Fifth Generation Communication Automotive Research and innovation

Deliverable D4.1
Initial design of 5G V2X system level architecture and security framework

Version: v1.0

2018-04-30

http://www.5g-ppp.eu/

Deliverable D4.1

Grant Agreement Number:	761510
Project Name:	Fifth Generation Communication Automotive Research and innovation
Project Acronym:	5GCAR
Document Number:	5GCAR/D4.1
Document Title:	Initial design of 5G V2X system level architecture and security framework
Version:	v1.0
Delivery Date:	2018-04-30
Editor(s):	Laurent Gallo, Thierry Lejkin, Bruno Tossou, Bernadette Villeforceix (Orange).
Authors:	Nadia Brahmi (Bosch), Ricardo Martínez, Ricard Vilalta (CTTC), Massimo Condoluci, YunXi Li (Ericsson), Apostolos Kousaridas, Panagiotis Spapis (Huawei), Toktam Mahmoodi, Jingjing Zhao (King’s College London), Rémi Theillaud (Marben), Vinh Van Phan, Ling Yu (Nokia Bell Labs), Sylvain Allio, Marc Bouillon, Laurent Gallo, Javier Gonzales Payo, Thierry Lejkin, Laurent Mussot, Pascal Rousset, Philippe Tamagnan, Bruno Tossou, Bernadette Villeforceix (Orange), Antonio Eduardo Fernandez Barciela (PSA), Per Lindberg (Volvo Cars).
Keywords:	5G V2X system design, 5G system architecture, 5G enablers, multi-link, multi-RAT, multi-operator, network slicing.
Status:	Final
Dissemination level:	Public

Disclaimer: This 5GCAR D4.1 deliverable is not yet approved nor rejected, neither financially nor content-wise by the European Commission. The approval/rejection decision of work and resources will take place at the next Review Meeting planned in September 2018, after the monitoring process involving experts has come to an end.

Abstract

Deliverable D4.1, entitled “Initial design of 5G V2X system level architecture and security framework” provides a first set of the mechanisms for Multi-Radio Access Technology (RAT) interworking, and Quality of Service (QoS) and traffic flow handling. Additionally, it is this WP’s first attempt to provide an overall security and privacy solution for 5G Vehicle-to-Everything (V2X) communications. Using these as inputs and building on the network slicing and cloud architectural paradigms Deliverable D4.1 describes the first version of the 5G V2X architecture provided in terms of this project addressing the multi-operator and multi-Original Equipment Manufacturer (OEM) scenarios. This deliverable serves as the basis for the remainder of the project and the forthcoming deliverables where the proposed architecture will be further refined and the described mechanisms will be evaluated and analyzed in more details. This deliverable is not intended to overlap ongoing V2X works in 3GPP but to be complementary to proposed solutions in this Standard Development Organisation.

Executive summary

The goal of the Work Package 4 in 5GCAR is to study and propose an evolution of the 5G Service-Based Architecture (SBA) that will be suitable to support the requirement of V2X applications. These include ultra-reliable and low-latency communications, critical security and privacy measures, as well as the need for QoS and flow management in a multi-link and multi-RAT configuration. This is not currently possible nowadays, because of the limitations of the current generation communication networks. The outcomes of the project will contribute to the evolution of the standardisation of the 5G architecture, in order to make it a true enabler for vehicular applications.

In this deliverable, we analyse the technical and non-technical issues brought by V2X requirements that challenge the current architectural conception, elaborate technical components each dealing with a specific limitation, and finally combine them into a preliminary proposition for the 5GCAR architecture.

We identify the general technical challenges brought by vehicular applications in five main technical areas, including V2X main characteristics (mobility, and simultaneous requirements for massive numbers of ultra-reliable, low latency, high bandwidth communications), multiple access network connectivity (including the management of multiple network slices), resilience requirements, security and data privacy, and roaming between operators. We then focus on the use cases specifically considered in 5GCAR, including the technical challenges associated to each of them, as well as their ecosystem and business implications.

Based on the outcomes of this analysis, 14 technical components (TC) have been elaborated: Road Side Unit (RSU) enabled Smart Zone (SM-Zone) for V2X communications, Self-Organising Network (SON) based multi-mode RSU for efficient QoS support and congestion control of V2X communications, several contribution addressing efficient SL and Uu multi-connectivity for high reliable and/or high data rate Vehicle-to-Vehicle (V2V) communication, Dynamic Selection of PC5 and Uu Communication modes, Infrastructure as a Service to automate deployment, 5GCAR Core Network mobility for V2X communications, Location aware scheduling, V2X service negotiation, Multi operator solutions for V2X communications, Evolution of Infrastructure-based communication for localized V2X traffic, Edge Computing in Millimetre Wave Cellular V2X Networks, Use case-aware Multi-RAT/Multi-Link connectivity, Security and privacy enablers for 5GCAR V2X.

These technical components are then integrated into a coherent first proposition of the 5GCAR architecture, and the impact of each technical solution on the existing Service-Based Architecture is presented. The purpose of such analysis is to identify which existing network functions need to be extended to support new functionalities, and which new Network Functions need to be added to the 5G architecture to serve the specific V2X requirements.

Contents

List of abbreviations and acronyms

3GPP	Third Generation Partnership Project
5GAA	5G Automotive Alliance
5G-PPP	5G Private Public Partnership
5GC	5G Core Network
5QI	5G QoS Indicator
AA	Authorisation Authority
AAA	Authentication, Authorisation and Accounting
AaSE	AIV-agnostic Slice Enabler
AF	Application Function
AIV	Air Interface Variant
AMF	core Access and Mobility management Function
AN-I	Access Network - Inner
AN-O	Access Network - Outer
API	Application Programming Interface
AS	Application Server
AT	Authorisation Ticket
AUSF	Authentication Server Function
B2B	Business-to-Business
BBF	Broadband Forum
BS	Base Station
BSM	Basic Safety Message
BSR	Buffer Status Report
CA	Carrier Aggregation
CAM	Cooperative Awareness Message
CN	Core Network
CP	Control Plane
CRL	Certificate Revocation List
CTL	Certificate Trust List
CUPS	Control and User Plane Separation
C-V2X	Cellular-V2X
D2D	Device-to-Device
DANE	DASH-Aware Network Element
DASH	Dynamic Adaptive Streaming over HTTP
DCI	Downlink Control Information
DDOS	Distributed Denial Of Service
DECOR	Dedicated Core Network
DENM	Decentralized Environmental Notification Message
DL	Downlink
DN	Data Network
DNS	Domain Name Server
DSRC	Dedicated Short Range Communications
E2E	End-to-End
EA	Enrolment Authority
EATA	European Automotive Telecom Alliance
ECA	Enrolment Certificate Authority
ECIES	Elliptic Curve Integrated Encryption Scheme
ECDSA	Elliptic Curve Digital Signature Algorithm
EE	End Entity
eDECOR	enhanced DECOR
eMBB	Enhanced MBB
eMBMS	evolved MBMS
eNB	evolved Node B
eV2X	Enhanced V2X
EPC	Evolved Packet Core
ETSI	European Telecommunications Standard Institute
FTP	File Transfer Protocol
GBR	Guaranteed Bit Rate
GSM	Global System for Mobile Communications
GSMA	GSM Association
gNB	g Node B
HARQ	Hybrid Automatic Repeat Request
HD	High Definition

HeNB	Home eNB
HO	Handover
H-PLMN	Home PLMN
HTTP	Hypertext Transfer Protocol
HSM	Hardware Security Module
HWN	Heterogeneous Wireless access Network
IaaS	Infrastructure as a Service
ICT	Information and Communication Technologies
IEEE	Institute of Electrical and Electronics Engineers
IoT	Internet of Things
IP	Internet Protocol
ITS	Intelligent Transportation System
ITU	International Telecommunication Union
KMF	Key Management Function
KPI	Key Performance Indicator
LBO	Local Breakout
LCCF	Local Chain Certificate File
LINP	Logically Isolated Network Partition
LIPA	Local IP Address
LOS	Line Of Sight
LPF	Local Policy File
LTCA	Long Term Certificate Authority
LTE	Long Term Evolution
L-GW	Local Gateway
MA	Misbehaviour Authority
MAC	Medium Access Control
MANO	Management and Orchestration
MBB	Mobile Broadband
MBMS	Multicast Broadcast Multimedia Service
MBSFN	Multicast Broadcast Single Frequency Network
MCS	Modulation and Coding Scheme
MEC	Multi-access Edge Computing
MIMO	Multi-Input Multi-Output
MME	Mobility Management Entity

mMTC	massive Machine Type Communications
MNO	Mobile Network Operator
MPTCP	Multi Path TCP
NAS	Non Access Stratum
NBI	Northbound Interface
NEF	Network Exposure Function
NF	Network Function
NFV	Network Functions Virtualisation
NR	New Radio
NRF	Network functions Repository Function
NSSF	Network Slice Selection Function
NG-C	New Generation Core network
OBU	On Board Unit
OEM	Original Equipment Manufacturer
OFDM	Orthogonal Frequency Division Multiplexing
ONAP	Open Network Automation Platform
OSS	Operations Support System
P2PCD	Peer-to-Peer Certificate Distribution
PCA	Pseudonym Certificate Authority
PCF	Policy Control Function
PCRF	Policy and Charging Rule Function
PDB	Packet Delay Budget
PDCCH	Physical Downlink Control Channel
PDCP	Packet Data Convergence Protocol
PDU	Protocol Data Unit
PGW	Packet data network Gateway
PGW-C	PGW-Control plane
PHY	Physical (layer)
PKI	Public Key Infrastructure
PLMN	Public Land Mobile Network
PLMNO	Public Land Mobile Network Operator
PoP	Point of Presence
PPPP	ProSe Per Packet Priority
PPPR	ProSe Per Packet Reliability
ProSe	Proximity Services
QAM	Quadrature Amplitude Modulation

QFI	QoS Flow ID
QoS	Quality of Service
RA	Registration Authority
RAN	Radio Access Network
RAT	Radio Access Technology
RBMT	Radio Bearers Mapping Table
REST	Representational State Transfer
RLC	Radio Link Control
RNTI	Radio Network Temporary Identifier
RRC	Radio Resource Control
RSS	Received Signal Strength
RSU	Road Side Unit
SaaS	Software as a Service
SAND	Server and Network Assisted DASH
SBA	Service-Based Architecture
SBI	Service-Based Interface
SCMS	Security Credential Management System
SC-PTM	Since Cell Point-to-Multipoint
SCS	Subcarrier Spacing
SDM-C	SDN Mobile Controller
SDM-O	SDN Mobile Orchestrator
SDM-X	SDN Mobile Coordinator
SDN	Software Defined Networking
SDO	Standard Development Organisation
SFC	Service Function Chaining
SIM	Subscriber Identity Module
SIPTO	Selective IP Traffic Offload
SL	Sidelink
SLA	Service Level Agreement
SMF	Session Management Function
SON	Self-Organizing Network
SPS	Semi Persistent Scheduling

SSP	Service-Specific Permission
SURB	Secondary Uu Resource Block
S-GW	Serving Gateway
TC	Technical Component
TCP	Transmission Control Protocol
TTI	Transmission Time Interval
UC	Use Case
UDM	Unified Data Management
UDR	Unified Data Repository
UE	User Equipment
UL	Uplink
UP	User Plane
UPF	User Plane Function
URL	Uniform Resource Locator
URLLC	Ultra-Reliable Low Latency Communications
USD	User Service Description
USDOT	United States Department Of Transportation
UTRAN	Universal Terrestrial Radio Access Network
V2I	Vehicle-to-Infrastructure
V2N	Vehicle-to-Network
V2P	Vehicle-to-Pedestrian
V2V	Vehicle-to-Vehicle
V2X	Vehicle-to-Everything
VNF	Virtualised Network Function
VoIP	Voice over IP
VRU	Vulnerable Road User
V-PLMN	Visited PLMN
WAVE	Wireless Access for Vehicular Environments
WG	Working Group

1 Introduction

The objectives of the deliverable D4.1, Initial design of 5G V2X system level architecture and security framework are to develop an overall 5G system architecture providing optimized end-to-end V2X network connectivity for highly reliable and low-latency V2X services, including security and privacy, QoS and traffic flow management in a multi-RAT and multi-link V2X communication system.

The first phase of this architecture study is the analysis of technical challenges related to V2X communications and services, then propose technical innovations dedicated to satisfy the specific requirements of vehicular applications, in terms of low latency, high reliability, scalability, flexibility and security/privacy.

The main outcome of the first D4.1 deliverable is the definition of an initial design of 5GCAR target architecture, the new network functions and new features offered by this target architecture will enable the 5GCAR project to contribute to the evolution of 3GPP architecture for eV2X (enhanced V2X, release 16 and beyond).

This document provides a preliminary version of the 5G V2X architecture which will be refined and enriched with more mature features to be presented in the final D4.1.

1.1 Objective and structure of the document

Chapter 2 entails the description of technical challenges and the identification of issues for the management of V2X applications, with the performance requirements of 5GCAR clearly identified in WP2 and described in D2.1. The overall technical challenges identified for future 5G V2X end-to-end architecture concern the provisioning of low latency and high reliability communications to handle 5GCAR use cases, where a new security and privacy architecture will guarantee the level of identities and data protection required in the network and within the vehicle.

Chapter 3 of this deliverable describes the drivers of these architecture studies, including a detailed description of 5GCAR use cases requirements, combined with the description of potential business models for the 5G ecosystem, and entails the description of the new technical components proposed to enhance the 5G architecture to sustain V2X.

Chapter 4 is the part of the document which gives a complete description of the technical components, including the identification of the problem, prior art related to the technical component, and the solution, with clear identification of the new features to be added for each standard 5G network function. For the initial design of the 5G architecture, 14 technical components have been proposed by 5GCAR WP4.

In chapter 5, a first target 5GCAR architecture will be proposed. This analysis includes, for each technical component, the description of the features and their impact on the current network functions standardized in the 5G architecture [3GPP17-23501], with detailed information about each NF, regarding control plane and user plane (interfaces, algorithms, with use case examples). For each technical component, a detailed message flow illustrates the described procedure.

Finally chapter 6 will conclude the document and highlight the main relative to the 5G architecture to support V2X communications.

2 Challenges and issue identification

2.1 V2X communications main characteristics and KPIs

The following characteristics were already specified as the target of 5GCAR [5GC17-D21]:

- Low end-to-end latencies below 5 ms [ETSI11-1028892]defined as the contribution by the radio network to the time from when the source sends a packet to when the destination receives it).
- Ultra-high reliability close to 10^{-5}, defined as the maximum tolerable packet loss rate at the application layer. A packet is considered lost if it is not received by the destination application within the maximum tolerable end-to-end latency for that application.
- Very large density of connected vehicles, defined as the maximum number of vehicles per unit area under which the specified reliability will be achieved. For urban environments, the vehicle density can reach 1000 to 3000 vehicles/km^2. The values can be estimated between 100 to 500 vehicles/km² for highway and 500 to 1000 vehicles/km² [ERF11] for suburban scenarios environments.
- Accurate positioning, measured in cm, and defined as the maximum positioning error tolerated by the application. To position a car on a lane, an accuracy of 30 cm allows for additional errors due to different car widths and lane widths. A positioning accuracy of 10 cm for a Vulnerable Road User (VRU), i.e., a pedestrian or a bicyclist, is requested while taking into consideration the energy efficiency of the end user devices.

The architecture is supposed to support different kind of services, which have various of characteristics/requirements. In [5GC17-D21], 5 typical Use Cases (UCs) were analysed, leading to following Key Performance Indicators (KPIs) specified:

Table 2.1 – Use Cases Key Performance Indicators

KPI	Lane merge	See-through	Network assisted vulnerable pedestrian protection	High definition local map acquisition	Remote driving for automated parking
Availability	V2I/V2N 99% and for V2V 99.9%	99%	99% to 99.99%	99% - 99.99%	99,999%

Mobility	0-150 km/h	0-30 km/h	0-100 km/h	0-250km/h	30-50 Km/h
Data rate	6,4 Mbps	15 to 29 Mbps, according to the configuration of the video streaming.	128 kbps	affected by traffic density	15-29 Mbps per camera
Latency	30 ms	50 ms	60 ms	30 ms	5 ms
Reliability	99.9%	99%	99% to 99.99%	99% - 99.99%	99.999%

The challenge of providing fast and reliable V2X communications in order to enable a broad range of applications and services, including those which require ultra-reliable and low-latency communications (URLLC), for vehicle users has long been identified as one of the key enablers for 5G [MET13-D11]. The most essential use cases of V2X are of course related to road safety and road efficiency but ultimately it is aimed to support autonomous or remote driving use cases in which handling of vehicles is delegated to vehicles themselves with possible remote control. In this regard, the targeted use cases of 5GCAR [5GC17-D21] are good representatives of challenging use cases and service requirements for 5G V2X. In addition, all infotainment or enhanced mobile broadband applications and services need to be provided to vehicle users.

Hence, it is anticipated that 5G networks will be able to support a massive amount of diverse V2X service data flows with different characteristics and QoS requirements in rather flexible, scalable and efficient ways. The basic characteristics of V2X service data flows include e.g.;

- different type of user devices and radio access modes for End-to-End (E2E) communications, e.g.: V, P, RSU; V2V, Vehicle-to-Infrastructure (V2I), Vehicle-to-Pedestrian (V2P), Vehicle-to-Network (V2N) over Sidelink (SL) or Uu with single- or multi-connectivity;
- different transmission periodicities, e.g.: ranging from a few milliseconds to hundreds of milliseconds, or dynamic event-triggered traffic;
- different packet sizes and per-packet access classes or priorities.

The QoS requirements are provided in terms of, e.g., data rate, packet delay or latency, and reliability for different access or priority classes. Note that the priority and reliability levels may be given as per a logical user connection, a service flow, a bearer service or as per an E2E application packet. The latter refers to, e.g., ProSe Per Packet Priority (PPPP) and ProSe Per Packet Reliability (PPPR) set by the application layer per an E2E application packet being passed down to the radio stack to transmit over SL at the transmitting User Equipment (UE) in 3GPP Rel-15 V2X [3GPP18-36300].

Hence, there is a need for flexible and robust schemes to manage such diverse V2X data flows to ensure an efficient support of challenging use cases such as those of 5GCAR. This may explore possible use of all applicable radio access technologies as well as access and transmission modes in a tight interworking fashion in order to enable and facilitate efficient and secure V2X communications. In particular, network slicing based architectures and interfaces, including SL and Uu radio interfaces, may need to be developed, enhanced and optimized for targeted V2X support. The following technical and practical issues need to be taken into considerations:

- multi-service: to support a broad range of applications and services including challenging URLLC as described above:
- multi-device: road users are equipped with different types and numbers of radio access devices with different capabilities including cellular access capabilities (with/without SIM – Subscriber Identity Module) for targeted applications;
- multi-operator or multi-tenancy: different user devices which need to communicate with each other on road may subscribe or register to different operators or tenants; also different operators or tenants may share common network infrastructure deployed on roads and spectrum resources (ITS – Intelligent Transportation System - infrastructure, roadside units, etc.);
- security and advanced networking services to ensure adequate authentication and authorisation of vehicles devices as well as integrity and, if applicable, confidentiality of information flows and contents.

2.2 Multiple access network connectivity, management of slices, and identity management

As we are slowly approaching the 5G era, providing ubiquitous connectivity in vehicles will be a challenging issue. 5G is expected to seamlessly integrate heterogeneous multi-tier networks with difference underlying network technologies with the aim of providing ubiquitous internet connectivity. Although quite a few of the recent studies have confirmed the viability of Heterogeneous Wireless access Networks (HWNs) to support V2X communications, the presence of multiple underlying networks with difference coverage areas, characteristics and QoS requirements pose a plethora of challenges when vehicles move at high-speed. For example, managing the frequent occurrences of vertical (inter-technology) handovers due to high speed vehicles crossing multiple diverse networks, is a considerable issue in terms of providing unhampered QoS and handover performance. To address the seamless connectivity in heterogeneous multi-RAT environments, the following technical issues need to be taken into consideration:

- Mobility management: While moving from one location to another, vehicles will need to cross a number of heterogeneous small networks and, due to their very high-speed of movement, vehicles, may quickly move in and out of the coverage areas of underlying networks performing handover activities. Ensuring seamless network connectivity requires faster and seamless vehicular handovers with the underlying network systems. Thus, more network discovery time implies more overall handover delay and deteriorated communication performance, which includes packet loss and potential session disruptions. Therefore, the efficient discovery of underlying networks is one of the key challenges in V2I communication over 5G networks.
- Data dissemination: In V2I communications, vehicles communicate with the RSUs through short message exchanges that are known as beacons. Through the beacons, vehicles can exchange diversified information, such as, their geographical locations, movement directions, acceleration, velocity etc., with the RSUs, which can periodically broadcast this information for other vehicles in the vicinity to know about them. In addition to these beacon messages, vehicles endowed with embedded sensors can collect environmental information and communicate it to ITS servers through RSUs. Moreover, vehicular communication also needs to provide appropriate support for different bandwidth-hungry applications such as mobile TV, 3D multimedia, gaming, VoIP, etc., which require effective handling of large volumes of very high speed multimedia traffic. As Cisco predicts, the global internet data growth is expected to reach 1.4 Zetabytes by 2017, which is likely to impose significant pressure on the existing V2X communication infrastructure. Vehicular network operators are thus left with an uphill task to find appropriate mechanisms to tackle this mammoth data capacity issue.
- Security issue: Future vehicular communication will support diversified cooperative applications and services and vehicles will need to transmit many kinds of sensitive data, such as vehicle ID, position and speed, which need to be private and secured enough for a broad acceptance of the whole communication system. In multi-RAT V2X communications, there are two broad categories associated with possible security problems: attacks on the user and attacks on the system. Attacks on the user seek to cause vehicle crashes, congestion, making the driver take a wring direction or reducing the user’s faith in the system due to unreliable messages. Attacks on a V2X communication system are aimed to track instantaneous locations for particular vehicles and to falsely generate reports of misconduct from a vehicle, causing innocent drivers to be revoked or receive administrative sanctions.

2.3 Network resilience

2.3.1 General definition of resilience

A general definition of network dependability can be the ability to maintain Service Level Agreement (SLA) and minimize service downtime during the whole lifecycle of a service.

However, the complexity of modern networks and the major place of telecommunication in everyday life is a severe challenge for dependability. Network operators have endured, during these past years, several major outages. Complex phenomena of disaster amplification are now a common place:

- snowball and domino effects where the congestion of one node results in overload of others, and progressively leads to global service failure
- impatience and massive user re-registering, protocol repetitions or buffer bloat effects
- oscillation phases of the whole network

As a response to that, many publications (e.g. [SH06]) state that moving from a system designed for robustness to one that supports resilience represents a significant strategic shift. Whilst systems have commonly been designed to be robust – systems which are designed to prevent failure – increasing complexity and the difficulty it poses to fail-proof planning have made a shift to “resilience” strategically imperative. A resilient system on the other hand accepts that failure is inevitable and focuses instead on early discovery and fast recovery from failure.

Accordingly, operators must reconsider the traditional techniques of robustness based on high-availability, double dimensioning, and absence of cooperation between nodes.

Based on such a paradigm, network “resilience” can be divided into several domains:

- to provide a highly available architecture end-to-end, including redundancy, connectivity and diversity,
- to identify challenges, possibly preventing faults, mitigating failures when possible,
- to control high load by relevant dimensioning, various abatement and/or inter-node coordination techniques in order to make the best use of available resources and maintain performance,
- to offer and maintain network manageability to operational teams,
- to heal manually or automatically injured resources and recover to a reference configuration,
- to manage in the long term the network resources (hardware, software) by appropriate software upgrades or updates, troubleshooting and tests.

2.3.2 New challenges of 5G and V2X

The arrival of 5G increases operators’ difficulties in providing sufficient network resilience, consequences of the new requirements of massive usage (“we want everything”), performance (“quickly”), ubiquity (“everywhere”), uncertainty of behaviour (“at any time”). These new constraints lead to reinforce the following 3 aspects:

- Monitoring: fault and performance monitoring, with almost mandatory failure prevention
- Reasoning: correlation, network awareness, decision (placement, remediation, recovery)

- Automation, orchestration and flexibility

In this approach, new technologies (new radio, service-oriented core network, and virtualisation) appear both as threats and opportunities

In the case of V2X applications, additional specificities can be identified:

- Environment strongly anchored in the public space, including core network sub-domains, with the inherent risks of disturbance (weather, geography, vandalism, etc.)
- Important number of use cases (e.g. lane crossing, takeover, and parking) which complicate each service V2X. Both (environment & use cases) broaden the scope of unknown unknowns
- Very high KPIs: bandwidth, latency, location accuracy etc.
- High degree of risk, in particular because of possible impacts on human life
- Exigency to alert end users early and in a relevant way
- Open ecosystem with an important number of stakeholders (automobile, I.T, telecom, communities)
- Need to take into account several resilience trade-offs (affordable costs, protection of private life, complexity, place available on the public environment)

As a consequence, each resilience domain should more specifically focus on:

- Architecture: redundancy (compute, storage, network), connectivity and diversity of any kind (geographical location, path, technology)
- Monitoring: fault detection and context-awareness, with automated fault remediation
- Load control: smart coordination of various techniques (scaling, throttling, overload notification, etc.)
- If possible, automated recovery, without forgetting long-term resilience (e.g. in-service software update).

All new technologies should indeed be associated in this approach: virtualisation (orchestration, infrastructure), radio and core network.

2.4 Security and data privacy

Security and Privacy for Cellular V2X (C-V2X) need to be considered across multiple domains (both within a domain and at domain external interfaces):

- User-equipment, radio access network, core network and external network domains.

- Infrastructure and virtualized resources or functions (either owned by mobile network operators or other tenants: 3rd parties/virtual network operators) domains.
- Software Defined Networking (SDN) / Network Function Virtualisation (NFV) management plane, control plane and data plane domains.

Some Security and Privacy threats are not V2X-specific, but more general in 5G context, though they certainly apply to V2X:

- Denial of Service attack, e.g. by jamming radio access network, overloading the control plane with signalling exchanges, installing malwares in devices (e.g. Botnets);
- Threats to anonymity and privacy of data across radio access network;
- SDN/NFV controllers/orchestrators being compromised or abused (e.g. tenants' impersonation) [ONF15-TR511]
- Network slice poor isolation leading either to data leakage, or to resource exhaustion.

Other security and privacy threats are V2X-specific, or are rather more critical in V2X context, e.g. for Safety or autonomous driving applications:

- Availability/resiliency, throughput and/or latency are of the utmost importance.
- Authentication, Authorisation and Accounting (AAA) are also critical security functions for verifying that any V2X actor (On Board Unit (OBU) or RSU UE, Multi-Access Edge Computing (MEC) application, etc.) behaves properly (in the current time and location context) and according to the permissions they have been granted by some V2X authorities.

  Such V2X permissions management may call for device enrolment (to global – E.U. or per-country - or delegate authorities), at which time they are granted permissions: regular vehicles vs. special vehicles (emergency vehicles, road-operator vehicles, police vehicles, public transport/transit vehicles, etc.); roadside units in toll zone, for traffic lights, for advertising roadworks, etc.
- V2X messages replay

Liability in a multi-operator environment, or more generally multi-tenants, multi-services, multi-countries/legislations, environment also needs to be considered.

2.5 Roaming between operators

V2X communication needs to de facto be addressed in a multi-operator environment, since the end devices (vehicles) cannot be assumed necessarily to operate under a single operator.

Almost all vehicles are roaming today, using home routing. Many features we are working on requires interaction with the serving Public Land Mobile Network (PLMN) on user plane.

Furthermore, a Visited PLMN (V-PLMN) does not normally honour QoS for roaming UEs using home routing.

- Maintain QoS

  To support a particular V2X services, some certain level of QoS is necessary. In roaming case, this would be problematic, or even impossible since the user plane is split between the serving PLMN and the Home PLMN (H-PLMN), and in most cases there would be an intermediate network in between the PLMNs. Furthermore, the user plane for a roaming user would reach an application server via the H-PLMN, meaning it would add in many cases a substantial delay for the communication.

  Low latency and high reliability QoS requirements that many V2X use cases have should be considered in a multi-operator environment, as well as when there is a need for roaming from one operator to the other.

- Monitor QoS

  The QoS being achieved needs to be monitored, so that the application can take proper action according the QoS situation. Similar as for above, the user plane is split on several networks making it difficult, or impossible to get the whole picture, so a serving PLMN for a roaming UE would only see a small fraction of the user plane due to the home routing. Furthermore, the V-PLMN would not normally apply the QoS requested from the H-PLMN.

- V2X User Service Description (USD) provisioning via V1 reference point

  Due to the home routing, user plane connectivity to Multicast Broadcast Multimedia Service (MBMS) servers in the V-PLMN is not possible, or requires complex connectivity from H-PLMN to servers located in V-PLMN domain, so user plane solutions to obtain for example information about MBMS channels in V-PLMN is hard/unrealistic to realize if home routing is applied.

- Local service information

  To access local service, the UE should interact with Domain Name Server (DNS) in serving PLMN, and potentially with servers in V-PLMN, this is hard/unrealistic to realize if home routing is applied.

  To just extend the roaming interfaces between PLMNs or introduce new roaming interfaces between PLMNs, with an accompanying roaming agreement would likely not be sufficient since roaming interfaces and agreements are complicated to get into place, and still for home routing user plane would be split between PLMNs, so if just 'normal' roaming procedures with home routing of user plane are applied, there would be routing challenges to get traffic from H-PLMN to servers in V-PLMN.
  In other words, roaming with Local BreakOut (LBO) in V-PLMN would be beneficial (if not really a must), however LBO is not commonly used today.

- OEM dedicated connection

Vehicle OEMs currently prefer a single protected connection between the vehicle and the OEM backend servers to have all communication with the vehicle via the backend servers, this to protect the vehicle since an OEM would be liable for malfunctions. The features we are working on would likely be negatively affected by the additional latency if this setup is continued. In other words, a second connection to the vehicle would likely be needed.

- Cross border V2X communications cases where the UE crosses the borders of a country, could be also considered as a multi-operator problem.

Nowadays, communicating mobile devices attach and operate under one operator only. For this reason during power-on they will search and select the operator with whom they have a subscription (H-PLMN) to attach to. In cases their operator is not present in an area (e.g., usually when the end device is abroad), the end device will perform roaming and attach to a V-PLMN.

This process was quite adequate up to now, however it is inefficient for vehicular communications handled by 5G cellular networks. In particular, it has been identified that vehicle-to-vehicle or vehicle-to-pedestrian communication needs to de facto be addressed for a multi-operator environment, since the end devices cannot be assumed necessarily to operate under a single operator. Obviously this requirement raises a number of technical challenges which need to be addressed by cellular networks, so as to offer low latency and high reliability QoS requirements. Also roaming case should be considered for vehicles that travel to other countries, including the cross border V2X communications cases where the UE crosses the borders of a country. In the latter case the service shouldn’t be disturbed by the crossing of the borders, while switching from one operator to the other due to the ending of the coverage.

2.6 Accurate user location

Accurate Location is a one of the key enablers for several use cases related to V2X and V2V systems. However, having a solution that can provide very accurate (centimetre level of precision) location estimate in a very short time (low-latency) is yet a challenge. Despite of the localisation technology (satellite-based or cellular-based), the problems rise from two front: 1) the errors affecting the position-related measurements due to the propagation of radio signals in a multipath environment and, 2) the delay to process and/or feed location-information to location server.

5GCAR work package 3, task 3.3 is currently leading studies to bring accurate positioning techniques thank to new triangulation algorithms running in multiple base stations with Multi-Input Multi-Output (MIMO) beamforming.

3 Drivers

3.1 Use case analysis

The definition of the 5G architecture must be motivated by the needs and the market orientation, more precisely in the case of V2X, this is translated into automotive use cases and an evolution of the current business model related to the connected and autonomous vehicle. Both topics are studied in WP2 of 5GCAR project and have served as an input for this deliverable.

On one side, for the use case definition, 5GCAR has selected 5 use cases to highlight the new challenges set by the automated and connected driving application on the communication systems. A short description of each of these use cases is provided in the table below. A detailed description can be found in D2.1 [5GC17-D21].

Table 3.1 - description of each use cases

Use Case	Short Description
Lane merge	A subject vehicle is coordinated with remote vehicles that are on the main lane in order to merge smoothly into the main lane without collisions and with minimal impact on the traffic flow. The trajectory recommendations are computed based on road user attributes such as position and speed. Connected cars continuously transmit their positions, speed, and driving direction. For the integration of non-connected road users, a fixed camera installation near the intersection is used. All road users (connected and non-connected) are detected and tracked in the camera images. The estimated road user attributes (positions, orientation, speed, etc.) are transferred to the global coordinate system and transmitted to the data fusion function which merges data originated from the camera system and from connected road users.
See-through	This is a form of a cooperative perception system that exploits the exchange of real-time sensor data (detected objects or video) between nearby vehicles via wireless communication. The data representing the scene in front of the vehicle ahead of a line is captured by a camera vision system and transferred to the rear vehicle to allow it to see through the forward vehicle and bypass the occluded area.
Network assisted vulnerable pedestrian protection	To detect the presence of vulnerable road users with the help of accurate positioning technology. Thanks to the exchange of local sensor/camera information between users via wireless communications the overall system will determine VRU positioning based on cellular radio signals, Global Navigation Satellite System or sensor/camera data. All this location information will be processed

	from multiple users for alert generation to vehicle drivers or autonomous driving vehicles with high accurate positioning.
High definition local map acquisition	The goal is to update the local dynamic map of vehicles on the move. An application server gathers all the information from different sources according to different layers starting from the map provider (static layer) to the cooperative sensing of the different vehicles available (temporary and dynamic layer) to build an optimal route map. This information is organized and divided in polygons for being distributed by push/ polling methods to the vehicles. Polling methods are used on a regular frequency and push methods by major changes or hazardous events.
Remote driving for automated parking	A remote cloud server drives remotely a vehicle from the "last mile" near a parking to the parking entrance to the parking spot without a human driver. The server provides to the vehicle that is remotely driven the appropriate trajectory and manoeuver instructions for the efficient and safe parking. The decisions and remote driving are supported by real time video streaming and sensor data that are sent from the remote vehicle and/or parking facilities to the cloud server.

The 5G V2X architecture defined will serve to these needs by respecting the main KPIs remarked in D2.1 deliverable and reflected in Section 2 of this document.

On the other side, use cases will only be deployed if there is a business model supporting their costs. This is way in Task 2.2 of the 5GCAR project, business model aspects of 5G and V2X have been analysed. This section provides a high level summary of the findings in the intermediate report from that work, [5GC18-D22].

3.2 Technical challenges associated with use cases

There are some new technological components in 5G, which are covered on an architecture point of view in Section 4 of this deliverable and which have been analysed and evaluated in how they could impact the business models in deliverable 2.2. The study has found that some new 5G technologies (such as network slicing, MEC, or the arrival of PC5) have the capacity to disrupt current eco-systems and value chains or enhance existing business models.

It must be taken into account that even if the use cases may be appealing and the new technological components are very promising in terms of a business model definition, all the development of the technology may be at stake if the 5G architecture is not able to solve the challenges described in Section 2 of this document For the service provisioning, a seamless transition between operators and countries is needed to enable bigger markets. These elements have been solved for Mobile Broadband (MBB) but are highly expected for V2X use cases. The coverage availability, and the combination of different access technologies (PC5, Uu, others) are also elements to boost or jeopardized new services depending on the service provided on the top of the 5G architecture for V2X.

3.3 Ecosystem and business perspectives

By analysing a small set of services, the eco-system and value chain around these services have been evaluated. From the study, it is clear that 5G could have a major impact in enabling new features in these services, and also enable new value chains. The study finds that value chains will change from being fairly linear with traditional customer/supplier roles, to being more dynamic and network oriented.

The automotive industry is a very competitive sector, where costs are analysed and innovations are adopted only if there is a business case supporting it, or through a regulatory mandate. As an example of a regulatory mandate, in Europe from April 2018, the SIM card will become a regulatory obligation at least for the e-call service. This means that all cars in the market will be equipped with SIM cards (physical or virtual) from now on. It is also important to remark that generally the features related to safety should be made available to everybody and enforced by regulation. Comfort driving features may be used as brand differentiators, but also need to be safe.

The study has found that for most of the services analysed, 5G will provide enhanced functionality that could contribute to an increased service value, thereby improving the business case. That value could be generated by e.g. a guaranteed quality of service, more efficient delivery of high data volumes, lower latency thereby enabling new types of services.

3.4 Summary

- There are technical features in 5G that can enable new business models for various stakeholders in the value chain. Network slicing and mobile edge computing are examples of such features.
- Existing services as well as autonomous driving features and convenience services may be enhanced by 5G technologies, thereby building added value in these services.
- The value chain as it looks today (Q1 2018), may be disrupted by 5G, driven by new 5G technologies, as well as changing eco-systems around the connected car, where a rapid digitization of existing industries is complemented by new types of digital and industrial stakeholders.

4 Technical analysis: use cases analysis and solutions

The technical propositions made in 5GCAR WP4 consider the 3GPP Release 15 Architecture, a description of which is provided in Section 5.1 of this document, as the baseline. Based on the technical challenges identified in Chapter 2 and on the performance requirements of the use cases analysed in Chapter 3, 14 Technical Components are introduced to enhance the 5G architecture. In this chapter, we provide for each technical component a description of the problem, followed by the study of prior art, by the proposed solution, and its impact on the baseline architecture.

4.1 RSU enabled Smart Zone (SM-Zone)

4.1.1 Problem

Most of V2X service data flows related to road safety and road efficiency are local or specific to individual roads which are originated or terminated by road users on the same road or road area. Thus, keeping V2X communications on a specific road local to the road by having, e.g., local radio access and local E2E service data flows not being routed via Core Network (CN) is rather desirable for efficient utilisation of spectrum and network resources as well as reducing E2E latency.

Device-to-Device (D2D) communications over SL between vehicles users, also referred to as V2V, is local to the road and therefore preferable. It is assumed that SL is based on transmitter oriented broadcast based D2D communications using either network scheduled or UE selected resources from a configured shared resource pool, similar to SL or PC5 specified in 3GPP. This allows a capable receiving device to receive D2D communications from a transmitting device in its proximity without a need of D2D discovery and connection establishment with the transmitting device beforehand. That is, as long as the receiving device is provided with a valid resource pool to monitor and receive D2D scheduling assignments sent by any transmitting devices within its proximity, the receiving device is able to receive D2D communications transmitted by those transmitting devices. There is no feedback control from the receiver side for radio transmissions. There is also no control plane needed over SL. This kind of SL therefore provides simple and instant packet access for proximity communications between user devices. However, there are some issues with this kind of SL as follows when considering applications requiring high reliability such as the advanced V2X applications of interest:

- the half-duplexing problem: if two user devices in proximity transmit simultaneously then they may not be able to hear one another;
- the problem of transmissions over SL without reassuring feedback control from the receiver side;

- the very well-known problem of collisions in the autonomous transmission mode when two user devices in proximity select the same resources and transmit simultaneously;
- the very well-known problem of hidden and exposed terminals [HL08]: two user devices in proximity are somewhat out of range or overhearing one another.

Thus, one of the main problems is how to adopt transmitter oriented broadcast based SL in a high mobility environment of vehicles with all the above problems coupled with the technical issues considered in Section 2 so as to ensure fast, reliable and efficient V2X communications for the targeted road safety and road efficiency related use cases of 5GCAR for examples.

Then, to enable and facilitate URLLC support for advanced V2X and to provide efficient V2X service flow and mobility management of vehicle user devices are part of the main objectives as well as the main problems.

4.1.2 Prior art

Some state of the arts on V2X communications are reported in [HL08], [HPY+14], [CDK+14] and [DHA+14]. There are many limitations and open issues with current IEEE 802.11p Dedicated Short Range Communications (DSRC) and IEEE 1609 Wireless Access for Vehicular Environment (WAVE) based V2X communication systems which use RSU for extended coverage. There is no clear evolution path to improve performance in term of latency, reliability, etc., and no URLLC support.

The most recent standardisation related development is Phase I and Phase II of V2X supports in 3GPP. Phase I has been completed for Long Term Evolution (LTE) in Rel-14 [3GPP17-23285]. Phase II is aimed for evolution towards 5G in Rel-15 and beyond [3GPP17-RP1740].

Phase I is specified for supporting limited messaging services, such as Cooperative Awareness Message (CAM) or Decentralised Environmental Notification Message (DENM). There is no advanced QoS support, no URLLC support. PC5 which is the radio interface specified for Proximity Services (ProSe) D2D SL in Rel-12 and Rel-13 is enhanced for high velocity, high UE density, shorter latency V2V support. PC5 enhancement for V2P is mainly from power efficiency point of view. The Uu interface is enhanced with Uplink (UL) Semi-Persistent Scheduling (SPS) and DL MBMS for V2X. Figure 4.1 and Figure 4.2 illustrate the overall network architecture and radio access network for supporting V2X in current LTE.

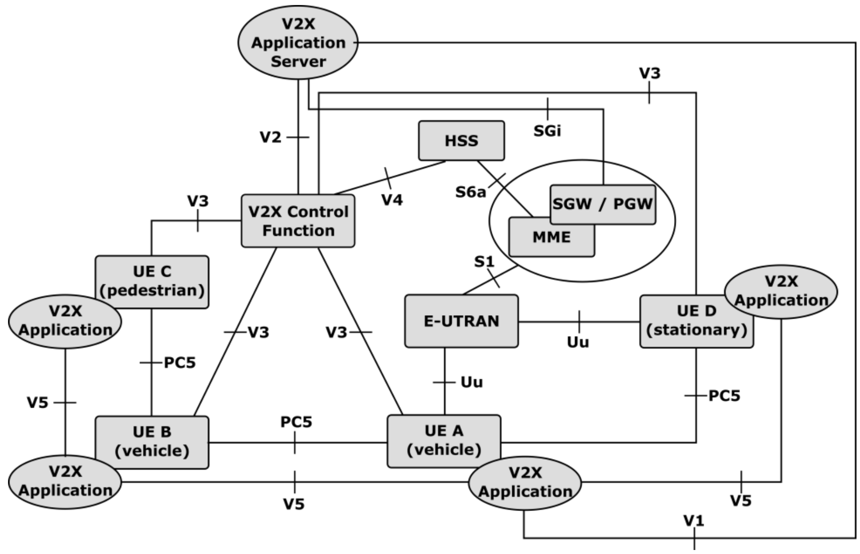


Figure 4.1: Network architecture for supporting V2X in LTE [3GPP17-23285]

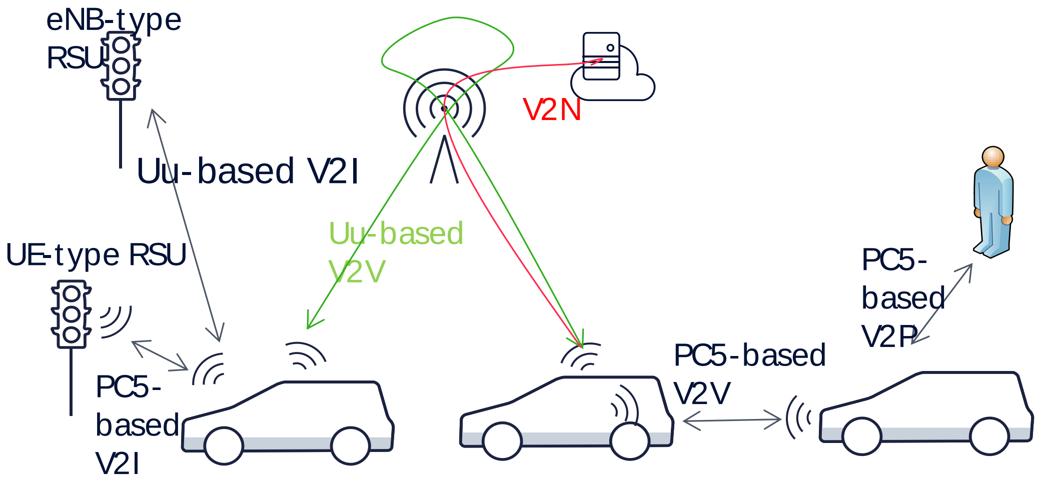


Figure 4.2: Radio Access Network (RAN) support of V2X in LTE

RSU, as defined in [3GPP16-23785] as “a stationary infrastructure entity supporting V2X applications that can exchange messages with other entities supporting V2X applications.” RSU is viewed as a V2X application receiving V2X messages via SGi, PC5 or LTE-Uu interface. 2 types of RSU are introduced: UE-type RSU and eNodeB (eNB)-type RSU. However, RSU is not considered as an architectural entity, but an implementation option. This is achieved by collocating a V2X application logic/server with some entities of the 3GPP system, as shown in Figure 4.3.

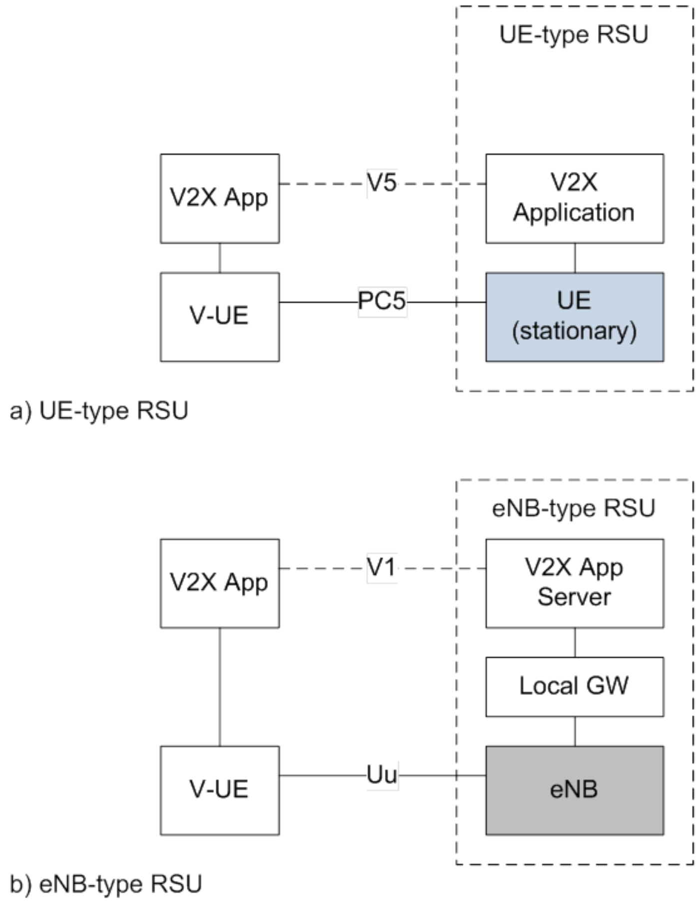


Figure 4.3: RSU implementation options: a) UE-type RSU and b) eNB-type RSU [23.285]

Phase II considers backward- and forward compatibility or coexistence of Rel-14 and Rel-15 and beyond UEs with possible resource sharing. The Carrier Aggregation (CA) and CA based packet duplication for SL are introduced for enhancing data rate and reliability over SL. Further latency reductions over SL in order to meet sub-10ms E2E latency requirements are also considered, including e.g. reduced SPS periodicities to sub-10ms, shortened Transmission Time Intervals (TTIs) to SL and higher Modulation and Coding Schemes (MCS), such as 64QAM, for SL.

4.1.3 Solution

Exploring smart use and deployment of RSU, a RSU enabled smart zone concept, denoted as SM-Zone, is introduced to address the issues described in Section 4.1.1. RSU is enhanced to be a functional entity of 5G network architecture for supporting advanced V2X communications. For instance, RSU may be considered as a small access point or base station enhanced with D2D capabilities so as to be capable of providing not only small cell coverage to cellular access capable devices but also SL for enhanced V2V communications. The working assumption is that a number of RSU is deployed along the road, e.g., mounted on the roadside lamp posts, as illustrated in Figure 4.5. The SM-Zone concept addresses network integration and deployment arrangement, network access and admission control along with features and enhancements introduced and specified to resolve at least the issues addressed in the previous sections for enabling fast, reliable and efficient vehicle communications in various mobility environments and road traffic conditions.

Figure 4.4 illustrates the targeted RSU with flexible capabilities and network functions, referred to as 5G-RSU. 5G-RSU functional options can be adapted to best serve targeted use cases using optimized combinations of D2D SL and Uu capabilities on C- and/or U-plane.

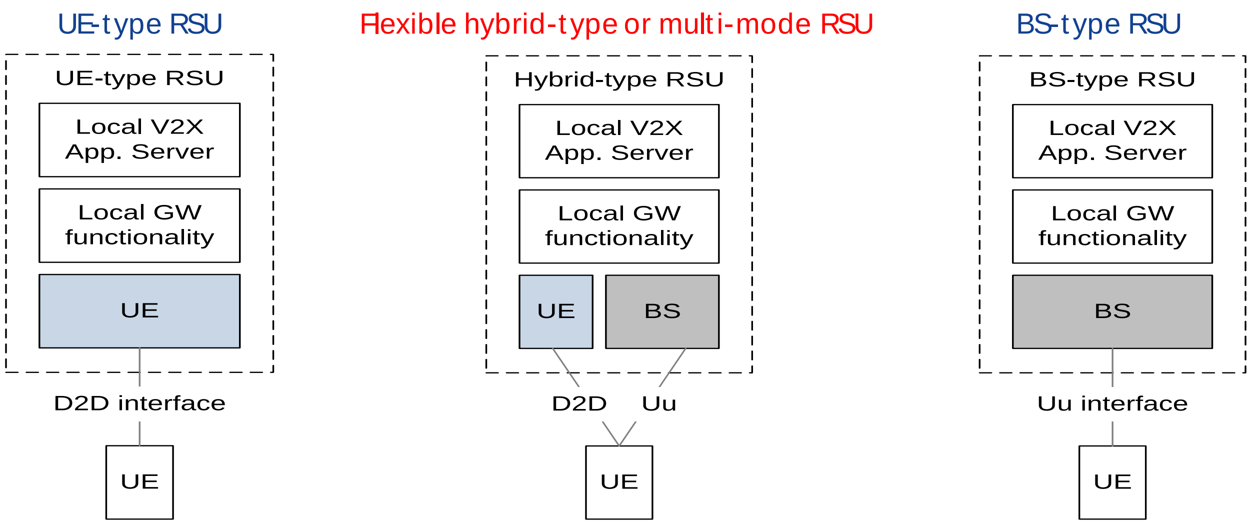


Figure 4.4: 5G-RSU with flexible capabilities and network functions

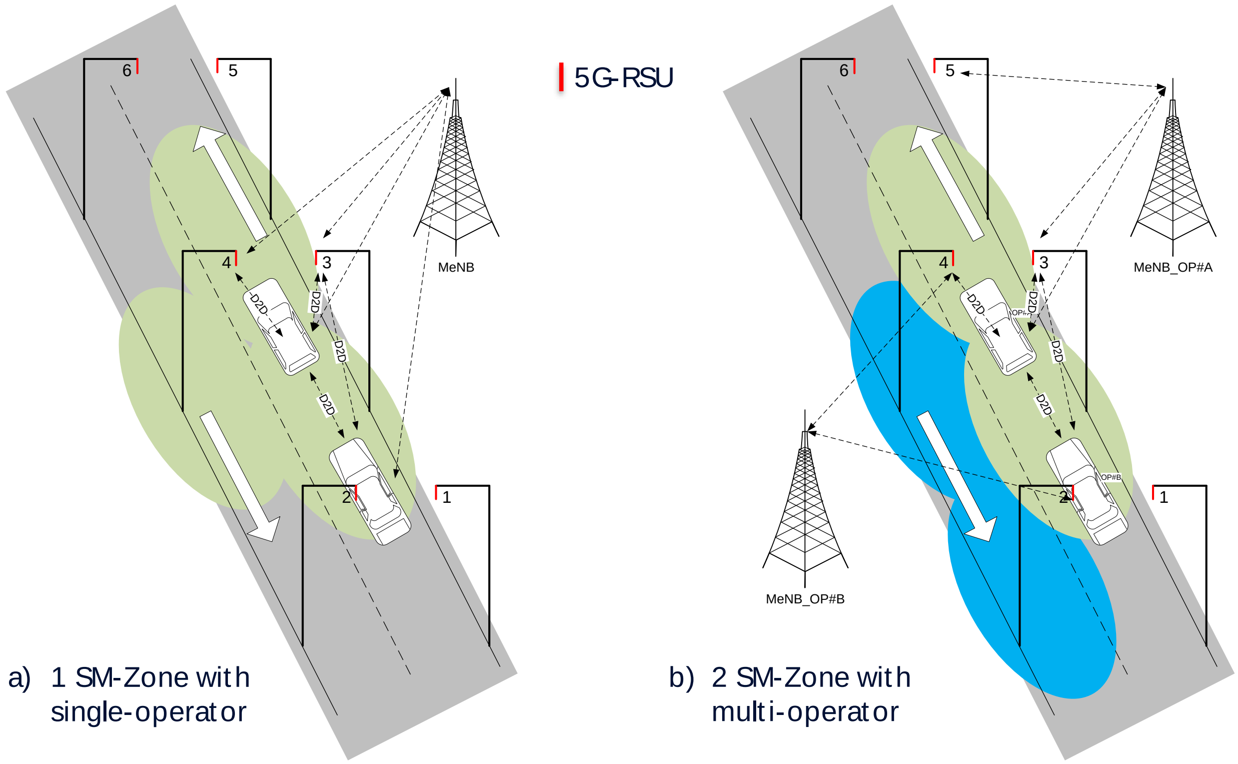


Figure 4.5: Illustration of 5G-RSU deployment and SM-Zone

The SM-Zone is defined and characterized by:

- 5G integrated radio service coverage layer local to road areas designated for V2X supports, in addition to operator specific macro-(micro/pico) cellular access layer(s)
- The road area radio coverage layer primarily provided by RSUs:
    - deployed along the roads, e.g., mounted on road side lamp posts with small footprints
    - functions and services adapted to enable and facilitate an optimized support of targeted challenging V2X use cases and requirements (e.g. URLLC related requirements, half-duplexing, performance or optimisation issues)
    - coordinated and controlled by a macro coverage 5G network
- UE, once admitted to a SM-Zone, may transmit/receive at least on the RSU provided road area radio coverage layer throughout the SM-Zone
    - SM-Zone based resource allocation and mobility control which may be decoupled with that of operator specific macro mobility cellular access layer(s) for simple and effective support of multi-operator, -service, -device, etc.

In a conceptual view, a SM-Zone is considered as a flexible and dynamically reconfigurable V2X service coverage which is local to a specific road or road area such as a highway, a suburban priority road. The local service coverage of a SM-Zone is provided by a subset of the deployed RSUs forming a road specific radio access layer under the general-purpose macro cellular coverage. The local deployed RSUs may also form a sub-network whether interconnected or not under assistance or control of a serving macro cellular network and/or under control of the local sub-network controller which is designated to provide a set of radio-access applications and services for targeted V2X users. The SM-Zone concept allows for flexible and efficient supports of V2V communications, adapted to diverse and demanding applications and services of vehicles communications as well as diverse and dynamic nature of road traffic conditions and patterns, as described in Section 2.

SM-Zone initial setup and reconfiguration

- Examples of SM-Zone and how SM-Zone solve the technical issues described above
- SM-Zone related network control and management procedures: adding, modifying or releasing a SM-Zone

UE related procedures

- SM-Zone access control
- Resource allocation and mobility management

4.1.4 Impact on architecture

This TC has certain impacts on the following 5G network functions and entities: the UE, the gNB, core Access and Mobility management Function (AMF), the User Plane Function (UPF), and the Application Function (AF). The following interfaces are impacted: Uu, Xn, PC5, and Nx. In addition, 5G RSU is introduced as an architecture entity with flexible and adaptable capabilities, functions and operations as described in Section 4.1.3, along with the SM-Zone service coverage. 5G RSU may communicate with UE via either Uu or PC5 or both; with other 5G RSU or gNB via Xn and towards UPF and AF via Nx, 5G RSU may provide the following enhanced network functions and services for examples:

- connection diversity over PC5:
    - relaying received D2D messages for UE individually;
    - rebroadcasting received D2D messages for a number of relevant UEs collectively;
- RSU-RSU data forwarding or distributing of local contents of V2X;
- MBMS localised along the road;
- connectionless packet access for V2X.

More details on the impact on the baseline architecture are provided in Section 5.4.1.

4.2 SON based multi-mode RSU for efficient QoS support and congestion control

4.2.1 Problem

First, let us consider some well-recognized road traffic scenarios concerning main roads in big cities like LA, Madrid, Beijing or Sydney which are often 2-way roads with 4-8 lanes per each direction and expected travelling speeds of 60-90km/h. In preferred road traffic situations, vehicles are expected to travel with speeds at around 80km/h and keep inter-car distance on the same lane at around 80m persistently. In rush hours, vehicles may have to travel with speeds at less than 10km/h and keep inter-car distance on the same lane at less than 10m. That is, in rush hours or congested situations, the road may experience some 10 folds increase of vehicle density compared to that of preferred road traffic situations. It is not unusual in big cities that vehicles and people aboard may be stuck in road traffic congestion for hours on the daily basis. On contrary the same road may be quite deserted during, e.g., night time.

The above implies that offered load of V2X communications for both local and remote access based E2E services of vehicles and users aboard, on certain roads varies greatly in different time periods on the daily basis.

- V2X service flows related to local access services such as V2V messaging and steaming over SL or Uu for road safety and road efficiency apps are specific and local to individual road which may cause local and temporary congestion on serving carrier(s).
- V2X service flows related to remote access services such as remote driving, map update or infotainment require extreme-MBB mobile connections which may cause high load to serving cellular network(s),

Thus, usage of SL connections for V2V communications and Uu connections for V2X may face capacity issues and need to be highly scalable and adaptable. This calls for a SON based reconfigurability on V2X support of the serving network: capacity and capability as well as access mode with fast enough reaction time.

4.2.2 Prior art

Referred to Section 4.1.2, two basic types of RSUs are considered in e.g., LTE: UE-type and eNB-type RSUs. RSU is adopted in IEEE 802.11p DSRC and IEEE 1609 WAVE based V2X communication systems. However, SON based adaptation of RSU modes on the fly in dependence of e.g. dynamic and time-varying road traffic scenarios, UE states or conditions, etc., is not considered.

4.2.3 Solution

Further exploring smart use and deployment of 5G-RSU described above, self-reconfigurable hybrid-type or multi-mode RSUs which are capable of operating in at least one of the following modes at a given time for examples: UE type RSU, Base Station (BS) type RSU, small-cell BS, and any combinations thereof are considered for scalable and efficient support of V2X. The main question is how to enable the serving network to configure and control adaptive operations of RSUs and corresponding UE behaviours on the fly, depending on dynamics of road traffic characteristics and service demands.

4.2.4 Impact on architecture

Referred to 4.1.4, this TC has no further impacts on the network architecture.

4.3 SL and Uu multi-connectivity for high reliable and/or high data rate V2V communication

4.3.1 Problem

It has been established that SL communications is necessary and preferred for many V2X applications and services, in terms of resource efficiency, latency reduction as well as out-of-coverage support. However, for V2V services that require high reliability and/or high data rate, the SL alone may not be sufficient to meet the requirements, especially in very complex environment of V2X communications, as characterized in Section 2. Therefore, possible use of multi-connectivity with SL and Uu link is explored for enhancing reliability as well as data rate for advanced V2X communications of interest in 5GCAR use cases.

4.3.2 Prior art

Dual- or, more generally, multi-connectivity in LTE and 5G allows UE to be configured two or more Uu radio connections with different BSs. This is different from the dual or multi-connectivity with SL and Uu connections for direct E2E communication in V2X.

Optimal data path routed via eNB using Uu links and possible switch between SL and optimal data path for V2V is discussed in 3GPP and literatures [GKA14], but not multi-connectivity with SL and Uu link.

In [YJS+17], the possible use of pre-established SL between a pair of given UEs for Hybrid Automatic Repeat Request (HARQ) diversity and therefore enhancing reliability of radio transmissions is considered. This is based on the eNB scheduled unicast D2D using SL which is out of the working assumption behind V2V which is connectionless transmitter oriented broadcast communications.

4.3.3 Solution

SL is established as the primary connection for E2E communication between vehicle devices. In addition, the secondary connection via BS may also be established for the same E2E communication, i.e., the UL transmission from one communication device is mapped to the DL transmission of another devices so that the E2E communication is routed via BS. Considering that there is a full radio protocol stack from PHY to Packet Data Convergence Protocol (PDCP) over the primary SL connection, the secondary Uu connection may not need to support full radio protocol layers in order to reduce processing overhead and delay over the secondary Uu connection for simplification and better performance.

4.3.4 Impact on architecture

Impacts of this TC are mainly on Radio Access Network (RAN) level including PC5 and Uu interfaces, as described in detail in Section 5.4.3.

4.4 Location aware scheduling

4.4.1 Problem

Wireless networks typically differentiate different types of traffic by assigning appropriate QoS parameters to different services, such as File Transfer Protocol (FTP), or Voice over IP (VoIP) Each QoS class is typically associated to network KPIs in terms of latency, throughput, relative priority, reliability, etc.

In 5G/NR, the concept of QoS is being extended by the use of “flows” which differentiate different traffic QoS parameters within a common radio bearer. As specified in [3GPP17-23501], the 5G/NR QoS model is based on QoS Flows, where QoS Flow is the finest granularity of QoS differentiation in the Protocol Data Unit (PDU) Session. A QoS Flow ID (QFI) is used to identify a QoS Flow in the 5G System. User Plane traffic with the same QFI within a PDU Session receives the same traffic forwarding treatment (e.g. scheduling, admission threshold). Moreover, a scalar 5G QoS Indicator (5QI) is specified that is used as a reference to 5G QoS characteristics. Standardized 5QI values have one-to-one mapping to a standardized combination of 5G QoS characteristics as specified in Table 5.7.4-1 of [3GPP17-23501].

The 5G QoS characteristics associated with 5QI describe the packet forwarding treatment that a QoS Flow receives in terms of the following performance characteristics: resource type (Guaranteed Bit Rate (GBR), delay critical GBR or Non-GBR), priority level, packet delay budget, packet error rate, averaging window, maximum data burst volume. Among these characteristics, Packet Delay Budget (PDB) is of particular interest for many V2X applications. As specified in [3GPP17-23501]3GPP17-23501, the PDB defines an upper bound for the time that a packet may be delayed between the UE and the UPF that terminates the N6 interface. Additionally, as given in Table 5.7.4-1 of [3GPP17-23501], the PDB associated with the current 5QI values varies from 5ms to 300ms.

However, in many V2X applications, the requirements of packet delivery should be interpreted in terms of geographical region or distance instead of absolute latency budget. Among various this type of applications, a typical example is High Definition (HD) map dissemination for V2X services [5GC17-D21]. In a practical way to distribute HD map, an off-board system, e.g., an application server, precisely aggregates and collects the context information, such as vehicles, pedestrians, road structure reference objects and so forth from different sources, and then uses the information to construct an HD map. The HD map is divided into sectors or polygons, which are then disseminated to the vehicles in the geographical areas corresponding to the sectors or polygons. In this way, the actual requirement of delivering HD map packet(s) is as follows: the packet(s) should be successfully delivered to the target vehicle(s) before the vehicle(s) enters the relevant area of the packet(s). Here relevant area is defined as a geographical region in which the vehicle(s) is interested in the corresponding map information.

If we directly apply the current QoS framework with the current 5QI values to this type of packet dissemination, the network or scheduler will schedule the packet based on generic QoS requirements that may be unnecessarily stringent. Note that the current PDB values vary from 5ms to 300ms. In this way, in case of large amount of data that requires reliable and urgent delivery, the network will have little freedom in assigning resources and thus the system performance will be degraded. To address this issue, new mechanisms are required to enable more efficient scheduling.

4.4.2 Prior art

Prior works on scheduling for vehicular environments mainly focused on guaranteeing low-latency and reliable communications, which requires additional efforts compared to legacy environments due to vehicles' mobility [SNK+16]. Among the recent contributions, [XZF+17] takes into consideration the scenario where vehicles communicate among each other relying their traffic through a base station (i.e., no direct communications) and formulates an optimisation problem to be applied at the base station to maximize the amount of transmitted data in the cell taking into consideration location and speed of vehicles. [ZCY+15] defines a scheduling algorithm running at the RSU based on the definition of a weight taking into consideration channel conditions, vehicle speeds and vehicle and/or data priority. A different approach can be found in [DSL+17], which proposes a two-fold scheme to reduce latency and improve reliability by jointly using semi-persistent scheduling and distributed power control. A more detailed list of works on this topic can be found in [SNK+16].

Works listed above as well as the largest portion of works in literature [SNK+16] considers that data to be delivered to vehicles are "local", meaning that data received by the vehicles will be immediately used once upon reception as they contain information (e.g., positions and speeds of other vehicles) relevant to the area where vehicles are moving in that specific time interval. As mentioned in Section 4.4.1, V2X applications deal with geographical data but in some cases (e.g., HD map dissemination) such data might be generated and available before the vehicle approaches the target area. This means that the network should take into account the geographical area of the data when scheduling data delivery. A possible approach is to use Multimedia/Broadcast Multicast Service. Although being suitable to this purpose as for instance analysed in [SSP+16] to improve the resource utilisation and in [PLY+17] for co-existence with direct vehicle-to-vehicle communications, MBMS has a main drawback in terms of high signalling for MBMS group creation, joining, notification, session start, etc., and this might impact the capability of the network in delivering the desired data before the vehicles approach reference area. In addition, MBMS is efficient when the number of UEs receiving the same content is reasonably high, and this condition might not happen in some V2X use cases as some geographical areas for data delivery might be small (e.g., in the order of a few tens of meters) with thus only few involved vehicles. Small geographical areas also bring additional issues in terms of higher number of groups to be managed and additional joining/leaving signalling.

4.4.3 Solution

To address the problem described in Section 4.4.1, we propose a solution that assigns or adjusts QoS provisioning for each individual packet based on position-related information.

The solution itself has a wide scope of applications. The general assumption or considered scenario of the solution is that, a data packet that is readily available in a distribution node needs to be delivered to another node, e.g., a vehicle, by use of a communication network before a certain condition is met. The data packet can consist, for example, of HD map information or of a traffic-related notification. With 'distribution node' we mean a generic node that may locate in an application server, a network node in the core, a radio node, a scheduler, a device, and so on. Moreover, the condition can also be defined in different ways. For example, if we consider the use case of HD map dissemination, the condition consists of delivering a data packet (i.e., information relevant to a map) before a vehicle approaches the relevant area of the packet. In another example of CAM message dissemination, the condition consists of delivering a data packet (i.e., a CAM message) before a vehicle has moved over a certain distance relative to when the data packet was generated.

In the following we will detail the solution by considering HD map dissemination as an exemplified use case. For HD map dissemination, a map packet (e.g., a warning of a detected slippery zone) may only be relevant for certain geographical region. In this case, the solution will assign each map packet with its specific QoS parameters and/or packet priority and/or packet delivery strategy (e.g., choice of which eNB/gNB, choice of bearer or flow), which depend on the relevant area of the packet and the position-related information of the target receiving vehicle.

An exemplified implementation of the proposed solution is illustrated in Figure 4.6. The input of the distribution node includes relevant area of the transmitted packet and location as well as velocity of the target receiving vehicle. With the information, the distribution node can estimate, e.g., the required latency budget of packet delivery as illustrated in Figure 4.7. Furthermore, with the estimated latency budget and some trajectory information of the vehicle, the distribution node or another network node can also select the optimal route (e.g., the optimal cell) for delivering the packet, as shown in Figure 4.8.

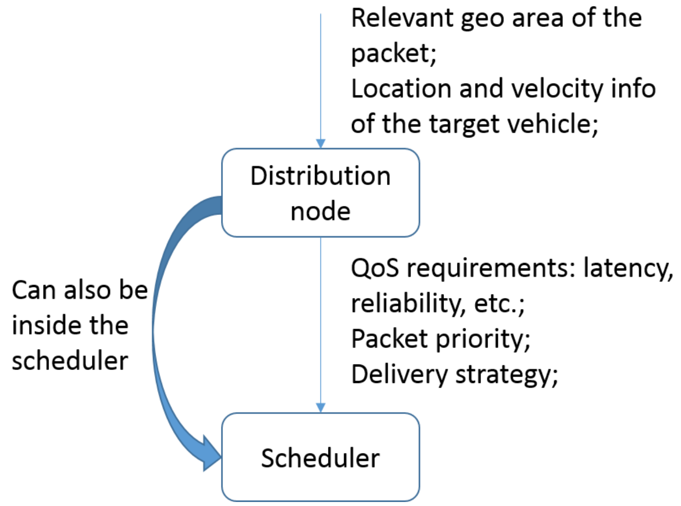


Figure 4.6: An exemplified implementation of location aware scheduling

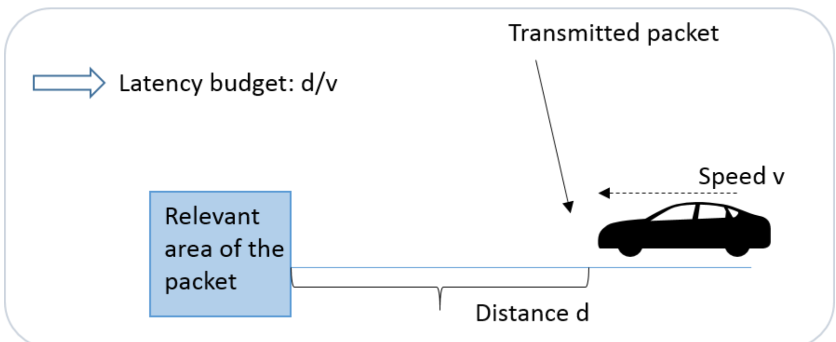


Figure 4.7: Example of latency budget estimation

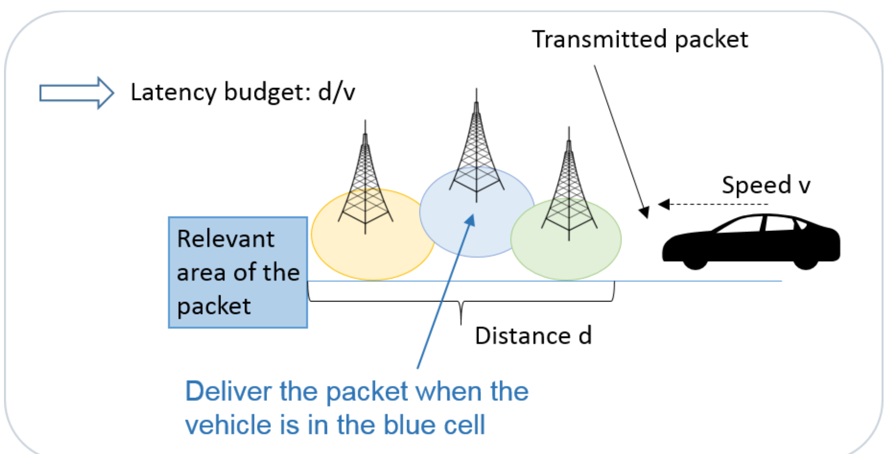


Figure 4.8: Example of selecting the optimal cell

4.4.4 Impact on architecture

The impact of this technical component to the architecture refers to the introduction of a distribution node, which needs to be located before (or co-located) with the packet scheduler of the network. The distribution node might be either a network entity or a network function that can be chained to other network functions (and, specifically, before the network function implementing scheduling decisions). The distribution node might be located in different locations of the network, for instance depending on the type of traffic or use cases. For instance, for traffic referring to large geographical areas or for traffic that could be sent much earlier the vehicle approaches the reference area, the distribution node can be for instance located at the entry point of the core network either as an external server or integrated within a network node (i.e., at the Packet data network Gateway (PGW) or UPF, considering 4G and 5G networks, respectively). For use cases dealing with more localised traffic, i.e., traffic referring to smaller geographical areas or in scenarios where vehicles are close to reach the reference area, the distribution node can be located closer to radio nodes or even as part of the scheduler within the base station or within a device. The flexibility of this component makes it suitable to be included within the architecture without requiring any specific adjustments to work with, e.g., multi-operator or local-breakout or multi-connectivity scenarios.

As a final remark, we highlight that the architecture is also impacted by means of introducing further information (e.g., geographical area of reference, mobility speed of the vehicle) in addition to data to the transmitted. Such information might be included as additional fields in the header of transmitted packets or as additional information provided by the source of the V2X traffic.

Further details on the impact on the baseline architecture are provided in Section 5.4.4.

4.5 Infrastructure as a Service (IaaS) for vehicular domain

4.5.1 Problem

Software system supporting vehicular services should be very complex to deploy, maintain, manage, configure, upgrade, repair due to the numerous of ITS services and number (a priori) of component induced. It will rely on servers distributed at several locations: at the edge on MEC server, at district level (big town, or group of town), at regional level, at country level, and more centralized (at the Internet for instance). This raises two challenges: in one hand well defining the whole system in a coherent way, and in another hand to be able to define a set of infrastructure services enabling to automate deployments of each component of the system.

Definition of the components of the system

WP2 has defined a set of use cases. Use cases could be extended dramatically starting from this base. We could add new ones, such as for instance platooning, crossroad traffic management, or city traffic optimisation. Some basic services have been defined in [ETSI14-3026373], [ETSI14-3026372], [ETSI09-102638], but the global system has not been defined already, as well as the precise cinematic of each use case to validate, improve facilities layer. We can image than a service provided by a piece of software (i.e. a software component) will be used by several use case. Several use-cases will be likely performed by the same software component which will be able to cope with several types of vehicle traffic management.

Some software components will depend on a specific local context such as Lane Merge for instance for which every crossroad will be different from each other, as well as the number, the location of the cameras. Thus parameters of lane merge component will have to be set. In other cases plug and play features could be achieved adding probably some complexity but simplifying management.

Some software component could be associated to a specific use case. For instance one software component deployment process will have to be connected with information system describing topology of roads, crossroad, road sign, RSU, etc., in order to set the parameters in the right way.

Another point to be considered is the dependency (data) between software components characterized by the fact that one component needs data produced by another one: for instance traffic regulation at large scale should rely on traffic measurement on several crossroads. ITS services could be organized in several levels corresponding to different level of geographical area.

This dependency could be present between two instances of the same component corresponding to the same use case. It will be for instance the case for platoon use case. The same platoon will be in sequence taken into account by several instances of platoon software components when these instances are distributed on several MEC servers. And finally the type of data used, produced by the software component will have to be taken into account when designing, implementing, deploying the system in order to well define data service linked to components. The input data could be for instance local at one component instance, shared locally as we mentioned before or global such as profile data (to distinguish the type of vehicle, such as a personal car from an ambulance), as well as identification / authentication data, etc.

Deployment options: impact on the cost and performances

As we mentioned before software components instances could be deployed at different levels of the network at the edge (MEC), regional, centralized on servers within datacentres of different size. The more the component instance will be instantiated high in the network, and the more the component instance will have to control use cases instances (for instance different cross roads for lane merge application). Thus scalability problems could appear leading to duplicate software components instances on several virtual machines mapped on different physical servers reducing in that case the interest of centralisation for resources saving allowed by the sharing of physical resources. Scalability issues could be solved by increasing the capacity of a component (scaling) for the same geographical area, some creating a new component with a new geographical repartition between components: these aspects are present in ETSI NFV at scaling for network services or Virtualised Network Functions (VNFs).

Furthermore, one common idea is the fact that implementing one critical component high in the network will have bad consequences on latency. It is not always the case, for instance if component has to retrieve centralized data (such as profile, or credentials): in that case centralisation could increase latency.

Implementing higher in the network could simplify the application development or corresponding deployment when data dependent is present.

Figure 4.9 shows the application of the previous principles to the platooning use case. One component instance implemented at MEC level is able to managed platoon along the highway corresponding to the distance covered by several antennas. Before leaving the area covered by last antenna, the corresponding component instance has to send information to the next instance implemented on the next MEC server relative area in which in entering the platoon. These pieces of information are related to the platoon (number, type, speed, destination, identification of cars, speed, lane, etc.). This procedure could be called soft handover.

If implemented in a centralized way, there will be no need of coordination if one instance is able to manage a given platoon along the same highway.

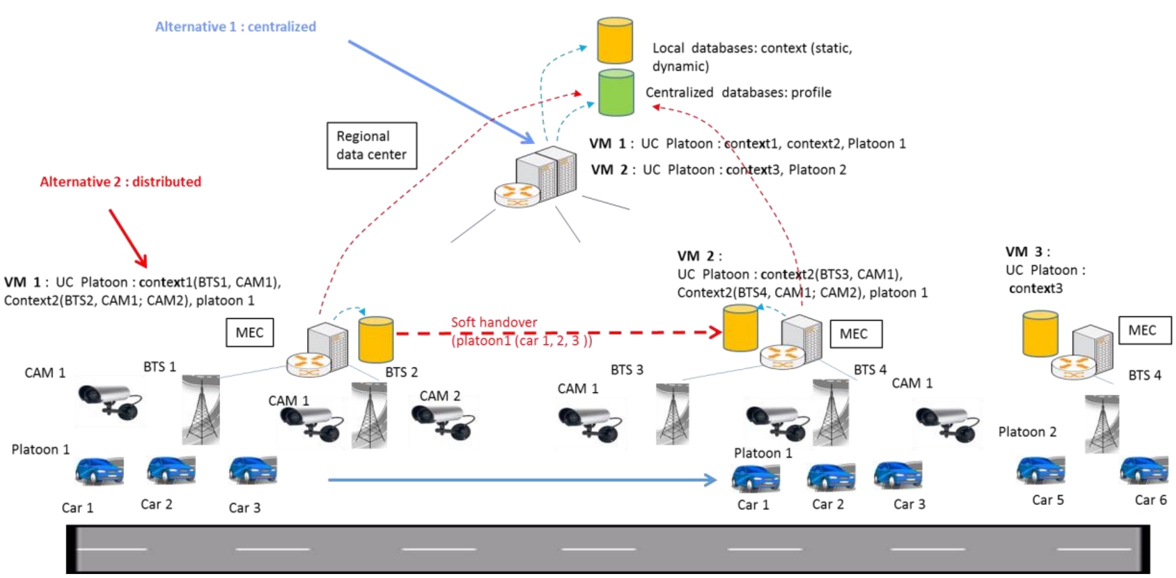


Figure 4.9: Soft handover and platoon use case.

When the architecture is distributed, the implementation of software components on MEC server, soft handover has to be performed when platoon 1 composed of car 1, 2, 3 is changing of zone. Following particular events (entering of a vehicle in a zone) during service, request to a centralized database is needed for authentication, access to profile: this could increase requests delay treatment, in particular when software components are distributed. Even if the implementation is centralized, software handover would be performed when component instance are implemented on different components.

4.5.2 Prior art

Virtualisation completed with Infrastructure as a Service allows a great automation all along the lifecycle management of an application. This automation rely on several orchestration level from End to End service orchestrator to component deployment orchestrator.

Ongoing studies done in telco world provide different open source NFV/SDN frameworks as:

- ETSI Open Source MANO (OSM) is an open source Management and Orchestration (MANO) framework with strong relationship with industry. Focus is on provision of Network Services on top of OpenStack/VMware/AmazonWS infrastructure. Its current release 4 is production ready, but it lacks support for containers and Service Function Chaining (SFC).
- Open Network Automation Platform (ONAP), which is defined as an open source software platform that delivers capabilities for the design, creation, orchestration, monitoring, and life cycle management of Virtual Network Functions. The carrier-scale Software Defined Networks that contain them and Higher-level services that combine the above. ONAP has consequently a wider perimeter than ETSI OSM. It is expected that release 3 (Cassablanca) is fully functional, but current deployment requires a large number of resources.

In other hand, the ITS software must be design in accordance with Cloud Ready principles [ODC-BP]. This adaptation to the features of the virtual resources and to the Cloud tools allows to reached the requirements of the ITS software notably the resiliency.

4.5.3 Solution

At deployment time

Main interest of virtualisation is on demand deployment. This deployment automation is also useful in roll out phase; for example, we can consider to deploy some ITS services independently in several towns or the ITS service can be dynamically deployed on the platform linked to the geographical area where the test car is.

The problem of placement of components in a distributed environment is a complex problem. An orchestrator will be used to deploy the components in order to satisfy on running time SLA constraints such as latency, reliability... Some open source projects such as OpenStack Tricircle [OST-TRI] deal with the problem of placement in a distributed environment.

Furthermore SDN will be an interesting tool: which allows reconfiguration at lower level. ONAP aims at integrating SDN in the framework.

The location of ITS services at the network edge, requires a novel architecture for edge computing resources in order to tackle specific requirements, such as lightweight virtualization, security and scalability. Thus, a novel edge computing architecture is proposed in order to cope with the requirements.

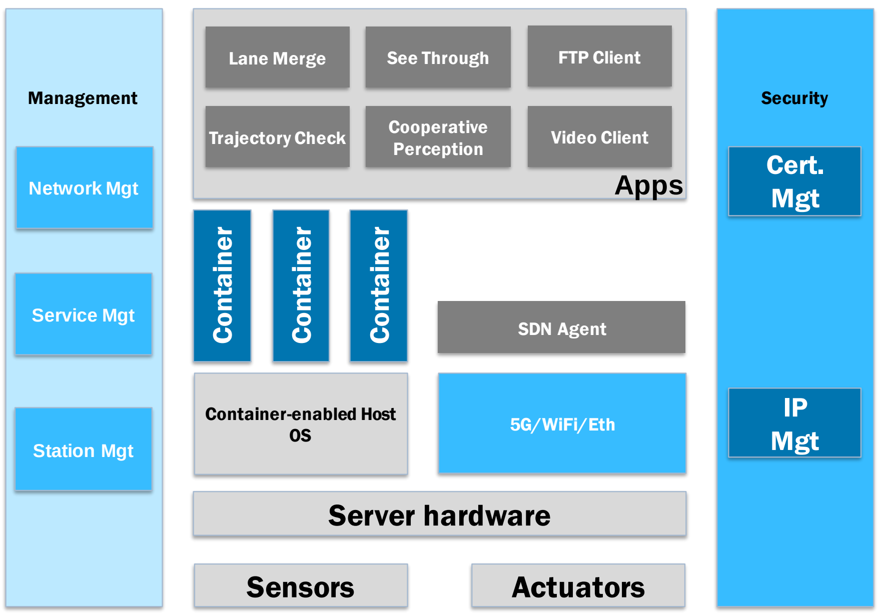


Figure 4.10: Edge computing architecture

The edge node architecture (see Figure 4.10) is the infrastructure that is responsible for provisioning the necessary computation, storage and networking resources at the network edge. Physically, an edge node might consist of a set of (mission critical) servers, hard disks and network interface cards (including WiFi, Ethernet or 5G connectivity). Sensor and actuators might be connected to the edge node. Logically, the edge node incorporates both a node manager in order to handle the offered virtual resources and a security manager, which is the responsible for handling the device authorization, authentication and encryption.

The Node Manager orchestrates the optimal allocation of the required virtual resources onto the node physical resources, which are the pool of existing resources within the node. The virtual allocated resources might include containers or virtual machines (depending on the capabilities of the node), allocated virtual memory, virtual disk space or virtual network interface cards and virtual switches.

Possible edge services abstract the access to the different smart things and devices. Example of these applications are lane merge, see through, FTP client, trajectory check, or video client.

Finally, in order to allow the usage of all these functionalities, the edge node offers a common Northbound Interface (NBI) that can be split into three different sections: i) Network Application Programming Interface (API), ii) Service API, and iii) Station API, which allows to control and manage both the edge node and the connected sensors and actuators.

At running time

As we mentioned scalability problems could appear when traffic is too high: this could lead to increase the treatment time at application level impacting latency. The system will have to be able to monitor the performance of the server or the software / service, being able to scale on the same machine, or to create a new instance dealing with a subset of contexts. The more difficult to achieve will be to be able to change configuration on the flow whiteout impacting ongoing treatments. The Cloud Ready design must allow adapting the service configuration to the deployment modification. (e.g. LBO configuration)

We will have to be able to dynamically change the architecture of the system (components placement) if failures are detected at resource or software level, or if performances are not coherent with SLA. For instance if automatic management detects a bad component behaviour, this component will be killed and corresponding traffic will be affected to safe components: self-healing allows creating automatically a new component in order to replace bad component: these features already exist in ONAP.

Some tools over ONAP have been defined in [SGH17] to dynamically define placement at running time using local information. Furthermore, ITS domain need tools being able to send applicative requests towards servers loaded under a given threshold (see [SGH18]) or for which network latency budget is acceptable. And finally we need being able to deploy software component at a short time period: this timescale has to be defined in order to achieve sufficient flexibility of the system.

Data tools

As we mentioned some data have a global scope, for instance profile data. When a component instance is implemented in a distributed way, retrieving centralized data could deteriorate the performance (i.e. latency). One interesting possibility would be using a service allowing duplicating data in an automatic way.

As we have seen data will have to be set between components: these connections could be offered by a primitive/API integrated in the software code.

4.5.4 Impact on architecture

The goal of NFV was to promote a system being able to cope with real time aspect of Telecom domain in which an important amount of data has to treatment in real time in a virtualisation environment. Some convergence could indeed be found with ITS domain in terms of scalability, resilience, and criticality. Supporting ITS solution on the same tools operator build their virtualized network would be an interesting principle to study.

Defining a IaaS for ITS domain is an important challenge. Indeed Over The Top solutions will not match with specific constraints of ITS domain: latency, resilience. IaaS raises the problem of business model: who will be the Business-to-Business (B2B) customers of IaaS? How to cope with multitenancy, how to integrated IaaS solution inside operator infrastructure?

Further details on the impact on the baseline architecture are provided in section 5.4.5.

4.6 Redundant mode PC5 + Uu

4.6.1 Problem

In order to achieve a very high reliability requested by Use Cases covered in 5GCAR, there are not so many options to guarantee the expected results and whatever the technology used, it could always fail at some point for unexpected reasons. Of course performing radio retransmissions in a more efficient way and design more reliable radio links such as URLLC is the first step but using a redundant system approach on the top of that could provide an alternative path in case of a radio link failure.

4.6.2 Prior art

In addition, it is worth noting that a possible switch mechanism between PC5 and Uu interfaces has been patented [KLV10]. However few studies have been carried out to see the gain in using concurrently both interfaces for a single flow instead of choosing the most appropriate one at a given time. A switch mechanism or any RAT selection method is generally based on specific KPIs and threshold values, like for instance when the radio level of some specific interface is decreasing. However radio conditions and vehicle positions could vary quite quickly leading to the point that a given RAT or interface selection decision may not be the most suitable for all situations, specifically when high reliability is the target.

4.6.3 Solution

Only for specific V2V services requiring very high reliability such as the Cooperative Manoeuvers Use Case type, a redundant scheduler could be implemented (at UE or RAN side) in order to perform the same transmission on both radio interfaces concurrently (PC5 +Uu) for the same V2V service, assuming that the receiver could receive the same message through either PC5 or Uu interface. Such concept is represented in Figure 4.11.

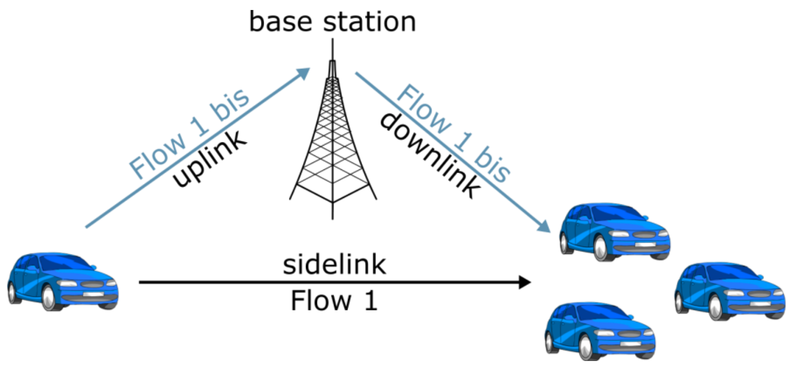


Figure 4.11: Redundant mode for a V2V service

If the same flow is transmitted through both interfaces at the same time on different frequencies, the reliability could be improved without the need of a complex scheduler or dynamic information coming from lower layers. Moreover it is worth noting that V2X concurrent inter-band configuration for Uu and PC5 is already defined since Rel14 [3GPP18-36101] allowing the possibility for the UE to transmit on both interfaces up to a maximum bandwidth of 40 MHz, with 20 MHz on Uu and 20 MHz on PC5. Such concept could also rely on an enhanced Mode 3 eNB scheduler, showing how to improve the reliability for V2V services thanks to Uu interface in complement of direct sidelink communications. However the flip side is that duplicated transmission will create overhead and will increase interferences, specifically with a high data rate service, so a first recommendation is given below:

Recommendation

The redundant scheduler for concurrent transmission on Uu+PC5 for same V2V application should mainly improve reliability for a cost of higher bandwidth. In consequence this feature shouldn't be applied to all V2X services as the global resources usage will not be optimized on both interfaces with a potential resources waste and interferences increase. However we could propose this feature for some specific V2V requiring:

- Higher reliability
- Lower data rate
- For a given limited time

According this statement and Service Requirements defined in WG2, this feature could be proposed and simulated for Cooperative Lane Merging UC.

It is worth noting that with a RSU using PC5 communications close to an eNB, this same solution could apply as well for V2I/N use cases.

Implementation options

There are different implementation's options concerning the location of this redundant scheduler, more specifically related to the layer level and involved entity.

Option A: Redundant scheduler at Application or Facilities level at UE side

The redundant scheduler could be implemented at the Application or Facilities level directly in the V2X UE, considering that all UEs are also compliant with ETSI ITS Stack from ITS European Standards [ETSI10-302665]. The Facilities Layer is described in [ETSI13-1028941] and could easily include a redundant scheduler responsible for the duplication of important messages. This option A following ETSI ITS architecture is represented in Figure 4.12.

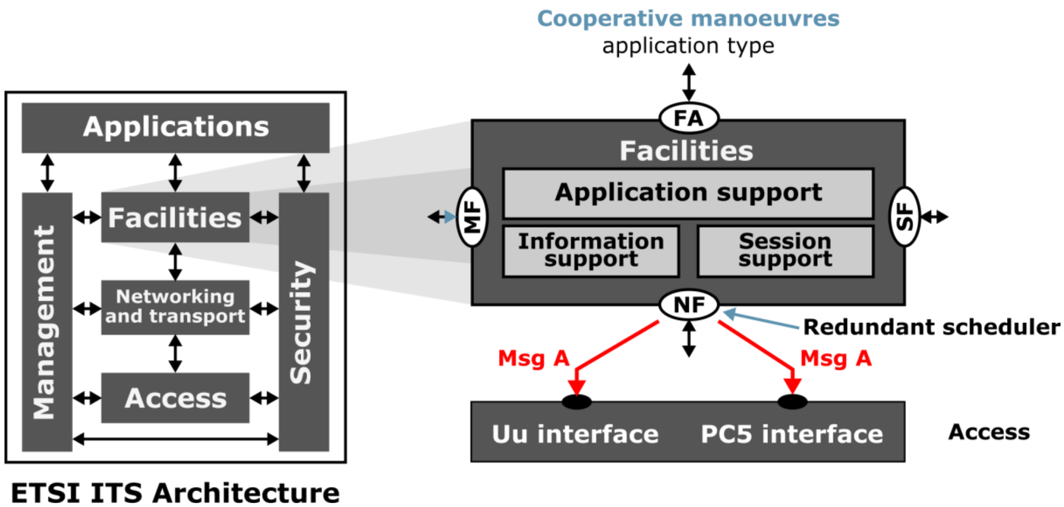


Figure 4.12: Option A - redundant scheduler at UE's Facilities level

A redundant scheduler implementation at Facilities Layer level should be easier and this feature could be restricted to specific applications types such as the Cooperative Manoeuvers Use cases. Moreover the communication interface selection, global messages management, support of repetitive transmission and relevance checking is already under the responsibility of the Facilities Layer in ETSI ITS Standards. However the flip side is that all V2X UE needs to be fully compliant with ETSI ITS specifications and it may not be the scope of 5GCAR project, focusing more on 3GPP standards.

Option B: Redundant scheduler at transport level of UE

Another alternative would be to implement the redundant scheduler at OSI transport level of transmitting UE. This approach is already possible with protocols like Multi Path TCP (MPTCP) [IETF18-RFC6824] but could be restricted only to unicast transmission and both transmitter and receiver UE will need to be compliant to the same transport protocol. An example with MPTCP implementation is represented in Figure 4.13.

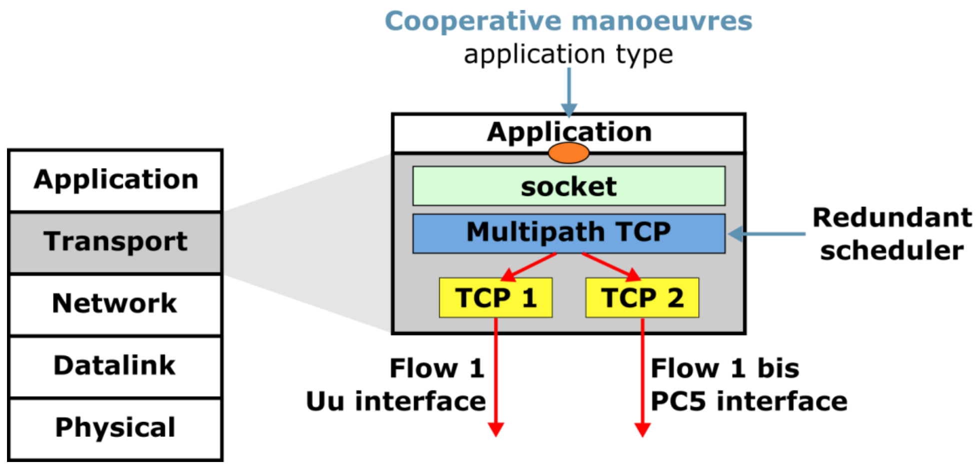


Figure 4.13: Option B - Redundant scheduler at UE Transport level

It is worth to mention that MPTCP is agnostic and can rely on different RATs easily with legacy well known TCP connections. Moreover a redundant scheduler code is already available for testing purposes [ICT17] as well as other scheduler policies. However an end-to-end TCP connection between a transmitter and a receiver UE is adding more delay than connectionless communications such as UDP and it works only in unicast communication with IP connectivity (excluding non IP messages such as CAM/DENM defined in ETSI ITS standards [ETSI14-3026372], [ETSI14-3026373]).

Option C: Redundant scheduler with RRC support

Figuring out all possible implementations for a redundant scheduler, the next option to consider is naturally the Radio Resource Control (RRC) layer with an implementation at RAN level (eNB or gNB). Indeed RRC messages can carry all necessary configurations signalling between UEs and eNB through the Uu interface and a cross-layer scheduling is already possible thanks to the mode 3 specified in 3GPP Rel. 14 [3GPP18-36213], allowing an eNB to schedule sidelink transmissions. The eNB sends a Downlink Control Information (DCI) format 5A (via PDCCH, Physical Downlink Control Channel) to indicate which resources should be used for the communication. Currently it can be used as a scheduling grant for SL (using the Radio Network Temporary Identifier (RNTI) “SL-V-RNTI”) in order to allow transmission through PC5 interface. Alternatively the same UE could receive from the eNB a SPS configuration in order to transmit specific periodic messages through Uu interface. However we miss a link between those two different communication modes if we want to duplicate the same message over the two different interfaces concurrently and for a given time. Some modifications on the 3GPP standards could be foreseen in order to introduce this “Enhanced” Mode 3 or eMode3.

Some guidelines and possible steps are described below:

<u>Step 1: Dedicated RRC reconfiguration signalling:</u>

1. SystemInformationBlockType21 detection from the UE including *sl-V2X-ConfigCommon* element.
2. RRC message sent from the UE to the eNB with *SidelinkUEInformation* element (v2x-CommTxResourceReq including a new element like “redundant” mode)
3. *RRCConnectionReconfiguration* message sent back from the eNB to the UE: *SL-V2XConfigDedicated/Scheduled* and/or *RadioResourceConfigDedicated/sps-Config/* (UL and SL Scheduled or Semi Persistent Scheduling configurations)

<u>Step 2: Physical Downlink Control Channel (PDCCH) scheduling grants with DCI (Downlink Control Information)</u>

4. The eNB gives to the UE a Semi Persistent Scheduling (SPS) grant for SL control and data (addressed to sl-SPS-V-RNTI) or direct grant for SL (addressed to SL-V-RNTI) concurrently with a SPS UL grant for data over Uu interface (addressed to ul-SPS-V-RNTI) for the same message in a given time.

<u>Step 3: Transmission</u>

5. The transmission of the same message is performed with SL (control information and data) and concurrently with UL in SPS.

All steps are represented in Figure 4.14.

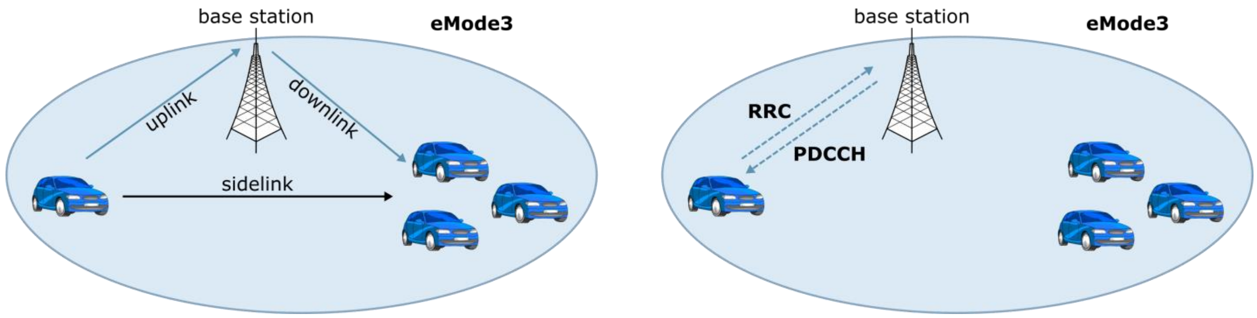


Figure 4.14: Option C - redundant scheduler at RAN level with RRC support

<u>Option D: Redundant scheduler at MAC Level of the UE</u>

Finally a redundant scheduler could be also implemented in a lower layer such as the UE Medium Access Control (MAC) entity following a scheme close to the UL Carrier Aggregation. Such approach is represented in Figure 4.15, with corresponding UL and SL channels.

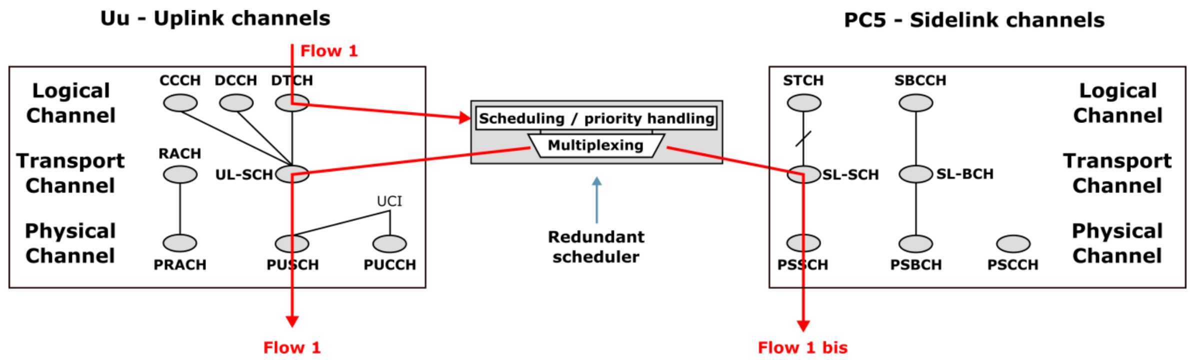


Figure 4.15: Option D - redundant scheduler at MAC Level of the UE

A new UE Mac entity performing Scheduling and Multiplexing could be added like for UL Carrier Aggregation. Scheduling at this level would allow other policies allowing better performances than a simple redundant approach. However this option will be certainly the most complex solution to implement and it is not yet clear how to disable or either allows redundancy for specific flows.

4.6.4 Impact on architecture

According to the implementation option chosen described in the chapter before the impact on architecture will be different. A summary on possible impact towards the global architecture is presented below according to the implementation option chosen:

Option A: Redundant scheduler at UE Application or Facilities level

All V2X UEs need to be compliant with ETSI ITS standards in order to implement a common facilities layer responsible for the transmission of V2X messages. No other impact on architecture is foreseen.

Option B: Redundant scheduler at UE Transport level

All V2X UEs need to be compliant with the same transport protocol. An example could be to implement MPTCP possibly already including a redundant scheduling policy. No other impact on architecture is foreseen.

Option C: Redundant scheduler at RAN level with RRC support

The V2X UEs will need to be compliant with a new RRC signalling as well as a new eMode3 scheduler within a eNB/gNB. No other impact on architecture is foreseen.

Option D: Redundant scheduler at UE MAC Level

The V2X UEs will need to include a new MAC entity with a redundant scheduler. No other impact on architecture is foreseen

A detailed description of the impact on the baseline architecture is provided in Section 5.4.6.

4.7 Evolution of infrastructure-based communication for localised V2X traffic

4.7.1 Problem

In many V2X use cases (e.g., cooperative collision avoidance, emergency trajectory alignment, sensor information sharing, video sharing) the information that is exchanged among the vehicles (V2V) is localised. This means that the communicating vehicles are located at the same geographical region, and there is no need to access a remote server (e.g., V2X Application Server, ITS cloud server), while in the context of the same service multiple transmission modes (unicast, broadcast, multicast) might be required.

For this type of communication both the cellular interface (i.e., when the end-devices communicate via the radio network infrastructure) and the sidelink interface (i.e. when the end-devices are directly connected via the radio interface) could be used taking into account radio conditions and the environment where the eV2X scenario takes place. Through the sidelink interface, the end-to-end latency can be reduced substantially, when devices are located in close proximity, while increasing the spectral efficiency of the transmission. On the other hand, the cellular interface (using uplink and downlink resources) does not suffer from the half-duplex constraint, as happens with the sidelink communication, a larger geographical area could be supported or guaranteed QoS (i.e., high reliability, low latency) could be provided in the case of no Line-of-sight (LOS) among communicating vehicles (e.g., intersections) or poor radio conditions in the sidelink interface.

However, many user plane and control plane procedures of the cellular interface have been designed taking into account the features and the Quality of Service (QoS) requirements of traditional services e.g., voice, voice over IP (VoIP), video and web data services. The low latency at the control plane (i.e. connection establishment, or new bearer establishment) was not a key requirement for traditional services and the existing systems have not been designed with the specific performance requirement. In addition, due to the nature of these services (i.e., a remote server or user participates) the core network entities are always involved for the setup of new bearers and for the data transmission, which increases the required communication and processing delay.

It is evident that existing technical solutions are not suitable to support the challenging performance requirements that localised V2X services have, which includes the need for fast and guaranteed transmission of localised data traffic together with the very fast establishment of radio bearers. Taking into consideration the localised nature of the V2X data traffic, the RAN and the New Radio (NR) Uu interface need to be redesigned in order to satisfy the demanding requirements for V2X reception and transmissions.

4.7.2 Prior art

The state-of-the-art proposals and the existing procedures for establishing and managing radio bearers have been designed taking into account mainly the requirements of traditional services e.g., voice, VoIP, video and web data services. Local IP Access (LIPA) and Selected IP Traffic Offload (SIPTO) are two potential solutions for the efficient handling of the localised data traffic. LIPA is an offloading technique allowing a direct connection between the user equipment (UE) and the local IP network using a femtocell (HeNB) with a co-located or a standalone Local Gateway (L-GW). The L-GW must support limited Packet data network Gateway (PGW) as well as Serving Gateway (SGW) functionalities such as interconnecting with the external IP networks. LIPA is only intended to allow the UEs to access their own Private Local Access Network via a femtocell. SIPTO is similar to the LIPA technology with the main difference that SIPTO can be used in both macrocells and femtocells. LIPTA or SIPTO are mainly used for traffic offloading from the mobile core network to allow a direct access to the public IP network via the fixed network. Both do not include any procedure for fast session establishment and fast transmission and there are no guarantees for QoS features, especially for demanding V2X Services. The MME, Mobility Management Entity, (or DNS) is involved for any control plane procedure [3GPP13-23859], [3GPP11-23829]. In addition, Single-cell Point-to-Multipoint (SC-PTM) is a complementary bearer type of evolved MBMS (eMBMS) [3GPP17-36300]. It is suitable for scenarios where broadcast/multicast service is expected to be delivered to a limited number of cells, to a group of UEs over shared Physical Downlink Shared Channel (PDSCH). The network architecture is the same as in Multicast Broadcast Single Frequency Network (MBSFN) but the SC-PTM provides more efficient allocation of resources comparing to the eMBMS. On the other hand SC-PTM, as happens also with eMBMS has a slow session establishment for a new service that is triggered by the involved devices with the involvement of the application server and focuses only on the downlink transmissions.

In the existing communication systems the control plane latency for establishing a new bearer is significantly larger compared to the requirements of many V2X use cases [3GPP17-22186] since core network entities are involved in the establishment of bearers, which increases the required communication and processing delay. More than 130 ms (control plane latency) are needed for the establishment of required bearers for a group of vehicles that participate in the same V2X service, when the vehicles are not connected to a BS (RRC IDLE state) [3GPP17-36331]. If the vehicles are connected to the same BS (i.e., can directly ask for resources) the control plane latency is higher than 80ms [3GPP17-36331]. Thus, these large control plane latency values are not the appropriate for many V2X use cases. The introduction of new state models for the Radio Access Networks, called “connected inactive”, where both the user equipment and the network maintain context information, enabling the quick and lightweight transition from inactive to active data transmission, could support the fast connection establishment problem [DMS16]. The proposed “connected inactive” could contribute to the reduction of the total control plane latency, but, it is not enough since the delay for adding new bearers, especially for services, where a group of devices participate, remains high (i.e., more than 80ms).

4.7.3 Solution

The formation of local end-to-end radio data paths over Uu interface is proposed to enable the fast and reliable transmission of localised data traffic among the involved devices, satisfying their QoS requirements and the features of the V2X services. The “end-to-end“ term denotes that the (user plane) radio data paths are established among the involved communicating end devices (i.e., vehicles), while the “local” term denotes that the paths are established by (and via) the BSs (i.e. the nodes of the core network do not participate in the user plane transmissions), since the data traffic is localised. Figure 4.16 provides an overview of the involved entities and interfaces.

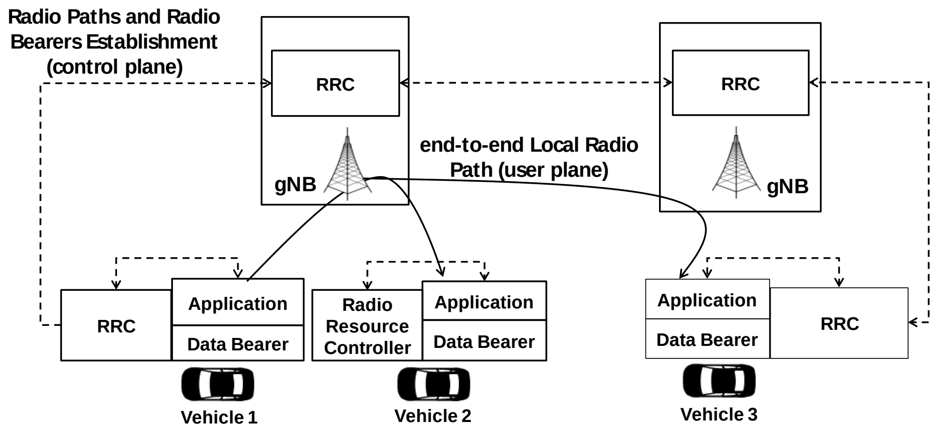


Figure 4.16: local end-to-end paths concept

New methods and signalling are introduced at the RAN for the efficient formation and management of the local end-to-end radio data paths that are presented in Figure 4.17. The Session Request is the initiating message, which is transmitted from the RRC module of a vehicle to the RRC of a BS. The scope of this message is to establish the end-to-end local radio paths among the group of vehicles that participate in the specific application-layer V2X service (e.g. cooperative collision avoidance), considering the required communication modes (unicast, multicast, broadcast) and performance requirements. The RRC entity of the vehicle receives this request from the application layer, based on the triggered events or services. A Session Request message that is transmitted by an individual vehicle triggers directly the establishment of the radio bearers and the configuration of the radio paths for all the vehicles that are involved in the specific V2X service.

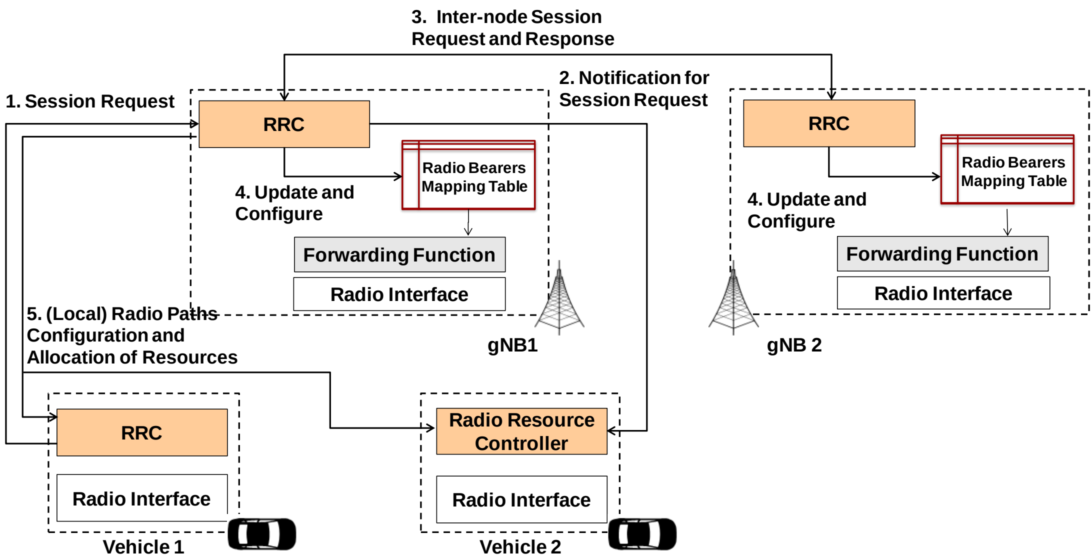


Figure 4.17 Introduced RAN methods and signalling for local e2e radio paths

The Radio Bearers Mapping Table (RBMT) is introduced at the BS and it is updated based on the received Session Requests messages for the formation of local end-to-end radio data paths. These end points of the radio paths could belong to a single cell or at different neighbouring cells (multi-cell radio data path). In the case that multiple cells are involved then an inter-cell coordination is needed. The RBMT of the BS maps and connects the uplink and downlink radio bearers to enable fast and reliable transmission of localised data traffic for different transmission modes (unicast, multicast, broadcast). The UL radio bearer could be linked with DL radio bearers, either unicast (Figure 4.18, case 1) or multicast e.g., SC-PTM radio bearer (Figure 4.18, case 2). The RBMT of the BS does not require IP addresses and there is no need for the MBMS to support the multicast/broadcast traffic, with the direct benefit of lower latency.

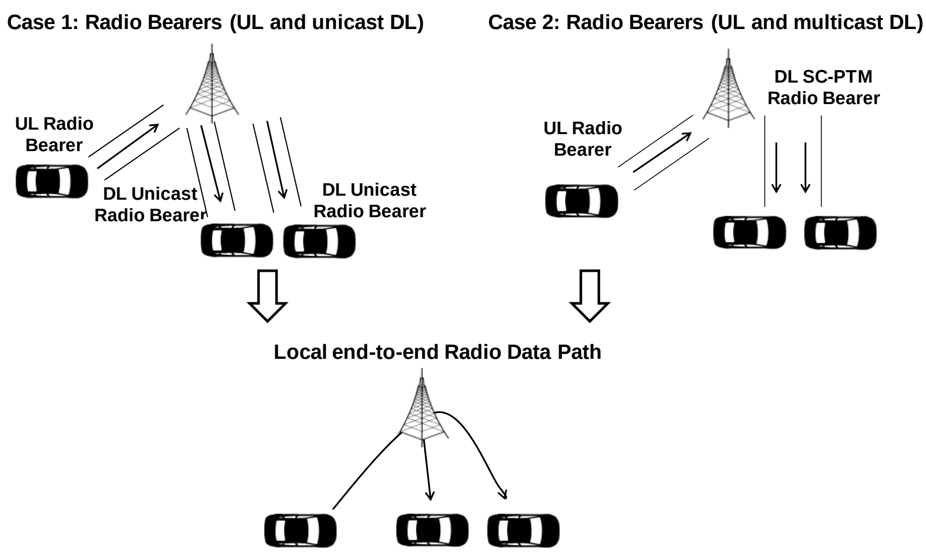


Figure 4.18 Example configuration of the Local end-to-end Radio Paths

4.7.4 Impact on architecture

The Network Functions impacted by this technical component are the UE, the gNB, and the AMF, as described in detail in section 5.4.7.

4.8 Use case-aware multi-RAT, multi-link connectivity

4.8.1 Problem

Bringing together the converge network solutions in the core network and the concept of dual connectivity in the radio access, here we pursue solutions to actively choose one or multiple of the available radio links or RAT for each of the served use cases. Given the diversity of use cases, combination of links and/or RATs that could deliver the requirement of those use case could also be diverse. For example, while cooperative safety is extremely time sensitive, and the success of cooperative manoeuver highly depends on persistent coverage, cooperative navigation is in real need of high bandwidth from the network. Selection of the best possible combination of links/RATs is also a changing choice, given mobility of the vehicles. This TC will provide solutions on where to connect to, for throughout execution of a use case, taking the nature of high mobility of vehicles into account.

4.8.2 Prior art

Dual connectivity as a solution has been defined in earlier standards (i.e. in Release 12 [3GPP12-36842]) as the simultaneous use of radio resources from two eNBs connected via non-ideal backhaul link over the X2 interface. There has been significant work, since then, on more flexible use of multiple connections in delivering single application data or multiple application data to the user. The exploitation of all connection availabilities has been extended to the utilisation of different radio technologies, including WiFi radio (widely considered in V2X communications today), and how different application or user traffic should use which combination of RATs has been subject of discussion in both 3GPP and Broadband Forum (BBF) [BBF16-348] [3GPP14-36842] [3GPP15-23793]. The question of where to connect to has also been addressed as a cell association problem in the literature but the solutions are often static.

4.8.3 Solution

The solution here is based on delivering KPI for a complete manoeuver rather than delivering KPI on a per data packet basis. The solution introduces a new concept, i.e. “completion of action”, where “action” can be defined depending on the use case.

While various technical components can be exploited in timely completion of action, this technical component, in particular, exploits the use multiple links/RATs, i.e. combined use of side link and infrastructure link dynamically.

Illustrative Example: focusing on a cooperative manoeuver use case, we can take the lane merge scenario as an example. The flow of actions in this scenario [5GC17-D21] would trigger the following communications, assuming both merging and in-lane vehicles are connected vehicles:

1. The merging vehicle (i.e. vehicle-1) transmits its intention messages and status information (e.g., position, speed) to the RSU or eNB.
2. The in-lane vehicle (i.e. vehicle-2) transmits its status information (e.g. position, speed) to the RSU or/and eNB.
3. Vehicle-2 also acquires the merging request of vehicle-1 and some other merging related information. In this case,
    a. Vehicle-2 can receive the merging request and merging related information via side-link and directly from vehicle-1, where the merging related information can include the status of vehicle-1, e.g., position, speed, etc.
    b. vehicle-2 can receive the merging request and merging related information from RSU or/and from eNB or an application server, where the merging related information can include the status of vehicle-1, e.g., position, speed, etc.
    c. vehicle-2 can also receive merging related information via combination of the side-link, and infrastructure connection.
    d. vehicle-2 can also receive the merging request and merging related information from RSU or/and from eNB or an application server, where the merging related information can include collision warning messages and manoeuver advise/control messages such as speed acceleration/deceleration profiles.
4. The vehicle-1 receives merging related information.
5. The two vehicles agree on the appropriate actions to enable a safe and efficient lane merging.
    a. This could be communicated via side-link or infrastructure or combined.

At this point, both vehicles apply the agreed action for safe merge. Depending on the (i) position of vehicle-1 when the first message is communicated (step 1), (ii) speed of vehicle-1 and speed of vehicle-2, the full transaction as explained above should be completed within a given time (i.e. k seconds). Therefore, the latency of each message communication should.

This analysis will be different for different use cases, although the concept remains the same.

4.8.4 Impact on architecture

The impact on architecture will mainly be due to the availability of some advance context, i.e. information related to the location, and information of the target manoeuver for the vehicle, as described in detail in Section 5.4.8. One option could be introduction of a new entity such as "advance context", while another option could be to incur no architectural impact but add new procedures and functionalities to the existing entities in eNB and in the UE.

4.9 Multi operator solutions for V2X communications

4.9.1 Problem

Nowadays, communicating mobile devices attach and operate under one operator only. For this reason during power-on they will search and select the operator with whom they have a subscription (Home Public Land Mobile Network - HPLMN) to attach to [3GPP16-23401]. In cases their operator is not present in an area (e.g., usually when the end device is abroad), the end device will perform roaming and attach to a visiting PLMN (VPLMN) [3GPP16-23122].

This process was quite adequate up to now, however it is inefficient for new use cases that will have to be supported in the near future with the arrival of the 5G cellular networks. More specifically, in [3GPP16-36885] it has been identified that V2X communication needs to de facto be addressed for a multi-operator environment, since the end devices cannot be assumed necessarily to operate under a single operator. Cross border cases where the UE crosses the borders of a country are problems in the similar area and require dealing with the multi operator problem; European Automotive Telecom Alliance (EATA) has already highlighted this issue [3GPP16-36885]. The stringent requirements of the V2X communications requiring often information to be exchanged in few ms and reliability of further complicate the problem. Figure 4.19 provides a typical deployment where a highway is covered by multiple operators. The vehicles in the area under consideration are attached to any of them.

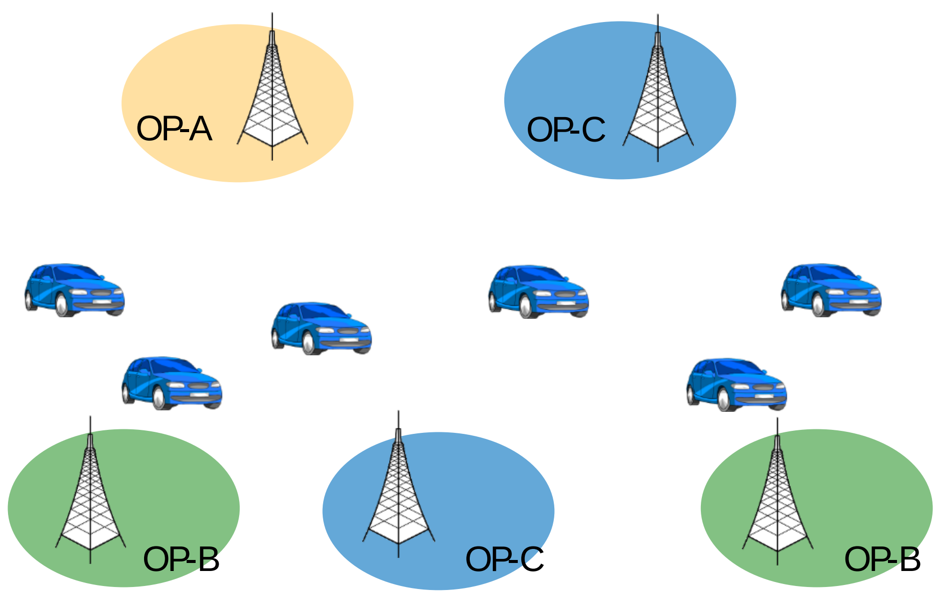


Figure 4.19 Highway scenario covered by multiple operators.

4.9.2 Prior art

Solutions available in the literature can be based on the use of the already available Uu interface and use of existing data paths, where the data for the two devices attached to different operators need to cross all domains (i.e., access, backhaul, and core) of the two operators, incurring a delay that is not acceptable for most of the cases of V2X communication. On the other hand use of PC5 interface requires coordination among the operators on the use of spectrum resources for ensuring proper reception of the information – without necessarily ensuring high reliability if the communication takes place in an uncoordinated manner (Mode 4 PC5). Other solutions can rely on use of schemes that facilitate the use of both interfaces but require either RAN sharing among the operators or to force the devices to roam from one operator to the other. The first approach which is based on RAN sharing requires agreements in all the coverage area so as to ensure proper operation of the network whereas the other based on force roaming introduces unacceptable delays for the transition from the one operator to the other because of the required attachment procedure. When edge computing is considered as well the previous problem is even more complex.

4.9.3 Solution

The aforementioned challenges related to increased delays or limited reliability can be solved if the multi operator problem may be simplified to single operator based on regional split among the operators. Splitting the overall area in regions where only one operator is responsible simplifies the multi-operator environment and enables efficient V2X communication, and offers a good business model among the operators which splits the profit opportunities fairly among them. PC5 communication is efficiently handled by one operator without requiring complex coordination among multiple operators. Also the edge cloud solutions do not require any further enhancements since a single operator offers such service. Figure 4.20 presents how this solution would be applied in the previously described case, where a highway is covered by multiple operators. As depicted, the vehicles in the left hand side of the border are attached to operator A and they operate under him. This facilitates reduced latency and PC5 communication in a controlled manner. Similarly, on the other side of the border all the vehicles operate under operator B. Once the vehicles cross the border they switch from operator A to operator B and vice-versa.

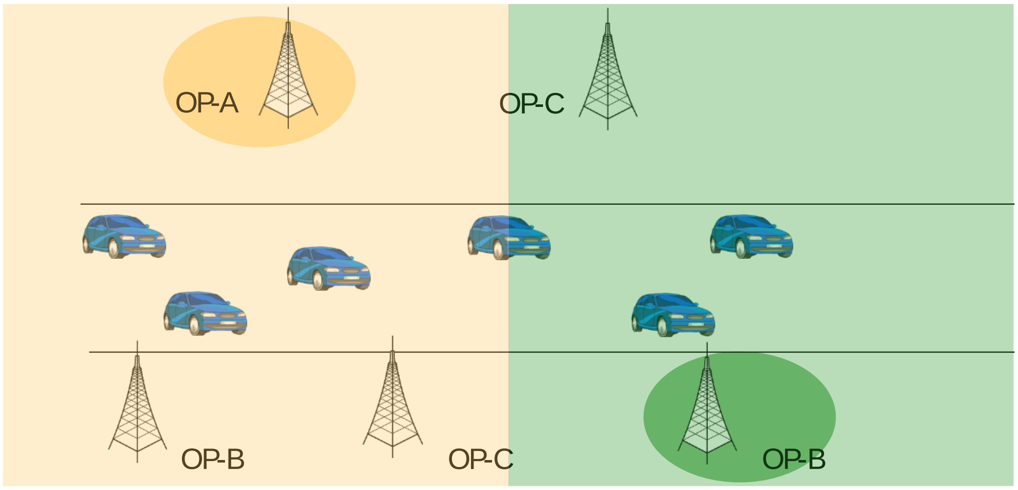
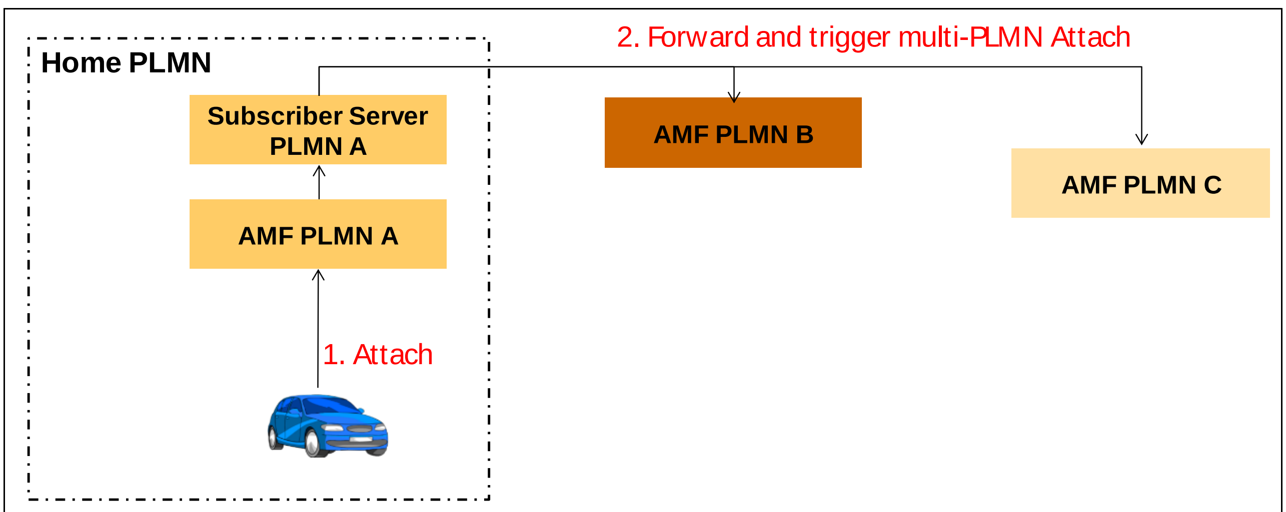
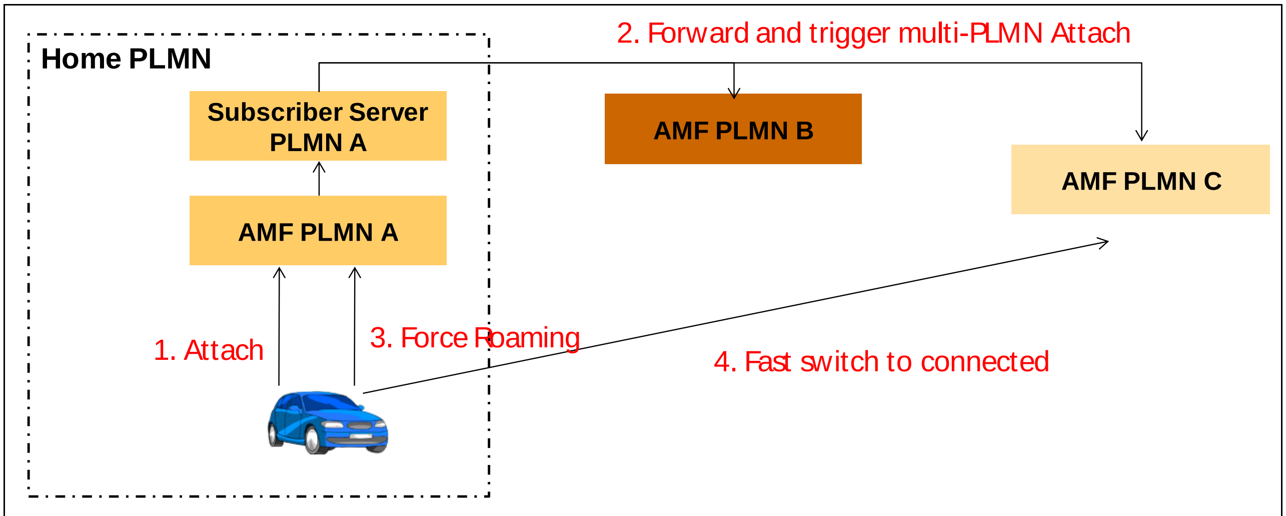


Figure 4.20 Regional split of a highway

However, the regional splitting for achieving acceptable delays requires certain enhancements. In particular the vehicular UE needs to register to all the operators in advance so as to reduce the required attachment delay (which according to [TGV15] is around 330 ms). This pre-registration allows to minimize the time needed to steer a device from one operator to another, because it will be RRC CONNECTED to the operator who is indicated to serve it and IDLE for all the other operators in the area under consideration (Figure 4.21(a)). It should be noted that the registration to all the operators takes place via interactions among the operators. The selection of through which operator a UE should communicate with the other UEs in the vicinity and the network depends on the geographical location of the device or specific instructions from the network. In the latter case the UE attaches to an operator, who instructs the UE to attach to another UE (Figure 4.21 (b)).



(a)



(b)

Figure 4.21 a) Pre-attach to operators and (b) force roaming procedures.

Following this approach the transition from one operator to the other instead of taking around 350 ms it may be achieved in few ms – which represent the delay required for changing from IDLE to CONNECTED RRC state. Assuming solutions based on the newly introduced RRC CONNECTED INACTIVE state [3GPP18-38300] the considered delay may be further reduced, with the cost of pre-transferring the context of the UE to the RAN.

4.9.4 Impact on architecture

The Network Functions impacted by this technical component are the UE, gNB, AMF, and the UDM, as described in detail in section 5.4.9.

4.10 V2X service negotiation

4.10.1 Problem

The effectively delivery of some V2X services might depend on the communication capabilities offered by the network in the time and space dimensions where the service is expected to be delivered. For example, the time needed to deliver a file being part of a HD map might vary depending on the capacity of the network in the specific area where the vehicle is located and also on the number of other devices currently active in that area. This might have an impact on the possibility of successfully delivering the file before the vehicle enters the geographical area of reference the file refers to, with consequent negative impact on the effective delivery of the V2X service. On the other hand, if the vehicle is not close to the reference area, the delivery of a file of the HD map might also delayed without affecting the delivery of the V2X service and with the positive consequence of avoiding unnecessary effort for the network if for instance loaded in the time instant when the file is requested to be delivered. A further example might be related to lane merge use case, the support of a specific inter-vehicle distance requires the fulfilment of specific QoS figures from the network. If the desired QoS cannot be met, then the inter-vehicle distance should be changed in order to optimize the service as well as, for instance, to do not affect safety.

Above listed examples underline the importance for V2X services to increase the awareness about network capabilities in order to optimize the behaviour of the service. Similarly, awareness about V2X service features and requirements would allow the network to optimize the delivery of the use case. In above examples, the meaning of awareness from a network point of view should be enhanced compared to information such as QoS needs. Indeed, additional information such as reference area of data content, involved vehicles with related position, mobility speed and trajectory would be necessary to increase the level of awareness of the network. From a V2X use case point of view, the awareness should comprise e.g. information if the service can be effectively supported or if an adaptation (e.g., higher inter-vehicle distance or delayed transmission) would be needed.

The technical component presented in this section aims at introducing the capability of exchanging information between V2X services and the network in order to improve the delivery of V2X use cases and its adaptation to network conditions.

4.10.2 Prior art

The feature of adapting the behaviour of a service according to network capabilities has been studied from different angles, mainly focusing on rate adaptation. A first approach has focused on adapting transmission rate on an end-to-end basis as for instance considered in [ALH12] and [JOH14], where sender and receiver perform some kind of measurement and the receiver transmit feedback to the sender to adapt the end-to-end data rate. Another example can be found in [TPH+14], which introduces a network-assisted rate adaptation approach where the rate adaptation for video traffic is enriched by information from the network to optimize the data rate. Rate adaptation with network has been also considered by 3GPP in [3GPP17-26247], which defines Progressive Download and Dynamic Adaptive Streaming over HTTP (3GP-DASH) for continuous media delivery. Within 3GP-DASH, Server and Network Assisted DASH (SAND) introduces messages between DASH clients and DASH-Aware Network Element (DANE) or between various network elements for the purpose to improve efficiency of streaming sessions by providing information about real-time operational characteristics of networks, servers, proxies, caches as well as DASH client's performance and status. Above listed approaches, although enabling network-application interaction, only focus on a reduced set of adaptation capabilities (i.e., rate) that might be too limited for V2X use cases. Furthermore, such approaches do not consider the availability of additional information regarding the service at the network as they only focus on the network providing information for rate adaptation.

Regarding the aspect of adapting the file delivery window to avoid congestion in the network, the most relevant example is represented by TCP_LEDBAT [RW13] that has been introduced with the goal to provide a “less than best effort” service delivery for applications that have loose constraints in terms of data delivery deadlines or throughput. Although such approaches represent a valid solution to deliver low-priority content (e.g., software updates) and might potentially enable a reduction in the cost of delivery, they do not take into consideration deadline aspects for file delivery. This aspect is of primary importance for some V2X services (e.g., HD map download) where data content should be delivered before the vehicle enters the geographical area the file refers to.

Recently, 3GPP in [3GPP17-23501] introduced the Application Function (AF), in charge of interacting with the 3GPP core network in order to support services such as influencing traffic routing or interacting with the Policy Framework for policy control. The AF can access the 3GPP network via the Network Exposure Function (NEF). The AF can also interact directly with other network functions if trusted by the operator. The NEF exposes capabilities and events to other network functions and from external applications. In addition, the NEF translates internal-external information when communicating with external functions. In [3GPP17-23502], the interaction between the AF and the NEF is exploited for some AF-network negotiation capabilities, in particular for negotiations for future background data transfer.

4.10.3 Solution

To address the problem described in Section 4.10.1 we propose a V2X service negotiation component which aims at introducing the capability of exchanging additional information between V2X services and the network to improve the mutual level of awareness thus optimizing or improving the service delivery. Examples of negotiation consist of, for example, informing the V2X service that a requested QoS cannot be supported in the time/space windows communicated by the service, in order to allow the service to promptly react (e.g., by sending an alert to the drivers of involved vehicles that the service is not supported). Additional examples, considering 5GCAR use cases, might be the following:

- lane merge, the V2X service increases the inter-vehicle distance and the gap for the merging vehicle if the latency and/or reliability is not sufficient to fulfil the requirements;
- see through, data rate and inter-vehicle distances can be adapted to react to QoS and capacity capabilities;
- pedestrian protection, increase human to vehicle safety distance if for instance position accuracy is not sufficient;
- HD local map acquisition, delivery time window of the map can be optimized according to network capacity all over the trajectory followed by the vehicle;
- Automated parking, speed can be adapted according to supported QoS in order to park the vehicle safely.

The general idea behind the V2X service negotiation component is depicted in Figure 4.22 and this is based on the following: the V2X service provides the network with a service descriptor containing additional information about the service and such information are used by the network to derive additional awareness on if/how the service can be delivered. The types of information exchanged between the V2X service and the network might depend on the use case as each use case has its own specific needs. Similarly, also the feedback provided by the network might depend on the use case. The negotiation could be triggered before the V2X service delivery (e.g., for lane-merge use case, the V2X service triggers the negotiation before the vehicles approach the reference area of the merging) or for an on-going service delivery to adapt in real-time the delivery of the service. This aspect might be service-specific and based on either periodic or trigger-based approaches.

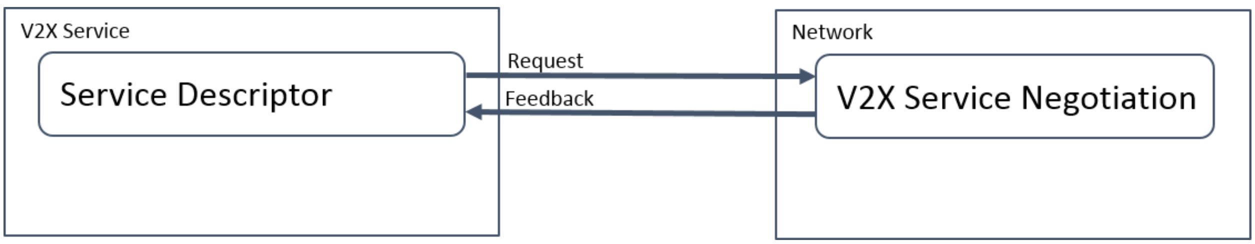


Figure 4.22: V2X service negotiation: general idea and main components

An example of possible benefits enabled by this component is depicted in Figure 4.23, taking into consideration the HD local map acquisition use case. In this case, the V2X service needs to deliver two different files of a different size to two different vehicles before they approach the reference area. The negotiation allows the network to estimate the most suitable time window for the file delivery potentially delaying the delivery (for instance, due to the fact that vehicles will cross base stations that are less loaded compared to the current one) but always with the constraint of delivering the file before vehicles approach the reference area.

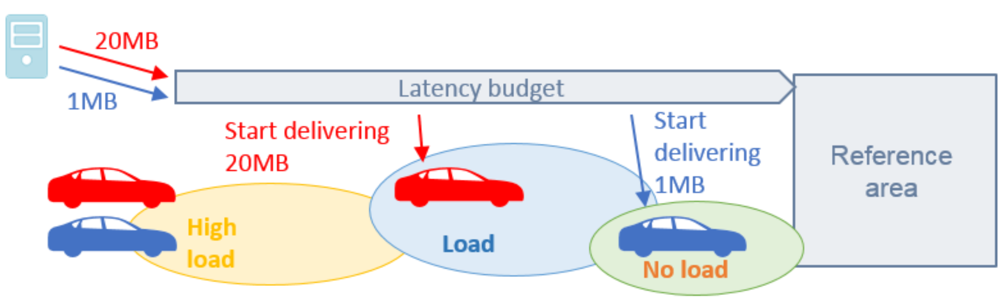


Figure 4.23: Negotiation applied to HD local map acquisition to optimize the delivery window with geographical delivery constraints

A more detailed explanation of the example in Figure 4.23 is given in Figure 4.24, which provides additional information about the types of information exchanged by the V2X service and the network. The V2X service provides the network with two types of information: (i) service descriptor that contains information about the overall size of data file to be transmitted as well as information about the area of reference (i.e., the area before which the file should be received); (ii) vehicle context with information about position, mobility speed and trajectory of intended receivers of the file. The area of reference and the vehicle context might also be combined together to derive the latency budget that will give an indication about the overall time window for delivery. The network will use the information about the vehicle trajectory and area of reference provided by the V2X service in order to estimate the set of candidate base stations to be potentially exploited for the file delivery. Then, the network could use information about the current status (e.g., load) of candidate BSs together with other information such as mobility patterns of involved UEs or statistical information to obtain the most suitable moment to trigger the file delivery taking into consideration the constraint of delivering the file before approaching the reference area. As a consequence, the feedback from the network might consist of e.g. an indication to start or pause/delay the delivery, as well as an indication that the file cannot be delivered if for instance the file is too large and there is not enough capacity available.

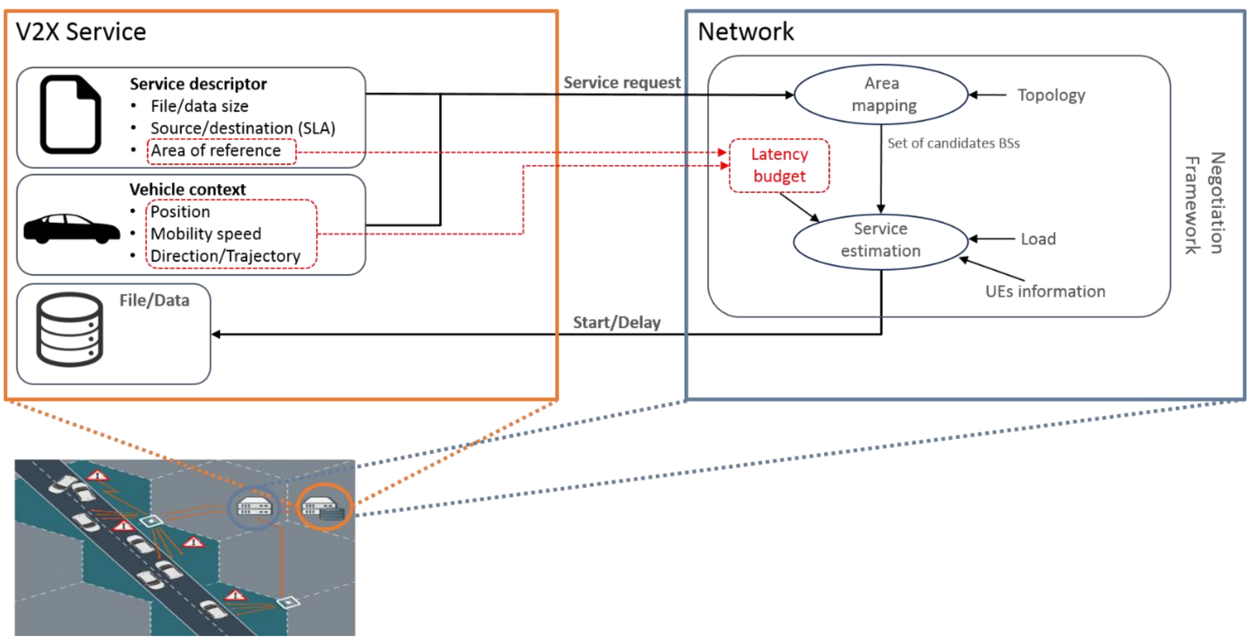


Figure 4.24: Example with information on V2X service negotiation referring to HD local map acquisition use case.

Future work will address other use cases to understand the parameters to be considered for negotiation as well as the effective impact on existing architecture and interfaces.

4.10.4 Impact on architecture

The impact on the architecture of this component is mainly due to:

- The introduction of a novel network functionality to handle the negotiation with the V2X service. Such functionality could be either integrated within existing network functions (e.g., NEF) or added as a novel function. This function could also be placed in different location within the network. For example, it could be placed closer to the network border for V2X services managed by external servers, or it could be placed closer to the edge for services exploiting edge-cloud capabilities or for services based on sidelink communications. Furthermore, such function handling negotiation should be also communicating with other network entities/functions by means of existing or new interfaces to retrieve information to handle the negotiation process.
- The introduction of a novel functionality for the V2X service to handle the negotiation with the network. Such functionality could be either integrated as a functionality of the V2X service or added as a novel functionality of the AF.
- The availability of an interface between the V2X service and the function handling the negotiation. Such interface could be either a dedicated interface or an existing one, and it could also be different for scenarios where the V2X service taking care of performing the negotiation is running on a server/cloud or on a vehicle (for e.g. sidelink-based V2X services). Such interface could be existing interface between the AF and the NEF, in case the negotiation is handled by these two network functions (i.e., the component of the V2X service taking care of negotiation is part of the AF).

It is worth to underline that the realization of the V2X service negotiation (and thus its impact on the architecture) might be use case specific, as different requirements from different use cases might require different architectural solutions.

A more detailed description of the impact on architecture is provided in Section 5.4.10.

4.11 Edge computing in millimetre Wave Cellular V2X networks

4.11.1 Problem

Edge computing, as an essential service technology, enables cloud computing at the edge of 5G networks for significant energy savings and ultra-low latency [MYZ+17]. Cellular and automotive industries and organisations including ETSI and 5GAA consider edge computing as a great value proposition for cellular V2X, to support various V2X services for connected vehicles and autonomous driving. In the meantime, millimetre wave (mmWave) is a key 5G radio-access technology, thanks to its much larger available bandwidths than sub-6 GHz [RSM+13] . In this work, we study edge computing in dense mmWave V2X networks, where there are a large number of vehicle connections and interference from nearby vehicles needs to be coordinated. Since vehicles may move at a high speed, the maximum allowable number of computing tasks shall be controlled such that vehicle’s offloaded tasks can be successfully executed within one RSU range before moving to the next RSU. Such control may also alleviate the load imbalance between RSUs, i.e., vehicles will not offload may tasks to one RSU, to avoid task computation failure resulting from large delay.

4.11.2 Prior art

Existing research has paid attention to the mmWave V2X networks such as beam design [VSB+16] and video delivery [KKC16] [VSB+16] presented an optimisation of beam design in terms of rate and studied the effect of the overlap on the average rate and outage. In [KKC16], the authors proposed feasible and satisfactory system design parameters to mitigate the impact of interference on real-time high-definition video streaming in I2V telematics platforms utilizing 60-GHz radiation and the corresponding IEEE 802.11ad baseband. To date, few works have studied edge computing in such networks, although there are many automotive use cases powered by edge computing [5GAA17-WP].

4.11.3 Solution

We consider a mmWave C-V2X network, where vehicles travel in a highway and RSUs provide edge computing services. While transmitting at the highest power can hypothetically provide the highest rate, and lower the communication latency, it would also increase the interference and inversely impact the reliability as well as latency of the communication. To this end, we formulate an optimisation problem to minimize energy consumption while keeping the network stable and ultra-low latency. The variables that we optimize are the offloaded computing task, transmit powers of the vehicles and its serving RSU at each time slot.

4.11.4 Impact on architecture

The network functions impacted by this technical component are the UE, and the RSU, as described in detail in section 5.4.11.

4.12 Dynamic selection of PC5 and Uu communication modes

4.12.1 Problem

5G communication systems provide both cellular (Uu) and sidelink (PC5) communication interfaces (modes) that have different transmission characteristics and features. For instance, the Cellular Uu has larger coverage area, while PC5 (e.g., mode 3) increases the system capacity through the spatial frequency reuse. As it is mentioned above the V2X use cases have different QoS requirements and the spatiotemporal dynamics of communication networks affect the QoS that the network can provide. For that reason, the dynamic selection of the most suitable communication mode/interface (Cellular (Uu), Sidelink (PC5)) in 5G systems to support the QoS requirements (delay, capacity, reliability) of a specific V2X service is introduced in order to utilize the benefits that each communication interface (mode) can provide at specific point of time and location.

4.12.2 Prior art

3GPP has enhanced its architecture to support the features and requirements of V2X services [3GPP17-23285], [3GPP18-23786], [KMA17]. As mentioned above, there are two modes of operation for V2X communication, namely over the PC5 interface and over the Uu interface. The Uu interface is used for the transmission and reception of V2X messages via the infrastructure. A User Equipment (UE)/vehicle can transmit and receive V2X messages either via Uu unicast downlink or via MBMS for multicast/broadcast reception by establishing the appropriate (radio and core network) bearers, according to the QoS requirements. The establishment of a Uu link for V2X traffic exchange takes place via RRC Connection Establishment messages or via Non-Access Stratum (NAS) messages in the case that a new Radio Bearer should be established. On both cases the existing methods and signaling focus only on the specific radio interface and there is no consideration of activating, establishing or configuring any sidelink (PC5) link.

Support of V2X services via the PC5 interface is provided by V2X sidelink communication, which is a mode of communication whereby UEs can communicate with each other directly over the PC5 interface. This communication mode is supported when the UE is served by network and when the UE is outside of network coverage. Only UEs authorized to be used for V2X services can perform V2X sidelink communication. A UE supporting V2X sidelink communication can operate in two modes for resource allocation:

- scheduled resource allocation (mode 3): The UE needs to be RRC_CONNECTED in order to transmit data. The UE requests transmission resources from the eNodeB (eNB), which schedules dedicated resources for transmission of Sidelink Control Information (SCI) and data.
- UE autonomous resource selection (mode 4): The UE on its own selects resources from resource pools and performs transport format selection to transmit SCI and data. If mapping between zones and transmission resource pools is (pre-)configured, the UE selects a resource pool based on the zone it is located in. The UE performs sensing for (re)selection of sidelink resources. Based on sensing results, the UE (re)selects specific sidelink resources and may reserve periodically recurring (i.e., semi-persistent) sidelink resources.

When a UE is in RRC_CONNECTED and intends to use the PC5 interface for communication, it sends a *SidelinkUEInformation* message to the serving cell in order to request assignment of dedicated sidelink resources [3GPP17-36331]. Thereinafter, the BS sends to the UE a *RRC Connection Reconfiguration* message *(*including *sl-V2X-ConfigDedicated information element)*, to provide to the UE the appropriate configuration (e.g., transmit V2X sidelink data based on sensing using one of the resource pools, Semi-Persistent Scheduling (SPS) sidelink transmission, V2X transmission based on sidelink specific Buffer Status Reports (BSR) from the UE).

From the above analysis, it is evident that the network and the control plane signalling (i.e., RRC Connection Establishment, RRC *SidelinkUEInformation*, NAS Service Request) does not provides the tools for dynamic selection of the appropriate communication interface (Uu or PC5).

4.12.3 Solution

The scope of this solution is to enable the dynamic selection of appropriate communication mode at the BS, taking into account the QoS requirements of the V2X service and the current network conditions (e.g., network load). More specifically, it is proposed to extend the RRC methods for fast mode selection at the BS-side. The application layer may not be the appropriate for making such a decision since it is too slow and does not have enough network layer information. The BS after the initial request by the UE/vehicle will decide whether a cellular (Uu) or sidelink (PC5) communication interface should be used for the corresponding request.

4.12.4 Impact on architecture

The Network Functions impacted by this technical component are the UE and the gNB, as described in detail in Section 5.4.12.

4.13 Security and privacy enablers

4.13.1 Problem

V2X messages exchanged over the air may trigger safety (or not safety) actions on receiving vehicles: actions initiated by the driver when notified of some alerts or triggered automatically by autonomous vehicles. The consequences of such actions can be very serious. It is therefore critical to ensure the authentication of the V2X message sender, its permissions for sending such a message with such a content, the relevance of the message (in space and time; e.g. this is not a replayed message), etc.

The sender of a V2X message must therefore provide some proof that the message is legitimate. One way is to sign the message.

Moreover, privacy is also a key requirement for V2X applications and messages. It must not be possible to identity or to track the sender of V2X messages.

Among all V2X security requirements, one requires the V2X device to use a tamper-proof HSM (Hardware Security Module) to sign messages. The procedure introduces some latency that may not suit all V2X applications.

4.13.2 Prior art

In today's C-ITS projects and pre-deployments, all V2X messages are signed in such a way that they allow every message receiver to perform different consistency and relevance verifications, in addition to the mere cryptographic signature verification itself. Such verifications rely on the signer's digital certificate attached to all messages. Such certificate specifies different applications/messages permissions, validity period and/or validity region. A receiver can therefore assert whether the sender of a message has permissions to send such a message, with such a content, in the region where it has been received, etc.

Digital certificates are themselves signed and assigned to UEs (OBUs, RSUs, VRUs, …) by Public Key Infrastructure (PKI) Authorities – see Annex B.

The European Commission C-ITS Certificate policies (see [EU17-SECFMWK]) allow multiple actors to operate their own PKIs and define the requirements such actors must fulfil to be added in the E.U. Certificate Trust List (CTL) . Such actors may be member states authorities, or private organisations such as car manufacturers or mobile network operators.

4.13.3 Solution #1: end-to-end security (all verifications performed by the receiving UE)

It is useful to distinguish two types of V2X messages:

- General-purpose V2X messages, such as awareness CAM (E.U.) and BSM (U.S.) messages. These messages need to be signed always because they can specify many optional elements for which specific permissions are required, and need be verified.
- V2X application-specific messages. Assuming the application exchanges are designed in such a way that a logical session is first setup, before going on with exchanging further messages, the following approach can take place:

    a. UEs send a signed message to an Application Server (AS). This AS verifies that message (including the UE permissions based on its signing certificate).

    b. The application server then allocates a new symmetric encryption key, referred to as a session key, and sends it – in a signed and encrypted message – to every UE involved in the session.

Key-generation must follow state-of-the-art algorithms to avoid, e.g., pseudo-random generators' lack of randomness pitfalls.

    c. Then each UE encrypts its messages using the session key: if a receiver can decrypt such messages, it proves that the sender knows the session key and therefore has been allowed in joining that application.

Note: the assumption here is that messages exchanged over the session are scoped to a specific application and cannot be used to send additional data that would trigger/impact other applications.

4.13.4 Solution #2: Network-side-based security

The solution proposed in the previous section does not take advantage of 5G capabilities. Another approach, when UEs are connected to the Network, is to rely on the Network itself to perform all consistency and relevance checks against a V2X message. This will alleviate the receiving UE of such work.

Different scenarios are being investigated:

- For traffic going through the Network across the Uu interface, the Network itself can inspect all V2X messages sent by a UE, and performs all consistency and relevance checks before forwarding the V2X messages to the receiving UE(s). So that the receiving UE simply trusts the received message.
- For traffic exchanged over the side-link, the Network can offer a function (Mobile Edge Computing approach) to allow the receiving UE requesting the verification of a received message. Some optimisation may be achievable on the Network side for V2X broadcasted messages whose verification will be requested by multiple receiving UEs.

4.13.5 Impact on architecture

The Network Functions impacted by this technical component are the UE, V2X AF, and the KMF, as described in detail in Section 5.4.13.

4.14 5G core network evolution for edge computing-based mobility

4.14.1 Problem

In conventional mobile networks, mobility of a mobile terminal is enabled by horizontal and vertical handover procedures when it changes serving access point (or base station). However, such handover mechanisms are insufficient to support reliable and punctual computation offloading in the environment of edge computing, where the access point additionally processes jobs uploaded from users. When a handover occurs, those pending jobs (e.g., jobs unprocessed, being processed, or processed but not fed back) must be recovered after the handover. Meanwhile, the new access point should allocate new computing resources to resume job processing. A time-predicted handover mechanism for vehicles is needed by leveraging available road information, associated with dedicated server embedded in the Edge of the network, to satisfy the demand for high mobility and reliability.

4.14.2 Prior art

3GPP Architecture group (SA2) is specifying 5G Architecture [3GPP17-23501] and several architecture options are under discussions. The main question that is proposed to handle in this section is how to adapt 3GPP 5G Architecture to 5GCAR requirements.

In 5G Release 15 Architecture, 3GPP proposes two different paths for the same PDU session accessing same Data Network (DN) [3GPP17-23501, § 5.6.4.1].

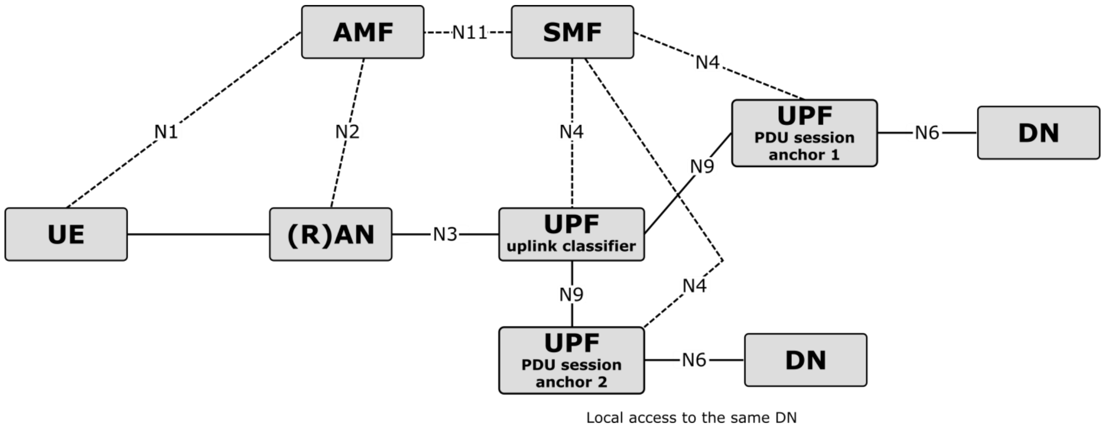


Figure 4.25: Two paths for same PDU sessions for same DN

Mobile-Edge computing architecture for future cellular vehicular network has also taking into account providing new aspect for optimizing core network inside slicing concept.

The basic idea for 5GCAR is to use the above 3GPP proposed architecture to make an architecture proposal for 5GCAR. We may consider two different slice types, eMBB and URLLC, which are already standardized by 3GPP (Table 5.15.2.2-1 [3GPP17-23501]).

4.14.3 Solution

In the proposed architecture for 5GCAR, we consider the slice 1 (eMBB) for accessing internet or vehicle maintenance that may be centralized with no specific constraint (anchor 1 in the figure).

Another slice (URLLC) for another Data Network in which dedicated servers need to be deployed to the Edge of the network (as close as possible of the AN) for automotive and driving help needing an ultra-low latency. This new Edge computing will allow to have functions close to the access in order to enhance the latency for vehicle use cases.

It is proposed then to adapt the 3GPP Rel.15 architecture by providing Edge computation (platform hosting automotive applications) associated with the UPF Uplink classifier as shown in Figure 4.26.

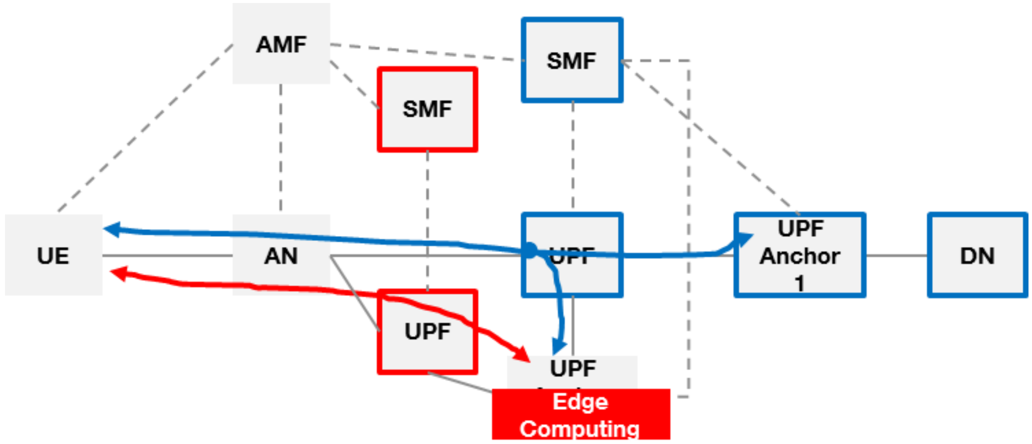


Figure 4.26: Two paths for different PDU sessions for different DNs

Edge Computing or MEC Server is added to 3GPP Architecture in order to response to 5GCAR requirements. One PDU session is associated with each slice.

Implementation proposals for UPF

Two different implementation for the UPF with classifier at the Edge:

1. Branching point IPv6 Multi Homing: In this case the UE is aware of two different IP address for two different destinations (two different DNs)
2. Uplink classifier: In this case UE is not aware that there are two different routes for traffic and the path will be decided by the UPF

Moreover, "UPF/SMF edge" are deployed for URLLC purpose, and, those "UPF edge" needs to have the function of Uplink classifier, controlled by associated "SMF edge". In this case UE is not aware that there are two different routes for traffic and the path will be decided by the UPF. The rules corresponding to the UPF Uplink Classifier are generated by the SMF.

Collocated to the "UPF Edge", dedicated Edge computing function is embedded providing Software as a Service (SaaS) to V2X UE.

The SaaS model is suitable for time-sensitive V2X applications conducted within a given geographical region, and the job processing service for that application is managed by one provider (for instance, the vehicle vendor).

In order to conduct smooth handovers between those "UPF Edge" embedding V2X application, a pre-migration is applied so that the migration is triggered prior to connection redirection, as seen in the picture above. The migration refers to the migration of user data and service logic from source edge to target edge.

On the UE side, the key module is the handover controller, which is responsible for:

1. Checking the channel states (e.g., RSS – Received Signal Strength) from different RANs and learning their changing tendencies

2. Checking the status of offloaded jobs (e.g., which job is “in flight”)

3. Send those pieces of information waiting for the Control Plane to make a decision on job migration and connection redirection

4. Interacting with UPF edge V2X application to advise job pre-migration prior to connection redirection and job recovery after connection redirection (not seen in the above data flows’ picture).

Simplified Procedure

Each UE (vehicle) periodically transmits its status information (NAS: including position, heading, speed, acceleration, etc.) to the gNB while traveling on the road, and then the gNB sends such information to the connected AMF.

The gNBs are supposed to be arranged linearly along the roads and the AMF has a good knowledge of their order, that is, the AMF is easy to determine the target cell for a specific vehicle which is about to perform HandOver (HO). For improving the HO latency, the delimitation of the “handover area” provides a more accurate area where vehicles are most likely to perform handover successfully. This area can be delimited through measuring the radio signal intensity of the roadside gNBs by leveraging detection UEs. Such “handover area” could be an excellent prediction of the handover location pertaining to all vehicles based on the prior knowledge.

The AMF determines the handover target cell based on the real-time information of UE (vehicle) status and the pre-stored cell information. Then it predicts the time when the vehicle arrives at the delimited handover area, whilst assigning a specific “handover slot” for the vehicle in consideration of the resource usage in the target cell. Such “handover slot” is contained in the “handover command” signalling which is sent to both SMF2 and the target SMF3 (see picture above). The handover vehicle can be also informed of such “handover slot” information and other configuration information by the AMF (through gNBs).

The serving/target gNB and the handover vehicle reach agreement under the control of the AMF, so as to work in a coordinated manner. At the “handover slot”, the vehicle disconnects from the S-gNB and smoothly accesses to the T-gNB, realizing seamless and fast handover.

More study will be performed to evaluation how much performance gain from this simplified procedure.

Mobility management in edge computing environment for V2X

Different mobility scenarios are possible with Edge Computing environment

V2X UE mobility without switching MEC server:
In this scenario when the V2X UE is moving, it changes Access Network but the anchoring point in the MEC Server remains the same

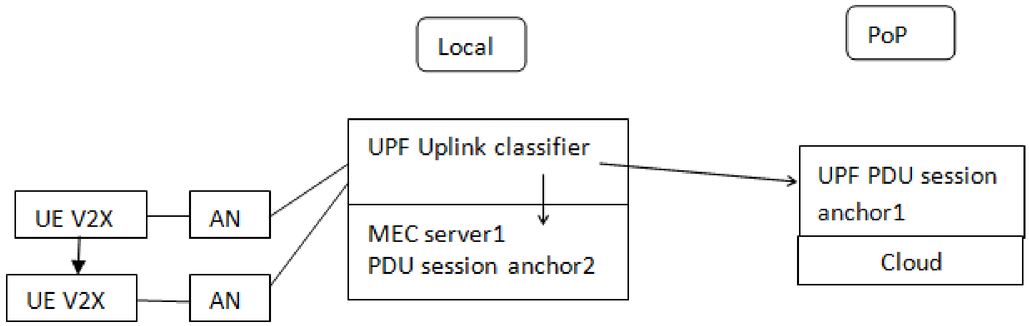


Figure 4.27: AN HO Management without switching MEC server (MEC as an HO controller, see Mobile Edge Computing-based Architecture for V2X)

The redesigned handover process takes specific characteristics and requirements of the vehicular networks into account. It should be pointed that such tailored network function of inter-cell handover can be created as a MEC service by utilizing the virtualisation infrastructure.

The MEC-based vehicular handover procedure is outlined below:

1. Premise preparation. The MEC server provides access to radio and network information, which contains not only a series of static information, like Cell-ID and the location of the connected BSs, but also the real-time resource allocation information of each connected BS. The BSs are supposed to be arranged linearly along the road and the connected MEC server has a good knowledge of their order, that is, the MEC server is easy to determine the target cell for a specific vehicle which is about to perform handover.

2. Delimitate the “handover area”. The “handover area”, a more accurate area where vehicles are most likely to perform handover successfully, can be delimited through measuring the wireless signal intensity of the roadside BSs by leveraging detection devices. Apparently, such “handover area” is an excellent prediction of the handover location pertaining to all vehicles based on the prior knowledge.

3. Report status information. Each vehicle periodically transmits its status information (including position, heading, speed, acceleration, etc.) to the serving BS while traveling on the road, and then the BS sends such information to the connected MEC server.

4. Assign the “handover slot”. The MEC server determines the handover target cell based on the real-time information of vehicle status and the pre-stored cell information. Then it predicts the time when the vehicle arrives at the delimited handover area, whilst assigning a specific “handover slot” for the vehicle in consideration of the resource usage in the target cell. Such “handover slot” is contained in the “handover command” signalling which is sent to both serving BS and the target BS. The handover vehicle can

be also informed of such "handover slot" information and other configuration information by the serving BS.

5. Handover execution. The serving/target BS and the handover vehicle reach agreement under the control of the MEC server, so as to work in a coordinated manner. At the "handover slot", the vehicle disconnects from the serving BS and smoothly accesses to the target BS, realizing seamless and fast handover.

V2X UE Mobility with MEC server switching:
In this scenario, the proposed Fog computing consist on transferring UE ongoing jobs on the current MEC Server to target MEC Server before V2X UE moves to the target AN (Make before Break). This mobility is managed by the MEC Server. In this context the MEC Server becomes:

- An hosting Entity for critical functions for the client
- But also a control and command point for the mobility

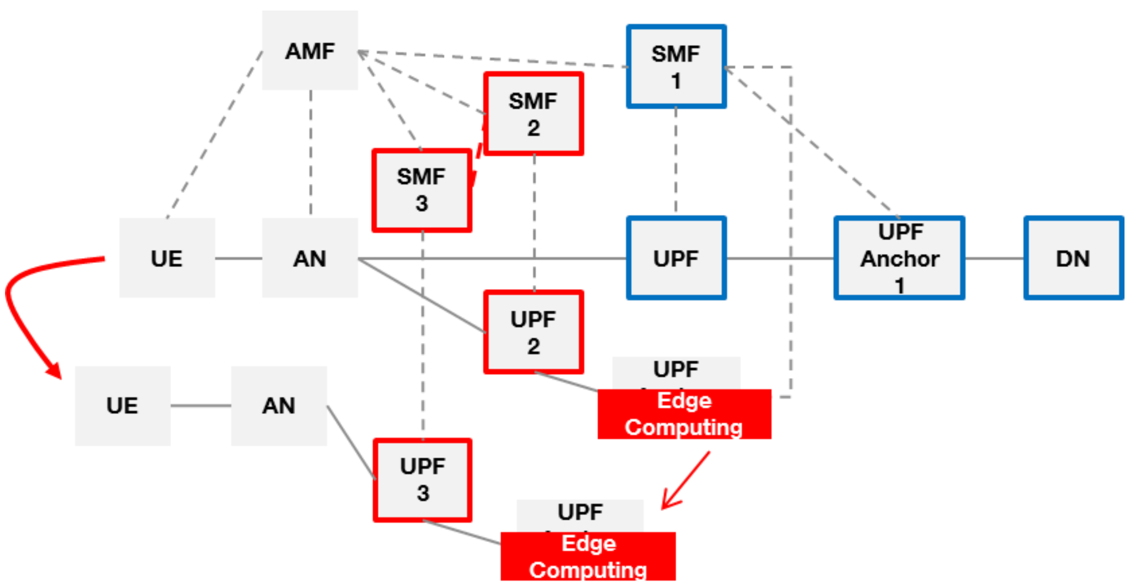


Figure 4.28: V2X UE Mobility with MEC server switching

4.14.4 Impact on architecture

The following upgrades are needed to the current 3GPP architecture, as described in more detail in Section 5.4.14:

- New interface between two SMFs for allowing communication between these entities
- New interface between two UPFs for allowing communication between these entities

5 E2E target architecture

In this section, the 5GCAR preliminary proposition for the E2E target architecture is presented. We first present the standardised 3GPP architecture for 5G in Section 5.1, describe the concept of network slicing and its enablers in Section 5.2, and study the specific challenges and solutions for RAN slicing in Section 5.2.2. Section 5.3 is then dedicated to the specific slicing approach we will adopt in 5GCAR.

Section 5.4 is finally dedicated to the detail study of the impact of the 5GCAR technical components on the 5G architecture. The new features by each 5G network functions are detailed for each of the technical component, and call flow diagrams are provided to visually illustrate the novel procedures introduced.

5.1 5G Service-based architecture

The 5G architecture breaks ground with respect to the previous generations starting from its very definition: the design of the control plane in 5G is centred around services rather than around interfaces, and Network Functions (NFs) are defined, instead of network entities. The 5G architecture is defined in [3GPP-TS23501], and illustrated in Figure 5.1 and Figure 5.2. The definition of each Network Function is provided in Table 5.1 and in Section 5.1.1.

In the 5G architecture, all of the Control Plane (CP) functions are interconnected to each other on a service-basis: each function is provided with an interface through which it can offer services to other functions, or subscribe to services offered by other functions. These services are web-based, utilizing Representational State Transfer (REST) APIs.

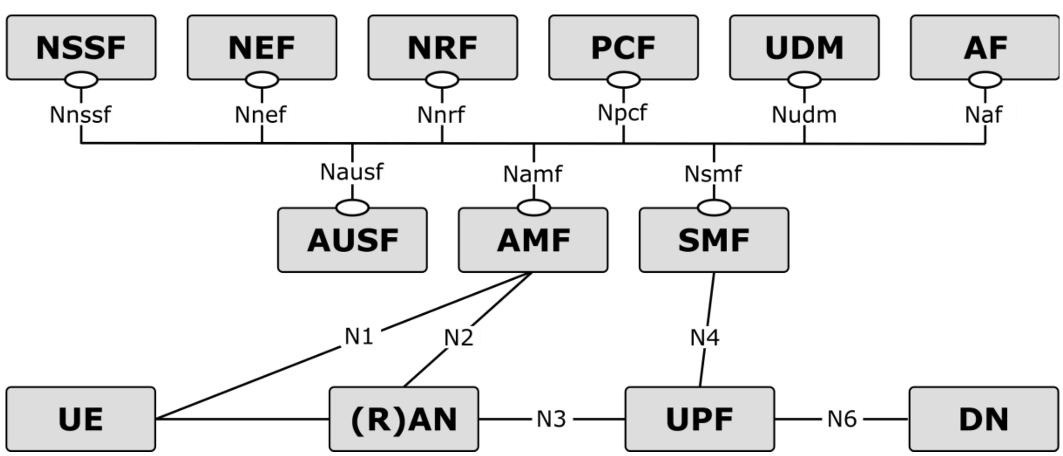


Figure 5.1: 5G System architecture [3GPP-TS23501]

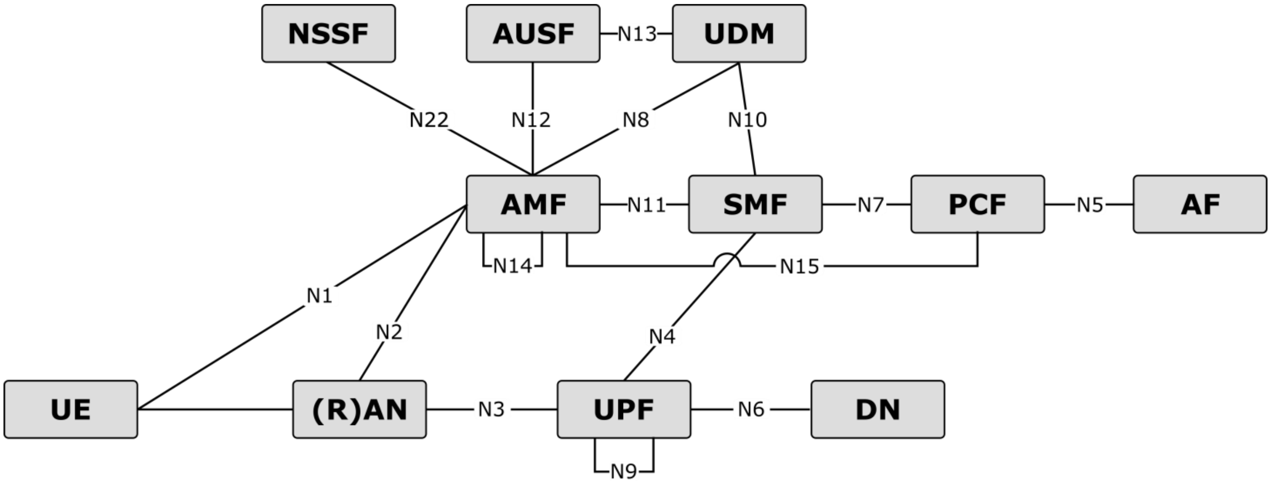


Figure 5.2: Reference point representation of the System architecture [3GPP-TS23501]

Table 5.1: 5G Network Functions

NF acronym	NF full name	Clause in [3GPP17-23501]	
AF	Application Function	6.2.10	Control Plane
AMF	core Access and Mobility management Function	6.2.1	
AUSF	Authentication Server Function	6.2.8	
NEF	Network Exposure Function	6.2.5	
NRF	Network functions Repository Function	6.2.6	
NSSF	Network Slice Selection Function	6.2.14	
PCF	Policy Control Function	6.2.4	
SMF	Session Management Function	6.2.2	
UDM	Unified Data Management	6.2.7	
DN	Data Network	-	User Plane (UP)
UPF	User Plane Function	6.2.3	
RAN	Radio Access Network	-	UP+CP
UE	User Equipment	-	

Interfaces and dedicated protocols are still in place for the interconnection between the control plane and the user plane.

The flexible definition of network functions and service-based architecture are of paramount importance to 5G, as they facilitate its evolution, in order to cover multitude of new services that arose recently and that will be defined in the coming years.

5.1.1 Description of 3GPP Release 15 network functions

In this section we provide a brief description of the network functions included in the base architecture scheme represented in Figure 5.1 and Figure 5.2. For a comprehensive list of the functionalities supported by each NF, we invite the interested reader to refer to the clauses of [3GPP17-23501] indicated in Table 5.1.

AF – Application Function

The Application Function is a control plane entity that influences traffic routing of PDU Sessions based on application-specific requirements. This is done by issuing AF Requests addressed to the PCF (see [3GPP17-23501, § 5.6.7]). According to the deployment configuration, AFs can be trusted or untrusted. In the former case, they may be located within the operator's network and directly connected to the other control plane functions via the Naf Service Based Interface (SBI). The latter deployment option is for AFs deployed in a public network: in this case, the AF cannot directly interact with the 5GC; instead, AF Requests need to be issued via the NEF.

AMF – core Access and Mobility management Function

The AMF is one of the principal control plane NFs, as it is the termination point of the Non Access Stratum, and is in charge of the UE registration management, connection management, reachability, and mobility management [3GPP-TS23.502, § 4.3].

Furthermore, the AMF is in charge of the user authentication and access authentication, and provides transport for session management messages coming from the SMF directed towards the RAN (via N11 and N2 interfaces, see).

The serving AMF for one UE is chosen by the (R)AN node to which the UE is attached following the procedure in [3GPP17-23501, § 6.3.5]; when a UE is subscribed to multiple network slices, the procedure is the one in [3GPP17-23501, § 5.15.5.2]. One AMF serves one UE at any given time, regardless of the number of slices the UE is subscribed to: in this case, the AMF manages the operations for all of the slices.

When applicable, the AMF also supports the aforementioned functions for non-3GPP connectivity.

The AMF covers the mobility management functions implemented by the Mobility Management Entity (MME) in the LTE Evolved Packet Core (EPC).

AUSF – Authentication Server Function

The AUSF is responsible for the authentication and security functions as determined by the 3GPP Services & System Aspects Working Group 3 (SA WG3).

NEF – Network Exposure Function

The NEF is responsible for exposing network functionalities and capabilities to untrusted external entities, such as AFs and edge computing nodes. In particular, the NEF as a gateway for external request to the 5G core network control plane, masking sensitive internal user information.

NRF – Network functions Repository Function

The NRF maintains the profiles and addresses of the NFs which are available at any time in the network. It supports NF service discovery: the NRF provides information about the NF instances available to support a given task, whenever a network procedure requires it.

NSSF – Network Slice Selection Function

The NSSF is a new network function, introduced in the 5G architecture specifically to support network slicing. For each UE, it determines the set of network slices the UE is allowed to access, the set of slices that serve the UE at any given time, and selects the AMF Set, or a list of candidate AMFs suitable to serve the UE, possibly by querying the NRF.

NWDAF – Network Data Analytics Function

The NWDAF, which was introduced in 5G in [3GPP18-29520], provides other network functions with information concerning the identifier of target network slices, and load level information concerning them. Despite it not being represented in the 3GPP base architecture schemes in Figure 5.1 and Figure 5.2, it has a role of paramount importance in 5GCAR, to support the technical components which require analytics data. From an architectural standpoint, it is connected to the PCF via the N23 interface, as illustrated in [3GPP18-29520, § 4].

PCF – Policy Control Function

The PCF is the network function in charge of the unified policy framework that governs the network behaviour. It is responsible to provide the other control plane functions with the Policy Rules to be enforced, and implements a front end interface for other functions to access subscription information relevant for policy decisions (stored in the UDR – Unified Data Repository).

The PCF covers the functions implemented by the Policy and Charging Rules Function (PCRF) in the LTE EPC.

SMF – Session Management Function

The SMF handles all of the aspects related to PDU Sessions, including establishment, modification, and release. It is responsible of maintaining tunnels between the (R)AN and UPF nodes, of the IP addresses allocation and management (for IPv4 and IPv6 PDU Sessions), and of the selection and control of the UPFs involved. It is in charge of determining the Session and Service Continuity (SSC) mode of a session (see [3GPP17-23501, § 5.6.9]).

The SMF also represents the termination interface towards the PCF, from which it receives the Policy rules, which it enforces by acting on the appropriate UPFs. It also represents the termination point of the Non Access Stratum session management messages from the UE, which are transported by the AMF.

UDM – Unified Data Management

The UDM is the function in charge of managing information related to the UE and to the subscriber profile, such as the generation of the 3GPP Authentication and Key Agreement credentials, user identification handling, and storing data related to registration management of network functions serving the UE. The UDM is in charge of supporting the session and service continuity, by keeping the assignment of SMF and DNN for an ongoing session. The UDM may use subscription and authentication data stored in the Unified Data Repository (UDR, see [3GPP17-23501, § 6.2.11]), a NF not included in the base architecture illustration in Figure 5.1 or Figure 5.2, which is responsible of storing and retrieving data related to subscription, policy, and structured data for exposure by the NEF.

DN – Data Network

The data network is a logical function block that may correspond to operator services, Internet access, or third party services. It is identified by a Data Network Name (DNN), which is equivalent to the Access Point Name (APN). The DNN may be used to select the appropriate SMF and UPFs, to select the N6 interface (connecting the UPF serving as the session anchor to the DN itself), and to determine policies to be applied for a PDU Session.

UPF – User Plane Function

UPFs are responsible for the User Plane functionalities, which include packet routing and forwarding, UP part of policy rules enforcement, and in general QoS enforcement for the user plane. Furthermore, the UPF serves as an anchor point for intra/inter-RAT mobility, and as external point of interconnect to a Data Network for PDU Sessions. The number of UPFs in the data path is not constrained by 3GPP specification [3GPP17-23501, §8.3.1]: beside the PDU Session anchor, there can be one, multiple, or no further UPF.

5.2 5G network slicing

Network slicing is a major new feature to be supported by the 5G core network and radio access network. A network slice, or 5G slice, a collection of 5G network functions and specific RAT settings that are combined together for the specific use case or business model [NGMN15]. Each slice can be seen an independent logical end-to-end network tailored to the needs of a specific set of use cases, running on a shared physical infrastructure. The concept of slicing was made possible by a number of innovations in networking, as presented in Section 5.2.1.

Slices span over the entire network architecture from end to end, including the core network and the radio access, each bringing in different sets of challenges. The core network, in fact, enables the allocation of specific network function instances to support the requirements of each slice, and a certain degree of overprovisioning, whereas the spectrum is a scarce resource, and the air interface needs to be shared among all of the slices. For this reasons, RAN slicing requires a special attention, as detailed in Section 5.2.2.

5.2.1 Network slicing enablers

On the core network side, slicing will be made possible in 5G by numerous recent developments in networking we refer to as enablers, which we describe in this section.

Network Functions Virtualisation

Nowadays, network entities are implemented on dedicated hardware running specifically targeted firmware. However, as technological evolution is becoming faster and customers more demanding, this approach is showing its limitations in terms of supply, flexibility and upgradability.

Network Function Virtualisation, on the other hand, is a paradigm that foresees the deployment of network functions as software modules running in a virtualized environment (virtual machines and/or containers) on general purpose servers. The advantages of NFV are manifold: network functions become more easily upgradable, speeding up the deployment of new functions and the maintenance of legacy ones. From an economic perspective, it opens the market to a larger variety of suppliers, driving down the costs of installing and operating a virtualized network infrastructure with respect to a traditional one.

Furthermore, NFV enables network function decomposition, a pattern wherein large and multipurpose network functions are disassembled into smaller interconnected modules. This is beneficial in multiple ways: first, only the necessary modules to support a specific use case can be implemented when needed; second, specific modules can be located in different areas of the network upon need. For instance, functions supporting ultra-low latency communications can be placed closer to the edge of the network.

In general, NFV enables the creation of network functions on demand, their optimized placement into the network infrastructure (i.e. in the core cloud or into distributed datacentres), and their interconnection to satisfy a specific and localised requirements in a dedicated manner.

Network Functions Virtualisation requires a dedicated architecture for the deployment, management and orchestration of interoperable virtualized network services, which has been specified and standardized by the ETSI in [ETSI14-MAN001].

Software Defined Networking

Alongside NFV, Software Defined Networking is a core technology involved in the transformation of the telecommunication industry. SDN drives the transformation of “boxed” network entities into network functions, and of protocols into APIs, leading towards the convergence of the telecom and the IT industries [5GPPP2017b].

One of the founding principles of SDN is the separation between the forwarding plane, i.e. the set of network resources dedicated to forwarding data traffic, and the control plane, which includes the functions that control one or more network devices. This principle, in combination with the softwarisation of the network control, provides the network administrators with a centralized control interface. Furthermore, since network control is made via standard APIs, it simplifies the interoperability between network functions provided by different technology providers.

CUPS: Control and User Planes Separation

CUPS - Control plane and user plane separation – is a concept specific to the 3GPP architecture first introduced in Release 14 [3GPP17-23214]. In previous releases, some core network entities (notably the SGW and the PGW) include both control plane and user plane functions, and are thus connected by reference points that support both control plane and user plane traffic.

Following the CUPS principle, these entities are split into their control part and data part; a new set of dedicated reference point is then defined to properly connect them, avoiding mixed interface. In this way, the CP and UP of these functions can evolve and be adapted and improved independently from each other. The most important advantage, however, stands in the fact that the user plane can scale independently from the control plane: a single control plane entity can in fact be connected to multiple, conveniently placed, user plane entities.

CUPS was introduced specifically to address shortcomings in the initial design of the LTE EPC. The separation of the control and user plane of the core network is one of the founding elements of the newly designed 5G Service Based Architecture, which is described in detail in the next section.

DECOR and eDECOR

Dedicated Core Network (DECOR) and enhanced Dedicated Core Network (eDECOR) are concepts standardized by the 3GPP in technical reports respectively in Release 13 [3GPP14-23707] and Release 14 [3GPP16-23711]. DECOR is a concept developed in LTE in order to enable the development of independent sub-networks assumed to serve only a subset of the subscriber. These sub-networks include control and user plane entities of the EPC (one or more MMEs, one or more SGW/PGW). eDECOR extends such concept, by improving the network selection procedures, and by allowing the same RAN to be connected to different core networks, operated by different PLMNOs, each possibly implementing dedicated core networks within themselves. The main difference between DÉCOR in 4G and slices in 5G is that a UE can only have one dedicated core network at a time but a UE can have multiple simultaneous slices in 5G.

The concepts of DECOR and eDECOR were instrumental to define some of the founding ideas of network slicing in 5G: each slice in 5G can in fact be seen as an independent logical network. With respect to (e)DECOR, the flexibility provided by the 5G architecture, by the virtualisation of network functions and by the SDN principle, allow multiple slices each composed by a custom set of tailored network functions to run on the same physical infrastructure.

5.2.2 RAN slicing

This section investigates the slicing concept in the 5G Radio Access Network (RAN). The objective is to identify the plausible options for implementing the slicing concept at the RAN level by the Mobile Network Operator (MNO) to respond to the needs of verticals. Considering the previously described three main services for 5Gs separately would lead to very different RAN designs and architectures. The slicing concept has then emerged as an efficient way for serving all these services with a multitude of requirements on a common infrastructure.

The typical C-V2X environment will be composed of a mix of these three types of services.

Slice multiplexing options

In [5GN17-D32], slice multiplexing options can be applied on different levels of the Radio Access Network protocol stack. This yields to four options for RAN slicing:

- Option 1: this option corresponds to very specific situations: a standalone network with its specific spectrum and infrastructure. It has to be discarded due to economically reasons.
- Option 2: in this option, only transmission functionality is share between network slices while all other functions are slice specific. This option can be used for example for mixing different radio access technologies.
- Option 3: in this option the PHY-MAC is common to all slices and the higher layer, starting from RLC, are slice specific. This option can be seen as an intermediate degree of isolation and customisation.
- Option 4: this option corresponds to a network implementing CN slicing with a shared RAN. It can be seen as a 3GPP multi operator core network using a QoS differentiation at the RAN level.

Hence three options (2, 3 and 4) can be considered for implementing slicing at RAN level and may coexist together, each providing different degrees of complexity and flexibility.

New radio RAN slicing enablers

To address the needs of specific services and requirements, the RAN part of network slice may have different radio configurations (combination of L1 resource component, L2 configuration/functions, and RAN infrastructure). The following new radio features of 5G help implementing the RAN slicing concept. Most of them will be studied in the 5GCar WP3:

- **Flexible numerology :** a first radio feature that is regarded as an essential enabler for RAN slicing is the flexible OFDM numerology concept for 5G. A numerology is defined by a Sub-Carrier Spacing (SCS) and a Cyclic Prefix overhead. The flexible numerology defines different subcarriers spacing’s (SCS). At higher SCS, symbol and TTI durations decrease, thus enabling lower latency for URLLC communications. Higher SCS means lower latency but it requires higher channel bandwidth.

- **Mini-slot:** this concept refers to the same concept of short TTI in LTE. In conventional LTE network, scheduling is performed at sub-frame basis and each sub-frame is 14 symbols length and 1 ms duration. In New radio mini-slot length can be 2, 4 and 7 symbols. Mini-slots may be used for fast flexible scheduling of services (this concept will be studied in 5GCar WP3 for C-V2X applications)
- **Downlink pre-emption:** flexible numerology and mini-slots are not sufficient for meeting the requirements of URLLC services. Indeed, URLLC applications are characterized by a very bursty traffic. Instead, a pre-emption mechanism can be used: in the downlink, as the scheduling is performed by the base station, already allocated resources used by eMMB transmissions can be pre-empt for URLLC urgent transmissions. (see “Efficient pre-emption-based multiplexing of services” in the future [5GC18-D31])
- **Grant free in uplink:** In a typical LTE network, the selection of shared data channels for uplink (UL) is scheduling based, and like in downlink the scheduling is controlled by the eNB. UE sends an UL scheduling request to the base station. When the eNB receives the scheduling request, it sends an UL resource allocation to the corresponding UE. Then the UE transmits data on the allocated resource. An issue with this mechanism is that the signalling overhead for the scheduling process can be large, especially in cases where the amount of data transmitted is small, and the initial delay is greater than 1 ms. A proposal of an uplink grant-free transmission scheme could solve these issues for urgent transmissions. The eNB blindly detects active UEs and decodes received uplink transmissions if possible. Furthermore for urgent transmissions, UE can use higher transmission power so that its transmission can be decoded at the receiver.

Network chaining for slices

A slice specific network chaining can be performed so that the QoS of the underlying service can be met. For example, for PDCP and the RLC, depending on network needs (message size, throughput and delay requirements) certain functions can be either suppressed (e.g., header compression, ciphering) or modified (e.g., segmentation, re-ordering, ciphering) in order to reduce the processing times and improve E2E network performances.

An important issue for slice implementation in the RAN is related to the identification, by RAN network functions, of the slice it belongs to and of the potential specific treatment that is to be applied to it. I would be very complex for a scheduler to handle directly the slice identifier, knowing the possibly large number of slices in the network (The support of network slice in RAN should allow to support a large set of slices in terms of hundreds or even thousands [R2-1800259]). A slice management entity has to facilitate this task for the scheduler (and other network functions).

Slice identification

An important issue for slice implementation in the RAN is related to the identification, by RAN network functions, of the slice it belongs to and of the potential specific treatment that is to be applied to it. I would be very complex for a scheduler to handle directly the slice identifier, knowing the possibly large number of slices in the network (The support of network slice in RAN

should allow to support a large set of slices in terms of hundreds or even thousands [R2-1800259]). A slice management entity has to facilitate this task for the scheduler (and other network functions).

RAN slicing issues

While the above detailed radio features facilitate the implementation of slicing, there are still many challenges to be overcome when it comes to resource allocation to slices, and to slices management

- Resource allocation: availability vs overprovisioning (Reserving wide-band resource for sporadic C-V2X traffic would cause a considerable waste).
- SLA monitoring : monitoring of QoS for both physical and virtual resources shared between slices
- Mobility management : availability of slices in a given geographical area

RAN openness to tenants: selection of network function that can be opened to third parties

5.3 Network slicing for V2X applications

5.3.1 The 5GCAR approach

Compared to previous and current 5GPPP projects, 5GCAR focuses entirely on vehicular applications and services, rather than just considering them as one among many other use cases for 5G. The advantage of this approach stands in the level of detail 5GCAR can attain in analysing and studying the requirements of the different services that can be described with the umbrella term “V2X”. In Figure 5.3 we give an illustrative example, wherein multiple of the innovative concepts considered in 5GCAR are included. In it, we considered a multi-tenancy scenario, with three tenants: the operator (which is also the mobile service provider, owning and operating the physical network infrastructure), an automaker offering connected services, and the cooperative perception, i.e. the service enabling road users to periodically exchange their status (position, speed, heading, etc.).

Each of the tenants offers different types of services, with custom requirements and routing paradigm, which are mapped onto separate slices.

The **operator** controls the evolved Mobile Broadband slice, which serves the on board infotainment, which involves the reception of data flows (audio, video) from servers on the public network. This slice is thus conceived to provide best effort wideband connectivity to the internet. Most of the control and user plane functions (packet routing and forwarding, session management and charging) are thus concentrated in the central cloud. The flexibility of the 5G architecture will enable, in the future, the implementation of advance caching of popular content in the edge cloud for optimizing the network efficiency; these features are however outside of the scope of 5GCAR.

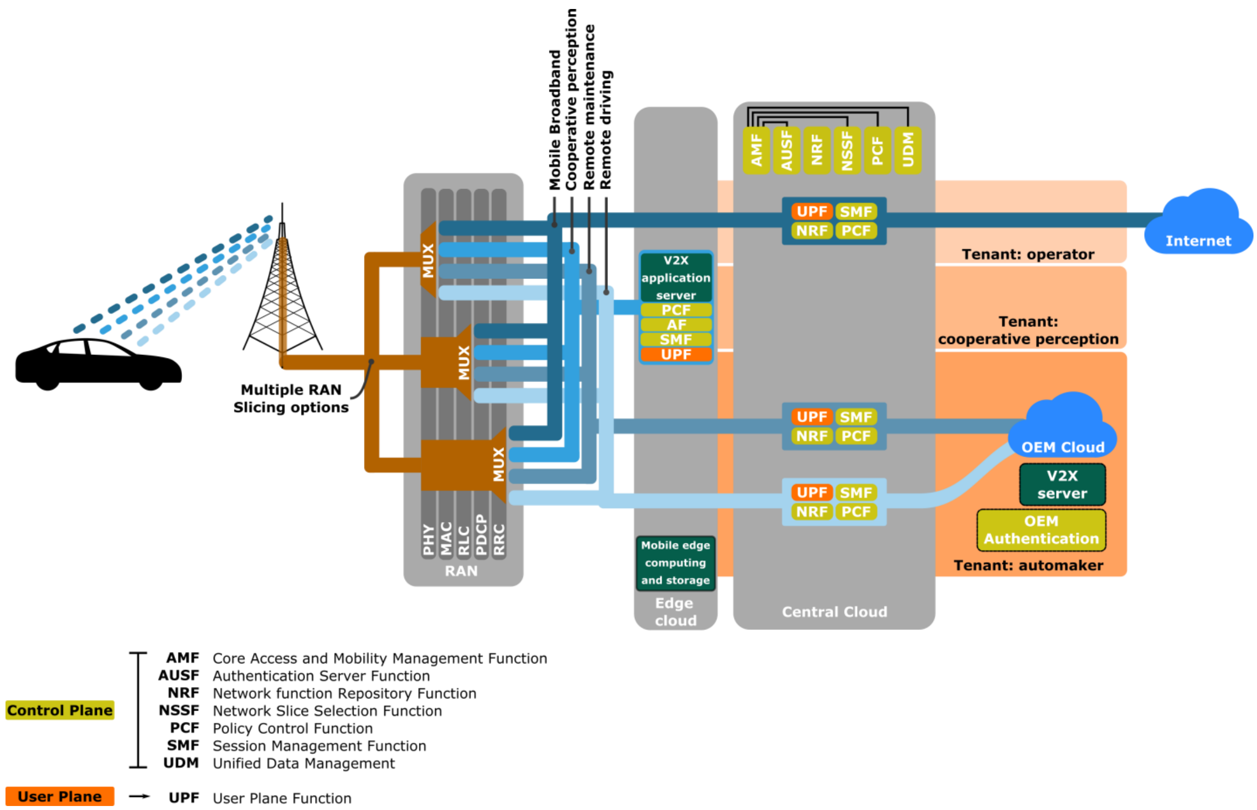


Figure 5.3: Example of 5G V2X sliced network scenario

The **cooperative perception** service is characterised by the local validity of the traffic generated by road users, as status messages are meant to be distributed to other road users in proximity in the shortest possible delay. For this reason, the cooperative perception slice only extends into the RAN and the edge cloud, wherein the V2X application server is located.

Automakers (herein also referred to as OEMs) are increasingly investing into developing tailored services to offer to their customers along with the vehicle. In this preliminary architecture, we consider two of them: remote maintenance, and remote driving, which come with different sets of requirements.

Remote maintenance could be to a certain extent associated to a mMTC-like type of service, wherein vehicles would upload to the manufacturer's cloud small amount of diagnostic data with relatively low frequency (reasonably, not more than daily), with low priority, updating the service with the current state of the vehicle's apparatus. Furthermore, the cloud feeds back maintenance information and alerts to the driver when an intervention on the vehicle is required. This type of traffic allows for a lower-cost deployment of network functions in the central cloud of the operator, which forwards the data to the OEM's own cloud.

On the opposite side of the spectrum, the remote driving use case may be associated to a URLLC-type communication. In this case, the automaker's V2X server issues real time commands to remotely drive the vehicle, which need to be received within few ms and with high reliability. This use case may also be associated to a real time video flow from the vehicle to the V2X function, to feedback in real time the result of the driving manoeuvers. These types of traffic require a very lean network architecture, wherein as few functions as possible are crossed by the data flow to reduce the processing delays to a minimum.

5.3.2 5GCAR: vehicular communication paths

Vehicular communications are meant to cover a very wide and diverse set of applications and requirements. Figure 5.4 provides a global overview of the V2X system architecture considered in 5GCAR, including all of the functional entities involved. The end-to-end specification and design of the V2X 5G architecture overarches all of them, and is expected to cover different communications among vehicles (and more in general road users), the network, and road-side infrastructure.

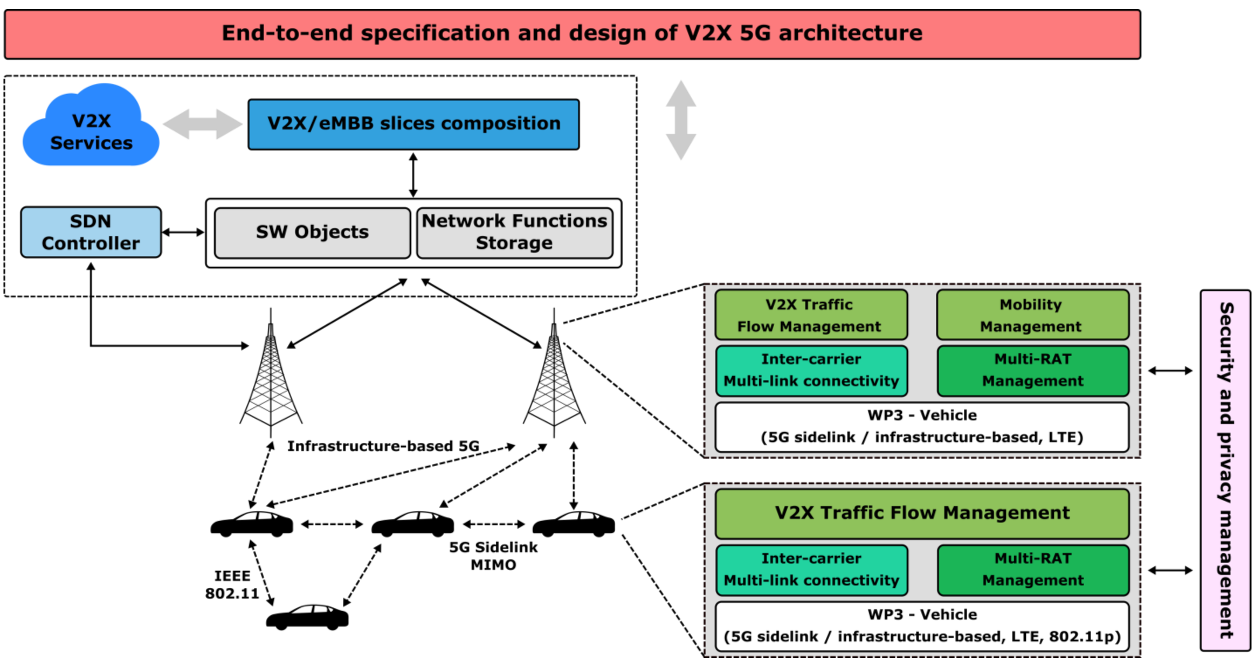


Figure 5.4: 5GCAR V2X system and architecture

These communication patterns, which are further highlighted in Figure 5.5, result from different types of message exchanges, each of which involve different actors. Different types of message may be directed to servers located in the Internet, locally routed by the infrastructure (in local breakout configuration), or be transmitted over direct V2V transmissions over the PC5 interface, these latter supporting both unicast and broadcast. This diversity of scenario hence requires a redefinition of the established concept of "end to end", since in 5GCAR a road user (either a vulnerable user or a vehicle) constitutes one end, whereas the other end can either be a remote server, a server located at the edge of the cellular network, or another road user in proximity.

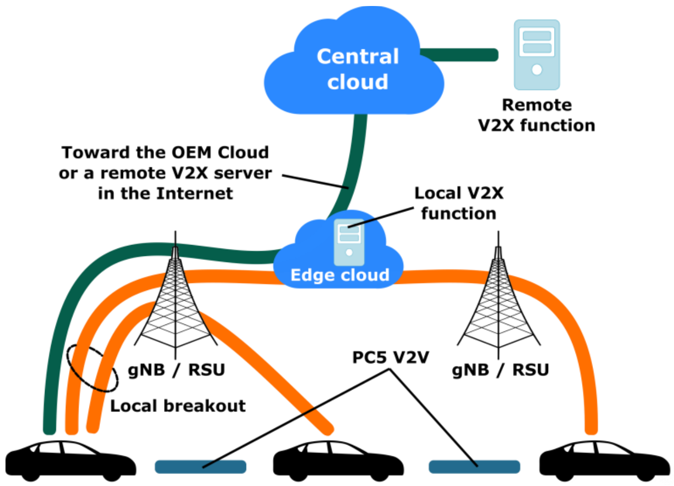


Figure 5.5: communication paths and end-2-end definitions in 5GCAR

5.4 New network features for V2X emerging from 5GCAR technical components

Using 3GPP Rel-15 5G network architecture [3GPP17-23501] as the baseline, the 5GCAR network architecture is being developed for supporting a set of targeted challenging V2X applications and services with a set of TCs incorporated. Those TCs which are being developed under the scope of 5GCAR WP4 are presented in Section 4. These TCs are not necessarily inter-related to or inter-dependent on each other, even though some TCs share certain commonality in terms of design target, network feature, and so forth. 5GCAR enabled efficient E2E network slices for targeted V2X support may benefit from the innovation brought by these TCs. It is noted that all the TCs are still under ongoing development. Further verification of individual TCs in terms of e.g. performance gain in terms of latency, reliability, feasibility and practicality is needed. These will be done for the final deliverable of 5GCAR WP4, D4.2. However, potential impacts of the individual TCs on the network architecture, e.g.: network elements, network functions or network features either on RAN, CN or E2E level as well as C-plane or U-plane, are preliminarily identified and presented in the following subsections. There is one subsection per each TC, starting with a summary table of potential network impacts and then followed by some network signalling procedure(s) illustrating for examples how to implement some of the impacts (dashed line is for C-plane, solid line is for U-plane and novelty is highlighted in red colour).

5.4.1 RSU enabled Smart Zone (SM-Zone)

TC description

Table 5.2: RSU enabled Smart Zone description

Network Entity/ Network Function	New Feature and/or Extension of existing Functionality
UE	Control plane (RRC, NAS): • SM-Zone discovery, selection and service request • SM Zone based resource allocation and mobility management for idle and active UEs User plane: • SL transmission mode using 5G RSUs with RSU based collective D2D relays
gNB	Control plane (RRC): • SM-Zone configuration and control towards UE • SM-Zone configuration and control towards RSU, especially with UE- or Hybrid-type RSU • SM Zone based resource allocation and mobility management for idle and active UEs Control plane (Xn and Nx AP): • SM-Zone driven coordination and control across multiple macro cells for configured SM-Zones, UEs and RSUs thereof User plane: • gNB-RSU packet forwarding
5G RSU	Control plane related to UE-type RSU (RRC): • SM-Zone configuration and control towards serving gNB with multi-operator support • Spatial reuse resource allocation for SL Tx Control plane related to gNB type RSU (RRC towards UEs and Xn AP towards gNB): • SM-Zone configuration and control towards UE • SM Zone based resource allocation and mobility management for idle and active UEs • SM-Zone driven coordination and control User plane: • SL transmission mode using 5G RSUs with RSU based collective D2D relays

	• relaying received D2D messages for UE individually; • rebroadcasting received D2D messages for a number of relevant UEs collectively; • RSU-RSU data forwarding or distributing of local contents of V2X; • MBMS localised along the road; • connectionless packet access for V2X with RSU packet forwarding and multi-operator support
AMF, SMF, AF	Management of RSUs and SM-Zones as well as UE contexts related to SM-Zones and V2X services
UPF	Distribution of local MBMS to RSUs, connectionless packet access for V2X, depending on RSU types and V2X services of SM-Zones

Call flows

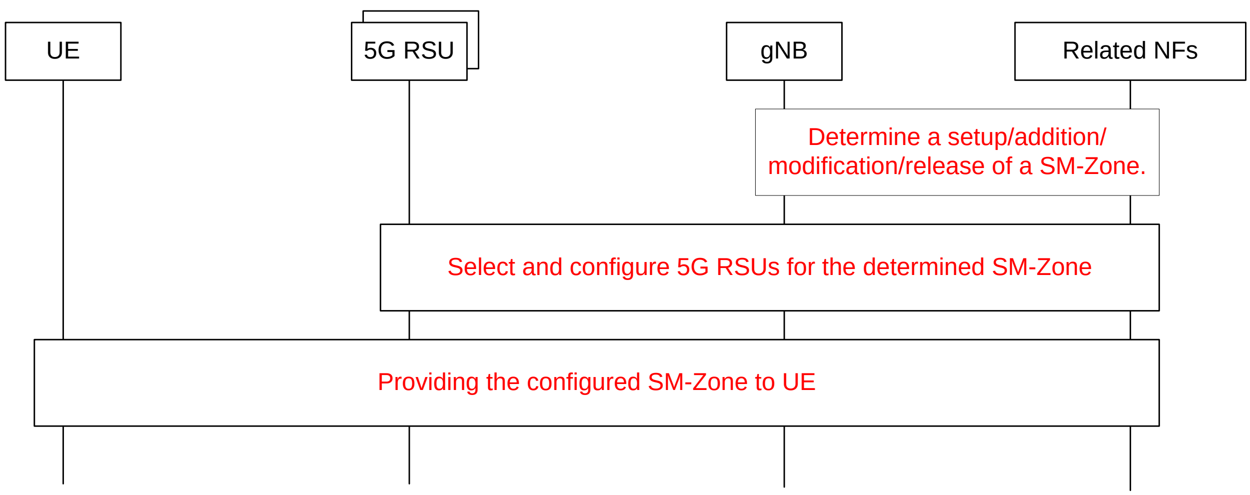


Figure 5.6: Overall RSU enabled SM-Zone concept

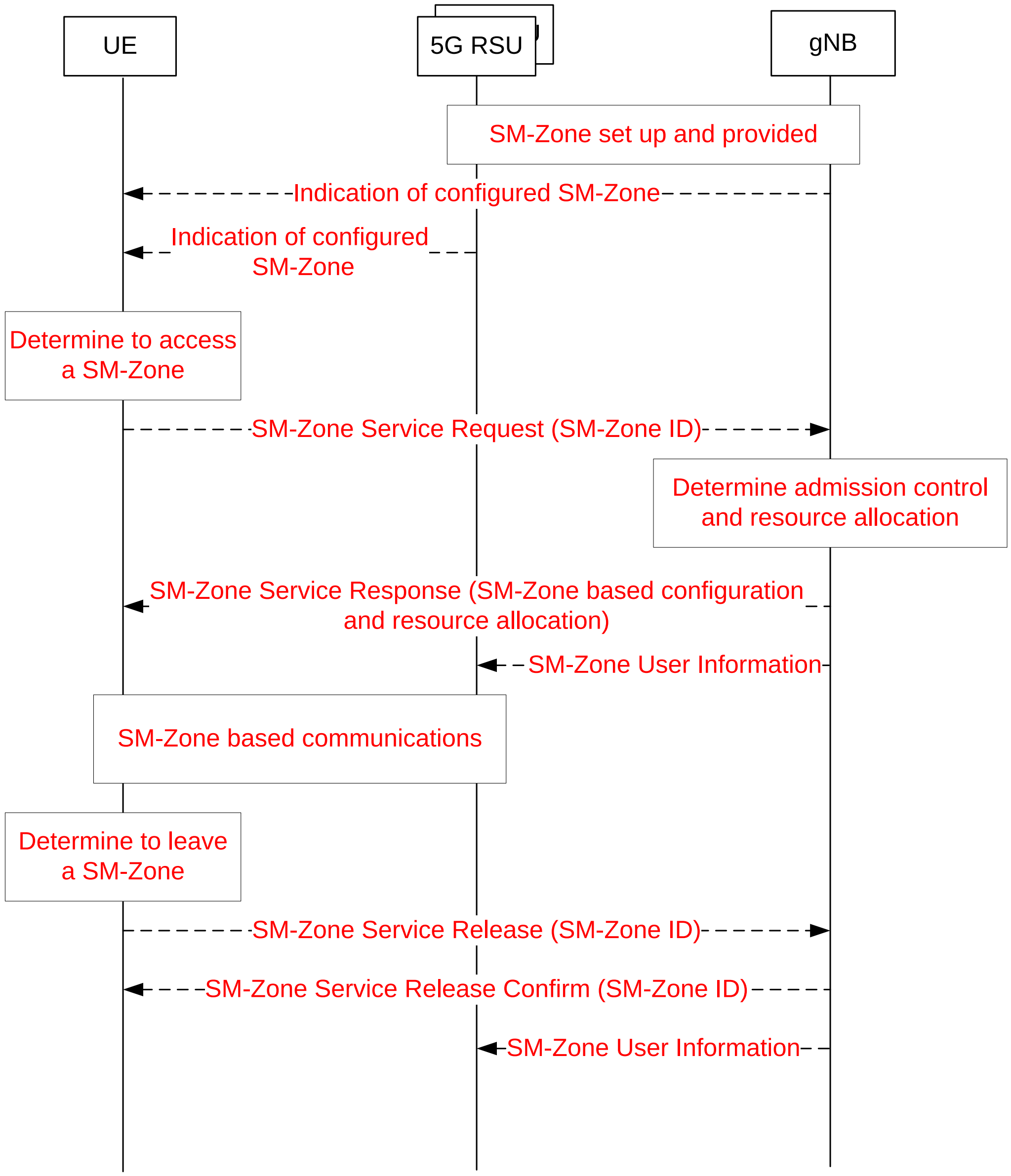


Figure 5.7: Example of C-plane SM-Zone configuration and UE access control

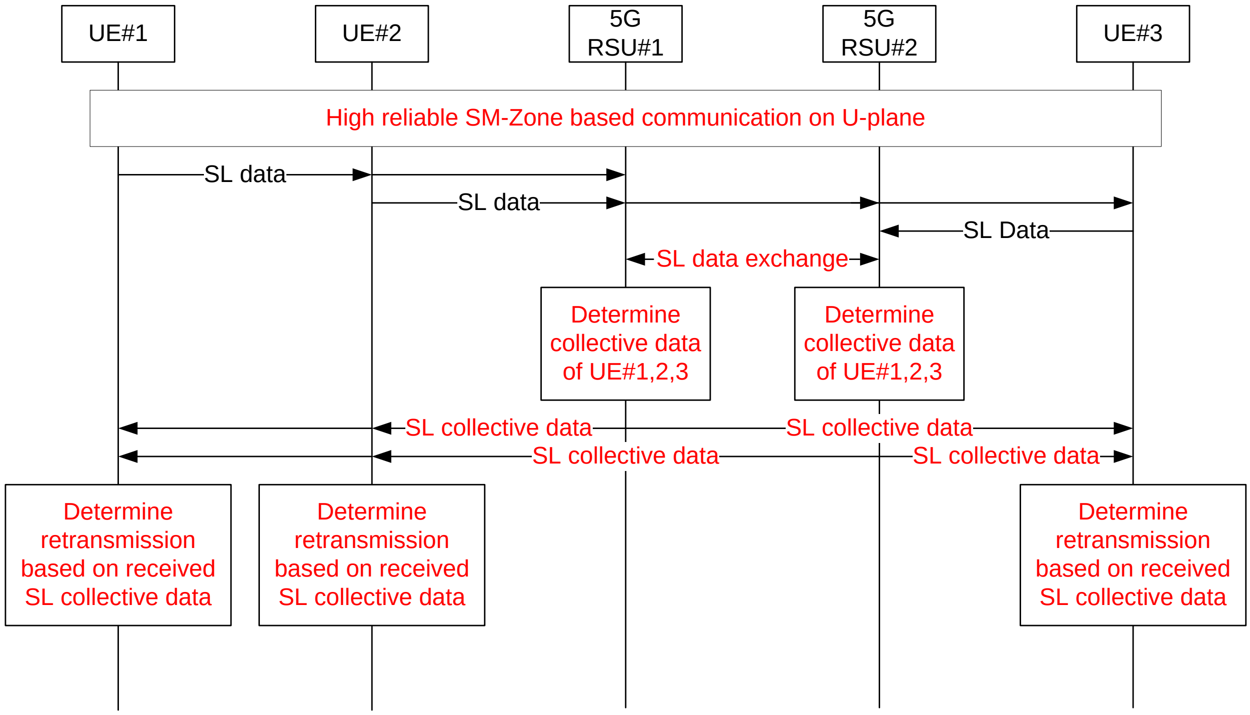


Figure 5.8: Example of U-plane SM-Zone based reliable SL communications with UE-type RSUs

5.4.2 SON based multi-mode RSU for efficient QoS support and congestion control

TC description

Table 5.3: SON based multi-mode RSU for efficient QoS support and congestion control description

Network Entity/ Network Function	New Feature and/or Extension of existing Functionality
UE	Control plane: (RRC, NAS): • Enhanced SL and V2X traffic measurement and reporting on road traffic conditions • Reception of RSU mode change notification
gNB	Control plane (RRC): • Enhanced UE and RSU measurement and reporting on road traffic conditions • Triggering and/or reconfiguring operation mode of RSUs
RSU	Control plane (RRC, NAS): • RSU measurement and reporting on road traffic conditions • SON based management of different RSU operation modes • Mode change notification towards UE and network
AMF, SMF, AF (SON related NFs)	Management of SON based RSUs instead of or in addition to that provided by gNB.

Call flows

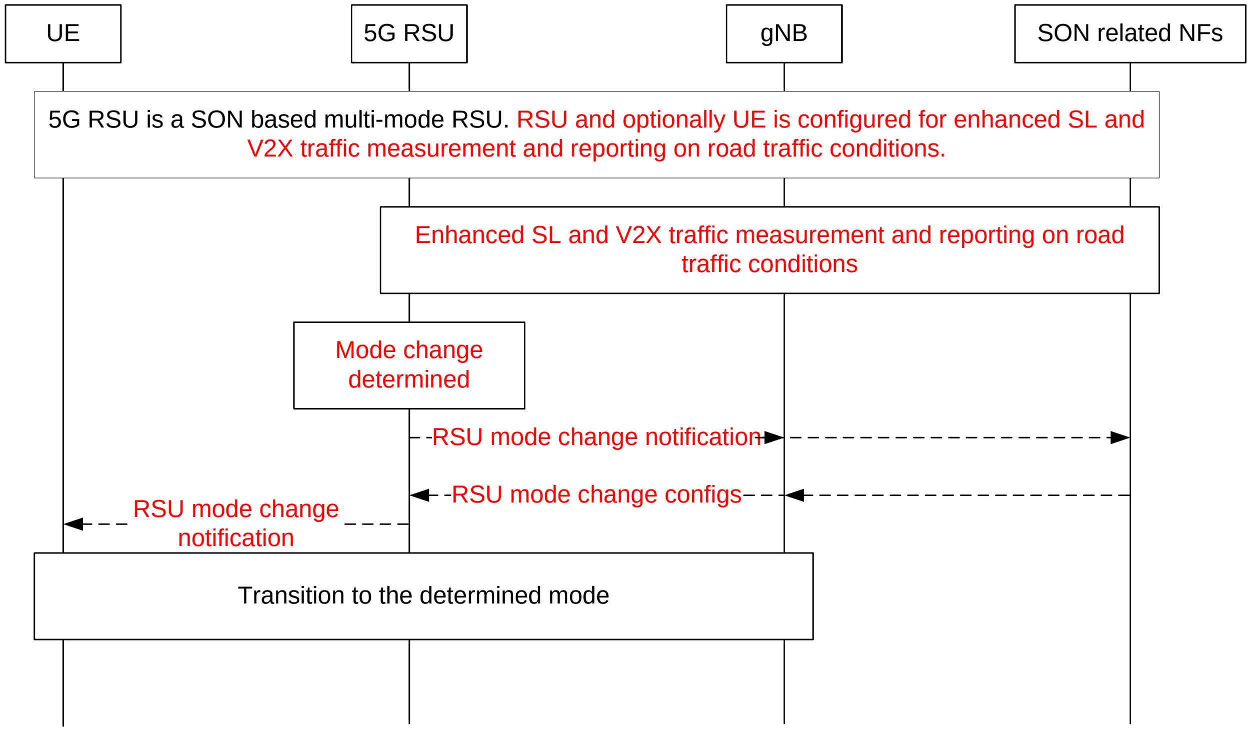


Figure 5.9: SON based adaptive reconfiguration of multi-mode RSU

5.4.3 SL and Uu multi-connectivity for high-reliable and/or high data rate V2V communication

TC description

Table 5.4: SL and Uu multi-connectivity for high-reliable and/or high data rate V2V communication description

Network Entity/ Network Function	**New Feature and/or Extension of existing Functionality**
UE	Control plane (RRC and MAC): • UE initiated establishment of a secondary radio connection with a serving gNB over Uu, also referred to as Secondary Uu RB (SURB) for the ongoing V2V services over SL as the primary connection; • Transmission of common SL/UL BSR for the primary SL and the SURB to serving gNB

	User plane (PDCP): • Transmission of PDCP Reception Status Report (PDCP_RSR) from receiving UE of the primary SL to serving gNB • UL transmission of SL data for duplication or split over SURB • DL reception of SL duplication or split data from the serving gNB on SURB
gNB	Control plane (RRC and MAC): • SURB configuration and control with one common virtual PDCP entity at the gNB side for all involved UEs of the primary SL: SURB has no PDU session towards CN • Reception of the common SL/UL BSR User plane (PDCP): • Reception of PDCP_RSR • UL reception of SL data for duplication or split over SURB from the Tx UE of the SL • DL transmission of SL duplication or split data on SURB

Call flows

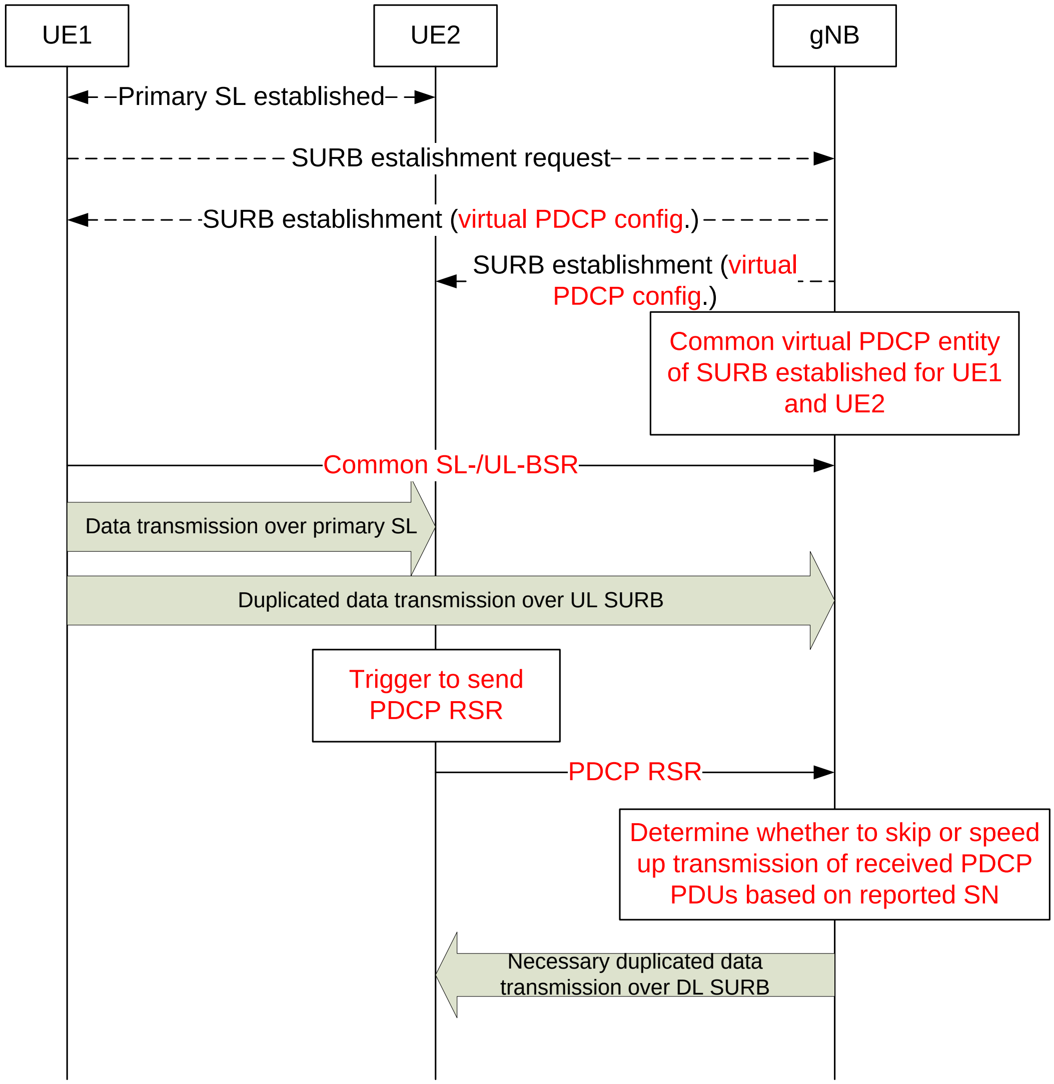


Figure 5.10: SL multi-connectivity with Primary SL and Secondary Uu RB

5.4.4 Location aware scheduling

TC description

Table 5.5: Location aware scheduling description

Network Entity/ Network Function	New Feature and/or Extension of existing Functionality
V2X app	Provide additional information about geographical area of relevance of packets Provide additional information about vehicle position, mobility speed and trajectory
gNB	For local traffic scheduled at the gNB (i.e., V2V via gNB or localised V2X traffic) New functionality to map the geographical constraints into QoS for packet delivery CP: trigger adaptation of parameters (QoS, priority, etc.) for packet transmission/reception considering location-aware QoS constraints
AMF	Provide information about RAN status (load, etc.) Trigger adaptation of parameters (QoS, priority, etc.) for packet transmission/reception considering location-aware QoS constraints Selection of the most suitable gNB for packet delivery considering location constraints Update and configuration of Radio Bearers to accommodate location-aware scheduling
UPF	For traffic scheduled at the UPF (i.e., from external V2X servers) New functionality to map the geographical constraint into QoS for packet delivery
SMF	PDU Session Modification to accommodate location-aware QoS constraints

Call flows

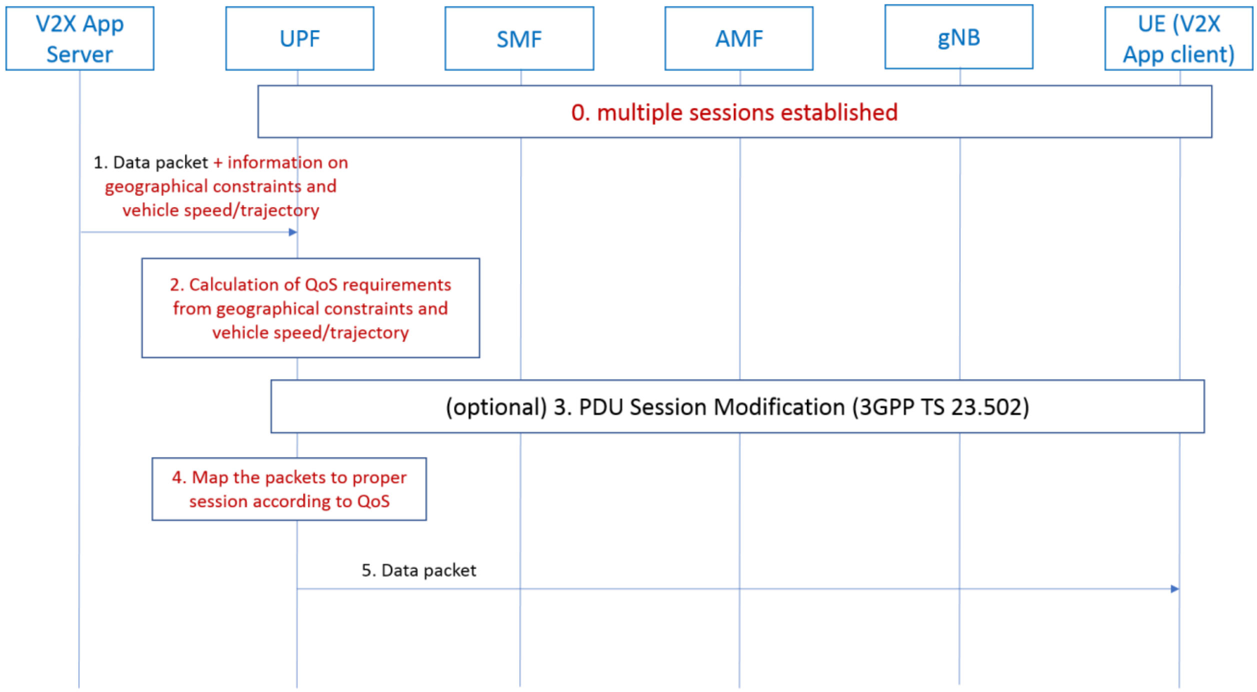


Figure 5.11: Example of location-aware scheduling realisation with capabilities of obtaining QoS information from geographical constraints of the packet implemented at the UPF

5.4.5 IaaS for vehicular domain

TC description

Table 5.6: IaaS for vehicular domain description

Network Entity/ Network Function	**New Feature and/or Extension of existing Functionality**
NEF	Exposition to AF (Nnef-Naf) of network information status useful for vehicular applicative purpose : Chanel quality info, UE Location, Cell-ID, Cell Load, Average Network Latency at access, core level, targeted Cell-ID (before handover), Interface between GMLC-NEF for location (already exists before 5G, should have to be activated)
UDR	Storing in UDR network information possibly averaged on a given windows which could be set.

eNB(RAN)	Send information (radio layer info) towards UDR
NWDAF	Provide analytics tools to identify network information status needed to be exposed to the NEF
BSS/OSS (API)	Exposition to service orchestrator of network info for service layer re-orchestration, re-configuration (optional)

Call flows

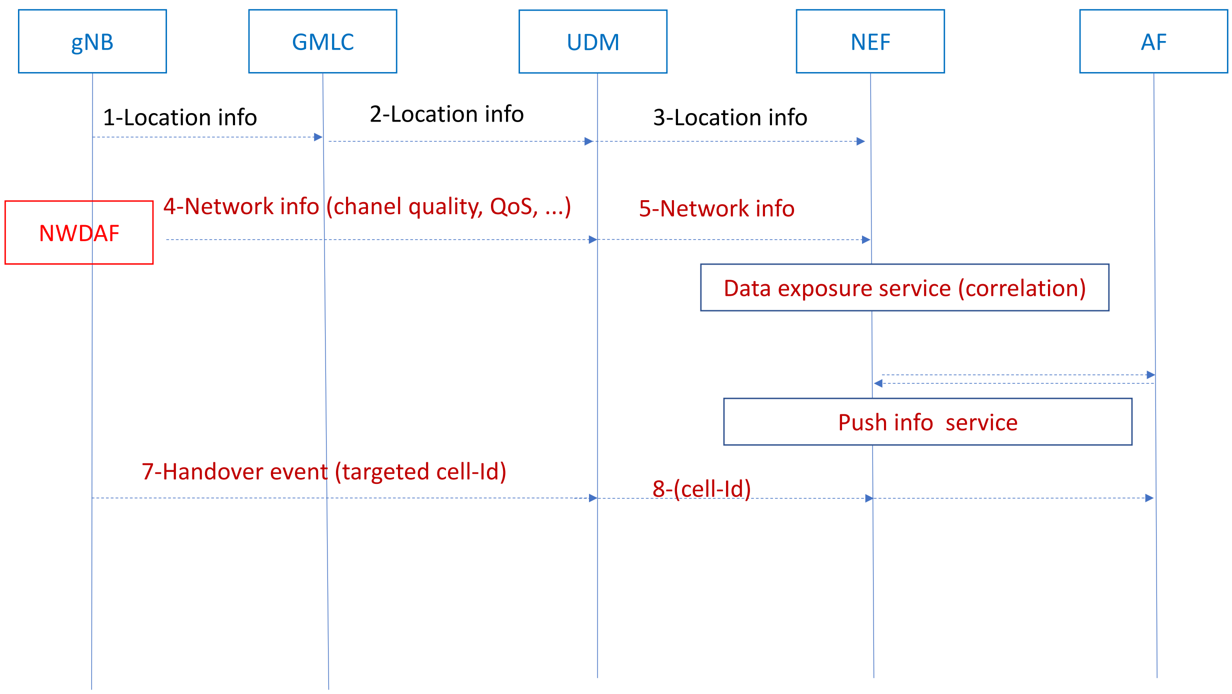


Figure 5.12: Example of exchange of network related information between gNB and NEF in order to be exposed to AF, through GMLC, UDM entities

5.4.6 Redundant mode PC5 + Uu

TC description

Table 5.7: Redundant mode PC5 + Uu description

Network Entity/ Network Function	New Feature and/or Extension of existing Functionality

UE	Control plane: • RRC: introduce a new redundant mode in Information Elements for TX and RX UE. • RRC: Modification of sl-V2X-ConfigCommon, SidelinkUEInformation, UEAssistanceInformation elements, and RRCConnectionReconfiguration message. User plane: • Transmit the same data packets concurrently over PC5 interface and over Uu interface for a given limited time. • Receive data packets from either PC5 or Uu interface (duplicate packets are dropped)
gNB	Control Plane • RRC: introduce an enhanced mode 3 scheduler for PC5 communications with a new redundant mode (PC5 + Uu). • RRC: Modification of sl-V2X-ConfigCommon, SidelinkUEInformation, UEAssistanceInformation elements, and RRCConnectionReconfiguration message. • PDCCH TX grants. SPS SL (addressed to SL-SPS-V-RNTI) or direct grant (addressed to SL-V-RNTI) concurrently with a SPS UL grant for the same data over Uu interface (addressed to UL-SPS-V-RNTI). • PDCCH DL assignment. Introduce a new RNTI (DL-SPS-V-RNTI) for V2X SPS DL over Uu concurrently with a SL reception for the same data over PC5 interface. User plane: • Forward packets received in UL over Uu to DL unicast or broadcast transmissions.
AMF	Possibly establish and Update the local e2e paths between the UEs over the gNB(s) in case of Handover.
V2X Control/PCF	Provision a V2X service to UEs for a given period with a redundant mode.

Call Flows

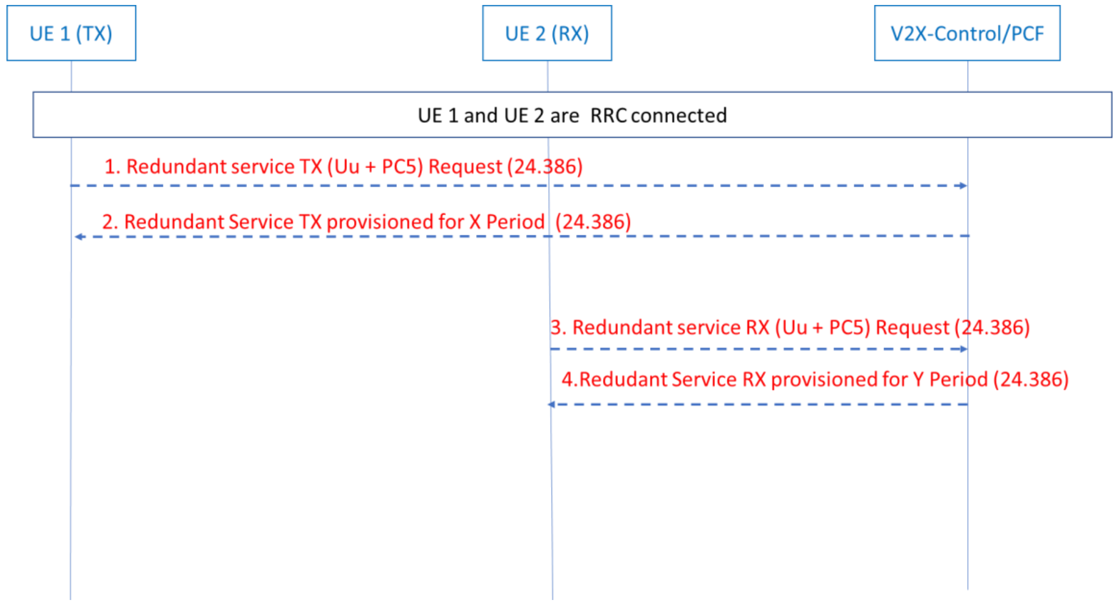


Figure 5.13: Redundant service provisioning

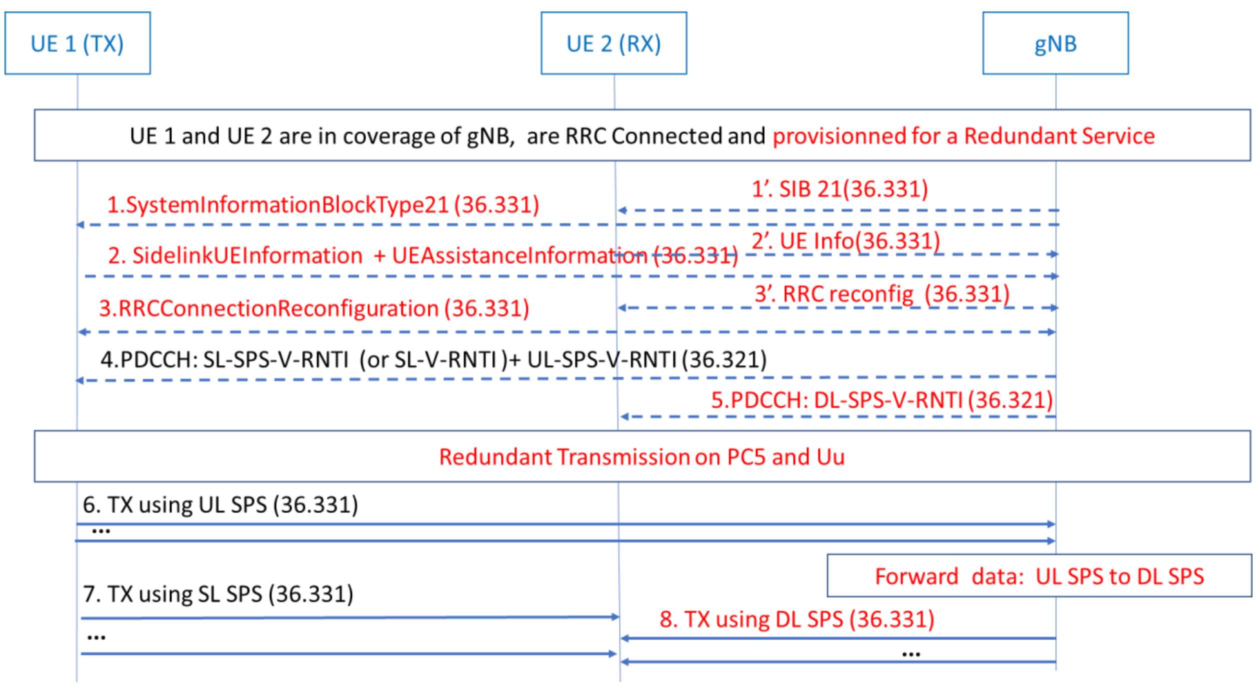


Figure 5.14: Redundant service call flow

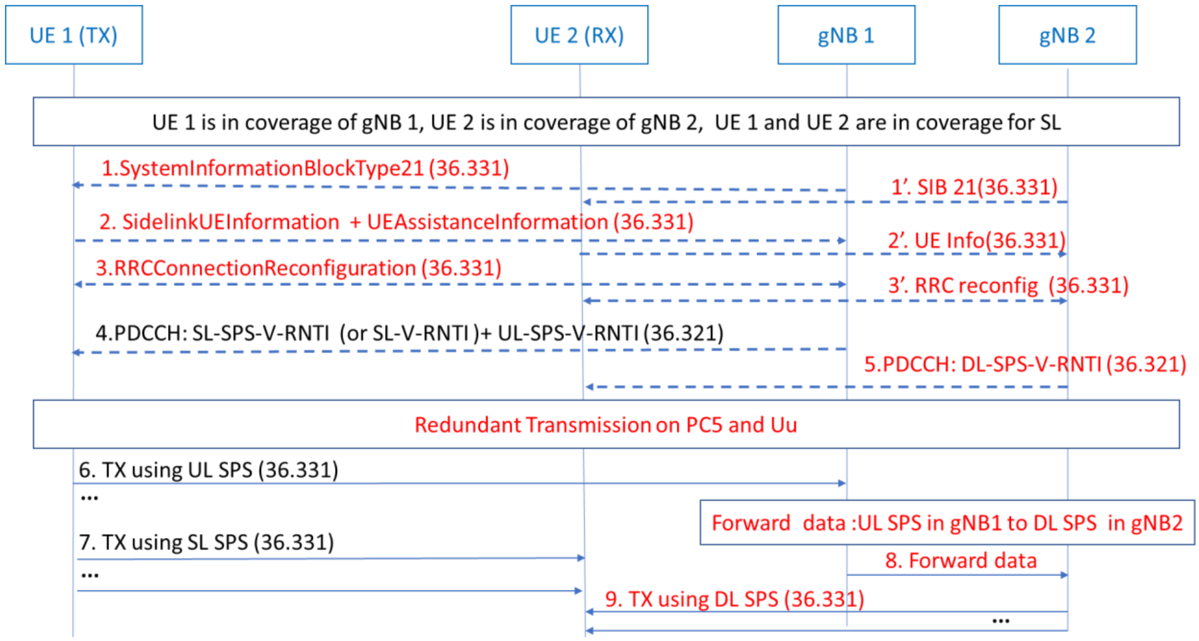


Figure 5.15: Redundant service call flow inter-gNB

5.4.7 Evolution of infrastructure-based communication for localised V2X traffic

TC description

Table 5.8: Evolution of infrastructure-based communication for localised V2X traffic description

Network Entity/ Network Function	New Feature and/or Extension of existing Functionality
UE	Control plane: • RRC, NAS: Request Establishment and update of the local e2e paths. User plane: • Transmit and Receive data packets over the local e2e paths.
gNB	Control plane: • RRC: Establish and Update the local e2e paths between the

	UEs over the gNB(s). • Update and Configure Radio Bearers Mapping Table (RBMT). • Inter gNBs coordination for local e2e paths establishment and update. User plane: • Forward packets over the local e2e paths, supporting multicast and unicast transmissions.
AMF	Establish and Update the local e2e paths between the UEs over the gNB(s). Update and Configure Radio Bearers Mapping Table.

Call flows

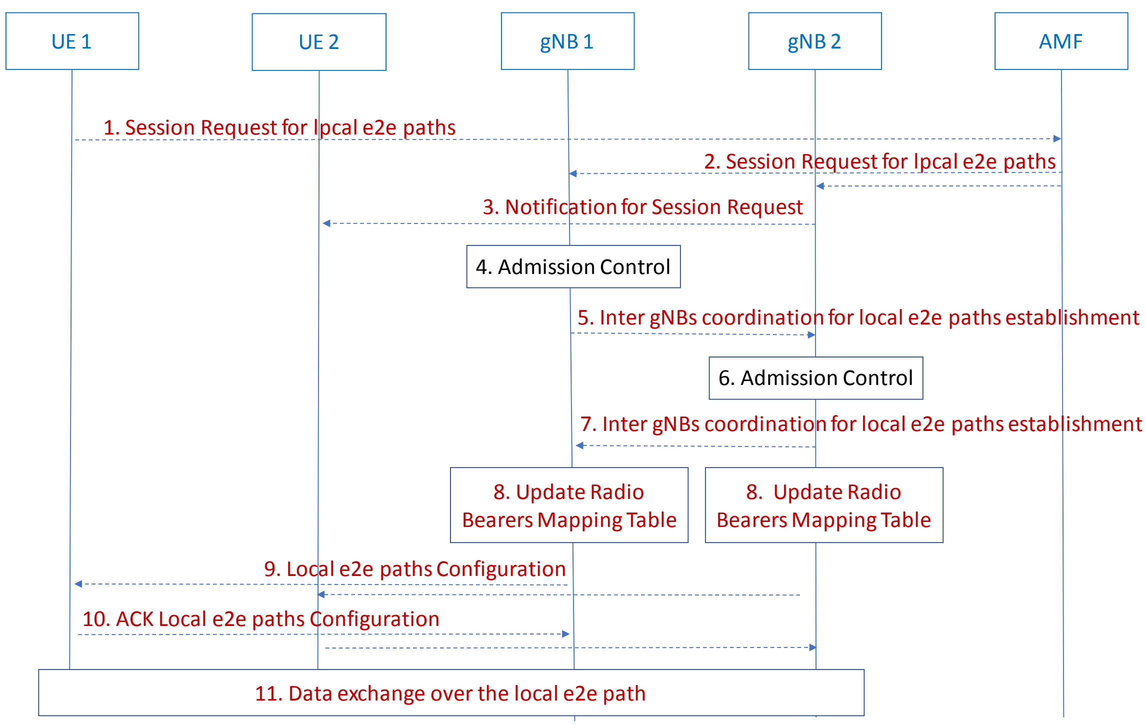


Figure 5.16: Evolution of infrastructure-based communication for localised V2X traffic

5.4.8 Use case-aware multi-RAT, multi-link connectivity

TC description

Table 5.9: Use case-aware multi-RAT, multi-link connectivity description

Network Entity/ Network Function	New Feature and/or Extension of existing Functionality
UE	Informing the network of the desired manoeuvre Computation of required completion time depending on the manoeuvre and the advance context information. Identification of the required network connectivity to complete manoeuvre time, through single or Multiple links or RATs.
gNB	Provide connectivity according to the requested criteria
NFs	Collection and Update of advance context information such as location and speed of all involved vehicles in the manoeuvre

Call flow

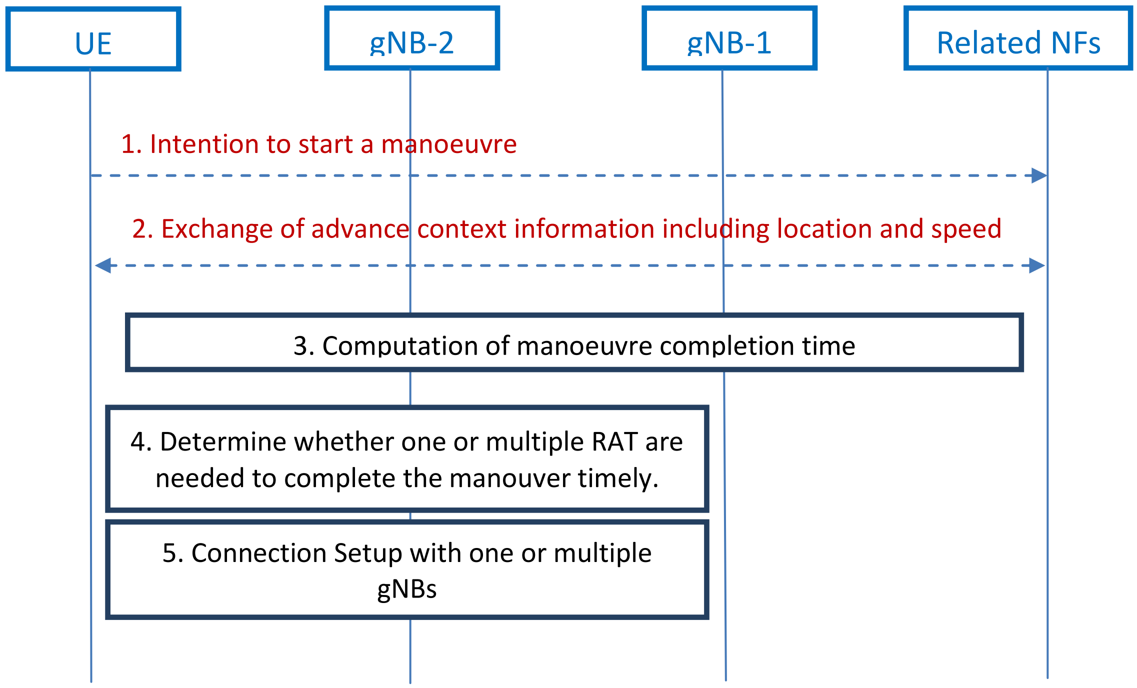


Figure 5.17: Use case-aware multi-RAT, multi-link connectivity

5.4.9 Multi-operator solutions for V2X communications

TC description

Table 5.10: Multi-operator solutions for V2X communications description

Network Entity/ Network Function	New Feature and/or Extension of existing Functionality
UE	Control Plane: • RRC: Switch between operators, inform about the restrictions in time/frequency • NAS: Receive the geographical information (operator split), multiple attach
gNB	Control plane: • RRC: use the restrictions in time/frequency provided by the UE, force roaming
AMF	Multiple attach, communication with other operators in case of operators' core network coordination, provision of geographical information (operator split)

UDM	Update in the locations, interactions with other operators' AMF

Call flows

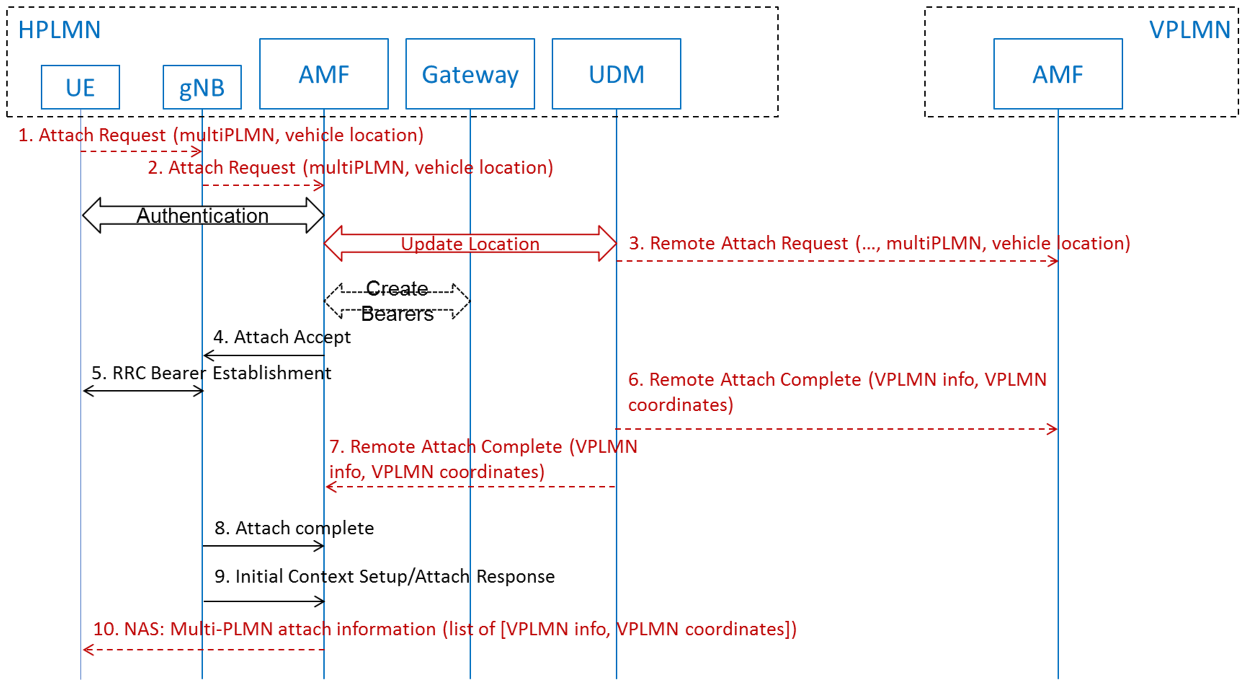


Figure 5.18: Multi-operator solution for V2X communications

5.4.10 V2X service negotiation

TC description

Table 5.11: V2X service negotiation description

Network Entity/ Network Function	New Feature and/or Extension of existing Functionality
AF	Provide the network with requests (type of service, QoS needs, file size, delivery deadlines, etc.) and additional information (vehicle mobility, trajectory, etc.) for negotiation Receive negotiation feedback
NEF	Receive negotiation request from the AF Provide the negotiation feedback to the AF according to the

	information received by other network functions of the system architecture
gNB	Provide information about AN status (current load, etc.)
AMF	Provide information about the RAN status (current load, etc.)
SMF	Provide information about the CN status according to reports of PDU sessions performance sent by the UPF
NWDAF	Provide statistical information about V2X-related slice or slices (such information might be obtained from information gathered by other network functions of the system architecture)

Call flows

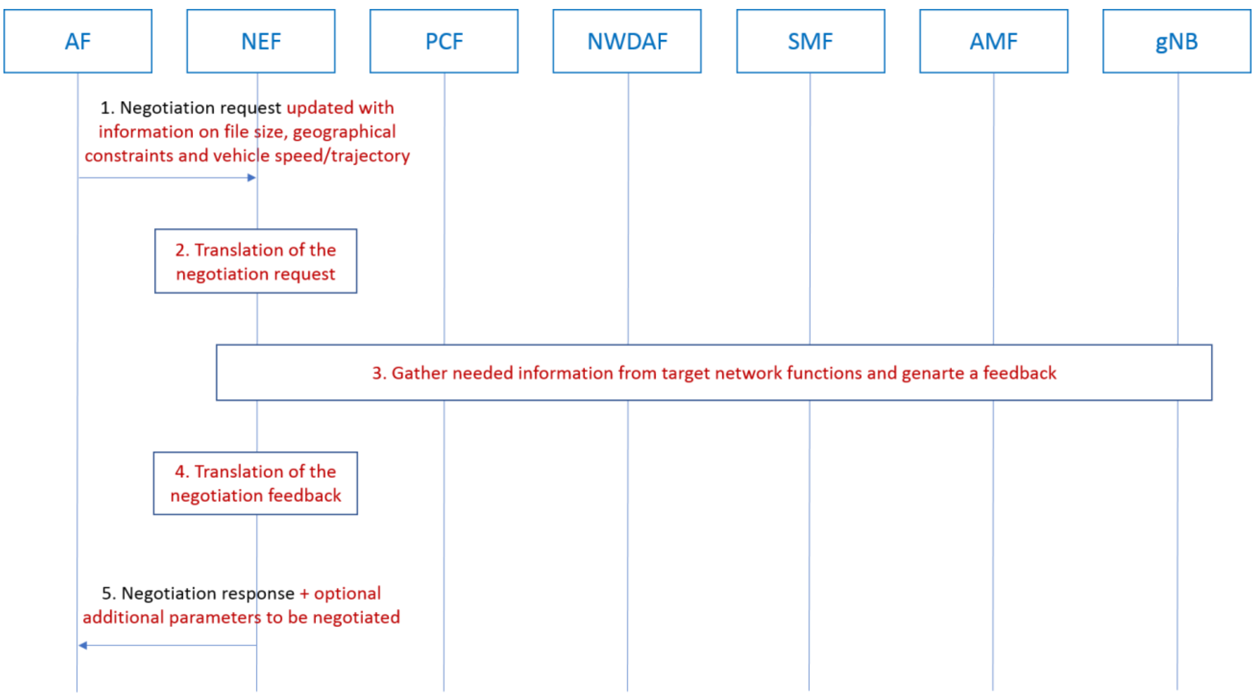


Figure 5.19: Example of a possible realization of service negotiation for HD local map acquisition where the negotiation request is sent by the AF and received by the NEF, which is in charge of transmitting the negotiation feedback to the AF

5.4.11 Edge Computing in millimetre Wave Cellular V2X Networks

TC description

Table 5.12: Edge Computing in millimetre Wave Cellular V2X Networks description

Network Entity/ Network Function	New Feature and/or Extension of existing Functionality
UE	Control plane: • Edge computing based resource allocation and mobility management for UEs User plane: • Transmit and receive data packets with RSU • Directional antennas with sectored beam pattern
RSU	Control plane: • Inter RSU coordination for edge computing services User plane: • Transmit and receive data packets with UE • Directional antennas with sectored beam pattern

Call flows

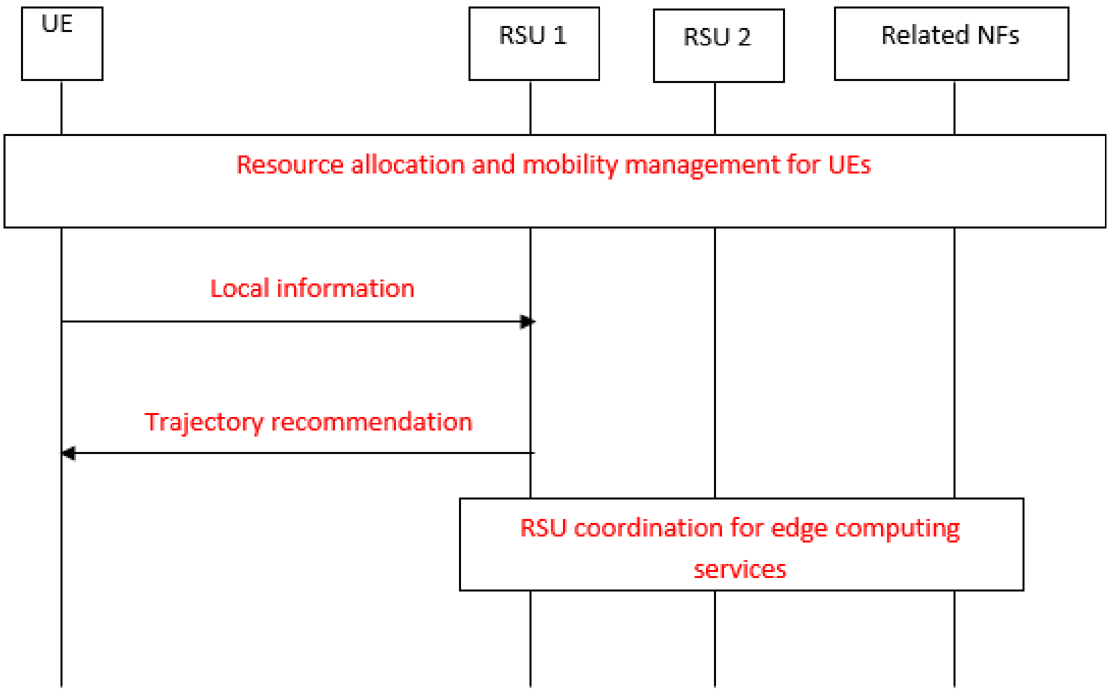


Figure 5.20: Edge Computing in millimeter wave cellular V2X networks

5.4.12 Dynamic selection of PC5 and Uu communication modes

TC description

Table 5.13: Dynamic selection of PC5 and Uu communication modes description

Network Entity/ Network Function	New Feature and/or Extension of existing Functionality
UE	Control plane: • RRC, NAS: Request selection of the appropriate communication mode. • Receive selected communication mode and the corresponding Uu or PC5 configuration from the gNB.
gNB	Control plane: • RRC: Dynamic selection of appropriate communication mode. • Notification to the UEs about the selected communication mode

	and provide the corresponding configuration.

Call flows

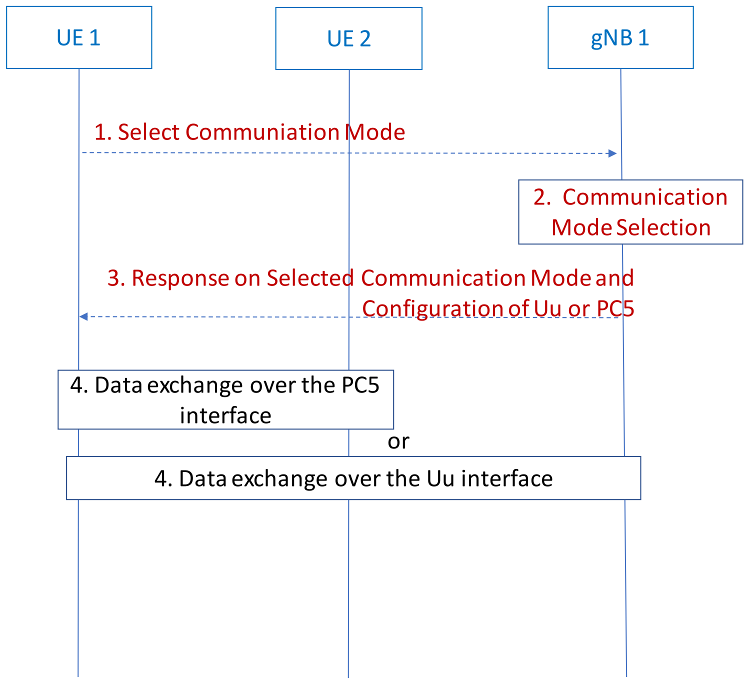


Figure 5.21: Dynamic selection of PC5 and Uu communication modes

5.4.13 Security and privacy enablers

TC description

Table 5.14: Security and privacy enablers description

Network Entity/ Network Function	**New Feature and/or Extension of existing Functionality**
UE	User Plane: • Signed an initial request to access a V2X Application Server with its C-ITS certificate. • Verifies the signed (+encrypted) message received from V2X Application Server. • Encrypt all further exchanges using the session key provided by the V2X AS.

V2X AS	Verify permissions of any UE requesting access to that application (through signed message) Invokes KMF function to get per-session symmetric encryption key Provides the session key to all involved UEs using signed + encrypted message Note: the exact sequence of messages depends on the V2X application (communication between two UEs only, between a group of UEs, etc.)
KMF	Allocate new session symmetric key

Call flows

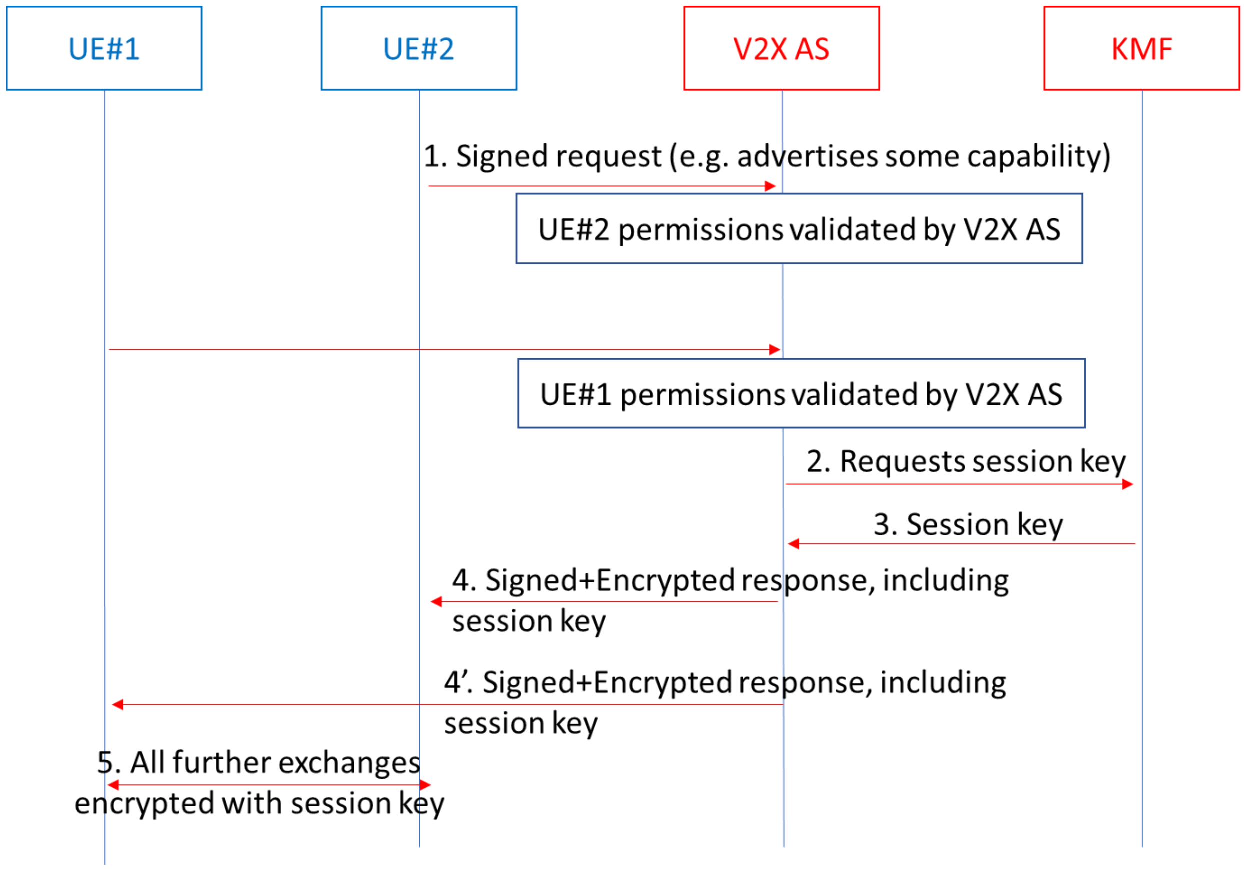


Figure 5.22: Communication between two UEs only (e.g. see-through use case)

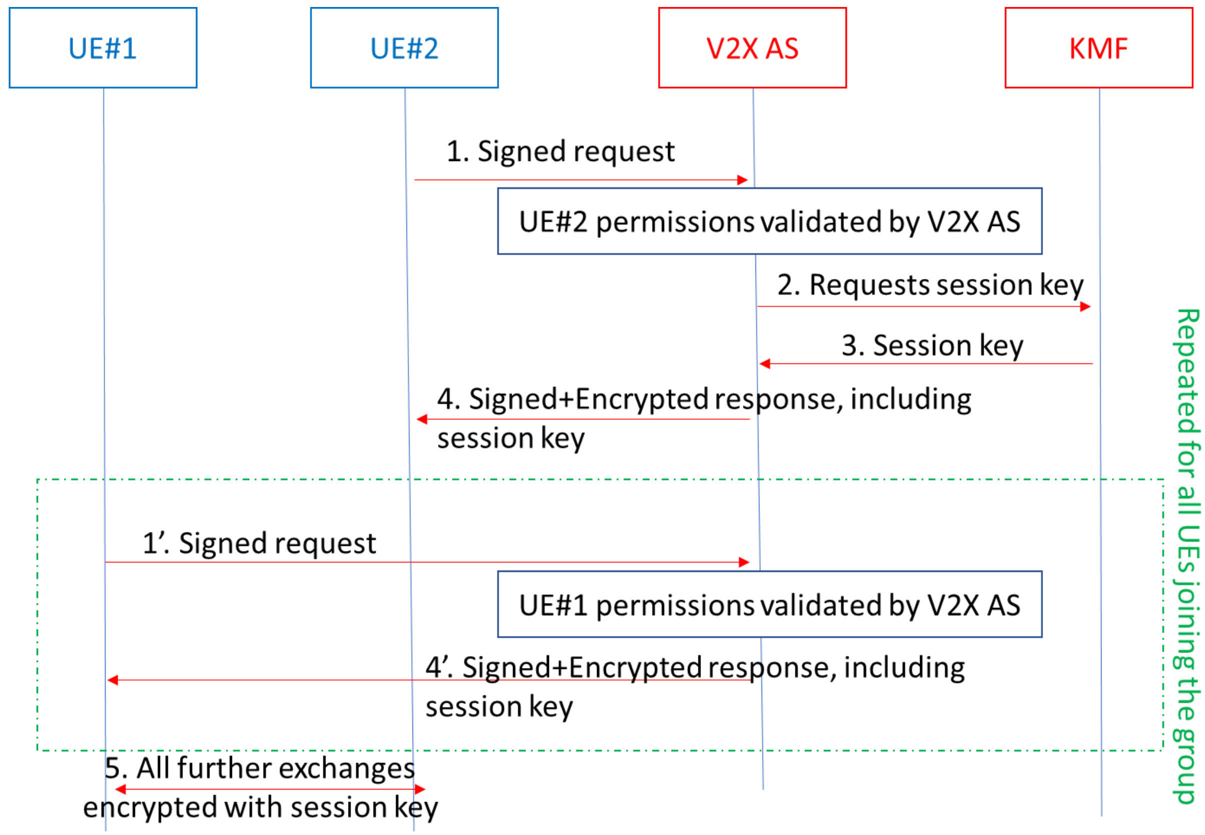


Figure 5.23: Two or more UEs joining a group

5.4.14 5G core network evolution for edge computing-based mobility

TC description

Table 5.15: 5G core network evolution for edge computing-based mobility description

Network Entity/ Network Function	**New Feature and/or Extension of existing Functionality**
UE	Periodically transmits its status information (NAS: including position, heading, speed, acceleration, etc.) Handover controller inside of the UE for following actions : Checking the channel states (e.g., RSSs) from different RANs

	Checking the status of offloaded jobs Send those information waiting for the Control Plane to make a decision on job migration and connection redirection
gNB	Forwards UE periodic status information to AMF.
NWDA	Performs algorithm to URLLC slice for predictive HO purpose
AMF	Based on NWDA algorithm AMF selects the target SMF Forwards jobs to SMFs for resources reservation
SMF	Interface between two SMFs to be defined
UPF	Interface between two UPFs to be defined

Call flows

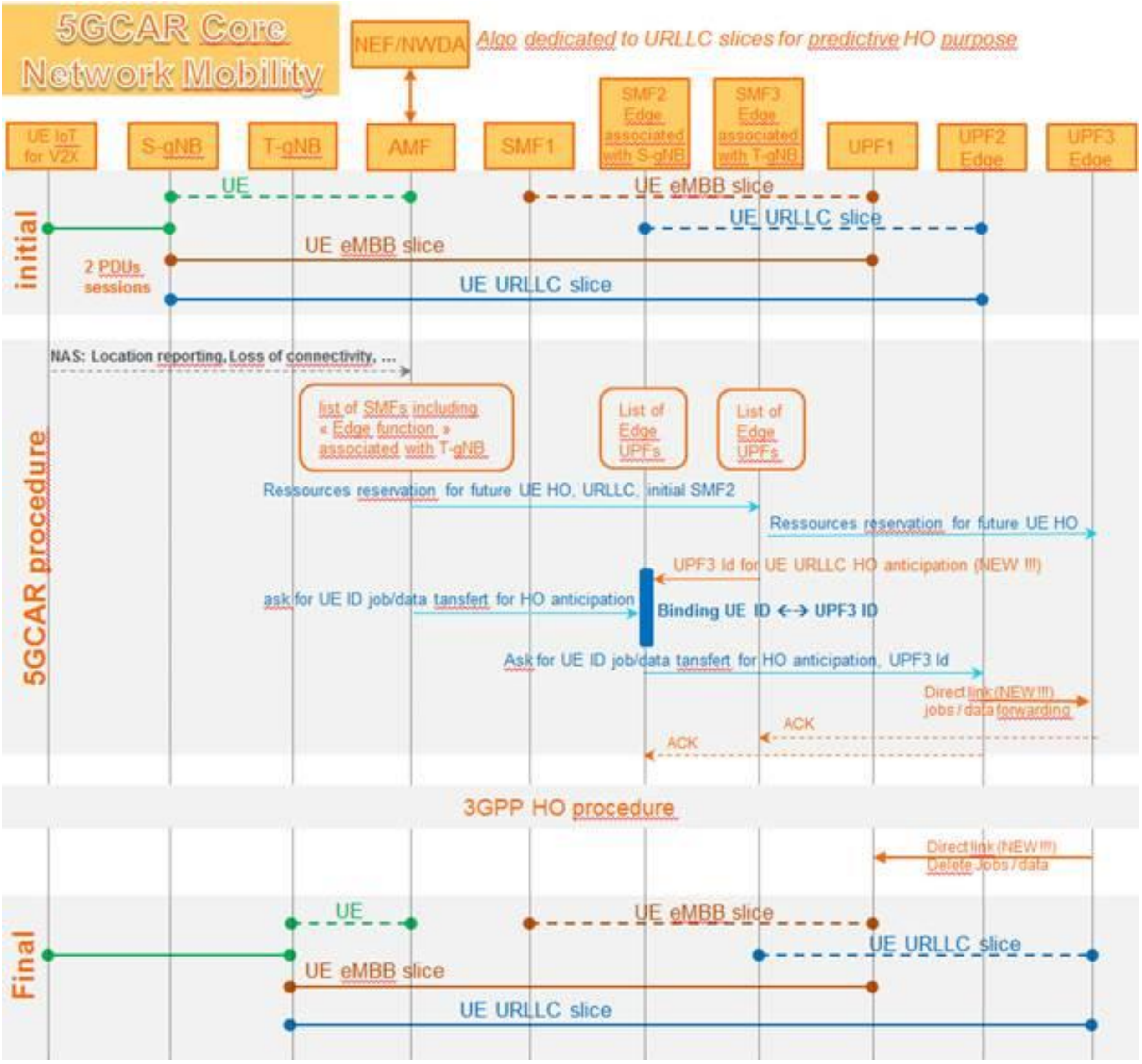


Figure 5.24: 5G core network evolution for edge computing-based mobility

6 Conclusions

This document is the first deliverable designing the 5GCAR architecture, of which we provided a preliminary description by using the existing 5G core network architecture defined in 3GPP as the baseline. In this deliverable we propose several innovations and enhancements to the 5G architecture, which could be grouped into two sets: V2X feature independent, and V2X feature specific.

V2X feature independent adaptations and enhancements are related to overall system design and they are not limited to V2X related scenarios/use cases. Overall, the service based architecture defined in 3GPP 5GC entails the network functions which are interconnected to each other on a service-basis; each function is provided with an interface through which it can offer services to other functions, or subscribe to services offered by other functions. Service based architecture facilitates its evolution, in order to cover multitude of new services that arose recently and that will be defined in the coming years.

Network slicing can be used to secure the QoS of various V2X services where three slices types (i.e., eMBB, OEM, and Safety) could be used to offer support for V2X services. Edge computing is essential to support latency services, but aspects related to access to data, multi operator support, etc. have to be dealt. Finally IaaS could be used to automate the software deployment, or other use cases

From a security and privacy perspective, this deliverable discusses V2X specificities in order to ensure security and privacy across all layers, and between all involved entities (OBU/RSU/VRU and V2X application servers). It builds on the European Commission certificate and security policies for C-ITS deployment, while proposing investigating several options to rely on 5G Network-based functions, 5G security architecture functions such as key management, and to take advantage of 5G capabilities such as Mobile Edge Computing.

Regarding the V2X feature specific architectural analyses, several proposals are made dealing with local traffic, reliability, multi operator support, scheduling and traffic optimization, service negotiation, etc. These preliminary analyses include the required architectural enhancements in terms of functionalities and signalling. The study of the technical components lead to the preliminary conclusions, that the function-based architecture defined in 3GPP 5GC will be used as a baseline.

Enhancements are however needed for the majority of the existing network functions, including V2X application, UE, gNB, AMF, SMF, UPF, UDM, AF, PCF, NEF, AF, NWDAF, and their corresponding interfaces.

This conclusion is based on the preliminary studies conducted in the first part of the project, which is the one covered by this deliverable. The solutions herein presented will furthermore need to be evaluated and validated by means of simulation, which will be done in the next phase of the project. All TCs are still under ongoing development, with only main message flow specified in high level. WP4 will focus on specifying all message flows of all TCs, with necessary parameter information specified.

Further studies will also be conducted in WP4, both concerning the finalization of the V2X feature independent part, and the V2X feature specific part, with this latter being the focus of next phase. Until now, in fact, each TC was studied individually and independently. Some further analysis on system level will be performed, focusing on the overlap, dependency and interactions among TCs. The evaluation on each TC will be performed considering performance, feasibility and other aspects.

Furthermore, additional TCs might be specified if the ongoing development will highlight new challenges currently not identified or not fully covered by the TCs presented in this document.

7 References

[3GPP11-23829] 3GPP Technical Report, “TR 23.829 Technical Specification Group Services and System Aspects; Local IP Access and Selected IP Traffic Offload (LIPA-SIPTO) (Release 10), v10.0.1”, October 2011.

[3GPP13-23859] 3GPP TR 23.859 V12.0.1: Local IP access (LIPA) mobility and Selected IP Traffic Offload (SIPTO) at the local network (Rel. 12), April 2013.

[3GPP14-36842] 3GPP TR 36.842, “Study on Small Cell Enhancements for E-UTRA and E-UTRAN; Higher Layer Aspects”, Sep. 2014.

[3GPP14-23707] 3GPP technical report, “TR 23.707 Technical Specification Group Services and System Aspects; Architecture Enhancements for Dedicated Core Networks; Stage 2 (Release 13) v13.0.0”, December 2014

[3GPP16-23122] 3GPP TS 23.122 V13.5.0: Non-Access-Stratum (NAS) functions related to Mobile Station (MS) in idle mode (Release 13), June 2016.

[3GPP16-23401] 3GPP TS 23.401 V14.0.0: General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access; (Release 4), June 2016.

[3GPP16-23711] 3GPP technical report, “TR 23.711 Technical Specification Group Services and System Aspects; Enhancements of Dedicated Core Network selection mechanisms; (Release 14) v14.0.0”, September 2016

[3GPP16-23785] 3GPP TR23.785, Study on architecture enhancements for LTE support of V2X services, 14.0.0, September 2016.

[3GPP16-36885] 3GPP TR 36.885 V2.0.0: Study on LTE-based V2X Services; (Release 14), May 2016.

[3GPP17-22185] 3GPP technical specification, “ TS 22.185 Technical Specification Group Services and System Aspects; Service requirements for V2X services; Stage 1 (Release 14) v14.3.0”, March 2017

[3GPP17-22186] 3GPP technical specification, “TS 22.186 Technical Specification Group Services and System Aspects; Enhancement of 3GPP support for V2X scenarios; Stage 1 (Release 15) v15.2.0”, September 2017

[3GPP17-23285] 3GPP TS23.285, Architecture enhancements for V2X services, 14.5.0, December 2017.

[3GPP17-23214] 3GPP technical specification, “TS 23.214 Technical Specification Group Services and System Aspects; Architecture enhancements for control and user plane separation of EPC nodes; Stage 2 (Release 15) v15.1.0”, December 2017

[3GPP17-23246] 3GPP, Multimedia Broadcast/Multicast Service (MBMS); Architecture and functional description (Release 15), v15.0.0, Dec. 2017.

[3GPP17-23285] 3GPP technical specification, "TS 23.285 Technical Specification Group Services and System Aspects; Architecture enhancements for V2X services (Release 14) v14.5.0", December 2017

[3GPP17-23501] 3GPP technical specification, "TS 23.501 Technical Specification Group Services and System Aspects; System Architecture for the 5G System; Stage 2 (Release 15) v15.0.0", December 2017

[3GPP17-23793] 3GPP TR 23.793, "Study on Access Traffic Steering, Switch and Splitting support in the 5G system architecture (Release 15)", Sep. 2017.

[3GPP17-24301] 3GPP TS 24.301 V14.3.0: Non-Access-Stratum (NAS) protocol for Evolved Packet System (EPS); (Release 14), March 2017.

[3GPP17-26247] 3GPP, "Dynamic Adaptive Streaming over HTTP (3GP-DASH) (Release 15)," TS 26.274, December 2017.

[3GPP17-28801] 3GPP technical report, "TR 28.801 Technical Specification Group Services and System Aspects; Telecommunication management; Study on management and orchestration of network slicing for next generation network (Release 15)", v 15.0.0, September 2017

[3GPP17-36300] 3GPP TS36.300, Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2; 14.5.0 December 2017.

[3GPP17-36331] 3GPP TS 36.331 V14.2.1: Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Resource Control (RRC); Protocol specification (Release 14), September 2017.

[3GPP17-RP1740] 3GPP TSG RAN RP-171740, "V2X phase 2 based on LTE," September 2017.

[3GPP18-23786] 3GPP TR 23.786 V0.4.0: Study on architecture enhancements for EPS and 5G System to support advanced V2X services (Release 16), March 2018.

[3GPP18-29520] 3GPP Technical Specification "TS 29.520: Technical Specification Group Core Network and Terminals; 5G System; Network Data Analytics Services; Stage 3 (Release 15) v0.4.0", March 2018

[3GPP18-36101] 3GPP Technical Specification "TS 36.101: Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) radio transmission and reception (Release 15) v15.1.0", January 2018

[3GPP18-36213] 3GPP Technical Specification, "TS 36.213 "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures (Release 14) v14.5.0", January 2018

[3GPP18-38300] 3GPP, TS 38.300 V15.0.0: NR and NG-RAN Overall Description, Stage 2 (Release 15), January 2018.

[5GAA17-WP] 5GAA White Paper: "Toward fully connected vehicles: Edge computing for advanced automotive communications," Dec. 2017.

[5GAM2017] 5G Americas whitepaper, "5G services and Use Cases", November 2017

[5GC17-D21] 5GCAR Deliverable 2.1, "5GCAR Scenarios, Use Cases, Requirements and KPIs", August 2017

[5GC18-D22] 5GCAR Deliverable 2.2, "Intermediate Report on V2X Business Models and Spectrum", March 2018

[5GC18-D31] 5GCAR Deliverable 3.1, "Intermediate 5G V2X Radio", May 2018

[5GN15-D31] 5G NORMA deliverable 3.1, "Functional Network Architecture and Security Requirements", December 2015

[5GN17-D32] 5G NORMA Deliverable D3.2 –"5G NORMA network architecture" - Intermediate report, January 2017

[5GN17-D33] 5G NORMA deliverable 3.3, "5G NORMA network architecture – Final report", October 2017

[5GPPP17-WP] 5GPPP white paper, "View on 5G Architecture (Version 2.0)", July 2017

[ALH12] H. Alvestrand, H. Lundin, and S. Holmer, "A Google Congestion Control Algorithm for Real-Time Communication on the World Wide Web," Tech. Rep. 3, IETF, 2012.

[BBF16-348] Broadband Forum, "Hybrid Access Broadband Network Architecture" TR-348, July 2016.

[CAMP17] CAMP llc, Security Credential Management System CV Pilots Documentation, May 2017. Available online: https://wiki.campllc.org/display/SCP/Environments+documentation

[CDK+14] Wai Chen, Luca Delgrossi, Timo Kosch, Tadao Saito, "Automotive Networking and Applications", IEEE Communications Magazine, Vol. 52, no. 6, June 2014.

[CHA16-D21] Charisma deliverable 2.1, "CHARISMA Initial Architecture Design and Interfaces", July 2016

[COH16-D22] Coherent deliverable 2.2, "System Architecture and Abstractions for Mobile Networks", July 2016

[DHA+14] Falko Dressler, Hannes Hartenstein, Onur Altintas, Ozan K. Tonguz, "Inter-Vehicle Communication: Quo Vadis", IEEE Communications Magazine, Vol. 52, no. 6, June 2014.

[DMS16] I. L. Da Silva, G. Mildh, M. Säily and S. Hailu, "A novel state model for 5G Radio Access Networks", IEEE International Conference on Communications Workshops (ICC), Kuala Lumpur, pp. 632-637, 2016.

[DSL+17] B. Di, L. Song, Y. Li and G. Y. Li, "Non-Orthogonal Multiple Access for High-Reliable and Low-Latency V2X Communications in 5G Systems," in IEEE Journal on Selected Areas in Communications, vol. 35, no. 10, pp. 2383-2397, Oct. 2017. doi: 10.1109/JSAC.2017.2726018

[ERF11] European Union Road Federation, "European Road Statistics", 10th edition, 2011.

[ETSI09-102638] ETSI TR 102 638, "Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Definitions", V1.1.1, 2009.

[ETSI10-302665] ETSI EN 302 665: "Intelligent Transport Systems (ITS); Communications Architecture v1.1.1", September 2010

[ETSI11-1028892] ETSI TR 102 889-2, "Electromagnetic compatibility and Radio Spectrum Matters (ERM); System Reference Document; Short Range Devices (SRD); Part 2: Technical characteristics for SRD equipment for wireless industrial applications using technologies different from Ultra-Wide Band (UWB)", V1.1.1, 2011.

[ETSI12-102941] ETSI, "Intelligent Transport Systems (ITS); Security; Trust and Privacy Management. Draft", June 2012

[ETSI13-1028941] ETSI TS 102 894-1 (2013-08): "Intelligent Transport Systems (ITS); Users and applications requirements; Part 1: Facility layer structure, functional requirements and specifications v1.1.1", August 2013

[ETSI14-3026372] ETSI EN 302 637-2 (V1.3.2 - 2014-11) Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Part 2: Specification of Cooperative Awareness Basic Service

[ETSI14-3026373] ETSI EN 302 637-3 (V1.2.2 - 2014-11) Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Part 3: Specifications of Decentralized Environmental Notification Basic Service

[ETSI14-MAN001] European Telecommunications Standard Institute, "GS NFV-MAN 001, Network Functions Virtualisation (NFV); Management and Orchestration", December 2014

[ETSI16-102940] ETSI, "Intelligent Transport Systems (ITS); Security; ITS communications security architecture and security management. V1.2.1", November 2016

[ETSI17-103097] ETSI, "Intelligent Transport Systems (ITS); Security; Security header and certificate formats. V1.3.1", October 2017

[EU17-SECFMWK] European Commission. Security Policy & Governance Framework for Deployment and Operation of European Cooperative Intelligent Transport Systems (C-ITS) Release 1 December 2017

[GKA14] Gunes T.T., Kwong U S.T., Afifi H. (2014) Hybrid Model for LTE Network-Assisted D2D Communications. In: Guo S., Lloret J., Manzoni P., Ruehrup S. (eds) Ad-hoc, Mobile, and Wireless Networks. ADHOC-NOW 2014. Lecture Notes in Computer Science, vol 8487. Springer, Cham, https://arxiv.org/ftp/arxiv/papers/1502/1502.00143.pdf.

[GSMA17] GSM Association, "An introduction to network slicing", 2017

[HL08] Hartenstein, H., and Laberteaux, K.P., "A Tutorial Survey on Vehicular Ad Hoc Networks" IEEE Communications Magazine, Vol. 46, no. 6, June 2008.

[HPY+14] Harding, J., Powell, G., R., Yoon, R., Fikentscher, J., Doyle, C., Sade, D., Lukuc, M., Simons, J., & Wang, J., "Vehicle-to-vehicle communications: Readiness of V2V technology for application." Report No. DOT HS 812 014, Washington, DC: National Highway Traffic Safety Administration, August 2014, http://www.nhtsa.gov/staticfiles/rulemaking/pdf/V2V/Readiness-of-V2V-Technology-for-Application-812014.pdf.

[ICT17] ICTEAM, *"MultiPath TCP - Linux Kernel implementation"*, August 2017. https://multipath-tcp.org/pmwiki.php/Users/ConfigureMPTCP

[IEEE16-16092] IEEE Standard for Wireless Access in Vehicular Environments—Security Services for Applications and Management Messages. IEEE Std 1609.2™-2016

[IETF18-RFC6824] IETF RFC 6824bis-10 *"TCP Extensions for Multipath Operation with Multiple Addresses"*, March 2018

[ITU15] ITU-R Recommendation M.2083, "IMT Vision – Framework and overall objectives of the future deployment of IMT for 2020 and beyond", September 2015

[ITU2012] ITU-T Recommendation Y.3011, "Framework of network virtualization for future networks", January 2012

[JOH14] I. Johansson, "Self-clocked rate adaptation for conversational video in LTE", ACM SIGCOMM Capacity Sharing Workshop, 2014

[KKC16] J. Kim, S. C. Kwon, and G. Choi, "Performance of video streaming in infrastructure-to-vehicle telematics platforms with 60-GHz radiation and IEEE 802.11ad baseband," IEEE Trans. Veh. Technol., vol. 65, no.12, pp. 101111-10115, Dec. 2016.

[KLV10] Soeng-Hun Kim, Gert-Jan van Lieshout, and Himke van der Velde, Patent N° 20170150490: *"Method and apparatus for switching communication interface in a wireless communication system."*, 2010

[KMA17] A. Kousaridas, D. Medina, S. Ayaz and C. Zhou, "Recent advances in 3GPP networks for vehicular communications," 2017 IEEE Conference on Standards for Communications and Networking (CSCN), pp. 91-97, 2017.

[MET217-D52] METIS-II deliverable 5.2, "Final Considerations on Synchronous Control Functions and Agile Resource Management for 5G", March 2017

[MYZ+17] Y. Mao, C. You, J. Zhang, K. Huang, and K. B. Letaief, "A survey on mobile edge computing: The communication perspective," IEEE Communications Surveys & Tutorials, vol. 19, no. 4, pp. 2322-2358, 2017.

[NFV-Slice17] ETSI NFV ISG, "NFV-EVE 012 report on network slicing support with ETSI NFV architecture framework," 2017.

[NGMN15] NGMN Alliance, "5G White Paper", February 2015

[NGMN16] NGMN Alliance white paper, "Perspectives on vertical industries and implications for 5G", June 2016

[NGMN17] NGMN Alliance, "NGMN Alliance Overview", August 2017

[ODC-BP] OPEN DATA CENTER ALLIANCE Best Practices: Architecting Cloud-Aware Applications Rev. 1.0

[ONF15-TR511] ONF, Principles and Practices for Securing Software-Defined Networks, January 2015, ONF TR-511

[OST-TRI] Openstack Tricircle, https://wiki.openstack.org/wiki/Tricircle

[PLY+17] H. Peng, D. Li, Q. Ye, K. Abboud, H. Zhao, W. Zhuang, and X. Shen, "Resource Allocation for Cellular-based Inter-Vehicle Communications in Autonomous Multiplatoons," in IEEE Transactions on Vehicular Technology, vol. 66, no. 12, pp. 11249-11263, Dec. 2017. doi: 10.1109/TVT.2017.2723430

[R2-1800259] 3GPP TSG-RAN WG2 Meeting AH 1801 - R2-1800259 – "What is RAN part of a network slice?" - Huawei, HiSilicon

[RSM+13] T.S. Rappaport et al., "Millimeter wave mobile communications for 5G cellular: It will work!" IEEE Access, vol.1, pp. 335-349, May 2013.

[RW13] D. Ros and M. Welzl, "Assessing LEDBAT's Delay Impact," in IEEE Communications Letters, vol. 17, no. 5, pp. 1044-1047, May 2013. doi: 10.1109/LCOMM.2013.040213.130137

[SGH17] F. Slim, F. Guillemin, A. Gravey and Y. Hadjadj-Aoul, "Towards a dynamic adaptive placement of virtual network functions under ONAP," *2017 IEEE Conference on Network Function Virtualization and Software Defined Networks (NFV-SDN)*, Berlin, 2017

[SGH18] F. Slim, F. Guillemin and Y. Hadjadj-Aoul, "CLOSE: A costless service offloading strategy for distributed edge cloud," *2018 15th IEEE Annual Consumer Communications & Networking Conference (CCNC)*, Las Vegas, NV, USA, 2018

[SH06] James P.G Sterbenz, David Hutchinson, "Resilinets: Multilevel Resilient and Survivable Networking Initiative", August 2006. Available online: https://www.ittc.ku.edu/resilinets/

[SNK+16] M. Sami, N. K. Noordin, M. Khabazian, F. Hashim and S. Subramaniam, "A Survey and Taxonomy on Medium Access Control Strategies for Cooperative Communication in Wireless Networks: Research Issues and Challenges," in IEEE Communications Surveys & Tutorials, vol. 18, no. 4, pp. 2493-2521, Fourthquarter 2016. doi: 10.1109/COMST.2016.2569601

[SSP+16] I. Safiulin, S. Schwarz, T. Philosof and M. Rupp, "Latency and Resource Utilization Analysis for V2X Communication over LTE MBSFN Transmission," WSA 2016; 20th International ITG Workshop on Smart Antennas, Munich, Germany, 2016, pp. 1-6.

[TGV15] R. Trivisonno, R. Guerzoni, I. Vaishnavi, D. Soldani, "Towards Zero Latency Software Defined 5G Networks", IEEE International Conference on Communication Workshop, pp. 2566-2571, 2015

[TPH+14] Ylva Timner, Jonas Pettersson, Hans Hannu, Min Wang, and Ingemar Johansson. 2014. Network assisted rate adaptation for conversational video over LTE, concept and performance evaluation. In Proceedings of the 2014 ACM SIGCOMM workshop on Capacity sharing workshop (CSWS '14). ACM, New York, NY, USA, 45-50. DOI: https://doi.org/10.1145/2630088.2630096

[USDOT16-SCMS] Security Credential Management System Proof–of–Concept Implementation. EE Requirements and Specifications Supporting SCMS Software Release 1.2.2, November 2016

[VSB+16] V. Va et al., "Beam design for beam switching based millimeter wave vehicle-to-infrastructure communications," in IEEE ICC, 2016, PP. 1-6.

[XZF+17] K. Xiong, Y. Zhang, P. Fan, H. C. Yang and X. Zhou, "Mobile Service Amount Based Link Scheduling for High-Mobility Cooperative Vehicular

Networks," in IEEE Transactions on Vehicular Technology, vol. 66, no. 10, pp. 9521-9533, Oct. 2017. doi: 10.1109/TVT.2017.2714863

[YJS+17] Ali Yaver, Maciej Januszewski, Moya Fernando Sanchez, Patrick Marsch, Matthias Hesse (2017) WO 2017088906 A1 Device to device communications, https://www.google.com/patents/WO2017088906A1?cl=en.

[ZCY+15] R. Zhang, X. Cheng, L. Yang, X. Shen and B. Jiao, "A Novel Centralized TDMA-Based Scheduling Protocol for Vehicular Networks," in IEEE Transactions on Intelligent Transportation Systems, vol. 16, no. 1, pp. 411-416, Feb. 2015. doi: 10.1109/TITS.2014.2335746

A. Annex 1: State of the art on 5G and network slicing

A.1 A brief introduction to 5G

Standardisation bodies, the industry, and the academia alike are currently at work for the the specification of 5G, the fifth generation of mobile networking, which will power wireless mobile communications for the upcoming decade. In Figure A.1, inspired by Figure 1 in [5GAM-2017], 5G is compared to the previous generations in terms of breakthrough new features, supported applications, and maximum achievable data rates.

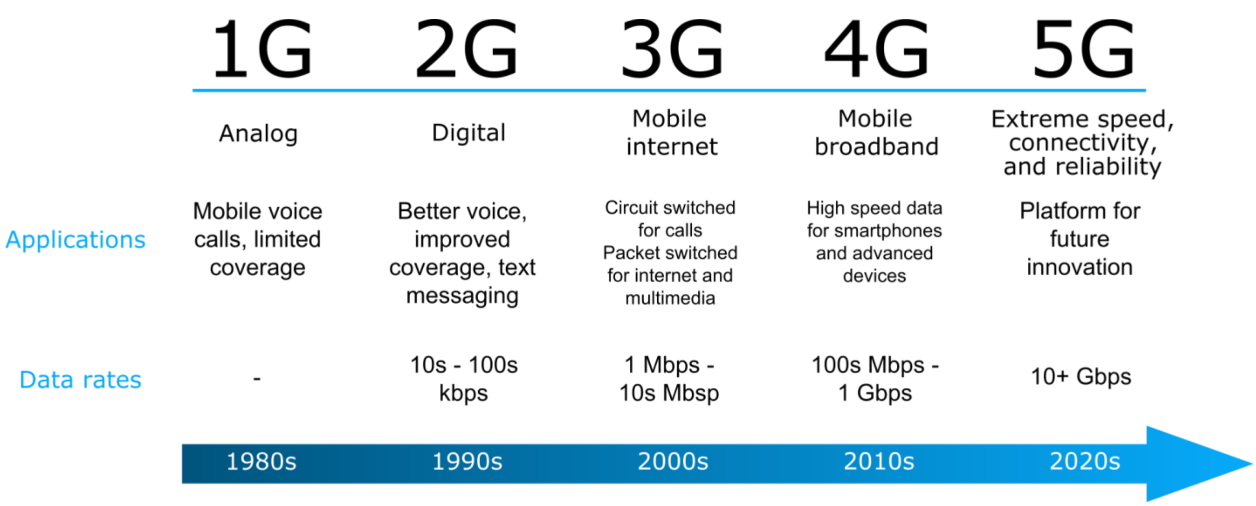


Figure A.1: evolution and key developments in mobile networking. Inspired by Fig. 1 in [5GAM-2017]

Beside introducing improvements over 4G in terms of coverage, spectrum availability, and maximum data rates, 5G will be differentiated from its predecessor by the native support of a multitude of use cases and connectivity scenarios which will be of paramount importance over the next few years. These scenarios were derived from the necessities of "verticals", which are ensembles of industries with common sets of requirements. Such verticals include [NGMN16]:

- the automotive industry
- transport and logistics

- health and wellness
- smart city and utilities
- agriculture

There is wide consensus around the three main classes of services that have been identified by the International Telecommunication Union (see for instance [ITU15]):

1. eMBB – enhanced Mobile Broadband, is the natural evolution of the connectivity supported today by the 4G network, offering increased data rate and improved Quality of Service for mobile portable devices;
2. mMTC – massive Machine-Type Communications is a paradigm that include Internet of Things devices. These are mostly static, low power devices that need to sporadically send small amount of data to the network. The objective is to support the connectivity of very large densities of devices, while guaranteeing very long battery lives;
3. URLLC – ultra Reliable Low Latency Communications are conceived to serve those applications that require almost-instantaneous channel access, ultra-low round trip times, and very high delivery reliability in order to serve safety-of-life and industrial control applications.

The ITU itself acknowledged in [ITU2012] that it is “difficult to satisfy the requirements of every service on a single network architecture”, while at the same time, it is “unrealistic to realize heterogeneous network architectures using multiple physical networks”.

The solution is then for a unique physical infrastructure to be able to run multiple independent networks, each tailored to support a specific set of use cases with comparable business value propositions and service requirements. This paradigm, commonly referred to as network slicing, will be described in the next section.

A.2 Network slicing

The evocative name “network slicing” describes the ability for a mobile service provider to partition their physical network infrastructure into slices, which are independent logical network dedicated to support a limited set of communication needs.

Network slicing is widely considered as one of the critical enablers for the development of the next generation of wireless communications, as recognized by the most important standard organisations and research partnerships. Each of them include the paradigm in their vision for 5G, and provide their own definition for it.

A.2.1 Definitions

NGMN alliance

The Next Generation Mobile Network alliance is a global partnership between operators, telco and IT vendors, research institutes established to ensure that functionality and performance of next generation mobile infrastructure, service platforms and devices will meet the requirements of operators, and ultimately, will satisfy end user demand and expectations [NGMN17].

In its 5G vision white paper [NGMN15], NGMN states that "a network slice […] supports the communication service of a particular connection type with a specifying way of handling the [control plane] and [user plane] for this service. To this end, a [network] slice is composed of a collection of 5G network functions and specific [radio access technology] settings that are combined together for the specific use case or business model. Thus, a [network] slice can span all domains of the network […]."

ITU

The International Telecommunication Union is the United Nations specialized agency for information and communication technologies.

The ITU first defined the concept of slicing in [ITU2012] as LINPs, Logically Isolated Network Partitions, i.e. virtual networks that coexist on a single physical infrastructure. Each LINP can provide the corresponding users with services similar to those provided by traditional networks. The users of LINPs are not limited to the users of services and applications, but can include service providers. LINPs implementation rely on the virtualisation of network functions, and provides the following characteristics: partitioning, abstraction, isolation, flexibility, programmability, and authentication, authorisation, and accounting, this latter meaning that usage of resources allocated to a LINP must be safe and secured.

GSMA

The GSM Association regroups and represents the interests of nearly 800 operators and 300 companies in the broader mobile ecosystem.

In [GSMA17], the GSMA defines network slicing from the point of view of business customers, and provides examples of characteristics and features offered by network slices. From an operator's point of view, a slice is defined as "independent end-to-end logical network that runs on a shared physical infrastructure, capable of providing a negotiated service quality".

5GPPP

The 5G Infrastructure Public Private Partnership is a joint initiative between the European Commission and European ICT industry, which aims to deliver solutions, architectures, technologies and standards for the ubiquitous next generation communication infrastructures of the coming decade.

The network slices is defined in [5GPPP17-WP] as a "composition of adequately configured network functions, network applications, and the underlying cloud infrastructure (physical, virtual, or even emulated resources, [radio access network] resources), that are bundled together to meet the requirements of a specific use case, e.g., bandwidth, latency, processing, and resiliency, coupled with a business purpose". It is further put into evidence that "network slicing is an end-to-end covering all the network segments including radio networks, wire access, core, transport, and edge networks".

3GPP

The 3rd Generation Partnership Project unites several telecommunication SDOs (Standard Development Organisations) known as "Organisational Partners". It covers cellular telecommunications network technologies, including radio access, the core transport network, and service capabilities, and thus provides complete system specifications.

In order to define the technical specifications that will be implemented into network equipment and user terminals, the 3GPP takes into account the results of studies produced by the aforementioned organisations. In [3GPP17-28801], a network slice instance is defined as "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 characteristics".

ETSI NFV

ETSI NFV ISG is actually highlighting the relationship between network services and Slices / SubnetSlices [NFV-Slice17]. This is important since the a lot of work has been performed in NFV and can be reused for Network Slicing.

It can be observed that the virtualised resources for the network slice and their connectivity to physical resources can be represented by the nested network service concept, or one or more VNFs and PNFs directly attached to the network service used by the network slice. ETSI states that "an NFV Network Service (network service) can thus be regarded as a resource-centric view of a network slice, for the cases where a Network Slice Instance (NSI) would contain at least one virtualised network function" [NFV-Slice17].

A.2.2 State of the art: 5GPPP Phase 1

Numerous 5GPPP Phase 1 projects have produced valuable results and propositions concerning the architecture for 5G, and Network Slicing, which will be briefly presented in this section.

5G NORMA (www.it.uc3m.es/wnl/5gnorma)

Focus: End to End.

5G-NORMA is the acronym for 5G Novel Radio Multiservice adaptative Network Architecture. The focus of the project is the development of network architectures capable of adapting to fluctuations in traffic demand resulting from heterogeneous and dynamically changing service portfolios.

Network slicing is of paramount importance in 5G NORMA, since it is instrumental to provide the flexibility required to provide adequate support to services with heterogeneous requirements and to enable multi-tenancy. Multi-tenancy is the principle according to which a slice could be hosted on the infrastructure owned by a Mobile Service Provider, but administered by a third party tenant, which is a general expression that could define a vertical, a specific company providing connectivity services, or a service itself.

5G NORMA focuses on 3 innovative enablers and 2 functionalities, in order to fulfil the architecture requirements [5GN15-D31]. The enablers are:

1. **software defined mobile network control and orchestration**: the softwarisation of network functions is widely considered as one of the critical enablers for the deployment of next generation networks. 5G NORMA thus defined a set of control and orchestration elements that enable the application of the SDN principle and of the ETSI NFV MANO paradigm to sliced 5G network. Specifically, two control entities are defined, the Software Defined Mobile Network Controller and Coordinator (SDM-C and SDM-X, respectively), wherein the former is in charge of interfacing with the network functions specific to each slice, whereas the latter is in control of the network functions that are shared among all of the network slices. This might be the case for instance, based on the deployment configuration, of RAN functions. In order to optimize the multiplexing gain, the radio interface may be shared dynamically by all of the slices, hence requiring functions common to all of them. To this end, the 5G NORMA supports the coexistence of both virtual network functions (VNFs) and physical network functions (PNFs), these latter being more suited for supporting the strict real-time requirements of the radio interface.
   A third element, known as the SDN Mobile Orchestrator (SDM-O), is in charge of administrating the network resources among the slices by having a global view over them. The SDM-O is part of a newly introduced Management and orchestration layer, which include the elements defined by the ESTI MANO framework, which the project has extended to support a sliced network architecture.
2. **adaptive composition and allocation of mobile network functions**: the modularisation of network entities is one of the key components of 5G NORMA, and of the 5G architecture alike. According to this paradigm, large network functions defined in 4G networks are split into smaller, single function modules. This approach is beneficial for multiple reasons: from an implementation perspective, it facilitates the development and maintenance of these functions. Furthermore, each module can implement only the subset of functionalities optimized to support a specific type of service. These functions can then be allocated dynamically either on the edge of the network, closer to the terminals, or in the central cloud, according to the use cases they are designed to serve.

   This, coupled with the deployment of network slices, enables the implementation of optimized and maintainable dedicated virtual networks.
3. **joint optimisation of mobile access and core**: the two previously described enablers permit a global optimisation of functions located both in the access network and in the

core network. The flexibility provided by adaptive (de)composition of NFs and their control via SDN enables end-to-end optimisation, which is the base for the support for new classes of use cases that will be defined in the long term.

The 2 functionalities identified by 5G NORMA are:

1. **mobile network multi-tenancy**: in 5G NORMA, the support of network slices is envisioned in order to enable third parties, services, or verticals to all offer dedicated services to their clients by deploying logically isolated slices sharing the same physical network. This approach is known as multi-tenancy.
2. **multi-service and context-aware adaptation of network functions**: in order to attain 5G NORMA objective in terms of flexibility and adaptability, a multitude of functional requirements need to be fulfilled. These include service-specific detection of traffic, service-specific derivation of service requirements, adaptation of NFs to provide service differentiation, context-aware NF allocation, and dynamic network monitoring.

CHARISMA (www.charisma5G.eu)
Focus: Radio Access Network.
Charisma stands for Converged Advanced 5G Cloud-RAN for Intelligent and Secure Media Access. The objective of the project is to provide ground-breaking low-latency services, high bandwidth and mobile cloud resilient security. This is achieved by proposing an RAN architecture based on para-virtualisation which performs intelligent hierarchical routing, with the purpose of offloading traffic with the shortest path nearest to the end user.

Every network location (from right to left: home gateway, cell site, local cloud, central cloud) is equipped with computing, storage, and connectivity resources which enables it to perform caching, and, most relevantly for the 5GCAR use cases, to perform local routing at the lowest possible level of the hierarchy when messages are directed towards terminals in proximity.

COHERENT (www.ict-coherent.eu)
Focus: Radio Access Network

Coherent stands for COordinated control and spectrum management for 5G HEterogeneous Radio access NeTwork. The architecture definition for the radio access network proposed by the Coherent project revolves around three key principles: control plane separation (which derives from the application of the SDN paradigm), the abstraction of network entities, and network slicing [COH16-D22] .

The application of the SDN paradigm in Coherent aims at centralizing the network control in order to achieve global optimisation. This task is however challenging when the RAN is concerned, due to the delay between the control entity and the radio entities: the Coherent proposal is thus to separate the control into two architectural sections, the central functionalities and real-time control. This latter part is responsible for handling rapidly varying wireless components. In this way, the benefits of centralized control can be maintained, while respecting hard real time constraints of the radio interface.

The abstraction of the radio access network components enables a coherent representation of the network state and infrastructure resources that facilitates the control and coordination of the network.

Finally, network slicing is of paramount importance for both speeding up the deployment and maintenance of new features, and for improving flexibility and scalability of mobile networks. In Coherent, the full-stack of network slices is considered, from the virtualized network resources up to the applications running on top of them.

METIS II (https://metis-ii.5g-ppp.eu)

Focus: Radio Access Network

METIS II is the acronym for Mobile and wireless communications Enablers for Twenty-twenty (2020) Information Society-II. The project aims at designing the radio interface for 5G, a system that will extend communications from humans to a plethora of machine type communications, including evolved tailored services for humans, such as transportation systems, eHealth, and industrial control to name a few.

METIS II integrates network slicing into its resource management framework focusing on slice support in presence of multiple Air Interface Variants (AIV). In order to do so, functional decomposition is applied to create a multi-layer resource management.

From an architectural standpoint, the system is split into two layers, Access Network Outer (AN-O) and Access Network Inner (AN-I), plus the 5G UE. Those two layers are respectively conceived to overarch multiple AIVs and to manage a specific AIV. Slicing overarching multiple AIVs is enabled by a novel entity residing in the AN-O layer, denominated AIV-agnostic Slice Enabler (AaSE), which performs traffic monitoring, scheduling, and traffic steering towards the appropriate AIV. Modules within each AIV will then perform short term radio resource management from inside the AN-I layer.

A.2.3 State of the 3GPP standardisation on V2X support

3GPP has recently embraced vehicular communications and many of their technical enablers that will make them in its specifications for the Release 15. In this section we will provide an overview, and highlight the benefits they bring for the 5G support of V2X services.

3GPP Standard support for V2X and eV2X

3GPP has been actively supporting vehicular communications since its Release 14. The first round of specifications revolved around the support of “day 1” safety and non-safety communications, under the name of V2X [3GPP17-22185]. Starting from Release 15, the standardisation has started on enhanced V2X (eV2X), which foresees the support of novel, more advanced applications [3GPP17-22186].

A novel architecture extension has been proposed in Release 14 [3GPP17-23285], illustrated in a simplified form in Figure A.2. Two new functional entities have been introduced: a control plane V2X Control Function, and a data plane V2X Application Server. Alongside them, a new set of interfaces was defined to interconnect them with the pre-existing entities, with the UEs, and with the V2X application running on the UEs.

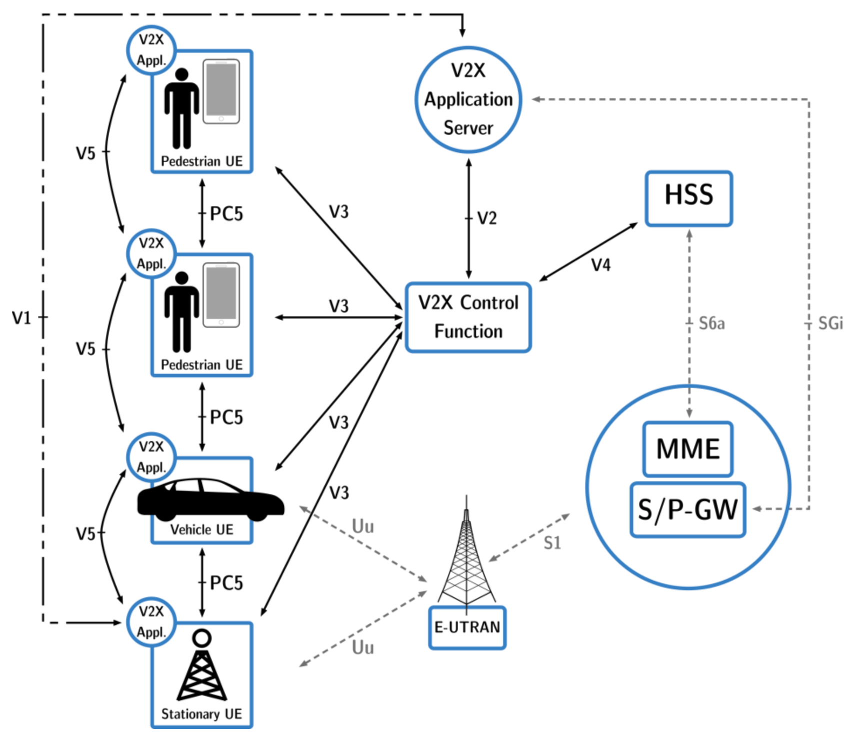


Figure A.2: V2X architecture, inspired by [3GPP17-23285].

The V2X control function is located within the home PLMN network, and is responsible for providing the UEs with parameters and configuration necessary for them to perform V2X communications.

The V2X application server is responsible for numerous data plane functions related to V2X, including receiving UL data transmitted by the UEs over the Uu interface, and delivering this data in a target area over unicast or MBMS.

B. Annex 2: U.S. / E.U. V2X security architecture

Both E.U. and U.S. security architectures follow similar principles and rely on Public Key Infrastructures (PKI) to deliver digital certificates to OBUs and RSUs (called EE "End Entities" hereafter – U.S. Security Credential Management System (SCMS) terminology).

Digital certificates formats for V2X are defined in [ETSI17-103097] for E.U., and [IEEE16-16092] for U.S. The latest [ETSI17-103097] standard is now based on [IEEE16-16092].

Every EE is assigned two kinds of certificates:

- A long-term certificate (a.k.a. enrolment certificate) which it gets during the enrolment phase: this long-term certificate is not used to sign V2X messages, but to sign requests sent to PKI Authorities (especially requests to get and/or download pseudonym certificates).
- Multiple short-term certificates (a.k.a. pseudonym certificates, or Authorisation Tickets AT), usually 20 per week, which are used by EEs to sign messages. For privacy purpose, the EE regularly changes its identity, which requires changing its current pseudonym certificate, hence the provision of 20 pseudonym certificates per week.

Digital certificate themselves are signed by a PKI authority: enrolment certificates are signed by an Enrolment Authority (EA, a.k.a. ECA, Enrolment Certificate Authority, or LTCA, Long Term Certificate Authority). Pseudonym certificates are signed by a Pseudonym Certificate Authority (PCA, a.k.a. Authorisation Authority AA). The mechanism is recursive, EA and PCA using their own certificates to sign enrolment and pseudonym certificates. The PKI architecture hence defines a tree of Authorities, one parent Authority signing the certificate of its child authorities. At the top is a Root Authority whose certificate is trusted by all.

A PKI defines a tree of Authorities. It also defines the protocol used by EEs to connect to different Authorities, and for Authorities to exchange with each other. See [ETSI16-102940] and [ETSI12-102941] for E.U., and [USDOT16-SCMS] for U.S.

A digital certificate specifies:

- A verification public key: this public key is used by the receiver of a V2X message to verify the signature of a message (if the verification is successful, it proves that the sender of the message owns the associated private key);
    - Note: this is true for explicit certificates. In U.S., enrolment and pseudonym certificates are implicit certificates, such certificates only specify a reconstruction value, instead of a verification public key and a (certificate) signature (hence an implicit certificate is shorter than an explicit one). The reconstruction value allows

reconstructing a public key, and if it allows successfully verifying a received message signature, then it will also prove that the certificate was issued by the expected Authority.

- An (optional) encryption public key: this public key can be used to encrypt data solely intended for the owner of that certificate (only the owner of the associated private key can decrypt the data);
- A geographical validity (e.g. one or multiple countries);
- A time validity;
- Service-specific permissions (SSP) stating whether the certificate allows signing a given V2X message, or a given content inside a V2X message;
  - When a signed V2X message is received, the received must not only cryptographically verifies the signature, but must also perform different relevance (about time and/or position for instance) checks and consistency (against SSP) checks.

Both E.U. and U.S. use elliptic curve based cryptographic operations: Elliptic Curve Digital Signature Algorithm (ECDSA) for signing, Elliptic Curve Integrated Encryption Scheme ECIES for encrypting. With 256 bits or 384 bits keys, hashes, …

All public/private key pairs are required to be generated by an Hardware Security Module (HSM), which ensure private keys cannot be disclosed. As a consequence, private key-based signing and decrypting cryptographic operations are performed by the HSM itself.

B.1 E.U. security architecture

shows the different E.U. PKI Authorities. See also [EU17-SECFMWK].

To get pseudonym certificates, an EE needs:

1. To be registered upon the EA. This is the “bootstrapping” procedure. For OBUs, it is expected that this registration will be performed by the car manufacturer itself, providing the vehicle’s canonical identifier (and associated public key);
2. To request an enrolment certificate to the EA; that request is signed with the public key associated to the registered canonical identifier;
3. To request pseudonym certificates, 20 per week, and for the weeks to come (so that the OBU can keep signing messages even though it cannot connect to PKI for some time); such requests are signed using the enrolment certificate;

It should be noted that the PKI protocol is designed in such a way that the AA, when answering the request for a pseudonym certificate (AT) does not know (and cannot track) the identity of the requesting EE. Hence the AA will not be able to tell whether two ATs have been assigned to the same EE. Similarly, the EA will not know which ATs have been generated for a given EE. Hence privacy is ensured against both EA and AA.

To sign a V2X message, an EE computes a signature and attaches its certificate to the message (signature and certificate are part of the security envelope defined in [ETSI17-103097]). For the 802.11p bandwidth sake, a specific rule applies to CAM messages (usually sent at 10Hz by OBUs): the certificate is not attached to every message, it can be replaced by its HashedId8 (i.e. 8 lower bytes of the certificate's hash). [ETSI17-103097] defines a mechanism by which one OBU can request to other OBUs a missing certificate.

There are two other critical functions supported by the PKI protocol:

1. It allows EEs regularly requesting Certification Revocation Lists (CRLs). A CRL allows identifying certificates that have been revoked (and thus messages signed by such certificates should not be accepted); A CRL specifies a list of certificate's HashedId10 (i.e. 10 lower bytes of the certificate's hash).
2. It allows EEs regularly requesting the last up-to-date Certificate Trust List (CTL) that provides PKI Authorities certificates and access points (e.g. URL).

A third important function (on-going standardisation) will allow misbehaviour reporting.

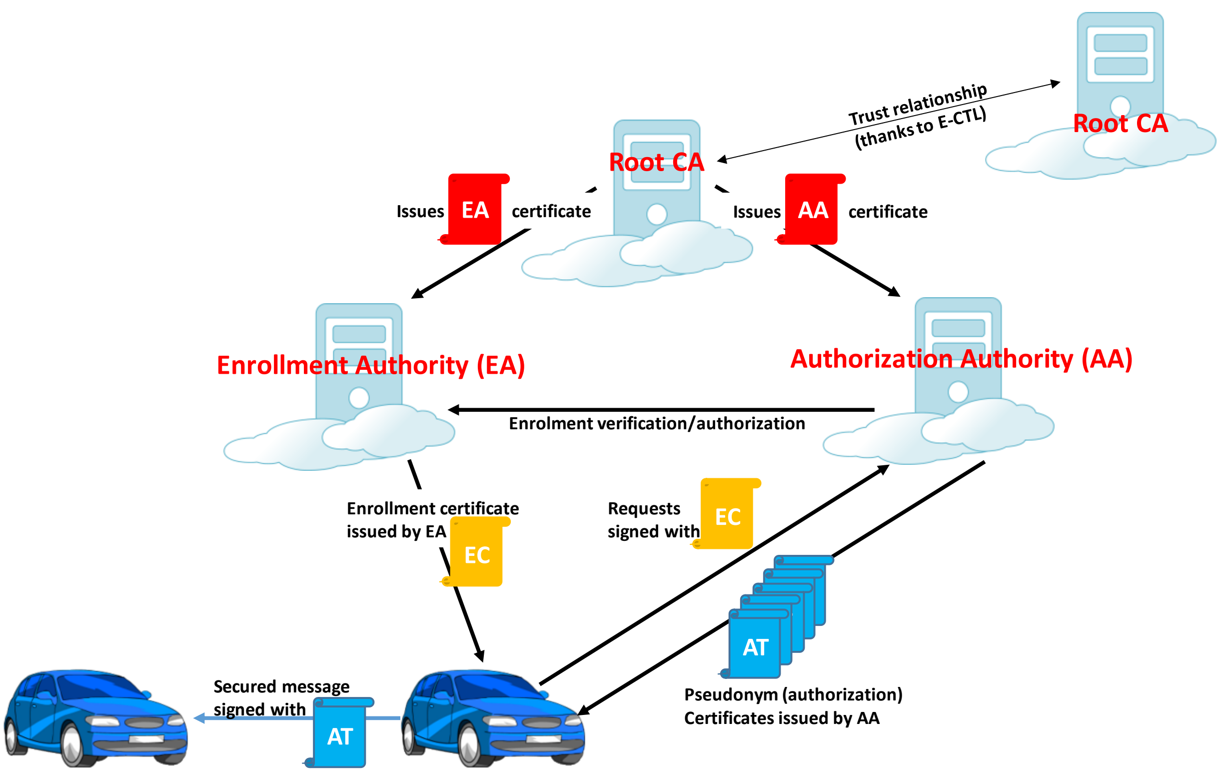


Figure B.1: E.U. Security architecture (source: [ETSI16-102940])

B.2 U.S. security architecture

Figure B.2 shows the different U.S. PKI (SCMS) Authorities.

To get pseudonym certificates, an EE needs:

1. To request an enrolment certificate; it is expected that this step will be triggered by the car manufacturer;
2. To trigger the generation of pseudonym certificates: this request is issued once (in normal situation), signed with the enrolment certificate, and requesting the generation of pseudonym certificates batches.

If it accepts that request the SCMS PKI (actually the RA and PCA together) will generate 3 years of per-week pseudonym certificates batches (each batch contains 20 certificates, or more exactly 20 encrypted – and signed - certificate responses).

The whole process is based on so-called butterfly keys: the initial request sent by the EE specifies two seed public keys (and some additional crypto materials) that will be used by the PKI to generate two sets of keys: one set to encrypt every certificate response; and one set to generate the public verification key found in each pseudonym certificate. The two private keys associated to the two seed public keys are required:

a. to reconstruct the private key allowing to decrypt each certificate response;
b. to reconstruct the private key associated to each pseudonym certificate (and therefore to sign messages with that certificate)

Note: the above description is over-simplified for the sake of this overview section. The important point is that only the EE whose HSM stores the two seed private keys can decrypt the certificate responses, and can use the pseudonym certificates.

3. To download per-week batches of pseudonym certificates. The EE can download as many batches as it wants, possibly all 3 years if it has enough storage capacity.

It should be noted that the PKI protocol is designed in such a way that the PCA cannot tell (and cannot track) the identity of the requesting EE. Hence the PCA will not be able to tell whether two pseudonym certificates have been assigned to the same EE. Similarly, the RA will not know which pseudonym certificates have been generated for a given EE. Hence privacy is ensured against both RA and PCA.

To sign a V2X message, an EE computes a signature and attaches its certificate to the message (signature and certificate are part of the security envelope defined in [IEEE16-16092]). For the 802.11p bandwidth sake, a specific rule applies to BSM messages (usually sent at 10Hz by OBUs): the certificate is not attached to every message, it can be replaced by its HashedId8 (i.e. 8 lower bytes of the certificate's hash). [IEEE16-16092] defines a mechanism (Peer-to-Peer Certificate Distribution, P2PCD) by which one OBU can request to other OBUs any missing certificate.

There are two other critical functions supported by the PKI protocol:

1. It allows EEs regularly requesting Certification Revocation Lists (CRLs). A CRL allows identifying certificates that have been revoked (and thus messages signed by such certificates should not be accepted);

A CRL specifies either certificates' HashedId10 (i.e. 10 lower bytes of the certificate's hash), or linkage value seeds. Linkage value seeds allow revoking a set of certificates at once. Two Linkage Authorities (two, so that one alone is not able to allow identifying all pseudonym certificates assigned to a given EE) are in charge of generating linkage values for generated pseudonym certificates. Disclosure of the two seeds (by the PKI Authority) allow verifying whether a certificate's linkage value matches and has therefore been revoked.

2. It allows EEs regularly requesting the last up-to-date Local Certificate Chain File (LCCF) that provides PKI Authorities certificates and access points (e.g. URL). It also allows EEs regularly requesting the last up-to-date Local Policy File (LPF).

A third important function (on-going standardisation) will allow misbehaviour reporting to the SCMS Misbehaviour Authority (MA).

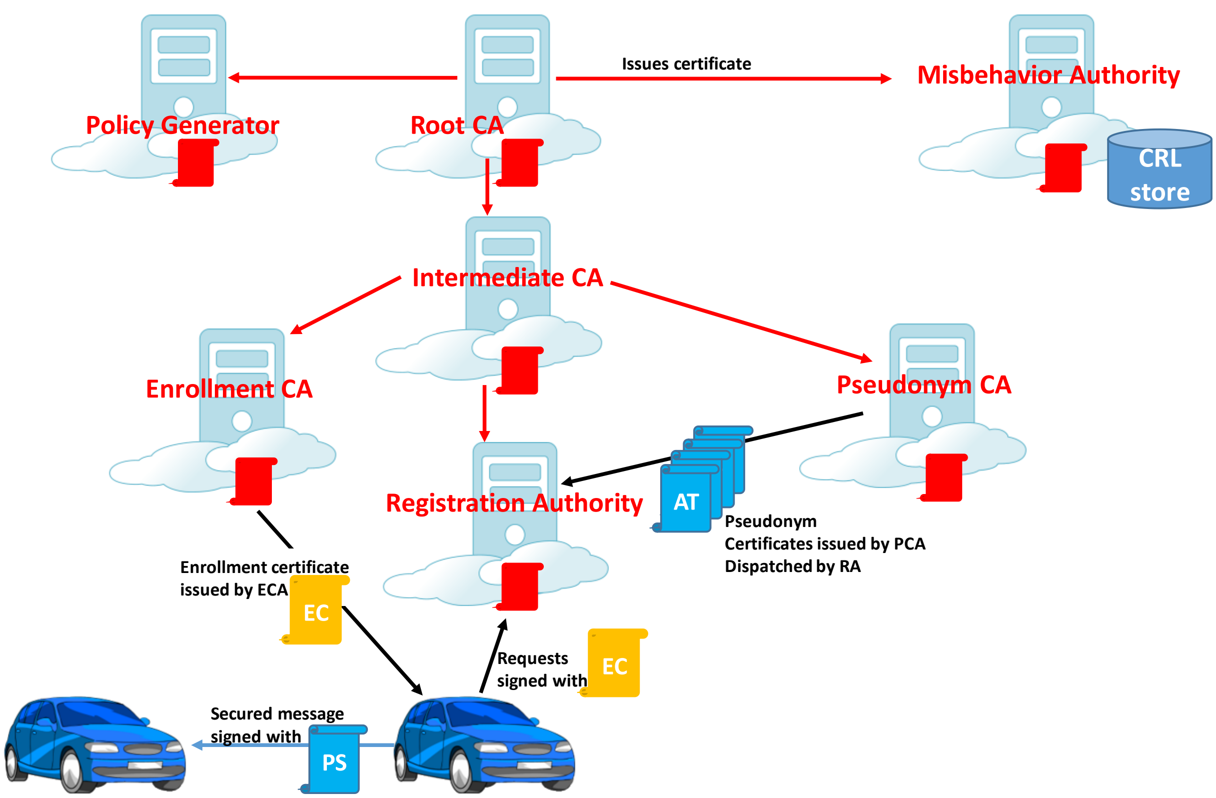


Figure B.2: SCMS QA environment (source: [CAMP17])

C. Annex 3: Synthetic description of technical components

C.1 RSU-enabled smart zone (SM-Zone)

Title of Technical Component	RSU enabled Smart Zone (SM-Zone) for V2X communications
Short Description of Problem	*The problem herein consists of a number of essential issues, including:* *To eliminate half-duplexing and out-of-range or hidden-terminal issues in direct V2V communications* *To enhance reliability and efficiency of PC5 like V2V transmissions which is in principle based on transmitter oriented broadcast without L1/L2 feedback* *To enable and facilitate URLLC support for advanced V2X* *To provide efficient traffic flow and mobility management of vehicles' devices (on-road traffic remains local and specific to individual road)*
Short Description of Solution	*Exploring smart use and deployment of road side units (RSU) which are enhanced to be network entities of 5G network architecture in supporting advanced V2X communications, RSU enabled SM-Zone is introduced to address the above issues:* *5G integrated radio service coverage layer local to individual road areas for V2X supports, primarily provided by RSUs which are:* *deployed along the roads, e.g., mounted on road side lamp posts with small footprints* *functions and services adapted to enable and facilitate an optimized support of targeted challenging V2X use cases and requirements* *coordinated and controlled by a macro coverage 5G network* *UE, once admitted to a SM-Zone, may transmit/receive at least to/from RSUs in RSU provided SM-Zone* *SM-Zone based resource allocation and mobility control which may be decoupled with that of operator specific macro mobility cellular access layer(s) for simple, effective and efficient supports of multi-operator, -service, -device, etc.*
Involved Entities	UE (Vehicle, Pedestrian, …), RSU, gNB, AMF, UPF, AF (toward ITS's facility and app level)

Involved Interfaces	Uu, PC5, and network interfaces between RSU and RSU, RSU and gNB, RSU towards CN (both edge clouds for local access and core cloud for remote access)
Target V2X Use Class(es)	Cooperative manoeuvre, Cooperative perception, Cooperative safety, Autonomous navigation, Remote driving, Others
Target V2X Use Case(s)	Lane merge, See-through, Network assisted vulnerable pedestrian protection, High definition local map acquisition, Remote driving for automated parking, Others
Target Performance Requirements for Optimisation	Latency, Reliability, Capacity, Data Rate, Position Accuracy,

C.2 SON based multi-mode RSU for efficient QoS support and congestion control

Title of Technical Component	**SON based multi-mode RSU for efficient QoS support and congestion control of V2X communications**
Short Description of Problem	*The problem is motivated by the following observations, besides the need of efficient QoS and congestion control in general:* *There is a wide range of V2X applications and services with diverse QoS requirements* *Traffic flows related to, e.g., road safety and road efficiency apps, are specific to individual road and should better be handled locally to individual road* *Traffic flows related to remote access services such as remote driving, map update or infotainment require extreme-MBB mobile connections causing high load to serving cellular network* *There is a large number of mobile devices on roads with various road traffic patterns, ranging from almost deserted to long lasting massive congested.*
Short Description of Solution	*Because using direct D2D or sidelink (SL) allows for maximizing localisation of local V2X traffic flows whereas using Uu via serving RAN allows for better control and multiplexing gain, flexible and adaptive use of multi-mode RSUs to best serve data traffic demands of actual or predictable road traffic situation enables an optimized QoS support and congestion control of V2X. Thus, the solution herein includes:* *To introduce enhanced multi-mode or hybrid-type RSU, UE-type RSU,*

	NB-type RSU, hybrid-type RSU or any combination thereof *To provide a SON based adaptive control and operation scheme for multi-mode RSU a configuration and control of multi-mode RSU to best serve individual local road where RSU is deployed*
Involved Entities	UE (Vehicle, Pedestrian, …), RSU, gNB, AMF, UPF, AF (toward ITS's facility and app level)
Involved Interfaces	Uu, PC5, and network interfaces between RSU and RSU, RSU and gNB, RSU towards CN (both edge clouds for local access and core cloud for remote access)
Target V2X Use Class(es)	Cooperative manoeuver, Cooperative perception, Cooperative safety, Autonomous navigation, Remote driving, Others
Target V2X Use Case(s)	Lane merge, See-through, Network assisted vulnerable pedestrian protection, High definition local map acquisition, Remote driving for automated parking, Others
Target Performance Requirements for Optimisation	Latency, Reliability, Capacity, Data Rate

C.3 SL and Uu multi-connectivity for high reliable and/or high data rate V2V communication

Title of Technical Component	**SL and Uu multi-connectivity for high reliable and/or high data rate V2V communication**
Short Description of Problem	*SL communication is beneficial and preferred for many V2X applications and services that generate only local data traffic in proximity. However, for vehicle services that require high reliability and/or have high data rate, the SL alone may not be sufficient to meet the requirements especially in very complex environment.*
Short Description of Solution	*Exploring the SL and Uu multi-connectivity in which the SL is established as the primary connection for E2E communication among vehicle devices and secondary connection via gNB is also established for the same E2E communication, the solution focuses on:* *efficient protocol design for SL and Uu multi-connectivity* *corresponding signalling mechanism and procedures for more efficient resource usage and power consumption, taking into account mobility management in multi-cell and multi-operator scenario*

Involved Entities	UE (Vehicle, Pedestrian, …), gNB, AMF, ….
Involved Interfaces	
Target V2X Use Class(es)	Cooperative manoeuver, Cooperative perception, Cooperative safety, Autonomous navigation, Others
Target V2X Use Case(s)	Lane merge, See-through, Network assisted vulnerable pedestrian protection, High definition local map acquisition, Others
Target Performance Requirements for Optimisation	Latency, Reliability, Capacity, Data Rate, …

C.4 Location aware scheduling

Title of Technical Component	**Location aware scheduling**
Short Description of Problem	Traffic load in the network can vary a lot over time and across different areas. In case of large amount of data that requires reliable and urgent delivery, the scheduler has little freedom in assigning resources and system performance is degraded.
Short Description of Solution	The solution proposes to adjust individual packet QoS provisioning based on location-related information.
Involved Entities	May involve application server, or a network node in the core, or gNB, or vehicle
Involved Interfaces	In general, involve the interface between packet-characteristic description and QoS provisioning. More specifically, may involve the interface between core network and RAN, or the interface between application server (or additional servers providing location-related information) and core network.
Target V2X Use Class(es)	Cooperative safety, Autonomous navigation
Target V2X Use Case(s)	High definition local map acquisition, CAM dissemination
Target Performance Requirements for Optimisation	Latency, Capacity, Reliability

C.5 IaaS for vehicular domain

Title of Technical Component	**Infrastructure as a Service for ITS domain**
Short Description of Problem	*Deployment of ITS software architecture at large scale*
Short Description of Solution	*The use of IaaS (infrastructure as a service) allows to automate the software deployment, update / upgrade, self - healing (back to the initial state after a failure) …*
Involved Entities	MEC, Regional PoP, Centralized PoP in countries, Centralized POP
Involved Interfaces	Internally at the system : OSS-BSS
Target V2X Use Class(es)	Cooperative manoeuver, Cooperative perception, Cooperative safety, Autonomous navigation, Remote driving, Others
Target V2X Use Case(s)	Lane merge, See-through, Network assisted vulnerable pedestrian protection, High definition local map acquisition, Remote driving for automated parking, Others
Target Performance Requirements for Optimisation	Latency, Reliability, Capacity,…

C.6 Redundant mode PC5 + Uu

Title of Technical Component	**Redundant mode PC5 + Uu**
Short Description of Problem	*Very short Latency and high Reliability are becoming a challenge to address for many Use Cases in 5G. In several domains a well know solution is to use a redundant system approach.*
Short Description of Solution	*In other to achieve higher reliability without introducing too much complexity a solution could be to implement a redundant mode using concurrently both radio interfaces (aka PC5 on 5.9 GHz band and Uu on Licensed band). The sender could be a Vehicle/Pedestrian and the receiver another Vehicle/Pedestrian in his close surroundings, both under network coverage. This solution is firstly mainly forecasted for V2V type of services. Alternatively, with a RSU using PC5*

	communications this solution could be also applied to V2I/V2N services. *This should be pointed out though that it will consume more global bandwidth because the same flow will be routed on both interfaces. In conclusion it should be restricted only for a limited usage in time and not for high throughput demanding services such as the See-through Use Case.*
Involved Entities	UE(Vehicle/Pedestrian), gNB, RSU (for V2I/V2N services)
Involved Interfaces	Uu and PC5
Target V2X Use Class(es)	Possibly all Classes: Cooperative manoeuver, Cooperative perception, Cooperative safety, Autonomous navigation, Remote driving
Target V2X Use Case(s)	For V2V & V2I/V2N: Use Case 1: Lane merge For V2I/V2N: Use Case 3: Network assisted vulnerable pedestrian protection, Use Case 4: High definition local map acquisition, Use Case 5: Remote driving for automated parking
Target Performance Requirements for Optimisation	Data rate < 6.4 Mbps Latency: <30ms Mobility: 30<km/h<130 Reliability: 99% to 99.999%

C.7 Evolution of infrastructure-based communication for localised V2X traffic

Title of Technical Component	**Evolution of infrastructure-based communication for localised V2X traffic**
Short Description of Problem	*The data traffic to be exchanged among the vehicles for many V2X services (Cooperative manoeuvers, cooperative perception) is expected to have localised significance (among devices located at the same geographical region, no need to access a remote server (e.g., ITS cloud server)) together with demanding QoS requirements (low latency, high reliability). Available schemes do not manage to handle local traffic with various transmission modes (unicast, multicast, broadcast) efficiently, because they require complex combinations of diverse implementations.*
Short Description of	*To establish and manage "local end-to-end (e2e)" paths with*

Solution	*guaranteed e2e link performance to address the fast and reliable transmission of localised data traffic among the involved devices. The (user plane) radio data paths are established among the involved communicating end devices (i.e., vehicles),* *The paths are established by (and via) the BSs (i.e. the nodes of the core network do not participate in the user plane transmissions), since the data traffic is localised.*
Involved Entities	UEs, gNBs, AMF
Involved Interfaces	Uu, Xn, PC5, NG-C
Target V2X Use Class(es)	Cooperative manoeuver, Cooperative perception, Cooperative safety, Autonomous navigation
Target V2X Use Case(s)	Lane merge, See-through, Network assisted vulnerable pedestrian protection
Target Performance Requirements for Optimisation	Latency, Reliability

C.8 Use case-aware multi-RAT, multi-link connectivity

Title of Technical Component	**Use case-aware Multi-RAT/Multi-Link connectivity**
Short Description of Problem	*Multi-RAT connectivity for reduced delay or increased reliability*
Short Description of Solution	*Use case specific radio access network association*
Involved Entities	Different BS/Aps, and side link
Involved Interfaces	
Target V2X Use Class(es)	all
Target V2X Use Case(s)	all
Target Performance Requirements for Optimisation	Latency, reliability, successful completion of a manoeuver.

C.9 Multi operator solutions for V2X communications

Title of Technical Component	**Multi operator solutions for V2X communications**
Short Description of Problem	*Communication among vehicles cannot be ensured that all the vehicles belong in the same operator. This causes increased delays or communication interruptions.*
Short Description of Solution	*Multiple registration of each vehicle to all operators may enable vehicles to switch from one operator to the other in a fast way. Regional split with this multiple registration facilitates the handling of other aspects of the problems as well (i.e., MEC coordination). Enhancements to the Radio Control protocols are required for ensuring the ability of the UE to communicate with the V2X network with multiple operators (in the transition area)*
Involved Entities	UEs, gNBs, AMF (MME)
Involved Interfaces	*Uu, PC5, [Inter-operator AMF?]*
Target V2X Use Class(es)	Cooperative manoeuver, Cooperative perception, Cooperative safety, Autonomous navigation, Remote driving), eMBB (mobility & Comfort), mMTC (remote maintenance, data collection)
Target V2X Use Case(s)	Lane merge, See-through, High definition local map acquisition, Remote driving for automated parking,
Target Performance Requirements for Optimisation	Latency, Reliability

C.10 V2X service negotiation

Title of Technical Component	**V2X service negotiation**
Short Description of Problem	A variation or unavailability of QoS or communication KPIs might impact the effectiveness of some V2X services. Currently, a service might react to the experienced QoS in real-time (e.g., adjust data rate) but the

	adaptation of V2X services in real-time is not optimal/ideal due to additional delays for accelerating/braking.
Short Description of Solution	A V2X service negotiation could allow OEM and MNO to exchange information about the service requirements (not only QoS, but with additional information such as vehicle speed, geographical relevance of data, etc.) and some sort of network capabilities (e.g., the service can be supported with a certain QoS) both in advance and in real-time to allow proper time to the OEM to adapt the service behaviour.
Involved Entities	May involve an application server, a network node or vehicles
Involved Interfaces	May involve the interface between a V2X application and the core network or the RAN.
Target V2X Use Class(es)	Potentially all use classes
Target V2X Use Case(s)	Potentially all use cases, main focus on lane merge and HD local map acquisition as representation of two use cases with different features
Target Performance Requirements for Optimisation	Latency, capacity, reliability

C.11 Edge computing in millimetre wave cellular V2X networks

Title of Technical Component	**Edge Computing in Millimetre Wave Cellular V2X Networks**
Short Description of Problem	*Edge Cloud access through mmWave radio*
Short Description of Solution	*an online algorithm for solving the stochastic optimisation problem of association of vehicles to mmWave gNB*
Involved Entities	Vehicle, gNB/RSU
Involved Interfaces	
Target V2X Use Class(es)	Cooperative manoeuver, Autonomous navigation
Target V2X Use Case(s)	Lane merge, High definition local map acquisition

Target Performance Requirements for Optimisation	Latency, precision of task

C.12 Dynamic selection of PC5 and Uu communication modes

Title of Technical Component	**Dynamic Selection of PC5 and Uu Communication modes**
Short Description of Problem	*5G communication systems will support both cellular (Uu) and sidelink (PC5) communication interfaces (modes) that have different transmission characteristics and features. Cellular Uu (e.g., local e2e paths) has larger coverage area, while PC5 (e.g., mode 3) increases the system capacity through the spatial frequency reuse.*
Short Description of Solution	*To dynamically select the appropriate communication modes for the vehicles that participate in V2X service*
Involved Entities	UEs, gNBs, AMF
Involved Interfaces	Uu, Xn, PC5, NG-C
Target V2X Use Class(es)	Cooperative manoeuver, Cooperative perception, Cooperative safety, Autonomous navigation, Remote driving, Others
Target V2X Use Case(s)	Lane merge, See-through, Network assisted vulnerable pedestrian protection, High definition local map acquisition, Remote driving for automated parking, Others
Target Performance Requirements for Optimisation	Latency, Reliability, Data Rate

C.13 Security and privacy enablers

Title of Technical Component	**Security and privacy enablers for C-V2X**

Short Description of Problem	*V2X messages exchanged between UEs (OBUs, RSUs, VRUs, …), or between UEs and V2X application servers, need to be secured, while ensuring privacy of the sender. Secured means that the sender (and possibly receiver) must be authenticated, and that his permissions to send a V2X message, with any particular content, and within a given context (time, geographical region, …) must be verified before that message may trigger any safety (or not) action on the receiver's side.*
Short Description of Solution	*Relying on C-ITS digital certificates, and a PKI architecture compliant with the European Commission C-IITS certificate policy, the relevance and consistency and cryptographic verifications may either be done by the receiver, or by the Network. To avoid the latency incurred by HSMs when signing, some application-specific messages will only be encrypted, with a symmetric key provided during some first signed exchanges. Though not applicable to all general-purpose awareness messages that always need to be signed.*
Involved Entities	UE (Vehicle, Pedestrian, …), UPF, ….
Involved Interfaces	
Target V2X Use Class(es)	All V2X Use Classes
Target V2X Use Case(s)	All V2X User Cases
Target Performance Requirements for Optimisation	Latency, Reliability, Privacy

C.14 5G core network for edge-computing-based mobility

Title of Technical Component	**5G core network for edge-computing-based mobility**
Short Description of Problem	*In conventional mobile networks, mobility of a mobile terminal is enabled by handover procedures when it changes serving access point (or base station).* *However, such handover mechanisms are insufficient to support reliable and punctual computation offloading in the environment of 5GCAR Edge computing.*
Short Description of	*The "service recovery" can be conducted in advance to substantially*

Solution	*reduce service interruptions. Then, it is possible to propose pre-migrates computation jobs when the handover is about to happen.*
Involved Entities	UE(Vehicle), gNB, AMF, UPF, new dedicated NF
Involved Interfaces	
Target V2X Use Class(es)	Cooperative manoeuver, Cooperative perception, Cooperative safety, Remote driving
Target V2X Use Case(s)	Lane Merging, See-through, pedestrian protection
Target Performance Requirements for Optimization	Latency: <30ms Mobility: 30<km/h<130