Technische Universität Braunschweig

INSTITUTE OF COMPUTER AND NETWORK ENGINEERING

Managing large dynamic real-time data in vehicles and beyond

Rolf Ernst, TU Braunschweig

with contributions from N. Sperling, J. Peeck, A. Bendrick, D. Tappe, D. Stöhrmann

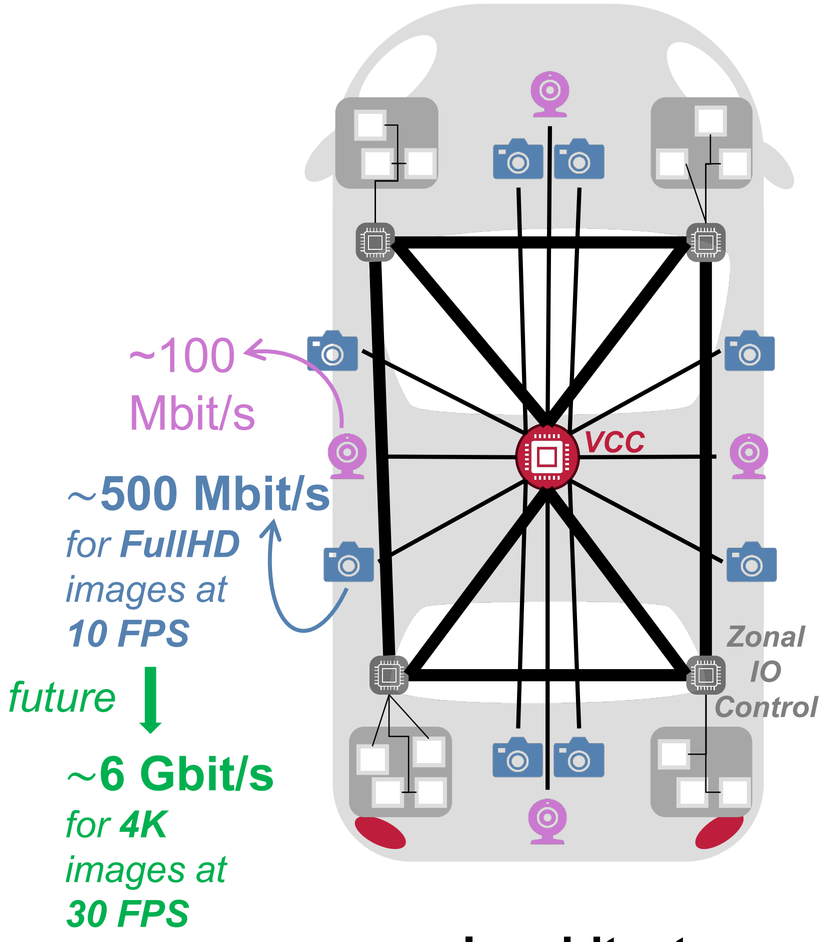
Motivation: growing critical data volume in a vehicle

- **higher vehicle sensor resolution in space and time**
  - radars, cameras, LIDARs
  - high resolution 3D real-time maps
  - vehicle application coordination via V2X

> *“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*

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

→ **pressure on communication and data management**
  - growing resolution → **higher data volume per sample**
  - reduced reaction time → **increased sample rate**



zonal architecture

Why increased sample rate?

- **smaller region of interest for motion detection**

- **perception latency depends on sample rate**

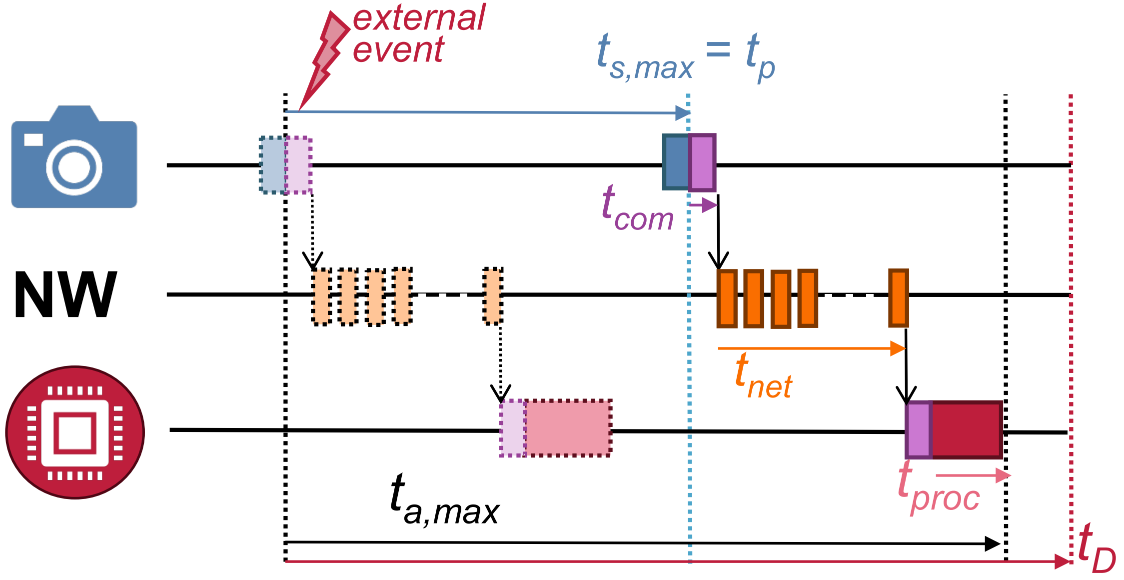
$$t_{a,max} = t_{s,max} + t_{net} + t_{com} + t_{proc}$$

  - typical perception deadline for high automation

$$t_D < 100\ ms$$

→ implement high sample rates with perception pipelining $t_p < t_D$

↔ in conflict with growing t_{net}



t_p = sampling period
$t_{s,max}$ = max. sampling delay
t_{net} = transport delay
t_{com} = com. stack delay
$t_{a,max}$ = max. input age
t_D = perception deadline
t_{proc} = processing time

Outline

- **Managing large real-time data objects – an application-based approach**
- From packet to application object deadlines
  - Application 1: in-vehicle stream coordination
  - Application 2: at run-time network re-configuration
  - Application 3: reliable wireless communication
- Application data caching for dynamic volume reduction
- Conclusion

Managing large real-time data objects – the communication mismatch

- **communication addresses packet (aka frame) transfer**
  - sensor packet sequence defined as **packet streams**

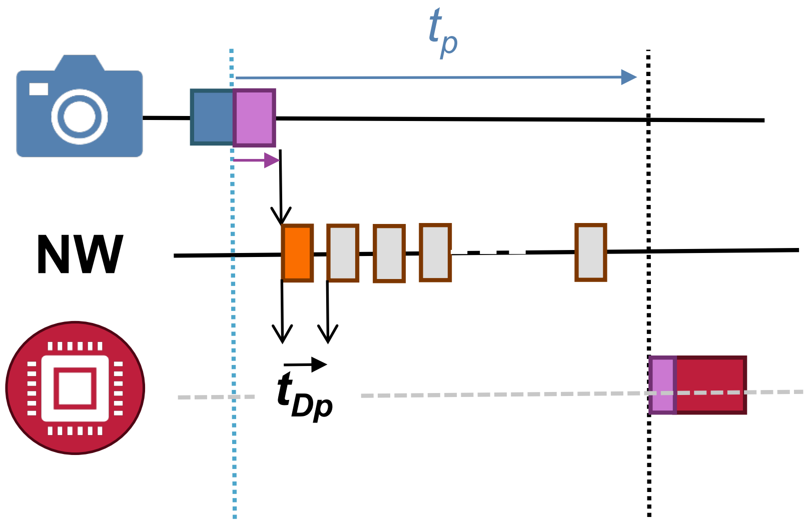
  → **packet-specific** deadlines t_{Dp}

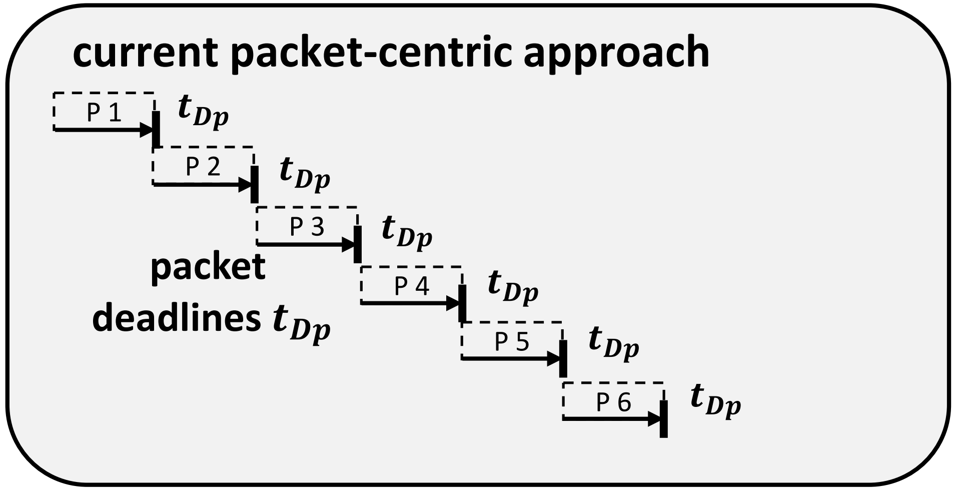
  → ***high timing requirements***

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.

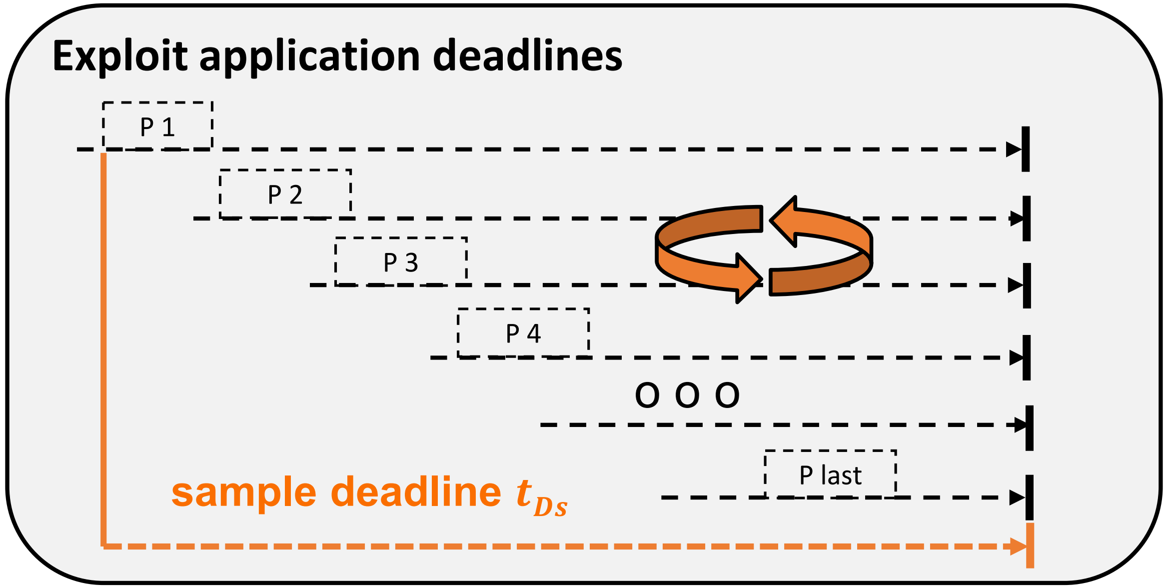
t_{Dp} < 1ms; t_p <= 10ms

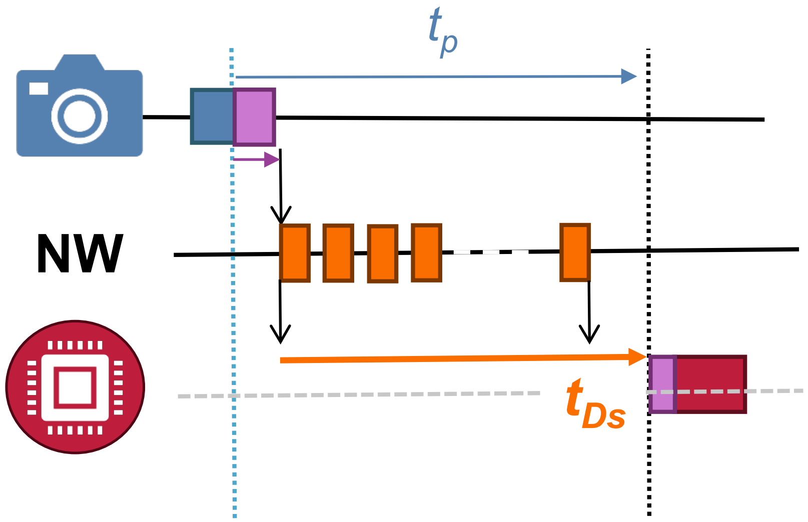




Managing large real-time data objects – the application perspective

- **transmit application objects, no packets!**
  - **sample-specific** deadlines t_{Ds}
  - $t_{Ds} >> t_{Dp}$
  - leads to extra communication slack
  - packet-level deadline asymmetry

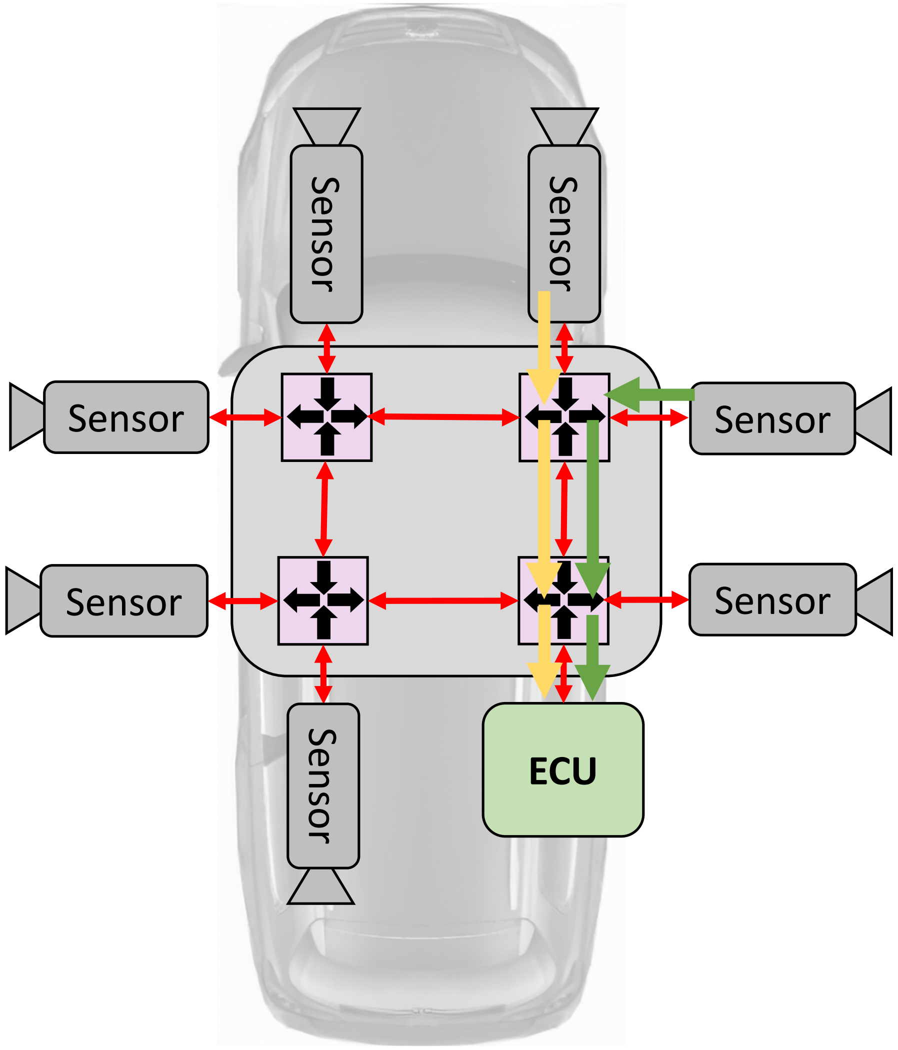


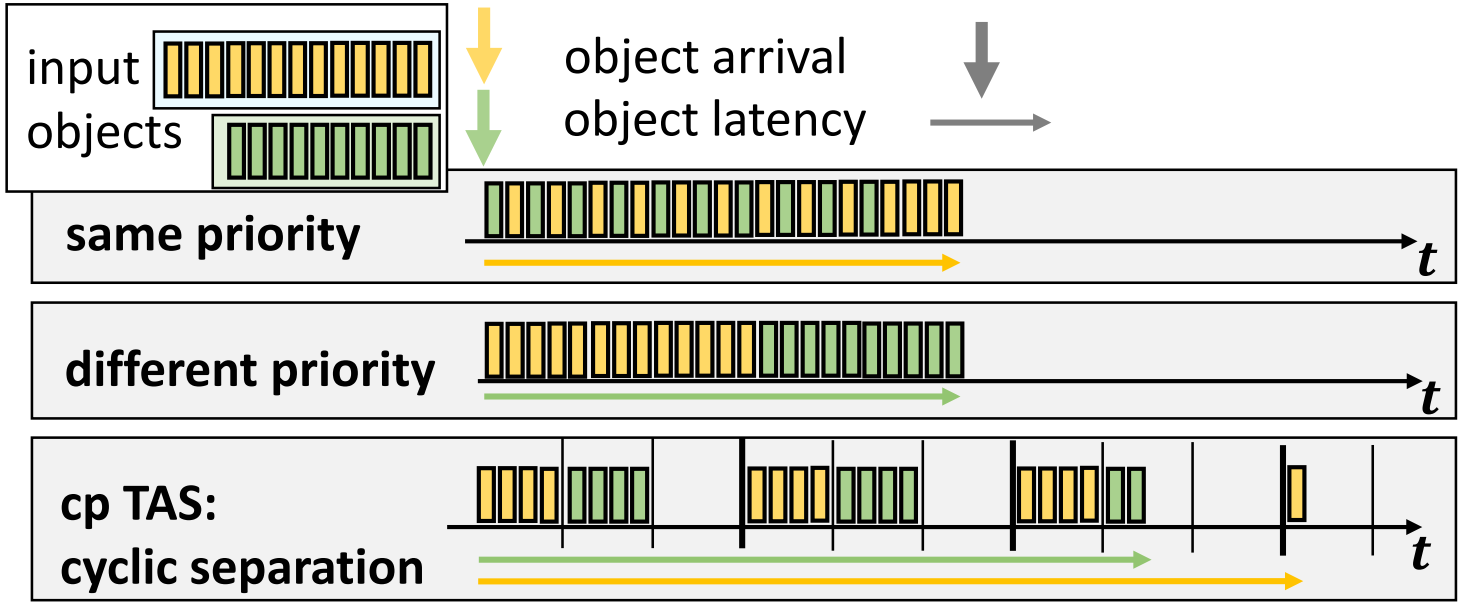


Outline

- Managing large real-time data objects – an application-based approach
- **From packet to application object deadlines**
  - **Application 1: in-vehicle stream coordination**
  - **Application 2: at run-time network re-configuration**
  - **Application 3: reliable wireless communication**
- Application data caching for dynamic volume reduction
- Conclusion

Application 1: In-vehicle camera stream coordination





- **SoA standard: Time Sensitive Network (TSN)**
- **TSN packet level scheduling**: latency always adds up

Coordinated burst transfer

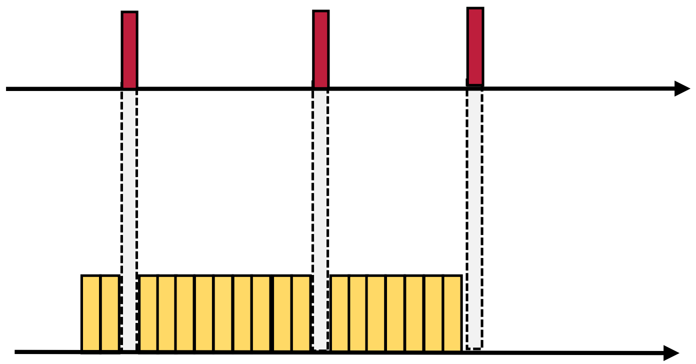
- **burst coordination challenge: *don't mess with critical short deadlines***

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

small control packets, t_{DP} < 1ms

camera samples

- **combine static prioritization with network access synchronization**



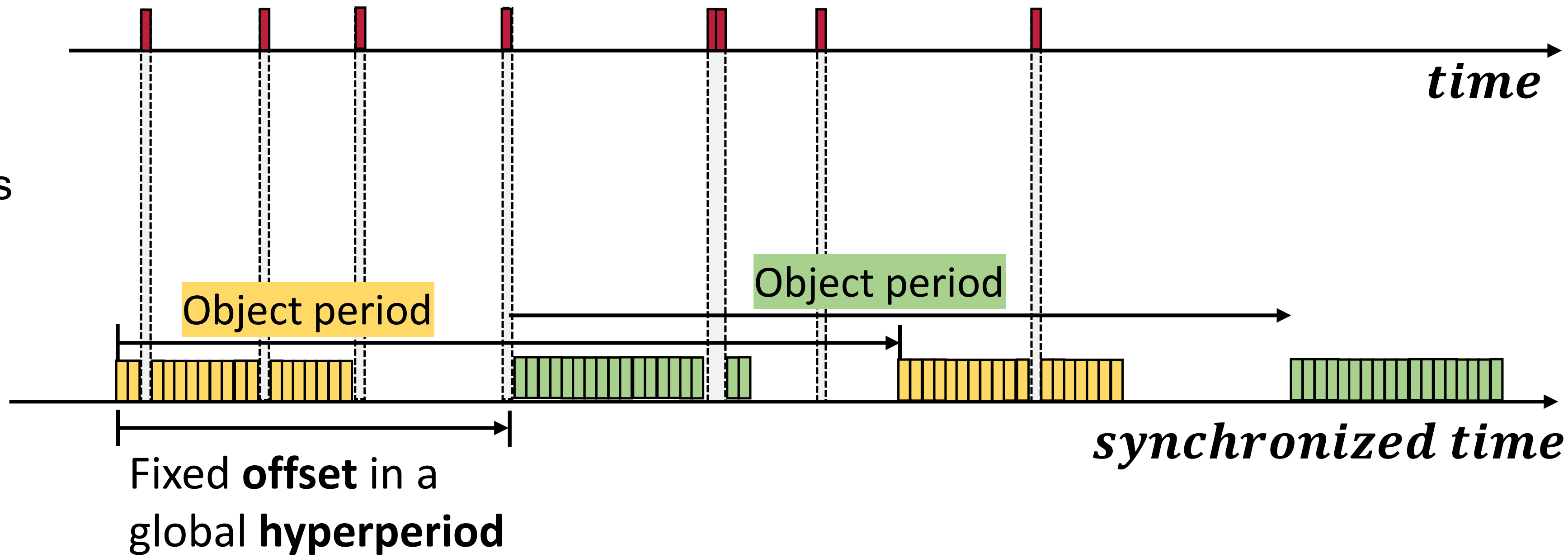
priority 1: small control packets
few % total utilization

priority 2: Large camera data objects
network access synchronization

Coordinated burst transfer → Mitigate Stream Interference

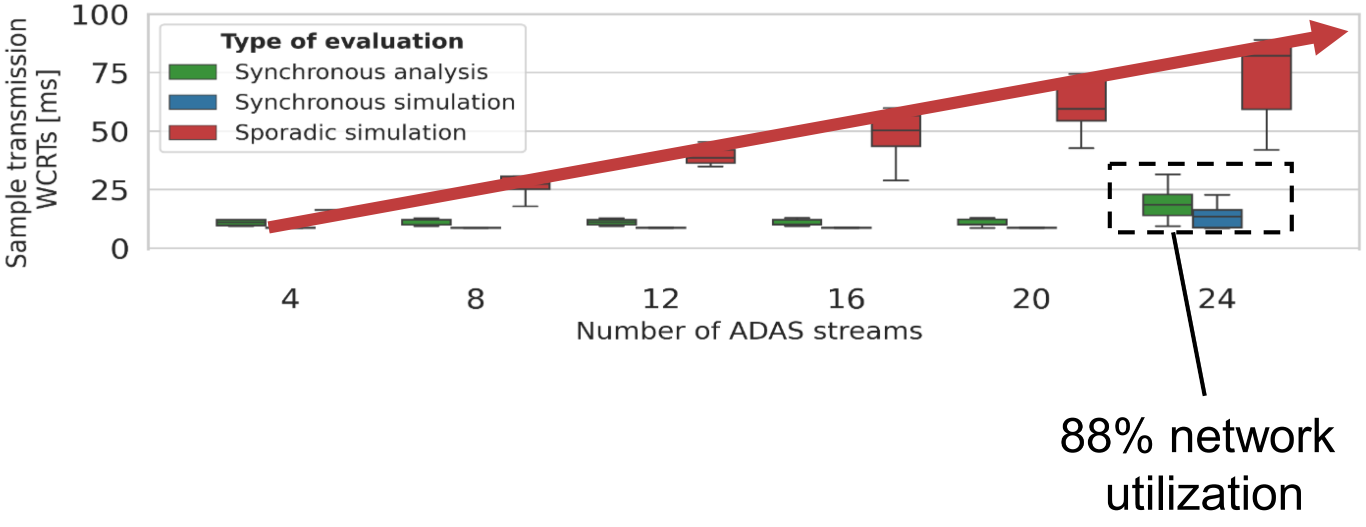
- **schedule application data objects – rather than individual packets**
  - synchronize **network access** with **fixed object offset**
  - **no shaping** in switches - IEEE 802.1Q prioritization only (possibly resynchronization in larger network)
  - simple network control

→ **no data object interference**
- short object latency
- accurate worst case analysis



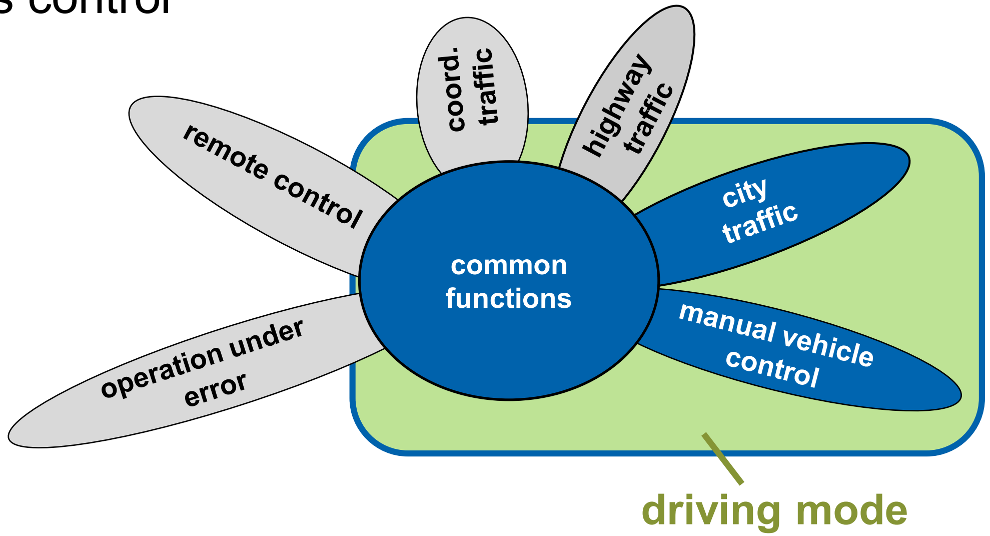
Results: The improvements of synchronous communication

- **non-coordinated communication**
  - → **worst-case object latency grows with utilization**
- **synchronized application bursts**
  - → **interference mitigated**
  - minimum latency
  - enables **efficient WCRT analysis**
- **time triggered object transmission not sufficient**
  - combination with prioritization essential

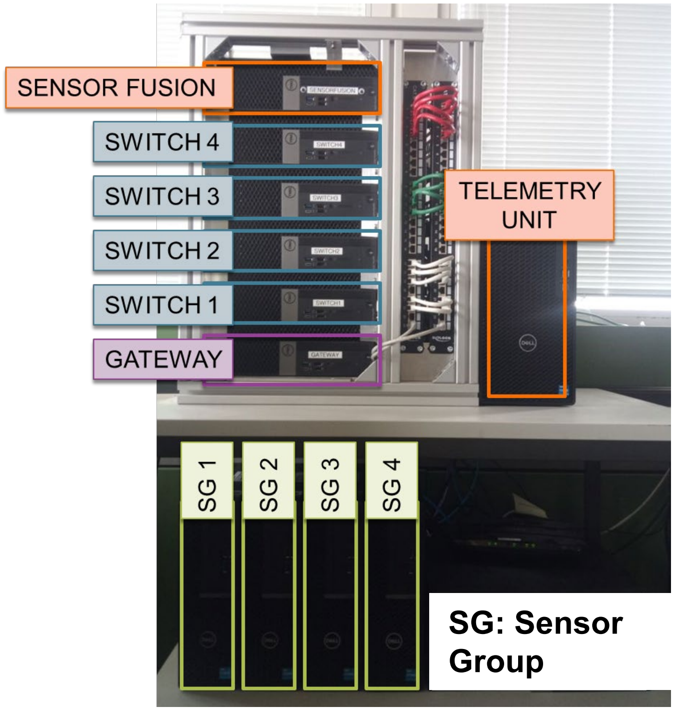


Application 2: At run-time network re-configuration

- **motivation: dynamic in-vehicle network management**
  - application and situation aware network adaptation
  - can be applied to subset of network nodes, parameters, and applications
  - more flexible design process – individual planning of situations and applications
- **requires safe re-configuration protocol**
  - *packet based* ***SDN too slow***
  - **exploit application data object deadline**
  - again: use network access control to avoid losses
  - time synchronized mode transition minimizes delay

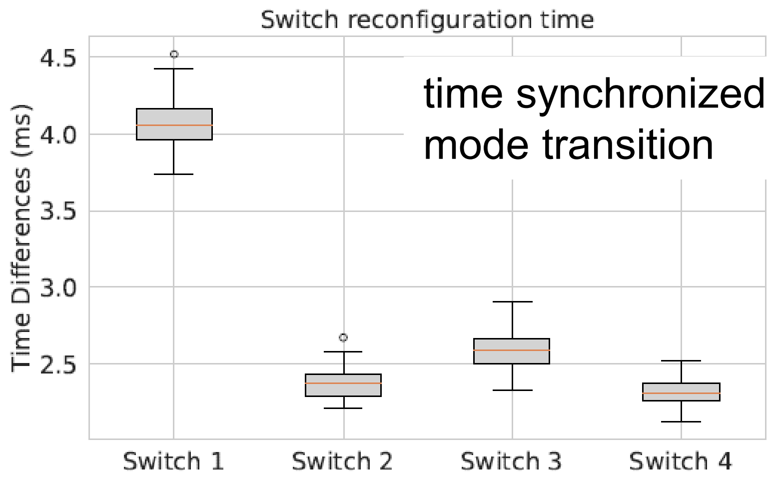


experimental platform



Coordinated switch re-configuration for time-synchronized mode transition

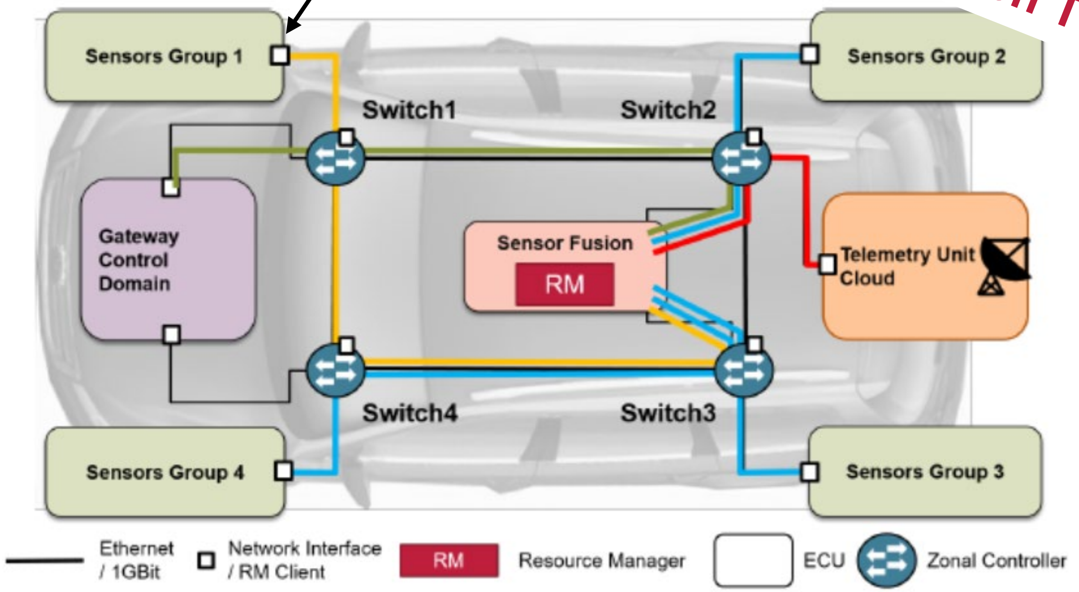
- **mode 1:**
  - red stream *disabled*
  - yellow stream routed over SW1→ SW2
- **mode 2:**
  - red stream *enabled*
  - yellow stream routed over SW1 → SW4 → SW3



time synchronized mode transition

paper on Friday, session T11.1

hold NW access during transition



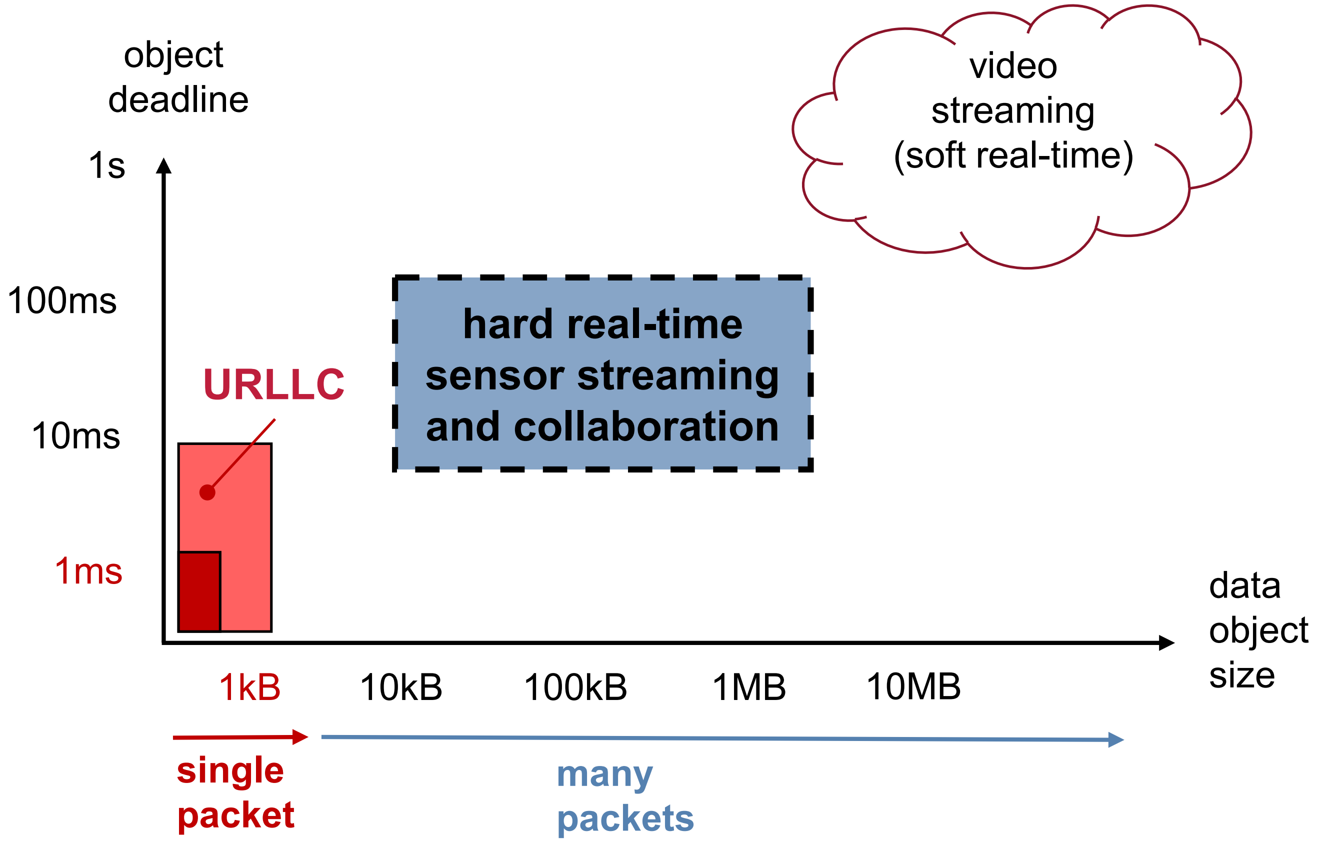
Switch1 … Switch4: Soft switches (Linux)

experimental platform configuration

Application 3: Reliable wireless communication

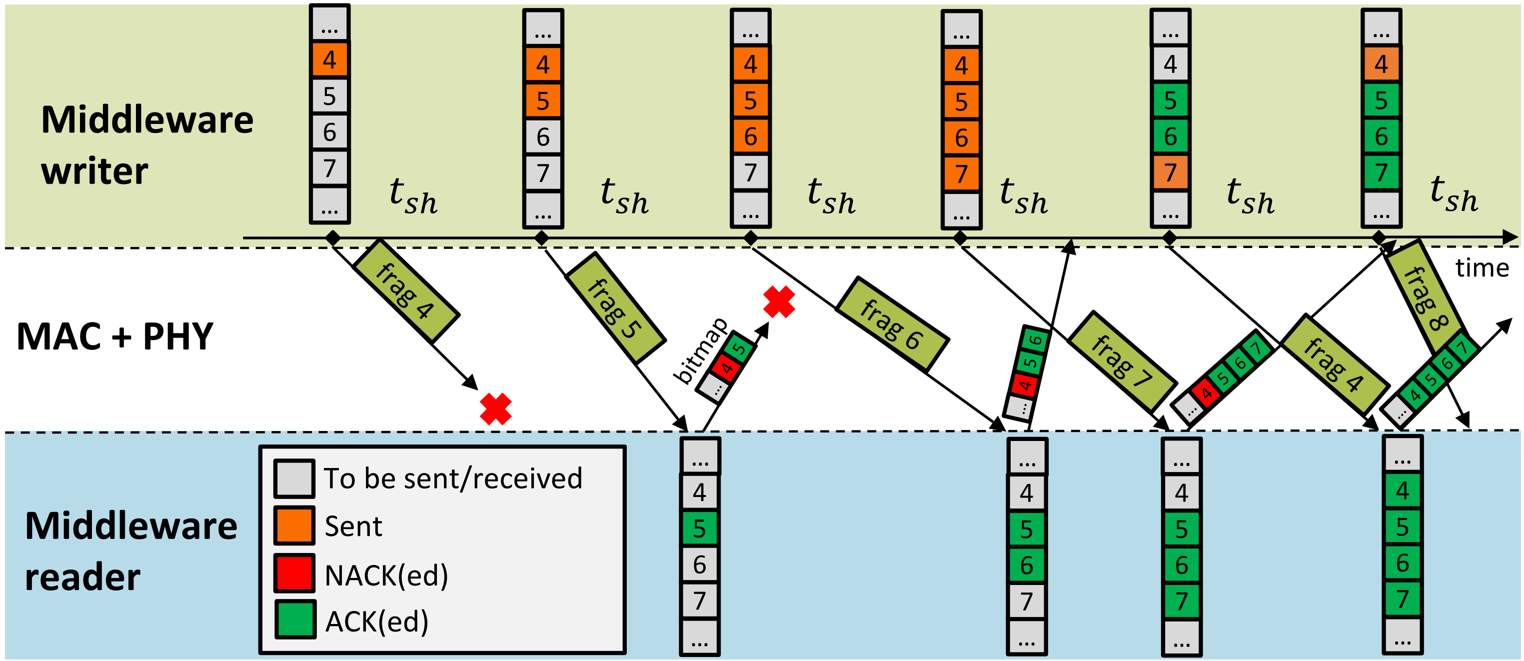
- **robust low latency guarantees only for single packets**
- **robust solutions for low latency sensor streaming and collaboration needed in addition**
  - remote vehicle control
  - collaborative sensing and perception
  - vehicle motion control

- **larger object deadlines enable efficient backward error correction (BEC)**
  - higher reliability than MAC-layer retransmissions
  - lower overhead than FEC
  - deadline asymmetry helps error statistics



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

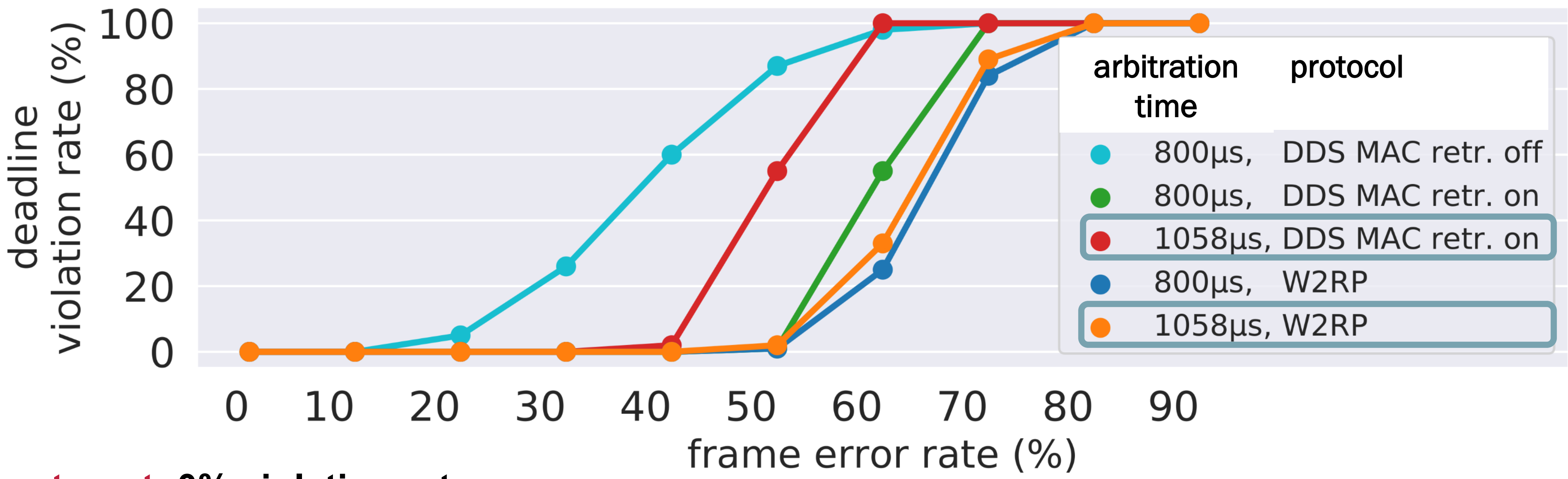
- **2-phase transmission scheme**
  - start with first object transmissions
  - bitmap-based retransmissions
- **media access shaping**
  - avoid growing round-trip times
  - reduce interference
  - controlled by resource management
- **exploit application-level deadlines**
  - avoid unnecessary error protection and retransmissions
  - survive short error bursts and frequent errors
  - low reliability overhead



W2RP efficiency

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

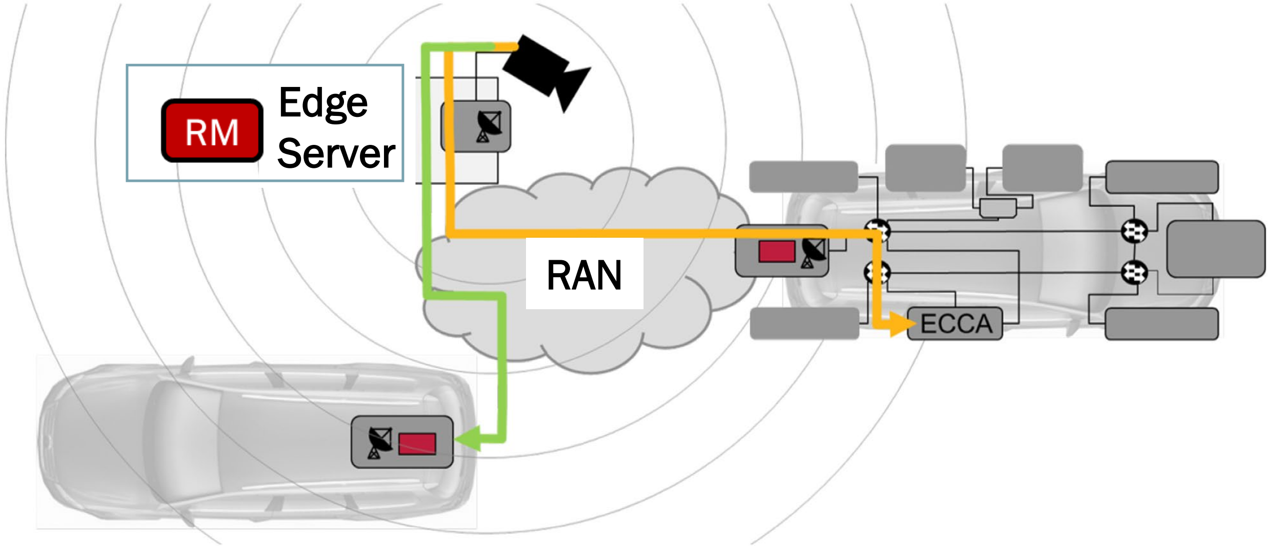
IEEE 802.11p
object size 20kB
period 100ms
deadline 100ms

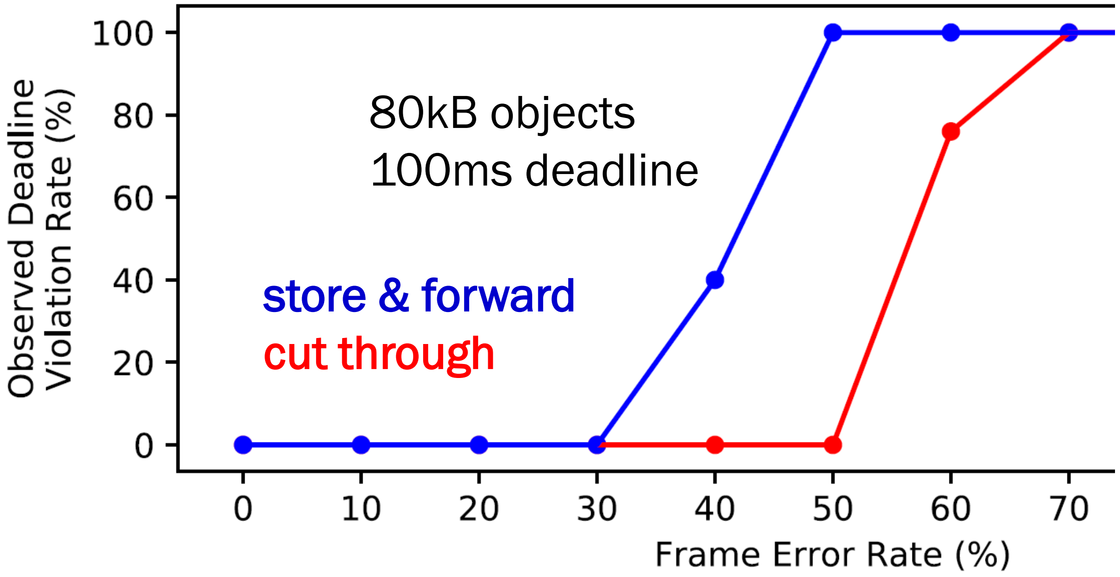


no-loss target: 0% violation rate

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 or cut through in Telemetry Unit
- wireless parameter control in edge server Resource Manager (RM) - hierarchical RM



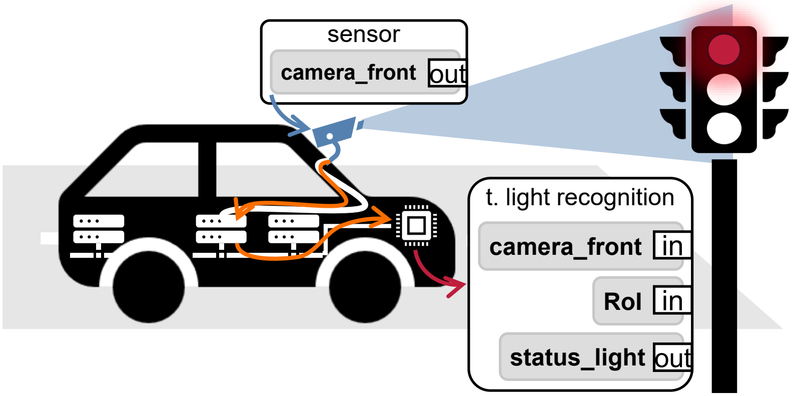


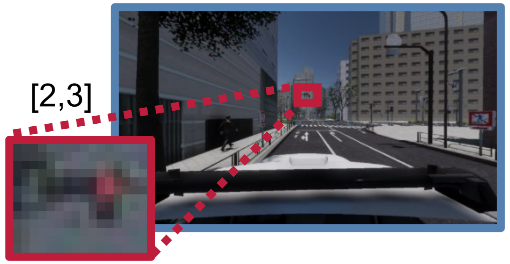
Outline

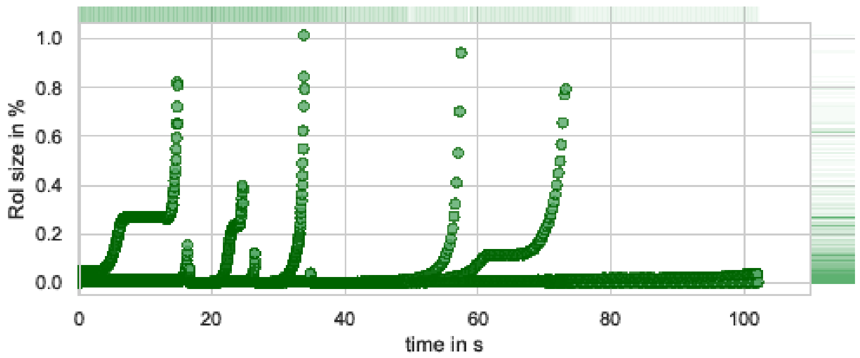
- Managing large real-time data objects – an application-based approach
- From packet to application object deadlines
  - Application 1: in-vehicle stream coordination
  - Application 2: at run-time network re-configuration
  - Application 3: reliable wireless communication
- **Application data caching for dynamic volume reduction**
- Conclusion

Managing large data objects - object size reduction

- **selective application-driven data transfer often sufficient**
  - example: selective data transfer for Region of Interest (RoI)
  - selection is feedback from sensor fusion or object tracking
    - AUTOWARE example: map-based RoI specification of traffic lights in camera images



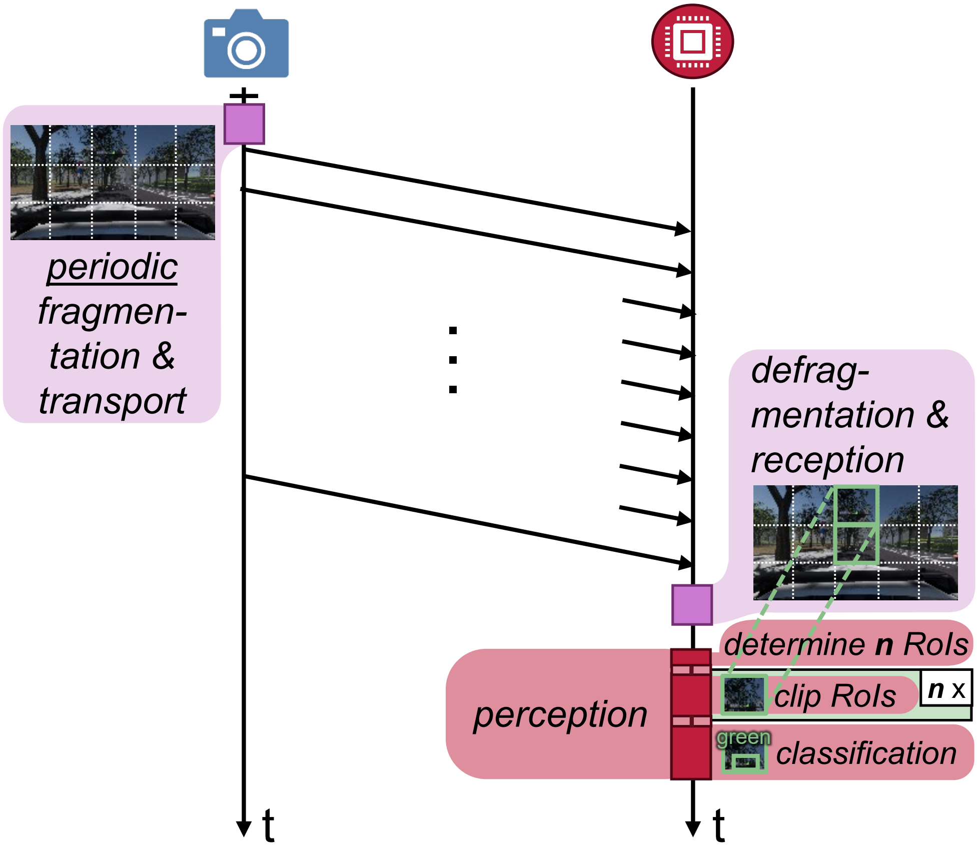




→ **large optimization potential**

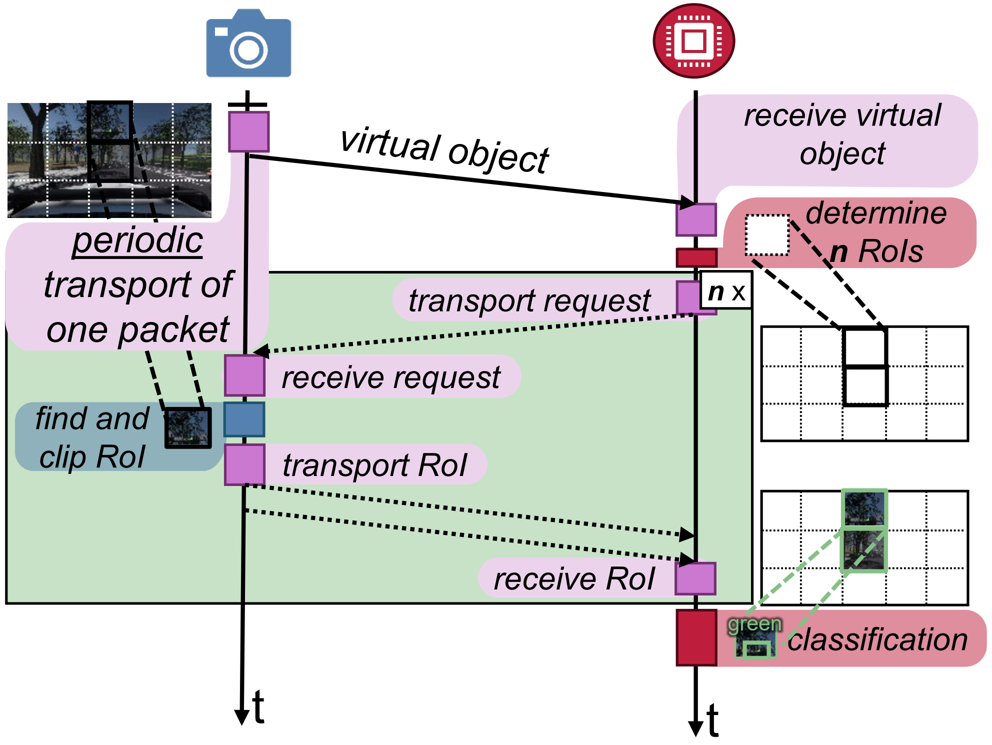
State-Of-The-Art Publish Subscribe Middleware (DDS)

- **large object are transmitted from the sensor node ( ) to the processing unit ( ) through a series of fragments**
  - example: **FullHD image** of 6,22 MB requires over **6000 TSN packets**
- **publish-subscribe (PS) middleware sends full sample**
  - **publisher controlled** communication
  - handles (de-)fragmentation, error protection, monitoring, …
  - example: de facto standard **DDS** (ROS2, AUTOSAR, …)
  - **application and network agnostic protocol**



Application-aware PS with subscriber-controlled software caching

- **sensor node announces sample as virtual object**
  - PS protocols reduced to **single packet**
- **subscriber application requests selected data blocks**
  - lean **software cache protocol**
  - saves network load and subscriber memory space
- **DDS compatible**
  - virtual object description uses standard data type
  - standard PS mechanism used for object announcement
  - software caching with independent companion protocol
  - *data object lifetime management stays with publisher (e.g. DDS history buffer)*



- ***made available to the ROS community – to be presented at ROSCon 2024***
- ***paper at VTC 2024: strong reduction of sample latency and object stream interference***

Outline

- Managing large real-time data objects – an application-based approach
- From packet to application object deadlines
  - Application 1: in-vehicle stream coordination
  - Application 2: at run-time network re-configuration
  - Application 3: reliable wireless communication
- Application data caching for dynamic volume reduction
- **Conclusion**

Conclusion

- **manage large real-time data at application level, not at packet level**
  - this talk: focus on application-communication mismatch
  - showed benefits in scheduling, network configuration, robustness, network latency
  - more examples: handover, service integration, …
- **application aware end-to-end management more important than individual network segments**
  - end-to-end network management needs an application layer
  - missing in current communication standards – too packet and packet stream oriented
- **established publish-subscribe data service treats all data objects equal**
  - inefficient for large data objects
  - software caching extension can be essential

Some related papers

Application 1:

- 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.

Application 2:

- 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.

Application 3:

- 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 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.
- Alex Bendrick, Jonas Peeck and Rolf Ernst, “On the effectiveness of W2RP in physical environments”, *Publications Institute of Computer and Network Engineering*, 2023, January. https://doi.org/10.24355/dbbs.084-202301231301-0
- 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.

Application Data Caching:

- Nora Sperling and Rolf Ernst, “Reducing Communication Cost and Latency in Autonomous Vehicles with Subscriber-centric Selective Data Distribution”, IEEE Vehicular Technology Conference (VTC), Singapore 2024.
- Nora Sperling, Alex Bendrick, Dominik Stöhrmann and Rolf Ernst, "Invited: Caching in Automated Data Centric Vehicles for Edge Computing Scenarios," 2023 60th ACM/IEEE Design Automation Conference (DAC), San Francisco, 2023.
- Kai Björn Gemlau, Jonas Peeck, Nora Sperling, Phil Hertha and Rolf Ernst, "A new design for data-centric Ethernet communication with tight synchronization requirements for automated vehicles," IECON 2019 - 45th Annual Conference of the IEEE Industrial Electronics Society, Lisbon, Portugal, 2019.