Why an additional Shaper

Marcel Kiessling
Distributed Embedded Systems

SIEMENS

UNIVERSITÄT PADERBORN
Die Universität der Informationsgesellschaft

Outline

- Recap: Industrial Requirements for Latency
  - Different Application requirements regarding Latency
- Sources of Latency
  - Reasons for the special treatment of Streams
- Summary

Recap: Industrial Automation Applications

Within industrial we have to differentiate two Systems:

- **Closed Systems**

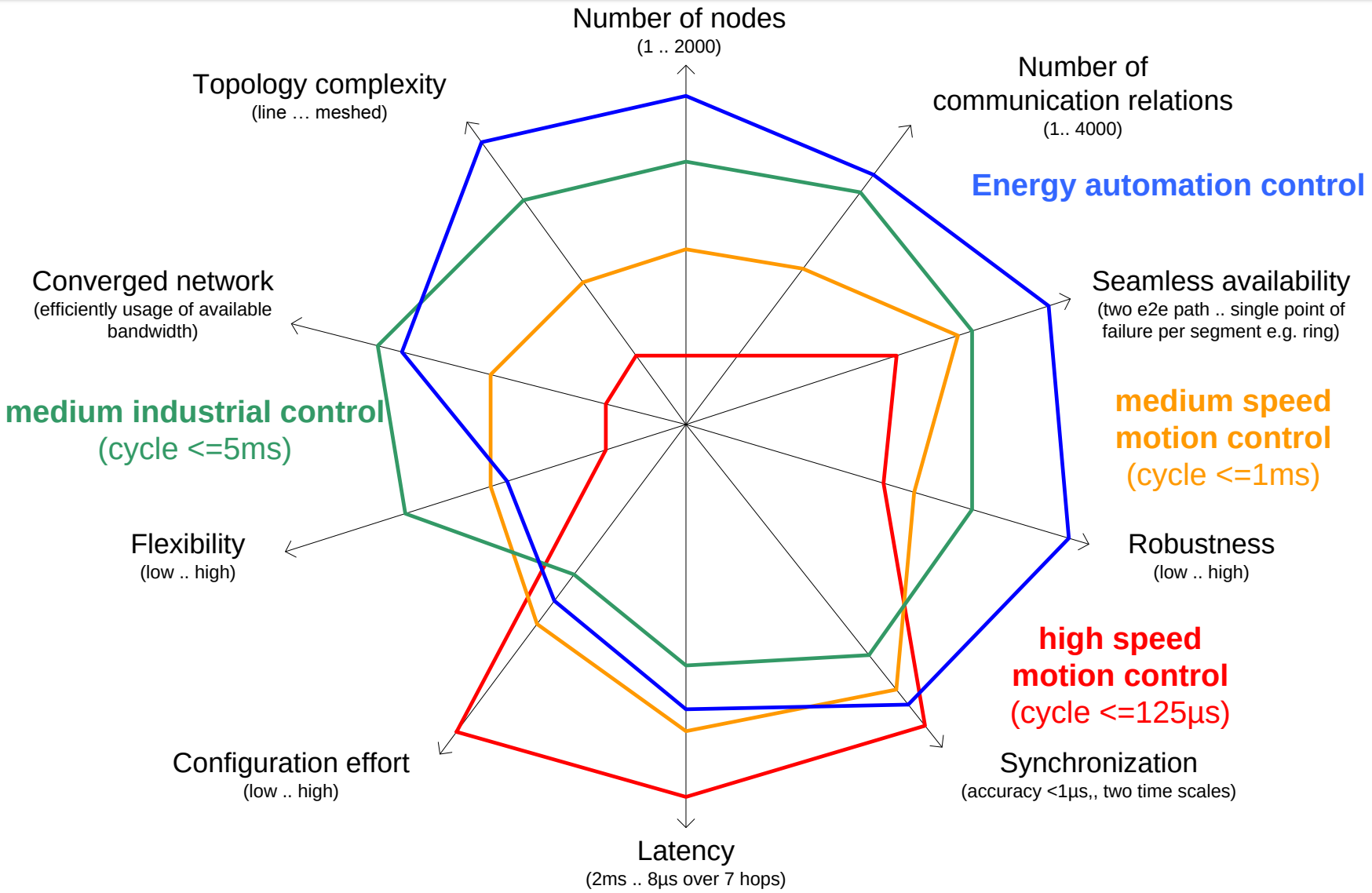
  Typical used for **"**Closed-Loop-Applications" like motion control system
  - One network for one application – this application is fixed
  - Fix topology – adapted to application
  - *Guaranteed QoS & guaranteed low latency*
    - *Highly Optimized, Static "PCE" case for high performance requirements*

- **Open Systems**

  Typical used for "Control-Applications" like assembly lines
  - *Multiple automation applications share the network*
  - *Topology can change when applications are added , changed or removed at runtime*
  - *Multiple Control-Data-Traffic Classes within one network (e.g. multiple transmission periods)*
  - *Guaranteed QoS & guaranteed low latency*
    - *Dynamic use case with multiple flexible traffic classes*

Industrial networks can also consist of one Closed and multiple Open Systems

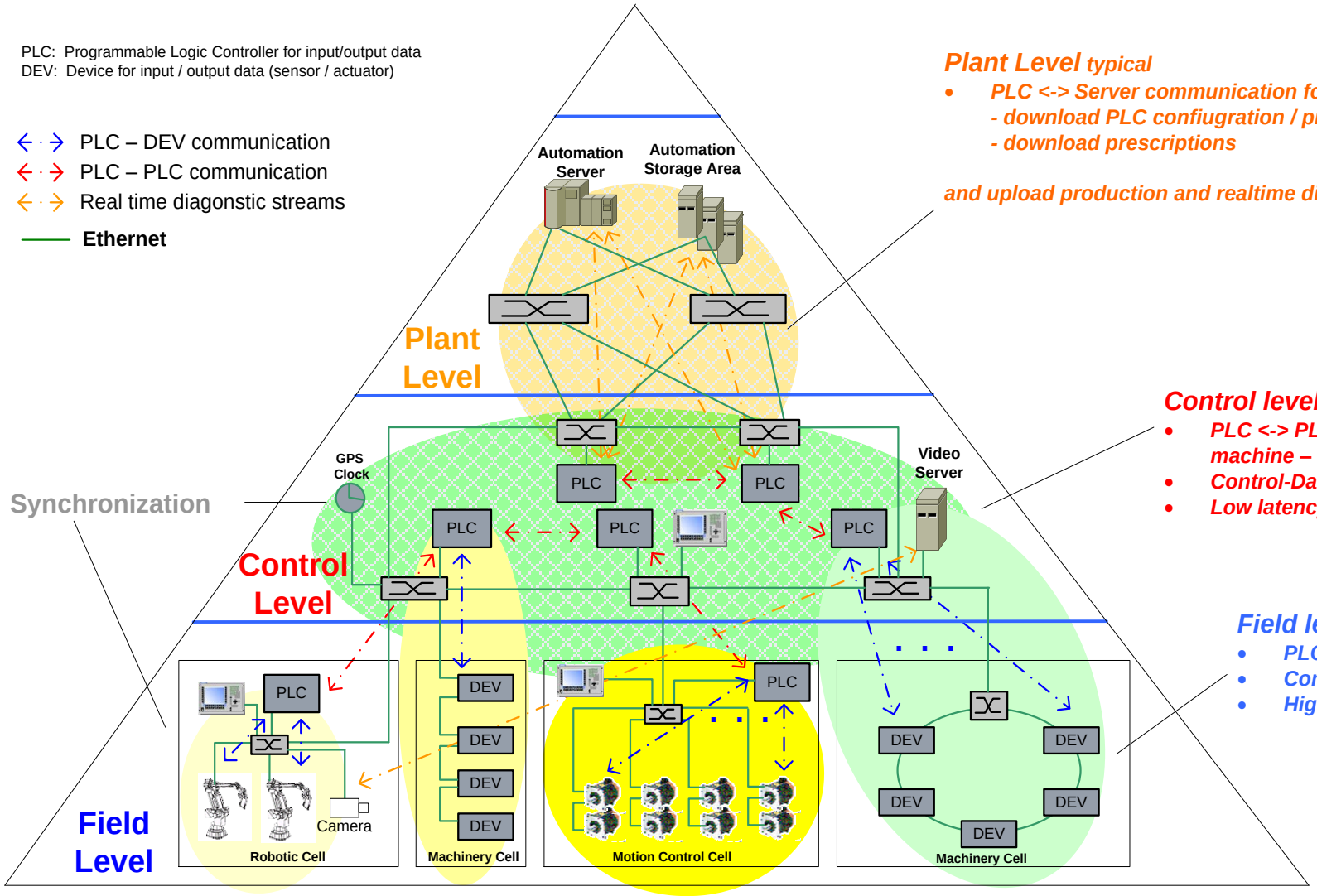
Recap: Different Applications with different Requirements



Different Applications within Hierarchical Automation Networks

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

PLC – DEV communication
PLC – PLC communication
Real time diagonstic streams
Ethernet



***Plant Level** typical*

- *PLC <-> Server communication for*
  - download PLC confiugration / program
  - download prescriptions

and upload production and realtime diagnostic data

***Control level** typical:*

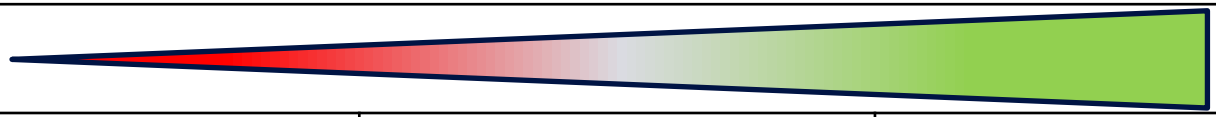
- *PLC <-> PLC / machine – machine communicaiton*
- *Control-Data-Traffic (CDT)*
- *Low latency*

***Field level** typical:*

- *PLC <-> DEV*
- *Control-Data-Traffic (CDT)*
- *Highest req. on latency*

Industrial Requirements on Communication at the Field- / Control- & Plant-Level (1/2)

SIEMENS

Features	Field-Level	Control-Level	Plant-Level
Link Speed	100 MBit or less	>= 1 GBit	
Max. Stream Bandwidth	< 50%	< 15%	?
Typical Traffic	Control-Data-Streams	slow Control-Data-Streams AV-Streams	AV-Stream for diagnostic and measurement
	… mixed with Synchronization, Network Control and Best-Effort-Traffic		
Transmission cycle **(TC)**	31,25 µs  10ms		
Low latency	Closed-Loop-Applications **~ <= 20% of TC** Control-Applications **~ TC** *(Open Systems)*	Control-Applications **~ TC** *(Open Systems)*	Application dependent **~ 2.. 10ms**
Transmission modes	scheduled and coordinated or scheduled	periodical or scheduled	periodical
Max. frame size	64 byte  600 byte		
Topology	simple topologies Daisy Chains or Rings (e.g. 64 2-Port devices)	more complex topologies e.g. combination with rings, coupled rings, trees and stars, …	

Industrial Requirements on Communication at the Field- / Control- & Plant-Level (2/2)

SIEMENS

<table>
<tr><th>Features</th><th>Field-Level</th><th>Control-Level</th><th>Plant-Level</th></tr>
<tr><td>Flexibility</td><td>Closed Systems with less requirements on flexibility</td><td colspan="2">High requirements on flexibility
Add and remove of streams at runtime
without any effect on established streams</td></tr>
<tr><td>Seamless Failover</td><td>single Rings</td><td colspan="2">ISIS-PCR for more complex topologies and flexibility</td></tr>
<tr><td rowspan="2">Path reservation</td><td>Closed-Systems
with Static path reservation</td><td colspan="2"></td></tr>
<tr><td colspan="3">Open-Systems with dynamic path reservation at runtime</td></tr>
<tr><td rowspan="2">Bandwidth & resource reservation</td><td colspan="3">is required for guaranteed QoS</td></tr>
<tr><td>Closed-Systems with static configuration
Open-Systems with dynamic configuration</td><td>Open-Systems with dynamic configuration</td><td>Dynamic configuration</td></tr>
</table>

Outline

- Recap: Industrial Requirements for Latency
  - Different Application requirements regarding Latency
- Sources of Latency
  - Reasons for the special treatment of Streams
- Summary

Guaranteed Latency

Maximum effect of all Latency Sources must be considered

AVB / TSN: Avoid or Limit effect for Streams

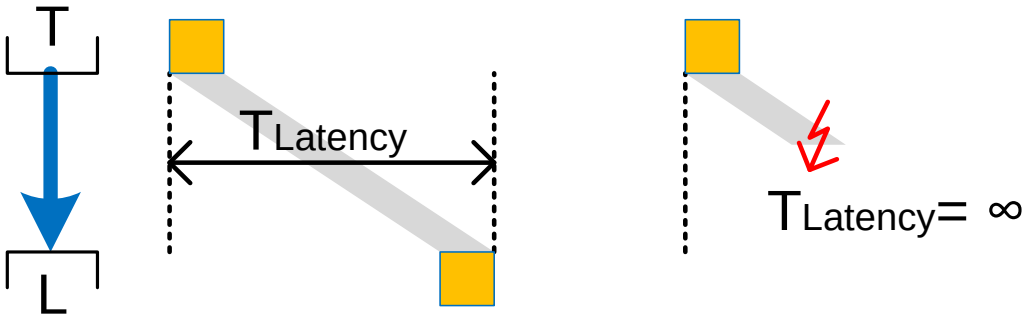
- Own class with preferred forwarding
- Reserved Resources for known Characteristic of Streams

Sources of Latency

- Loss of Frames – Infinite End-to-End Latency
- Priority – Traffic from classes with higher priority
- Priority Inversion – Traffic from classes with lower priority
- In-Class Interference – Traffic from the same class
- Bridge Delay and other HW dependent effects
- others …

Loss of Frames

Loss of frames is the worst case latency (e.g. due to congestion)

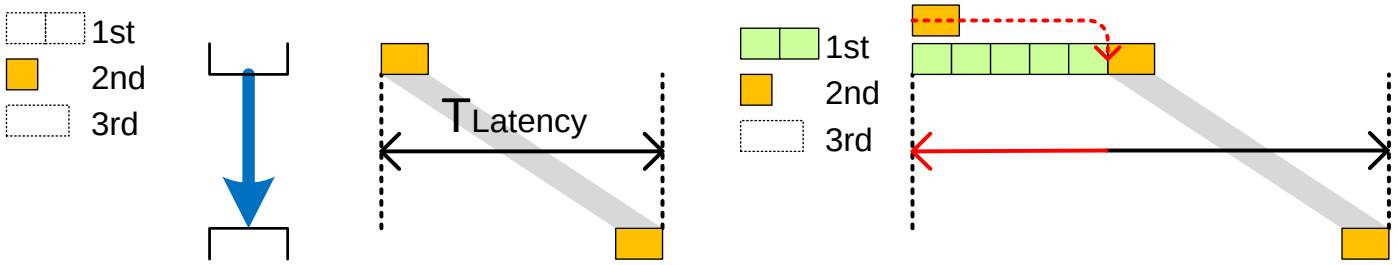


Standardized AVB mechanisms avoid congestion loss

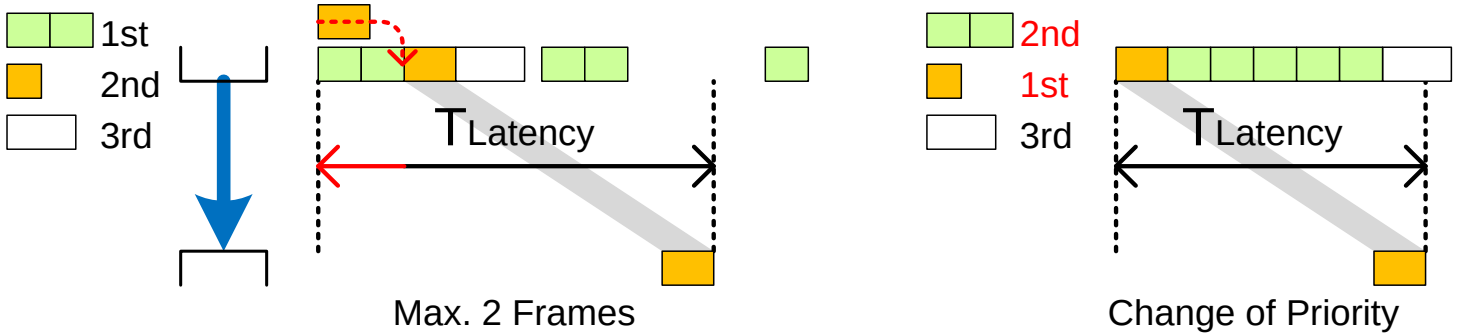
- Preferred Forwarding - Own Class (with highest Priority) for Streams
- Reservation of Resources along the path
- CBSA Shaper to align the Stream with the reserved Resources (special kind of Policing for streams)

Priority

Other Frames are transmitted first (higher priority)

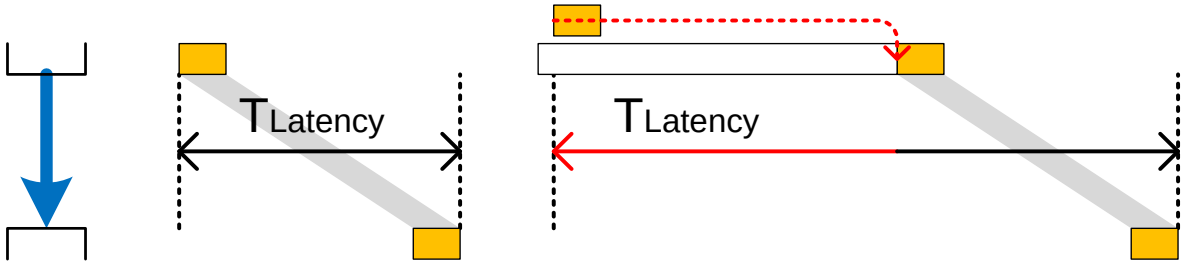


- **Limited Burst-Size**- to limit the effect of Interference
  (CBSA in normal operation spreads the traffic to avoid bursts)
- **Highest Priority** - to avoid higher priorities

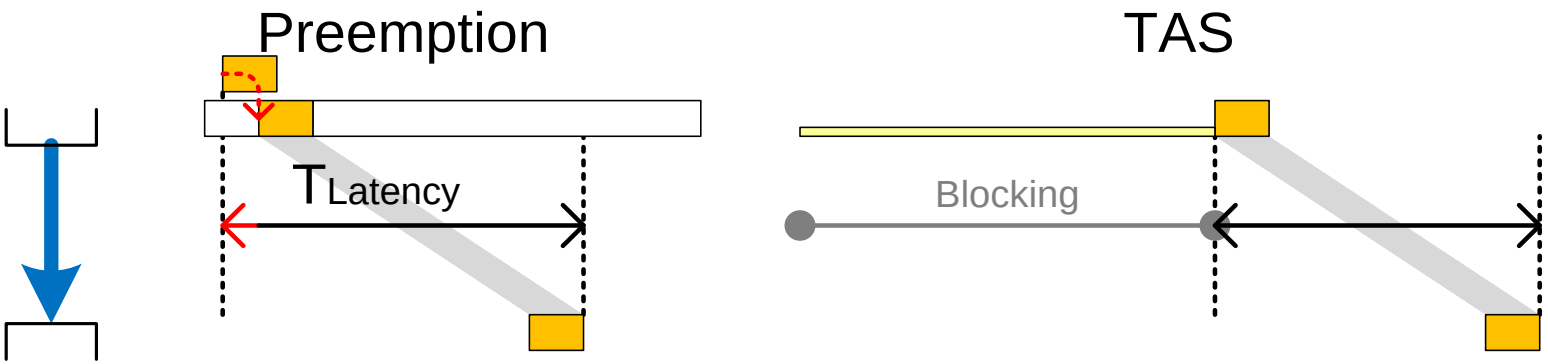


Priority Inversion

Other Frames are transmitted – busy target Link

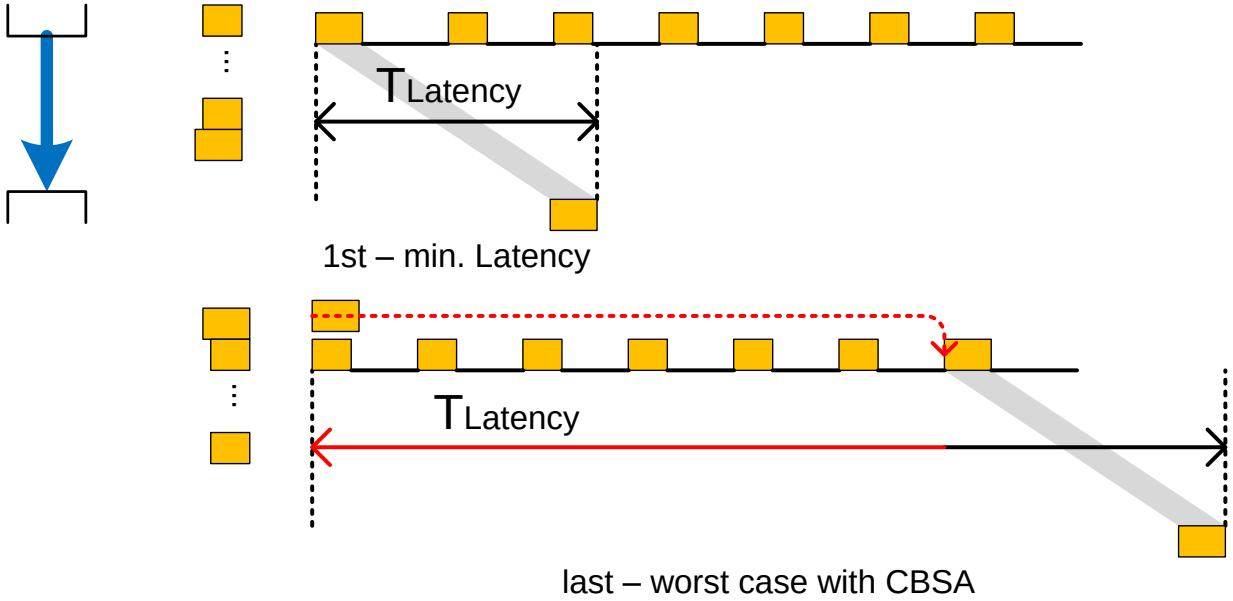


- **Preemption** - (802.3 br) to shorten the interference
- **TAS** - (802.1 Qbv) to avoid the Interference
  (block the link for other traffic in advance)



In-Class Interference

Other Frames of the same class can be transmitted first



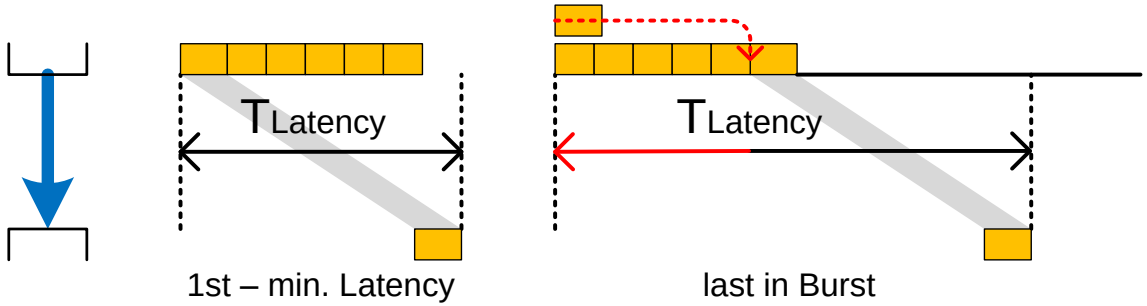
AVB Maximum Latency based on the Worst-Case
(additional **CBSA** latency from spreading between frames)

New **TSN** Shapers to avoid the additional spreading:
UBS / TAS with coordination

In-Class Interference - BLS

- **BLS** – Burst Limiting Shaper *(Policer)*

Other Frames of the same class can be transmitted first



Strict Priority releases all frames in one burst

(Basically Strict Priority in normal operation – robustness for lower traffic classes in case of errors by dropping frames)

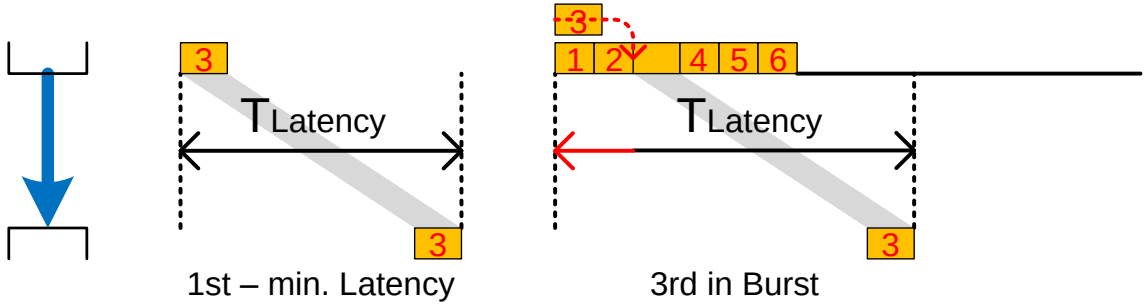
BLS as Policing to ensure the max. size of a burst

(Limit the size of interference to a maximum amount for lower priorities in case of an error)

In-Class Interference - UBS

- **UBS** – Urgency based Shaper

Frames are transmitted according to their sub-priority
(Urgency from a connection defines a sub-priority for the class)



UBS lowers the worst-case for urgent frames
(frames with higher priority are preferred in case of congestion)

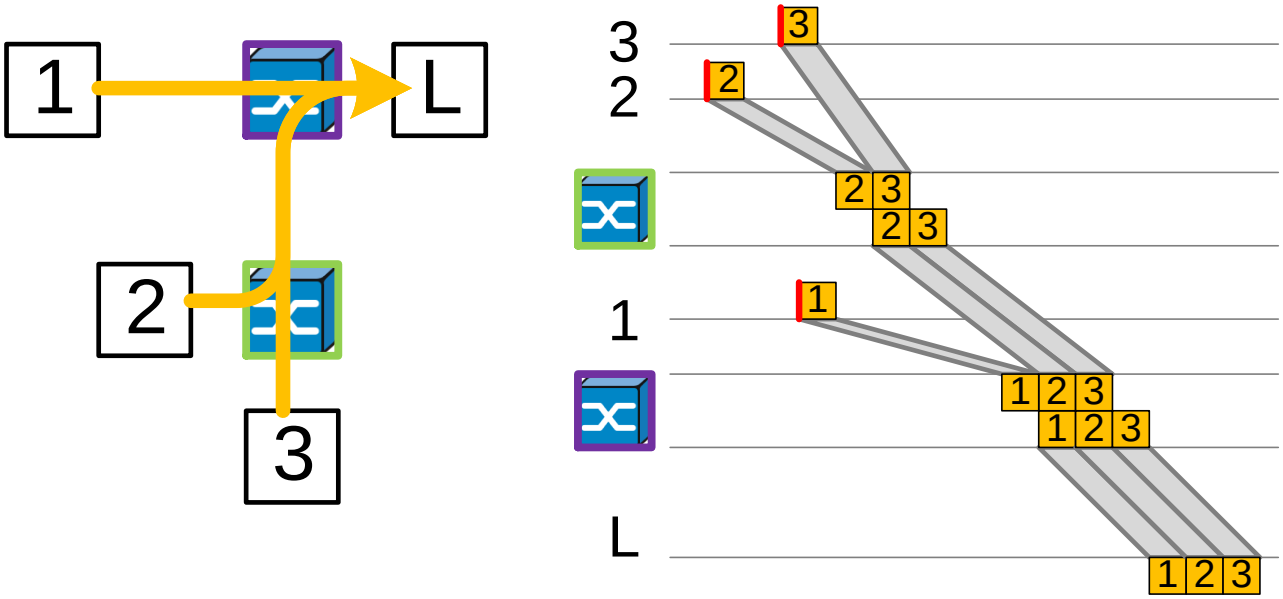
Improved UBS to cover automotive and industrial requirements
(e.g. Groups of Streams for scalability)
http://www.ieee802.org/1/files/public/docs2014/new-tsn-specht-ubs-status-update-0514-v01.pdf

In-Class Interference - Coordination

SIEMENS

- **Coordination** of transmit times for Streams

Frames are transmitted according to a planned schedule
(Interference can be avoided by shift of transmit time)

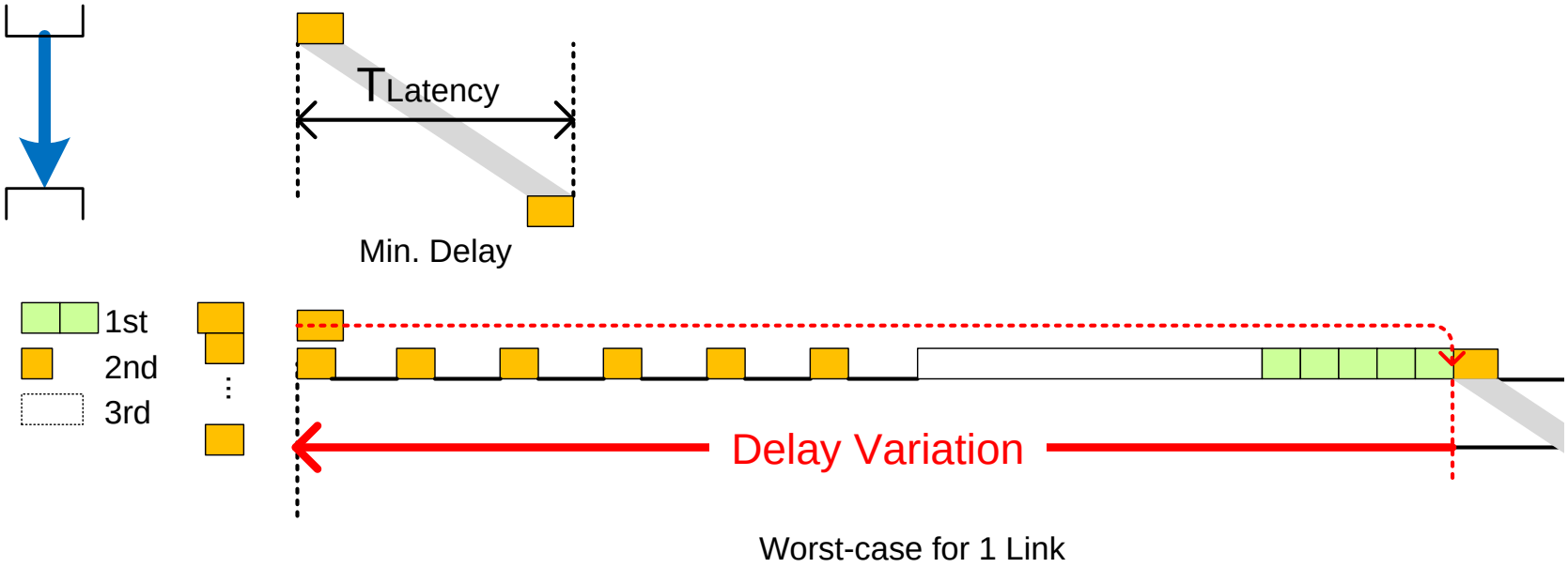


Can be combined with other mechanism

e.g. **TAS** can be used to avoid interference (see Priority Inversion)

Jitter - Delay Variation

Jitter is the variation of the Delay, End-to-End (E-2-E) Jitter or Delay must be accumulated along the complete path



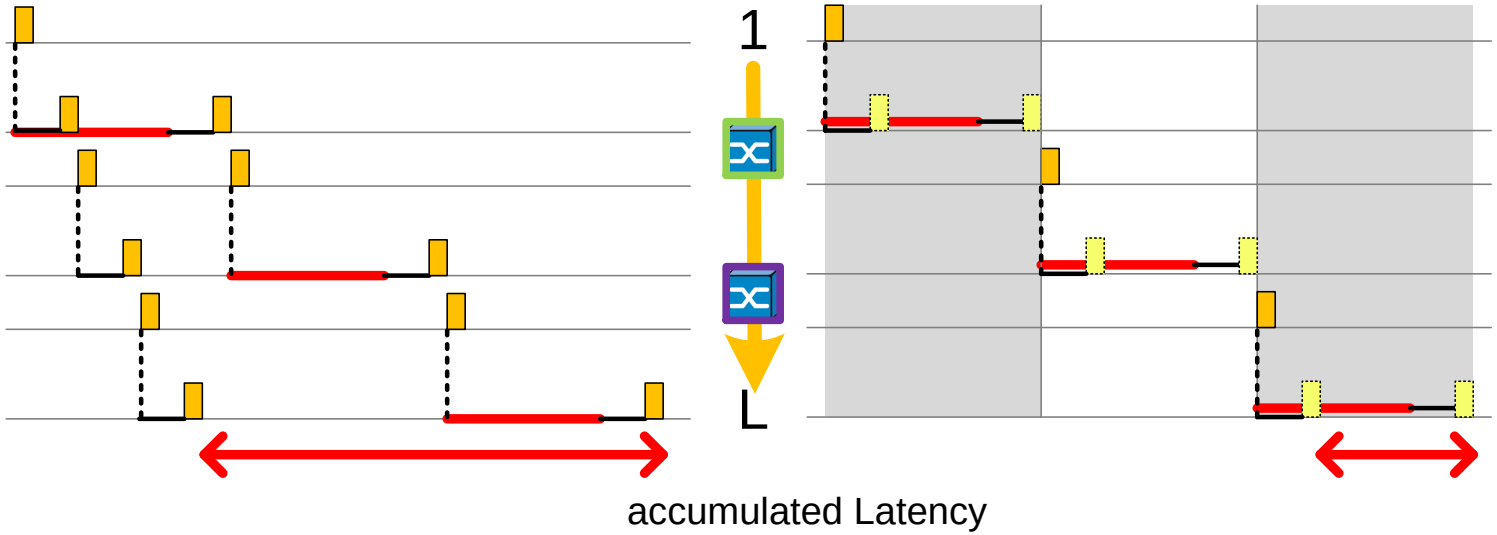
Applications must handle Latency between min. and max.

CB avoids loss of frames in case of a network failure (∞-Latency)

TAS avoids **/Preemption** limits Jitter from lower-class interference

E-2-E Jitter - SQF

Jitter can be lowered by increasing the min. Latency



Applications must handle a smaller possible arrival windows

SQF can be combined with other mechanisms

Outline

- Recap: Industrial Requirements for Latency
  - Different Application requirements regarding Latency
- Sources of Latency
  - Reasons for the special treatment of Streams
- Summary

Overview of Mechanisms

High Priority for lower latency

Reservation to avoid congestion causing loss of frames

CB can avoid loss of frames in case of a failure

Coordination to influence congestion and in-class interference

Preemption to lower Frame Interference
(improve bandwidth usage in case of TAS)

Shaper influence Delay and Jitter

Low Latency and Shapers are converse requirements

Policing for robustness in case of failures
(robustness in case of failures – no delaying in normal operation)

Overview of Shapers

Shaper influence Delay and Jitter:

- ~~***BLS***~~ *see policing – requirements considered in UBS*
- **UBS** for low latency with limited In-class interference (sub-priority dependent worst-case)
- **TAS** to avoid Jitter caused by Frame Interference
- **SQF** to minimize Jitter by increasing min. Latency *(replacement for CBSA?)* http://www.ieee802.org/1/files/public/docs2014/new-tsn-mjt-peristaltic-shaper-0114.pdf

Features	Field-Level	Control-Level	Plant-Level
Which predominant requirement?	**Closed Systems:** TAS with Coordination for **lowest latency** or **Open Systems:** **Flexibility** (with UBS)	**Flexibility** (with UBS)	**Flexibility** (with UBS) or SQF for **Low-Jitter** AV Streams

Overview of Mechanisms

Current Status:

TAS for closed systems (802.1 Qbv)

UBS and **SQF** are currently “homeless” (no PAR)

Policing … requirements from automotive and industrial automation for stability

Additional Information

Some Sources of Latency and Jitter

http://www.ieee802.org/1/files/public/docs2013/bv-kiessling-Some-sources-of-Latency-and-Jitter-0513-v01.pdf

How Many Transmission Selection Algorithms Do We Need?

http://www.ieee802.org/1/files/public/docs2013/new-tsn-boiger-new-shaper-0513.pdf

Why an additional Shaper?

http://www.ieee802.org/1/files/public/docs2013/new-tsn-kiessling-Why-an-additional-Shaper-0313-v3.pdf