Switched Ethernet solutions for embedded real time networks

Jean-Luc Scharbarg
Université de Toulouse - IRIT - ENSEEIHT/INPT

August 2015

ETR'2015

Agenda

1. Context
   - Avionics network architecture
   - Automotive network architecture
2. Ethernet
   - Ethernet Evolution
   - AFDX
   - Ethernet-AVB
   - TTEthernet
3. Some open issues

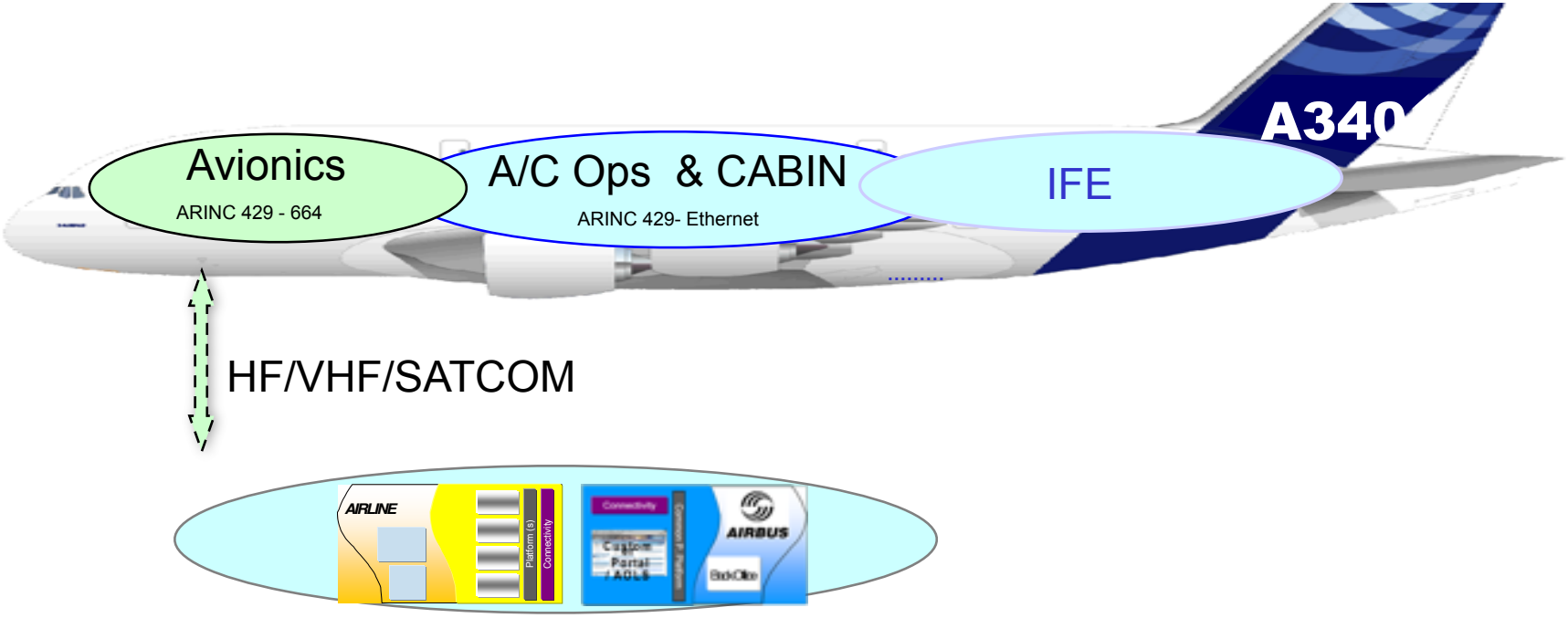
Agenda

Historical evolution of embedded networks

- Early days: one single computer for all embedded functions
- Then distribution of the functions on computing nodes
  - Transmission of data between these functions
  - Dedicated point-to-point links for deterministic transmissions
- Two many links ⇒ shared communication medium
  - Timing constraints have to be insured
  - Specific technologies have been defined, e.g. CAN
  - Timing analysis have been developed
- More and more functions ⇒ huge increase of communication needs
  - Specific technologies are no more sufficient
  - Ethernet technologies are candidate
    - ★ Pros: large bandwidth, mature technology
    - ★ Cons: Ethernet is not deterministic
- Real-time Ethernet widely used for industrial communications, e.g. Profinet
- Candidates for embeded communications: AFDX, Ethernet-AVB, TTEthernet, ...

Agenda

Civilian aircraft communications



Ground Information System (Airlines, Airbus,…)

Avionics systems characteristics

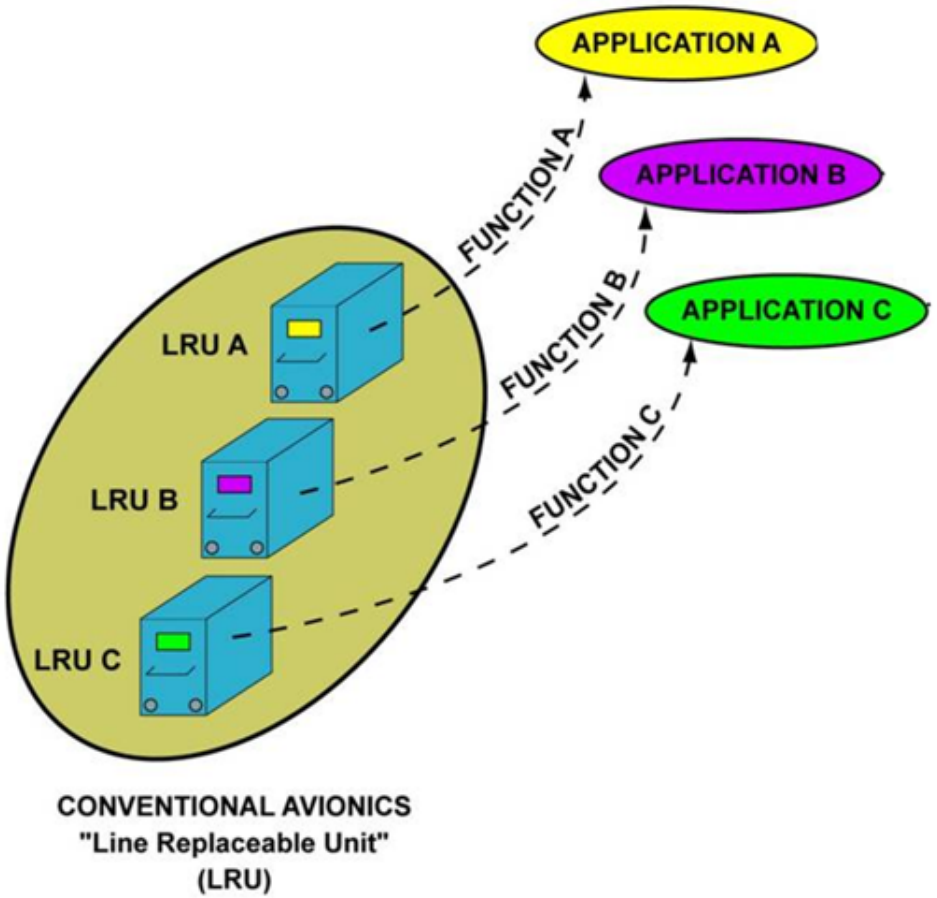
- Avionics: electronic systems installed in an aircraft
  - Calculators and their software, sensors and actuators
  - Communication links between these elements
- Example of avionic systems: autopilot, navigation, flight control, ...
- Constraints on the avionics systems
  - Volume and weight limitations
  - Correct behavior in severe conditions: heat, vibrations, electromagnetic interferences, ...
  - Safety level required, depending on the system (ARP 4754): catastrophic (failure = loss of the aircraft), hazardous, major, minor, no effect
  - Segregation between critical functions
- Aeronautical Radio INCorporated (ARINC): leadership in the development of specifications and standards for avionics equipment
- Airlines Electronic Engineering Committee (AEEC): development and maintenance of specifications and standards (airlines, governments, ARINC)
- ADN: Aircraft Data Network

Avionics systems evolution

- In the 1950's
  - Very simple standalone systems
  - One single calculator for the execution of all the functions
- In the 1960's
  - The beginning of modern avionics
  - Replacement of analog devices by their digital equivalent
- Since then
  - Huge increase in the complexity of avionics
  - More and more digital equipments for existing or new functions
  - New needs inherent to the evolution of civil aviation or better answers to existing problems (e.g. fly-by-wire command system)
- Increase in the number of embedded systems and functions $\Rightarrow$ increase in communication needs
- Integrated Modular Avionics (IMA) in the 1990's: better sharing of execution and communication resources
  - First implementation of IMA for the Boeing 777, based on the ARINC 629 multiplexed data bus
  - IMA in the Airbus A380, based on the ARINC 664 (switched Ethernet)

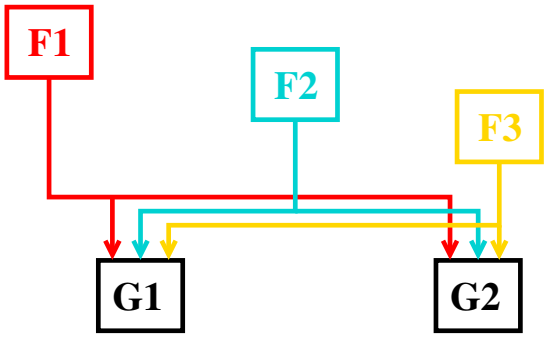
Classical avionics architecture (up to A340)

- Each equipment (LRU) dedicated to a system function (braking, flight computers, ...): sensors, actuators, calculators ...
- natural segregation between the equipments



Data transmission in classical avionics

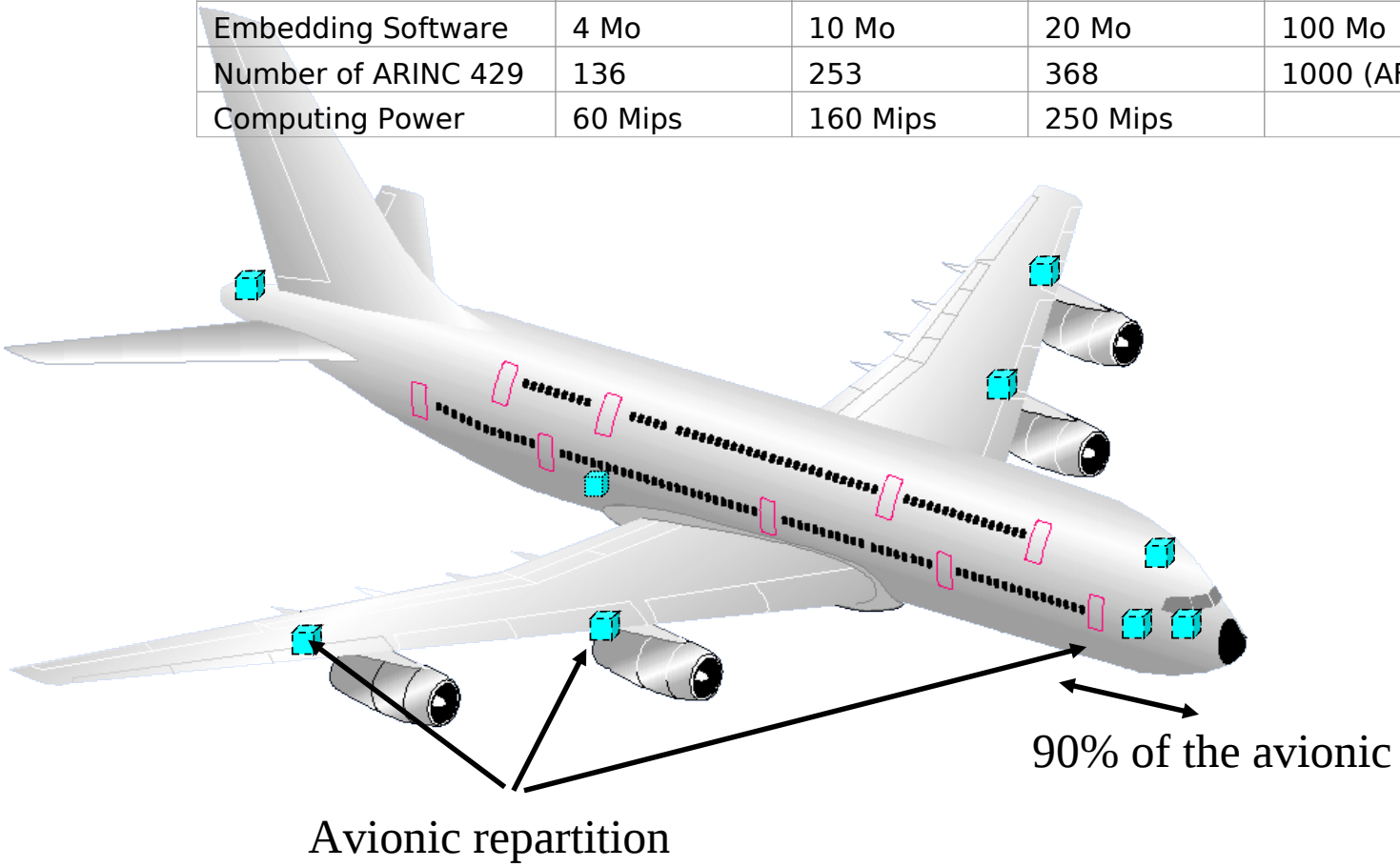
- A network of equipments



- Each data is transmitted from its source to every equipment that needs the data
- One dedicated line for each data
- Repetitive transmission of the data on each line
- Each line: a mono-emitter ARINC 429 data bus
- Each data individually identified by a label
- Low bit rate: 12 Kbits/s to 100 Kbits/s
- At most 20 receivers per line

Classic avionics evolution

	A310 (1983)	A320 (1988)	A340 (1993)	A380 (2005)
Avionic Volume	745 litres	760 litres	830 litres	
Number of equipment	77	102	115	147
Embedding Software	4 Mo	10 Mo	20 Mo	100 Mo
Number of ARINC 429	136	253	368	1000 (AFDX)
Computing Power	60 Mips	160 Mips	250 Mips	



IMA characteristics

- Hardware components
  - LRM (Line Replaceable modules): resources in common cabinets
    - ★ Core modules for the execution of the applications
    - ★ Input/output modules for communications with non-IMA equipments
    - ★ Gateway modules for communications between cabinets
  - LRU (Line Replaceable Unit): existing non-IMA equipments
  - Backplane databus: ARINC 659
- Sharing of execution resources
  - An avionics subsystem: a partition with an assigned time window to execute its application on a shared module
  - Guarantee isolation between subsystems ⇒ robust partitioning concept
    - ★ Spatial isolation: limitation and protection of the address space of each partition, e.g. by a MMU
    - ★ Temporal isolation: static allocation of a time slice for the execution of each partition, based on WCET
  - Communications between partitions via ports (APEX: APplication EXecutive)
    - ★ Sampling ports: only the last value of data is stored
    - ★ Queueing ports: all the values of data are stored
  - Logical channel: multicast link between ports, independent of the communication technology

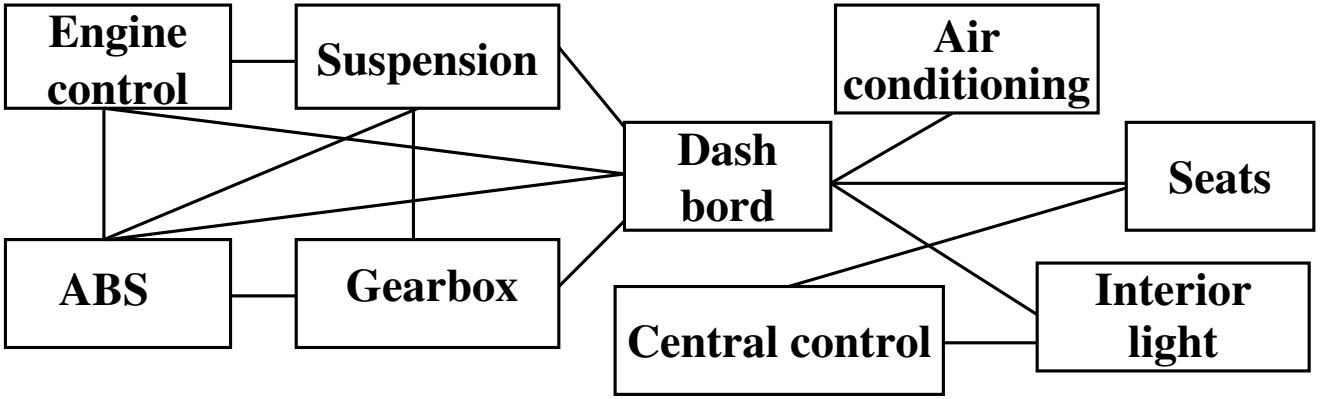
The reasons for Ethernet in avionics

- ARINC 429 is no more sufficient: too many buses are needed
- ARINC 629 is too expensive
- Why Ethernet technology ?
  - High throughput offered to the connected units (100 Mbits/s)
  - High connectivity offered by the network structure
  - A mature industrial standard
  - Low connection cost
- But Ethernet CSMA/CD is not deterministic
  - Potential collision on the physical medium
  - Binary Exponential Back off retransmission algorithm
- The solution adopted for ARINC 664
  - switched Ethernet technology: units directly connected by point-to-point links to Ethernet switches $\Rightarrow$ possible collision domain = single link between two elements
  - Full duplex links $\Rightarrow$ no more collisions
- The problem is shifted to the switch level

Agenda

From point-to-point to multiplexed communications

- Before 1970, mechanical and hydraulic systems for engine control, brake system, gear system, ...
- After 1970, huge increase in the number of electronic systems
- Performance and reliability of hardware components + software technologies ⇒ implementation of complex functions that improve the comfort and safety (Antilock Braking System, Electronic Stability Program, active suspension, GPS, doors, entertainment, ...)
- Early days of automotive electronics: 1 function = 1 Electronic Control Unit (ECU: a microcontroller and sensors and actuators)
- Exchanges of data between functions via dedicated links (e.g. speed estimated by the engine control and used by the suspension control)



From point-to-point to multiplexed communications

- A function is either too complex or too small for a single ECU ⇒ functions distributed over several ECUs or several functions on the same ECU
- Too many information exchanges ⇒ need for a shared communication medium.
  - Reduces the weight of the electronic systems: reduction of 15 Kg with the replacement of the dedicated links by a shared bus for the control of the doors of a BMW
  - Necessary to guarantee that the communication delays can be bounded
- First automotive bus: Controller Area Network (Bosch, 1980s)
- First embedded CAN bus: Mercedes class S (1991)
- Today: a complex architecture with several communication technologies

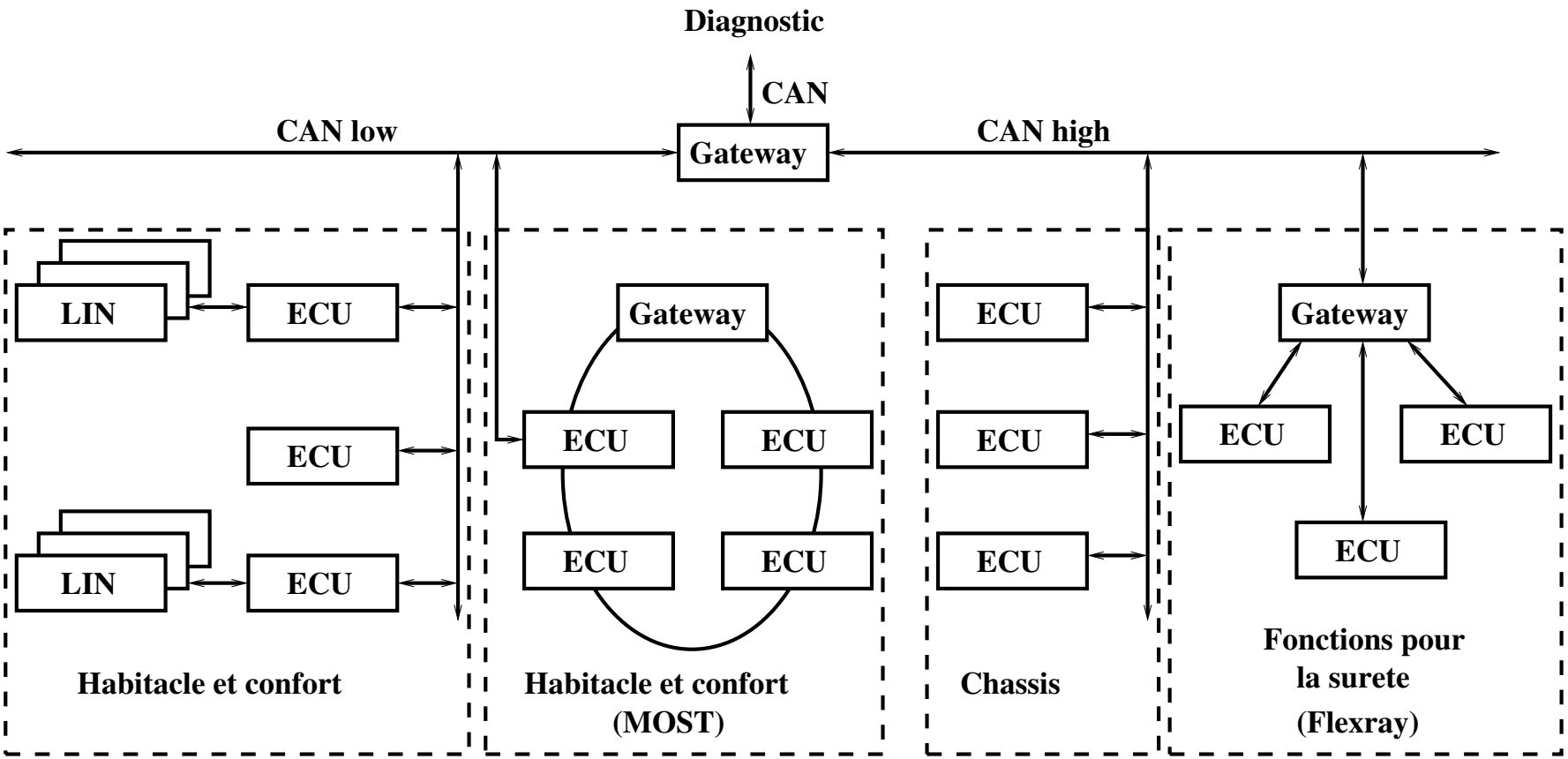
Automotive functional domains

- Different applications with different domains
- Powertrain
  - Control of engine and transmission
  - Complex control laws implemented in microcontrollers with high computing power
  - Stringent time constraints
  - Frequent data exchanges with other car domains
  - Closed-loop control systems ⇒ implentation is moving toward a time-triggered approach
- Chassis
  - Control of the vehicle's according to steering, braking and driving conditions
  - Part of the vehicle active safety systems (Antilock Braking System, Adaptive Cruise Control, ...)
  - Communication requirements similar to those for the powertrain, more emphasis on safety

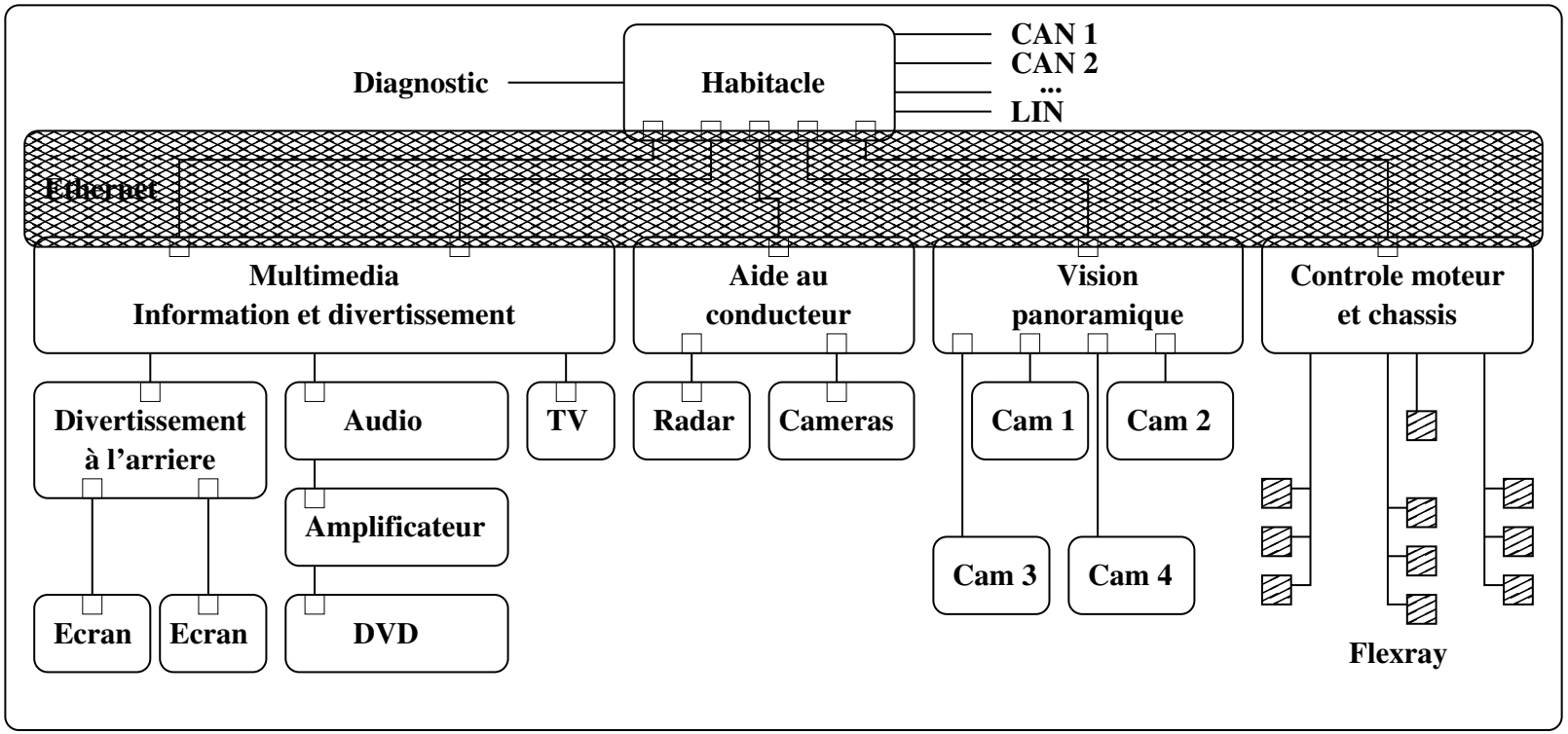
Automotive functional domains

- Body and comfort
  - Software-based systems that control dashboard, wipers, lights, doors, windows, seats, mirrors, air conditioning
  - Many exchanges of small pieces of information
  - Activation of body functions triggered by occupants' solicitations ⇒ event-triggered communication
- Multimedia/infotainment
  - Audio CD, DVD players, rear seat entertainment
  - Hand-free phones, in-car navigation systems, traffic information systems, remote vehicle diagnostic, vehicle tracking, fleet management, ...
  - Large amounts of data to be exchanged within the vehicle and with the external world
  - QoS of multimedia data stream, bandwidth, security

Automotive network architecture in 2006



Vision for future in-car networks



Agenda

Agenda

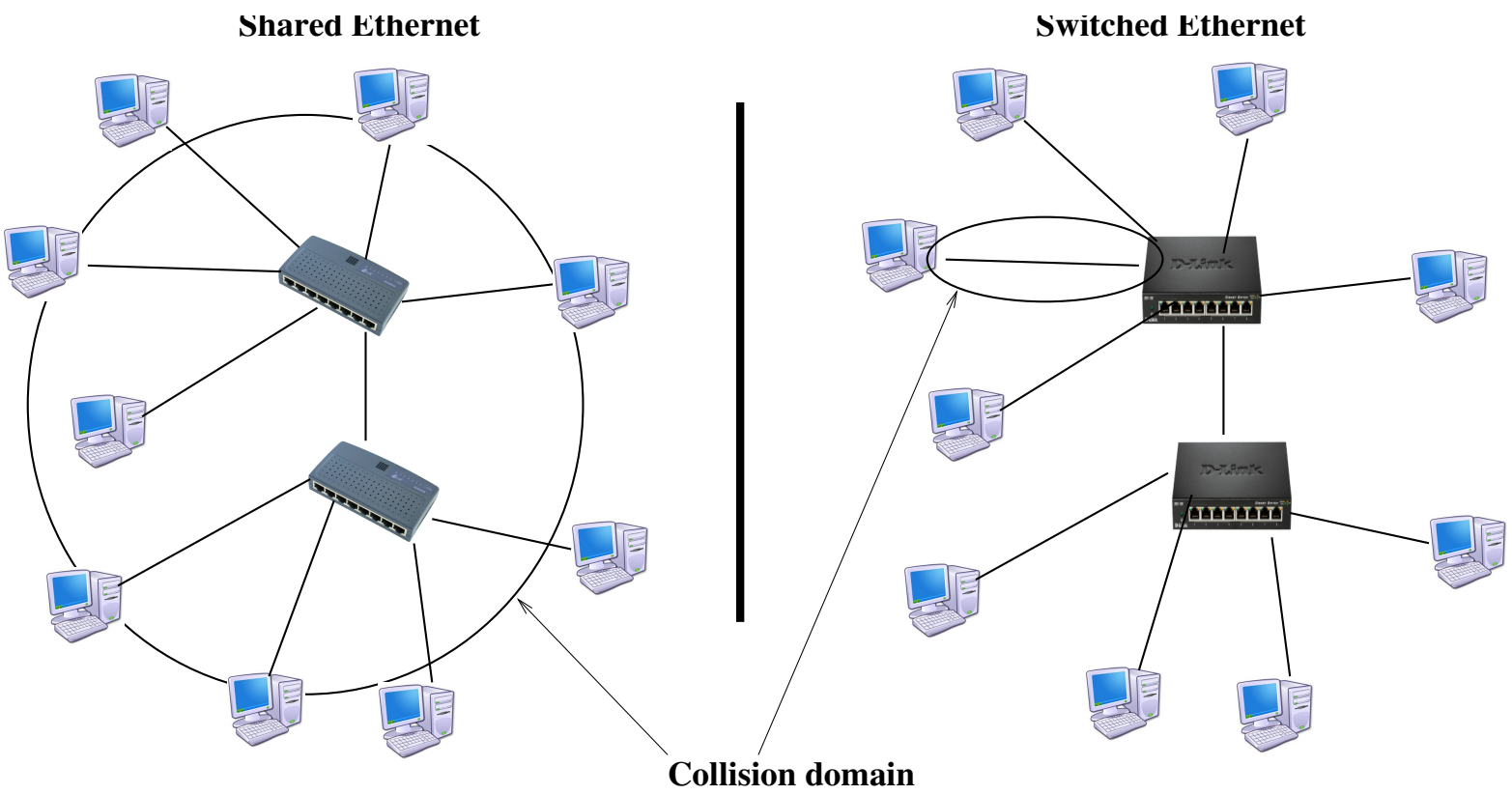
Shared Ethernet

- A set of stations share a single communication channel
- A frame from a station is received by all the other stations (eventually through hubs)
- Collisions may occur ⇒ CSMA/CD
  - CSMA (Carrier Sense Multiple Access): listen to the medium, send only when medium is detected idle
  - CD (Collision Detection): if a collision is detected during transmission
    - ★ Transmission is stopped
    - ★ Station waits a random delay (exponential backoff)
    - ★ Station tries to transmit again
- At least 64 bytes in order to detect a collision
- Number of collisions: major impact on Ethernet performance
  - Performance decreases very quickly when number of collisions increases
  - Collision domain: the whole network

Switched Ethernet

- Division of the network with separate collision domains ⇒ more than one frame being transmitted at a given time without collision
- A switch:
  - A set of input and output ports with buffers for pending frames
  - A switch fabric which "moves" each frame from its input port to its destination output ports
  - A service discipline for the scheduling of frames in output ports
- Two modes for a switch
  - Store and forward: reception and verification of the whole frame before it is forwarded to its output port(s)
  - Cut through: frame is transfered as soon as destination is decoded
    - ★ Smaller latency
    - ★ Invalid frame might be transferred

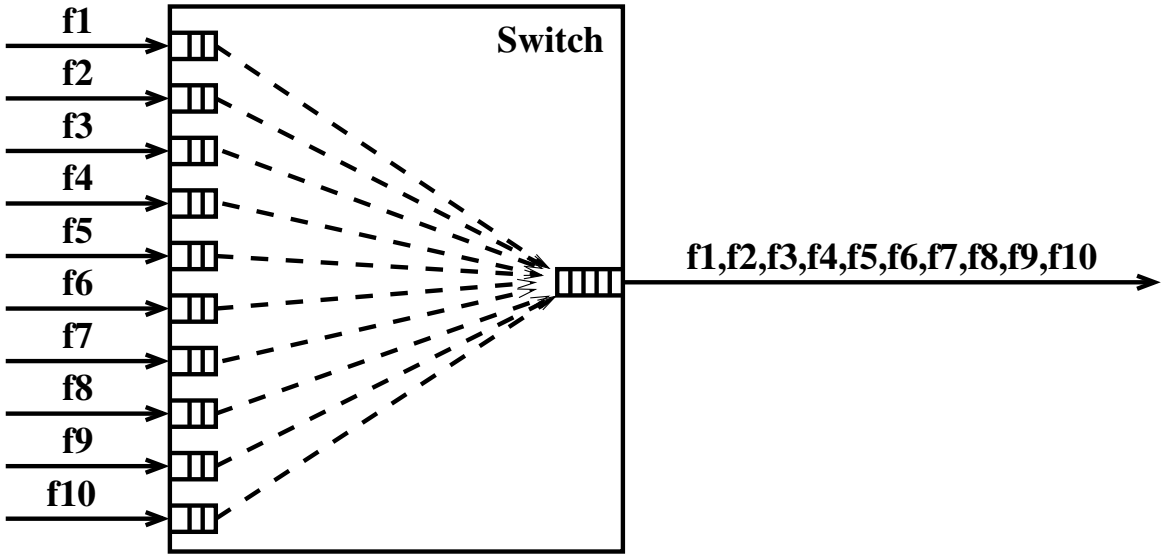
Shared Ethernet Vs switched Ethernet



- Full duplex links ⇒ no more collisions

Full duplex switched Ethernet is not deterministic

- Temporary congestion on an output port
  - Increase of the waiting delay of frames in the Tx buffer
  - Frame loss by overflow of the Tx buffer
- An illustrative example

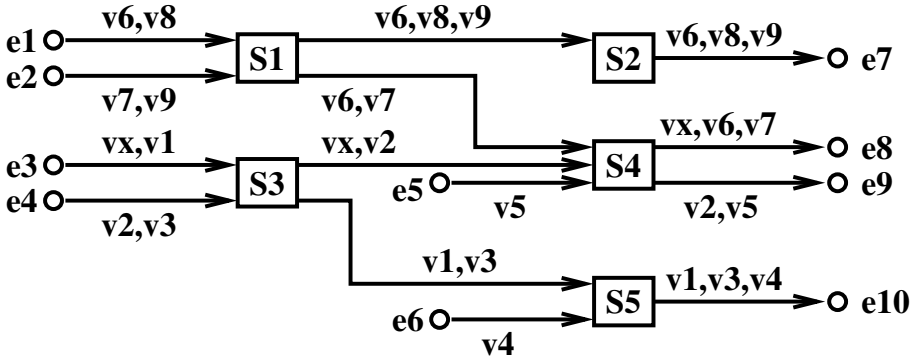


  - Five frames at the same time ⇒ one frame waits until the transmission of the four other ones
  - More than five frames at the same time ⇒ at least one frame is lost
- Addition of dedicated mechanisms to classical full duplex switched Ethernet in order to guarantee the determinism of transmissions

Agenda

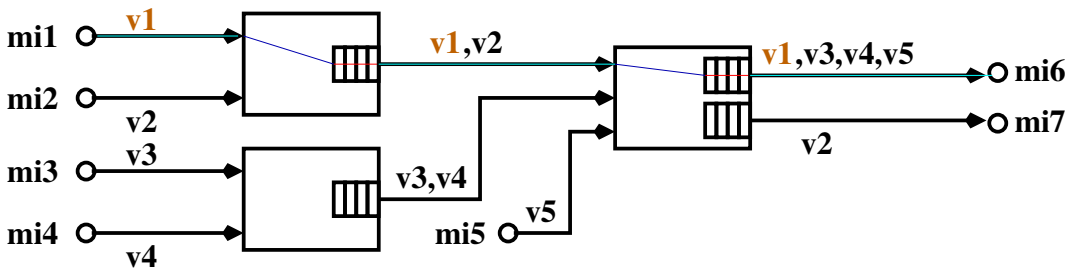
The AFDX key characteristics

- Static full duplex Switched Ethernet: no CSMA/CD, no spanning tree, static 802.1D tables
- Two FIFO buffers per output port (high and low priorities)
- Traffic characterization based on the Virtual Link (VL) concept
  - Static definition of the flows which enter the network
  - Multicast path with deterministic routing
  - Mono transmitter assumption
  - Sporadic flows (*BAG*: Bandwidth Allocation Gap)
    - ★ Discrete values: 1, 2, 4, 8, 16, 32, 64, 128 ms
    - ★ Traffic shaping on each emitting end system
  - Minimum (S_{min}) and maximum (S_{max}) frame lengths for each VL
  - *BAG* and S_{max} define the maximum bandwidth allocated to a VL
    - ★ Example: $BAG = 16$ ms and $S_{max} = 128$ bytes $\Rightarrow$ 64000 bits/s
  - Scheduling of VLs on end systems $\Rightarrow$ (bounded) jitter



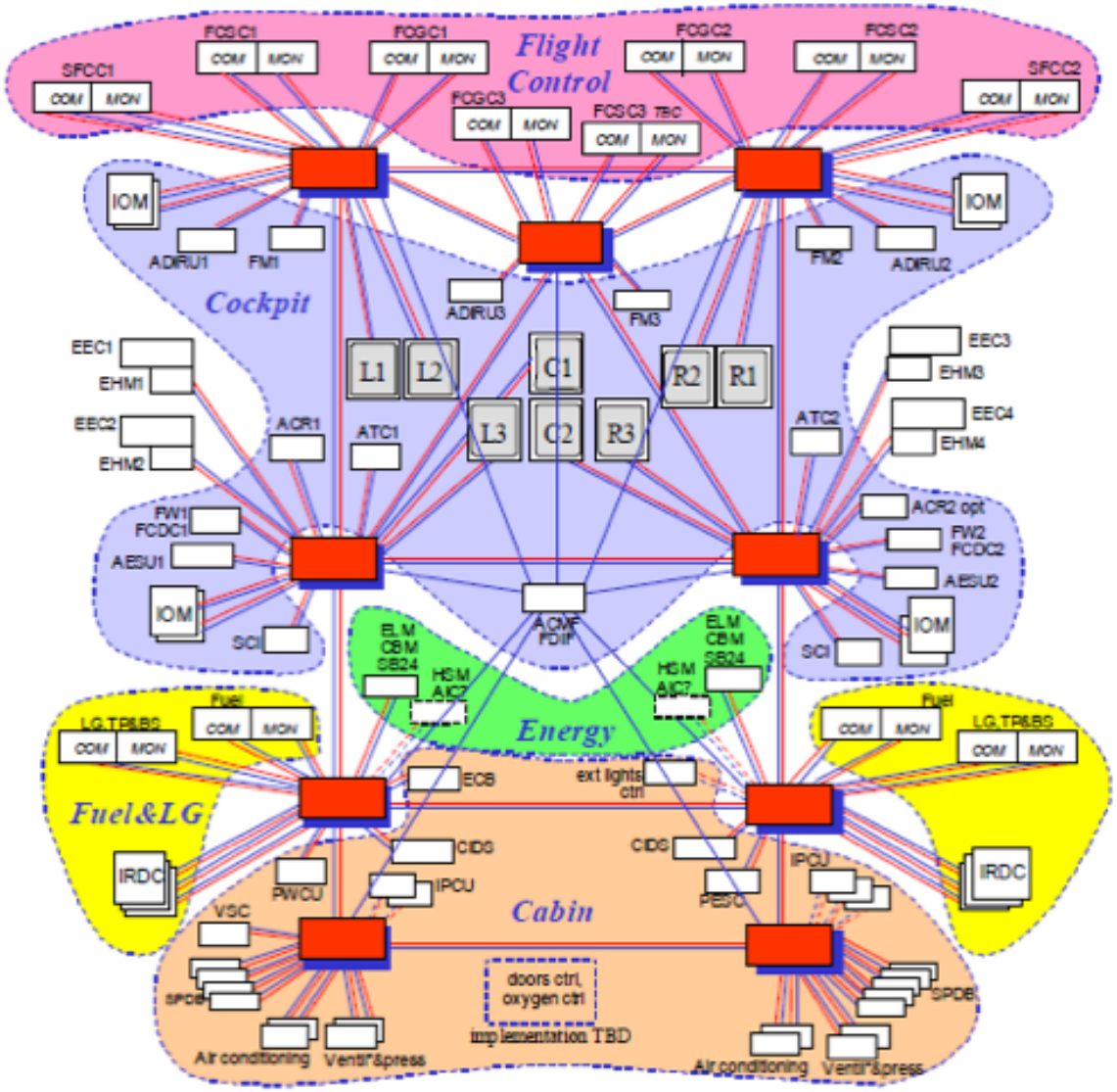
Worst-case end-to-end delay on an AFDX network

- Three parts in the delay
  - Transmission times on links (known): link bandwidth, frame size
  - Switching latency (known)
  - Waiting time in FIFO queues (unknown)



- Upper bound on the waiting time in FIFO queues
  - Network calculus
    - ★ based on modelling with arrival curves and service curves
    - ★ Certification of AFDX on board A380 and A350
  - Trajectory approach
    - ★ Maximum workload of interfering frames on the considered path

An industrial AFDX configuration



An industrial AFDX configuration

- 2 redundant networks, 9 switches per network, 24 ports per switch, 123 end systems
- 984 VLs, 6412 paths
- 100 Mbits data rate
- BAG and maximum frame length

Bag (ms)	Number of VL
2	20
4	40
8	78
16	142
32	229
64	220
128	255

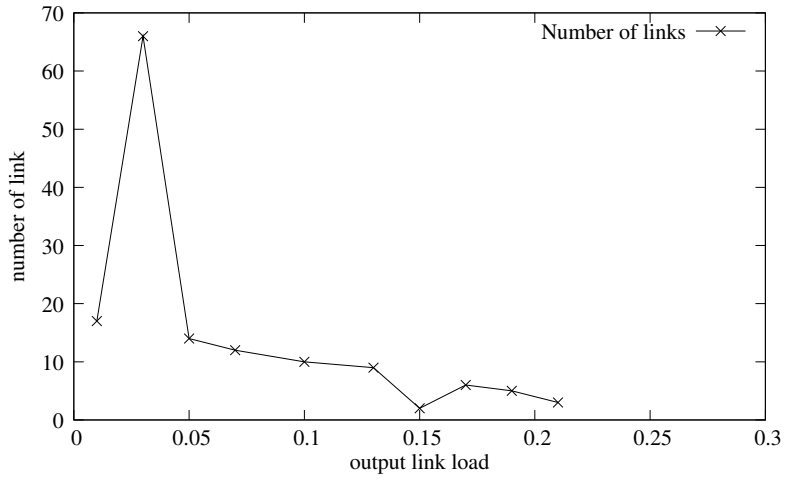
S_{max} (bytes)	Number of VL
0-150	561
151-300	202
301-600	114
601-900	57
901-1200	12
1201-1500	35
> 1500	3

An industrial AFDX configuration

- Length of VL paths

Number of crossed switches	Number of paths
1	1797
2	2787
3	1537
4	291

- Load of the links



Need for QoS support in embedded networks

- in the context of avionics
  - AFDX is purely event-triggered with no global synchronization
  - AFDX only support two priority levels which are not significantly used
  - AFDX is lightly loaded ⇒ share the network with non avionic flows ⇒ non avionic flows should not disturb avionic ones
- In the context of automotive
  - Hetergeneous flows with heterogeneous constraints ⇒ needs for differentiated service disciplines

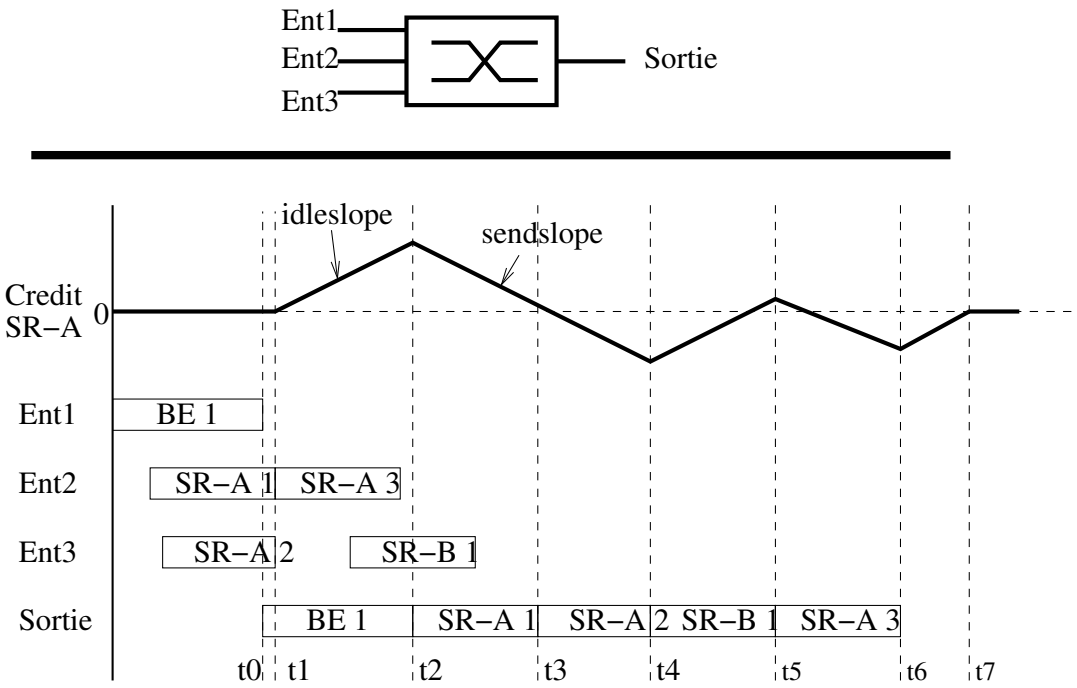
Agenda

Ethernet AVB (IEEE Audio Video Bridging)

- Set of standards that provide the specification for time-synchronized low latency streaming services through IEEE 802.1 networks
  - IEEE 802.1AS: timing and synchronization for time-sensitive applications
    - ★ Variation of IEEE 1588
    - ★ Precise time synchronization with an accuracy better than 1 μs
  - IEEE 802.1Qat: stream Reservation protocol
    - ★ Resevation of resources along the path between the talker and the listener
  - IEEE 802.1Qav: forwarding and queueing enhancements for time-sensitive streams
    - ★ Different classes for time-critical and non-time-critical traffic
    - ★ Two Stream Reservation classes for time-critical traffic: SR-A and SR-B
    - ★ Credit-Based Shaper (CBS) to prevent traffic bursts
    - ★ SR-A and SR-B can start transmission only if their credit is not negative

Ethernet-AVB Credit-Based Shaper

- Credit decreases at sendslope rate during the transmission of frames of the associated class
- Credit increases at idleslope rate when frames of the associated class are waiting or no frame of the associated class is waiting, but credit is negative
- Credit reset to 0 when credit greater than 0 and no frame of the associated class is waiting

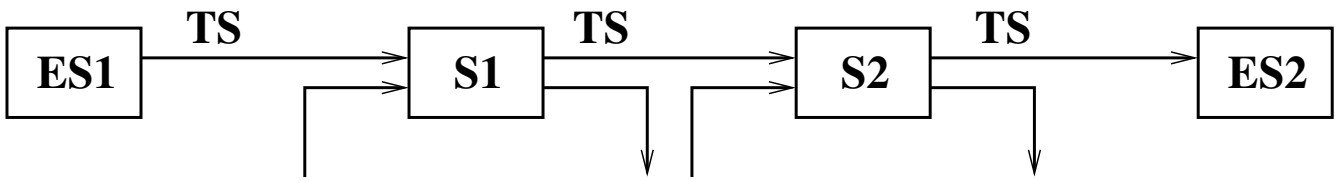


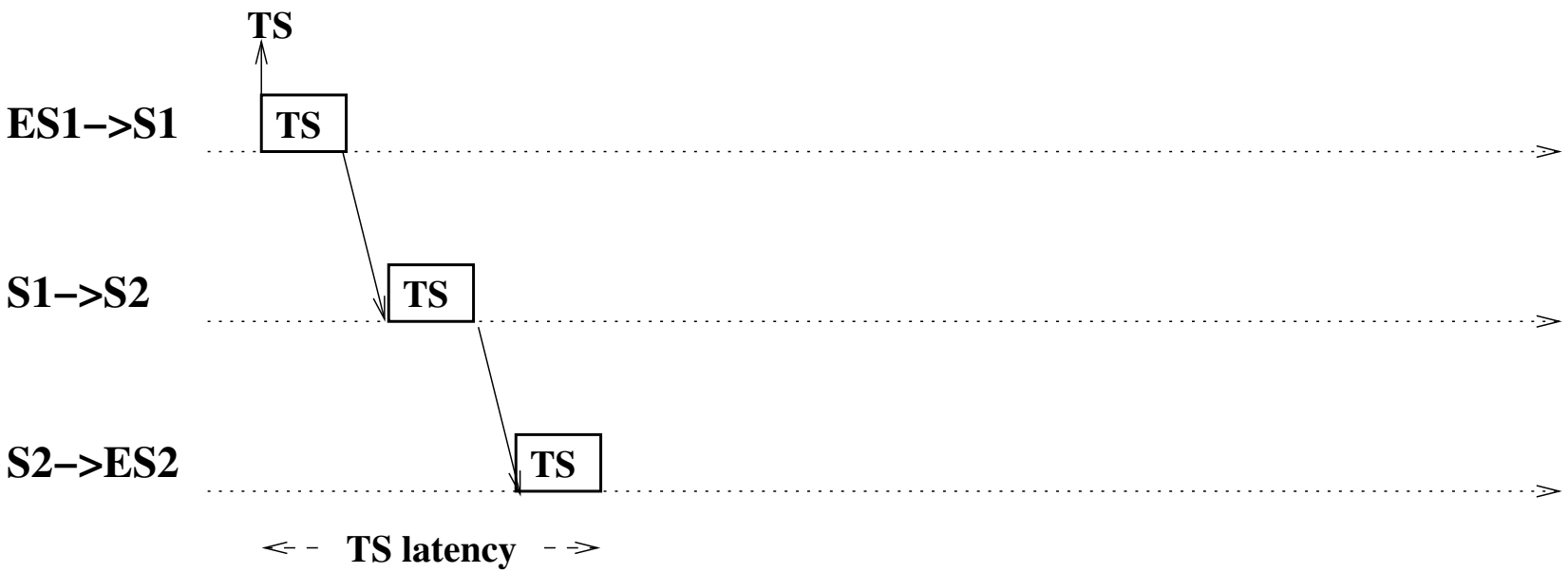
Scheduled traffic over AVB

- Ongoing project within the Time Sensitive Networking Task Group
  - Procedures to suspend the transmission of a non-time critical frame and allows for one or multiple critical frames to be transmitted
  - Policies to support scheduled traffic
- Scheduled traffic is a traffic class that
  - Requires to schedule frame transmission based on time
  - Is time-sensitive and requires a bounded latency
    - Ex: delay-sensitive command and control traffic which requires deterministic and very small delays

Latency assessments for time-sensitive traffic

- Without interfering traffic
  - Maximum latency of a time-sensitive frame mainly depends on
    - Transmission time of the time-sensitive frame on links (computed from the frame size)

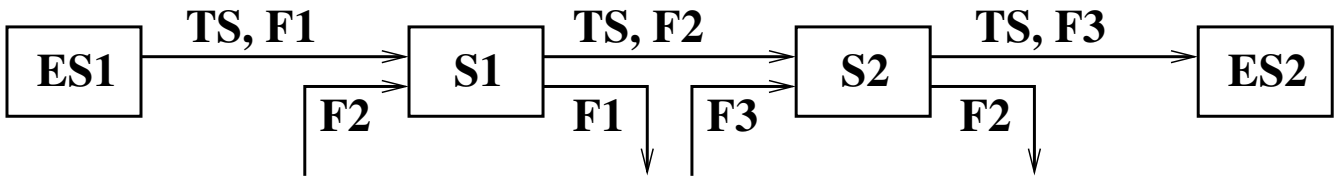


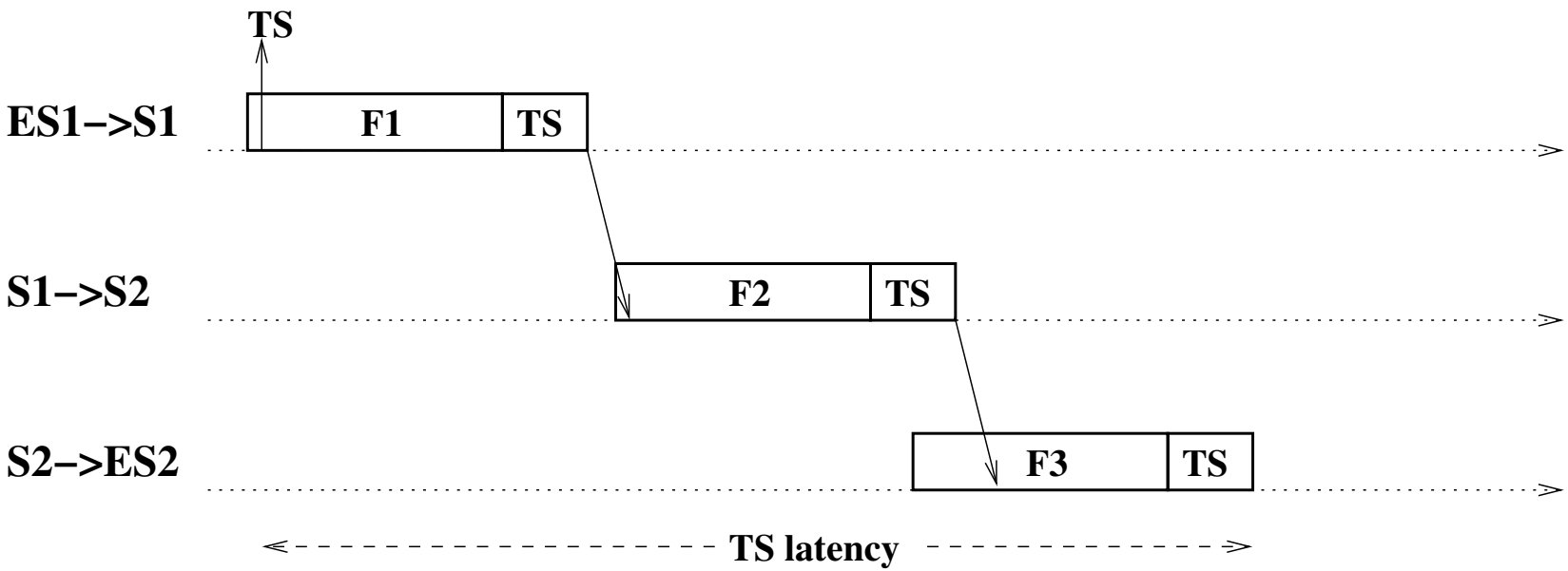


- Good news since time-sensitive flows typically feature small frames

Latency assessments for time-sensitive traffic

- With interfering traffic
  - Maximum latency of a time-sensitive frame also depends on
    - ★ Maximum size of the interfering frames, due to non-preemptive scheduling at the egress port of each network device

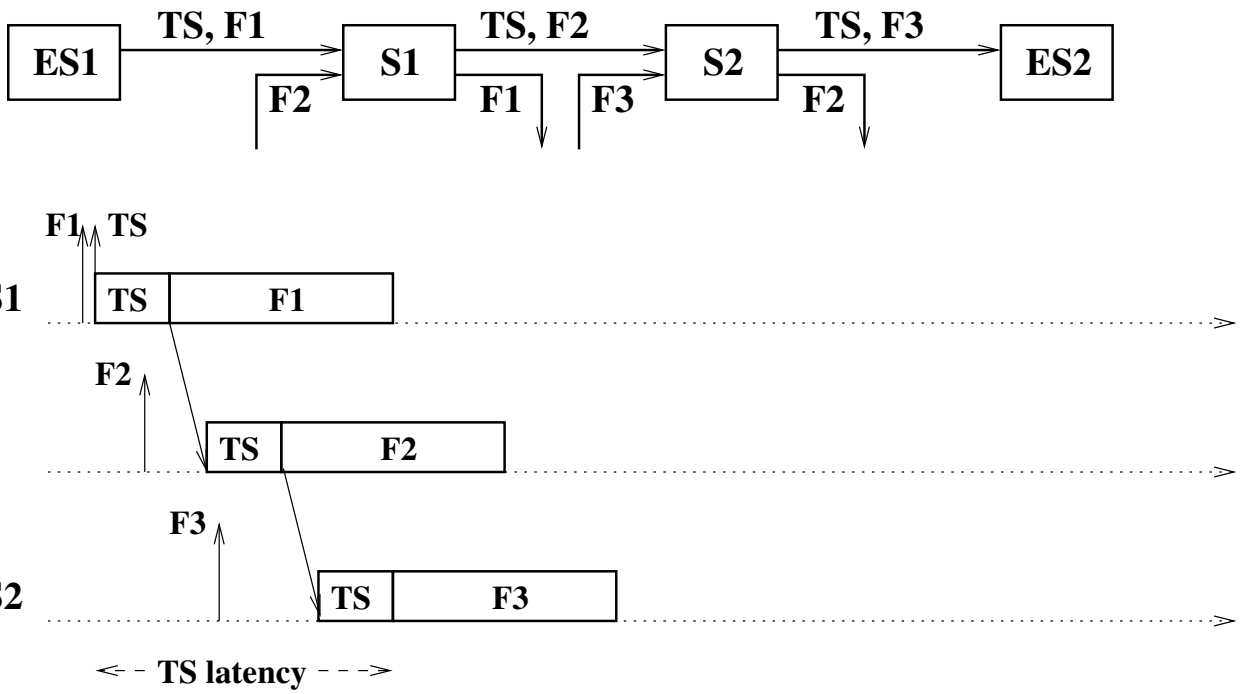




- Bad news since non time-sensitive frames might be large

Time-Aware Shapers (TAS)

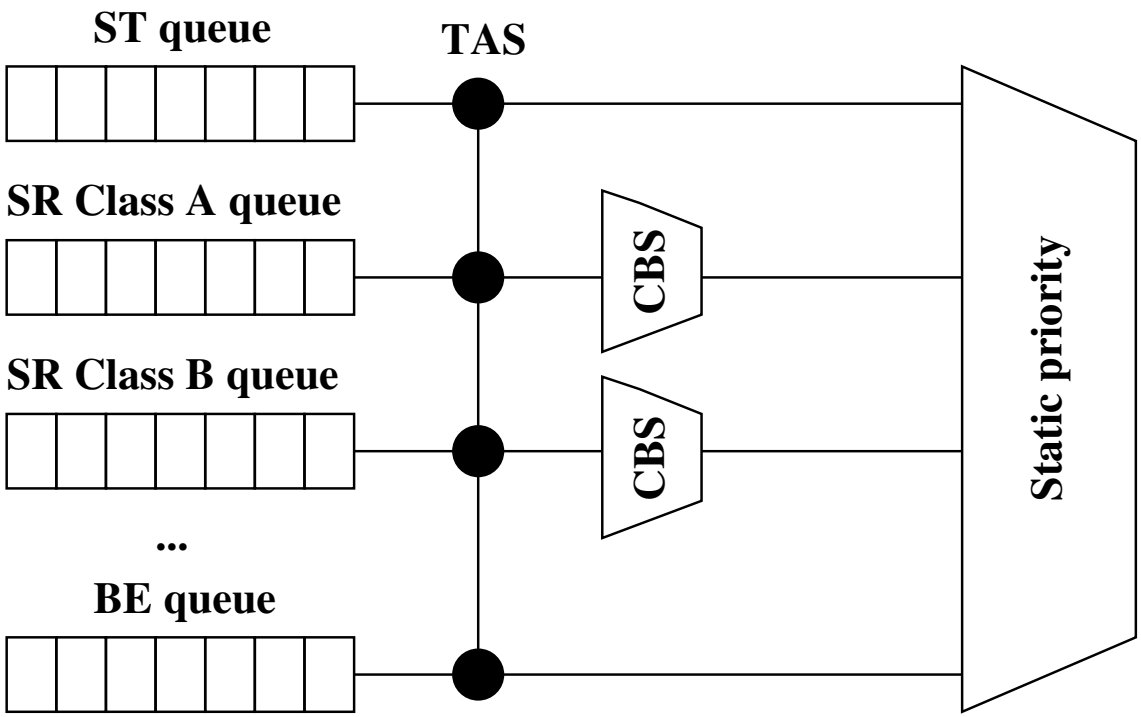
- Time-Aware Shapers have been defined to deal with interfering frames
- Based on the knowledge of the next arrival time for scheduled time-sensitive frames
  - Synchronization is provided by IEEE 802.1AS
  - Time-sensitive traffic typically follows a regular pattern
- TAS blocks any lower priority frame that would interfere with an upcoming time-sensitive frame
  - Time-sensitive traffic mapped onto SR Class A



An alternative solution to support scheduled traffic

- TAS alone are not enough to support scheduled traffic
  - Only two real-time traffic classes (A and B) ⇒ time-sensitive control traffic often has to share Class A with other real-time traffic
  - It might interfere significantly with time-sensitive frames
  - Increased effect due to credit-based shaping
- AVB_ST (Scheduled Traffic) approach
  - A new ST on top of AVB SR class A and B
  - ST traffic is handled in a separate queue
  - ST traffic does not undergo credit-based shaping
  - Offline scheduling of ST frames (a priori known periods and frame sizes)

AVB_ST


ST queue
TAS
SR Class A queue
CBS
SR Class B queue
CBS
...
BE queue
Static priority

Agenda

TTEthernet very short summary

- Time is shared between time-triggered and event-triggered traffic
- Three types of traffic
  - TT flows: scheduled flows, possible thanks to the clock synchronization
  - Rate Constrained flows (RC): very similar to AFDX VLs
  - Best Effort flows (BE)
- Priority order: TT, then RC, then BE
- RC and BE frames might interfere with TT frames
- Three solutions when a RC or BE frame is ready before a TT slot
  - RC or BE frame starts transmission and might delay TT frame
  - RC or BE frame starts transmission and is preempted if TT frame is ready
  - RC or BE frame starts transmission only if there is enough time before TT slot

Agenda

Some open issues

- Comparative analysis of candidate switched Ethernet solutions
- Design and analysis of heterogeneous network architectures
  - Interconnection of fieldbuses through an Ethernet backbone
- Probabilistic analysis
  - Up to now, certification process based on guaranteed delay upper bound
  - Worst-case delay scenario: very rare event
  - Leads to over-dimensionning of the network
  - Do we really need a guaranteed upper bound?
  - Is it ok if the probability to exceed the deadline is 10^{12}?
  - Avionics functions designed to behave correctly even if some (rare) frames are lost $\Rightarrow$ computing a probabilistic upper bound makes sense

Thank you for your attention