Challenges and opportunities stemming from the adoption of radio resource slicing and edge computing paradigms

Vincenzo Mancuso

IMDEA Networks Institute, Madrid, Spain

vincenzo.mancuso@imdea.org

Introduction

What am I going to talk about?

- Communication networks and cars!

Why?

- Well, there is a growing interest in connected cars and extended networks services for “verticals”, beyond the network operator’s scope
- This goes beyond networking and involves dynamic infrastructure management and flexible computing

How is that possible?

- Thanks to techniques like network slicing and multi-access edge computing (MEC)

Slicing what?

- Slicing has been proposed for computer infrastructures
  - To enable and isolate multiple services
- Network slicing techniques extend the concept to network infrastructures
  - To create services with a pool of resources
  - To enable multi-tenancy
  - It is a key point of 5G
  - Recently addressed by a number of works
  - Little has been done in terms of slicing and handling **radio resources**

MEC for what?

- Edge computing: moving computing close to the end-user
- ETSI has defined a multi-access edge computing (MEC) architecture
  - For mobile networks and network services
  - For virtualization of network functions
  - Again, this is a 5G key point
- Many use-cases have been defined
  - but there is yet no clear assessment of its performance and no clear vision on how to use MEC to enhance services in concrete and dynamic wireless network scenarios.

Focus on

- Slicing and MEC for wireless connected cars
  - Cars that connect the network and networks that connect cars
- Models for radio slicing, for the coexistence heterogeneous “verticals”
  - human-type and machine-type traffic

Challenges and opportunities offered by

5G-style network slicing in the radio access network

and by the deployment of a distributed ETSI MEC

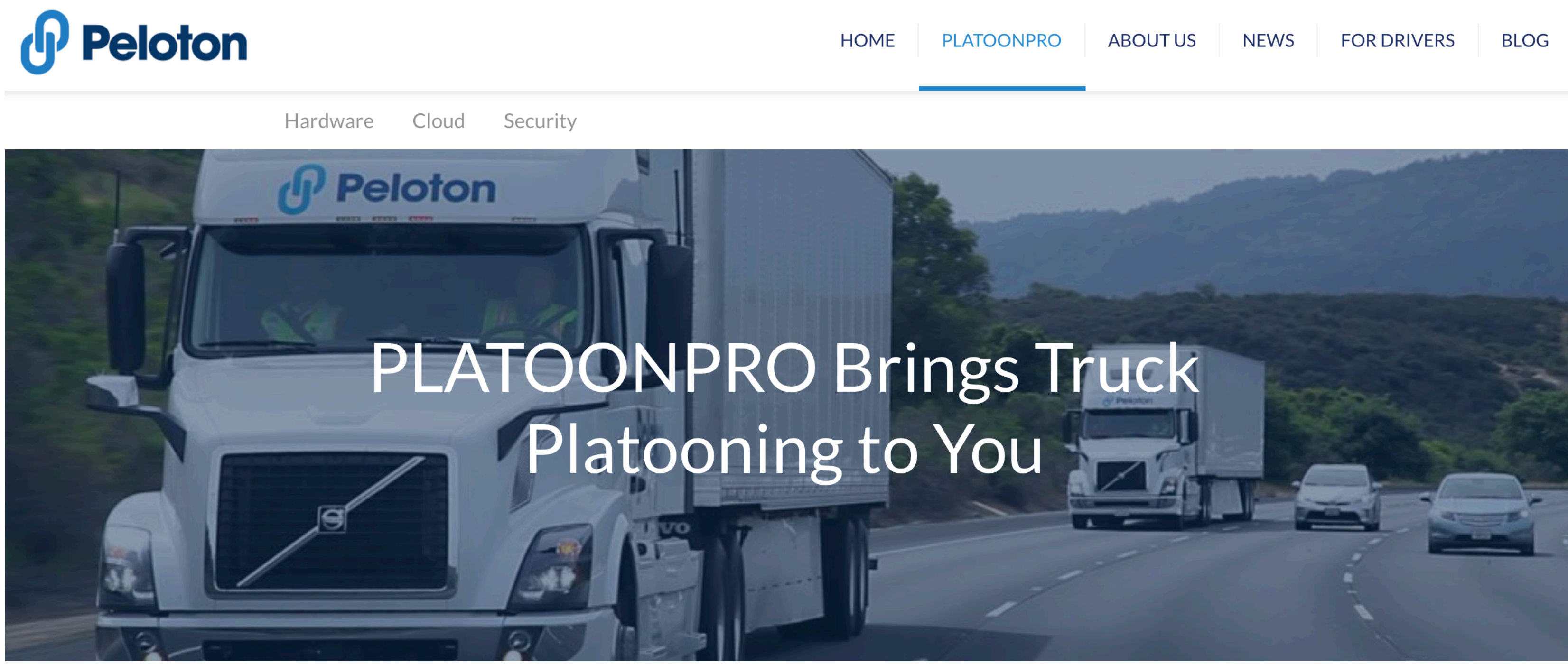
Cars and Networking

Networked vehicles / Connected cars

- Many possible (network-challenging) applications
  - Assisted driving
  - Platooning
  - Safety for drivers and pedestrians
  - Etc.
- An opportunity for expand network coverage and services
  - Relay/backhauling
  - Cell extension
  - Info spreading
  - Cooperative positioning/localization
  - Etc.

Example: vehicle communications for platooning

- E.g., coordinated braking and acceleration of vehicles traveling together
  - With V2V and V2X
- Challenges
  - Delay
  - Bandwidth
- Already commercial!
  - However, it currently works with a fistful of vehicles (2 trucks!)



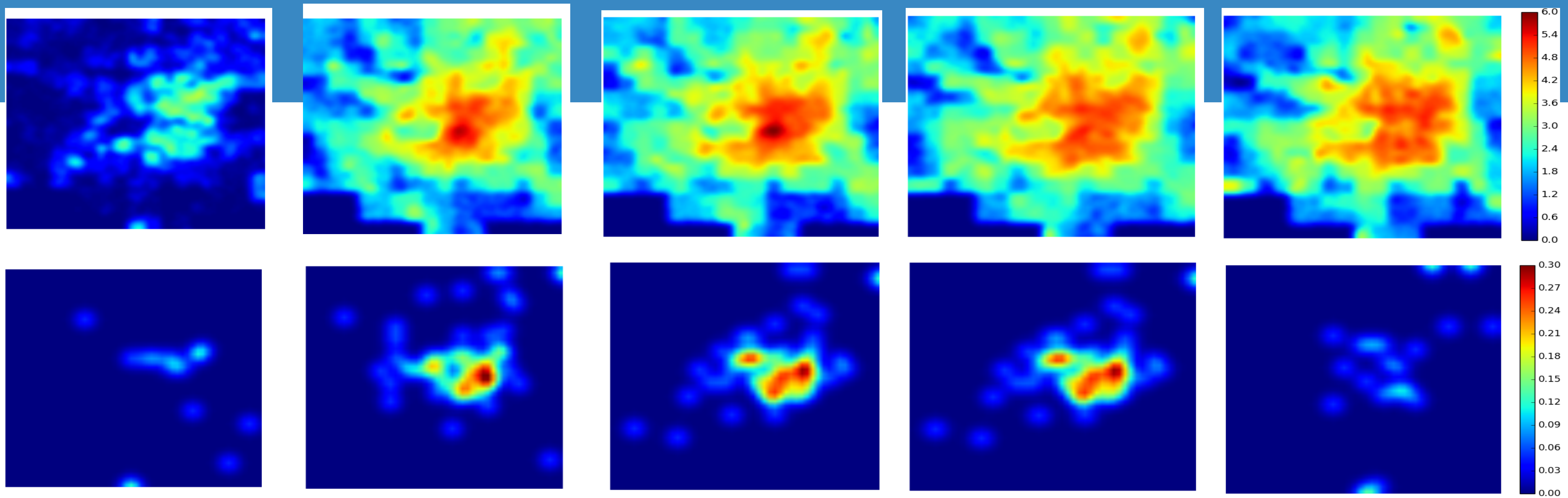
*"**Peloton**'s truck platooning system uses Vehicle to Vehicle (V2V) communication to connect the braking and acceleration between the two trucks. The V2V link allows the lead truck to control the acceleration and braking of both trucks virtually simultaneously, reacting faster than a human or even radar sensors could."*

Example: vehicles for network densification

- Radio access network densification is necessary to increase capacity
  - Network densification with fixed small cell base stations is costly
  - Network densification with drones/UAVs is computationally challenging

- *Alternative unsupervised approach*: adaptive densification using mobile small-cell on ground vehicle
  - temporary network densification where and when traffic demand peaks
  - adaptive wireless bandwidth increase

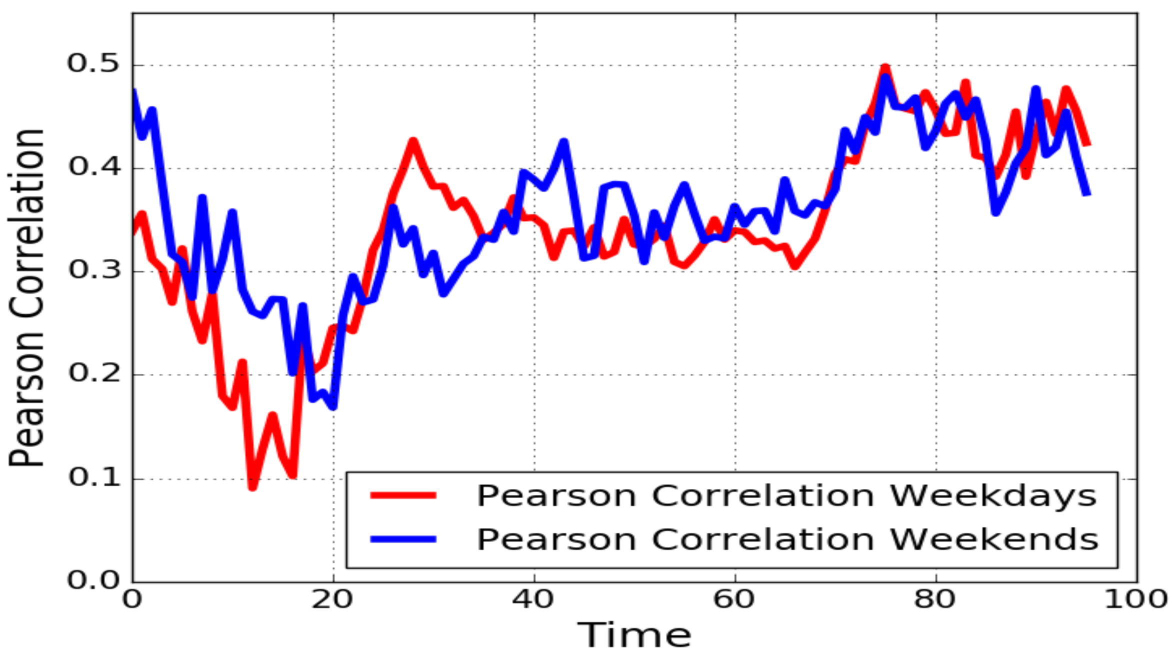
Telecom traffic & vehicular traffic

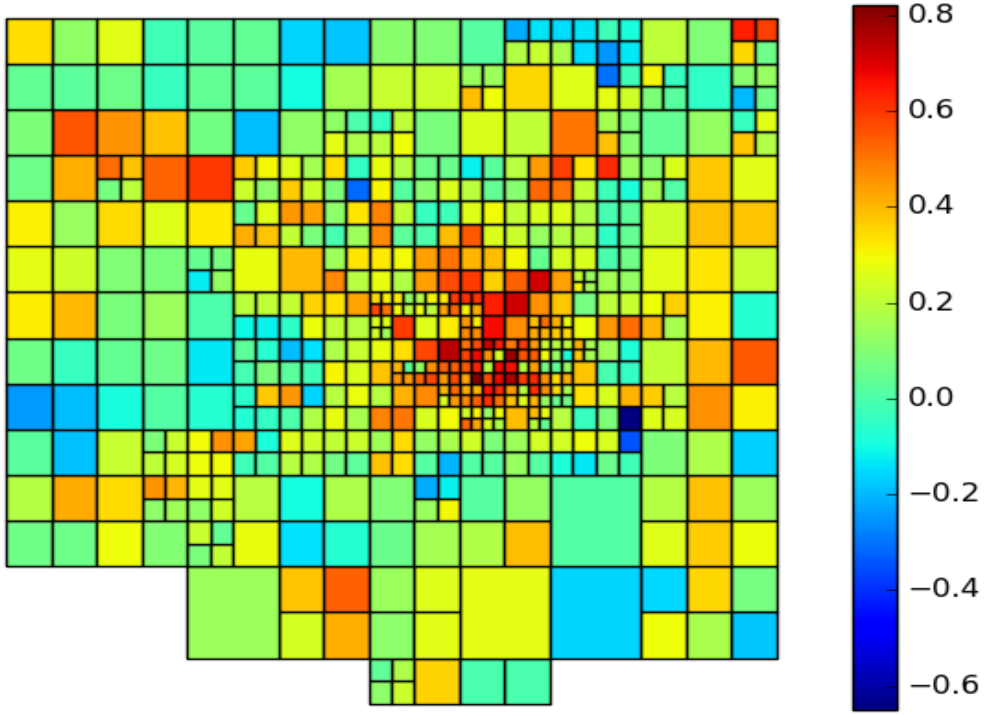


- Example heat-maps at 5 different hours: 5 am, 10 am, 3 pm, 8 pm and midnight
- Telecom traffic in comparison to vehicular traffic rises earlier in the morning, then it remains high later in the evening and is more pervasive
- Traffic peaks coincide with the city center in both cases

F. Mohammadnia, M. Fiore and M. Ajmone Marsan, "Adaptive densification of mobile networks: Exploring correlations in vehicular and telecom traffic," ***Med-Hoc-Net*****, Capri, 2018, pp. 1-8.**

Example of traffic correlation



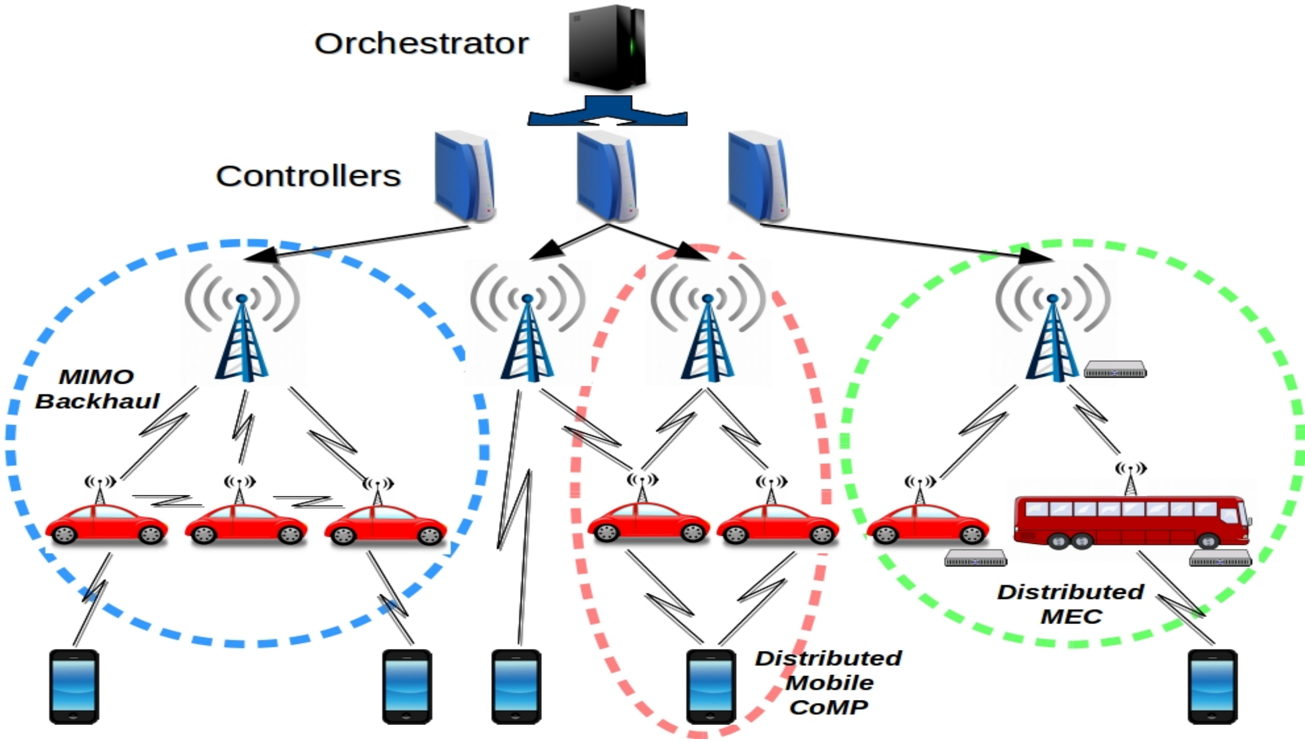


F. Mohammadnia, M. Fiore and M. Ajmone Marsan, "Adaptive densification of mobile networks: Exploring correlations in vehicular and telecom traffic," *Med-Hoc-Net*, Capri, 2018, pp. 1-8.

So, shall we use base stations on vehicles?

➔ ***Idea: equip some vehicles with base stations, <u>if a fast backhaul is available</u>***

- Advantages: low CAPEX, new markets, better coverage, etc.
- Challenges: stable mobile backhaul, dynamic re-optimization, privacy, network control, etc.
- ***New research opportunities:***
  - V2V/V2I/V2N cooperation, opportunistic BS access (fast and frequent handover, access to multiple BSs), use of new technologies (mmWave, VLC inside the vehicle, D2D, etc.)



M. Ajmone Marsan, F. Mohammadnia, C. Vitale, M. Fiore, V. Mancuso, "Towards mobile radio access infrastructures for mobile users", Ad Hoc Networks, Volume 89, 2019

Networks tools for connected vehicles applications

Can 5G help?

Many heterogeneous challenges!

- Bandwidth
- Delay
- Computation
- Backhaul
- Privacy
- …

- A part for bandwidth, 5G will offer tools for
  - Service tailoring
  - 3rd party services
  - Verticals

- Slicing
  - Assign resources to verticals
  - But also to flows, groups, etc.
  - Dedicated resources? More or less!
- Virtualization of network functions
  - Flexible control and composition of functions
  - Distributed and reconfigurable implementation
- MEC
  - Edge computing in the loop!
  - Bring computational tools very close to users
  - Deploy services and virtualize network functions through it

Network slicing in 5G

- Network slicing is one of the defining features of the 5G technology
  - “network slice instance: a set of network functions and the resources for these network functions which are arranged and configured, forming a complete logical network to meet certain network characteristics.” (3GPP TR 28.801)
  - It will allow the presence of several tenants on one infrastructure and the effective coexistence of multiple services with different characteristics and requirements
- Several papers and projects already studied issues related to resource orchestration
- Less work available on slicing radio resources

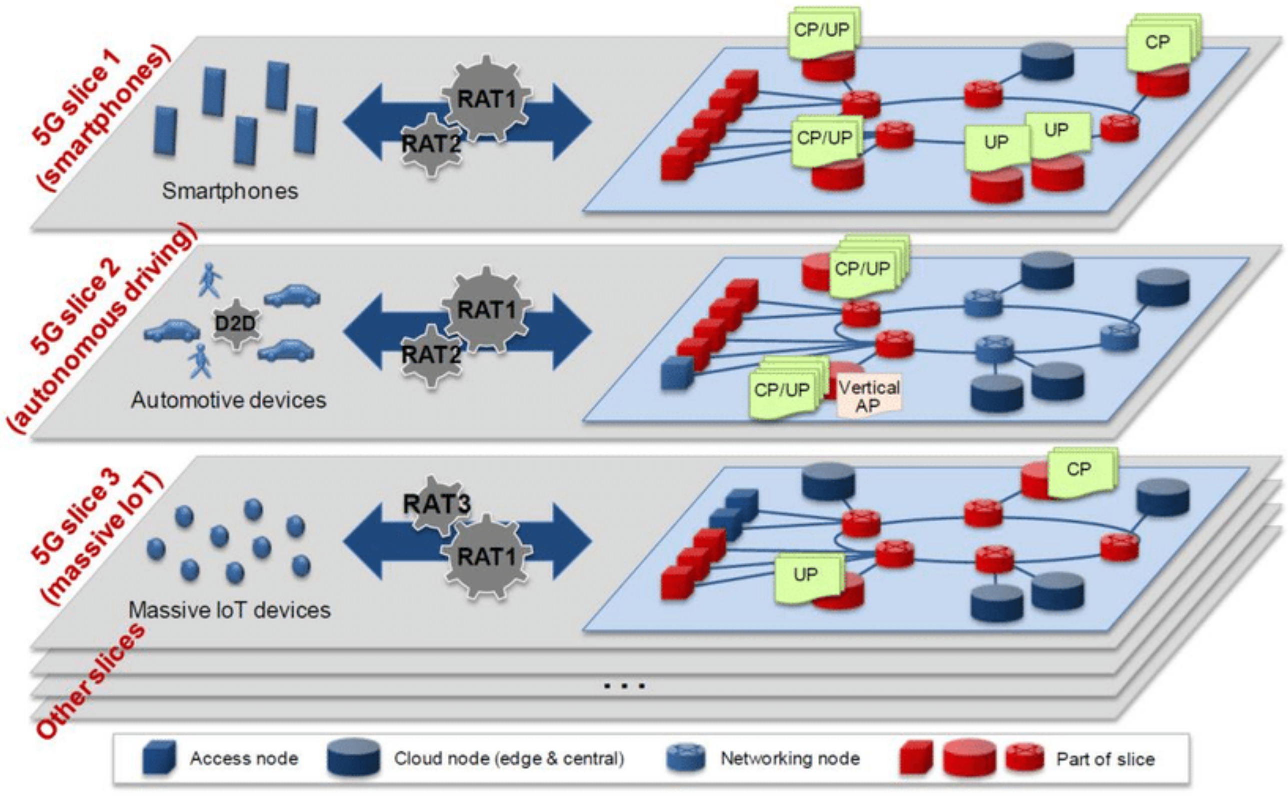
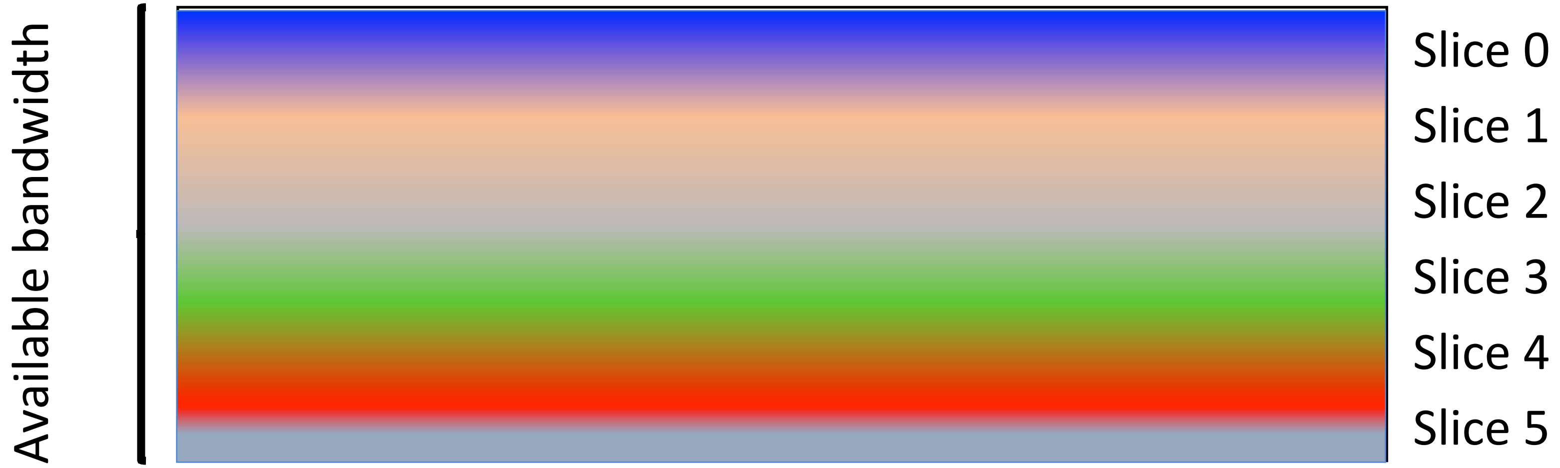


Figure extracted from Deliverable D1.1 of the H2020 METIS project

Network slicing in the RAN

- Logical subdivision of Radio Access Network (RAN) resources
  - With shared radio scheduler and (partially, dynamically) shared physical resources
- It leverages on Software Defined Network (SDN) for control and Network Function Virtualization (NFV) for building the user data plane
  - With a resource orchestrator that manages resources and monitors KPIs



What to actually slice in the RAN?

Many kinds of resources could be **sliced** in view of multi-tenancy and to handle traffic flows in general, e.g.:

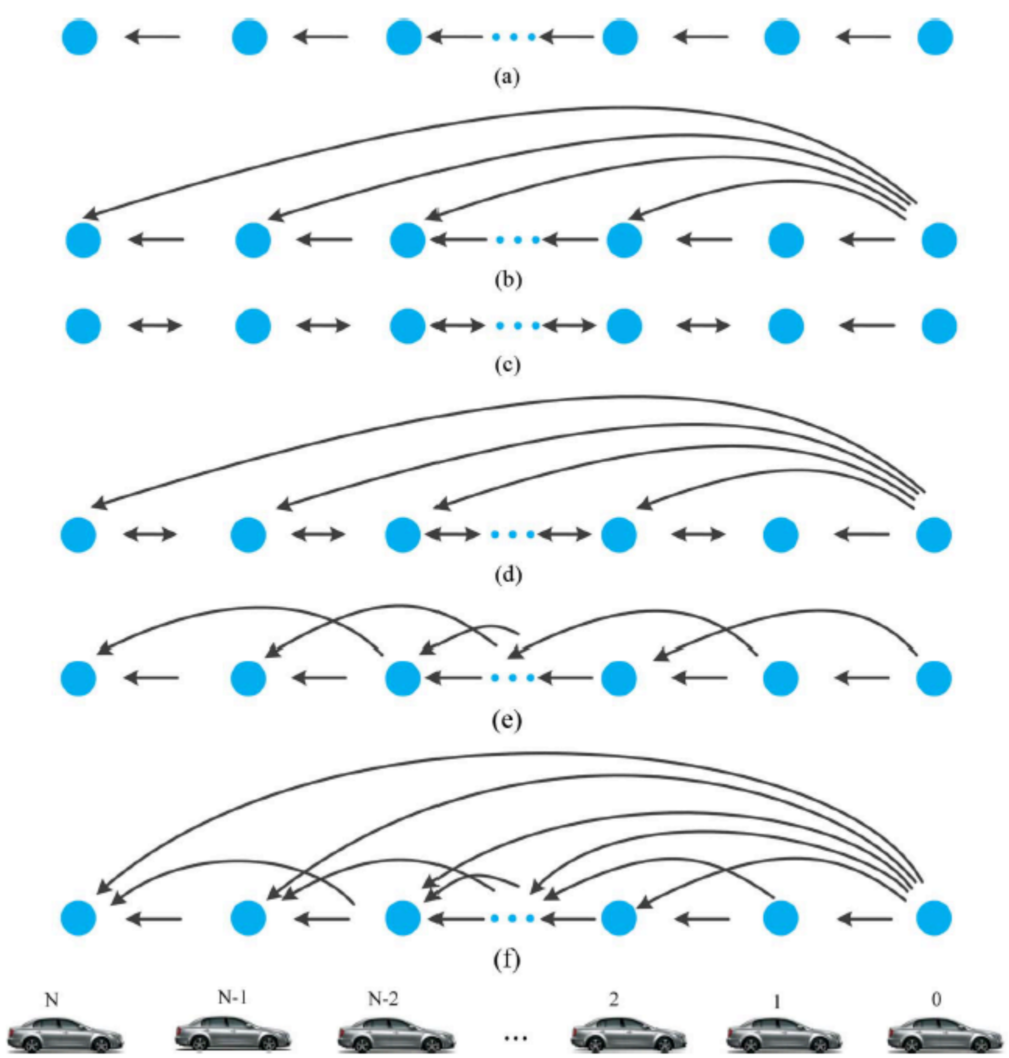
- Bandwidth in general
- Physical resources (e.g., minislots, OFDMA blocks, antennas, radios, etc.)
- Number of users under services (RRC_CONNECTED)
- RACH preambles
- Number of ACKs per RAO
- Etc.

Different levels of complexity are possible

- Slice-type-dependent
- Vertical industry-dependent

RAN slicing for connected vehicles?

- Support various types of communications
  - V2V
  - V2N
  - V2I
- Support various types and combination of traffic requirements:
  - mMTC for high numbers of machine-type users
  - eMBB for high bandwidth
  - URLLC for low delay & reliability



Y. Zheng, S. Eben Li, J. Wang, D. Cao and K. Li, "*Stability and Scalability of Homogeneous Vehicular Platoon: Study on the Influence of Information Flow Topologies*," in *IEEE Trans on Intelligent Transportation Systems*, vol. 17, no. 1, 2016.

V2X in 3GPP release 15 (“5G”). What’s beyond?

V2V and V2I in 5G is performed via:

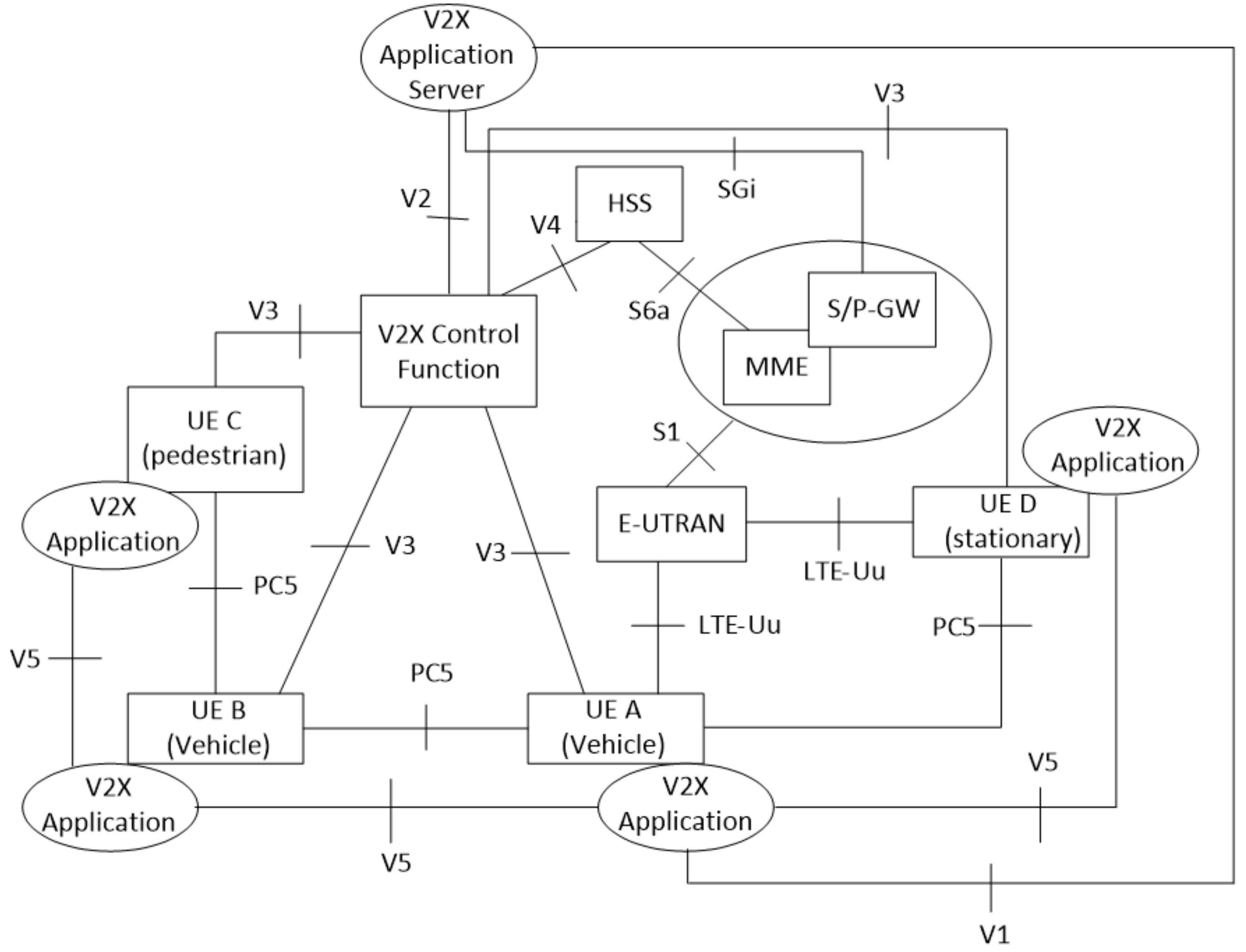
1) the ProSe PC5 interface

2) or 802.11p, with Dedicated Short-Range Communication

RSUs are seen as stationary UEs

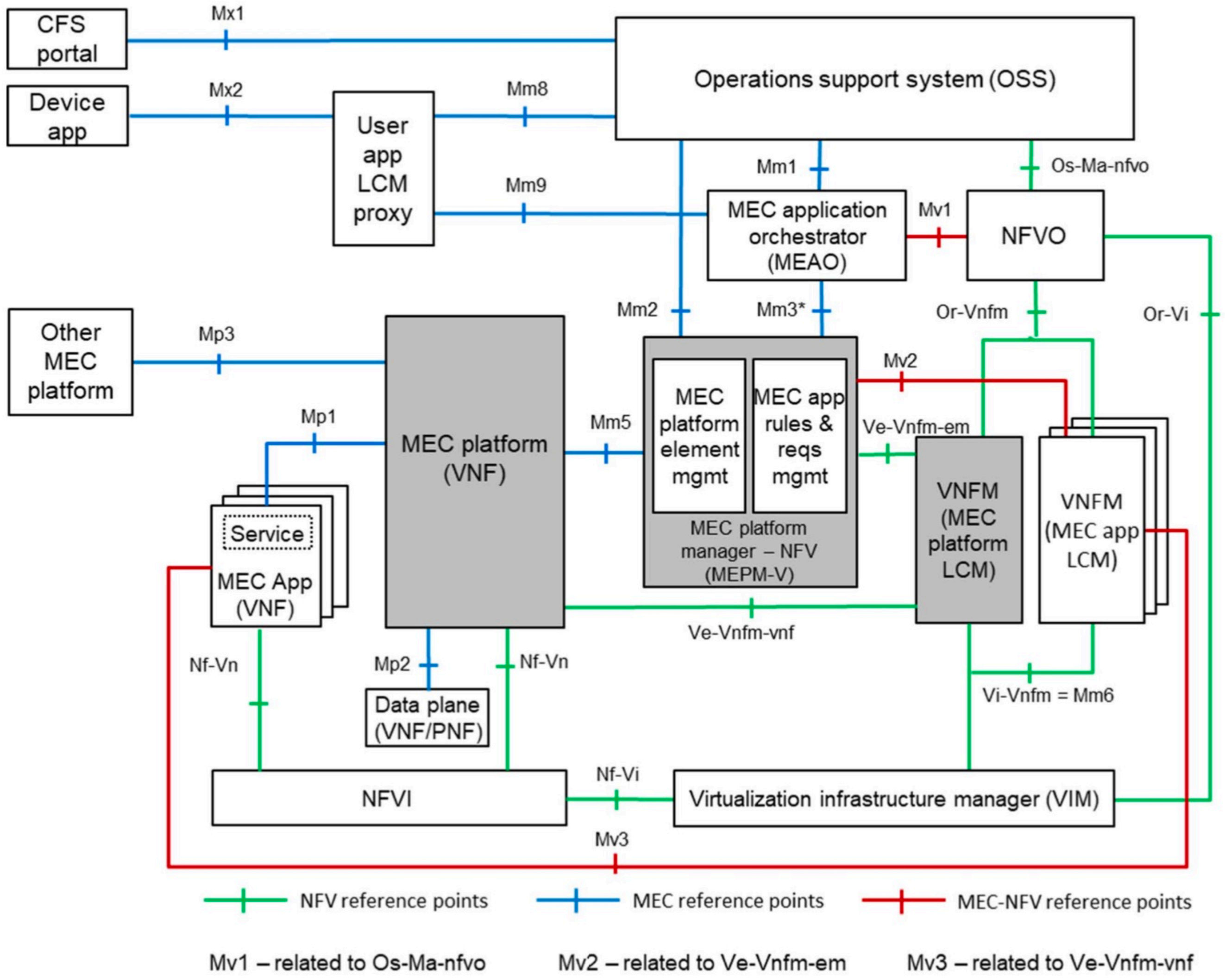
Latency and bandwidth are still unsolved problems:

- **Use V2N instead of V2V, leveraging on, e.g., eMBB + URLLC?**
- **Use mMTC to collect sensor data?**

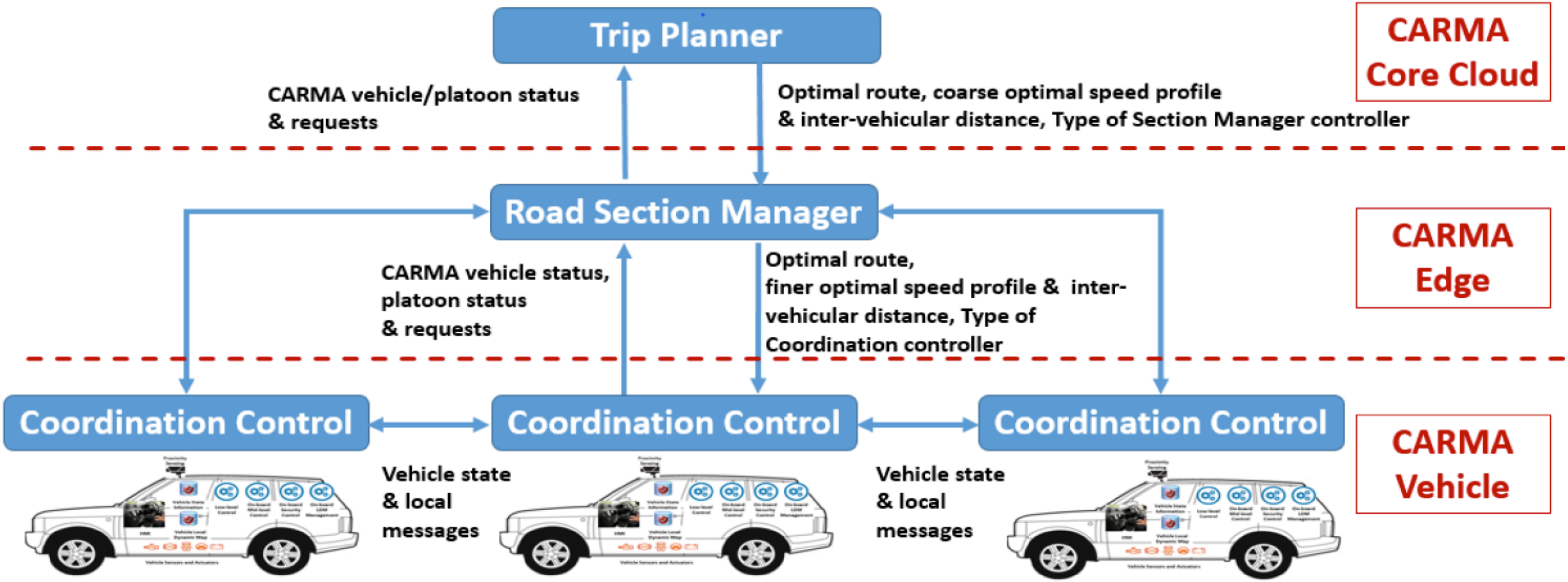


Multi-access edge computing (MEC)

- Cloud functionalities for network virtualization at the edge
  - Low latency
  - High bandwidth
  - real-time access to radio network information
  - Multi-tenancy
  - *MEC offers application developers and content providers cloud-computing capabilities and an IT service environment at the edge... (ETSI)*
- Joint ETSI standardization effort towards MEC + VNF
  - MEC extends and complements the virtualized infrastructure of the network



Cloud and Edge computing for connected cars?



Montanaro, Umberto & Fallah, Saber & Dianati, Mehrdad & Oxtoby, David & Mizutani, Tom & Mouzakitis, Alexandros. (2018). *Cloud-Assisted Distributed Control System Architecture for Platooning*. 1258-1265. 10.1109/ITSC.2018.8569295.

ETSI MEC use cases for connected cars!

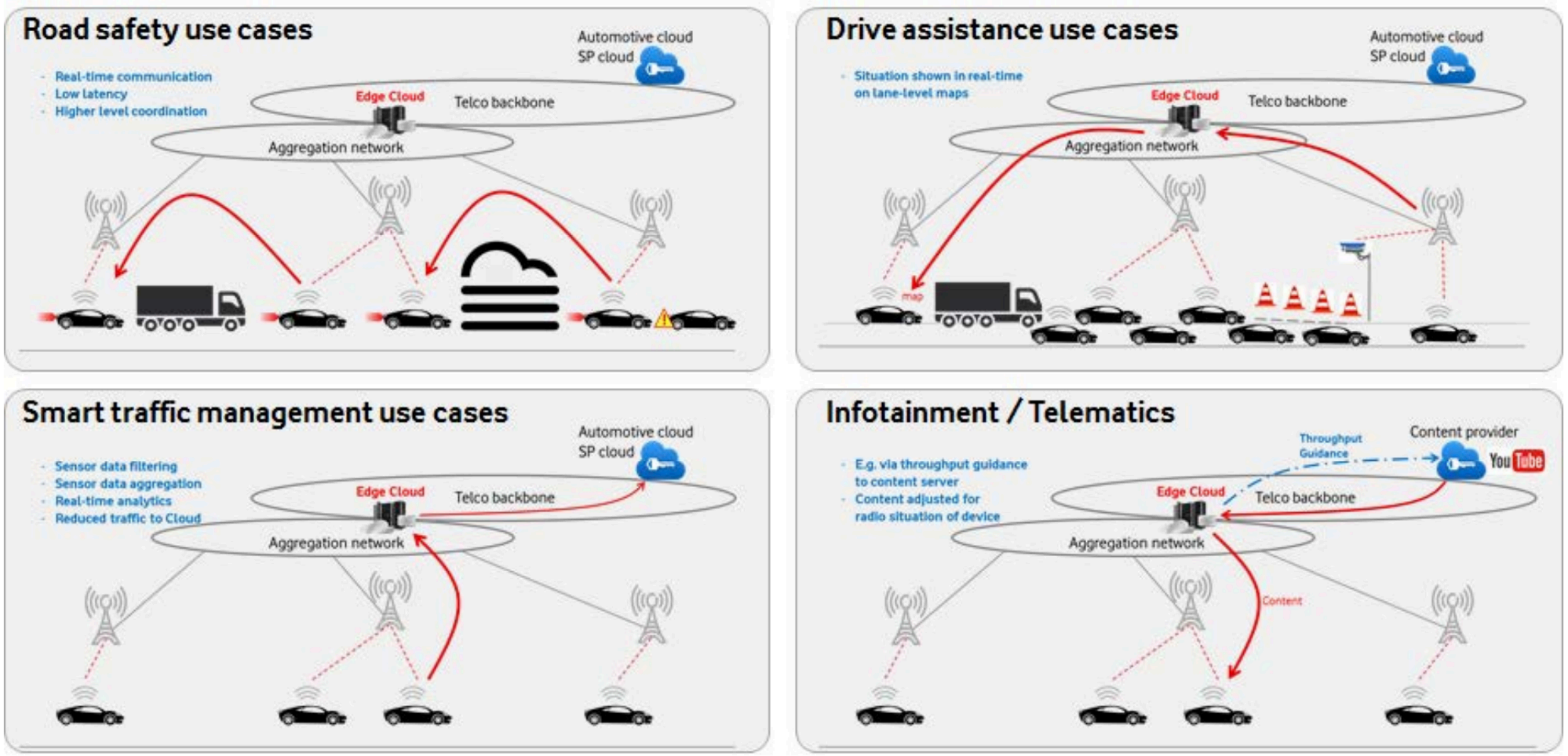
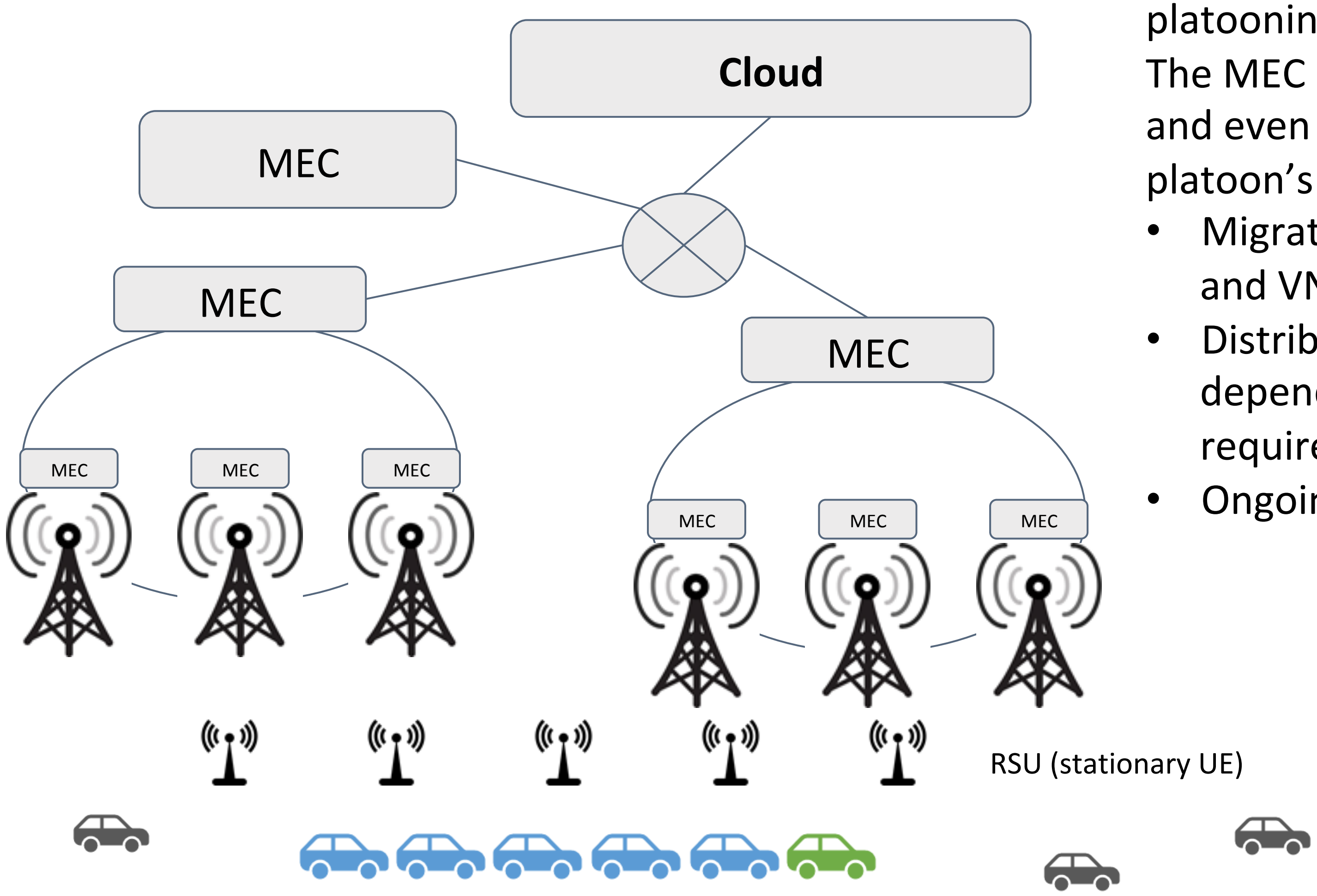


Figure 6.1.1-1: Example MEC support for V2X use cases

Flexible MEC location

- The MEC can be distributed
- MEC machines can be attached close or far from the UPF
- Networking:
  - Integrate operator's VNFs
  - Support for specific slices
- Computing:
  - Run data fusion services
  - Service caching
  - OTT service control
- Tradeoff between latency and computational power



Example for platooning:
The MEC can follow and even anticipate platoon's moves

- Migration of MEC and VNFs
- Distributed MEC, depending on requirements
- Ongoing work

Slice and service coexistence in the presence of VNFs for connected vehicles

- According to ETSI, MEC with connected vehicles is meant for:
  - Convenience applications
  - Safety applications
  - Advanced driving assistance
  - Protect vulnerable road users (pedestrians and cyclists)
- This requires not only MEC …
  - **Fast computation** at the edge
  - Data collection and **data fusion**
  - Manage **MEC** resources and **service chains** smartly
- …but also slicing
  - Use of **multiple slices and slice types**
  - Manage **RAN** resources efficiently

Focus on wireless networking

RAN slicing model

Let's start with requirements (e.g., for platooning)

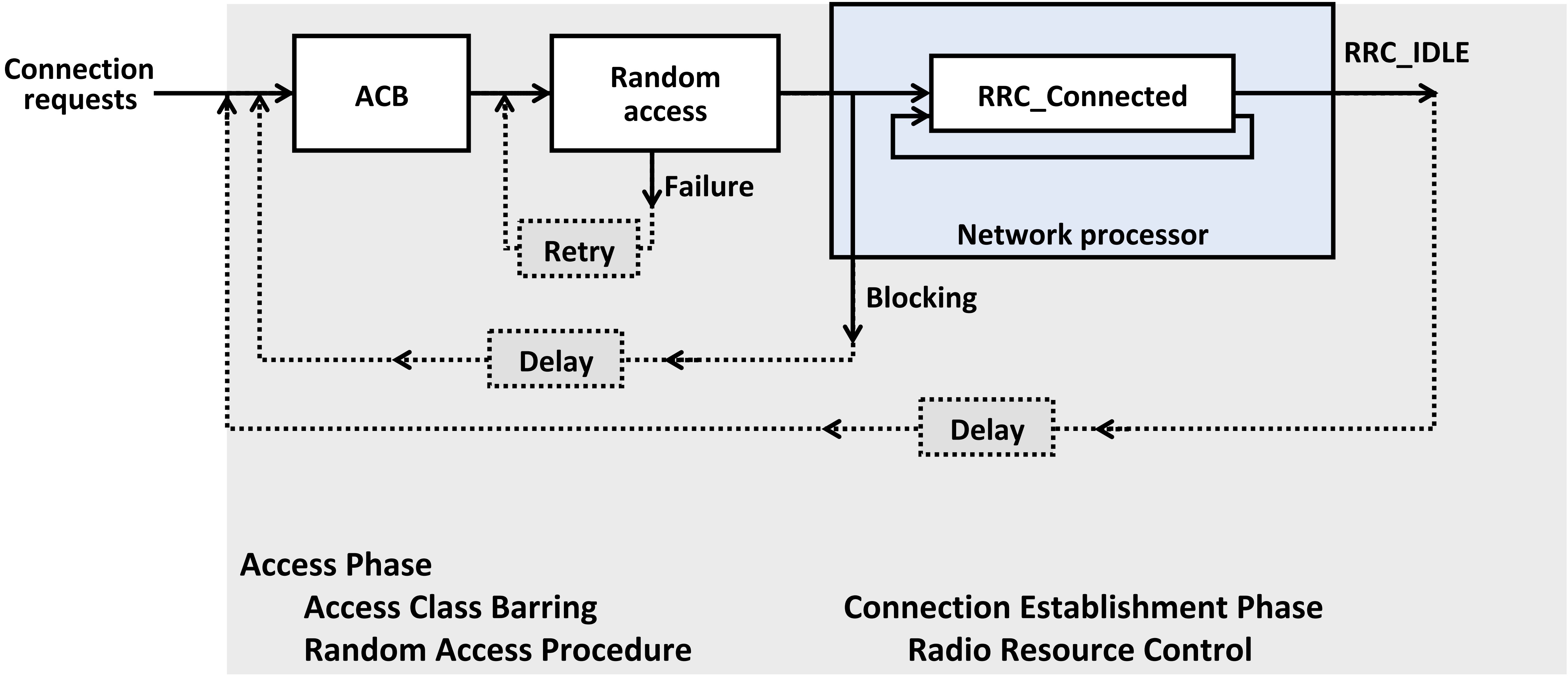
- Platoon Messages CAM (Cooperative Awareness Message)
  - 50-1000 bytes platoon message uplink
  - 40-100 Hz according to the platoon parameters
  - latency 10-25 ms according to the platoon parameters (**URLLC needed**), up to 100 ms for coarse driving intentions (e.g. speed adaptation) (**can use mMTC**)
  - loss 10^{-3} - 10^{-5} based on the platoon parameters
- Sensor sharing uploaded to MEC to refine the context
  - uplink intensive 10-50 Mb/s (**it requires eMBB**)
  - 10-20 Hz
  - latency up to 100ms (**can require URLLC in addition to another slice**)
  - loss 10^{-1} - 10^{-3}
  - High connection density

→ **Multiple slice types are needed, how would they coexist?**

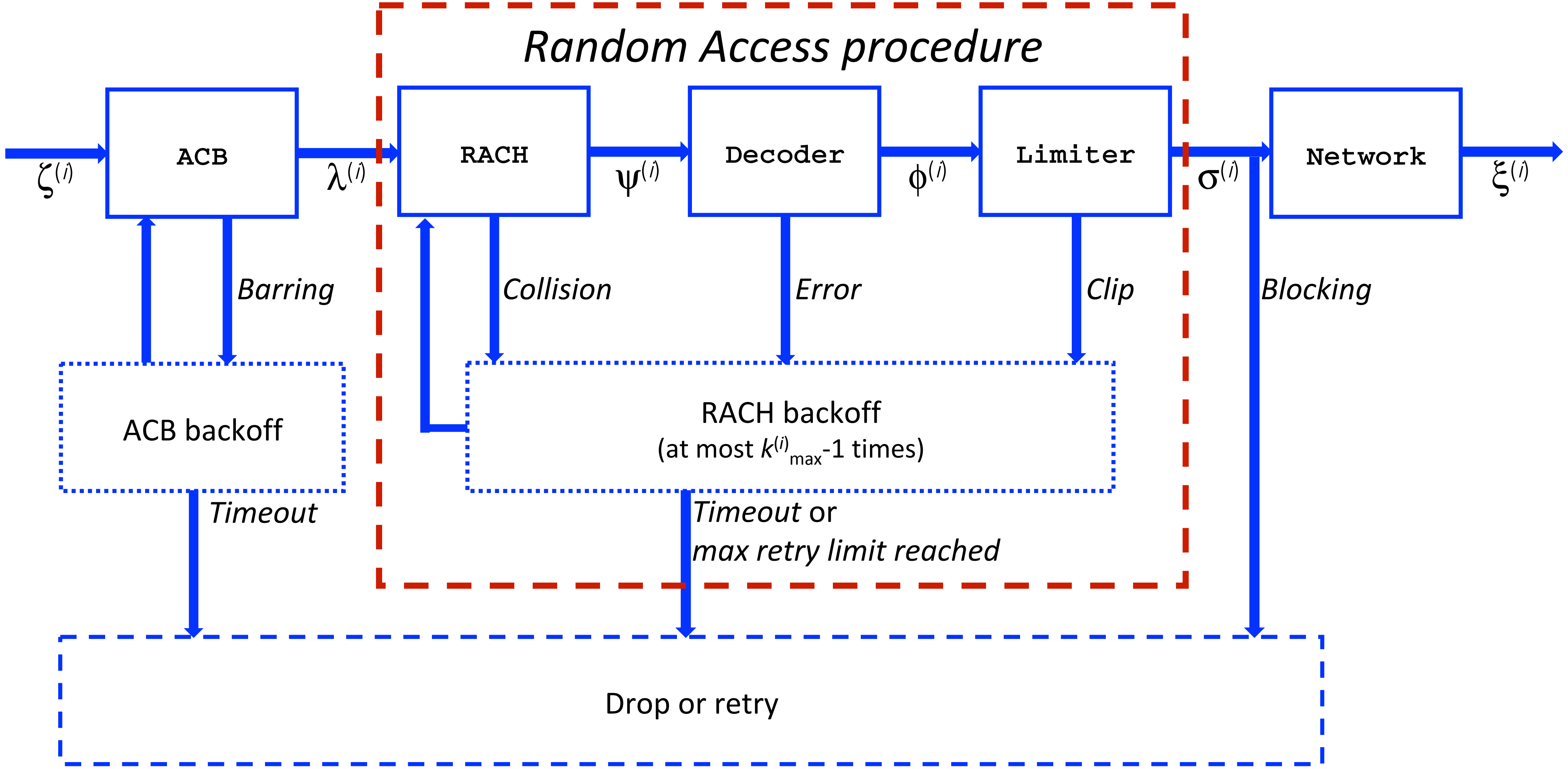
RAN slicing model

- We developed a detailed stochastic model of the behavior of the radio access network in a sliced cell
  - Access class barring (ACB)
  - Random Access procedure
    1. Preamble transmission on the Random Access CHannel (RACH)
    2. Preamble decoding
    3. Random Access Response (RAR)
  - Network service (network processor)
    - Radio Resource Control (RRC) connection and its limitations
    - Transmission
- **RACH and Network processor resources are sliced**
  - Parametric policies for allocating RACH/Network resources to slices
  - Part of resources is assigned to a single slice for exclusive use
  - Part of resources form a common pool for all slices

Accessing network resources



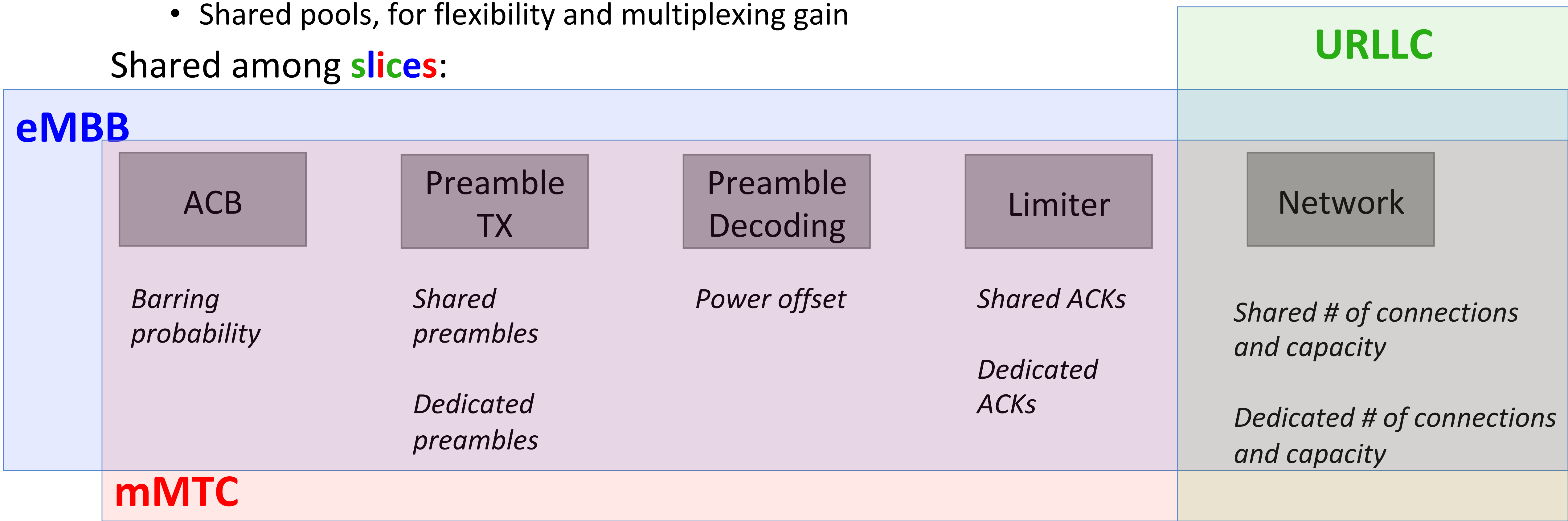
RACH more in detail



Sharing the RAN infrastructure

- Configuration of each system component is slice-dependent
- In each component resources can be parametrically divided in
  - Reserved for slices, to guarantee required QoS
  - Shared pools, for flexibility and multiplexing gain

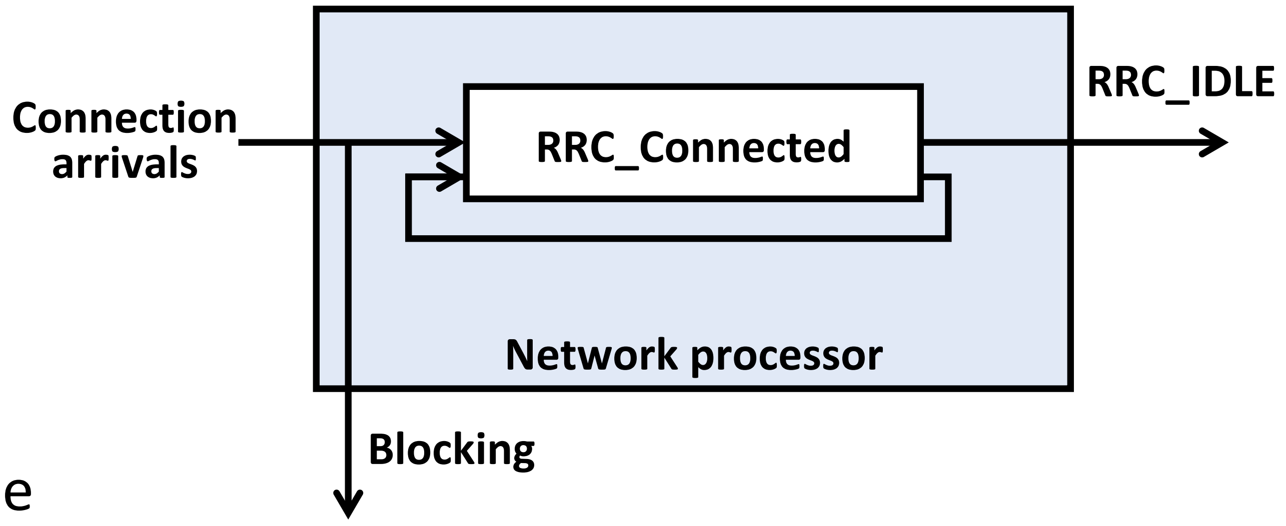
Shared among **slices**:



Network processor of a cell, with multiple slices

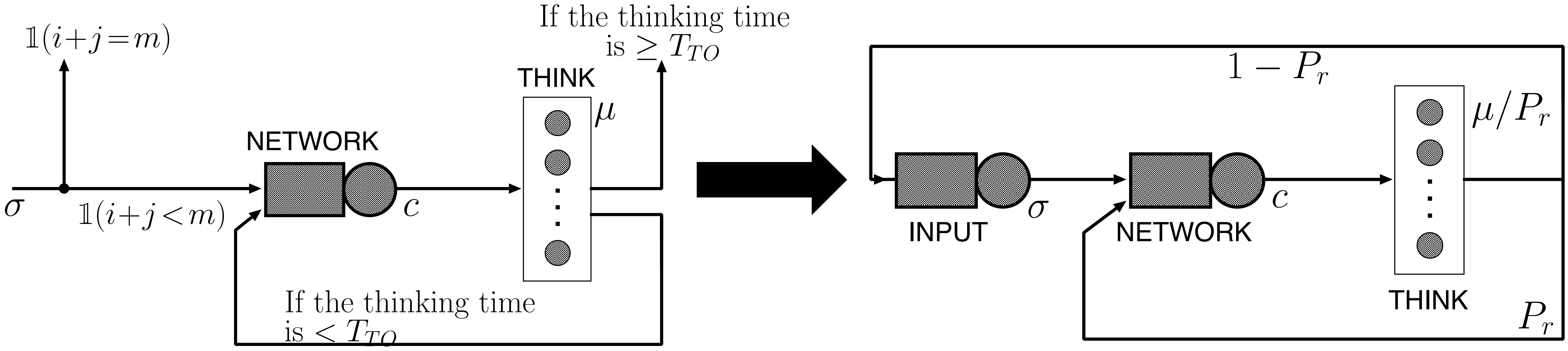
- Resources reserved for
  - eMBB
  - mMTC
  - URLLC
  - Shared services

- URLLC can steal resources
  - mostly from eMBB?
  - Easy to model: FIFO virtual queue with priority over the rest
  - No RRC_CONNECT/IDLE



Network processor models

- The network processor is an open queueing system
- However, it can be studied by means of closed queue systems
  - There is a simple and elegant product form solution



One slide in isolation (Infocom'18): plain system description and equivalent close queueing system

Modeling methodology

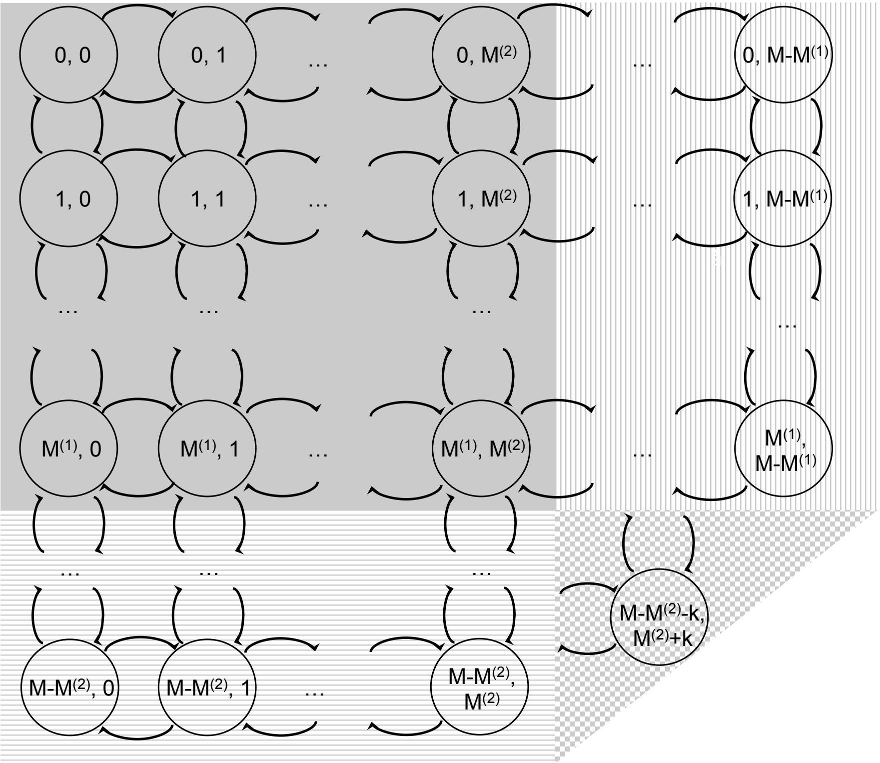
Given the exogenous traffic demand, and the sharing parameters, we derive **access delay distributions** and stochastic expressions for:

- All event **probabilities** (timeout, drop, success, etc.)
- All **flows** (steady state average values)

Computing **network blocking** requires solving a CTMC for **eMBB** and **mMTC** slices

- Special shape of the MC, due to slice **coupling**
- Very large state space
- We use SoTA solvers (few seconds for millions of states)

(results appeared in Infocom'19)



URLLC is seen as an external player whose role consists in determining how many resources are left for the other slices

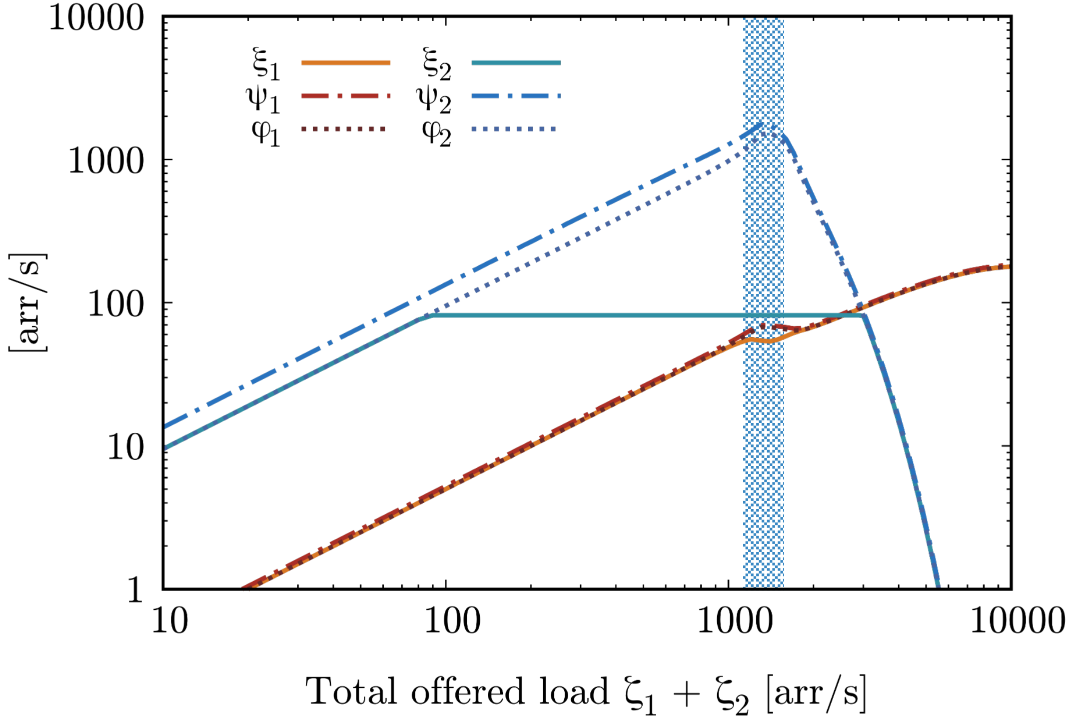
Simple study case: HTC and MTC coexistence (no URLLC here)

- The model can be used to study a large range of slice types and policies
- Slice types: We tested few key configurations for a cell hosting two slices
  - Machine Type Communications (**MTC**) and Human Type Communications (**HTC**) slices
- Policies: Resources are allocated to slices to provide a minimum QoS
  - Additional resources are shared among all slices

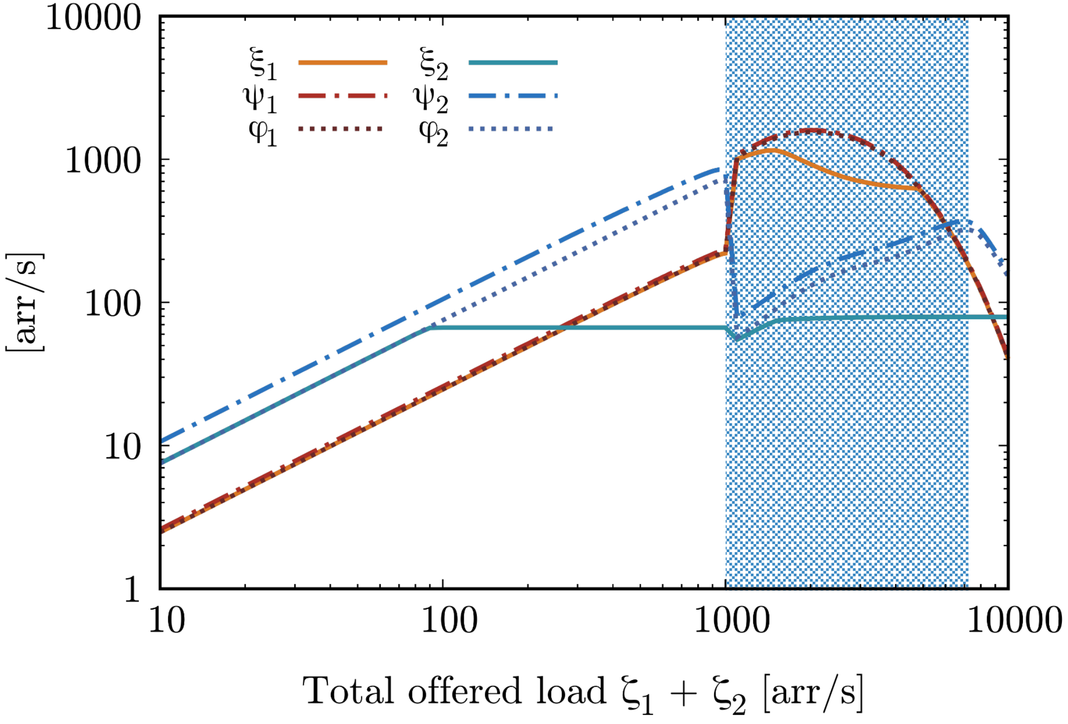
Scenario	Traffic share (%)		Dedicated capacity (%)		Dedicated preambles N (out of 54)		Dedicated connections (out of 200)		Timeout [s]	
	HTC	**MTC**	**HTC**	**MTC**	**HTC**	**MTC**	**HTC**	**MTC**	**HTC**	**MTC**
Sparse IoT	95	5	80	2	40	5	100	10	5	5
Dense IoT	5	95	75	5	10	40	40	100	3	1
Small Factory	75	25	50	20	30	10	50	50	5	0.1
Big Factory	30	70	10	50	10	30	20	150	5	0.1

IoT in a urban scenario

Sparse IoT



Dense IoT

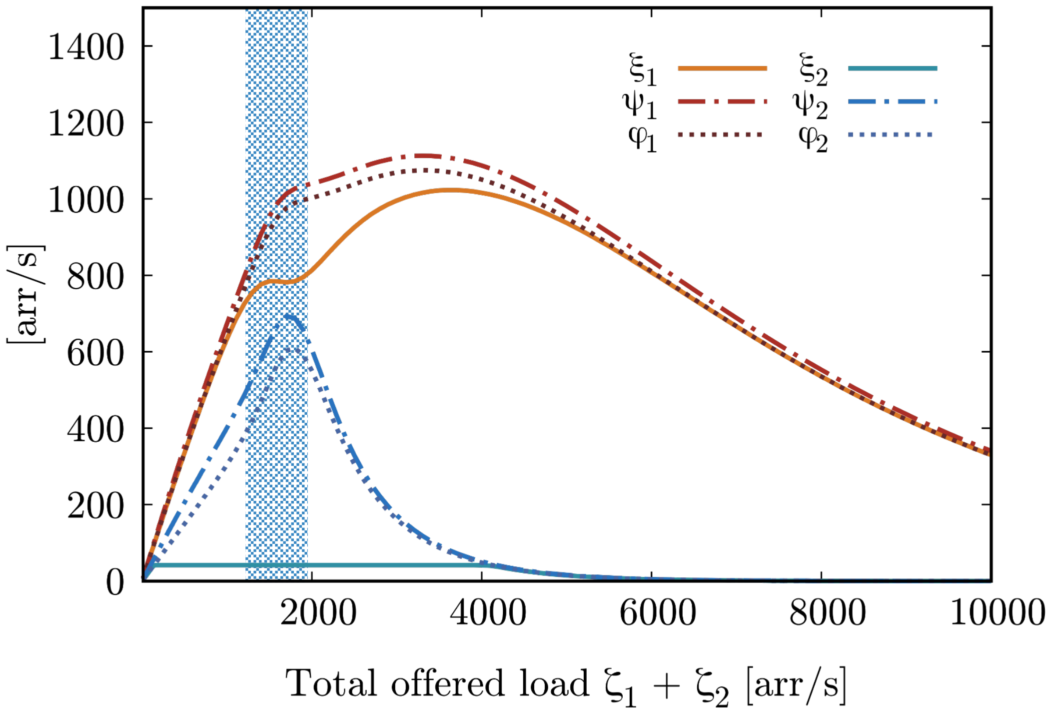


Sparse IoT: a BS serving a urban cell with mostly **HTC** traffic, and a small slice for IoT (**MTC**) traffic

Dense IoT: a BS serving a cell with mostly **MTC** connections with low traffic load, although a large part of the capacity is used by a few **HTC** devices

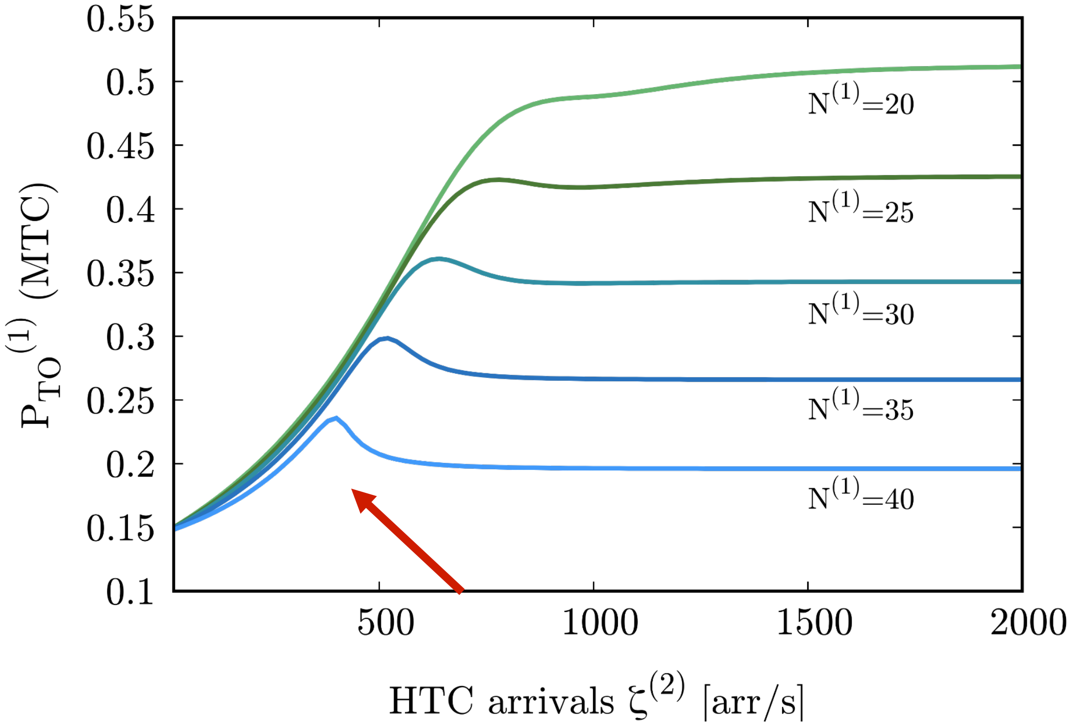
MTC in industry

Big Factory



Big Factory: A BS serving a private area with mostly **MTC** traffic, and a slice to handle **HTC** traffic

Timeout probability in Big Factory



Timeout probability for **MTC** at fixed **MTC** arrival rate ζ(1)=1000 arrivals/s.

Conclusions: opportunities and challenges

- Connected vehicles: a new adventure (at least for me, it just started)!

My future papers

  - RAN slicing + MEC assisted services
  - With 3GPP (5G and beyond)
  - Mobility seen as a service-enabler, beyond opportunistic networking
- ETSI, 3GPP and 5G groups are developing fundamental tools for connected vehicles

New services

  - Virtualization of network functions, MEC & Slicing
  - Key use cases for connected cars are under study
- Applications for connected vehicles are challenging network developers

New research opportunities

  - **Currently not enough bandwidth, delay guarantees, reliability, service control, etc.**
  - **Bandwidth in the radio access will increase, efficient and slice-aware mechanisms are needed**
  - **Wireless backhauling resources are still not enough, although the use of mmWaves is promising a concrete and solid solution**
  - **URLLC is key for delay and reliability, but it currently supports only a few users per cell**
  - **MEC services are also key, but they need to “follow” the users efficiently, more work is needed**

Thanks!

Contact:

Vincenzo Mancuso

IMDEA Networks Institute, Madrid, Spain

vincenzo.mancuso@imdea.org