Predictive QoS and V2X Service Adaptation

5GAA Automotive Association
Technical Report

CONTACT INFORMATION:

Executive Manager – Thomas Linget
Email: liaison@5gaa.org

MAILING ADDRESS:

5GAA c/o MCI Munich
Neumarkter Str. 21
81673 München, Germany
www.5gaa.org

VERSION:	Version 2
DATE OF PUBLICATION:	26 January, 2023
DOCUMENT TYPE:	Technical Report
EXTERNAL PUBLICATION:	Yes
DATE OF APPROVAL BY 5GAA BOARD:	23 January, 2023

Contents

Foreword

This Technical Report has been produced by 5GAA.

The contents of the present document are subject to continuing work within the Working Groups (WG) and may change following formal WG approval. Should the WG modify the contents of the present document, it will be re-released by the WG with an identifying change of the consistent numbering that all WG meeting documents and files should follow (according to 5GAA Rules of Procedure):

x-nnzzzz

(1) This numbering system has six logical elements:

(a) x: a single letter corresponding to the working group:

where x =

T (Use cases and Technical Requirements)

A (System Architecture and Solution Development)

P (Evaluation, Testbed and Pilots)

S (Standards and Spectrum)

B (Business Models and Go-To-Market Strategies)

(b) nn: two digits to indicate the year. i.e. ,17,18 19, etc

(c) zzzz: unique number of the document

(2) No provision is made for the use of revision numbers. Documents which are a revision of a previous version should indicate the document number of that previous version

(3) The file name of documents shall be the document number. For example, document S-160357 will be contained in file S-160357.doc

1 Scope

Predictive Quality of Service (QoS) enables the mobile network to provide notifications[1] about the probability of QoS changes to the Vehicle-to-Everything (V2X) applications and other consumers. Based on the predicted values of QoS parameters, V2X applications (e.g. related to safety, efficiency and convenience) are able to weigh the need to adapt their behaviour or configuration in advance of the predicted connectivity conditions, and to react in terms of providing updated Service-Level Requirements (SLR). It is clear that any adaptation that may be performed by the application does not depend only on the information about the QoS prediction: the analytics information that is provided by the network is intended to be used by the V2X application to complement all other available contextual information.

In previous work items [3] [4], 5GAA had worked on the specification of interfaces, the content of the QoS prediction message, the Key Performance Indicators (KPIs) that can be predicted, as well as the functional requirements and required architectural enhancements. In such activities, 5GAA had also defined a basic framework for defining application adaptation.

The Predictive QoS and V2X Service Adaptation (PRESA) Work Item (WI) and this Technical Report (TR) enhance the framework for the analysis of the potential adaptations of V2X applications and V2X services when a QoS change (e.g. data rate degradation, reliability degradation, latency increase) is predicted and such information is provided to the application. The scope of this TR is to investigate the application and system reactions/adaptations, based on predicted changes of QoS. The methodology for the analysis of V2X use cases that use predictive QoS information is introduced in order to investigate various aspects, such as identifying interactions, QoS parameters of interest for prediction (e.g. downlink minimum data rate) for various use cases, related QoS prediction time horizons, thresholds for triggering prediction notification, etc.

Three use cases have been selected (Tele-operated Driving, Infrastructure Assisted Environment Perception, Cooperative Lane Merge) to study application and system reactions, using the introduced methodology with the objective to identify:

a) different types of application and system reactions for selected use cases specified in 5GAA, according to predicted QoS change,

b) steps of different application and system reactions, considering various parameters (e.g. specific use cases, vehicle operation, road environment, etc.),

c) application reaction times and QoS prediction notification periods for different types of application reactions to different predicted SLRs.

Lastly, this TR studies different enhancements to and requirements for interfaces, signalling and architecture of the QoS prediction, including but not limited to edge, cloud and interoperability aspects among different MNOs, OEMs, etc. In this context, since the latest 3GPP work has already introduced further enhancements to Predictive QoS functionality [13], this TR focuses on the areas 5GAA previously identified for improving the 3GPP Solution ([3], [4]), providing a revision of such areas of improvement in light of the new use case analysis and new projected functionality of the 3GPP System.

2 References

[1] 5GAA White paper, 'C-V2X Use Cases Volume II: Examples and Service Level Requirements', October 2020. Available at: https://5gaa.org/wp-content/uploads/2020/10/5GAA_White-Paper_C-V2X-Use-Cases-Volume-II.pdf

[2] 5GAA White paper, 'Making 5G Proactive and Predictive for the Automotive Industry', December 2019. Available at: https://5gaa.org/wp-content/uploads/2020/01/5GAA_White-Paper_Proactive-and-Predictive_v04_8-Jan.-2020-003.pdf

[3] 3GPP Technical Specification 23.287, 'Architecture enhancements for 5G System (5GS) to support Vehicle-to-Everything (V2X) services (Release 17)', June 2022. Available at: https://3gpp.org/DynaReport/23287.htm

[1] Such notifications are also defined as QoS Potential Notifications (QPNs), In-advance QoS Notifications (IQN), potential QoS notifications, or notification on potential QoS change.

[4] 3GPP Technical Specification 23.288, 'Architecture enhancements for 5G System (5GS) to support network data analytics services (Release 17)', June 2022. Available at: https://3gpp.org/DynaReport/23288.htm

[5] 3GPP Technical Specification 23.501, 'System architecture for the 5G System (5GS) (Release 17)', June 2022. Available at: https://3gpp.org/DynaReport/23501.htm

[6] 3GPP Technical Specification 23.286, 'Application layer support for Vehicle-to-Everything (V2X) services; Functional architecture and information flows', June 2022. Available at: https://3gpp.org/DynaReport/23286.htm

ETSI Group Report MEC 022 V2.1.1, 'Study on MEC Support for V2X Use Cases', September 2018. Available at: https://www.etsi.org/deliver/etsi_gr/mec/001_099/022/02.01.01_60/gr_mec022v020101p.pdf

[7] 5GAA Technical Report A-190003, 'Architectural Enhancements for Providing QoS Predictability in C-V2X', 2019

[8] 5GAA Technical Report A-200094, 'V2X Application Layer Reference Architecture', June 2020. Available at: https://5gaa.org/wp-content/uploads/2020/06/5GAA_A-200094_V2XSRA-Application-Layer-Reference-Architecture-final.pdf

[9] 5GAA Technical Report A-200146 'MEC4AUTO: Use Cases and initial test specifications review'. Available at: https://5gaa.org/wp-content/uploads/2021/07/5GAA_MEC4AUTO.pdf

[10] 5GAA Technical Report A-200150 "MEC4AUTO Task 2: MEC-for-Automotive-in-Multi-Operator-Scenarios'. Available at: https://5gaa.org/wp-content/uploads/2021/03/5GAA_A-200150_MEC4AUTO_Task2_TR_MEC-for-Automotive-in-Multi-Operator-Scenarios.pdf

[11] ETSI Group Specification MEC 030 'MEC V2X Information Service API' v2.1.1. Available at: https://www.etsi.org/deliver/etsi_gs/MEC/001_099/030/02.01.01_60/gs_MEC030v020101p.pdf

[12] 5GAA Technical Report A-200055, '5GS Enhancements for Providing Predictive QoS' in C-V2X, 2020. Available at: https://5gaa.org/wp-content/uploads/2020/05/5GAA_A-200055_eNESQO_TR_final.pdf

[13] 3GPP Technical Report 23.700-81, 'Study of Enablers for Network Automation for 5G); Phase 3 (Release 18). Available at: https://www.3gpp.org/DynaReport/23700-81.htm

[14] ETSI Revised Group Specification MEC-0030v311V2XAPI Work Item 'Multi-access Edge Computing (MEC); V2X Information Service API'. Available at: https://portal.etsi.org/webapp/WorkProgram/Report_WorkItem.asp?WKI_ID=62539

[15] 3GPP Technical Report 23.700-80, 'Study on 5G System Support for AI/ML-based Services' (Release 18). Available at: https://www.3gpp.org/DynaReport/23700-80.htm

[16] 3GPP SP-221320, 'Reply LS on QoS Sustainability analytics and V2X service adaptations', Release 18, Release 19.

3 Definitions and abbreviations

3.1 Definitions

For the purposes of the present document, the following definitions apply:

Predictive QoS: A mechanism that enables the mobile network to provide notifications about predicted QoS changes to interested consumers in order to adjust the application behaviour in advance.

Potential QoS change notification or **In-advance QoS change Notification, IQN** or **QoS Prediction Notification, QPN**): As defined in [7], the message containing the QoS Prediction which is delivered to consumers. In the 3GPP system, this is defined as the analytics output in cl. 6.9 of [4] for an analytics target period in the future.

Tdel: The time when the potential QoS change notification is delivered to the intended consumer, such as the V2X Application [7].

Tarc: The time when the V2X Application has completed the reaction to receiving the potential QoS change notification [7].

3.3 Abbreviations

For the purposes of the present document, the following abbreviations apply:

5GS	5G System
5QI	5G Quality of Service Indicator
AF	Application Function
API	Application Programming Interface
ARCI	Application Reaction Completion Interval
AS	Application Server
C-V2X	Cellular Vehicle-to-Everything
DL	Down Link
E2E	End-to-End
eNESQO	Enhanced E2E Network Slicing and Predictive QoS
ETSI	European Telecommunications Standards Institute
GBR	Guaranteed Bit Rate
HD	High Definition
HV	Host Vehicle
IP	Internet Protocol
IQN	In-advance Quality of Service Notification
ISG	Industry Specification Group
IVE	In-Vehicle Entertainment
KPI	Key Performance Indicator
MEC	Multi-Access Edge Computing
MNO	Mobile Network Operator
ms	Millisecond
N3 interface	5GS interface between the RAN (gNB) and the (initial) UPF
N6 interface	5GS interface between the User Plane Function (UPF) and any other external (or internal) networks or service platforms, such as the Internet, the public cloud or private clouds
NAS	Non-Access Stratum
NEF	Network Exposure Function
NESQO	Predictive QoS and End-to-end Network Slicing for Automotive Use Cases
NF	Network Function
NG-RAN	Next-Generation Radio Access Network
NHTSA	National Highway Traffic Safety Administration
NWDAF	Network Data Analytics Function
OAM	Operations, Administration and Maintenance
OEM	Original Equipment Manufacturer
PDB	Packet Delay Budget
PER	Packet Error Rate
PF	Prediction Function
PQoS	Predictive QoS
PRESA	Predictive QoS and V2X Service Adaptation
QoE	Quality of Experience
QoS	Quality of Service
RRC	Radio Resource Control
RV	Remote Vehicle
SL	Sidelink
SLA	Service Level Agreement
SLR	Service Level Requirement
S-NSSAI	Single – Network Slice Selection Assistance Information
SP	Service Provider
TAI	Tracking Area Identifiers
ToD	Tele-operated Driving
UC	Use Case
UE	User Equipment
UL	Up Link
UPF	User Plane Function
Uu	Interface for cellular communication between device and base station
V2N2I	Vehicle-to-Network-to-Infrastructure
V2N2V	Vehicle-to-Network-to-Vehicles

V2X	Vehicle-to-Everything
VAE	V2X Application Enabler
WG	Work Group
WI	Work Item

4 Overview

4.1 Predictive QoS

In 5GAA Working Group 1 (use cases and technical requirements) various V2X use cases for different categories (e.g. safety, automated driving, traffic efficiency) and road configurations have been presented. According to [1], different use cases have different performance requirements for their smooth and reliable operation, while in most cases the safety and automated driving use cases have the most demanding QoS requirements on the wireless communication system.

The experienced QoS may be affected by various factors, such as interference, mobility, coverage, network conditions (e.g. network load, configuration), terminal characteristics (e.g. number of antennas), handover and roaming transitions. The avoidance of a sudden session interruption due to QoS degradation is a key requirement especially for critical V2X services (e.g. safety, automated driving, cooperative manoeuvres). A feature that many V2X services have is an inner capability to adapt to different QoS conditions, provided that the QoS variation is notified in advance to the application, with sufficient time and confidence to give the application the ability to implement the adaptation. Such capability means that some V2X applications may operate with different configurations, each one corresponding to different QoS requirements (e.g. latency, packet error rate, data rate, and jitter). As an example, due to this feature, when a QoS degradation is predicted, the applications can continue to operate by selecting another configuration that corresponds to an alternative (often lower) QoS requirement and make the transition before the degradation occurs. The above-mentioned characteristics of the V2X applications and the specific automotive requirements create the need to predict the change of the QoS level of one or more QoS parameters, as well as to provide early notifications to the vehicles about the expected decrease or increase of the QoS. This notification can help the V2X application quickly adjust ahead of the QoS change, instead of having to adapt after the QoS change has happened and already affected the V2X service.

As it is described in [2], **predictive QoS** *is a mechanism that enables the mobile network to provide notifications about predicted QoS changes to interested consumers in order to adjust the application behaviour in advance*. Such prior notifications, whenever predictions are made with sufficient confidence, should be delivered with some time in advance (a notice period) before the new predicted QoS is experienced. The notice period depends on the specific application and use case, and should be long enough to give the application sufficient time to adapt to the new predicted QoS.

4.2 Predictive QoS for Automotive

The spatio-temporal dynamics of wireless networks and high mobility of vehicles may lead to sudden QoS changes. These sudden changes and especially a degradation in QoS may affect the functioning of many V2X applications (e.g. the sudden release of a data bearer may affect the V2X service performance, safety and/or efficiency). Hence, affected V2X applications can be sent a notification by the network informing them of an expected or estimated change of QoS before the actual change occurs, allowing them to modify their configuration (e.g. move from automated assisted driving to manual mode, increase inter-vehicle gap, decrease speed, etc.) in an agile and timely way. For each application-level configuration, a different QoS level (e.g. data rate, latency, packet error rate, jitter, etc.) may be required. With this prediction, sudden QoS changes at the application level can be avoided or mitigated by informing the applications (whether in the vehicle, in an edge hosting environment or in the cloud) about the connectivity parameters and imminent changes to maintain safety and efficiency of V2X services.

In 5GAA there are several use cases identified that could benefit from QoS prediction notifications. A non-exhaustive list includes: Tele-operated Driving, Tele-operated Support, Automated Intersection Crossing, Cooperative Lane Merge, Coordinated/Cooperative Driving Manoeuvre, Intersection Movement Assist, High-Density Platooning, Emergency Brake Warning, Lane-Change Warning, UC Awareness Confirmation, Cross-Traffic Left-Turn Assist, Cooperative Lateral Parking, Software Update, and High-Definition (HD) mapping. Because of time limits, not all these use cases have been studied in this TR, however a subset has been selected and described in Section 5.2.

4.3 Previous Work

5GAA has conducted extensive analysis and design work on predictive QoS, in the context of the **Predictive QoS and End-to-end Network Slicing for Automotive Use Cases** (**NESQO)** Work Item (WI) and **Enhanced E2E Network Slicing and Predictive QoS (eNESQO)** WI. The NESQO work has proposed a number of enhancements in the 5G

System (5GS) aiming at introducing Predictive QoS as a new proactive behaviour and enabled by in-advance notifications with QoS predictions. V2X applications can better address the consequences of potential QoS degradations thanks to the prediction notification. This mechanism has obvious advantages compared to the reactions that could be performed after QoS degradation has already happened and been detected by the application. IQN can also help network functions in the 5GS; similar considerations apply to how the network reacts to the potential QoS degradation as well. For predictive QoS, NESQO defined a set of requirements, high-level procedures, message content as well as a proposed KPIs which could be relevant for the QoS prediction. NESQO also identified two approaches to such predictions – network-level and application level – and provided a proposal for which KPIs can be predicted by each approach.

The follow-up work under eNESQO has further developed these results in two main directions:

1) By further detailing aspects and mechanisms for making QoS predictions: those aspects were analysed according to network-level and application-level predictions, while an analysis of the characteristics of the two approaches (including advantages and disadvantages) has been provided,

2) By further detailing how automotive applications may take advantage of the QoS prediction: examples of prediction-centric use case descriptions, with detailed logical flows of actions that may be performed by the application when an IQN is received, have been described. Moreover, a high-level framework to be used to define application and network reactions to QoS predictions has been introduced.

This study capitalises on the work that has been conducted in the context of the NESQO WI and eNESQO WI and further investigates the impact of QoS prediction notifications on application and system reactions, as well further developing and enhancing the interfaces, signalling and architecture of the predictive QoS System.

4.4 Standardisation Status

In 3GPP Service and System Aspects Working Group 2 (SA2), an architectural solution has been introduced, in Release 16, about the notifications on potential QoS changes [3]. The goal is to enable 5G communication systems to provide notifications regarding potential QoS changes upon request from a V2X Application Server (AS), which also provides the network location information in the form of a 'path of interest' or geographical area time window indicating the period to which the notification applies and threshold(s) indicating level(s) which, if crossed, indicate that the potential QoS change (improvement or worsening) could occur.

The utilisation of this procedure for V2X applications is discussed in TS 23.287 [3] and TS 23.288 [4]. It is considered that a V2X AS may request notifications on QoS Sustainability Analytics for an indicated geographic area and time interval in order to adjust the application behaviour in preparation for a potential QoS change. The V2X Application Server acts as an Application Function (AF), which communicates with the Network Exposure Function (NEF) which is a Network Function (NF) consumer of the 'QoS Sustainability' analytics procedure, defined in TS 23.288. The V2X AS can either subscribe to notifications from the NEF (i.e. a Subscribe-Notify model) or request a single notification from the NEF (i.e. a Request-Response model). The current procedure to provide early QoS notifications (or 'QoS Sustainability' analytics) is outlined below and illustrated in Figure 1:

1. The V2X AS collects application layer information (e.g. V2X service, path, path start time and QoS requirements and thresholds).
2. The V2X AS requests or subscribes to analytics information on 'QoS Sustainability' provided by the Network Data Analytics Function (NWDAF) via the NEF. The NWDAF is responsible for on-demand provision of analytics to other network entities (i.e. consumers). The parameters included in the request or subscription message are the following: Analytics ID = 'QoS Sustainability', Analytics Filter Information (QoS requirements, Location information, Observation Period, Threshold(s), Single Network Slice Selection Assistance Information (S-NSSAI). The latter (S-NSSAI) is an optional parameter.
3. The NWDAF collects statistics provided by the Operations, Administration and Maintenance (OAM) entity (explained in 3GPP TS 23.288 [4]). OAM is responsible for management plane activities, including network performance monitoring. The NWDAF verifies whether the triggering conditions are met and derives the requested analytics or predictions about any expected QoS change. The NWDAF can detect the need for notification by comparing the requested analytics of the target 5QI [5] collected from the OAM against the threshold(s) provided by the V2X AS in any cell over the requested Observation Period.
4. The NWDAF provides a response or notification on 'QoS Sustainability' to the V2X AS, again via the NEF. The structure of the response is as follows:
    - Applicable Area: A list of Tracking Area Identifiers (TAIs) or Cell IDs within the Location information that the analytics applies to.
    - Applicable Time Period: The time period within the Observation Period that the analytics applies to.
    - Crossed Threshold(s): The threshold(s) that are met or exceeded by the statistics value or the expected value of the QoS KPI.

- Confidence: Confidence of the analytics.

5. Based on a received notification by the network, the V2X application adjustment may take place.

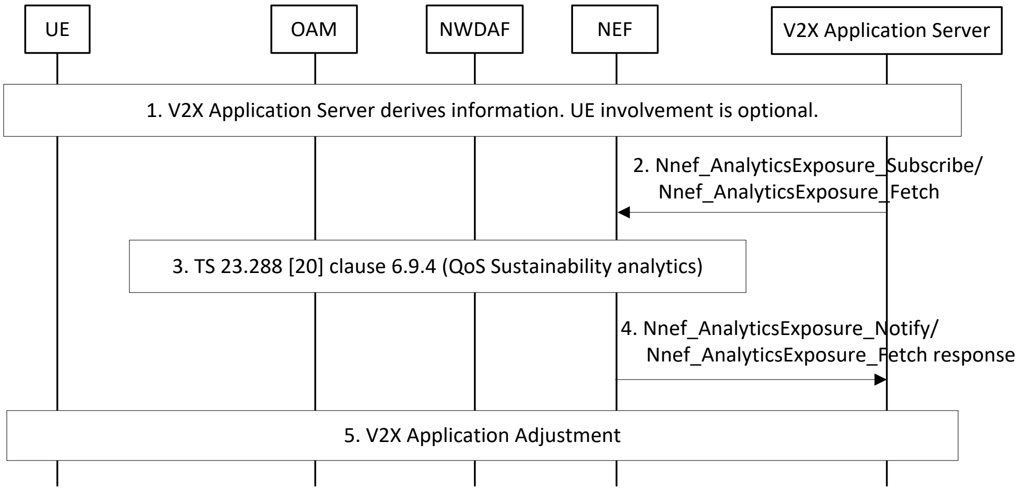


Figure 1: Notification on QoS Sustainability Analytics to the V2X Application Server [3]

Some enhancements have been made in Release 17 on the input features of the QoS Sustainability analytics, while in the context of Release 18 there is – at the time of writing – ongoing discussion about improving the granularity of QoS Sustainability analytics, which in Release 17 is limited to cell-level. That means the 3GPP System can provide statistically valid QoS predictions under certain conditions and based on statistics collected on specific KPIs and averaged across all User Equipment (UE) in the cell. Such granularity does not enable the 3GPP System to consider key characteristics of the UE context which may affect the QoS experienced, and therefore limits the accuracy of the QoS prediction. It has been concluded that Release 18 QoS will support more granular and detailed QoS Sustainability analytics, e.g. below cell level. In addition, 3GPP SA6 has specified the procedure where the V2X Application Enabler (VAE) server sends notification of the network monitoring information (including QoS Sustainability) to the V2X UE [6].

Predictive QoS support has also been identified as one of the key solutions for mobility and Quality of Experience (QoE) support issues, described by the European Telecommunications Standards Institute (ETSI), in the context of the Multi-Access Edge Computing (MEC) framework [11].

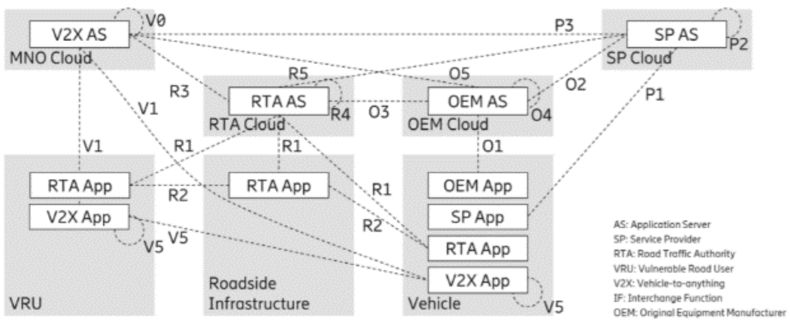
5 Application and System Reactions

5.1 Use Case Analysis Methodology

This section presents the methodology in the use case analysis for Predictive QoS-related Service Level Requirements. It helps to identify the parameters and the aspects of each use case that should be taken into consideration as well as to determine the Application Reaction Completion Interval (ARCI) [7] of each use case, which is useful for estimating the so-called QoS Prediction Time Horizon.

Each V2X use case has a list of potential different reactions that could be triggered when the application receives a prediction about one or more SLRs that cannot be satisfied. For each use case, it is proposed to use the following table to collect the list of potential application reactions according to one or more identified SLR changes.

Table 1: Potential Application Reactions per Use Case

Use Case Title	Indicate the title of the use case, according to the latest WG1 list
User Story	Indicate the selected user story of the use case, according to the latest WG1 list
Type of Service	Indicate the type of the C-V2X service selection, i.e. one of the following: ‘Short duration event-driven’, ‘Medium/Long duration session-based’, or ‘Periodic’
Predicted QoS/Service Level Requirements of Interest per communication link	For each considered communication link used in the corresponding use case, please present the following: 1) the QoS parameter(s) to be predicted: (Service Level) Latency, (Service Level) Reliability, Data Rate, Positioning Accuracy… 2) the type of communication link (e.g. PC5, Uu) 3) the transmission mode for the communication link e.g. unicast, multicast, broadcast 4) Type/description of data traffic (e.g. video streaming or audio streaming)/data traffic management (e.g. see-through – discover the car/ find the cameras) The goal is to indicate the association per QoS parameter the communication link, as well as per mode/interface
Considered Interface(s)	Specify the considered interfaces according to the Deployment view of V2X System Application Layer Reference Architecture with MNO stakeholder for V2X AS (no Interchange Function) [8] 
QoS Prediction Recipient and/or Orchestrator for Adaptation?	Specify the entity or entities that receive the QoS prediction request and/or the entity responsible for the orchestration of the application adaptation, based on QoS prediction notification. Use the entities mentioned in the Deployment view of V2X System Application Layer Reference Architecture
Involved Entities	e.g. Vehicle, V2X Application Server, SP Application Server, OEM Application Server, RTA Application Server

Reaction Scenario #	Potential Application Reactions	Entity that Triggers the Application Reaction	SLR Change (e.g. Data Rate, Coverage, Latency, Reliability, Position Accuracy)
1.	e.g. Release of the service		e.g. Coverage loss…
2.	e.g. Switch to local sensor		e.g. Decrease of Data Rate…
3.	…		
	…		
	…		

NOTE: for the same change of SLR, reaction can be different per different vehicles.

We have to identify different scenarios of QoS prediction outputs in terms of accuracy, notification period, etc.:

- Areas with capability to predict the QoS accurately
- Areas that may not be able to estimate accurately
- Areas that a minimum communication capability can be estimated

The application reactions may include several steps (sub-processes) that constitute the ARCI. Hence, for the calculation of ARCI, it is needed to identify the duration of each step/sub-process of an application reaction. The following tables are used to present the Time Breakdown Analysis per Application Reaction. It is useful also to include the steps that are involved for QoS prediction request and delivery. In the scenarios analysed below, it is assumed that the vehicle and/or application service has subscribed to the QoS prediction service in order to receive QoS prediction notifications.

Table 2: Time Breakdown Analysis per Application Reaction for Reaction Scenario #1

Application Reaction: Scenario #1		
Road Environment Type: Indicate the type of road that the use case and application reaction is applied		
Vehicle Type: Indicate the type of vehicle that the use case and application reaction is applied		
Human Intervention: yes/no/…		
List Steps	Latency (ms)[2]	Explanation
1. Description of step 1	e.g. 10	Justification of estimated latency
2. Description of step 2	e.g. 100	
3. …		
4.		
5.		
Total Sum:		

Table 3: Time Breakdown Analysis per Application Reaction for Reaction Scenario #2

Application Reaction: Scenario #2		
Road Environment Type: Indicate the type of road that the use case and application reaction is applied		
Vehicle Type: Indicate the type of vehicle that the use case and application reaction is applied		
Human Intervention: yes/no/…		
List Steps	Latency (ms)	**Explanation**
1.		
2.		
3.		
4.		
5.		
Total Sum:		

[2] We consider a worst-case latency estimation for each identified step.

For ARCI calculation, one important factor is the identification whether Machine or Human is involved in the Reaction after a QoS prediction notification and whether that human reaction can be simple and fast ('hit the brakes!') or complex and slow ('resume driving in a complex situation'). The following reactions types are considered:

1) If QoS is degraded but still feasible for autonomous operation, then inter-vehicle distance can be recalculated and readjusted before the QoS actually changes. **The communication, processing, decision, and reaction/adjustment times would fit between Tdel and Tarc and occur at machine speeds.**

2) On the other hand, if QoS cannot support continued autonomous operation, then two things must occur in parallel:

   - **Driver must be re-engaged** to resume control.

     a) Attentive and trained 'safety driver,' this could be **less than 1 second to start a simple action (just starting to hit the brakes).**

     b) Engaged driver not expecting immediate danger, this is **typically closer to 1.7-2 seconds to start the action** (e.g. National Highway Traffic Safety Administration, NHTSA, test observations when using back-up camera, time from object detection to start of braking).

     c) For a driver who has experienced Primary Task Reversal (e.g. if the driver is watching a video on their smartphone), some information indicates the time until re-engagement and **complete driving task resumption could be 20-30 seconds**, while other information indicates several seconds (+/-4 seconds) might be suitable in realistic but unexpected scenarios.

   - The automated driving or ADAS system can help to increase reaction time available to the human driver by starting to take action at machine reaction times.

Table 4 provides an overview of the estimated QoS prediction time horizons, considering the different types of application adaptations of the ToD service. It should be noted that the QoS prediction time horizon is the sum of the ARCI and a guard prediction interval that is introduced. The guard prediction interval can be specified considering different factors that are important to ensure that the QoS prediction time horizon is long enough for the requirements of each application. Such factors can be, for example, the interval needed to determine and transfer the prediction, as well as other implementation-specific factors.

Table 4: QoS Prediction Time Horizon of Different Application Reactions

Use Case Title	**Indicate the title of the use case, according to the latest WG1 list**		
User Story	Indicate the selected user story of the Use Case, according to the latest WG1 list		
Application Reaction	**Application Reaction Completion Interval (ARCI)** **(in ms)**	**Guard Prediction Interval (in ms)**	**QoS Prediction Time Horizon[3] (in ms)**
Indicate the Title of the Identified Application Reaction	It should be clarified that it is needed to indicate the minimum (acceptable) QoS prediction time horizon, considering safety requirements, e.g. 10 seconds		Indicate how fast the QoS prediction notification should be sent/received

[3] The QoS prediction time horizon is equal to the QPN Notice Period and Transfer Interval that have been specified in NESQO TR (5GAA TR A-190003).

	(It is equal to the 'Analytics Target Period' term used in 3GPP)		
….			

During the use case selection phase and later in the use case analysis phase assumptions are made on what machine/human reactions (and latency) would need to be considered. As an example, one potential human reaction could be to choose and select an alternative route (e.g. in the case of Tele-operated Driving). The time for the completion of such actions is one of the inputs for such analysis and needs to be determined based on some assumptions. Other assumptions (left as an exercise for the reader), may also be valid under different conditions and therefore alternative outcomes, based on those assumptions, are possible. For the sake of efficiency in this study, a limited number of use cases have been selected based on given assumptions. However, the methodology developed in PRESA and described in this TR could be used in the future to perform further analysis of the same use cases under different assumptions or completely new ones with the purpose of determining the requirements for predictive QoS for potentially any use case.

In Table 5, Alternative QoS levels could be considered for a specific use case, if needed. In addition, it should be noted that different implementations of a specific use case may consider different threshold values. Hence, no need to 'stick' to specific values or consider them as something static for all implementations. In those cases something more generic could be used to describe the QoS thresholds, such as a percentage of a QoS value.

Table 5: QoS Change Thresholds of QoS Predictions

Use Case Title	**…**		
User Story	…		
Predicted QoS/Service Level Requirements	**Considered Interface(s)**	**QoS/SLR value (5GAA TRs)**	**QoS Change Thresholds for Triggering Prediction Notification[4]**
e.g. UL Throughput	e.g. Uu	e.g. ToD UL Data rate 20Mbps	e.g. UL Throughput lower than 30 Mbps

5.2 Use Case Analysis

In the PRESA WI, the following use cases have been selected to analyse predictive QoS and apply the methodology that has been introduced in Section 5.1.

- Tele-operated Driving
- Infrastructure Assisted Environment Perception
- Cooperative Lane Merge

The selected use cases help us to derive conclusions that could be applicable to services with similar features. For instance, the ToD finding is a very good example of a medium/long duration session-based service that is realised through the Uu interface. The Infrastructure Assisted Environment Perception use case is a periodic service where both Uu and PC5 could

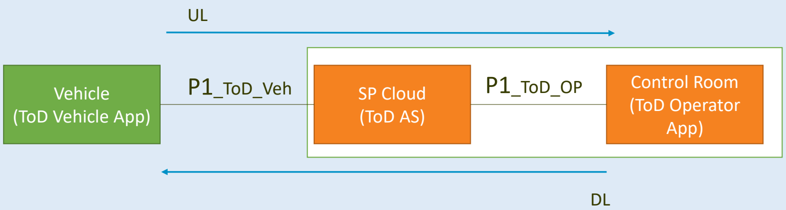
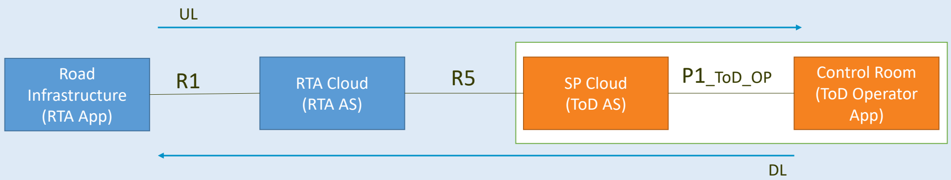
[4] More information is available in sections 5.1.1, 5.5.2 of eNESQO TR: https://5gaa.org/wp-content/uploads/2020/05/5GAA_A-200055_eNESQO_TR_final.pdf

be used. While the Cooperative Lane Merge use case is a short duration event-driven service where vehicles are interacting via the PC5 or alternatively the Uu interface (i.e. V2N2V).

5.2.1 Tele-operated Driving

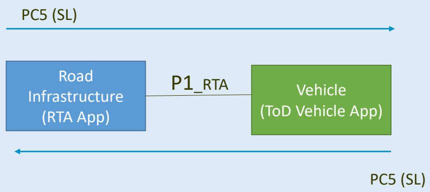
This section presents the analysis of QoS prediction parameters using the Tele-operated Driving use case.

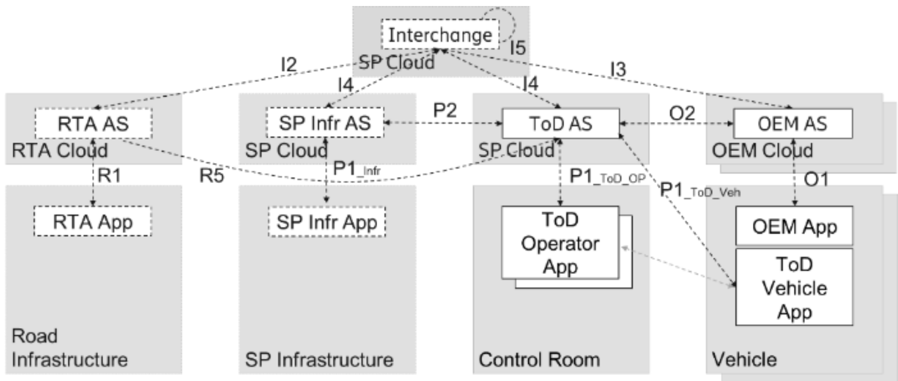
Table 6: Potential Application Reactions of ToD Use Case

Use Case Title	**Tele-operated Driving**
User Story	Based on the perceived environment, the Remote Driver provides to the remotely driven vehicle appropriate manoeuvre instructions to navigate to the destination efficiently and safely (**Direct Control ToD**)
Type of Service	Medium/Long duration ‘session-based’ service
Predicted QoS (Service Level Requirements) of Interest per Communication Link	F1. Vehicle – Control Room: - F1.1: Uu Interface, UL (from the Vehicle to the Control Room): Data Rate , latency, jitter and reliability (unicast)[5] o F1.1.1: Video streaming o F1.1.2: Sensor information o …. - F1.2: Uu Interface, DL (from the Control Room to the Vehicle): Reliability and latency (unicast) o F1.2.1: Manoeuvre commands  F2. Roadside Infrastructure – Control Room (optional): - F2.1: Uu Interface, UL (from Service Provider, SP (Roadside) Infrastructure to Control Room): Data Rate, latency, jitter and reliability (unicast) o F2.1.1 Sensor information o F2.1.2 Video Streaming …  F3. Roadside Infrastructure – vehicle (optional):[6]

[5] The current working assumption is that different QoS flows will be used for the different type of information being exchanged over an interface, e.g. video streaming, sensor information, etc. The same could be applied even between the UL and DL traffic, since they have different QoS requirements.

[6] A notification of predicted QoS at the specific interface could be provided at the vehicle and/or the remote driver. But, does not necessarily mean that always an adaptation will be triggered due to the predicted QoS of this interfaces, since this may not be a mandatory interface for the direct control mode of ToD use case.

	- F3.1: PC5 Interface, (from SP (Roadside) to vehicle): Data Rate, latency, jitter and reliability (unicast)
	  ○ F3.1.1: Sensor information
	  ○ …
	

Considered Interface(s)	1. P1_ToD_veh 2 R1 3. P1_RTA 
QoS Prediction Recipient and/or Orchestrator for Adaptation?	QoS Prediction Recipient: • 'Vehicle – Control Room' Interface (F1): Vehicle or Control Room (ToD Operator App) • 'Road Roadside Infrastructure – Control Room' Interface[7] (F2): control room (ToD AS) • 'Road Roadside Infrastructure – Vehicle' Interface (F3): Vehicle Orchestrator for application adaptation: Vehicle, Control Room (Remote Driver) The type of adaptation that will be selected depends on the actual prediction quality, the road environment, and the thresholds that the application developer/owner sets or even on some regulatory decisions.
Involved Entities	Vehicle, SP Cloud (ToD AS), Control Room, Road Infrastructure, RTA Cloud

Reaction Scenario #	Potential Application Reactions	Entity that Triggers the Application Reaction	SLR Change (e.g. Data Rate, Coverage, Latency, Reliability, Position Accuracy)
1.	The Remote Driver safely stops the vehicle and releases the ToD service	Vehicle or Control Room (Remote Driver)	Vehicle – Control Room: • Coverage Loss and/or

[7] The Remote Operator can be informed about the QoS prediction of this interface and trigger (if needed) an application adaptation that at the end the vehicle can accept (/validate) or not.

			- Reduction of UL Data Rate (F1.1) below defined thresholds and/or - Increase of DL Latency (F1.2) above defined thresholds and/or - Reduction of DL Reliability (F1.2) below defined thresholds
2.	The Remote Driver hands over the control to the vehicle (using local sensors) and releases the ToD service	Vehicle or Control Room (Remote Driver)	Vehicle – Control Room: - Coverage Loss and/or - Reduction of UL Data Rate (F1.1) below defined thresholds and/or - Increase of DL Latency (F1.2) above defined thresholds and/or - Reduction of DL Reliability (F1.2) below defined thresholds
3.	The Remote Driver asks the vehicle to park itself and releases the ToD service	Vehicle, Control Room (Remote Driver)	Vehicle – Control Room: - Coverage Loss and/or - Reduction of UL Data Rate (F1.1) below defined thresholds and/or - Increase of DL Latency (F1.2) above defined thresholds and/or - Reduction of DL Reliability (F1.2) below defined thresholds
4.	The Remote Driver hands over the control to the driver in car and releases the ToD service	Vehicle, Control Room (Remote Driver)	Vehicle – Control Room: - Coverage Loss and/or - Reduction of UL Data Rate (F1.1) below defined thresholds and/or - Increase of DL Latency (F1.2) above defined thresholds and/or - Reduction of DL Reliability (F1.2) below defined thresholds
5.	Detour	Vehicle, Control Room (Remote Driver)	Vehicle – Control Room: - Coverage Loss and/or - Reduction of UL Data Rate (F1.1) below defined thresholds and/or - Increase of DL Latency (F1.2) and/or - Reduction of DL Reliability (F1.2) below defined thresholds
6.	Change sensor set and sensor properties and/or video configuration	Vehicle, Control Room (Remote Driver)	Vehicle – Control Room: Reduction (/Increase) of UL Data Rate (F1.1) below

			(/above) defined thresholds and/or • Increase (Reduction) of UL Latency (F1.1), jitter (F1.1) above (/below) defined thresholds and/or • Reduction (/Increase) of DL Reliability (F1.2) below (/above) defined thresholds
7.	Change driving properties e.g. speed	Vehicle, Control Room (Remote Driver)	Vehicle – Control Room: • Reduction (/Increase) of UL Data Rate (F1.1) below (/above) defined thresholds and/or • Increase (Reduction) of DL Latency (F1.2) above (/below) defined thresholds and/or • Reduction (/Increase) DL Reliability (F1.2) below (/above) defined thresholds

The following tables present the Time Breakdown Analysis for each potential application reaction of the ToD use case.

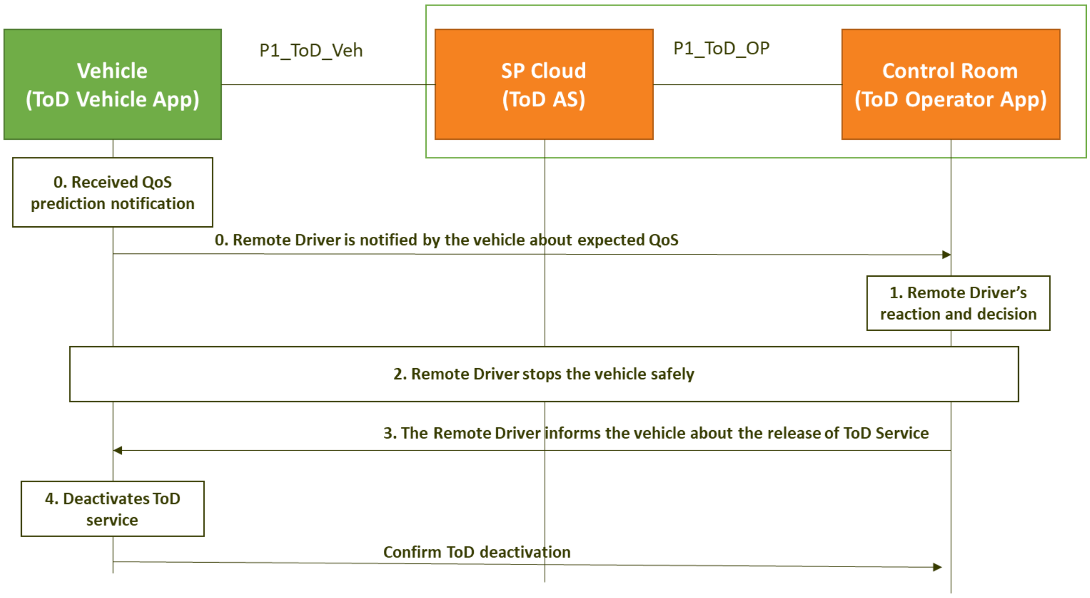


Figure 2: Example Flow Chart for ToD Reaction, Scenario #1: The Remote Driver stops the vehicle safely and releases the ToD service

Table 7: Time Breakdown Analysis for ToD Reaction, Scenario #1: The Remote Driver stops the vehicle safely and releases the ToD service

Application Reaction: Scenario #1 – The Remote Driver stops the vehicle safely and releases the ToD service
Road Environment Type: Highway, Rural, Urban

Vehicle Type: Passenger vehicle		
Human Intervention: Yes		
List Steps	**Latency (ms)**	**Explanation**
1. The Remote Driver is notified about expected QoS degradation or Coverage Loss and decides his/her reaction **It should be noted that Step 0 could be omitted in the event that the notification is sent directly by the network to the Control Room (ToD Operator App)**	2,000	Attentive and trained 'safety driver,' this could be less than 2 second to start a simple action (just starting to hit the brakes)
2. The Remote Driver stops safely the vehicle (Braking Distance)	7,000	Depends on the speed of the vehicle; in the ToD scenario, an assumed maximum speed of 50km/h and the vehicle can stop at the right side of the road, without searching for appropriate parking spot Other factors e.g. lane that the vehicle is located, weather conditions, surface, vehicle type may affect this distance. The worst-case deceleration is $-4m/s^2$ according to the ISO 23793-Intelligent transport systems – Minimal Risk Manoeuvre (MRM) for automated driving – Part 1: Framework, straight-stop and in-lane stop.
3. The Remote Driver informs the vehicle about the release of ToD Service	500	Time that Remote Driver needs to de-activate ToD and communication time from Remote Driver to the vehicle that is remotely driven
4. Remote Vehicle deactivates ToD application and informs the Remote Driver	500	Time needed by the vehicle to de-activate any ToD process and switch back to normal mode
Total Sum:	10,000	

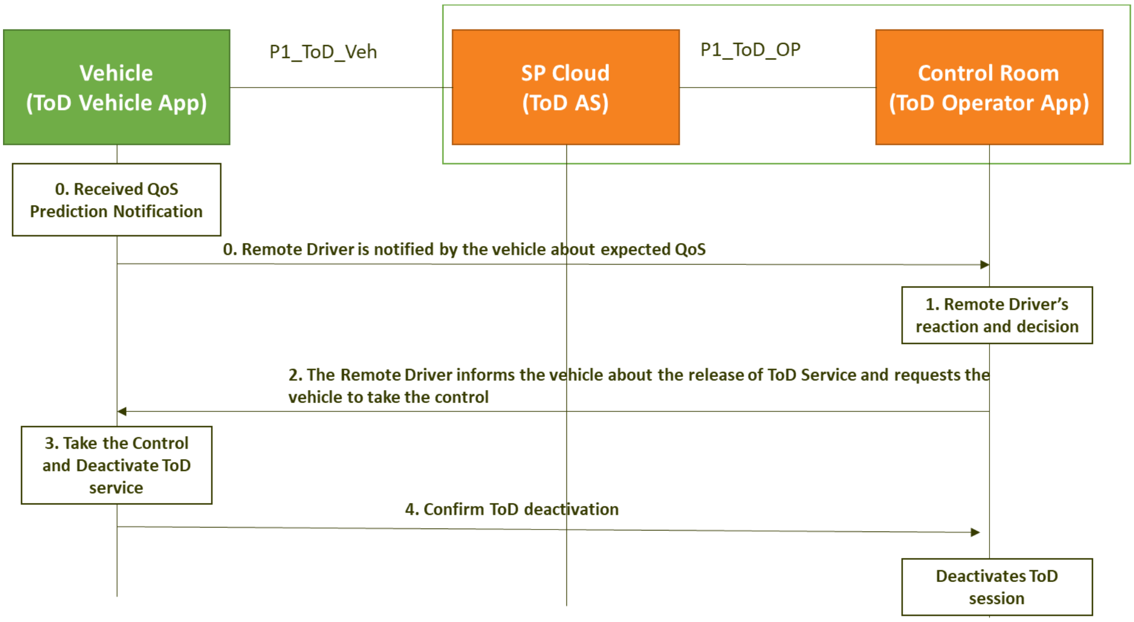


Figure 3: Example Flow Chart for ToD Reaction, Scenario #2: The Remote Driver hands over the control to the vehicle (using local sensors) and releases the ToD service

Table 8: Time Breakdown Analysis for ToD Reaction, Scenario #2: The Remote Driver hands over the control to the vehicle (using local sensors) and releases the ToD service

Application Reaction: Scenario #2, The Remote Driver hands over the control to the vehicle (using local sensors) and release the ToD service		
Road Environment Type: Highway, Rural, Urban		
Vehicle Type: Passenger vehicle		
Human Intervention: Yes		
List Steps	**Latency (ms)**	**Explanation**
1. The Remote Driver is notified about expected QoS degradation or Coverage Loss and decides his/her reaction **It should be noted that Step 0 could be omitted in the event that the notification is sent directly by the network to the Control Room (ToD Operator App)**	2,000	Attentive and trained 'safety driver,' this could be less than 2 second to start a simple action (just starting to hit the brakes)
2. The Remote Driver informs the vehicle about the release of ToD Service and requests the vehicle to take the control	500	Time that Remote Driver needs to de-activate ToD and communication time from Remote Driver to the vehicle that is remotely driven
3. Remote Vehicle takes control of the vehicle, using local sensors, deactivates ToD application, and informs the Remote Driver	500	Time needed by the vehicle to de-activate any ToD process and switch to a mode of autonomous driving includes time needed to build required environmental awareness to drive autonomously
Total Sum:	3,000	

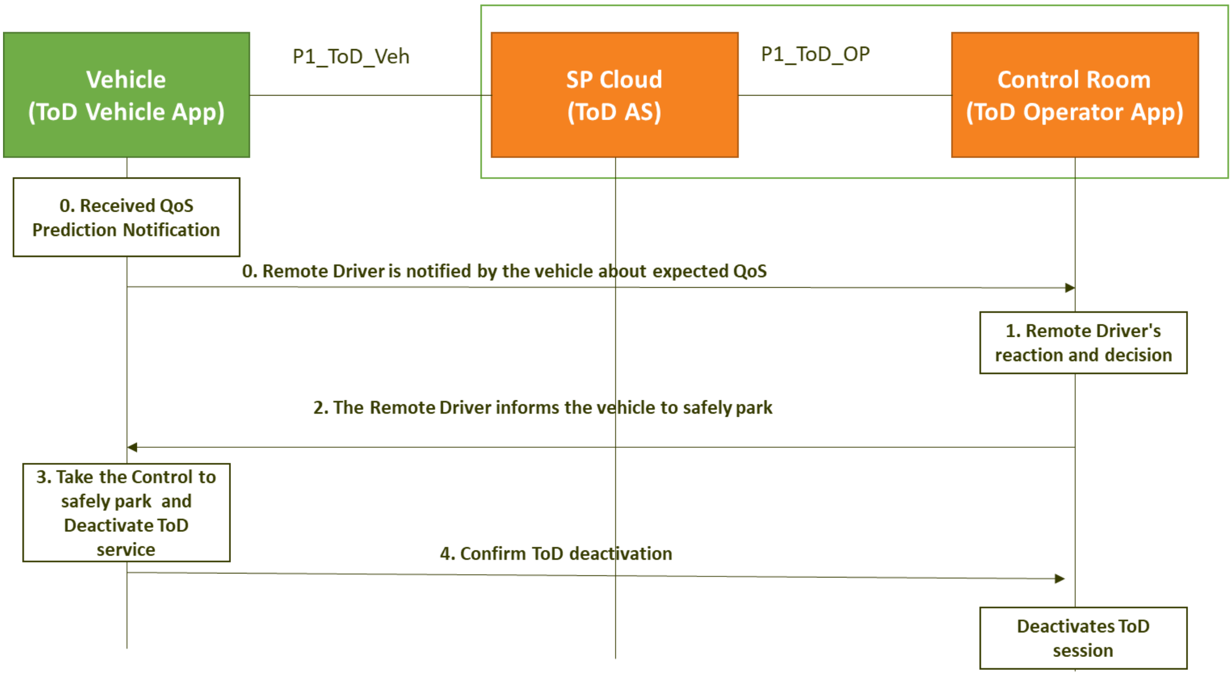


Figure 4: Example Flow Chart for ToD Reaction, Scenario #3: The Remote Driver asks the vehicle to safely stop itself and release the ToD service

Table 9: Time Breakdown Analysis for ToD Reaction, Scenario #3: The Remote Driver asks the vehicle to safely stop itself and release the ToD service

Application Reaction: Scenario #3, The Remote Driver asks the vehicle to safely stop itself and release the ToD service		
Road Environment Type: Highway, Rural, Urban		
Vehicle Type: Passenger vehicle		
Human Intervention: Yes		
List Steps	**Latency (ms)**	**Explanation**
1. The Remote Driver is notified about expected QoS degradation or Coverage Loss and initiates his/her reaction **It should be noted that Step 0 could be omitted in the event that the notification is sent directly by the network to the Control Room (ToD Operator App)**	2,000	Attentive and trained 'safety driver', this could be less than 2 second to start a simple action (just starting to hit the brakes)
4. The Remote Driver informs the vehicle to park itself and the release of ToD Service	500	Time that Remote Driver needs to communicate to the Remote Vehicle the decision to park safely itself
5. The Remote Vehicle takes control, starts the parking system and bulids the environmental awareness.	500	Time needed by the vehicle to de-activate any ToD process and switch to a mode of autonomous driving includes time needed

		to build required environmental awareness to drive autonomously
6. The Remote Vehicle stops safely (Braking Distance)	7,000	Depends on the speed of the vehicle; in ToD scenario, the assumed maximum speed of 50km/h and also that the vehicle can stop at the right side of the road, without searching for appropriate parking spot Other factors, such as weather conditions, surface, vehicle type affect this distance
7. Remote Vehicle deactivates ToD application and informs the Remote Driver	500	Time needed by the vehicle to de-activate any ToD process and switch back to normal mode
Total Sum:	10,500	

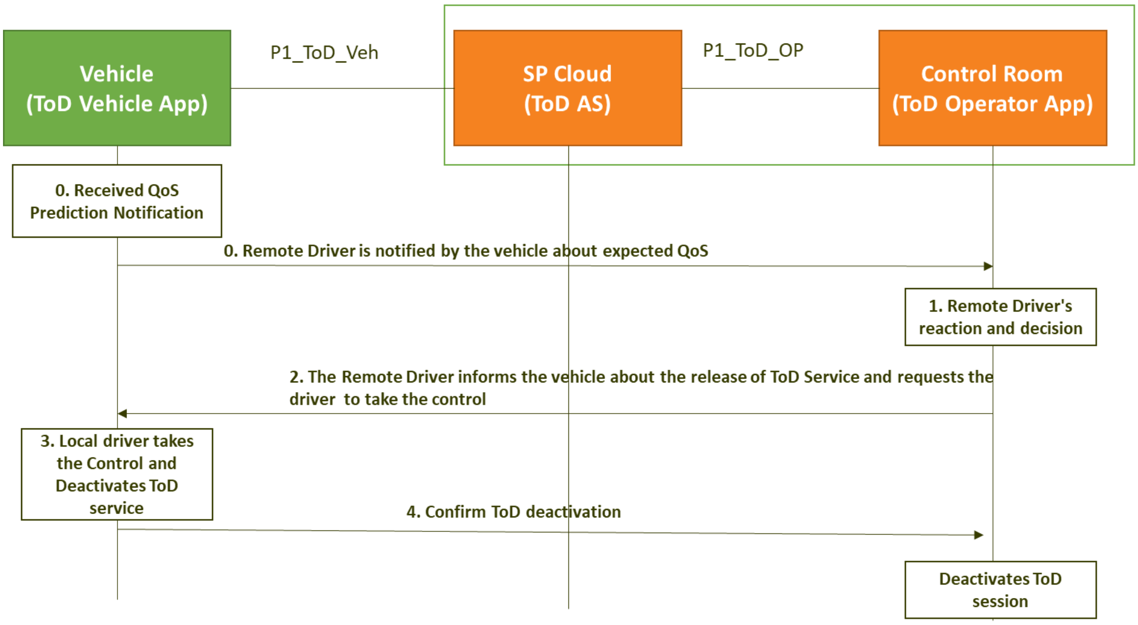


Figure 5: Example Flow Chart for ToD Reaction, Scenario #4: The Remote Driver hands over the control to the driver in car and releases the ToD service

Table 10: Time Breakdown Analysis for ToD Reaction, Scenario #4: The Remote Driver hands over the control to the driver in car and releases the ToD service

Application Reaction: Scenario #4, The Remote Driver hands over the control to the driver in car and releases the ToD service		
Road Environment Type: Highway, Rural, Urban		
Vehicle Type: Passenger vehicle		
Human Intervention: Yes		
List Steps	**Latency (ms)**	**Explanation**
1. The Remote Driver is notified about expected QoS degradation or	2,000	Attentive and trained 'safety driver', this could be less than 2 second to start a

Coverage Loss and initiates his/her reaction **It should be noted that Step 0 could be omitted in the event that the notification is sent directly by the network to the Control Room (ToD Operator App)**		simple action (just starting to hit the brakes)
2. The Remote Driver informs the vehicle about the release of ToD Service and requests the driver to take the control	500	Time that remote driver needs to de-activate ToD and communication time from Remote Driver to the vehicle that is remotely driven
3. The local driver takes over the control of the vehicle	4,000	Driver must be re-engaged to resume control; several seconds (+/-4 seconds) might be suitable in realistic but unexpected scenarios If within 4 seconds there is no reaction by the local driver on this request, then the scenario 1 can be activated (i.e. safe stop) The assumed duration of this step can be larger in a worst-case scenario
4. Remote Vehicle has deactivated the ToD application and informs the Remote Driver	500	
Total Sum:	7,500	

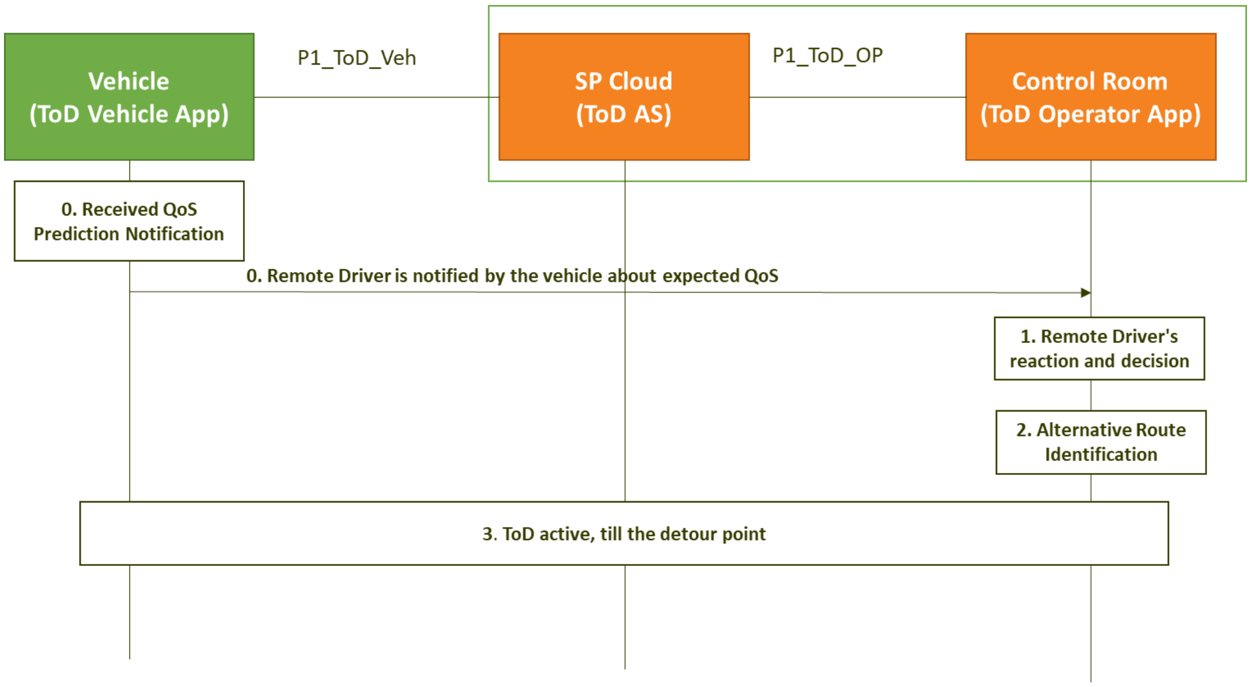


Figure 6: Example Flow Chart for ToD Reaction, Scenario #5: Detour of the remotely driven vehicle via another route

Table 11: Time Breakdown Analysis for ToD Reaction, Scenario #5: Detour of the remotely driven vehicle via another route

Application Reaction: Scenario #5, Detour of the remotely driven vehicle via another route		
Road Environment Type: Highway, Rural, Urban		
Vehicle Type: Passenger vehicle		
Human Intervention: Yes		
List Steps	**Latency (ms)**	**Explanation**
1. The Remote Driver is notified about expected QoS degradation or Coverage Loss and initiates his/her reaction **It should be noted that Step 0 could be omitted in the event that the notification is sent directly by the network to the Control Room (ToD Operator App)**	2,000	Attentive and trained 'safety driver,' this could be less than 2 second to start a simple action (just starting to hit the brakes)
2. Time needed for Remote Driver to find alternative route, e.g. using navigation system	10,000	In this application reaction, it is considered that the remote driver gets three alternative route suggestions to choose; thus, within the assumed period the remote driver analyses the suggested routes on the map, decides the most appropriate one and activates it If for example only one alternative route was suggested the considered time could be lower because then the remote driver would only have to analyse a single suggested track and accept or reject it If none of the routes is accepted by the Remote Driver, another application reaction could follow, e.g. adapting driving properties (see below)
Total Sum:	12,000	

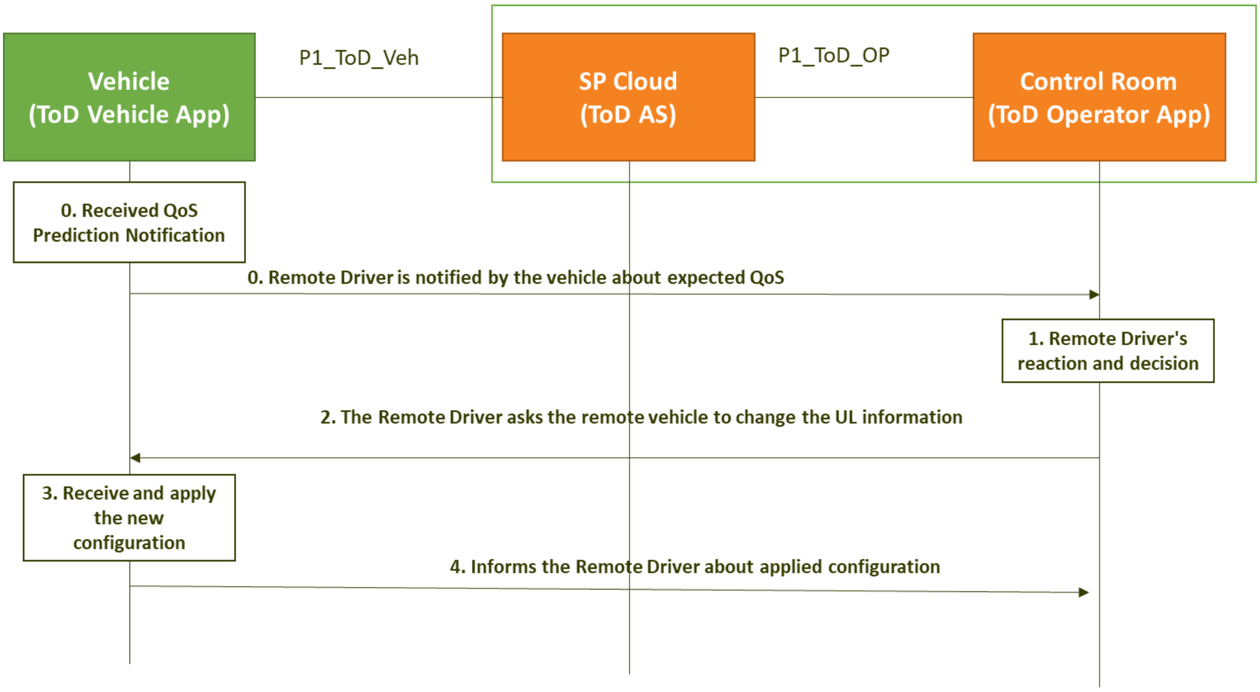


Figure 7: Example Flow Chart for ToD Reaction, Scenario #6: Change sensor set and sensor properties and/or video configuration

Table 12: Time Breakdown Analysis for ToD Reaction, Scenario #6: Change sensor set and sensor properties and/or video configuration

Application Reaction: Scenario #6, Change sensor set and sensor properties and/or video configuration		
Road Environment Type: Highway, Rural, Urban		
Vehicle Type: Passenger vehicle		
Human Intervention: Yes		
List Steps	**Latency (ms)**	**Explanation**
1. The Remote Driver is notified about expected QoS degradation and initiates his/her reaction **It should be noted that Step 0 could be omitted in the event that the notification is sent directly by the network to the Control Room (ToD Operator App)**	2,000	Attentive and trained 'safety driver,' this could be less than 2 second to start a simple action (just starting to hit the brakes)
2. The Remote Driver asks the Remote Vehicle to change the UL information received from sensors or cameras and/or related configuration (e.g. use video with lower resolution, deactivate one of the cameras), based on new QoS	10,000	Time that the Remote Driver (or the Remote Vehicle) needs to select from alternative configurations (e.g. from a list of preconfigured configurations)

Alternatively, the vehicle could automatically perform the appropriate application adaptation		
3. Remote Vehicle receives and applies the new configuration	2,000	Time that vehicles needs to apply re-configuration of ToD service (software adaptation)
4. Remote Vehicle informs the Remote Driver and continues driving with new reporting configuration	500	
Total Sum:	14,500	

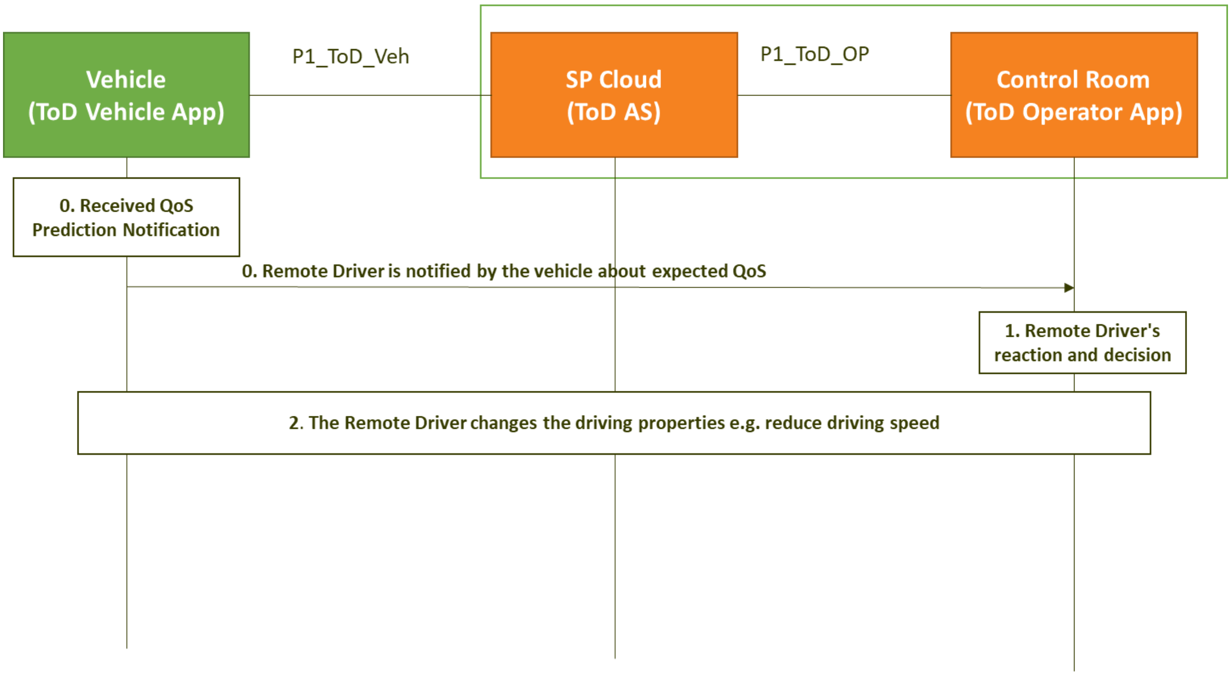


Figure 8: Example Flow Chart for ToD Reaction, Scenario #7: Change driving properties

Table 13: Time Breakdown Analysis for ToD Reaction, Scenario #7: Change driving properties

Application Reaction: Scenario #7, Change driving properties e.g. speed		
Road Environment Type: Highway, Rural, Urban		
Vehicle Type: Passenger vehicle		
Human Intervention: Yes		
List Steps	**Latency (ms)**	**Explanation**
1. The Remote Driver is notified about expected QoS degradation and initiates his/her reaction **It should be noted that Step 0 could be omitted in the event that the notification is sent directly by the**	2,000	Attentive and trained ‘safety driver,’ this could be less than 2 second to start a simple action (just starting to hit the brakes)

network to the Control Room (ToD Operator App).		
2. The Remote Driver changes the driving properties e.g. reduces driving speed	7,000	Depends on the speed of the vehicle; in ToD scenario, it is assumed maximum speed of 50km/h and that we are braking to 10km/h from 50km/h Other factors such as lane that the vehicle is located, weather conditions, surface, vehicle type may affect this distance The worst-case deceleration is $-4m/s^2$ according to the ISO 23793-Intelligent transport systems – Minimal Risk Manoeuvre (MRM) for automated driving – Part 1: Framework, straight-stop and in-lane stop
Total Sum:	9,000	

Table 14 provides an overview of the estimated QoS prediction time horizons, considering the different types of application adaptations of the ToD service.

Table 14: QoS Prediction Time Horizon of Different Application Reactions of ToD Use Case

Use Case Title	**Tele-operated Driving**		
User Story	Based on the perceived environment, the Remote Driver provides to the vehicle that is remotely driven the appropriate manoeuvre instructions to navigate to the destination efficiently and safely (**Direct Control)**		
Application Reaction	**Application Reaction Completion Interval (ARCI) (in ms)**	**Guard Prediction interval (in ms)**	**QoS Prediction Time Horizon[8] (in ms)**
Scenario #1 – The Remote Driver stops safely the vehicle and Release the ToD service	10,000	3,000	13,000
Scenario #2 – The Remote Driver hands over the control to the vehicle (using local sensors) and Release the ToD service	3,000	3,000	6,000
Scenario #3 – The Remote Driver asks the vehicle to park itself and Release the ToD service.	10,500	3,000	13,500
Scenario #4 – The Remote Driver hands	7,500	3,000	10,500

[8] The QoS prediction time horizon is equal to the QPN Notice Period and Transfer Interval that have been specified in NESQO TR [7].

over the control to the driver in car and Release the ToD service			
Scenario #5 – Detour	12,000	3,000	15,000
Scenario #6 – Change sensor set and sensor properties and/or video configuration	14,500	3,000	17,500
Scenario #7 – Change driving properties e.g. speed	9,000	3,000	12,000

Table 15 **provides examples of QoS Change Thresholds** of the Interface F1 (Uu) that could trigger the QoS prediction notification. For instance, the developer or operator of the ToD application can specify the thresholds for degradation (or improvement) of QoS parameters that are acceptable for the efficient, smooth and safe operation of the ToD service. It should be noted that different implementations of a specific ToD service can consider different threshold values. Hence, generic threshold values cannot be defined that apply to different implementations or deployments. Also, alternative QoS levels could be considered for a specific use case, if needed.

An alternative option is that a prediction notification is not based on specific QoS change thresholds that will trigger the transmission (or not) of a notification. But rather the application subscribes to receive periodic notifications (e.g. every 5 seconds) and the application (or the vehicle) can decide to adapt or not according to the received notification.

Table 15: QoS Change Thresholds of QoS Predictions for ToD Use Case

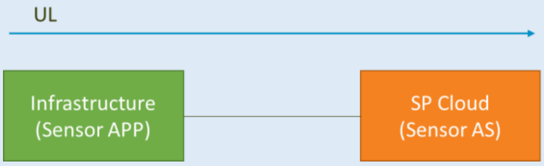
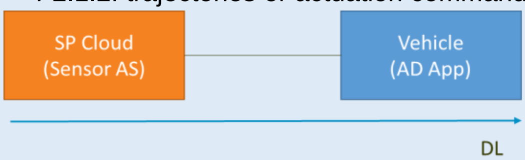
Use Case Title	**Tele-operated Driving**		
User Story	Based on the perceived environment, the Remote Driver provides to the vehicle that is remotely driven the appropriate manoeuvre instructions to navigate to the destination efficiently and safely (**Direct Control**)		
Predicted QoS/Service Level Requirements	**Considered Interface(s)**	**QoS/SLR value (5GAA TRs)**	**QoS Change Thresholds for Triggering Prediction Notification**
UL Throughput	Uu (F1.1)	• Video (for human ToD operator) Up to 32 Mb/s	e.g. UL Throughput lower than 30 Mbps
		• Object information Up to 4 Mb/s	e.g. UL Throughput lower than 3.5 Mbps
		• Audio ~96 kb/s	-
		• Vehicle information (e.g. speed, acceleration, vehicle position) ~0.2 Mb/s	-
UL Latency	Uu (F1.1)	• 40ms	e.g. UL latency higher than 45ms
UL Reliability	Uu (F1.1)	• Video (for human ToD operator) 99%	e.g. UL reliability lower 98%
		• Object information 99%	e.g. UL reliability lower 98%
DL Latency	Uu (F1.2)	• Vehicle maneuverer commands (Direct Control ToD type): 20 ms	e.g. UDL latency higher than 21ms

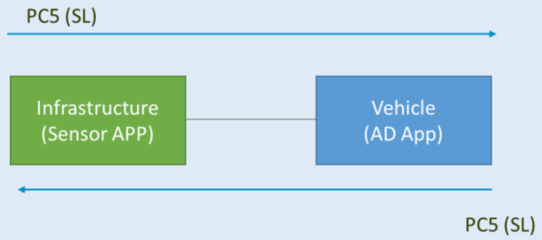
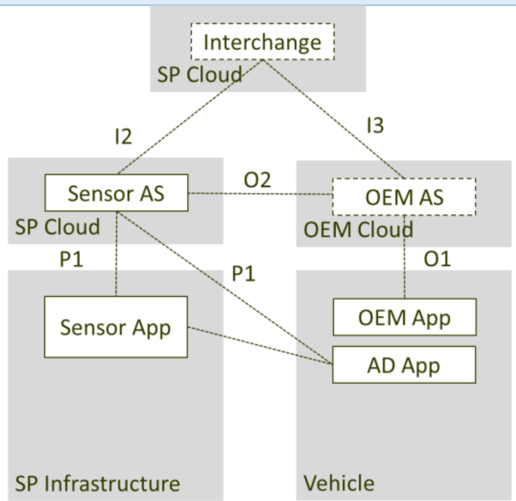
DL Reliability	Uu (F1.2)	• Vehicle maneuverer commands (Direct Control ToD type): 99.9% or higher	e.g. DL reliability lower than 99.8%

5.2.2 Infrastructure Assisted Environment Perception

This section presents the analysis of QoS prediction parameters using as an example the Infrastructure Assisted Environment Perception use case.

Table 16: Potential Application Reactions of Infrastructure Assisted Environment Perception Use Case

Use Case Title	Infrastructure Assisted Environment Perception
User Story	When an automated vehicle enters a section of the road covered by infrastructure sensors, it enrols to receive information from the infrastructure containing environment data provided by dynamic and static objects on the road; this data is used to increase the trust level of the car's own sensor observations and extends its viewing range
Type of Service	Periodic service
Predicted QoS (Service Level Requirements) of Interest per Communication Link	F1. Infrastructure – Application Server: - F1.1: Uu Interface, UL (from Infrastructure to Application Server): data rate, latency, jitter and reliability (unicast) o F1.1.1: Sensor information o F1.1.2: Video streaming o F1.1.3: Object lists o ….  F2. Application Server – Vehicle: - F2.1: Uu Interface, DL (from Application Server to Vehicle): data rate, reliability, latency and jitter o F2.1.1: Object lists (broadcast) o F2.1.2: trajectories or actuation commands (unicast)  F3. Infrastructure – Vehicle: - F3.1: PC5 (Sidelink) Interface, (from the Infrastructure to the Vehicle): Data Rate, latency, jitter, reliability (broadcast) o F3.1.1: Object lists

	
Considered Interface(s)	 Considered Interface Architecture[9]
QoS Prediction Recipient and/or Orchestrator for Adaptation?	QoS Prediction Recipient: • Infrastructure – Application Server Interface (F1.1): Application Server or/and Infrastructure • Application Server – Vehicle Interface (F2.1): Application Server or/and Vehicle • Infrastructure – Vehicle Interface (F3.1): Infrastructure or/and Vehicle Orchestrator for application adaptation: Vehicle, Infrastructure, Application Server The type of adaptation that will be selected depends on the actual prediction quality, the road environment, and the thresholds that the application developer/owner sets or even on some regulatory decisions
Involved Entities	Vehicle, Application Server, Infrastructure

Reaction Scenario #	**Potential Application Reactions**	**Entity that Triggers the Application Reaction**	**SLR Change (e.g. Data Rate, Coverage, Latency, Reliability, Position Accuracy)**
1.	The AV changes the confidence level of data from the Infrastructure via network	Vehicle	Uu Interface: • Reduction of UL Data Rate (F1.1) below defined thresholds and/or

[9] Reference: WG2_WI_V2XSRA 5GAA_A-200094_V2XSRA Application Layer Reference Architecture.docx

			• Increase of Latency (F1.1&F2.1) above defined thresholds and/or • Reduction of Reliability (F1.1&F2.1)below defined thresholds PC5 Interface: • Reduction of Sidelink Data Rate (F3.1) below defined thresholds and/or • Increase of Latency (F3.1) above defined thresholds and/or • Reduction of Reliability (F3.1)below defined thresholds
2.	The AV ignores Infrastructure's data and drives only with its own sensor observations	Vehicle	Uu Interface: • Coverage Loss and/or • Reduction of UL Data Rate (F1.1) below defined thresholds and/or • Increase of Latency (F1.1&F2.1) above defined thresholds and/or • Reduction of Reliability (F1.1&F2.1)below defined thresholds PC5 Interface: • Connection Quality Decreasing due to increased range and/or • Reduction of Data Rate (F3.1) below defined thresholds and/or • Increase of Latency (F3.1) above defined thresholds and/or • Reduction of Reliability (F3.1)below defined thresholds
3.	The AV changes driving properties e.g. speed	Vehicle	Uu Interface: • Coverage Loss and/or • Reduction of UL Data Rate (F1.1) below defined thresholds and/or • Increase of Latency (F1.1&F2.1) above defined thresholds and/or • Reduction of Reliability (F1.1&F2.1) below defined thresholds PC5 Interface: • Connection Quality Decreasing due to increased range and/or • Reduction of Data Rate (F3.1) below defined thresholds and/or • Increase of Latency (F3.1) above defined thresholds and/or • Reduction of Reliability (F3.1)below defined thresholds

4.	The AV hands over the control to the driver in car and releases the automated driving system	Vehicle	Uu Interface: • Coverage Loss and/or • Reduction of UL Data Rate (F1.1) below defined thresholds and/or • Increase of Latency (F1.1&F2.1) above defined thresholds and/or • Reduction of Reliability (F1.1&F2.1) below defined thresholds PC5 Interface: • Connection Quality Decreasing due to increased range and/or • Reduction of Data Rate (F3.1) below defined thresholds and/or • Increase of Latency (F3.1) above defined thresholds and/or • Reduction of Reliability (F3.1)below defined thresholds
5.	The AV stops safely and releases the automated driving system	Vehicle	Uu Interface: • Coverage Loss and/or • Reduction of UL Data Rate (F1.1) below defined thresholds and/or • Increase of Latency (F1.1&F2.1) above defined thresholds and/or • Reduction of Reliability (F1.1&F2.1)below defined thresholds PC5 Interface: • Connection Quality Decreasing due to increased range and/or • Reduction of Data Rate (F3.1) below defined thresholds and/or • Increase of Latency (F3.1) above defined thresholds and/or • Reduction of Reliability (F3.1)below defined thresholds
6.	The Infrastructure only transmits object lists, without raw sensor data	Infrastructure	Uu Interface: • Reduction of UL Data Rate (F1.1) below defined thresholds PC5 Interface: • Reduction of Data Rate (F3.1) below defined thresholds
7.	The Infrastructure stops transmitting any data and informs Application Server and Vehicle	Infrastructure, Vehicle	Uu Interface: • Coverage Loss and/or • Reduction of UL Data Rate (F1.1) below defined thresholds and/or

			• Increase of Latency (F1.1&F2.1) above defined thresholds and/or • Reduction of Reliability (F1.1&F2.1) below defined thresholds PC5 Interface: • Connection Quality Decreasing due to increased range and/or • Reduction of Data Rate (F3.1) below defined thresholds and/or • Increase of Latency (F3.1) above defined thresholds and/or • Reduction of Reliability (F3.1) below defined thresholds
8.	The Infrastructure/Application Server increases the frequency of messages for Vehicle	Infrastructure or Application Server	Uu Interface: • Reduction of Reliability (F1.1&F2.1) below defined thresholds PC5 Interface: • Reduction of Reliability (F3.1) below defined thresholds
9.	The Application Server only sends object lists, and stops the feature sending trajectories or actuation commands for Vehicle	Application Server	Uu Interface: • Reduction of Reliability (F1.1&F2.1) below defined thresholds

NOTE: for the same change of SLR, reaction can be different per different vehicles.

The following tables present the Time Breakdown Analysis for each potential application reaction of the Infrastructure Assisted Environment Perception use case.

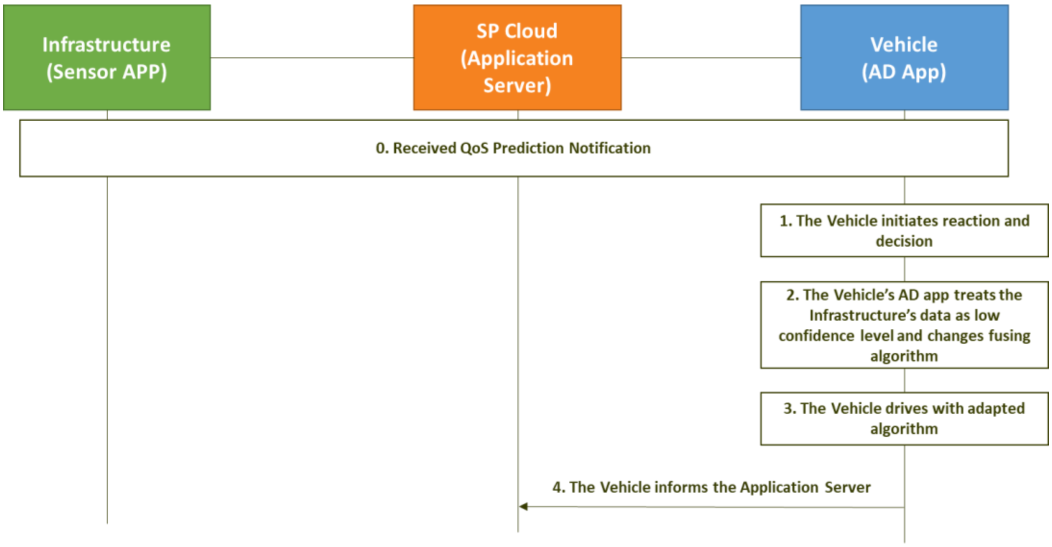


Figure 9: Example Flow Chart for Infrastructure Assisted Environment Reaction, Scenario #1: The AV changes the confidence level of data from the Infrastructure via network

Table 17: Time Breakdown Analysis for Infrastructure Assisted Environment Reaction, Scenario #1: The AV changes the confidence level of data from the Infrastructure via network

Application Reaction: Scenario #1 – The AV changes the confidence level of data from the Infrastructure via network		
Road Environment Type: Highway, Rural, Urban		
Vehicle Type: Passenger vehicle		
Human Intervention: No		
List Steps	**Latency (ms)**	**Explanation**
6. The Vehicle's AD Application is notified about expected QoS degradation and initiates its reaction	500	Time for the Vehicle to receive the P-QoS message and analyse the degradation level/threshold to determine a suitable reaction
7. The Vehicle's AD Application treats the Infrastructure's data with low confidence level and changes fusing algorithm	500	During normal mode, the Vehicle uses both Infrastructure sensor data and its own sensor data with a 'high confidence' coefficient to generate the environment perception input for the AV system; when the SLR degrades, the Infrastructure's sensor data may be longer latency or lower reliability, so the Vehicle treats the Infrastructure's data as 'low confidence' Time for the Vehicle to determine the degraded level of Infrastructure data and change the AV algorithm
8. The Vehicle drives with adapted algorithm	2,000	The Vehicle actives the adapted AV algorithm and change to the adapted AV driving mode includes time needed to

		rebuild environmental awareness with new confidence coefficients, and the decision-making and motion-planning system adapts with the rebuilt awareness inputs, then the vehicle mechanical parts adjust following the AV system's new algorithm
9. The Vehicle informs the Application Server	200	Time for the Vehicle to generate the notification message and send it to the Application Server
Total Sum:	3,200	

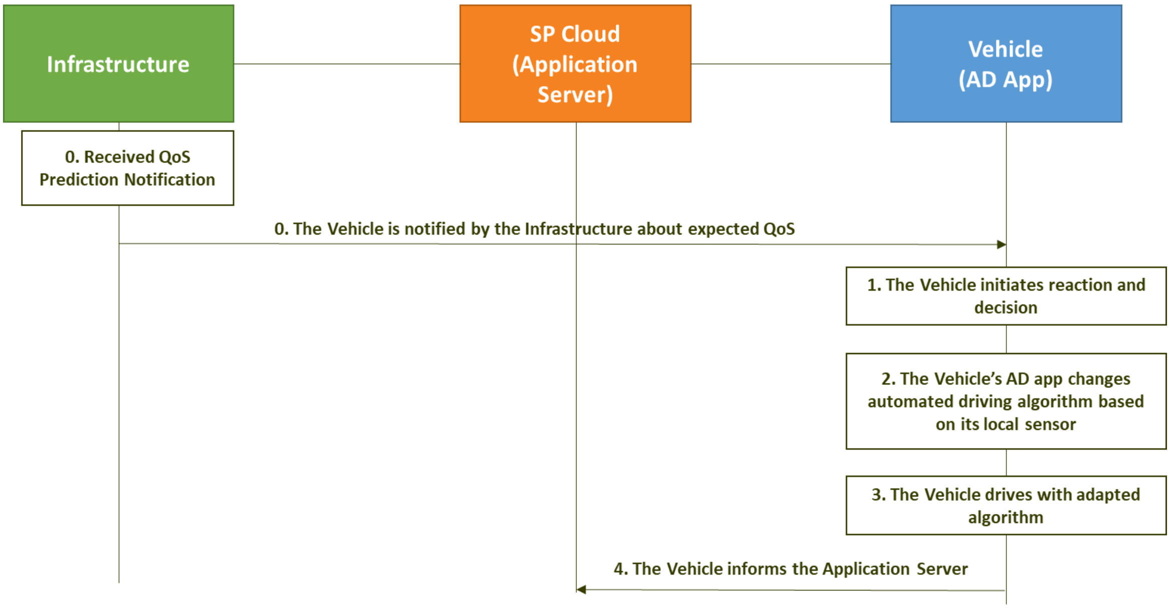


Figure 10: Example Flow Chart for Infrastructure Assisted Environment Reaction, Scenario #2: The AV ignores Infrastructure's data and drives only with its own sensor observations

Table 18: Time Breakdown Analysis for Infrastructure Assisted Environment Reaction, Scenario #2: The AV ignores Infrastructure's data and drives only with its own sensor observations

Application Reaction: Scenario #2, The AV ignores Infrastructure's data and drives only with its own sensor observations		
Road Environment Type: Highway, Rural, Urban		
Vehicle Type: Passenger vehicle		
Human Intervention: No		
List Steps	**Latency (ms)**	**Explanation**
1. The Vehicle's AD Application is notified about expected QoS degradation or Coverage Loss and initiates its reaction	500	Time for the Vehicle to receive the P-QoS message and analyse the degradation level/threshold to determine a suitable reaction
2. The Vehicle's AD Application changes automated driving algorithm based on its local sensor	500	In normal mode, the Vehicle uses both Infrastructure sensor data and its own sensor data with a 'high confidence'

		coefficient to generate the environment perception input for AV system; when the SLR degrades, the Infrastructure's sensor data may be delayed too much or unreliable, so the Vehicle treats the Infrastructure's data as useless messages and drives only based on its own sensors Time for the Vehicle to determine the degraded level for Infrastructure data and change the AV algorithm
3. The Vehicle drives with adapted algorithm	2,000	The Vehicle actives the adapted AV algorithm and changes to the adapted AV driving mode including time needed to rebuild environmental awareness with its own sensors only, and for the decision-making and motion-planning systems to adapt to the rebuilt awareness inputs, then the Vehicle mechanical parts adjusts following the AV system's new algorithm
4. The Vehicle informs the Application Server	200	Time for the Vehicle generate the notify message and send it to the Application Server
Total Sum:	3,200	

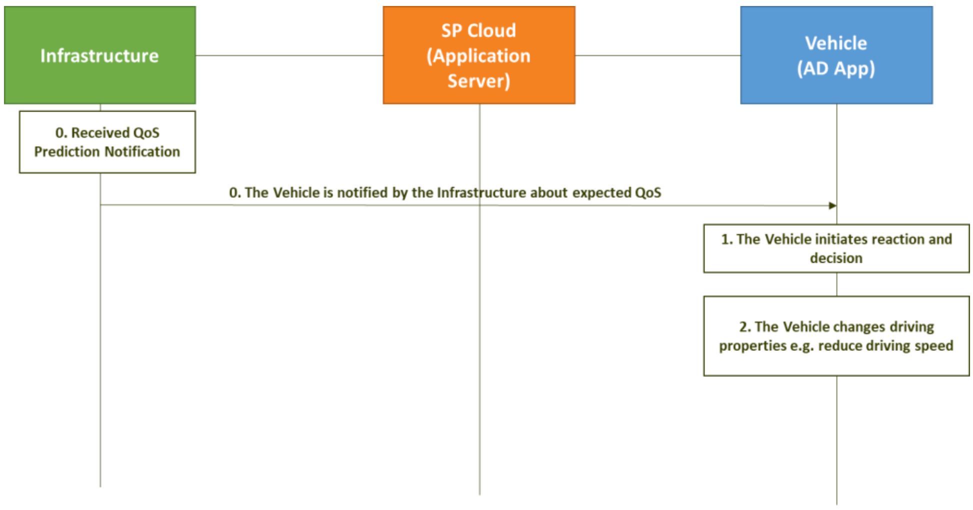


Figure 11: Example Flow Chart for Infrastructure Assisted Environment Reaction, Scenario #3: The AV changes driving properties, e.g. speed

Table 19: Time Breakdown Analysis for Infrastructure Assisted Environment Reaction, Scenario #3: The AV changes driving properties, e.g. speed

Application Reaction: Scenario #3, The AV changes driving properties, e.g. speed
Road Environment Type: Highway, Rural, Urban
Vehicle Type: Passenger vehicle

Human Intervention: No		
List Steps	Latency (ms)	Explanation
1. **The Vehicle's AD Application is notified about expected QoS degradation or Coverage Loss and initiates its reaction**	500	Time for the Vehicle to receive the P-QoS message and analyse the degradation level/threshold to determine a suitable reaction
2. **The Vehicle changes driving properties, e.g. reduce driving speed**	7,000	Depends on the speed of the Vehicle; in the AD scenario, it was assumed braking from 50km/h to 10km/h Other factors, such as the lane the Vehicle is located in, weather conditions, surface, vehicle type may affect this distance
Total Sum:	7,500	

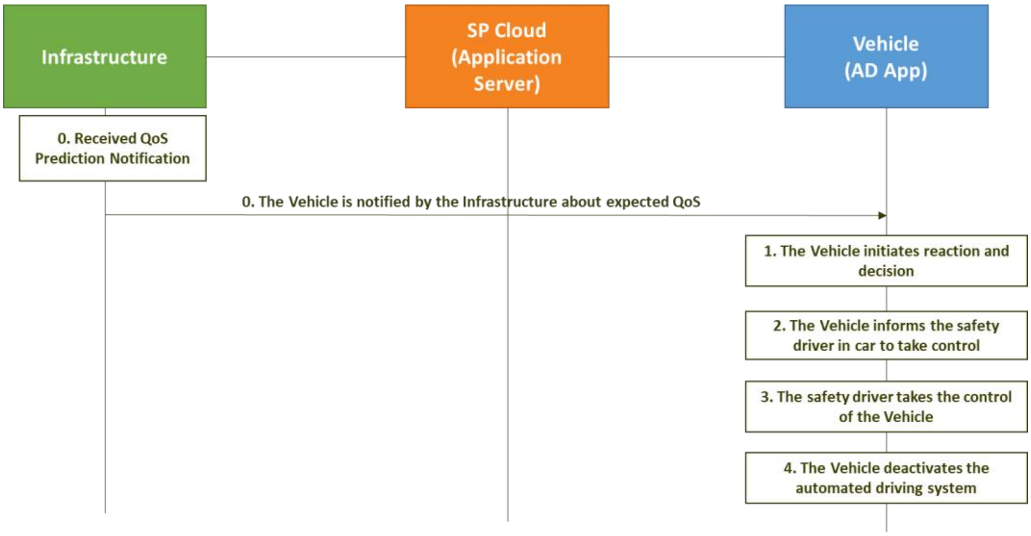


Figure 12: Example Flow Chart for Infrastructure Assisted Environment Reaction, Scenario #4: The AV hands over the control to the driver in car and releases the automated driving system

Table 20: Time Breakdown Analysis for Infrastructure Assisted Environment Reaction, Scenario #4: The AV hands over the control to the driver in car and releases the automated driving system

Application Reaction: Scenario #4, The AV hands over the control to the driver in car and releases the automated driving system		
Road Environment Type: Highway, Rural, Urban		
Vehicle Type: Passenger vehicle		
Human Intervention: Yes		
List Steps	Latency (ms)	Explanation
1. **The Vehicle's AD app is notified about expected QoS degradation or Coverage Loss and initiates its reaction**	500	Time for the Vehicle to receive the P-QoS message and analyse the degradation

		level/threshold to determine a suitable reaction
2. The Vehicle informs the normal driver in car to take control	500	Time needed for the Vehicle to generate the message for the driver and show symbols or words on the HMI screen or vibrate to inform the driver
3. The driver takes control of the Vehicle	10,000	Driver must be re-engaged to resume control; 10 seconds might be suitable in realistic but unexpected scenarios, and the timing value is also considered based on UNR-157 If within 10 seconds, there is no reaction by the local driver on this request, then scenario 5 can be activated (i.e. safe stop)
4. The Vehicle deactivates the automated driving system	500	
Total Sum:	11,500	

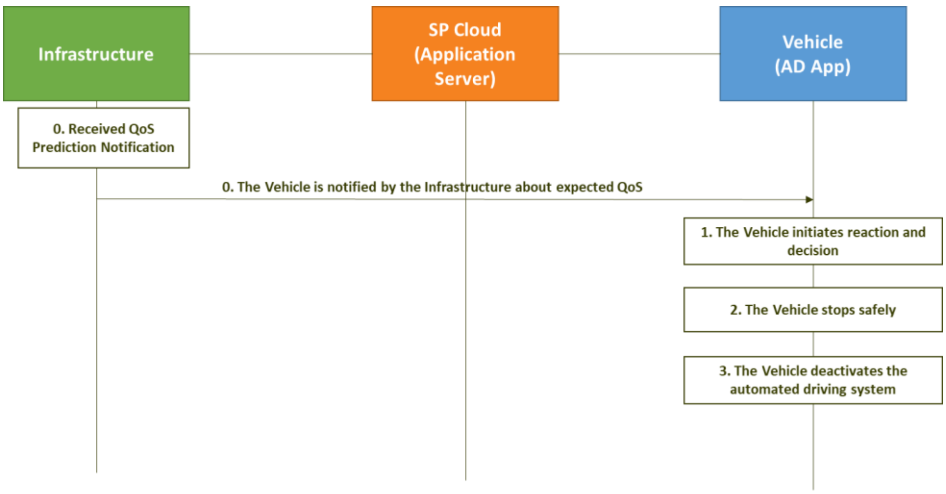


Figure 13: Example Flow Chart for Infrastructure Assisted Environment Reaction, Scenario #5: The AV stops safely and releases the automated driving system

Table 21: Time Breakdown Analysis for Infrastructure Assisted Environment Reaction, Scenario #5: The AV stops safely and releases the automated driving system

Application Reaction: Scenario #5, The AV stops safely and releases the automated driving system		
Road Environment Type: Highway, Rural, Urban		
Vehicle Type: Passenger vehicle		
Human Intervention: No		
List Steps	**Latency (ms)**	**Explanation**
1. The Vehicle's AD app is notified about expected QoS degradation or	500	Time for the Vehicle to receive the P-QoS message and analyse the degradation

Coverage Loss and initiates its reaction		level/threshold to determine a suitable reaction
2. The Vehicle stops safely	7,000	Depends on the speed of the Vehicle; in the AD scenario, it is assumed the maximum speed is 50km/h and also that the Vehicle can stop at the right side of the road, without searching for appropriate parking spot Other factors, such as the lane the Vehicle is located in, weather conditions, surface, vehicle type may affect this distance If the Vehicle needs to search a safety parking spot, the time needed could be longer
3. The Vehicle deactivates the automated driving system	500	
Total Sum:	8,000	

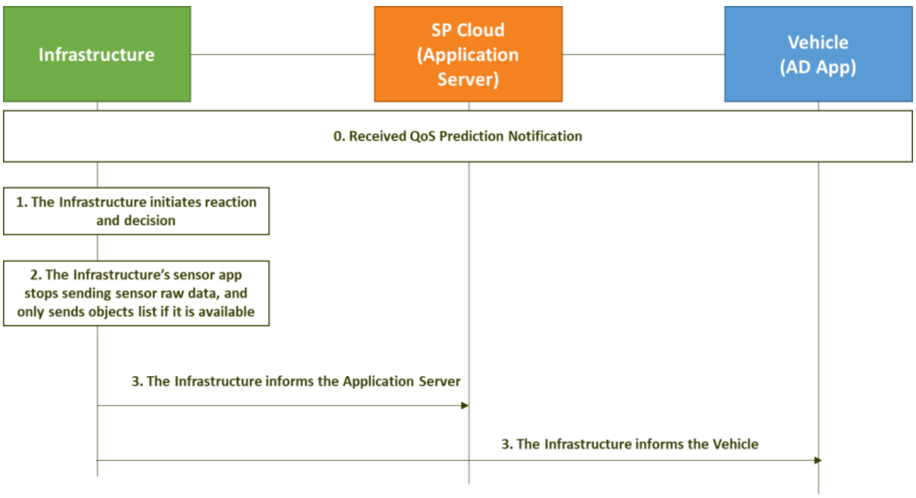


Figure 14: Example Flow Chart for Infrastructure Assisted Environment Reaction, Scenario #6: The Infrastructure only transmits object lists, without raw sensor data

Table 22: Time Breakdown Analysis for Infrastructure Assisted Environment Reaction, Scenario #6: The Infrastructure only transmits object lists, without raw sensor data

Application Reaction: Scenario #6, The Infrastructure only transmits object lists, without raw sensor data		
Road Environment Type: Highway, Rural, Urban		
Vehicle Type: Passenger vehicle		
Human Intervention: No		
List Steps	**Latency (ms)**	**Explanation**
1. The Infrastructure is notified about expected QoS degradation and initiates its reaction	500	Time for the Infrastructure to receive the P-QoS message and analyse the

		degradation level/threshold to determine a suitable reaction
2. The Infrastructure's sensor app stops sending sensor raw data, and only sends object lists if it is available **(In this case, the pre-condition is that the Infrastructure sensors could generate both raw data and object lists)**	500	Since transmitting the raw sensor data needs a higher date rate, only object lists could be sent to the Vehicle when data rate is deduced to be below defined thresholds
3. The Infrastructure informs the Application Server and Vehicle	500	Time needed for the Infrastructure generating the message to transmit it to the Application Server and Vehicle
Total Sum:	1,500	

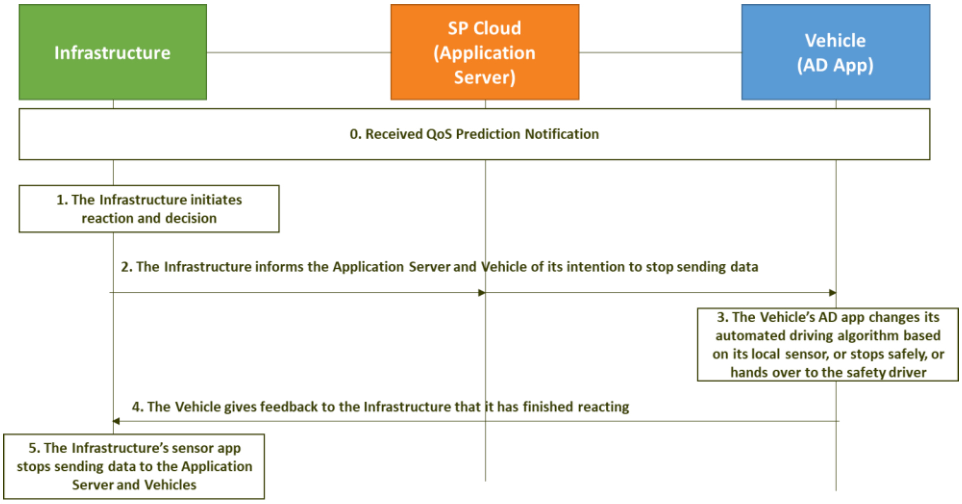


Figure 15: Example Flow Chart for Infrastructure Assisted Environment Reaction, Scenario #7: The Infrastructure stops transmitting data and informs the Application Server and Vehicle

Table 23: Time Breakdown Analysis for Infrastructure Assisted Environment Reaction, Scenario #7: The Infrastructure stops transmitting data and informs the Application Server and vehicle

Application Reaction: Scenario #7, The Infrastructure stops transmitting data and informs the Application Server and Vehicle		
Road Environment Type: Highway, Rural, Urban		
Vehicle Type: Passenger vehicle		
Human Intervention: No		
List Steps	**Latency (ms)**	**Explanation**
1. The Infrastructure is notified about expected QoS degradation and initiates its reaction	500	Time for the Infrastructure to receive the P-QoS message and analyse the degradation level/threshold to determine a suitable reaction

2. The Infrastructure's sensor app intends to stop sending any data to Application Server and Vehicle, and the Infrastructure informs the Application Server and Vehicle its intention in advance	500	Time needed for the Infrastructure generating the message and transmitting it to the Application Server and Vehicle, and for the Application Server or Vehicle de-coding the message
3. The Vehicle's AD Application changes the automated driving algorithm based on its local sensor, or stops safely, or hand over to the safety driver (scenario #2/4/5)	7,000	Refer to scenario #2/4/5, the maximum time could be 7 seconds
4. The Vehicle hands over feedback to the Infrastructure that it has finished reacting	500	
5. The Infrastructure's sensor app stops sending data to the Application Server and Vehicle	200	Time needed for the Infrastructure to de-activate the sensor function
Total Sum:	8,700	

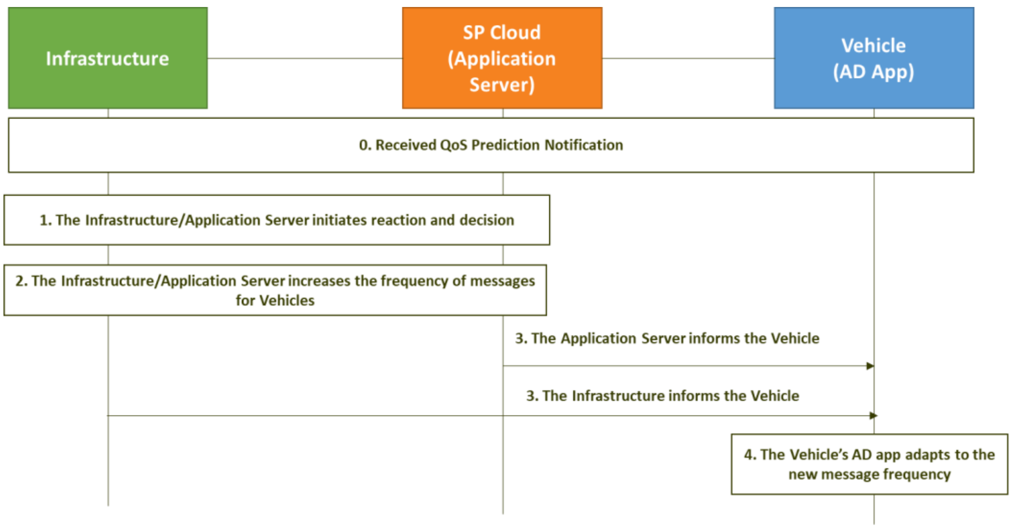


Figure 16: Example Flow Chart for Infrastructure Assisted Environment Reaction, Scenario #8: The Infrastructure/Application Server increases the frequency of messages for Vehicles

Table 24: Time Breakdown Analysis for Infrastructure Assisted Environment Reaction, Scenario #8: The Infrastructure/Application Server increases the frequency of messages for Vehicles

Application Reaction: Scenario #8 – The Infrastructure/Application Server increases the frequency of messages for Vehicles		
Road Environment Type: Highway, Rural, Urban		
Vehicle Type: Passenger vehicle		
Human Intervention: No		
List Steps	**Latency (ms)**	**Explanation**
1. The Infrastructure/Application Server is notified about expected	500	Time for the Infrastructure/application server to receive the P-QoS message and

QoS degradation and initiates its reaction		analyse the degradation level/threshold to determine a suitable reaction
2. The Infrastructure/Application Server increases the frequency of messages for Vehicles	500	In order to compensate for the possible message loss because of reduced transmission reliability, the Infrastructure/Application Server increases the frequency of object list messages
3. The Infrastructure/application informs the Vehicle	500	Time needed for Infrastructure to generate the message and transmit it to the Application Server and Vehicle
4. The Vehicle's AD Application adapts to the new message frequency	2,000	The Vehicle activates the adapted AV algorithm, and changes to the adapted AV driving mode include time needed for decision-making and motion-planning system to adapt with new message frequency, and the Vehicle mechanical parts to adjust following the AV system's new algorithm
Total Sum:	3,500	

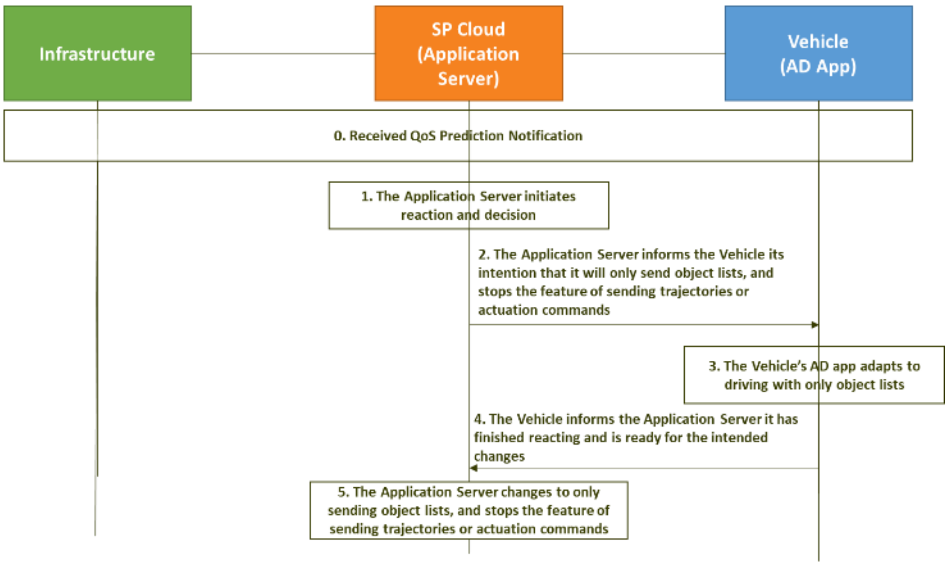


Figure 17: Example Flow Chart for Infrastructure Assisted Environment Reaction, Scenario #9: The Application Server only sends object lists, and stops the feature of sending trajectories or actuation commands for Vehicles

Table 25: Time Breakdown Analysis for Infrastructure Assisted Environment Reaction, Scenario #9: The Application Server only sends object lists, and stops the feature of sending trajectories or actuation commands for Vehicles

Application Reaction: Scenario #9, The Application Server only sends object lists, and stops the feature of sending trajectories or actuation commands for Vehicles
Road Environment Type: Highway, Rural, Urban
Vehicle Type: Passenger vehicle
Human Intervention: No

List Steps	Latency (ms)	Explanation
1. c	500	Time for the Application Server to receive the P-QoS message and analyse the degradation level/threshold to determine a suitable reaction
2. The Application Server plans to send object lists, and stops the feature of sending trajectories or actuation commands for Vehicles; and the Application Server informs the Vehicle its intention in advance **(In this case, the pre-condition is that the Application Server could send both object lists and trajectories or actuation commands for the Vehicle)**	500	Since transmitting the trajectories or actuation commands needs a higher SLRs, only object lists are sent to the Vehicle when the SLR degrades Time needed for the Application Server to generate the message and transmit it to the Vehicle
3. The Vehicle's AD Application adapts to driving with only object lists	2,500	Time needed for decision-making and motion-planning systems to adapt with limited messages, and the Vehicle's mechanical parts to adjust following the AV system's new algorithm
4. The Vehicle informs the Application Server it has finished reacting and is ready for the intended changes	500	Time need for the Vehicle to generate the message and transmit it to the Application Server
5. The Application Server changes to only sending object lists, and stops the feature of sending trajectories or actuation commands	500	
Total Sum:	4,500	

Table 26 provides an overview of the estimated QoS prediction time horizons, considering the different types of application adaptations of the Infrastructure Assisted Environment Perception.

Table 26: QoS Prediction Time Horizon of different Application Reactions of Infrastructure Assisted Environment Perception Use Case

Use Case Title	Infrastructure Assisted Environment Perception		
User Story	When an automated vehicle enters a section of the road covered by infrastructure sensors it enrols to receive information from the infrastructure containing environment data provided by dynamic and static objects on the road. This data is used to increase the trust level of the car's own sensor observations and extends its viewing range.		
Application Reaction	**Application Reaction Completion Interval (=ARCI) (in ms)**	**Guard Prediction interval (in ms)**	**QoS Prediction Time Horizon (in ms)**
Scenario #1 – The AV changes automated algorithm fusing with the Infrastructure's data as low confidence	3,200	3,000	6,200

Scenario #2 – The AV ignores the Infrastructure's data and drives only with its own sensor observations	3,200	3,000	6,200
Scenario #3 – The AV changes driving properties, e.g. speed	7,500	3,000	10,500
Scenario #4 – The AV hands over the control to the driver in car and releases the automated driving system	11,500	3,000	14,500
Scenario #5 – The AV stops safely and releases the automated driving system	8,000	3,000	11,000
Scenario #6 – The Infrastructure only transmits object lists, without raw sensor data	1,500	3,000	4,500
Scenario #7 – The Infrastructure stops transmitting any data and informs Application Server and Vehicle	8,700	3,000	11,700
Scenario #8 – The Infrastructure/Application Server increases the frequency of messages for Vehicles	3,500	3,000	6,500
Scenario #9 – The Application Server only sends object lists, and stops the feature of sending trajectories or actuation commands for Vehicles	4,500	3,000	7,500

Table 27 provides examples of QoS Change Thresholds of the Interfaces that could trigger the QoS prediction notification. The QoS/SLR values and related thresholds for triggering prediction notification are given as examples, and the values are mainly referenced to 5GAA White Paper C-V2X UC&SLR VOL II. It needs to be noted that some SLR values could be ranges, for example, the data rate of RADAR is 40~160kbps (refer to the UC&SLR White Paper), because the detected targets included in the RADAR messages vary over time. In Table 4 the threshold values defined refer to the related SLR upper limits as examples in the event that the network QoS could not meet all dynamic requirements. Indeed, the threshold values could be defined differently based on different deployments or operators' considerations.

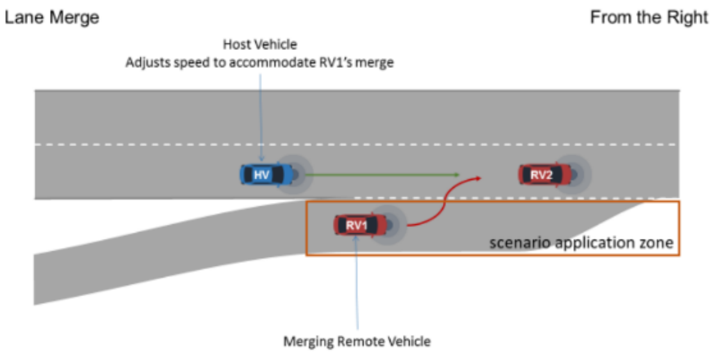
Table 27: QoS Change Thresholds of QoS Predictions for Infrastructure Assisted Environment Perception Use Case

Use Case Title	Infrastructure Assisted Environment Perception
User Story	When an automated vehicle enters a section of the road covered by infrastructure sensors it enrols to receive information from the infrastructure containing environment data provided by dynamic and static objects on the

Predicted QoS/Service Level Requirements	Considered Interface(s)	QoS/SLR value (5GAA TRs)	QoS Change Thresholds for triggering prediction notification
	road. This data is used to increase the trust level of the car's own sensor observations and extends its viewing range.		
Uu – UL Throughput	Uu (F1.1)	• Uncompressed/unprocessed Video streaming up to 120Mbps • Compressed video ~8Mbps	e.g. UL throughput lower than 110Mbps for unprocessed video streaming or throughput lower than 7Mbps for compressed video
		• RADAR up to 160kbps	e.g. throughput lower than 145kbps
		• LIDAR 35Mbps	e.g. throughput lower than 32Mbps
		• Object lists up to 4000kbps	e.g. UL throughput lower than 3600kbps
Uu – UL Latency	Uu (F1.1)	• 50ms	e.g. latency higher than 55ms or Jitter larger than ±5ms
Uu – UL Reliability	Uu (F1.1)	• Sensor data 99%	e.g. reliability lower 98%
		• Object lists 99.99%	e.g. reliability lower 99.98%
Uu – DL Throughput	Uu (F2.1)	• Object lists up to 4000kbps (broadcast)	e.g. DL throughput lower than 2000kbps
		• Trajectories or Actuation Commands 8Mbps (unicast)	e.g. DL throughput lower than 7Mbps
Uu – DL Latency	Uu (F2.1)	• 50ms	e.g. latency higher than 55ms or Jitter larger than ±5ms
Uu – DL Reliability	Uu (F2.1)	• Object lists 99.99%	e.g. reliability lower 99.98%
		• Trajectories or Actuation Commands 99.999%	e.g. reliability lower 99.998%
SL – Throughput	PC5 (F3.1)	• Object lists up to 4000kbps	e.g. throughput lower than 3600kbps
SL – Latency	PC5 (F3.1)	• 100ms	e.g. latency higher than 110ms or Jitter larger than ±5ms
SL – Reliability	PC5 (F3.1)	• Object lists 99.99%	e.g. reliability lower 99.98%

5.2.3 Cooperative Lane Merge

In this section, we apply the template and the methodology for Predictive QoS-related SLRs on the Cooperative Lane Merge [1].



The use story of this use case includes interaction among vehicles, between Remote Vehicle 1 or RV1 (i.e. the vehicle merging into the HV's traffic lane) and Host vehicle or HV (i.e. the vehicle accommodating RV1's manoeuvre). Three messages are exchange before **initiating the Cooperative Lane Merge** manoeuvre**.**

HV receives RV1's intention to apply a lane-merging manoeuvre, providing location, speed and manoeuvre information.

- If there is not a lead vehicleRV2 then
  - HV uses RV1's location and dynamics and the length of the merge to adjust the speed of the HV, such that at the end of the manoeuvre, the HV is positioned in a safe following distance from RV1
  - RV1 is made aware of the HV's intention to accommodate the manoeuvre; it adapts accordingly (if needed) its speed and notifies the HV of its acceptance before initiating the manoeuvre
- If there is a lead vehicle RV2 then
  - HV uses RV1's location and dynamics, RV2's location and dynamics and the length of the merge to adjust the speed of the HV, such that at the end of the manoeuvre, the HV is positioned a safe following distance from RV1 and RV1 is positioned a safe following distance from RV2
  - RV1 is made aware of HV's intention to accommodate the manoeuvre; it adapts accordingly (if needed) its speed and notifies the HV for acceptance before initiating the manoeuvre

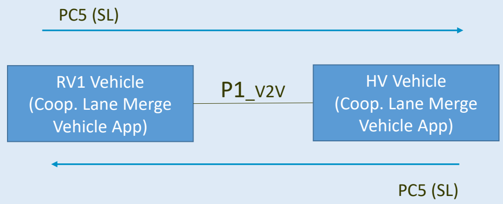
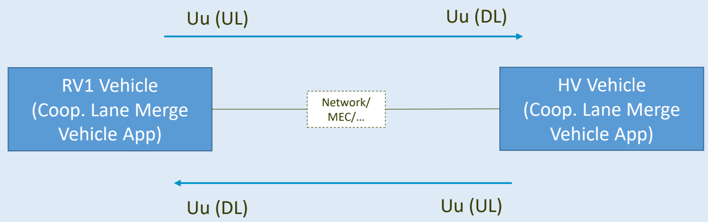
After the merge, the HV is positioned a safe following distance from RV1 and RV1 is positioned a safe following distance from any lead RV2.

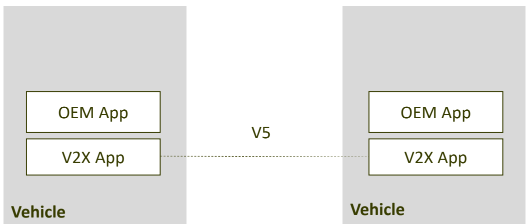
This use case is considered as a 'Short duration event-driven service' and there is no demanding data rate requirement, but the challenge of this use case is the very low latency (10ms) and service level reliability of exchanged messages.

The Cooperative Lane Merge use case has a list of potential applications reactions that could be triggered when the application receives a prediction about one or more SLRs that cannot be satisfied. Different predictions about SLR changes can trigger different types of application reactions if the SLR/QoS cannot be fulfilled.

Table 28: Potential Application Reactions of Cooperative Lane Merge Use Case

Use Case Title	**Cooperative lane merge**
User Story	A host vehicle accommodates a Remote Vehicle that is merging into the HV's traffic lane
Type of Service	Short duration event-driven service

Predicted QoS (Service Level Requirements) of Interest per Communication Link	F1 Vehicle – Vehicle:
	- F1.1: PC5 Interface, (from the RV1 to HV): latency, and reliability, data rate (low priority requirement), (unicast or broadcast)
	    ○ Intention to apply a lane-merging manoeuvre; acceptance before initiating the manoeuvre
	- F1.2: PC5 Interface (from the HV to RV1): latency, and reliability, Data Rate (low priority KPI), (unicast or broadcast)
	    ○ Intention to accommodate the manoeuvre
	
	Alternative Implementation
	F2 Vehicle – Vehicle:
	- F2.1: UL Interface of RV1 and DL Interface of HV (from the RV1 to HV latency, and reliability, data rate (low priority requirement) (unicast or broadcast)
	    ○ Intention to apply a lane-merging manoeuvre; acceptance before initiating the manoeuvre
	- F2.2: UL Interface of HV and DL Interface of RV1 (from the HV to RV1): latency, and reliability, Data Rate (low priority KPI) (unicast or broadcast)
	    ○ Intention to accommodate the manoeuvre
	

Considered Interface(s)	1. V5 
QoS Prediction Recipient and/or	QoS Prediction Recipient: • RV1 to HV Interface (F1.1 or F2.1): Remote Vehicle, RV1 • HV to RV1 Interface (F1.2 or F2.2): Host Vehicle, HV

<table>
<tr><td>Orchestrator for Adaptation?</td><td colspan="3">Orchestrator for application adaptation: Vehicle that initiates the manoeuvre</td></tr>
<tr><td>Involved Entities</td><td colspan="3">Vehicles (RV1, HV)</td></tr>
<tr><td>Reaction Scenario #</td><td>Potential Application Reactions</td><td>Entity that triggers the Application Reaction</td><td>SLR Change (e.g. Data Rate, Coverage, Latency, Reliability, Position Accuracy)</td></tr>
<tr><td>6.</td><td>The AV does not Initiate the cooperative lane merge manoeuvre and drives with its own sensor observations</td><td>Vehicle, RV1</td><td>PC5 Interface:
• Reduction of SL Data Rate (F1.1 or F1.2) below defined thresholds and/or
• Increase of Latency (F1.1 or F1.2) above defined thresholds and/or
• Reduction of Reliability (F1.1 or F1.2) below defined thresholds

Uu Interface (UL Interface of RV1 and DL Interface of HV):
• Reduction of Uu Data Rate (F2.1 or F2.2) below defined thresholds and/or
• Increase of Uu Latency (F2.1 or F2.2) above defined thresholds and/or
• Reduction of Uu Reliability (F2.1 or F2.2) below defined thresholds</td></tr>
<tr><td>7.</td><td>The AV changes driving properties e.g. speed.</td><td>Vehicle, RV1</td><td>PC5 Interface:
• Reduction of SL Data Rate (F1.1 or F1.2) below defined thresholds and/or
• Increase of Latency (F1.1 or F1.2) above defined thresholds and/or
• Reduction of Reliability (F1.1 or F1.2) below defined thresholds

Uu Interface (UL Interface of RV1 and DL Interface of HV):
• Reduction of Uu Data Rate (F2.1 or F2.2) below defined thresholds and/or
• Increase of Uu Latency (F2.1 or F2.2) above defined thresholds and/or
• Reduction of Uu Reliability (F2.1 or F2.2) below defined thresholds</td></tr>
<tr><td>8.</td><td>The AV stops safely and releases the automated driving system.</td><td>Vehicle, RV1</td><td>PC5 Interface:</td></tr>
</table>

			• Reduction of SL Data Rate (F1.1 or F1.2) below defined thresholds and/or • Increase of Latency (F1.1 or F1.2) above defined thresholds and/or • Reduction of Reliability (F1.1 or F1.2) below defined thresholds Uu Interface (UL Interface of RV1 and DL Interface of HV): • Reduction of Uu Data Rate (F2.1 or F2.2) below defined thresholds and/or • Increase of Uu Latency (F2.1 or F2.2) above defined thresholds and/or • Reduction of Uu Reliability (F2.1 or F2.2) below defined thresholds

NOTE: For the same change of SLR, reaction can be different per different vehicles.

In the event that a QoS prediction notification is sent to the AV (RV1) then the latter can decide to apply or not apply the cooperative Lane-merging manoeuvre (scenario #1). For this type of decision an estimation of the expected QoS is needed for a period of time equal to the expected duration to apply the lane-merging manoeuvre. However, the calculation of this period of time depends on the length of the ramp, its geometry, speed etc. For instance, assuming a ramp of 100 metres and a driving speed of 30km/h then 12 seconds would be needed for the vehicle to traverse the ramp and reach the highway. In that case the QoS prediction time horizon should be of 12 seconds or more.

Other forms of adaptation are also possible (as presented via scenarios #2/#3). Probably, the notification of a QoS change, before the actual initiation of this use case, could be more beneficial to decide the initiation or not of the lane merge. It may not always be easy to adapt the application during the realisation of the use case, e.g. when the notification of a QoS change arrives during the operation of this type of service. The vehicle(s) may have to enter in a safe state phase (e.g. safe stop, switch back to non-cooperative driving).

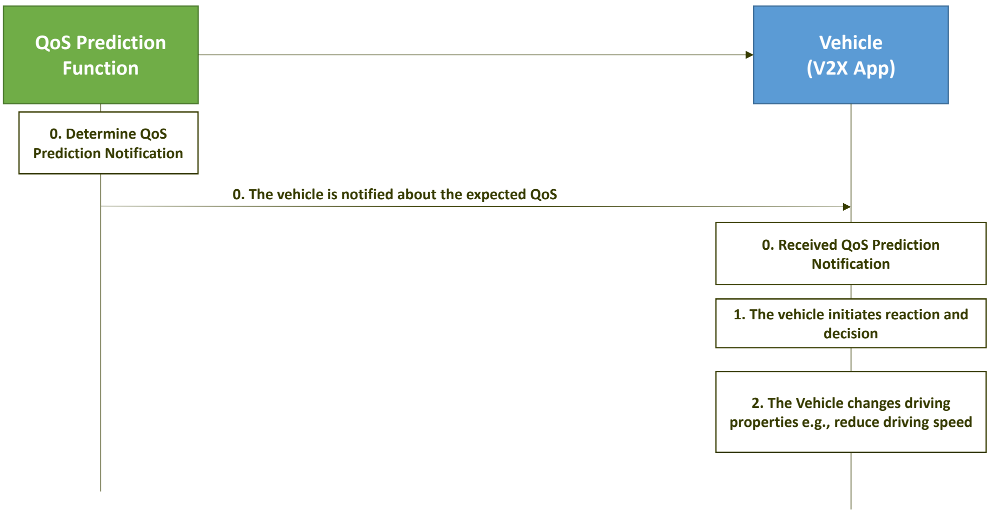


Figure 18: Example Flow Chart for Cooperative Lane Merge Reaction, Scenario #2: The AV changes driving properties, e.g. speed

Table 29: Time Breakdown Analysis for Cooperative Lane Merge Reaction, Scenario #2: The AV changes driving properties, e.g. speed

Application Reaction: Scenario #2, The AV (RV1) changes driving properties, e.g. speed		
Road Environment Type: Highway, Rural, Urban		
Vehicle Type: Passenger vehicle		
Human Intervention: No		
List Steps	**Latency (ms)**	**Explanation**
3. The Vehicle's AD app is notified about expected QoS degradation or Coverage Loss and initiates its reaction	500	Time for the vehicle to receive the P-QoS message and analyse the degradation level/threshold to determine a suitable reaction
4. The Vehicle changes driving properties, e.g. reduce driving speed	7,000	Depends on the speed of the vehicle; in the AD scenario, the assumed braking is from 50km/h to 10km/h Other factors, such as the lane the vehicle is located in, weather conditions, surface, vehicle type may affect this distance
Total Sum:	7,500	

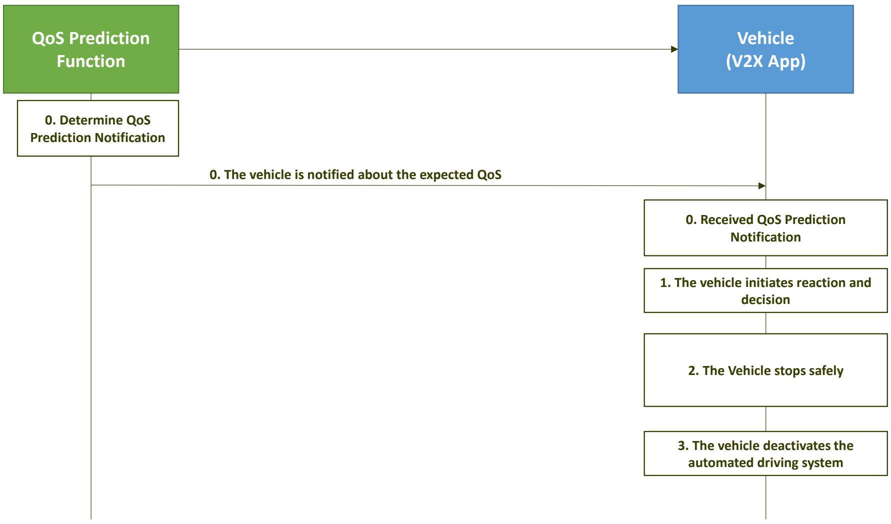


Figure 19: Example Flow Chart for Cooperative Lane Merge Reaction, Scenario #3: The AV (RV1) stops safely and releases the automated driving system

Table 30: Time Breakdown Analysis for Cooperative Lane Merge Reaction, Scenario #3: The AV (RV1) stops safely and releases the automated driving system

Application Reaction: Scenario #3, The AV (RV1) stops safely and releases the automated driving system		
Road Environment Type: Highway, Rural, Urban		
Vehicle Type: Passenger vehicle		
Human Intervention: No		
List Steps	**Latency (ms)**	**Explanation**

4. The Vehicle's AD app is notified about expected QoS degradation or Coverage Loss and initiates its reaction	500	Time for the vehicle to receive the P-QoS message and analyse the degradation level/threshold to determine a suitable reaction
5. The Vehicle stops safely	7,000	Depends on the speed of the vehicle; in the AD scenario, it was assumed the maximum speed is 50km/h and also that the vehicle can stop at the right side of the road, without searching for appropriate parking spot. Other factors, such as the lane that the vehicle is located in, weather conditions, surface, vehicle type may affect this distance If the vehicle needs to search a safety parking spot, the time needed could be longer
6. The vehicle deactivates the automated driving system	500	
Total Sum:	8,000	

Table 3 provides an overview of the estimated QoS prediction time horizons, considering the different types of application adaptations of the Cooperative Lane Merge.

Table 31: QoS Prediction Horizon of different Application Reactions of Cooperative Lane Merge Use Case

Use Case Title	Cooperative Lane Merge		
User Story	A host vehicle accommodates a Remote Vehicle that is merging into the HV's traffic lane		
Application Reaction	**QoS Prediction Time Horizon Application Reaction Completion Interval (=ARCI) (in ms)**	**Notice Period Guard Prediction interval (in ms)**	**QoS Prediction Time Horizon (in ms)**
Scenario #1 – The AV does not initiate the Cooperative Lane Merge manoeuvre and drives with its own sensor observations	12,000	3,000	15,000
Scenario #2 – The AV changes driving properties, e.g. speed.	7,500	3,000	10,500
Scenario #3 – The AV stops safely and releases the automated driving system	8,000	3,000	11,000

Table 32 provides examples of QoS Change Thresholds of the Interfaces that could trigger the QoS prediction notification. The QoS/SLR values and related thresholds for triggering prediction notification are given as examples, and the values are mainly referenced to 5GAA White Paper C-V2X UC&SLR VOL II.

Table 32: QoS Change Thresholds of QoS Predictions for Cooperative Lane Merge Use Case

Use Case Title	Cooperative Lane Merge		
User Story	A host vehicle accommodates a remote vehicle that is merging into the HV's traffic lane.		
Predicted QoS/'Service Level Requirements'	**Considered Interface(s)**	**QoS/SLR value** **(5GAA TRs)**	**QoS Change Thresholds for triggering prediction notification**
SL – Delay **or** **Uu Delay (UL and DL)**	PC5 (F1.1 and F1.2) Or Uu (F2.1 and F2.2)	20 ms	e.g. latency higher than 20 ms
SL – Reliability **or** **Uu Reliability (UL and DL)**	PC5 (F1.1 and F1.2) Or Uu (F2.1 and F2.2)	99.9%	e.g. reliability lower 99.9%

5.3 Summary of Application and System Reaction Analysis

Three use cases were selected to conduct the analysis regarding QoS prediction requirements and application reactions/adaptations.

- The ToD use case, which is a medium or long duration 'session-based' service. In this use case, QoS prediction information is needed mainly for the Uu interface especially for demanding UL (data rate) and DL (latency and reliability) requirements. QoS prediction information provided at or to the vehicle or the remote driver.
- The infrastructure-assisted environment perception use case, which is a 'periodic' service. In this use case, both Uu and PC5 interfaces are used and several entities can be involved (Vehicle, Infrastructure or Application Server). QoS prediction notifications may be needed for both unicast and broadcast interfaces.
- The Cooperative Lane Merge use case, which is a short duration 'event-driven' service. In this use case vehicles are interacting via the PC5 interface. Alternatively, the Uu interface could be used (i.e. V2N2V) with demanding latency and reliability requirements for the exchanged messages among vehicles.

Different types of application reactions were identified on the above use cases to address predicted QoS changes. Different types of V2X services have different features and consequently different types of application reactions may be needed, when a notification about an expected QoS change is provided. As it was presented in the above use cases, even in the context of the same V2X service, different options for potential application reactions are available. Of course, the type of adaptation that will be selected depends on the actual prediction quality, the road environment, the thresholds that the application developer/owner sets or even on some regulatory decisions. In all the analysed use cases, it has been shown how the QoS prediction notification enables the implementation of application reactions that can mitigate the effect of a sudden QoS change. The application should of course evaluate the confidence of the prediction when implementing application reactions as triggered by a QoS prediction notification.

For the medium/long duration 'session-based' services (e.g. ToD, High-Density Platooning, Infotainment), different application reactions could be considered and selected by the vehicle or other involved stakeholders (e.g. stop the vehicle safely and release the service, change sensor set and sensor properties and/or video configuration, give the control to a driver in car – if present). The same applies to the 'periodic' services that are always active (e.g. Intersection Movement Assist, Infrastructure-Assisted Environment Perception). In the short duration 'event-driven' services (e.g. Cooperative Lane Merging, Cooperative Driving Manoeuvre, Emergency Brake Warning), it may not be easy to adapt the application during the execution of the use case, e.g. when the notification of a QoS change arrives during the operation of this type of service. The vehicle(s) may have to enter in a safe state phase (e.g. safe stop, switch back to non-cooperative driving, change driving properties). For the short duration event-driven service the notification of a potential QoS change could be more useful when it is received by the application before the actual initiation of the service, i.e. to decide whether to initiate the service or not.

According to the above analysis, different QoS prediction time horizons were estimated for each potential application adaptation. For the three use cases analysed above the estimated QoS prediction time horizon were in a range from 6 to 18 seconds. As it is obvious, the QoS prediction time horizon of each V2X use case should be set with the worst-case scenario (or more demanding, in terms of time required, application adaptation). For instance,

- the ToD service should set a QoS prediction time horizon of 17.5 seconds (i.e. due to scenario #6: change sensor set and sensor properties and/or video configuration),
- the Infrastructure-Assisted Environment Perception service should set a QoS prediction time horizon of 14 seconds (i.e. due to scenario #4: the AV hands over the control to the driver in car and releases the automated driving system),
- the Cooperative Lane Merge service should set a QoS prediction time horizon of 15 seconds (i.e. due to scenario #1: the AV does not initiate the cooperative lane merge manoeuvre and drives with its own sensor observations).

An important factor is identifying whether a 'machine' or a human is involved in the reaction after a QoS prediction notification and whether that human reaction can be simple and fast ('hit the brakes!') or complex and slow ('resume driving in a complex situation'). Many of the considered use cases could be implemented with or without the involvement of a human. In addition, factors such as speed, acceleration, vehicle density, road topology, road environment and assumptions on human reactions affect the application ARCI and consequently the QoS prediction time horizon.

The QoS change thresholds that could trigger the QoS prediction notification or even each adaptation are implementation specific. For instance, the developer or operator of an application can specify the thresholds for degradation (or improvement) of QoS parameters that are acceptable for the efficient, smooth and safe operation of a service. It should be noted that different implementations of a specific use case which utilise QoS prediction features like ToD can consider different threshold values. Hence, generic threshold values cannot be defined that apply to different implementations or deployments. Also, alternative QoS levels could be considered for a specific use case, if needed. An alternative option is that a prediction notification is not based on specific QoS change thresholds that will trigger the transmission (or not) of a notification. But the application subscribes to receive periodic notifications (e.g. every 5 seconds) and the application (or the vehicle) can decide whether to adapt or not, according to the information received in the QoS prediction notification.

Meanwhile, it is also necessary to consider the case that the actual QoS change may not occur as predicted since the confidence level is not (and does not always need to be) 100%. Therefore, when the application layer prepares certain reaction(s) upon receiving the QoS prediction notification, it may also be necessary to re-evaluate whether the conditions for the reaction(s) are still valid before carrying out these reactions. This can be realised by continuously monitoring the network QoS during the adaption procedures. This is related to certain V2X use cases and should be part of V2X service adaptation implementation. There can be two cases, i.e. normal and abnormal ones. The normal case refers to QoS changes consistent with predictions. In such case, the vehicle or other V2X entities receive the QoS prediction notification and then trigger the reaction. All the reaction flows will be completed before the QoS prediction takes place or becomes effective. The abnormal case is when the reaction is different compared with the previous QoS prediction, i.e. the previous use case reaction was no longer suitable. The vehicle or other related entities need to discontinue previous reaction flows and take actions based on the actual QoS change (i.e. stopping the planned reaction and taking other measures). It should be noted that the two cases above are just illustrations and there could be other ways of handling, for example, to extend QoS prediction delays. The QoS prediction time range may need to be extended if there is no incoming new notification. The vehicle or the system may also record the 'state change' histories if the application needs to go back to a previous state for the sake of robustness.

6 System Enhancements and Requirements

6.1 Architecture and Interfaces

6.1.1 Application Deployment

The impact of ToD Deployment Architecture on QoS Prediction is discussed below. As show in Figure 20 cited from ToD TR[10], the interface between ToD Vehicle Application and ToD Application Server consists of three parts (Uu

[10] *5GAA_A_200145_ToD_D2_System_Requirements_Architecture_v1.02_For publication*

interface, N3 interface, N6 interface). To enable e2e QoS prediction for a ToD service, all the three aspects need to be taken into account. The QoS of Uu and N3 interfaces can be guaranteed or predicted by 5G QoS mechanisms, while QoS of N6 (UPF to ToD Application Server) is outside the 5G system and thus cannot utilise 5G QoS mechanism and is affected by different deployment options. Here, 5G QoS mechanisms refers to 5QI based QoS mechanism over Uu interface and GTP-u tunnel based QoS mechanism within the 5G core network. It is needed to consider how the different deployment architectures can influence QoS prediction. In the following, three main deployment options have been analysed:

- Option 1: ToD Application Server deployed inside MNO network
- Option 2: ToD Application Server deployed outside MNO network
- Option 3: ToD Application Server deployed inside MNO network in MEC infrastructure, multiple UPFs applied

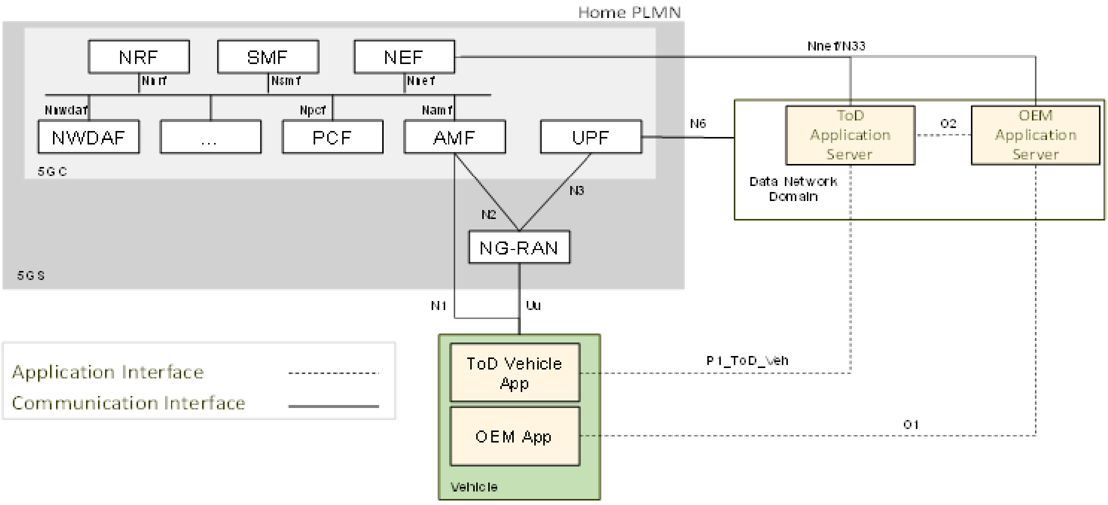


Figure 20: 5G System Architecture

Option 1: ToD Application Server deployed inside MNO network

This is a relatively simple case, and all the ToD system components (ToD Vehicle Application, ToD Application Server, ToD Operator Application) are deployed in the realm of an MNO network. The link between ToD Vehicle Application and ToD Application Server can receive QoS guarantees via MNO's 5G network and the end-to-end (E2E) QoS on the Uu, N3, N6 can be theoretically predicted using MNO mechanisms defined in the 3GPP System (see reference [4] cl. 6.9), even though those mechanisms are typically limited to the network segment between the UE and the UPF node. Such assumption is based on the fact that the potential QoS issues on the N3 and N6 interfaces and on the data network are considered neglectable.

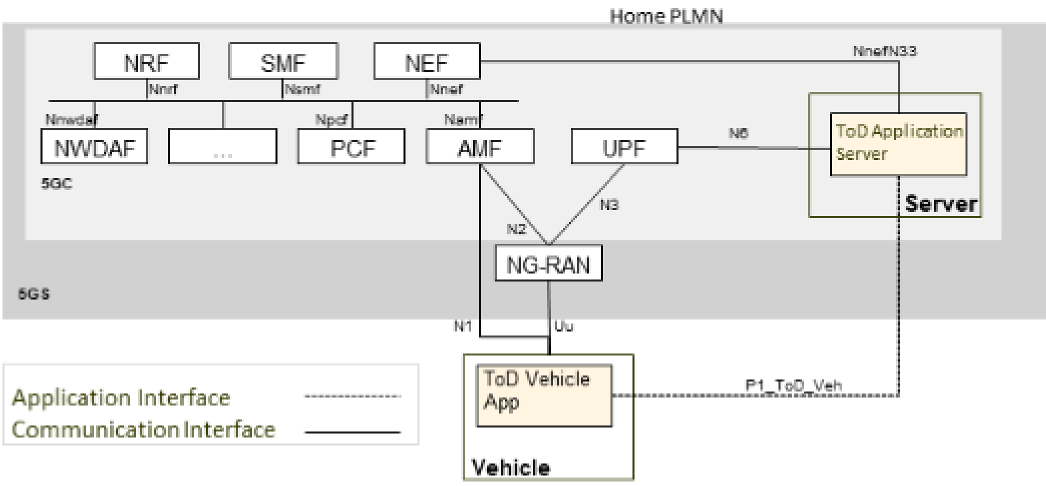


Figure 21: Architecture of ToD deployed with Option 1

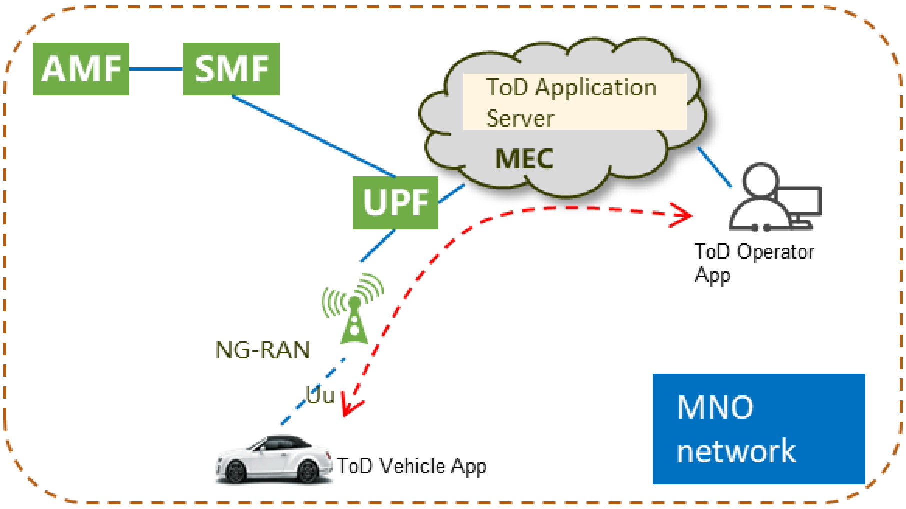


Figure 22: Example deployment of Option 1

Option 2: ToD Application Server deployed outside MNO network

The ToD Application Server is deployed in a public cloud and outside the domain of an MNO. In this case the QoS inside the MNO (Uu, N3 interfaces) could be guaranteed by 5G QoS mechanisms, but the path between the MNO's UPF and the ToD server may not be managed by the MNO's QoS management procedures. Similarly, the QoS prediction function of the MNO can foresee potential QoS changes related to Uu and N3 interfaces, but not the path from the UPF to the ToD Application Server.

NOTE: Data network analytics, as defined in cl. 6.14 of reference [4] may be used to predict potential issues, however the format of such analytics is not consistent with the format of QoS prediction provided by cl. 6.9 of the same reference. Therefore it will be difficult for the application to combine such different information to retrieve an end-to-end view of the QoS including N6.

There may be multiple routers/switches along the link and the network performance may be affected by various factors outside MNO control. Service level agreements are needed for this option to satisfy the E2E QoS requirement of the ToD service. Alternatively, the application may need to use an end-to-end QoS prediction service also covering N6 and the network segment including the ToD Application Server.

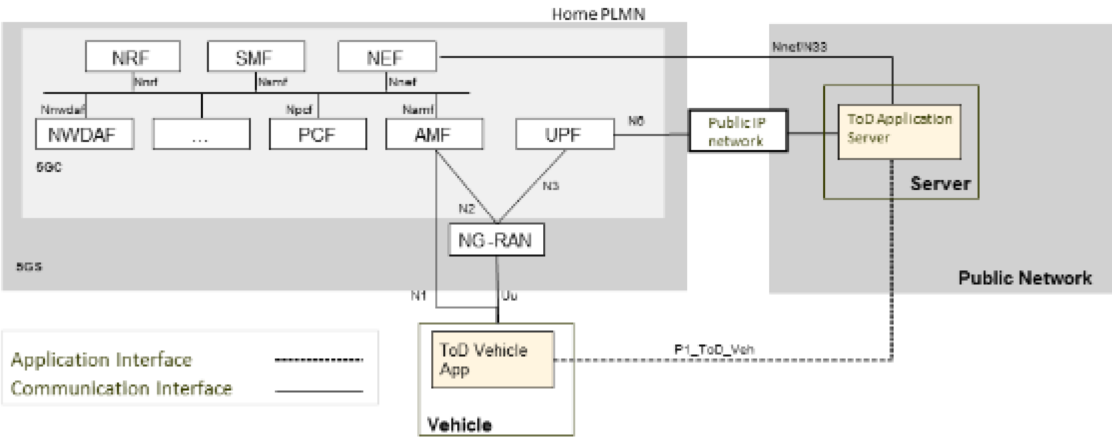


Figure 23: Architecture of ToD deployed with Option 2

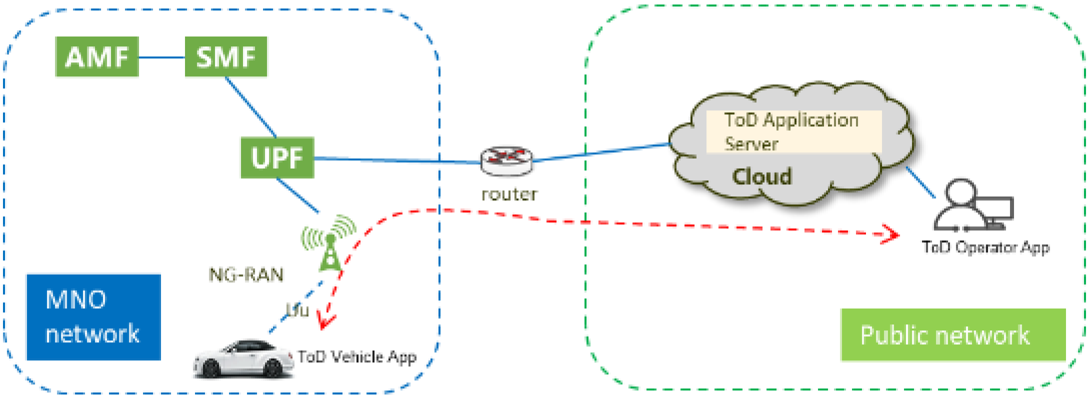


Figure 24: Example deployment of Option 2

Option 3: ToD Application Server deployed inside MNO and multiple MECs applied

When the ToD Application Server is deployed within MEC infrastructure and the remotely driven vehicle moves across different MEC infrastructure sites, multiple UPFs may be involved in the ToD service operation. The local UPF may change due to vehicle migration, but the PDU session anchor (PSA) UPF can remain the same. To enable the ToD system's QoS prediction mechanism, SLAs between the UPF and the ToD Server should consider the hierarchical topology of UPFs.

To generate accurate QoS predictions as the serving UPF is changing, the prediction function (e.g. NWDAF) may need to combine information from the serving UPF and the MEC hosting environment where the Application Server is located. The overall E2E network performance may demonstrate considerable variation if the UPF node selected to serve the PDU session varies.

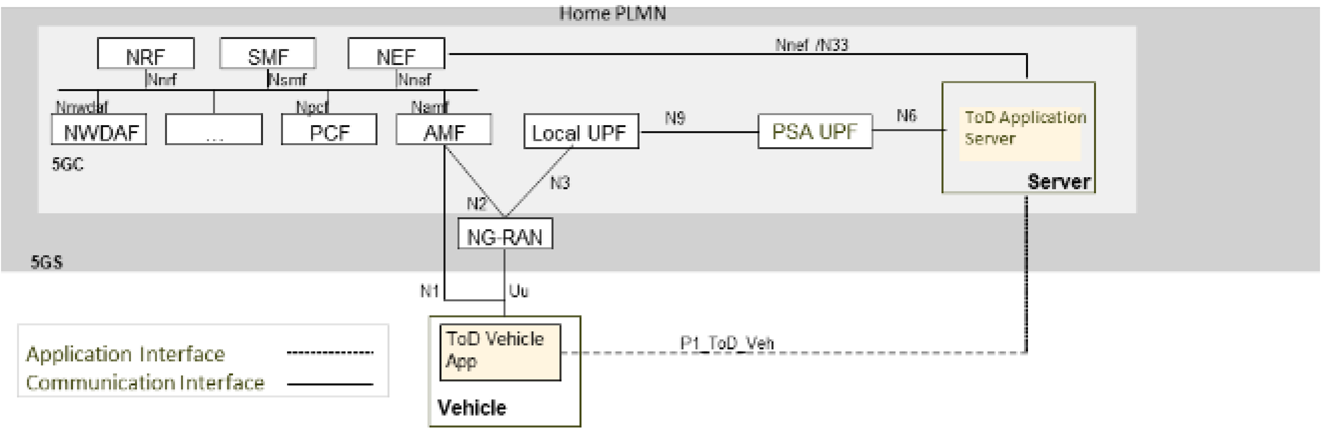


Figure 25: Architecture of ToD deployed with Option 3

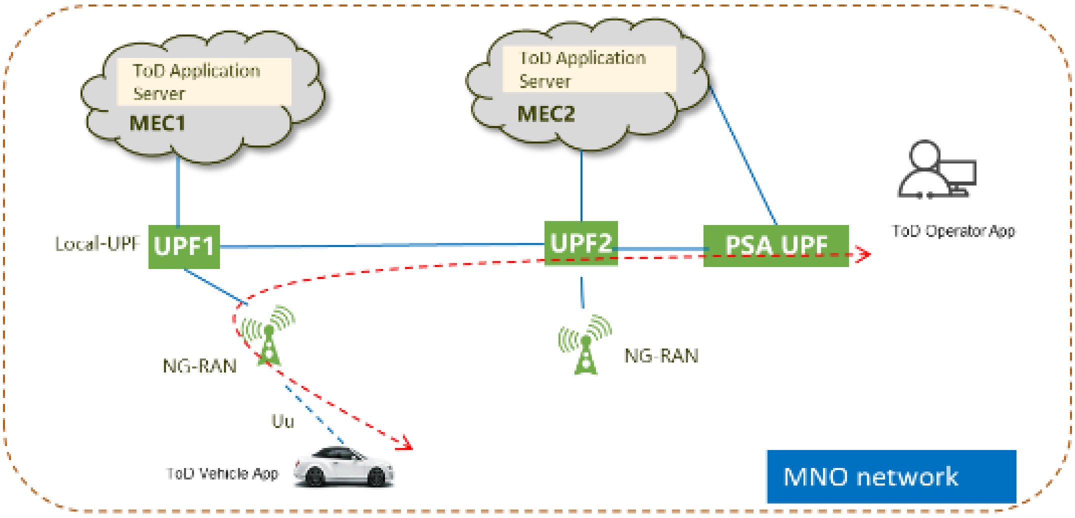


Figure 26: Example deployment of Option 3

In order to support QoS prediction in all three identified scenarios, it is necessary to consider both the 5G QoS prediction mechanisms deployed within the MNO domain and also the QoS prediction mechanisms for network segments outside the 5GS. In most cases, such network segments could be handled as wireline. Option 1 and Option 3 can rely on QoS prediction as provided by the 3GPP standard. However, in comparison, Option 2 still faces some challenges due to lack of standardisation for QoS prediction related to the network segment between the UPF and third-party Application Server.

There have been various QoS mechanisms over IP networks that can be utilised to guarantee QoS, but working out how to support these activities taking place outside the MNO is an open issue. In real deployment situations enabling QoS support and prediction, it is possible that the MNO and third party can make an SLA to reserve sufficient network resources. Meanwhile, there needs to be some measurement and collection of network entities like routers and switches along the path, and prediction algorithms are needed to provide predicted KPIs such as data rate, latency, jitter, etc. Although QoS provisioning over a wireline network is easier than wireless, in order to achieve support for end-to-end QoS prediction, this open issue also needs to be considered. This is an important aspect before rolling out QoS prediction solutions in smart transport and automated driving industries.

6.1.2 QoS Prediction in Edge Computing Deployments

Available MEC APIs for V2X [11] may help V2X applications by providing additional QoS prediction services.

Mobile Edge Computing or MEC is a key enabler of several C-V2X applications that require ultra-low latency and high reliability. In general, MEC supports those applications needing to process large amounts of data which could benefit from the use of MEC for near-vehicle processing, instead of uploading the data to the cloud, causing additional round-trip delays. The use cases analysed by [9] are some examples that can benefit from edge computing deployments.

Some specific enhancements that could be addressed to existing ETSI MEC APIs and that relate to QoS prediction [11] are currently being studied by ETSI MEC ISG [14].

6.1.2.1 Multi-domain (Inter-MNO and Inter-OEM) Scenarios

5GAA MEC4AUTO studied different deployment options of MEC infrastructure, where several MNOs and OEMs may share part of the MEC resources such as the MEC platform or the MEC application. Those scenarios are summarised in the MEC4AUTO Task 2 TR [10]. It is relevant for automotive applications to study those scenarios for the purpose of identifying how network analytics (such as QoS prediction) may be produced and consumed by the V2X applications, considering the specifics of edge deployments. The study has considered the following main cases:

1. Scenario 1 – Both MNO A and MNO B have MEC platform and MEC application X.
2. Scenario 2 – Both MNO A and MNO B have MEC platform, but MEC application X is available only in MNO A.
3. Scenario 3 – Only MNO A has MEC platform and MEC application X is available only in MNO A.
4. Scenario 4 – Analytics on the user plane link up with a third-party.

In the above four scenarios it is possible to make the following observations:

- The application instances deployed in the vehicle may implement V2N2V or V2N2I modes of communication by traversing multiple network domains, usually belonging to a specific MNO, OEM or third party. This means that the application endpoints involved in the communication (e.g. in the vehicle or road infrastructure) will likely establish an end-to-end communication link that traverses networks belonging to different organisations. The QoS of the E2E communication link depends on the QoS achievable on each of the involved network segments.
- Each of those network domains may support the generation of network analytics, such as QoS prediction, within the single network domain. However, there is currently no standard network analytics service capable of supporting the generation of network analytics for the whole E2E user plane link through multiple domains.
- As the network domains belong to different organisations (OEMs, MNOs, RTAs, etc.) it may be difficult to consider that those entities will share the network analytics externally.
- 3GPP has yet to provide a mechanism for issuing analytics across multiple operator networks.
- ETSI MEC V2X Information Services (VIS) API [11] could address this issue by providing a consolidated view of QoS prediction across multiple domains, such as several operator networks.

6.1.2.2 VIS API to Support a Cooperative Framework for QoS Prediction in Multi-domain Scenarios

When it comes to edge deployments, ETSI MEC GS 030 [11] already provides a service for journey-specific QoS predictions. The V2X Information Service API (VIS) includes an existing service where a V2X application – either a V2X AS deployed in the cloud or in the edge-hosting environment, or an application instance deployed in the UE and accessing MEC services via a MEC application – can utilise the VIS services typically deployed in a MEC platform. The QoS prediction provided by VIS may refer to a specific journey (e.g. from one location to another) which may be specified via one or more waypoints connecting the point of 'origin' from the 'destination'. When requesting QoS prediction, the V2X application may also specify the time when the vehicle can be at each waypoint, making it possible to request a prediction for a desired time in the future, either for one location or for a set of locations in the journey. The VIS QoS prediction service in this case configures as an alternative QoS prediction service to the one provided by the 3GPP System, which is implemented by the NWDAF, usually accessible via the Network Exposure Function or NEF. In order to understand how the ETSI MEC VIS and the 3GPP NWDAF prediction services can be accessed by a V2X application, it is possible to consider the Synergised Mobile Edge Cloud architecture supported by 3GPP and ETSI ISG MEC, which is summarised in the Figure 27.

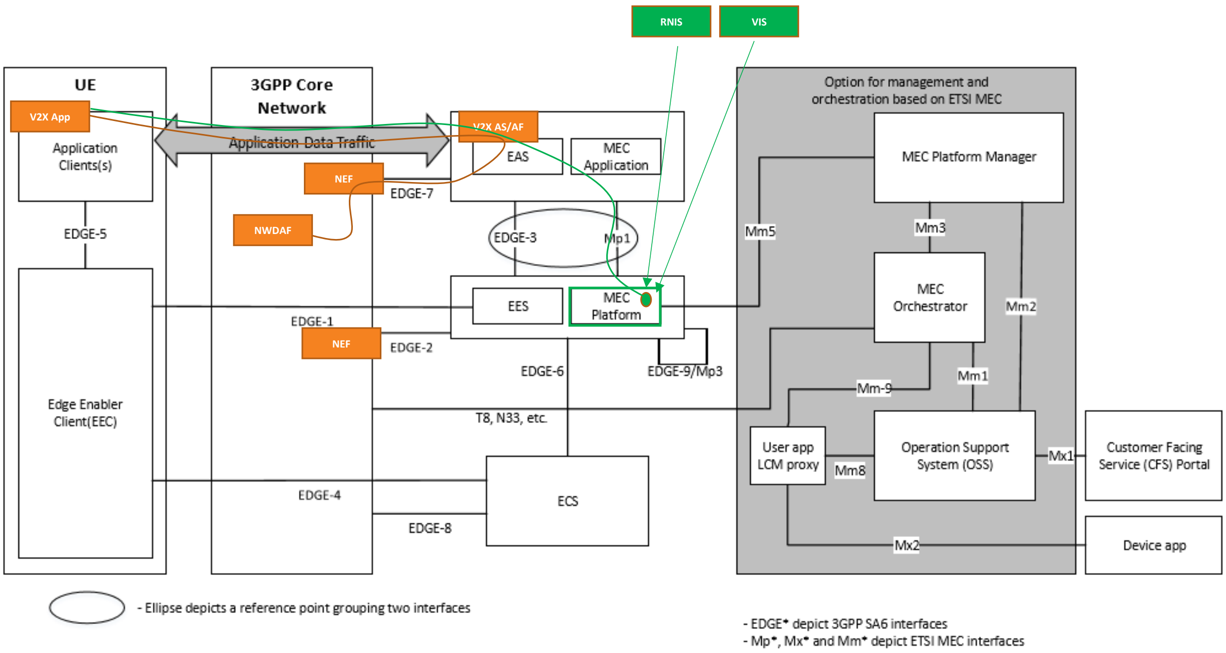


Figure 27: Synergised Mobile Edge Cloud architecture, supported by 3GPP and ETSI ISG MEC

As it can be seen in the figure, a V2X Application may receive QoS prediction information via the 'orange' route to the 3GPP QoS prediction service and the 'green' route to the ETSI ISG MEC service, ETSI GS MEC 030 (V2X Information Services API).

Depending on the specific MEC deployment scenario, the VIS service and the MEC platform where it resides could be hosted by one of the following:

1. The MNO serving the UE where the V2X application is deployed.
2. Another MNO (e.g. the MNO serving the other UE in the V2N2V case) as in one of the MEC resource sharing deployment scenarios [10].
3. In a third-party hosting environment, including a shared hosting scenario.

VIS may be used to implement a cooperative QoS prediction service where information from multiple sources/inputs can be used by the API to simplify the process of delivering information to the interested V2X application instances residing in different domains (MNOs, OEMs, etc.).

One interesting aspect is that VIS may complement the 3GPP QoS prediction service for the scenarios in which the following conditions apply:

1. The application needs to retrieve the QoS prediction of the **E2E user plane link between two application end-points**, and not just a portion of the link. In the V2N2V, the two endpoints are two vehicles, while in the V2N2I the endpoints are one vehicle and the road infrastructure.
2. The end-to-end user plane link is traversing several network domains which refer to different entities. For example **multiple MNO networks or multiple OEM networks**.

As an example of the two conditions described above it is possible to consider the use case Cooperative Lane Merge (V2N2V) with the referenced MEC Scenario 1 [10], as represented in the figure below. According to [10], in such a MEC scenario each vehicle may exchange V2X messages with the other vehicles via the Uu interface and MEC infrastructure. Each vehicle has a different MNO providing network services. Each MNO has deployed a MEC platform and MEC application in its own network. Each MNO can also provide QoS prediction for its own 3GPP network, according to [3] and [4]. In this use case, the blue vehicle (HV) needs to make decisions based on the QoS prediction of the end-to-end link between the HV and the red vehicle RV1.

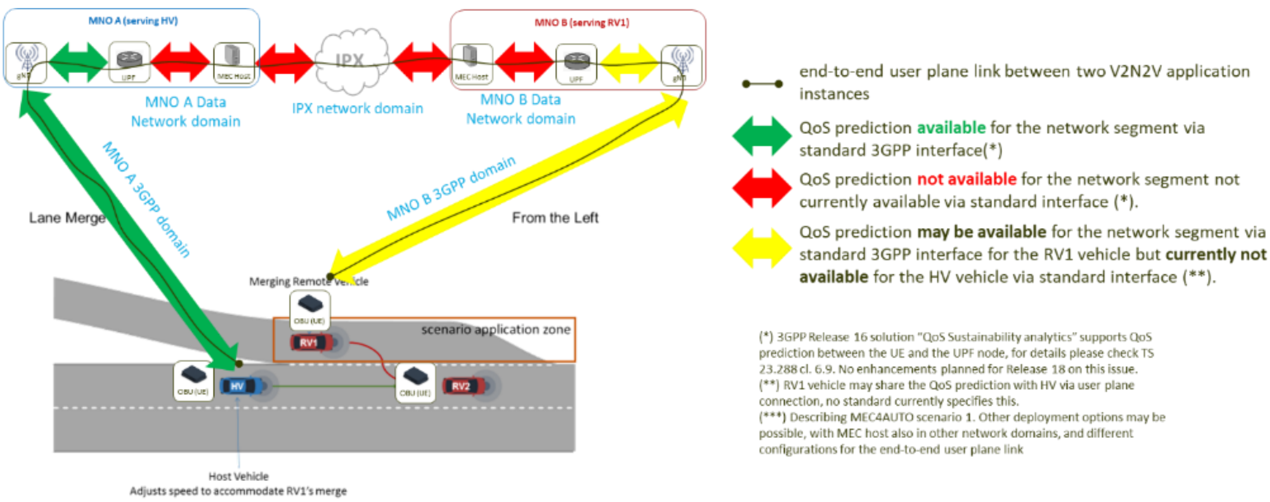


Figure 28: Using QoS prediction for the end-to-end user plane link between two application instances (case V2N2V) in Cooperative Lane Merge, assuming the deployment is according to MEC4AUTO Scenario 1

As it can be seen in the picture, the HV may retrieve QoS predictions for the network segment shown in green (MNO A 3GPP domain). It may retrieve QoS predictions for the network segment shown in yellow (MNO B 3GPP domain) if such information is provided by the MNO B to the application instance serving the RV1 vehicle, as long as such application instances share the information to the HV. At the time of writing there is no standard procedure for such sharing to happen. Also the network segments between the UPF of the MNO A and the UPF of MNO B (shown in red) are not covered by any standard QoS prediction provider, even if it is possible to envisage that there could be entities supporting those network segments at least partially. This means specifically that:

- 'DN Performance' analytics described in cl. 6.14 of [3] could provide information related to the QoS prediction of the data network hosting the MEC host of MNO A, in the event it implements such a service.
- 'DN Performance' analytics described in cl. 6.14 of [3] could provide information related to the QoS prediction of the data network where the MEC host of MNO B is located, in the event it implements such service. It is possible to assume that such information may be available to an Application Server interconnected to the MNO

B core network and serving RV1 vehicle, and that such information could be shared with the HV, even if there is no current standard procedure for the sharing of such information.
- An OTT Link Performance Prediction (LPP) or equivalent non-standard QoS prediction service could provide information related to the QoS prediction of the IPX network domain interconnecting the two operators, again only if the related network domain implements such a service.

VIS with such enhancements could support QoS prediction of the end-to-end link shown in the Figure 28 through cooperating acquisition of the QoS prediction from the multiple network domains involved, provided that some specific enhancements are performed to VIS specifications.

6.2 Open Issues

Requirements or areas of improvement for QoS prediction have been introduced in the context of NESQO and eNESQO WIs [12]. An overview of those requirements is provided below:

- **Area of improvement 1:** Delivery of potential QoS change notifications to the vehicle (UE-side) for application adaptation.
- **Area of improvement 2:** Support prediction for the following end-to-end KPIs: latency for GBR or Non-GBR QoS Flows, Packet Delivery Ratio for GBR or Non-GBR QoS Flows, Uplink Throughput for GBR QoS Flow and Non-GBR QoS Flows Downlink Throughput for GBR QoS Flow and Non-GBR QoS Flows, coverage and capability.
- **Area of improvement 3:** Support the prediction of the actual KPI value (either average or median or the CDF) as an alternative or in addition to the value range for the KPI, according to a set of thresholds in the QoS prediction notification.
- **Area of improvement 4:** Additional input data, such as from the application layer (vehicle/server), RAN, CN NF and third parties, could also be used for generating prediction notifications, and investigated in order to increase the quality of the prediction.
- **Area of improvement 5:** Finer granularity in determining the group of UEs, PDU Sessions and QoS Flows for which the prediction is applicable could be achieved if further dimensions can be added to the data collected, potentially also exploiting additional input data sources.
- **Area of improvement 6:** Provide support for an NF consumer to specify a Notice Period, or a time by when the prediction notification is to be received by the NF consumer in relation to the time when the potential QoS change event is predicted.

Taking into consideration the above areas of improvement and other requirements identified, the PRESA WI has proposed requirements to 3GPP SA2 to be considered in the Release 18 work. An overview of the requirements and the current status in 3GPP standardisation body is provided in the table below, Table 33 together with the PRESA evaluation on current Release 18 status, as explained below.

Table 33: QoS Prediction Requirements and Standardisation Status

Requirement No.	Requirements of QoS Prediction	3GPP Status	PRESA Evaluation on Current Release 18 Status
Req.1	Increase granularity of QoS predictions (e.g. below cell level, more KPI reporting, richer information, aware of the specific UE context) also by means of more input data to the analytics enabler (e.g. from core network, RAN, etc.)	Under discussion in SA2 SI 'Study on Enablers for Network Automation for 5G – phase 3', via the following Key Issue: Key Issue #7: Enhancements on QoS Sustainability analytics	High
Req. 2	Improve freshness of QoS predictions (e.g. notice period)	Could be addressed via the above Key Issue #7	High
Req. 3	Enable QoS predictions also of the PC5/Sidelink Interface	Not considered in Rel. 18	Not Applicable

Req. 4	Investigate QoS prediction in Multi-MNO/Cross-border environments	Under discussion in SA2 SI 'Study on Enablers for Network Automation for 5G – phase 3', via the following Key Issue: Key Issue #3: Data and analytics exchange in roaming case (The above task is applicable to all analytics ID, hence including QoS prediction)	Low
Req. 5	Investigate the need for coordinating notifications of QoS prediction among V2X Servers (e.g. ordering, prioritisation)	Not considered in Rel. 18	Not Applicable
Req. 6	Investigate harmonisation of QoS change notifications for V2X application (e.g. QoS change notification sent to the UE via NAS/RRC)	Under discussion in SA2 'Study on 5G System Support for AI/ML-based Services' to explore the possible architectural and functional extensions to support the Application layer AI/ML operations defined in TS 22.261	Low

3GPP Release 18 work aiming at enhancing QoS prediction functionality is ongoing [13] and proper evaluations can only be done after such work and the following normative phase is completed. However, some initial conclusions have been reached in [13] and therefore it is possible – with a degree of approximation – to project what potential improvements could be seen with new solutions in Release 18, as described in the table above. In this context, it is important for 5GAA to evaluate the current outlook for QoS prediction functionality in Release 18 in light of the areas of improvements and the requirements that have been identified and described above. Such an evaluation should be done according to the following scale:

- **High** – most of the 5GAA requirements are expected to be fulfilled, or could be potentially fulfilled with the improvements planned for Release 18.
- **Medium-High** – some of the 5GAA requirements are expected to be fulfilled, or could be potentially fulfilled with the improvements planned for Release 18. The service is improved, but some issues may need to be solved in later releases.
- **Medium** – some improvements have been delivered, but important work is required in future 3GPP releases.
- **Low** – little or no improvement is observed in the current 3GPP release. The related aspects will need to be solved in future 3GPP releases.

In the scope or **Requirement 1**, the current status can be analysed by firstly evaluating the related area of improvement under 5, and then areas of improvement expected under 2, 3 and 4, as follows:

- Area of improvement 5, aiming at *increasing the granularity of the prediction* (e.g. sub-cell level, for areas smaller than a cell and aware of the specific UE context) also by means of more input data:

  The evaluation about the current outlook/potential outcome of current Release 18 work is evaluated as **High**.

  The solutions that have been selected in [13] support further granularity of the QoS prediction for an area of interest or a UE path that is smaller than a cell. Such improvement has been possible since the NWDAF can now determine – when calculating the analytics – the list of the UEs in the area of interest or required UE path. Once such list is determined, the NWDAF can initiate collection of additional data from the 5G core in order to retrieve the actual QoS for those UEs. Such data can later be used to infer QoS predictions. The selection of input data only for the relevant UEs can provide more accurate analytics. Moreover, the proposed solutions can collect information on the UE context and subscription (e.g. serving UPF, type of UE, UE speed, etc.) in order to achieve even greater accuracy. Acquiring more information on the UE from the GSMA database is needed in order to filter collected data in a more meaningful way and support analytics filters that can be specific for UE models. This is because it is not possible to assume that actual QoS is independent from the type of UE and its characteristics. Subscription is also very important in terms of the final impact of the QoS as MNOs may use policy mechanisms to provide different type of services and build different business models.

Some of the aspects may require further study, for example the following:

- The UE list needs to be constantly updated as new UEs entering the area of interest or UE path need to be considered in the measurements, while UEs that are leaving the area need to be excluded from the data used to calculate the analytics.
- How to collect the UE location and speed has not been decided yet. UE position in the network may not be accurate, or not updated/not in sync with the QoS measurements collected in the network. This may cause potential inaccuracies in determining the UE list and, in turn, the analytics.
- The additional input data could be collected either for all relevant UEs in the area of interest, or the UE path where a highly accurate QoS prediction is required, or only for the UEs in the area of interest/UE path experiencing QoS failures. While in the second case the advantage is to limit the impact on signalling, in the first case the 3GPP System can make predictions with higher accuracy when data is collected for a specific UE type. In addition, the first option may support higher accuracy prediction as well as more KPIs for non-GBR QoS flows.
- The additional input data is only collected after the application has requested/subscribed to the analytics. This means that the Application Server may need to request/subscribe to the analytics weeks or months before accurate predictions can be produced. The UEs in the area of interest/UE path when the application/request subscription is received by the NWDAF may not be the ones relevant to the QoS prediction in question. Alternatively, the MNO will need to use other mechanisms (e.g. non-standard) to collect input data prior to application request/subscription and pre-load it into the NWDAF via some proprietary interface.
- The generation of more accurate prediction depends on the application subscription/request. Changes in the area of interest (e.g. new lane is built for the road, a new road segment/parking lot is built) will need to re-trigger collection of additional data and take some time before highly accurate QoS prediction is again supported for that area of interest/UE path.

- Area of improvement 2, aiming at *introducing support of additional QoS KPIs for GBR and optionally for non-GBR QoS flows*:

  The evaluation about the current outlook/potential outcome of current Release 18 work is **High**.

  Solutions currently selected for normative use have the potential to cover such requirements with additional data to be retrieved from the UPF, SMF, AMF and UDM. As a reminder, the 3GPP scope is limited to the segment between the UE and the UPF, therefore metrics such as the E2E throughput cannot be supported if end-to-end refers to network segments beyond the UPF (e.g. N6 interface and data network, where the Application Server is usually located). At the same time, 3GPP is not expected to support QoS KPIs other than what is supported in the 5G QoS model. Because of the possibility to limit collection of input data only for events of QoS failures it is expected that the Release 18 solution can provide better support for QoS flows for GBR rather than for non-GBR flows.

- Area of improvement 3, aiming at providing *more details in the potential QoS change notification, with respect to the existing threshold-based information* (range in which the value of the QoS KPI is predicted):

  This aspect has not been covered in Release 18. However, PRESA analysis in Section 5 has demonstrated that the current threshold-based mechanism is sufficient for the applications to adapt. This is true at least for the use cases already analysed so far and no use case that requires a different output mechanism has been identified as yet.

  As current support status seems sufficient, even if the specific requirement is not fulfilled the evaluation of the current outlook/potential outcome of the current Release 18 work is considered **High**.

- Area of improvement 4, aiming at adding *additional input data available for QoS prediction* to the analytics enabler.

  The evaluation about the current outlook/potential outcome of the current Release 18 work is **Medium**.

  While additional data from the Core Network has been made available to the analytics enabler (e.g. input data collected from the UDM, AMF, SMF, LMF/GMLC, UPF, PCF) and from the OAM, there is no possibility for the NWDAF to source data collection directly from the RAN. It is not possible at the moment to estimate the impact of such a limitation. Another issue is the inability of the NWDAF to collect input data for finer granular prediction before the Application Server has subscribed to the analytics. The effects of such a limitation have been described above.

Taking into account the evaluation of area of improvement 1 (High), area of improvement 2 (High), area of improvement 3 (High) and area of improvement 4 (Medium), the overall evaluation for Requirement 1 can be considered as **High**.

In the scope or **Requirement 2** and former Area of Improvement 6, aiming at *improving the freshness of QoS predictions (e.g. notice period)* this can be considered covered by the parameter 'time when analytics information is needed', which is an optional input as part of the Nnwdaf_AnalyticsInfo_Request or Nnwdaf_AnalyticsSubscription_Subscribe service operation and supported since Release 16.

In normal circumstances, the Application Server will subscribe to QoS prediction for specific locations/areas corresponding to road segments, and for a specific time interval (analytics target period), so it is expected that the Application Server will set the parameter 'time when analytics information is needed' early enough before the analytics target period starts, and depending on the QoS prediction time horizon that has been estimated in the use case analysis. Implementation aspects may impact how early the QoS prediction may be produced and delivered.

According to [16] the "Time when analytics information is needed" or "periodic reporting", should not be set to a value less than the "Supported Analytics Delay per Analytics ID". The determination of "Supported Analytics Delay per Analytics ID" is NWDAF implementation specific. 3GPP SA2 work in Rel-18 does not guarantee that the "Supported Analytics Delay per Analytics ID" is less than 6-18 seconds. However, 3GPP SA2 confirms that "Time when the analytics are needed" + "notice period" should be set to a time in the future which is earlier or equal to the start time of the "Analytics target period" in the future.

From a standardisation perspective, the evaluation about the current outlook/potential outcome of current Release 18 work is considered as **High**.

In the scope of **Requirement 3**, aiming at *enabling QoS predictions also of the PC5/Sidelink Interface,* since it is not in the scope of Release 18, the evaluation is considered as **Not Applicable**.

In the scope of **Requirement 4**, aiming at *providing support for QoS prediction in Multi-MNO/Cross-border environments,* several solutions have been proposed but several concerns still exist and it is too early to understand whether a conclusion will be reached. Resistance from the MNO side to share analytics with other third parties may also limit implementation of such solutions. Current solutions are considered generic for analytics exposure, not specifically for QoS prediction or QoS sustainability analytics.

According to those considerations, the evaluation about the current outlook/potential outcome of current Release 18 work is considered as **Low**.

In the scope of **Requirement 5**, aiming at investigating the need for coordinating notifications of QoS prediction among V2X Servers (e.g. ordering, prioritisation)), since it is not in the scope of Release 18, the evaluation is considered as **Not Applicable.**

In the scope or **Requirement 6** and former area of improvement 1, aiming at supporting delivery of potential QoS change notifications to the vehicle for UE-side application adaptation, while no definitive solution has been found to this issue, 3GPP Release 18 study [15] has evaluated solutions to share analytics with the UE for the purpose of AI/ML application adaptation, independently from V2X use cases. None of the solutions of the above study have been selected for normative in Release 18. It is expected that 3GPP will focus on this topic during Release 19.

As already discussed in [12], solutions that can deliver information on potential QoS change to the UE-side of the application are needed especially for the time horizons in the order of a few seconds and become particularly important under critical radio conditions, since the user plane may not always be available, while control plane radio resources are expected to be available in most cases even when user plane resources have been deallocated. As described in [12] , the delivery of potential QoS change notification using a similar mechanism as the existing QoS change notification (e.g. via NAS) may bring advantages in terms of UE-side adaptation for short time horizons. Moreover, a solution that can harmonise the delivery of potential QoS change notification and existing 'actual' QoS change notification (e.g. QoS notification control) that are delivered when the QoS is actually changed may provide further benefits. For example, it can enable the application correlating two notifications for the same QoS change event (potential and 'actual') and facilitate implementation of the adaptation. The application may decide to implement the adaptation in two steps, anticipating some actions when the QoS change is only predicted and completing the adaptation later when the actual QoS change happens. Other benefits relate to the possibility of the application evaluating the reliability of the QoS prediction service, e.g. to understand if predicted QoS changes are in fact happening, and measure potential deviations. Another open problem concerning the aspect of analytics delivered to the UE is whether the UE should subscribe directly for those analytics or whether the Application Server in the cloud should subscribe on behalf of the UE. In general, Application Server-based subscriptions (with respect to UE triggered subscriptions) avoid too much impact on signalling and avoid the risk of congestion in the control plane radio resources.

Based on these arguments, the evaluation about the current outlook/potential outcome of current Release 18 work is considered **Low**. Therefore this remain an important issue to be solved in Release 19.

In the PRESA WI based on analysis conducted in the previous sections additional areas of improvement have been defined, as summarised below:

- **Area of improvement 7** – The network needs to be capable of supporting end-to-end QoS prediction in V2N2V or V2N2I modes of communication, where multiple network domains are traversed by the user plane link between two V2X application instances.

  This should not be confused with Requirement 4, since it has a broader scope. While Requirement 4 relates to the support of analytics across several MNO networks, but still within 3GPP scope (e.g. QoS prediction of the network segment between UE and UPF) improvement 7 is aimed at supporting QoS prediction of the whole user plane link (i.e. beyond the network segment between UE and UPF), while considering other network domains such as data networks, MEC platforms and MEC applications, IP interconnections between operators, OTT domains, RTA domains, etc.

  At the time of writing, PRESA has not identified any available standard service capable of supporting the generation of network analytics for the whole E2E user plane link through multiple domains (e.g. MNO, OEM, application service providers, etc.). Several solutions could address this requirement, such as:

  - In the 3GPP domains, alignment of 'DN Performance' analytics and 'QoS Sustainability' analytics provided by NWDAF to support homogenous QoS prediction. Such alignment needs to be related and not limited to the format of the analytics request or subscription originated by the consumer, as well as the output of the analytics.
  - In the case of ToD application deployment considering Option 2, the QoS prediction service could take into account the network segment outside the MNO domain, which enables QoS prediction for the end-to-end path between the vehicle and Application Server.
  - In the case of MEC deployments, the VIS API provided by ETSI GS MEC 030 [11] could evolve to provide support for end-to-end QoS prediction.
- **Area of improvement 8** – The network needs to be capable of supporting prediction of jitter.

  Jitter is not currently supported by the current QoS prediction. Support of jitter may be provided also without changes to QoS prediction, e.g. by monitoring the change of latency which can already been predicted.

- **Area of improvement 9** – The network needs to provide mechanisms for the application to constantly monitor QoS KPIs consistent with the QoS prediction functionality. Through combined use of QoS monitoring and QoS prediction, the application needs to be capable to correlate observed QoS changes and previously received predictions in relation to QoS changes, in order to evaluate the reliability and consistency of the QoS prediction service as well.