REAL-TIME NETWORKS
Introduction

Prof. J.-D. Decotignie
CSEM Centre Suisse d'Electronique et de Microtechnique SA
Jaquet-Droz 1, 2007 Neuchâtel
jean-dominique.decotignie@csem.ch

Outline

- Definition
- Examples
- Models
- Course content and objective

Definitions

- Response time

Interval between the instant at which one or more inputs are changed and the instant at which the matching response at outputs can be perceived

- Real-time system

« A computer system for which correctness of calculations not only depends on the logical behavior but also on the instant at which the result is produced. If the temporal constraints are not fulfilled, it is said that a failure has occurred », real-time systems FAQ

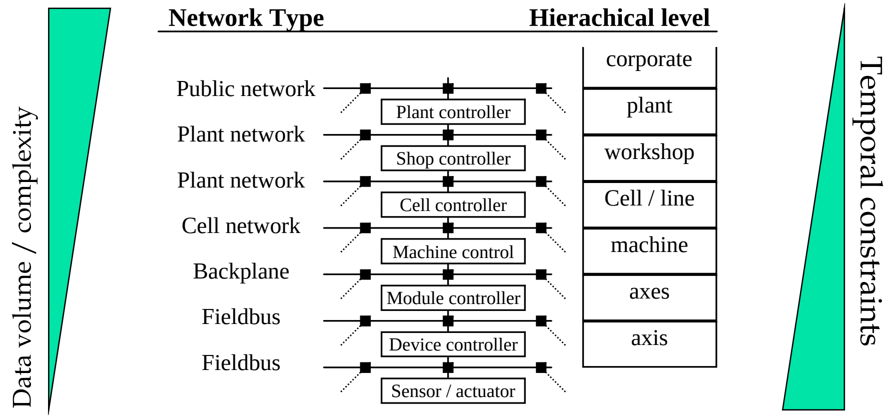
Definitions (2)

« hence, it is essential that the timing constraints of the system are guaranteed to be met. Guaranteeing timing behavior requires that the system be predictable. It is also desirable that the system attain a high degree of utilization while satisfying the timing constraints of the system », Real-time systems FAQ

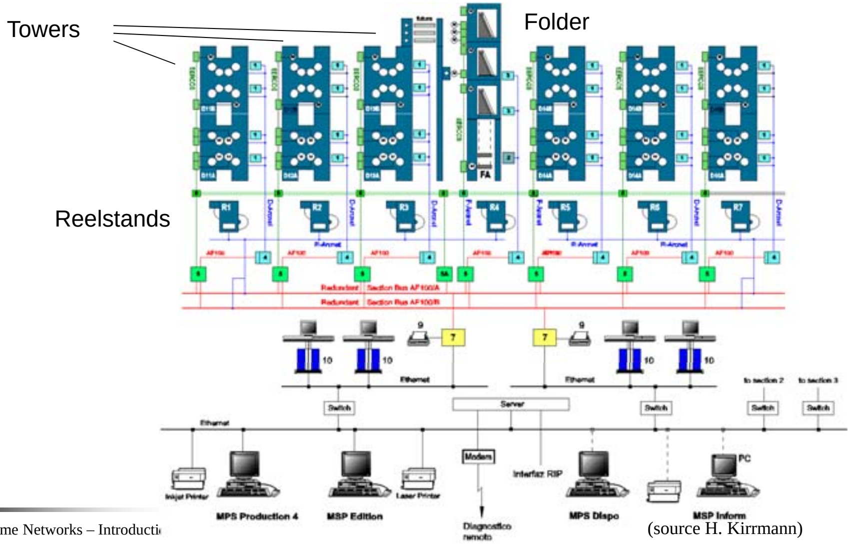
Hard vs soft real-time

- Hard real-time
  - in case of violation of one or more temporal constraints, there is full loss of functionality
- Soft real-time
  - the functionality is not lost but its quality is degraded
  - Still some reward after deadline (value given)
- Firm real-time
  - Same but no value is given if deadline missed

Production hierarchy

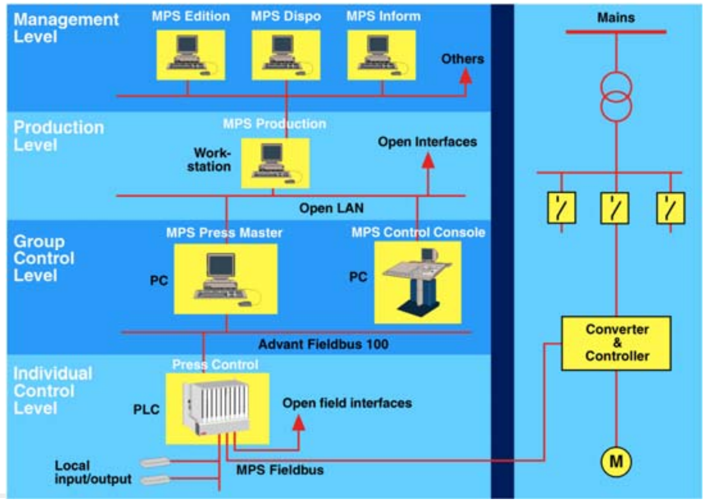


Printing facility



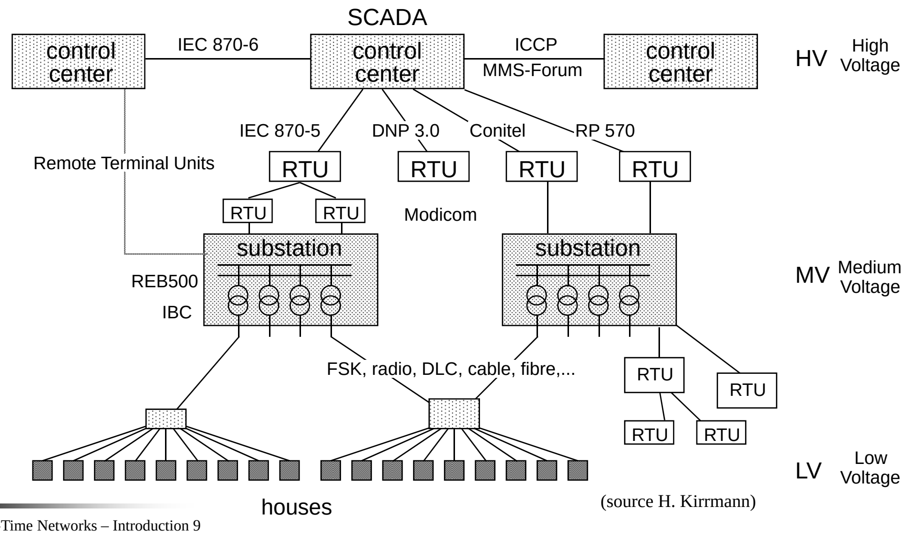
(source H. Kirrmann)

Printing facility architecture



(source H. Kirrmann)

Electricity control network



(source H. Kirrmann)

Data volume in plants

- Power Plant 25 years ago
  - 100 measurement and action variables (called "points")
  - Analog controllers, analog instruments
  - one central "process controller" for data monitoring and protocol.
- Coal-fired power plant today
  - 10'000 points, comprising
    - 8'000 binary and analog measurement points and
    - 2'000 actuation points
    - 1'000 microcontrollers and logic controllers
- Nuclear Power Plant
  - three times more points than in conventional power plants
- Electricity distribution network
  - 100’000 - 10’000’000 points

(source H. Kirrmann)

Distances

- 1-1000 km: Transmission & Distribution.
  - Control/supervision of large distribution nets: water, gas, oil, electricity
- 1-5 km: Power Generation.
  - Out of primary energy sources: waterfalls, coal, gas, oil, nuclear ...
- 50m – 3km: Industrial Plants.
  - Manufacturing and transformation plants: cement works, steel works, food silos, printing , paper, pulp processing, glass plants, harbors, ...
- 500m – 2km: Building Automation.
  - energy, air conditioning, fire, intrusion, repair, ...
- 1m – 1km: Manufacturing.
  - flexible manufacturing cells – robots
- 1m – 800m: Vehicles.
  - Locomotives, trains, streetcars, (trolley) buses, cars, planes, spacecraft, ..

(source H. Kirrmann)

Reaction times

- 10 µs: positionning of offset cylinder in offset printing (0,1 mm at 20 m/s)
- 46 µs: sensor synchro. in bus-bar protection for substations (1º @ 60Hz)
- 100 µs: resolution of clock for a high-speed vehicle (1m at 360 km/h)
- 1.6 ms: sampling rate for protection algorithms in a substation
- 10 ms : resolution of events in the processing industry
- 20 ms: time to close or open a high current breaker
- 200 ms: acceptable reaction to an operator's command (hard-wire feel)
- 1s: acceptable refresh rate for the data on the operator's screen
- 3 s: acceptable set-up time for a new picture on the operator's screen
- 1 min: general query for refreshing the process data base in case of major crash

(source H. Kirrmann)

Application model

- Time-triggered
  - Application work cyclically
    - Wait beginning of period
    - Sample inputs
    - Compute some algorithm (using inputs and parameters)
    - Make the results available at outputs
  - Special cases: acquisition, distribution
  - Periodicity is not always mandated
    - If yes, limited jitter may be required
- Event-triggered

Application model (2)

- Time-triggered
- Event-triggered
  - Activated upon event occurrence
    - Computes some algorithm using system state and parameters
    - Sends the results (local or remote)
  - Need to know the order of occurrence of events
  - Maximum end-to-end delay between event occurrence and answer (result of algorithm)
  - Actions (computations) may need to be synchronized
    - Axis movements for instance (see printing facility)

Data Model

- process data,
- configurations and parameters,
- programs

- Properties
  - Temporal consistency
    - Absolute
    - relative
  - Spatial consistency

Temporal consistency

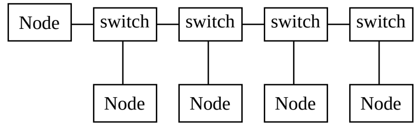
- Let $[a, t_a, v_a]$ and $[b, t_b, v_b]$ be the internal representations of variables a and b where t_a and t_b indicate the instants at which the values v_a and v_b of a and b have been acquired.
- At instant t, v_a is said absolutely consistent if and only if: $t - t_a < A$ (A: absolute consistency threshold for a)
  - freshness
- v_a and v_b are said relatively consistent if and only if: $|t_a - t_b| < R$ (R: relative consistency threshold)

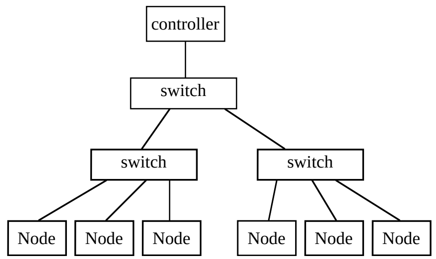
Spatial Consistency

- Applies when copies of the same information is made available on different locii of control (nodes)
- Spatial consistency = copies are identical

Network model

- Single sub-network
- Traffic may come from or go to external world but has no RT constraints
- Free topology





Definitions

- Data circuit (ISO 7498) “a common path in the physical media for OSI among two or more physical-entities together with the facilities necessary in the physical layer for the transmission of bits on it”;
- Real subnetwork (ISO 7498) “a collection of equipments & physical media which form an autonomous whole & can be used to interconnect real systems for the purpose of data transfers”
- subnetwork (ISO 7498) – “an abstraction of a real subnetwork”;
- data link (IEC 8802.2) – “an assembly of two or more terminal installations and the interconnecting communication channel operating according to a particular method that permit information to be exchanged. In this context, the term terminal installation does not include the data source and the data sink”.

Definitions (2)

- LAN – a data link using the same physical layer and medium access control protocols.
- segment – synonymous to data circuit when the nodes are connected through wires.
- Cell – synonymous to data circuit but in the case of wireless medium.

Traffic model

- Real-time periodic
  - $M_{p,i}=\{T_i, D_i. C_i\}$
    - C_i length of message, T_i period of transfer, D_i relative deadline from beginning of period (absolute deadlines are $d_{n,i} = n\ T_i + D_i$)
- Real-time sporadic
  - $M_{s,i}=\{T_i, D_i. C_i\}$
    - T_i min. interarrival time, D_i relative deadline from arrival time (absolute deadlines are $d_{n,i} = arr_{n,i} + D_i$)
- Best effort
  - Configuration data
  - File transfer
- Multimedia

Traffic model (2)

- File transfer
  - $M_{ft,i}=\{T_{i,inner}, T_{i,outer}, n_i, D_i. C_i\}$
    - File transfer flow I happens at most every $T_{i,outer}$. When it occurs, n_i packets (messages) are transferred at a rate given by $T_{i,inner}$
- Multimedia
  - Characterized by uneven arrivals and varying packet lengths
  - Not real-time but we would like to provide more than best effort
  - Solution is often to provide some share of bandwidth
  - $M_{m,i}=\{T_i, C_i\}$
    - C_i bits each T_i
- Best effort traffic may not play the game (tries to send the maximum)

Error model

- Most studies assume “errorless” networking
  - Close to reality for networks on cables
  - Not representative for wireless networks
- Error model depends on type of media
  - Wireless -> interferences (bursts) and fading (long)
  - Wireline -> well modeled by packet error rate with no correlation
- Model
  - Omissions: fail to receive a packet
  - Duplications: receive same packet twice or more
  - Collisions: emissions from two or more nodes at same time
  - Babbling idiot: node starts to send without limits
  - Node failure

Quality of Service (QoS)

- Observable properties of the network
  - Transfer delay bounds
  - Transfer delay variations (jitter)
  - Throughput

Course Objective

- Master the main problems and solutions related to communications under real-time QoS constraints
  - in transportation systems
  - in the control of industrial processes
  - In multimedia
- Deep insight into how
  - to guarantee QoS at the protocol level
  - to assess the QoS properties
- Introduction to research in the domain

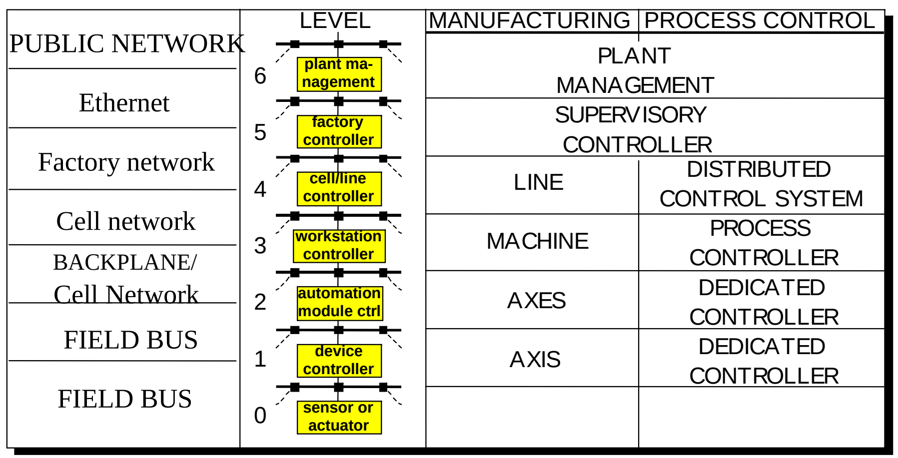
Course Content

- Introduction
  - Requirements (delay, jitter, predictability, topology, cost, etc.)
- Layers and impact on the temporal behaviour
  - Physical layer impact
  - Medium Access Control and Logical Link Control
  - Other layers (network, transport, application, clock synchronization, network management)
- Real-Time performance assessment (scheduling, without error, in presence of errors)
- Fieldbusses and analysis of the main solutions (Profibus, FIP, MVB, CAN, ASi, etc.) and how they fulfill the requirements
- Ethernet and the many ways to offer real-time performances
- Wireless solutions (802.11, Zigbee, Bluetooth, sensor networks)

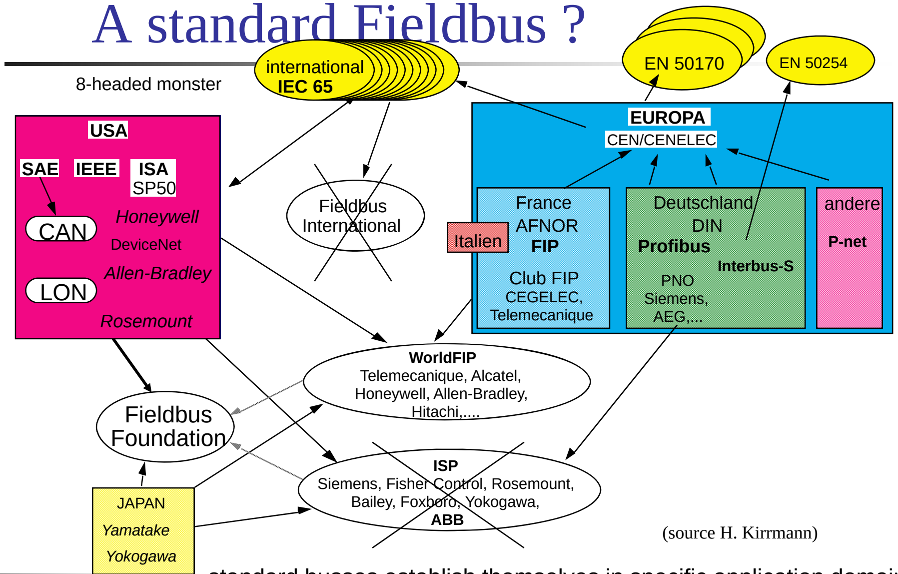
Real-Time Networks

- A-bus
- Arcnet
- Arinc 625
- * ASI
- Batibus
- Bitbus
- * CAN
- ControlNet
- DeviceNet
- DIN V 43322
- DIN 66348 (Meßbus)
- FAIS
- EIB
- Ethernet
- Factor
- * Fieldbus Foundation
- FIP
- Hart
- IEC 61158
- IEEE 1118 (Bitbus)
- Instabus
- * Interbus-S
- ISA SP50
- IsiBus
- IHS
- ISP
- J-1708
- J-1850
- LAC
- * LON
- MAP
- Master FB
- MB90
- MIL 1553
- MODBUS
- MVB
- P13/42
- P14
- Partnerbus
- P-net
- * Profibus-FMS
- Profibus-PA
- Profibus-DP
- PDV
- * SERCOS
- SDS
- Sigma-i
- Sinec H1
- Sinec L1
- Spabus
- Suconet
- VAN
- * WorldFIP
- ZB10
- ...

Communication hierarchy



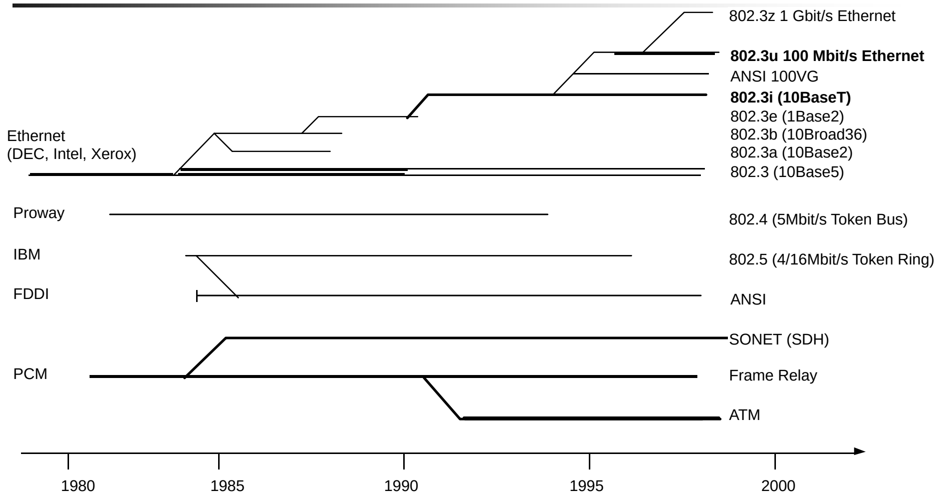
A standard Fieldbus ?



(source H. Kirrmann)

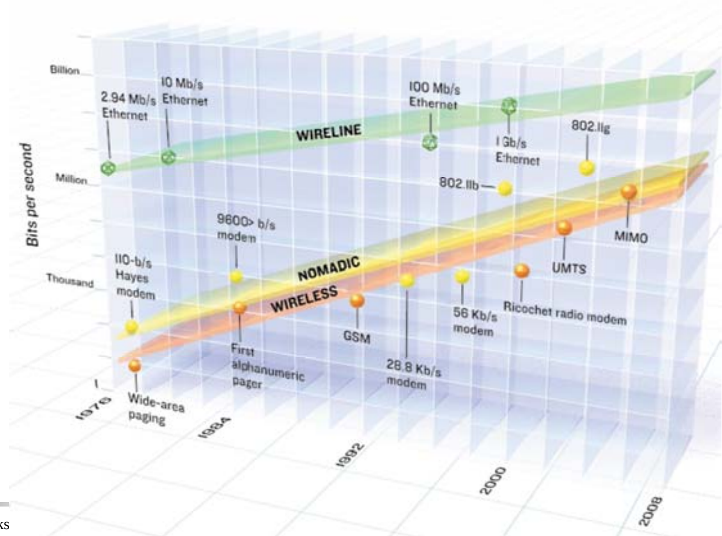
standard busses establish themselves in specific application domains

LAN evolution



(source H. Kirrmann)

Edholm's law [Cherry 04]



General references

- Books
  - B. Forouzan, « Data Communications and networking », McGraw-Hill, 2004, ISBN: 0072923547
  - W. Stallings, « Data and Computer Communications », Prentice Hall, 7th ed., 2003, ISBN: 0131006819
  - D. Comer, R. Droms, « Computer Networks and Internets », Prentice Hall, 4th ed. 2003, ISBN: 0131433512
- Papers
  - J.-D. Decotignie, « Ethernet Based Real-Time and Industrial Communications », Proc. of the IEEE, vol. 93, issue 6, pp. 1102-17.
  - S. Cherry, "Edholm's law of bandwidth", IEEE Spectrum, Vol. 41 , Issue: 7 , July 2004, pp.58 - 60

REAL-TIME NETWORKS
Layers and impact on QoS

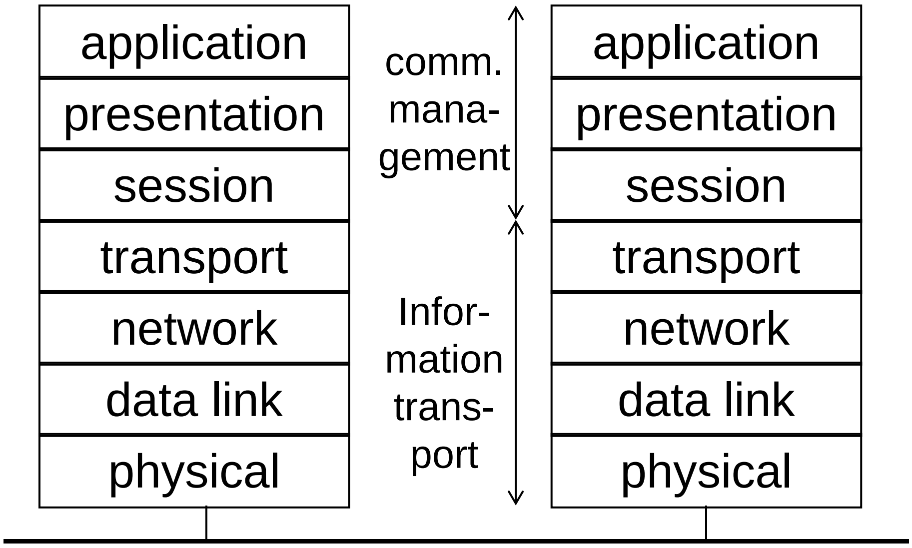
Prof. J.-D. Decotignie
CSEM Centre Suisse d’Electronique et de Microtechnique SA
Jaquet-Droz 1, 2007 Neuchâtel
jean-dominique.decotignie@csem.ch

Outline

- ISO Open Systems Interconnection (OSI) model
- Physical layer impact
- Data link layer impact
  - Medium access control
  - Logical link control
- Network and transport layers impact
- Application layer impact
  - Interaction models

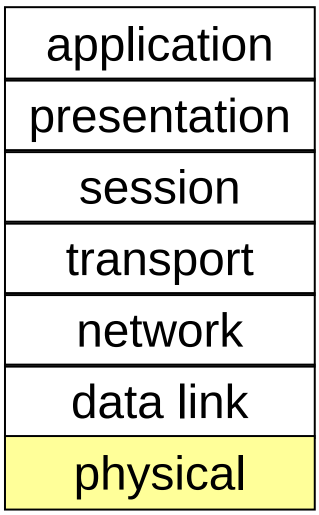
ISO OSI Model

- ISO: International Standards Organization
- OSI: Open System Inteconnection



Physical Layer

- Transport of bits
- Characteristics
  - Mechanical
  - Electrical (voltages, currents, impedance, baud rate, modulation, bit encoding, synchronisation, etc.)
  - Functional (topology, repeaters, etc.)



Data Link Layer

- Groups bits in frames
- Frame synchronization
- Detection (correction) of errors
- Flow control
- Management of access to medium

- Is often dependent on the physical layer

application
presentation
session
transport
network
data link
physical

Network layer

- Routing of packets across links
- Flow congestion / control
- Gives a unique address over the network

application
presentation
session
transport
network
data link
physical

Transport layer

- End to end reliable and transparent transport of information on a network
  - Checking and correcting errors
  - Flow regulation
- Establishment (release) of virtual circuits
- Multiplexing of virtual circuits

application
presentation
session
transport
network
data link
physical

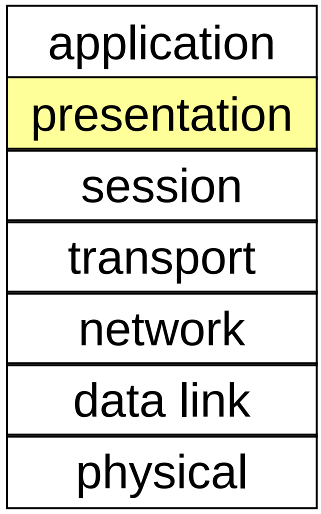
Session layer

- Management of dialog
  - Definition of synchronization points
  - Return to known state

application
presentation
session
transport
network
data link
physical

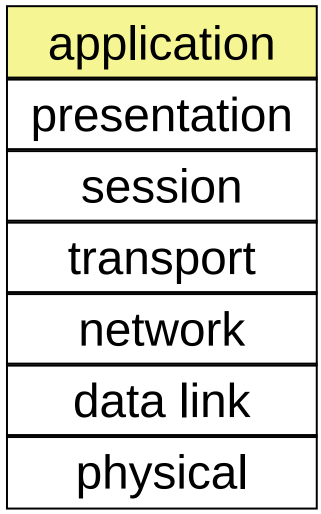
Presentation layer

- Format conversion
  - To and from transfer syntax
- Ciphering
- Data compression



Application layer

- The only one visible to the application
- Add semantics to the information transfers
  - Defines concepts
  - Provides services
- Ex. FTP, SNMP, HTTP



Interconnections

- Repeaters
  - physical layer
- Bridges
  - data link layer
- Routers
  - network layer
- Gateways
  - application layer

See [Perlman, 2000]

Repeaters

- Used when the protocols on all layers are identical on both sides
- Connect two data circuits
- Expand the distance covered by (or the number of devices connected to) a data link whether wireless or wired.
- Regenerate the signals received on one side and transmit them on the other side and vice-versa.
- On some occasions, may also be used to interconnect a wireless cell to a wired link
  - Word repeaters
- Ethernet hubs are an example of repeaters.

Bridges

- Interconnect subnetworks using the same layer protocols above the data link layer
- Interconnect data links
- Both sides must also use compatible addressing information
- Examples:
  - IEEE 802.11 base stations interconnect an Ethernet based link and a wireless cell.
  - an Ethernet switch is used to interconnect two or more Ethernet links.

Routers

- Operate at the network layer level
- Their task is to find a route to convey a message from a source to a destination
  - Exchange information between themselves in order to find such a route
  - Can thus find an optimum path between two nodes
    - whereas bridges only use a subset of the available topology.
- Difference with bridges
    - Bridges are transparent, routers are not
    - Routers modify the packets they forward in particular their address fields

Gateways

- Used when the protocols at the application layer are different on both sides
- Translate the messages from one protocol to the other one.
- Examples:
  - Connecting a Profibus or a CAN Open network to the Internet using HTTP over TCP/IP, requires a gateway because the protocols are different at all layers.
- Sometimes called “proxies”

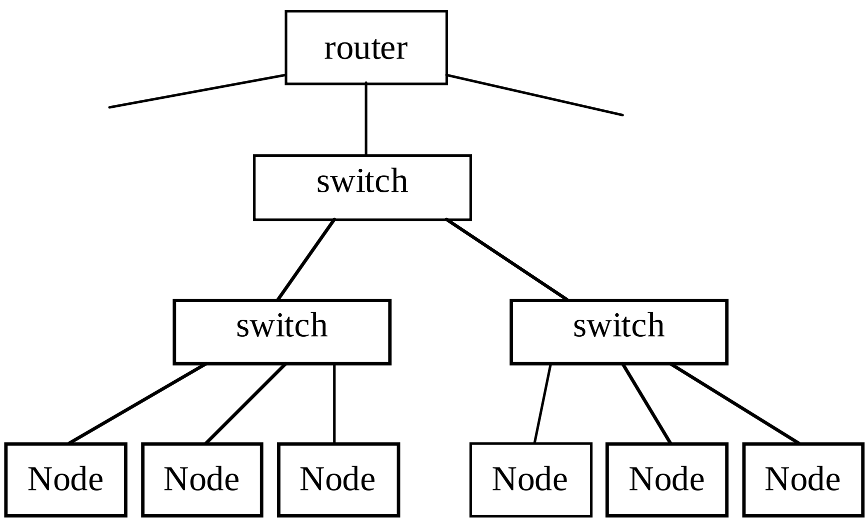
Impact of layers on QoS

- Observable properties of the network
  - Transfer delay bounds
  - Transfer delay variations (jitter)
  - Throughput
- All layers have an impact but some more than others

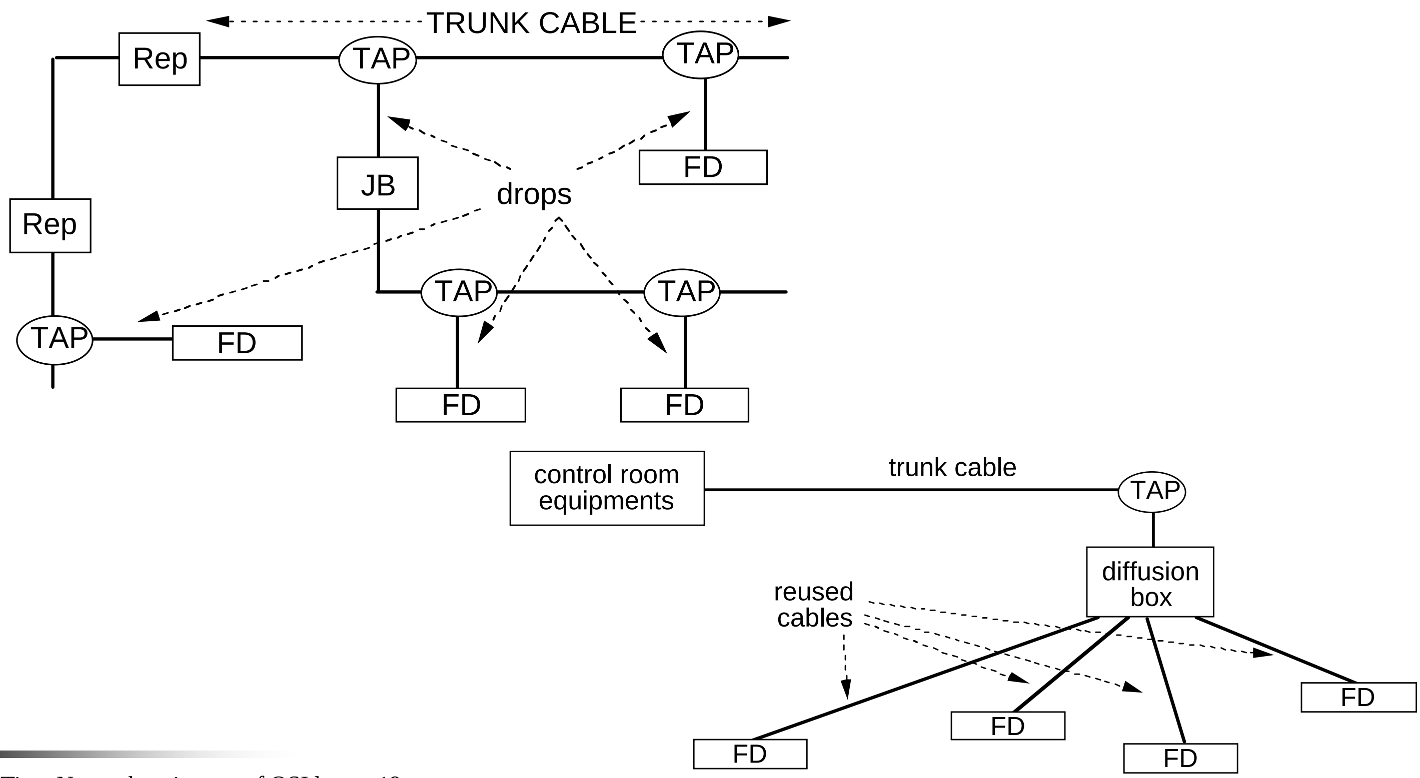
Physical layer

- Topology and physical limitations
  - How many nodes may be reached in one hop ?
- Bit rate (not Baud rate)
- Signal to noise ratio
  - Bit error rate
- Resilience to interferences
  - Bit error rate
  - Bursts of errors on bits

Topologies in offices



Topologies in factories



Medium Access Control

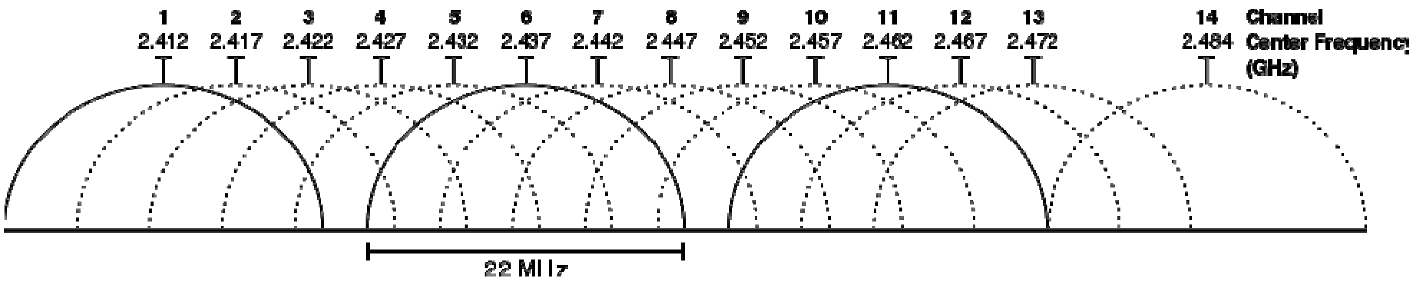
- Access mechanism
  - May be influenced by priorities
- Error detection scheme
  - Performance of error detection
- Error correction scheme
  - Automatic Repeat reQuest (stop and wait, selective repeat, Go back N)
  - FEC
  - Hybrid FEC-ARQ
- Packet delimitation
  - Packet error rate

Access mechanism

- How to isolate the emissions from different sources
- 3 basic choices
  - Use different frequency bands
    - Frequency Division Multiple Access (FDMA)
  - Emit at different instants
    - Time Division Multiple Access (TDMA)
  - Use a combination of both
    - Code Division Multiple Access (CDMA)
      - CDMA is also used to spread the spectrum of emission

FDMA

- Different transmitters use different channels
  - There is often some overlap between adjacent channels
    - Example: 802.11 (14 channels, but no more than 4 at any given place)



- Hardly used in wired LANs
- Hopping used in some WLANs to mitigate interferences (for instance DECT, wirelessHART)

TDMA

- All nodes use the same frequency but at different instants
- Some temporal synchronisation is thus required
- Advantages
  - The bandwidth can be adapted according to the emitter
  - It is possible to power off the emitter in absence of emission
- Drawbacks
  - Additional load due to synchronisation
  - More problems (than with CDMA) with multiple paths

CDMA

- Separation in time and frequency
- Two principles
  - Direct sequence: each bit is converted into a sequence of chips
  - Frequency hopping: each transmission is performed at a different carrier frequency (used in Bluetooth, Wireless HART and ISA100.11a)
- Advantages
  - Difficult to spy, rather insensitive to perturbations, no need for synchronisation, cells may use the same frequency band
- Drawbacks
  - Complex, requires control of emission power, requires a large frequency band

FDD and TDD

- 2 ways to handle full duplex operations
  - FDD (Frequency Division Duplexing)
    - Each direction uses a different band
  - TDD (Time Division Duplexing)
    - Both directions use the same band but at different instants

TDMA

- Predetermined
  - Each node has one (or more) slots in time
  - Usually called “TDMA” or “pure TDMA”
- Centralised access control
  - Polling, probing
- Decentralised techniques
- Reservation

Centralised access

- One master station / N slave stations
  - A slave station may only transmit as a response to the master station
- Advantages
  - Simple, the master is the unique point of coordination
  - Easy to adapt polling to slaves needs
  - Worst case polling time can be calculated
    - Good point for real-time applications
- Drawbacks
  - The master is a hot point for reliability
  - The master is used in each transfer -> additional delays
  - Not very efficient when few slaves are active (or numerous slaves)
    - Can be improved by probing

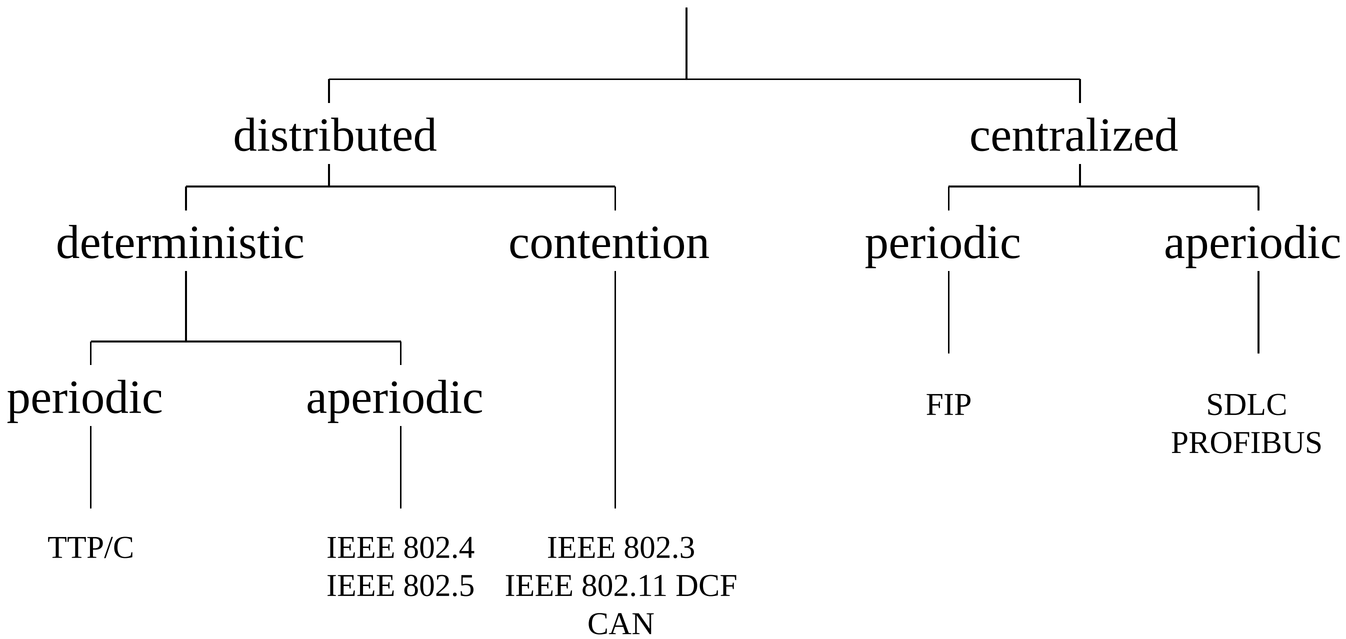
Distributed access

- Appealing as compared to centralized techniques
  - More reliable
  - Access delays often shorter
  - Better use of the bandwidth
  - No need for planning (i.e. in case of multiple wireless cells)
- drawbacks
  - Often more complex
  - Not always easy to predict temporal properties

Distributed access techniques

- Static (predetermined)
- Distributed probing
- ALOHA
- Carrier Sense Multiple Access (CSMA)
- Ethernet
- CSMA/CA
- Token bus
- Token ring

Classification of some solutions



Reservation

- When a node wants to transmit (for a long period)
  - Gets access and signals its request
  - Request is granted and resources are allocated
    - Resources may be slots or medium for a given duration
  - When the node no longer needs the resources, it releases them (may be automatic)
- There is no conflict on the resource use
- There might conflicts in the requests

- Widely used technique (cellular phones, 802.11, …)
- Interesting from the QoS perspective

Logical Link Control

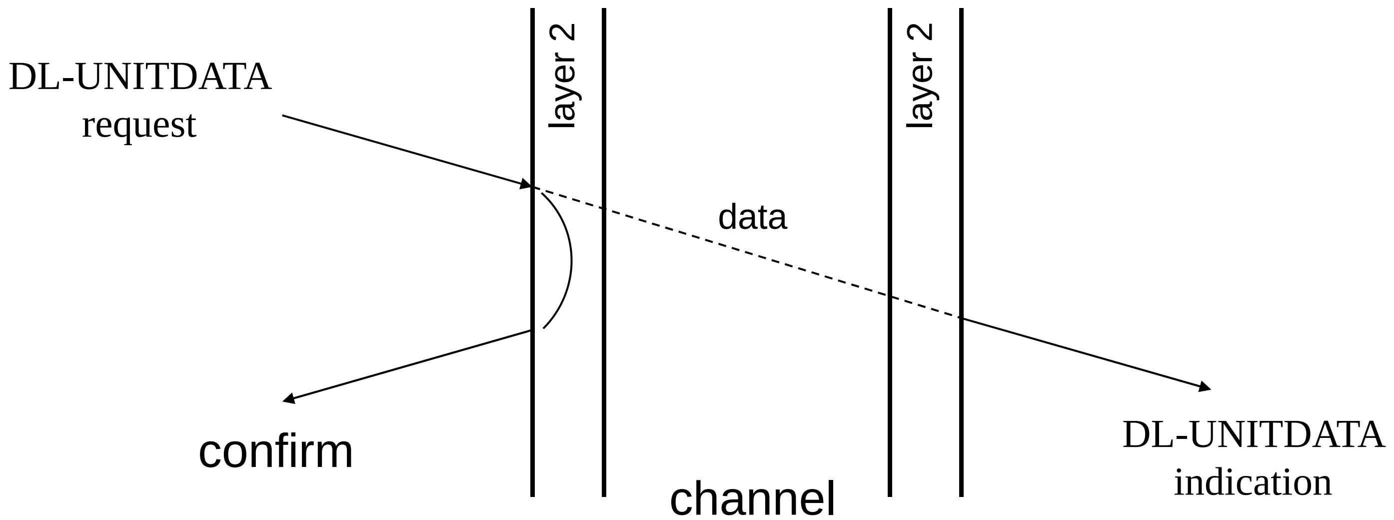
- Connectionless services
  - QoS: priority
  - SDN (Send Data with No ack)
    - Unacknowledged connectionless-mode data transfer
      - DL-UNITDATA request - DL-UNITDATA indication
  - SDA (Send Data with Ack)
    - Acknowledged connectionless-mode data unit transmission service
      - DL-DATA-ACK request, DL-DATA-ACK indication, DL-DATA-ACK-STATUS indication
  - RDR (Request Data with Reply) or SDR (Send Data with Reply)
    - Acknowledged connectionless-mode data unit exchange service
      - DL-REPLY request - DL-REPLY indication - DL-REPLY-STATUS indication
      - DL-REPLY-UPDATE request DL-REPLY-UPDATE-STATUS indication

Logical Link Control (2)

- Connection oriented service
  - QoS: priority
  - Connection establishment
    - DL-CONNECT request - DL-CONNECT indication - DL-CONNECT response - DL-CONNECT confirm
  - Data transfer
    - DL-DATA request -- DL-DATA indication
  - Termination
    - DL-DISCONNECT request -- DL-DISCONNECT indication
  - Reset
    - DL-RESET request -- DL-RESET indication -- DL-RESET response -- DL-RESET confirm
  - Flow control
    - DL-CONNECTION-FLOWCONTROL request - DL-CONNECTION-FLOWCONTROL indication (parameter: amount of data allowed)

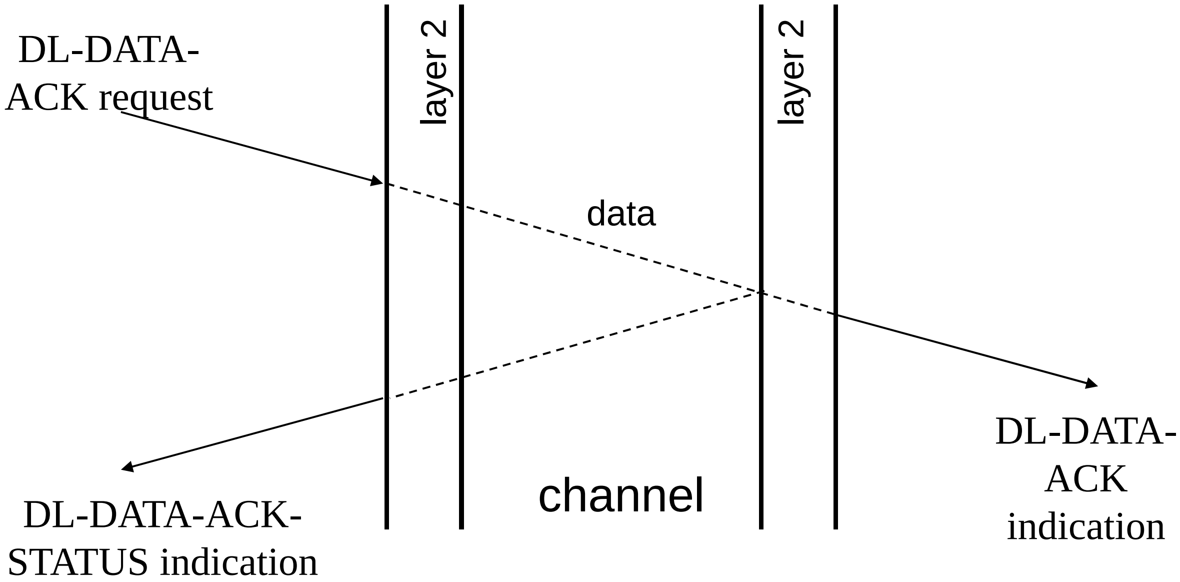
Send Data No acknowledge (SDN)

- No temporal problem (except access control)
- Possible response is separated (adds time)
- May be used to synchronise (multicast or broadcast)



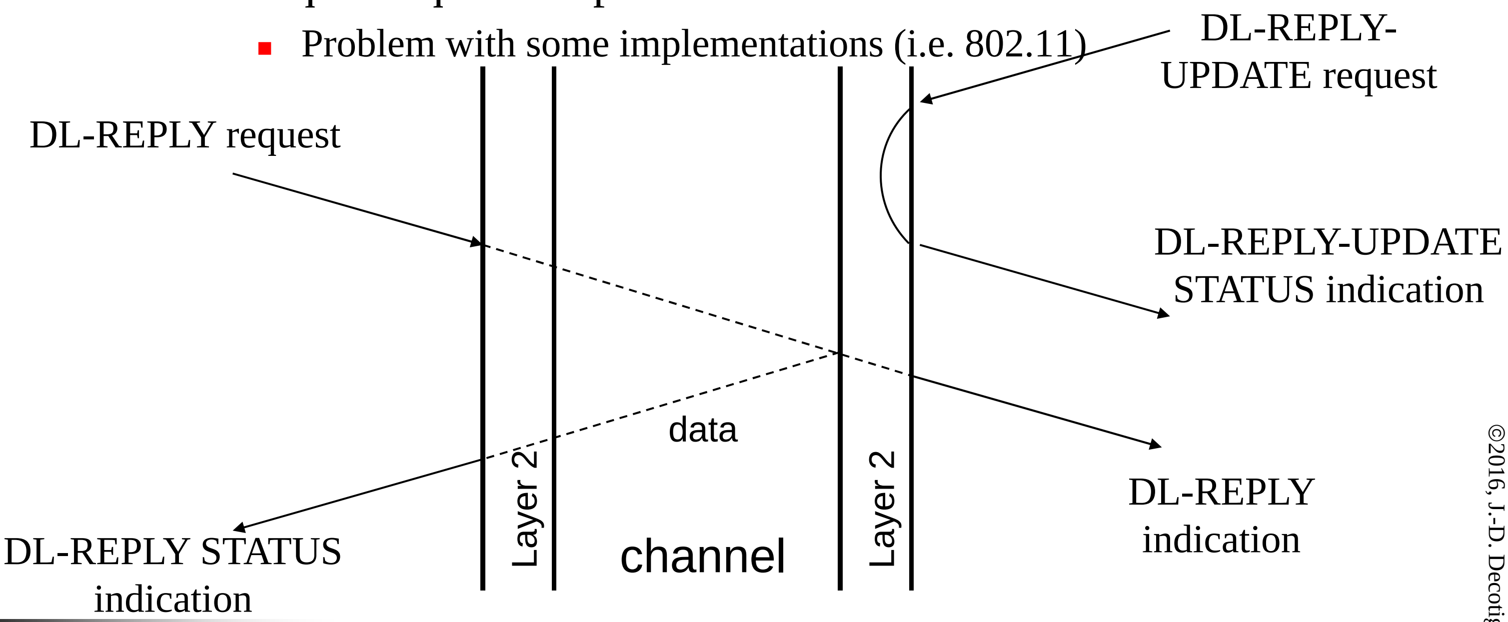
Send Data with Ack. (SDA)

- No temporal problem (except access control)
- Possible response is separated (adds time)



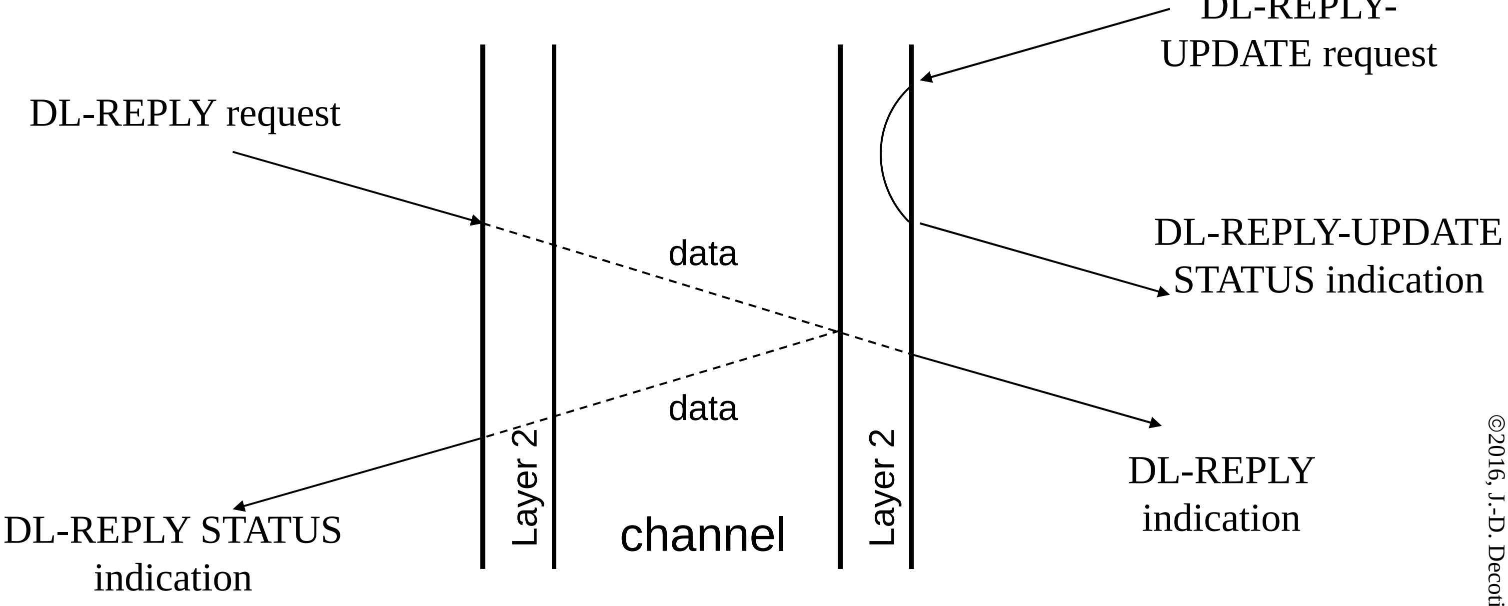
Request Data with Response (RDR)

- Good to decouple requester from provider applications
- Requires quick response
  - Problem with some implementations (i.e. 802.11)



Send Data with Response (SDR)

- Good to decouple requester from provider applications
- Requires quick response



Network Layer

- QoS negotiation and admission control
- Resource reservation
- Packet buffering and scheduling
- Resource management
- Routing table management
  - See [Pragyansmita]

- Metrics
  - Bandwidth, delay, delay variation (jitter)

Network Layer (2)

- Enabling QoS routing of data
  - Consider various metrics to select the best route
  - Provide a fair bandwidth to non QoS flows
  - Graceful performance degradation
- Approaches
  - Statefull: manage per flow state & perform per flow operations
    - Intserv + RSVP
  - Stateless:
    - DiffServ (different behavior between core and edge routers)

Transport Layer

- Connection establishment and release
- Flow control mechanisms
- Error control mechanisms

ISO transport layer QoS parameters

- Connection establishment delay: max. acceptable time between a transport connection being requested and its confirmation being received by the user
- Connection establishment failure probability: probability that a connection cannot be established within the max. delay Connection release delay: max. acceptable delay between a user initiating release of a connection and actual release at peer user
- Throughput: number of bytes of user data sent per unit of time
- Transit delay: elapsed time between submission and delivery
- Residual error rate: ratio of incorrect, lost and duplicate TSDUs to the total number sent
- Transfer failure probability: ratio of total transfer failures to total transfer samples during a given window

ISO transport layer QoS parameters

- Connection Release Failure Probability: fraction of connection release attempts that did not complete within the connection release delay interval (as agreed)
- Protection: used by the user sender to specify interest in having the transport protocol provide protection against unauthorized third parties reading or modifying the transmitted data.
- Priority: used to specify the relative importance of transport connections. In case of congestions or the need to recover resources, lower-priority connections are degraded or terminated before higher-priority ones.
- Resilience: probability that the transport protocol will spontaneously terminate a connection due to internal or network problems [Iren 99]

Session and Presentation Layers

- Session layer
  - Check points
    - Frequency of check pointing impact time lost for recovery
- Presentation layer
  - Compression
  - Transfer syntax compactness

Application Layer

- Interaction model [Thomesse 93]
  - Client-server
    - Need to wait until server responds
  - Publish-subscribe
    - Temporal decoupling between the publisher and the user
  - Producer-consumer

References

- R. Perlman, “Interconnections – bridges and routers”, 2nd edition, Addison Wesley, Reading, 2000
- S. Iren et al., “The transport layer: tutorial and survey”, ACM Computing Surveys, vol. 31, no. 4, pp. 360-404., Dec. 1999.
- J.-P. Thomesse, “Time and industrial local area networks”, in Proc. of 7th COMPEURO'93, Paris-Evry, France, May 24-27, 1993, pp. 365-374.
- ISO/IEC 7498:1: 1996. Information Processing Systems - Open systems interconnection - Basic reference model : the Basic model
- ISO/IEC 8802.2: 1998. Information technology - Telecommunications and information exchange between systems - Local and metropolitan area networks - Specific requirements. Part 2: Logical link control

References (2)

- Pragyansmita Paul and S.V.Raghavan, “survey on QoS Routing”, 15th International Conference on Computer Communications, August 12-14, 2002, Mumbai.
- M. El-Gendy, A. Bose, K. Shin, “Evolution of the Internet QoS and Support for Soft Real-Time Applications”, proc. of the IEEE, vol. 91 (7), July 2003, pp. 1086-1104.
- ISO/IEC 15802-3: 1998, ANSI/IEEE Std 802.1D, 1998 , Information technology - telecommunications and information exchange between systems – local and metropolitan area networks - common specifications. Part 3: Media Access Control (MAC) bridges, 10 Dec. 1998

REAL-TIME NETWORKS
Controller Area Network

Prof. J.-D. Decotignie
CSEM Centre Suisse d'Electronique et de Microtechnique SA
Jaquet-Droz 1, 2007 Neuchâtel
jean-dominique.decotignie@csem.ch

Outline

- CAN (Controller Area Network) history and use
- Architecture and variants
- Physical layer
- Data link layer
- Temporal behavior
- Response time analysis
- Improving the real-time behavior

History

- 1983: Initial development by Bosch
- 1986: Official introduction of CAN protocol
- 1991: CAN specification 2.0 published
- 1991: first application layer (CAN Kingdom)
- 1992: 2nd AL, CAN Application Layer (CAL) by CiA
- 1992: 1st cars from Mercedes-Benz using CAN
- 1993: becomes an international standard (ISO 11898)
- 1994: modifications for industrial use :DeviceNet by Allen-Bradley and SDS by Honeywell
- 1995: CAN open protocol from CiA (CAN in Automation)
- 2000: Time-triggered comm. protocol for CAN (TTCAN)
- 2012: CAN-FD (payload up to 64B @ bit rate 8 x arbitration)

Market

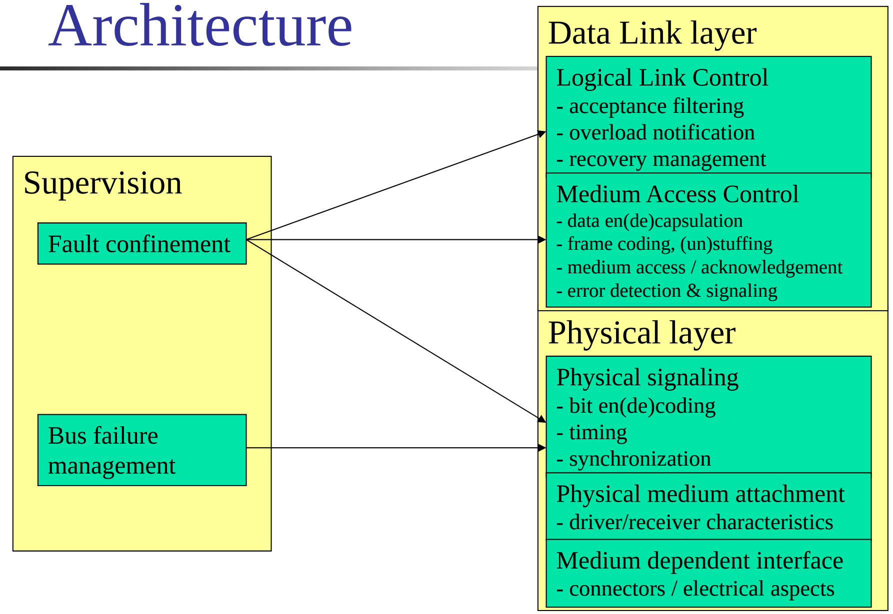
- Initially development for the car industry
  - In competition with others
- Attracted a lot of interest outside initial market
  - Due to availability of inexpensive silicon
- Today:
  - Actually used in cars but not for safety critical functions
  - Some use in the industry
  - Widely available in micro-controllers

Architecture

- 3 layer collapsed model
- Only first 2 layers are standard
- A few proposals for AL
  - CAL
  - CAN Kingdom
  - DeviceNet
  - CAN Open
  - SAE J1939
  - ….

application
presentation
session
transport
network
data link
physique

Architecture



Variants

- SDS
- DeviceNet
- TT-CAN
- FTT-CAN
- CAN-FD

Physical layer

- Topology
  - Terminated bus
- Number of stations
  - In principle limited to 30 (depends on drivers)
- Medium
  - Twisted pair, single wire (FO possible but not standard)
- Range
  - Signaling speed and propagation speed dependent: 40m at 1Mbit/s
  - Drop length limited to 30 cm
- Signaling and bit encoding
  - 10 kbit/s to 1 Mbit/s, NRZ
  - Up to 8 Mbit/s in payload with CAN-FD

Physical layer (2)

- Synchronization
  - Uses signal edges (implies bit stuffing with NRZ)
    - After 5 consecutive ones, a zero is inserted
    - After 5 consecutive zeros, a one is inserted
    - This rules includes a possible stuffing bit inserted before
- Signals
  - Recessive: logical “1”
  - Dominant: logical “0”
    - When 2 stations compete on a bit by bit basis, the station that emits a dominant bit imposes this level on the bus

Medium Access Control - frame

- RTR: data or request frame
- IDE: normal or extended format
- DLC: data field length
- EOF: End of frame (7 recessive bits)
- Ack: global ack by all connected nodes
- IFS: inter frame silence (3 recessive bits)
- R0: reserved (dominant)

SOF	Identifier	RTR	IDE	R0	DLC	Data	CRC	ACK	EOF	IFS
1	11	1	1	1	4	0..8 x 8	16	2	7	3

Extended addressing

- RTR: data or request frame
- IDE: extended format (recessive)
- DLC: data field length
- EOF: End of frame (7 recessive bits)
- Ack: global ack by all connected nodes
- IFS: inter frame silence (3 recessive bits)
- R0,R1: reserved (dominant)
- SRR: (recessive)

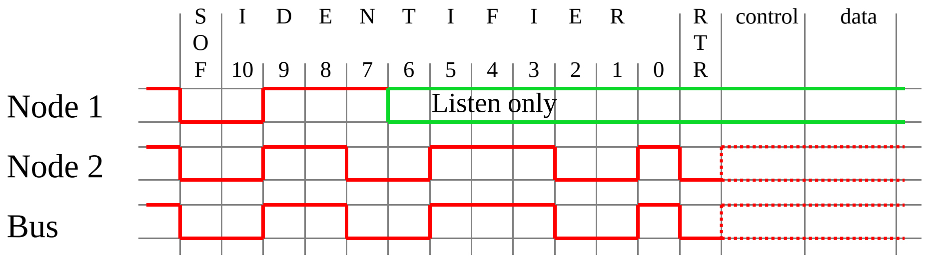
SOF	Identifier	SRR	IDE	Identifier extension	RTR	R1	R0	DLC	Data	CRC	ACK	EOF	IFS
1	11	1	1	18	1	1	1	4	0..8 x 8	16	2	7	3

Addressing

- Single 11 or 29 bit identifier par frame
  - If used to identify a node
    - Source (data) or destination (request) of the message
  - Normally used to identify the payload
    - Also called “Broadcast source addressing”
  - A lower value gives higher priority in contention

Medium Access Control

- CSMA with collision resolution
  - Each node observes bus while transmitting
  - If level different from what it has put, withdraws
  - Dominant bit overwrites recessive one



Error detection

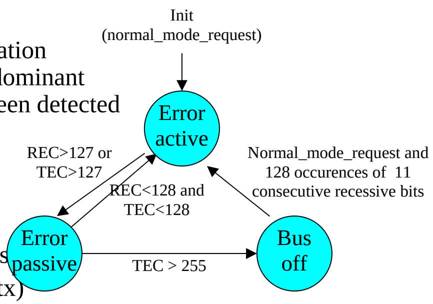
- Several means
  - Bit error
    - When what is one the bus is different from what was emitted
      - Except when a recessive bit was emitted during arbitration or ack slot
  - Cyclic Redundancy Check (CRC)
  - Frame check (the frame structure is checked)
  - ACK errors (absence of a dominant bit during the ack slot)
  - Monitoring (each node which transmits also observes the bus level and thus detects differences between the bit sent and the bit received)
  - Bit stuffing (checking adherence to the stuffing rule.)
- A frame is valid for
  - A transmitter if there is no error until the end of EOF
  - A receiver if there is no error until the next to last bit of EOF

Behaviour in case of error

- In case of stuff, bit, form or acknowledge errors
  - An error flag is started at the next bit
- In case of CRC error
  - An error frame is send after the ack delimiter
- Fault confinement
  - Each time an reception error occurs, REC is incremented
  - Each time a frame is received correctly, REC is decremented
  - Same for the emission errors with TEC
  - The values of TEC and REC may trigger mode changes

Connection modes

- To enforce fault confinement, nodes may be in one of 3 modes
  - Error active
    - Normally takes part to the communication and may send an active error flag (6 dominant consecutive bits) when an error has been detected
  - Error passive
    - Takes part in communication but must not send an active error flag. Instead, it shall send a passive error flag (6 recessive consecutive bits)
    - Some restrictions (silence between 2 tx)
  - Bus off
    - Cannot send or receive any frame.
    - A node is in this state when it is switched of the bus due to a request from a fault confinement entity. May exit from this state only by a user command.



Error frame

- 2 fields: error flag and error delimiter
- Error flag
  - Active: 6 dominant bits
  - Passive: 6 recessive bits
  - As all nodes monitor the bus and the flag violates stuffing rules, they will send error flags too
    - The error flag will last from 6 to 12 bits
- Error delimiter (8 recessive bits)
  - After sending an error flag, a node shall send recessive bits
    - As soon as it senses a recessive bit, it sends 7 recessive bits

Error recovery

- Automatic retransmission
  - of all frames that have lost arbitration
  - of all frames have been disturbed by errors during transmission

Medium Access Control

- All messages are sent in broadcast
- Nodes filter according to their interest
- All messages are acknowledged including by nodes that are not interested by the message
  - Acknowledge just means “message well received by all receivers”
    - It does not mean “intended receiver received it”
- Node that does not receive message correctly sends an error bit sequence
- Node that is too busy may send an overload bit seq.
  - MA_OVLD.request/indication/confirm
  - Same principle as an error frame (overload frame = 6 dominant + 8 recessive bit)

Logical Link Control

- 2 types of services (connectionless)
  - Send Data with no ack
    - L_DATA.request, L_DATA.indication, L_DATA.confirm
    - Uses a data frame
  - Request Data
    - L_REMOTE.request, L_REMOTE.indication, L_REMOTE.confirm
    - Uses a remote frame (same as a data frame but data field is empty)
- Flow control using the overload bit sequence

Temporal analysis

- CAN is essentially unfair
  - Lower identifier frames get priority
  - There exists schemes to overcome this
    - [Cena 2001]
    - DeviceNet to some extent
- However
  - Given a set of traffic needs
  - Given that all nodes comply with the expressed needs
    - It is possible to check whether a network will comply with the requirements

Traffic requirements

- Real-time periodic message streams
  - $M_{p,m}=\{T_m, D_m, C_m, J_m\}$
    - C_m duration of emission of message, T_m period of transfer, D_m relative deadline from beginning of period (absolute deadlines are $d_{n,m} = n\,T_m + D_m$)
    - J_m is the arrival jitter (variability in queuing instants periodicity)
- Real-time sporadic message streams
  - $M_{s,m}=\{T_m, D_m. C_m\}$
    - T_m min. interarrival time, D_m relative deadline from arrival time (absolute deadlines are $d_{n,m} = arr_{n,m} + D_m$)

Response time analysis

- The worst case response time of a message R_m is the longest time between
  - Queuing of the message
  - Arrival time of the message at destination
- A set of traffic requirements can be handled by a CAN networks iif $R_m \leq D_m$ for all message streams
- $R_m = C_m + J_m + I_m$
  - Where I_m is the interference time

Response time analysis

- Response time [Tindell 94&95]

$$R_m = C_m + J_m + I_m \text{ with } I_m = B_m + \sum_{\forall j \in hp(m)} \left\lceil \frac{I_m + J_m + \tau_{bit}}{T_j} \right\rceil C_j$$

  - B_m is the transmission of the longest lower priority message + S the duration of the interframe silence
  - Assumes no new message queued before previous sent
- Different from Joseph and Pandya

$$R_m = C_m + \sum_{\forall j \in hp(m)} \left\lceil \frac{R_m}{T_j} \right\rceil C_j$$

Message emission time

- Influence of bit stuffing
  - Normal addressing

$$C_m = \left(47 + 8b + \left\lfloor \frac{34 + 8b - 1}{4} \right\rfloor\right)\tau_{bit}$$

  - Extended addressing

$$C_m = \left(65 + 8b + \left\lfloor \frac{52 + 8b - 1}{4} \right\rfloor\right)\tau_{bit}$$

    - In the original paper 5 was used instead of 4 in the division. This is a mistake due to a lack of understanding of the stuffing mechanism that also includes stuff bits

Example - SAE benchmark

		Bits	Period [ms]	Typ	D[ms]	source
1,2	Traction battery voltage / current	8/8	100/100	P/P		Battery
3,5	Traction battery temp. average / max	8/8	1000/1000	P/P		Battery
4,6	Auxiliary battery voltage / current	8/8	100/100	P/P		Battery
7,17	Accelerator position / switch	8/2	5/50	P/S	-/20	driver
8,9	Brake pressure master cylinder / line	8/8	5/5	P/P		brakes
10,11	Transaxle lubrication/Trans. clutch line pressure	8/8	100/5	P/P		Transmission
12,18	Vehicle speed / brake switch	8/1	100/20	P/S	-/20	brakes
13,36	Traction battery ground fault / idem test	1/2	1000/1000	P/P		Battery/ V-C
14,28	HI&LO contactor open-close / interlock	4/1	50/50	S/S	5/20	Battery
15,16	Key switch run / start	1/1	50	S/S	20/20	driver
19,20	Emergency brake / shift lever	1/3	50/50	S/S	20/20	driver
22,26	Speed control / brake mode / SOC reset	3/1/1	50/50/50	S/s/s	All 20	driver
21	Motor/trans over temperature	2	1000	P		transmission
23-25	12V Power ack vehicle ctrl / inverter / I-M cont	1/1/1	50/50/50	S/S/s	All 20	battery

Example - SAE benchmark

		Bits	Period [ms]	Typ	D[ms]	source
29,30	High / low contactor control	8/8	10/10	P/P		V-C
31,35	Reverse & 2nd gear clutches / 12V power relay	2/1	50/50	S/S	20/20	V-C
32,33	Clutch pressure ctrl / DC-DC converter	8/1	5/1000	P/P		V-C
34,37	DC-DC converter current ctrl / brake solenoid	8/1	50/50	S/S	20/20	V-C
38,39	Backup alarm / warning lights	1/7	50/50	S/S	20/20	V-C
40	Key switch	1	50	S	20	V-C
42,43	Torque command / torque measured	8/8	5/5	P/P		V-C / I-M C
41,45	Main contactor close / Fwd-rev ack	1/1	50/50	S/S	20/20	I-M C
44,46	Fwd-Rev / idle	1/1	50/50	S/S	20/20	V-C
47,49	Inhibit / Processed motor speed	1/8	50/5	S/P	20/-	I-M C
48,53	Shift in progress / main contactor ack	1/8	50/50	S/S	20/20	V-C
50	Inverter temperature status	2	50	S	20	I-M C
51	Shutdown	8	50	S	20	I-M C
52	Status-malfunction	1	50	S	20	I-M C

Timing analysis [Tindell 95]

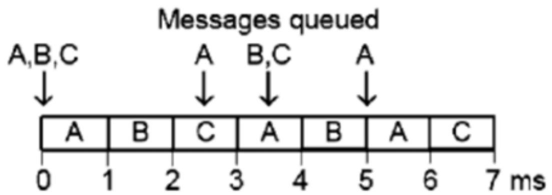
Signal number	Byte	Period [ms]	Type	D[ms]	R [ms]	source
14	1	50	S	5	1.544	Battery
8,9	2	5	P	5	2.128	Brakes
7	1	5	P	5	2.632	Driver
43,49	2	5	P	5	3.216	I-M C
11	1	5	P	5	3.72	Trans
32,42	2	5	P	5	4.304	V-C
31,34,35,37-40,44,46,48,53	6	10	P	10	5.192	V-C
23,24,25,28	1	10	P	10	8.456	Battery
15-17,19,20,22,26,27	2	10	P	10	9.04	Driver
41,45,47,50-52	2	10	P	10	9.624	I-M C
18	1	100	S	20	10.128	Brakes
1,2,4,6	4	100	P	100	18.944	Battery
12/10/21	1/1/1	100/100/1000	P/P/P	=T	19.44/19.55/29.19	Br/trans
3,5,13/33,36	3/1	1000/1000	P/P	=T	20.608/29.696	Batt/V-C

@125 kbit/s

Flaw in Tindell's analysis

- Consider

Message	Priority	Period	Deadline	Tx time
A	3	2.5 ms	2.5 ms	1 ms
B	2	3.5 ms	3.25 ms	1 ms
C	1	3.5 ms	3.25 ms	1 ms



New Analysis

- Use the definition of busy period. Level-m busy period is defined as:
  - It starts at some time t^s when a message of priority m or higher is queued ready for transmission and there are no messages of priority m or higher waiting to be transmitted that were queued strictly before time t^s .
  - It is a contiguous interval of time during which any message of priority lower than m is unable to start transmission and win arbitration.
  - It ends at the earliest time t^e when the bus becomes idle, ready for the next round of transmission and arbitration, yet there are no messages of priority m or higher waiting to be transmitted that were queued strictly before time t^e.
- All messages of priority m or higher, queued strictly before the end of the busy period, are transmitted during the busy period.
- These messages cannot therefore cause any interference on a subsequent instance of message m queued at or after the end of the busy period.
- Maximal busy period start at the critical instant when all message m is queued simultaneously with all higher priority msg and each of these is subsequently queued at highest speed

Where is the flaw ?

- Tindell's analysis

$$R_m = C_m + J_m + I_m \text{ with } I_m = B_m + \sum_{\forall j \in hp(m)} \left\lceil \frac{I_m + J_m + \tau_{bit}}{T_j} \right\rceil C_j$$

- Implicitely assumes that level-m busy period will end at or before T_m.
  - Would be true for preemptive scheduling as on completion of message m, no higher priority message would be pending
  - However in CAN, on completion of a message transmission, a higher priority message may be pending
    - Level-m busy period may extend beyond T_m.

New Analysis

- Busy period length $t_m = B_m + \sum_{\forall j \in hep(m)} \left\lceil \frac{t_m + J_m}{T_j} \right\rceil C_j$
- Number of instances of message m that becomes ready for transmission before the end of the busy period

$$Q_m = \left\lceil \frac{t_m + J_m}{T_m} \right\rceil$$

- The longest time from the start of busy period to instance q (q=0 is first one) starting transmission is

$$I_m(q) = B_m + qC_m + \sum_{\forall j \in hp(m)} \left\lceil \frac{I_m + J_m + \tau_{bit}}{T_j} \right\rceil C_j$$

Conclusion on response time analysis

- Takes profit of fixed priority nature of transmission
- Provided all nodes play the game (do not exceed traffic announced)
  - Allows to assess if traffic can be handled in absence of errors

Improving the real-time behaviour

- Provide fairness guarantees
  - MUST and CAN [Cena 01]
- Better predictability of periodic traffic
  - TTCAN [Leen 02]
- Same with handling sporadic traffic
  - FTT-CAN [Almeida 02]
- Comparing CAN and TTCAN in presence of errors
  - [Broster 04]

Updated response time analysis

- Response time [Broster 04]

$$R_m = t_m + J_m \quad \text{with} \quad t_m = B_m + C_m + I_m(t_m) + E_m(t_m)$$

  - $E_m(t)$ is the worst case overhead due to network faults and extra frames occurring before t
  - B_m is the transmission of the longest lower priority message + S the duration of the interframe silence
  - $I_m(t) = \sum_{\forall j \in hp(m)} \left\lceil \frac{t - C_m + J_j + \tau_{bit}}{T_j} \right\rceil (S + C_j)$
  - Assumes no new message queued before previous sent

Taking faults into account

- Here we only consider network faults
  - All nodes are well behaved (according to protocol)
  - Caused by electromagnetic interferences or physical faults
    - It is possible for 2 nodes on the bus to simultaneously read different values from the bus
- Scenarios
  - Error during a data frame $\Rightarrow$ error frame (duration E is transmitted): additional delay = C+E+S
  - Error during error frame: additional delay = E+S
  - Error during bus idle (false start): add. delay = E+S
  - Burst of interferences of duration Z: add. delay = Z+C+E+S

Sporadic fault model

- Faults are always separated by a minimum inter-arrival time T_F.
- n_{burst} is the maximum number of faults that can occur in a succession during a burst

$$E_m(t) = \left(n_{burst} + \left\lceil \frac{t}{T_F} \right\rceil \right)\left(E + \max_{j \in hep(m)} C_j \right)$$

  - hep(m) is the set of messages with priority higher or equal to m

Probabilistic fault model

- Probability that k faults occur in interval t is given by

$$p(F = k, t) = \frac{(\lambda t)^k e^{-\lambda t}}{k!}$$

- Each fault causes the maximum length error frame and occurs on the last bit of the longest frame

$$M_m = E + \max_{\forall j \in hep(m)} C_j$$

- The error overhead function is a random distribution $E_m = kM_m$ with probability $p(F = k, t)$
- So we get a set of pairs $R_m = \langle t_m, p(t_m) \rangle$

Analysis

- $R_m = \langle t_m, p(t_m) \rangle$ gives us the response time in case of k errors (denoted as $R_{m|k}$) and the probability that k faults occurred between times 0 and t_m
- However, this cannot be taken as is:
  - For instance $R_{m|1}$ is the response time when one fault occurred before $R_{m|1}$.
  - However this fault must have occurred before $R_{m|0}$
    - Otherwise we would have had a response time of $R_{m|0}$

$$P(R_{m|1}) = p(1,[0, R_{m|1}]) - P(R_{m|0})\, p(1,]R_{m|0}, R_{m|1}])$$

Scenari

Response Time	Possible Scenarios (Shorthand)	Number of Scenarios
$R_{i\|0}$	[0]	1
$R_{i\|1}$	[10]	1
$R_{i\|2}$	[200], [110]	2
$R_{i\|3}$	[3000], [2100], [2010], [1200], [1110]	5
$R_{i\|4}$	[40000], [31000], [30100], [30010], [21100], [21010], ...	14

Source: I. Broster, 2004

Analysis (2)

- With 2 faults
  - We take the probability to have 2 faults before $R_{m,2}$
  - minus the case in which there is 0 fault before $R_{m,0}$
  - Minus the case in which there is 1 fault before $R_{m,1}$

$$P(R_{m|2}) = p(2,[0,R_{m|2}]) - P(R_{m|1})\,p\big(1,]R_{m|1},R_{m|2}]\big) - P(R_{m|0})\,p\big(2,]R_{m|0},R_{m|2}]\big)$$

Analysis (3)

- For k faults, this generalizes to:

$$P(R_{m|k}) = p(k,[0,R_{m|k}]) - \sum_{j=0}^{k-1} P(R_{m|j})\,p\big(k-j,]R_{m|j},R_{m|k}]\big)$$

- With

$$P(R_{m|0}) = p(1,[0,R_{m|0}])$$

- The probability of a deadline failure is then given by:

$$p_m(R_m > D_m) = 1 - \sum_{\forall j|R_{m|j}<D_m} p(R_{m|j})$$

Case study

- SAE benchmark difficult to use because very tight
- Mobile robot message set [Borster 04]
  - bus at 256 kbit/s, 41% use, λ=30 faults/s

Priority	Length [ms]	Period [ms]	Deadline [ms]	WCRT [ms]	Probability of failure	signal
6	0.288	2	2	0.828	$1.5\ 10^{-5}$	Motor control
5	0.328	4	4	1.168	$1.6\ 10^{-9}$	Wheel 1
4	0.328	4	4	1.508	$8.7\ 10^{-8}$	Wheel 2
3	0.528	8	8	2.048	$2.7\ 10^{-9}$	Radio In
2	0.248	12	12	2.608	$2.1\ 10^{-12}$	Proximity
1	0.528	240	240	2.32	$< 10^{-20}$	Logging

Summary

- Efficient for sporadic traffic
  - No control / non destructive distributed access
- Low response time in case of low traffic
- Periodicity difficult to achieve without jitter and some additional control
- A lot of care for error detection and handling
- Provides spatial consistency (global ack)
- Multicast easily implemented (broadcast source addr)
- Bounds can be derived for all kind of traffic
  - Under some assumptions
- MAC essentially unfair but there exists improvements

References

- Standard
  - ISO 11898, « Road Vehicles – Interchange of digital information – Controller area network (CAN) for high-speed communications », 1993.
- Sources of information www.can-cia.org
- Response time analysis
  - K. Tindell, A. Burns and A. J. Wellings, Calculating controller area network (can) message response times, Control Engineering Practice, vol. 3(8), pp. 1163-69, 1995.
  - K. Tindell et al. « Analyzing real-time communications: Controller Area Network (CAN) » , 1994
  - R. Davis et al., « Controller Area Network (CAN) schedulability analysis: Refuter, revisited and revised », Real-Time Systems, vol. 35, pp. 239-272, 2007.

References (2)

- Response time analysis in case of errors
  - N. Navet, Y. -Q. Song and F. Simonot, Worst-case deadline failure probability in real-time applications distributed over controller area network , Journal of Systems Architecture, vol. 46 (7), pp. 607-617.
  - I. Broster, A. Burns, G. Rodriguez-Navas, Comparing real-time communication under electromagnetic interference, ECRTS 2004. pp. 45 – 52
  - Broster, I.; Burns, A.; Rodriguez-Navas, G.; Probabilistic analysis of CAN with faults, 23rd IEEE Real-Time Systems Symposium, 3-5 Dec. 2002, pp. 269 – 278
  - L. Pinho et al., « Integrating inaccessibility in response time analysis of CAN networks », Proc. 3rd Workshop on Factory Comm. Systems, Porto, Sept. 2000, pp. 77-84.

References (3)

- improvements
  - G. Cena, A. Valenzano, "Integrating the CAN and MUST access techniques in a single fieldbus", Proc. 8th int. conf. on Emerging Technologies and Factory Automation, 15-18 Oct. 2001, pp.231 – 239.
  - G. Leen, D. Heffernan, "TTCAN: a new time-triggered controller area network", Microprocessors and microsystems, Vol. 26, Issue 2, Pages 51-94 (17 March 2002), Pages 77-94.
  - L. Almeida et al.,"The FTT-CAN protocol: why and how", IEEE Transactions on Industrial Electronics, vol. 49 , Issue: 6 , Dec. 2002, pp.1189 - 1201 .
  - SDS: http://content.honeywell.com/sensing/prodinfo/sds/
  - I. Broster et al., « Comparing real-time communications under electromagnetic interference », Proc. 16th Euromicro Conf. On real-time systems, Catania, Italy, July 2002, pp.
  - CAN-FD: http://www.can-cia.org/can-knowledge/can/can-fd/

Real-Time Networks
WorldFIP

Prof. J.-D. Decotignie
CSEM Centre Suisse d'Electronique et de Microtechnique SA
Jaquet-Droz 1, 2007 Neuchâtel
jean-dominique.decotignie@csem.ch

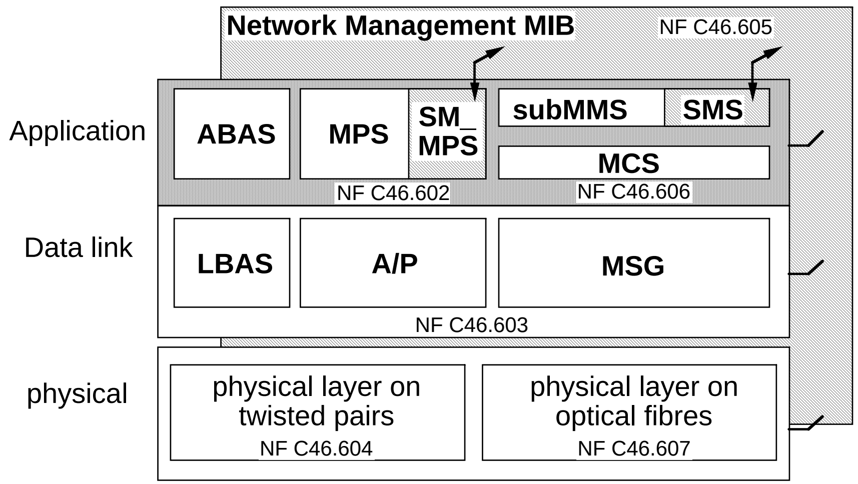
Outline

- WorldFIP history
- Architecture
- Physical layer
- Data link layer
  - Medium access control
  - Variable transfers
  - Message transfers
  - Consistency
- Temporal behavior
- Response time analysis

History

- 1984: White paper
  - D. Galara, J.-P. Thomesse, "Groupe de réflexion FIP: proposition d'un système de transmission série multiplexée pour les échanges entre des capteurs, des actionneurs et des automates réflexes", French Ministry of Industry, Paris, May 1984.
- 1989: French Standard (NF C46601-46607)
- 1996: European Standard (EN 50170)
- 2003: ISO/IEC standard (IEC 61158)

Architecture



Basic Choices

- Real-time traffic consists of state information transferred in a cyclic or periodic manner but not event transmission
- All information is broadcast
- For real-time traffic,
  - only variables are identified (identifiers). Network nodes or application processes are not identified
  - the communication model is the Producer-Distributor-Consumer (PDC) Model where variable values transferred on the network are neither queued at production nor at reception
  - There is no retransmission and no acknowledge for real-time traffic
  - Consumers are responsible for checking transmission status and taking appropriate actions when problems arise

Basic Choices (2)

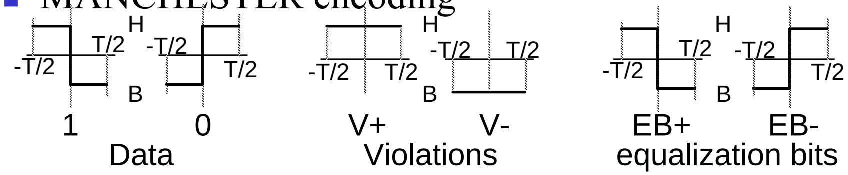
- All variables transferred in real-time and their characteristics (period, relation with other transfers) are known in advance
- MAC handled by a unique active distributor. Backup allowed
- The response time is bounded between 10 and 70 bit times.
- Transfers are either periodic or sporadic.
  - The transfer period for the values corresponding to a variable may differ from the period associated to another variable. The only restriction is that periods should be integer multiples of a basic period.
- Mechanisms to indicate temporal validity as well as temporal and spatial consistency of variable values built in the protocol.
- The non real-time traffic handled using the conventional client-server model and the corresponding application is a subset of MMS (Manufacturing Message Specification).

Physical layer

- twisted pairs (C46-604), optical fibers (C46-607)
- 3 speed classes
  - S1 (31.25 Kbits/s) / S2 (1 Mbits/s) / S3 (2.5 Mbits/s)
- Number of elements and range limited by
  - Propagation time Tp
    - normal case: Tp < 20 Tbit / extended: Tp < 40 Tbit
  - Losses and distortion on the cable
- 3 conformance classes
  - CH: high level class, allows long distance transmission at high data rates
  - CM: medium power class
  - CL: low power class for application with intrinsic safety requirements.
- recommended speeds
  - S1 or S2 for CL, S2 for CM, S2 or S3 for CH

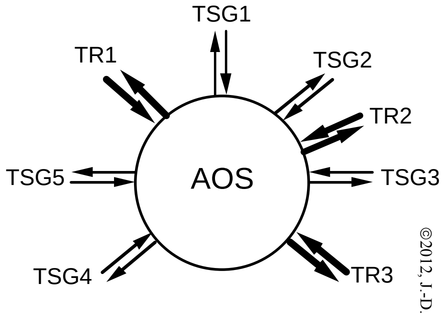
Coding

- balanced (bipolar) transmission
- MANCHESTER encoding

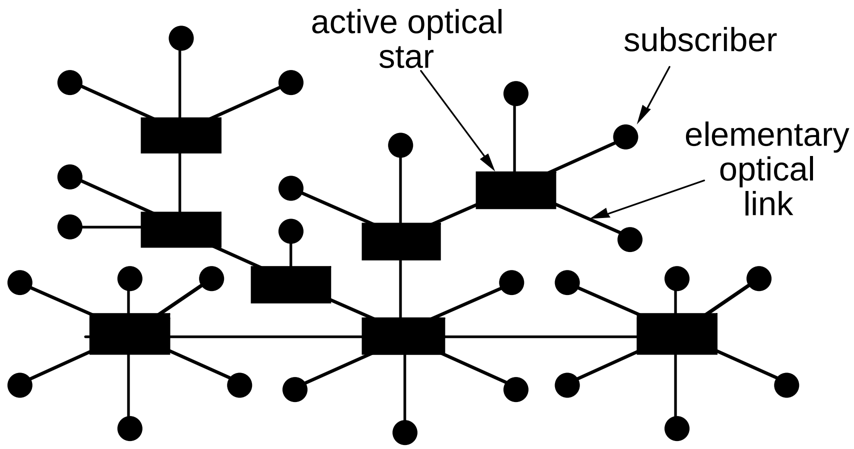


Optical physical layer

- identical to twisted pair one for:
  - Services, coding and transmission speeds
- restricted topology by exclusive use of active optical stars and point to point links
  - ⇒ special confirmation signal between 2 stars
- amplitude modulation of the optical signal (H symbol at higher level than B symbol)

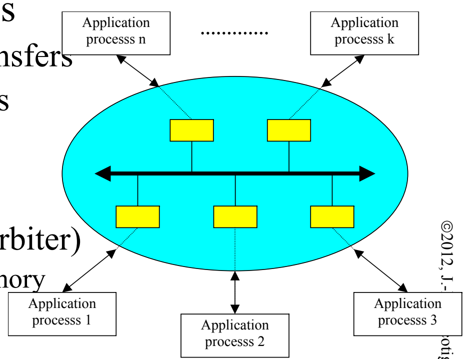


Topology using optical fibers



Data Link Layer

- makes use of cyclic nature of exchanges
  - nearly systematic use of broadcast
- centralized medium access control (bus arbiter)
- 2 types data transfer services
  - services used for variable transfers
  - services for message transfers
- 2 transfer types
  - on request
  - cyclic (triggered by the bus arbiter)
    - seems a distributed shared memory

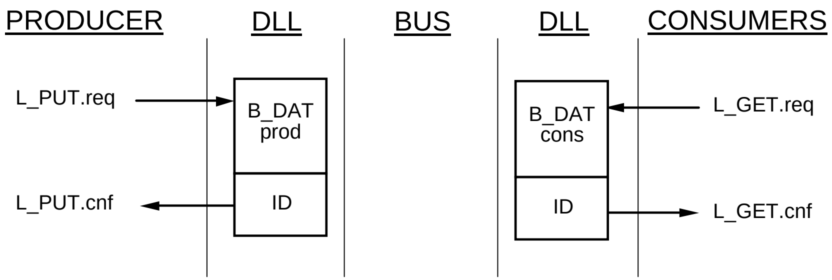


Addressing

- 2 different addressing spaces
  - For variables
    - identified by a unique identifier (16 bits) for each variable
    - address is not related to physical location
    - each variable producer and consumer(s) know the identifier of the variable
  - For messages
    - bear a source and a destination address (24 bits)
    - address is divided into a segment number (8 bits) and a LSAP (16 bits)

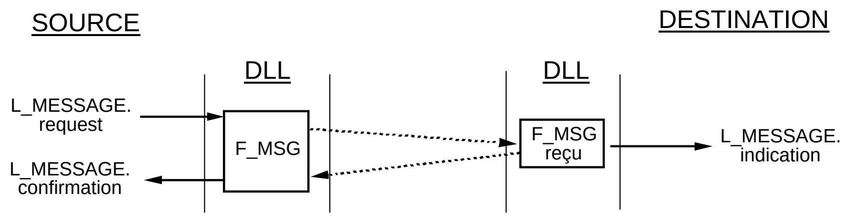
Variable transfers

- the DLL associates
  - a buffer (B_DATprod) to each identifier of produced variable
  - a buffer (B_DATcons) to each identifier of consumed variable



Message transfers

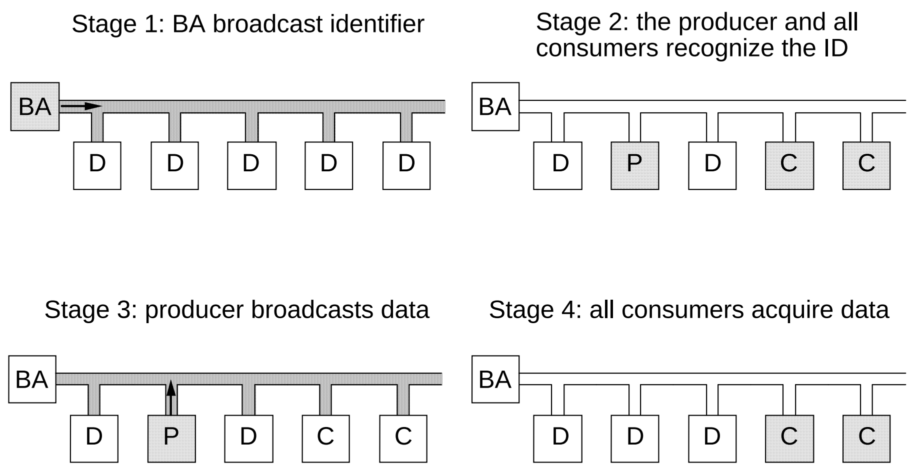
- the DLL associates
  - a queue (F_MSG) for the messages it needs to send
  - a queue (F_MSGreçu) for the messages that it receives



Behavior

- variables
  - writing (reading) a variable does not trigger directly any traffic on the network
  - a new write operation overwrites the previous variable value
  - reception and emission of a value is signaled
    - → flow control possible (not necessary)
- messages
  - stored in a queue → no overwrite
  - transfer is signaled → flow control
  - duplications are avoided using alternating bit protocol

Medium access control



Bus Arbiter (BA)

- control medium access
- a single one is active, however several may be in backup
- triggers all transactions
  - cyclic variable and message transfers
    - decided at configuration
    - configuration may be changed at run time
  - on-request variable and message transfers

> As a station may not emit without being invited to, each station must possess a mean to inform the Bus Arbiter that it requests a transfer

Signalling a request

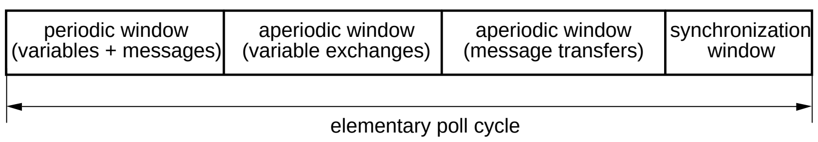
- uses the periodic transfer of variables
- in its response to a periodic variable transfer invitation (from the Bus Arbiter), the producer DLL will indicate its request

> a station that does not produce any periodic variable cannot explicitly require the transfer of a variable or a message

  - ⇒ message transfers are related to a variable identifier
- BA will later poll the requesting station to get the list of identifiers of the variables or messages for which a transfer is requested

Operating principle

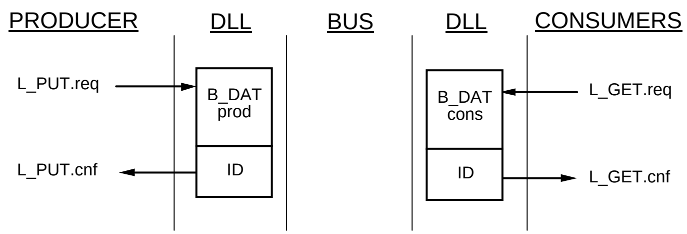
- according to configuration and explicit requests, the BA will establish a polling order



- the duration of the elementary polling cycle
  - is often constant (sampled data systems)
  - cannot exceed the duration of the smallest polling period requested by the users at configuration

Read / write buffer services

- L_PUT/GET.request(identifier, value)
- L_PUT/GET.confirmation(identifier, status)
  - status gives the result:
    - result OK, unknown identifier, identifier invalidated by Network management, buffer access conflict, data length non compatible with buffer

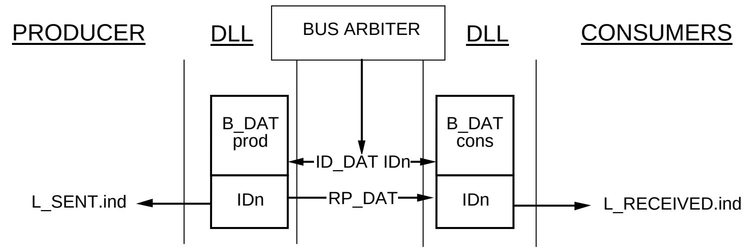


Buffer transfers

- 2 data units

ID_DAT	IDENTIFIER	FCS

RP_DAT	VARIABLE VALUE	FCS



Explicit buffer transfer requests

- request signaled when responding to a periodic buffer transfer request (ID_DAT)
- 2 types
  - specified explicit request (identifier specified by DLL user)
  - free explicit request (identifier chosen by DLL)
- requester may be neither producer nor consumer of the variable
- the same identifier may not be used for both types of requests. It must be configured for a type of service

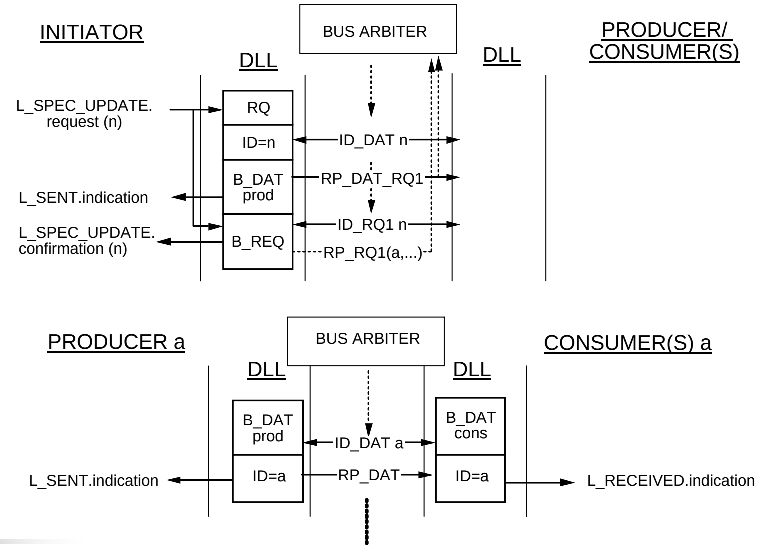
Explicit buffer transfer requests (2)

- BA later polls the requesting node for the list of identifiers of variables and messages for which a transfer is requested
- BA establishes a list of pollings (ID_MSG or ID_DAT) to fulfill the requests
  - If the same ID is requested while already requested, only one request is assumed

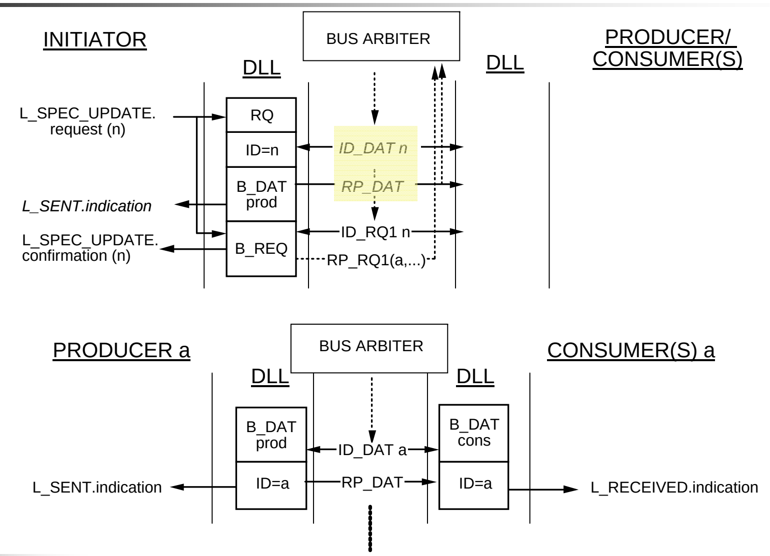
Explicit transfer requests

- specified explicit request
  - requests are not queued inside DLL (buffer B_REQ)
  - and are served either
    - immediately in the periodic window (RQ_INHIBE true)
    - or in an aperiodic window (RQ_INHIBE false) with urgent priority
- free explicit request
  - requests are queued (file F_REQi) according to priority (i=1 or 2)
  - are served according to the priority (1: urgent or 2: normal) in an aperiodic window

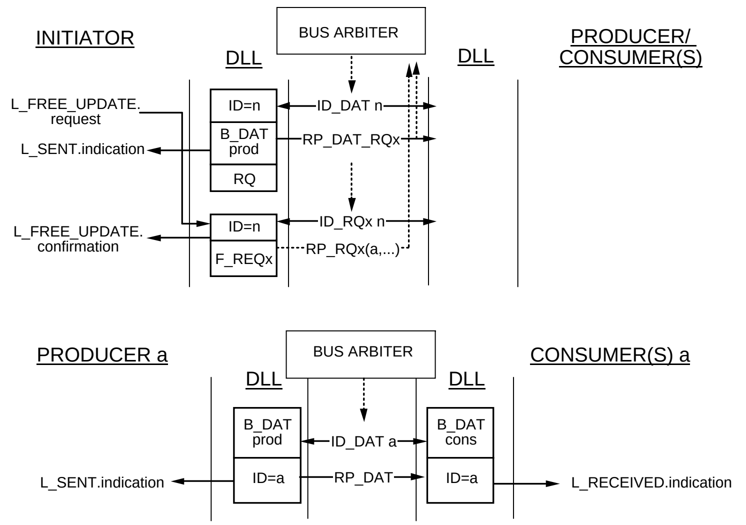
Specified explicit request (RQ_Inhibe false)



Specified explicit request (RQ_Inhibe true)



Free explicit requests



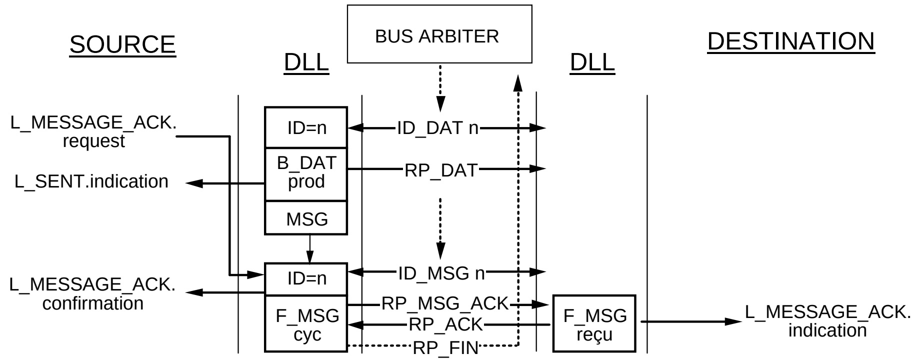
Message transfer services

- with (point to point only) or without acknowledge (protection against loss and duplication by alternating bit)
- requests are served in
  - the periodic window (cyclic msg transfers)
    - → resources allocated at configuration
  - the aperiodic window (aperiodic message transfers)
    - → requests queued in file F_MSGaper
- multisegment transfers
  - only point to point
  - go through a bridge (ack sent by bridge)

Cyclic message transfers

- use one or more variable identifiers configured for this type of transfer. The same Ids may be used for buffer transfer.
- a file (queue F_MSGcyc) is associated to each identifier
- messages stored in the F_MSGcyc queue associated to the identifier indicated in the request
- request served in the periodic window (even if no msg pending)
- transfer is performed with or without ack

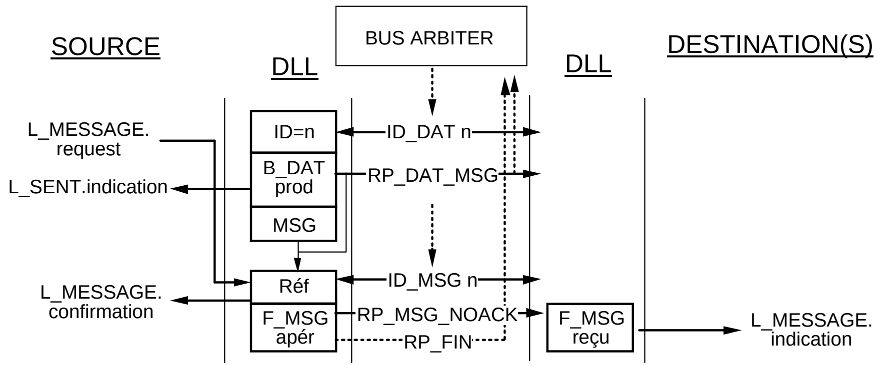
Cyclic message transfers (2)



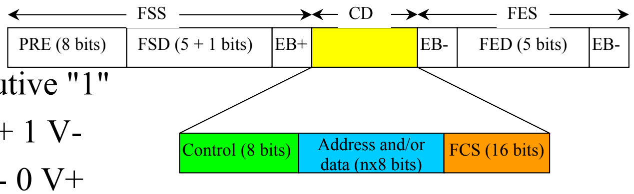
Aperiodic message transfers

- no identifier is indicated in the request
- identifier is chosen by the DLL
  - it must be configured for this type of service
  - it must correspond to a variable produced by the given station(source of message)
  - it must be already associated to a message in the queue
- the request is signaled to the bus arbiter in the response to an ID_DAT frame with the selected identifier
- association with the identifier is cut when the message has been transferred. The ID may then be used for another message in the queue

Aperiodic message transfer



Frame Format



- PRE = 8 consecutive "1"
- FSD = V+ V- V+ 1 V-
- FED = V- V+ V- 0 V+
- FCS= CRC 16 (CCITT V41) on CD field

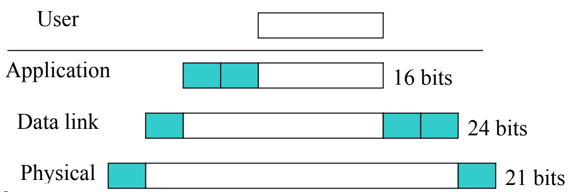
Frame Type	Data / address field (CD field)
identifier (ID_xyz)	Identifier (2 bytes)
Response request (RP_RQx)	List of identifier of requested var. transfers (≤ 64 ids)
Response message (RP_MSG)	Source + destin. Addr. (2x24 bits) + message (≤256B)
Response data (RP_DAT_z)	Data (up to 128 bytes)
Other frames	empty

Timers

Name	Localization	Function	Value
T0		Return time	
T1	Active BA	Absence of RPxx after IDxx	T0
T2	Active BA	Filling synchronization window	EC time
T3	Potential BAs	Absence of IDxx after RPxx	K*N*T0
T4	Consumer	no RP_DAT_xx after ID_DAT	T0
T5	Active BA	no RP_FIN after ID_MSG	2*T0
T6	Source of msg	no RP_ACK after RP_MSG_ACK	T0

Importance of timer T0

- T0 = TR Tmac with $10 \leq TR \leq 70$
- number of bits in a frame = 61 bits + 8 n
- transaction duration = 61 Tmac + (61+8n) Tmac + 2 T0



T0	n	Efficiency	Thoughput (@1Mbit/s)	Duration [µs]
10	2	10.1 %	101.3 Kbit/s	158
70	2	5.8%	57.6 Kbit/s	278
10	10	36%	360 Kbit/s	222
70	10	23%	234 Kbit/s	342

Spatio-temporal aspects

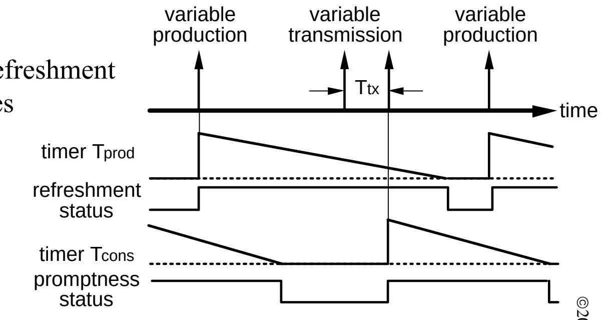
- for periodicity
  - centralized medium access control
  - mandatory maximum return time
- for validity duration
  - refreshment status
  - promptness status
- for temporal consistency
  - based upon refreshment and promptness statuses
- for spatial consistency
  - broadcast of consistency variable

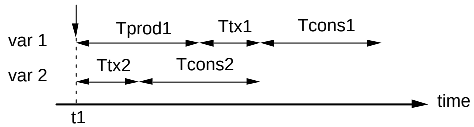
List concept

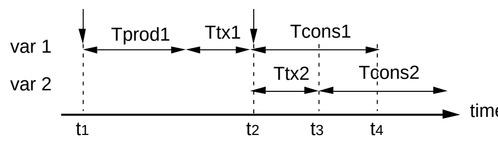
- ordered collection of consumed variables
- lists should not intersect in the system
- variables may be periodic or not
- periods may be different
- temporal consistency may be defined on a list

Temporal consistency

- asynchronous status
  - $Ttx + Tcons \leq Tvalidity \leq Ttx + Tprod + Tcons$
- temporal consistency
  - = logical AND of all refreshment and promptness statuses
    $\Rightarrow$ problems

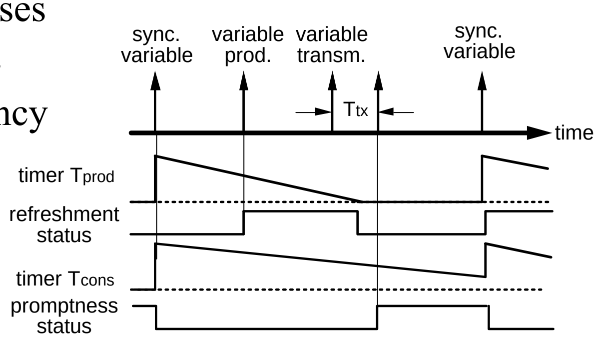






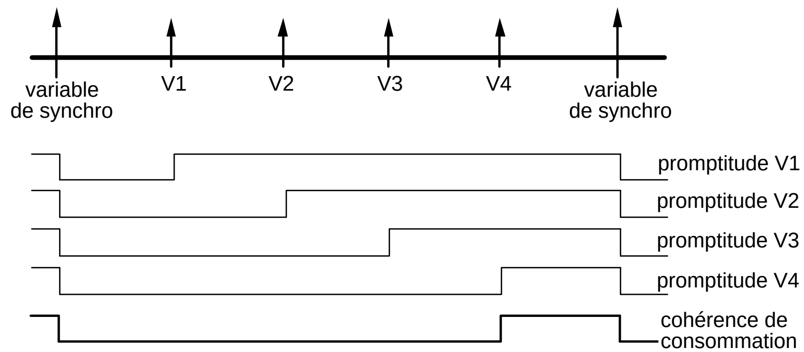
Temporal consistency (2)

- synchronous statuses
  - $Tvalidity = Tcons$
- temporal consistency
  - = logical AND of all refreshment and promptness statuses



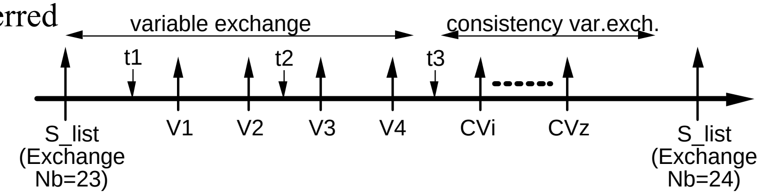
$\Rightarrow$ OK as long as the same sync. variable is used for all

Temporal consistency (3)



Spatial consistency

- may only be elaborated on lists
- based upon two types of variables
  - a synchronization variable
    - its value corresponds to the exchange Nb
  - a consistency variable for each consumer
    - bears the exchange Number / indicates for each variable in the list
  - either the exchange number of the available variable value (maintained detection) or the indication of reception in the current exchange cycle (instantaneous detection)
- at the end of variable transfers, the consistency variables are transferred



Temporal behaviour

- Central polling
  - Guaranteed transaction time (ID_xxx + RP_yyy) provided each transactions is bounded
    - Timer T0
  - Scheduling algorithm is not specified
- Periodic operations possible
  - Also with multiple periods
- No retransmission in case of errors
  - Except for messages with ack
  - Oversampling may be necessary to tolerate absence of a value in case of error (vacant sampling)

Temporal behaviour (2)

- Sporadic traffic
  - Best effort (except for specified explicit requests handled in periodic window)
  - Handling order not specified (may be FIFO)
  - 2 priorities
  - No admission control

Schedulability analysis

- Periodic traffic
  - Assumed to be guaranteed so worst case is to be considered
    - Variable transfers + non acknowledged messages
    - Acknowledged messages
- Aperiodic (sporadic) traffic
  - May be analysed in worst case or statistically
    - Statistics are interesting to take errors (retries) into account

Periodic traffic

- $P_i=\{C_p, D_i, T_i\}$; length, deadline, period
- Elementary cycle (EC): duration T_{mc}
  - $GCD(T_i)$ or using the cyclic executive principle [Bak89]
- Macrocyle = LCM (T_i)
- May be scheduled off line or on line
  - Enough time should be left to accomodate a single aperiodic transfer of the longuest duration T_{data}^{max}
  - Number of transfers that can be scheduled in an EC

$$N_p = \left\lfloor \frac{T_{mc} - T_{data}^{max}}{C_p} \right\rfloor$$

Transaction duration (@ 1 Mbit/s)

- Message with acknowledgement (K-1 retries)
  - Max length (254 byte message)

$$T_{data}^{max} = \{ID\_MSG + RP\_FIN\}T_{mac} + (2K+2)T_0$$
$$+ K(RP\_MSG\_ACK + RP\_ACK)T_{mac}$$
$$= \{61+45\} + 8 \bullet 70 + 3(2141+45) = 7224\mu s$$

  - Max length (8 byte message, 2 retries) = 1320µs
- Variable
  - Max. length (8 byte payload) = 61+125+2*70=326 µs

Schedulability of periodic traffic

- Assume N_p transactions per EC
- Priorities according to RM or DM
- If there are no gaps, the m-th transaction will be handled at: $t = \left\lfloor \frac{m-1}{N_p} \right\rfloor T_{mc} + \{(m \, MOD \, N_p) + 1\} C_p T_{mac}$
- The first transaction of the i-th periodic flow will be the m-th transaction where m is given by

$$m = 1 + \sum_{j \in hp(i)} \left\lceil \frac{t}{T_j} \right\rceil$$

Example [Tov99a]

- Example used for Profibus
- Assume high priority is made of synchronous messages
  - Period is equal to deadline

	Prod 1	Prod 2	Prod 3	Prod 4	Prod 5	Prod 6
$Dh_1(m)$	50 ms	90 ms	120 ms	60 ms	60 ms	80 ms
$Dh_2(m)$	100 ms	80 ms	130 ms	200 ms	100 ms	80 ms
$Dh_3(m)$		140 ms	110 ms	140 ms	100 ms	100 ms
C	1.32	1.32	1.32	1.32	1.32	1.32
nlp	3	3	3	3	3	3

Example (2)

- HCF= 10 ms
- LCM = 3603.600s
- Implementation using FIP messages (max. 2 retries)
  - Synchronous window= 4 ms (3 periodic messages per EC)
    - 1’081’080 messages in LCM (need is 1’007’279)
  - Deadlines (priorities according to DM)

	Prod 1	Prod 2	Prod 3	Prod 4	Prod 5	Prod 6
Dh_1(m)	50 /1.32	90 / 21.32	120 / 41.32	60 / 2.64	60 / 3.96	80 / 11.32
Dh_2(m)	100 / 22.64	80 / 12.64	130 / 42.64	200 / 53.96	100 / 23.96	80 / 13.96
Dh_3(m)		140 / 43.96	110 / 33.96	140 / 52.64	100 / 31.32	100/ 32.64
C	1.32	1.32	1.32	1.32	1.32	1.32

Response time analysis

- Asumptions [Ped97]
  - Periodic variables (no message)
  - Aperiodic (sporadic traffic) variables (no message)
    - Deadline ≤ period (no more than
  - Free explicit requests (2 priorities)
  - Requests are queued in FIFO order at BA
  - All periodic traffic has the same length C_p

$$C_p = (len(ID\_DAT) + len(RP\_DAT))T_{mac} + 2T_0$$

  - All aperiodic traffic has the same length C_a
  - Duration of periodic window is fixed T_s^{max}

Worst case

- Queueing time
  - When a variable (message) transfer request is queued in BA after all other variables and messages
- Dead interval (maximum interval between 2 pollings of a node) σ_k
  - When a request arrives at a node just after the periodic variable transfer took place

Temporal analysis

- Request at time t
- Will be signalled to BA at time t+σ_k at latest
- Identifier will be known to BA after the list will have been transferred
  - $\delta_{pa}(k)$ is the time to collect the list of identifiers for that station
- Pending aperiodic traffic must be transferred
- Periodic traffic is transferred in between

$$R_n = \sigma_k + \delta_{pa} + T_s^{max}\left\lceil \frac{R_n - \sigma_k}{T_{mc}} \right\rceil + \sum_{\forall j \in S_x} C_a$$

Response time parameters

$$R_n = \sigma_k + \delta_{pa} + T_s^{max} \left\lceil \frac{R_n - \sigma_k}{T_{mc}} \right\rceil + \sum_{\forall j \in S_x} C_a$$

- δ_{pa}

$$\delta_{pa} = n_s (ID\_RQ + RP\_RQ) T_{mac} + 2 n_s T_0$$

  - n_s number of stations that generate sporadic traffic
- S_x is the set of generated sporadic traffic between t and t+Rn
  - All urgent sporadic variables if we evaluate Rn for an urgent stream
  - All sporadic variables otherwise

Example 2

- 6 nodes, 3 streams each, 50 ms worst case interrival time (is also deadline)
- High priority traffic implemented as variable transfers with oversampling 3 times (all periods divided by 3)
- Buffer transfer (variable) time = 326 µs (appx 1/3 ms)
- EC duration T_{mc}= 10/ 3 ms; hyperperiod = 1201.2 s

	Prod 1	Prod 2	Prod 3	Prod 4	Prod 5	Prod 6
$Dh_1(m)$	16.67 /0.33	30 / 7	40 / 13.67	20 / 0.67	20 / 1	26.67 / 3.67
$Dh_2(m)$	33.3 / 7.33	26.67 / 4	43.3 / 14	66.7 / 17.67	33.3 / 7.67	26.67 / 4.33
$Dh_3(m)$		46.67 / 14.3	36.67 / 11	46.67 / 17.3	33.3 / 10.33	33.3/ 10.67
C	0.33	0.33	0.33	0.33	0.33	0.33

- Synchronous window: T_s^{max} = 1ms

Example 2 (2)

$$R_n = \sigma_k + \delta_{pa} + T_s^{max} \left\lceil \frac{R_n - \sigma_k}{T_{mc}} \right\rceil + \sum_{\forall j \in S_x} C_a$$

- δ_{pa} : 6 nodes, 1 Mbit/s, ID_RQ length= 61 bits
  - RP_RQ length = 61 + 16 * Nb_Identifiers = 109 bits
  - (61+109+2*70)*6=1860

	Prod 1	Prod 2	Prod 3	Prod 4	Prod 5	Prod 6
σ_k	16.67	26.67	36.67	20	20	26.67
δ_{pa}	1.86	1.86	1.86	1.86	1.86	1.86
R_n	53.3	63.3	73.3	56.5	56.6	63.3
C	1.32	1.32	1.32	1.32	1.32	1.32

Analysis

- Compared to Profibus
  - FIP suffers from the dead time for sporadic traffic response deadline
    - This is true only for traffic coming from a master station
  - Provided the token rotation time is set adequately, performances are comparable if we allow retries
  - FIP can provide periodicity and multiple periods

Summary

- FIP is interesting to handle periodic traffic
- Bounds can be derived for sporadic traffic
  - But are directly linked to polling periods
- FIP has mechanisms to indicate temporal validity and consistency
  - Validity may not be needed in the case of other solutions
- Thanks to broadcast source addressing, multicast is easily implemented
- Timer T0 proves to be critical in the performances

References

- [Alm02] L. Almeida et al., « Schedulability analysis of real-time traffic in WorldFIP networks: an integrated approach », IEEE Trans. On Industrial Electronics 49 (5), Oct. 2002, pp. 1165 - 1174 (interesting for its review of previous work and introduction to FIP)
- [Ped97] P. Pedro, A. Burns, « Worst case response time analysis of hard real-time sporadic traffic in FIP networks », 9th Euromicro Workshop on Real-Time Systems, 11-13 June 1997, pp. 3 - 10 (some misconceptions i.e. the idea of slots)
- Y. Song, « Performance analysis of periodic and aperiodic real-time message transmission in FIP networks »,
- J.-D. Decotignie, P. Prasad, "Spatio-Temporal Constraints in Fieldbus: Requirements and Current Solutions", 19th IFAC/IFIP Workshop on Real-Time Programming, Isle of Reichnau, June 22-24, pp.9-14, 1994.

References (2)

- [Bak89] T. Baker, A. Shaw, « The cyclic executive model and Ada », Real-time systems, vol. 1 (1), pp.7-25, 1989.

Real-Time Networks
WorldFIP

Prof. J.-D. Decotignie
CSEM Centre Suisse d'Electronique et de Microtechnique SA
Jaquet-Droz 1, 2007 Neuchâtel
jean-dominique.decotignie@csem.ch

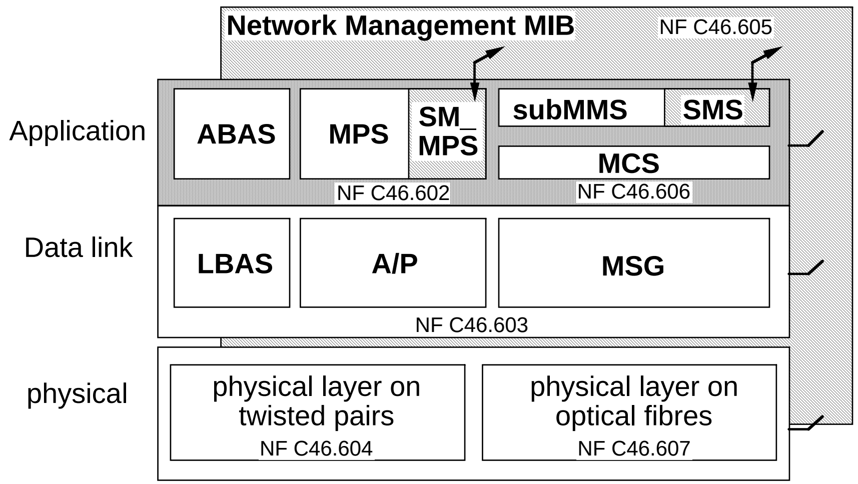
Outline

- WorldFIP history
- Architecture
- Physical layer
- Data link layer
  - Medium access control
  - Variable transfers
  - Message transfers
  - Consistency
- Temporal behavior
- Response time analysis

History

- 1984: White paper
  - D. Galara, J.-P. Thomesse, "Groupe de réflexion FIP: proposition d'un système de transmission série multiplexée pour les échanges entre des capteurs, des actionneurs et des automates réflexes", French Ministry of Industry, Paris, May 1984.
- 1989: French Standard (NF C46601-46607)
- 1996: European Standard (EN 50170)
- 2003: ISO/IEC standard (IEC 61158)

Architecture



Basic Choices

- Real-time traffic consists of state information transferred in a cyclic or periodic manner but not event transmission
- All information is broadcast
- For real-time traffic,
  - only variables are identified (identifiers). Network nodes or application processes are not identified
  - the communication model is the Producer-Distributor-Consumer (PDC) Model where variable values transferred on the network are neither queued at production nor at reception
  - There is no retransmission and no acknowledge for real-time traffic
  - Consumers are responsible for checking transmission status and taking appropriate actions when problems arise

Basic Choices (2)

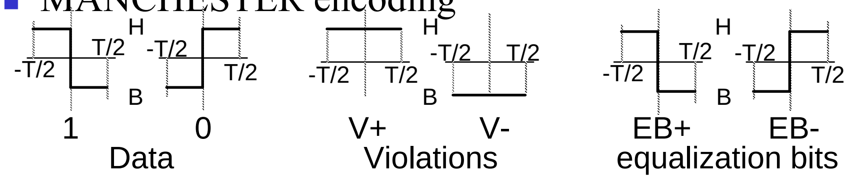
- All variables transferred in real-time and their characteristics (period, relation with other transfers) are known in advance
- MAC handled by a unique active distributor. Backup allowed
- The response time is bounded between 10 and 70 bit times.
- Transfers are either periodic or sporadic.
  - The transfer period for the values corresponding to a variable may differ from the period associated to another variable. The only restriction is that periods should be integer multiples of a basic period.
- Mechanisms to indicate temporal validity as well as temporal and spatial consistency of variable values built in the protocol.
- The non real-time traffic handled using the conventional client-server model and the corresponding application is a subset of MMS (Manufacturing Message Specification).

Physical layer

- twisted pairs (C46-604), optical fibers (C46-607)
- 3 speed classes
  - S1 (31.25 Kbits/s) / S2 (1 Mbits/s) / S3 (2.5 Mbits/s)
- Number of elements and range limited by
  - Propagation time Tp
    - normal case: Tp < 20 Tbit / extended: Tp < 40 Tbit
  - Losses and distortion on the cable
- 3 conformance classes
  - CH: high level class, allows long distance transmission at high data rates
  - CM: medium power class
  - CL: low power class for application with intrinsic safety requirements.
- recommended speeds
  - S1 or S2 for CL, S2 for CM, S2 or S3 for CH

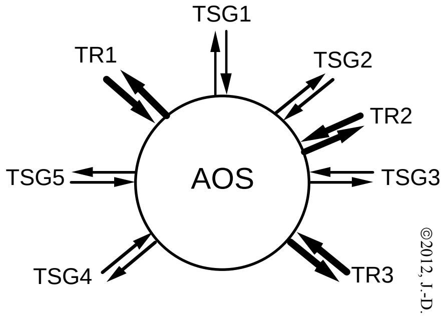
Coding

- balanced (bipolar) transmission
- MANCHESTER encoding

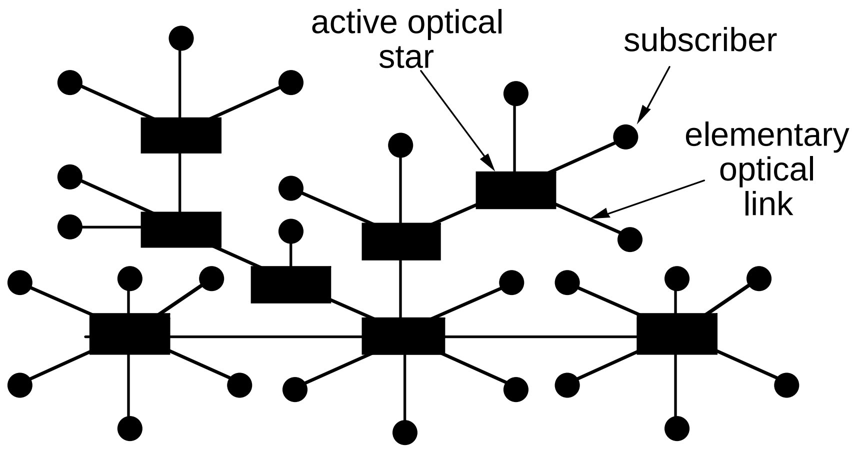


Optical physical layer

- identical to twisted pair one for:
  - Services, coding and transmission speeds
- restricted topology by exclusive use of active optical stars and point to point links
  - ⇒ special confirmation signal between 2 stars
- amplitude modulation of the optical signal (H symbol at higher level than B symbol)

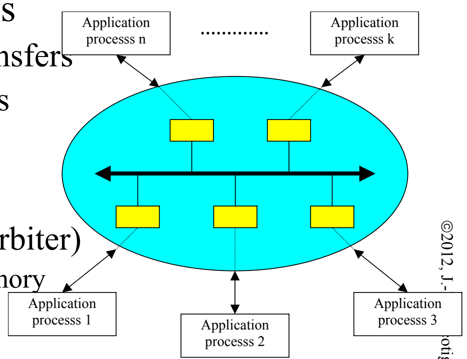


Topology using optical fibers



Data Link Layer

- makes use of cyclic nature of exchanges
  - nearly systematic use of broadcast
- centralized medium access control (bus arbiter)
- 2 types data transfer services
  - services used for variable transfers
  - services for message transfers
- 2 transfer types
  - on request
  - cyclic (triggered by the bus arbiter)
    - seems a distributed shared memory

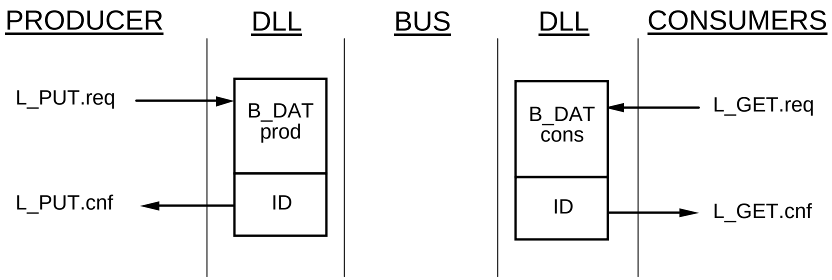


Addressing

- 2 different addressing spaces
  - For variables
    - identified by a unique identifier (16 bits) for each variable
    - address is not related to physical location
    - each variable producer and consumer(s) know the identifier of the variable
  - For messages
    - bear a source and a destination address (24 bits)
    - address is divided into a segment number (8 bits) and a LSAP (16 bits)

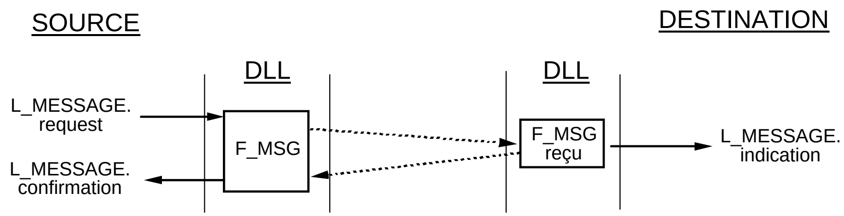
Variable transfers

- the DLL associates
  - a buffer (B_DATprod) to each identifier of produced variable
  - a buffer (B_DATcons) to each identifier of consumed variable



Message transfers

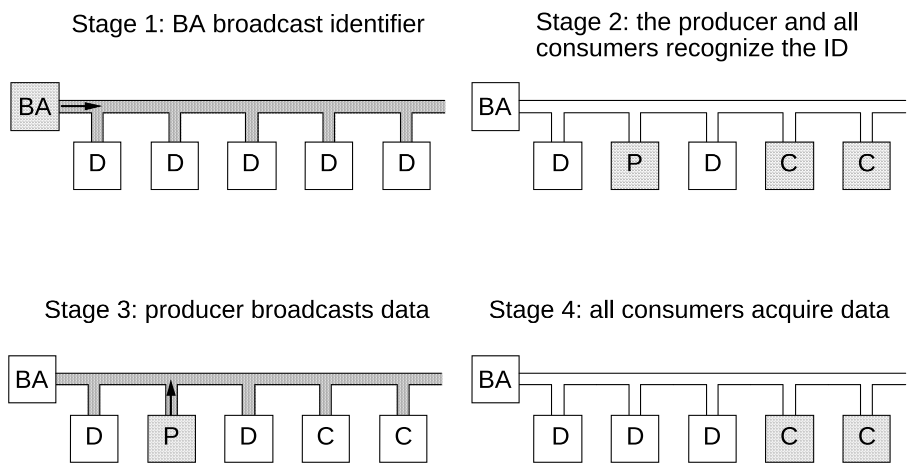
- the DLL associates
  - a queue (F_MSG) for the messages it needs to send
  - a queue (F_MSGreçu) for the messages that it receives



Behavior

- variables
  - writing (reading) a variable does not trigger directly any traffic on the network
  - a new write operation overwrites the previous variable value
  - reception and emission of a value is signaled
    - → flow control possible (not necessary)
- messages
  - stored in a queue → no overwrite
  - transfer is signaled → flow control
  - duplications are avoided using alternating bit protocol

Medium access control



Bus Arbiter (BA)

- control medium access
- a single one is active, however several may be in backup
- triggers all transactions
  - cyclic variable and message transfers
    - decided at configuration
    - configuration may be changed at run time
  - on-request variable and message transfers

As a station may not emit without being invited to, each station must possess a mean to inform the Bus Arbiter that it requests a transfer

Signalling a request

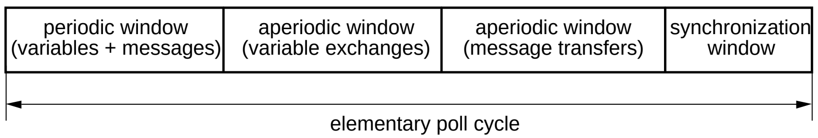
- uses the periodic transfer of variables
- in its response to a periodic variable transfer invitation (from the Bus Arbiter), the producer DLL will indicate its request

a station that does not produce any periodic variable cannot explicitly require the transfer of a variable or a message

  - ⇒ message transfers are related to a variable identifier
- BA will later poll the requesting station to get the list of identifiers of the variables or messages for which a transfer is requested

Operating principle

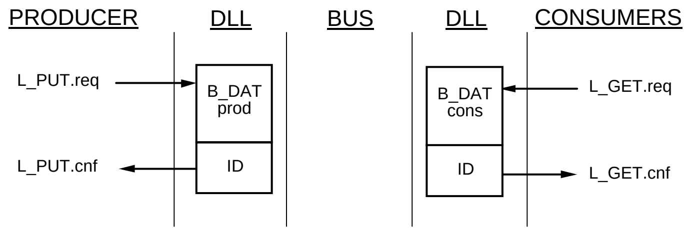
- according to configuration and explicit requests, the BA will establish a polling order



- the duration of the elementary polling cycle
  - is often constant (sampled data systems)
  - cannot exceed the duration of the smallest polling period requested by the users at configuration

Read / write buffer services

- L_PUT/GET.request(identifier, value)
- L_PUT/GET.confirmation(identifier, status)
  - status gives the result:
    - result OK, unknown identifier, identifier invalidated by Network management, buffer access conflict, data length non compatible with buffer

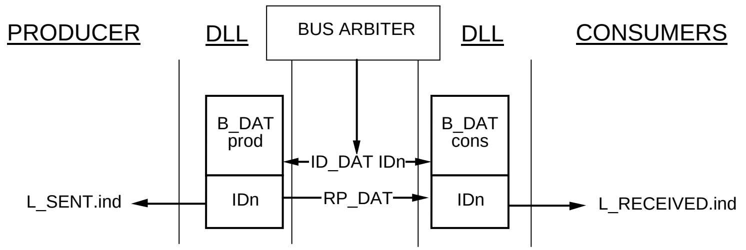


Buffer transfers

- 2 data units

ID_DAT	IDENTIFIER	FCS

RP_DAT	VARIABLE VALUE	FCS



Explicit buffer transfer requests

- request signaled when responding to a periodic buffer transfer request (ID_DAT)
- 2 types
  - specified explicit request (identifier specified by DLL user)
  - free explicit request (identifier chosen by DLL)
- requester may be neither producer nor consumer of the variable
- the same identifier may not be used for both types of requests. It must be configured for a type of service

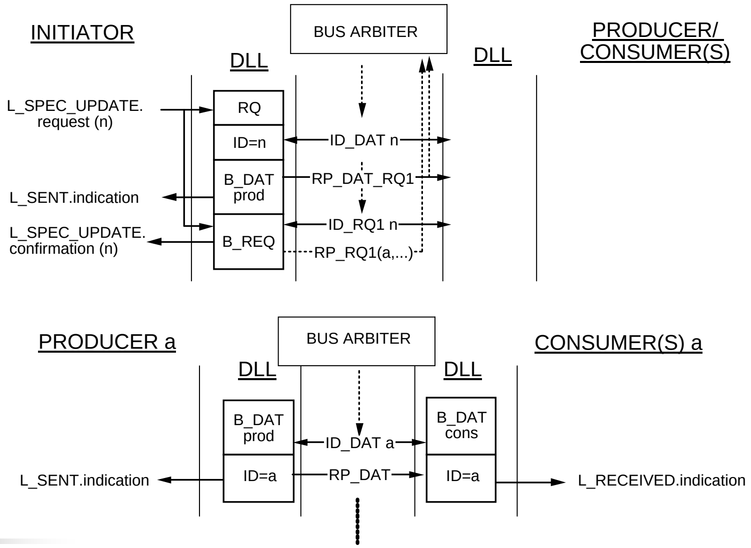
Explicit buffer transfer requests (2)

- BA later polls the requesting node for the list of identifiers of variables and messages for which a transfer is requested
- BA establishes a list of pollings (ID_MSG or ID_DAT) to fulfill the requests
  - If the same ID is requested while already requested, only one request is assumed

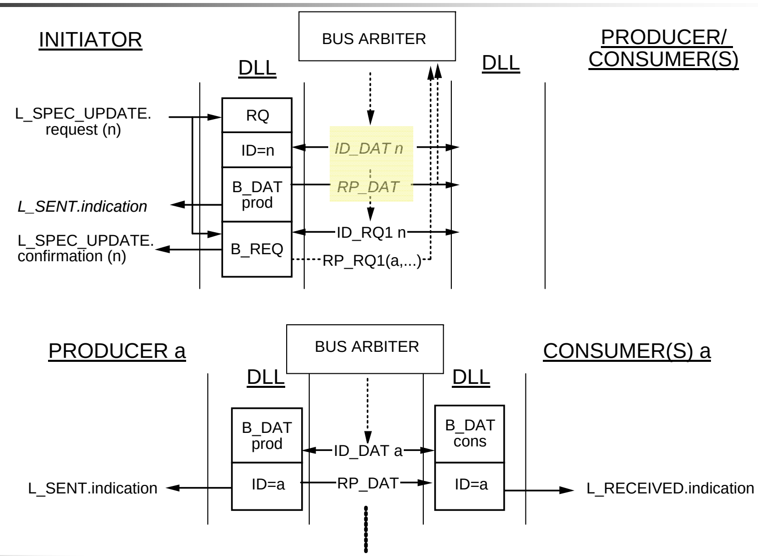
Explicit transfer requests

- specified explicit request
  - requests are not queued inside DLL (buffer B_REQ)
  - and are served either
    - immediately in the periodic window (RQ_INHIBE true)
    - or in an aperiodic window (RQ_INHIBE false) with urgent priority
- free explicit request
  - requests are queued (file F_REQi) according to priority (i=1 or 2)
  - are served according to the priority (1: urgent or 2: normal) in an aperiodic window

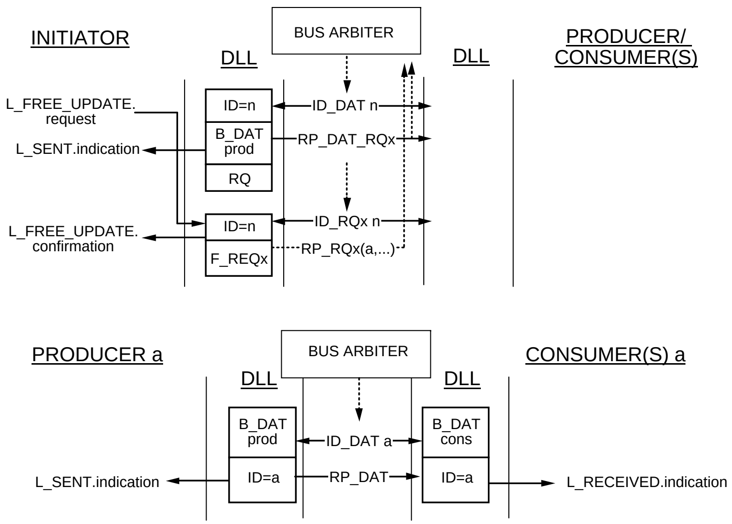
Specified explicit request (RQ_Inhibe false)



Specified explicit request (RQ_Inhibe true)



Free explicit requests



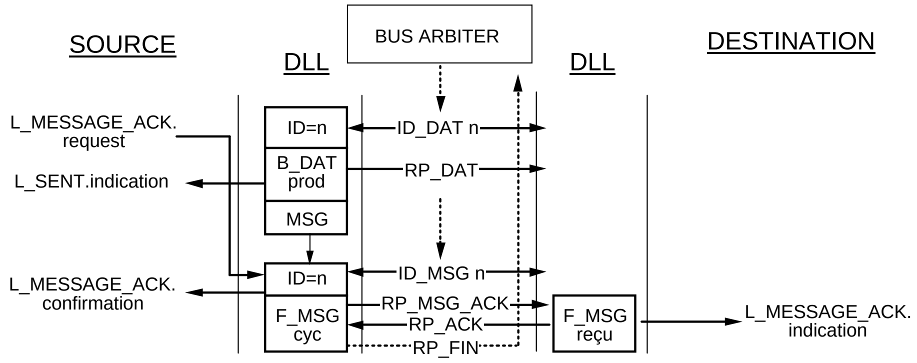
Message transfer services

- with (point to point only) or without acknowledge (protection against loss and duplication by alternating bit)
- requests are served in
  - the periodic window (cyclic msg transfers)
    - → resources allocated at configuration
  - the aperiodic window (aperiodic message transfers)
    - → requests queued in file F_MSGaper
- multisegment transfers
  - only point to point
  - go through a bridge (ack sent by bridge)

Cyclic message transfers

- use one or more variable identifiers configured for this type of transfer. The same Ids may be used for buffer transfer.
- a file (queue F_MSGcyc) is associated to each identifier
- messages stored in the F_MSGcyc queue associated to the identifier indicated in the request
- request served in the periodic window (even if no msg pending)
- transfer is performed with or without ack

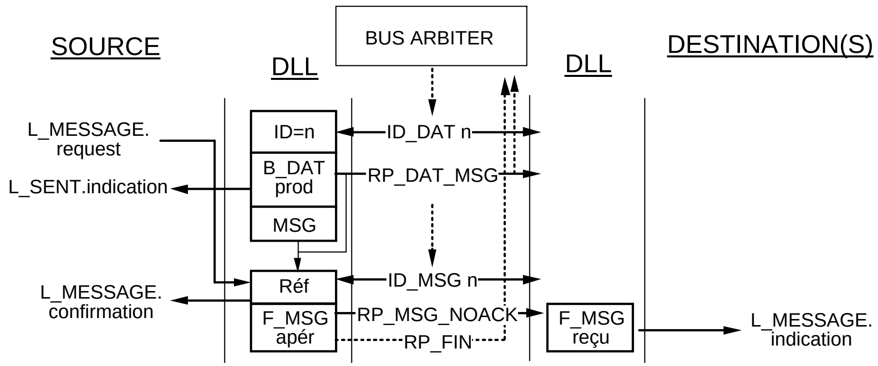
Cyclic message transfers (2)



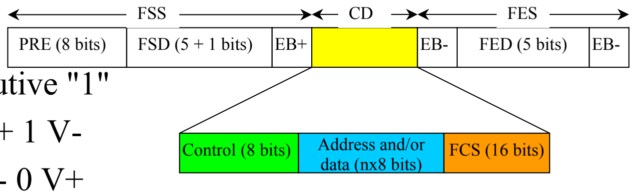
Aperiodic message transfers

- no identifier is indicated in the request
- identifier is chosen by the DLL
  - it must be configured for this type of service
  - it must correspond to a variable produced by the given station(source of message)
  - it must be already associated to a message in the queue
- the request is signaled to the bus arbiter in the response to an ID_DAT frame with the selected identifier
- association with the identifier is cut when the message has been transferred. The ID may then be used for another message in the queue

Aperiodic message transfer



Frame Format



- PRE = 8 consecutive "1"
- FSD = V+ V- V+ 1 V-
- FED = V- V+ V- 0 V+
- FCS= CRC 16 (CCITT V41) on CD field

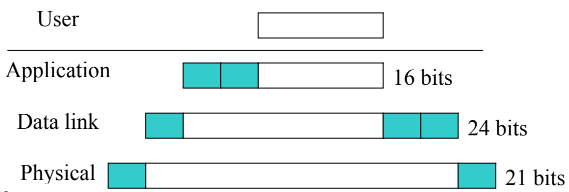
Frame Type	Data / address field (CD field)
identifier (ID_xyz)	Identifier (2 bytes)
Response request (RP_RQx)	List of identifier of requested var. transfers (≤ 64 ids)
Response message (RP_MSG)	Source + destin. Addr. (2x24 bits) + message (≤256B)
Response data (RP_DAT_z)	Data (up to 128 bytes)
Other frames	empty

Timers

Name	Localization	Function	Value
T0		Return time	
T1	Active BA	Absence of RPxx after IDxx	T0
T2	Active BA	Filling synchronization window	EC time
T3	Potential BAs	Absence of IDxx after RPxx	K*N*T0
T4	Consumer	no RP_DAT_xx after ID_DAT	T0
T5	Active BA	no RP_FIN after ID_MSG	2*T0
T6	Source of msg	no RP_ACK after RP_MSG_ACK	T0

Importance of timer T0

- T0 = TR Tmac with $10 \leq TR \leq 70$
- number of bits in a frame = 61 bits + 8 n



- transaction duration = 61 Tmac + (61+8n) Tmac + 2 T0

T0	n	Efficiency	Thoughput (@1Mbit/s)	Duration [µs]
10	2	10.1 %	101.3 Kbit/s	158
70	2	5.8%	57.6 Kbit/s	278
10	10	36%	360 Kbit/s	222
70	10	23%	234 Kbit/s	342

Spatio-temporal aspects

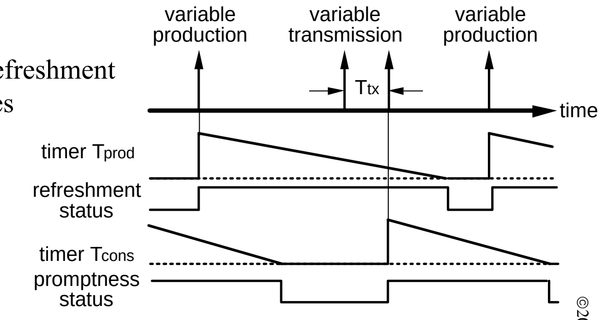
- for periodicity
  - centralized medium access control
  - mandatory maximum return time
- for validity duration
  - refreshment status
  - promptness status
- for temporal consistency
  - based upon refreshment and promptness statuses
- for spatial consistency
  - broadcast of consistency variable

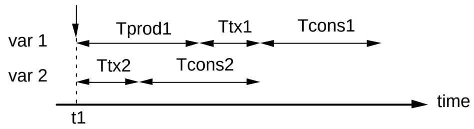
List concept

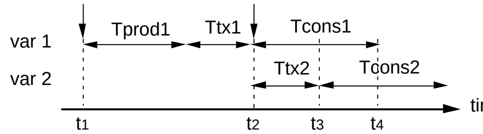
- ordered collection of consumed variables
- lists should not intersect in the system
- variables may be periodic or not
- periods may be different
- temporal consistency may be defined on a list

Temporal consistency

- asynchronous status
  - Ttx + Tcons ≤ Tvalidity ≤ Ttx + Tprod + Tcons
- temporal consistency
  - = logical AND of all refreshment and promptness statuses ⇒ problems

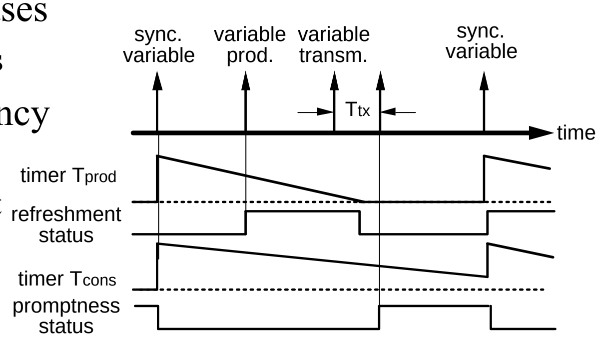






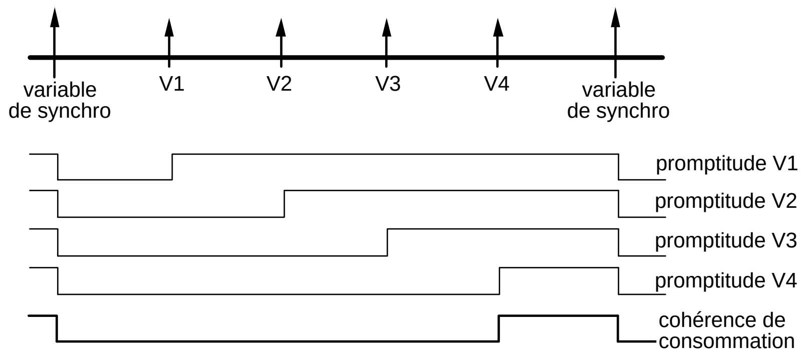
Temporal consistency (2)

- synchronous statuses
  - Tvalidity = Tcons
- temporal consistency
  - = logical AND of all refreshment and promptness statuses



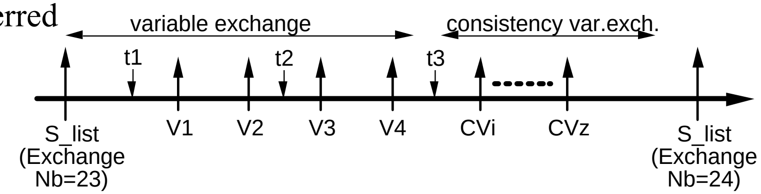
⇒ OK as long as the same sync. variable is used for all

Temporal consistency (3)



Spatial consistency

- may only be elaborated on lists
- based upon two types of variables
  - a synchronization variable
    - its value corresponds to the exchange Nb
  - a consistency variable for each consumer
    - bears the exchange Number / indicates for each variable in the list
  - either the exchange number of the available variable value (maintained detection) or the indication of reception in the current exchange cycle (instantaneous detection)
- at the end of variable transfers, the consistency variables are transferred



Temporal behaviour

- Central polling
  - Guaranteed transaction time (ID_xxx + RP_yyy) provided each transactions is bounded
    - Timer T0
  - Scheduling algorithm is not specified
- Periodic operations possible
  - Also with multiple periods
- No retransmission in case of errors
  - Except for messages with ack
  - Oversampling may be necessary to tolerate absence of a value in case of error (vacant sampling)

Temporal behaviour (2)

- Sporadic traffic
  - Best effort (except for specified explicit requests handled in periodic window)
  - Handling order not specified (may be FIFO)
  - 2 priorities
  - No admission control

Schedulability analysis

- Periodic traffic
  - Assumed to be guaranteed so worst case is to be considered
    - Variable transfers + non acknowledged messages
    - Acknowledged messages
- Aperiodic (sporadic) traffic
  - May be analysed in worst case or statistically
    - Statistics are interesting to take errors (retries) into account

Periodic traffic

- $P_i=\{C_p, D_i, T_i\}$; length, deadline, period
- Elementary cycle (EC): duration T_{mc}
  - $GCD(T_i)$ or using the cyclic executive principle [Bak89]
- Macrocyle = LCM (T_i)
- May be scheduled off line or on line
  - Enough time should be left to accomodate a single aperiodic transfer of the longuest duration T_{data}^{max}
  - Number of transfers that can be scheduled in an EC

$$N_p = \left\lfloor \frac{T_{mc} - T_{data}^{max}}{C_p} \right\rfloor$$

Transaction duration (@ 1 Mbit/s)

- Message with acknowledgement (K-1 retries)
  - Max length (254 byte message)

$$\begin{aligned} T_{data}^{max} &= \{ID\_MSG + RP\_FIN\}T_{mac} + (2K+2)T_0 \\ &\quad + K(RP\_MSG\_ACK + RP\_ACK)T_{mac} \\ &= \{61+45\} + 8 \bullet 70 + 3(2141+45) = 7224\mu s \end{aligned}$$

  - Max length (8 byte message, 2 retries) = 1320µs
- Variable
  - Max. length (8 byte payload) = 61+125+2*70=326 µs

Schedulability of periodic traffic

- Assume N_p transactions per EC
- Priorities according to RM or DM
- If there are no gaps, the m-th transaction will be handled at: $t = \left\lfloor \frac{m-1}{N_p} \right\rfloor T_{mc} + \{(m\ MOD\ N_p) + 1\} C_p T_{mac}$
- The first transaction of the i-th periodic flow will be the m-th transaction where m is given by

$$m = 1 + \sum_{j \in hp(i)} \left\lceil \frac{t}{T_j} \right\rceil$$

Example [Tov99a]

- Example used for Profibus
- Assume high priority is made of synchronous messages
  - Period is equal to deadline

	Prod 1	Prod 2	Prod 3	Prod 4	Prod 5	Prod 6
$Dh_1(m)$	50 ms	90 ms	120 ms	60 ms	60 ms	80 ms
$Dh_2(m)$	100 ms	80 ms	130 ms	200 ms	100 ms	80 ms
$Dh_3(m)$		140 ms	110 ms	140 ms	100 ms	100 ms
C	1.32	1.32	1.32	1.32	1.32	1.32
nlp	3	3	3	3	3	3

Example (2)

- HCF= 10 ms
- LCM = 3603.600s
- Implementation using FIP messages (max. 2 retries)
  - Synchronous window= 4 ms (3 periodic messages per EC)
    - 1’081’080 messages in LCM (need is 1’007’279)
  - Deadlines (priorities according to DM)

	Prod 1	Prod 2	Prod 3	Prod 4	Prod 5	Prod 6
Dh_1(m)	50 /1.32	90 / 21.32	120 / 41.32	60 / 2.64	60 / 3.96	80 / 11.32
Dh_2(m)	100 / 22.64	80 / 12.64	130 / 42.64	200 / 53.96	100 / 23.96	80 / 13.96
Dh_3(m)		140 / 43.96	110 / 33.96	140 / 52.64	100 / 31.32	100/ 32.64
C	1.32	1.32	1.32	1.32	1.32	1.32

Response time analysis

- Asumptions [Ped97]
  - Periodic variables (no message)
  - Aperiodic (sporadic traffic) variables (no message)
    - Deadline ≤ period (no more than
  - Free explicit requests (2 priorities)
  - Requests are queued in FIFO order at BA
  - All periodic traffic has the same length C_p

$$C_p = (len(ID\_DAT) + len(RP\_DAT))T_{mac} + 2T_0$$

  - All aperiodic traffic has the same length C_a
  - Duration of periodic window is fixed T_s^{max}

Worst case

- Queueing time
  - When a variable (message) transfer request is queued in BA after all other variables and messages
- Dead interval (maximum interval between 2 pollings of a node) σ_k
  - When a request arrives at a node just after the periodic variable transfer took place

Temporal analysis

- Request at time t
- Will be signalled to BA at time t+σ_k at latest
- Identifier will be known to BA after the list will have been transferred
  - δ_{pa}(k) is the time to collect the list of identifiers for that station
- Pending aperiodic traffic must be transferred
- Periodic traffic is transferred in between

$$R_n = \sigma_k + \delta_{pa} + T_s^{max} \left\lceil \frac{R_n - \sigma_k}{T_{mc}} \right\rceil + \sum_{\forall j \in S_x} C_a$$

Response time parameters

$$R_n = \sigma_k + \delta_{pa} + T_s^{\max}\left\lceil \frac{R_n - \sigma_k}{T_{mc}} \right\rceil + \sum_{\forall j \in S_x} C_a$$

- δ_{pa}

$$\delta_{pa} = n_s (ID\_RQ + RP\_RQ) T_{mac} + 2 n_s T_0$$

  - n_s number of stations that generate sporadic traffic
- S_x is the set of generated sporadic traffic between t and t+Rn
  - All urgent sporadic variables if we evaluate Rn for an urgent stream
  - All sporadic variables otherwise

Example 2

- 6 nodes, 3 streams each, 50 ms worst case interrival time (is also deadline)
- High priority traffic implemented as variable transfers with oversampling 3 times (all periods divided by 3)
- Buffer transfer (variable) time = 326 µs (appx 1/3 ms)
- EC duration T_{mc}= 10/ 3 ms; hyperperiod = 1201.2 s

	Prod 1	Prod 2	Prod 3	Prod 4	Prod 5	Prod 6
$Dh_1(m)$	16.67 /0.33	30 / 7	40 / 13.67	20 / 0.67	20 / 1	26.67 / 3.67
$Dh_2(m)$	33.3 / 7.33	26.67 / 4	43.3 / 14	66.7 / 17.67	33.3 / 7.67	26.67 / 4.33
$Dh_3(m)$		46.67 / 14.3	36.67 / 11	46.67 / 17.3	33.3 / 10.33	33.3/ 10.67
C	0.33	0.33	0.33	0.33	0.33	0.33

- Synchronous window: $T_s^{\max}$ = 1ms

Example 2 (2)

$$R_n = \sigma_k + \delta_{pa} + T_s^{\max}\left\lceil \frac{R_n - \sigma_k}{T_{mc}} \right\rceil + \sum_{\forall j \in S_x} C_a$$

- δ_{pa} : 6 nodes, 1 Mbit/s, ID_RQ length= 61 bits
  - RP_RQ length = 61 + 16 * Nb_Identifiers = 109 bits
  - (61+109+2*70)*6=1860

	Prod 1	Prod 2	Prod 3	Prod 4	Prod 5	Prod 6
σ_k	16.67	26.67	36.67	20	20	26.67
δ_{pa}	1.86	1.86	1.86	1.86	1.86	1.86
R_n	53.3	63.3	73.3	56.5	56.6	63.3
C	1.32	1.32	1.32	1.32	1.32	1.32

Analysis

- Compared to Profibus
  - FIP suffers from the dead time for sporadic traffic response deadline
    - This is true only for traffic coming from a master station
  - Provided the token rotation time is set adequately, performances are comparable if we allow retries
  - FIP can provide periodicity and multiple periods

Summary

- FIP is interesting to handle periodic traffic
- Bounds can be derived for sporadic traffic
  - But are directly linked to polling periods
- FIP has mechanisms to indicate temporal validity and consistency
  - Validity may not be needed in the case of other solutions
- Thanks to broadcast source addressing, multicast is easily implemented
- Timer T0 proves to be critical in the performances

References

- [Alm02] L. Almeida et al., « Schedulability analysis of real-time traffic in WorldFIP networks: an integrated approach », IEEE Trans. On Industrial Electronics 49 (5), Oct. 2002, pp. 1165 - 1174 (interesting for its review of previous work and introduction to FIP)
- [Ped97] P. Pedro, A. Burns, « Worst case response time analysis of hard real-time sporadic traffic in FIP networks », 9th Euromicro Workshop on Real-Time Systems, 11-13 June 1997, pp. 3 - 10 (some misconceptions i.e. the idea of slots)
- Y. Song, « Performance analysis of periodic and aperiodic real-time message transmission in FIP networks »,
- J.-D. Decotignie, P. Prasad, "Spatio-Temporal Constraints in Fieldbus: Requirements and Current Solutions", 19th IFAC/IFIP Workshop on Real-Time Programming, Isle of Reichnau, June 22-24, pp.9-14, 1994.

References (2)

- [Bak89] T. Baker, A. Shaw, « The cyclic executive model and Ada », Real-time systems, vol. 1 (1), pp.7-25, 1989.

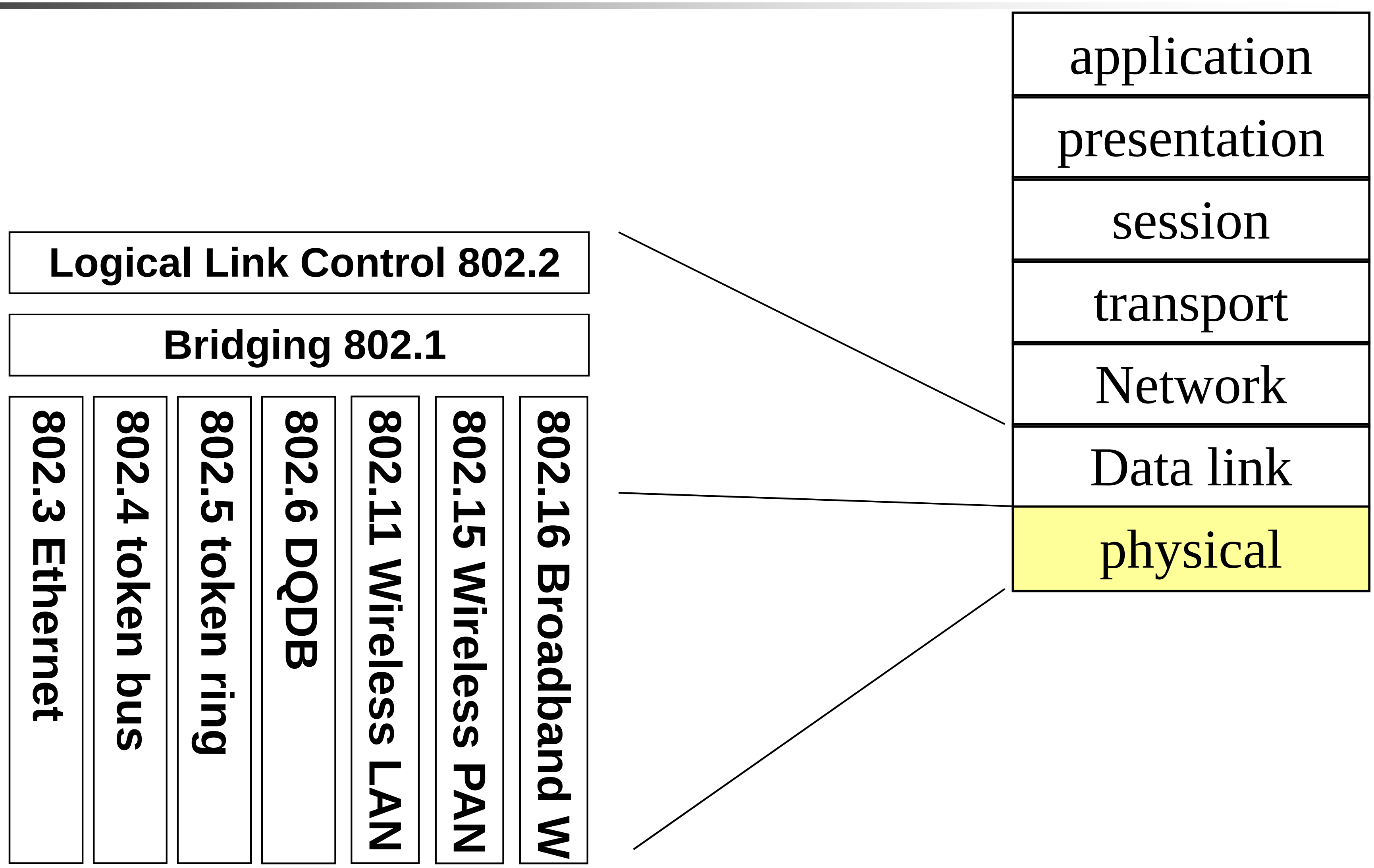
REAL-TIME NETWORKS
Ethernet

Prof. J.-D. Decotignie
CSEM Centre Suisse d'Electronique et de Microtechnique SA
Jaquet-Droz 1, 2007 Neuchâtel
jean-dominique.decotignie@csem.ch

Outline

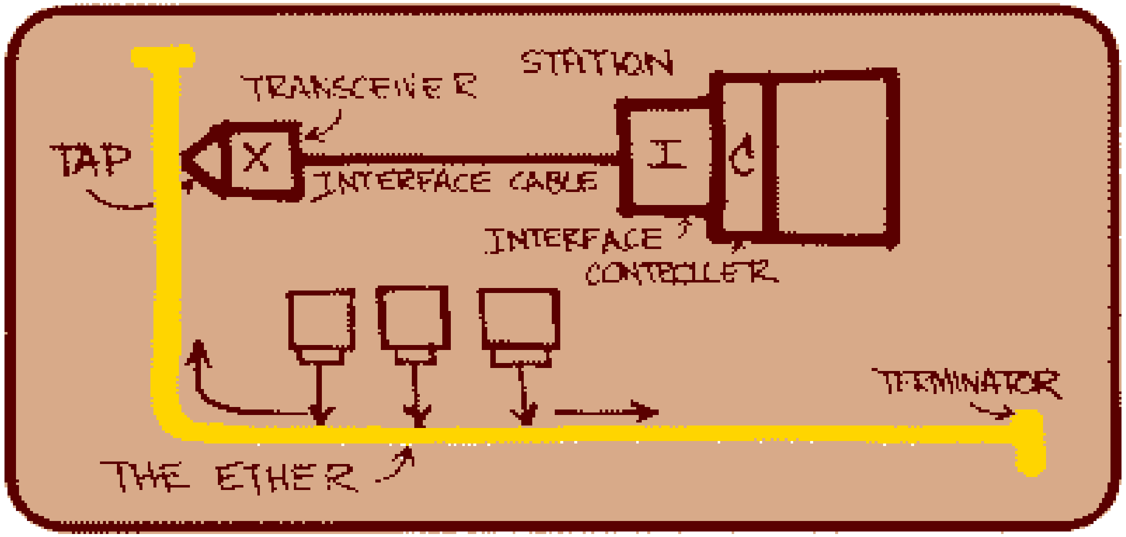
- The good old Ethernet
- Myths and realities
- Companion protocols
- Improvements
- Switched Ethernet
- How to improve temporal behaviour
- Industrial Ethernet
- Experiments on Switched Ethernet
- Conclusion

IEEE 802 standards



« Vintage Ethernet »

- One segment of coaxial cable
- All nodes are hooked to the same cable
- A medium access control
  - Every station may listen what is emitted on the bus (cable segment)
  - Each station may listen to what it emits
- Packet transmission



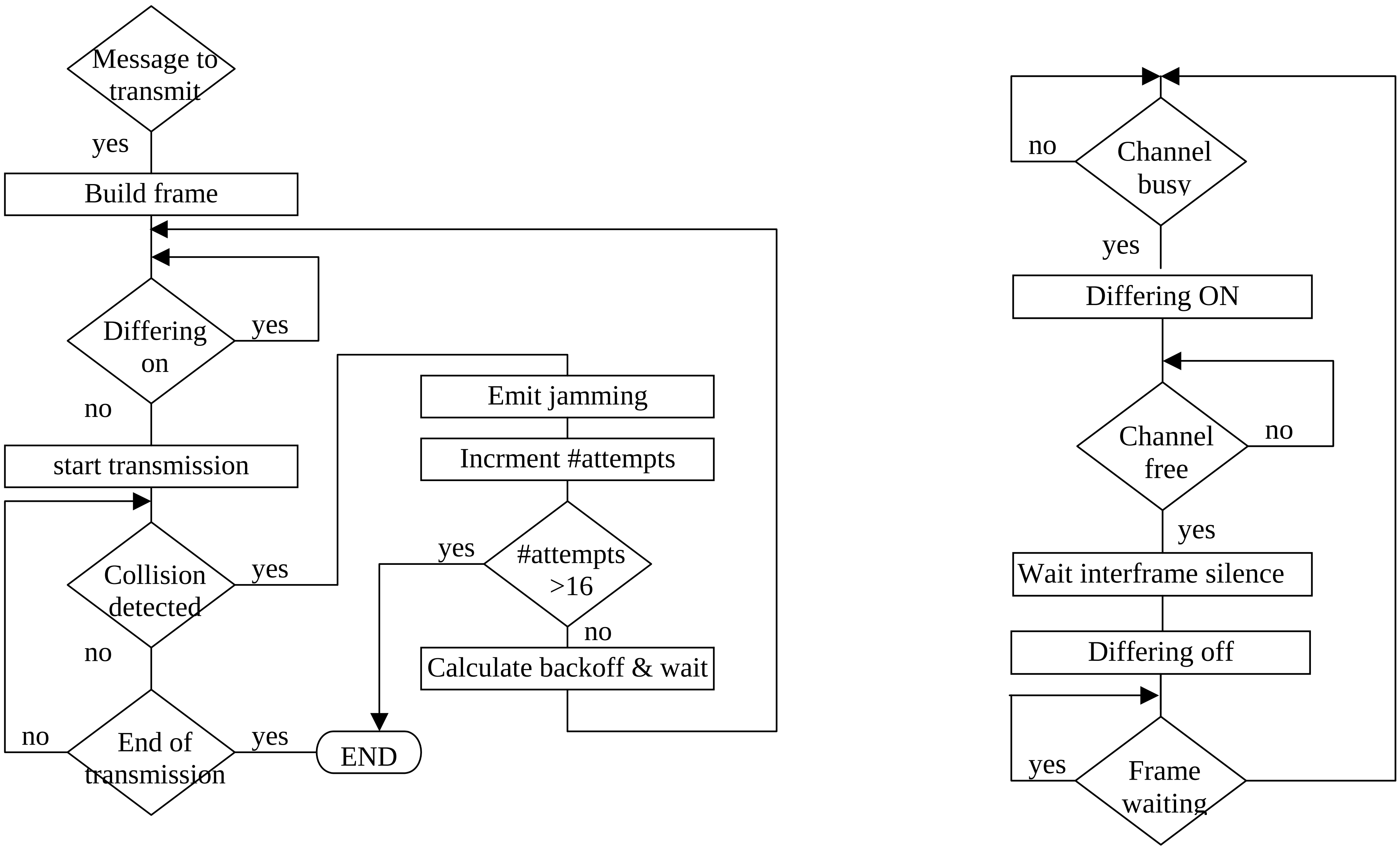
Basic principles - emission

- A station that wants to transmit
  - Listen the transmission medium
  - If free during a given duration
    - Emits its message
    - While listening what is sensed on the cable
    - If there is a collision, it sends a jamming sequence et prepares to retry
  - If the medium is busy, the station waits until it becomes free and does as indicated above
    - 1-persistent CSMA/CD

Basic principles – retry in case of collision

- Not done immediately
- A number is chosen randomly in the backoff interval
- This number is multiplied by the slot duration (5µs à 100Mbit/s)
- Starts to transmit after the given duration (if medium is busy)
- If there is a collision, the backoff window is doubled
  - Maximum of 0..1023
  - After 16 attempts, transmission is cancelled
- When transmission is successful, the backoff interval returns to its original value [0..1]

CSMA/CD for Ethernet

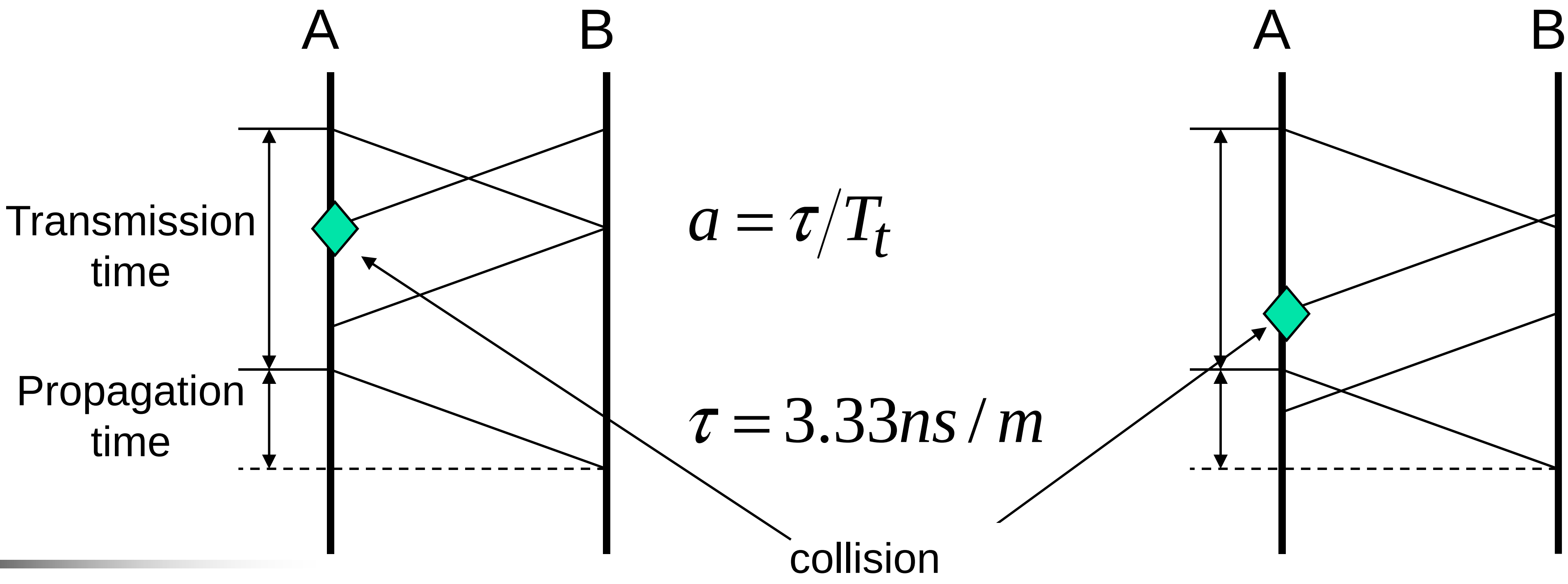


Backoff

- In case of collision
  - Transmission is stopped after sending the jamming sequence
  - the backoff interval is calculated
    - $[0..2^m-1]$ where m is the number of successive collisions
- A number is selected randomly in this interval
- Waiting time = selected number t . « slot time »
  - Slot time = 512 bit time (4096 @ 1Gbit/s)
- After each collision, the interval is doubled
  - Maximum interval of [0..1023] after 10 attempts
- In case of success, m is reset to 0

Collisions

- Collision zone = 2 x propagation time
  - This defines the « slot time »
  - Limits the cable length and defines the minimal size of a packet



$$a = \tau / T_t$$

$$\tau = 3.33 ns / m$$

Parameters (10base5)

- 50 ohms coaxial cable (thick ethernet)
- 10 Mbit/s
- Interframe silence = 96 bit time (9.6µs)
- « slot time » = 512 bits (51.2µs)
- Distance < 2500 m. between 2 stations
  - 5 segments of 500 m and 4 repeaters
- Minimum frame size = 512 bits
- Maximum frame size = 1518 octets
- Jamming sequence size = 32 to 48 bits

History

- 1970: ALOHAnet is developped at Hawai University
- 1976: Bob Metcalfe and David Boggs make a talk at the National Computer Conference
- 13 déc. 1977: U.S. Patent #4,063,220 - Multipoint data communication system with collision detection
- 1982: 1st version published by Xerox, Intel and DEC
- 1985: first 802.3 standard

Conventions

Speed	Transmission	Medium
1	BASE	-2
10	BROAD	-5
100		-36
1000		-T, -T2
10G		-F, -FB, -FL
		-X

Ethernet Frame

preamble	SFD	Destination address	Source address	Type / length	802.2 header and data	FCS	ESD
7 bytes	1 byte	6 bytes	6 bytes	2 bytes	46-1500 bytes	4 bytes	1 byte

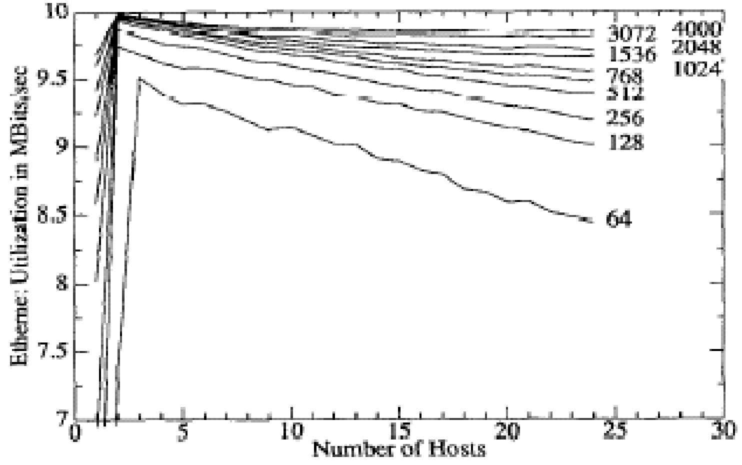
- MAC address
  - 24 bits: identification of manufacturer (Organization Unique Indentifier)
  - 24 bits: assigned by the manufacturer or the vendor

Vintage Ethernet Performances

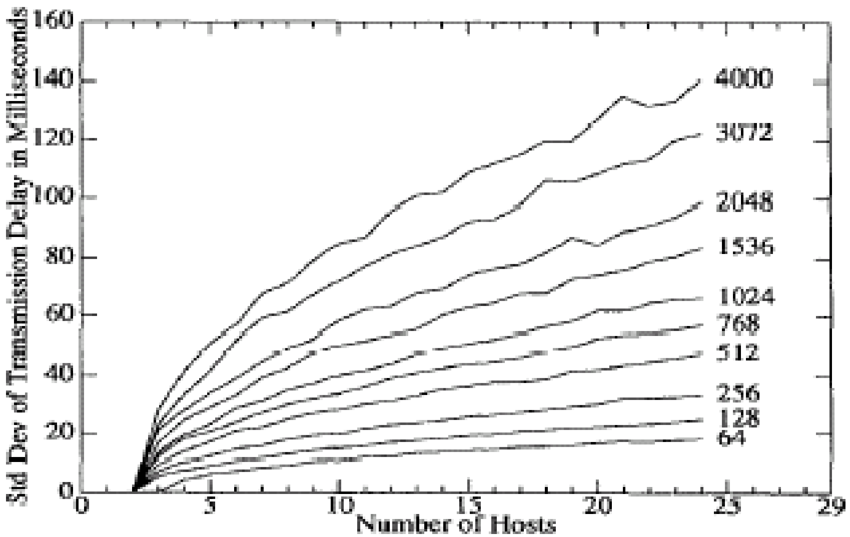
- Difficult to establish in theory
- A maximum throughput of 37% is often mentionned
  - This is only valid under simplified asumptions and small packets
- In practice
  - Access delay are minimal in case of low traffic
  - Transmission delay increases linearly with the size of the packets and the number of nodes
  - The delay standard deviation also increases but more slowly

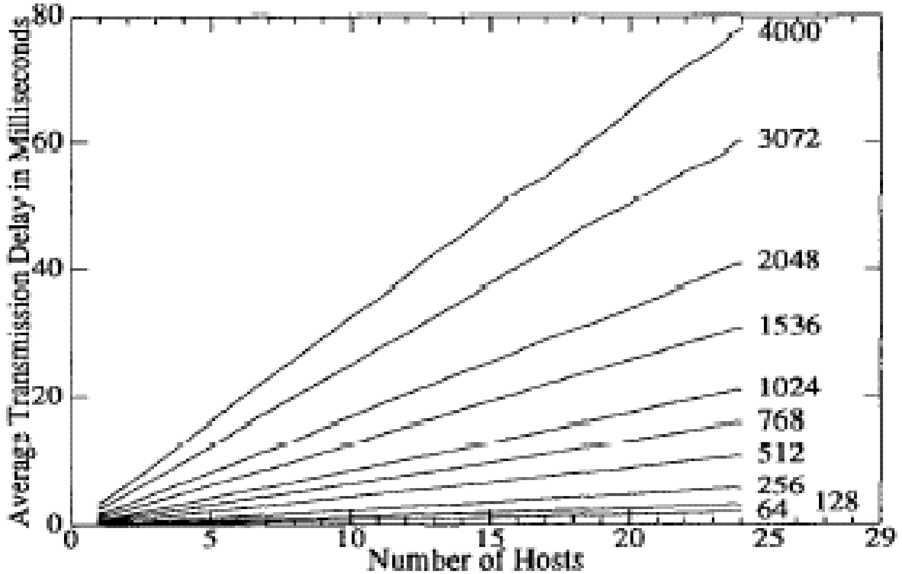
Measured performances at 10Mbit/s

- Mean delay
- Its standard deviation



Source Boggs, 1988





What makes a difference between theory and practice

- Calculations are complicated $\Rightarrow$ simplified asumptions
  - Poisson arrival law
    - In practice, it is less smooth
  - Infinite population and always waiting packets
  - Packets are dropped due to buffer overflow
- In reality
  - Cable lengths are much smaller than the maximum allowed
  - Packet sizes according to a bimodal distribution
    - Lengths close to min and lengths close to max.

Problems

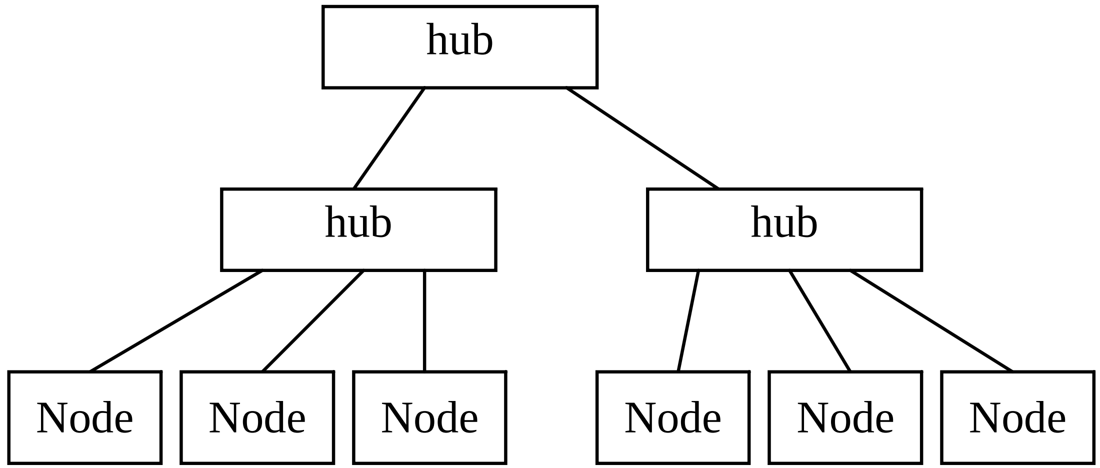
- « capture effect »
  - Assume 2 stations A and B have a lot of messages to send
  - At a given time, their emissions collide
  - A chooses 1 as a random number and B elects 0
  - Thus B succeeds at its first attempt
  - As B has more traffic, it attempts to send the next message right the way
  - This will collide with A second attempt
  - A has to double the backoff interval while B has the initial one
  - B has hence more chance to succeed than A, ….. And so on
- Unbounded transmission time
  - Not exact but !!!

Companion protocols

- IP: Internet Protocol RFC 791
- TCP: Transport Control Protocol, RFC 793
- ARP: Address Resolution Protocol, RFC 826
- IGMP: Internet Group Management Protocol
- ICMP: Internet Control Message Protocol,
- RARP: Reverse Address Resolution Protocol, RFC 903
- DHCP: Dynamic Host Configuration Protocol, RFC 2131-2
- IPv6, BGP, OSPF, RIP, NAT, …..

Improvements

- 1987: 1BASE5: 1Mbit/s on twisted pairs with hub
- 1990: 10 Mbit/s
- 1995: 100 Mbit/s
- 1998: 1 Gbit/s
- 10 Gbit/s
- 1997: full duplex
  - Hub replaced by a bridge (switch)



Parameters

- Distances on twisted pairs < 100m.
- Distances with optical fibers < 40 km (10G)

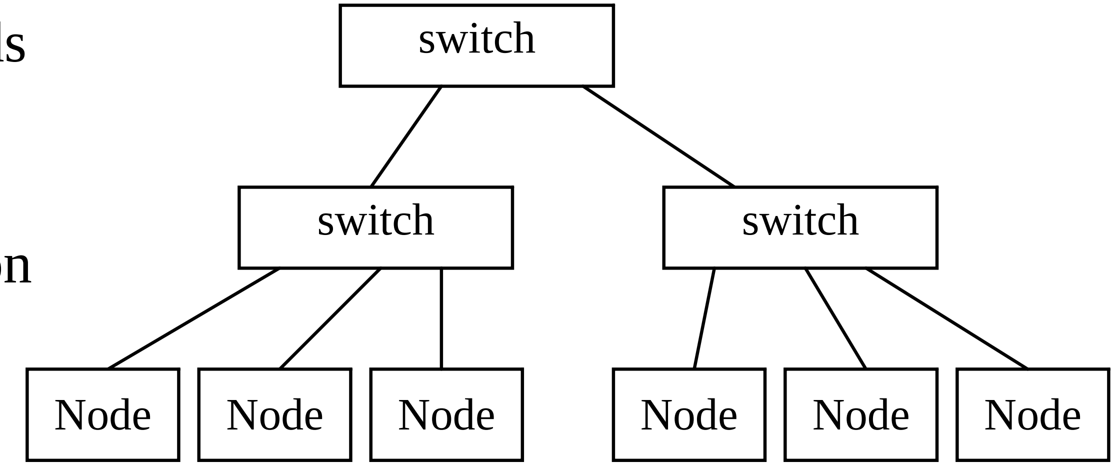
Speed	Interframe space	Slot time
10	96 bits	512 bits
100	96 bits	512 bits
1000	96 bits	4096 bits
10G	96 bits	NA

Switched Ethernet

- The original Ethernet operates in a half duplex mode
  - One station transmits or another but not both at the same time (except in case of collision)
- With a hub (repeater), there are only 2 nodes on a cable (hub + end node)
  - However the behavior must be the same as if all nodes share the same cable (collision domain)
  - This is necessary to detect possible collisions
- A switch
  - Receives on a port
  - Re-transmit on the port(s) toward the destination(s) of the packet
  - In between, it stores the packet

Switched Ethernet (2)

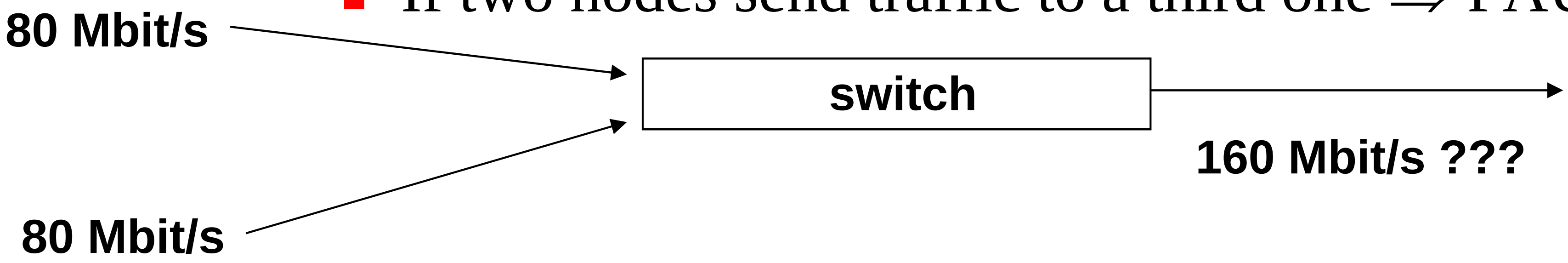
- With a switch, there are no longer collisions
  - An end node no longer needs to detect the collisions
  - As in a twisted pair cable, there a pair for each direction
    - Both directions may be exploited simultaneously



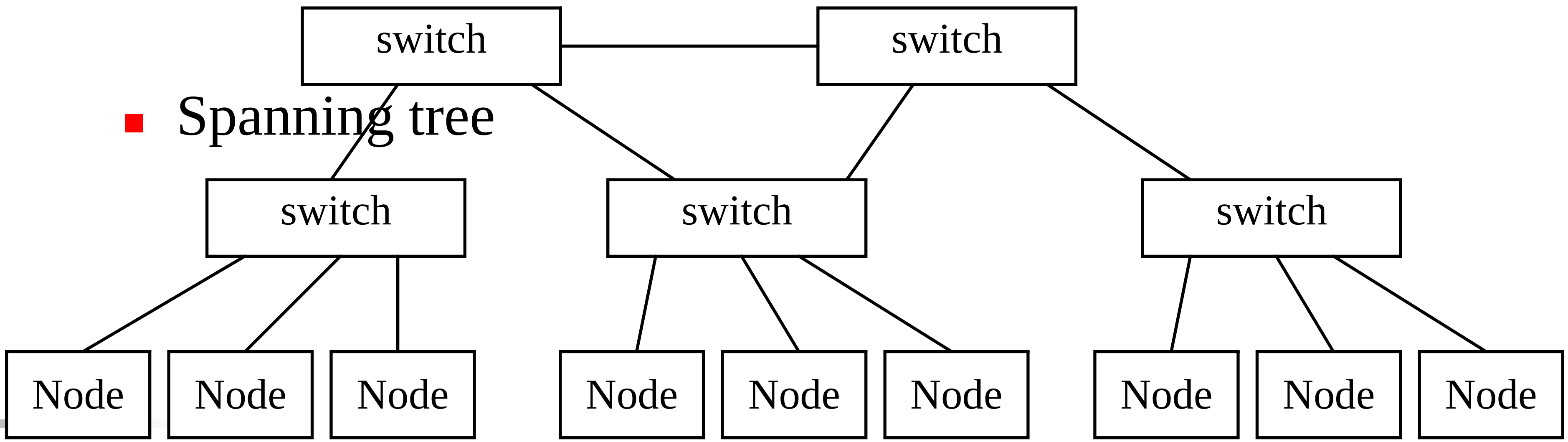
- Each link may hence be potentialy exploited at twice the speed
- The available bandwidth is potentialy higher (multiple domains)
- That does not mean that all problems are solved

Switched Ethernet (3)

- Switch buffers may overflow
  - If two nodes send traffic to a third one ⇒ PAUSE frame



- What about multiple paths between two nodes
  - Spanning tree

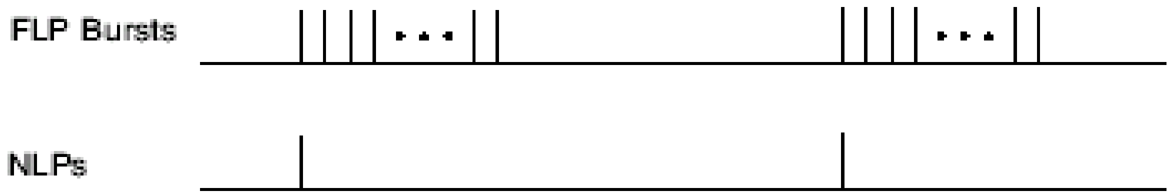


Hubs and switches

- Sometimes misleading
  - Hub 10/100 is a switch
- Multilayer switches
  - Include a routing module and a switching module

Complementaty aspects

- Auto negociation
  - every 16ms +/-8



- VLAN – IEEE 802.1Q
  - A way to isolate traffic among a group of nodes
  - Uses the « type/length » field
  - Add a LAN identifier (12 bits) and a priority field (3 bits)
- Quality of service – IEEE 802.1D (includes 802.1p)
- Self learning
  - Switches may learn on which ports the nodes are reachable
  - Multicast raises a problem
    - IGMP snooping

Quality of service management

- According 802.1D (MAC bridges)
  - Not only for Ethernet
  - 8 traffic priorities
    - Most switches only implement « best effort », « real-time » and « management » on each egress
    - Packets arriving in a ingress port are inserted in the queue corresponding to the priority at the egress port (FIFO order within a priority level)
    - When the output link is free, the oldest packet of the highest priority is transmitted
      - This may cause priority inversions which are quite long (1500 bytes)
    - Packets that have no priority field or those that enter through selected ports may be forced at a given priority level
      - Very useful to filter the traffic coming from outside
- !!! This is not to be confused with QoS at network level such as IntServ or DiffServ

How to improve QoS in Ethernet ?

- Access time
  - Modify the MAC
  - Add another MAC on top of CSMA/CD
    - Pure TDMA, master/slave, token, reservation, etc.
  - Adapt the traffic
  - Use of switches
- Synchronisation, dating, consistency
  - Clock synchronization algorithms

Add another MAC

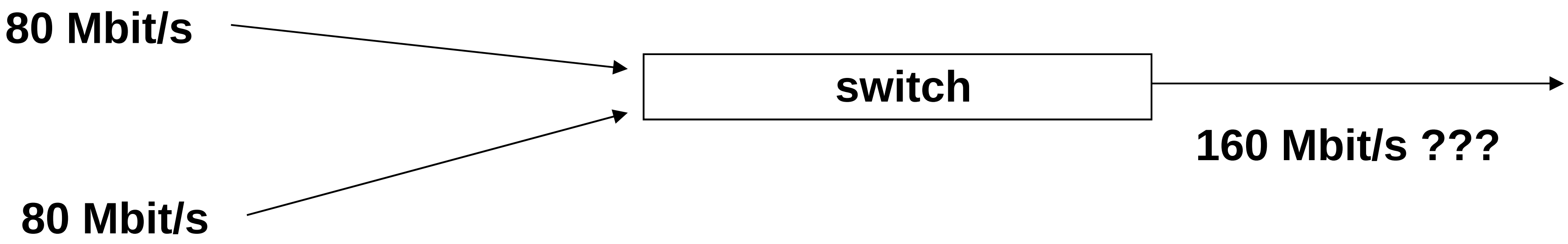
- A number of such proposals
- Nearly all MAC types have been suggested
- The MAC used in many fieldbussses has been used
- Generaly speaking, rather inefficient
  - Pure TDMA @ 1Gbit/s $\Rightarrow$ 4% utilization

Traffic adaptation

- Avoid bursts due to non real-time traffic
  - smoothing (limit traffic instantaneous intensity by delaying part of it)
  - shaping (periodic emissions by blocks)
- This only adds statistical guarantees
  - Reduces drop rate and jitter
  - Increases mean delay

Use of switches

- Avoids collisions
- Queues in switches may however overflow



- Increases mean delays and increases jitter

Clock synchronization

- Using messages to synchronize and provide simultaneous sampling is not possible (unless exception)
- Accuracy γ of synchronization depends on the uncertainty ε in the transmission delay
  - $\gamma \geq \varepsilon (1 - 1/n)$ where n is the number of participating nodes

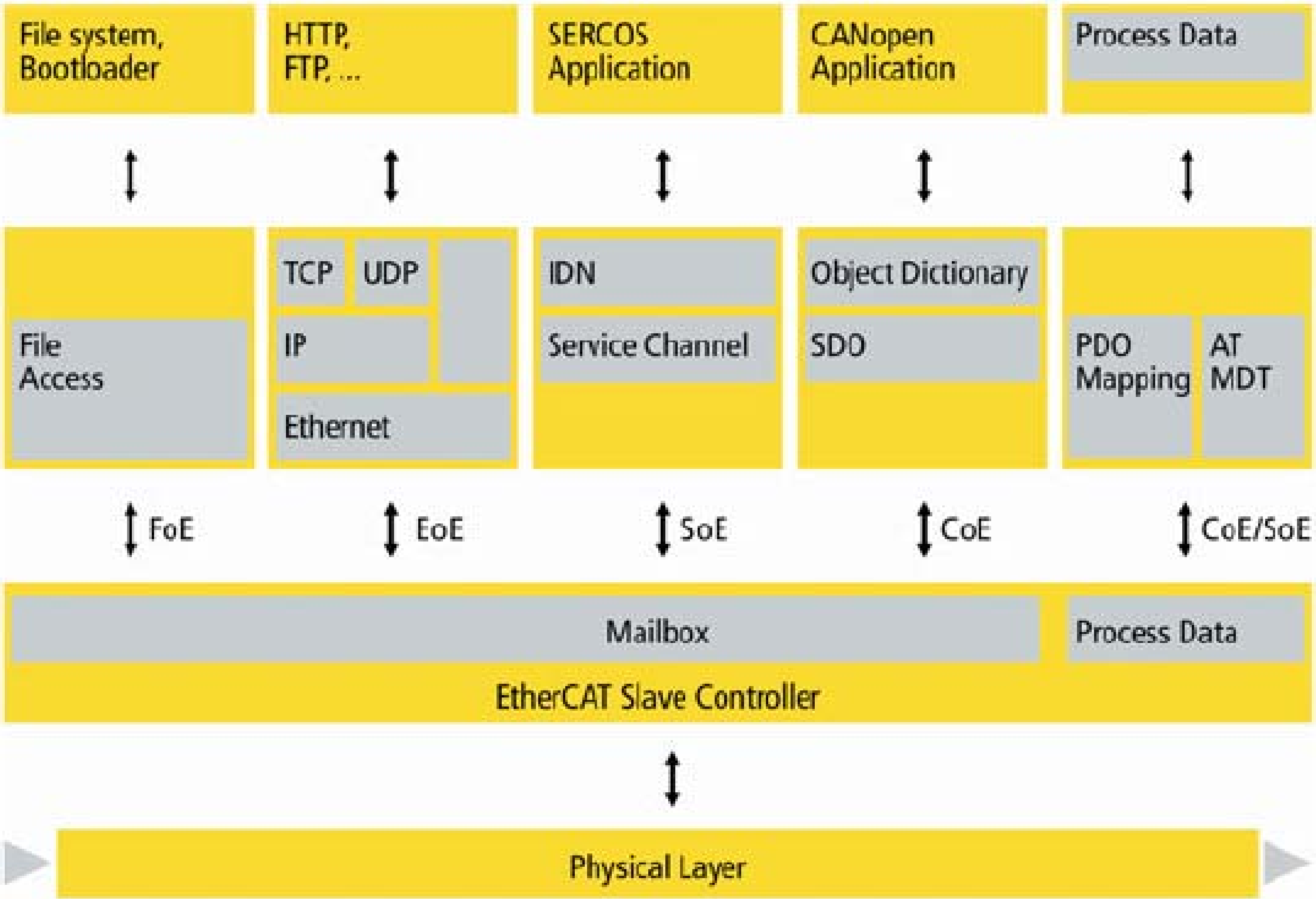
Current RT Ethernet contenders

- Non standard hardware
  - EtherCAT
  - SynqNet
  - SERCOS III
  - PowerDNA ?
- Standard hardware but not compatible with regular 802.3 nodes
  - Ethernet Powerlink
  - FTT-Ethernet
  - TTP over Ethernet (TTE)
- Standard hardware, compatible with regular 802.3 nodes
  - PROFINET
  - JetSync
  - EtherNet/IP (+CIP)
  - Modbus-TCP
  - Real-Time Publish-Subscribe
- And a number of academic proposals

EtherCat

- General topology with 65535 nodes / segment
  - Wide choice of cable from twisted pair to optical fibers
- Synchronisation based on distributed synchronized clocks
  - Claim accuracy to microseconds
- Quite performing
  - 100 axes in 100µs; 1000 I/O points in 30 µs
- Compliant with IEEE 802.3u Fast Ethernet
  - All protocols based on IP are supported (TCP, UDP,...)
  - Dedicated hardware (except for the master)
- Integration with CANopen and SERCOS profiles for compatibility with legacy
- Node to node communication only via master

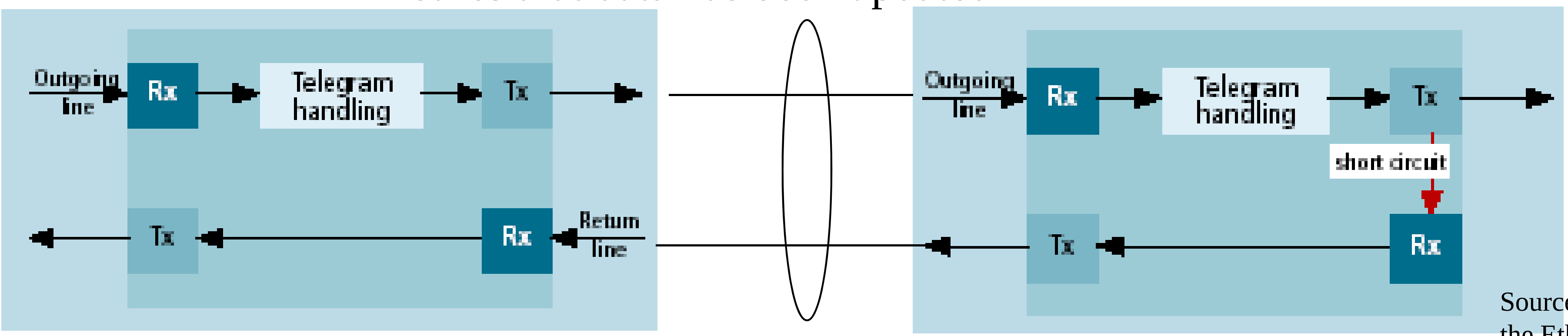
EtherCat – Architecture & protocols



EtherCat technical introduction and overview, dec. 2004

EtherCat – Principles

- Master sends an Ethernet frame
  - That contains one or more EtherCat datagrams
    - Contains data for the outputs and has room for the input data
    - Nodes read and update part of the frame on the fly
    - Location in datagram is independent from the physical localization
- Inside each slave, there is a sync manager
  - That informs of the reception of new data
  - That manages the data read and write order
    - Ensures that data has been updated



Ethernet cable

Source: EtherCat the Ethernet fieldbus,

EtherCat - analysis

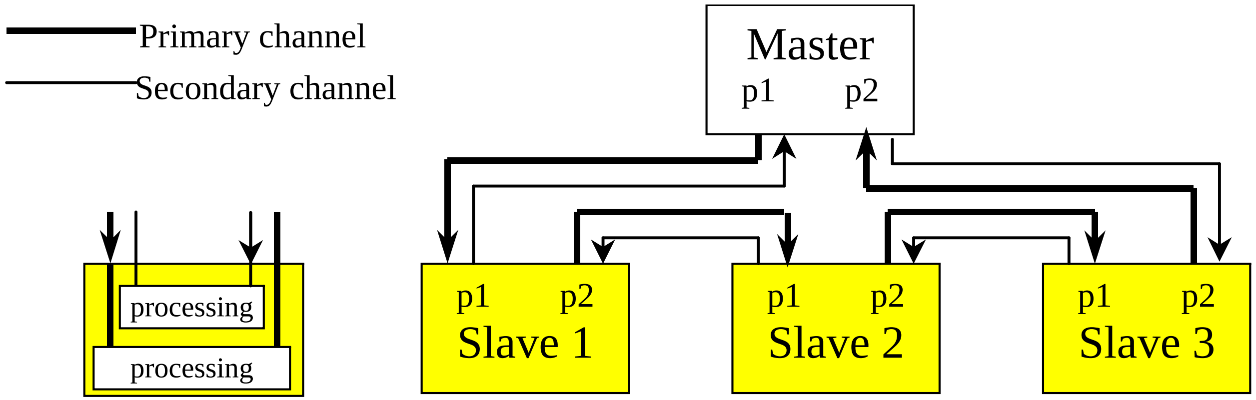
- Advantages
  - Very fast (is this useful ?)
  - Standard hardware for the master
  - Compatible with legacy CAN Open or SERCOS
- Drawbacks
  - Targeted to inputs and outputs which is not sufficent in many cases
  - Incompatible with Ethernet
  - Requires specialized hardware
  - Restrictive topology

SERCOS III

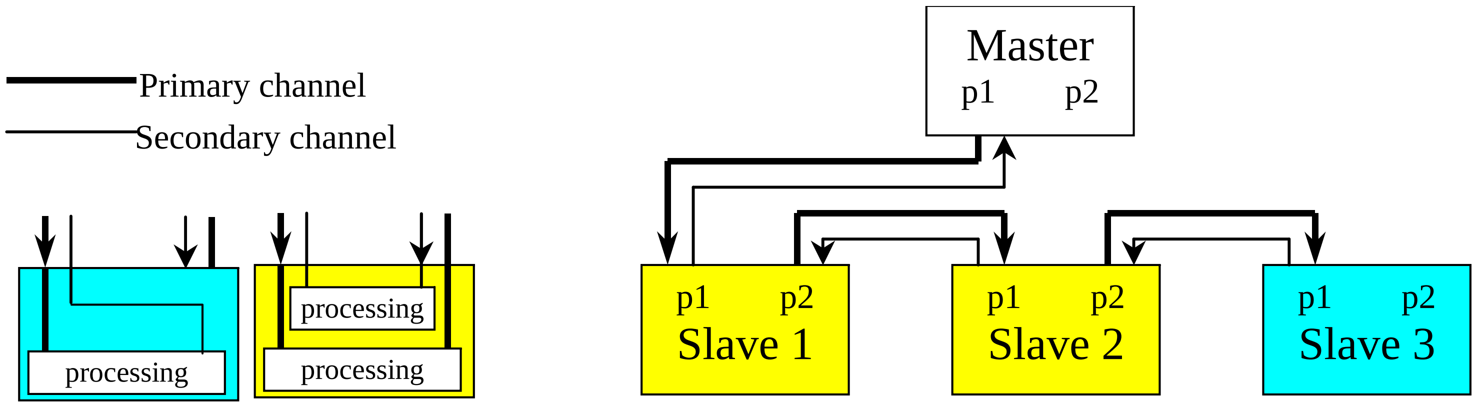
- Up to 254 nodes on a single segment
  - Line or ring topology
    - Medium redundancy capability
- So-called « guaranteed and deterministic communication »
  - Cycle time from 31.25µs (62.5, 125, 250 and n x 250µs)
  - jitter <1µs (or 50µs for the low performance class)
- Compliant with IEEE 802.3u Fast Ethernet (fiber and copper)
  - IP based protocols are supported (TCP, UDP,...)
  - Implementation on dedicated hardware
- Integration SERCOS profiles
- Possibility of real-time node to node communication

SERCOS III - Topology

- Ring



- Line

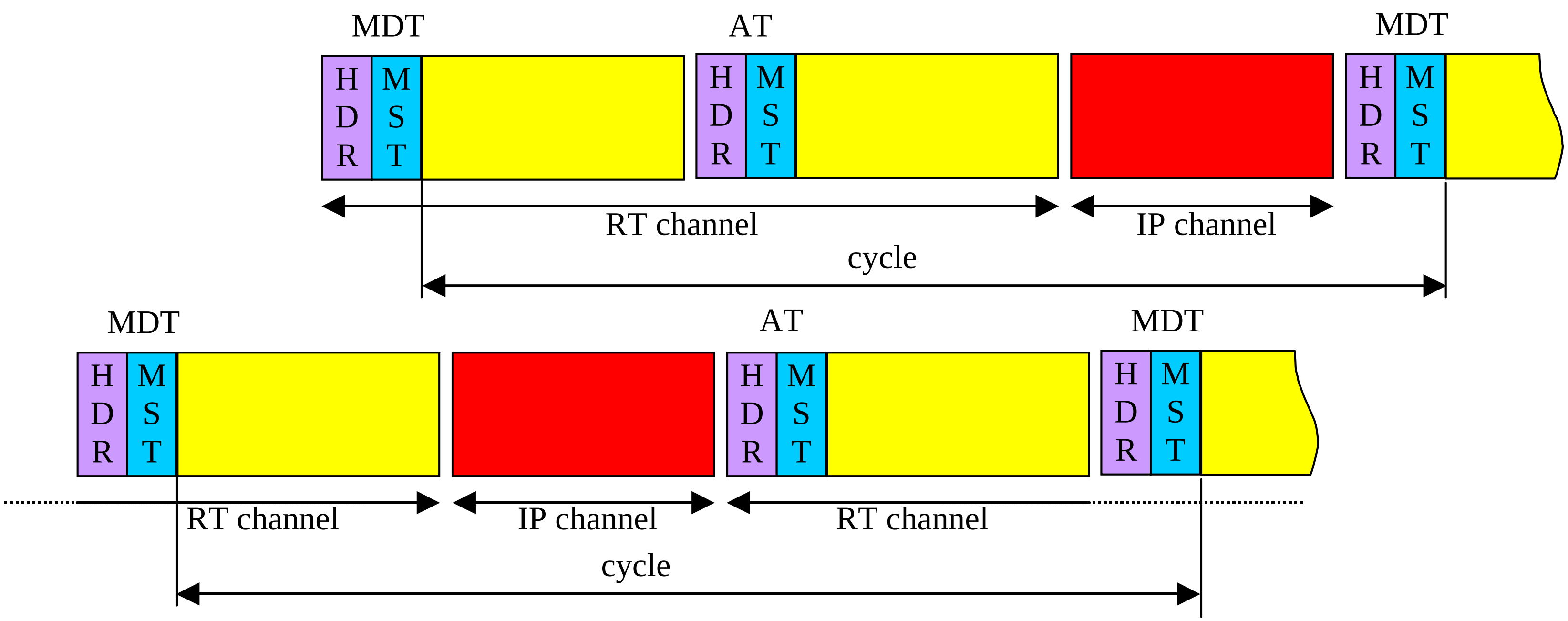


SERCOS - cycle

- Real-time channel present in each cycle to
  - Synchronize the slaves with the master
  - Exchange data and commands between the master and the slaves
    - In each cycle (real-time)
    - On request (service channel)
  - Exchange data between slaves
- IP channel for on demand transfers (optionnal)
  - Base on IP
  - Permet to exchange data between master and slaves
    - Operator display data, files, configurations, …
  - Exchange between slaves
  - Transparent communication with standard Ethernet nodes (PC, …)

SERCOS III – cycle

- Up to 4 Master Data Telegrams
- Up to 4 AT Device Telegrams
- A number of IP messages (max. duration)



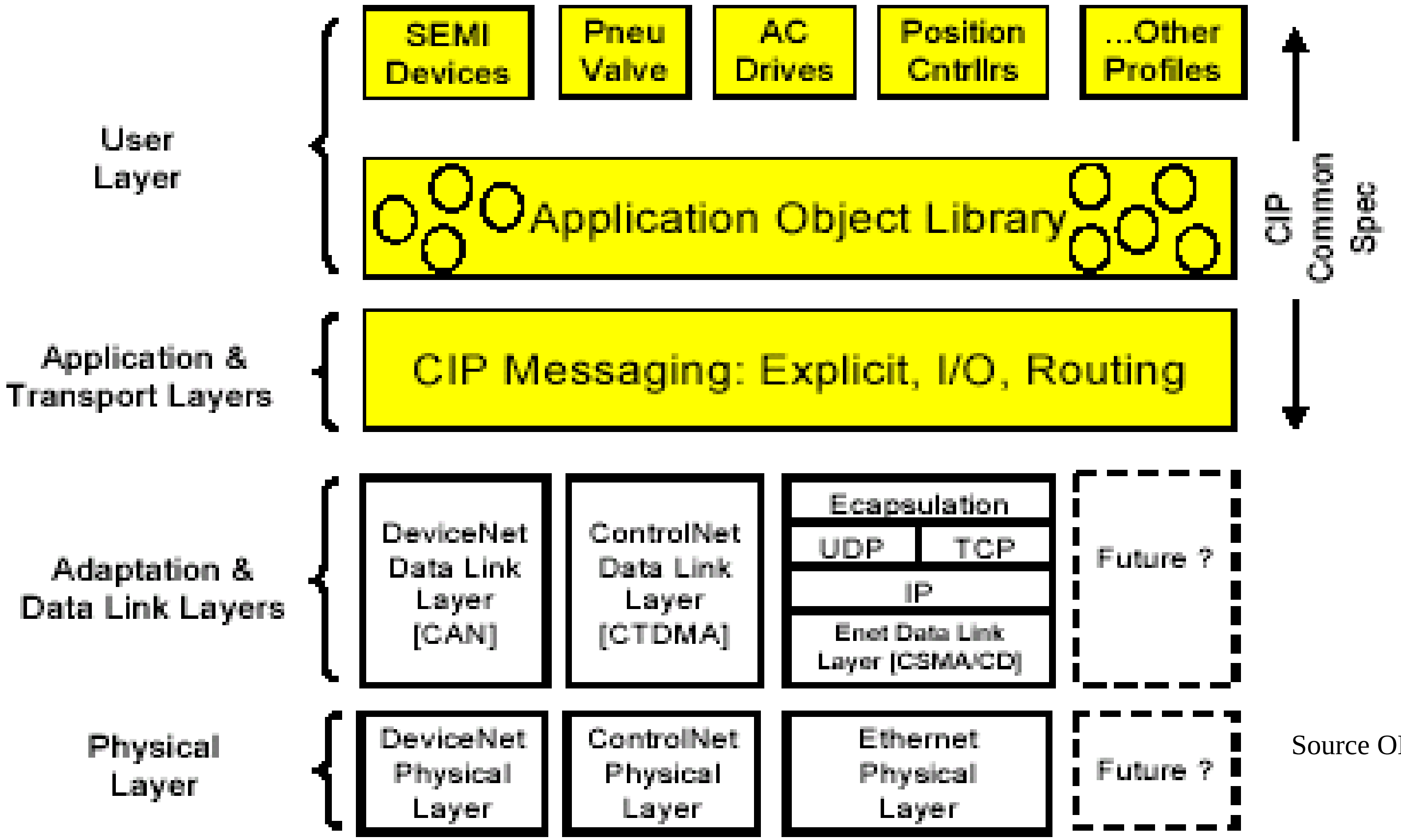
SERCOS III - analysis

- Advantages
  - Efficient / possibility of line topology
  - hotplug
  - Redundancy capability
  - A node may be hooked to a regular Ethernet network (one port must be left open)
  - Guranteed cycle time and reduced jitter
  - Uses regular IP frames (no encapsulation)
- Drawbacks
  - Protected by patents
  - Requires non standard hardware
  - Real-time behavior impossible in presence of non compliant nodes
  - Requires a bridge to regular Ethernet

Ethernet IP with time synchronization

- No limit in node number
  - Topology according to Ethernet
- No modification to 802.3
  - It is possible to use 802.1p priorities
  - Synchronisation according to IEEE 1588
- Compliant with IEEE 802.3u Fast Ethernet
  - IP based protocols are supported (TCP, UDP,...)
  - Implementation on standard 802.3 hardware
- Integration with CIP for compatibility with DeviceNet and ControlNet
- Node to node communication along a producer/consumer model

Ethernet IP



Source ODVA

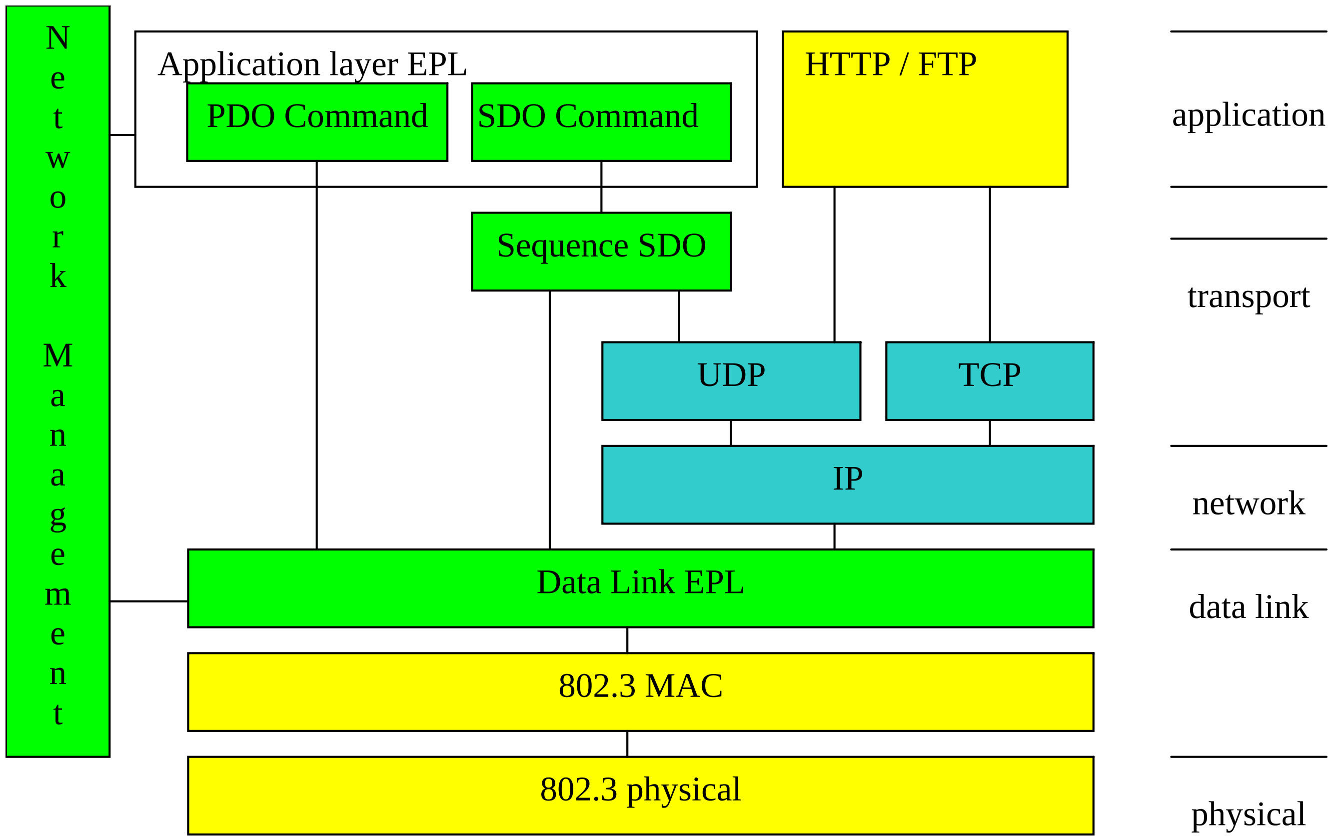
Ethernet IP - analysis

- Advantages
  - Standard hardware
  - Most of the used protocols are standard
  - Can be mixed with « pure » Ethernet nodes
  - Compatibility with DeviceNet and ControlNet
  - Synchronisation based on IEEE 1588
- Drawbacks
  - Cycle time is not constant, jitter is not controlled
  - Only statistical guarantees
    - Using priorities in switches
    - Filtering external traffic

Ethernet Powerlink in short

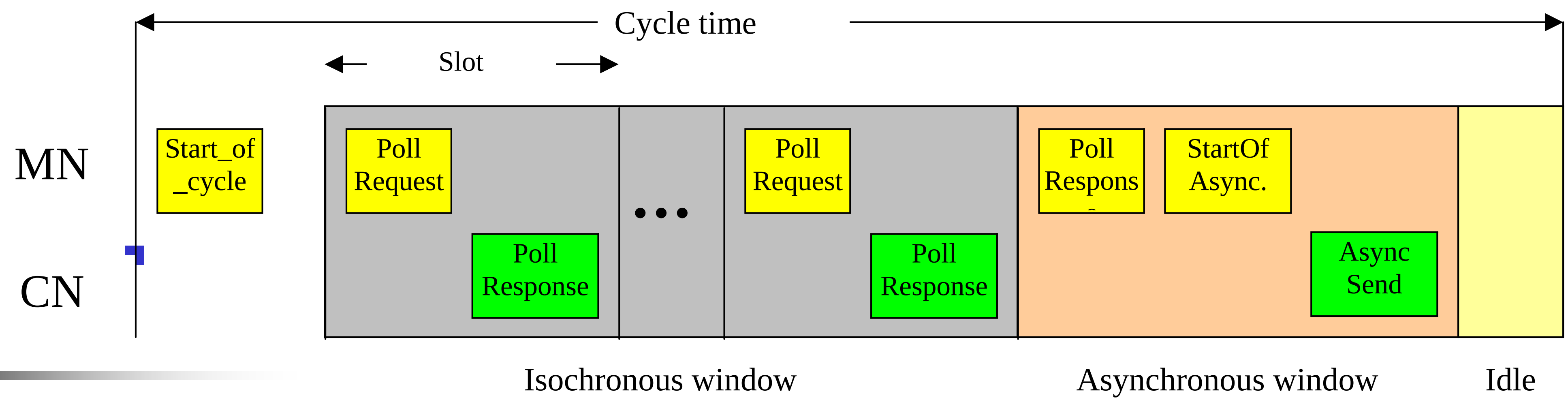
- Up to 240 nodes on a segment
  - Topology according to Ethernet + line topology possible
- So-called « guaranteed and deterministic communication »
  - Cycle time from 200µs
  - jitter <1µs for precise node synchronisation
  - IAONA real-time class 4 (highest performance one)
- Compliant with IEEE 802.3u Fast Ethernet
  - IP based protocols are supported (TCP, UDP,...)
  - Implementation on regular 802.3 hardware
- Integration with CANopen EN50325-4 for interoperability
- Node to node communication along a producer/consumer mode

Ethernet Powerlink – Model



Ethernet Powerlink – cycle

- Synchronous part made of slots
  - Continuous: CN is polled at each cycle
  - Multiplexed: a given CN is polled every n cycles
  - Synchronous exchanges
    - Sequence of PollRequest followed by PollResponse
    - Response is sent in broadcast (available to all)
      - Producer -> consumer
- Idle part (to keep a constant duration to the cycle)



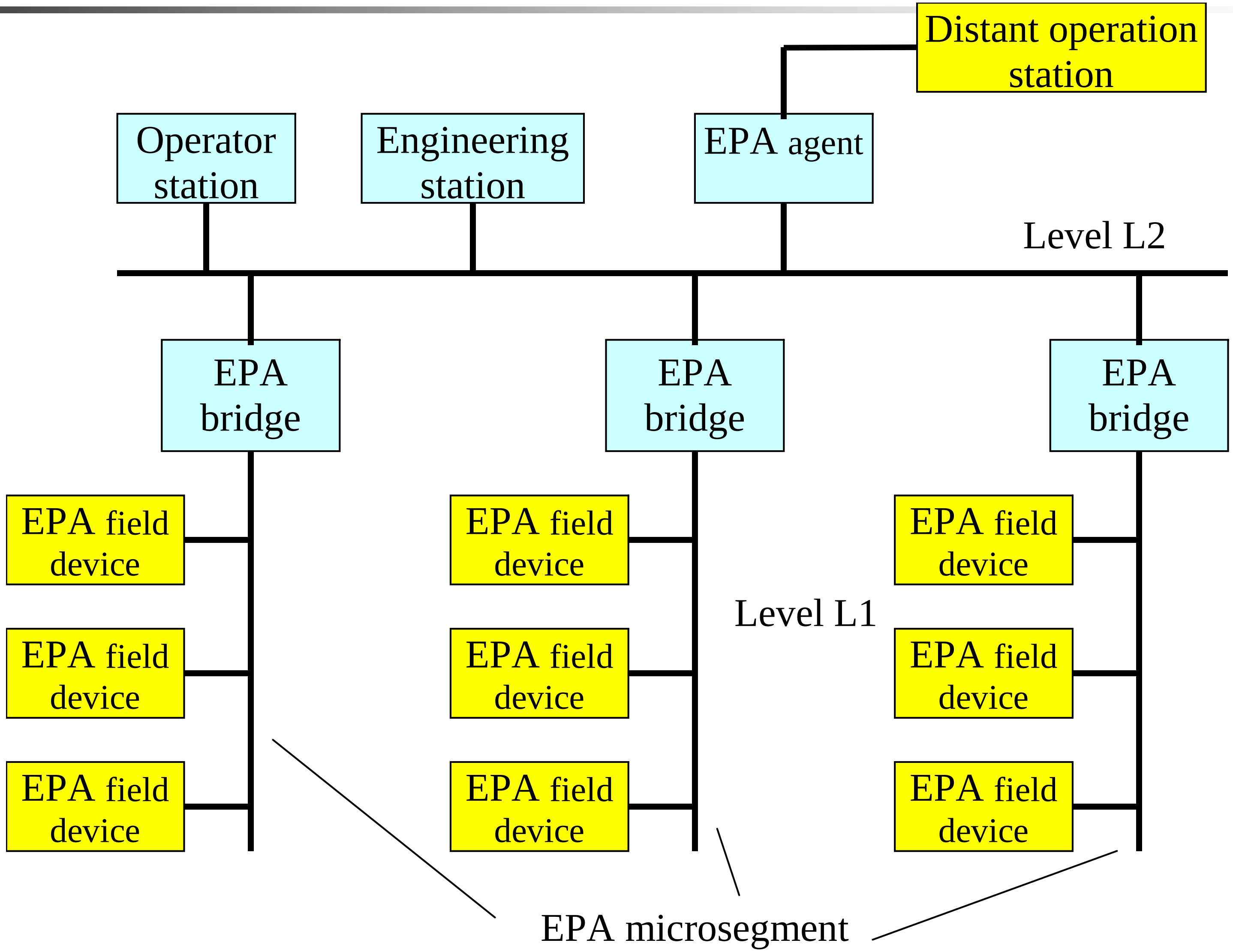
Ethernet Powerlink - Analysis

- Advantages
  - Constant cycle duration (in absence of error)
  - Line topology possible
  - Presence of an application layer
- Drawbacks
  - Only contains a limited subset of the necessary concepts (cf. WorldFIP)
  - Low efficiency (25µs lost in each transaction)
  - Cannot coexist with regular Ethernet nodes
  - Requires specialy designed routers and hubs
  - Model does not take errors into account
  - Quite sensitive to timing errors (collisions !)

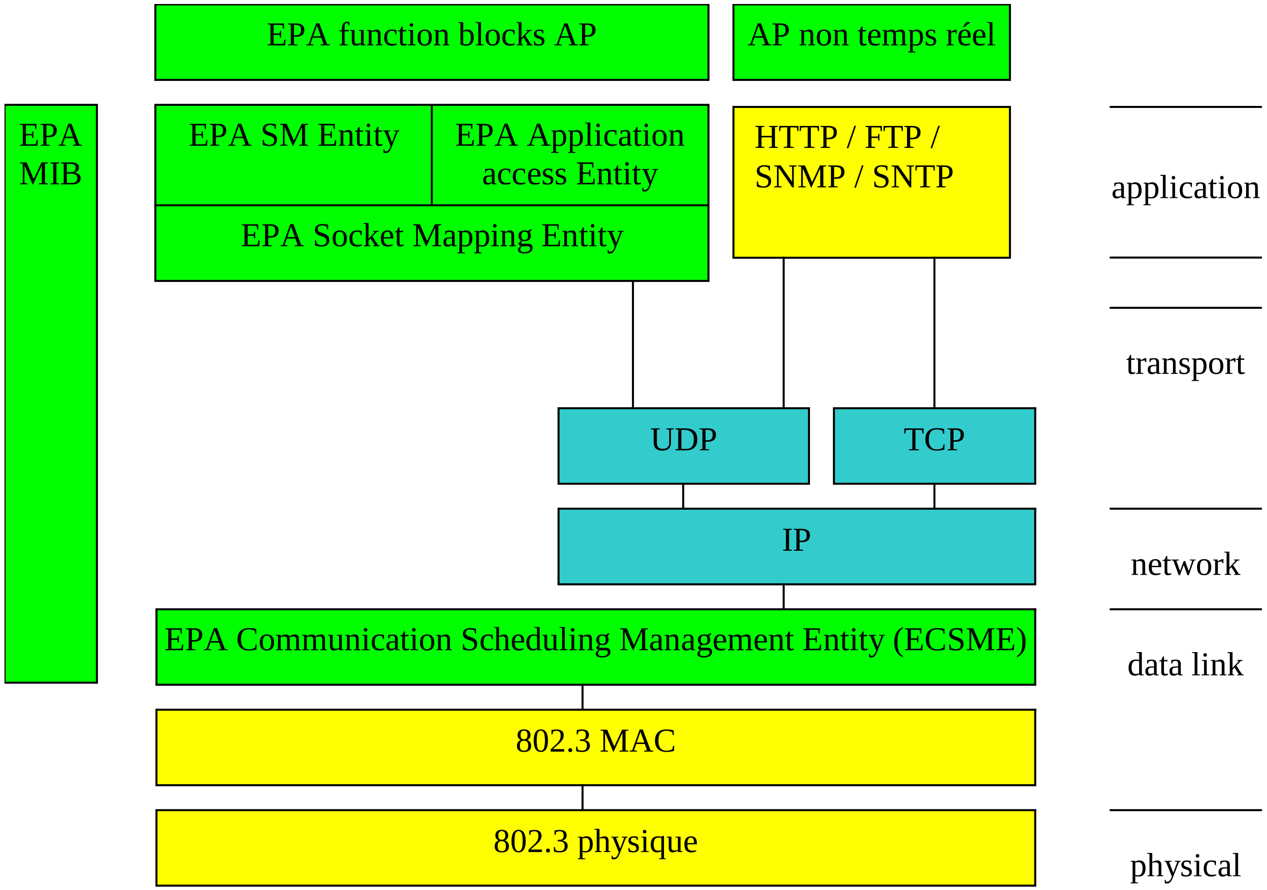
Real-Time Ethernet EPA - Ethernet for Plant Automation – in short

- Number of nodes only limited by physical considerations
  - Topology according to Ethernet (switch and hub)
- So-called « guaranteed and deterministic communication »
  - Cycle time in multiples of milliseconds
  - jitter depends on the selected IEEE 1588 class
- Compliant with IEEE 802.3 (all variants)
  - IP based protocols are used (TCP, UDP,...)
  - Implementation on regular 802.3 hardware
- Based upon IEEE 1588 for clock synchronisation
- Device capabilities described in XML
- Conventional client-server communication model
  - Process variables communicated according to IEC 61499 and IEC 61804 (function blocks)

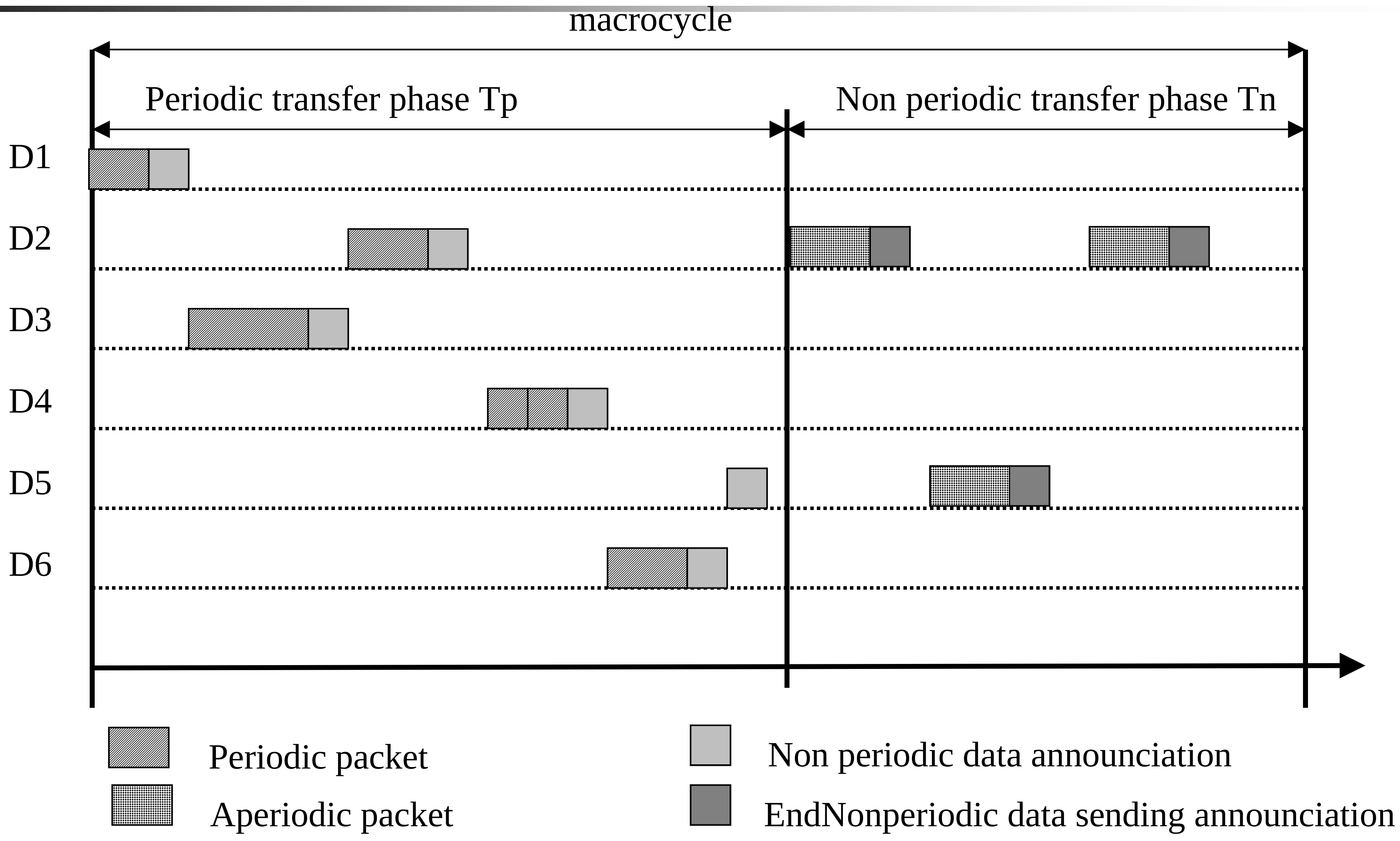
EPA – Topology



Real-Time Ethernet EPA - architecture



EPA - scheduling



Real-Time Ethernet EPA - analysis

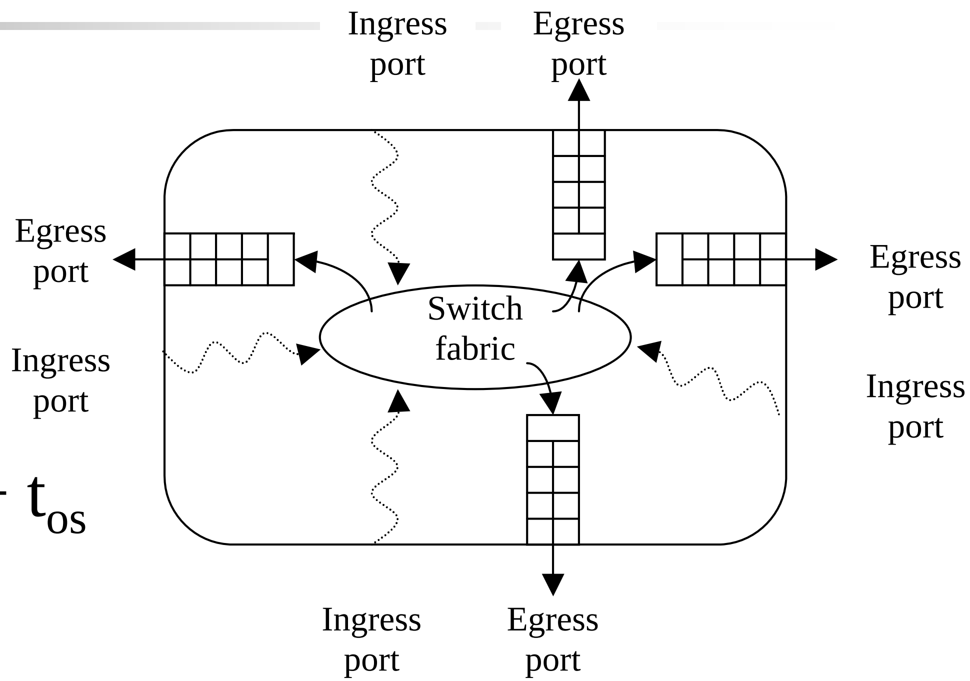
- Advantages
  - Support functional blocks
  - suitable application layer
- Drawbacks
  - Only client-server
  - Single polling cycle
  - Difficult to calculate guarantees (even in absence of error)
  - Some parts of the standard are quite obscure
  - Configuration must be centralized
    - To ensure that parameters are compatible and consistent

Others

- Standards
  - Profinet
    - V3 introduced real-time guarantees
  - AFDX Avionics Full Duplex Switched Ethernet
  - P-NET on IP
    - PNET is a virtual token MAC
    - Same on IP without any guarantee
  - Real-Time publish-subscribe
    - Just a publish subscribe layer on top of TCP without any guarantee
  - MODBUS on TCP/IP
    - Nothing real-time, just the protocol on top of TCP
  - Vnet/IP
    - Based on temporal windows / Proposed by Yokogawa (www.yokogawa.com/us)
  - Time Critical Control Network
- Non standard (so far): TT-Ethernet

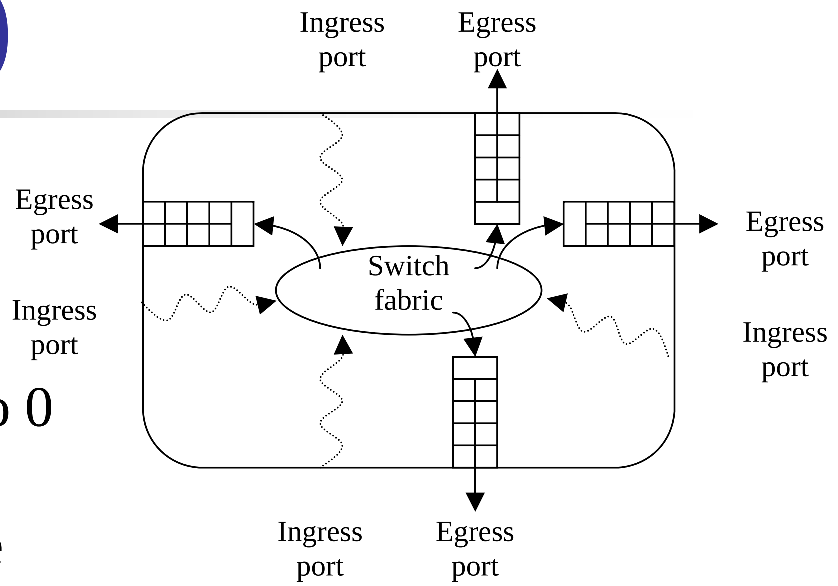
Switched Ethernet

- Temporal definitions
  - t_{mux} : multiplexing delay
  - t_{queue}: queueing delay
  - t_{trans} : packet transmission delay = $t_{frame} + t_{mux} + t_{queue} + t_{os}$
  - t_{os} : operating system delay
  - t_{frame}: frame tx delay
    - 121µs for 1514 bytes @ 100Mbit/s
- How to calculate the delay ?
  - Use of the network calculus



4 port switch

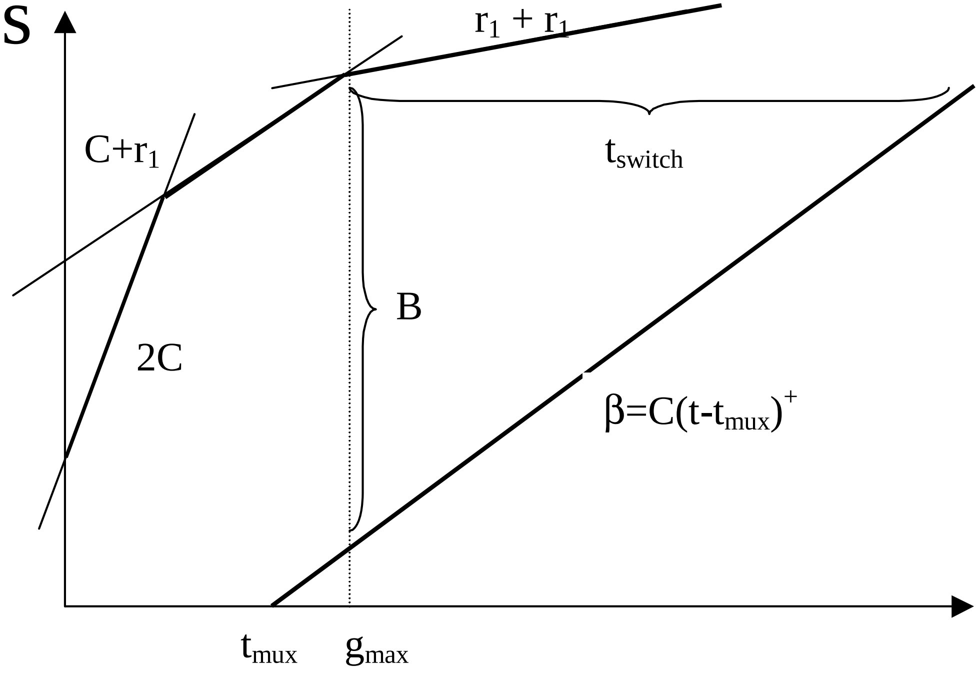
Switched Ethernet (2)



- Let us look at one output port
  - Service curve $\beta(t) = C\,(t-t_{mux})+$
    - + means that the function is equal to 0 for $t-t_{mux} < 0$
    - C is the maximum transmission rate on the output port
  - Arrival curve
    - Sum of arrival curves of all incoming traffic whose destination is that output port
    - One station on each ingress port : $\alpha(t) = \min\,[Ct + M, r_k t + b_k]$
      - M maximum frame size
      - r_k long term average rate ($\Sigma\, r_k \leq C$ should hold for output)
      - b_k burstiness of the traffic ($M < b_k$ and $r_k \leq C$)

Switched Ethernet (3)

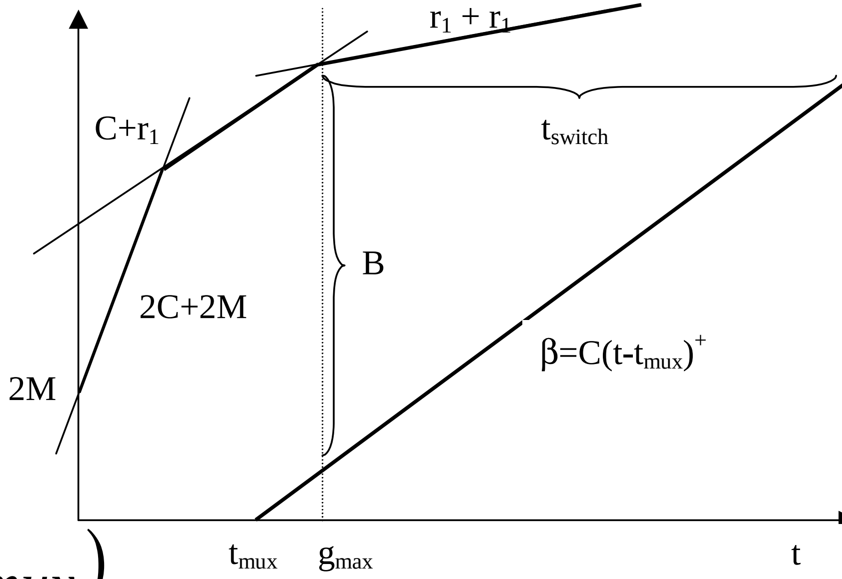
- Case of 2 incoming traffics
  - Maximum backlog = max. vertical distance between arrival and service curves
  - Maximum delay (FIFO order) = max. horizontal deviation between arrival and service curves



Switch backlog (FIFO order)

- Inflexion point

$$g_i = \frac{b_i - M}{C - r_i}$$



- Backlog

$$B = \sum_{k=1}^{N} b_k + \sum_{k=1}^{N} r_k g_{\max} - C\left(g_{\max} - t_{mux}\right)$$

$$B = \sum_{k=1}^{N} b_k - g_{\max}\left(C - \sum_{k=1}^{N} r_k\right) + C t_{mux}$$

- Backlog estimation

$$B_{est} = \sum_{k=1}^{N} b_k + C t_{mux}$$

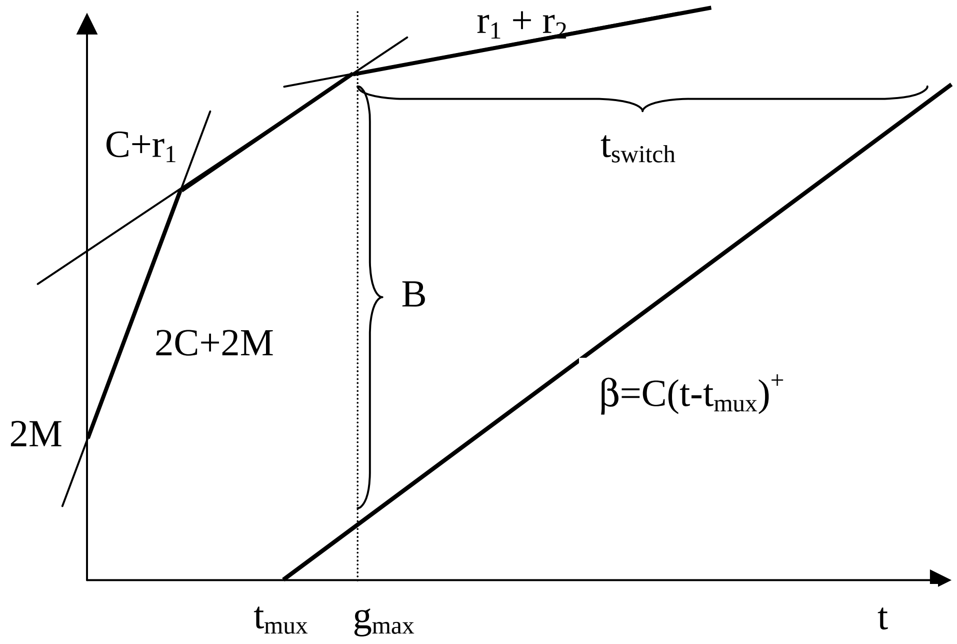
Switch delay (FIFO order)

- Distance at g_{max}
- Delay

$$t_{switch} = \sum_{k=1}^{N} \frac{b_k}{C} - g_{\max}\left(1 - \sum_{k=1}^{N} \frac{r_k}{C}\right) + t_{mux}$$

- Can be approximated

$$t_{est} = \sum_{k=1}^{N} \frac{b_k}{C} + t_{mux}$$

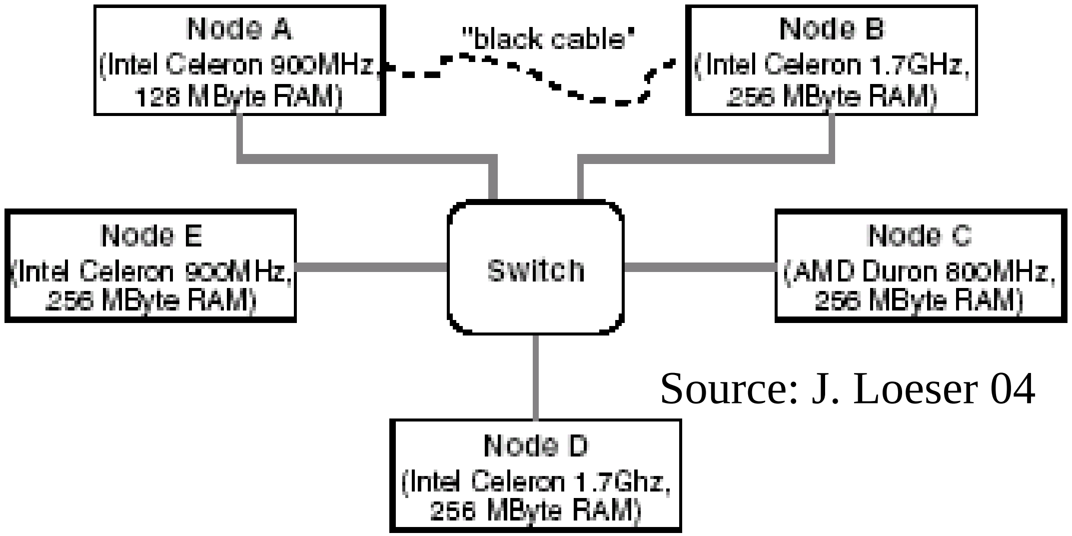


Remarks

- Arrival curve implies traffic shaping
  - This can be done using the token bucket algorithm
    - Bucket size b (If token arrives when the bucket is full, it is discarded)
    - Fill rate r (1 token is added to the bucket every 1/r seconds)
      - Alternately S tokens added every S/r seconds
    - Maximum packet size M
    - When a packet of n bytes arrives, n tokens are removed from the bucket, and the packet is sent to the network (waits otherwise).
- Backlog is per egress queue
  - Is a way to calculate necessary buffer size
- Analysis is valid for one switch (May be extended)

Measurement setup

- 3 different switches
  - 3Com office connect
  - Level-One FSW-2108TX
  - Intel Netstructure 470F
- NIC
  - Intel EEPro/100
  - 3Com 3C985B-SX
- Delays measured from with min. size frames on UDP



Source: J. Loeser 04

Switch buffer size

- Measured by sending bursts of traffic
  - Increasing burst length until missing frames
  - Using network calculus to derive buffer size

Switch	# 1514 byte frames	Size [Kbyte]
100M 3Com	14	20.5
100M level-One	87	127.4
1000M Intel	200	293

Source: J. Loeser 04

Influence of shaping interval

- Bucket size: $b = r\,T_s + M$

Node	C (40MB/s)	D (32MB/s)	E (20MB/s)
T_s = 10ms	51514 bytes	41514 bytes	26514 bytes
T_s = 1ms	6515 bytes	5514 bytes	4014 bytes
T_s = 0.1ms	2014 bytes	1914 bytes	1764 bytes

Source: J. Loeser 04

- CPU usage

Node	C (40MB/s)	D (32MB/s)	E (20MB/s)
T_s = 10ms	4.1%	2.9%	2.3%
T_s = 1ms	11%	9%	7.2%
T_s = 0.1ms	21.2%	17.2%	11.9%

Buffer bounds and delays

- Packet transmitted from A to B

	buffer bound	t_{max}	t_{est}	t_{obs} max.
T_s = 10ms	111.8 KB	9.357 ms	9.731 ms	8.759 ms
T_s = 1ms	15.7KB	1.38 ms	1.345 ms	1.3 ms
T_s = 0.1ms	6.1 KB	0.582 ms	0.506ms	0.438 ms

Source: J. Loeser 04

Analysis of the experiment

- Max. bandwidth on a link 98.6 Mb/s
- Experiment used 92 Mb/s ≈ 93% link use
  - Max. delay from 0.58ms, 1.4 ms to 9.4 ms at 100 Mb/s
  - Requires traffic shaping
- Results valid
  - for a single switch
    - Must be adapted for cascaded switches (traffic shapes)
  - For FIFO behavior
  - Provided all nodes play the game

Can be extended to multiple switches, see J. Specht, S. Samii, Urgency-Based Scheduler for Time-Sensitive Switched Ethernet Networks, ECRTS 2016, pp.75-85

Conclusion

- Vintage Ethernet was not so bad
- Switched Ethernet
  - Gives the possibility to use up 100% of the bandwidth
    - With switches that have large enough buffers
    - With traffic shaping
    - Provided all nodes stick to the traffic shaping rules
  - Is not so good in terms of topology
  - Still suffer from delay variations

The solution ?

- For satisfying real-time constraints
  - Switched Ethernet
  - Use of message priorities in switches (IEEE 802.1Q)
  - Traffic smoothing / shaping
  - IEEE 1588 for clock synchronization
- Look at IEEE Time Sensitive Networking group (https://en.wikipedia.org/wiki/Time-Sensitive_Networking)
- Security: 802.1x, IPSec, SSL
- Application layer (MMS or similar)

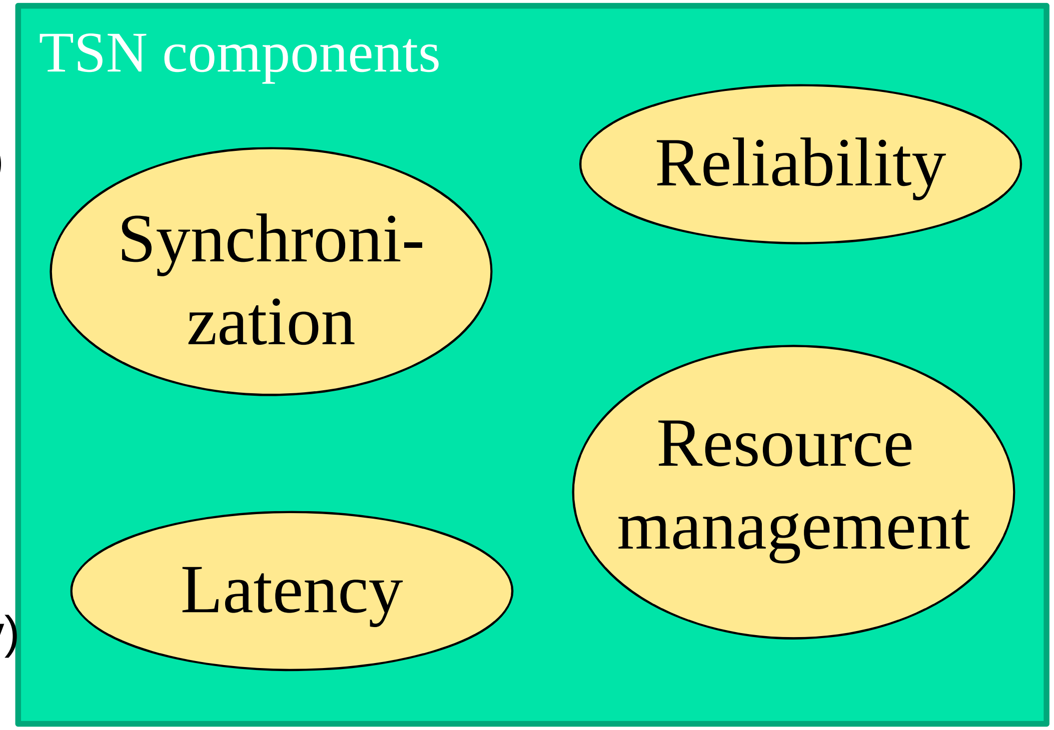
Work on Ethernet

- Still an active subject of research
  - Mainly studying the switch scheduling policies and the corresponding schedulability analysis
- Improving the standard
  - Time Sensitive Networking working groups under the IEEE 802.1 project
    - See next slide
  - Audio Video Bridging (AVB)
    - Started in 2005. First efforts to bring real-time to standard Ethernet
      - See U. Bordoloi et al, Schedulability Analysis of Ethernet AVB Swiches, RTCSA 2014.
    - Now generalized and renamed to TSN

Time Sensitive Networking (TSN)

Timing and sync (802.1AS)

Preemption (802.1Qbu)
Scheduled traffic (802.1Qbv)
Cyclic queuing (802.1 Qch)
Asyn. Shaping (802.1Qcr)
Credit-based shaper (802.1 Qav)



Frame replacement & elimination (P802.1CB)
Path Control (802.1Qca)
Per-Stream Filtering (802.1 Qci)
Time Sync (P802.1AS-Rev)

Stream Reserv Prot (802.1Qat)
TSN configuration P802.1Qcc)
YANG (P802.1Qcp)
Link-local Reserv Prot P802.1CS)

Standards for TSN (Time Sensitive Networking)

- IEEE Std 802.1Qbu-2016 - IEEE Standard for Local and Metropolitan Area Networks -- Bridges and Bridged Networks -- Amendment 26: Frame Preemption. It allows a Bridge Port to suspend the transmission of non time critical frames while one or more time critical frames are transmitted
- IEEE Std 802.1Qbv-2015 - IEEE Standard for Local and Metropolitan Area Networks -- Bridges and Bridged Networks -- Amendment 25: Enhancements for Scheduled Traffic. It specifies time aware queue draining to schedule the transmission of frames relative to a known time scale.
- • IEEE Std 802.1Qca-2015 - IEEE Standard for Local and Metropolitan Area Networks -- Bridges and Bridged Networks -- Amendment 24: Path Control and Reservation. It extends the application of Intermediate System to Intermediate System (IS-IS) to bridged networks in order to provide explicit trees for data traffic.

All 3 are now in IEEE Std 802.1Q-2018

AVB standards (Audio Video Bridging)

- IEEE Std 802.1AS-2011 - IEEE Standard for Local and Metropolitan Area Networks -- Timing and Synchronization for Time-Sensitive Applications in Bridged Local Area Networks. It provides a Layer 2 time synchronizing service that is appropriate for the most stringent requirements of consumer electronics applications.
- IEEE Std 802.1Qat-2010 - IEEE Standard for Local and Metropolitan Area Networks -- Virtual Bridged Local Area Networks - Amendment 14: Stream Reservation Protocol (SRP). Has been rolled into IEEE Std 802.1Q-2014.
- IEEE Std 802.1Qav-2009 - IEEE Standard for Local and Metropolitan Area Networks -- Virtual Bridged Local Area Networks - Amendment 12: Forwarding and Queueing Enhancements for Time-Sensitive Streams, which specifies the Credit Based Shaper. Has been rolled into IEEE Std 802.1Q
- • IEEE Std 802.1BA-2009 - IEEE Standard for Local and Metropolitan Area Networks -- Audio Video Bridging (AVB) Systems. It specifies a set of usage-specific profiles to help interoperability between networked devices using the AVB specifications.

TSN Standards under study

- P802.1CS – Link-local Registration Protocol
- P802.1DC – Quality of Service Provision by Network Systems
- P802.1DF – TSN Profile for Service Provider Networks
- P802.1DG – TSN Profile for Automotive In-Vehicle Ethernet Communications
- P802.1AS-Rev – Timing and Synchronization for Time-Sensitive Applications
- P802.1AX-Rev – Link Aggregation Revision
- P802.1ABcu – LLDP YANG Data Model
- P802.1CBcv – FRER YANG Data Model and Management Information Base Module
- P802.1CBdb – FRER Extended Stream Identification Functions
- P802.1CMde – Enhancements to Fronthaul Profiles to Support New Fronthaul Interface, Synchronization, and Syntonization Standards

TSN Standards under study (2)

- P802.1Qcj – Automatic Attachment to Provider Backbone Bridging (PBB) services
- P802.1Qcr – Bridges and Bridged Networks Amendment: Asynchronous Traffic Shaping
- P802.1Qcw – YANG Data Models for Scheduled Traffic, Frame Preemption, and Per-Stream Filtering and Policing
- P802.1Qcx – YANG Data Model for Connectivity Fault Management
- P802.1Qcz – Congestion Isolation
- P802.1Qdd – Resource Allocation Protocol

More info at https://1.ieee802.org/tsn/#Ongoing_TSN_Projects

References

- G. Held, « Ethernet networks », 4th edition, Wiley, 2003, ISBN 0-470-84476-0
- D. Boggs, J. Mogul, C. Kent, “ Measured Capacity of an Ethernet : myths and reality “, in Proc. SIGCOMM’88, Stanford, Ca, Aug.16-19, 1988, pp.222-234.
- A. Tanenbaum, « Networks », 4e ed., Pearson, ISBN 2-7440-7001-7
- J. Loeser, H. Haertig, « Low-latency Hard real-time communication over switched Ethernet », Proc. ECRTS 04, Catania, It., pp. 13-22.
- J.-D. Decotignie, « Ethernet Based Real-Time and Industrial Communications », Proc. of the IEEE, Vol. 93 (6), June 2005
  - Look at the references given in the paper

Information Sources

- Modbus (www.modbus.org) specs are online
- P-Net
- Ethernet Powerlink (www.ethernet-powerlink.org) few docs
- EtherCat (www.ethercat.org) some brochures
- Profinet (www.profibus.com)
  - PROFINET, « PROFINET CBA Architecture Description and Specification », Version 2.02, May 2004
- Ethernet IP (www.ethernet-ip.org)
- SERCOS (www.sercos.de)
- Fachhochschule Reutlingen website has a lot of information (www-pdv.fh-reutlingen.de/rte)

REAL-TIME NETWORKS
intro to wireless, IEEE 802.11 and Bluetooth

Prof. J.-D. Decotignie

CSEM Centre Suisse d'Electronique et de Microtechnique SA

Jaquet-Droz 1, 2007 Neuchâtel

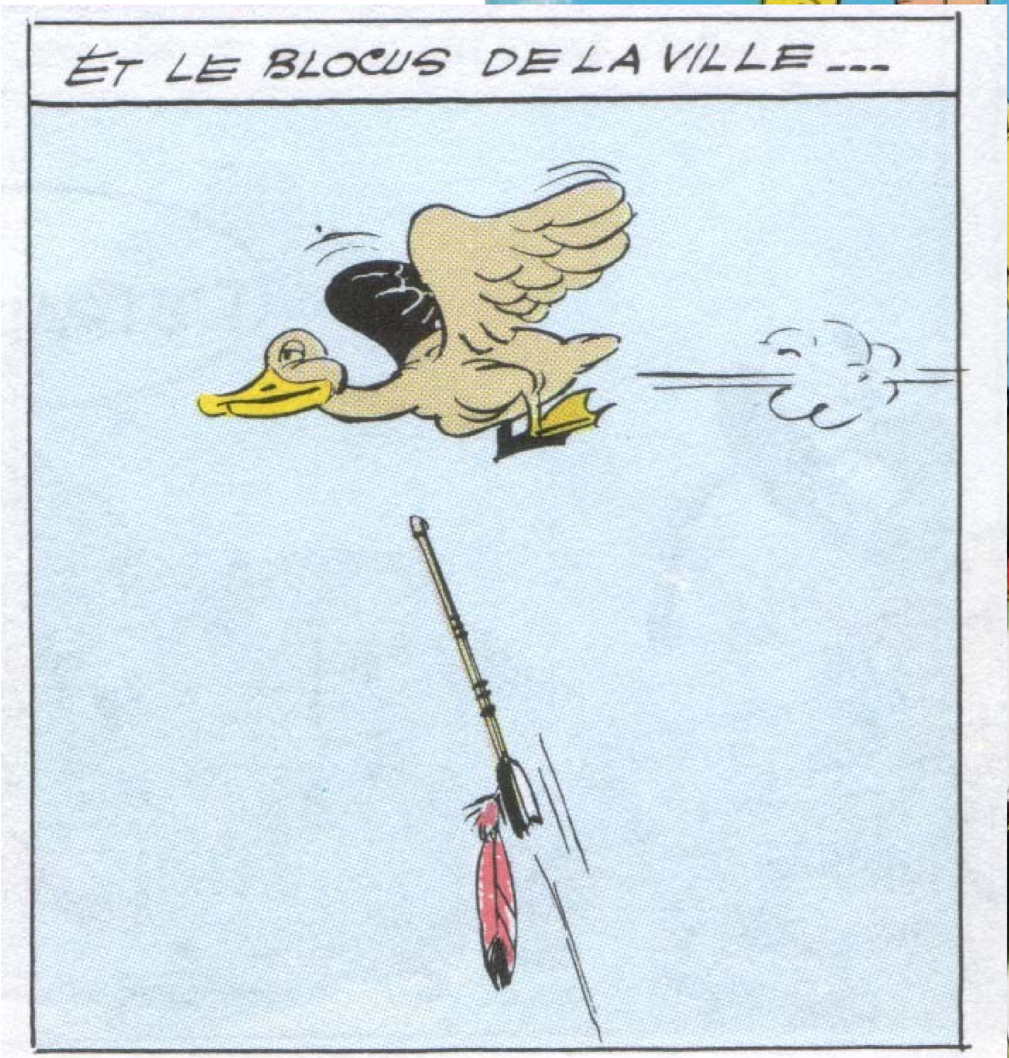
jean-dominique.decotignie@csem.ch

The Troy war

Troy –
Lemnos -
Mount Athos -
Euboea -
Euripos -
Plain of Asopos -
Mount Kithairon -
Saronic Gulf -
Argos

Chappe

More Recently



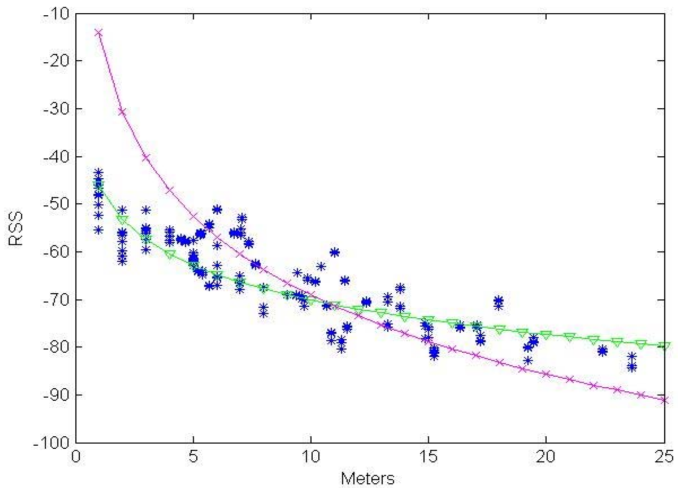
assumptions

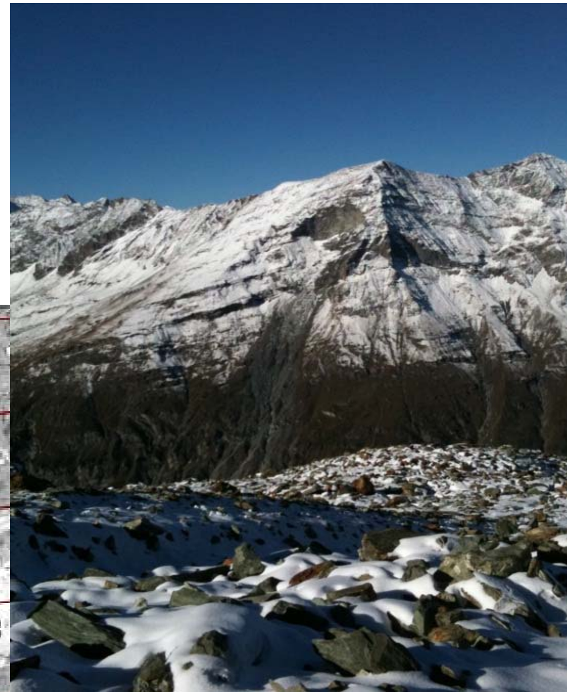
Myths about wireless transmission

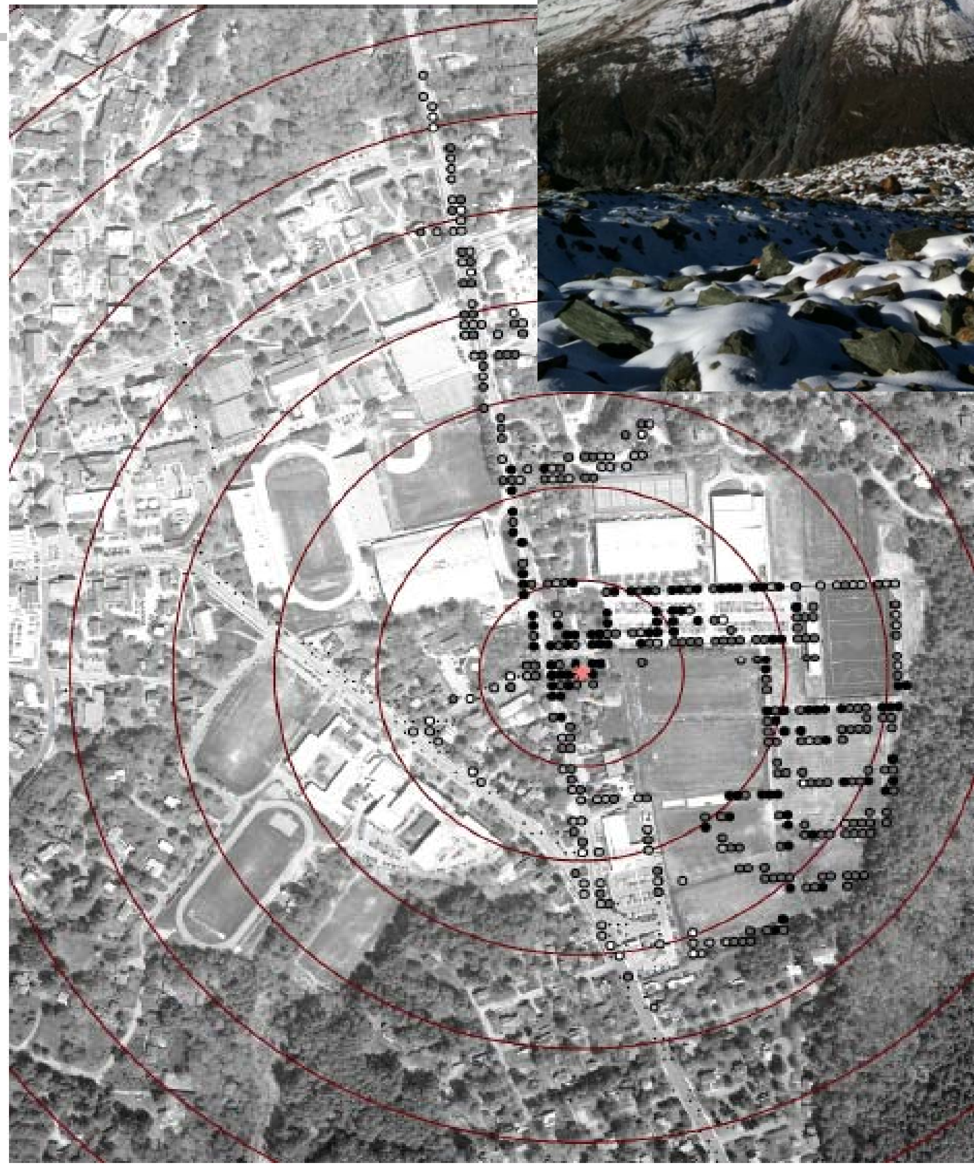
- The world is flat & radio transmission area is circular
  - signal strength is a simple function of distance
- All radios have equal range
- Link quality does not change
  - if I can hear you, you can hear me & if I can hear you at all, I can hear you perfectly
- The only source of packet loss is collision
- Broadcast is for free
- Energy is proportional to the number of packets and their size
- Duty cycling is the only way to reduce energy consumption

“transm. area is circular” “the world is flat”

- radio coverage is not at all circular
  - obstacles, height, fading, …
- signal strength is loosely related with distance

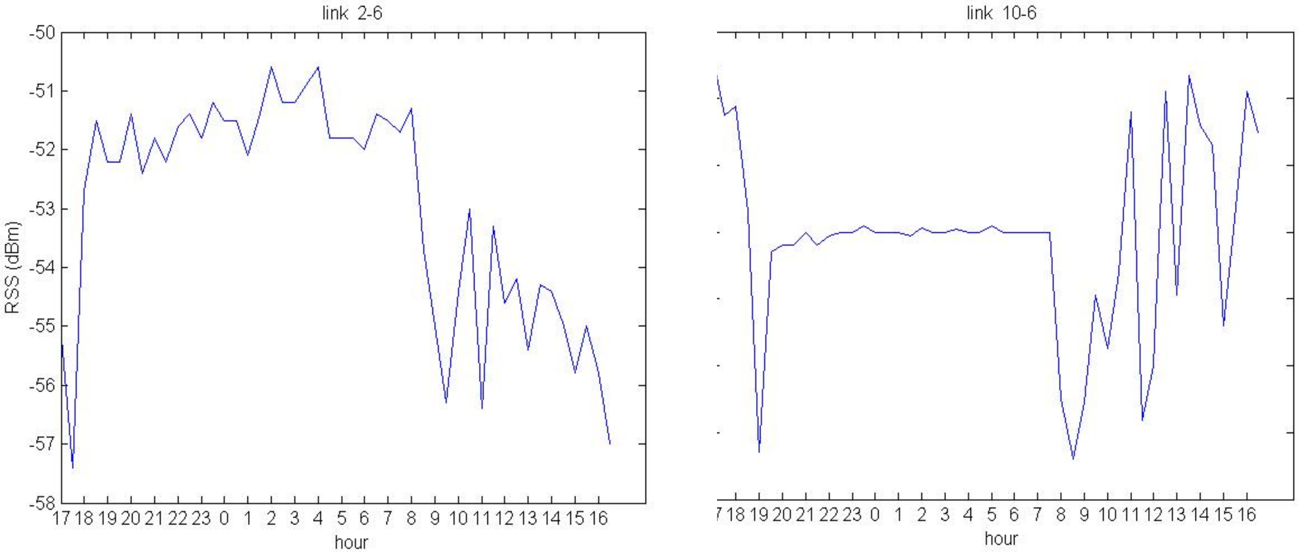


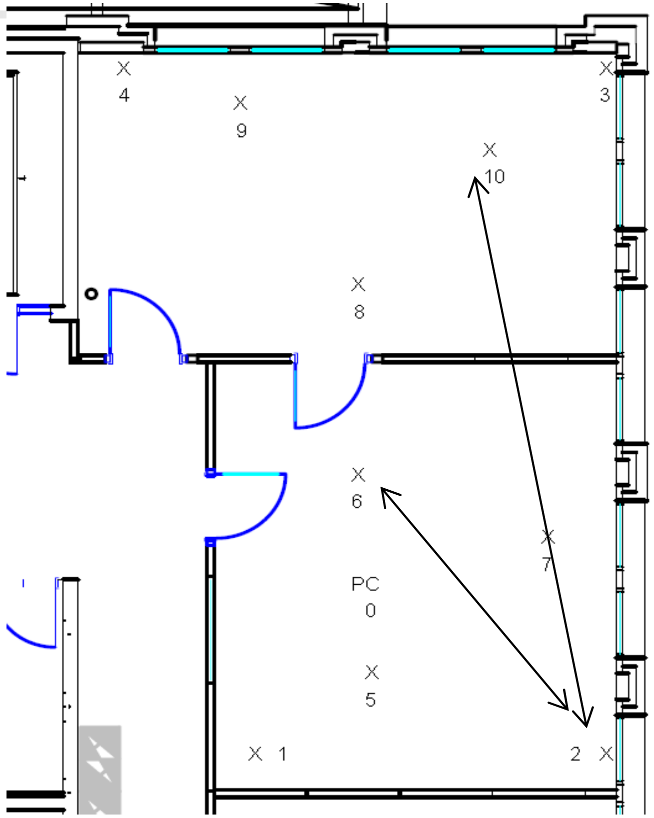




source: D. Kotz et al., 2003

"link quality does not change"

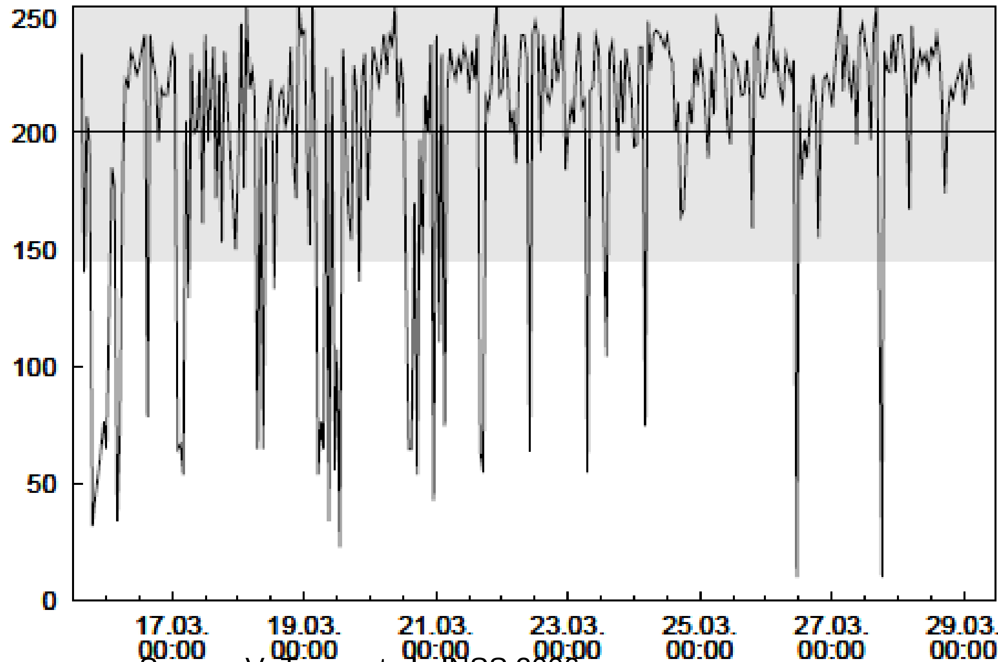




- links fall into 3 categories
  - connected, transitional, disconnected
- transitional links are often unreliable and asymmetric (even for static nodes)

"The only source of packet loss is collision"

- packet error does not mean collision
  - Coexistence: What if there are other people on the earth ????
  - Link quality change
- It is often counterproductive to retry immediately
  - At least on same channel
- There are other techniques than retry to correct errors
- Hidden / exposed terminal



Source: V. Turau et al., INSS 2006

A few words about energy

- sources of energy waste at the MAC layer:

idle listening ➔ listening when no data is available

overhearing ➔ listening to data dedicated to others

oversending ➔ emitting while there is no receiver

collisions ➔ two parties are sending at the same time

protocol overhead ➔ data that is not directly used for the application

“Broadcast is for free” / “Energy ~ to number of packets & their size”

- Broadcast means all nodes must be synchronized in time (and frequency)
  - Synchronization is not free
- Packet transmission means synchronization between sender and receiver(s)
  - There is an overhead per packet (can be large)
  - It varies with sending interval
- Turning off nodes for long periods of time
  - Introduces long latencies
  - There are other techniques (e.g. preamble sampling)

In addition

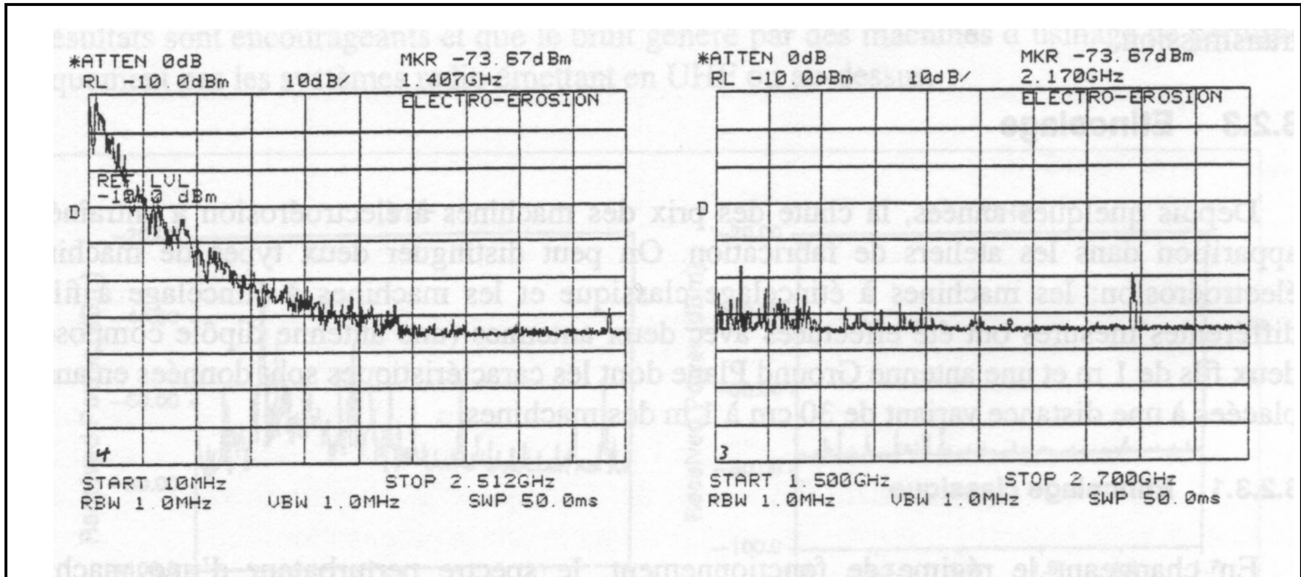
- Severe resource constraints
  - energy, bandwidth, memory size, processing
- Network dynamics
  - Nodes come and go, link go up and down
- Scalability (along number of nodes, traffic, error
- Multiple traffic requirements
  - periodic, sporadic, critical, non critical, …
  - Often unbalanced (to sink)
  - and also changing with time
- Regulations (e.g. ETSI)
- Dependability (many sources of failure)

Context - radio transmission

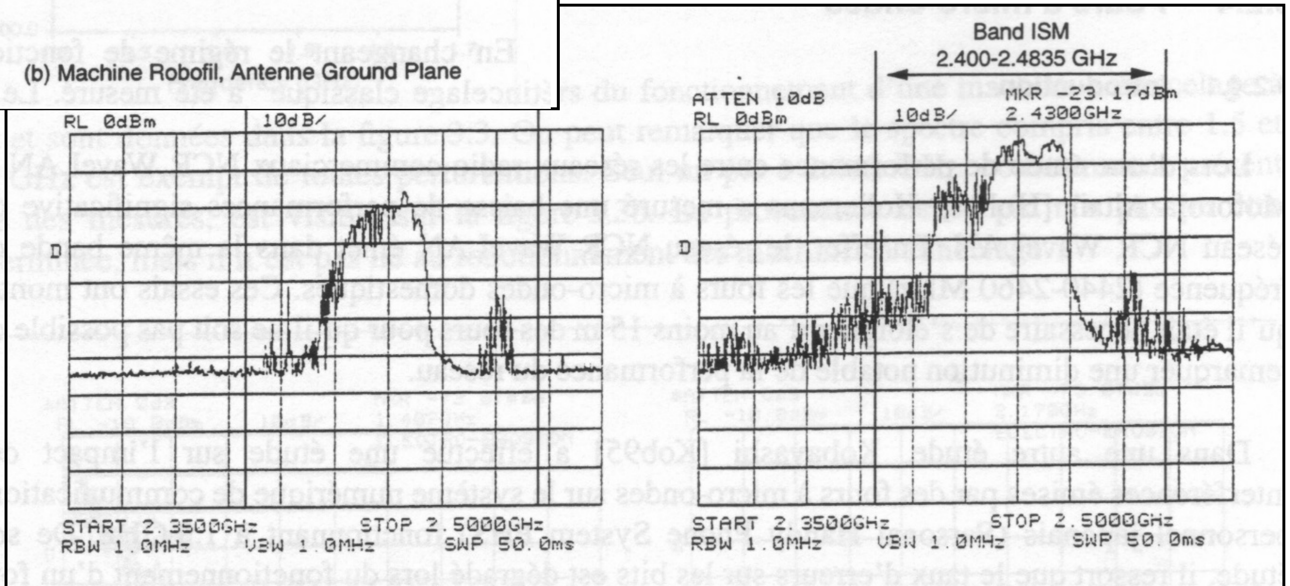
- higher BER
- lower signalling rate
- limited possibility to detect collisions
- low spatial reuse
- prone to interference
- lower distances
- security concerns
- remote powering

- radio transmission
  - fading
  - incompatible regulations
  - free use of ISM bands
  - higher cost
  - longer turn on and switching times
  - hidden terminal effect
- light transmission
  - line of sight
  - sensitive to heat
  - health concerns

Noise Sources



(a) Machine Robofil, Antenne dipôle (b) Machine Robofil, Antenne Ground Plane



(a) Porte complètement fermée (b) Porte entrouverte de 1.6 mm

source: Ph. Morel, 1996

Implications of wireless transmission properties

- MAC
  - master-slave (switching time ⇒ longer timeouts)
  - bus arbiter (hidden node ⇒ limitation in broadcast, reliable detection of silence ⇒ BA redundancy)
  - tokens (hidden node ⇒ token loss, switching time ⇒ longer timeouts)
  - virtual token (reliable detection of silence ⇒ token passing)
  - CSMA (no collision detection ⇒ use timeouts)
  - TDMA (switching time ⇒ longer gaps)

Implications of wireless transmission properties (2)

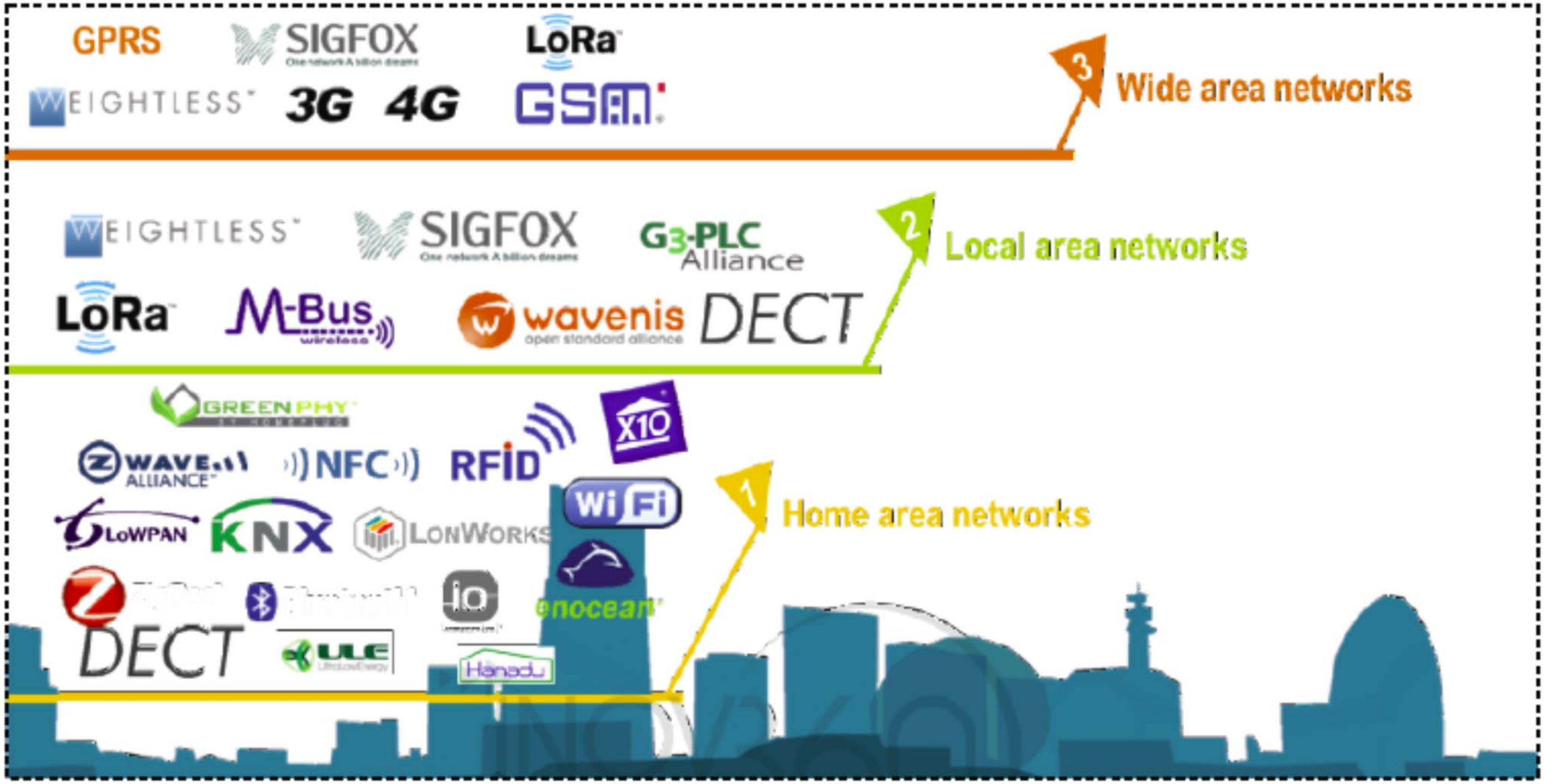
- Error recovery
  - immediate retransmission
    - lower bandwidth, impact on higher layers
  - no immediate retransmission (cyclic transmission)
    - likelihood that errors will last
  - use forward error correction codes to lower apparent FER

source: Ph. Morel, EPFL 1996

Time and Networking

- Several layers play a role in QoS
- Physical layer: robustness
- Data Link Layer: error detection/correction & guarantees at MAC and ack at LLC
- Network: classes of traffic
- Transport layer: retransmission schemes
- Application: interaction model

Wireless landscape

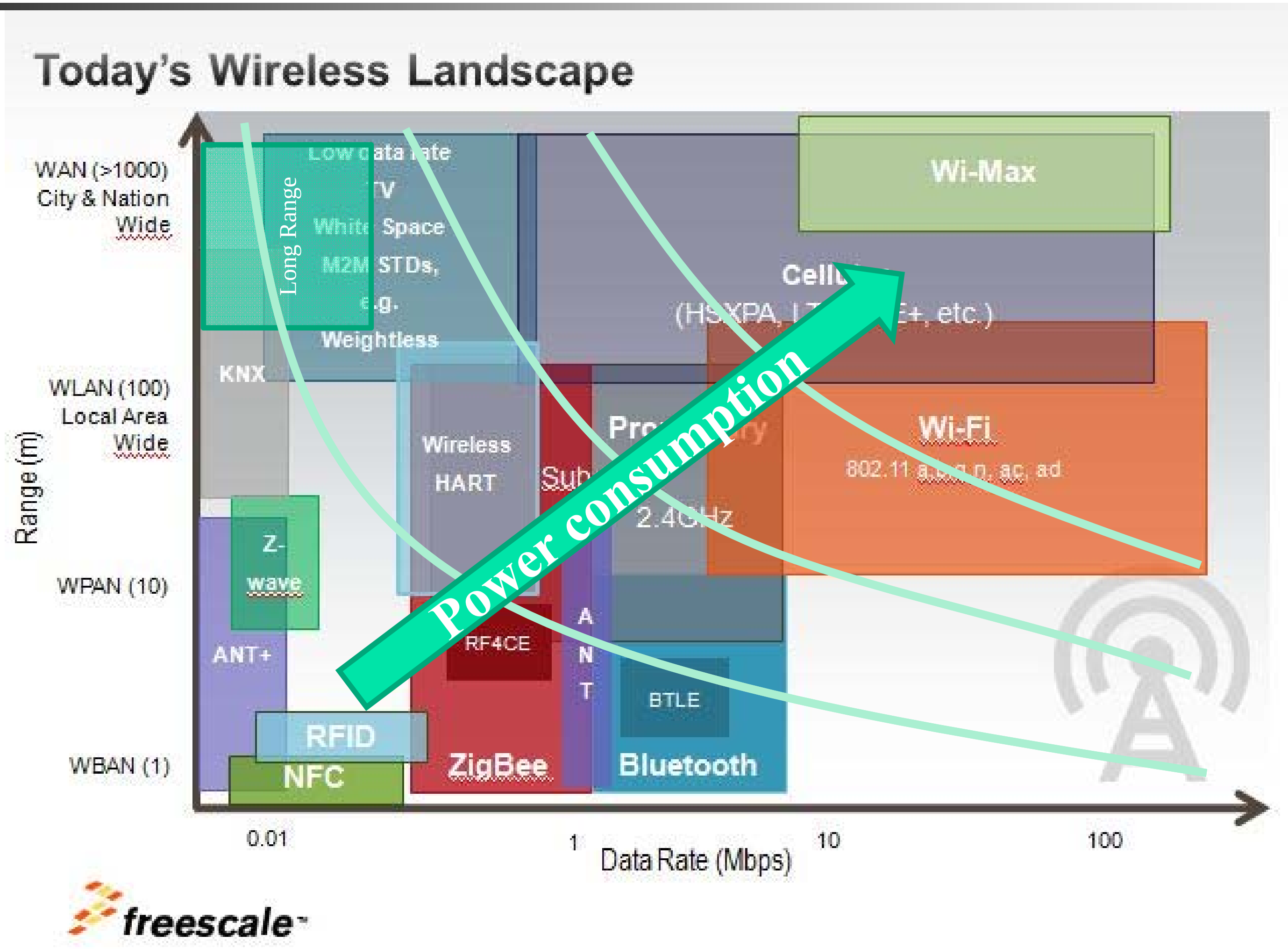


See also: http://literature.cdn.keysight.com/litweb/pdf/5992-1217EN.pdf?id=2773109

Source: Xebia, P. Antoine, S. ben Fredj

EPFL

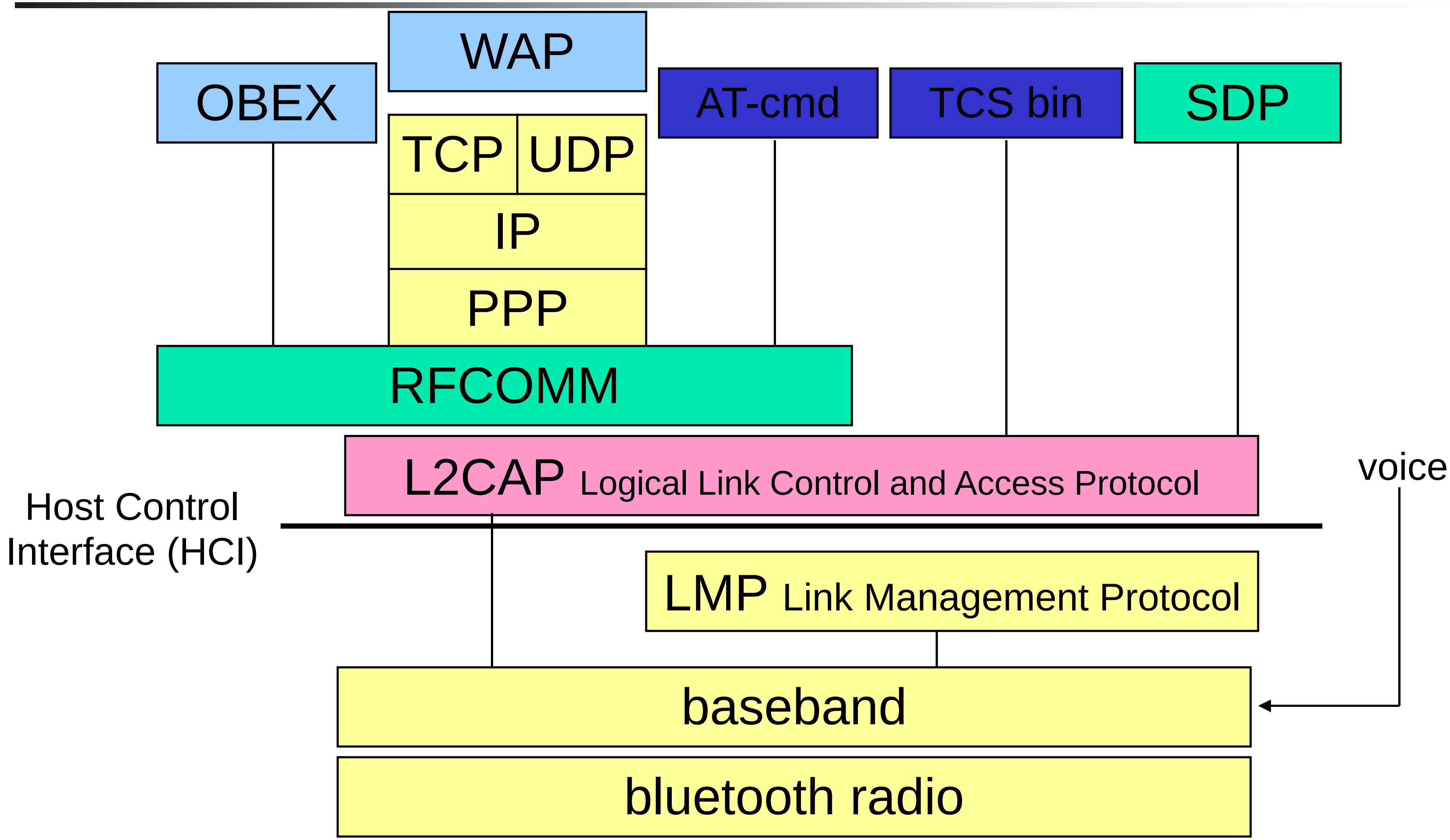
Wireless Landscape (2)



Bluetooth (IEEE 802.15.1)

- open specifications
- data and voice communication
- ISM band 2.4 GHz / FH-CDMA (1600 hops/s, 79 ch.)
- power 1mW (10m) option for 100mW
- cells with max. 8 participants (1 master - 7 slaves)
  - max. 3 voice communications (from/to master)
  - or 1 voice communication and 1 data communication
  - or 723.2+57.6 kb/s asymmetric, 433.9 kb/s symmetric
- TDD, 1 Mbit/s raw bit rate

Bluetooth Architecture

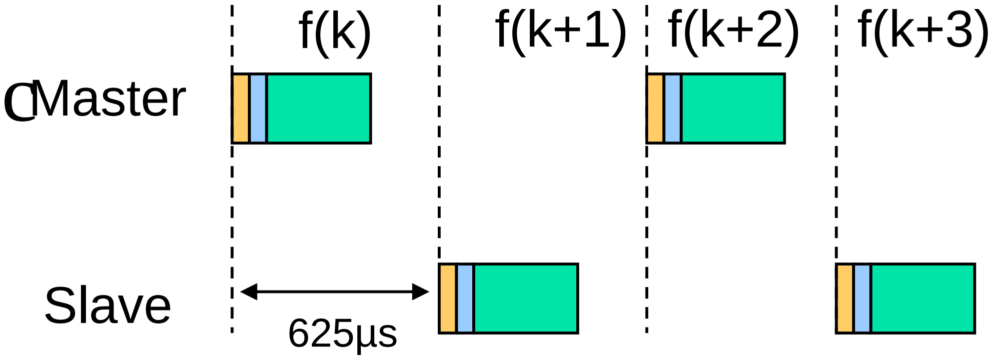


Bluetooth - MAC

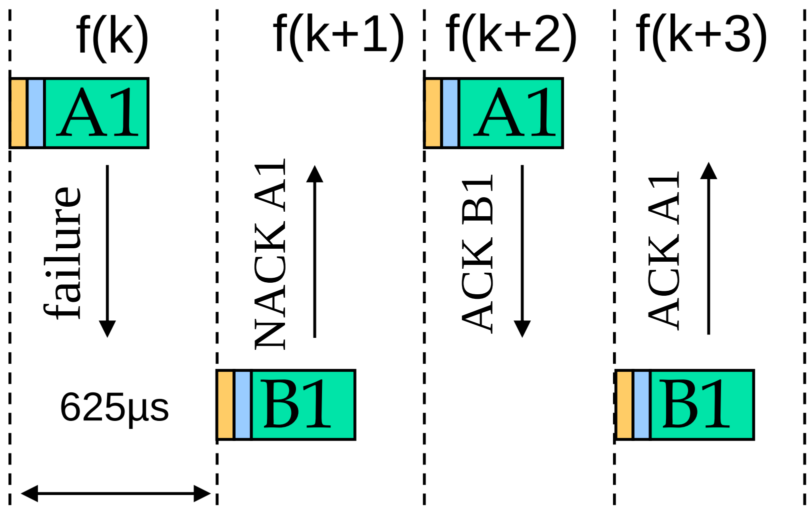
- Frequency hopping spread spectrum (1600 hops/s)
  - hopping sequence based on master identity and clock phase
  - around 23 hours duration
- TDD (Time Division Duplexing) full duplex
  - each 625µs window is used alternately by the master and a slave (frequency hop at each new slot)
  - master-slave communication (request - response)
  - Master is the one that initiated the connection

Bluetooth - TDD

- Master-slave asynchronous traffic

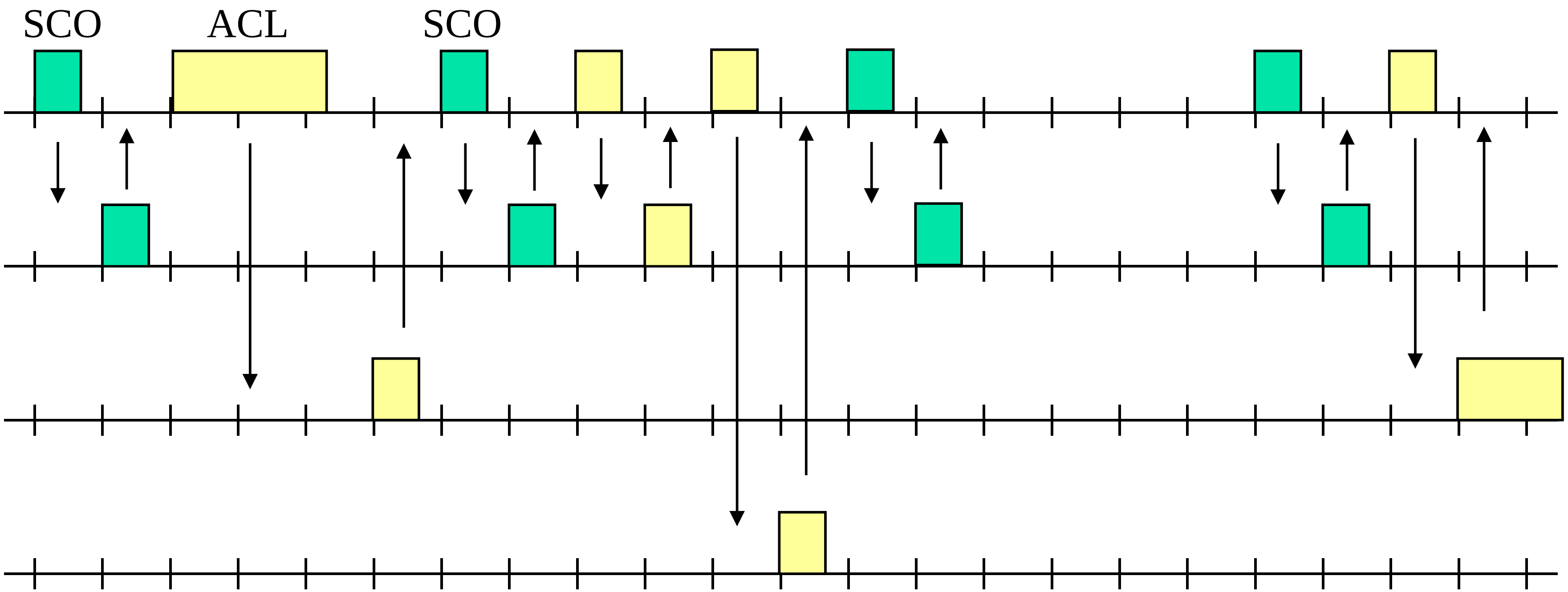


- Error recovery

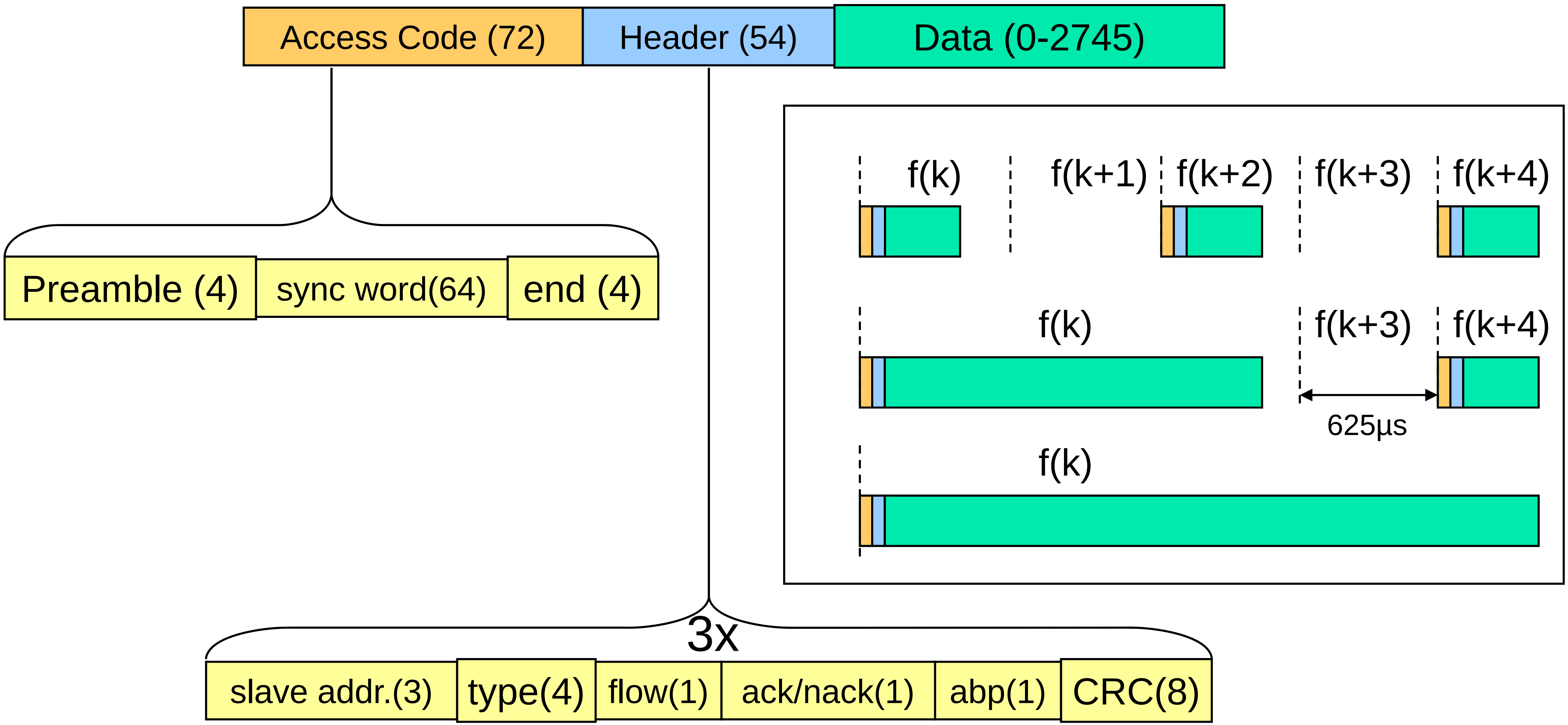


Bluetooth – Synchronous Traffic

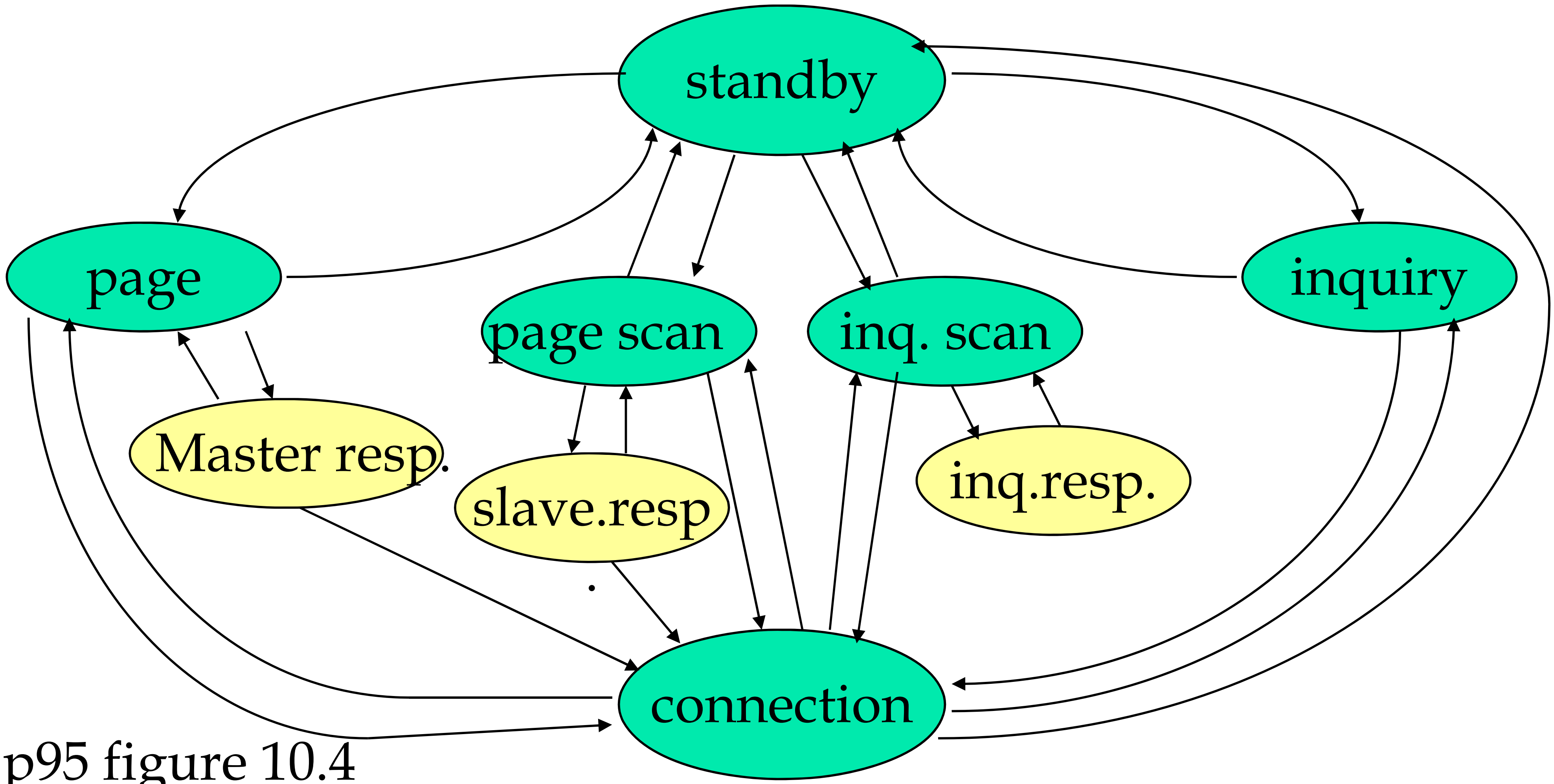
- When used synchronous (SCO) traffic is regular
- Asynchronous (ACL) traffic is interleaved



Bluetooth - packets



Bluetooth - States



p95 figure 10.4

Bluetooth –scheduling

- Local inside node
  - Seems to be FIFO
- Intra piconet
  - None specified
- Inter piconet
  - None specified

Bluetooth - Reduced Traffic Modes

- Sniff mode
  - slave needs to listen only at Tsniff interval
  - each time listens during Nsniff_attempts slots
  - each time, it receives a packet the listening time may be extended
- hold mode
  - does not handle ACL traffic for a given duration
  - SCO traffic is still supported
  - returns to normal mode after the negotiated duration
  - may be used by a station to participate to another piconet

Bluetooth - Reduced Traffic Modes (2)

- Park mode. In this mode, a slave
  - no longer participates in the piconet traffic
  - remains synchronised with the master (master broadcasts a beacon at regular intervals)
  - gives up its active member address (AM_ADDR) and receives 2 new addresses
    - Park Mode address (PM_ADDR)
    - Access Request address (AR_ADDR)
  - may be unparked by master (indicated in beacon)
  - may request to be unparked (access window after beacon)
  - virtually no limit in number of parked slaves

Bluetooth - private protocols

- LM (Link Management)
  - authentication and ciphering / parameter negotiation
  - controls power mode
- L2CAP (Logical Link Control and Adaptation Prot.)
  - adaptation to higher layer protocols
  - segmentation / re assembly (max. 64 Kbytes)
  - connection oriented / connection less services (async. Data)
  - multiplexing and group abstraction
- SDP (Service Discovery Protocol)
  - information on device capability

Bluetooth - Piconet & Scatternet

- Piconet
  - group of max. 8 participants (1 master, max. 7 slaves)
  - a station may be master in a single piconet at any given time
- Scatternet
  - set of piconets in the same geographical area
  - a station may pertain to more than a single piconet
    - must synchronize alternately of all piconets
    - HOLD allows to leave temporarily a piconet

Bluetooth - Link Manager

- Setup, control and security of links
- offers services to
  - authenticate and pair devices
  - setup encryption
  - switch role (master-slave)
  - change mode (park, sniff, hold)
  - manage paging
  - manage SCO links
  - control power
  - supervise link

Bluetooth - Link Manager (2)

- LM messages have higher priority than user data
- Max. response time 30 seconds
- Messages are always single slot packets
- First 2 bits in header indicate LM PDUs
- Flow bit = 0 (ignored)
- 1st byte of body = transaction id. (1), opcode (7)
  - Id = 0, if transaction initiated by master (=1 by slave)
- PDU sent alone (DM1) or in voice packets (HV1)

Bluetooth - Security

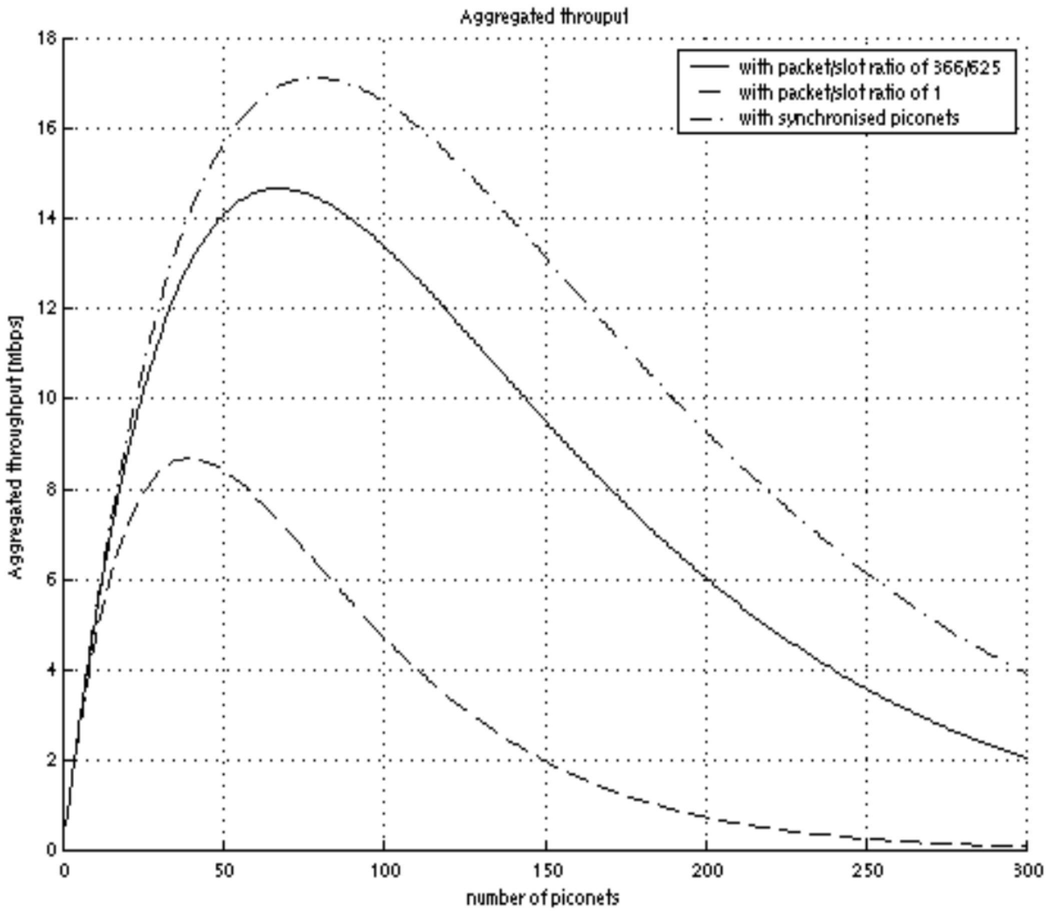
- Authentication
  - based on challenge-response scheme
  - key can be established on line or pre owned
- Encryption
  - can be used or not
  - key size can be negociated

Interference between Bluetooth piconets

- Bluetooth
  - Frequency Hopping, 79 Frequencies
  - Hop every 625 us, Packet length 366 us
- Assumptions
  - collocated piconets
  - interference => packet loss

Aggregated Throughput

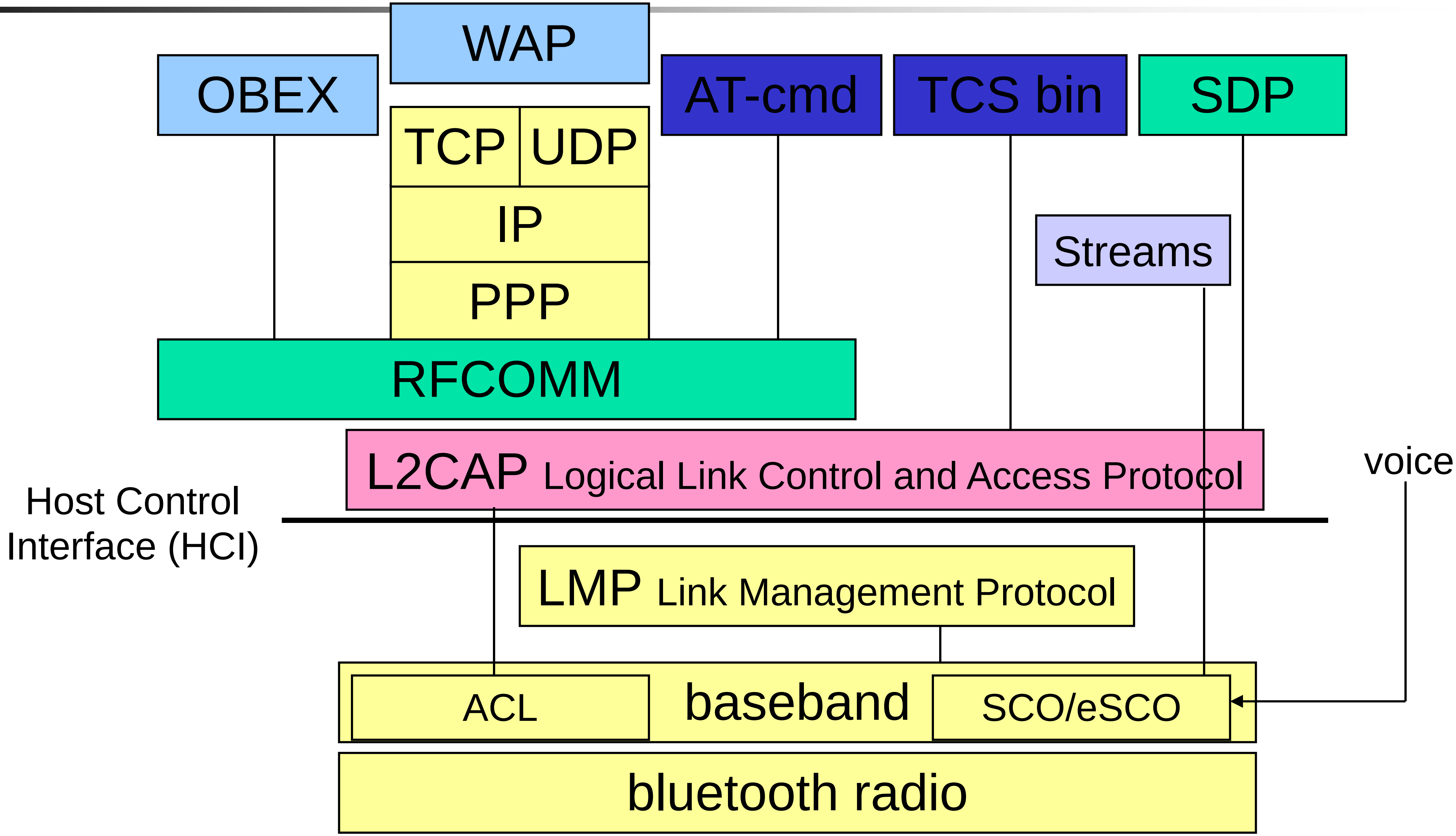
$$S_a(n) = n \cdot P_s(n)$$
$$= n \cdot a^{n-1}$$



Bluetooth - Pros and cons

- Pros
  - low interference (microwave ovens ?) and fading
  - no planning, low cost, authentication and encryption
  - power management possible
  - device discovery protocol
- Cons
  - point to multipoint, short distance
  - no real-time capability for data
  - limited capacity (# devices, throughput)
  - long connection time (up to 10.24 s)

Bluetooth - extensions



Bluetooth – Extensions (2)

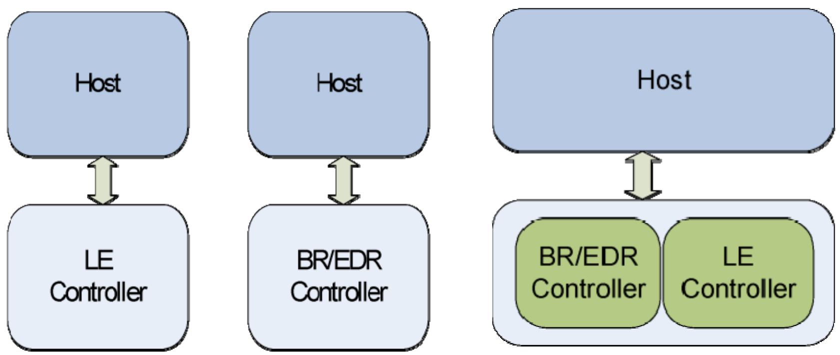
- Version 1.2 (2003)
  - Extended SCO links
  - QoS
  - Better flow management (windows)
- Version 2.0 (2004)
  - 3 Mbit/s
- Version 3.0 (2009)
  - Support for high speed alternate physical layer (802.11)
- Version 4.0 (2010)
  - Low energy version (BT Low Energy / BT Smart)
- Version 5.0 (2016)
  - Long range, 2 Mbit/s for LE, high duty cycle, BLE Mesh (2017)

Bluetooth - QoS

- On SCO and eSCO links
  - Constant bit rate / content is free (not managed by Bluetooth)
    - Management may have a higher priority
  - Error correction may be performed using retransmission
  - Only one link per slave
- On ACL links
  - Managed according to « tokens bucket » algorithm
    - Mean throughput with some peaks
  - Only lower priority than SCO and eSCO
  - May be subject to admission control

BTLE objectives

- Targets, principally low-power and low-latency, applications for wireless devices within short range (<50 m)
- To operate more than a year on a button cell battery
- lower power consumption not achieved by nature of the active radio transport, but by design of the protocol to allow low duty cycles, and the use cases envisaged.
- Designed to be lowest cost and easy to implement
- Node types

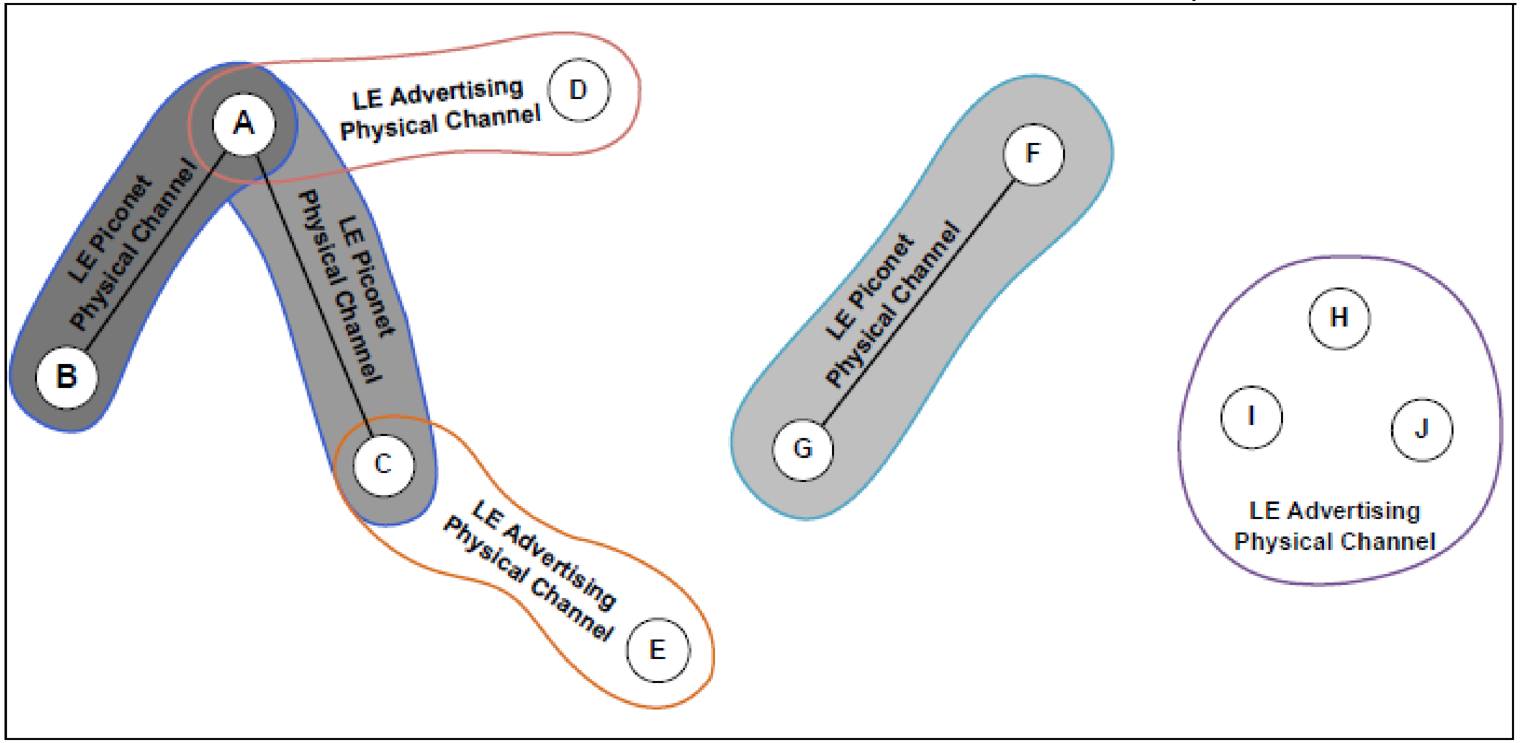


BTLE novelties

- Mostly new PHY
- New advertisement mechanism
  - => ease of discovery connection
- Asynchronous connection-less MAC: used for low latency, fast transactions (e.g. 3ms from start to finish)
  - No carrier sense before transmitting
  - Fast interactions with channel diversity and random waits
  - Connections with regular channel hopping
- New Generic Attribute Profile
  - to simplify devices and the software that uses them.

BTLE - Examples of interactions

Source: Bluetooth specification version 4, 30.6.2010.



- A master with B & C as slaves / F master with G as slave
- D advertizer with A initiator / E scanner with C advertizer
- H advertizer with I & J as scanners

BTLE advertisement

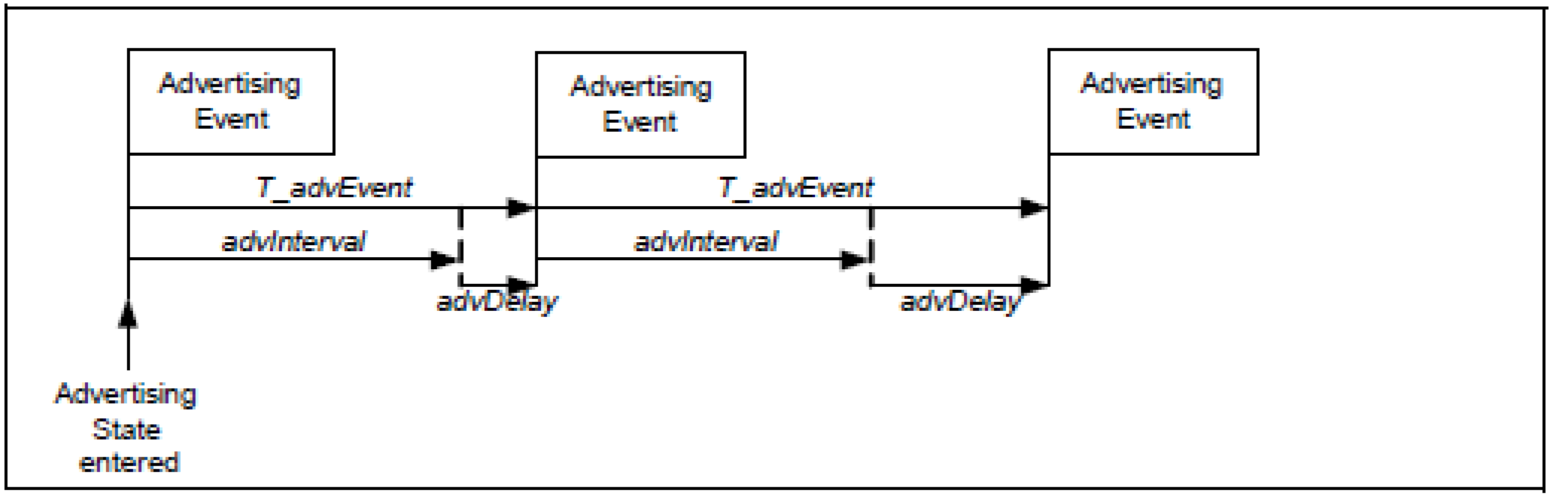


Figure 4.1: Advertising events perturbed in time using advDelay

Source: Bluetooth specification version 4, 30.6.2010.

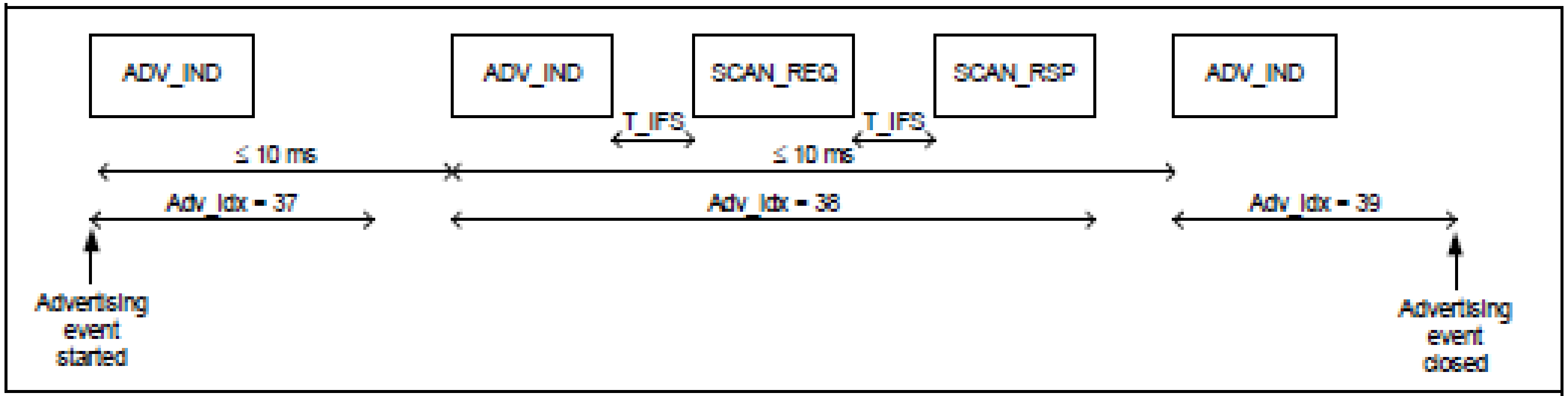


Figure 4.3: Connectable undirected advertising event with SCAN_REQ and SCAN_RSP PDUs in the middle of an advertising event

BTLE connections

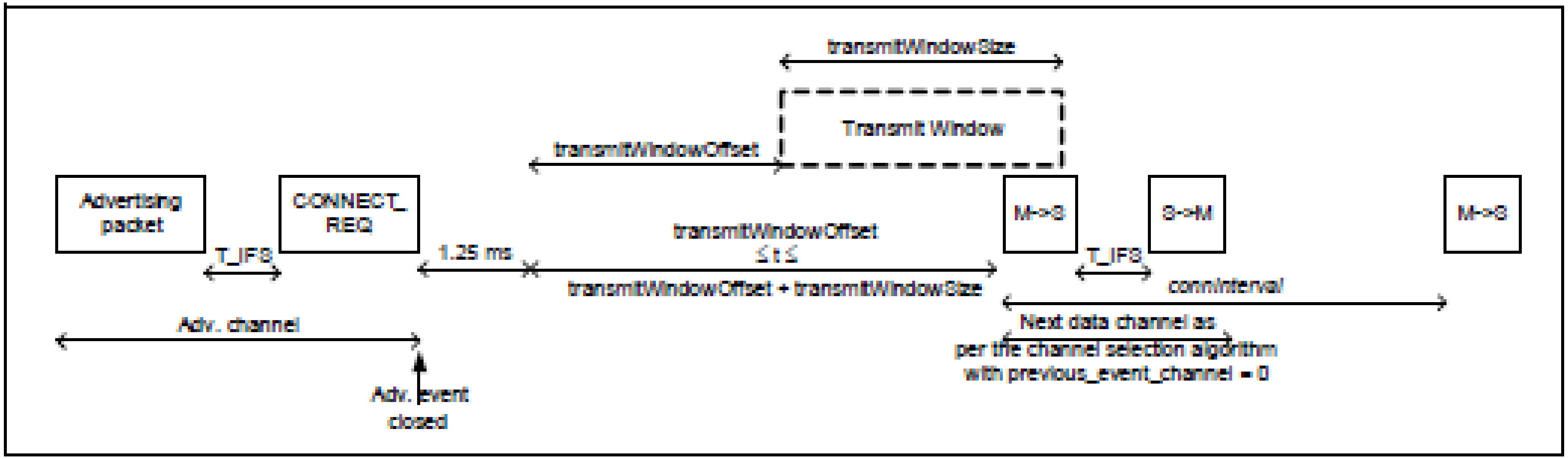


Figure 4.11: Master's view on LL connection setup with a non-zero transmitWindowOffset

Source: Bluetooth specification version 4, 30.6.2010.

BTLE 4.2 (BT Smart) - Analysis

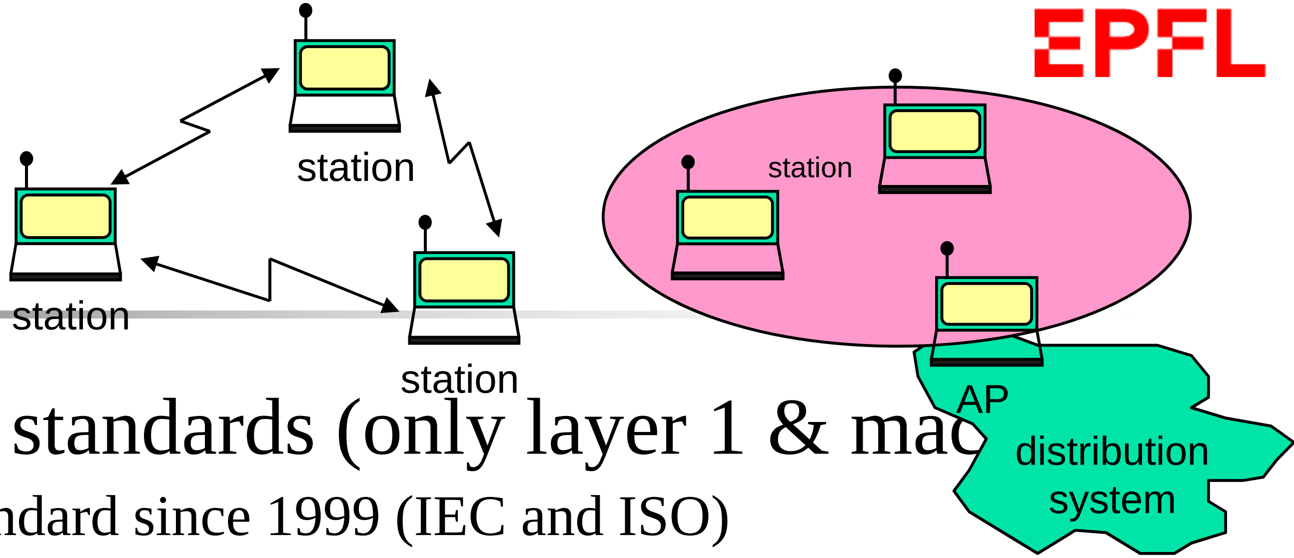
- Pros
  - Targets low energy
  - Good coexistence (channel hopping)
  - Advertisement can be fast (good for spurious interactions)
  - Security
  - Can be implemented on resource limited devices
- Cons
  - No real-time guarantee
  - Limited throughput (max. around 80-90 Kbps)

Bluetooth 5

- 2 Mbit/s for BTLE
- LE long range (300m)
- Connection-less broadcasting up to 255B
- Lower duty cycle

- Bluetooth Mesh (2017)
  - Mesh networking, based on publish-subscribe
  - M. Woolley, S. Schmidt, “Bluetooth mesh networking - An Introduction for Developers”, 2017, https://www.bluetooth.com/bluetooth-technology/topology-options/le-mesh/mesh-tech

IEEE 802.11



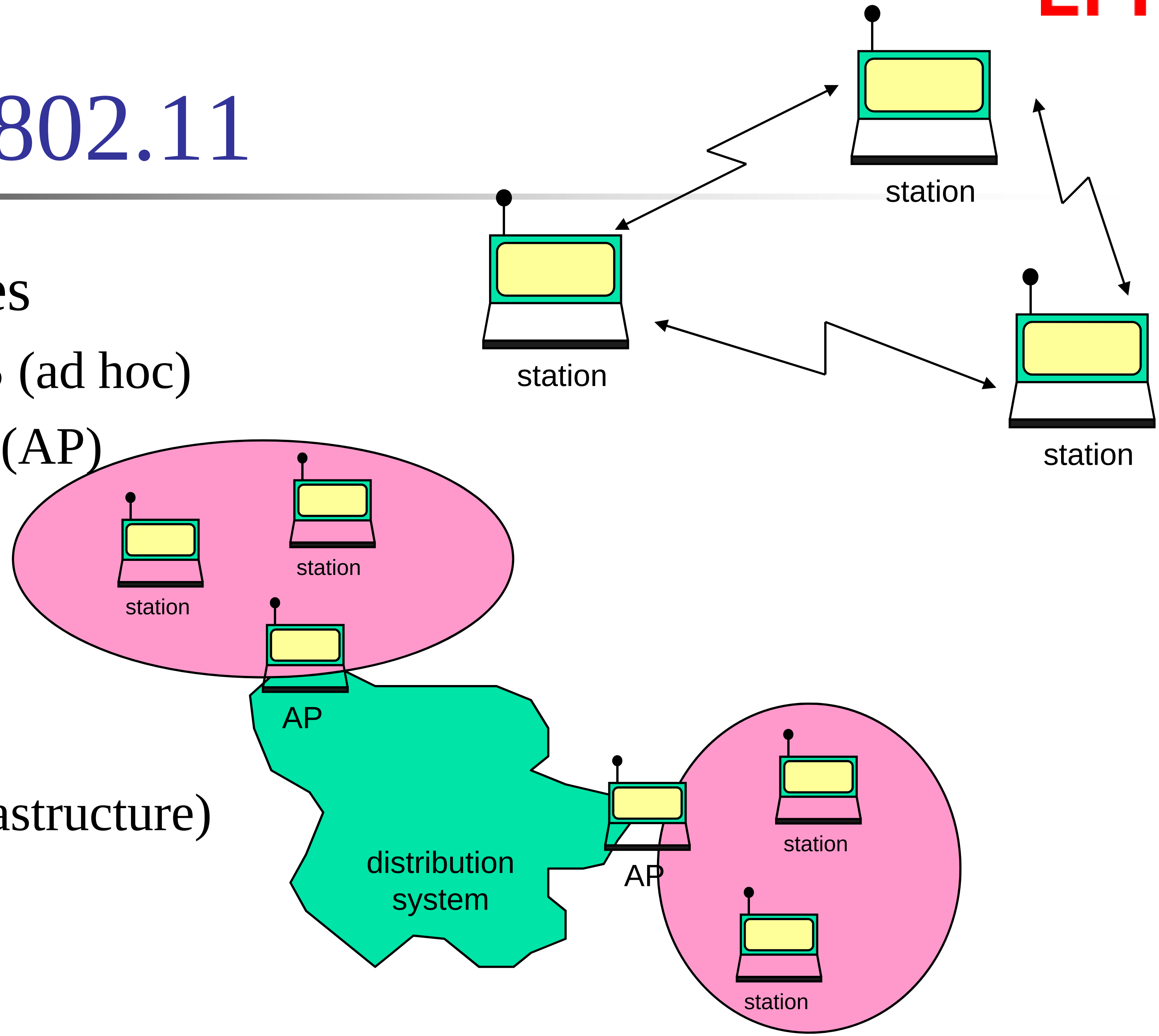
- Part of IEEE 802 standards (only layer 1 & mac
  - international standard since 1999 (IEC and ISO)
- wireless LAN with 2 operating modes
  - station to station without coordination (ad hoc network or DCF)
  - coordinated by a single base station per cell (PCF)
- 3 physical layer options (2.4 GHz radio FH & DS, IR)
  - DS SS(11 chips, 30 MHz between channels) 1, 2, 5.5 and 11 Mbit/s
  - FH SS (79 channels, 3 sets of 26 hopping sequences, > 2.5 hops/s)
    - range (30m indoor, 200m outdoor, 30km directive)
  - IR pulse position modulation, 1 & 2 Mbit/s, diffuse communication
- MAC: CSMA/CA + contention-less period (PCF)

A large family…

- 802.11b: 2.4GHz band, DSSS, 1,2,5.5,11 Mbit/s
- 802.11a: 5 GHz band, OFDM, up to 54 Mbit/s
- 802.11g: same but in 2.4 GHz band
- 802.11f: recommandations for inter AP protocols
- 802.11i: AES security
- 802.11h/: 5GHz band operations in Europe / Japan
- 802.11e: QoS again
- 802.11n/ac: 135 Mbit/s (2.4GHz) / 780 Mbit/s (5GHz)
- 802.11ad: 6.75Gbit/s (60 GHz) – 100 Gbit/s planned
- Projects still active (http://www.ieee802.org/11/Reports/802.11_Timelines.htm)

IEEE 802.11

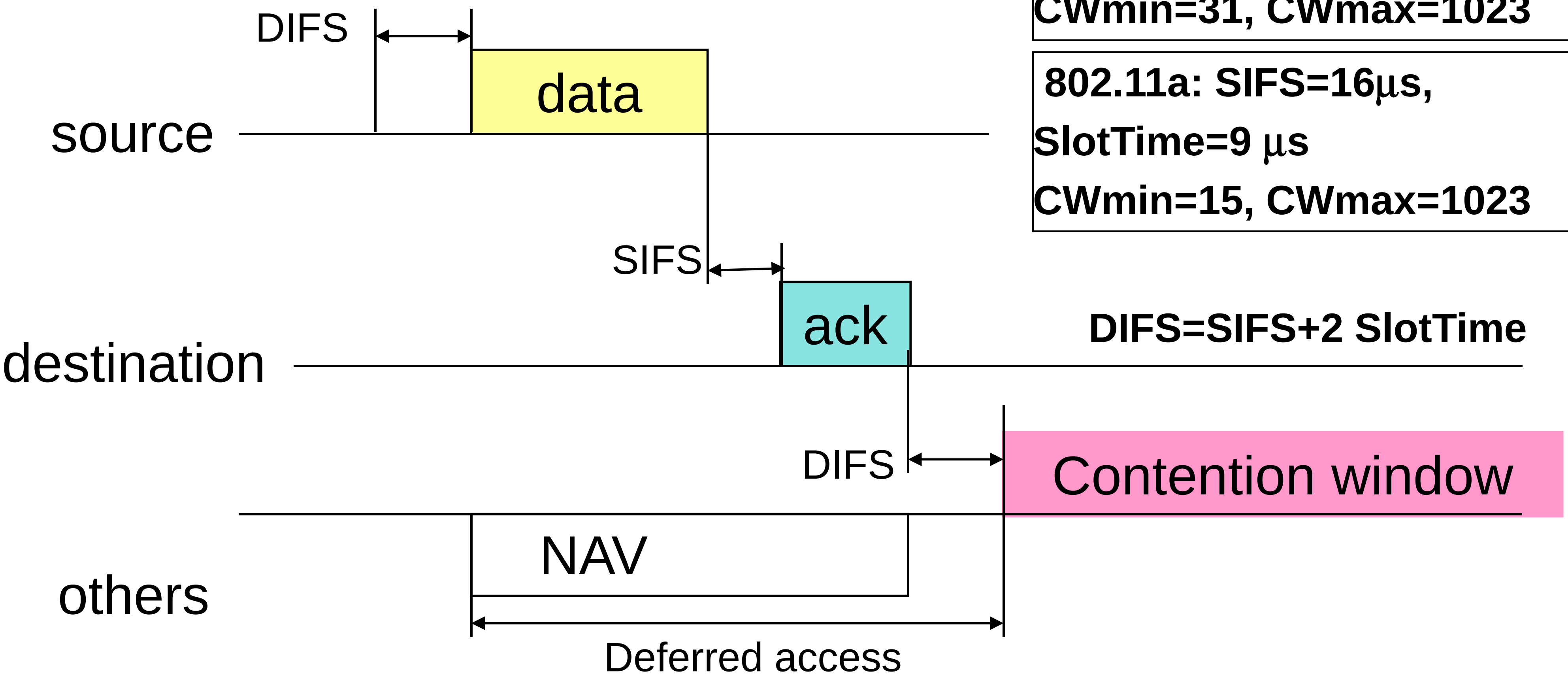
- 3 modes
  - IBSS (ad hoc)
  - BSS (AP)
  - ESS (infrastructure)



IEEE 802.11 - DCF (distributed coordination function)

- Asynchronous transfer without guaranty
- CSMA / CA
  - Physical carrier detection and logical carrier detection
    - Each frame carries time required to transmit remaining traffic
    - Is used to update the Network Allocation Vector (NAV)
  - If no carrier when data arrives, immediate transmission
  - If carrier, wait until no carrier during DIFS
    - Computes backoff interval, start decrementing (freeze decrement when medium busy), and transmits when 0
  - If failure, increment backoff window (up to Cwmax)

IEEE 802.11 - DCF (2)



802.11b: SIFS=10μs,
SlotTime=20 μs
CWmin=31, CWmax=1023

802.11a: SIFS=16μs,
SlotTime=9 μs
CWmin=15, CWmax=1023

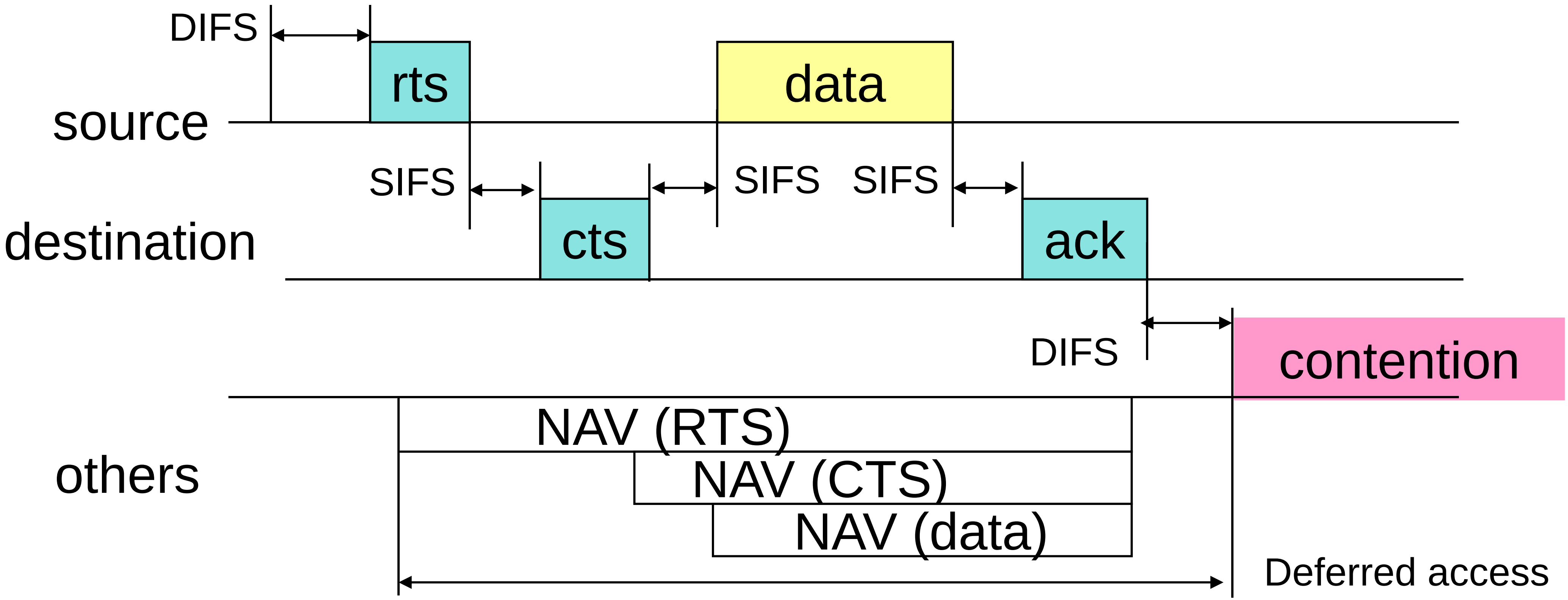
DIFS=SIFS+2 SlotTime

Transmission of long frames (MPDU)

- Maximum size 2436 bytes (20ms @1Mbit/s)
- Loss of efficiency because collision may not be detected
  - RTS/CTS mechanism
- High probability of corruption
  - Segmentation to avoid retransmission of the whole frame
  - The whole frame is transmitted without other transmissions (blocking time)

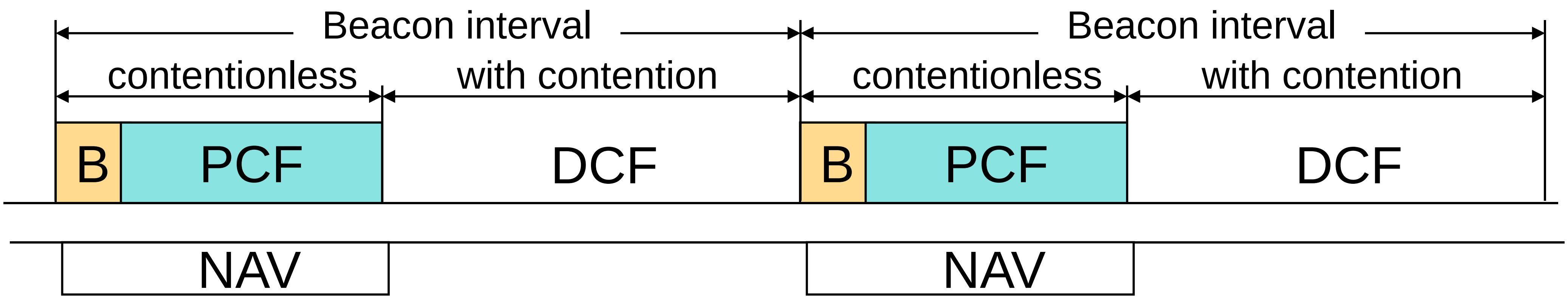
IEEE 802.11 - DCF (3)

- Collisions cannot be detected -> avoid sending long frames



IEEE 802.11 - Point Coordination Function

- Optionnal, Connection oriented
- Supports transfers without contention
- Based on a special function "point coordinator" handled by the access point AP



IEEE 802.11 - PCF

- Transfers according to a polling list
  - AP sends beacon after PIFS (SIFS+SlotTime)
  - AP sends poll to STA (optional data & ack of previous resp.)
  - STA responds with ack and optional data
- Supports station to station transfers
  - STA response is addressed to another STA (not AP)
  - Other STA must ack.
- CFP duration
  - Min: 2 times tx of the max. duration MPDU + beacon + CFP end
  - Max: beacon interval – tx time of the maximum duration MPDU
- The beacon may be delayed if a frame under transmission

IEEE 802.11 QoS limitations

- By QoS, we mean throughput, delay, jitter
- DCF
  - Best effort, no traffic differentiation
- PCF
  - Centralized traffic (even if some STA to STA possible)
  - Strong requirements on response time (SIFS ~10μs)
  - Jitter in beacon because STA may transmit across TBTT (Target Beacon Tx Time)
  - Transmission instant of polled STA variable

Possible improvements

- Station based differentiation in DCF
  - AC scheme: different backoff increase, DIFS & frame length
  - Distributed fair scheduler
  - Virtual MAC
- Station based differentiation in PCF
  - Priority based polling
  - Distributed TDMA

Possible improvements (2)

- Queue based service differentiation in DCF
  - EDCF, AEDCF
- Queue based service differentiation in PCF
  - HCF
- Error control based schemes
  - Selective repeat ARQ
  - Go back N ARQ
  - FEC
  - Hybrid FEC-ARQ

802.11e

- Included in 2007 version
  - QAP – QoS Access Point
  - QBSS – QoS Basic Service Set
  - QIBSS – QoS Independent BSS
  - QSTA – QoS Station
  - nQAP, nQBSS, nQIBSS, nQSTA – non QoS …
  - QSTA may associate to nQAP but will not provide any QoS

IEEE 802.11e (2)

- Core QoS facilities
  - EDCA (Extended DCF Access)
    - 4 access categories
  - HCCA (HCF controlled access)
    - 8 traffic streams
- Optional QoS facilities
  - Block Acknowledgement function,
  - Direct Link Set-up (DLS) and
  - Automatic Power-save Delivery (APSD)
  - Contention Free Period (CFP)

IEEE 802.11e (3)

- Introduces a Hybrid Coordination Function (HCF)
- 2 medium (channel) access mechanisms
  - Contention-based channel access (Ext. DCF access or EDCA)
    - Up to 4 backoff entities in a given station (queues)
  - Controlled channel access (HCF controlled access or HCCA)
- 2 periods, CP & CFP: EDCA in CP, HCCA in both
- 1 station coordinates a QoS supporting BSS (QBSS)
  - Hybrid coordinator (HC) (usually the QAP)
- QoS data frames carry the size of the waiting queues
- Traffic may be subjected to admission control (per class/stream)

Extended DCF Access (EDCA)

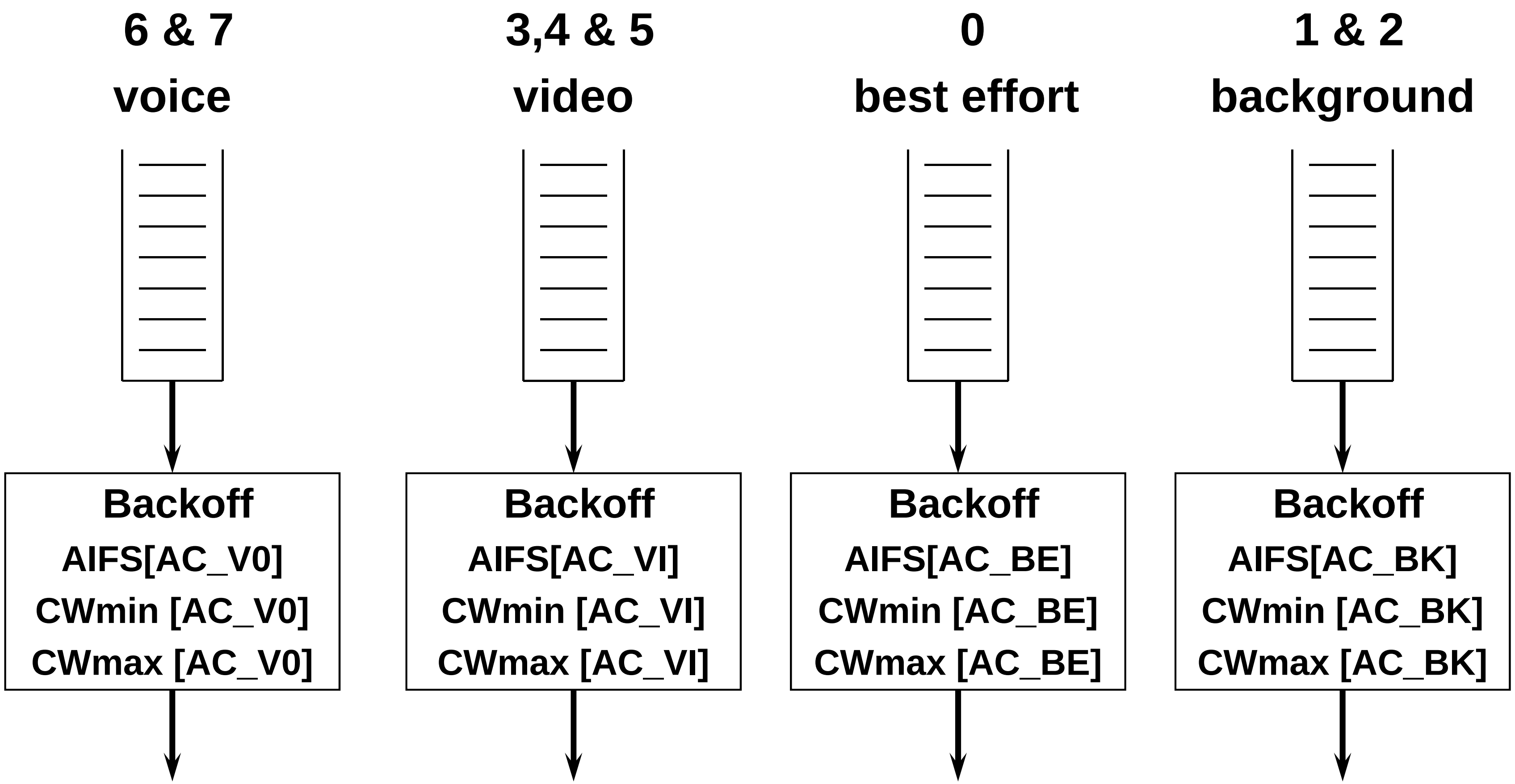
- May be used without any AP
  - Provides QoS support in ad-hoc mode (IBSS)
- Traffic differentiation based on
  - Amount of time STA senses channel idle before backoff or transmission
  - Length of contention window
  - Duration during which a station may transmit once it has acquired the channel
- May be subjected to admission control (in QBSS)

EDCA - Backoff entity rules

- Must not use radio resource more than a given limit TXOP
  - EDCA –TXOP is given by the HC (within beacons)
- Cannot transmit across TBTT (not true for HCCA)
- A frame can be sent to any other backoff entity (not only AP)
  - Requires establishment of a direct link using DL protocol
- MSDU maximum life time (dropped wo being tx)

Backoff entities in a station

802.1D priorities

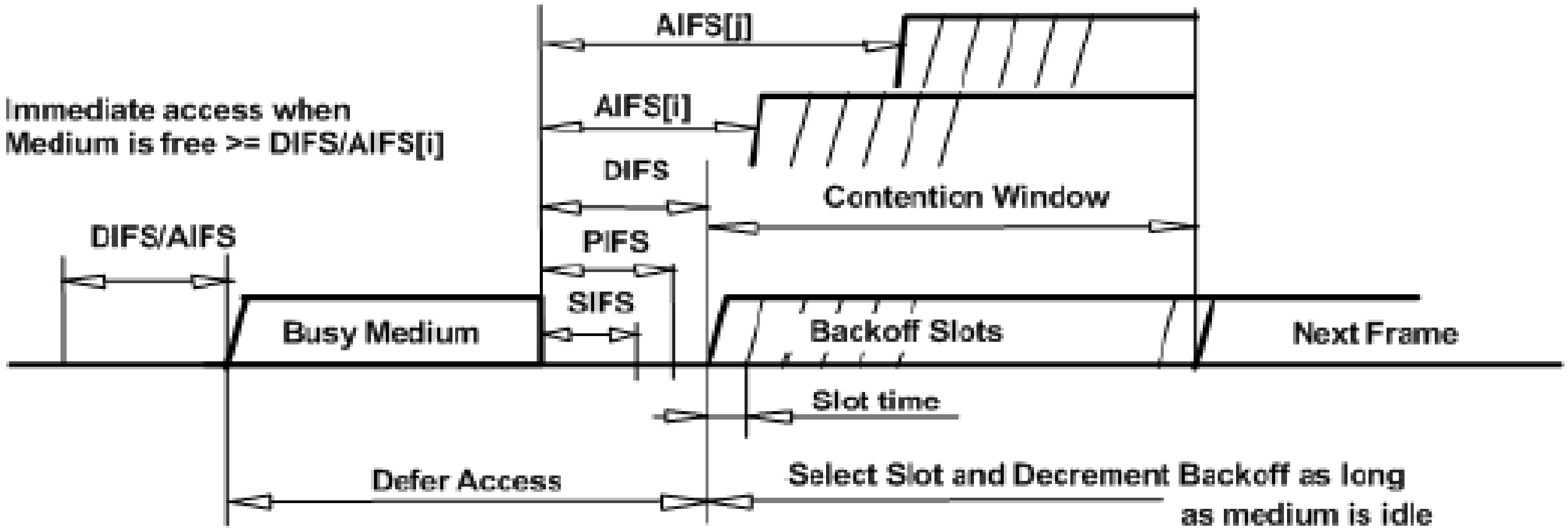


In case of parallel access on the same slot, the highest priority wins

EDCA per Access Category

- Values of AIFSN, Cwmin, Cwmax and TXOPLimit are announced by the QAP in beacon/association frames
  - Fixed in QIBSS
  - QAP may use a different set of values for itself
- AIFS[AC]=AIFSN[AC]*SlotTime+SIFS
- AIFSN >=2 for non QAP
- When frame arrives at empty queue and medium has been idle for longer than AIFS[AC]+SlotTime, it is transmitted immediately
- If medium busy, wait until free and then differ for AIFS[AC]+SlotTime

Relationship between Interframe gaps



EDCA – additional rules

- Once STA has gained channel
  - If may send a sequence of consecutive MSDUs
  - As long as the elapsed time does not exceed TXOPlimit[AC]
- Admission ctrl may be mandatory for some ACs
  - Admission is based on requests (ADDTS) from QSTA to QAP

    QAP responds with average time per period

HCCA

- Allows for reservation of transmission opportunities
  - Based on request from non AP STA (up and down)
    - Must establish a traffic stream by exchanging TSPECs
  - Governed by admission control (vendor dependent)
    - Once admitted cannot be changed by QAP unless a new request is made
  - Traffic scheduled by HC collocated to QAP
    - From STA: using polls set according to QSTA requests
    - From AP: according to actual traffic

HCCA (2)

- The HC may obtain a TXOP via the controlled medium access
- It may allocate TXOP to
  - Itself by sending MSDU after medium idle for PIFS (no backoff)
  - QSTA by sending QoS CF-Poll under same rule
- New rules remove direct relationship between beacon frequency and polling frequency

HCCA traffic scheduling

- Based on TSPECs

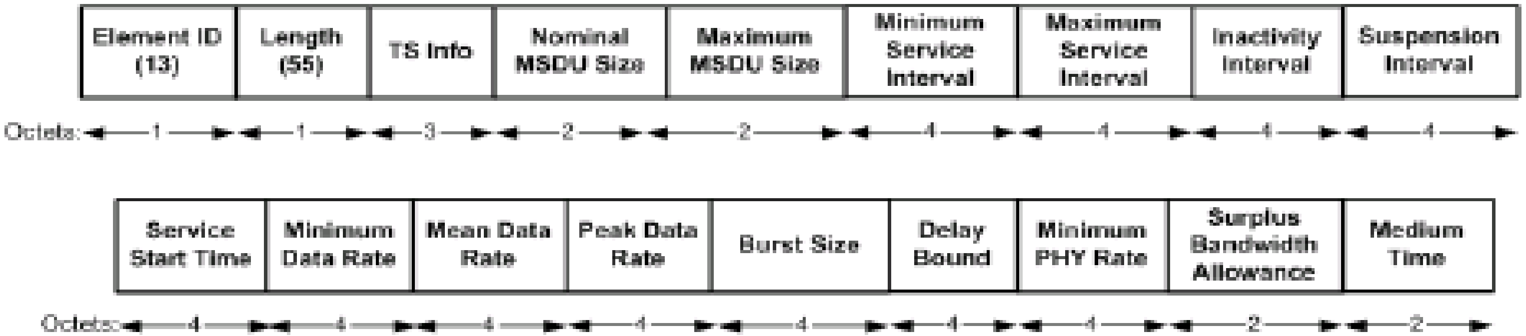


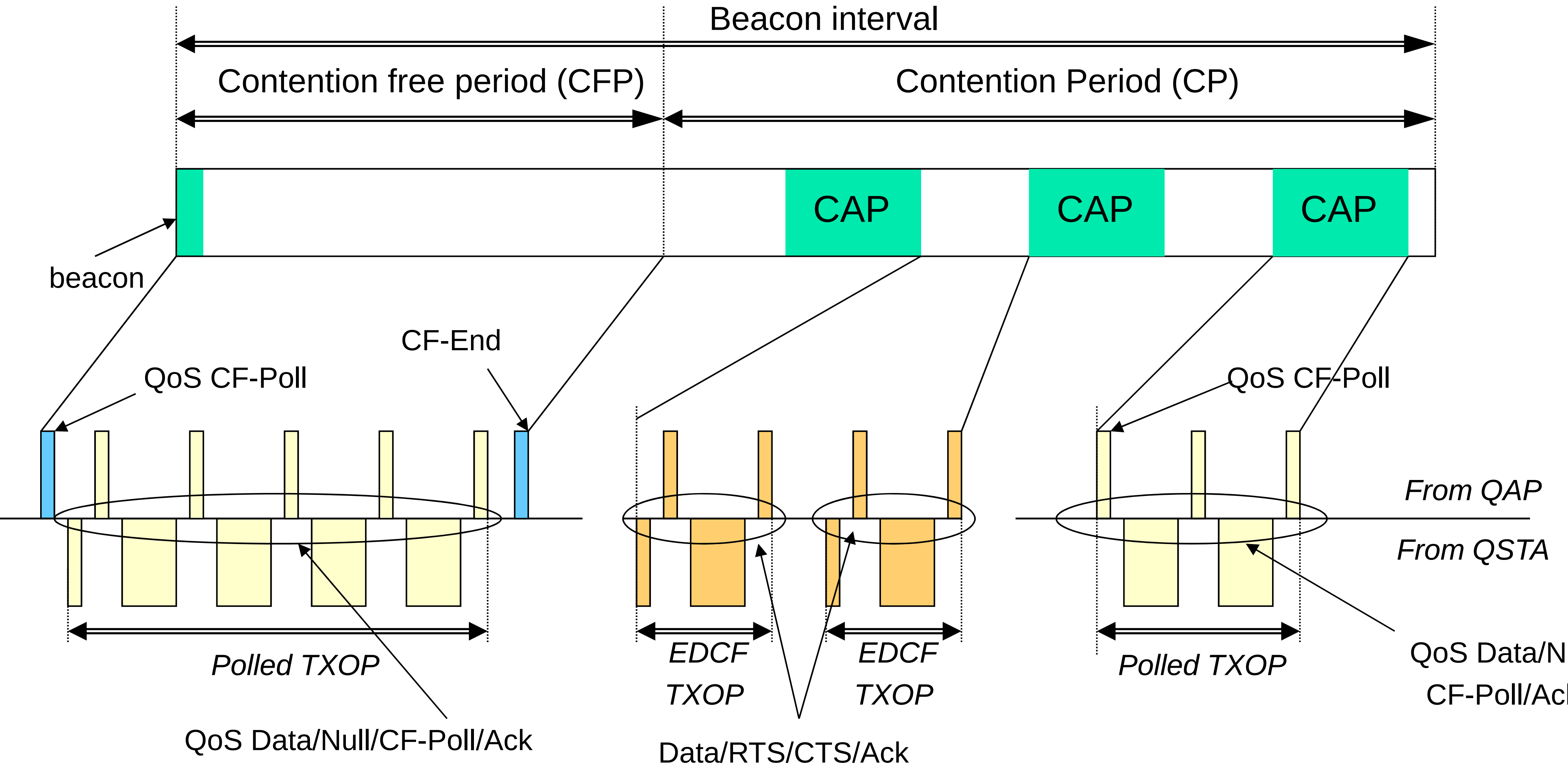
Figure 46.7—Traffic Specification element format

- QAP scheduler Computes duration of polled-TXOP in each QSTA
- Scheduler in each QSTA allocates TXOP for different TS queue according to priority order

Access rules for STA

- During CP
  - Under EDCA rules
  - In response to QoS CF-Poll
- During CFP
  - Only in response to QoS CF-Poll
    - May send multiple frames separated by SIFS as long as elapse time does not exceed TXOPLimit

HCF beacon interval and traffic



CAP: Controlled Access Phase

Other features

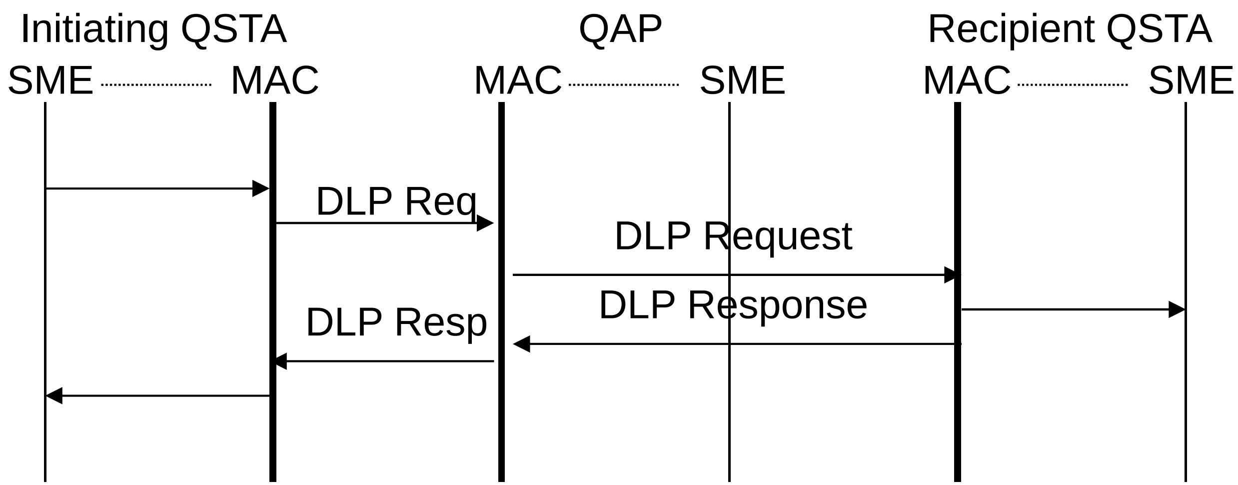
- Block acknowledgment
- Direct Link Protocol
- Power Management
  - Already in legacy standard
  - Traffic indication maps in beacon frames
- Support for time synchronisation (multicast + time since reception of a given part of the frame)
- Broadcast and multicast offered but frames must be sent once at a time
- Piggy back acknowlegments

Block acknowledgement

- Up to 64 data frames can be sent in a row (separated by SIFS) before an ack is sent back
  - Subject to TXOP duration limit
  - May be spread over several TXOPs
- Requires initial setup to check capability and reserve resources (i.e. to store blocks before reassembly)
- Ack can be immediate or delayed
- Applies both to polled TXOPs and HCCA

Direct Link Protocol

- Allows direct QSTA to QSTA transfers in QBSS
- Setup goes through QAP



- Any access mechanism may be used
  - Polled TXOP, EDCA, block ack, Traffic streams

Conclusion

- Bluetooth does not offer any QoS mechanism
  - However, there are a few possibilities starting at V2
- Bluetooth Low Energy
  - Introduced in BT V4.0
  - As name states, done for ultra-low-power devices
  - No temporal guarantees
- IEEE 802.11
  - Offers a number of possibilites for QoS
  - Much more to explore
  - Room is left open for improvements (scheduling traffic)

References

- IEEE 802.11-2012 IEEE Standard for Information technology-Telecommunications and information exchange between systems-Local and metropolitan area networks-Specific requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications
- Q. Ni et al., « A survey of QoS Enhancements for IEEE 802.11 Wireless LANs », J. of Wireless Comm. And Mobile Comp. 4 (5), pp. 547-66, 2004.
- S. Mangold et al., « Analysis of IEEE 802.11E for QoS Support in Wireless LANs », IEEE Wireless Comm. 10 (6), pp.40-50, 2003.

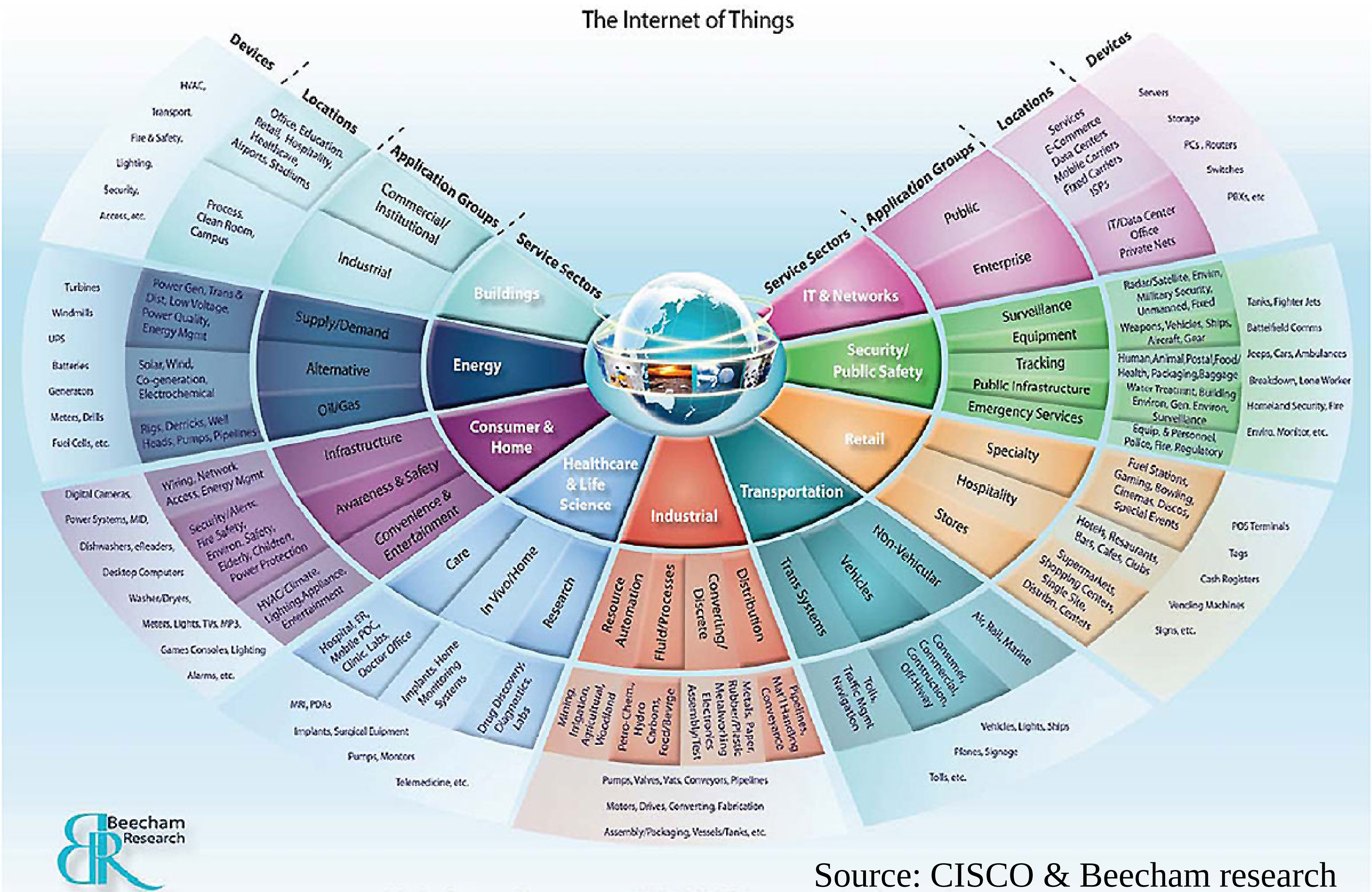
References (2)

- Bluetooth
  - http://www.bluetooth.org
  - J. Haartsen, "The Bluetooth Radio System", IEEE Personal Communications 7 (1), pp.28-36, 2000.
  - R. Shorey, « The Bluetooth Technology: Merits and Limitations », Proc. Int. Conf. On Personal Wireless Communications ICPWC'2000, pp.80-84.

REAL-TIME NETWORKS Wireless sensor networks (aka Internet of Things – IoT)

Prof. J.-D. Decotignie
CSEM Centre Suisse d’Electronique et de Microtechnique SA
Jaquet-Droz 1, 2007 Neuchâtel
jean-dominique.decotignie@csem.ch

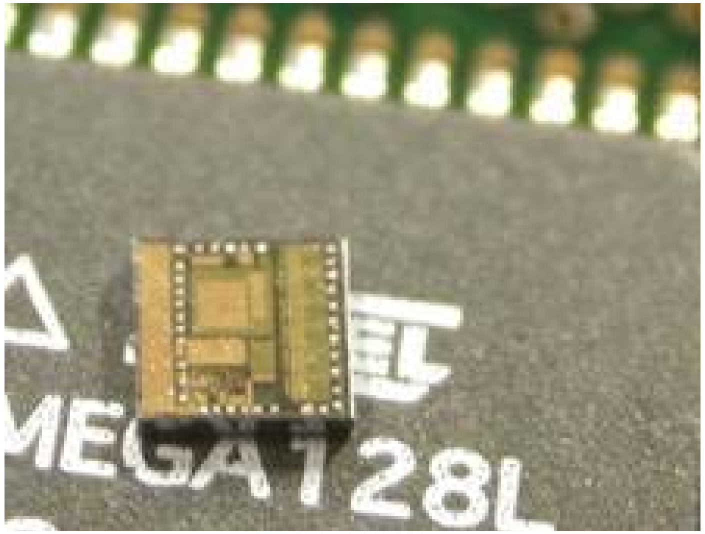

A new device has been found
Device:Airbus A310
Run Auto-configuration?
Start
Cancel

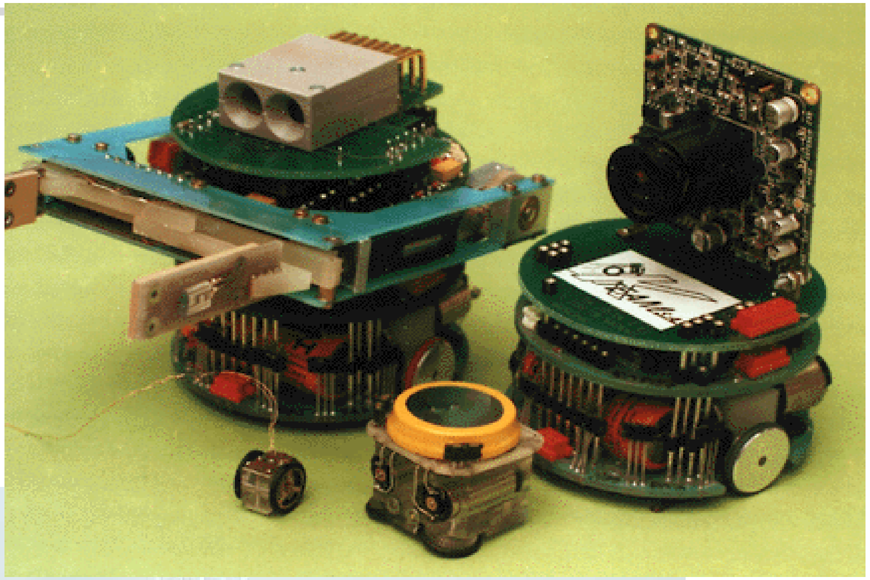


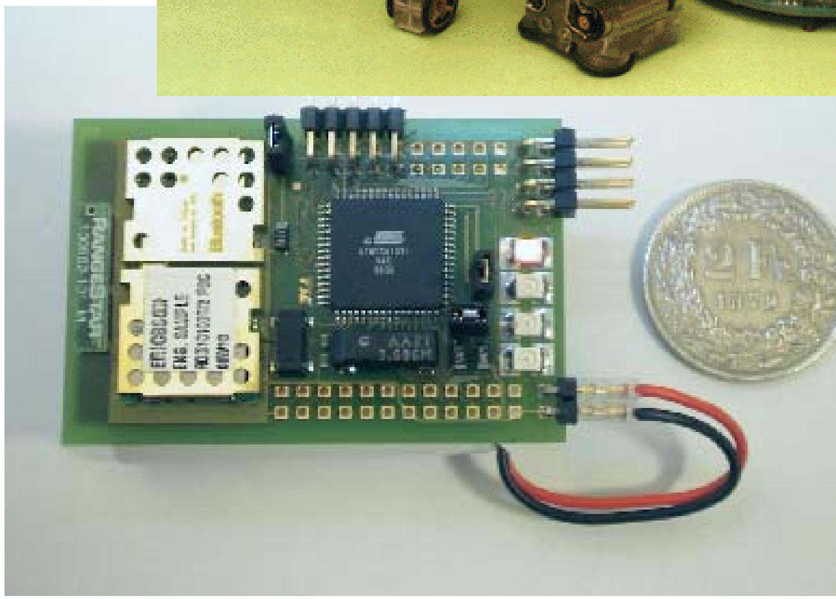
Source: CISCO & Beecham research

tignie

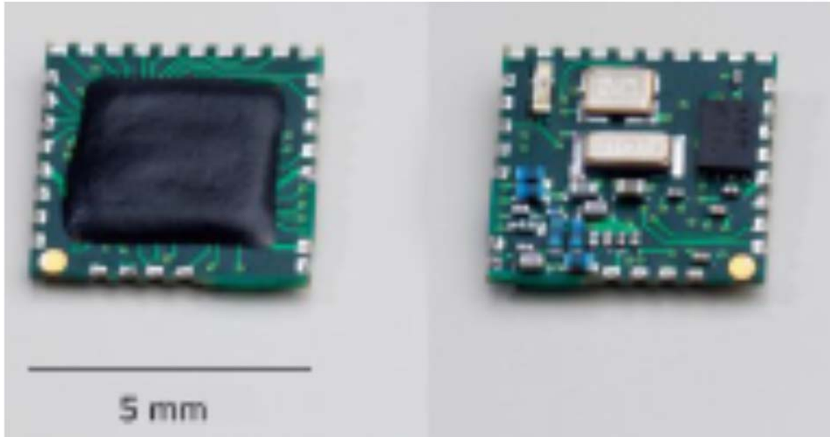
Y2K devices: Motes, BT nodes, Smart-Its, Cyclope, …



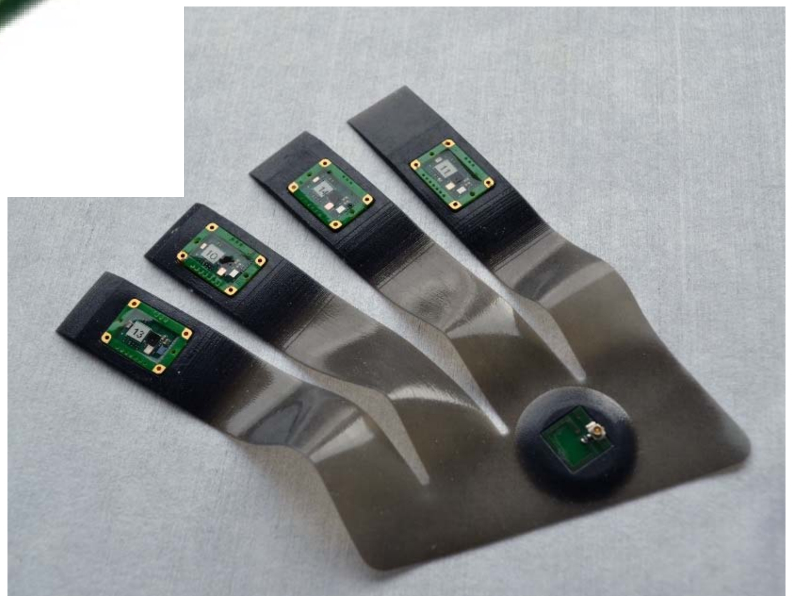




More recent example



Source: CSEM



Outline

- Characteristics of wireless transmission
- Features of wireless sensor networks
- Medium access control schemes for WSNs
- WiseMAC
- Comparison with other protocols
- Conclusion

Characteristics of radio transmission

- higher BER
- lower signalling rate
- limited possibility to detect collisions
- low spatial reuse
- prone to interference
- lower distances
- security concerns
- remote powering

- radio transmission
  - fading
  - incompatible regulations
  - higher cost
  - longer turn on and switching times
  - hidden terminal effect
- light transmission
  - line of sight
  - sensitive to heat
  - health concerns

Wireless tx properties implications

- MAC
  - master-slave (switching time ⇒ longer timeouts)
  - bus arbiter (hidden node ⇒ limitation in broadcast, reliable detection of silence ⇒ BA redundancy)
  - tokens (hidden node ⇒ token loss, switching time ⇒ longer timeouts)
  - virtual token (reliable detection of silence ⇒ token passing)
  - CSMA (no collision detection ⇒ use timeouts)
  - TDMA (switching time ⇒ longer gaps)

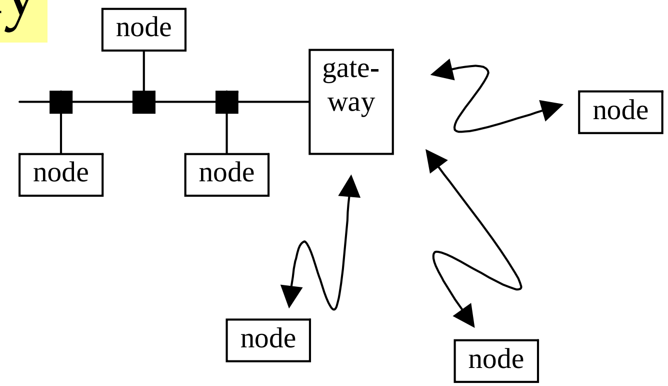
Implications of wireless tx prop (2)

- Error recovery
  - immediate retransmission
    - lower bandwidth, impact on higher layers
  - no immediate retransmission (cyclic transmission)
    - likelihood that errors will last
  - use forward error correction codes to lower apparent FER

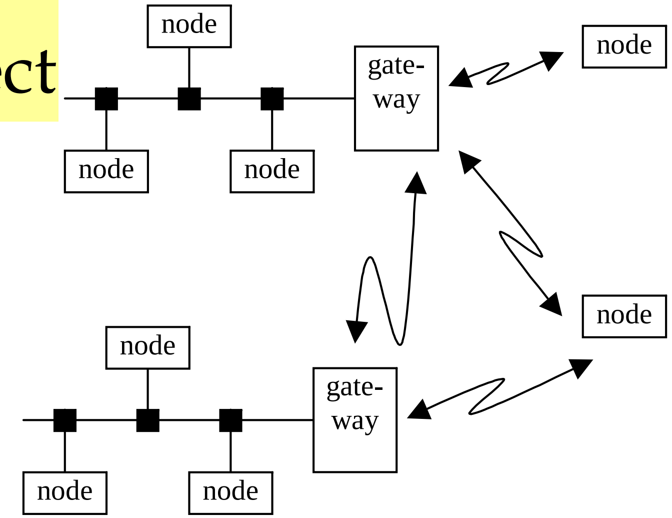
source: Ph. Morel, EPFL 1996

Examples of mixed architecture

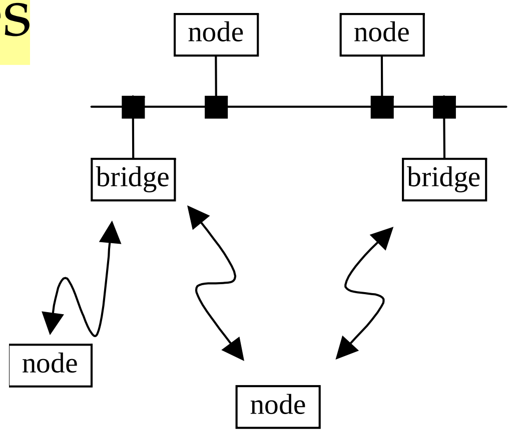
gateway



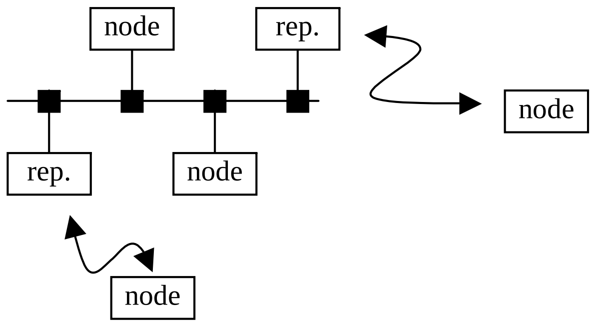
gateway interconnect



isolated nodes



repeater



source: Decotigne et al., FeT'2001

Possible architectures

- single or multiple wireline segments
- single or multiple wireless cells
- single or multiple points of connection
- connection through repeater, bridge, router or gateway
- ad hoc (self organising) wireless cell or base station
- single hop / multiple hops
- use of satellites

Approaches

- put a wireless physical layer to an existing solution
  - IEC 61158 fieldbus standard
  - ESPRIT project OLCHFA
  - LON
  - European Project R-Fieldbus
- adapt (or use) an existing wireless solution
- build a new solution

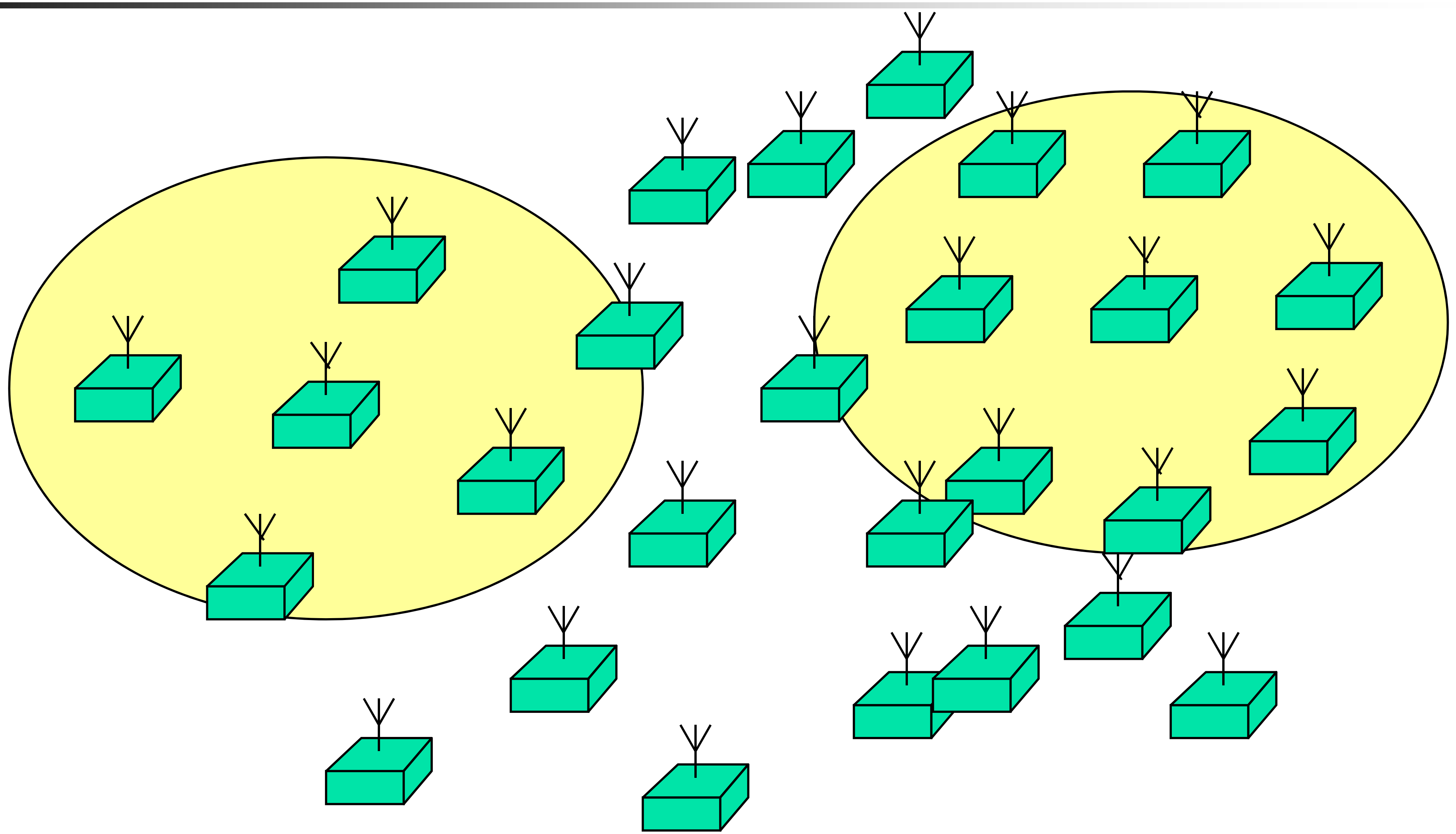
The change cannot be done so easily. It implies a number of changes at higher layers

Use an existing solution

- Radio modems
- Bluetooth
- ZigBee
- 802.11
- DECT
- IrDA
- Proprietary (Ant+, ..

Solutions can be used to solve a number of problems,
but none fulfills entirely the requirements
(additions are needed)

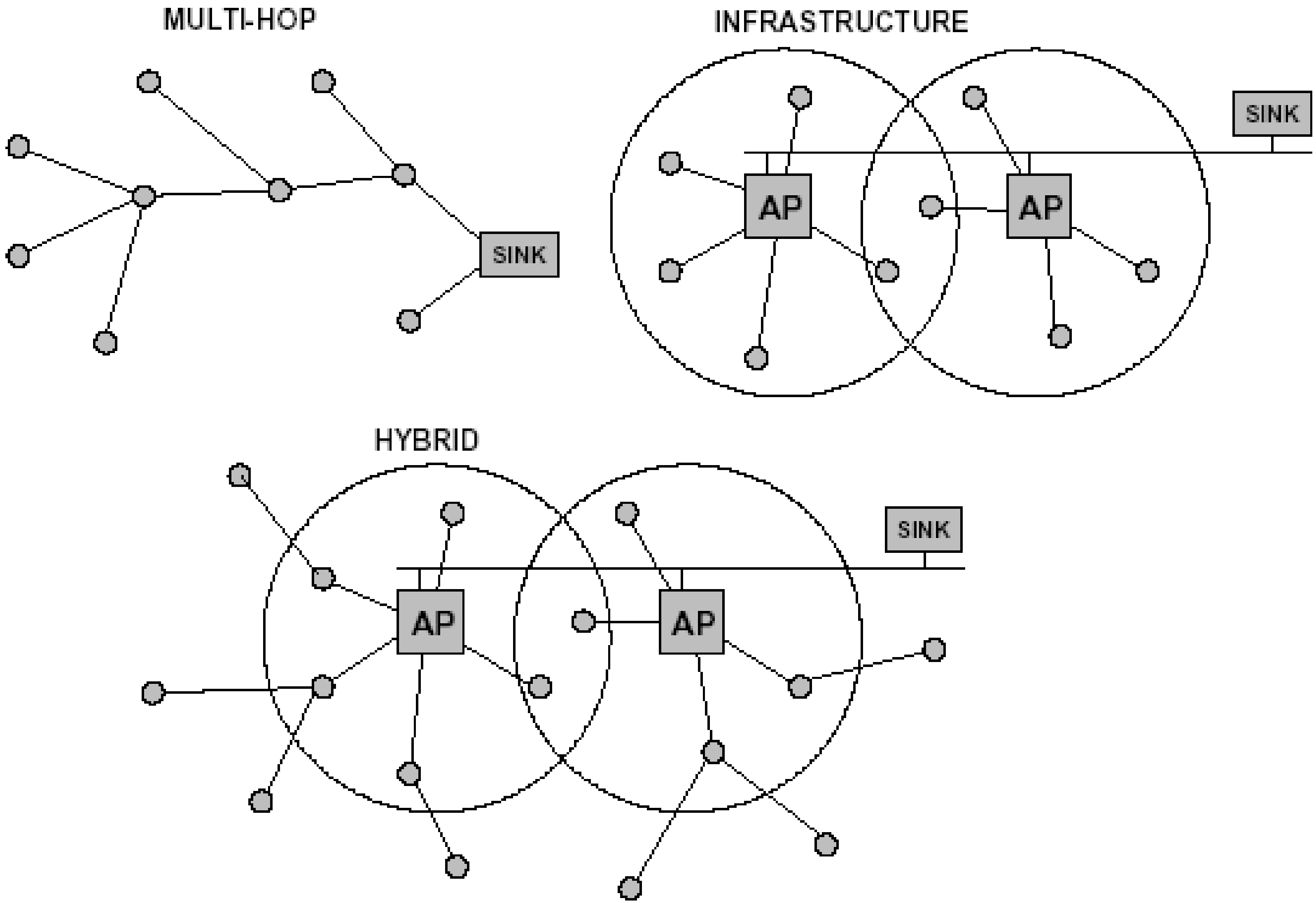
Wireless Sensor Networks



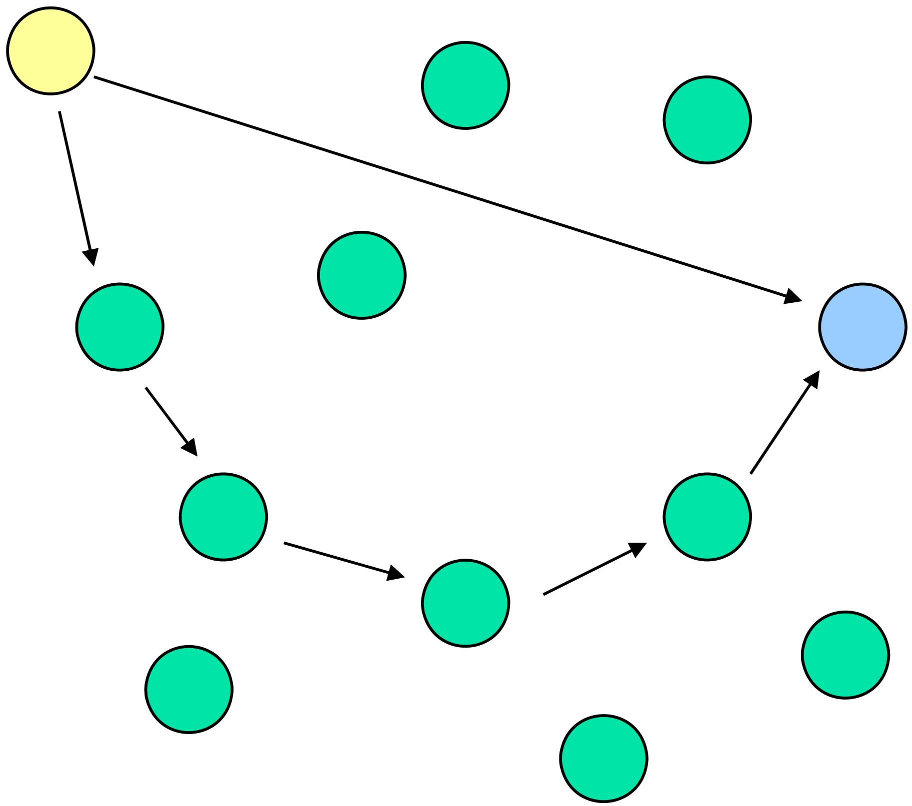
WSN expected features

- self organized
- no infrastructure (no base station, etc.)
- battery operated (low energy)
- multihop transmission
- small (< $1cm^3$), low cost (target <0.5$)
- low data rate (up to 10 Kbit/s/node)
- large number of nodes (0.05 to 1 nodes/m^2)
- sensor information temporal consistency

Topologies



The energy problem



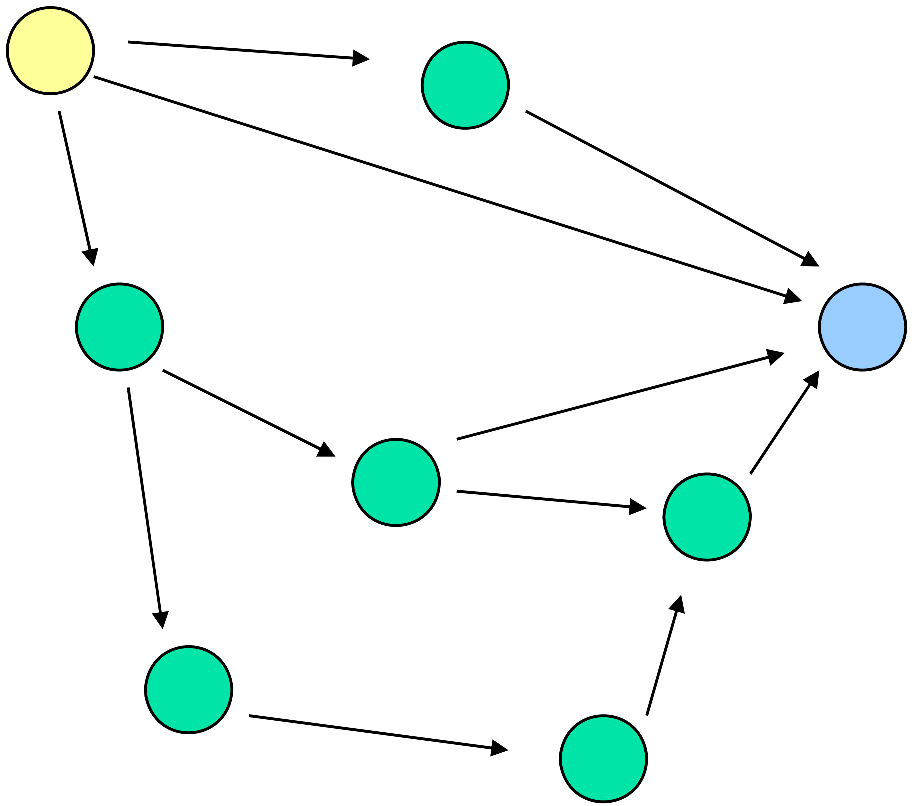
- 1 hop over 50m
  - 1.25 nJ/bit (thermal limit)
- 5 hops of 10m each
  - 5 x 2 pJ/bit = 10 pJ/bit

Assumes ~R^4 losses

ignores overhead and retransmissions

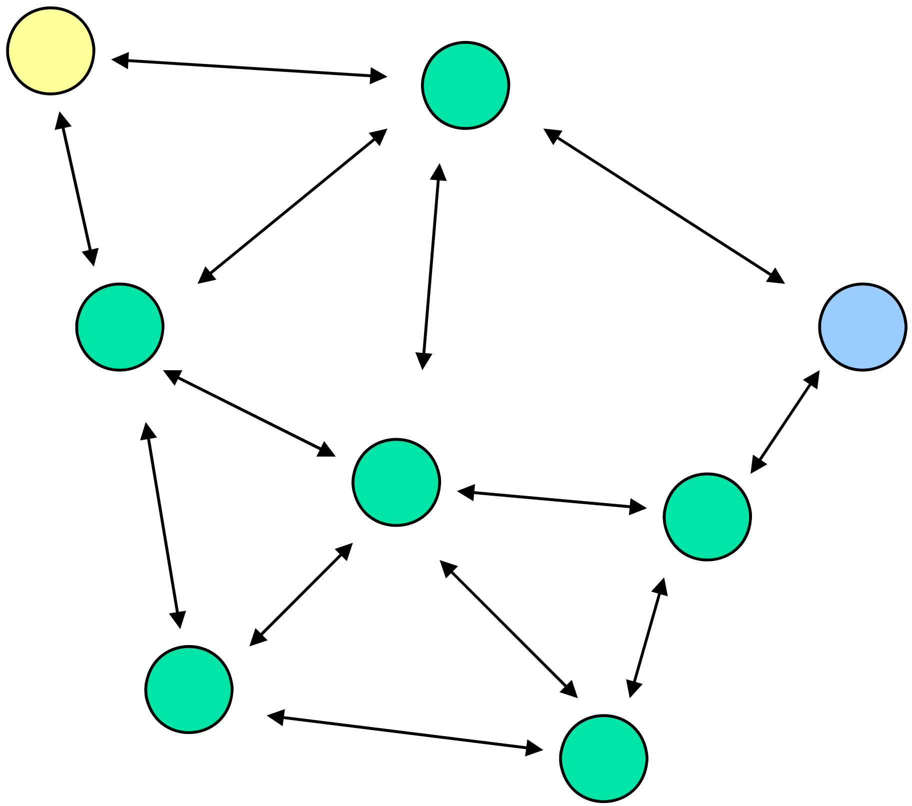
source: J. Rabaey, UCB

Fading and reliability



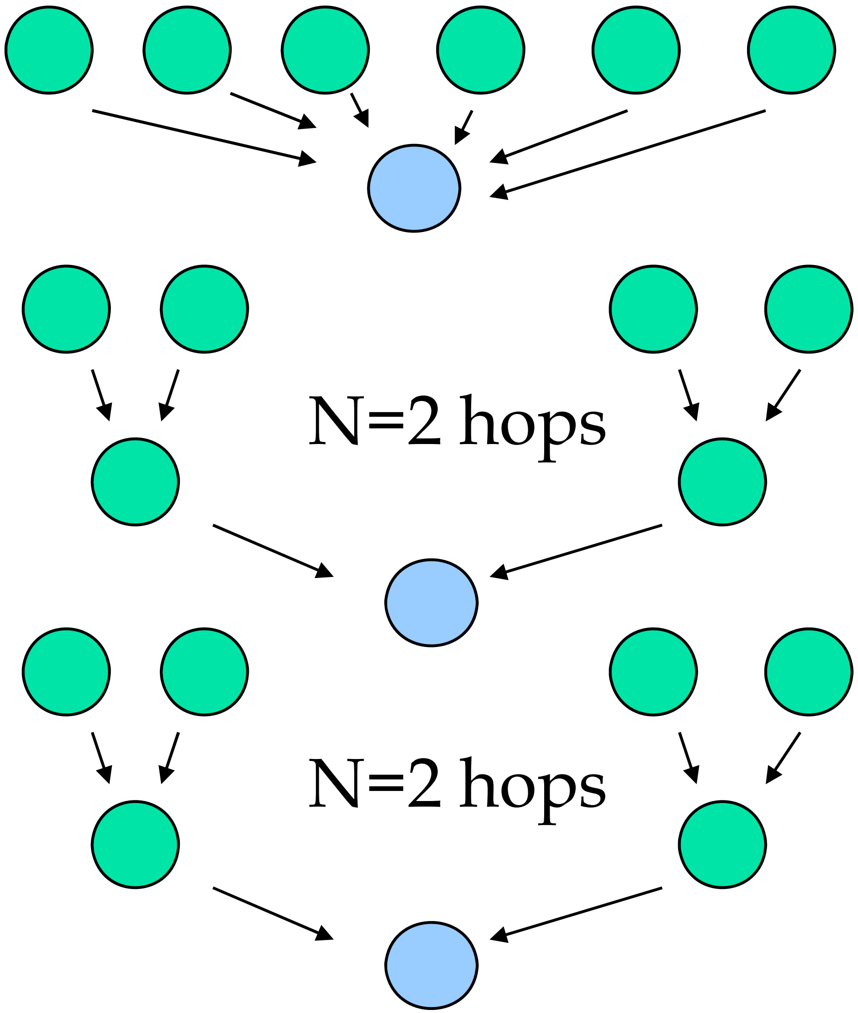
- Fading
  - There will exist multiple paths
- Reliability
  - multiple paths
  - no central point of failure
  - node redundancy

Localization



- relative positioning
- need for some know points (anchor points)

What about delay ?



- 2^{N+1}-2 sensors
- time for one message

$$D = T_{sync} + (L_o + L_M)T_{bit}$$

- central consumer, direct link

$$T = (2^{N+1} - 2)D$$

- central consumer multihop
  - serialized

$$T = (2^{N+1} - 2)D + \left((N-1)2^{N+2} + 4\right)L_M T_{bit}$$

  - in parallel

$$T = 2ND + 4L_M T_{bit}\left(2^N - N - 1\right)$$

Some Projects on WSNs

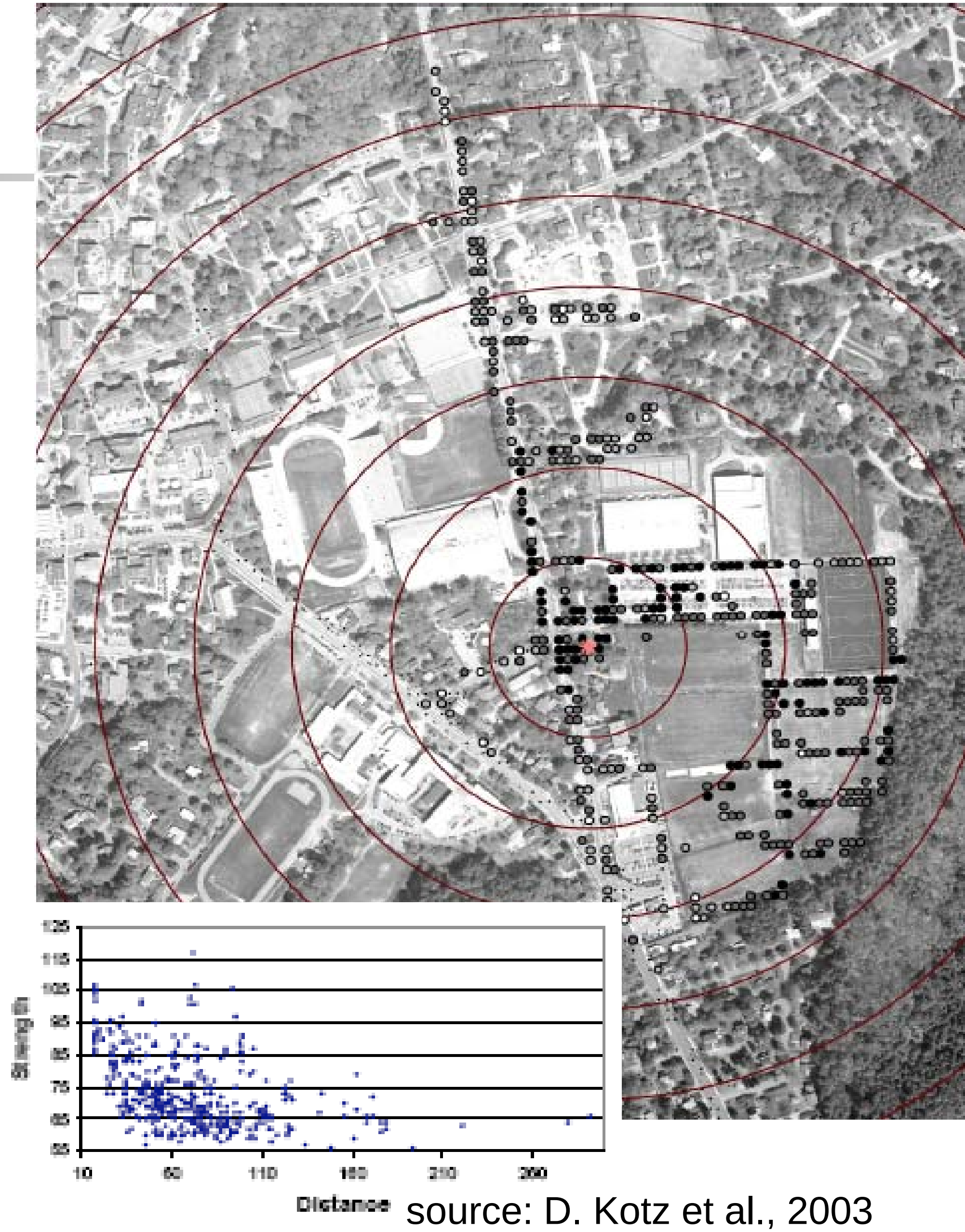
- Early projects
  - PicoRadio: Berkeley Wireless Research Center
  - Wireless Integrated Network Sensors (WINS): UCLA
  - Wireless Network of Devices (WIND): MIT
  - Wisenet: CSEM
- EU projects: EYES, WiseNts, Cruise, eSense, …
- MICS project led by EPFL

7 myths about wireless transmission

- the world is flat
- a radio transmission area is circular
- all radios have equal range
- if I can hear you, you can hear me
- if I can hear you at all, I can hear you perfectly
- signal strength is a simple function of distance
- link quality does not change

Some truths about wireless communication

- links fall into 3 categories
  - connected, transitional, disconnected
- transitional links are often unreliable and asymmetric (even for static nodes)
- packet error does not mean collision
- radio coverage is not at all circular
  - obstacles, height, fading, …
- signal strength is loosely related with distance



source: D. Kotz et al., 2003

Other challenges

- severe resource constraints
  - energy, bandwidth, memory size, processing
- unbalanced traffic
  - sink nodes
- network dynamics
- scalability
- multiple traffic requirements

So what's new for RTN research ?

- not only MAC
  - mulithop must be taken into account
- impossible to ignore errors
- highly dynamic cases
- limited resources
  - energy (means good models for that)
  - memory (buffers)

MAC for Wireless Sensor Networks

- Same objectives as a conventional MAC
  - Fair access
  - Minimize waiting time
  - Maximize network utilisation

- In addition
  - preserve energy

Sources of Energy Waste (@MAC)

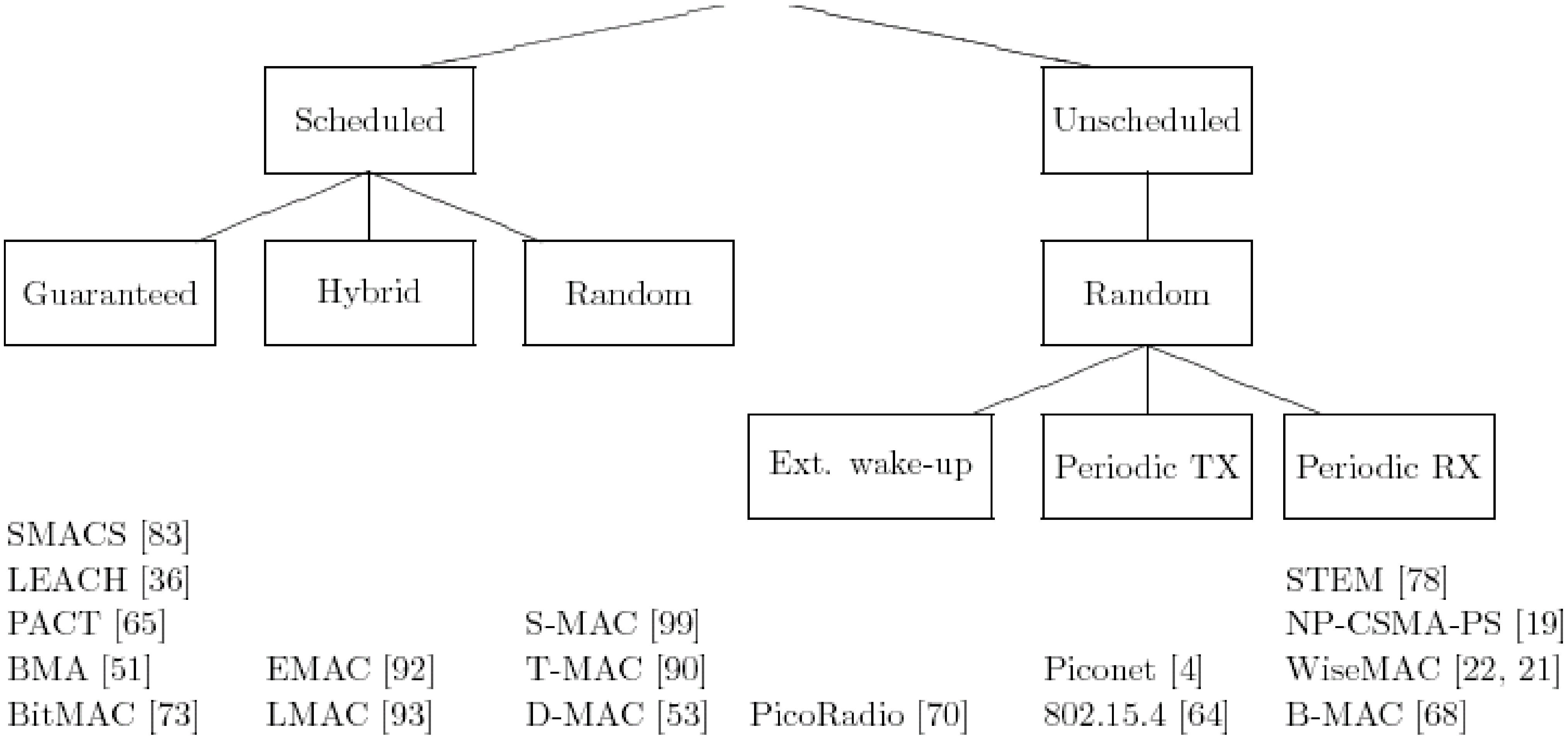
- Idle listening
  - Channel expected to be idle during long periods in sensor networks
- Overhearing
  - Not to be underestimated. Can become important in case of dense ad-hoc networks. Limits scalability in infrastructure sensor networks
- Collisions
  - To be avoided as retransmissions cost energy
- Over Emitting
  - Sending while there is no receiver ready to get the information
- Protocol Overhead
  - Required frame header and signaling to implement the MAC

Intuition

- Transmission does not cost too much energy
  - as long as there are not too many collisions
  - As long as we do not need to retransmit because the receiver did not receive the message

- Sleep most of the time
- Wake up when there is something to receive

MAC for Wireless Sensor Networks

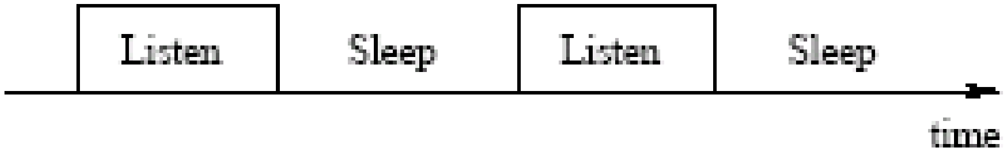


source: El-Hoiydi, 2005

See also K. Langendoen, A. Meier. 2010.

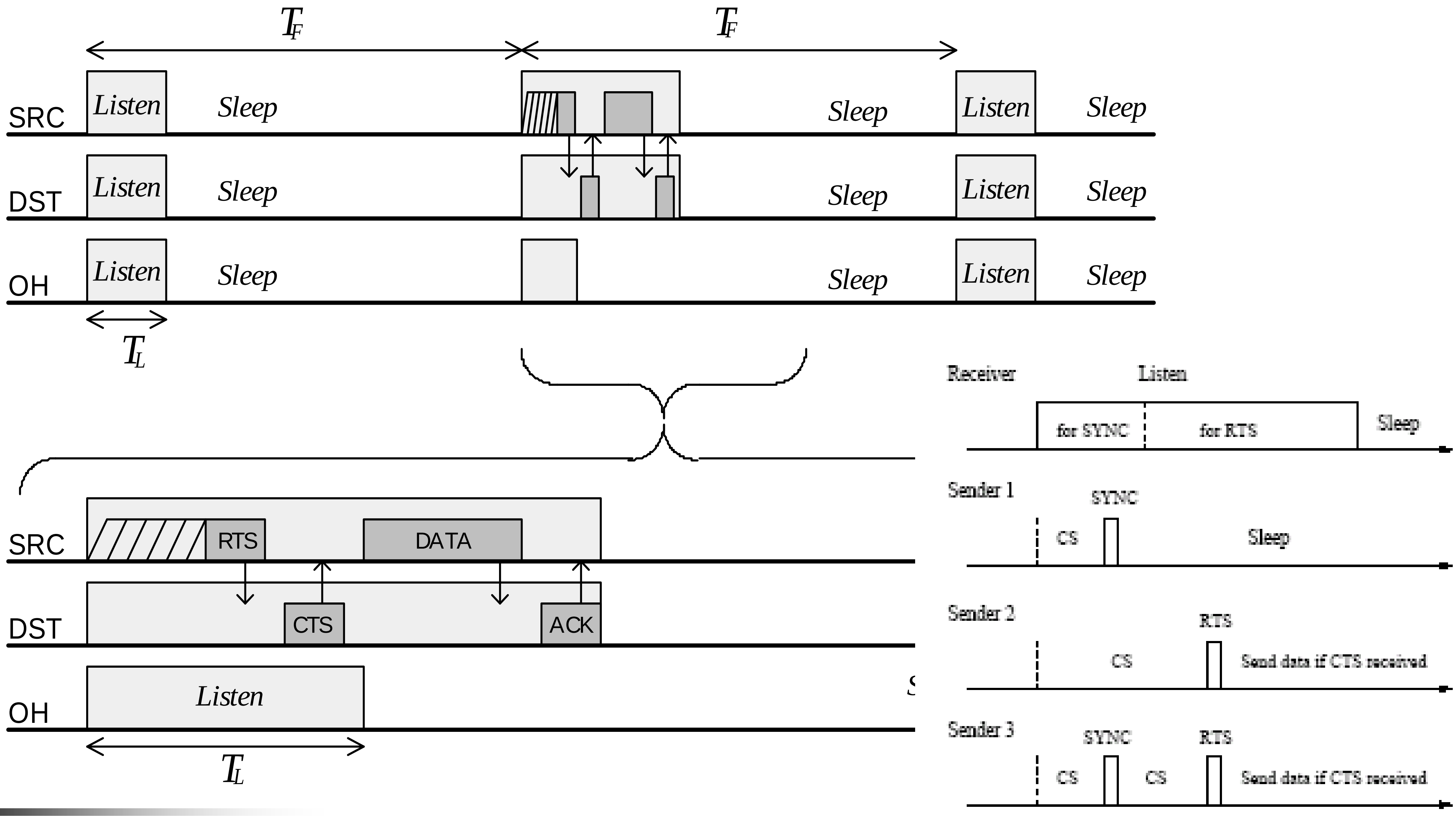
S-MAC

- Nodes have a listen/sleep periodic schedule



- A starting node
  - Listens to for a neighbor schedule
    - If found, adopts it / if not, selects one randomly (if it hears another schedule late, it adopts both)
  - then regularly broadcasts its schedule (SYNC packets)
- CSMA protocol with RTS/CTS to reduce collisions

SMAC



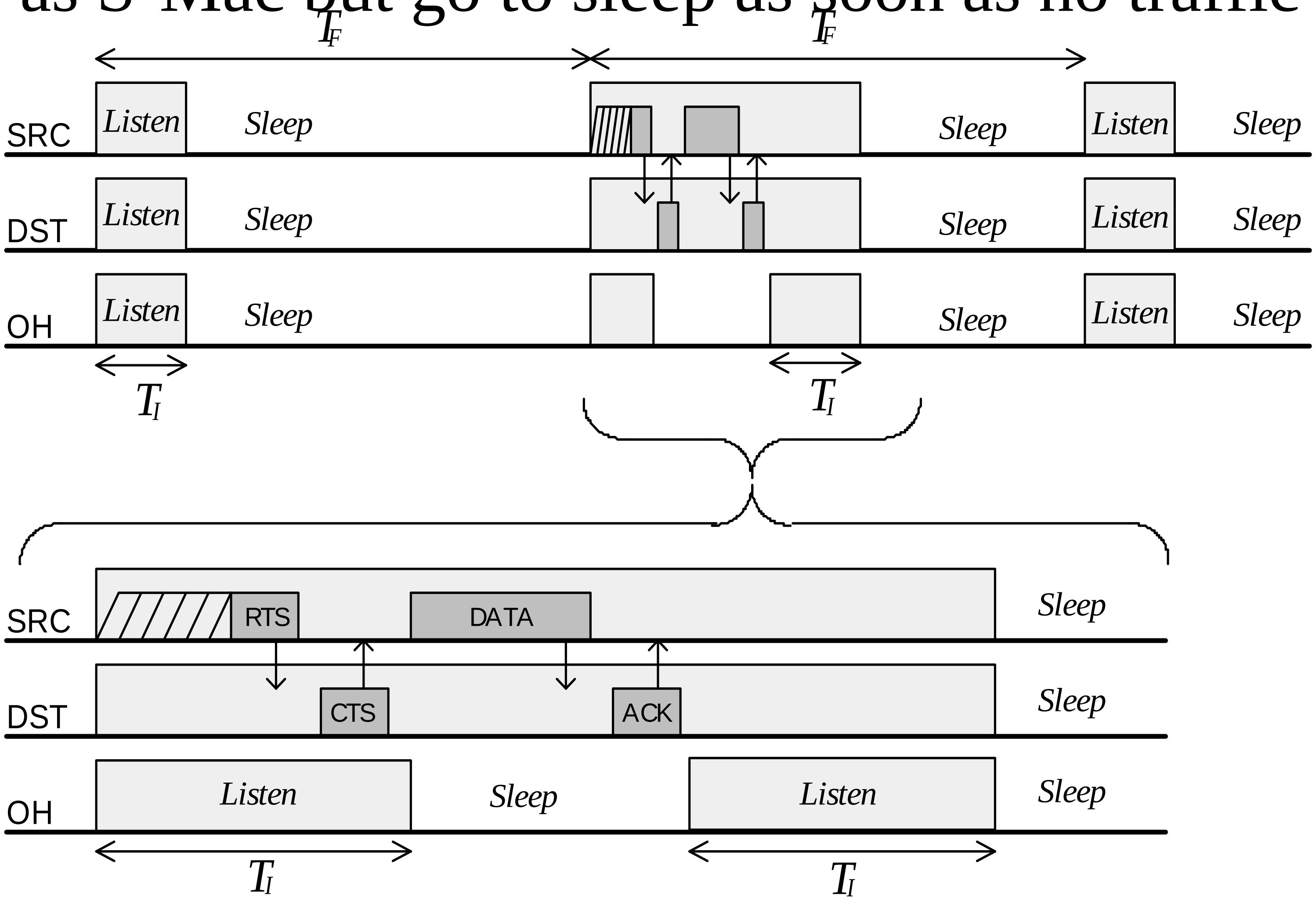
T_F
T_F
SRC
Listen
Sleep
Sleep
Listen
Sleep
DST
Listen
Sleep
Sleep
Listen
Sleep
OH
Listen
Sleep
Sleep
Listen
Sleep
T_L
SRC
RTS
DATA
DST
CTS
ACK
OH
Listen
T_L
Receiver
Listen
for SYNC
for RTS
Sleep
Sender 1
SYNC
CS
Sleep
Sender 2
RTS
CS
Send data if CTS received
Sender 3
SYNC
RTS
CS
CS
Send data if CTS received

S-MAC Analysis

- No guarantee
- Average latency is half of listen/sleep period
- Power consumption directed linked to fixed duty cycle
  - Does not adapt to traffic
- Mitigates
  - overhearing (RTS/CTS)
  - Idle listening (listen/sleep)
  - Collisions (RTS/CTS)
    - But higher due to concentration of traffic in the listen period

TMAC

- Same as S-Mac but go to sleep as soon as no traffic

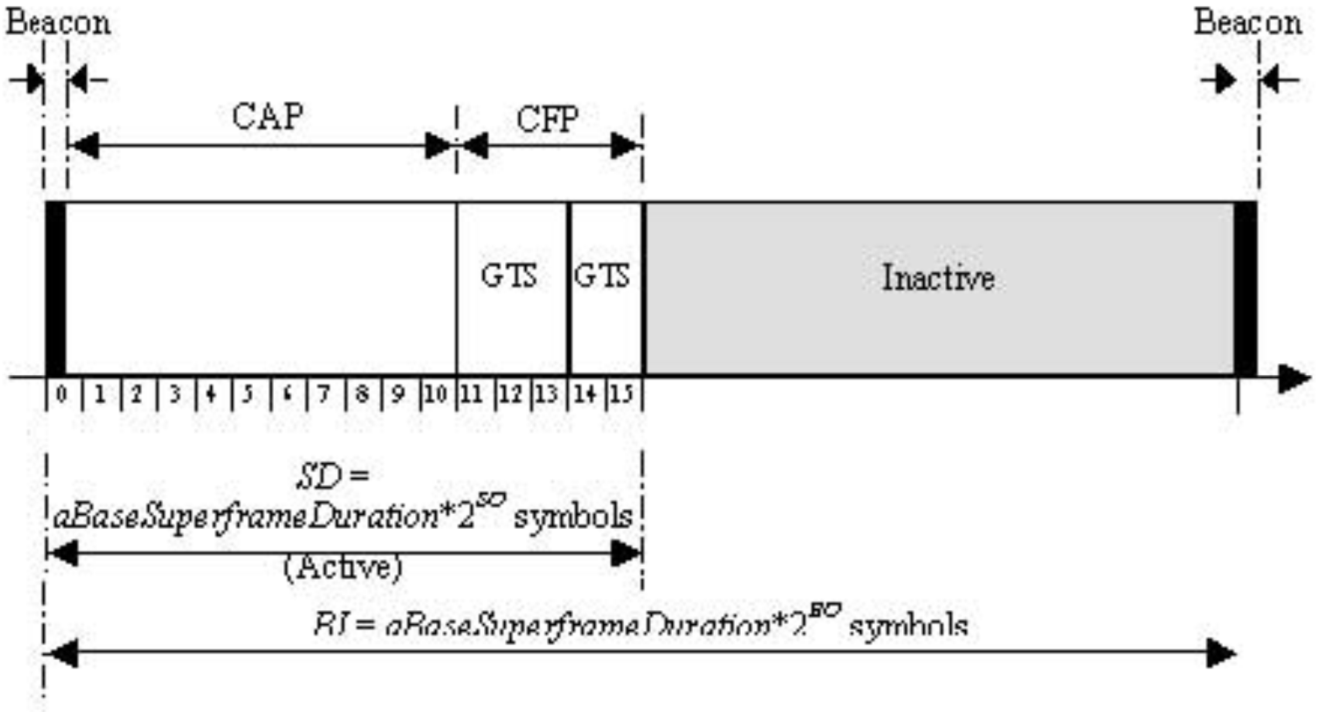


IEEE 802.15.4

- Part of IEEE 802 standards (only layer 1 & mac) international standard
- wireless LAN with 2 operating modes
  - station to station without coordination (ad hoc network or DCF)
  - coordinated by a single base station per cell
  - Provision for multihop (beacon offsets)
- 3 physical layer options (2.4 GHz – 16 ch., 902MHz and 868MHz)
- MAC: CSMA/CA + contention-less period
- Lower layers of many solutions
  - ZigBee industrial standard
  - ISA 100 / WirelessHART, ….

IEEE 802.15.4 (improperly called ZigBee)

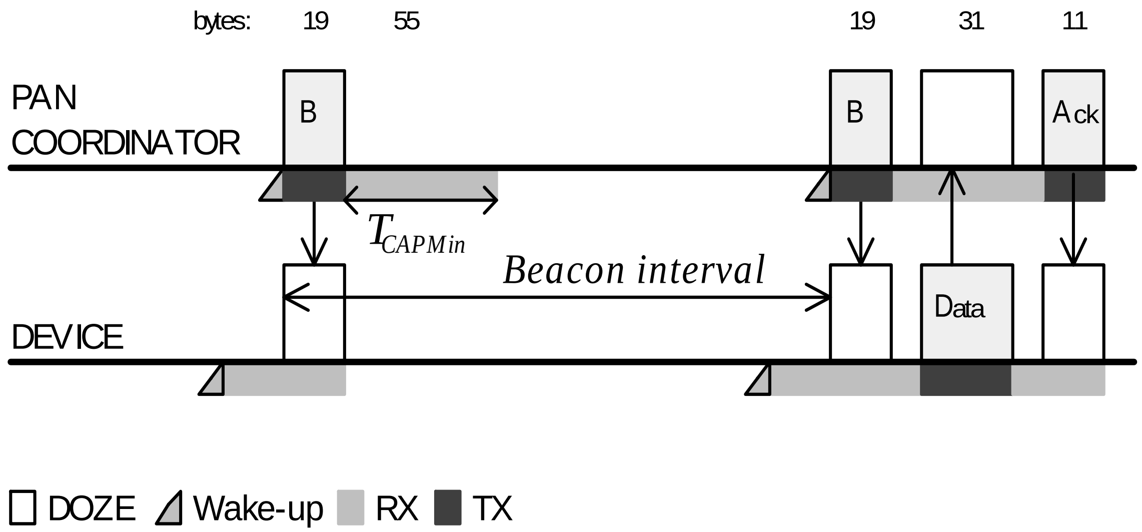
- 2 modes: peer-to-peer and star



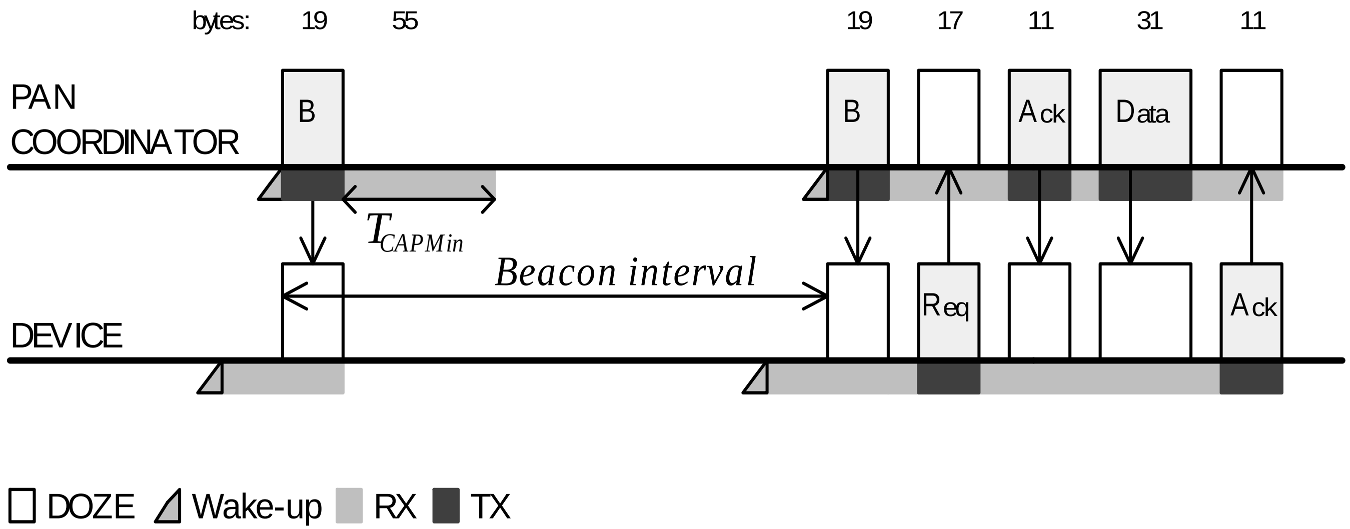
IEEE 802.15.4 – low power

- Duty cycle
- Power save mode
  - Nodes goes to sleep as soon as no traffic for a given time
- Traffic indication map
  - Coordinator indicates downlink traffic in beacon
  - Nodes may wake-up at their convenience, listen for beacon and check if incoming traffic

IEEE 802.15.4



bytes: 19 55 19 31 11
PAN COORDINATOR
B
B
Ack
T_{CAPMin}
Beacon interval
DEVICE
Data
DOZE
Wake-up
RX
TX



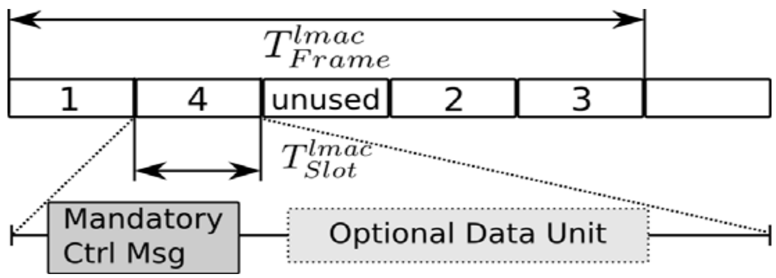
bytes: 19 55 19 17 11 31 11
PAN COORDINATOR
B
B
Ack
Data
T_{CAPMin}
Beacon interval
DEVICE
Req
Ack
DOZE
Wake-up
RX
TX

802.15.4 Analysis

- Star with beacon
  - Guarantees with GTS
  - Low consumption for “slaves”
- Peer-to-peer without beacon
  - No guarantee
  - Rather high idle listening (CSMA)
- Multihop with beacons
  - High energy consumption in relay nodes (coordinators)
  - Protocol not yet stabilized

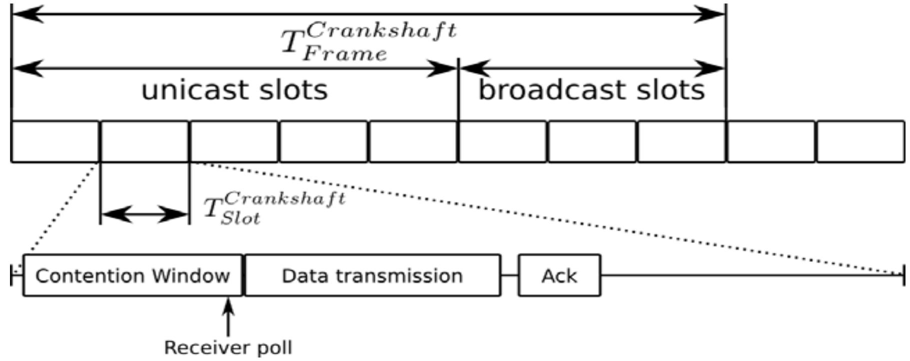
LMAC (Hoesel 2004)

- TDMA, time divided in
  - frames
  - and then in slots
- one node
  - owns one slot
    - Assigned in a distributed way
    - Slots are reused spatially
  - has to send a message at each instance of its slot
    - message = control part + data part
    - control part include id, slot id, slot used, slot with collision, dest,
  - must listen to control part of all neighbours
- The gateway starts the network



L. van Hoesel and P. Havinga, "A Lightweight Medium Access Protocol (LMAC) for Wireless Sensor Networks," INSS 2004.

Crankshaft (Halkes 2007)



- For dense networks
- TDMA (frames and slots)
  - requires clock sync.
  - Nu unicast slots
    - assigned to a receiving node (MAC address modulo Nu)
    - 2 windows : contention and data transfer
    - Senders contend to access medium (3 retries in subsequent frames)
  - Nb broadcast slots
    - all nodes must be listening
- synchronization
  - Every node has to send a broadcast packet every Tsync

G. P. Halkes, K. G. Langendoen. Crankshaft: an energy-efficient MAC-protocol for dense wireless sensor networks. EWSN'07, 228-244.

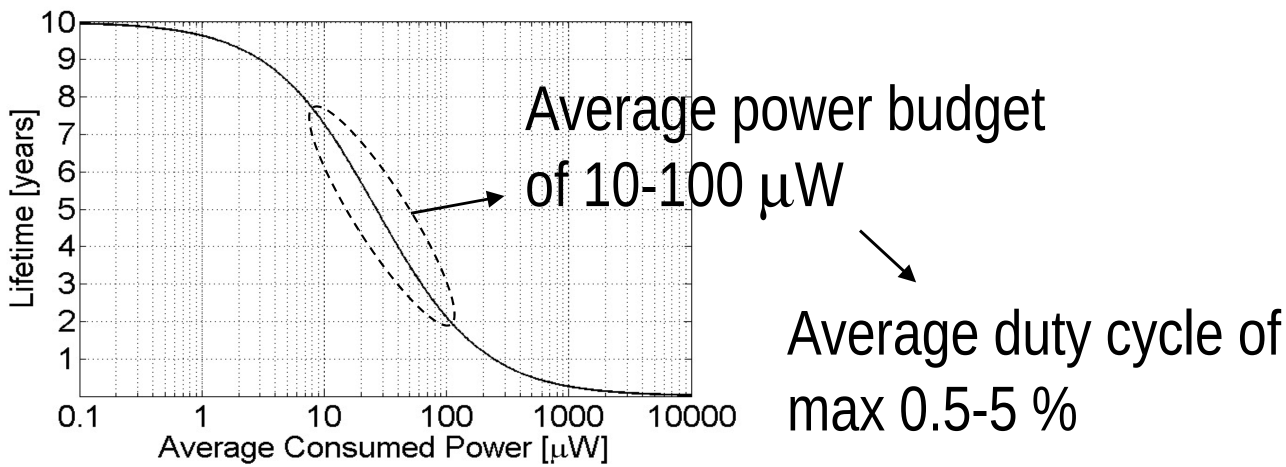
TDMA analysis

- All require tight synchronization
  - Often chicken and egg problem
- Slot time must be long enough to accommodate longest packet
- LMAC is in principle able to sustain RT traffic
  - In absence of errors
- Suited for regular traffic
  - How to accommodate different periods ?
  - How to accommodate varying demands ?

WiseMAC Req. & Assumptions

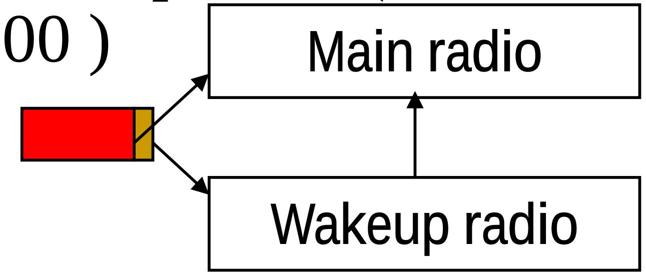
- Designed for the WiseNET System-on-a-chip
  - FSK Radio, 2 mA RX, 800 µs setup time, 35 mA TX, 5 µW DOZE
  - SoC operates down to 0.9V to require a single battery
  - 8 bits CoolRisc µC, little memory available
- Target lifetime of years on a single AA alkaline battery
- Multi-hop, Low average data rate, High latencies tolerated

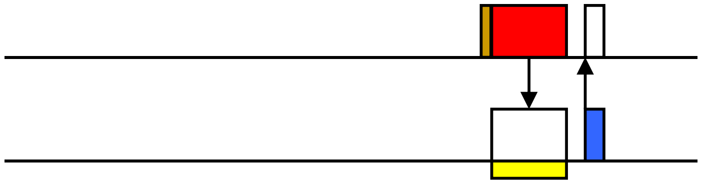
AA alkaline battery
- 2.6 Ah
- Power leakage of 27 µW



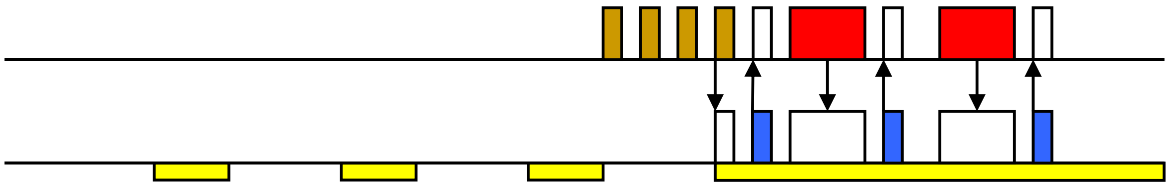
Wake-Up Schemes

- Wake-up radio (Picoradio, Rabaey et al, IEEE Computer Mag., 2000)

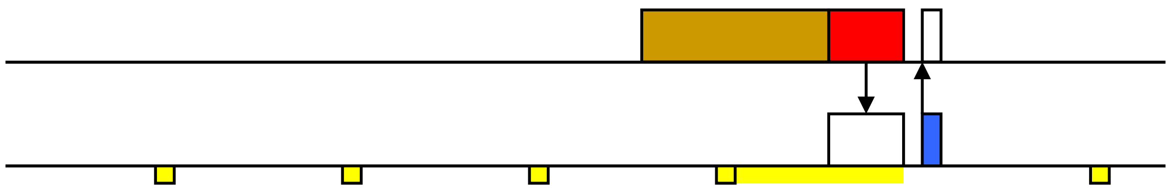




- Wake-up channel (STEM, Schurgers et al, IEEE Aerospace Conf. 2002)

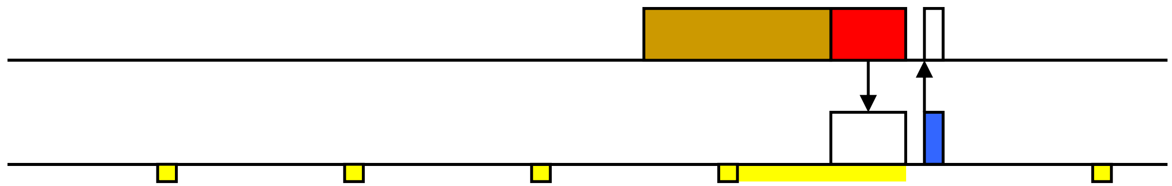


- Preamble Sampling (Aloha with P. S., El-Hoiydi, IEEE ICC 2002)



B-MAC

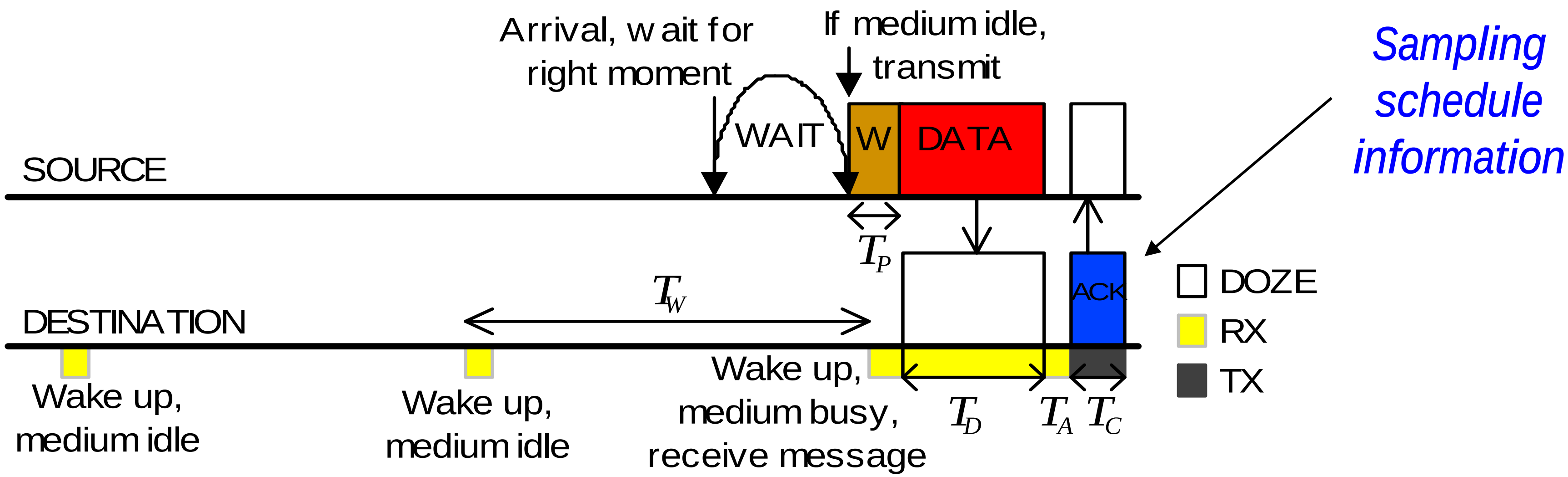
- Same principles as Aloha with Preamble Sampling
- Meant to be a generic MAC by allowing to change parameters by upper layers
  - CCA and ACK on/off, sampling period, preamble length



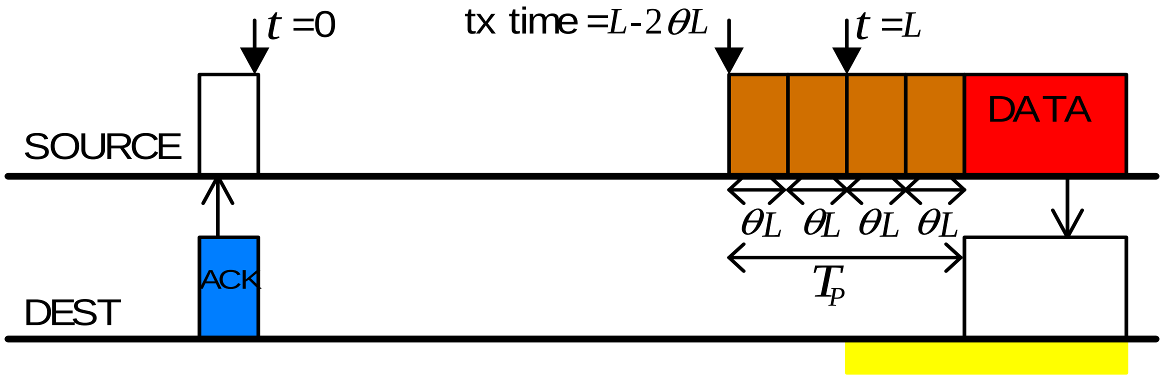
J. Polastre, J. Hill, D. Culler. Versatile low power media access for wireless sensor networks. SenSys '04, , 95-107.

WiseMAC

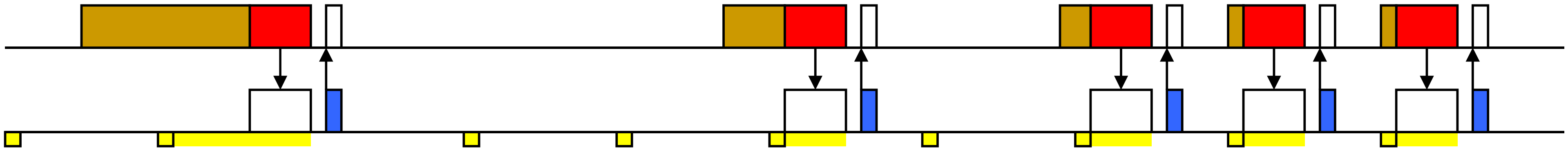
- Based on Synchronized CSMA with Preamble Sampling
  - Preamble sampling to minimize consumption of idle listening
  - wake-up preamble minimized by exploiting knowledge of sampling schedule of direct neighbors, less transmit, receive & overhearing overheads.
  - No setup-signalling. Self local synchronisation and re-synchronization
  - Choice of T_W is a trade-off between energy and latency



Minimum Preamble Length

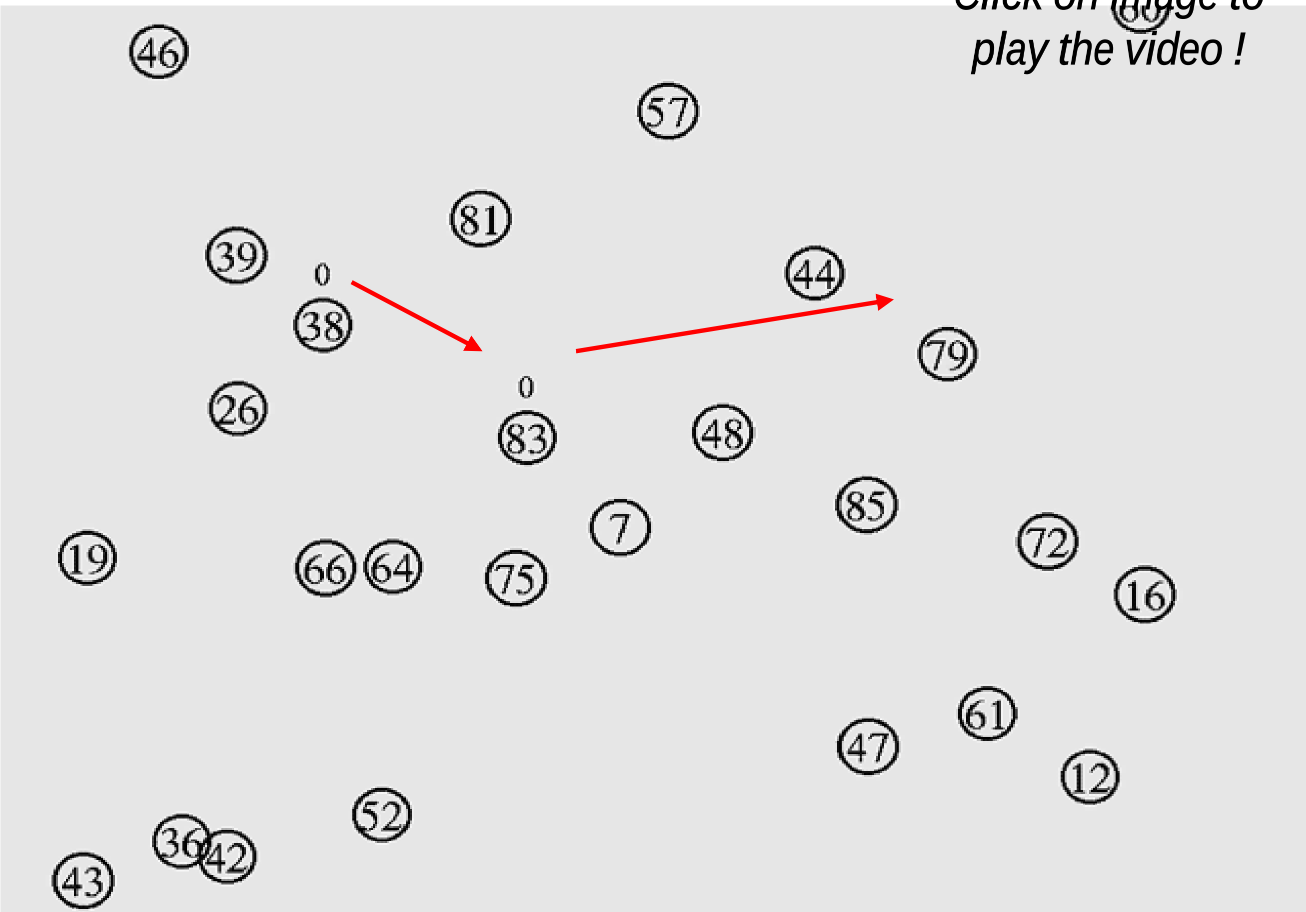


- Sampling schedule information inserted in every ack.
- Clocks drift of at most $\pm\theta$ per second (typically θ =30 ppm)
- With L seconds in between two communications, the length of the preamble must be $T_P = \min(4\theta L, T_W)$



Preamble Length Reduction

Click on image to play the video !



Node
- Black: Sleeping
- Green flash: Sampling
- Yellow: Listening
- Purple: Waiting for the right tx time
- Circled: Backoff

Message
- Brown: Preamble
- Red: Data
- Dark Blue: Ack

Sequence takes 1 second in real time (40 times faster)

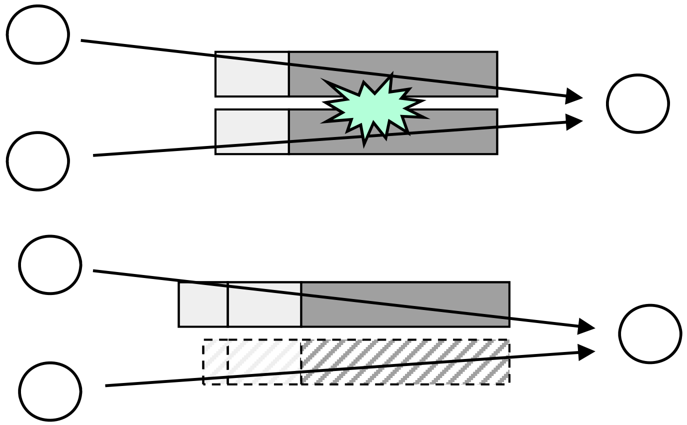
1st transmission: no sampling schedule info, using long preamble

2nd transmission: sampling schedule info gained through previous communication, using short preamble → less overhearers !

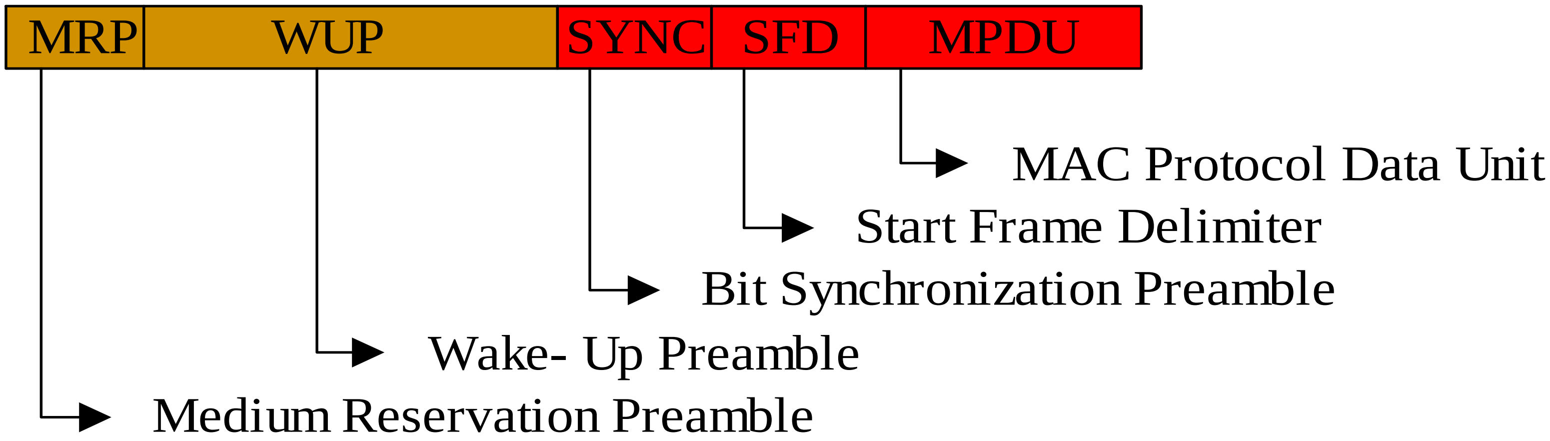
WiseMAC – Further Design Choices

- Medium reservation
  - To avoid collisions when using short preambles
- ‘More’ bit in data header
  - To allow the transport of traffic bursts
- Data frame repeated in preamble when preamble large
  - To reduce the frame error rate, mitigate overhearing, differentiate interferences from wake-up preambles
- Receive threshold above sensitivity
  - To avoid useless wake-ups caused by noise or weak signals (at the cost of the range)
- Carrier sensing range larger than interference range
  - To mitigate the hidden node effect (at the cost of the capacity)

Medium Reservation

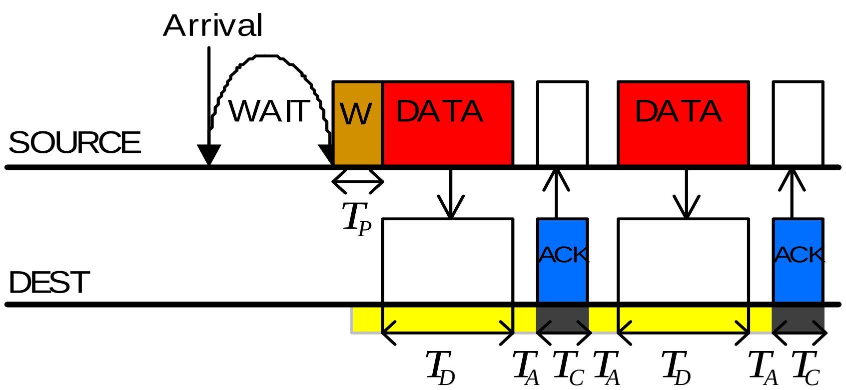


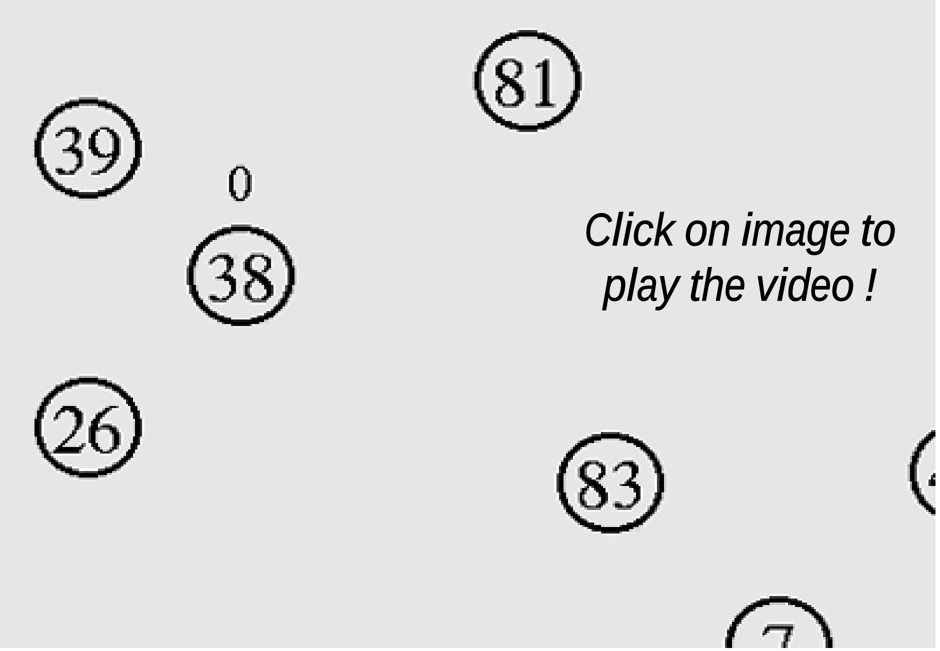
To avoid collisions between transmissions to the same node and at the same target sampling time



'More' Bit

- 'More' bit in data header indicates follow-up packets
- Allows the transport of busty traffic with low delays
- Cost of wake-up preamble shared among several packets



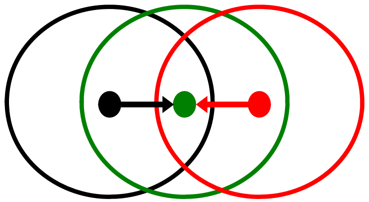


Preamble made of repeated packets

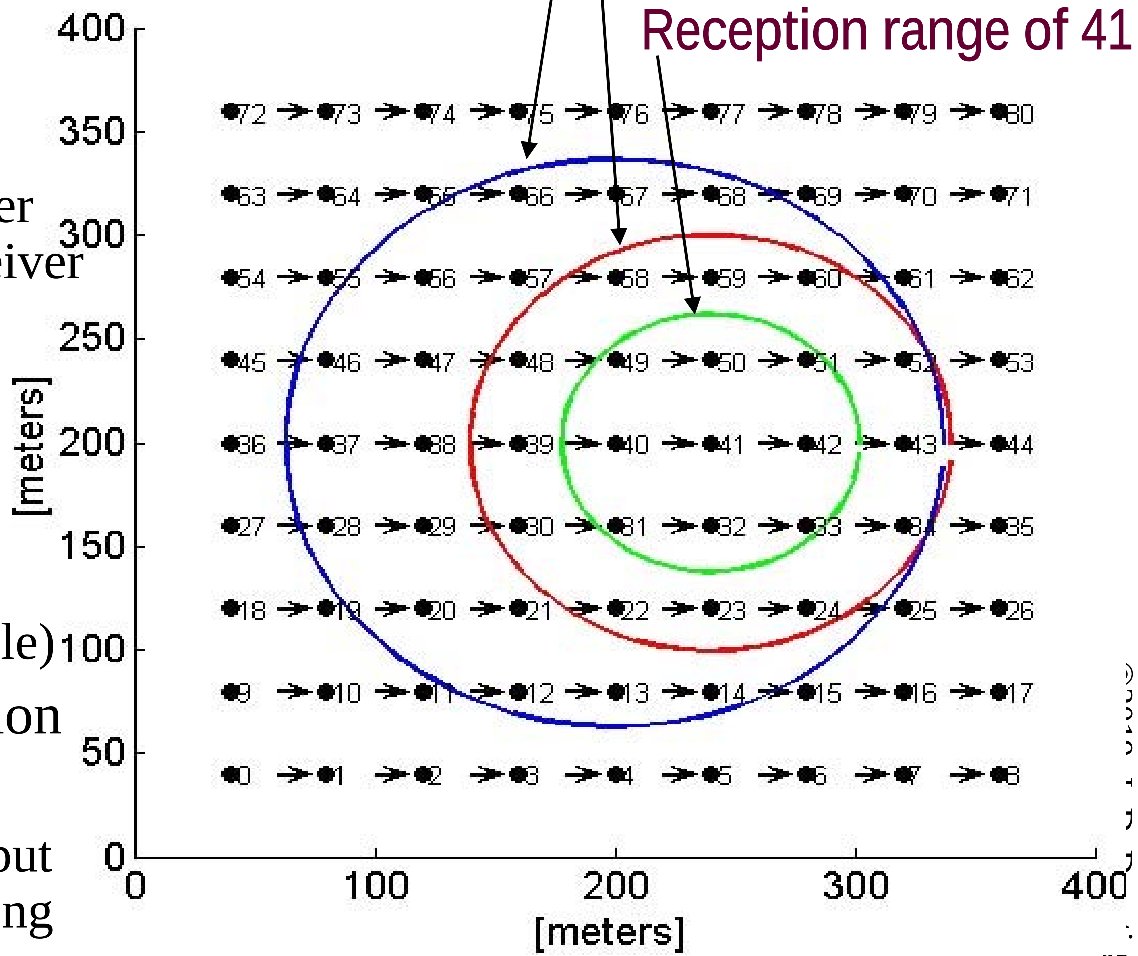
- Instead of having a preample made of a busy tone
- May be done by repeating
  - the data packet
  - A packet with only the address of the destination
- Modern radios impose a silence between 2 consecutive packets
  - Preamble sampling must be repeated twice
- Busy tone (carrier) is not always possible on some radios
  - Preamble is a multiple of packet size
    - Some degradation is one wants an efficient reservation preamble

Carrier Sensing Range Extension

- Hidden Node Effect

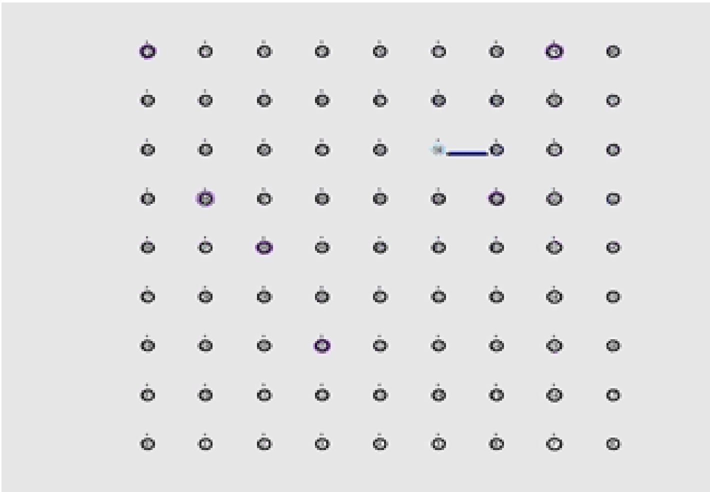


  - Hidden node is far from sender (not sensed), but close to receiver (interfering with reception).
- RTS-CTS
  - Not useful when data packets are small
  - Requires broadcast communication (long preamble)
- Carrier Sensing Range Extension
  - Mitigates hidden node effect
  - But reduces the max throughput (channel capacity shared among more nodes with CSMA)

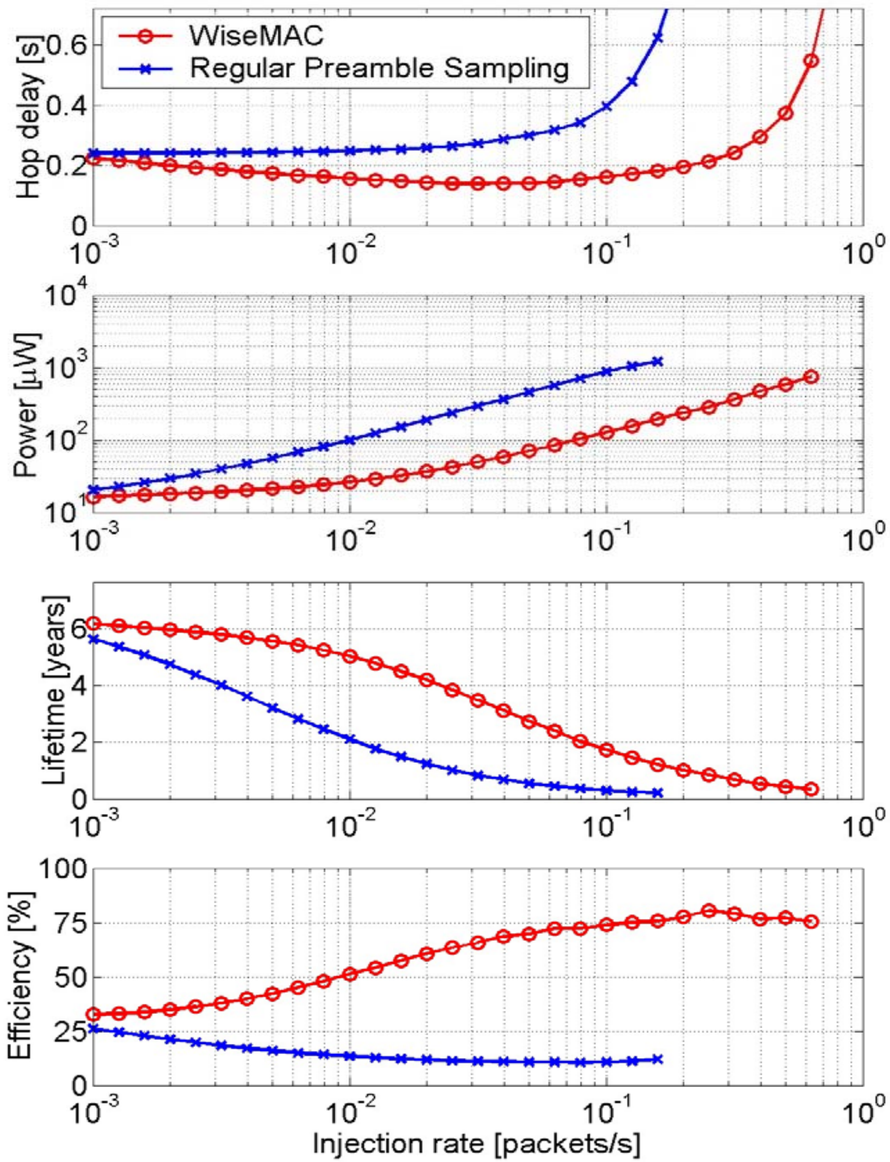


Performances

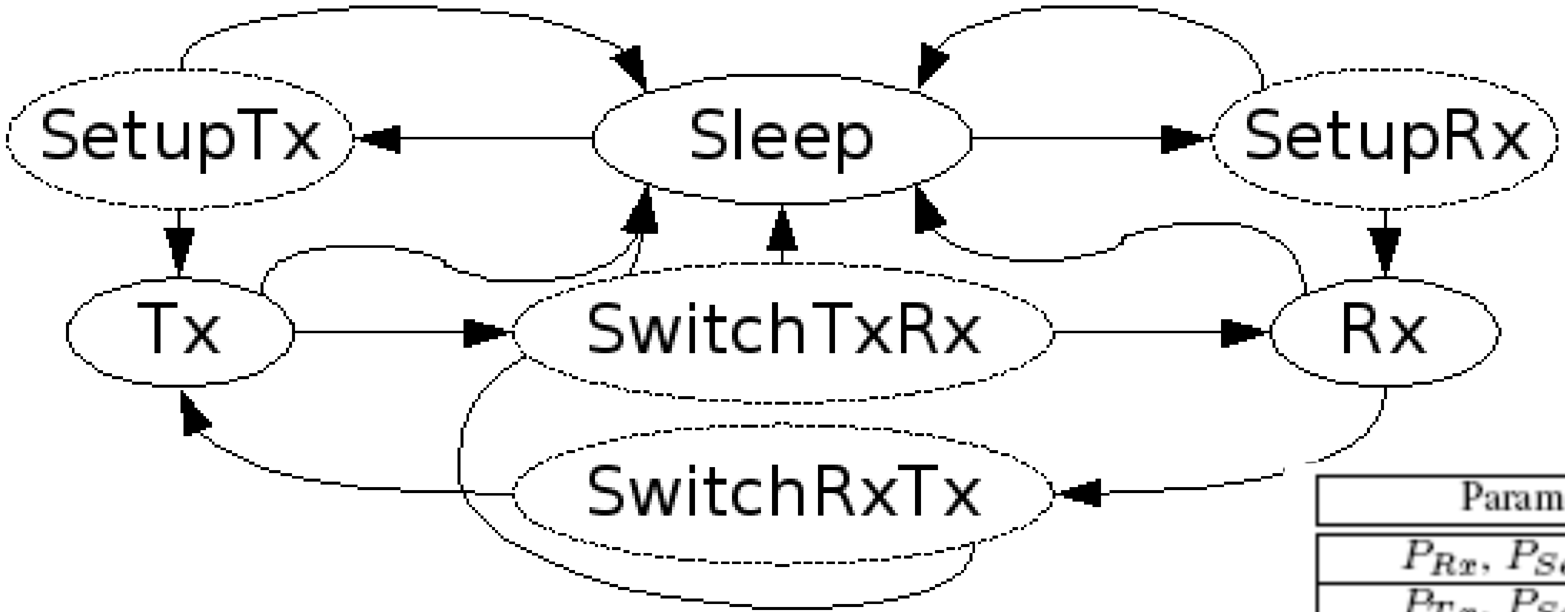
- Lattice Network



  - Traffic inserted in left-side nodes
  - Forwarded towards the right
  - Statistics collected on central node
  - Models infinitely large sensor net
- WiseMAC adaptive to the traffic
  - Ultra-low power consumption in low traffic conditions, high energy efficiency in high traffic condition



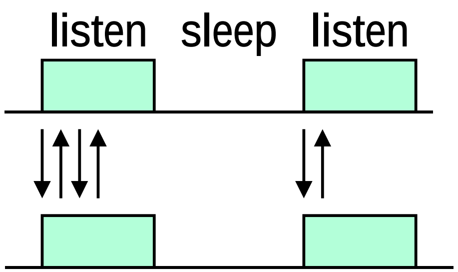
Radio state model

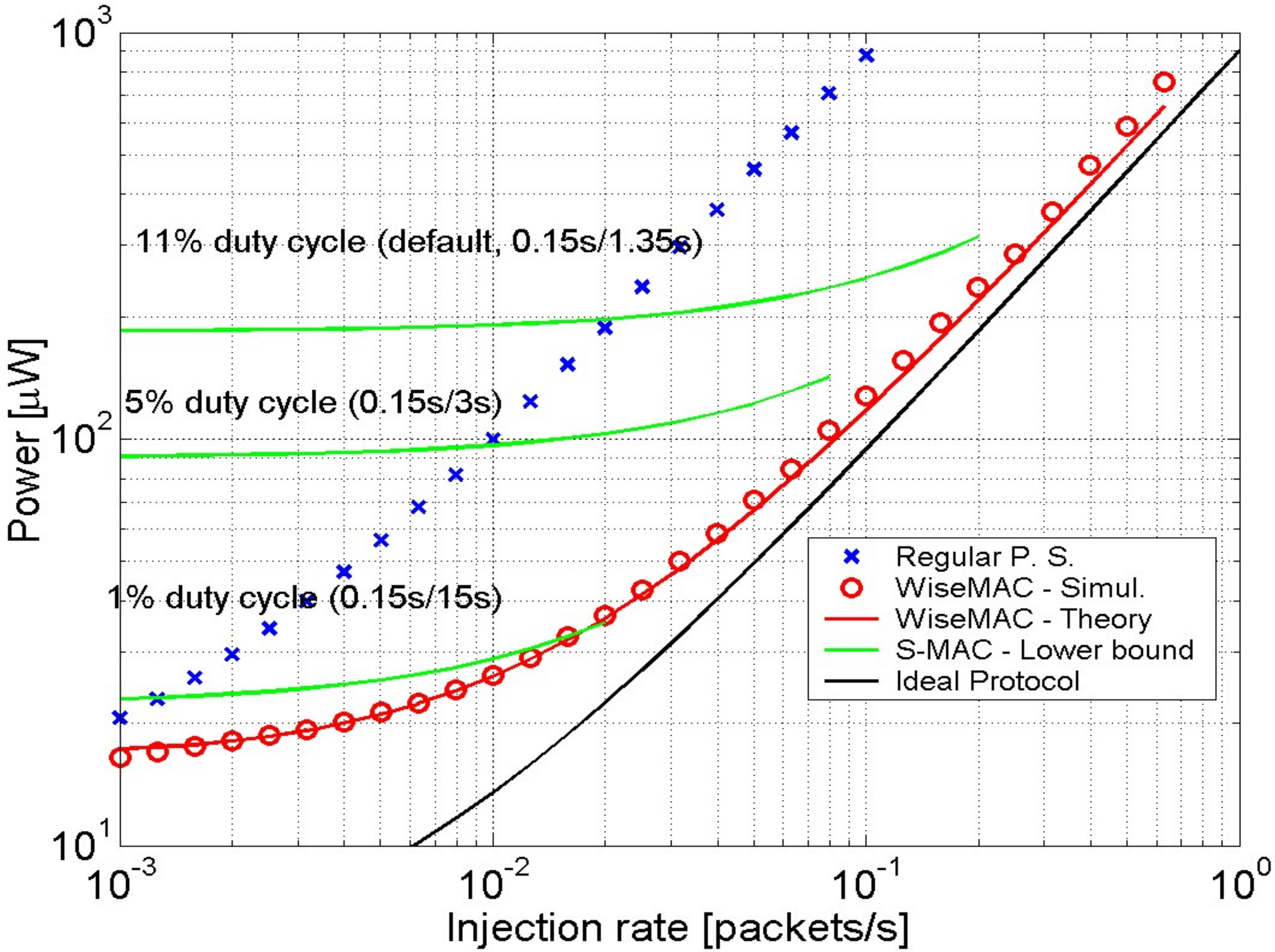


Parameter	Value
$P_{Rx}, P_{SetupRx}$	56.4 mW
$P_{Tx}, P_{SetupTx}$	25.5 mW
P_{Sleep}	0.06 mW
P_{SwTxRx}, P_{SwRxTx}	54.3 mW
T_{SwTxRx}	160 µs
$T_{SetupRx}$	192 µs
$T_{SetupTx}$	12 µs
T_{CCA}	128 µs
T_{Sync}	160 µs
Θ	$30 * 10^{-6}$
Bit rate	250 kbps

Comparison with S-MAC*

$$
\begin{aligned}
E_{WiseMAC} \approx\ & \frac{1}{\lambda}\left(\frac{P_{RX} T_{Se}}{T_{WU}} + P_{DOZE}\right) \\
& + P_{TX}\left(T_{MR} + T_P + T_M\right) \\
& + P_{RX}\left(\frac{T_P}{2} + T_M\right) \\
& + 7\frac{P_{RX}\left(T_{MR} + T_P + T_M\right)^2}{2T_{WU}}
\end{aligned}
$$





* S-MAC, Ye, Heidemann, Estrin, *INFOCOM* 2002.

SCP-MAC

- uses periodic Preamble Sampling
- neighbours are synchro-nized (SMAC technique)
- Adaptation to traffic with multihop optimization
- Collision mitigation

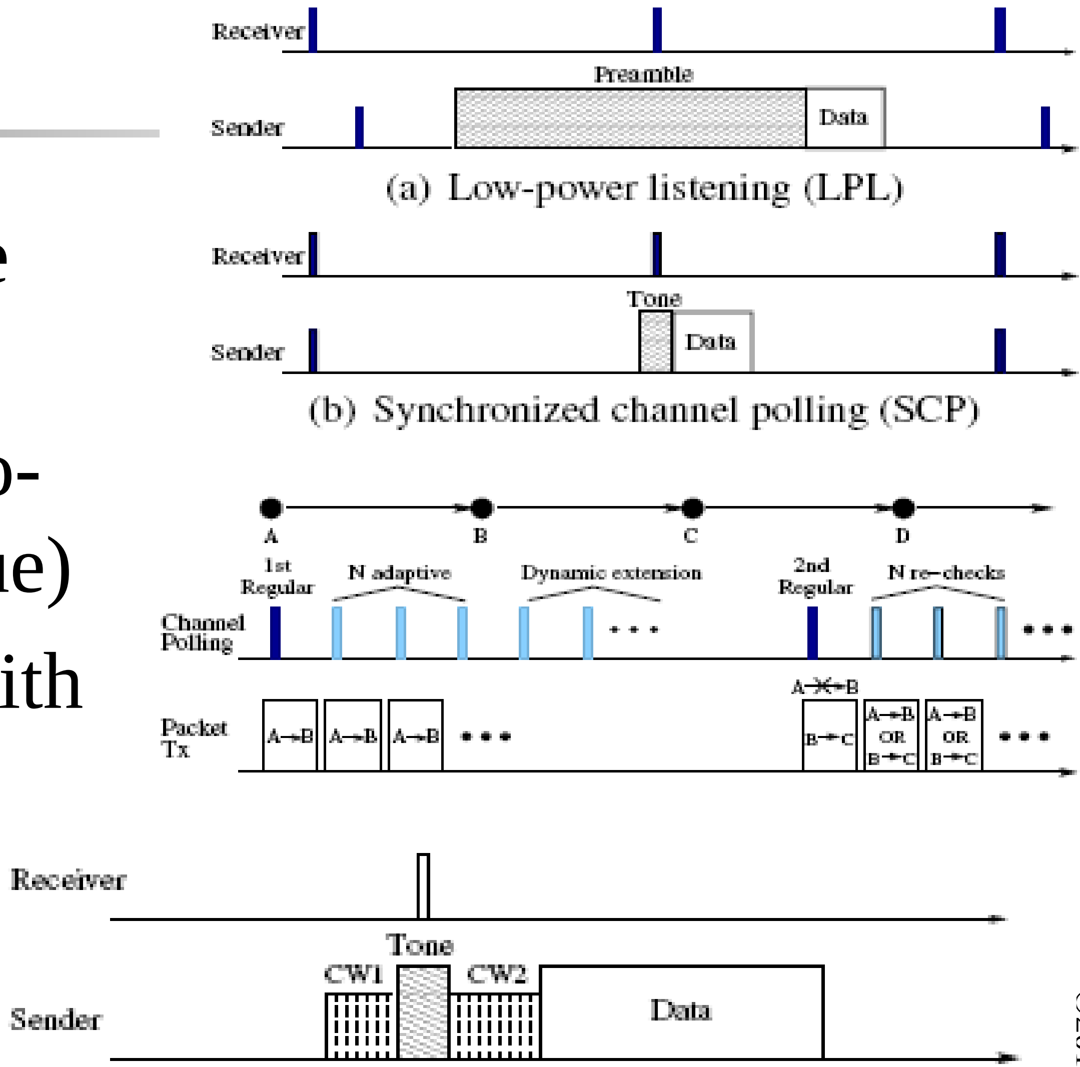
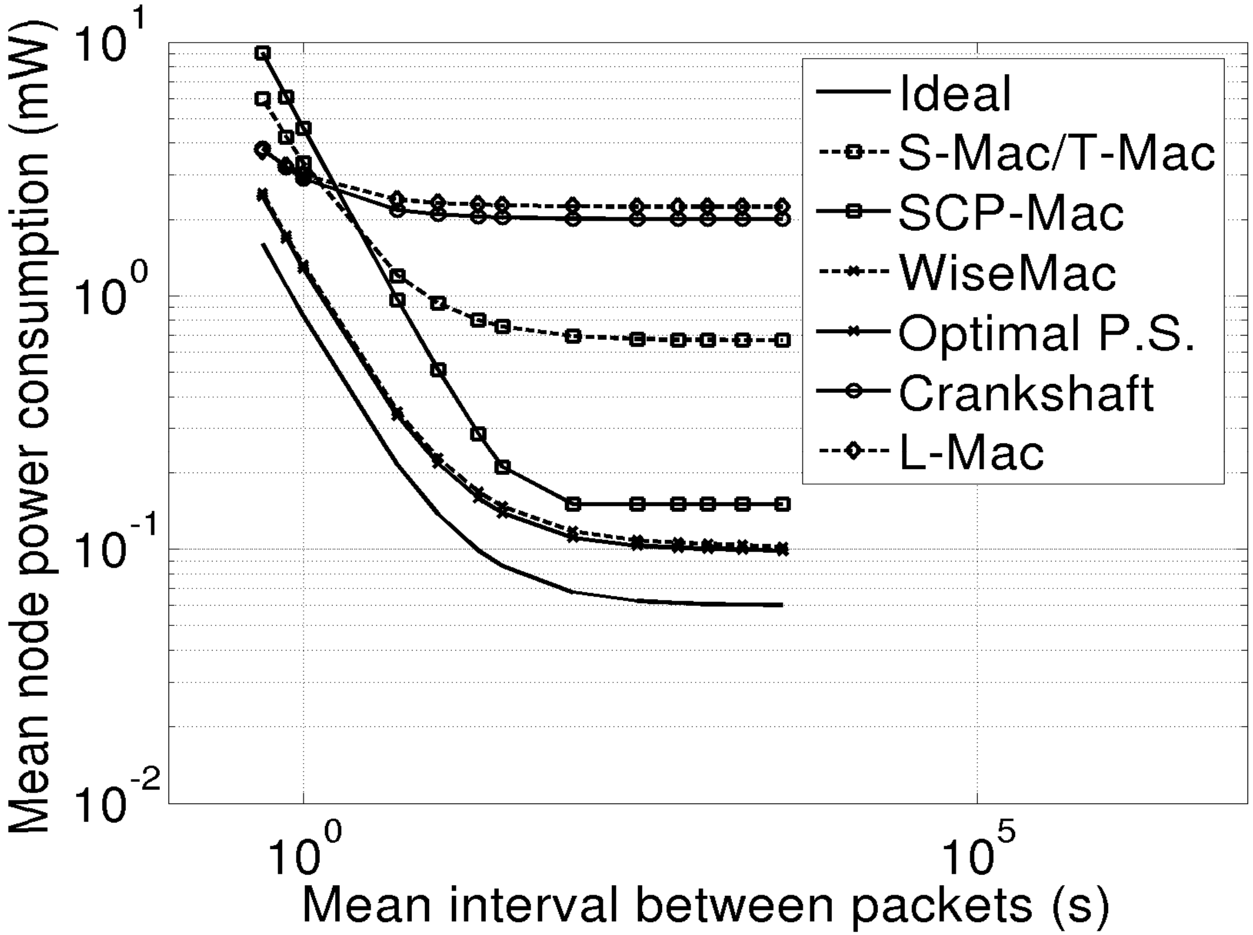


Figure 3. Two-phase contention in SCP-MAC.

Ye, W., et al. Ultra-low duty cycle MAC with scheduled channel polling. Proc. SenSys '06. pp. 321-334.

Comparisons

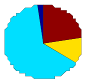
- each node sends and receives the same traffic
- all nodes are visible (10 nodes)



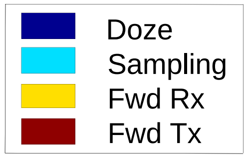
Comparison with 802.15.4

Forwarding a 32 bytes packet every 30 seconds in a multi-hop network.
Same weak-up period T_W = 250 ms, same hardware (XE1203 Radio).

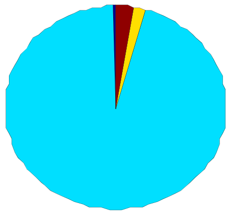
WiseMAC
(from CSEM)



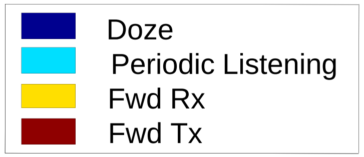
Total 304 uW



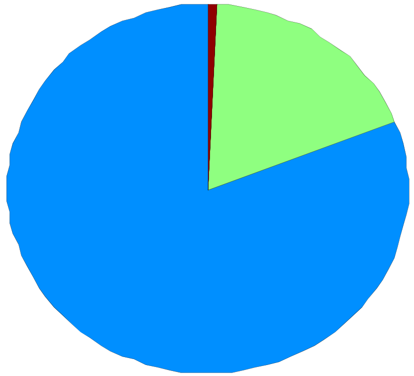
S-MAC
(from UCLA)



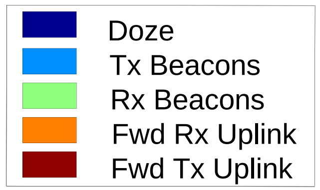
Total 2318 uW (x 7)



IEEE 802.15.4
ZigBee standard



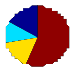
Total 7659 uW (x 25)



Comparison using an IC that reduces receiving energy

Forwarding a 32 bytes packet every 30 seconds in a multi-hop network. *Same weak-up period T_W = 250 ms, same hardware (WiseNET SoC).*

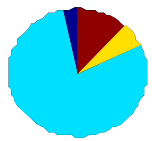
WiseMAC
(from CSEM)



Total 30 uW

Doze
Sampling
Fwd Rx
Fwd Tx

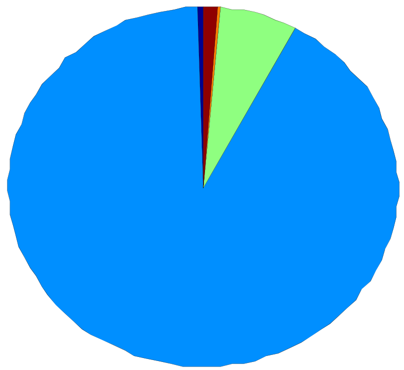
S-MAC
(from UCLA)



Total 139 uW (x 5)

Doze
Periodic Listening
Fwd Rx
Fwd Tx

IEEE 802.15.4 ZigBee standard

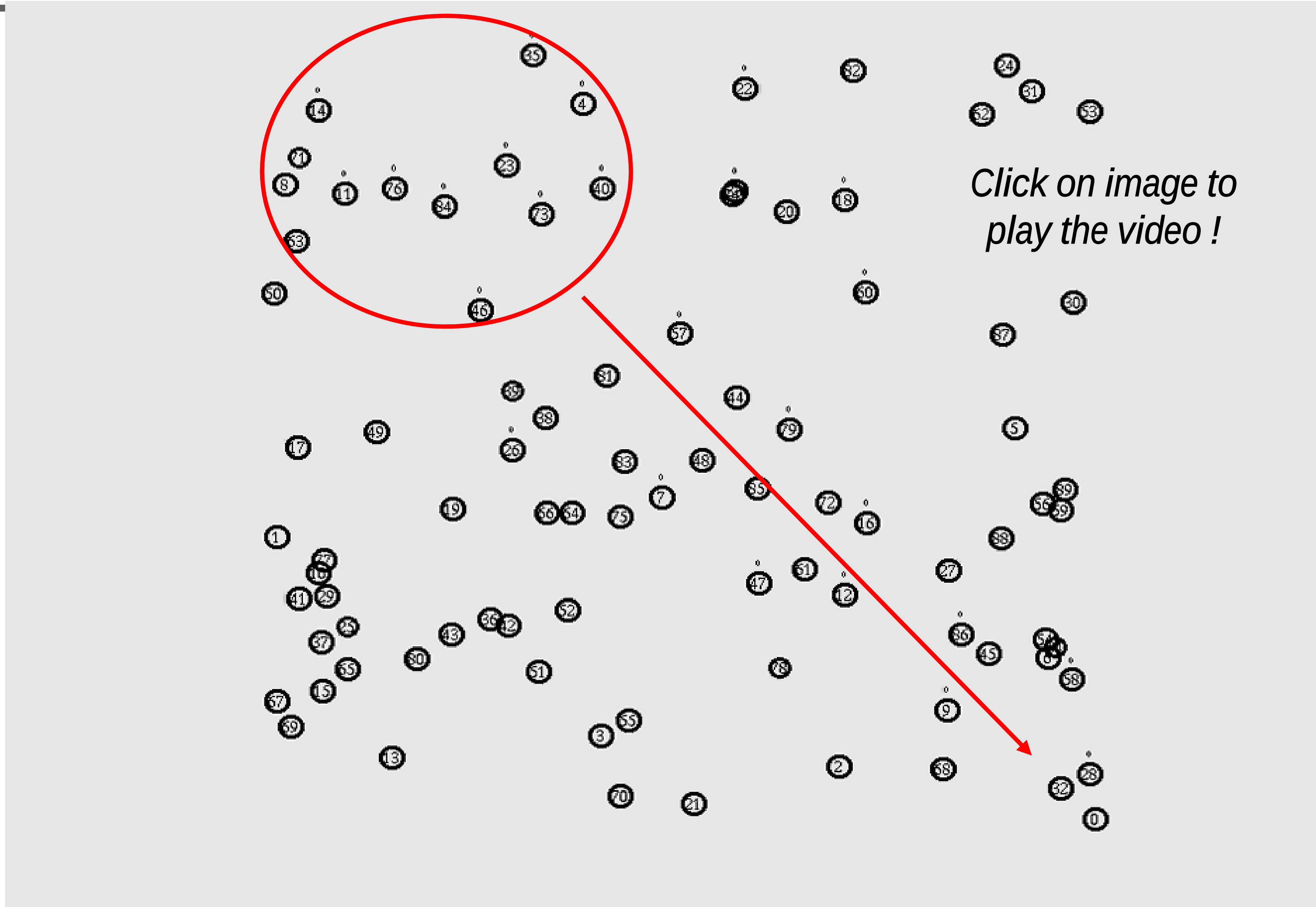


Total 1092 uW (x 36)

Doze
Tx Beacons
Rx Beacons
Fwd Rx Uplink
Fwd Tx Uplink

Event Detection in a Random Plane Network

- Transmission of initial report
- followed by the periodic transmission of measurements



First Conclusions

- WiseMAC is a single channel contention MAC protocol based on synchronized CSMA with preamble sampling
  - No setup-signaling. Self local synchronisation and re-synchronisation.
  - Adaptive to traffic: ultra-low power consumption in low traffic conditions and high energy efficiency in high traffic condition
  - Supports sporadic, periodic and bursty traffic
  - Provides years of autonomy with a single AA alkaline battery under traffic loads up to 1 message every 10 seconds
- “the WiseMAC protocol showed a remarkable consistent beha-vior across a wide range of operational conditions, always achieving the best, or second-best performance.” Langendoen 2010.
- A number of modifications of WiseMAC have been proposed
  - SyncWUF, CSMA-MPS, X-MAC, MX-MAC, …

LPL / PS practical issues

- Radios are not able to create preambles
  - Not a problem as preambles are only repeated packets
- Some radios go immediately into receive mode after sending -> Some silence between packets
  - Sample preamble twice with a delay > silence
- Noisy environment -> False wakeups (See M. Sha, et al., Energy-efficient LPL for wireless sensor networks in noisy environments. IPSN '13, pp.277-288)
  - Adjust level / sample medium twice / repeat packets
  - Adapt preamble (as in WiseMAC)

Broadcast in LPL

- Broadcasts are not acknowledged
- No possibility to get timing information
  - Thus always, long preamble

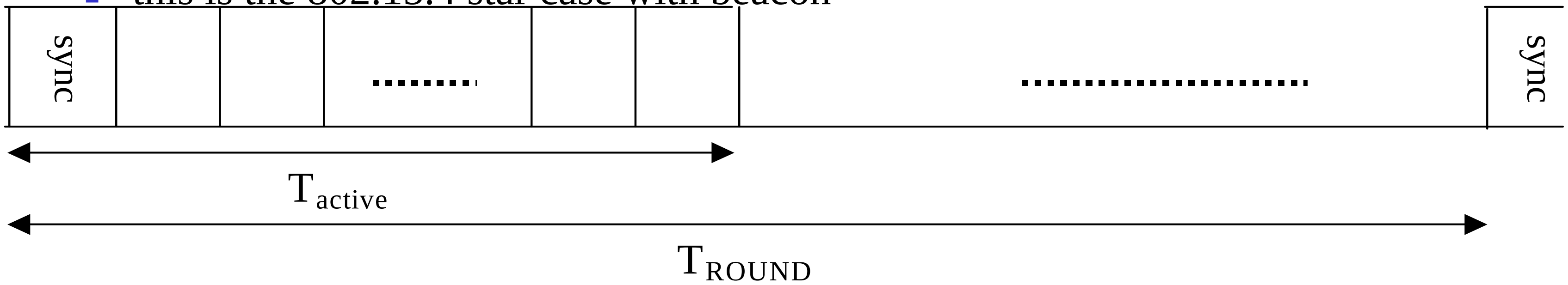
- How to reduce this overhead
  - Like with "more bit"
  - Send broadcast packets "back to back"

What about pure TDMA ?

- we know that pure TDMA
  - Is very good from the safety point of view
  - Is not flexible
  - Does not scale
  - Does not support mobility
  - Hardly supports fluctuating links (because of desynchronisation)
- but
  - It is energy efficient
    - no collisions, nodes are turned on just the necessary time
  - It is the best choice in case of high loads
- Is that so ?

What about pure TDMA ? (2)

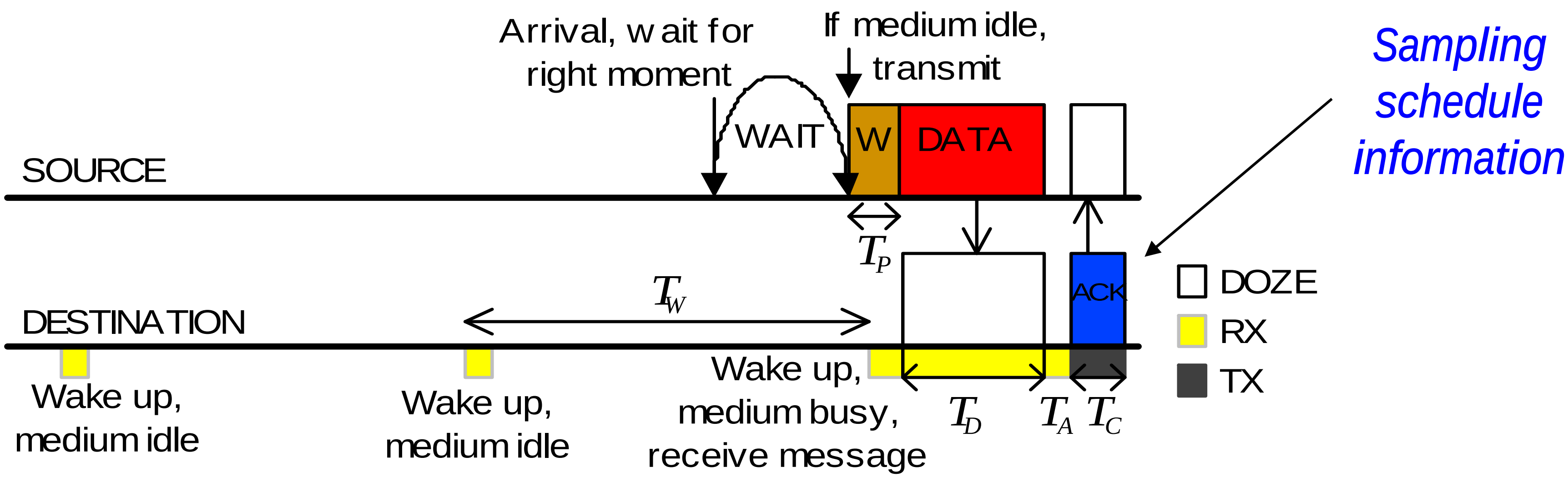
- simple case
  - all traffic has same period
    - this is the 802.15.4 star case with beacon



    - multihop network: each node receives 1 packet and transmits 1 packet per round
      - energy per round = $E_{rec} + E_{tx} + E_{rxsync} + \{E_{txsync} + 4\theta T_{ROUND}\, P_{tx}\} / N$
      - N number of participating nodes, θ clock drift
      - E_{txsync} (E_{rxsync}): energy to send (resp. receive) a sync. packet
      - E_{tx} (E_{rec}) : energy to send (resp. receive) a packet and receive (resp. send) ack
- adaptable case

What about pure TDMA ? (3)

- let us use a non TDMA protocol, Wisemac



  - assume same low traffic (1 msg sent and 1 msg receive par T_{round})
  - energy per round = $E_{rec} + E_{tx} + \{T_{ROUND} / T_W - 1\}E_{PS} + 4\theta T_{ROUND}\, P_{tx}$
  - as compared to $E_{rec} + E_{tx} + E_{rxsync} + \{E_{txsync} + 4\theta T_{ROUND}\, P_{tx}\} / N$

What about pure TDMA ? (4)

- adaptable case
  - TDMA rounds with slots assigned to links
  - a node need not emit at each round
  - one node sends a synchronisation message every K rounds
  - all nodes transmits 1 msg and receives 1 msg every L rounds
  - energy per L units = $E_{rec} + E_{tx} + L\ E_{rxsync}/K + \{L / T_{ROUND} - 1\}E_{PS} + 2\theta L\ P_{tx} + \{E_{txsync} + 4\theta K T_{ROUND}\ P_{tx}\} / N$
  - energy for Wisemac = $E_{rec} + E_{tx} + \{L / T_W - 1\}E_{PS} + 4\theta L\ P_{tx}$
  - of course, this is only valid for low traffic conditions

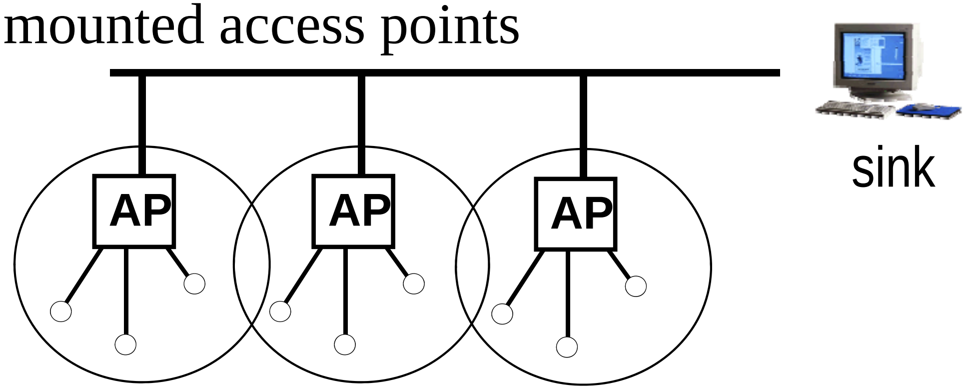
What about pure TDMA ? (5)

- what about high traffic conditions ?
  - we need some margin between the slots for desync.
  - we need some margin for retries
  - we need some margin to accommodate the longest packet

- asynchronous protocols are also able to avoid collisions by suppressing arbitration is some cases
  - more bit (Wisemac)
  - TXOP (IEEE 802.11)

The special case of Infrastructure Based Wireless Sensor Network

- Sensor nodes communicate with access points
- Access points are connected to an infrastructure network
- Access points are energy unconstrained
- Topology is made of multiple star networks. Examples:
  - Temperature monitoring in an office building, using Ethernet infrastructure
  - Alarm system at home, using Powerline infrastructure
  - Solar powered or vehicle mounted access points



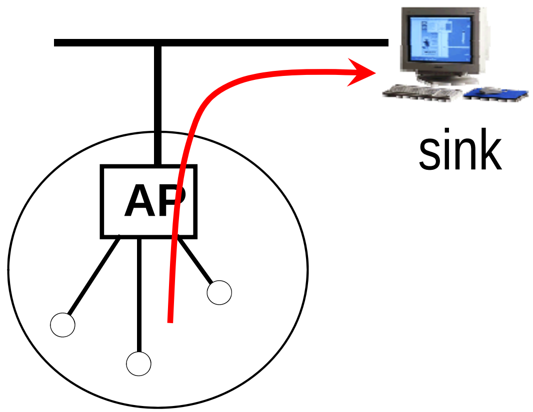
Can we design a better solution in star networks ?

- To save energy, the radio of sensor nodes must be turned off when not communicating.
- MAC protocol should minimize idle listening, overhearing and collisions.

  How to wake-up sensor nodes for receiving traffic ?
- Base station is energy unconstrained ➘
  - May listen to the channel all the time.
  - May transmit any amount of data and signaling traffic.

How to exploit the unlimited energy at the base station to save energy at the sensor node ?

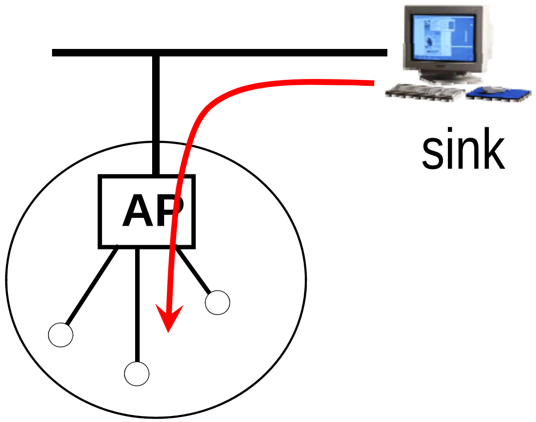
Low Power MAC Protocol for Uplink



- Kind of traffic: Periodic measurements, event detection reports
- Challenge: Minimize collisions
- Base station can listen all the time
- Traffic is far from congestion

Non persistent CSMA

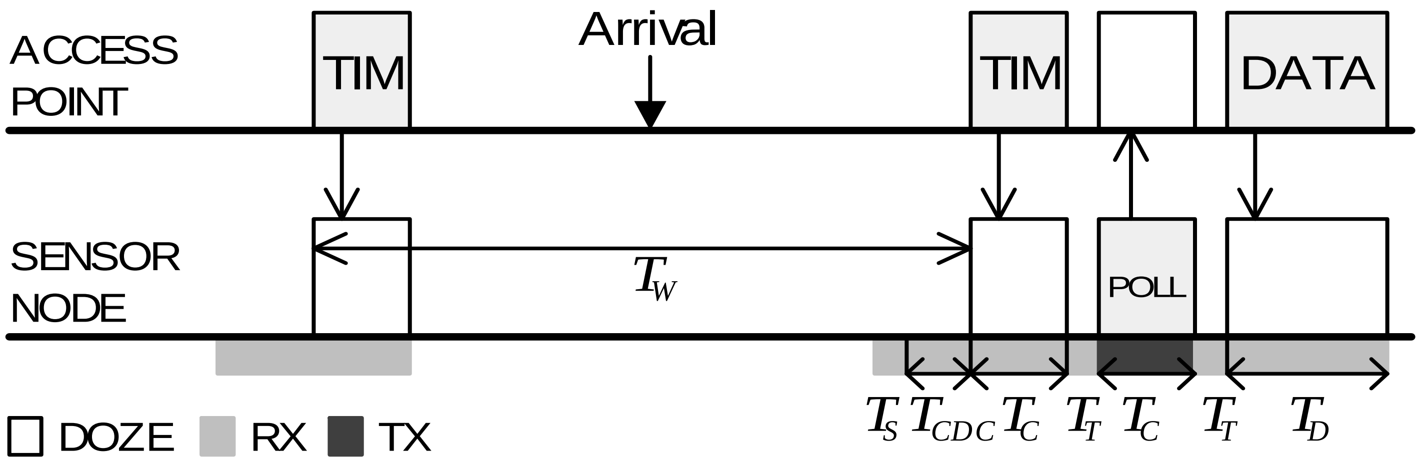
Low Power MAC Protocols for Downlink



- Kind of traffic: Queries, configuration commands.
- Challenge: Minimize idle listening and overhearing.
- MAC Protocols under consideration:
  - WiseMAC (Wireless Sensor MAC)
  - PTIP (Periodic Terminal Initiated Polling)
  - PSM (IEEE 802.11 and IEEE 802.15.4 Power Save Mode)

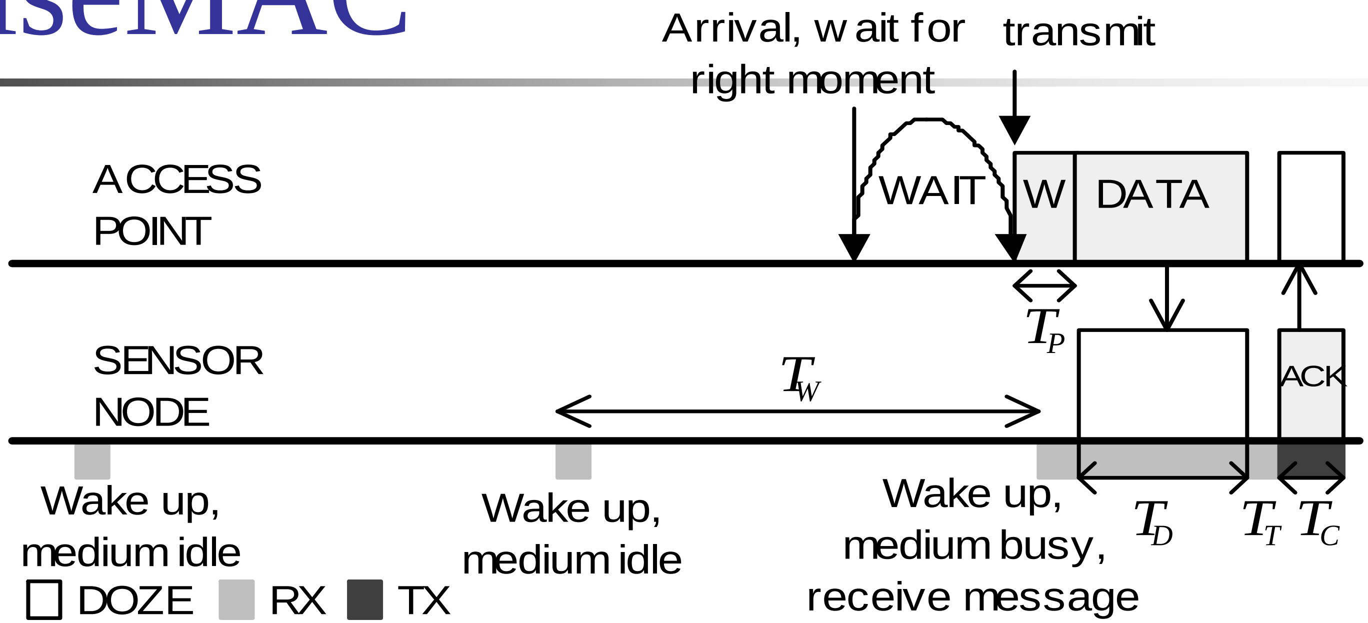
Original contributions

PSM



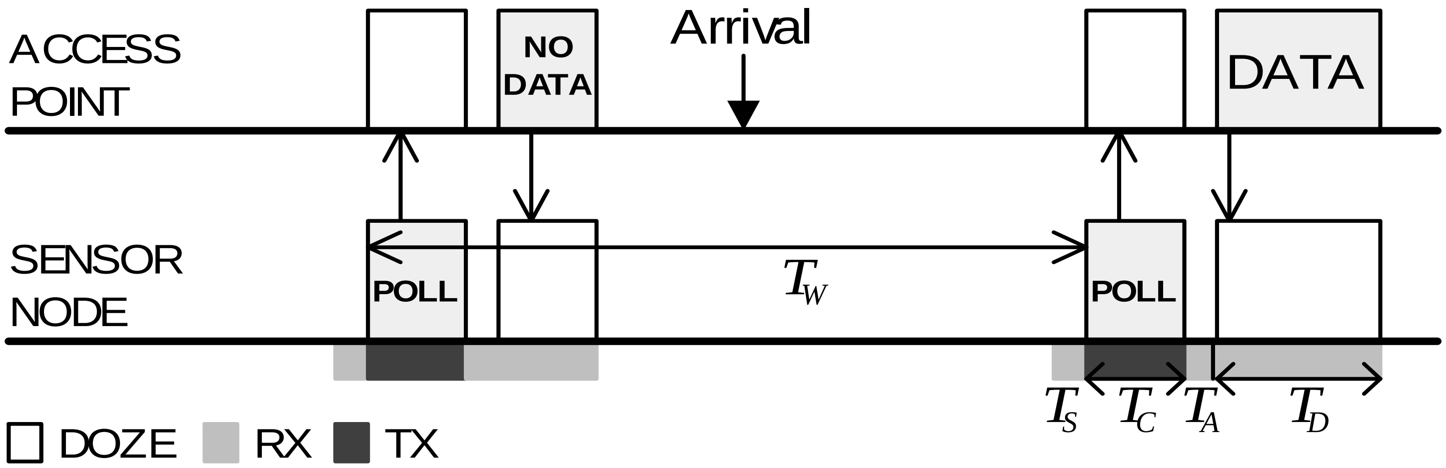
- Sensor nodes regularly wake-up to receive the traffic indication map.
- If a sensor node finds its address in the TIM, it polls the base station to receive the data packet(s).

WiseMAC



- Sensor nodes sample the medium to detect a wake-up preamble.
- Access points send wake-up preamble in front of data packets.
- APs learn the sampling schedule of sensor nodes (through ACK messages) & send messages at the right time, with a wake up preamble of minimized size.

PTIP



- *Sensor nodes periodically poll the access point to check for potential data packet(s).*

Performance Evaluation - Metrics

- Performance metrics:
  - Average power consumption P, and transmission delay D
- Protocol parameter:
  - Wake-up period T_W
- System parameters:
  - Power consumption in DOZE, RX and TX modes P_Z , P_R , P_T
  - Setup and turn-around time T_S , T_T
  - Quartz drift rate θ

Analytical Expressions

$$P_{WiseMAC} = P_Z + \frac{\hat{P}_R(T_S + 1/B)}{T_W} + \frac{\hat{P}_R(T_P/2 + T_D + T_T) + \hat{P}_T T_C}{L} + \hat{P}_R(N-1)\frac{(T_P + T_D)^2}{2LT_W}$$

$$D_{WiseMAC} = T_W/2 + T_P + T_D$$

$1/\lambda << T_D + T_T + T_C$ *(queuing negligible)*

$$P_{PTIP} = P_Z + e^{-\frac{T_W}{L}}\frac{\hat{P}_T T_C + \hat{P}_R(T_S + T_T + T_C)}{T_W} + \frac{\hat{P}_T T_C + \hat{P}_R(T_S + T_T + T_D)}{L}$$

$$D_{PTIP} = T_W/2 + T_T + T_D$$

$T_W << NT_C$ *(collisions negligible)*

$$P_{PSM} = P_Z + 2\theta\hat{P}_R + \frac{\hat{P}_R(T_S + T_C)}{T_W} + \frac{\hat{P}_T T_C + \hat{P}_R(T_S + 2T_T)}{L}$$

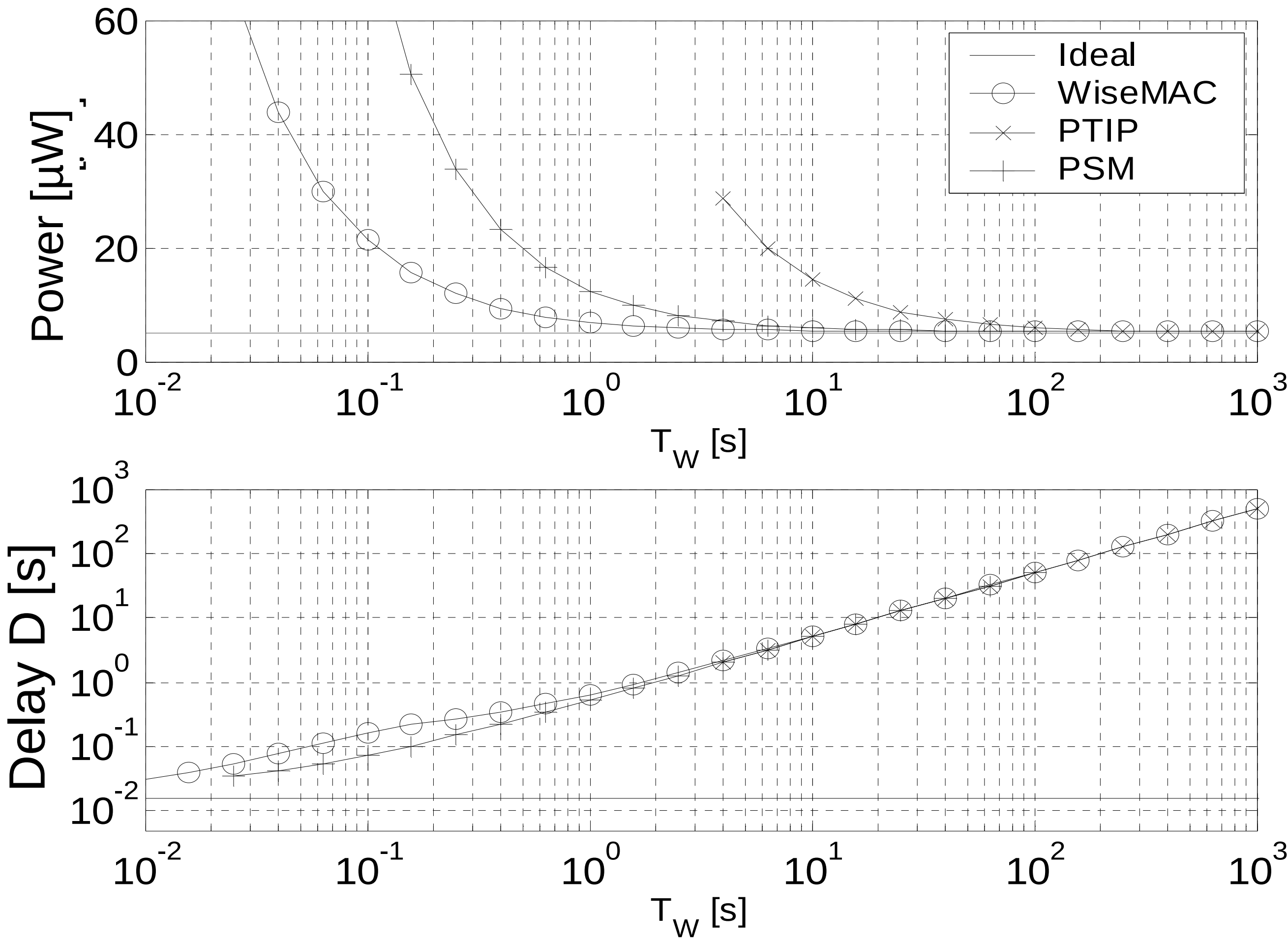
$$D_{PSM} = T_W/2 + 2T_C + 2T_T + T_D$$

$T_W \le 1/\lambda$ *(collisions negligible)*

Performance Evaluation - Results

- N=10,
- L=1000 s
- Data: 50 bytes
- Control: 10 bytes

- CSEM WiseNET® low power radio:
- P_R =1.8 mW
- P_T =27 mW



Conclusions on infrastructure (star) case

- In low traffic conditions, WiseMAC is significantly more energy efficient than PSM.

- When high latencies can be tolerated by an application, the very simple PTIP protocol becomes energy efficient.

Conclusions

- Protocols that were created for wireline comm. are not suited
- It is possible to achieve at the same time
  - Very low power consumption and low latency
- Pure TDMA may not be the solution because asynchronous solutions give similar results without the synchronisation pb.
- In WSN, communication is often the largest source of consumption
- One solution may be to switch between protocols depending on the traffic (having different modes)

- Routing has a large impact on consumption and latency

Choice

Criterion	Comments
Traffic model (deadline, period, inter arrival, ...)	Load evaluation for elimination or ranking.
Temporal guarantees	YES/NO ? Under which conditions ?
Reliability constraints	allows to reject solutions without retries
Maximum distance between nodes	allows to reject single hop solutions
Mobility or Immobility	allows to reject solutions based on long associations
Coexistence with other systems	allows to reject solutions that need planning
Dependence on Infrastructure or not	Allows to reject protocols that rely on this when this is not available
Single, Multiple sinks or Other patterns	Allows to reject protocols that do not support multiple sinks when this is needed by application
Energy constraints	If the constraint is on all nodes, this eliminates solutions with special coordination roles
Position referenced nodes	Allows reject protocols that need it when this is not available on the nodes
Simplicity	Ranking criterion

General references

- J.-D. Decotignie et al. « Architectures for the interconnection of wireless and wireline fieldbusses », Proc. FeT 2001, Nancy, Nov. 15-16, pp.285-290
- J.-D. Decotignie, « Wireless fieldbusses - a survey of issues and solutions », In Proc. 15th IFAC World Congress on Automatic Control (IFAC 2002), 2002.
- Philippe Morel, « Intégration d'une liaison radio dans un réseau industriel », Thèse EPFL, no 1571 (1996).
- D. Kotz et al.. "The mistaken axioms of wireless-network research", Dartmouth College CS Technical Report TR2003-467, July 18, 2003
- Amre El-Hoiydi et al., « Low Power MAC Protocols for Infrastructure Wireless Sensor Networks », In Proc. European Wireless (EW'04), pages 563-569, Feb. 2004.
- G. P. Halkes et al. , "Comparing Energy-Saving MAC Protocols for Wireless Sensor Networks", Mobile Networks and Applications, volume 10(5), oct. 2005, pp. 783 - 79
- R. Serna Oliver, Gerhard Fohler. Timeliness in Wireless Sensor Networks: Common Misconceptions. In Proce. RTN2010, Brussels, Belgium, July 2010
- K. Langendoen and A. Meier. « Analyzing MAC protocols for low data-rate applications ». *ACM Trans. Sen. Netw.* 7, 1, Article 10 (August 2010), 34 p.
- Rousselot, J.; El-Hoiydi, A.; Decotignie, J.-D.; "Low power medium access control protocols for wireless sensor networks, 14th European Wireless Conf, pp.1-5, 2008.

Reference (dedicated protocols)

- W. Ye et al. « Medium Access Control With Coordinated Adaptive Sleeping for Wireless Sensor Networks », IEEE/ACM Trans. on Networking, vol. 12, issue 3, June 2004, pp. 493- 506
- SMAC: W. Ye, et al., “An energy-efficient MAC protocol for wireless sensor networks”, in INFOCOM 2002, pp. 1567-76, June 2002.
- WiseMAC: A. El-Hoiydi, « Energy efficient medium access control for wireless sensor networks », PhD thesis, EPFL, Lausanne, 2005
- Crankshaft: G. Halkes, K. Langendoen: Crankshaft: An Energy-Efficient MAC-Protocol for Dense Wireless Sensor Networks. EWSN 2007: 228-244
- LMAC: L. van Hoesel and P.. Havinga, A lightweight medium access protocol for wireless sensor networks, 1st int conf. On networked sensing systems, Tokyo, 2004
- WiseMAC: A. El-Hoiydi, J.-D. Decotignie, « WiseMAC: An Ultra Low Power MAC Protocol for Multi-hop Wireless Sensor Networks », In Proc. ALGOSENSORS 2004, LNCS 3121, pp. 18-31. Springer-Verlag, July 2004.
- IEEE Std 802.15.4-2011 , IEEE Standard for Local and metropolitan area networks--Part 15.4: Low-Rate Wireless Personal Area Networks (LR-WPANs) , pp. 1 - 314

REAL-TIME NETWORKS
Real-Time Wireless Industrial Networks

Prof. J.-D. Decotignie
CSEM Centre Suisse d'Electronique et de Microtechnique SA
Jaquet-Droz 1, 2007 Neuchâtel
jean-dominique.decotignie@csem.ch

Outline

- Introduction
- IEEE 802.15.4e
- Wireless HART
- 6TiSCH and IETF proposals
- Link quality issues
- Conclusion

Introduction

- Industry does not like probabilistic approaches
  - One of the main reasons is certification
  - Tendency to use periodic traffic
- A number of pure TDMA solutions have been proposed
  - To IEC: WirelessHART, ISA 100.11a, WIA/PA, WIA/FA
  - To IETF: 6TiSCH
  - In scientific papers: RT-WiFi

IEEE 802.15.4e

- Amendment to IEEE 802.15.4
- Main innovations
  - 3 pure TDMA options (LLDN, TSCH, DSME)
  - 2 low energy options
    - one similar to WiseMAC (CSL: coordinated sampled listening)
    - receiver initiated transmissions (RIT)
  - Information Elements (IEs)
    - Together with multipurpose frames and enhanced beacons

Time Synchronized Channel Hopping - TSCH

- Fixed duration time slots (value is left to implementer)
  - Enough for sending one packet and receiving ack
- No superframe, no regular beacon, general topology
  - Each node has the notion of slotframe (repetition period) for each piece of information it sends or receives

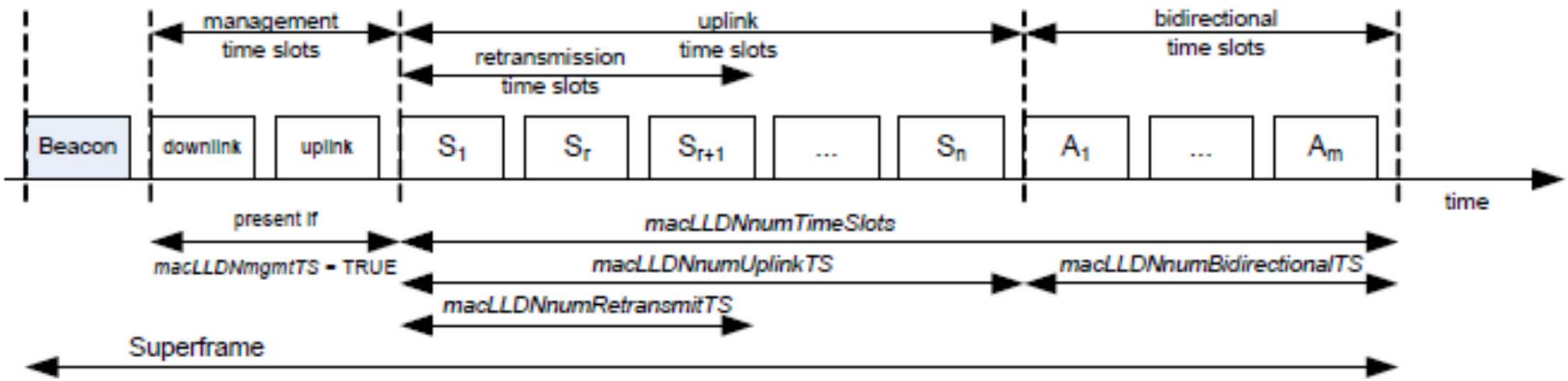
	ASN=0	ASN=1	ASN=2	ASN=3	ASN=4	ASN=5	ASN=6	ASN=7	
Slotframe 1 5 slots	TS 0	TS 1	TS 2	TS 3	TS 4	TS 0	TS 1	TS 2	...
Slotframe 2 3 slots	TS 0	TS 1	TS 2	TS 0	TS 1	TS 2	TS 0	TS 1	...

Source: IEEE 802.15.4 std

- See also [Deji Chen 2014]

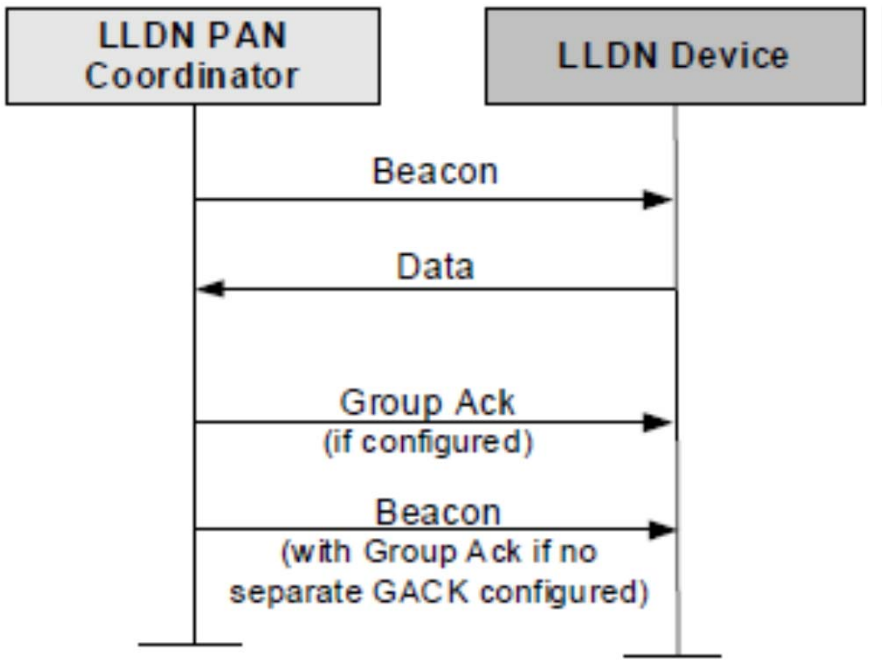
Low Latency Deterministic Networks - LLDN

- Star network with a single coordinator
  - Send the beacon in the first slot of the superframe
  - All other slots are assigned to nodes
    - Some may be assigned to multiple nodes (shared)
    - Slot type (uplink or bidir) is indicated in beacon
    - No ack in slot but in next beacon / slots reserved for retries dyn. alloc



LLDN group ack

- No immediate ack after transmission



- Slots for retransmissions allocated in a distributed manner

LLDN (cont.)

- Slot size

$$t_{TS} = (p \cdot sp + (m+n) \cdot sm + IFS)/v$$

with

- p number octets of the PHY header
- sp number of symbols per octet in PHY header
- m number octets for MAC overhead
- sm number of symbols per octet in PSDU
- n maximum number octets in data payload = Timeslot Size field value
- IFS = macSIFSPeriod symbols if m + n ≤ aMaxSIFSFrameSize octets, or macLIFSPeriod symbols if m + n > aMaxSIFSFrameSize octets
- v symbol rate

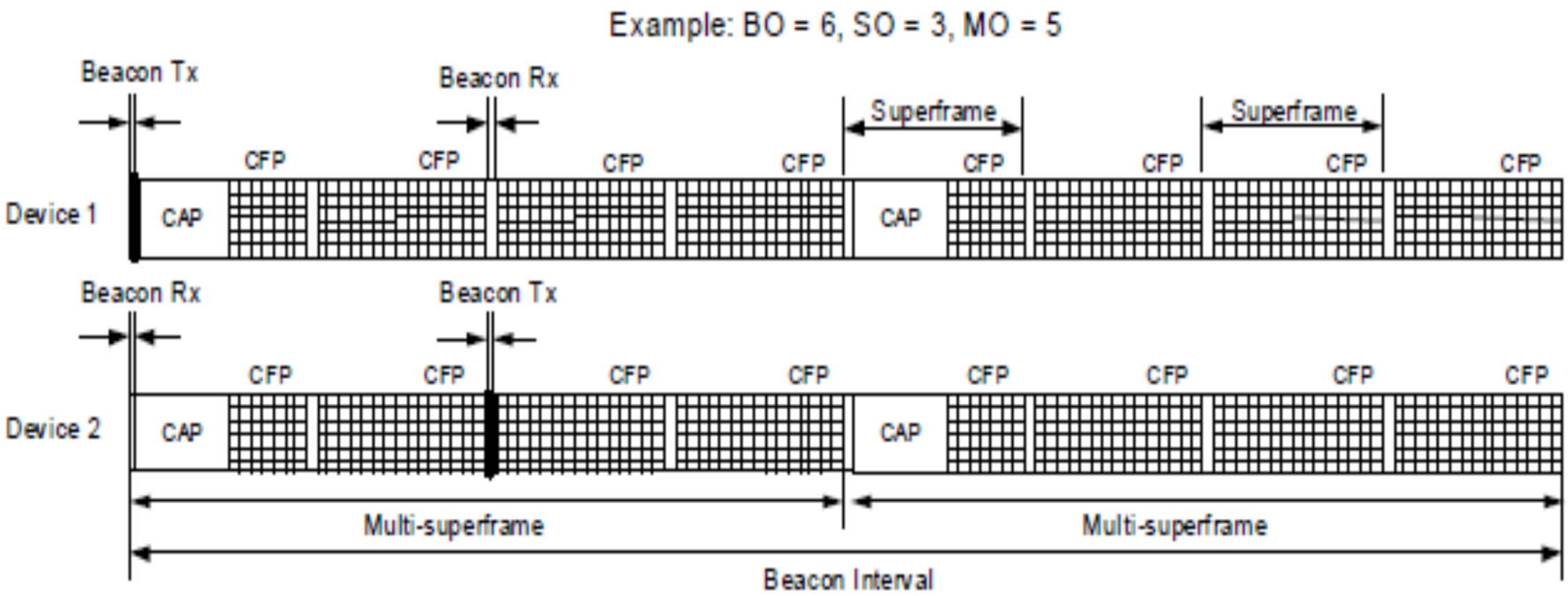
- Variable slot assignment is interesting

Deterministic and Synchronous Multi-channel Extension - DSME

- Multi-channel, multi-superframe, mesh extension to GTS for deterministic latency, flexibility & scalability
- Group acknowledgment option for high reliability and efficiency
- Distributed beacon scheduling and distributed slot allocation for robustness and scalability
  - Deferred beacon with offset indication
- Two channel diversity modes (channel adaptation and channel hopping) for robustness and high reliability even in dynamic channel conditions

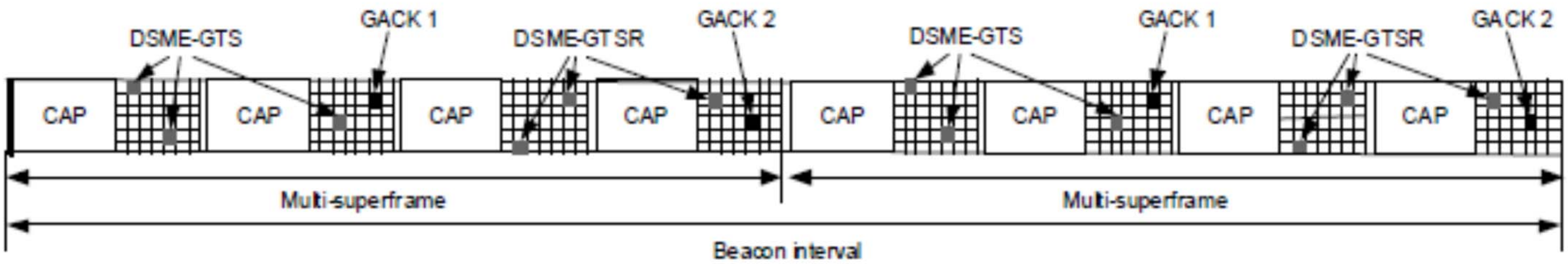
DSME (cont.)

- keeps the constraints established in 802.15.4-2011
  - 16 slots in any superframe



DSME retransmissions

- Slots may be reserved for group acks
  - GACK1 for all packets until then
  - GACK2 for all packets from GACK1 until then



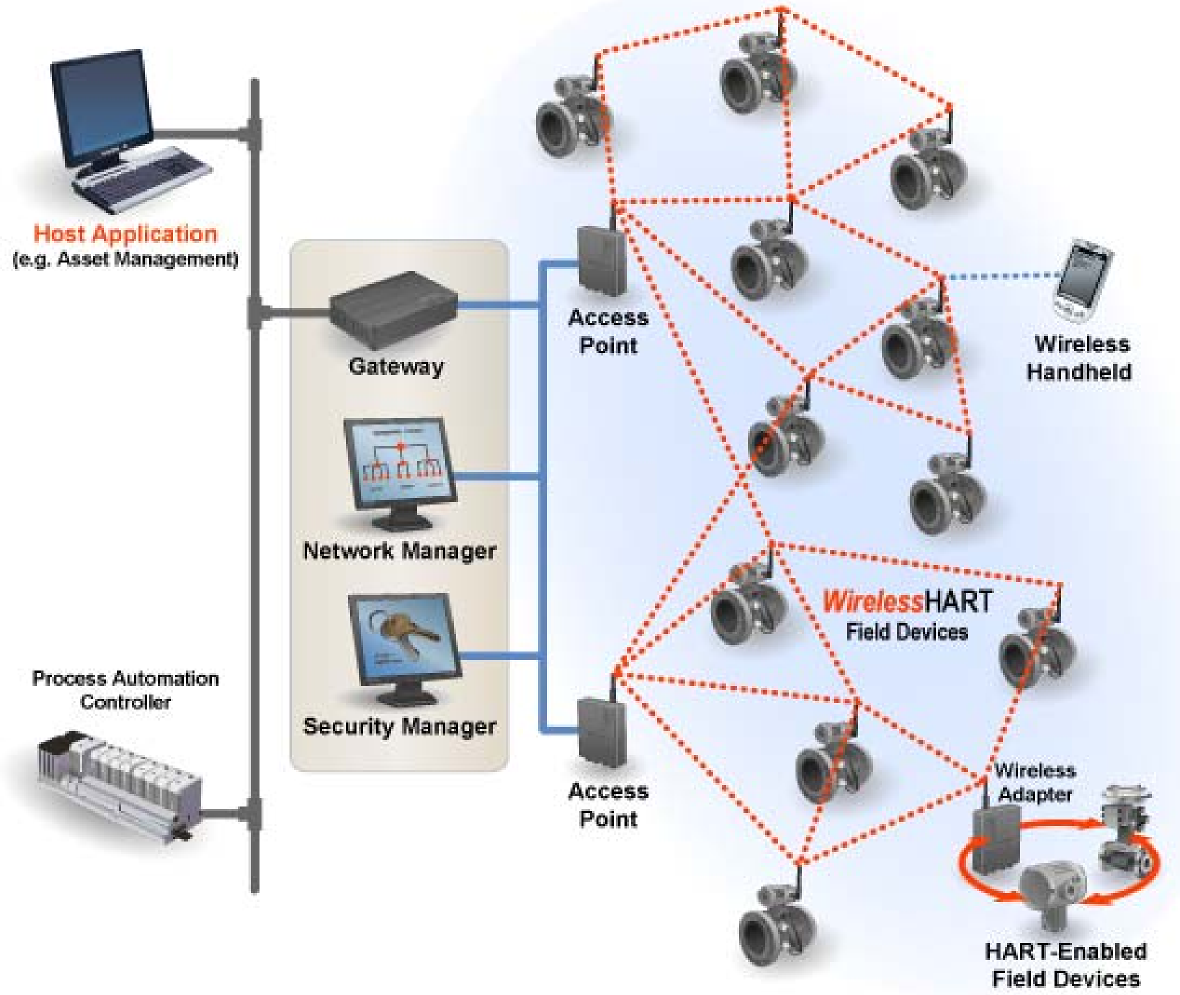
  - “The devices shall allocate an additional DSME-GTSR (i.e., GTS for Retransmission) per each allocated DSMEGTS for transmission to that coordinator”

Wireless HART

- Industry initiative (HART foundation)
- Provides a counterpart of HART using radio
- Meant for process control (rather slow)
- Now an international standard – IEC 62591
- Full solution (not only MAC)
- MAC similar to IEEE 802.15.4e TSCH
  - See [Deji Chen 2014] for the differences

Elements of a WHART network

- Field devices
  - Source
  - Sink
  - routers
- Security manager
- Network manager
  - Redundancy possible
- Gateway(s)
  - Access points



Wireless HART in short

- Uses IEEE 802.15.4 physical layer and PDUs
- TDMA MAC
  - Time and frequency diversity
  - 100 hops /s
  - Entirely configured from the network manager
- Mesh network with route redundancy
  - All field devices are possible routers
- Prioritized traffic possible
- Application layer compatible with HART
- Security based on AES-128

Physical Layer

- Physical layer: 802.15.4-2011
- Services
  - PH-CCA
  - PH-RECV-SD (start delimiter indication)
  - PH-DATA
  - PHM-SET/GET

Medium Access Control

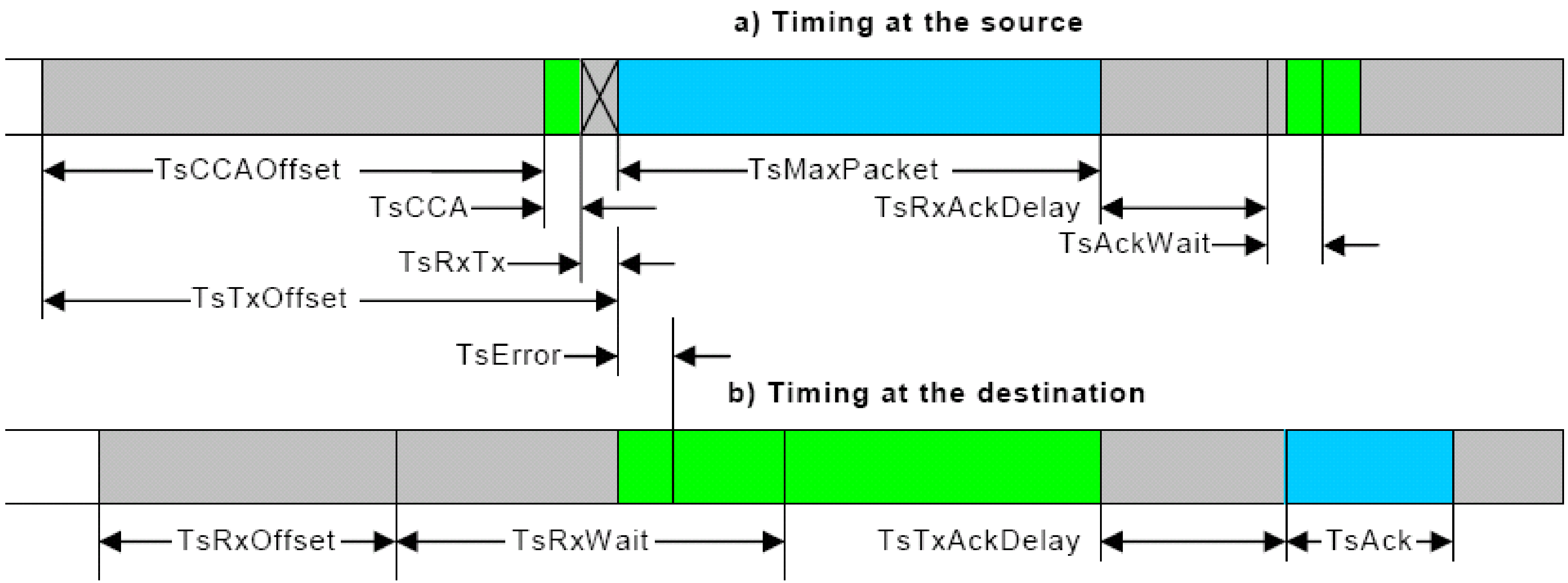
Frame 0 5 slots	TS0	TS1	TS2	TS3	TS4	TS0	TS1	TS2	TS3	TS4	TS0	TS1
Frame 1 3 slots	TS2	TS0	TS1	TS2	TS0	TS1	TS2	TS0	TS1	TS2	TS0	TS1

- TDMA
- One or more superframes of a fixed number of slots
- Channel hopping at each slot
  - There a list of used channels (same for all, some may be blacklisted)
    - 16 channels for 2.4 GHz operations
  - There is counter of the slot number since startup (ASN)
    - Active channel = (Channel-offset + Absolute Slot Nb) mod Nb_of_active_channels
- Slots (fixed size): Send / receive / shared / broadcast / join
- (superframe, slot #, channel offset) = link
  - Normal, broadcast, join, discovery
  - One or more links per device

Medium Access Control (2)

- Slot in superframe obtained by
  - SlotNb = ASN mod (Number_of_slots_in_superframe)
  - There might be more than one link in a given slot
    - Belong to different superframes
    - Transmit has priority over receive
- Rules in case more than one packet is to be sent in a given slot
- Slots may be with one source or shared
  - Shared slots are accessed using CSMA/CA
- Each device maintains a list of neighbors
  - Keep_alive and advertise PDUs

Time slot



- TsMaxPacket = 4.256 ms
- Max. 7.3ms used (43% efficiency)

Traffic specifications

- Timetables
  - For periodic transfers
    - Specifies transfer period on a connection
    - End-to-end latency is assumed not to exceed 1/3 of period
  - For sporadic (intermittent) transfers
    - Specifies maximum end-to-end latency
  - Used to filter traffic
    - Traffic that exceed what is in time table will be rejected
- Slots are assigned by NM

Traffic scheduling

- Defined in superframes
  - made of one or more slots
  - Collection of links assigned to time slots
  - One or more superframes
  - Superframe periods should follow an harmonic chain
  - Associated with a graph ID
- All superframes start at time 0+N.T (T = period)
- In case a node has to transmit (receive) on 2 or more slots at the same time, it elects the slot in the superframe with lowest ID

Traffic scheduling (2)

- A data transfer has one slot + 1 slot for retry + 1 slot on another path for 2nd retry
- There is a slot for each en-route device (router) + 2 slots for retries
- There is a management superframe (6400 slots)
  - Slots for keep_alive (3 slots each 15 minutes per node)
  - Slots for join request/resp + en-route relaying (no retry)
    - Shared with network management commands
  - Slots for ad-hoc request and response traffic (>=1 slot/min/device)
  - Special purpose slots (block transfers, hand held)
- Gateway superframe (40 slots, alternate tx/rx, shared)
  - All slots should be allocated (if no traffic then advertise)

Addressing

- Source + destination
- 8 bytes IEEE EUI 64
- Or 2 bytes unique address within a network
- 2 bytes network ID
- Broadcast address is 5 bytes long (all =0) ???

Data Link Layer

- Data receive service – DL_Receive (.ind)
- Data transfer service – DL-Transmit (.req, .ind., .cnf)
  - Includes retries / supports multicast and broadcast
  - May carry the next hop neighbors as param (graph)
  - If broadcast indicates superframe id to be used
- Event service (connect, disconnect, path failure,
- Management service (Set, Get, Action
- Security (no encryption but authentication)
- QoS
  - Priority (Command, process data, Normal, alarm)
    - Traffic below a given priority may be prohibited (Priority_Threshold)
  - timeout

DLE / NLE tables

- Superframe table
- Link table
  - Each link belongs to one and only one supeframe
- Neighbor table
  - Initialy sent by network manager, updated by the node
- Graph table
  - Next hop destinations for upward or downward traffic
  - Directed list of paths

DLL maintenance

- Discovery
  - Advertise DLPDU
    - ASN, graph ID, list of supeframes, list of links
    - Sent periodically using a random period between 0 to discovery_time
  - Keep Alive DLPDU
  - Nodes continuously listen for advertisements (in discovery links) and join requests
    - Update the neighbor list, Communicate list of neighbors to NM
- Time keeping
  - Some nodes are clock references
  - Nodes measure difference between expected receive time & effective one
    - Difference is transmitted in ack packets
    - Used by non reference nodes to adjust their clock

Network layer

- End to end security (3 keys: join, network mgr, gateway)
  - Activated by network management
  - Keys distributed by NM
- Transmit – NL-Transmit
- Management – SET/GET/Action
  - Action
    - Add/delete session, add/delete route, add/delete timetable, default route, reset

Routing

- Defined by network manager
- 4 types
  - Graph route
    - Any of the next hops indicated in the graph
    - Graph_ID selected by initial source device upper layers
  - Source route
    - Route is decided by source entity (contained in NPDU)
    - Used to test the routes
  - Broadcast
    - NL indicates superframe to be used (any broadcast slot in that superframe may then be used)
  - Proxy (joining device)

Graph routing

- Normal routing technique
- Multiple graphs in a single network
- Directed links
  - Graph is undirectionnal
- Redundant
  - Link is choosen localy for each NPDU. How ???
- NPDU conveys graph id
  - Each intermediate node must have the local view of the graph
  - This is stored in connection tables
- Superframe routing (superframe_id instead of graph_id)
  - Forwards to any neighbor that has a link in the superframe

Transport layer

- Data transfer
  - Unacknowledged – TL-DATA-Transfer
  - End-to-end acknowledged - TL-DATA-Exchange
    - No loss, duplication, reordering
    - Only one pending transaction at a time
  - There is a maximum time to complete and a maximum number of retries (TL parameter)
- Management
  - SET/GET

Acknowledged transport

- Uses pipes (optional for non ack)
  - Unicast data exchange slave with NM
  - Broadcast data exchange slave with NM
  - Unicast data exchange slave with gateway
  - Master for event notification to NM
  - Master for event notification to the Gateway

Time synchronisation

- Some nodes are time synchronisation sources
  - Used to synchronize the other devices
  - How are they synchronized ????
- Each device records the difference between the expected time of arrival and the actual one
  - This is used returned in the ack packet
  - This is used to correct local clock
- Keep_alive pdus may be used in case there is not enough traffic (max. each 30s)
  - 10 ppm clock drift per device

Application layer

- Client-server model
  - A number of services to access/modify variables and configuration
- Pre-defined basic types, structured types
- Rules for encoding variables and services
- Possibility to publish data (to client)
- Possibility for event notification (to client)

Data Types

- Fixed length
  - Integer8,16,24,32 - Unsigned8,16,24,32,40 - Float32,64
  - Data, Tim
  - Enumeration, Bit Field
  - Security Key, Unique ID, Engineering Unit
- String
  - Packet ASCII
  - ISO Latin-1
- Structure, Array
  - Nesting in allowed
- Encoding in defined in IEC 61158-6-20

Application Layer

- Virtual Field Device ASE
  - identity
- Variable ASE
  - Services to read/write variable objects + information report
- Action ASE
  - Read, write, reset, self-test,
- Device application services
  - Access and modify device attributes
- Layer management services
  - Set and read node parameters

Application relationship

- ARs are composed of a set of endpoints of compatible classes. One endpoint of each AR has to be Master class and the other end has to be Slave class.
- Each device can have only one instance of a Slave class AREP.
- A device can use more than one Master class AREP to communicate with several Slave class AREPs.
- The user at Master class AREP provides the identification of the Slave class AREP as a parameter in the request primitive.

Conclusion on Wireless HART

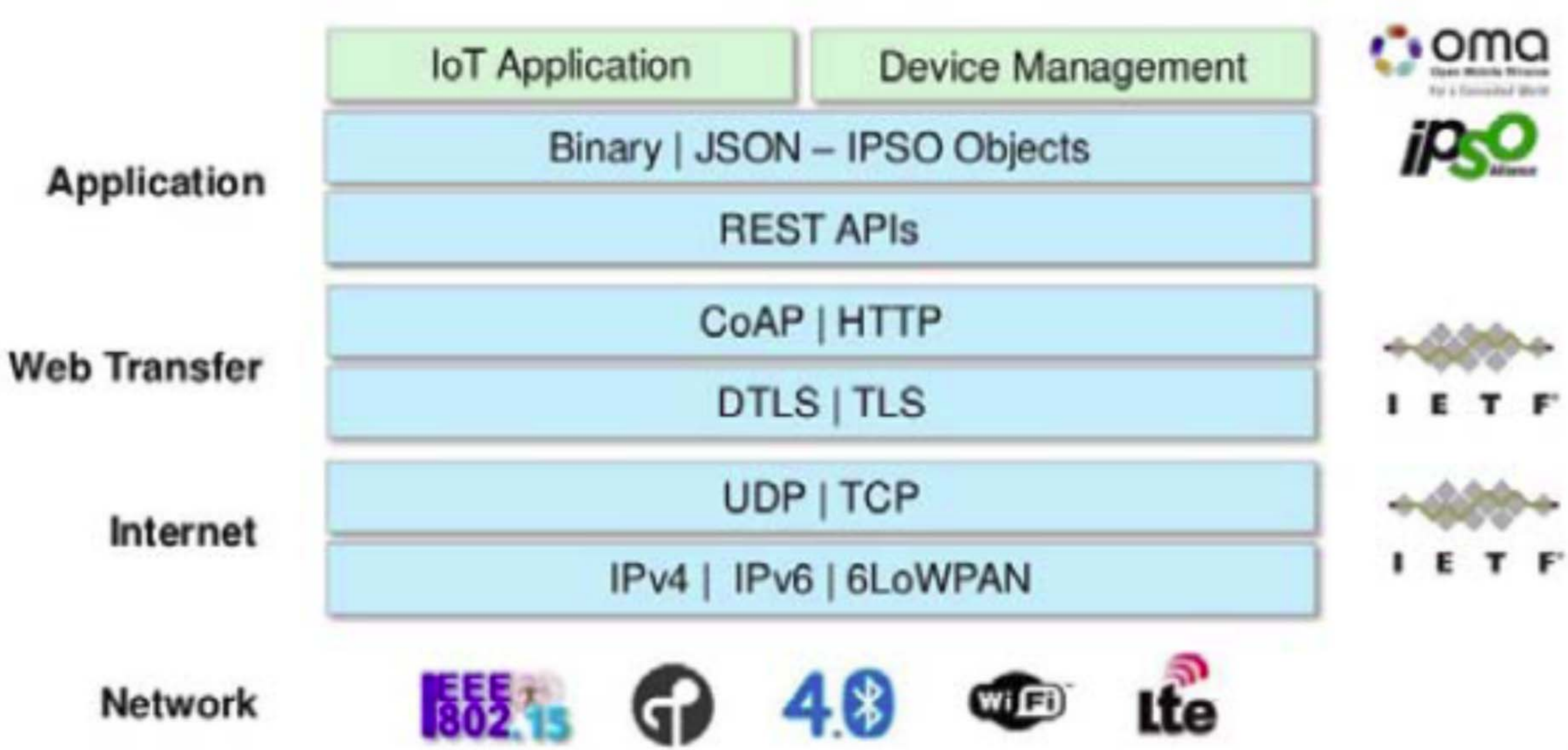
- pros
  - A complete solution
  - Good robustness to interferences and errors
    - Route diversity, retransmissions, channel hopping
- Cons
  - Complex (difficult to schedule)
  - Rather inefficient (fixed size slots, resources allocated for 2 retries)
  - Centralized (single sink)

Wireless HART References

- IEC 62591 Ed. 1.0: Industrial communication networks – Wireless communication network and communication profiles – WirelessHART™, document IEC 65C/587/FDIS
- S. Han et al, Reliable and Real-Time Communication in Industrial Wireless Mesh Networks, IEEE RTAS 2011, pp. 3-12
- Deji Chen et al. WirelessHART and IEEE 802.15.4e, IEEE ICIT, pp. 760-5, 2014.

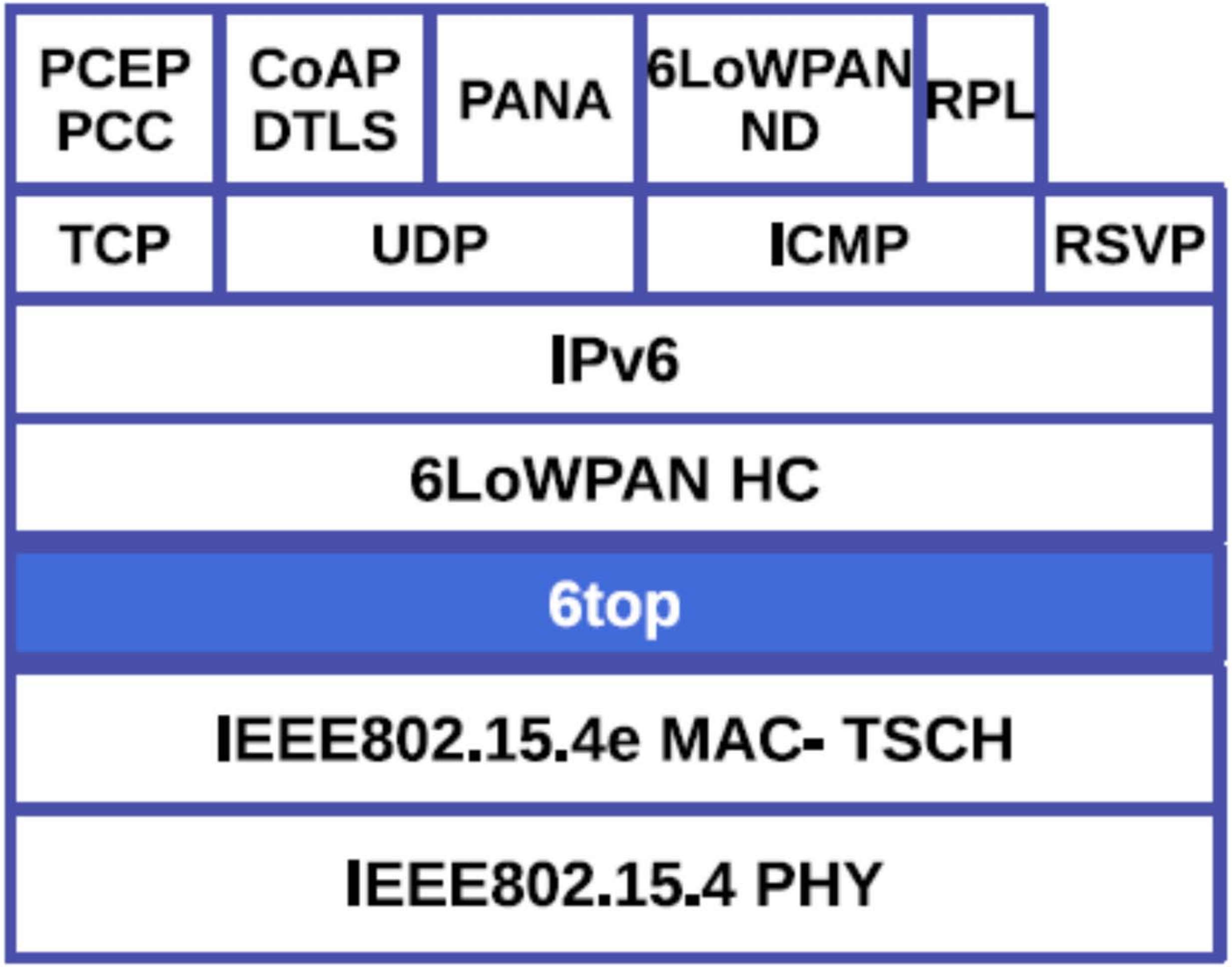
Internet Engineering Task Force

- Low-power lossy networks: CoAP, RPL, 6LoWPAN



Source: http://postscapes.com/internet-of-things-protocols#graphics

IETF & 6TiSCH



Source: Palattella, 2016

From System Engineering Guidelines IEC 62591 WirelessHART, Engineering Guidelines 00809-0100-6129, Rev AB, February 2016

- It is recommended that wireless field devices used for control and high speed monitoring have a higher path stability than general monitoring devices with updates slower than two seconds.
- Path Stability is the measure of successfully transmitted messages on any given path relative to the attempted transmissions. General requirements are 60 percent path stability, but 70 percent is recommended for control and high speed monitoring. The addition consideration provided in this text ensures higher path stability that can be confirmed once the network is deployed. Most WirelessHART vendors provide the means to verify after installation.

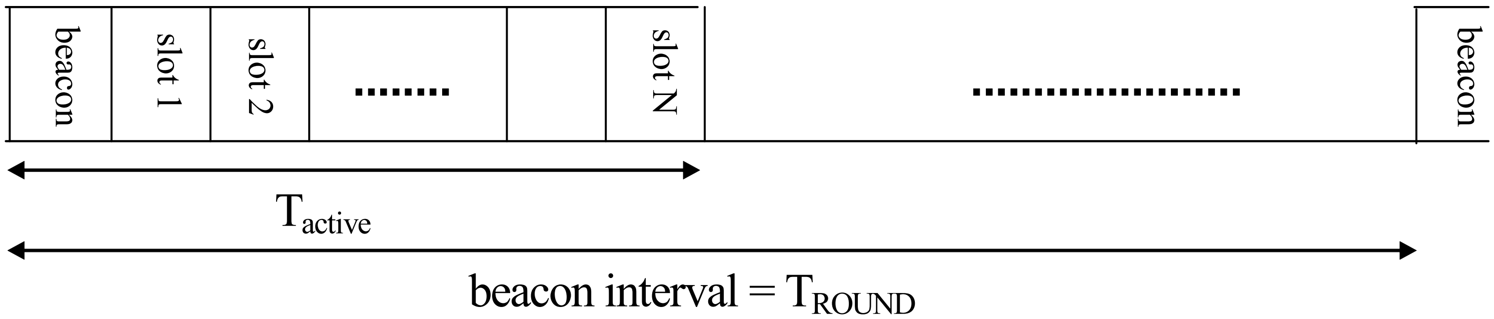
Conclusion on industrial solutions

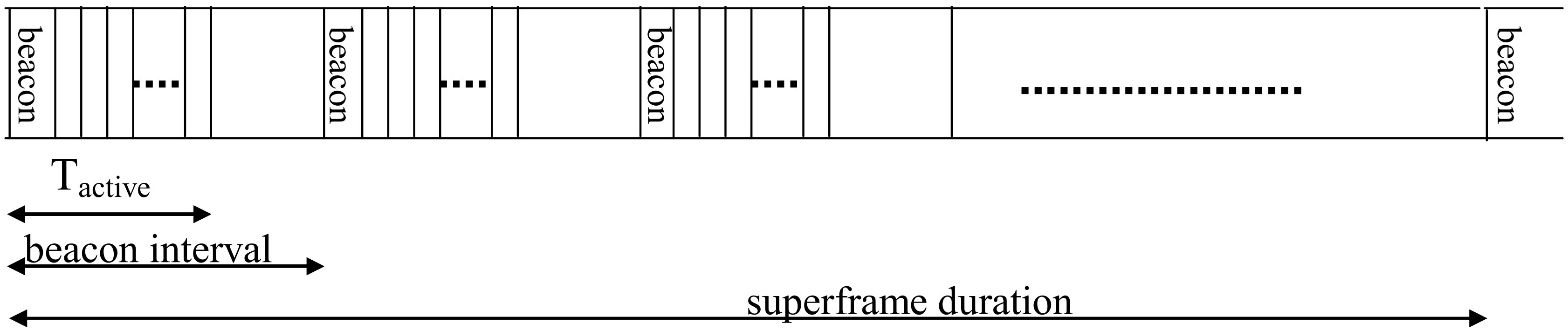
- Having a good protocol is considered as not sufficient
- It is necessary to install the devices (or at least the antennas) so that links are good and stable

- But, what if we cannot have good links ?

Retransmission scheme

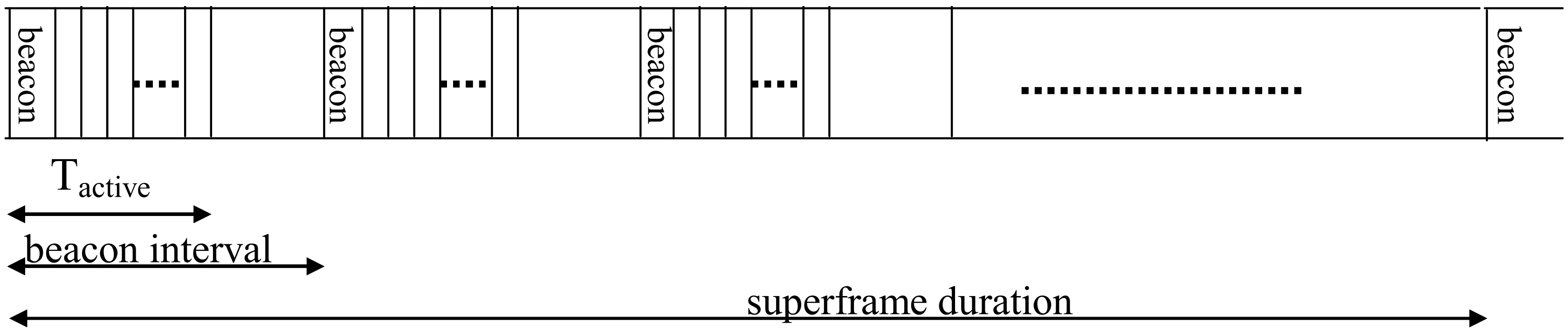
- Wireless transmission prone to errors (e.g. BER 10-4)
- Having an efficient retransmission scheme is important





Retransmission schemes

- Assign slots for retries in a fixed manner
  - 1 slot for transmission and KR slots for retry (in KR beacon interval)

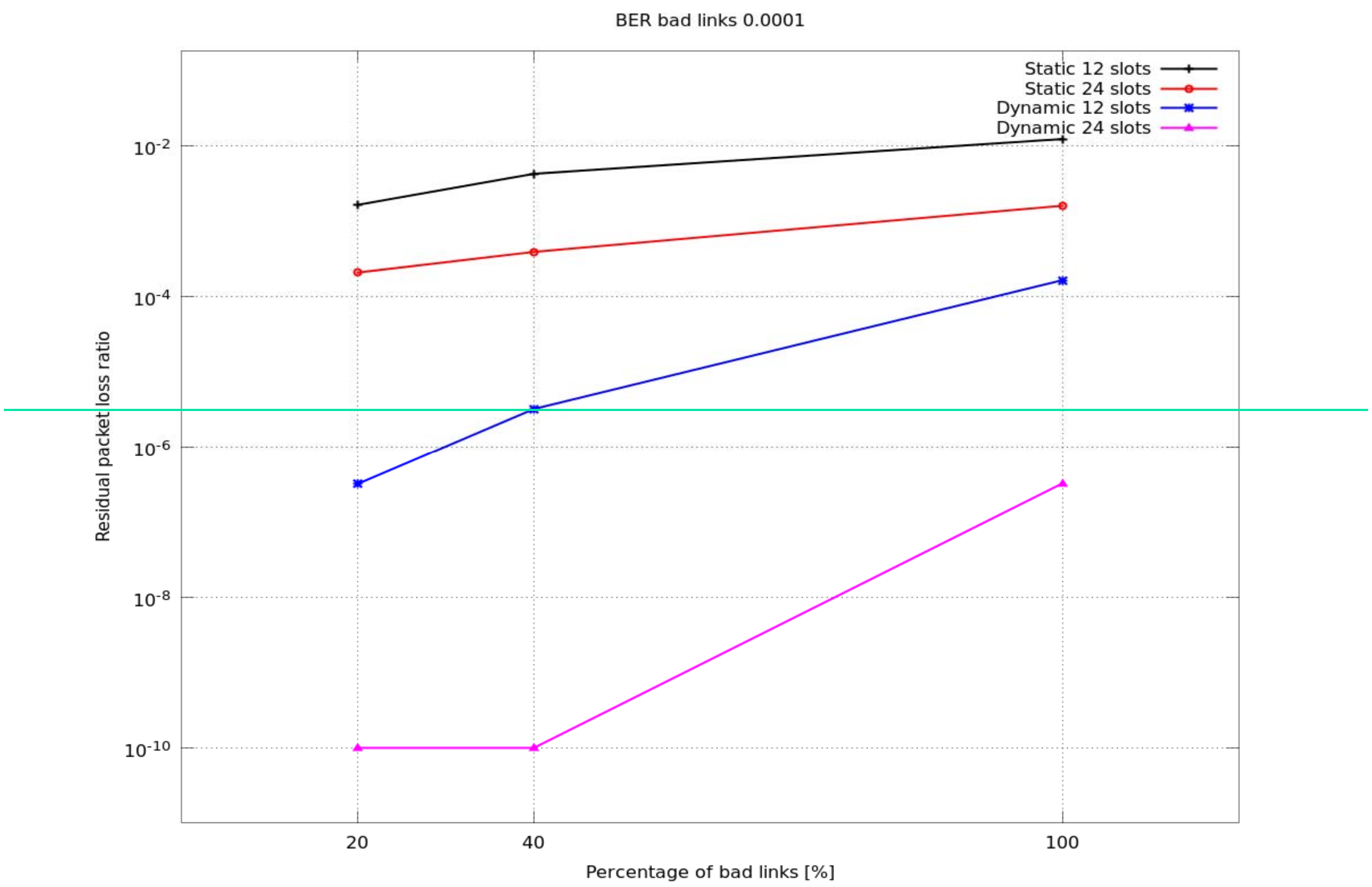


- Assign slots for retries in a dynamic manner
  - In each beacon interval, there are NSR slots for retries
  - The slots are assigned dynamically to recover from failures in previous beacon interval

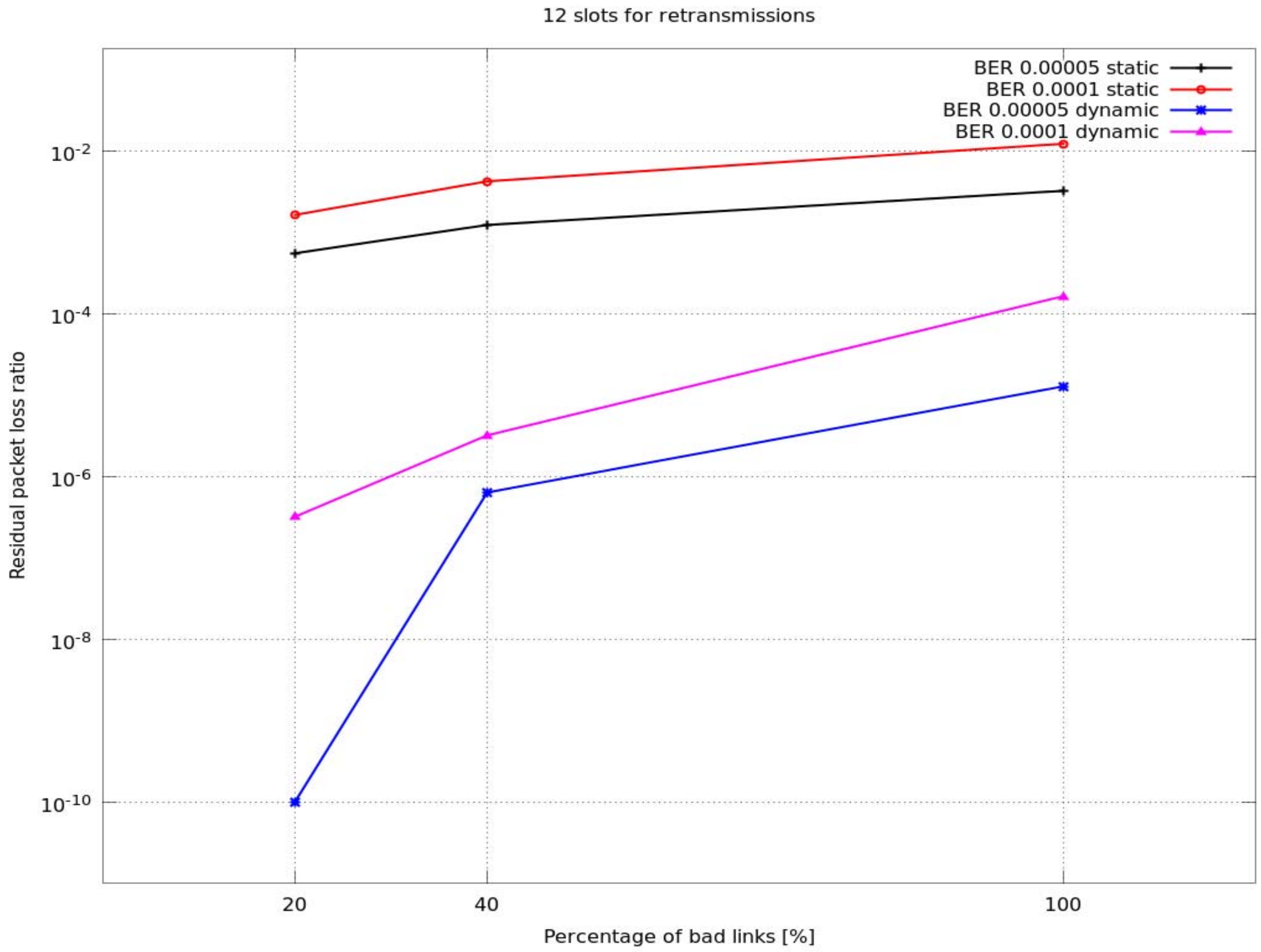
Parameters

- slot time = 150 µs, 127 Byte packet size
- Application period (superframe duration) = 12 ms
- Beacon interval: TROUND=3 ms
- Number of sensor nodes: N=12.
  - 12 slots in the 1^{st} beacon interval in addition to the beacon.
- IR-UWB packet: PHY bitrate:
  - 1 Mb/s for the short preamble, 27 Mbit/s for the data.
- good links that have a BER of 10-6
- bad links that have a BER varying from 5.10-5 to 10-4.

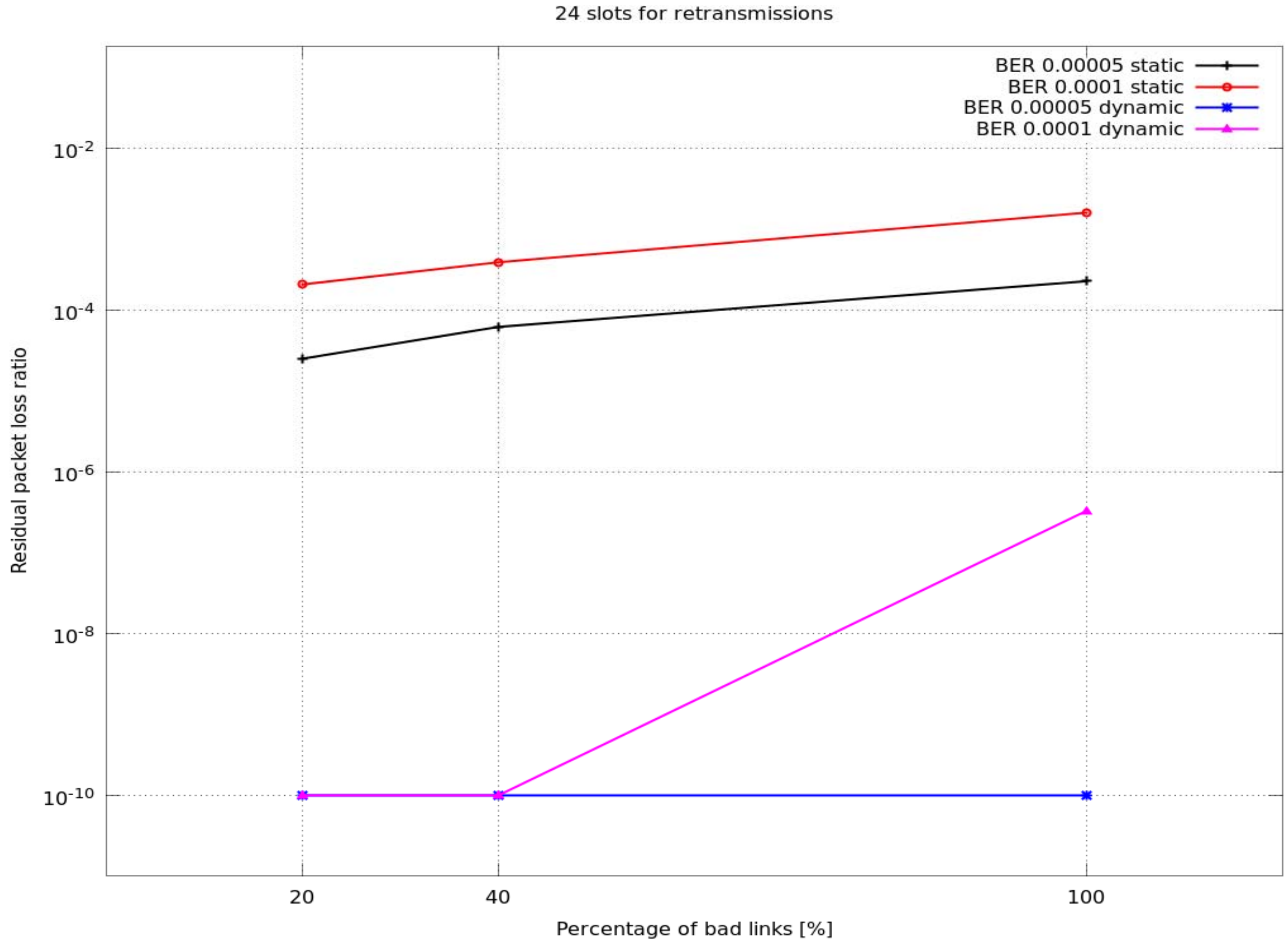
Residual PER as a function of retry policy (BER bad links = 10-4)



Residual PER as a function of BER

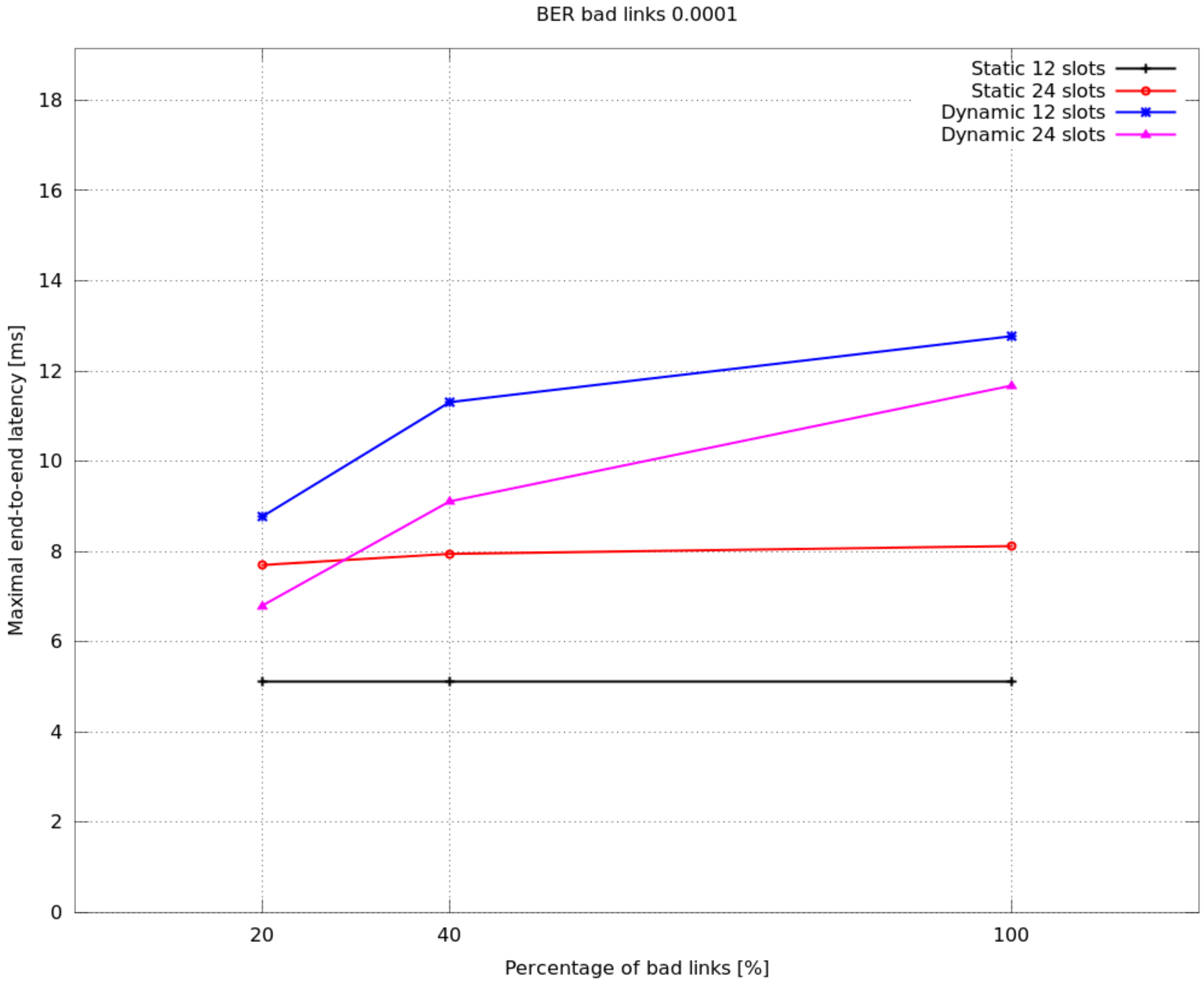


Residual PER as a function of BER



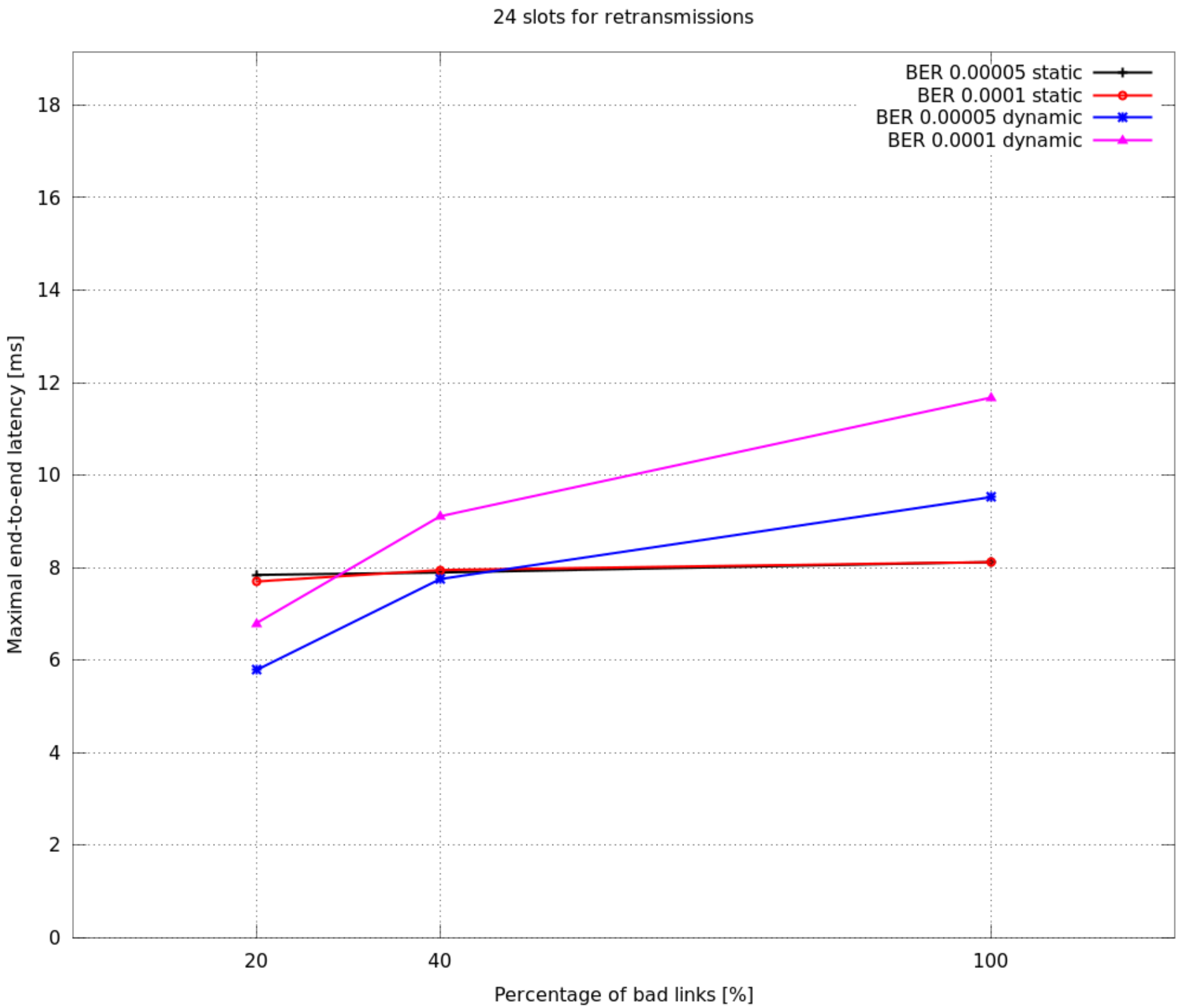
Latency as a function of retry policy

- Max ->



- Mean latency from 1.3 to 1.6ms

Maximum latency as a fct of retry policy



Conclusion

- Pure TDMA has been heavily used in industrial communications
- Different solutions with different application scopes
  - Hard to have “one size fits all”
- Most retransmission schemes are static
  - Not very efficient in particular when link reliability is uneven

References

- P. Dallemagne, J.-D. Decotignie et al., "Suitability of the IEEE 802.15.4e extensions for spacecraft and launcher communications", DASIA 2014
- Deji Chen et al., “WirelessHART and IEEE 802.15.4e”, 2014 IEEE Int. Conf. on Industrial Technology (ICIT), p 760-5, 2014
- G. Alderisi et al., “Simulative assessments of the IEEE 802.15.4e DSME and TSCH in realistic process automation scenarios”, INDIN 2015, pp. 948-55
- G. Patti et al., “Introducing multi-level communication in the IEEE 802.15.4e protocol: The MultiChannel-LLDN”, ETFA 2014, pp. 1-8.
- M. Palattella et al., “On-the-Fly Bandwidth Reservation for 6TiSCH Wireless Industrial Networks”, IEEE Sensors Journal, v 16, n 2, p 550-60, Jan. 2016
- Yi-Hung Wei et al., “RT-WiFi: Real-Time High-Speed Communication Protocol for Wireless Cyber-Physical Control Applications”, RTSS 2013, pp. 140-9

References - standards

- IEEE Std 802.15.4e-2012 (Amendment to IEEE Std 802.15.4-2011), IEEE Standard for Local and metropolitan area networks--Part 15.4: Low-Rate Wireless Personal Area Networks (LR-WPANs) Amendment 1: MAC sublayer, pp. 1-225
- IEC 62591:2016, Industrial networks - Wireless communication network and communication profiles - WirelessHART™
- IEC 62601:2015, Industrial networks - Wireless communication network and communication profiles - WIA-PA
- IEC 62734:2014, Industrial networks - Wireless communication network and communication profiles - ISA 100.11a
- www.ietf.org for RPL, CoAP and 6LoWPAN

REAL-TIME NETWORKS

Recapitulation

Prof. J.-D. Decotignie

CSEM Centre Suisse d'Electronique et de Microtechnique SA

Jaquet-Droz 1, 2007 Neuchâtel

jean-dominique.decotignie@csem.ch

Wired networks covered

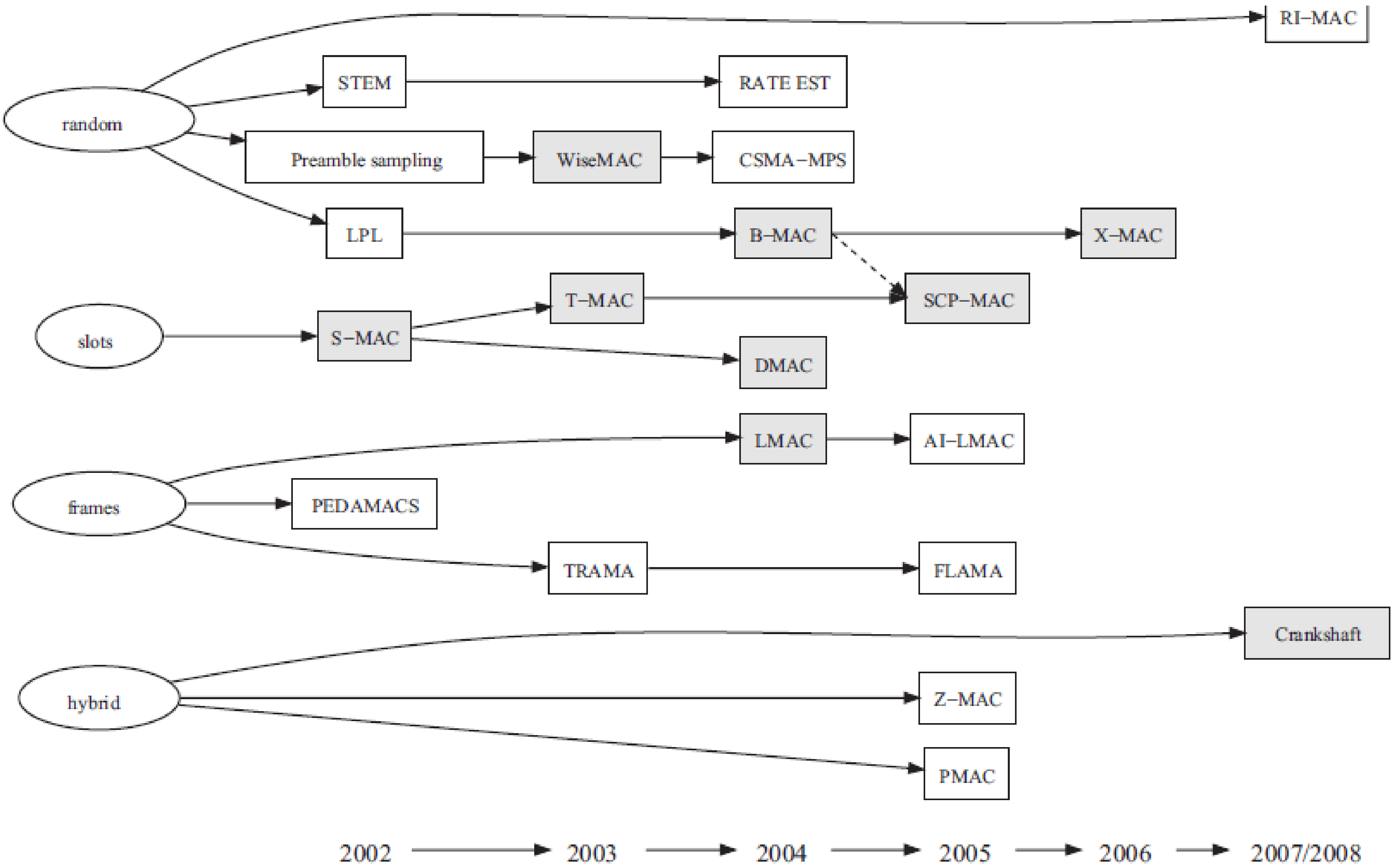
- CAN
- WorldFIP
- Ethernet
  - “pure”
  - Additional rules (e.g. traffic shaping)
  - Additional MAC

Wired Networks

- Most of the real-time properties at MAC level
- Do not forget inaccessibility
- The higher the line speed, the higher the overhead of request/response protocols
- Switched networks are able to exploit both directions simultaneously

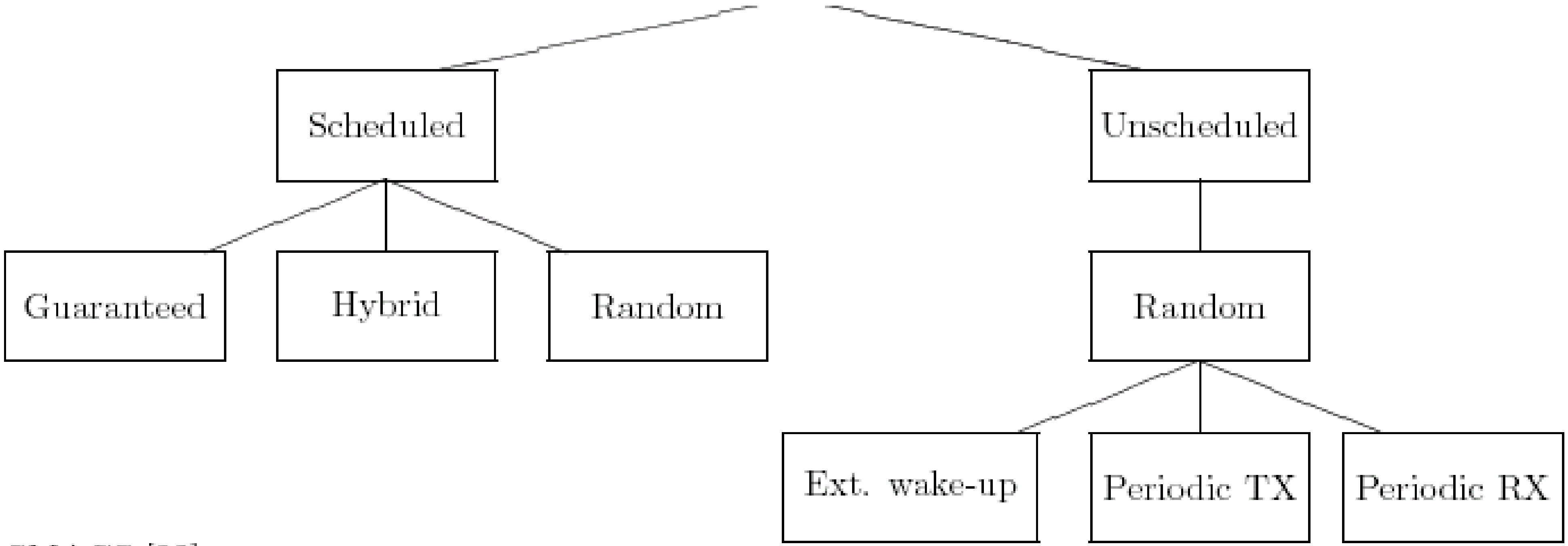
- Time Sensitive Networking (TSN) is the current effort to make Ethernet real-time

Wireless sensor networks



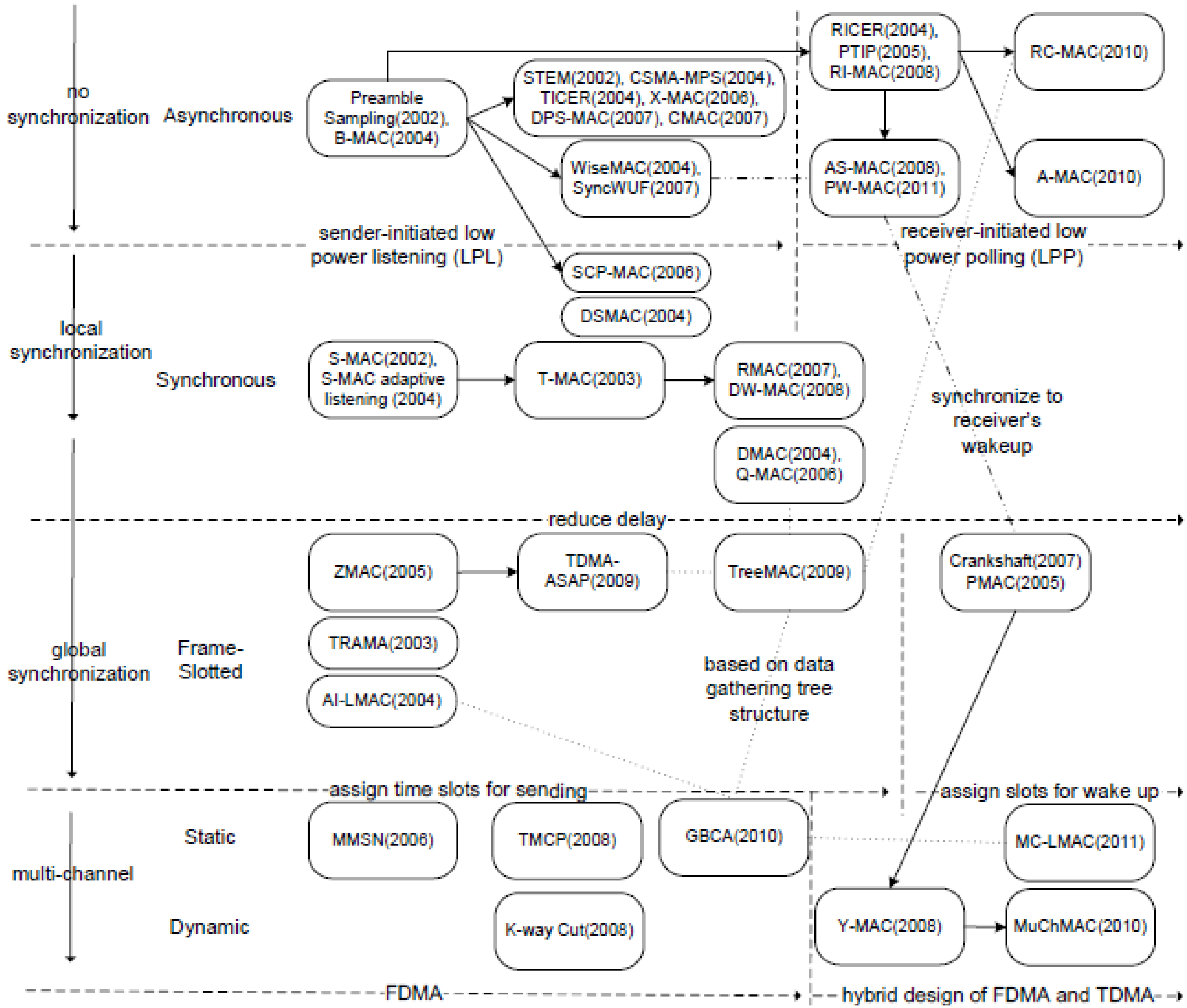
Source K. Langendoen and A. Meier. « Analyzing MAC protocols for low data-rate applications ». *ACM Trans. Sen. Netw.* 7, 1, Article 10 (August 2010)

Another WSN taxonomy



SMACS [83]					
LEACH [36]				RI-MAC,	BMAC
PACT [65]		S-MAC [99]		802.15.4e	WiseMAC
BMA [51]	EMAC [92]	T-MAC [90]			X-MAC….
BitMAC [73]	LMAC [93]	D-MAC [53]	PicoRadio [70]		

More recent survey



P. Huang et al, The Evolution of MAC Protocols in Wireless Sensor Networks: A Survey, IEEE Communications Surveys & Tutorials, vol. 15 (1), pp. 101-120, 2013

Myths

Wireless transmission issues

- The world is flat & radio transmission area is circular
  - signal strength is a simple function of distance
- All radios have equal range
- Link quality does not change
  - if I can hear you, you can hear me & if I can hear you at all, I can hear you perfectly
- The only source of packet loss is collision
- Broadcast is for free
- Energy is proportional to the number of packets and their size
- Duty cycling is the only way to reduce energy consumption

Propagation !

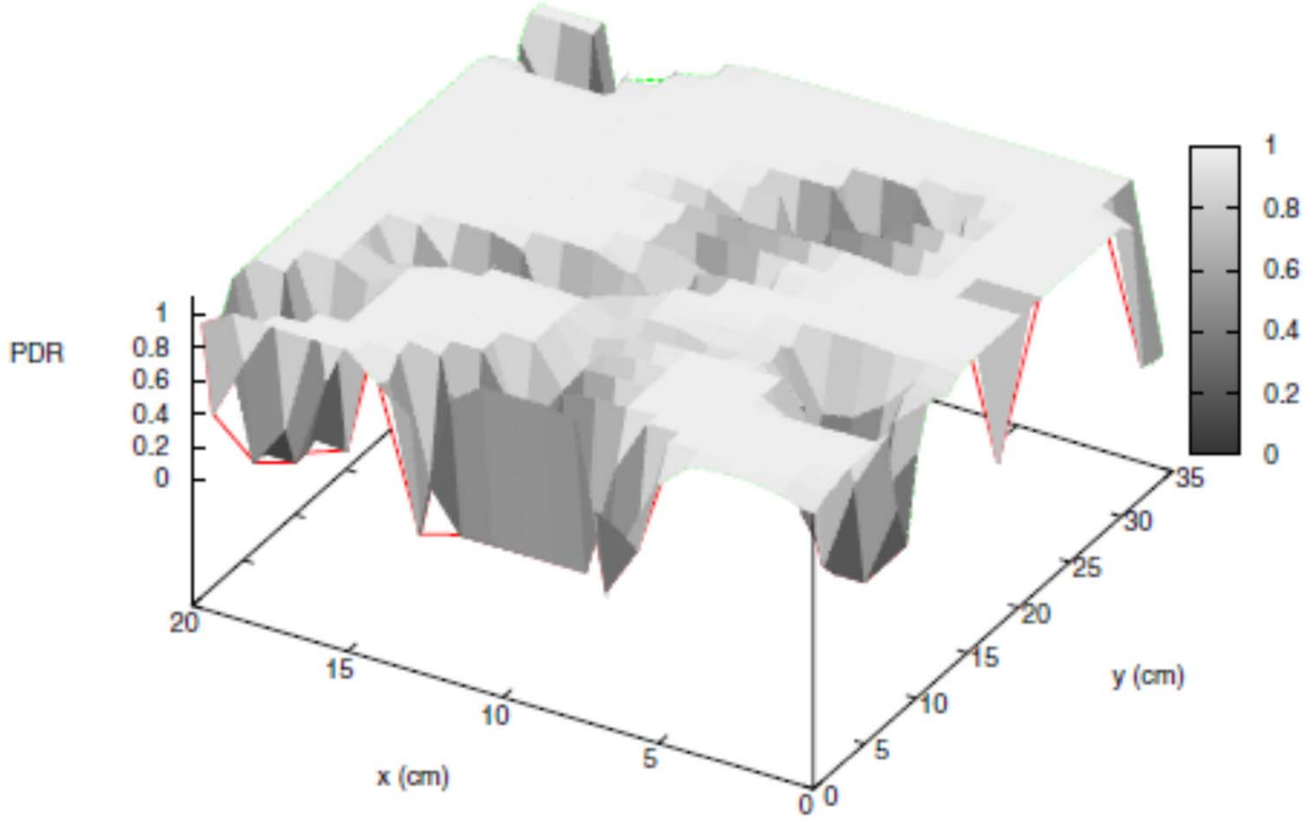
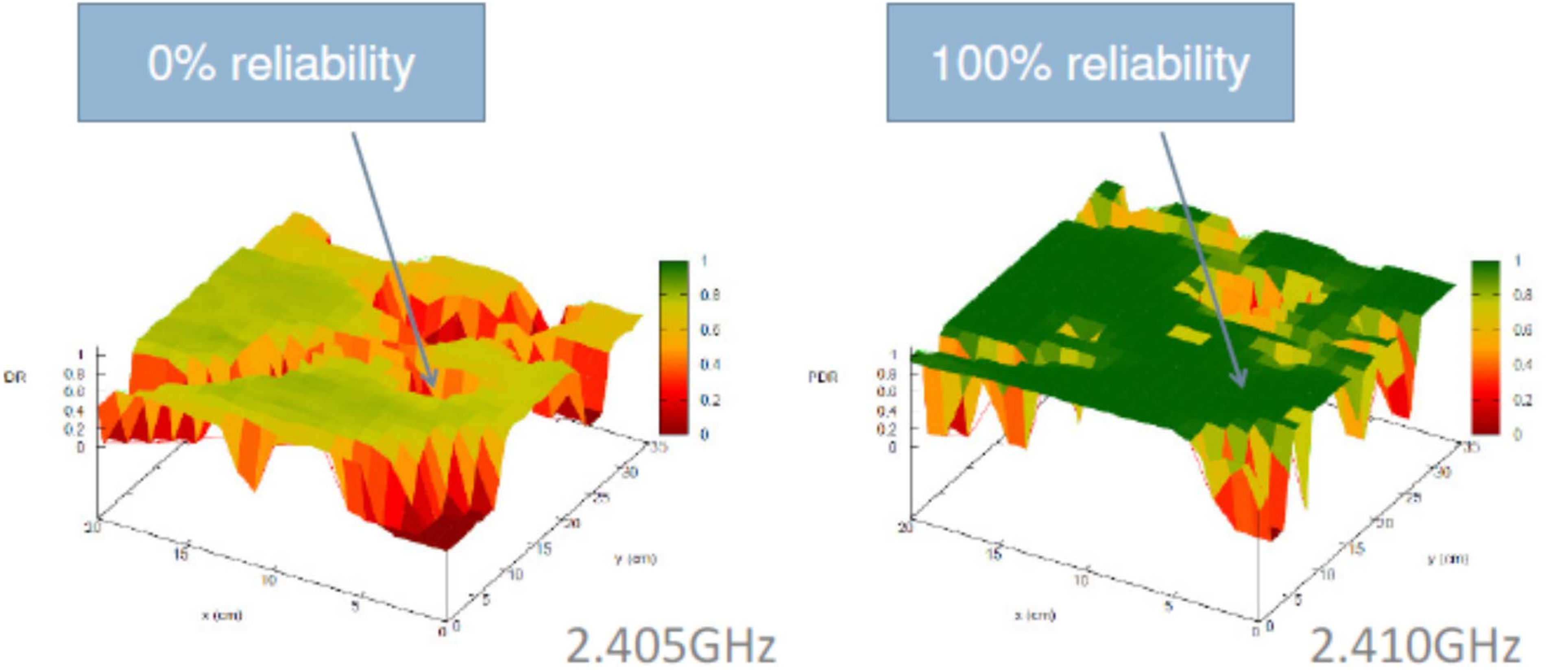


Fig. 1. Witnessing multipath fading. The x and y coordinates represent the position of the transmitter on a $20cm \times 34cm$ area; the receiver is static. The z axis (and the shade) represent the Packet Delivery Ratio, PDR. Results obtained for sender and receiver communicating on IEEE802.15.4 channel 20 (2.450GHz) while separated by 1m; transmission power is set to -16dBm.

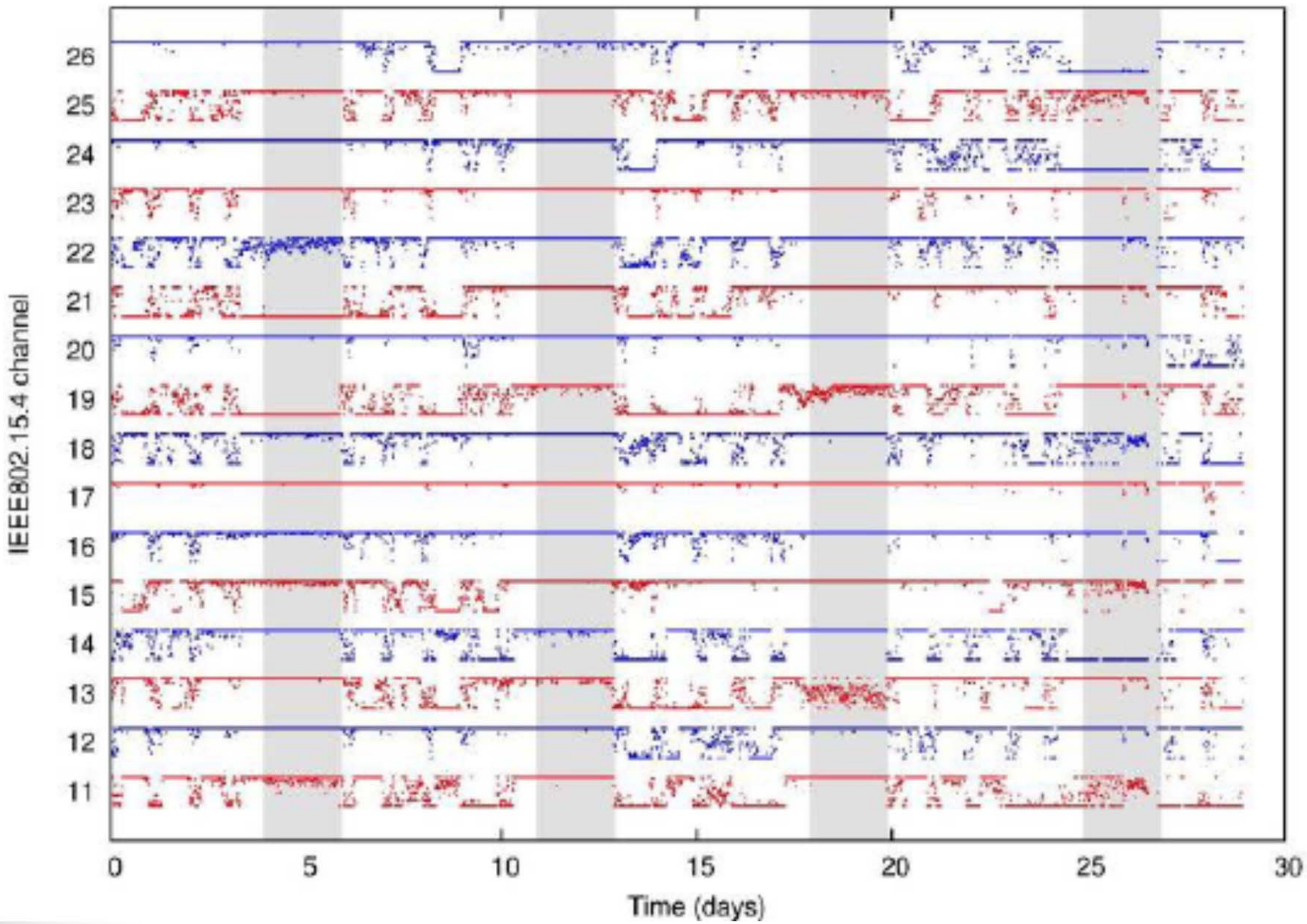
Source T. Watteyne, ICC 2010, p.1

Propagation !



Source T. Watteyne, DREAM Seminar 2014

Propagation over time



Source T. Watteyne, DREAM Seminar 2014

Wireless Networks issues

- RT guarantees
  - Often fail in case of coexistence
  - CS and systematic channel hopping are good ways to support coexistence
    - CDMA should also be but requires power control
  - Routing should also be considered
- Energy consumption
  - Be aware that traffic is seldom constant
    - Fixed assignments are an overkill when traffic varies a lot
    - Future is likely to be in adaptive protocols or clever combinations of protocols (see S. Mo et al., Self-Adapting MAC Layer for Wireless Sensor Networks, RTSS 2013, pp.192-201)

Wireless Networks options

- TDMA
  - Good when traffic is constant (and also centralized)
  - Beware of coexistence (e.g. by having systematic hopping)
  - Beware of mobility
  - Is able to reach high channel utilization (with some care)
- Asynchronous (random) protocols
  - Good in terms of energy when traffic is low
  - Good for coexistence
    - Careful with coping with channel noise (M. Sha, et al., Energy-efficient low power listening for wireless sensor networks in noisy environments. IPSN '13, pp.277-288)

Choice

Criterion	Comments
Traffic model (deadline, period, inter arrival, ...)	Load evaluation for elimination or ranking.
Temporal guarantees	YES/NO ? Under which conditions ?
Reliability constraints	allows to reject solutions without retries
Maximum distance between nodes	allows to reject single hop solutions
Mobility or Immobility	allows to reject solutions based on long associations
Coexistence with other systems	allows to reject solutions that need planning
Dependence on Infrastructure or not	Allows to reject protocols that rely on this when this is not available
Single, Multiple sinks or Other patterns	Allows to reject protocols that do not support multiple sinks when this is needed by application
Energy constraints	If the constraint is on all nodes, this eliminates solutions with special coordination roles
Position referenced nodes	Allows reject protocols that need it when this is not available on the nodes
Simplicity	Ranking criterion

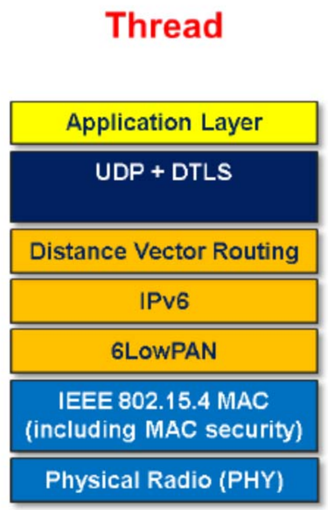
Other aspects

- Higher layers in particular application layer
  - May limit real-time (e.g. web services)
- Dependability (fault tolerance)
- Safety

- Tools
  - For deployment
  - For test
  - For management

Most serious contenders

- IETF (6LoWPAN, CoAP, ROLL/RPL)
  - MAC agnostic
  - Thread (Google) bears some similarities
- IEEE 802.15.4 (different options including with QoS)
- ZigBee (uses 802.15.4)
- Bluetooth Low Energy (multihop since V4.2)
- Wireless Industrial (Wireless HART, ISA 100.11a)
- Home automation: Ant+, KnX, EnOcean, Z-Wave, WirelessMbus
- Long range low bandwidth (SigFox, LORA, Weightless)
- MQTT (-SN) is quite popular to connect to the cloud



RFC 1925 Fundamental Truths of Networking

- (3) With sufficient thrust, pigs fly just fine. However, this is not necessarily a good idea. It is hard to be sure where they are going to land, and it could be dangerous sitting under them as they fly overhead.

RFC 1925 Fundamental Truths of Networking (2)

- (7a) (corollary). Good, Fast, Cheap: Pick any two (you can't have all 3).

- (12) In protocol design, perfection has been reached not when there is nothing left to add, but when there is nothing left to take away.