802.3 Call For Interest (CFI)
Frame Preemption

MAC Services in support of
Frame Preemption

IEEE 802.3 Plenary, March 11~16, 2012

CFI Requestor: Michael Grimwood
CFI Moderator: Michael Johas Teener, 802.1 AVB TG Chair

Call For Interest (CFI) Objectives

- To gauge the interests in forming a study group that supports Frame Preemption.
- This meeting will NOT
  - Explore the complete problem
  - Debate nor choose a solution
  - Create PAR and the Five Criteria
- The above are parts of Study Group work.

CFI – "Preemption" Agenda

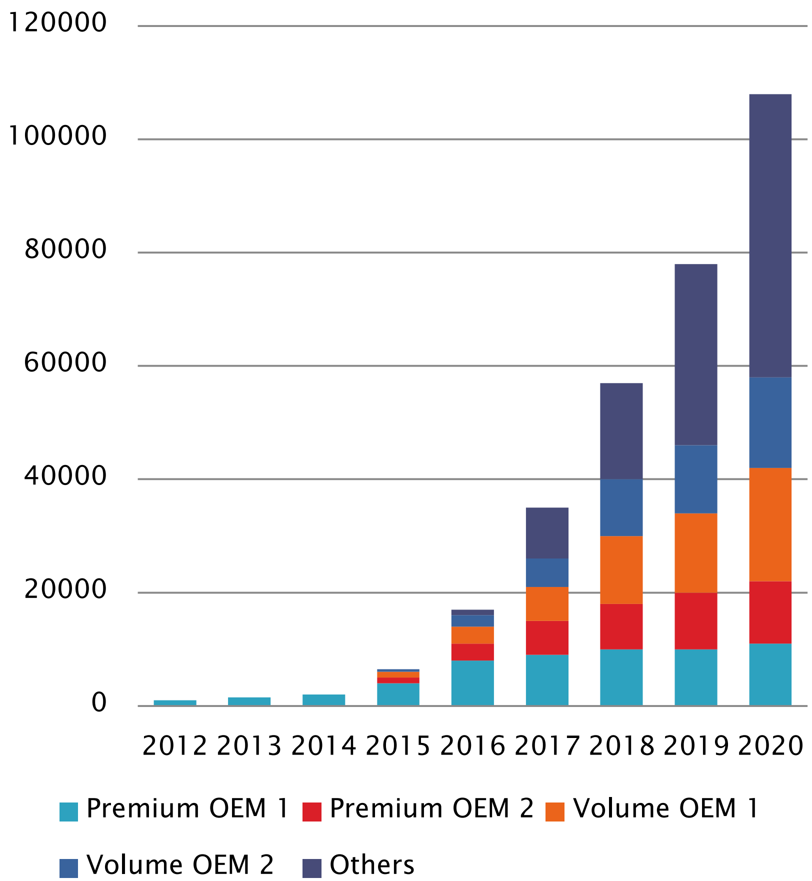
- The Need
  - Automotive and Industrial Needs – User Perspective.
  - Tech/Standards Needs
- Why Now
- 802.1 & 802.3 Preemption Coordination.
- Vote to gauge interests

Supporters (37 Individuals from 25 Companies)

- Hugh Barrass, Cisco
- Robert Boatright, Harman
- Christian Boiger, Deggendorf Univ.
- Joseph Chou, Realtek
- Rodney Cummings, National Instrument
- Nick Difiore, Xilinx
- Ali Dixon, XMOS
- Norm Finn, Cisco
- Masahiro Fukuda, Renesas
- Geoff Garner, Self
- Franz-Josef Goetz, Siemens
- Mike Grimwood, Broadcom,
- Craig Gunther, Harman
- Mark Gustlin, Xilinx
- Thomas Hogenmüller, Bosch
- Mitsuru Iwaoka, Yokogawa
- Tony Jeffree, Self
- Markus Jochim, General Motors
- Michael Johas Teener, Broadcom
- Max Kicherer, BMW
- Dongok Kim, Hyundai Kia Motors
- Yong Kim, Broadcom
- Oliver Kleinberg, Hirschman
- Ludwig Leurs, Bosch Rexroth
- Kirsten Matheus, BMW
- Anatoly Moldovansky, Rockwell
- Dave Olsen, Harman
- Kevin Stanton, Intel
- Wilfried Steiner, TT Tech
- Mario Traeber, Lantiq
- Albert Tretter, Siemens
- Jose Villanueva, Renault,
- Karl Weber, ZHAW (Zurich Univ.)
- Hans Weibel, ZHAW (Zurich Univ.)
- Ludwig Winkel, Siemens
- Helge Zinner, Continental
- Paul Zoratti, Xilinx

Market Served – Broad Market Potential

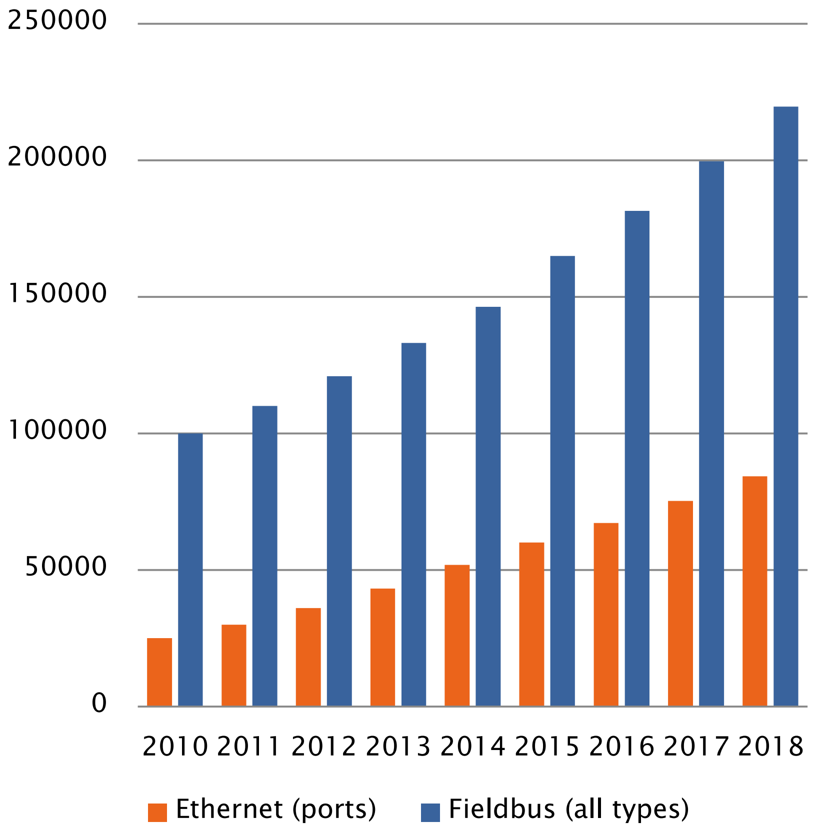
Automotive Ethernet (K ports)



Source: Strategy Analytics, 11/2011 Munich

Industrial Bus

(K Ports)

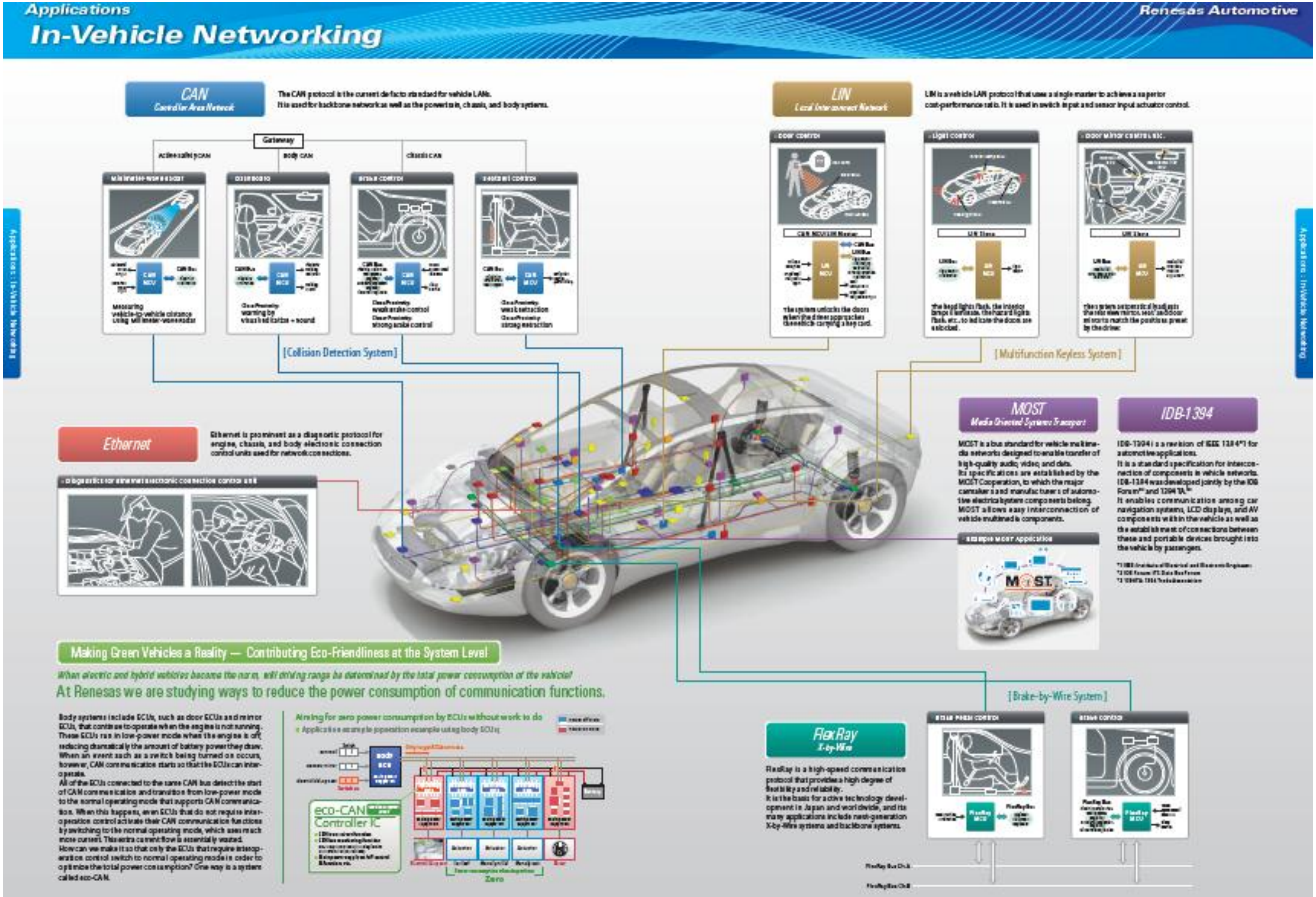


Source: Contributions from Hirschmann, Siemens and Broadcom

The Need

Automotive

In-Vehicle Networking Introduction



http://resource.renesas.com/lib/eng/ebook/01/pageview/pageview.html#page_num=47

Ethernet Backbone Needs – Introduction

- **There are multitude of in-vehicle bus systems.**
  - **LIN**: Multi-drop "UART-like" with synchronization, <=19.2Kbps
  - **CAN**: Widely available CSMA/CR bus system, <=1 Mbps
  - **FlexRay**: Time-triggered TDMA Bus and Star system, <=10 Mbps
  - **MOST**: Synchronous TDMA Ring, 25, 50, and 150 Mbps, Shared.
  - **Ethernet**: Switched Full-Duplex (modern) Star system, 100 Mbps +, switched.

- **Vehicle Communication Domains**
  - **Powertrain**: Engine, transmission
  - **Chassis**: Steering, ABS, Tire pressure
  - **Body**: Doors, Lamps, Seats, A/C
  - **Safety**: Air-bags, Sensors, Actuators, Occupant Safety System
  - **Infotainment** & **Driver Assist**: Navigation, Telematics, TV/Radio/CD/DVD, RSE, Cameras

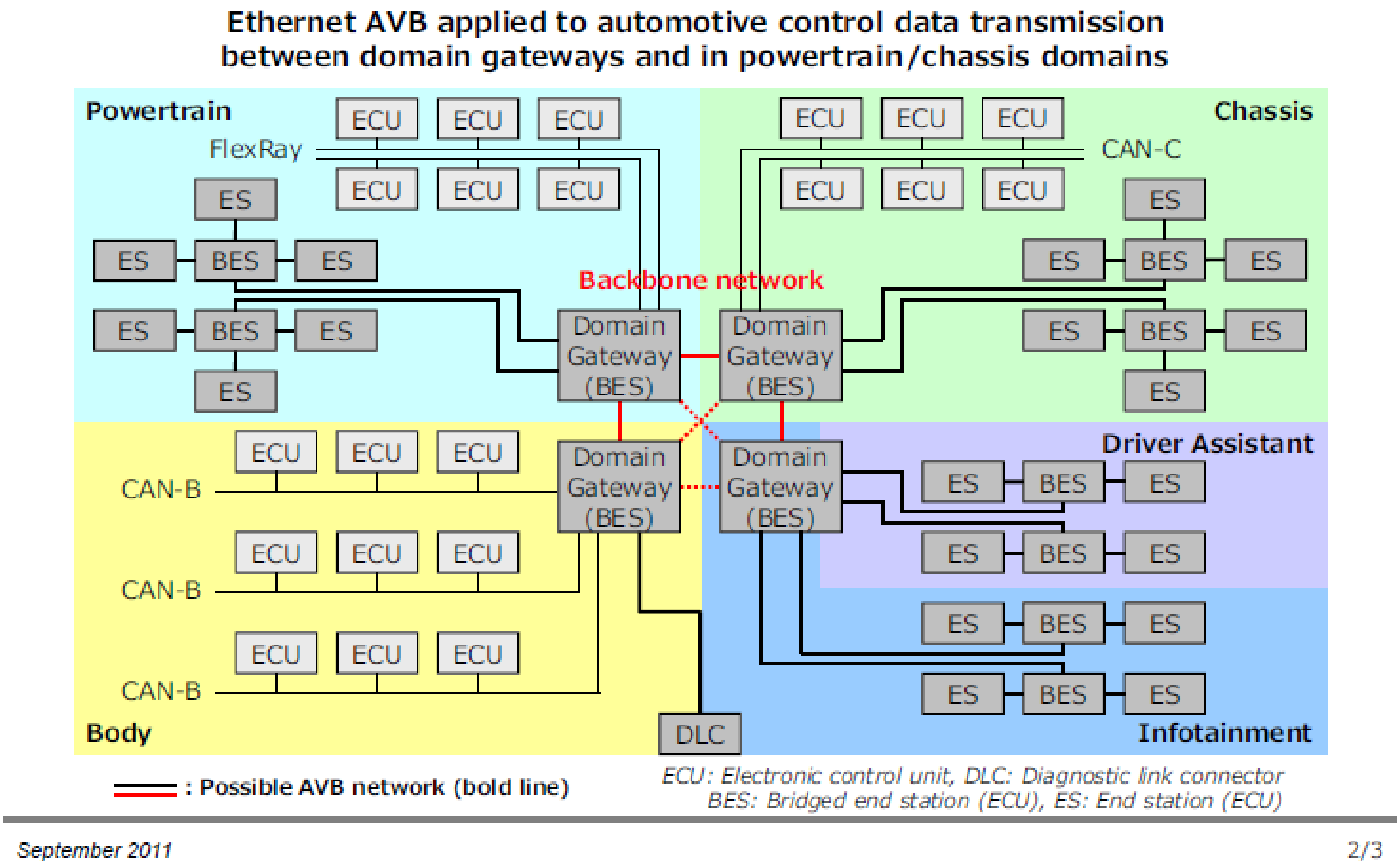
- **Trends in in-vehicle communication needs**
  - Infotainment and Driver Assist drives higher bandwidth (graphic panels, cameras, WLAN, BT)
  - Communication convergence: More buses are connected through use of [bus] gateways
  - Information sourced in one zone and used in many zones (e.g. camera, sensor data)
  - Vehicular Diagnostics interface is standardized to be over Ethernet and IP and already adopted.

- **There is a need for a converged in-vehicle backbone network**

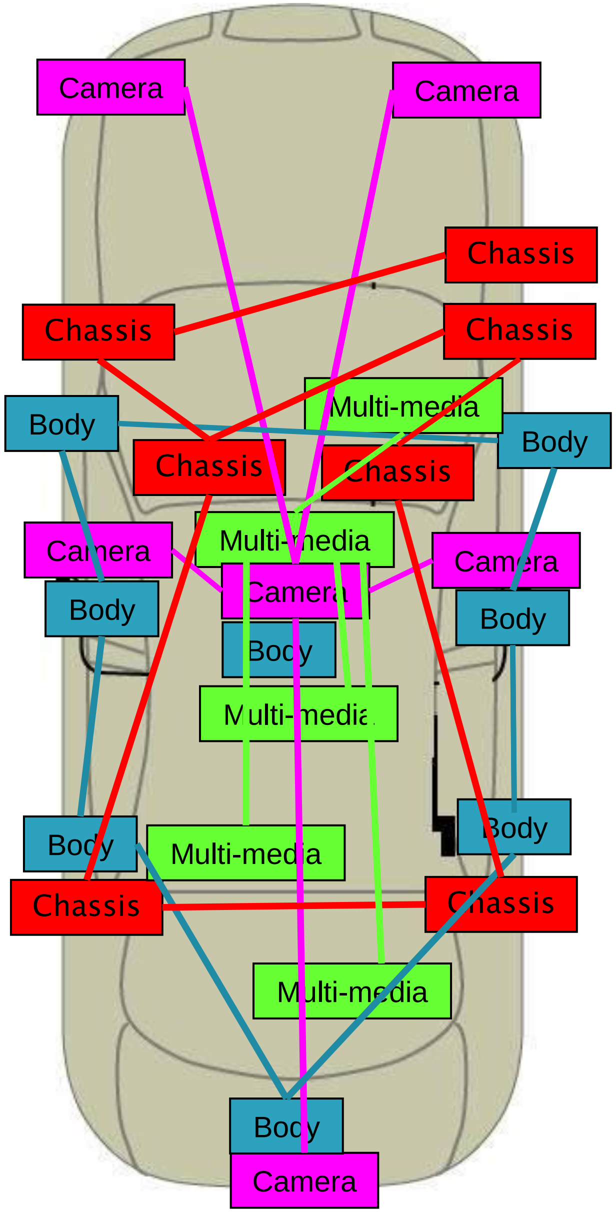
Use Case – Automotive 1

from new-avb-nakamura-automotive-backbone-requirements-0907-v02.pdf

Example next-generation automotive network architecture



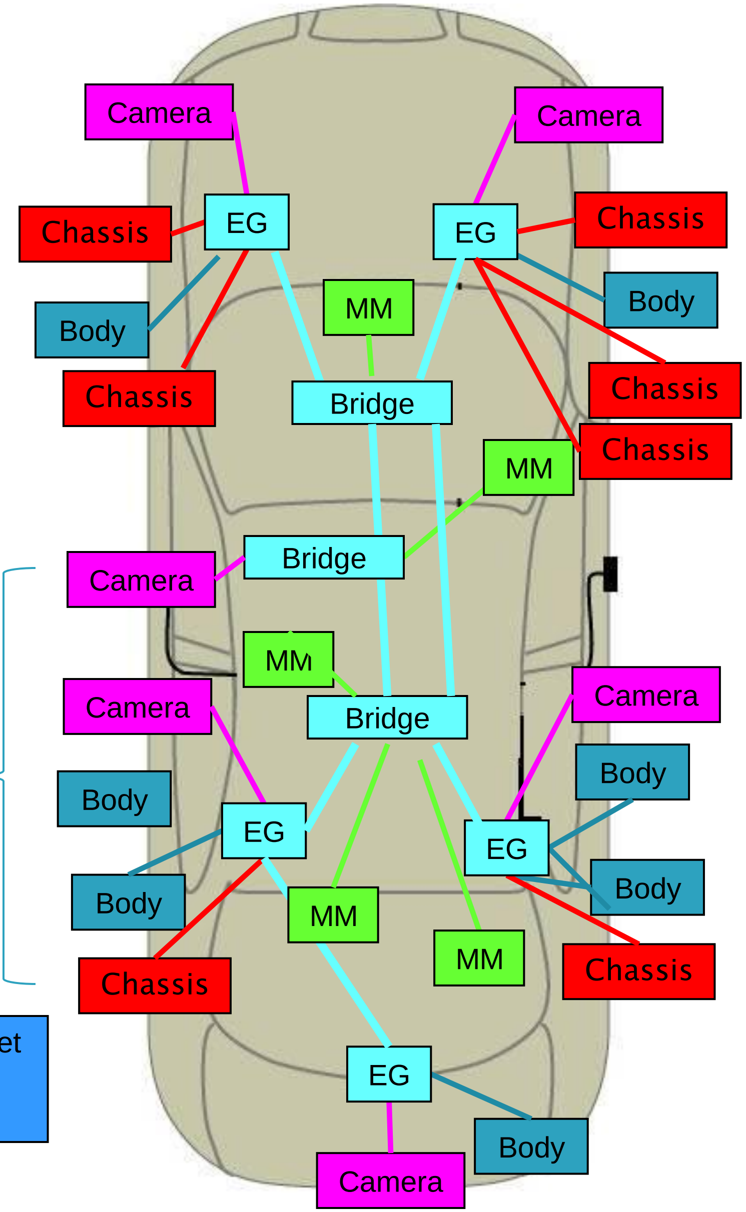
Use Case – Automotive 1 Spatial



40~60 uC in a mid-range car
Over 100 uC in a high-end car.

Each dedicated domain network with separate wiring

Ethernet Backbone Network with bus gateways



MM: Multi-media

EG: Ethernet [Bus] Gateway

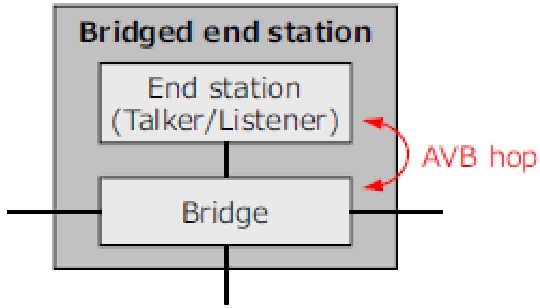
Use Case – Automotive 2

from new-avb-nakamura-automotive-backbone-requirements-0907-v02.pdf

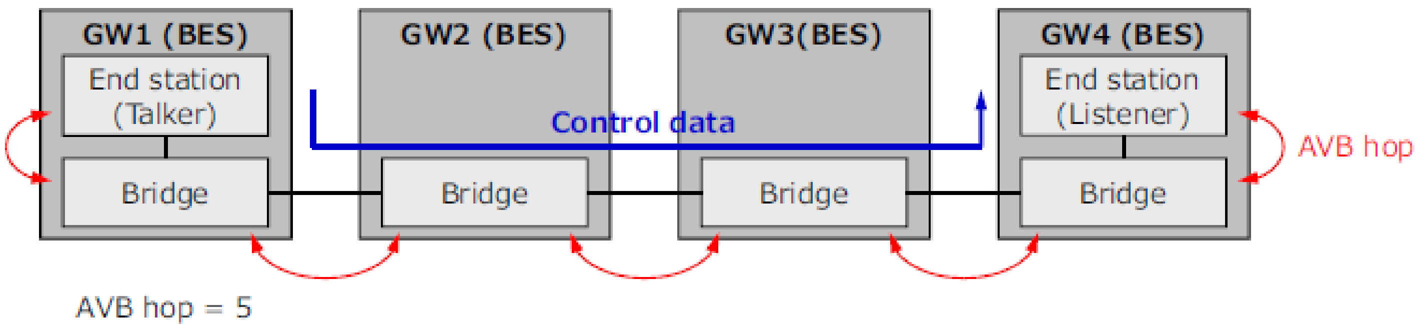
Consideration of AVB hop counts

Assumption for bridged end stations

- A bridged end station consists of a bridge and a end station (talker/listener).
- One AVB hop needs to be counted inside of bridged end stations.



Example control data transmission in a backbone network



September 2011 3/3

Use Case – Automotive Requirements

from new-avb-nakamura-automotive-backbone-requirements-0907-v02.pdf

QoS requirements for automotive control data class

Performance requirements for automotive control data class

- Maximum latency: 100 us / 5 AVB hops
  - Guaranteed latency
  - Topology independent
  - Automotive control data class to have higher priority than SR classes
    - Maximum 2 priority classes (e.g. Control data class and SR class A)

Preconditions for performance requirements

- Network type: Dedicated network in a vehicle
- Network attributes
  - Maximum AVB hop count: 7
  - Maximum number of nodes (bridged end station & end stations): 32
  - Maximum cable length: 24 m
  - Maximum end-to-end cable length: 30 m
- Automotive control data class attributes
  - Maximum data size (payload size): 128 bytes @FE ~ 256 bytes @GE
  - Maximum number of simultaneous transmission: 8 @FE ~ 32 @GE
  - Transmission period: 500 us
- Payload size for other/lower traffic classes: 256 bytes @FE ~ 1500 bytes @GE

These are our best estimates derived from multiple assumptions of the current and future automotive applications.

September 2011 1/3

The Need

Industrial

Industrial Automation

Ethernet captures more and more Industrial Applications

Traditional Markets

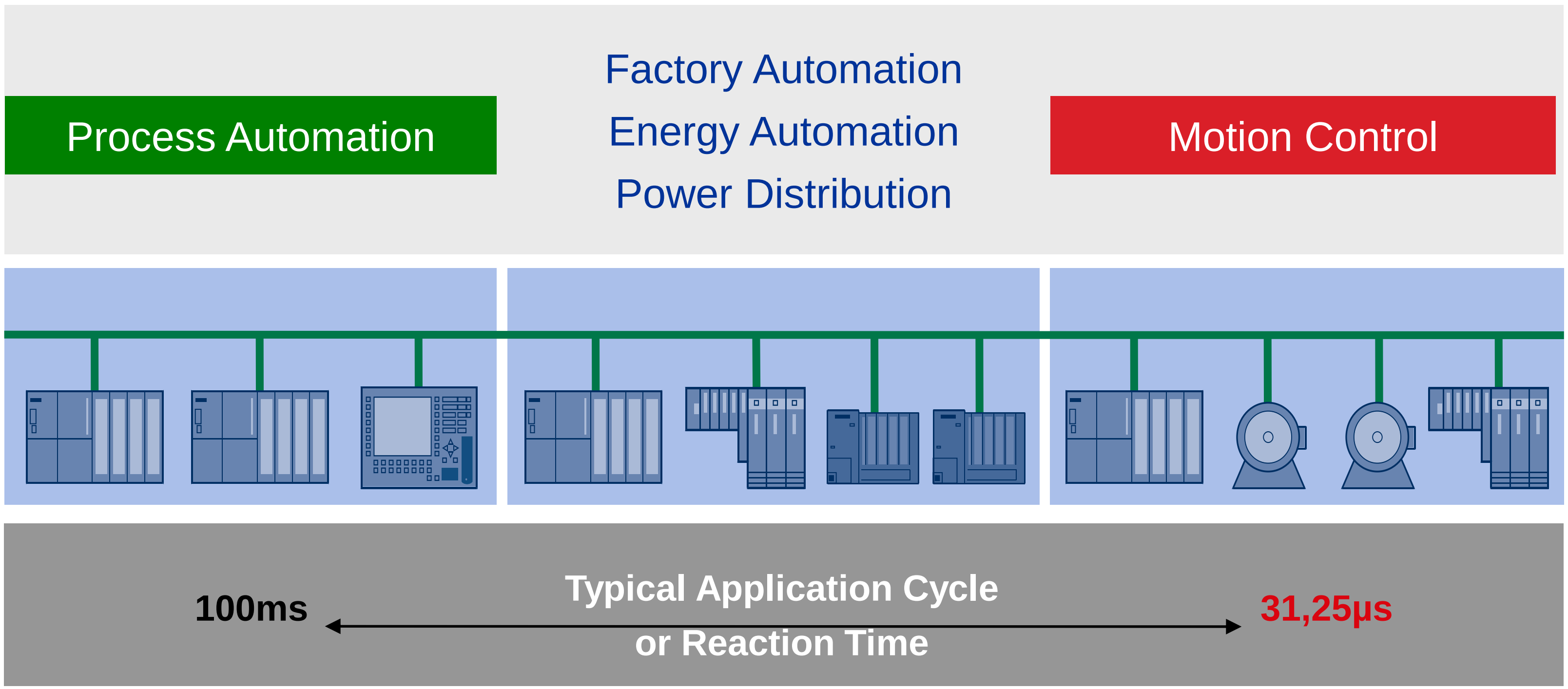
- Industrial Automation
  - Factory Automation
    - e.g. PLC, Motion Control, Robots
  - Process Automation
    - e.g. Oil, Gas, Chemical / Petrochemical, Food & Beverage
- Energy Automation
  - Power Generation
    - e.g. Fossil Power Plants, Wind Turbines
  - Power Transmission and Distribution
- Building Automation
  - Climate Control
  - Fire Safety

New Markets

- Avionics
  - Fly-by-Wire
- Railway Systems
  - Train Control
  - Railway Traffic Management Systems
- Medical

Industrial Automation

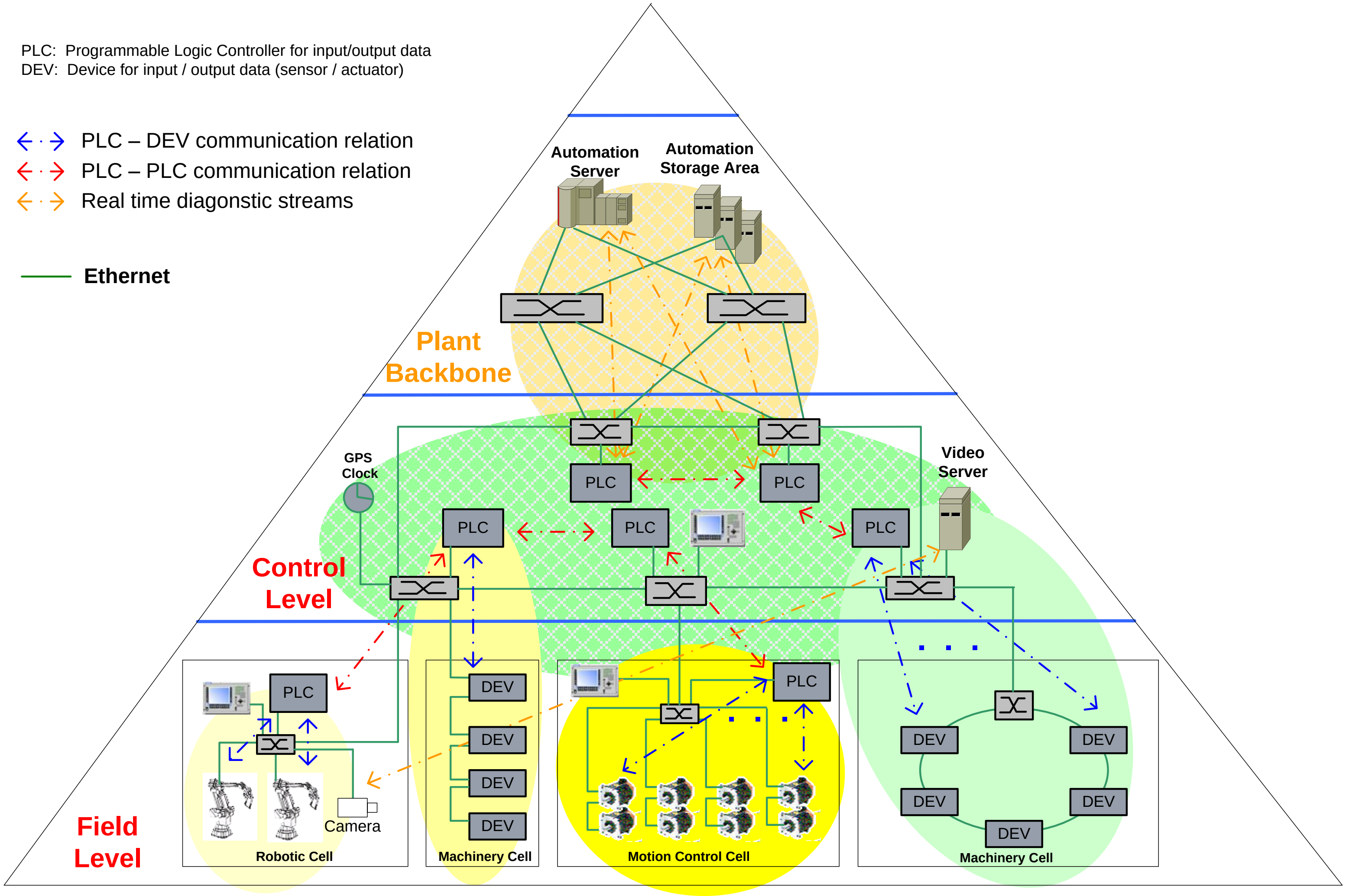
Wide Range of Industrial Performance Requirements to be met



Industrial Automation

The automation pyramid:Hierarchical layering in one industrial automation networked system

PLC: Programmable Logic Controller for input/output data
DEV: Device for input / output data (sensor / actuator)

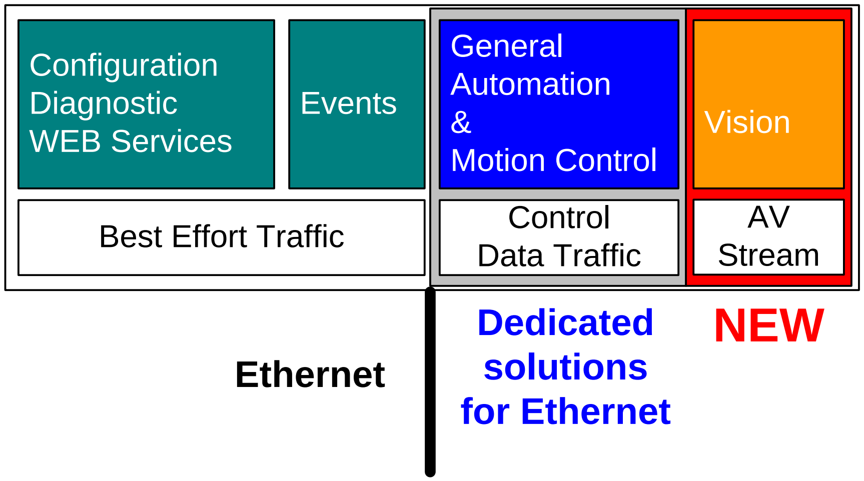


Industrial Automation

Industrial Communications Services

Services:

- **Best Effort Traffic**
  - Configuration
  - Diagnostic
  - Web Services
  - Events
  - …
- **AV Streams**
  - Real Time Diagnostic
    e.g. inspection, identification, tracking, counting and measurement
  - Vision Systems
- **Control Data Traffic**
  - General Automation to exchange typical analog an digital values
    e.g. manufacturing and process industry
  - Motion Control to exchange typical analog and digital values
    from actuators and sensors based on synchronized processes

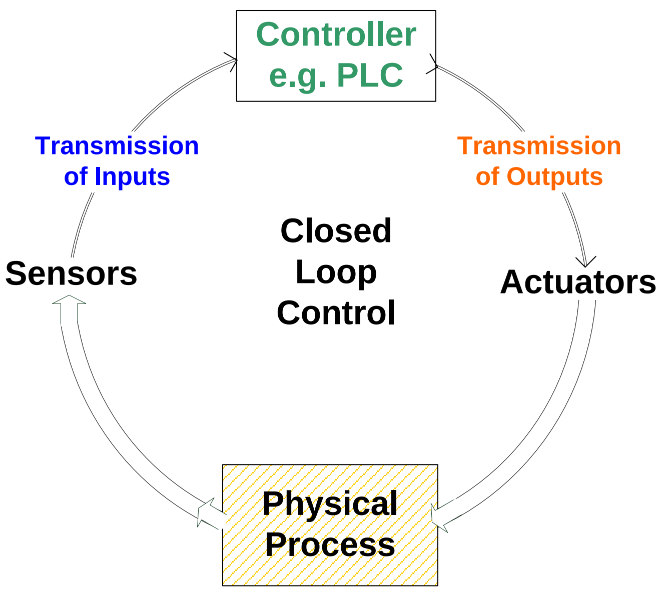


	Today		Future
Best Effort Traffic:	No guaranteed bandwidth		Guaranteed amount of bandwidth
AV Streams:	Separate network		Guaranteed QoS for AV Streams in one network
Control Data Traffic:	Dedicated solutions to guarantee min latency, resources and bandwidth		Standardized solution to guarantee min latency, resources and bandwidth

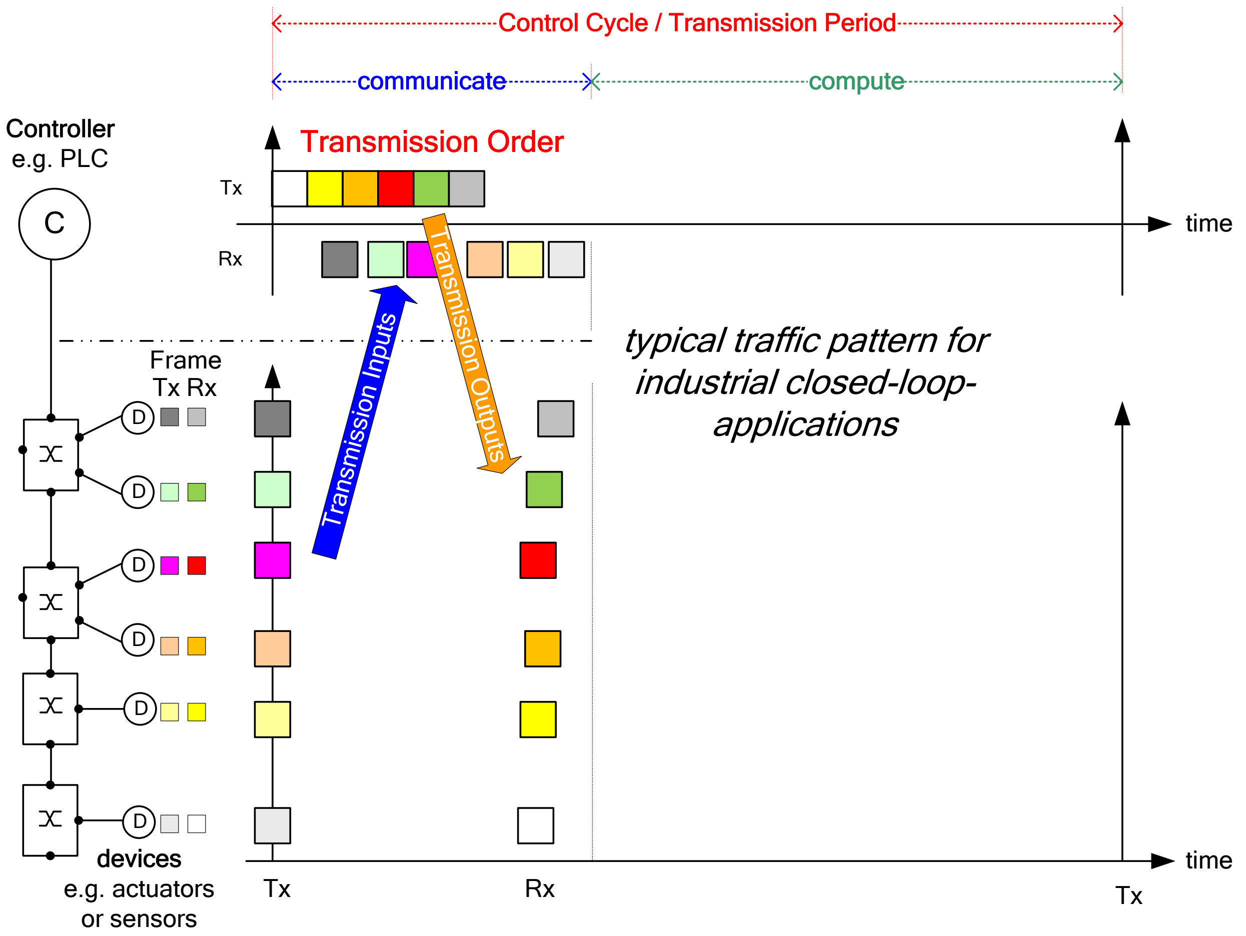
Typical Application

Closed Loop application

The application model for control traffic <u>is different</u>



Typical traffic pattern for control traffic



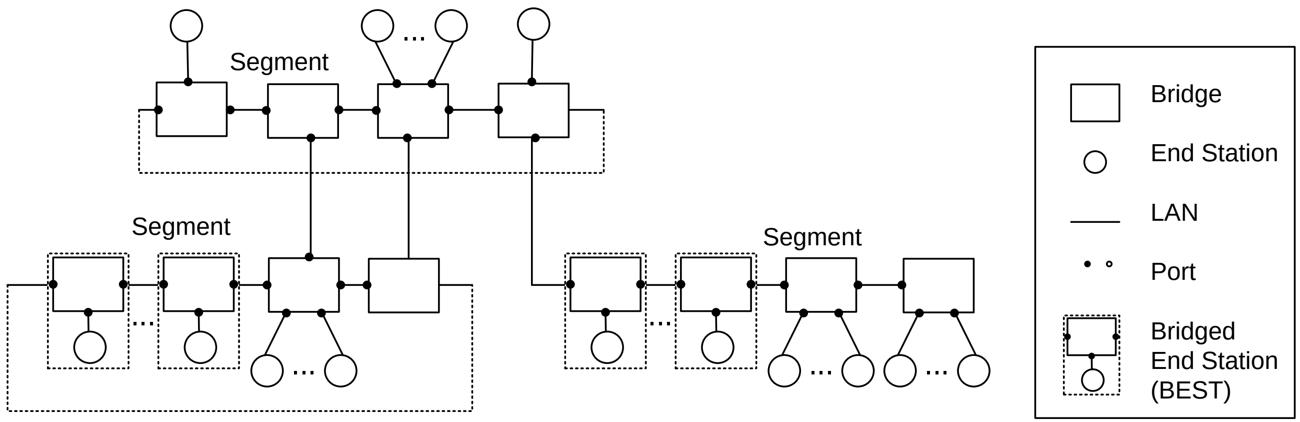
typical traffic pattern for industrial closed-loop-applications

Topologies

Linear Topologies (daisy chain, ring)

Core

Distribution
+ Access



Reasons for linear topologies:

- Applications are normally physically distributed in a linear manner , i.e. production lines
- The network is tailored to the application, i.e. line topologies
- To reduce total cost, 2-port-bridges are integrated into the end stations
- Ring topologies are used to provide high availability
- Minimize wiring effort -> avoid thick cable harness –> cable channel
- Automatic tool changer by robots ->reduce number of connectors

Best effort traffic and AV Streams cause additional delay for control data traffic on each hop

Motivation for LOW Latency

Low Latency for Control Data Traffic

Reasons for low latency:

- Increase rate or speed of production
- Increase product quality by increasing product accuracy
- Reduce resources

=> Minimize control cycle

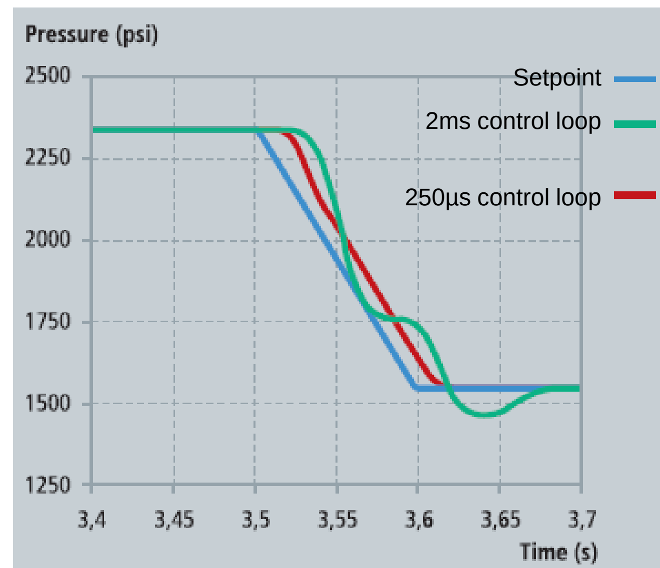
Reduce Resources Example (plastic moulding)

Control cycle optimization through minimization of the response time and constant dead time is not an end in itself

In addition to enhanced quality, the aim is to save material through reduced wall thickness:

In the application example a reduction in part weight by only 2 grams results in the following savings for the machine user:

- **Cup weight today (g): 22**
- **Cup weight target (g): 20**
- **Parts / year: 54,568,421**
- **Annual material saving (kg): >109,137**
- **Savings per year ($ US): >200,000**



Use Case 1 - Motion Control

Trend of mechatronic solutions

TREND

Historical solution

- expensive mechanical components
- rigid coupled

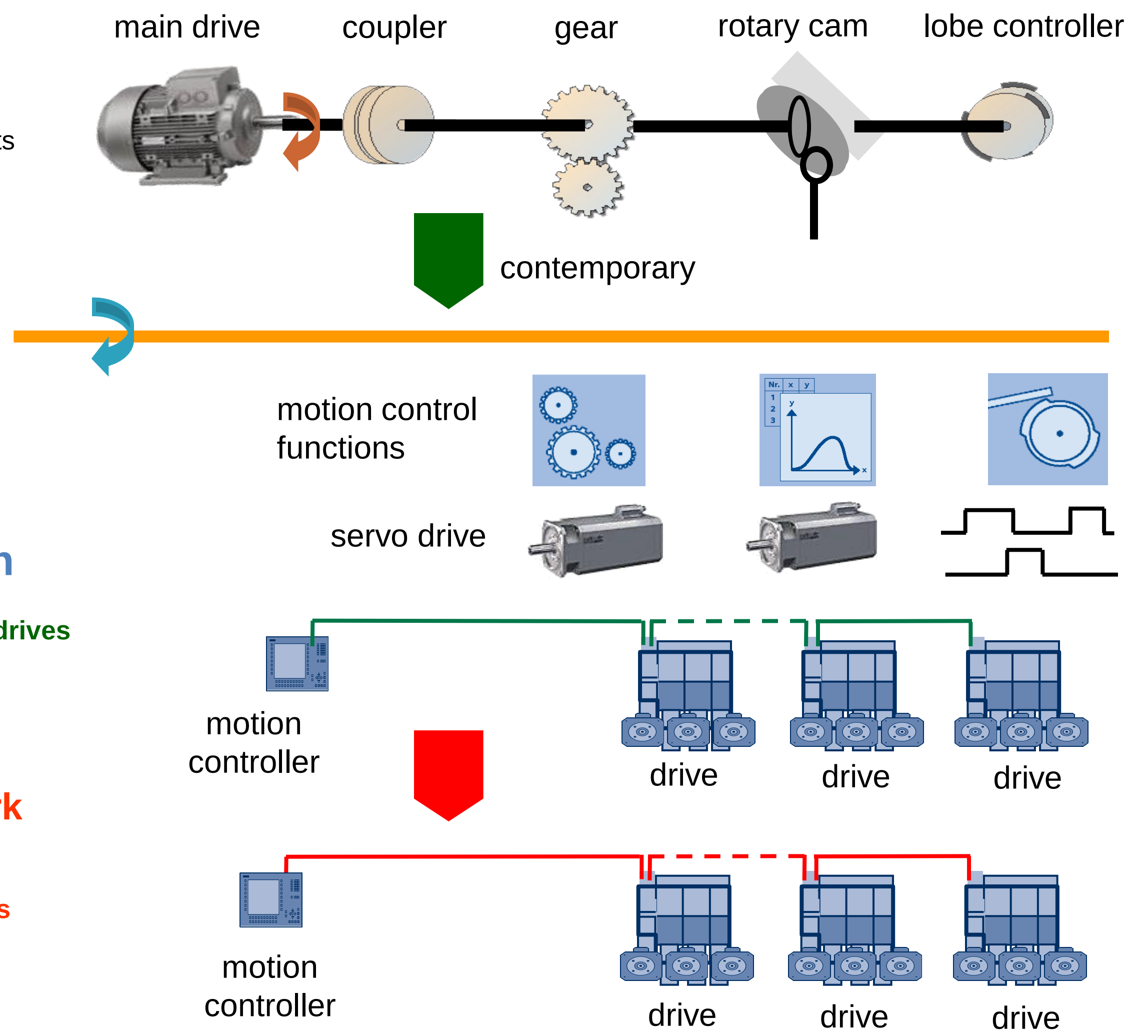
Virtual Master Axis

Mechatronic solution

- **dedicated Ethernet network for drives**
- **separate network for diagnostic**

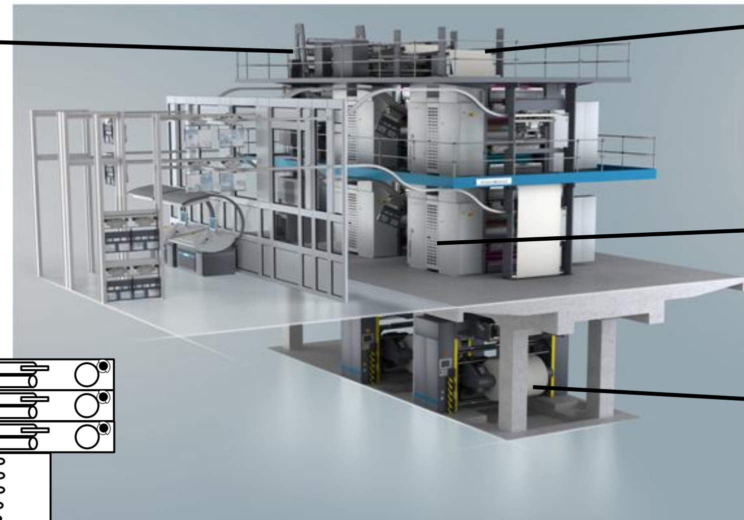
Unified Ethernet network

- **for drives (control data)**
- **<u>and</u> real time diagnostic streams**



Use Case 2 – Printing Machine

Finishing Unit
- Turner Bars
- Folder superstructure
- Folding Unit

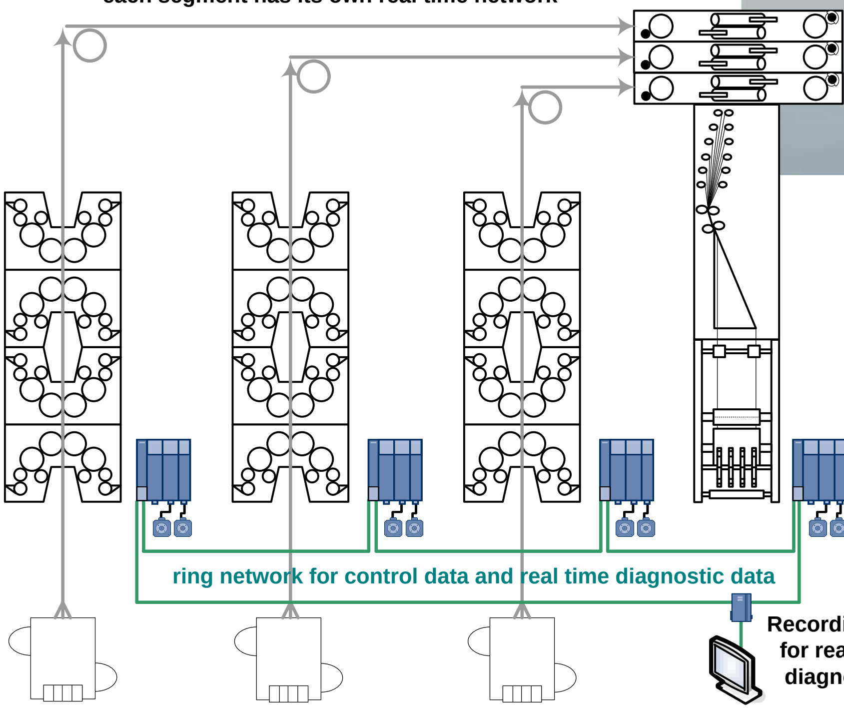


Guide Rollers
(to route Paper Webs)

Printing Unit

Reel Stands
and Paper Rolls

- multiple printing units per column
- up to 10 axes per column
- a segment can consist up to 10 column
- each segment has its own real time network

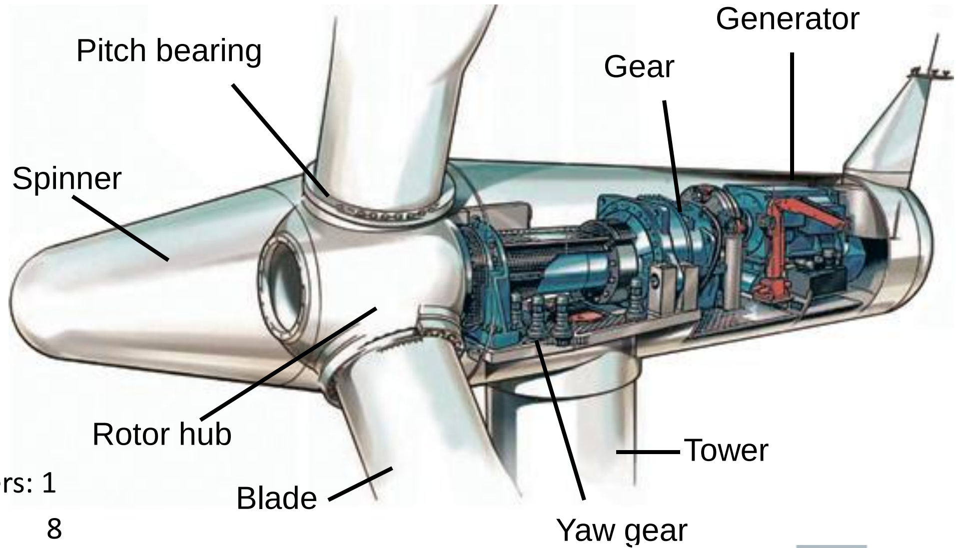


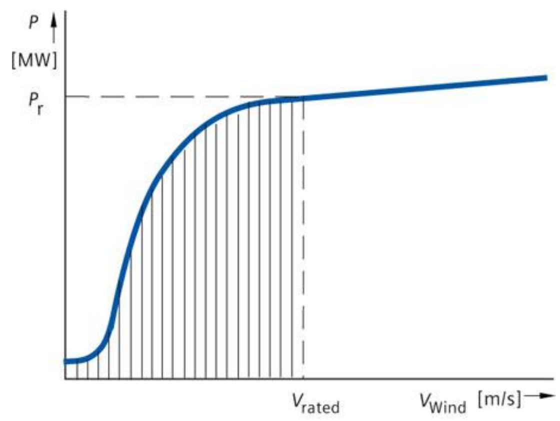
Requirements:

- Transmission Period: 1ms – 125µs, in future 62,5
- Typical control data size: 12 – 60 bytes/axis
- Maximum number of axes: 100 in one segment
  - 1000 axes for huge machines (e.g. huge printing machine)
- Topology: daisy chain, ring, comb in combination with star
- High available network – seamless redundancy
- Real time diagnostic for monitoring systems
  - Payload size of streams: ~400 bytes
- Unified network with control data class, AV streams, and legacy traffic
  - Payload size of legacy traffic : ~1500 bytes

Control network will benefit form pre-emption mechanism

Use Case 3 - Pitch Control for Wind Turbines

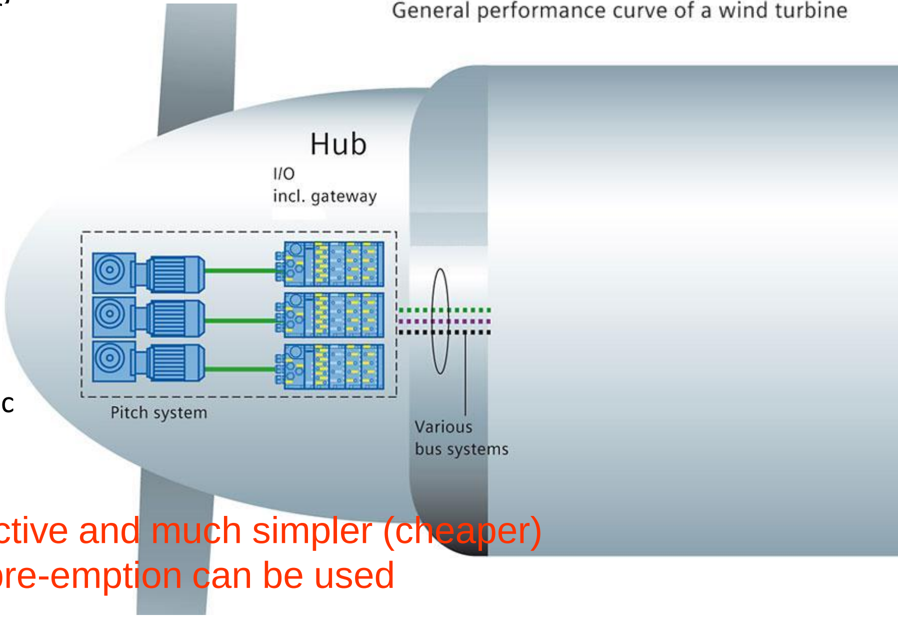




General performance curve of a wind turbine

Requirements:

- Number of controllers: 1
- Number of devices: 8
- Typical size of control data:
  - Input: 40 bytes
  - Output: 40 bytes
- Transmission Period: 125us in future 62,5µs
- Topology: daisy chain
- Communication relationship: one-to-one
- Real time diagnostic for monitoring systems (e.g. vibration)
  - Payload size of stream: ~400 bytes
- Unified network with control data class, AV streams, and legacy traffic
  - Payload size of legacy traffic: ~1500 bytes



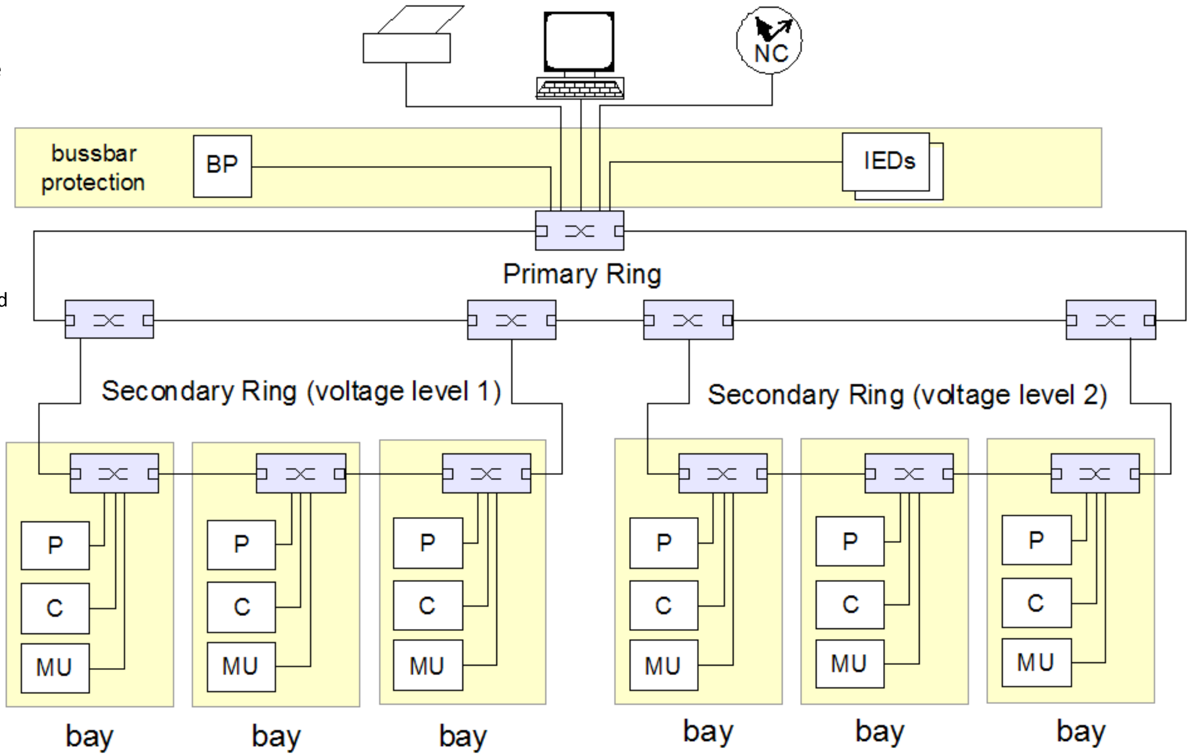
Future implementation can be more effective and much simpler (cheaper)
if standard-based solution like pre-emption can be used

Use Case 4 – IEC 61850 Substation Automation System

Substation Automation according to IEC 61850 is an integral part in many Smart Grid Initiatives around the world!

From TR IEC 61850-90-4 draft

IED = Intelligent Electronic Device
MU = Merging Unit
P = Protection Relay
BP = Busbar Protection Relay
C = Bay Controller
NC = Network Clock
GOOSE – Generic Object Oriented Substation Event.



Traffic characteristics:

High Priority

- **L2 event trip messages (GOOSE)**
- **L2 cyclic traffic (Sampled Values)**

Low Priority

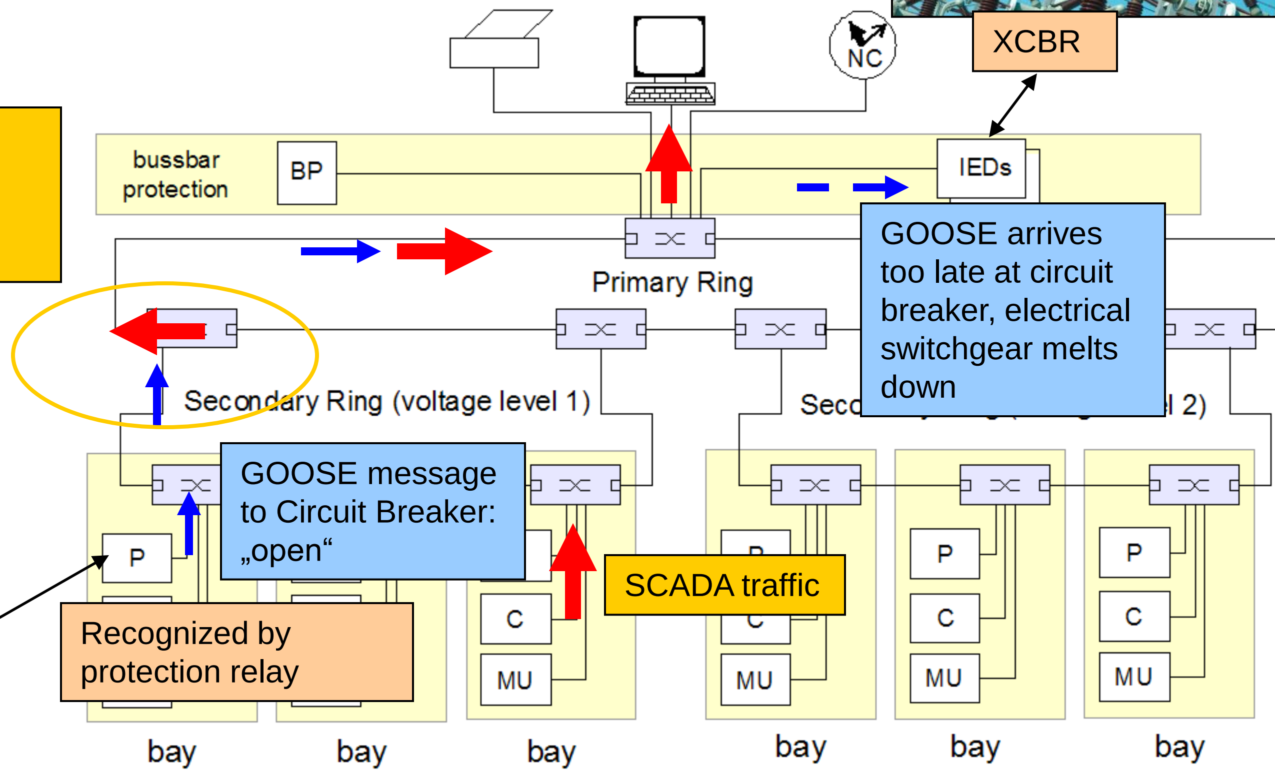
MMS (Manufacturing Messaging Specification)
File up-/downloads, device configuration

- **IEC 61850 networks carry time-critical traffic (GOOSE and SV) for communication between IEDs as well as non time-critical traffic, e.g. SCADA file uploads/downloads.**
- **MMS traffic, e.g. (config) file up-/downloads increase the latency for time-critical GOOSE and SV traffic**
- **Pre-emption enables running both SCADA/MMS and time-critical traffic across one convergent network simultaneously, minimizing the risk of missing a critical deadline**

Use Case 4 – IEC 61850 Substation Automation System

Substation Automation according to IEC 61850 is an integral part in many Smart Grid Initiatives around the world

Best effort SCADA traffic (file upload/download, reporting, etc.) impairing GOOSE latency



Earth fault introduced

Industrial Automation Requirements

from new-avb-tretter-requirements-of-industrial-applications-0711.pdf

QoS requirements for industrial control data class

Performance requirements for industrial control data class

- Maximum latency: < 3µs per hop @ 1 GE
    - Guaranteed latency
    - Topology independent
    - Industrial control data has higher priority than AV streams
        - -Unified network with control data class, AV streams, and legacy traffic

Preconditions for performance requirements

- Network topology: star, line, ring and combinations
- Network attributes
    - Maximum 64 hops
    - Maximum number of nodes (bridged end stations & end stations): 2000
    - Maximum cable length: 100m
    - High available network - seamless redundancy for control data class
- Industrial control data class attributes
    - Typical data size (payload size): 40 - 300 bytes
    - Range of transmission period: 31,25µs – 1ms
    - Payload size for real diagnostic data: ~400 bytes
    - Payload size for best effort traffic: 1500 bytes

These are our best estimates derived from multiple assumptions of the current and future industrial applications

The Need

Preemptive Transmission advantages – Is it Worth the Effort?

Technology Aspect
Norm Finn
(nfinn at cisco dot com)

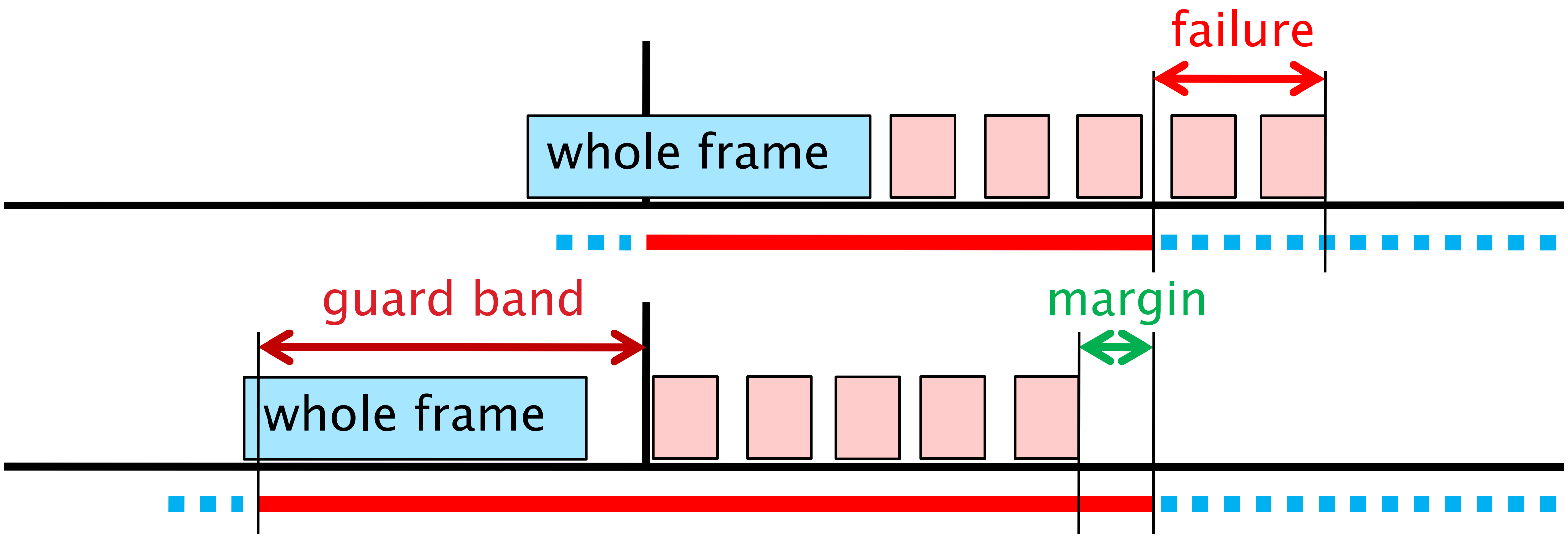
http://www.ieee802.org/1/files/public/docs2012/new-avb-nfinn-preempt-advantage-0112-v02.pdf

Scheduling is required for real-time nets

- The real-time network scheduling model is: communicate, compute, communicate, compute, …

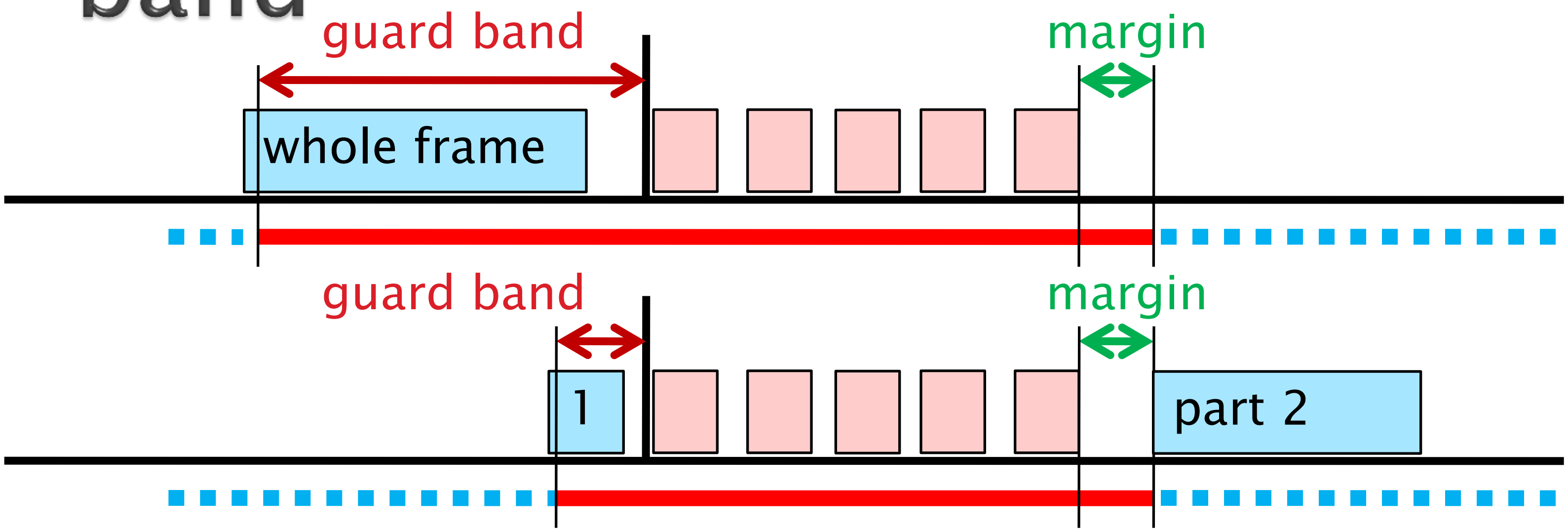
- Communication occurs at specified times.
- The scheduled cycle timing is driven by the requirements of the critical application.
  - Scheduled cycle does not scale with link bandwidth
- Only by **strict scheduling** can we guarantee, **no matter what happens**, that we will respond to external events in a timely manner.
- There is an IEEE 802.1 PAR that addresses this.

Guard bands are necessary



- If an interfering frame starts transmission just before the start of a reserved time period, it can extend critical transmissions outside the window.
- Therefore, a **guard band** is required before the window starts, equal in size to the largest possible interfering frame.

Preemption shrinks the guard band



- If preemption is used, the guard band need only be as large as the largest possible interfering fragment, instead of the largest possible interfering frame.
- It is easy to see that **the smaller the size of the time-reserved windows, the larger the impact of preemption**.

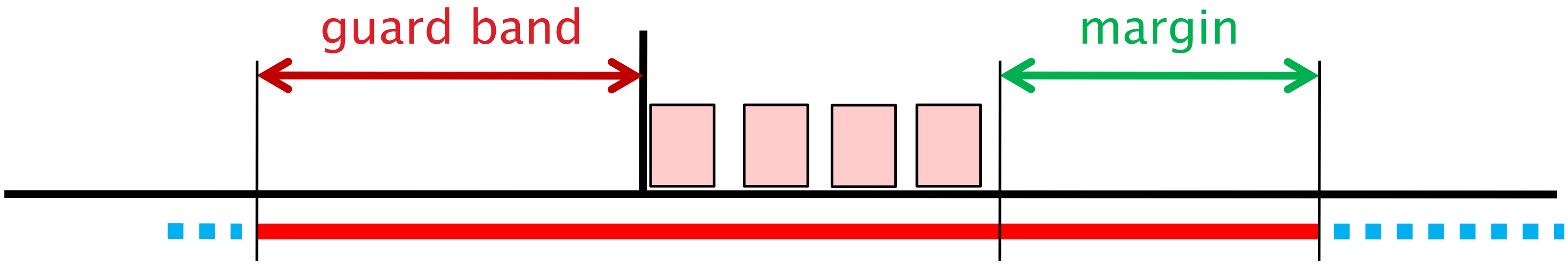
Some numbers (Example Assumption)

- Let us assume that the time-critical data frames are typically small, say 128 bytes.
- Let us assume that standard 1522-byte data frames are permitted for all other traffic.
- We will use the standard 20 bytes for preamble, start of frame delimiter, and inter-frame gap.
- Let us assume that preempting a frame adds only an extra 20 bytes; this is the minimum practical penalty.
- We will assume that the worst case frame size is 127 bytes, which cannot be preempted. A 128-byte frame could be preempted and separated into two 64-byte fragments.

Example 1

- In the first example, the time window is sized for **four 128-byte frames** with a margin of 4 more such frames (50% utilization of the window).
- The basic window size is 8 * (128 + 20) = 1184 byte times.
- Without preemption, we require a (1522 + 20) byte guard band, for a total window size of 2726 bytes.
- With preemption, we require a (127 + 20) byte guard band, for a total window size of 1331 bytes.
- Thus, whatever percentage of the total bandwidth is allocated to time-critical traffic, it requires **more than twice as much time** (2726/1331) be reserved for that traffic if **preemption** is **not utilized**.

Example 1

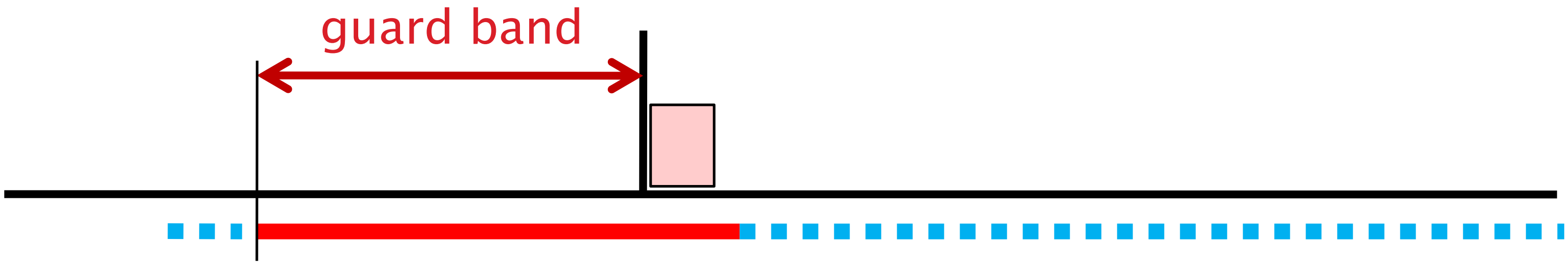


- 1184 byte window, including margin, plus guard band.
- Bandwidth is scaled up by increasing windows per second, not by making windows larger, because the applications determine the window size, not the available bandwidth.

Example 2

- In the second example, the time window is sized for **one 128-byte frame** with no margin. This is perfectly possible if we assume that the switch will store a time-critical frame very briefly, and then transmit it at the appropriate moment.
- The basic window size is (128 + 20) = 148 byte times.
- Without preemption, we require a (1522 + 20) byte guard band, for a total window size of 1690 bytes.
- With preemption, we require a (127 + 20) byte guard band, for a total window size of 295 bytes.
- Thus, whatever percentage of the total bandwidth is allocated to time-critical traffic, it requires **more than 5.7 times as much time** (1690/295) be reserved for that traffic if **preemption** is **not utilized**.

Example 2



- 148 byte window, no margin, plus guard band.
- Bandwidth is scaled up by increasing windows per second, not by making windows larger, because the applications determine the window size, not the available bandwidth.

Percent of time reserved for critical data

Link speed	Mbits/sec required for critical data alone	(no guard band)†		with preemption		no preemption	
		1184 B* window	148 B* window	1184 B* window + guard	148 B* window + guard	1184 B* window + guard	148 B* window + guard
100 Mbits	0.1	0.23%	0.12%	0.26%	0.23%	0.53%	1.32%
	1	2.31%	1.16%	2.60%	2.30%	5.32%	13.20%
	10	23.13%	11.56%	26.00%	23.05%	53.24%	132.03%
	30	69.38%	34.69%	77.99%	69.14%	159.73%	396.09%
1 Gbit	1	0.23%	0.12%	0.26%	0.23%	0.53%	1.32%
	10	2.31%	1.16%	2.60%	2.30%	5.32%	13.20%
	100	23.13%	11.56%	26.00%	23.05%	53.24%	132.03%
	300	69.38%	34.69%	77.99%	69.14%	159.73%	396.09%

* Half of 1184-byte window is reserved for margin; none of 148-byte is margin.
† "No guard band" shows wastage from margin, preamble and inter-frame gap.

Technical Needs Summary

- Mixing time-scheduled, bandwidth reserved, and best-effort traffic is the goal.
- Guard-bands are required to ensure time-scheduled transmissions happen on time, in the presence of bandwidth reserved and best-effort traffic.
- But, guard-bands have a big impact on the bandwidth available for bandwidth reserved and best-effort traffic.
- Preemption drastically reduces the size of the guard-bands.
- **Mixing these three traffic types is much more practical if preemption is available.**

Why Now?

Why Now? IEEE 802.1 PARs

- IEEE 802.1Qbu “Preemption” PAR, submitted to EC for approval (March 2012), needs support from the 802.3 to complete its system solution.
  - PAR Title:
    “802.1Q Amendment: Frame Preemption”

- IEEE 802.1Qbv “Time Aware Scheduler” PAR, submitted to EC for approval (March 2012), will have an effect to this study.
  - PAR Title:
    “802.1Q Amendment: Time Aware Scheduler”

Why Now? Market Served

Convergence

- IEEE 802.1 AVB & IEEE 802.3bf provides Audio and Video real-time traffic over mainstream 802.3 Networks (Done!).
- IEEE 802.1 AVB "2.0" would provides scheduled time-critical traffic over 802.3 Networks.
  - Scheduled time-critical traffic need both preemption (802.1Qbu) and time-aware scheduler (802.1Qbv).
  - Preemption PAR requires this frame preemption support in 802.3 MAC to deliver a complete and optimized system solution.
  - IEEE 802.1 PARs generated in support of 802.1 AVB "2.0" objectives.

Scheduled time-critical traffic use.

- Industrial Automation, Motion Control, Power Stations, manufacturing automation, and other control systems.
- Automotive systems, chassis, body, driver assist, and other control system bus backbone applications and future direct attached control systems applications.

Preemption Work

802.3 and 802.1 Coordination

802.1 & 802.3 Preemption Coordination

- The Problem: Re-stated "MAC Service interface (boundary between 802.1 and 802.3) as defined today does not support preemption"
- Proposed Coordination:
  1. Joint 802.1/802.3 Scope: 802.1 (ISS/EISS) and 802.3 MAC Service Interface studied to support preemption.
     - 802.3 members has specific expertise to understand the implication of this interface.
     - 802.1 members has specific expertise to convey requirements for preemption MAC services.
  2. 802.1 Scope: 802.1 Services determine frames that are preemptable, or, preempting.
     - Effect of AVB shaper, transmit selection, congestion management (if relevant) TBD during standardization.
     - Layer management generated frame types are to be considered and classified as preempting or preemptable during standardization process.
  3. 802.3 Scope: 802.3 to consider preemption related enhancements.
     - Meeting the requirements of 802.1 preemption in a 802.3 PAR process.

Thank you!