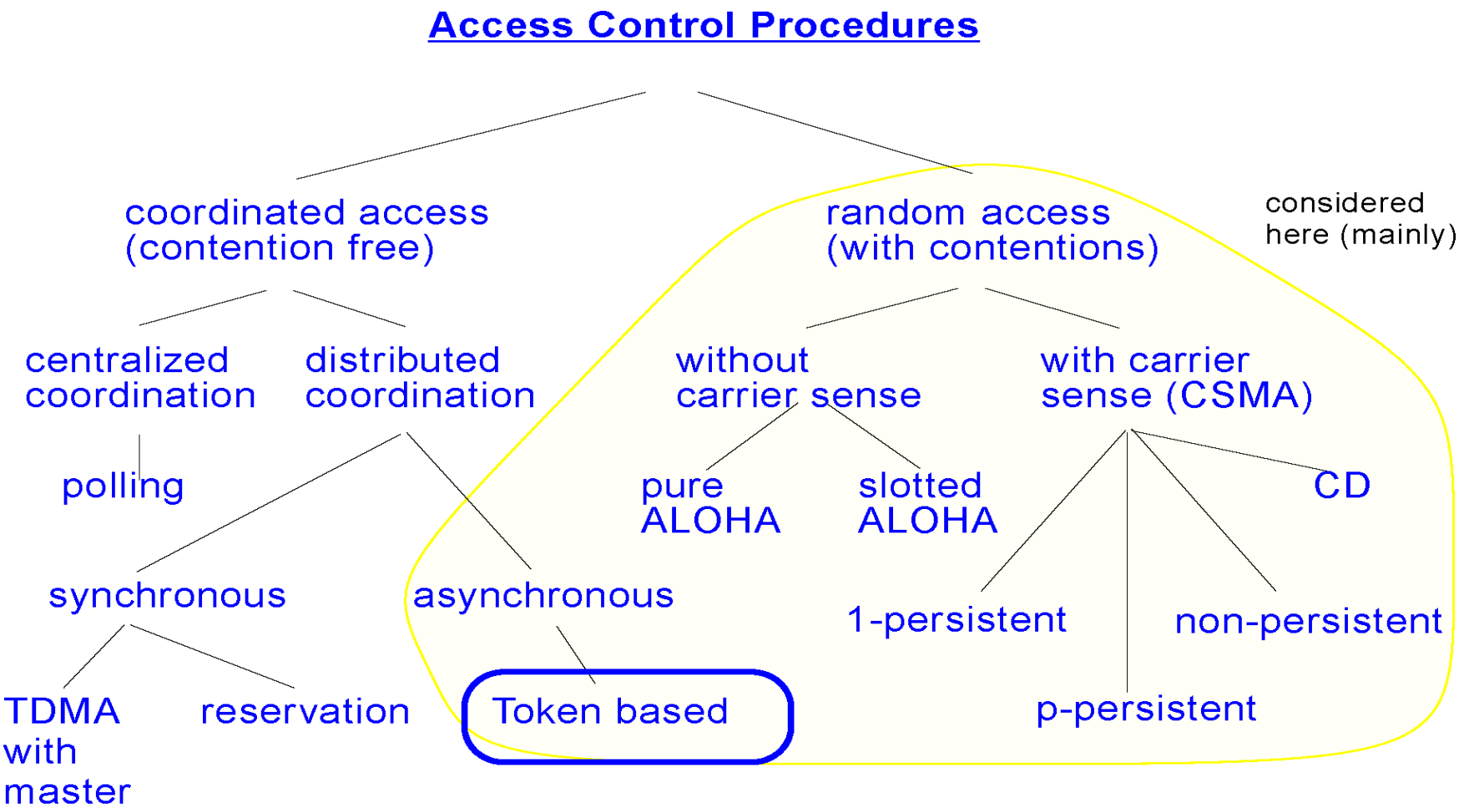
IEEE 802.3ba: 40Gb/s and 100Gb/s Ethernet

Requirements

- To support full-duplex operation only
- To preserve the 802.3 frame format utilizing the 802.3 MAC
- To preserve minimum and maximum FrameSize of current 802.3 standard
- To support a bit error ratio (BER) better than or equal to 10^{-12} at the MAC service interface

MAC sublayer
Token Ring

IEEE 802.5: Token Ring



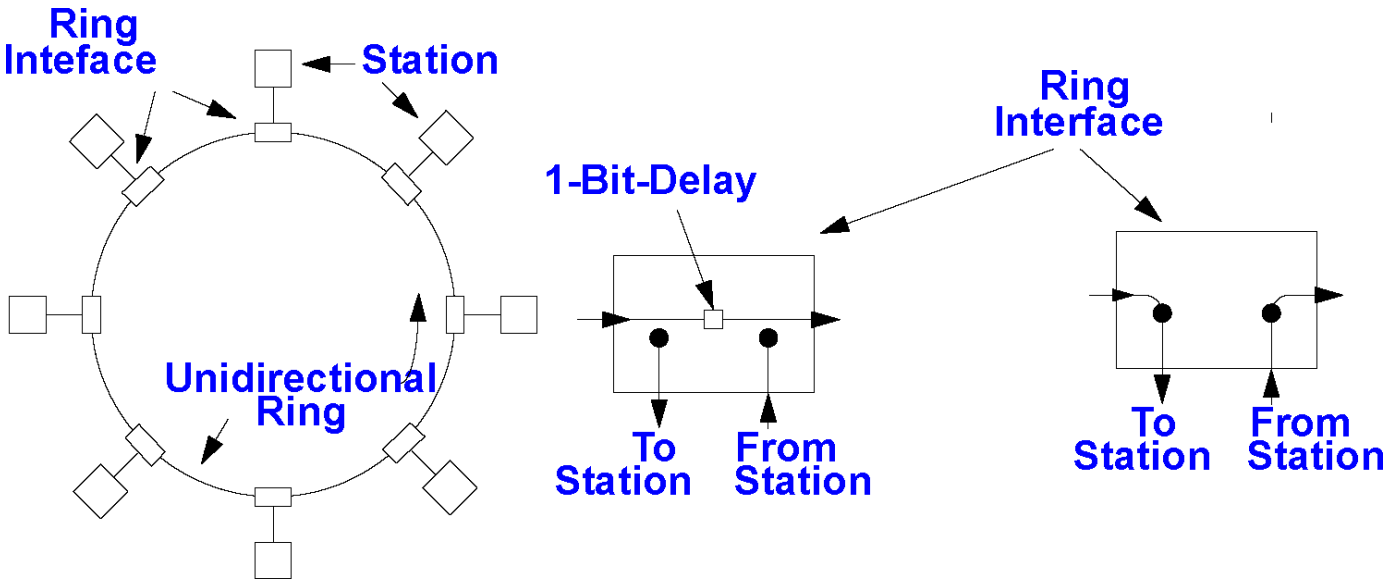
802.5: Ring Topology

Ring

- not really a broadcast medium, but
  - a multitude of point-to-point lines

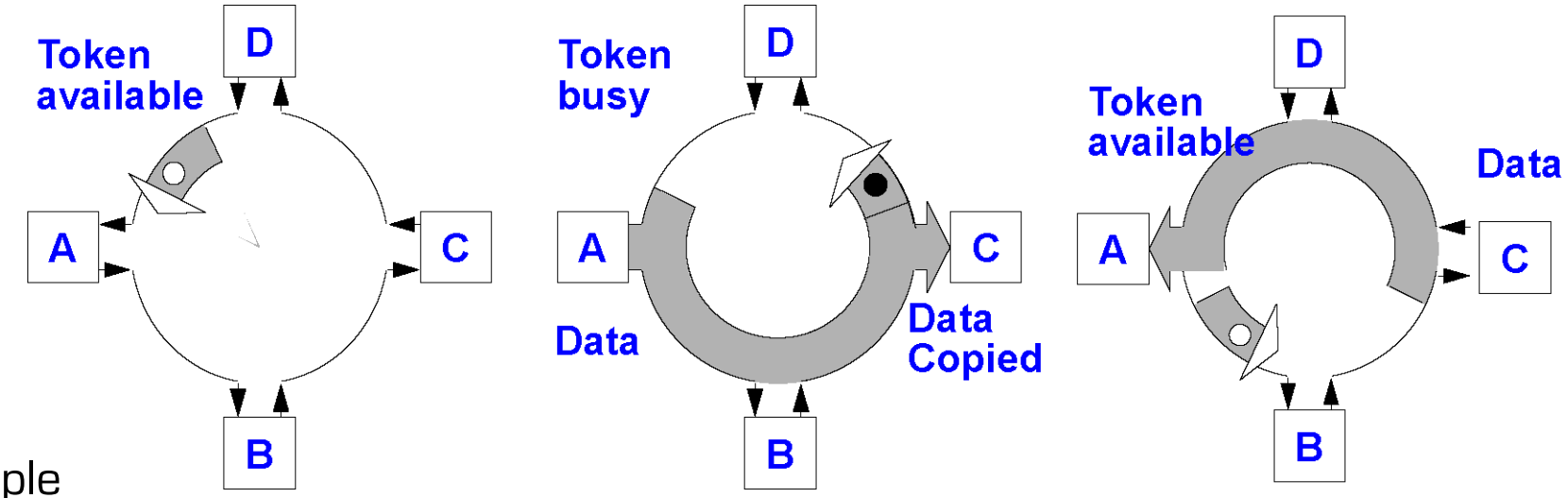
Station

- copies information bit by bit from one line to the next (active station)



802.5: MAC Protocol

Token Protocol



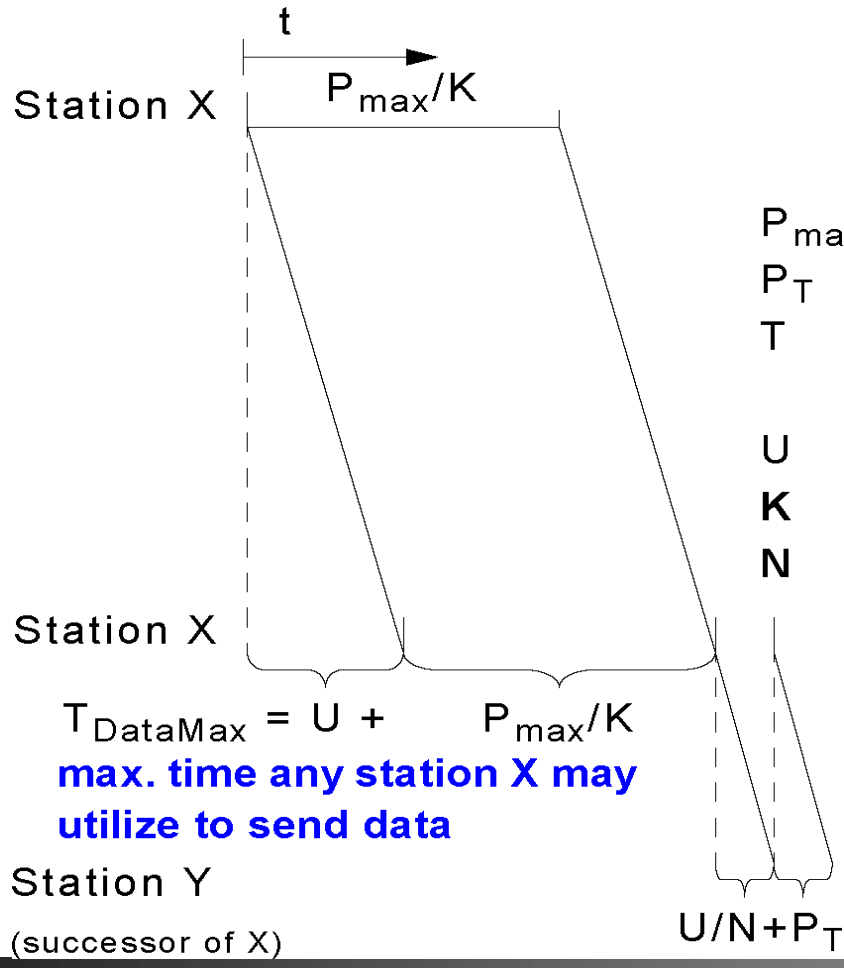
Principle

- Token
  - frame with special bit pattern
- one token circulates on the ring
  - 1: before station is permitted to send
    - it must own and remove the token from the ring
  - 2: station may keep the token for a pre-defined time and may send several frames
  - 3: after receiving its own data back completely
    - the station generates a new token

802.5: Maximum Waiting Period

What is the maximum waiting period for a station before it receives permission to send again?

- i.e. all stations want to send with the max. amount of allowed time



P_{max} = max. packet length [bit]
P_T = token length [bit]
T = transport time for token to adjacent node
U = ring circulation time
K = transmission capacity
N = number of stations

802.5: Maximum Waiting Period

What is the maximum waiting period for a station before it receives permission to send again?

W = maximum waiting period:

W = all others are sending + token rotates x-times

$$= (N-1)(P_{max}/K + U) + N(P_T/K + U/N)$$

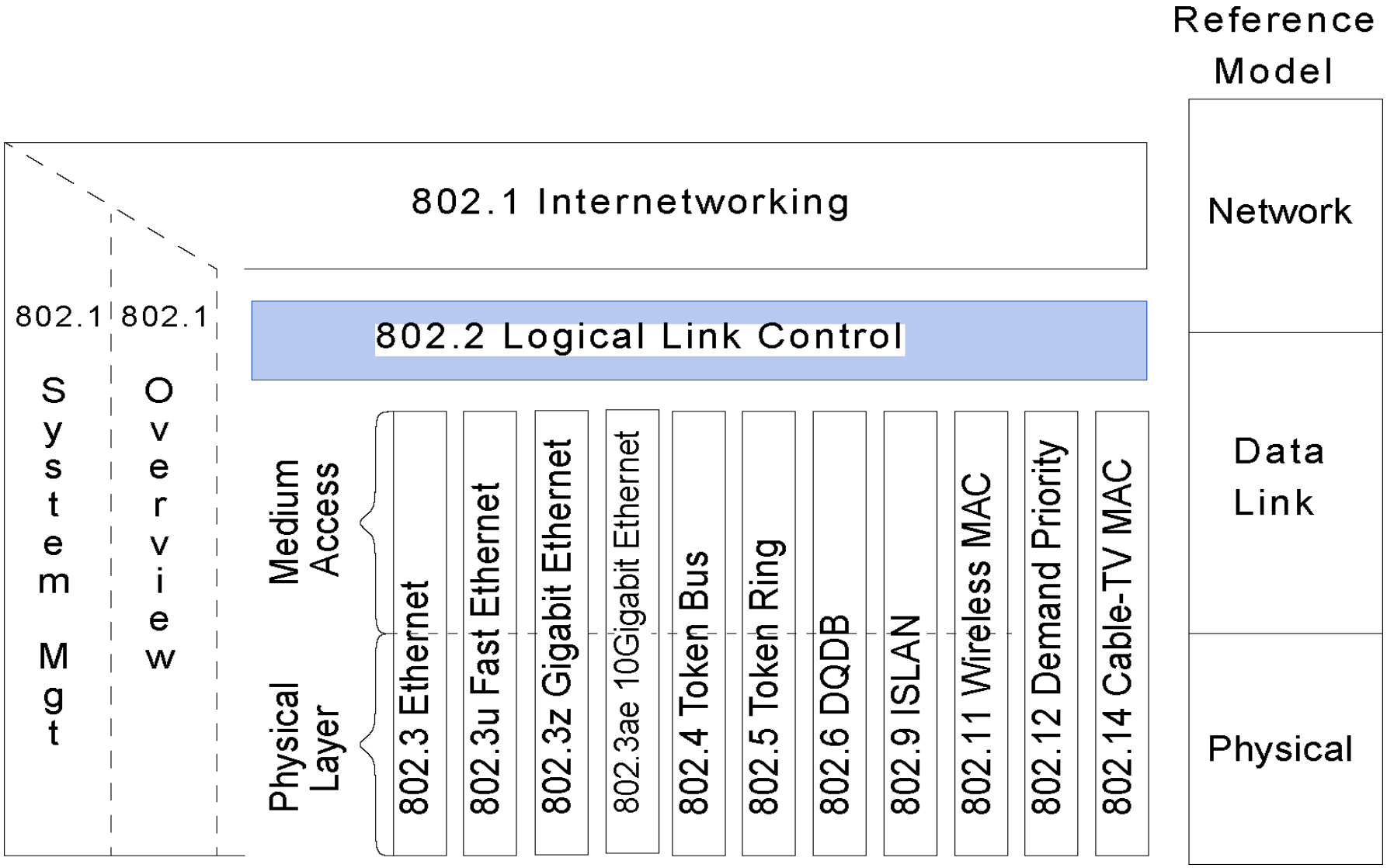
$$= (N-1)(P_{max}/K + U) + NP_T/K + U$$

$$\approx (N-1)(P_{max}/K + U) + U$$

Note: $NP_T/K = 0$ **for** $P_T \ll P_{max}$

LLC sublayer
IEEE 802.2

802.2: Logical Link Control



802.2: Logical Link Control

- Function
  - subset of HDLC
    - High Level Data Link Control HDLC
  - common interface
    - to L3 for all underlying LAN/MAN/WAN components

- Services
  - unacknowledged connectionless (unreliable datagram)
    - upper layers ensure
      - that sequence is maintained, error correction, flow control
  - acknowledged connectionless (acknowledged datagram)
    - each datagram is followed by exactly one acknowledgement
  - connection oriented
    - connect and disconnect
    - data transmission incl. acknowledgement, guaranteed delivery to receiver
    - maintaining the sequence
    - flow control

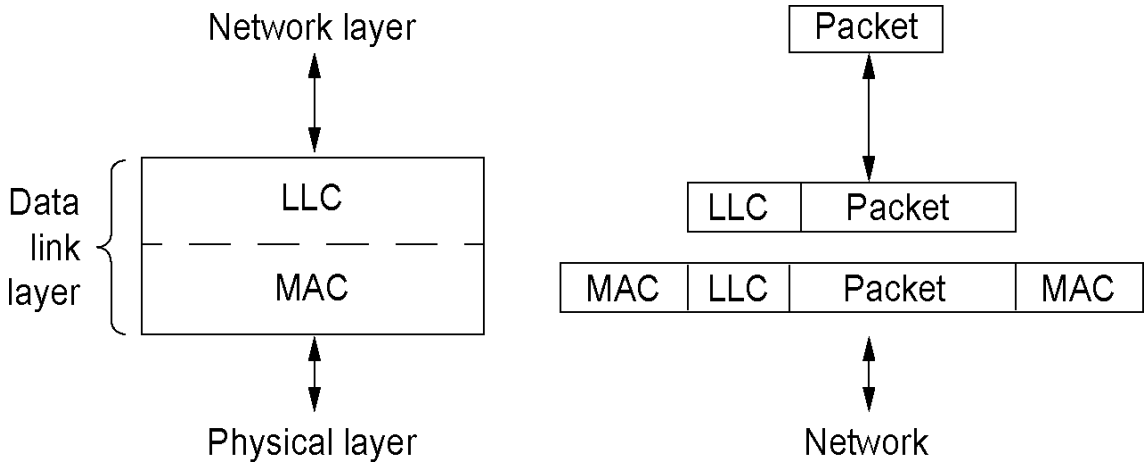
802.2: Logical Link Control

- Function
  - common interface to L3 for all underlying LAN/MAN/WAN components

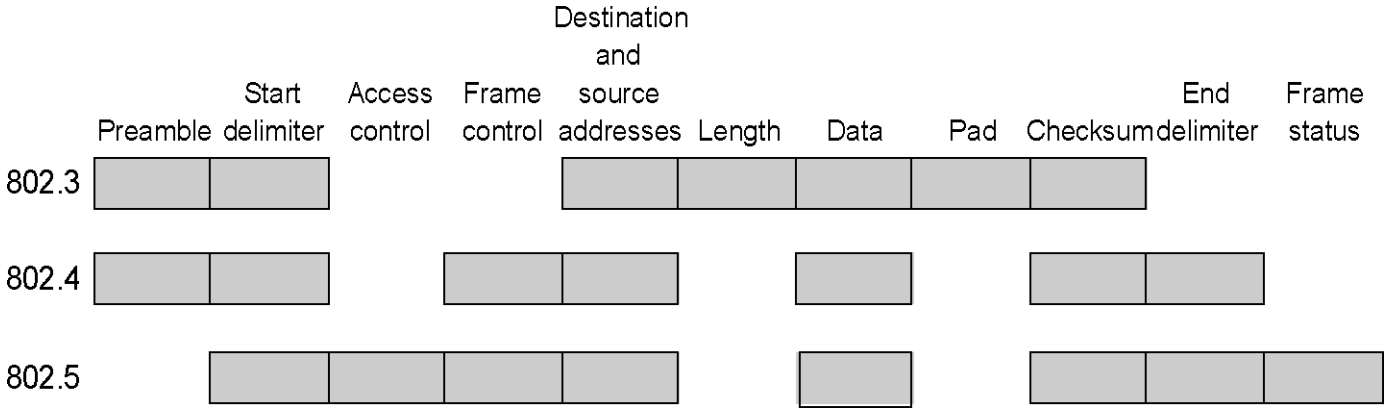
- Services
  - unacknowledged connectionless (unreliable datagram)
    - upper layers ensure
      - that sequence is maintained, error correction, flow control
  - acknowledged connectionless (acknowledged datagram)
    - each datagram is followed by exactly one acknowledgement
  - connection oriented
    - connect and disconnect
    - data transmission incl. acknowledgement, guaranteed delivery to receiver
    - maintaining the sequence
    - flow control

LLC Frame

- Format
  - includes LLC Service Access Points SAPs for source and destination



- Varying AC frames:
  - formats



Flow Control in wired LAN

Ethernet does not have any flow control (usually)

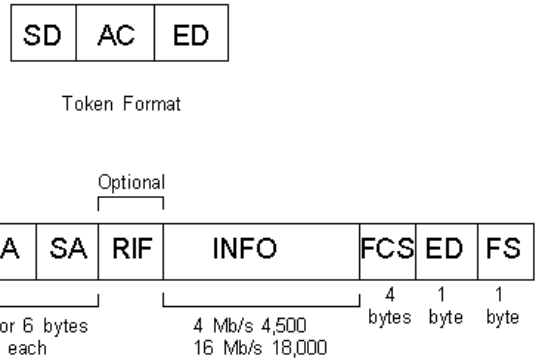
1. usual operation
   - bit error rate on wired Ethernet connections is very low
   - Ethernet provides ordered, but not lossless service to L3
   - therefore, Ethernet does not have to perform retransmissions
   - if a frame arrives with errors, it is discarded
2. PAUSE frame
   - there is a rarely implemented mode that allows a receiver to send PAUSE frames to throttle a sender
   - “priority flow control” (even rarer) can PAUSE only one Type-of-Service
3. data center operations
   - networking in compute clusters should be lossless
   - computer clusters should use GB Ethernet due to cost
   - priority flow control and a few other enhancements make this possible

Flow Control in wired LAN

Token Ring

- speed is no reason for flow control in Token Ring
- if the receiver copies the frame successfully, it confirms reception in the frame itself
- the sender must always yield the token after one frame, whether it has been received or not

- the semantics are weaker than stop-and-wait
  - the FS field is not checksummed
  - Token Ring does never retransmit, this is a higher layer decision



Flow Control in wired LAN

Powerline G.hn (ITU-T G.9961)

- the electrical infrastructure in a household can be quite wild
- no shielding against noise, electrical noise from electrical devices consuming power is the rule rather than the exception
- G.hn supports unicast, multicast and broadcast at the link layer

- unicast and multicast support selective ACK (LLC sublayer)
- a MAC frame contains several LLC subframes (LDPUs) because of the high likelihood of noise
  - 64 - 1500 bytes
  - 16 bit sequence number (65536), window size 1024 (for data), 32 (for control)
  - An ACK can contain an LSSN (lowest segment sequence number) acting as cumulative ACK, and several SSNs the ACK frames after the loss

Flow Control in wired MAN

DOCSIS – Data-Over-Cable Service Interface Specifications

- for Internet over Cable TV
- pretty long distance, very asymmetric bandwidth, strictly hierarchical branching, “channel bonding groups” to increase bandwidth, CMTS (provider’s modem pool) talks to CM (customer’s modem)
- features
  - does not provide lossless service to L3
  - does provide ordered delivery, although one L2 entity uses several L1 channels and packets may be reordered
- packet sequence number
  - 16 bits long
  - plus 1 bit “sequence change count”
  - allows reordering at the receiver
  - the sequence change count is flipped when PSN wraps, so it is actually the 17th bit, ensuring a correct sliding window interpretation

Flow Control in wired WAN

HDLC – High Level Data Link Control

- extremely flexible framing format
- mostly used in WAN connections (SONET/SDH)
  - SONET/SDH transfer multiple digital bit streams synchronously over optical fiber
- sliding window protocol
  - with ACKs
    - RR – cumulative ACK
    - RNR – cumulative ACK but stop transmission
  - NACKs
    - REJ – retransmit 1 frame
    - SREJ – retransmit several sequences of frames
- choice of sequence number spaces
  - sequence number space may be 3 bits (8), 7 bits (128), 31 bits (32 768) or 63 bits (2 147 483 648)
  - window size is negotiable at link establishment for up to N-1 bits

Flow Control in wired WAN

LAPB - Link Access Procedure, Balanced

- Link layer of the (very old) WAN protocol suite X.25
- preceded HDLC

- control frames do not have sequence numbers
- data frames do have sequence numbers
  - 3 bits, 7 bits, or 31 bits
- supports ACKs and NACKs
  - RR, RNR, REJ as above, SREJ (for 31 bits, optional for 7 bits)

Flow Control in wired WAN

L2TP - Layer Two Tunneling Protocol

- Used by ISPs to emulate an L2 service over an authenticated, possibly encrypted long-distance connection
- main motivation: ISP rents part of their network from other ISPs, still want to identify their users and get paid
- provides ordered but unreliable service to L3
- for control information, sliding window is used
  - 16 bit sequence number
  - both go-back-N and selective repeat are explicitly allowed
- for data, sequence number is only used for reordering, not for retransmission

Flow Control in wired LAN

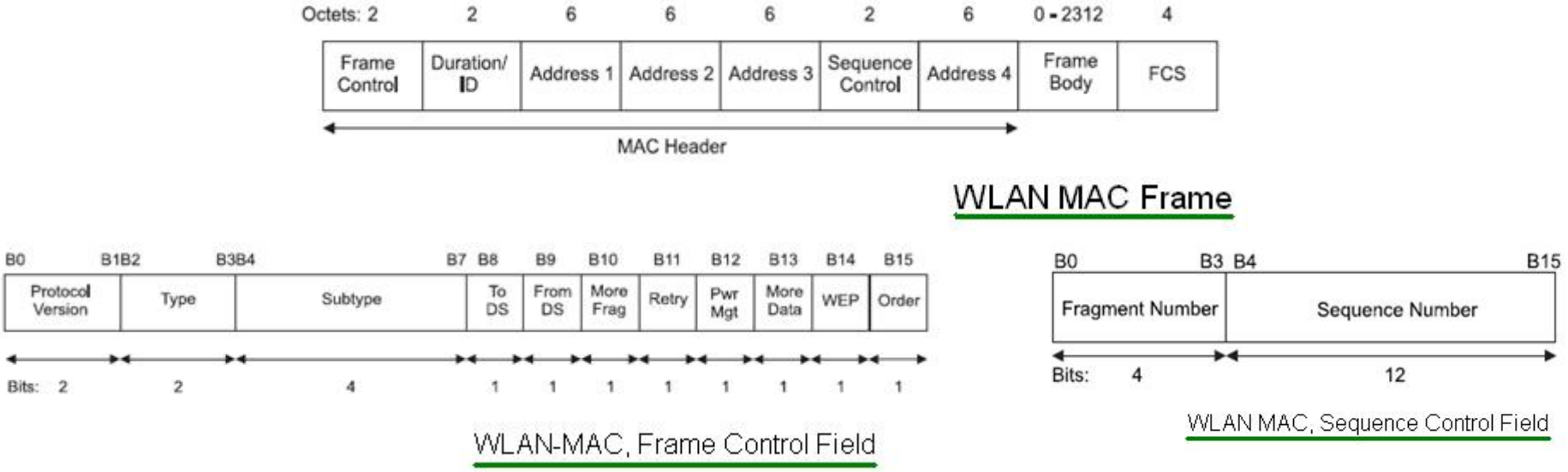
PCI Express

- originally a serial replacement for busses that interconnect components inside a computer
  - data travels between a tree of components
- extended to interconnect computers via non-transparent bridges (NTBs)
  - cluster communication protocol for distances of some meters
- link layer provides reliable, ordered service
  - frames are called TLP (transaction layer packets)
  - flow control credits limit the sending speed per receiver: configurable with maximum 2048 credits, 16 bytes/credit
  - frames have 12 bit sequence numbers
    - note that 2^11=2048 (max credit), so sliding window with selective repeat works
  - support for ACKs and NAKs, timeouts handle a lack of ACKs
  - no cumulative ACKs

Flow Control in wireless LAN

802.11 WiFi

- a real wireless LAN, in the large IEEE 802 family

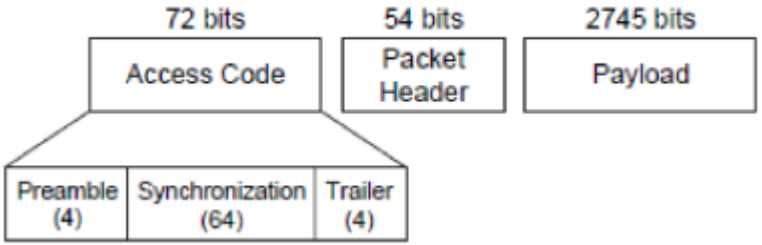


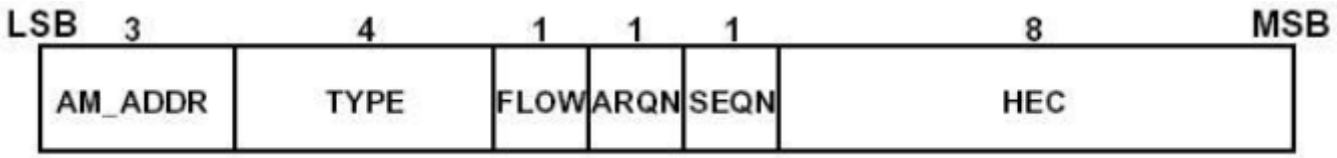
- 12-bit sequence number and 4 bits for fragments
- sequence number space is 4096, ACK for each packet
- retry bit allows sender to indicate that a frame is a retransmit
- this is a classical sliding window with selective repeat

Flow Control in wireless LAN

Bluetooth

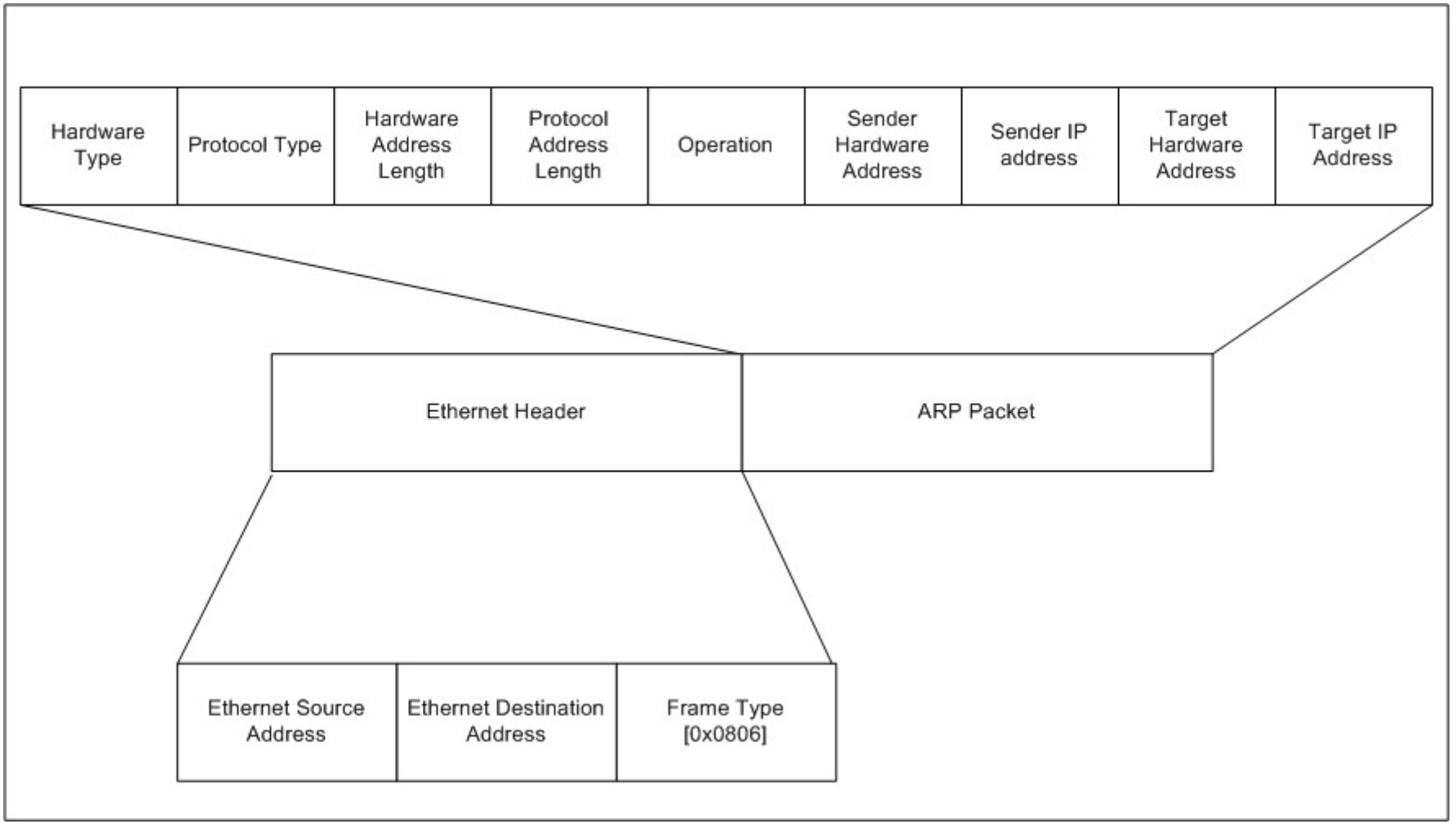
- acts like a wireless replacement of a serial wired line
- used for headphones, keyboard, printers, etc.





- one-bit sequence number (SEQN) and one-bit ACK (ARQN) control indication
- so this is a classical Stop-and-Wait

What is ARP ?


Hardware Type
Protocol Type
Hardware Address Length
Protocol Address Length
Operation
Sender Hardware Address
Sender IP address
Target Hardware Address
Target IP Address
Ethernet Header
ARP Packet
Ethernet Source Address
Ethernet Destination Address
Frame Type [0x0806]