Technische Universität Braunschweig

INSTITUTE OF COMPUTER AND NETWORK ENGINEERING

Worst-case design revisited

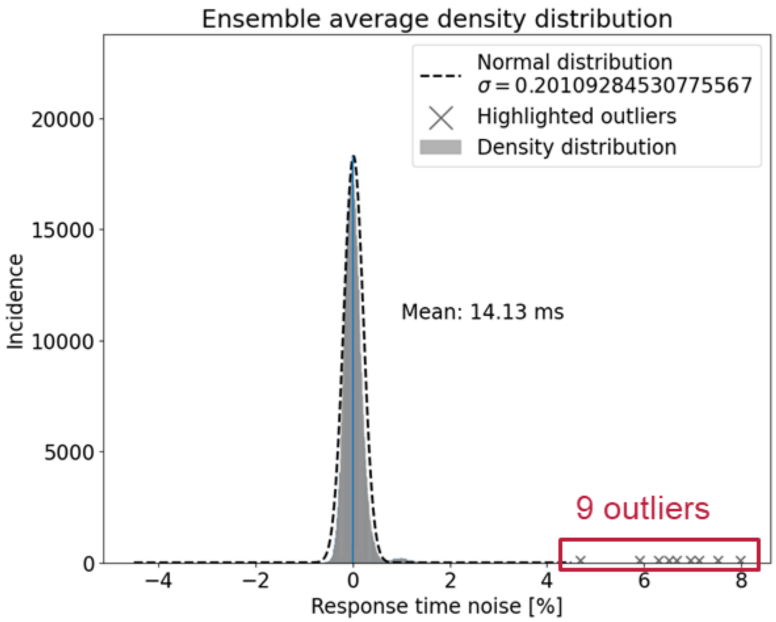
Rolf Ernst, TU Braunschweig

with support by Jonas Peeck, Nora Sperling, Daniel Tappe, Robin Hapka, and Alex Bendrick

MpSoC 2024: Worst-case execution timing - Wrap-up

- **NN on embedded HPC platforms**
  - no repeatable WCET
  - Gaussian execution time distribution + rare outliers
- **provided novel tight worst-case timing model**
  - elaborated worst-case design solutions
  - methods for rare deadline miss compensation
- **update 2025**
  - controller architecture for application in safety critical applications (avionics)
  - evaluation of further NN implementations

- nvidia ORIN AGX (2 cores + GPU)
- YOLOv5



This year: Design for worst-case communication timing

- **current approach**
  - packet-level worst-case timing guarantees (WCCT)
  - application data agnostic – unified design approach
  - example real-time Ethernet (TSN): Automotive profile 802.1 DG
  - example wireless networks: URLLC

- **alternative: application data object level guarantees**
  - worst-cast timing guarantees at data object level

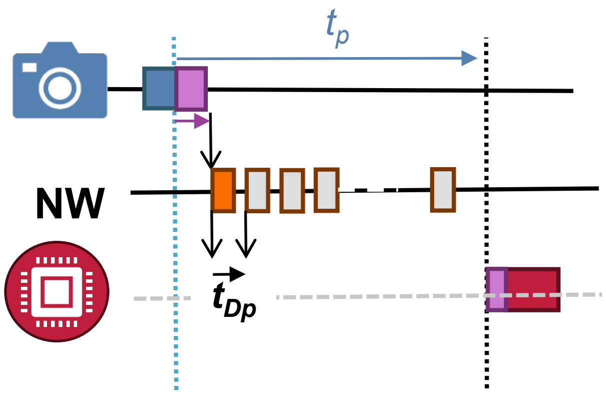
Example 1: Automotive profile 802.1 DG

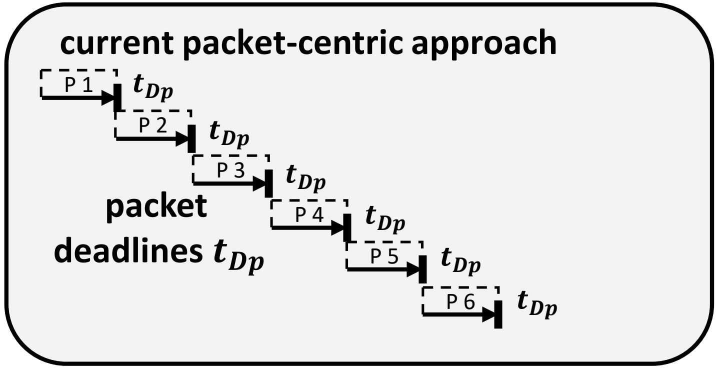
- **profile addresses packet (aka frame) transfer**
  - **sensor packet sequence defined as packet streams**
  - **→ packet-specific deadlines t_{Dp}**
  - ***→ high timing requirements***

Traffic Type	Period	Guarantee[4]	Tolerance to Loss[5]	Frame Size	Criticality
Safety-relevant Control: see 3.4.1.2	<= 20ms	Deadline based Reserved w/Latency < 1ms	No	64 bytes	High
Safety-relevant Media: see 3.4.1.3	<= 10ms	Bandwidth based Reserved w/Latency < 1ms	No	64 to max frame size[6] (w/1500 data bytes)	High
Network Control:	50ms to 1s	Sporadic Highest priority	Yes	64 to 512[7]	High

Automotive Profile in PANNELL, Don, et al. Use Cases-IEEE P802. 1DG V0. 4 . Accessed: Aug 5, 2024.

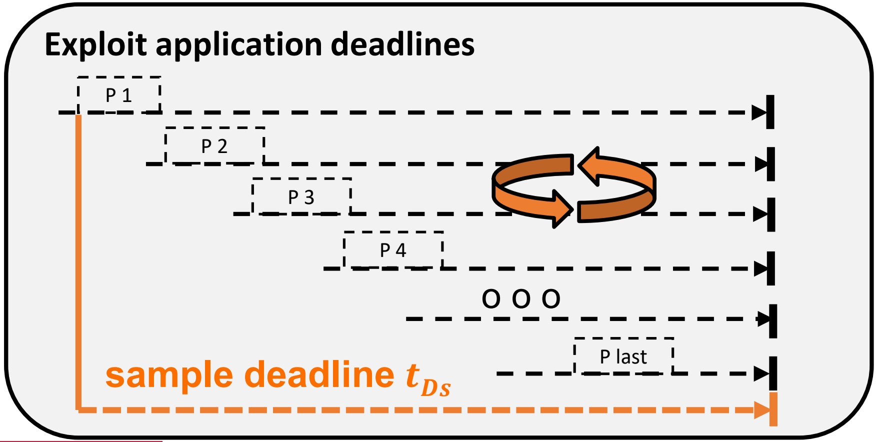
t_{Dp} < 1ms; t_p <= 10ms

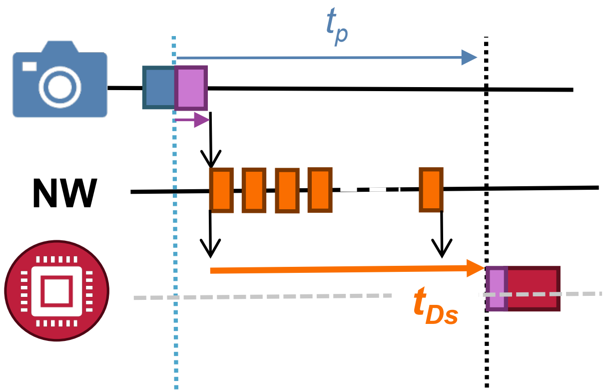




Large real-time data objects – the application perspective

- manage **application data objects, not packets!**
  - sample-specific deadlines t_{Ds}
  - $t_{Ds} >> t_{Dp}$
  - leads to extra **communication slack**
  - packet-level deadline asymmetry



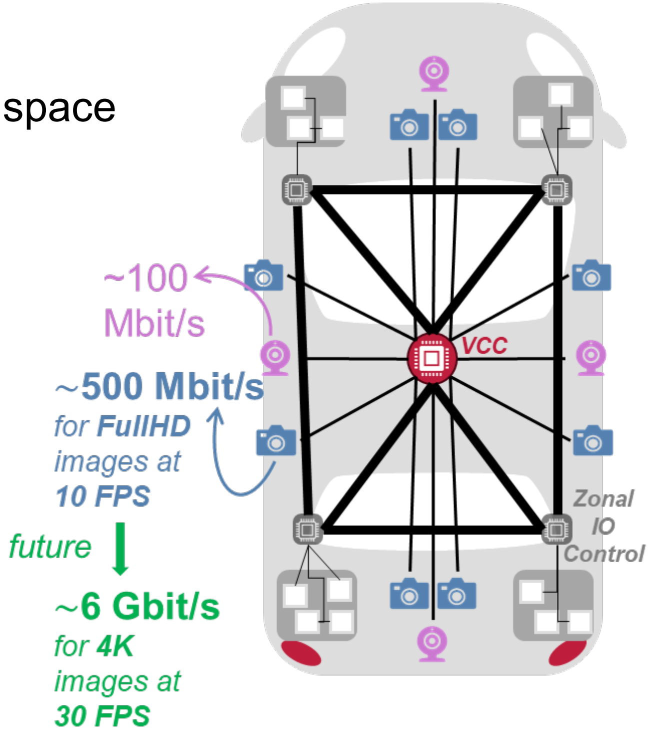


Application to vehicle TSN

- **mobile systems increase sensor resolution**
  - radars, cameras, LIDARs – growing resolution in time and space
  - high resolution 3D real-time maps

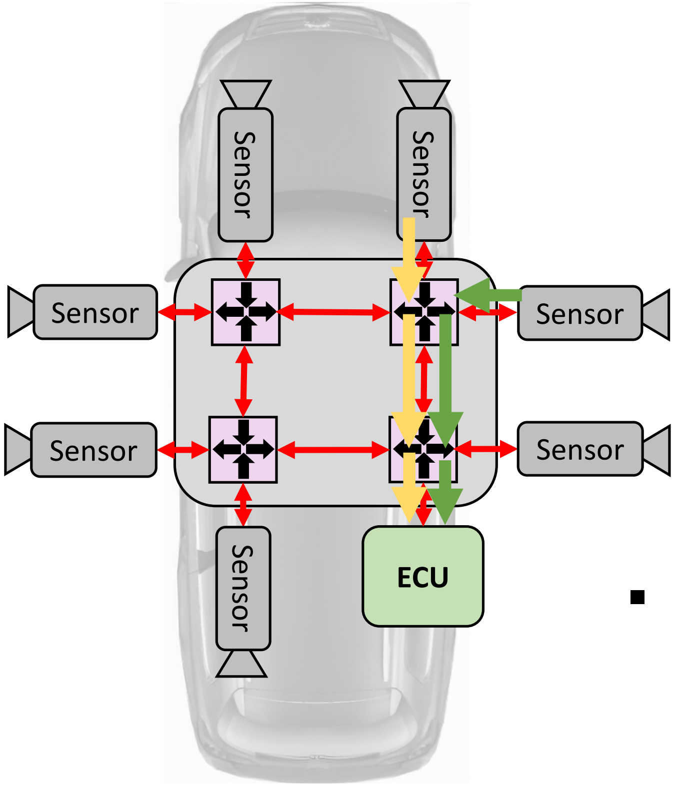
> *"Advanced driver assistance systems (ADAS) and autonomous driving use cameras for all-around visibility of the vehicle's surroundings.*
> *Each camera sends about* ***500-3,500 Mbps*** *of video frame data."*
> *Hwee Yng Yeo, Keysight Technologies*

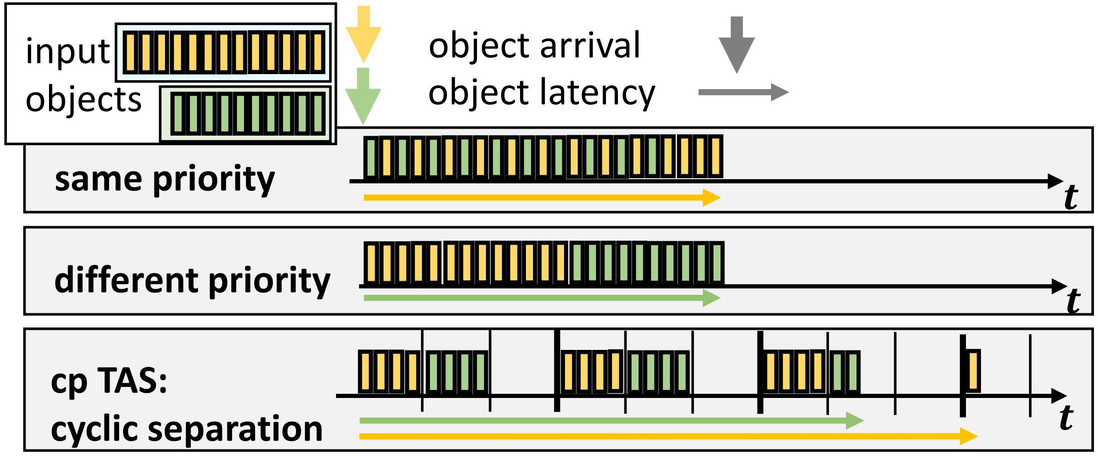
- **vehicle: large share of critical real-time data**
  - higher levels of driving automation
  - reduced vehicle reaction time



zonal vehicle architecture

In-vehicle camera stream coordination





- **SoA standard: TSN packet level scheduling**
  - latency always adds up

Limitation: No long reservation intervals

Traffic Type	Period	Guarantee[4]	Tolerance to Loss[5]	Frame Size	Criticality
Safety-relevant Control: see 3.4.1.2	<= 20ms	Deadline based Reserved w/Latency < 1ms	No	64 bytes	High
Safety-relevant Media: see 3.4.1.3	<= 10ms	Bandwidth based Reserved w/Latency < 1ms	No	64 to max frame size[6] (w/1500 data bytes)	High
Network Control:	50ms to 1s	Sporadic Highest priority	Yes	64 to 512[7]	High

Automotive Profile in PANNELL, Don, et al. Use Cases-IEEE P802. 1DG V0. 4 .
Accessed: Aug 5, 2024.

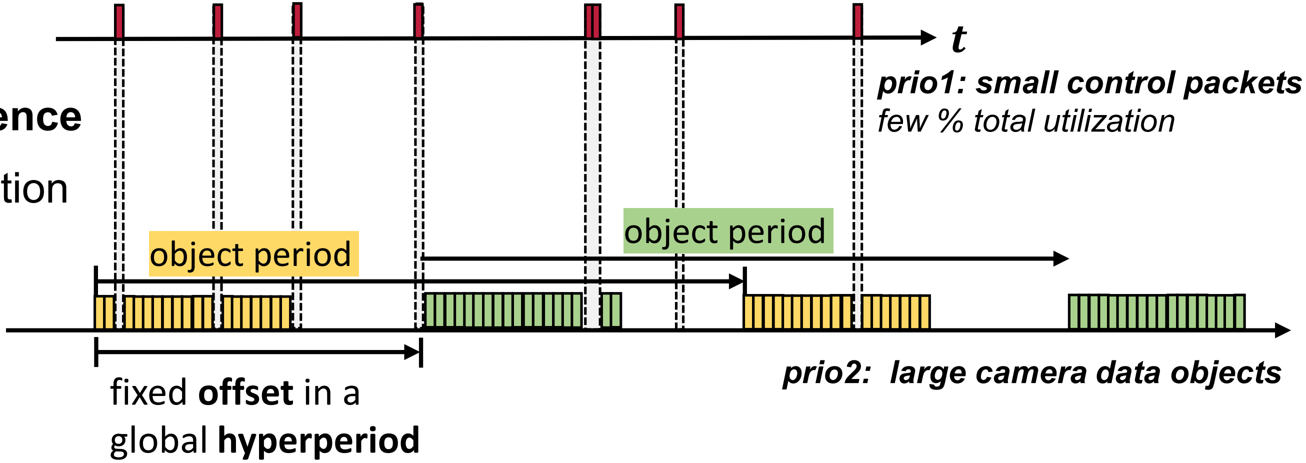
- **safety messages (control packets) are latency constrained (1 ms)**
  - **time triggered shaping (TAS) insufficient for combination of profiles**

Coordinated burst transfer → Mitigate Stream Interference

- **schedule application data objects – rather than individual packets**
  - synchronize network access with fixed object offset
  - no shaping in switches - IEEE 802.1Q prioritization only (resync in larger networks)
  - simple network control

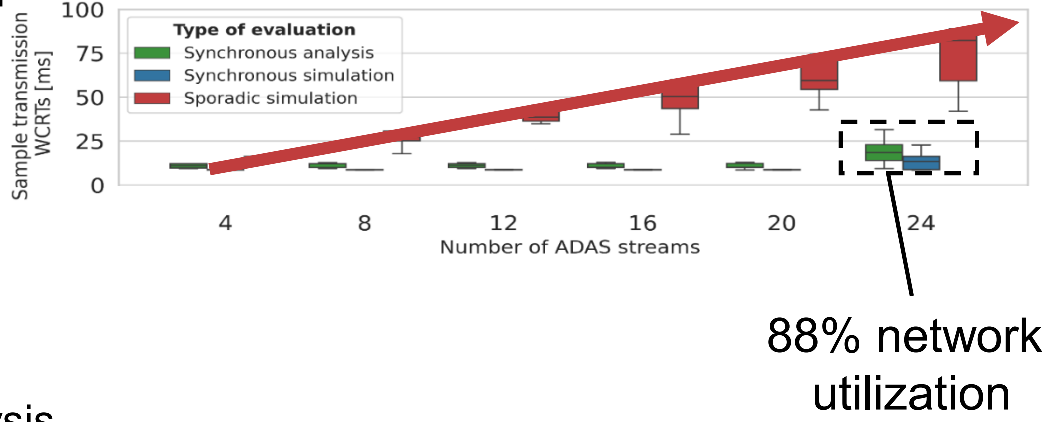
→no data object interference

  - short object communication latency
  - accurate worst case analysis



Results: Improvements of synchronized object communication

- **non-coordinated communication**
  - → worst-case object latency grows with utilization



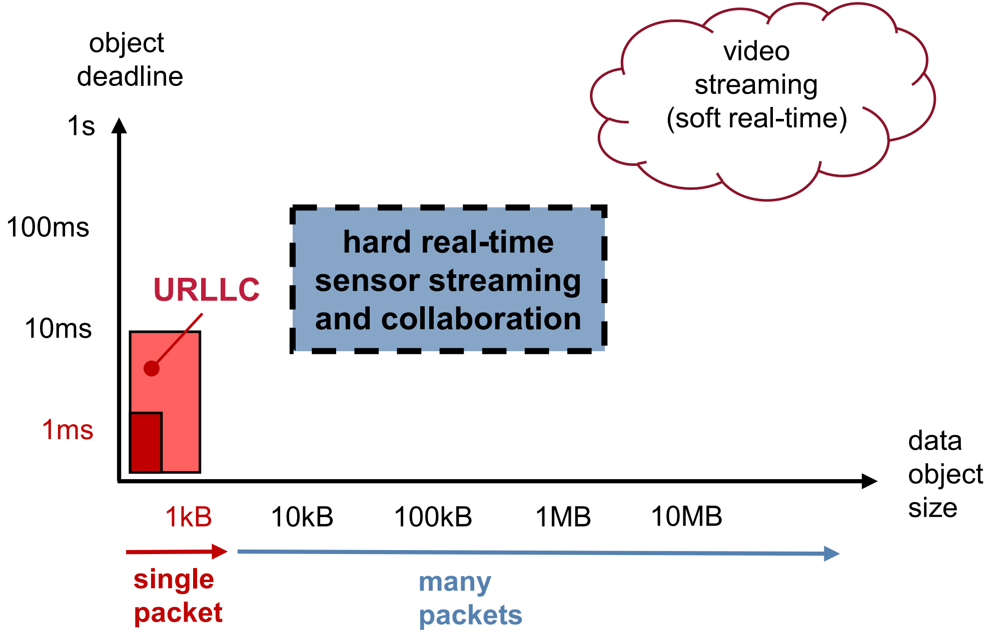
- **synchronized application bursts**
  - → interference mitigated
    - minimum latency
    - enables efficient object WCCT analysis

↔ **time triggered object transmission not sufficient**

- combination with prioritization essential

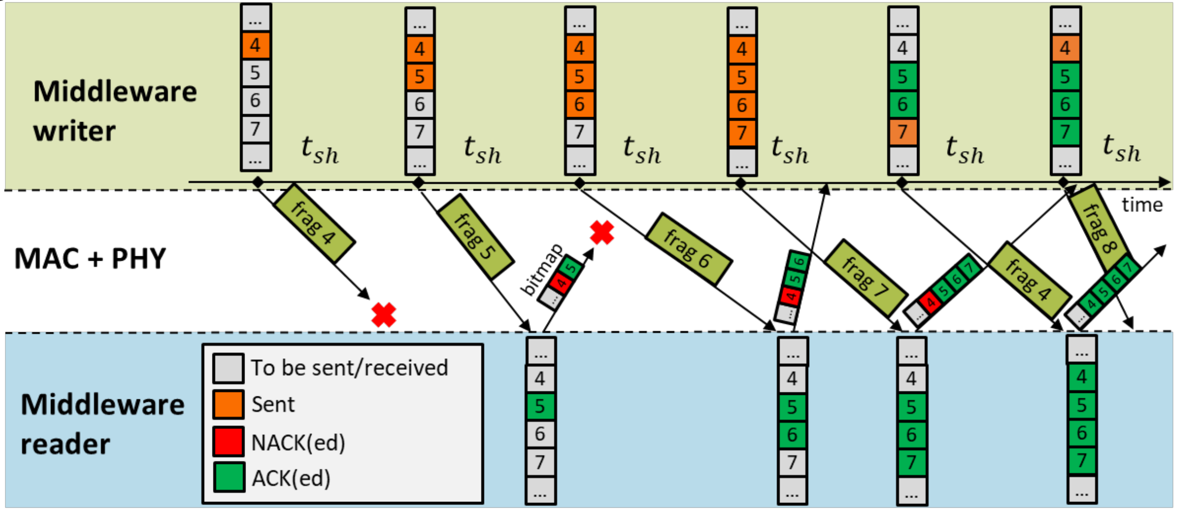
Example 2: Reliable wireless low-latency communication

- **SoA: robust low latency guarantees for single packets**
- **robust solutions for low-latency large object streaming still missing**
  - vehicle teleoperation
  - collaborative sensing and perception
- **utilize object slack for backward error correction (BEC)**
  - higher reliability than (bounded) MAC-layer retransmissions
  - lower overhead than FEC
  - deadline asymmetry helps error statistics



Wireless Reliable Real-Time Protocol (W2RP) – key concepts

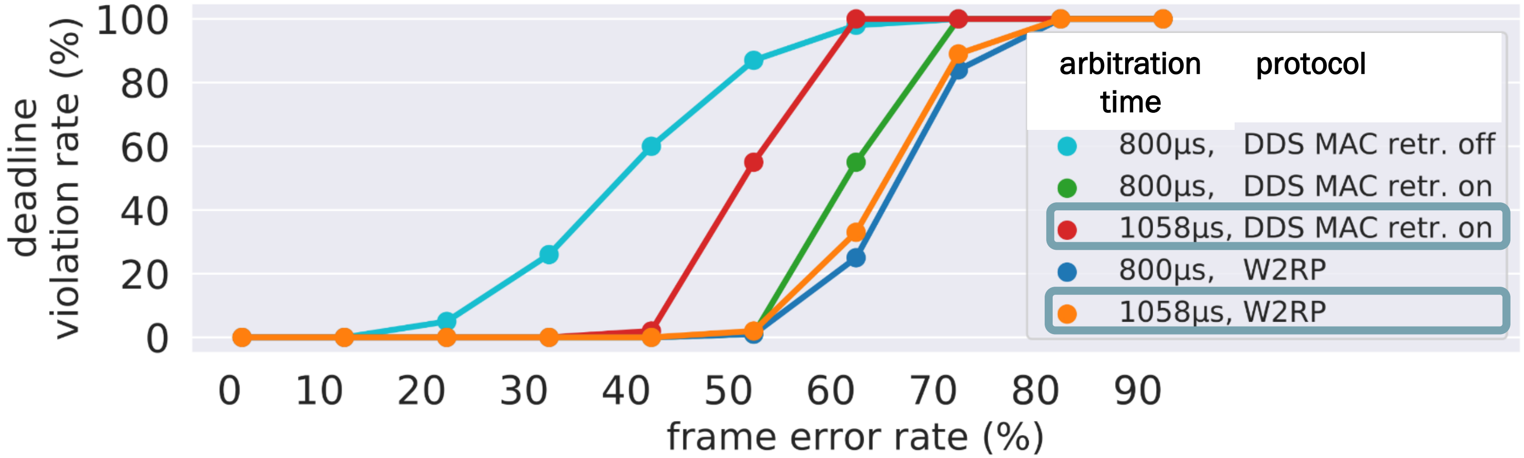
- **2-phase transmission scheme**
  - ph1: object transmissions
  - ph2: bitmap-based retransmissions
- **media access shaping**
  - avoid growing round-trip times
  - resource management with arbitration time adaptation
- **extensions**
  - multicast
  - overlapping object transmisions
  - include MCS management, …



W2RP efficiency

- **violation rate comparison with SoA protocols**
  - **with DDS middleware and MAC layer retransmission**
  - **for different access shaping (arbitration time)**

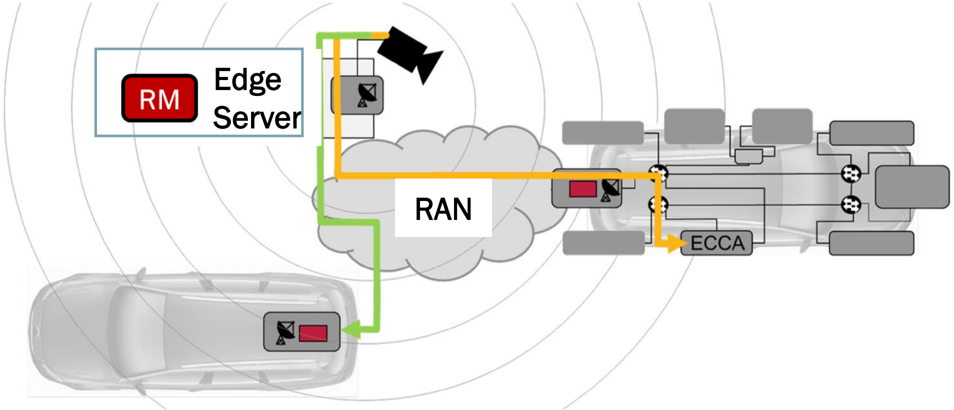
IEEE 802.11p
object size 20kB
period 100ms
deadline 100ms

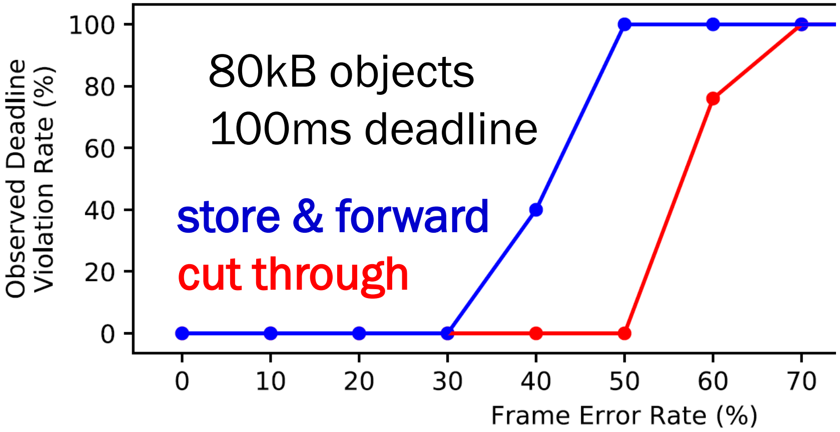


no-loss target: 0% violation rate

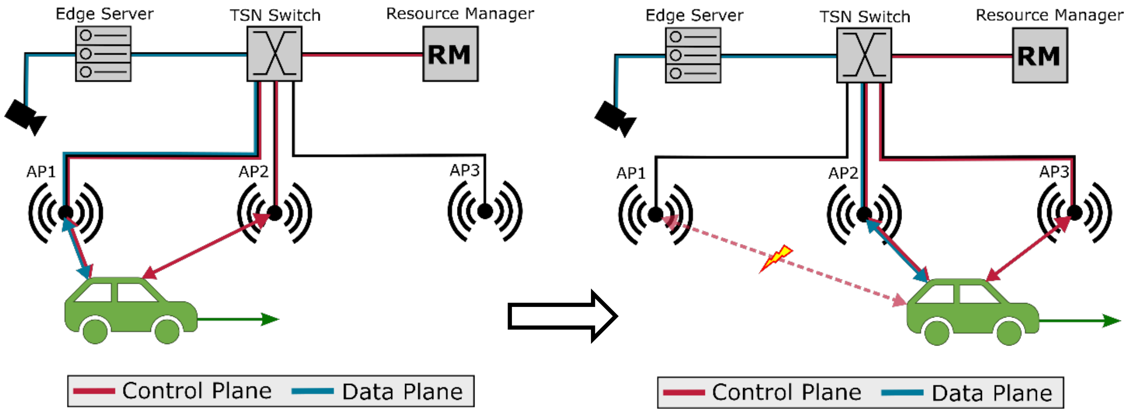
End-to-end handling of large data objects – TSN + W2RP (802.11p)

- e2e transfer of camera stream to sensor fusion unit (ECCA)
- W2RP + TSN with store&forward vs. cut through in Telemetry Unit
- wireless parameter control in hierarchical Resource Management (RM)

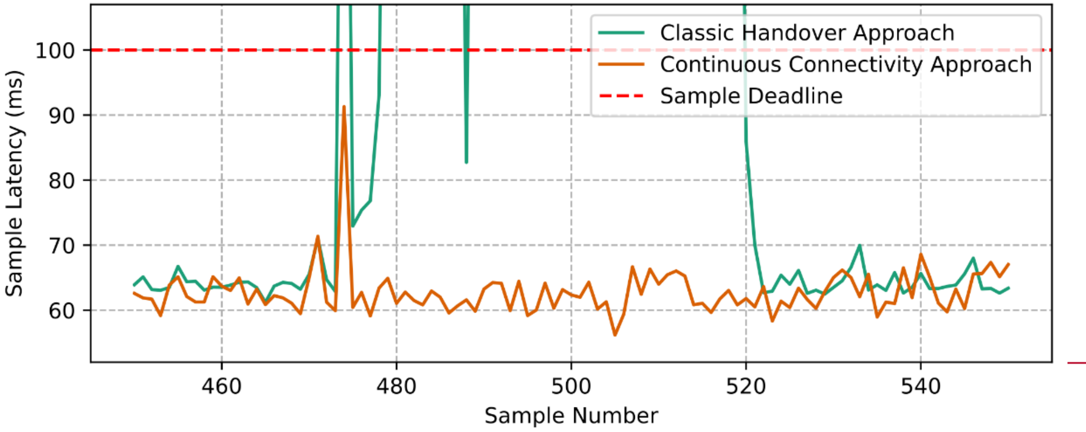




Object slack and W2RP for continuous connectivity



- **multi-link control plane**
- **single link data plane**
  - fast re-route with dynamic TSN re-conf.
  - fast repetition of lost packets via W2RP



Conclusion

- **worst-case design of real-time systems revisited**
  - **computation**
    - from logical worst-case instruction sequences to platfrom design with physical and variance effects
  - **communication**
    - from individual packets to application data object stream communication
    - from network segments to application driven end-to-end management

Some related papers

- TSN Synchronized Burst Transfer and Reconfiguration
  - Jonas Peeck and Rolf Ernst. "Improving Worst-case TSN Communication Times of Large Sensor Data Samples by Exploiting Synchronization." EMSOFT 2023 and ACM Transactions on Embedded Computing Systems.
  - Kostrzewa, Adam, and Rolf Ernst. "Achieving safety and performance with reconfiguration protocol for Ethernet TSN in automotive systems." Journal of Systems Architecture 118 (2021): 102208.
  - Kostrzewa, Adam, and Rolf Ernst. "Fast failover in Ethernet-based automotive networks." 2020 IEEE 23rd International Symposium on Real-Time Distributed Computing (ISORC). IEEE, 2020.
  - Dominik Stöhrmann, and Rolf Ernst. "Fast Vehicular TSN Network Reconfiguration with Application Aware Network Synchronization" ETFA 2024.
- W2RP and Extensions
  - Alex Bendrick, Daniel Tappe, Dominik Stohrmann and Rolf Ernst, "Synchronized Lossfree Reconfiguration of Safety-critical V2X Streaming Applications", *IEEE Transactions on Vehicular Technology*, pp. 1-16, 2024. Daniel Tappe, Alex Bendrick and Rolf Ernst, "Continuous multi-access communication for high-resolution low-latency V2X sensor streaming", In *2024 IEEE Intelligent Vehicles Symposium (IV)*, 2024, June.
  - Alex Bendrick, Daniel Tappe and Rolf Ernst, "Ultra Reliable Hard Real-Time V2X Streaming with Shared Slack Budgeting", In *2024 IEEE Intelligent Vehicles Symposium (IV)*, 2024, June.
  - Rolf Ernst, Dominik Stöhrmann, Alex Bendrick and Adam Kostrzewa, "Application-centric Network Management - Addressing Safety and Real-time in V2X Applications", *ACM Transactions on Embedded Computing Systems (Perspective Paper)*, vol. 22, no. 2, pp. 20:1–20:25, 2023, January.
  - Jonas Peeck and Rolf Ernst, "Enabling multi-link data transmission for collaborative sensing in open road scenarios", In *Proceedings of the IEEE 31st International Conference on Real-Time Networks and Systems*, pp. 76–86, 2023, June.
  - Alex Bendrick, Jonas Peeck and Rolf Ernst, "An Error Protection Protocol for the Multicast Transmission of Data Samples in V2X Applications", *ACM Transactions on Cyber-Physical Systems*, 2023, August.
  - Alex Bendrick and Rolf Ernst, "Hard Real-Time Streaming of Large Data Objects with Overlapping Backward Error Correction", In *IECON 2023-49th Annual Conference of the IEEE Industrial Electronics Society*, pp. 1-8, 2023, October.
  - Jonas Peeck, Mischa Möstl, Tasuku Ishigooka and Rolf Ernst, "A Protocol for Reliable Real-Time Wireless Communication of Large Data Samples", *IEEE Transactions on Vehicular Technology*, vol. 72, no. 10, pp. 13146–13161, 2023, October.
  - Jonas Peeck, Mischa Möstl, Tasuku Ishigooka and Rolf Ernst, "A Middleware Protocol for Time-Critical Wireless Communication of Large Data Samples", In *2021 IEEE Real-Time Systems Symposium (RTSS)*, pp. 1–13, 2021, December.
  - D. Tappe, A. Bendrick and R. Ernst, "Continuous multi-access communication for high-resolution low-latency V2X sensor streaming," 2024 IEEE Intelligent Vehicles Symposium (IV), Jeju Island, Korea.
  - A. Bendrick et al., "Teleoperation as a Step Towards Fully Autonomous Systems," 2025 Design, Automation & Test in Europe Conference (DATE), Lyon, France, 2025, pp. 1-8

Related Projects

- **funded by German DFG**
  - **IPF 2.0: https://www.ida.ing.tu-bs.de/en/forschung/ipf-20**
  - **LOTUS: https://ida-tubs.github.io/lotus/**
- **funded by German BMWE**
  - **MC-ADAMS: https://www.ida.ing.tu-bs.de/forschung/projekte/mc-adams**