AECC

AUTOMOTIVE EDGE
COMPUTING CONSORTIUM

AECC Technical Report v2.0

Automotive Edge Computing Consortium (AECC)

Technical Solution Working Group (WG2)

Driving Data to the Edge: The Challenge of Data Traffic Distribution

Contents

Executive Summary

Connected Vehicles are swiftly transforming the automotive industry, with emerging services driving the requirement for high volumes of data communication. New Connected Vehicle services are expected to make the Automotive sector the industry segment with the fastest-growing demand for mobile machine-to-machine connectivity. In principle, every new vehicle that is manufactured will be continuously connected; it will also generate massive volumes of data that need to be transmitted from the vehicle to the cloud.

Considering the global distribution of vehicle fleets and therefore the global nature of this challenge, stakeholders, including vehicle manufacturers, technology solution vendors, network operators, cloud infrastructure and service providers, should consider how communication networks and computing resources can be orchestrated securely and cost-effectively to enable these new Mobility Services.

The members of the Automotive Edge Computing Consortium (AECC) are working to articulate an architecture capable of addressing the critical industry needs. To address this challenge, the AECC has proposed a “Distributed Computing on Localized Networks” solution concept and architecture that will offer the service flexibility and efficiency required to support the evolution of the automotive industry into and beyond the Connected Vehicle era. The solution consists of three main aspects: Localized Networks, Distributed Computing and Local Data Integration. This technical report, which provides a summary of the in-depth study performed by AECC members, focuses on six of the key issues and corresponding solutions as well as initial high-level results for the Distributed Computing aspect. Specifically, these key issues are:

Edge Data Offloading. How Cellular Networks can support the offloading of data from the cellular network to appropriate localized distributed computing infrastructure in an efficient and flexible manner, considering the mobility of vehicles and service requirements.

MSP Server Selection. How Connected Vehicles are able to select Mobility Service instances in the distributed computing infrastructure.

Vehicle System Reachability. How Connected Vehicle can be awakened and contacted, despite vehicle mobility and network topology changes.

Access Network Selection. How Connected Vehicles can select appropriate Access Networks according to service requirements and capabilities of the access networks.

Provisioning and Configuration Update. How configuration parameters and policies can be provisioned to Connected Vehicles in a dynamic and distributed environment.

Opportunistic Data Transfer. How cellular network resources can be used to offer latency-tolerant data transfer with minimal interference to existing services.

To help stakeholders consider how to address these issues, the AECC has identified and evaluated a range of potential solutions, with this technical report containing the resulting recommendations.

The AECC continues to work on additional key issues and solutions relating to data transfer between vehicle and cloud, with further studies to be released in the future. The AECC welcomes feedback on the findings and recommendations contained within this report and will take that feedback into consideration in its ongoing work.

Terms and Abbreviations

Terms:

Access Network
The network used to connect to the immediate service provider. It may contain a WLAN, Cellular Network and/or wired network.

Cloud
A logical, location independent computing platform that hosts services to store, manage and process data and which is implemented on a set of remote servers.

Cellular Network
A network that is typically defined by 3GPP and GSMA and operated by an MNO.

Center Server
A hardware and/or software platform to hosts center Mobility Services.

Computing Infrastructure
The resources and services on which other systems and services are built.

Computing Facility
A physical facility that hosts computing infrastructure and related resources (including compute, network and storage, power, cooling etc.).

Connected Vehicle
A network-attached vehicle that shares data with other network-attached devices and servers.

Distributed Computing
A computing paradigm that divides a problem into many tasks that can be addressed by many computers.

Edge Computing
A type of distributed computing system where applications, memory and processing power are allocated to other computers in order to provide desired service levels.

Edge Server
A localized hardware or software platform that hosts edge Mobility Services.

Enterprise Network
A network connecting Center and Edge Servers for a specific enterprise.

Intelligent Driving
A service that augments an Advanced Driver Assistance System (ADAS) or an Automated Driving System with strategic decisions based on predictions of conditions along route alternatives that are gathered using connectivity to external sources.

Local Data Integration
A platform that integrates data from Localized Networks and the Distributed Computing system.

Localized Network
A local network that covers a limited domain such as a defined geographical area.

Mobility as a Service
Integration of various forms of transport services into a single mobility service accessible on demand.

Mobility Service — A service provided to the passengers, the drivers or the vehicle manufacturers (e.g., telematics, traffic, map, car/ride sharing, insurance, etc.).

Mobility Service Provider — A platform-independent provider that provides customers with access to one or more Mobility Services.

MSP Server — A Center Server or an Edge Server operated by a Mobility Service Provider.

Network Edge — One or more locations within a network domain in close adjacency to the source of the data producer/consumer.

V2Cloud — Communication between a vehicle and applications or services hosted on a cloud.

Vehicle System — A system composed of a computing platform, applications, services and other components residing in the Connected Vehicle.

Abbreviations:

3GPP	The 3rd Generation Partnership Project
5G	5th Generation
5GMS	5G Media Streaming
5GS	5G System
ACDC	Application-specific Congestion control for Data Communication
ADAS	Advanced Driver Assistance System
AF	Application Function
AMF	Access Management Function
ANDSF	Access Network Discovery and Selection Function
API	Application Programming Interface
APN	Access Point Name
BDT	Background Data Transfer
BSF	Bootstrapping Server Function
C-V2X	Cellular Vehicle-to-Everything
CAN	Controller Area Network
CUPS	Control User Plane Separation
DHCP	Dynamic Host Configuration Protocol
DNS	Domain Name System
DRB	Data Radio Bearer
DSRC	Dedicated Short Range Communication
ECU	Electronic Control Unit
eNB	Evolved Node B (LTE Base Station)
EPC	Evolved Packet Core
EPS	Evolved Packet System
EN-DC	E-UTRA-NR Dual Connectivity
FQDN	Fully Qualified Domain Name
GMA	Generic Multi-Access
gNB	Next Generation Node B (5G Base Station)
GPRS	General Packet Radio Services

GSMA	Global System for Mobile communications Association
GTP	GPRS Tunneling Protocol
HSS	Home Subscriber Server
HTTP	HyperText Transfer Protocol
IaaS	Infrastructure as a Service
IMEI	International Mobile Equipment Identity
IoT	Internet of Things
ISMP	Inter-System Mobility Policy
ISRP	Inter-System Routing Policy
IPSec	Internet Protocol Security
LTE	Long Term Evolution (the 4th Generation Mobile Communication Radio)
LWA	LTE WLAN Aggregation
MAC	Medium Access Control
M2M	Machine to Machine
MaaS	Mobility as a Service
MME	Mobility Management Entity
MNO	Mobile Network Operator
MSP	Mobility Service Provider
MPQUIC	Multi-Path QUIC
MPTCP	Multi-Path TCP
N3IWF	Non-3GPP Inter-Working Function
NAPT	Network Address Protocol Translation
NAT	Network Address Translation
NEF	Network Exposure Function
NF	Network Function
NR	New Radio (the 5th Generation Mobile Communication Radio)
OMA	Open Mobile Alliance
OS	Operating System
OTA	Over the Air
PaaS	Platform as a Service
PCC	Policy and Charging Control
PCF	Policy Control Function
PCRF	Policy and Charging Rules Function
PDN	Packet Data Network
PDU	Packet Data Unit
P-GW	Packet Gateway
P-GW-U	P-GW User Plane
QoS	Quality of Service
QUIC	Quick UDP Internet Connections
RAN	Radio Access Network
RAT	Radio Access Technologies
RRC	Radio Resource Control
RSD	Route Selection Descriptor
RTT	Round Trip Time
SD-WAN	Software-Defined Wide Area Network
S-GW	Serving Gateway

S-GW-U	S-GW User Plane
SCEF	Service Capability Exposure Function
SINR	Signal to Interference plus Noise Ratio
SIPTO	Selected IP Traffic Offload
SLA	Service Level Agreement
SMF	Session Management Function
SSC	Session and Service Continuity
SSID	Service Set Identifier
TCP	Transmission Control Protocol
UDP	User Datagram Protocol
UDR	User Data Repository
UDT	Unattended Data Traffic
UE	User Equipment
UL-TFT	Uplink Traffic Flow Template
UPF	User Plane Function
URSP	UE Route Selection Priority
USIM	Universal Subscriber Identification Module
UUID	Universally Unique Identifier
WLAN	Wireless Local Area Network
XMPP	Extensible Messaging and Presence Protocol

1 Introduction

Connected Vehicles are anticipated to be a significant factor contributing to the growth of communications data volumes, with forecasts projecting every new vehicle produced being “connected” by 2025 [1]. Millions of cars are already connected using 4G cellular access, and cellular broadband IoT connectivity (4G/5G) is expected to grow significantly through 2024 as outlined in the Ericsson Mobility Report [2] and visualized in the Ericsson Mobility Visualizer [3]. According to the Cisco Visual Networking Index [4], Connected Vehicles that incorporate applications such as fleet management, in-vehicle entertainment, internet access, roadside assistance, vehicle diagnostics, navigation and advanced driver assistance services will be the fastest-growing industry segment with respect to machine-to-machine connections [4]. Furthermore, many emerging automotive services, such as intelligent driving assistance and Mobility as a Service (MaaS), work on the expectation that vehicles will be connected to cloud computing facilities. The AECC estimates that the related data traffic globally has the potential to exceed 10 exabytes per month by 2025, a volume 1,000 times larger than the present numbers, as described in an AECC white paper [5].

Stakeholders in this ecosystem, such as vehicle manufacturers, technology solution vendors, network operators, cloud infrastructure and service providers, must establish a practical platform architecture capable of supporting the variety of Vehicle-to-Cloud (V2Cloud) services, considering that the global distribution of vehicle fleets, vehicle data communications and the processing of that data at scale presents a significant challenge to the currently deployed architectures.

The Cellular Network is one of the major Access Networks for Connected Vehicles, and many functions and mechanisms have been standardized in the 3rd Generation Partnership Project (3GPP). However, the present work within 3GPP has not fully addressed the challenge of automotive “big data,” and there is a high risk that future network deployments and business models will fail to support the emerging needs of connected vehicles. The cellular vehicle-to-everything (C-V2X) communication considered in 3GPP, for example, mainly covers latency-sensitive safety applications and does not address the forecast data volume growth between vehicles and the cloud. The 5G cellular system will provide improved functionality for both capacity and low latency, but the automotive industry will continue to use a mix of cellular Access Network technologies for the foreseeable future. In addition, increased data volume aggregated into data centers will cause network and processing congestion that degrades the user experience of Connected Vehicles.

The AECC believes that the current mobile communication network architectures and cloud computing deployments are not fully optimized to effectively handle emerging requirements of Connected Vehicles at a global scale. In response, the AECC proposes the use of a “Distributed Computing on Localized Networks” solution concept to solve these issues, with a proposal for a system architecture able to accommodate the predicted volumes of data traffic from Connected Vehicles. The concept focuses on three main aspects, which are Localized Networks, Distributed Computing and Local Data Integration.

The system architecture provides a framework that supports the distribution of computation processes across a set of Localized Networks, shown in Figure 1.

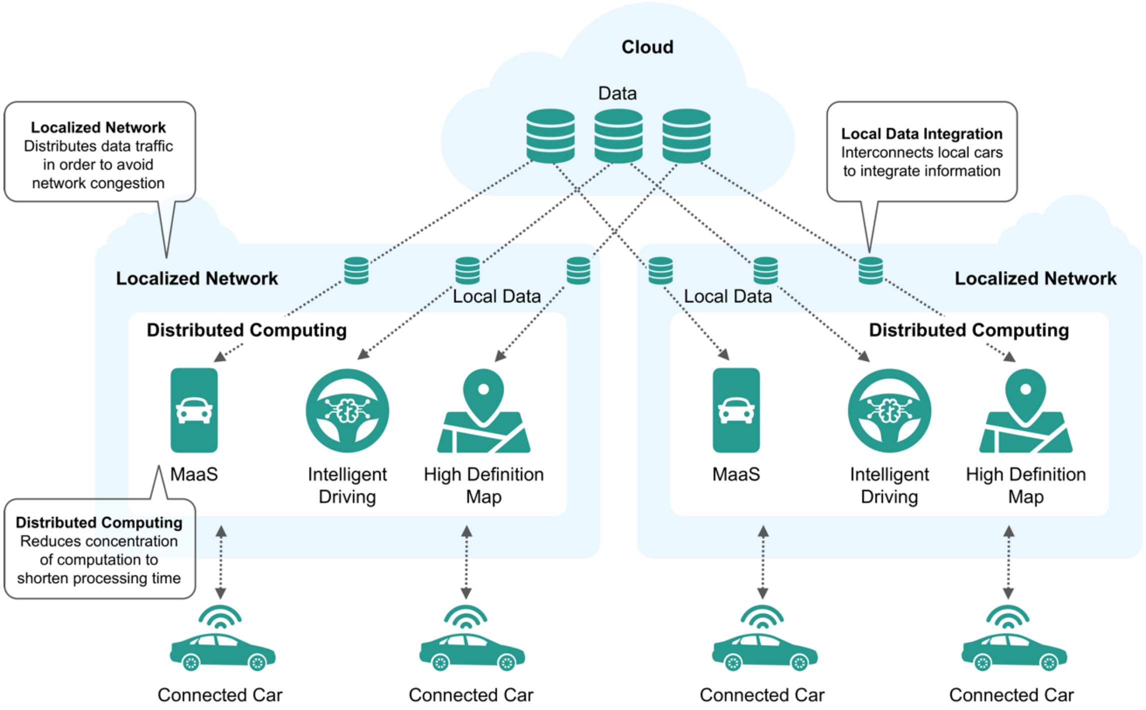

Figure 1. Three pillars of the AECC concept.

The AECC has identified a set of key issues, six of which were prioritized for study, as shown in Figure 2.

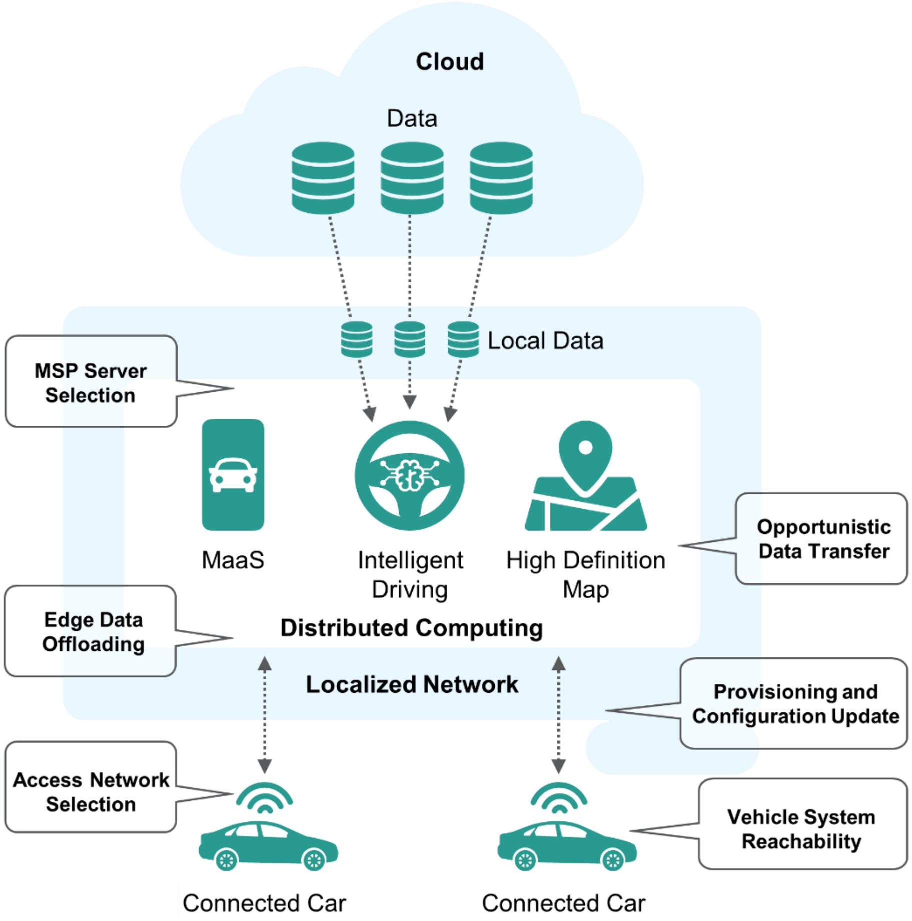

Figure 2. Six key issues addressed by this technical report.

1. Edge Data Offloading – Cellular access is an important network access method for Connected Vehicles. Traditional Cellular Networks are designed to have a data gateway, through which all data to and from the Connected Vehicle will need to pass. The traditional approach creates a bottleneck and can prevent the efficient routing of data traffic into the distributed computing architecture. Discussion of this key issue describes how Cellular Networks can support the offloading of data to appropriate local computing infrastructure in an efficient and flexible manner.
2. MSP Server Selection – Mobility Service application instances are expected to make use of the capabilities of the distributed computing architecture. As application instances will be distributed across both Edge and Centralized computing infrastructure, this brings a new challenge for Connected Vehicles – how to determine which is the most appropriate Mobility Service instance to communicate with. Discussion of this key issue describes how Connected Vehicles will be able to select and use Mobility Service instances.
3. Vehicle System Reachability – Connected Vehicles exhibit very different usage patterns and encounter more challenging connectivity conditions than most typical mobile devices such as smartphones. With vehicle manufacturers responsible for globally distributed fleets of vehicle with heterogeneous network coverage, discussion of this key issue describes how Connected Vehicles can be contacted across heterogeneous network access methods.

4. Access Network Selection – To meet the requirements of Mobility Services, a Connected Vehicle is expected to use a mix of different access technologies and access networks. Determining which network to use may consider aspects such as network bandwidth, capacity, coverage and reliability. Discussion of this key issue introduces the ways in which a Connected Vehicle can select from two or more Access Networks according to service requirements and network capabilities.
5. Provisioning and Configuration Update – With Mobility Services needing to be able to be provisioned in a flexible manner, adjusting to the changes within the numbers of location of Connected Vehicles, it is essential to be able to direct how Connected Vehicles should use the available services. The required policy may also need to adapt, reflecting the changing services environment. Discussion of this key issue describes how configuration parameters and policies can be provisioned to Connected Vehicles.
6. Opportunistic Data Transfer – Bandwidth is a resource that needs to be carefully managed in Cellular Networks. Discussion of this key issue investigates how cellular network resources can be best leveraged providing data transfer for latency-tolerant Mobility Services used by Connected Vehicles while minimizing the impact to other users of the cellular network.

In order to address these issues, the AECC identified and evaluated a range of potential solutions and has provided recommendations with respect to the different options.

The AECC also identified a preliminary distributed computing architecture to support the distributed data processing needs that are envisaged. A high-level description relating to this architecture is contained within this technical report, with further details to be published in future AECC reports.

2 System Overview

2.1 Architectural Requirements

The AECC proposes a system that will support the deployment and execution of Mobility Services using a distributed computing and networking architecture, including the Vehicle System, Networks (Cellular Network, WLAN, Fixed access, IP-based, etc.) and Center/Edge Servers. The primary goal of the architecture is to provide computing resources closer to where vehicle fleets are operating, enabling the processing of data to take place “in-region,” instead of pulling data back to centralized cloud environments or centralized infrastructure operated by vehicle manufacturers and vehicle fleet operators, etc.

The AECC System architectural requirements related to the Cellular Network shall apply to both 4G and 5G systems. Also, they shall apply to non-standalone 5G deployments, where the core network is 4G and the 5G New Radio (NR) is used as radio access for data communication.

The Vehicle System may connect to the AECC System using a WLAN when applicable. The WLAN may use Wi-Fi (IEEE802.11) as an access technology or 3GPP-based access technology (e.g., NR in millimeter wave carriers). The WLAN can connect to the cellular core network or to an internet service provider network.

The AECC System includes a distributed computing environment that will be used to support the various Mobility Services. Mobility Services are executed within the Vehicle System and Computing Facilities that may be composed of Center Server and/or Edge Servers located outside the Vehicle System.

2.2 Reference Architecture for Hierarchical Data Traffic Distribution

Taking the Intelligent Driving service as an example, the AECC System introduces a hierarchical data processing architecture as shown in Figure 3, where the data will be processed not only in the Vehicle System and the Cloud but also at the Edge Servers located between the Cloud and the Vehicle System.

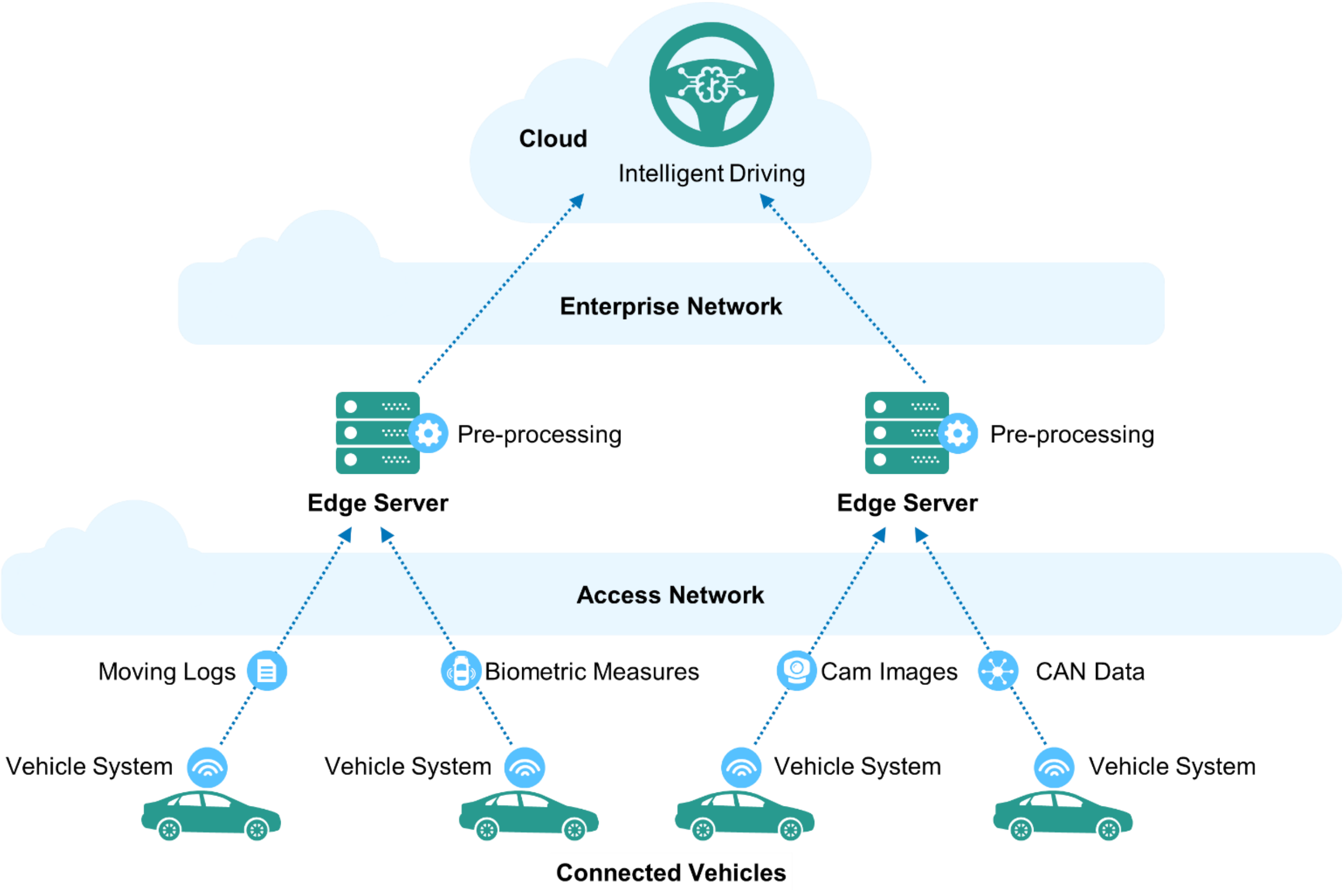

Figure 3. An example of the AECC System hierarchical data processing architecture.

2.2.1 Cellular Network Reference Models

In all cases of using Cellular Networks, the AECC System assumes the use of 3GPP reference models. This paper is written with an assumption that the reader is familiar with 3GPP reference models and associated terminology.

For the case where 4G is the only radio access used, the Evolved Packet System (EPS) architecture reference model is assumed [3].

For the case where non-standalone 5G New Radio (NR) is used as the radio access, the EPS architecture reference model according to 3GPP Release 15 and later (see [6]) with a non-standalone architecture option is assumed; i.e., 5G radio access is used in conjunction with an Evolved Packet Core (EPC). Figure 4 shows a non-roaming architecture as an example.

The 5G System (5GS) architecture reference model 3GPP 5GS Release 15 (see [7]) or later is assumed. This includes standalone NR deployment and multiple access dual connectivity options, with both NR and LTE radio connection to a 5G Core (5GC). Figure 5 shows a non-roaming architecture as an example.

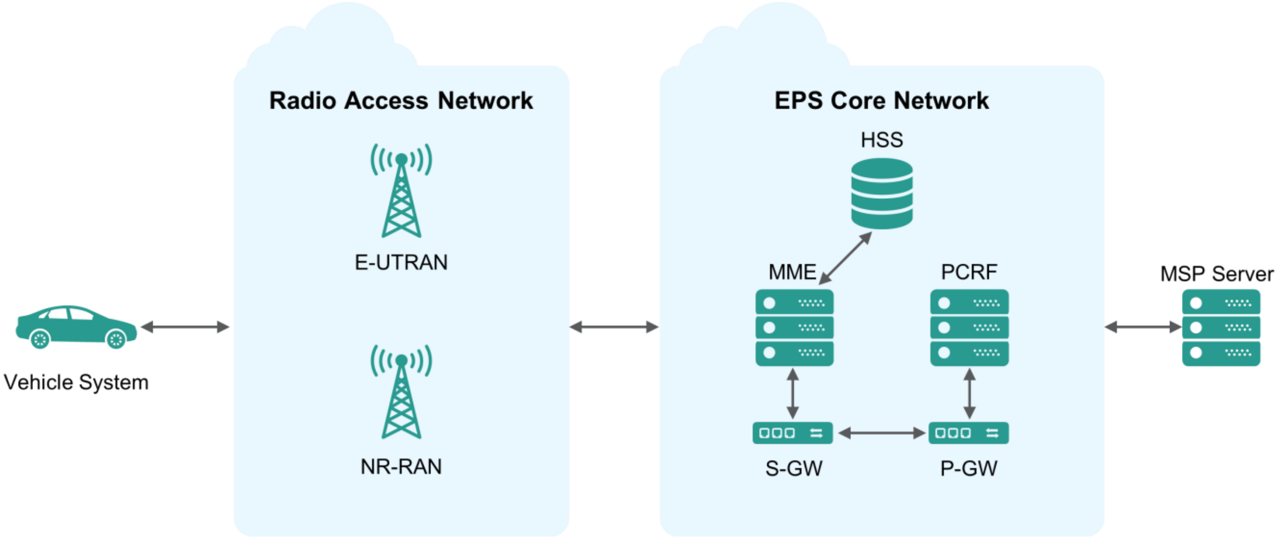

Figure 4. EPS architecture reference models with non-standalone NR radio access.

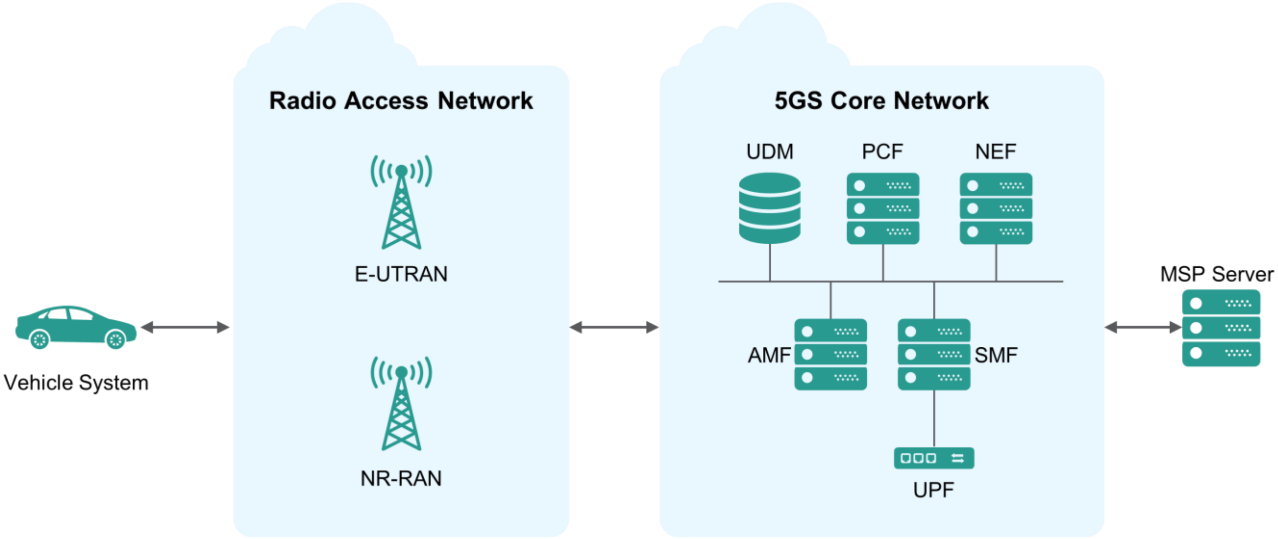

Figure 5. 5GS architecture reference models.

2.2.2 Distributed Computing Reference Model

A preliminary distributed computing reference model has been studied by the AECC and is briefly presented in this section, with further details presented in Section 4.2, “Distributed Computing: New Paradigm for Mobility Services." Further studies on this topic will be undertaken for inclusion in future technical reports.

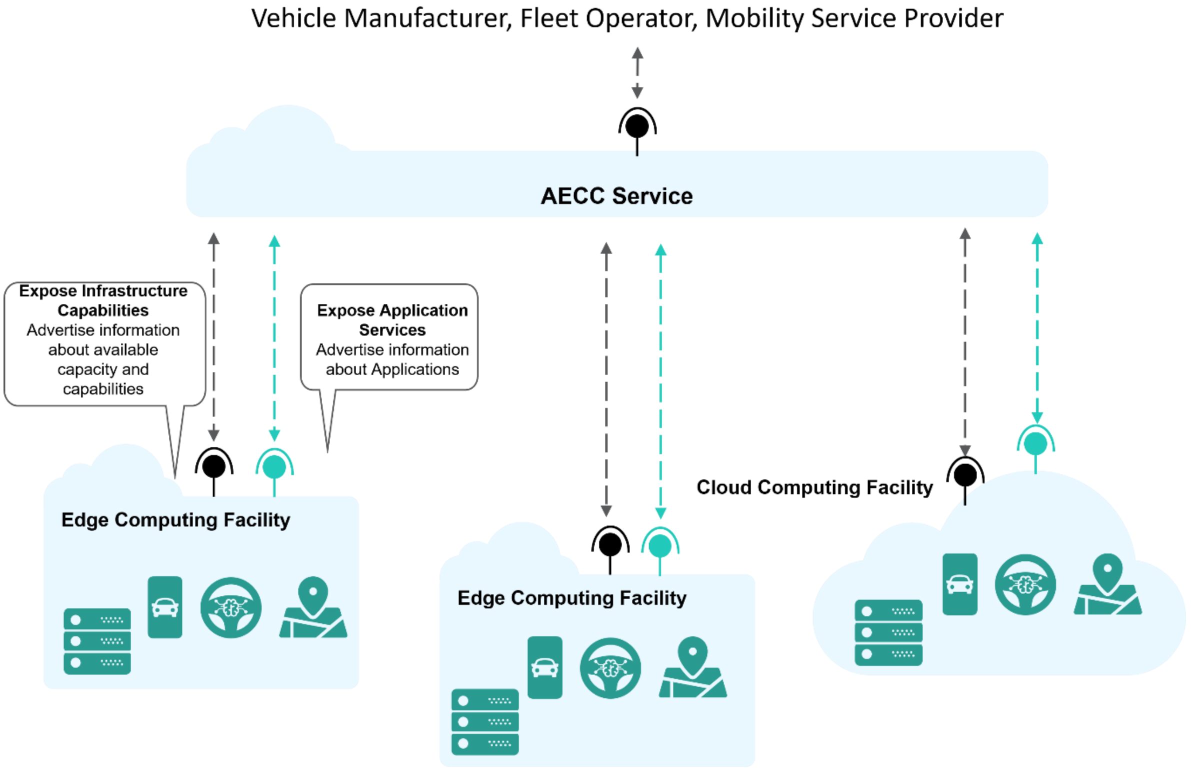
Vehicle Manufacturer, Fleet Operator, Mobility Service Provider
AECC Service
Expose Infrastructure Capabilities
Advertise information about available capacity and capabilities
Expose Application Services
Advertise information about Applications
Edge Computing Facility
Edge Computing Facility
Cloud Computing Facility

Figure 6. Infrastructure capabilities and application services.

As shown in Figure 6, in order to create a platform to support the Mobility Services, Computing Infrastructure operators with appropriate resources will expose information about their resources to an AECC Service. When applications are instantiated on the computing facilities, information will be exposed to the AECC Service as well.

The AECC Service provides a view of the available resources (computing, network, storage) as infrastructure capabilities to Mobility Service Providers, and a view of the set of applications (Intelligent Driving, HD Mapping, etc.) as application services that can be leveraged by vehicle manufacturers, fleet operators and so on.

2.2.3 Vehicle System Reference Model

The diagram below outlines the key components that will be present within a Vehicle System.

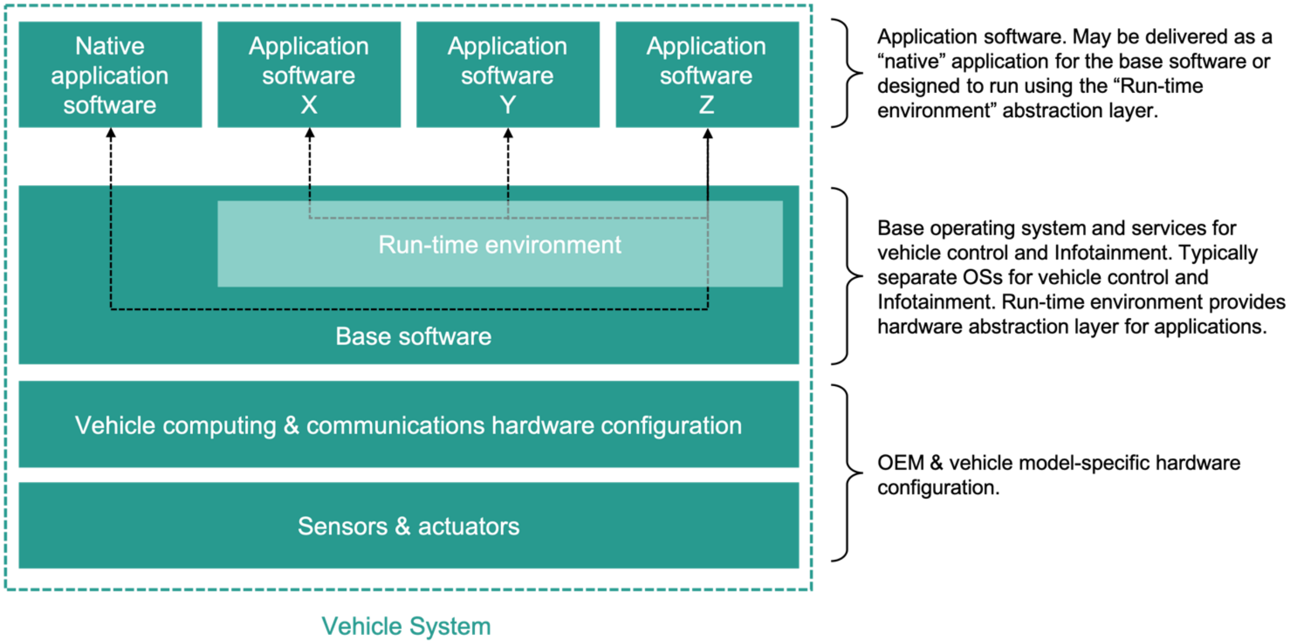
Native application software
Application software X
Application software Y
Application software Z
Run-time environment
Base software
Vehicle computing & communications hardware configuration
Sensors & actuators
Vehicle System
Application software. May be delivered as a "native" application for the base software or designed to run using the "Run-time environment" abstraction layer.
Base operating system and services for vehicle control and Infotainment. Typically separate OSs for vehicle control and Infotainment. Run-time environment provides hardware abstraction layer for applications.
OEM & vehicle model-specific hardware configuration.

Figure 7. Vehicle System components.

A Vehicle System will be composed of a specific set of hardware components including sensors, actuators, discrete electronic control units (ECUs) and functional subsystems, such as the Telecommunications Control Unit (TCU). There may be one or more computers present within the vehicle. The base software will include an operating system (or multiple operating systems) on top of which application software will be able to run. Applications will be able to make use of the functions within the Vehicle System in order to gain access to remote services hosted on infrastructure described in the AECC System architecture.

3 Key Issues and Solutions

3.1 Edge Data Offloading

3.1.1 Key Issue

In traditional cellular networks, a device connected to the network will have a fixed data offloading point in the network, through which all data to and from the device has to pass. With connected vehicles, the movement may require frequent changes of the data offloading points and the serving MSP Servers.

The AECC System architecture envisages the use of distributed computing facilities. To use these facilities in an efficient and effective manner, cellular networks must also be able to change the selected data offloading point for a Vehicle System as it moves, to provide an optimal communications path between the Vehicle System and MSP Edge Servers. Furthermore, to optimize utilization of the cellular network’s edge infrastructure, only traffic designated to MSP Edge Servers should be offloaded at such edge data offloading points, while communication intended for other destinations, such as the MSP Center Server, should use more appropriate data offloading points, if available.

By using Edge Data Offloading, cellular network operators will be able to take data communications traffic off the cellular network sooner, relieving potential congestion. In addition, Edge Data Offloading is key to being able to direct data communications to local application instances, rather than to centralized instances.

In order to alleviate the impact of high-volume data transfers on cellular networks as identified in the AECC white paper [7], the Cellular Network (both EPS and 5GS), shall support data offloading to the designated MSP Edge Servers, as shown in Figure 8. The MSP Edge Servers are connected to the MSP Center Server via the Enterprise Network as defined in the AECC distributed computing architecture reference model (Section 4.2.1). In a Cellular Network, all traffic must enter and leave the network at specific data offloading points. According to the deployment of MSP Edge Server instances, these data offloading points shall be placed at appropriate locations in the Cellular Network to meet the service requirements on latency and capacity.

Note 1: the traffic flows of different services may selectively offload to different MSP Edge Servers to meet the various requirements of service use cases.

Note 2: the case of data offloading when using different Access Networks, such as WLANs, is discussed in Section 3.4, Access Network Selection.

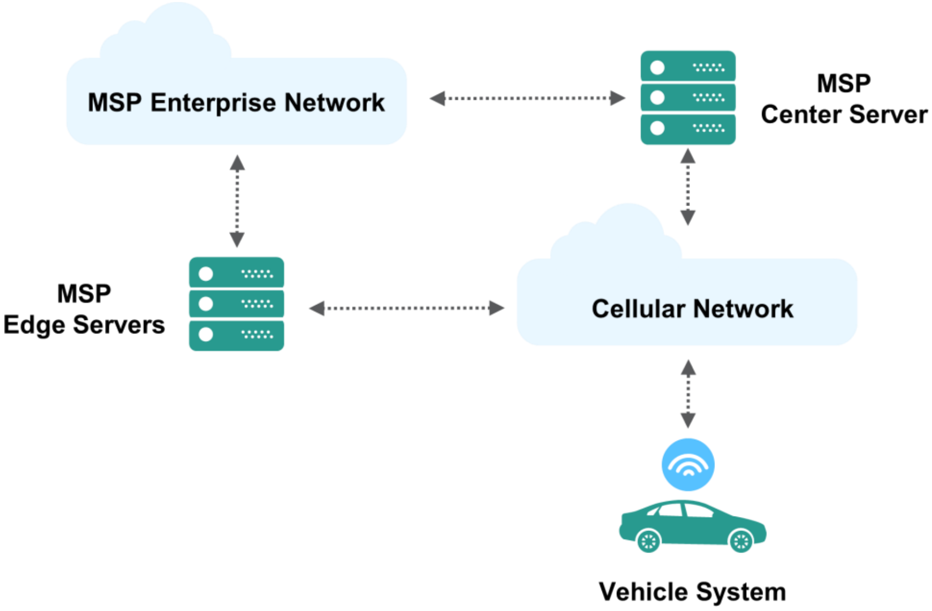

Figure 8. Connectivity between the Vehicle System and MSP Servers can be provided over a Cellular Network, in which case appropriate data offloading points must be selected in the Cellular Network.

3.1.2 Potential Solutions

The following solutions are described for this key issue.

- Solution 1 – Data Offload with Single PDN Connection in EPS
- Solution 2 – Data Offload with Multiple PDN Connections in EPS
- Solution 3 – S1/N3 GTP Packet Filtering in EPS and 5GS
- Solution 3 – S1/N3 GTP Packet Filtering in EPS and 5GS
- Solution 4 – Data Offload with a single PDU Session in 5GS
- Solution 5 – Data Offload with multiple PDU Sessions in 5GS
- Solution 6 – Uplink Classifier
- Solution 7 – IPv6 Multi-Homing

The first section below introduces the general problem statement and the applicability of different solutions to different situations. Afterwards, all solutions are defined one by one.

3.1.2.1 Solutions Overview

Edge Data Offloading in the Evolved Packet System

In 4G, or more specifically in the Evolved Packet System (EPS), PDN-Gateways (P-GWs) act as data offloading points of the Cellular Network. The Control and User Plane Separation (CUPS) feature provides more flexibility when it comes to data offloading and splits the P-GW functionality into a control plane entity (P-GW-C) and a user plane entity (P-GW-U). Once a PDN Connection from User Equipment (the UE being the communications module within the Vehicle System) to such a P-GW-U is established, it cannot be re-anchored to a different P-GW. Provided the PDN Connection persists, all traffic will be offloaded to the initial P-GW-U (see Figure 9).

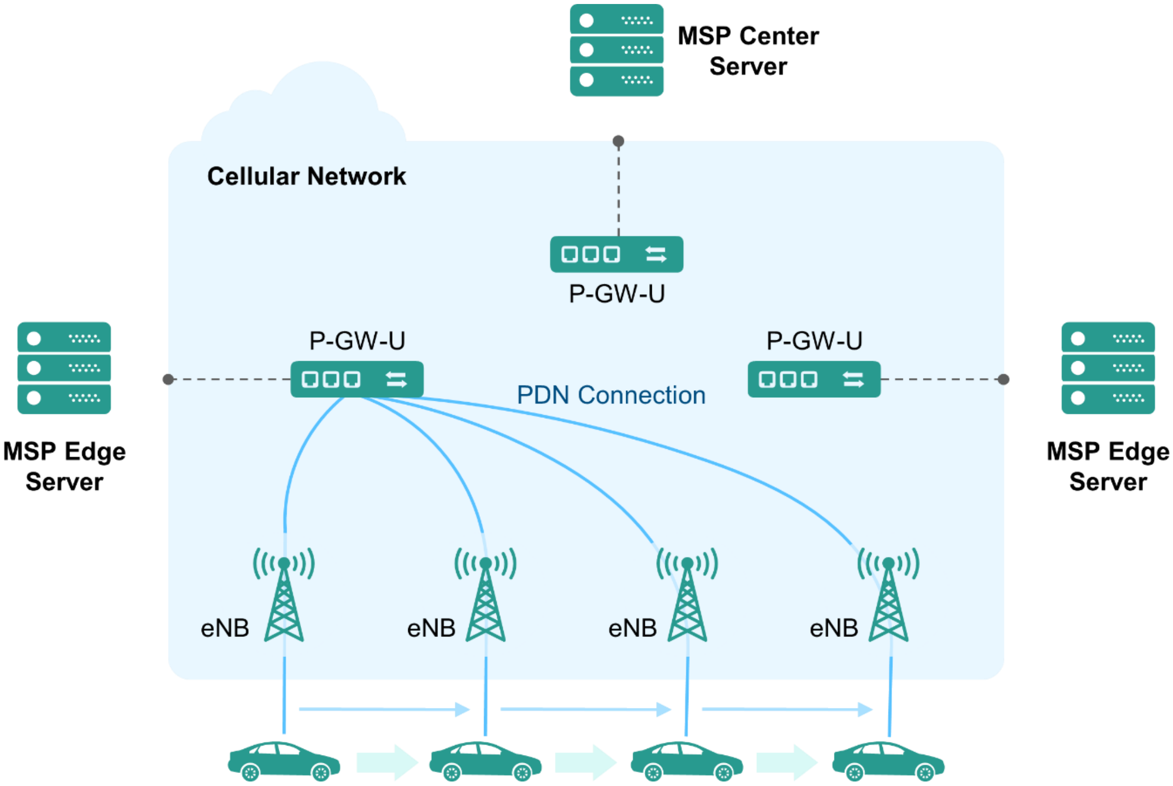

Figure 9. Vehicle moving with a persisting PDN Connection in EPS.

For re-initiating the P-GW-U selection procedure, the PDN Connection must be reactivated or a new PDN Connection must be established. The latter is done during a re-attach procedure, which automatically happens when connectivity is re-established after it was lost, such as due to large radio coverage white spots. During the attach procedure, the MME will then select an appropriate P-GW-U for the PDN Connection, based on the operator’s configuration, such as tracking area.

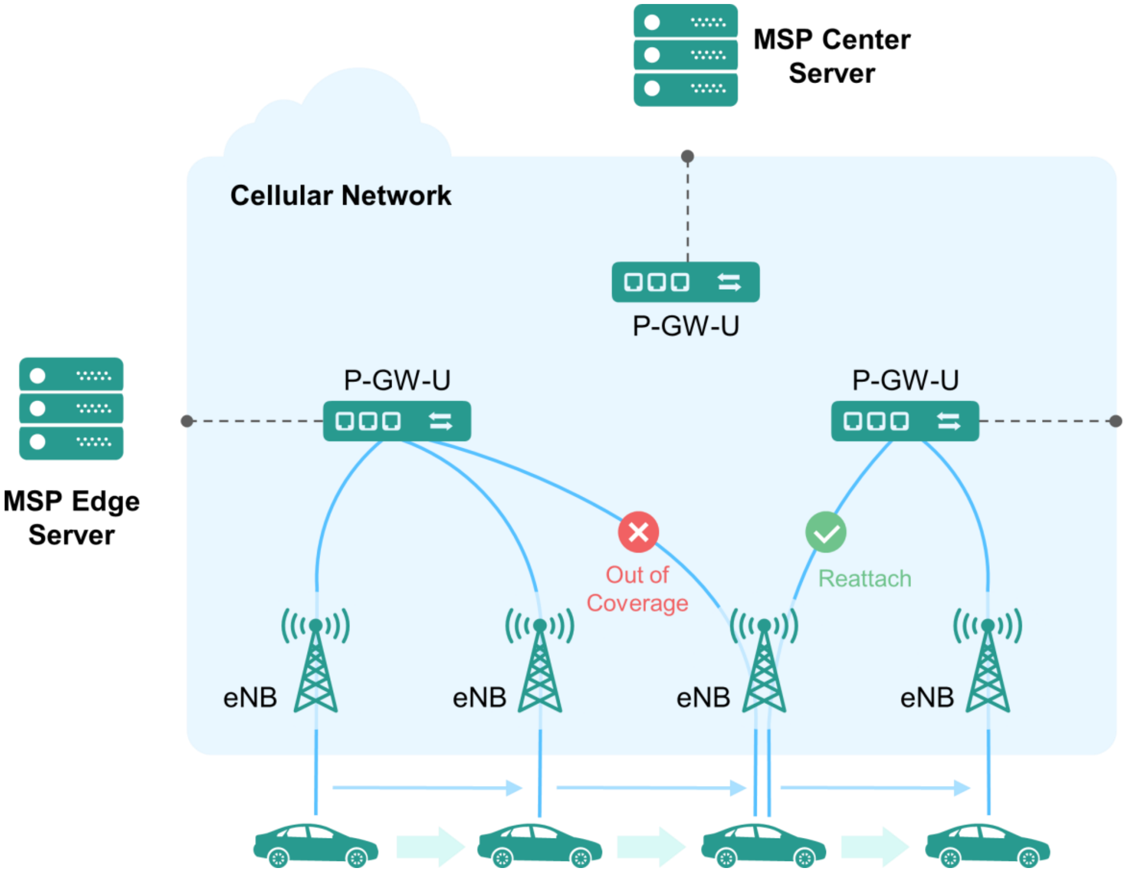

Figure 10. Vehicle moving with a changing PDN Connections in EPS.

When running into a large coverage white spot, the PDN Connection is terminated and a new PDN Connection is created when returning to coverage. This assures proper selection of an appropriate P-GW-U. If a Cellular Network has several white spots between areas that should be served by different P-GW-Us,[1] movement between areas served by different P-GW-Us might not be an issue, as the P-GW-U is automatically reselected on a regular basis. However, with many P-GWs deployed and more continuous coverage, a PDN Connection reactivation must be triggered by the system. For this purpose, EPS provides an optional feature ("Selective IP Traffic Offload [SIPTO] above RAN") where the MME requests a PDN Connection deactivation with a subsequent reactivation to the UE (see Figure 11).

[1] Note: this is a typical case for spotty coverage in rural areas with a low degree of P-GW-U distribution.

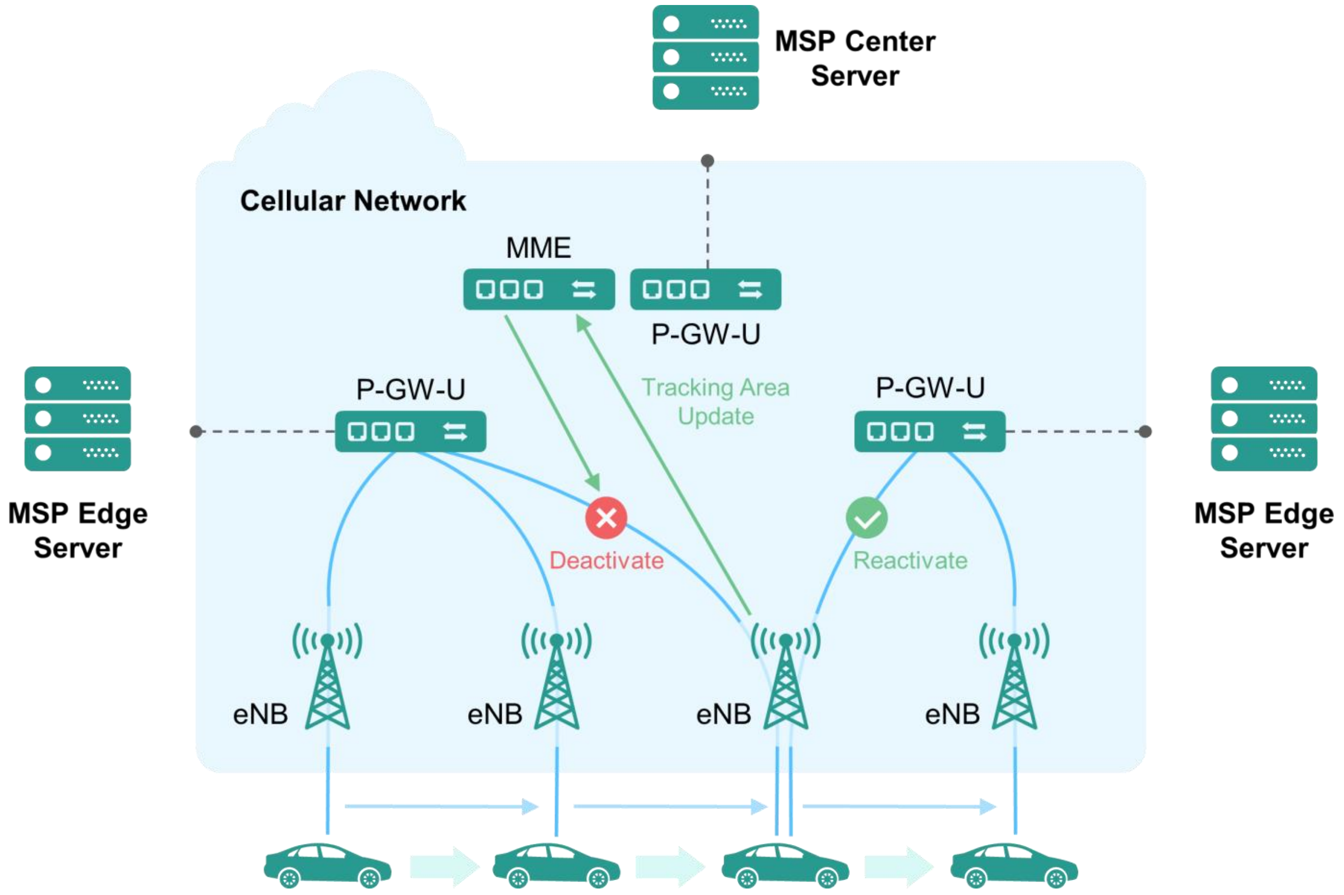

Figure 11. Reactivation of the PDN Connection with SIPTO.

When using only a single PDN Connection per UE, data exchanges between a Vehicle System and an MSP center site will go via the current edge P-GW-U (see Figure 12, right car). This may require over-dimensioning the edge P-GW-U if the share of data is relevant and geographically varying. In that case, the load on the edge P-GW-Us can be reduced by using a second PDN Connection on the same UE, with a different APN that is configured for anchoring at a central P-GW-U (see Figure 12, left car). In that case, two IP interfaces must be managed on the Vehicle System, one per APN.

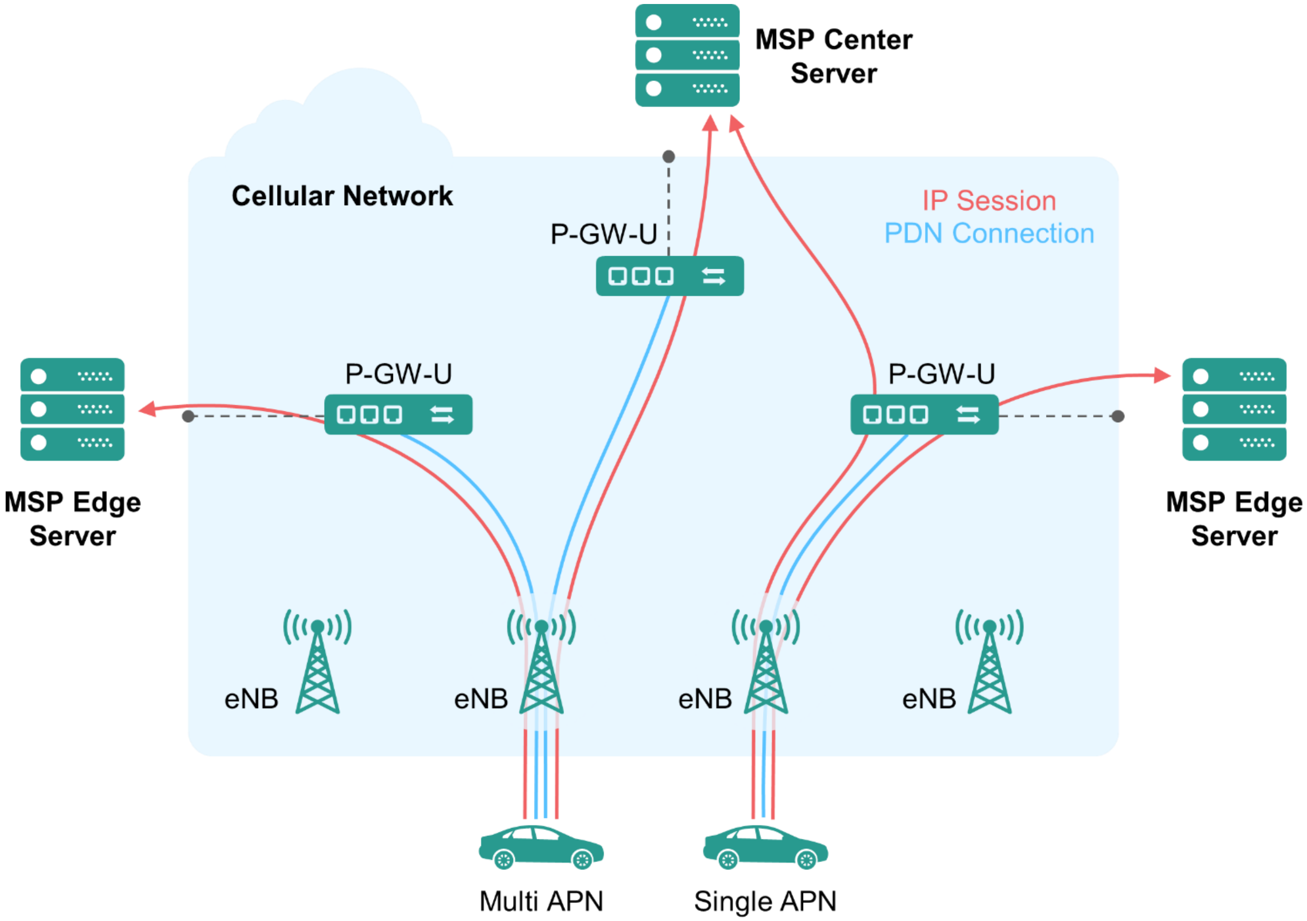

Figure 12. Offloading edge GWs with a second PDN Connection.

Edge Data Offloading in the 5G System

The 5G System (5GS) has the same paradigm of an existing central anchor point (the PDU Session Anchor in the User Plane Function [UPF]). This requires similar solutions with a single PDU Session that offloads at the edge. However, compared to the EPS, the 5GS supports multiple Session and Service Continuity (SSC) modes that control behavior during mobility between anchor points (UPFs). Depending on the SSC mode, a PDU Session will persist until running out of coverage (SSC mode 1), a PDU Session may be released and immediately re-established to a new UPF (SSC mode 2, SIPTO-like behavior) or a new PDU Session may be established before releasing the old PDU Session (SSC mode 3). Just as in EPS, an additional PDU Session, anchored at a central UPF, can be used to reduce the load on edge UPFs.

While the features described above enable a Vehicle System to always have connectivity via appropriate data offloading points, ongoing sessions are interrupted during re-attachment procedures, meaning that data communication has to be re-established, resulting in a temporary interruption in service. The 5GS also offers mechanisms to maintain connectivity while re-anchoring the PDU Session at a different UPF. For example, an uplink classifier (ULCL) policy can be provisioned in a UPF to offload the selected traffic to an Edge Server, and the SMF can dynamically insert and remove an uplink classifier into the data path of the PDU Session (see Figure 13). Typically, IP 5-tuples are used in such policies to decide which packets to offload to which edge site.

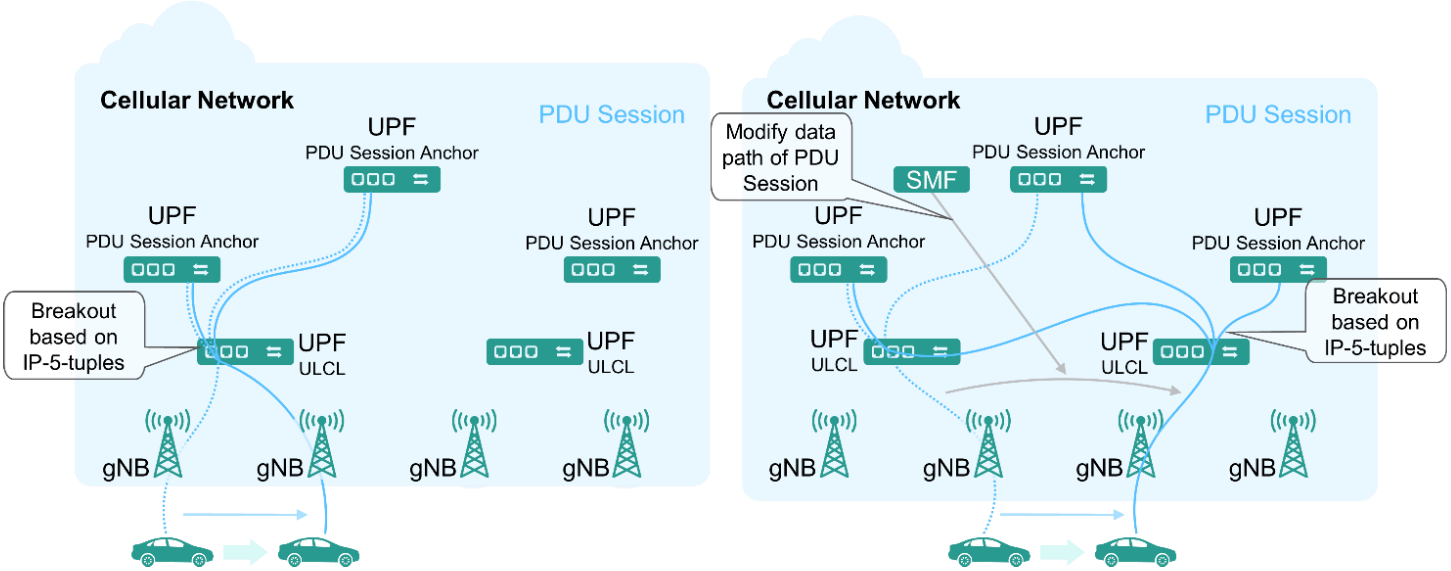

Figure 13. Selectively routing packets with uplink classifiers.

In this approach, while there are multiple PDU Session anchors, there is only one IP anchor. Even when moving the PDU Session to a different PDU Session Anchor (UPF), the IP session is maintained while traffic to the old uplink classifier UPF is tunneled. This communication link is deactivated when not used anymore; e.g., when using timeout procedures (see Figure 14). For downlink traffic, the Vehicle System is reachable using the same IP via all UPFs, which must be taken into account when considering IP Routing and path selection between the MSP Servers, the intermediate IP network(s) and the UPFs themselves.

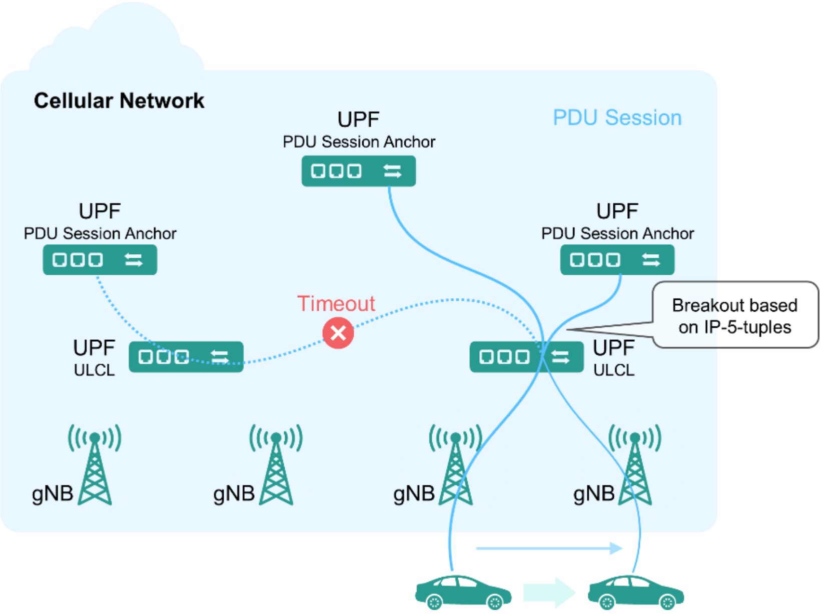

Figure 14. Terminating an unused link with timeout procedures.

3.1.2.2 Solution 1 – Data Offload with a Single PDN Connection in EPS

In EPS, SIPTO above RAN enables dynamic reselection of GWs (S-/P-GW), and selection of GW-Us in the case of Control and User Plane Separation (CUPS). Those GWs or GW-Us are geographically/topologically close to the UE. The selection mechanism can consider UE location in the network (Tracking Area), APN or other parameters.

The UE does not need to be aware of whether the PDN Connection corresponds to the MSP Edge or Center server. One AECC-dedicated APN can be provisioned to the UE for all AECC traffic flows – including traffic with and without offloading. When the UE uses this AECC APN, it does not need to know whether or not traffic on this APN will be offloaded. The network (MME) will choose appropriate GWs for AECC APN traffic purely based on the MNO’s configuration; e.g., based on defined groups of cells (tracking area). The GW (S-/P-GW or S-/P-GW-U) selection will be based on information such as SIPTO permission information per subscription per APN, UE location information and so on. The MME can also decide to move a PDN Connection from one GW to another (e.g., from a GW serving the MSP Edge Server to a GW serving the MSP Center Server) for AECC APN if needed.

The different types of AECC traffic flow will be offloaded to the applications on the MSP Edge Server or pass through to the MSP Center Server as shown in Figure 15.

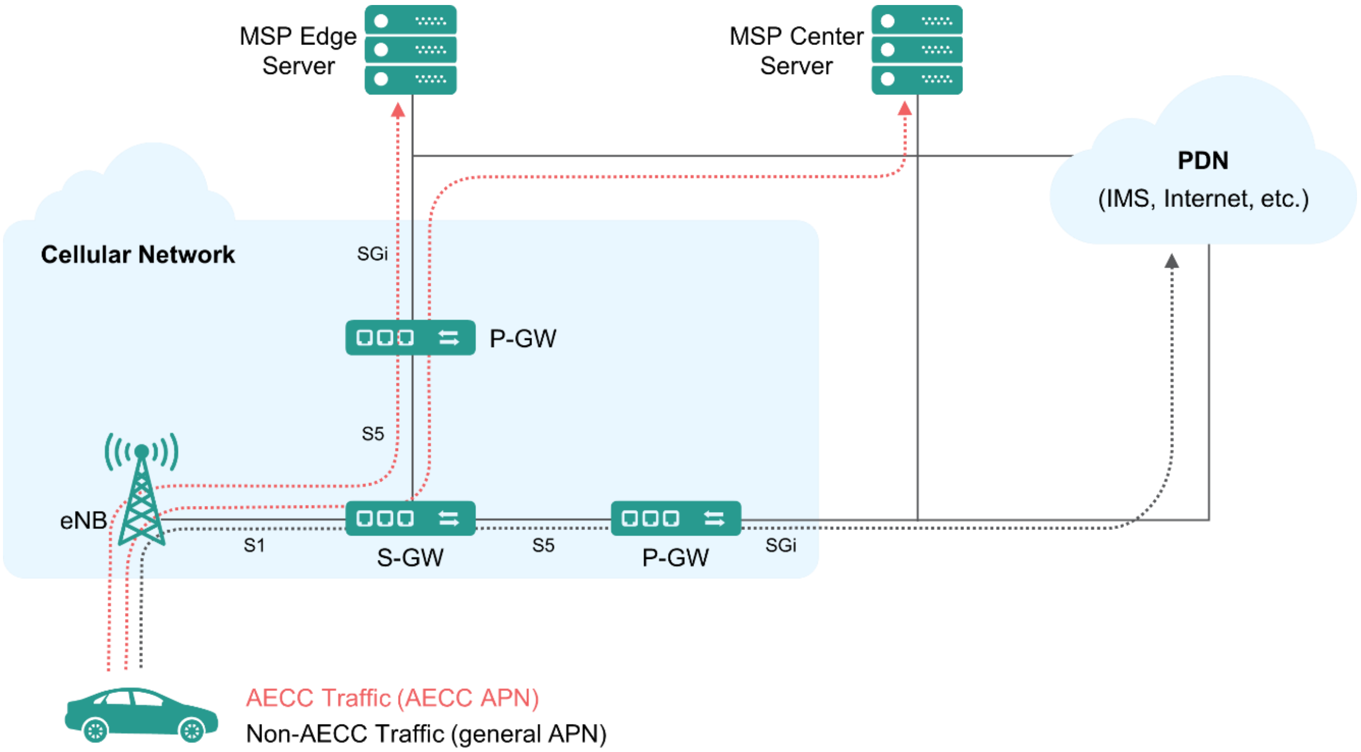

Figure 15. Data offload with a single PDN Connection for AECC-related traffic (red) and other traffic (black) in EPS.

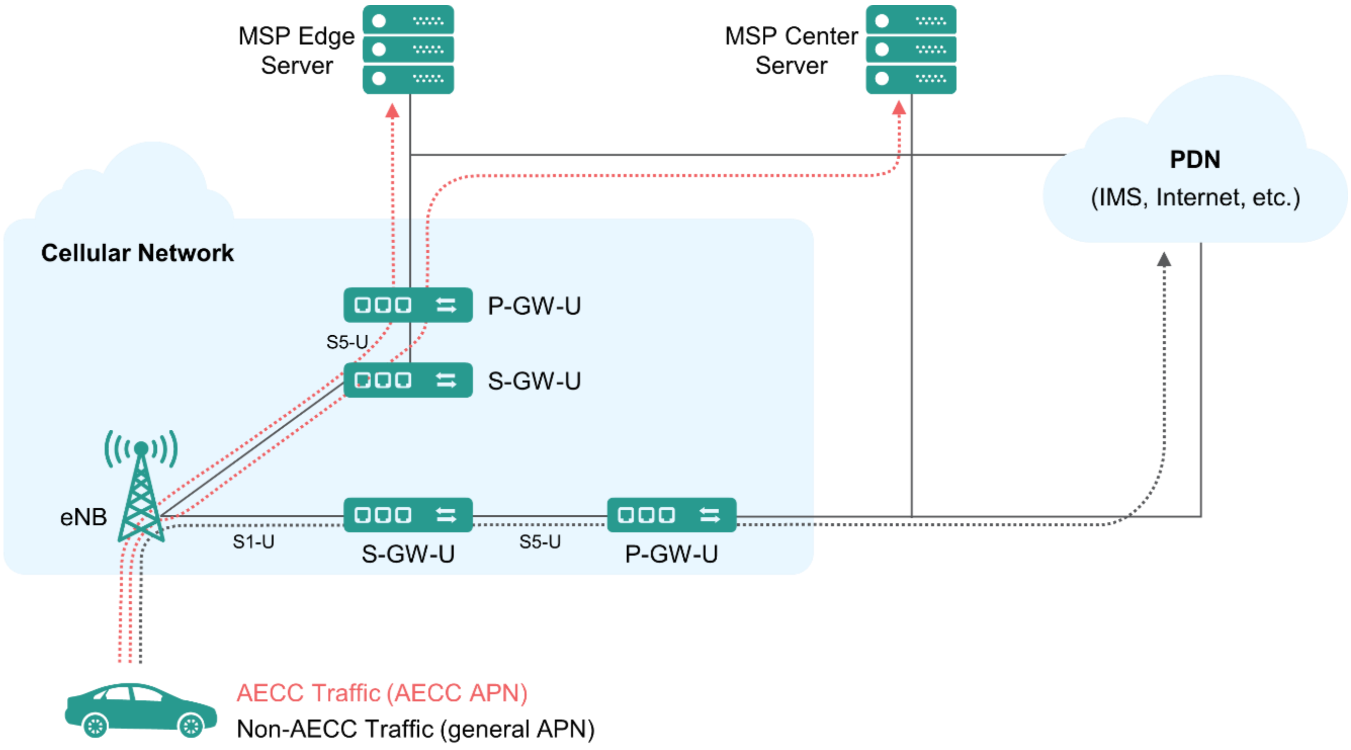

Figure 16. Data offload with a single PDN Connection for AECC-related traffic in EPS in the case where CUPS is deployed.

3.1.2.3 Solution 2 – Data Offload with Multiple PDN Connections in EPS

SIPTO above RAN enables the Cellular Network to offload data traffic sent through a PDN Connection for an APN configured for SIPTO to a designated MSP Edge Server via a selection of P-GW (or P-GW-U in the case of CUPS). It is a function introduced and standardized since 3GPP LTE Release 10. SIPTO requires a dedicated APN for offloading the selected data traffic to the Edge Server, so it needs to support multiple APNs for different PDN Connections at the UE side to achieve a selective data offload. Consequently, the Vehicle System needs to implement support for multiple outbound IP interfaces (each corresponding to a different PDN Connection), and corresponding routing functionality (which includes static routing or “hardwiring”) in order to place the IP traffic onto the correct interface.

Due to movement of the UE in the network, the serving MME may need to redirect a PDN Connection to a different P-GW that is more appropriate for the location of the UE, based on the tracking area of the Vehicle System. In this case, this solution cannot maintain session continuity while changing to the new P-GW.

The GW selection is configured by the MNO through tracking area configuration and mapping to GWs; that is, no direct control to external entities is provided in the current solution. For selecting which PDN Connection to use for a packet in the uplink, one can either push the selection process to the application layer in the UE or use UL-TFTs for doing an automatic mapping based on IP 5-tuples. For the downlink, one can select a PDN Connection by using the respective IP address assigned by the corresponding GW.

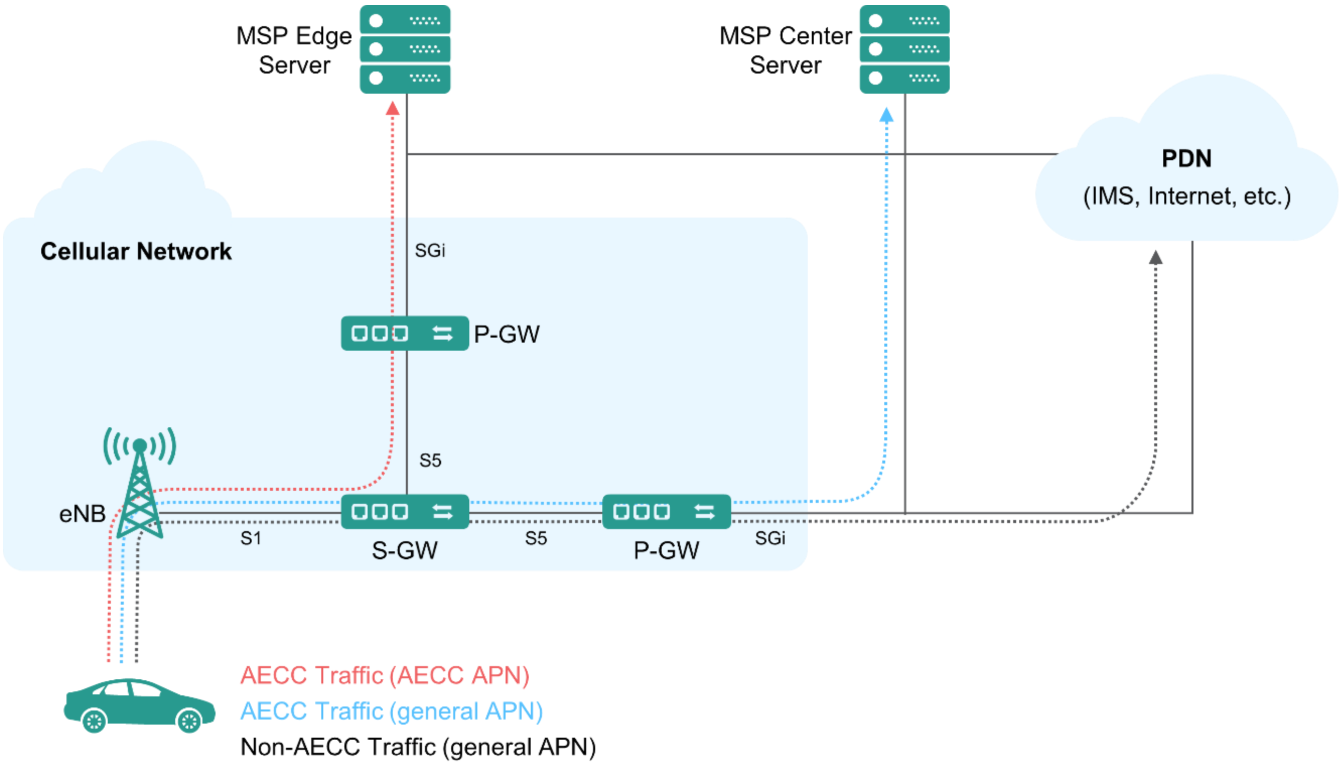

Figure 17. Data offload with multiple PDN Connections for AECC-related traffic (red and blue) and other traffic (black) in EPS.

3.1.2.4 Solution 3 – S1/N3 GTP Packet Filtering in EPS and 5GS

Edge data offloading can be conducted based on S1 GTP packet filtering mechanisms. The whole solution should include:

- Packet filtering, also called a traffic filter policy
- Packet filtering management

To support edge offloading in the Cellular Network, a traffic filter solution to handle the traffic offloading of the data plane can be used. For the packet filtering policy, a set of traffic filter policies composed of traffic rules and traffic filters corresponding to the rules is shown in Figure 18.

For example, such a TrafficFilter policy could use IP header information (IP Address, Port, L4 protocol) or consider L4/L7 parameters when feasible. If the edge offloading resides on the S1 interface inside the Cellular Network, it can also support filtering based on GTP tunnel information such as GTP-U TEID and so on. The policy also includes actions such as forward, drop, passthrough and duplicate.

For the packet filter policy (TrafficFilter policy) management, the system uses an architecture solution to address policy management for edge offloading inside 4G or 5G Cellular Networks. The AECC System can generate traffic rules in order to exert influence over the data plane within the packet filter in the Cellular Network.

The solution for the packet filter is shown in Figure 18, where packet filtering must be implemented on a device on the S1 interface between eNBs/gNBs and the core network.

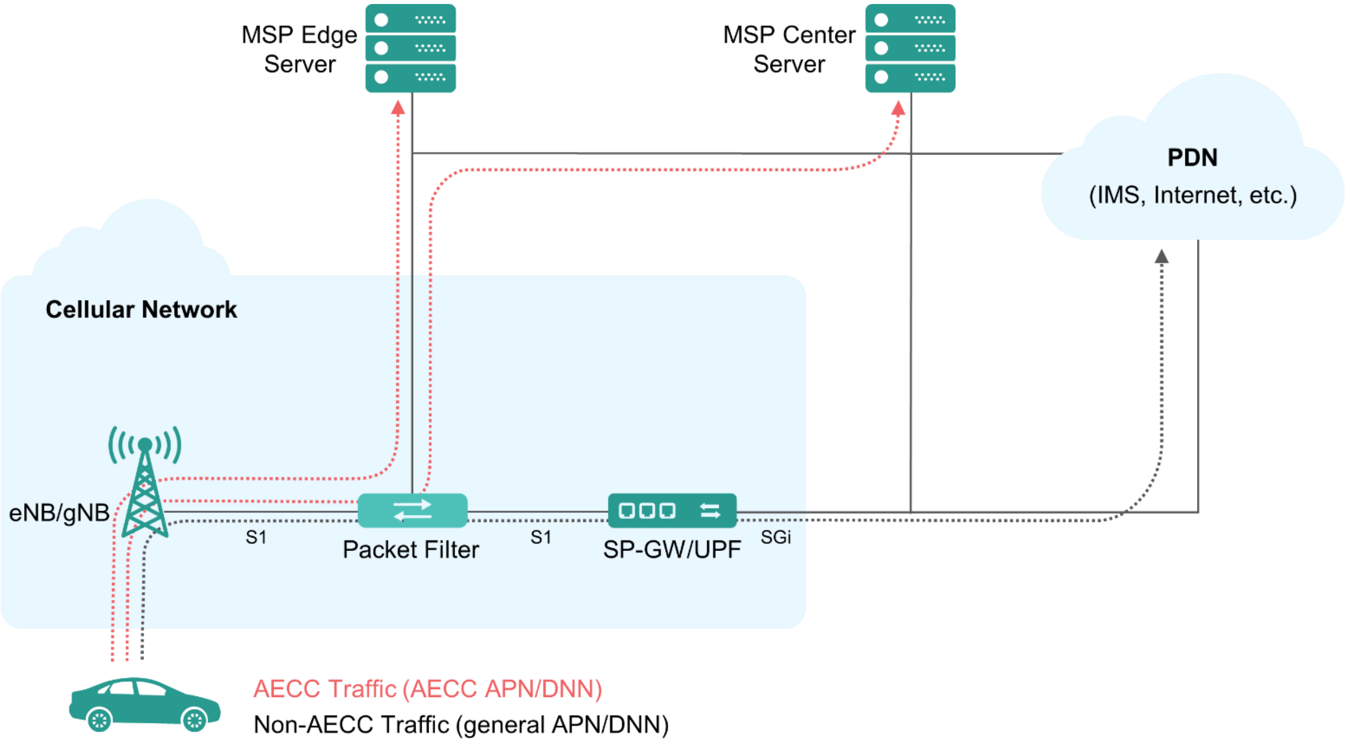

Figure 18. GTP packet filter implementing traffic filter policy.

In Figure 18, a packet filter policy based on port numbers, for example, can be used to apply different edge data offloading behaviors.

The key issue is solved in this case by intercepting S1-U (in EPS) or N3 (in 5GS) traffic and communicating directly with the MSP Servers from the intercepting entity.

This solution is applicable to both EPS and 5GS.

3.1.2.5 Solution 4 – Data Offload with a Single PDU Session in a 5GS

In a 5GS [7], the SMF is in charge of selecting or reselecting a UPF for a PDU Session and can consider a number of parameters for the selection process. Among these is the tracking area identifier, which allows for a cell-specific UPF selection. In this approach, the selection process is used to offload all traffic to an edge UPF. Likewise, downlink traffic would always pass through this UPF. Different traffic flows might still be offloaded to different MSP Servers but would use the same breakout/UPF.

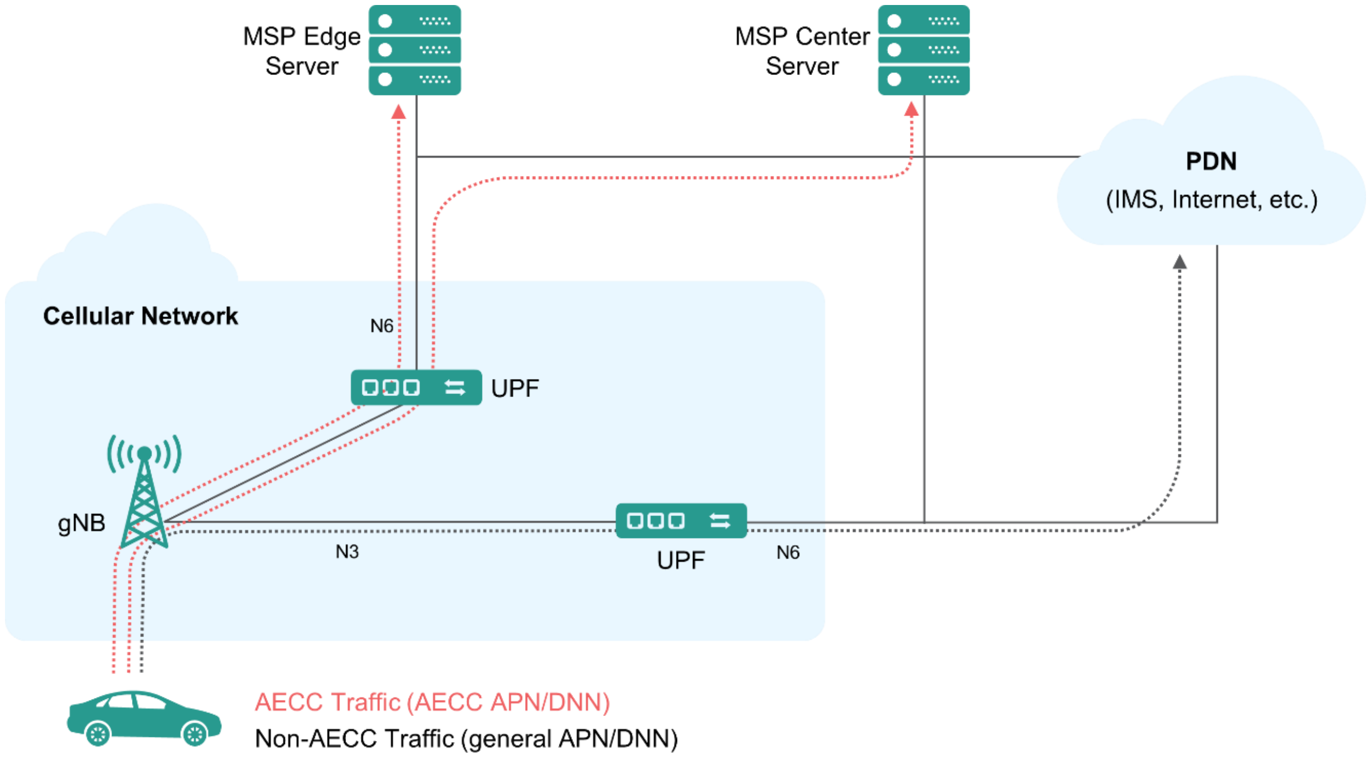

Figure 19. Data offload with a single PDU Session for AECC-related traffic (red) and other traffic (black) in a 5GS.

3.1.2.6 Solution 5 – Data Offload with Multiple PDU Sessions in a 5GS

The UE creates multiple PDU Sessions using the same procedure as for the edge breakout with a single PDU Session. While one PDU Session offloads traffic to edge UPFs, the other PDU Session uses a central UPF. For downlink traffic, the two external IP addresses of the respective PDU Sessions are used to select the corresponding UPF. Again, the Vehicle System needs to implement support for multiple IP interfaces (each corresponding to a different PDU Session) and corresponding routing functionality.

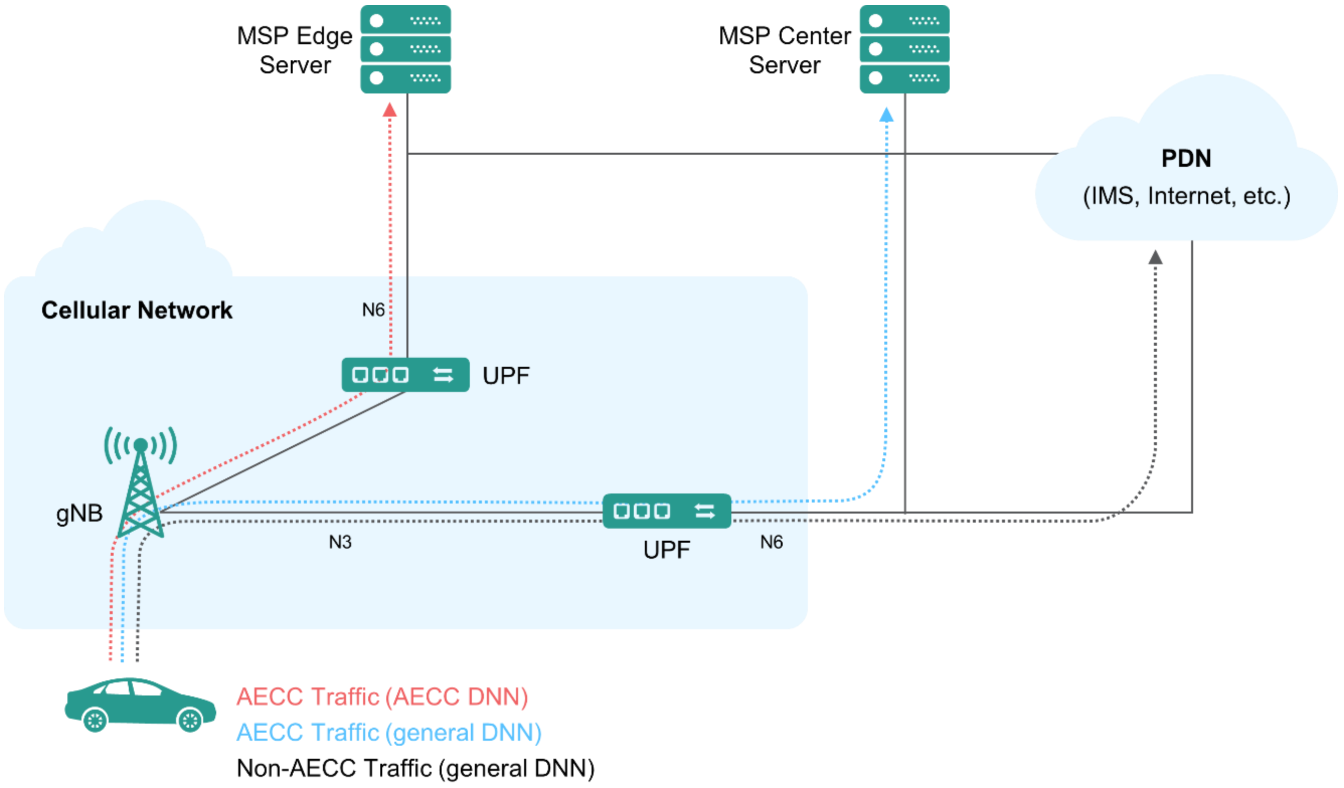

Figure 20. Data offload with multiple PDU Sessions for AECC-related traffic (red and blue) and other traffic (black) in a 5GS.

3.1.2.7 Solution 6 – Uplink Classifier

In a 5GS, an uplink classifier policy can be provisioned in a UPF to offload the selected traffic to an Edge Server. The insertion and removal of an uplink classifier policy is controlled by the SMF. The SMF may include multiple UPFs with uplink classifier policies in the traffic data path and may modify this UPF chain dynamically. This solution is only supported in a 5GS, and IP 5-tuples are used as traffic filters in the UPFs that, when matched, trigger local offload of the respective traffic.

In this approach, while there are multiple PDU Session anchors, there is only one IP anchor; that is, the IP address of the UE is assigned by only one UPF and preserved during the lifetime of the PDU Session. Even when the PDU Session anchor changes (e.g., due to movement of the vehicle system), the IP session is maintained, while traffic to the old uplink classifier UPF is tunneled. For downlink traffic, the Vehicle System is reachable using the same IP via all UPFs, which must be considered in the IP configuration of the MSP Servers and the corresponding IP network(s).

In order to forward data to the appropriate PDU Session anchors, uplink classifiers must be configured accordingly, based on information on IP subnets and location of MSP Servers, in order to know which PDU Session anchor is most appropriate for the UE in a given cell, and the IP subnet with which it communicates. This information can be configured and updated manually, or it can be dynamically exposed to an MNO by an MSP. Typically, defining how such information is exchanged is described as part of an SLA.

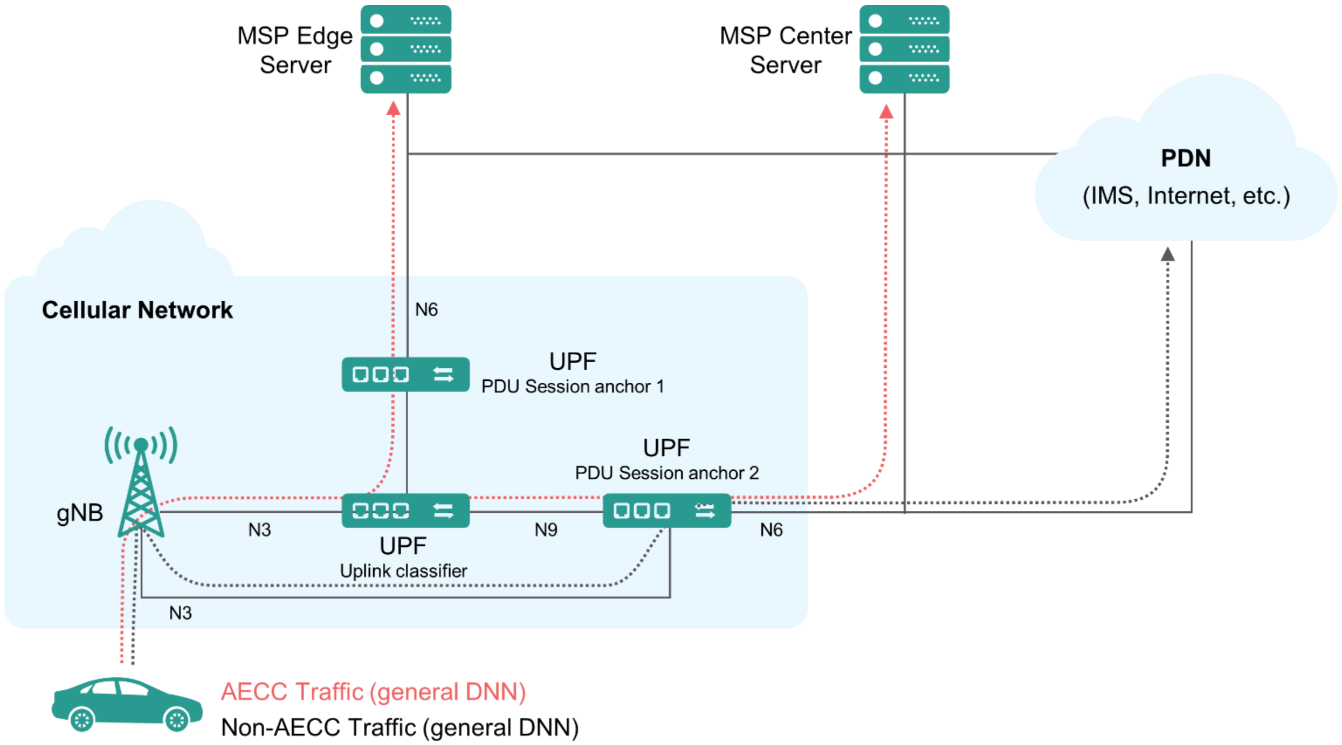

Figure 21. Data offload with a single PDU Session for AECC-related traffic (red) and other traffic (black) in a 5GS, using uplink classifiers to selectively offload traffic based on traffic filters (usually IP 5-tuples).

3.1.2.8 Solution 7 – IPv6 Multi-homing

In a 5GS, the PDU Session from a Vehicle System may be associated with multiple IPv6 prefixes. Selected traffic can be offloaded to the designated Edge Server as configured by the SMF, using a specific IPv6 prefix. In the traffic data path, the common UPF acts as a branching point, where the uplink traffic is split to different destinations and downlink traffic is merged to the Vehicle System. The UE selects the source IPv6 prefix according to rules pre-configured in the UE or received from the network.

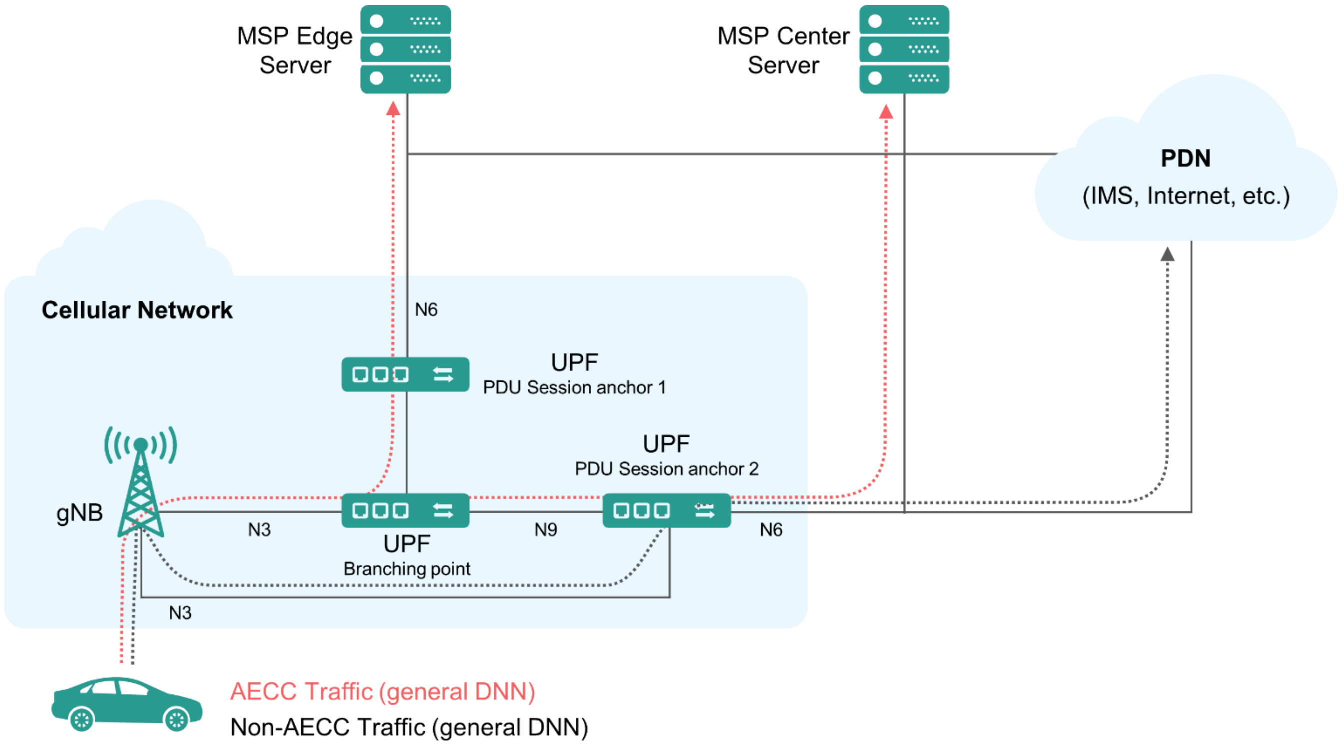

Figure 22. Data offload with a single PDU Session for AECC-related traffic (red) and other traffic (black) in a 5GS, using different IPv6 prefixes to selectively offload traffic.

3.1.3 Conclusions

For a 5GS, the recommended first choice of the AECC is Solution 6 (Uplink Classifier). It leaves IP session termination during anchor point changes to the applications, maintains IP connectivity via the central breakout point irrespective of the current edge breakout configuration, avoids complexity in the Vehicle System and allows for the deployment of edge breakout capabilities with reduced capacity by comparison to solutions where all data must pass through the particular edge breakout function. On the other hand, as there is only a single IP anchor on the network side, IP-addressing has increased complexity and needs additional consideration for downlink traffic. This solution involves optional components (specifically, uplink classifier functionality in UPFs deployed at the edge) of the 5G architecture, so it might not be deployed with global coverage.

As a fallback solution, Solution 4 (Data Offload with a Single PDU Session) is recommended. This solution is based on distributed anchors for the AECC PDU session with re-anchoring as needed. Therefore, it consists only of mandatory 5GS features while limiting complexity in the Vehicle System. The feasibility of this solution depends on the mobile network deployment and the degree of cloud distribution. For a low degree of distribution, this solution may in fact be an excellent choice due to its low complexity. However, inefficiency for this solution rises with higher degrees of distribution combined with large data volumes to or from the MSP Center Server.

Solution 6 and Solution 4 require the same functionality in the Vehicle System, which is why no MNO-specific functionality needs to be applied. One difference in behavior is that Solution 4, in the most basic configuration (SSC mode 1), forcefully terminates PDU Sessions when changing the anchor point, while Solution 6 keeps PDU Sessions up even during handover.

Finally, as Vehicle System capabilities evolve, Solution 5 (Data Offload with Multiple PDU Sessions) is more feasible and can then be applied as a complement to Solution 4 or Solution 6.

For EPS, the AECC recommendation is Solution 1 (SIPTO with a Single PDN Connection) as the first choice for keeping Vehicle System complexity low. Furthermore, SIPTO ensures connectivity to the most appropriate offloading point, by reacting to movement in the network. However, as the PDN Connection is terminated and re-established during such movement, all IP connectivity (including connectivity to the MSP Center Server) is unavailable for a short time.

Solution 2 (SIPTO with Multiple PDN Connections) is a possible enhancement for reducing the load on the edge breakouts and maintaining connectivity to the MSP Center Server at all times while the Vehicle System is in coverage. However, the Vehicle System needs to manage two separate PDN Connections with different APNs in this case, as opposed to Solution 1.

3.2 MSP Server Selection

3.2.1 Key Issue

The AECC System is expected to support the execution of software applications that will be used by Vehicle Systems of different types and from differing manufacturers. The working assumption is that applications will be delivered utilizing IPv4 or IPv6-based communications protocols and that in keeping with today’s modern cloud deployment platforms, a dynamic mechanism will be required that will be able to inform the Vehicle System of the resources that are available to it and then direct the Vehicle System’s application software to use the most appropriate application server instance. The selection function, hereafter referred to as the MSP Server selection service, forms part of the overall set of services provided by the AECC System, enabling the exchange of data between applications executing in the Vehicle Systems and MSP Servers.

The working assumption adopted by the AECC is that where there are multiple concurrent applications in use within a Vehicle System, the Vehicle System may connect to multiple MSP Servers as shown in Figure 23, since different applications may be hosted on different servers. For example, in this figure, Vehicle System A connects to the MSP Center Server, MSP Edge Server 1 and MSP Edge Server 2A. Vehicle System B connects with MSP Edge Server 2A, 2B and the MSP Center Server.

The objective of the MSP Server selection service will vary, depending on each service scenario. Information such as vehicle geolocation, Access Network topology, server load, network performance and policy may be contributed as part of the selection process. The function that the MSP Server selection service performs is to collect, process and distribute information about the available MSP Servers, enabling the applications within the Vehicle System to connect to the most appropriate application server instance.

An AECC System may have a highly dynamic network topology. For this reason, the use of names (such as FQDN) is more flexible than endpoints using IP addressing. A naming scheme is therefore necessary as part of an overall resource scheme within the AECC System that the MSP Server selection service can then utilize.

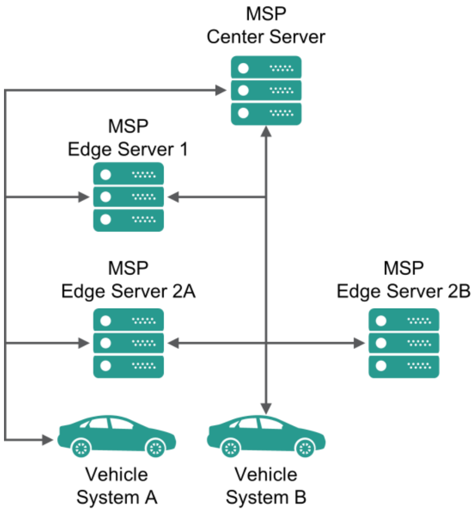

Figure 23. Topology of hierarchical distributed computing, with lines and arrows showing the possible data exchange paths between all entities.

3.2.2 Potential Solutions

There are four solutions to the issue, as described in the following sections.

1. Cellular Network-based MSP Server Assignment
2. IP Network-based MSP Server Assignment
3. MSP Server Assignment by a Selection Function
4. Vehicle System-based MSP Server Assignment

The solutions are based on the entity that will select the target MSP Server. The solution mapping is shown in Figure 24.

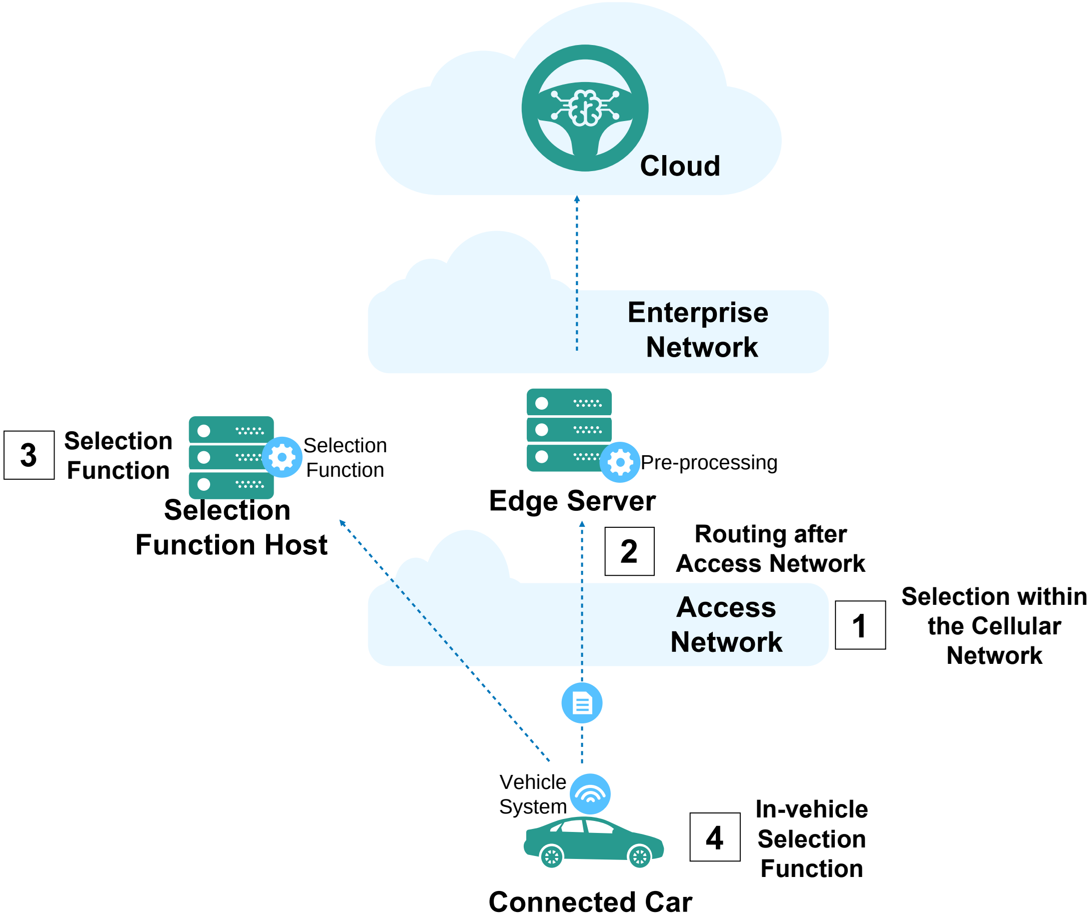

Figure 24. Mapping of potential solutions for MSP Server selection.

3.2.2.1 Solution 1 – Cellular Network-based MSP Server Assignment

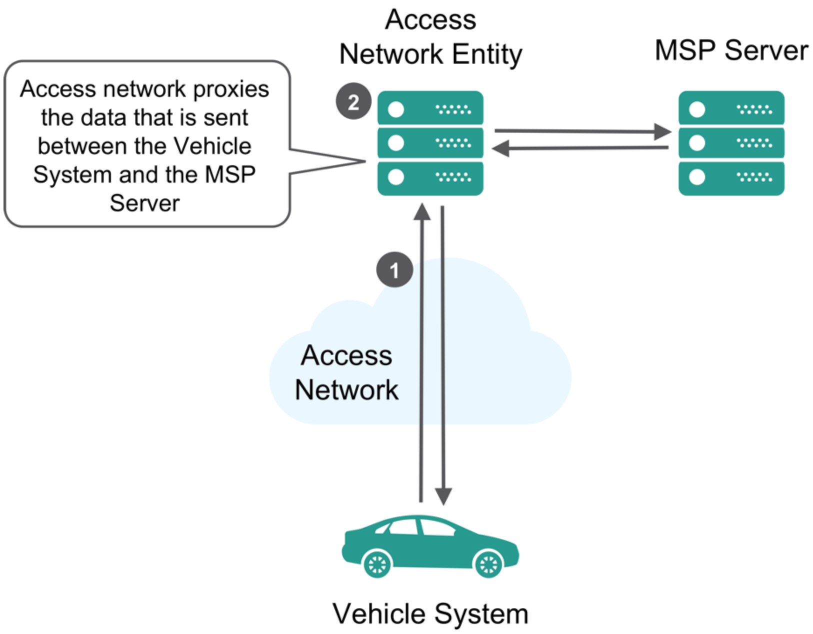
Access Network Entity
MSP Server
Access network proxies the data that is sent between the Vehicle System and the MSP Server
2
1
Access Network
Vehicle System

Figure 25. Cellular Network-based MSP Server Assignment.

The location of a vehicle that connects to the Cellular Network may be known to a certain degree by the Cellular Network. Cellular Networks also have SCEF (LTE) and NEF (5G) components that enable communication with the application instance. In this approach, the Cellular Network entity becomes a proxy and a control agent for communication between the Vehicle System and the MSP Server; for applications that are dependent on the Vehicle System's movement, the Vehicle System is oblivious to the MSP Server Assignment procedure. The approach assumes that the Cellular Network operator and the MSP have mutual agreement on how the assignment should occur.

1) The Vehicle System sends data through the Cellular Network. This enables the Cellular Network to identify the vehicle's location through the base station to which the vehicle is connected.
2) The Cellular Network entity chooses the most suitable MSP Server based on the agreement with the MSP and other criteria such as current server load. This process includes the selected hostname resolving.
3) The vehicle connects and is able to communicate with the target MSP Server.
4) Response from the MSP Server is then routed back to the Vehicle System through the Cellular Network.

3.2.2.2 Solution 2 – IP Network-based MSP Server Assignment

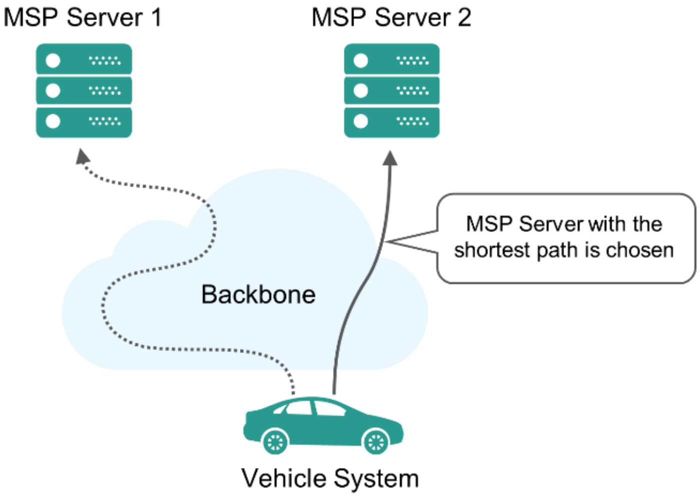

Figure 26. IP network-based MSP Server Assignment.

In this approach, both the MSP Server and the Vehicle System do not have any logic to decide upon the appropriate MSP Server. The routing scheme may leverage IP anycast, so that traffic from the Vehicle System will be forwarded by routers within the IP network to the MSP Server with the shortest path. Application instances are deployed in a distributed manner, with application instances being provisioned with predetermined IP addresses. No selection function takes place, and instead network topology-based routing to the MSP Server is used instead. In this approach, all required information is located in the application layer and not shared with the network, thus making it agnostic regarding the Access Network.

3.2.2.3 Solution 3 – MSP Server Assignment by a Selection Function

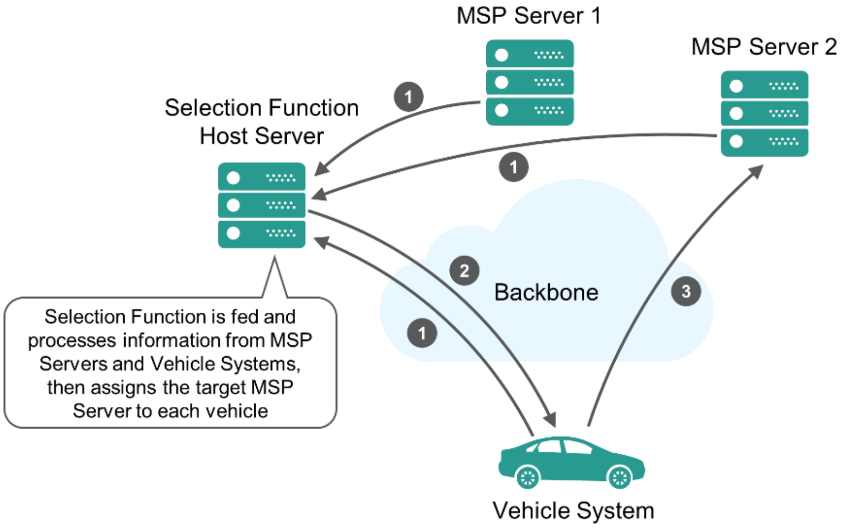
MSP Server 1
MSP Server 2
Selection Function Host Server
1
1
1
2
3
Backbone
Selection Function is fed and processes information from MSP Servers and Vehicle Systems, then assigns the target MSP Server to each vehicle
Vehicle System

Figure 27. Selection Function-based MSP Server Assignment.

In this approach, a Selection Function receives information from MSP Servers and the Vehicle System. The Selection Function then processes the information and informs the Vehicle System of which MSP Server to use, allowing the Vehicle System to initiate a session with the selected MSP Server. For example, implementation of the Selection Function could be by a DNS Server, where the algorithm runs while resolving a host name (Fig. 27), although implementation is possible independent of the DNS system. Specific configuration of the Selection Function may allow processing of information shared by Vehicle Systems and MSP Servers, including but not limited to geolocation and/or server health check. This approach is agnostic with respect to the Access Network.

1) The Selection Function accepts information from MSP Servers and Vehicle Systems.
2) The Selection Function executes the selection algorithm and assigns the target MSP Server.
3) The Vehicle System connects to the target MSP Server.

Figure 28 shows an example of implementing Edge Server assignment by a Selection Function using DNS.

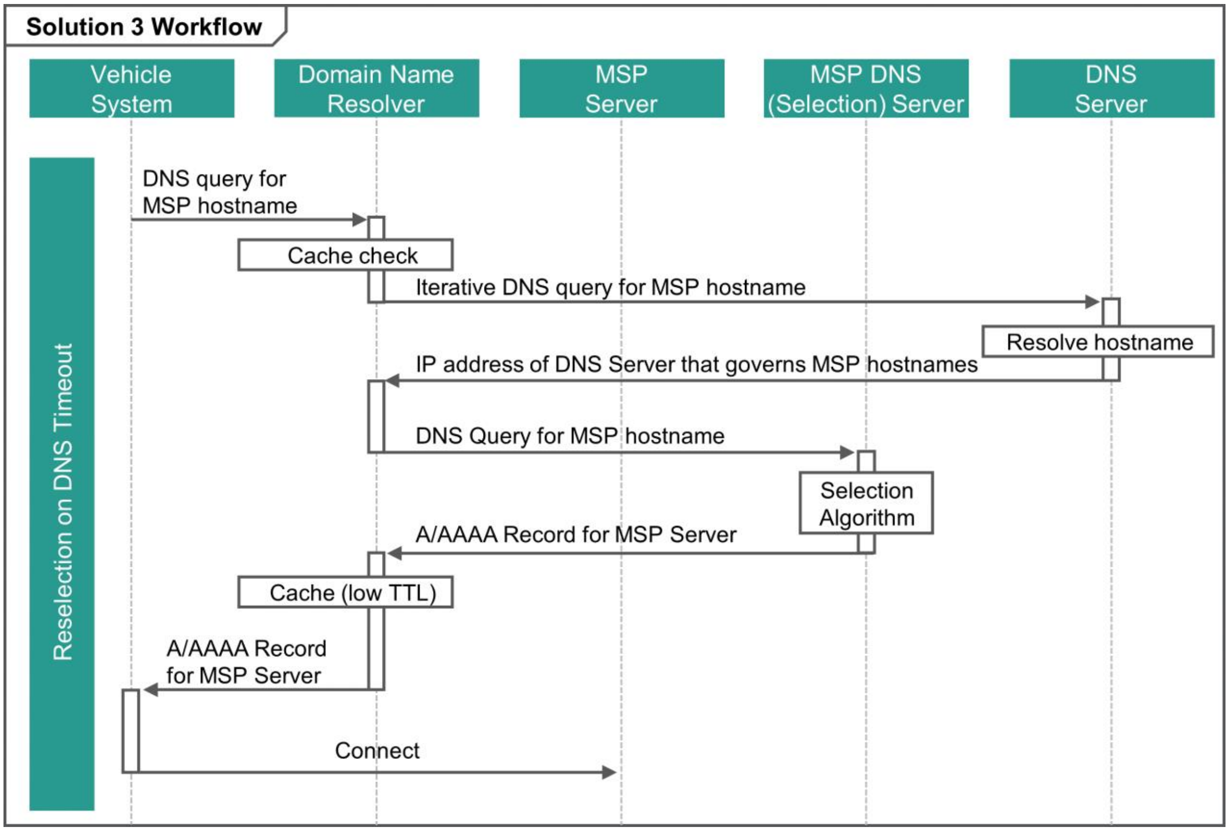
Solution 3 Workflow
Vehicle System
Domain Name Resolver
MSP Server
MSP DNS (Selection) Server
DNS Server
Reselection on DNS Timeout
DNS query for MSP hostname
Cache check
Iterative DNS query for MSP hostname
Resolve hostname
IP address of DNS Server that governs MSP hostnames
DNS Query for MSP hostname
Selection Algorithm
A/AAAA Record for MSP Server
Cache (low TTL)
A/AAAA Record for MSP Server
Connect

Figure 28. Example of an implementation of Solution 3.

Notes on the implementation:

- Due to the Vehicle System's movement within the environment, MSP Server reselection may be required. To enable responsive MSP Server reselection, appropriate DNS cache timeouts should be selected.

- When an application is deployed in a manner such that it is capable of covering only a certain geographic region, the Selection Function will need the vehicle's geolocation information. The region or area supported by a particular application instance may differ based on each service or as a result of the computing resources assigned to an application instance. This may need to be taken into consideration with respect to the DNS query.
- The selection sequence takes into account that the resolved IP address may be an address of a load-balancer, thus triggering Solution 5. When using a 5GS, additional functions may be needed to enable the correct Edge Selection when coupled with the offloading function discussed in the previous key issue, Edge Data Offloading.

3.2.2.4 Solution 4 – Vehicle System-based MSP Server Assignment

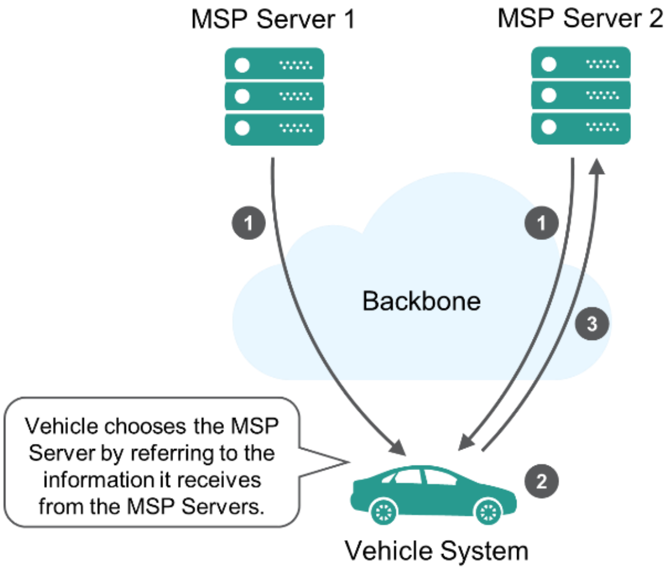
MSP Server 1
MSP Server 2
1
1
Backbone
3
2
Vehicle chooses the MSP Server by referring to the information it receives from the MSP Servers.
Vehicle System

Figure 29. Vehicle System-based MSP Server Selection.

This solution can be combined with the Selection Function solution, allowing MSP Server assignment without the Vehicle System having to share sensitive information, such as vehicle geolocation. In this approach, the Vehicle System will choose its MSP Server. The Vehicle System may select the appropriate MSP Server based on in-vehicle information such as physical vehicle location and/or additional information provided by potential MSP Servers. This approach is agnostic with respect to the Access Network.

1) The Vehicle System requests information from MSP Servers.
2) Based on the information received by the Vehicle System, the Vehicle System selects an MSP Server.
3) The Vehicle System connects to the target MSP Server.

Figure 30 shows an example of implementing the combination of a Selection Function and vehicle-based assignment.

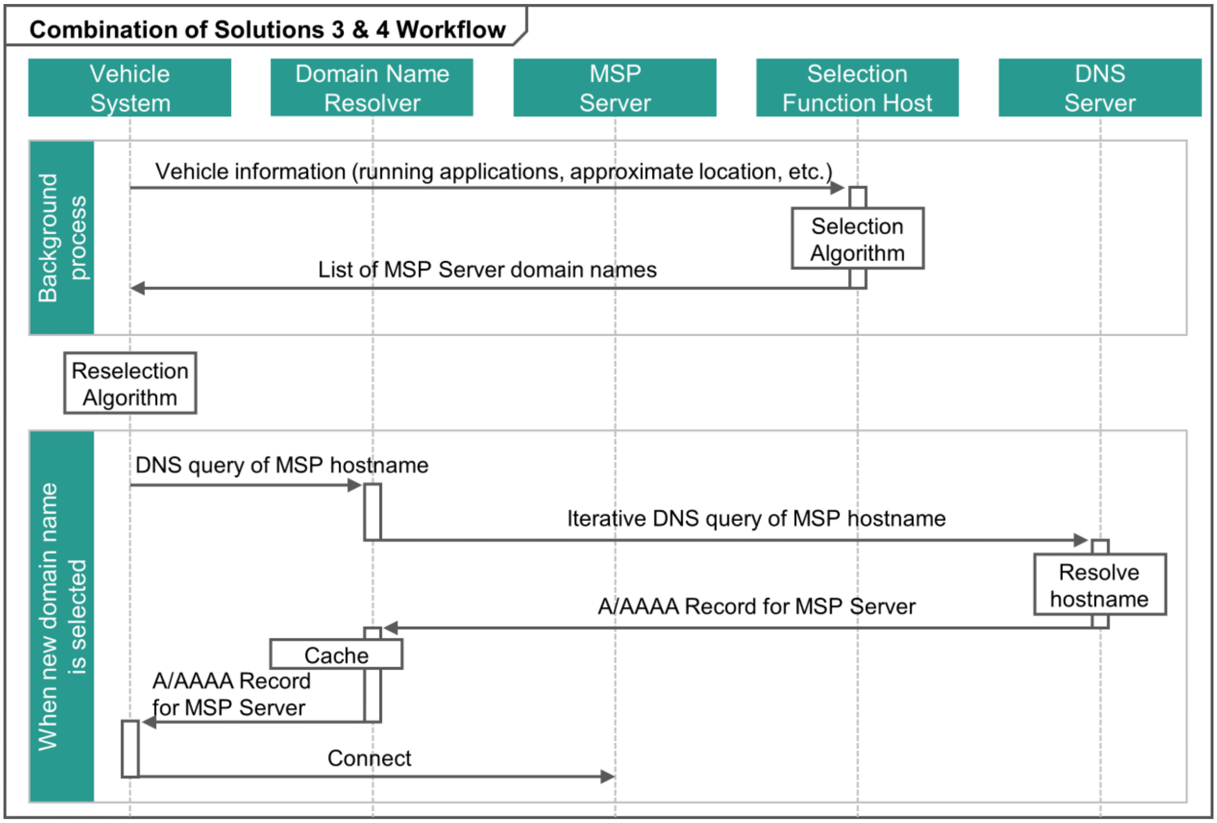
Combination of Solutions 3 & 4 Workflow
Vehicle System
Domain Name Resolver
MSP Server
Selection Function Host
DNS Server
Background process
Vehicle information (running applications, approximate location, etc.)
Selection Algorithm
List of MSP Server domain names
Reselection Algorithm
When new domain name is selected
DNS query of MSP hostname
Iterative DNS query of MSP hostname
Resolve hostname
A/AAAA Record for MSP Server
Cache
A/AAAA Record for MSP Server
Connect

Figure 30. Example of an implementation of the combination of Solution 3 and Solution 4.

3.2.2.5 Solution 5 – Combination with Load Balancer for MSP Server Selection

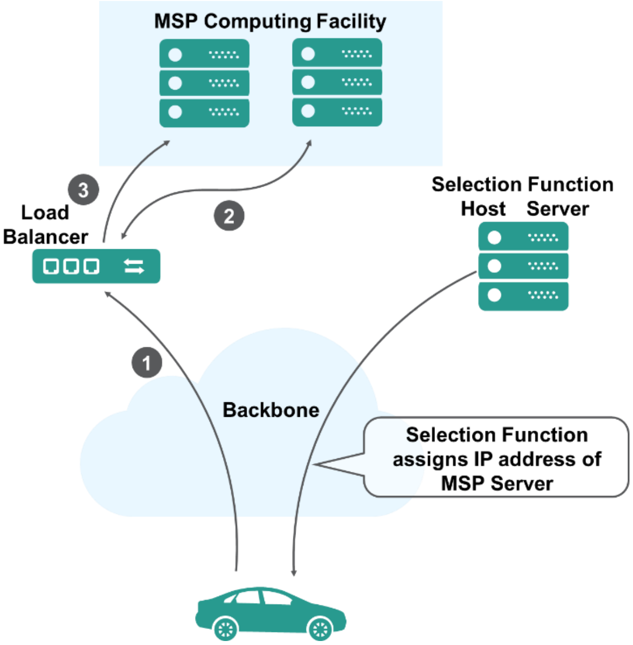
MSP Computing Facility
3
Load Balancer
2
Selection Function Host Server
1
Backbone
Selection Function assigns IP address of MSP Server

Figure 31. Example illustrating a load balancer coupled with selection on a server.

This method combines one of the selection methods with an application-aware load balancer in order to determine which application instance should be used by the Vehicle System. One must be aware that the MSP Edge Server is not a single server but rather a set of servers in one or more data centers composed of the necessary functionality to support the various service scenarios. In the diagram above, the Vehicle System will request IP address resolution for a particular service, with the result being an appropriate IP address that happens to be allocated to a load-balancer device.

1) The Vehicle System initiates a connection toward the load balancer.
2) The load balancer determines which MSP Server to use.
3) The load balancer will manage the session on behalf of the Vehicle System toward the selected application instance within the MSP data center.

3.2.2.6 Parameters for MSP Server Assignment

The following parameters may be used for MSP Server Assignment.

- IP Ping **Round Trip Time (RTT) –** The RTT between the Vehicle System and reachable MSP Servers.
- **Completion Rate –** The quality of this parameter is inversely proportional to the number of packet losses and timeouts between the Vehicle System and the MSP Server.
- **Hops –** Number of hops needed to route the data between the Vehicle System and the MSP Server.
- **Vehicle System Physical Location –** To address localized contents/process, the Vehicle System's physical location may be required. Attention to security and privacy considerations is required.

- **Request SLA –** By identifying the dataflow, the system is able to identify the required SLA of the request, including the required time needed to complete the processing of a particular data flow.
- **Server Turnaround Time –** The time needed for an application in the MSP Server to complete a process queried by the Vehicle System.
- **Server Load –** The load of an MSP Server; this might include CPU and memory utilization.

3.2.2.7 Considerations across Key Issues

In current deployments of LTE and WLAN networks, when dynamic IP address assignment and configuration are used, the network typically provisions the IP address of a DNS server. The DNS server may be located close to the Network Edge and configured in a manner so as to provide results that are aware of resources that are close to it. Conversely, a DNS server may be located in a remote network and therefore be unaware of resources located within the local network. In 5G, an MNO's DNS Server may serve clients from multiple anchor points.

Consider now the combination of a DNS-based solution with an uplink classifier-based solution, which is the preferred Edge Data Offloading solution for a 5GS. As the user has multiple possible anchor points of presence that can be used based on IP header filtering (typically destination IP address), the Cellular Network must ensure that a DNS query is sent to a Domain Name Resolver at a point of presence that is deemed feasible for the corresponding FQDN and UE location in the network.

As a solution, the mobile network operator operating the respective Cellular Network should provide a central DNS stub resolver that, when receiving a DNS query, forwards this query to an appropriate Domain Name Resolver. This action is based on the tracking area (looked up using standard 5G core functionality) of the UE, and potentially on additional knowledge related to the FQDN, such as rules about feasibility of different breakouts for specific domain names. In addition, the DNS stub resolver may aggregate multiple records in the reply.

3.2.3 Conclusions

Recommendations on the technologies to be used for MSP Server Selection are as follows:

- Solution 3 is preferable for deployments in the near future, since this solution does not involve major modifications to the Vehicle System. Furthermore, it minimizes the customization effort for the Access Network. Dedicated DNS Servers (MSP DNS Servers) that are authoritative for the corresponding DNS zone(s) are recommended to be deployed for MSP Server Selection, to enable seamless integration with existing deployed systems, due to their wide adoption on the internet, in Access Networks and on existing clients.
- Solution 4 or the combination of Solution 4 and Solution 3 allows the Vehicle System to select MSP Servers without the Vehicle System having to share information related to the vehicle. This combination allows the MSP Server Selection service to give possible options to the vehicle and the final decision to be made by the vehicle. This option is also transparent to the Access Network. A specific selection module must be implemented within each Vehicle System.
- Solution 1 allows a deployment that is transparent to the Vehicle System and MSP Servers.

- In conclusion, Solution 3 is recommended as a baseline solution, preferably reusing existing DNS functionality and infrastructure, while both Solutions 1 and 4 are feasible to offer enhanced functionality for the MSP Server selection process.

3.3 Vehicle System Reachability

3.3.1 Key Issue

In the AECC distributed computing architecture, MSP Servers – Center or Edge – are required in many use cases to send data to the Vehicle System. However, it is challenging for the MSP Servers to effectively reach the Vehicle System.

1) The IP anchor point changes.
2) The IP mapping changes due to network functions such as an NAT/NAPT timer’s expiration, resulting in a change of the Global IP address.
3) A service outage occurs when the Vehicle System moves into an area without network coverage.
4) Handover between different Access Networks occurs.

All these issues can cause a Vehicle System IP address change that results in MSP Servers not being able to reach the Vehicle System. Therefore, the AECC System needs to deploy a specific mechanism to ensure the reachability of a Vehicle System according to service requirements.

Note 1: depending on the service requirements, the Vehicle System reachability issue could be handled differently by the application layer or by both the application and network layer.

Note 2: both IP and non-IP based solutions shall be considered for this key issue.

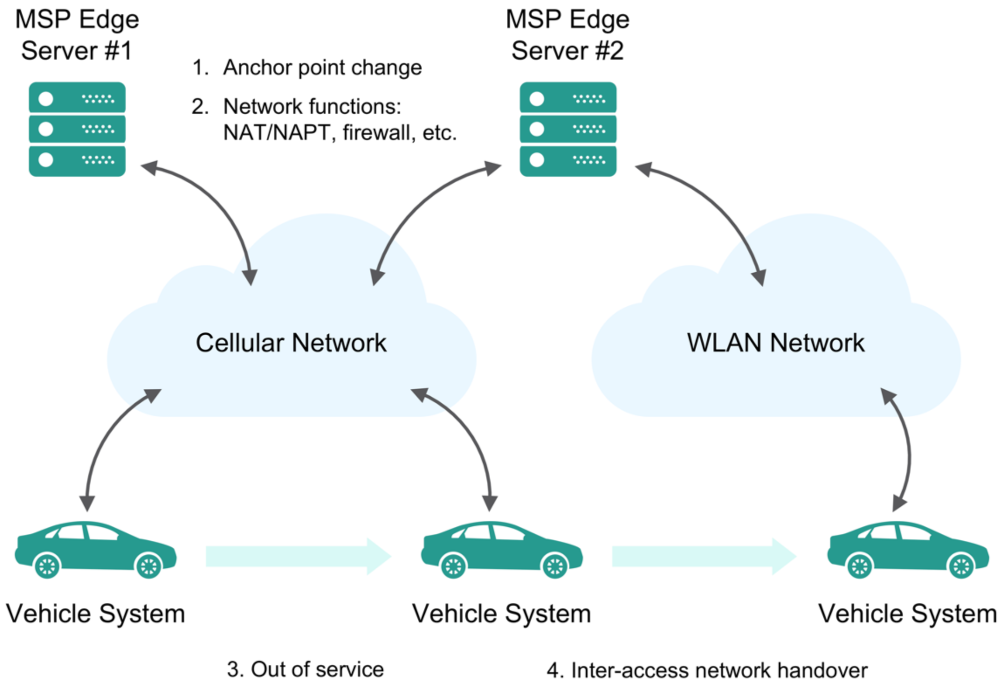

Figure 32. Typical causes of the Vehicle System reachability issue.

3.3.2 Potential Solutions

Three potential solutions are investigated in the following sections.

1. SMS Push.
2. Push Notifications.
3. Vehicle System Triggering via Network Exposure Function.

3.3.2.1 Solution 1 – SMS Push

SMS Push is an SMS "call-in" trigger message common in many current Over-the-Air (OTA) software delivery systems. The trigger message is sent to the target Vehicle System by a campaign manager application. The application maintains links to a database operated by the vehicle manufacturer that contains the IMEI number of each Vehicle System within the fleet. Database fields enable the vehicle manufacturer to identify subsections of the vehicle fleet based on parameters such as vehicle type, country and so on. When the SMS trigger message is received, the Vehicle System initiates the appropriate application. The application will then establish a connection to the OTA delivery system. Once a Vehicle System connects, it is marked as having received the message and taken the action. This allows the OTA system to identify those vehicles that have not yet connected.

Since the application in the Vehicle System initiates the connection into the OTA delivery system, the IPv4/IPv6 address of the Vehicle System will be obtained when the connection is established.

3.3.2.2 Solution 2 – Push Notifications

Push notifications are a common method for maintaining connectivity used in mobile consumer device platforms, such as smartphone applications.

A push notification is a message that is "pushed" automatically from a backend server or application to remote clients. These notifications are sent from the application to a remote server, which acts as an intermediary. Each client application needs to be registered with the remote server using a unique key or UUID. The remote server then sends the message against the unique key and delivers the message to the client application via an agreed client/server protocol such as HTTP or XMPP.

The Vehicle System's IPv4/IPv6 address is not required since the backend server or application communicates with the remote server, and the client-side application in the Vehicle System registers with the remote server.

3.3.2.3 Solution 3 – Vehicle System Triggering via Network Exposure Function

Network Exposure Functions, such as SCEF (defined in 3GPP TS 23.682) and NEF (defined in 3GPP TS 29.522), specify various Network APIs for third parties. The Vehicle System could be triggered via a control plane message to establish its IP connection to the MSP Servers with its currently used IP address.

One solution example demonstrated in Figure 33 shows how the oneM2M framework leverages the exposure functions of Cellular Networks [8]. In this example, the oneM2M function in the MSP Server will continuously maintain the ID-IP binding between the vehicle oneM2M ID and its IP address. The binding can be updated periodically or by an event that uses the Cellular Network exposure function to trigger the oneM2M function in the vehicle system for updating its IP address. The applications using the oneM2M framework in the MSP Server will only need to know the unique ID to wake and reach the Vehicle System.

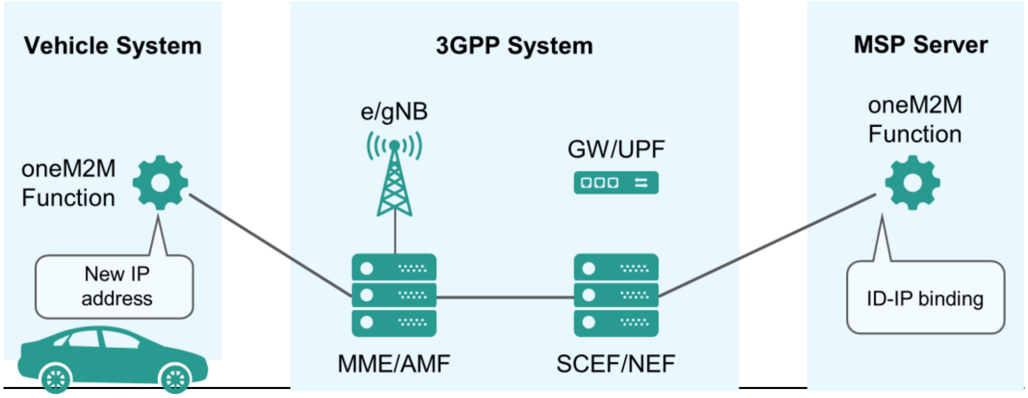

Figure 33. Example of Network Exposure Functions.

3.3.3 Conclusions

Recommendations on the technologies to be used for Vehicle System Reachability are as follows.

- Solution 2 is preferable for supporting both Cellular Network and WLAN access types, which is an architectural requirement in the AECC System. It is recommended to deploy this solution as an AECC function, but it is also flexible enough to be implemented by a third-party application that is outside the AECC System.
- Solution 1 could be a fallback solution when the Cellular Network is the only access to be used for push services, since Solution 1 already has good adoption in many existing industry sectors, such as automotive and other IoT systems.
- Solution 3 would be promising to replace Solution 1 in Cellular Networks that support the SCEF/NEF, due to the advantages of less Vehicle System complexity and better extensibility as compared to Solution 1.

3.4 Access Network Selection

3.4.1 Key Issue

As shown in Figure 34, the Vehicle System is expected to use a mix of different wireless access technologies, including cellular and WLAN, to connect to MSP Servers. It may be preferable for the Vehicle System to use multiple Access Networks simultaneously to access increased bandwidth or to improve reliability. For example, a Vehicle System may be traveling through an area where service coverage from network operators improves and degrades. How should the Vehicle System adapt to the changing communications environment? At the same time data flows with different QoS requirements or from different applications may be required to go through different Access Networks in order to meet AECC service requirements. Besides the Access Network status, information such as policies, service requirements, network connectivity and the Vehicle System’s movement can be considered as part of the process. Therefore, a mechanism is needed to enable the Vehicle System or the AECC System to select from multiple Access Networks and steer traffic over different connections. The elements involved in this process are Access Networks, the Vehicle System and MSP Servers.

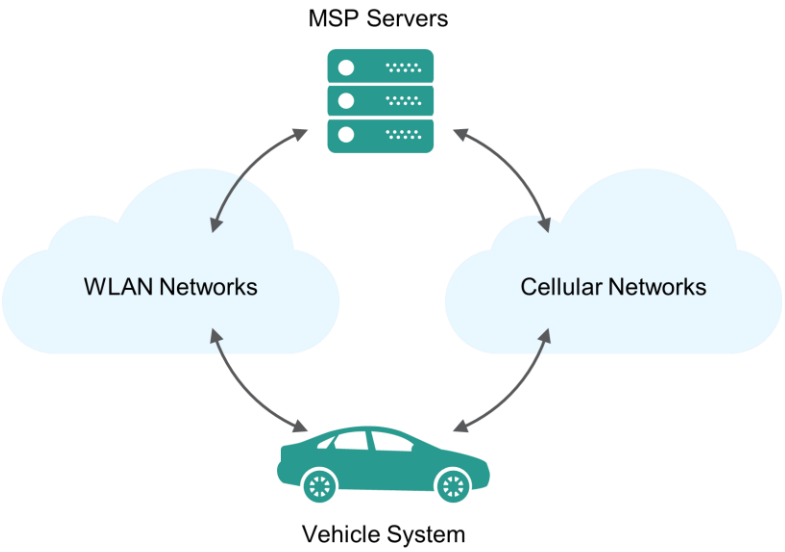

Figure 34. Access Network selection.

3.4.2 Potential Solutions

Access Network selection can be divided into two steps: connection selection and traffic steering. Connection selection focuses on how to select one or multiple connections based on the end-to-end system capabilities and other information from the Vehicle System, the Cellular Network(s), the WLAN(s) and MSP Servers. Examples of relevant information are wireless signal strength, application requirements and the movement of the Vehicle System through the environment. Traffic steering focuses on how to steer traffic to multiple connections and enable different applications to use different connections simultaneously. Access Network selection can be triggered periodically or by events.

Solutions for Access Network Selection are listed below.

Solutions for connection selection:

- Solution 1 – Vehicle System-based Solutions for Connection Selection
  - Solution 1.1 – Application layer solution on the vehicle
  - Solution 1.2 – System layer solution on the vehicle
- Solution 2 – MSP Server-based Solutions
  - Solution 2.1 – Application layer solutions on the MSP Server
  - Solution 2.2 – System layer solutions on the MSP Server
- Solution 3 – Access Layer Solutions
  - Solution 3.1 – ANDSF
  - Solution 3.2 – P-GW level convergence based on S2a/S2b interfaces

Solutions for traffic steering:

- Solution 4 – System Layer Solutions
  - Solution 4.1 – MPTCP/MPQUIC
  - Solution 4.2 – Generic Multi-Access (GMA)/Multi-Access Management Services (MAMS)
- Solution 5 – Access Layer Solutions
  - Solution 5.1 – LTE WLAN Aggregation (LWA)
  - Solution 5.2 – LTE WLAN radio-level integration over IPSec tunnel (LWIP)
  - Solution 5.3 – Access Traffic Steering, Switch and Splitting (ATSSS)
  - Solution 5.4 – Multi Radio Dual Connectivity (MR-DC)

The first section below provides the solution overview. Afterwards, all solutions are defined one by one.

3.4.2.1 Solution Overview

As shown in Figure 35, the basic approach for both connection selection and traffic steering takes multiple inputs from the Vehicle System, MSP Servers and Access Networks to make decisions about selecting one or multiple connections, or steering traffic over the selected connections.

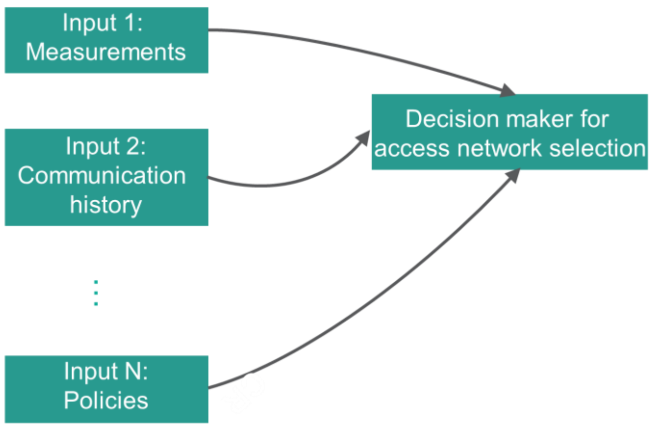

Figure 35. Basic approach for Access Network selection.

The inputs for Access Network selection may include:

- Policy: this refers to the policies used to govern Access Network selection, such as
  - A condition regarding channel state for selecting an Access Network. For example, a threshold of the SINR can be defined to decide when to switch between Access Networks.
  - A condition regarding the Vehicle System's location for selecting an Access Network. For example, the Vehicle System may switch to a WLAN once it arrives home.
  - A condition regarding service or traffic types for selecting an Access Network
  - Operational policies such as using a primary Access Network if it is available
- Data: this refers to the data or information used to assist Access Network selection (note that the inputs are not mandatory and depend on availability), such as
  - Radio link measurements about different RATs, such as signal strength, SINR
  - Communication history
  - Access Network charging information
  - Network status, such as congestion or Round Trip Time (RTT)
  - Location information of the Vehicle System or MSP Servers
  - Service-related information, such as service type, data rate, traffic behavior, traffic statistics, priority
  - Application context, such as the user's preference

As shown in Figure 36, the Access Network selection assistance function can be located within different layers on the Vehicle System or the MSP Servers:

- The access layer generally consists of cellular and WLAN networks, including both radio access networks and core networks, and other infrastructure that provides connectivity between the Vehicle System and MSP Servers.
- The system layer generally includes a function layer, a platform layer, an infrastructure layer and computing, network and storage on the Vehicle System and MSP Servers.
- The application layer includes the applications running on the Vehicle System and MSP Servers.

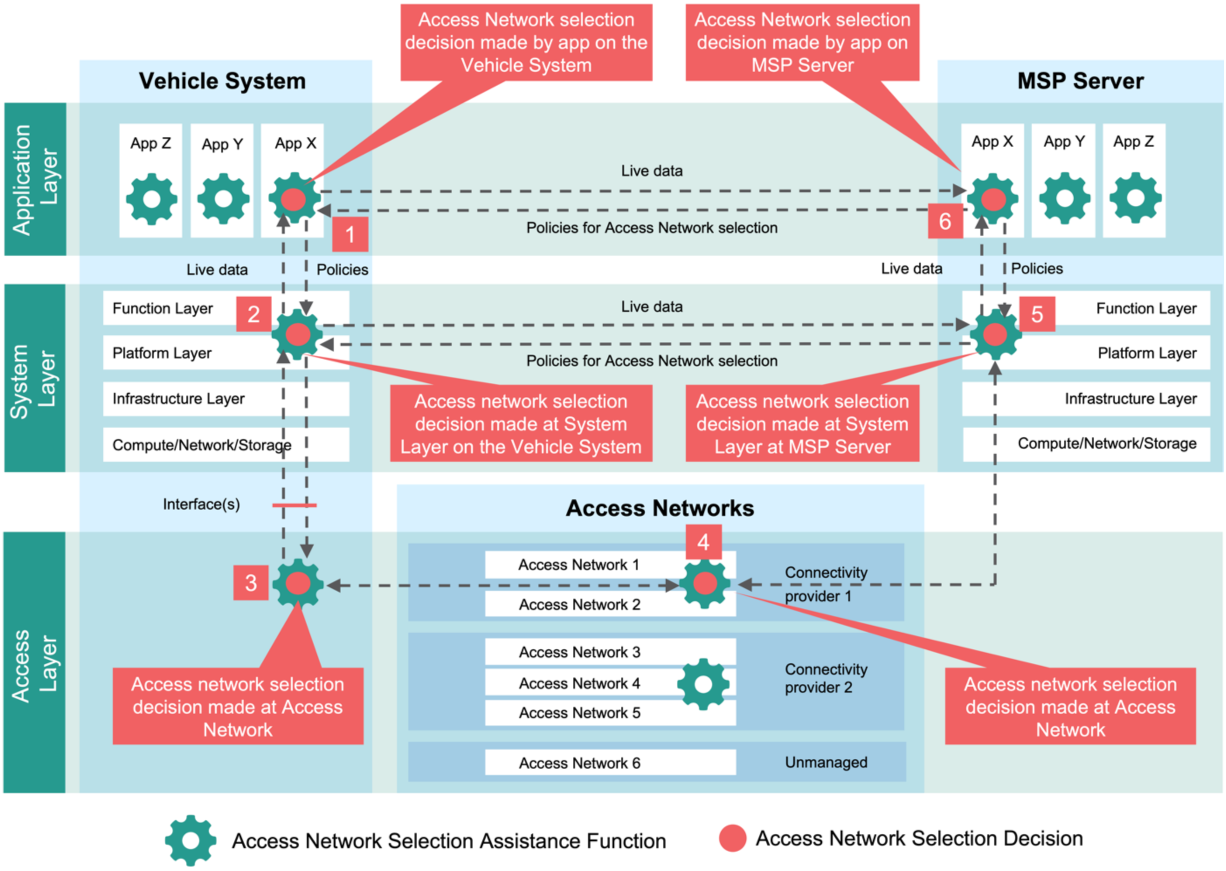
Access Network selection decision made by app on the Vehicle System
Access Network selection decision made by app on MSP Server
Vehicle System
MSP Server
Application Layer
App Z
App Y
App X
App X
App Y
App Z
Live data
Policies for Access Network selection
1
6
Live data
Policies
Live data
Policies
System Layer
Function Layer
Platform Layer
Infrastructure Layer
Compute/Network/Storage
2
5
Live data
Policies for Access Network selection
Function Layer
Platform Layer
Infrastructure Layer
Compute/Network/Storage
Access network selection decision made at System Layer on the Vehicle System
Access network selection decision made at System Layer at MSP Server
Interface(s)
Access Networks
4
Access Layer
3
Access Network 1
Access Network 2
Connectivity provider 1
Access Network 3
Access Network 4
Access Network 5
Connectivity provider 2
Access Network 6
Unmanaged
Access network selection decision made at Access Network
Access network selection decision made at Access Network
Access Network Selection Assistance Function
Access Network Selection Decision

Figure 36. Possible information sharing and decision making for Access Network selection.

Figure 36 illustrates the possible information exchange between functions that may assist the selection process. The application layer may provision policies to the system layer, which again may provision policies to the access layer. Access Network selection assistance functions outside the Vehicle System may provision policies to the Vehicle System, within the respective layer. In addition, the access layer may share live data with the system layer, and the system layer may share live data with the application layer. The Vehicle System may also share live data with Access Network selection assistance functions outside the Vehicle System, within the respective layer. For example, this could include metrics relating to throughput. For a practical solution, only a subset of the mentioned information is shared, depending on where the selection algorithm is instantiated and the requirements for Access Network selection.

While Access Network selection in the system and application layers deals with the selection of Access Networks exposed by the modem in the Vehicle System (this could be a single Access Network per connectivity provider), selection in the access layer deals with the selection of Access Networks by a single connectivity provider, and the two selection processes may be combined.

Figure 36 also illustrates how the described layers relate to the reference architecture. The key point is that on both the Vehicle System and MSP Server sides, Access Network selection assistance functions in different layers can work together to enable different granularities of Access Network selection. Placing the application layer selection assistance function will be application-specific, while placing the selection assistance function in the system layer is Vehicle System-specific.

The decision on Access Network selection includes the decision on connection selection and the decision on traffic steering. The decision on connection selection may be made at a different layer from where the decision on traffic steering is made. However, the design of the whole system shall make sure that the mechanisms of connection selection and traffic steering can work together.

3.4.2.2 Solutions for Connection Selection

Connection selection for links includes two stages:

- Stage 1: initial connection selection for selecting a connection without any existing connection
- Stage 2: connection reselection for selecting a connection with one or more existing connections

Initial connection selection would be largely based on in-vehicle information and provisioned policies. Connection reselection could be conducted by an MSP Server or the Vehicle System, or assisted by a Cellular Network with the assumption that one or more connections is successfully established.

Types of connection selection include:

- Type 1: only one connection is selected at one time.
- Type 2: multiple connections are selected and can be activated simultaneously.

Type 1 connection selection generally can be based on policies such as using a WLAN if available and using a Cellular Network otherwise. It can also be provided by Access Network Discovery and Selection Function (ANDSF) policies, which will be introduced in Solution 3.1 in this section.

Type 2 connection selection could consider the end-to-end system capabilities of traffic steering. For example, the vehicle could select the following combinations of Access Networks: independent Access Networks such as primary cellular, secondary cellular and WLAN by different operators, or Access Networks that can support some interworking features, such as a Cellular Network and a WLAN connected to the same core network. ANDSF also defines policies for selecting multiple Access Networks.

The connection selection can be made at the application, system or access layers on the Vehicle System or MSP Servers. The decision on selecting a connection at different layers will result in selecting one or multiple physical links or logical links. However, a logical link can be mapped to a physical link, and eventually to an Access Network.

In Table 1, the solutions for connection selection are summarized based on where the connection selection decision is made, conditions and requirements, and the selection types and stages as defined above. The details of each solution are introduced afterwards.

Table 1. Summary of solutions for connection selection.

Solutions	Layers where decision is made	Option number in Figure 36	Assumptions and requirements	Selection Type and Stage
Solution 1.1 -- Application layer solution on the Vehicle System	At application layer on the Vehicle System	1	Northbound interface from the system layer to applications and the southbound interface to the Access Network may be required on the Vehicle System.	Stage 1 and Stage 2 for Type 1 or Type 2
Solution 1.2 -- System layer solution on the Vehicle System	At system layer on the Vehicle System	2	Northbound interface from the system layer to applications and the southbound interface to the Access Network may be required on the Vehicle System.	Stage 1 and Stage 2 for Type 1 or Type 2
Solution 2.1 -- Application layer solutions on the MSP Server	At application layer on the MSP Server	6	Northbound interface from the system layer to applications and the southbound interface to the Access Network may be required on the MSP Server.	Stage 2 for Type 1 or Type 2
Solution 2.2 -- System layer solutions on the MSP Server	At system layer on the MSP Server	5	Northbound interface from the system layer to applications and the southbound interface to the Access Network may be required on the MSP Server.	Stage 2 for Type 1 or Type 2
Solution 3.1 -- ANDSF	At access layer	3	Both the Vehicle System and the Access Network have to support ANDSF.	Stage 2 for Type 1 or Type 2
Solution 3.2 -- P-GW-level convergence based on S2a/S2b interfaces	At access layer	4	Applied to LTE network	Reselection for a WLAN for Type 2

Solution 1 – Vehicle System-based Solutions for Connection Selection

The Vehicle System-based solutions apply to the scenarios where the Vehicle System makes the decision on connection selection.

Solution 1.1 – Application layer solution on the Vehicle System

Connection selection can be implemented as part of the application on the Vehicle System that receives inputs from different layers, such as the application, system and access layers on both the Vehicle System and MSP Servers, to make decisions, shown as Option 1 in Figure 36. For example, Access Networks could provide measurements of signal strength and an available network list or policies from an ANDSF to the connection selection function in the application in order to make a decision. In addition, the connection selection functions in the application layer on the MSP Server can provide policies for connection selection. This solution can also integrate the user's preference to make a selection decision.

To facilitate this solution, it may be possible for the Vehicle System to provide northbound interface APIs with connection selection functions at the application layer. Functions of the northbound interface may include the Vehicle System providing information about the AECC System's capabilities in supporting Access Networks, and providing information (e.g., Access Network conditions, location) for the application to make a selection decision. The Vehicle System could also provide a southbound interface to the Access Network. Functions of the southbound APIs may include Access Network measurement reports and the Access Network charging policy. However, the Vehicle System may have no or limited support for the aforementioned capabilities.

Solution 1.2 – System layer solution on the Vehicle System

The decision on connection selection can be made at the system layer on the Vehicle System, which can get inputs from the connection selection assistance functions in the application and access layers on both the Vehicle System and MSP Servers, shown as Option 2 in Figure 36. In this solution, the connection selection decision can be agnostic about applications.

To facilitate this solution, the Vehicle System's system layer could provide a northbound interface to applications. Functions of the northbound APIs may include getting the policies or user preference for connection selection. The Vehicle System could provide a southbound interface to the Access Network. Functions of southbound APIs may include Access Network measurement reports and the Access Network charging policies.

Solution 2 – MSP Server-based Solutions

MSP Server-based solutions apply to the scenarios where MSP Servers make the decision on connection selection once an initial connection has been established by the Vehicle System. The selected Access Networks can be either Cellular Networks or WLANs. The MSP Server-based solutions apply only to connection reselection and require capabilities from the MSP Server to have related interfaces and get information about Access Networks.

Solution 2.1 – Application layer solutions on the MSP Server

The decision on connection selection can be made at the application layer on the MSP Server, which can get inputs from the connection selection assistance functions in the access layer and system layer on both the Vehicle System and MSP Servers, shown as Option 6 in Figure 36. The information from the Access Network can be sent to the application layer or system layer on the MSP Server via an application layer interface between the Vehicle System and MSP Server or via the southbound interface between the MSP Server and the Access Networks.

A primary connection needs to be established for the communication between the Vehicle System and MSP Servers**.** To facilitate this solution, the AECC System could provide northbound interface APIs to the application. Functions of the northbound APIs include providing information on the AECC System's capabilities in supporting Access Networks, providing information (e.g., Access Network conditions, location) for the application to make a selection decision. The AECC System shall provide a southbound interface to the Access Network. Functions of southbound APIs may include providing Access Network measurement reports and the Access Network charging policy.

Solution 2.2 – System layer solutions on the MSP Server

The decision on connection selection can be made at the system layer on the MSP Server, which can make use of the connection selection assistance functions in the application and access layers on both the Vehicle System and MSP Servers, shown as Option 5 in Figure 36. In this case, connection selection assistance functions at the system layer can be provided as standardized or proprietary APIs. A primary connection needs to be established for the communication between the Vehicle System and MSP Servers for the Vehicle System to provide information to the MSP Servers in order to execute the connection selection assistance function and notify the Vehicle System of the result.

To facilitate this solution, the AECC System could provide northbound interface APIs to applications. Functions of the northbound APIs may include getting the policies or user preference on network selection. The connection selection assistance function may only be able to use "self-monitored" information such as traffic statistics. If the MSP Servers have interfaces to the Access Networks, information exchanged with the Access Network may be leveraged. The AECC System may provide a southbound interface to the Access Network. Functions of the southbound APIs may include providing Access Network measurement reports and the Access Network charging policy.

Solution 3 – Access Layer Solutions

In these solutions, connection selection is assisted by the Cellular Network, and the Cellular Network needs to support certain capabilities to enable Access Network selection.

Solution 3.1 – ANDSF

In EPS or 5GS, an ANDSF can provide an available network list or connection selection-related policies as specified in a 3GPP specification [10]. The decision on connection selection is made in the Vehicle System based on the policies received from ANDSF, shown as Option 3 in Figure 36. The input to the ANDSF is the UE's profile, which specifies what output is expected from the ANDSF.

The UE may retain and use the Access Network discovery information provided by the ANDSF until new/updated information is retrieved. The ANDSF communicates with the UE over the S14 reference point, which is essentially a synchronization of an OMA-DM management object (MO) specific to the ANDSF.

This solution requires that the Vehicle System can receive the ANDSF policy from the cellular MNO and that both the Vehicle System and Cellular Networks support the ANDSF. This solution applies to both initial connection selection and connection reselection. The ANDSF does not provide RAN-level information but mostly provides information based on the Vehicle System's location.

Solution 3.2 – P-GW level convergence based on S2a/S2b interfaces

In EPS, S2a/S2b can provide Core Network (CN)-based interworking between cellular and WLAN networks. The S2a or S2b is the interface between the P-GW and a trusted Non-3GPP IP access or an untrusted Non-3GPP IP access to offload traffic to a non-3GPP Access Network such as a WLAN. The decision on connection selection is made by the Access Network based on policies such as user preference, RAN-level policies or WLAN service provider policies, shown as Option 4 in Figure 36.

Connection selection between 3GPP access and a WLAN is supported using an ANDSF or using RAN rule procedures without an ANDSF. As in Solution 3.1, an ANDSF can provide a list of available WLAN networks and related information to facilitate discovery and connection establishment. However, it does not contain information at RAN level, such as signal strength. When the Vehicle System has valid 3GPP subscription credentials (i.e., a valid USIM) and WLANSP policies, the Vehicle System can perform WLAN selection based on WLANSP policies, the applicable user preferences and the corresponding procedures as specified in a 3GPP specification [9]. This solution requires that the Access Networks integrate with the cellular core network, and it only applies to connection reselection for a WLAN network.

3.4.2.3 Solutions for Traffic Steering

Based on how traffic steering is supported, the solutions can be classified as:

- Non-seamless multi-access mode: in this case, multiple-Access Networks operate independently and no coordination is provided.
- 3GPP seamless multi-access: in this case, traffic steering relies on the 3GPP network to provide seamless multi-access service. Convergence between the Cellular Network and WLAN can be provided at CN level or RAN level. Generally, two Access Networks are supported; i.e., one Cellular Network and one WLAN. A single IP address may be used, independent of the underlying Access Networks.
- Non-3GPP seamless multi-access: in this case, a higher-layer solution needs to be supported for cellular and WLAN network convergence. Convergence can be provided to more than two Access Networks simultaneously. A single IP address may be used, independent of the underlying Access Networks.

Traffic steering enablers are listed below:

- **Enabler 1 – APN**
  One or more APNs can be allocated to an application to identify different flows as agreed with the operators. In this case, an application may be presented with endpoints (such as next-hop IP addresses) that can be used to place traffic onto a particular APN. This technology requires agreement with the MNO to assign APNs and applies only to Cellular Networks for non-seamless multi-access mode.
- **Enabler 2 – Socket Binding**
  Socket binding can bind a socket to a specific network interface IP address. For example, if an application opens a socket for the TCP protocol and explicitly binds it to the WLAN interface, then the socket sends and accepts data only from the WLAN interface. This allows an application to transfer data only when the WLAN is available. This solution applies to both the Cellular Network and WLAN for non-seamless multi-access mode.

- **Enabler 3 – SD-WAN**
  Established Software-Defined Wide Area Network (SD-WAN) solutions are being enhanced with host functionality, which can then be operated in the vehicle system layer. Established SD-WAN policies define traffic steering policies, including whether to route packets directly to the internet or via secured tunnels. These SD-WAN policies can be enhanced to enable traffic steering based on Access Network characteristics. This solution applies to cellular and WLAN networks for seamless multi-access mode.

In Table 2, the solutions for traffic steering are summarized based on where the traffic steering decision is made, conditions and requirements. The details of each solution are introduced afterwards.

Table 2. Summary of solutions for traffic steering.

Solutions	Layers where traffic steering decision is made	Option number indicated in Figure 36	Assumptions and requirements
Solution 4.1 – MPTCP/MPQUIC	At system layer	2 and 5	Both the Vehicle System and MSP Server support MPTCP/MPQUIC or there is a function to do the mapping between TCP/QUIC and MPTCP/MPQUIC.
Solution 4.2 – Generic Multi-Access (GMA)/Multi-Access Management Services (MAMS)	At system layer	2 and 5	Both the Vehicle System and MSP Server support GMA/MAMS.
Solution 5.1 – LTE WLAN Aggregation (LWA)	At access layer	4	Applies only to LTE network; the Cellular Network has to support LWA.
Solution 5.2 – LTE WLAN Radio-level integration over IPSec tunnel (LWIP)	At access layer	4	Applies only to LTE network; the Cellular Network has to support LWIP.
Solution 5.3 – Access Traffic Steering, Switch and Splitting (ATSSS)	At access layer	4	Applies only to 5G network; the Vehicle System and the Cellular Network have to support ATSSS.
Solution 5.4 – Multi Radio Dual Connectivity (MR-DC)	At access layer	4	Both the Vehicle System and the Cellular Network have to support MR-DC.

Solution 4 – System Layer Solutions

Solution 4.1 – MPTCP/MPQUIC

Multipath TCP (MPTCP) is a major modification to TCP that allows multiple paths to be used simultaneously by a single transport connection. The MPTCP protocol has been standardized by the IETF in RFC 6824. MPTCP allows multiple sub-flows to be set up for a single MPTCP session. An MPTCP session starts with an initial sub-flow. Then, after the first MPTCP sub-flow is set up, additional sub-flows can be established that are bound to the existing MPTCP session [11]. Data for the connection can then be sent over any of the active sub-flows that have the capacity to take it.

Quick UDP Internet Connection (QUIC) is a multiplexed and secure transport protocol that runs on top of UDP and combines functions of HTTP/2, TLS and TCP. QUIC is targeted at reducing the latency of client-server communication, providing an alternative to the conventional layered HTTP/TLS/TCP protocol stack used by the web. One of the rationales for the development of QUIC has been the constraints experienced by TCP that is implemented in operating system kernels and middlebox firmware, making significant changes to TCP (e.g., multipath capability) very challenging to deploy [11]. Being based on UDP, QUIC doesn’t suffer from such limitations and hence is able to incorporate new features without having to upgrade legacy systems. Multipath QUIC (MPQUIC) is an extension to the QUIC protocol that enhances the migration capabilities to enable support of a single connection over multiple paths.

In this solution, the decision is made by the MPTCP/MPQUIC layer, which may be requested by the application to set up multi-paths. This solution applies to both Cellular Networks and WLANs for non-3GPP seamless mode. If ATSSS is supported in the Cellular Network, MPTCP can be regarded as a “higher-layer” implementation of ATSSS. This solution requires MPTCP or MPQUIC capability at both the vehicle and MSP Servers.

Solution 4.2 – Generic Multi-Access (GMA)/Multi-Access Management Services (MAMS)

GMA/MAMS is a mechanism to steer traffic over different connections. GMA convergence [12] provides a radio-agnostic IP-layer solution to support seamless traffic splitting, switching and steering over multiple connections/paths in the above examples. It can be configured and managed by over-the-top control messages, e.g., MAMS [13], and consists of the following two sublayers:

- Convergence sublayer: this layer performs multi-access-specific tasks, e.g., multi-link (path) aggregation, splitting/reordering, lossless switching/retransmission, fragmentation, concatenation, etc.
- Adaptation sublayer: this layer performs functions to handle tunneling, network layer security and NAT (network address translation).

The convergence sublayer operates on top of the adaptation sublayer in the protocol stack (see Figure 37), and uses a new lightweight trailer-based encapsulation protocol [12] for inserting control information, such as Sequence Number or Timestamp, into each data packet. The adaptation sublayer uses existing protocols such as IPsec, DTLS, UDP or NAT.

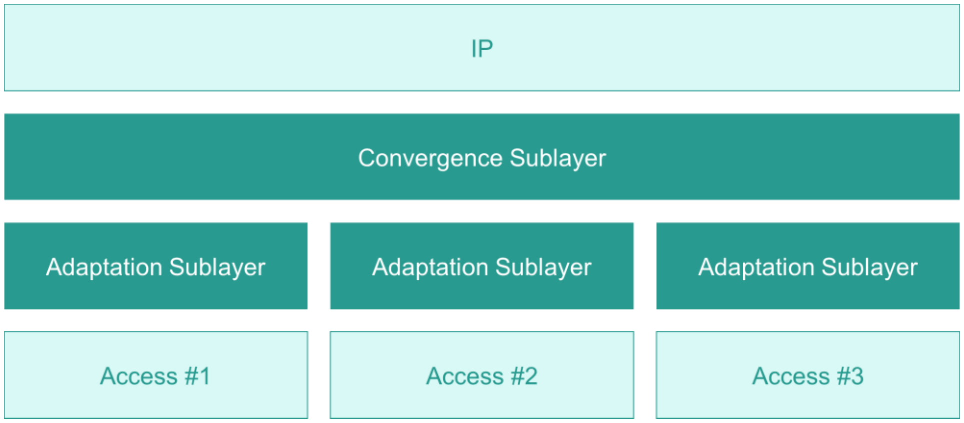

Figure 37. GMA protocol stack.

In this solution, the decision can be made by MAMS negotiation between the Network Connection Manager (NCM) and Client Connection Manager (CCM) or it can be implementation dependent, which further configures the GMA paths. This solution requires that the GMA layer be implemented on both the vehicle and MSP Servers, and it applies to both cellular and WLAN networks for non-3GPP seamless mode.

Solution 5 – Access Layer Solutions

Solution 5.1 – LTE WLAN Aggregation (LWA)

LWA is a 3GPP Release-13 feature to enable LTE and WLAN interworking at RAN level. The LWA radio protocol architecture for the scenarios of colocated eNB and WLAN Termination (WT) is shown in Figure 38 [14].

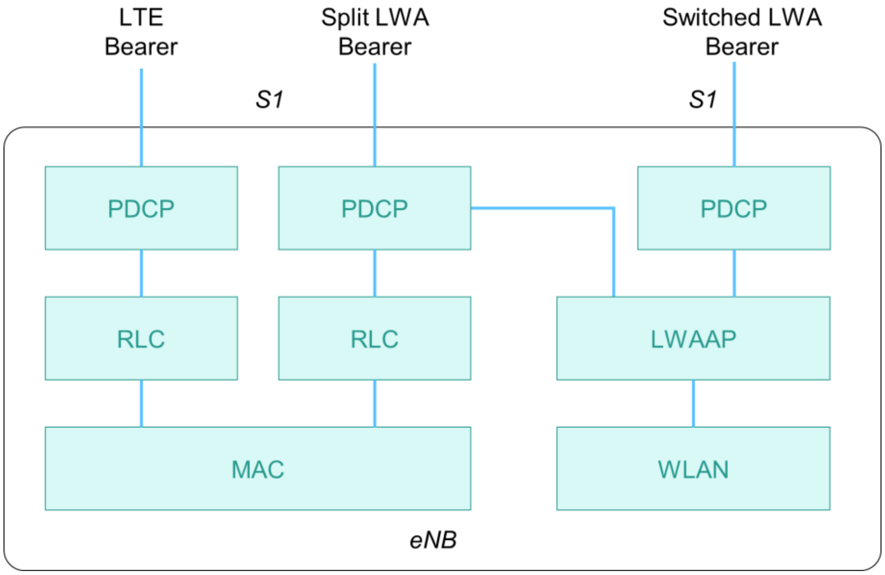

Figure 38. LWA radio protocol architecture stack in colocated scenario.

A sublayer LTE-WLAN Aggregation Adaptation Protocol (LWAAP) is introduced between a split LWA bearer and the WLAN. Functions of the LWAAP sublayer are performed by LWAAP entities. For an LWAAP entity configured at the eNB, there is a peer LWAAP entity configured at the UE. For all LWA bearers, there is one LWAAP entity in the eNB and one LWAAP entity in the UE. A new interface Xw is defined between the eNB and WT in the non-colocated scenario.

In the control plane, the eNB is responsible for LWA activation, deactivation and the decision about which bearers are offloaded to the WLAN. It does so using WLAN measurement information reported by the UE. Once LWA is activated, the eNB configures the UE with a list of WLAN identifiers (referred to as the WLAN Mobility Set) within which the UE can move without notifying the network.

In the data plane, for PDUs sent over a WLAN in LWA operation, the LWAAP entity, as specified in a 3GPP specification [14], generates an LWAAP PDU containing a Data Radio Bearer (DRB) identity, and the WT uses a defined sequence for forwarding the data to the UE over the WLAN.

In this solution, the decision is made by the eNB to select a split bearer based on RAN-level information such as measurements and the UE's preference. This solution only applies to an LTE network for 3GPP-based seamless mode between LTE and WLAN networks, and it requires an LWAAP entity configured at the vehicle and an eNB at the Cellular Network.

Solution 5.2 – LTE WLAN radio-level integration over IPSec tunnel (LWIP)

LWIP is also a 3GPP feature to enable LTE and WLAN networks to securely interwork at the RAN level. The LWIP feature allows UE in RRC_CONNECTED to be configured by the eNB to utilize WLAN radio resources via IPsec tunneling [16]. The end-to-end protocol architecture for LWIP is illustrated in Figure 39. Connectivity between the eNB and the LWIP-SeGW is provided by the Xw interface introduced in Solution 5.1.

The IP Packets transferred between the UE and the LWIP-SeGW are encapsulated using IPsec, as specified in a 3GPP specification [16], in order to provide security to the packets that traverse the WLAN. The IP packets are then transported between the LWIP-SeGW and eNB via the Xw interface. The end-to-end path between the UE and eNB via the WLAN network is referred to as the LWIP tunnel.

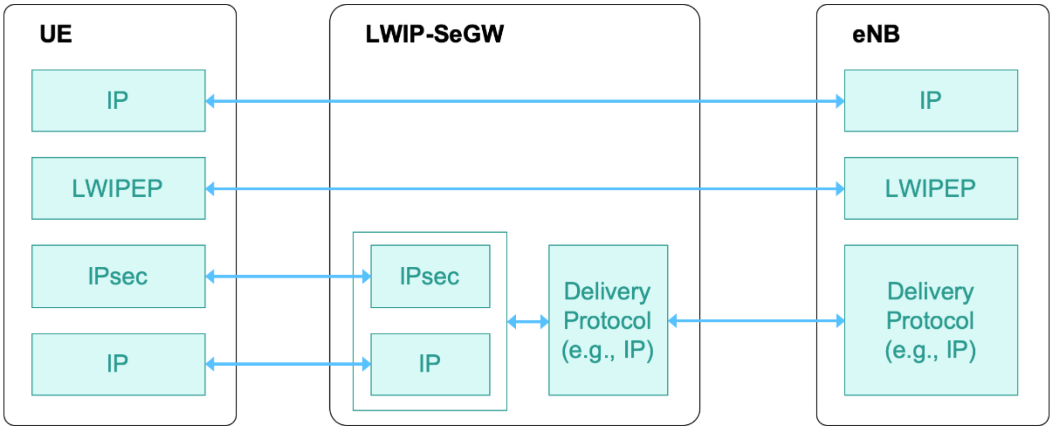

Figure 39. Protocol architecture for LWIP.

A single IPSec tunnel is used per instance of UE for all the data bearers that are configured to send and/or receive data over the WLAN. The data corresponding to each IPSec Tunnel is transported over the Xw interface on a single GTP-U tunnel. Each data bearer may be configured so that traffic for that bearer can be routed over the IPsec tunnel in only downlink, only uplink or both uplink and downlink over the WLAN. SRBs are carried over the LTE network only. The eNB configures specific bearers to use the IPsec tunnel.

For the downlink of a data bearer, the packets received from the IPsec tunnel are forwarded directly to upper layers. For the UL, the eNB configures the UE to route the uplink data either via LTE or WLAN networks using Radio Resource Control (RRC) signaling. If routed via the WLAN, then all UL traffic of the data bearer is offloaded to the WLAN.

In this solution, the decision is made by the eNB to select a split bearer based on RAN-level information such as measurements and the UE's preference. This solution requires LWIPEP entities configured at both the UE and the eNB, and requires the WLAN AP to be enhanced to support the 3GPP-defined Xw interface. This solution only applies to LTE networks for 3GPP-based seamless mode between LTE and WLAN networks.

Solution 5.3 – Access Traffic Steering, Switch and Splitting (ATSSS)

ATSSS provides a way to manage 3GPP access and non-3GPP access such as WLANs, and to manage the traffic from both networks when the UE alternates between these networks, as defined in a 3GPP specification [17]. The reference architecture is shown in Figure 40 [17]. The WLAN connects to the cellular core network by the Non-3GPP Inter-Working Function (N3IWF).

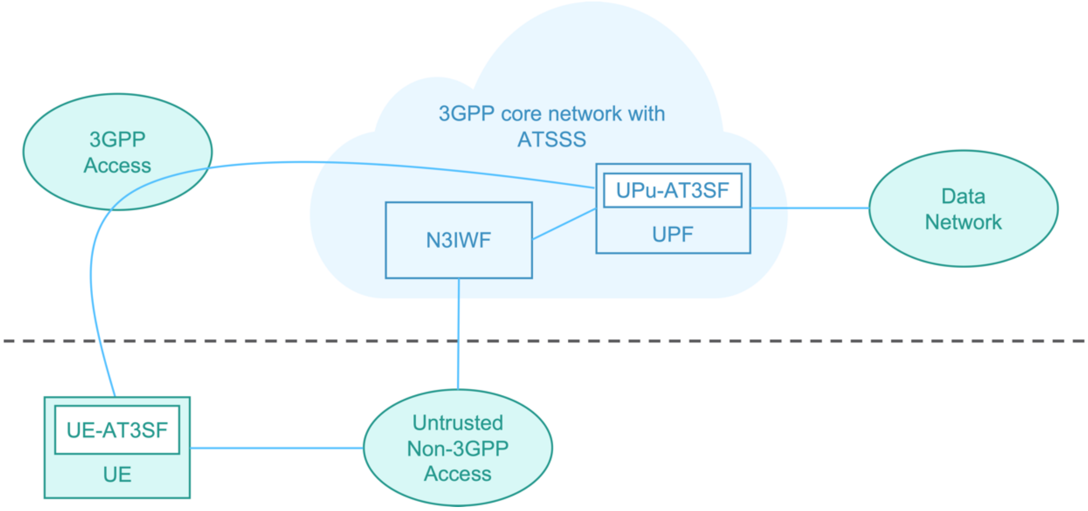

Figure 40. ATSSS reference architecture.

In ATSSS, a Multi-Access PDU (MA-PDU) session is created by bundling together two separate PDU sessions, which are established over different accesses. ATSSS policies are provisioned in the UE and the UPF, which can be generated in the PCF to provide the access switching rule and put one access in stand-by mode.

Within the same MA-PDU session, if MPTCP is enabled, it is possible to steer the MPTCP flows by using the MPTCP protocol (or the MPTCP function) and, simultaneously, to steer all other flows by using lower-layer steering functionality, called the ATSSS function. This is schematically illustrated in Figure 41 [17] for the UE. Note that the same set of ATSSS rules is applied to configure the MPTCP function and the ATSSS function. During the process of setting up the MA-PDU session, the UE shall exchange capability information about support of MPTCP with the core network as specified in a 3GPP specification [17].

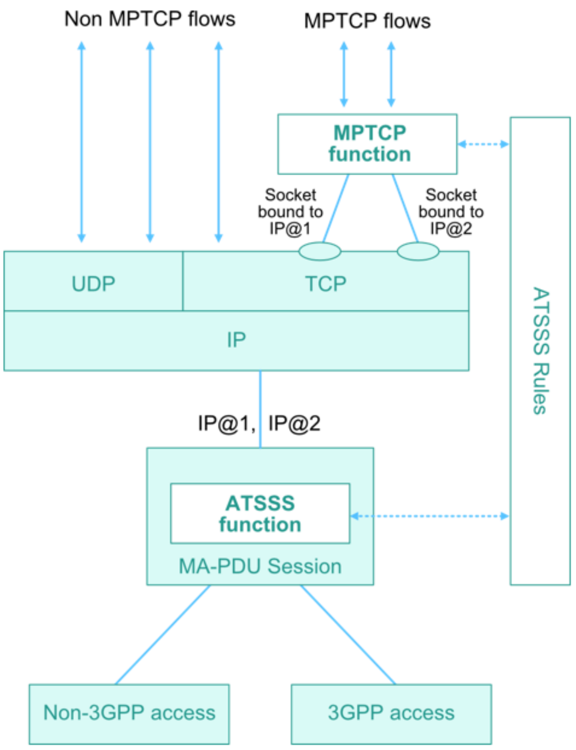

Figure 41. Example of UE supporting an MPTCP function and an ATSSS function.

In this solution, the UE or UPF can initiate the MA-PDU session to trigger ATSSS, and the traffic is managed by the ATSSS policies provided by the PCF. ATSSS is a solution to enable 3GPP-based seamless traffic steering over a cellular and a WLAN network. This solution applies to 5G networks and requires the UE and cellular core network to support ATSSS capabilities. MPTCP can be regarded as a higher-layer implementation of ATSSS.

Solution 5.4 – Multi Radio Dual Connectivity (MR-DC)

MR-DC is an extension of dual connectivity that allows UE to simultaneously connect to two nodes in the network [18]. In MR-DC, multiple Rx/Tx UE may be configured to utilize resources provided by two different nodes connected by non-ideal backhaul. One node acts as the Master Node (MN) and the other as the Secondary Node (SN). The MN and SN are connected via a network interface and at least the MN is connected to the core network. The MN may control the connectivity and data splitting toward the two nodes. Note that “non-ideal backhaul” generally means backhauls with latency of 5~30 ms or even higher.

MR-DC can generally increase user throughput, provide mobility robustness and support load balancing among eNBs/gNBs.

In this solution, the decision about where to steer the traffic is made by the MN based on information such as measurements. This solution requires MR-DC capability at the vehicle and the Cellular Network, and it applies to both LTE and 5G networks for 3GPP-based seamless mode.

3.4.3 Conclusions

We recommend the following combination of solutions for Access Network selection.

Connection selection

The AECC recommends taking the decision about connection selection in the system layer of the Vehicle System, as described in Solution 1.2, where the connection algorithm considers information such as signal strength, cost and policies. Applications running in the Vehicle System can provide policies to the connection selection function. The connection selection function then interacts with communication modules in the access layer to connect to and disconnect from available Access Networks. Solution 1.2 can reduce the complexity of the vehicle system by a unified interface to applications to avoid duplications of Access Network selection assistance functions at the application layer.

Traffic steering

When multiple connections are available and used, the Vehicle System steers traffic on the transport layer using MPTCP/MPQUIC, as described in Solution 4.1. This allows decoupling the traffic steering task from which Access Networks are used. Consequently, this solution works in the same way for different combinations of Access Networks, such as using multiple Cellular Networks, or a combination of a WLAN and a Cellular Network.

3.5 Provisioning and Configuration Update

3.5.1 Key Issue

The AECC System embraces high-volume data loads of different varieties. As the Vehicle System moves, the location, environment, network availability and generated data change dramatically. During the data exchange process among the Vehicle System, Access Networks and the MSP Servers, a large number of parameters and policies are involved. Some parameters and policies can be fairly static, but some can be very dynamic. In order to prepare the MSP Servers, Access Networks and the Vehicle Systems to meet AECC service requirements in a dynamic environment, provisioning and configuration update needs to be supported as shown in Figure 42. The AECC System should be aware of these changes either based on Access Network or vehicle reports, and adjust the policy/configuration accordingly. The Cellular Network can also update its related parameters and policies to the Vehicle System. In this section, the parameters and policies related to the AECC System are characterized and possible solutions are summarized.

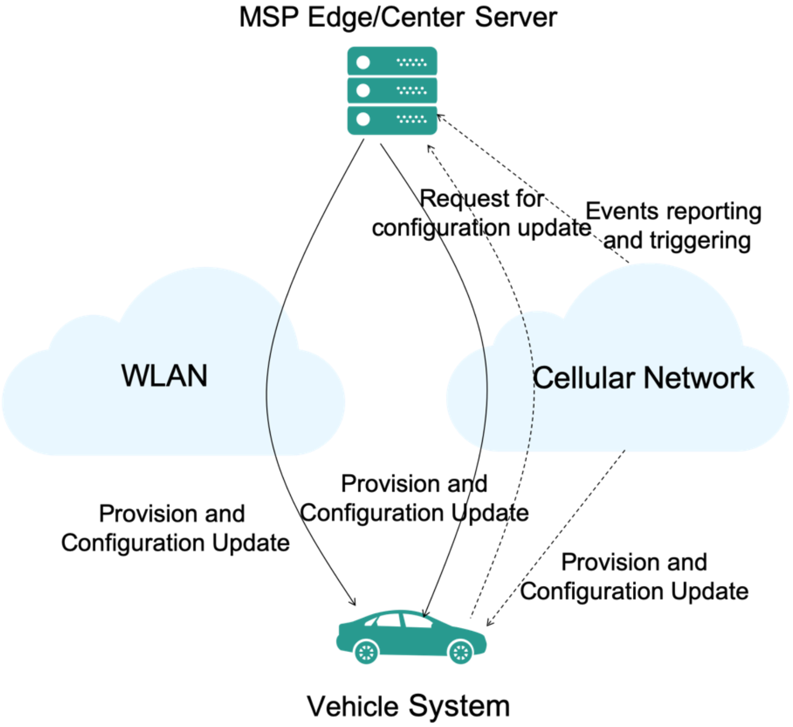
MSP Edge/Center Server
Request for configuration update
Events reporting and triggering
WLAN
Cellular Network
Provision and Configuration Update
Provision and Configuration Update
Provision and Configuration Update
Vehicle System

Figure 42. Provisioning and configuration update.

Provisioning and configuration for the Vehicle System mainly include provisioning parameters and policies; for example:

- Credentials and configurations of Access Networks, e.g., Access Point Name (APN) and Quality of Service (QoS).
- Non-Access-Network-related configurations, e.g., hostnames and addresses.

- Policies, e.g., edge offloading policies.

3.5.2 Potential Solutions

Solutions for provisioning and configuration update are listed below.

- Solution 1 – Pre-configuration
- Solution 2 – Configured through the Access Network
  - Solution 2.1 – Configured by cellular subscription
  - Solution 2.2 – Configured through the Cellular Network
- Solution 3 – Provisioned by the AECC Server at the Application Level
- Solution 4 – Provision through a Generic AECC Configuration Function
- Solution 5 – Provision a Bootstrap URL of Config Function
  - Solution 5.1 – Bootstrap URL provisioned by the DHCP server
  - Solution 5.2 – Pre-configured bootstrap URL

In Table 3, the solutions for provisioning and configuration update to the vehicle are summarized, with applicable examples of parameters and polices. Note that this is not an exhaustive list but is instead examples.

Table 3. Solutions for provision and configuration update.

Solutions	Examples of parameters and policies	Static or dynamic
Solution 1 – Pre-configuration	Credentials for cellular or WLAN networks from cellular or WLAN MNOs; IP addresses or domain names, URLs of potential MSP Servers, non-AECC servers from the AECC System; user preference from the user; APNs and service QoS from the Access Network and MSP	Generally static
Solution 2.1 – Configured by cellular subscription	Cellular subscription information such as device category, billing information, parameters to access the network, data rate, data transmission window, regional regulatory requirements agreed between MNOs and MSPs or manufacturers	Generally static
Solution 2.2 – Configured through the Cellular Network	Update UE-related information such as UE's behavior and its communication groups; negotiate a policy, such as a background data transfer policy; influence the traffic, such as data offloading, etc.	Generally dynamic
Solution 3 – Provisioned by the AECC Application Server at the Application Level	Application-level parameters on the Vehicle System from applications on the AECC servers	Both static and dynamic; out of the scope of the AECC
Solution 4 – Provision through the AECC Configuration Function	AECC System-related parameters such as IP addresses or domain names, URLs of potential MSP Servers, non-AECC servers; can integrate user preference	Both static and dynamic

Solution 5.1 – Bootstrap URL provisioned by the DHCP server	Appropriate parameters and policies in a configuration file sent together with the IP configurations when the vehicle first accesses the network	Both static and dynamic
Solution 5.2 – Pre-configured bootstrap URL	Appropriate parameters and policies in a configuration file sent as bootstrap URL to the vehicle	Both static and dynamic

3.5.2.1 Solution 1 – Pre-configuration

The vehicle can be pre-configured (e.g., by its manufacturer) with the policies for the AECC System, which may include:

- Credentials for cellular or WLAN network domains, agnostic about the AECC System, such as SIM card or WLAN account information. This provision information is expected to come from cellular or WLAN operators.
- Configurations for Mobility Services such as IP addresses or domain names, URLs of potential MSP Servers, non-AECC servers, agnostic about the cellular or WLAN network. This provision information is expected to come from the AECC System.
- Preferences at the application level that may have an impact on Mobility Services, such as user settings, account and billing information or data transmission preference. This information is expected to come from the user.
- Configuration for the cellular or WLAN network to fulfill Mobility Services such as APN and service QoS. This information is expected to come from the Access Network and MSP.

This solution includes the information that shall be pre-configured in the application or platform of the Vehicle System. This solution applies to both cellular and WLAN Access Networks.

3.5.2.2 Solution 2 – Configured through the Access Network

The solutions described in this section mostly target policies relating to Cellular Network communication.

Solution 2.1 – Configured by cellular subscription

Relatively static parameters and policies can be provisioned to the Vehicle System as part of the Cellular Network subscription information as agreed between the network operator and vehicle manufacturer (or MSP). This may include device category, billing information, parameters to access the network, data rate, data transmission window, regional regulatory requirements and so on. Subscription information is provisioned by the Access Network, which is different from Solution 1.

This solution applies to the cellular Access Network only.

Solution 2.2 – Configured through the Cellular Network

In one alternative version of this solution, the MSP Server provisions the Cellular Network via the SCEF/NEF over the T8 reference point or Nnef interface. The Cellular Network can then enforce the vehicle system’s policy via the UE’s configuration update procedure as defined in a 3GPP specification [22]. The parameters and policies can be updated by the MSP Servers on a periodic, event-triggered basis. This applies to the Access Network-related parameters and policies.

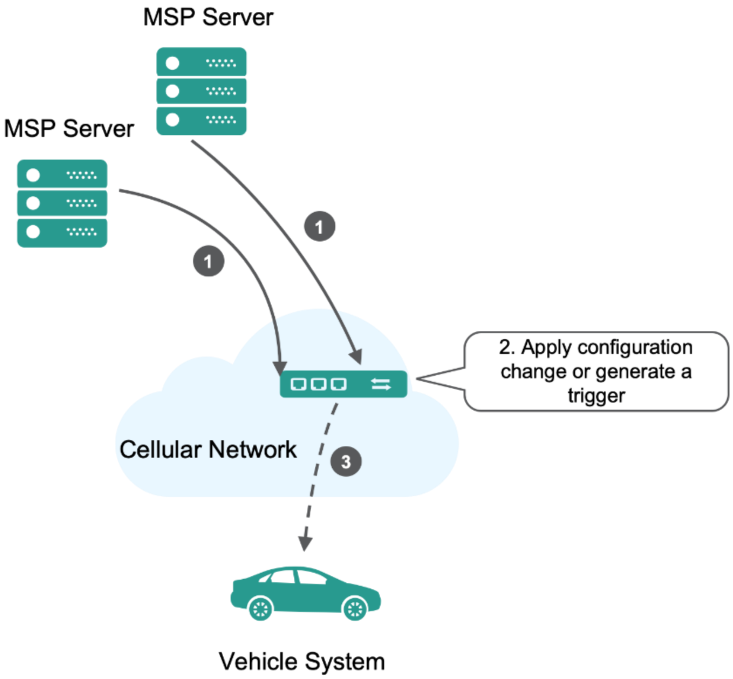

Figure 43. Procedure for provision configuration via Cellular Network.

The parameter and policy update enforced by the Cellular Network is shown in Figure 43. The procedure is described as follows:

1) The MSP Center Server or Edge Server sends configuration parameters through APIs to the SCEF/NEF in the Cellular Network via T8 or Nnef.
2) The SCEF/NEF applies configuration changes to Cellular Network entities, which may include a PCRF/PCF, HSS/UDM, UDR, etc. As stated in a 3GPP specification [22] clause 4.15.6, the MSP servers can leverage an AF to externally provision parameters to UDM/UDR.
3) The updated parameters are delivered to the Vehicle System via N1/S1 or Uu. This step is optional: that is, the Vehicle System does not need to be specifically configured; the policies will be applied in the network when the UE performs access or mobility-related procedures.

The SCEF/NEF can expose capabilities for MSP Servers to

- Update UE-related information such as UE’s behavior and its communication groups.
- Negotiate a policy, such as a background data transfer policy.
- Influence the traffic, such as data offloading, etc.

More policies are defined in a 3GPP specification [22]. Figure 44 shows the procedure for an AF requesting traffic influence on the UE by modifying session-related policies in the PCF.

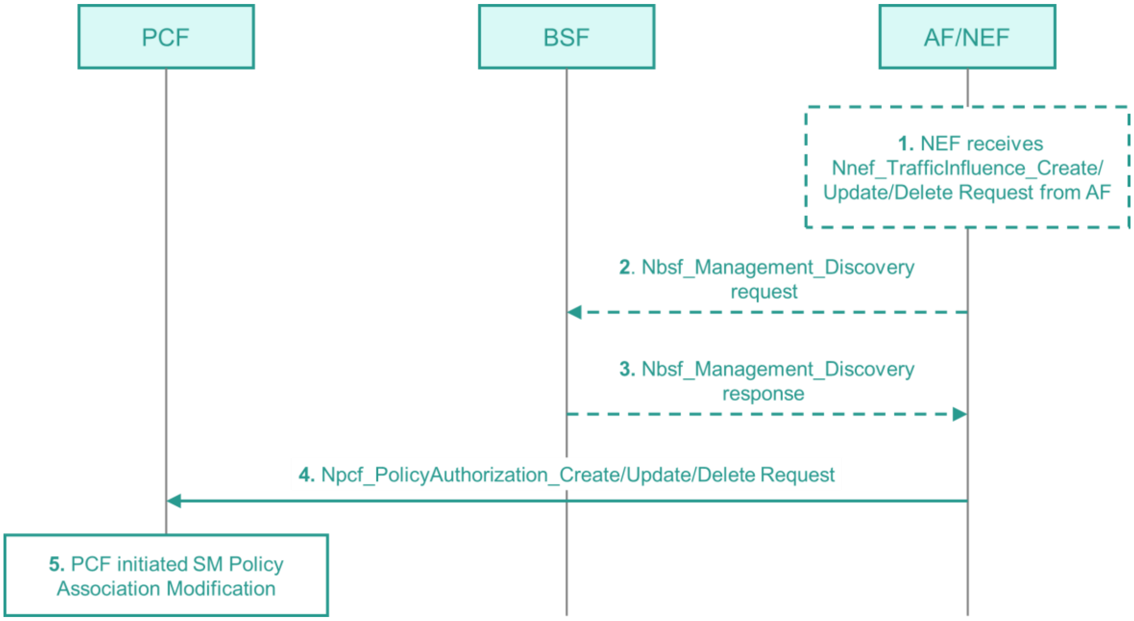
PCF
BSF
AF/NEF
1. NEF receives Nnef_TrafficInfluence_Create/ Update/Delete Request from AF
2. Nbsf_Management_Discovery request
3. Nbsf_Management_Discovery response
4. Npcf_PolicyAuthorization_Create/Update/Delete Request
5. PCF initiated SM Policy Association Modification

Figure 44. AF requesting a configuration/policy update.

1) The AF requests to influence the UE's traffic routing through the SCEF/NEF by the UE's address.
2) [Optional] The SCEF/NEF initiates the discovery of a related PCF by the UE's address.
3) [Optional] The Bootstrapping Server Function (BSF) provides the PCF's address in the discovery response.
4) The AF/NEF initiates the policy authorization to create/update/delete the requested policy.
5) The PCF starts to create/update/delete the related policies.

In addition, the PCF can update UE access selection and PDU Session selection-related policy information in the UE configuration, as shown in Figure 45.

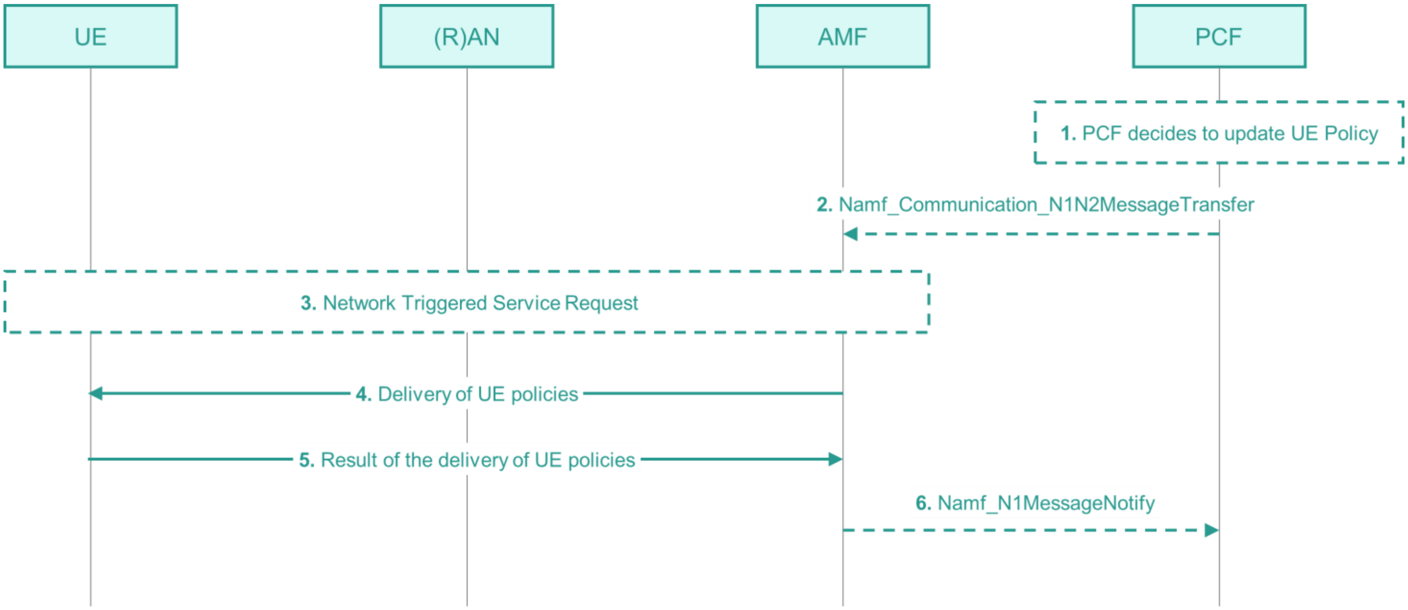
UE
(R)AN
AMF
PCF
1. PCF decides to update UE Policy
2. Namf_Communication_N1N2MessageTransfer
3. Network Triggered Service Request
4. Delivery of UE policies
5. Result of the delivery of UE policies
6. Namf_N1MessageNotify

Figure 45. UE configuration update procedure.

1) The PCF decides to update UE policy procedures based on triggering conditions.
2) The PCF invokes the Namf_Communication_N1N2MessageTransfer service operation provided by the AMF.
3) [Optional] The UE may need to be triggered to initiate a service request procedure for policy delivery.
4) The AMF delivers related policies to the UE.
5) The UE confirms the policy delivery.
6) [Optional] The AMF notifies the PCF about the policy delivery.

In another alternative version of this solution, the MSP Server requests an SMS trigger from the Cellular Network that includes the Vehicle System-related parameters. Upon receiving the MSP request (via the SCEF/NEF), the Cellular Network will generate the SMS and send it to the Vehicle System. The Cellular Network does not need to understand the configuration update. The procedure is described as follows.

1) The MSP Servers request a trigger from the Cellular Network that includes the Vehicle System-related parameters. The MSP request is sent through the NEF.
2) The Cellular Network generates an SMS with server-configured parameters.
3) The Cellular Network sends the SMS message to the Vehicle System.

Such a solution can be used to provision small configuration files to Vehicle Systems, for example.

3.5.2.3 Solution 3 – Provisioned by the AECC Server at the Application Level

At the application level, the parameter and policy provisioning is done through the interface between the applications residing on the Vehicle System (e.g., an AECC platform service or a third-party application hosted by the AECC System) and the applications residing on MSP Servers. If the applications running on the Vehicle System side and on the MSP Server side are third-party applications, the technical realization of the application layer interface is outside the scope of the AECC.

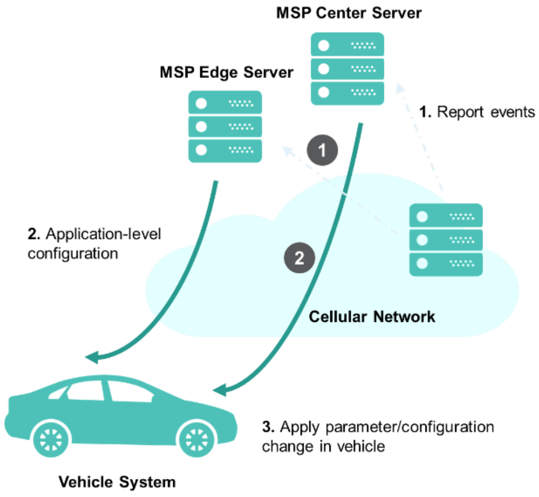
MSP Center Server
MSP Edge Server
1. Report events
1
2. Application-level configuration
2
Cellular Network
3. Apply parameter/configuration change in vehicle
Vehicle System

Figure 46. Application-level parameter/configuration update.

Figure 46 shows the application-level parameter and policy update.

1) The Cellular Network optionally reports network events to MSP Servers. The network events could have been configured previously through the NEF. When using other types of Access Network, entities within those Access Networks may also provide reports to MSP Servers.
2) The MSP Servers update parameters and configuration through the AECC application, which is agnostic about the change of Access Network. Alternatively, the MSP Servers can send a push notification with provisioned parameters, which has been discussed in Section 3.3.1, the Key Issue in Vehicle System Reachability.
3) The Vehicle System applies the related parameter change.

The monitoring events can include loss of connectivity, location reporting, number of instances of UE presented in a geographical area, etc. as defined in Section 4.15.3 in a 3GPP specification [22].

3.5.2.4 Solution 4 – Provision through a Generic AECC Configuration Function

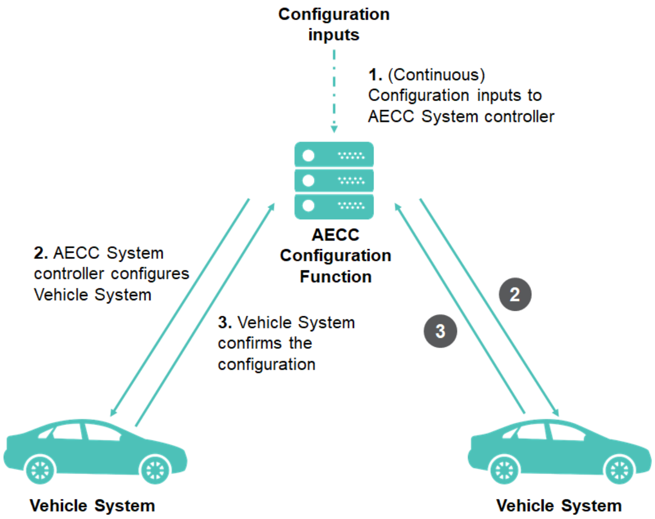
Configuration inputs
1. (Continuous) Configuration inputs to AECC System controller
AECC Configuration Function
2. AECC System controller configures Vehicle System
3. Vehicle System confirms the configuration
2
3
Vehicle System
Vehicle System

Figure 47. Configuration by the AECC configuration function (AECC configuration function-triggered).

The Vehicle System is configured through the AECC configuration function if there is an interface between the AECC configuration function and the Vehicle System. In one alternative version of the solution, the Vehicle System configuration is triggered by the AECC configuration function, as shown in Figure 47. The steps are as follows:

1) [Optional] Configuration parameters are input into the AECC configuration function. The inputs can be provided by external users or applications through interfaces provided by the AECC configuration function. This step can be event-triggered or periodic.

2) The AECC configuration function provides provision and configuration updates to the Vehicle Systems.

3) The Vehicle System confirms to the AECC configuration function when the configuration updates are received.

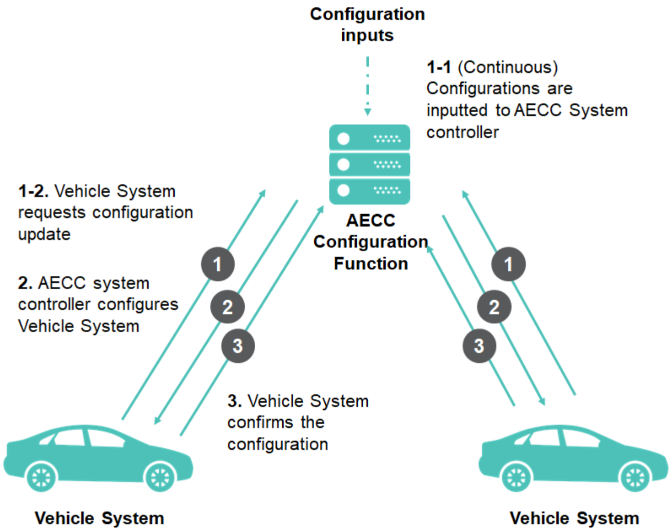
Configuration inputs
1-1 (Continuous) Configurations are inputted to AECC System controller
AECC Configuration Function
1-2. Vehicle System requests configuration update
2. AECC system controller configures Vehicle System
1
2
3
3. Vehicle System confirms the configuration
Vehicle System
Vehicle System

Figure 48. Configuration by the AECC configuration function (Vehicle System-triggered).

In another alternative version of the solution, the Vehicle System configuration is triggered by the configuration function, as shown in Figure 48. The steps are as follows:

1-1)[Optional] Configuration parameters are input into the AECC configuration function. The inputs can be provided by external users or applications through interfaces provided by the AECC configuration function. This step can be event-triggered or periodic.

1-2)The Vehicle System sends a configuration update request to the AECC configuration function. Note that the order of sequence between Step 1-1 and Step 1-2 is interchangeable.

2) The AECC configuration function provides provision and configuration updates to the Vehicle System.

3) The Vehicle System confirms to the AECC configuration function when the configuration updates are received.

3.5.2.5 Solution 5 – Provision a Bootstrap URL of Config Function

This solution includes a two-step provisioning mechanism, where the first step targets provisioning a URL (bootstrap URL) from which a more extensive configuration file is downloaded in the second step. The configuration server can be outside the AECC System.

Solution 5.1 – Bootstrap URL provisioned by the DHCP server

In this solution, the location of the configuration file is sent together with the IP configurations when the Vehicle System first accesses the network. This scheme works for configuration that applies directly to the Vehicle System upon the Vehicle System's startup. The DHCP bootstrap option is described in IETF RFC 2132 [19] with additional information in RFC 5970 for DHCPv6 [20]. One condition for applying this solution is that the DHCP server needs to be informed of the config function location. If the DHCP server and the server hosting the config function are provided by different service providers, inter-service provider agreement is needed.

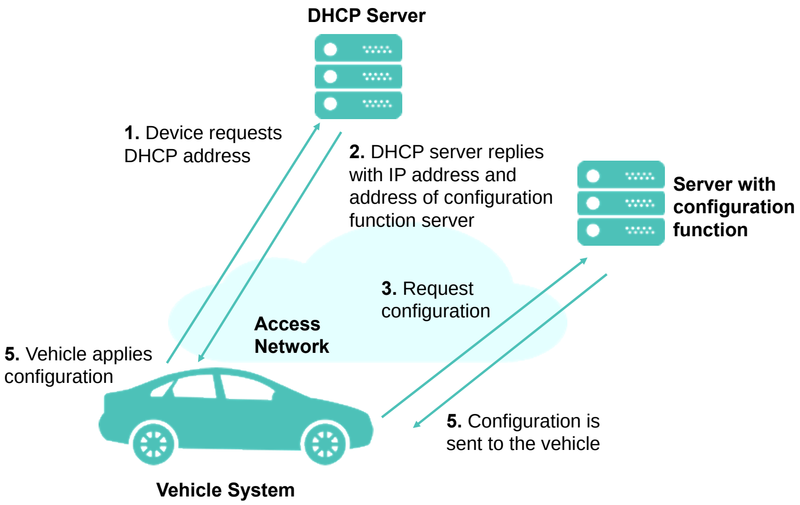
DHCP Server
1. Device requests DHCP address
2. DHCP server replies with IP address and address of configuration function server
Server with configuration function
3. Request configuration
Access Network
5. Vehicle applies configuration
5. Configuration is sent to the vehicle
Vehicle System

Figure 49. Configuration file distribution through the DHCP.

Figure 49 shows configuration file distribution using the DHCP.

1) The Vehicle System requests the IP address from the DHCP server.

2) The DHCP server sends IP configurations along with a config function location.

3) The vehicle system accesses the configuration file from the location provided previously by the DHCP server.

4) The configuration file server sends the configuration file to the vehicle system.

5) The vehicle system applies the configurations.

Solution 5.2 – Pre-configured bootstrap URL

In this scheme, the location of the configuration file server is pre-configured in the Vehicle System and the configuration files are downloaded when there is connectivity to the configuration function. Typically, the URL would contain a hostname that needs to be resolved using DNS, before contacting the config function.

Here, “pre-configured” does not necessarily mean permanently hard-coded, but that the bootstrap URL is always available in the Vehicle System, possibly by having an initial hard-coded URL, which can be updated periodically during the lifetime of the Vehicle System using other provisioning mechanisms.

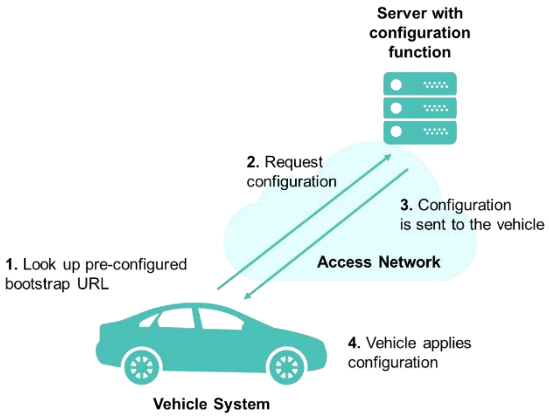
Server with configuration function
2. Request configuration
3. Configuration is sent to the vehicle
Access Network
1. Look up pre-configured bootstrap URL
4. Vehicle applies configuration
Vehicle System

Figure 50. Configuration file at pre-configured location.

Figure 50 shows the configuration look-up procedure.

1) The Vehicle System looks up the bootstrap URL.

2) The Vehicle System requests the configuration file from the pre-configured location.

3) The configuration file server sends the configuration file to the Vehicle System.

4) The Vehicle System applies the configuration.

While the described procedure assumes a Vehicle System as target for the provisioning, the solution is also applicable for other entities, such as MSP Servers.

This solution applies to both cellular and WLAN Access Networks.

3.5.3 Conclusions

We recommend the following solutions for the respective layers:

- For provisioning and configuration updates in the system layer, the AECC recommends using Solution 1 pre-configuration for static parameters and policies and for initialization. Solution 5.2 (Pre-configured bootstrap URL) is recommended for semi-static provisioning and configuration updates, possibly updating pre-configured parameters. As a more advanced method, Solution 4, Provision through a Generic AECC Configuration Function, can be used for semi-static and dynamic provisioning and configuration updates, given that corresponding interfaces are available and configured.
    - One example of provisioning in the system layer would be provisioning an MSP Server FQDN to a Vehicle System, which might be pre-configured initially (Solution 1), with the option to update the FQDN later using Solution 4 and/or Solution 5.2.
- For provisioning and configuration updates in the Access Network layer, in the case of Cellular Networks, we recommend Solution 2.1 (Configured by cellular subscription), and Solution 2.2 (Configured through the Cellular Network), depending on the parameters. The condition for applying Solution 2.2 is that the T8 (in EPS) or Nnef/N33 (in 5GS) interfaces and related APIs may be needed in the access layer. Specific mechanisms for provisioning concrete parameters regarding the Cellular Network are specified by 3GPP. For

any Access Network type, Solution 1 (Pre-configuration) is recommended for static parameters, and for default values.

  - APNs and DNNs are pre-configured in the Vehicle System (Solution 1).
  - UE categories are configured as part of the cellular subscription (Solution 2.1).

Provisioning in the application layer (i.e., from the MSP Server application layer to the Vehicle System application layer) is outside the scope of the AECC.

3.6 Opportunistic Data Transfer

3.6.1 Key Issue

Network capacity planning has become a major challenge, especially for MNOs, due to the exponential increase in mobile data traffic. Meanwhile, new vertical markets such as connected vehicles will further drive mobile traffic growth to a new level. It causes a critical need for MNOs and MSPs to provide sufficient network capacity to meet the traffic demands of new mobility services.

On the other hand, data traffic of non-latency-sensitive mobility services, such as vehicle data collection [5], could be transferred opportunistically, which implies that it will be delivered in a lower priority without affecting other services delivered by the network. The goal of opportunistic data transfer is to deliver certain types of delay-tolerant data, such as vehicular data collection, without degrading the service quality of other data traffic.

Note 1: due to the different dynamics of mobile data traffic in different time-scales, opportunistic data transfer could be managed differently. For instance, mobile data traffic usually fluctuates drastically in small-scale time periods (e.g., a second), while it exhibits a comparatively static pattern in large-scale time (e.g., a day or week). Different solutions can be used and combined to make use of these two effects.

Note 2: Opportunistic Data Transfer shall also apply to other Access Networks defined in the AECC System, such as WLANs. However, as the capacity issues are more urgent to solve with respect to Cellular Networks, this key issue focuses on the Cellular Network.

3.6.2 Potential Solutions

The following solutions are described for this key issue.

- Solution 1 – Access Control and Barring
  - Solution 1.1 – Application-specific Congestion control for Data Communication (ACDC)
  - Solution 1.2 – Unattended Data Traffic (UDT)
- Solution 2 – Background Data Transfer (BDT)
  - Solution 2.1 – 3GPP Network-based Background Data Transfer (BDT)
  - Solution 2.2 – 3GPP UE-based BDT
- Solution 3 – Dynamic Policy Adaptation

3.6.2.1 Solution 1 – Access Control and Barring

Access control methods are generally considered in order to block or defer certain user groups/services at base stations according to the access class pre-assigned to different users and services (e.g., data and voice) when the network is overloaded. The basic operation is that the UE will block or defer the traffic transmission request according to the broadcasted access control policy and parameters from base stations. There are many access control mechanisms that are defined in 3GPP specifications [21], including Access Class Barring (ACB), Service Specific Access Control (SSAC), Extended Access Barring (EAB), Smart Congestion Mitigation (SCM), Application-specific Congestion control for Data Communication (ACDC) and Unattended Data Traffic (UDT). By applying these mechanisms only to the non-latency-sensitive traffic, disturbing other communication is avoided, and opportunistic data transfer is achieved.

Solution 1.1 – Application-specific Congestion control for Data Communication (ACDC)

In 3GPP Release 13, ACDC is standardized, and applicable to solve the opportunistic data transfer problem. ACDC is an access control mechanism for the operator to allow/prevent new access attempts from particular operator-identified applications in the UE in idle mode. ACDC does not apply to the UE in connected mode. In ACDC, the network can handle access control policy per application category, thus preventing network congestion due to the large amount of delay-tolerant data transfer from vehicles.

Solution 1.2 – Unattended Data Traffic (UDT)

UDT is defined in 3GPP as data traffic that the user is unaware he/she initiated; for example, based on the screen/keypad lock being activated, length of time since the UE last received any input from the user, or specific types of applications. The IE (Information Element) *SystemInformationBlockType2* contains radio resource configuration information that is common for all UE. The UE can restrict access attempts due to unattended data traffic via the radio resource configuration.

3.6.2.2 Solution 2 – Background Data Transfer (BDT)

The Cellular Network and MSP Servers can negotiate a time slot (e.g., non-busy hours) to deliver non-time-critical data via the method of background data transfer. It is a proactive approach to solving this key issue, which needs advanced functions of the Cellular Network, such as APIs for scheduled transmission. Such methods would typically be used for bulk data distribution or upload.

Solution 2.1 – 3GPP Network-based Background Data Transfer (BDT)

3GPP specifies that a northbound interface is between the core network entities and third-party application servers using the Service Capability Exposure Function (SCEF, EPC) and Network Exposure Function (NEF, 5GC). This interface specifies APIs that allow a third-party application server to access the network service and capability provided by 3GPP network entities. It is named as the T8 reference point in LTE. In a 5GS, the NEF reuses most of the T8 APIs. An AF could negotiate a BDT policy through the SCEF/NEF and then related policies can be generated in the PCF to be enforced by other network functions, such as AMF for access control or SMF for session management, as shown in Figure 51.

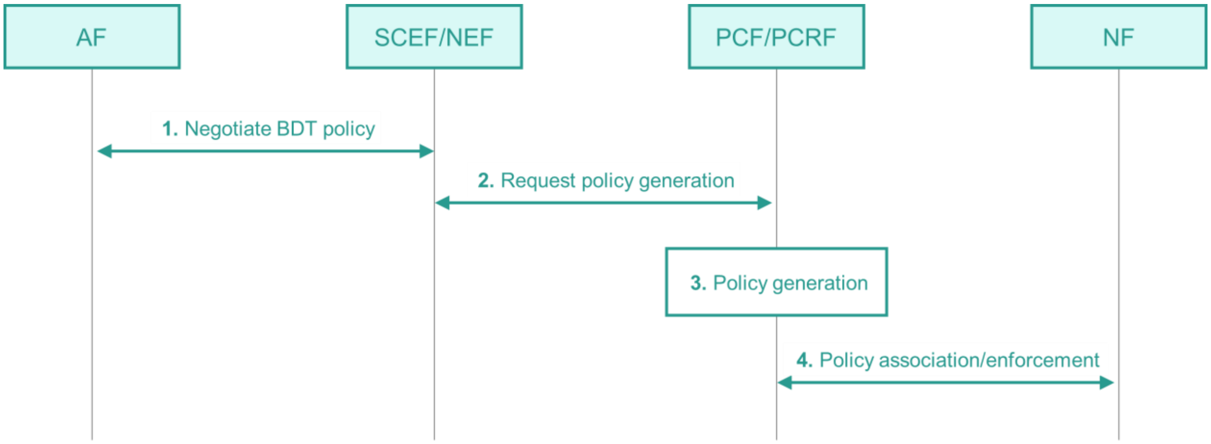

Figure 51. Negotiation for future background data transfer in Cellular Networks.

The workflow is as follows:

- Step 1: The AF requests to negotiate BDT policy with the SCEF/NEF based on different parameters.
- Step 2: The SCEF/NEF requests policy creation from the PCF/PCRF.
- Step 3: BDT policy generation in the PCF/PCRF.
- Step 4: Policies are applied to related Cellular Network functions.

Opportunistic Data Transfer policy can be based on BDT policy. A transfer policy consists of a recommended time window for the background data transfer, a reference to a charging rate for this time window and optionally a maximum aggregated bitrate.

The start/stop can also be assisted by SCEF/NEF event monitoring, as shown in Figure 52.

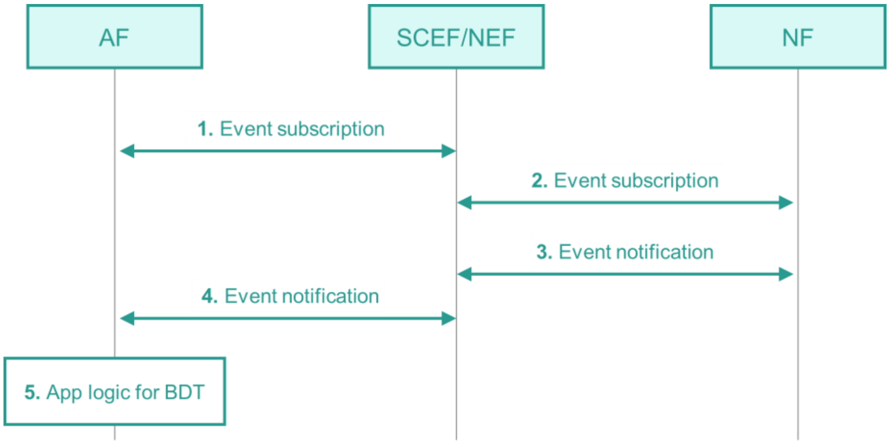

Figure 52. Cellular event triggering for dynamic BDT.

The workflow is as follows:

- The AF requests event monitoring capability from the SCEF/NEF for a vehicle or a vehicle group.

- The SCEF/NEF subscribes to the vehicle's network status.
- A network event is detected at the NF and notified to the SCEF/NEF.
- The AF is notified of the event.
- The AF applies application-related logic for BDT.

The network events can include:

- Vehicle loss of connection.
- Vehicle's location.
- Network load.

For example, the ReportingNetworkStatus API can expose the network status to the MSP Server, and thus the MSP Server can manage the data transfer more efficiently. The congestion value in the API indicates whether there is congestion and what the congestion level is. The mechanism can be: first, the MSP Server subscribes to the network status report according to periodic mode or based on events; second, the MSP Server receives the congestion indicator from the network status report via the cellular northbound APIs. Finally, according to the congestion indicator, the MSP Server can decline or continue data transfer to the vehicle system.

Solution 2.2 – 3GPP UE-based BDT

UE can support the 3GPP system to optimize its use of the wireless resources based on policies (such as time window, network area information, etc.) for background data transfer if the policies are sent from the core network or the servers lying outside the 3GPP system.

This solution enables the 5GS to support delivery of the policies for background data transfer to the UE. In particular, the 5G network is able to provide policies for background data transfer to the UE, so that the 5GS can optimally use the control plane and/or user plane resources by directly managing the transmission behavior of the UE. If Background Data Transfer policy information (e.g., time window and location criteria) is not going to be sent to the UE as part of the URSP rule, then, at the time the background data transfer is about to start, the AF provides for each instance of UE the Background Data Transfer reference ID together with the AF session information to the PCF (via the N5 interface). The PCF retrieves the corresponding background transfer policy from the UDR and derives the PCC rules for the background data transfer according to this transfer policy.

The policy information content will define the time window and location criteria that need to be met for background data transfer. It defines how and when the PCF activates/distributes the policies related to background data transfer to the UE. A single dedicated PDU Session might be used for background data transfer that is established and released based on the background data transfer policies. In addition, the policy should consider avoiding a large quantity of UE sending data and/or signaling to the network at the same time.

In order to ensure that data is sent only in the designated time window, an additional (implementation-dependent) interface between modem and applications is needed in the Vehicle System, to deny delivery requests outside this time window. Further, the URSP rules need to be configured in such a way that delivery attempts outside the designated time window are not sent on different PDU Sessions by default (e.g., due to “match all” URSP rules).

3.6.2.3 Solution 3 – Dynamic Policy Adaptation

This solution assures that data is received on time but at the same time exploits the relaxed delay constraints of non-real-time communication. Besides describing the data communication aspect, this solution also presents means for the application and network to negotiate background data transmission policies according to a subscription agreement.

It is unlikely that all MNOs will support the same policy enforcement mechanism. The choice of enforcement might have an impact on the client implementation in the vehicles. As a result, harmonizing the policy provisioning and activation procedures is important in order to minimize client implementation variants.

The section below, after definitions of some key terms, illustrates how policies are provisioned and activated, while the material following gives one example of how these policies can be enforced. In these sections, 5GS terms are used, but the solution is also applicable to EPS, when using eNB/P-GW/PCRF instead of gNB/UPF/PCF.

Policy provisioning and activation

An application client in the Vehicle System, which can, for example, be an HD Map application, determines the need for fetching the next HD Map tile, depending on the current vehicle location, vehicle speed and vehicle route.

The 5G Media Streaming Application Function authorizes application policy requests. An MSP Server may have the choice of different policies and the application decides, based on need, which policy to activate for the upcoming transaction. While it is specified in the context of media streaming, it is applicable for any traffic type. Different collaboration options are supported. The 5GMS AF is operated by the MNO (trusted AF, Figure 53) or by the MSP (external AF, Figure 54).

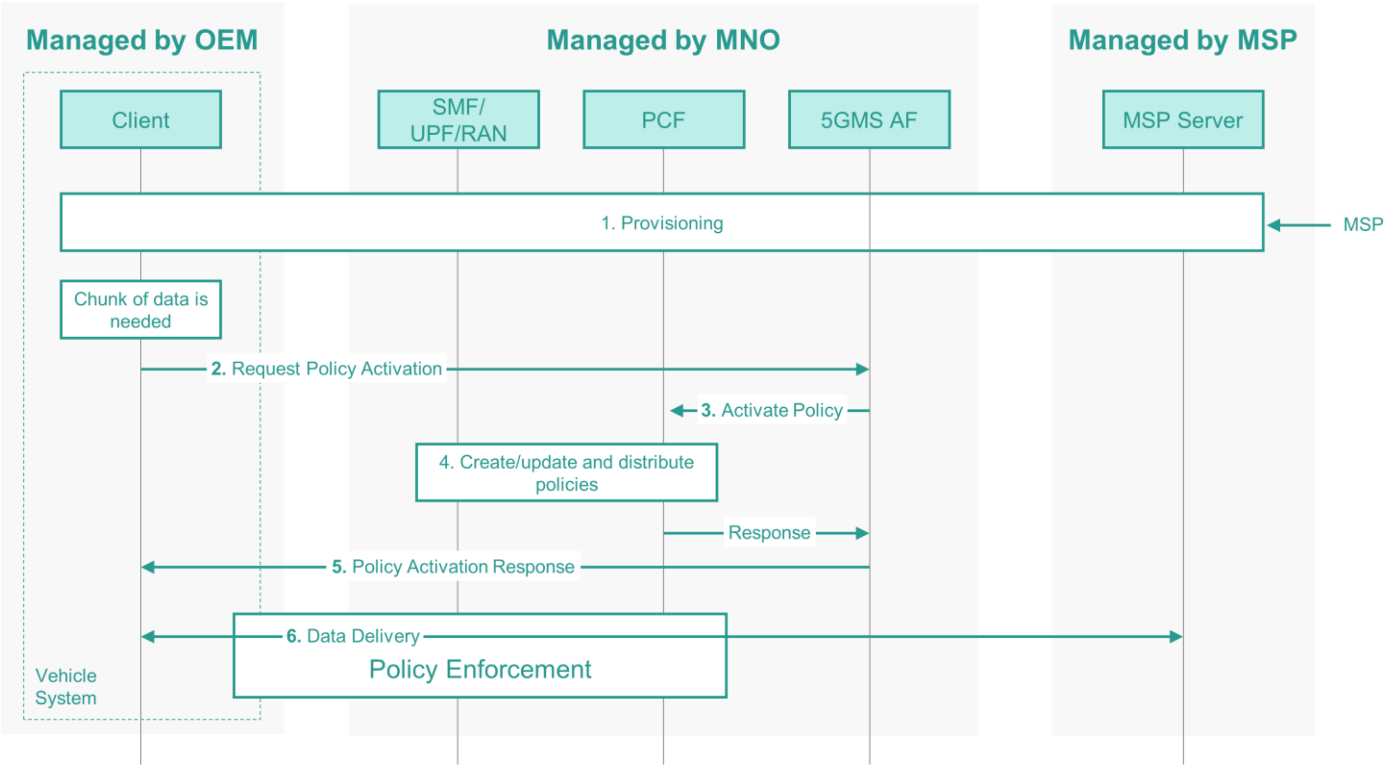

Figure 53. Dynamic policy adaptation – a 5GMS-AG managed by the MNO.

The workflow is as follows:

1 The MSP agrees with a Mobile Network Operator (MNO) to use a set of different policies. The policies to be used can be anchored directly as part of an SLA or, in a more modern approach, using dynamic provisioning mechanisms and exposure APIs. As a result, the MSP Server has a list of policies to choose from and is authorized to trigger them.

2 The client detects that a chunk of data should be uploaded or downloaded within a certain time frame (e.g., provided in seconds). The client sends a Policy Activation Request for a certain policy to the 5GMS AF.

3 The 5GMS AF activates the selected policy in the 5GS via the NEF. The NEF interacts with the PCF. If the 5GMS AF is a trusted entity for the 5GS (Figure 53), it can instead interact with the PCF directly, and thus has more possible interactions.

4 The PCF creates/updates and distributes corresponding dynamic PCC rules to relevant SMFs and other network functions. Depending on which type of enforcement the mobile network operator uses, other entities are involved. For enforcement on the MAC layer in the RAN, the gNB and UE are involved as well.

5 The 5GMS AF sends a response, indicating that the policy request was granted, and potentially includes policy constraints.

6 The client and MSP Server exchange data (upload and/or download), while the mobile network takes care of the policy enforcement.

Steps 2-5 might be repeated during delivery, if the need to adapt the policy arises (e.g., when changing from a lower-than-best-effort policy back to a best-effort or even prioritized policy).

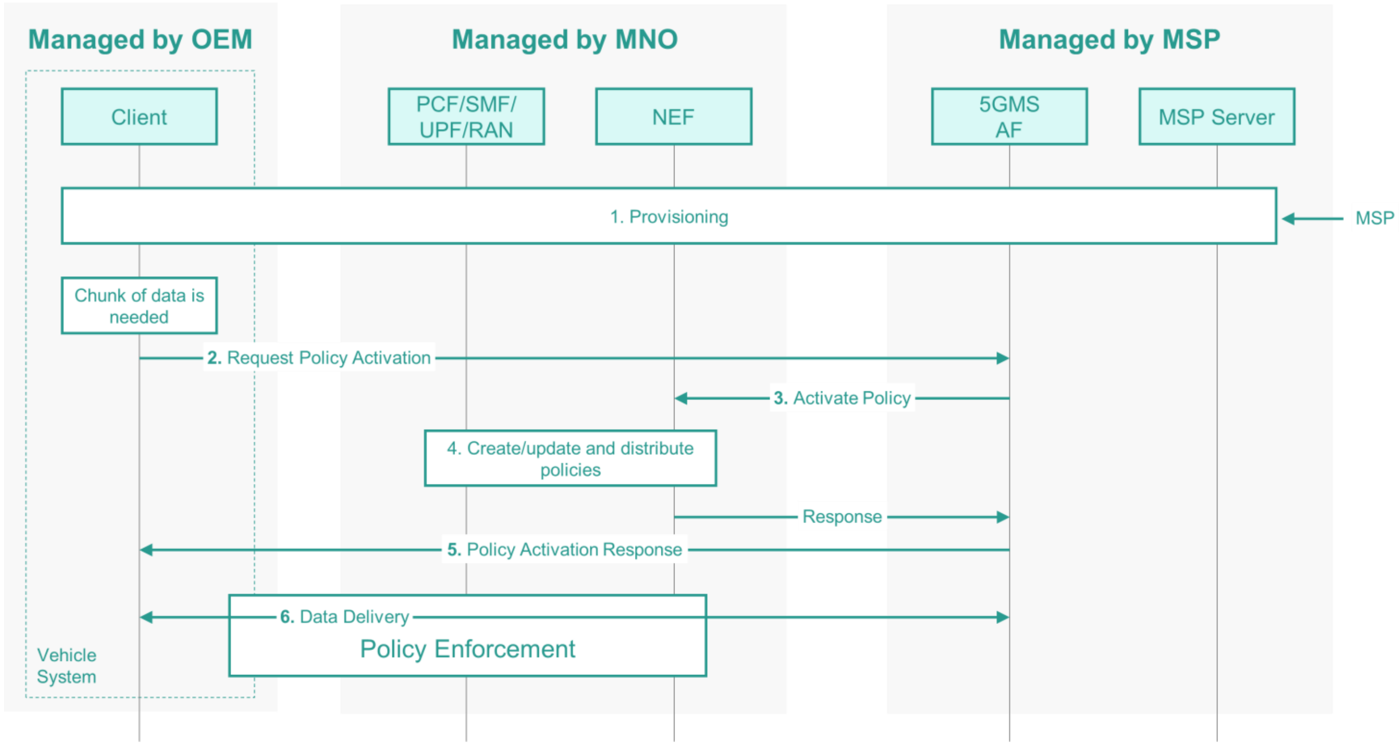

Figure 54. Dynamic policy adaptation – a 5GMS AF managed by the MSP.

Policy Enforcement

Several enforcement mechanisms are feasible, and different mobile network operators will have different preferences. Examples would be:

- Using the built-in QoS framework of 4G and 5G networks.
- Using a TCP proxy for traffic shaping.
- Using a QUIC spin bit for pausing the client or server.
- Traffic shaper on IP layer (e.g., leaky bucket).

As one example for policy enforcement, traffic shaping using a transparent TCP proxy is illustrated in Figure 55, where the TCP Proxy detects high load conditions (e.g., by monitoring RTTs) and throttles the corresponding TCP connection. To this end, the TCP proxy would be dynamically configured to adapt read/write operations for reducing throughput. Using this mechanism, the throughput can be throttled to a negligible minimum, by forcing the receive window to zero at the proxy (i.e., stop forwarding any data). In this case, the client only sees a bad connection, and it is up to the client how to react to this (pause delivery, try to reconnect, change Access Network, etc.). In any case, no specific requirements are put on the client, and it is up to the MNO how to implement the details, without impact on other entities.

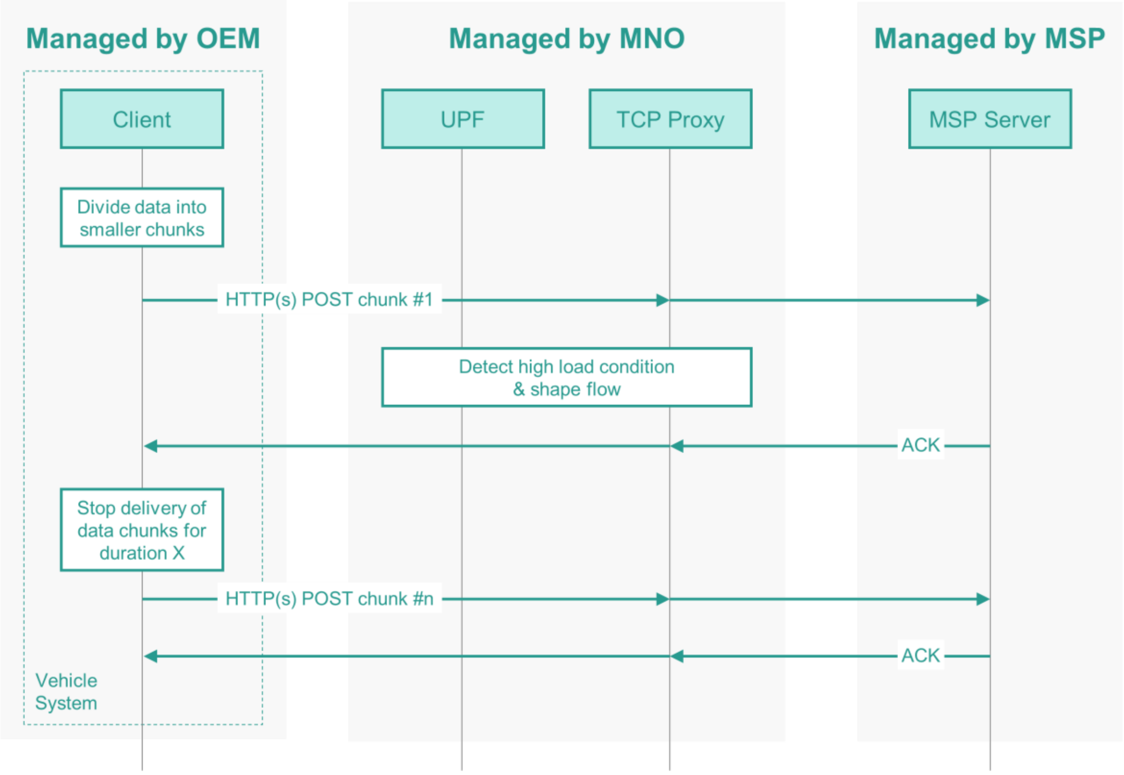

Figure 55. Policy enforcement using a transparent TCP proxy.

3.6.3 Conclusions

The combination of Solutions 2.2 (3GPP UE-based BDT) and 3 (Dynamic Policy Adaptation) is recommended for the future deployment of 5G systems. This solution combination can deal with various requirements on delay tolerance ranging from seconds to days per application. Solution 2.2 and Solution 3 provide the ability for MNOs to dynamically configure policies for UE and easily differentiate between applications. It also provides flexibility for MNOs to optimize efficiency by considering radio resource availability and network conditions. While these solutions provide the most appropriate functionality, their adoption is currently limited; however, since it is targeting the new deployment of 5G systems with less impact on UE and RAN, and because policy negotiations between MSPs and MNOs are based on standardized interfaces, adoption should not be a problem.

In the case of existing deployed EPS, industry adoption and solution impact on UE and RAN will play a significant role. The combination of Solutions 1.1 (Application-specific Congestion Control for Data Communication) and 2.1 (3GPP Network-based Background Data Transfer) is recommended to provide opportunistic data transfer on PDN Connections of specific APNs pre-defined by MNOs. The benefit of this combination is its high industry adoption, which implies less deployment effort. However, it comes with certain limitations on flexibility for controlling data transfer per application (by mapping application flows to APNs in the Vehicle System) and optimizing network resources.

Solution 1.1 can only control data transfer based on preconfigured and limited number of application categories, and Solution 2.1 policy enforcement at the network side is not as resource efficient as Solution 2.2. Consequently, upgrading deployed EPS to adopt Solution 2.2 and Solution 3 is recommended in the long run.

4 A Path Forward for New Solutions

The sets of solutions for the key issues presented in this version of the AECC technical report further detail the path on how to drive data to the edge from a networking and distributed computing perspective. The AECC believes that this proposal should play a significant role in supporting new automotive service scenarios.

In the future, the AECC will continue to identify new key issues and design solutions around distributed computing architecture and connectivity solutions to meet the growing data volumes with high demands on security, sovereignty and efficiency of the delivered data of the automotive industry.

4.1 New Key Issues

The AECC plans to study more issues in order to propose potential solutions. Two examples are given below.

4.1.1 Service Continuity

In the AECC distributed computing architecture, a Vehicle System may be served by different MSP Servers and communication anchor points where handovers may be triggered during mobility.

Different services put different requirements on service continuity: some may require keeping transport layer sessions undisrupted, while for other delay-tolerant services this may not be required. A challenge is that there are communication sessions on several levels that have different impacts on service continuity;

- Access layer session: a Cellular Network or WLAN could be used between the Vehicle System and the serving MSP Servers, and mobility may cause handover to occur between different Access Networks, or within the same Access Network.
- Transport layer session: on the transport layer there is typically a transport session established between the Vehicle System and the MSP Server, which can be disrupted by the transition within and between Access Networks or due to change of the Vehicle System network address.
- Application layer session: on the application layer, stateful sessions may also be affected during mobility.

The AECC System should be defined such that adequate service continuity can be preserved during handovers.

4.1.2 Data Identification

In the AECC System, different applications exist, and they may use different traffic types to communicate through the system. The communication endpoints, and possibly also other entities along the delivery path, should be able to identify specific data types for associating them with the corresponding application, or for application-specific treatment. In addition, assuming that some grouping is in place for Vehicle Systems (e.g., by vehicle vendor and model), the communication endpoints should be enabled to differentiate the same type of data from different groups of Vehicle Systems.

4.2 Distributed Computing: New Paradigm for Mobility Services

Manufacturers that are challenged with supporting globally distributed fleets of Connected Vehicles will face the issues related to gathering and processing data. The solution proposed by the AECC envisages the ability to perform localized data processing, closer to the location of the vehicle systems, rather than attempting to pull data back to a few centralized locations.

Vehicle manufacturers may choose to deploy their own data processing and computing facilities around the world or may look to leverage computing facilities offered by an ecosystem of partners. Further, in an ecosystem where Mobility Services such as HD Mapping and Intelligent Driving may be provided not only by the vehicle manufacturer but also by other entities, referred to as Mobility Service Providers (MSPs), how can those MSPs identify locations where their applications can be deployed? The challenge that is to be addressed is how to access the computing facilities in a common manner, abstracting away from the individual implementations that a computing facility provider may offer, as well as being able to discover and utilize the Mobility Services hosted within the AECC System.

The distributed computing architecture described in the following sections outlines an approach where the AECC System provides services that expose information about available computing facilities and the Mobility Services running on those facilities.

More details relating to the architecture, operational models and associated key issues will be presented in future technical reports.

4.2.1 Distributed Computing Architecture Reference Model

In order to create a platform to support the Mobility Service scenarios, Computing Infrastructure Providers with appropriate resources will expose information about their resources to an AECC Service. The AECC Service provides an implementation-agnostic method for utilizing the capabilities of the infrastructure in order to manage the lifecycle of applications and other services. Such information may include computing availability, storage capacity, network functions and so on. In addition, the information may include the notion of “location.” The AECC Service interface may expose this information in the form of service offerings.

With information provided via the AECC Service, Mobility Service Providers will be able to determine where to place applications in order to meet the requirements of their customers.

When applications are instantiated on the computing facilities, information will be exposed to the AECC Service. Such information may include the set of applications currently available, the state of the applications and so on. It does not provide application-specific interfaces to each application, although it should provide information about instantiated applications, including their supported interfaces.

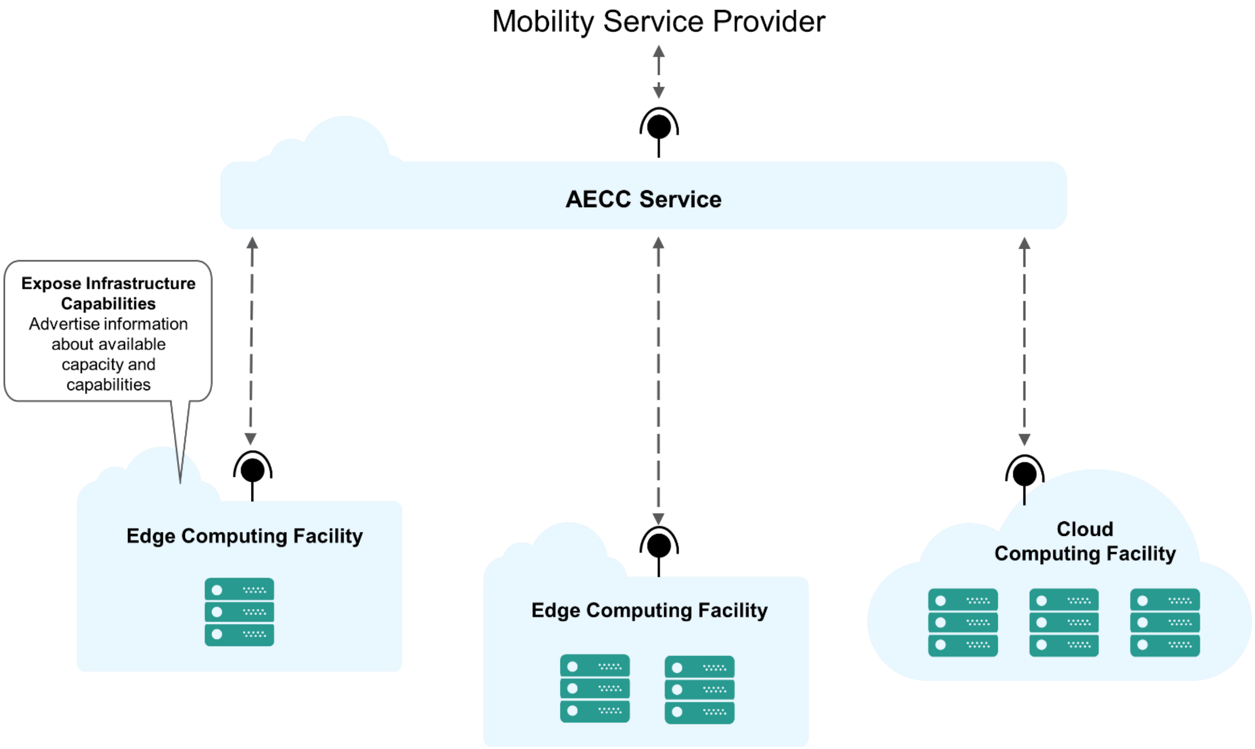
Mobility Service Provider
AECC Service
Expose Infrastructure Capabilities
Advertise information about available capacity and capabilities
Edge Computing Facility
Edge Computing Facility
Cloud Computing Facility

Figure 56. Infrastructure capabilities.

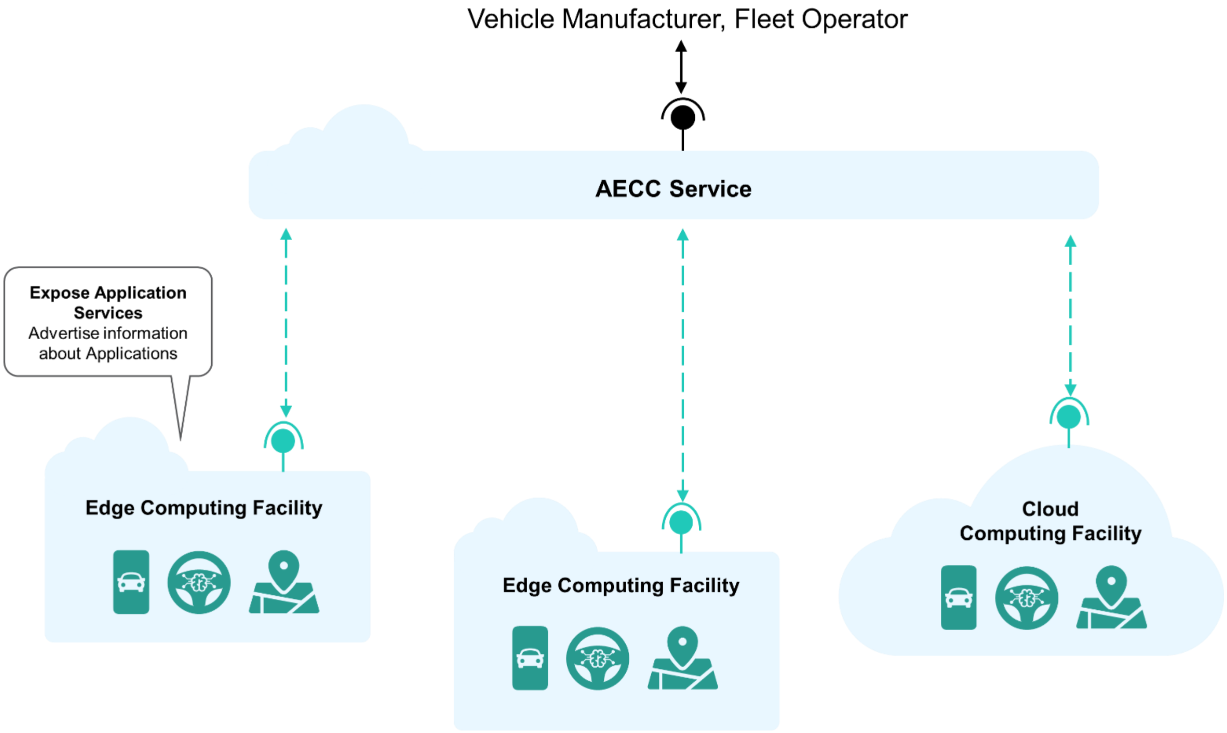
Vehicle Manufacturer, Fleet Operator
AECC Service
Expose Application Services
Advertise information about Applications
Edge Computing Facility
Edge Computing Facility
Cloud Computing Facility

Figure 57. Application services.

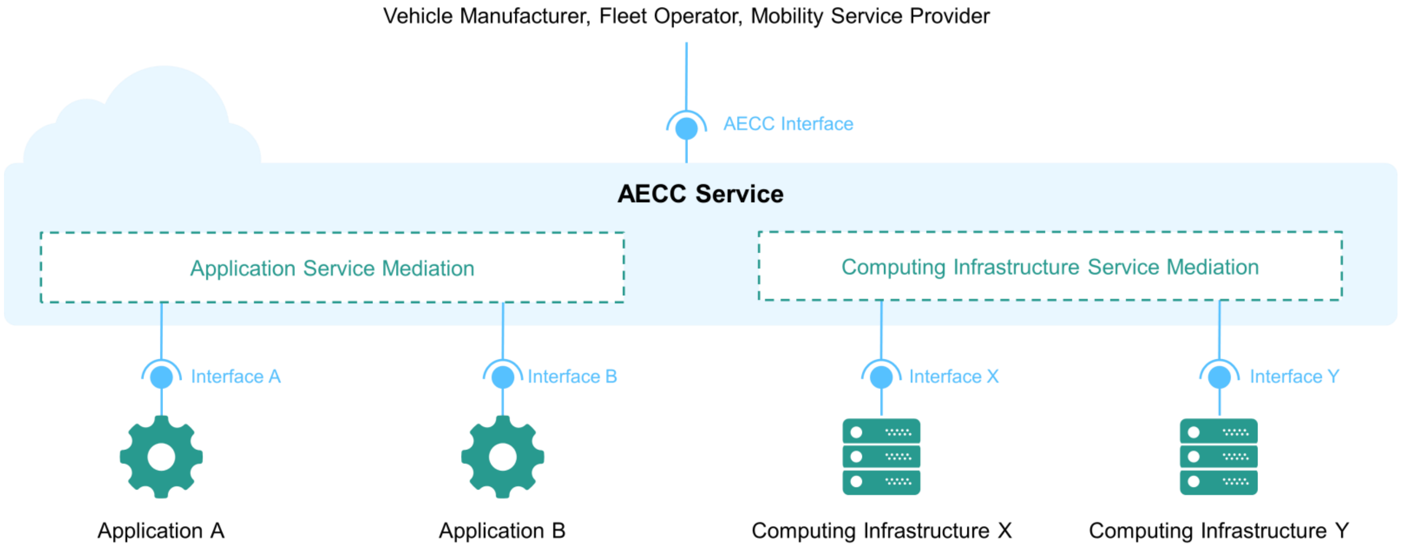
Vehicle Manufacturer, Fleet Operator, Mobility Service Provider
AECC Interface
AECC Service
Application Service Mediation
Computing Infrastructure Service Mediation
Interface A
Interface B
Interface X
Interface Y
Application A
Application B
Computing Infrastructure X
Computing Infrastructure Y

Figure 58. An AECC Service.

The AECC Service interface provides access to both Computing Infrastructure service information and Application service information and is responsible for mediation to the underlying infrastructure. For example, a Computing Infrastructure operator will expose information toward the AECC Service. This interface is likely to be proprietary, subject to the underlying system that the Computing Infrastructure Operator is using; for example, VM-based or container-based implementations. Computing Infrastructure information will be mediated by the AECC Service and exposed to Mobility Service Providers wishing to utilize the infrastructure. The MSP will use the AECC Service in order to perform application lifecycle operations (deploy, operate, monitor, terminate). A vehicle manufacturer or fleet operator will be able to use the AECC Service in order to obtain information about the set of Application Services that are being provided but will not (necessarily) be able to obtain information about the underlying infrastructure.

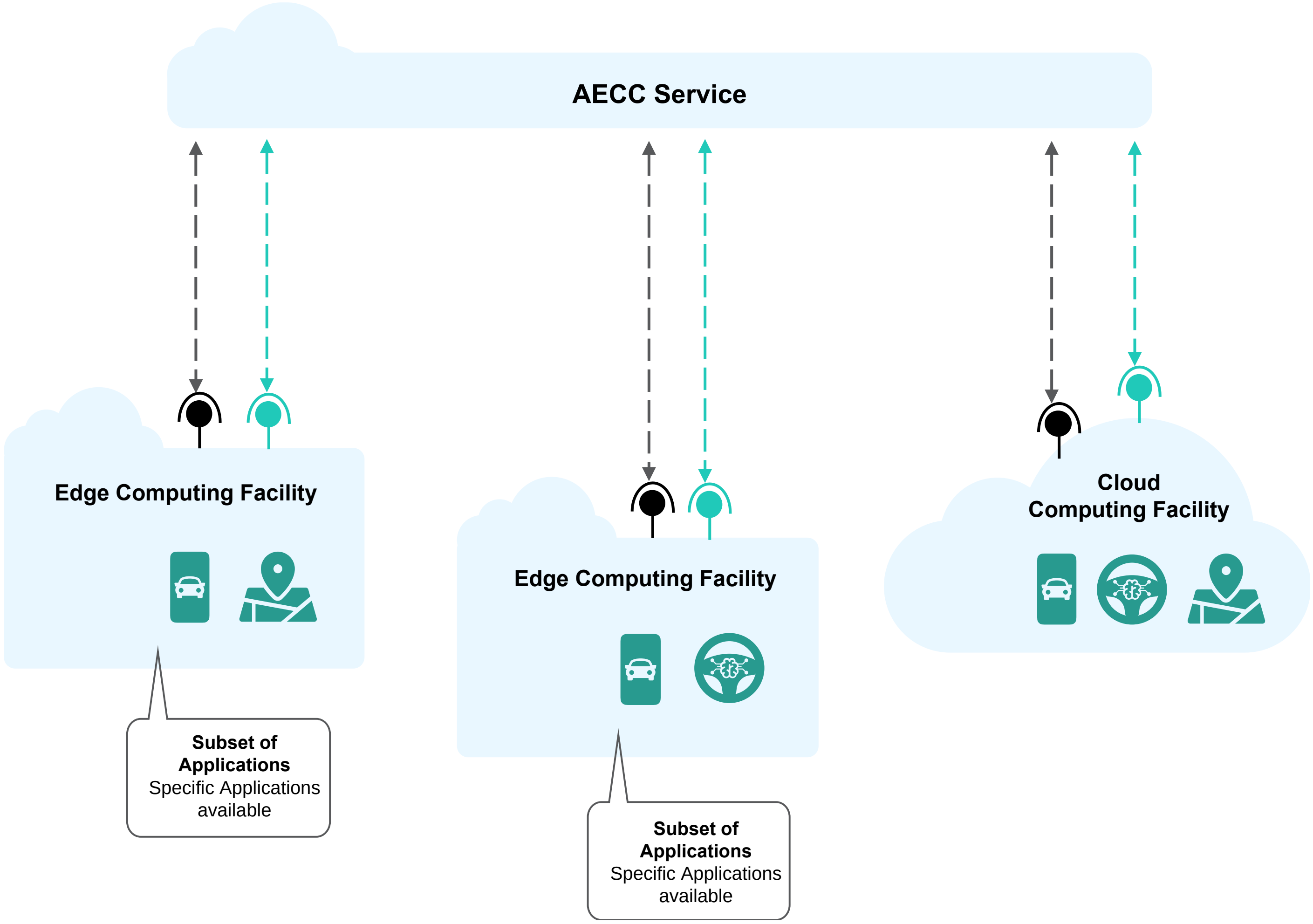

Figure 59. A subset of applications.

It is expected that the set of resources available to the AECC System will vary from facility to facility and will depend on the capabilities offered by the implementing ecosystem. An Edge computing facility is likely to have reduced capacity when compared to a Cloud computing facility. Similarly, it is expected that some applications will be present and instantiated in some facilities but not in others, as shown in the diagram above.

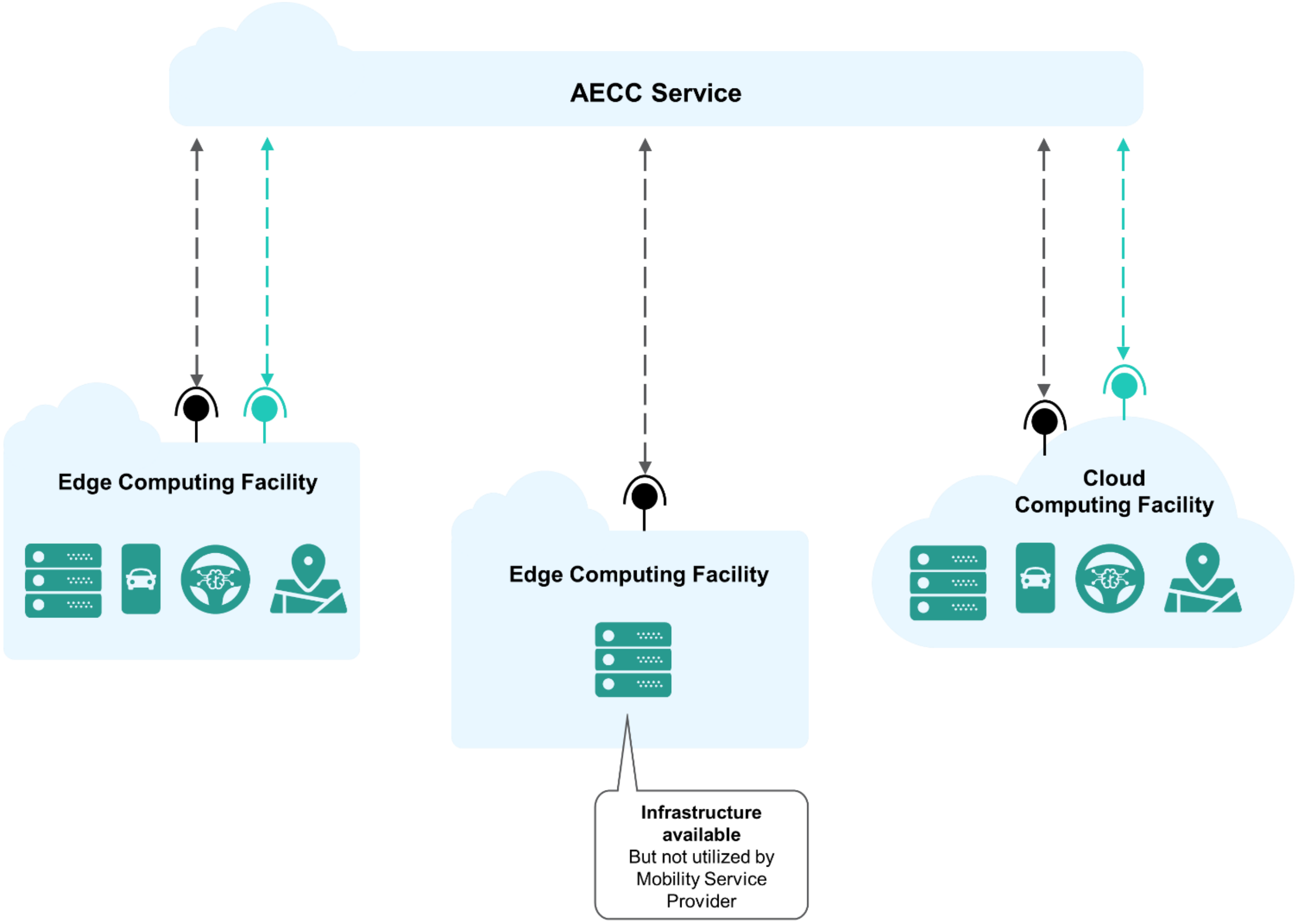

Figure 60. Unutilized capacity.

With the application lifecycle expected to include application instance instantiation and termination, it is expected that at some points in time infrastructure may be unutilized. The infrastructure will remain as part of the overall AECC System, ready for an application to be deployed at some future point.

As shown in Figure 60, the AECC System will allow AECC stakeholders to access information regarding the AECC cloud infrastructure and the applications that are deployed through the Expose Infrastructure Capabilities and Expose Application Services respectively.

4.2.2 Infrastructure Resources and Services

While such services rely on their underlying infrastructure, the implementation details of how or where a service is instantiated are generally out of scope. Rather, the AECC is concerned with the level of service being delivered. Thus, location is primarily important for its effects on the delivered service.

Two categories of service are particularly relevant. In the first are services or applications that provide support for the mobility features of a vehicle, or for the transportation and business functions that affect the infrastructure that supports the use of vehicles. In the second are infrastructure services that support the execution of those applications. Infrastructure services operate at many different levels of implementation, including bare metal, virtualization-based systems and container-based systems.

The distinguishing ability of an infrastructure service is to run other services or applications.

For AECC purposes, there are several roles that infrastructure services participate in: Center, Edge and Roadside Unit (RSU).

Center services support the aggregate services provisioned by Mobility Service Providers. Edge services support a subset of services that support the services exposed by the Cloud Computing Facility. There may be many Edge services, each handling a portion of the work required to support the Center services. Both Center and Edge services are implemented on real or virtual servers within computing facilities.

Composition of an AECC System may be iterative so that that a service acting as a Center may at the same time be acting as an Edge to another Center service.

Roadside services are similar to Edge services, with the difference that they are located within roadside facilities, rather than in data center facilities. Applications running on a Roadside service can have a temporal relationship to services running on a vehicle.

4.2.3 Infrastructure Resources Model

The AECC System is intended to be able to support Vehicle Systems from multiple vehicle manufacturers and applications from multiple Mobility Service Providers operating on an AECC System simultaneously. The model is intended to accommodate a number of different approaches to performing the core task of creating a distributed computing environment where applications can be executed in a distributed manner.

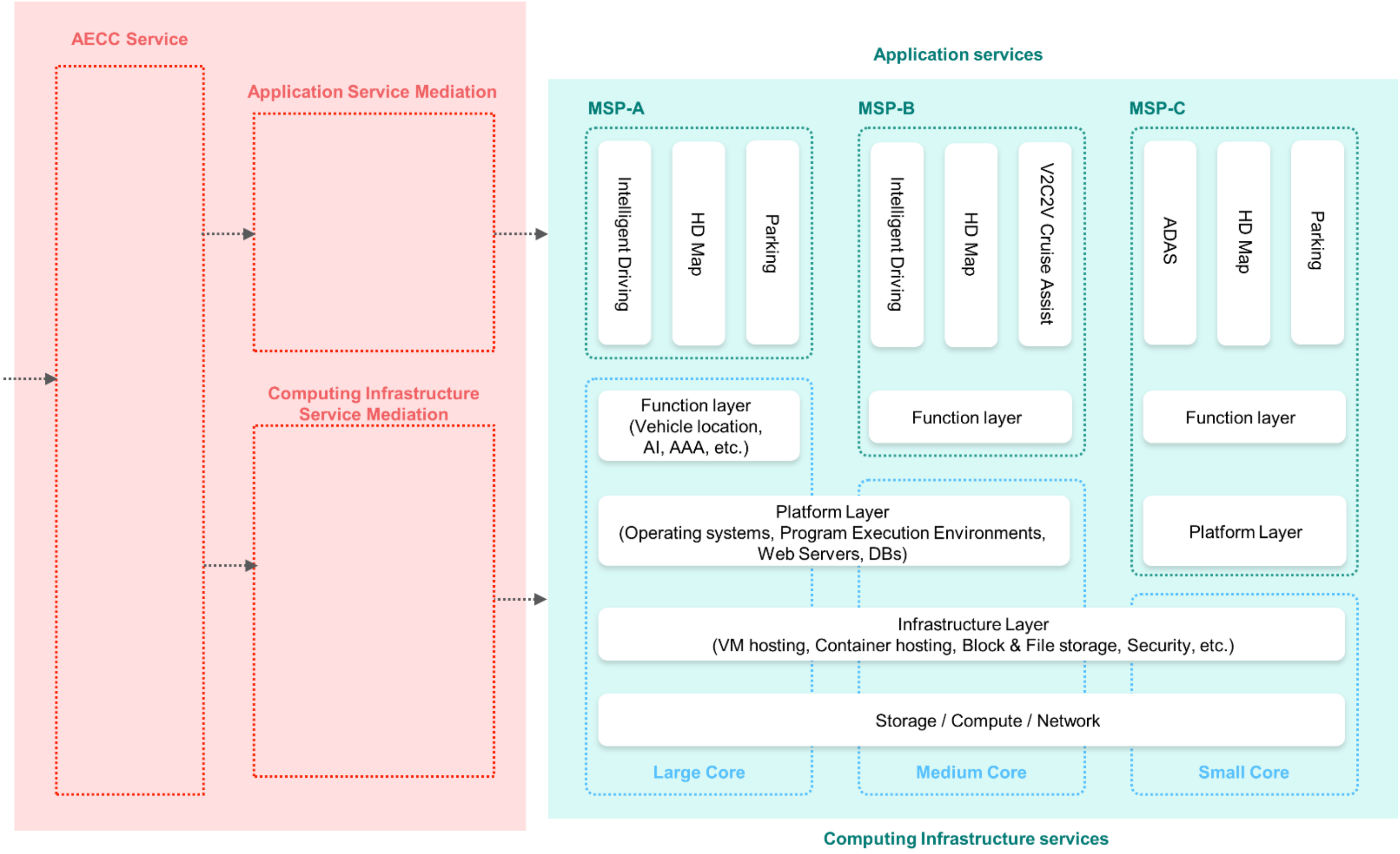
AECC Service
Application Service Mediation
Computing Infrastructure Service Mediation
Application services
MSP-A
Intelligent Driving
HD Map
Parking
MSP-B
Intelligent Driving
HD Map
V2C2V Cruise Assist
MSP-C
ADAS
HD Map
Parking
Function layer (Vehicle location, AI, AAA, etc.)
Function layer
Function layer
Platform Layer (Operating systems, Program Execution Environments, Web Servers, DBs)
Platform Layer
Infrastructure Layer (VM hosting, Container hosting, Block & File storage, Security, etc.)
Storage / Compute / Network
Large Core
Medium Core
Small Core
Computing Infrastructure services

Figure 61. Infrastructure resources model architecture.

The architecture in the figure above illustrates the potential set of capabilities that an AECC System would offer for use by Mobility Service Providers.

Computing Infrastructure operators would provide resources to an AECC System. In the "small-core" approach, Computing Infrastructure operators offer a base set of capabilities covering the provision of resources in the form of storage, computing and network services. An Infrastructure Layer handles the provision and operation of services on top of the underlying resources in an Infrastructure-as-a-Service (IaaS) approach. The small-core approach enables MSPs to build their own platform and function layers, on top of which they are able to run their applications.

In the "large-core" approach, the Computing Infrastructure operators would deliver further value by offering a function layer containing sets of services; for example, vehicle location finding, artificial intelligence libraries, security functions and so on. The large-core approach enables MSPs to build their own applications using the functions and capabilities provided by the platform. Each MSP can create its own application suite, which can then be deployed on the large core.

The architecture figure also illustrates the AECC Service element. As described previously in this document, the purpose of the AECC Service is to provide capabilities that Mobility Service Providers, running applications on the AECC System, can leverage to perform application lifecycle management functions.

The “core” model is intended to appeal to the broadest set of potential parties to ease adoption. The approach also assists with system design longevity, enabling components within each of the layers to be replaced as new technologies and solutions become available, without requiring a redesign of the system architecture.

References

[1] GSMA, Connected Car Forecast: Global Connected Car Market to Grow Threefold Within Five Years. Feb. 2013.

https://www.gsma.com/IoT/wp-content/uploads/2013/06/cl_ma_forecast_06_13.pdf

[2] Ericsson, Ericsson Mobility Report, June 2019 edition.

https://www.ericsson.com/49d1d9/assets/local/mobility-report/documents/2019/ericsson-mobility-report-june-2019.pdf

[3] Ericsson, Ericsson Mobility Visualizer, June 2019.

https://www.ericsson.com/en/mobility-report/mobility-visualizer?f=14&ft=2&r=1&t=21,22,23&s=14&u=1&y=2018,2024&c=1s

[4] Cisco, Cisco Visual Networking Index: Forecast and Trends, 2017–2022.

https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white-paper-c11-741490.html

[5] Automotive Edge Computing Consortium, White Paper, General Principle and Vision, Version 2.0.0, Dec. 2018.

https://aecc.org/wp-content/uploads/2018/02/AECC_White_Paper.pdf

[6] 3GPP, Technical Specification 23.401: General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access, Version 16.3.0, Jun. 2019.

http://www.3gpp.org/ftp//Specs/archive/23_series/23.401/23401-g30.zip

[7] 3GPP, Technical Specification 23.501: System Architecture for the 5G System, Version 16.1.0, Jun. 2019.

http://www.3gpp.org/ftp//Specs/archive/23_series/23.501/23501-g10.zip

[8] oneM2M, Technical Specification 0001: Functional Architecture, Version 4.1.0, Jun. 2019.

http://member.onem2m.org/Application/documentapp/downloadLatestRevision/default.aspx?docID=30069

[9] 3GPP, Technical Specification 23.402: Architecture enhancements for non-3GPP accesses, Version 15.3.0, Mar. 2018.

http://www.3gpp.org/ftp//Specs/archive/23_series/23.402/23402-f30.zip

[10] 3GPP, Technical Specification 24.312: Access Network Discovery and Selection Function (ANDSF) Management Object (MO). Version 15.0.0, Jun. 2018.

http://www.3gpp.org/ftp//Specs/archive/24_series/24.312/24312-f00.zip

[11] 3GPP, Technical Specification 24.302: Access to the 3GPP Evolved Packet Core (EPC) via non-3GPP Access Networks; Stage 3. Version 16.0.0, Mar. 2019.

http://www.3gpp.org/ftp//Specs/archive/24_series/24.302/24302-g00.zip

[12] IETF, Internet-Draft: Generic Multi-Access (GMA) Convergence Encapsulation Protocols.

https://tools.ietf.org/html/draft-zhu-intarea-gma-03

[13] IETF, Internet-Draft: Multi-Access Management Services.

https://www.ietf.org/id/draft-kanugovi-intarea-mams-framework-04.txt

[14] 3GPP, Technical Specification 36.300: Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2. Version 15.7.0, Sep. 2019.

http://www.3gpp.org/ftp//Specs/archive/36_series/36.300/36300-f70.zip

[15] 3GPP, Technical Specification 33.401: 3GPP System Architecture Evolution (SAE); Security architecture. Version 16.1.0, Dec. 2019.

http://www.3gpp.org/ftp//Specs/archive/33_series/33.401/33401-g10.zip

[16] 3GPP, Technical Specification 36.361: Evolved Universal Terrestrial Radio Access (E-UTRA); LTE-WLAN Radio Level Integration Using Ipsec Tunnel (LWIP) encapsulation; Protocol specification. Version 15.0.0, Jul. 2018.

http://www.3gpp.org/ftp//Specs/archive/36_series/36.361/36361-f00.zip

[17] 3GPP, Technical Report 23.793: Study on access traffic steering, switch and splitting support in the 5G System (5GS) architecture. Version 16.0.0, Dec. 2018.

http://www.3gpp.org/ftp//Specs/archive/23_series/23.793/23793-g00.zip

[18] 3GPP, Technical Specification 37.340: NR; Multi-connectivity; Overall description; Stage-2. Version 15.7.0, Sep. 2019.

http://www.3gpp.org/ftp//Specs/archive/37_series/37.340/37340-f70.zip

[19] IETF, RFC 2132: DHCP Options and BOOTP Vendor Extensions.

https://tools.ietf.org/html/rfc2132

[20] IETF, RFC 5970: DHCPv6 Options for Network Boot.

https://tools.ietf.org/html/rfc5970

[21] 3GPP, Technical Specification 22.011: Service Accessibility. Version 17.0.0, Dec. 2019.

http://www.3gpp.org/ftp//Specs/archive/22_series/22.011/22011-h00.zip

[22] 3GPP, Technical Specification 23.502: Procedures for the 5G System; Stage 2. Version 16.3.0, Dec. 2019.

http://www.3gpp.org/ftp//Specs/archive/23_series/23.502/23502-g30.zip

Acknowledgements

This document is provided by the Automotive Edge Computing Consortium (AECC) Technical Solution Working Group (WG2). The following people have provided valuable contributions to this document:

ORGANIZERS

Leifeng Ruan (WG Chair)

Mikael Klein (WG Vice-Chair)

Lei Zhong (Chief Editor)

CONTRIBUTORS (in alphabetical order)

Lidwina Andarini

Hirochika Asai

Bastian Cellarius

Zongrui Ding

George Ericson

Toru Furusawa

Brian Keller

Seiichi Koizui

Tomoyuki Kurobe

Clara Li

Morgan Lindqvist

Thorsten Lohmar

Takuya Miyasaka

Koya Mori

Joel Obstfeld

Ryokichi Onishi

Ryo Tamura

Xiaopeng Tong

Wenlei Wu

Version History

Date	Version	Description
2019/09	1. 0.0	Initial AECC Reference Architecture, three key issues and corresponding solution recommendations for Edge Data Offloading, MSP Server Selection and Vehicle System Reachability.
2020/07	2.0.0	Added three key issues of Access Network Selection, Provisioning and Configuration Update and Opportunistic Data Transfer. Added initial Distributed Computing Reference Model into AECC Reference Architecture.

About the Automotive Edge Computing Consortium (AECC)

The AECC is a consortium of leaders across industries focused on driving the evolution of edge network architectures and computing infrastructures to support high-volume data services in a smarter, more efficient connected-vehicle future. AECC members are key players in the automotive, high-speed mobile network, edge computing, wireless technology, distributed computing and artificial intelligence markets.

Our mission is to develop an open, technology-agnostic framework to support the transfer of data and communications between the vehicle and local networks near the source of the data, and then to a centralized cloud in a seamless, safe, reliable and optimized manner. Our members collaborate on the development of use cases, technical solutions and reference architectures.

- **Use Cases** – Create use cases and requirements in networking and computing for connected services in cars.
- **Technical Solutions** – Provide inputs to standards and open source communities on best practices for deploying distributed and layered networking and computing solutions for connected cars.
- **Reference Architectures** – Develop architectures for next-generation mobile networks and distributed computing that are suitable for automotive use cases.

We invite you to join the AECC and add your insights and influence. Visit https://aecc.org/ to learn more.